

# A1 Erste Schritte mit R

Jan-Philipp Kolb

11 Januar 2019

# Disclaimer/ Informationen vorab

Normalerweise gibt es große Unterschiede bei Vorkenntnissen und Fähigkeiten - bitte gebt Bescheid, wenn es zu schnell oder zu langsam geht oder etwas unklar geblieben ist.

- Wenn es Fragen gibt - immer fragen
- In diesem Kurs gibt es viele **Übungen**, denn das Programmieren / die Nutzung von R lernt man am Ende nur allein.
- Ich habe viele **Beispiele** - probiert sie aus
- R macht mehr Spaß zusammen - arbeitet zusammen!

## Motivation allgemein

- Raumbezug herstellen/nutzen
- Sekundäranalyse für bestehenden Daten
- Analysepotentiale der Geokodierung vorstellen
- Verbindung von sozial- mit raumwissenschaftlichen Daten

## Warum die Darstellung in Karten

# Was heißt das für diesen Kurs

## Vorgestellt werden:

- Möglichkeiten für den Download, den Import, die Verarbeitung und die Visualisierung von Geodaten
- Quellen für Geodaten
- Bspw. die wichtigsten Prognosen Betreff Erhalten Größe Kategorien Henze, Jost AG Umsetzung Führungsleitbild 21.12.2018 22 KB rammierschnittstellen (APIs) um die Daten zu bekommen
- R-Pakete um diese Daten zu verarbeiten und zu visualisieren

## Georeferenzierung von Daten

Situation und Zukunft der  
Geodatenlandschaft in  
Deutschland

Herausgegeben vom Rat für Sozial- und Wirtschaftsdaten

RatSWD ■

Rat für Sozial- und  
Wirtschaftsdaten

GEFÖRDERT VOM



Bundesministerium  
für Bildung  
und Forschung

# Gründe R zu nutzen...

- ... R ist eine **quelloffene Sprache**
- ... hervorragende **Grafiken, Grafiken, Grafiken**
- ... **R kann in Kombination mit anderen Programmen verwendet werden** - z.B. zur **Verknüpfung von Daten**
- ... R kann **zur Automatisierung** verwendet werden
- ... Breite und aktive Community - **Man kann die Intelligenz anderer Leute nutzen ;-)**

# R kann in Kombination mit anderen Programmen genutzt werden...

Use R!

Richard M. Heiberger  
Erich Neuwirth

R Through Excel



SASmixed



rPython R package

Statistics and Computing

Robert A. Muenchen · Joseph M. Hilbe

R for  
Stata Users

[IBM SPSS Statistics Essentials for R: Project Web Hosting - Open Source Software](#)

**IBM SPSS Statistics Essentials for R**

Users

[Download IBM SPSS Statistics  
Essentials for R files](#)

[Donate money](#)

[Project detail and discuss](#)

[Get support](#)

Not what you're looking for?

Open Source Software

# Die Beliebtheit von R-Paketen



# Download R:

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



*CRAN*  
[Mirrors](#)  
[What's new?](#)  
[Task Views](#)  
[Search](#)

*About R*  
[R Homepage](#)  
[The R Journal](#)

*Software*  
[R Sources](#)  
[R Binaries](#)  
[Packages](#)  
[Other](#)

## The Comprehensive R Archive Network

### Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

### Source Code for all Platforms

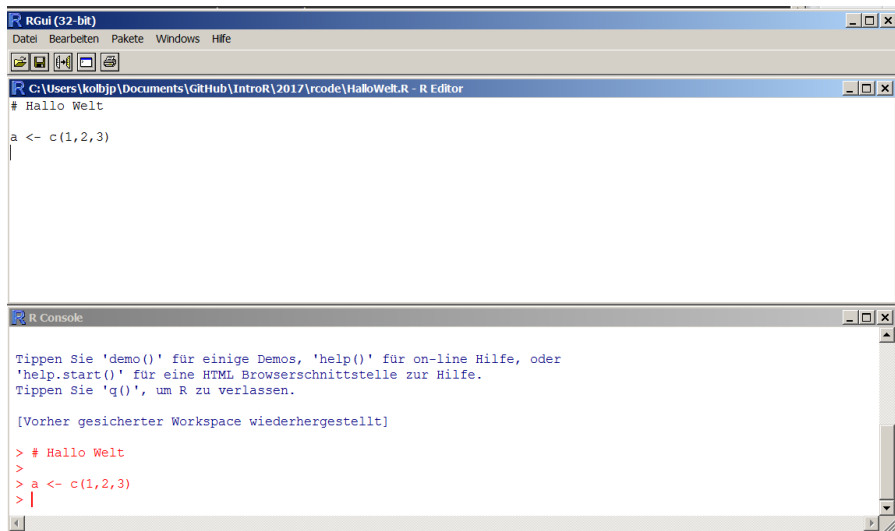
Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (Friday 2017-04-21, You Stupid Darkness)  
[R-3.4.0.tar.gz](#), read [what's new](#) in the latest version.



# Open Source Programm R

## Das ist das Basis-R:



The screenshot displays the R GUI (32-bit) interface. The top window, titled "R GUI (32-bit)", contains a menu bar with "Datei", "Bearbeiten", "Pakete", "Windows", and "Hilfe", and a toolbar with icons for file operations. Below this is the "R C:\Users\kolbjp\Documents\GitHub\IntroR\2017\rcode\HalloWelt.R - R Editor" window, which shows the following R code:

```
# Hallo Welt  
  
a <- c(1,2,3)
```

The bottom window, titled "R Console", shows the output of the code execution:

```
Tippen Sie 'demo()' für einige Demos, 'help()' für on-line Hilfe, oder  
'help.start()' für eine HTML Browserschnittstelle zur Hilfe.  
Tippen Sie 'q()', um R zu verlassen.  
  
[Vorher gesicherter Workspace wiederhergestellt]  
  
> # Hallo Welt  
>  
> a <- c(1,2,3)  
> |
```

# Graphical user interface

Viele Leute nutzen ein **Graphical User Interface** (GUI) oder ein **Integrated Development Interface** (IDE).

Aus den folgenden Gründen:

- Syntax-Hervorhebung
- Auto-Vervollständigung
- Bessere Übersicht über Graphiken, Pakete, Dateien, ...

# RStudio

The screenshot displays the RStudio IDE interface. The main editor window shows an R script with the following content:

```
42 c("A1 Getting started", "A2 How to get help", "A3 Data import", "A4 The GESIS panel data", "A5 data export")
43
44 c("B1 Basic data analysis", "B2 The use of the survey package", "B3 Graphics", "B4 linear regression", "B5 logistic regression")
45
46
47 c("C1 Understanding error messages", "C2 Hierarchical/Multi-level models")
48
49
50 ## {r,echo=F}
51 library(knitr)
52 sched <- xlsx::read.xlsx2("../orga/schedule.xlsx", 1)
53 data <- sched[1:7, c(1, 4, 5)]
54 data <- data[, which(data$Part == "Break"), ]
55
56 kable(data, row.names = F)
57
58
59
60 ## Overview - advantages of R
61
```

The Environment pane on the right shows the current environment with variables like `misc/addon.general.Rmd`, `misc/psw_pres.Rmd`, `misc/psw_pres.pdf`, `slides/A1_GettingStarted.Rmd`, `slides/A1_GettingStarted.pdf`, `slides/A2_How2gethelp.pdf`, and `slides/A2_How2gethelp.pdf`.

The File Explorer pane on the right shows the file structure of the project, including folders like `slides` and files like `A1_GettingStarted.html`, `A1_GettingStarted.Rmd`, `A1_GettingStarted.pdf`, `A1_GettingStarted.tex`, `A2_How2gethelp.html`, `A2_How2gethelp.Rmd`, `A2_How2gethelp.pdf`, `A2_How2gethelp.tex`, `A3_GESISPanel.pdf`, `A3_GESISPanel.Rmd`, `A3_GESISPanel.tex`, `A3_GESISPanel.cache`, `A3_GESISPanel_files`, `A4_DataExport.pdf`, `A4_DataExport.Rmd`, `A4_GESISPanel.cache`, `A4_GESISPanel_files`, `A5_DataExport.pdf`, `A5_DataExport.Rmd`, `B1_DataProcessing.html`, and `B1_DataProcessing.Rmd`.

The Console pane at the bottom shows the output of the script, including the knitr chunk output and the final output file created:

```
ordinary text without R code
|-----| 96%
label: unnamed-chunk-13 (with options)
list of 1
$ eval: symbol F
|-----| 100%
ordinary text without R code

"Dr:/Programme/RStudio/bin/pandoc/pandoc" -RTS -K512m -RTS A2_How2gethelp.utf8.md --to beamer -
-from markdowntoautolink_bare_uris+asci identifiers+tex_math_single_backslash+implicit_figures
--output A2_How2gethelp.tex --variable theme=CambridgeUS --variable colortheme=beaver --varia
ble fonttheme=structurebold --highlight-style tango --pdf-engine pdflatex --self-contained
output file: A2_How2gethelp.knit.md

Output created: A2_How2gethelp.pdf
```

# A1A Übung - Vorbereitung

- Schau, ob R auf dem Computer installiert ist
- Wenn nicht, lade **R** herunter und installiere es.
- Prüfe ob Rstudio installiert ist.
- Wenn nicht - **installiere** Rstudio.
- Starte RStudio. Gehe in die Konsole (meistens Fenster unten links) und tippe

3+2

- Wenn noch kein Skript geöffnet im oberen linken Teil von Rstudio geöffnet ist, gehe zum Menü und öffne ein neues Skript. Checks das Datum mit `date()` und die R version mit `sessionInfo()`.

```
date()
```

```
sessionInfo()
```

# R ist eine objektorientierte Sprache.

## Vektoren und Zuweisungen

- <- ist der Zuweisungsoperator

```
b <- c(1,2) # create an object with the numbers 1 and 2
```

- Auf dieses Objekt kann eine Funktion angewendet werden:

```
mean(b) # computes the mean
```

```
## [1] 1.5
```

Mit diesen Funktionen können wir etwas über die Eigenschaften des Objekts erfahren:

```
length(b) # b has the length 2
```

# A1B Übung - Zuweisungen und Funktionen

Erstellen Sie einen Vektor `b` mit den Zahlen von 1 bis 5 und berechnen Sie. . . .

- 1 den Mittelwert
- 2 die Varianz
- 3 die Standardabweichung
- 4 die Quadratwurzel aus dem Mittelwert

# Wo man Routinen findet

- Viele Funktionen sind in Basis-R enthalten.
- Viele spezifische Funktionen sind in zusätzliche Bibliotheken integriert.
- R kann modular durch sogenannte Pakete oder Bibliotheken erweitert werden.
- Die wichtigsten Pakete, die auf CRAN gehostet werden (13610 at Mo Jan 07)
- Weitere Pakete findet man z.B. unter **bioconductor**

# Übersicht R-Pakete





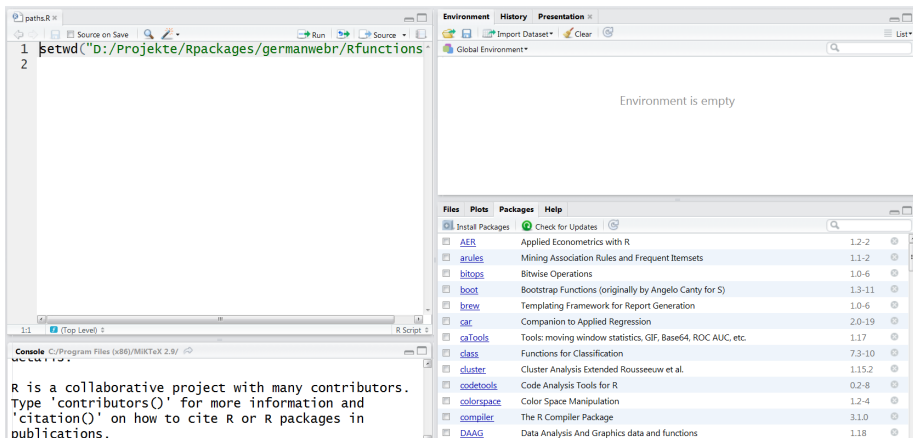
# Installation von Paketen

- Die Anführungszeichen um den Paketnamen herum sind für den Befehl `install.packages` notwendig.
- Sie sind optional für den Befehl `library`.
- Man kann auch `require` anstelle von `library` verwenden.

```
install.packages("raster")
```

```
library(raster)
```

# Installation von Paketen mit RStudio



The screenshot displays the RStudio interface with the following components:

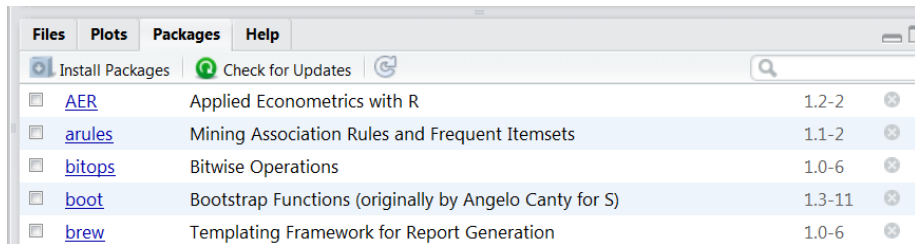
- Script Editor:** Contains the R code:

```
1 setwd("D:/Projekte/Rpackages/germanwebr/Rfunctions")  
2
```
- Environment Pane:** Shows "Global Environment\*" and the message "Environment is empty".
- Packages Pane:** Lists installed and available packages. The installed packages are highlighted in blue.

Package	Description	Version
AER	Applied Econometrics with R	1.2-2
arules	Mining Association Rules and Frequent Itemsets	1.1-2
bitops	Bitwise Operations	1.0-6
boot	Bootstrap Functions (originally by Angelo Canty for S)	1.3-11
brew	Templating Framework for Report Generation	1.0-6
car	Companion to Applied Regression	2.0-19
caTools	Tools: moving window statistics, GIF, Base64, ROC AUC, etc.	1.17
class	Functions for Classification	7.3-10
cluster	Cluster Analysis Extended Rousseeuw et al.	1.15.2
codetools	Code Analysis Tools for R	0.2-8
colorspace	Color Space Manipulation	1.2-4
compiler	The R Compiler Package	3.1.0
DAAG	Data Analysis And Graphics data and functions	1.18

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.

# Bestehende Pakete und Installation



# Übersicht über viele nützliche Pakete:

- Luhmann - Übersicht mit vielen nützlichen Paketen

## Weitere interessante Pakete:

- Das Paket `sf` - bietet Zugang zu **simple features**.
- Mit dem Paket `leaflet` kann man interaktive Karten erstellen.
- Das Paket `tmap` kann zur Erstellung von thematischen Karten genutzt werden.
- **Paket `maptools` um Karten zu erzeugen**



# Pakete aus verschiedenen Quellen installieren

## Pakete vom CRAN Server installieren

```
install.packages("lme4")
```

## Pakete vom Bioconductor Server installieren

```
source("https://bioconductor.org/biocLite.R")  
biocLite(c("GenomicFeatures", "AnnotationDbi"))
```

## Pakete von Github installieren

```
install.packages("devtools")  
library(devtools)  
  
install_github("hadley/maptools")
```

# Wie bekomme ich einen Überblick?

- Entdecke Pakete, die kürzlich auf den **CRAN** Server hochgeladen wurden
- Nutze eine Shiny Web-App, die **Pakete anzeigt, die kürzlich von CRAN** heruntergeladen wurden.
- Werfe einen Blick auf eine **Quick-Liste nützlicher Pakete**
- . . . . , oder auf eine Liste mit den **besten Paketen für die Datenverarbeitung und -analyse**, . . . . .
- . . . . , oder schaue unter **die 50 meistgenutzten Pakete**

# CRAN Task Views

- Bezüglich mancher Themen gibt es einen Überblick über alle wichtigen Pakete - (**CRAN Task Views**)
- Momentan gibt es 35 Task Views.
- Alle Pakete einer Task-View können mit folgendem Befehl installiert werden: **command**:

```
install.packages("ctv")  
library("ctv")  
install.views("Spatial")
```

## CRAN Task Views

<a href="#">Bayesian</a>	Bayesian Inference
<a href="#">ChemPhys</a>	Chemometrics and Computational Physics
<a href="#">ClinicalTrials</a>	Clinical Trial Design, Monitoring, and Analysis
<a href="#">Cluster</a>	Cluster Analysis & Finite Mixture Models
<a href="#">DifferentialEquations</a>	Differential Equations
<a href="#">Distributions</a>	Probability Distributions
<a href="#">Econometrics</a>	Econometrics
<a href="#">Environmetrics</a>	Analysis of Ecological and Environmental Data

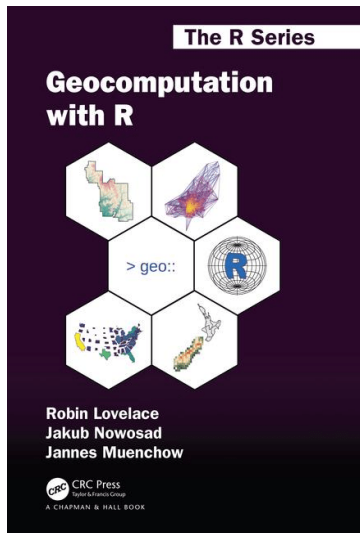
# A1C Übung - zusätzliche Pakete

Geh bspw. auf <https://cran.r-project.org/> und suche nach Paketen. . .

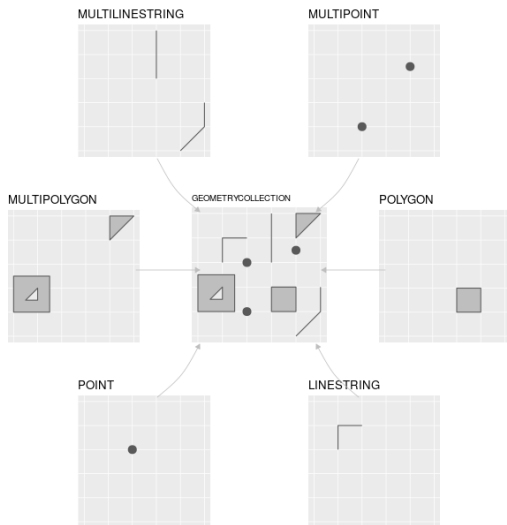
- die sich für interaktive Karten eignen.
- mit denen man thematische Karten erstellen kann
- mit denen man die räumliche Distanz berechnen kann
- mit denen man eine Satellitenkarte bekommen kann



# Robin Lovelace - Geocomputation with R



# Typen von simple feature die von sf unterstützt werden.



# Erste Weltkarten

```
library(sf)
```

```
## Linking to GEOS 3.6.1, GDAL 2.2.3, PROJ 4.9.3
```

```
library(spData)
```

```
## To access larger datasets in this package, install the spData  
## package with: `install.packages('spDataLarge',  
## repos='https://nowosad.github.io/drat/', type='source')`
```

```
plot(world)
```

```
## Warning: plotting the first 9 out of 10 attributes; use max  
## plot all
```

iso\_a2

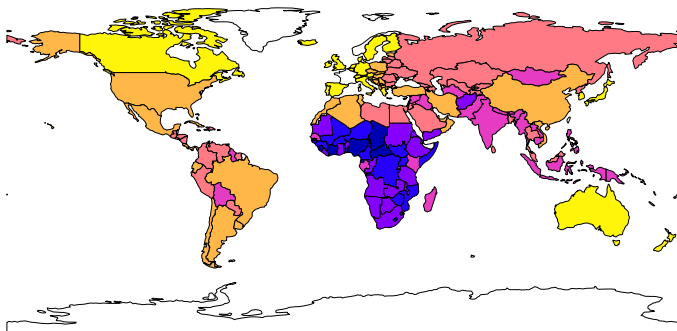
name\_long

continent

# Nur einen Indikator plotten

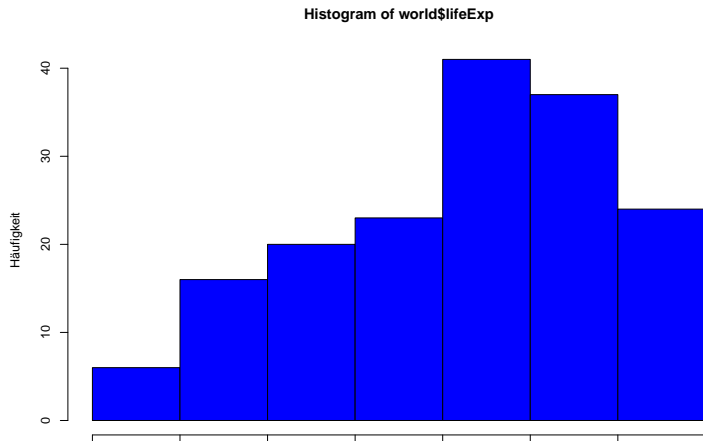
```
plot(world["lifeExp"],main="Lebenserwartung in Jahren")
```

Lebenserwartung in Jahren



# Eine andere Darstellungsform

```
hist(world$lifeExp,col="blue",xlab="Lebenserwartung",  
      ylab="Häufigkeit")
```



# Rasterdaten

**A. Cell IDs**

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

**B. Cell values**

22	74	28	91
72	84	NA	85
NA	92	24	53
31	62	56	5

**C. Colored cell values**

