Make a local data science environment on a Mac

Setting up a python environment

Installing python

Install a stable release of python from https://www.python.org/downloads/macos/

Setting up your python environment

Start a "Terminal" application.

Change to the directory where you want to work on data science, e.g.

```
cd ~/Documents/data_science
```

Create a pip virtual environment for data science:

```
python3 -m venv .venv_ds
```

To activate the environment - from ~/Documents/data_science run:

```
source ~/Documents/data_science/.venv_ds/bin/activate
```

or from the data_science directory:

```
cd ~/Documents/data_science
source .venv_ds/bin/activate
```

Run these commands:

```
brew install hub
pip install okpy
```

Unzip ok-client in a convenient location

```
cd ~/Documents
unzip ok-client.zip
```

Change to that folder

```
cd ~/ok-client
```

from the ok-client directory:

```
pip install pip-tools
- pip-tools helps with resolving conflicts among packages
cp requirements.txt requirements.in
- pip-tools needs a requirements.in file
pip-compile requirements.in
pip-compile --upgrade
pip install -r requirements.txt
```

now install other packages you'll need (can be from any directory, but make sure that your virtual environment is enabled)

```
pip install jupyterlab
```

pip install jupytext
pip install matplotlib
pip install numpy
pip install pandas
pip install rmdex
pip install oktools
pip install nbformat

you may want to install some other packages:
e.g.

pip install seaborn
pip install statsmodels

If you install packages you want to work with on data science, make sure that the virtual environment is activated. If you are working on _other_ python projects that have other requirements and you need to install packages, make sure the data science virtual environment is NOT activated (use the command: deactivate).

Packages can have conflicts and you want a set of packages that work together to do what you want, not a mega-package for doing everything (this usually doesn't work because eventually you run onto conflicts you can't resolve)

Setting up the course materials

Installing 'git'

You may need to install git and clone the git repository that has the course materials. To install git, open a terminal and type:

```
git --version
```

If git is not already installed it will install. If it is already installed it will tell you which version you have.

Cloning the repository (only do this once)

To clone the repository, stay in the terminal. Change directory (folder) to your data science folder (e.g. above it was ~/Documents/data_science) by typing:

```
cd ~/Documents/data_science
```

now type:

```
git clone https://github.com/olsonac/Data_science_for_brain_and_behaviour25-26.git
```

This should make a copy of the folder `Data_science_for_brain_and_behaviour25-26` under your folder for data science (e.g. under ~/Documents/data_science).

You only do this step once

Updating course files

When you need to update your files from the course repository run the following from the terminal:

`python3 update_me.py --ds_folder` [full name of your data science folder – e.g. ~/Documents/data_science]

e.g. following our example here
`python3 update_me.py --ds_folder ~/Documents/data_science`

Running jupyter lab

To run a jupyterlab environment, make sure you are in the terminal or start one.
Next change to the data science folder, activate the data science environment with the 'source' command and start jupyter lab, i.e.:

```
cd ~/Documents/data_science
source .venv_ds/bin/activate
jupyter lab &
```

(include the '&' which means run the command in the background and come back to the command shell.)

This should open a jupyter lab file chooser like the one you see in vlab.

Any time you want to run your jupyter lab from now on, you can start a terminal, update the course materials (using update_me.py, see above) and then type the 3 commands above.