

gl04_hoermann

Paul Hörmann

11/23/2019

```
getwd()
```

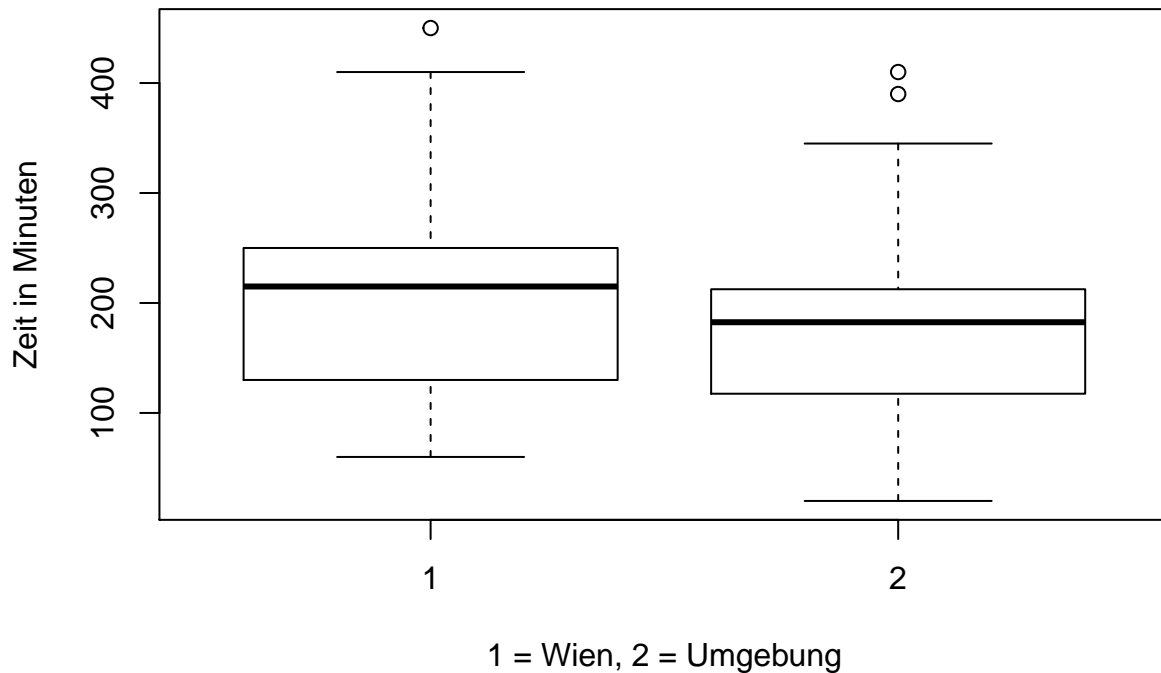
```
## [1] "/home/paul/git/fhhgb_ws19/0_ASC/r/ue05"
```

Aufgabe 1

```
monitorData = read.csv("./monitor.csv", sep=";")  
head(monitorData)
```

```
##   zeit stadt  
## 1  230     1  
## 2  170     1  
## 3  180     1  
## 4  220     1  
## 5   85     1  
## 6  130     1
```

```
boxplot(monitorData$zeit~monitorData$stadt,ylab = "Zeit in Minuten",xlab = "1 = Wien, 2 = Umgebung")
```



H0: Zeit des fernsehens ist gleich H1: Jugendliche in der Stadt schauen länger

```
# teste auf gruppe 1 größer gruppe 2
```

```
wilcox.test(monitorData$zeit~monitorData$stadt, alternative = "greater")
```

```
##
```

```
## Wilcoxon rank sum test with continuity correction
```

```
##
## data: monitorData$zeit by monitorData$stadt
## W = 1141, p-value = 0.1269
## alternative hypothesis: true location shift is greater than 0
```

Weil p kleiner alpha akzeptieren wir H0.

Aufgabe 2

Daten:

```
politData = matrix(c(120, 23, 11, 182),nrow=2)
rownames(politData) = c("Partei A", "Partei B")
colnames(politData) = c("Partei A", "Partei B")
politData
```

```
##           Partei A Partei B
## Partei A      120      11
## Partei B       23     182
```

Test Ergebnis:

```
politTestResult = mcnemar.test(politData)
politTestResult
```

```
##
## McNemar's Chi-squared test with continuity correction
##
## data: politData
## McNemar's chi-squared = 3.5588, df = 1, p-value = 0.05923
```

Weil p größer als alpha akzeptieren wir H0.

Aufgabe 3

Daten:

```
library(readxl)
babyData = read_excel("./Babydaten.xlsx")
head(babyData)
```

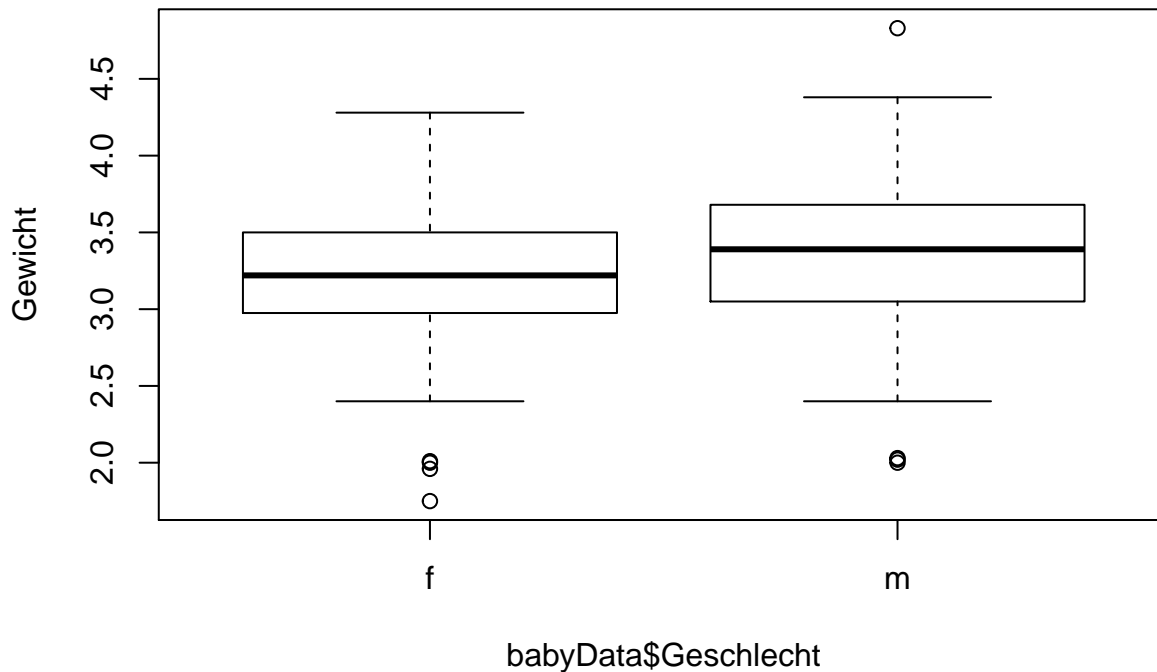
```
## # A tibble: 6 x 6
##       ID Geburtsdatum Geschlecht Gewicht Größe Blutgruppe
##   <dbl> <dtm>         <chr>      <dbl> <dbl> <chr>
## 1  1059 2007-03-20 12:52:00 f         3.68   52 0
## 2  1060 2007-03-24 18:33:00 m         3.52   52 0
## 3  1061 2007-03-25 11:08:00 f         4.03   52 A
## 4  1062 2007-03-21 17:02:00 m         3.41   48 A
## 5  1063 2007-04-01 16:34:00 m         2.92   48 AB
## 6  1064 2007-03-21 09:31:00 m         2.66   47 A
```

```
nval = nrow(babyData)
```

b)

Test: H0: Gewicht ist unabhängig vom Geschlecht. H1: Gewicht von Buben ist größer.

```
boxplot(babyData$Gewicht~babyData$Geschlecht,ylab = "Gewicht")
```

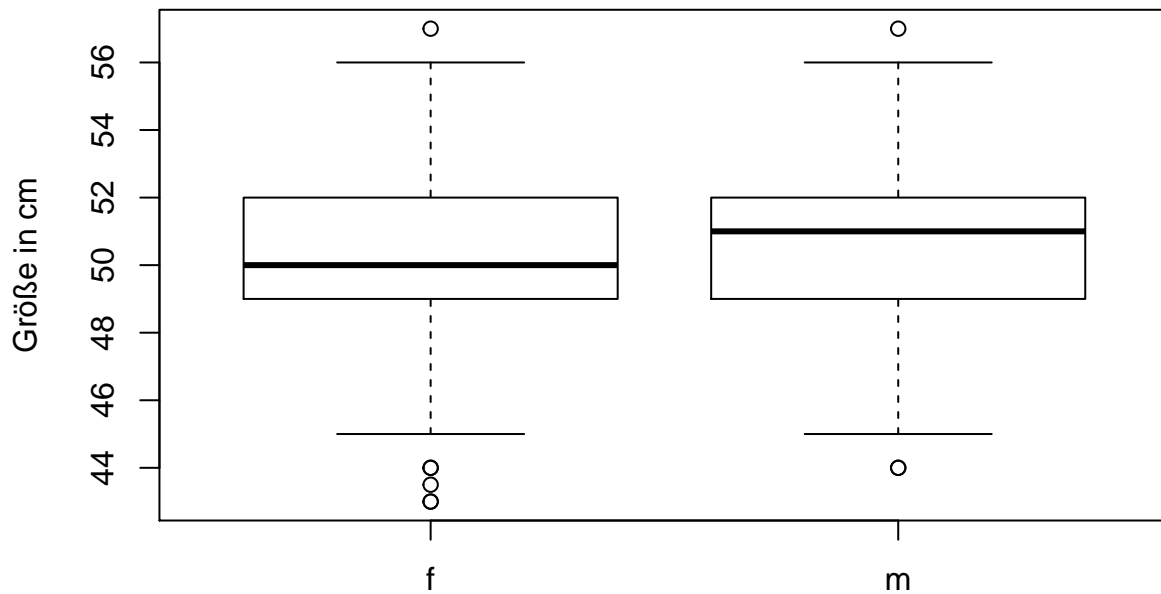


```
wilcox.test(babyData$Gewicht~babyData$Geschlecht,alternative = "less")
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: babyData$Gewicht by babyData$Geschlecht
## W = 14396, p-value = 0.002072
## alternative hypothesis: true location shift is less than 0
```

Da p größer als alpha können wir H0 zugunsten von H1 verwerfen. ### a) Test: H0: Größe ist unabhängig vom Geschlecht. H1: Größe von Buben ist größer.

```
boxplot(babyData$Größe~babyData$Geschlecht,ylab = "Größe in cm")
```



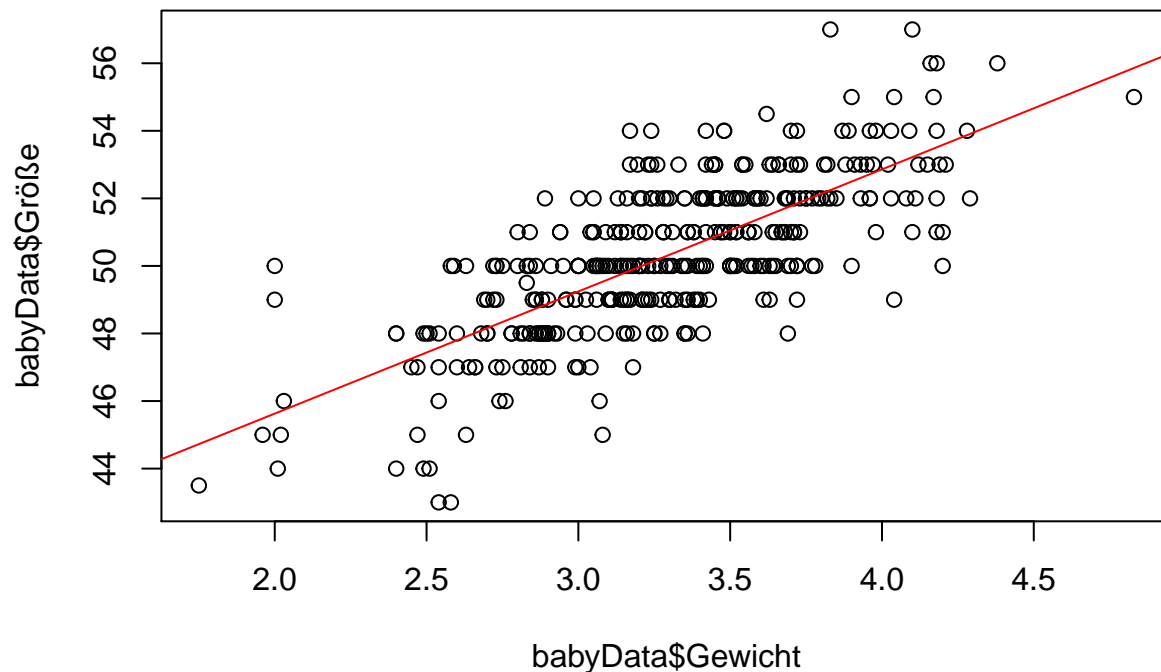
babyData\$Geschlecht

```
wilcox.test(babyData$Größe~babyData$Geschlecht,alternative="less")
```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: babyData$Größe by babyData$Geschlecht
## W = 14527, p-value = 0.002803
## alternative hypothesis: true location shift is less than 0
```

Da p größer als alpha können wir H0 zugunsten von H1 verwerfen. ### c)

```
plot(babyData$Gewicht, babyData$Größe) + abline(lm(babyData$Größe~babyData$Gewicht), col="red")
```



```
## integer(0)
cor.test(babyData$Gewicht, babyData$Größe, alternative="two.sided",
         method = "kendall")
```

```
##
## Kendall's rank correlation tau
##
## data: babyData$Gewicht and babyData$Größe
## z = 15.076, p-value < 2.2e-16
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0.5557917
```

Da p kleiner als alpha können wir H0 zugunsten von H1 verwerfen.

Aufgabe 4

a)

```
library(foreign)
dmobile = read.dta(file = "./mobil-1.dta")
head(dmobile)
```

```
##      dauer      kosten geschlecht tarif      alter
## 1 332.6756 141.89270      Frau  Wenig 18.66894
## 2 347.0924 150.21748      Mann  Wenig 16.15220
## 3 133.3460  73.30383      Mann  Wenig 24.74806
## 4 162.0536 114.47440      Frau   Viel 23.91885
## 5 162.6686  89.74941      Frau   Viel 25.40752
## 6 474.6637 155.90486      Frau Normal 18.38577
```

```
dmobile[which(dmobile$dauer == max(dmobile$dauer)),]
```

```
##      dauer      kosten geschlecht tarif      alter
## 194 499.8094 104.6147      Frau   Viel 23.87184
```

Durchschnittliche Telefonat dauer:

```
mean(dmobile$dauer)
```

```
## [1] 276.5569
```

Frauen telefonieren im Schnitt:

```
mean(dmobile[which(dmobile$geschlecht == "Frau"),]$dauer)
```

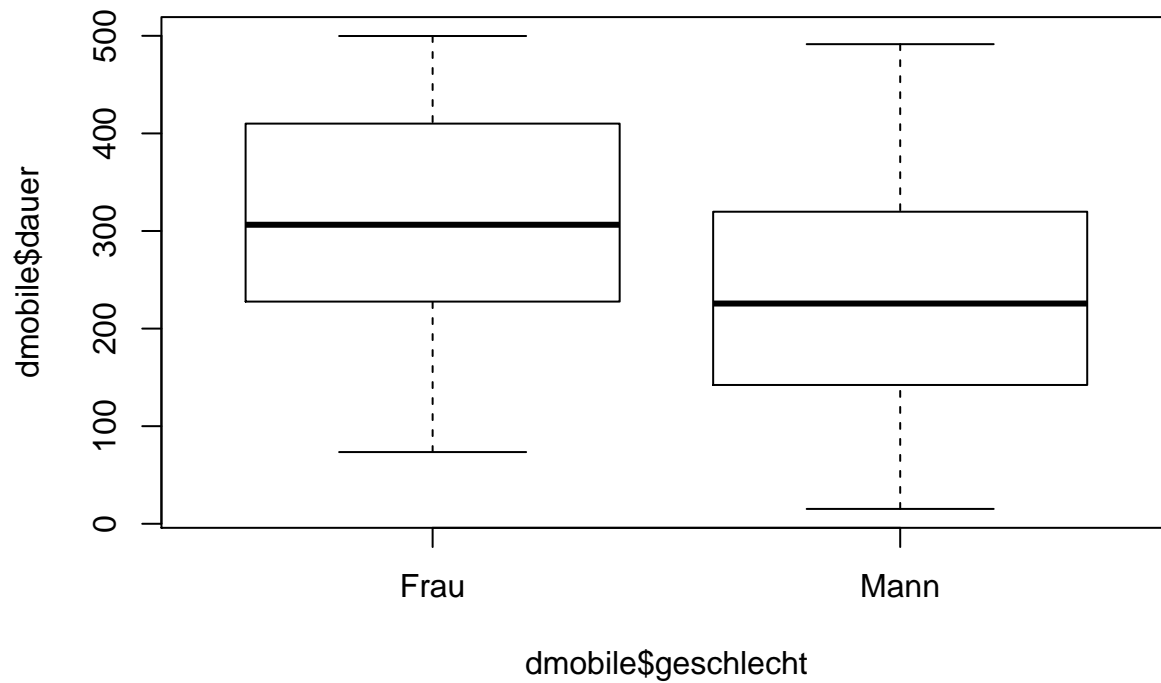
```
## [1] 313.5014
```

Männer telefonieren im Schnitt:

```
mean(dmobile[which(dmobile$geschlecht == "Mann"),]$dauer)
```

```
## [1] 233.41
```

```
boxplot(dmobile$dauer~dmobile$geschlecht)
```

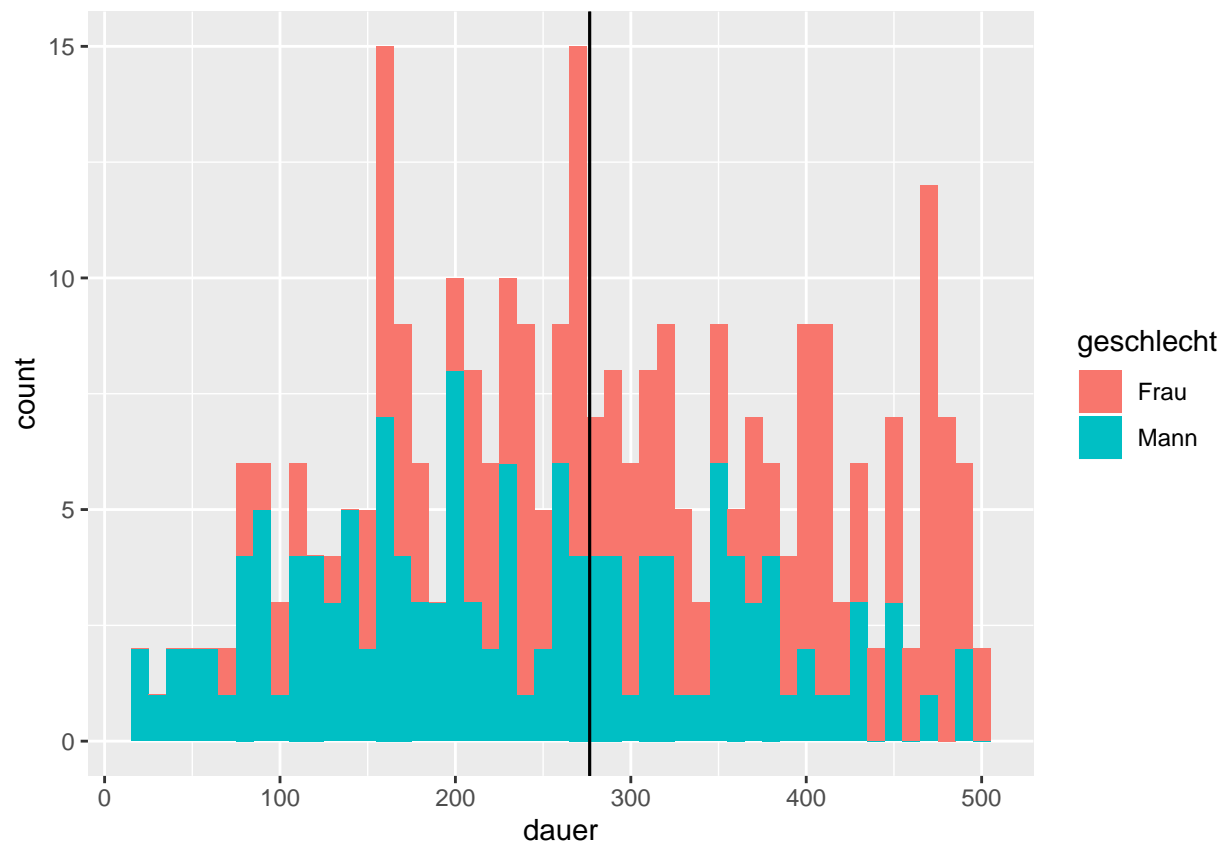


Telefondauer unabhängig vom Geschlecht H1: Frauen telefonieren länger

H0:

```
library(ggplot2)
ggplot(dmobile, aes(x=dauer, fill=geschlecht), col = "red") + geom_histogram(binwidth = 10) + geom_vline
```

Warning: Ignoring unknown aesthetics: fill



```
wilcox.test(dmobile$dauer~dmobile$geschlecht, alternative = "greater")
```

```
##  
## Wilcoxon rank sum test with continuity correction  
##  
## data: dmobile$dauer by dmobile$geschlecht  
## W = 15014, p-value = 1.964e-08  
## alternative hypothesis: true location shift is greater than 0
```

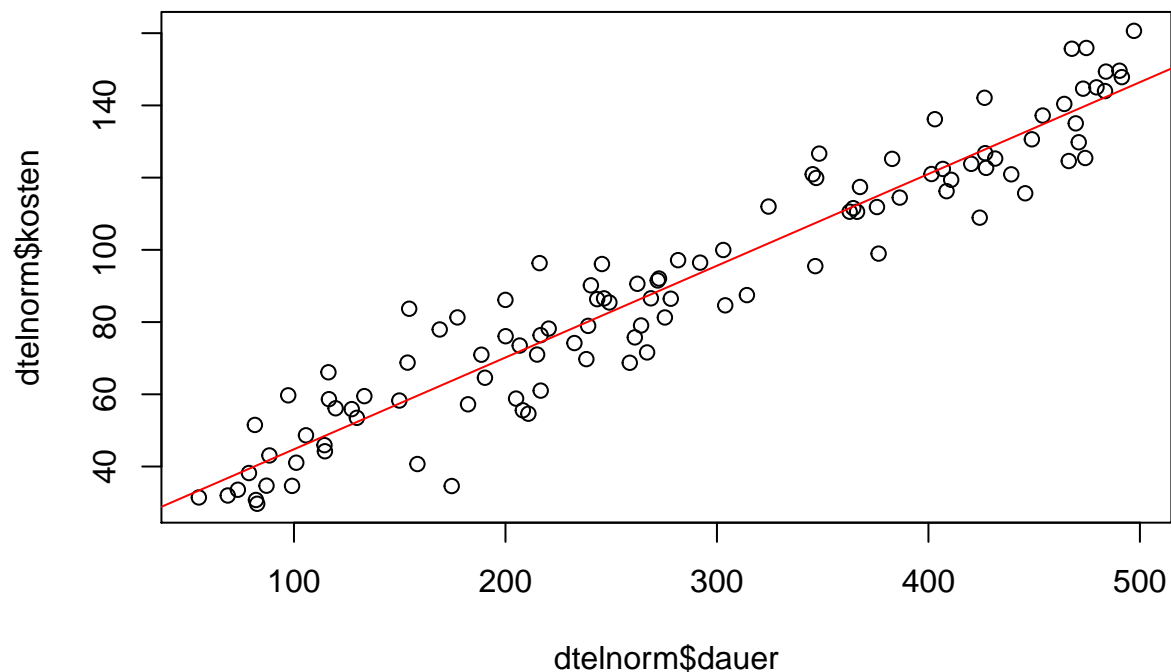
Das heißt Frauen telefonieren im Schnitt länger. ### c)

```
dtelnorm = dmobile[dmobile$tarif == "Normal",]  
dtelviel = dmobile[dmobile$tarif == "Viel",]  
dtelwenig = dmobile[dmobile$tarif == "Wenig",]  
ltelnorm = lm(dtelnorm$kosten~dtelnorm$dauer)  
ltelviel = lm(dtelviel$kosten~dtelviel$dauer)  
ltelwenig = lm(dtelwenig$kosten~dtelwenig$dauer)
```

```
ltelnorm
```

```
##  
## Call:  
## lm(formula = dtelnorm$kosten ~ dtelnorm$dauer)  
##  
## Coefficients:  
## (Intercept) dtelnorm$dauer  
## 19.3661 0.2541
```

```
plot(dtelnorm$dauer, dtelnorm$kosten) + abline(ltelnorm, col = "red")
```



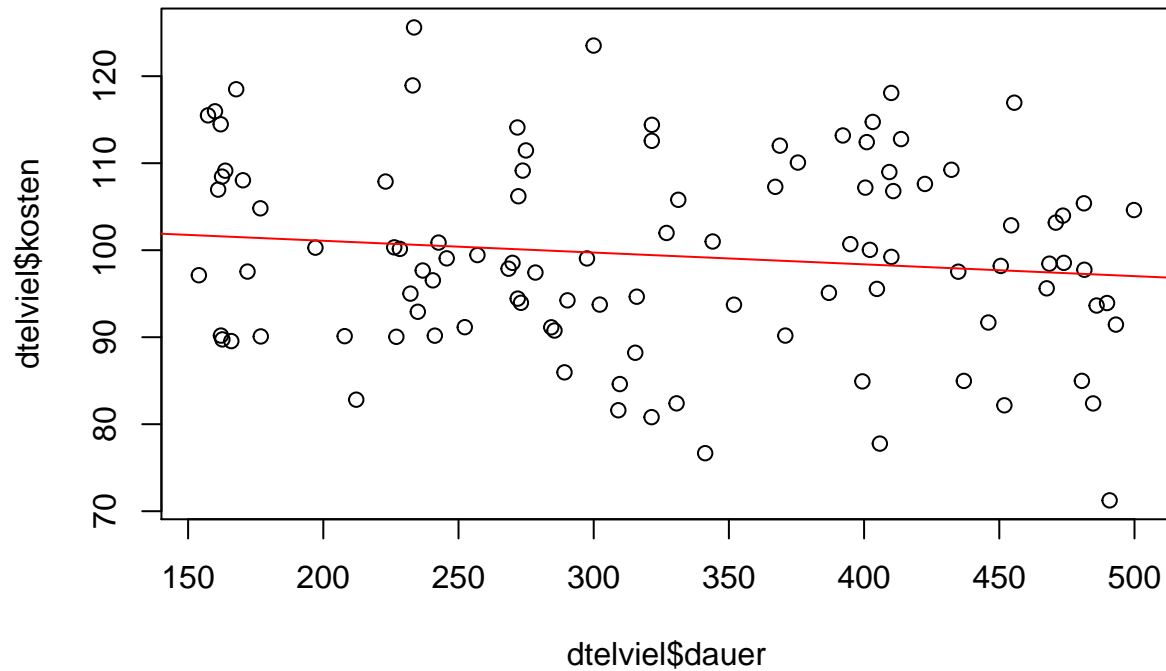
```
## integer(0)
```

```
ltelviel
```

```
##  
## Call:
```

```
## lm(formula = dtelviel$kosten ~ dtelviel$dauer)
##
## Coefficients:
##      (Intercept)  dtelviel$dauer
##      103.78868      -0.01354
```

```
plot(dtelviel$dauer, dtelviel$kosten) + abline(ltelviel, col = "red")
```

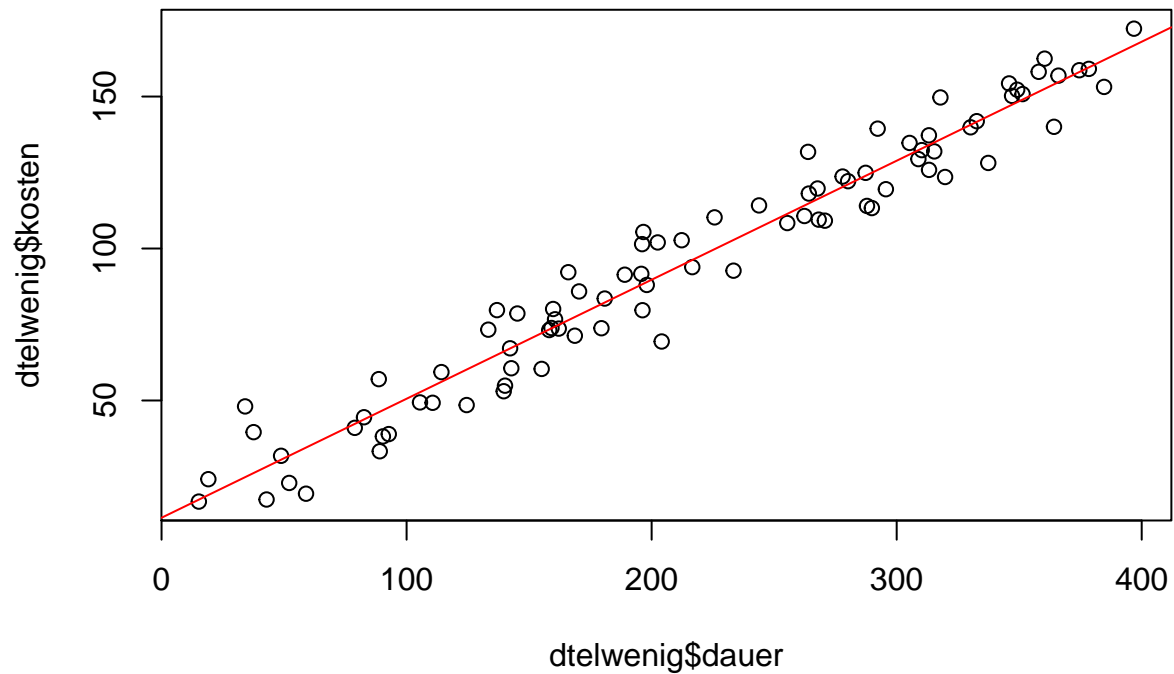


```
## integer(0)
```

```
ltelwenig
```

```
##
## Call:
## lm(formula = dtelwenig$kosten ~ dtelwenig$dauer)
##
## Coefficients:
##      (Intercept)  dtelwenig$dauer
##      11.4255      0.3915
```

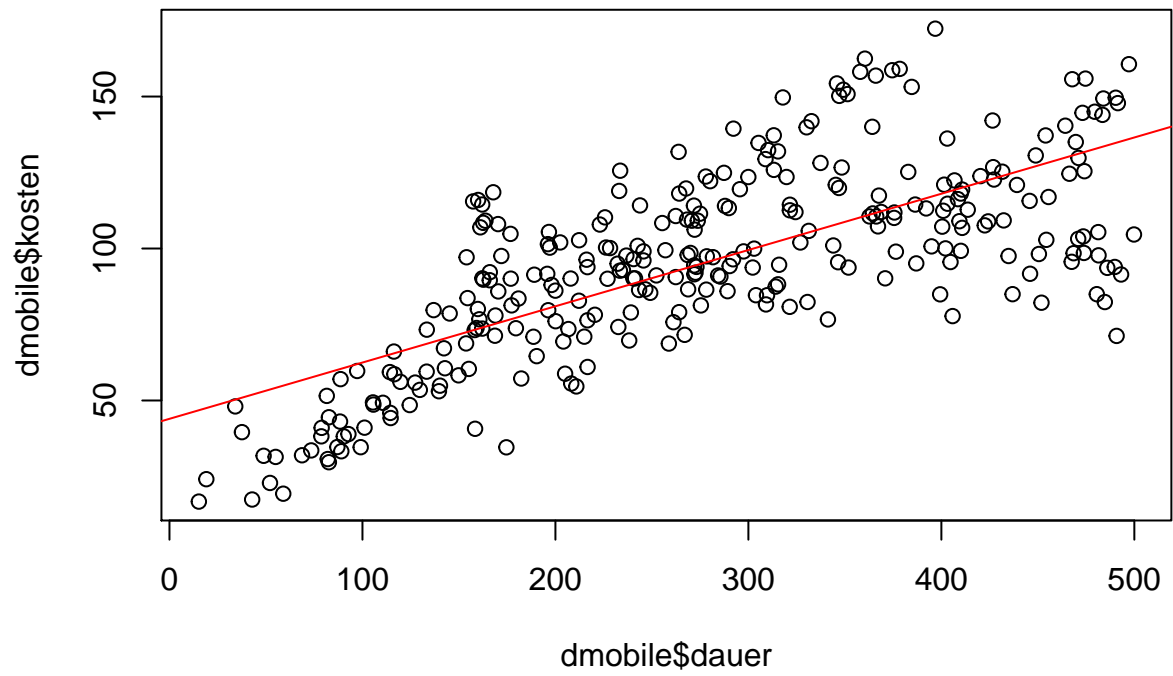
```
plot(dtelwenig$dauer, dtelwenig$kosten) + abline(ltelwenig, col = "red")
```

```
## integer(0)
attach(dmobile)
ltelmultivar = lm(kosten~dauer)
detach(dmobile)
ltelmultivar

##
## Call:
## lm(formula = kosten ~ dauer)
##
## Coefficients:
## (Intercept)      dauer
##    43.9676    0.1851

plot(dmobile$dauer, dmobile$kosten) + abline(ltelmultivar, col = "red")
```



```
## integer(0)
```