# EDDA -Prctice
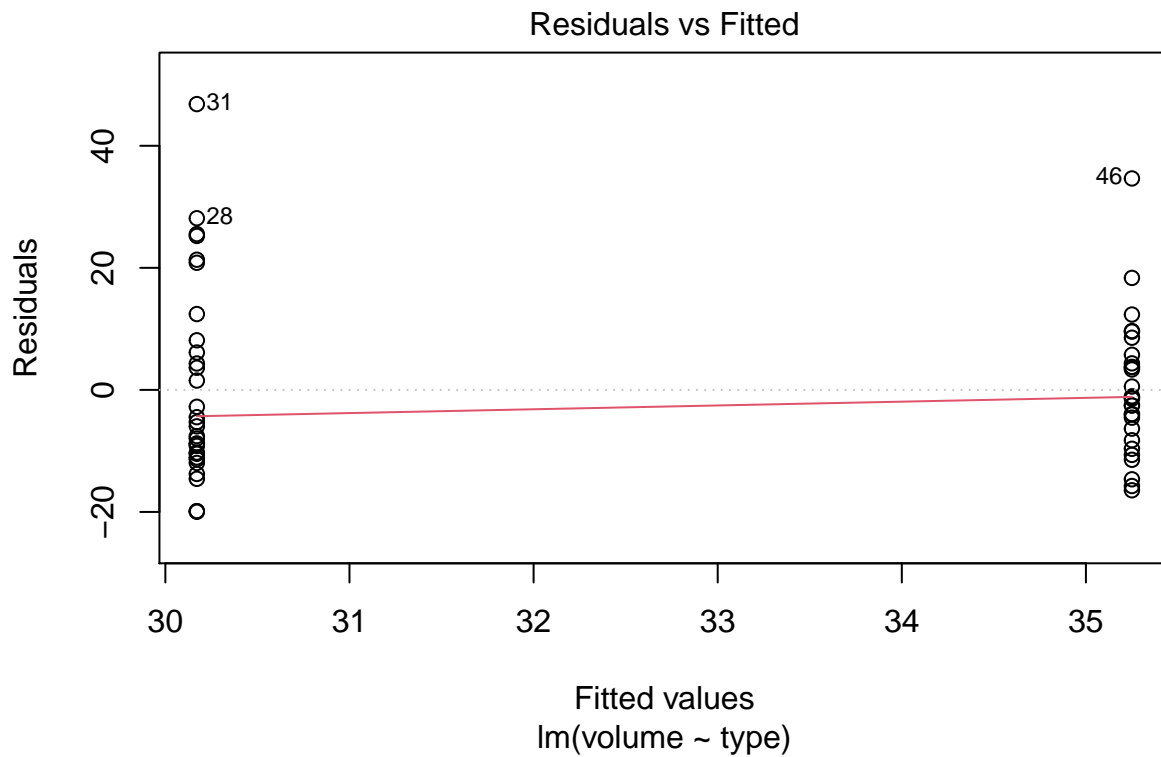
Kamiel Gülpen
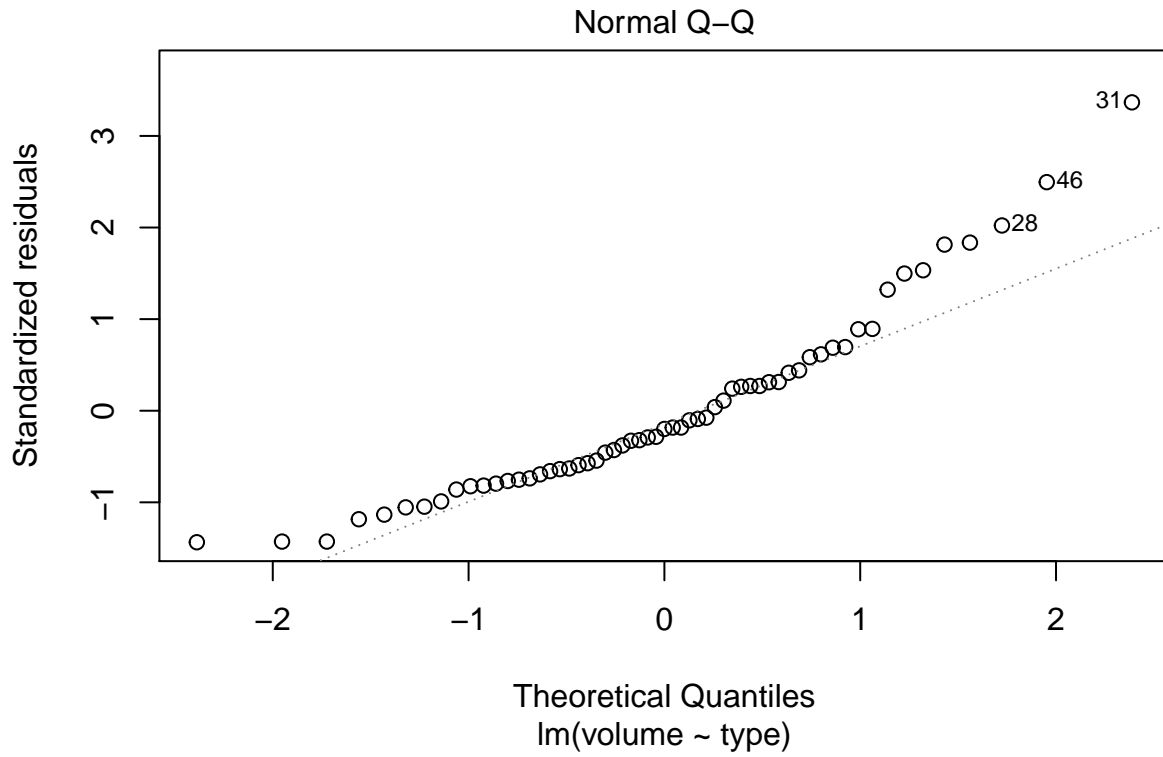
## Exercise 1

```
data <- read.table(file="treeVolume.txt", header=TRUE)


data$type=as.factor(data$type)
dataaov=lm(volume~type,data=data)
plot(dataaov, 1)
```



```
plot(dataaov, 2)
```

## Normal Q–Q



Theoretical Quantiles
lm(volume ~ type)

```
anova(dataaov)
```

```
## Analysis of Variance Table
##
## Response: volume
##           Df Sum Sq Mean Sq F value Pr(>F)
## type       1    380     380     1.9   0.17
## Residuals 57  11395     200
```
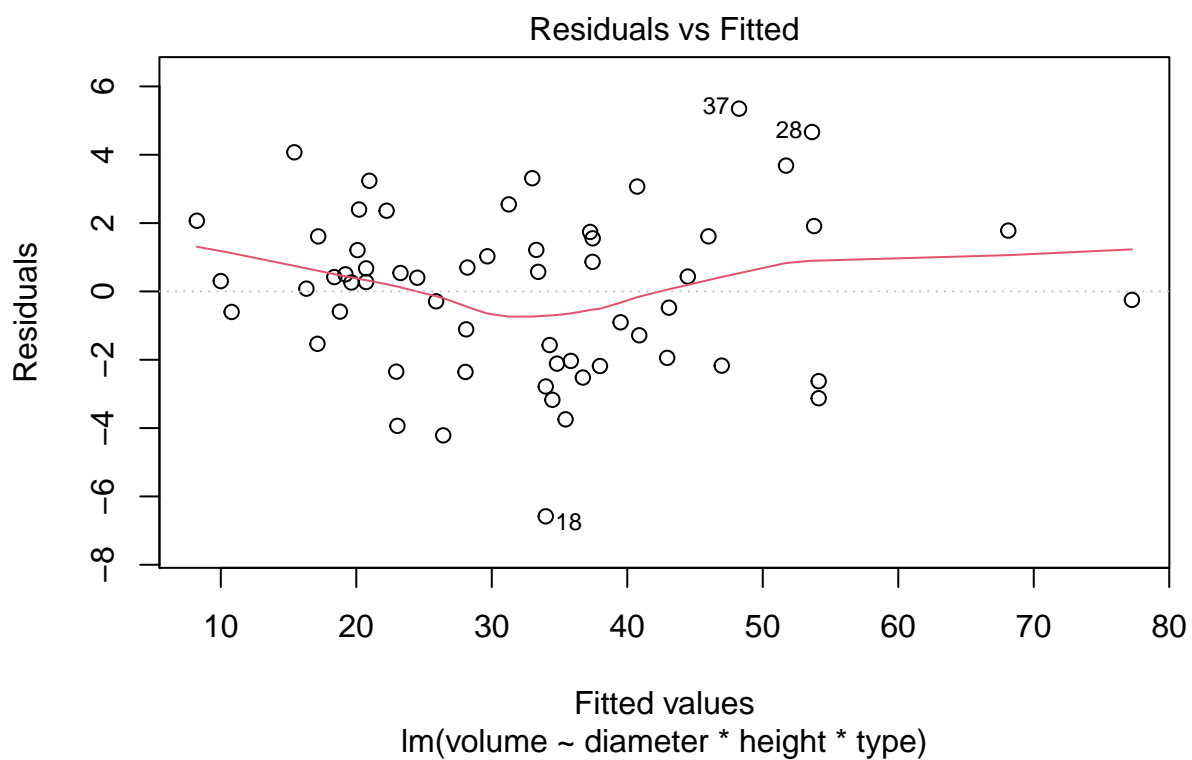
```
summary(dataaov)
```

```
##
## Call:
## lm(formula = volume ~ type, data = data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -19.97  -9.96  -2.77   5.94  46.83
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    30.17       2.54   11.88   <2e-16 ***
## typeoak         5.08       3.69    1.38     0.17
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
## Residual standard error: 14.1 on 57 degrees of freedom
## Multiple R-squared:  0.0322, Adjusted R-squared:  0.0153
## F-statistic:  1.9 on 1 and 57 DF,  p-value: 0.174
```

The estimated volumes are for Beech 30.17 and for oak 35.26.

```
fiber1=lm(volume~diameter+height+type,data=data)
fiber1=lm(volume~diameter*height*type,data=data)
plot(fiber1, 1)
```



Residuals vs Fitted

```
plot(fiber1, 2)
```

## Normal Q–Q



Theoretical Quantiles
lm(volume ~ diameter * height * type)
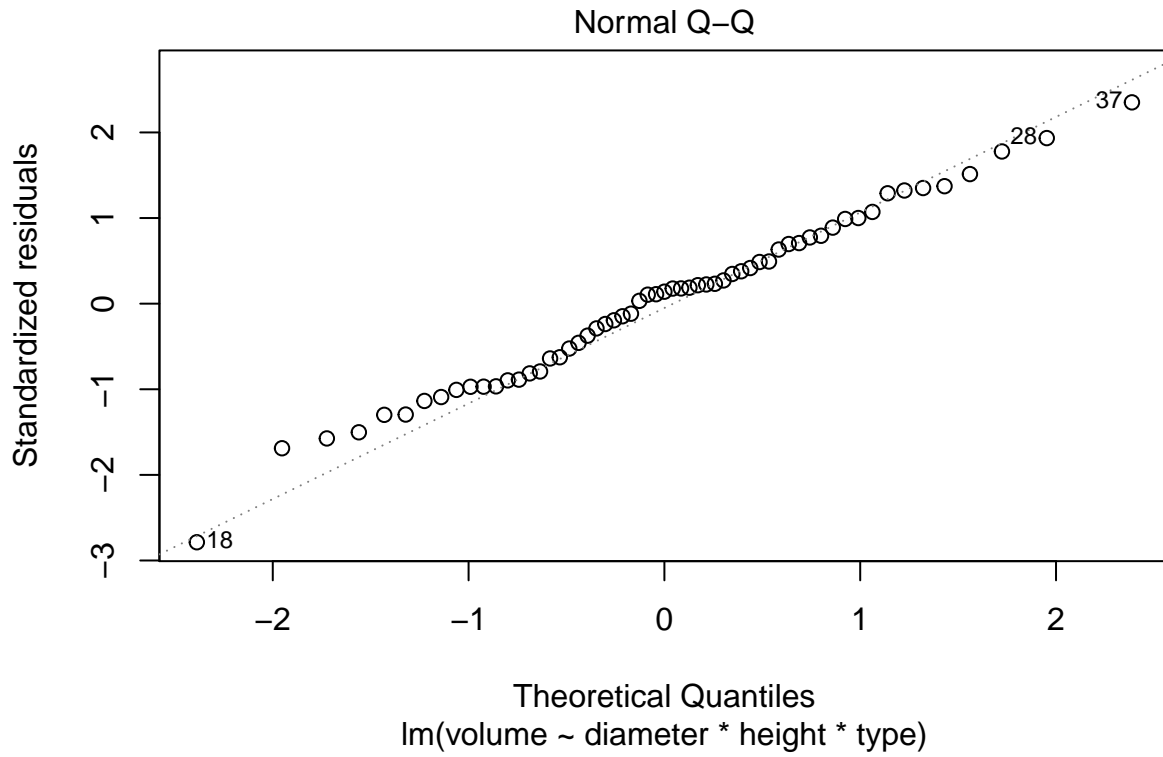
```
drop1(fiber1,test="F")
```

```
## Single term deletions
##
## Model:
## volume ~ diameter * height * type
##                     Df Sum of Sq RSS AIC F value Pr(>F)
## <none>                           332 118
## diameter:height:type  1    42.2 374 123    6.47  0.014 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(fiber1)
```

```
##
## Call:
## lm(formula = volume ~ diameter * height * type, data = data)
##
## Residuals:
##    Min    1Q Median    3Q    Max
## -6.582 -1.989  0.303  1.611  5.349
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)                   69.3963    22.4603    3.09    0.0032 **
## diameter                      -5.8558     1.8105   -3.23    0.0021 **
## height                        -1.2971     0.2920   -4.44  4.8e-05 ***
## typeoak                     -144.8393    56.9011   -2.55    0.0140 *
## diameter:height               0.1347     0.0230    5.86  3.4e-07 ***
## diameter:typeoak             10.6396     3.8515    2.76    0.0080 **
## height:typeoak                1.8116     0.7765    2.33    0.0236 *
## diameter:height:typeoak      -0.1331     0.0523   -2.54    0.0140 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.55 on 51 degrees of freedom
## Multiple R-squared:  0.972,  Adjusted R-squared:  0.968
## F-statistic:  251 on 7 and 51 DF,  p-value: <2e-16
```

```
# Taking the averages
a_height = sum(data$height)/59
a_diameter = sum(data$diameter)/59

a_height
```

```
## [1] 75.8
```

```
a_diameter
```

```
## [1] 13.9
```

```
Oak = -63.7814 + 1 * -1.3046 + a_diameter * 4.6981 + a_height * 0.4172

Beech = -63.7814 + a_diameter * 4.6981 + a_height * 0.4172

Oak
```
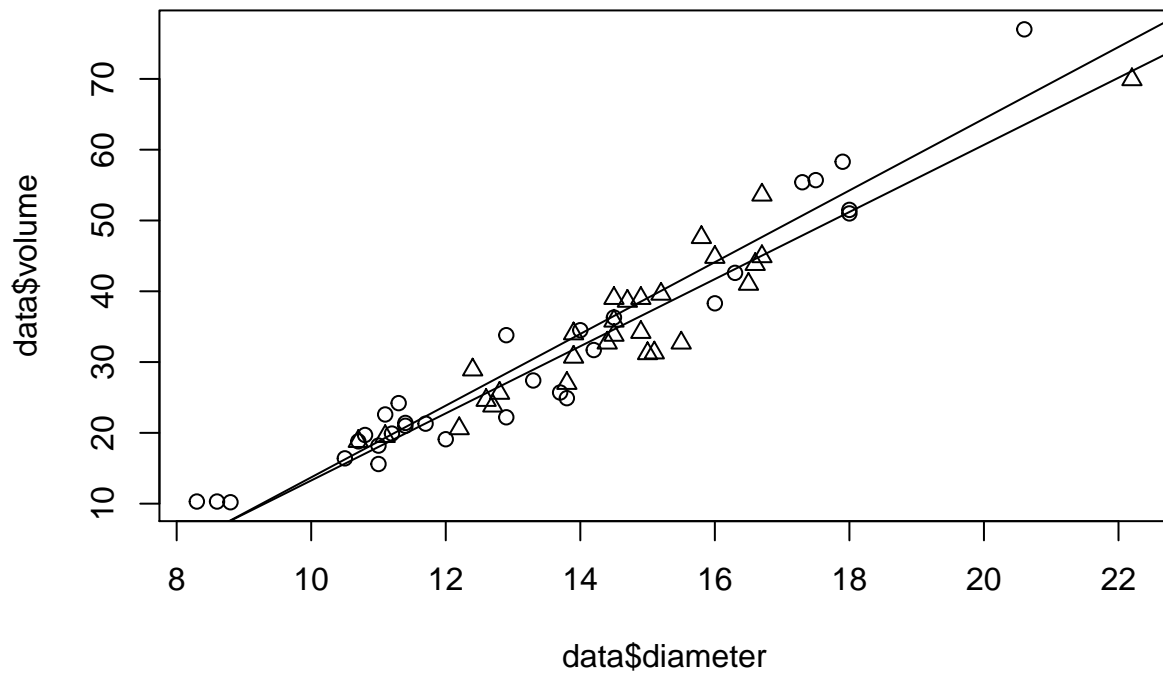
```
## [1] 31.9
```
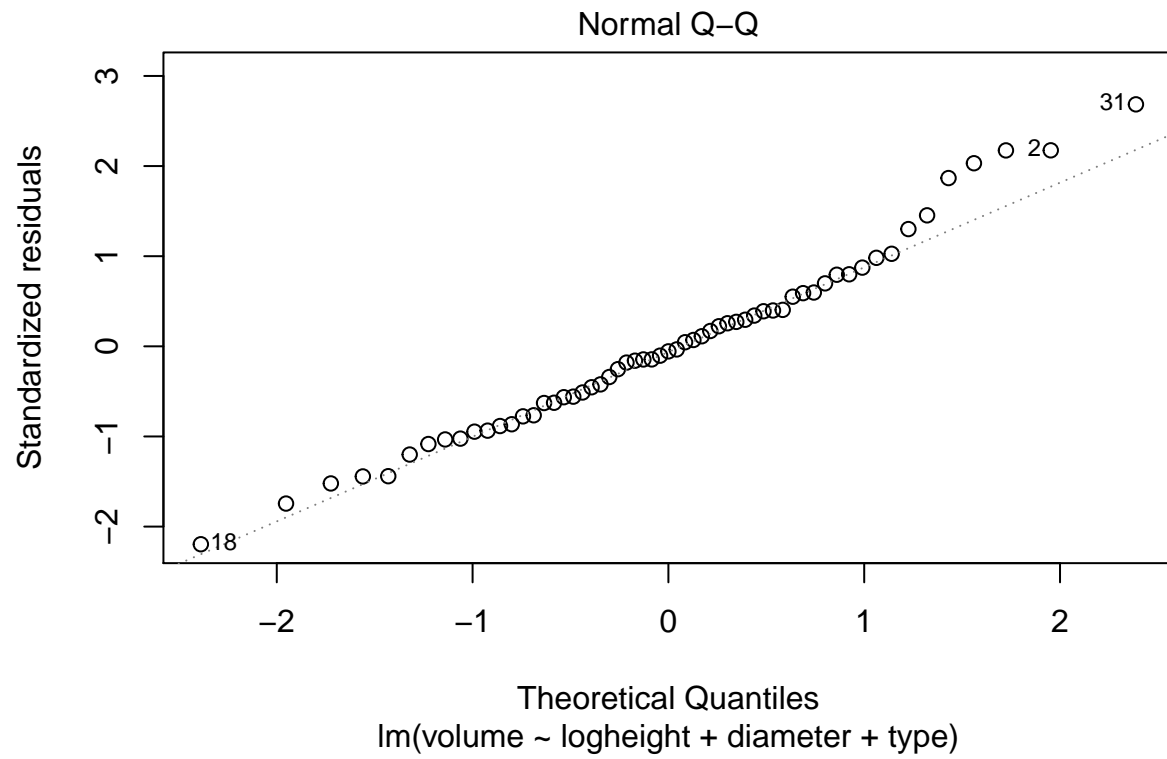
```
Beech
```

```
## [1] 33.2
```

Alsso in this model the type is insignificant. The estimate of the tree can be found by doing the following: Y = -63.7814 + typeoak * -1.3046 + diameter * 4.6981 + height * 0.4172. This means tha ta Oak has a volume of 31.9 and Beech a volume of 33.2

```
trees = c("beech", "oak")
plot(data$volume~data$diameter,pch=unclass(data$type));for (i in trees){abline(lm(volume~diameter,data=
```

We see that the diameter has a strong effect on the volume of both trees, a little bit more on Beech than Oak

```
data$logheight = log(data$height)
fiber1=lm(volume~logheight+diameter+type,data=data)
plot(fiber1, 2)
```

Normal Q–Q

Standardized residuals

Theoretical Quantiles
lm(volume ~ logheight + diameter + type)

```
plot(fiber1, 1)
```

Residuals vs Fitted

Fitted values
lm(volume ~ logheight + diameter + type)