

R Stats

M Alvarez

Der Kurs

Einführung in
R

RStudio

R als Taschen-
rechner

Daten in R

Statistik und Graphiken mit R

[Termin 1]

Miguel Alvarez

27. Februar 2024

Der Kurs

R Stats

M Alvarez

Der Kurs

Einführung in
R

RStudio

R als Taschen-
rechner

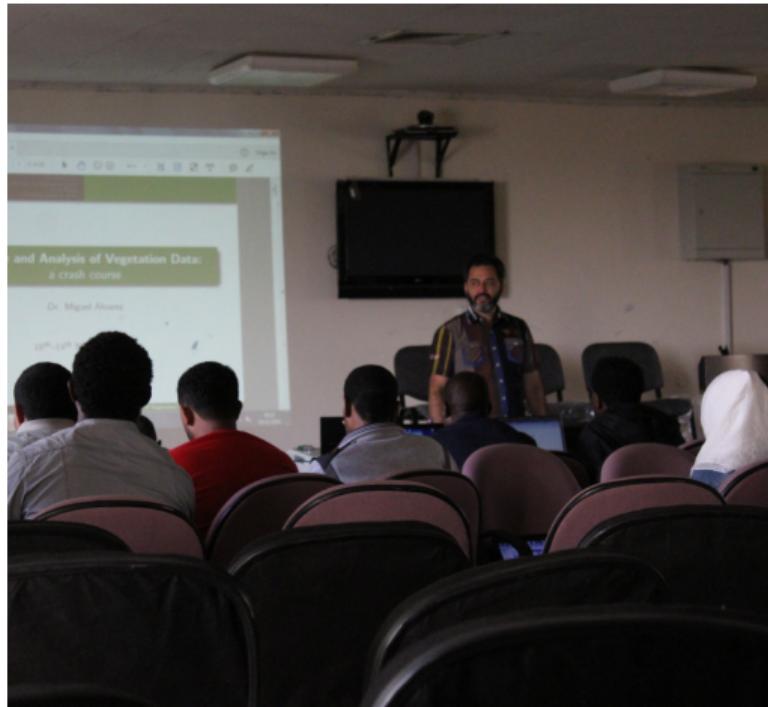
Daten in R

18:00 – 20:30

Pause 19:15 – 19:30

- ▶ Methoden
 - ▶ Vortrag
 - ▶ Life-Codierung
 - ▶ Übungen

<https://kamapu.github.io/GrundkursR/>



R Stats

M Alvarez

Der Kurs

Einführung in
R

RStudio

R als Taschen-
rechner

Daten in R

► **Termin 1 & 2**

- ▶ Grundlagen
- ▶ Datentypen

► **Termin 3 & 4**

- ▶ Objekten (Datenstrukturen)
- ▶ Lesen und Schreiben

► **Termin 5 & 6**

- ▶ Statistiken
- ▶ Graphiken (1)

► **Termin 7 & 8**

- ▶ Graphiken (2)
- ▶ Fortgeschrittenes Programmieren
- ▶ Abschluss

Der Kurs

R Stats

M Alvarez

Der Kurs

Einführung in
R

RStudio

R als Taschen-
rechner

Daten in R

Miguel Alvarez



- ▶ Vegetationsökologist
- ▶ Datenwissenschaftler
- ▶ KI-Experte
- ▶ R (und Python) Programmierer
 - ▶ Umwelt und Biodiversität
 - ▶ Geographische Informationssysteme
 - ▶ Datenbanken
 - ▶ Reproduzierbarkeit

Open Source Aktivist

R Stats

M Alvarez

Der Kurs

Einführung in
R

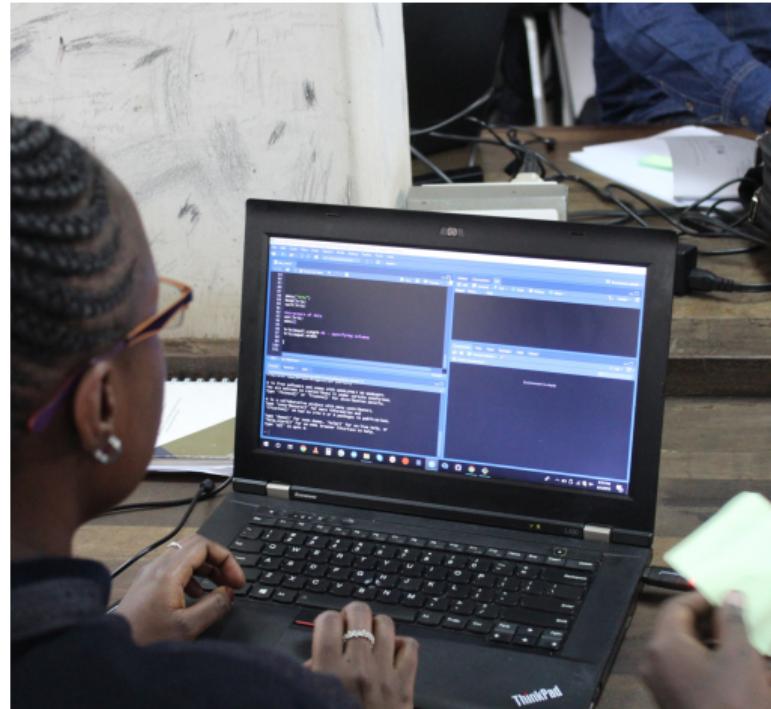
RStudio

R als Taschen-
rechner

Daten in R

Die Teilnehmer?

- ▶ Warum R?
- ▶ Erwartung von dem Kurs



Geschichte

- ▶ S (1975)
 - ▶ S-PLUS
 - ▶ TIBCO Spotfire S+
- ▶ R (1992)

R ist die kostenlose Alternative zu **S**

Ross Ihaka

Robert Gentleman



R Stats

M Alvarez

Der Kurs

Einführung in
R

RStudio

R als Taschen-
rechner

Daten in R

Was ist R?

- ▶ Programmiersprache
 - ▶ Open Source (offene Quelle)
 - ▶ Freeware (kostenlos)
- ▶ Statistische Umgebung
 - ▶ Interface
 - ▶ Terminal
 - ▶ Editoren

<https://www.r-project.org/>



[Home]

Download

CRAN

R Project

About R

Logo

Contributors

What's New?

Reporting Bugs

Conferences

Search

Get Involved: Mailing Lists

Get Involved: Contributing

Developer Pages

R Blog

R Foundation

Foundation

Board

Members

Donors

Donate

Help With R

Getting Help

Documentation

Manuals

FAQs

The R Journal

Books

Certification

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and Mac OS. To [download R](#), please choose your preferred CRAN mirror.

If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers](#) to frequently asked questions before you send an email.

News

- [R version 4.2.2 \(Innocent and Trusting\) prerelease versions](#) will appear starting Friday 2022-10-21. Final release is scheduled for Monday 2022-10-31.
- [R version 4.2.1 \(Funny-Looking Kid\)](#) has been released on 2022-06-23.
- [R version 4.1.3 \(One Push-Up\)](#) was released on 2022-03-10.
- Thanks to the organisers of useR! 2020 for a successful online conference. Recorded tutorials and talks from the conference are available on the [R Consortium YouTube channel](#).
- You can support the R Foundation with a renewable subscription as a supporting member

News via Twitter

⌚ The R Foundation Retweeted

 [R Contributo...](#) @R_Contributo... · Oct 10 
[R Contribution Working Group meeting Oct 18, 18:30 - 19:30 UTC](#)

We'll discuss progress on current issues (github.com/r-devel/rcontr_), including

- Translation hackathons (LatinR, AsiaR)
- Office hours
- Code of conduct

All welcome. Zoom registration:
us02web.zoom.us/meeting/register_...

Was ist R?

- ▶ Statistisches Programmieren
 - ▶ Mathematik
 - ▶ Grafiken
- ▶ Kommandozeilen
- ▶ REPL (Read-Eval-Print-Schleife)

```
(5 + 100) * 25
```

```
## [1] 2625
```

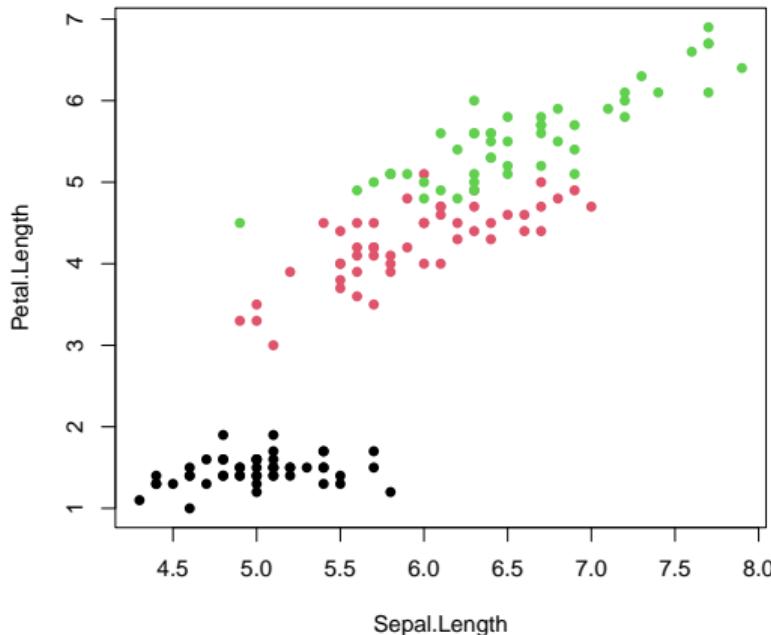
```
50/10
```

```
## [1] 5
```

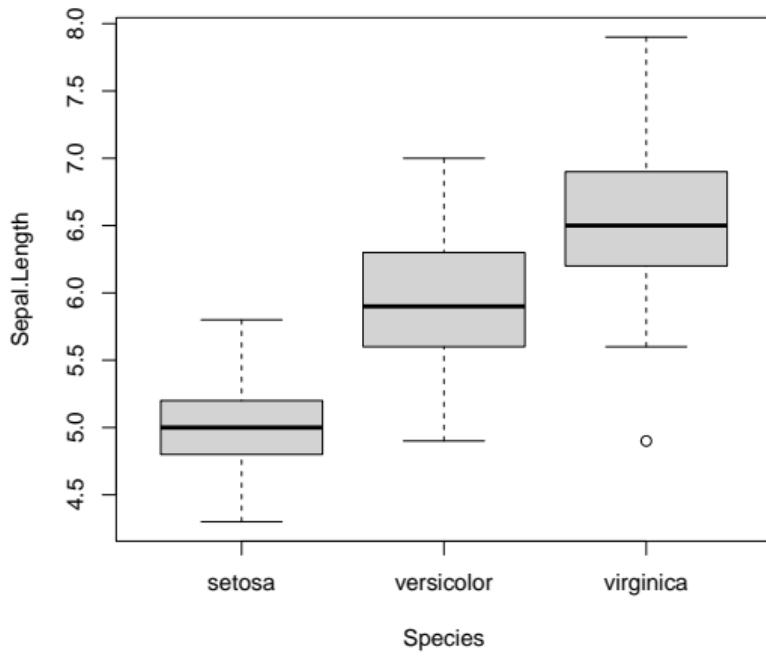
```
60 > 15
```

```
## [1] TRUE
```

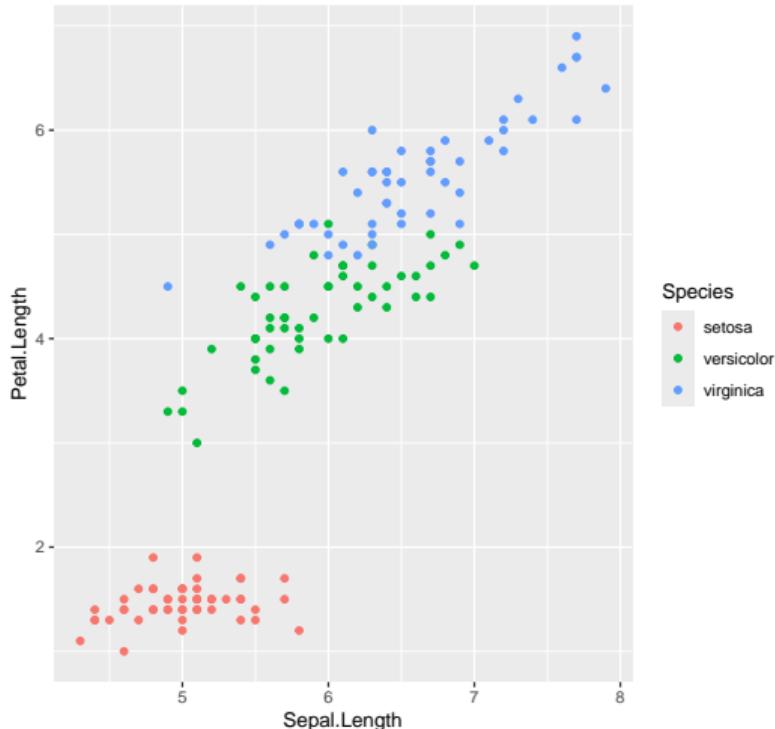
```
plot(Petal.Length ~ Sepal.Length,  
      pch = 16, col = Species,  
      data = iris)
```



```
boxplot(Sepal.Length ~ Species,  
       data = iris)
```



```
library(ggplot2)
ggplot(iris,
       aes(x = Sepal.Length,
            y = Petal.Length,
            color = Species)) +
  geom_point()
```



Warum R?

- ▶ Kostet nichts
- ▶ Steuerung von Analysen
- ▶ Skripten
 - ▶ Protokolle
 - ▶ Reproduzierbarkeit
- ▶ Vielseitig

```
# Wir laden zunächst die Daten
data(iris)
```

```
# Wir berechnen die Mittelwerte für
# Länge und Breite von Petalen
# per Schwertlilienart
aggregate(cbind(Petal.Length, Petal.Width) ~ Species,
          FUN = mean, data = iris)

##      Species Petal.Length Petal.Width
## 1      setosa       1.462      0.246
## 2  versicolor       4.260      1.326
## 3 virginica        5.552      2.026
```

Wichtige Elemente

- ▶ Konsole
- ▶ Sitzung
- ▶ Workspace (Environment)
- ▶ Arbeitsverzeichnis (working directory)
- ▶ Skript
- ▶ Paket (Package)

R Stats

M Alvarez

Der Kurs

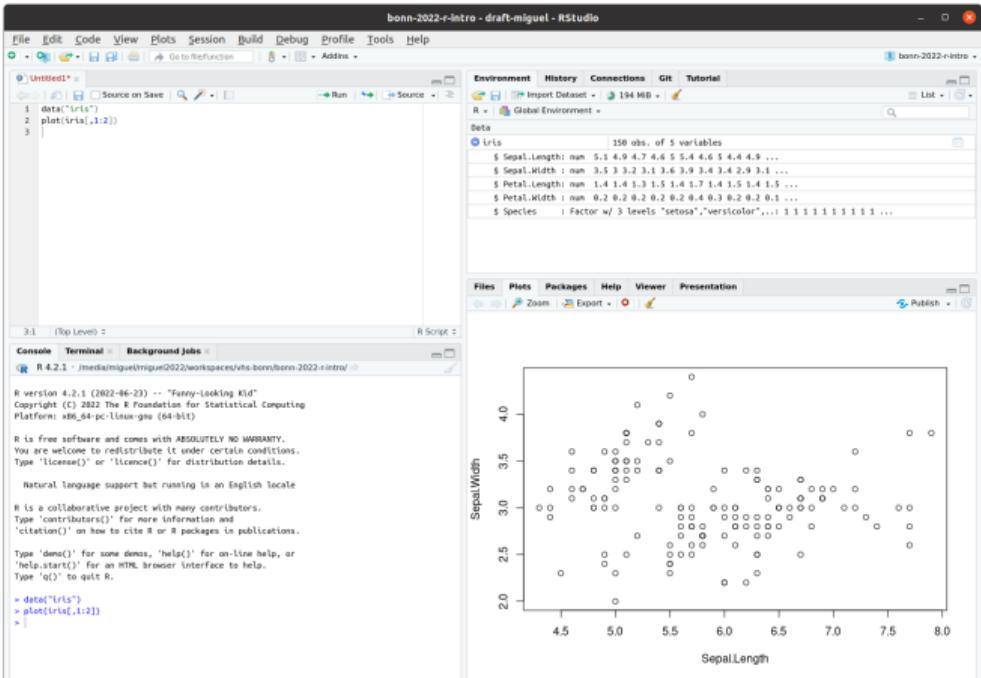
Einführung in R

RStudio

R als Taschenrechner

Daten in R

- ▶ Text Editor
- ▶ IDE (integrated development environment)



R Stats

M Alvarez

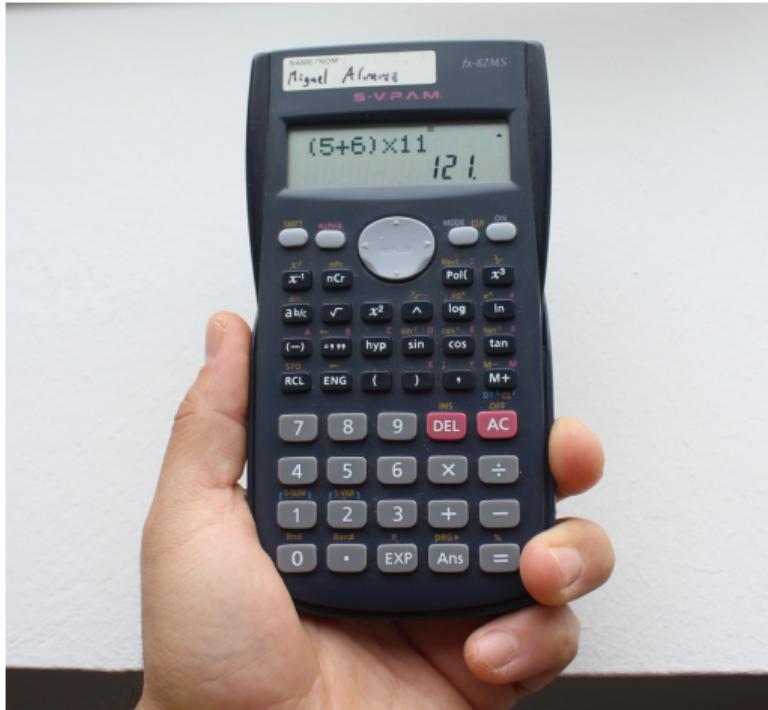
Der Kurs

Einführung in R

RStudio

R als Taschenrechner

Daten in R



- ▶ Mathematische Operatoren
- ▶ Klammern

 $5+6*11$

```
## [1] 71
```

 $(5+6)*11$

```
## [1] 121
```

- ▶ Zuweisung `<-`

```
M <- 5
```

```
M + 2
```

```
## [1] 7
```

R Stats

M Alvarez

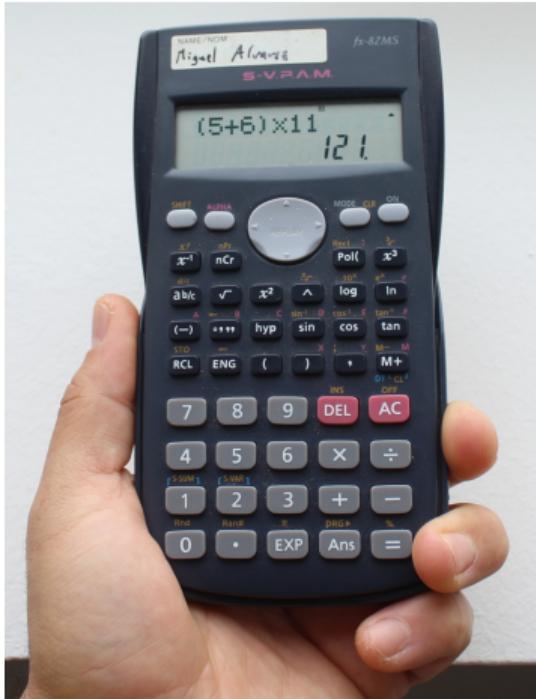
Der Kurs

Einführung in
R

RStudio

R als Taschenrechner

Daten in R



Aufgabe 1

Haben/Hatten Sie einen Taschenrechner?
Erzählen Ihre Geschichte.

Aufgabe 2

Taschenrechnen mit R

R Stats

M Alvarez

Der Kurs

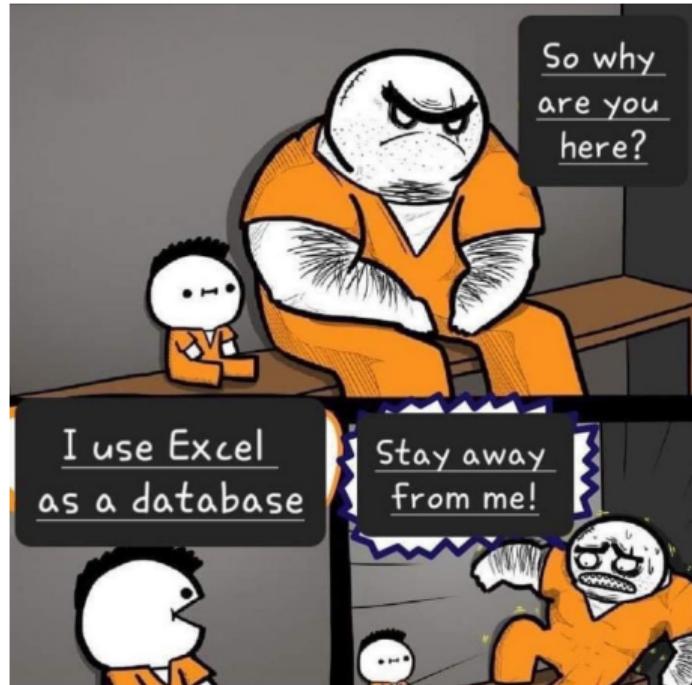
Einführung in R

RStudio

R als Taschenrechner

Daten in R

`==``!=``>``>=``<``<=``&``|``%in%``!``any()``all()``10 > 15``## [1] FALSE``10 < 15``## [1] TRUE`

[R Stats](#)[M Alvarez](#)[Der Kurs](#)[Einführung in R](#)[RStudio](#)[R als Taschenrechner](#)[Daten in R](#)

Which types of data do you know?

Vektoren

Der Vektor ist die grundlegende Datenstruktur in R

- ▶ Länge `length()`
- ▶ Typ `class()`
- ▶ Evtl. Namen `names()`

```
c(1:10)
## [1] 1 2 3 4 5 6 7 8 9 10
rep(5, times = 10)
## [1] 5 5 5 5 5 5 5 5 5 5
LETTERS[1:5]
## [1] "A" "B" "C" "D" "E"
```

Indexieren

- ▶ Eckige Klammern
- ▶ Index
 - ▶ integer
 - ▶ logical (Bedingung)
 - ▶ character (Namen)

```
# Mit integer
letters[15]
## [1] "o"
```

```
# Mit logischen Werten
letters[!letters %in% c("a", "b", "c")]
## [1] "d" "e" "f" "g" "h" "i" "j" "k" "l" "m"
## [20] "w" "x" "y" "z"
```

```
# Mit Namen
names(letters) <- letters
letters["m"]
##     m
## "m"
```

R Stats

M Alvarez

Der Kurs

Einführung in
R

RStudio

R als Taschen-
rechner

Daten in R

Datentypen

- ▶ integer
- ▶ numeric
- ▶ logical
- ▶ factor
- ▶ character

```
A <- c(1:10)
is.numeric(A)
## [1] TRUE
```

```
B <- as.character(A)
B
## [1] "1"  "2"  "3"  "4"  "5"  "6"  "7"  "8"
is.numeric(B)
## [1] FALSE
```

Sonderklassen

- ▶ NA
- ▶ NaN
- ▶ NULL
- ▶ Inf
- ▶ -Inf

```
5/0
## [1] Inf
log(0)
## [1] -Inf
sqrt(-1)
## Warning in sqrt(-1): NaNs produced
## [1] NaN
```

Vielen Dank!

```
library(fortunes)
fortune(283)

##
## The good way to do it is to include the following comment at the beginning:
## # This is a holy Script, please edit it not
## -- Kenn Konstabel (on "... how to protect R Script files from inadvertent
##       editing by users.")
##       R-help (April 2011)
```