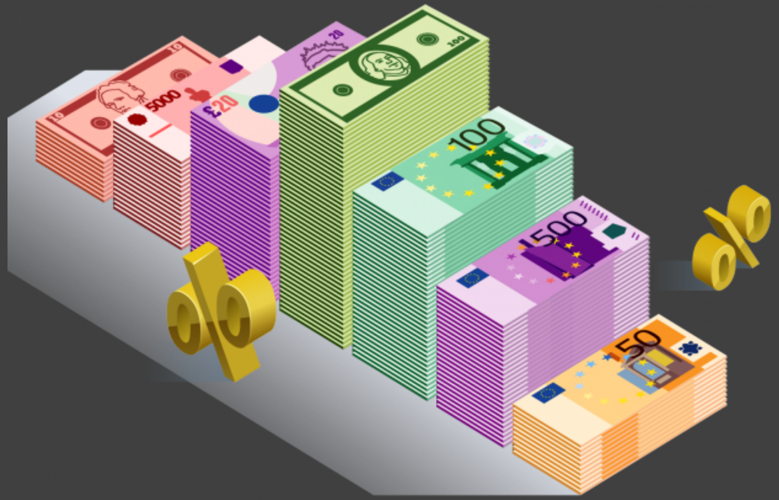


Credit Risk Analysis

BFS Capstone

1. Harsha Ravi
2. Amani Prasad
3. Shefali P
4. Vighnesh Desai

Executive Summary

CredX -a leading credit card provider gets thousands of credit card applications every year. In the past few years, CredX is experiencing an increase in credit loss. The CEO believes that the best strategy to mitigate credit risk is to ‘acquire the right customers’.

.

Help CredX identify the right customers using predictive models. Build models on the bank’s past applicants data and determine the factors affecting credit risk, create strategies to mitigate the acquisition risk and assess the financial benefit to the bank

The aim is to identify patterns and deriving variables, which indicate if an applicant is likely to default on credit, which may be used for deciding whether the credit card should be issued to the applicant or not

Assumption – The applicants with no information on CC utilization are assumed to have no prior credit history and hence we will consider those applicants with zero CC utilization

Implement exploratory data analysis on the demographic and credit bureau dataset to gain an understanding of risk analytics in a business environment.

Problem Solving Methodology

A. Business Context

Develop understanding of credit risk analysis and how to mitigate the risk

B. Load and Profile Data, Filtering & Preparation

Understand variables in **demographic** and **credit bureau** data through data dictionary, data cleaning, treating NA values, addressing outliers and Identifying relevant columns for analysis

C. Data Exploration & Analysis

Understanding the trends in categorical and quantitative variables through EDA. Perform univariate ,bivariate and multi- variate analysis and identify the factors affecting credit risk

Check the information value for all the variables to aid in feature selection

D. Analysis Outcomes

Build predictive models to identify right customers. Tune the model hyperparameters and choose the best fit model to validate and evaluate the solution

E. Road Map

Present the insights from our analysis and provide an approach to further solve the problem by building application score card to decide on whether CC needs to be issued or not

Key findings regarding demographic data

Data Profile

1. Demographic dataset includes information provided by the applicants at the time of credit card application. It contains customer-level information on age, gender, income, marital status, etc.
2. 12 initial set of data attributes.
3. Contains 71295 observations
4. We see 1425 observations lack Performance tag. These are the applicants which were rejected CC by the bank, we subset these records separately to evaluate the model on

12 variables

Variables	Description
Application ID	Unique ID of the customers
Age	Age of customer
Gender	Gender of customer
Marital Status	Marital status of customer
No of dependents	No. of children of customers
Income	Income of customers
Education	Education of customers
Profession	Profession of customers
Type of residence	Type of residence of customers
No of months in current residence	No of months in current residence of customers
No of months in current company	No of months in current company of customers
Performance Tag	Status of customer performance (" 1 "Default")

Issues with Data

1. Data issues observed in age variable which has negative and no. lesser than 18 but as we know the minimum age requirement for issuing a credit card is 18 so we impute age less than 18 with 18
2. Income cannot take negative value so we replace those with zero
3. Education details is not available for 118 applicants ,we impute it with education category 'Others' which takes up less no. of events
4. Other variables such as profession, gender, marital status has very minimal records with null values , we impute those null values with biggest category for that variable

Key findings regarding credit bureau data

Data Profile

1. Credit bureau dataset includes information provided by credit bureau and contains variables such as 'number of times 30 DPD or worse in last 3/6/12 months', 'outstanding balance', 'number of trades', etc.
2. 19 initial set of data attributes.
3. Contains 71295 observations
4. We see 1425 observations lack Performance tag. These are the applicants which were rejected CC by the bank, we subset these records separately to evaluate the model on

19 variables

Variable	Description
No of times 30/60/90 DPD or worse in last 6 months	Number of times customer has not payed dues since 30/60/90days in last 6 months
No of times 30/60/90 DPD or worse in last 12 months	Number of times customer has not payed dues since 30/60/90 days days last 12 months
Avgas CC Utilization in last 12 months	Average utilization of credit card by customer
No of trades opened in last 6/12 months	Number of times the customer has done the trades in last 6/12 months
No of PL trades opened in last 6/12 months	No of PL trades in last 6/12 month of customer
No of Inquiries in last 6/12 months (excluding home & auto loans)	Number of times the customers has inquired in last 6/12 months
Presence of open home loan	Is the customer has home loan (1 represents "Yes")
Outstanding Balance	Outstanding balance of customer
Total No of Trades	Number of times the customer has done total trades
Presence of open auto loan	Is the customer has auto loan (1 represents "Yes")

Issues with Data

1. Null values observed in variable 'Average CC Utilization in last 12 months' for 1023 applicants. We assume that these applicants don't have any prior credit history hence we impute the null values with 0.
2. Variable 'Presence of home loan' and 'Outstanding balance' has 272 records with null value. Both these variables looks correlated from the data dictionary description and we impute null values in both these columns with zero

Weight of Evidence & IV Analysis

Credit Portfolio Analysis based on available records

Who are the risky applicants

Approx. 4.2% of all issued credit cards have resulted in defaults

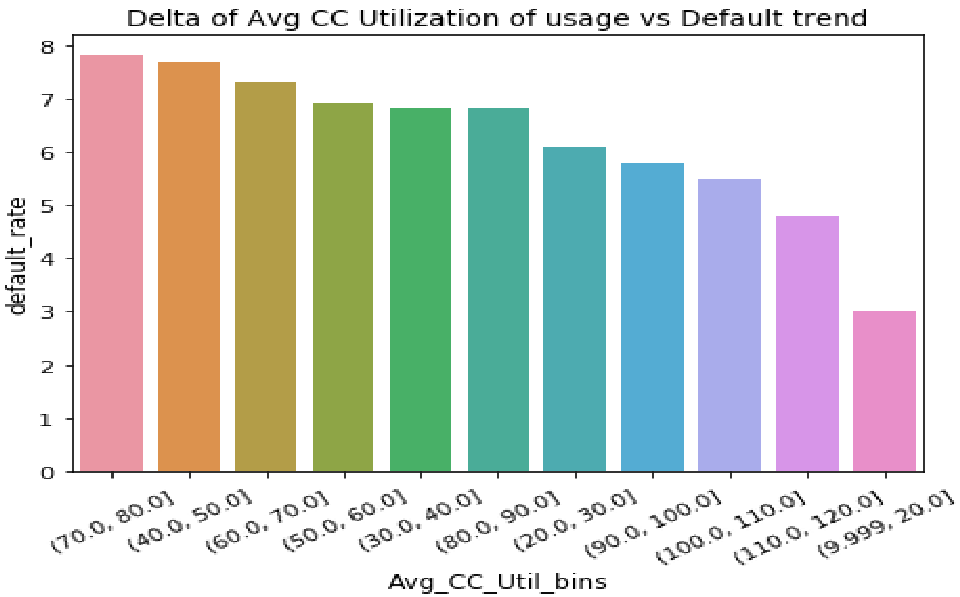
Performance Tag	Percentage
Not Defaulted	95.8 %
Defaulted	4.2%

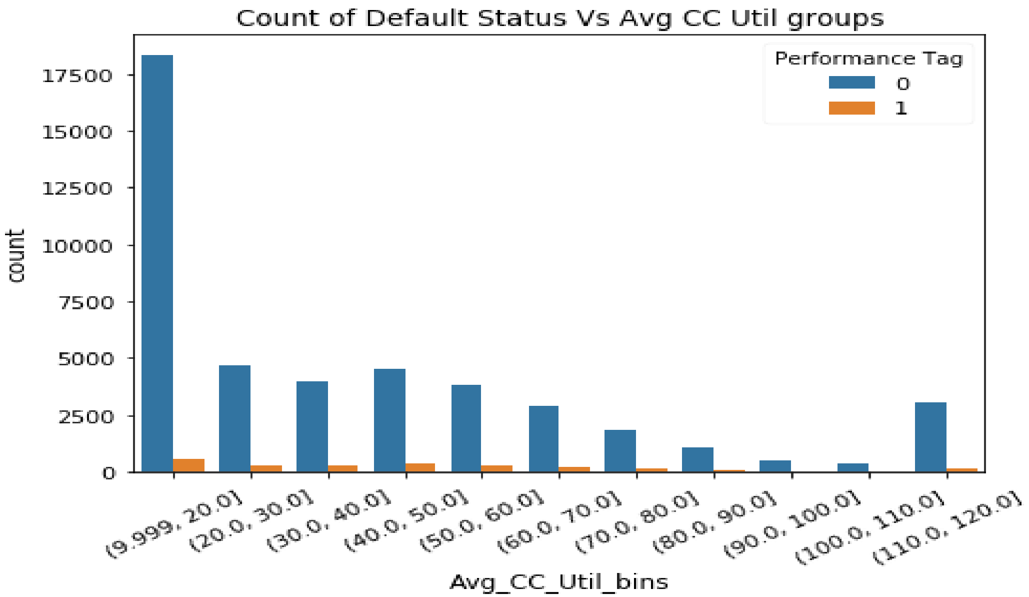
Variables	Information Value
No of times 90 DPD or worse in last 6 months	0.16265
No of times 60 DPD or worse in last 6 months	0.211263
No of times 30 DPD or worse in last 6 months	0.244237
No of times 90 DPD or worse in last 12 months	0.215644
No of times 60 DPD or worse in last 12 months	0.188225
No of times 30 DPD or worse in last 12 months	0.218599
Avgas CC Utilization in last 12 months	0.299205
No of trades opened in last 6 months	0.191861
No of trades opened in last 12 months	0.293588
No of PL trades opened in last 6 months	0.224242
No of PL trades opened in last 12 months	0.258558
No of Inquiries in last 6 months	0.11335
No of Inquiries in last 12 months	0.245238
Outstanding Balance	0.245102
Total No of Trades	0.232287

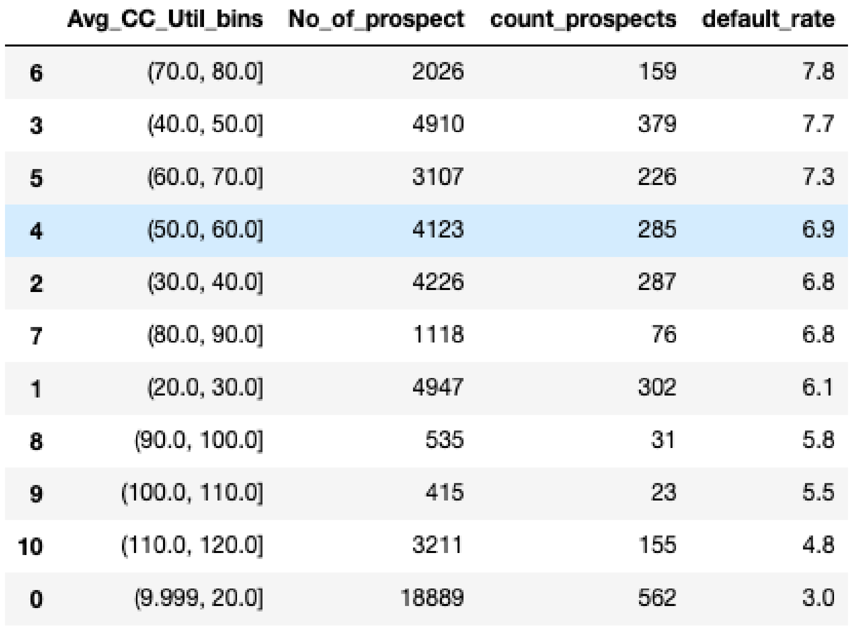
As per WOE and IV analysis, none of the variables from Demographic data have predictive value above 0.02 So we can say that applicant data alone cannot provide good insights and credit bureau data is required to further evaluate factors influencing the credit risk analysis

Exploratory Analysis

Observation - 1





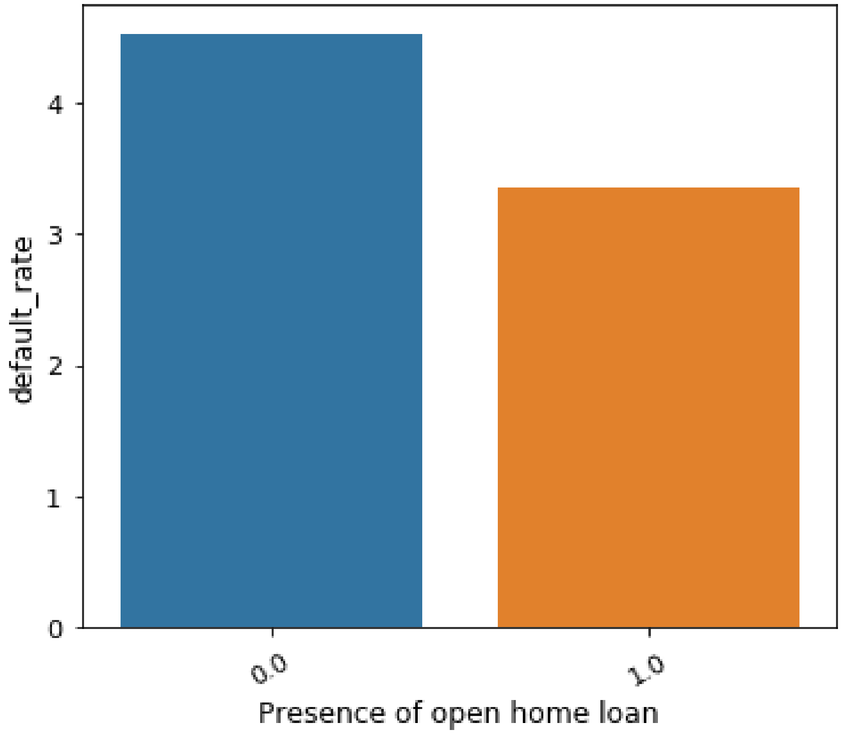


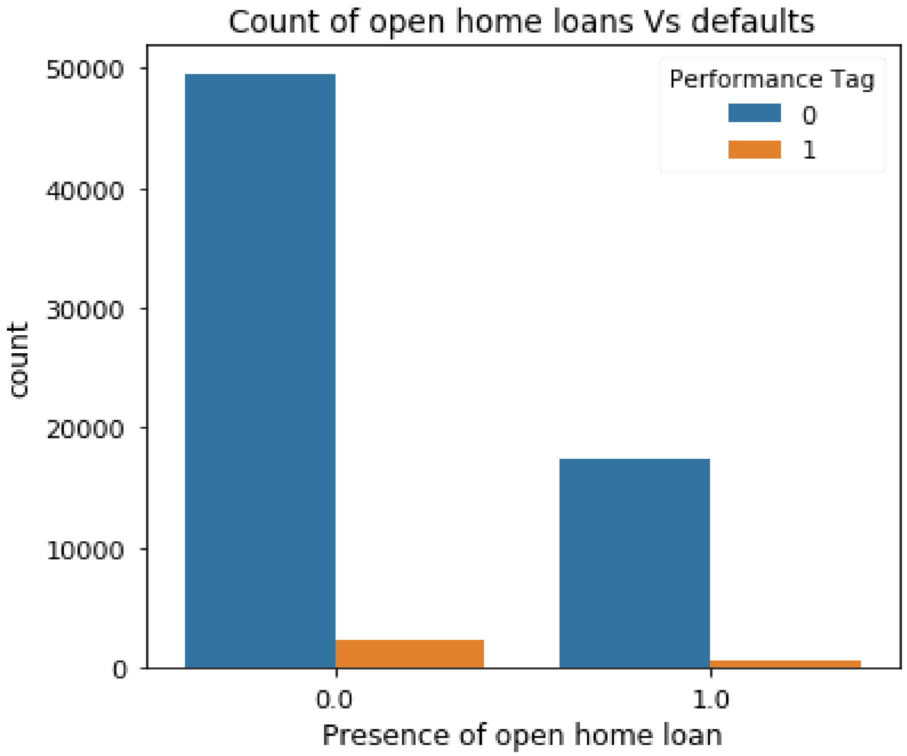
	Avg_CC_Util_bins	No_of_prospect	count_prospects	default_rate
6	(70.0, 80.0]	2026	159	7.8
3	(40.0, 50.0]	4910	379	7.7
5	(60.0, 70.0]	3107	226	7.3
4	(50.0, 60.0]	4123	285	6.9
2	(30.0, 40.0]	4226	287	6.8
7	(80.0, 90.0]	1118	76	6.8
1	(20.0, 30.0]	4947	302	6.1
8	(90.0, 100.0]	535	31	5.8
9	(100.0, 110.0]	415	23	5.5
10	(110.0, 120.0]	3211	155	4.8
0	(9.999, 20.0]	18889	562	3.0

Insight1 (Average CC Utilization in last 12 months)

Although the highest number of credit card applications are issued to customers with average CC utilization between 10-20 , From the Plot it is clear that the highest percentage of defaulters are from group where CC utilization rate between 70-80 lies. So we can see CC utilization rate in mid range i.e. 30 to 90 contributes major percent of defaults

Observation - 2



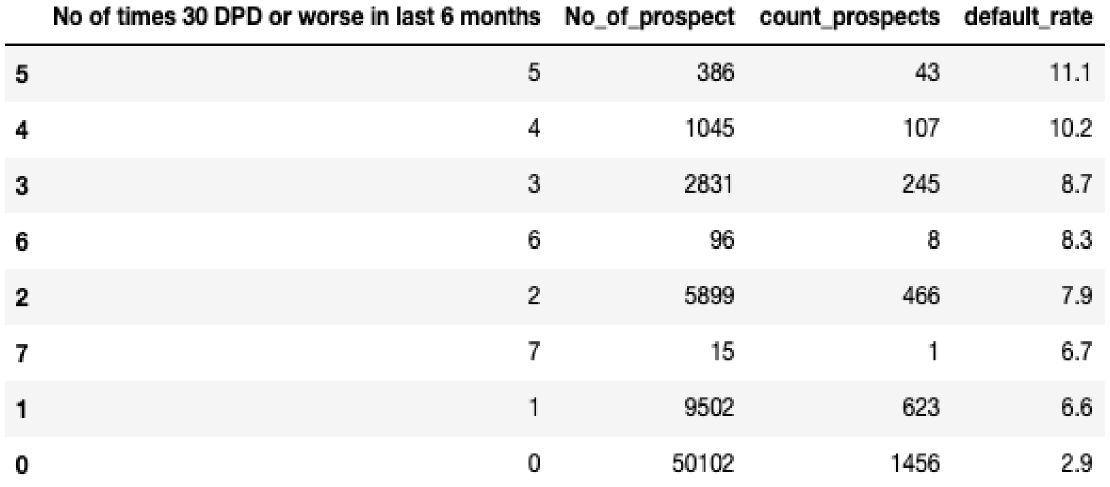


Insight2 (**Presence of open home loan**)

From the table it is clear that the highest number of CC are issued to applicants without any prior home loan . However, from the plot it is evident that they tend to default more marginally than people with home loan.

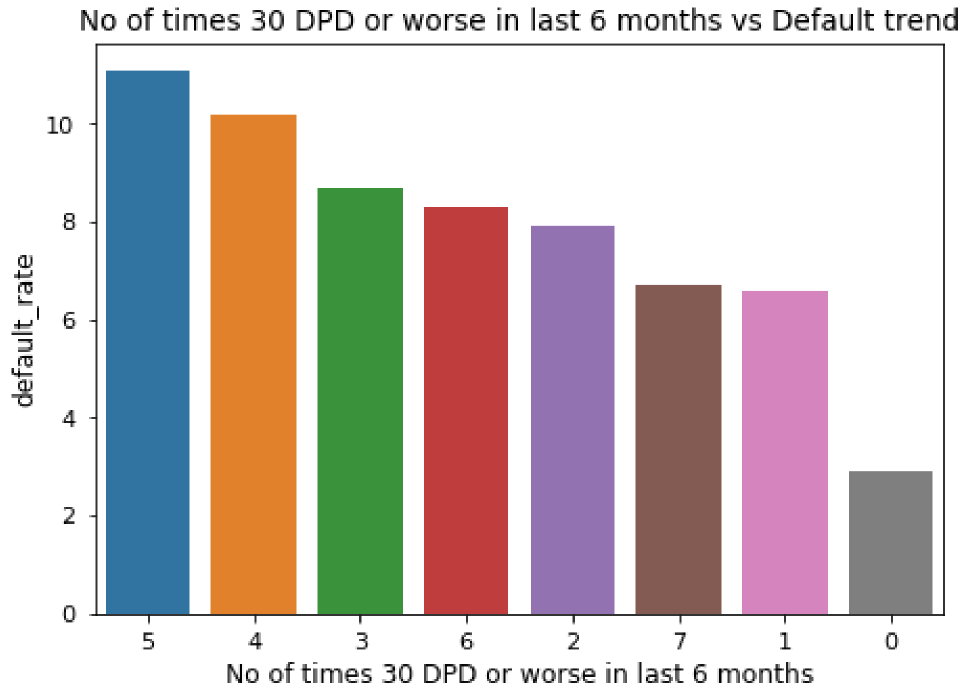
Presence of open home loan	Default prospects	No of prospects	Default rate
No	2342	51805	4.52%
Yes	607	18071	3.36%

Observation - 3

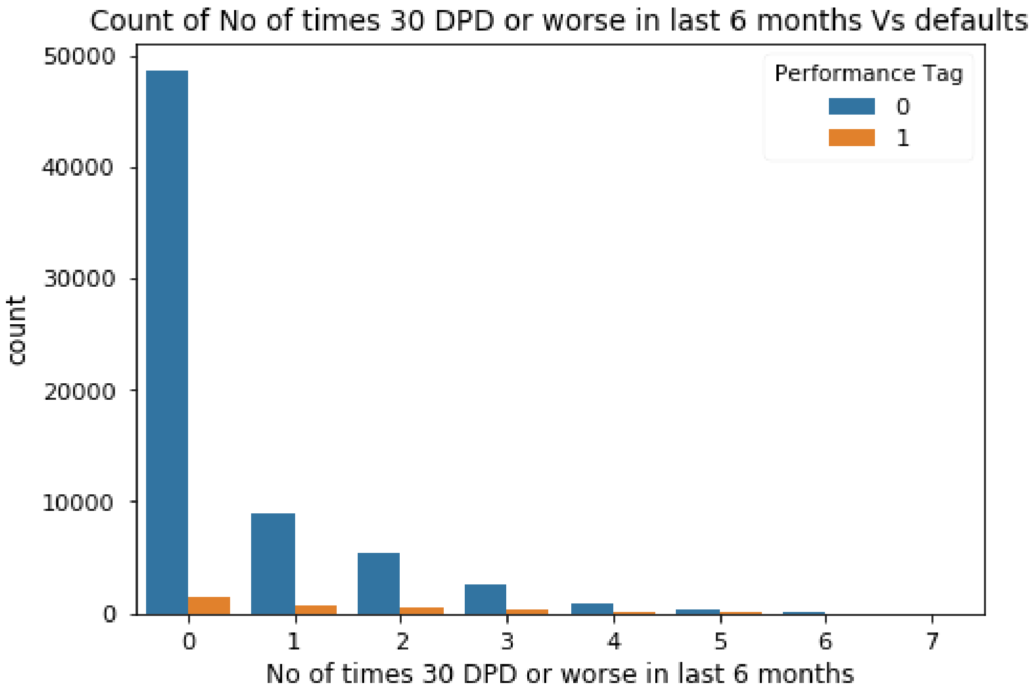


	No of times 30 DPD or worse in last 6 months	No_of_prospect	count_prospects	default_rate
5	5	386	43	11.1
4	4	1045	107	10.2
3	3	2831	245	8.7
6	6	96	8	8.3
2	2	5899	466	7.9
7	7	15	1	6.7
1	1	9502	623	6.6
0	0	50102	1456	2.9

Insight3 (**No of times 30 DPD or worse in last 6 months**)
Majority of credit cards are issued for applicants with no history of DPD in last 6 months. However, from the plot it is clear that a % of CC issued to customers with no history of DPD in past has a significantly lesser chance of resulting in credit loss.

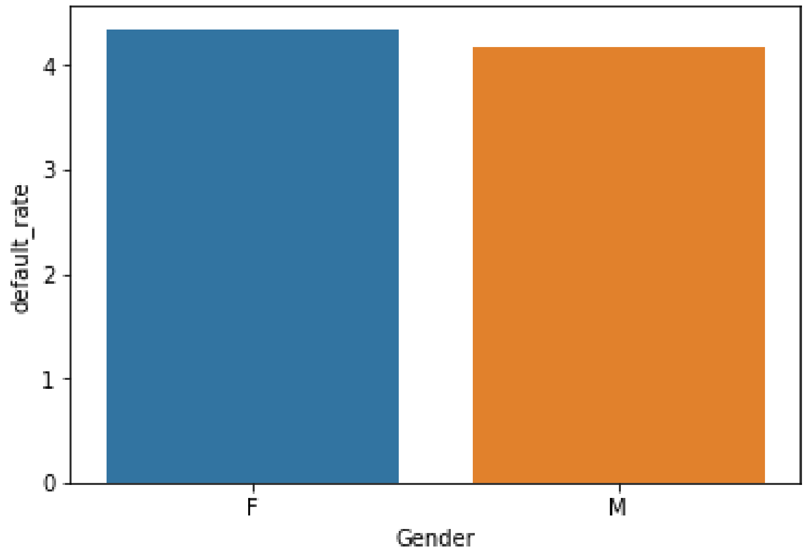


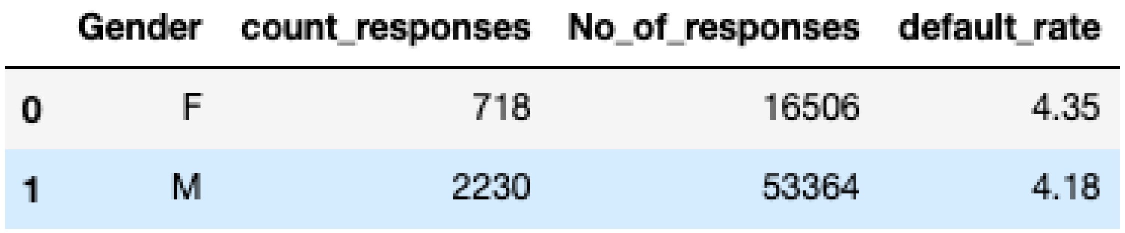
No of times 30 DPD or worse in last 6 months vs Default trend



Count of No of times 30 DPD or worse in last 6 months Vs defaults

Observation - 4

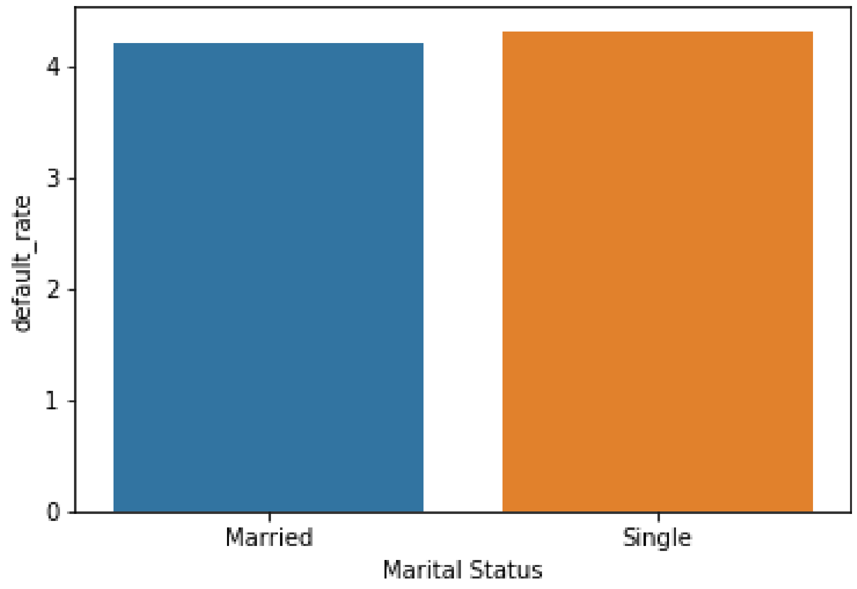


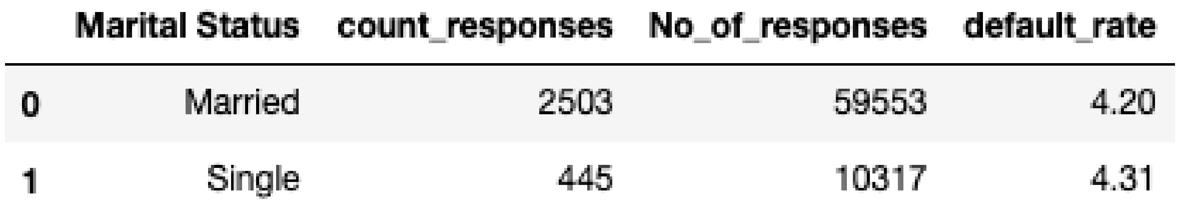


	Gender	count_responses	No_of_responses	default_rate
0	F	718	16506	4.35
1	M	2230	53364	4.18

Insight4

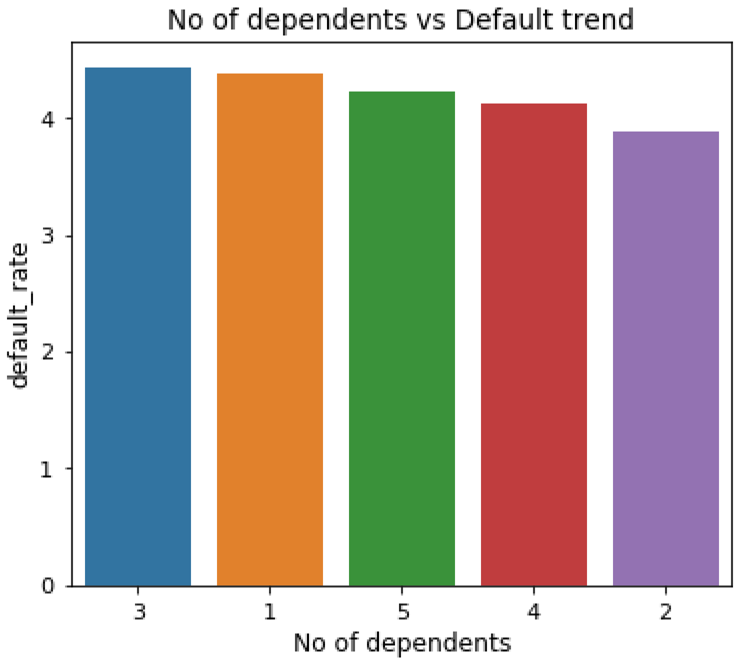
As seen from the % table, demographic features such as Gender and Marital Status do not exhibit any trend with respect to default rate. All the categories almost contribute similar number of default cases.

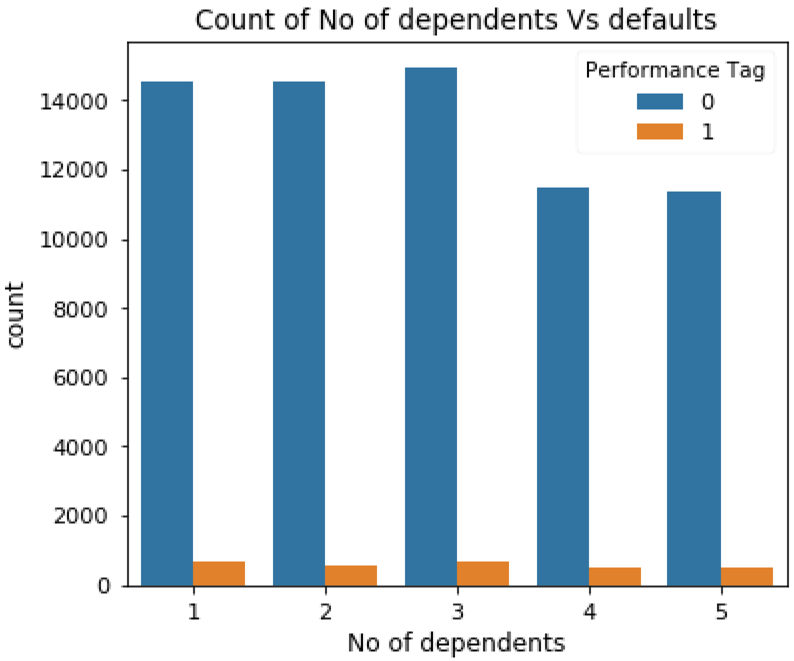


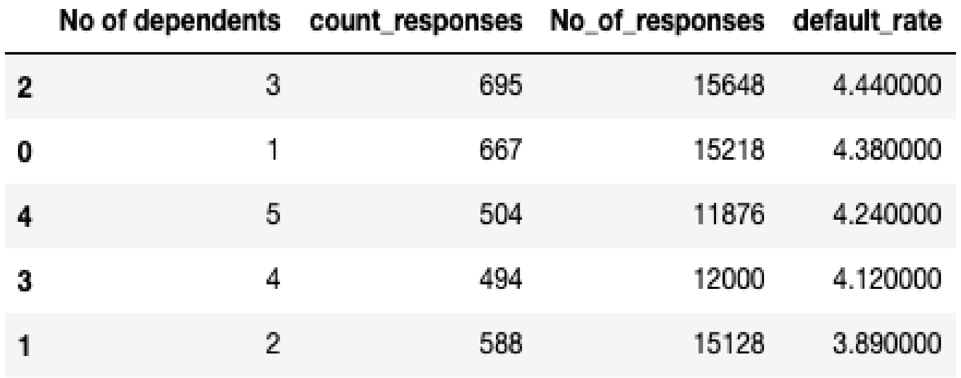


	Marital Status	count_responses	No_of_responses	default_rate
0	Married	2503	59553	4.20
1	Single	445	10317	4.31

Observation - 5


No of dependents vs Default trend


Count of No of dependents Vs defaults

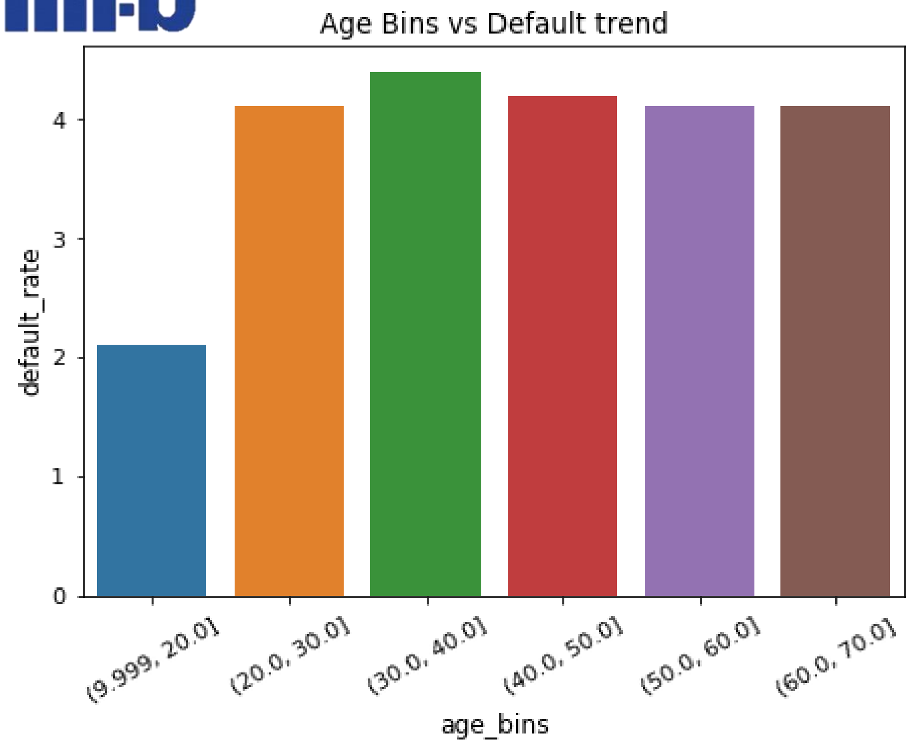


	No of dependents	count_responses	No_of_responses	default_rate
2	3	695	15648	4.440000
0	1	667	15218	4.380000
4	5	504	11876	4.240000
3	4	494	12000	4.120000
1	2	588	15128	3.890000

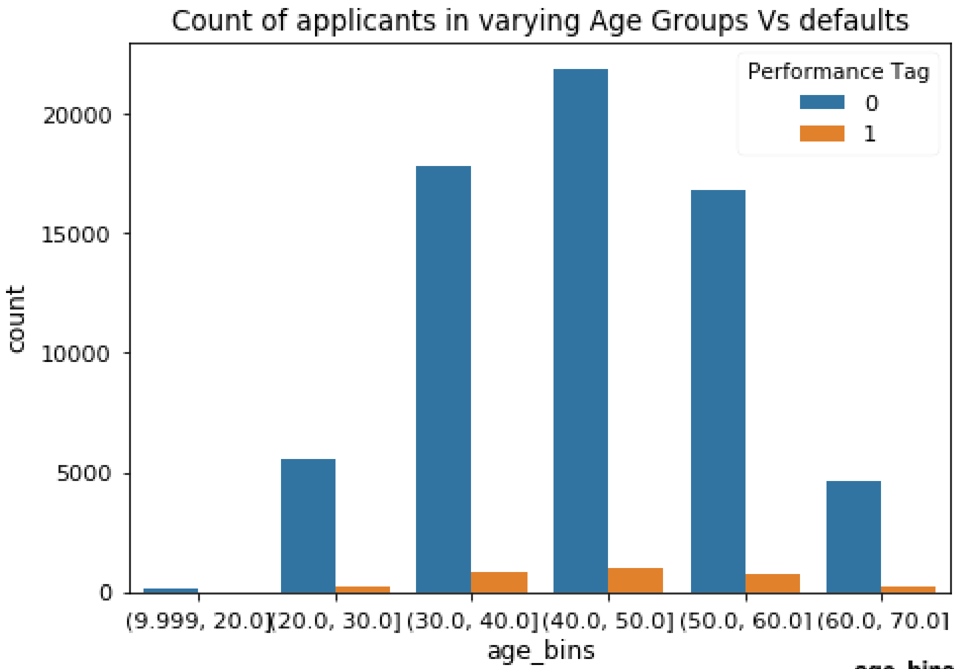
Insight5 (**No of dependents**)

From the proportionality plot , it is seen the default rate is almost same for different number of dependents. Thereby no of children does not seem to be an influencing factor for default rate

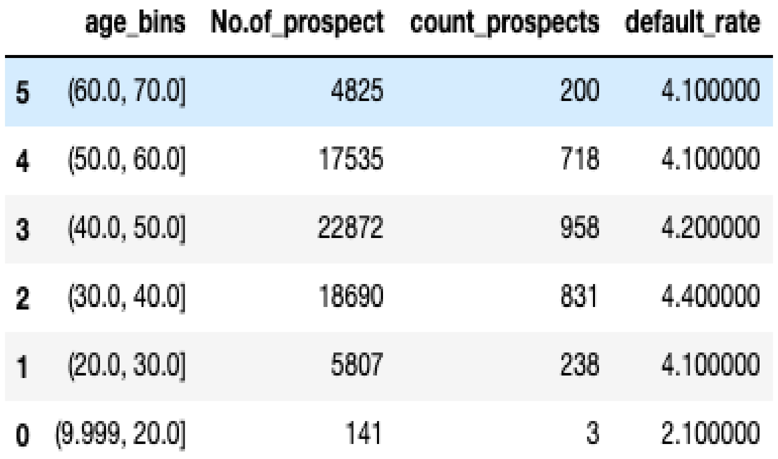
Observation - 6



Age Bins vs Default trend



Count of applicants in varying Age Groups Vs defaults

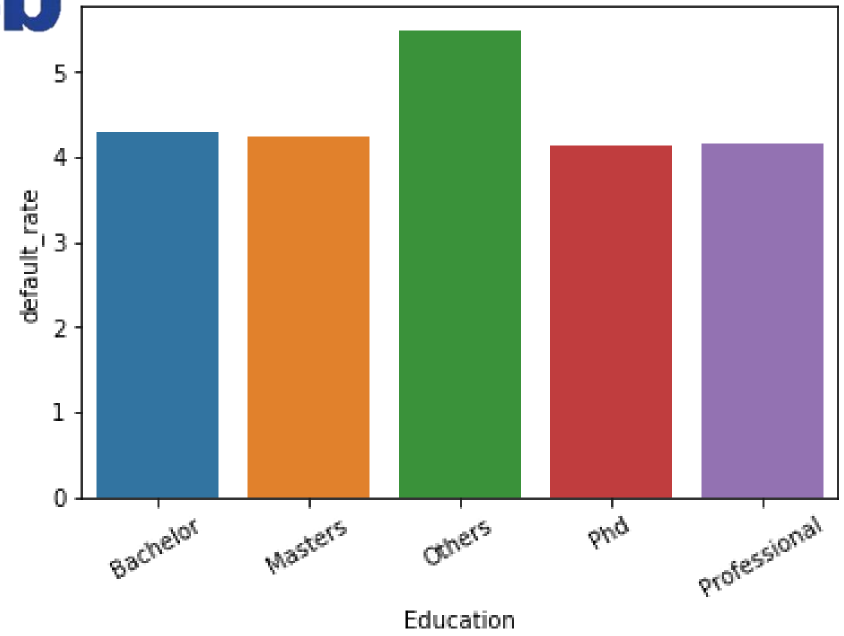


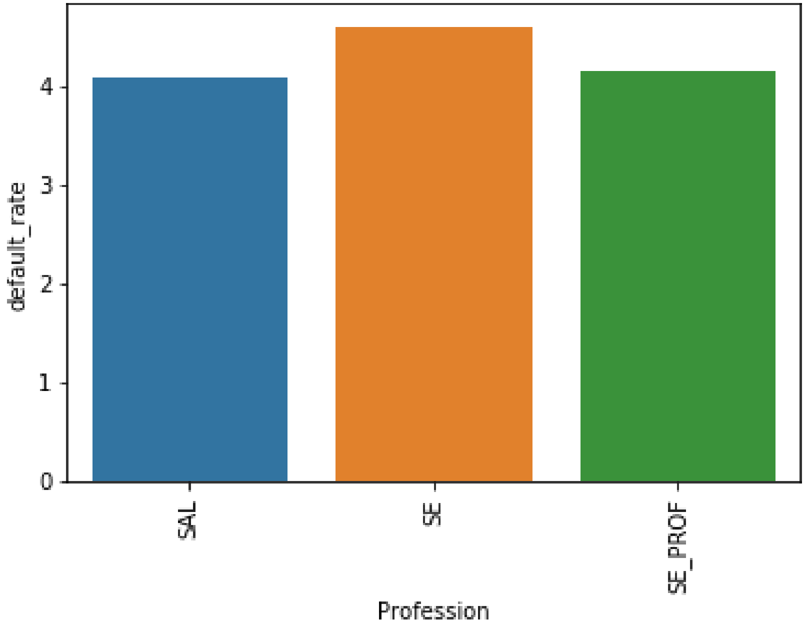
	age_bins	No.of_prospect	count_prospects	default_rate
5	(60.0, 70.0]	4825	200	4.100000
4	(50.0, 60.0]	17535	718	4.100000
3	(40.0, 50.0]	22872	958	4.200000
2	(30.0, 40.0]	18690	831	4.400000
1	(20.0, 30.0]	5807	238	4.100000
0	(9.999, 20.0]	141	3	2.100000

Insight6 (Age)

From the table it is clear that highest frequency of CC belong to customers in the age group of 40 to 50. However, the default percentage is almost in same range for all the age groups from 20 to 70. There are not many applicants in 10 to 20 range since the minimum age requirement for issuing a credit card to an applica is 18 years.

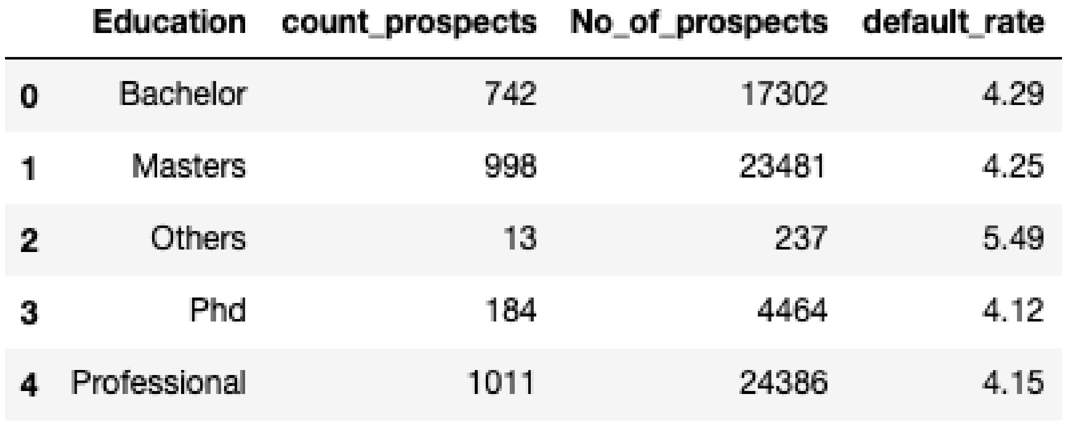
Observation - 7



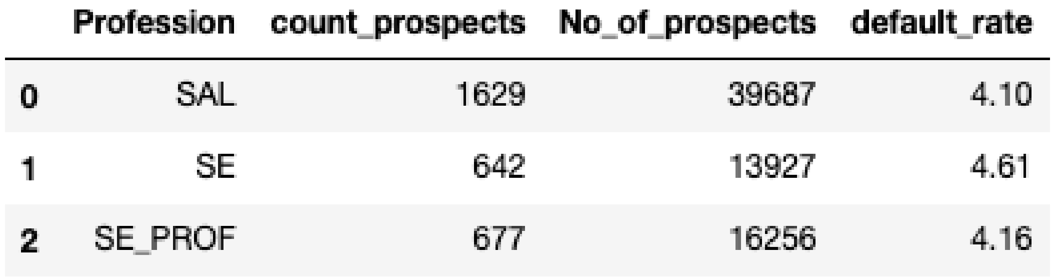


Insight7 (**Education & Profession**)

We can see there is no trend seen on plotting demographic factors such as education and profession against the default rate.

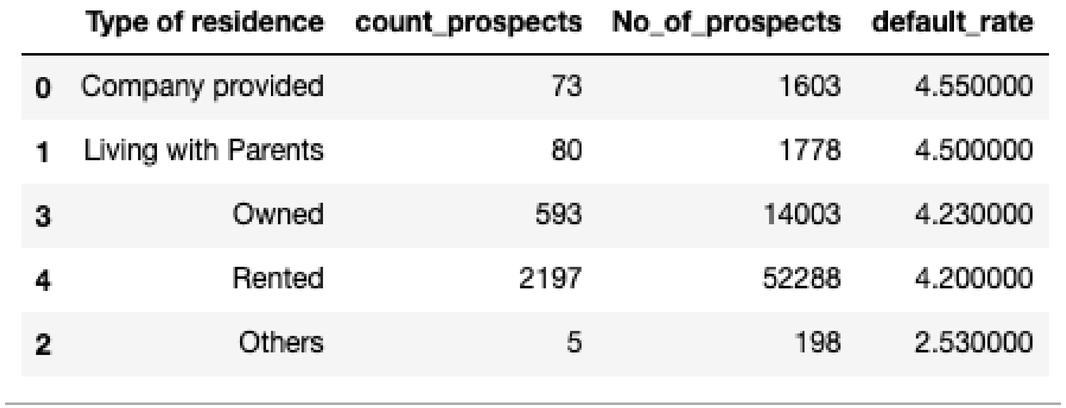


	Education	count_prospects	No_of_prospects	default_rate
0	Bachelor	742	17302	4.29
1	Masters	998	23481	4.25
2	Others	13	237	5.49
3	Phd	184	4464	4.12
4	Professional	1011	24386	4.15



	Profession	count_prospects	No_of_prospects	default_rate
0	SAL	1629	39687	4.10
1	SE	642	13927	4.61
2	SE_PROF	677	16256	4.16

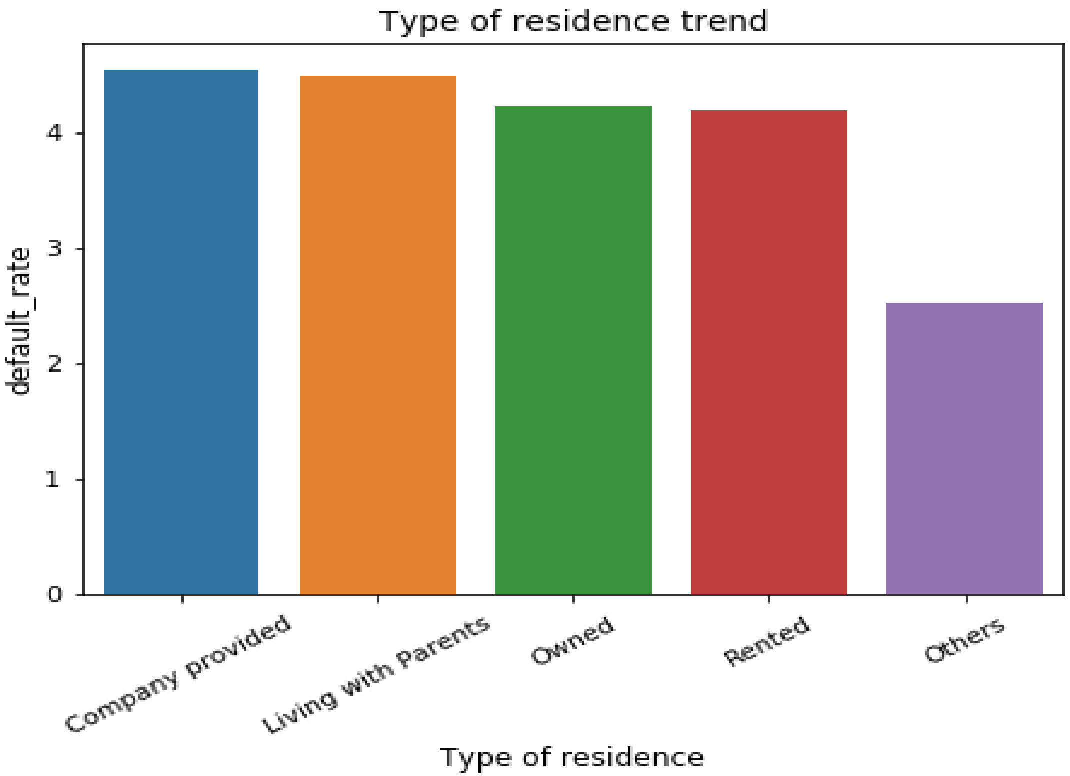
Observation - 8

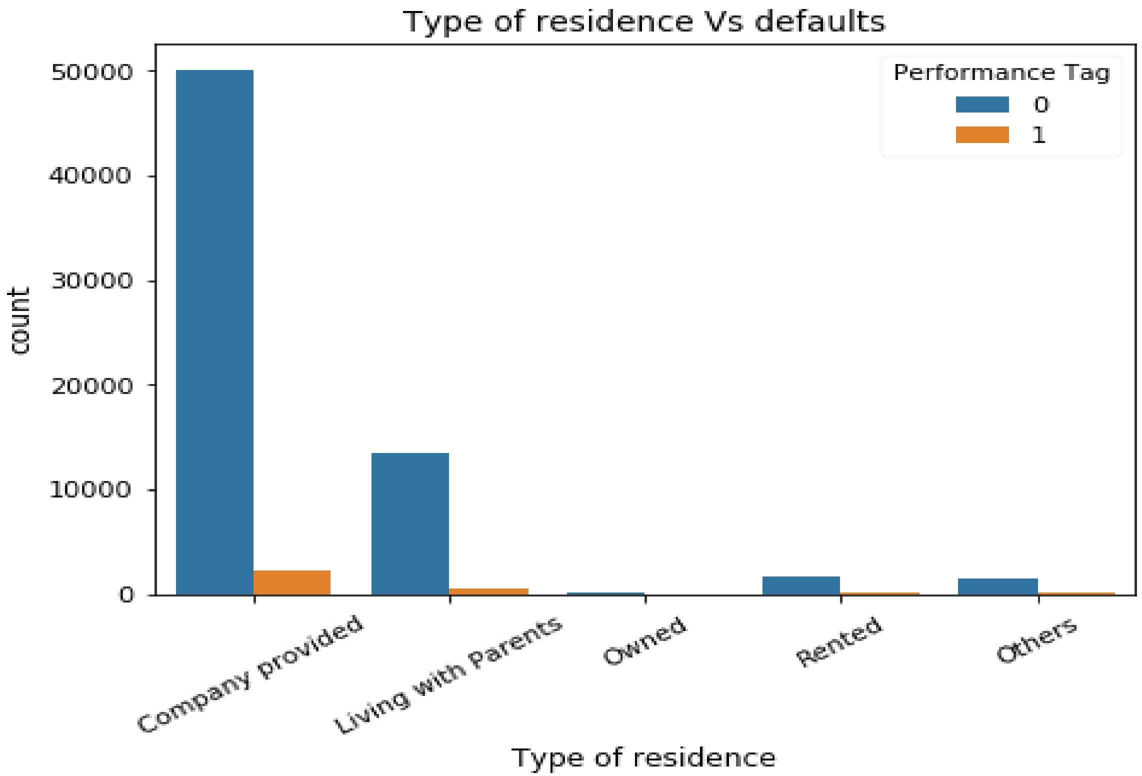


	Type of residence	count_prospects	No_of_prospects	default_rate
0	Company provided	73	1603	4.550000
1	Living with Parents	80	1778	4.500000
3	Owned	593	14003	4.230000
4	Rented	2197	52288	4.200000
2	Others	5	198	2.530000

Insight8(**Type of residence**):

People who live in rented, owned or company provided accommodation tend to default in similar manner

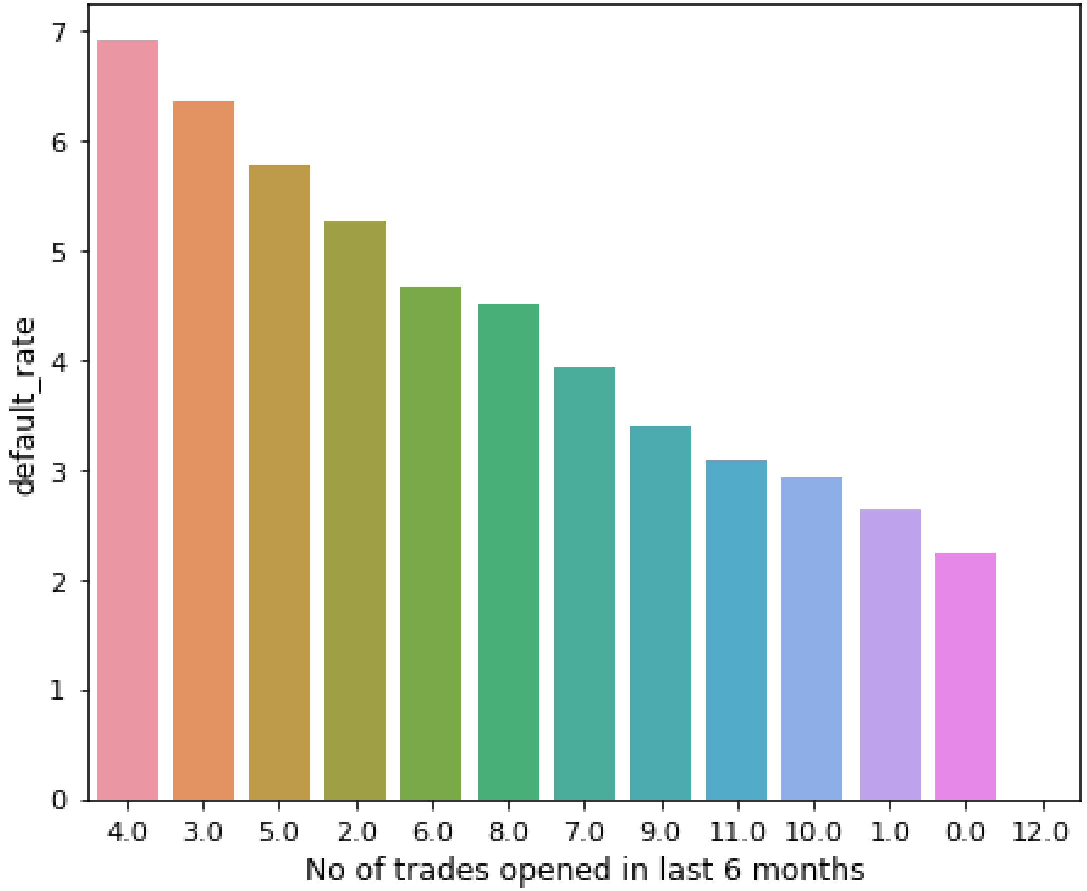


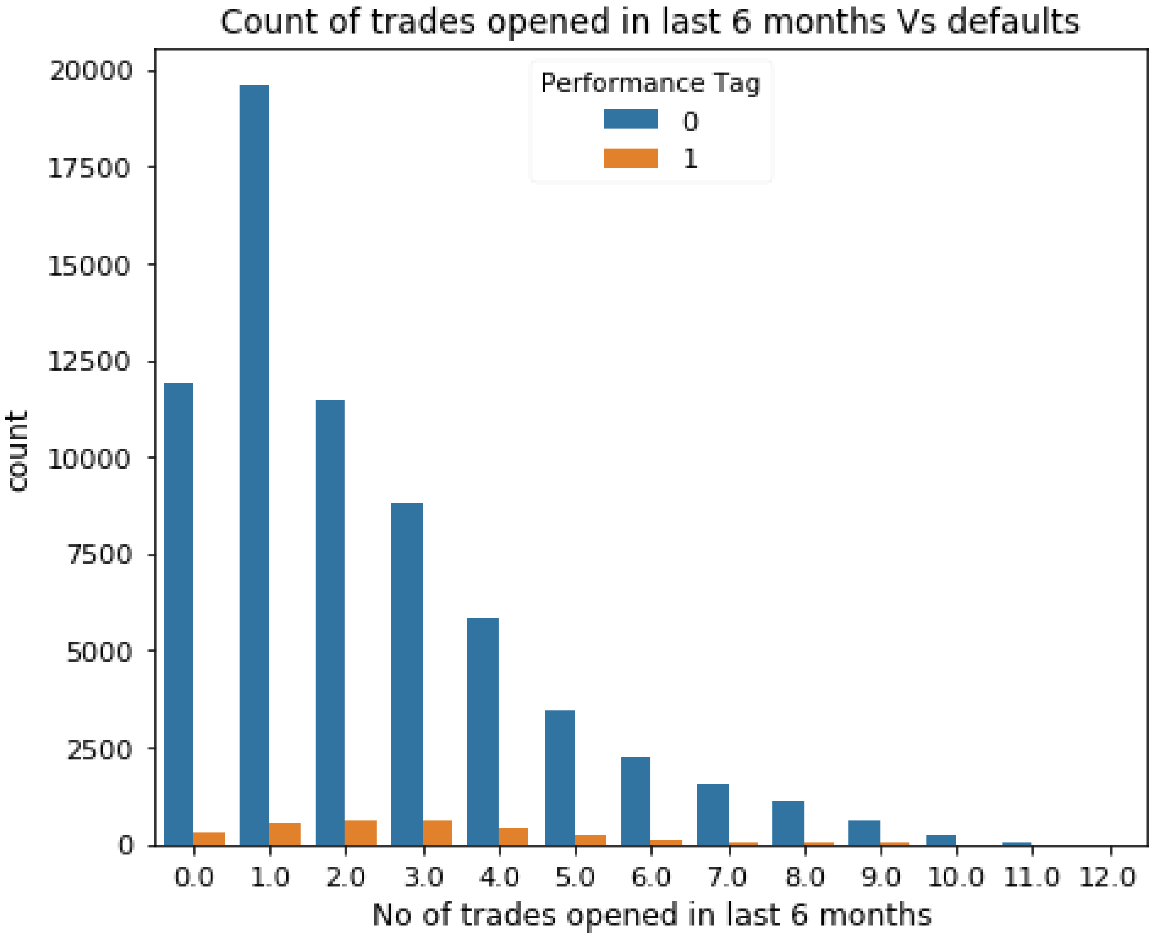


Observation - 9

Insight9 (Number of trades opened in last 6 months):

Applicants who have opened trades in last 6 months tend to default more than customers who have not done any trading since 6 months

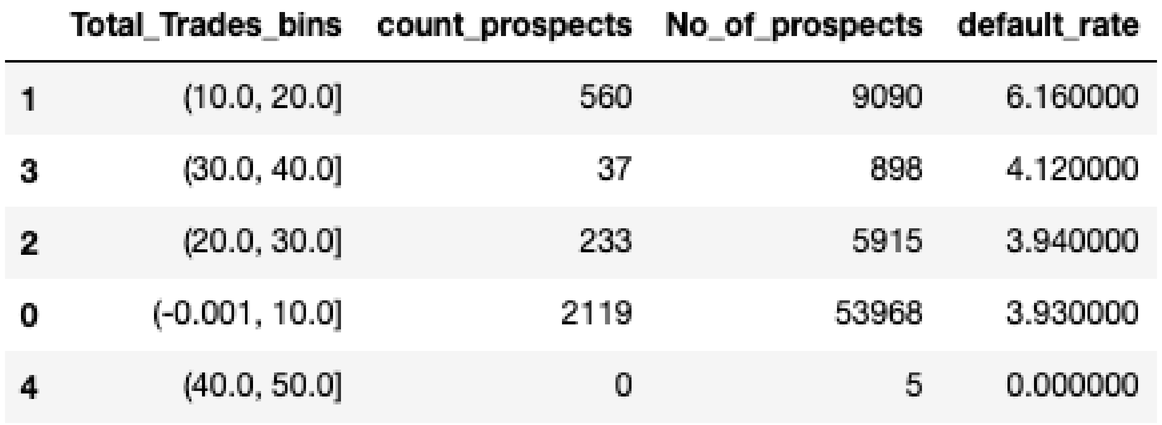



Count of trades opened in last 6 months Vs defaults

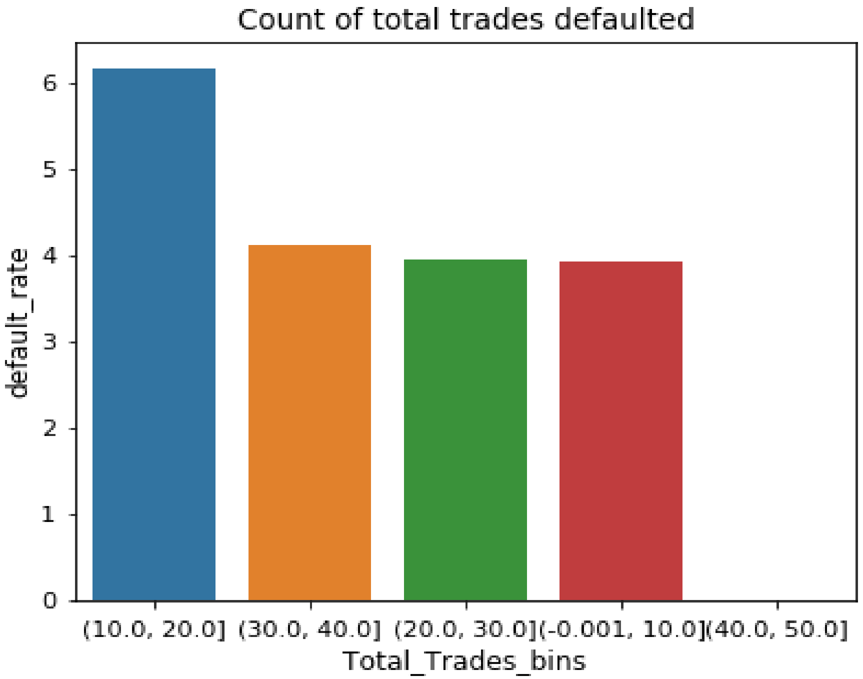
Observation - 10

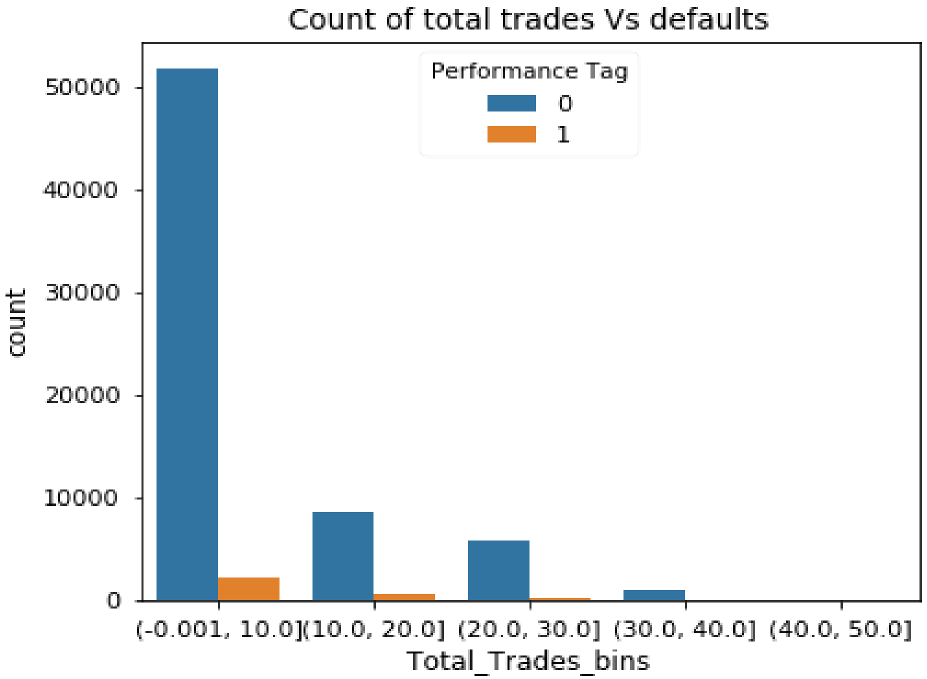
Insight10 (Total No of Trades):

Applicants who are having lesser no of trades are mainly issued with credit card compared to those having higher no of trades. From the plot, more % of defaults are seen in the bin of 10-20

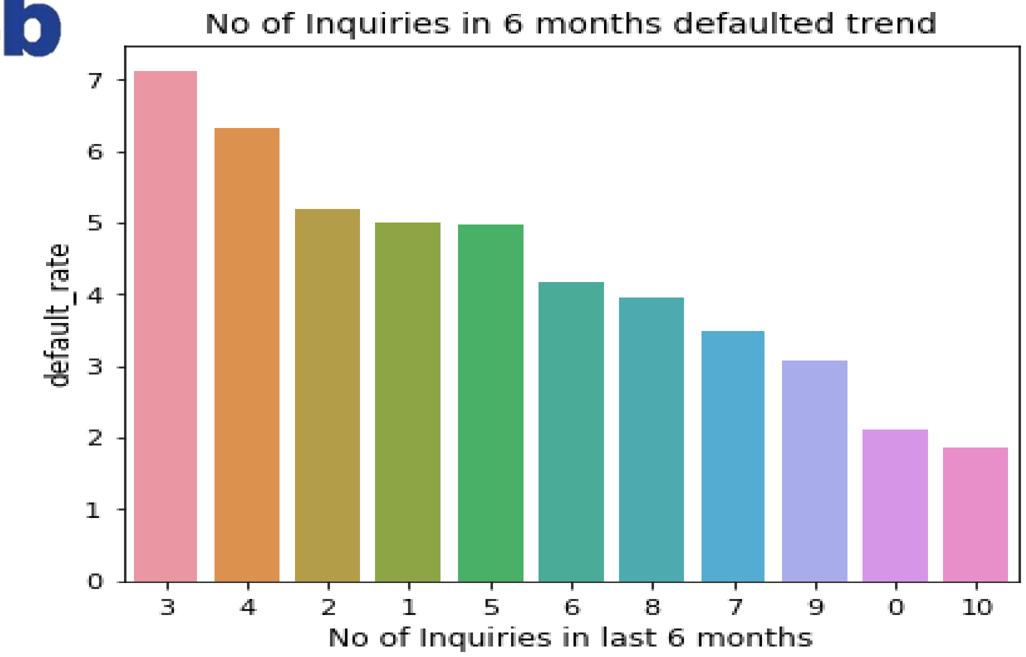


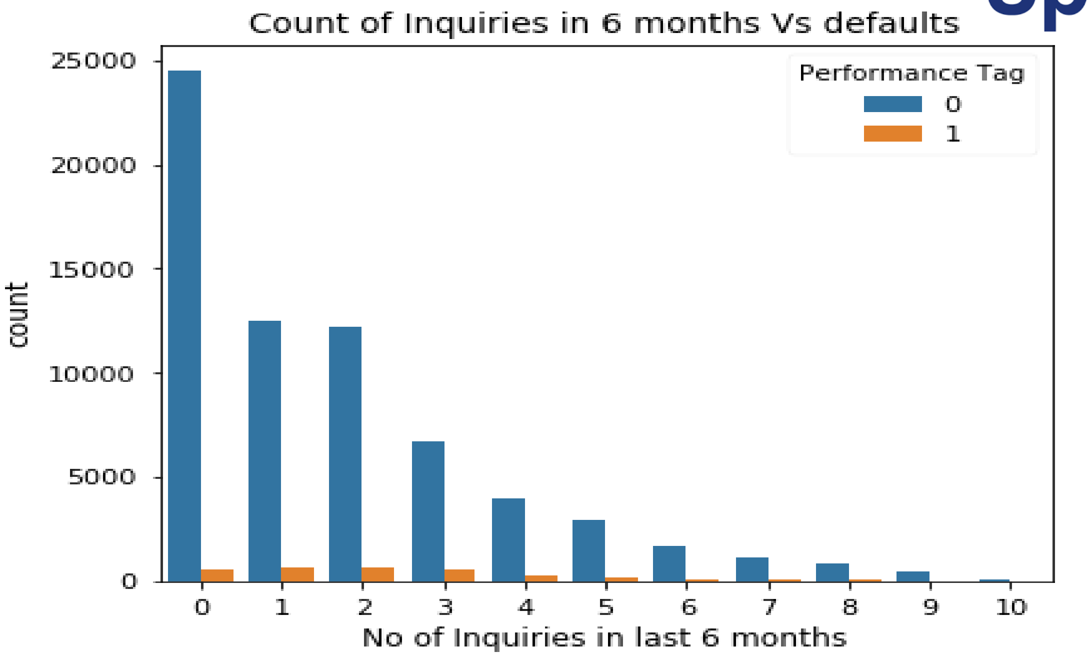
	Total_Trades_bins	count_prospects	No_of_prospects	default_rate
1	(10.0, 20.0]	560	9090	6.160000
3	(30.0, 40.0]	37	898	4.120000
2	(20.0, 30.0]	233	5915	3.940000
0	(-0.001, 10.0]	2119	53968	3.930000
4	(40.0, 50.0]	0	5	0.000000

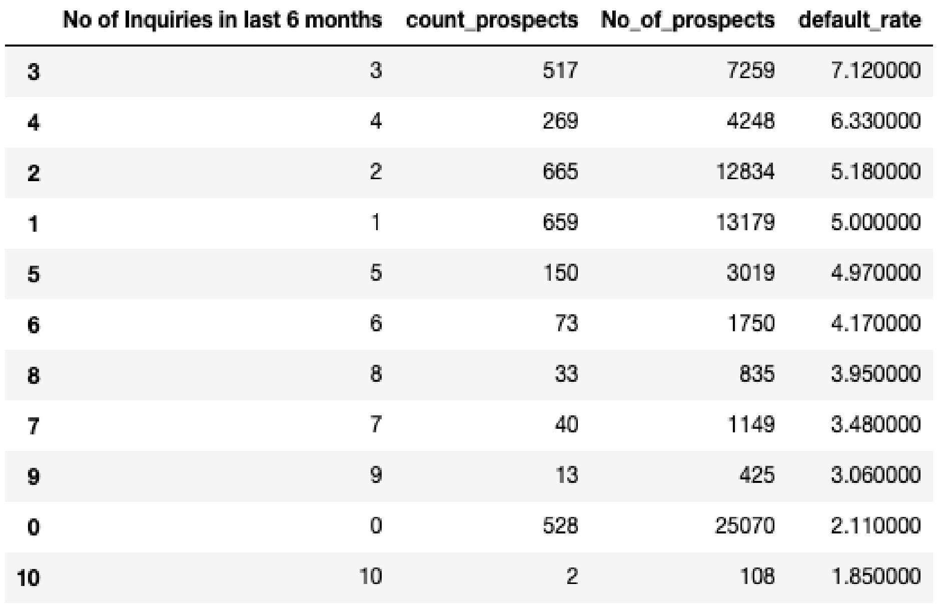




Observation - 11





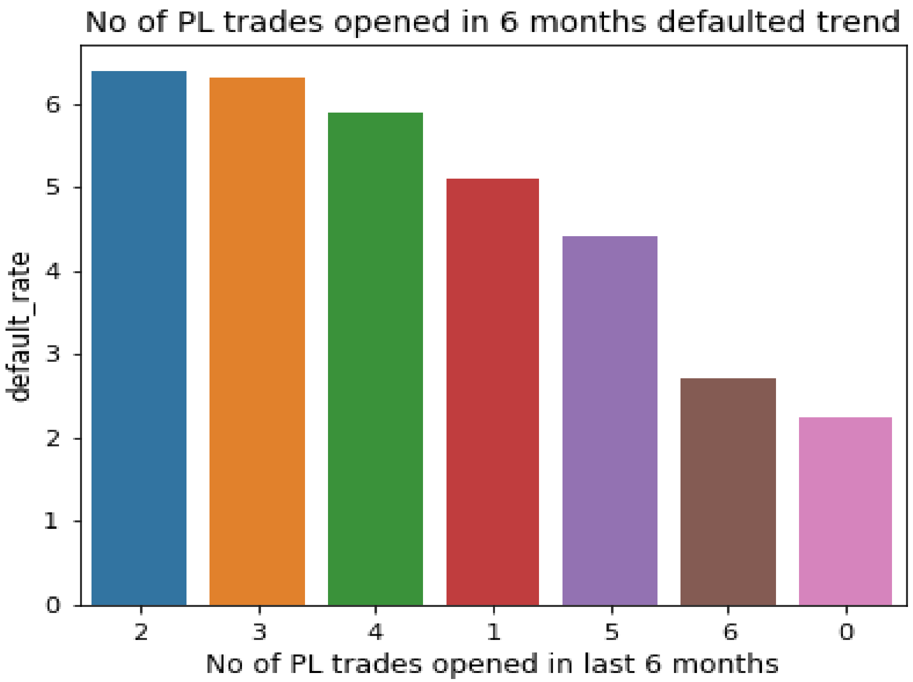


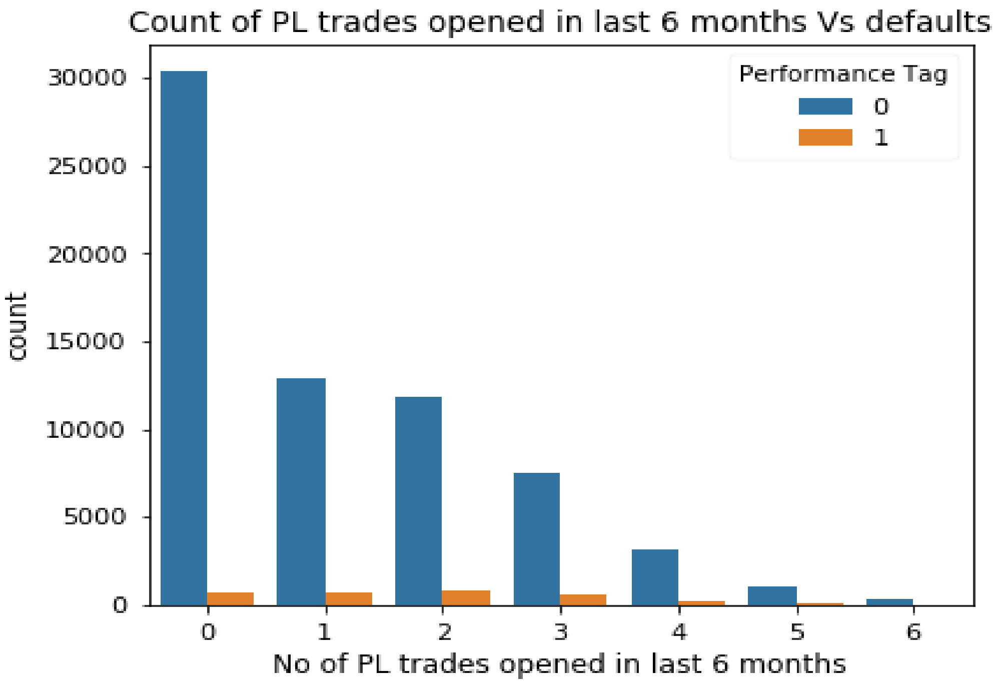
	No of Inquiries in last 6 months	count_prospects	No_of_prospects	default_rate
3	3	517	7259	7.120000
4	4	269	4248	6.330000
2	2	665	12834	5.180000
1	1	659	13179	5.000000
5	5	150	3019	4.970000
6	6	73	1750	4.170000
8	8	33	835	3.950000
7	7	40	1149	3.480000
9	9	13	425	3.060000
0	0	528	25070	2.110000
10	10	2	108	1.850000

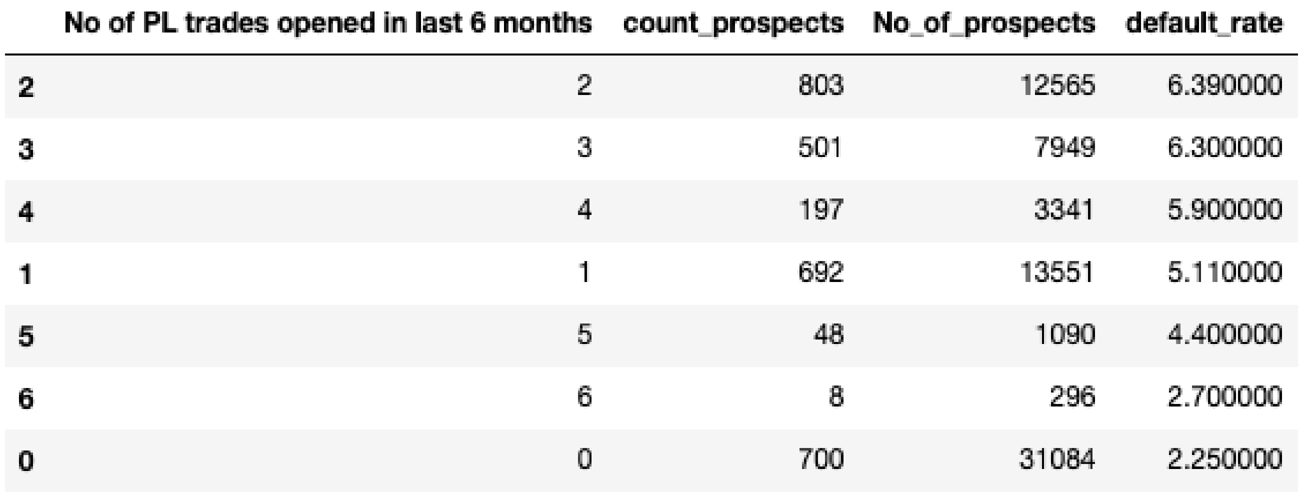
Insight11 (No of Inquiries in last 6 months):

Applicants having no inquiry in past 6 months are more in number to have been issued with credit card than those having inquiries. From the plot, more % of defaults are seen for people having inquired thrice. This seems to be an influencing factor from the plot

Observation - 12


No of PL trades opened in 6 months defaulted trend


Count of PL trades opened in last 6 months Vs defaults



	No of PL trades opened in last 6 months	count_prospects	No_of_prospects	default_rate
2	2	803	12565	6.390000
3	3	501	7949	6.300000
4	4	197	3341	5.900000
1	1	692	13551	5.110000
5	5	48	1090	4.400000
6	6	8	296	2.700000
0	0	700	31084	2.250000

Insight12 (No of PL trades opened in last 6 months):

Applicants with no PL trades opened in past 6 months are more in number to have issued with credit card than those having opened PL trades. From the plot, more % of defaults are seen for people having opened 2 to 4 PL trades. This seems to be an influencing factor from the plot

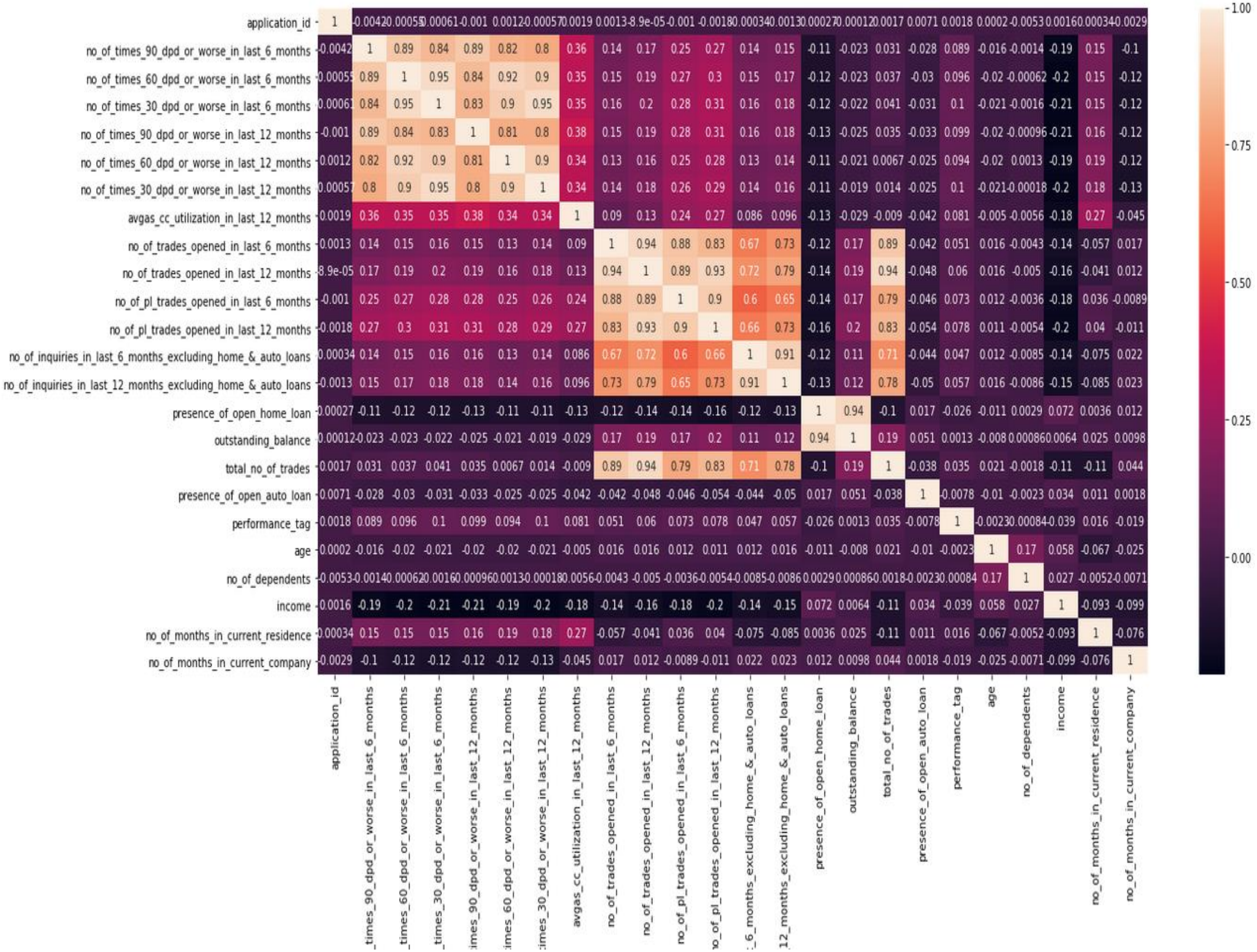
Observation - 13

1. All the DPD (Day Past Due) variables are highly correlated.

2. Number of Trades, the Total number of trades in last 6 & 12 months, Number of PL trades in last 6 and 12 months are highly correlated.

3. Number of Inquiries in last 6 and 12 months excluding home & auto loans are highly correlated.

4. Number of Inquiries in the last 12 months excluding home & auto loans are highly correlated with Number of trades open in the last 12 months and Total number of trades



Observation - 14

1. The default rate is 4.22 % and it shows that the predicted value "performance tag ” is highly imbalanced, This will cause our classification models to have low predictive accuracy for the minority class. This means, our predictive model will not be able to predict customers who might default with high accuracy.

2. We will be doing the Over-sampling the minority class, this will create synthetic (not duplicate) samples of the minority class. Hence making the minority class equal to the majority class.

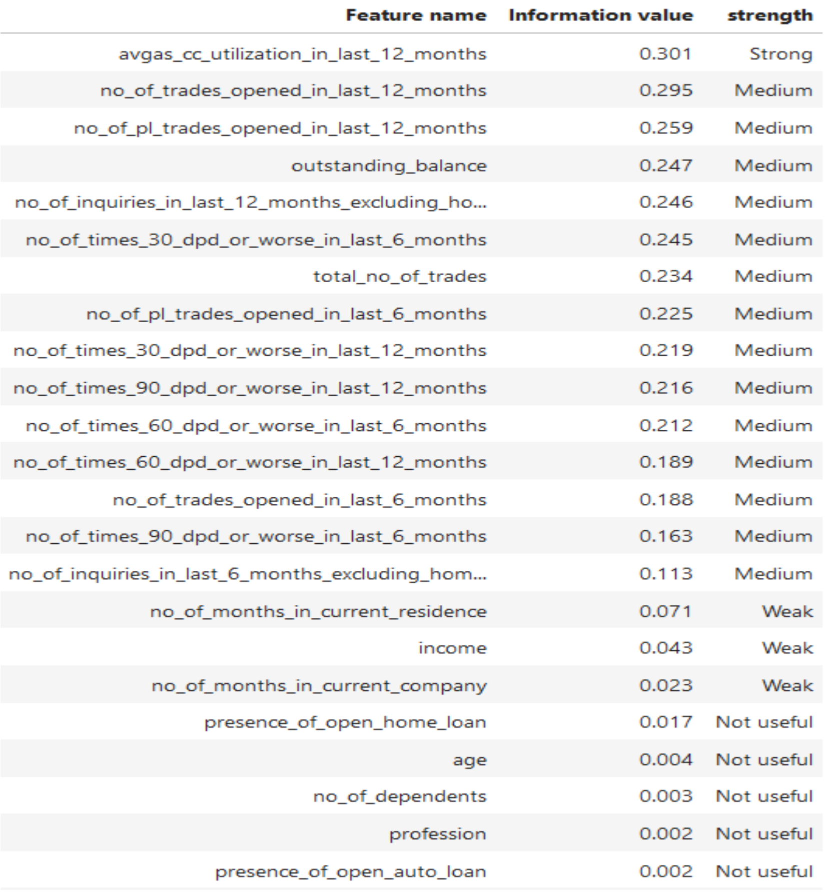
3. We have observed that merge master data contains 1425 Null value for "performance tag" and these are the rejected applicant. We will be creating a separate data set for these applicants and compare the score with defaulted applicants.

4. We can see 3 applicant’s data duplicated in the master data, since this amounts to a very negligible percent and won’t lead to any data loss we decide to remove the duplicate applicant data from the master data

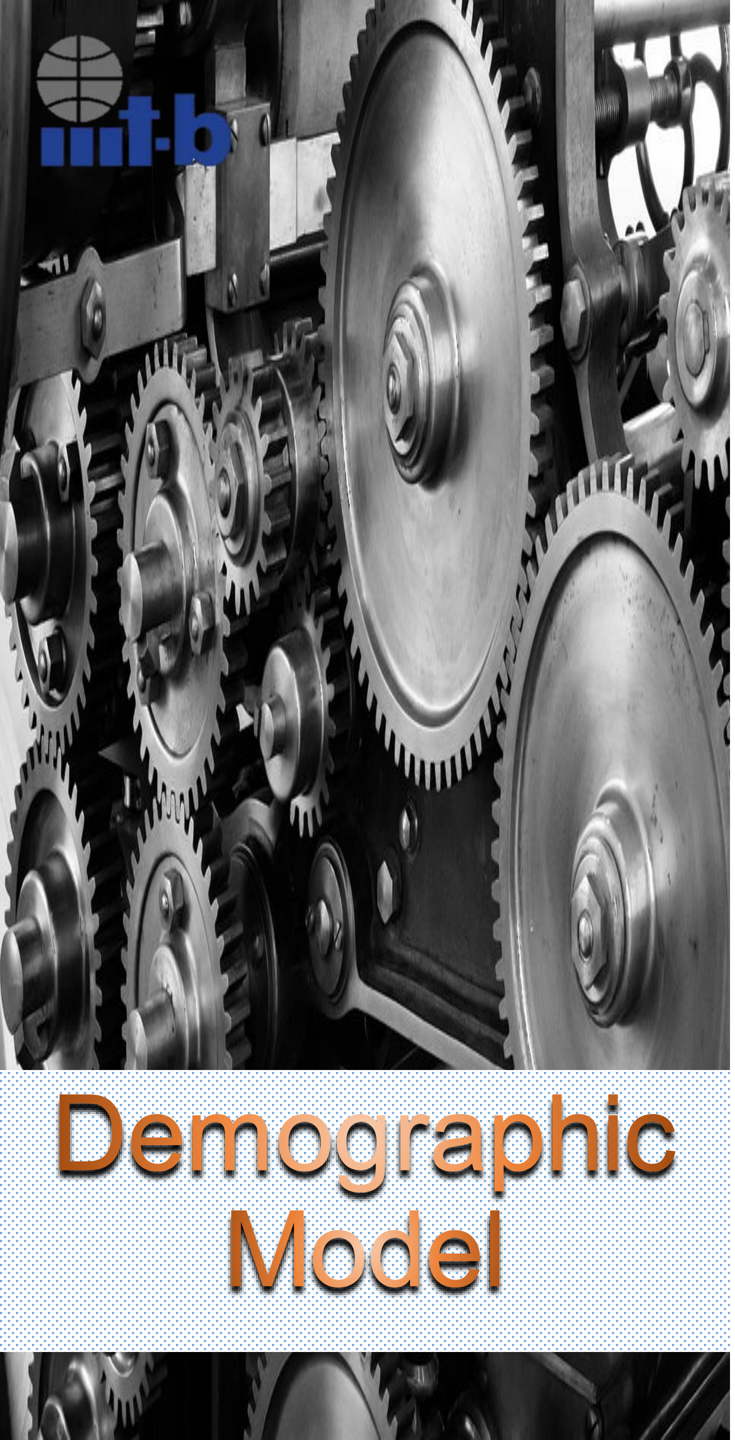
5. We will be building 3 models: Merge data set without WOE, Merge data set with WOE and demographic data set.

Analysis Outcome

These are the important predictor found using the weight of evidence analysis and we found almost all the variables were significant during EDA. We will only consider the "strong & medium" predictors for model building.



Feature name	Information value	strength
avgas_cc_utilization_in_last_12_months	0.301	Strong
no_of_trades_opened_in_last_12_months	0.295	Medium
no_of_pl_trades_opened_in_last_12_months	0.259	Medium
outstanding_balance	0.247	Medium
no_of_inquiries_in_last_12_months_excluding_ho...	0.246	Medium
no_of_times_30_dpd_or_worse_in_last_6_months	0.245	Medium
total_no_of_trades	0.234	Medium
no_of_pl_trades_opened_in_last_6_months	0.225	Medium
no_of_times_30_dpd_or_worse_in_last_12_months	0.219	Medium
no_of_times_90_dpd_or_worse_in_last_12_months	0.216	Medium
no_of_times_60_dpd_or_worse_in_last_6_months	0.212	Medium
no_of_times_60_dpd_or_worse_in_last_12_months	0.189	Medium
no_of_trades_opened_in_last_6_months	0.188	Medium
no_of_times_90_dpd_or_worse_in_last_6_months	0.163	Medium
no_of_inquiries_in_last_6_months_excluding_hom...	0.113	Medium
no_of_months_in_current_residence	0.071	Weak
income	0.043	Weak
no_of_months_in_current_company	0.023	Weak
presence_of_open_home_loan	0.017	Not useful
age	0.004	Not useful
no_of_dependents	0.003	Not useful
profession	0.002	Not useful
presence_of_open_auto_loan	0.002	Not useful



Demographic Model

The logistic model built on demographic variables yields accuracy of 56% which is as much of a random model.

From the table , we can see even the information value of demographic variables is very low. As per heuristic, variables having IV < 0.02 are not suitable for prediction.

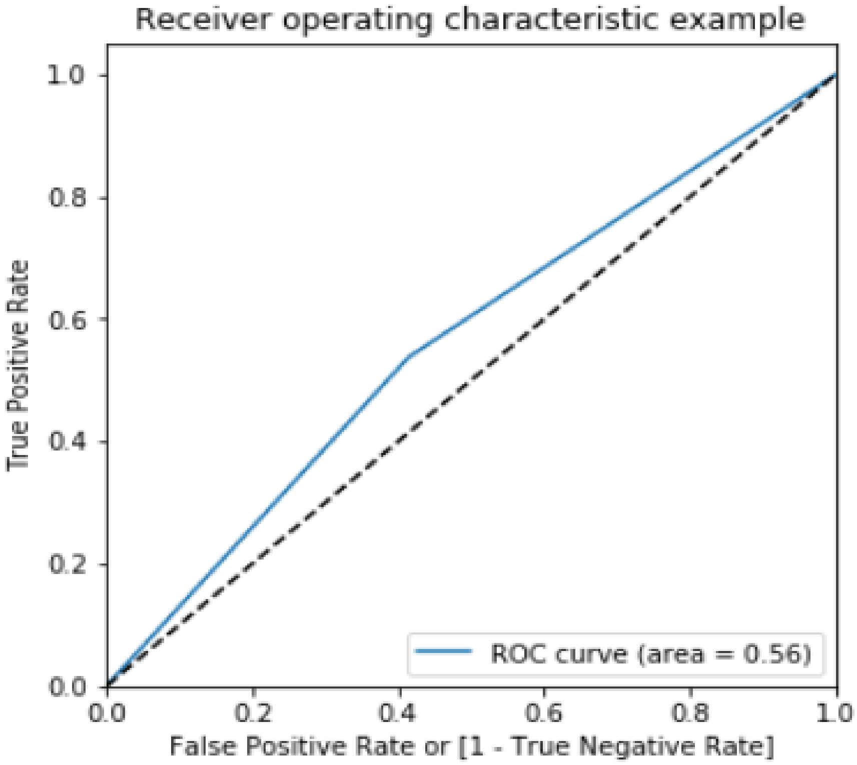
We can deduce here that demographic variables alone does not have enough information and we need to build model on merged dataset of demographic and credit data to get hidden insights

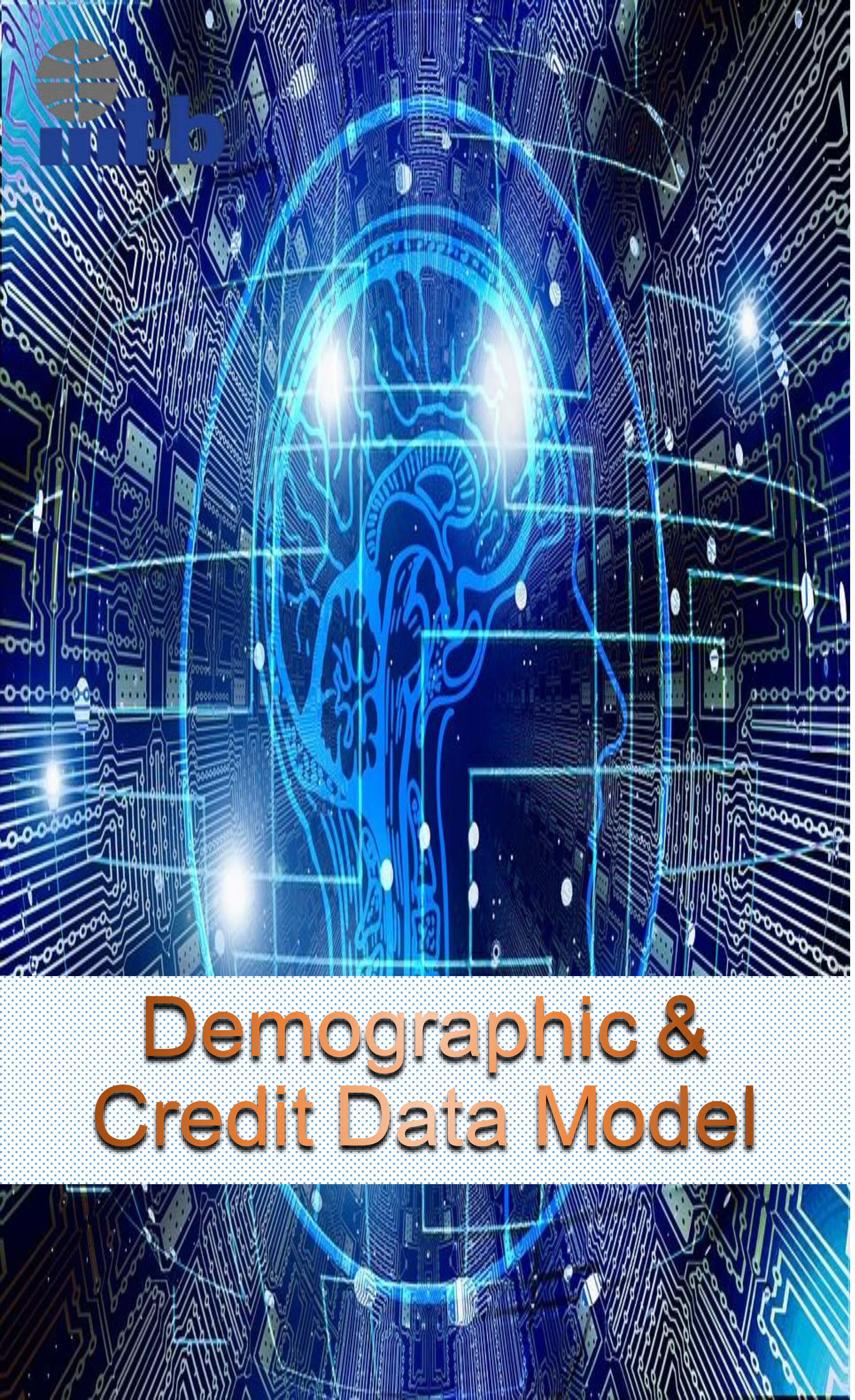
Variables	IV
Application ID	0.001487
Age	0.004168
Gender	0.000320
Marital Status	0.000093
No of dependents	0.002653
Income	0.042834
Education	0.000539
Profession	0.002289
Type of residence	0.000918
No of months in current residence	0.070902
No of months in current company	0.022714

Evaluation Matrics- Demographic data

1. These are the evaluation metrics obtained from the model built on balanced data set.
2. The model built using demographic data is not enough to predict the applicant who will default as it's having very less metrics values

Evaluation Matrics	Value
Accuracy	56%
Precision	57%
Recall	54%
F1-Score	55%
AUC	56%





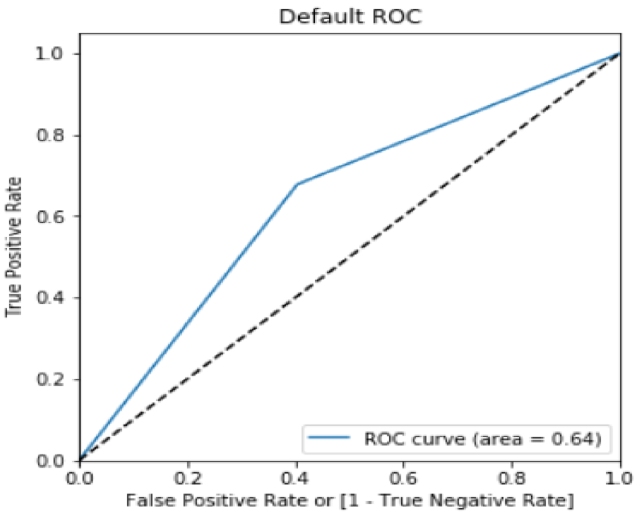
Demographic & Credit Data Model

UpGrad

The logistic model built on master data variables yields an accuracy of 64%.

Since the accuracy is not high we will try the Random forest or other classification algorithm to get good evaluation metrics

Evaluation Matrics	Value
Accuracy	64%
Precision	63%
Recall	68%
F1-Score	65%
AUC	64%

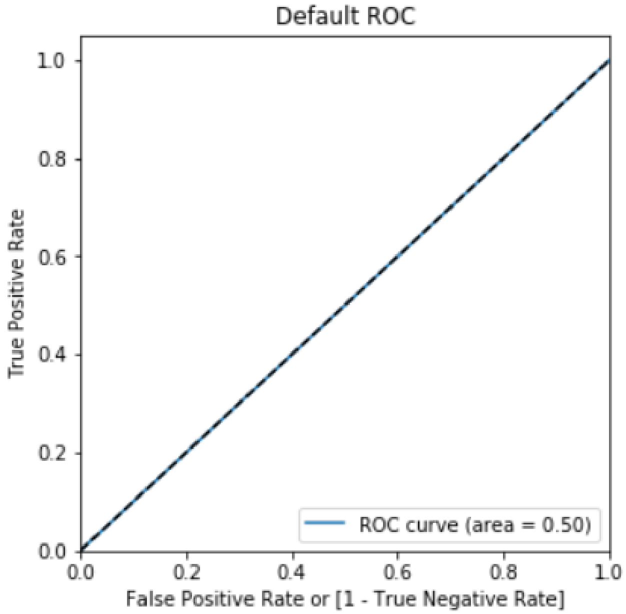


Model Building & Evaluations Metrics

Logistic Regression using Demographic data

- We have highly imbalanced data and We tried building the model using the actual data after EDA.
- Once we built the model, we evaluated the model using confusion metrics such as accuracy, sensitivity and specificity were captured for the test data.
- We got the accuracy of .96 for the model with very low sensitivity and specificity for predicting the default.
- We analyzed the below metrics and found that highly imbalanced data leads to an unstable model

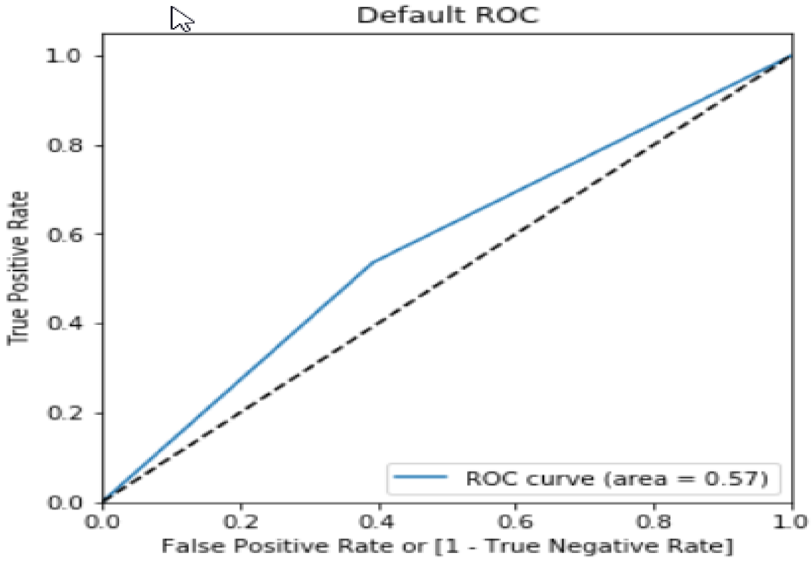
Evaluation Metrics	Value
Accuracy	95%
Precision	00%
Recall	00%
F1-Score	00%
AUC	50%



Logistic Regression using Oversample Demographic data

- We balanced the data using oversample method to make minority class equal to the majority class.
- Once we built the model, we evaluated the model using confusion metrics such as accuracy, sensitivity and specificity were captured for the test data.
- We got the accuracy of 57%
- We analyzed the below metrics and found that only demographic data is not enough to predict the defaults

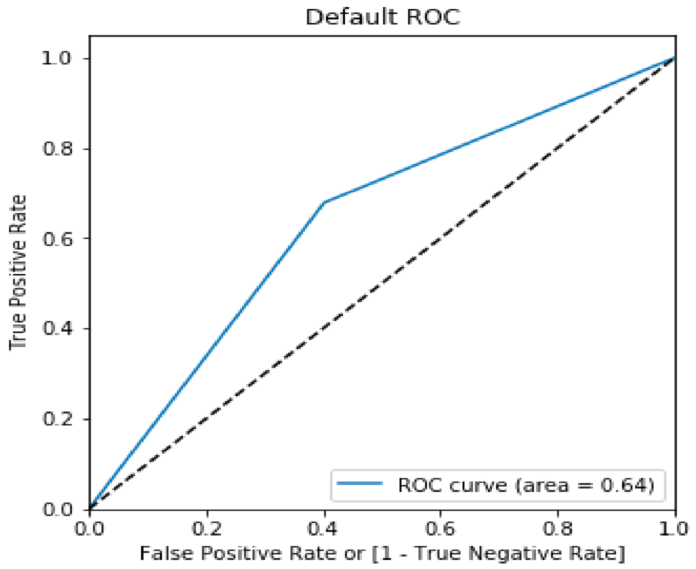
Evaluation Metrics	Value
Accuracy	57%
Precision	58%
Recall	54%
F1-Score	56%
AUC	.57



Logistic Regression using Merged data

- We merged the demographic data and credit data on application id for Model building. Selecting only the important variable using Information value(IV)
- Initially the model is being built on unbalanced data and later built using the oversample data
- The important variables are replaced with their WOE value and balanced using oversampling technique to make minority class equal to the majority class.
- Once we built the model, we evaluated the model using confusion metrics such as accuracy, sensitivity and specificity were captured for the test data.
- We have used RFE to select the 10 features and built the model with an accuracy of 63%
- We analyzed the below metrics and found that only demographic data is not enough to predict the defaults

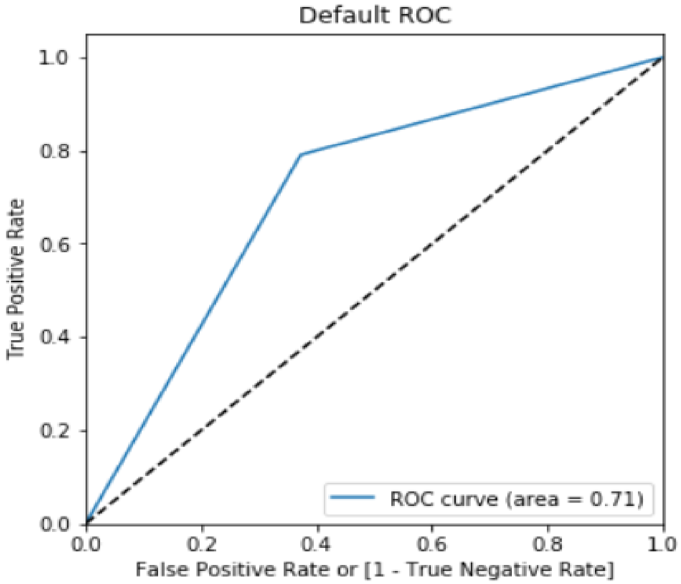
Evaluation Metrics	Value
Accuracy	63%
Precision	63%
Recall	67%
F1-Score	65%
AUC	.63



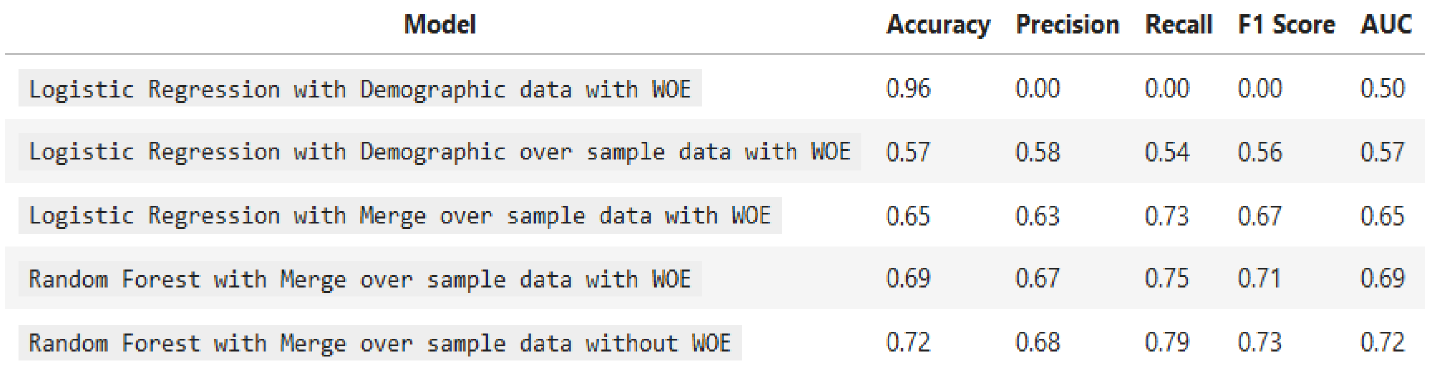
Random Forest using Oversample Merged data

- We have built two random forest models on the merged data.
- Random forest using WOE balanced data using identified variables by the information value
- Random forest using raw balanced data on complete data set
- Initially, the model was built using the unbalanced data set and later built using oversample data by making minority class equal to the majority class.
- Once we built the model, we evaluated the model using confusion metrics such as accuracy, sensitivity and specificity were captured for the test data.
- We found the model built on raw balanced data set performed better than other random forest models.

Evaluation Metrics	Value
Accuracy	72%
Precision	68%
Recall	79%
F1-Score	73%
AUC	.71



Model Selection



Model	Accuracy	Precision	Recall	F1 Score	AUC
Logistic Regression with Demographic data with WOE	0.96	0.00	0.00	0.00	0.50
Logistic Regression with Demographic over sample data with WOE	0.57	0.58	0.54	0.56	0.57
Logistic Regression with Merge over sample data with WOE	0.65	0.63	0.73	0.67	0.65
Random Forest with Merge over sample data with WOE	0.69	0.67	0.75	0.71	0.69
Random Forest with Merge over sample data without WOE	0.72	0.68	0.79	0.73	0.72

- `**Sensitivity**` in our context can be defined as a measure of total no of actual Default correctly predicted out of the total no of actual Default. It is also referred to as `**True Positive Rate**` or simple as `**Recall**`.

- The business goal here is to identify Default more accurately than to identify the non-Default. Hence the evaluation metric that satisfies the business goal is Sensitivity.

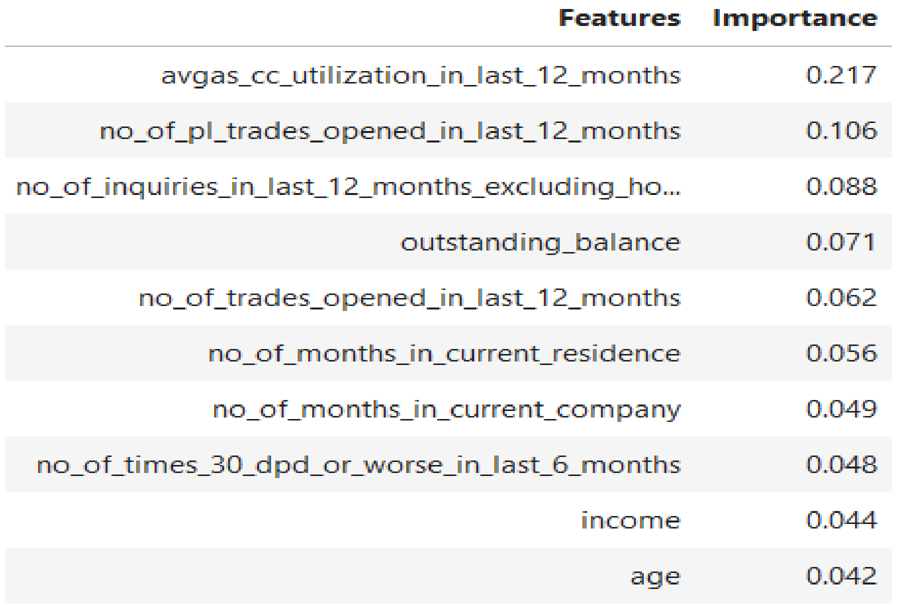
- From the above table, we see that the Random Forest Model has the best value for **Sensitivity(79%)**.ie. the model can **accurately predict 79%** of the Default customers.

- Model is able to **correctly predict 99.96% of the rejected population** as expected to default

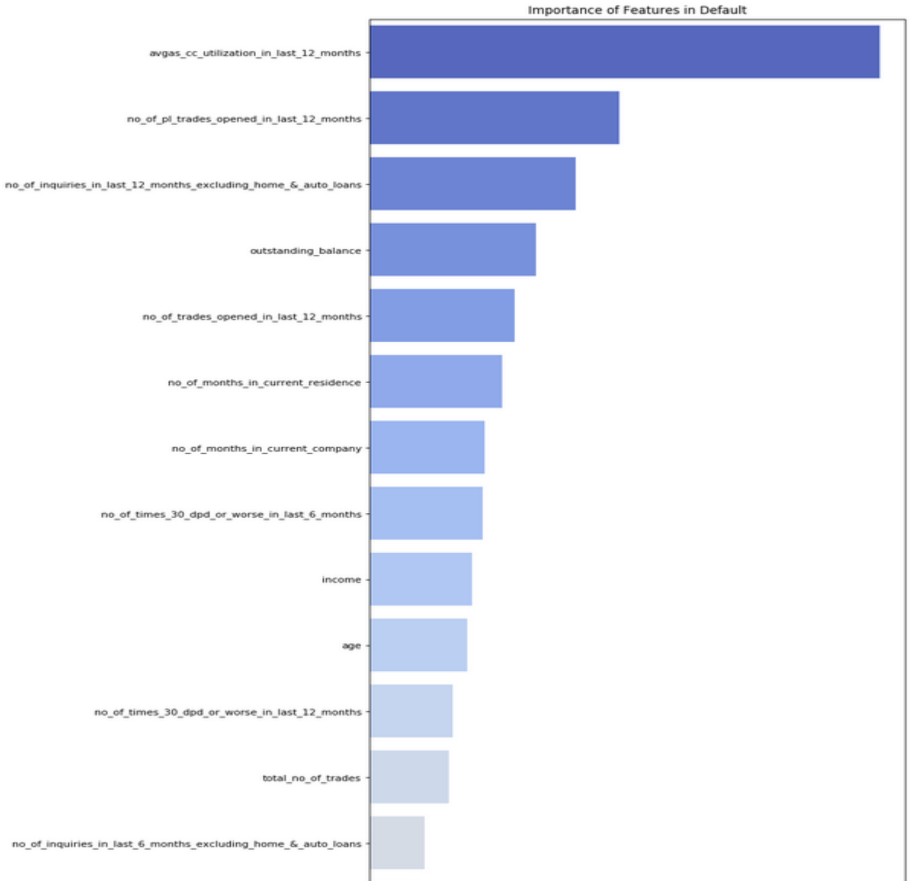
- Model is able to correctly predict 79% of the applicants who have defaulted.

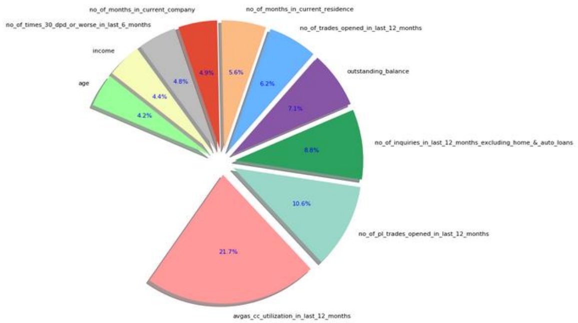
Important feature from the final Model

Important variable from the final model that we have built using the Random Forest on Raw balanced data set. Below are the few graphic visualizations to show the importance of a variable.



Features	Importance
avgas_cc_utilization_in_last_12_months	0.217
no_of_pl_trades_opened_in_last_12_months	0.106
no_of_inquiries_in_last_12_months_excluding_ho...	0.088
outstanding_balance	0.071
no_of_trades_opened_in_last_12_months	0.062
no_of_months_in_current_residence	0.056
no_of_months_in_current_company	0.049
no_of_times_30_dpd_or_worse_in_last_6_months	0.048
income	0.044
age	0.042



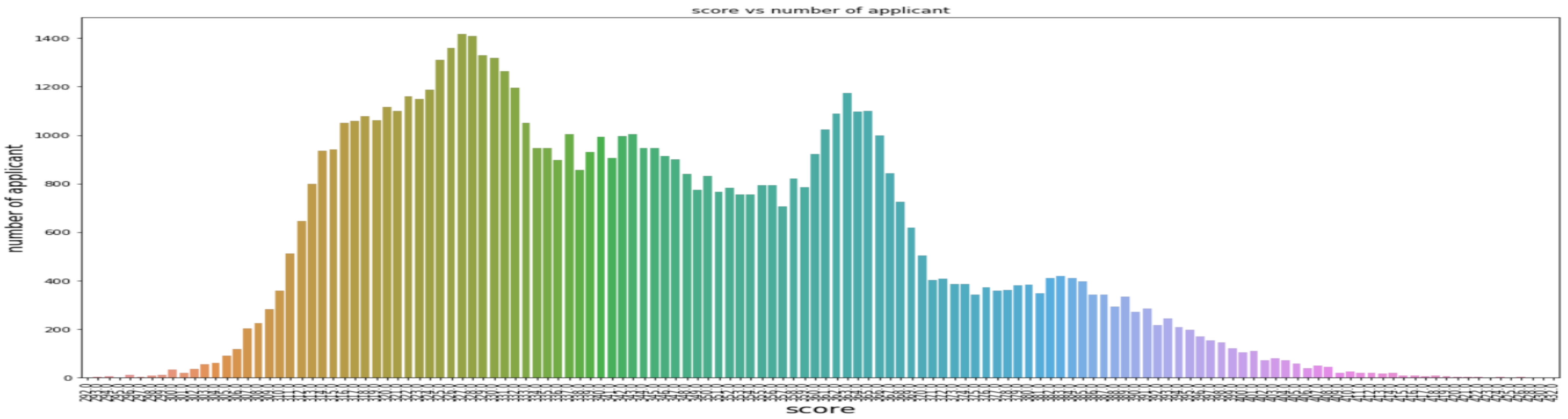


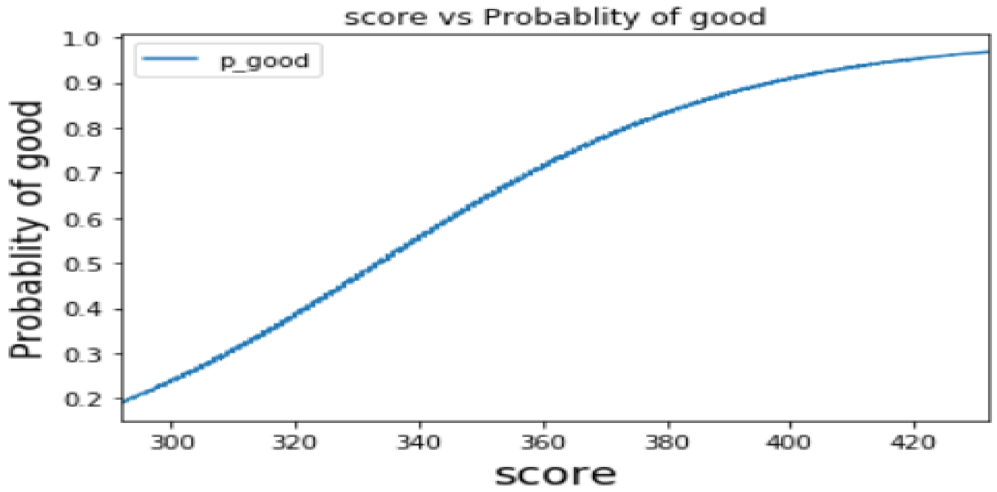
Application Score-Card

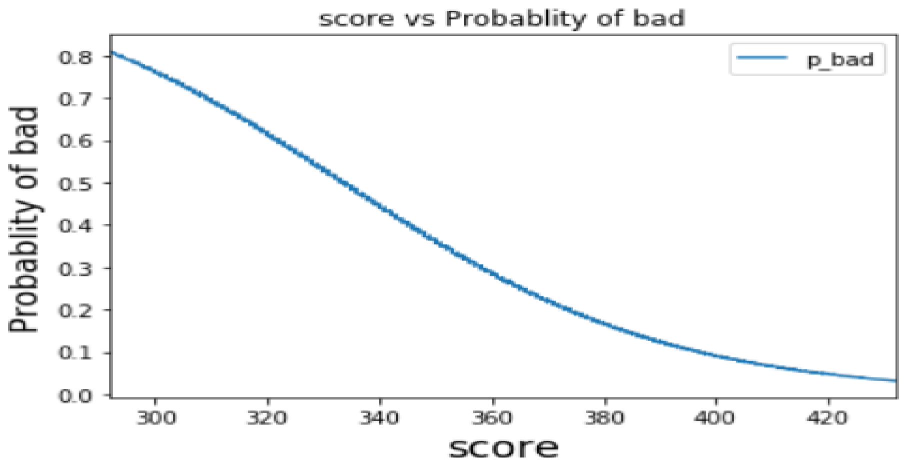
- The Score card was built based on the prediction of final selected model and using the data where rejected customer was removed.
- Point to Double, Base Score & good to bad odds was taken as 20, 400 & 10 respectively. After which the offset and scores were calculated for each applicant.
- We have identified the cut off score using the accuracy, sensitivity and specificity of the final model
- The cut-off score identified is **327.** The applicants with score greater then 327 are expected to not default and any applicant less than the cut-off score are likely to default. So we will not grant credit cards to applicants with score less than 327 since they are the risky customers.
- Score card was built on the initial rejected customer base and applicants securing score greater than 327 can be actually be granted a credit card by CredEx since they are very less likely to default
- **Model predicted 26.0% applicant as defaulter** than the **data set 4.22 % customer defaulted.** that is quite good number that CredEx can save from credit loss
- Scores built in the application scorecard range from 292 to 432 and the median score is 344.
- According to the model **51568** applicants has score greater the cutoff score likely to not default and **17995** applicants likely to default

Score-Card

Below visualization shows score distribution among the number of applicant, probability of good and bad applicants.







Financial Benefit

Model performance for financial benefits

The model was evaluated on the complete data set and the results are below:

- Total applicants who have defaulted and model has predicted as likely to default is 1337
- Total applicants who have defaulted and model has predicted as not likely to default is 1602
- Total applicants who have not defaulted and model has predicted as not likely to default is 49966
- Total applicants who have not defaulted and model has predicted as likely to default is 16658

The above data shows the model performance for the CredEx data set, the next slide explains the financial benefits obtained from the model Vs without model

Assessing Financial Benefits for CredEx

Credit Loss Avoided:

Current Percentage of defaulters according to the given data is 4.22 % and by using the model, default rate based on model prediction wherein model rejected the bad customers is 1.94%

Total Credit loss saved using Model = 4.22 - 1.94 = 2.28%

Potential Revenue Loss due to rejection of good customers:

The bad customer likely to default according to model out of good customer base from the data set

Total revenue loss without using Model : 25.33 %

Conclusion : We can see above the power of predictive modeling for the CredEx compared to the conventional method to identify the applicant for default. Scores help the financial institution to clearly understand the applicant and accelerate the process of granting credit card. The model can also help the institution to identify the right applicant and helps in deciding whom to accept/reject.

Thank You