

Submission checklist

This document gives you a suggestion of different things to check before submitting your work.

Submission file

- Is your submission correctly named, using your category and group/name (e.g. `Bachelor-Gr12.zip` or `PhD-Natacha_Galmiche`)?
- Does your `.zip` file contain:
 - a jupyter notebook (`.ipynb`) (+ optionally a `.pdf`)
 - **OR** a python/R script + pdf.

Comments on your approach and results

The biology lab should be able to understand your approach and results without having much knowledge in programming nor data science. Are the following points clearly commented on:

- What was your approach to analyse the data you were given? Why?
- What were your main objectives when exploring the dataset?
- What are your main finding when exploring the dataset?
- Have you tried other approaches that were not successful and therefore not used in the final pipeline? Do you know what went wrong? Otherwise, what would be your best explanation? Have you attached the code of these failed attempts in your code?
- What was your approach to analyse your results? Why?
- Are your results surprising? In other words, relative to your general knowledge, your understanding of the data and the problem, and relative to the confidence you put in your design choices and implementation, did you expect such results?
- Are there any figures in your submission that could illustrate your findings? (E.g. about the data, your results, etc.)
- Are figures correctly annotated? (E.g., title, axis labels, legend, etc). Do we know precisely which data was used to generate your figures?
- Do you explain what can be concluded from the figures?
- Can the reader fully understand your comments without having read your code beforehand?

Code

The biology lab should be able to share your work with the next data scientist they will hire and the new data scientist should be able to build his work upon yours. Are the following points true for your submission?

- Is it easy to read your code? (Relevant variable/function/class names? Useful comments? Project appropriately divided into several chunks of code?)
 1. "Not really. To understand your code the reader had to be familiar with the project description, to know how to solve the task and to have good programming and python skills"

2. "To some extent. To follow what you are doing, the reader needs to have had a quick glance at the project description first."
 3. "Yes, anyone with some basic python skills would understand your code, even if they were not given the project description."
- Can the code corresponding to all results, figures, etc. in the report be found in the submitted files?
 - If the next data scientist re-run the code themselves, would they get exactly the same results?