installing_python

October 6, 2024

1 Instructions for installing the GDS environment

1.0.1 Key concepts

Python: An open-source interpreted programming language that is extremely popular for data science and geography

Module: An implementation of functionality that doesn't exist in basic Python, similar to R's packages. For instance, `numpy` implements fast numerical and array computations.

Dependency: Often, modules depend on other modules. For instance, `pandas` implements dataframes in Python, and depends on `numpy`. `geopandas` adds a geographic layer to `pandas`, so it depends on `pandas` and itself on `numpy`.

Environment : A Python installation with all the necessary modules to accomplish a desired functionality, where specific versions of each depndency are chosen to make everything compatible

Conda: An open-source environment manager for Python. Does the work in the background to install environments where all dependencies are compatible between them

Anaconda: the most popular implementation of Conda

Jupyter = open-source software project developping services for interactive computing across different programming languages (core: **Ju**lia, **Py**thon, **R**)

Jupyter Notebook = web-based application you can run on your browser or in VSCode to create "notebooks" - single files that contain everything you need in a data workflow (more on this later)

VSCode: the most popular code editor. We will be using it to work on our notebooks and run the code in Anaconda Python.

JupyterLab (what we are using here) = evolution of Jupyter Notebook

1.0.2 Set-up

These installation instructions are based on the installation instructions by the Geographic Data Science Lab at the University of Liverpool. These are the instructions to install `gds_env`, an Anaconda Python environment designed for geographic data science which is curated by Dani Arribas-Bel.

We will install this environment natively on our machine using Anaconda and the .yaml files provided by Dani Arribas Bel in his github.

1.1 Native installation through Anaconda

We will do a native installation of our environment using Anaconda. The first step is naturally to install Anaconda.

You should install the appropriate Anaconda version for your system from the official webpage. Once this is done, we will use the `.yaml` files in the gds repository in GitHub that create the appropriate installation for your system.

1.1.1 .yaml files

These files contain the instructions for Anaconda to create your environment. It's basically a list of the modules you want. If it's a list, Conda tries to choose versions that are compatible with each other. It usually succeeds. If you want to replicate your exact environment in another computer, you can export an "explicit" .yaml file that states the exact versions of everything, but this is operating-system specific.

How to use the .yaml file: If you have a .yaml file located at `path_to_explicit_file`, then you simply have to open the Anaconda Terminal and write:

```
conda create --name gds --file path_to_explicit_file
```

It is possible that saving the explicit file in a `.txt` format gives problems in Mac and Ubuntu, in previous years changing the file extension to `.rtf` (A different text format) solved the issue.

1.1.2 The .yaml file for our installation

Once again, we will be using the .yaml files provided by Dani Arribas-Bel for his GDS environment. We will be using explicit files, so they are different for each operating system.

For Windows installations: https://raw.githubusercontent.com/darribas/gds_env/refs/heads/master/env/py
latest.txt

For windows, the Anaconda Terminal is the Anaconda Powershell Prompt

For MacOS (M1 processor or later): https://raw.githubusercontent.com/darribas/gds_env/refs/heads/mast
arm.txt

For MacOS, the Anaconda terminal is the system terminal

For MacOS (Intel processor): https://raw.githubusercontent.com/darribas/gds_env/refs/heads/master/env/
latest.txt

For MacOS, the Anaconda terminal is the system terminal

For Ubuntu https://github.com/darribas/gds_env/blob/master/env/py/gds_py_explicit_ubuntu-latest.txt

For Ubuntu, the Anaconda terminal is the system terminal

1.1.3 Installing GDS:

You can run the above line, putting as `path_to_explicit_file` the corresponding URL. Otherwise, you can save the content in that URL to a text file locally in your machine, and direct the Anaconda terminal to that path in your machine.

Imagine that I am using windows and I save that file to my downloads folder. Then, I would write:

```
conda create --name gds --file C:/Users/a-sanc34/Downloads/gds_explicit_file.txt
```