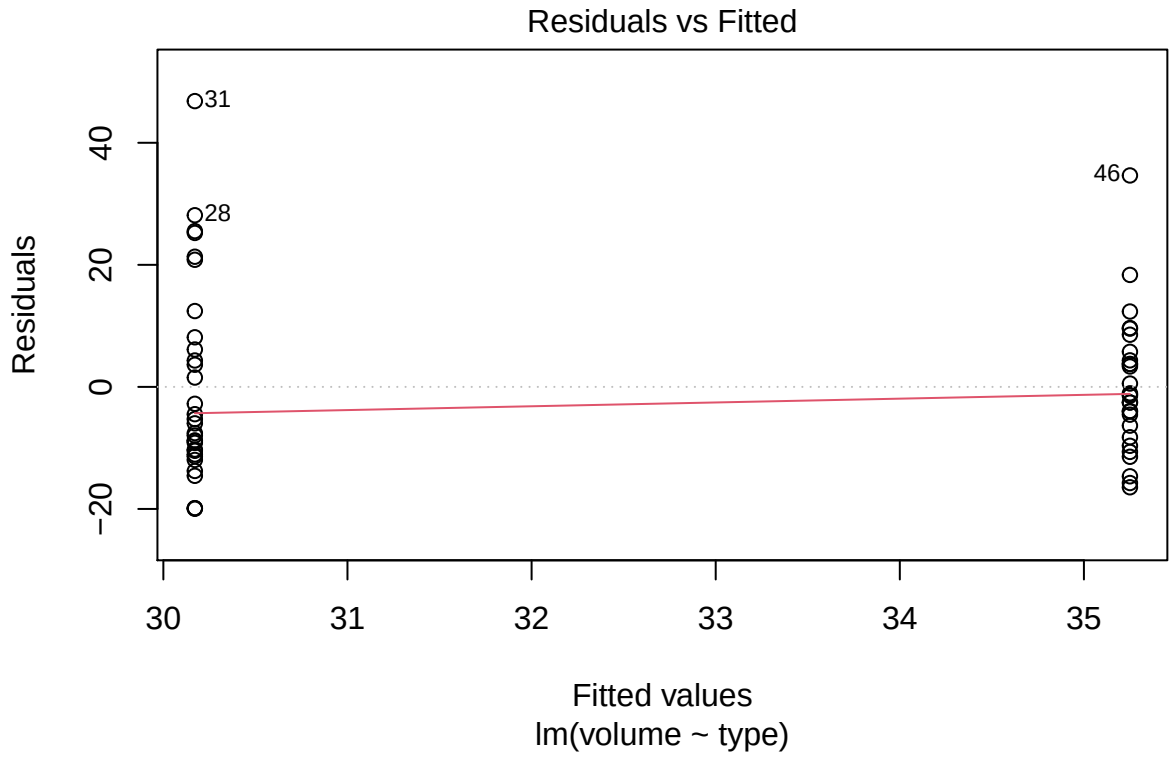
EDDA -Prctice

Kamiel Gülpen

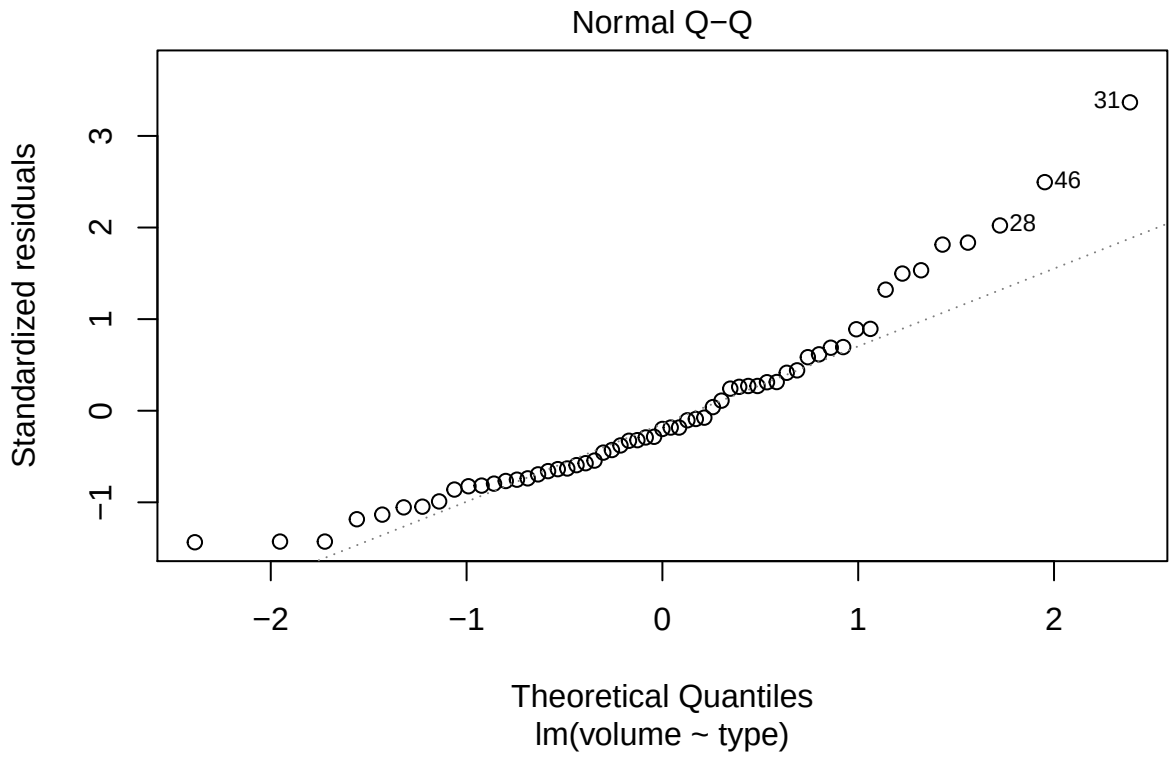
Exercise 1

```
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)
```




```
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
```

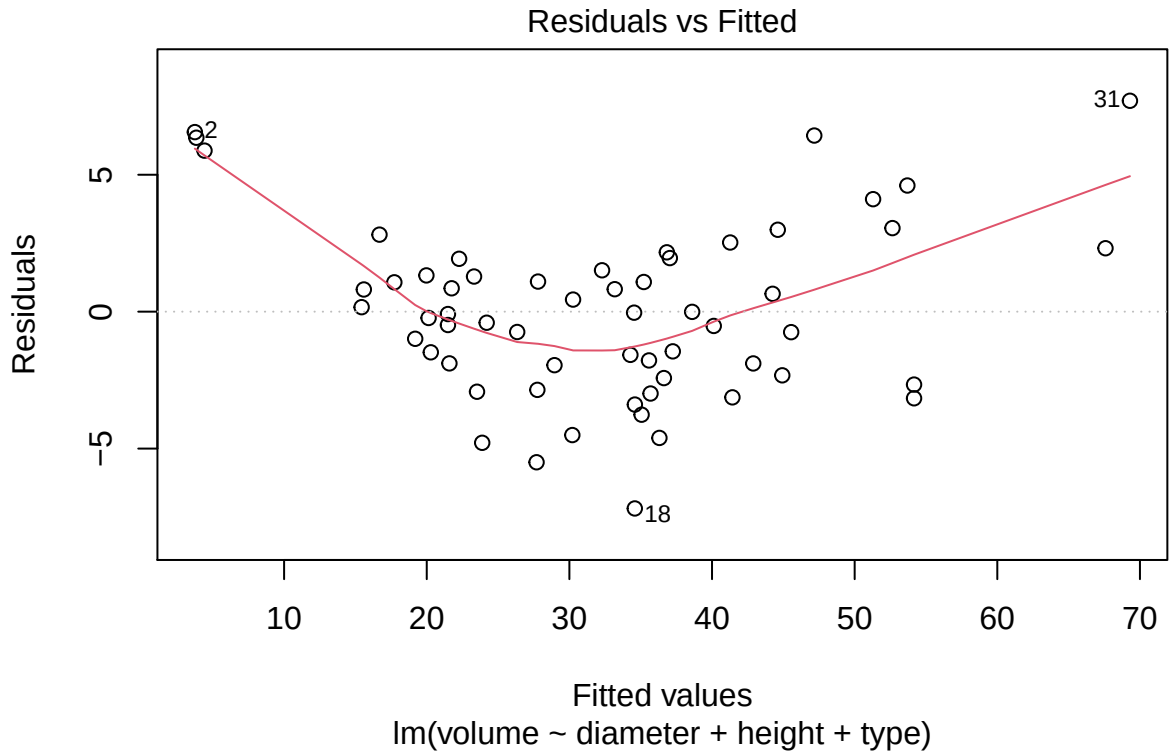
```
summary(dataaov)
```

```
##
## 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
```

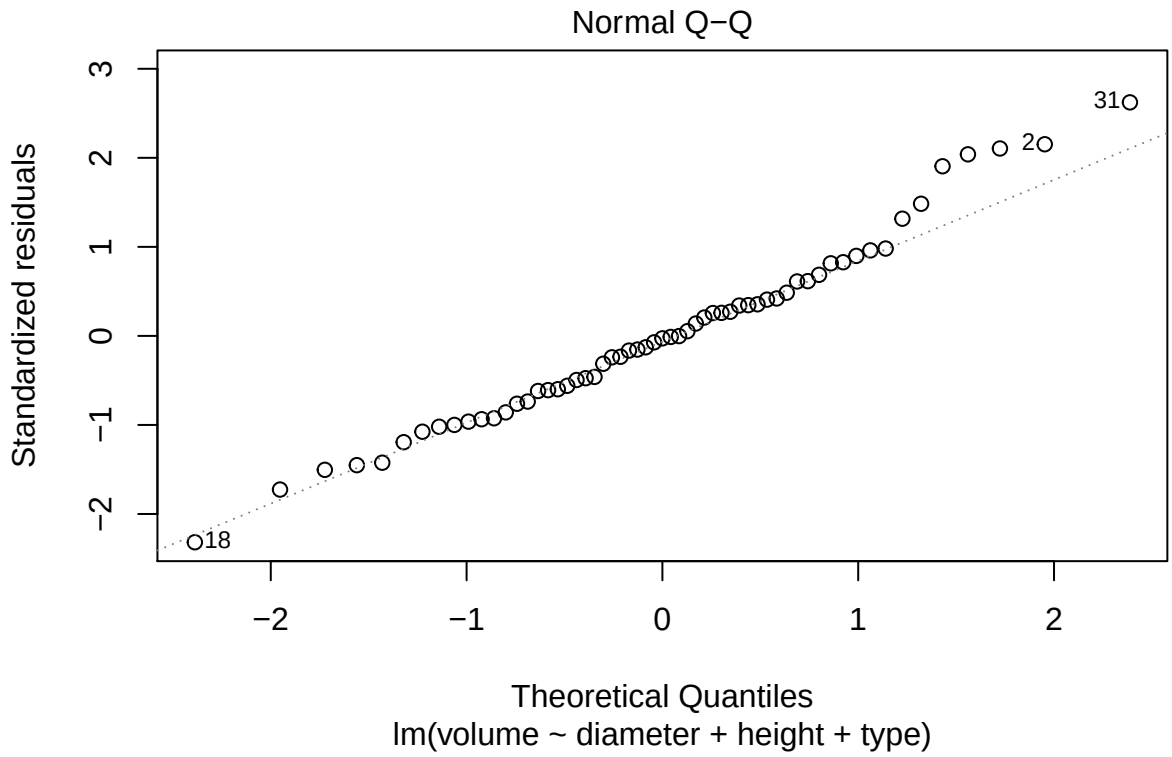
```
##
## Residual standard error: 14.1 on 57 degrees of freedom
## Multiple R-squared:  0.0322, Adjusted R-squared:  0.0153
## F-statistic:  1.9 on 1 and 57 DF,  p-value: 0.174
```

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

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




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




```
drop1(fiber1,test="F")
```

```
## Single term deletions
##
## Model:
## volume ~ diameter + height + type
##          Df Sum of Sq  RSS AIC F value  Pr(>F)
## <none>                 578 143
## diameter  1      8577 9155 304  815.61 < 2e-16 ***
## height    1       324  903 167   30.82 8.4e-07 ***
## type      1        23  602 143    2.21    0.14
## ---
## 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
## -7.186 -2.140 -0.087  1.721  7.701
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -63.7814     5.5129  -11.57  2.3e-16 ***
## diameter      4.6981     0.1645   28.56  < 2e-16 ***
## height        0.4172     0.0752    5.55  8.4e-07 ***
## typeoak      -1.3046     0.8779   -1.49     0.14
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.24 on 55 degrees of freedom
## Multiple R-squared:  0.951,  Adjusted R-squared:  0.948
## F-statistic:  355 on 3 and 55 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
```

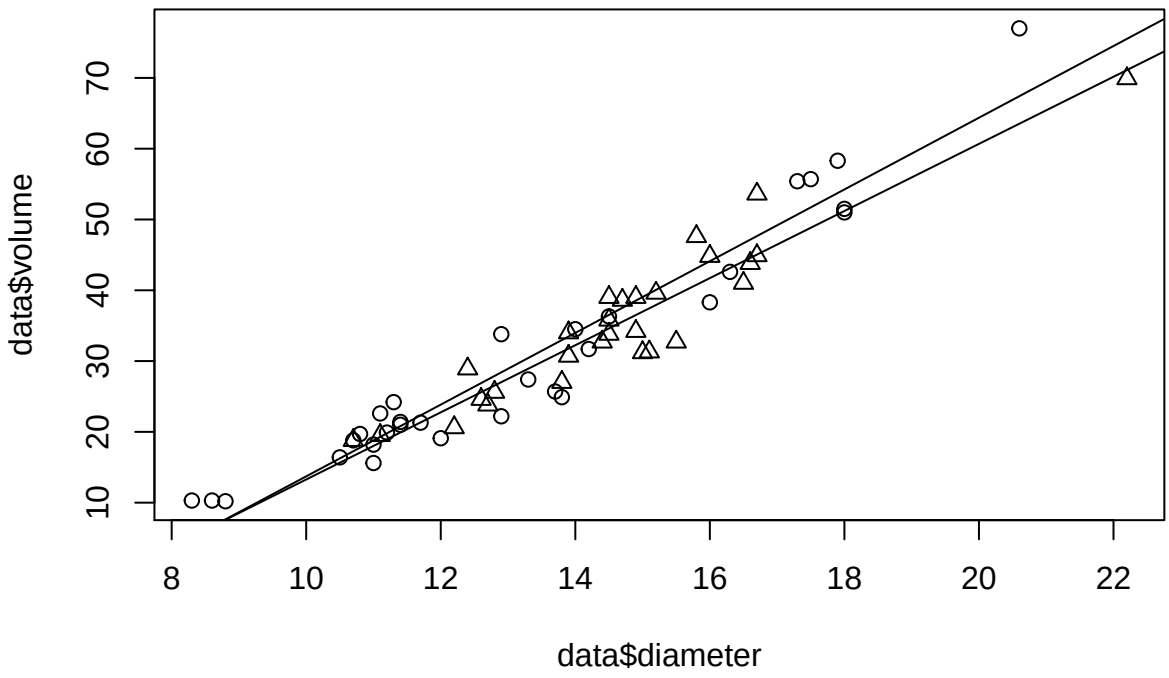
```
## [1] 31.9
```

```
Beech
```

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

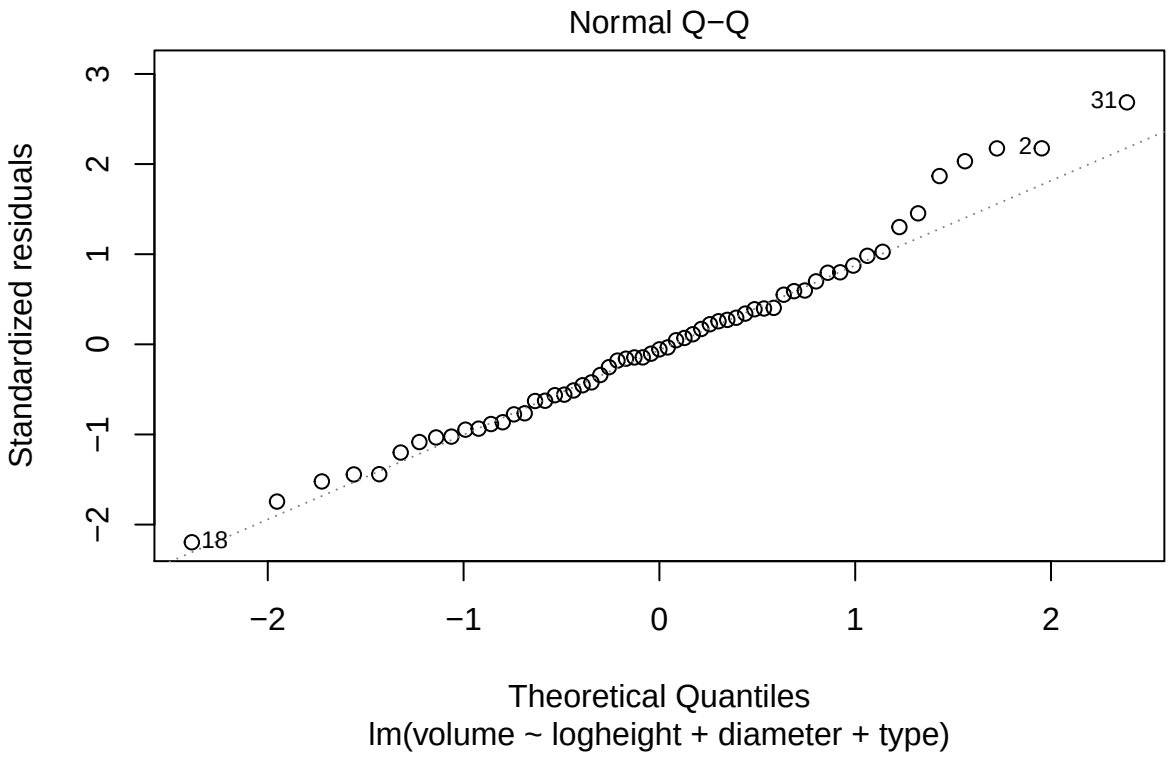
Alsso in this model the type is insignificant. The estimate of the tree can be found by doing the following: Y = -63.7814 + typeoak * -1.3046 + diameter * 4.6981 + height * 0.4172. This means tha ta Oak has a volume of 31.9 and Beech a volume of 33.2

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

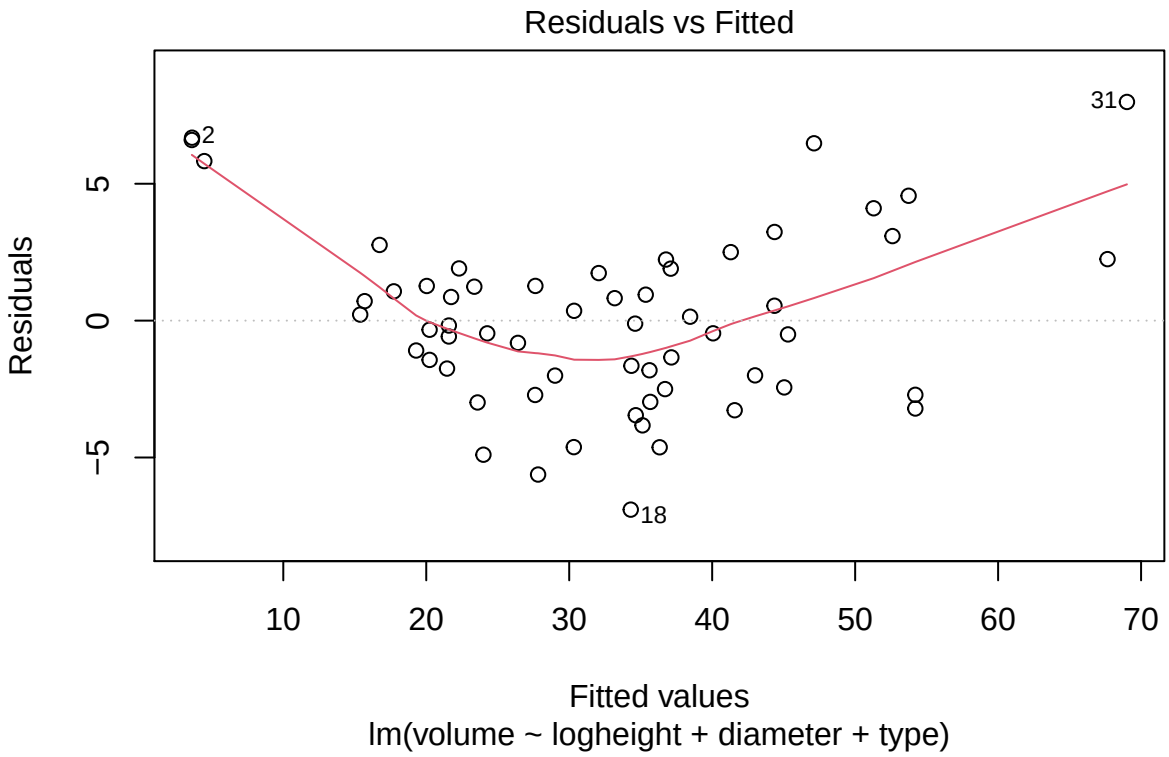


We see that the diameter has a strong effect on the volume of both trees, a little bit more on Beech than Oak

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




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


Residuals vs Fitted
Residuals
Fitted values
lm(volume ~ logheight + diameter + type)
2
31
18
5
0
−5
10
20
30
40
50
60
70