

# PM10

Nicole

20220226

```
library(MPV)
```

```
## Loading required package: KernSmooth
## KernSmooth 2.23 loaded
## Copyright M. P. Wand 1997-2009
```

```
library(olsrr)
```

```
## Warning: package 'olsrr' was built under R version 4.0.3
##
## Attaching package: 'olsrr'
## The following object is masked from 'package:MPV':
##
##      cement
## The following object is masked from 'package:datasets':
##
##      rivers
```

```
library(glmnet)
```

```
## Warning: package 'glmnet' was built under R version 4.0.3
## Loading required package: Matrix
## Loaded glmnet 4.0-2
```

```
library(ggpubr)
```

```
## Warning: package 'ggpubr' was built under R version 4.0.5
## Loading required package: ggplot2
```

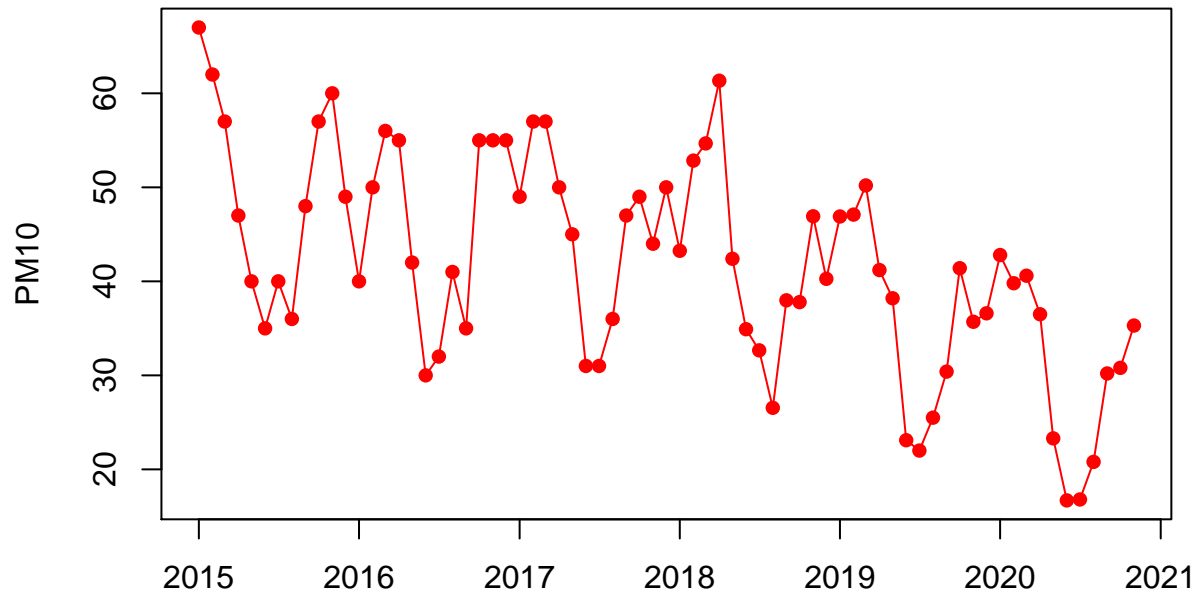
```
library(psych)
```

```
## Warning: package 'psych' was built under R version 4.0.3
##
## Attaching package: 'psych'
## The following objects are masked from 'package:ggplot2':
##
##      %+%, alpha
```

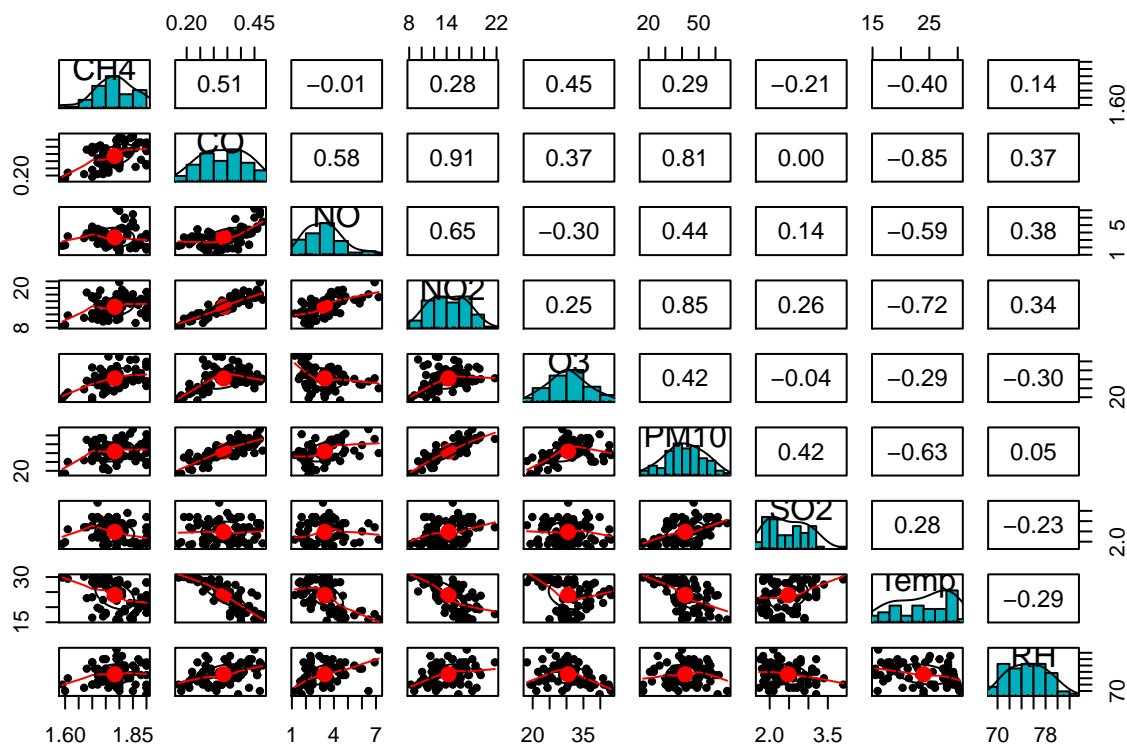
```
data=read.csv('C:/Users/user/OneDrive/data.csv',header=T)[,c(1,2,3,4,5,6,8,10,11)]
time = seq.Date(from = as.Date("2015/01/01",format = "%Y/%m/%d"),
                by = "month", length.out = nrow(data))
```

```
data = cbind(time,data)
plot(x=data$time,y=data$PM10,type='l',xlab = '',ylab='PM10',
     main='Variations of monthly PM10 at the monitoring stations',col='red')
points(x=data$time,y=data$PM10,pch=16,col='red')
```

## Variations of monthly PM10 at the monitoring stations



```
data=data[, -1]
pairs.panels(data,method = "pearson",hist.col = "#00AFBB",density = TRUE,ellipses = TRUE)
```



```
std_data=as.data.frame(scale(data))
```

```
set.seed(123)
```

```
train.index = sample(x=1:nrow(std_data), size=ceiling(0.8*nrow(std_data)))
```

```
train = std_data[train.index, ]
```

```
test = std_data[-train.index, ]
```

```
Model=lm(PM10 ~ .,data=train)
```

```
summary(Model)
```

```
##
```

```
## Call:
```

```
## lm(formula = PM10 ~ ., data = train)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -0.68490 -0.24219  0.03736  0.22593  0.76705
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -0.03673    0.04657  -0.789  0.434194
```

```
## CH4          -0.04875    0.06360  -0.767  0.447074
```

```
## CO           0.41495    0.20084   2.066  0.044248 *
```

```
## NO          -0.25361    0.08456  -2.999  0.004282 **
```

```
## NO2          0.17051    0.14819   1.151  0.255614
```

```
## O3           0.01128    0.07852   0.144  0.886374
```

```
## S02          0.49183    0.07597    6.474 4.71e-08 ***
## Temp        -0.48708    0.13516   -3.604 0.000743 ***
## RH          -0.07970    0.06531   -1.220 0.228315
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3454 on 48 degrees of freedom
## Multiple R-squared:  0.8959, Adjusted R-squared:  0.8786
## F-statistic: 51.65 on 8 and 48 DF,  p-value: < 2.2e-16
```

```
ols_vif_tol(Model)
```

```
## Variables Tolerance VIF
## 1 CH4 0.47480279 2.106138
## 2 CO 0.05033268 19.867807
## 3 NO 0.27037042 3.698630
## 4 NO2 0.09662854 10.348909
## 5 O3 0.33272173 3.005515
## 6 S02 0.35481048 2.818406
## 7 Temp 0.10696173 9.349138
## 8 RH 0.49674409 2.013109
```

```
X=as.matrix(train[,-6])
lambda=eigen(t(X)%*%X)$values
k=max(lambda)/min(lambda)
k
```

```
## [1] 105.7372
```

```
k1=ols_step_best_subset(Model)
k1
```

```
## Best Subsets Regression
## -----
## Model Index Predictors
## -----
## 1 NO2
## 2 CO S02
## 3 O3 S02 Temp
## 4 CO NO S02 Temp
## 5 CO NO NO2 S02 Temp
## 6 CO NO NO2 S02 Temp RH
## 7 CH4 CO NO NO2 S02 Temp RH
## 8 CH4 CO NO NO2 O3 S02 Temp RH
## -----
```

```
## Subsets Regression Summary
## -----
## Model R-Square Adj. R-Square Pred R-Square C(p) AIC SBIC SBC MSE
## -----
## 1 0.7158 0.7106 0.692 78.0870 94.0284 -70.6302 100.1575 16.2042
## 2 0.8217 0.8151 0.8011 31.2435 69.4563 -94.3327 77.6285 10.3583
## 3 0.8527 0.8444 0.8308 18.9336 60.5606 -102.5378 70.7758 8.7206
## 4 0.8865 0.8778 0.8607 5.3358 47.6920 -113.1916 59.9504 6.8500
## 5 0.8905 0.8798 0.8596 5.4870 47.6421 -112.6021 61.9435 6.7402
## 6 0.8946 0.8820 0.8589 5.5892 47.4581 -111.9374 63.8025 6.6192
```

```
##      7      0.8959      0.8810      0.8549      7.0206      48.7872      -110.0499      67.1747      6.6780
##      8      0.8959      0.8786      0.8498      9.0000      50.7627      -107.6916      71.1932      6.8172
## -----
## AIC: Akaike Information Criteria
## SBIC: Sawa's Bayesian Information Criteria
## SBC: Schwarz Bayesian Criteria
## MSEP: Estimated error of prediction, assuming multivariate normality
## FPE: Final Prediction Error
## HSP: Hocking's Sp
## APC: Amemiya Prediction Criteria

cat("Predictors: ", k1$predictors[k1$cp == min(k1$cp)])

## Predictors:  CO NO SO2 Temp
modell1=lm(PM10~CO+NO+SO2+Temp+RH, train)
summary(modell1)

##
## Call:
## lm(formula = PM10 ~ CO + NO + SO2 + Temp + RH, data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.60257 -0.25671 -0.02359  0.22719  0.74995
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.04004     0.04642  -0.863  0.392397
## CO           0.49049     0.11984   4.093  0.000152 ***
## NO          -0.23248     0.06523  -3.564  0.000804 ***
## SO2          0.55448     0.06429   8.624  1.56e-11 ***
## Temp        -0.52894     0.13088  -4.041  0.000180 ***
## RH          -0.06765     0.05792  -1.168  0.248245
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3453 on 51 degrees of freedom
## Multiple R-squared:  0.8895, Adjusted R-squared:  0.8786
## F-statistic: 82.09 on 5 and 51 DF,  p-value: < 2.2e-16

full=lm(PM10 ~ ., data = train)
null=lm(PM10 ~ 1, data = train)
forward.lm=step(null, scope = list(lower=null,upper=full), direction="forward",test="F")

## Start:  AIC=-0.03
## PM10 ~ 1
##
##      Df Sum of Sq  RSS      AIC  F value    Pr(>F)
## + NO2    1    39.371 15.635 -69.731 138.4952 < 2.2e-16 ***
## + CO     1    35.743 19.263 -57.836 102.0516 3.874e-14 ***
## + Temp   1    24.631 30.376 -31.876  44.5970 1.269e-08 ***
## + O3     1    10.649 44.357 -10.294  13.2040 0.0006153 ***
## + NO     1     9.772 45.234  -9.178  11.8820 0.0010935 **
## + SO2    1     8.013 46.994  -7.003   9.3776 0.0033960 **
## + CH4    1     5.873 49.134  -4.465   6.5739 0.0131109 *
```

```

## <none>          55.006 -0.029
## + RH      1      0.422 54.585   1.532   0.4248 0.5172585
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: AIC=-69.73
## PM10 ~ NO2
##
##      Df Sum of Sq    RSS      AIC F value    Pr(>F)
## + RH      1    2.90741 12.728 -79.458 12.3351 0.0009067 ***
## + O3      1    2.01991 13.615 -75.615  8.0111 0.0065159 **
## + SO2     1    1.96552 13.670 -75.388  7.7644 0.0073374 **
## + NO      1    1.23308 14.402 -72.413  4.6233 0.0360351 *
## + CO      1    0.64841 14.987 -70.145  2.3363 0.1322240
## <none>          15.635 -69.731
## + CH4     1    0.34784 15.287 -69.013  1.2287 0.2725754
## + Temp    1    0.18574 15.450 -68.412  0.6492 0.4239308
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: AIC=-79.46
## PM10 ~ NO2 + RH
##
##      Df Sum of Sq    RSS      AIC F value    Pr(>F)
## + CO      1    1.38012 11.348 -84.000  6.4459 0.01409 *
## + SO2     1    0.68805 12.040 -80.625  3.0288 0.08760 .
## + O3      1    0.68129 12.047 -80.593  2.9974 0.08922 .
## + NO      1    0.51364 12.214 -79.805  2.2288 0.14139
## + Temp    1    0.44027 12.288 -79.464  1.8990 0.17398
## <none>          12.728 -79.458
## + CH4     1    0.41751 12.310 -79.359  1.7975 0.18573
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: AIC=-84
## PM10 ~ NO2 + RH + CO
##
##      Df Sum of Sq    RSS      AIC F value    Pr(>F)
## + SO2     1    3.15399  8.1938 -100.561 20.0161 4.206e-05 ***
## + NO      1    0.48851 10.8593  -84.508  2.3393  0.1322
## <none>          11.3478  -84.000
## + O3      1    0.10505 11.2427  -82.530  0.4859  0.4889
## + Temp    1    0.07683 11.2710  -82.387  0.3545  0.5542
## + CH4     1    0.00519 11.3426  -82.026  0.0238  0.8780
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: AIC=-100.56
## PM10 ~ NO2 + RH + CO + SO2
##
##      Df Sum of Sq    RSS      AIC F value    Pr(>F)
## + Temp    1    0.69912  7.4947 -103.645  4.7574 0.03381 *
## + NO      1    0.59239  7.6014 -102.839  3.9745 0.05156 .
## <none>          8.1938 -100.561

```

```

## + O3      1    0.28084 7.9130 -100.549  1.8101 0.18445
## + CH4     1    0.07247 8.1213  -99.068  0.4551 0.50298
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: AIC=-103.64
## PM10 ~ NO2 + RH + CO + SO2 + Temp
##
##           Df Sum of Sq    RSS      AIC F value    Pr(>F)
## + NO      1    1.69924 5.7954 -116.30 14.6601 0.0003597 ***
## + O3      1    0.66592 6.8287 -106.95  4.8759 0.0318489 *
## <none>                    7.4947 -103.64
## + CH4     1    0.00055 7.4941 -101.65  0.0037 0.9518673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: AIC=-116.3
## PM10 ~ NO2 + RH + CO + SO2 + Temp + NO
##
##           Df Sum of Sq    RSS      AIC F value Pr(>F)
## <none>                    5.7954 -116.30
## + CH4     1    0.067809 5.7276 -114.97  0.5801 0.4499
## + O3      1    0.000176 5.7953 -114.30  0.0015 0.9694
model2=lm(PM10~CO+NO+NO2+SO2+Temp+RH, train)
summary(model2)

##
## Call:
## lm(formula = PM10 ~ CO + NO + NO2 + SO2 + Temp + RH, data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.68895 -0.24384  0.00912  0.20457  0.79337
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.03949    0.04577  -0.863 0.392395
## CO           0.33353    0.15495   2.152 0.036212 *
## NO          -0.24997    0.06529  -3.829 0.000360 ***
## NO2          0.21218    0.13549   1.566 0.123655
## SO2          0.49295    0.07458   6.609 2.45e-08 ***
## Temp        -0.51134    0.12954  -3.947 0.000247 ***
## RH          -0.08066    0.05772  -1.397 0.168438
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3405 on 50 degrees of freedom
## Multiple R-squared:  0.8946, Adjusted R-squared:  0.882
## F-statistic: 70.76 on 6 and 50 DF,  p-value: < 2.2e-16
PRESS(model1)

## [1] 7.747941

```

```
summary(model2)
```

```
##
## Call:
## lm(formula = PM10 ~ CO + NO + NO2 + SO2 + Temp + RH, data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.68895 -0.24384  0.00912  0.20457  0.79337
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.03949    0.04577  -0.863 0.392395
## CO           0.33353    0.15495   2.152 0.036212 *
## NO          -0.24997    0.06529  -3.829 0.000360 ***
## NO2          0.21218    0.13549   1.566 0.123655
## SO2          0.49295    0.07458   6.609 2.45e-08 ***
## Temp        -0.51134    0.12954  -3.947 0.000247 ***
## RH          -0.08066    0.05772  -1.397 0.168438
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3405 on 50 degrees of freedom
## Multiple R-squared:  0.8946, Adjusted R-squared:  0.882
## F-statistic: 70.76 on 6 and 50 DF,  p-value: < 2.2e-16
```

```
M1.anova = anova(model1)
```

```
M1.anova
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: PM10
```

```
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## CO           1 35.743  35.743 299.8337 < 2.2e-16 ***
## NO           1  0.180   0.180   1.5141  0.2242
## SO2          1 10.151  10.151  85.1529 1.856e-12 ***
## Temp         1  2.690   2.690  22.5621 1.695e-05 ***
## RH           1  0.163   0.163   1.3642  0.2482
## Residuals   51  6.080   0.119
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sst1 = sum(M1.anova$'Sum Sq')
```

```
pred.r.squared1 = 1 - PRESS(model1)/(sst1)
```

```
pred.r.squared1
```

```
## [1] 0.859145
```

```
ols_vif_tol(model1)
```

```
##   Variables Tolerance    VIF
## 1          CO 0.1412880 7.077740
## 2          NO 0.4540747 2.202281
## 3          SO2 0.4950935 2.019821
## 4         Temp 0.1140078 8.771332
## 5          RH 0.6312225 1.584227
```



```
M2.anova = anova(model2)
sst2 = sum(M2.anova$'Sum Sq')
pred.r.squared2 = 1 - PRESS(model2)/(sst2)
pred.r.squared2
```

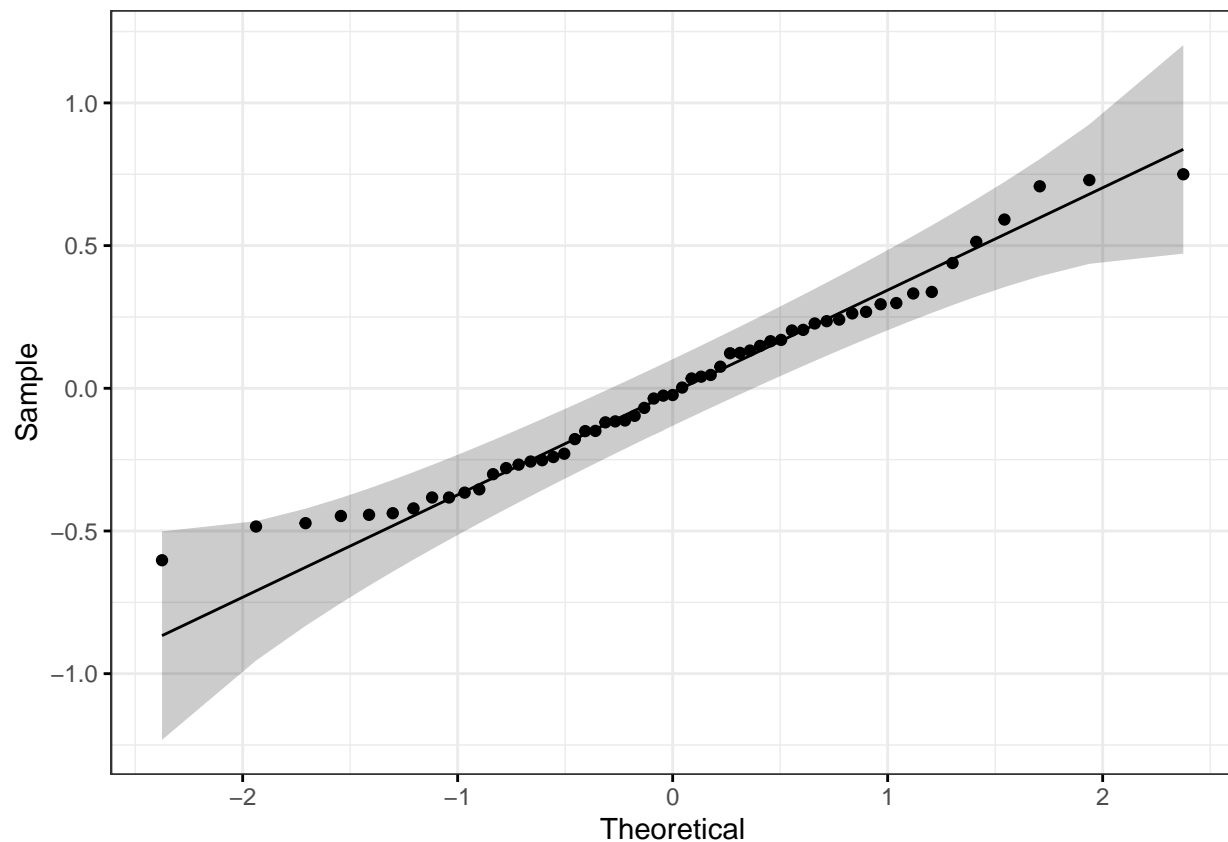
```
## [1] 0.8589037
```

```
ols_vif_tol(model2)
```

```
## Variables Tolerance VIF
## 1 CO 0.08217479 12.169182
## 2 NO 0.44077781 2.268717
## 3 NO2 0.11233621 8.901849
## 4 SO2 0.35769675 2.795664
## 5 Temp 0.11314985 8.837837
## 6 RH 0.61815549 1.617716
```

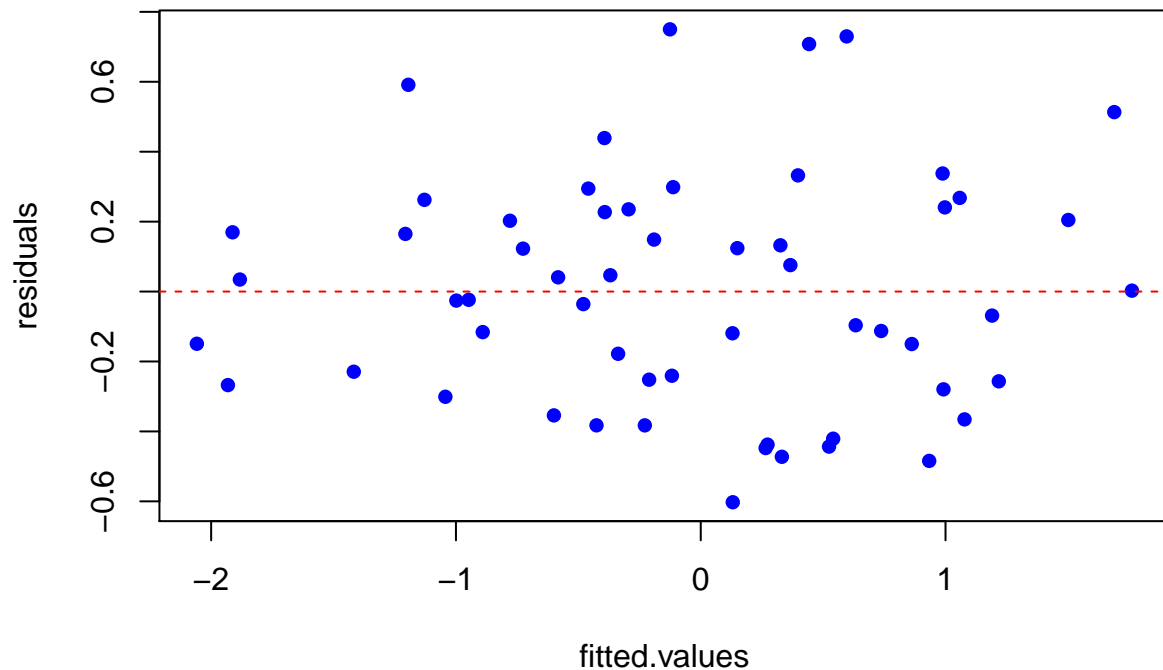
```
R.stdres=rstudent(model1)
```

```
ggqqplot(model1$residuals,ggtheme = theme_bw())
```



```
#dwtest(model1)
```

```
plot(x=model1$fitted.values, y=model1$residuals, xlab='fitted.values', ylab='residuals', pch=16, col='blue')
abline(0,0,col="red",lty=2)
```

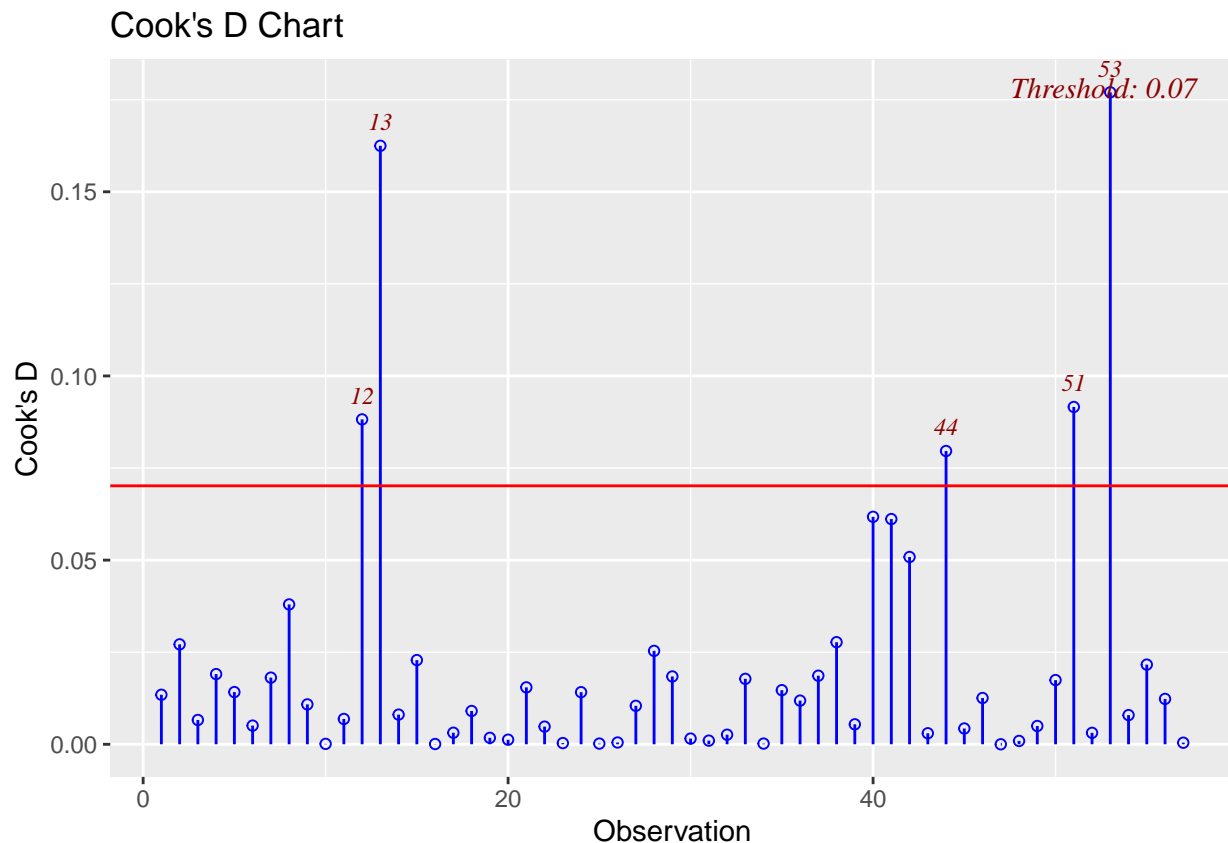


```
backward.lm=step(full, scope = list(lower=null,upper=full), direction="backward",test="F")
```

```
## Start:  AIC=-113
## PM10 ~ CH4 + CO + NO + NO2 + O3 + SO2 + Temp + RH
##
##          Df Sum of Sq    RSS      AIC F value    Pr(>F)
## - O3      1    0.0025  5.7276 -114.972  0.0206 0.8863737
## - CH4     1    0.0701  5.7953 -114.303  0.5877 0.4470738
## - NO2     1    0.1579  5.8831 -113.446  1.3238 0.2556145
## - RH      1    0.1776  5.9028 -113.255  1.4891 0.2283152
## <none>
##          5.7252 -112.996
## - CO      1    0.5091  6.2343 -110.140  4.2684 0.0442483 *
## - NO      1    1.0728  6.7980 -105.206  8.9946 0.0042822 **
## - Temp    1    1.5491  7.2742 -101.347 12.9873 0.0007434 ***
## - SO2     1    4.9995 10.7247  -79.219 41.9162 4.71e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step:  AIC=-114.97
## PM10 ~ CH4 + CO + NO + NO2 + SO2 + Temp + RH
##
##          Df Sum of Sq    RSS      AIC F value    Pr(>F)
## - CH4     1    0.0678  5.7954 -116.301  0.5801 0.4499198
## - NO2     1    0.1669  5.8945 -115.334  1.4280 0.2378467
## <none>
##          5.7276 -114.972
## - RH      1    0.2432  5.9708 -114.602  2.0805 0.1555538
```

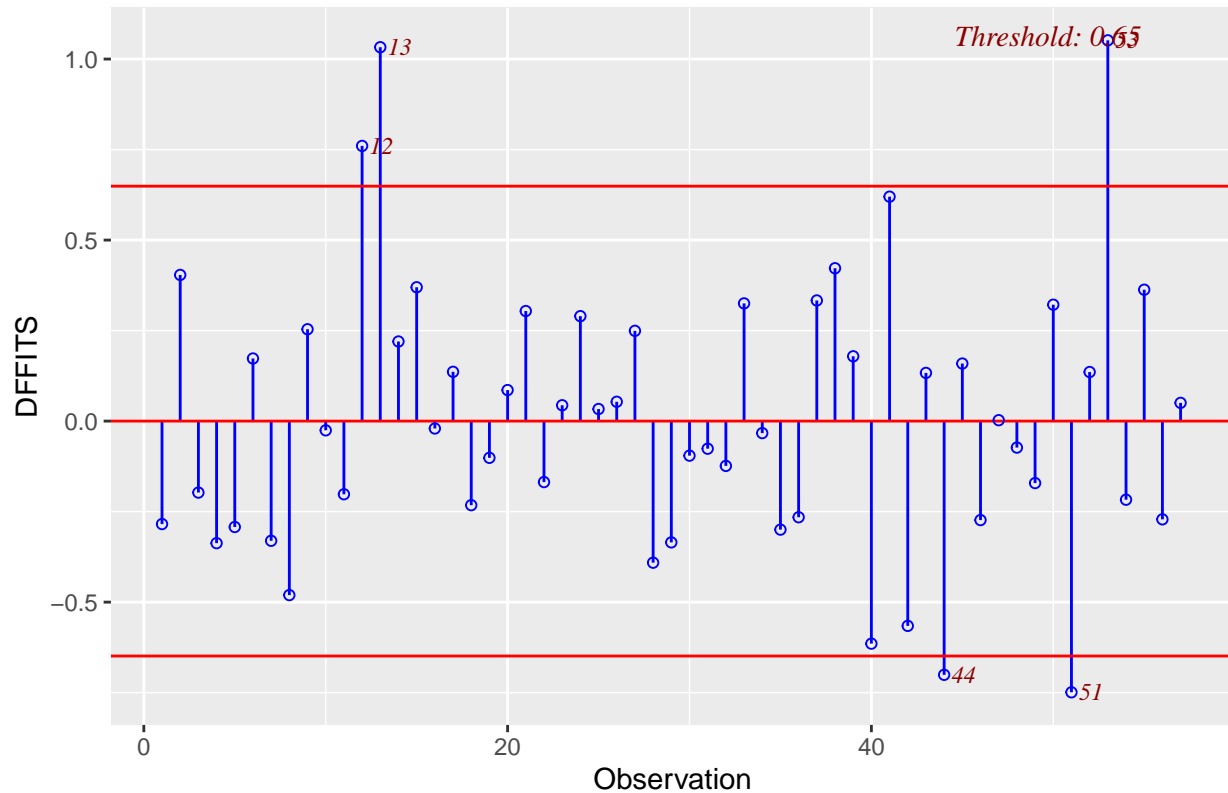
```
## - CO      1      0.5531  6.2807 -111.717  4.7320 0.0344607 *
## - Temp    1      1.5573  7.2849 -103.263 13.3223 0.0006361 ***
## - NO      1      1.7665  7.4941 -101.649 15.1124 0.0003047 ***
## - SO2     1      5.0154 10.7431 -81.121  42.9071  3.3e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: AIC=-116.3
## PM10 ~ CO + NO + NO2 + SO2 + Temp + RH
##
##          Df Sum of Sq    RSS    AIC F value    Pr(>F)
## <none>                5.7954 -116.301
## - RH      1      0.2264  6.0218 -116.117  1.9530 0.1684375
## - NO2     1      0.2843  6.0797 -115.572  2.4524 0.1236546
## - CO      1      0.5370  6.3325 -113.250  4.6331 0.0362119 *
## - NO      1      1.6992  7.4947 -103.645 14.6601 0.0003597 ***
## - Temp    1      1.8060  7.6014 -102.839 15.5810 0.0002475 ***
## - SO2     1      5.0632 10.8587 -82.511  43.6830 2.452e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ols_plot_cooksd_chart(model1)
```



```
ols_plot_dffits(model1)
```

## Influence Diagnostics for PM10



```
inflm.SR = influence.measures(model1)
inflm.SR # all
```

```
## Influence measures of
## lm(formula = PM10 ~ CO + NO + SO2 + Temp + RH, data = train) :
##
##      dfb.1_    dfb.CO    dfb.NO    dfb.SO2    dfb.Temp    dfb.RH    dffit cov.r
## 31 -0.139131  0.116984 -0.049145 -0.12579  0.033301 -5.05e-02 -0.28440 1.055
## 51  0.185575  0.179525 -0.200314 -0.09265  0.078459  1.29e-01  0.40371 1.143
## 14 -0.047050  0.118253  0.094519 -0.07494  0.167397 -3.07e-02 -0.19739 1.292
## 67 -0.090833  0.018343 -0.128313  0.18030 -0.075733  1.82e-01 -0.33710 1.206
## 42 -0.163241  0.027122  0.180390 -0.12272  0.036940 -3.69e-02 -0.29250 1.026
## 50  0.072792  0.044992 -0.096449 -0.02739  0.003375  7.73e-02  0.17295 1.298
## 43 -0.165494  0.142337  0.166613 -0.20919  0.109709 -1.72e-01 -0.33059 1.040
## 37 -0.153809  0.312021  0.081633 -0.18792  0.354602 -2.55e-01 -0.48060 1.045
## 52  0.124786  0.045079 -0.151416 -0.01658  0.016775  1.45e-01  0.25368 1.191
## 69 -0.011222 -0.014712 -0.004813  0.01858 -0.016784  1.44e-02 -0.02549 1.246
## 54 -0.101057 -0.050503 -0.104508  0.12584 -0.119213  5.68e-03 -0.20205 1.156
## 25  0.256314 -0.454123 -0.106711  0.01865 -0.497092  2.06e-01  0.76011 0.651
## 26  0.196769 -0.728947 -0.142003  0.37569 -0.851306 -7.11e-02  1.03269 0.691
## 27  0.131137 -0.041031 -0.037563  0.11303 -0.095959  4.21e-02  0.21984 1.048
## 5  0.145065 -0.185540 -0.084909  0.20435 -0.126918  3.19e-01  0.36971 1.184
## 70 -0.007656  0.005915  0.005007  0.00558  0.005390  5.76e-03 -0.02038 1.220
## 58  0.064837  0.075402 -0.037747 -0.06030  0.053753 -6.52e-02  0.13614 1.202
## 28 -0.108037  0.007538  0.022412 -0.13168  0.045064  6.74e-02 -0.23221 1.115
## 9  -0.050006 -0.051063 -0.017145 -0.02184 -0.051339  1.31e-02 -0.10145 1.248
## 29  0.050804  0.044982  0.023258 -0.01394  0.043873 -5.42e-02  0.08566 1.169
```

```

## 35  0.134868 -0.144334 -0.143374  0.12562 -0.148714  2.66e-01  0.30400 1.134
## 8  -0.081252  0.037819 -0.040654 -0.07194 -0.006574 -7.73e-02 -0.16842 1.195
## 59  0.012200 -0.008614 -0.006349 -0.01501 -0.011911 -2.09e-02  0.04375 1.264
## 7   0.085352 -0.210066 -0.092557  0.22257 -0.172948  9.54e-02  0.28987 1.239
## 68  0.014251 -0.002190  0.009876 -0.01767  0.009649  3.87e-03  0.03334 1.241
## 53  0.027136  0.017555 -0.010568 -0.00791  0.018601  3.37e-02  0.05337 1.280
## 19  0.092228 -0.039591  0.127447 -0.03083  0.043477 -1.09e-01  0.24935 1.148
## 36 -0.119047  0.229680  0.162698 -0.23368  0.325975 -3.50e-02 -0.39106 1.087
## 62 -0.166942 -0.056767 -0.136608  0.16471 -0.023793  9.23e-02 -0.33502 0.966
## 17 -0.058610 -0.023911  0.006234 -0.03436 -0.030279 -3.00e-02 -0.09540 1.190
## 39 -0.028995 -0.045457 -0.012018  0.01644 -0.026371  5.68e-02 -0.07622 1.268
## 12 -0.051663 -0.061735 -0.088867  0.04610 -0.060586  3.34e-02 -0.12376 1.254
## 15  0.118943  0.099411  0.223624 -0.04597  0.106129 -6.35e-03  0.32509 1.247
## 32 -0.014216  0.008420 -0.010999 -0.01262 -0.000994  8.98e-04 -0.03317 1.231
## 48 -0.167559  0.106745  0.125068 -0.04884  0.178796 -2.02e-02 -0.29978 0.931
## 46 -0.088039 -0.016841  0.128227  0.04122  0.037290  1.25e-01 -0.26551 1.192
## 10  0.161578  0.187991 -0.109045  0.08431  0.126769  5.07e-02  0.33321 1.203
## 23  0.330653  0.137239  0.003168  0.07714  0.103204  1.38e-01  0.42222 0.684
## 55  0.065564  0.014724  0.066698 -0.12218  0.069217 -5.76e-02  0.17900 1.222
## 61 -0.232146 -0.417653 -0.328755  0.41626 -0.371196  1.70e-01 -0.61466 1.058
## 21  0.294845  0.259881  0.382038 -0.38421  0.446640 -1.33e-02  0.61990 0.842
## 64 -0.256065 -0.128281  0.321756  0.14571  0.032548  8.65e-02 -0.56564 0.820
## 71  0.071537 -0.013053  0.031992 -0.07850  0.003232 -2.27e-02  0.13310 1.133
## 33 -0.263104 -0.141321  0.026482 -0.37414 -0.128099  6.34e-02 -0.70081 1.018
## 6   0.042030 -0.107761  0.015028  0.08977 -0.067705  1.45e-02  0.15898 1.298
## 60 -0.089377  0.030040 -0.195001  0.10917 -0.005554  3.77e-02 -0.27342 1.181
## 2   0.000989 -0.000316 -0.000347  0.00132 -0.001109 -8.58e-05  0.00253 1.245
## 57 -0.046835 -0.009319  0.018373  0.02633 -0.016580  4.09e-03 -0.07316 1.162
## 66 -0.048823  0.060839 -0.035732  0.06333  0.007395  3.70e-02 -0.17120 1.252
## 40  0.102587  0.123004 -0.137779  0.07747  0.025398 -1.41e-01  0.32155 1.326
## 13 -0.221718  0.055430 -0.406215  0.10605 -0.017752 -2.74e-01 -0.74910 1.119
## 18  0.061134 -0.045487  0.063423 -0.00323  0.009748 -2.12e-04  0.13552 1.177
## 1   0.201715  0.257174  0.599302 -0.13828  0.142744 -7.64e-01  1.05174 1.061
## 38 -0.099029  0.024059  0.048634 -0.04518  0.096673 -3.16e-02 -0.21710 1.133
## 41  0.184124  0.126160 -0.108066 -0.06837  0.105644 -1.70e-01  0.36292 0.983
## 44 -0.146927  0.003933  0.096713  0.01635 -0.033609 -1.68e-01 -0.27134 1.112
## 47  0.036065  0.020789 -0.010252  0.00103  0.014179  2.22e-02  0.05031 1.177
##      cook.d      hat inf
## 31 1.34e-02 0.0668
## 51 2.71e-02 0.1323
## 14 6.59e-03 0.1509
## 67 1.91e-02 0.1406
## 42 1.42e-02 0.0610
## 50 5.07e-03 0.1498
## 43 1.81e-02 0.0755
## 37 3.80e-02 0.1190
## 52 1.08e-02 0.1107
## 69 1.10e-04 0.0972
## 54 6.88e-03 0.0793
## 25 8.82e-02 0.0924
## 26 1.62e-01 0.1556  *
## 27 8.05e-03 0.0460
## 5   2.29e-02 0.1395
## 70 7.06e-05 0.0772

```

```
## 58 3.14e-03 0.0850
## 28 9.04e-03 0.0713
## 9 1.75e-03 0.1072
## 29 1.24e-03 0.0519
## 35 1.55e-02 0.1003
## 8 4.79e-03 0.0897
## 59 3.25e-04 0.1109
## 7 1.41e-02 0.1437
## 68 1.89e-04 0.0937
## 53 4.84e-04 0.1222
## 19 1.04e-02 0.0898
## 36 2.54e-02 0.1080
## 62 1.84e-02 0.0583
## 17 1.54e-03 0.0675
## 39 9.87e-04 0.1167
## 12 2.60e-03 0.1145
## 15 1.78e-02 0.1562
## 32 1.87e-04 0.0866
## 48 1.47e-02 0.0431
## 46 1.19e-02 0.1147
## 10 1.86e-02 0.1379
## 23 2.77e-02 0.0367
## 55 5.42e-03 0.1075
## 61 6.18e-02 0.1589
## 21 6.11e-02 0.1011
## 64 5.08e-02 0.0840
## 71 2.99e-03 0.0473
## 33 7.96e-02 0.1680
## 6 4.28e-03 0.1477
## 60 1.26e-02 0.1116
## 2 1.09e-06 0.0955
## 57 9.08e-04 0.0441
## 66 4.96e-03 0.1225
## 40 1.74e-02 0.1935
## 13 9.16e-02 0.2121
## 18 3.11e-03 0.0706
## 1 1.77e-01 0.2619 *
## 38 7.92e-03 0.0737
## 41 2.16e-02 0.0695
## 44 1.23e-02 0.0819
## 47 4.30e-04 0.0486
```

```
which(apply(inflm.SR$is.inf, 1, any)) # which
```

```
## 26 1
## 13 53
```

```
#observations 'are' influential
summary(inflm.SR) # only these
```

```
## Potentially influential observations of
## lm(formula = PM10 ~ CO + NO + SO2 + Temp + RH, data = train) :
##
## dfb.1_ dfb.CO dfb.NO dfb.SO2 dfb.Temp dfb.RH dffit cov.r cook.d hat
## 26 0.20 -0.73 -0.14 0.38 -0.85 -0.07 1.03_* 0.69 0.16 0.16
```

```
## 1    0.20    0.26    0.60   -0.14    0.14    -0.76    1.05_*    1.06    0.18    0.26
```

```
model3=lm(PM10~CO+NO+S02+Temp+RH, test)
summary(model3)
```

```
##
## Call:
## lm(formula = PM10 ~ CO + NO + S02 + Temp + RH, data = test)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.58678 -0.14860  0.05575  0.13827  0.32108
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.23043    0.11078   2.080  0.07110 .
## CO           0.76652    0.18295   4.190  0.00304 **
## NO           0.20731    0.19639   1.056  0.32198
## S02          0.34059    0.16268   2.094  0.06962 .
## Temp        -0.09279    0.22508  -0.412  0.69100
## RH          -0.37328    0.13308  -2.805  0.02303 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3252 on 8 degrees of freedom
## Multiple R-squared:  0.9396, Adjusted R-squared:  0.9019
## F-statistic: 24.89 on 5 and 8 DF,  p-value: 0.0001115

pred=predict(model3, data = test)
error=test$PM10-pred

MSE=sum(error^2)/10
MSE

## [1] 0.08460076
```