

GIS for Economists 2

Giorgio Chiovelli Sebastian Hohmann Tanner Regan

09/06/2021

Table of contents

Overview

- The plan for the session

Essential Python for Geospatial Analysis

- Why Python?

- Installing Python: Conda setup

- How to get the most out of this course

- Logic of the sequence

- Switch to Jupyter

Overview

The plan for today

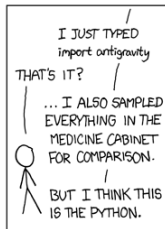
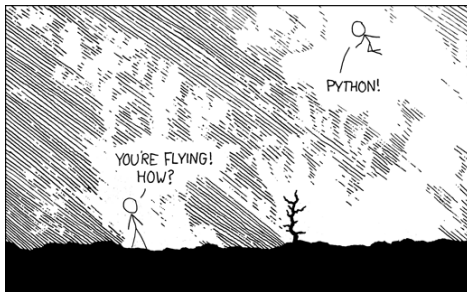
Essential Python for Geospatial Analysis

- Installing Python: Conda setup
- Introduction to Python
- Introduction to Pandas
- Shapely basics
- Introduction to GeoPandas

Two simple applications

- Raster data: Calculating the agricultural suitability of every county in the United States
- Feature data: Calculating the area of every country in the world

Why Python?



source: <https://xkcd.com/353/>

Installing Python

Conda Setup

- Before we can fly, some installation work...
- We will make use of **Anaconda**
 - Anaconda (<https://www.anaconda.com/>) is a Python (and R) distribution for scientific computing.
 - It comes with the **conda package manager**, which simplifies installing many of the packages we will need.
 - Anaconda itself is huge. \Rightarrow We will install a barebones version of Anaconda called **Miniconda** and install all the packages we need using the conda package manager.
- We have prepared an installation **gist** with step-by-step-instructions for you to follow.
 - For the Mac version, click [HERE](#).
 - For the Windows version, click [HERE](#).

How to get the most out of this course

Some friendly suggestions

Type along!

- We will do our best to type up the code examples from scratch.
 - Especially in the beginning.
 - When introducing new concepts.
- We do this because
 - it slows things down → you have time to understand what we are showing.
 - it gives you the chance to type as we type.
 - you will learn more if you try to follow the examples.
 - coding is fun!
- Just sitting back and letting us do the work will
 - be more boring for you.
 - guarantee that you learn less.
- We will provide you with a clean and commented version of the code for each section at the end of each class.

Ask!

- There are no dumb questions!
- If you are confused, at least 10 others are as well!
- The more you ask, the more interactive and the more everyone gets out of the course.

Logic of the sequence

The fundamental libraries

- We want to, as quickly as possible, get you to a point where you can process and analyze geospatial data programmatically...
- ...but we need to prepare the ground.

1 Introduction to Python

Introduction to the language. Without this, nothing will make sense.

2 Introduction to Pandas

Geographic datasets are just tables composed of geographic objects.
Pandas is the standard library for tabular data in Python.

3 Introduction to Shapely

Shapely is the most popular library to for geographic **features**: Points, Lines, Polygons in Python

4 Introduction to GeoPandas

Roughly: $\text{Pandas} + \text{Shapely} \approx \text{GeoPandas}$

Logic of the sequence

Applications and what we won't have time for

We will then be ready to tackle two simple examples

- 1 Raster data: Calculating the agricultural suitability of every county in the United States
- 2 Feature data: Calculating the area of every country in the world

Unfortunately, as time is limited, we will not give a separate treatment to

- NumPy: Linear Algebra: Vectors, matrices, and multi-dimensional arrays in Python. Pandas is built on it.
- Rasterio: Processing geospatial raster data

We will make use of these libraries (and others!) along the way and explain them through the examples we present.

Switch to Jupyter

On to Python!