

Statistik und Graphiken mit R

[Termine 1 & 2]

Miguel Alvarez

27. & 29. Februar 2024

▶ Termin 1 & 2

- ▶ Grundlagen
- ▶ Datentypen

▶ Termin 3 & 4

- ▶ Objekten (Datenstrukturen)
- ▶ Lesen und Schreiben

▶ Termin 5 & 6

- ▶ Statistiken
- ▶ Graphiken (1)

▶ Termin 7 & 8

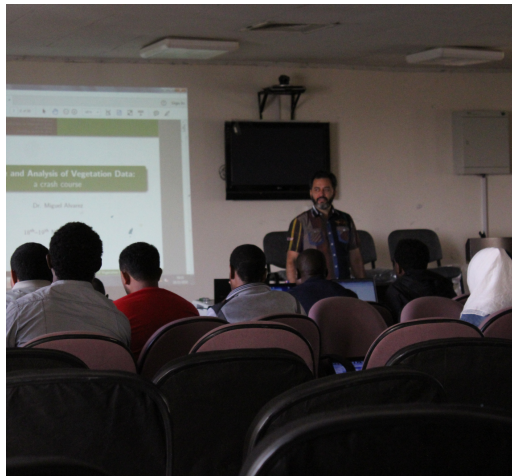
- ▶ Graphiken (2)
- ▶ Fortgeschrittenes Programmieren
- ▶ Abschluss

18:00 – 20:30

Pause 19:15 – 19:30

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

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



Miguel Alvarez

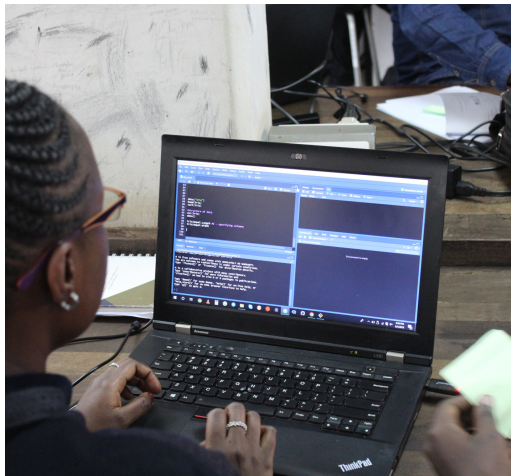


- ▶ Vegetationsökologist
- ▶ Datenwissenschaftler
- ▶ R (und Python) Programmierer
 - ▶ Umwelt und Biodiversität
 - ▶ Geographische Informationssysteme
 - ▶ Datenbanken
 - ▶ Reproduzierbarkeit

Open Source Aktivist

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



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 MacOS. 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/joining/regist...

Was ist R?

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

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

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

```
50/10
```

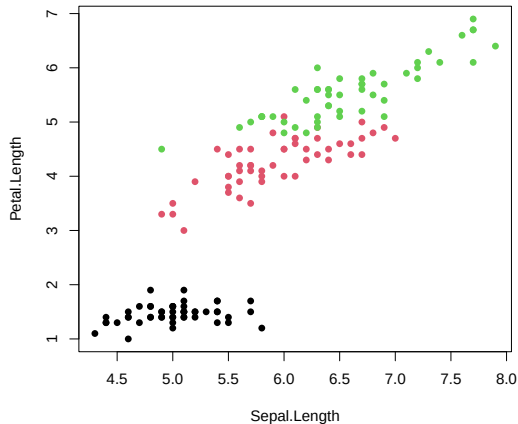
```
## [1] 5
```

```
60 > 15
```

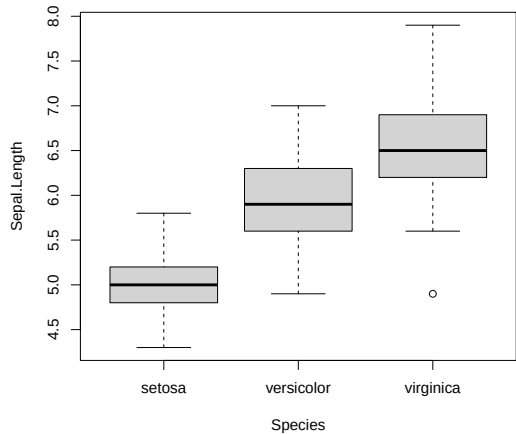
```
## [1] TRUE
```



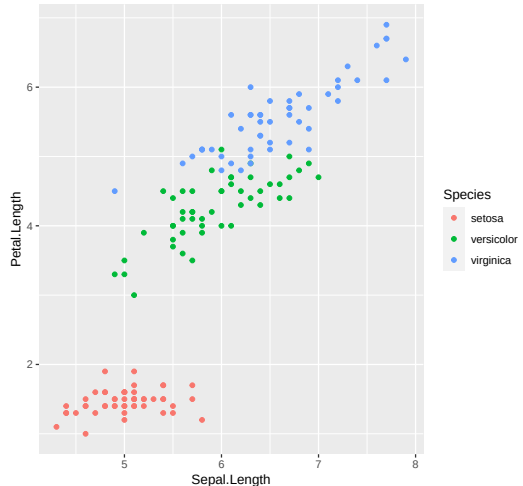
```
plot(Petal.Length ~ Sepal.Length,  
     pch = 16, col = Species,  
     data = iris)
```



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



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



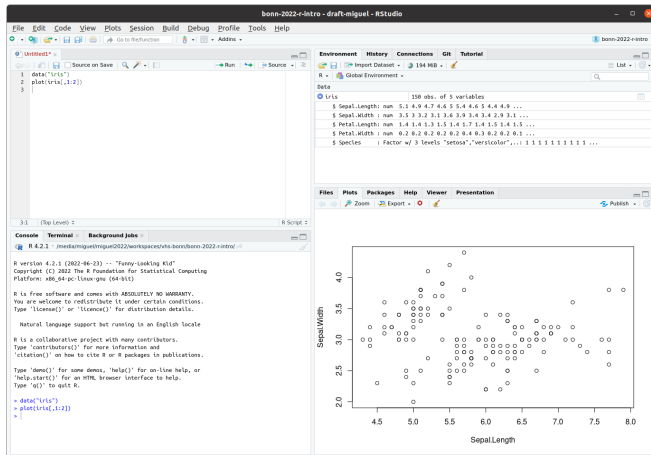
Warum R?

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

Jargon

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

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



eclipse (StatET)

The screenshot shows the Eclipse IDE with the RStudio plugin. The Project Explorer on the left displays a file tree for a project named 'vhs-bonn - bonn-2022-r-intro'. The Editor in the center shows the R code for 'FolienRKurs-12.Rmd', which includes a section for 'RStudio' and 'R als Taschenrechner'. The Console at the bottom shows the R startup message. The right-hand side of the IDE displays a scatter plot of 'Sepal.Length' vs 'Petal.Length' for the 'Iris' dataset.

```

322 *** (r echoFALSE, out_width="0.7\\textwidth")
323 include_graphics("images/Rstudio-screen.png")
324 ***
325
326
327 ## RStudio
328
329 **eclipse (StatET)**
330
331 \\vspace{0.3cm}
332
333 *** (r echoFALSE, out_width="0.7\\textwidth")
334 include_graphics("images/StatET-screen.png")
335 ***
336
337
338 ## R als Taschenrechner
339
340
341 - Mathematische Operatoren
342 - Logische Operatoren
343 - Klammern

```

R [R Console] R / R [Oct 16, 2022, 5:45:24 PM] - ... / vhs-bonn/bonn-2022-r-intro

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

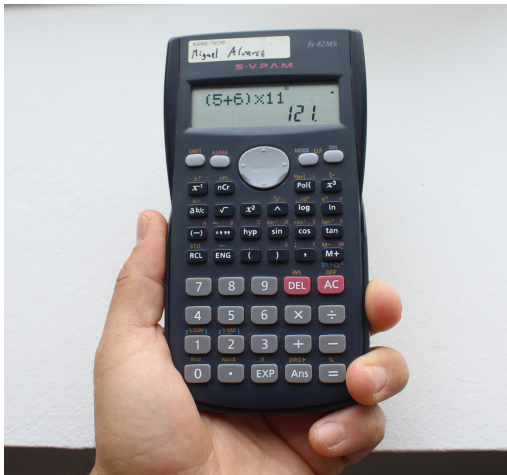
```

> data(iris)
> plot(iris[, 1:2])

```

Submit

Scatter plot showing Sepal.Length (x-axis, 4.5 to 8.0) vs Petal.Length (y-axis, 2.0 to 4.0) for the Iris dataset.



- ▶ Mathematische Operatoren
- ▶ Klammern

```
5+6*11
```

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

```
(5+6)*11
```

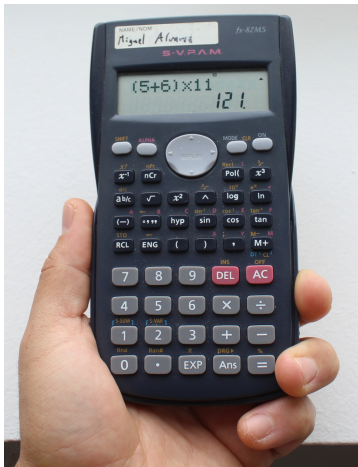
```
## [1] 121
```

- ▶ Zuweisung <-

```
M <- 5
```

```
M + 2
```

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

Aufgabe 1

Haben/Hatten Sie einen Taschenrechner?
Erzählen Ihre Geschichte.

Aufgabe 2

Taschenrechnen mit R

==

!=

>

>=

<

<=

&

|

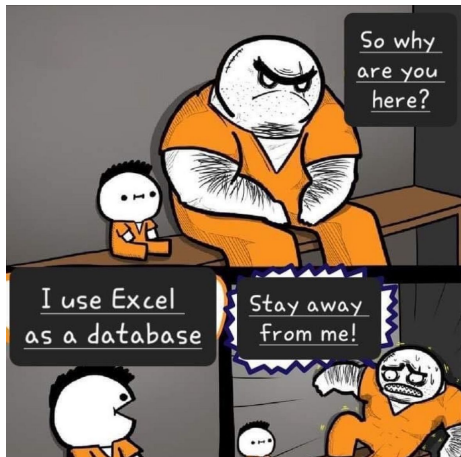
%in%

!

any()

all()

`10 > 15``## [1] FALSE``10 < 15``## [1] TRUE`



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"
```

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)
```