

Einführung in die Datenanalyse mit R: Datenaufbereitung

Marco Wähler



Gebäude 37.03.03.14



marco.waehner@hhu.de



Sprechstunde nach Vereinbarung



NRW-FORSCHUNGSKOLLEG
ONLINE-PARTIZIPATION

hhu Heinrich Heine
Universität
Düsseldorf



RESEARCH
FOR THE
DIGITAL AGE

Troubleshooting

- Zu den häufigsten Fehlerquellen zählen
 - Fehlende Klammer „()“ am Anfang oder am Ende eines Befehls
 - Fehlende Anführungszeichen „“““ am Anfang oder am Ende
 - Fehlendes Komma „ , “
 - Typos (R ist case-sensitive)
 - Package nicht installiert oder aktiviert
 - Dateipfad oder Directory ist falsch
 - Variable hat die falsche Klasse
 - Das Problem sitzt i.d.R. vor dem Bildschirm
 - liegt das Problem auf der Seite des Systems, einfach neustarten
- R gibt in der Console eine Fehlermeldung (error) aus, z.B.
 - Error: unexpected string constant in "a <- c("a", "b", "c" "d"“
 - Error: object 'b' not found
 - Error in a + b : non-numeric argument to binary operator

debugging



1.
I got this.



2.
Huh. Really
thought that
was it.



3.
(...)



4.
Fine. Restarting.



5.
OH WTF.



6..
Zombie
meltdown



7.



8.
A NEW HOPE!



9.
[insert awesome
theme song]



10.
I ♥ CODING!

- Die Lösung liegt irgendwo zwischen Stufe 6 (Zombie meltdown) und Stufe 8 (new hope)
- Ressourcen und Strategien zur Lösungssuche
 - Auf Typos achten
 - Alle Klammern geschlossen?
 - Fehlermeldung lesen (und versuchen zu verstehen)
 - Fehlermeldung googeln!
 - Foren wie [Stackoverflow](#) durchsuchen

- Weitere Ressourcen

- [Getting help with R](#)
- Help-Funktion „?“
- Insbesondere für packages
 - Package Dokumentation
 - Package Vignette
- R-Suchmaschine: <https://rseek.org/>

Datenaufbereitung II (Kapitel 4)

- Datenaufbereitung mit tidyverse

„The tidyverse is a coherent system of packages for data manipulation, exploration and visualization that share a common design philosophy.” ([Rickert, 2017](#))
- Installiert automatisch eine ganze Reihe weiterer Packages
 - ggplot2 (Visualisierung)
 - dplyr
 - magrittr
- Pipe-Operator als kleiner Zauberstab
 - `df %>%` Funktion
 - Shortcut strg + shift + m (windows) oder cmd + shift + m (Mac)



Praxis

Hands-On

Inglehart-Index (Lösung)

Vielen Dank!

Marco Wähner



Gebäude 37.03.03.14



marco.waehner@hhu.de



Sprechstunde nach Vereinbarung