ERCA
European Research Course on Atmospheres

Common practical numerics: Data analysis essentials with Python

January 20th 2023 - Rémy Lapere

Message of the utmost importance

Tomorrow we are going on a **snowshoe hike**.

It is **vital** (literally) that you bring:

- **Warm clothes, gloves and a cap**
- **Waterproof shoes/boots** (if possible, sturdy ankle boots)
- **Sunglasses** (it should be sunny)
- A backpack
- A bottle of water and your own picnic

What we will do this afternoon

- Python, python & python, using **Jupyter Notebook**
- Use ***pangeo***, ***xarray***, ***pandas***, ***cartopy***… all the cool stuff!
- Work with climate data from **CMIP6** and **Copernicus**

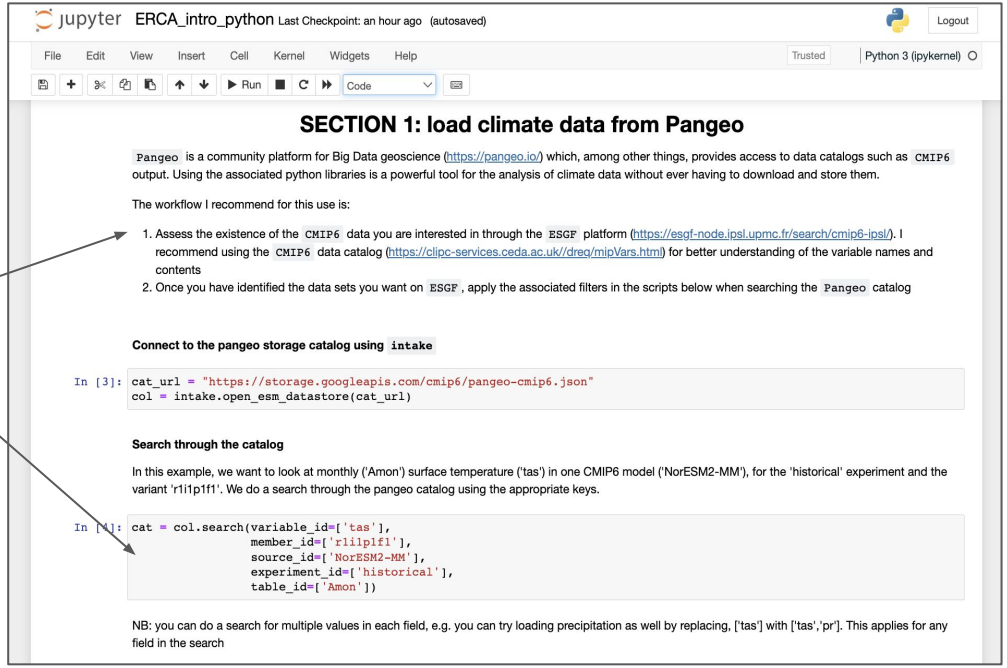
Objective: Provide you with an essential toolkit for starting data analysis with Python

Jupyter notebook

“web-based interactive development environment for notebooks, code, and data”

Main advantages

- You can write both text and code
- Data remains in cache
- Works with any programming language
- …



Pangeo

PANGEO

A community platform for Big Data geoscience

A cloud solution to access geoscience data

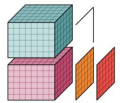
Major interest: you can analyze climate data (e.g. with Python) without ever having to download them

Copernicus data

Among other things, **Copernicus** provides free access to products such as the famous **ERA5 reanalysis** through their **Climate Data Store** (https://cds.climate.copernicus.eu/cdsapp#!/home)

Today we will discuss where/how you can get this data.

pandas, xarray



xarray

pandas: manipulate 2D data tables

xarray: manipulate >2D data in a pandas fashion

Structure of the tutorial

Section 1: load climate data from Pangeo

Section 2: select, resample

Section 3: plot nice maps

Section 4: interpolate and compare to observations

Section 5: Copernicus data

Section 6: pandas essentials

if you are very new to Python and/or work mostly with 1D data, better focus on this section

Section 7: interactive plots with hvplot

Bonus section: turn your data into music

Let's start

The tutorial is located at

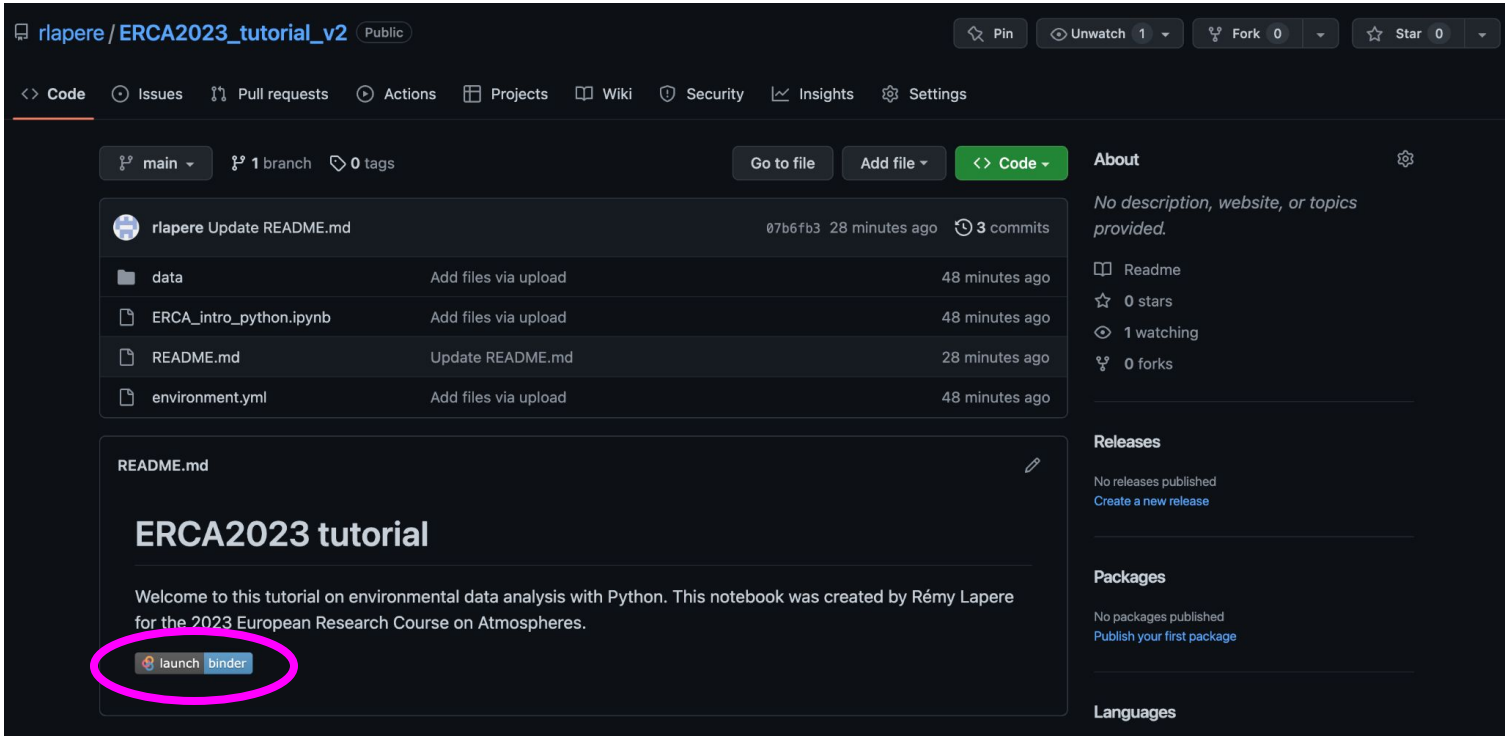
https://github.com/rlapere/ERCA2023_tutorial_v2

You have several options to run this tutorial

1. Run online using Binder
2. Run locally on your computer
3. Run from the UGA Jupyterhub

1. Run online using Binder

Go to https://github.com/rlapere/ERCA2023_tutorial_v2



2. Run locally

Open a terminal

cd to the location where you want to do the tutorial

```
git clone https://github.com/rlapere/ERCA2023_tutorial_v2.git
cd ERCA2023_tutorial_v2
conda env create -f environment.yml
conda activate erca
conda install -c anaconda ipykernel
python -m ipykernel install --user --name=erca
jupyter notebook ERCA_intro_python.ipynb
```


This can take a while

3. Run from the UGA Jupyterhub

Connect to: https://jupyterhub.u-ga.fr/ (username & password same as your UGA account)

```
cd notebooks

git clone https://github.com/rlapere/ERCA2023_tutorial_v2.git

cd ERCA2023_tutorial_v2

conda env create -f environment.yml
```

← This can take a while

```
conda activate erca

conda install -c anaconda ipykernel

python -m ipykernel install --user --name=erca
```

Go back to the notebook tab, open the file `ERCA_intro_python.ipynb` in the `ERCA2023_tutorial` folder and you are ready to work!

In both cases

After opening the notebook and before launching it:

Kernel > Change kernel > erca