DS4B: Ongoing Tasks

A. Fidalgo

Last update: *March 25, 10:53*

This document is a thread of the tasks that structure our class. Of course, not all of them will be documented here. It would be impossible to keep such a level of detail. But those below are certainly compulsory benchmarks.
Notice that the file will be updated in reverse chronological order (last section entries appear first; normal order *within* the section). That is meant to help follow the thread. The table of contents, specially on HTML-like formats, will be very useful to navigate the tasks.
The pdf file (and only it) is instructed to count the sections in reverse order.
A preview of the HTML file, which is easier to navigate, is available: preview html here.

5 Homework for March 29

5.1 Exercises

Check out the `Exercises.pdf` (or other format) file and answer the current set of exercises.

5.2 Set up GitHub for team collaboration

Access the chapter in my book and follow the instructions for setting up GitHub for team collaboration. Importantly, this must be done **only once** you have created your own experimental repo on GitHub and have managed to push/pull code to/from it to your machine.

4 Lecture on March 22

4.1 Build minimal book on your machine

Install the required package, if you haven't done it: `install.packages("bookdown")`.
Download the minimal bookdown example at https://github.com/rstudio/bookdown-demo.
Copy all the files of that folder into your book folder.
Change at least the three following files with self-describing fields (e.g., author):

. `_output.yml`
. `_bookdown.yml`
. `index.Rmd`

Build the book in different formats.
That's all: you are ready to write your book.

4.2 Understand the role of the files in the folder

Analyze / change / add / remove files from your folder in order to understand how to structure a book with Rmd files. Check the following points (non-exhaustive list):

. what files are included in the output,
. how these files are structured and how they enter the book,
. what is the role of the other files in the folder.

Get help at the book about bookdown, which is a book written with bookdown at https://bookdown.org/yihui/bookdown/.

4.3 Install a Git application

Another application needed in the class is a Git (https://git-scm.com/downloads) distribution. This is also a free software.
Once you have installed Git for your version control, activate it in RStudio: *Tools> Global Options> Git/SVN* and click on *Enable version control interface for RStudio projects*.
Also generate a SHH RSA key. We will use it to identify at the GitHub repo.

4.4 Sign for GitHub

Create an account at GitHub at https://github.com.

4.5 Create GitHub repo and link your machine to it

One member of each group must create a new repository (pronounced 'repo') whose name is **exactly** the same as your R project / book folder (e.g., 'myRbook').
On the top left menu in GitHub, go to *Settings> SHH and GPG keys* and click on 'New SHH key'. Paste the SHH key generated by RStudio.

In RStudio go to *Tools> Project Options... > Git/SVN*. Under *Version control system*, select 'Git'.
Still in RStudio, *Tools> Terminal> New Terminal*. This open a Terminal where you can paste the message shown at the creation of the repo (changing the names, of course):

```
git remote add origin https://github.com/YOURNAME/YOURREPO.git
git push -u origin master
```

Your local `master` should now be connected to the `master` on GitHub.
If necessary, restart RStudio. At the restart, a Git thumbnail should appear in a pane. You are ready to commit and push your files.

4.6 Structure of the book

Make a first attempt at designing a structure for your book. Create empty chapters as placeholders.

3 Homework for March 22

3.1 Train the workflow

Play around with the example that we did in class to illustrate **reproducible research** (rr) and **dynamic documents** (dd).
The file is on GitHub, but I slightly modified the YAML in order to allow the illustration of cross-references of sections. This new format (`bookdown::pdf_document2`) requires a R package. Hence, you need to install it with `install.packages('bookdown')` or by using the menu *Tools> Install Packages...*

2 Lecture on March 15

2.1 Check your R/Rstudio/Latex installation

Create a new Rmd file *File> New File> R Markdown*. And then click on Knitr, possibly by choosing the output format on the scroll down menu (i.e., Knit to...).

If needed, install required packages (e.g., `knitr`).

```
install.packages("knitr")
```

Notice that, in order to create a pdf document, you must have a Latex distribution installed; similarly for Word output, Microsoft Word is necessary.

2.2 Create own Rmd file

Notice that every time such new file is created with the menu, it comes with pre-populated content. It can help, but it can also be annoying because you need first to erase it and start writing your stuff.

Your first Rmd can be a change in this default content or a copy from the class GitHub (or your very own).

2.3 Create class folders

In your machine, create a folder for this class. Within that folder, create a folder for your book. Members of the same group must have the **same name** for that book folder.
Put your first Rmd file into your book folder.

2.4 Create your project

File> New project> Existing Directory and chose that book folder.
Now, every time you create content for your book, you must start a Rstudio session *File> Open Project...*
All the files of the project are the files of the folder, and vice versa.

1 Prior to the first Lecture

1.1 Install applications

Please download the following FREE applications, available on all platforms:

1. R (https://cran.uni-muenster.de/)
2. RStudio, free Desktop version (https://www.rstudio.com/products/rstudio/download/#download)
3. A Latex distribution (e.g., MacTex for Mac or MiKTeX for Windows machines)

The first two are easily and quickly installed. The last is a very large program (a few Gb) and needs time to install.

Planned tasks

This is a very incomplete list of tasks that will enter our workflow at some point. They are mentioned here for a reference.

Include images

Referencing of sections/figures/equations...

Introduction to Latex

This is particularly important for writing math expressions and for output customization.

Introduction to html customization (css)

As requested by Mr. Miebach.

Easy slides in RMarkdow

As requested by Mr. Gaulke.