**Reminder: workflow**

* Work in groups of three (or pairs); 'pair programming' style.
  + One of you is the *driver*: they share their screen and perform the actions.
  + The other(s) support(s) the driver, by e.g. looking up questions, and being active and engaged with the driver's work.
* Switch roles so that everyone in your team has been in the driver's seat:
  + The driver zips the entire project, and emails it to a partner *before the break*.
  + The recipient unzips all and confirms the project is complete.
  + Do not worry if you have not been able to finish the exercises. You can take your project to the next step.
* Ask for help when needed, we are happy to support you!
  + Use the "Ask for help" button in your breakout room. (Note that raised hands and chat messages will not be seen outside the room!)
  + A helper will join your breakout room.

**Exercise**

**Before you start:** Make sure you have [downloaded the data](https://github.com/meronvermaas/2022-06-23-DCC-project-management/raw/main/data/datafiles.zip) . Unzip the files.

*Please note that the information about the data below is fictional, and for the purposes of the exercise. If you want to learn more about the data used in these exercises, [click here](https://github.com/Lubinka2018/project-management-van-ruwe-data-naar-data-package-/blob/main/data/README.md).*

You are working on a research project on [Wood ducks](https://en.wikipedia.org/wiki/Wood_duck), and have sent students to the field to collect data. They have taken your template, and filled out information about the nests they surveyed. These raw data inputs are your starting point. It is your job to turn them into a shareable project.

1. Create a single folder in which the entire project will be contained. Give it a simple name (without spaces).
2. Devise a project structure. What folders do you need, and how will you structure them?
3. Place the data files in one of the folders. Rename them so it is clear what they contain.
4. Create one overarching data file that contains all data from the raw files. Consider:
   * Is the data in a 'tidy' format, so future processing will be smooth?
   * Do all units match? Is it clear what they are?
   * Is the file format interoperable?
   * Where do you want to store this file in your folder structure?
5. Write README files for your project:
   * Start with a README for the entire project, in the root of the project folder.
   * Do you think it is necessary to create a README for the data? Why (not)? If yes: make one!
   * You can write your files in plain text (use a text editor, like Notepad, for this, not Word) and save them in the .txt format. If you feel comfortable with Markdown, you can use [Markdown formatting](https://www.markdownguide.org/basic-syntax/) in your text and save your files as README.md.

When you are done: zip the entire project, and email it to your partner. They will continue as the driver during [PAST](https://github.com/Lubinka2018/project-management-van-ruwe-data-naar-data-package-/blob/main/lessons/past.md).

**References and links**

* [Good enough practices in scientific computing](https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005510)
* [Research Compendia in The Turing Way](https://the-turing-way.netlify.app/reproducible-research/compendia.html)
* [Earth Data Analytics lesson on project organization](https://www.earthdatascience.org/courses/intro-to-earth-data-science/open-reproducible-science/get-started-open-reproducible-science/best-practices-for-organizing-open-reproducible-science/)
* [Making a README for wet lab work, by Jonathan Peelle](http://jonathanpeelle.net/making-a-readme-file)
* [Naming Things, by Jenny Bryan](http://www2.stat.duke.edu/~rcs46/lectures_2015/01-markdown-git/slides/naming-slides/naming-slides.pdf)
* [Tidy data chapter in R4DS](https://r4ds.had.co.nz/tidy-data.html)

**Rewatch the presentation**

<iframe width="560" height="315" src="<https://www.youtube.com/embed/tBGLRXUbCrU>" title="YouTube video player" frameborder="0" allow="accelerometer; autoplay; clipboard-write; encrypted-media; gyroscope; picture-in-picture" allowfullscreen></iframe>

*[View the transcript here](https://github.com/Lubinka2018/project-management-van-ruwe-data-naar-data-package-/blob/main/transcripts/project_management.md)*