



An introduction to geodata in Python

Margot Ridderikhoff & Amber Mulder

+

25-10-2023

RIO MECAYA · Putumayo, Colombia · September 2020



**HURRICANE IAN'S AFTERMATH** · Gasparilla Sound, Florida, USA · September 30, 2022

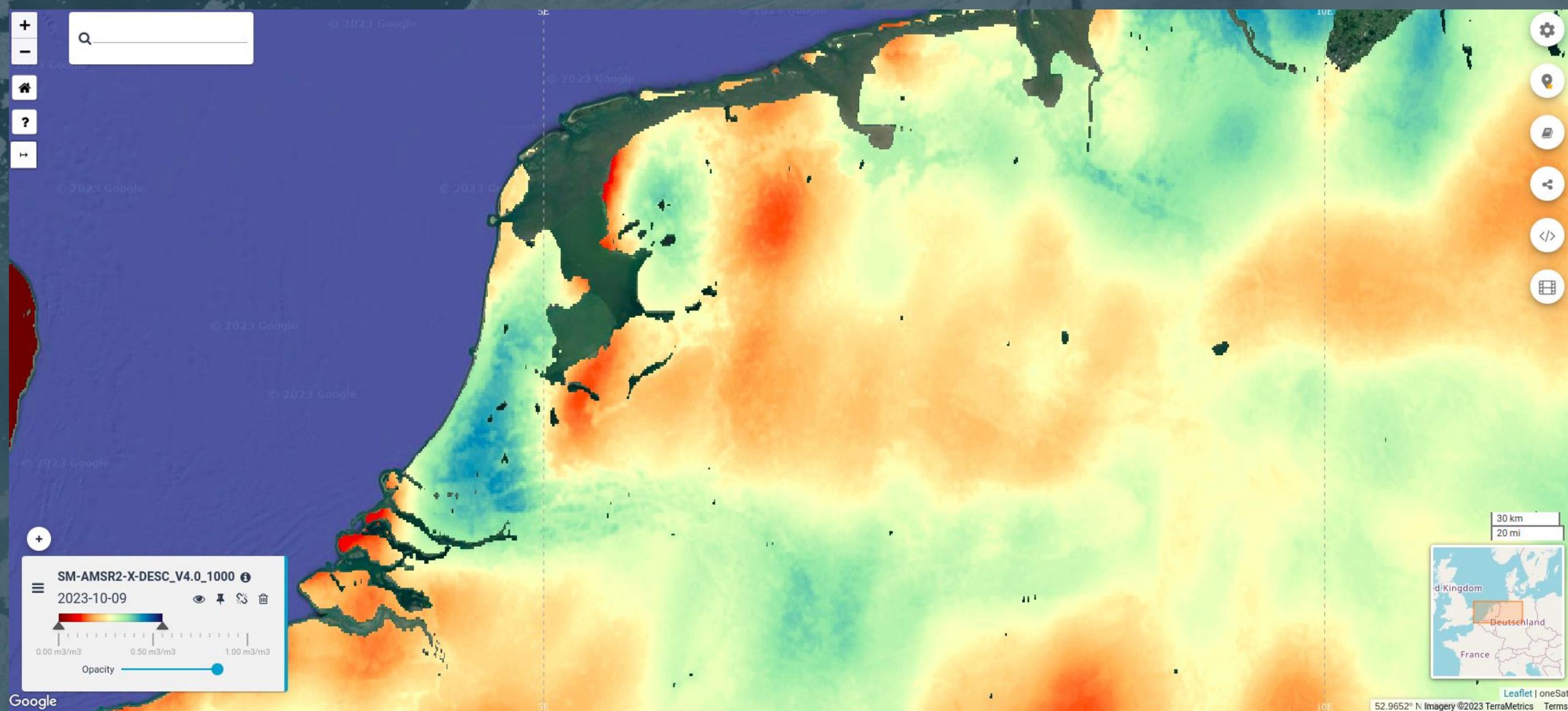


Margot Ridderikhoff  
Software engineer @ Planet



Amber Mulder  
Software engineer @ Planet

# Soil Water Content



ICE SHELF COLLAPSE · Ellesmere Island, Canada · July 31, 2020



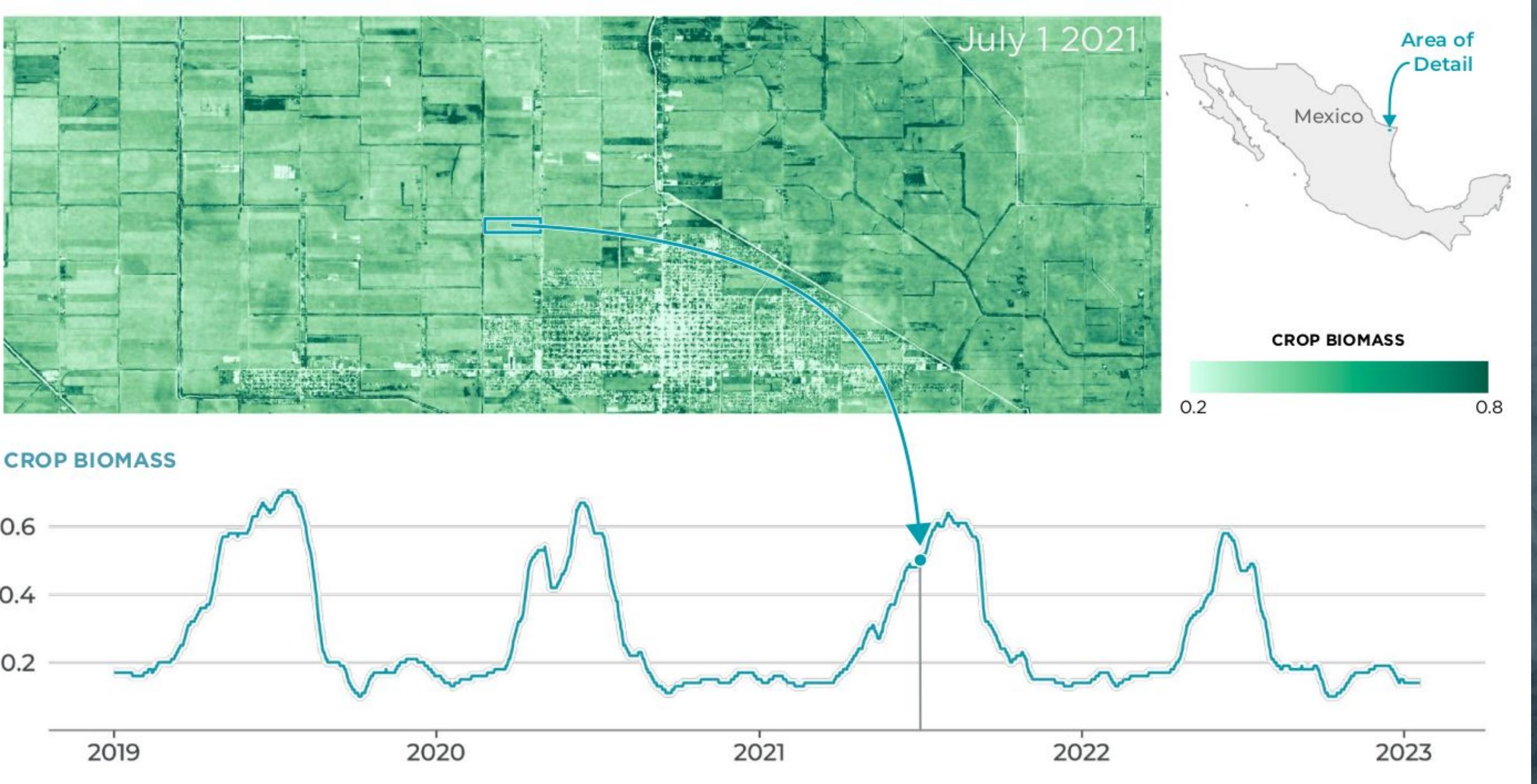


Margot Ridderikhoff  
Software engineer @ Planet



Amber Mulder  
Software engineer @ Planet

# Crop Biomass



ICE SHELF COLLAPSE · Ellesmere Island, Canada · July 31, 2020





“Using Space to Help Life on Earth”

# What will we do today?



# What will we do today?

- Geospatial data theory
- Python libraries for geodata
- Hands-on: Jupyter notebook exercises





# Case study

 planet

# Case study

Which forest patch in our study area is the most affected by the wildfire?

# Case study

Which forest patch in our study area is the most affected by the wildfire?

- **Part 1: Raster data**
  - **Satellite imagery in Python**
- Part 2: Vector data
  - Extract forest areas from land use data
- Part 3: Bringing it together
  - Locate the forest area impacted the most by the wildfire

# Case study

Which forest patch in our study area is the most affected by the wildfire?

- Part 1: Raster data
  - Satellite imagery in Python
- **Part 2: Vector data**
  - **Extract forest areas from land use data**
- Part 3: Bringing it together
  - Locate the forest area impacted the most by the wildfire

# Case study

Which forest patch in our study area is the most affected by the wildfire?

- Part 1: Raster data
  - Satellite imagery in Python
- Part 2: Vector data
  - Extract forest areas from land use data
- **Part 3: Bringing it together**
  - **Locate the forest area impacted the most by the wildfire**



# What is geospatial data?

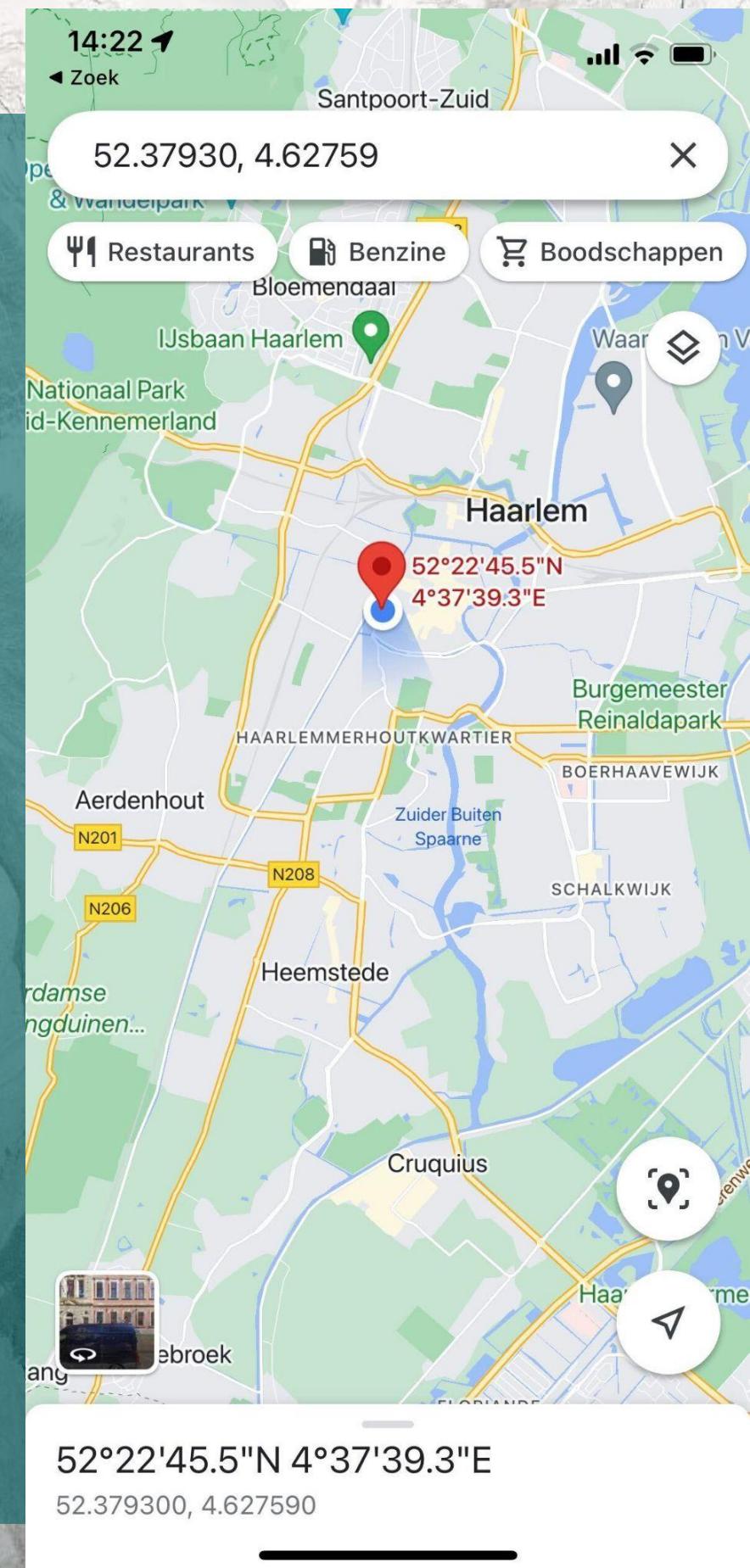
RECEDING GLACIERS • Sorata Massif, Bolivia • October 21, 2021





# What is geospatial data?

“Information that is tied to a specific location on earth”



Source: [Google Maps](#)

RECEDING GLACIERS • Sorata Massif, Bolivia • October 21, 2021





← Bogota

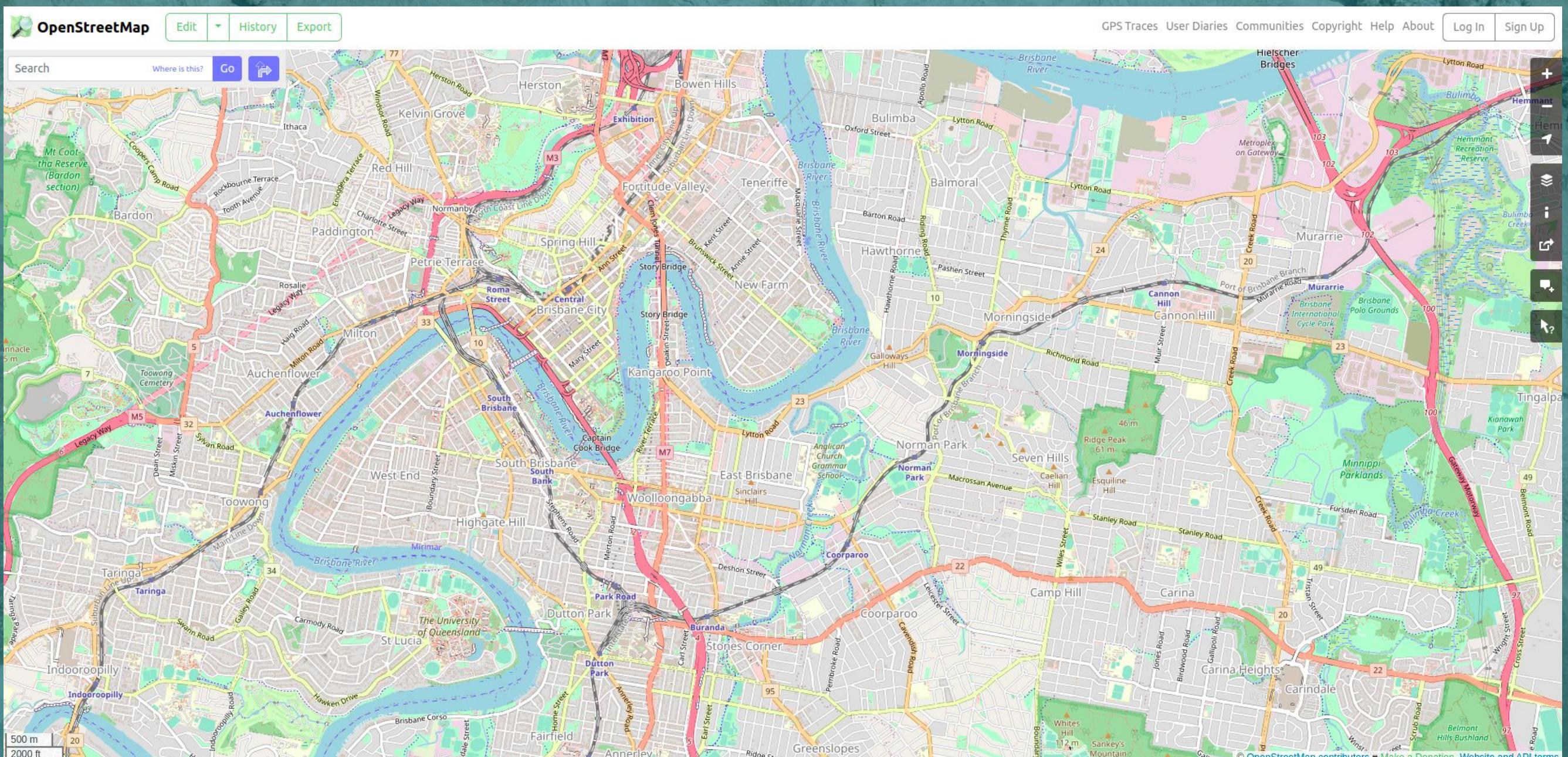
48 hours **14 days**

## 14-day forecast

The weather for the next 14 days in Bogota

	weather	min/ max	precipitation	wind	figure	
Thursday Oct 5		8 °/22 °	0.5mm		<span>7</span>	
Friday Oct 6		8 °/21 °	34.5mm		<span>5</span>	
Saturday Oct 7		10 °/21 °	12.7mm		<span>5</span>	
Sunday Oct 8		10 °/20 °	6.1mm		<span>6</span>	
Monday Oct 9		9 °/21 °	3.3mm		<span>6</span>	
Tuesday Oct 10		10 °/21 °	3.3mm		<span>6</span>	
Wednesday Oct 11		10 °/21 °	0.2mm		<span>7</span>	

Source: [Weeronline](#)



Source: [OpenStreetMap](https://www.openstreetmap.org)

RECEDING GLACIERS • Sorata Massif, Bolivia • October 21, 2021





Source: [Planet Labs](#)

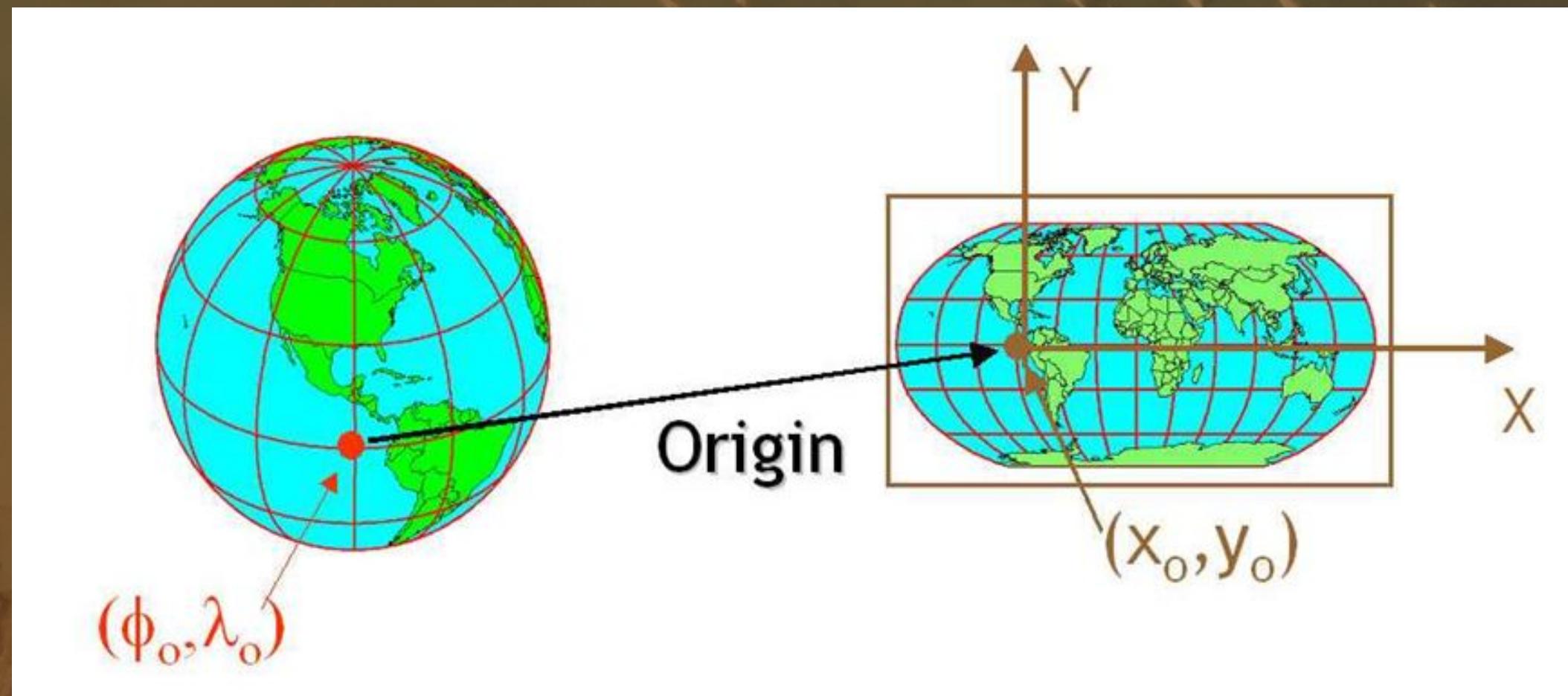
**RECEDING GLACIERS** • Sorata Massif, Bolivia • October 21, 2021





# Coordinate Reference System (CRS)

- Datum
- Projection

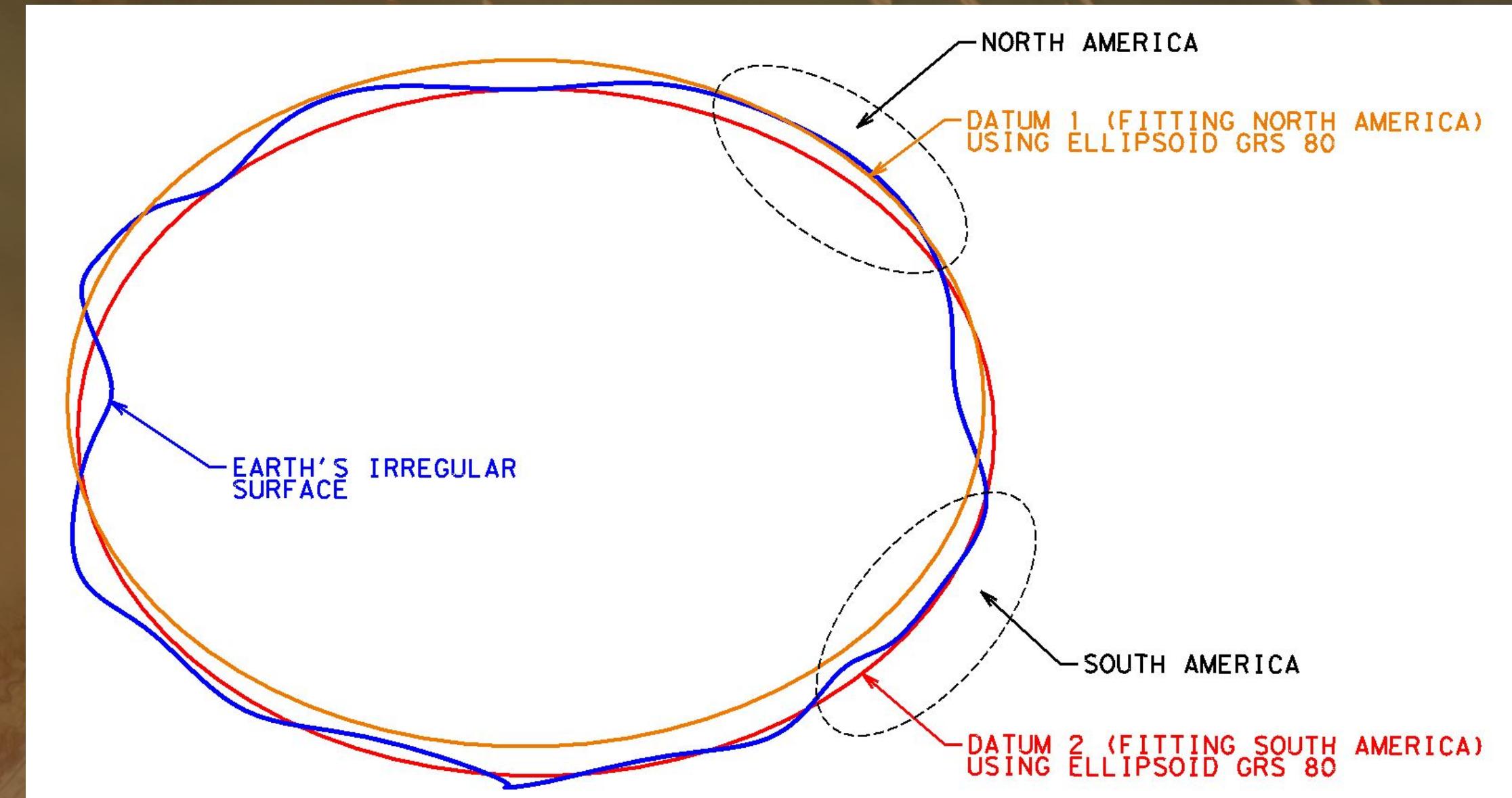


Source: [Anything Mapping](#)



# Coordinate Reference System (CRS)

- **Datum**
- Projection

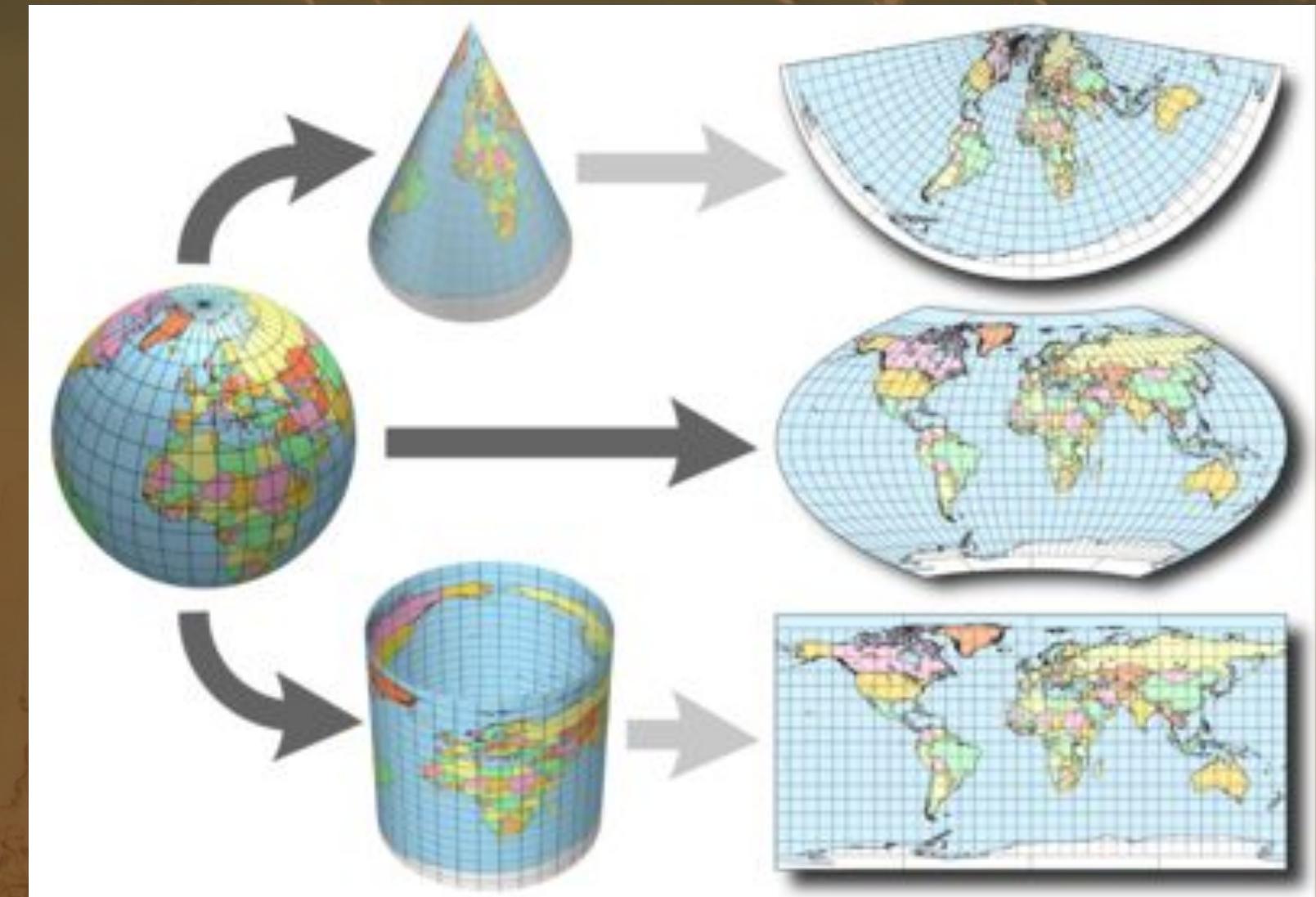


Source: [Krest Engineers](#)



# Coordinate Reference System (CRS)

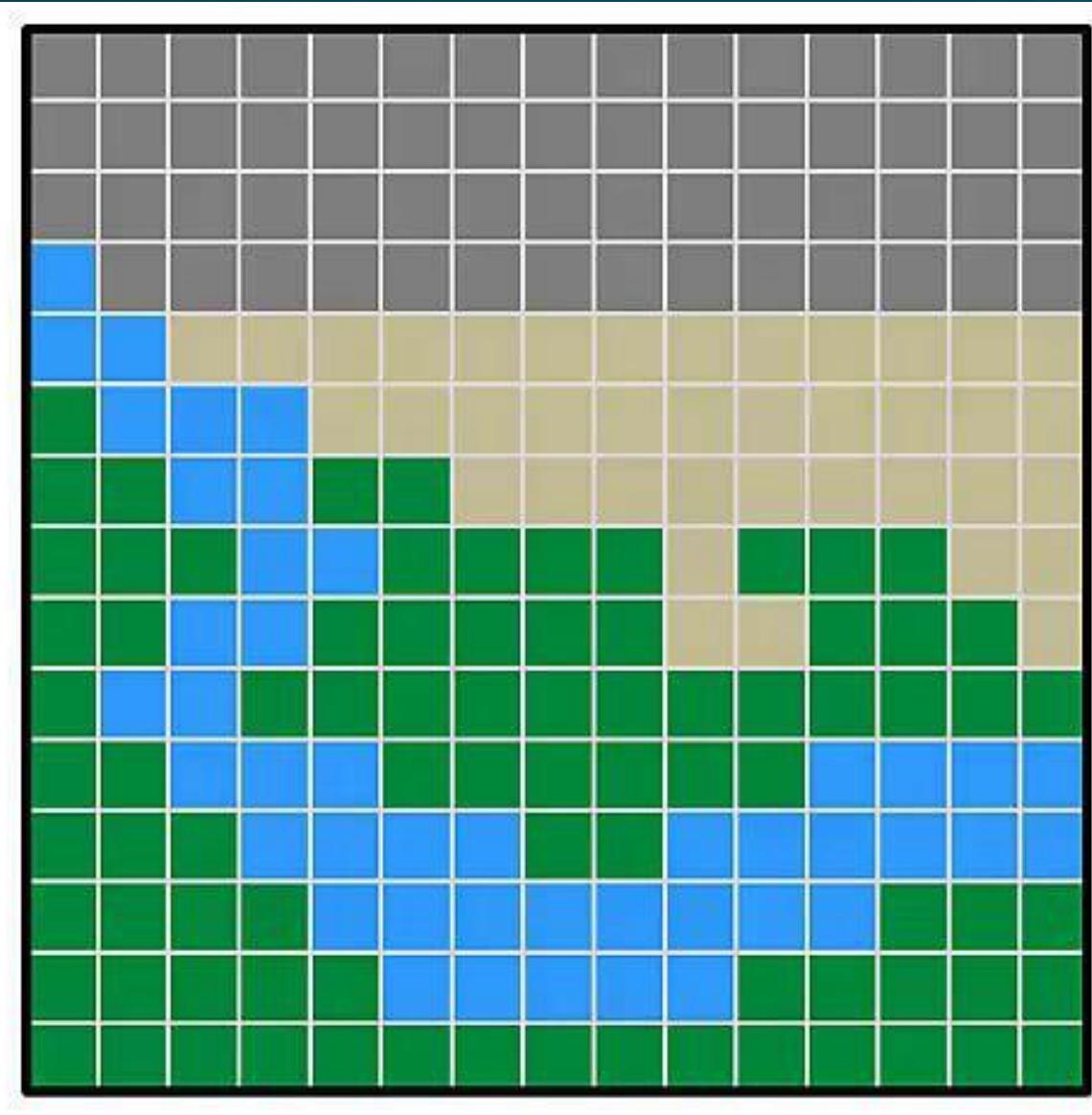
- Datum
- **Projection**



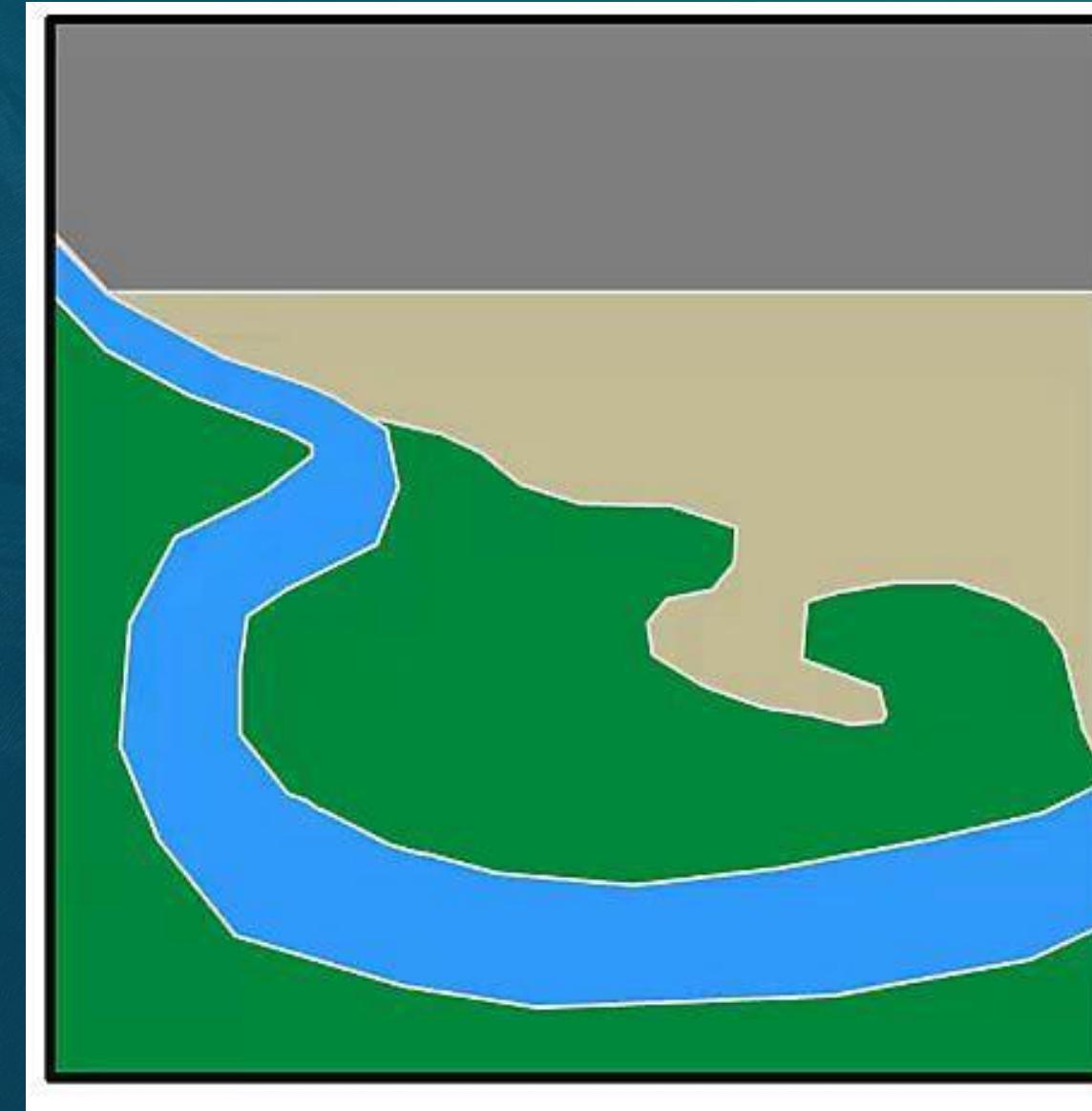
Source: [Earth Lab](#)



# Types of geodata



Raster

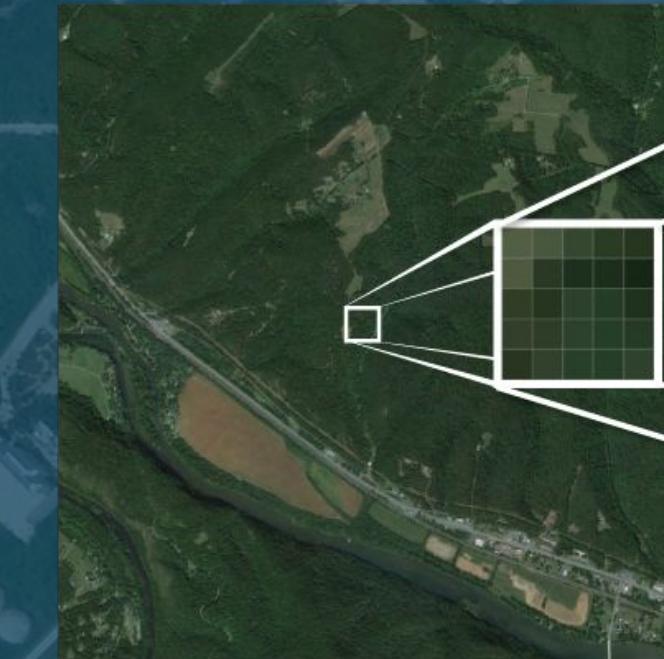


Vector

Source: [Learn Data Science](#)



# Raster data



0,0                    5,0

1m {

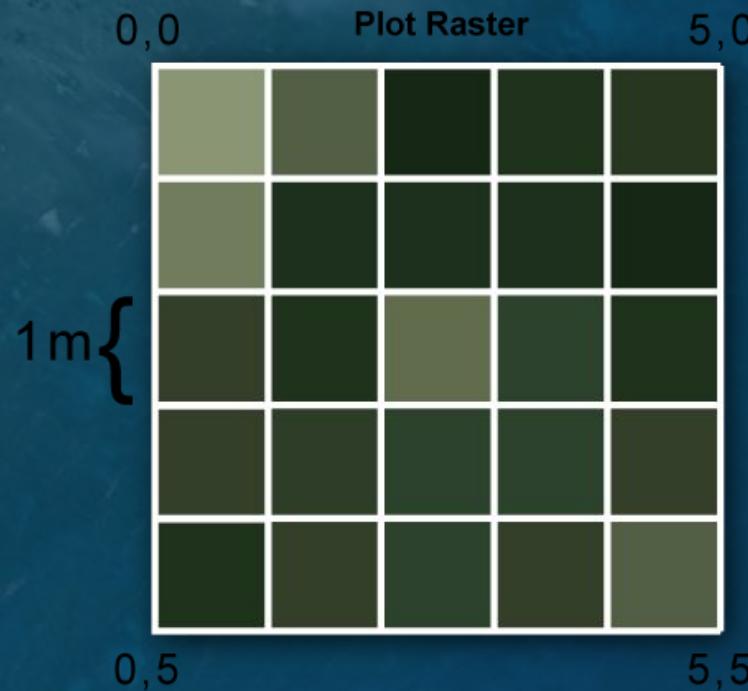
1	3	9	7	7
2	8	7	7	8
6	7	3	5	7
7	6	5	5	6
8	6	5	6	4

0,5                    5,5



Legend

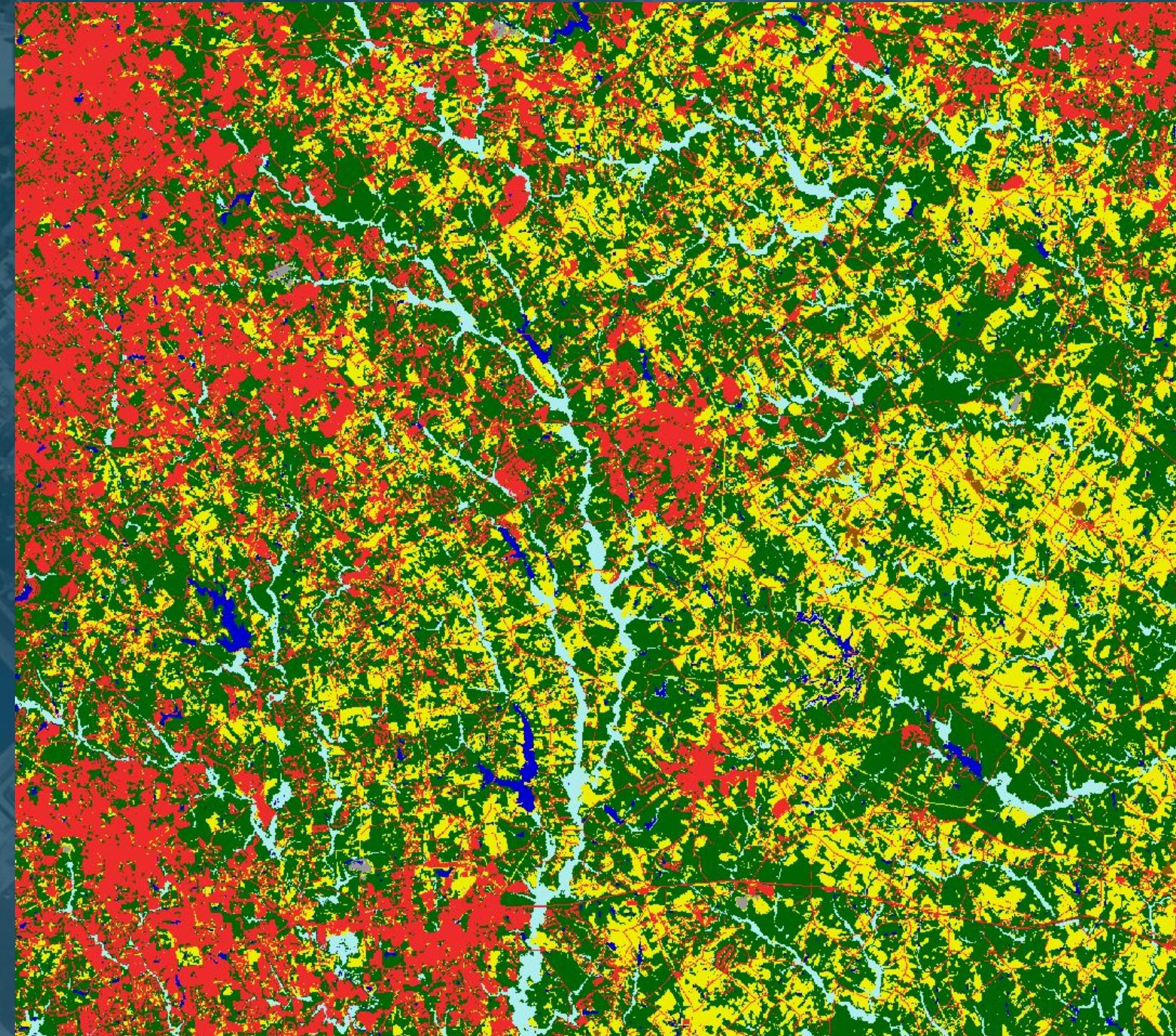
0  
1  
2  
3  
4  
5  
6  
7  
8  
9



Source: [National Ecological Observatory Network \(NEON\)](#)



## Land cover map

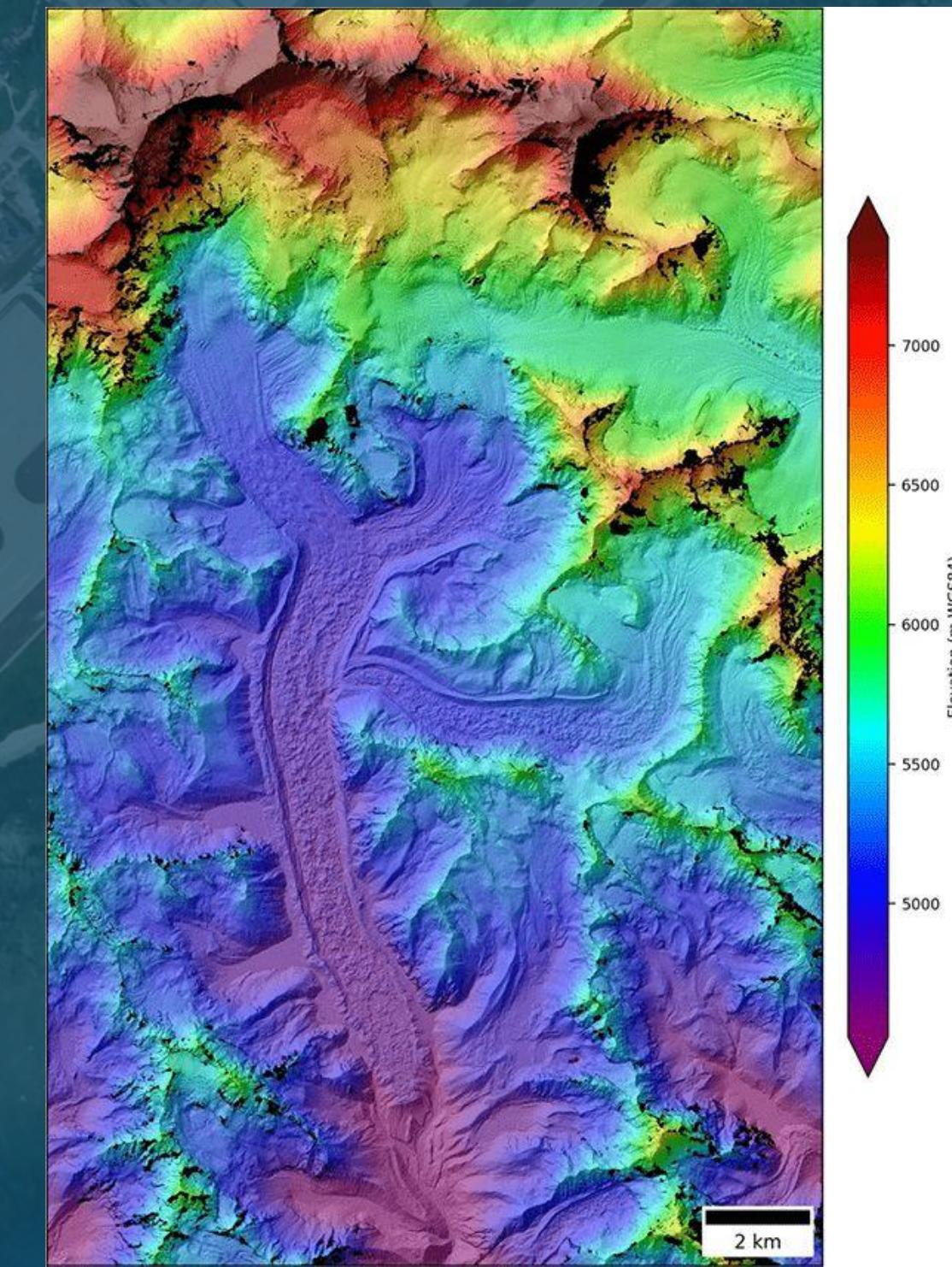


Land cover
Water
Developed
Barren
Forest
GrassShrub
Cropland
Wetland

Source: [Bookdown](#)

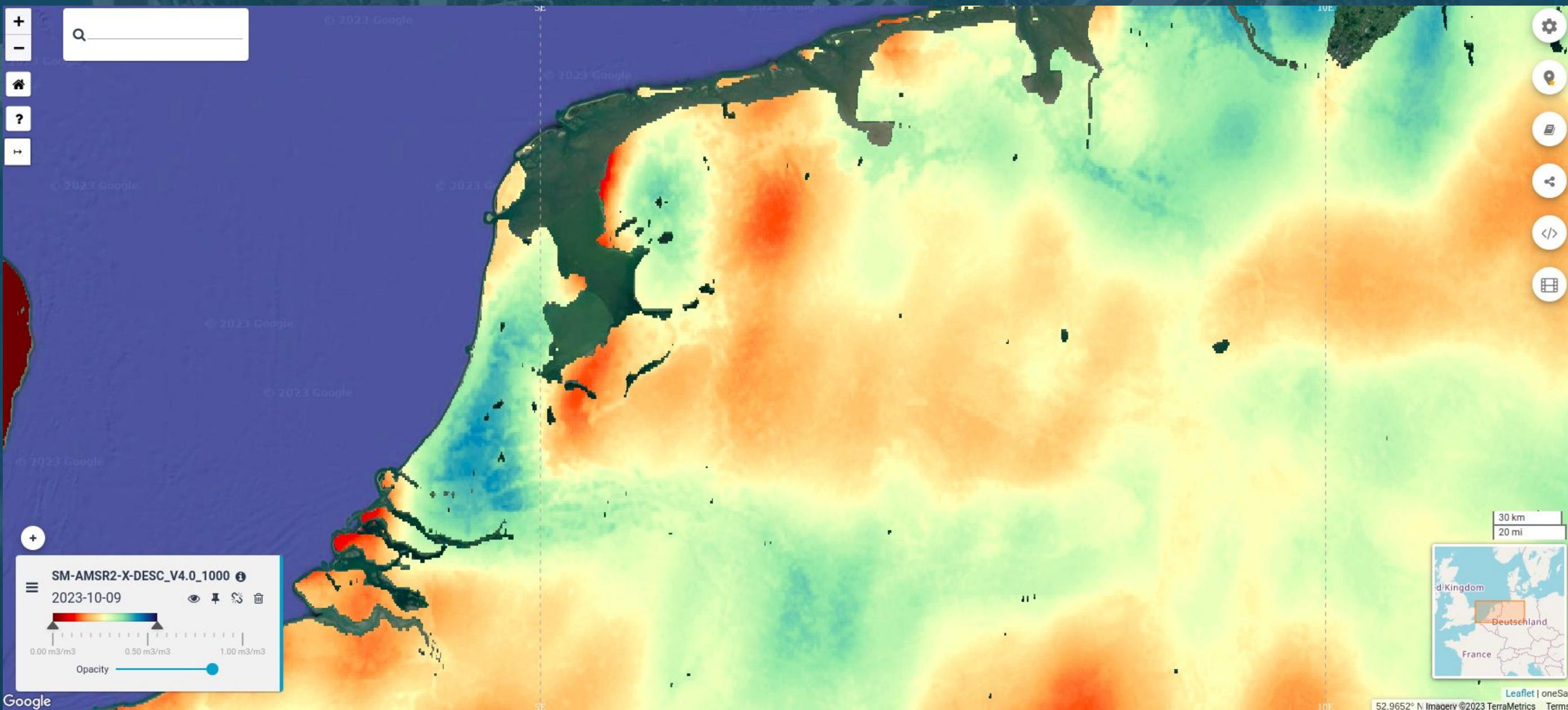


# Digital Elevation Model (DEM)



Source: [Equator Studios](#)

# Soil Water Content data



Source: [Planet Labs 2023](#)

WOODY ISLAND · South China Sea · March 28, 2018





Extent

X min, Y max

X max, Y max

X min, Y min

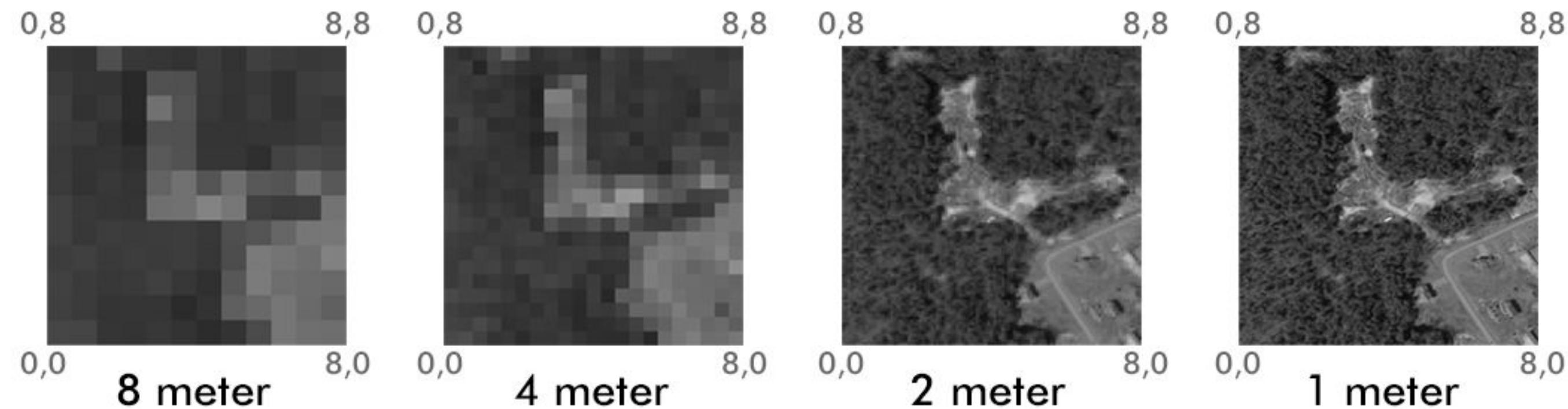
X max, Y min



Source: [National Ecological Observatory Network \(NEON\)](#)

# Resolution

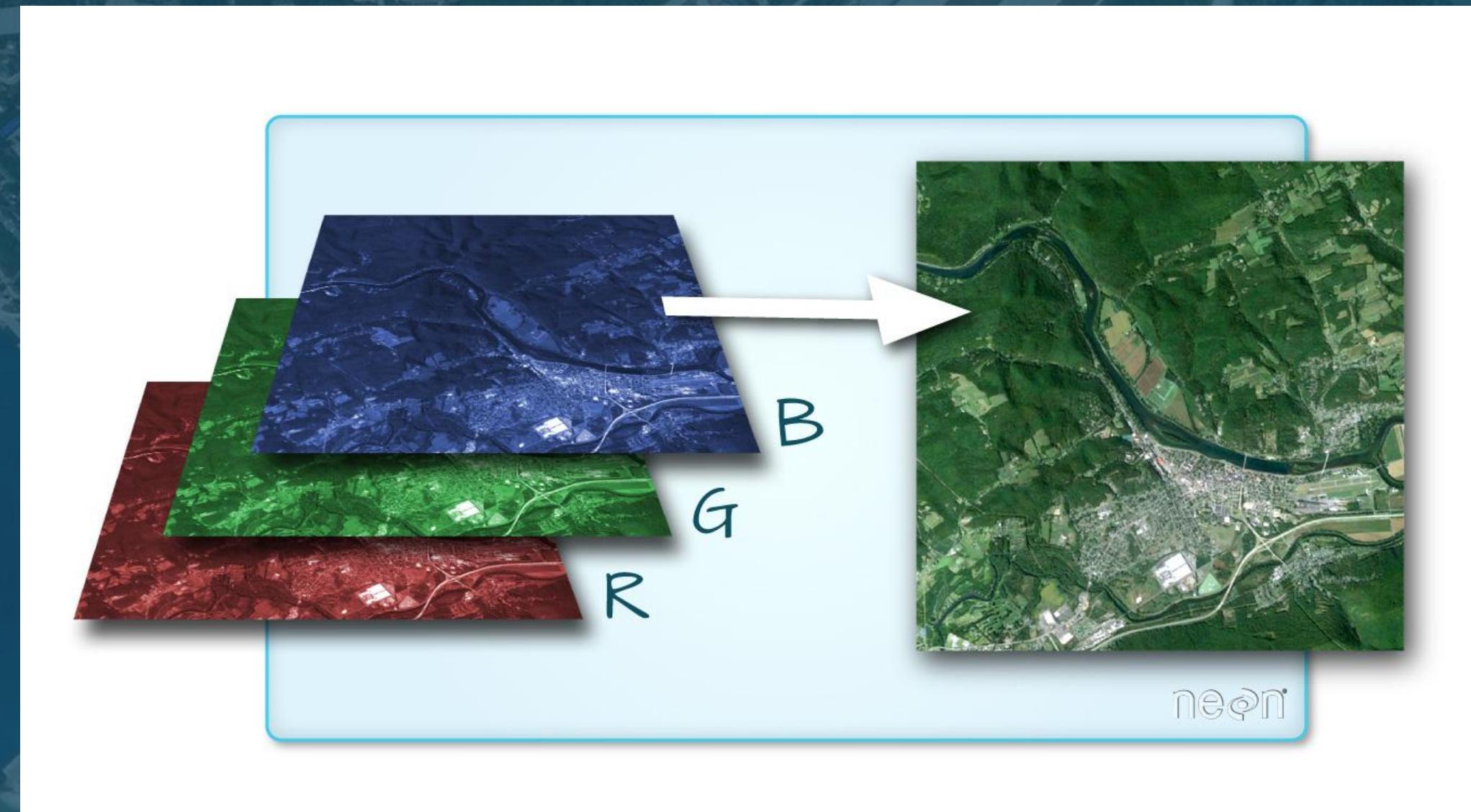
Raster over the same extent, at 4 different resolutions



Source: [National Ecological Observatory Network \(NEON\)](#)



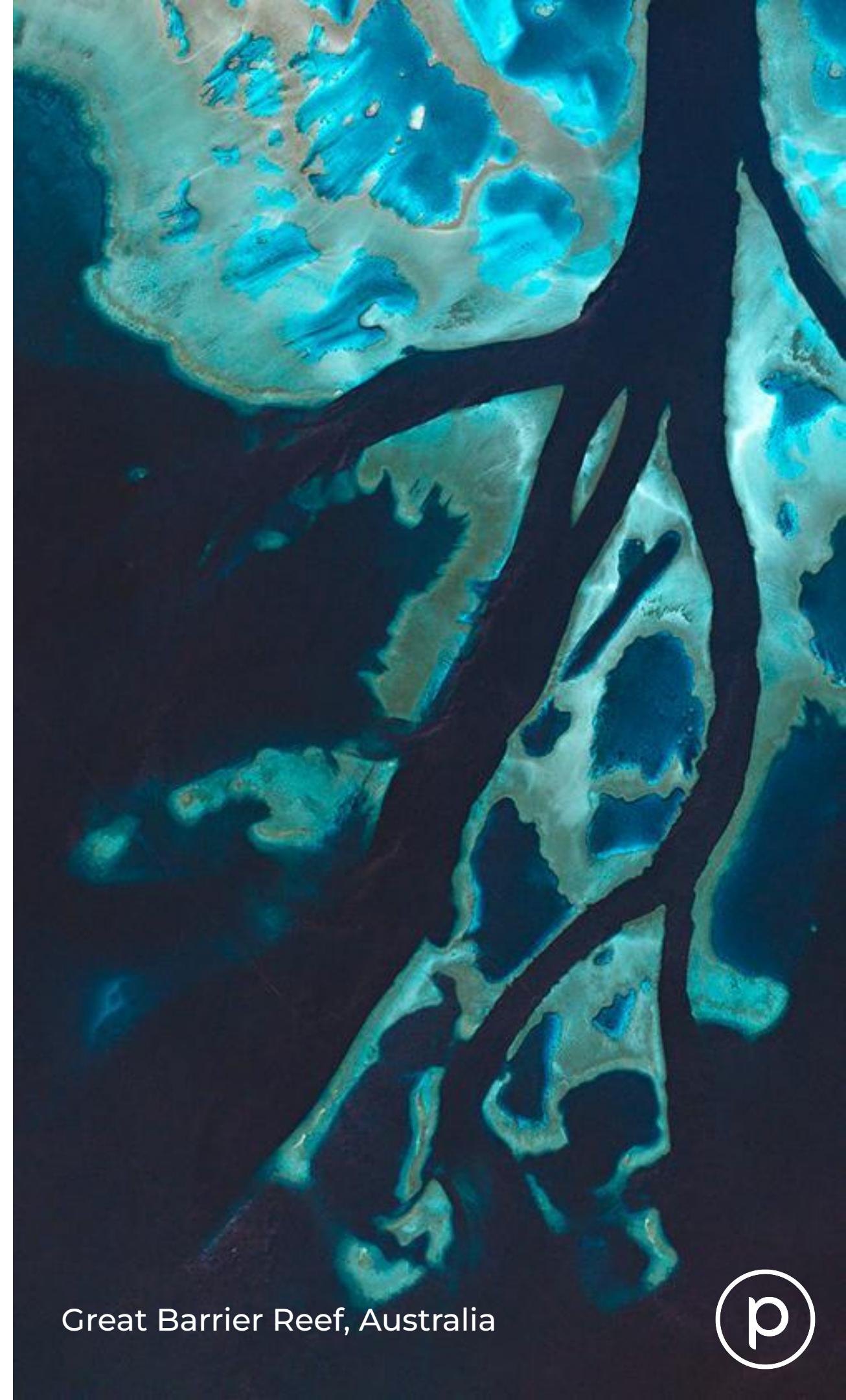
## Multi-band rasters



Source: [National Ecological Observatory Network \(NEON\)](#)

# Python libraries for raster data

- [GDAL](#)
- [Rasterio](#)
- [RSGISLib](#)
- [Rasterstats](#)
- [Rasterframes](#)



# Python libraries for raster data

- [GDAL](#)
- [Rasterio](#)
- [RSGISLib](#)
- [Rasterstats](#)
- [Rasterframes](#)

“Your Passport to Geospatial Marvels, Empowering Python Developers to Seamlessly Read, Write, and Transform Diverse Earth Data with Precision and Efficiency!”

- ChatGPT



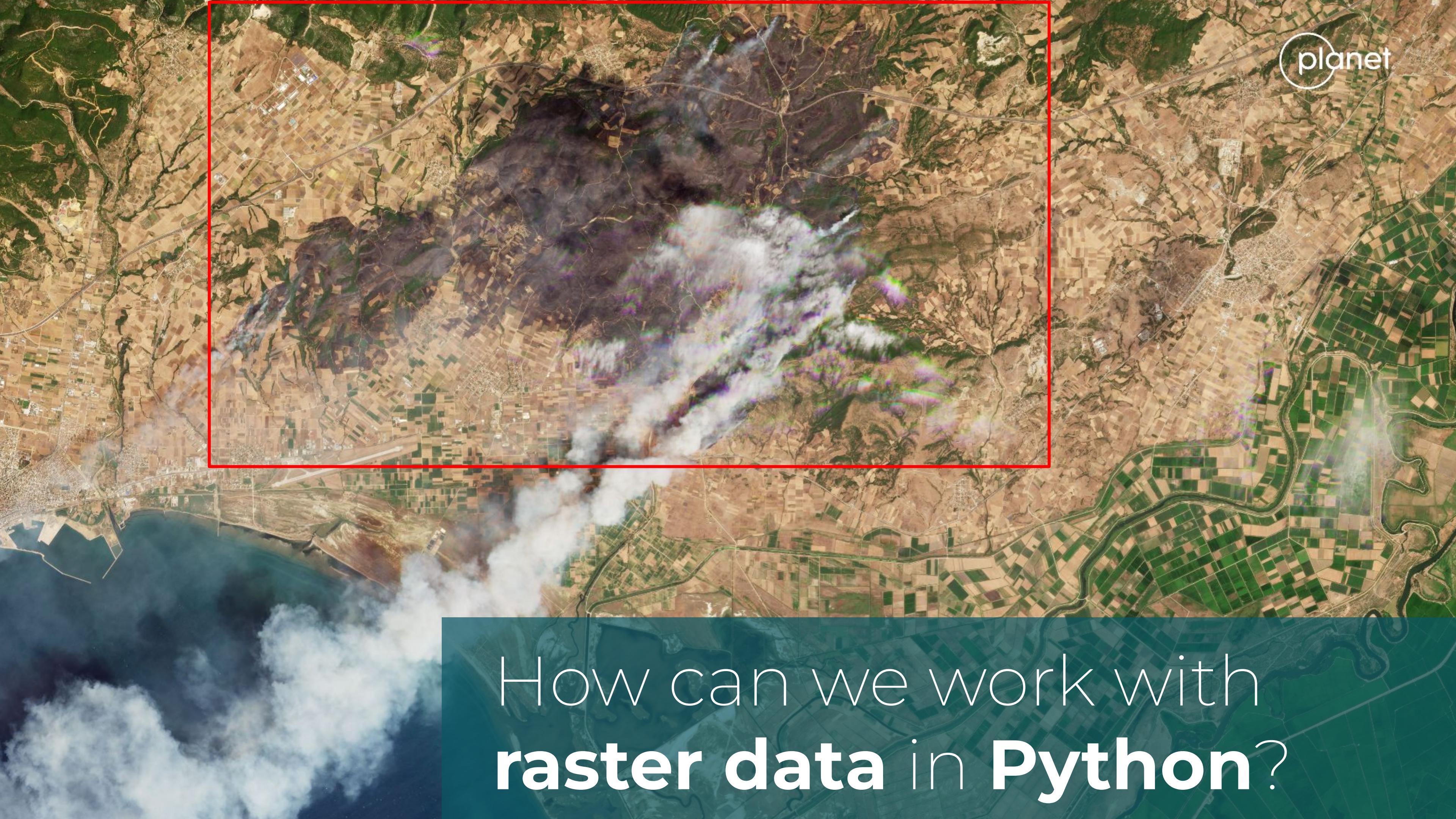
# Python libraries for raster data

- [GDAL](#)
- [Rasterio](#)
- [RSGISLib](#)
- [Rasterstats](#)
- [Rasterframes](#)

"Unleashing the Power of Pixels, Empowering Python Developers to Manipulate and Visualize Satellite and Geospatial [raster] Data with Ease!"

- ChatGPT





How can we work with  
raster data in Python?

+

# Part 1 - Raster data

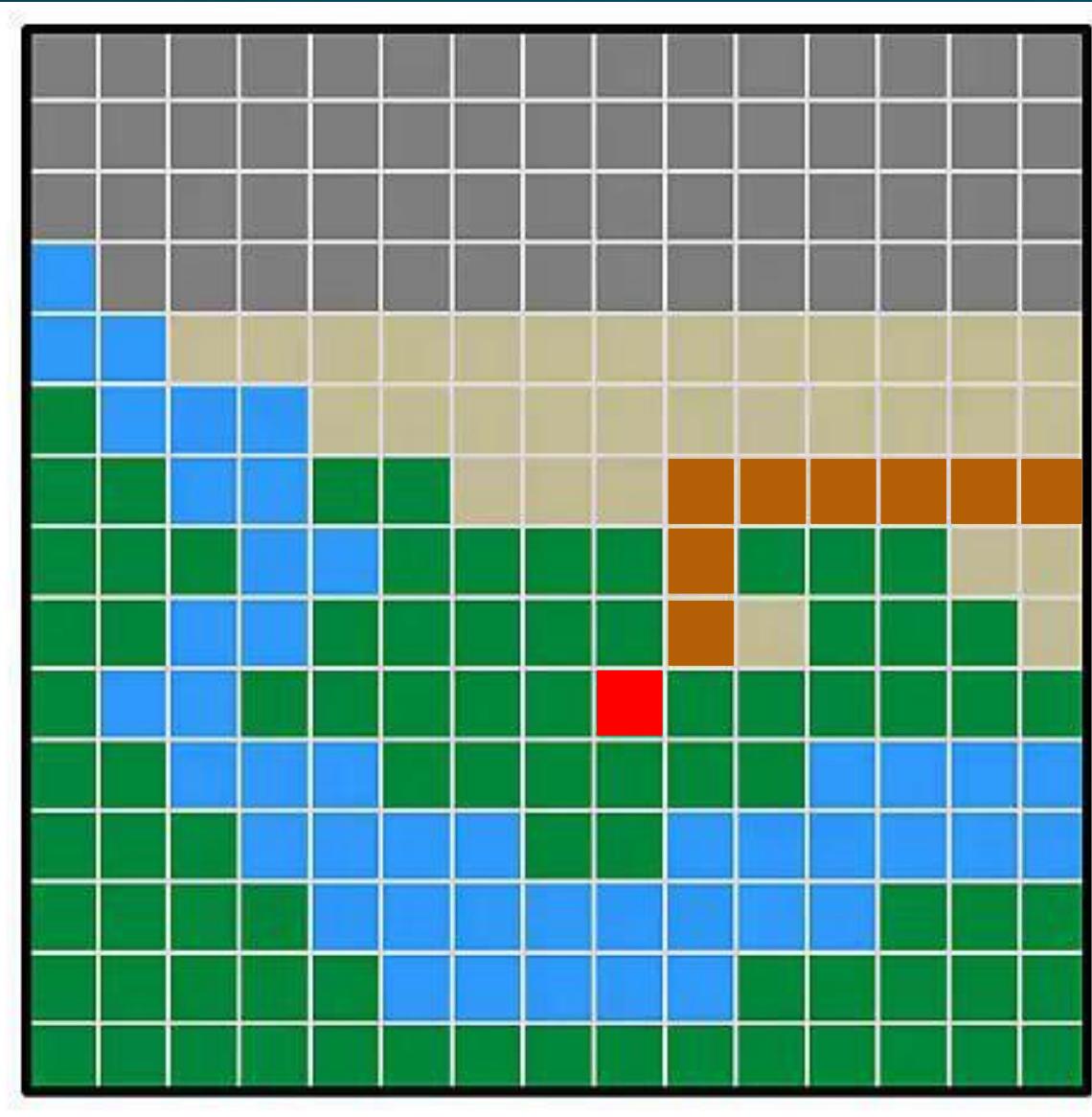


CAPE TOWN · South Africa · September 8, 2021

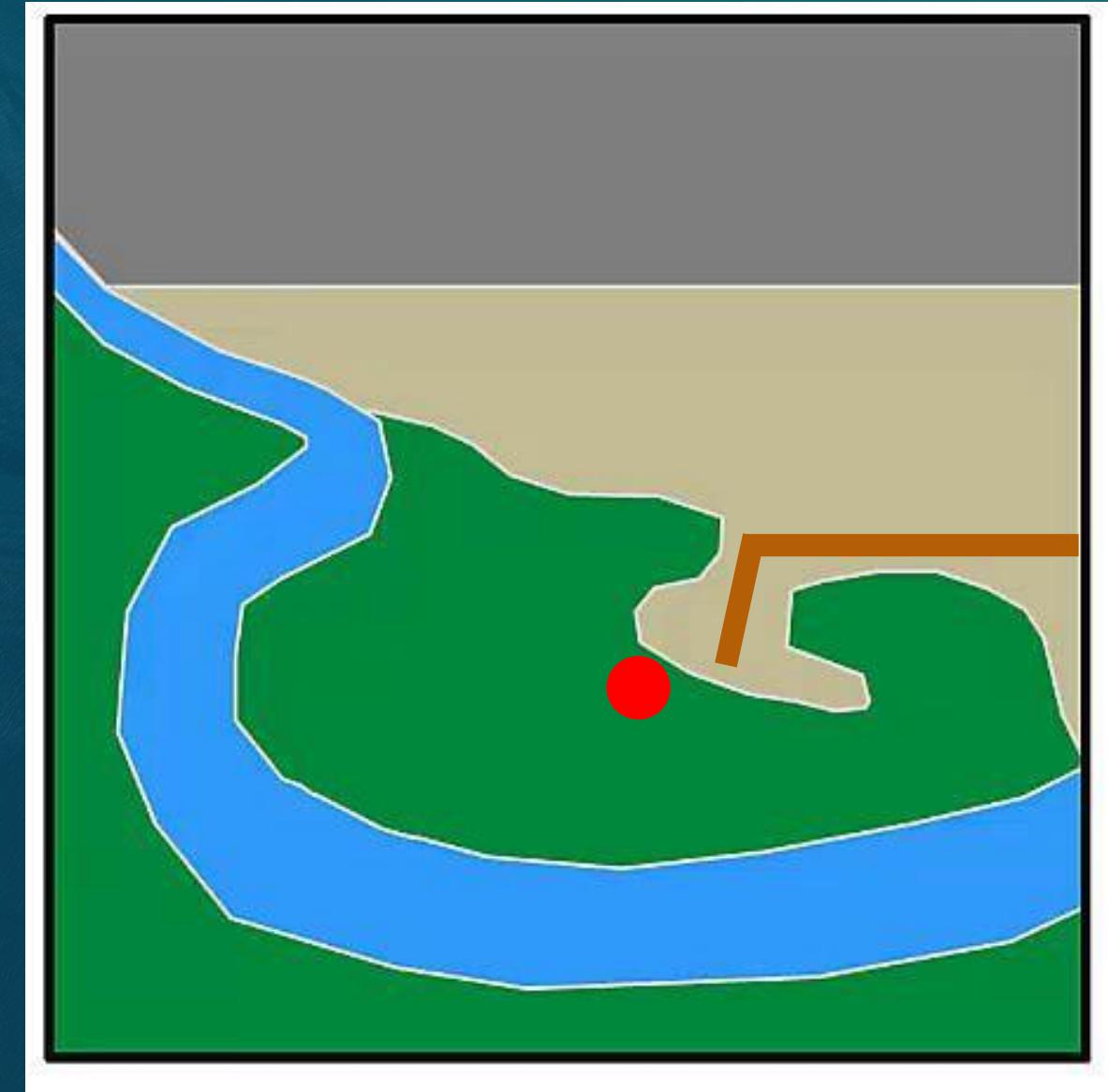
p



# Vector data



Raster

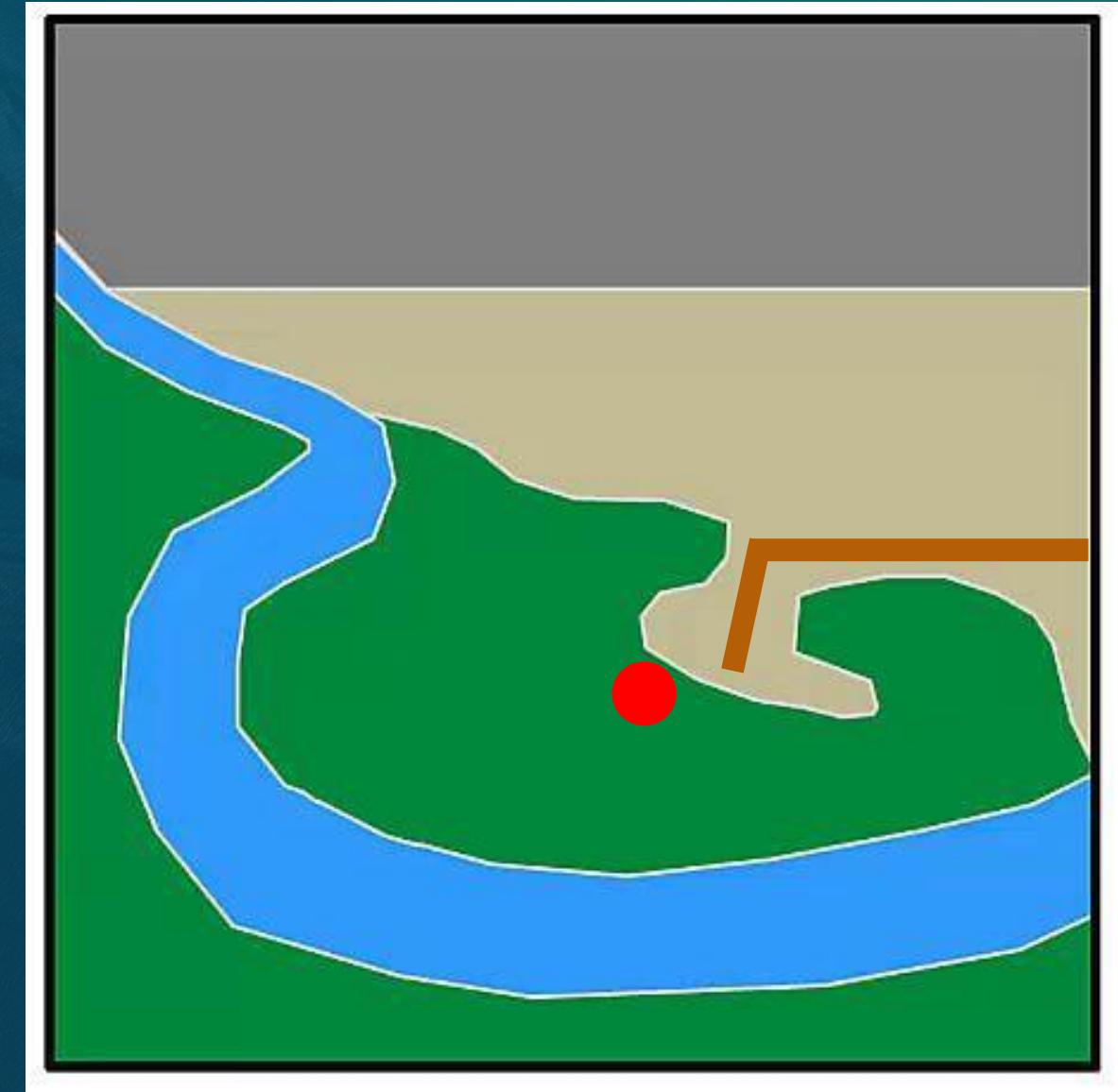
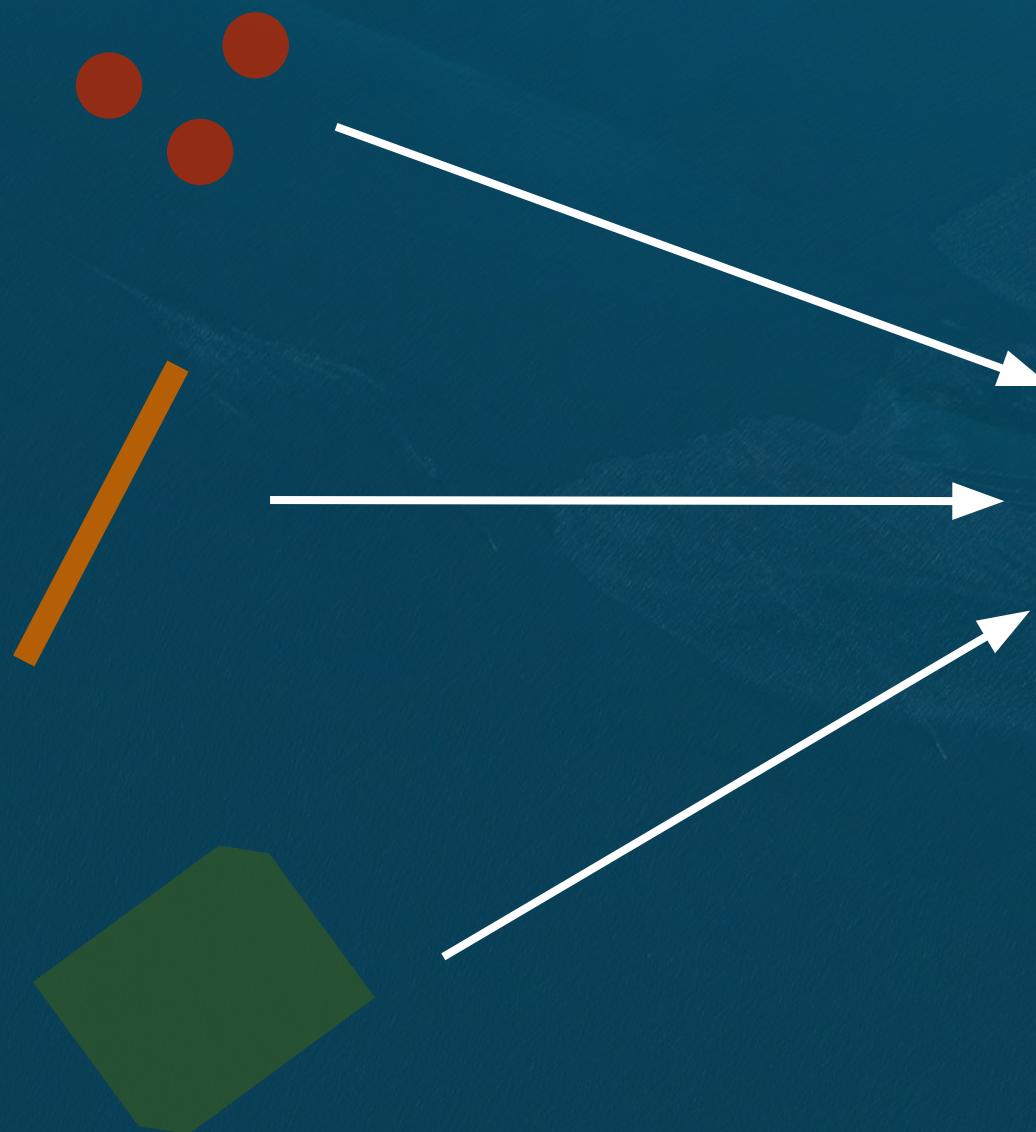


Vector

Source: [Learn Data Science](#)



# Vector data



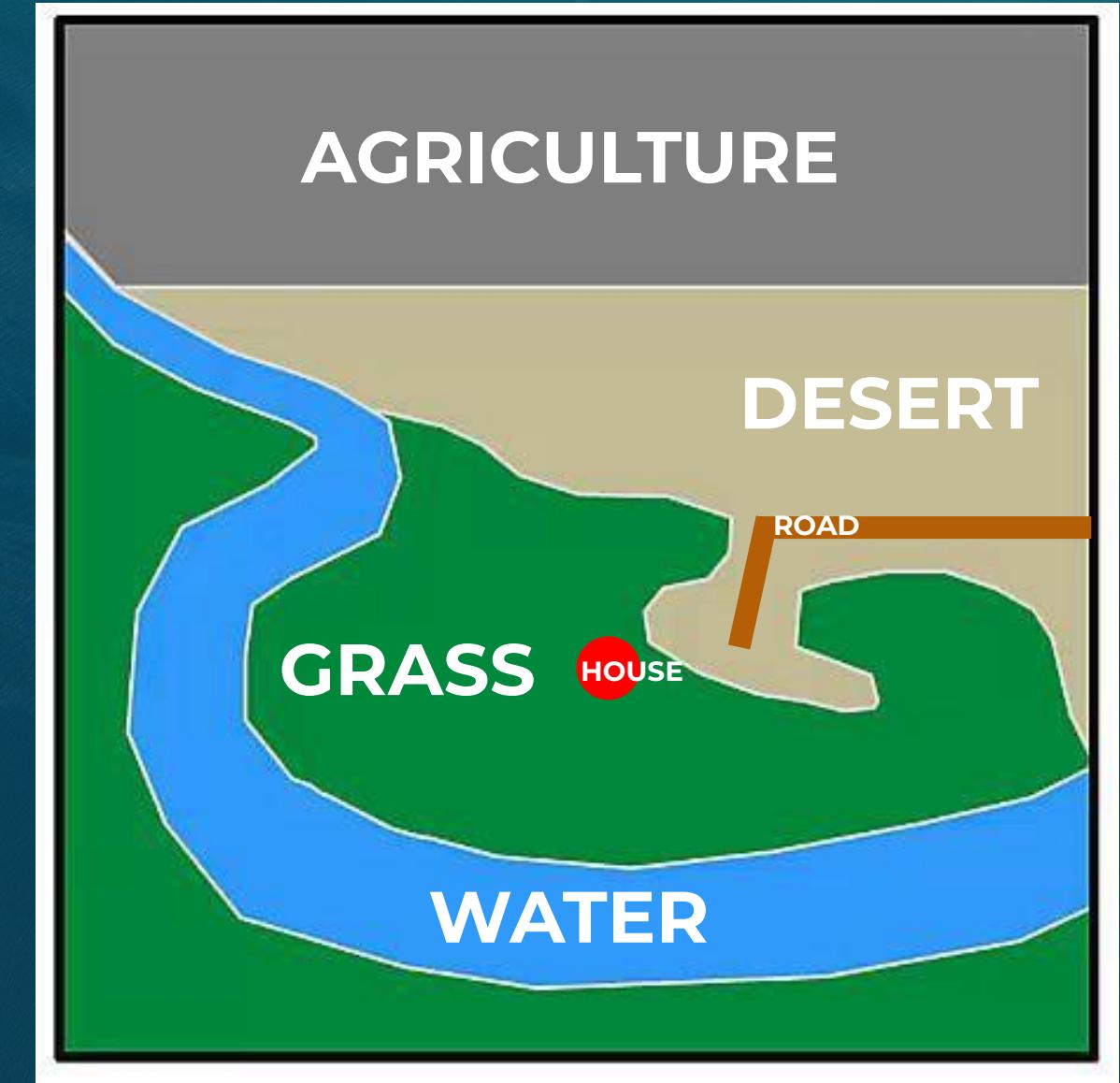
Vector

Source: [Learn Data Science](#)



# Attributes

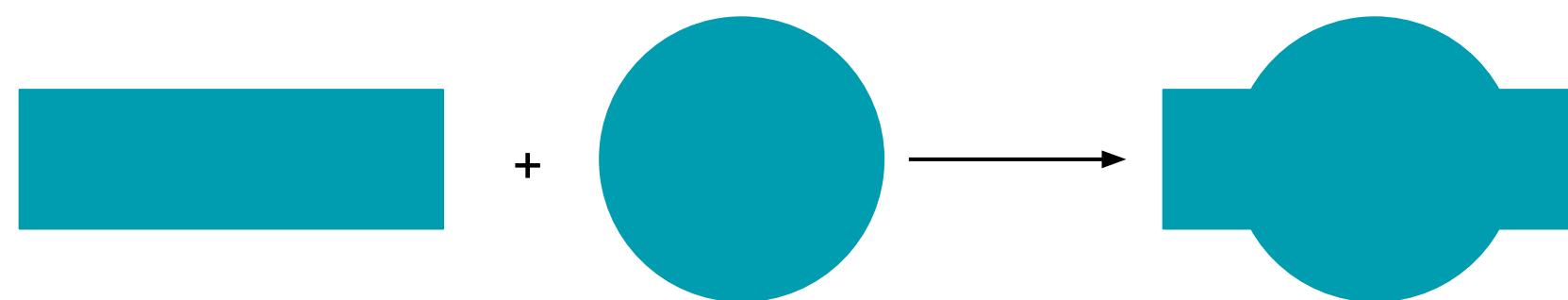
ID	CLASS	AREA [ha]	INHABITANTS
0	AGRICULTURE	10,000	1000
1	DESERT	25,000	100
2	GRASS	15,000	600
3	GRASS	12,000	800
4	WATER	5,000	0
5	HOUSE	100	3
6	ROAD	900	0



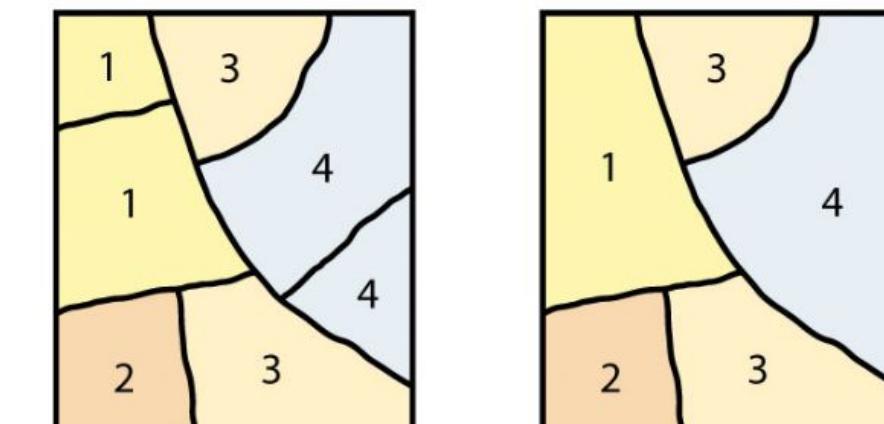
Source: [Learn Data Science](#)



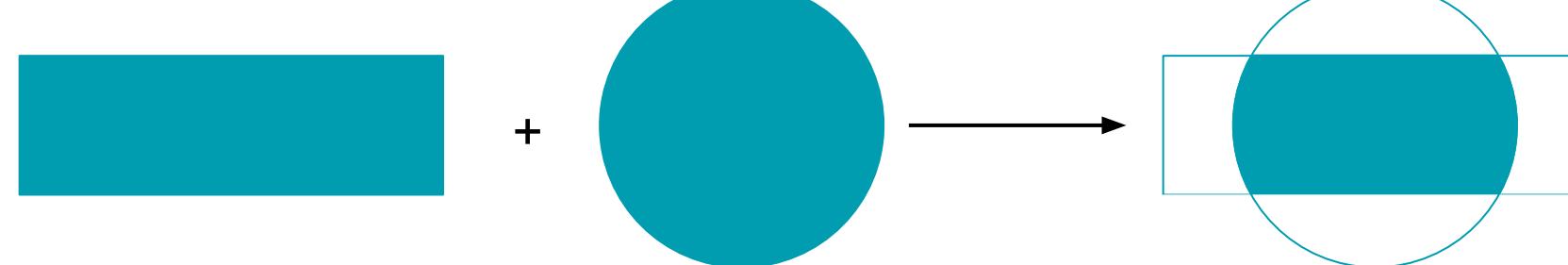
# Common vector operations



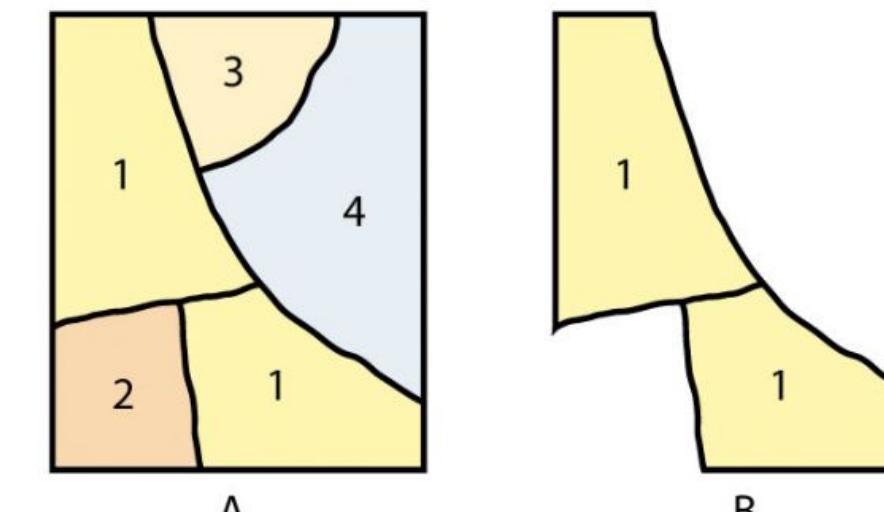
Union



Dissolve



Intersect



Select



# Vector data formats

Most commonly used

- (Geo)JSON
- ESRI shapefiles
- Geographic Markup Language (GML)
- Google Keyhole Markup Language (KML/KMZ)
- GPS eXchange (GPX)
- OpenStreetMap (OSM)





# Vector data formats

Most commonly used

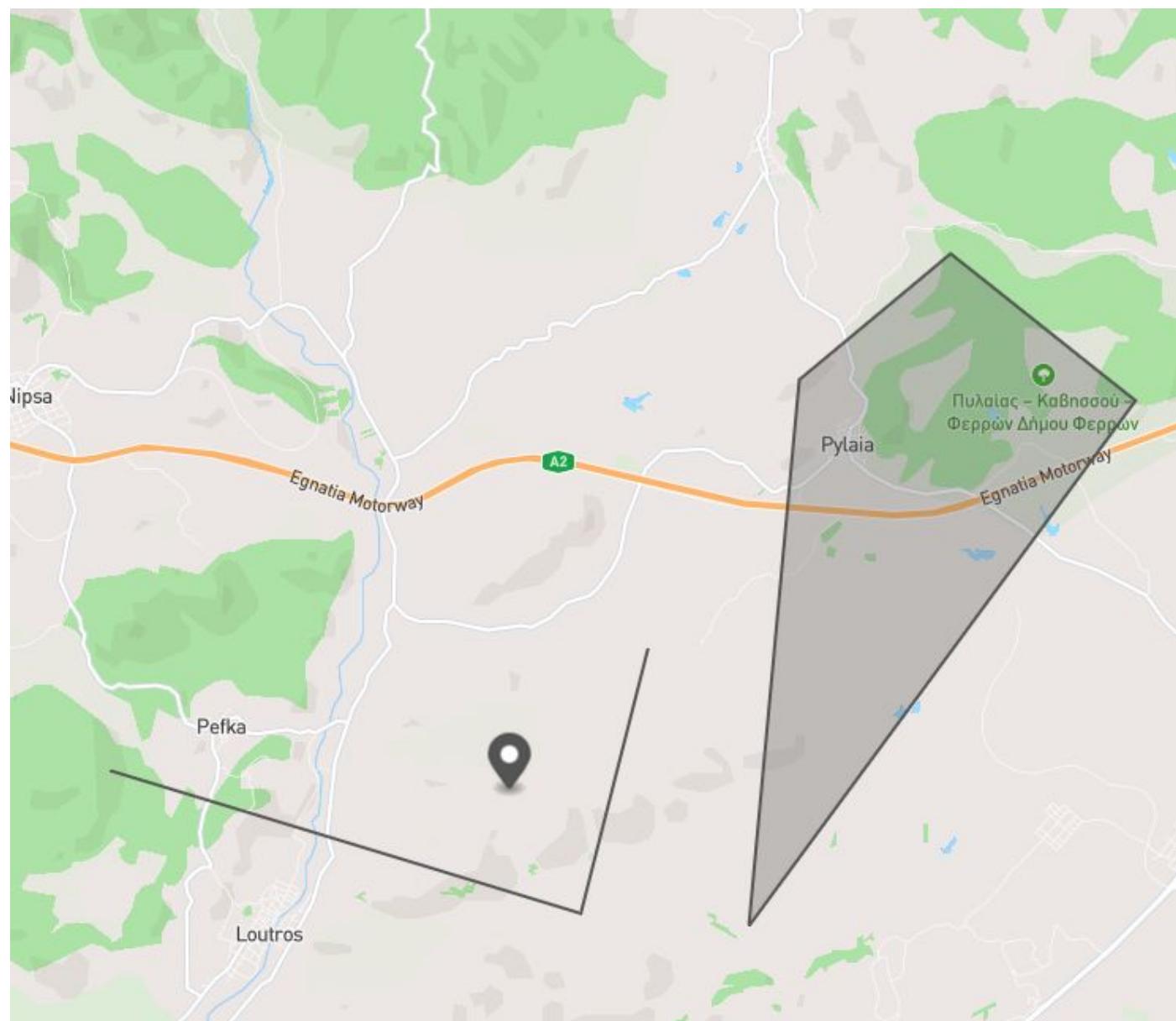
- **(Geo)JSON**
- ESRI shapefiles
- Geographic Markup Language (GML)
- Google Keyhole Markup Language (KML/KMZ)
- GPS eXchange (GPX)
- OpenStreetMap (OSM)





# Vector data formats

## GeoJSON



```
{ "type": "FeatureCollection",
  "features": [
    {"type": "Feature",
      "geometry": {
        "coordinates": [
          26.07576412038469, 40.8945601342609],
        "type": "Point" } },
    {
      "type": "Feature",
      "geometry": {
        "coordinates": [
          [26.02767985445459, 40.8960905438488 ],
          [26.084368673234593, 40.8830809338711],
          [26.092467075917625, 40.90718495453754]],
        "type": "LineString"}},
    {
      "type": "Feature",
      "geometry": {
        "coordinates": [
          [[26.110688481953787, 40.93166258553666],
          [26.104614679942188, 40.88193290430556],
          [26.151180495367697, 40.92975059704111],
          [26.128909887990005, 40.943133354584944],
          [26.110688481953787, 40.93166258553666]]],
        "type": "Polygon"}]
  }
```

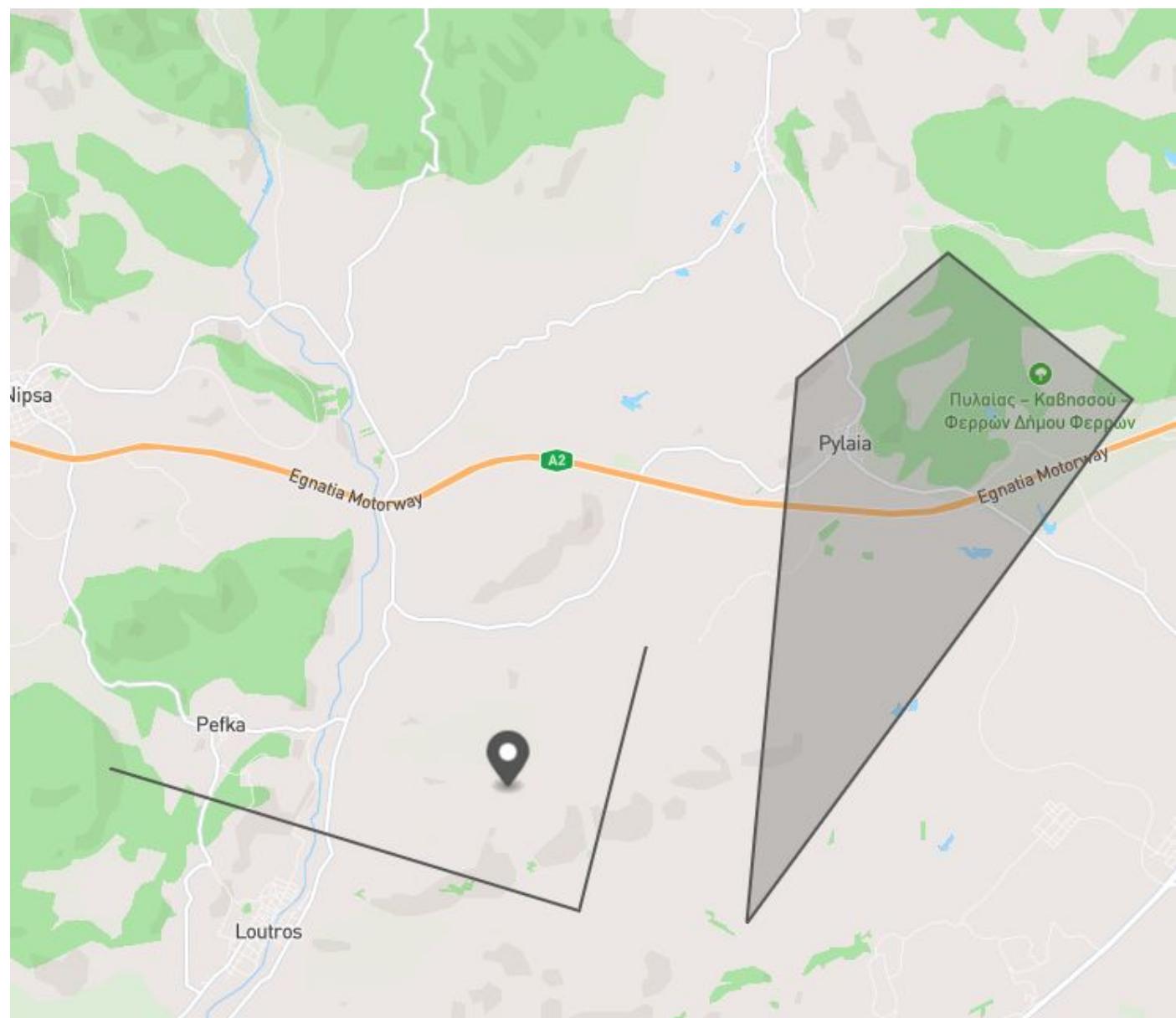


LAKE OKEECHOBEE  
Florida, USA



# Vector data formats

## GeoJSON



[Draw your own geometries!](#)

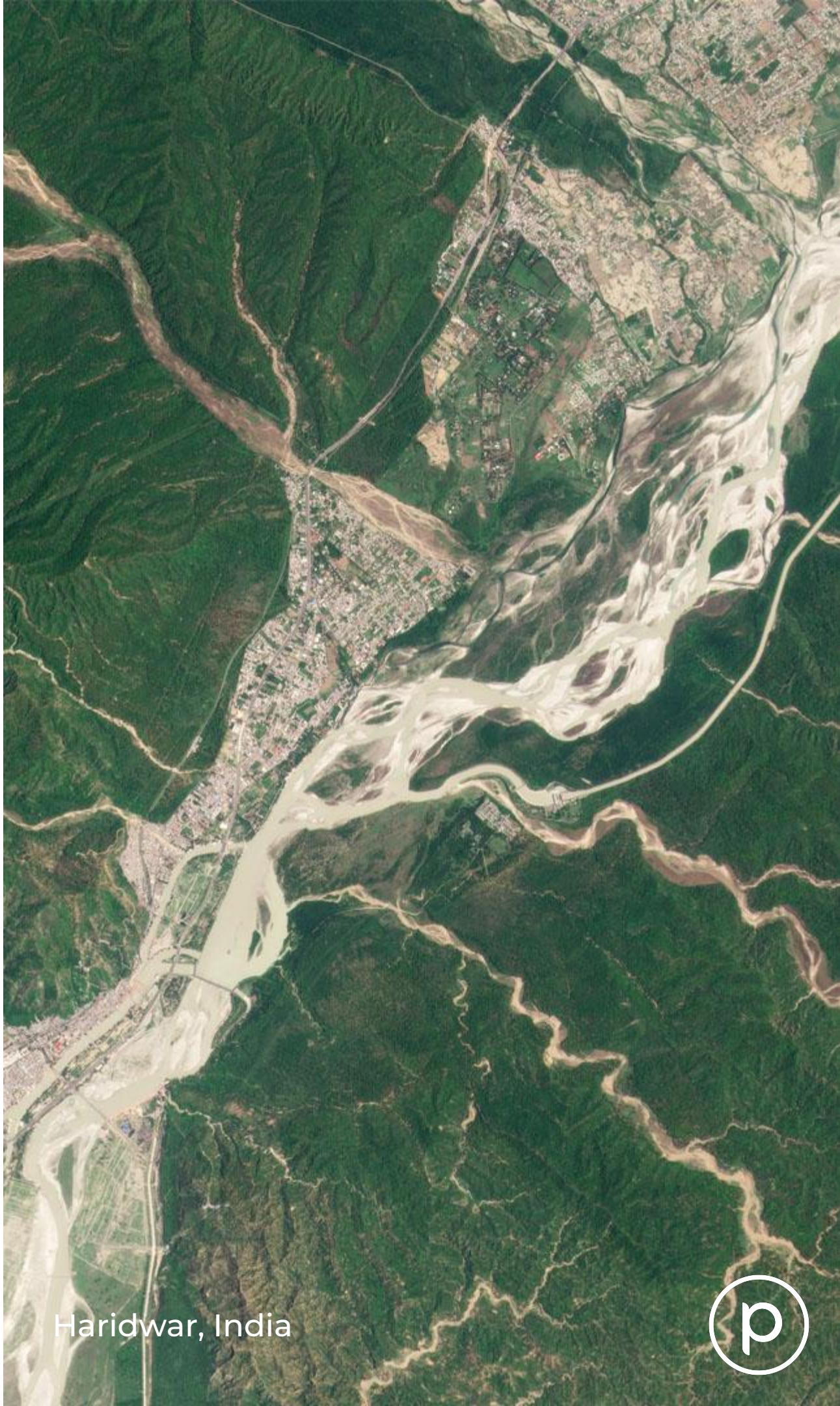
```
{ "type": "FeatureCollection",
  "features": [
    { "type": "Feature",
      "geometry": {
        "coordinates": [
          26.07576412038469, 40.8945601342609],
        "type": "Point" }},
    {
      "type": "Feature",
      "geometry": {
        "coordinates": [
          [26.02767985445459, 40.8960905438488 ],
          [26.084368673234593, 40.8830809338711],
          [26.092467075917625, 40.90718495453754]],
        "type": "LineString"}},
    {
      "type": "Feature",
      "geometry": {
        "coordinates": [
          [[26.110688481953787, 40.93166258553666],
          [26.104614679942188, 40.88193290430556],
          [26.151180495367697, 40.92975059704111],
          [26.128909887990005, 40.943133354584944],
          [26.110688481953787, 40.93166258553666]]],
        "type": "Polygon"}}
  ]}
```





# Python libraries for vector data

- [GeoPandas](#)
- [Shapely](#)
- [Fiona](#)
- [PyProj](#)



Haridwar, India

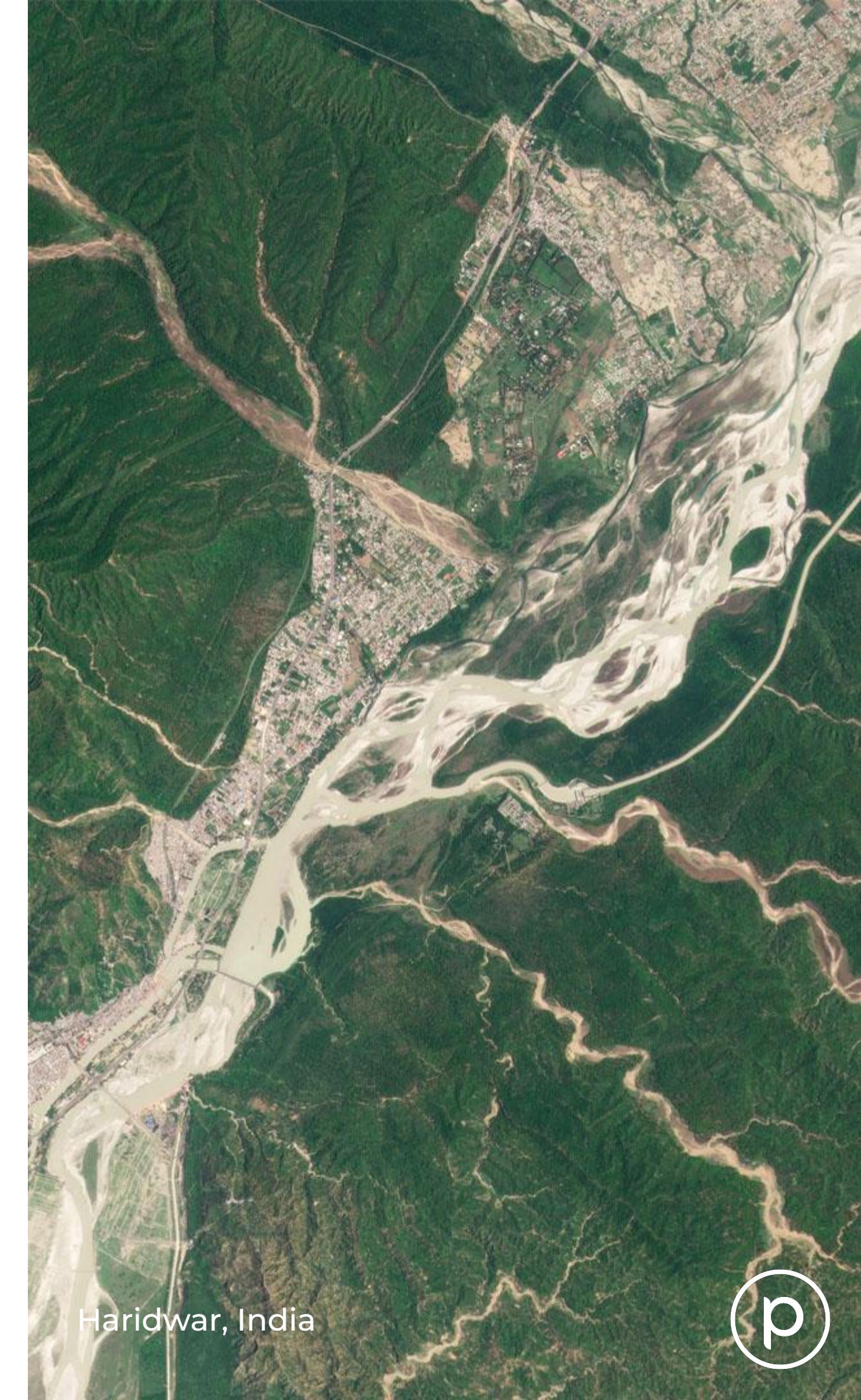


# Python libraries for vector data

- [\*\*GeoPandas\*\*](#)
- [Shapely](#)
- [Fiona](#)
- [PyProj](#)



[Source](#)



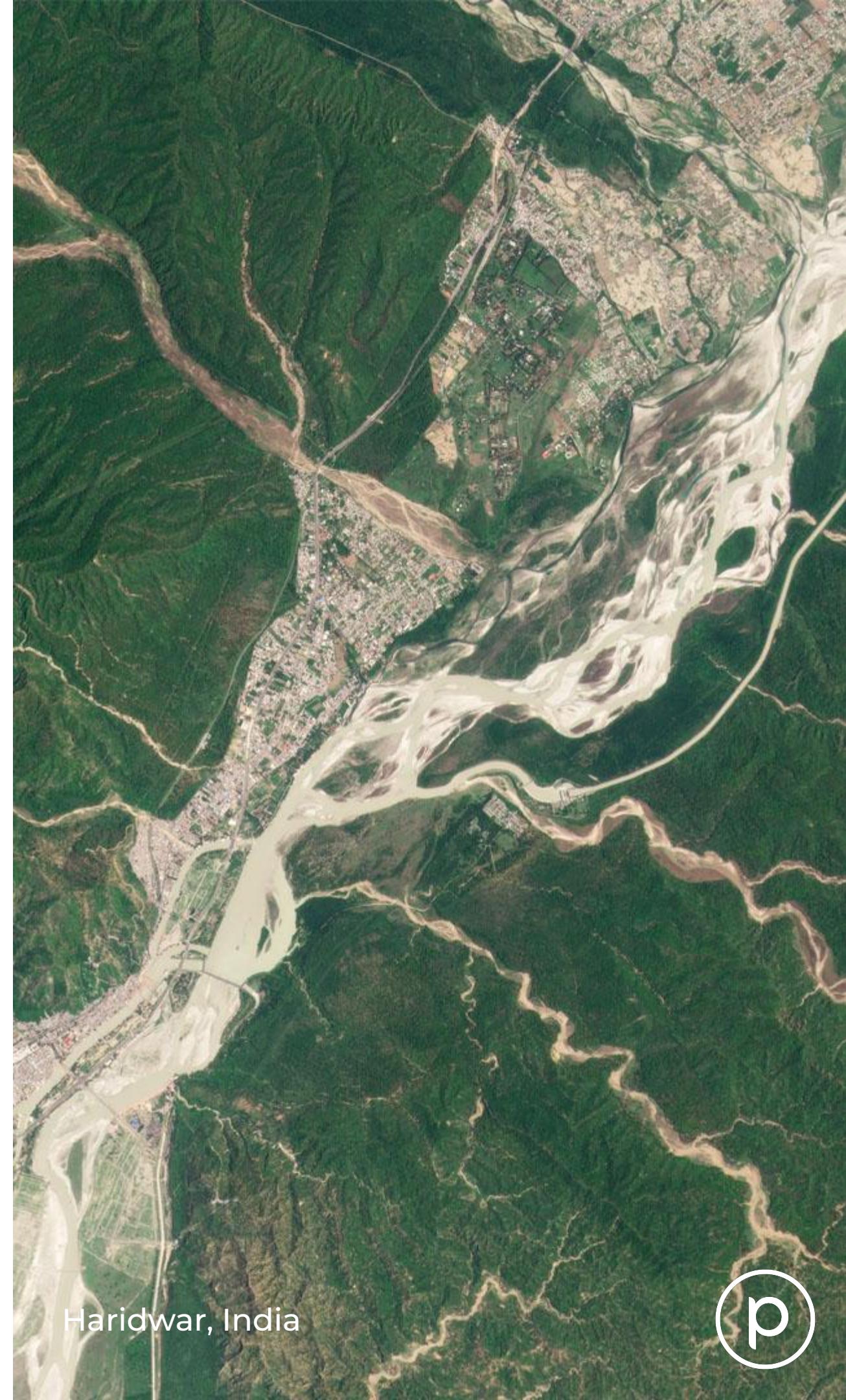
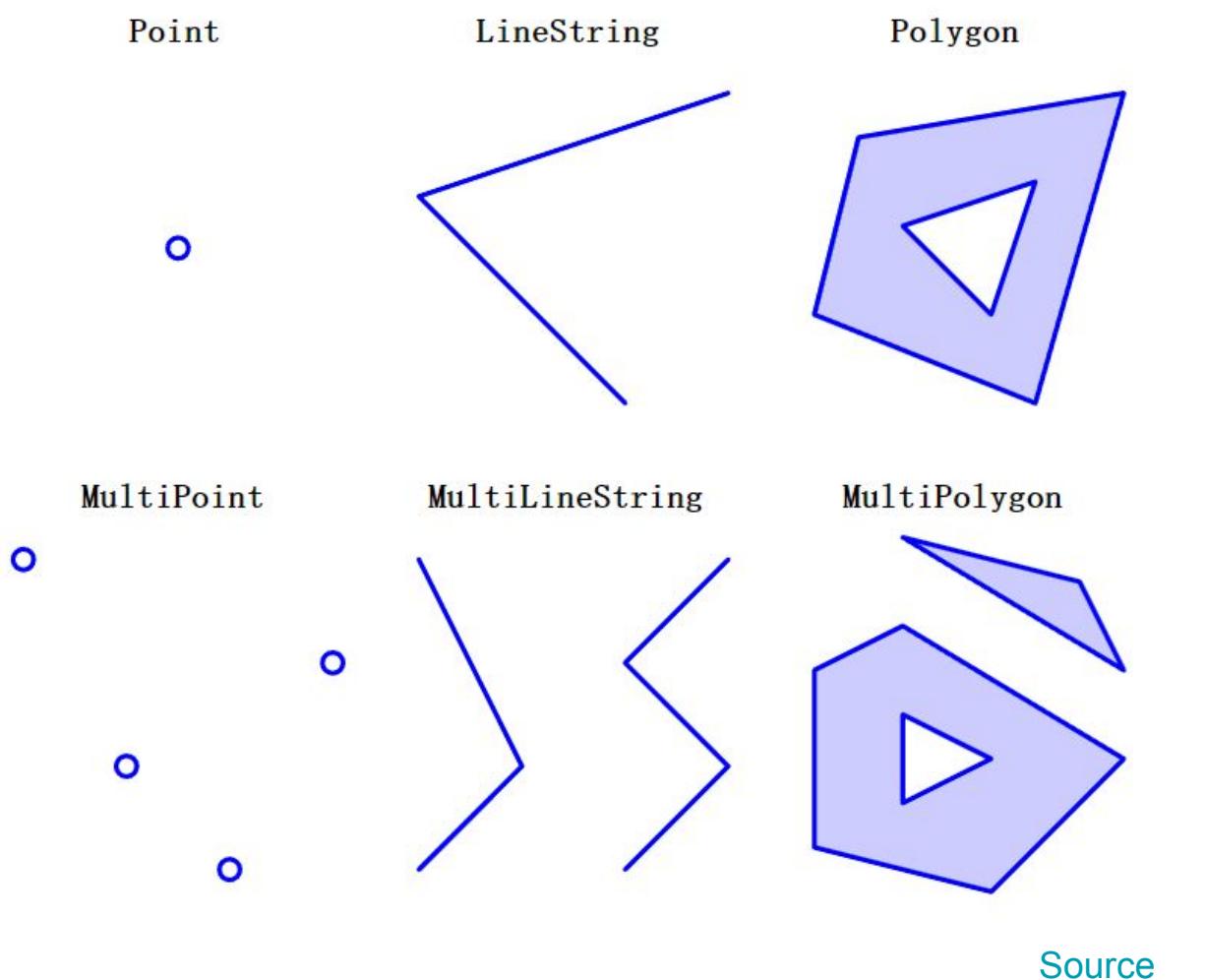
Haridwar, India





# Python libraries for vector data

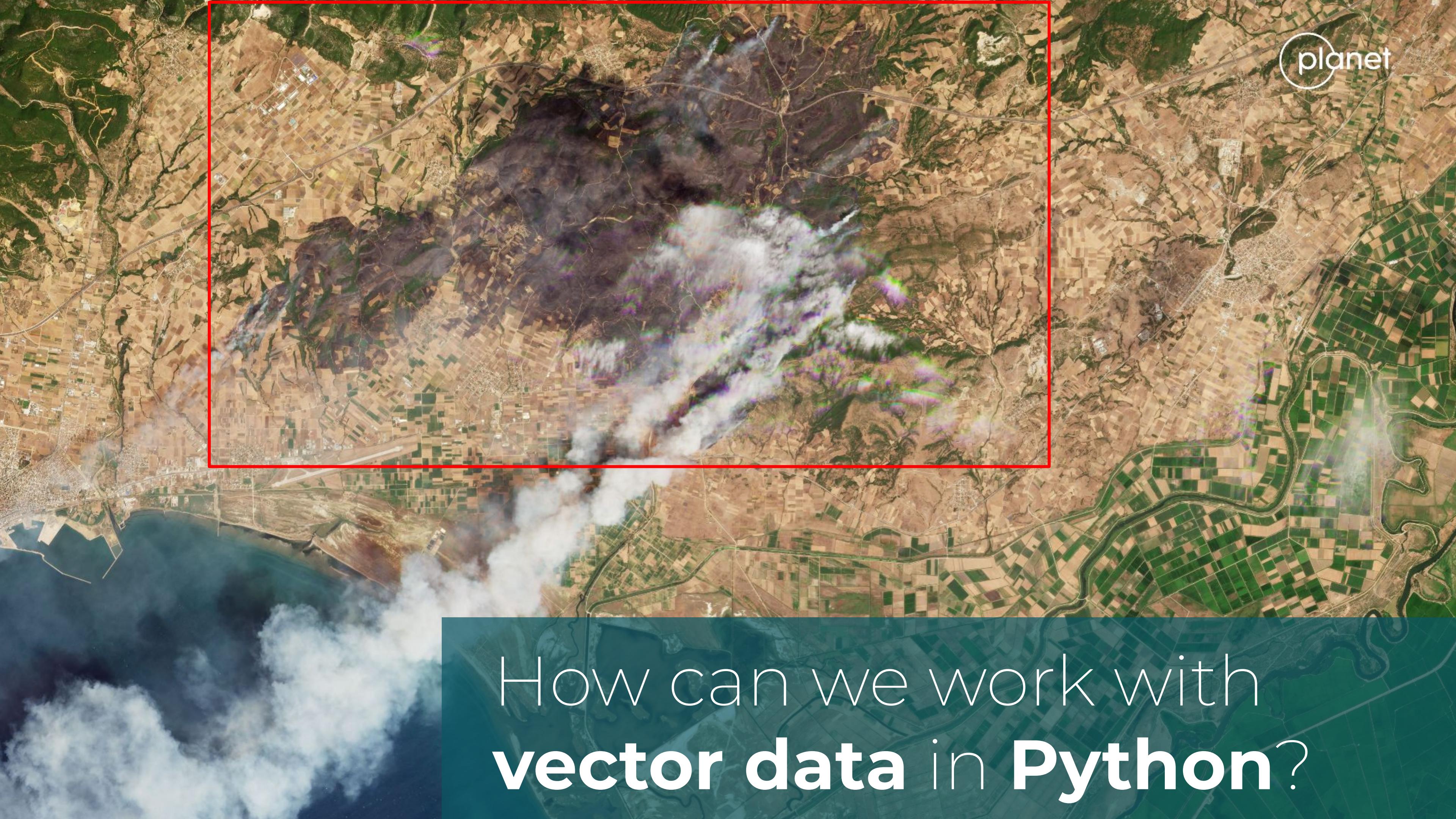
- [GeoPandas](#)
- [Shapely](#)
- [Fiona](#)
- [PyProj](#)



Haridwar, India

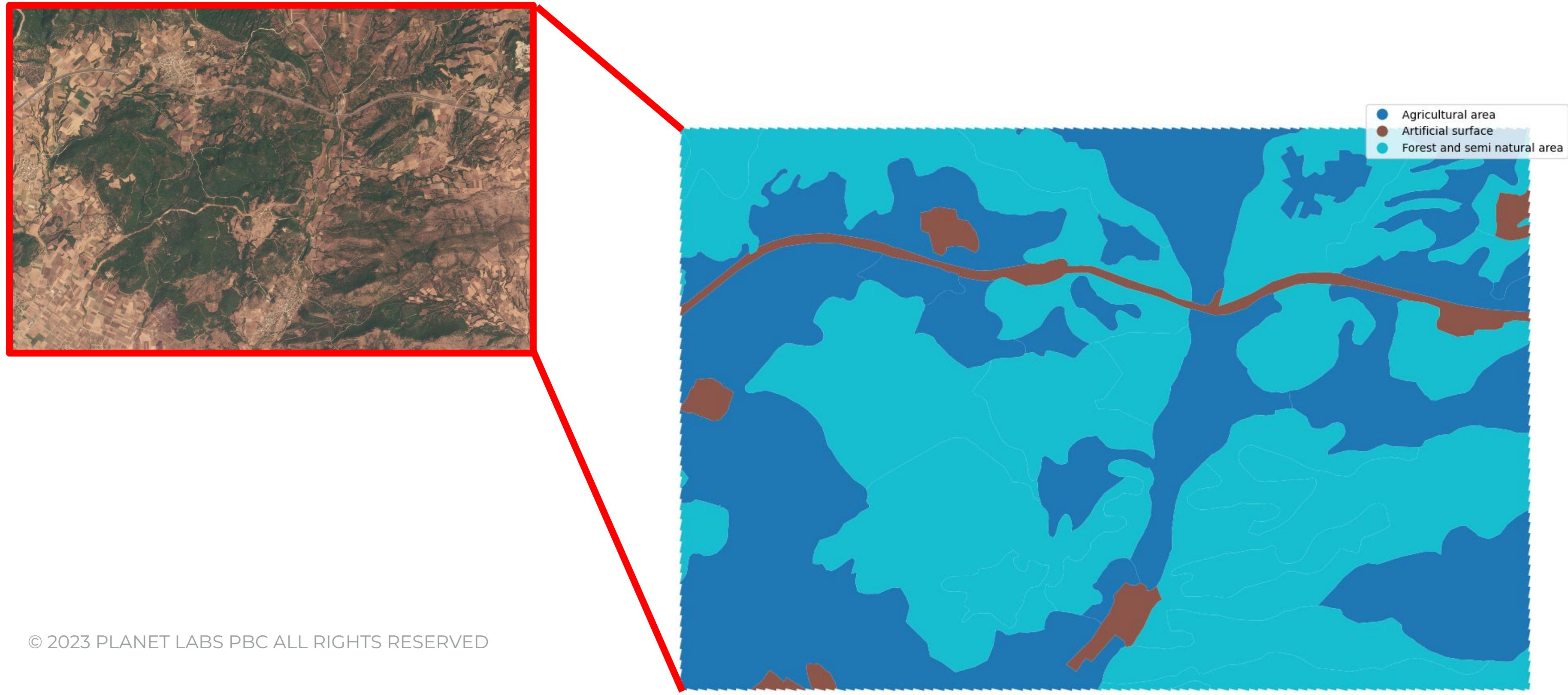


How can we work with  
**vector data** in **Python**?



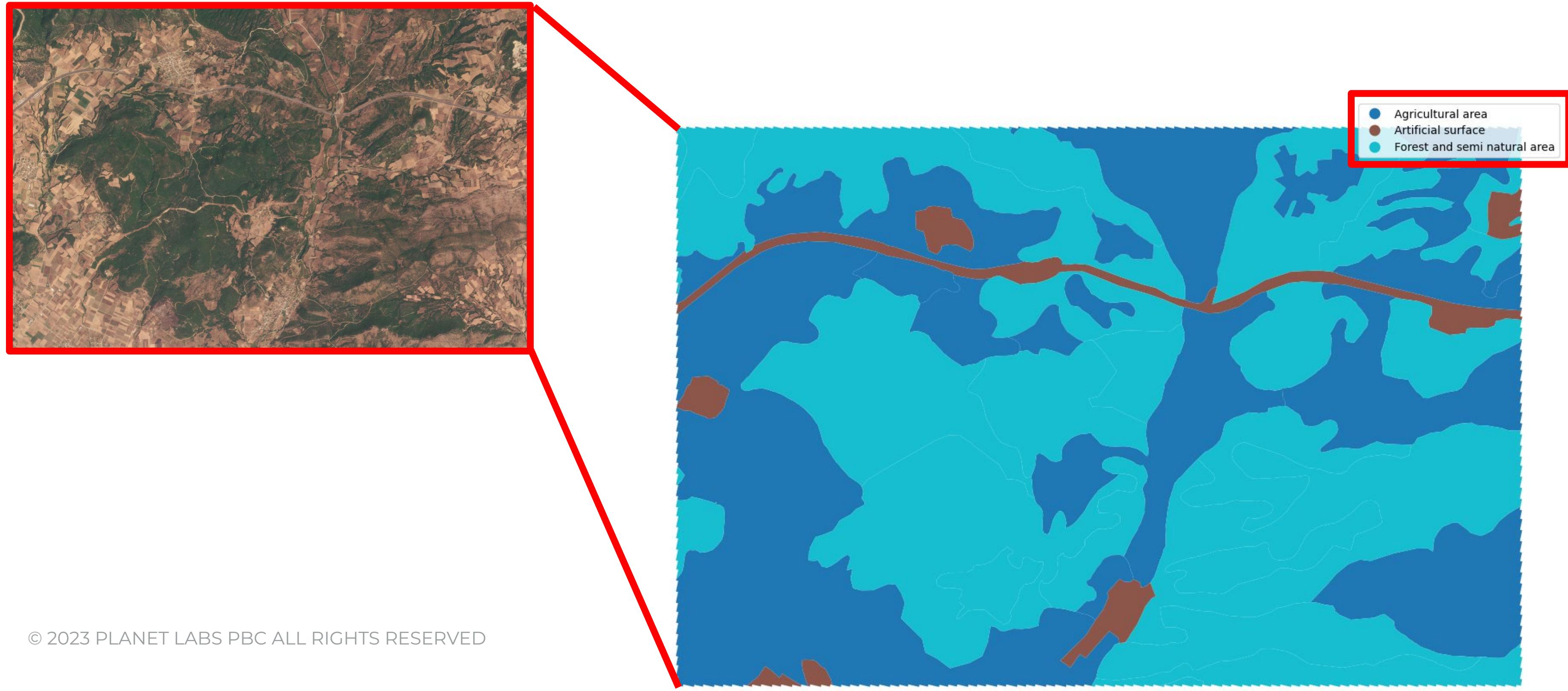


# Understanding our study area



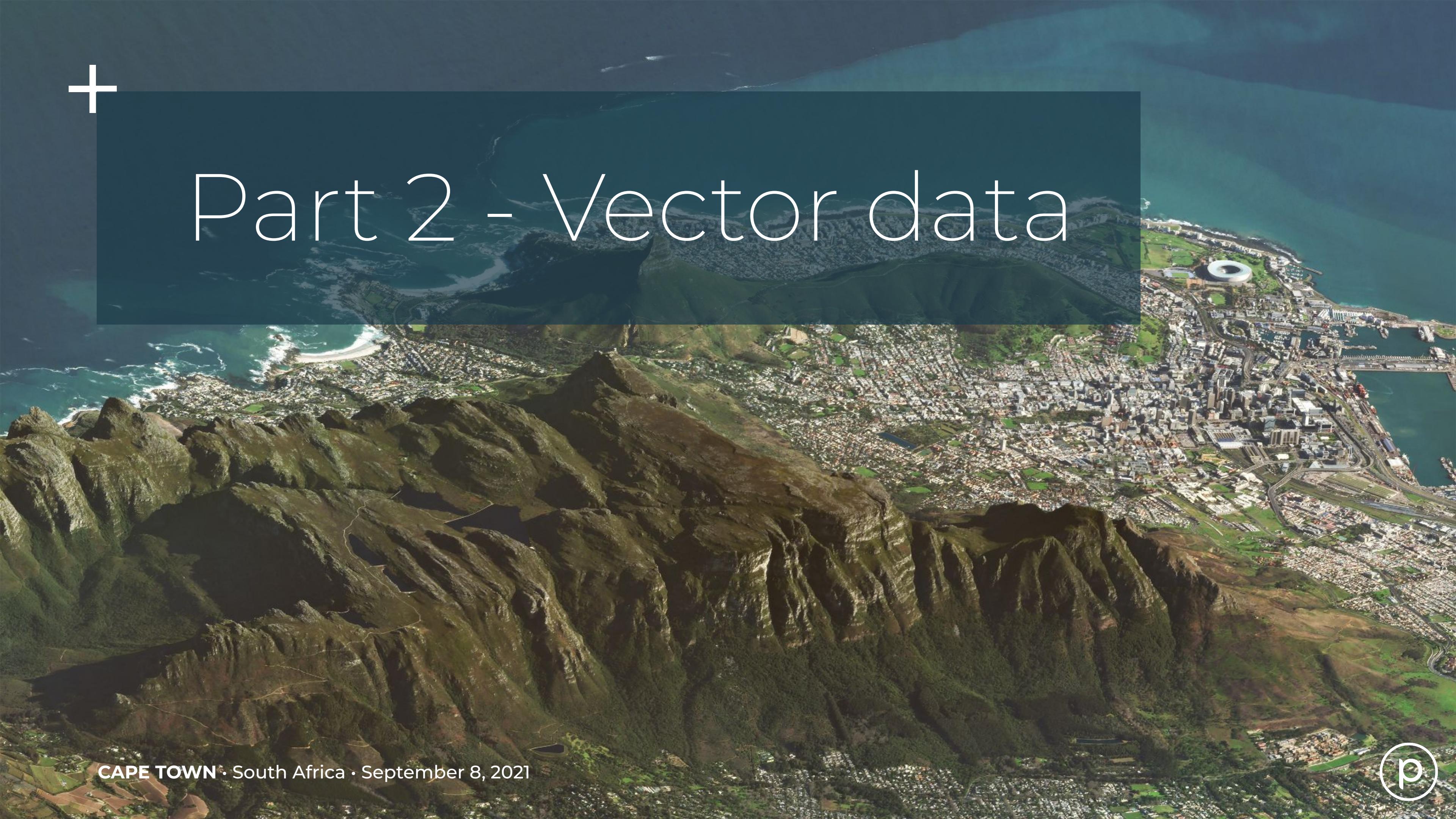


# Understanding our study area

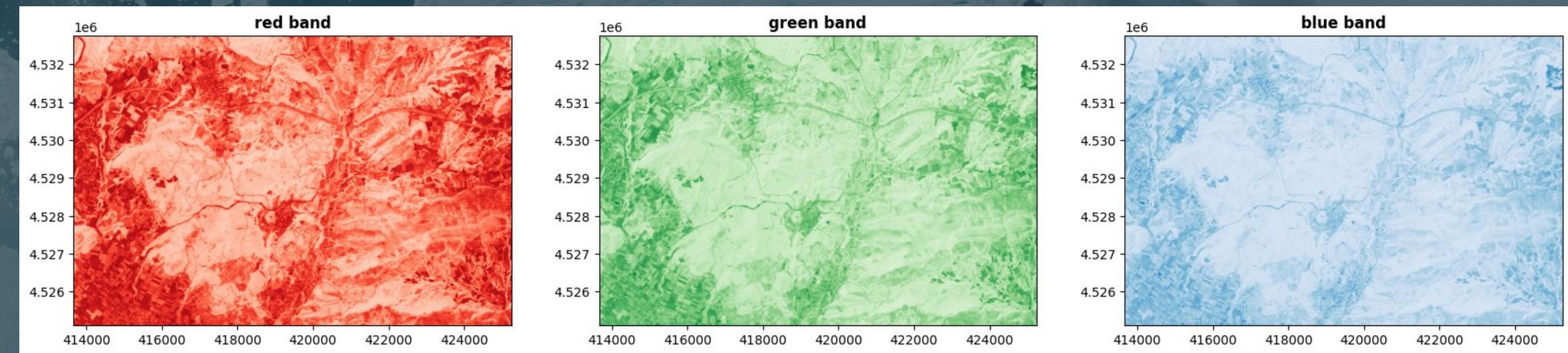


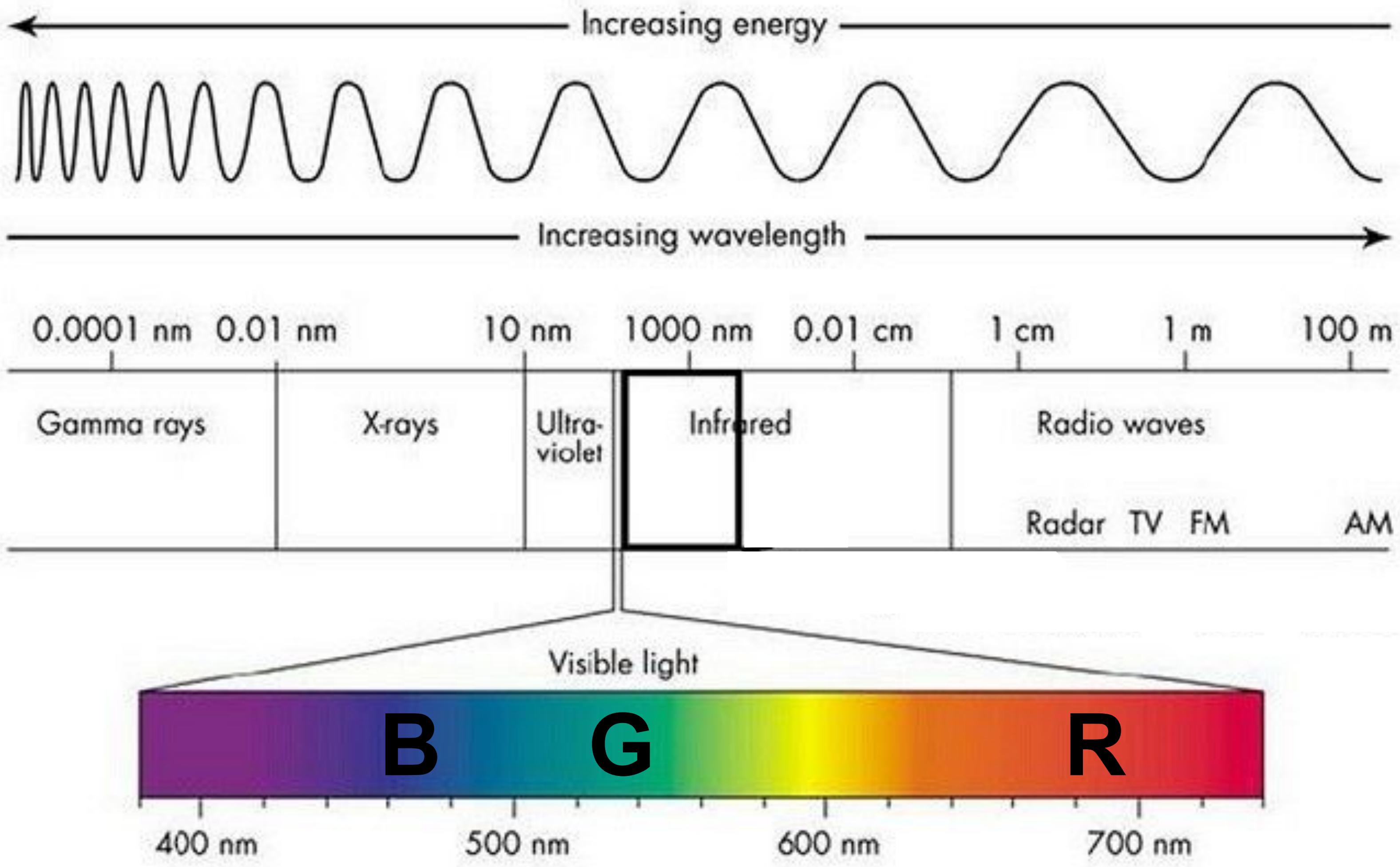
+

# Part 2 - Vector data

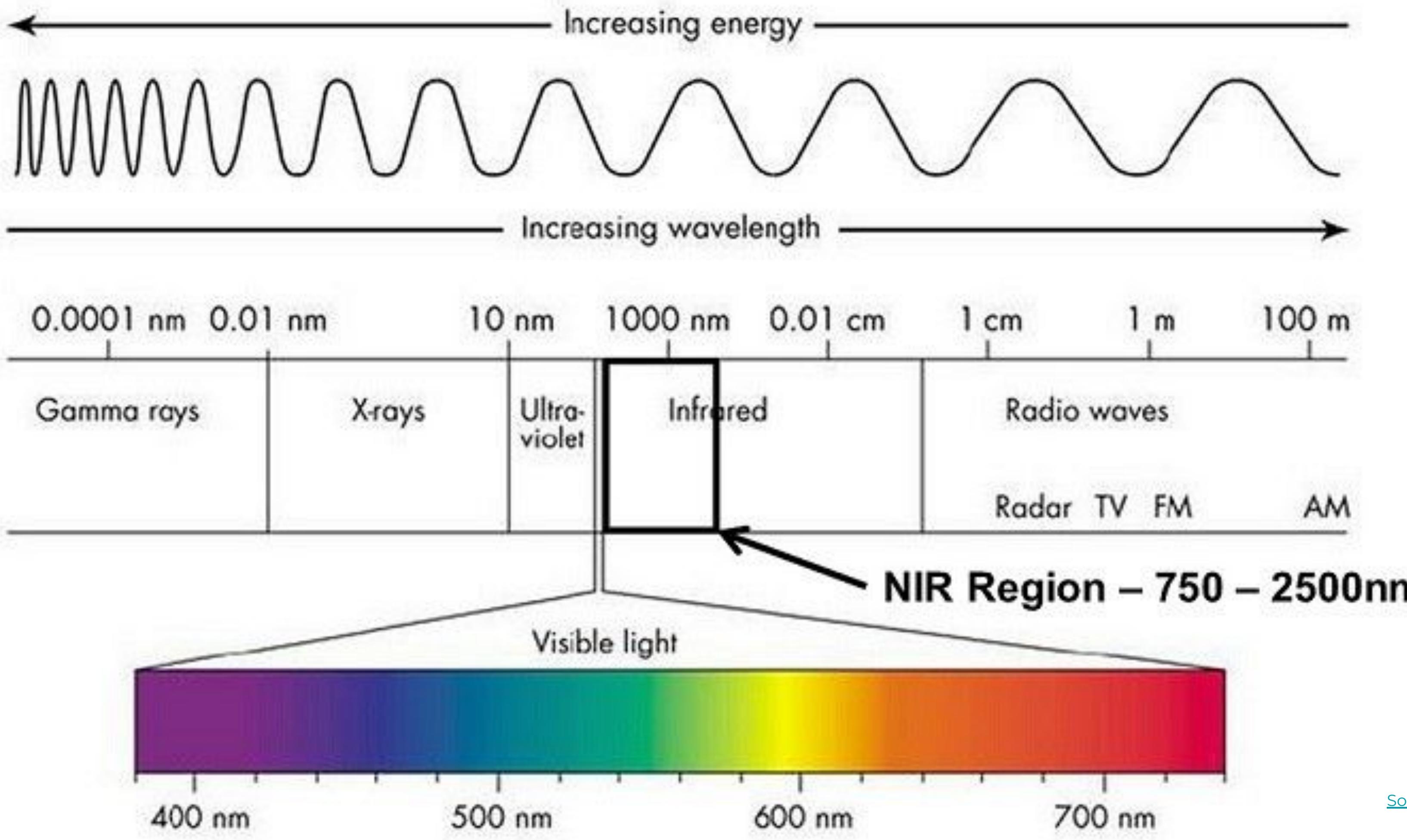


# Working with multiple bands





[Source](#)

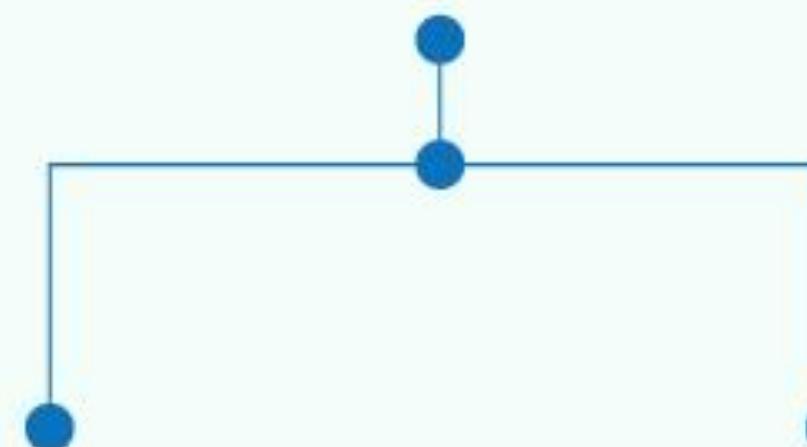


[Source](#)

# Normalized Difference Vegetation Index (NDVI)

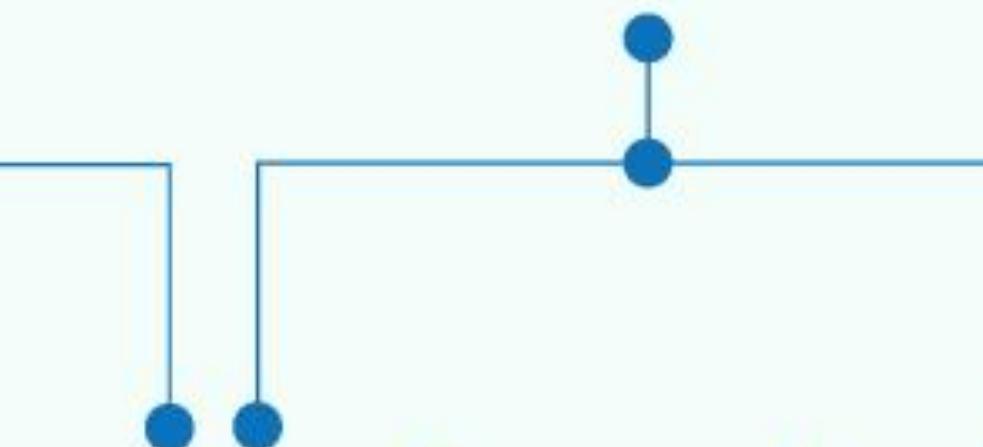
**-1 - 0**

Dead plant or inanimate object



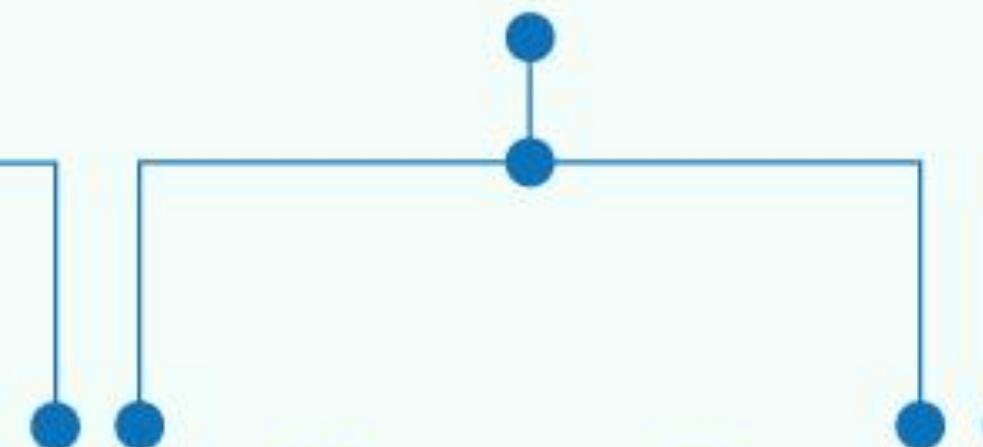
**0 - 0.33**

Diseased plant



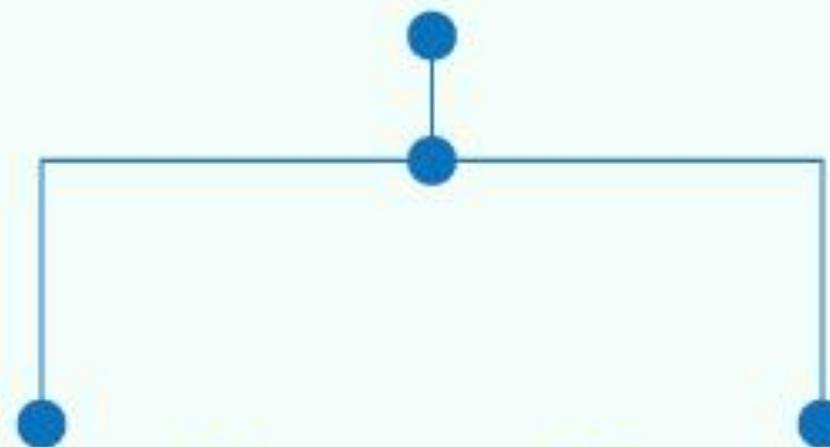
**0.33 - 0.66**

Moderately healthy plant



**0.66 - 1**

Very healthy plant





# Combining vector and raster

Locating the forest area impacted the most by the wildfire



RGB image before fire



RGB image after fire



# Part 3 - Impact of wildfires

CAPE TOWN · South Africa · September 8, 2021





# What we have learned

BEGA FIRE · New South Wales, Australia · August 8, 2018





# What we have learned

- Geodata is information that is tied to a **specific location on earth**



# What we have learned

- Geodata is information that is tied to a **specific location on earth**
- Geodata can support **informed decision making**



# What we have learned

- Geodata is information that is tied to a **specific location on earth**
- Geodata can support **informed decision making**
- Use case: Wildfires in Greece
  - Working with **raster** and **vector** data in **Python**
  - Satellite imagery can aid in wildfire **monitoring & prediction**



planet.

+ Thank you for your attention



# Contact information

Margot Ridderikhoff  
[margot@planet.com](mailto:margot@planet.com)

Amber Mulder  
[amber.mulder@planet.com](mailto:amber.mulder@planet.com)