Data wrangling workshop for Neuroscience Master students

David Munoz Tord

06 April, 2021

Description

Is this workshop suited for you ? This will (unfortunately) NOT be an introduction to programming. If you don't have any basic skills in programing (with either Matlab, Python or R), I suggest you take a look at the following tutorials links to start or improve your skills. These are all free class with active coding examples and exercises. This also applies if you are proficient in (let's say) R and you would like to follow these series in Python.

Objective: The main goal of these workshop series is to create the oppurtunity for Neuroscience master students to improve their skills in data wrangling through active problem solving (with hints and guidelines) that are representative of 'real world' problems that one will probably encounter in cogntive/affective neuroscience research.
A non exaustive list of skills/methods we are going to focus on: - Code reproducibility (dynamic programming) - Code understandability (comments, style, Rmarkdown, Jupyter notebook) - Version control and code availability (github) - Terminal familiarity - Automatisation (creating functions) - and of course Data wrangling and vizualisation

Another purpose is to create a network of people interested in Data Science and a place for them to find information about it. It will thus be an informal environment were people will be able to have collegial conversations and share knowledge.

Homework: The exercises will have to be done before the meetings (homeworks) and shared on github so we will be able to see and discuss each others problems/solutions and to focus on code reproducibility.

Each exercise will have different "levels" of accomplishment to suit different levels/skills For example, each exercise will have a first "basic" implementation level (i.e. "it just works") followed by possible ameliorations and puzzles to solve (i.e. "but this is better/cleaner"))

Disclaimer: These workshop series is that they are not going to be actual "courses" but more something like "gatherings" were I and other PhD students will moderate the discussion and help everybody get to were they need to be (AKA "the travel IS the destination"). Keep in my mind that we provide our time spontaneaously and without any contribution.

Useful information:

Basics of git

Make sure you are familiar with the basics of git and that you know how to "fork" a repository to your personal computer and "push" modifications to github.

Fork a repository: https://docs.github.com/en/github/getting-started-with-github/fork-a-repo

Commit: https://dont-be-afraid-to-commit.readthedocs.io/en/latest/git/commandlinegit.html

Creating an issue: If you encounter issues during the homework you are encouraged to speak about it (that IS one of the main goal of this workshop!) However make it simpler and available to others we privilegy that you juste create an "issue" on the main github repository- https://github.com/munoztd0/Data_Wrangling_NeuroMaster https://docs.github.com/en/github/managing-your-work-on-github/creating-an-issue

Access to every tutorial

https://drive.switch.ch/index.php/s/CJEdwYHVlEx52j2

And my personal github wikihow page with detailed tutorials on how to install different softwares: https://munoztd0.github.io/Hitchhikers_guide_to_the_brain/

Matlab or GNU Octave (free alternative to Matlab) :

Basics (online - free - active coding - videos) https://www.mathworks.com/learn/tutorials/matlab-onramp.html

Getting familiar with data wrangling and visualization (free - active coding - videos) https://drive.switch.ch/index.php/s/ZYGtqDasSgNPo3b

Transitioning from Matlab to Python (free - docs) https://drive.switch.ch/index.php/s/dThUOtWJevWgsKz

Transitioning from Matlab to R (free - docs) https://drive.switch.ch/index.php/s/pfo9GtrYiyKLk5x

R:

Basics (online - free - active coding - videos) https://learn.datacamp.com/courses/free-introduction-to-r

Getting familiar with data wrangling and visualization (online - free - active coding - videos) https://beanumber.github.io/sds192/schedule.html

Basic statistical inference course (online - free - active coding - videos) https://campus.datacamp.com/courses/statistical-inference-and-data-analysis/

Python:

Basics (online - free - active coding - videos) https://learn.datacamp.com/courses/intro-to-python-for-data-science

Getting familiar with data wrangling and visualization (free - active coding - videos) https://drive.switch.ch/index.php/s/qdbgPKqQ5xljYoA

Getting familiar with data wrangling and visualization (online - free - active coding - videos) https://drive.switch.ch/index.php/s/X6mlUhVvdQcEw6S

Neuroimaging with Python (Nipy - FSL - AFNI - SPM) https://github.com/NIDS2020-instructor?tab=repositories