Assignment 1

working title

October 7, 2016

```
Coal <- read.csv("https://www.eia.gov/totalenergy/data/browser/csv.cfm?tbl=T06.01")

library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
CoalProduction <- Coal %>% filter(MSN == "CLPRPUS")
library(stringr)
CoalProduction <- CoalProduction %>% filter(str_sub(as.character(YYYYMM),5 ) == "13")
CoalProduction <- CoalProduction %>% select(YYYYMM, Value)
names(CoalProduction) <- c("RawYear", "ProductionKShortTon")
CoalProduction$ProductionKShortTon <- as.numeric(as.character(CoalProduction$ProductionKShortTon))
CoalProduction <- CoalProduction  %>% mutate(Year = as.numeric(str_sub(as.character(RawYear),0,4 )))
library(Quandl)
```

```
## Loading required package: xts

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

##
## Attaching package: 'xts'

## The following objects are masked from 'package:dplyr':
##
##     first, last
```

```r
CPI <- Quandl("UIFS/CPI_USA")
Prices <- Quandl("EPI/152")

summary(CPI)
```

```
##       Date            CPI ALL ITEMS CITY AVERAGE (INDEX NUMBER)
##  Min.   :1948-12-31   Min.   : 12.20
##  1st Qu.:1964-03-31   1st Qu.: 15.73
##  Median :1979-07-01   Median : 35.29
##  Mean   :1979-07-01   Mean   : 46.62
##  3rd Qu.:1994-09-30   3rd Qu.: 75.42
##  Max.   :2009-12-31   Max.   :110.25
##
##  CPI % CHANGE (PERCENT PER ANNU)
##  Min.   :-1.412
##  1st Qu.: 1.859
##  Median : 2.991
##  Mean   : 3.708
##  3rd Qu.: 4.317
##  Max.   :13.509
##  NA's   :1
```

```r
summary(Prices)
```

```
##       Year            Price (U.S. Dollars)
##  Min.   :1949-01-01   Min.   :16.78
##  1st Qu.:1963-01-01   1st Qu.:20.19
##  Median :1977-01-01   Median :25.02
##  Mean   :1976-12-31   Mean   :27.85
##  3rd Qu.:1991-01-01   3rd Qu.:31.52
##  Max.   :2005-01-01   Max.   :50.92
```

```r
CoalProduction$RawYear <- as.numeric(str_sub(CoalProduction$RawYear,0,4))
CPI$Date <- as.numeric(str_sub(CPI$Date,0,4))
names(CoalProduction) <- c("Year", "ProductionKShortTon", "RawYear")
CoalProduction$RawYear <- NULL
names(CPI) <- c("Year", "CPI Value", "% change")

summary(CoalProduction)
```

```
##       Year      ProductionKShortTon
##  Min.   :1949   Min.   : 420423
##  1st Qu.:1966   1st Qu.: 558547
##  Median :1982   Median : 829700
##  Mean   :1982   Mean   : 796953
##  3rd Qu.:1998   3rd Qu.:1033239
##  Max.   :2015   Max.   :1171809
```

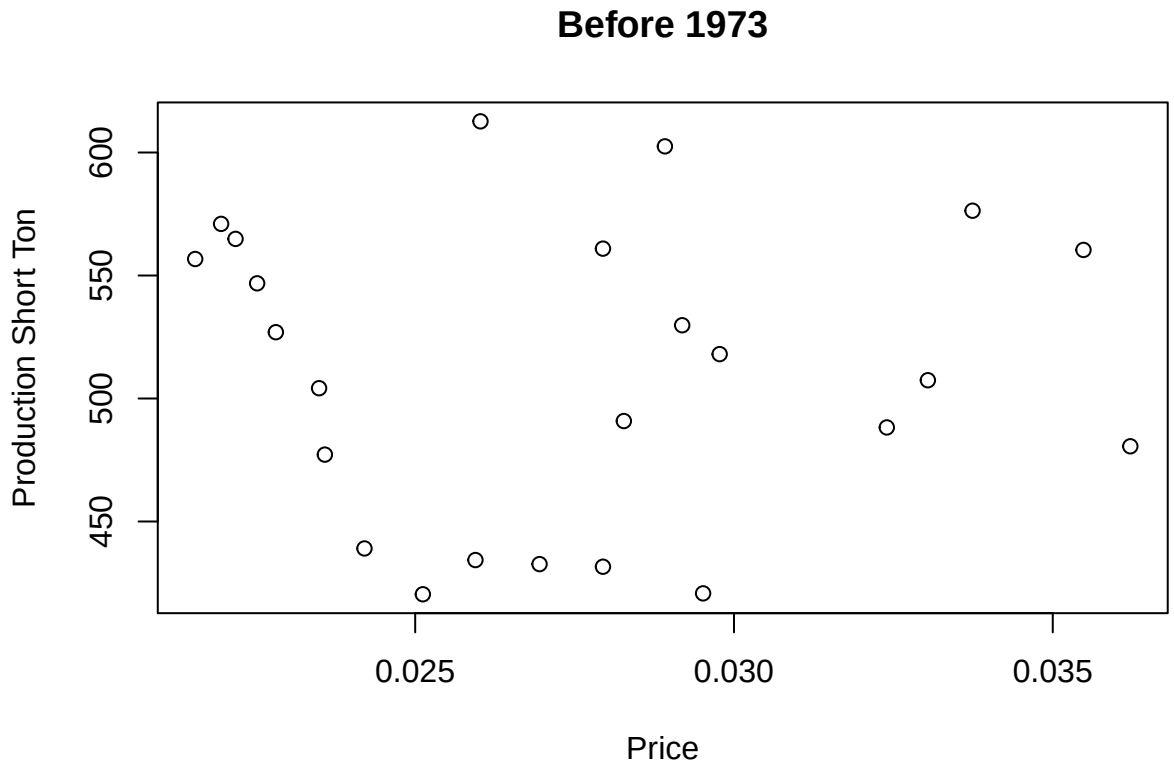
```r
CoalProduction <- inner_join(CPI, CoalProduction, by = "Year")
summary(CoalProduction)
```

```
##       Year        CPI Value         % change      ProductionKShortTon
##  Min.   :1949   Min.   : 12.32   Min.   :-1.412   Min.   : 420423
##  1st Qu.:1964   1st Qu.: 15.88   1st Qu.: 1.859   1st Qu.: 546822
##  Median :1979   Median : 37.17   Median : 2.991   Median : 781134
##  Mean   :1979   Mean   : 47.18   Mean   : 3.708   Mean   : 775698
##  3rd Qu.:1994   3rd Qu.: 75.90   3rd Qu.: 4.317   3rd Qu.:1032974
##  Max.   :2009   Max.   :110.25   Max.   :13.509   Max.   :1171809
```

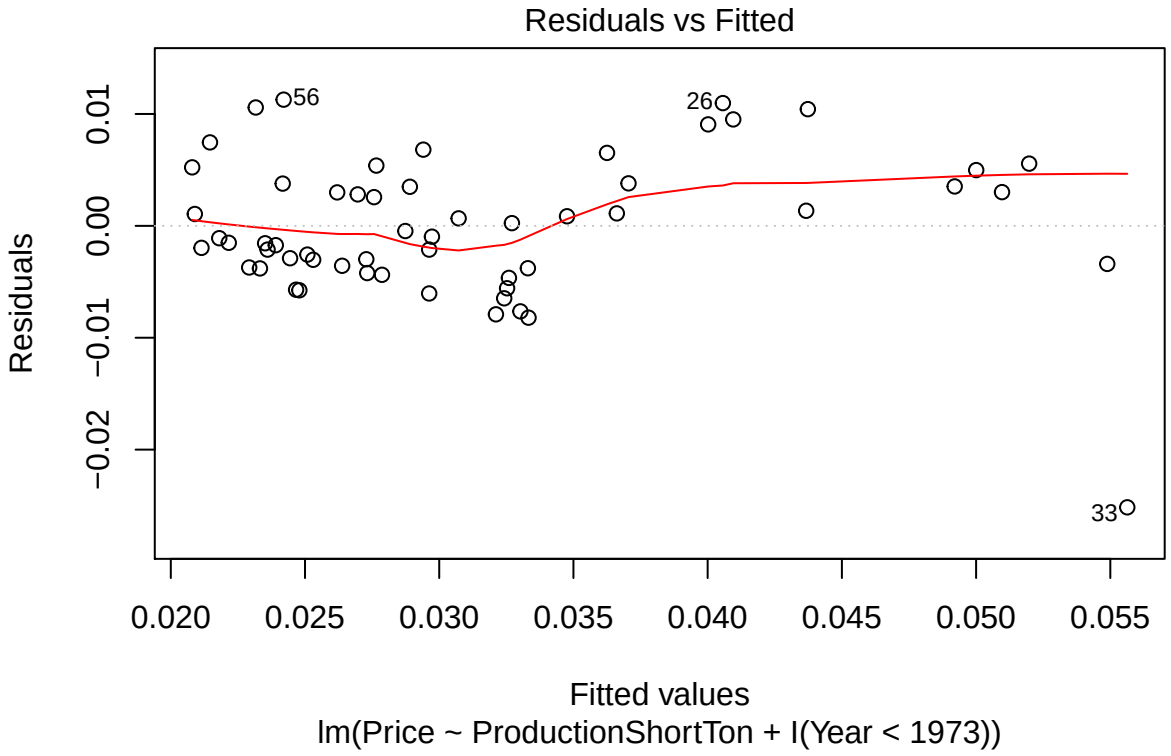
```
Prices$Year <- as.numeric(str_sub(Prices$Year,0,4))
names(Prices) <- c("Year", "Price")
CoalMarket <- inner_join(Prices, CoalProduction, by ="Year")

summary(CoalMarket)
```

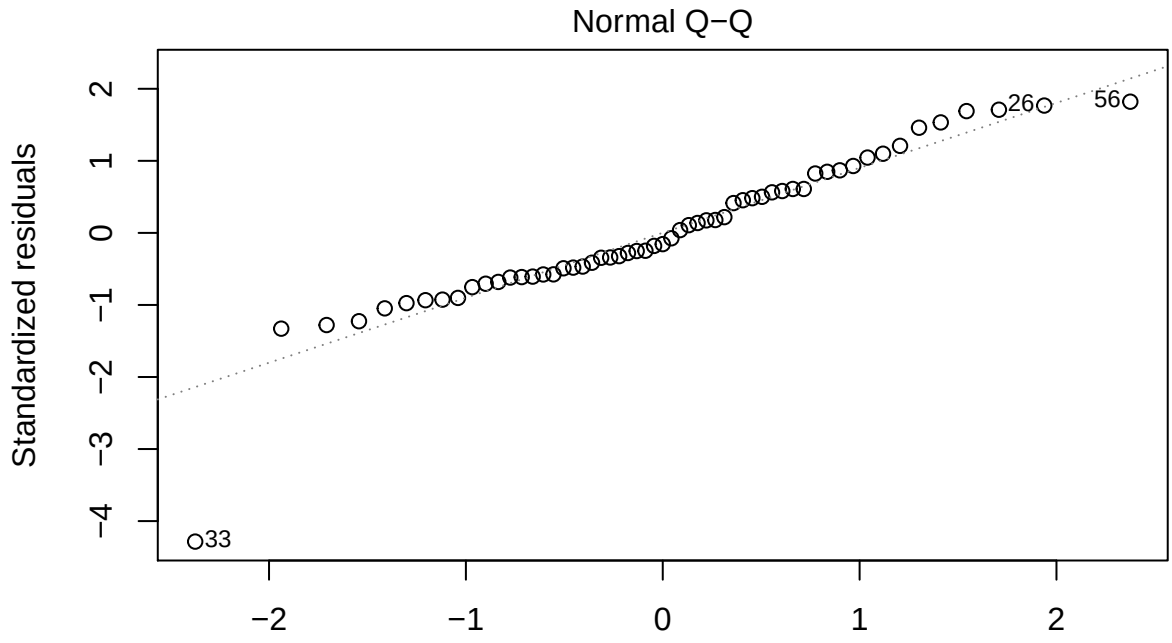
```
##       Year          Price         CPI Value         % change
##  Min.   :1949   Min.   :16.78   Min.   : 12.32   Min.   :-1.412
##  1st Qu.:1963   1st Qu.:20.19   1st Qu.: 15.68   1st Qu.: 1.859
##  Median :1977   Median :25.02   Median : 31.03   Median : 2.991
##  Mean   :1977   Mean   :27.85   Mean   : 42.96   Mean   : 3.800
##  3rd Qu.:1991   3rd Qu.:31.52   3rd Qu.: 69.74   3rd Qu.: 4.827
##  Max.   :2005   Max.   :50.92   Max.   :100.00   Max.   :13.509
##  ProductionKShortTon
##  Min.   : 420423
##  1st Qu.: 529774
##  Median : 684913
##  Mean   : 750200
##  3rd Qu.: 995984
##  Max.   :1131498
```

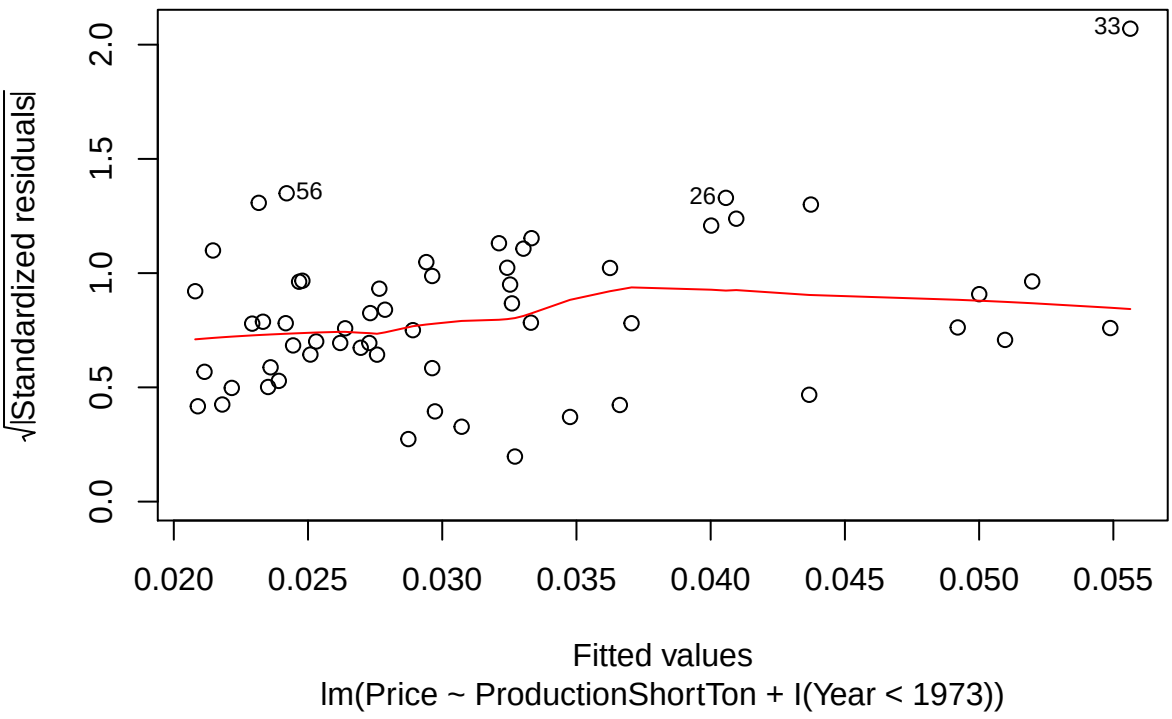
```
library(dplyr)
CoalMarket$Price <- ((CoalMarket$Price)*1.13)/1000
CoalMarket$ProductionKShortTon <- ((CoalMarket$ProductionKShortTon)/1000)
names(CoalMarket)<- c("Year", "Price", "CPI Value", "% Change", "ProductionShortTon")

#Years 1973 to 1977 are excluded due to oil crisis (OPEC market manipulaion)
CoalMarket$OPEC_Years <- factor(with(CoalMarket, ifelse((Year > 1972  & Year < 1978), 0, 1)))

#Before OPEC market manipulation
CoalRegressionA <-  lm(Price ~ ProductionShortTon +I(Year<1973), data = CoalMarket)
plot((CoalMarket$Price)[(CoalMarket$Year)<1973], (CoalMarket$ProductionShortTon)[(CoalMarket$Year)<1973]
```

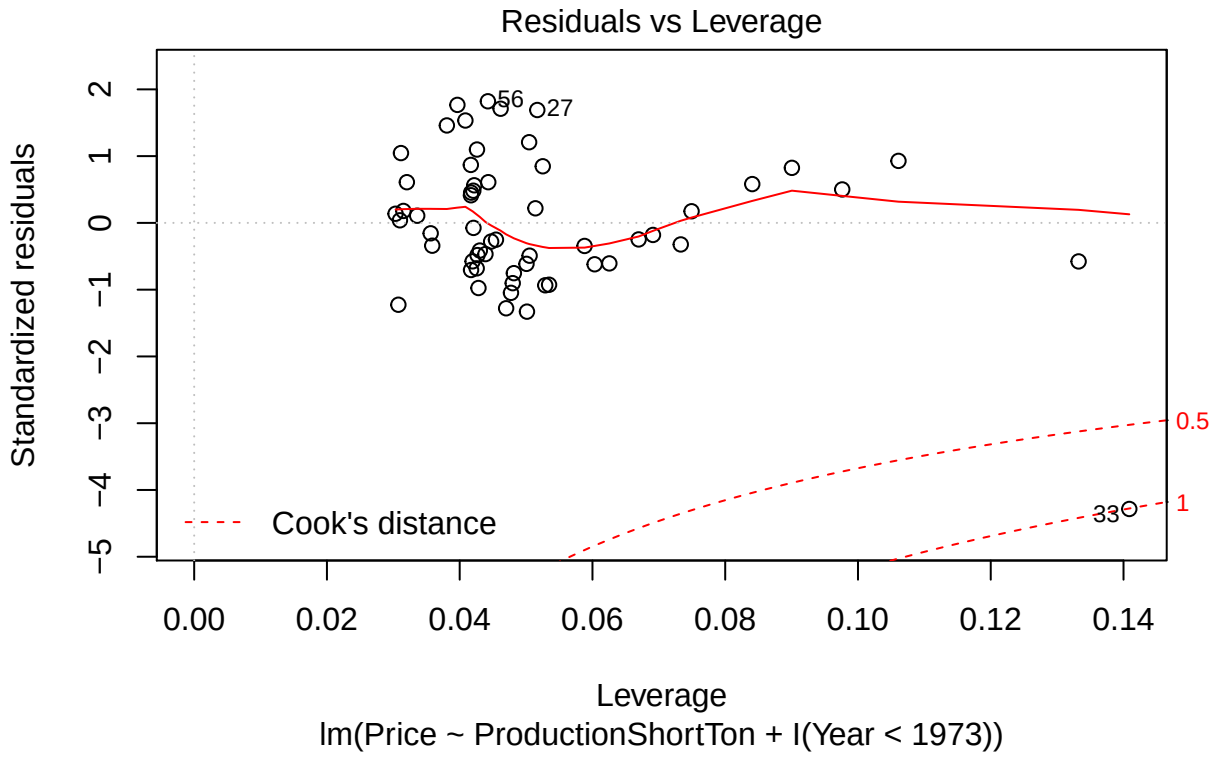



```
plot(CoalRegressionA)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Price ~ ProductionShortTon + I(Year < 1973))
33
26
56

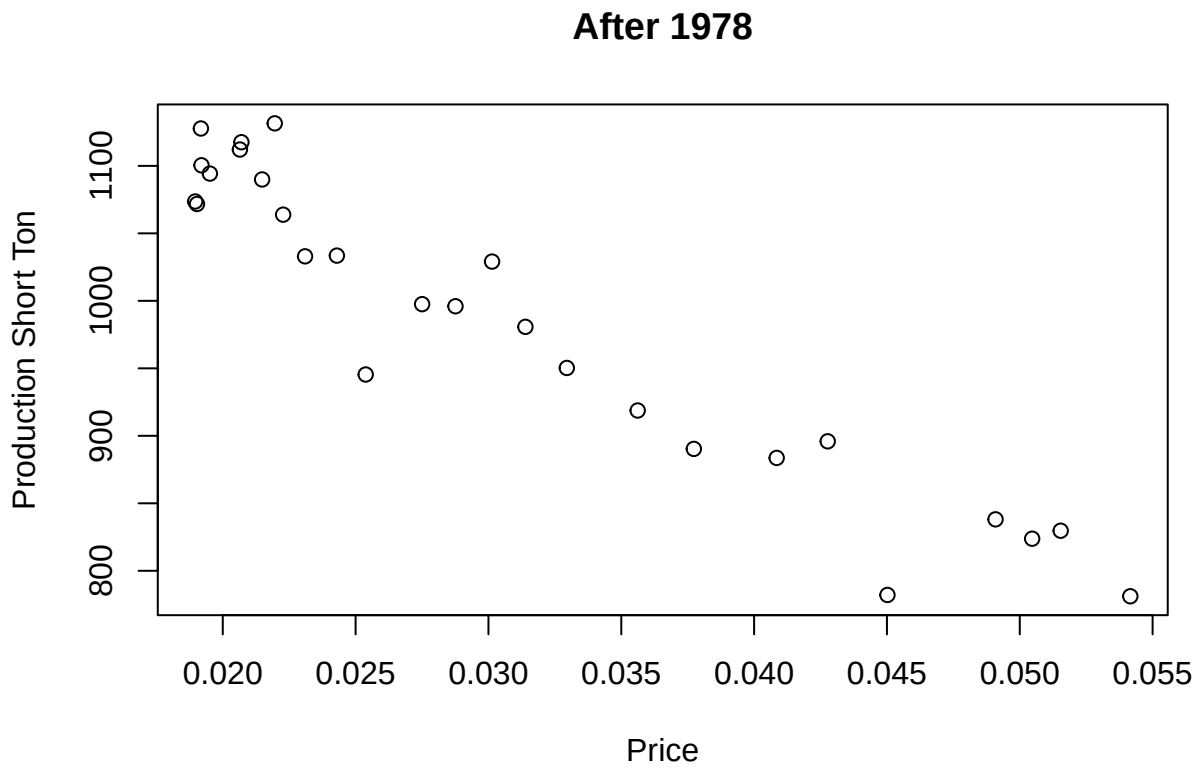

Scale−Location
√|Standardized residuals|
Fitted values
lm(Price ~ ProductionShortTon + I(Year < 1973))
33
56
26



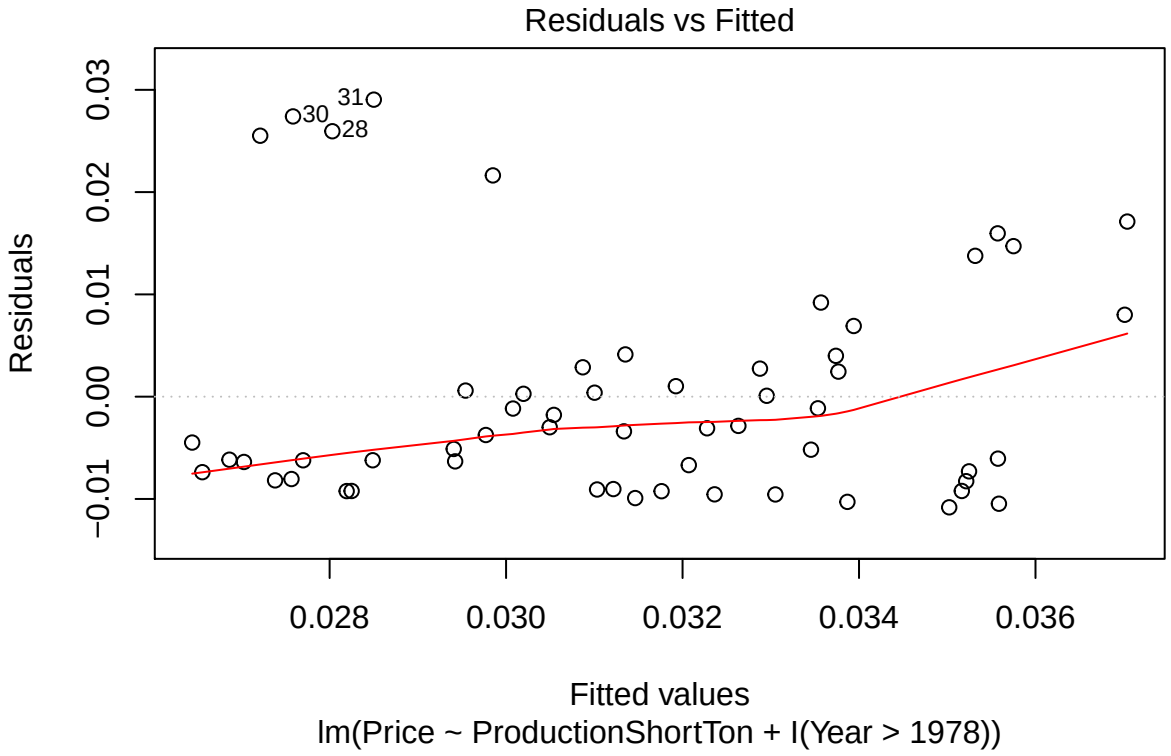

```
summary(CoalRegressionA)
```

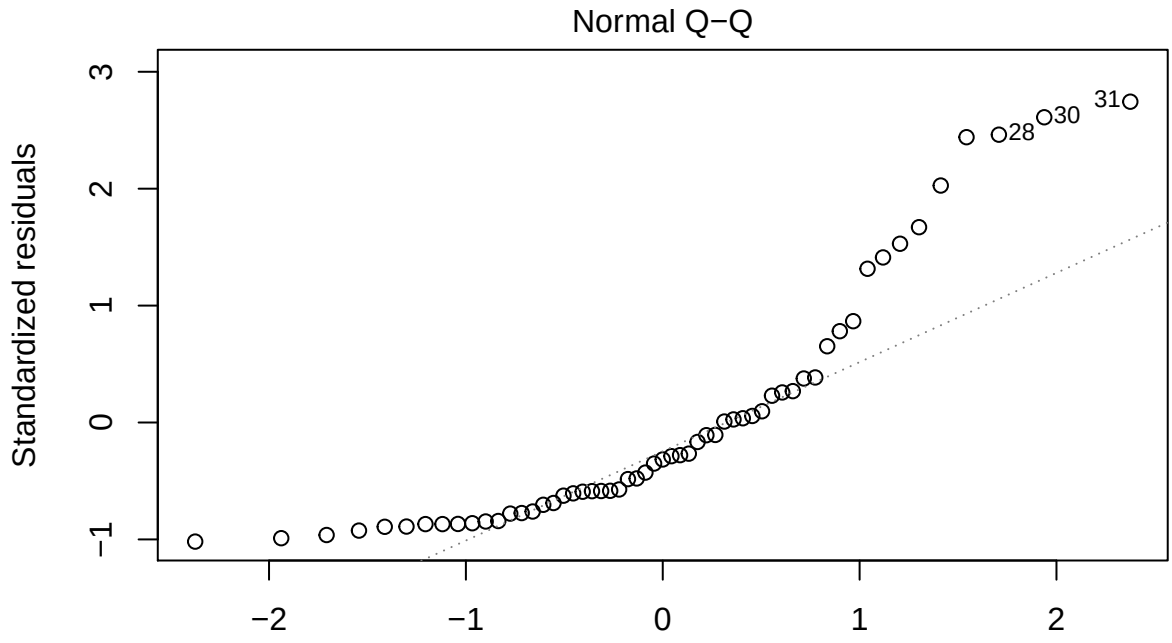
```
##
## Call:
## lm(formula = Price ~ ProductionShortTon + I(Year < 1973), data = CoalMarket)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0251563 -0.0037215 -0.0009698  0.0037760  0.0112785
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          9.465e-02  6.076e-03   15.58  < 2e-16 ***
## ProductionShortTon -6.518e-05  6.464e-06  -10.09 5.07e-14 ***
## I(Year < 1973)TRUE -3.392e-02  3.169e-03  -10.70 5.86e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.006335 on 54 degrees of freedom
## Multiple R-squared:  0.6867, Adjusted R-squared:  0.6751
## F-statistic: 59.19 on 2 and 54 DF,  p-value: 2.452e-14
```

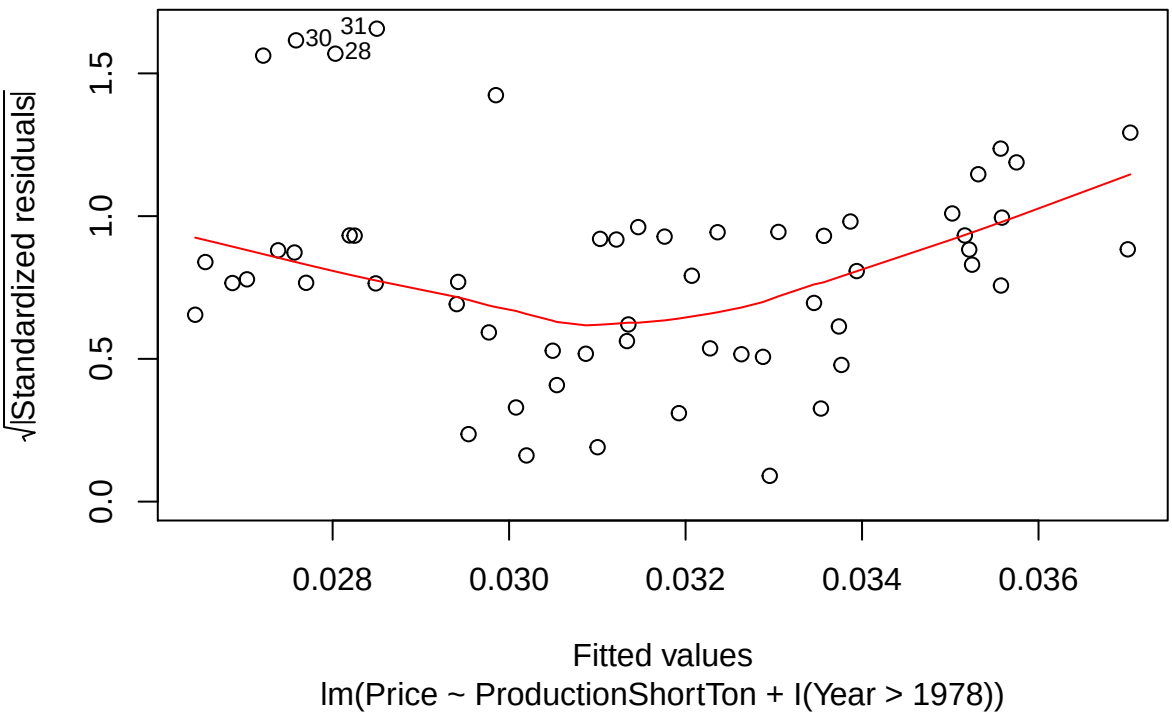
```
#After OPEC market manipulation
CoalRegressionB <-  lm(Price ~ ProductionShortTon +I(Year>1978), data = CoalMarket)
plot((CoalMarket$Price)[(CoalMarket$Year)>1978], (CoalMarket$ProductionShortTon)[(CoalMarket$Year)>1978]
```

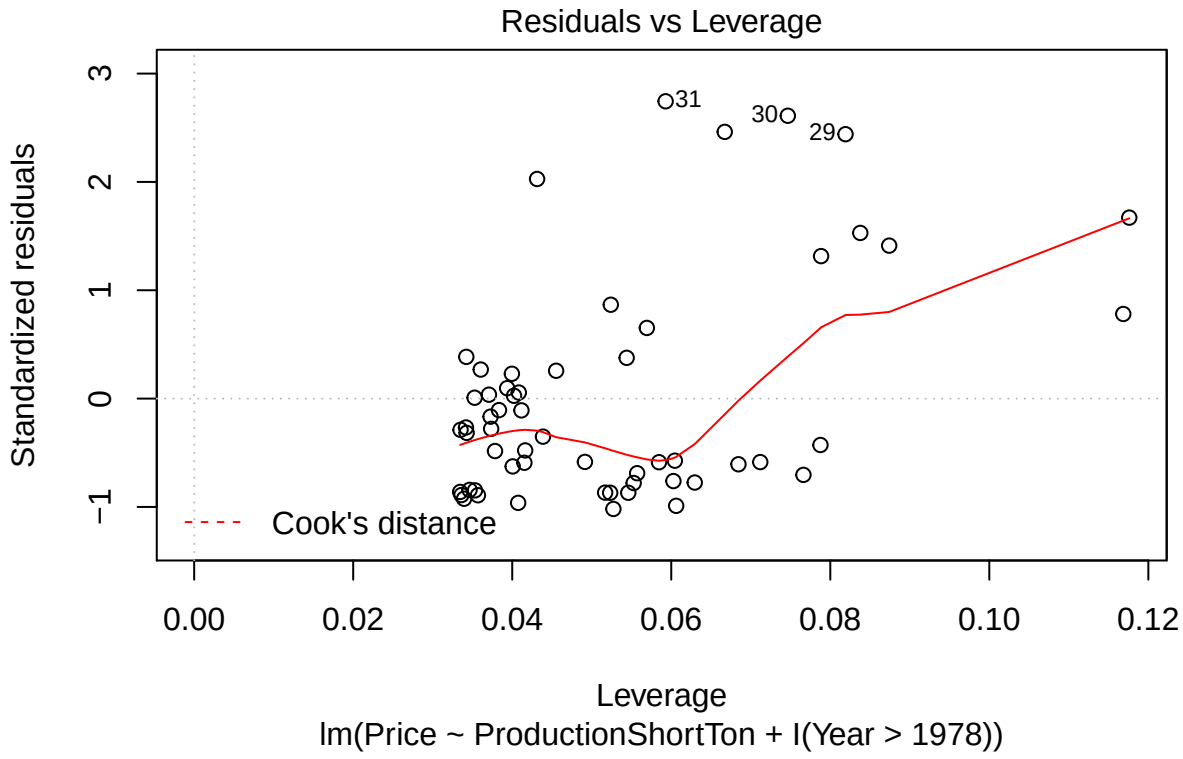



```
plot(CoalRegressionB)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Price ~ ProductionShortTon + I(Year > 1978))
31
30
28


Scale−Location
√|Standardized residuals|
Fitted values
lm(Price ~ ProductionShortTon + I(Year > 1978))
31
30
28




```
summary(CoalRegressionB)
```

```
##
## Call:
## lm(formula = Price ~ ProductionShortTon + I(Year > 1978), data = CoalMarket)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.010818 -0.008053 -0.003391  0.002872  0.029039
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)          4.830e-02  8.431e-03   5.729 4.62e-07 ***
## ProductionShortTon -3.025e-05  1.520e-05  -1.990   0.0517 .
## I(Year > 1978)TRUE  1.237e-02  7.369e-03   1.678   0.0991 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01091 on 54 degrees of freedom
## Multiple R-squared:  0.07074,    Adjusted R-squared:  0.03632
## F-statistic: 2.055 on 2 and 54 DF,  p-value: 0.1379
```