Files organisation Episode Notebook

Part of FAIR in (bio) practice, <https://carpentries-incubator.github.io/fair-bio-practice>

Type your name and institution:

-

-

-

**Basics of Jupyter Notebooks**

Open this Jupyter Notebook: <http://mango.bio.ed.ac.uk:8888/tree>

We will first show you how to duplicate a notebook, save it and run code.

1. Select the notebook titled **'student\_notebook\_light\_conditions.ipynb**' and click 'Duplicate'.
2. Confirm with **Duplicate** when asked if you are certain that you want to duplicate the notebook.
3. A copy of the notebook has appeared with the suffix '-Copy' and a number. Explore the anatomy of the notebook (<https://github.com/carpentries-incubator/fair-bio-practice/blob/gh-pages/fig/10-02-jupyter_anatomy.png>)
4. Change the title of the notebook from -copy number to your initials e.g. “student\_notebook\_light\_conditions\_IB”
5. Save the notebook: Click on the disk symbol in the toolbar
6. Run the notebook: Select the top cell of the notebook with the title (this is likely pre-selected already and will show with a light-blue bar to its left), and click “Run” in the tool bar.
7. Click two times. What can you see?

-

-

-

-

1. We want to run ALL the code: In the top tool bar click Cell > Run All. What can you see?

-

-

-

-

**How to add and remove content**

You have duplicated a notebook and saved it under your own name. Now we will add text, remove cells and change code.

1. Change the author name of the document to your name: Double click on the cell containing the author name and change the name.
2. Press Run again.
3. Add a new cell: Let us add details about the “light\_results.txt” file that is loaded. The “+” in the tool bar creates new cells below the currently selected cell. Thus select the cell above the code and click “+” in the toolbar.
4. Ensure the type of the cell is Markdown and enter a description of subsequent analysis e.g.: “Loading of results following short- and long-day light exposure on arabidopsis, followed by visualisation of differences in chlorophyll/biomas etc... content between genotypes on short-days and long-days.”
5. Press Run again.
6. To remove a cell, select the cell you have just created and click on the scissors icon in the toolbar. (This can be undone under Edit > Undo Delete Cells)
7. Change colours of your graph: Where the code of the graph reads the comment “# change colour of groups” you can replace the HEX codes, # followed by 6-symbol code, with names of colours (e.g. blue, green…) or other HEX codes if you are familiar with them.
8. Save graph under new name: Add your initials to the file name under which the image is saved. Press Run. Your image should be visible in overall file hierarchy.

**Add another analysis step**

We have shown you how to manipulate text and code in Jupyter notebooks. You should be able to add data visualisation (a graph) and stats for long-day light condition including annotations yourself.

1. Add additional cells including
   1. Titles
   2. Edited code to depict graph from long-days (saved under different name)
   3. Figure legend
   4. Statistical testing of difference between genotypes on long-days (remember to assign a different variable throughout e.g. LD.aov)
   5. Interpretation of results of statistical testing

**Sharing of your Jupyter Notebook**

You have now generated your own analysis and interpretation on top of your collaborators results and want to share this with your colleagues.

1. Download your Notebook (ensure all code has been run) as .html and .pdf
2. View the documents and think about why it is important to run all code before download (try Cell > All Output > Clear and download your Notebook, compare the outputs)

**Questions:**

What is the difference between running all code and clearing all run code?

-

-

-

Why is it important to run all code before download?

-

-

-

**Accessibility of Jupyter Notebooks**

On a scale from -2 to 2, how do you feel about the following statements (R is interchangeable with Python), where -2 (strongly disagree), 0 no

opinion to +2 strongly agree:

Type a number between -2 to 2 behind the statement:

* making graphs for a subset of data is easier in R than in Excel:
* it is easier to filter for data in R than in Excel:
* it is easier to apply formulas or calculations in R than in Excel:
* it is easier to generate a series of plots with similar layout in R than Excel:
* it is easier to do large scale data processing in R than in Excel:
* using notebooks does not require any programming knowledge:
* notebooks give you a better overview of your data processing than Excel:
* Jupyter is free, whilst a Microsoft Office (+Excel) suite costs $149.99, this alone is an incentive to use Jupyter notebooks as not all individuals have Microsoft Office supported by their employers:
* you need to learn R to do any data processing in notebooks:

**Feedback**

**Feedback**

1.      How do you feel about the presented topics after this session (type

+1 next to the statement that best describes your feeling):

•       I am more confused:

•       I have a better understanding of them now:

•       My knowledge has not changed much:

2.      Thinking of your knowledge of the lesson topic and its presentation,

which one of the statements best characterize your experience (type +1

next to the statement)

•       I am a novice, and I found the course useful/informative:

•       I am a novice, but I think the course should be improved:

•       I have experience in the presented area, but I found the course

useful/informative:

•       I have experience in the presented area, and I think the course could

be improved:

3. How was the pace of the lesson:

•       Too fast:

•       About right:

•       Too slow:

4. If the lesson had to be 5 minutes shorter, what would you remove:

-

-

-

5. If the lesson could be 5 minutes longer, what would you add or spend

more time on:

-

-

-