# CAPSTONE BFSI


Group Name:

1. Abinash Panda
2. Lipsa Satapathy
3. Prabhudatta Praharaj
4. Jai Shankar Bhagat

# Business Objective

## Project Background

- CredX is a leading credit card provider that gets thousands of credit card applicants every year.
- Develop a model to mitigate credit risk is to 'acquire the right customers'.

## Problem Statement

- The company receives 1000s of credit cards application every year from different demography and customer types.
- in the past few years, it has experienced an increase in credit loss.
- Using past data of the bank's applicants, you need to determine the factors affecting credit risk, create strategies to mitigate the acquisition risk and assess the financial benefit of your project

## Business Objective

- The objective is to develop a robust credit risk predictive model by analyzing credit bureau and demographic data.

# Model Building Methodology

- Business Understanding and Data Understanding
- Data Preparation and Data Cleaning
- Exploratory Data Analysis ( Graphs & plots )
- Data Transformation ( Outliers treatment , missing value treatment , Numeric conversation , Data Removal )
- Model Development ( Logistic Regression, Random Forest)
- Model Evaluation and Testing
- Model Acceptance or Rejection
- Application Score card based on the best model

# Data Preparation & Cleaning

We are provided with 2 datasets : 1> Credit Bureau Data 2> Demographic data

Post study of the 2 given datasets we have done the following Validation before EDA

## DATA Validation

- We have verified common column is Application ID in all the data set.
- There are 3 duplicate application IDs, which are removed.

## DATA Preparation

- For EDA and model building activites we have merged all the relevant data into a single Dataframe.
- We have created few calculated columns like No of overtime instances for each employee.
- We have prepared box plot for Outliers treatment , however did not find any significant ones which can be excluded.

# Data Manipulation & Transformation

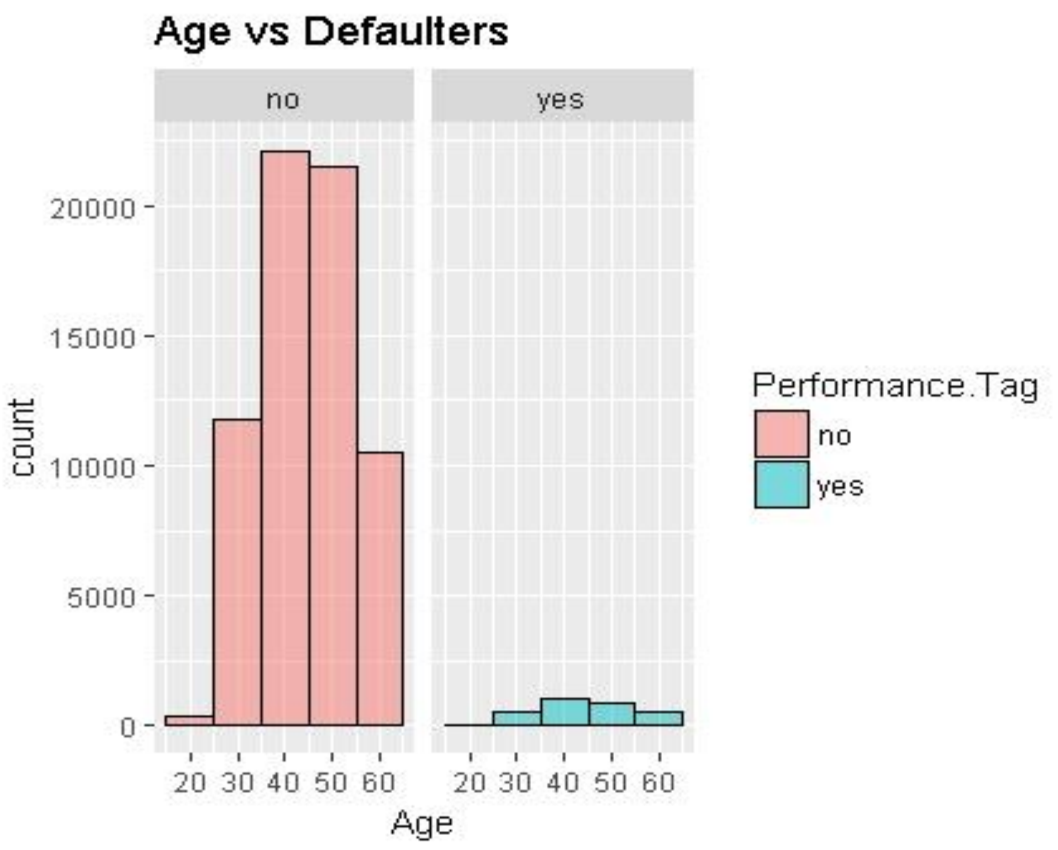
Post Data cleaning we have done the following data transformation

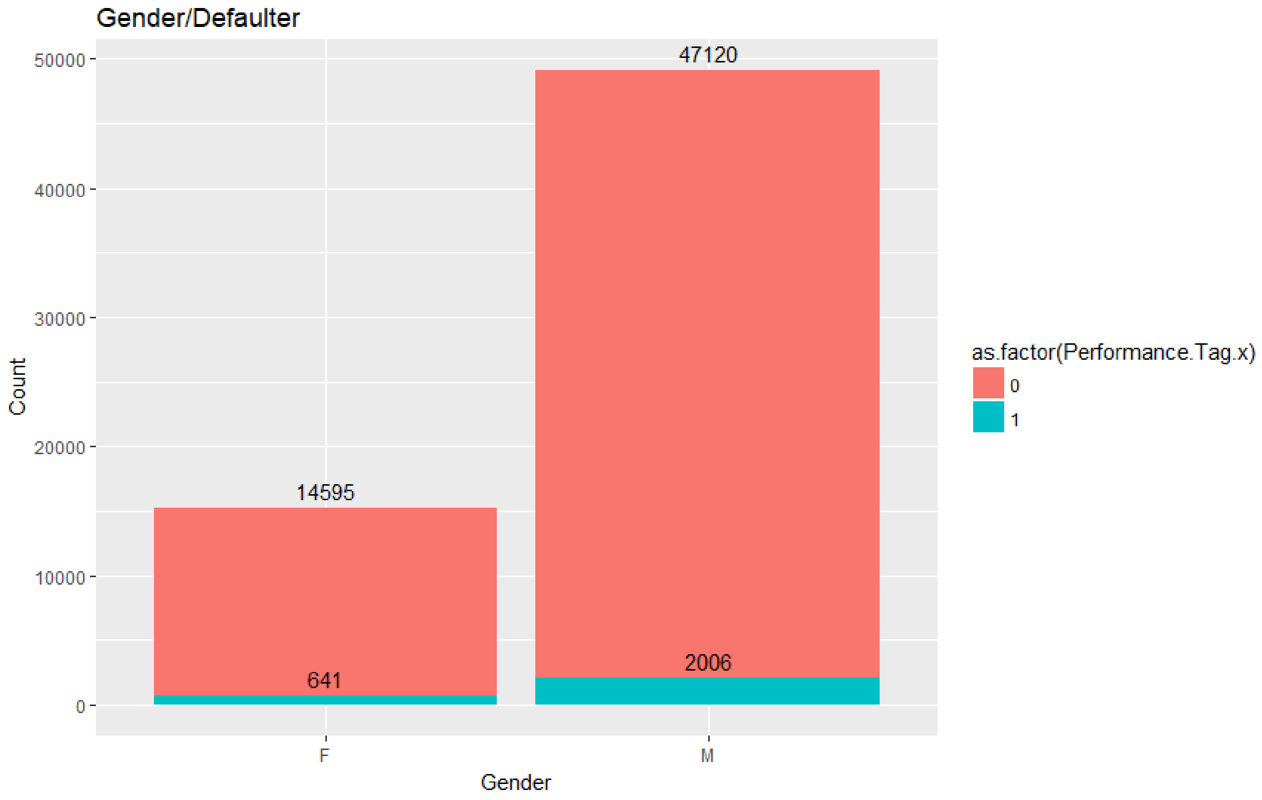
## Missing Value Treatment

- We found NA/missing values in various columns demographic. Since the count is very less compared we are removing those rows.
- There are NA values in credit bureau as well. One of the variable "Avgas.CC.Utilisation" has high no NA values. Since this variable
- has high IV value we cannot remove the NAs. Hence we will be using WOE binning to change this variable to categorical.

# Exploratory Data Analysis

We have plotted the following graphs to identify possible significant variables through EDA



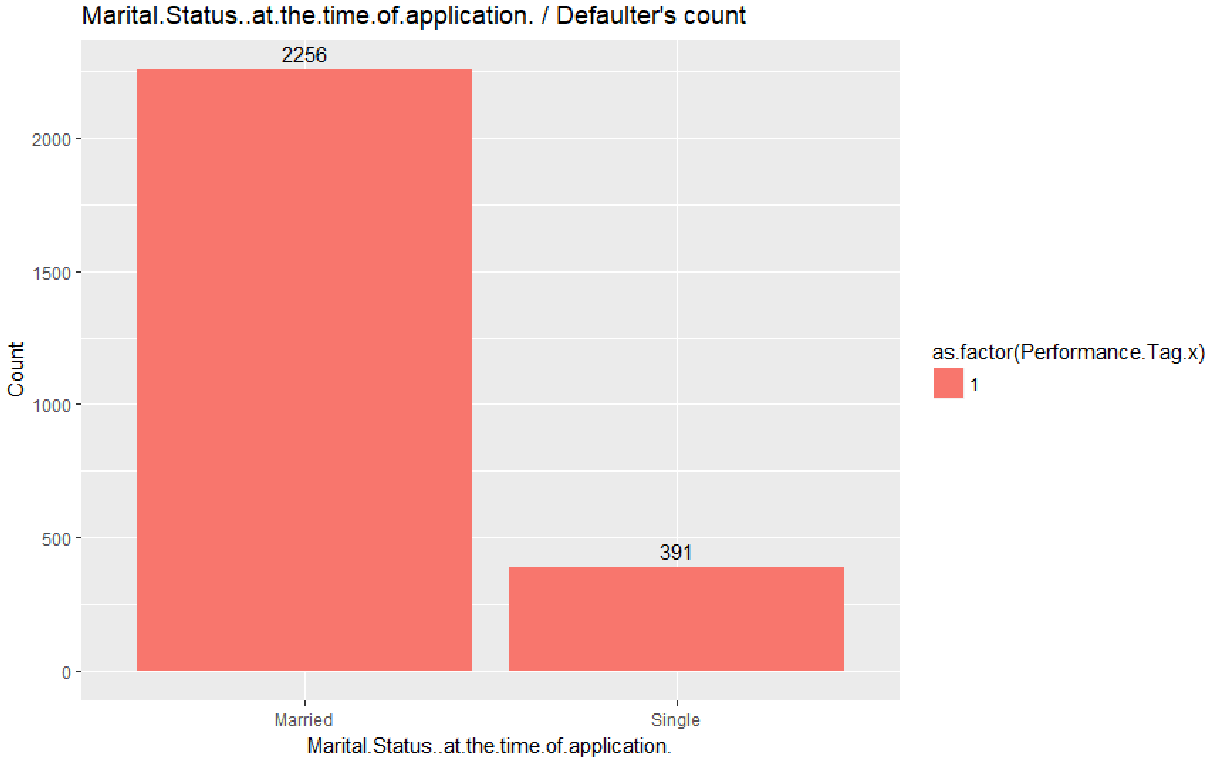

**Observation :** *Age group between 40-55 tend to default the most*

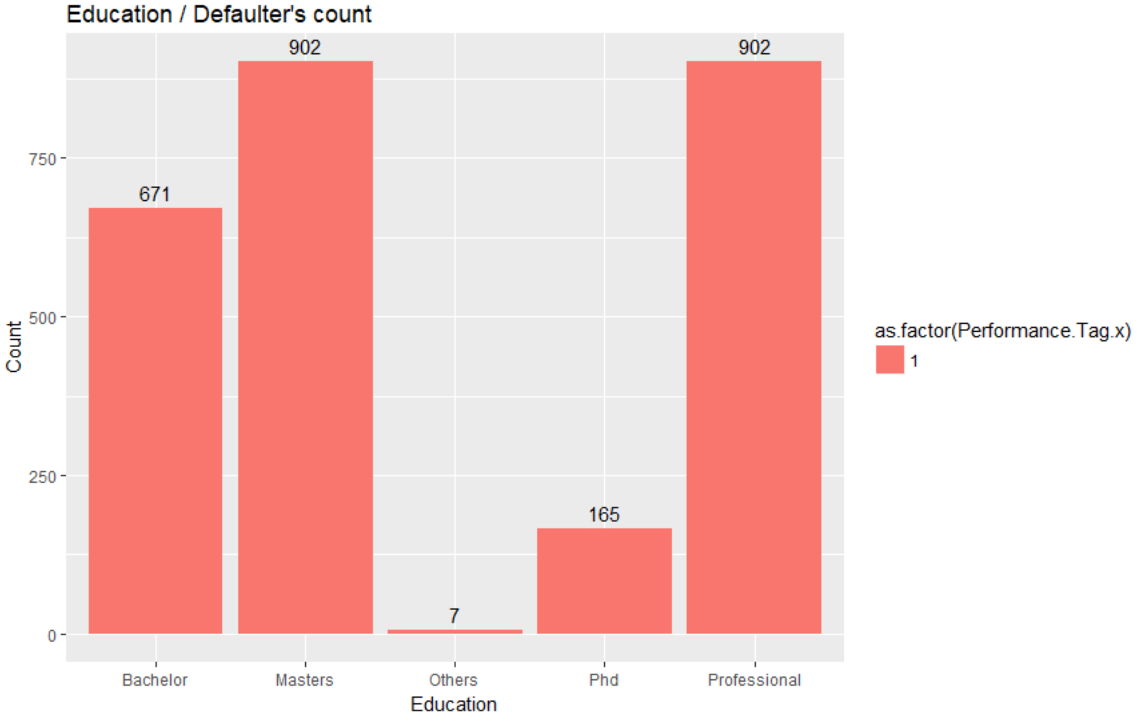
**Observation :** *Males seems to default more than females.*

EDA Continue .....

# Exploratory Data Analysis ….

We have plotted the following graphs to identify possible significant variables through EDA



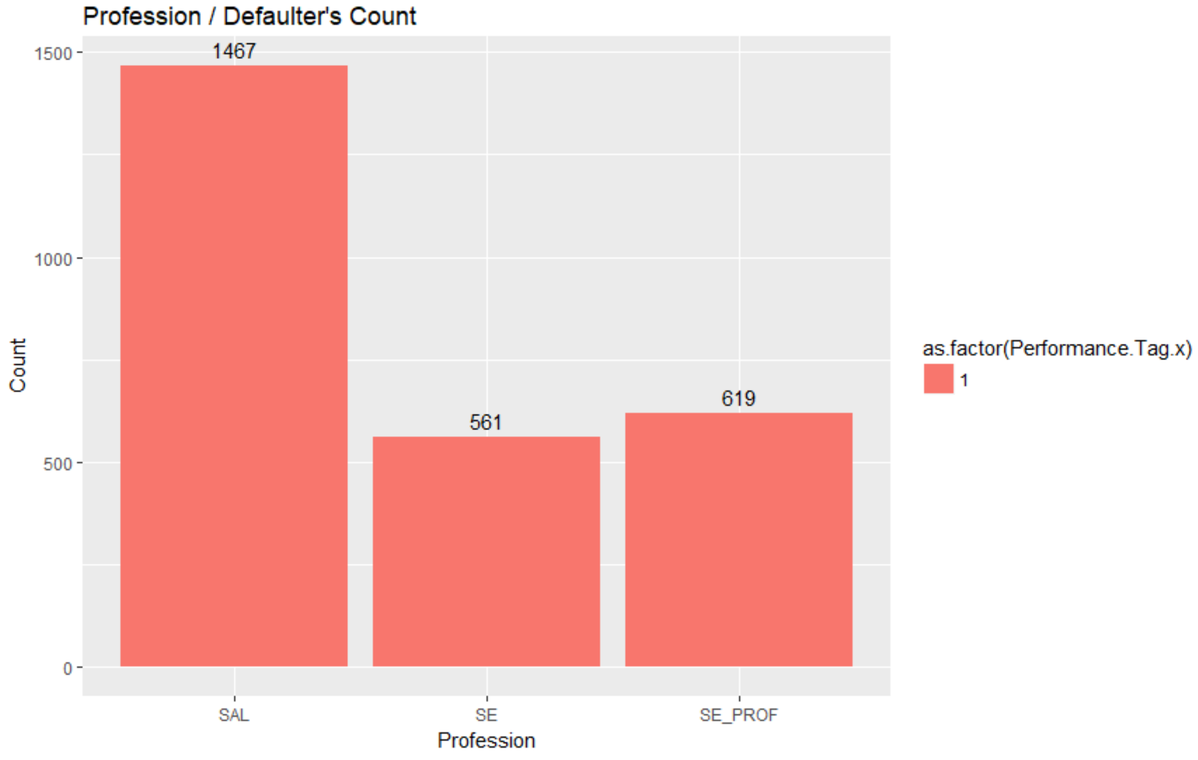



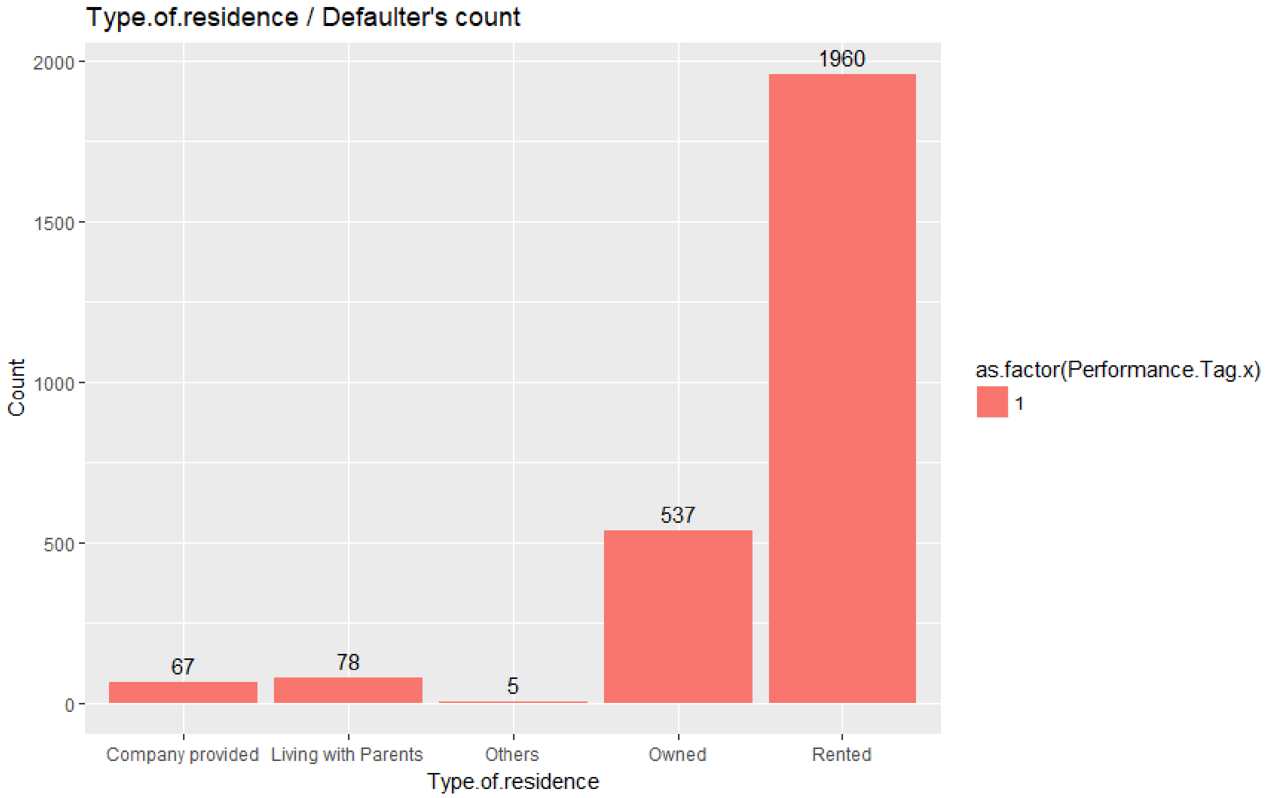

**Observation :** *Applicants having marital status married has high risk of defaulting*

**Observation :** *Applicants with Masters or Professional Education Qualification has higher risk of defaulting.*

EDA Continue …..

# Exploratory Data Analysis ....



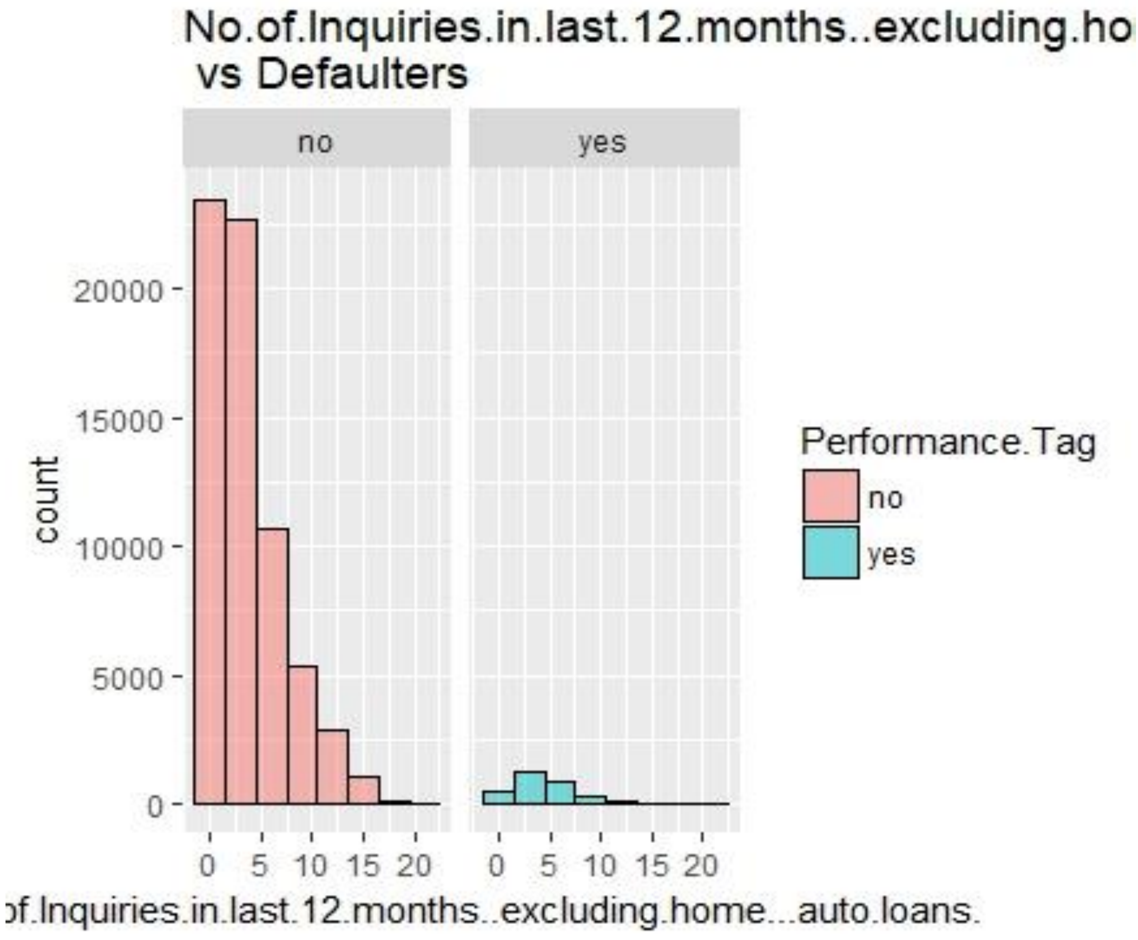

**Observation :**

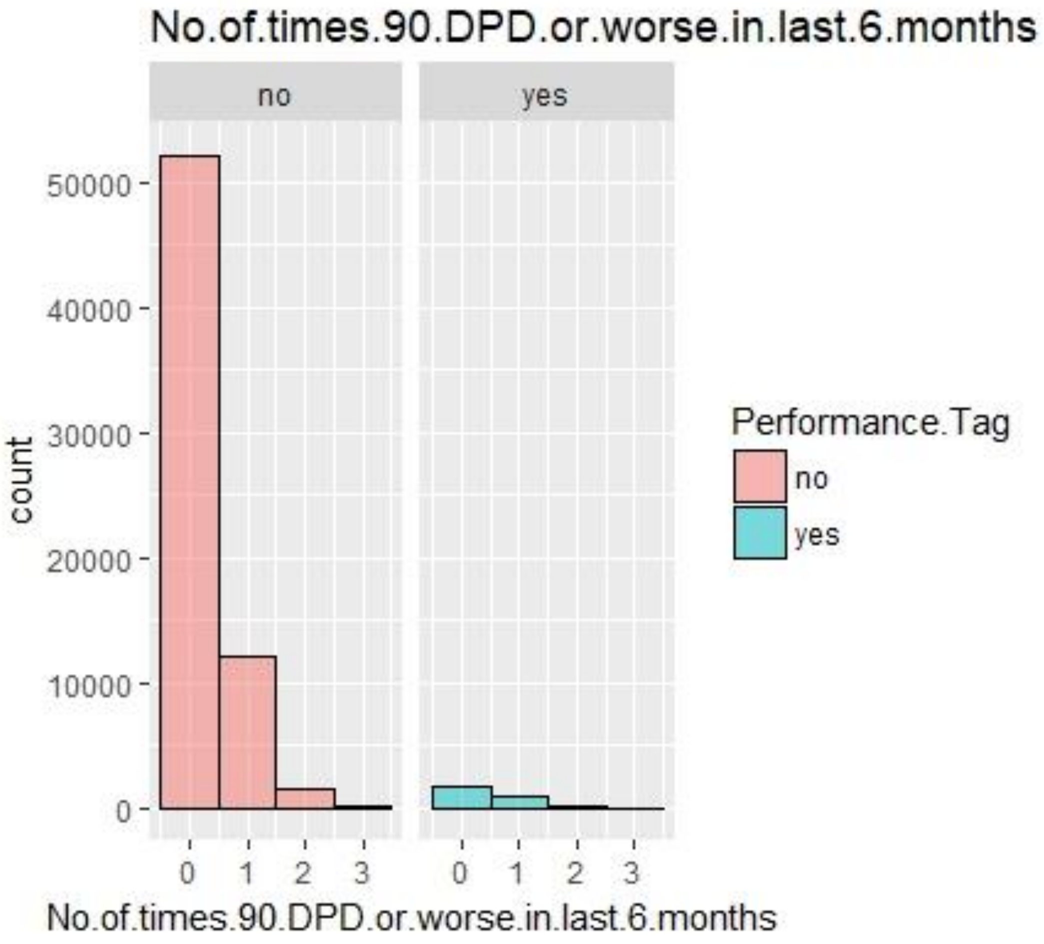
*Salaried Applicants are the ones who default the most.*

**Observation :**

*Rented ones are having high default chances .*

# Exploratory Data Analysis ....

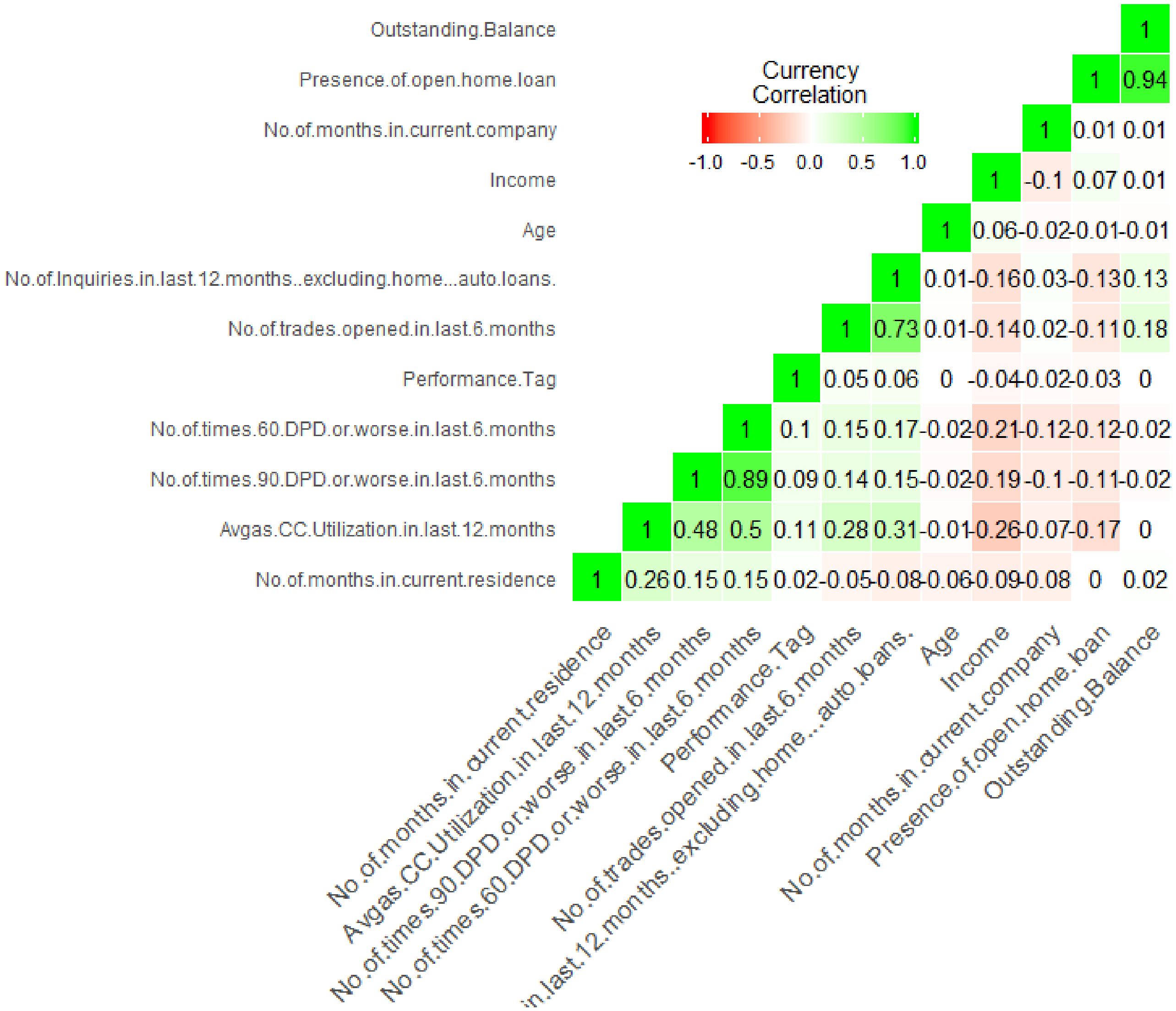



**Observation :** *There is a high defaulter rate between 3 to 7 inquiries.*

**Observation :** *No of delinquency 1 and 2 are more likely to default..*

# Exploratory Data Analysis ….




**Observation :** *Through the heat-map of the correlation matrix ,we found the Variables which are effecting the dependent variable "**Performance. Tag**".*

# Observations from EDA

**Attributes with missing values**

- No of dependents=3 na's
- Performance Tag=1425
- Average Credit Card Utilization=1058
- No of trades opened in last six months=1
- Presence of open Home Loan=272
- Outstanding Balance=272

**Attributes with outliers**

- Age has negative values as well as less than 18 which is not permissible for application.
- Income has negative values.

**Attributes with Blanks**

- Gender
- Marital Status
- Education
- Profession
- Residence

# Observations from EDA

**Steps taken to handle missing values and outliers**

- The Na in no of dependents were eliminated since they were very few and would not have a significant effect.
- The **Na in performance Tag corresponds to the applicants whose application was rejected**. Hence it will be handled later.
- The no of Na in Average Credit Card Utilization are 1058. Hence **eliminating them** can **introduce bias**. Hence we will **replace them by the Weight of Evidence Values**.
- The Na in no of trades opened in last six months is eliminated since there is only 1 na.
- The Na in **presence of open Home Loan** will be replaced by **WOE values.**
- The Na in **outstanding balance** will be replaced by **WOE values.**
- The **values less than 18 in age is replaced by Na** since **no applicant below the age of 18** can **apply** for the credit card application. These Na values are then **replaced by WOE values**.
- Similarly we will do the **same for income** which also has some negative values.
- The row containing blanks in Gender are eliminated because there is only 1 row and hence it has very little significance.
- The rows containing **Marital status are eliminated since there are only 5 rows** and will have little significance.
- The **blanks in Education are replaced by Missing which will be replaced by WOE later.** We converted the Blank values to missing in Education since there was a lot of rows and replacing them by mode could introduce bias.
- The **blanks in Profession and residence are replaced by their respective modes.** This is done because the number of rows in Profession and residence are not very less and are not very high thus if we replace them by mode their will be very less bias that will be included.

# WEIGHT OF EVIDENCE AND INFROMATION VALUE

## 1. Information Value

- Information values are used to find out the predictive power of different attributes on the target attribute.

## 2 . Steps taken to find IV and Woe

- First **we will split data between rejected and selected candidates**. To do this we will remove all the rows which had Na's in Performance Tag because we only need selected candidates to assess which attributes are necessary to predict the Performance Tag of the selected candidate.
- **IV <- create_infotables(data=woe_data, y="Performance.Tag", bins=10, parallel=FALSE)** will create the information Table and tell Woe values curresping to each bins.

# WEIGHT OF EVIDENCE AND INFROMATION VALUE

**Following output is the details related to IV for few of the independent variables :**

```
$`Tables`$Age
        Age     N      Percent          WOE           IV
1        NA    62 0.0008875528 -0.988343026 0.0005656994
2   [18,30]  5883 0.0842173073 -0.035001357 0.0006672362
3   [31,35]  6926 0.0991482356  0.034503159 0.0007871513
4   [36,38]  6923 0.0991052895  0.069043818 0.0012748119
5   [39,41]  7128 0.1020399399  0.068265187 0.0017654759
6   [42,44]  7004 0.1002648343 -0.037674595 0.0019053600
7   [45,47]  6828 0.0977453296 -0.003832285 0.0019067930
8   [48,50]  6742 0.0965142080 -0.012653877 0.0019221577
9   [51,53]  6841 0.0979314294 -0.137084765 0.0036512826
10  [54,57]  7619 0.1090687853  0.043225929 0.0038591568
11  [58,65]  7899 0.1130770883 -0.010192745 0.0038708500

$`Tables`$Gender
  Gender     N   Percent         WOE           IV
1      F 16502 0.2362322  0.03224836 0.0002493307
2      M 53353 0.7637678 -0.01017016 0.0003279621

$`Tables`$Marital.Status..at.the.time.of.application.
  Marital.Status..at.the.time.of.application.     N   Percent         WOE           IV
1                                     Married 59539 0.8523227 -0.00407893 1.415421e-05
2                                      Single 10316 0.1476773  0.02324868 9.482898e-05

$`Tables`$No.of.dependents
  No.of.dependents     N   Percent          WOE          IV
1                1 15216 0.2178226  0.040043342 0.000355746
2                2 15126 0.2165342 -0.085263404 0.001869896
3                3 15643 0.2239353  0.054109893 0.002542038
4                4 11997 0.1717415 -0.025296795 0.002650676
5                5 11873 0.1699664  0.004387421 0.002653954

$`Tables`$Income
     Income     N     Percent         WOE          IV
1        NA    81 0.001159545 -0.55376983 0.000278031
2     [0,5]  6245 0.089399470  0.31064671 0.010242299
3    [6,10]  6509 0.093178727  0.27573411 0.018291803
4   [11,16]  7922 0.113406342  0.06604176 0.018801656
5   [17,21]  6802 0.097373130  0.08077721 0.019461038
6   [22,26]  6827 0.097731014  0.02503774 0.019523011
7   [27,31]  6817 0.097587861  0.07846934 0.020145962
8   [32,36]  6830 0.097773960 -0.15613435 0.022366272
9   [37,41]  6722 0.096227901 -0.26370672 0.028306536
10  [42,48]  7784 0.111430821 -0.17704285 0.031529592
11  [49,60]  7316 0.104731229 -0.36054243 0.043107488
```

# WEIGHT OF EVIDENCE AND INFROMATION VALUE

**Following output is the details related to IV for few of the independent variables :**

```
$`Tables`$No.of.times.60.DPD.or.worse.in.last.6.months
  No.of.times.60.DPD.or.worse.in.last.6.months     N   Percent        WOE         IV
1                                        [0,0] 51863 0.7424379 -0.3364066 0.07221892
2                                        [1,5] 17992 0.2575621  0.6226574 0.20588944

$`Tables`$No.of.times.30.DPD.or.worse.in.last.6.months
  No.of.times.30.DPD.or.worse.in.last.6.months     N   Percent        WOE         IV
1                                        [0,0] 50091 0.7170711 -0.3868273 0.09020276
2                                        [1,1]  9499 0.1359817  0.4642520 0.12659426
3                                        [2,7] 10265 0.1469472  0.7430911 0.24162386

$`Tables`$No.of.times.90.DPD.or.worse.in.last.12.months
  No.of.times.90.DPD.or.worse.in.last.12.months     N   Percent        WOE         IV
1                                         [0,0] 50485 0.7227113 -0.3566695 0.07832097
2                                         [1,1] 11661 0.1669315  0.5088281 0.13313093
3                                         [2,5]  7709 0.1103572  0.7222275 0.21392778

$`Tables`$No.of.times.60.DPD.or.worse.in.last.12.months
  No.of.times.60.DPD.or.worse.in.last.12.months     N   Percent        WOE         IV
1                                         [0,0] 45862 0.6565314 -0.3519656 0.06942840
2                                         [1,1] 12814 0.1834371  0.2141391 0.07871496
3                                         [2,7] 11179 0.1600315  0.6942958 0.18556159

$`Tables`$No.of.times.30.DPD.or.worse.in.last.12.months
  No.of.times.30.DPD.or.worse.in.last.12.months     N   Percent        WOE         IV
1                                         [0,0] 44851 0.6420585 -0.3764375 0.07683608
2                                         [1,2] 17588 0.2517787  0.2804936 0.09939425
3                                         [3,9]  7416 0.1061628  0.7999062 0.19833516

$`Tables`$Avgas.CC.Utilization.in.last.12.months
   Avgas.CC.Utilization.in.last.12.months    N    Percent        WOE           IV
1                                      NA 1022 0.01463031  0.1123205 0.0001943648
2                                   [0,4] 5523 0.07906377 -0.8017531 0.0359837127
3                                   [5,6] 5471 0.07831938 -0.8016817 0.0714308313
4                                   [7,8] 6869 0.09833226 -0.7947025 0.1152909675
5                                  [9,11] 9596 0.13737027 -0.6724790 0.1614694245
6                                 [12,14] 6593 0.09438122 -0.4678696 0.1782340284
7                                 [15,21] 6853 0.09810321 -0.0790318 0.1788250902
8                                 [22,37] 7118 0.10189679  0.4754680 0.2075805310
9                                 [38,51] 6746 0.09657147  0.5844097 0.2509357541
10                                [52,71] 7017 0.10045093  0.5635516 0.2924575395
11                               [72,113] 7047 0.10088040  0.3814642 0.3099863082

$`Tables`$No.of.trades.opened.in.last.6.months
  No.of.trades.opened.in.last.6.months     N    Percent        WOE         IV
1                                [0,0] 12193 0.17454728 -0.6577239 0.05648058
2                                [1,1] 20120 0.28802520 -0.4796436 0.10997480
3                                [2,2] 12112 0.17338773  0.2330302 0.12046166
4                                [3,3]  9402 0.13459309  0.4349446 0.15164232
5                                [4,4]  6293 0.09008661  0.5247803 0.18334375
6                               [5,12]  9735 0.13936010  0.1367842 0.18612074
```

```
$`Tables`$No.of.PL.trades.opened.in.last.12.months
  No.of.PL.trades.opened.in.last.12.months     N    Percent        WOE        IV
1                                    [0,0] 25821 0.36963711 -0.8938719 0.2002405
2                                    [1,1]  6641 0.09506836 -0.1311962 0.2017820
3                                    [2,2]  6827 0.09773101  0.2516248 0.2087341
4                                    [3,3]  8129 0.11636962  0.4122478 0.2326947
5                                    [4,4]  7899 0.11307709  0.5004390 0.2684655
6                                    [5,5]  6188 0.08858349  0.4261426 0.2880834
7                                   [6,12]  8350 0.11953332  0.2429782 0.2959801

$`Tables`$No.of.Inquiries.in.last.6.months..excluding.home...auto.loans.
  No.of.Inquiries.in.last.6.months..excluding.home...auto.loans.     N   Percent         WOE        IV
1                                                          [0,0] 25066 0.3588290 -0.71828758 0.1349888
2                                                          [1,1] 13174 0.1885907  0.17697266 0.1413979
3                                                          [2,2] 12829 0.1836519  0.21608183 0.1508734
4                                                          [3,4] 11502 0.1646554  0.50999440 0.2052141
5                                                         [5,10]  7284 0.1042731  0.01252292 0.2052306

$`Tables`$No.of.Inquiries.in.last.12.months..excluding.home...auto.loans.
  No.of.Inquiries.in.last.12.months..excluding.home...auto.loans.     N    Percent         WOE        IV
1                                                           [0,0] 20578 0.29458163 -1.06756799 0.2122267
2                                                           [1,1]  3899 0.05581562 -0.06195388 0.2124349
3                                                           [2,2]  7907 0.11319161  0.14196536 0.2148704
4                                                           [3,3]  8975 0.12848042  0.16452153 0.2186222
5                                                           [4,4]  7111 0.10179658  0.24822309 0.2256578
6                                                           [5,5]  4927 0.07053182  0.58800126 0.2577678
7                                                           [6,8]  8948 0.12809391  0.48431138 0.2954306
8                                                          [9,20]  7510 0.10750841  0.01366460 0.2954508

$`Tables`$Presence.of.open.home.loan
  Presence.of.open.home.loan     N    Percent         WOE           IV
1                       <NA>   272 0.00389378 -0.37397672 0.0004603996
2                          0 51516 0.73747047  0.07368874 0.0046027573
3                          1 18067 0.25863575 -0.23660820 0.0176122254

$`Tables`$Outstanding.Balance
   Outstanding.Balance    N    Percent        WOE           IV
1                   NA  272 0.00389378 -0.3739767 0.0004603996
2             [0,6843] 6957 0.09959201 -0.7703167 0.0426233040
3         [6847,25509] 6959 0.09962064 -0.9203742 0.0992222746
4       [25522,386809] 6958 0.09960633 -0.1343724 0.1009141364
5      [386813,585402] 6958 0.09960633  0.2543889 0.1081654771
6      [585423,774228] 6959 0.09962064  0.4532107 0.1334406289
7      [774241,972455] 6958 0.09960633  0.4342381 0.1564334722
8     [972456,1357300] 6958 0.09960633  0.4049362 0.1761532842
9    [1357399,2960987] 6958 0.09960633 -0.3873877 0.1887163424
10   [2960994,3282013] 6959 0.09962064 -0.8233079 0.2358460390
11   [3282027,5218801] 6959 0.09962064  0.2958411 0.2458466821
```

# WEIGHT OF EVIDENCE AND INFROMATION VALUE

**3. After we find the WOE values we can map all our attribute values with corresponding WOE values.**

**4. Next using IV we find the predictive power of independent attributes:**

```
> IV$Summary
                                                    Variable           IV
18           Avgas.CC.Utilization.in.last.12.months 3.099863e-01
20            No.of.trades.opened.in.last.12.months 2.980502e-01
22         No.of.PL.trades.opened.in.last.12.months 2.959801e-01
24 No.of.Inquiries.in.last.12.months..excluding.home...auto.loans. 2.954508e-01
26                              Outstanding.Balance 2.458467e-01
14         No.of.times.30.DPD.or.worse.in.last.6.months 2.416239e-01
27                             Total.No.of.Trades 2.366873e-01
21          No.of.PL.trades.opened.in.last.6.months 2.197590e-01
15        No.of.times.90.DPD.or.worse.in.last.12.months 2.139278e-01
13         No.of.times.60.DPD.or.worse.in.last.6.months 2.058894e-01
23  No.of.Inquiries.in.last.6.months..excluding.home...auto.loans. 2.052306e-01
17        No.of.times.30.DPD.or.worse.in.last.12.months 1.983352e-01
19             No.of.trades.opened.in.last.6.months 1.861207e-01
16        No.of.times.60.DPD.or.worse.in.last.12.months 1.855616e-01
12         No.of.times.90.DPD.or.worse.in.last.6.months 1.601793e-01
10                No.of.months.in.current.residence 7.904415e-02
6                                            Income 4.310749e-02
11                  No.of.months.in.current.company 2.176757e-02
25                        Presence.of.open.home.loan 1.761223e-02
2                                               Age 3.870850e-03
5                                  No.of.dependents 2.653954e-03
8                                        Profession 2.226299e-03
28                        Presence.of.open.auto.loan 1.658166e-03
1                                    Application.ID 1.499617e-03
9                                 Type.of.residence 9.229593e-04
7                                         Education 7.839395e-04
3                                            Gender 3.279621e-04
4             Marital.Status..at.the.time.of.application. 9.482898e-05
>
```

| Information Value | Variable Predictiveness |
|---|---|
| Less than 0.02 | Not useful for prediction |
| 0.02 to 0.1 | Weak predictive Power |
| 0.1 to 0.3 | Medium predictive Power |
| 0.3 to 0.5 | Strong predictive Power |
| >0.5 | Suspicious Predictive Power |

# Model Building/Selection

## LOGISTIC REGRESSION ON MASTER FILE:

- Accuracy obtained from the model is **95.82%**.
- However when observed closely we see that **postive prediction** has **95.82%** accuracy whereas **negative prediction accuracy is 0%**. This is because of the **imbalance in the dataset.**
- Hence we **first balance the data by using Synthetic Minority Over Sampling Technique.**
- Now we have a balanced master file which **has almost equal number of 0's and 1's in performance Tag.**

## LOGISTIC REGRESSION ON BALANCED MASTER FILE:

- **Accuracy=70%.** The model does not fit well on the test data as it has very less accuracy hence we will try some other supervised learning algorithm**.**
- Overall statistics of the model:
  - Accuracy : 0.7098
  - Sensitivity : 0.51083
  - Specificity : 0.71851

| Prediction | No | Yes |
|---|---|---|
| No | 14427 | 429 |
| Yes | 5652 | 448 |

# Model Building/Selection

## RANDOM FOREST ON BALANCED DATA :

- We tried to tune the Random forest hyperparameters to achieve the best results. The hyperparameters we got after tuning the RF model is:
  - ntree=483
  - mtry=8
  - nodesize=20
- We trained the model with above hyperparameters on the balanced data and got significant improvements in the results to that of LR model.
- Overall statistics of RF model:
  - Accuracy : 0.7377
  - Sensitivity : 0.74720
  - Specificity : 0.73728

| Predicted | No | Yes |
|---|---|---|
| No | 49330 | 745 |
| Yes | 17578 | 2202 |

- Hence from this algorithm we have received the highest accuracy than other models. Therefore we will use this model for generating application scorecards and predicting defaulters.

# Model Evaluation

## EVALUATING THE MODEL ON REJECTED CANDIDATES:

- We **assume that since rejected candidates were rejected by the bank because of their high likelihood of doing default. Hence we can assume that Performance Tag will be 1 for all the rejected applicants.**
- We predict the **random forest model on rejected candidates** and compare how many of them are predicted correctly.
- **86% of the rejected candidates are classified as defaulters.**
- RESULT: SINCE THE REJECTED CANDIDATES WERE REJECTED BECAUSE OF THE RISK TO DEFAULT IN FUTURE HENCE **OUR MODEL PREDICTS THAT MORE THAN 86% OF CANDIDATES WILL BE DEFAULTERS IF THEY WOULD HAVE BEEN ACCEPTED.**
- The model **predicts that 14% of rejected candidates may not have defaulted** if they would have been accepted.

# Application Score Card

## BUILDING APPLICATION SCORECARD:-

We have to build application scorecard with the good to bad odds of 10 to 1 at a score of 400 doubling every 20 points.

**> Hence, factor= 20/ln(2)**

**> Hence, score=400+factor*log(probability of not doing default/probability of doing default)**

**> Following formulae were used in the codes to derive the scorecards:**

- **Factor=20/log(2)**
- **Offset=400-(Factor*log(10))**
- **score=Offset+(Factor*log((1-predict_unb_final)/predict_unb_final))**

**Steps to find Probability of default:**

- First we **replace all the Na's in the performance Tag with 1** because these are rejected candidates and would have defaulted if they would have been accepted.
- We observe that the **master file has imbalanced Classification**.
- Hence we will **first balance** the master file with **Synthetic Minority Over Sampling Technique.**
- Next we will apply **Random Forest** on the balanced master file, as it is able to make **equally good positive and negative prediction** unlike the Logistic Regression Model.
- We will **use this model** to predict the probability of default **on the original unbalanced dataset.**
- We will **use this probability in the above formula.**

# Application Score Card

## IDENTIFYING THE CUTOFF SCORE:

- We want to find the **score below which maximum candidates in the rejected dataset is rejected and above which maximum candidates in the selected dataset are selected.**
- **Steps to identify the cutoff:-**
  1. First of all we **will split the final merged data into selected and rejected candidates.**
  2. Next we find the minimum and maximum score in each rejected and selected candidates.

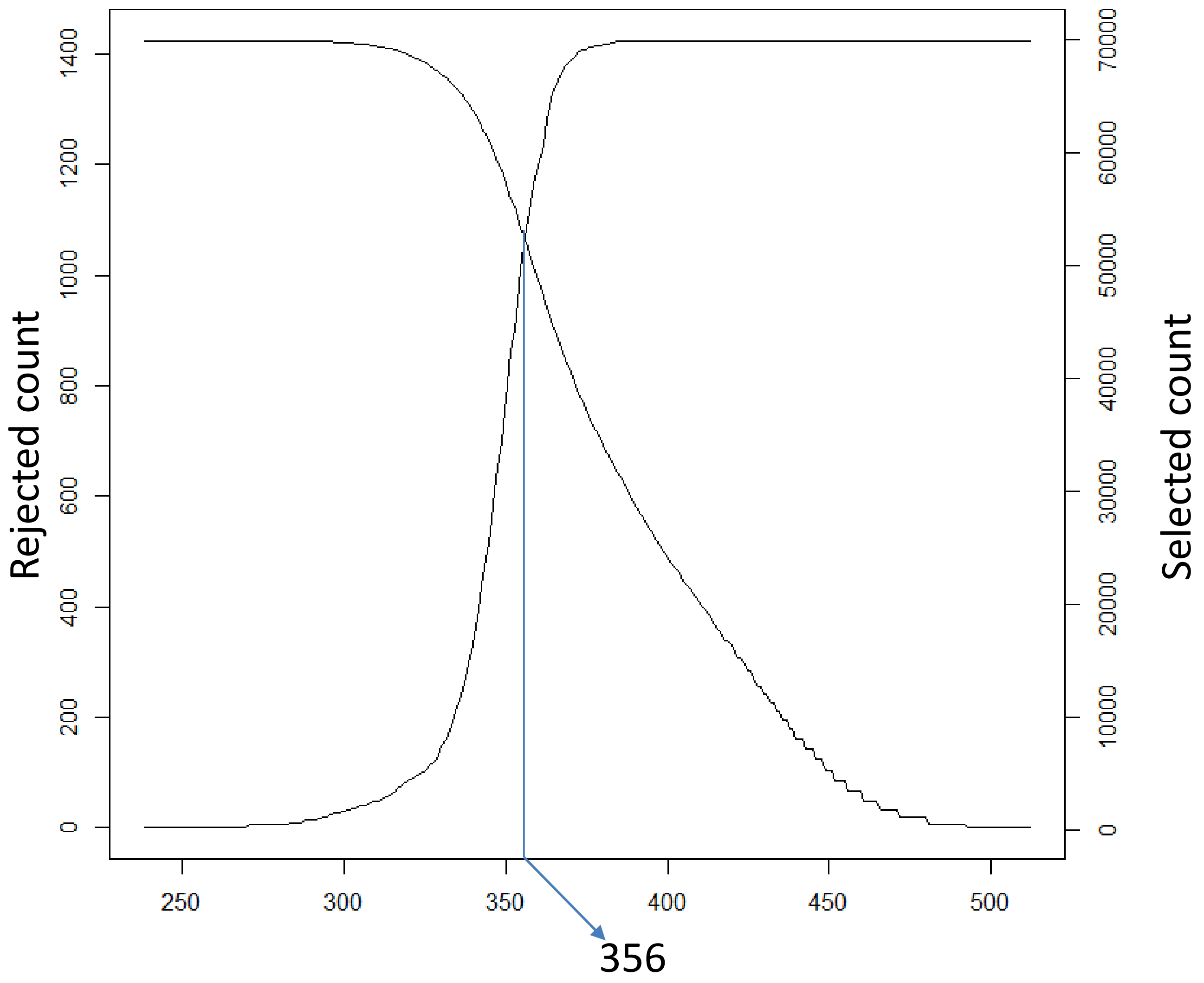
```
> summary(rejected_score$score)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  253.1   340.4   349.1   346.7   356.0   384.8
> summary(selected_score$score)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  250.8   355.7   378.9     Inf   416.3     Inf
```

  3. We observe that **minimum score is 253.1 and the maximum score is 511.81.**
  4. Now we will **create a vector ranging from 201 to 512.**
  5. We will **use this values in this vector to find out how many candidates are selected and how many are rejected.**
  6. The score below which there are highest number of rejected candidates and above which there are highest number of selected candidates are there will be the cutoff score.

# Application Score Card

7. In the graph to the right, the **left y axis label is the rejected count** which was below the cutoff score and the **right y axis label is the selected count** which was above the cutoff point.

8. The **cutoff score is the point at which both the lines are intersecting. The cutoff score is 356.**



# Financial Benefits of the Model

## AUTO APPROVAL AND REJECTION:

```
> table(rejected_score$score<356)

FALSE  TRUE
  423  1002
> table(selected_score$score>356)

FALSE  TRUE
17030 52825
```

As we can see only **33.19%(423/1425) of candidates rejected by the bank are accepted by the model** and **24.37%(17030/69855) of candidates accepted by the bank are rejected by the model.**

# POTENTIAL CREDIT LOSS AVOIDED BY THE MODEL

- The candidates who have been selected by the bank and have defaulted are responsible for the credit loss to the bank. Our model has rejected 24% of selected candidates.
- **We need to find out how many candidates rejected by the model have defaulted.**
- Total number of candidates selected by the bank but defaulted – 2947.
  - 2947/69855 = 4.21%
- No of candidates selected by the model and who defaulted – 884.
- No of candidates selected by the model – 52825
- % of candidates selected by the model and defaulted
  - 884/52825 = 1.67%
- % of employees selected and defaulting before model=4.21%
- % of employees selected and defaulting after model=1.67%
- **Credit loss saved** = 2.54%

# POTENTIAL CREDIT LOSS AVOIDED BY THE MODEL

## Revenue loss

- No of candidates rejected by the model who didn’t default – 14967.
- Total No of candidates who didn’t default – 66908
- % of good candidates rejected by our model – 22.36%.

- **About 22% percent is the revenue loss where we have identified good customers as bad .**

# CONCLUSION

- We identified few important variables that can be used to identify good customers from Logistic Regression Model:
  - Avgas.CC.Utilization.in.last.12.months
  - Outstanding.Balance
  - No.of.times.30.DPD.or.worse.in.last.6.months
  - No.of.times.90.DPD.or.worse.in.last.12.months
  - No.of.times.60.DPD.or.worse.in.last.6.months
  - No.of.Inquiries.in.last.6.months..excluding.home...auto.loans.,

  So we can use these variables while inquiring about customer before giving them loan.

- We can conclude that the model has accurately predicted 74% of the performance Tag in the dataset.

- By tuning the parameters of the model we can increase the accuracy of the model and it can be used to predict whether who will default and who will not default. This can reduce a lot of hours and save a lot of resources at the same time increasing efficiency.

- By this we found out that credit loss % was decreased when we used this model. Hence it is accurate in rejecting the candidate who may default in future.

- Hence this can save a lot of hours, money of the bank and at the same time increase the efficiency and resources of the bank.

# Thank You