CREDIT EDA ASSIGNMENT

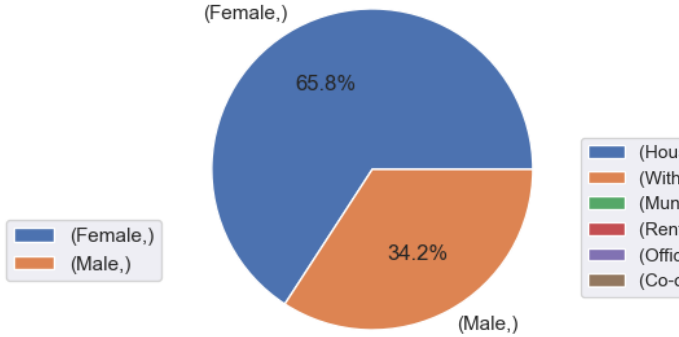
IIITB | Data Science Program January 2024

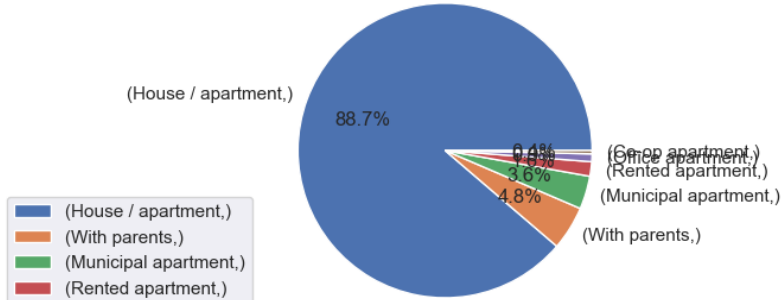
CONTENTS OF THIS ANALYSIS

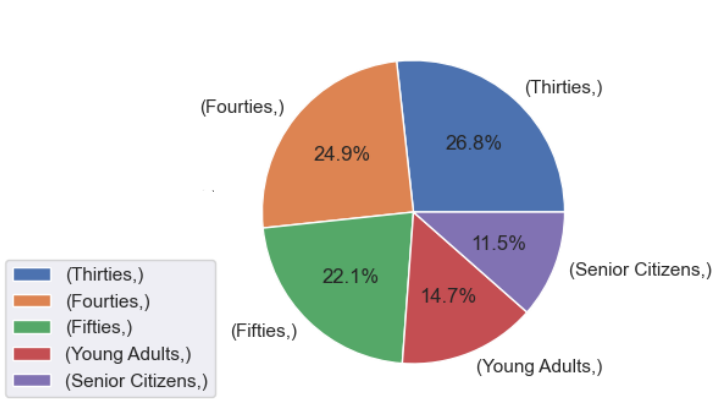
This is an analysis based on the data provided. The Aim is to figure out what are good and bad indicators of defaults.
We have two Target variables. The defaulters themselves & the Status of the loan application.

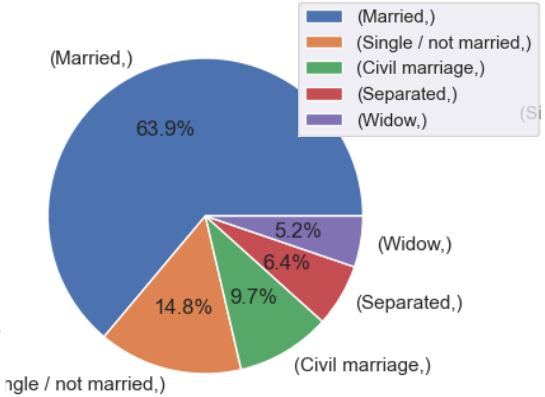
APPLICANTS DATA	Analysis of the personal data. (AGE, gender, family count etc.)
NUMERICAL DISTRIBUTION	Analysis on the amount, type of loan w.r.t the defaulters. Incl. Income, Credit amt. etc.
DEFAULTER'S DATA	Broad analysis on Defaults by personal info and groups.
CONTRACT STATUS	Analysis on the Status of application w.r.t the payers & defaulters.
CORELATION WITH PAYERS	Determining the positive indigators for which the loan was approved
CORELATION WITH BOTH TARGET	Heat maps shwoing the correlation bewteen the Defaulters & the status of loan.

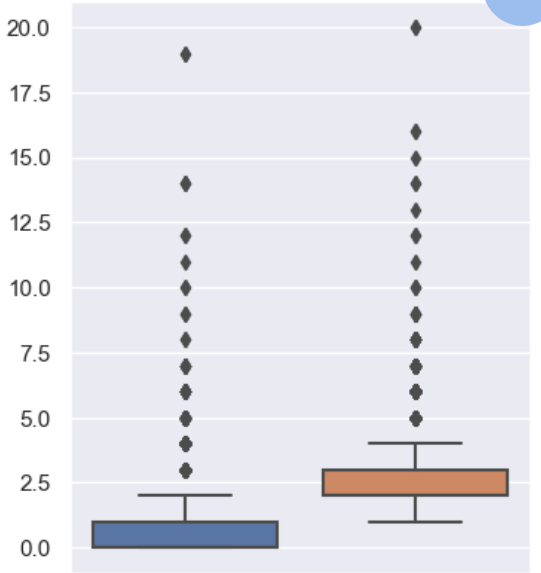
DATA OF APPLICANTS



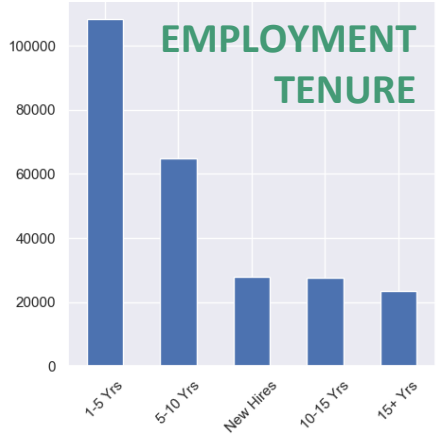


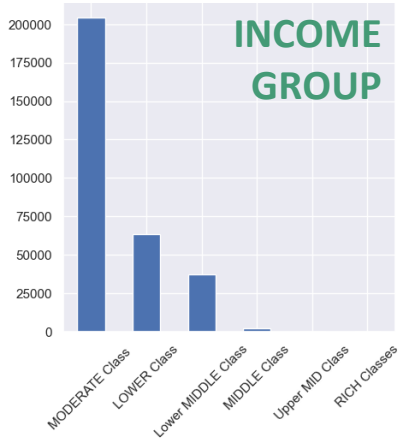


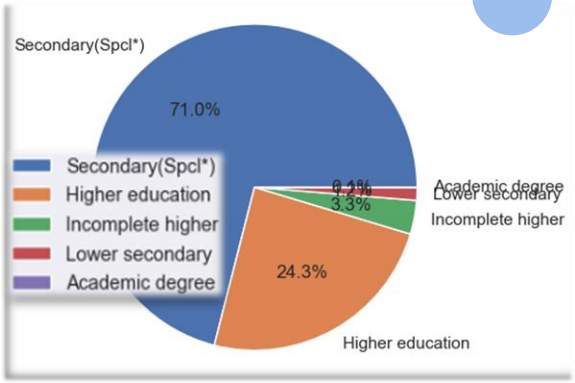




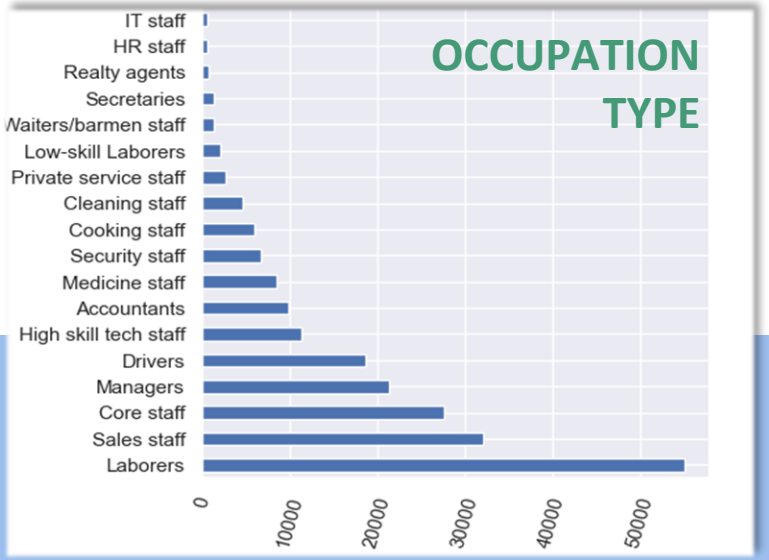
Children & Family Members Count

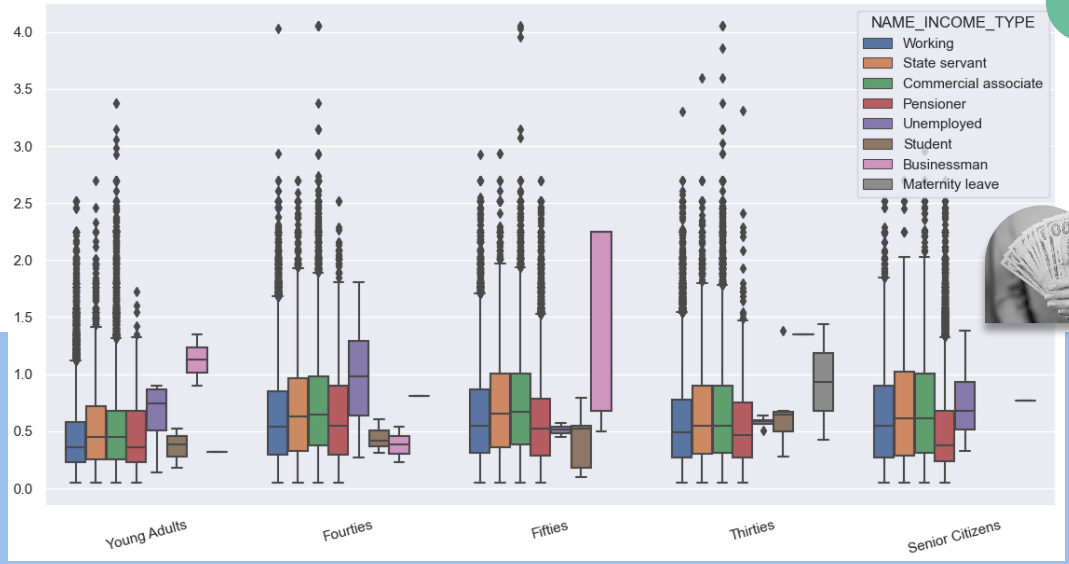






GROUPING OF APPLICANTS





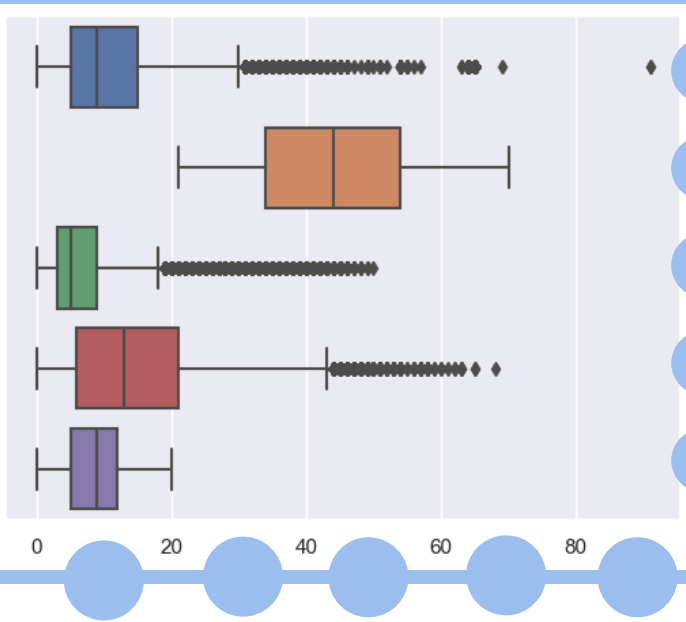
CAR'S AGE

APPLICANT'S AGE

EMPLOYMENT YEARS

REGISTRATION CHANGE YEAR

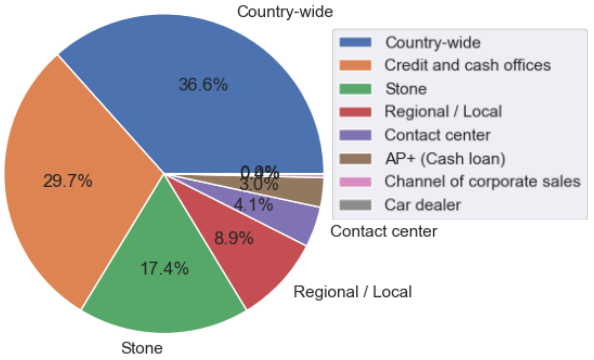
ID CHANGE YEAR



CONCLUTION by ANALYSIS

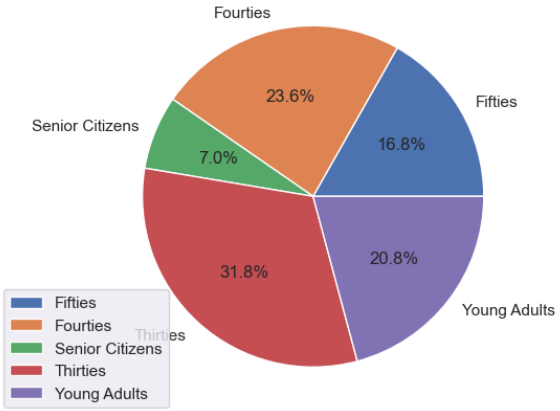
- It was found that the majority of Female have applied for loan i.e. 65.8%
- Married Applicants are higher in numbers.
- Majority of applicants has income in between 1Lakh – 2.5Lakhs
- Businessman in 50s has high amount of loans.
- Majority of loans are taken by the Laborers
- Office Apartments applicants are in least numbers

CHANNEL BY CLIENT'S APPLICATION WAS ACQUIRED.

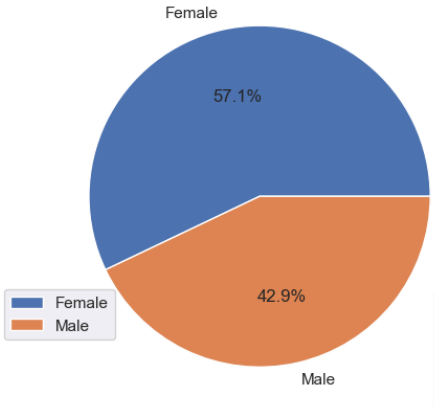


DEFAULTER'S DATA >>>> Next Slide

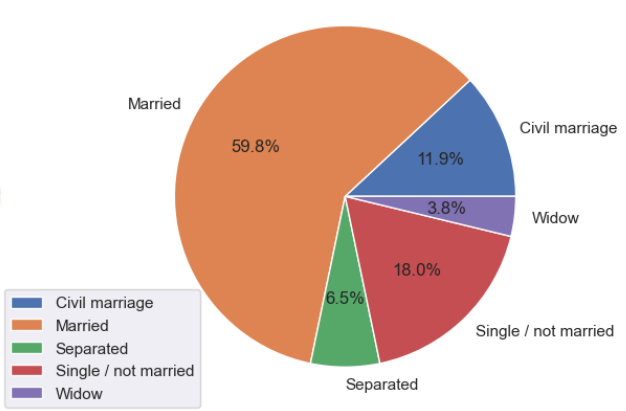
Deafulters by AGE_GROUP



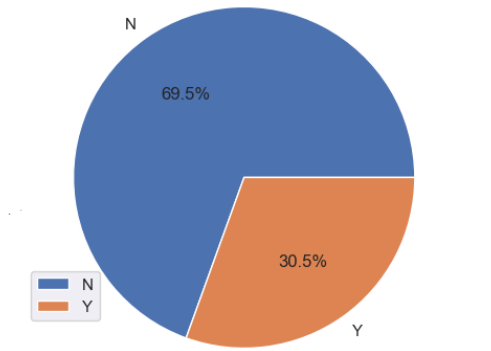
Deafulters by GENDER



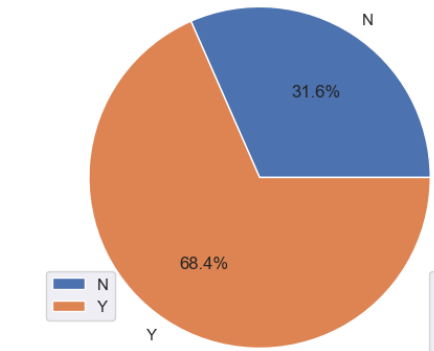
Deafulters by NAME_FAMILY_STATUS



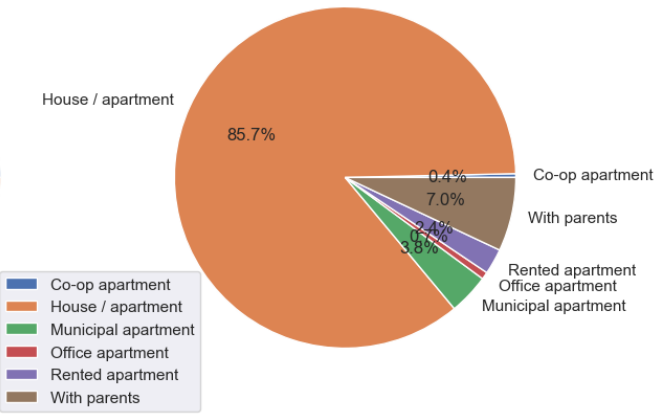
Deafulters by FLAG_OWN_CAR

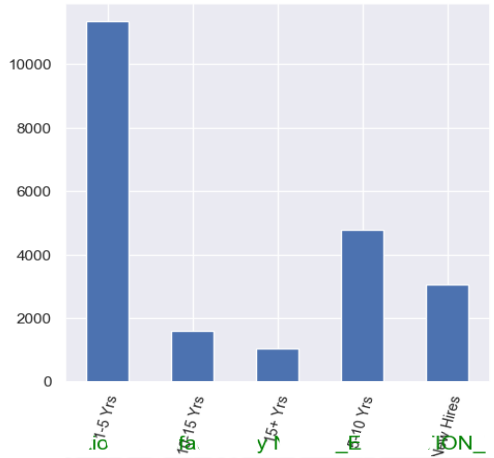


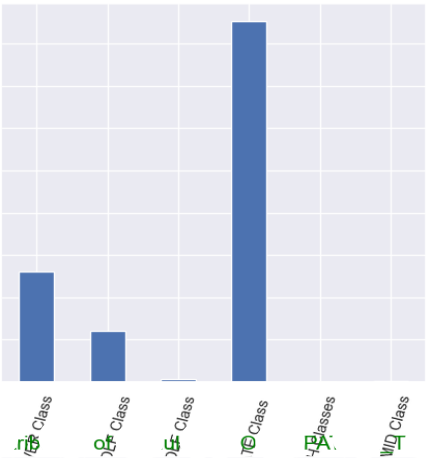
Deafulters by FLAG_OWN_REALTY

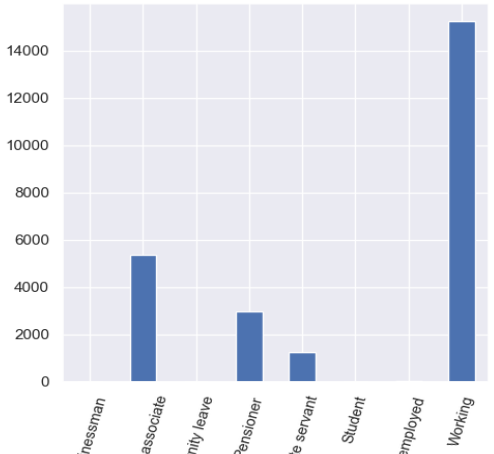


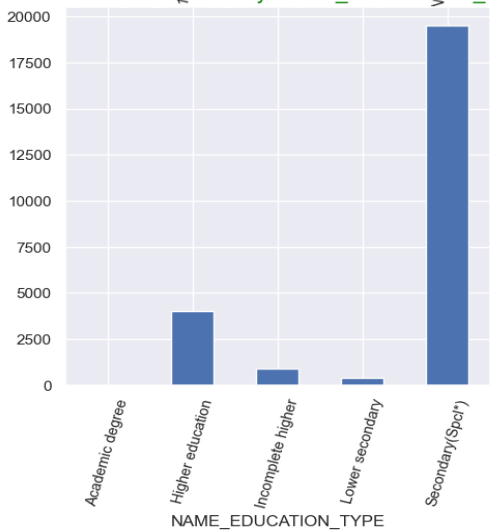
Deafulters by NAME_HOUSING_TYPE

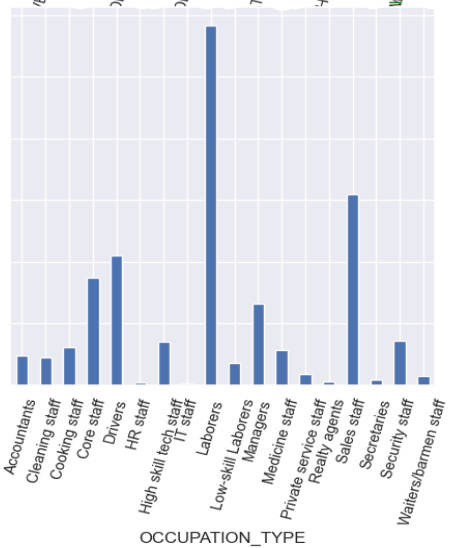



Distribution of Defaulter by EMP_EXP_GROUP
10000
8000
6000
4000
2000
0


Distribution of Defaulter by INCM_GROUP
16000
14000
12000
10000
8000
6000
4000
2000
0


Distribution of Defaulter by NAME_INCOME_TYPE
14000
12000
10000
8000
6000
4000
2000
0
Businessman
Commercial associate
Maternity leave
Pensioner
State servant
Student
Unemployed
Working


20000
17500
15000
12500
10000
7500
5000
2500
0
Academic degree
Higher education
Incomplete higher
Lower secondary
Secondary(Spcl*)
NAME_EDUCATION_TYPE


6000
5000
4000
3000
2000
1000
0
Accountants
Cleaning staff
Cooking staff
Core staff
Drivers
HR staff
High skill tech staff
IT staff
Laborers
Low-skill Laborers
Managers
Medicine staff
Private service staff
Realty agents
Sales staff
Secretaries
Security staff
Waiters/barmen staff
OCCUPATION_TYPE

CONCLUTION by ANALYSIS

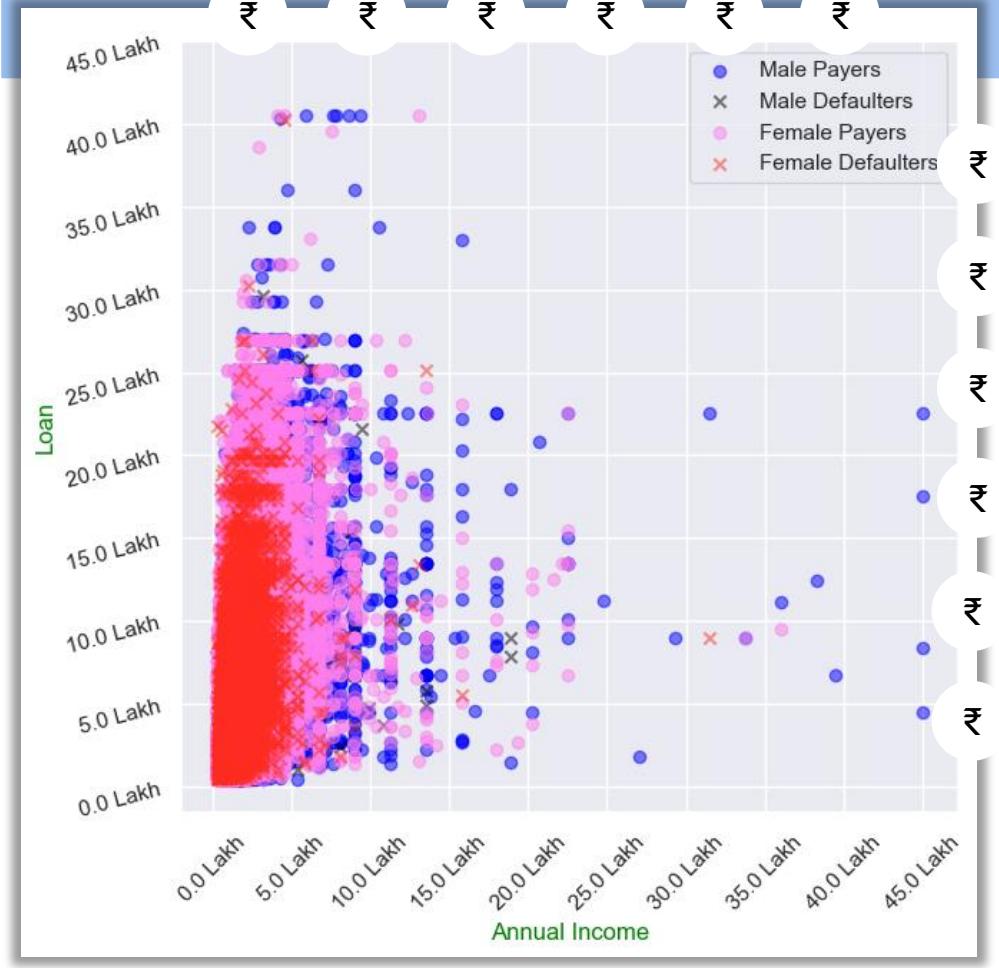
- Majority of Defaulters are Females.
- Married Applicants are higher in defaults.
- Majority of applicants are between 1-5 Years of Employment.
- Moderate class applicants are high in default numbers.
- Working Class has high number of defaults.

NUMERICAL DISTRIBUTION

>>>>

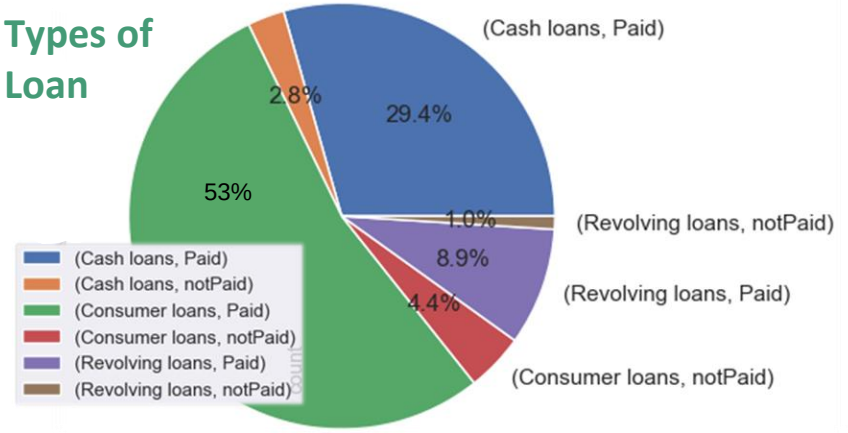
Next Slide

CREDIT TAKEN VS ANNUAL INCOME



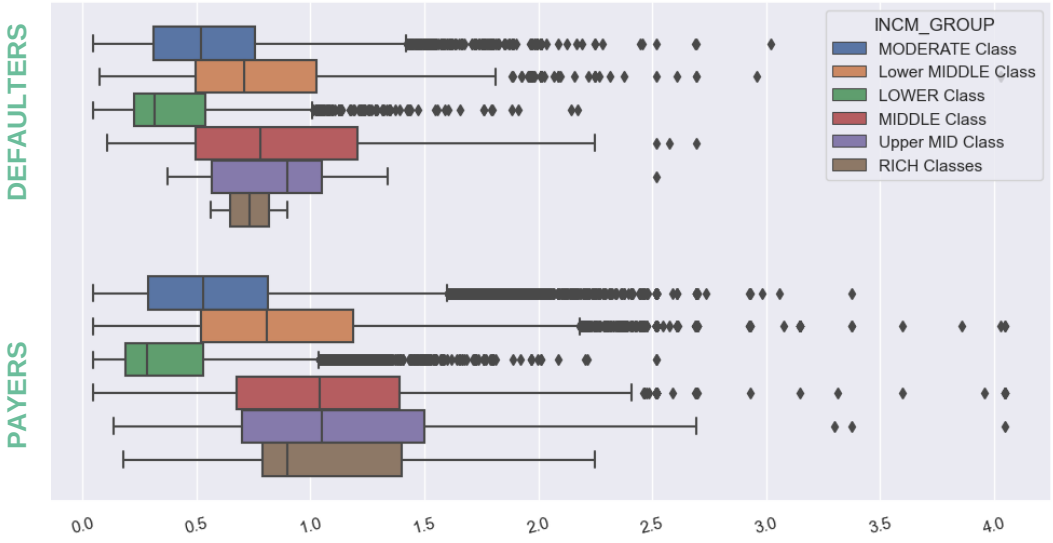
Mean of INCOME, CREDIT, ANNUITY & GOODS PRICE

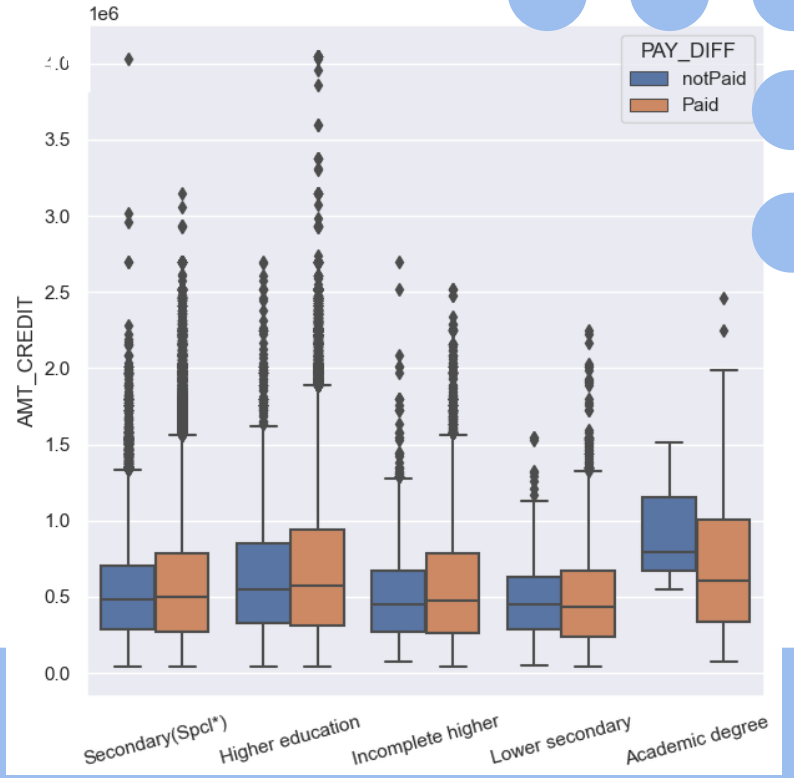
Distribution of Types of Loan

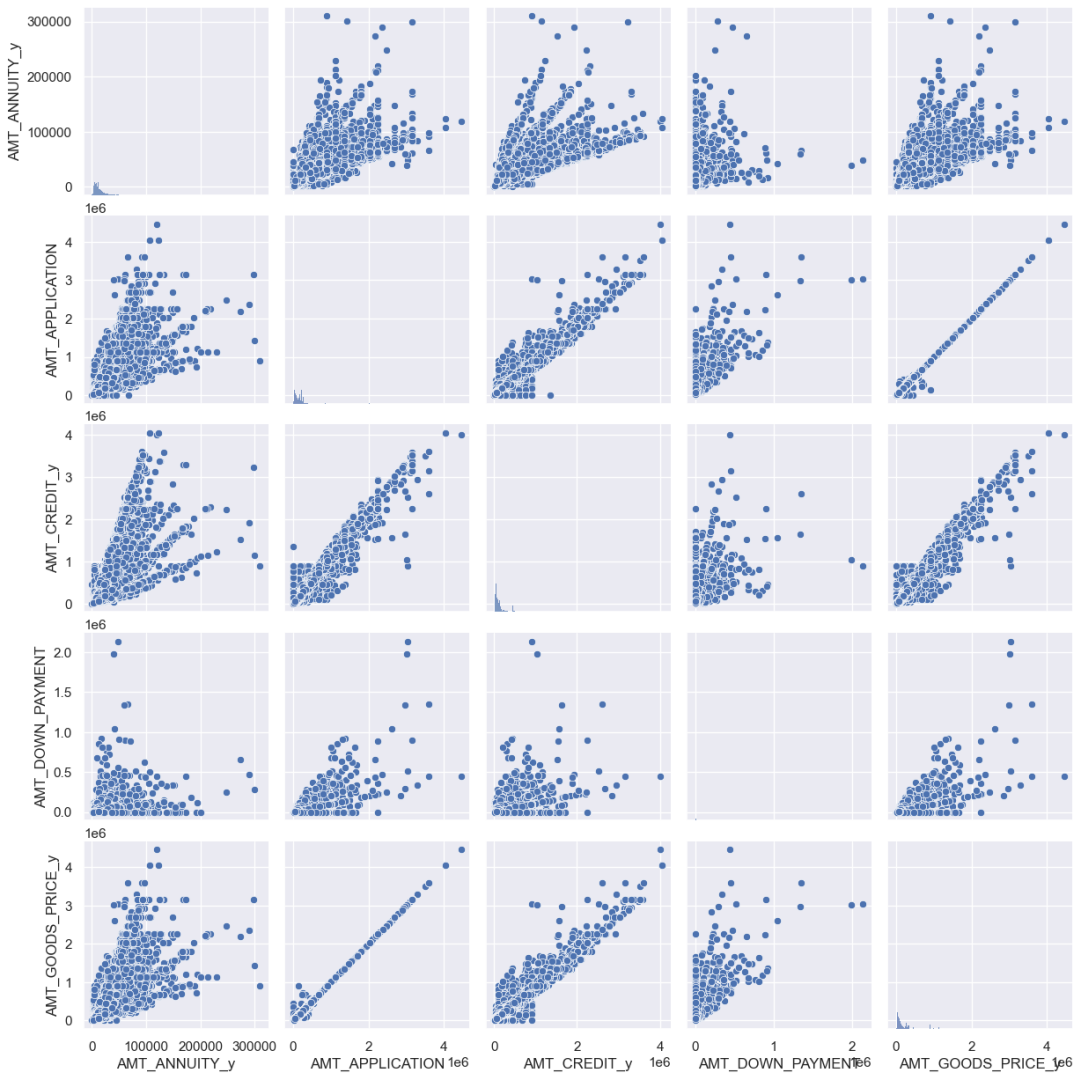


INCOME & ACEDEMIC distribution of Payer & Defaulters? W.r.t. Amount Credit

1.0 Unit
= 10Lakhs

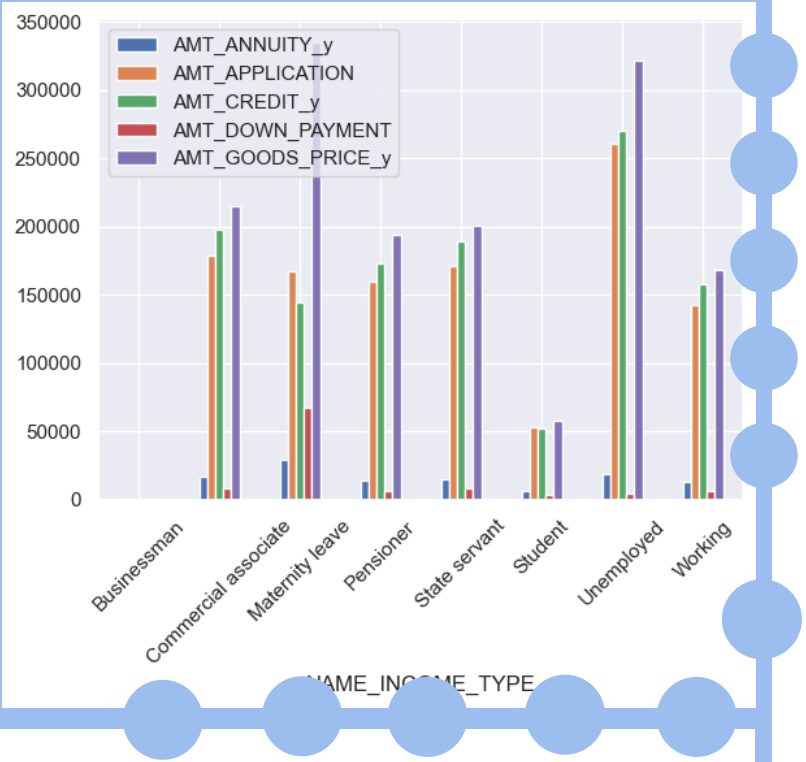






Correlation Check on Numeric values

- Strong Linear Correlation between Goods amt. &A Application amt.
- Good linear relationship between Good amt &Credit amt.
- Moderate relationship between Annuity & Credit amt. & Application amount



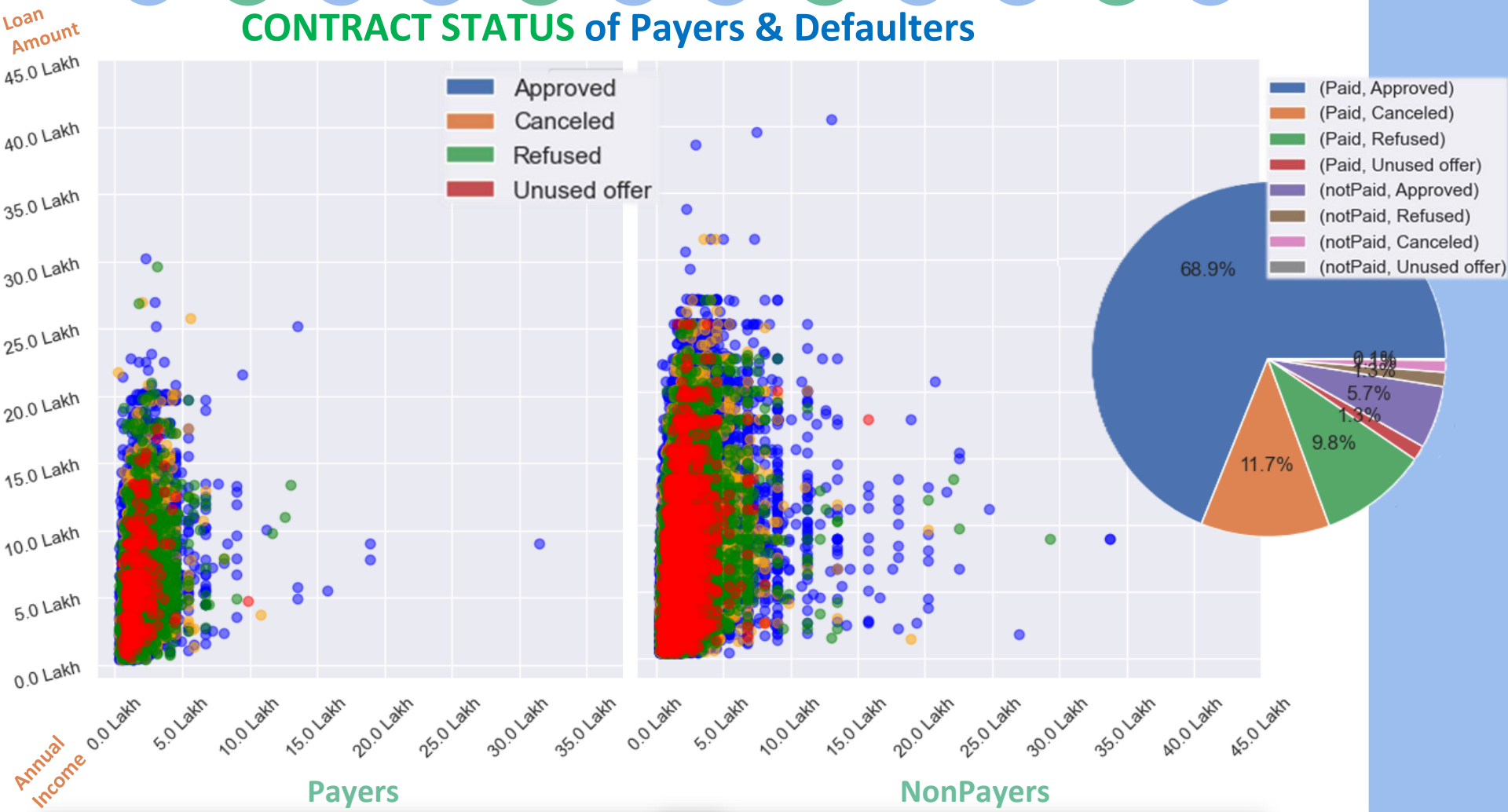
CONCLUTION by ANALYSIS

- Consumer loans are highest unpaid
- Middle Class Defaulter has taken more loans.
- Upper middle Class has taken High loans however has no issues with payments.

CONTRACT STATUS

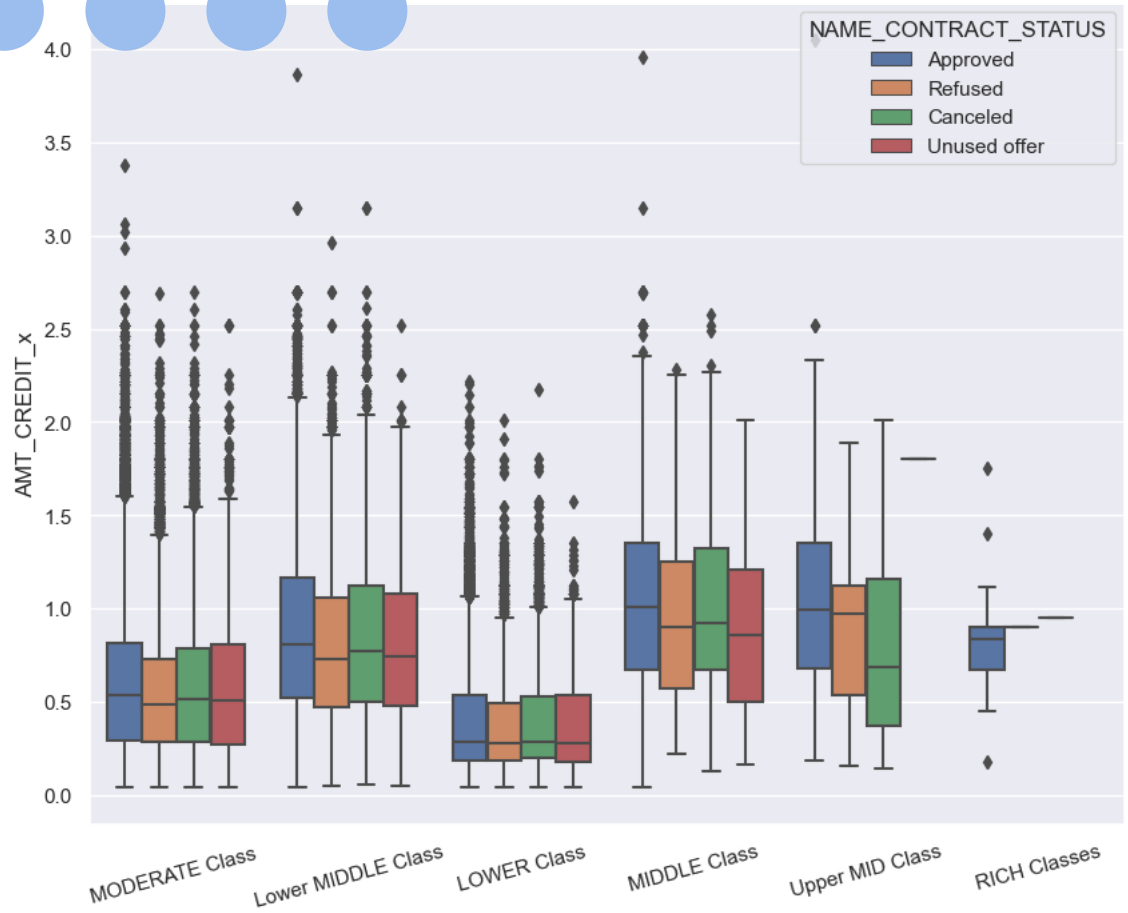
>>>>
Next Slide

CONTRACT STATUS of Payers & Defaulters

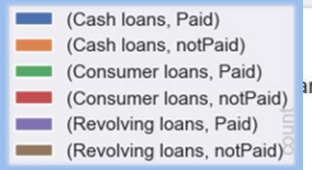


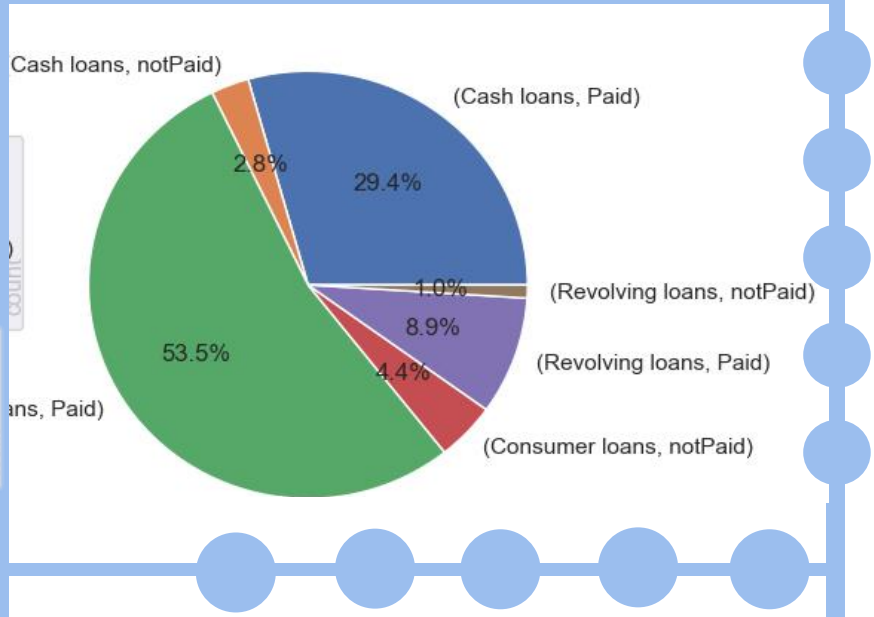
CONTRACT STATUS classified into INCOME CLASS

with CREDIT AMT.



Distribution of Types of Loan





APPLICANTS's CONTRACT STATUS (Total)	
Approved	217090
Canceled	37284
Refused	32411
Unused offer	4272

CONCLUTION by ANALYSIS

- 5.7% of Defaulters loan application was approved. That comes under the risky portfolios.
- 1.3% non-Defaulter has unused offers. Bank should give them better offers. An Opportunity.
- 4272 are number of unused offers, should check with the Credit enquiry data. If they are looking for loans outside.
- Consumer loans has majority of non-payments.

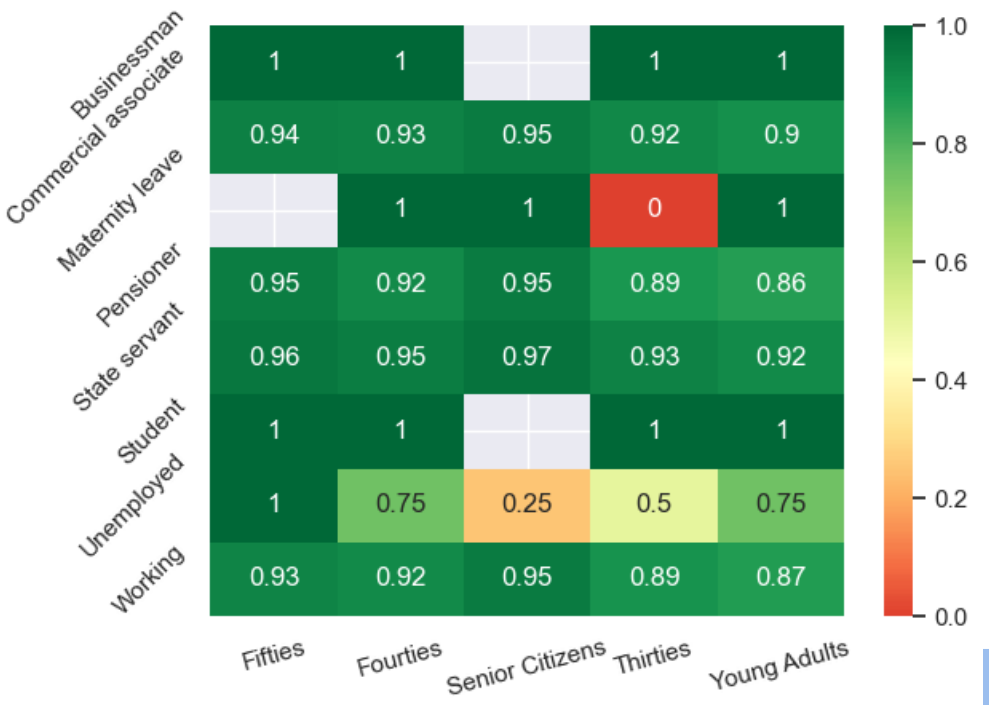
CORELATION WITH PAYERS >>>>

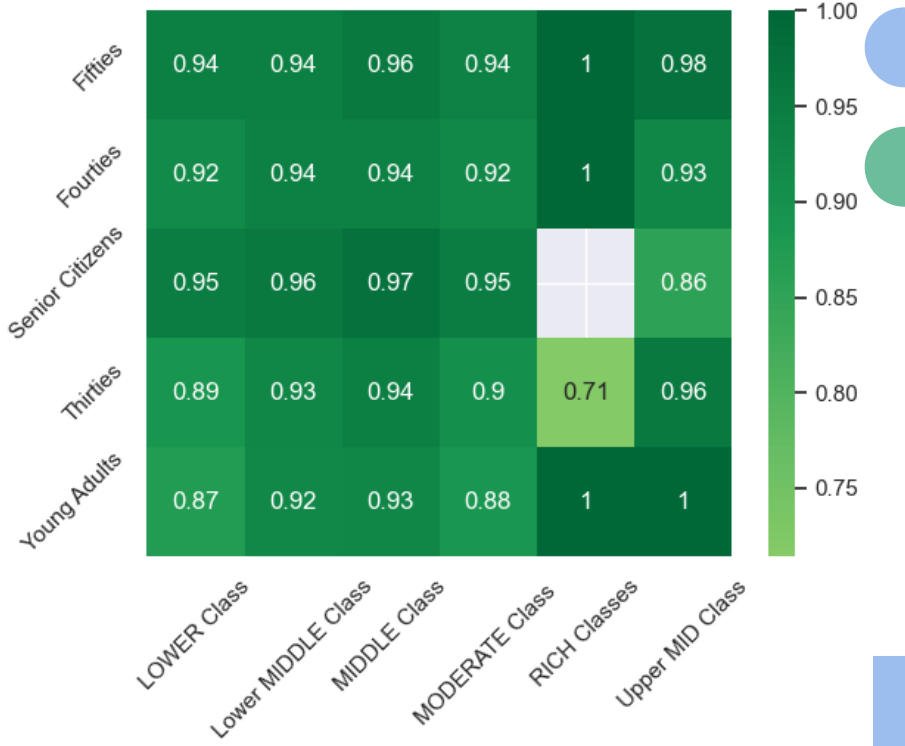
Next Slide

Positive Correlation for Age group with

Income Type & Income Group

w.r.t non-Targets

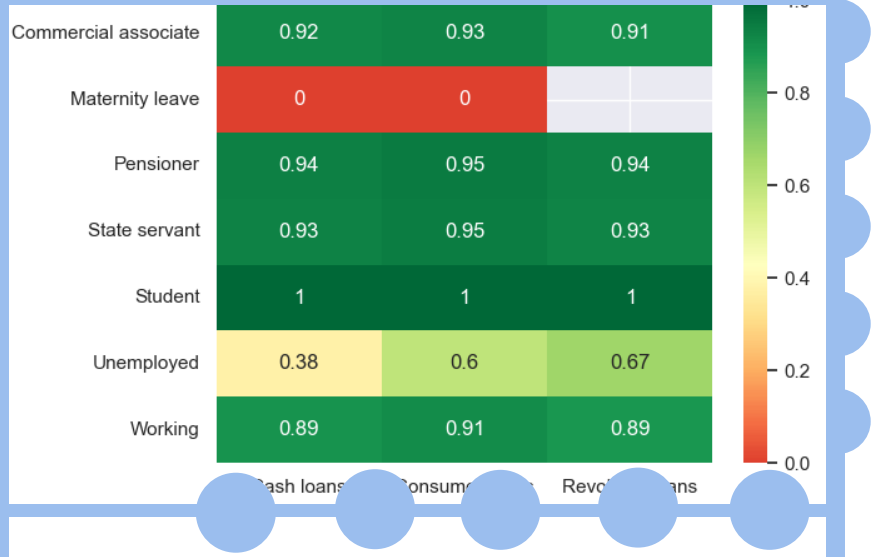




Positive Correlation b/w

Contract Type & Income Group

w.r.t non-Targets



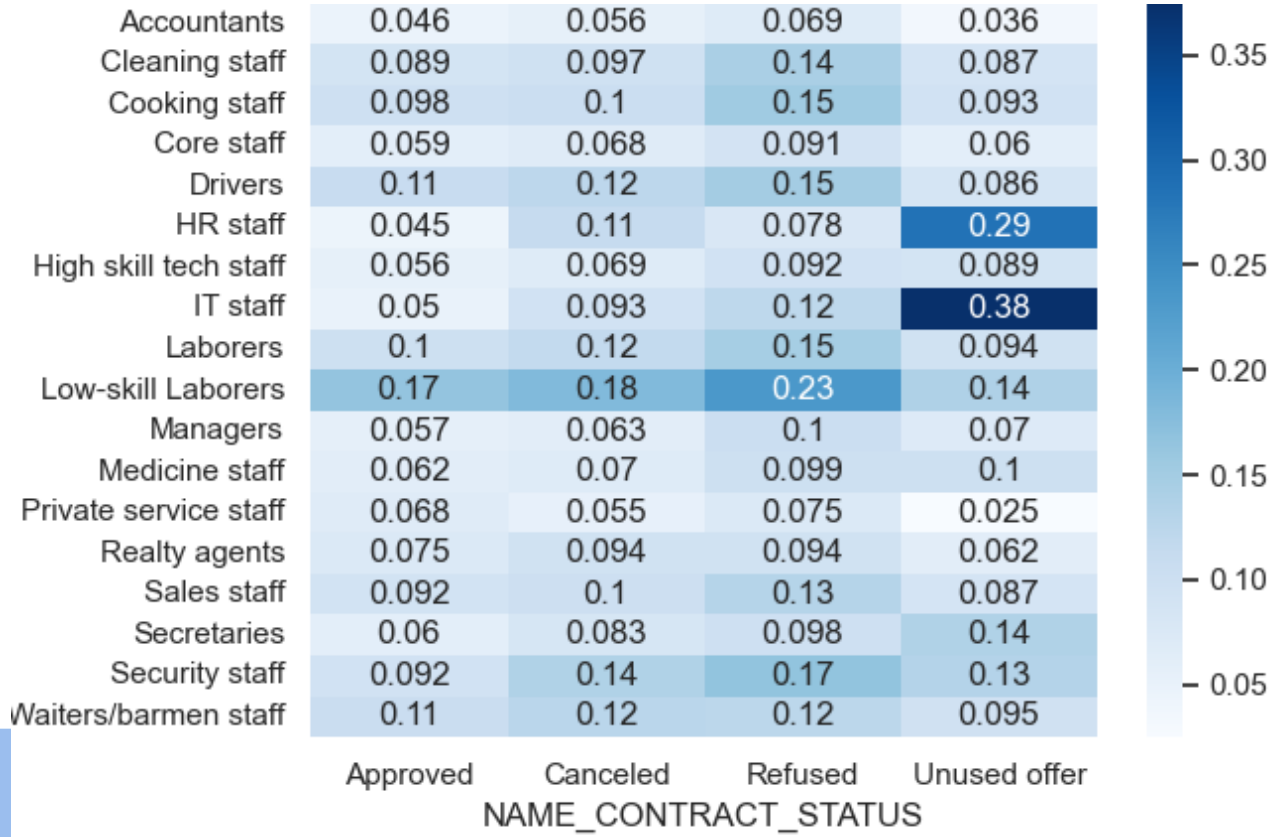
CONCLUTION by ANALYSIS

- Businessmen are good payers, on the other hand the Unemployed has high in number of defaults across all loan category.
- Working class has moderate amount of good payment.
- Senior Citizen State Servants are best performing.
- Students & Pensioners are also good payers
- across all loan types

CORELATION WITH BOTH TARGET >>>> Next Slide

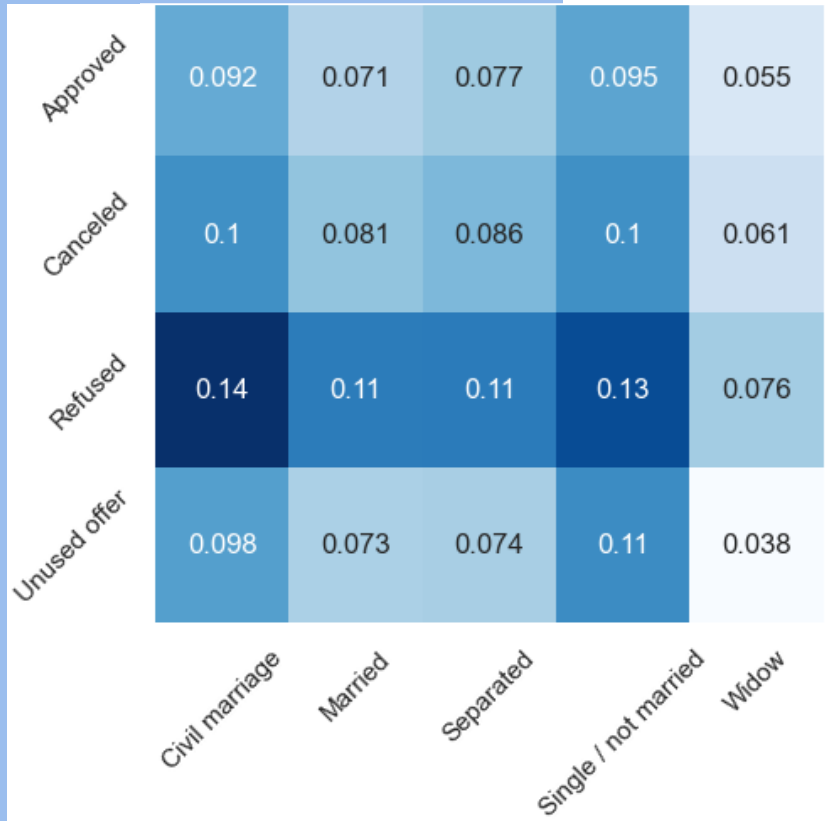
Correlation w.r.t both TARGETS!

- By Occupation

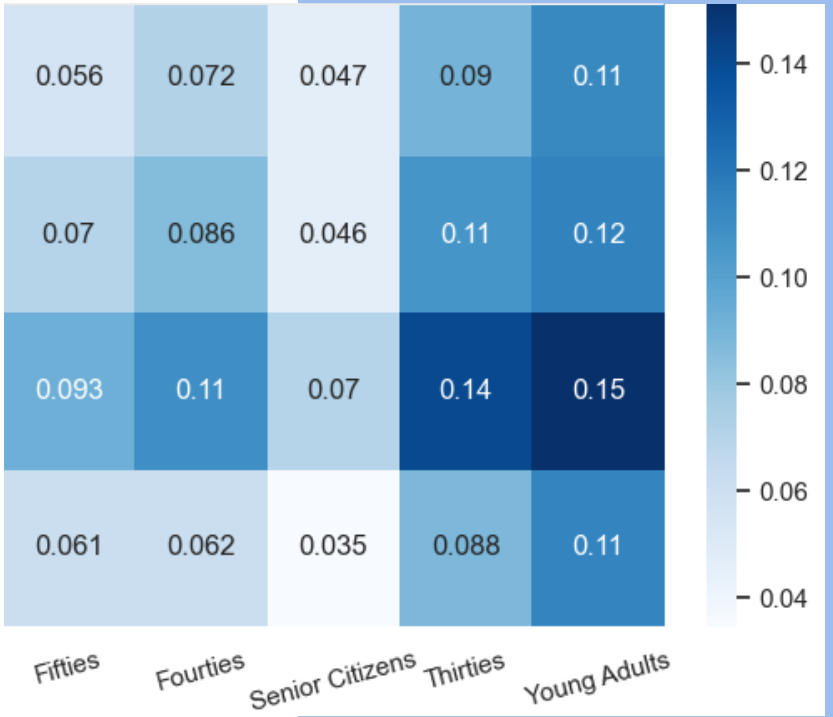


Loan Status Correlation w.r.t Defaulters, by-

FAMILY STATUS

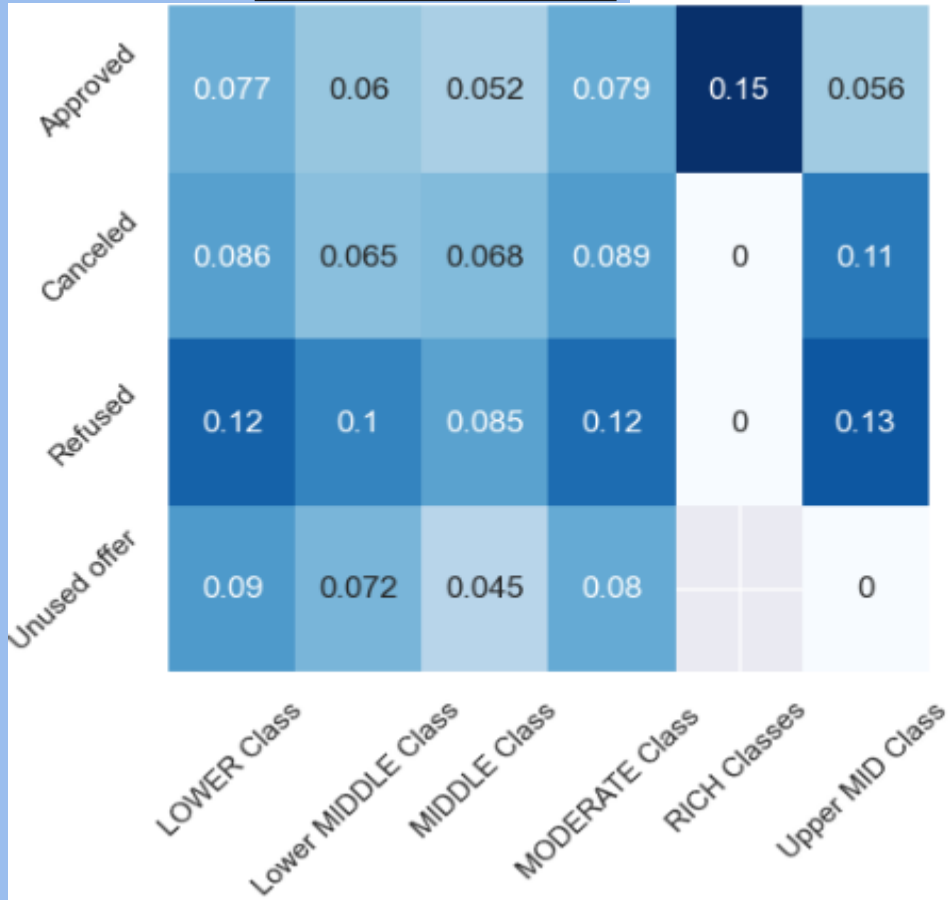


AGE GROUP

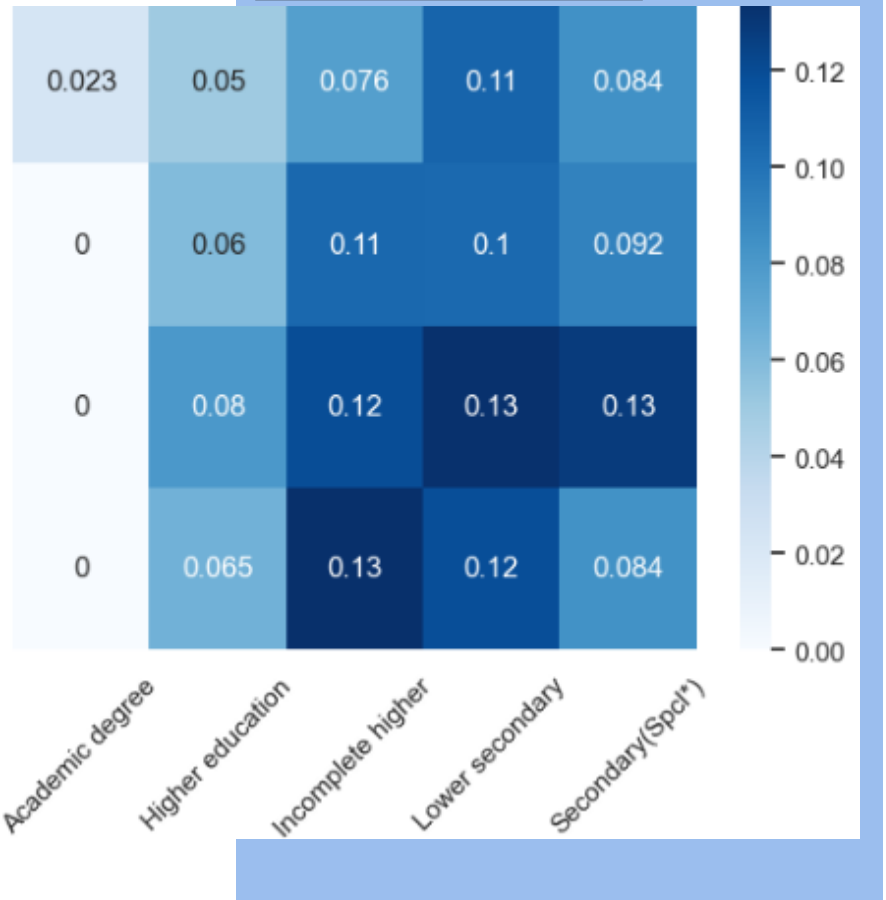


Loan Status Correlation w.r.t Defaulters , by-

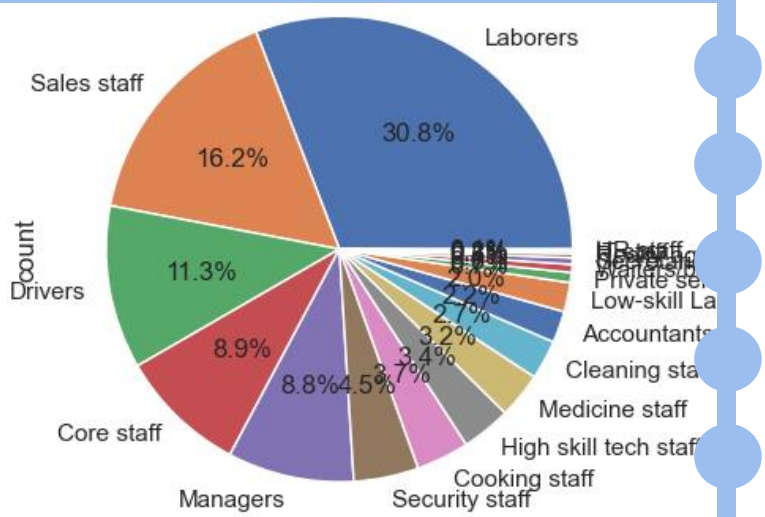
INCOME GROUP



EDUCATION TYPE



Numbers of refusal Percentage



CONCLUTION by ANALYSIS

- Low skilled Labors has high % of refusals.
- IT & HR Staff defaulters are wise as the have unused offer.
- Civil marriage defaulters has high amount of refusal, followed by singles.
- Young adult and Thirties defaulters (<40) has high % of refusal
- Lower Secondary Education Type has high refusal %
- 50+ & Senior Citizen Age_Gorup have
- least refusal rates. (>0.1)

Final Analysis >>>> Next Slide

FINAL CONCLUTION by ANALYSIS

Green Focus area:-

- Upper middle Class has taken High loans however has no issues with payments.
- 1.3% non-Defaulter has unused offers. Bank should give them better offers. An Opportunity.
- State Servants across all Age groups are highly considerable.
- Businessmen are good payers
- Students turned out to be a very good payers (poss. because of many EDU Institution are their payers)
- Pensioners have good performers poss. because the loans are paid via their fixed pensions.
- Upper middle Class has taken High loans however has no issues with payments.
- IT & HR Staff defaulters are wise they are not taking offered loan knowing they have few defaults.
- All 50+ Aged Applicants are considerable.

Risky areas:-

- Moderate Income class applicants are high in default numbers. (1 Lakh-2.5 Lakhs PA).
- 5.7% of Defaulters loan application was approved. That comes under the risky portfolios. (Deep Dives require).
- Working Class has high number of defaults.
- Unemployed Applicants has high in number of defaults across all loan category.
- Applicants in thirties who's income is based on maternity leave also exhibits high defaults.
- Low Skilled Labors has high refusal.
- Civil married Applicant's defaulting rate is high.
- Young adult and Thirties defaulters (All<40) has high % of refusal.
- Lower Secondary Education Type has high refusal %.

THANKS!

Pallav Rajput

ribhu.s18@gmail.com
+91 9888 143 002