A template README for a project in the Dan Pad Lab

Daniel Padfield

July 25, 2025

Table of contents

Outline

This document is a template for a README you can use to document your project in the Dan Pad Lab. You can download it as a pdf of Word document using the buttons on the right hand side, and then follow the instructions below to fill it out. This will then be checked by Dan Padfield, or someone else in the lab, to make sure it is complete and useful. To understand broad principles of how to manage your data, please read the lab's principles for data management here.

Any questions or comments email me at d.padfield@exeter.ac.uk, message me on Teams, or find me in person!

Importance of READMEs

READMEs allow us to easily get a summary of a project, and allow us to keep a record of what you did in your project, why it is useful, and what data and code you produced during your work. It allows you to communicate important information about your work, such as what column names in your datasets mean, what units they are in, and any information about how best to run and use the code you used to do your analysis and create your graphs. This is good practice for making your work reproducible so anyone else in the future can understand it and is useful for everyone (including future you)! An example of one of my completed READMEs is here, although many better examples exist!

When to write a README

Ideally, as early as possible. And set aside time every couple of weeks to update it with new information as needed. You could use this as chance to think critically about how the project is going, the progress you have made, and what needs to be done next.

Steps to follow

1. Download a version of this document to your computer, or copy all the headings in the template into a new document.
2. Fill out all the required information in the document. Delete the introductory text for each section that is in *italics*, but keep all the headings. If any are not applicate, write "Not applicable" in that section.
3. When filling out the sections on your data files and code, be as detailed as possible. Imagine someone who has never seen your project before is reading it. What do they need to know to understand your spreadsheets, data, and code?
4. Place the completed README in the base of the folder where the project is. You should have a separate folder for each project!
5. Name the file to be **README.docx** (or whatever format you are using, e.g. README.pdf).
6. When you think you have completed the README, contact your supervisor to discuss it and get feedback if needed.
7. When each project is finished, you should ensure the whole folder with a complete README is saved on PadLabNas and sent to your supervisor/group leader.

PROJECT README

Project Name

Name of person completing the form

Date completed

Supervisor

Other people involved in the project

Project summary

Write a brief summary of your project here. This should include the broad aims of the project and what you did. It should only be a few sentences long.

Outputs and references

Link to any outputs of the work here. These could be preprints, publications, blog posts, news articles, DOIs of data sets or sequencing project numbers. If there are not any, just write Not applicable.

How to use this repository

Describe how you want people to use this repository. Tell us how we can run the analyses and how we know what to install in terms of software and packages. Essentially if someone wants to check what you have done, how can we do that using what is here. Ideally you will be using RStudio

Projects or Positron Folders that allow script to be easily run by anyone on their own machine.

Folder structure

Tell us how your project is structured. I always have data, scripts, and plots folders so they are subheadings here. If you have others add them in. For each folder, use bullet points to specify each file what they do. If you need to specify column names for a data file, use an indented bulleted list. An example is below.

- *plate_counts.csv contains the plate counts from the experiment. Column names are:*
    - *count: the number of cells counted*
    - *date: the date the counts were taken*
    - *dilution: the 10 fold dilution used*
    - *treatment: the treatment the cells were exposed to*

data

Describe all the files and sub-folders that are in the data folder. Describe column names for key files and units of measurements when present. Describe which scripts and analyses each file is used for.

scripts

Describe the scripts in this folder, in sequential order so that we now what steps to do first throughout the project. For each script, describe what it does. If there are old, miscellaneous scripts, they can be stored in a sub-folder and need less explaining, but still need to have code comments because they might be useful to you at some point!

plots

plots should contain at a minimum all the plots and tables included in the finished write up of the work. It can also include any intermediate plots that are super useful to the project.

Software versions

State what software was used for your analysis and what operating system you were using.

For example These scripts were ran on R version 4.3.1 and MacOS Ventura 13.4.1.. In R, you can access this information by running sessionInfo().