Einführung in R und RStudio

[Termine 5 & 6]

Miguel Alvarez

23. & 24. März 2023

Der Kurs

- **Termin 1 & 2**
  - Grundlagen
  - Datentypen
- **Termin 3 & 4**
  - Objekten
  - Lesen und Schreiben
- **Termin 5 & 6**
  - Statistiken
  - Graphiken (1)
- **Termin 7 & 8**
  - Graphiken (2)
  - Fortgeschrittenes Programmieren
  - Abschluss

Daten

- Datenformat
  - csv (comma-separated values)
  - ods (OpenDocument spreadsheet)
  - xlsx (Excel spreadsheet)

- Datenstruktur
  - Spalten-orientierte Tabelle (Datenbank-Liste)
  - Pivot-Tabelle (Kreuztabelle)

Daten

Pivot Tabellen und Zusammenfassungen

- Packae `tidyr`
  - `pivot_longer()`
  - `pivot_wider()`
- Package `stats`
  - `aggregate()`

```
aggregate(Gesamt ~ StadtBezirk,
    data = Bevoelkerung,
    FUN = sum)
```

```
##     StadtBezirk Gesamt
## 1          Bonn 466564
## 2 Bad Godesberg 228717
## 3         Beuel 203412
## 4     Hardtberg 103845
```

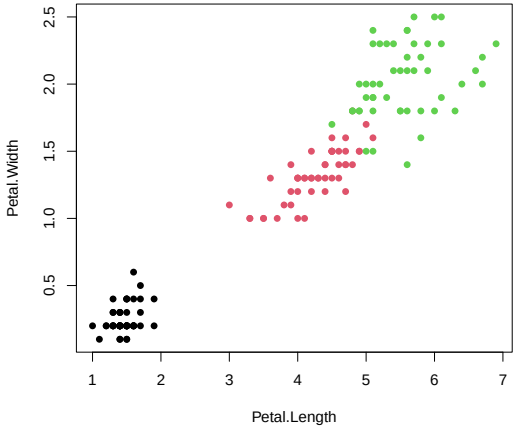
Statistik

- Deskriptive Statistik
  - Explorative Datenauswertung
- Hypothesen Test
  - Parametrische Statistik
  - Nicht-parametrische Statistik
- Modellierung
- Multivariate Statistik

Plots (1)

scatterplots

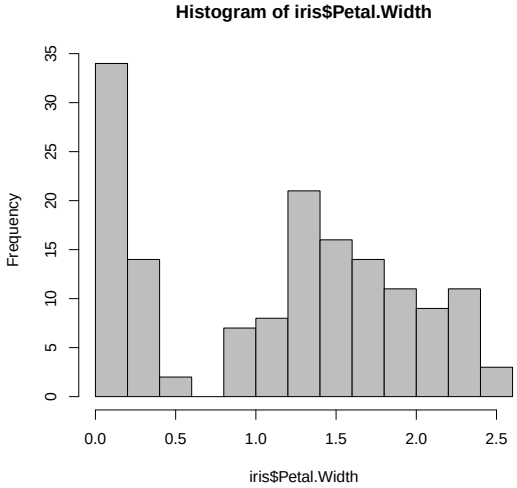
```
plot(iris[, c("Petal.Length","Petal.Width")],
     col = iris$Species, pch = 16)
```



Plots (1)

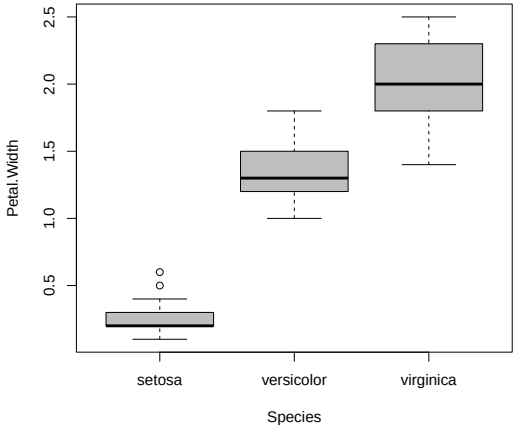
histograms

```
hist(iris$Petal.Width, col = "grey")
```



box plots

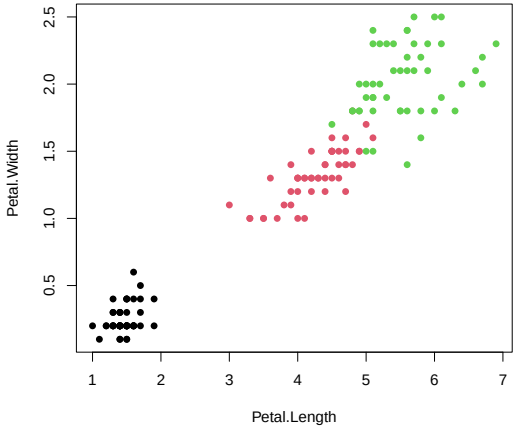
```
boxplot(Petal.Width ~ Species,
    data = iris, col = "grey")
```



Plots (1)

High Level Functions

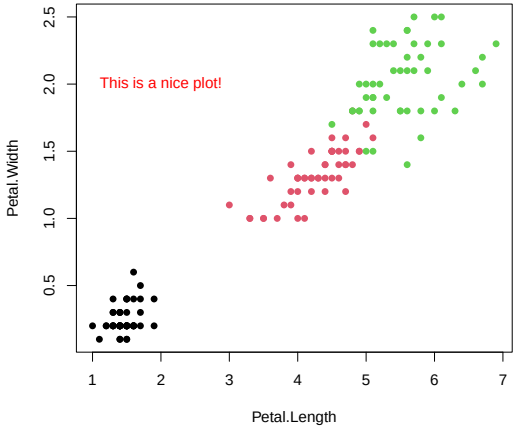
```
plot(iris[, c("Petal.Length","Petal.Width")],
        col = iris$Species, pch = 16)
```



Plots (1)

Low Level Functions

```
text(2, 2, label = "This is a nice plot!",
  col = "red")
```



Plots (1)

- High Level Functions
  - `plot()`
  - `hist()`
  - `barplot()`
  - `pie()`
  - `boxplot()`
- Low Level Functions
  - `text()`
  - `points()`
  - `axis()`
  - `mtext()`

Some Links

R Base Graphics

R Color Chart

Vielen Dank!

```
library(fortunes)
fortune(269)
```

```
##
## Knut Krueger: Is there any function available to combine those p values?
## Stephan Kolassa: ?"+"
##    -- Knut Krueger and Stephan Kolassa (about ways to combine p values from
##       different tests)
##       R-help (July 2010)
```