Introduction to Spatial Analysis in R

BOLDER Africa Gap Analysis Workshop

Wanjiku Githu

b.wanjiku@cgiar.org

January 28, 2025

Cali, Colombia

Time: 2 hrs 30 mins

Overview

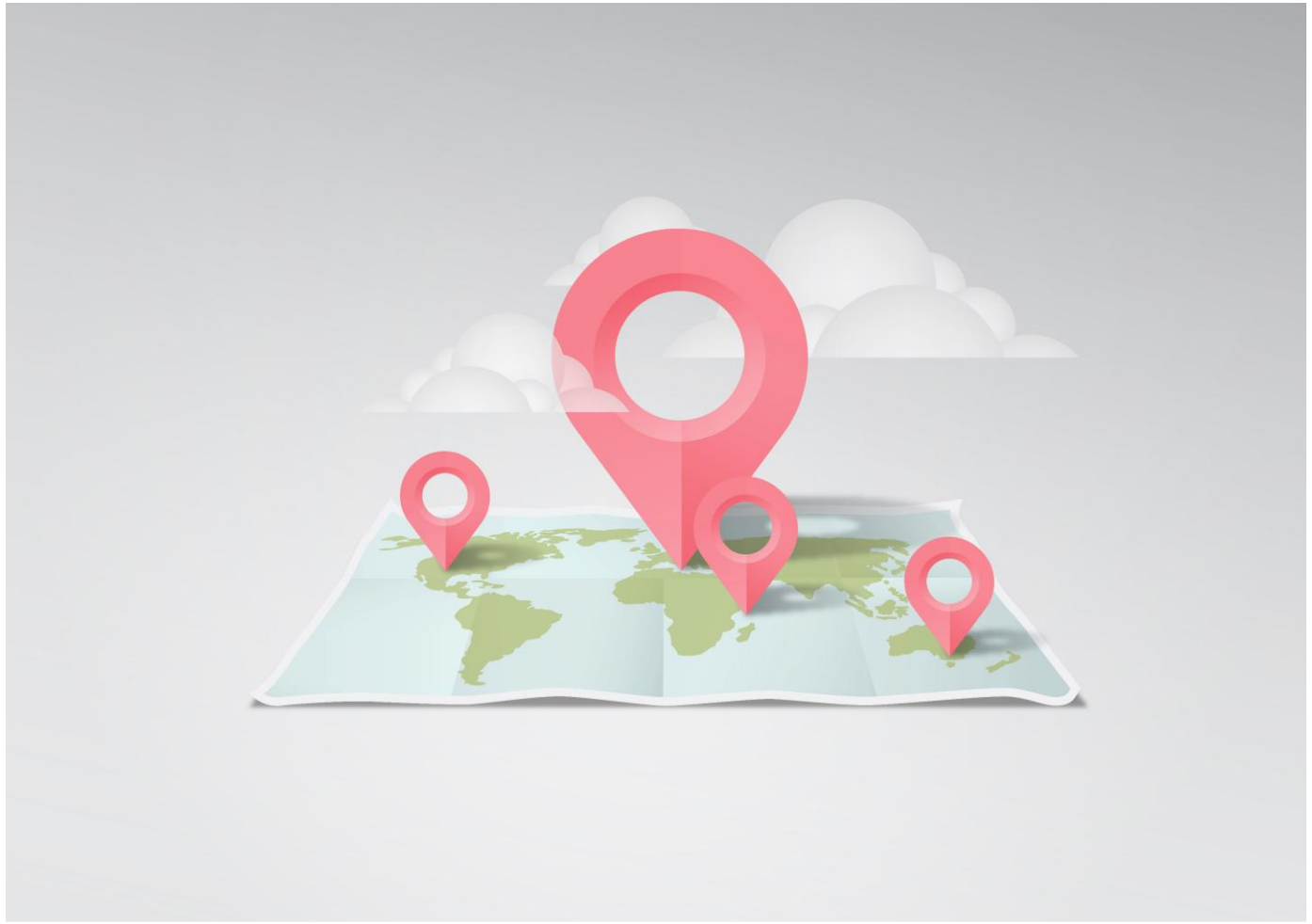
- Introduction
- Spatial data formats
- Important spatial concepts
- Basic spatial data operations
- Further resources
- R Practical Session

Key questions

- What is spatial data?
- What are the common spatial data formats?
- What are the important spatial concepts that you need to know?
- How do you work with vector and raster data?
- What basic operations can we perform on spatial data?

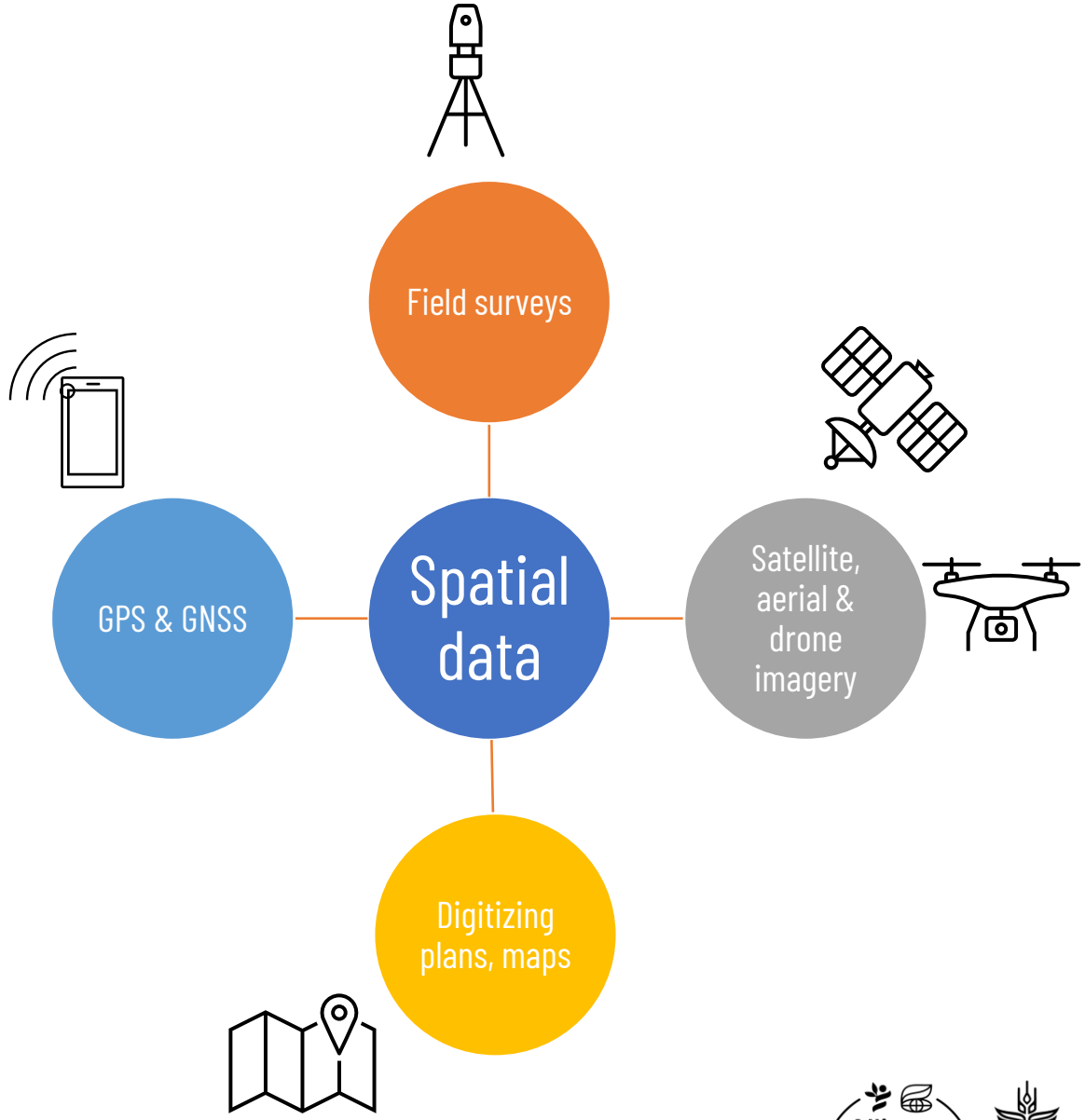
Spatial data is ...

- Data that has location information about a specific area on the earth's surface
- Referred to as geospatial data or geographic information
- Used in environmental science, geography, urban planning etc.

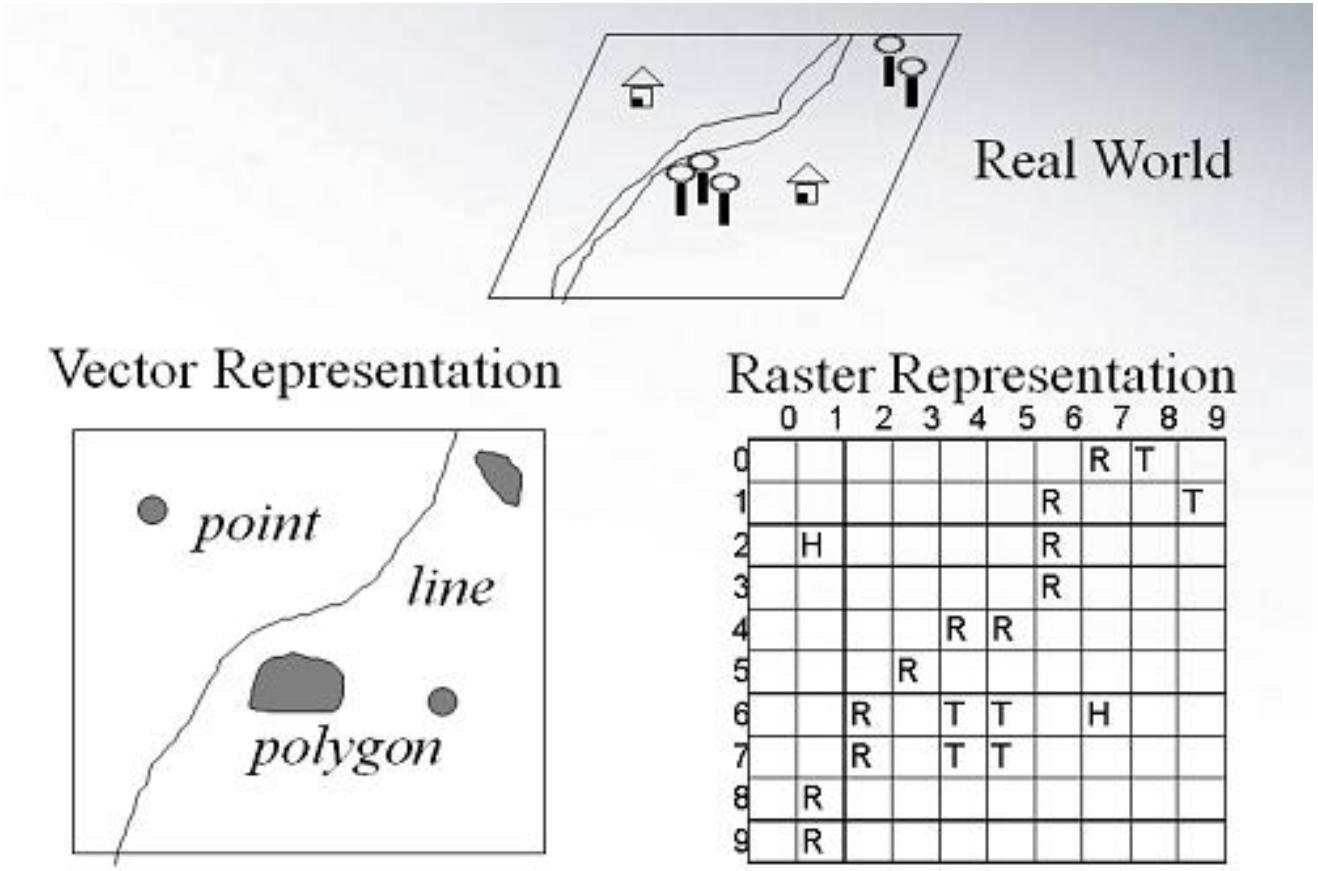


It is collected from various sources.....

- With technology, sources and methods have diversified
- Examples: ground based surveys, GPS, Satellite and Drone imagery, Aerial imagery, Crowdsourced data, digitizing paper files like plans and maps etc.
- We use GIS (technology and tools) to generate, edit, manage, analyze and visualize spatial data

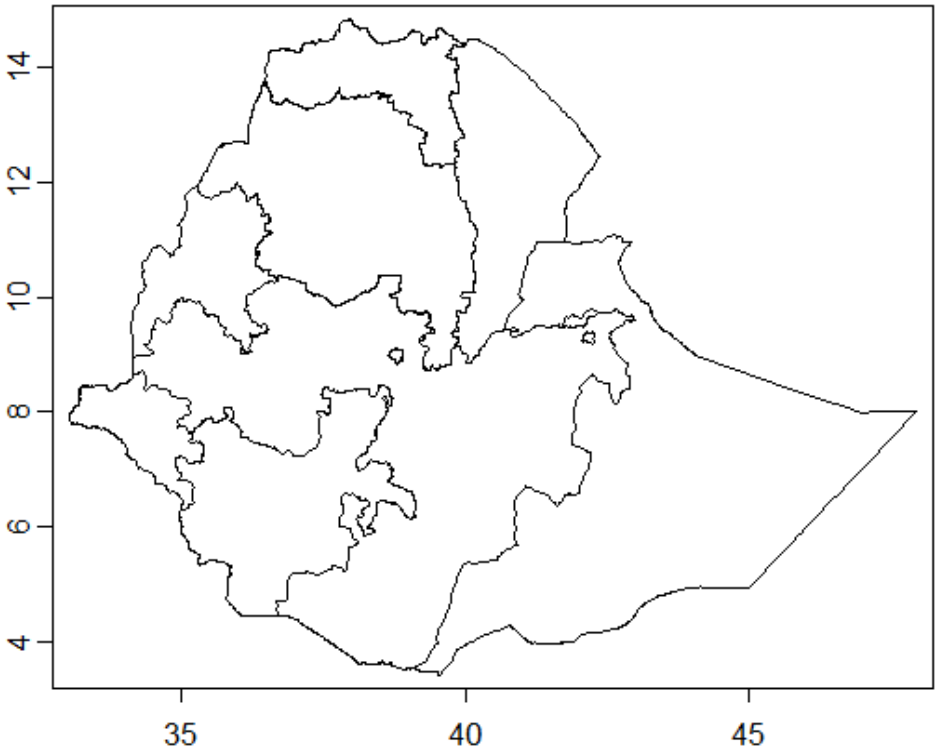


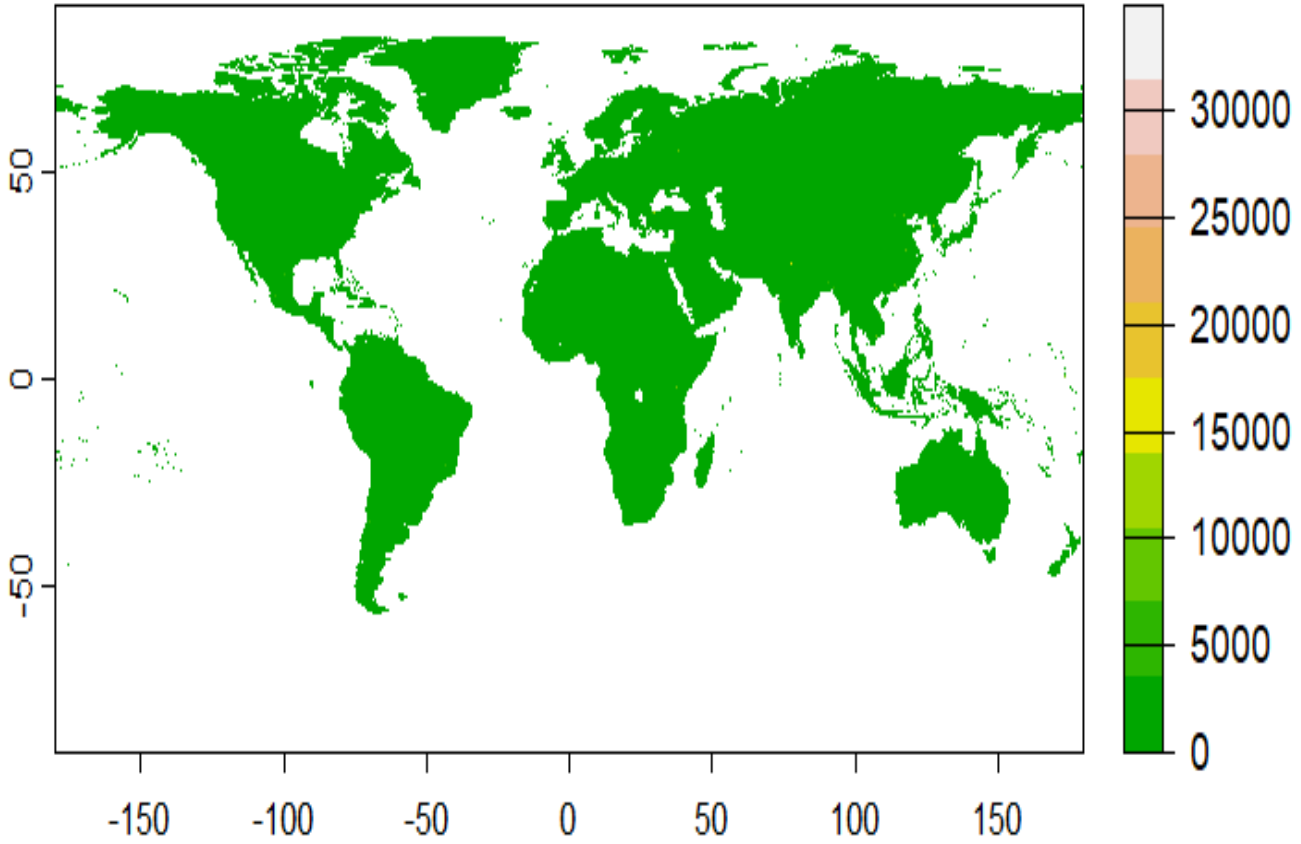
Spatial data types



- Represent the real world onto a surface
- Vector data: represented in points, lines & polygons
  - E.g. Town & city centers, rivers, roads, admin boundaries, buildings
- Raster: represented as a grid of regularly sized cells or pixels
  - Continuous data e.g temperature distribution, elevation etc.

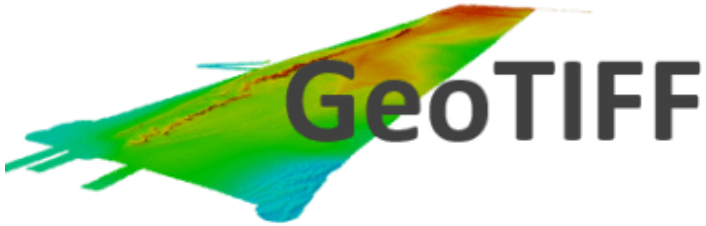
Spatial data types: vector & raster





Spatial data formats

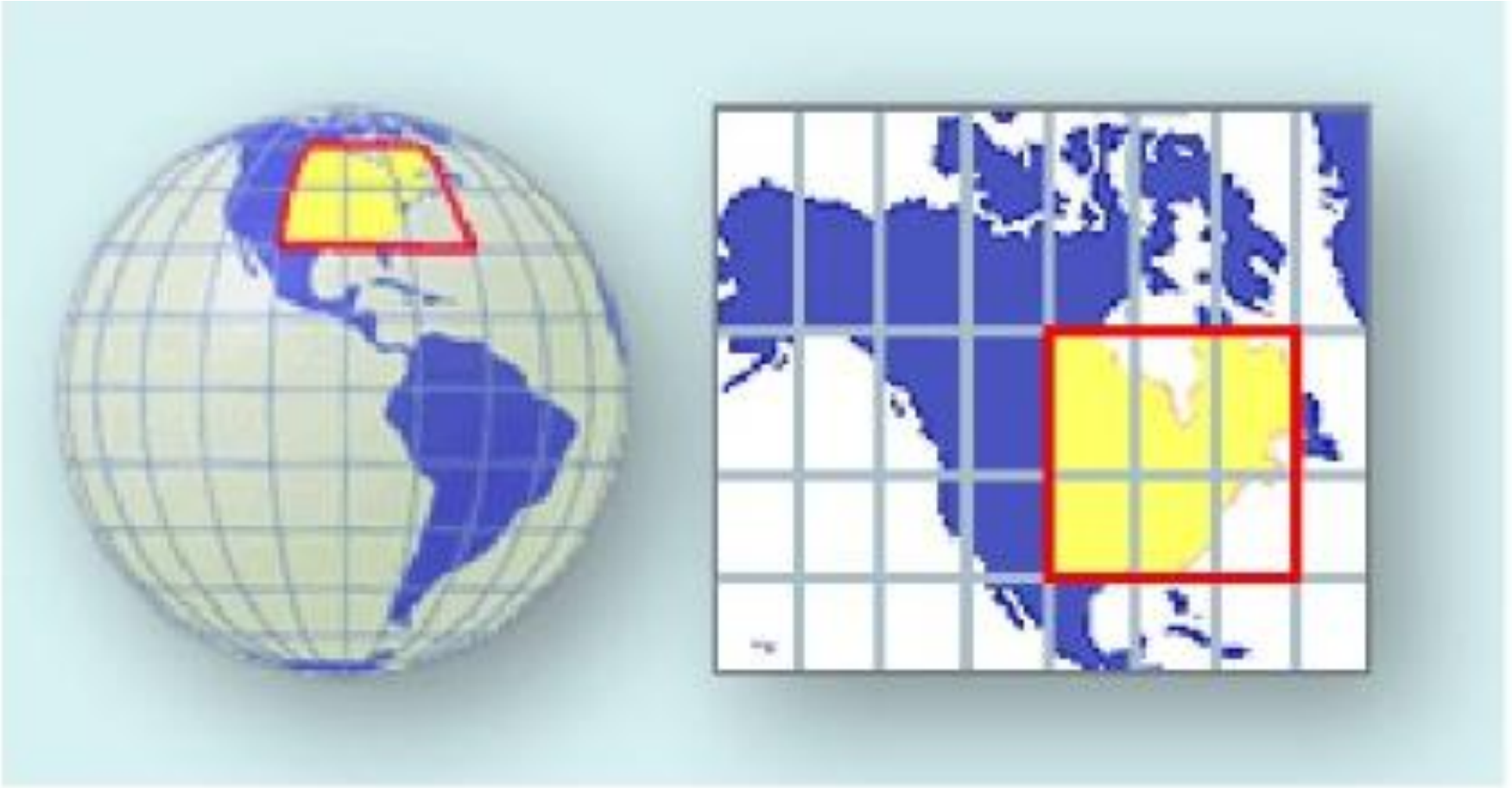
- Vector
  - Geopackage – more recent
  - Esri Shapefile – most common filetype. Made of 3 mandatory files
    - .shp stores the feature geometry
    - .shx is the shape index position
    - .dbf is the attribute data
    - .prj is the projection system metadata
    - .xml is the associated metadata
    - .sbn is the spatial index for optimizing queries
    - .sbx optimizes loading times
  - Others geojson, kml etc.
- Raster
  - GeoTIFF



Important Spatial Concepts

Map Projections

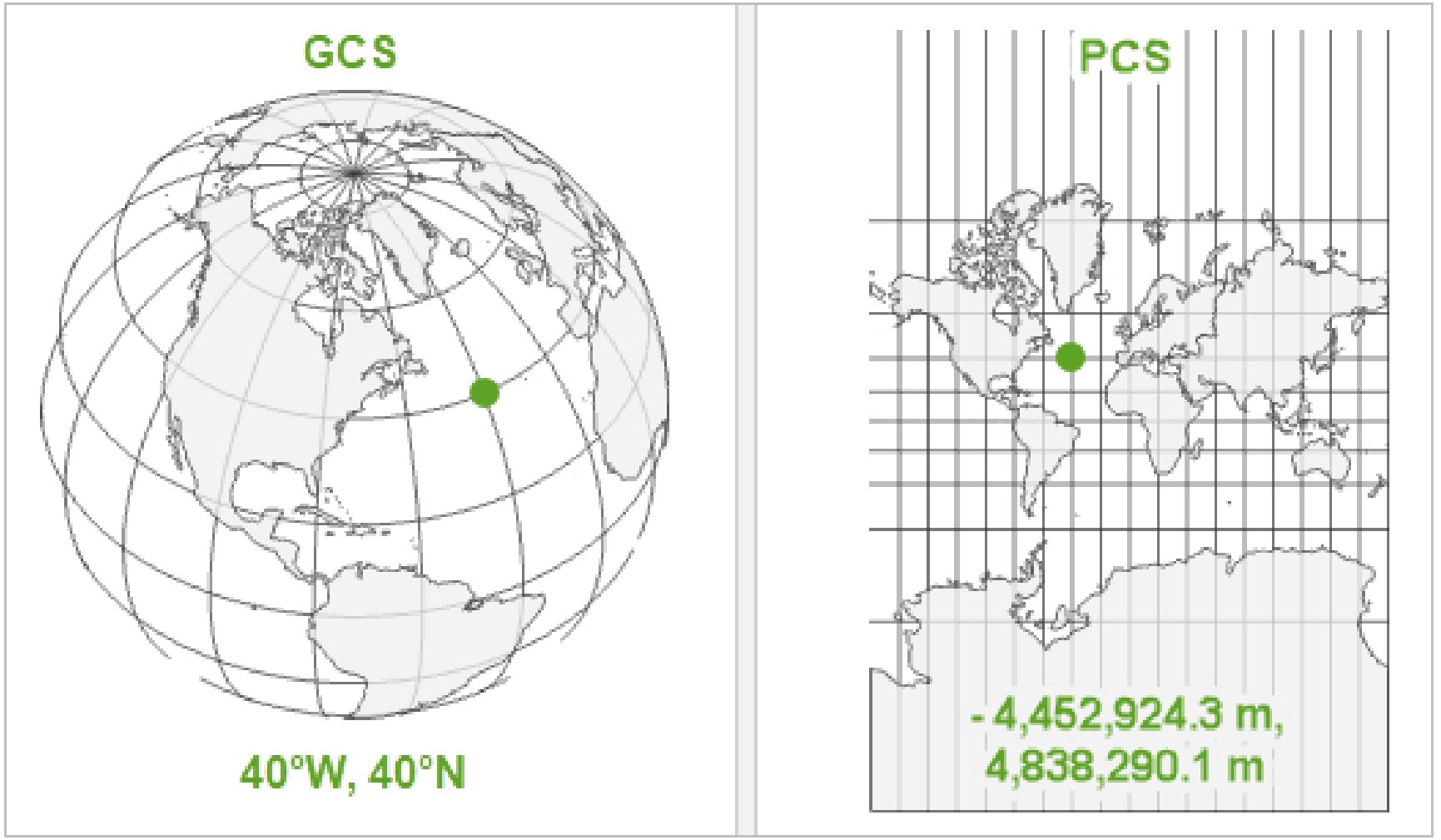
Method of representing part of the earth's surface on a plane surface



Coordinate Reference Systems

- CRS then defines how the two-dimensional map relates to real places on the earth.
- We can define CRS as systems that define how spatial data aligns with the earth's surface
- Divided into:
  - Geographic coordinate systems: They use degrees of latitude and longitude and describe a location on the earth's surface. The most popular is called WGS 84. Basically, shows you where you are on earth. Units are in angular degrees
  - Projected coordinate system tells us how to draw on a flat surface, like on a paper map or a computer screen. Units are linear e.g metres which enables measurements
- CRS are important as they ensure that data aligns correctly when overlaid.

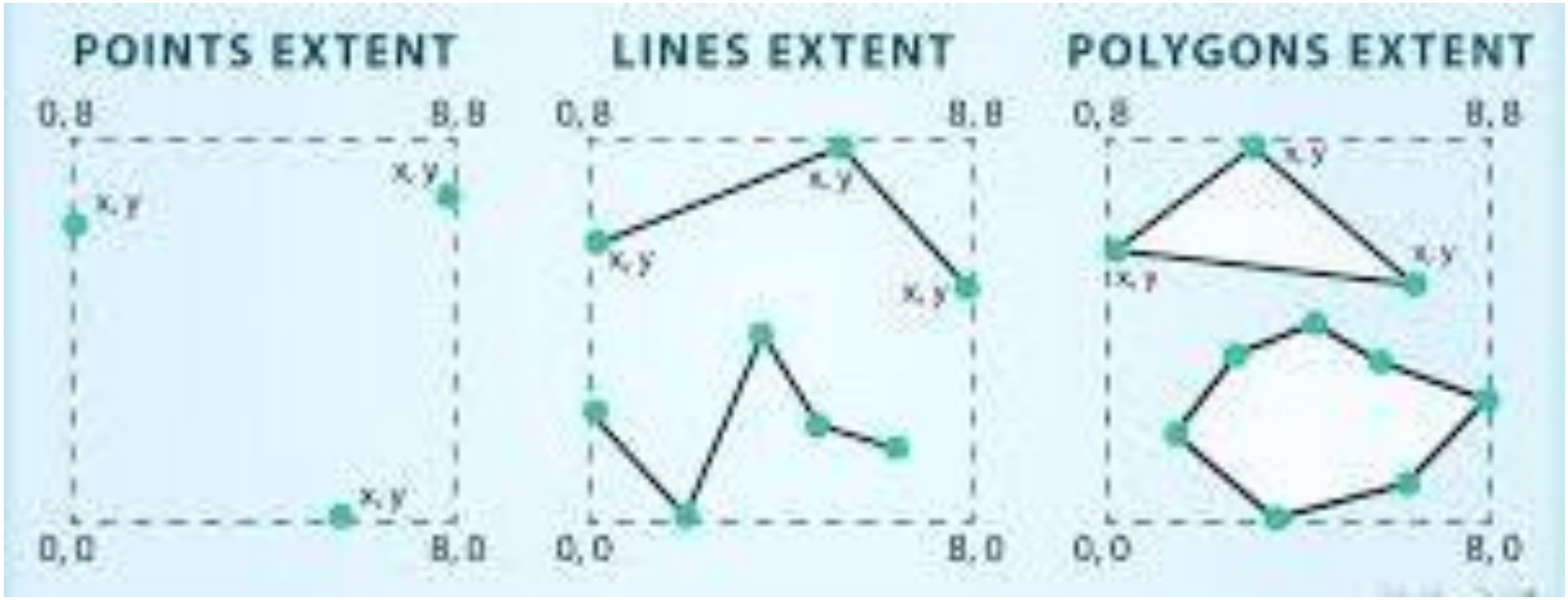
Coordinate Reference systems

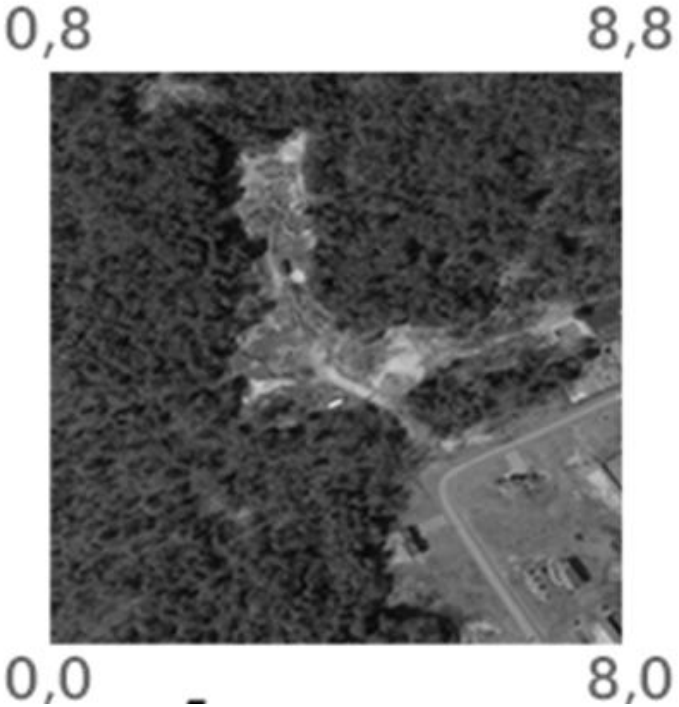


If you have the same dataset saved in two different projections, these two files won't line up correctly when rendered together.

Extent

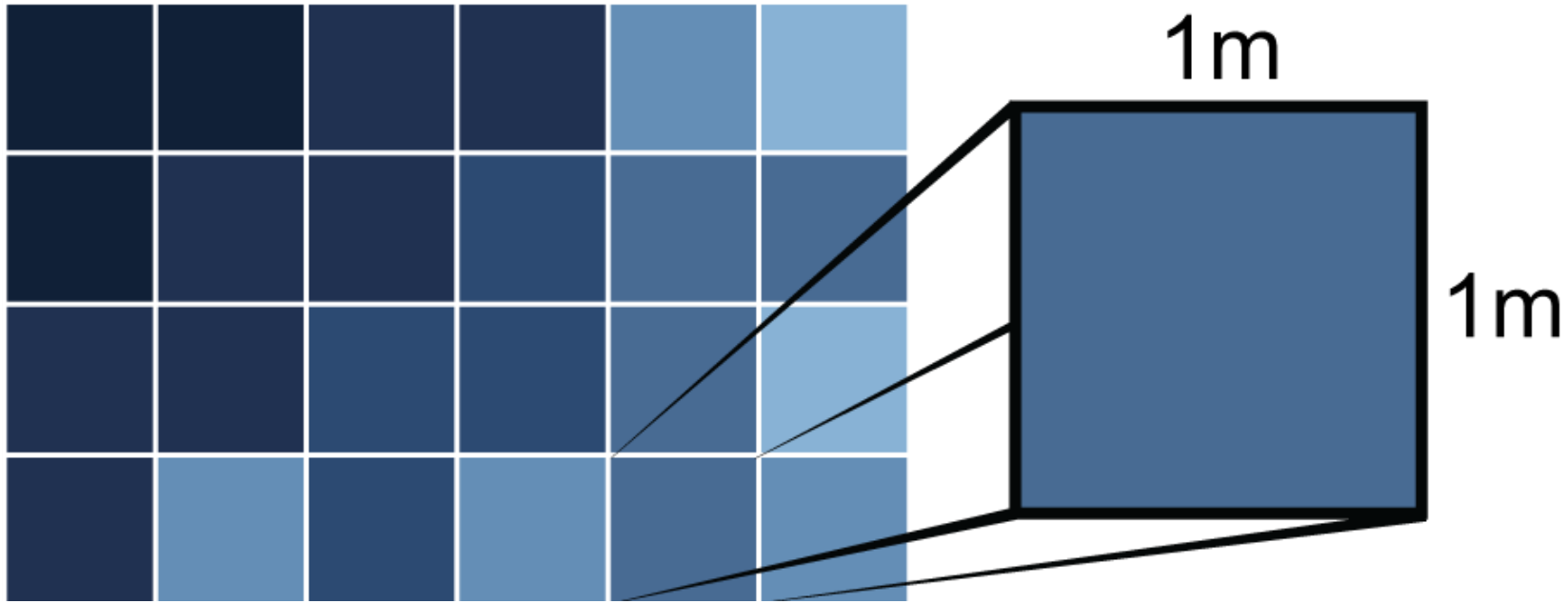
- The geographic boundary/coverage of spatial data (xmin, xmax, ymin, ymax).
  - Determines the area covered by spatial data.



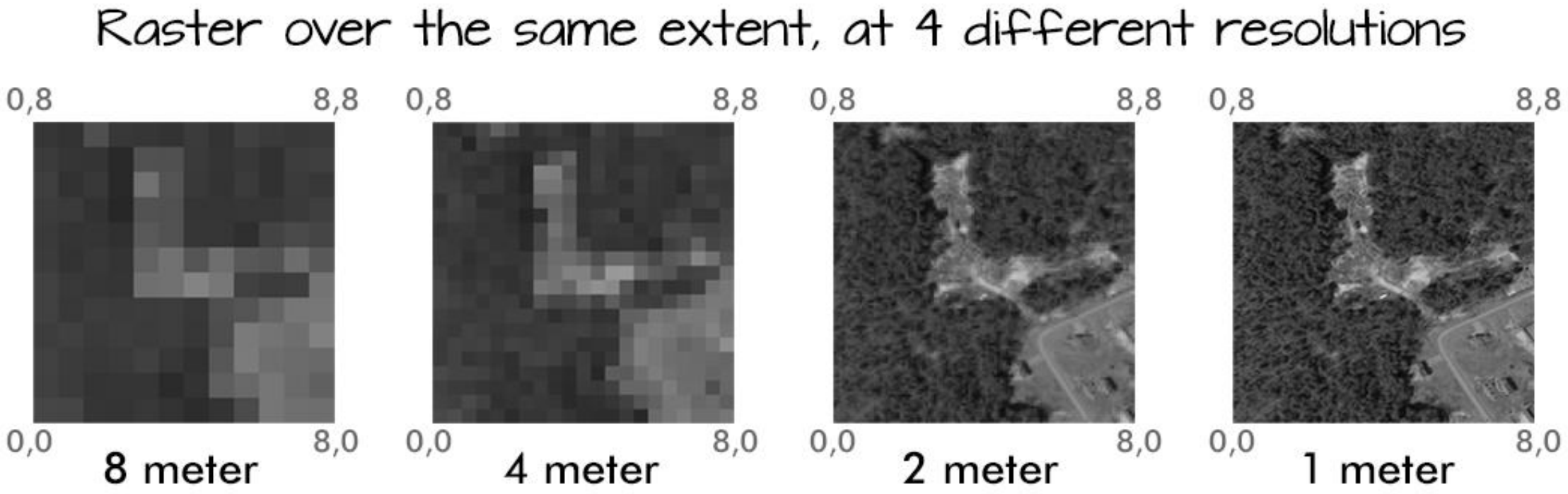


Raster Spatial resolution

The size of the area on the surface that each pixel covers is known as the spatial resolution of the image. For instance, an image that has a 1 m spatial resolution means that each pixel in the image represents a 1 m x 1 m area.



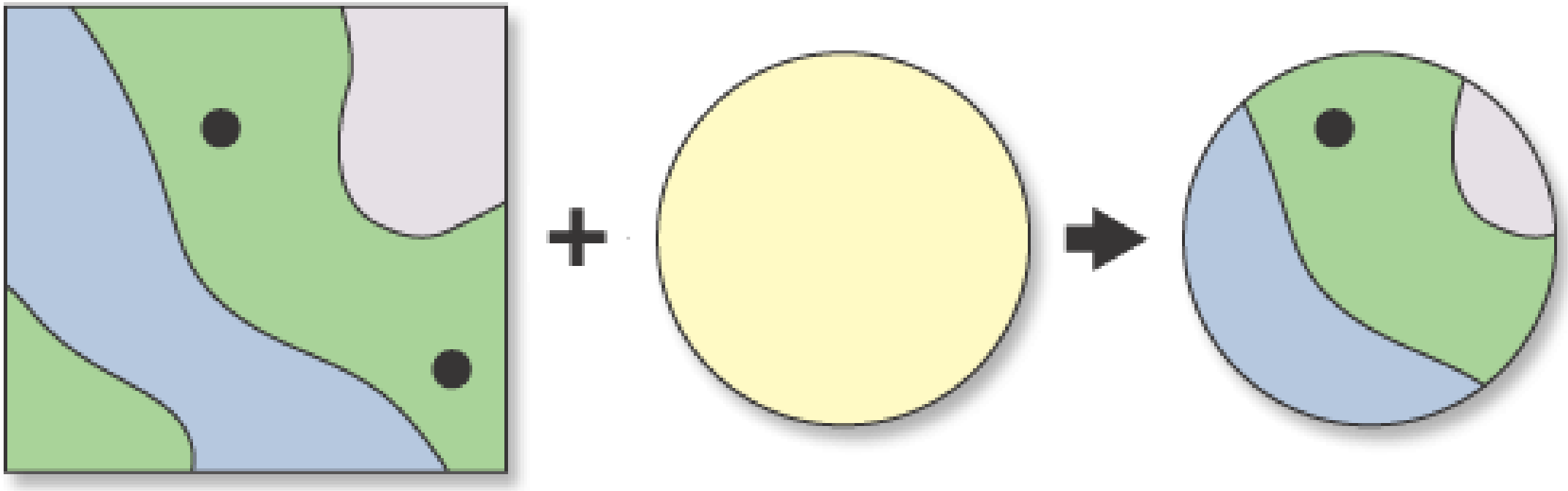
Raster Extent and Resolution



Basic Operations

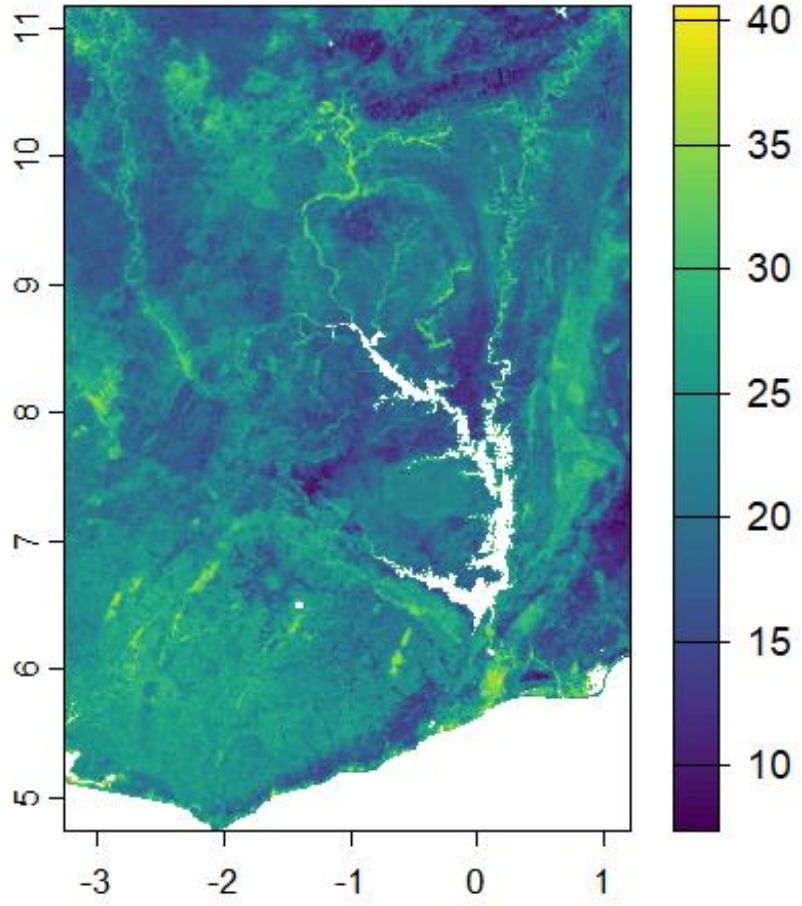
- Spatial data is often larger than we need it to be and we need to manipulate the data to get the desired output
- And we are often only interested in the portion of that data that is located inside our study region
- To get the piece of spatial data that we want for our maps or analysis we have to the select the data or crop it to our areas of interest

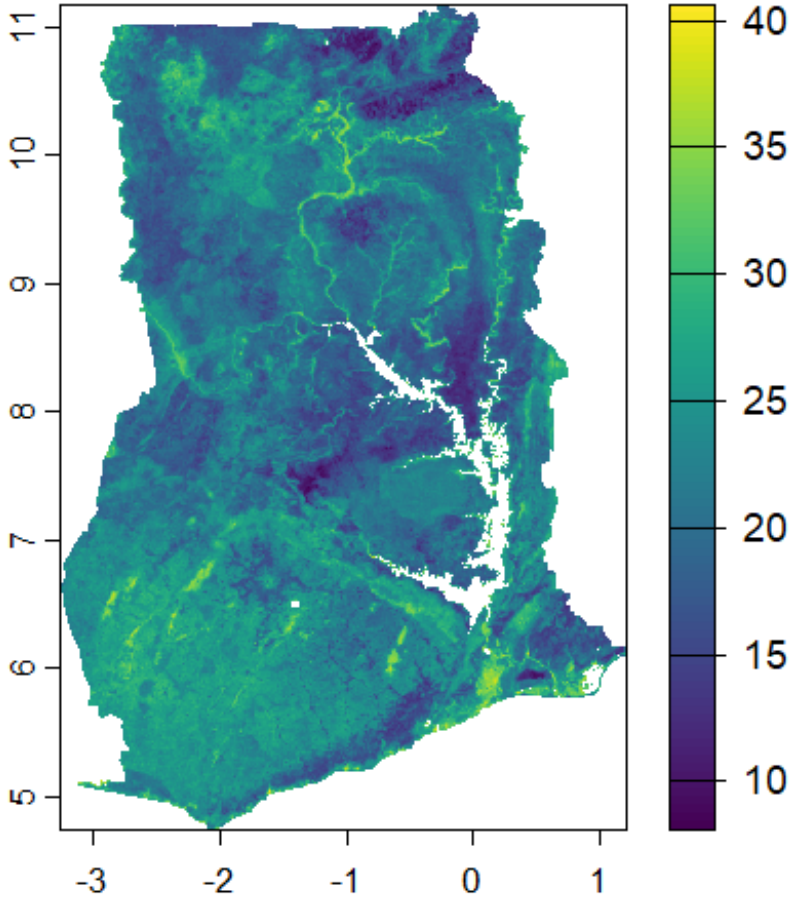
Basic Operations: Vector Clip



Raster Crop and Mask

Raster data is typically a large rectangular block of data so we need to crop to get to the desired output





R Practical Session

- Rspatial libraries
  - terra, sf
- Data sources – GADM, geodata package etc.
- Reading and writing vector and raster data
- Plotting spatial data

Resources

- Materials
  - Design principles for cartography
  - Map Projections & Coordinate Reference Systems
  - Cropping & masking data
- Data sources
  - GADM
  - Humanitarian Data Exchange
    - Water bodies, roads & admin boundaries
  - Geodata package in R
- Spatial analysis in R
  - rspatial.org

Thanks!

Map elements

- Frame line & neat lines
- Title
- Map area
- Insets
- Legend
- Data source
- Scale
- Map orientation
- Use tmap and ggplot2 packages to create visually appealing maps