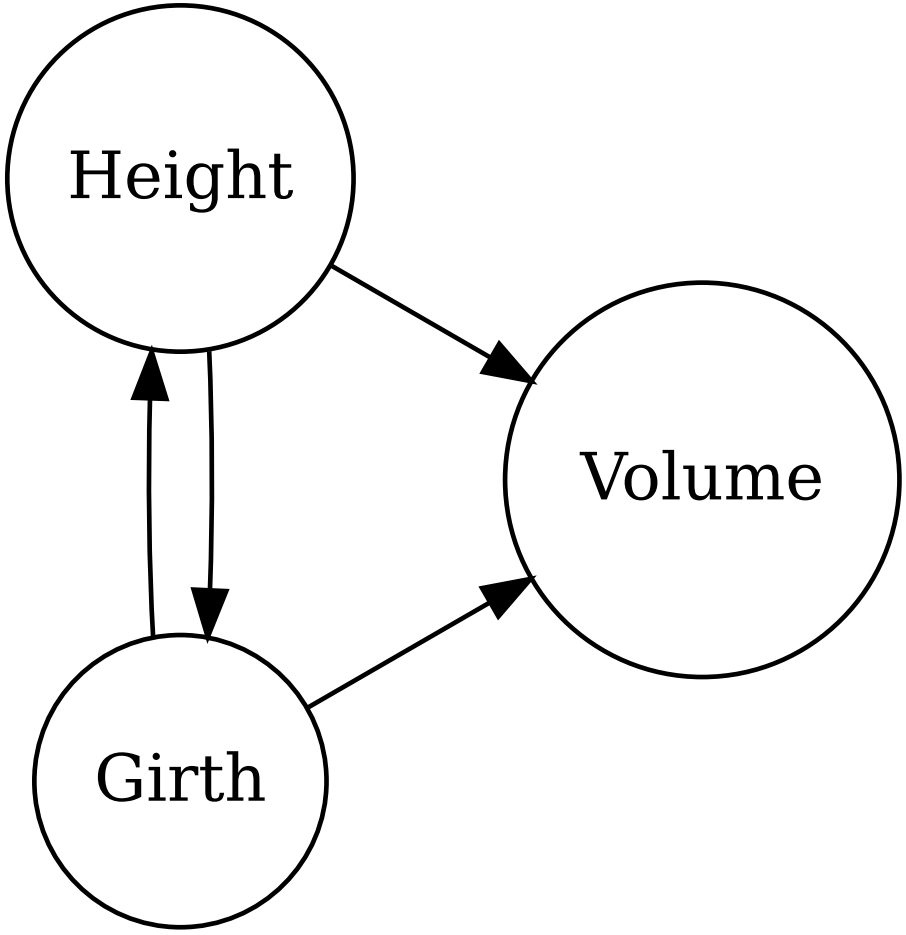
Tutorial / Lösungshinweise (RStudio)

Dennis Klinkhammer (2022)

Aufgabe: TREES

Analysemodell

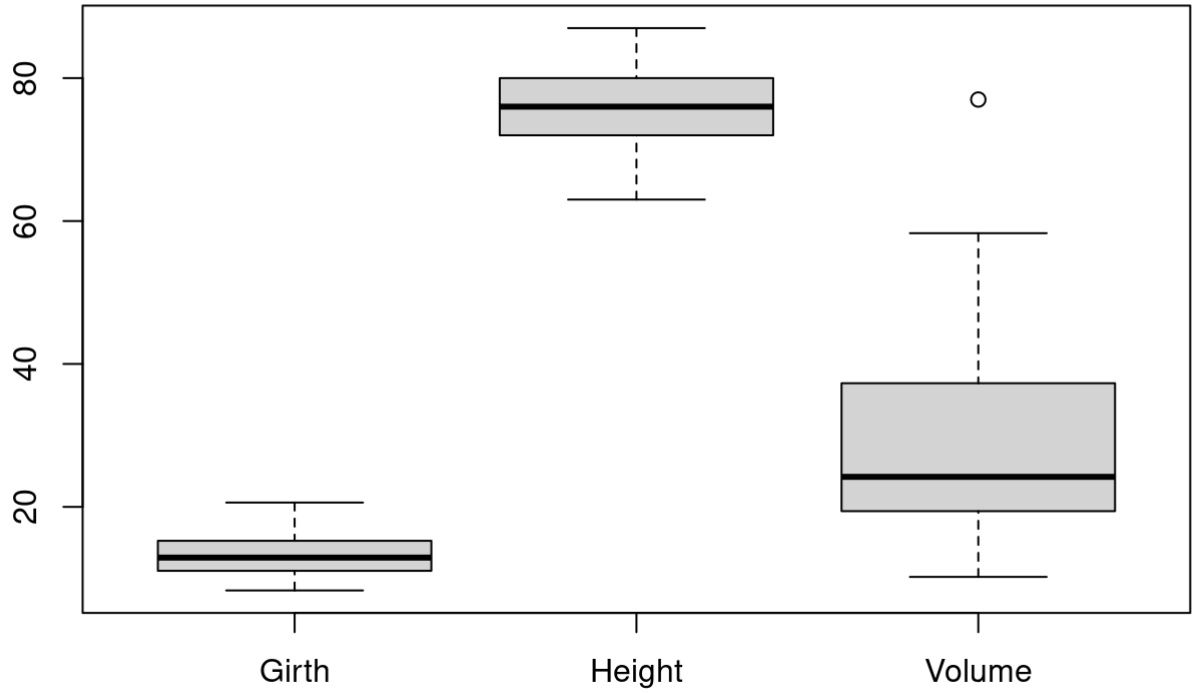


Univariate Analyse

```
summary(trees)
```

```
##      Girth           Height       Volume     
##  Min.   : 8.30   Min.   :63   Min.   :10.20  
##  1st Qu.:11.05   1st Qu.:72   1st Qu.:19.40  
##  Median :12.90   Median :76   Median :24.20  
##  Mean   :13.25   Mean   :76   Mean   :30.17  
##  3rd Qu.:15.25   3rd Qu.:80   3rd Qu.:37.30  
##  Max.   :20.60   Max.   :87   Max.   :77.00
```

```
boxplot(trees)
```

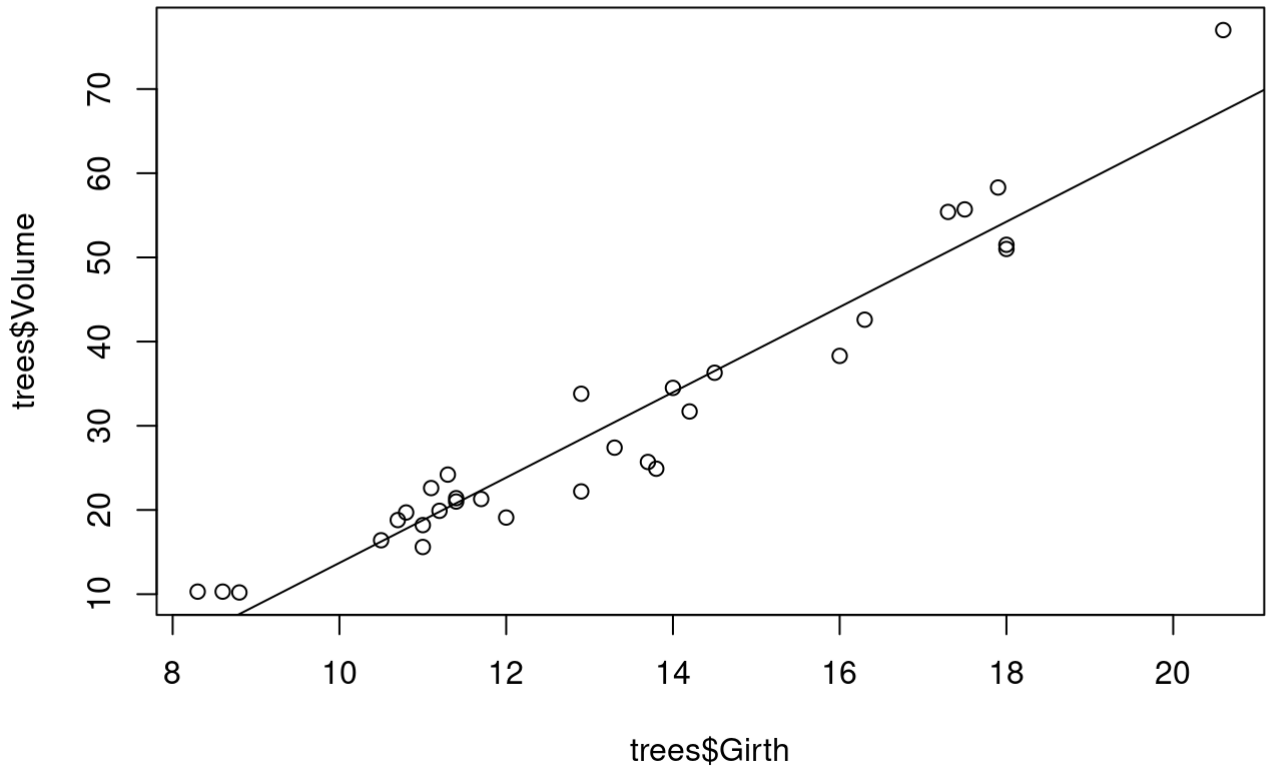


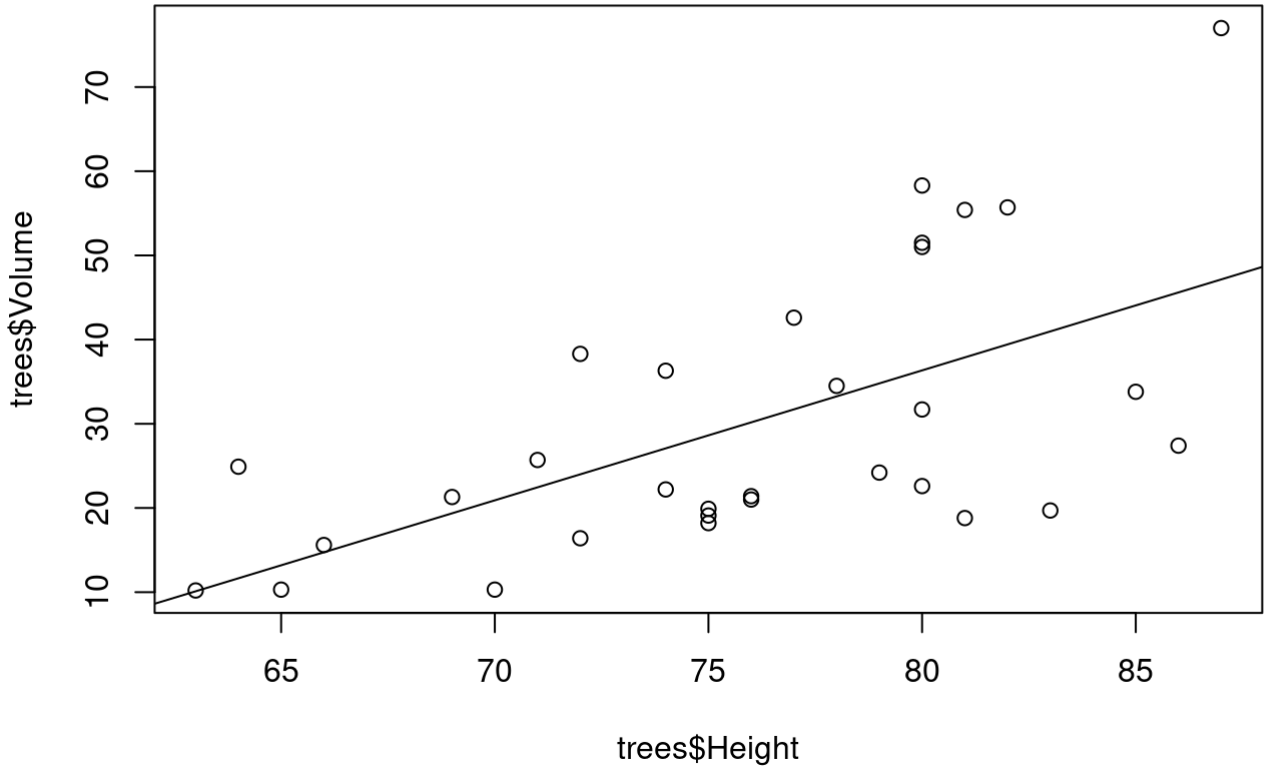
Bivariate Analyse

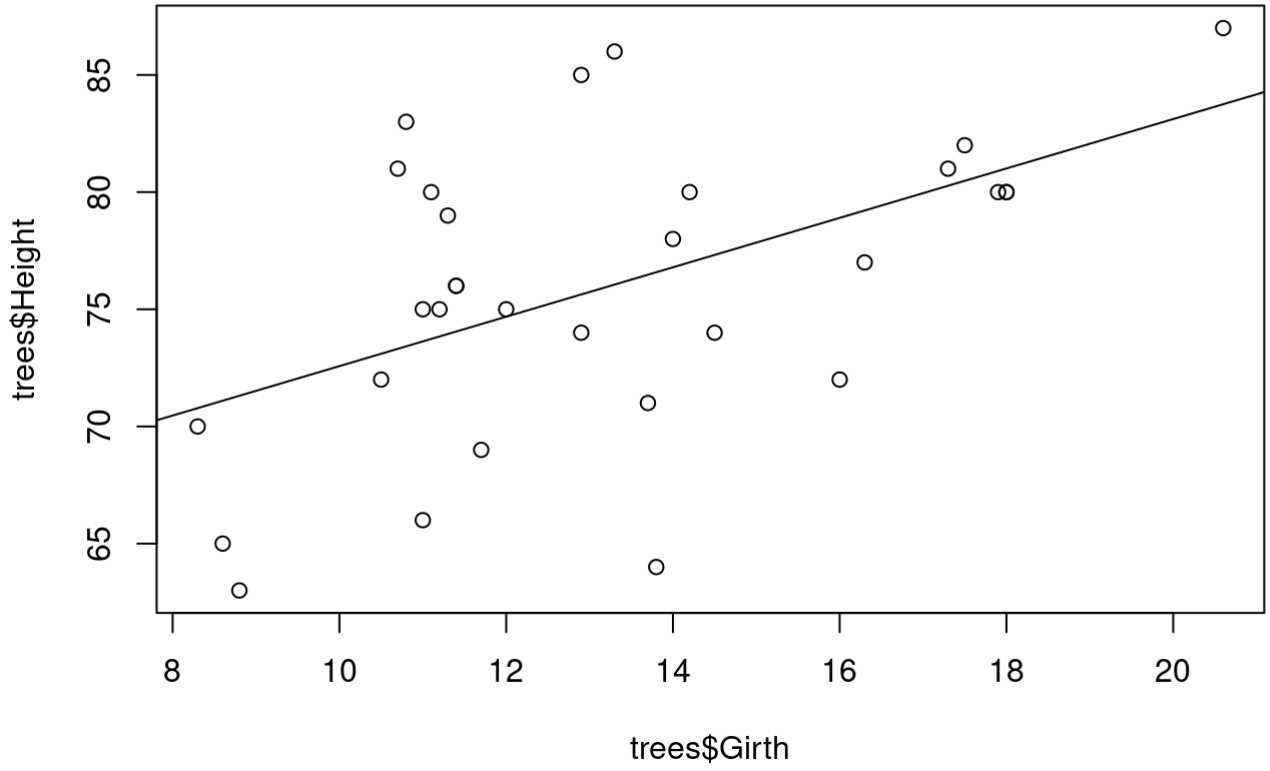
```
cor(trees)
```

```
##            Girth    Height    Volume
## Girth  1.0000000 0.5192801 0.9671194
## Height 0.5192801 1.0000000 0.5982497
## Volume 0.9671194 0.5982497 1.0000000
```

```
plot(trees$Volume~trees$Girth)
abline(lm(trees$Volume~trees$Girth))
```




trees$Volume
10
20
30
40
50
60
70
65
70
75
80
85
trees$Height



Multivariate Analyse

```
summary(lm(Volume~Girth+Height, data=trees))
```

```
##
## Call:
## lm(formula = Volume ~ Girth + Height, data = trees)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.4065 -2.6493 -0.2876  2.2003  8.4847
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -57.9877     8.6382  -6.713 2.75e-07 ***
## Girth         4.7082     0.2643  17.816  < 2e-16 ***
## Height        0.3393     0.1302   2.607   0.0145 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.882 on 28 degrees of freedom
## Multiple R-squared:  0.948,  Adjusted R-squared:  0.9442
## F-statistic:   255 on 2 and 28 DF,  p-value: < 2.2e-16
```

Überprüfung der Vorhersagegenauigkeit

```
trees[13,]
```

```
##    Girth Height Volume
## 13  11.4     76   21.4
```

```
# VOLUME ~ -57.98 + GIRTH * 4.71 + HEIGHT * 0.34
# 21.4 ~ -57.98 + 11.4 * 4.71 + 76 * 0.34
# 21.4 ~ -57.98 + 53.69 + 25.84
# 21.4 ~ 21.55
```