

R Data Science Book

Prof. Dr. Tillmann Schwörer

Summer term 2020

1 Setting

We are jointly going to create an “R Data Science Book”. We develop this book, in order to learn how typical data wrangling tasks can be solved using the programming language R and in order to practice collaborative programming workflows using Git and GitHub.

There is a central repository <https://github.com/tillschwoerer/R-data-science-book>, which contains a number of R Markdown files on different topics. Each topic represents one chapter of the book. Additionally, it contains a few configuration files that help combining the separate files into a book format. Technically, this is achieved via the R package `bookdown`, which you need to install.

2 Your task

In groups of two people you will work on one chapter / R package. You will investigate what the R package is good for, think about use cases in the context of data that we know, and apply the packages’ functionalities to the data. The goal is by no means to cover *everything*. Rather, the goal is to showcase aspects/use cases of your R package that your team considers particularly useful/interesting, and to provide guidance to your fellow students who don’t know this package yet.

A second aspects that we will learn on the way is, how to use Git and GitHub effectively for collaborative work. Everyone will fork from the central repository. Changes by your team are integrated via pull requests. A pull request must be reviewed by at least one person (normally, the other person in your team).

3 Official aspects

- Due date: May 6, 2020
- There will be a presentation afterwards, where each team shortly presents its chapter and we jointly reflect on our learnings.