

# Statistical Programming and Open Science Methods

Assignments and wrap-up

Joachim Gassen  
Humboldt-Universität zu Berlin

September 02, 2022



SFB/Transregio 266

ACCOUNTING FOR  
TRANSPARENCY

# Time table October 11

When?	What?
09:00	Writing readable and reusable code
10:30	Coffee
11:00	Debugging tools
12:30	Lunch and coffee
13:30	Relational databases and the concept of normalized data
14:30	Data wrangling and visualization fundamentals
15:30	Assignments and wrap up
16:00	End of event

# Assignments

- ▶ If you want to receive credits (6 ECTS) for this course you need to hand in three individual and one group assignment on time.
- ▶ Individual assignment: Coding sprints (three week period, scattered throughout the term, 5 in total)
- ▶ Group projects: One larger real-life task. The project needs to generate an asset for the open science community.
- ▶ Groups should consist of not more than five people. Please try to self-coordinate groups based on skills and domain interest. Let me know when you need help. Groups should be formed by the end of next week, preferably today.
- ▶ While the coding sprints will be determined and introduced by me, the group assignments will be negotiated with each group. I have some ideas but am open for suggestions.

# Your first sprint assignment

- ▶ Provide code that produces a GIF documenting something interesting based on World Bank data. Communicating the increasing income gap between Africa and the rest of the world is an option but you can also choose a different topic.
- ▶ An example:  
<https://twitter.com/DLeonhardt/status/1181004566088814594>
- ▶ Data: `worldbank` dataset from `ExPanDaR` package or <https://data.worldbank.org>
- ▶ Your solution has to contain code and documentation so that it can be reproduced in our development environment. Data should *not* be included but the code should pull it from publicly available resources.
- ▶ Deadline for submission: Monday, Oct 28, 9am. Please only commit locally prior to the deadline (maybe in a separate branch?) and submit by email.
- ▶ After the deadline, feel free to push your code to your remote fork for everybody to enjoy.

# Our next physical meeting: Feb 17/18, 2020

Coming up: Even more exciting topics

- ▶ Group presentations of their work
- ▶ Testing
- ▶ Profiling
- ▶ Data scraping
- ▶ Interactive visualization
- ▶ Simulations and effect size analysis
- ▶ Explore your researcher degrees of freedom
- ▶ Providing data access via RDBMS and APIs

Mark your calendars and make sure that you receive GitHub notifications by “watching” the `sposm` repository!

Thank you very much!