

Lösung zu Kapitel 10: Beispiel 3

Die Datei `vwgh.csv` enthält unter anderem die Variablen `senat` und `dauer3`, die hier verwendet werden sollen um die Verfahrensdauer in verschiedenen Senaten zu vergleichen.

- Wir lesen `vwgh.csv` in das Objekt `vwgh` ein und verwenden `attach` um die Variablen direkt ansprechen zu können. Alle `-9999` in `dauer3` stehen für fehlende Werte und müssen daher recodiert werden, was wir mit der Funktion `ifelse` erreichen. Die Ergebnisse speichern wir in das Objekt `d3`.

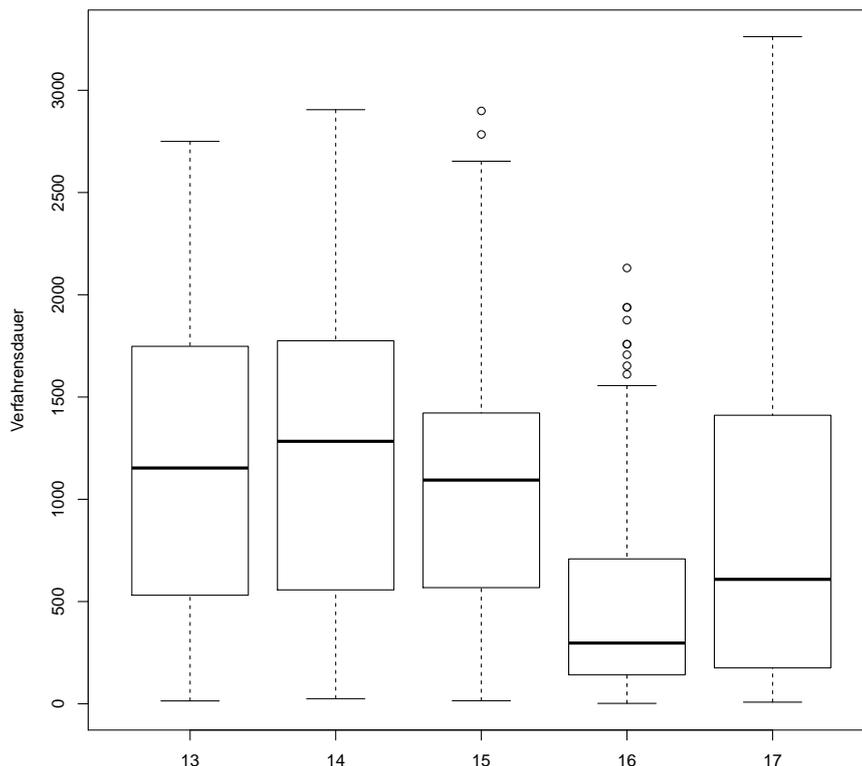
R

```
> vwgh <- read.csv2("vwgh.csv", header = TRUE)
> attach(vwgh)
> d3 <- ifelse(dauer3 == -9999, NA, dauer3)
```

- Die Verfahrensdauer nach Senat kann man mit einem Boxplot veranschaulichen, wobei man hier sieht, dass die Verteilungen schief sind und wahrscheinlich heterogene Varianzen vorliegen.

R

```
> boxplot(d3 ~ senat, ylab = "Verfahrensdauer")
```



- Aufgrund der schiefen Verteilungen und unterschiedlichen Varianzen berechnen wir statt einer Varianzanalyse einen Kruskal-Wallis-Test. Wir sehen, dass $p < .001$ ist, wodurch wir die Nullhypothese verwerfen und davon ausgehen, dass es Unterschiede zwischen den Senaten hinsichtlich der Verfahrensdauer gibt.

R

```
> kruskal.test(d3 ~ senat)
```

Kruskal-Wallis rank sum test

data: d3 by senat

Kruskal-Wallis chi-squared = 631.0894, df = 4, p-value < 2.2e-16

R

```
> detach(vwgh)
```