

Spatial Thinking with Python

Sangarshanan



{{ ME_IRL }}

Sangarshanan

Recently graduated from VIT Vellore

Working at Grofers

Obsessed Memes and Astrophysics



[Github](#)



[Twitter](#)

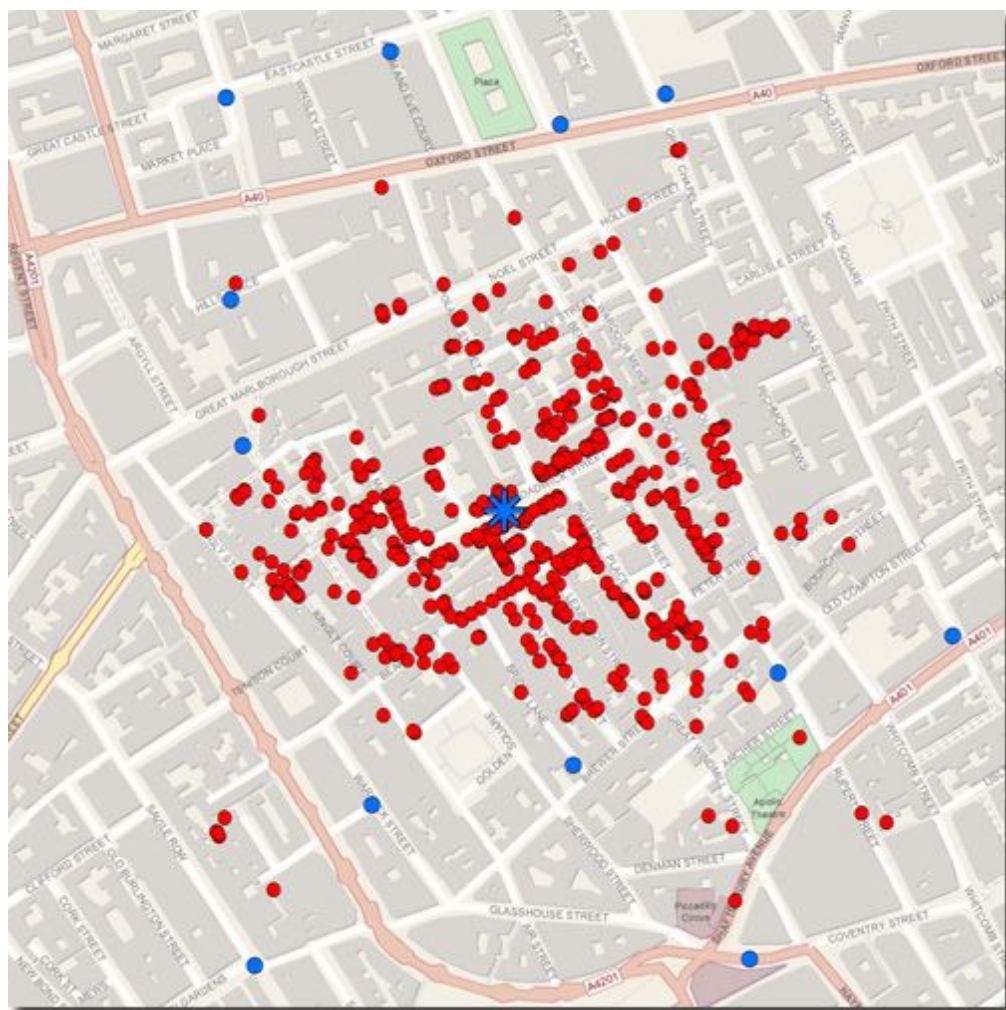


**That one person
who appreciates
memes**

The geostory begins

When you realize that John Snow was basically the first person to effectively make use of geospatial data





Location Intelligence

Intelligence you acquire from spatial thinking

When you realize that spatial
data doesn't get the attention it
deserves



WHAT IS GEOSPATIAL DATA ?

Geospatial data is any data with a geographic component that can be layered onto a map

**Types of these
geographic components ?**

Vector

Points / Geometries
/Shapes

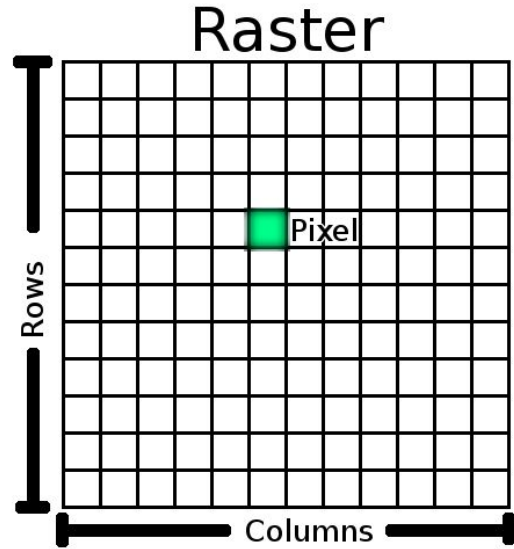
eg: Shapefiles, GEOJSON

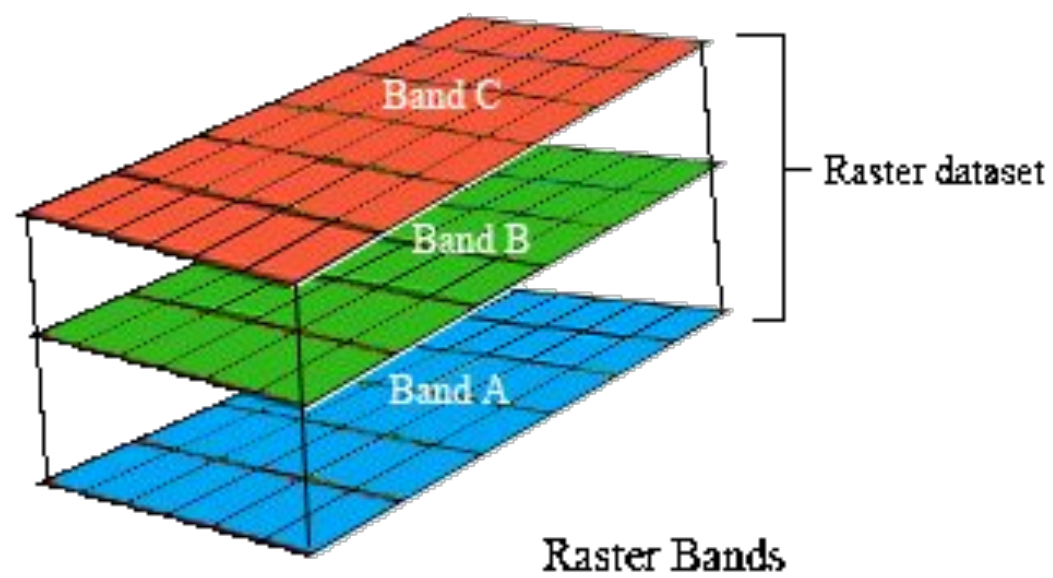
Raster

Pixels with or without
spatial metadata

eg: GEOTIFF, TIFF

Raster Data



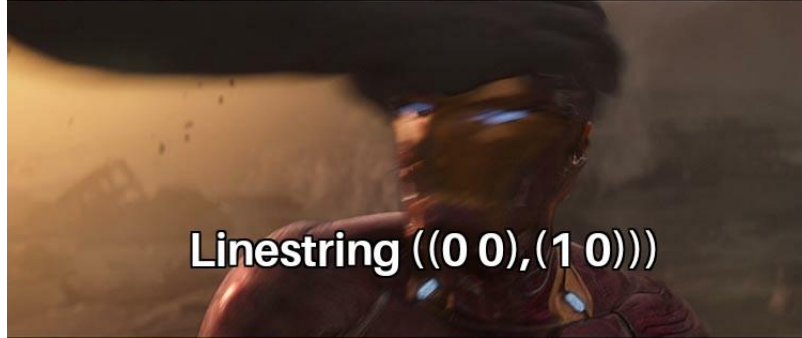
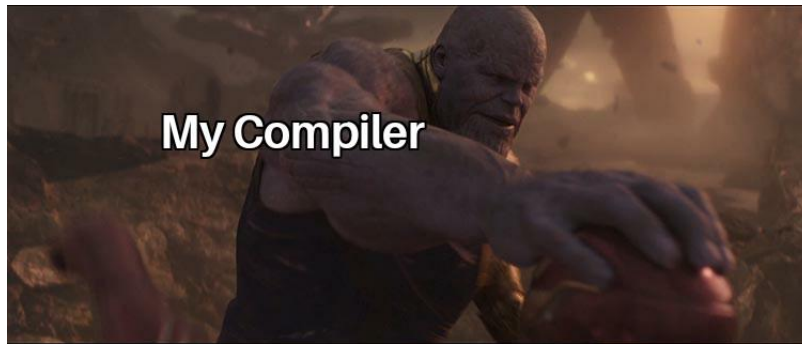


We can get our Raster's degree now



We can get our Raster's degree now

- Monitoring and predicting natural phenomena like Hurricanes, Forest fires etc
- Do tons of cool analysis over time like calculating the Affluence of the area, Extent of vegetation, Types of buildings, Road network, Population, Nightlights.

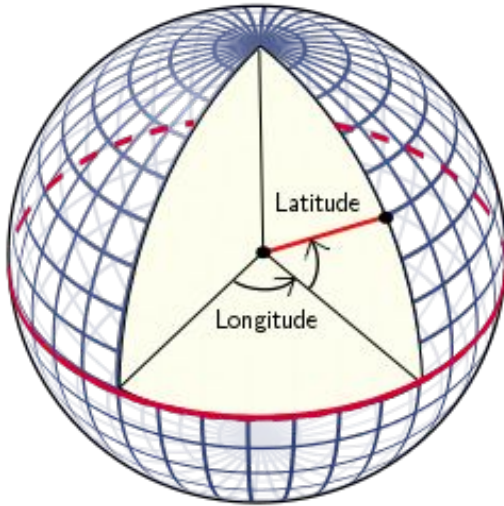


Well Known text

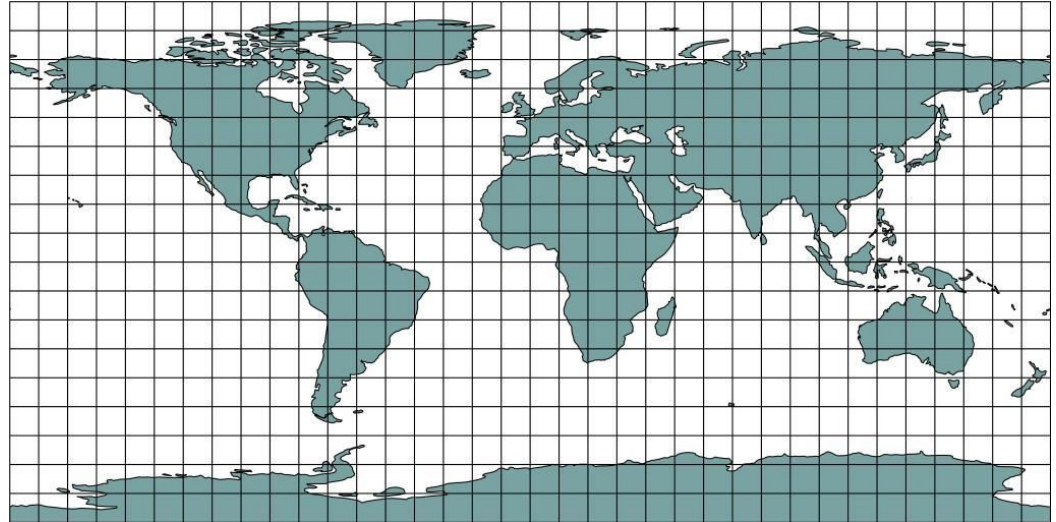
Well known binary

Spatial Reference System

Geographic coordinate systems



Projected coordinate systems



Sentinel 2

Geotiff

KMZ

KML

Carto

SHP

Geojson

SHP

KML

GPX

ESRI

GeoRSS

GML

CSV

GSC

Read/ Write/ Analyze

Number of geospatial
python libraries jumping
from 10 to 1000



GDAL / OGR

Geospatial Data Abstraction
Library

YEAR 1998



154 raster and 93 vector
geospatial data formats

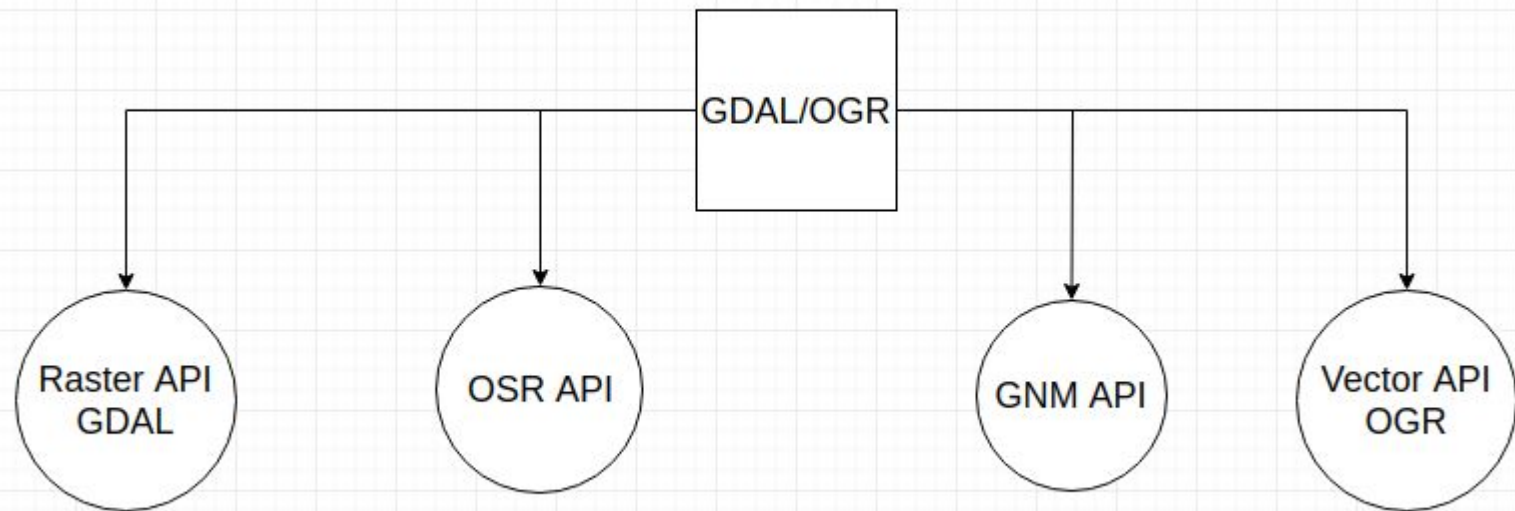
ogr2ogr can convert data in PostGIS to KML

```
$ ogr2ogr -f "KML" \  
neighborhoods.kml \  
PG:"host=myhost user=myloginname dbname=mydbname password=mypassword" \  
-sql \  
"select gid, name, the_geom from neighborhoods" \  

```

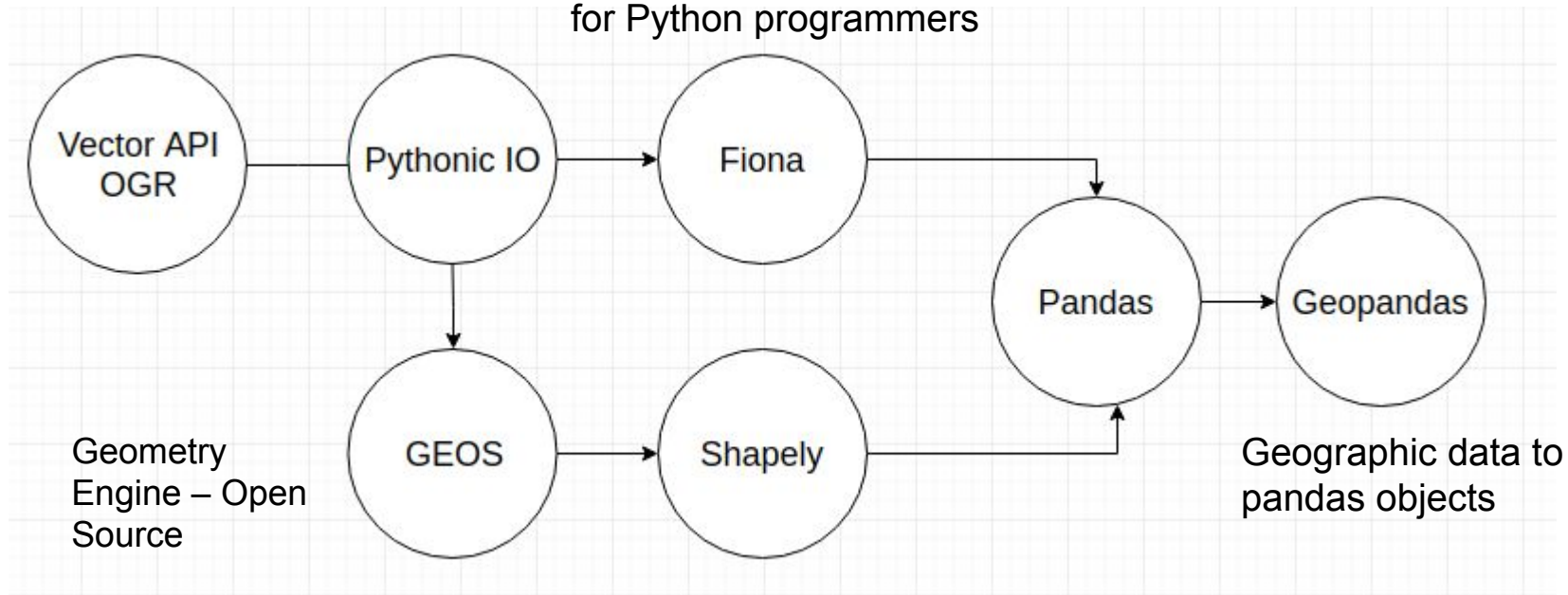
ogrinfo lists information about the data

```
$ ogrinfo data/EXAMPLE.NTF
```

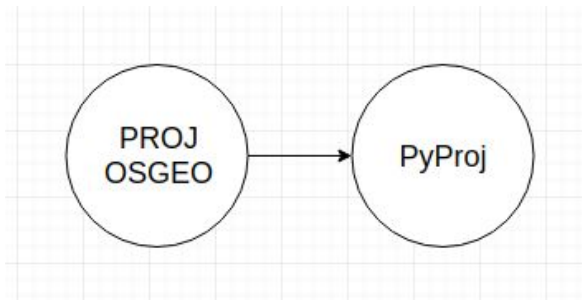


Vector Data

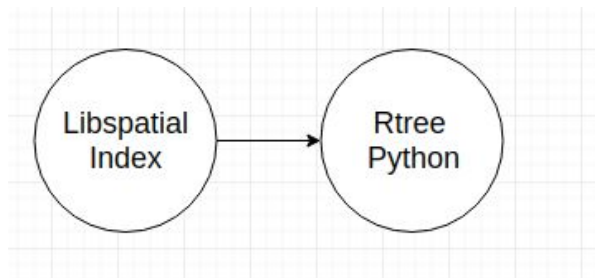
OGR's neat and nimble API
for Python programmers



OSR Projections / Rtree



Pythonic cartographic projections and coordinate transformations library



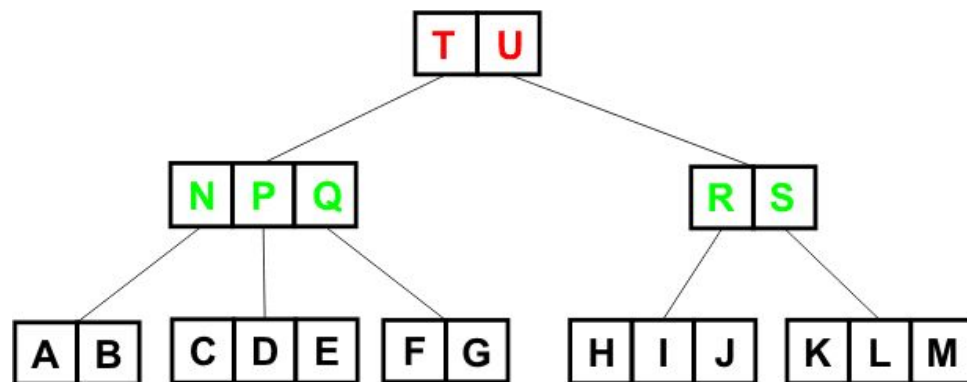
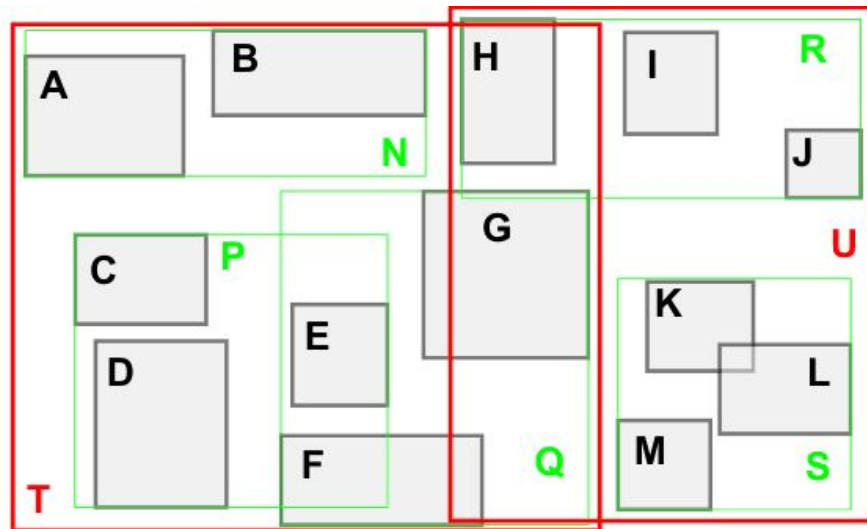
Advanced spatial indexing features

Fiona ,Shapely, Geopandas

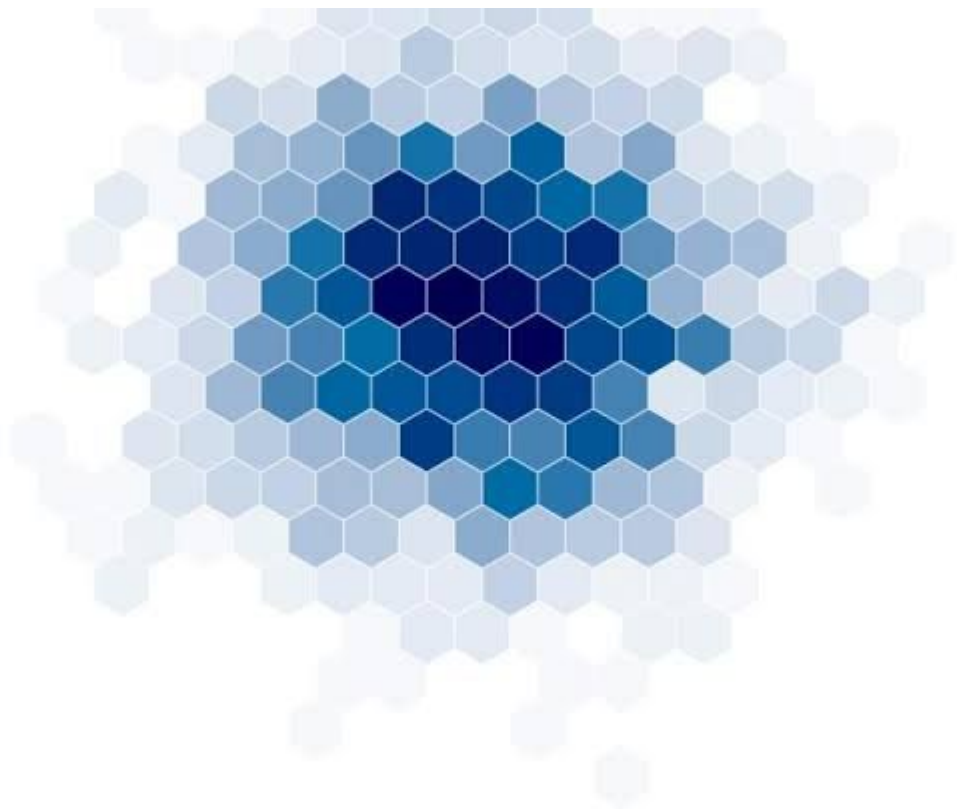
Indexing Geospatial Data

Suppose you want to find all the geospatial points in a given radius
Are you gonna iterate? HELL NAHHH

Use Spatial indices provided by the spatial extensions of traditional databases like postgres (Postgis uses Rtrees)



Hexagonal grid indexing (Uber H3)



```
SELECT superhero.name  
FROM city, superhero  
WHERE ST_Contains(city.geom, superhero.geom)  
AND city.name = 'Gotham';
```

Raster Data



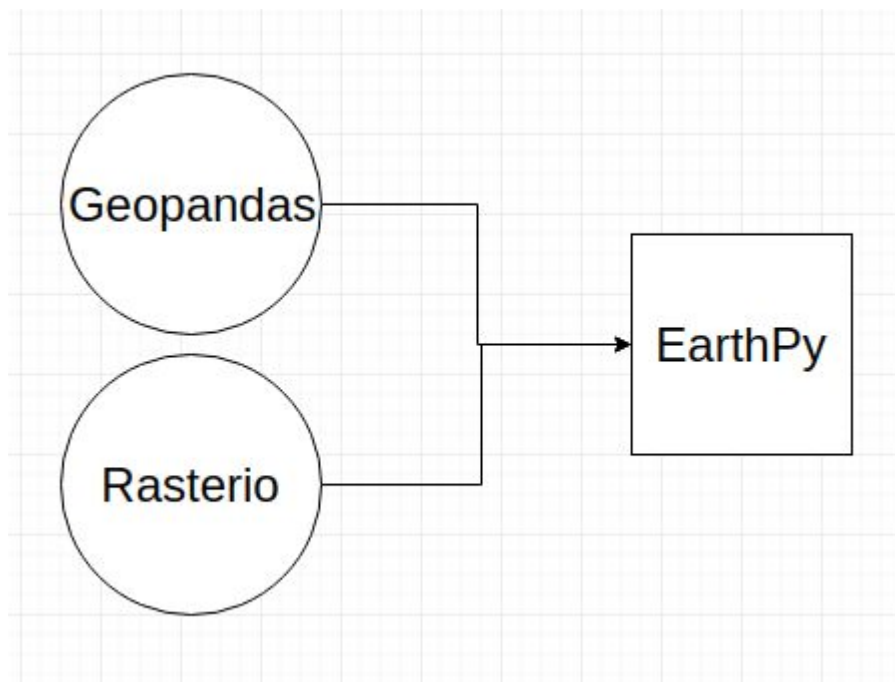
Rasterio reads and writes raster formats and provides a Python API based on N-D arrays.



There are also packages than combine these existing packages that are actually wrappers on already existing packages (like inception)

Vector data

Raster data



OSMNX (OSM + Networkx)

Work with road network data from osm
using networkx

Analyze and visualize street networks,
routing , travel times etc

Visualize

When you plot and visualize all
your spatial layers on a basemap





You need to know javascript to create
such cool maps

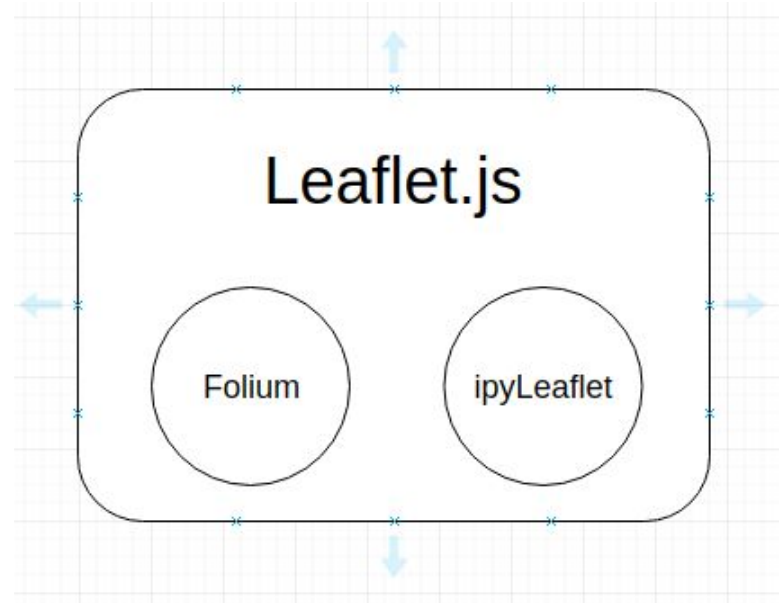


Just Plot em

Matplotlib (for everything)

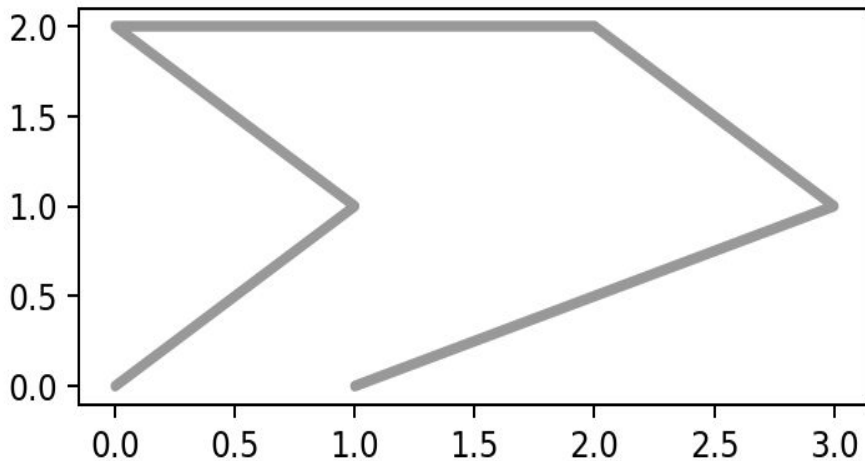
Leaflet / Openlayers / Mapbox

Plotly + Mapbox / Mapboxgl



Geometry + Matplotlib = Descartes

Enables plotting of
shapely geometries as
matplotlib paths/ patches.




```
In [109]: import folium
import pandas as pd

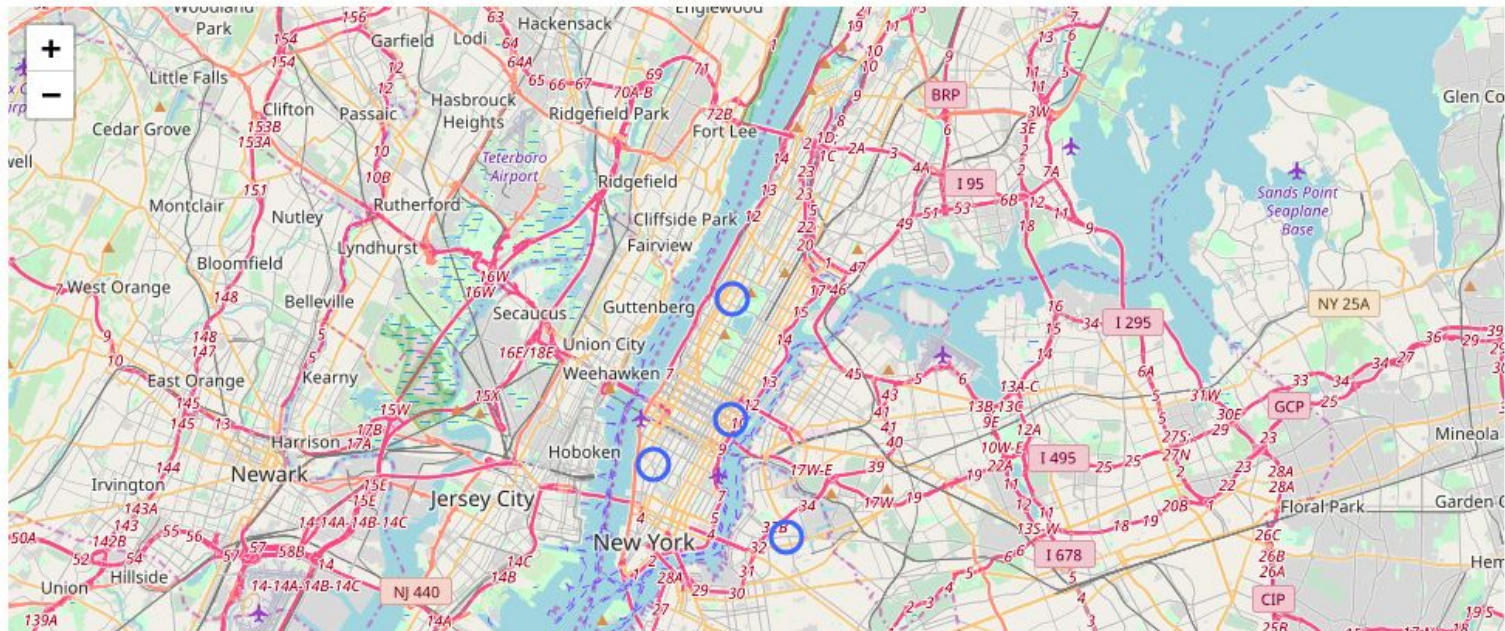
train_df = pd.DataFrame({'Latitude': [40.7145, 40.7947, 40.7388, 40.7539],
                          'Longitude': [-73.9425, -73.9667, -74.0018, -73.9677] })

map_osm = folium.Map(location=[40.742, -73.956], zoom_start=11)

train_df.apply(lambda row: folium.CircleMarker(location=[row["Latitude"], row["Longitude"]],
                                                radius=10)
               .add_to(map_osm), axis=1)

map_osm
```

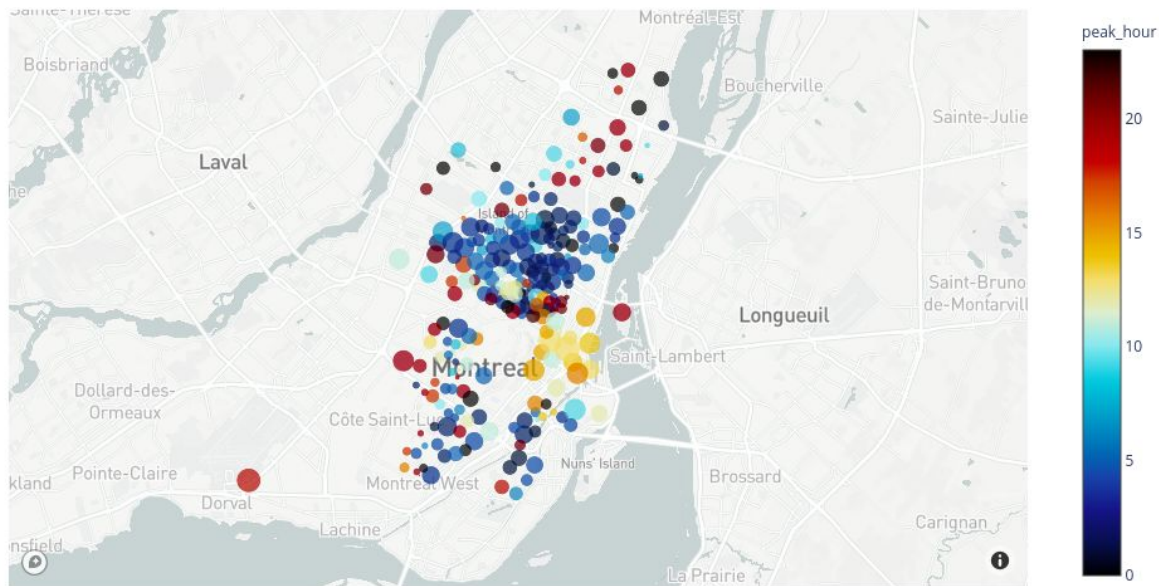
Out[109]:



```

import plotly.express as px
px.set_mapbox_access_token(open(".mapbox_token").read())
carshare = px.data.carshare()
fig = px.scatter_mapbox(carshare, lat="centroid_lat", lon="centroid_lon", color="peak_hour", size="car_hours",
                        color_continuous_scale=px.colors.cyclical.IceFire, size_max=15, zoom=10)
fig.show()

```



It's Code Time



Spatial Data Science

PySAL: Python Spatial Analysis Library

Spatially constrained Clustering

Spatio-temporal data analysis

Spatial regression and Statistical modeling

What's the POINT(12.22,73.32) ?

Marketing

GEO - Marketing

PennyWISE Decision Making



Burger King ✓
@BurgerKing

how do you order a Whopper sandwich for a penny
“at” McDonald’s? here’s how.

[#WhopperDetour burgerking.app.link/Dh8MTXeemS](https://burgerking.app.link/Dh8MTXeemS)



12:30 AM · Dec 5, 2018 · [Twitter Media Studio](#)

2.1K Retweets **6.5K** Likes

GEO - Surge

That was UnUbercool

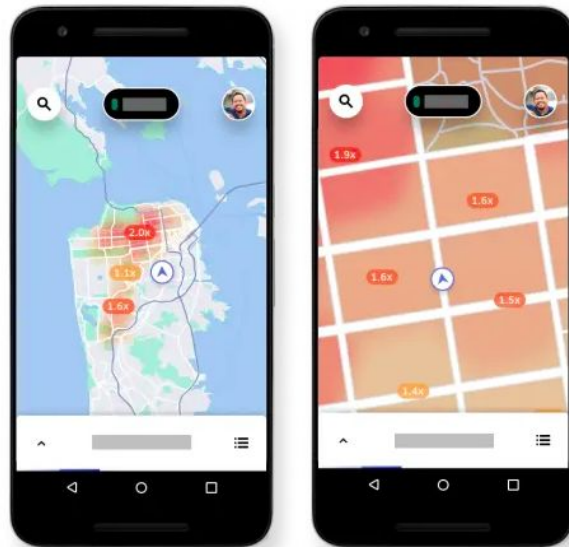
How to identify surge in the app

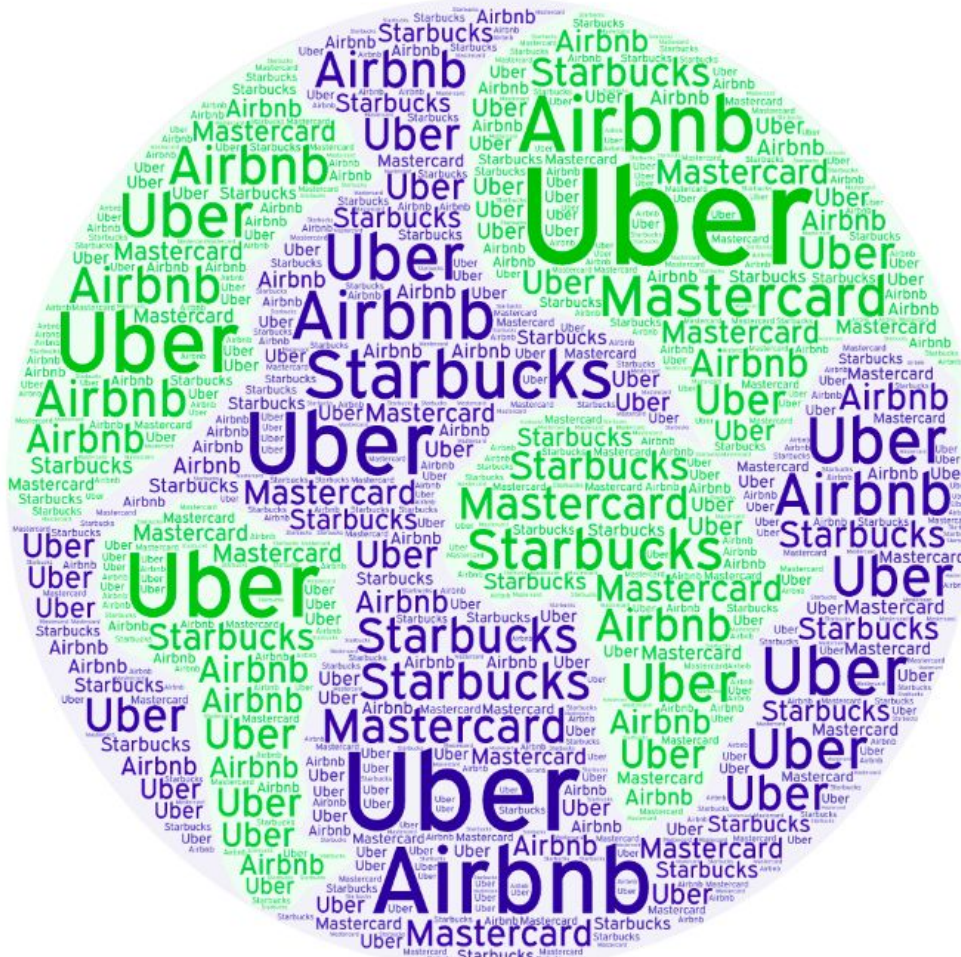
When demand increases in a specific area, that neighborhood will change color. You can zoom into colored areas of your app's city map to see current surge pricing.

The colored areas of the map will range from light orange to dark red. Light orange areas represent small multipliers while dark red areas indicate large multipliers.

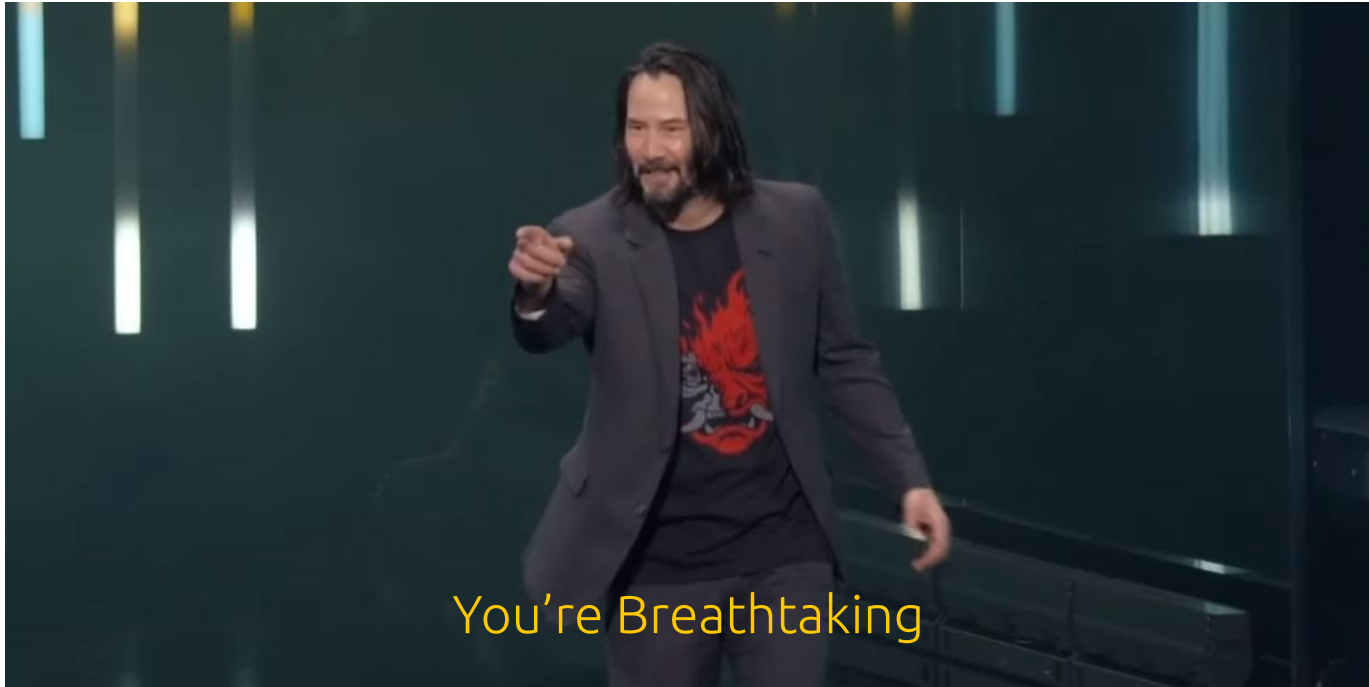
You can see the amount of surge pricing that will apply to a potential trip when you are deciding whether or not to accept a trip in the driver app.

How are surge prices calculated?





Hey You.... Yes you :)



Blogs I stole content from

- <https://blog.mapbox.com/a-dive-into-spatial-search-algorithms-ebd0c5e39d2a>
- <https://github.com/pcjericks/py-gdalogr-cookbook>
- <https://medium.com/locale-ai>
- <https://geoffboeing.com/>
- <https://towardsdatascience.com/geospatial-indexing-with-ubers-h3-766399b690c>
- <https://medium.com/@chrieke/essential-geospatial-python-libraries-5d82fcc38731>
- <https://pysal.readthedocs.io/en/latest/>