

# rstatsZH - Datenwissenschaften mit R

## Angebotsüberblick

Lars Schöbitz

2021-06-23

# Hallo Zusammen 🙌

- Lars Schöbitz
  - Umweltingenieur
  - Toiletten-Forscher
  - Datenwissenschaftler
- R Nutzer seit 2016
- R Trainer seit 2020
- Twitter: <https://twitter.com/larnsce>



# rstatsZH - moderner NutzeR

Ein **moderner NutzeR** zu werden bedeutet, Fähigkeiten, Gewohnheiten und Denkweisen für eine reproduzierbare und kooperative Arbeitsweise zu entwickeln.

Wir lernen:

1. Coding mit bewährten Verfahren (R/RStudio/tidyverse)
2. Kooperatives Arbeiten (Git/GitHub)
3. Reproduzierbare Dokumente schreiben und veröffentlichen (R Markdown)

# Bewährte Verfahren

## R - der Motor



## RStudio IDE - das Dashboard



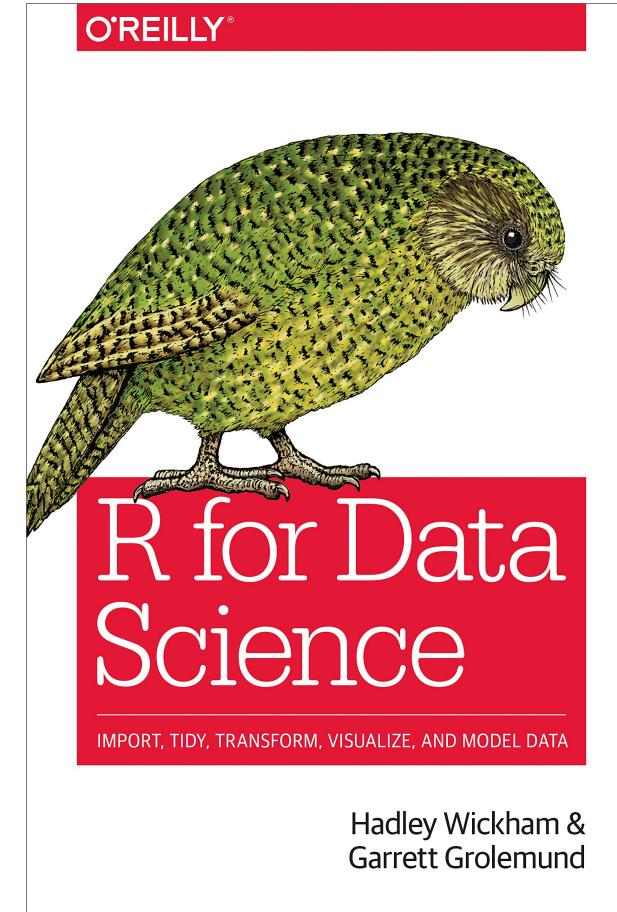
# Bewährte Verfahren - Tidyverse

- Ein logisch zusammenhängendes System zur
  - Daten Aufbereitung
  - Daten Exploration
  - Daten Visualisierung
  - Daten Modellierung
- Eine Designphilosophie, Grammatik und Daten Struktur
- Entwickelt um Produktivität zu unterstützen
- Haupt Entwickler: Hadley Wickham



# Bewährte Verfahren - R for Data Science

- Das Buch für den Kurs
- Kostenfrei
- Geschrieben von Hadley Wickham
- Nutzt die Philosophie des Tidyverse



# Kollaboratives Arbeiten - Git/GitHub/R Markdown



# Reproduzierbarkeit - Ein Beispiel



[← Zurück zur Übersicht](#)



## **Corona verändert Alltagssorgen nur wenig**

Mitteilung 06.08.2020

**Die «Dargebotene Hand» bietet als Sorgentelefon Unterstützung für die Bevölkerung. Während seit jeher die Mehrheit der Anrufenden Frauen sind, melden sich im Corona-Jahr 2020 mehr Männer als im Vorjahr. Kaum etwas geändert hat sich bei den Beratungsinhalten, wobei das Thema «Suchtverhalten» die grösste Veränderung aufweist.**

# Reproduzierbarkeit - Ein Beispiel



**Kanton Zürich**

## Datenanalyse Dargebotene Hand

---

Datenquelle: [Dargebotene Hand](#)

Mitteilung: [Corona verändert Alltagssorgen nur wenig](#)

Code: [R Markdown](#)

# Reproduzierbarkeit - Ein Beispiel

## Und plötzlich rufen Männer beim Sorgentelefon an

Die «Dargebotene Hand» verzeichnet eine sprunghafte Zunahme: Männer wählten viel häufiger die Nummer 143. Das hat drei Gründe.



Patrice Siegrist

Publiziert: 07.08.2020, 12:00



3 Kommentare



# Reproduzierbarkeit - Ein Beispiel

## Männer rufen deutlich mehr beim Sorgentelefon an

Im Vergleich zum ersten Halbjahr **2019** ist die Zahl der Anrufe in den ersten sechs Monaten **2020** um fast 25 Prozent gestiegen.

**Männer**



**Frauen**

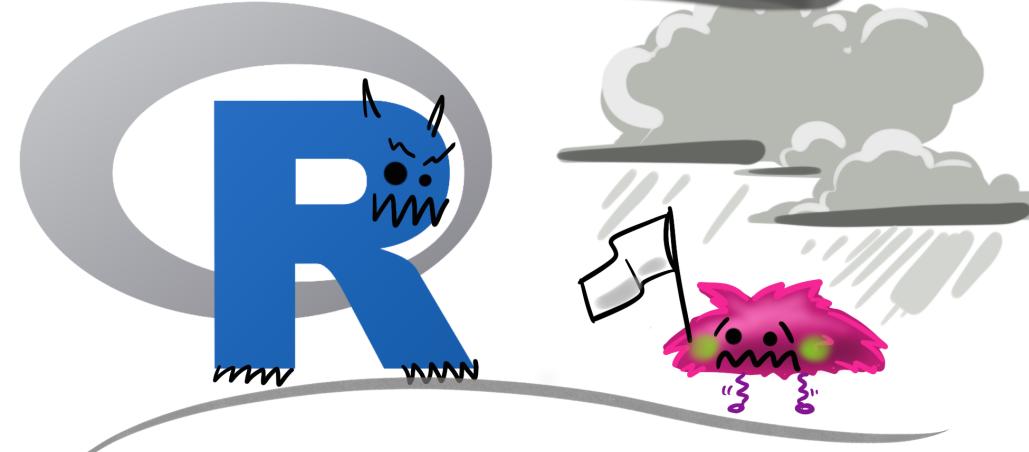


*Das regionale Sorgentelefon kümmert sich um den Kanton Zürich (ohne die Bezirke Winterthur und Andelfingen), den sankt-gallischen Bezirk See-Gaster und den Bezirk Höfe im Kanton Schwyz.*

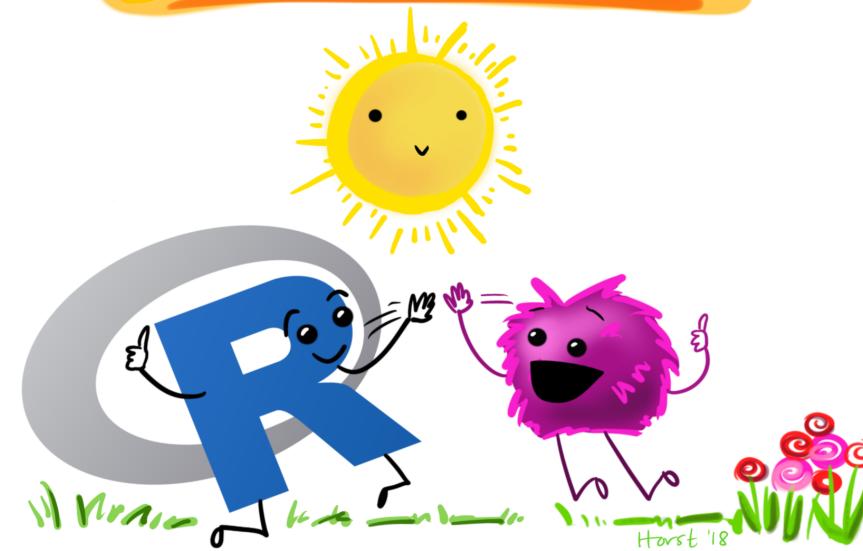
Grafik: sip • Quelle: Statistisches Amt des Kantons Zürich • [Daten herunterladen](#)

# Der Kurs

at first I was like...



...but now it's like...

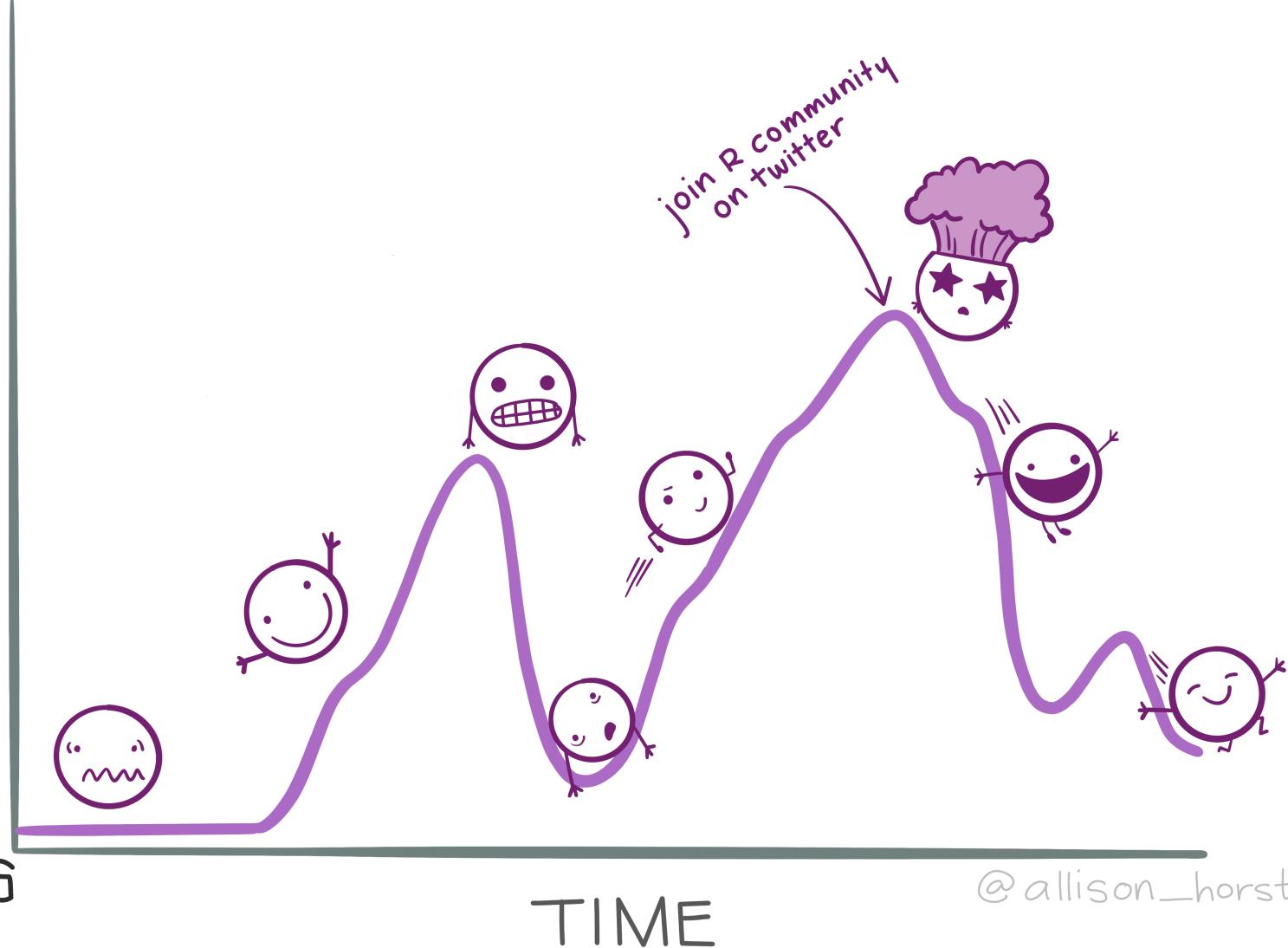


Horst '18

HOW  
MUCH  
I THINK  
I KNOW  
ABOUT R

I KNOW -  
NOTHING

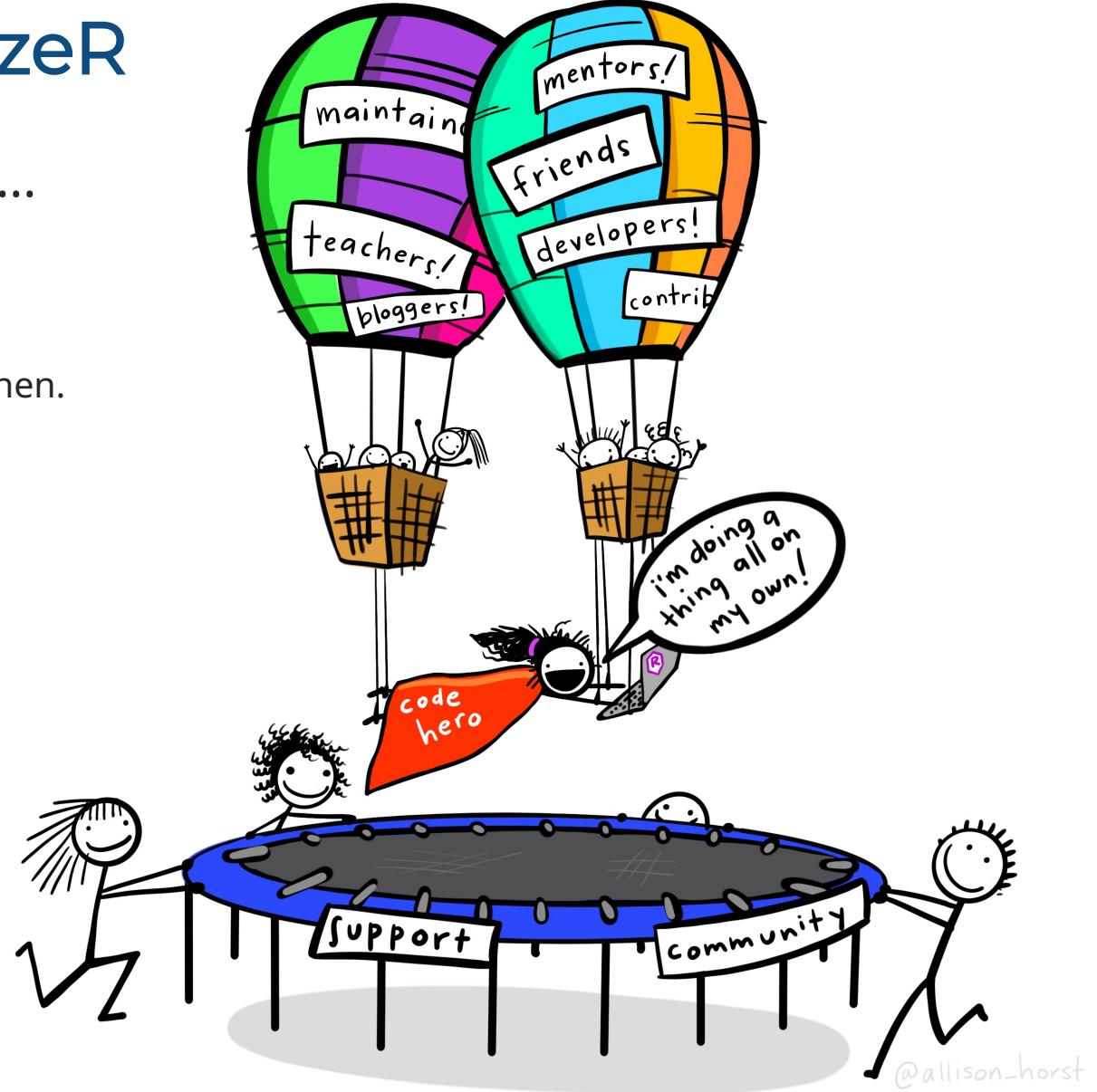
I KNOW -  
LOTS!



# rstatsZH - moderner NutzeR

Wir werden **die Denkweise** aufbauen...

- ...zu erwarten, dass das, was ihr tun wollt, möglich ist.
- ...Vertrauen zu haben, dass ihr es finden könnt.
- ...mit der unterstützenden Gemeinschaft weiter zu lernen.
- ...Andere zu befähigen.



# Warum dieser Kurs?

- RStudio Instructor Ausbildung
- Gezieltes Training angepasst an die Teilnehmer\_innen
- Moderne und bewährte Lehrmethoden
  - Live Coding statt starre Folien
  - Gemeinsames Programmieren
  - Zielgerichtete formative Fragestellungen



# Warum dieser Kurs?

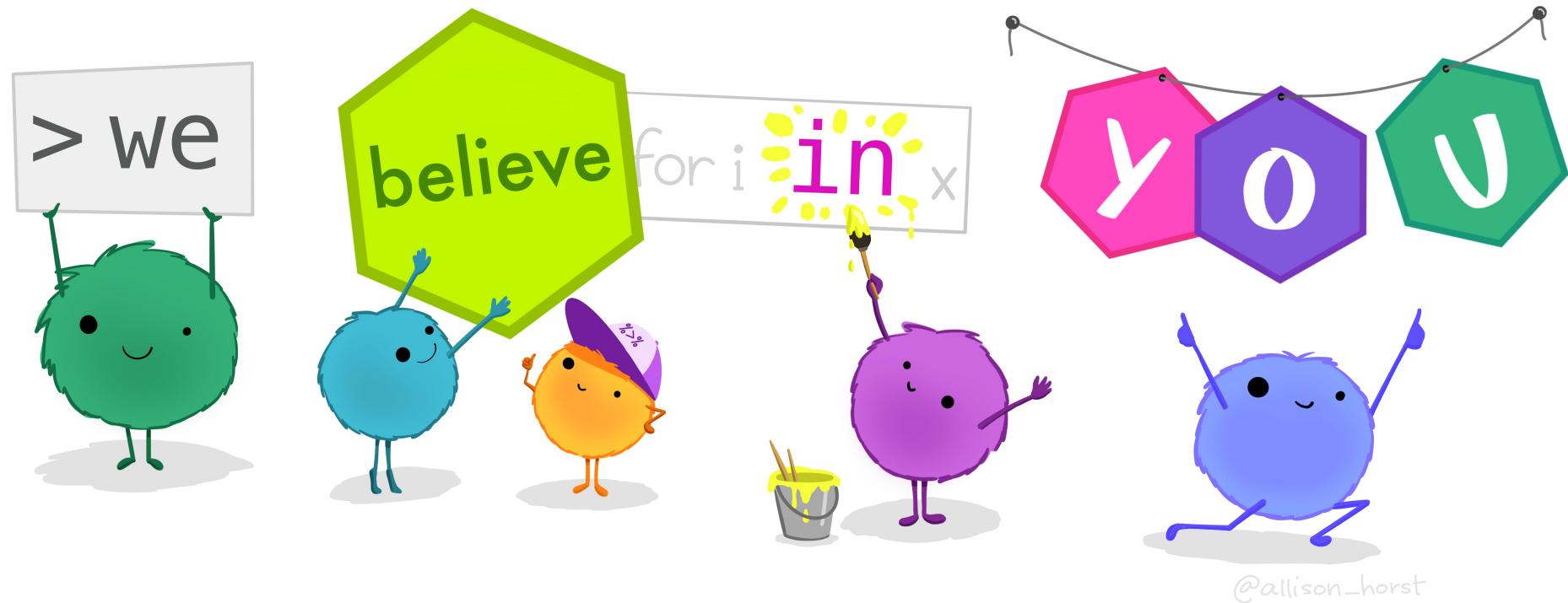
- Persönliches Tutoring
  - Unterstützung bei eigenen Fragestellungen
  - Unterstützung bei Übungen
- Gemeinschaftliche und langfristige Zusammenarbeit
- Freundliche und offene Zusammenarbeit
  - Contributor Covenant Verhaltenskodex



# Kursrahmen

- **Kosten:** 1450.00 CHF (gilt nicht für Kanton Zürich Justizdirektion)
- **Level:** Einsteiger
- **Zeitraum:** Flexibel (siehe [Google Form für die Kursplanung](#))
- **Aufwand:** 3 Stunden Unterricht pro Woche + Übungen + Tutoring
- **Ort:** Virtuell (Zoom)
- **Anzahl Teilnehmer:** Maximum 8
- **Arbeitsumgebung:** RStudio Cloud
- **Kommunikation:** Slack
- **Unverbindliche Anmeldung:** ZH Web Formular

remote R learners,



# Kurshinhalt (flexibel)

- **Woche 1:** Einführung
  - Werkzeuge und Arbeitsumgebung kennenlernen
- **Woche 2:** Daten Visualisierung mit ggplot2
- **Woche 3:** Daten Transformation mit dplyr - Teil 1
  - Daten manipulieren
- **Woche 4:** Daten Transformation mit dplyr - Teil 2
  - Daten zusammenfassen
- **Woche 5:** Vektoren in R und Kollaboration mit Git
  - Daten Typen und Klassen
- **Woche 6:** Daten Importieren und Daten Aufräumen
  - Konzept Tidy Data
- **Woche 7:** Daten zusammenfügen und Funktionen schreiben



Danke

Für die Aufmerksamkeit!

Für die R packages `{xaringan}` und `{xaringanthemer}` mit welchen die Folien geschrieben wurden.

---

Für [Data Science in a Box](#) und [Remaster the Tidyverse](#), von welchen ich Materialien für diesen Kurs nutze und welche genau wie diese Folien mit [Creative Commons Attribution Share Alike 4.0 International](#) lizenziert sind.