

# sold\_units\_complete

Francisco Finochietto and Jerónimo Pissinis

2/27/2022

## Analysis of the factors related with the number of units sold per year

```
#Importing the packages
```

```
library(readr)
library(car)
```

```
## Loading required package: carData
```

```
library(glmnet)
```

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

```
## Loaded glmnet 4.1-3
```

```
library(leaps)
library(lmvar)
```

Importing the data

```
file_path<-"../raw/sold_units_complete.csv"
sold_units<-read_csv(file_path)
```

```
##
## -- Column specification -----
## cols(
##   Año = col_double(),
##   'Unidades Vendidas' = col_double(),
##   'ITCRB Estados Unidos Promedio' = col_double(),
##   'Importacion de autos' = col_double(),
##   'Crisis Semiconductores' = col_double(),
##   'Devaluacion Interanual' = col_double(),
##   Inflacion = col_double(),
##   'Restriccion de importaciones' = col_double(),
##   'PIB (Millones de US$ a precios actuales)' = col_double(),
##   'Reservas Internacionales' = col_double(),
##   'PIB/reservas' = col_double(),
##   'Brecha Cambiaria' = col_double(),
##   'Diferencia Trade Balance Industria' = col_number()
## )
```

```

#Dropping the year column.
sold_units<-sold_units[,-1]

#Centering the variables to reduce structural multicollinearity
sold_units[,8]<-scale(sold_units[,8],scale=FALSE)
sold_units[,9]<-scale(sold_units[,9],scale=FALSE)
sold_units[,10]<-scale(sold_units[,10],scale=FALSE)

#Renaming the columns
my_names<-c("num_units", "itcrb", "imported_cars", "semiconductor_crisis",
            "devaluacion_interanual", "inflation", "import_restriction",
            "PIB", "reserves", "PIB_over_reserves", "exchange_difference",
            "industry_trade_balance_diference")
names(sold_units)<-my_names

```

Building the model

```

sold_units_model<-lm(sold_units)
summary(sold_units_model)

```

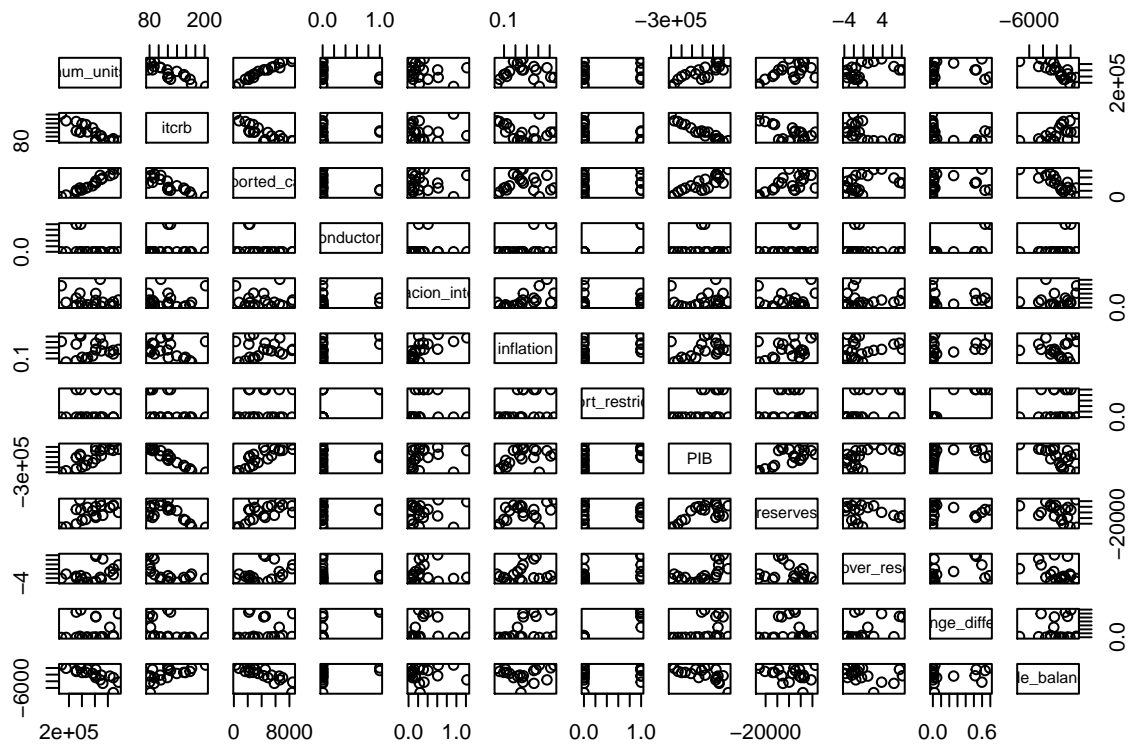
```

##
## Call:
## lm(formula = sold_units)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -62015 -12576  -1454   19485   66203
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    6.445e+05  4.954e+05   1.301   0.2295
## itcrb          -3.134e+03  3.815e+03  -0.822   0.4351
## imported_cars    8.011e+01  2.615e+01   3.063   0.0155 *
## semiconductor_crisis -8.198e+04  1.099e+05  -0.746   0.4770
## devaluacion_interanual -8.825e+04  5.990e+04  -1.473   0.1789
## inflation        5.412e+03  1.883e+05   0.029   0.9778
## import_restriction -3.926e+04  1.058e+05  -0.371   0.7202
## PIB             5.299e-01  8.115e-01   0.653   0.5321
## reserves        -8.287e+00  1.045e+01  -0.793   0.4506
## PIB_over_reserves -3.020e+04  3.780e+04  -0.799   0.4474
## exchange_difference  8.548e+04  2.097e+05   0.408   0.6942
## industry_trade_balance_diference 1.064e+01  1.747e+01   0.609   0.5594
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 43590 on 8 degrees of freedom
## Multiple R-squared:  0.9866, Adjusted R-squared:  0.9682
## F-statistic: 53.58 on 11 and 8 DF,  p-value: 2.911e-06

```

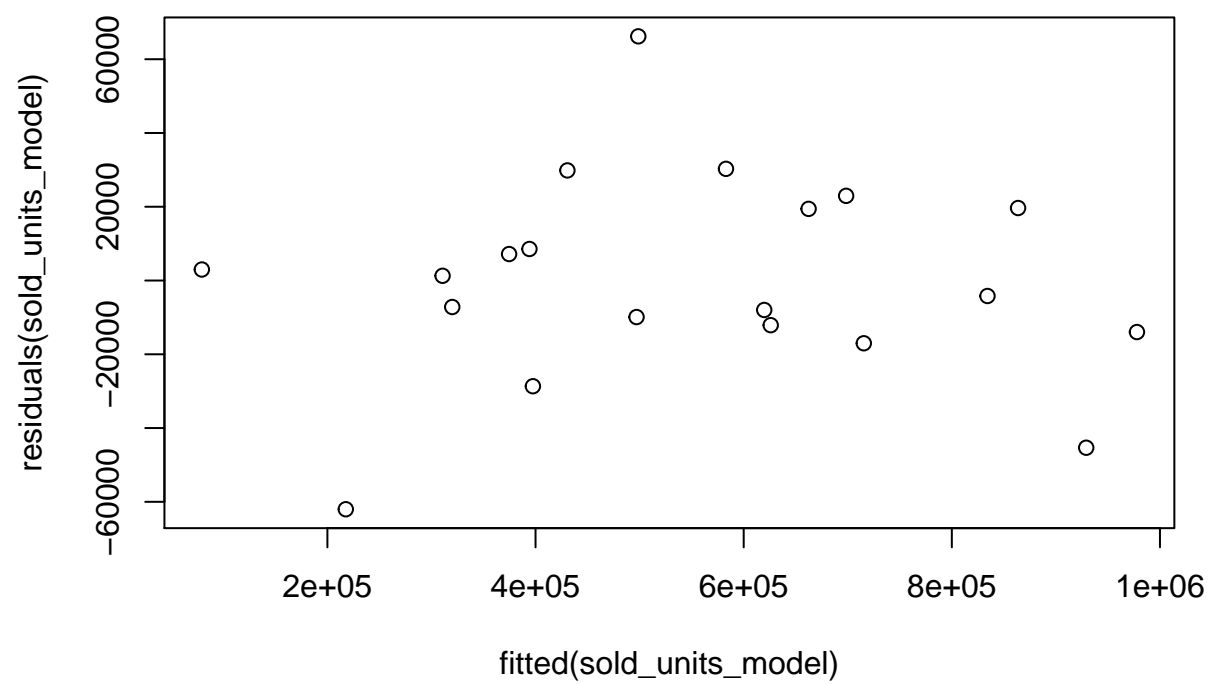
Pairwise plots of the features

```
pairs(sold_units)
```

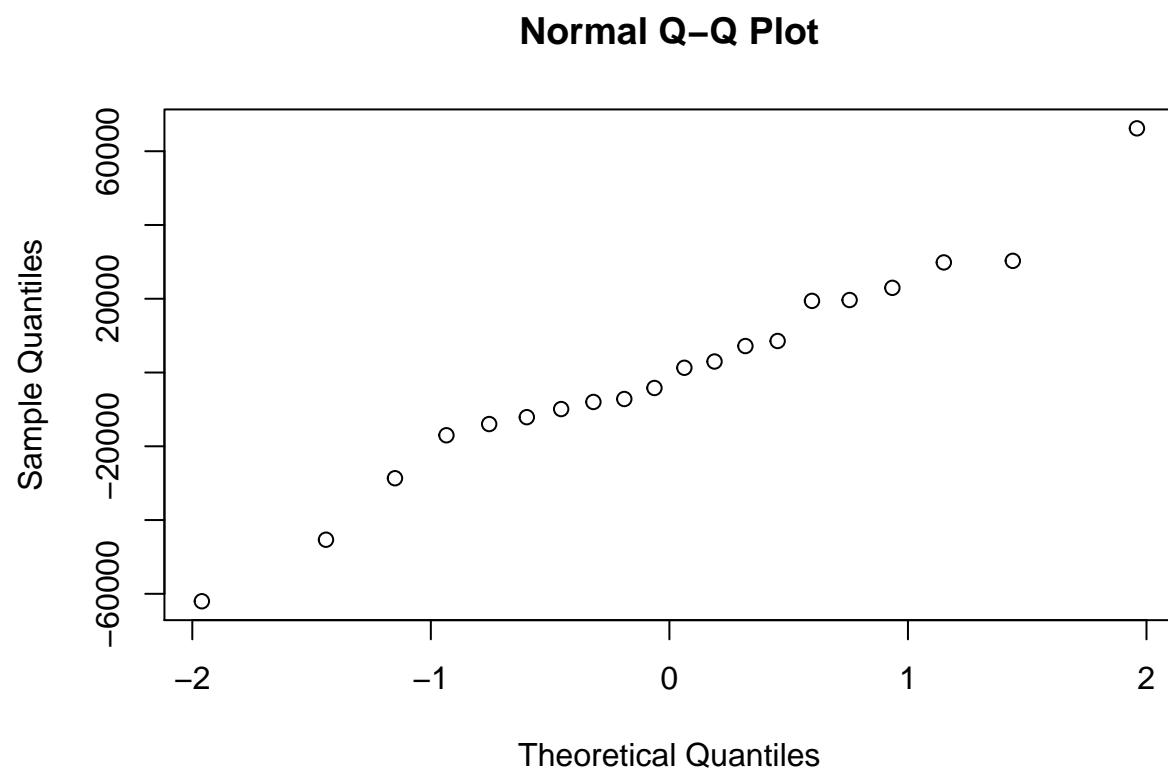


Analyzing the residuals

```
plot(fitted(sold_units_model),residuals(sold_units_model))
```

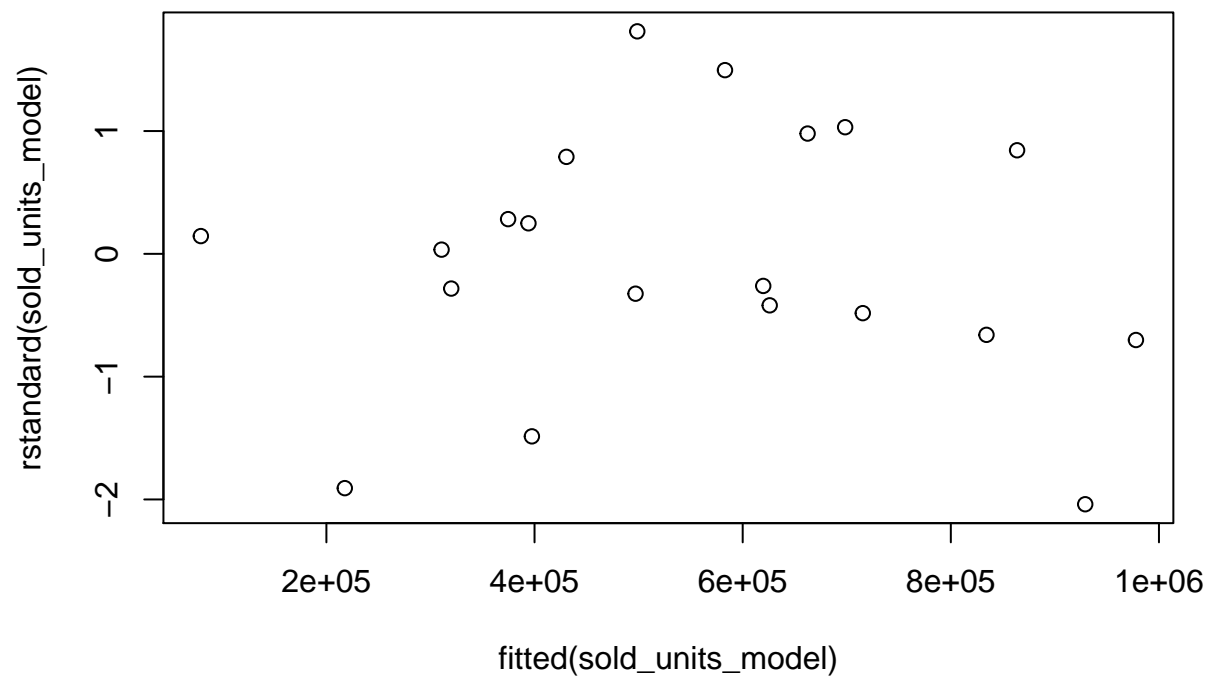


```
qqnorm(residuals(sold_units_model))
```

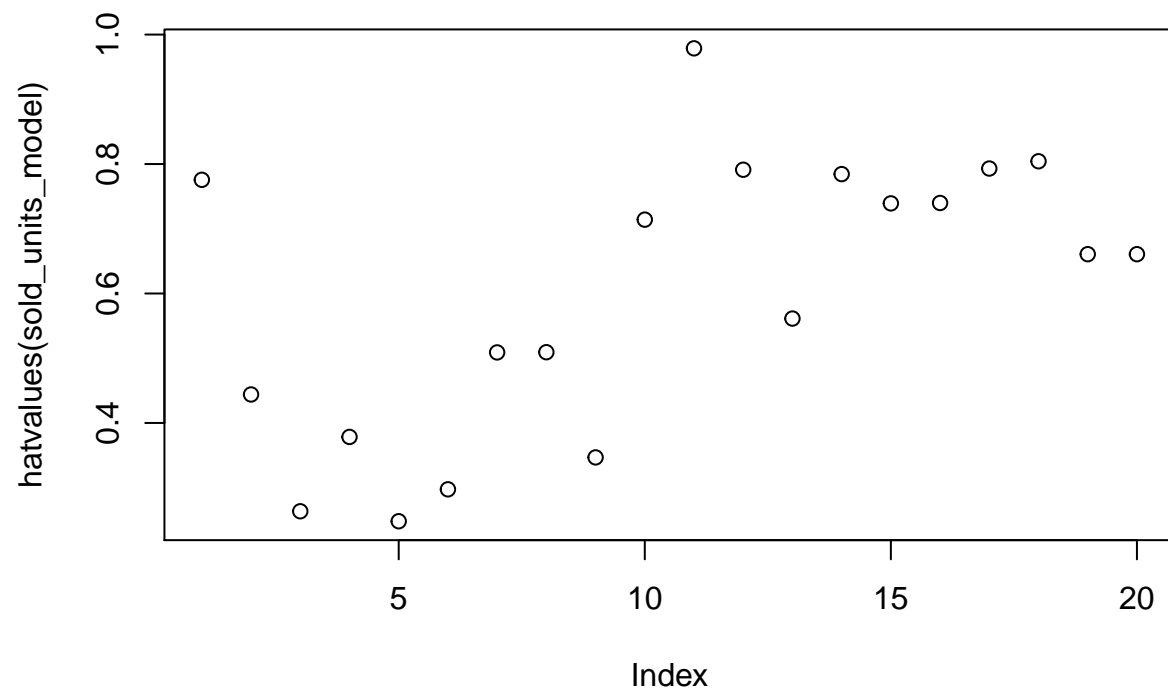


Looking for outliers and high leverage points

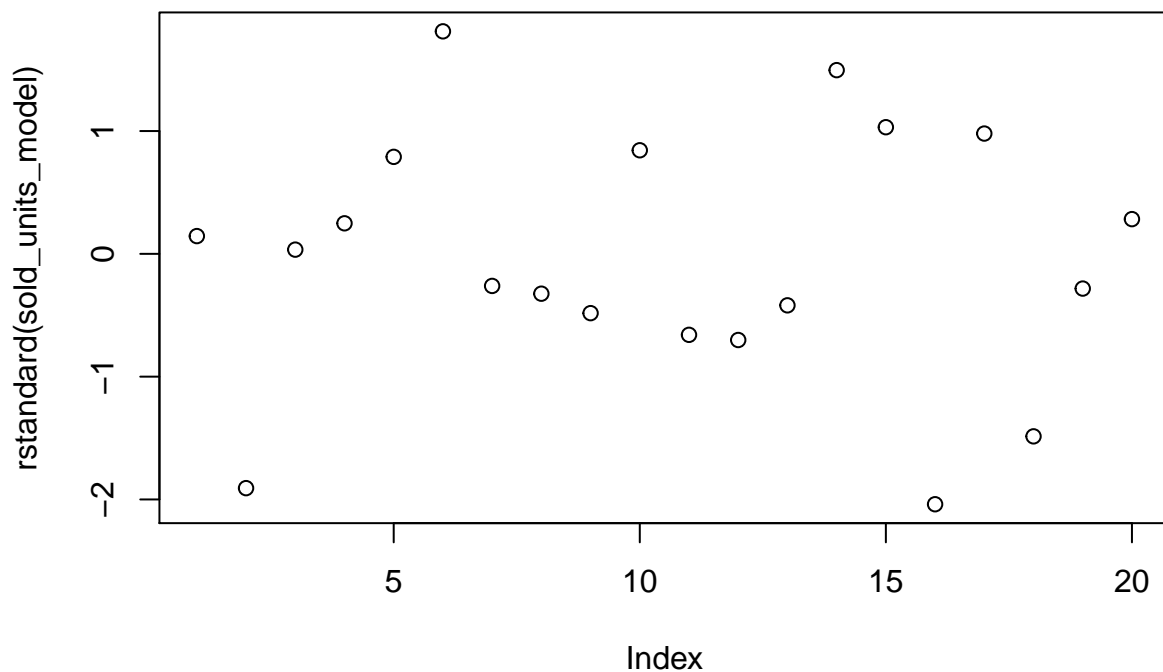
```
plot(fitted(sold_units_model),rstandard(sold_units_model))
```



```
plot(hatvalues(sold_units_model))  
abline(h=length(coef(sold_units_model))/nrow(sold_units)*2,  
       col = "red",lty = 2)
```



```
high_leverage_points<-hatvalues(sold_units_model)>
  (length(coef(sold_units_model))/nrow(sold_units)*2)
plot(rstandard(sold_units_model),
     col = factor(high_leverage_points))
```



Looking for colinearity Correlation matrix

```
cor(sold_units[, -1])
```

```
##               itcrb imported_cars semiconductor_crisis
## itcrb          1.00000000 -0.84998083      0.031025294
## imported_cars -0.84998083      1.00000000     -0.275463784
## semiconductor_crisis 0.03102529 -0.27546378      1.000000000
## devaluacion_interannual 0.04051446 -0.02721305      0.005059507
## inflation        -0.27166538  0.10875007      0.391827027
## import_restriction -0.45615253  0.15017680      0.509175077
## PIB              -0.97038025  0.84102366      0.011897026
## reserves         -0.64341540  0.59313786      0.113498077
## PIB_over_reserves -0.57391965  0.39056882     -0.123565165
## exchange_difference -0.40454741  0.09171044      0.650902612
## industry_trade_balance_diference 0.62115112 -0.85146240      0.258351230
##
##               devaluacion_interannual  inflation
## itcrb                0.040514456 -0.27166538
## imported_cars        -0.027213050  0.10875007
## semiconductor_crisis  0.005059507  0.39182703
## devaluacion_interannual 1.000000000  0.65528084
## inflation             0.655280837  1.00000000
## import_restriction    0.060402026  0.31212355
## PIB                   0.125177055  0.42376923
## reserves              0.081914320  0.41050289
## PIB_over_reserves     0.092155529  0.14588345
```



```

## exchange_difference          0.073825750  0.38737953
## industry_trade_balance_diference  0.079954149  0.08132427
##                               import_restriction      PIB      reserves
## itcrb                        -0.45615253 -0.97038025 -0.64341540
## imported_cars                 0.15017680  0.84102366  0.59313786
## semiconductor_crisis          0.50917508  0.01189703  0.11349808
## devaluacion_interanual         0.06040203  0.12517705  0.08191432
## inflation                     0.31212355  0.42376923  0.41050289
## import_restriction            1.00000000  0.42912174  0.02503357
## PIB                           0.42912174  1.00000000  0.65862991
## reserves                     0.02503357  0.65862991  1.00000000
## PIB_over_reserves             0.53395398  0.58138567 -0.21014720
## exchange_difference            0.95207008  0.39340556  0.03724469
## industry_trade_balance_diference  0.06360759 -0.64427806 -0.33854868
##                               PIB_over_reserves exchange_difference
## itcrb                        -0.57391965          -0.40454741
## imported_cars                 0.39056882           0.09171044
## semiconductor_crisis          -0.12356516           0.65090261
## devaluacion_interanual         0.09215553           0.07382575
## inflation                     0.14588345           0.38737953
## import_restriction            0.53395398           0.95207008
## PIB                           0.58138567           0.39340556
## reserves                     -0.21014720           0.03724469
## PIB_over_reserves             1.00000000           0.48228198
## exchange_difference            0.48228198           1.00000000
## industry_trade_balance_diference -0.37317620           0.08784890
##                               industry_trade_balance_diference
## itcrb                        0.62115112
## imported_cars                -0.85146240
## semiconductor_crisis         0.25835123
## devaluacion_interanual       0.07995415
## inflation                     0.08132427
## import_restriction           0.06360759
## PIB                          -0.64427806
## reserves                    -0.33854868
## PIB_over_reserves            -0.37317620
## exchange_difference           0.08784890
## industry_trade_balance_diference 1.00000000

```

Variance inflation factors

```
vif(sold_units_model)
```

```

##                               itcrb                               imported_cars
##                               188.647668                               40.041749
##                               semiconductor_crisis                 devaluacion_interanual
##                               11.435279                               3.626339
##                               inflation                               import_restriction
##                               7.827097                               24.727457
##                               PIB                               reserves
##                               182.555051                               224.155138
##                               PIB_over_reserves                 exchange_difference
##                               196.605341                               29.934856

```

```
## industry_trade_balance_diference
## 12.545070
```

Eigenvalues of the correlation matrix

```
eigen(cor(sold_units[, -1]))$values
```

```
## [1] 4.518840497 2.690677954 1.652199919 1.222258483 0.450025634 0.278666946
## [7] 0.121728329 0.033478040 0.023565463 0.007116211 0.001442524
```

## Feature selection

Applying best subset selection

```
sold_units_all<-regsubsets(sold_units$num_units~.,sold_units,nvmax = 12)
summary(sold_units_all)
```

```
## Subset selection object
## Call: regsubsets.formula(sold_units$num_units ~ ., sold_units, nvmax = 12)
## 11 Variables (and intercept)
##
```

|                                     |  | Forced in | Forced out |
|-------------------------------------|--|-----------|------------|
| ## itcrb                            |  | FALSE     | FALSE      |
| ## imported_cars                    |  | FALSE     | FALSE      |
| ## semiconductor_crisis             |  | FALSE     | FALSE      |
| ## devaluacion_interannual          |  | FALSE     | FALSE      |
| ## inflation                        |  | FALSE     | FALSE      |
| ## import_restriction               |  | FALSE     | FALSE      |
| ## PIB                              |  | FALSE     | FALSE      |
| ## reserves                         |  | FALSE     | FALSE      |
| ## PIB_over_reserves                |  | FALSE     | FALSE      |
| ## exchange_difference              |  | FALSE     | FALSE      |
| ## industry_trade_balance_diference |  | FALSE     | FALSE      |

```
## 1 subsets of each size up to 11
## Selection Algorithm: exhaustive
##
```

|             |     | itcrb | imported_cars | semiconductor_crisis | devaluacion_interannual |
|-------------|-----|-------|---------------|----------------------|-------------------------|
| ## 1 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 2 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 3 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 4 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 5 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 6 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 7 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 8 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 9 ( 1 )  | " " | " "   | " "           | " "                  | " "                     |
| ## 10 ( 1 ) | " " | " "   | " "           | " "                  | " "                     |
| ## 11 ( 1 ) | " " | " "   | " "           | " "                  | " "                     |

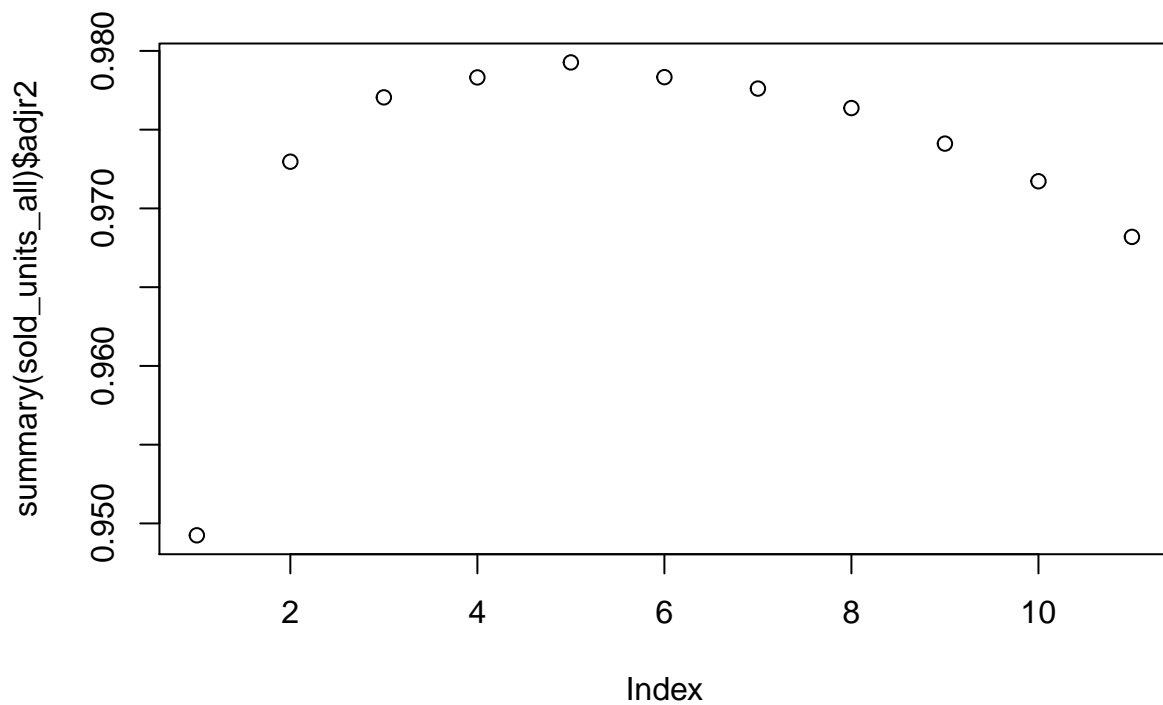
```
##
```

|            |     | inflation | import_restriction | PIB | reserves | PIB_over_reserves |
|------------|-----|-----------|--------------------|-----|----------|-------------------|
| ## 1 ( 1 ) | " " | " "       | " "                | " " | " "      | " "               |
| ## 2 ( 1 ) | " " | " "       | " "                | " " | " "      | " "               |
| ## 3 ( 1 ) | " " | " "       | " "                | " " | " "      | " "               |
| ## 4 ( 1 ) | " " | " "       | " "                | " " | " "      | " "               |

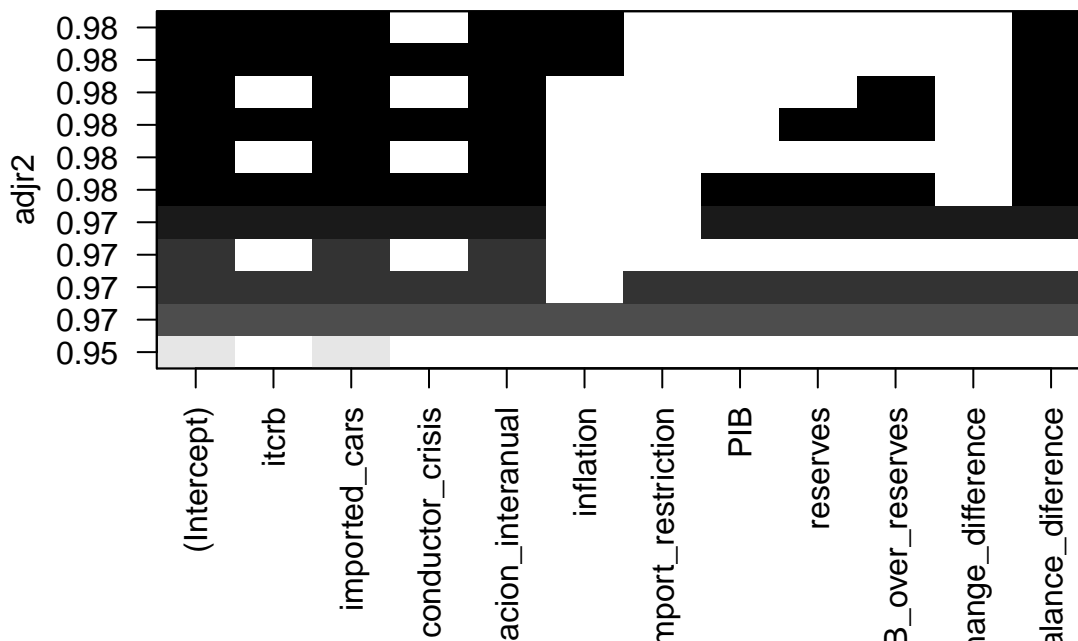
```
## 5 ( 1 ) "*"      " "      " " " "      " "
## 6 ( 1 ) "*"      " "      " " " "      " "
## 7 ( 1 ) " "      " "      " " "*"      "*"
## 8 ( 1 ) " "      " "      "*" "*"      "*"
## 9 ( 1 ) " "      " "      "*" "*"      "*"
## 10 ( 1 ) " "      "*"      "*" "*"      "*"
## 11 ( 1 ) "*"      "*"      "*" "*"      "*"
##
##      exchange_difference industry_trade_balance_difference
## 1 ( 1 ) " "      " "
## 2 ( 1 ) " "      " "
## 3 ( 1 ) " "      "*"
## 4 ( 1 ) " "      "*"
## 5 ( 1 ) " "      "*"
## 6 ( 1 ) " "      "*"
## 7 ( 1 ) " "      "*"
## 8 ( 1 ) " "      "*"
## 9 ( 1 ) "*"      "*"
## 10 ( 1 ) "*"      "*"
## 11 ( 1 ) "*"      "*"

```

```
plot(summary(sold_units_all)$adjr2)
```



```
plot(sold_units_all, scale = "adjr2")
```



```
best_adjr2<-which.max(summary(sold_units_all)$adjr2)
subset_coef<-names(coef(sold_units_all, best_adjr2))
```

Building the selected model

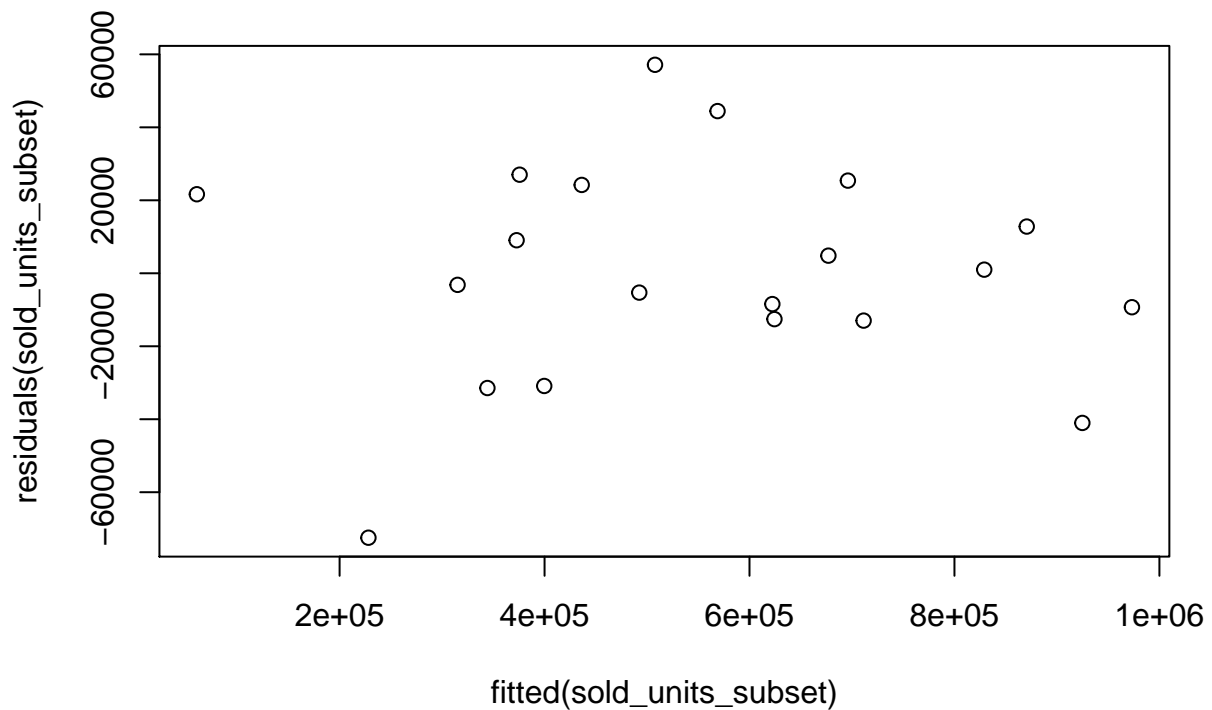
```
sold_units_subset<-
  lm(sold_units[,names(sold_units)%in%
    c("num_units",subset_coef)], y = TRUE, x = TRUE)
summary(sold_units_subset)
```

```
##
## Call:
## lm(formula = sold_units[, names(sold_units) %in% c("num_units",
##   subset_coef)], x = TRUE, y = TRUE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -72464 -12668  -1086   22298   57133
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.343e+05  9.583e+04   3.489  0.00361 **
## itcrb          -8.901e+02  5.036e+02  -1.767  0.09894 .
## imported_cars    9.820e+01  1.028e+01   9.553 1.64e-07 ***
## devaluacion_interanual -8.505e+04  3.654e+04  -2.327  0.03547 *
```

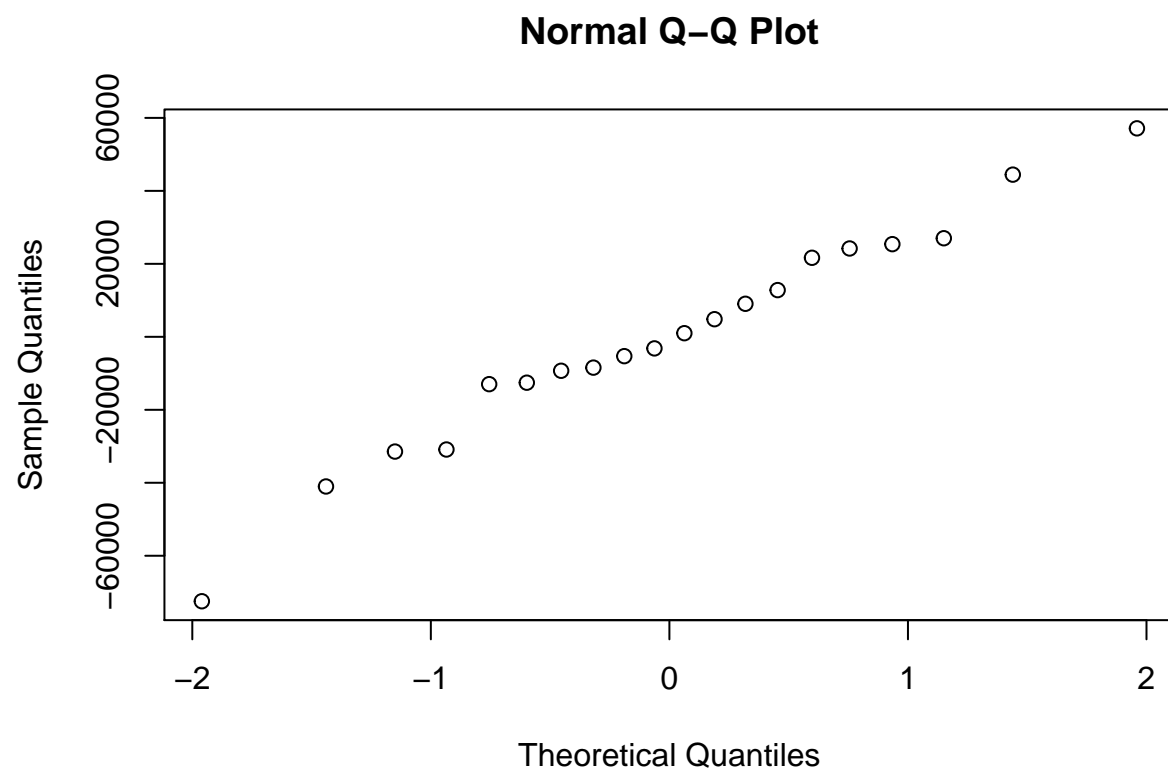
```
## inflation -1.214e+05 8.538e+04 -1.421 0.17709
## industry_trade_balance_diference 1.457e+01 8.434e+00 1.728 0.10602
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 35190 on 14 degrees of freedom
## Multiple R-squared:  0.9847, Adjusted R-squared:  0.9793
## F-statistic: 180.5 on 5 and 14 DF,  p-value: 3.385e-12
```

Analyzing the residuals

```
plot(fitted(sold_units_subset),residuals(sold_units_subset))
```

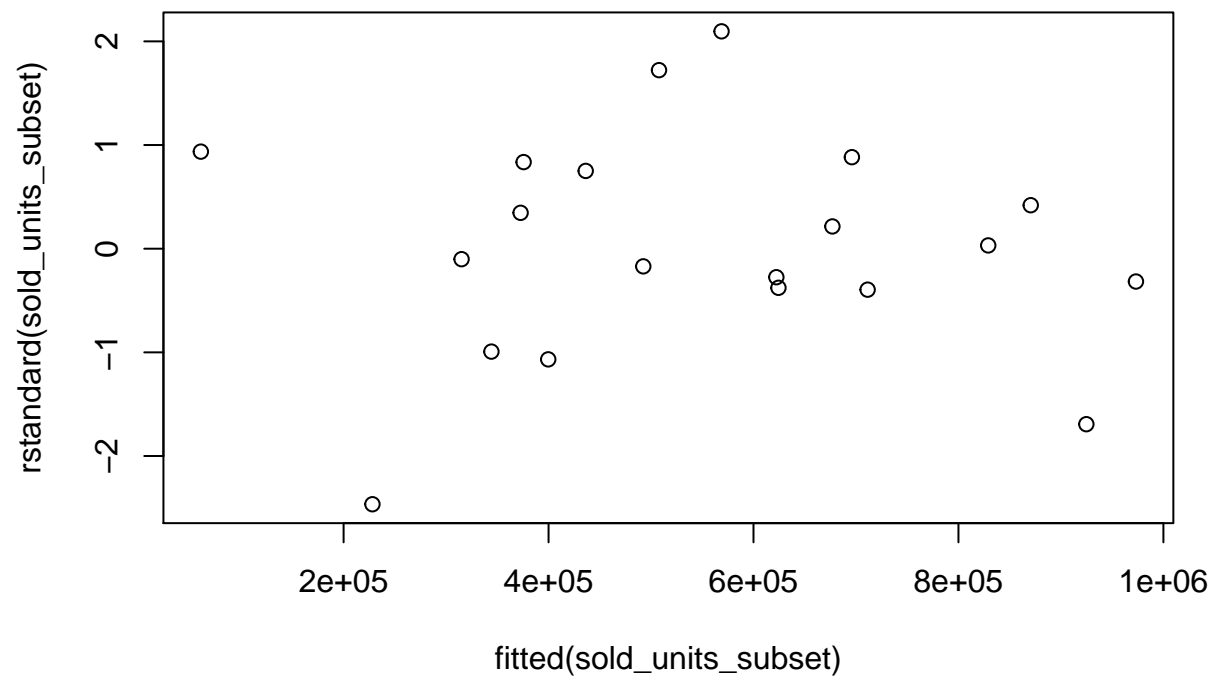


```
qqnorm(residuals(sold_units_subset))
```

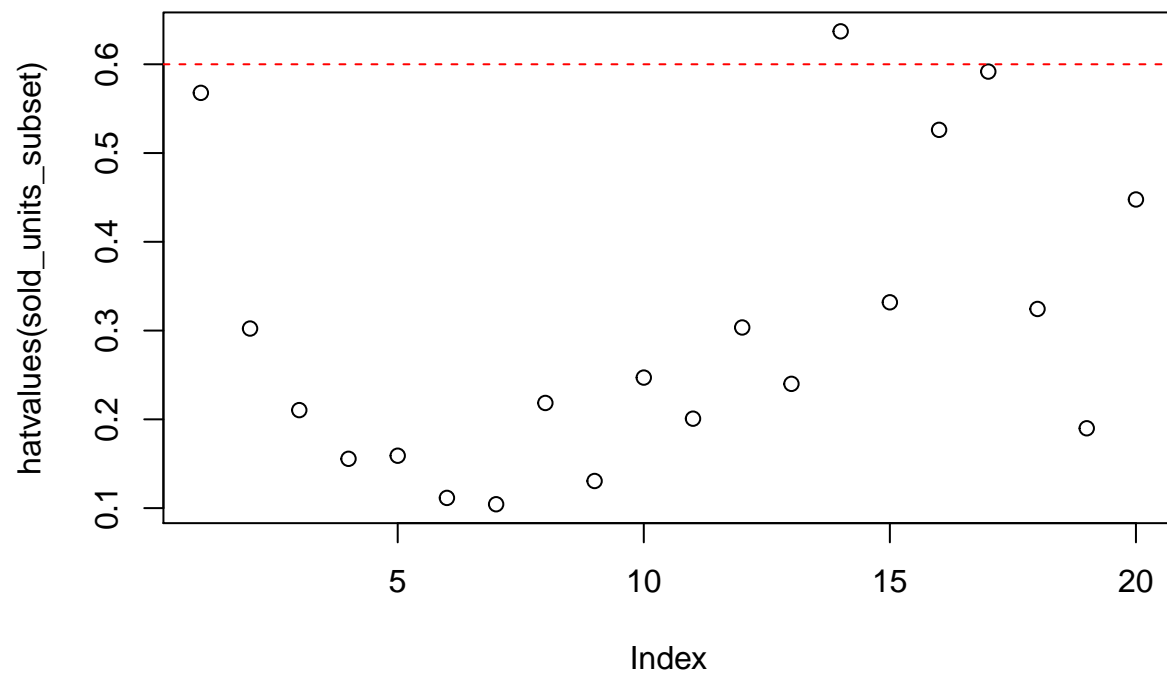


Looking for outliers and high leverage points

```
plot(fitted(sold_units_subset),rstandard(sold_units_subset))
```

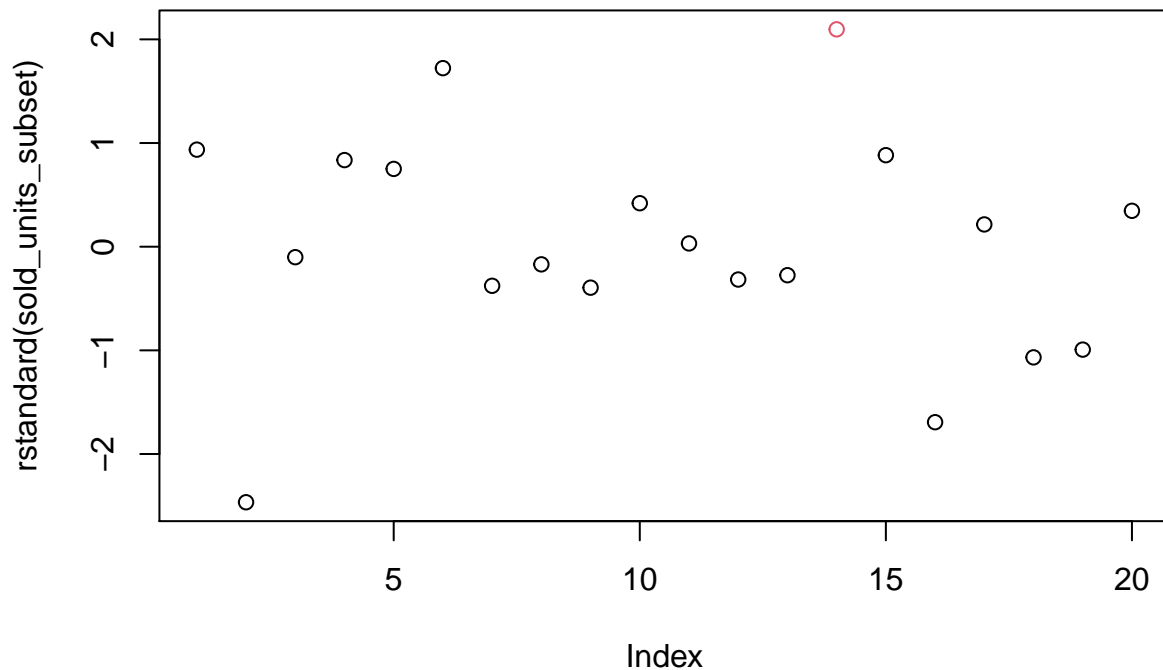


```
plot(hatvalues(sold_units_subset))
abline(h=length(coef(sold_units_subset))/nrow(sold_units)*2,
       col = "red", lty = 2)
```



```
high_leverage_points<-hatvalues(sold_units_subset)>
  (length(coef(sold_units_subset))/nrow(sold_units)*2)
plot(rstandard(sold_units_subset),
     col = factor(high_leverage_points))
```





Looking for colinearity Correlation matrix and its eigen values

```
subset_coef_cor<-cor(sold_units[,names(sold_units)%in%subset_coef])
subset_coef_cor
```

```
##               itcrb imported_cars
## itcrb          1.00000000 -0.84998083
## imported_cars  -0.84998083  1.00000000
## devaluacion_interannual  0.04051446 -0.02721305
## inflation      -0.27166538  0.10875007
## industry_trade_balance_diference  0.62115112 -0.85146240
##               devaluacion_interannual  inflation
## itcrb          0.04051446 -0.27166538
## imported_cars  -0.02721305  0.10875007
## devaluacion_interannual  1.00000000  0.65528084
## inflation      0.65528084  1.00000000
## industry_trade_balance_diference  0.07995415  0.08132427
##               industry_trade_balance_diference
## itcrb          0.62115112
## imported_cars  -0.85146240
## devaluacion_interannual  0.07995415
## inflation      0.08132427
## industry_trade_balance_diference  1.00000000
```

```
eigen(subset_coef_cor)$values
```

```
## [1] 2.5725655 1.6807642 0.4957628 0.1849049 0.0660026
```

Variance inflation factors

```
vif(sold_units_subset)
```

```
##               itcrb               imported_cars
##               5.045267               9.491226
##      devaluacion_interanual               inflation
##               2.071411               2.468309
## industry_trade_balance_diference
##               4.483669
```

Removing the high leverage outlier

```
sold_units_subset_rm<-
  lm(sold_units[!(high_leverage_points &
    (rstandard(sold_units_subset)>2)),
    names(sold_units)%in% c("num_units",subset_coef)],
    y = TRUE, x = TRUE)
summary(sold_units_subset_rm)

##
## Call:
## lm(formula = sold_units[!(high_leverage_points & (rstandard(sold_units_subset) >
##      2)), names(sold_units) %in% c("num_units", subset_coef)],
##      x = TRUE, y = TRUE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -52469 -20381   4372  16455  48379
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    9.729e+04  1.274e+05   0.764  0.45869
## itcrb          3.519e+02   6.684e+02   0.527  0.60736
## imported_cars   1.172e+02   1.178e+01   9.951  1.9e-07 ***
## devaluacion_interanual -1.492e+05  4.096e+04  -3.642  0.00298 **
## inflation       8.452e+03   9.066e+04   0.093  0.92715
## industry_trade_balance_diference 2.090e+01  7.699e+00   2.714  0.01770 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 30250 on 13 degrees of freedom
## Multiple R-squared:  0.9895, Adjusted R-squared:  0.9854
## F-statistic: 244.6 on 5 and 13 DF,  p-value: 2.218e-12
```

```
summary(sold_units_subset)
```

```
##
## Call:
## lm(formula = sold_units[, names(sold_units) %in% c("num_units",
##   subset_coef)], x = TRUE, y = TRUE)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -72464 -12668  -1086   22298   57133
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.343e+05  9.583e+04   3.489  0.00361 **
## itcrb          -8.901e+02  5.036e+02  -1.767  0.09894 .
## imported_cars    9.820e+01  1.028e+01   9.553 1.64e-07 ***
## devaluacion_interanual -8.505e+04  3.654e+04  -2.327  0.03547 *
## inflation       -1.214e+05  8.538e+04  -1.421  0.17709
## industry_trade_balance_diference 1.457e+01  8.434e+00   1.728  0.10602
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 35190 on 14 degrees of freedom
## Multiple R-squared:  0.9847, Adjusted R-squared:  0.9793
## F-statistic: 180.5 on 5 and 14 DF,  p-value: 3.385e-12
```

```
vif(sold_units_subset_rm)
```

```
##              itcrb              imported_cars
##              11.106352              16.856451
##      devaluacion_interanual              inflation
##              3.343531              3.766226
## industry_trade_balance_diference
##              5.044398
```

Testing the model selected with best subset selection using cross-validation

```
cv_sold_units_subset<-cv.lm(sold_units_subset, k=5)
cv_sold_units_subset
```

```
## Mean absolute error      : 33132.86
## Sample standard deviation : 22566.13
##
## Mean squared error       : 2041989851
## Sample standard deviation : 2082479934
##
## Root mean squared error  : 39268.21
## Sample standard deviation : 24999.94
```

Applying LASSO

```
sold_units_lasso<-glmnet(as.matrix(sold_units[,-1]),
                        as.matrix(sold_units[,1]),alpha=1)
sold_units_lasso
```

```
##
## Call:  glmnet(x = as.matrix(sold_units[, -1]), y = as.matrix(sold_units[,      1]), alpha = 1)
##
##      Df  %Dev Lambda
## 1    0  0.00 232400
## 2    1 16.16 211800
## 3    1 29.58 193000
## 4    1 40.72 175800
## 5    1 49.97 160200
## 6    1 57.65 146000
## 7    1 64.02 133000
## 8    1 69.31 121200
## 9    1 73.71 110400
## 10   1 77.35 100600
## 11   1 80.38  91680
## 12   1 82.90  83530
## 13   1 84.98  76110
## 14   1 86.72  69350
## 15   1 88.16  63190
## 16   2 89.38  57580
## 17   2 90.45  52460
## 18   2 91.34  47800
## 19   2 92.08  43550
## 20   2 92.69  39680
## 21   3 93.38  36160
## 22   3 94.17  32950
## 23   3 94.82  30020
## 24   3 95.37  27350
## 25   3 95.82  24920
## 26   3 96.19  22710
## 27   3 96.50  20690
## 28   3 96.76  18850
## 29   3 96.97  17180
## 30   3 97.15  15650
## 31   3 97.30  14260
## 32   3 97.42  12990
## 33   4 97.54  11840
## 34   4 97.66  10790
## 35   4 97.76   9830
## 36   4 97.85   8957
## 37   4 97.91   8161
## 38   4 97.97   7436
## 39   5 98.02   6776
## 40   5 98.06   6174
## 41   5 98.10   5625
## 42   6 98.14   5125
## 43   6 98.19   4670
## 44   6 98.22   4255
## 45   6 98.25   3877
```

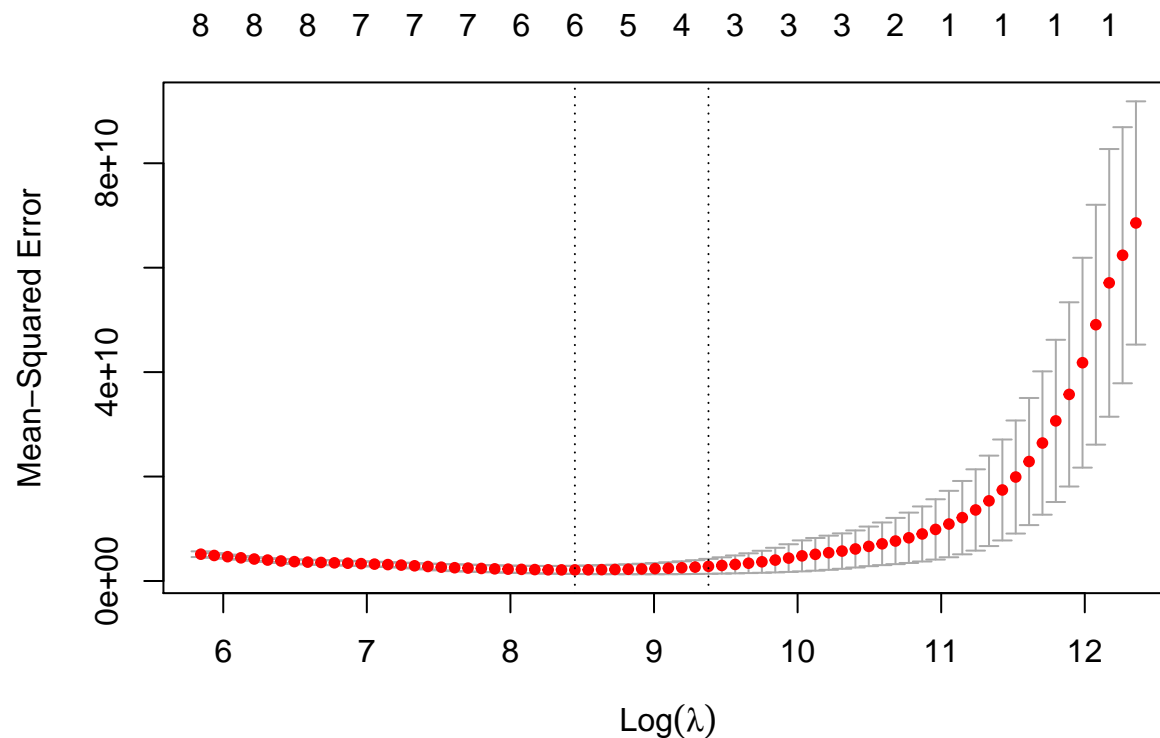
```
## 46 6 98.28 3533
## 47 6 98.30 3219
## 48 6 98.32 2933
## 49 6 98.33 2672
## 50 7 98.34 2435
## 51 7 98.38 2219
## 52 7 98.41 2022
## 53 7 98.43 1842
## 54 7 98.45 1678
## 55 7 98.47 1529
## 56 7 98.48 1393
## 57 7 98.49 1270
## 58 7 98.50 1157
## 59 7 98.51 1054
## 60 7 98.52 960
## 61 7 98.52 875
## 62 7 98.53 797
## 63 8 98.53 726
## 64 8 98.54 662
## 65 8 98.54 603
## 66 8 98.54 550
## 67 8 98.54 501
## 68 8 98.55 456
## 69 8 98.55 416
## 70 8 98.55 379
## 71 8 98.55 345
```

```
#selecting lambda using cross-validation
cv_sold_units_lasso<- cv.glmnet(as.matrix(sold_units[,-1]),
                                as.matrix(sold_units[,1]),
                                type.measure = c("mse"),
                                alpha=1,nfolds = 5)

cv_sold_units_lasso
```

```
##
## Call: cv.glmnet(x = as.matrix(sold_units[, -1]), y = as.matrix(sold_units[, 1]), type.measure = "mse")
##
## Measure: Mean-Squared Error
##
##      Lambda Index  Measure      SE Nonzero
## min   4670     43 2.128e+09 7.816e+08      6
## 1se  11840     33 2.803e+09 1.431e+09      4
```

```
plot(cv_sold_units_lasso)
```



```
best_lambda <- cv_sold_units_lasso$lambda.min
sold_units_lasso_best <- glmnet(as.matrix(sold_units[, -1]),
                               as.matrix(sold_units[, 1]), alpha = 1,
                               lambda = best_lambda)
sold_units_lasso_best
```

```
##
## Call:  glmnet(x = as.matrix(sold_units[, -1]), y = as.matrix(sold_units[, 1]), alpha = 1, lambda = best_lambda)
##
##      Df %Dev Lambda
## 1    6 98.19  4670
```

```
coef(sold_units_lasso_best)
```

```
## 12 x 1 sparse Matrix of class "dgCMatrix"
##                                     s0
## (Intercept)                   351145.69706
## itcrb                        -938.22246
## imported_cars                   83.35042
## semiconductor_crisis          -27781.45560
## devaluacion_interanual        -99385.23128
## inflation                     -10853.72857
## import_restriction             3814.48410
## PIB                           .
## reserves                       .
```

```
## PIB_over_reserves      .
## exchange_difference     .
## industry_trade_balance_diference .
```

Comparing the MSE of the best subset and LASSO models

```
mse_lasso<-min(cv_sold_units_lasso$cvm)
mse_subset<-cv_sold_units_subset$MSE$mean
mse_lasso
```

```
## [1] 2127914386
```

```
mse_subset
```

```
## [1] 2041989851
```

```
sqrt(mse_lasso)
```

```
## [1] 46129.32
```

```
sqrt(mse_subset)
```

```
## [1] 45188.38
```

```
user_predictors<-c("num_units", "itcrb",
                  "devaluacion_interanual",
                  "inflation", "import_restriction",
                  "PIB_over_reserves", "exchange_difference",
                  "industry_trade_balance_diference")
sold_units_user<-sold_units[,names(sold_units)%in%user_predictors]
str(sold_units_user)
```

```
## tibble[,8] [20 x 8] (S3: tbl_df/tbl/data.frame)
## $ num_units      : num [1:20] 82345 155666 311961 402690 460478 ...
## $ itcrb          : num [1:20] 202 172 168 158 155 ...
## $ devaluacion_interanual : num [1:20] 0.9418 0.2234 0.0956 0.0675 0.0253 ...
## $ inflation      : num [1:20] 0.409 0.037 0.061 0.123 0.098 0.148 0.257 0.114 0.23...
## $ import_restriction : num [1:20] 0 0 0 0 0 0 0 0 0 0 ...
## $ PIB_over_reserves : num [1:20] -2.09 -0.192 -1.17 -2.436 -1.644 ...
## $ exchange_difference : num [1:20] -0.00632 -0.00437 -0.00255 -0.00318 0.00013 ...
## $ industry_trade_balance_diference: num [1:20] 881 -91 -1029 -1364 -1321 ...
```

```
sold_units_user_model<-lm(sold_units_user)
summary(sold_units_user_model)
```

```
##
## Call:
## lm(formula = sold_units_user)
##
```

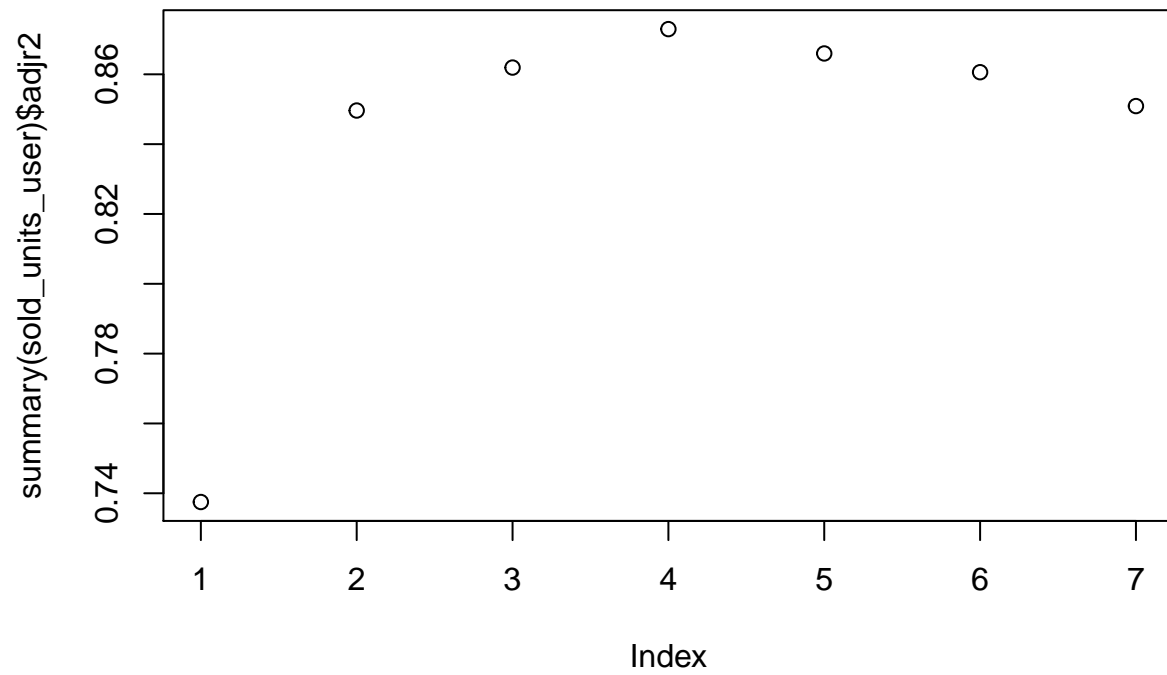
```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -124601  -42315   -6945   53967  164597
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                1105471.28  198492.04   5.569 0.000122 ***
## itcrb                      -4934.98    1162.30  -4.246 0.001136 **
## devaluacion_interannual    -40994.99  104431.39  -0.393 0.701532
## inflation                  -122750.81  258636.29  -0.475 0.643587
## import_restriction         128014.29   174850.95   0.732 0.478143
## PIB_over_reserves          -9225.65    8175.45  -1.128 0.281183
## exchange_difference        -261616.69  304320.68  -0.860 0.406815
## industry_trade_balance_diference    -46.38     17.21  -2.696 0.019469 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 94380 on 12 degrees of freedom
## Multiple R-squared:  0.9058, Adjusted R-squared:  0.8509
## F-statistic: 16.49 on 7 and 12 DF,  p-value: 2.818e-05
```

```
sold_units_user<-regsubsets(sold_units$num_units~.,sold_units_user,nvmax = 9)
summary(sold_units_user)
```

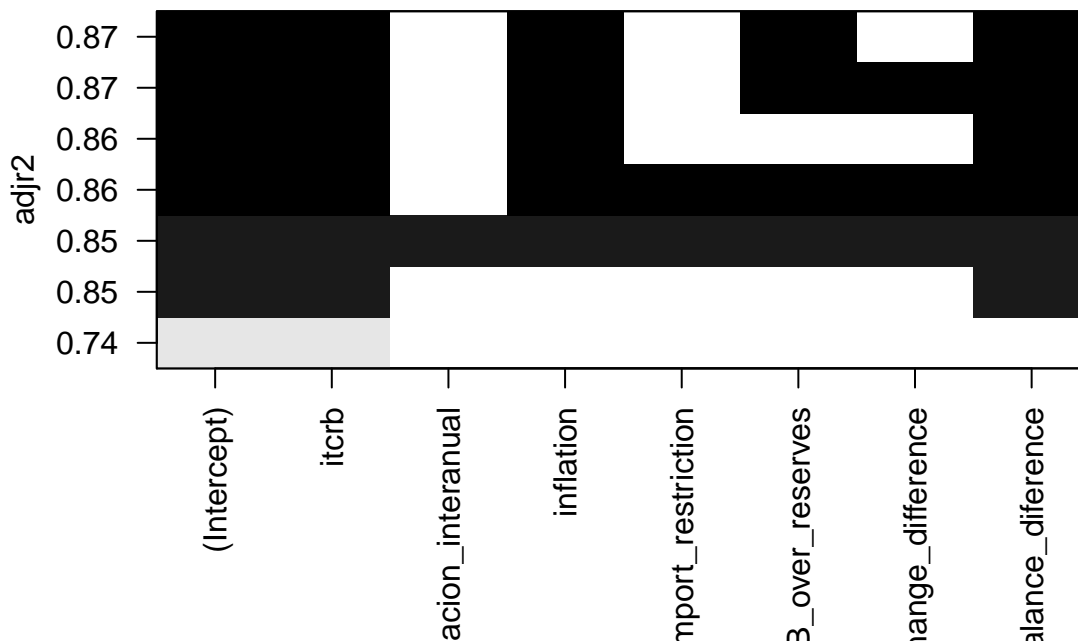
```
## Subset selection object
## Call: regsubsets.formula(sold_units$num_units ~ ., sold_units_user,
##      nvmax = 9)
## 7 Variables (and intercept)
##
##              Forced in Forced out
## itcrb              FALSE      FALSE
## devaluacion_interannual    FALSE      FALSE
## inflation            FALSE      FALSE
## import_restriction        FALSE      FALSE
## PIB_over_reserves         FALSE      FALSE
## exchange_difference        FALSE      FALSE
## industry_trade_balance_diference    FALSE      FALSE
## 1 subsets of each size up to 7
## Selection Algorithm: exhaustive
##
##      itcrb devaluacion_interannual inflation import_restriction
## 1  ( 1 ) "*"  " "                  " "          " "
## 2  ( 1 ) "*"  " "                  " "          " "
## 3  ( 1 ) "*"  " "                  "*"          " "
## 4  ( 1 ) "*"  " "                  "*"          " "
## 5  ( 1 ) "*"  " "                  "*"          " "
## 6  ( 1 ) "*"  " "                  "*"          "*"
## 7  ( 1 ) "*"  "*"                  "*"          "*"
##
##      PIB_over_reserves exchange_difference industry_trade_balance_diference
## 1  ( 1 ) " "          " "          " "
## 2  ( 1 ) " "          " "          "*"
## 3  ( 1 ) " "          " "          "*"
## 4  ( 1 ) "*"          " "          "*"
## 5  ( 1 ) "*"          "*"          "*"
## 6  ( 1 ) "*"          "*"          "*"
## 7  ( 1 ) "*"          "*"          "*"
##
```



```
plot(summary(sold_units_user)$adjr2)
```



```
plot(sold_units_user, scale = "adjr2")
```



```
best_adjr2<-which.max(summary(sold_units_user)$adjr2)
user_subset_coef<-names(coef(sold_units_user, best_adjr2))
sold_units_user_subset<-lm(sold_units[,names(sold_units)%in%
                           c("num_units", "itcrb","inflation",
                             "PIB_over_reserves",
                             "exchange_difference",
                             "industry_trade_balance_difference")])
summary(sold_units_user_subset)
```

```
##
## Call:
## lm(formula = sold_units[, names(sold_units) %in% c("num_units",
##           "itcrb", "inflation", "PIB_over_reserves", "exchange_difference",
##           "industry_trade_balance_difference")])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -129523  -47420   -2623    56492   155470
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1182397.40   159240.99    7.425 3.23e-06 ***
## itcrb        -5345.13     954.15   -5.602 6.52e-05 ***
## inflation   -230768.04   155643.52   -1.483  0.1603
## PIB_over_reserves -8777.76    7379.61   -1.189  0.2540
```

```
## exchange_difference      -50892.32  108607.48  -0.469   0.6466
## industry_trade_balance_difference  -41.91    15.34  -2.733   0.0162 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 89480 on 14 degrees of freedom
## Multiple R-squared:  0.9012, Adjusted R-squared:  0.866
## F-statistic: 25.55 on 5 and 14 DF,  p-value: 1.425e-06
```

```
vif(sold_units_user_subset)
```

```
##                itcrb                inflation
##                2.800831                1.268531
##                PIB_over_reserves        exchange_difference
##                1.778757                1.905707
## industry_trade_balance_difference
##                2.293935
```