INDIAN CREDIT RISK ANALYSIS

BUSINESS REPORT

MS VIJAYALAKSHMI G
CA-INTER, B.COM

GREAT LAKES INSTITUTE OF MANAGEMENT
PGP-BABI

PROBLEM STATEMENT

Data description - Please watch the video- Default Risk Estimation-Part-1. 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)

Dependant 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.

Indian Credit Risk Analysis

Vijayalakshmi

3/21/2020

```r
library(glmnet)

library(bnstruct)

library(mice)

library(car)

library(tidyverse)

library(corrplot)

library(MASS)

library(psych)

library(GPArotation)

library(car)

library(psy)

library(caret)


raw_data =read.csv('raw-data.csv')

dim(raw_data)
## [1] 3541    52

names(raw_data)

##  [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"
```

```
summary(raw_data)
```

```
##       Num         Networth.Next.Year  Total.assets         Net.worth
## Min.   :   1   Min.   :-74265.6   Min.   :      0.1   Min.   :     0.0
## 1st Qu.: 886   1st Qu.:    31.7   1st Qu.:     91.3   1st Qu.:    31.3
## Median :1773   Median :   116.3   Median :    309.7   Median :   102.3
## Mean   :1772   Mean   :  1616.3   Mean   :   3443.4   Mean   :  1295.9
## 3rd Qu.:2658   3rd Qu.:   456.1   3rd Qu.:   1098.7   3rd Qu.:   377.3
## Max.   :3545   Max.   :805773.4   Max.   :1176509.2   Max.   :613151.6
##
##   Total.income       Change.in.stock     Total.expenses
## Min.   :      0.0   Min.   :-3029.40   Min.   :     -0.1
## 1st Qu.:    106.5   1st Qu.:   -1.80   1st Qu.:     95.8
## Median :    444.9   Median :    1.60   Median :    407.7
## Mean   :   4582.8   Mean   :   41.49   Mean   :   4262.9
## 3rd Qu.:   1440.9   3rd Qu.:   18.05   3rd Qu.:   1359.8
```

```
##    Max.   :2442828.2    Max.   :14185.50    Max.   :2366035.3
##    NA's   :198          NA's   :458         NA's   :139
##    Profit.after.tax        PBDITA              PBT
##    Min.   : -3908.30   Min.   :  -440.7    Min.   : -3894.80
##    1st Qu.:     0.50   1st Qu.:     6.9    1st Qu.:     0.70
##    Median :     8.80   Median :    35.4    Median :    12.40
##    Mean   :   277.36   Mean   :   578.1    Mean   :   383.81
##    3rd Qu.:    52.27   3rd Qu.:   150.2    3rd Qu.:    71.97
##    Max.   :119439.10   Max.   :208576.5    Max.   :145292.60
##    NA's   :131         NA's   :131         NA's   :131
##     Cash.profit        PBDITA.as...of.total.income PBT.as...of.total.income
##    Min.   : -2245.70   Min.   :-6400.000           Min.   :-21340.00
##    1st Qu.:     2.90   1st Qu.:    5.000           1st Qu.:     0.55
##    Median :    18.85   Median :    9.660           Median :     3.31
##    Mean   :   392.07   Mean   :    4.571           Mean   :   -17.28
##    3rd Qu.:    93.20   3rd Qu.:   16.390           3rd Qu.:     8.80
##    Max.   :176911.80   Max.   :  100.000           Max.   :   100.00
##    NA's   :131         NA's   :68                  NA's   :68
##    PAT.as...of.total.income Cash.profit.as...of.total.income
##    Min.   :-21340.00        Min.   :-15020.000
##    1st Qu.:     0.35        1st Qu.:     2.020
##    Median :     2.34        Median :     5.640
##    Mean   :   -19.20        Mean   :    -8.229
##    3rd Qu.:     6.34        3rd Qu.:    10.700
##    Max.   :   150.00        Max.   :   100.000
##    NA's   :68               NA's   :68
##    PAT.as...of.net.worth     Sales            Income.from.financial.services
##    Min.   :-748.72       Min.   :      0.1    Min.   :    0.00
##    1st Qu.:   0.00       1st Qu.:    112.7    1st Qu.:    0.40
##    Median :   7.92       Median :    453.1    Median :    1.80
##    Mean   :  10.27       Mean   :   4549.5    Mean   :   80.84
##    3rd Qu.:  20.19       3rd Qu.:   1433.5    3rd Qu.:    9.68
##    Max.   :2466.67       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
```

```
head(raw_data)
```

```
##   Num Networth.Next.Year Total.assets Net.worth Total.income
## 1   1            8890.6      17512.3    7093.2      24965.2
## 2   2             394.3        941.0     351.5       1527.4
## 3   3              92.2        232.8     100.6        477.3
## 4   4               2.7          2.7       2.7           NA
## 5   5             109.0        478.5     107.6       1580.5
## 6   6             688.6       2434.4     675.8       2648.6
##   Change.in.stock Total.expenses Profit.after.tax PBDITA    PBT
## 1           235.8        23657.8           1543.2 2860.2 2417.2
## 2            42.7         1454.9            115.2  283.0  188.4
## 3            -5.2          478.7             -6.6    5.8   -6.6
## 4              NA             NA               NA     NA     NA
## 5           -17.0         1558.0              5.5   31.0    6.3
## 6            62.3         2636.4             74.5  200.1   74.5
##   Cash.profit PBDITA.as...of.total.income PBT.as...of.total.income
## 1      1872.8                       11.46                     9.68
```

```
## 2           158.6                        18.53                   12.33
## 3             0.3                         1.22                   -1.38
## 4              NA                         0.00                    0.00
## 5            11.9                         1.96                    0.40
## 6           146.9                         7.55                    2.81
##    PAT.as...of.total.income Cash.profit.as...of.total.income
## 1                      6.18                             7.50
## 2                      7.54                            10.38
## 3                     -1.38                             0.06
## 4                      0.00                             0.00
## 5                      0.35                             0.75
## 6                      2.81                             5.55
##    PAT.as...of.net.worth    Sales Income.from.financial.services
## 1                  23.78 24458.0                            158.0
## 2                  38.08  1504.3                              4.0
## 3                  -6.35   475.6                              1.5
## 4                   0.00      NA                               NA
## 5                   5.25  1575.1                              3.9
## 6                  21.78  2639.5                              6.4
##    Other.income Total.capital Reserves.and.funds
## 1         297.2         423.8             6822.8
## 2          15.9         115.5              257.8
## 3           0.2          81.4               19.2
## 4            NA           0.5                2.2
## 5           0.9           6.2              161.8
## 6           0.2          33.8              972.0
##    Deposits..accepted.by.commercial.banks. Borrowings
## 1                                       NA       14.9
## 2                                       NA      272.5
## 3                                       NA       35.4
## 4                                       NA         NA
## 5                                       NA      193.1
## 6                                       NA      717.1
##    Current.liabilities...provisions Deferred.tax.liability
## 1                            9965.9                  284.9
## 2                             210.0                   85.2
## 3                              96.8                     NA
## 4                                NA                     NA
## 5                             112.8                    4.6
## 6                             555.9                   54.4
##    Shareholders.funds Cumulative.retained.profits Capital.employed TOL.TNW
## 1              7093.2                       6263.3           7108.1    1.33
## 2               351.5                        247.4            624.0    1.23
## 3               100.6                         32.4            136.0    1.44
## 4                 2.7                          2.2              2.7    0.00
## 5               107.6                         82.7            300.7    2.83
## 6               698.2                        317.7           1415.3    1.80
##    Total.term.liabilities...tangible.net.worth
## 1                                         0.00
## 2                                         0.34
## 3                                         0.29
## 4                                         0.00
## 5                                         1.59
## 6                                         0.37
##    Contingent.liabilities...Net.worth.... Contingent.liabilities
```

```
## 1                                       14.80                      1049.7
## 2                                       19.23                        67.6
## 3                                       45.83                        46.1
## 4                                        0.00                          NA
## 5                                       34.94                        37.6
## 6                                       36.28                       245.2
##   Net.fixed.assets Investments Current.assets Net.working.capital
## 1           1900.2      1069.6        13277.5              3588.5
## 2            286.4         2.2          563.9               203.5
## 3             38.7         4.3          167.5                59.6
## 4              2.5          NA            0.2                 0.2
## 5             94.8         7.4          349.7               215.8
## 6            864.9        22.7         1296.2               278.5
##   Quick.ratio..times. Current.ratio..times. Debt.to.equity.ratio..times.
## 1                1.18                  1.37                          0.00
## 2                0.95                  1.56                          0.78
## 3                1.11                  1.55                          0.35
## 4                  NA                    NA                          0.00
## 5                1.41                  2.54                          1.79
## 6                0.48                  1.27                          1.09
##   Cash.to.current.liabilities..times.
## 1                                0.43
## 2                                0.06
## 3                                0.21
## 4                                  NA
## 5                                0.00
## 6                                0.11
##   Cash.to.average.cost.of.sales.per.day Creditors.turnover
## 1                                 68.21                3.62
## 2                                  5.96                9.80
## 3                                 17.07                5.28
## 4                                    NA                0.00
## 5                                  0.00               13.00
## 6                                 15.78                6.50
##   Debtors.turnover Finished.goods.turnover WIP.turnover
## 1             3.85                  200.55        21.78
## 2             5.70                   14.21         7.49
## 3             5.07                    9.24         0.23
## 4             0.00                      NA           NA
## 5             9.46                   12.68         7.90
## 6            21.13                   10.14         8.38
##   Raw.material.turnover Shares.outstanding Equity.face.value   EPS
## 1                  7.71           42381675                10 35.52
## 2                 11.46           11550000                10  9.97
## 3                    NA            8149090                10 -0.50
## 4                  0.00              52404                10  0.00
## 5                 17.03             619635                10  7.91
## 6                  4.74            1141718                10 30.57
##   Adjusted.EPS Total.liabilities PE.on.BSE
## 1         7.10           17512.3     27.31
## 2         9.97             941.0      8.17
## 3        -0.50             232.8     -5.76
## 4         0.00               2.7        NA
## 5         7.91             478.5        NA
## 6        15.28            2434.4        NA
```

```r
Default=ifelse(raw_data$Networth.Next.Year>0,0,1)

summary(as.factor(Default))

## 0 1
## 3298 243

### This means that there are 3298 non defaulting Companies whereas 243 defaulting
Companies in the given dataset.


val =read.csv('validation_data.csv')

str(raw_data)

## 'data.frame':    3541 obs. of  52 variables:
##  $ Num                                  : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Networth.Next.Year                   : num  8890.6 394.3 92.2 2.7 109 ...
##  $ Total.assets                         : num  17512.3 941 232.8 2.7 478.5
...
##  $ Net.worth                            : num  7093.2 351.5 100.6 2.7 107.6
...
##  $ Total.income                         : num  24965 1527 477 NA 1580 ...
##  $ Change.in.stock                      : num  235.8 42.7 -5.2 NA -17 ...
##  $ Total.expenses                       : num  23658 1455 479 NA 1558 ...
##  $ Profit.after.tax                     : num  1543.2 115.2 -6.6 NA 5.5 ...
##  $ PBDITA                               : num  2860.2 283 5.8 NA 31 ...
##  $ PBT                                  : num  2417.2 188.4 -6.6 NA 6.3 ...
##  $ Cash.profit                          : num  1872.8 158.6 0.3 NA 11.9 ...
##  $ PBDITA.as...of.total.income          : num  11.46 18.53 1.22 0 1.96 ...
##  $ PBT.as...of.total.income             : num  9.68 12.33 -1.38 0 0.4 ...
##  $ PAT.as...of.total.income             : num  6.18 7.54 -1.38 0 0.35 2.81 0
0.72 8.29 -2.88 ...
##  $ Cash.profit.as...of.total.income     : num  7.5 10.38 0.06 0 0.75 ...
##  $ PAT.as...of.net.worth                : num  23.78 38.08 -6.35 0 5.25 ...
##  $ Sales                                : num  24458 1504 476 NA 1575 ...
##  $ Income.from.financial.services       : num  158 4 1.5 NA 3.9 6.4 NA NA
7.3 NA ...
##  $ Other.income                         : num  297.2 15.9 0.2 NA 0.9 ...
##  $ Total.capital                        : num  423.8 115.5 81.4 0.5 6.2 ...
##  $ Reserves.and.funds                   : num  6822.8 257.8 19.2 2.2 161.8
...
##  $ Deposits..accepted.by.commercial.banks. : logi  NA NA NA NA NA NA ...
##  $ Borrowings                           : num  14.9 272.5 35.4 NA 193.1 ...
##  $ Current.liabilities...provisions     : num  9965.9 210 96.8 NA 112.8 ...
##  $ Deferred.tax.liability               : num  284.9 85.2 NA NA 4.6 ...
##  $ Shareholders.funds                   : num  7093.2 351.5 100.6 2.7 107.6
...
##  $ Cumulative.retained.profits          : num  6263.3 247.4 32.4 2.2 82.7
...
##  $ Capital.employed                     : num  7108.1 624 136 2.7 300.7 ...
##  $ TOL.TNW                              : num  1.33 1.23 1.44 0 2.83 1.8
0.03 5.17 1.05 3.25 ...
##  $ Total.term.liabilities...tangible.net.worth: num  0 0.34 0.29 0 1.59 0.37 0.03
```

```
0.94 0.3 0.54 ...
##  $ Contingent.liabilities...Net.worth....    : num  14.8 19.2 45.8 0 34.9 ...
##  $ Contingent.liabilities                    : num  1049.7 67.6 46.1 NA 37.6 ...
##  $ Net.fixed.assets                          : num  1900.2 286.4 38.7 2.5 94.8
...
##  $ Investments                               : num  1069.6 2.2 4.3 NA 7.4 ...
##  $ Current.assets                            : num  13277.5 563.9 167.5 0.2 349.7
...
##  $ Net.working.capital                       : num  3588.5 203.5 59.6 0.2 215.8
...
##  $ Quick.ratio..times.                       : num  1.18 0.95 1.11 NA 1.41 0.48
NA 0.54 0.59 0.39 ...
##  $ Current.ratio..times.                     : num  1.37 1.56 1.55 NA 2.54 1.27
NA 1.15 1.58 0.5 ...
##  $ Debt.to.equity.ratio..times.              : num  0 0.78 0.35 0 1.79 1.09 0.32
2.31 0.94 3.13 ...
##  $ Cash.to.current.liabilities..times.       : num  0.43 0.06 0.21 NA 0 0.11 NA
0.04 0.19 0 ...
##  $ Cash.to.average.cost.of.sales.per.day     : num  68.21 5.96 17.07 NA 0 ...
##  $ Creditors.turnover                        : num  3.62 9.8 5.28 0 13 ...
##  $ Debtors.turnover                          : num  3.85 5.7 5.07 0 9.46 ...
##  $ Finished.goods.turnover                   : num  200.55 14.21 9.24 NA 12.68
...
##  $ WIP.turnover                              : num  21.78 7.49 0.23 NA 7.9 ...
##  $ Raw.material.turnover                     : num  7.71 11.46 NA 0 17.03 ...
##  $ Shares.outstanding                        : num  42381675 11550000 8149090
52404 619635 ...
##  $ Equity.face.value                         : num  10 10 10 10 10 10 10 NA 10 10
...
##  $ EPS                                       : num  35.52 9.97 -0.5 0 7.91 ...
##  $ Adjusted.EPS                              : num  7.1 9.97 -0.5 0 7.91 ...
##  $ Total.liabilities                         : num  17512.3 941 232.8 2.7 478.5
...
##  $ PE.on.BSE                                 : num  27.31 8.17 -5.76 NA NA ...
```

Univariate Analysis

```r
plot(raw_data$Networth.Next.Year,raw_data$Total.capital, main="Histogram of Networth
Next Year",

xlab="Net worth Next year")
```

Histogram of Networth Next Year
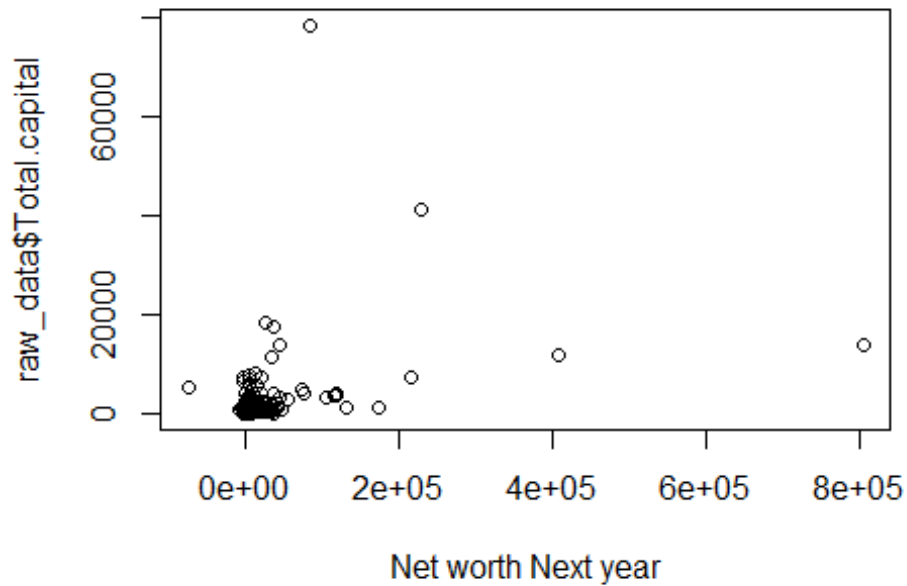

```
raw1 =raw_data[,c(3:52)]

raw1 =cbind(Default,raw1)

raw1[,c(2:51)] =sapply(raw1[,c(2:51)],  function(x)

  (x-mean(x[!is.na(x)]))/sd(x[!is.na(x)]))

colSums(is.na(raw1))
```

```
##                             Default
##                                   0
##                        Total.assets
##                                   0
##                           Net.worth
##                                   0
##                        Total.income
##                                 198
##                     Change.in.stock
##                                 458
##                      Total.expenses
##                                 139
##                    Profit.after.tax
##                                 131
##                              PBDITA
##                                 131
##                                 PBT
##                                 131
##                         Cash.profit
```

```
##                                           131
##                PBDITA.as...of.total.income
##                                            68
##                   PBT.as...of.total.income
##                                            68
##                   PAT.as...of.total.income
##                                            68
##           Cash.profit.as...of.total.income
##                                            68
##                        PAT.as...of.net.worth
##                                             0
##                                        Sales
##                                           259
##               Income.from.financial.services
##                                           935
##                                 Other.income
##                                          1295
##                                Total.capital
##                                             4
##                            Reserves.and.funds
##                                            85
##      Deposits..accepted.by.commercial.banks.
##                                          3541
##                                    Borrowings
##                                           366
##              Current.liabilities...provisions
##                                            96
##                        Deferred.tax.liability
##                                          1140
##                            Shareholders.funds
##                                             0
##                   Cumulative.retained.profits
##                                            38
##                              Capital.employed
##                                             0
##                                       TOL.TNW
##                                             0
## Total.term.liabilities...tangible.net.worth
##                                             0
##         Contingent.liabilities...Net.worth....
##                                             0
##                        Contingent.liabilities
##                                          1188
##                              Net.fixed.assets
##                                           118
##                                   Investments
##                                          1435
##                                Current.assets
##                                            66
##                          Net.working.capital
##                                            32
##                         Quick.ratio..times.
##                                            93
##                       Current.ratio..times.
##                                            93
```

```
##                 Debt.to.equity.ratio..times.
##                                            0
##          Cash.to.current.liabilities..times.
##                                           93
##        Cash.to.average.cost.of.sales.per.day
##                                           85
##                            Creditors.turnover
##                                          333
##                              Debtors.turnover
##                                          328
##                        Finished.goods.turnover
##                                          740
##                                  WIP.turnover
##                                          640
##                          Raw.material.turnover
##                                          361
##                            Shares.outstanding
##                                          692
##                             Equity.face.value
##                                          692
##                                           EPS
##                                            0
##                                  Adjusted.EPS
##                                            0
##                              Total.liabilities
##                                            0
##                                      PE.on.BSE
##                                         2194
```

```r
raw_imp =mice(raw1,m=5,maxit=20,defaultMethod ="pmm",seed=500)
```

```
## Warning: Number of logged events: 3506
```

```r
raw2 =mice::complete(raw_imp,1)

raw2 =raw2[ , -which(names(raw2) %in%c("Deposits..accepted.by.commercial.banks."))]

raw2 =raw2[complete.cases(raw2), ]

write.csv(raw2,"raw.csv", row.names =FALSE)


val[,c(3:52)] =sapply(val[,c(3:52)],  function(x)

  (x-mean(x[!is.na(x)]))/sd(x[!is.na(x)]))

val1 =val[,c(2:52)]

val_imp <-mice(val1,m=5,maxit=20,meth='pmm',seed=500)
```

```
## Warning: Number of logged events: 3506
```

```r
val2 =mice::complete(val_imp,1)

val2 =val2[ , -which(names(val2) %in%c("Deposits..accepted.by.commercial.banks."))]
```

```r
val2 =val2[complete.cases(val2), ]

write.csv(val2,"val_final.csv", row.names =FALSE)

full_model =glm(Default~.,data = raw2, family='binomial')

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

step_model =stepAIC(full_model, direction ='both')

raw3 =raw2[ , which(names(raw2)
%in%c("Default","Total.assets","Net.worth","Total.income","PBDITA","PBT","PAT.as...of
.total.income","PAT.as...of.net.worth","Sales","Income.from.financial.services","Tota
l.capital","Reserves.and.funds","Borrowings","Current.liabilities...provisions","Defe
rred.tax.liability","TOL.TNW","Total.term.liabilities...tangible.net.worth","Net.fixe
d.assets","Investments","Quick.ratio..times.","Current.ratio..times.","Finished.goods
.turnover","WIP.turnover","PE.on.BSE"))]

val3 =val2[ , which(names(val2)
%in%c("Default...1","Total.assets","Net.worth","Total.income","PBDITA","PBT","PAT.as.
..of.total.income","PAT.as...of.net.worth","Sales","Income.from.financial.services","
Total.capital","Reserves.and.funds","Borrowings","Current.liabilities...provisions","
Deferred.tax.liability","TOL.TNW","Total.term.liabilities...tangible.net.worth","Net.
fixed.assets","Investments","Quick.ratio..times.","Current.ratio..times.","Finished.g
oods.turnover","WIP.turnover","PE.on.BSE"))]

fit_model =glm(Default ~Total.assets +Net.worth +Total.income +PBDITA+
PBT +PAT.as...of.total.income +PAT.as...of.net.worth +
Sales +Income.from.financial.services +Total.capital +
Reserves.and.funds +Borrowings +Current.liabilities...provisions +
Deferred.tax.liability +TOL.TNW +Total.term.liabilities...tangible.net.worth +
Net.fixed.assets +Investments +Quick.ratio..times. +Current.ratio..times. +
Finished.goods.turnover +WIP.turnover +PE.on.BSE,data = raw2,family ='binomial')

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

vif(fit_model)

##                           Total.assets
##                             544.405919
##                              Net.worth
##                              20.636014
##                           Total.income
##                             923.586623
##                                 PBDITA
##                              29.146225
##                                    PBT
##                               9.720235
##               PAT.as...of.total.income
##                               1.354144
##                 PAT.as...of.net.worth
##                               1.233509
##                                  Sales
##                             798.882455
```

```
##             Income.from.financial.services
##                            4.549312
##                        Total.capital
##                           23.361141
##                    Reserves.and.funds
##                           12.862346
##                            Borrowings
##                          235.267617
##          Current.liabilities...provisions
##                           38.111351
##                 Deferred.tax.liability
##                            2.081500
##                              TOL.TNW
##                           14.504278
## Total.term.liabilities...tangible.net.worth
##                           14.548891
##                     Net.fixed.assets
##                            8.767035
##                          Investments
##                            1.627623
##                   Quick.ratio..times.
##                            5.308946
##                 Current.ratio..times.
##                            5.645261
##              Finished.goods.turnover
##                            1.093182
##                         WIP.turnover
##                            1.125075
##                            PE.on.BSE
##                            1.007328
```

```
fit_model2 =glm(Default ~Net.worth +
PAT.as...of.total.income +PAT.as...of.net.worth +
Sales +Income.from.financial.services +Total.capital +
Borrowings +Current.liabilities...provisions +
Deferred.tax.liability+TOL.TNW +
Investments +Quick.ratio..times. +Current.ratio..times. +
Finished.goods.turnover +WIP.turnover +PE.on.BSE,data = raw2,family ='binomial')
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
vif(fit_model2)
```

```
##                   Net.worth          PAT.as...of.total.income
##                   32.445241                          1.514689
##           PAT.as...of.net.worth                         Sales
##                    1.092941                          5.726098
##    Income.from.financial.services                 Total.capital
##                    5.478106                          3.753918
##                  Borrowings Current.liabilities...provisions
##                   12.834215                          6.746077
##         Deferred.tax.liability                        TOL.TNW
##                    6.370873                          1.066156
##                 Investments           Quick.ratio..times.
##                    6.017383                          4.646934
```

```
##             Current.ratio..times.          Finished.goods.turnover
##                         4.818059                          1.082121
##                     WIP.turnover                          PE.on.BSE
##                         1.133133                          1.008097

raw4 =raw2[ , which(names(raw2)
%in%c("Default","Net.worth","PAT.as...of.total.income","PAT.as...of.net.worth","Sales
","Income.from.financial.services","Total.capital","Borrowings","Current.liabilities.
..provisions","Deferred.tax.liability","TOL.TNW","Net.fixed.assets","Investments","Qu
ick.ratio..times.","Current.ratio..times.","Finished.goods.turnover","WIP.turnover","
PE.on.BSE"))]

val4 =val2[ , which(names(val2)
%in%c("Default...1","Net.worth","PAT.as...of.total.income","PAT.as...of.net.worth","S
ales","Income.from.financial.services","Total.capital","Borrowings","Current.liabilit
ies...provisions","Deferred.tax.liability","TOL.TNW","Net.fixed.assets","Investments"
,"Quick.ratio..times.","Current.ratio..times.","Finished.goods.turnover","WIP.turnove
r","PE.on.BSE"))]

cor(raw4[2:18])

##                                      Net.worth PAT.as...of.total.income
## Net.worth                         1.000000000               0.0077083835
## PAT.as...of.total.income          0.007708384               1.0000000000
## PAT.as...of.net.worth             0.017823649               0.0902314134
## Sales                             0.796423550               0.0056371094
## Income.from.financial.services    0.580129939              -0.0395902875
## Total.capital                     0.392756109               0.0042830397
## Borrowings                        0.887319420               0.0071949997
## Current.liabilities...provisions  0.725949978               0.0060020957
## Deferred.tax.liability            0.952065008               0.0035787896
## TOL.TNW                          -0.014096311              -0.0195192726
## Net.fixed.assets                  0.972868326               0.0060136011
## Investments                       0.880145668               0.0005368384
## Quick.ratio..times.              -0.006168826              -0.0301049746
## Current.ratio..times.            -0.006336952              -0.0243525908
## Finished.goods.turnover          -0.003600006               0.0001702175
## WIP.turnover                     -0.008724276              -0.0256738344
## PE.on.BSE                        -0.003838827               0.0037315711
##                                   PAT.as...of.net.worth          Sales
## Net.worth                                  0.017823649    0.796423550
## PAT.as...of.total.income                   0.090231413    0.005637109
## PAT.as...of.net.worth                      1.000000000    0.013298358
## Sales                                      0.013298358    1.000000000
## Income.from.financial.services             0.008372735    0.832404354
## Total.capital                              0.003689746    0.288987946
## Borrowings                                 0.007663734    0.855990787
## Current.liabilities...provisions           0.020881326    0.827225004
## Deferred.tax.liability                     0.007870282    0.851555060
## TOL.TNW                                   -0.089490986   -0.005992928
## Net.fixed.assets                           0.011193485    0.787835262
## Investments                                0.010182964    0.895556065
## Quick.ratio..times.                       -0.011278507   -0.008042847
## Current.ratio..times.                     -0.011470702   -0.007360653
## Finished.goods.turnover                    0.029188896   -0.007420215
```

```
## WIP.turnover                          -0.012792247 -0.004760999
## PE.on.BSE                              0.010113966 -0.002647748
##                                 Income.from.financial.services
## Net.worth                                          0.580129939
## PAT.as...of.total.income                          -0.039590288
## PAT.as...of.net.worth                              0.008372735
## Sales                                              0.832404354
## Income.from.financial.services                     1.000000000
## Total.capital                                      0.294010868
## Borrowings                                         0.676620040
## Current.liabilities...provisions                   0.797331890
## Deferred.tax.liability                             0.664778258
## TOL.TNW                                           -0.002184978
## Net.fixed.assets                                   0.535537871
## Investments                                        0.737360000
## Quick.ratio..times.                                0.006736201
## Current.ratio..times.                              0.007461886
## Finished.goods.turnover                            0.011323710
## WIP.turnover                                      -0.006199066
## PE.on.BSE                                         -0.003539680
##                                 Total.capital    Borrowings
## Net.worth                          0.392756109  0.887319420
## PAT.as...of.total.income           0.004283040  0.007195000
## PAT.as...of.net.worth              0.003689746  0.007663734
## Sales                              0.288987946  0.855990787
## Income.from.financial.services     0.294010868  0.676620040
## Total.capital                      1.000000000  0.329943252
## Borrowings                         0.329943252  1.000000000
## Current.liabilities...provisions   0.316460268  0.745752999
## Deferred.tax.liability             0.397729188  0.898760246
## TOL.TNW                            0.004182240 -0.001921664
## Net.fixed.assets                   0.328944369  0.891819623
## Investments                        0.218453216  0.895828058
## Quick.ratio..times.               -0.002222723 -0.009447987
## Current.ratio..times.             -0.004004821 -0.009107308
## Finished.goods.turnover           -0.001614468  0.014195664
## WIP.turnover                       0.011535313 -0.003308503
## PE.on.BSE                         -0.003915815 -0.004560432
##                                 Current.liabilities...provisions
## Net.worth                                            0.7259499781
## PAT.as...of.total.income                             0.0060020957
## PAT.as...of.net.worth                                0.0208813265
## Sales                                                0.8272250041
## Income.from.financial.services                       0.7973318899
## Total.capital                                        0.3164602683
## Borrowings                                           0.7457529990
## Current.liabilities...provisions                     1.0000000000
## Deferred.tax.liability                               0.8118119215
## TOL.TNW                                              0.0027693617
## Net.fixed.assets                                     0.7106331258
## Investments                                          0.7501019888
## Quick.ratio..times.                                 -0.0092861883
## Current.ratio..times.                               -0.0095206647
## Finished.goods.turnover                             -0.0073527972
## WIP.turnover                                        -0.0002437027
```

```
## PE.on.BSE                                            -0.0051645747
##                                       Deferred.tax.liability        TOL.TNW
## Net.worth                                        0.952065008 -0.014096311
## PAT.as...of.total.income                         0.003578790 -0.019519273
## PAT.as...of.net.worth                            0.007870282 -0.089490986
## Sales                                            0.851555060 -0.005992928
## Income.from.financial.services                   0.664778258 -0.002184978
## Total.capital                                    0.397729188  0.004182240
## Borrowings                                       0.898760246 -0.001921664
## Current.liabilities...provisions                 0.811811922  0.002769362
## Deferred.tax.liability                           1.000000000  0.014303166
## TOL.TNW                                          0.014303166  1.000000000
## Net.fixed.assets                                 0.955913828 -0.008161322
## Investments                                      0.859730788 -0.008446857
## Quick.ratio..times.                             -0.008650166 -0.019583038
## Current.ratio..times.                           -0.008510131 -0.017973348
## Finished.goods.turnover                         -0.004767182  0.001889868
## WIP.turnover                                    -0.001296868  0.014430287
## PE.on.BSE                                       -0.004141897 -0.003438302
##                                       Net.fixed.assets    Investments
## Net.worth                                  0.972868326  0.8801456683
## PAT.as...of.total.income                   0.006013601  0.0005368384
## PAT.as...of.net.worth                      0.011193485  0.0101829636
## Sales                                      0.787835262  0.8955560647
## Income.from.financial.services             0.535537871  0.7373599997
## Total.capital                              0.328944369  0.2184532155
## Borrowings                                 0.891819623  0.8958280577
## Current.liabilities...provisions           0.710633126  0.7501019888
## Deferred.tax.liability                     0.955913828  0.8597307878
## TOL.TNW                                   -0.008161322 -0.0084468567
## Net.fixed.assets                           1.000000000  0.8364733253
## Investments                                0.836473325  1.0000000000
## Quick.ratio..times.                       -0.008469338 -0.0066170654
## Current.ratio..times.                     -0.008083859 -0.0065688691
## Finished.goods.turnover                   -0.008227970  0.0085984485
## WIP.turnover                              -0.004435649  0.0058152286
## PE.on.BSE                                 -0.003445401 -0.0037955947
##                                       Quick.ratio..times. Current.ratio..times.
## Net.worth                                    -0.006168826          -0.006336952
## PAT.as...of.total.income                     -0.030104975          -0.024352591
## PAT.as...of.net.worth                        -0.011278507          -0.011470702
## Sales                                        -0.008042847          -0.007360653
## Income.from.financial.services                0.006736201           0.007461886
## Total.capital                                -0.002222723          -0.004004821
## Borrowings                                   -0.009447987          -0.009107308
## Current.liabilities...provisions             -0.009286188          -0.009520665
## Deferred.tax.liability                       -0.008650166          -0.008510131
## TOL.TNW                                      -0.019583038          -0.017973348
## Net.fixed.assets                             -0.008469338          -0.008083859
## Investments                                  -0.006617065          -0.006568869
## Quick.ratio..times.                           1.000000000           0.979788668
## Current.ratio..times.                         0.979788668           1.000000000
## Finished.goods.turnover                       0.047735906           0.048591950
## WIP.turnover                                  0.003207906          -0.003101069
## PE.on.BSE                                     0.003015215           0.008062199
```

```
##                                      Finished.goods.turnover  WIP.turnover
## Net.worth                                   -0.0036000058 -0.0087242758
## PAT.as...of.total.income                     0.0001702175 -0.0256738344
## PAT.as...of.net.worth                        0.0291888962 -0.0127922474
## Sales                                       -0.0074202145 -0.0047609992
## Income.from.financial.services              0.0113237096 -0.0061990662
## Total.capital                              -0.0016144676  0.0115353126
## Borrowings                                  0.0141956637 -0.0033085033
## Current.liabilities...provisions           -0.0073527972 -0.0002437027
## Deferred.tax.liability                     -0.0047671821 -0.0012968682
## TOL.TNW                                      0.0018898680  0.0144302867
## Net.fixed.assets                           -0.0082279698 -0.0044356492
## Investments                                 0.0085984485  0.0058152286
## Quick.ratio..times.                         0.0477359062  0.0032079062
## Current.ratio..times.                       0.0485919501 -0.0031010685
## Finished.goods.turnover                     1.0000000000  0.2406780827
## WIP.turnover                                0.2406780827  1.0000000000
## PE.on.BSE                                   -0.0059511197 -0.0070653213
##                                       PE.on.BSE
## Net.worth                            -0.003838827
## PAT.as...of.total.income              0.003731571
## PAT.as...of.net.worth                 0.010113966
## Sales                                -0.002647748
## Income.from.financial.services       -0.003539680
## Total.capital                        -0.003915815
## Borrowings                           -0.004560432
## Current.liabilities...provisions     -0.005164575
## Deferred.tax.liability               -0.004141897
## TOL.TNW                              -0.003438302
## Net.fixed.assets                     -0.003445401
## Investments                          -0.003795595
## Quick.ratio..times.                   0.003015215
## Current.ratio..times.                 0.008062199
## Finished.goods.turnover              -0.005951120
## WIP.turnover                         -0.007065321
## PE.on.BSE                             1.000000000
```

```r
fac_raw4 <-factanal(raw4[2:18],
factors =4,                # number of factors to extract
rotation="promax",lower =0.1)     # 'varimax' is an ortho rotation
```

```r
fac_raw4
## Call:
## factanal(x = raw4[2:18], factors = 4, rotation = "promax", lower = 0.1)
##
## Uniquenesses:
##                      Net.worth          PAT.as...of.total.income
##                          0.100                             0.997
##            PAT.as...of.net.worth                           Sales
##                          1.000                             0.100
##    Income.from.financial.services                   Total.capital
##                          0.100                             0.719
##                      Borrowings Current.liabilities...provisions
##                          0.107                             0.180
##          Deferred.tax.liability                         TOL.TNW
```

```
##                          0.100                              0.999
##              Net.fixed.assets                         Investments
##                          0.100                              0.100
##          Quick.ratio..times.               Current.ratio..times.
##                          0.100                              0.100
##       Finished.goods.turnover                        WIP.turnover
##                          0.996                              1.000
##                        PE.on.BSE
##                          1.000
##
## Loadings:
##                                Factor1 Factor2 Factor3 Factor4
## Net.worth                       0.988          -0.164
## PAT.as...of.total.income
## PAT.as...of.net.worth
## Sales                           0.774           0.354
## Income.from.financial.services  0.352           0.719
## Total.capital                                           0.484
## Borrowings                      0.974
## Current.liabilities...provisions 0.469          0.414   0.268
## Deferred.tax.liability          0.879                   0.168
## TOL.TNW
## Net.fixed.assets                0.994          -0.215
## Investments                     0.990           0.137  -0.218
## Quick.ratio..times.                     0.969
## Current.ratio..times.                   0.969
## Finished.goods.turnover
## WIP.turnover
## PE.on.BSE
##
##                Factor1 Factor2 Factor3 Factor4
## SS loadings      5.614   1.884   0.911   0.413
## Proportion Var   0.330   0.111   0.054   0.024
## Cumulative Var   0.330   0.441   0.495   0.519
##
## Factor Correlations:
##          Factor1  Factor2  Factor3 Factor4
## Factor1  1.0000 -0.01071 -0.44920  0.5551
## Factor2 -0.0107  1.00000 -0.00325 -0.0162
## Factor3 -0.4492 -0.00325  1.00000 -0.1914
## Factor4  0.5551 -0.01618 -0.19136  1.0000
##
## Test of the hypothesis that 4 factors are sufficient.
## The chi square statistic is 9234.7 on 74 degrees of freedom.
## The p-value is 0

scree.plot(fac_raw4$correlation)
```

Scree Plot
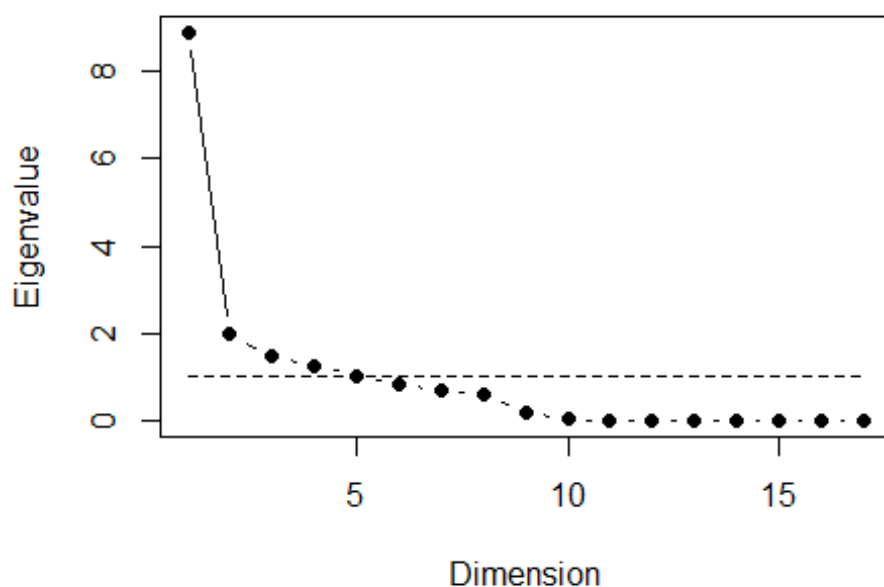

```r
cor(raw4$Net.worth,raw4$Sales)

## [1] 0.7964235

f1_r
=raw4$Net.worth+raw4$Sales+raw4$Borrowings+raw4$Deferred.tax.liability+raw4$Net.fixed
.assets+raw4$Investments
f2_r =raw4$Quick.ratio..times.+raw4$Current.ratio..times.
f3_r =raw4$Income.from.financial.services
f4_r =raw4$Total.capital

raw_fac =data.frame("Default"=raw4$Default,"Company Size"
=f1_r,"Liquidity"=f4_r,"Profitability" =f3_r,"Leverage" =f2_r)

f1_v
=val4$Net.worth+val4$Sales+val4$Borrowings+val4$Deferred.tax.liability+val4$Net.fixed
.assets+val4$Investments
f2_v =val4$Quick.ratio..times.+val4$Current.ratio..times.
f3_v =val4$Income.from.financial.services
f4_v =val4$Total.capital

val_fac =data.frame("Default"=val4$Default...1,"Company Size"
=f1_v,"Liquidity"=f4_v,"Profitability" =f3_v,"Leverage" =f2_v)

log_fit_raw_fac =glm (Default~.,data = raw_fac,family ='binomial')

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

summary(log_fit_raw_fac)
```

```
##
## Call:
## glm(formula = Default ~ ., family = "binomial", data = raw_fac)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8711  -0.3641  -0.3469  -0.3147   4.4825
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -3.07696    0.12082 -25.468   <2e-16 ***
## Company.Size  -0.11045    0.05779  -1.911   0.0560 .
## Liquidity      0.04406    0.04754   0.927   0.3540
## Profitability  0.39527    0.18980   2.083   0.0373 *
## Leverage      -1.65478    0.53802  -3.076   0.0021 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1423.4  on 3281  degrees of freedom
## Residual deviance: 1402.5  on 3277  degrees of freedom
## AIC: 1412.5
##
## Number of Fisher Scoring iterations: 8

soft_pred =predict(log_fit_raw_fac,newdata = val_fac[2:5],type ='response')
hard_pred =as.factor(ifelse(soft_pred>0.075,1,0)) # Tried different values of p* to
get the value that gives best accuracy.
confusionMatrix(hard_pred,as.factor(val_fac$Default))

## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##          0 627  39
##          1   3   0
##
##                Accuracy : 0.9372
##                  95% CI : (0.9161, 0.9544)
##     No Information Rate : 0.9417
##     P-Value [Acc > NIR] : 0.7236
##
##                   Kappa : -0.0084
##
##  Mcnemar's Test P-Value : 6.641e-08
##
##             Sensitivity : 0.9952
##             Specificity : 0.0000
##          Pos Pred Value : 0.9414
##          Neg Pred Value : 0.0000
##              Prevalence : 0.9417
##          Detection Rate : 0.9372
##    Detection Prevalence : 0.9955
##       Balanced Accuracy : 0.4976
```

```
## 
##          'Positive' Class : 0
## 

val_pred =cbind(soft_pred,val_fac)

answer <-function(test){
  test2 =test[order(-test$soft_pred),]
  val
=c(sum(test2[1:67,"Default"]),sum(test2[68:134,"Default"]),sum(test2[135:201,"Default
"]),
sum(test2[202:268,"Default"]),sum(test2[269:335,"Default"]),
sum(test2[336:402,"Default"]),sum(test2[403:469,"Default"]),
sum(test2[470:536,"Default"]),sum(test2[537:603,"Default"]),sum(test2[604:669,"Defaul
t"]))
print(val)
}

answer(val_pred)

##  [1] 13  8  3  4  3  3  2  0  3  0
```

The accuracy derived is 93% due to class imbalance. The number above states that 13 out of the above class have been identified correctly.

OTHER OBSERVATIONS RELATED TO GIVEN DATASET

Based on Dr. Sarkar's Default risk estimation, following exercise was carried out for further understanding of dataset:

Default rate
243/(3298+243)
Default rate is 0.06862468
summary(Default)

```
summary(Default)
    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.00000 0.00000 0.00000 0.06862 0.00000 1.00000
```

plot(`Total income`, `Networth Next Year`)
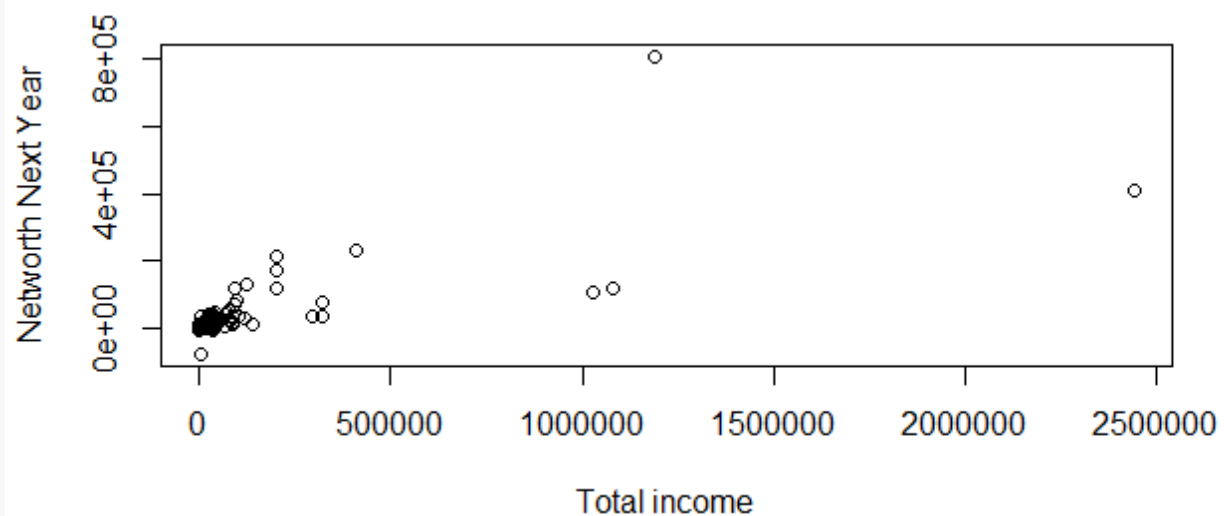
plot(`PBT as % of total income`, `Networth Next Year`)
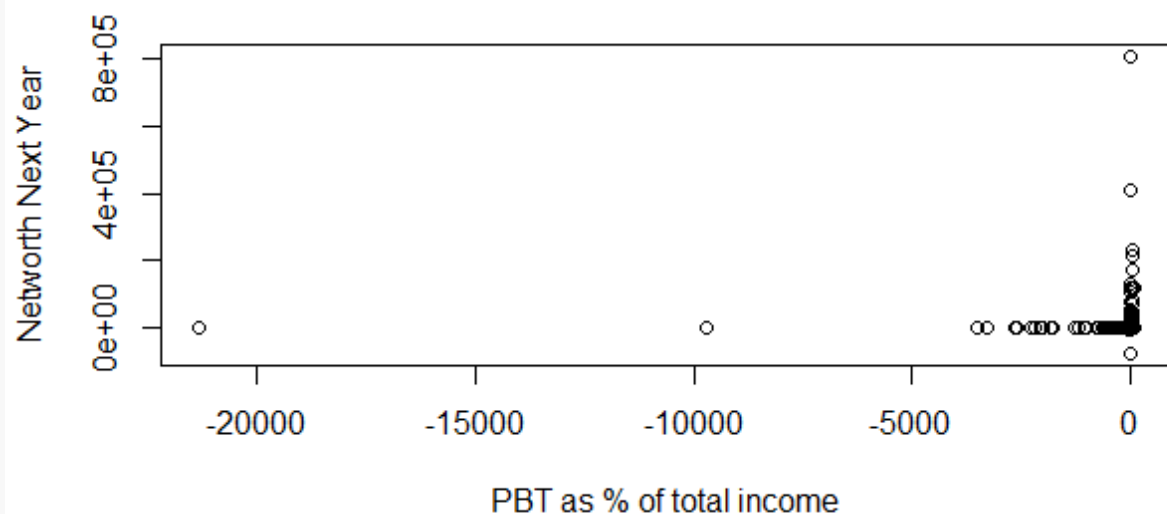
```
summary(`PBT as % of total income`)
```

```
 Min.   1st Qu.   Median      Mean  3rd Qu.      Max.      NA's
-21340.00     0.55      3.31    -17.28      8.80    100.00        68
```

```
summary(`PBT as % of total income`[Default==0])
```

```
  Min.   1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
-9700.00      0.93      3.69    -3.97      9.29   100.00       52
```

```
summary(`PBT as % of total income`[Default==1])
```

Logistic regression
Default Model 1 (PROFITABILITY)
```
Default.model1 = glm(as.factor(Default)~`PBT as % of total income`,
          family=binomial)
summary(Default.model1)
```

```
Call:
glm(formula = as.factor(Default) ~ `PBT as % of total income`,
    family = binomial)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-4.1063  -0.3611  -0.3605  -0.3591   2.3921

Coefficients:
                            Estimate Std. Error z value Pr(>|z|)
(Intercept)                 -2.696222   0.069883 -38.582  < 2e-16 ***
`PBT as % of total income`  -0.001147   0.000248  -4.626 3.72e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1677.3  on 3472  degrees of freedom
Residual deviance: 1647.0  on 3471  degrees of freedom
  (68 observations deleted due to missingness)
AIC: 1651
```

```
Number of Fisher Scoring iterations: 5
```

The analysis Of the summary clearly shows that the variable PBT as % Total
Income is significant in impacting the dependant variable Net worth Next
Year.
Hence the equation derived through Logistic regression for the variable PBT
as % of Total Income is significant:
-2.696222 + -0.001147(PBT % of Total Income)

summary(Default.model1$fitted.values)

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.05674 0.06260 0.06297 0.06536 0.06316 1.00000
```

Fitted values refer to the probabilities of a company to do the default.
The summary says that the least risky company has a predicted default rate
of 0.05674, a typical company has a median default rate of 0.06297, and
the worst of the companies show a very high predicted default rate of 1.

plot(as.factor(Default.model1$y), Default.model1$fitted.values)
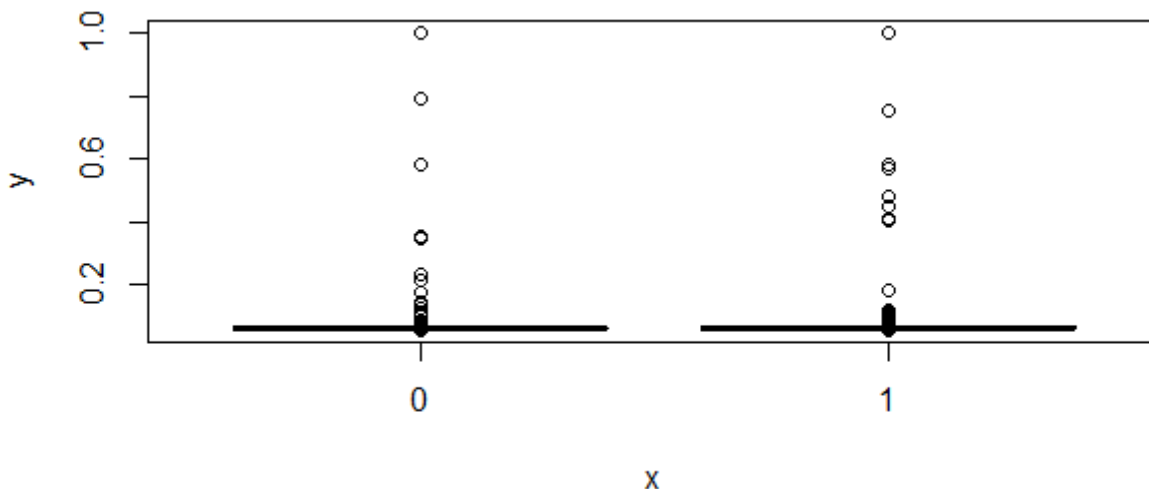
The plot reads that there is not a high discrimination between our actual
probabilities and the fitted values. WWe need to build a richer model to
understand the variables showing a reasonable discrimination between the
actual probabilities and the fitted values being predicted probabilities.
Let us consider more variables and check the impact on the model.

Default Model 2 (PROFITABILITY + LEVERAGE)
Default.model2=glm(as.factor(Default)~`PBT as % of total income`+
          `Debt to equity ratio (times)`, family=binomial)
summary(Default.model2)

```
Call:
glm(formula = as.factor(Default) ~ `PBT as % of total income` +
    `Debt to equity ratio (times)`, family = binomial)
```

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-6.1326  -0.3363  -0.3268  -0.3210   2.4836

Coefficients:
                                Estimate Std. Error z value Pr(>|z|)
(Intercept)                   -2.9417330  0.0784617 -37.493  < 2e-16 ***
`PBT as % of total income`    -0.0010333  0.0002601  -3.973 7.10e-05 ***
`Debt to equity ratio (times)`  0.0637388  0.0083463   7.637 2.23e-14 ***
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1677.3  on 3472  degrees of freedom
Residual deviance: 1504.9  on 3470  degrees of freedom
  (68 observations deleted due to missingness)
AIC: 1510.9

Number of Fisher Scoring iterations: 6
```

Debt to equity Ratio as a variable is 7.637 standard deviations away from
0. Hence, this is even more significant than the PBT ratio as a % of Total
Income.
Equation for this model is as follows:
-2.9417330 - (0.0010333 * PBT as % of Total Income) + (0.0637388 * Debt to
Equity Ratio)
summary(Default.model2$fitted.values)

```
Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.04543 0.05061 0.05251 0.06536 0.05581 1.00000
```

plot(as.factor(Default.model2$y), Default.model2$fitted.values
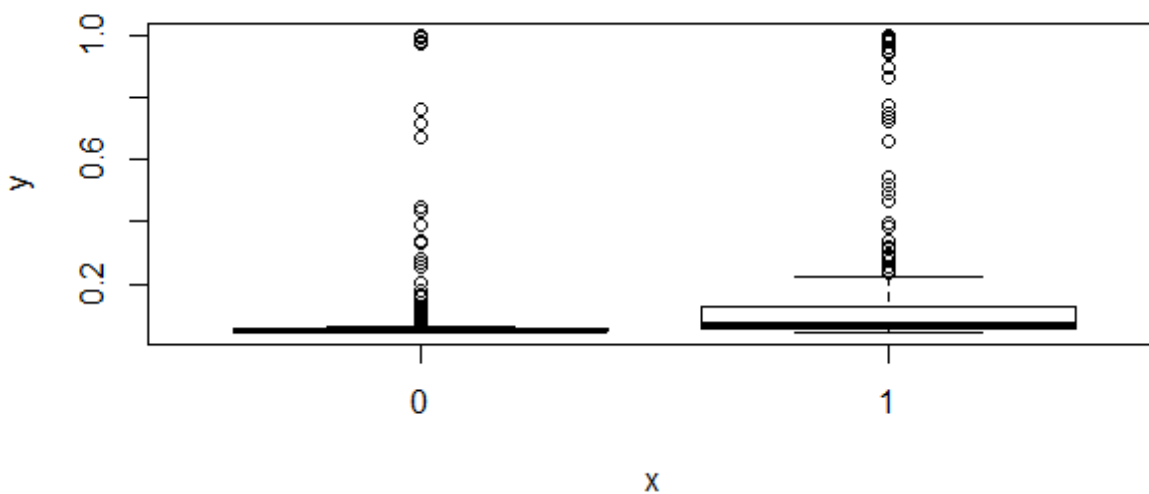
It is seen that the 0s and 1s are separated more here. The upper Quartile
of the predicted probabilities for the defaulting companies is more than

```
Default.prediction2=ifelse(Default.model2$fitted.values>0.0686,1,0)
table(Default.model2$y, Default.prediction2)
Default.prediction2
        0    1
  0  3113  133
  1   107  120
```

```
Default.prediction2=ifelse(Default.model2$fitted.values>0.06,1,0)
table(Default.model2$y, Default.prediction2)
Default.prediction2
        0    1
  0  2929  317
  1    81  146
```

```
146/(81+146)
[1] 0.6431718
```

```
2929/(2929+317)
0.9023413
```

```
(2929+146)/(2929+317+81+146)
0.8854017
```

```
1-(2929+146)/(2929+317+81+146)
```

```
0.1145983
```

The rate of mis-classification is 11.45%. The mis-classification can be
classified into following 2 categories:
1. Defaulters misclassified as Non-defaulters (Given loan might
not come back)
2. Non-defaulters misclassified as Defaulters (Loan is not
extended to such customers based on their mis-classified
defaulting nature, hence, we lose out on prospective Income
earning genres that were open to us - thereby placing hot
cake like opportunities onto the lap of Competitors.)

Choosing a good threshold depends on numerous factors other than the
statistical knowhow and technique usage, such as, Business type, nature,
statutory, legal, economical, political, social and geographical acumen.
After taking into consideration all the pros and cons and regulatory
compliances, then and only then, can a good threshold be defined and
implemented through evaluations and ellaboratively analysed and tested
processes.

For theoretical purposes, all things remaining constant, let us tighten
the default rate as our base parameter:
Default.prediction2=ifelse(Default.model2$fitted.values>0.05,1,0)
table(Default.model2$y, Default.prediction2)
```
   Default.prediction2
        0    1
  0   436 2810
  1    10  217
```

217 of companies have been correctly classified, however, essential
numbers of good companies have been classified as defaulters.
This ruins our Customer base.

Default.prediction2=ifelse(Default.model2$fitted.values>0.058,1,0)
table(Default.model2$y, Default.prediction2)
```
   Default.prediction2
        0    1
  0  2815  431
  1    69  158
```

Whether the table with default rate of 0.058 is better than the table
for 0.06, is a business decision based on significant factors affecting
such business. However, the universal principle of Investments states
clearly that Reducing risk is directly related to reducing returns or
opportunities.

Default Model 3 (DECISION TREE)

```
rpart(as.factor(Default)~`PBT as % of total income`+
        `Debt to equity ratio (times)`)
n= 3541

node), split, n, loss, yval, (yprob)
      * denotes terminal node

 1) root 3541 243 0 (0.93137532 0.06862468)
   2) Debt to equity ratio (times)< 5.17 3307 127 0 (0.96159661 0.03840339)
     4) PBT as % of total income>=-4.19 3067  69 0 (0.97750245 0.02249755) *
     5) PBT as % of total income< -4.19 240  58 0 (0.75833333 0.24166667)
      10) Debt to equity ratio (times)< 2.455 198  33 0 (0.83333333
0.16666667) *
      11) Debt to equity ratio (times)>=2.455 42  17 1 (0.40476190
0.59523810) *
   3) Debt to equity ratio (times)>=5.17 234 116 0 (0.50427350 0.49572650)
     6) PBT as % of total income>=0.03 87   9 0 (0.89655172 0.10344828) *
     7) PBT as % of total income< 0.03 147  40 1 (0.27210884 0.72789116) *
```
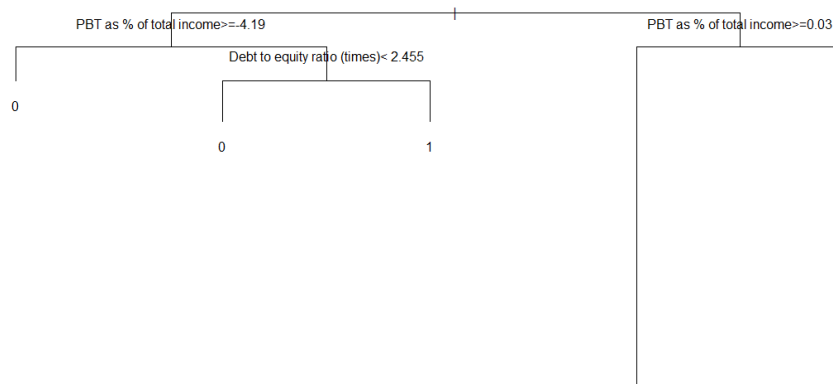
plot(rpart(as.factor(Default)~`PBT as % of total income`+
        `Debt to equity ratio (times)`))
text(rpart(as.factor(Default)~`PBT as % of total income`+
        `Debt to equity ratio (times)`))

Let us understand the Decision Tree. Out of 3541 companies, 243 had default.
The first criteria asked was whether Debt Equity Ratio is more or less than
5.17. The default rate considered is 6.86%. If Debt equity ratio is less
than 5.17,default rate plummets down to 3.84%, whereas, If Debt equity ratio
is more than 5.17, default rate hikes up to 49.57%.
Secondly, if companies are classified based on Debt equity ratio, the next
criterion is to question whether the profitability is low or high.
Such rules can be used as an alternative to logistic regression and
Discriminant analysis.