Statistik und Graphiken mit R

[Termin 2]

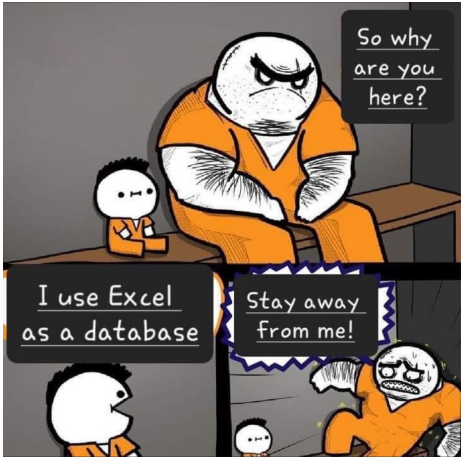
Miguel Alvarez

29. Februar 2024

Der Kurs

- **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

Vektoren



Which types of data do you know?

Vektoren

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"
```

Vektoren

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
```

Vektoren

Sonderklassen

- NA
- NaN
- NULL
- Inf
- -Inf

```
5/0
```

```
## [1] Inf
```

```
log(0)
```

```
## [1] -Inf
```

```
sqrt(-1)
```

```
## Warning in sqrt(-1): NaNs produced
## [1] NaN
```

Objekte

Funktion

`foo(par1 = arg1, ..., parn = argn)`

Funktionen und Argumente (Parameter) werden dokumentiert.
Achte auf Standardeinstellungen (default values).

```
A <- c(1, NA, 3, 5)
mean(A)
```

```
## [1] NA
```

```
mean(A, na.rm = TRUE)
```

```
## [1] 3
```

Objekte

Matrix

- Typ von Inhalt (`mode()`).
- Zwei Dimensionen.

```
M <- matrix(1:20, nrow = 4)
M
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    5    9   13   17
## [2,]    2    6   10   14   18
## [3,]    3    7   11   15   19
## [4,]    4    8   12   16   20
```

```
class(M)
```

```
## [1] "matrix" "array"
```

```
mode(M)
```

```
## [1] "numeric"
```

```
length(M)
```

```
## [1] 20
```

```
dim(M)
```

```
## [1] 4 5
```

Objekte

Liste

Liste (Sammlung) von Objekten, inklusive Listen.
Achte, dass `data.frame` eine spezielle Form von `list` ist.

```
MeineListe <- list(
    A = 1:10,
    B = matrix(1:10, nrow = 2),
    C = "Dies ist eine Liste")
MeineListe
```

```
## $A
##  [1]  1  2  3  4  5  6  7  8  9 10
##
## $B
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    3    5    7    9
## [2,]    2    4    6    8   10
##
## $C
## [1] "Dies ist eine Liste"
```

Objekte

Datensatz

Spaltenorientierte Tabelle (`data.frame`)

```
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal
## 1          5.1         3.5          1.4
## 2          4.9         3.0          1.4
## 3          4.7         3.2          1.3
## 4          4.6         3.1          1.5
## 5          5.0         3.6          1.4
## 6          5.4         3.9          1.7
```

```
str(iris)
```

```
## 'data.frame':    150 obs. of  5 variables:
##  $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4
##  $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9
##  $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.
##  $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.
##  $ Species     : Factor w/ 3 levels "setosa"
```

Pakete

CRAN

(Comprehensive R Archive Network)

- `install.packages()`
- `update.packages()`

```
install.packages("ade4")
update.packages(ask = FALSE)
```

https://cran.r-project.org/

Available CRAN Packages By Name

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

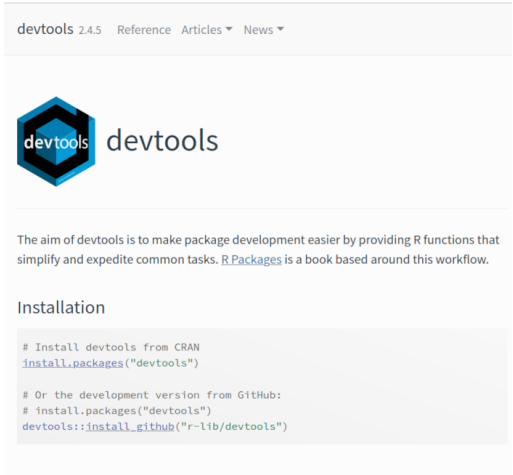
A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
AATtools	Reliability and Scoring Routines for the Approach-Avoidance Task
ABACUS	Apps Based Activities for Communicating and Understanding Statistics
abbreviate	Readable String Abbreviation
abbyyR	Access to Abbyy Optical Character Recognition (OCR) API
abc	Tools for Approximate Bayesian Computation (ABC)
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
ABC.RAP	Array Based CpG Region Analysis Pipeline
abcADM	Fit Accumulated Damage Models and Estimate Reliability using ABC
ABCanalysis	Computed ABC Analysis
abclass	Angle-Based Large-Margin Classifiers
ABCoptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2
abcrf	Approximate Bayesian Computation via Random Forests
abcrlda	Asymptotically Bias-Corrected Regularized Linear Discriminant Analysis
abctools	Tools for ABC Analyses
abd	The Analysis of Biological Data
abdiv	Alpha and Beta Diversity Measures
abe	Augmented Backward Elimination
abess	Fast Best Subset Selection
abglasso	Adaptive Bayesian Graphical Lasso
ABHgenotypeR	Easy Visualization of ABH Genotypes
abind	Combine Multidimensional Arrays
abjData	Databases Used Routinely by the Brazilian Jurimetrics Association
abjutils	Useful Tools for Jurimetrical Analysis Used by the Brazilian Jurimetrics Association
abmR	Agent-Based Models in R
abn	Modelling Multivariate Data with Additive Bayesian Networks
abnormality	Measure a Subject's Abnormality with Respect to a Reference Population
abodOutlier	Angle-Based Outlier Detection
ABPS	The Abnormal Blood Profile Score to Detect Blood Doping
abstr	R Interface to the A/B Street Transport System Simulation Software
abstractr	An R-Shiny Application for Creating Visual Abstracts
abtest	Bayesian A/B Testing
abundant	High-Dimensional Principal Fitted Components and Abundant Regression
Ac3net	Inferring Directional Conservative Causal Core Gene Networks
ACA	Abrupt Change-Point or Aberration Detection in Point Series
academictwitteR	Access the Twitter Academic Research Product Track V2 API Endpoint

Pakete

devtools

- `install()`
- `install_github()`

https://ropensci.org/



Import/Export

- readLines()
- read.table()
  - read.csv()
  - read.csv2()

```
Bonn2021 <- read.csv("Bevoelkerung-2021.csv")
str(Bonn2021)
```

```
## 'data.frame':    67 obs. of  13 variables:
##  $ BezirkNr          : int  110 111 112 113 114 115 1
##  $ BezirkName        : chr  "Zentrum-Rheinviertel" "Z
##  $ Gesamt            : int  2343 3161 6768 8906 5157
##  $ DichteKm2         : int  6508 6585 11874 16193 433
##  $ Maenner           : int  1166 1537 3189 4575 2481
##  $ MaennerProzent    : num  49.8 48.6 47.1 51.4 48.1
##  $ Frauen            : int  1177 1624 3579 4331 2675
##  $ FrauenProzent     : num  50.2 51.4 52.9 48.6 51.9
##  $ Zuwanderer        : int  753 1092 1762 2732 1873 2
##  $ ZuwandererProzent : num  32.1 34.5 26 30.7 36.3 35
##  $ Auslaender        : int  494 813 1145 2010 1235 12
##  $ AuslaenderProzent : num  65.6 74.5 65 73.6 65.9 55
##  $ AuslaenderProzent2: logi  NA NA NA NA NA NA ...
```

Import/Export

- `readLines()`
- `read.table()`
  - `read.csv()`
  - `read.csv2()`

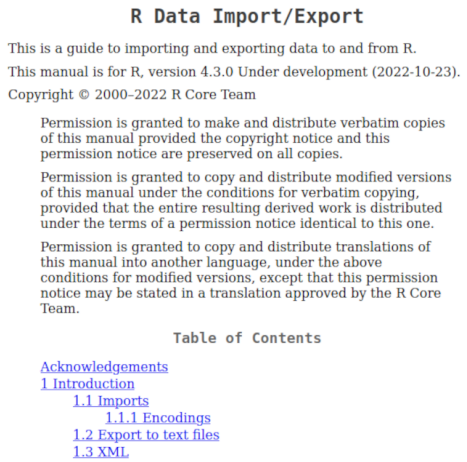
- `write.table()`
  - `write.csv()`
  - `write.csv2()`

```
write.csv(iris, file = "iris.csv")
write.csv2(iris, file = "iris2.csv")
```

Import/Export

Pakete können eigene Funktionen für Importieren und Exportieren anbieten.

- **xlsx**
  - `read.xlsx()`
  - `write.xlsx()`
- **readODS**
  - `read_ods()`
  - `write_ods()`



R Data Import/Export

This is a guide to importing and exporting data to and from R.

This manual is for R, version 4.3.0 Under development (2022-10-23).

Copyright © 2000–2022 R Core Team

Permission is granted to make and distribute verbatim copies of this manual provided the copyright notice and this permission notice are preserved on all copies.

Permission is granted to copy and distribute modified versions of this manual under the conditions for verbatim copying, provided that the entire resulting derived work is distributed under the terms of a permission notice identical to this one.

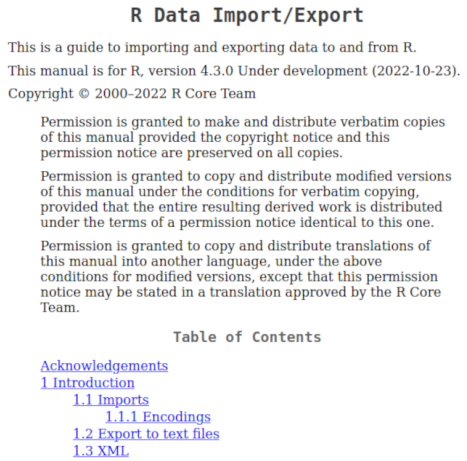
Permission is granted to copy and distribute translations of this manual into another language, under the above conditions for modified versions, except that this permission notice may be stated in a translation approved by the R Core Team.

Table of Contents

Import/Export

Pakete können eigene Funktionen für Importieren und Exportieren anbieten.

- **xlsx**
  - `read.xlsx()`
  - `write.xlsx()`
- **readODS**
  - `read_ods()`
  - `write_ods()`



R Data Import/Export

This is a guide to importing and exporting data to and from R.

This manual is for R, version 4.3.0 Under development (2022-10-23).

Copyright © 2000–2022 R Core Team

Permission is granted to make and distribute verbatim copies of this manual provided the copyright notice and this permission notice are preserved on all copies.

Permission is granted to copy and distribute modified versions of this manual under the conditions for verbatim copying, provided that the entire resulting derived work is distributed under the terms of a permission notice identical to this one.

Permission is granted to copy and distribute translations of this manual into another language, under the above conditions for modified versions, except that this permission notice may be stated in a translation approved by the R Core Team.

Table of Contents

Import/Export

R-Images

- Workspace
  - `save()`
  - `load()`
  - Dateierweiterung **.rda** oder **.RData**
- Einzelnes Objekt
  - `saveRDS()`
  - `readRDS()`
  - Dateierweiterung **.rds**

Vielen Dank!

```
library(fortunes)
fortune(10)
```

```
##
## Overall, SAS is about 11 years behind R and S-Plus in statistical capabilities
## (last year it was about 10 years behind) in my estimation.
##    -- Frank Harrell (SAS User, 1969-1991)
##       R-help (September 2003)
```