,

Financial Risk Analytics Project

Name: Abheer Bandodker

G2-PGPBABIO-May

Batch A

Problem Statement

Dear Participants,

Please find below the graded individual assignment.

You are requested to create an India credit risk(default) model, using the data provided in the spreadsheet raw-data.xlsx, and validate it on validation_data.xlsx. Please use the logistic regression framework to develop the credit default model.

Hints :

Data description - Please direct them to the video - Default Risk Prediction. After removing variables for multicollinearity, we should try to take at least one variable for creating the model from each of the 4 factors namely -

1) Profitability

2) Leverage

3) Liquidity

4) Company's size

In Dr. Sarkar's video of Default Risk Estimation, he has clearly bifurcated all the variables in different buckets.

Creation of new variables - This is an important step in the project as the company which is the biggest in size, will also have bigger asset size, cash flows, etc. (Hint: We need to think in terms of ratios - Equity to asset ratio, debt to equity ratio, etc)

Dependent variable - We need to create a default variable which should take the value of 1 when net worth is negative & 0 when net worth is positive.

Validation Dataset - We need to build the model on the raw dataset and check the model performance measures on the validation dataset.

This project requires you to understand what mode of transport employees prefers to commute to their office. The attached data 'Cars.csv' includes employee information about their mode of transport as well as their personal and professional details like age, salary, work exp. We need to predict whether or not an employee will use Car as a mode of transport. Also, which variables are a significant predictor behind this decision?

Data Importing:

```
> setwd("C:/Users/abheer/Desktop/Data science/Financial_Risk_Analytics/Pro
ject FRA")
> library(readxl)
> company <- read_excel("raw-data.xlsx")
> names(company)
 [1] "Num"
 [2] "Networth Next Year"
 [3] "Total assets"
 [4] "Net worth"
 [5] "Total income"
 [6] "Change in stock"
 [7] "Total expenses"
 [8] "Profit after tax"
 [9] "PBDITA"
[10] "PBT"
[11] "Cash profit"
[12] "PBDITA as % of total income"
[13] "PBT as % of total income"
[14] "PAT as % of total income"
[15] "Cash profit as % of total income"
[16] "PAT as % of net worth"
[17] "Sales"
[18] "Income from financial services"
[19] "Other income"
[20] "Total capital"
[21] "Reserves and funds"
[22] "Deposits (accepted by commercial banks)"
[23] "Borrowings"
[24] "Current liabilities & provisions"
[25] "Deferred tax liability"
[26] "Shareholders funds"
[27] "Cumulative retained profits"
[28] "Capital employed"
[29] "TOL/TNW"
[30] "Total term liabilities / tangible net worth"
[31] "Contingent liabilities / Net worth (%)"
[32] "Contingent liabilities"
[33] "Net fixed assets"
[34] "Investments"
[35] "Current assets"
[36] "Net working capital"
[37] "Quick ratio (times)"
[38] "Current ratio (times)"
[39] "Debt to equity ratio (times)"
[40] "Cash to current liabilities (times)"
[41] "Cash to average cost of sales per day"
[42] "Creditors turnover"
[43] "Debtors turnover"
[44] "Finished goods turnover"
[45] "WIP turnover"
[46] "Raw material turnover"
[47] "Shares outstanding"
[48] "Equity face value"
[49] "EPS"
[50] "Adjusted EPS"
[51] "Total liabilities"
[52] "PE on BSE"
> attach(company)
```

```
> company$`Creditors turnover` = as.numeric(company$`Creditors turnover`)
> company$`Debtors turnover` = as.numeric(company$`Debtors turnover`)
> company$`Finished goods turnover` = as.numeric(company$`Finished goods t
urnover`)
> company$`WIP turnover` = as.numeric(company$`WIP turnover`)
> company$`Raw material turnover` = as.numeric(company$`Raw material turno
ver`)
> company$`Shares outstanding` = as.numeric(company$`Shares outstanding`)
> company$`Equity face value` = as.numeric(company$`Equity face value`)
> company$`PE on BSE` = as.numeric(company$`PE on BSE`)
>
> summary(company)
      Num        Networth Next Year  Total assets
 Min.   :   1   Min.   :-74265.6   Min.   :      0.1
 1st Qu.: 886   1st Qu.:    31.7   1st Qu.:     91.3
 Median :1773   Median :   116.3   Median :    309.7
 Mean   :1772   Mean   :  1616.3   Mean   :   3443.4
 3rd Qu.:2658   3rd Qu.:   456.1   3rd Qu.:   1098.7
 Max.   :3545   Max.   :805773.4   Max.   :1176509.2

   Net worth          Total income        Change in stock
 Min.   :     0.0   Min.   :      0.0   Min.   :-3029.40
 1st Qu.:    31.3   1st Qu.:    106.5   1st Qu.:   -1.80
 Median :   102.3   Median :    444.9   Median :    1.60
 Mean   :  1295.9   Mean   :   4582.8   Mean   :   41.49
 3rd Qu.:   377.3   3rd Qu.:   1440.9   3rd Qu.:   18.05
 Max.   :613151.6   Max.   :2442828.2   Max.   :14185.50
                    NA's   :198         NA's   :458
 Total expenses      Profit after tax         PBDITA
 Min.   :     -0.1   Min.   : -3908.30   Min.   :  -440.7
 1st Qu.:     95.8   1st Qu.:     0.50   1st Qu.:     6.9
 Median :    407.7   Median :     8.80   Median :    35.4
 Mean   :   4262.9   Mean   :   277.36   Mean   :   578.1
 3rd Qu.:   1359.8   3rd Qu.:    52.27   3rd Qu.:   150.2
 Max.   :2366035.3   Max.   :119439.10   Max.   :208576.5
 NA's   :139         NA's   :131         NA's   :131
      PBT              Cash profit        PBDITA as % of total income
 Min.   : -3894.80   Min.   : -2245.70   Min.   :-6400.000
 1st Qu.:     0.70   1st Qu.:     2.90   1st Qu.:    5.000
 Median :    12.40   Median :    18.85   Median :    9.660
 Mean   :   383.81   Mean   :   392.07   Mean   :    4.571
 3rd Qu.:    71.97   3rd Qu.:    93.20   3rd Qu.:   16.390
 Max.   :145292.60   Max.   :176911.80   Max.   :  100.000
 NA's   :131         NA's   :131         NA's   :68
 PBT as % of total income PAT as % of total income
 Min.   :-21340.00        Min.   :-21340.00
 1st Qu.:     0.55        1st Qu.:     0.35
 Median :     3.31        Median :     2.34
 Mean   :   -17.28        Mean   :   -19.20
 3rd Qu.:     8.80        3rd Qu.:     6.34
 Max.   :   100.00        Max.   :   150.00
 NA's   :68               NA's   :68
 Cash profit as % of total income PAT as % of net worth
 Min.   :-15020.000               Min.   :-748.72
 1st Qu.:     2.020               1st Qu.:   0.00
 Median :     5.640               Median :   7.92
 Mean   :    -8.229               Mean   :  10.27
 3rd Qu.:    10.700               3rd Qu.:  20.19
 Max.   :   100.000               Max.   :2466.67
 NA's   :68
     Sales           Income from financial services
 Min.   :      0.1   Min.   :    0.00
 1st Qu.:    112.7   1st Qu.:    0.40
 Median :    453.1   Median :    1.80
 Mean   :   4549.5   Mean   :   80.84
 3rd Qu.:   1433.5   3rd Qu.:    9.68
 Max.   :2384984.4   Max.   :51938.20
 NA's   :259         NA's   :935
  Other income      Total capital     Reserves and funds
```

```
Min.   :    0.00   Min.   :    0.1   Min.   : -6525.9
1st Qu.:    0.40   1st Qu.:   13.1   1st Qu.:     5.0
Median :    1.40   Median :   42.1   Median :    54.8
Mean   :   41.36   Mean   :  216.6   Mean   :  1163.8
3rd Qu.:    5.97   3rd Qu.:  100.3   3rd Qu.:   277.3
Max.   :42856.70   Max.   :78273.2   Max.   :625137.8
NA's   :1295       NA's   :4         NA's   :85
Deposits (accepted by commercial banks)   Borrowings
Mode:logical                              Min.   :     0.10
NA's:3541                                 1st Qu.:    23.95
                                          Median :    99.20
                                          Mean   :  1122.28
                                          3rd Qu.:   352.60
                                          Max.   :278257.30
                                          NA's   :366
Current liabilities & provisions Deferred tax liability
Min.   :     0.1                 Min.   :    0.1
1st Qu.:    17.8                 1st Qu.:    3.2
Median :    69.4                 Median :   13.4
Mean   :   940.6                 Mean   :  227.2
3rd Qu.:   261.7                 3rd Qu.:   50.0
Max.   :352240.3                 Max.   :72796.6
NA's   :96                       NA's   :1140
Shareholders funds Cumulative retained profits Capital employed
Min.   :     0.0   Min.   : -6534.3           Min.   :     0.0
1st Qu.:    32.0   1st Qu.:     1.1           1st Qu.:    60.8
Median :   105.6   Median :    37.1           Median :   214.7
Mean   :  1322.1   Mean   :   890.5           Mean   :  2328.3
3rd Qu.:   393.2   3rd Qu.:   202.3           3rd Qu.:   767.3
Max.   :613151.6   Max.   :390133.8           Max.   :891408.9
                   NA's   :38
   TOL/TNW         Total term liabilities / tangible net worth
Min.   :-350.480   Min.   :-325.600
1st Qu.:   0.600   1st Qu.:   0.050
Median :   1.430   Median :   0.340
Mean   :   3.994   Mean   :   1.844
3rd Qu.:   2.830   3rd Qu.:   1.000
Max.   : 473.000   Max.   : 456.000

Contingent liabilities / Net worth (%) Contingent liabilities
Min.   :    0.00                       Min.   :     0.1
1st Qu.:    0.00                       1st Qu.:     6.3
Median :    5.33                       Median :    38.0
Mean   :   53.94                       Mean   :   932.9
3rd Qu.:   30.76                       3rd Qu.:   192.7
Max.   :14704.27                       Max.   :559506.8
                                       NA's   :1188
Net fixed assets    Investments        Current assets
Min.   :     0.0   Min.   :     0.00   Min.   :     0.1
1st Qu.:    26.0   1st Qu.:     1.00   1st Qu.:    36.2
Median :    93.5   Median :     8.35   Median :   145.1
Mean   :  1189.7   Mean   :   694.73   Mean   :  1293.4
3rd Qu.:   344.9   3rd Qu.:    64.30   3rd Qu.:   502.2
Max.   :636604.6   Max.   :199978.60   Max.   :354815.2
NA's   :118        NA's   :1435        NA's   :66
Net working capital Quick ratio (times) Current ratio (times)
Min.   :-63839.0    Min.   :  0.000     Min.   :  0.00
1st Qu.:    -1.1    1st Qu.:  0.410     1st Qu.:  0.93
Median :    16.2    Median :  0.670     Median :  1.23
Mean   :   138.6    Mean   :  1.401     Mean   :  2.13
3rd Qu.:    84.2    3rd Qu.:  1.030     3rd Qu.:  1.71
Max.   : 85782.8    Max.   :341.000     Max.   :505.00
NA's   :32          NA's   :93          NA's   :93
Debt to equity ratio (times) Cash to current liabilities (times)
Min.   :  0.00               Min.   :  0.0000
1st Qu.:  0.22               1st Qu.:  0.0200
Median :  0.79               Median :  0.0700
Mean   :  2.78               Mean   :  0.4904
3rd Qu.:  1.75               3rd Qu.:  0.1900
```

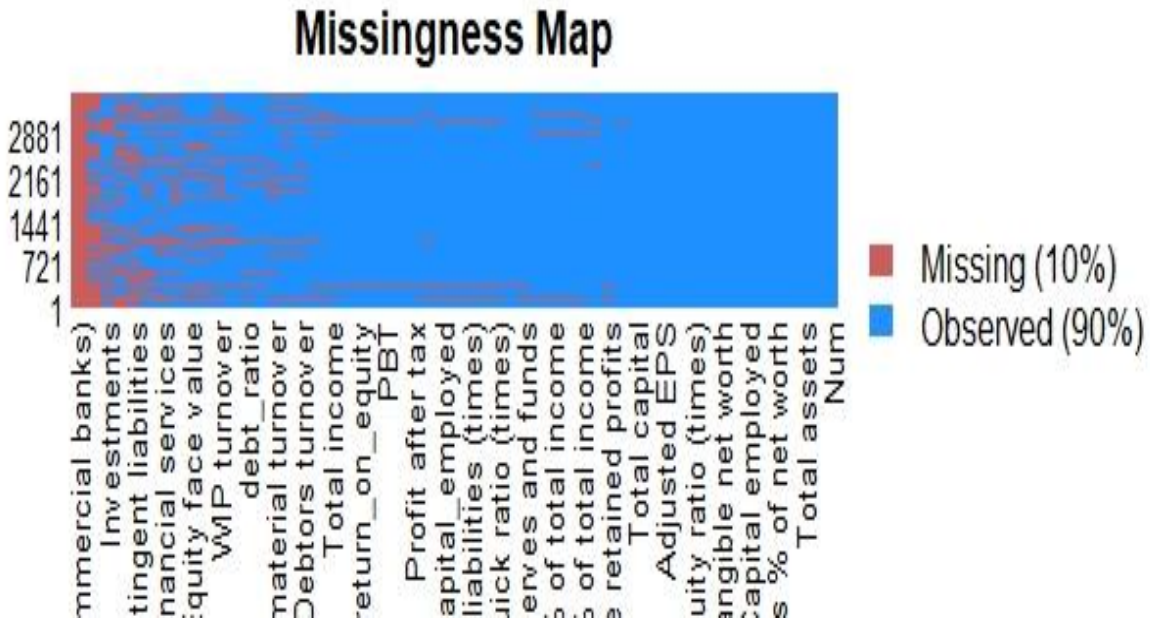
```
Max.   :456.00                Max.   :165.0000
                              NA's   :93
Cash to average cost of sales per day Creditors turnover
Min.   :     0.00                     Min.   :   0.000
1st Qu.:     2.79                     1st Qu.:   3.700
Median :     8.03                     Median :   6.095
Mean   :   158.44                     Mean   :  15.446
3rd Qu.:    21.79                     3rd Qu.:  11.490
Max.   :128040.76                     Max.   :2401.000
NA's   :85                            NA's   :333
Debtors turnover  Finished goods turnover  WIP turnover
Min.   :   0.00   Min.   :   -0.09        Min.   :  -0.18
1st Qu.:   3.76   1st Qu.:    8.20        1st Qu.:   5.10
Median :   6.32   Median :   17.27        Median :   9.76
Mean   :  17.04   Mean   :   87.08        Mean   :  27.93
3rd Qu.:  11.68   3rd Qu.:   40.35        3rd Qu.:  20.24
Max.   :3135.20   Max.   :17947.60        Max.   :5651.40
NA's   :328       NA's   :740             NA's   :640
Raw material turnover Shares outstanding   Equity face value
Min.   :   -2.00      Min.   :-2.147e+09   Min.   :-999999
1st Qu.:    2.99      1st Qu.: 1.316e+06   1st Qu.:     10
Median :    6.40      Median : 4.672e+06   Median :     10
Mean   :   19.09      Mean   : 2.207e+07   Mean   :  -1334
3rd Qu.:   11.85      3rd Qu.: 1.065e+07   3rd Qu.:     10
Max.   :21092.00      Max.   : 4.130e+09   Max.   : 100000
NA's   :361           NA's   :692          NA's   :692
     EPS             Adjusted EPS       Total liabilities
Min.   :-843181.8   Min.   :-843181.8   Min.   :      0.1
1st Qu.:      0.0   1st Qu.:      0.0   1st Qu.:     91.3
Median :      1.4   Median :      1.2   Median :    309.7
Mean   :   -220.3   Mean   :   -221.5   Mean   :   3443.4
3rd Qu.:      9.6   3rd Qu.:      7.5   3rd Qu.:   1098.7
Max.   :  34522.5   Max.   :  34522.5   Max.   :1176509.2

  PE on BSE
Min.   :-1116.64
1st Qu.:    3.27
Median :    9.10
Mean   :   63.91
3rd Qu.:   17.79
Max.   :51002.74
NA's   :2194
```

Creating new variables :

```
> # Creating a new variable debt_ratio
>
> company$Total_capital_employed <- company$`Total assets`/ company$`Curre
nt liabilities & provisions`
> company$debt_ratio <- company$Borrowings / company$Total_capital_employe
d
> # Creating a new variable return_on_equity
>
> company$return_on_equity <- company$`Profit after tax` / company$`Net wo
rth`
>
>
> # Creating a new variable debt_ratio
>
> company$Total_capital_employed <- company$`Total assets`/ company$`Curre
nt liabilities & provisions`
> company$debt_ratio <- company$Borrowings / company$Total_capital_employe
d
```

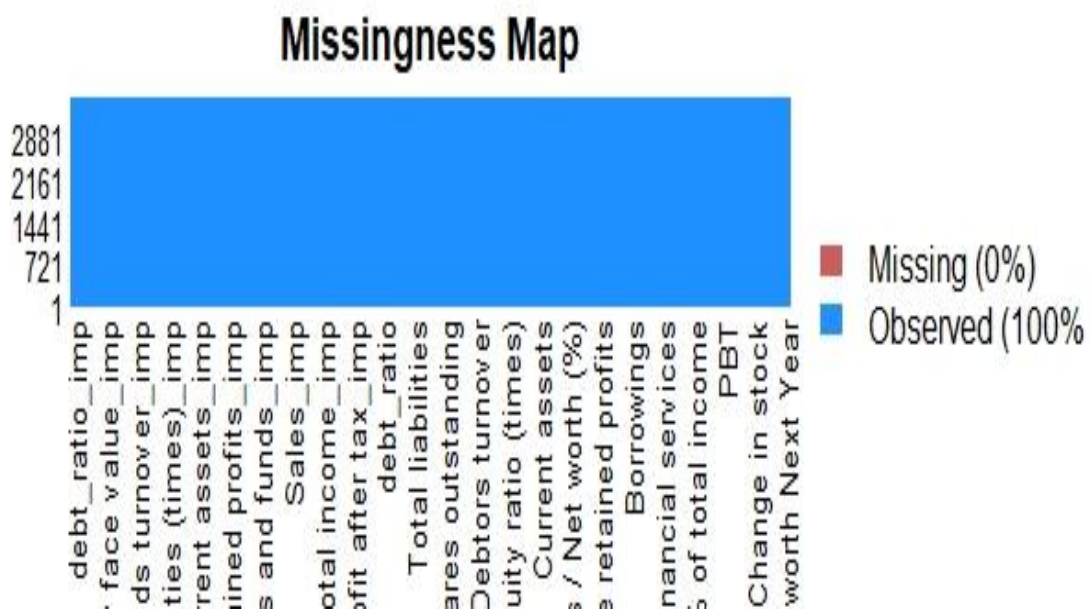
Treating Missing values:

```
> # MIssing value teartment
>
> # Treating missing values
> library(mice)
> library(VIM)
>
> # Displaying a graph to detect any missing data in the dataset
> missmap(company)
```




```
> df <- company[,-c(1,22)]
>
>
> df1 <- kNN(df,
+                  variable = c(
+                                "Total income","Change in stock",
+                                "Total expenses","Profit after tax",
+                                "PBDITA","PBT","Cash profit",
+                                "PBDITA as % of total income",
+                                "PBT as % of total income",
+                                "PAT as % of total income",
+                                "Cash profit as % of total income",
+                                "Sales","Income from financial services",
+                                "Other income",
+                                "Total capital","Reserves and funds",
+                                "Borrowings","Current liabilities & provisi
ons",
+                                "Deferred tax liability",
+                                "Cumulative retained profits",
+                                "Contingent liabilities",
+                                "Net fixed assets","Investments",
+                                "Current assets",
+                                "Net working capital","Quick ratio (times)"
,
+                                "Current ratio (times)",
+                                "Cash to current liabilities (times)",
```

```
+                                "Cash to average cost of sales per day",
+                                "Creditors turnover","Debtors turnover",
+                                "Finished goods turnover","WIP turnover",
+                                "Raw material turnover","Shares outstanding
",
+                                "Equity face value","PE on BSE",
+                                "return_on_equity",
+                                "Total_capital_employed",
+                                "debt_ratio"
+                                ),
+                  k = 6)
>
> missmap(df1)
```



Outlier treatment:

```
> # Outlier treatment
>
> boxplot(df1$`Total assets`, horizontal = T)
> bench1 = 1098.7 + 1.5 * IQR(df1$`Total assets`)
> bench1
[1] 2609.8
> df1$`Total assets`[df1$`Total assets` > bench1] <- bench1
> boxplot(df1$`Total assets`, horizontal = T)
>
> boxplot(df1$`Net worth`, horizontal = T)
> bench2 = 377.3 + 1.5 * IQR(df1$`Net worth`)
> bench2
[1] 896.3
> df1$`Net worth`[df1$`Net worth` > bench2] <- bench2
> boxplot(df1$`Net worth`, horizontal = T)
>
> boxplot(df1$`Total income`, horizontal = T)
> bench3 = 1440.9 + 1.5 * IQR(df1$`Total income`)
> bench3
```

```
[1] 3337.05
> df1$`Total income`[df1$`Total income` > bench3] <- bench3
> boxplot(df1$`Total income`, horizontal = T)
>
> boxplot(df1$`Change in stock`, horizontal = T)
> bench4 = 13.60 + 1.5 * IQR(df1$`Change in stock`)
> bench4
[1] 35.65
> df1$`Change in stock`[df1$`Change in stock` > bench4] <- bench4
> boxplot(df1$`Change in stock`, horizontal = T)
>
> boxplot(df1$`Total expenses`, horizontal = T)
> bench5 = 1284.6 + 1.5 * IQR(df1$`Total expenses`)
> bench5
[1] 3098.475
> df1$`Total expenses`[df1$`Total expenses` > bench5] <- bench5
> boxplot(df1$`Total expenses`, horizontal = T)
>
> boxplot(df1$`Profit after tax`, horizontal = T)
> bench6 = 48.1 + 1.5 * IQR(df1$`Profit after tax`)
> bench6
[1] 119.8
> df1$`Profit after tax`[df1$`Profit after tax` > bench6] <- bench6
> boxplot(df1$`Profit after tax`, horizontal = T)
>
> boxplot(df1$PBDITA, horizontal = T)
> bench7 = 139.1 + 1.5 * IQR(df1$PBDITA)
> bench7
[1] 340.1
> df1$PBDITA[df1$PBDITA > bench7] <- bench7
> boxplot(df1$PBDITA, horizontal = T)
>
> boxplot(df1$PBT, horizontal = T)
> bench8 = 67.5 + 1.5 * IQR(df1$PBT)
> bench8
[1] 168.15
> df1$PBT[df1$PBT > bench8] <- bench8
> boxplot(df1$PBT, horizontal = T)
>
> boxplot(df1$`Cash profit`, horizontal = T)
> bench9 = 86.8 + 1.5 * IQR(df1$`Cash profit`)
> bench9
[1] 214.15
> df1$`Cash profit`[df1$`Cash profit` > bench9] <- bench9
> boxplot(df1$`Cash profit`, horizontal = T)
>
> boxplot(df1$`PBDITA as % of total income`, horizontal = T)
> bench10 = 16.3 + 1.5 * IQR(df1$`PBDITA as % of total income`)
> bench10
[1] 33.505
> df1$`PBDITA as % of total income`[df1$`PBDITA as % of total income`
> bench10] <- bench10
> boxplot(df1$`PBDITA as % of total income`, horizontal = T)
>
> boxplot(df1$`PBT as % of total income`, horizontal = T)
> bench11 = 8.65 + 1.5 * IQR(df1$`PBT as % of total income`)
> bench11
[1] 20.935
> df1$`PBT as % of total income`[df1$`PBT as % of total income` > benc
h11] <- bench11
> boxplot(df1$`PBT as % of total income`, horizontal = T)
>
> boxplot(df1$`PAT as % of total income`, horizontal = T)
> bench12 = 6.27 + 1.5 * IQR(df1$`PAT as % of total income`)
> bench12
[1] 15.285
> df1$`PAT as % of total income`[df1$`PAT as % of total income` > benc
h12] <- bench12
> boxplot(df1$`PAT as % of total income`, horizontal = T)
```

```
>
> boxplot(df1$`Cash profit as % of total income`, horizontal = T)
> bench13 = 10.5  + 1.5 * IQR(df1$`Cash profit as % of total income`)
> bench13
[1] 23.52
> df1$`Cash profit as % of total income`[df1$`Cash profit as % of tota
l income` > bench13] <- bench13
> boxplot(df1$`Cash profit as % of total income`, horizontal = T)
>
> boxplot(df1$`PAT as % of net worth`, horizontal = T)
> bench14 = 20.19  + 1.5 * IQR(df1$`PAT as % of net worth`)
> bench14
[1] 50.475
> df1$`PAT as % of net worth`[df1$`PAT as % of net worth` > bench14] <
- bench14
> boxplot(df1$`PAT as % of net worth`, horizontal = T)
>
> boxplot(df1$Sales, horizontal = T)
> bench15 = 1314.7  + 1.5 * IQR(df1$Sales)
> bench15
[1] 3175
> df1$Sales[df1$Sales > bench15] <- bench15
> boxplot(df1$Sales, horizontal = T)
>
> boxplot(df$`Income from financial services`, horizontal = T)
> bench16 = 5.80  + 1.5 * IQR(df1$`Income from financial services`)
> bench16
[1] 14.05
> df1$`Income from financial services`[df1$`Income from financial serv
ices` > bench16] <- bench16
> boxplot(df1$`Income from financial services`, horizontal = T)
>
> boxplot(df1$`Other income`, horizontal = T)
> bench17 = 2.90  + 1.5 * IQR(df1$`Other income`)
> bench17
[1] 6.8
> df1$`Other income`[df1$`Other income` > bench17] <- bench17
> boxplot(df1$`Other income`, horizontal = T)
>
> boxplot(df1$`Total capital`, horizontal = T)
> bench18 = 100.3  + 1.5 * IQR(df1$`Total capital`)
> bench18
[1] 231.25
> df1$`Total capital`[df1$`Total capital` > bench18] <- bench18
> boxplot(df1$`Total capital`, horizontal = T)
>
> boxplot(df1$`Reserves and funds`, horizontal = T)
> bench19 = 263.2  + 1.5 * IQR(df1$`Reserves and funds`)
> bench19
[1] 652.6
> df1$`Reserves and funds`[df1$`Reserves and funds` > bench19] <- benc
h19
> boxplot(df1$`Reserves and funds`, horizontal = T)
>
> boxplot(df1$Borrowings, horizontal = T)
> bench20 = 303.5  + 1.5 * IQR(df1$Borrowings)
> bench20
[1] 730.7
> df1$Borrowings[df1$Borrowings > bench20] <- bench20
> boxplot(df1$Borrowings, horizontal = T)
>
> boxplot(df1$`Current liabilities & provisions`, horizontal = T)
> bench21 = 249.1  + 1.5 * IQR(df1$`Current liabilities & provisions`)
> bench21
[1] 600.25
> df1$`Current liabilities & provisions`[df1$`Current liabilities & pr
ovisions` > bench21] <- bench21
> boxplot(df1$`Current liabilities & provisions`, horizontal = T)
>
```

```
> boxplot(df1$`Deferred tax liability`, horizontal = T)
> bench22 = 32.9  + 1.5 * IQR(df1$`Deferred tax liability`)
> bench22
[1] 79.175
> df1$`Deferred tax liability`[df1$`Deferred tax liability` > bench22]
<- bench22
> boxplot(df1$`Deferred tax liability`, horizontal = T)
>
> boxplot(df1$`Shareholders funds`, horizontal = T)
> bench23 = 393.2  + 1.5 * IQR(df1$`Shareholders funds`)
> bench23
[1] 935
> df1$`Shareholders funds`[df1$`Shareholders funds` > bench23] <- benc
h23
> boxplot(df1$`Shareholders funds`, horizontal = T)
>
> boxplot(df1$`Cumulative retained profits`, horizontal = T)
> bench24 = 199.4  + 1.5 * IQR(df1$`Cumulative retained profits`)
> bench24
[1] 497.3
> df1$`Cumulative retained profits`[df1$`Cumulative retained profits`
> bench24] <- bench24
> boxplot(df1$`Cumulative retained profits`, horizontal = T)
>
> boxplot(df1$`Capital employed`, horizontal = T)
> bench25 = 767.3  + 1.5 * IQR(df1$`Capital employed`)
> bench25
[1] 1827.05
> df1$`Capital employed`[df1$`Capital employed` > bench25] <- bench25
> boxplot(df1$`Capital employed`, horizontal = T)
>
> boxplot(df1$`TOL/TNW`, horizontal = T)
> bench26 = 2.830  + 1.5 * IQR(df1$`TOL/TNW`)
> bench26
[1] 6.175
> df1$`TOL/TNW`[df1$`TOL/TNW` > bench26] <- bench26
> boxplot(df1$`TOL/TNW`, horizontal = T)
>
> boxplot(df1$`Total term liabilities / tangible net worth`, horizonta
l = T)
> bench27 = 1  + 1.5 * IQR(df1$`Total term liabilities / tangible net
worth`)
> bench27
[1] 2.425
> df1$`Total term liabilities / tangible net worth`[df1$`Total term li
abilities / tangible net worth` > bench27] <- bench27
> boxplot(df1$`Total term liabilities / tangible net worth`, horizonta
l = T)
>
> boxplot(df1$`Contingent liabilities / Net worth (%)`, horizontal = T
)
> bench28 = 30.76  + 1.5 * IQR(df1$`Contingent liabilities / Net worth
(%)`)
> bench28
[1] 76.9
> df1$`Contingent liabilities / Net worth (%)`[df1$`Contingent liabili
ties / Net worth (%)` > bench28] <- bench28
> boxplot(df1$`Contingent liabilities / Net worth (%)`, horizontal = T
)
>
> boxplot(df1$`Contingent liabilities`, horizontal = T)
> bench29 = 94  + 1.5 * IQR(df1$`Contingent liabilities`)
> bench29
[1] 229.45
> df1$`Contingent liabilities`[df1$`Contingent liabilities` > bench29]
<- bench29
> boxplot(df1$`Contingent liabilities`, horizontal = T)
>
> boxplot(df1$`Net fixed assets`, horizontal = T)
```

```
> bench30 = 328.8  + 1.5 * IQR(df1$`Net fixed assets`)
> bench30
[1] 787.5
> df1$`Net fixed assets`[df1$`Net fixed assets` > bench30] <- bench30
> boxplot(df1$`Net fixed assets`, horizontal = T)
>
> boxplot(df1$Investments, horizontal = T)
> bench31 = 23.65  + 1.5 * IQR(df1$Investments)
> bench31
[1] 57.4
> df1$Investments[df1$Investments > bench31] <- bench31
> boxplot(df1$Investments, horizontal = T)
>
> boxplot(df1$`Current assets`, horizontal = T)
> bench32 = 485.9  + 1.5 * IQR(df1$`Current assets`)
> bench32
[1] 1165.7
> df1$`Current assets`[df1$`Current assets` > bench32] <- bench32
> boxplot(df1$`Current assets`, horizontal = T)
>
> boxplot(df1$`Net working capital`, horizontal = T)
> bench33 = 81.6  + 1.5 * IQR(df1$`Net working capital`)
> bench33
[1] 205.5
> df1$`Net working capital`[df1$`Net working capital` > bench33] <- be
nch33
> boxplot(df1$`Net working capital`, horizontal = T)
>
> boxplot(df1$`Quick ratio (times)`, horizontal = T)
> bench34 = 1.060  + 1.5 * IQR(df1$`Quick ratio (times)`)
> bench34
[1] 2.035
> df1$`Quick ratio (times)`[df1$`Quick ratio (times)` > bench34] <- be
nch34
> boxplot(df1$`Quick ratio (times)`, horizontal = T)
>
> boxplot(df1$`Current ratio (times)`, horizontal = T)
> bench35 = 1.740  + 1.5 * IQR(df1$`Current ratio (times)`)
> bench35
[1] 2.955
> df1$`Current ratio (times)`[df1$`Current ratio (times)` > bench35] <
- bench35
> boxplot(df1$`Current ratio (times)`, horizontal = T)
>
> boxplot(df1$`Debt to equity ratio (times)`, horizontal = T)
> bench36 = 1.75  + 1.5 * IQR(df1$`Debt to equity ratio (times)`)
> bench36
[1] 4.045
> df1$`Debt to equity ratio (times)`[df1$`Debt to equity ratio (times)
` > bench36] <- bench36
> boxplot(df1$`Debt to equity ratio (times)`, horizontal = T)
>
> boxplot(df1$`Cash to current liabilities (times)`, horizontal = T)
> bench37 = 0.2  + 1.5 * IQR(df1$`Cash to current liabilities (times)`
)
> bench37
[1] 0.47
> df1$`Cash to current liabilities (times)`[df1$`Cash to current liabi
lities (times)` > bench37] <- bench37
> boxplot(df1$`Cash to current liabilities (times)`, horizontal = T)
>
> boxplot(df1$`Cash to average cost of sales per day`, horizontal = T)
> bench38 = 21.27  + 1.5 * IQR(df1$`Cash to average cost of sales per
day`)
> bench38
[1] 49.26
> df1$`Cash to average cost of sales per day`[df1$`Cash to average cos
t of sales per day` > bench38] <- bench38
> boxplot(df1$`Cash to average cost of sales per day`, horizontal = T)
```

```
>
> boxplot(df1$`Creditors turnover`, horizontal = T)
> bench39 =  10.81 + 1.5 * IQR(df1$`Creditors turnover`)
> bench39
[1] 21.445
> df1$`Creditors turnover`[df1$`Creditors turnover` > bench39] <- benc
h39
> boxplot(df1$`Creditors turnover`, horizontal = T)
>
> boxplot(df1$`Debtors turnover`, horizontal = T)
> bench40 = 10.88  + 1.5 * IQR(df1$`Debtors turnover`)
> bench40
[1] 22.265
> df1$`Debtors turnover`[df1$`Debtors turnover` > bench40] <- bench40
> boxplot(df1$`Debtors turnover`, horizontal = T)
>
> boxplot(df1$`Finished goods turnover`, horizontal = T)
> bench41 = 47.54  + 1.5 * IQR(df1$`Finished goods turnover`)
> bench41
[1] 106.76
> df1$`Finished goods turnover`[df1$`Finished goods turnover` > bench4
1] <- bench41
> boxplot(df1$`Finished goods turnover`, horizontal = T)
>
> boxplot(df1$`WIP turnover`, horizontal = T)
> bench42 = 19.490  + 1.5 * IQR(df1$`WIP turnover`)
> bench42
[1] 42.0575
> df1$`WIP turnover`[df1$`WIP turnover` > bench42] <- bench42
> boxplot(df1$`WIP turnover`, horizontal = T)
>
> boxplot(df1$`Raw material turnover`, horizontal = T)
> bench44 = 11.21  + 1.5 * IQR(df1$`Raw material turnover`)
> bench44
[1] 23.51
> df1$`Raw material turnover`[df1$`Raw material turnover` > bench44] <
- bench44
> boxplot(df1$`Raw material turnover`, horizontal = T)
>
> boxplot(df1$`Shares outstanding`, horizontal = T)
> bench45 =  8.570e+06 + 1.5 * IQR(df1$`Shares outstanding`)
> bench45
[1] 19965308
> df1$`Shares outstanding`[df1$`Shares outstanding` > bench45] <- benc
h45
> boxplot(df1$`Shares outstanding`, horizontal = T)
>
> boxplot(df1$`Equity face value`, horizontal = T)
> bench46 = 10  + 1.5 * IQR(df1$`Equity face value`)
> bench46
[1] 10
> df1$`Equity face value`[df1$`Equity face value` > bench46] <- bench4
6
> boxplot(df1$`Equity face value`, horizontal = T)
>
> boxplot(df1$EPS, horizontal = T)
> bench47 = 9.6  + 1.5 * IQR(df1$EPS)
> bench47
[1] 24.03
> df1$EPS[df1$EPS > bench47] <- bench47
> boxplot(df1$EPS, horizontal = T)
>
> boxplot(df1$`Adjusted EPS`, horizontal = T)
> bench48 = 7.5  + 1.5 * IQR(df1$`Adjusted EPS`)
> bench48
[1] 18.72
> df1$`Adjusted EPS`[df1$`Adjusted EPS` > bench48] <- bench48
> boxplot(df1$`Adjusted EPS`, horizontal = T)
>
```

```
> boxplot(df1$`Total liabilities`, horizontal = T)
> bench49 = 1098.7  + 1.5 * IQR(df1$`Total liabilities`)
> bench49
[1] 2609.8
> df1$`Total liabilities`[df1$`Total liabilities` > bench49] <- bench4
9
> boxplot(df1$`Total liabilities`, horizontal = T)
>
> boxplot(df1$`PE on BSE`, horizontal = T)
> bench50 = 14.41  + 1.5 * IQR(df1$`PE on BSE`)
> bench50
[1] 29.5225
> df1$`PE on BSE`[df1$`PE on BSE` > bench50] <- bench50
> boxplot(df1$`PE on BSE`, horizontal = T)
>
> boxplot(df1$return_on_equity, horizontal = T)
> bench51 = 0.19  + 1.5 * IQR(df1$return_on_equity)
> bench51
[1] 0.4591579
> df1$return_on_equity[df1$return_on_equity > bench51] <- bench51
> boxplot(df1$return_on_equity, horizontal = T)
>
> boxplot(df1$Total_capital_employed, horizontal = T)
> bench52 = 8.653  + 1.5 * IQR(df1$Total_capital_employed)
> bench52
[1] 17.21909
> df1$Total_capital_employed[df1$Total_capital_employed > bench52] <-
bench52
> boxplot(df1$Total_capital_employed, horizontal = T)
>
> boxplot(df1$debt_ratio, horizontal = T)
> bench53 = 65.46  + 1.5 * IQR(df1$debt_ratio)
> bench53
[1] 159.123
> df1$debt_ratio[df1$debt_ratio > bench53] <- bench53
> boxplot(df1$debt_ratio, horizontal = T)

 >
```

Treating multicollinearity:

```
> # Treating multicollinearity
>
> library(faraway)
>
> mymodel = lm(`Networth Next Year` ~`Total assets` + `Net worth`+
+                `Total income` +  `Change in stock` +
+                `Total expenses` + `Profit after tax` +
+                `PBDITA` + `PBT` + `Cash profit` +
+                `PBDITA as % of total income` +
+                `PBT as % of total income` +
+                `PAT as % of total income` +
+                `Cash profit as % of total income` +
+                `Sales` + `Income from financial services` +
+                `Other income` +
+                `Total capital` + `Reserves and funds` +
+                `Borrowings` + `Current liabilities & provisions` +
+                `Deferred tax liability` +
+                `Shareholders funds`+
+                `Cumulative retained profits` +
+                `Capital employed`+
+                `TOL/TNW`+
+                `Total term liabilities / tangible net worth`+
+                `Contingent liabilities / Net worth (%)`+
+                `Contingent liabilities` +
+                `Net fixed assets` + `Investments` +
+                `Current assets` +
+                `Net working capital` + `Quick ratio (times)` +
+                `Current ratio (times)` +
```

```
+                `Debt to equity ratio (times)`+
+                `Cash to current liabilities (times)` +
+                `Cash to average cost of sales per day` +
+                `Creditors turnover` + `Debtors turnover` +
+                `Finished goods turnover` + `WIP turnover` +
+                `Raw material turnover` + `Shares outstanding` +
+                `Equity face value` + `EPS`+`Adjusted EPS` +
+                `Total liabilities`+
+                `PE on BSE` +
+                `return_on_equity` +
+                `Total_capital_employed` +
+                `debt_ratio`, data = df1)
> summary(mymodel)

Call:
lm(formula = `Networth Next Year` ~ `Total assets` + `Net worth` +
    `Total income` + `Change in stock` + `Total expenses` + `Profit after
tax` +
    PBDITA + PBT + `Cash profit` + `PBDITA as % of total income` +
    `PBT as % of total income` + `PAT as % of total income` +
    `Cash profit as % of total income` + Sales + `Income from financial se
rvices` +
    `Other income` + `Total capital` + `Reserves and funds` +
    Borrowings + `Current liabilities & provisions` + `Deferred tax liabil
ity` +
    `Shareholders funds` + `Cumulative retained profits` + `Capital employ
ed` +
    `TOL/TNW` + `Total term liabilities / tangible net worth` +
    `Contingent liabilities / Net worth (%)` + `Contingent liabilities` +
    `Net fixed assets` + Investments + `Current assets` + `Net working cap
ital` +
    `Quick ratio (times)` + `Current ratio (times)` + `Debt to equity rati
o (times)` +
    `Cash to current liabilities (times)` + `Cash to average cost of sales
per day` +
    `Creditors turnover` + `Debtors turnover` + `Finished goods turnover`
+
    `WIP turnover` + `Raw material turnover` + `Shares outstanding` +
    `Equity face value` + EPS + `Adjusted EPS` + `Total liabilities` +
    `PE on BSE` + return_on_equity + Total_capital_employed +
    debt_ratio, data = df1)

Residuals:
    Min      1Q  Median      3Q     Max
-236957   -1369     -31     973  437080

Coefficients: (1 not defined because of singularities)
                                          Estimate Std. Error
(Intercept)                             -2.183e+03  1.019e+03
`Total assets`                          -3.421e+00  2.597e+00
`Net worth`                              6.631e-01  7.770e+00
`Total income`                          -5.848e+00  3.560e+00
`Change in stock`                       -2.492e+01  2.692e+00
`Total expenses`                        -4.449e+00  3.711e+00
`Profit after tax`                      -8.683e+00  1.957e+01
PBDITA                                  -1.370e+01  7.279e+00
PBT                                      1.971e+01  1.796e+01
`Cash profit`                            1.609e+00  9.412e+00
`PBDITA as % of total income`           -1.667e-01  2.248e+00
`PBT as % of total income`               1.871e+00  6.741e+00
`PAT as % of total income`              -1.317e+00  5.839e+00
`Cash profit as % of total income`      -7.478e-02  4.580e+00
Sales                                    1.075e+01  3.315e+00
`Income from financial services`         8.195e+01  7.741e+01
`Other income`                           1.570e+02  1.393e+02
`Total capital`                          5.356e+00  7.188e+00
`Reserves and funds`                    -6.072e-01  2.543e+00
Borrowings                               5.078e+00  3.950e+00
`Current liabilities & provisions`       2.698e+00  5.155e+00
```

```
`Deferred tax liability`                        -5.693e+00  1.900e+01
`Shareholders funds`                            -5.300e-01  8.077e+00
`Cumulative retained profits`                    7.484e-01  2.408e+00
`Capital employed`                               3.670e-01  3.427e+00
`TOL/TNW`                                       -2.091e+02  1.782e+02
`Total term liabilities / tangible net worth`    2.382e+02  2.018e+02
`Contingent liabilities / Net worth (%)`        -1.025e+01  1.330e+01
`Contingent liabilities`                         4.983e+00  5.963e+00
`Net fixed assets`                               2.514e+00  2.874e+00
Investments                                      2.729e+01  1.578e+01
`Current assets`                                 3.017e+00  2.427e+00
`Net working capital`                           -5.677e+00  1.332e-01
`Quick ratio (times)`                            9.040e+02  9.592e+02
`Current ratio (times)`                          4.470e+02  6.500e+02
`Debt to equity ratio (times)`                  -1.109e+02  2.767e+02
`Cash to current liabilities (times)`           -2.992e+02  2.909e+03
`Cash to average cost of sales per day`          9.345e+00  2.322e+01
`Creditors turnover`                            -2.507e+01  4.821e+01
`Debtors turnover`                               8.524e+01  4.925e+01
`Finished goods turnover`                        8.570e+00  9.675e+00
`WIP turnover`                                  -2.666e+01  2.758e+01
`Raw material turnover`                         -4.597e+01  3.993e+01
`Shares outstanding`                             3.923e-05  7.435e-05
`Equity face value`                             -8.503e-02  1.599e-01
EPS                                              2.572e+00  5.716e+01
`Adjusted EPS`                                  -2.570e+00  5.717e+01
`Total liabilities`                                     NA         NA
`PE on BSE`                                     -4.941e-01  7.648e+00
return_on_equity                                -4.541e+01  1.170e+02
Total_capital_employed                          -5.621e-01  6.622e+01
debt_ratio                                       5.952e+00  1.723e+01

                                                t value Pr(>|t|)
(Intercept)                                      -2.143  0.03218 *
`Total assets`                                   -1.317  0.18780
`Net worth`                                       0.085  0.93199
`Total income`                                   -1.643  0.10052
`Change in stock`                                -9.254  < 2e-16 ***
`Total expenses`                                 -1.199  0.23059
`Profit after tax`                               -0.444  0.65725
PBDITA                                           -1.882  0.05986 .
PBT                                               1.097  0.27269
`Cash profit`                                     0.171  0.86430
`PBDITA as % of total income`                    -0.074  0.94087
`PBT as % of total income`                        0.278  0.78137
`PAT as % of total income`                       -0.225  0.82163
`Cash profit as % of total income`               -0.016  0.98697
Sales                                             3.244  0.00119 **
`Income from financial services`                  1.059  0.28987
`Other income`                                    1.127  0.25972
`Total capital`                                   0.745  0.45622
`Reserves and funds`                             -0.239  0.81127
Borrowings                                        1.286  0.19867
`Current liabilities & provisions`                0.523  0.60078
`Deferred tax liability`                         -0.300  0.76455
`Shareholders funds`                             -0.066  0.94769
`Cumulative retained profits`                     0.311  0.75599
`Capital employed`                                0.107  0.91473
`TOL/TNW`                                        -1.174  0.24050
`Total term liabilities / tangible net worth`     1.180  0.23807
`Contingent liabilities / Net worth (%)`         -0.771  0.44101
`Contingent liabilities`                          0.836  0.40344
`Net fixed assets`                                0.875  0.38173
Investments                                       1.729  0.08388 .
`Current assets`                                  1.243  0.21389
`Net working capital`                           -42.612  < 2e-16 ***
`Quick ratio (times)`                             0.942  0.34605
`Current ratio (times)`                           0.688  0.49169
`Debt to equity ratio (times)`                   -0.401  0.68862
`Cash to current liabilities (times)`            -0.103  0.91809
```

```
`Cash to average cost of sales per day`         0.402  0.68735
`Creditors turnover`                           -0.520  0.60313
`Debtors turnover`                              1.731  0.08363 .
`Finished goods turnover`                       0.886  0.37575
`WIP turnover`                                 -0.967  0.33375
`Raw material turnover`                        -1.151  0.24969
`Shares outstanding`                            0.528  0.59775
`Equity face value`                            -0.532  0.59499
EPS                                             0.045  0.96412
`Adjusted EPS`                                 -0.045  0.96414
`Total liabilities`                               NA       NA
`PE on BSE`                                    -0.065  0.94849
return_on_equity                               -0.388  0.69807
Total_capital_employed                         -0.008  0.99323
debt_ratio                                      0.345  0.72979
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13740 on 3490 degrees of freedom
Multiple R-squared:  0.3898,   Adjusted R-squared:  0.381
F-statistic: 44.58 on 50 and 3490 DF,  p-value: < 2.2e-16

>
> vif(mymodel)
                              `Total assets`
                                1.030603e+02
                                 `Net worth`
                                1.110723e+02
                               `Total income`
                                3.089888e+02
                            `Change in stock`
                                1.085817e+00
                             `Total expenses`
                                2.936991e+02
                           `Profit after tax`
                                8.958397e+01
                                      PBDITA
                                1.508390e+01
                                         PBT
                                8.901194e+01
                               `Cash profit`
                                1.704703e+01
                  `PBDITA as % of total income`
                                2.000882e+00
                     `PBT as % of total income`
                                1.512235e+02
                     `PAT as % of total income`
                                1.158394e+02
               `Cash profit as % of total income`
                                3.576276e+01
                                       Sales
                                2.446192e+02
             `Income from financial services`
                                2.948757e+00
                              `Other income`
                                2.075025e+00
                             `Total capital`
                                5.587222e+00
                         `Reserves and funds`
                                9.530070e+00
                                  Borrowings
                                1.944308e+01
               `Current liabilities & provisions`
                                2.195072e+01
                    `Deferred tax liability`
                                5.195326e+00
                        `Shareholders funds`
                                1.304110e+02
                  `Cumulative retained profits`
```

```
                                7.025004e+00
                          `Capital employed`
                                8.791978e+01
                                     `TOL/TNW`
                                2.515665e+01
`Total term liabilities / tangible net worth`
                                2.403297e+01
     `Contingent liabilities / Net worth (%)`
                                2.342293e+00
                  `Contingent liabilities`
                                4.720159e+00
                       `Net fixed assets`
                                1.161945e+01
                                 Investments
                                2.138944e+00
                         `Current assets`
                                1.804040e+01
                    `Net working capital`
                                1.046971e+00
                    `Quick ratio (times)`
                                5.563249e+00
                   `Current ratio (times)`
                                4.466040e+00
           `Debt to equity ratio (times)`
                                2.223272e+00
     `Cash to current liabilities (times)`
                                4.045483e+00
   `Cash to average cost of sales per day`
                                2.852000e+00
                     `Creditors turnover`
                                1.817289e+00
                       `Debtors turnover`
                                1.980499e+00
                `Finished goods turnover`
                                2.187897e+00
                           `WIP turnover`
                                2.429937e+00
                 `Raw material turnover`
                                1.427621e+00
                    `Shares outstanding`
                                5.474308e+02
                     `Equity face value`
                                5.415892e+02
                                        EPS
                                1.243568e+07
                           `Adjusted EPS`
                                1.243577e+07
                     `Total liabilities`
                                8.941062e+02
                                  `PE on BSE`
                                2.562821e+02
                          return_on_equity
                                3.558553e-01
                    Total_capital_employed
                                1.334674e-01
                                  debt_ratio
                                3.884350e-01
```

```
> df2 <- df1[,-c(54:93)]
> names(df2)
 [1] "Networth Next Year"
 [2] "Total assets"
 [3] "Net worth"
 [4] "Total income"
 [5] "Change in stock"
 [6] "Total expenses"
 [7] "Profit after tax"
 [8] "PBDITA"
```

Credit Risk Default Dataset

```
 [9] "PBT"
[10] "Cash profit"
[11] "PBDITA as % of total income"
[12] "PBT as % of total income"
[13] "PAT as % of total income"
[14] "Cash profit as % of total income"
[15] "PAT as % of net worth"
[16] "Sales"
[17] "Income from financial services"
[18] "Other income"
[19] "Total capital"
[20] "Reserves and funds"
[21] "Borrowings"
[22] "Current liabilities & provisions"
[23] "Deferred tax liability"
[24] "Shareholders funds"
[25] "Cumulative retained profits"
[26] "Capital employed"
[27] "TOL/TNW"
[28] "Total term liabilities / tangible net worth"
[29] "Contingent liabilities / Net worth (%)"
[30] "Contingent liabilities"
[31] "Net fixed assets"
[32] "Investments"
[33] "Current assets"
[34] "Net working capital"
[35] "Quick ratio (times)"
[36] "Current ratio (times)"
[37] "Debt to equity ratio (times)"
[38] "Cash to current liabilities (times)"
[39] "Cash to average cost of sales per day"
[40] "Creditors turnover"
[41] "Debtors turnover"
[42] "Finished goods turnover"
[43] "WIP turnover"
[44] "Raw material turnover"
[45] "Shares outstanding"
[46] "Equity face value"
[47] "EPS"
[48] "Adjusted EPS"
[49] "Total liabilities"
[50] "PE on BSE"
[51] "return_on_equity"
[52] "Total_capital_employed"
[53] "debt_ratio"
>
> # Dropping variables with high vif values
>
> # Dropping variables with vif value higher than 8
>
> # Dropping Profit after tax, PBT, Reserves and funds, capital employed
> # and Total liablities
> df3 <- df2[,-c(7,9,20,26,49)]
>
> # Again running multicollinearity test after dropping variables
>
> mymodel2 = lm(`Networth Next Year`~.,data = df3)


> vif(mymodel2)
                              `Total assets`
                                4.838728e+01
                                  `Net worth`
                                1.106614e+02
                               `Total income`
                                2.935703e+02
                            `Change in stock`
                                1.084116e+00
                             `Total expenses`
```

```
                                    2.915394e+02
                                          PBDITA
                                    1.140538e+01
                                `Cash profit`
                                    3.719596e+00
                `PBDITA as % of total income`
                                    2.004532e+00
                   `PBT as % of total income`
                                    1.509884e+02
                   `PAT as % of total income`
                                    1.155436e+02
           `Cash profit as % of total income`
                                    3.566064e+01
                     `PAT as % of net worth`
                                    1.662655e+00
                                           Sales
                                    2.269966e+02
             `Income from financial services`
                                    2.944113e+00
                               `Other income`
                                    2.064636e+00
                             `Total capital`
                                    5.465492e+00
                                      Borrowings
                                    1.650903e+01
           `Current liabilities & provisions`
                                    1.960926e+01
                     `Deferred tax liability`
                                    5.185654e+00
                        `Shareholders funds`
                                    1.235328e+02
                `Cumulative retained profits`
                                    2.610568e+00
                                       `TOL/TNW`
                                    2.495522e+01
`Total term liabilities / tangible net worth`
                                    2.381776e+01
     `Contingent liabilities / Net worth (%)`
                                    2.331981e+00
                     `Contingent liabilities`
                                    4.699105e+00
                          `Net fixed assets`
                                    1.142183e+01
                                     Investments
                                    2.128891e+00
                            `Current assets`
                                    1.799804e+01
                       `Net working capital`
                                    1.044105e+00
                       `Quick ratio (times)`
                                    5.563561e+00
                     `Current ratio (times)`
                                    4.463978e+00
               `Debt to equity ratio (times)`
                                    2.202337e+00
        `Cash to current liabilities (times)`
                                    4.041475e+00
      `Cash to average cost of sales per day`
                                    2.846677e+00
                        `Creditors turnover`
                                    1.813918e+00
                          `Debtors turnover`
                                    1.993276e+00
                   `Finished goods turnover`
                                    2.187405e+00
                              `WIP turnover`
                                    2.425439e+00
                     `Raw material turnover`
                                    1.433874e+00
```

```
                          `Shares outstanding`
                                  5.454653e+02
                           `Equity face value`
                                  5.396141e+02
                                           EPS
                                  1.237726e+07
                                `Adjusted EPS`
                                  1.237732e+07
                                  `PE on BSE`
                                  1.093256e+00
                             return_on_equity
                                  1.292449e+00
                       Total_capital_employed
                                  1.960244e+00
                                    debt_ratio
                                  1.681835e+01
>
> # Dropping variables with vif value higher than 4
>
> # Dropping total assets, Total capital, Deferred tax liability,
> # Contingent liabilities, Quick ratio (times),
> # Current ratio (times), Cash to current liabilities (times),
> # Shares outstanding and Equity face value,
> names(df3)
 [1] "Networth Next Year"
 [2] "Total assets"
 [3] "Net worth"
 [4] "Total income"
 [5] "Change in stock"
 [6] "Total expenses"
 [7] "PBDITA"
 [8] "Cash profit"
 [9] "PBDITA as % of total income"
[10] "PBT as % of total income"
[11] "PAT as % of total income"
[12] "Cash profit as % of total income"
[13] "PAT as % of net worth"
[14] "Sales"
[15] "Income from financial services"
[16] "Other income"
[17] "Total capital"
[18] "Borrowings"
[19] "Current liabilities & provisions"
[20] "Deferred tax liability"
[21] "Shareholders funds"
[22] "Cumulative retained profits"
[23] "TOL/TNW"
[24] "Total term liabilities / tangible net worth"
[25] "Contingent liabilities / Net worth (%)"
[26] "Contingent liabilities"
[27] "Net fixed assets"
[28] "Investments"
[29] "Current assets"
[30] "Net working capital"
[31] "Quick ratio (times)"
[32] "Current ratio (times)"
[33] "Debt to equity ratio (times)"
[34] "Cash to current liabilities (times)"
[35] "Cash to average cost of sales per day"
[36] "Creditors turnover"
[37] "Debtors turnover"
[38] "Finished goods turnover"
[39] "WIP turnover"
[40] "Raw material turnover"
[41] "Shares outstanding"
[42] "Equity face value"
[43] "EPS"
[44] "Adjusted EPS"
[45] "PE on BSE"
```

```
[46] "return_on_equity"
[47] "Total_capital_employed"
[48] "debt_ratio"
> df4 <- df3[,-c(2,17,20,26,31,32,34,41,42)]
>
> # Again running multicollinearity test after dropping variables
>
> mymodel3 = lm(`Networth Next Year`~.,data = df4)


> vif(mymodel3)
                                  `Net worth`
                                1.075250e+02
                               `Total income`
                                2.879463e+02
                             `Change in stock`
                                1.077314e+00
                              `Total expenses`
                                2.886025e+02
                                      PBDITA
                                1.124951e+01
                               `Cash profit`
                                3.645238e+00
                  `PBDITA as % of total income`
                                1.999056e+00
                     `PBT as % of total income`
                                1.508962e+02
                     `PAT as % of total income`
                                1.154042e+02
             `Cash profit as % of total income`
                                3.563261e+01
                         `PAT as % of net worth`
                                1.656407e+00
                                       Sales
                                2.253928e+02
               `Income from financial services`
                                2.918358e+00
                              `Other income`
                                2.050746e+00
                                Borrowings
                                1.404886e+01
             `Current liabilities & provisions`
                                1.611905e+01
                        `Shareholders funds`
                                1.106090e+02
                  `Cumulative retained profits`
                                2.326087e+00
                                   `TOL/TNW`
                                2.422591e+01
`Total term liabilities / tangible net worth`
                                2.311767e+01
     `Contingent liabilities / Net worth (%)`
                                1.281591e+00
                             `Net fixed assets`
                                6.763861e+00
                                 Investments
                                2.093433e+00
                              `Current assets`
                                1.560046e+01
                           `Net working capital`
                                1.036694e+00
                `Debt to equity ratio (times)`
                                2.130993e+00
      `Cash to average cost of sales per day`
                                1.205896e+00
                           `Creditors turnover`
                                1.721375e+00
                             `Debtors turnover`
                                1.517837e+00
```

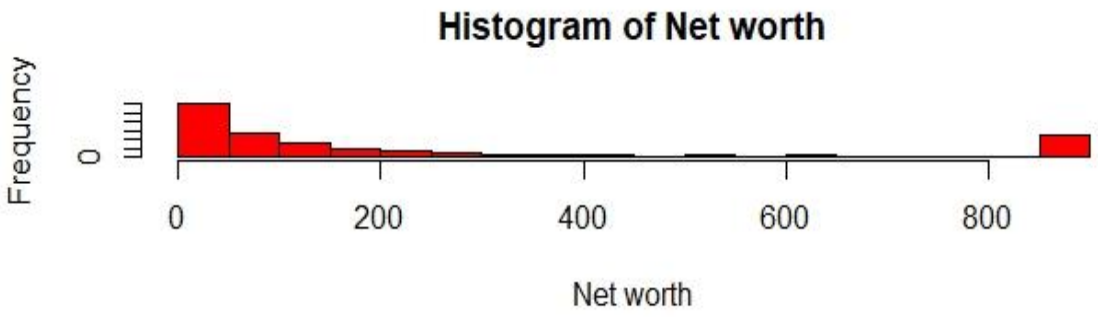
```
                    `Finished goods turnover`
                                 2.171568e+00
                                `WIP turnover`
                                 2.338768e+00
                      `Raw material turnover`
                                 1.379100e+00
                                          EPS
                                 1.205812e+07
                               `Adjusted EPS`
                                 1.205812e+07
                                  `PE on BSE`
                                 1.084546e+00
                            return_on_equity
                                 1.290604e+00
                     Total_capital_employed
                                 1.788498e+00
                                  debt_ratio
                                 1.638325e+01
>
> # Dropping variables with vif value higher than 4
>
> # Dropping net fixed assets,
>
> names(df4)
 [1] "Networth Next Year"
 [2] "Net worth"
 [3] "Total income"
 [4] "Change in stock"
 [5] "Total expenses"
 [6] "PBDITA"
 [7] "Cash profit"
 [8] "PBDITA as % of total income"
 [9] "PBT as % of total income"
[10] "PAT as % of total income"
[11] "Cash profit as % of total income"
[12] "PAT as % of net worth"
[13] "Sales"
[14] "Income from financial services"
[15] "Other income"
[16] "Borrowings"
[17] "Current liabilities & provisions"
[18] "Shareholders funds"
[19] "Cumulative retained profits"
[20] "TOL/TNW"
[21] "Total term liabilities / tangible net worth"
[22] "Contingent liabilities / Net worth (%)"
[23] "Net fixed assets"
[24] "Investments"
[25] "Current assets"
[26] "Net working capital"
[27] "Debt to equity ratio (times)"
[28] "Cash to average cost of sales per day"
[29] "Creditors turnover"
[30] "Debtors turnover"
[31] "Finished goods turnover"
[32] "WIP turnover"
[33] "Raw material turnover"
[34] "EPS"
[35] "Adjusted EPS"
[36] "PE on BSE"
[37] "return_on_equity"
[38] "Total_capital_employed"
[39] "debt_ratio"
> df5 <- df4[,-c(23)]
>
> # Again running multicollinearity test after dropping variables
>
> mymodel4 = lm(`Networth Next Year`~.,data = df5)
```

```
> vif(mymodel4)
                                    `Net worth`
                                  1.073079e+02
                                 `Total income`
                                  2.862631e+02
                               `Change in stock`
                                  1.077266e+00
                                `Total expenses`
                                  2.886020e+02
                                         PBDITA
                                  1.093829e+01
                                 `Cash profit`
                                  3.645238e+00
                     `PBDITA as % of total income`
                                  1.999051e+00
                        `PBT as % of total income`
                                  1.507862e+02
                        `PAT as % of total income`
                                  1.152953e+02
                `Cash profit as % of total income`
                                  3.563189e+01
                           `PAT as % of net worth`
                                  1.643162e+00
                                           Sales
                                  2.230898e+02
                  `Income from financial services`
                                  2.909201e+00
                                `Other income`
                                  2.020953e+00
                                  Borrowings
                                  1.208969e+01
                `Current liabilities & provisions`
                                  1.566808e+01
                             `Shareholders funds`
                                  1.086679e+02
                     `Cumulative retained profits`
                                  2.318032e+00
                                     `TOL/TNW`
                                  2.415486e+01
`Total term liabilities / tangible net worth`
                                  2.305454e+01
     `Contingent liabilities / Net worth (%)`
                                  1.279264e+00
                                   Investments
                                  2.063975e+00
                                `Current assets`
                                  1.487065e+01
                           `Net working capital`
                                  1.036017e+00
                  `Debt to equity ratio (times)`
                                  2.126768e+00
      `Cash to average cost of sales per day`
                                  1.203293e+00
                           `Creditors turnover`
                                  1.705155e+00
                             `Debtors turnover`
                                  1.495582e+00
                     `Finished goods turnover`
                                  2.169022e+00
                                `WIP turnover`
                                  2.337560e+00
                       `Raw material turnover`
                                  1.377496e+00
                                           EPS
                                  1.204885e+07
                               `Adjusted EPS`
                                  1.204884e+07
                                  `PE on BSE`
                                  1.084508e+00
```

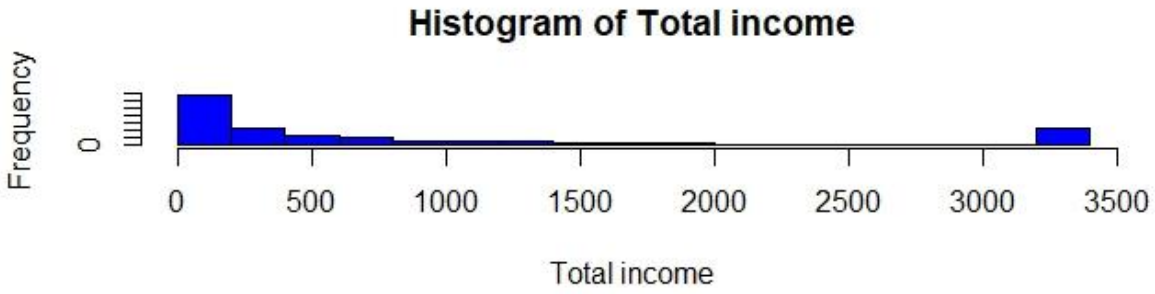
```
                              return_on_equity
                                  1.290463e+00
                        Total_capital_employed
                                  1.781003e+00
                                    debt_ratio
                                  1.620950e+01
>
> # There is no more multicollinearity withing the remaining variables
>
> names(df5)
 [1] "Networth Next Year"
 [2] "Net worth"
 [3] "Total income"
 [4] "Change in stock"
 [5] "Total expenses"
 [6] "PBDITA"
 [7] "Cash profit"
 [8] "PBDITA as % of total income"
 [9] "PBT as % of total income"
[10] "PAT as % of total income"
[11] "Cash profit as % of total income"
[12] "PAT as % of net worth"
[13] "Sales"
[14] "Income from financial services"
[15] "Other income"
[16] "Borrowings"
[17] "Current liabilities & provisions"
[18] "Shareholders funds"
[19] "Cumulative retained profits"
[20] "TOL/TNW"
[21] "Total term liabilities / tangible net worth"
[22] "Contingent liabilities / Net worth (%)"
[23] "Investments"
[24] "Current assets"
[25] "Net working capital"
[26] "Debt to equity ratio (times)"
[27] "Cash to average cost of sales per day"
[28] "Creditors turnover"
[29] "Debtors turnover"
[30] "Finished goods turnover"
[31] "WIP turnover"
[32] "Raw material turnover"
[33] "EPS"
[34] "Adjusted EPS"
[35] "PE on BSE"
[36] "return_on_equity"
[37] "Total_capital_employed"
[38] "debt_ratio"
```

Univariate Analysis:

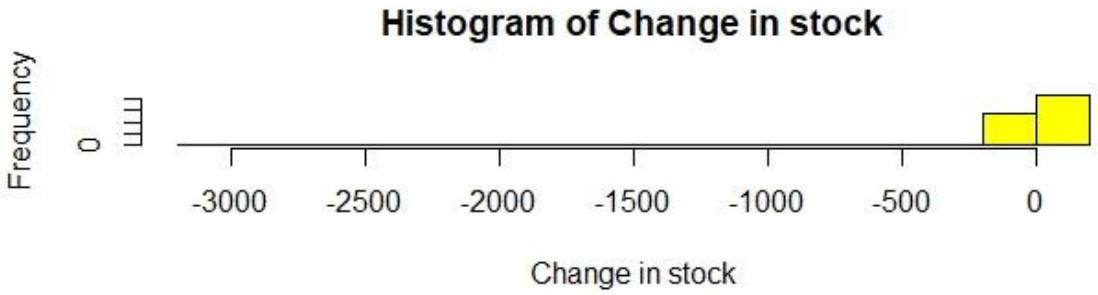
```
> # Univariate analysis
> attach(df5)
> hist(`Net worth`, col = "Red")
```



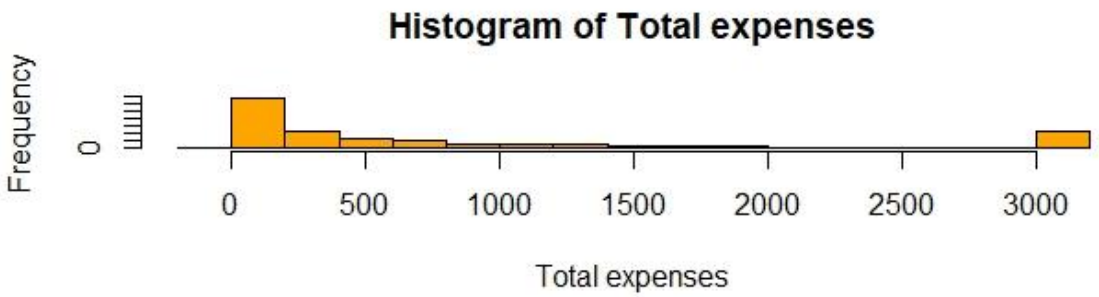

```
> hist(`Total income`, col = "Blue")
```




```
> hist(`Change in stock`, col = "Yellow")
```



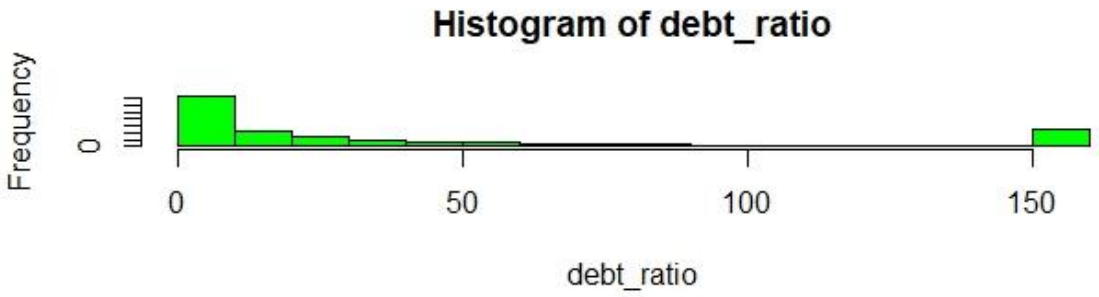

```
> hist(`Total expenses`, col = "Orange")
```



Similarly the univariate analysis has been conducted for various variables

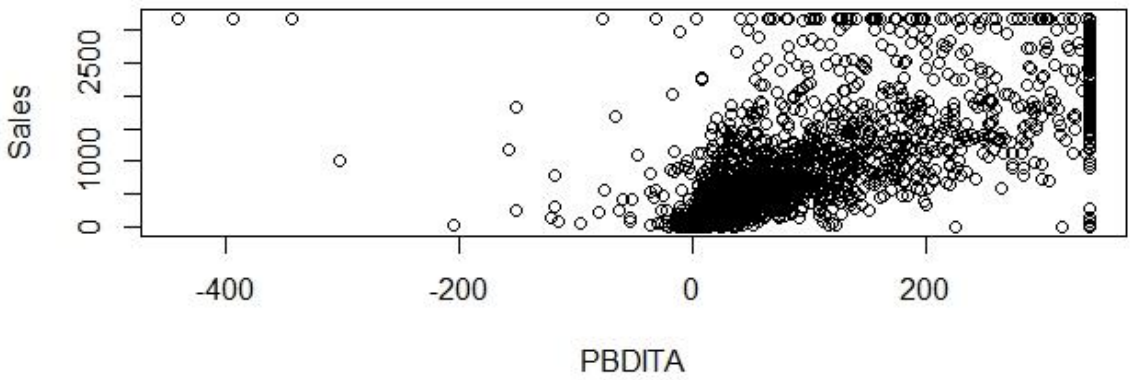
```
> hist(PBDITA, col = "Red")
> hist(`Cash profit`, col = "Blue")
> hist(`PBDITA as % of total income`, col = "Yellow")
> hist(`PBT as % of total income`, col = "Orange")
>
> hist(`PAT as % of total income`, col = "Red")
> hist(`Cash profit as % of total income`, col = "Blue")
> hist(`PAT as % of net worth`, col = "Yellow")
> hist(Sales, col = "Orange")
>
> hist(`Income from financial services`, col = "Red")
> hist(`Other income`, col = "Blue")
> hist(Borrowings, col = "Yellow")
> hist(`Current liabilities & provisions`, col = "Orange")
>
> hist(`Shareholders funds`, col = "Red")
> hist(`Cumulative retained profits`, col = "Blue")
> hist(`TOL/TNW`, col = "Yellow")
> hist(`Total term liabilities / tangible net worth`, col = "Orange")
>
> hist(`Contingent liabilities / Net worth (%)`, col = "Red")
> hist(Investments, col = "Blue")
> hist(`Current assets`, col = "Yellow")
> hist(`Net working capital`, col = "Orange")
>
> hist(`Debt to equity ratio (times)`, col = "Red")
> hist(`Cash to average cost of sales per day`, col = "Blue")
> hist(`Creditors turnover`, col = "Yellow")
> hist(`Debtors turnover`, col = "Orange")
>
> hist(`Finished goods turnover`, col = "Red")
> hist(`WIP turnover`, col = "Blue")
> hist(`Raw material turnover`, col = "Yellow")
> hist(EPS, col = "Orange")
>
> hist(`Adjusted EPS`, col = "Red")
> hist(`PE on BSE`, col = "Blue")
> hist(return_on_equity, col = "Yellow")
> hist(Total_capital_employed, col = "Orange")
```

```
> hist(debt_ratio, col = "Green")
```



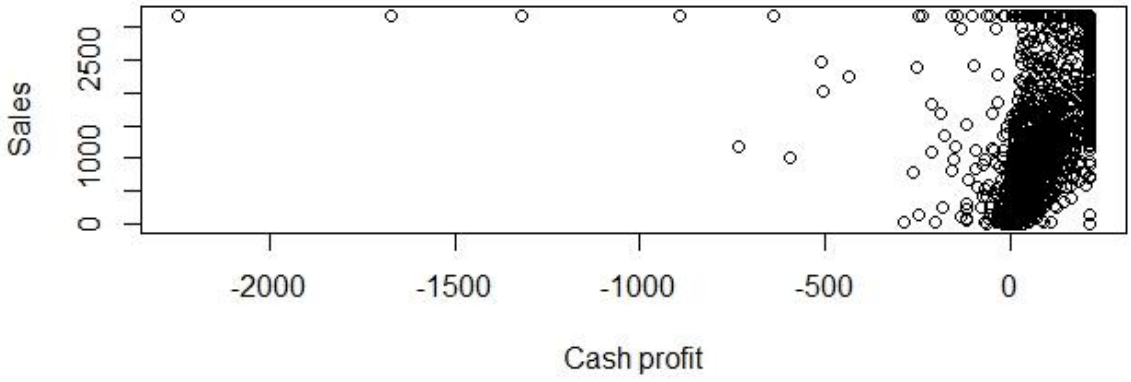
Bivariate Analysis:

```
> # bi-variate analysis
>
> plot(Sales~PBDITA, data = df5)
```



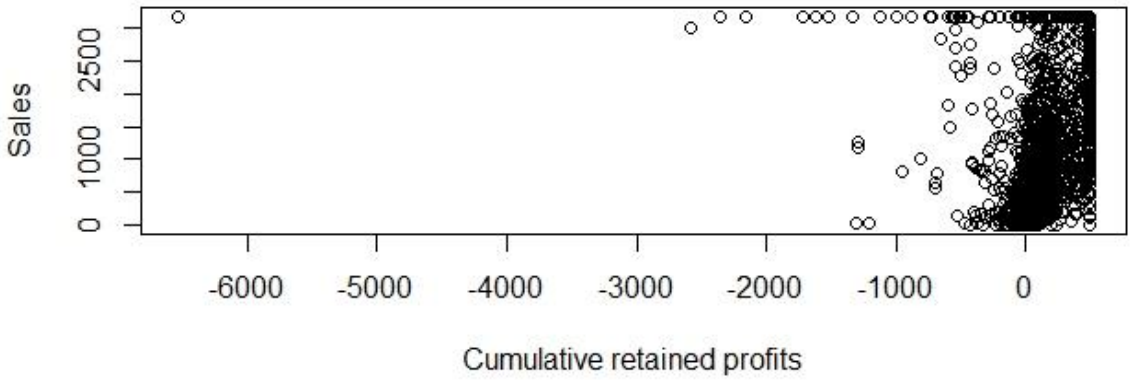
As the PBDITA increases so does sales they are both highly corelated

```
> plot(Sales~`Cash profit`, data = df5)
>
```



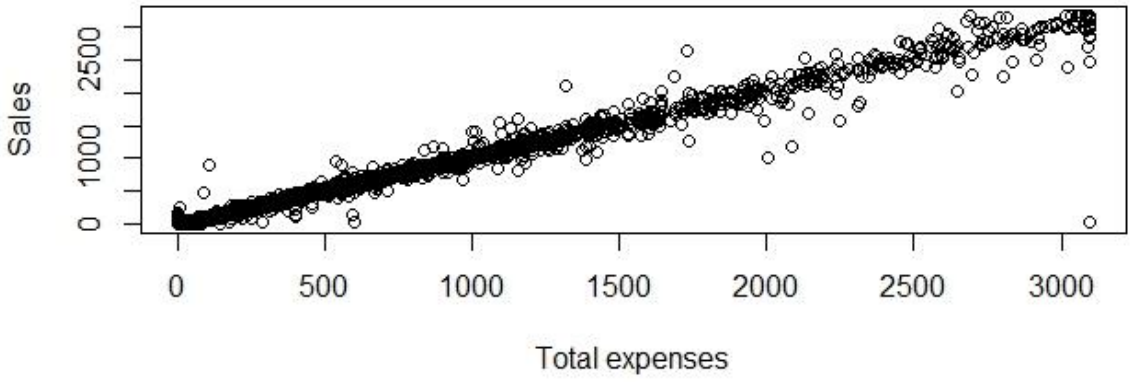
As the Cash profit increases so does sales they are both highly corelated

```
> plot(Sales~`Cumulative retained profits`, data = df5)
```



As the Cumulative retained profits increases so does sales they are both highly corelated

```
> plot(Sales~`Total expenses`, data = df5)
```



Balancing the uneven dataset using smote :

```
> Default <- ifelse(df5$`Networth Next Year`>0,0,1)
Warning messages:
1: In doTryCatch(return(expr), name, parentenv, handler) :
  display list redraw incomplete
2: In doTryCatch(return(expr), name, parentenv, handler) :
  invalid graphics state
3: In doTryCatch(return(expr), name, parentenv, handler) :
  invalid graphics state
>
> summary(as.factor(Default))
   0    1
3298  243
>
> 243/(243+3298)
[1] 0.06862468
>
> # This is an unbalances dataset as the number of default scenerios are
> # only 6.8%
>
> # Dropping the Networth Next Year column
> df6 <- df5[,-c(1)]
>
> names(df6)
 [1] "Net worth"
 [2] "Total income"
 [3] "Change in stock"
 [4] "Total expenses"
 [5] "PBDITA"
 [6] "Cash profit"
 [7] "PBDITA as % of total income"
 [8] "PBT as % of total income"
 [9] "PAT as % of total income"
[10] "Cash profit as % of total income"
[11] "PAT as % of net worth"
[12] "Sales"
[13] "Income from financial services"
[14] "Other income"
[15] "Borrowings"
[16] "Current liabilities & provisions"
[17] "Shareholders funds"
[18] "Cumulative retained profits"
[19] "TOL/TNW"
[20] "Total term liabilities / tangible net worth"
[21] "Contingent liabilities / Net worth (%)"
[22] "Investments"
```

```
[23] "Current assets"
[24] "Net working capital"
[25] "Debt to equity ratio (times)"
[26] "Cash to average cost of sales per day"
[27] "Creditors turnover"
[28] "Debtors turnover"
[29] "Finished goods turnover"
[30] "WIP turnover"
[31] "Raw material turnover"
[32] "EPS"
[33] "Adjusted EPS"
[34] "PE on BSE"
[35] "return_on_equity"
[36] "Total_capital_employed"
[37] "debt_ratio"
> default_dataset <- cbind(df6,Default)
> names(default_dataset)
 [1] "Net worth"
 [2] "Total income"
 [3] "Change in stock"
 [4] "Total expenses"
 [5] "PBDITA"
 [6] "Cash profit"
 [7] "PBDITA as % of total income"
 [8] "PBT as % of total income"
 [9] "PAT as % of total income"
[10] "Cash profit as % of total income"
[11] "PAT as % of net worth"
[12] "Sales"
[13] "Income from financial services"
[14] "Other income"
[15] "Borrowings"
[16] "Current liabilities & provisions"
[17] "Shareholders funds"
[18] "Cumulative retained profits"
[19] "TOL/TNW"
[20] "Total term liabilities / tangible net worth"
[21] "Contingent liabilities / Net worth (%)"
[22] "Investments"
[23] "Current assets"
[24] "Net working capital"
[25] "Debt to equity ratio (times)"
[26] "Cash to average cost of sales per day"
[27] "Creditors turnover"
[28] "Debtors turnover"
[29] "Finished goods turnover"
[30] "WIP turnover"
[31] "Raw material turnover"
[32] "EPS"
[33] "Adjusted EPS"
[34] "PE on BSE"
[35] "return_on_equity"
[36] "Total_capital_employed"
[37] "debt_ratio"
[38] "Default"
>
> default_dataset$Default = as.factor(default_dataset$Default)
>
> # Using the smote function to make the dataset evenly distrubuted
> library(DMwR)
> Default.smote<- SMOTE(Default ~ ., default_dataset,
+                       perc.over = 700, perc.under=150)
>
> table(Default.smote$Default)

   0    1
2551 1944
>
> 1944/(2551+1944)
```

Credit Risk Default Dataset

```
[1] 0.4324805
>
>
> # 43% of the dataset consists of default scenerios
```

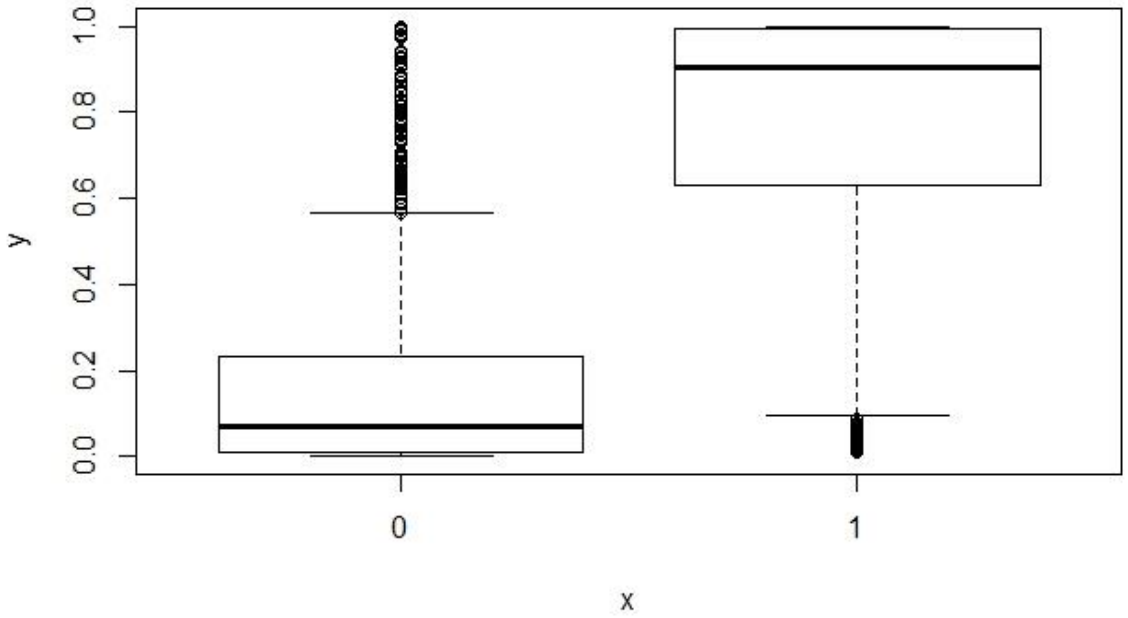
Implementing logistic regression:

```
> # Implementing logistic regression to develop a model
>
> Default.smote$Default = as.factor(Default.smote$Default)
> default_logistic1 <- glm(Default~., data=Default.smote,
+                          family=binomial)
Warning message:
glm.fit: fitted probabilities numerically 0 or 1 occurred
>
> summary(default_logistic1)

Call:
glm(formula = Default ~ ., family = binomial, data = Default.smote)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.0233  -0.4324  -0.0321   0.2086   2.7554

Coefficients:
                                                  Estimate Std. Error
(Intercept)                                     -0.5050477  0.1678839
`Net worth`                                      0.0009433  0.0021712
`Total income`                                   0.0064298  0.0023079
`Change in stock`                                0.0091258  0.0033361
`Total expenses`                                -0.0012095  0.0021163
PBDITA                                          -0.0005841  0.0028663
`Cash profit`                                   -0.0164670  0.0034529
`PBDITA as % of total income`                    0.0014121  0.0011079
`PBT as % of total income`                       0.0355904  0.0082698
`PAT as % of total income`                      -0.0349621  0.0082493
`Cash profit as % of total income`              -0.0029032  0.0009713
`PAT as % of net worth`                         -0.0324230  0.0040430
Sales                                           -0.0050085  0.0023408
`Income from financial services`                -0.0669538  0.0301586
`Other income`                                   0.1092311  0.0440862
Borrowings                                      -0.0003779  0.0009186
`Current liabilities & provisions`               0.0041528  0.0018271
`Shareholders funds`                            -0.0024645  0.0020215
`Cumulative retained profits`                   -0.0038295  0.0004973
`TOL/TNW`                                        0.2221806  0.0368200
`Total term liabilities / tangible net worth` -0.3726707  0.0885110
`Contingent liabilities / Net worth (%)`         0.0025909  0.0023022
Investments                                     -0.0036614  0.0050661
`Current assets`                                 0.0008506  0.0009443
`Net working capital`                            0.0003906  0.0008923
`Debt to equity ratio (times)`                   0.5072237  0.0517371
`Cash to average cost of sales per day`         -0.0071203  0.0031464
`Creditors turnover`                            -0.0491959  0.0126053
`Debtors turnover`                              -0.0084574  0.0110153
`Finished goods turnover`                        0.0004673  0.0022261
`WIP turnover`                                  -0.0038471  0.0066927
`Raw material turnover`                         -0.0769518  0.0114383
EPS                                              0.0454295  0.0357457
`Adjusted EPS`                                  -0.0470476  0.0359784
`PE on BSE`                                     -0.0035671  0.0017967
return_on_equity                                -0.8879059  0.2393585
Total_capital_employed                           0.0027296  0.0121407
debt_ratio                                      -0.0271017  0.0059427
                                                z value Pr(>|z|)
(Intercept)                                      -3.008 0.002627 **
`Net worth`                                       0.434 0.663961
```

```
`Total income`                                    2.786 0.005336 **
`Change in stock`                                 2.735 0.006229 **
`Total expenses`                                 -0.572 0.567654
PBDITA                                           -0.204 0.838522
`Cash profit`                                    -4.769 1.85e-06 ***
`PBDITA as % of total income`                     1.275 0.202473
`PBT as % of total income`                        4.304 1.68e-05 ***
`PAT as % of total income`                       -4.238 2.25e-05 ***
`Cash profit as % of total income`               -2.989 0.002800 **
`PAT as % of net worth`                          -8.020 1.06e-15 ***
Sales                                            -2.140 0.032383 *
`Income from financial services`                 -2.220 0.026415 *
`Other income`                                    2.478 0.013224 *
Borrowings                                       -0.411 0.680767
`Current liabilities & provisions`                2.273 0.023033 *
`Shareholders funds`                             -1.219 0.222790
`Cumulative retained profits`                    -7.701 1.35e-14 ***
`TOL/TNW`                                         6.034 1.60e-09 ***
`Total term liabilities / tangible net worth`    -4.210 2.55e-05 ***
`Contingent liabilities / Net worth (%)`          1.125 0.260414
Investments                                      -0.723 0.469842
`Current assets`                                  0.901 0.367721
`Net working capital`                             0.438 0.661582
`Debt to equity ratio (times)`                    9.804  < 2e-16 ***
`Cash to average cost of sales per day`          -2.263 0.023635 *
`Creditors turnover`                             -3.903 9.51e-05 ***
`Debtors turnover`                               -0.768 0.442616
`Finished goods turnover`                         0.210 0.833745
`WIP turnover`                                   -0.575 0.565416
`Raw material turnover`                          -6.728 1.73e-11 ***
EPS                                               1.271 0.203762
`Adjusted EPS`                                   -1.308 0.190988
`PE on BSE`                                      -1.985 0.047103 *
return_on_equity                                 -3.710 0.000208 ***
Total_capital_employed                            0.225 0.822113
debt_ratio                                       -4.561 5.10e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6149.2  on 4494  degrees of freedom
Residual deviance: 2521.6  on 4457  degrees of freedom
AIC: 2597.6

Number of Fisher Scoring iterations: 14

>
> # Building a 2nd iteration of the model comprising only of the
> # significant variables
>
> names(Default.smote)
 [1] "Net worth"
 [2] "Total income"
 [3] "Change in stock"
 [4] "Total expenses"
 [5] "PBDITA"
 [6] "Cash profit"
 [7] "PBDITA as % of total income"
 [8] "PBT as % of total income"
 [9] "PAT as % of total income"
[10] "Cash profit as % of total income"
[11] "PAT as % of net worth"
[12] "Sales"
[13] "Income from financial services"
[14] "Other income"
[15] "Borrowings"
[16] "Current liabilities & provisions"
[17] "Shareholders funds"
```

```
[18] "Cumulative retained profits"
[19] "TOL/TNW"
[20] "Total term liabilities / tangible net worth"
[21] "Contingent liabilities / Net worth (%)"
[22] "Investments"
[23] "Current assets"
[24] "Net working capital"
[25] "Debt to equity ratio (times)"
[26] "Cash to average cost of sales per day"
[27] "Creditors turnover"
[28] "Debtors turnover"
[29] "Finished goods turnover"
[30] "WIP turnover"
[31] "Raw material turnover"
[32] "EPS"
[33] "Adjusted EPS"
[34] "PE on BSE"
[35] "return_on_equity"
[36] "Total_capital_employed"
[37] "debt_ratio"
[38] "Default"
>
> # Dropping the non significant variables such as
> # Net Worth, total expenses, PBDITA,
> # PAT as % of total income,
> # PBT as % of total income, Income from financial services,
> # Other income, Borrowings, Current liabilities & provisions,
> # Shareholders funds, Contingent liabilities / Net worth (%),
> # Investments, Current assets, Net working capital,
> # Debtors turnover, Finished goods turnover, WIP turnover,
> # EPS, Adjusted EPS, Total_capital_employed
>
> Default.smote2 <- Default.smote[,-c(1,4,5,8,9,13:17,21:24,28:30,32,33,35
:37)]
>
> # Implementing logistic regression to develop a model
>
> Default.smote2$Default = as.factor(Default.smote2$Default)
> default_logistic2 <- glm(Default~., data=Default.smote2,
+                          family=binomial)
```

```
> plot(as.factor(default_logistic2$y), default_logistic2$fitted.values)

> # As shown in the model the boxplot has a very high distinctive and
> # predictive power as the boxplots differ in a larger manner
```




```
> summary(default_logistic2)

Call:
glm(formula = Default ~ ., family = binomial, data = Default.smote2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.1575  -0.4579  -0.0387   0.3189   3.0195

Coefficients:
                                                 Estimate Std. Error
(Intercept)                                    -0.5102773  0.1035609
`Total income`                                  0.0045340  0.0013485
`Change in stock`                               0.0068819  0.0026660
`Cash profit`                                  -0.0132167  0.0020528
`PBDITA as % of total income`                   0.0004856  0.0010509
`Cash profit as % of total income`             -0.0021625  0.0008801
`PAT as % of net worth`                        -0.0422173  0.0028971
Sales                                          -0.0051315  0.0014064
`Cumulative retained profits`                  -0.0032682  0.0004166
`TOL/TNW`                                       0.2739667  0.0336045
`Total term liabilities / tangible net worth` -0.4467180  0.0795611
`Debt to equity ratio (times)`                  0.4301305  0.0426708
`Cash to average cost of sales per day`        -0.0076443  0.0029098
`Creditors turnover`                           -0.0480963  0.0104157
```

```
`Raw material turnover`                          -0.0745107  0.0104733
`PE on BSE`                                      -0.0035884  0.0016986
                                                 z value Pr(>|z|)
(Intercept)                                       -4.927 8.34e-07 ***
`Total income`                                     3.362 0.000773 ***
`Change in stock`                                  2.581 0.009842 **
`Cash profit`                                     -6.438 1.21e-10 ***
`PBDITA as % of total income`                      0.462 0.644022
`Cash profit as % of total income`                -2.457 0.014005 *
`PAT as % of net worth`                          -14.572  < 2e-16 ***
Sales                                             -3.649 0.000264 ***
`Cumulative retained profits`                     -7.846 4.31e-15 ***
`TOL/TNW`                                          8.153 3.56e-16 ***
`Total term liabilities / tangible net worth`     -5.615 1.97e-08 ***
`Debt to equity ratio (times)`                    10.080  < 2e-16 ***
`Cash to average cost of sales per day`           -2.627 0.008613 **
`Creditors turnover`                              -4.618 3.88e-06 ***
`Raw material turnover`                           -7.114 1.12e-12 ***
`PE on BSE`                                       -2.113 0.034636 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6149.2  on 4494  degrees of freedom
Residual deviance: 2686.2  on 4479  degrees of freedom
AIC: 2718.2

Number of Fisher Scoring iterations: 8
```

Analysing coefficient & their signs

The interpretation of estimates above:

- **The total income is positively corelated to the dependent variable and as it increases the chances of the customer defaulting also increases**
- **The change in stock is positively corelated to the dependent variable and as it increases the chances of the customer defaulting also increases**
- **The cash profit is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The PBDITA as % of total income is positively corelated to the dependent variable and as it increases the chances of the customer defaulting also increases**
- **The Cash profit as % of total income is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The PAT as % of net worth is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The sales are negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The Cumulative retained profits is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The TOL/TNW is positively corelated to the dependent variable and as it increases the chances of the customer defaulting also increases**

- **The Total term liabilities / tangible net worth is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The Debt to equity ratio (times)is positively corelated to the dependent variable and as it increases the chances of the customer defaulting also increases**
- **The Cash to average cost of sales per day is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The Creditors turnover is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The Raw material turnover is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**
- **The PE on BSE is negatively corelated to the dependent variable and as it increases the chances of the customer defaulting also decreases**

Importing the validation dataset to conduct stress test:

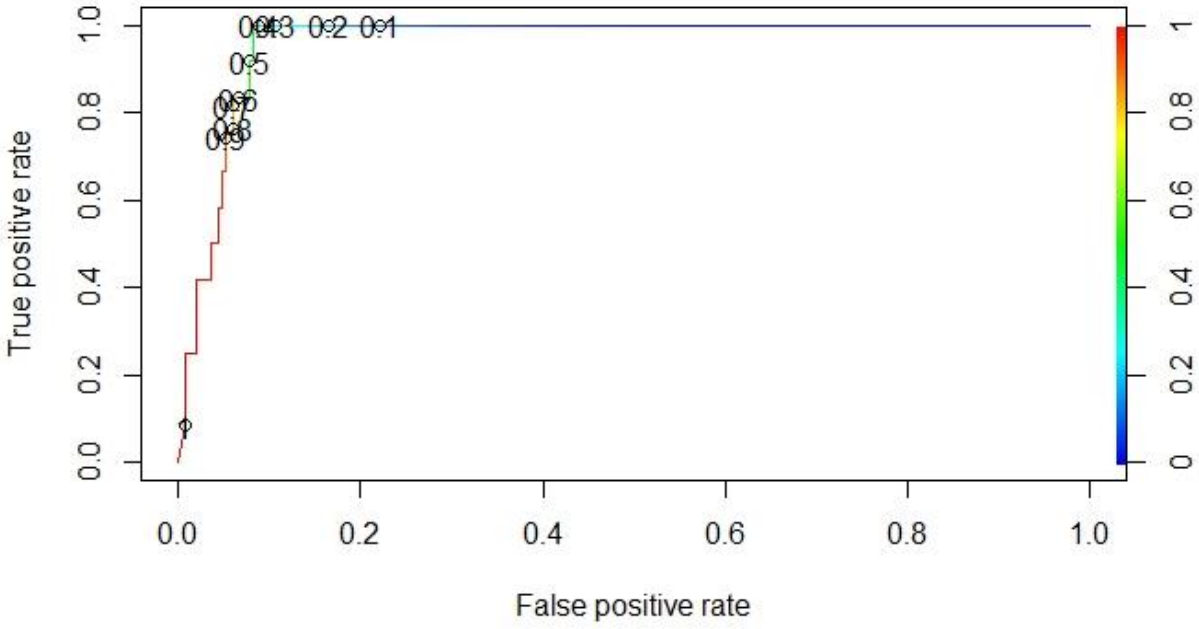
```
> validation_dataset <- read_excel("validation_data.xlsx")
> names(Default.smote2)
 [1] "Total income"
 [2] "Change in stock"
 [3] "Cash profit"
 [4] "PBDITA as % of total income"
 [5] "Cash profit as % of total income"
 [6] "PAT as % of net worth"
 [7] "Sales"
 [8] "Cumulative retained profits"
 [9] "TOL/TNW"
[10] "Total term liabilities / tangible net worth"
[11] "Debt to equity ratio (times)"
[12] "Cash to average cost of sales per day"
[13] "Creditors turnover"
[14] "Raw material turnover"
[15] "PE on BSE"
[16] "Default"
> names(validation_dataset)
 [1] "Num"
 [2] "Default - 1"
 [3] "Total assets"
 [4] "Net worth"
 [5] "Total income"
 [6] "Change in stock"
 [7] "Total expenses"
 [8] "Profit after tax"
 [9] "PBDITA"
[10] "PBT"
[11] "Cash profit"
[12] "PBDITA as % of total income"
[13] "PBT as % of total income"
[14] "PAT as % of total income"
[15] "Cash profit as % of total income"
[16] "PAT as % of net worth"
[17] "Sales"
[18] "Income from financial services"
[19] "Other income"
[20] "Total capital"
[21] "Reserves and funds"
[22] "Deposits (accepted by commercial banks)"
[23] "Borrowings"
```

```
 [24] "Current liabilities & provisions"
 [25] "Deferred tax liability"
 [26] "Shareholders funds"
 [27] "Cumulative retained profits"
 [28] "Capital employed"
 [29] "TOL/TNW"
 [30] "Total term liabilities / tangible net worth"
 [31] "Contingent liabilities / Net worth (%)"
 [32] "Contingent liabilities"
 [33] "Net fixed assets"
 [34] "Investments"
 [35] "Current assets"
 [36] "Net working capital"
 [37] "Quick ratio (times)"
 [38] "Current ratio (times)"
 [39] "Debt to equity ratio (times)"
 [40] "Cash to current liabilities (times)"
 [41] "Cash to average cost of sales per day"
 [42] "Creditors turnover"
 [43] "Debtors turnover"
 [44] "Finished goods turnover"
 [45] "WIP turnover"
 [46] "Raw material turnover"
 [47] "Shares outstanding"
 [48] "Equity face value"
 [49] "EPS"
 [50] "Adjusted EPS"
 [51] "Total liabilities"
 [52] "PE on BSE"
```

```
> validation_dataset$`Creditors turnover` = as.numeric(validation_dataset$
`Creditors turnover`)
> validation_dataset$`Raw material turnover` = as.numeric(validation_datas
et$`Raw material turnover`)
> validation_dataset$`PE on BSE` = as.numeric(validation_dataset$`PE on BS
E`)

> res <- predict(default_logistic2,validation_dataset,
+                type = "response")
>
> table(validation_dataset$`Default - 1`,res>0.5)

    FALSE TRUE
  0   226   19
  1     1   11
>
> # Accuracy of the model
> (229+11)/(229+11+16+1)
[1] 0.9338521
>
> # The accuracy is 0.92 as in 93.3%
```

Using the ROC curve to improve the accuracy of the model

```
 > library(ROCR)
 >
 > ROCRPred <- prediction(res,validation_dataset$`Default - 1`)
 > ROCRPref <- performance(ROCRPred,"tpr","fpr")
 > plot(ROCRPref, colorize = TRUE, print.cutoffs.at=seq(0.1,by=0.1))

 >
```




```
> table(validation_dataset$`Default - 1`,res>0.3)

    FALSE TRUE
  0   220   25
  1     0   12
> # With the implementation of ROC curve
> # Accuracy of the model is 8.8%
> (219+12)/(219+12+26+0)
[1] 0.8988327
>
> # Sensitiviy of the model is 89.3%
> 219/(219+26)
[1] 0.8938776
>
> # Specificty of the model is 100%
> 12/(12+0)
[1] 1
>
> # As it is bank default we have alloted high preference to increase
> # specificity as the defaults should be reduced
```