CBM 122020

Kartheek Raj

1/2/2020

Problem : Find the features impacting target variable i.e GT_Turbine_decay_state_coefficie and the best model yields better root mean square error.

Setting working directory

```
getwd()
```

```
## [1] "C:/Users/mkart/OneDrive/Documents/Projects/ProcessingBigData/CBM/CBM"
```

```
setwd("C:/Users/mkart/OneDrive/Documents/Projects/ProcessingBigData/CBM/CBM")
```

Data pulling from local machine for analysis.

```
data <-read.delim("data.txt", header = FALSE, sep = "", quote = "\"",
           dec = ".", fill = F,stringsAsFactors= T)
cbm<-data.frame(data)
```

Checking first 10 rows

```
head(cbm,10)
```

```
##       V1 V2        V3       V4       V5      V6      V7       V8  V9     V10
## 1  1.138  3   289.964 1349.489 6677.380   7.584   7.584  464.006 288 550.563
## 2  2.088  6  6960.180 1376.166 6828.469  28.204  28.204  635.401 288 581.658
## 3  3.144  9  8379.229 1386.757 7111.811  60.358  60.358  606.002 288 587.587
## 4  4.161 12 14724.395 1547.465 7792.630 113.774 113.774  661.471 288 613.851
## 5  5.140 15 21636.432 1924.313 8494.777 175.306 175.306  731.494 288 645.642
## 6  6.175 18 29792.731 2307.404 8828.360 246.278 246.278  800.434 288 676.397
## 7  7.148 21 38982.180 2678.086 9132.429 332.077 332.077  854.747 288 699.954
## 8  8.206 24 50996.808 3087.561 9318.562 437.989 437.989  952.122 288 741.770
## 9  9.300 27 72763.329 3560.395 9778.528 644.905 644.905 1115.797 288 789.094
## 10 1.138  3   379.880 1355.375 6683.916   7.915   7.915  464.017 288 550.985
##      V11   V12    V13   V14    V15   V16  V17   V18
## 1  1.096 0.998  5.947 1.019  7.137 0.082 0.95 0.975
## 2  1.331 0.998  7.282 1.019 10.655 0.287 0.95 0.975
## 3  1.389 0.998  7.574 1.020 13.086 0.259 0.95 0.975
## 4  1.658 0.998  9.007 1.022 18.109 0.358 0.95 0.975
## 5  2.078 0.998 11.197 1.026 26.373 0.522 0.95 0.975
## 6  2.501 0.998 13.356 1.030 35.760 0.708 0.95 0.975
## 7  2.963 0.998 15.679 1.035 45.881 0.908 0.95 0.975
## 8  3.576 0.998 18.632 1.040 62.440 1.236 0.95 0.975
## 9  4.498 0.998 22.811 1.049 92.556 1.832 0.95 0.975
## 10 1.100 0.998  5.963 1.019  3.879 0.079 0.95 0.976
```

coloumn names not defined as required, assing coloumn names as inteded.

```
colnames(cbm)<-c("Lever position (lp)",
                 "Ship speed (v) [knots]",
                 "Gas Turbine (GT) shaft torque (GTT) [kN m]",
                 "GT rate of revolutions (GTn) [rpm]",
                 "Gas Generator rate of revolutions (GGn) [rpm]",
```

```
                "Starboard Propeller Torque (Ts) [kN]",
                "Port Propeller Torque (Tp) [kN]",
                "Hight Pressure (HP) Turbine exit temperature (T48) [C]",
                "GT Compressor inlet air temperature (T1) [C]",
                "GT Compressor outlet air temperature (T2) [C]",
                "HP Turbine exit pressure (P48) [bar]",
                "GT Compressor inlet air pressure (P1) [bar]",
                "GT Compressor outlet air pressure (P2) [bar]",
                "GT exhaust gas pressure (Pexh) [bar]",
                "Turbine Injecton Control (TIC) [%]",
                "Fuel flow (mf) [kg/s]",
                 "GT Compressor decay state coefficient",
                 "GT_Turbine_decay_state_coefficient")
```

checking last 10 rows

```
tail(cbm,10)
```

```
##       Lever position (lp) Ship speed (v) [knots]
## 11925               9.300                     27
## 11926               1.138                      3
## 11927               2.088                      6
## 11928               3.144                      9
## 11929               4.161                     12
## 11930               5.140                     15
## 11931               6.175                     18
## 11932               7.148                     21
## 11933               8.206                     24
## 11934               9.300                     27
##       Gas Turbine (GT) shaft torque (GTT) [kN m]
## 11925                                   72773.934
## 11926                                    3131.810
## 11927                                    3881.963
## 11928                                    8375.659
## 11929                                   14718.321
## 11930                                   21624.934
## 11931                                   29763.213
## 11932                                   39003.867
## 11933                                   50992.579
## 11934                                   72775.130
##       GT rate of revolutions (GTn) [rpm]
## 11925                           3560.398
## 11926                           1406.601
## 11927                           1345.441
## 11928                           1386.737
## 11929                           1547.453
## 11930                           1924.342
## 11931                           2306.745
## 11932                           2678.052
## 11933                           3087.434
## 11934                           3560.400
##       Gas Generator rate of revolutions (GGn) [rpm]
## 11925                                      9742.081
## 11926                                      6673.118
```

```
## 11927                                    6761.168
## 11928                                    7063.239
## 11929                                    7729.639
## 11930                                    8470.013
## 11931                                    8800.352
## 11932                                    9120.889
## 11933                                    9300.274
## 11934                                    9742.950
##       Starboard Propeller Torque (Ts) [kN] Port Propeller Torque (Tp) [kN]
## 11925                              644.843                           644.843
## 11926                               10.692                            10.692
## 11927                               22.478                            22.478
## 11928                               60.319                            60.319
## 11929                              113.743                           113.743
## 11930                              175.239                           175.239
## 11931                              245.954                           245.954
## 11932                              332.389                           332.389
## 11933                              438.024                           438.024
## 11934                              644.880                           644.880
##       Hight Pressure (HP) Turbine exit temperature (T48) [C]
## 11925                                              1039.245
## 11926                                               510.408
## 11927                                               519.295
## 11928                                               563.854
## 11929                                               617.242
## 11930                                               681.658
## 11931                                               747.405
## 11932                                               796.457
## 11933                                               892.945
## 11934                                              1038.411
##       GT Compressor inlet air temperature (T1) [C]
## 11925                                          288
## 11926                                          288
## 11927                                          288
## 11928                                          288
## 11929                                          288
## 11930                                          288
## 11931                                          288
## 11932                                          288
## 11933                                          288
## 11934                                          288
##       GT Compressor outlet air temperature (T2) [C]
## 11925                                       767.743
## 11926                                       552.868
## 11927                                       555.401
## 11928                                       573.260
## 11929                                       599.109
## 11930                                       628.950
## 11931                                       658.853
## 11932                                       680.393
## 11933                                       722.029
## 11934                                       767.595
##       HP Turbine exit pressure (P48) [bar]
## 11925                                4.532
```

```
## 11926                                1.213
## 11927                                1.231
## 11928                                1.391
## 11929                                1.663
## 11930                                2.087
## 11931                                2.512
## 11932                                2.982
## 11933                                3.594
## 11934                                4.531
##       GT Compressor inlet air pressure (P1) [bar]
## 11925                                      0.998
## 11926                                      0.998
## 11927                                      0.998
## 11928                                      0.998
## 11929                                      0.998
## 11930                                      0.998
## 11931                                      0.998
## 11932                                      0.998
## 11933                                      0.998
## 11934                                      0.998
##       GT Compressor outlet air pressure (P2) [bar]
## 11925                                     22.491
## 11926                                      6.451
## 11927                                      6.566
## 11928                                      7.416
## 11929                                      8.830
## 11930                                     10.990
## 11931                                     13.109
## 11932                                     15.420
## 11933                                     18.293
## 11934                                     22.464
##       GT exhaust gas pressure (Pexh) [bar] Turbine Injecton Control (TIC) [%]
## 11925                                1.052                              86.168
## 11926                                1.019                               0.088
## 11927                                1.020                              20.720
## 11928                                1.021                              11.419
## 11929                                1.023                              16.151
## 11930                                1.027                              23.803
## 11931                                1.031                              32.671
## 11932                                1.036                              42.104
## 11933                                1.043                              58.064
## 11934                                1.052                              86.067
##       Fuel flow (mf) [kg/s] GT Compressor decay state coefficient
## 11925                 1.706                                    1
## 11926                 0.146                                    1
## 11927                 0.172                                    1
## 11928                 0.226                                    1
## 11929                 0.320                                    1
## 11930                 0.471                                    1
## 11931                 0.647                                    1
## 11932                 0.834                                    1
## 11933                 1.149                                    1
## 11934                 1.704                                    1
##       GT_Turbine_decay_state_coefficient
```

```
## 11925                              0.999
## 11926                              1.000
## 11927                              1.000
## 11928                              1.000
## 11929                              1.000
## 11930                              1.000
## 11931                              1.000
## 11932                              1.000
## 11933                              1.000
## 11934                              1.000
```

checking the data types in cbm data.we found all features numeric data type.

```
str(cbm)
```

```
## 'data.frame':    11934 obs. of  18 variables:
##  $ Lever position (lp)                                  : num  1.14 2.09 3.14 4.16 5.14 ...
##  $ Ship speed (v) [knots]                               : num  3 6 9 12 15 18 21 24 27 3 ...
##  $ Gas Turbine (GT) shaft torque (GTT) [kN m]           : num  290 6960 8379 14724 21636 ...
##  $ GT rate of revolutions (GTn) [rpm]                   : num  1349 1376 1387 1547 1924 ...
##  $ Gas Generator rate of revolutions (GGn) [rpm]        : num  6677 6828 7112 7793 8495 ...
##  $ Starboard Propeller Torque (Ts) [kN]                 : num  7.58 28.2 60.36 113.77 175.31 ...
##  $ Port Propeller Torque (Tp) [kN]                      : num  7.58 28.2 60.36 113.77 175.31 ...
##  $ Hight Pressure (HP) Turbine exit temperature (T48) [C]: num  464 635 606 661 731 ...
##  $ GT Compressor inlet air temperature (T1) [C]         : num  288 288 288 288 288 288 288 288 288 1
##  $ GT Compressor outlet air temperature (T2) [C]        : num  551 582 588 614 646 ...
##  $ HP Turbine exit pressure (P48) [bar]                 : num  1.1 1.33 1.39 1.66 2.08 ...
##  $ GT Compressor inlet air pressure (P1) [bar]          : num  0.998 0.998 0.998 0.998 0.998 0.998 0
##  $ GT Compressor outlet air pressure (P2) [bar]         : num  5.95 7.28 7.57 9.01 11.2 ...
##  $ GT exhaust gas pressure (Pexh) [bar]                 : num  1.02 1.02 1.02 1.02 1.03 ...
##  $ Turbine Injecton Control (TIC) [%]                   : num  7.14 10.65 13.09 18.11 26.37 ...
##  $ Fuel flow (mf) [kg/s]                                : num  0.082 0.287 0.259 0.358 0.522 ...
##  $ GT Compressor decay state coefficient                : num  0.95 0.95 0.95 0.95 0.95 0.95 0.95 0
##  $ GT_Turbine_decay_state_coefficient                   : num  0.975 0.975 0.975 0.975 0.975 0.975 0
```

checking the quantity of features and variables in cbm data.

```
dim(cbm)
```

```
## [1] 11934    18
```

numerical summary of the cbm dataset loike mean,median and mode.

```
summary(cbm)
```

```
##  Lever position (lp) Ship speed (v) [knots]
##  Min.   :1.138       Min.   : 3
##  1st Qu.:3.144       1st Qu.: 9
##  Median :5.140       Median :15
##  Mean   :5.167       Mean   :15
##  3rd Qu.:7.148       3rd Qu.:21
##  Max.   :9.300       Max.   :27
##  Gas Turbine (GT) shaft torque (GTT) [kN m] GT rate of revolutions (GTn) [rpm]
##  Min.   :  253.6                            Min.   :1308
##  1st Qu.: 8375.9                            1st Qu.:1387
##  Median :21630.7                            Median :1924
##  Mean   :27247.5                            Mean   :2136
##  3rd Qu.:39001.4                            3rd Qu.:2678
```

```
##  Max.   :72784.9                              Max.   :3561
##  Gas Generator rate of revolutions (GGn) [rpm]
##  Min.   :6589
##  1st Qu.:7058
##  Median :8482
##  Mean   :8201
##  3rd Qu.:9133
##  Max.   :9797
##  Starboard Propeller Torque (Ts) [kN] Port Propeller Torque (Tp) [kN]
##  Min.   :  5.304                      Min.   :  5.304
##  1st Qu.: 60.317                      1st Qu.: 60.317
##  Median :175.268                      Median :175.268
##  Mean   :227.336                      Mean   :227.336
##  3rd Qu.:332.365                      3rd Qu.:332.365
##  Max.   :645.249                      Max.   :645.249
##  Hight Pressure (HP) Turbine exit temperature (T48) [C]
##  Min.   : 442.4
##  1st Qu.: 589.9
##  Median : 706.0
##  Mean   : 735.5
##  3rd Qu.: 834.1
##  Max.   :1115.8
##  GT Compressor inlet air temperature (T1) [C]
##  Min.   :288
##  1st Qu.:288
##  Median :288
##  Mean   :288
##  3rd Qu.:288
##  Max.   :288
##  GT Compressor outlet air temperature (T2) [C]
##  Min.   :540.4
##  1st Qu.:578.1
##  Median :637.1
##  Mean   :646.2
##  3rd Qu.:693.9
##  Max.   :789.1
##  HP Turbine exit pressure (P48) [bar]
##  Min.   :1.093
##  1st Qu.:1.389
##  Median :2.083
##  Mean   :2.353
##  3rd Qu.:2.981
##  Max.   :4.560
##  GT Compressor inlet air pressure (P1) [bar]
##  Min.   :0.998
##  1st Qu.:0.998
##  Median :0.998
##  Mean   :0.998
##  3rd Qu.:0.998
##  Max.   :0.998
##  GT Compressor outlet air pressure (P2) [bar]
##  Min.   : 5.828
##  1st Qu.: 7.447
##  Median :11.092
```

```
##  Mean   :12.297
##  3rd Qu.:15.658
##  Max.   :23.140
##  GT exhaust gas pressure (Pexh) [bar] Turbine Injecton Control (TIC) [%]
##  Min.   :1.019                       Min.   : 0.00
##  1st Qu.:1.020                       1st Qu.:13.68
##  Median :1.026                       Median :25.28
##  Mean   :1.029                       Mean   :33.64
##  3rd Qu.:1.036                       3rd Qu.:44.55
##  Max.   :1.052                       Max.   :92.56
##  Fuel flow (mf) [kg/s] GT Compressor decay state coefficient
##  Min.   :0.0680        Min.   :0.950
##  1st Qu.:0.2460        1st Qu.:0.962
##  Median :0.4960        Median :0.975
##  Mean   :0.6624        Mean   :0.975
##  3rd Qu.:0.8820        3rd Qu.:0.988
##  Max.   :1.8320        Max.   :1.000
##  GT_Turbine_decay_state_coefficient
##  Min.   :0.9750
##  1st Qu.:0.9810
##  Median :0.9875
##  Mean   :0.9875
##  3rd Qu.:0.9940
##  Max.   :1.0000
```

Checking for missing values.Found no missing values in the data.

```
sum(is.na(cbm))
```

```
## [1] 0
```

Checking for null values.Found no missing values in the data.

```
sum(is.null(cbm))
```

```
## [1] 0
```

checking **correlation** between feautures.

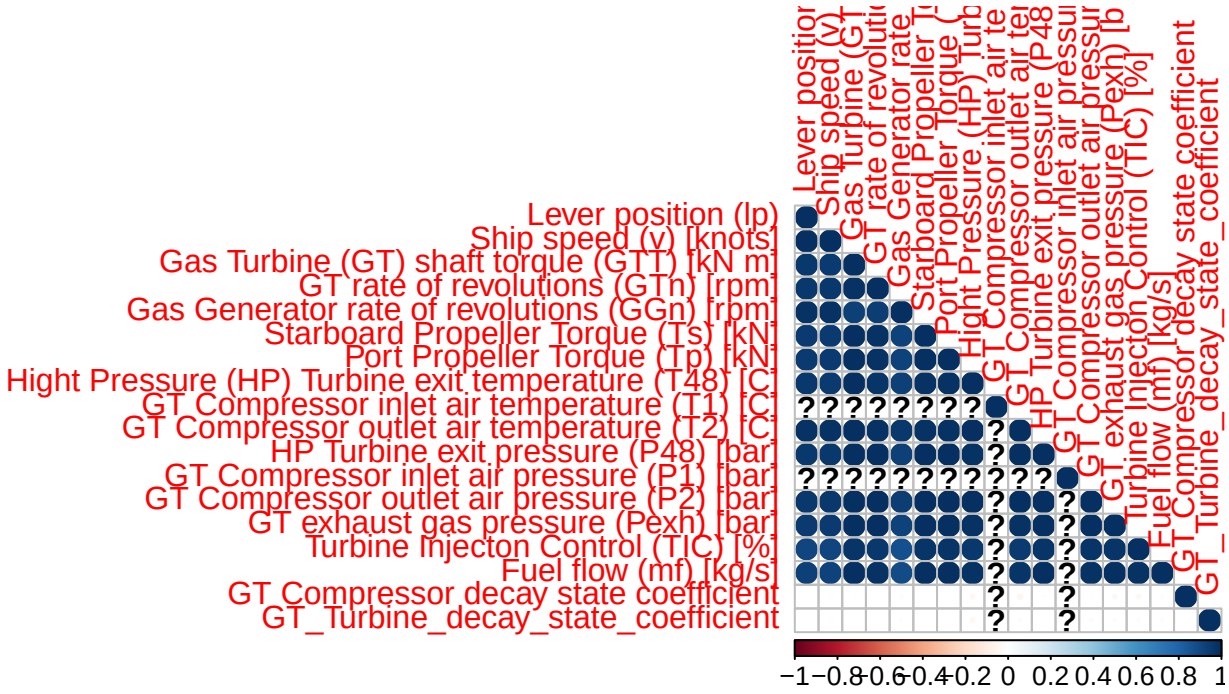
```
par(mfrow=c(1,1),bg="lightyellow")
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```
cbmcorplot<-cor(cbm)
```

```
## Warning in cor(cbm): the standard deviation is zero
```

```
corrplot(cbmcorplot,method = "circle",type="lower")
```



Insight 1 : GT Compressor inlet air temprature and GT compressor inlet air pressure after plot corrplot its need inspection. Rest all are shows postive corelation to each other.we found data is refelcting multicolinearity but we proceed to regression models.

```
var(cbm$`GT Compressor inlet air temperature (T1) [C]`)
```

```
## [1] 0
```

```
var(cbm$`GT Compressor inlet air pressure (P1) [bar]`)
```

```
## [1] 0
```

```
sum(cbm$`Starboard Propeller Torque (Ts) [kN]`-cbm$`Port Propeller Torque (Tp) [kN]`)
```

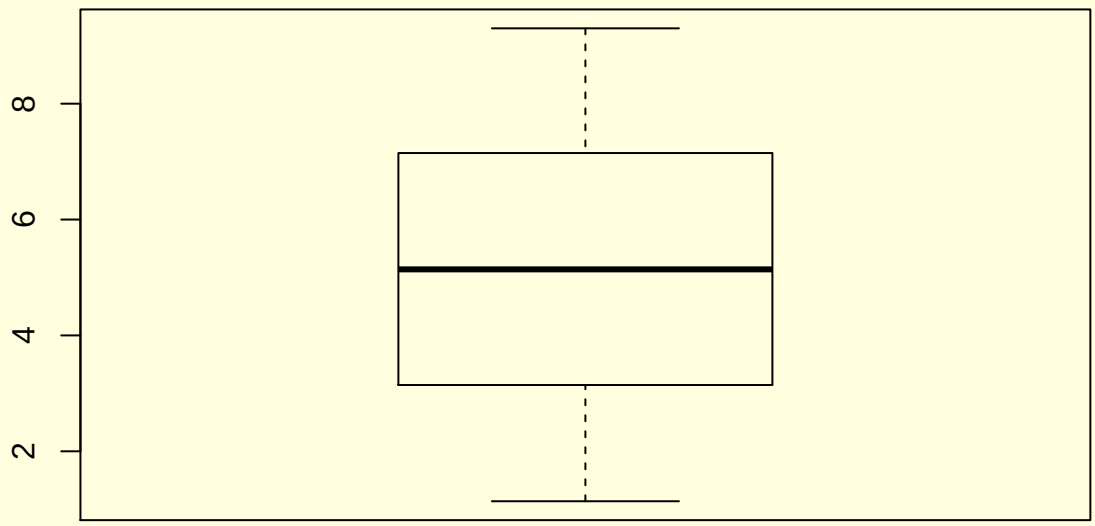
```
## [1] 0
```

Insight 2 : As variance is zero we will remove those features from the dataset.

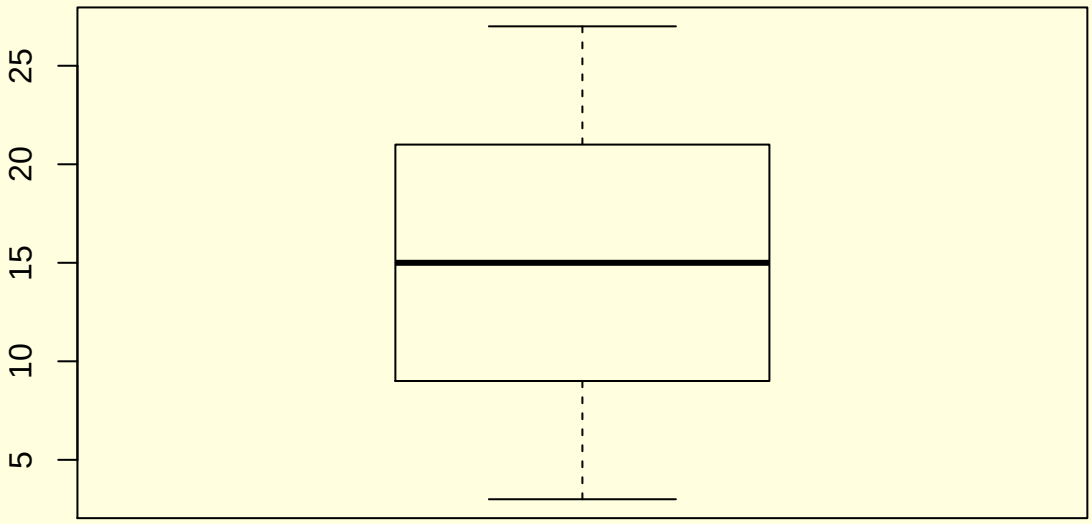
Exploiry data analysis

You can also embed plots, for example:

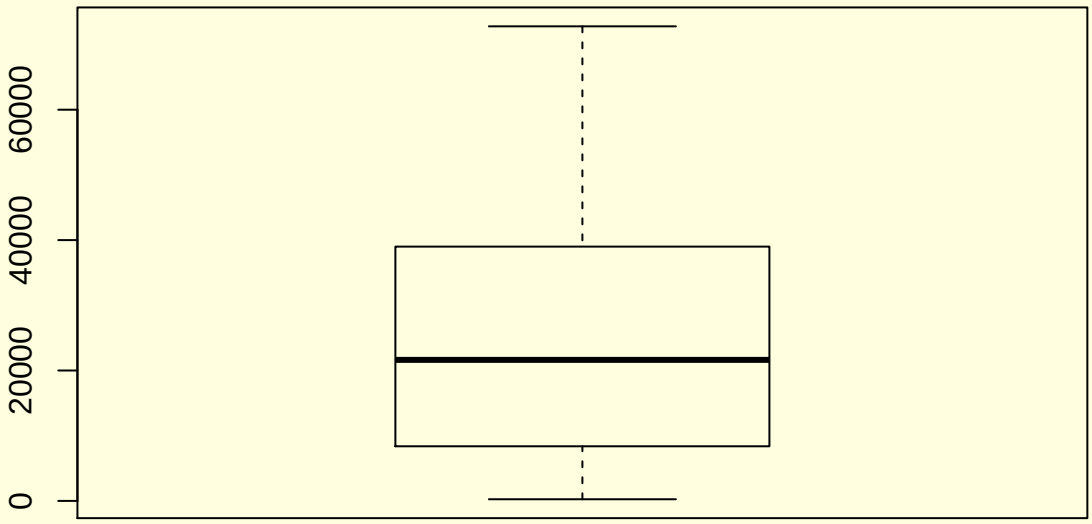
```
par(mfrow=c(1,1),bg="lightyellow")
boxplot(cbm$`Lever position (lp)`)
```



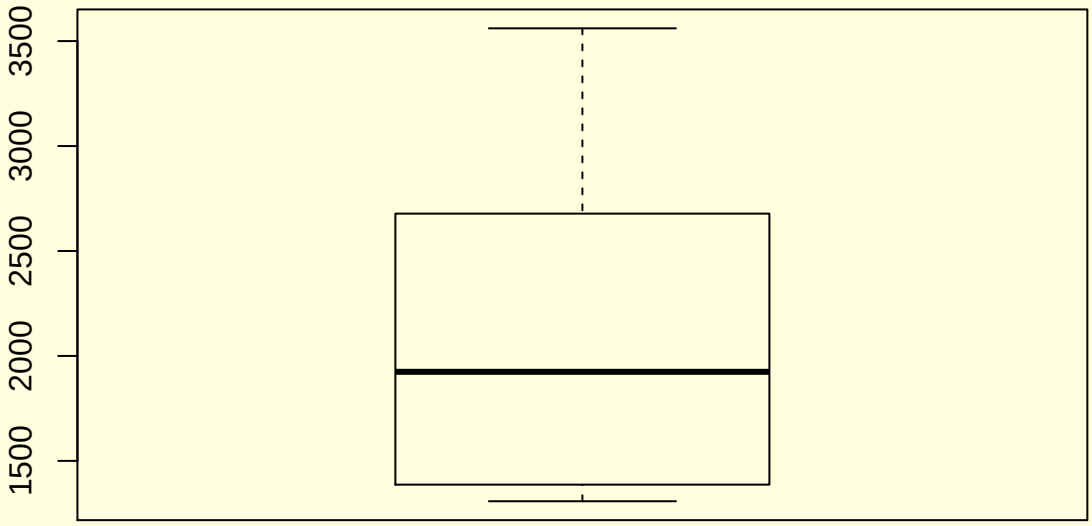

```
boxplot(cbm$`Ship speed (v) [knots]`)
```



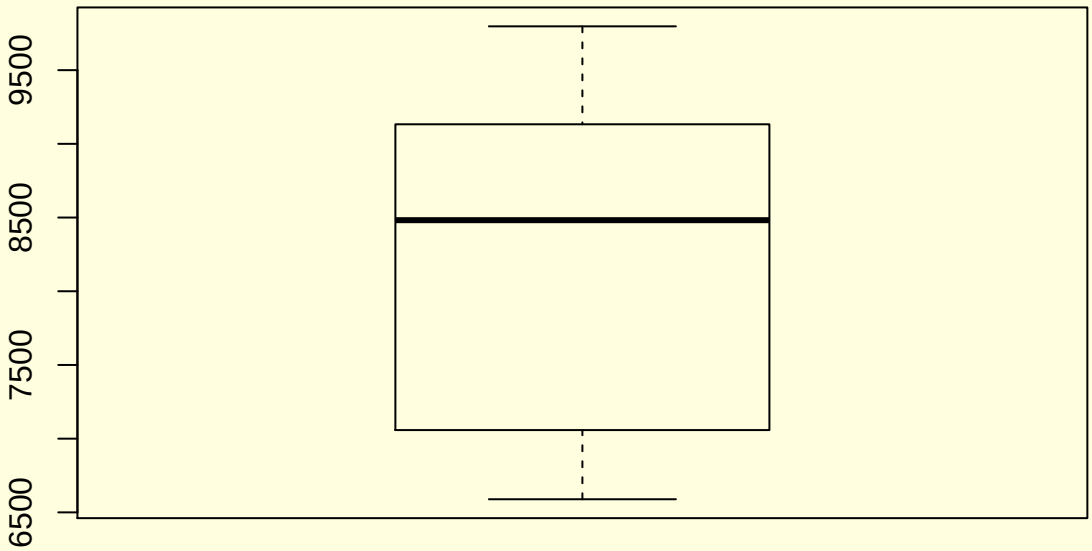

```
boxplot(cbm$`Gas Turbine (GT) shaft torque (GTT) [kN m]`)
```



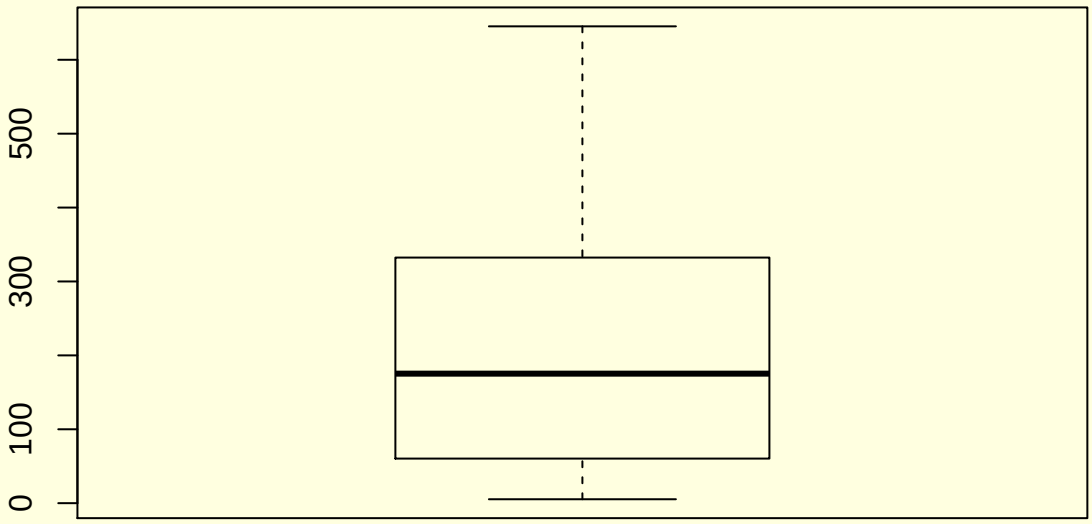

```
boxplot(cbm$`GT rate of revolutions (GTn) [rpm]`)
```



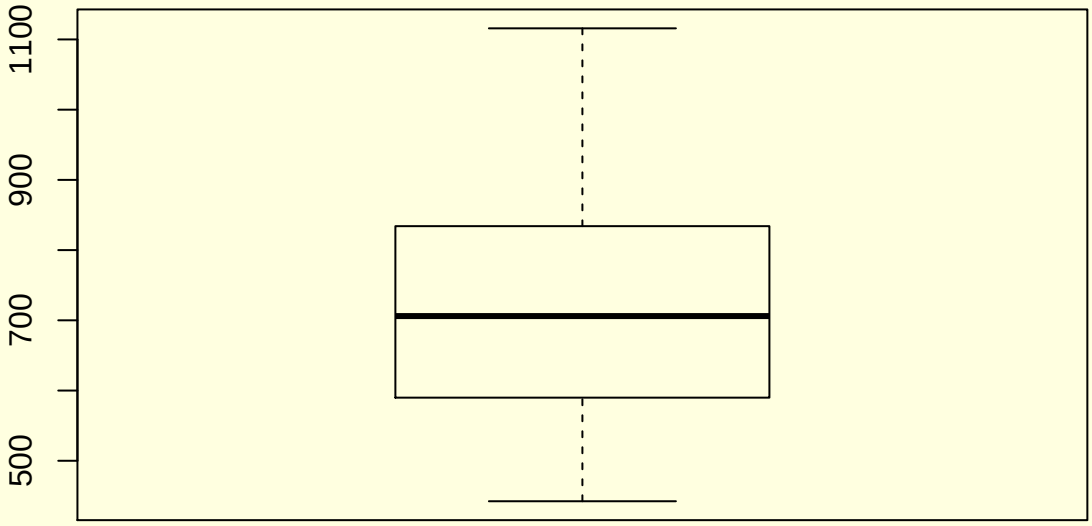

```
boxplot(cbm$`Gas Generator rate of revolutions (GGn) [rpm]`)
```



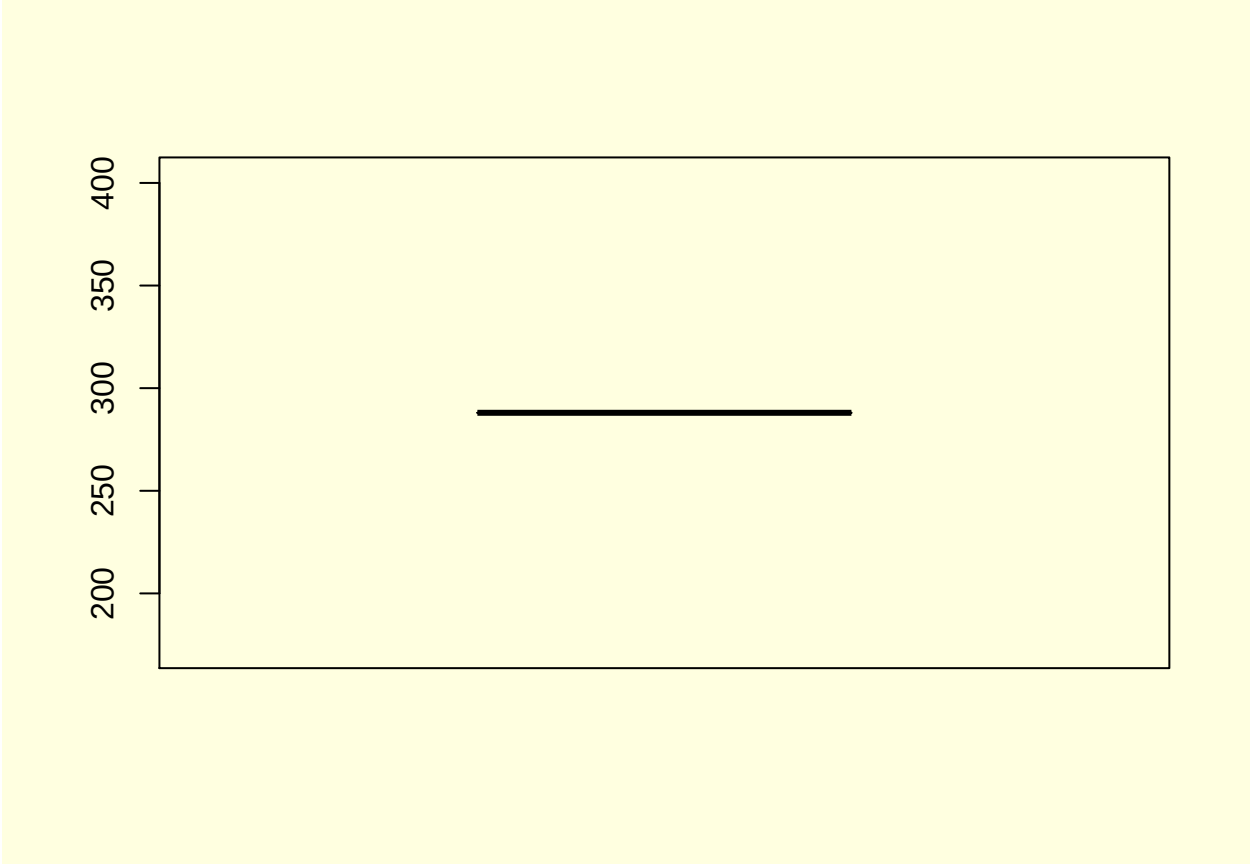

```
boxplot(cbm$`Starboard Propeller Torque (Ts) [kN]`)
boxplot(cbm$`Port Propeller Torque (Tp) [kN]`)
```



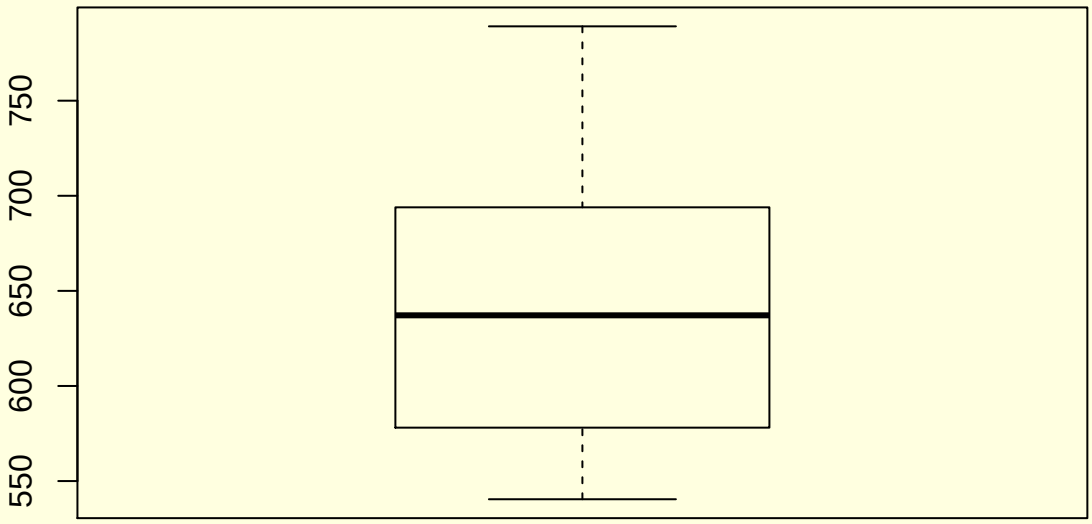

```
boxplot(cbm$`Hight Pressure (HP) Turbine exit temperature (T48) [C]`)
```



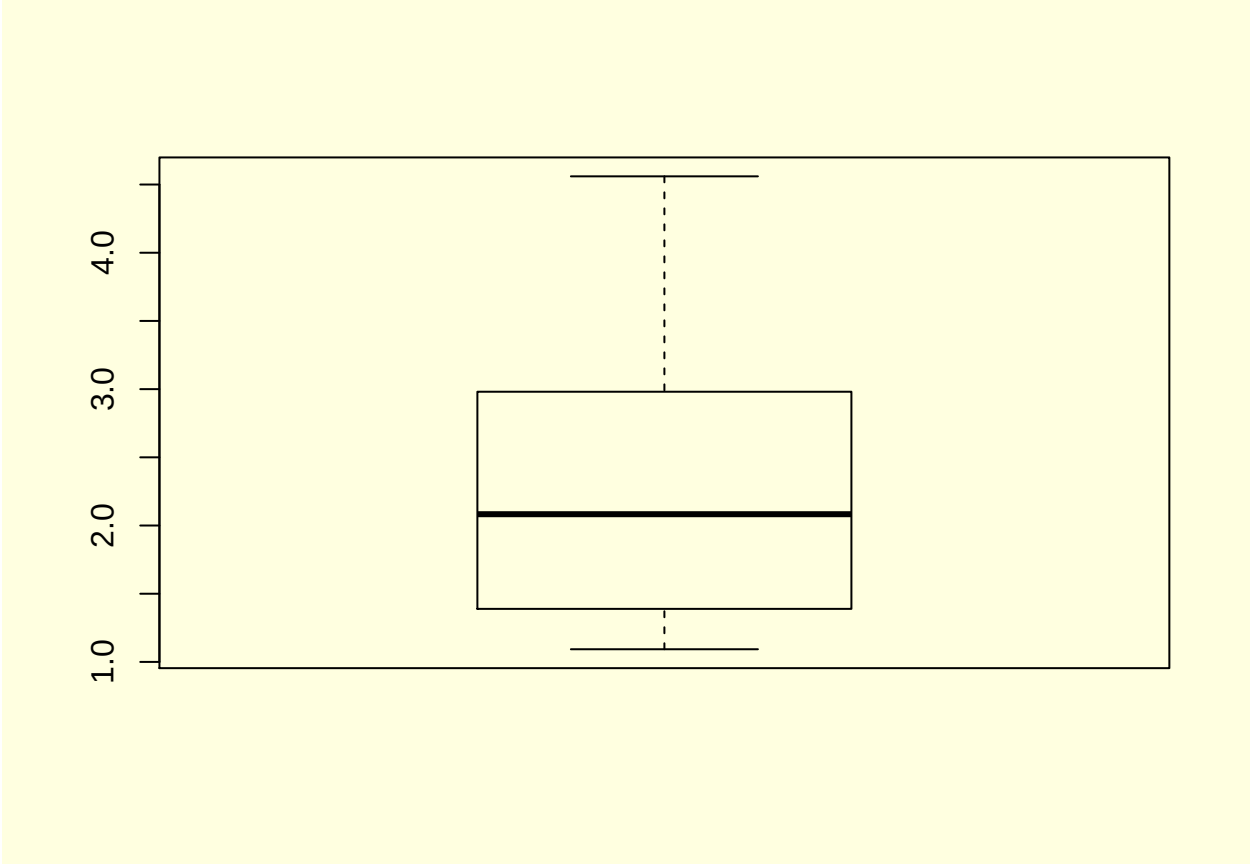

```
boxplot(cbm$`GT Compressor inlet air temperature (T1) [C]`)
```



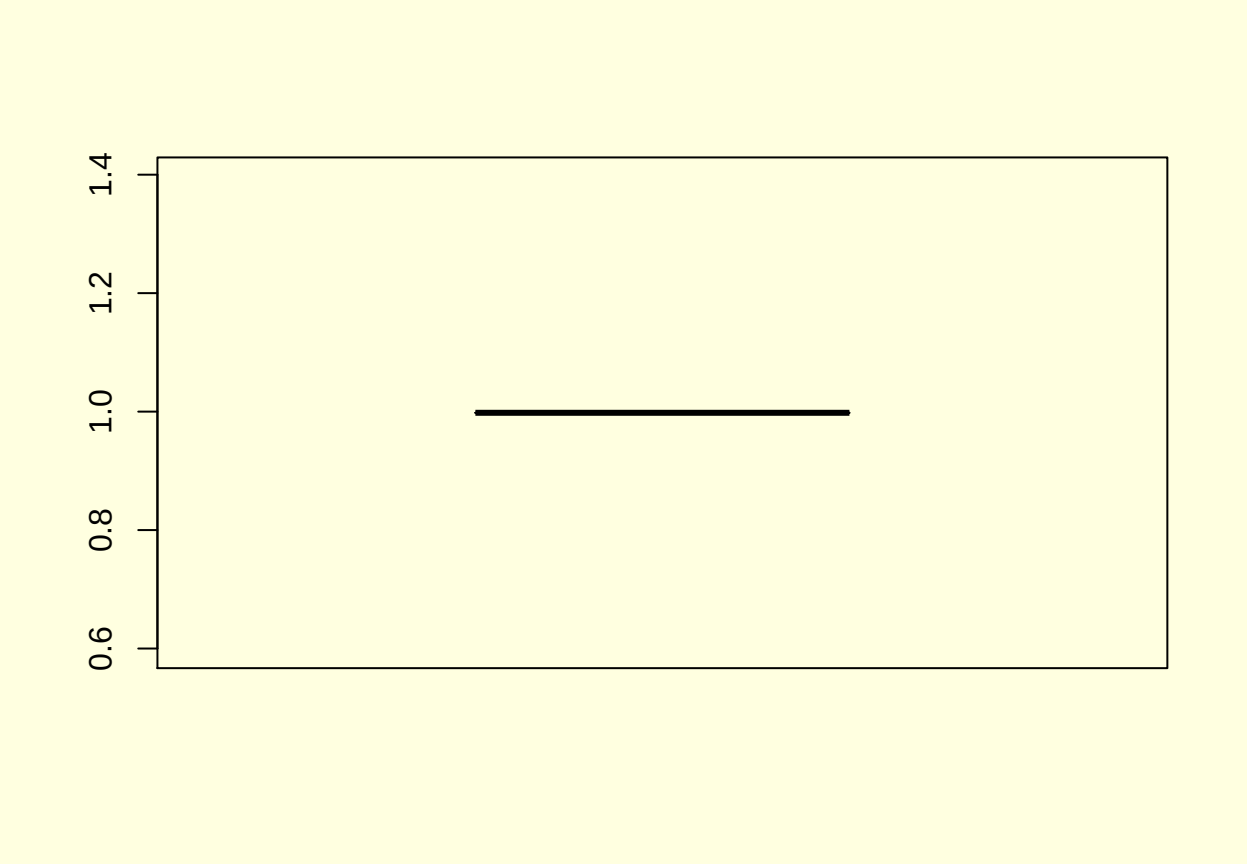

```
boxplot(cbm$`GT Compressor outlet air temperature (T2) [C]`)
```



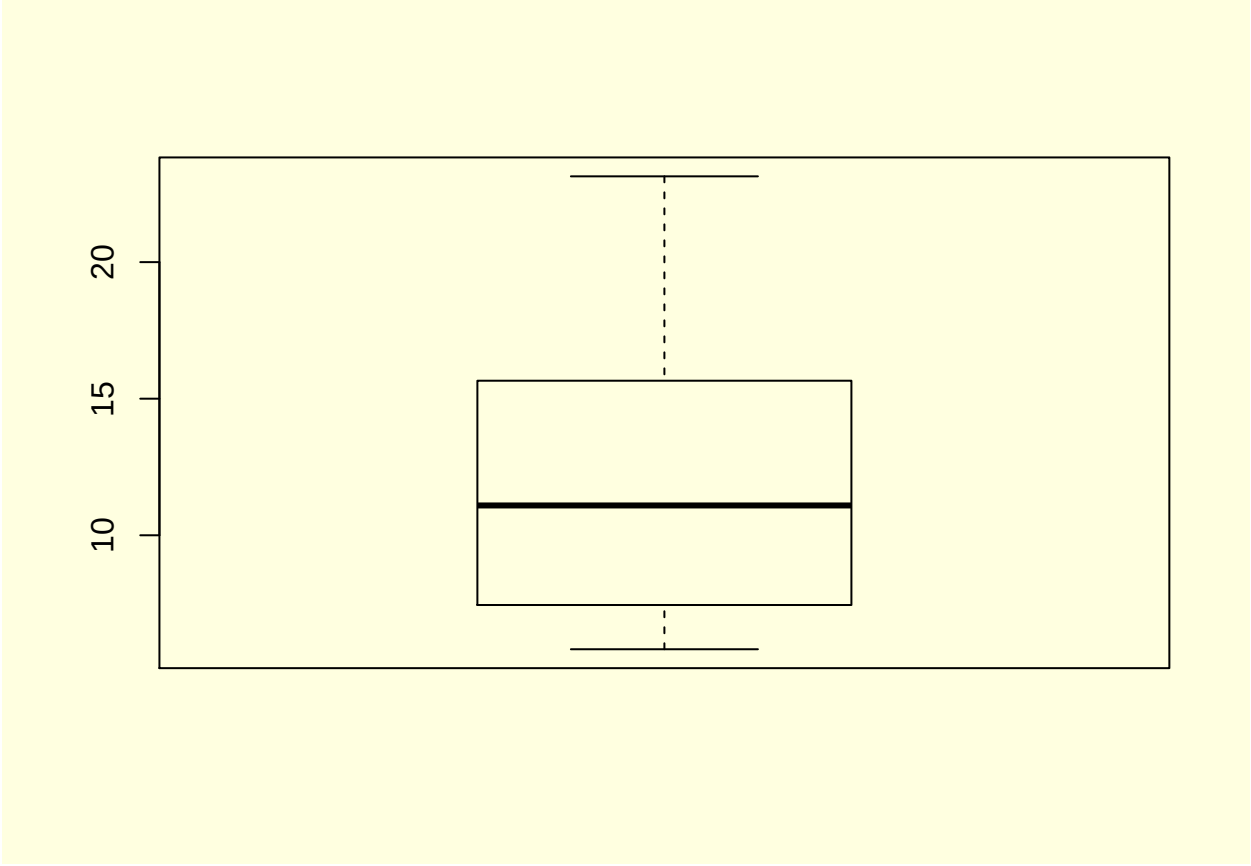

```
boxplot(cbm$`HP Turbine exit pressure (P48) [bar]`)
```



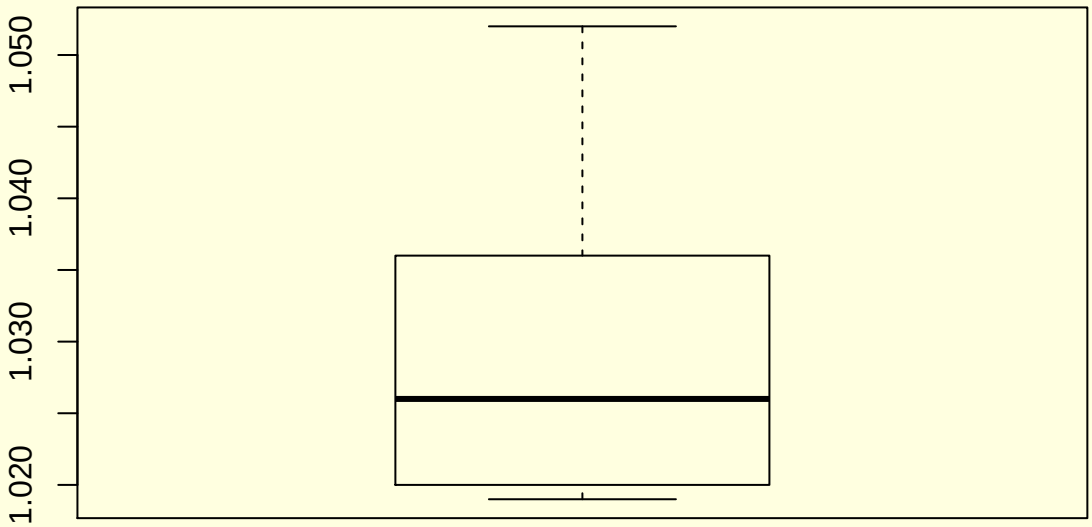

```
boxplot(cbm$`GT Compressor inlet air pressure (P1) [bar]`)
```



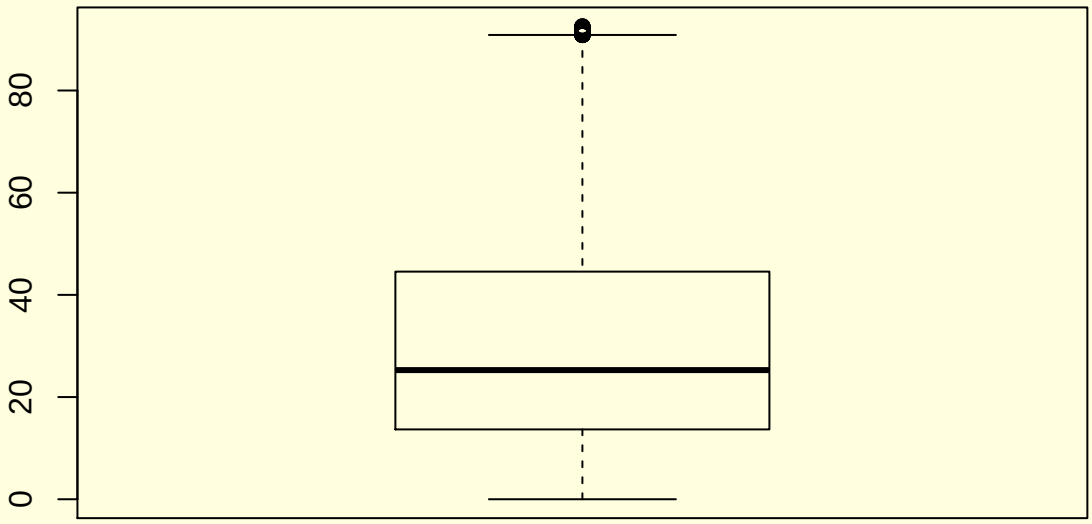

```
boxplot(cbm$`GT Compressor outlet air pressure (P2) [bar]`)
```



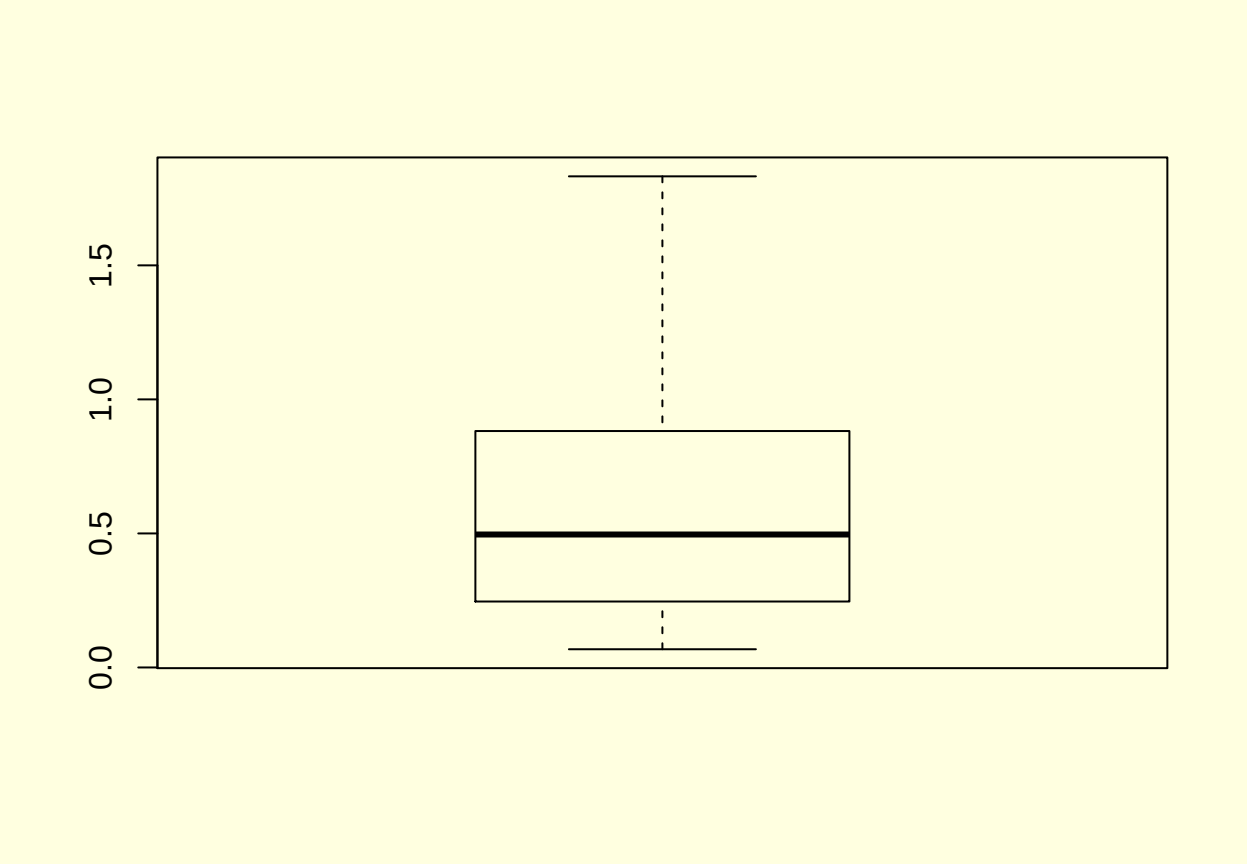

```
boxplot(cbm$`GT exhaust gas pressure (Pexh) [bar]`)
```



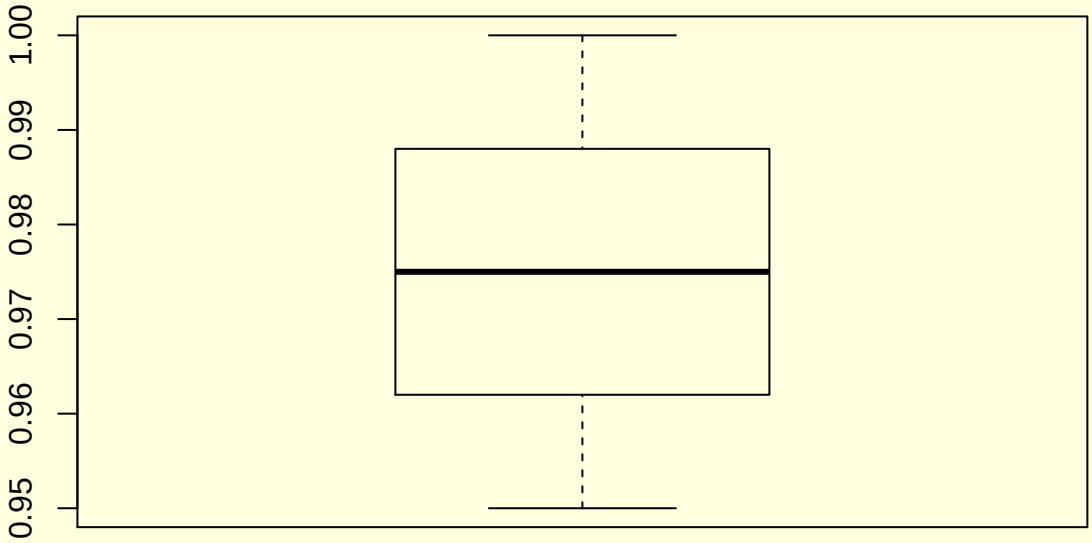

```
boxplot(cbm$`Turbine Injecton Control (TIC) [%]`)
```



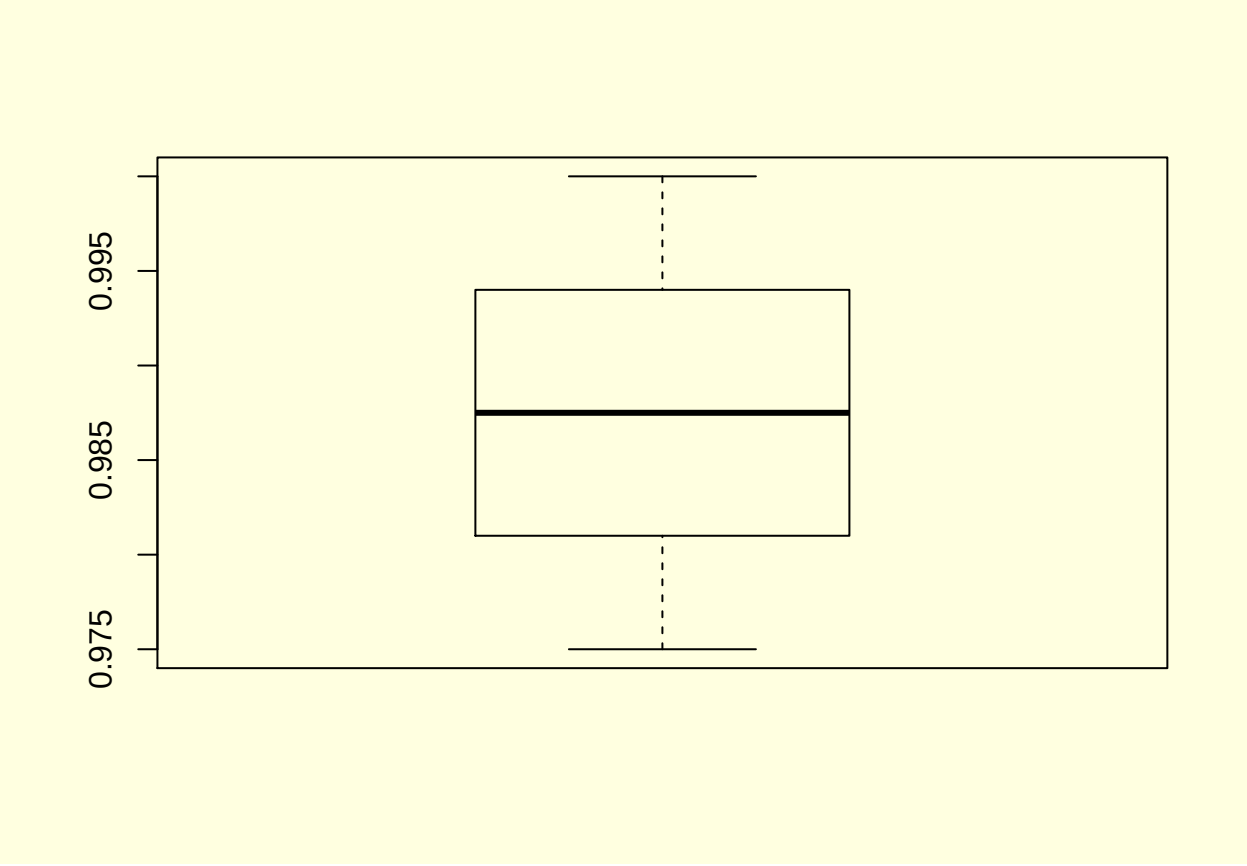

```
boxplot(cbm$`Fuel flow (mf) [kg/s]`)
```




```
boxplot(cbm$`GT Compressor decay state coefficient`)
```



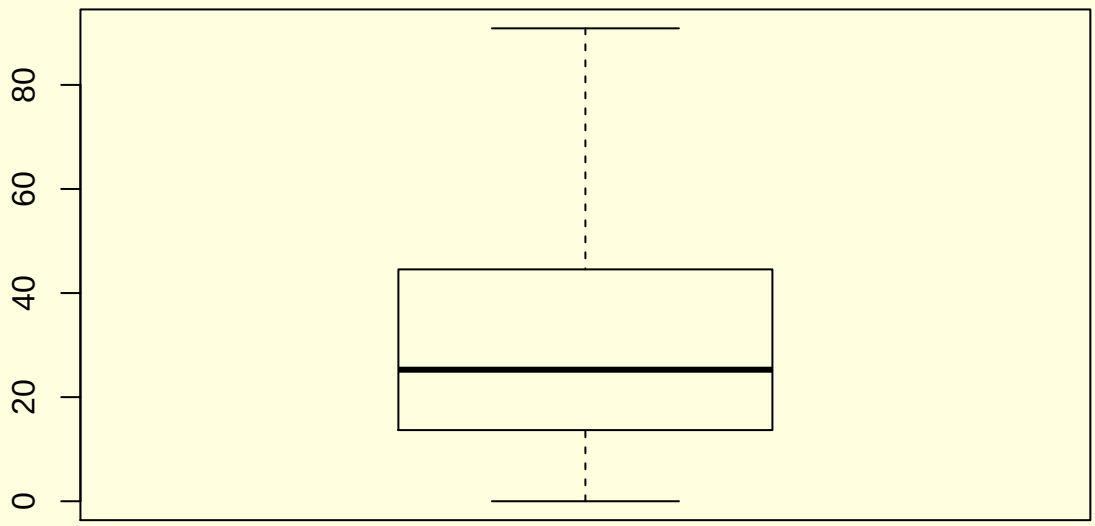

```
boxplot(cbm$`GT_Turbine_decay_state_coefficient`)
```



Insight 3: Tutbine injection control (TIC) is the feature has 188 outliers out of 11934 records is small.Even very small quantity still we handle outliers replacing by by upper whisker.

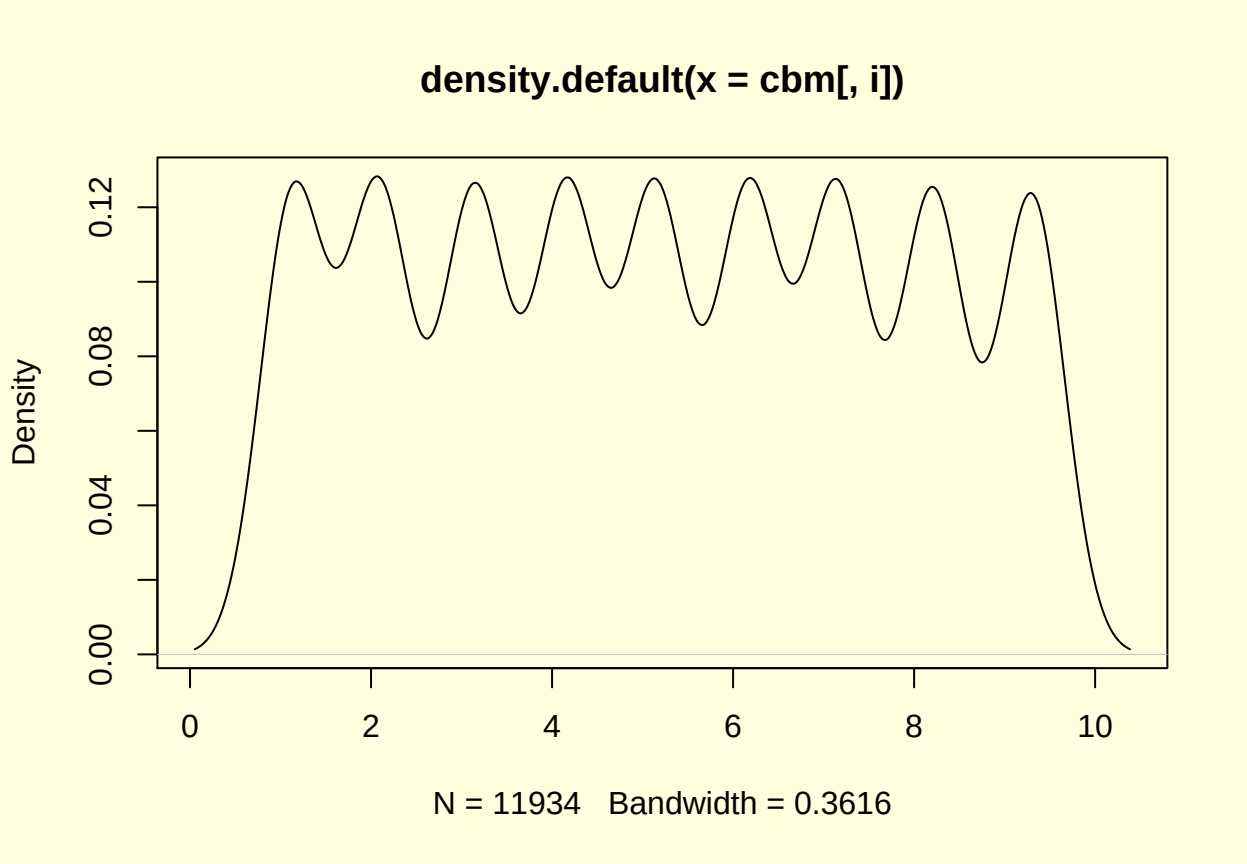
Handling outliers in the TIC feauture

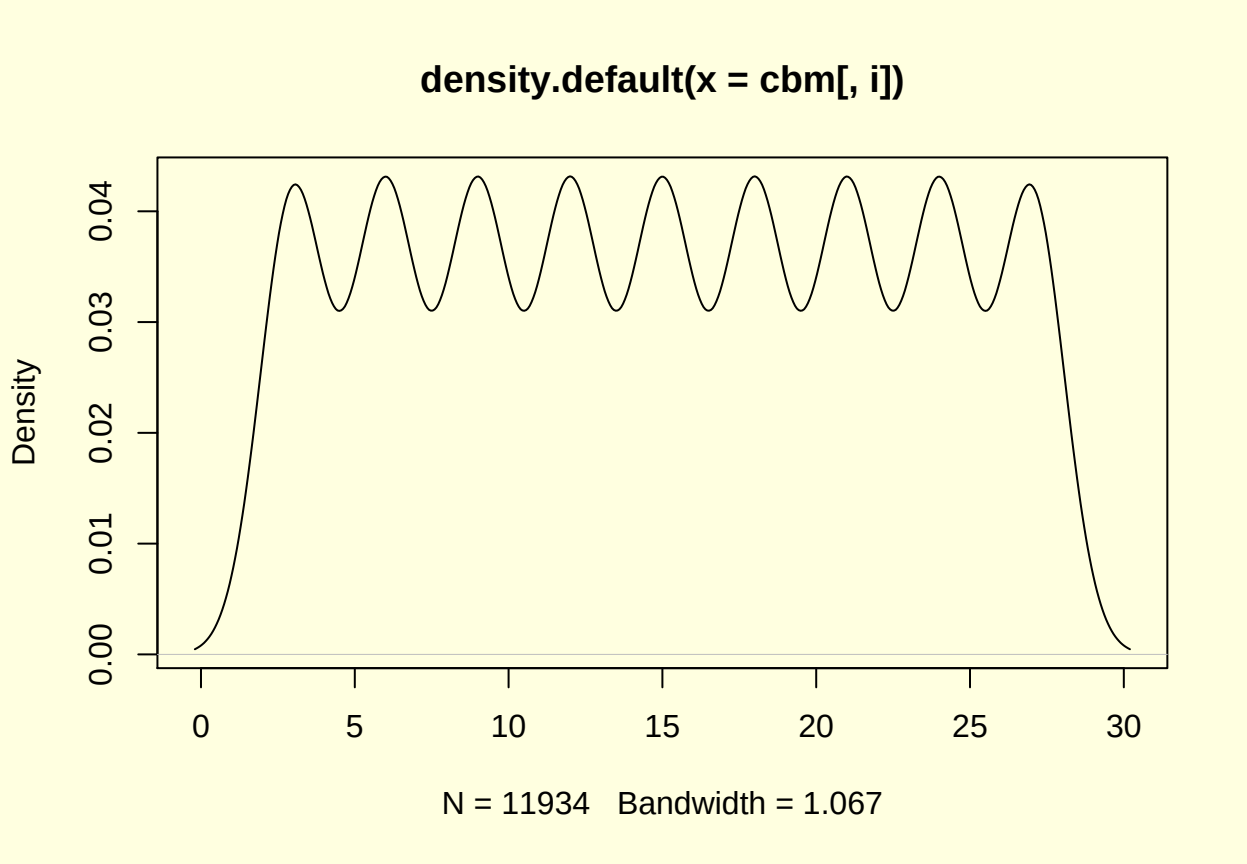
```
par(mfrow=c(1,1),bg="lightyellow")
#Function for replace outliers
replace_outliers = function(x,na.rm= TRUE){

qnt = quantile(x, probs = c(.25,.75))
outlier = 1.5*IQR(x)
x[x < (qnt[1]-outlier) ] <- qnt[1]
x[x > (qnt[2]+outlier) ] <- qnt[2]
return(x)
}
cbm$`Turbine Injecton Control (TIC) [%]`<-replace_outliers(cbm$`Turbine Injecton Control (TIC) [%]`)
boxplot(cbm$`Turbine Injecton Control (TIC) [%]`)
```

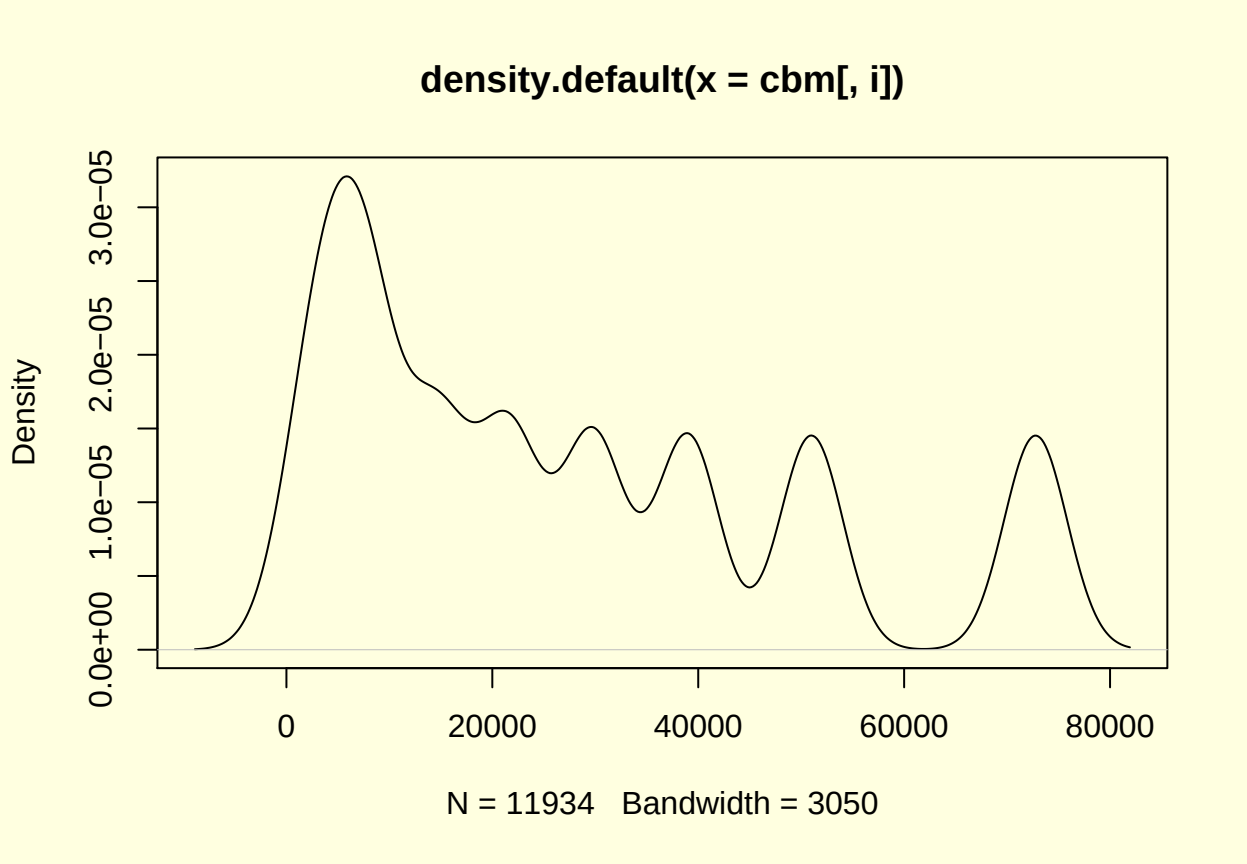


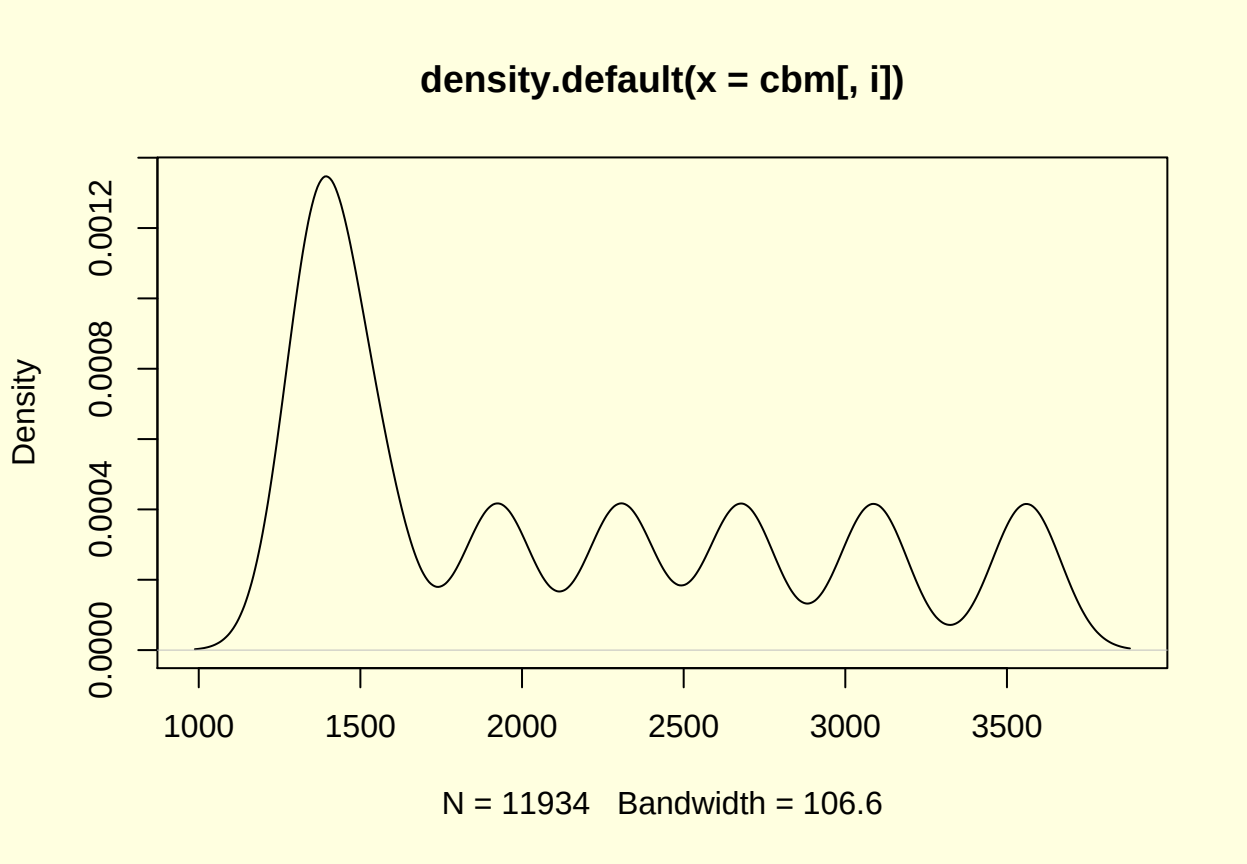
plotting **histogram** feautures.

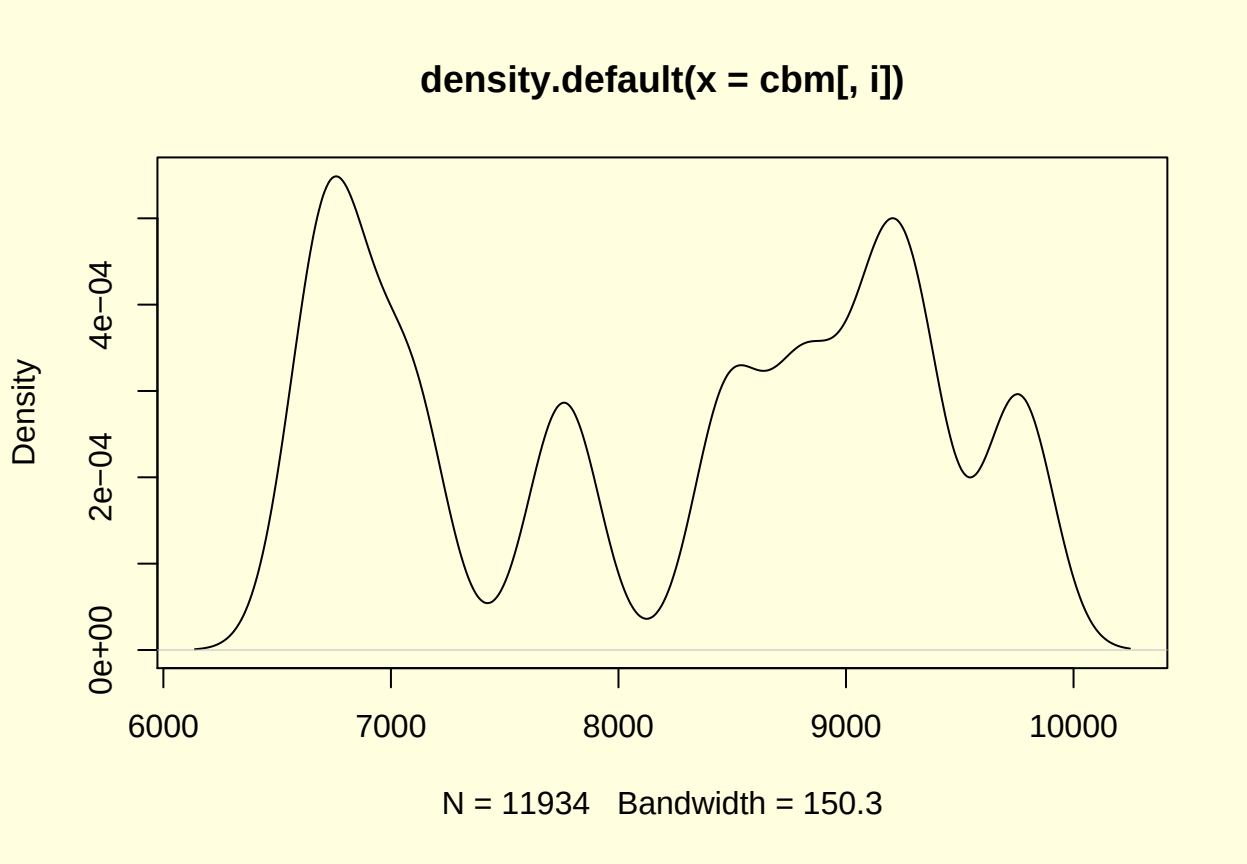
```
par(mfrow=c(1,1),bg="lightyellow")
for (i in 1:ncol(cbm)) {
  d<-density(cbm[,i])
  plot(d)

}
```

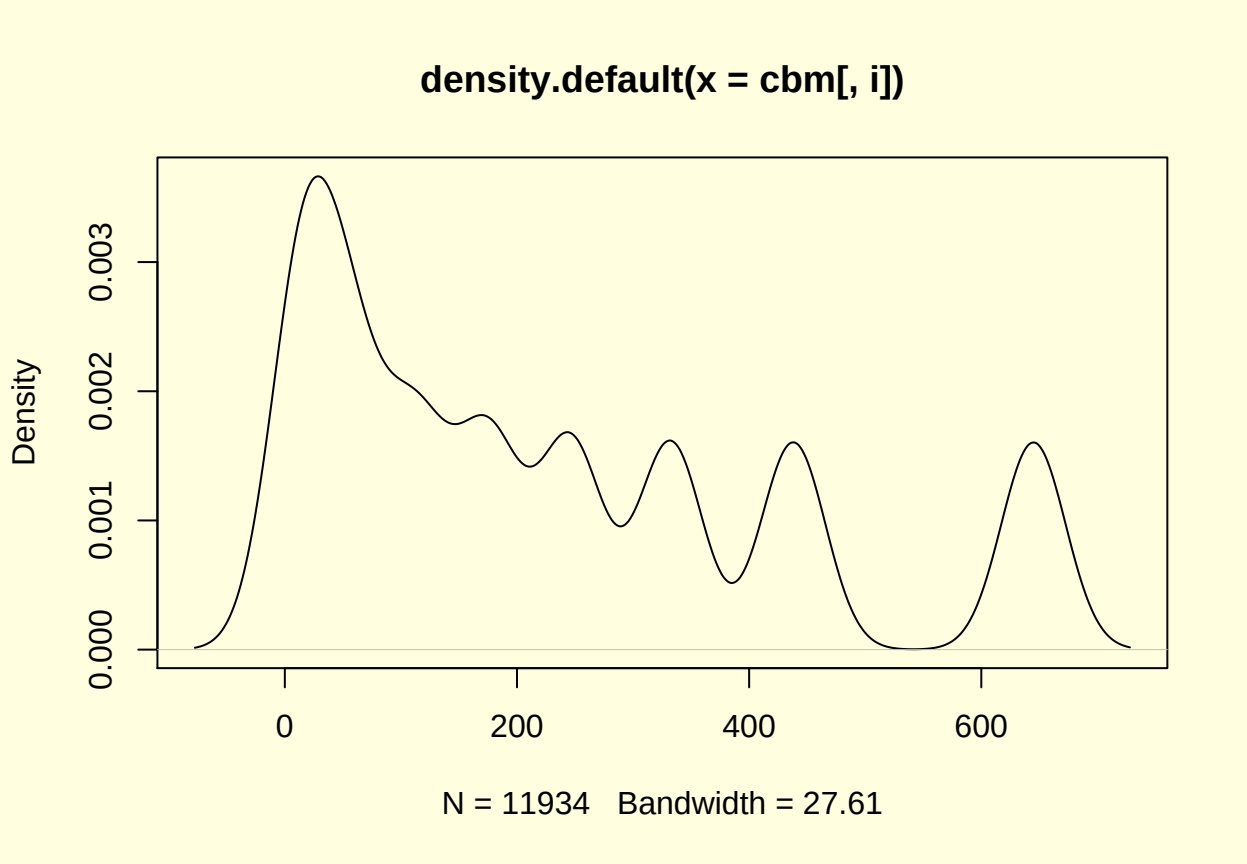

density.default(x = cbm[, i])
Density
0.00
0.04
0.08
0.12
0
2
4
6
8
10
N = 11934 Bandwidth = 0.3616

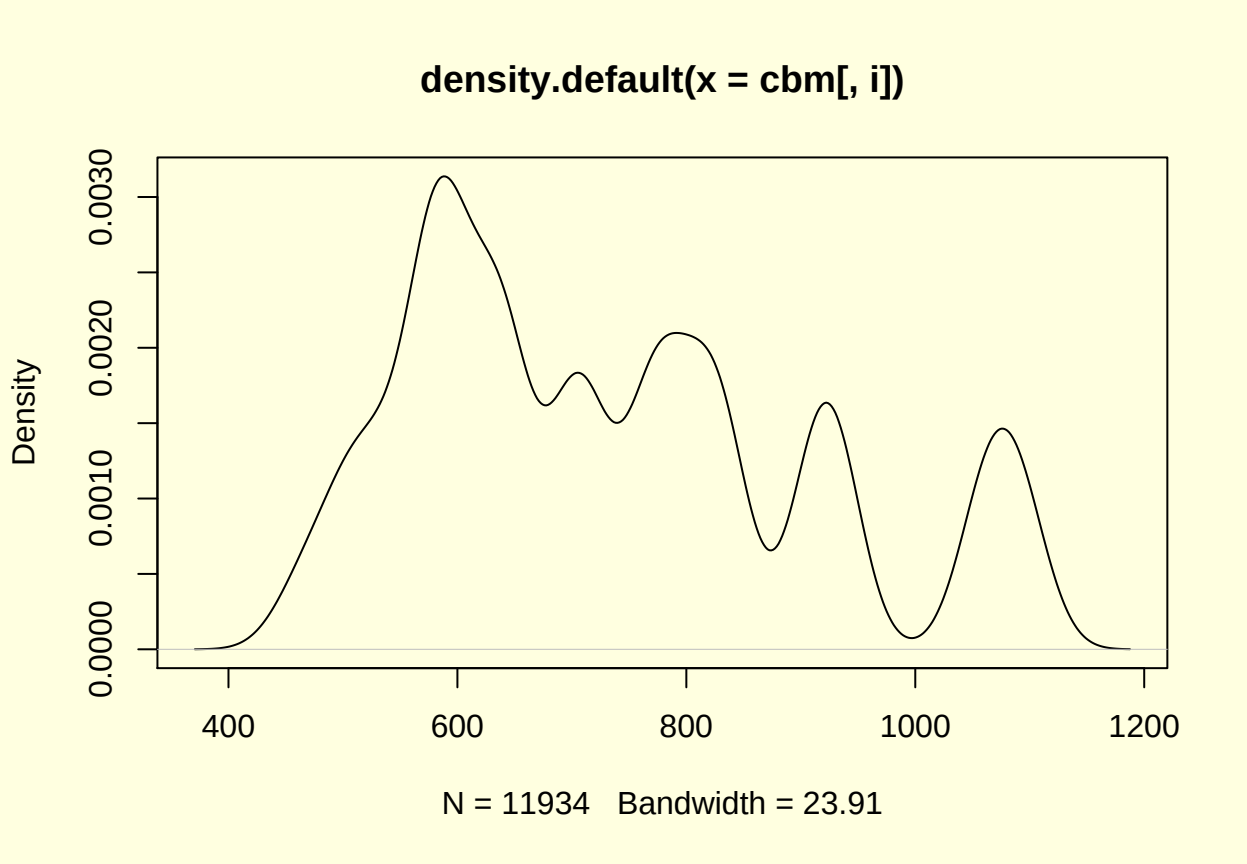

density.default(x = cbm[, i])
Density
0.00
0.01
0.02
0.03
0.04
0
5
10
15
20
25
30
N = 11934 Bandwidth = 1.067

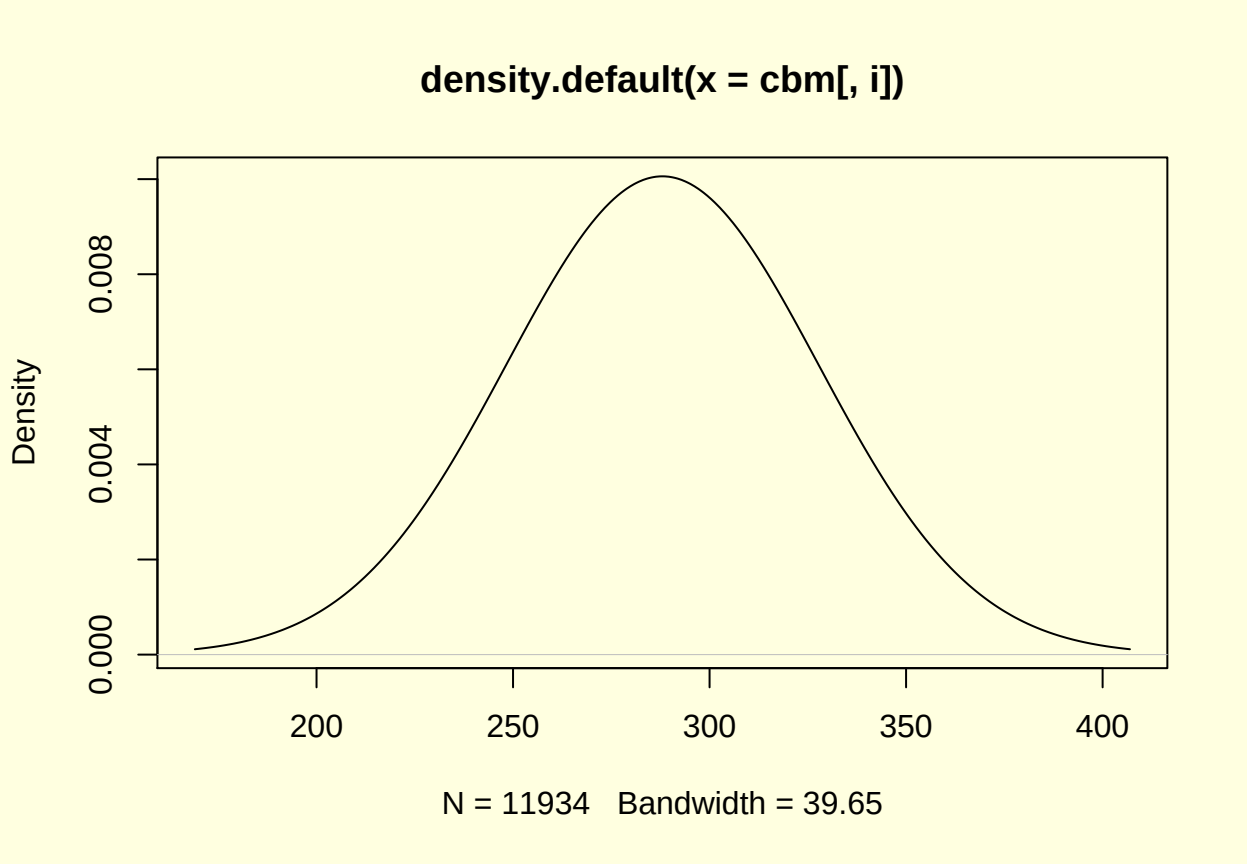

density.default(x = cbm[, i])
Density
0.0e+00
1.0e−05
2.0e−05
3.0e−05
0
20000
40000
60000
80000
N = 11934 Bandwidth = 3050

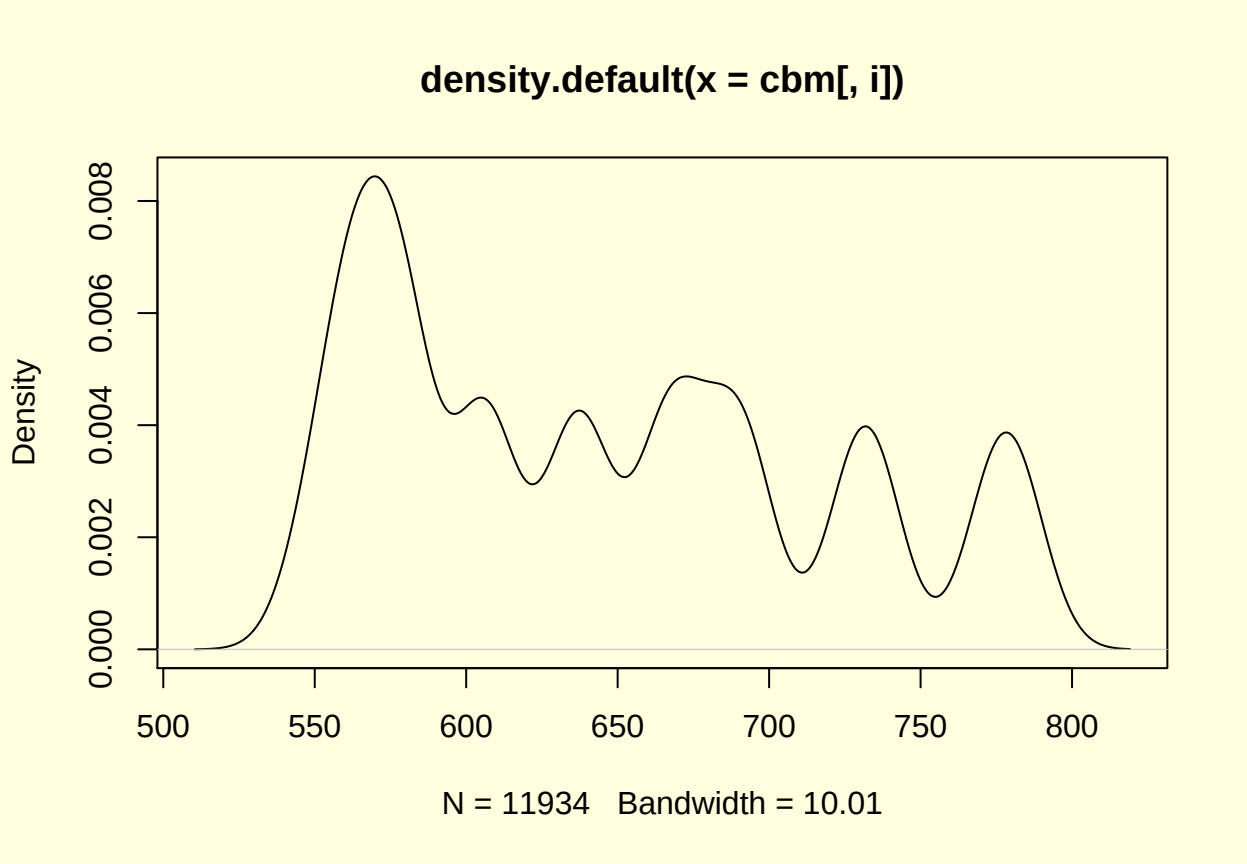

density.default(x = cbm[, i])
Density
0.0000
0.0004
0.0008
0.0012
1000
1500
2000
2500
3000
3500
N = 11934 Bandwidth = 106.6

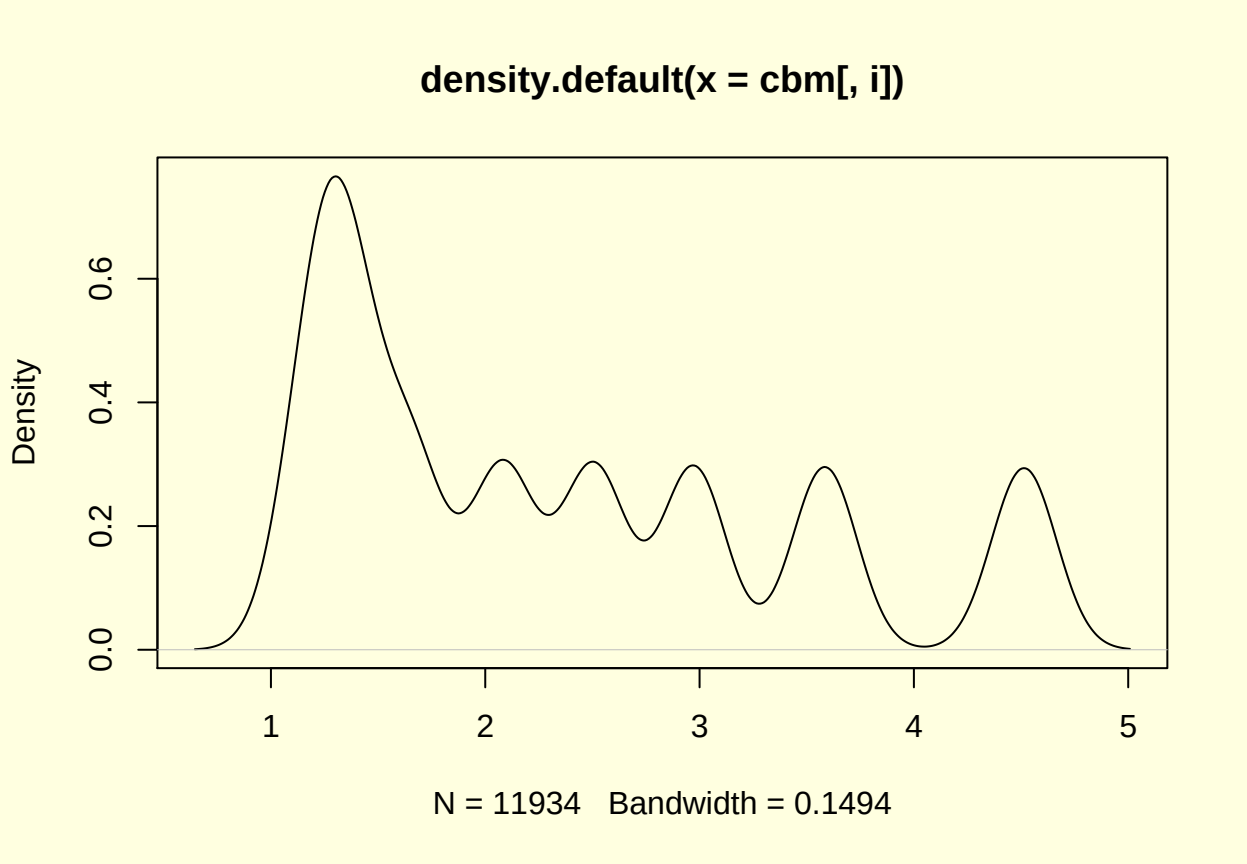

density.default(x = cbm[, i])
Density
0e+00
2e−04
4e−04
6000
7000
8000
9000
10000
N = 11934 Bandwidth = 150.3

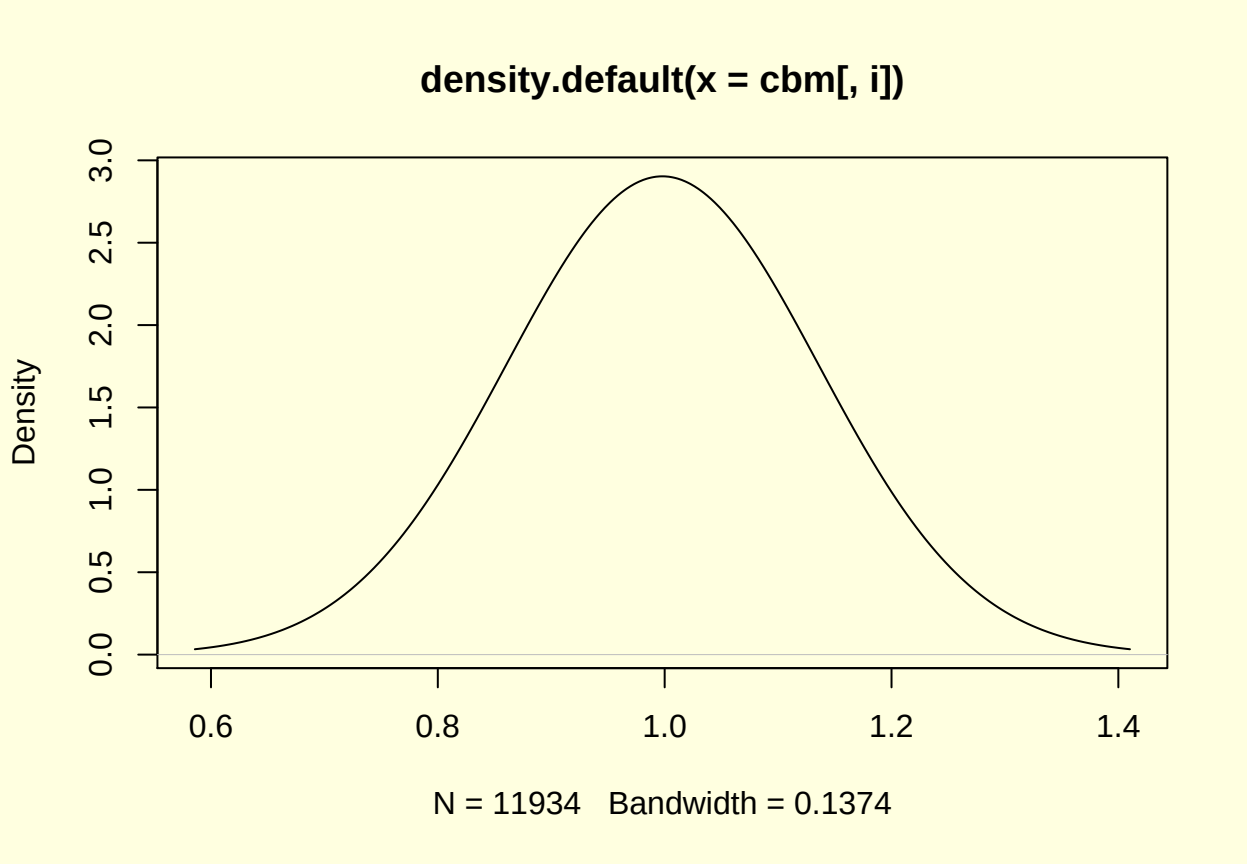

density.default(x = cbm[, i])
Density
0.000
0.001
0.002
0.003
0
200
400
600
N = 11934 Bandwidth = 27.61

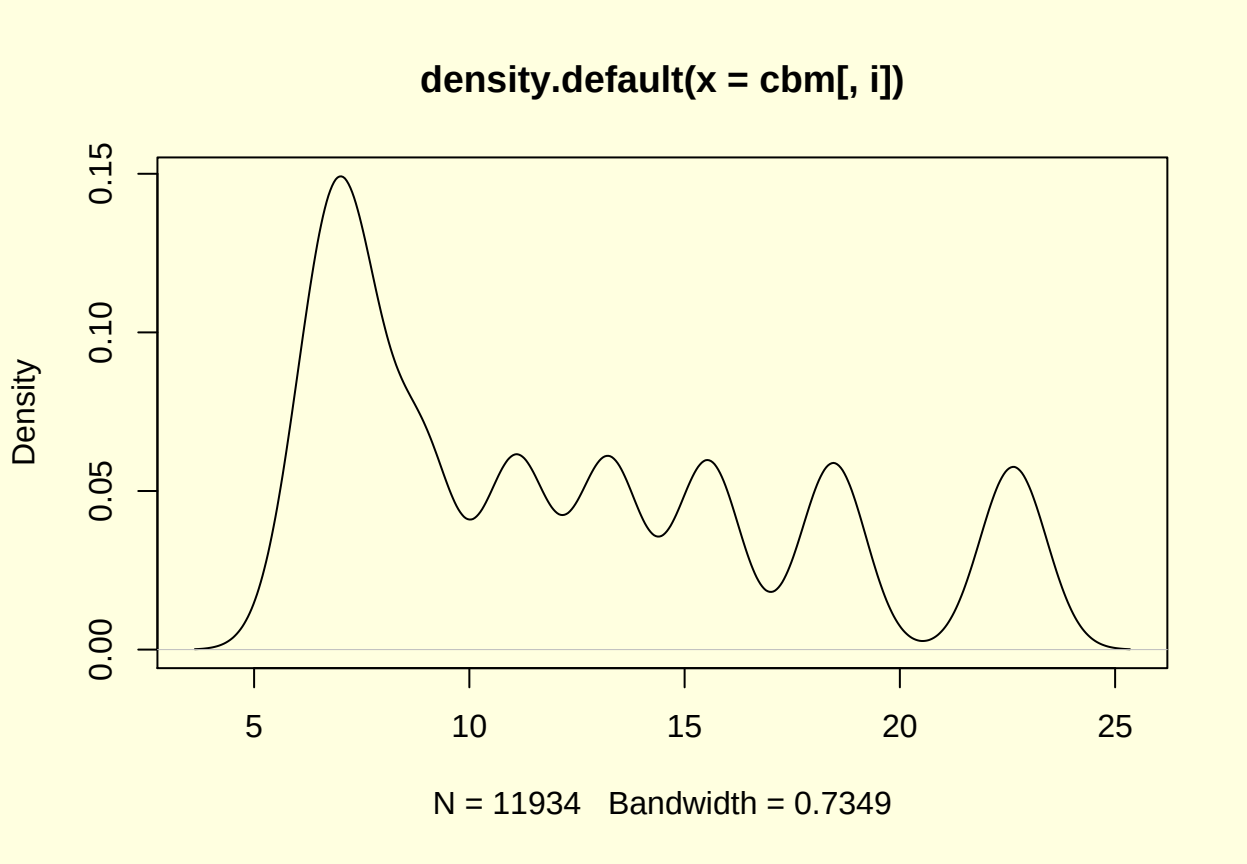

density.default(x = cbm[, i])
Density
0.0000
0.0010
0.0020
0.0030
400
600
800
1000
1200
N = 11934 Bandwidth = 23.91

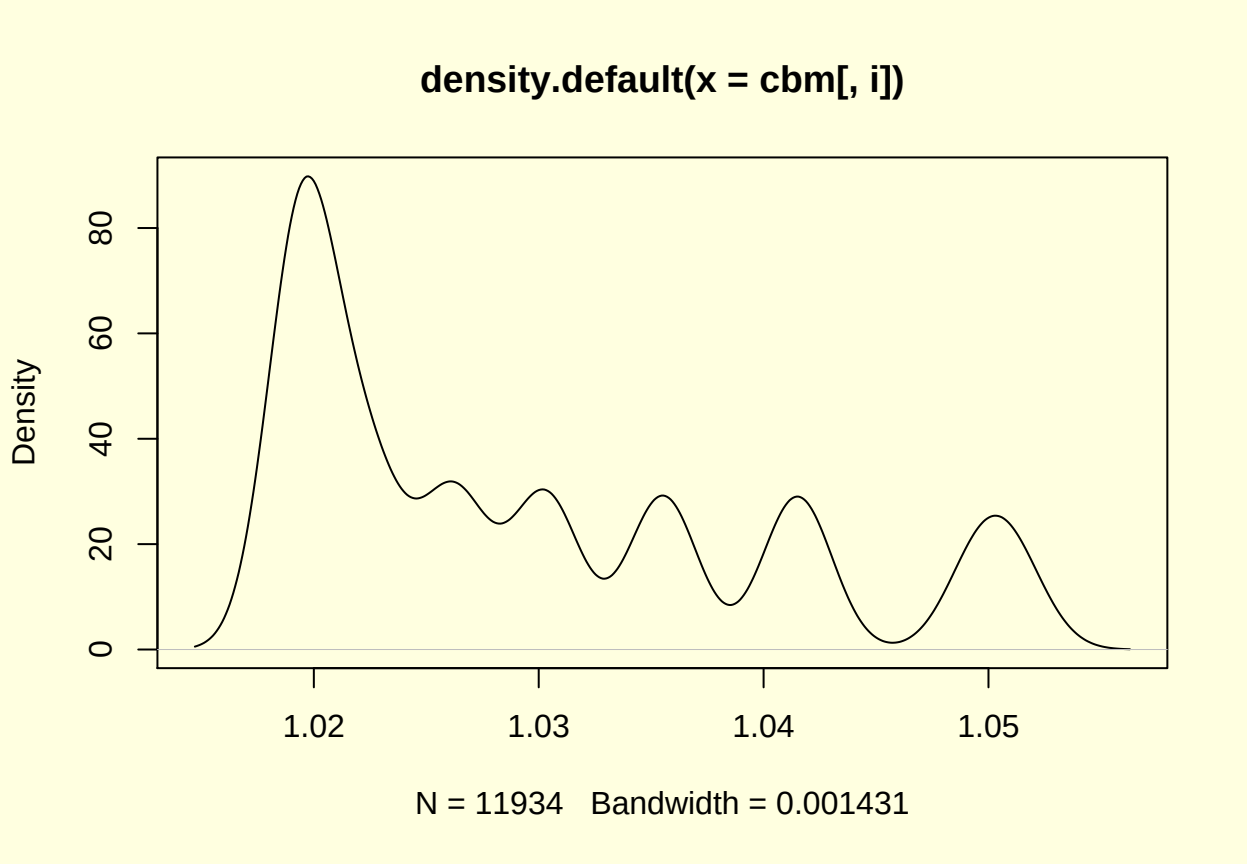

density.default(x = cbm[, i])
Density
0.000
0.004
0.008
200
250
300
350
400
N = 11934 Bandwidth = 39.65

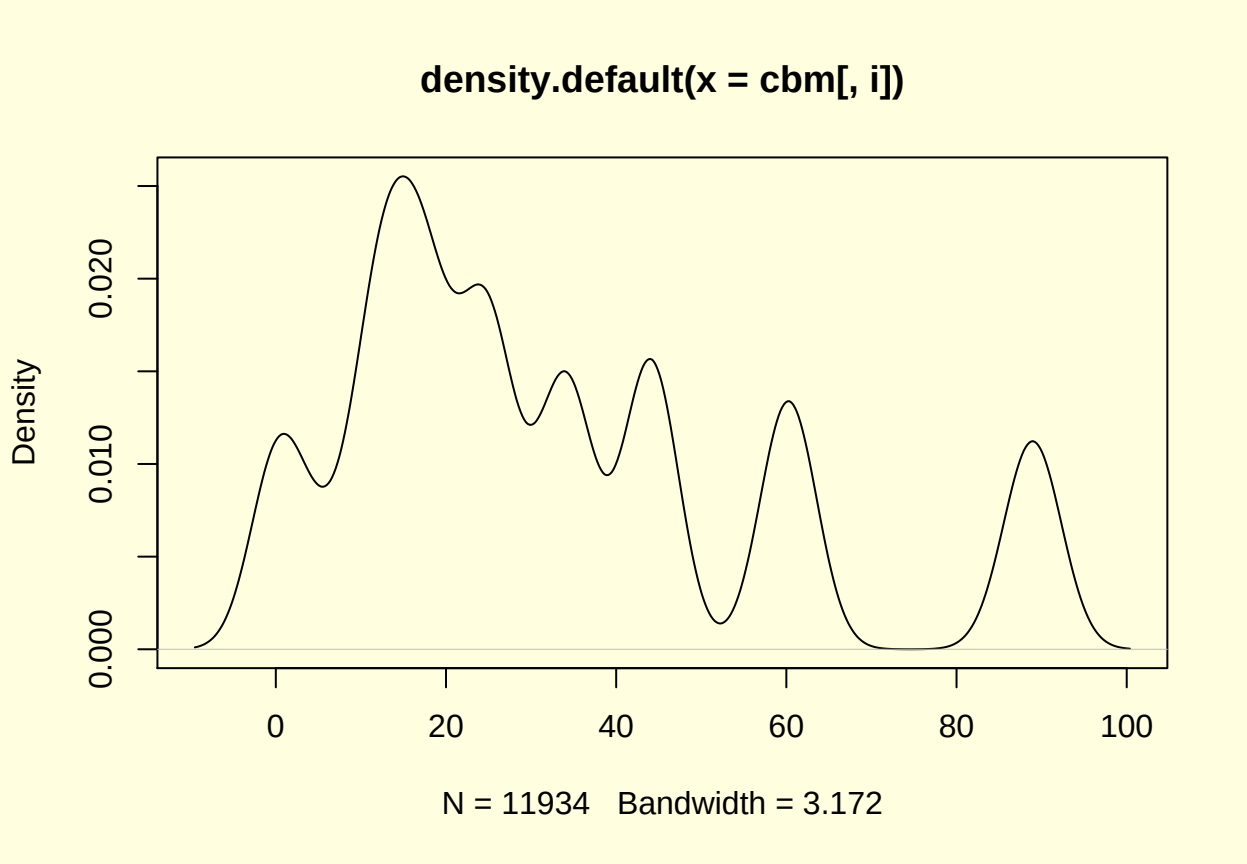

density.default(x = cbm[, i])
Density
0.000
0.002
0.004
0.006
0.008
500
550
600
650
700
750
800
N = 11934 Bandwidth = 10.01

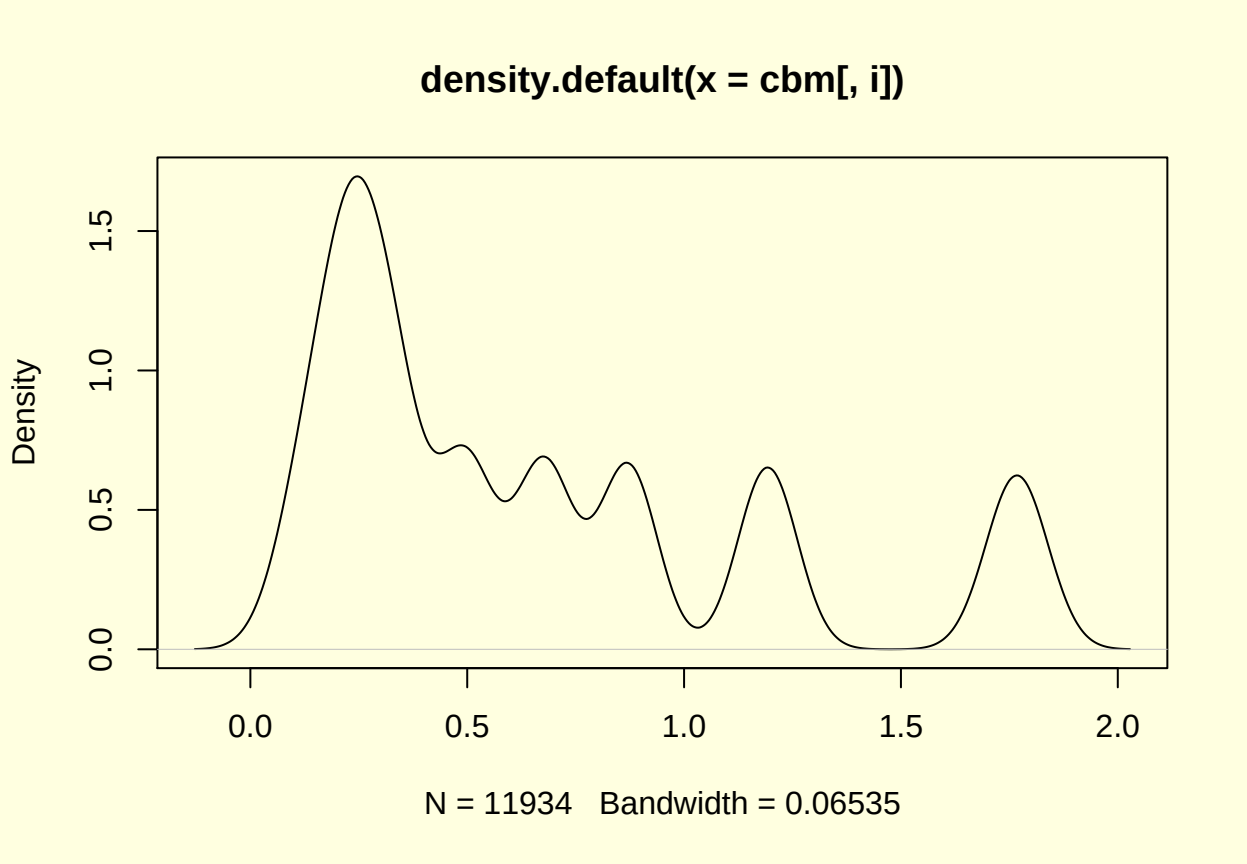

density.default(x = cbm[, i])
Density
0.0
0.2
0.4
0.6
1
2
3
4
5
N = 11934 Bandwidth = 0.1494

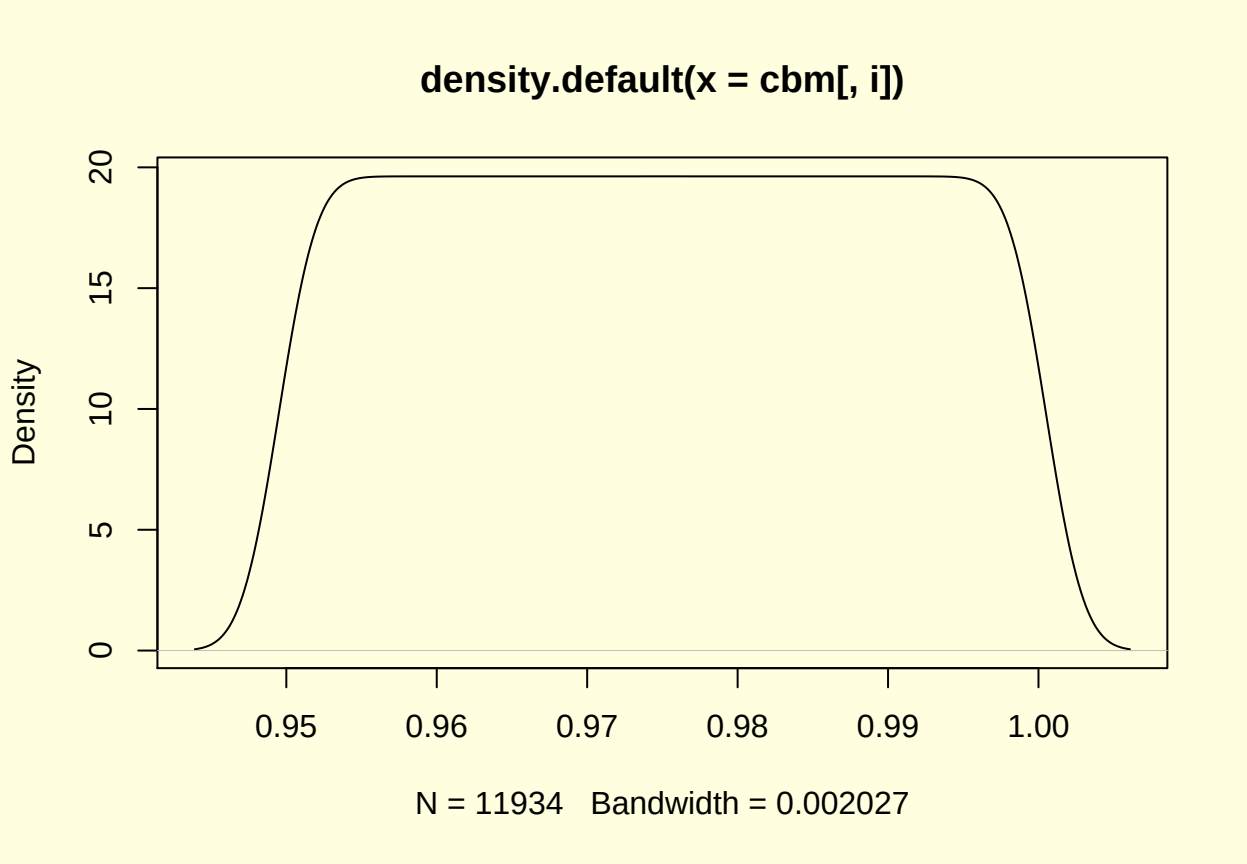

density.default(x = cbm[, i])
Density
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0.6
0.8
1.0
1.2
1.4
N = 11934 Bandwidth = 0.1374

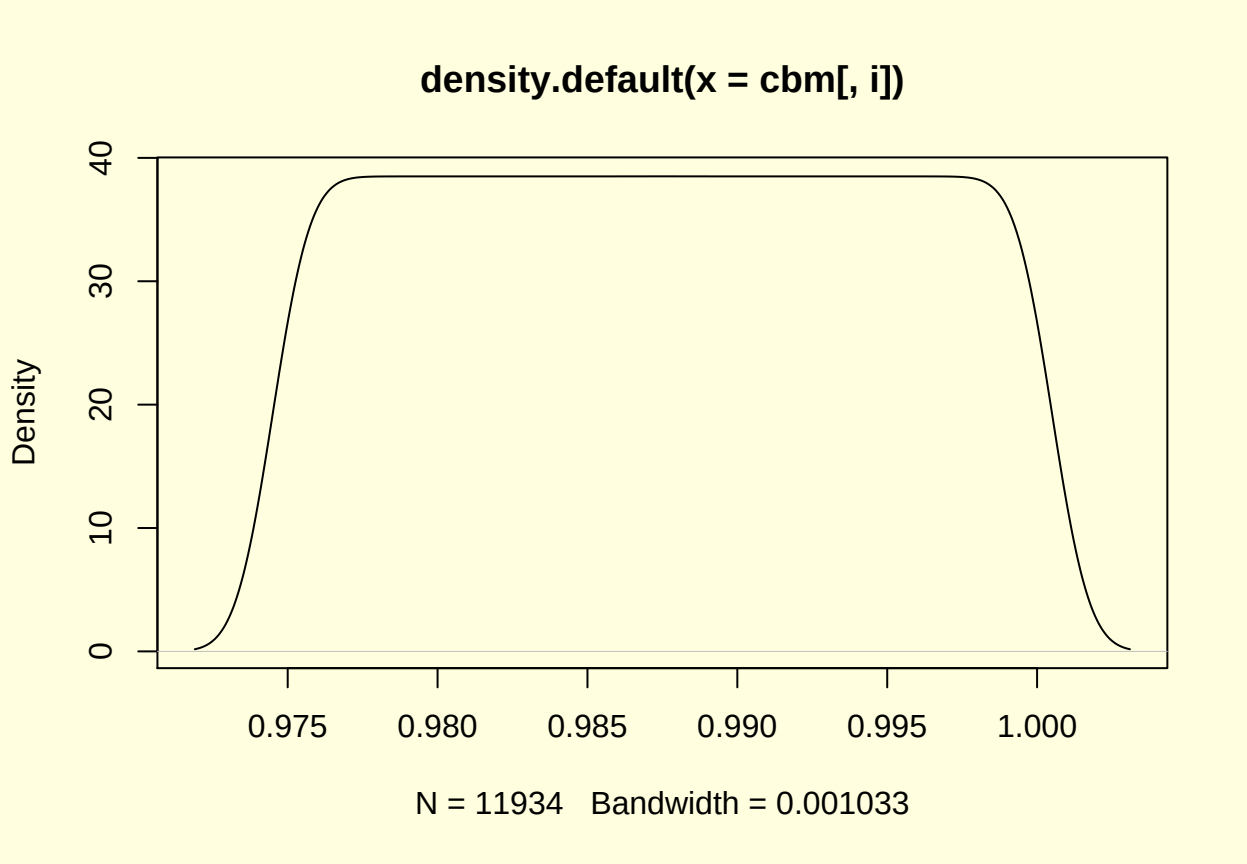

density.default(x = cbm[, i])
Density
0.00
0.05
0.10
0.15
5
10
15
20
25
N = 11934 Bandwidth = 0.7349


density.default(x = cbm[, i])
Density
0
20
40
60
80
1.02
1.03
1.04
1.05
N = 11934 Bandwidth = 0.001431


density.default(x = cbm[, i])
Density
0.000
0.010
0.020
0
20
40
60
80
100
N = 11934 Bandwidth = 3.172


density.default(x = cbm[, i])
Density
0.0
0.5
1.0
1.5
0.0
0.5
1.0
1.5
2.0
N = 11934 Bandwidth = 0.06535


density.default(x = cbm[, i])
Density
0
5
10
15
20
0.95
0.96
0.97
0.98
0.99
1.00
N = 11934 Bandwidth = 0.002027



Removing the features as found insight 2

```
cbmu<-cbm[,-c(7,9,12,17)]
colnames(cbmu)
```

```
##  [1] "Lever position (lp)"
##  [2] "Ship speed (v) [knots]"
##  [3] "Gas Turbine (GT) shaft torque (GTT) [kN m]"
##  [4] "GT rate of revolutions (GTn) [rpm]"
##  [5] "Gas Generator rate of revolutions (GGn) [rpm]"
##  [6] "Starboard Propeller Torque (Ts) [kN]"
##  [7] "Hight Pressure (HP) Turbine exit temperature (T48) [C]"
##  [8] "GT Compressor outlet air temperature (T2) [C]"
##  [9] "HP Turbine exit pressure (P48) [bar]"
## [10] "GT Compressor outlet air pressure (P2) [bar]"
## [11] "GT exhaust gas pressure (Pexh) [bar]"
## [12] "Turbine Injecton Control (TIC) [%]"
## [13] "Fuel flow (mf) [kg/s]"
## [14] "GT_Turbine_decay_state_coefficient"
```

performing **train and test data split**.

```
set.seed(1234)
ids = sample(nrow(cbmu), nrow(cbmu)*0.6)
train<-data.frame(cbmu[ids,])
test<-data.frame(cbmu[-ids,])
```

z score Scaling the data.

```r
train_withou_Y<-train[,-c(14)]
test_withou_Y<-test[,-c(14)]

train_zscale<-data.frame(scale(train_withou_Y))
test_zscale<-data.frame(scale(test_withou_Y))
head(train_zscale)
```

```
##      Lever.position..lp. Ship.speed..v...knots.
## 7452           1.5819556              1.5574706
## 8016           0.3935042              0.3970256
## 7162           0.7635404              0.7838406
## 8086          -0.3724290             -0.3766044
## 7269           0.3935042              0.3970256
## 9196           0.7635404              0.7838406
##      Gas.Turbine..GT..shaft.torque..GTT...kN.m.
## 7452                                  2.0602075
## 8016                                  0.1212794
## 7162                                  0.5374577
## 8086                                 -0.5567321
## 7269                                  0.1212597
## 9196                                  0.5379594
##      GT.rate.of.revolutions..GTn...rpm.
## 7452                          1.8483556
## 8016                          0.2298238
## 7162                          0.7091705
## 8086                         -0.7505758
## 7269                          0.2298625
## 9196                          0.7091098
##      Gas.Generator.rate.of.revolutions..GGn...rpm.
## 7452                                     1.4359502
## 8016                                     0.5555773
## 7162                                     0.8554823
## 8086                                    -0.4083156
## 7269                                     0.5537797
## 9196                                     0.8479048
##      Starboard.Propeller.Torque..Ts...kN.
## 7452                            2.0860876
## 8016                            0.1005187
## 7162                            0.5300546
## 8086                           -0.5574466
## 7269                            0.1005088
## 9196                            0.5306122
##      Hight.Pressure..HP..Turbine.exit.temperature..T48...C.
## 7452                                              1.8846280
## 8016                                              0.2386647
## 7162                                              0.4877293
## 8086                                             -0.5791260
## 7269                                              0.2795192
## 9196                                              0.5007668
##      GT.Compressor.outlet.air.temperature..T2...C.
## 7452                                     1.7760527
## 8016                                     0.2906860
## 7162                                     0.5832129
## 8086                                    -0.5689311
```

```
## 7269                                    0.3159750
## 9196                                    0.5646585
##      HP.Turbine.exit.pressure..P48...bar.
## 7452                            1.9957672
## 8016                            0.1550815
## 7162                            0.5814134
## 8086                           -0.6276012
## 7269                            0.1550815
## 9196                            0.5915423
##      GT.Compressor.outlet.air.pressure..P2...bar.
## 7452                                    1.9084597
## 8016                                    0.2027858
## 7162                                    0.6124994
## 8086                                   -0.6245025
## 7269                                    0.2145775
## 9196                                    0.6433823
##      GT.exhaust.gas.pressure..Pexh...bar. Turbine.Injecton.Control..TIC.....
## 7452                           2.07936173                          2.22139947
## 8016                           0.05970584                          0.06692786
## 7162                           0.63675038                          0.43596782
## 8086                          -0.61351279                         -0.63929452
## 7269                           0.05970584                          0.08472942
## 9196                           0.63675038                          0.44933912
##      Fuel.flow..mf...kg.s.
## 7452            2.12749875
## 8016            0.04529687
## 7162            0.40151477
## 8086           -0.63958617
## 7269            0.06104131
## 9196            0.41529115
```

```
head(test_zscale)
```

```
##   Lever.position..lp. Ship.speed..v...knots.
## 2          -1.1894658             -1.1794258
## 3          -0.7866480             -0.7914044
## 4          -0.3987070             -0.4033830
## 6           0.3695458              0.3726599
## 7           0.7407028              0.7606813
## 8           1.1442835              1.1487027
##   Gas.Turbine..GT..shaft.torque..GTT...kN.m. GT.rate.of.revolutions..GTn...rpm.
## 2                                 -0.9302193                          -0.9974135
## 3                                 -0.8659906                          -0.9837187
## 4                                 -0.5787971                          -0.7759140
## 6                                  0.1032227                           0.2067310
## 7                                  0.5191536                           0.6860442
## 8                                  1.0629571                           1.2155190
##   Gas.Generator.rate.of.revolutions..GGn...rpm.
## 2                                    -1.2772653
## 3                                    -1.0167766
## 4                                    -0.3908699
## 6                                     0.5613221
## 7                                     0.8408660
## 8                                     1.0119862
##   Starboard.Propeller.Torque..Ts...kN.
```

```
## 2                          -1.00813539
## 3                          -0.84728111
## 4                          -0.58006109
## 6                           0.08280625
## 7                           0.51202615
## 8                           1.04186378
##   Hight.Pressure..HP..Turbine.exit.temperature..T48...C.
## 2                                             -0.5884885
## 3                                             -0.7581664
## 4                                             -0.4380241
## 6                                              0.3640081
## 7                                              0.6774784
## 8                                              1.2394833
##   GT.Compressor.outlet.air.temperature..T2...C.
## 2                                    -0.9030554
## 3                                    -0.8213302
## 4                                    -0.4593074
## 6                                     0.4028263
## 7                                     0.7275358
## 8                                     1.3039272
##   HP.Turbine.exit.pressure..P48...bar.
## 2                           -0.9567478
## 3                           -0.9031865
## 4                           -0.6547728
## 6                            0.1237130
## 7                            0.5503564
## 8                            1.1164440
##   GT.Compressor.outlet.air.pressure..P2...bar.
## 2                                   -0.9544418
## 3                                   -0.8996498
## 4                                   -0.6307560
## 6                                    0.1853077
## 7                                    0.6212046
## 8                                    1.1753173
##   GT.exhaust.gas.pressure..Pexh...bar. Turbine.Injecton.Control..TIC.....
## 2                           -1.02288442                          -0.9047142
## 3                           -0.92653174                          -0.8065829
## 4                           -0.73382637                          -0.6038213
## 6                            0.03699507                           0.1086903
## 7                            0.51875847                           0.5172410
## 8                            1.00052187                           1.1856722
##   Fuel.flow..mf...kg.s.
## 2            -0.7527232
## 3            -0.8080963
## 4            -0.6123128
## 6             0.0798509
## 7             0.4753730
## 8             1.1240294
```

performing **minmax scaling** on the dataset

```
minmax<-function(x){
  return((x-min(x))/max(x)-min(x))
}
```

```r
train_minmax<-data.frame(minmax(train_withou_Y))
test_minmax<-data.frame(minmax(test_withou_Y))
```

Linear regression

Fitting linear regression models to both train data sets

```r
Y_train<-train$GT_Turbine_decay_state_coefficient
Y_test<-test$GT_Turbine_decay_state_coefficient

train_zscale_withY<-cbind(train_zscale[,],Y_train)
train_minmax_withY<-cbind(train_minmax[,],Y_train)

test_zscale_withY<-cbind(test_zscale[,],Y_test)
test_minmax_withY<-cbind(test_minmax[,],Y_test)


lmode_z<-lm (Y_train~.,data=train_zscale_withY)
lmode_minmax<-lm (Y_train~.,data=train_minmax_withY)
summary(lmode_z)
```

```
##
## Call:
## lm(formula = Y_train ~ ., data = train_zscale_withY)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.009473 -0.001369  0.000018  0.001410  0.007945
##
## Coefficients:
##                                                    Estimate Std. Error
## (Intercept)                                       9.875e-01  2.708e-05
## Lever.position..lp.                              -1.996e-02  5.411e-03
## Ship.speed..v...knots.                            5.953e-02  5.041e-03
## Gas.Turbine..GT..shaft.torque..GTT...kN.m.        2.925e-01  3.723e-03
## GT.rate.of.revolutions..GTn...rpm.                4.223e-02  1.060e-03
## Gas.Generator.rate.of.revolutions..GGn...rpm.     2.868e-02  3.700e-04
## Starboard.Propeller.Torque..Ts...kN.             -2.914e-01  2.730e-03
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C. -1.049e-01  1.508e-03
## GT.Compressor.outlet.air.temperature..T2...C.     1.768e-02  1.057e-03
## HP.Turbine.exit.pressure..P48...bar.              1.274e-01  4.739e-03
## GT.Compressor.outlet.air.pressure..P2...bar.      -2.941e-01  2.137e-03
## GT.exhaust.gas.pressure..Pexh...bar.              8.577e-03  9.601e-04
## Turbine.Injecton.Control..TIC.....               -2.092e-03  1.039e-04
## Fuel.flow..mf...kg.s.                             1.390e-01  2.185e-03
##                                                    t value Pr(>|t|)
## (Intercept)                                      36460.551  < 2e-16 ***
## Lever.position..lp.                                 -3.689 0.000227 ***
## Ship.speed..v...knots.                              11.810  < 2e-16 ***
## Gas.Turbine..GT..shaft.torque..GTT...kN.m.          78.561  < 2e-16 ***
## GT.rate.of.revolutions..GTn...rpm.                  39.817  < 2e-16 ***
## Gas.Generator.rate.of.revolutions..GGn...rpm.       77.534  < 2e-16 ***
## Starboard.Propeller.Torque..Ts...kN.              -106.747  < 2e-16 ***
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C.  -69.539  < 2e-16 ***
```

```
## GT.Compressor.outlet.air.temperature..T2...C.             16.727  < 2e-16 ***
## HP.Turbine.exit.pressure..P48...bar.                      26.875  < 2e-16 ***
## GT.Compressor.outlet.air.pressure..P2...bar.            -137.662  < 2e-16 ***
## GT.exhaust.gas.pressure..Pexh...bar.                       8.934  < 2e-16 ***
## Turbine.Injecton.Control..TIC.....                       -20.134  < 2e-16 ***
## Fuel.flow..mf...kg.s.                                     63.596  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002292 on 7146 degrees of freedom
## Multiple R-squared:  0.906,  Adjusted R-squared:  0.9058
## F-statistic:  5299 on 13 and 7146 DF,  p-value: < 2.2e-16
```

```
summary(lmode_minmax)
```

```
##
## Call:
## lm(formula = Y_train ~ ., data = train_minmax_withY)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.009473 -0.001369  0.000018  0.001410  0.007945
##
## Coefficients:
##                                                    Estimate Std. Error
## (Intercept)                                       2.102e-01  9.438e-02
## Lever.position..lp.                              -5.525e+02  1.498e+02
## Ship.speed..v...knots.                            5.587e+02  4.731e+01
## Gas.Turbine..GT..shaft.torque..GTT...kN.m.        9.596e-01  1.221e-02
## GT.rate.of.revolutions..GTn...rpm.                3.968e+00  9.966e-02
## Gas.Generator.rate.of.revolutions..GGn...rpm.     1.909e+00  2.462e-02
## Starboard.Propeller.Torque..Ts...kN.             -1.056e+02  9.889e-01
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C. -4.388e+01  6.310e-01
## GT.Compressor.outlet.air.temperature..T2...C.     1.769e+01  1.057e+00
## HP.Turbine.exit.pressure..P48...bar.              8.537e+03  3.176e+02
## GT.Compressor.outlet.air.pressure..P2...bar.     -4.007e+03  2.911e+01
## GT.exhaust.gas.pressure..Pexh...bar.              6.004e+04  6.721e+03
## Turbine.Injecton.Control..TIC.....               -6.131e+00  3.045e-01
## Fuel.flow..mf...kg.s.                             1.991e+04  3.131e+02
##                                                   t value Pr(>|t|)
## (Intercept)                                         2.227 0.025945 *
## Lever.position..lp.                                -3.689 0.000227 ***
## Ship.speed..v...knots.                             11.810  < 2e-16 ***
## Gas.Turbine..GT..shaft.torque..GTT...kN.m.         78.561  < 2e-16 ***
## GT.rate.of.revolutions..GTn...rpm.                 39.817  < 2e-16 ***
## Gas.Generator.rate.of.revolutions..GGn...rpm.      77.534  < 2e-16 ***
## Starboard.Propeller.Torque..Ts...kN.             -106.747  < 2e-16 ***
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C.  -69.539  < 2e-16 ***
## GT.Compressor.outlet.air.temperature..T2...C.      16.727  < 2e-16 ***
## HP.Turbine.exit.pressure..P48...bar.               26.875  < 2e-16 ***
## GT.Compressor.outlet.air.pressure..P2...bar.     -137.662  < 2e-16 ***
## GT.exhaust.gas.pressure..Pexh...bar.               8.934  < 2e-16 ***
## Turbine.Injecton.Control..TIC.....               -20.134  < 2e-16 ***
## Fuel.flow..mf...kg.s.                             63.596  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002292 on 7146 degrees of freedom
## Multiple R-squared:  0.906,  Adjusted R-squared:  0.9058
## F-statistic:  5299 on 13 and 7146 DF,  p-value: < 2.2e-16
```

```
lmode_z$coefficients
```

```
##                                           (Intercept)
##                                           0.987497486
##                                  Lever.position..lp.
##                                          -0.019961249
##                                Ship.speed..v...knots.
##                                           0.059532973
##             Gas.Turbine..GT..shaft.torque..GTT...kN.m.
##                                           0.292488172
##                    GT.rate.of.revolutions..GTn...rpm.
##                                           0.042225463
##          Gas.Generator.rate.of.revolutions..GGn...rpm.
##                                           0.028683249
##                   Starboard.Propeller.Torque..Ts...kN.
##                                          -0.291373873
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C.
##                                          -0.104872232
##          GT.Compressor.outlet.air.temperature..T2...C.
##                                           0.017681427
##                   HP.Turbine.exit.pressure..P48...bar.
##                                           0.127373243
##           GT.Compressor.outlet.air.pressure..P2...bar.
##                                          -0.294119771
##                   GT.exhaust.gas.pressure..Pexh...bar.
##                                           0.008577395
##                     Turbine.Injecton.Control..TIC.....
##                                          -0.002091655
##                                Fuel.flow..mf...kg.s.
##                                           0.138984688
```

predicting on **test data** using previous models.

```
lmode_z_predictions<-predict(lmode_z, newdata=test_zscale_withY)
lmode_minmax_predictions<-predict(lmode_minmax, newdata=test_minmax_withY)
```

Root mean square error function

```
rmse<-function(x){

testerror<- (Y_test-x)
testerror_sq <- testerror ** 2
rmse<-sqrt(mean(testerror_sq))
rmse

}

rmse(lmode_z_predictions)
```

```
## [1] 0.002299194
```

```
rmse(lmode_minmax_predictions)
```

```
## [1] 0.00230002
```

we can clearly evident r-square value is 0.90 but the **feature significance** post **close values**.To handel this we **perform ridge regeressoion as handels multicolinearity**.

Ridge regression

```
par(mfrow=c(1,1),bg="lightyellow")


library(glmnet)
```

```
## Warning: package 'glmnet' was built under R version 3.6.2
```

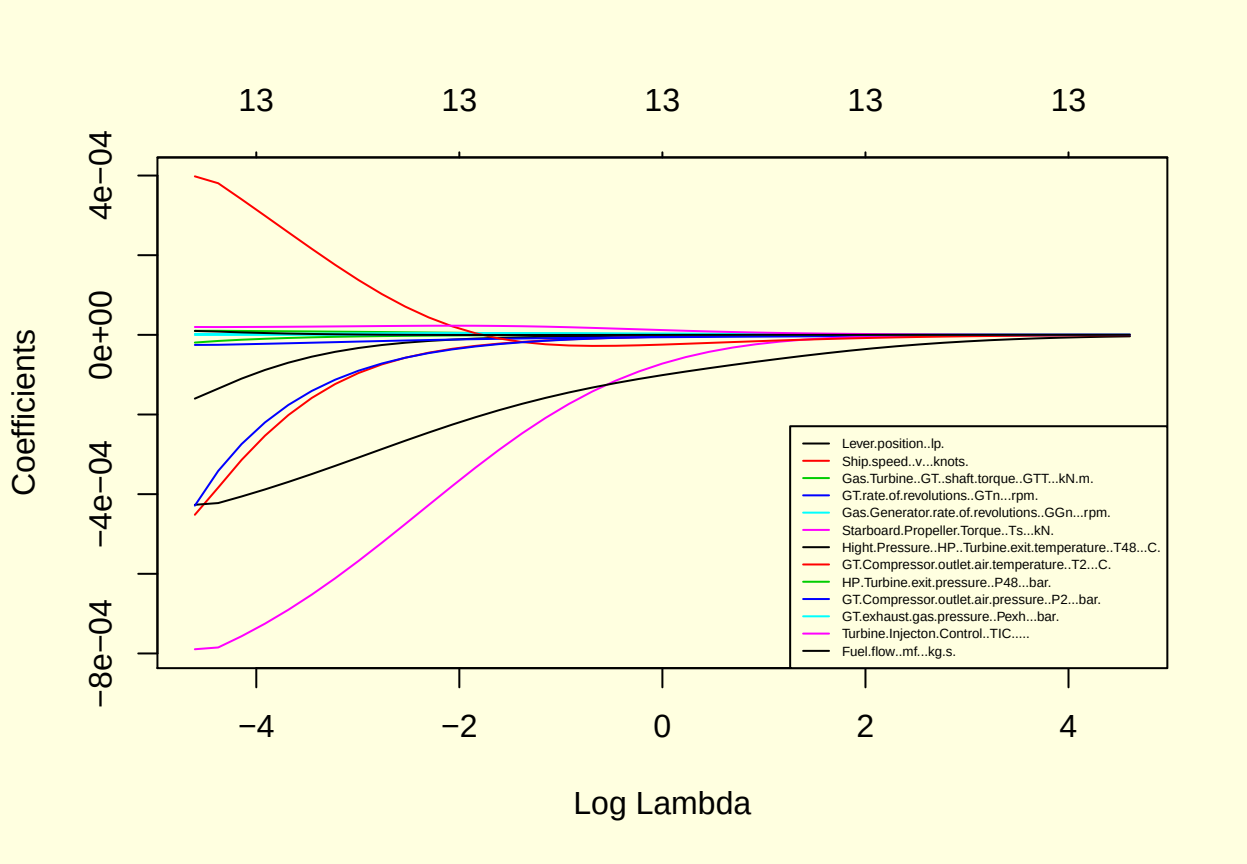
```
## Loading required package: Matrix
```

```
## Loaded glmnet 3.0-2
```

```
x_reg<-data.matrix(train_withou_Y)
y_reg<-Y_train
lambda_seq <- 10^seq(2, -2, by = -.1)

ridge_model <- glmnet(x_reg, y_reg, alpha = 0, lambda  = lambda_seq,standardize = FALSE)
summary(ridge_model)
```

```
##           Length Class     Mode
## a0         41    -none-    numeric
## beta      533    dgCMatrix S4
## df         41    -none-    numeric
## dim         2    -none-    numeric
## lambda     41    -none-    numeric
## dev.ratio  41    -none-    numeric
## nulldev     1    -none-    numeric
## npasses     1    -none-    numeric
## jerr        1    -none-    numeric
## offset      1    -none-    logical
## call        6    -none-    call
## nobs        1    -none-    numeric
```

```
plot(ridge_model, xvar = "lambda")
legend("bottomright", lwd = 1, col = 1:6, legend = colnames(x_reg), cex = .45)
```

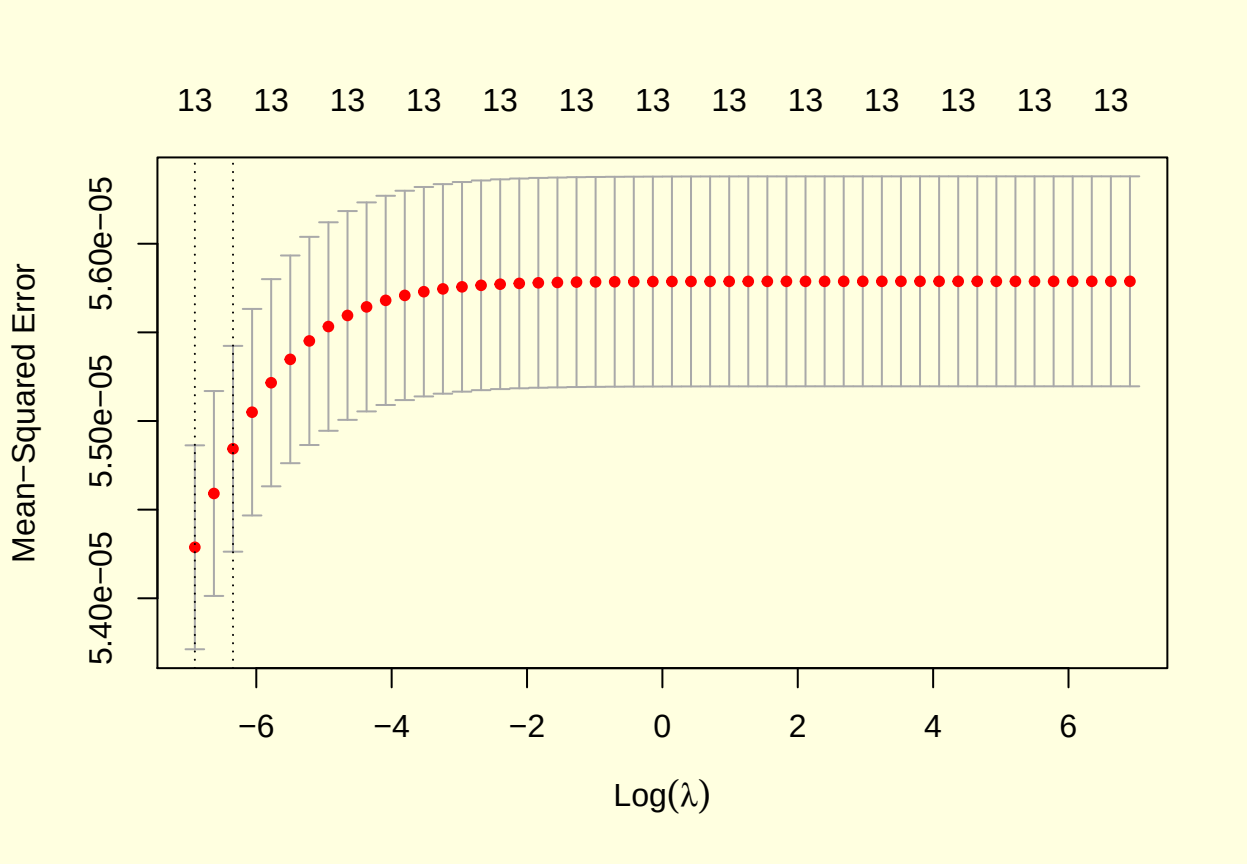


Finding the **optimal lambda**

```
par(mfrow=c(1,1),bg="lightyellow")

lambdas_to_try <- 10^seq(-3, 3, length.out = 50)
ridge_cv <- cv.glmnet(x_reg, y_reg, alpha = 0,lambda = lambdas_to_try ,nfolds = 100)
best_lambda<-ridge_cv$lambda.min
best_lambda
```

```
## [1] 0.001
```

```
plot(ridge_cv)
```



Final ridge regression model with optimal lambda value

```
x_reg<-data.matrix(train_withou_Y)
y_reg<-Y_train
lambda_seq <- 10^seq(2, -2, by = -.1)

final_ridge_model <- glmnet(x_reg, y_reg, alpha = 0, lambda  = best_lambda ,standardize = FALSE)
coef(final_ridge_model)
```

```
## 14 x 1 sparse Matrix of class "dgCMatrix"
##                                                                 s0
## (Intercept)                                          9.070571e-01
## Lever.position..lp.                                 -5.175049e-04
## Ship.speed..v...knots.                              -1.076006e-03
## Gas.Turbine..GT..shaft.torque..GTT...kN.m.           1.229369e-05
## GT.rate.of.revolutions..GTn...rpm.                  -2.593418e-05
## Gas.Generator.rate.of.revolutions..GGn...rpm.        3.373632e-06
## Starboard.Propeller.Torque..Ts...kN.                -9.319132e-04
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C. -5.159662e-04
## GT.Compressor.outlet.air.temperature..T2...C.        6.604703e-04
## HP.Turbine.exit.pressure..P48...bar.                -1.522194e-04
## GT.Compressor.outlet.air.pressure..P2...bar.         -3.568094e-03
## GT.exhaust.gas.pressure..Pexh...bar.                -2.459583e-06
## Turbine.Injecton.Control..TIC.....                   2.357019e-05
## Fuel.flow..mf...kg.s.                                9.470750e-05
```

Insight3 : Rmse value improved slightly and best TIC,Fuel FLow ,GG rate,GT compressor temp and GT

shaft. Top3 features follows GT compressor temp,GG rate,TIC

```
y_predicted <- predict(final_ridge_model , s = best_lambda, newx = data.matrix(test_withou_Y))

rmse(y_predicted)
```

```
## [1] 0.00551595
```

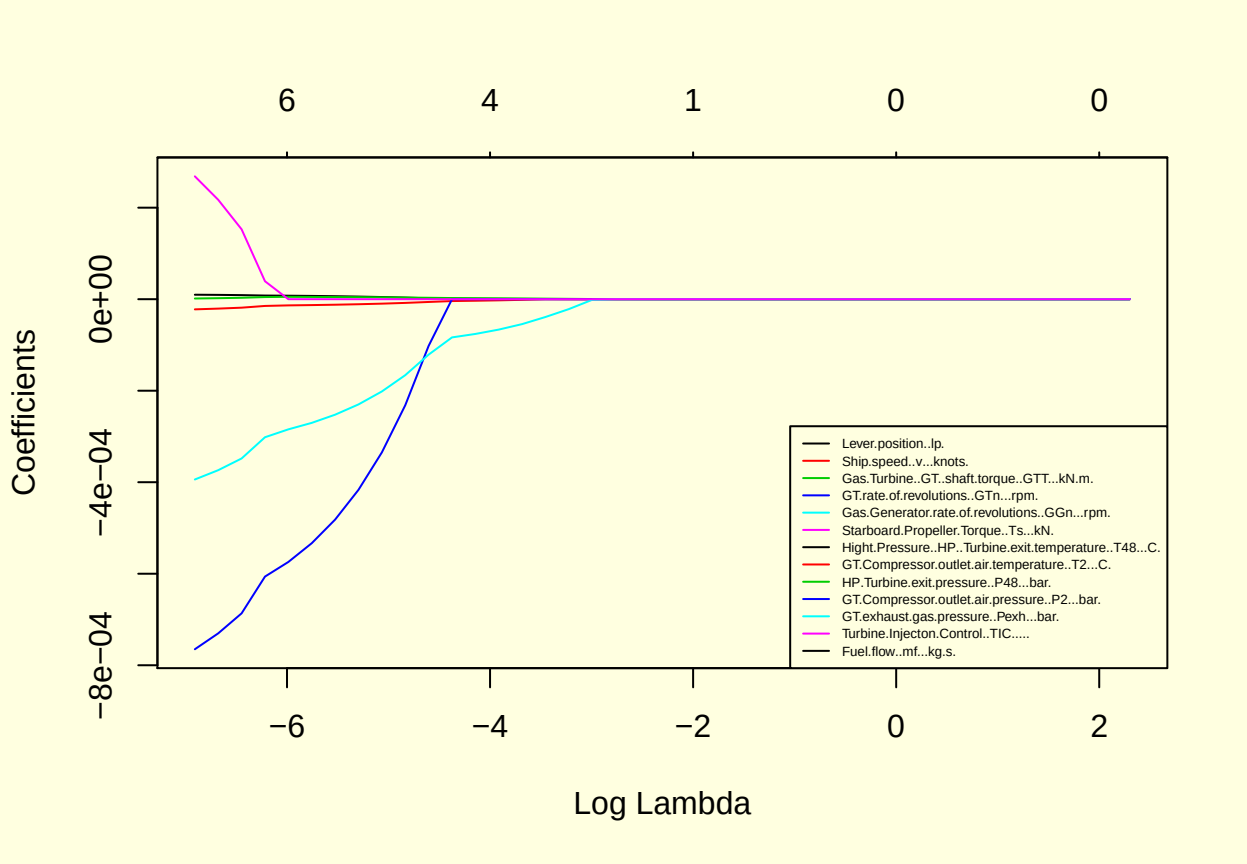
#Lasso Regression

next we will perform lasso regression.

```
par(mfrow=c(1,1),bg="lightyellow")

library(glmnet)

x_reg<-data.matrix(train_withou_Y)
y_reg<-Y_train
lambda_seq <- 10^seq(1, -3, by = -.1)

lasso_model <- glmnet(x_reg, y_reg, alpha = 1, lambda  = lambda_seq,standardize = FALSE)
summary(lasso_model)
```

```
##           Length Class     Mode
## a0         41    -none-    numeric
## beta      533    dgCMatrix S4
## df         41    -none-    numeric
## dim         2    -none-    numeric
## lambda     41    -none-    numeric
## dev.ratio  41    -none-    numeric
## nulldev     1    -none-    numeric
## npasses     1    -none-    numeric
## jerr        1    -none-    numeric
## offset      1    -none-    logical
## call        6    -none-    call
## nobs        1    -none-    numeric
```

```
plot(lasso_model, xvar = "lambda")
legend("bottomright", lwd = 1, col = 1:6, legend = colnames(x_reg), cex = .45)
```



Finding the **optimal lambda** for **losso regression** using **cross validation.**

```
par(mfrow=c(1,1),bg="lightyellow")
set.seed(12345)

lambdas_to_try <- 10^seq(-6, 1, length.out = 50)
lasso_cv <- cv.glmnet(x_reg, y_reg, alpha = 1,lambda = lambdas_to_try ,nfolds = 100)
```

```
## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned
```

```
## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
```

```
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
```

```
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
```

```
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
```

```
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
```

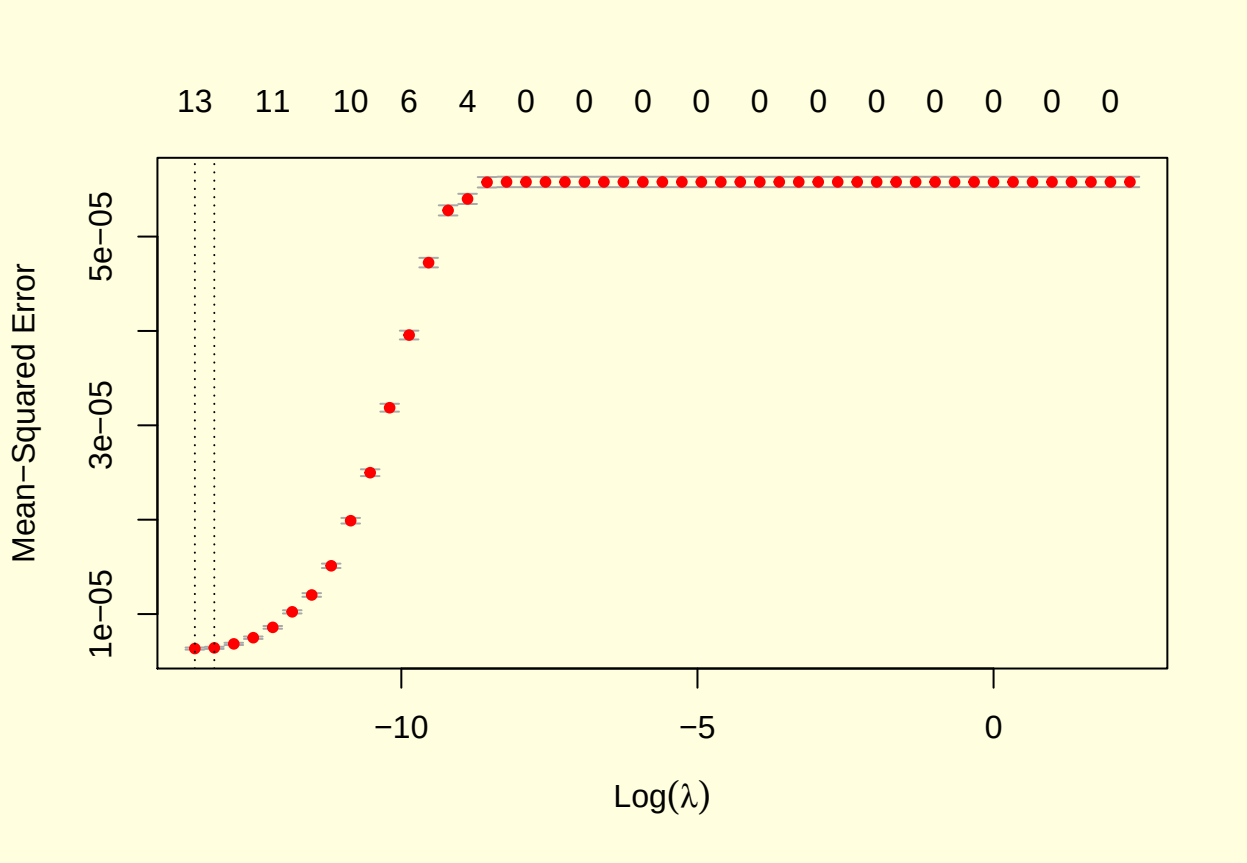
```
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned
```

```
## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -49); Convergence for 49th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned

## Warning: from glmnet Fortran code (error code -50); Convergence for 50th lambda
## value not reached after maxit=100000 iterations; solutions for larger lambdas
## returned
```

```
best_lambda_lasso<-ridge_cv$lambda.min
best_lambda_lasso
```

```
## [1] 0.001
```

```
plot(lasso_cv )
```



Final lasoo regression model with optimal lambda value

```
set.seed(12345)

x_reg<-data.matrix(train_withou_Y)
y_reg<-Y_train

final_lasso_model <- glmnet(x_reg, y_reg, alpha = 1, lambda  = best_lambda_lasso ,standardize = FALSE)
coef(final_lasso_model)
```

```
## 14 x 1 sparse Matrix of class "dgCMatrix"
##                                                                      s0
## (Intercept)                                                1.046939e+00
## Lever.position..lp.                                        .
## Ship.speed..v...knots.                                     .
## Gas.Turbine..GT..shaft.torque..GTT...kN.m.                 9.786350e-06
## GT.rate.of.revolutions..GTn...rpm.                        -2.204397e-05
## Gas.Generator.rate.of.revolutions..GGn...rpm.              1.571053e-06
## Starboard.Propeller.Torque..Ts...kN.                      -7.632473e-04
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C. -3.918647e-04
## GT.Compressor.outlet.air.temperature..T2...C.              2.628906e-04
## HP.Turbine.exit.pressure..P48...bar.                       .
## GT.Compressor.outlet.air.pressure..P2...bar.               .
## GT.exhaust.gas.pressure..Pexh...bar.                       .
## Turbine.Injecton.Control..TIC.....                         .
```

```
## Fuel.flow..mf...kg.s.                                        .
```

Insight4 : Rmse improved very littel and the features significance as follows GG rate,GT shaft torque,GT compressor T2

```
set.seed(12345)

y_predicted_lasso <- predict(final_lasso_model , s = best_lambda_lasso, newx = data.matrix(test_withou_

rmse(y_predicted_lasso)
```

```
## [1] 0.006061261
```

Random Forest

we will perform ensemble methods on the same dataset.

```
set.seed(12345)
library(randomForest)
```

```
## Warning: package 'randomForest' was built under R version 3.6.2
```

```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
rftree<-randomForest(GT_Turbine_decay_state_coefficient~.,data=train,mtry=10,ntree=100,importance=T)
rftree$importance
```

```
##                                                          %IncMSE
## Lever.position..lp.                                   5.019978e-07
## Ship.speed..v...knots.                                1.189112e-06
## Gas.Turbine..GT..shaft.torque..GTT...kN.m.            3.445186e-05
## GT.rate.of.revolutions..GTn...rpm.                    7.906223e-05
## Gas.Generator.rate.of.revolutions..GGn...rpm.         6.865953e-05
## Starboard.Propeller.Torque..Ts...kN.                  1.089972e-05
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C. 4.558177e-06
## GT.Compressor.outlet.air.temperature..T2...C.         8.962702e-06
## HP.Turbine.exit.pressure..P48...bar.                  6.790741e-06
## GT.Compressor.outlet.air.pressure..P2...bar.          1.055039e-04
## GT.exhaust.gas.pressure..Pexh...bar.                  2.869298e-05
## Turbine.Injecton.Control..TIC.....                    1.233223e-05
## Fuel.flow..mf...kg.s.                                 1.293143e-05
##                                                      IncNodePurity
## Lever.position..lp.                                     0.0004284791
## Ship.speed..v...knots.                                  0.0005177402
## Gas.Turbine..GT..shaft.torque..GTT...kN.m.              0.0408322698
## GT.rate.of.revolutions..GTn...rpm.                      0.0405940700
## Gas.Generator.rate.of.revolutions..GGn...rpm.           0.0560852780
## Starboard.Propeller.Torque..Ts...kN.                    0.0111886765
## Hight.Pressure..HP..Turbine.exit.temperature..T48...C.  0.0067983515
## GT.Compressor.outlet.air.temperature..T2...C.           0.0139185269
## HP.Turbine.exit.pressure..P48...bar.                    0.0070084898
## GT.Compressor.outlet.air.pressure..P2...bar.            0.1526067270
## GT.exhaust.gas.pressure..Pexh...bar.                    0.0231309284
## Turbine.Injecton.Control..TIC.....                      0.0255984401
## Fuel.flow..mf...kg.s.                                   0.0206211418
```

Finding best mtry and ntree for the random forest regressor.

```
set.seed(12345)

 rmse_updated<-function(x) {
  h=x
  predictvalues<-predict(h,newdata=test)
  mean((predictvalues-test$GT_Turbine_decay_state_coefficient)^2)
}

 mtrail<- function(x){
h=x
carrftree<-randomForest(GT_Turbine_decay_state_coefficient~.,data=train,mtry=h,importance=T)
 rmse_updated(carrftree)

 }

 for (i in 1:10){
  mtrail(i)
  print(mtrail(i))

 }
```

```
## [1] 8.020788e-06
## [1] 1.684727e-06
## [1] 1.100441e-06
## [1] 9.23354e-07
## [1] 8.254999e-07
## [1] 7.635149e-07
## [1] 7.2256e-07
## [1] 7.005129e-07
## [1] 6.892878e-07
## [1] 6.6291e-07
```

```
 ntree<- function(x){
h=x
carrftree<-randomForest(GT_Turbine_decay_state_coefficient~.,data=train,mtry=3,ntree=h,importance=T)
 rmse_updated(carrftree)

 }


 for (i in 1:10){
  ntree(i)
  print(ntree(i))

 }
```

```
## [1] 4.803744e-06
## [1] 2.847083e-06
## [1] 2.239733e-06
## [1] 2.167531e-06
## [1] 1.997725e-06
## [1] 1.708186e-06
## [1] 1.740071e-06
## [1] 1.612491e-06
```

```
## [1] 1.508718e-06
## [1] 1.555065e-06
```

final random tree

```
set.seed(12345)
final_rftree<-randomForest(GT_Turbine_decay_state_coefficient~.,data=train,mtry=3,ntree=4,importance=T)
rmse_updated(final_rftree)
```

```
## [1] 1.88992e-06
```

Insight5 : Rmse lowered and the top 3 features significance as follows GG rate,GT rate,GT compressor T2

Extra Tree regression

```
options(java.parameters = "-Xmx4g")

library(extraTrees)
```

```
## Warning: package 'extraTrees' was built under R version 3.6.2
```

```
## Loading required package: rJava
```

```
et <- extraTrees(data.matrix(train_withou_Y), Y_train, nodesize=3, mtry=3, numRandomCuts=3)

yhat <- predict(et, data.matrix(test_withou_Y))
rmse(yhat)
```

```
## [1] 0.0007042053
```

Final Insight : Random forest is the best models yields low Root mean squared error and the best features are subjected to models but accross the GGrate ,GT compressor T2. Lasso model is better at elimanation of features and poised GGrate ,GT compressor T2 and GT rate.