

A large, soft pink brushstroke graphic that starts from the top left and sweeps across the top half of the slide, creating a textured, painterly effect.

Credit EDA - Case Study

SUBMITTED BY - SREELAKSHMI THANKACHAN

Problem Statement

In this case study, the company wants to understand the driving factors (or driver variables) behind loan default, that is the variables which are strong indicators of default.

Two types of risks are associated with the company's decision:

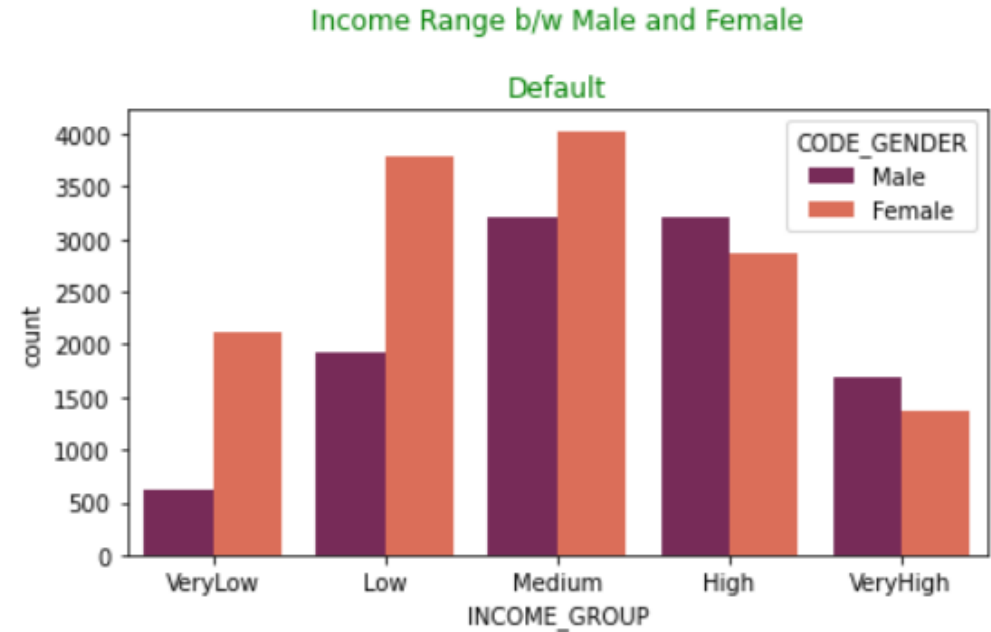
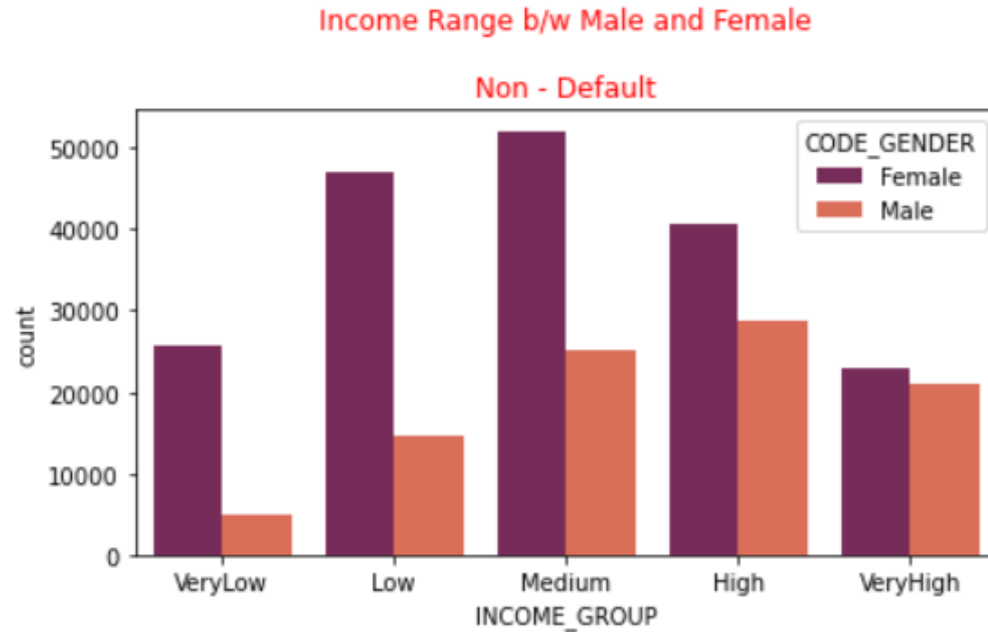
- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company**
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.**



Analysing steps:

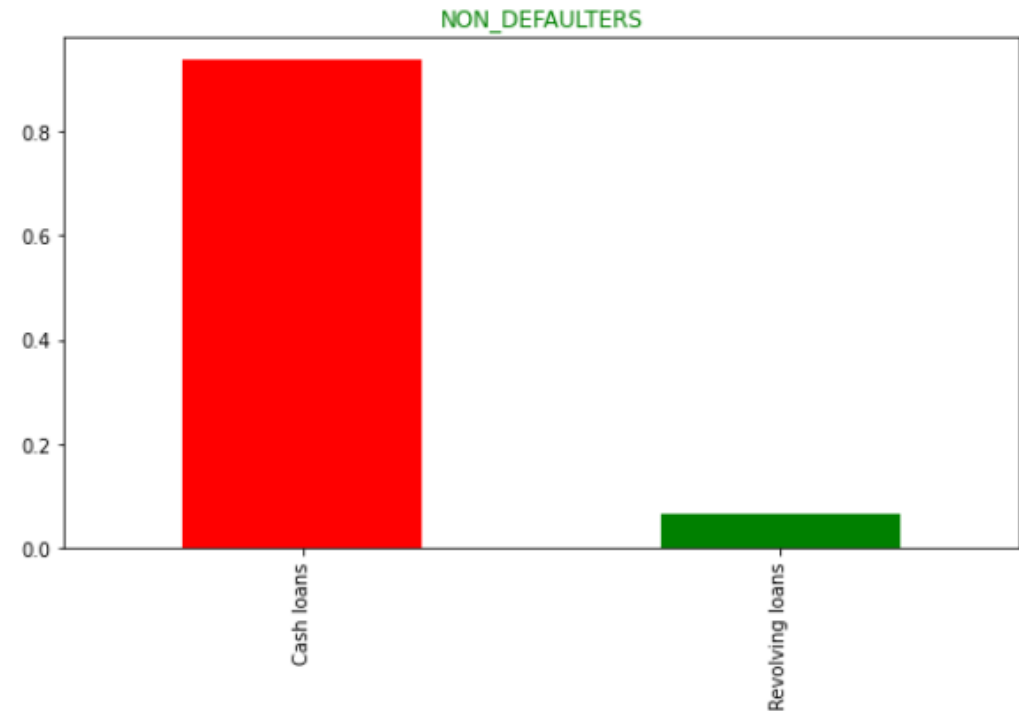
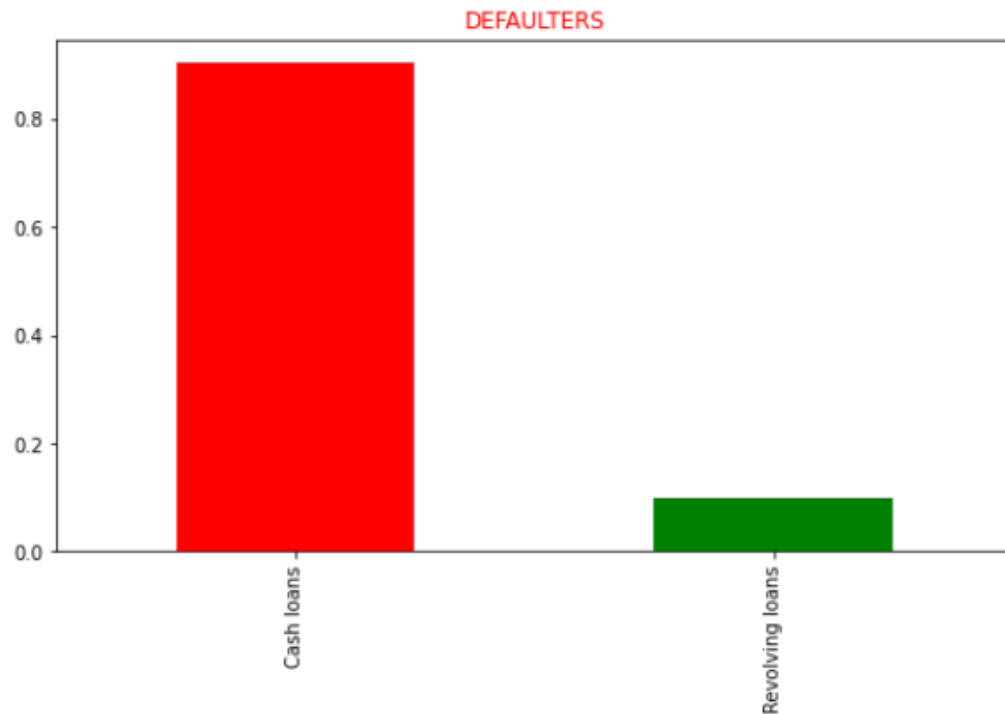
- Treatment of missing values
- Handling of outliers
- Univariate, Bivariate and Multivariate Analysis
- Top correlations
- Case - study Insights
- Case - Study Conclusions

Analysis of Gender Based on Income Range



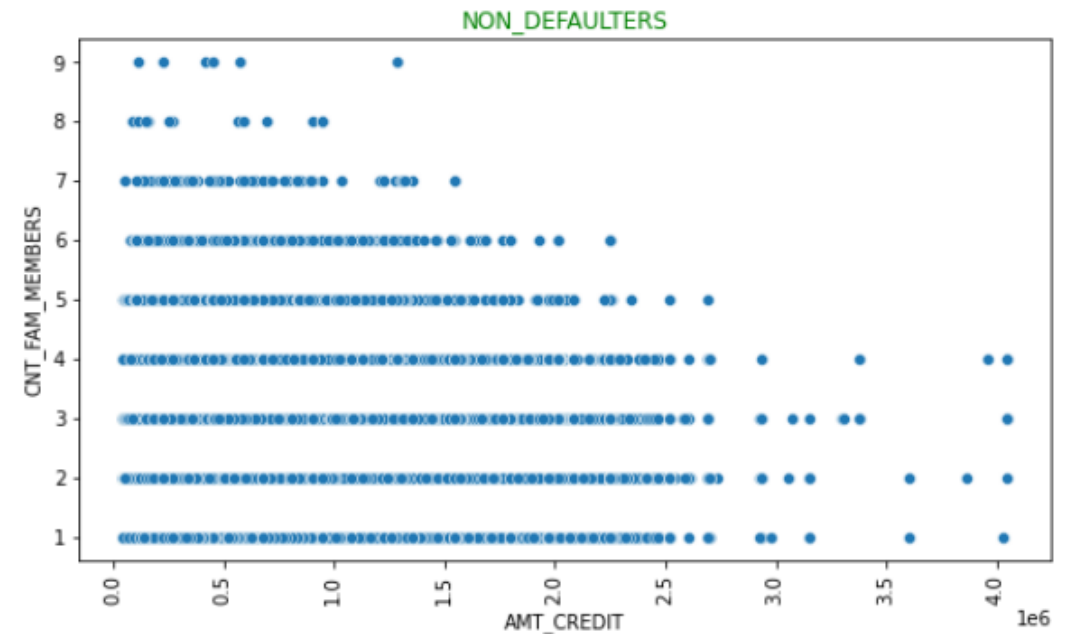
- Male clients in low and very low income range are tend to be Defaulters
- More females are applying for loans than Males

Contract Type Analysis



- The people who took Revolving loans are tend to be a Defaulter
- Most of the people are applying for Cash loans rather than Revolving loans

Analysis of Count of Family Members

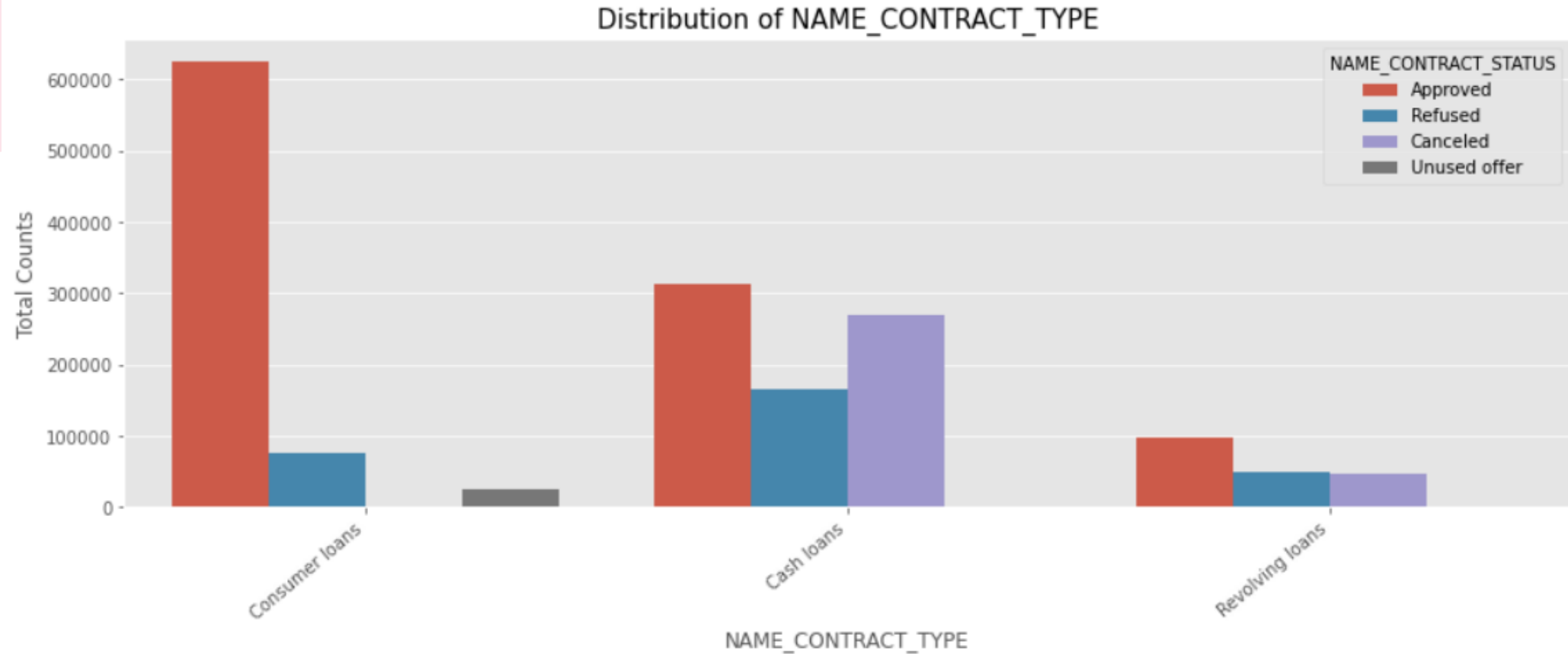


- Families with less number of people default often with low credit amount
- Families with large number of people default less often and the amount of credit is low for them
- Families with less number of people are applying for loans more often

Top 10 Correlation for Defaulters

	VAR1	VAR2	CORRELATION	CORR_ABS
30	AMT_CREDIT	AMT_GOODS_PRICE	0.983104	0.983104
12	AMT_ANNUITY	AMT_GOODS_PRICE	0.752693	0.752693
32	AMT_CREDIT	AMT_ANNUITY	0.752184	0.752184
27	DAYS_REGISTRATION	DAYS_EMPLOYED	0.192439	0.192439
20	DAYS_EMPLOYED	AMT_ANNUITY	-0.081194	0.081194
13	AMT_ANNUITY	AMT_INCOME_TOTAL	0.046416	0.046416
31	AMT_CREDIT	AMT_INCOME_TOTAL	0.038123	0.038123
6	AMT_INCOME_TOTAL	AMT_GOODS_PRICE	0.037574	0.037574
26	DAYS_REGISTRATION	AMT_ANNUITY	-0.034262	0.034262
34	AMT_CREDIT	DAYS_REGISTRATION	0.025874	0.025874

Contract Type Analysis



- Most of the applications are for 'Cash loan' and 'Consumer loan'.
- The cash loans are refused and Cancelled more often than others.
- Unused offer is very less

Client Type Analysis

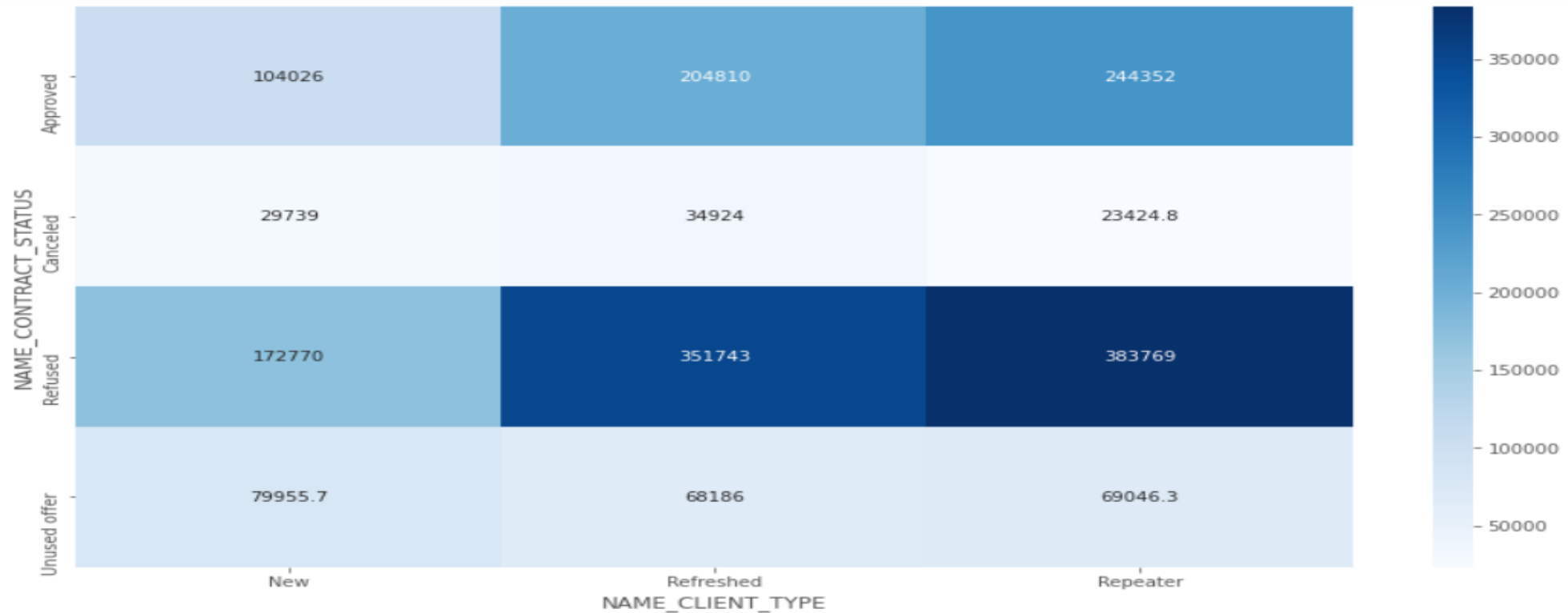


- Most of the loan applications are from repeat customers.
- Repeat customers get refused more often.
- The percentage of new customers is low and most of them get Approved

Top correlations on Previous Application Data

	Var1	Var2	Correlation
4	AMT_CREDIT	AMT_APPLICATION	0.975824
9	DAYS_DECISION	AMT_CREDIT	0.133763
8	DAYS_DECISION	AMT_APPLICATION	0.133660
14	SELLERPLACE_AREA	DAYS_DECISION	0.018382
13	SELLERPLACE_AREA	AMT_CREDIT	0.009567
12	SELLERPLACE_AREA	AMT_APPLICATION	0.007649

Contract Status - Credit Amount



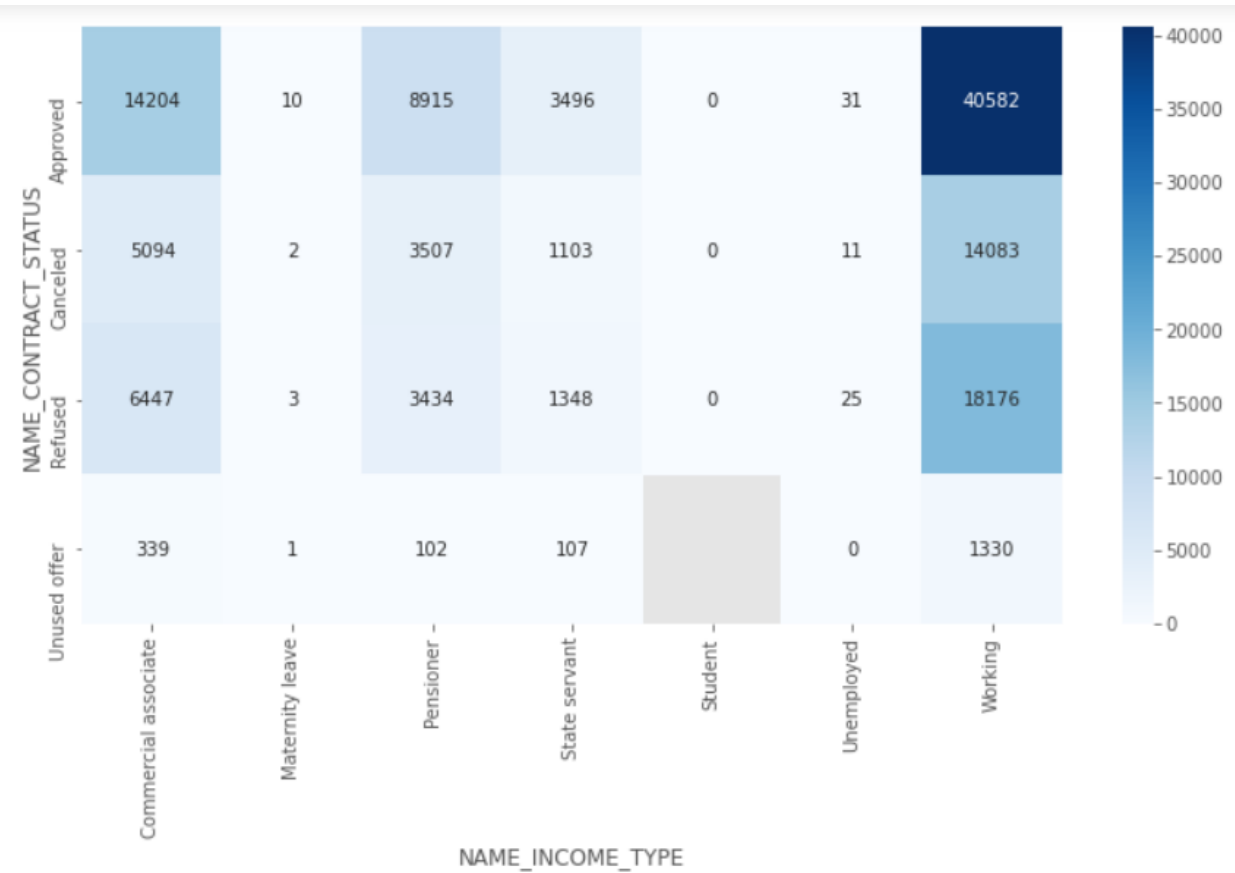
- For "unused offer" the CREDIT AMOUNT is low. This may be the reason for customer is not using the offer
- Previously Refused 'Refreshed' and 'Repeater' client has high Credit amount

Previously Approved and Currently Defaulted Data Correlation

Previously Approved And Currently Defaulted Clients belong to

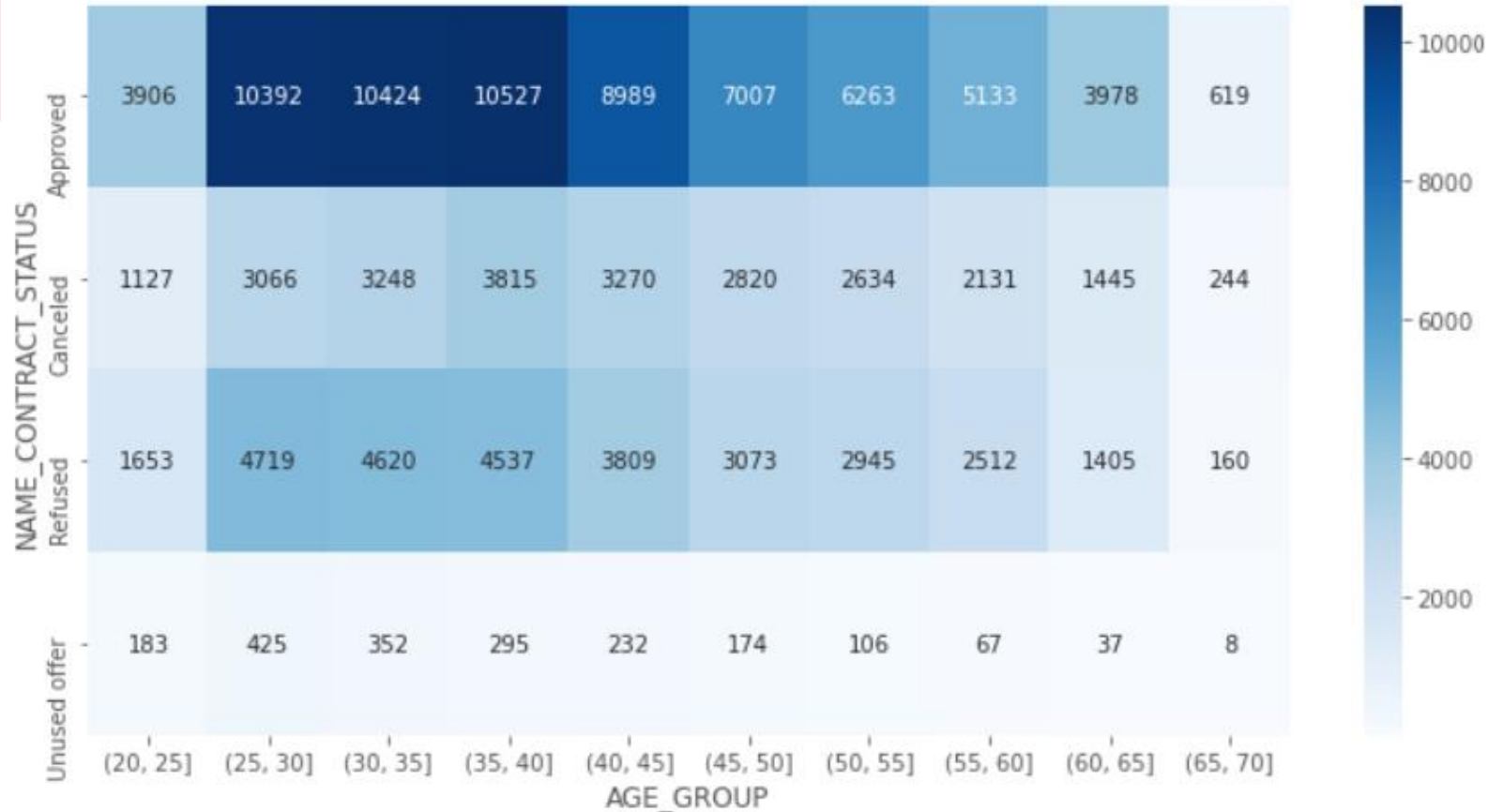
- INCOME_GROUP - Medium
- AGE_GROUP - Age ranging from 20 to 40
- NAME_INCOME_TYPE - Working
- OCCUPATION_TYPE - Laborers, Sales staff, Drivers
- ORGANIZATION_TYPE - Business Entity type

CONTRACT STATUS AND INCOME TYPE ANALYSIS



- HERE HIGHER VALUE INDICATE CORRELATION TO DEFAULT
- WORKING AND PREVIOUSLY APPROVED CLIENTS ARE THE ONES DEFAULTED MORE
- HERE PREVIOUSLY REFUSED, CANCELLED, UNUSED OFFER ALSO HAS DEFAULTERS BUT COMPARATIVELY LOW TO PREVIOUSLY APPROVED
- SHOULD BE CAREFUL ABOUT THE CLIENTS WHO ARE PREVIOUSLY APPROVED

Contract Status - Age Group Correlation



- Here higher value indicate correlation to default
- Previously approved category of people in age ranging from 25 to 40 have more defaulters
- Previously Refused, cancelled, Unused offer also have defaulted in current data, but less compare to Previously approved.

Case Study - Insights

Driving Factors to default:

Income Group :
Medium

Age Group : Age in
range 25 to 40

Gender : Males are
tend to be defaulters

Occupation :
Laborers, Sales staff,
Drivers

Contract type :
Revolving loans are
more likely to
become a defaulter

Family Status :
Married and widow

Region Rating :
People living in
better areas (rating
3)

Contract type :
Student ,pensioner
and Businessman

Working and
previously approved
clients are the ones
defaulted more
often

Clients who don't
have a car are tend
to be a defaulter

Family with less
number of members
and children are
tend to be defaulter

People with higher
education who have
lesser income are
tend to be defaulter

Secondary /
secondary special
with low income are
tend to be a
defaulter.



Conclusions

Bank should be careful while approving loans to previously approved clients especially the one belongs to "Working" category

Should focus more on contract type Student ,pensioner and Businessman

If the clients are living with parents then give preference to them since they are less tend to be a defaulter

Banks should focus more on Higher educated married people with higher Income

Should focus on unused offer. In unused offers the credit amount is low, this may be the reason that the client refused to accept the offer.

Most of the clients of banks are repeated so bank should focus to get new customers to improve the business

Previous applications with Refused, Cancelled, Unused loans have cases where Payments are on time in current application. So bank should reconsider that data to minimize loss

Should give more weightage to female customers since they default less