models_diversos_datasets.R

jcortes

2023-11-30

```
rm(list=ls())
```

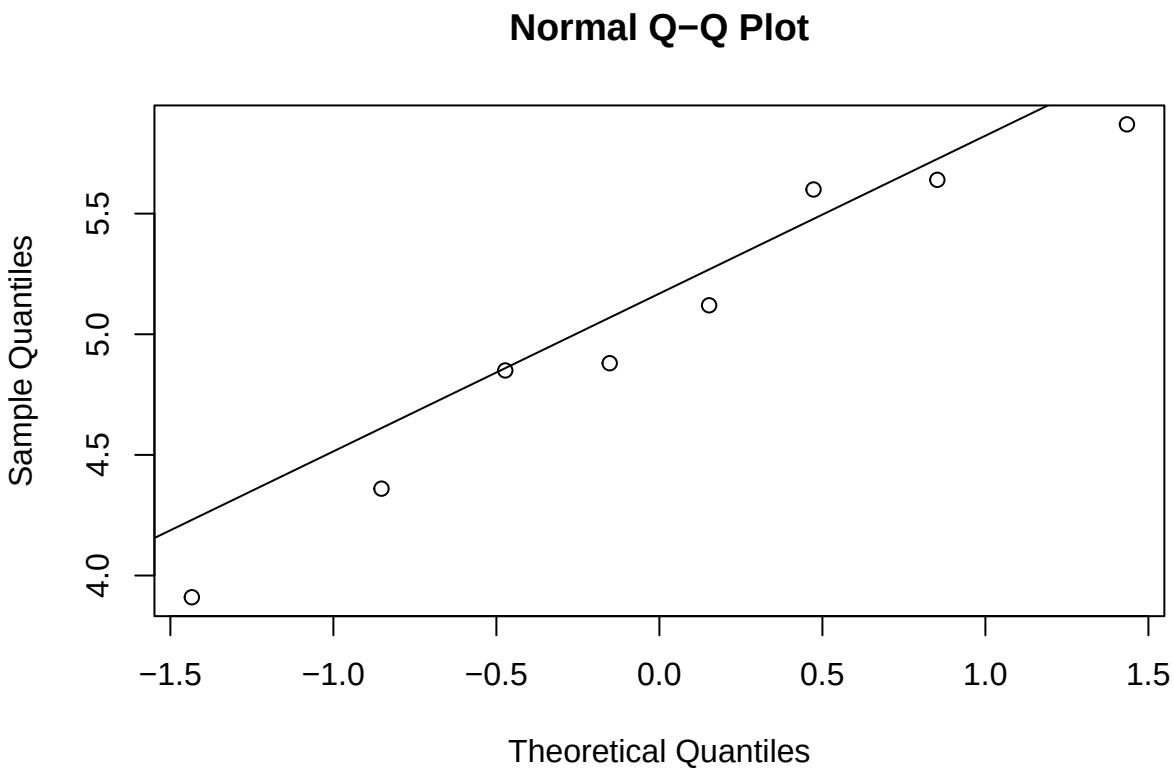
```
library(emmeans)
```

```
d <- read.csv('https://raw.githubusercontent.com/jordicortes40/PE_Bloc_D/main/Dades/temps_compressors.cs
# d <- read.csv('../Dades/temps_compressors.csv')

mod <- lm(temps~1,d)
summary(mod)
```

```
##
## Call:
## lm(formula = temps ~ 1, data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11875 -0.30125 -0.02875  0.58125  0.84125
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.0287     0.2379   21.14 1.33e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6728 on 7 degrees of freedom
```

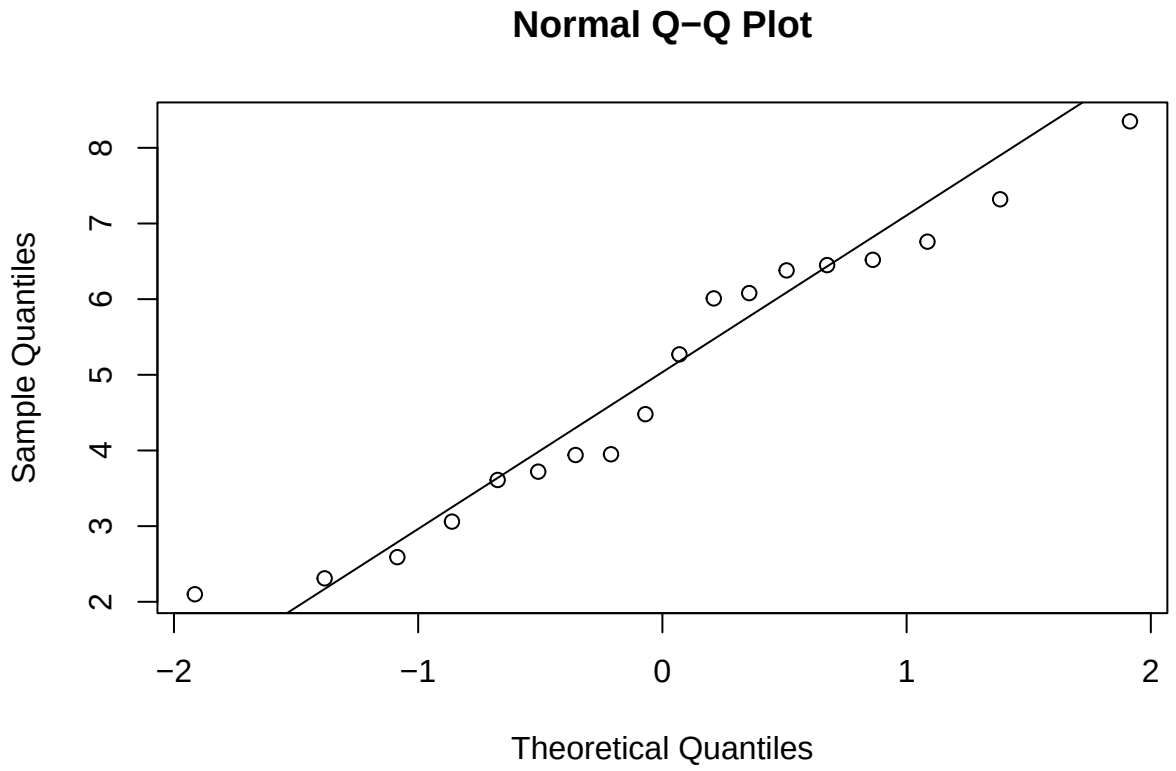
```
# Premissa normalitat
qqnorm(d$temps)
qqline(d$temps)
```



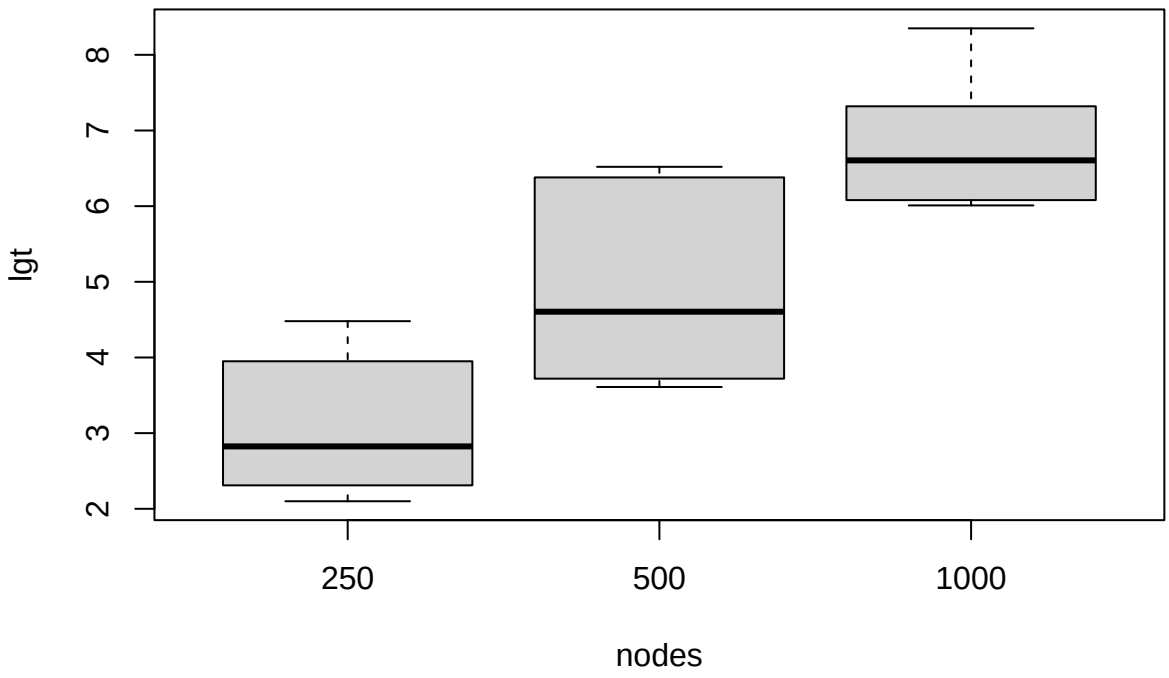

```
d <- read.csv('https://raw.githubusercontent.com/jordicortes40/PE_Bloc_D/main/Dades/algoritme_dijkstra.
# d <- read.csv('../Dades/algoritme_dijkstra.csv')
mod <- lm(lgt~as.factor(nodes),d)
summary(mod)
```

```
##
## Call:
## lm(formula = lgt ~ as.factor(nodes), data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2967 -0.8067 -0.2233  0.7742  1.6133
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)            3.0817     0.4394   7.014 4.18e-06 ***
## as.factor(nodes)500    1.8250     0.6214   2.937   0.0102 *
## as.factor(nodes)1000   3.7467     0.6214   6.030 2.31e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.076 on 15 degrees of freedom
## Multiple R-squared:  0.708,  Adjusted R-squared:  0.669
## F-statistic: 18.18 on 2 and 15 DF,  p-value: 9.789e-05
```

```r
# Premissa normalitat
qqnorm(d$lgt)
qqline(d$lgt)
```




```r
# Premissa homoscedasticitat
boxplot(lgt~nodes,d)
```




```
# IC95% per les mitjanes dels temps dels nodes
emmeans(mod,~nodes)
```

```
##  nodes emmean    SE df lower.CL upper.CL
##    250   3.08 0.439 15     2.15     4.02
##    500   4.91 0.439 15     3.97     5.84
##   1000   6.83 0.439 15     5.89     7.76
##
## Confidence level used: 0.95
```

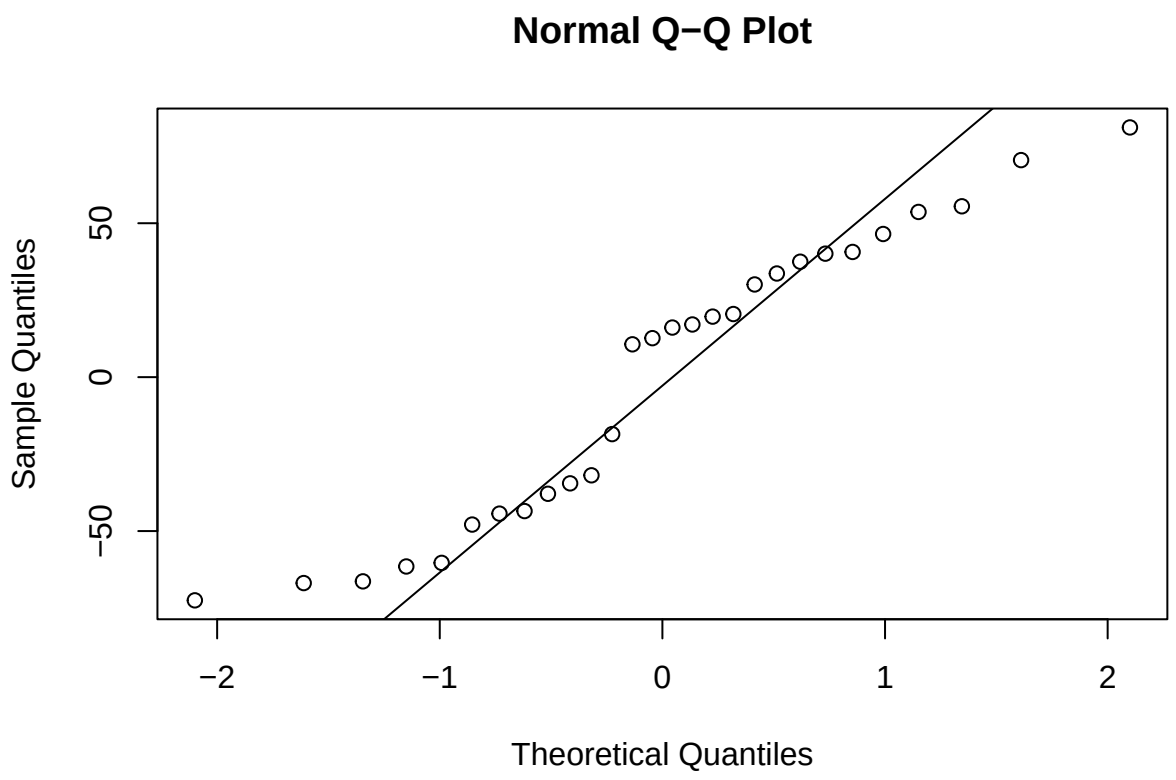
```
d <- read.csv('https://raw.githubusercontent.com/jordicortes40/PE_Bloc_D/main/Dades/recorre_arbres.csv')
# d <- read.csv('../Dades/recorre_arbres.csv')
```

```
mod1 <- lm(Temps ~ as.factor(metode),d)
summary(mod1)
```

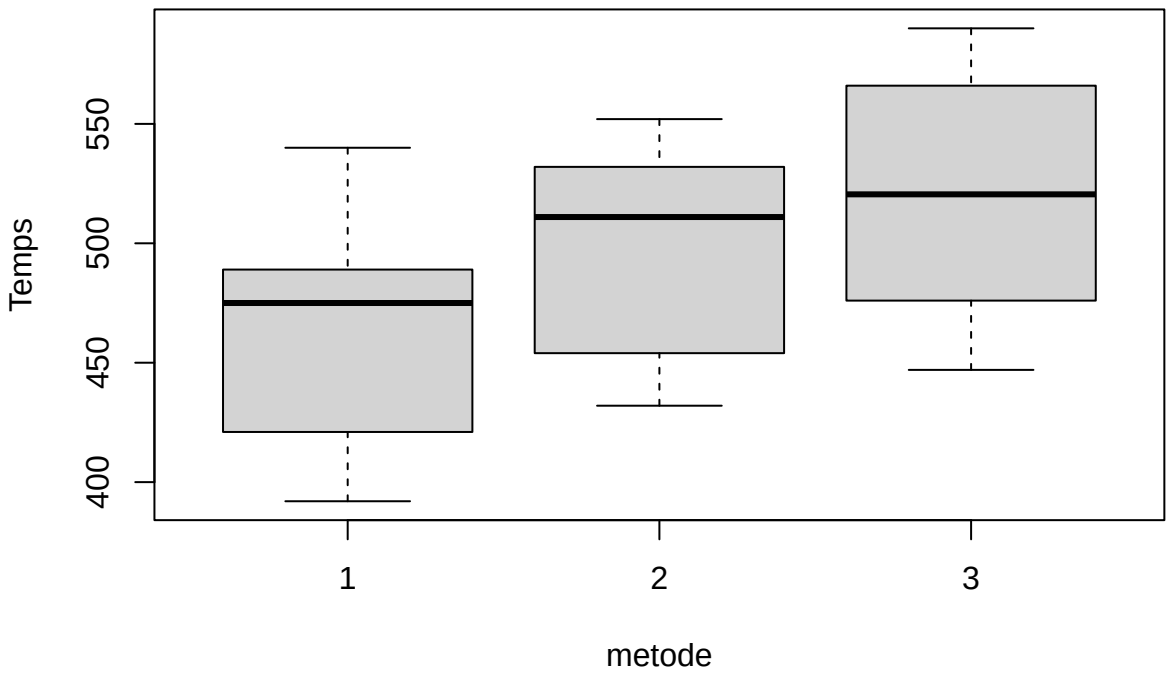
```
##
## Call:
## lm(formula = Temps ~ as.factor(metode), data = d)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -72.50 -43.71  14.39  38.15  81.11
##
```

```
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           458.89      16.29  28.171   <2e-16 ***
## as.factor(metode)2     39.44      23.04   1.712   0.0992 .
## as.factor(metode)3     60.61      22.45   2.699   0.0123 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 48.87 on 25 degrees of freedom
## Multiple R-squared:  0.2292, Adjusted R-squared:  0.1675
## F-statistic: 3.716 on 2 and 25 DF,  p-value: 0.03864
```

```r
# Premissa normalitat residus
qqnorm(resid(mod1))
qqline(resid(mod1))
```



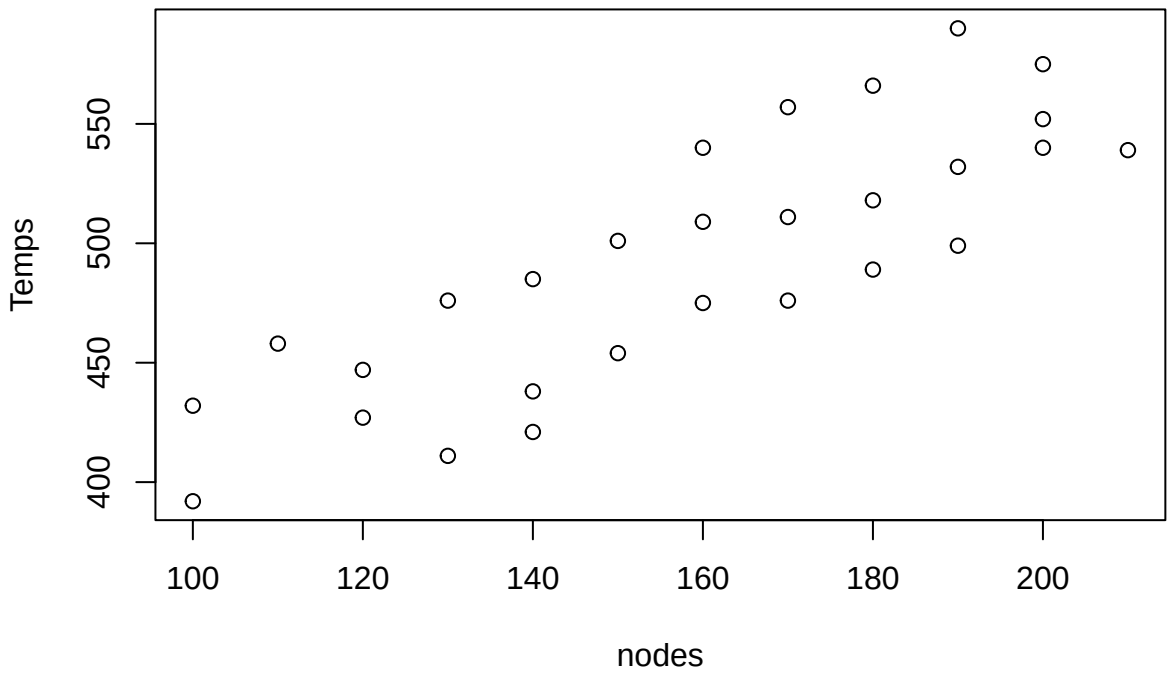

```r
# Premsissa homoscedasticitat
boxplot(Temps~metode,d)
```



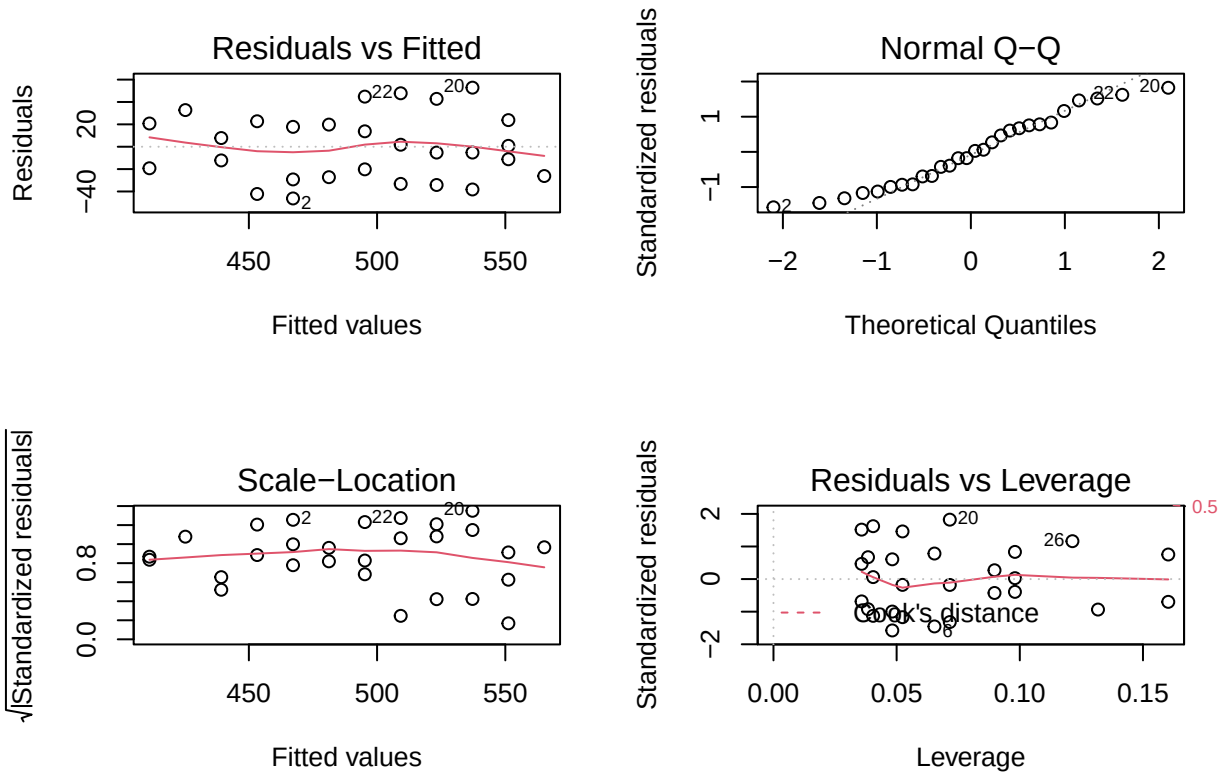

```
mod2 <- lm(Temps~nodes,d)
summary(mod2)
```

```
##
## Call:
## lm(formula = Temps ~ nodes, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -46.222 -26.449  -2.199  21.265  52.799
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 271.2801    29.2824   9.264 1.02e-09 ***
## nodes         1.3996     0.1812   7.726 3.38e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 30.06 on 26 degrees of freedom
## Multiple R-squared:  0.6966, Adjusted R-squared:  0.6849
## F-statistic: 59.69 on 1 and 26 DF,  p-value: 3.379e-08
```

```
plot(Temps~nodes,d)
```



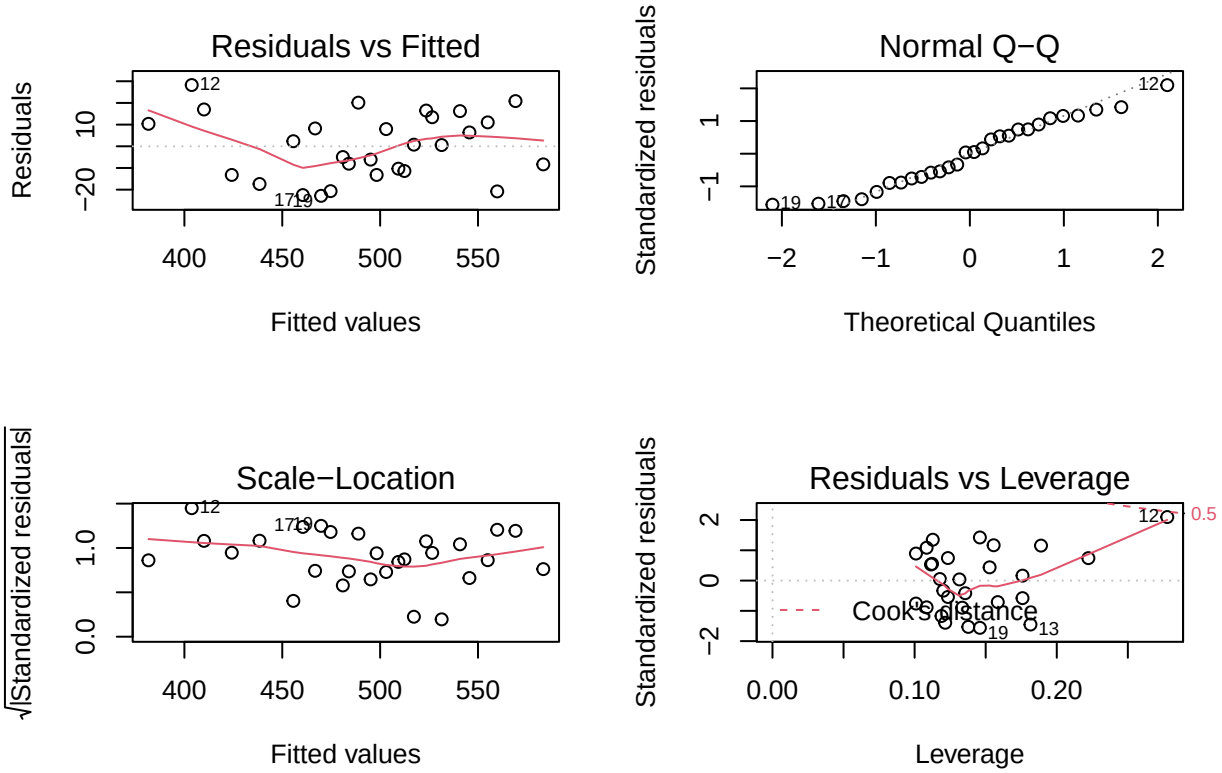

```
# Premisses
par(mfrow=c(2,2))
plot(mod2,ask=FALSE)
```




```
mod3 <- lm(Temps ~ nodes + as.factor(metode),d)
summary(mod3)
```

```
##
## Call:
## lm(formula = Temps ~ nodes + as.factor(metode), data = d)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -22.8461 -11.8574   0.6575  11.6263  28.2456
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        239.78109   15.88211  15.098 9.43e-14 ***
## nodes                1.41868    0.09699  14.628 1.87e-13 ***
## as.factor(metode)2  22.10498    7.56022   2.924  0.00743 **
## as.factor(metode)3  59.82295    7.27785   8.220 1.95e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.84 on 24 degrees of freedom
## Multiple R-squared:  0.9223, Adjusted R-squared:  0.9125
## F-statistic: 94.91 on 3 and 24 DF,  p-value: 1.892e-13
```

```r
# Premisses
par(mfrow=c(2,2))
plot(mod3,ask=FALSE)
```



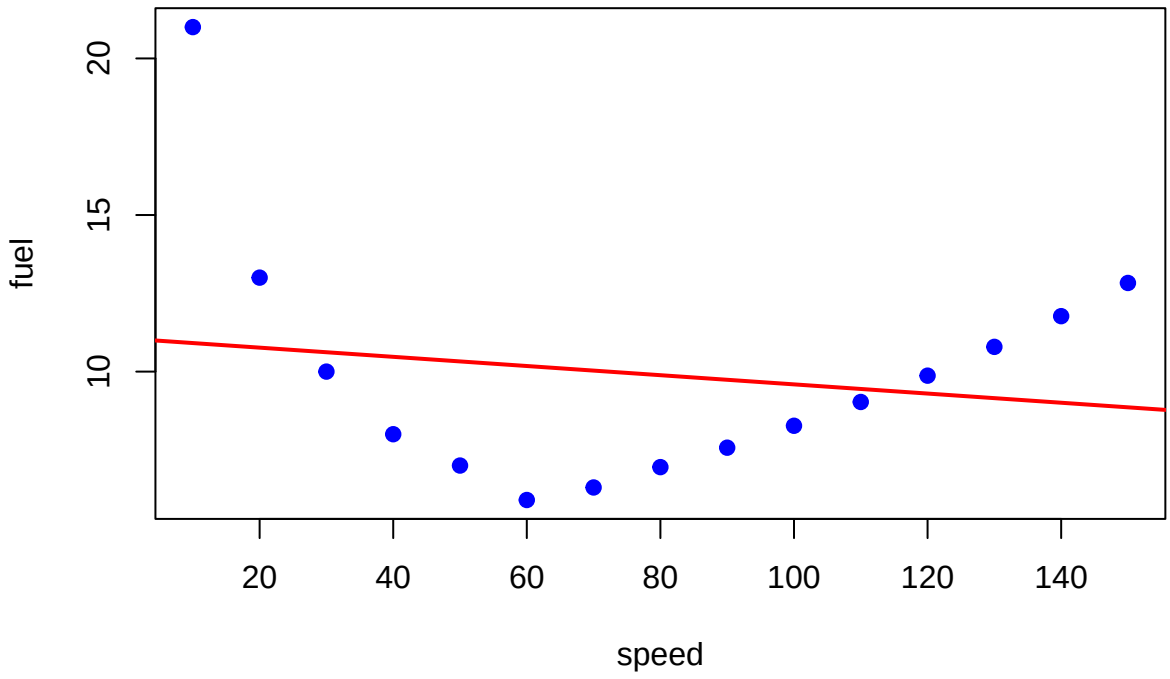

```r
d <- read.csv('https://raw.githubusercontent.com/jordicortes40/PE_Bloc_D/main/Dades/benzina_velocitat.cs
# d <- read.csv('../Dades/benzina_velocitat.csv')

# Model
mod <- lm(fuel~speed,d)
summary(mod)
```

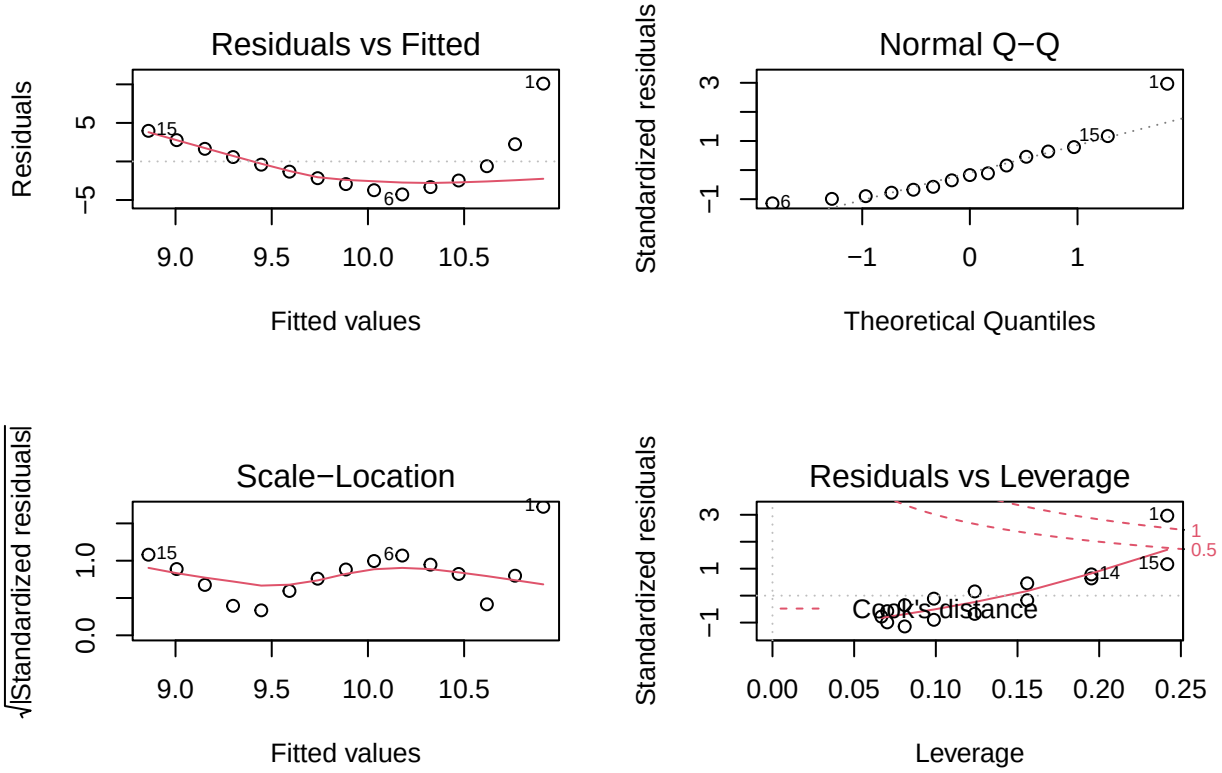
```
##
## Call:
## lm(formula = fuel ~ speed, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.2785 -2.7035 -0.6182  1.9364 10.0887
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 11.05790    2.12168   5.212 0.000168 ***
## speed       -0.01466    0.02334  -0.628 0.540810
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 3.905 on 13 degrees of freedom
## Multiple R-squared:  0.02945,    Adjusted R-squared:  -0.0452
## F-statistic: 0.3945 on 1 and 13 DF,  p-value: 0.5408
```

```r
# Grafic
par(mfrow=c(1,1))
plot(fuel~speed,d, pch=19, col='blue')
abline(mod,lwd=2,col='red')
```




```r
# Premisses
par(mfrow=c(2,2))
plot(mod,ask=FALSE)
```

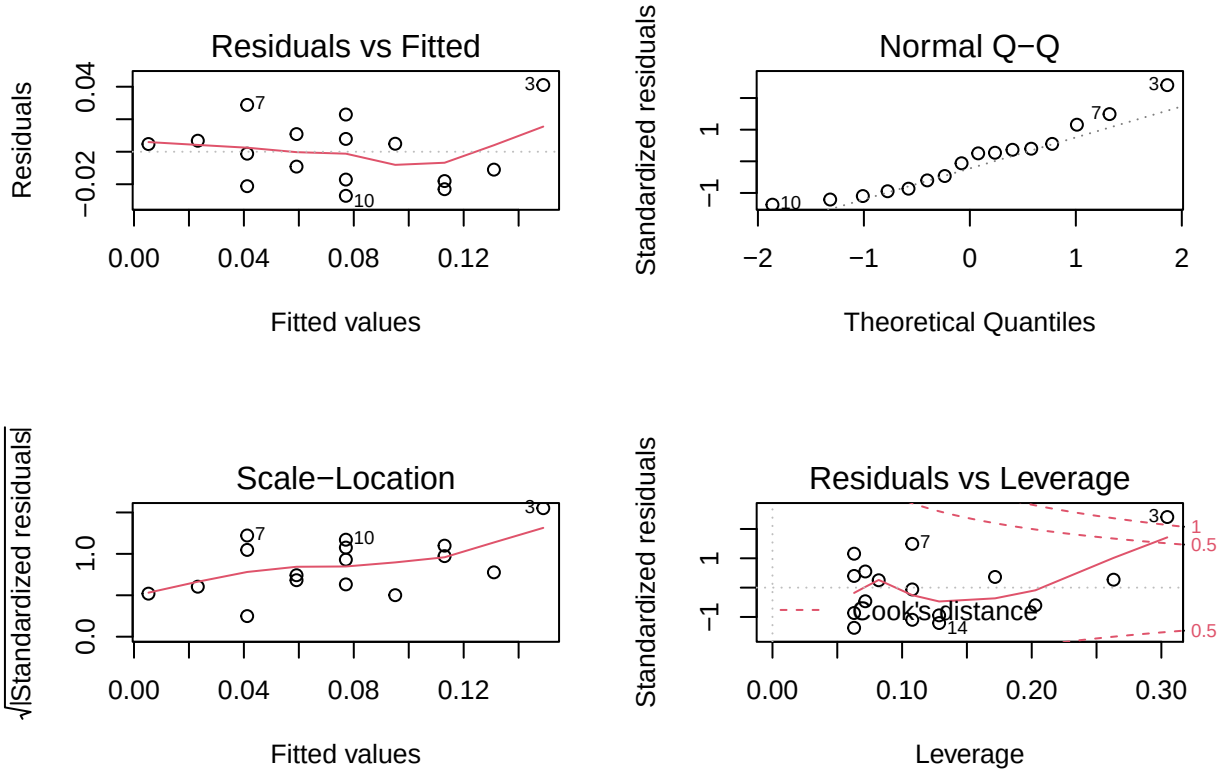



```r
d <- read.csv('https://raw.githubusercontent.com/jordicortes40/PE_Bloc_D/main/Dades/cervesa_alcohol.csv
# d <- read.csv('../Dades/cervesa_alcohol.csv')

# Model
mod <- lm(alc~n.cerv,d)
summary(mod)
```

```
##
## Call:
## lm(formula = alc ~ n.cerv, data = d)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.027118 -0.017350  0.001773  0.008623  0.041027
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.012701   0.012638  -1.005    0.332
## n.cerv       0.017964   0.002402   7.480 2.97e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02044 on 14 degrees of freedom
## Multiple R-squared:  0.7998, Adjusted R-squared:  0.7855
## F-statistic: 55.94 on 1 and 14 DF,  p-value: 2.969e-06
```

```r
# Premisses
par(mfrow=c(2,2))
plot(mod,ask=FALSE)
```



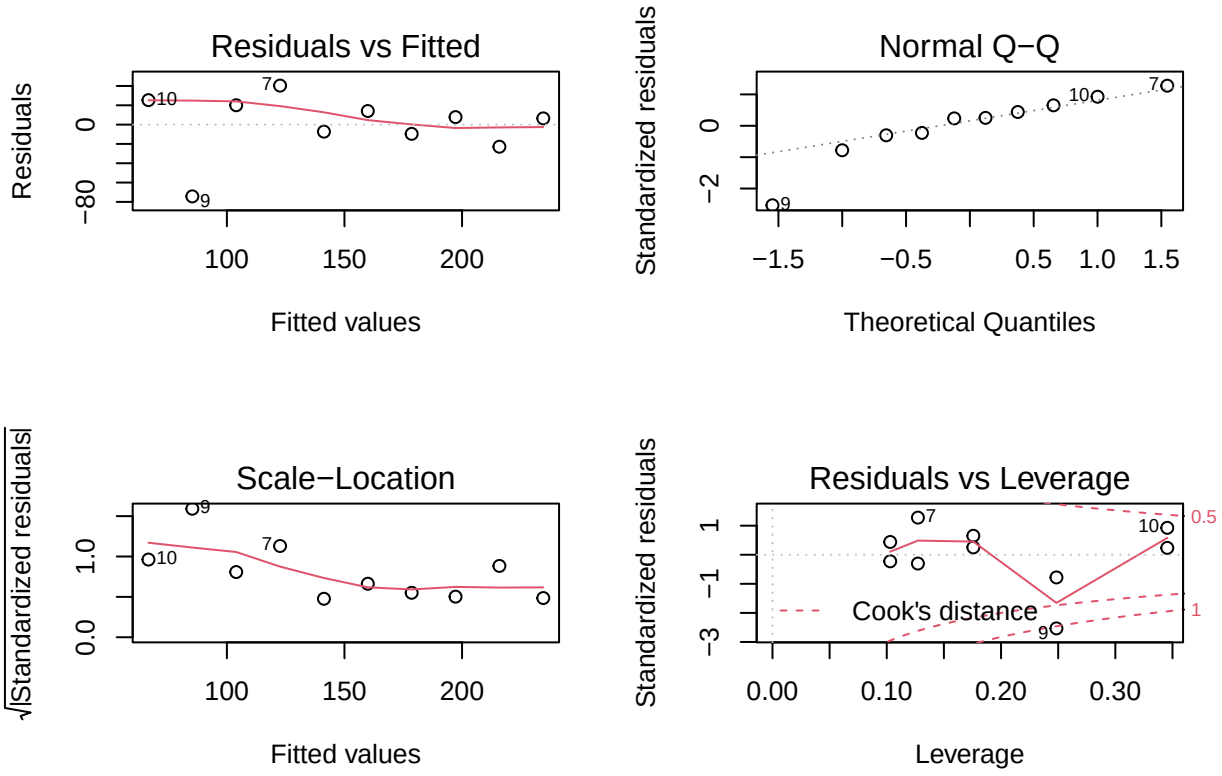

```r
d <- read.csv('https://raw.githubusercontent.com/jordicortes40/PE_Bloc_D/main/Dades/brillantor_durada.cs
# d <- read.csv('../Dades/brillantor_durada.csv')

# Model
mod <- lm(Durada~Brillantor,d)
summary(mod)
```

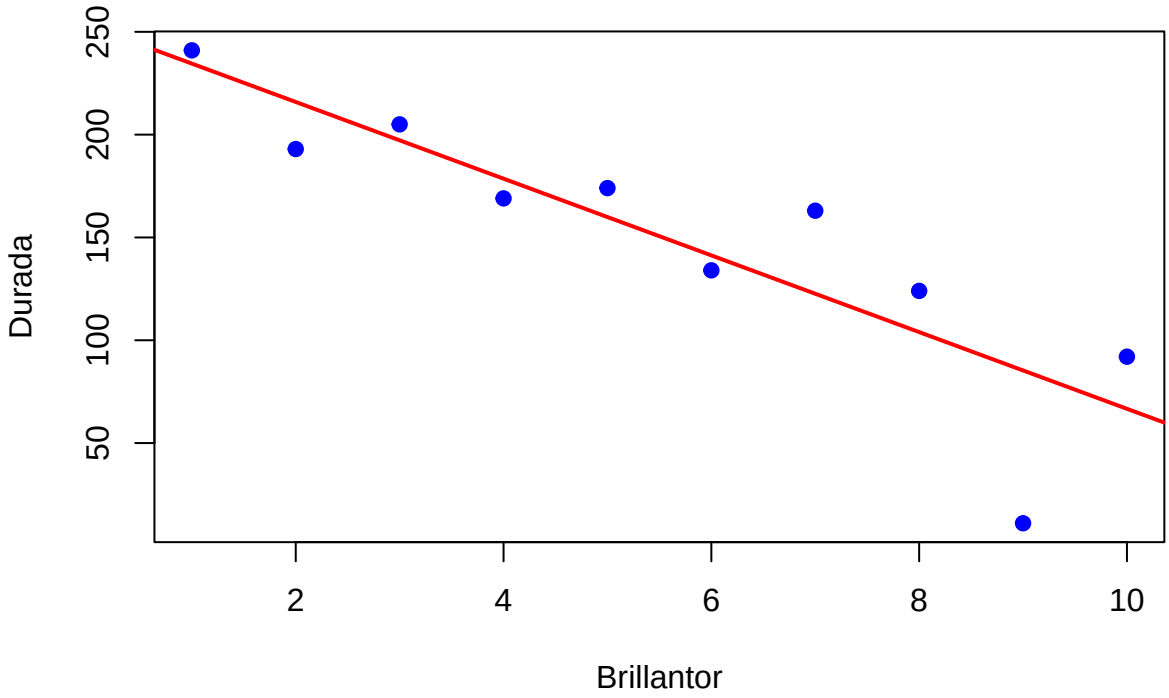
```
##
## Call:
## lm(formula = Durada ~ Brillantor, data = d)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -74.309  -9.005   7.109  18.545  40.382
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  253.200     23.122  10.951 4.29e-06 ***
## Brillantor   -18.655      3.726  -5.006  0.00104 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 33.85 on 8 degrees of freedom
## Multiple R-squared:  0.758,  Adjusted R-squared:  0.7278
## F-statistic: 25.06 on 1 and 8 DF,  p-value: 0.001045
```

```r
# Premisses
par(mfrow=c(2,2))
plot(mod,ask=FALSE)
```




```r
# Grafic
par(mfrow=c(1,1))
plot(Durada~Brillantor,d, pch=19, col='blue')
abline(mod,lwd=2,col='red')
```


Durada
250
200
150
100
50
2
4
6
8
10
Brillantor