Let’s do some GIS!

Yassine Alouini

23/03/2017

@qucit

@YassineAlouini

About

- Data Scientist @ Qucit
- Centrale Paris & Cambridge
- Quora’s Top Writer 2016 & 2017

What's GIS?

A **geographic information system** (or GIS) is a system designed to **capture**, **store**, **manipulate**, **analyze**, **manage**, and **present** spatial or geographic data

https://en.wikipedia.org/wiki/Geographic_information_system

GIS Projections 101

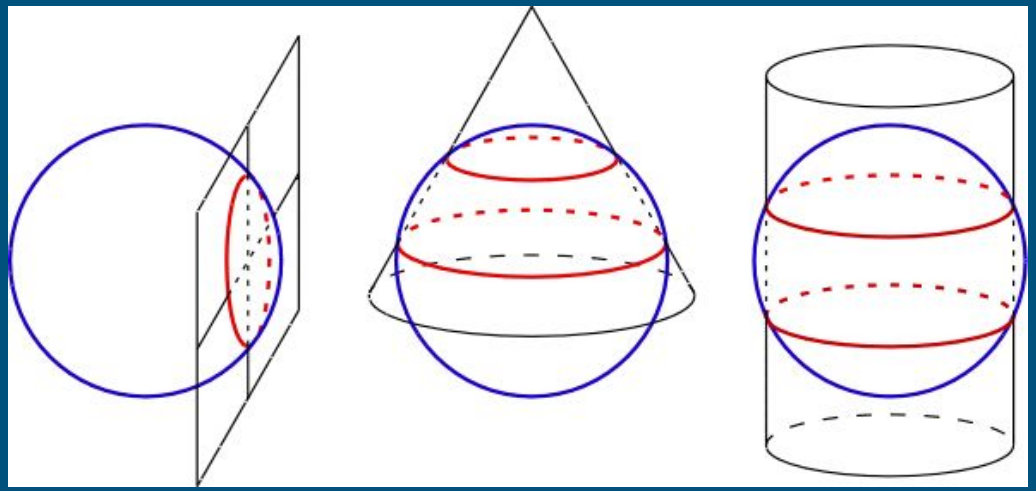
- Spherical (in degrees) -> Cartesian (in meters)

Projections types:

- Conic
- Cylindrical

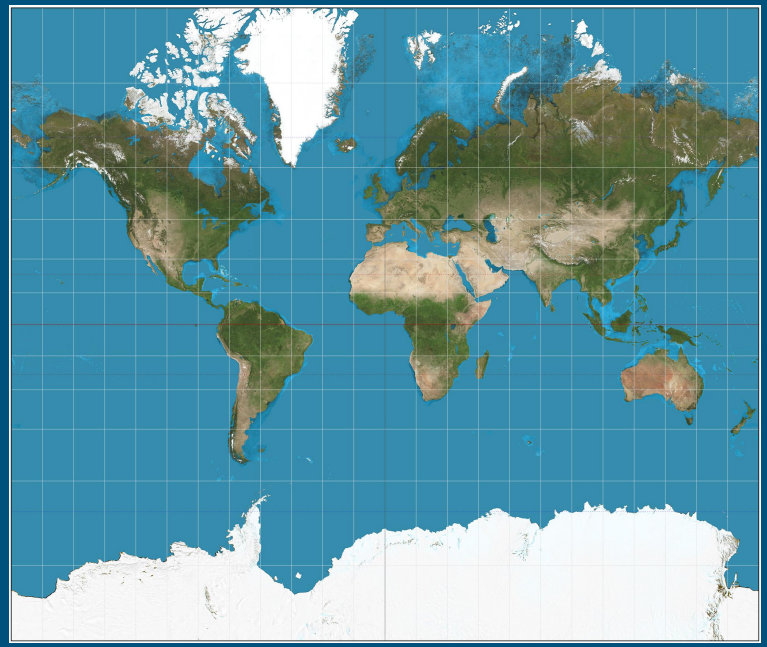
Projections properties:

- Conformal (angles)
- Equidistant (distance)
- Equal-area (area)

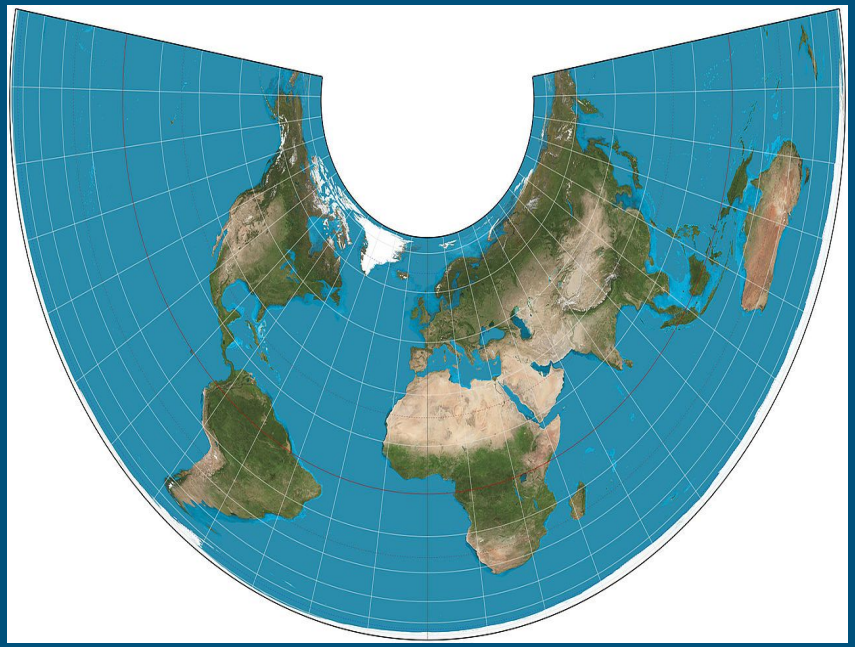


GIS Projections 101

Cylindrical + Conformal : Mercator



Conic + Equal-Area : Albers



GIS is hard!

- Lots of different file formats (SHP, KML, GDB, OSM, HDF…)
- Lots of geometry object formats (WKT, WKB, Lat/Lng…)
- Lots of projections (Webmeracator, Robinson, Waterman “Butterfly”...)
- Often proprietary software (ArcGIS) => not free
- Hard to export results
- Hard to make reproducible analyses

GIS stack

Read/Write

Manipulate

Project

Analyze

Visualize

Read/Write

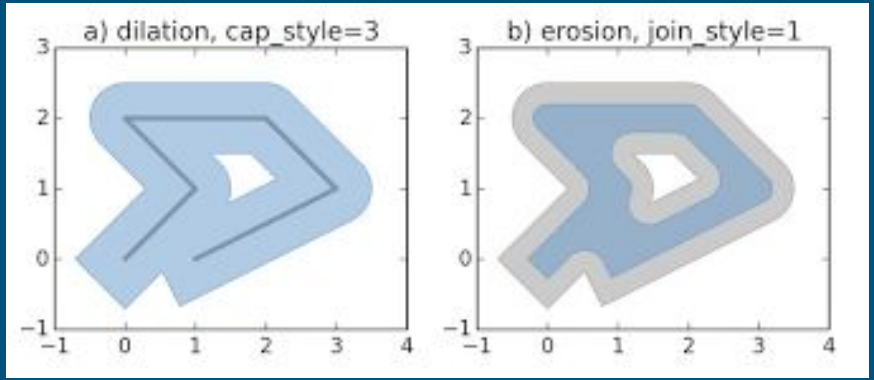
Fiona

Fiona reads and writes spatial data files

Manipulate

Shapely

Manipulation and analysis of geometric objects in the Cartesian plane.



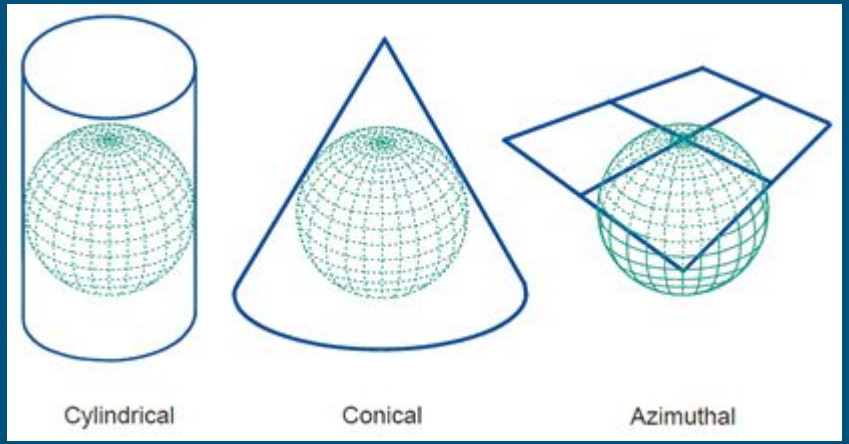
Manipulate

```
from shapely.geometry import Point
from math import pi

radius = 10.0
approximate_circle = Point(0.0, 0.0).buffer(radius, resolution=256)

approximate_circle.area 314.15729403670866

pi * radius ** 2 314.1592653589793
```

Project

Pyproj

Performs cartographic transformations and geodetic computations.



Project

```
from functools import partial
import pyproj
from shapely.ops import transform
from shapely.geometry import Point


# g = partial(f, a, b) => g(x, y) = f(a, b, x, y)

project = partial(
    pyproj.transform,
    pyproj.Proj(init='epsg:4326'),  # GPS
    pyproj.Proj(init='epsg:3857'))  # Webmercator

bordeaux_center = Point(-0.5667, 44.8333)  # Longitude, Latitude
projected_bordeaux_center = transform(project, bordeaux_center)

# In degrees
str(bordeaux_center) 'POINT (-0.5667 44.8333)'
# In meters
str(projected_bordeaux_center) 'POINT (-63084.75543254755 5595316.049417061)'
```

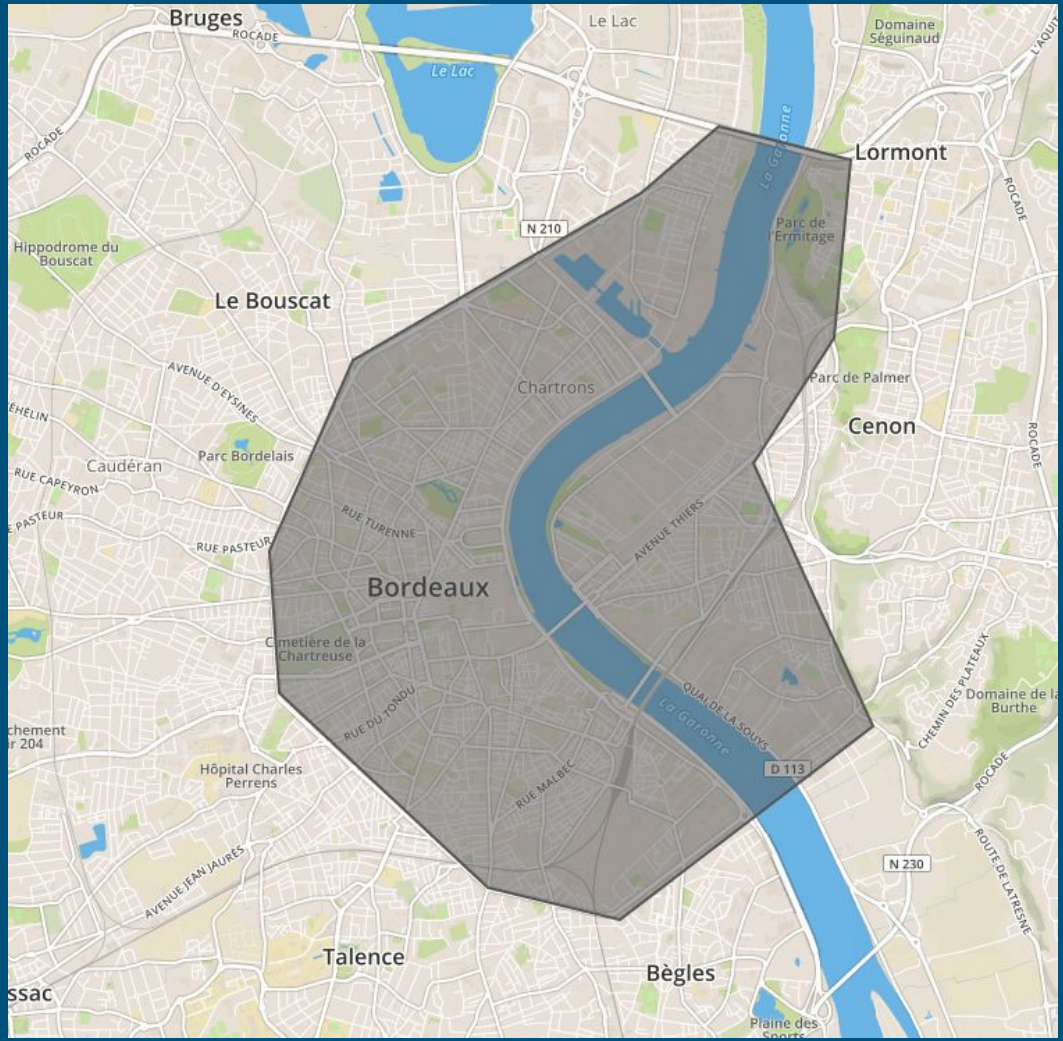
Analyze

Geopandas

GeoPandas is an open source project to make working with geospatial data in python easier

Analyze

Area?



Analyze

```
import geopandas as gpd ✓

GEOJSON_PATH = '../.data/bordeaux.geojson' ✓

gdf = gpd.read_file(GEOJSON_PATH) ✓

gdf.crs {'init': 'epsg:4326'}

# In angular degrees

float(gdf.geometry.area) 0.0033993639905879865


# Webmercator area (in Km^2)

float(gdf.to_crs({'init': 'epsg:3857'}).area / (1000 * 1000)) 59.413413009358756


# Equal Area Cylindrical (in Km^2)


float(gdf.to_crs({'proj': 'cea'}).area / (1000 * 1000)) 29.865862037143522
```

Visualize

Geoplot

geoplot is a high-level Python geospatial plotting library. It's an extension to cartopy and matplotlib which makes mapping easy: like seaborn for geospatial

Live Demo

Key points

- Use open source tools whenever possible
- Build on previous work - Don’t reinvent the wheel!
- Make your workflow reproducible **=>** scale

Questions?

Thanks for your attention

Qucit is hiring!

Resources

- GIS Wikipedia article
- GIS Python libraries:
  - https://pypi.python.org/pypi/Fiona
  - https://pypi.python.org/pypi/Shapely
  - https://github.com/jswhit/pyproj
  - https://github.com/geopandas/geopandas
  - https://github.com/ResidentMario/geoplot
- A nice GIS notebook analyzing Airbnb data
- A video presentation of the Geospatial Python tools
- A StackOverflow answer with an exhaustive list of Python GIS libraries.
- Different world maps: http://www.naturalearthdata.com/
- OSM data: http://openstreetmapdata.com/data/land-polygons
- Live demo code is here