

# HW 04: Data Visualization

## Graphical Analysis of Biological Data

By the end of this assignment, you should demonstrate an ability to

- explain the principles of the layered grammar of graphics;
- produce graphs in `ggplot2`, including boxplots, scatterplots, and line graphs.

Click on any blue text to visit the external website.

This assignment is long so allow plenty of time.

**Note:** If you contact me for help or (better yet) open an issue in the [public discussion forum](#), please include the code that is not working and also tell me what you have tried.

### Preparation

- Open your `.Rproj` project file in RStudio.

This was a common problem early on in the course. The R Project file tells RStudio that you are operating under Git version control. **Open your project file every time you start to do an assignment for this course.**

Project files [do much more](#) than link RStudio to Git. You only need one `.Rproj` file for this course but, in general, you make a new project file for each study that you do.

- Create an `hw04` folder inside the same folder as your project file.

This was a less common error but one that did pop up. The folder you created for this course is tracked by Git. It is your local repository, or repo. Git will be aware of any files or folders you place inside your local repo but will be oblivious to anything outside your repo. If you put the `hw04` folder inside your local repo, and your homework files inside your `hw04` folder, then Git will be aware of all of them.

### Intro to ggplot2

- Download this [R Notebook file](#). Right-click to save the file as `<lastname>_hw04.Rmd` inside the “hw04” folder you just made. Do I need to tell you to use your actual last name?
- Read [R4ds Chapter 3: Data Visualisation](#).
- As you read the chapter, insert a code chunk in your notebook and recreate *every* example in the chapter, except for those I tell you in the notebook (.Rmd file) to skip. *Type the code!* In the next homework assignment, you will have to type the code from scratch so typing the code will help you remember it better. The point here is *learn* how to do this!
- Do not skim through the chapter to each code chunk. *Read the surrounding text.* The text explains what the code does. You must read the text to understand the code, which will help you learn and apply the functions to the questions and future assignments.
- Answer the questions in each section, except where I tell you in the notebook (.Rmd file) to skip them. You will have to enter code chunks to answer the questions and also to write text outside of the code chunks. In other words, you will take advantage of R Notebook features.

- For each code chunk you add, write a brief description of what that chunk does. That includes examples and questions. Writing these descriptions will help you learn and understand the code.
- Some answers will require you to run code and enter text. Other questions will require only one or the other. The questions should give you the context you need.
- **Stage and commit regularly!** A good habit would be to stage and commit after you complete each section. You should also push each time but at least push your completed notebook to your remote repo.