



Fakultät für
**Kultur- und
Sozialwissen-
schaften**

ZLI Tech Talk: R/RStudio

8. Mai, 12 Uhr

Fakultät für **Kultur- und Sozialwissenschaften**
Arbeitsstelle Quantitative Methoden
Markus Tausendpfund

Fahrplan

1. Ausgangslage: Von SPSS/Stata zu R/RStudio
 2. Vorstellung R/RStudio
 3. Herausforderungen
 4. ChatGPT und R/RStudio
 5. Fragen, Diskussion und Ausblick
- Ressourcen zu R

Ausgangslage: Von SPSS/Stata zu R/RStudio

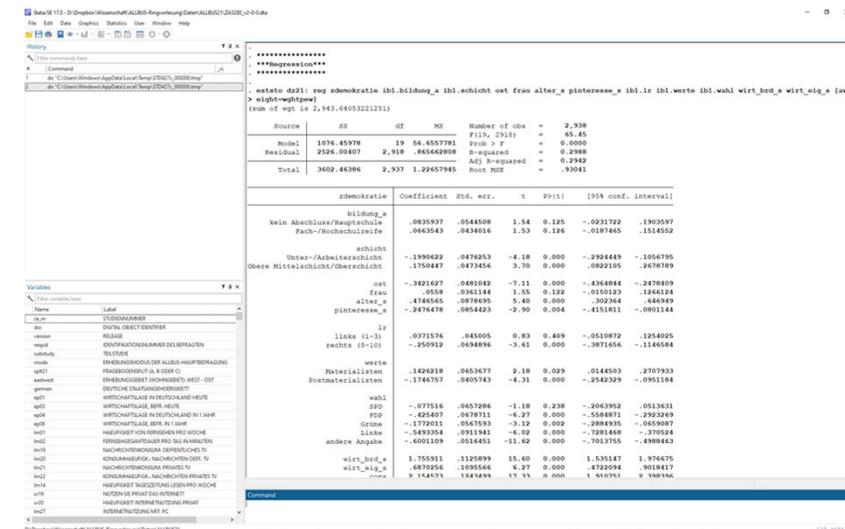
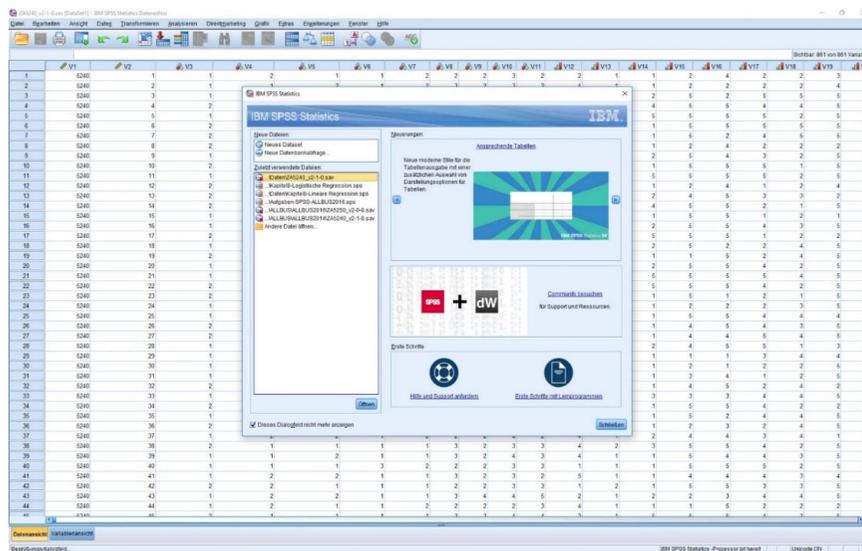
- Klassische Softwarepakete (Fokus: Sozialwissenschaften)



1968



1985



Ausgangslage: Von SPSS/Stata zu R/RStudio

- Klassische Perspektive
 - Befragung als ein „klassisches Instrument der Datenerhebung“ (Reinecke 2022, S. 949)
 - Befragung als „Königsweg der praktischen Sozialforschung“ (König 1974, S. 27)

- Standardisierte Befragungen als Grundlage für die Mehrheit empirischer Studien
 - Nationale Datenerhebungen: ALLBUS, GLES, SOEP, Politbarometer....
 - Europäische Datenerhebungen: ESS, Eurobarometer, EVS...
 - Weltweite Datenerhebungen: WVS

Ausgangslage: Von SPSS/Stata zu R/RStudio

- Struktur von Befragungsdaten
 - rechteckige Datenmatrix (strukturierte Daten)

Variablen

Untersuchungsobjekte

za_nr	doi	version	respid	eastwest	german	ep01	ep03	ep04	ep06	lm01	lm02	lm19	lm20	lm21
1	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	1	1	1	1	2	2	3	6.0	75	1	6.0	
2	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	2	2	1	2	4	3	3	7.0	240	1	7.0	
3	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	3	1	1	2	2	3	2	2.0	120	1	1.0	
4	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	4	2	1	2	2	3	3	7.0	180	1	2.0	
5	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	5	2	1	3	2	3	2	7.0	60	1	7.0	
6	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	6	1	1	1	3	3	2	3.0	120	1	3.0	
7	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	7	1	1	3	2	3	3	7.0	120	1	3.0	
8	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	8	1	1	2	3	3	2	2.0	60	2	-10.0	
9	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	9	1	1	1	2	4	3	3.0	60	1	2.0	
10	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	10	2	1	2	2	2	3	7.0	240	1	7.0	
11	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	11	2	1	3	3	3	3	7.0	180	2	-10.0	
12	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	12	1	1	3	3	4	3	3.0	210	2	-10.0	
13	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	13	1	1	1	2	3	3	7.0	120	2	-10.0	
14	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	14	2	1	2	2	3	3	7.0	300	1	7.0	
15	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	15	1	1	3	3	4	3	6.0	75	1	5.0	
16	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	16	1	1	1	4	3	2	5.0	60	1	5.0	
17	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	17	1	3	2	2	2	1	5.0	60	1	3.0	
18	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	18	1	1	3	2	1	3	2.0	60	1	2.0	
19	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	19	1	1	3	4	4	3	1.0	90	1	1.0	
20	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	20	1	1	3	3	3	3	7.0	90	1	7.0	
21	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	21	1	1	2	2	3	4	7.0	180	1	7.0	
22	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	22	2	1	3	2	-8	3	7.0	120	1	7.0	
23	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	23	1	1	4	1	3	3	1.0	120	2	-10.0	
24	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	24	1	3	1	3	4	2	1.0	90	2	-10.0	
25	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	25	1	1	2	2	3	2	7.0	60	1	7.0	
26	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	26	1	1	2	2	4	3	7.0	150	1	7.0	
27	5270 doi:10.4232/1.13250	2.0.0 (2019-03-26)	27	2	1	2	2	3	3	0	-10	-10	-10.0	-1

Ausgangslage: Von SPSS/Stata zu R/RStudio

- Verbreitung der Softwarepakete
 - Befragung von Simon Munzert 2015; politikwissenschaftlich ausgerichtete Fakultäten

Tabelle 1: Einsatz von Datenanalysesoftware in Forschung und Lehre. Angaben in Prozent der befragten Institute.

Software	Einsatz in Lehre	Einsatz in Forschung
SPSS	64	82
Stata	64	80
Excel	44	56
R	42	69
MAXQDA	40	62
ATLAS.ti	13	29
MPlus	4	7
Tosmana	4	4
Python	2	16
QDA Miner		13
NVivo		11
SAS		9
Eviews		2



Quelle: Munzert, Simon. 2018. Auf dem Weg zu einer fundierten Softwareausbildung in der Politikwissenschaft. In *Computational Social Science*, Hrsg. Andreas Blätke, Joachim Behnke, Kai-Uwe Schnapp und Claudius Wagemann, 379-402. Baden-Baden: Nomos.

Ausgangslage: Von SPSS/Stata zu R/RStudio

- „Digitalisierung“ der Gesellschaft
 - Neue Datenbestände
 - Digitale Verhaltensdaten: (Twitter/X), Facebook, YouTube, Reddit, ...
 - Big Data: „große“ Datenmengen
 - Neue Analysestrategien
 - unstrukturierte Datenbestände
 - neue Analyseverfahren
 - Neue Studiengänge entstehen
 - Computational Social Science (interdisziplinäres Wissenschaftsfeld)
 - Data Science

Ausgangslage: Von SPSS/Stata zu R/RStudio



Neue Forschungsmöglichkeiten für die Sozialwissenschaften

Was sind digitale Verhaltensdaten?

Der Begriff „digitale Verhaltensdaten“ (DVD) umfasst digitale Beobachtungen menschlichen und algorithmischen Verhaltens, wie sie z.B. von Online-Plattformen (wie Google, Facebook oder dem World Wide Web) oder Sensoren (wie Smartphones, RFID-Sensoren, Satelliten oder Street View-Kameras) erfasst werden. Es gibt Schnittmengen mit den Konzepten von „Big Data“ – insofern digitale Verhaltensdaten häufig deren Eigenschaften teilen – oder „digitale Spurdaten“ – insofern diese die Grundlage für DVD bilden. „Digitale Verhaltensdaten“ fokussiert jedoch immer auf die gesellschaftlich bedeutsamen Aspekte menschlichen und algorithmischen Verhaltens und die damit verbundenen Forschungsperspektiven.

Warum untersuchen wir digitale Verhaltensdaten?

Die Digitalisierung hat unser Leben in fast allen Bereichen verändert: sei es unsere Kommunikation, unser Arbeiten, Lernen und Lehren, Einkäufe, Reisen, die Suche nach einem Job, einem Partner oder Informationen zu jedem beliebigen Thema. Wir alle interagieren in und mit sozio-technischen Systemen, in denen Algorithmen eine entscheidende Rolle spielen. Damit sind digitale Verhaltensdaten zu einer wichtigen Quelle für die Beobachtung und Analyse der strukturellen Entwicklung von Gesellschaften und des gesellschaftlichen Verhaltens von Menschen geworden. Mit DVD können wir klassische sozialwissenschaftliche Forschungsfragen mit neuen Datentypen angehen, wir bekommen Daten für Bereiche, für die es keine oder nur wenige quantitative Daten aus herkömmlichen Quellen gibt, und vor allem können wir die grundlegenden strukturellen Veränderungen des privaten und des öffentlichen Lebens, die durch die Digitalisierung der Gesellschaft erzeugt werden, *in vivo* erforschen.

Warum engagiert sich GESIS in diesem Bereich?

Die Datenstandards in den Sozialwissenschaften sind hoch, und die Besonderheiten digitaler Verhaltensdaten stellen einzelne Wissenschaftler*innen, die sie in ihre Forschungsdesigns integrieren möchten, vor große Herausforderungen. Um den Einsatz von DVD in der sozialwissenschaftlichen Forschungsgemeinschaft und darüber hinaus zu ermöglichen und weiter voranzutreiben, sind zuverlässige Infrastrukturen für das Sammeln, Vorverarbeiten, Analysieren und Bereitstellen digitaler Verhaltensdaten erforderlich. GESIS ist die größte Infrastruktureinrichtung für Sozialwissenschaften in Europa, Daten sind unsere Kernkompetenz und GESIS verfügt – seit der Gründung der Abteilung Computational Social Science im Jahr 2013 – über langjährige Erfahrung mit digitalen Verhaltensdaten. Seit 2022 können wir dieses Engagement – dank zusätzlicher Mittel der Leibniz-Gemeinschaft („Sonderatbestand“) – abteilungsübergreifend zu einer tragenden Säule unseres Instituts ausbauen.

<https://www.gesis.org/institut/digitale-verhaltensdaten>

Ausgangslage: Von SPSS/Stata zu R/RStudio

- Keine Lizenz erforderlich
 - SPSS und Stata kommerzielle Anbieter (proprietär), R freie Software
- Neue Datenbestände und neue Analyseverfahren
 - Grenzen von SPSS und/oder Stata
- Methoden-agnostisch
 - Anwendungen für quantitative und qualitative Analyseverfahren (z.B. QCA)
- Zukunftssicher
 - ständige Weiterentwicklung, aber auch Alternativen (z.B. Python, Julia)

Ausgangslage: Von SPSS/Stata zu R/RStudio



Literatur

- Munzert, Simon. 2018. Auf dem Weg zu einer fundierten Softwareausbildung in der Politikwissenschaft. In *Computational Social Science*, Hrsg. Andreas Blätte, Joachim Behnke, Kai-Uwe Schnapp und Claudius Wagemann, 379-402. Baden-Baden: Nomos.
- Munzert, Simon. 2014. Big Data in der Forschung! Big Data in der Lehre? Ein Vorschlag zur Erweiterung der bestehenden Methodenausbildung. *Zeitschrift für Politikwissenschaft* 24 (1-2): 205-220.

Fahrplan

1. Ausgangslage: Von SPSS/Stata zu R/RStudio
2. **Vorstellung R/RStudio**
3. Herausforderungen
4. ChatGPT und R/RStudio
5. Fragen, Diskussion und Ausblick

Was ist R?

- freie Programmiersprache für (statistische) Datenanalyse und Grafiken
 - <https://www.r-project.org>
- in den 1990er Jahren von Ross Ihaka und Robert Gentleman entwickelt
- unter der General Public Licence (GNU) veröffentlicht
- für verschiedene Betriebssysteme frei verfügbar
- Weiterentwicklung durch R Core Team
- kein „geschlossenes“ Programm, sondern das Fundament vieler „Programme“ (Pakete)

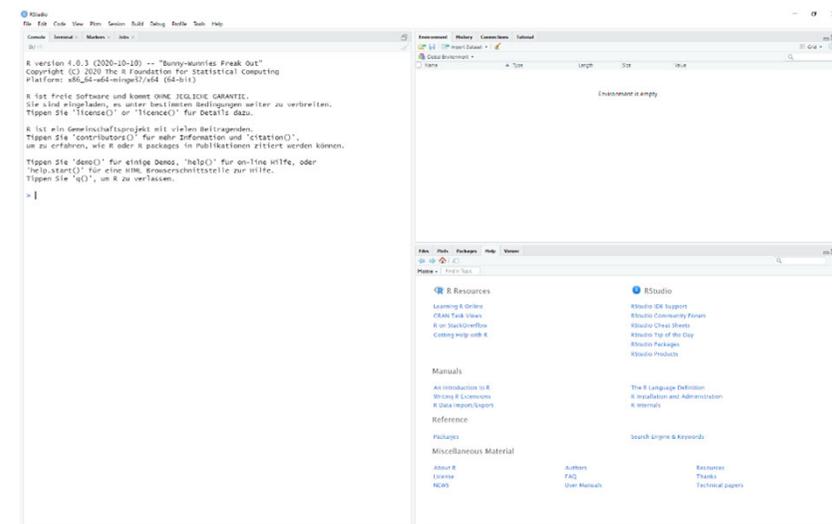
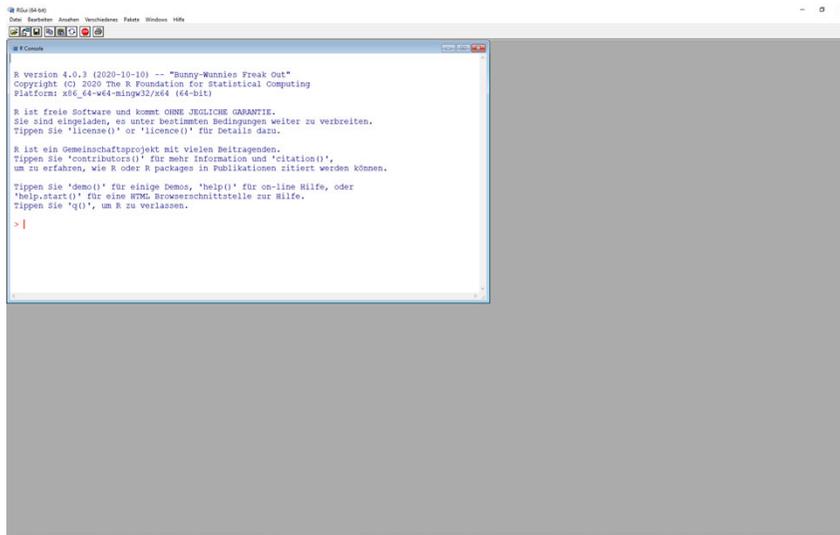
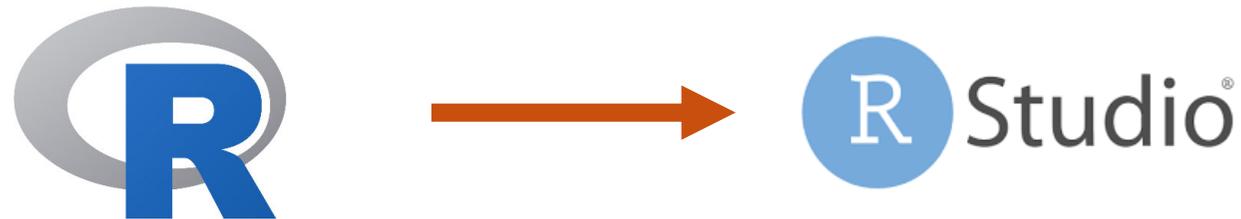


Was ist RStudio?

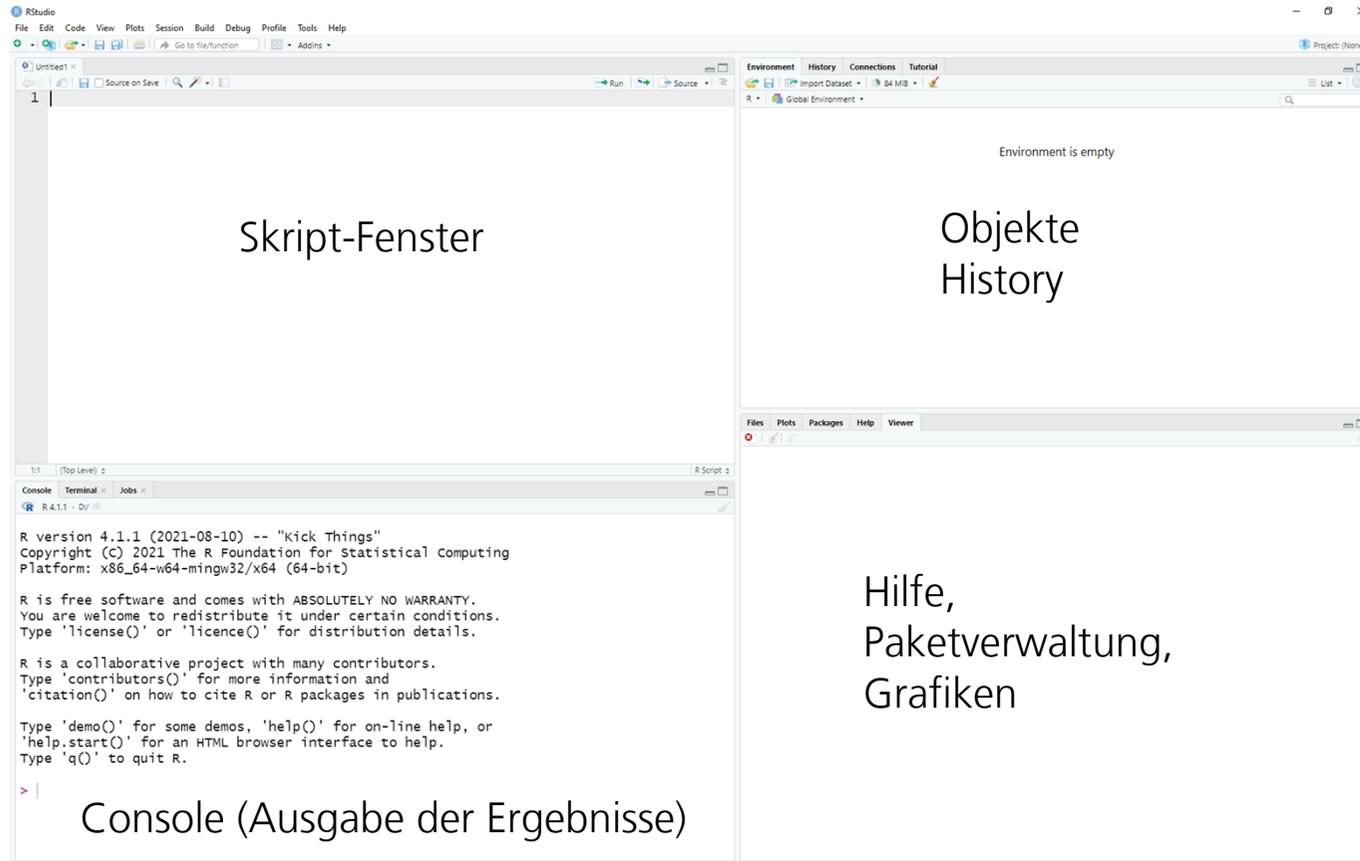
- grafische Benutzeroberfläche für R (entwickelt von der Firma posit, früher RStudio)
 - <https://posit.co/downloads/>
- ebenfalls für verschiedene Betriebssysteme frei verfügbar
- erleichtert die Arbeit mit R (Skripteditor, Paketverwaltung, Hilfebereich)
- Alternativen zu RStudio
 - RCommander



R und RStudio



RStudio



The screenshot shows the RStudio interface with the following components labeled:

- Skript-Fenster**: The main editor window for writing R scripts.
- Objekte History**: The Environment pane showing the current state of the R environment (empty).
- Hilfe, Paketverwaltung, Grafiken**: The Files, Plots, Packages, Help, and Viewer panes.
- Console (Ausgabe der Ergebnisse)**: The Console pane showing the R version information and the prompt.

```
R version 4.1.1 (2021-08-10) -- "Kick Things"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

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.

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.

> |
```

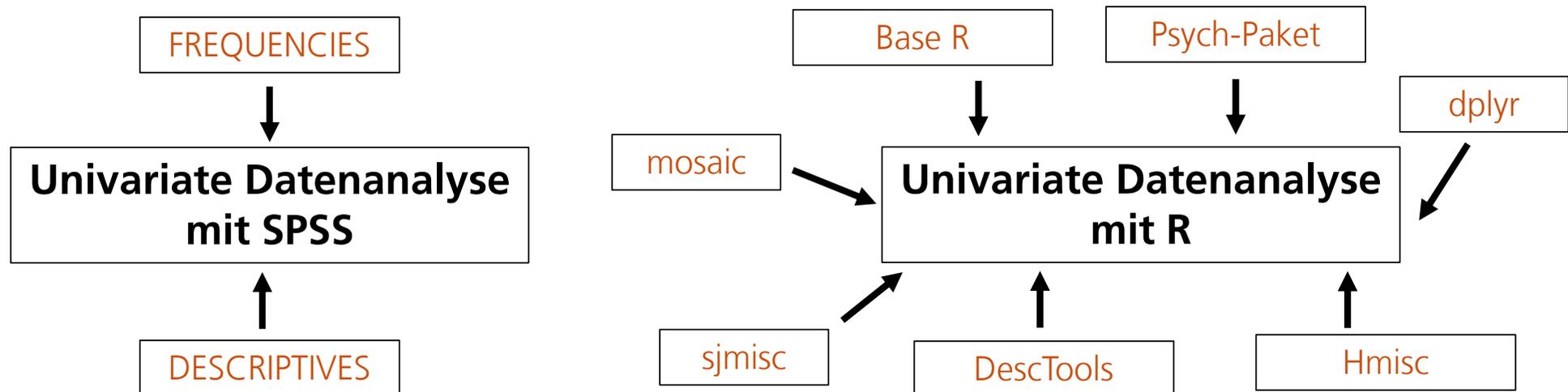
Herausforderungen bei der Arbeit mit R und RStudio

- Updates
 - R (seit 24. April ist Version 4.4 verfügbar)
 - RStudio
 - Packages

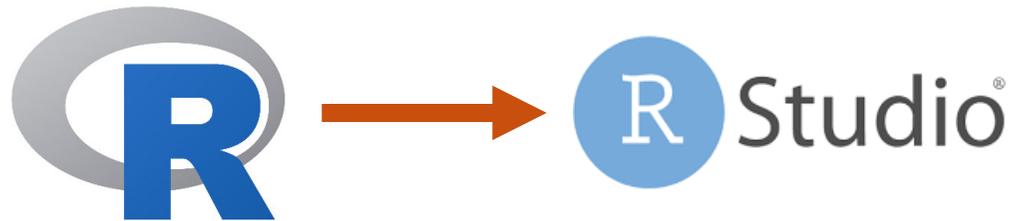
- Packages
 - erweitern den Funktionsumfang
 - sehr viele (April 2024: > 20.000 allein auf CRAN; Überblick via TaskViews)
 - nicht (unbedingt) aufeinander abgestimmt (hmisc / psych)
 - viele Pakete für die gleiche Aufgabe verfügbar

Pakete

- Viele Wege führen zum (gleichen) Ergebnis



R, RStudio und Pakete („Packages“)



„Welt der Pakete“

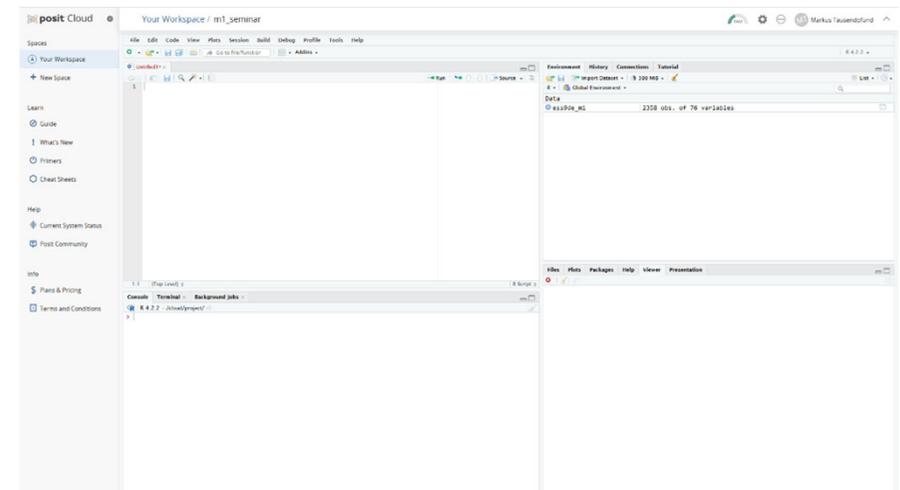


Fahrplan

1. Ausgangslage: Von SPSS/Stata zu R/RStudio
2. Vorstellung R/RStudio
3. **Herausforderungen**
4. ChatGPT und R/RStudio
5. Fragen, Diskussion und Ausblick

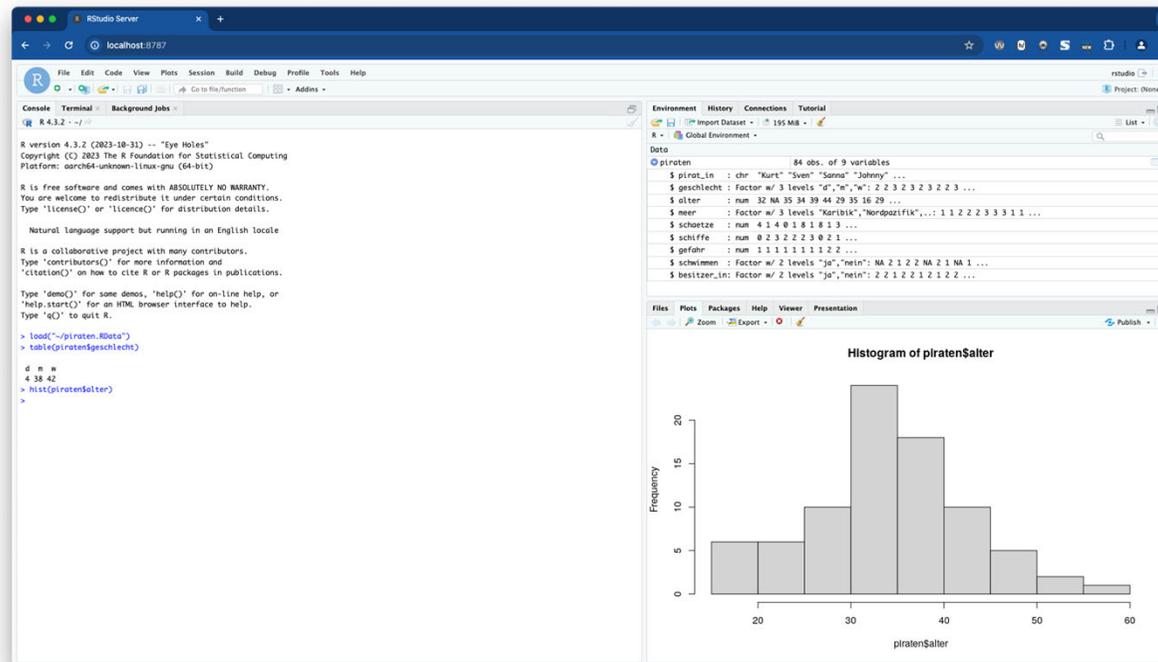
Studierende können R/RStudio nicht installieren

- Warum?
 - Studierende haben keine Admin-Rechte zur Installation von R/RStudio
 - Betriebssystem unterstützt keine aktuelle R-Version (≥ 4.0)
 - Computer zu leistungsschwach
 -
- Mögliche Lösung: Posit Cloud
 - <https://posit.cloud>
 - Cloud Free (begrenzte Nutzung)

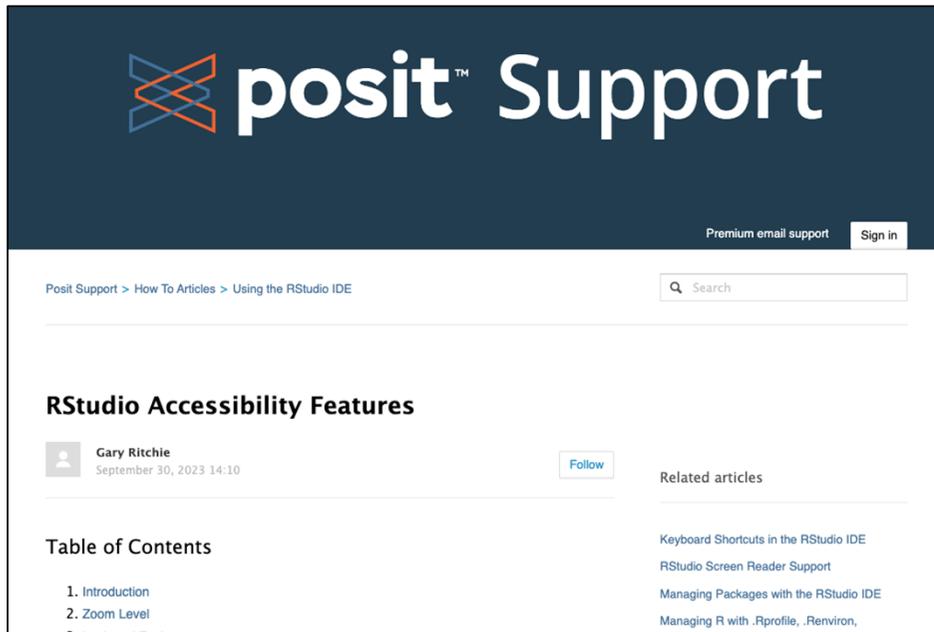


Studierende ohne Internet-Zugang

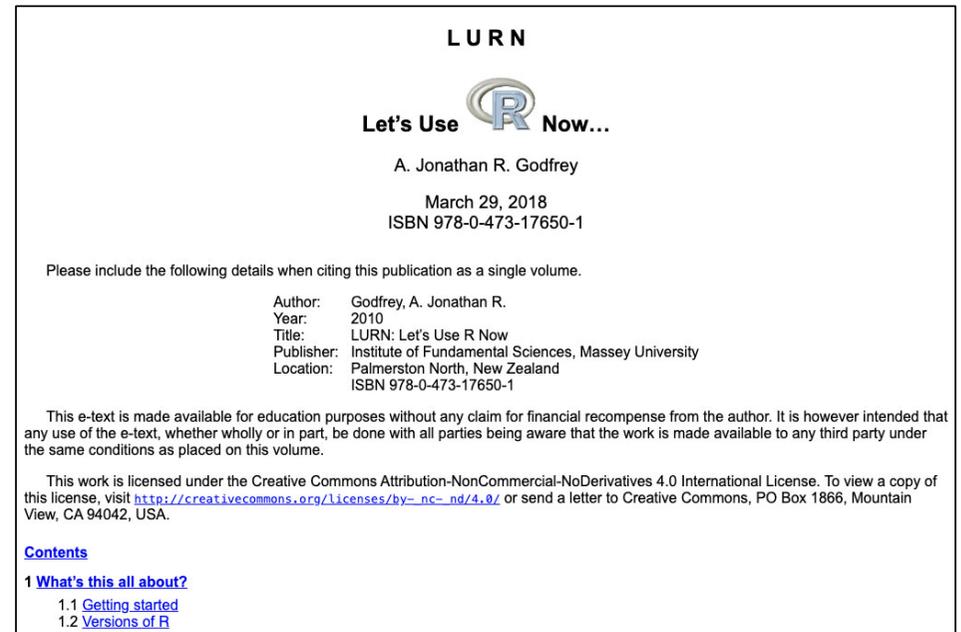
- RStudio als Docker-Container



Informationen für Studierende mit Behinderung



The screenshot shows the Posit Support website. At the top, there is a dark blue header with the Posit logo and the word "Support" in white. Below the header, there is a navigation bar with "Premium email support" and a "Sign in" button. The main content area has a breadcrumb trail: "Posit Support > How To Articles > Using the RStudio IDE". A search bar is visible. The article title is "RStudio Accessibility Features" by Gary Ritchie, dated September 30, 2023. A "Follow" button is present. Below the title, there is a "Table of Contents" section with the following items: 1. Introduction, 2. Zoom Level, and 3. Keyboard Features. To the right of the article, there is a "Related articles" section with the following items: Keyboard Shortcuts in the RStudio IDE, RStudio Screen Reader Support, Managing Packages with the RStudio IDE, and Managing R with .Rprofile, .Renviron, and .Rprofile.



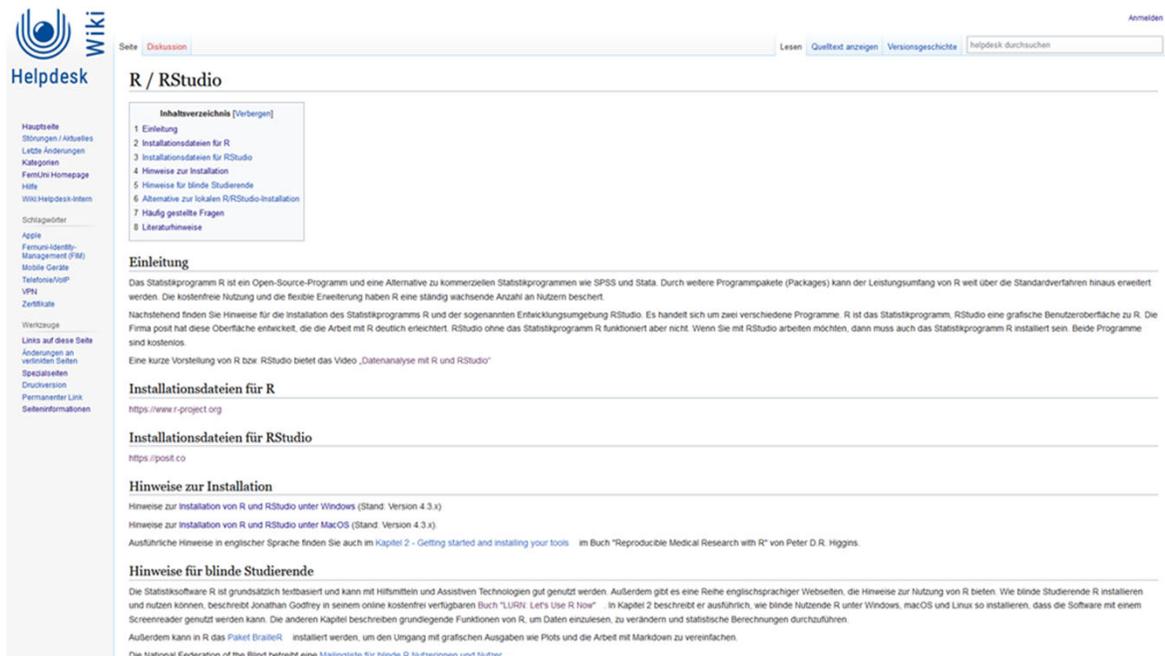
The screenshot shows the LURN publication page. At the top, there is the LURN logo and the title "Let's Use R Now..." with the R logo. Below the title, the author is listed as "A. Jonathan R. Godfrey", the date as "March 29, 2018", and the ISBN as "ISBN 978-0-473-17650-1". A section titled "Please include the following details when citing this publication as a single volume." lists the following information: Author: Godfrey, A. Jonathan R., Year: 2010, Title: LURN: Let's Use R Now, Publisher: Institute of Fundamental Sciences, Massey University, Location: Palmerston North, New Zealand, ISBN 978-0-473-17650-1. Below this, there is a paragraph stating: "This e-text is made available for education purposes without any claim for financial recompense from the author. It is however intended that any use of the e-text, whether wholly or in part, be done with all parties being aware that the work is made available to any third party under the same conditions as placed on this volume." Another paragraph states: "This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA." Below this, there is a "Contents" section with the following items: 1 What's this all about?, 1.1 Getting started, and 1.2 Versions of R.

Aller Anfang ist schwer...

- Nutzung von R/RStudio in (mindestens) 4 von 5 Fakultäten...
- ... aber kein zentraler Basis-Support (z.B. typische Probleme bei der Installation)
- ... durchaus überraschend, da SPSS-Installationsanleitung im Wiki des helpdesk
 - https://wiki.fernuni-hagen.de/helpdesk/index.php?title=SPSS_28_Installation

R/RStudio im wiki des helpdesk

- Erste Lösungsschritte (Danke an Björn Fisseler)



The screenshot shows a wiki page titled "R / RStudio" on the "Helpdesk" website. The page includes a navigation menu on the left, a search bar at the top right, and a table of contents. The main content area contains an introduction, installation instructions for Windows and MacOS, and a section for blind students.

Inhaltsverzeichnis [Verbergen]

- 1 Einleitung
- 2 Installationsdateien für R
- 3 Installationsdateien für RStudio
- 4 Hinweise zur Installation
- 5 Hinweise für blinde Studierende
- 6 Alternative zur lokalen R/RStudio-Installation
- 7 Häufig gestellte Fragen
- 8 Literaturhinweise

Einleitung

Das Statistikprogramm R ist ein Open-Source-Programm und eine Alternative zu kommerziellen Statistikprogrammen wie SPSS und Stata. Durch weitere Programmpakete (Packages) kann der Leistungsumfang von R weit über die Standardverfahren hinaus erweitert werden. Die kostenfreie Nutzung und die flexible Erweiterung haben R eine ständig wachsende Anzahl an Nutzern beschert.

Nachstehend finden Sie Hinweise für die Installation des Statistikprogramms R und der sogenannten Entwicklungsumgebung RStudio. Es handelt sich um zwei verschiedene Programme. R ist das Statistikprogramm, RStudio eine grafische Benutzeroberfläche zu R. Die Firma posit hat diese Oberfläche entwickelt, die die Arbeit mit R deutlich erleichtert. RStudio ohne das Statistikprogramm R funktioniert aber nicht. Wenn Sie mit RStudio arbeiten möchten, dann muss auch das Statistikprogramm R installiert sein. Beide Programme sind kostenlos.

Eine kurze Vorstellung von R bzw. RStudio bietet das Video „Datenanalyse mit R und RStudio“

Installationsdateien für R

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

Installationsdateien für RStudio

<https://posit.co>

Hinweise zur Installation

Hinweise zur Installation von R und RStudio unter Windows (Stand: Version 4.3.x)

Hinweise zur Installation von R und RStudio unter MacOS (Stand: Version 4.3.x)

Ausführliche Hinweise in englischer Sprache finden Sie auch im Kapitel 2 - Getting started and installing your tools im Buch "Reproducible Medical Research with R" von Peter D. R. Higgins.

Hinweise für blinde Studierende

Die Statistiksoftware R ist grundsätzlich textbasiert und kann mit Hilfsmitteln und Assistiven Technologien gut genutzt werden. Außerdem gibt es eine Reihe englischsprachiger Webseiten, die Hinweise zur Nutzung von R bieten. Wie blinde Studierende R installieren und nutzen können, beschreibt Jonathan Godfrey in seinem online kostenfrei verfügbaren Buch "LURR: Let's Use R Now". In Kapitel 2 beschreibt er ausführlich, wie blinde Nutzende R unter Windows, MacOS und Linux so installieren, dass die Software mit einem Screenreader genutzt werden kann. Die anderen Kapitel beschreiben grundlegende Funktionen von R, um Daten einzulesen, zu verändern und statistische Berechnungen durchzuführen.

Außerdem kann in R das Paket **BrailleR** installiert werden, um den Umgang mit grafischen Ausgaben wie Plots und die Arbeit mit Markdown zu vereinfachen.

Die National Federation of the Blind betreibt eine [Mailinliste für blinde R-Nutzerinnen und Nutzer](#).



<https://e.feu.de/helpdesk-r>

<https://wiki.fernuni-hagen.de/helpdesk/index.php?title=R / RStudio>

Fahrplan

1. Ausgangslage: Von SPSS/Stata zu R/RStudio
2. Vorstellung R / RStudio
3. Herausforderungen
4. **ChatGPT und R/RStudio**
5. Fragen, Diskussion und Ausblick

ChatGPT und R/RStudio

 **Sie**
Was ist R und RStudio?

 **ChatGPT**
R ist eine Programmiersprache und Umgebung für statistische Berechnungen und Grafiken. Sie wird häufig in Bereichen wie Datenanalyse, Statistik, maschinelles Lernen und Datenvisualisierung verwendet. R ist eine Open-Source-Sprache und bietet eine breite Palette von Bibliotheken und Paketen, die von der Gemeinschaft entwickelt wurden, um verschiedene Aufgaben zu erleichtern.

RStudio ist eine integrierte Entwicklungsumgebung (IDE) für R, die entwickelt wurde, um die Entwicklung von R-Anwendungen zu erleichtern. Es bietet Funktionen wie Syntax-Hervorhebung, Code-Vervollständigung, Grafikfenster, Paket-Manager und viele andere Werkzeuge, die den Entwicklungsprozess erleichtern. RStudio ist ebenfalls Open Source und hat dazu beigetragen, die Verbreitung von R als Programmiersprache zu fördern, indem es eine benutzerfreundliche Oberfläche für die Entwicklung und Ausführung von R-Code bereitstellt.



ChatGPT und R/RStudio

TA Sie

Du hast die Altersangaben von 10 Befragten. Wie kann ich in R den Mittelwert berechnen?

ChatGPT

Um den Mittelwert der Altersangaben von 10 Befragten in R zu berechnen, kannst du die Funktion `mean()` verwenden. Hier ist ein Beispielcode:

```
R Copy code
# Altersangaben der Befragten
alter <- c(25, 30, 35, 40, 45, 50, 55, 60, 65, 70)

# Mittelwert berechnen
mittelwert <- mean(alter)

# Mittelwert ausgeben
print(mittelwert)
```

In diesem Beispiel wird zunächst ein Vektor `alter` mit den Altersangaben der 10 Befragten erstellt. Dann wird die Funktion `mean()` verwendet, um den Mittelwert dieser Altersangaben zu berechnen. Das Ergebnis wird in der Variable `mittelwert` gespeichert und mit `print()` ausgegeben.

ChatGPT und R/RStudio



Sie

Warum funktioniert das folgende R-Skript nicht?

```
alter <- c(25, 30, 35, 40, NA)
mean(alter)
```



ChatGPT

Das R-Skript funktioniert nicht, weil der Vektor `alter` einen fehlenden Wert (`NA`) enthält. Die Funktion `mean()` kann standardmäßig nicht mit fehlenden Werten umgehen und gibt daher auch ein `NA` als Ergebnis zurück.

Um den Mittelwert zu berechnen und dabei fehlende Werte zu ignorieren, kannst du den Parameter `na.rm = TRUE` verwenden. Hier ist der korrigierte Code:

```
R Copy code
alter <- c(25, 30, 35, 40, NA)
mittelwert <- mean(alter, na.rm = TRUE)
print(mittelwert)
```

Durch Hinzufügen von `na.rm = TRUE` wird die Funktion `mean()` angewiesen, fehlende Werte zu ignorieren und den Mittelwert nur aus den vorhandenen Werten zu berechnen.



ChatGPT und R/RStudio

Sie
Was ist der Pipe-Operator? Warum ist der Pipe-Operator nützlich?

ChatGPT
Der Pipe-Operator (`%>%`) ist ein Operator in R, der aus dem Paket `magrittr` stammt und oft in Verbindung mit dem Paket `dplyr` verwendet wird. Der Pipe-Operator ermöglicht eine einfache und intuitive Verkettung von Funktionen, indem er das Ergebnis einer Funktion als Eingabe für die nächste Funktion verwendet.

Anstelle von verschachtelten Funktionen oder Zwischenspeichern von Zwischenergebnissen ermöglicht der Pipe-Operator einen klaren und lesbareren Code, indem er den Datenfluss von links nach rechts abbildet.

Hier ist ein einfaches Beispiel, das den Pipe-Operator zeigt:

```
R Copy code
# Ohne Pipe-Operator
result <- mean(sqrt(abs(c(1, 2, 3, 4, 5))))

# Mit Pipe-Operator
library(dplyr)
result <- c(1, 2, 3, 4, 5) %>%
  abs() %>%
  sqrt() %>%
  mean()
```

Wie du sehen kannst, ermöglicht der Pipe-Operator eine klare und intuitive Darstellung des Datenflusses. Dies erleichtert das Verständnis des Codes, insbesondere bei komplexeren Analysen oder Datenmanipulationen.

ab R-Version 4.1

Pipe-Alternative: `|>`

ChatGPT und R/RStudio

Sie
Wie kann ich in ggplot2 ein Streudiagramm mit eingezeichneter OLS-Gerade erstellen?

ChatGPT
Um ein Streudiagramm mit einer eingezeichneten OLS (ordinary least squares) Geraden in ggplot2 zu erstellen, kannst du den `geom_smooth()`-Layer verwenden und `method = "lm"` setzen, um die lineare Regression anzugeben. Hier ist ein Beispielcode:

```
R Copy code  
  
library(ggplot2)  
  
# Beispiel Daten  
daten <- data.frame(x = c(1, 2, 3, 4, 5),  
                    y = c(2, 3, 5, 4, 6))  
  
# Streudiagramm mit OLS-Gerade  
ggplot(daten, aes(x = x, y = y)) +  
  geom_point() + # Streudiagramm  
  geom_smooth(method = "lm", se = FALSE) # OLS-Gerade ohne Standardfehler
```

In diesem Beispiel werden die Daten in einem Streudiagramm dargestellt (`geom_point()`), und dann wird die OLS-Gerade mit `geom_smooth()` hinzugefügt, wobei `method = "lm"` für die lineare Regression steht und `se = FALSE` gesetzt wird, um die Standardfehler der Schätzung nicht anzuzeigen.

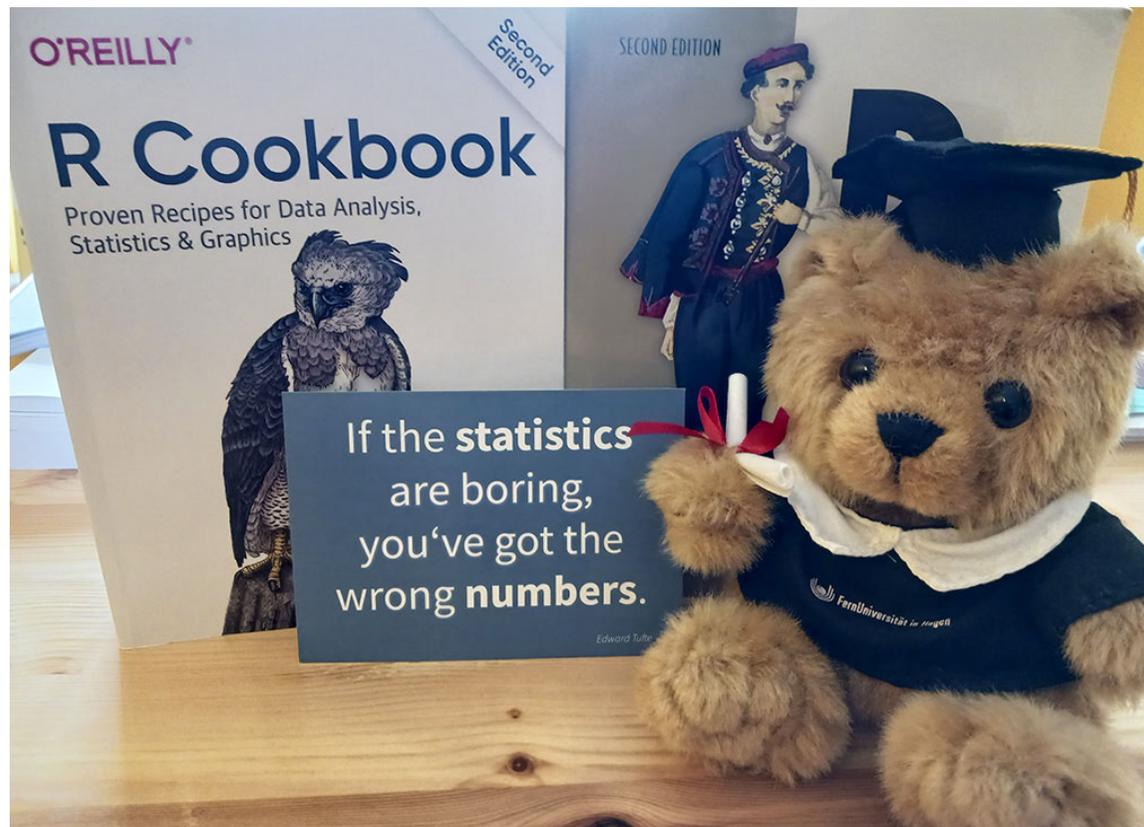
Du kannst dieses Beispiel anpassen, indem du deine eigenen Daten für `daten` verwendest und ggplot2 an deine spezifischen Anforderungen anpasst.

Fahrplan

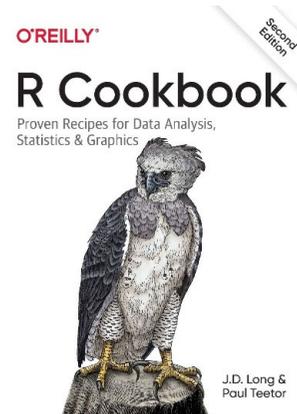
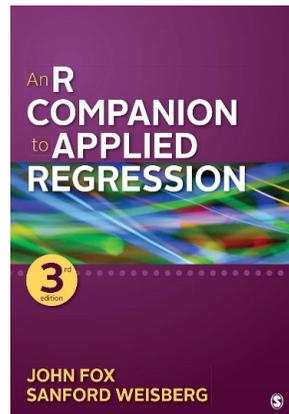
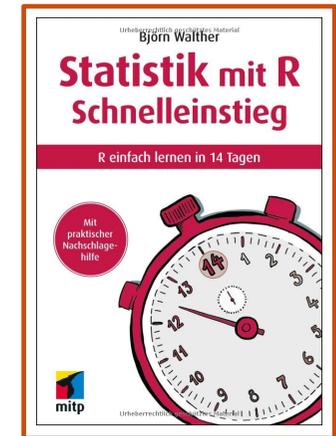
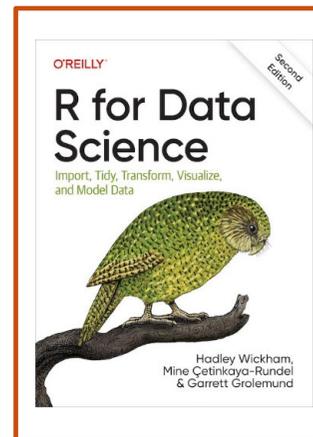
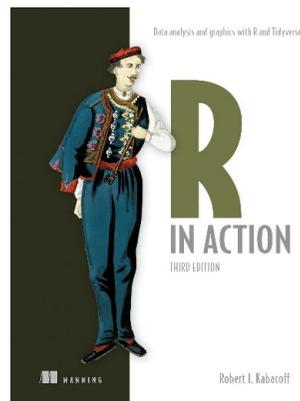
1. Ausgangslage: Von SPSS/Stata zu R/RStudio
2. Vorstellung R/RStudio
3. Herausforderungen
4. ChatGPT und R/RStudio
5. **Fragen, Diskussion und Ausblick**

Fragen, Diskussion und Ausblick

- Information: Rahmenvertrag für SPSS-Lizenzen läuft zum SoSe 24 aus
 - Studierende: keine SPSS-Lizenzen mehr über die FernUniversität
 - Mitarbeitende: ???



Literatur zu R und RStudio



Videos zu R und RStudio (nach „R“ und/oder „RStudio“ suchen)

- in Moodle / Video-Plattform (video.fernuni-hagen.de)

The screenshot shows a Moodle search results page for the category 'Quantitative Methoden'. The page displays 55 articles found across 7 pages. The results are organized into a grid of video thumbnails, each with a title, description, author (Markus Tausendpfund), view count, and date (27.7.2021). The videos cover topics such as 'Was ist R? Was ist RStudio?', 'Hilfe in R', 'Skripte in R strukturieren und kommentieren', 'Einfache Berechnungen in R', 'RStudio kennenlernen', 'Univariate Datenanalyse in R', 'R-Pakete installieren und laden', and 'Häufigkeitstabellen in R'. The page also includes navigation elements like a search bar, home icon, and pagination controls at the bottom.

Videos zu R und RStudio

- YouTube

It is the mark of a truly intelligent person to be moved by statistics.
George Bernard Shaw

yuzar Data Science
@yuzar-Data-Science · 6830 Abonnenten · 63 Videos
Welcome to my VLOG! My name is Yury Zablotzki & I love to use R for Data Science - "yuzar..."

Abonniert Mitglied werden

Übersicht Videos Playlists Community

Für mich

MULTIVARIABLE LINEAR MODEL
Multivariable Linear Regression in R: Everything You Need to Know!
2118 Aufrufe · vor 2 Wochen

FIND THE BEST MODEL
R package reviews | gmult | Find The Best Model!
10.029 Aufrufe · vor 1 Jahr

DPLYR FOR BEGINNERS
Top 10 Must Know (dplyr) Commands for Data Wrangling in R!
5471 Aufrufe · vor 1 Jahr

QUANTILE REG
Quantile Regression as The Most for Ordinary Linear Regression
12.444 Aufrufe · vor 1 Jahr

Videos Alle wiedergeben

MULTIVARIABLE LINEAR MODEL **CATEGORICAL PREDICTOR LINEAR MODEL** **LINEAR MODEL NUMERIC PREDICTOR** **UNLEASH QUANTILE REGRESSION RESULTS** **COMBINE PLOTS** **BEFORE AFTER**

Analysen in R SPSS X
Statistik am PC
bjornwalther.com

Statistik am PC
@StatistikamPC_BjornWalther · 33.100 Abonnenten · 717 Videos
Auf der Suche nach Hilfestellungen zu statistischen Auswertungen? Dann wird euch hier m...

amazon.to/4741AYx und 8 weitere Links
Abonniert Mitglied werden

Übersicht Videos Shorts Playlists Community Shop

Für mich

6 Gründe für nicht signifikante Ergebnisse
5:10

Multiple Regression
21:37

Mediation interpretieren Fünf „Typen“
5:21

Regress Welche Wert muss ich ange
5:21

HILFE! Meine Ergebnisse sind NICHT SIGNIFIKANT - 6 mögliche Gründe
7978 Aufrufe · vor 2 Jahren

Multiple lineare Regression in R [ALL IN ONE] - Voraussetzungen, Durchführung und Interpretation
25.968 Aufrufe · vor 1 Jahr

Mediation einordnen - 5 Typen der Mediation
3249 Aufrufe · vor 2 Jahren

Lineare Regression - welche Ergebnisse berichten?
68.889 Aufrufe · vor 5 Jahren

Analysieren von Daten in R Alle wiedergeben
Alle Videos zu Datenimport und analytischen Tests in der kostenlosen Analysesoftware R.

und viele mehr...

Webseiten zu R

- <https://bookdown.org/ndphillips/YaRrr>
- <https://r-intro.tadaa-data.de>
- <https://www.tidyverse.org/learn>

- <https://moderndive.netlify.app>

- <https://www.bookdown.org>
- <https://www.r-bloggers.com>
- <https://www.bigbookofr.com>

Fortbildungen zu R/RStudio

- <https://www.gesis.org/gesis-training>
- <https://www.datacamp.com>
- <https://statisticalhorizons.com>