UNIVERSITETET I OSLO

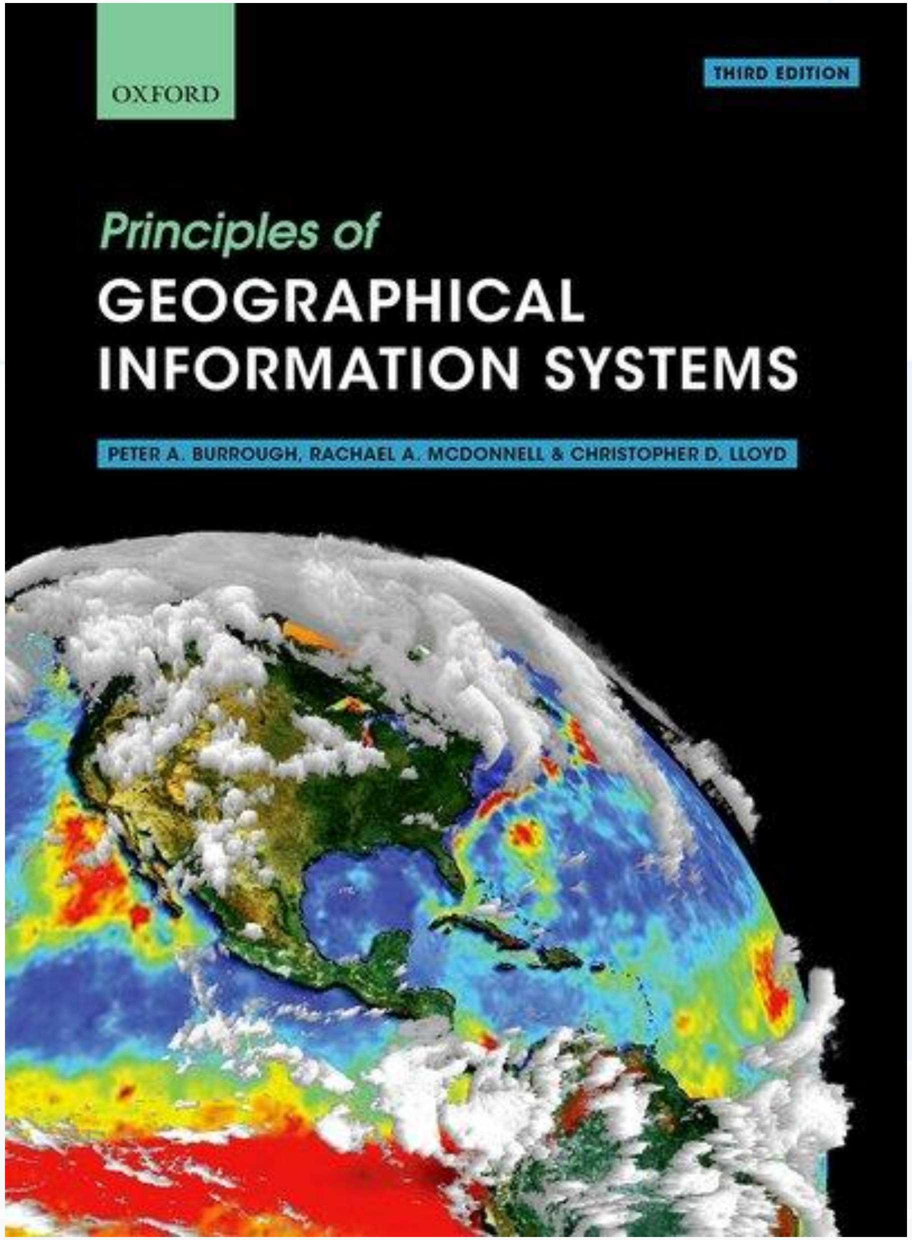
GEO3460 – Geografiske informasjonssystemer (GIS) og geografisk datainnsamling – vår 2025

GEO3460 - Geographical Information Systems (GIS) and Geographical Data Acquisition - spring 2025

Spatial Analysis Data structure Topology

Luc Girod (luc.girod@geo.uio.no)

Based on the lectures by Livia Piermattei and Martin Lund



- Reference textbook:

"*Geographical information systems*"

Learning Objectives

Basic concepts for describing GIS data, raster and vector data, their attribute and level of measurement, and the file format.

1

GIS

- Spatial – temporal dimension
- What is a GIS
- GIS application

2

Digital representation of spatial data

- Geographic data
- Continuous vs discrete data
- Raster, Vector & Topology
- Attributes

3

File Format

- Raster
- Vector
- Software

Today's topics

GIS

1

GIS

- Spatial – temporal dimension
- What is a GIS
- GIS application

2

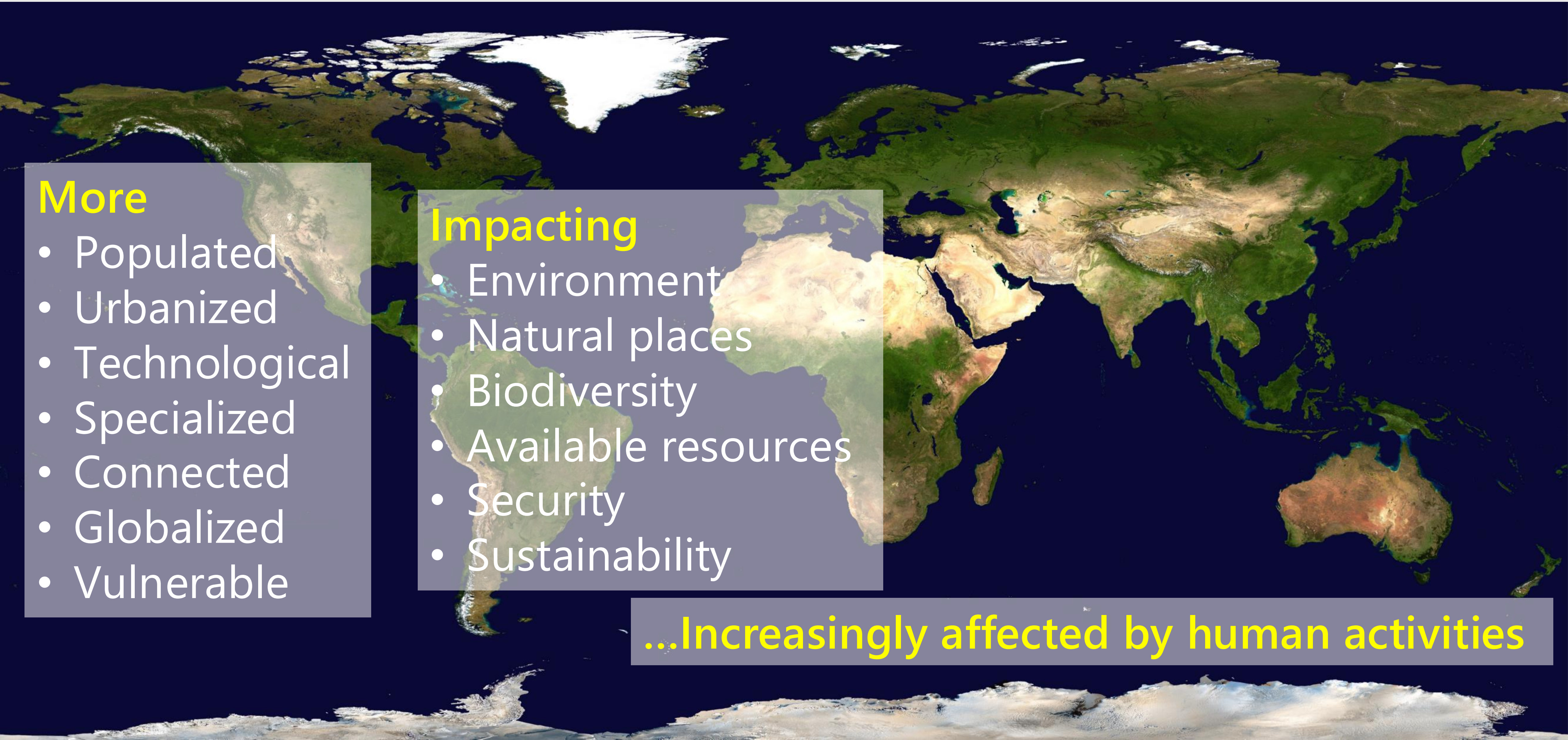
Digital representation of spatial data

- Geographic data
- Continuous vs discrete data
- Raster, Vector & Topology
- Attributes

3

File format

Our World is Becoming . .

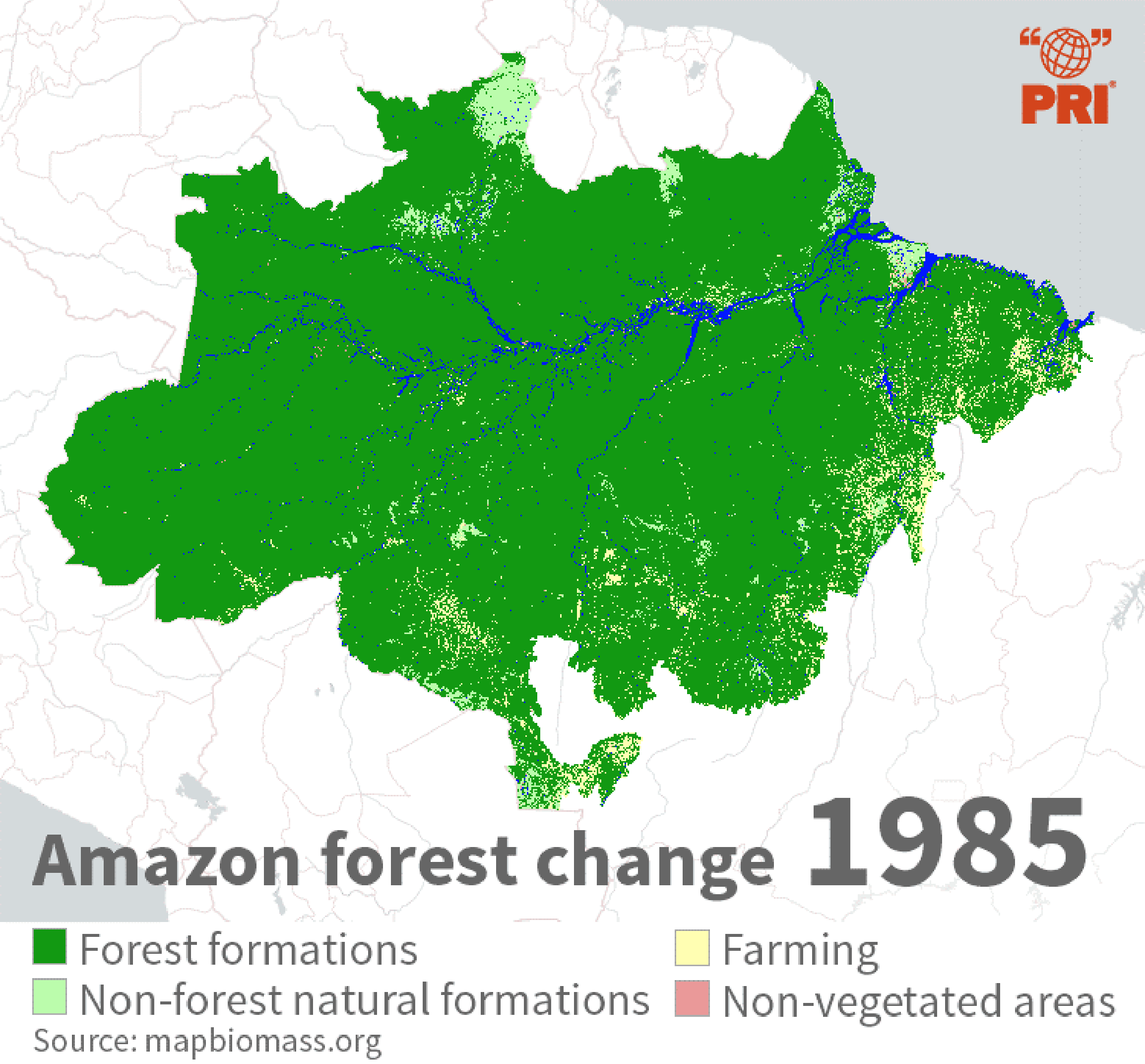


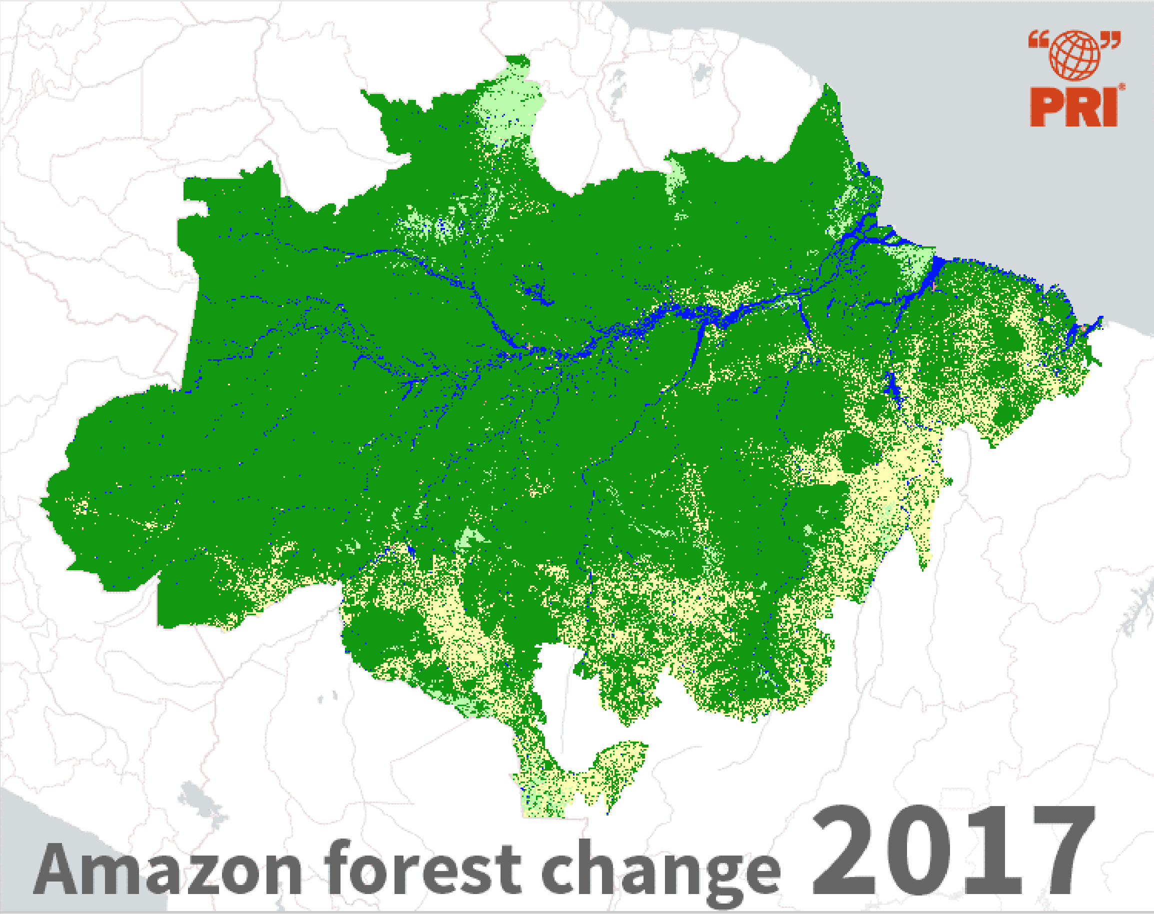
More
- Populated
- Urbanized
- Technological
- Specialized
- Connected
- Globalized
- Vulnerable

Impacting
- Environment
- Natural places
- Biodiversity
- Available resources
- Security
- Sustainability

…Increasingly affected by human activities

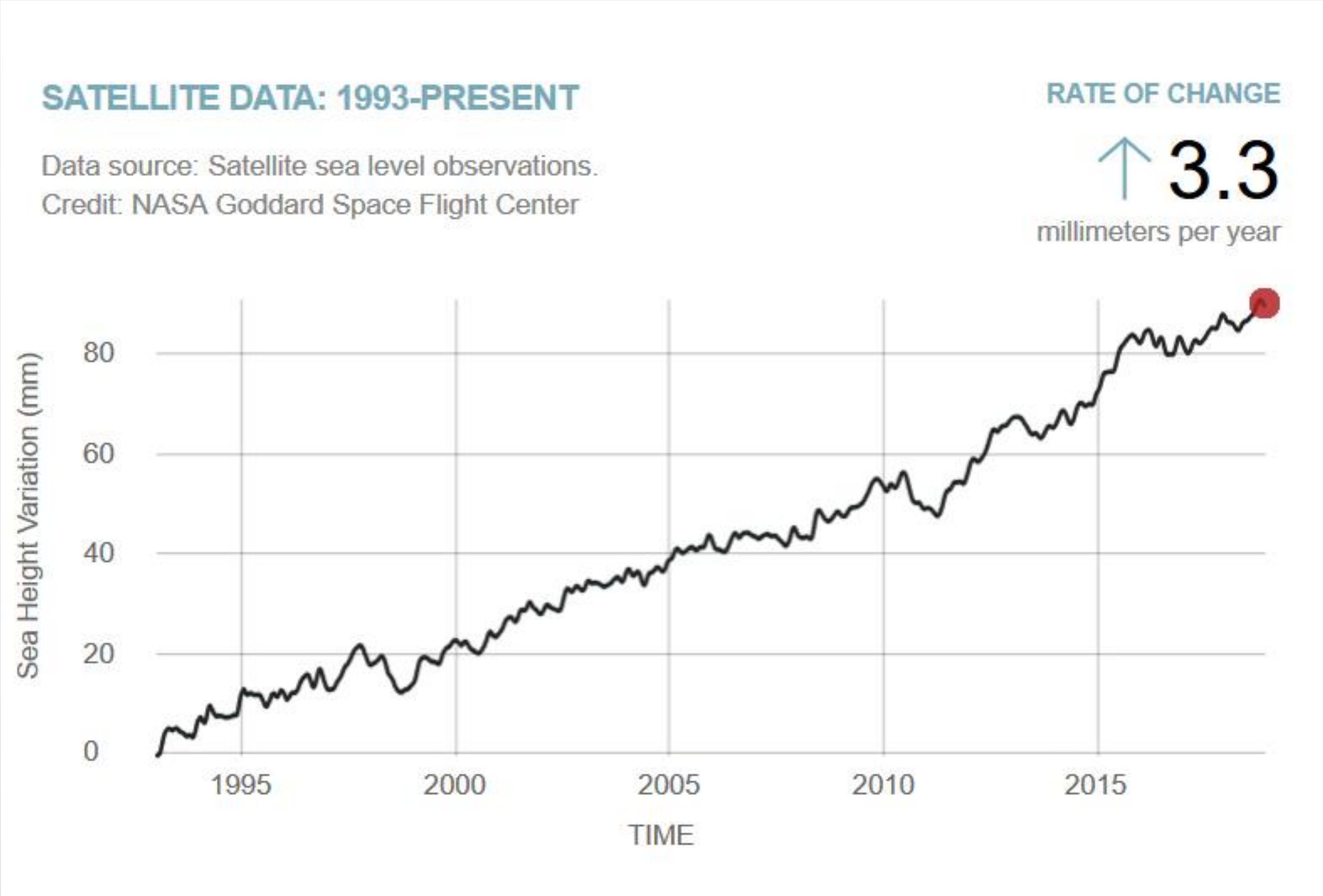
The evidence is everywhere

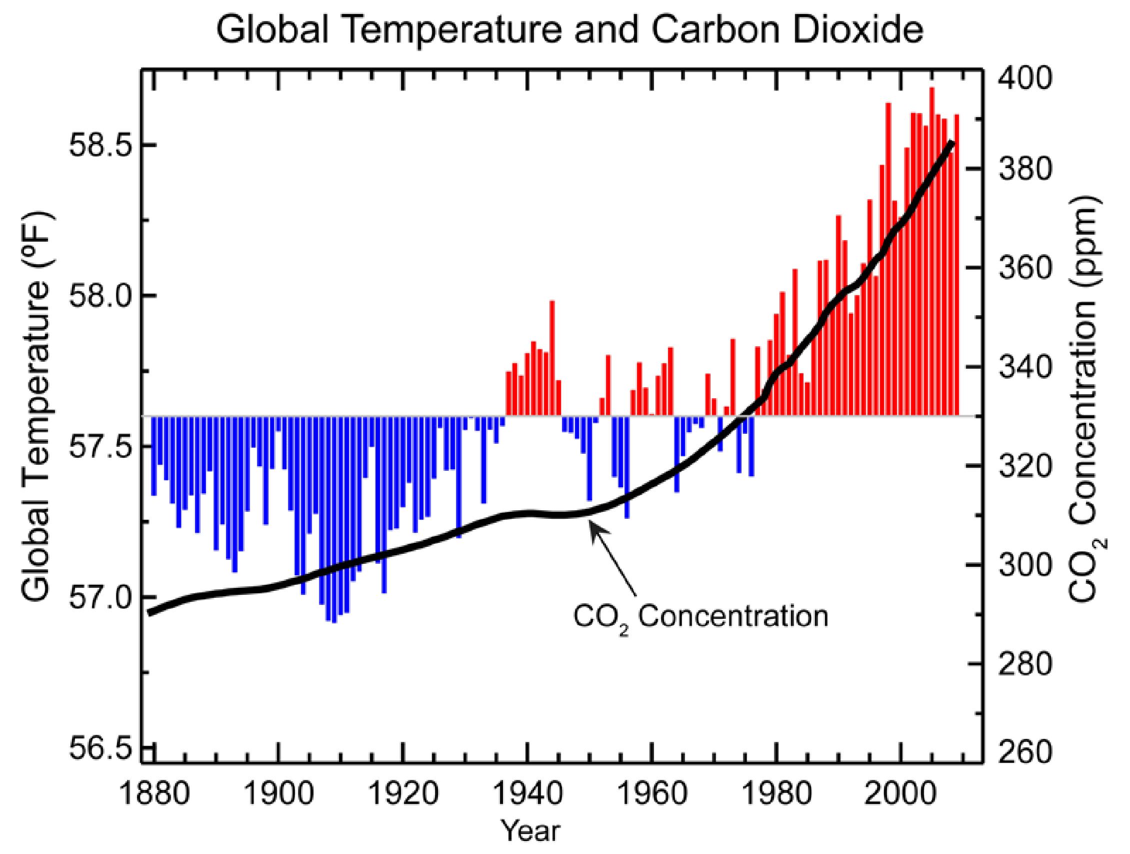




…We need to learn how to better manage our world

The evidence is everywhere



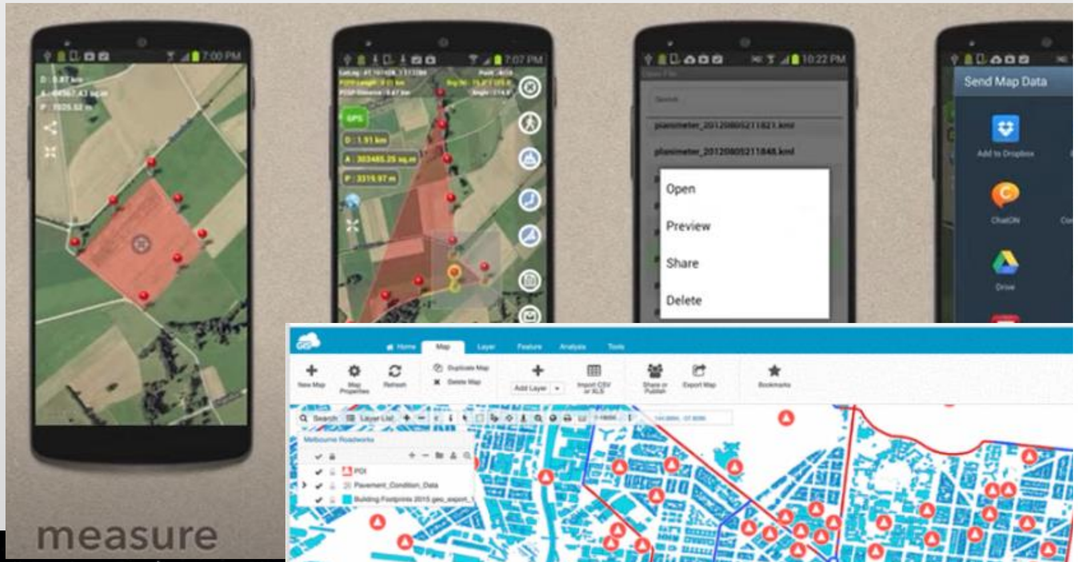


Source: NOAA (National Oceanographic and Atmospheric Administration)

…We need to learn how to better manage our world

We are part of a Geodata society

- Satellite imagery
- GPS
- Mobile mapping
- Real time monitoring
- Online GIS services
- Web portals
- . . .



...with more geospatial information

Some fundamental observations

Different processes dominate at different spatial scales

Some example of landscape scales:

- Continent, ocean basin, climatic zone (~10,000,000 km^2)
- Mountain range (~1,000,000 km^2)
- Isolated sea (~100,000 km^2) (Norway: 385,207 km^2)
- Massif, group of related landforms (~10,000 km^2)
- River valley (~1,000 km^2)
- Individual mountain, volcano, small valleys (~100 km^2)
- Hillslopes, stream channels, estuary, individual glacier (~10 km^2)
- Gully, small glaciers, rock glaciers (~1 km^2)
- Meter-sized features

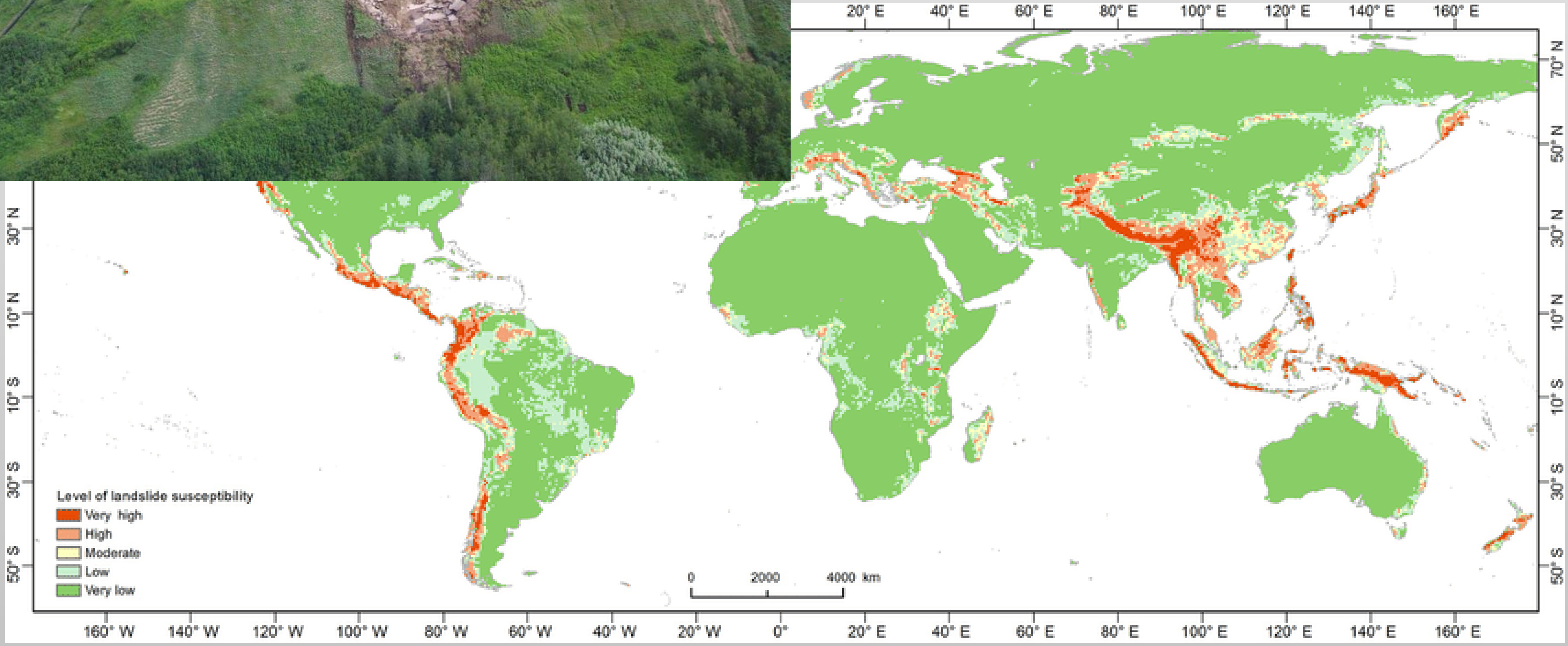
... and temporal scales

Some fundamental observations

Landslide (local scale)

Spatial and

temporal-dimension



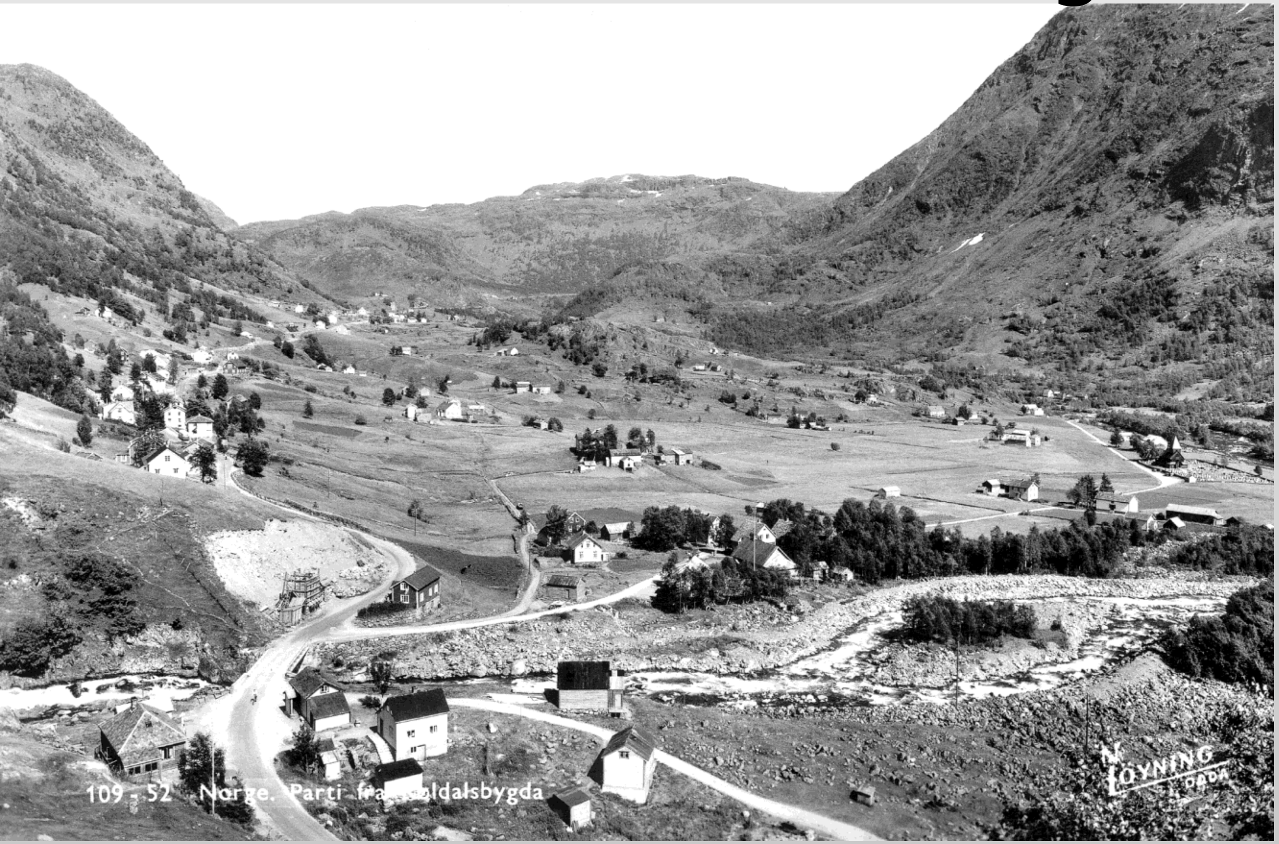
From local to global scale

(Lin et al., 2017.)

Some fundamental observations

Changes in ~50 years

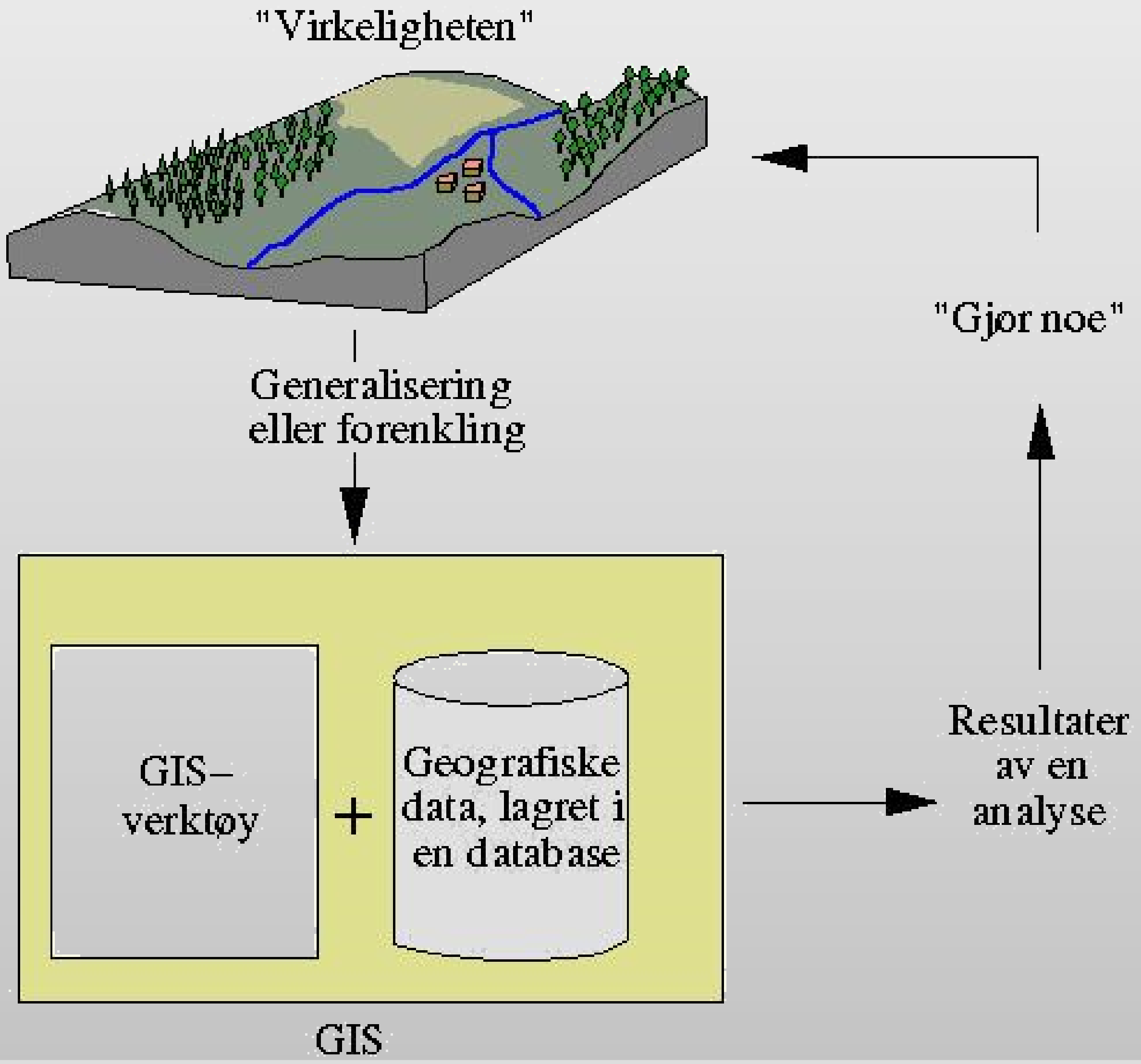
Spatial and temporal-dimension



Source: N. Løyning. Røldalsbygda, Odda. 1952

Source: Oskar Puschmann, NIBIO. Røldal, Odda. 2003

So What is GIS???



G: Geographic: Interest in SPATIAL IDENTITY or locality of certain entities ON, UNDER or ABOVE the SURFACE of the EARTH.

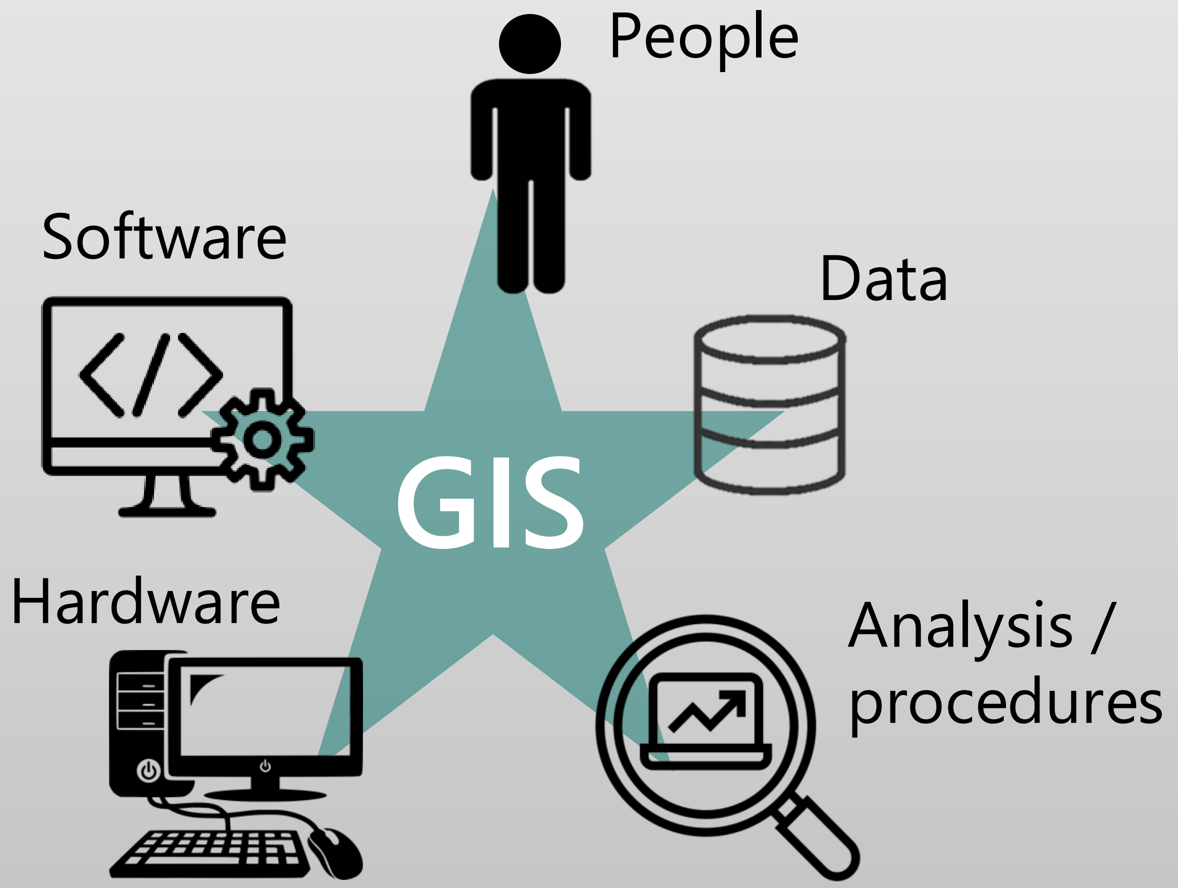
I: Information: Data, processes, phenomena, facts are interpreted to CREATE INFORMATION that is useful for decision-making.

S: System: STAFF, COMPUTER HARDWARE and PROCEDURES, which can produce the information required for decision-making (data collection, processing, and presentation).

So What is GIS???

"A powerful set of tools for collecting, storing, retrieving at will, transforming, and displaying spatial data from the real world for a particular set of purposes"
- Burrough et al. 2015

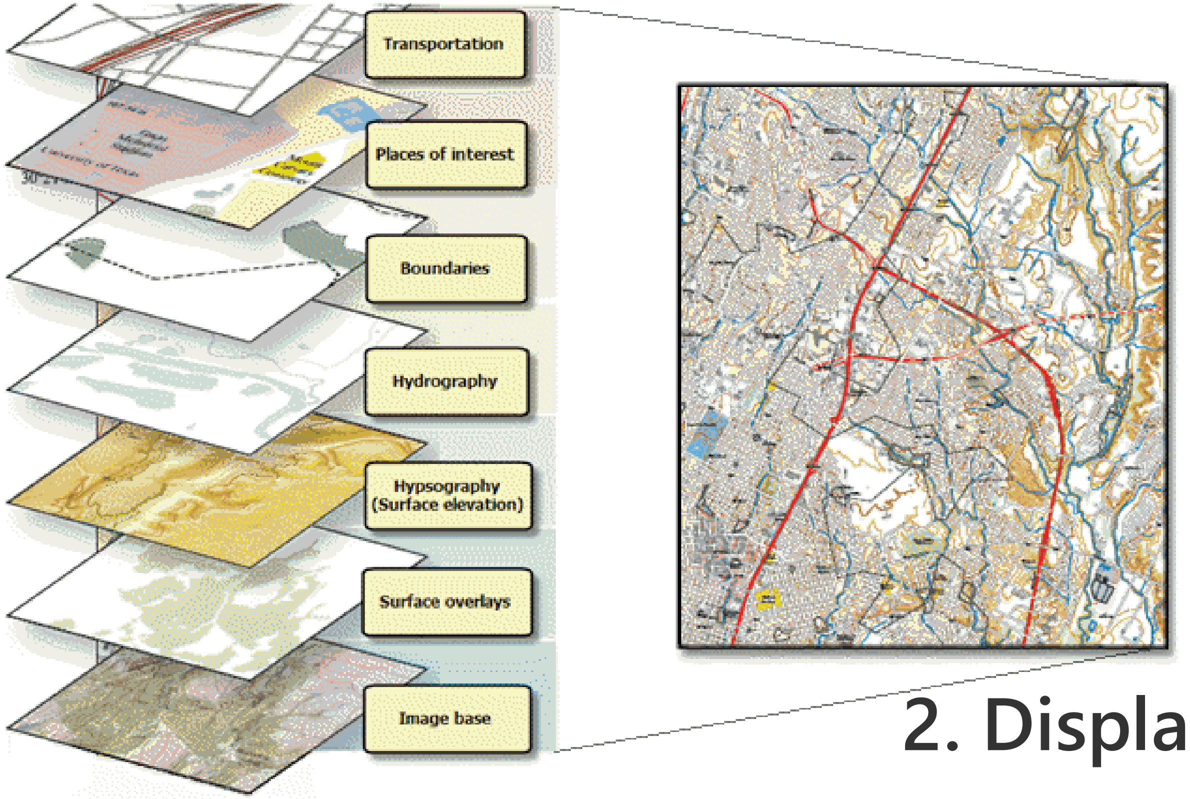
An integration of these five components



So What is GIS???

Answer questions by comparing different layer of data

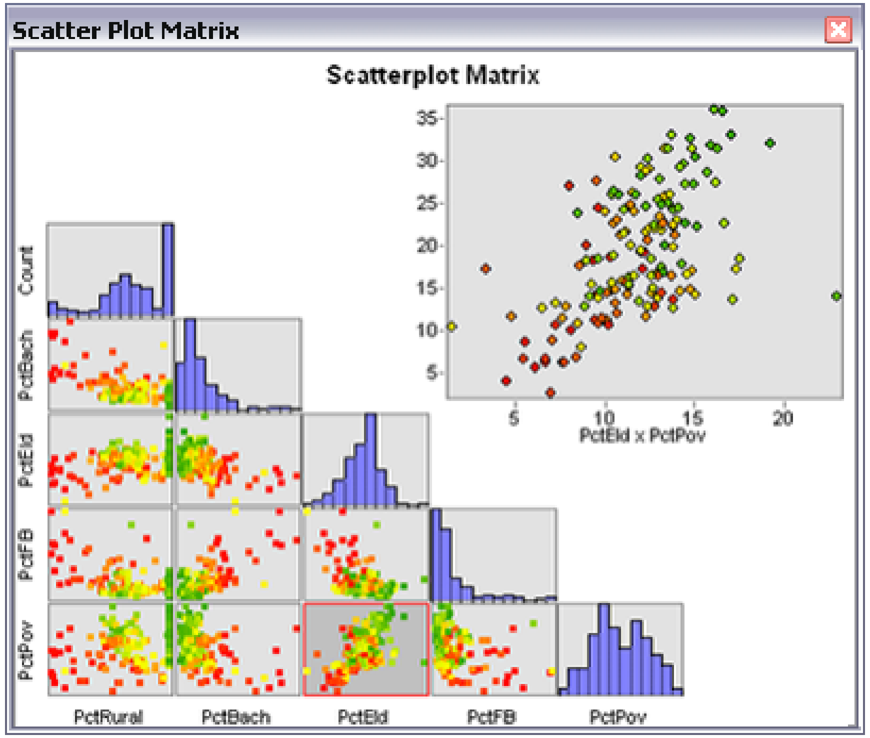
1. Organize spatial data



2. Display

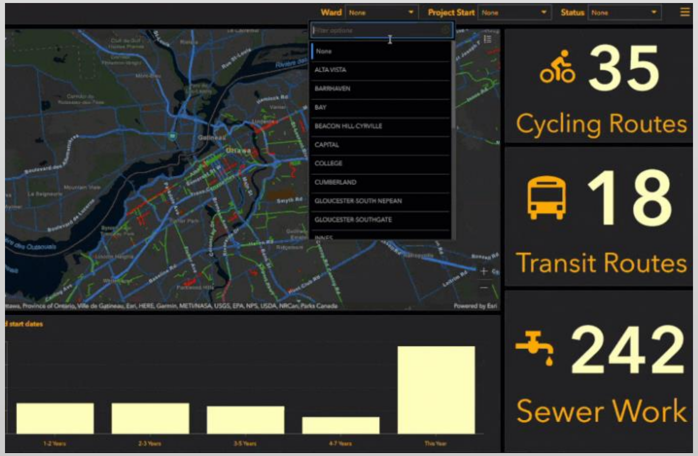
A graphical explanation of the concept of *layer*

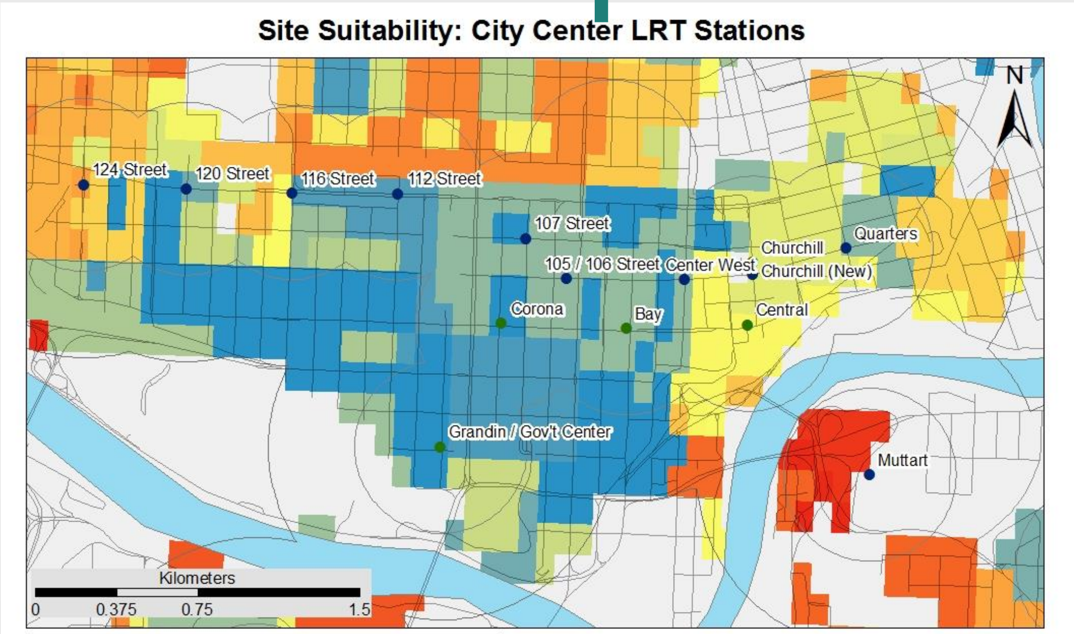
3. Graphs, report



GIS use: Planning & economic development

- Urban areas, smart city
- Mobility
- Conservation
- Land use, land cover
- Web-based suitability modelling
- Site selection
- Customer analysis
- Real time data
- . . .



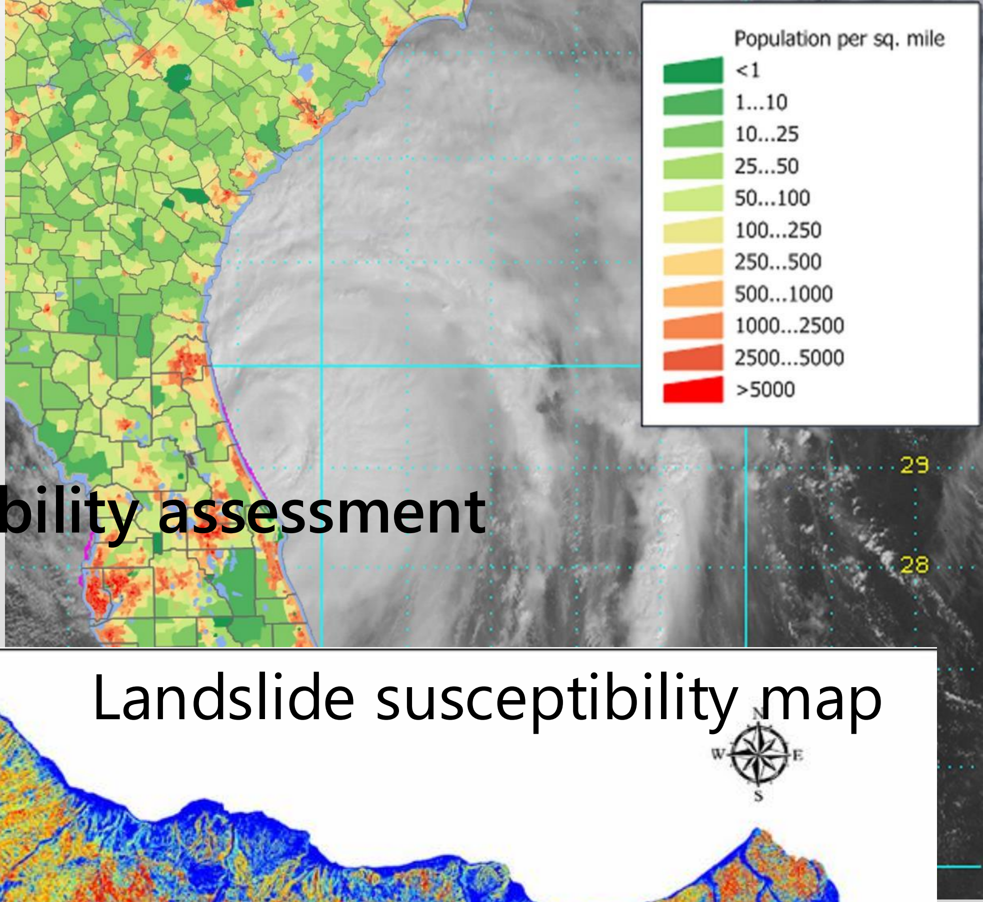


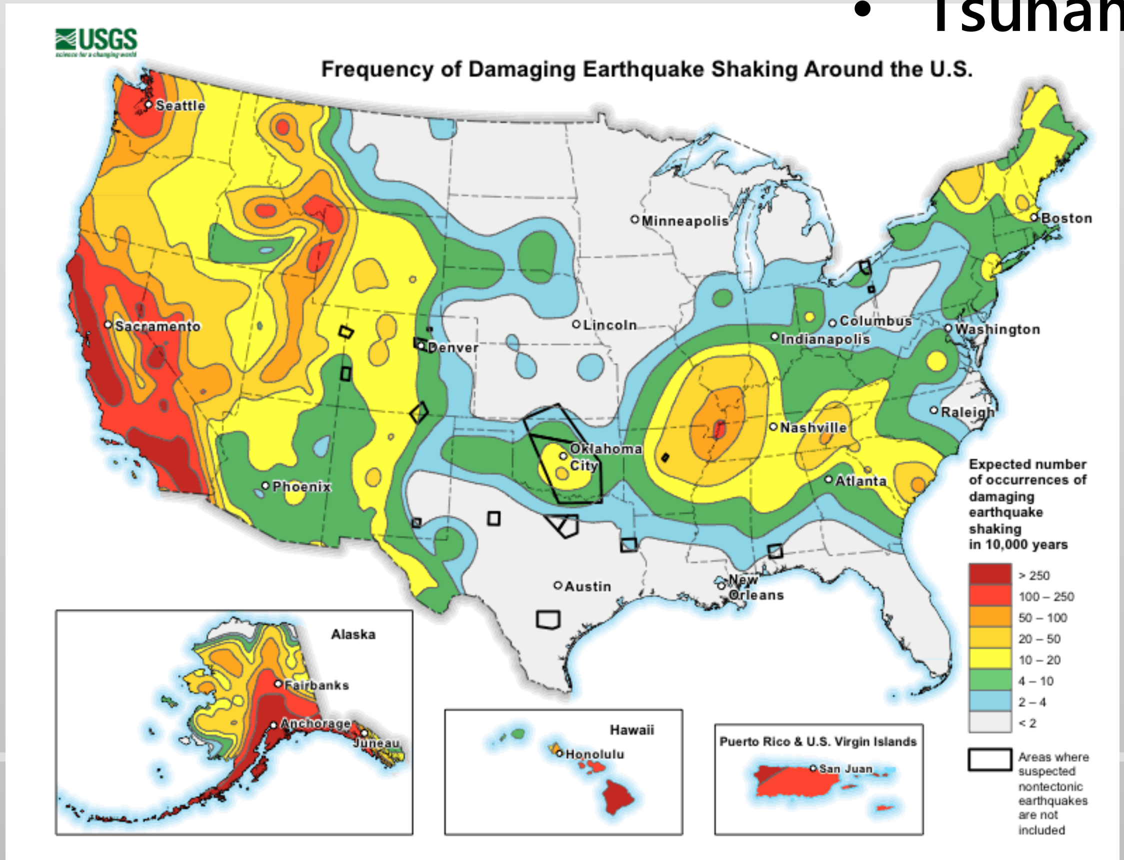
GIS use: Natural hazard

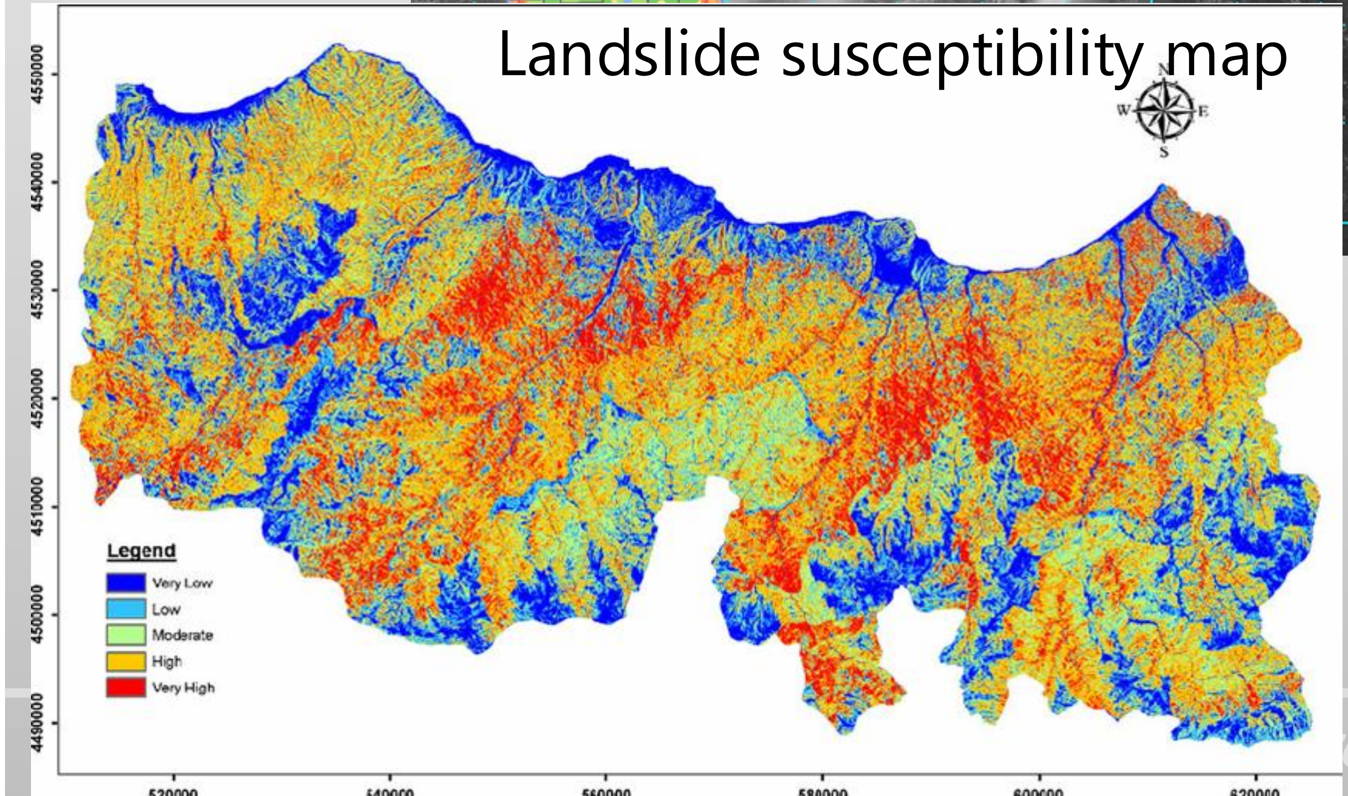
- Monitoring
- Damage assessment
- Modelling
- Real time tracking

Phenomena:

- Hurricane
- Earthquake
- Landslides, slope stability assessment
- Tsunami

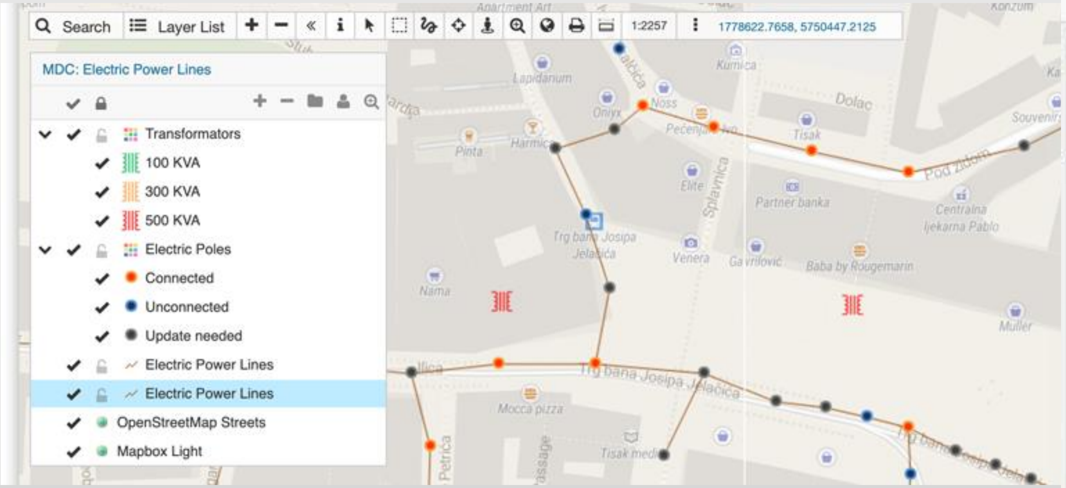


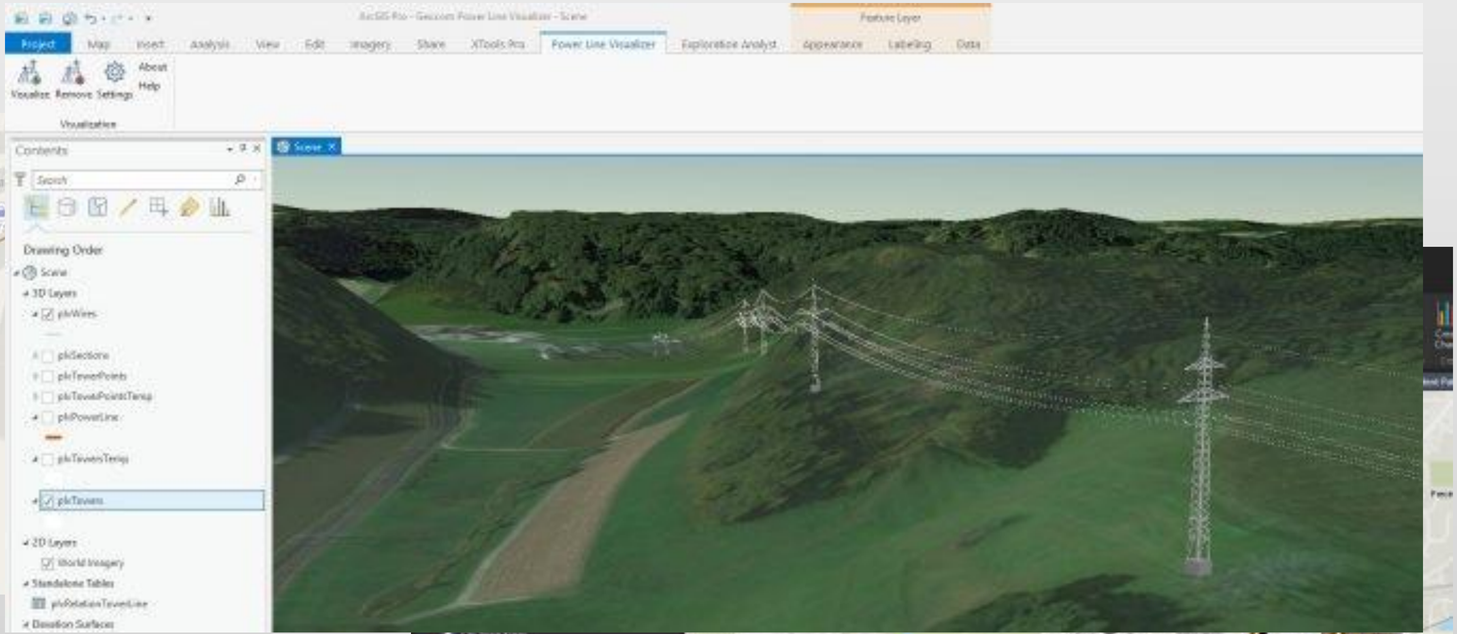




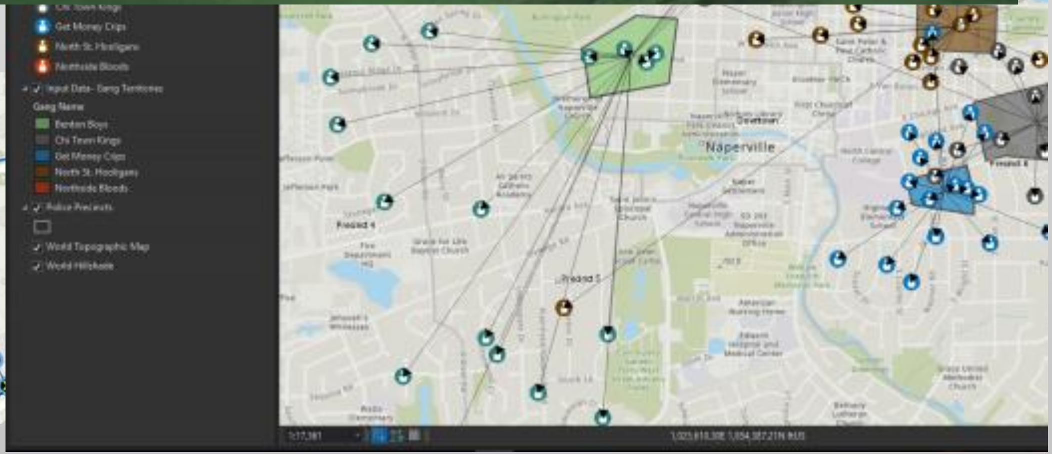
Landslide susceptibility map

GIS use: Managing Utilities network, Infrastructure



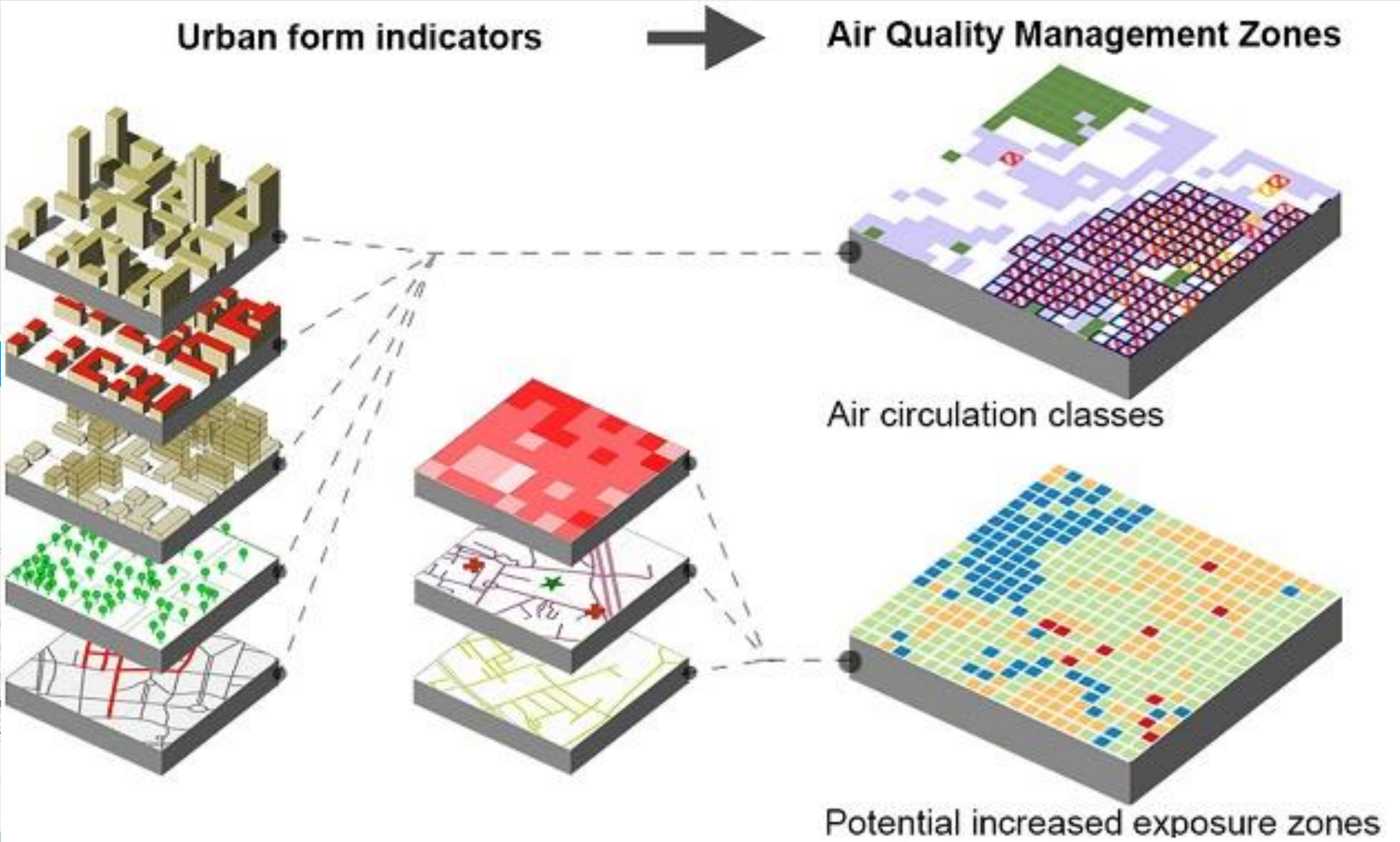


- Electricity network, power line
- Water distribution
- Agriculture
- Property
- Crime analysis
- . . .



GIS use: Human health & Environment

- Epidemiology
- Air pollution monitoring
- Contamination
- . . .



Digital representation of spatial data

1

GIS

- Spatial-temporal
- What is a GIS
- GIS applications

2

Digital representation of spatial data

- Geographic data
- Continuous vs discrete data
- Raster, Vector & Topology
- Attributes

3

File format

Basic GIS questions

- **What** is it?
  - What is its **relation** to other **entities**?
- **Where** is it?
- **When** is it?

Basic GIS questions

Each geographical object has 3 basic components:

- **What?**
  - Phenomena (with attribute)
  - Measured values
    - Quantitative (value with rank), ex. height
    - Qualitative (characteristic, class), ex. land use
- **Where?**
  - Coordinates (2D, 3D), projection, coordinate system
- **When? (Time)**
  - Snapshot (e.g. topographic map)
  - Time object (e.g. glacier front in different time periods)
  - Cyclic time ("repetitive" phenomena, e.g. season)

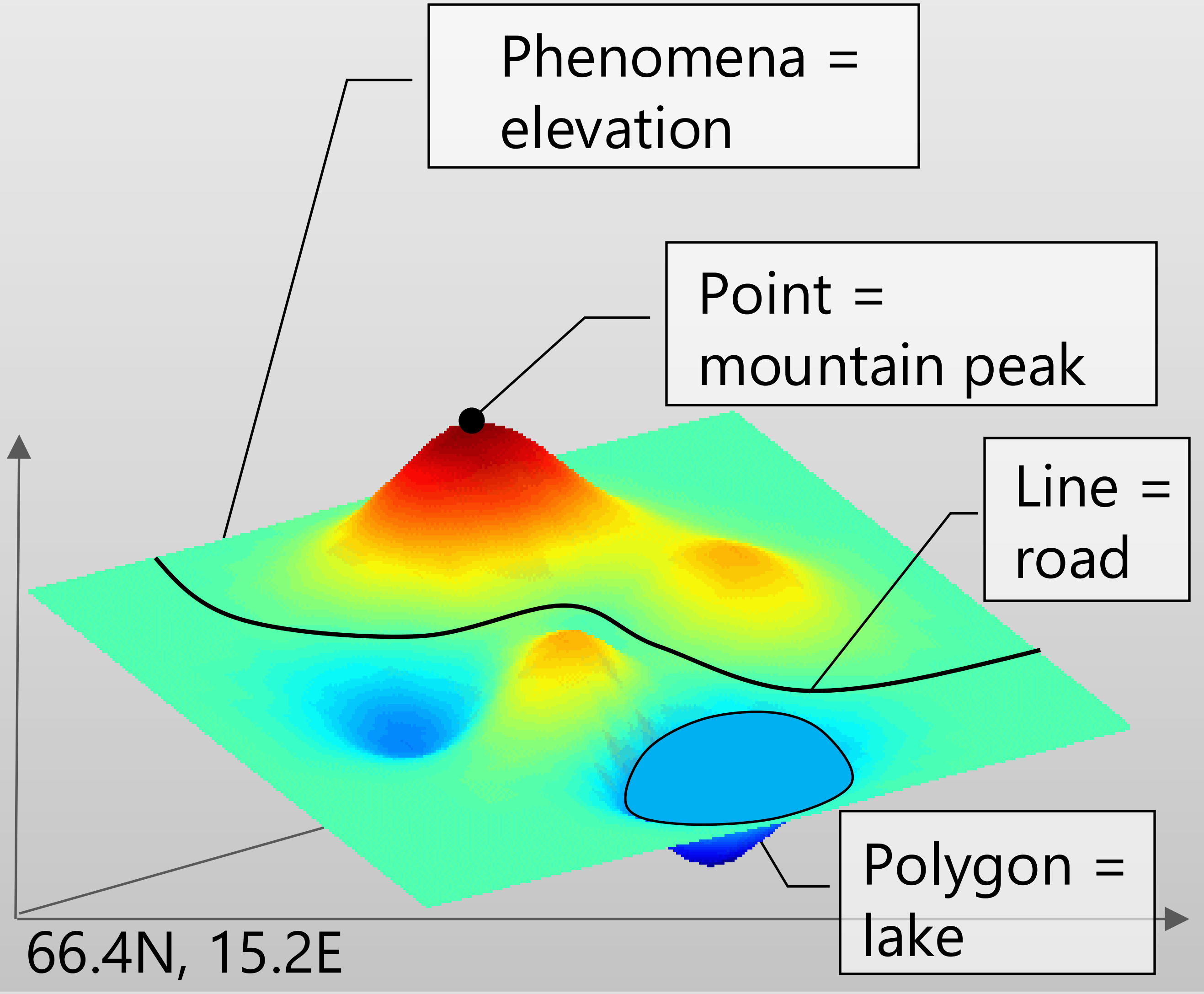
Basic characteristics of geographic data

How do we represent the real world and geographical phenomena?

Data in a GIS represent a simplified view of physical entities or phenomena

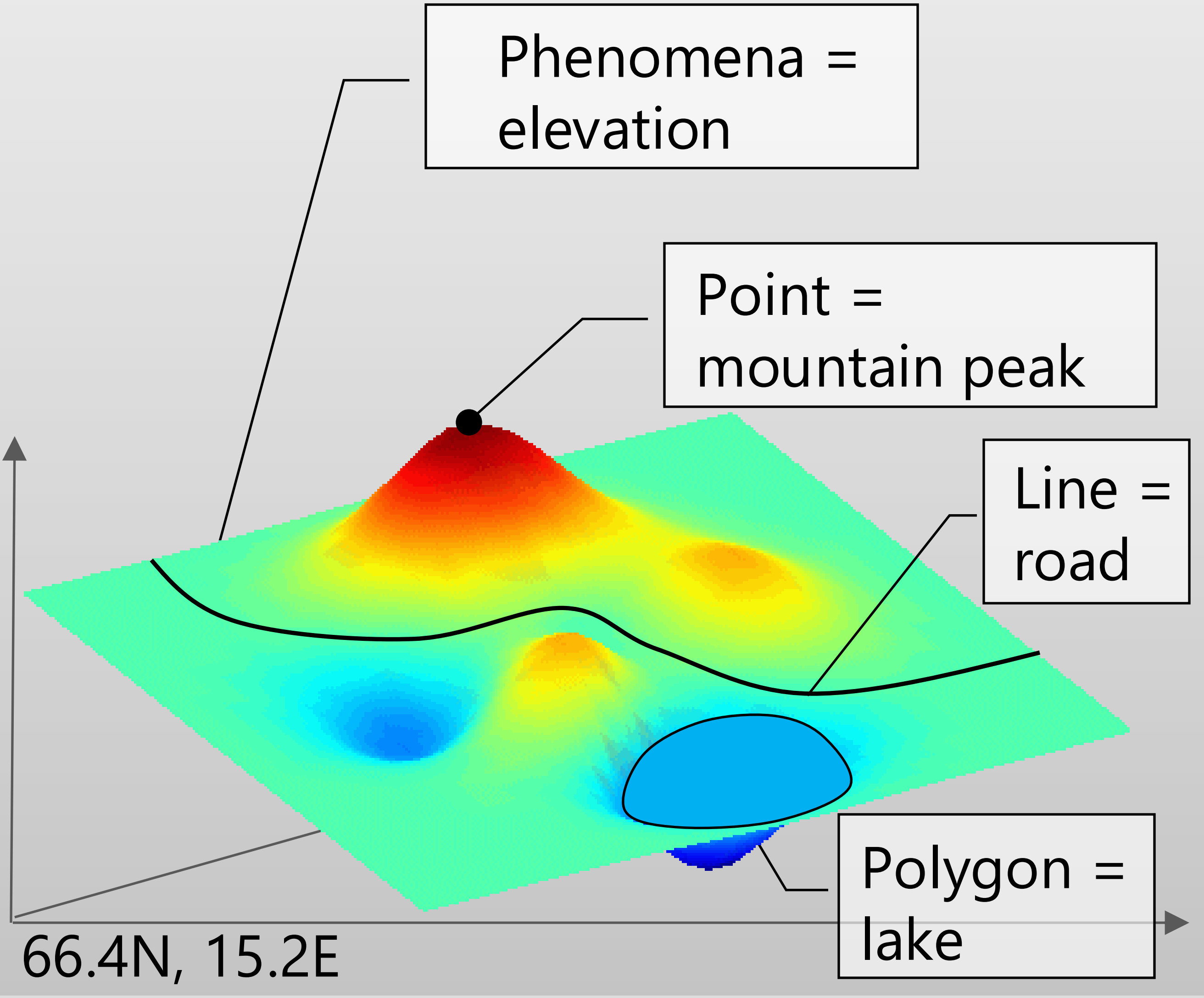
Basic characteristics of geographic data

- **Geographic phenomena** in the real world exist as:
  - Objects/entities (**discrete**):
    - Points, Lines, Polygons
  - **Continuous** fields



Basic characteristics of geographic data

- **Geographic phenomena** in the real world exist as:
  - Objects/entities (discrete):
    - Points, Lines, Polygons
  - Continuous fields
- **Spatial data have two distinct components**
  - Place (geographical location)
  - Attribute (properties)



Basic characteristics of geographic data

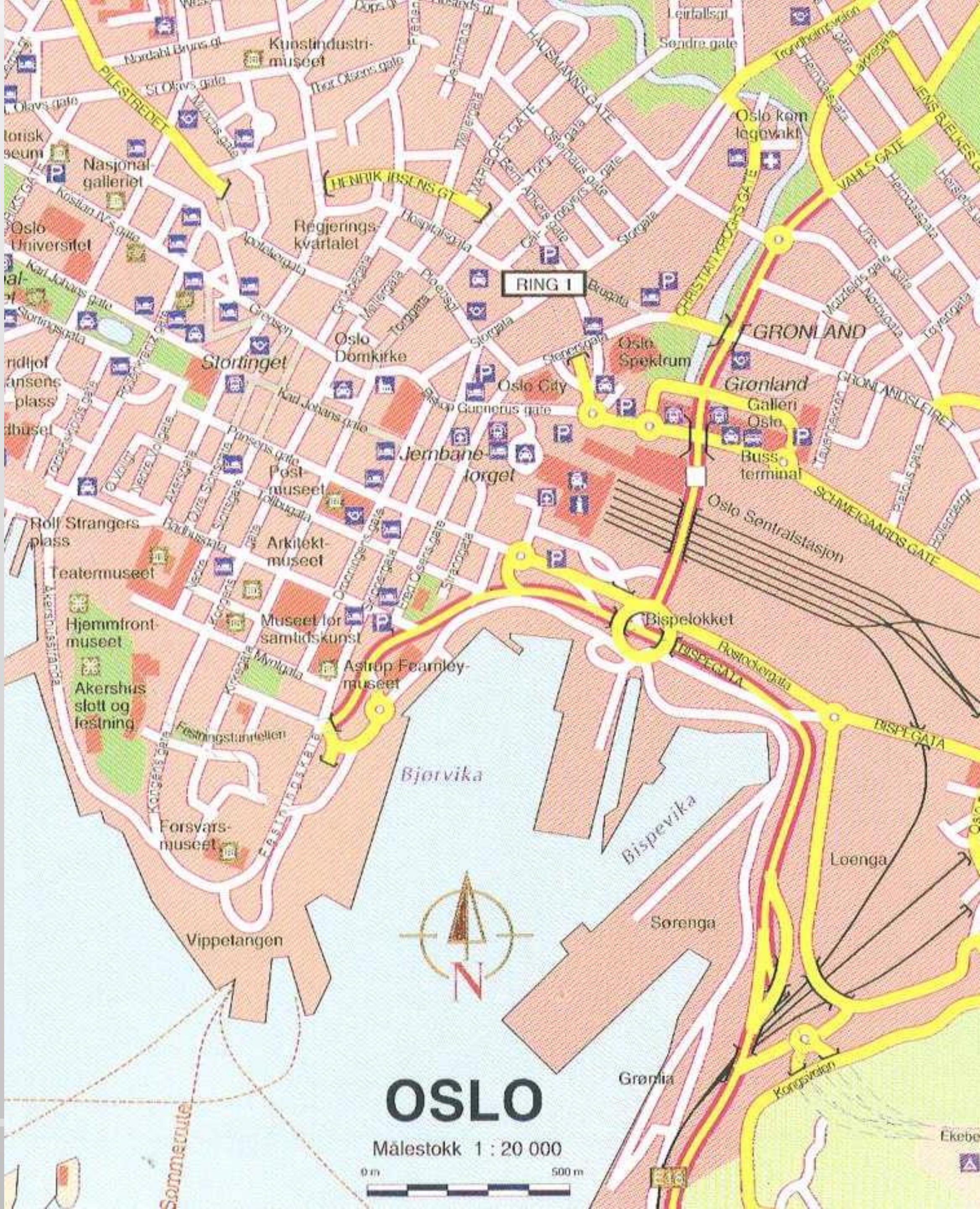
- Spatial location
- Continuity (continuous vs discrete)
- Attributes (level of measurement)

Basic characteristics of geographic data

- Spatial location
- Continuity (continuous vs discrete)
- Attributes (level of measurement)

Geographical location

- ☐ **Absolute reference system**
  - – Georeferencing systems
    - Ellipsoids
    - Geoids
    - Datum
    - Projections
    - Coordinate systems
- ☐ **Relative reference system**
  - – Local coordinate systems
- ☐ **Geographical information**
  - – E.g. postcodes, administrative units



Basic characteristics of geographic data

- Spatial location
- Continuity (continuous vs discrete)
- Attributes (level of measurement)

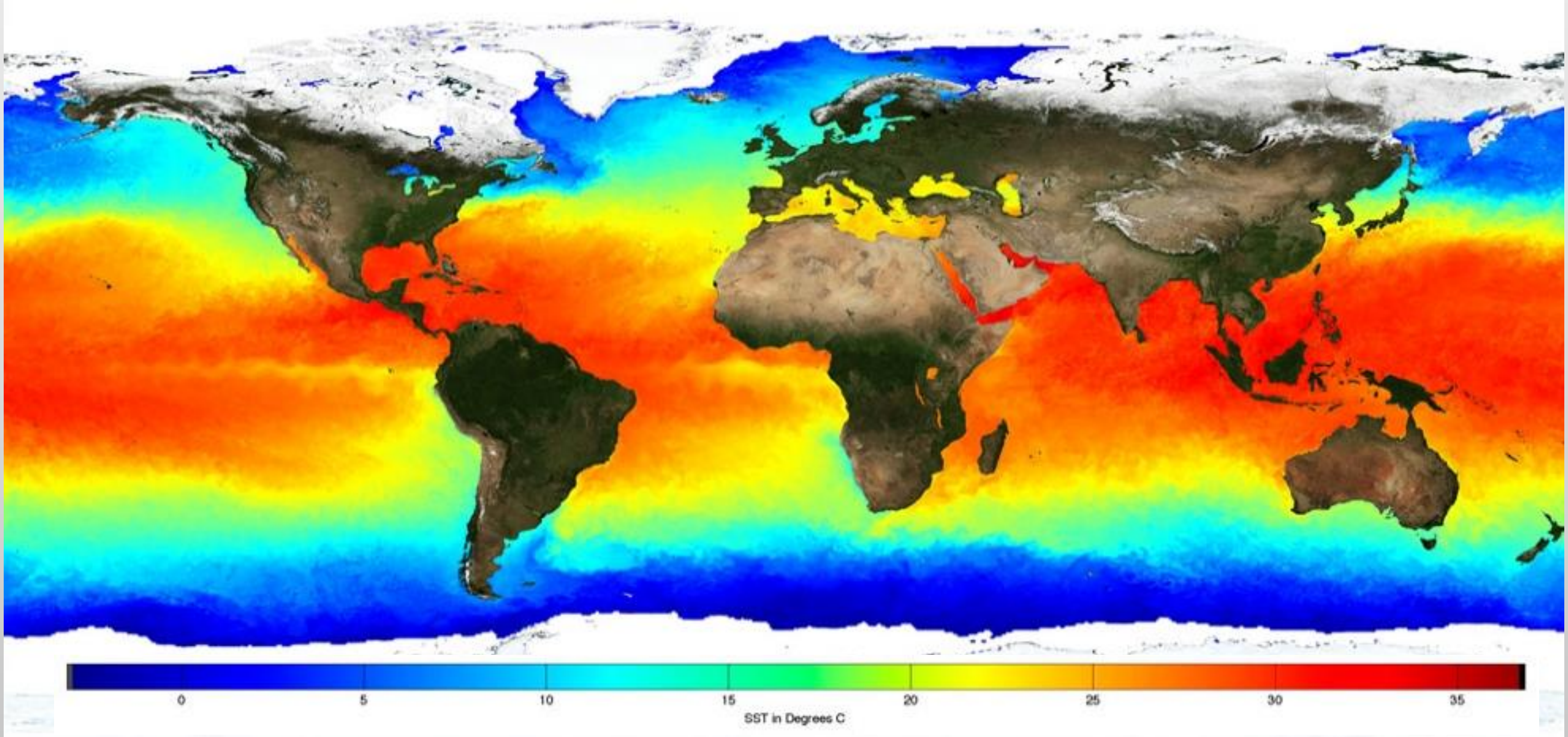
Continuous vs. discrete data

- **Discrete object**
  - The space is empty except where there are objects
    The object has well-defined delimitation.
    - Objects can be counted. Examples:
      - Points: Town, power pole
      - Line: river stream, highway
      - Polygon (area): national park, countries
- **Continuous object**
  - Data values distributed across a surface without interruption
  - Each variable is defined at every possible position
    - Example: Temperature, elevation

Continuous vs. discrete data

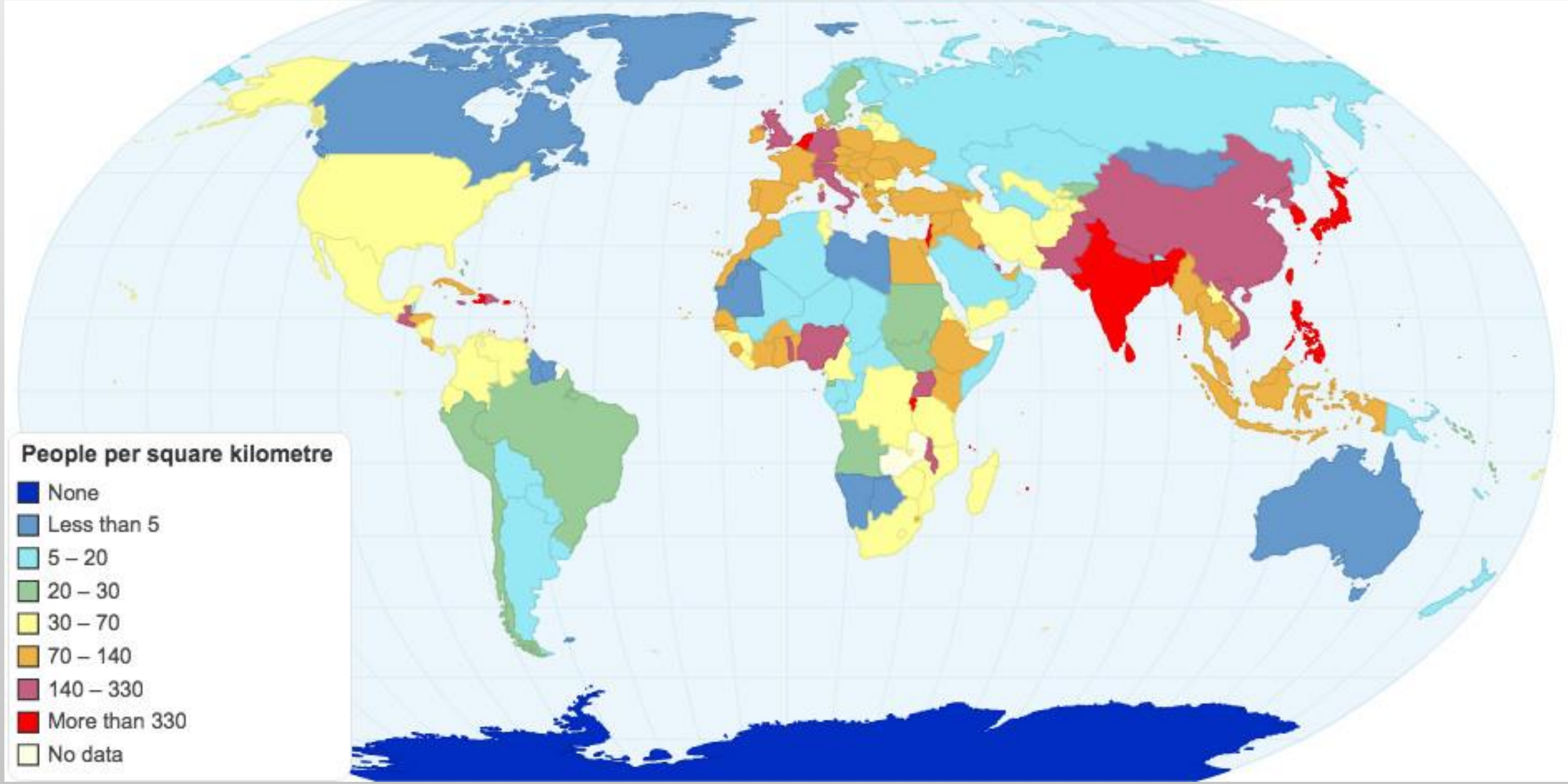
- **Some data types may be presented as either discrete or continuous**
  - Example:
    - Population at a point (discrete)
    - Population density surface for an area (continuous)

Continuous vs. discrete data?



Sea Surface Temperature (SST) (°C)

Continuous vs. discrete data?



Digital representation of spatial data

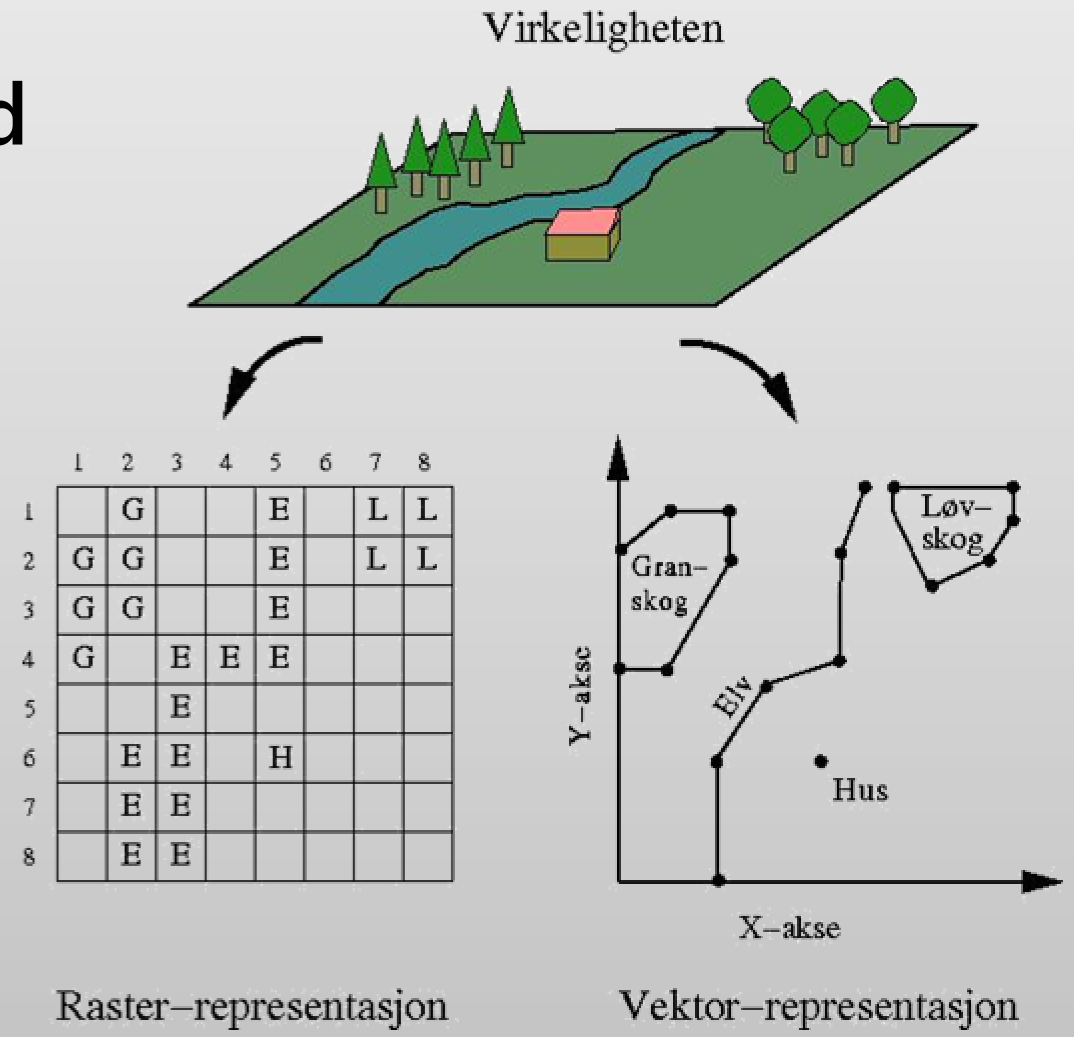
How would you represent geographical object?

Digital representation of spatial data

- Raster and vector data are used for representing continuous and discrete objects

Raster representations: divide the world into <u>arrays of cells</u> and assign attributes to the cells.

Vector data uses <u>sets of coordinates</u> and associated attribute data to define <u>discrete objects</u>. There are three basic types of vector objects: <u>points, lines, and polygons</u>.

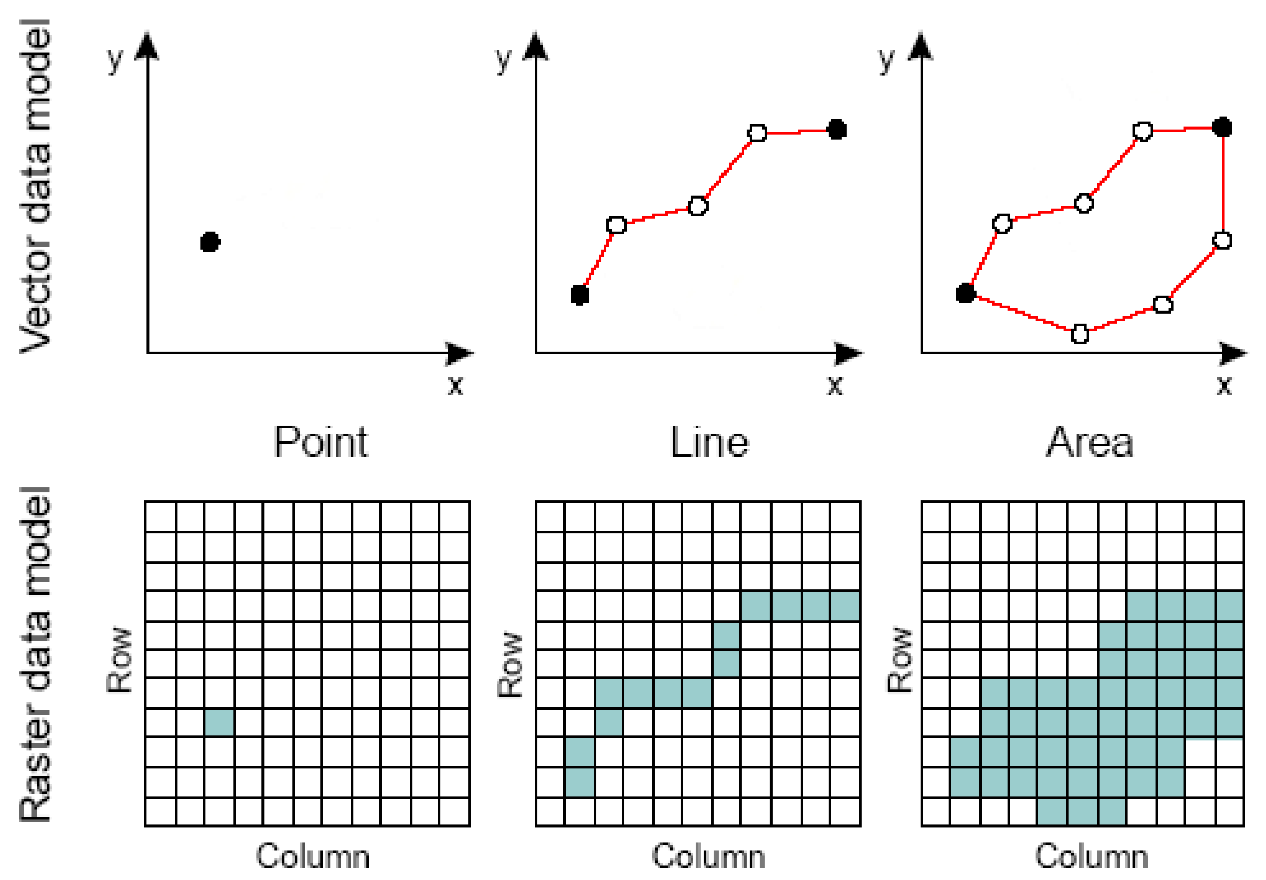


Digital representation of spatial data

- Raster and vector data are used for representing continuous and discrete objects

Raster representations: divide the world into arrays of cells and assign attributes to the cells.

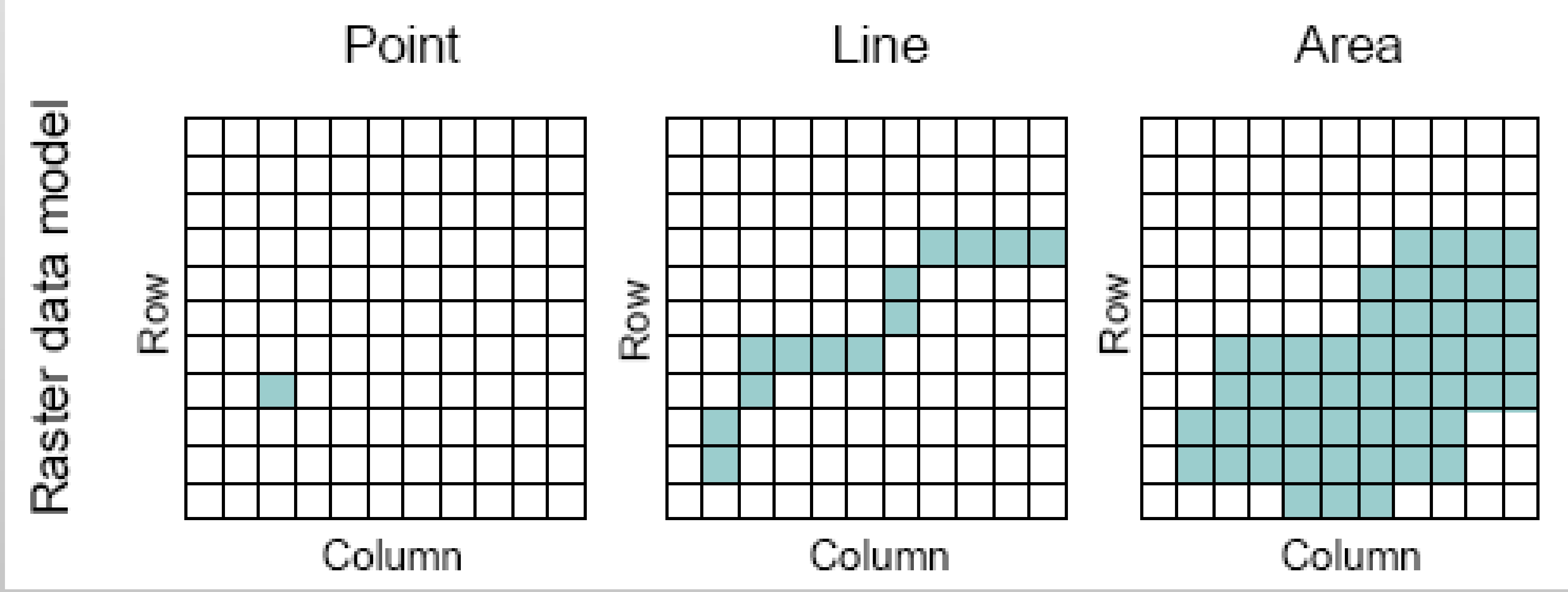
Vector data uses sets of coordinates and associated attribute data to define discrete objects. There are three basic types of vector objects: points, lines, and polygons.



Digital representation of spatial data

- **Raster data**

  Define space as an array of equally sized cells arranged in rows and columns. Each cell contains an attribute value and location coordinates.

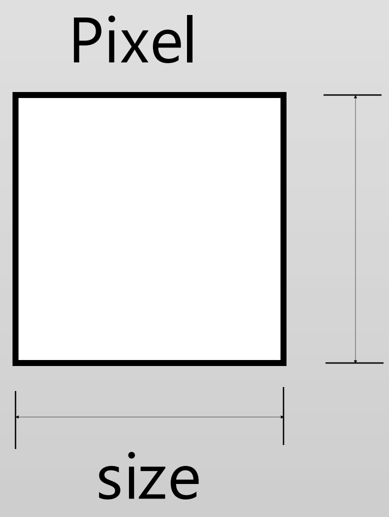


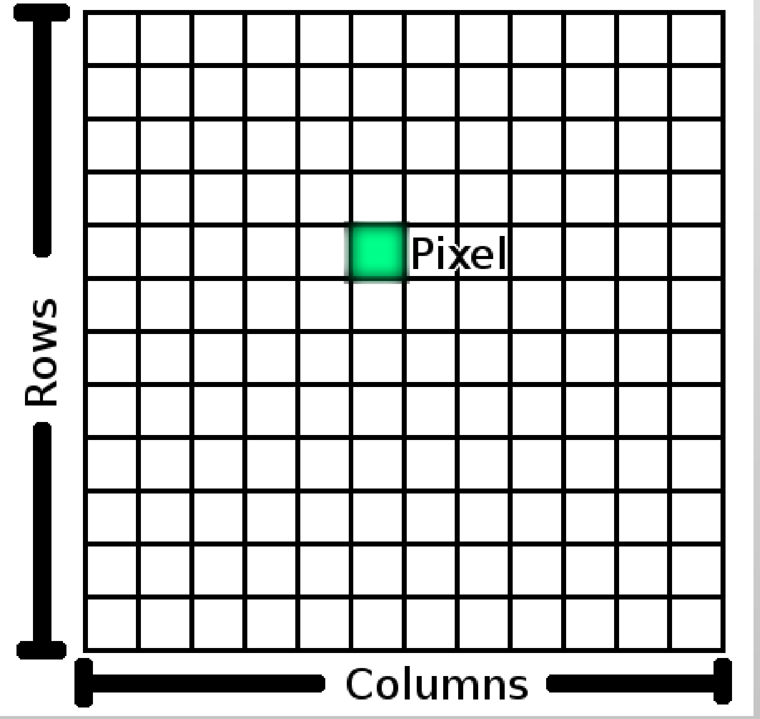
Digital representation of spatial data

☐ **Raster data**

Define space as an array of equally sized cells arranged in rows and columns. Each cell contains an attribute value and location coordinates

- Matrix of pixels or cells
  - Rows
  - Columns



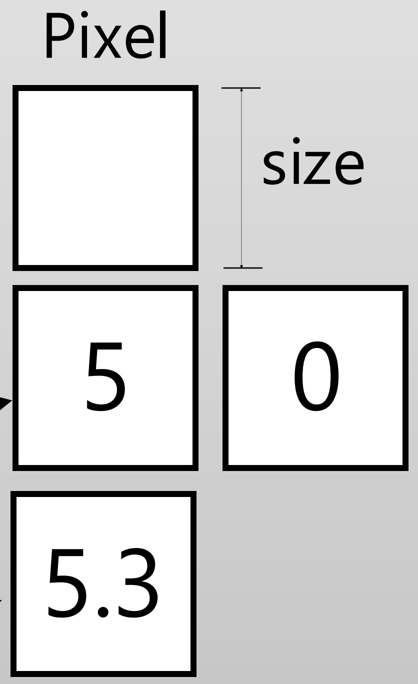


Digital representation of spatial data

Raster data

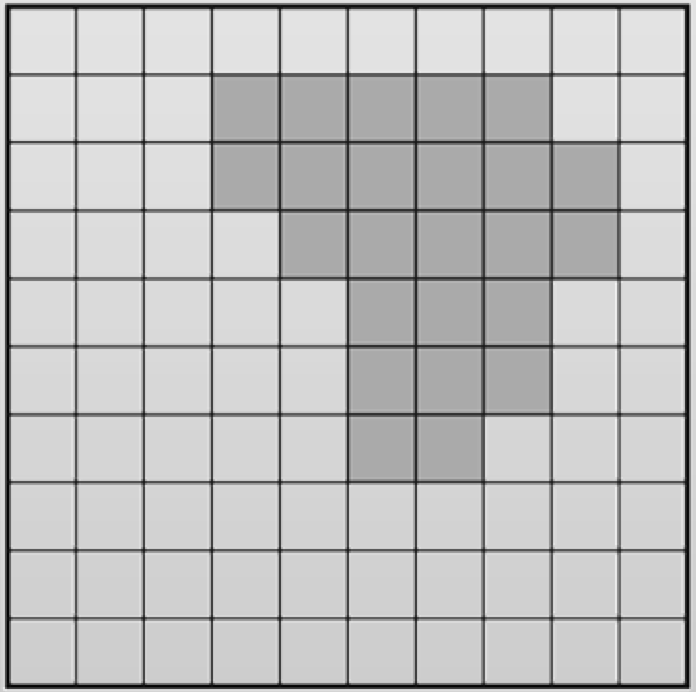
Define space as an array of equally sized cells arranged in rows and columns. Each cell contains an attribute value and location coordinates

- **Data type**
  - Integer for discrete or bolean (0,1 data)
  - Float for continuous data
  - Nodata (nan, NoData, -9999, 99999)
  - Discrete or continuous



Digital representation of spatial data

- Raster data: Data type



entity model

0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	1	1	0	0
0	0	0	1	1	1	1	1	1	0
0	0	0	0	1	1	1	1	1	0
0	0	0	0	0	1	1	1	0	0
0	0	0	0	0	1	1	1	0	0
0	0	0	0	0	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

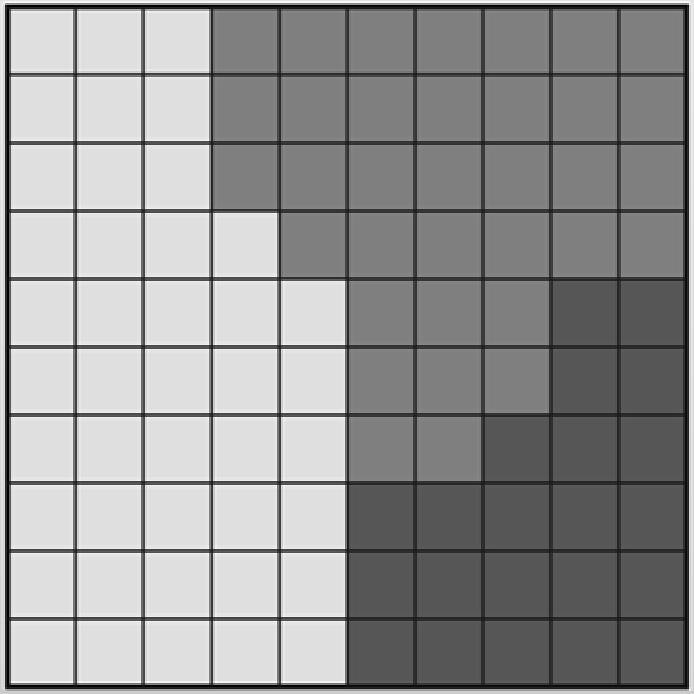
cell values (bolean)

10,10,1
0,0,0,0,0,0,0,0,0,0
0,0,0,1,1,1,1,1,0,0
0,0,0,1,1,1,1,1,1,0
0,0,0,0,1,1,1,1,1,0
0,0,0,0,0,1,1,1,0,0
0,0,0,0,0,1,1,1,0,0
0,0,0,0,0,1,1,0,0,0
0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0

File structure

Digital representation of spatial data

- Raster data: Data type



entity model

1	1	1	2	2	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2
1	1	1	1	2	2	2	2	2	2
1	1	1	1	1	2	2	2	3	3
1	1	1	1	1	2	2	2	3	3
1	1	1	1	1	2	2	3	3	3
1	1	1	1	1	3	3	3	3	3
1	1	1	1	1	3	3	3	3	3
1	1	1	1	1	3	3	3	3	3

cell values (integer)

10,10,1
1,1,1,2,2,2,2,2,2,2
1,1,1,2,2,2,2,2,2,2
1,1,1,2,2,2,2,2,2,2
1,1,1,1,2,2,2,2,2,2
1,1,1,1,1,2,2,2,3,3
1,1,1,1,1,2,2,2,3,3
1,1,1,1,1,2,2,3,3,3
1,1,1,1,1,3,3,3,3,3
1,1,1,1,1,3,3,3,3,3
1,1,1,1,1,3,3,3,3,3

File structure

Digital representation of spatial data

- Raster data: Data type

.txt file of a Digital Elevation Model (DEM)

cell values float

Header

Grid with elevation (m)

```
ncols         1350
nrows         1300
xllcorner     609950
yllcorner     8799950
cellsize      100
NODATA_value  -9999
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.04
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.9700
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3.177062
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.3820248 1.71
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.04774284 2.621858
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.5602636 3.134378 5.708
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.072784 3.646899 6.221014 8
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.73428 4.15942 6.733535 9.30765
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.8791168 5.108694 9.338271 10.6236
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3.043989 5.576283 4.875862 4.491128 8.483109
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3.495815 7.044259 6.48873 5.852359 9.10017 13.00581 15
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 4.270318 9.768658 11.9959 16.54611 19.88974 19.33421 18.5831
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3.777967 14.15788 17.71671 25.71457 27.63547 27.12899 29.5964 32.7
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3.483173 12.59066 22.20958 31.78131 43.13511 43.34449 45.89667 45.50227
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.944654 9.266139 19.51434 28.50399 38.42742 49.26828 64.05908 72.11234 68.722
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 5.49435 5.346334 11.46607 23.03997 31.36146 41.41116 51.69438 71.24174 85.87129 97.5
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 6.313496 13.63164 16.97084 20.03208 33.50088 46.96968 58.98118 66.47193 93.57722 100 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.333815 5.638436 12.87361 20.19175 27.50988 28.59808 42.06689 56.38108 71.9068 81.24949 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2.268702 6.39661 10.52452 12.53972 19.69508 26.85044 34.07 40.07184 50.72955 66.25526 81.78098
0 0 0 0 0 0 0 0 0 0 3.895014 6.644197 6.868384 7.072284 7.276185 7.331496 11.4594 15.58731 19.49526 20.73676 26.59637 33.75173
0 0 0 0 0 0 0 3.870367 8.283019 12.6985 17.18199 18.49135 18.95608 19.15998 19.2402 17.90228 20.65011 24.77801 27.8558 29.0973
0 0 0 0 0.3992912 8.576537 13.06844 17.85626 22.02965 25.98548 29.68417 30.3385 30.99282 31.24768 30.35506 28.83609 29.84081 3
0 0 0 6.337076 14.51432 22.69157 27.22748 31.84216 35.77629 39.27246 41.53132 42.18565 42.83997 43.33537 41.46992 39.76989 40.
0 0.4963841 15.02564 23.9315 29.21978 36.8066 41.84476 45.86507 49.52292 52.28429 51.34307 51.60319 51.86331 52.08657 51.11317
0 19.62345 33.74644 41.58176 47.7405 53.0591 58.23998 63.03893 64.06525 59.20271 56.05276 56.31288 56.57299 56.83311 55.89993
0 39.488 51.23852 59.16027 66.42404 71.75256 76.70193 75.84619 70.98365 66.12112 61.25858 61.02257 61.28268 61.5428 60.68669 6
0 73.88699 78.20627 78.74113 84.75336 90.36492 87.62714 82.7646 77.90206 73.03953 68.17699 65.73225 65.99237 66.25249 65.47346
0 100 100 100 100 99.40809 94.54555 89.68301 84.82047 79.95793 75.0954 70.44194 70.70206 70.96218 70.26022 71.478 73.82849 74.
0 100 100 100 100 100 100 96.60142 91.73888 86.87634 82.01381 77.15127 75.41175 75.67187 75.04697 76.69225 79.16434 79.95953 8
0 100 100 100 100 100 100 100 98.6573 93.79475 88.93221 84.06968 80.12144 80.38155 79.83374 81.9065 84.37859 85.73479 86.18106
0 100 100 97.93361 100 100 100 100 100 100 95.85062 90.98809 86.12555 85.09125 84.64867 87.12075 89.59284 91.51003 91.95303 92
```

Digital representation of spatial data

☐ Raster data: Data type

Discrete raster:

- Relative few possible values with boundaries.

Example: Land use and land cover map (raster), soil types, etc.

Continuous raster:

- Numeric values range smoothly from one location to another.

Example: Digital elevation model (DEM), temperature, remote sensing images

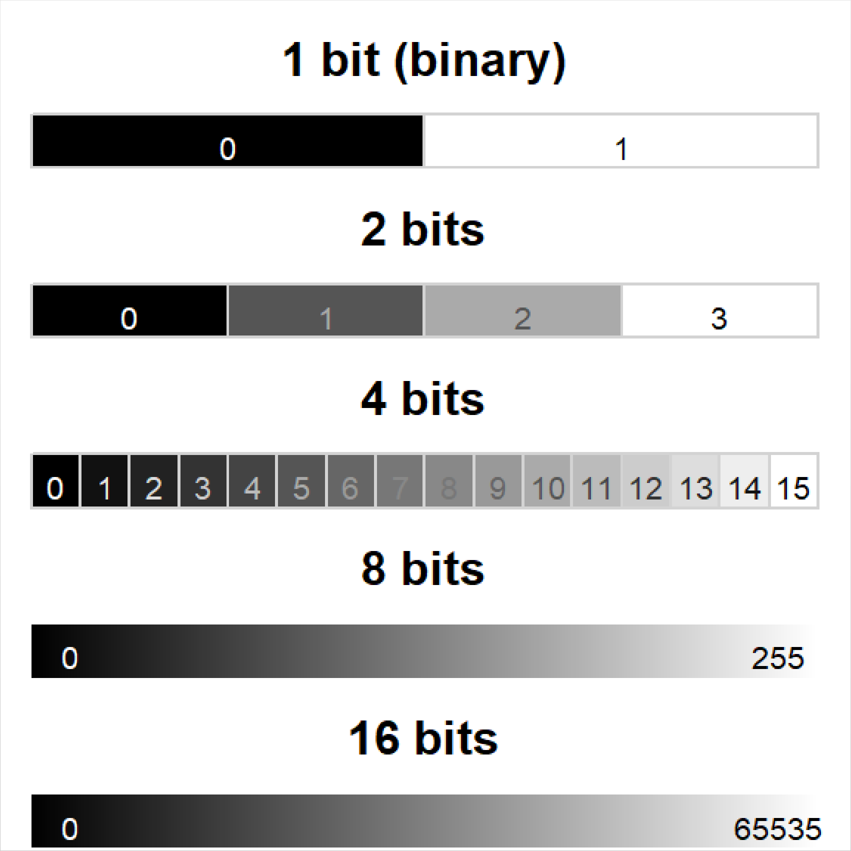
Digital representation of spatial data

Raster data: Data type

Bit depth: number of possible values / colours for each cell

- 1 byte / cell = 8 bits / cell = 2 ^ 8 = 256 shades of gray (from 0 to 255)
  Ex. RGB image, elevation raster

- 2 bytes / cell = 16 bits / cell = 2 ^ 16 = 65536 grayscale (from 0 to 65535).
  Ex. satellite image (multispectral)

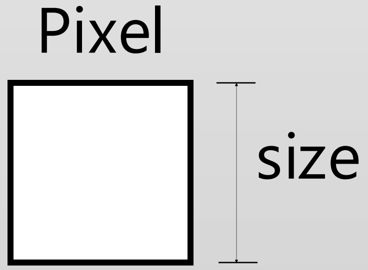


Digital representation of spatial data

☐ **Raster data:**

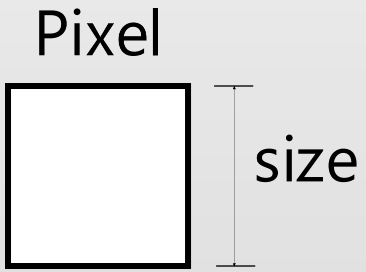
Define space as an array of equally sized cells arranged in rows and columns. Each cell contains an attribute value and location coordinates

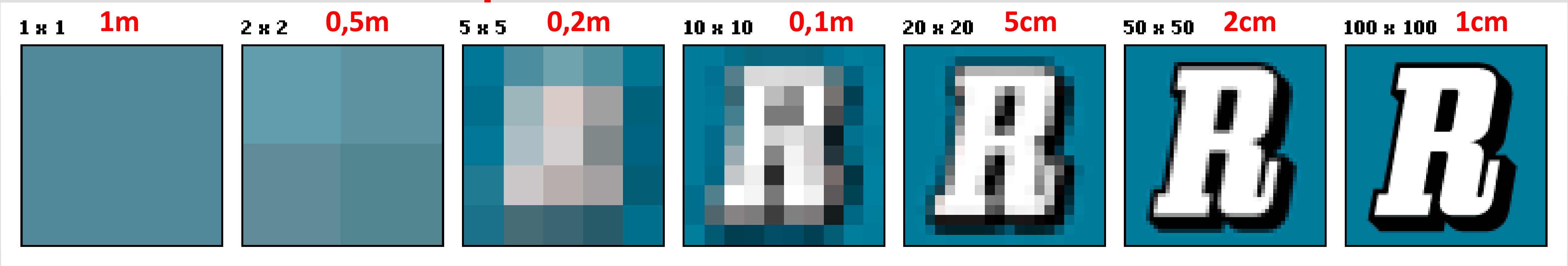
- **Spatial resolution (pixel/cell size)**
  - The level of detail in a raster or image
  - The pixel size on the ground (not the pixel on your screen)
  - Number of pixels in a given frame



Digital representation of spatial data

- Raster data: Spatial resolution,
  Ground Sampling Distance (GSD)
  → (pixel/cell size)





Grov oppløsning

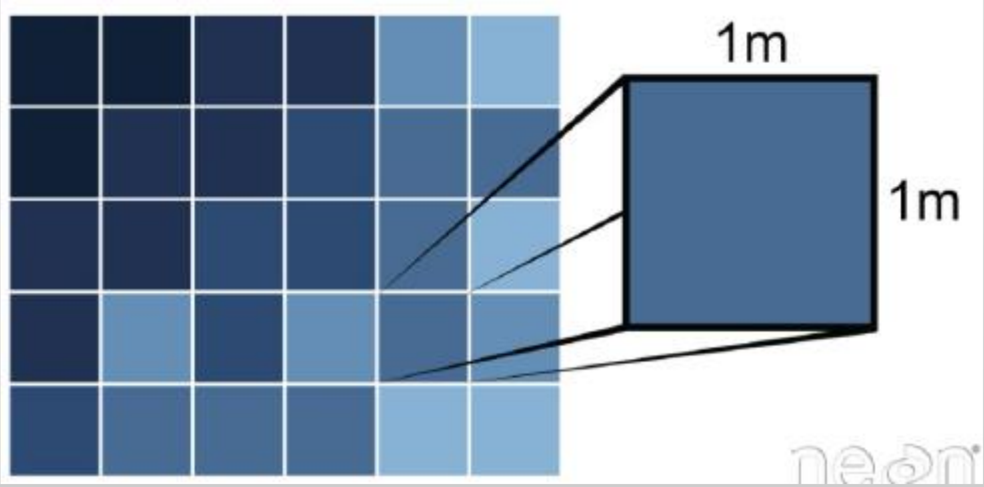
Fin oppløsning

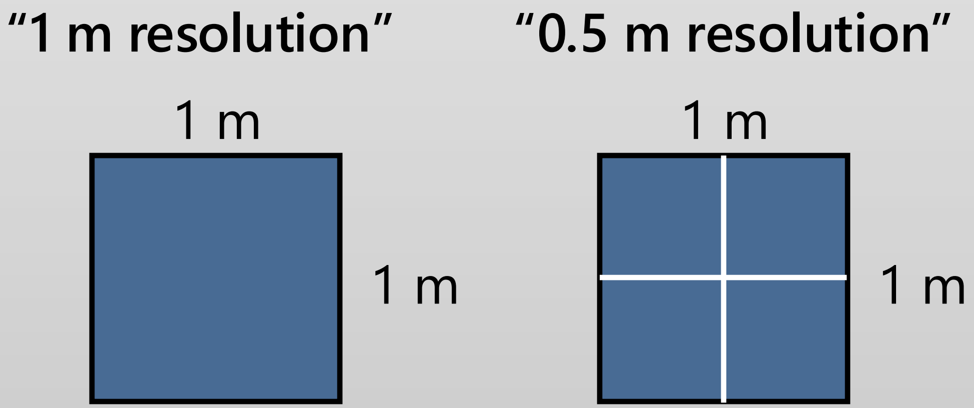
Spatial resolution should be less than half the size of the smallest object you want to map

Digital representation of spatial data

- Raster data: Spatial resolution (pixel/cell size)

Remember: raster resolution is a tradeoff between the **information you want to convey** and the **size (and processing time)** of the file you are dealing with.

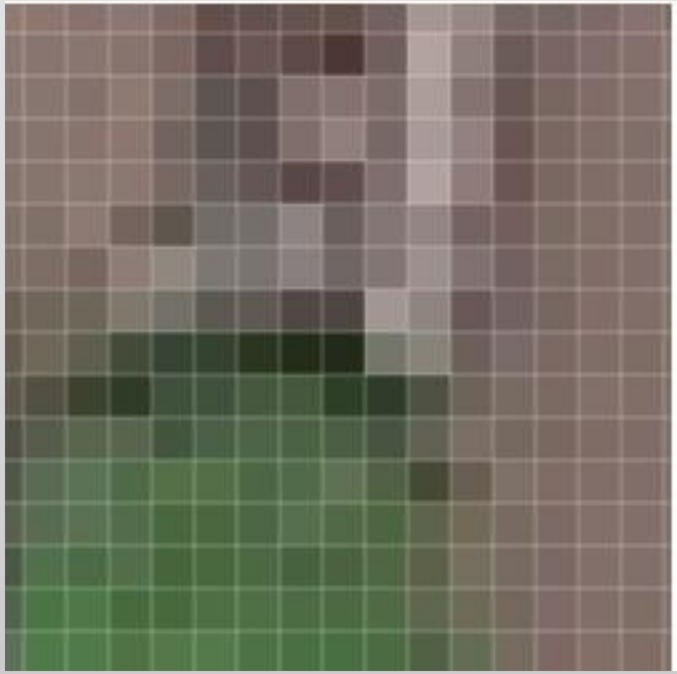




Note: To **DOUBLE** the resolution, we have to **QUADRUPLE** the cells. *So be very careful, because files can get really big.*

Digital representation of spatial data

☐ Raster data: Spatial resolution (pixel/cell size)



Example: Orthophoto

- Spatial resolution should be less than half the size of the smallest object you want to map
- Cell resolution is given as its size on the ground.

Digital representation of spatial data

- **Rasters as basemaps**
  - Orthophotos displayed underneath other layers
    - Example: orthophotos from aerial photography, satellite imagery, and scanned maps

Important: map layers must be spatially aligned

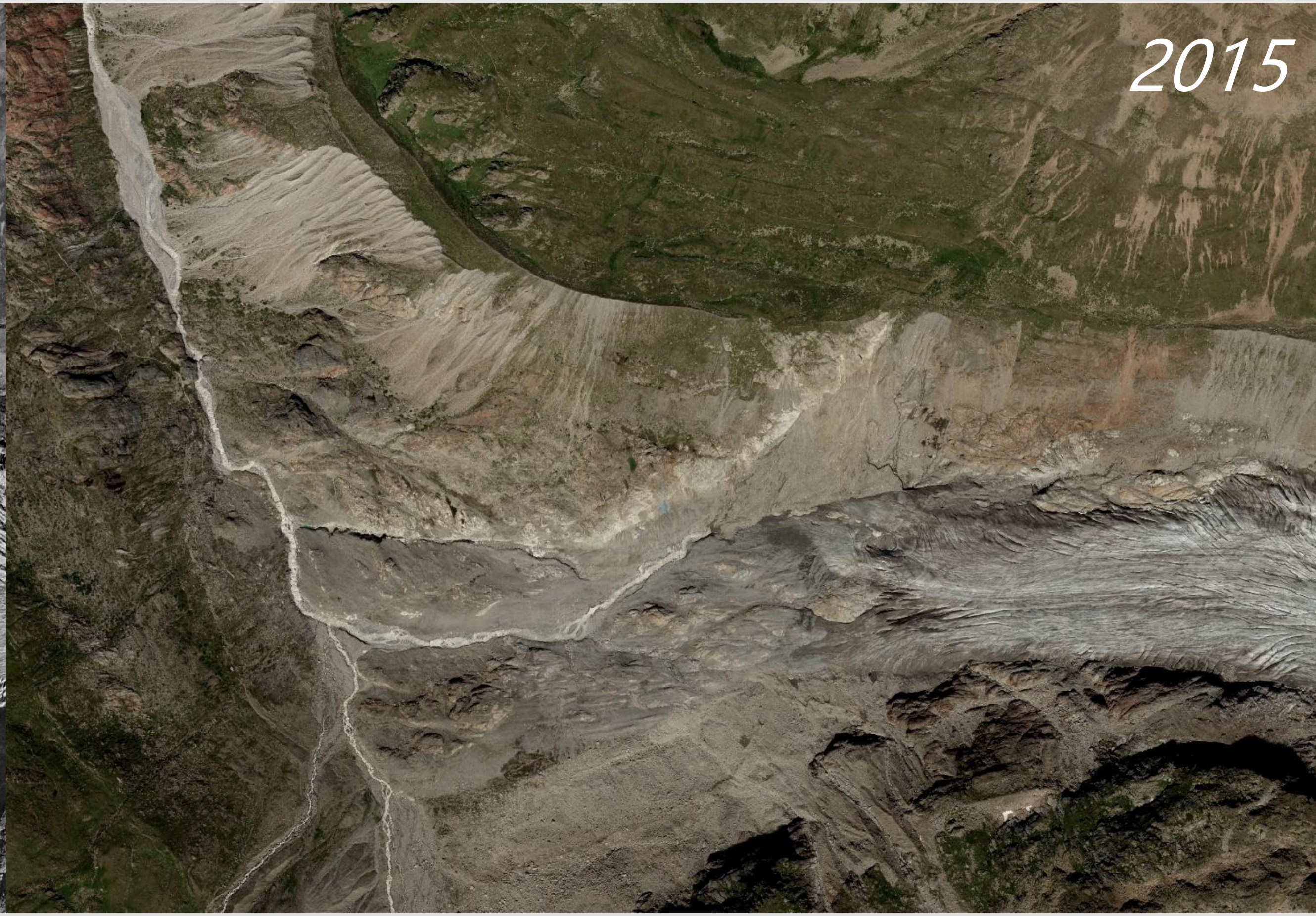
- **Rasters as surface maps:**
  - Representing data that changes continuously across a landscape (surface)
  - Method of storing the continuity as a surface
    - Example: Elevation values, temperature, rainfall, concentration

- **Rasters as thematic maps:**
  - Representing thematic data
    - Example: classified satellite image by land-cover categories

Digital representation of spatial data

- Rasters as basemaps

Example: Orthophoto as a base to map features

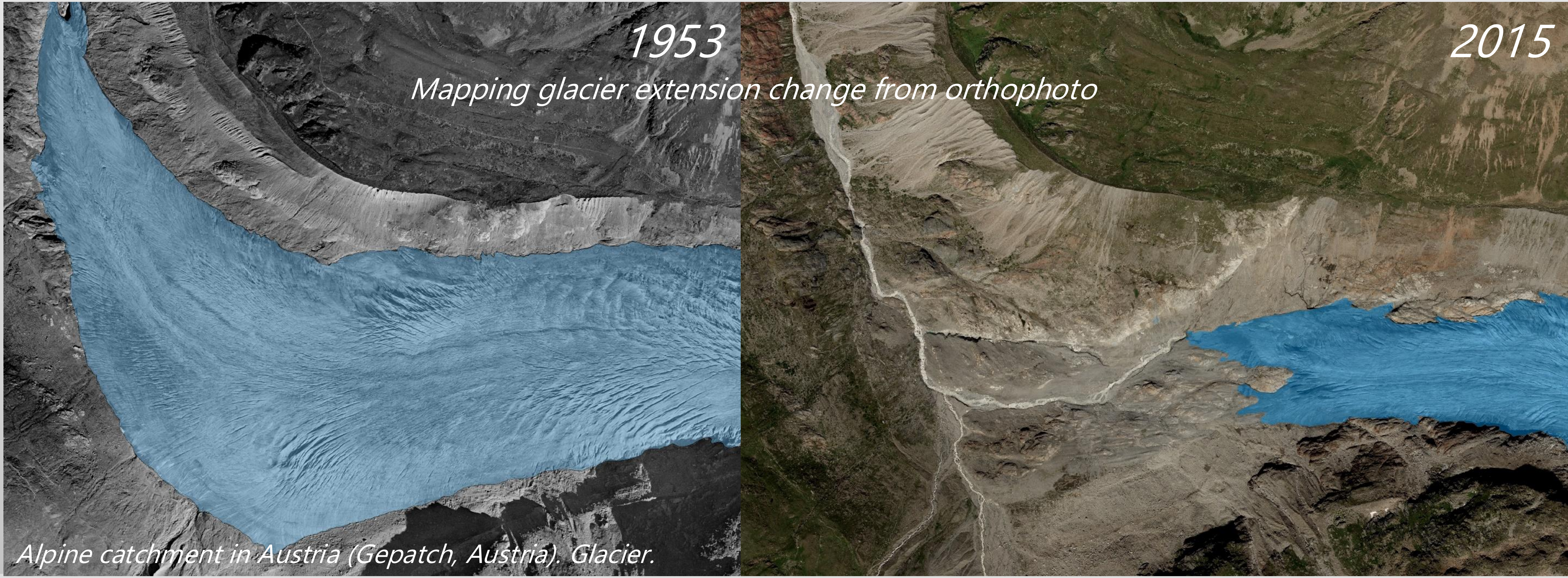


Alpine catchment in Austria (Gepatch, Austria). Glacier.

Digital representation of spatial data

- Rasters as basemaps

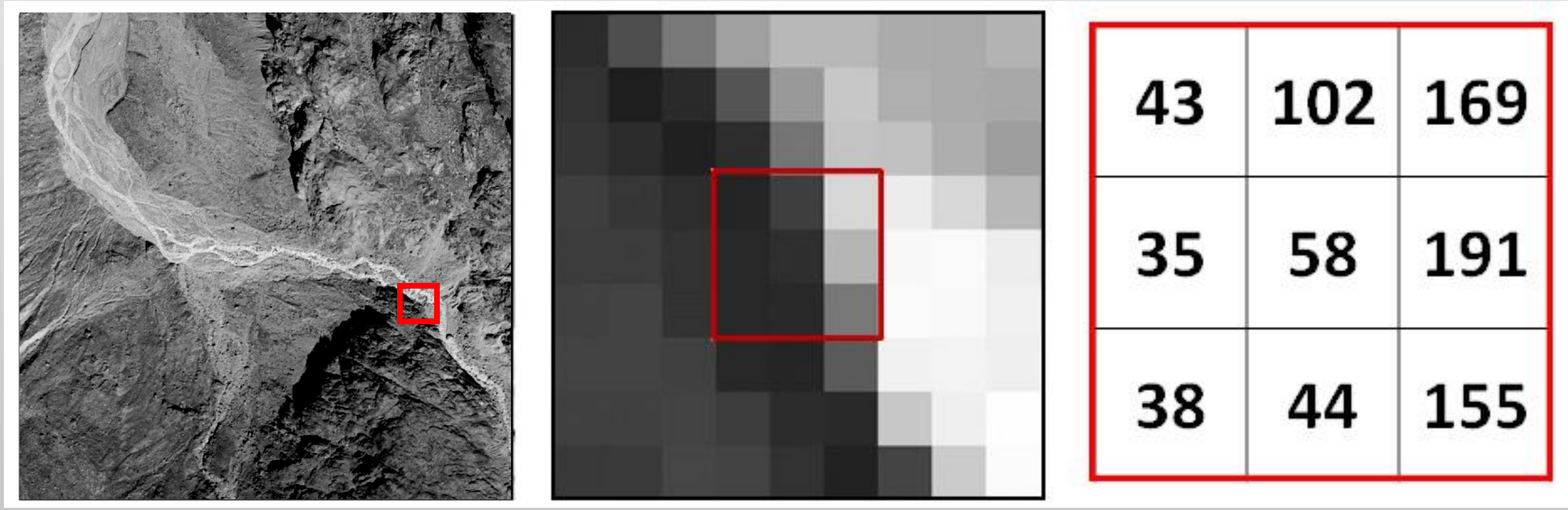
Example: Orthophoto as a base to map features



Digital representation of spatial data

- Rasters as basemaps

Remember: maps and orthophoto are numbers first, images later



Digital representation of spatial data

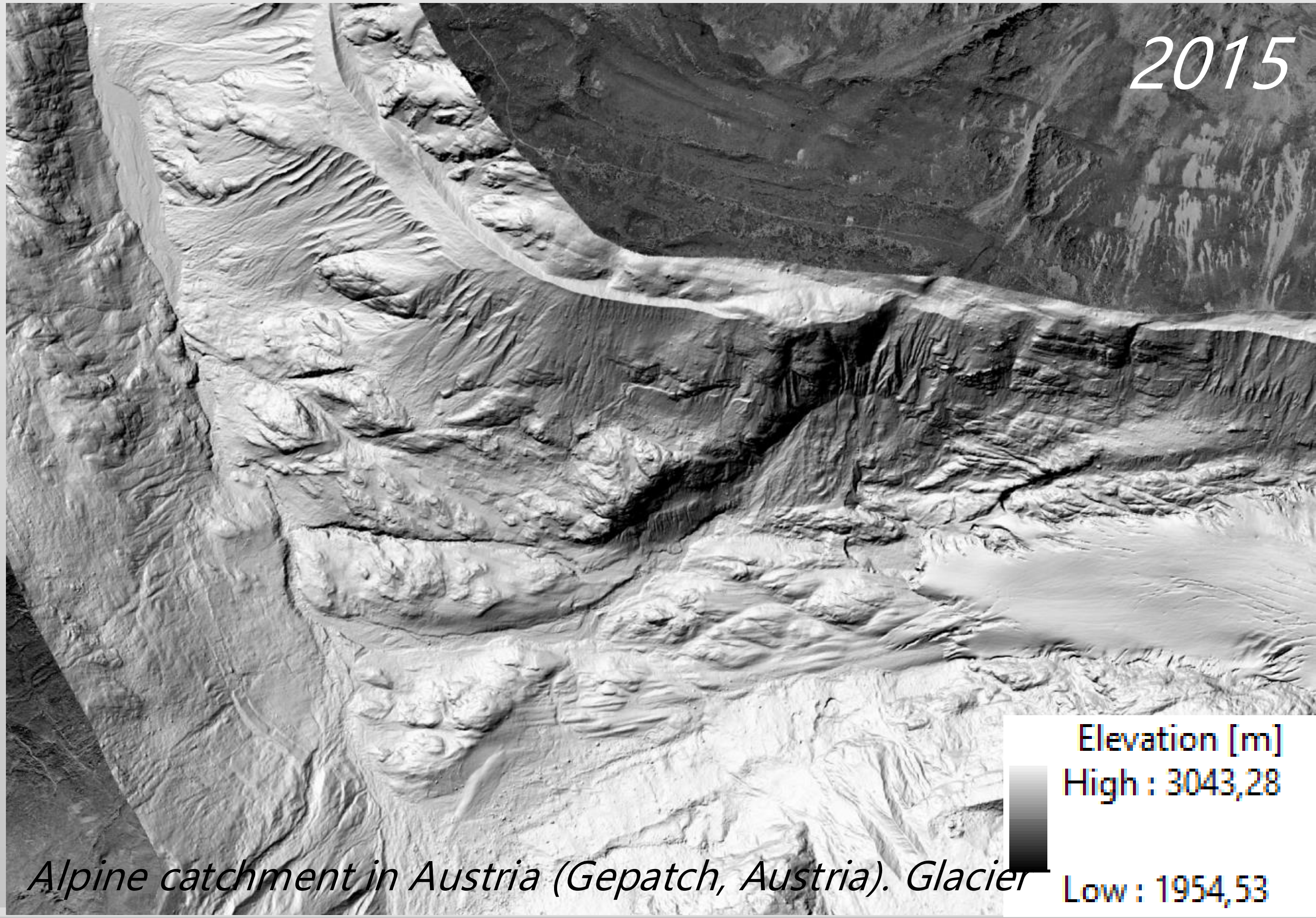
Rasters as surface maps

Elevation values measured from the Earth's surface are the most common application of surface maps

Other values can also define surfaces and be spatially analyzed

- Rainfall
- Temperature
- Concentration
- population density

Example: Digital elevation model (DEM)

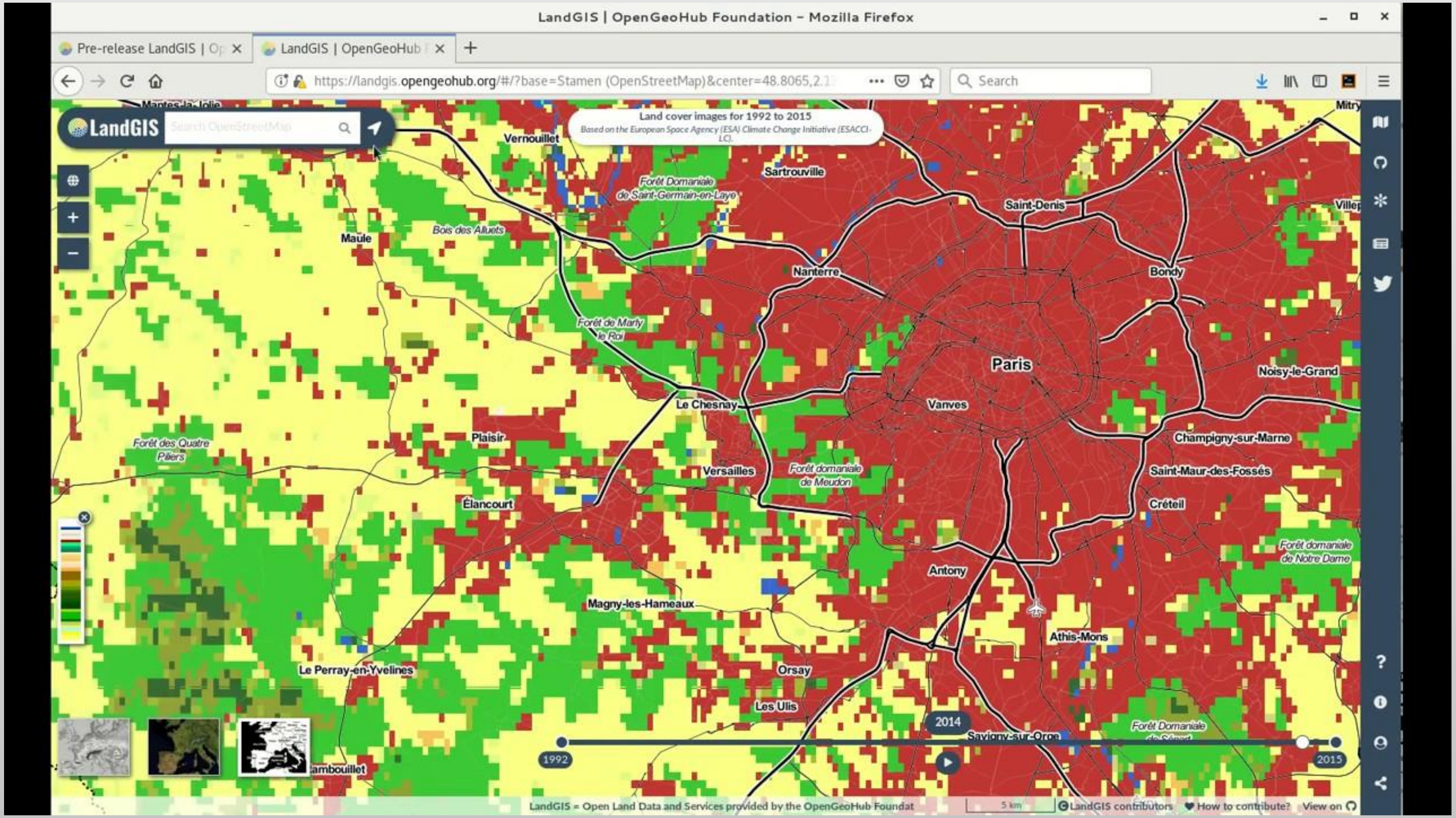


Alpine catchment in Austria (Gepatch, Austria). Glacier

Digital representation of spatial data

- Rasters as thematic maps

Example: Land use map (raster)



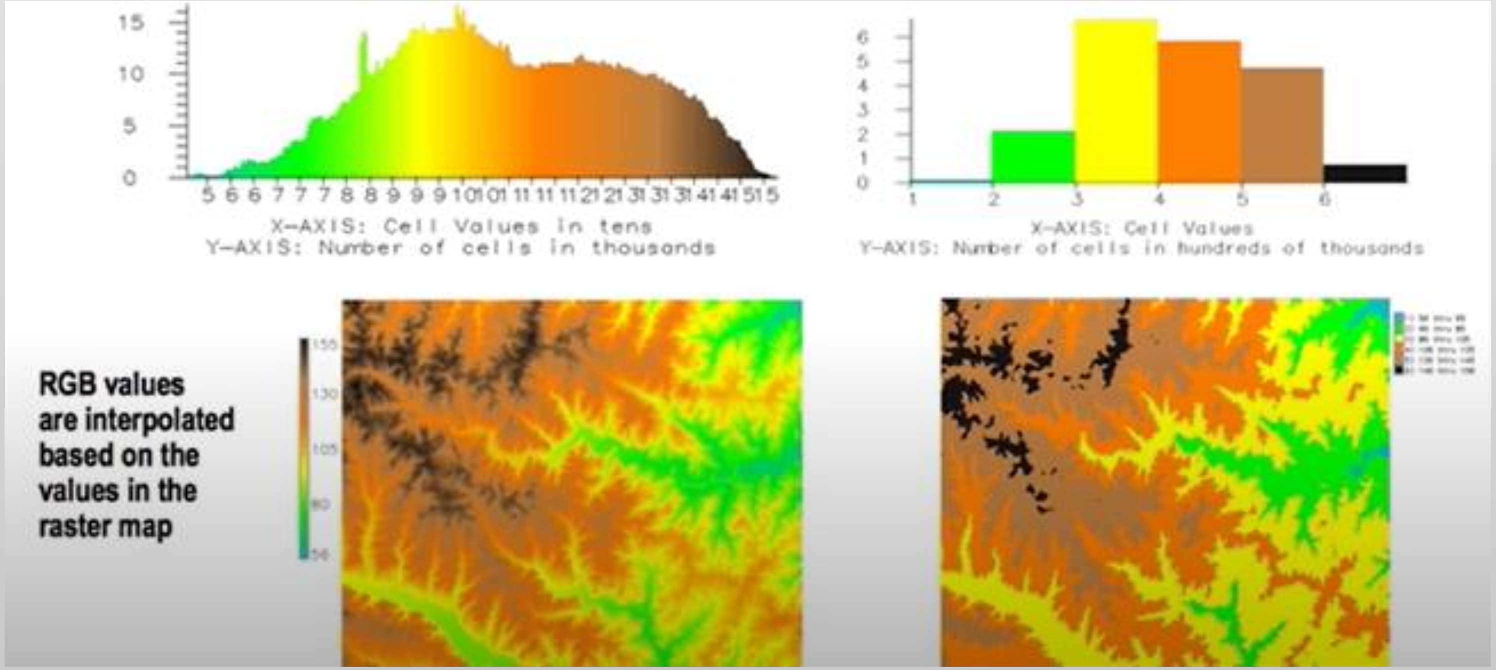
How do you see that this is a raster!?

Digital representation of spatial data

☐ Raster data: Visualization

Colour table:

continuous discrete



In a GIS, the raster values are often assigned to grayscale / colour automatically when loading (8 bit, 0-255 values)

The user can adjust the colour balance using:

- <u>stretching</u> (continuous representation)
- <u>histogram</u> adjustment
- <u>classification</u> (discrete representation)

Digital representation of spatial data

- **Raster data: Accuracy and Precision**
  - Accuracy: how close is a value to the "Truth"?
    - Coordinate accuracy - right place
    - Attribute accuracy - correct attribute value
  - Precision: how close is a value to the other values?

> *"Accuracy is telling the truth . . . Precision is telling the same story over and over again."*
> *Yiding Wang*

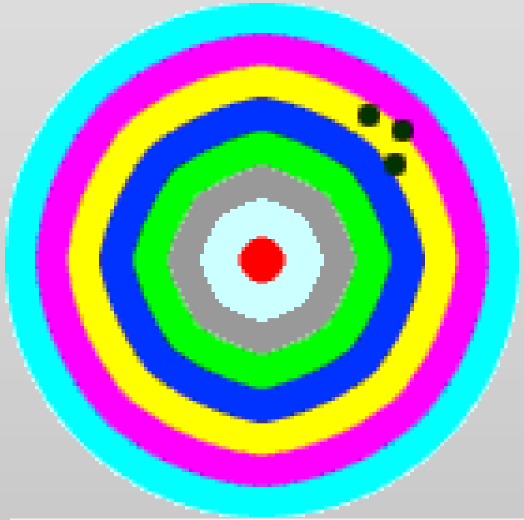


Figure 1: Precision

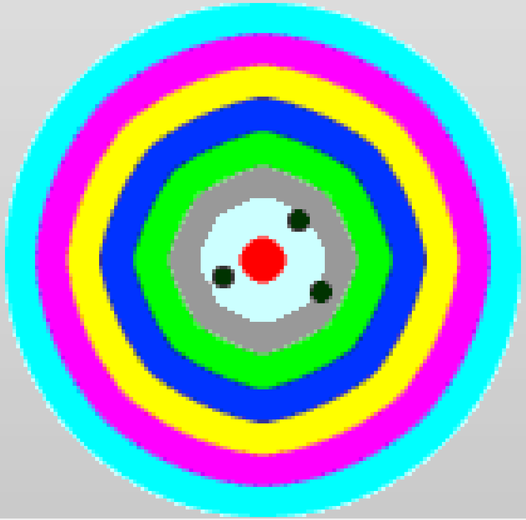


Figure 2: Accuracy

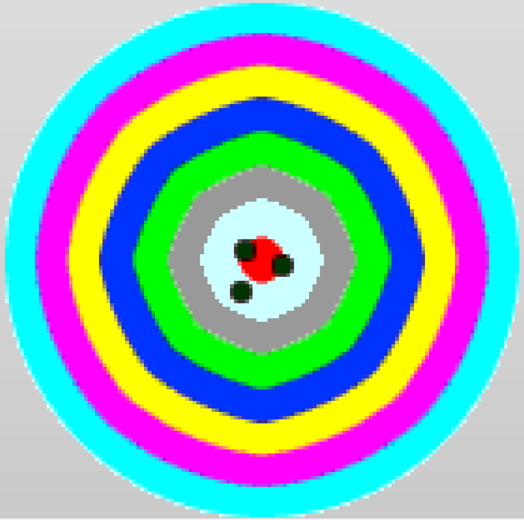


Figure 3: Accuracy with Precision

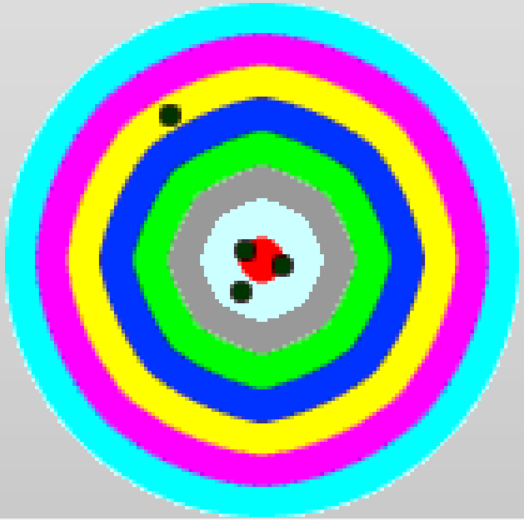


Figure 4: Accuracy with Precision and blunder

Digital representation of spatial data

☐ Raster data

Advantages

- Simple data structure
- Easy overlay
- Various kinds of spatial analysis
- Uniform size and shape

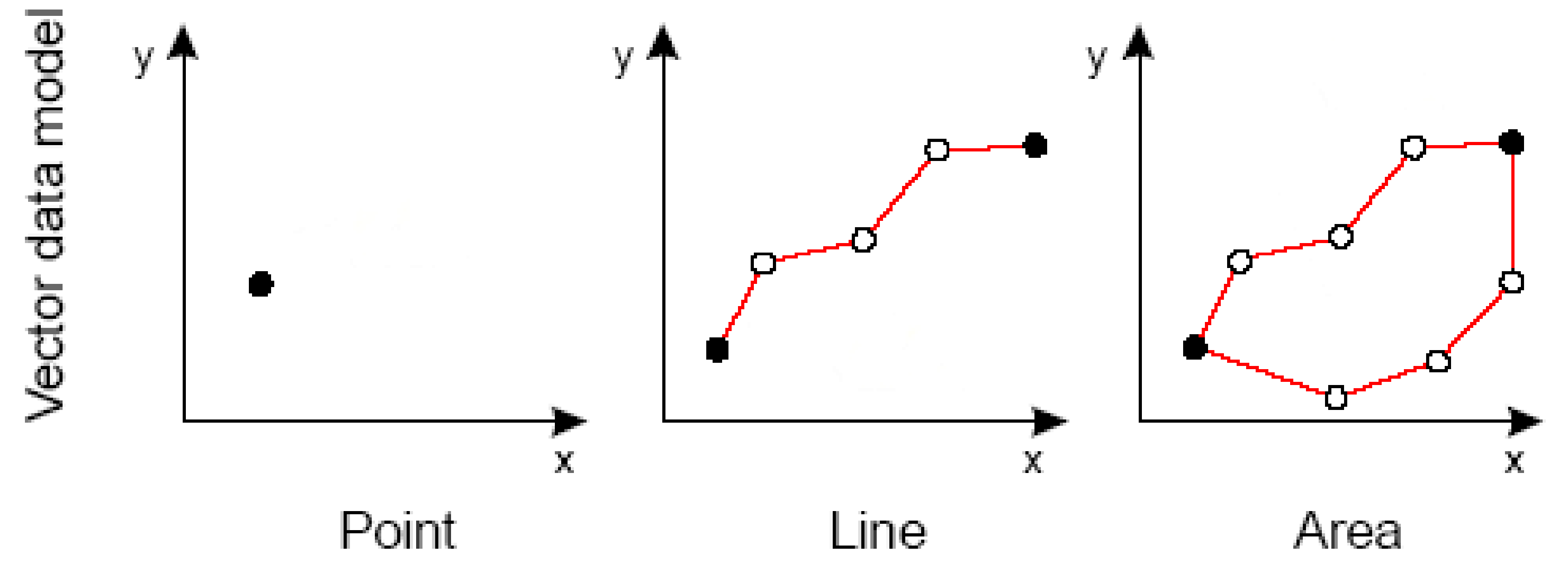
Disadvantages

- Large amount of data
- Less "pretty"
- Projection transformation is difficult
- Different scales between layers can be a nightmare
- May lose information due to generalization

Digital representation of spatial data

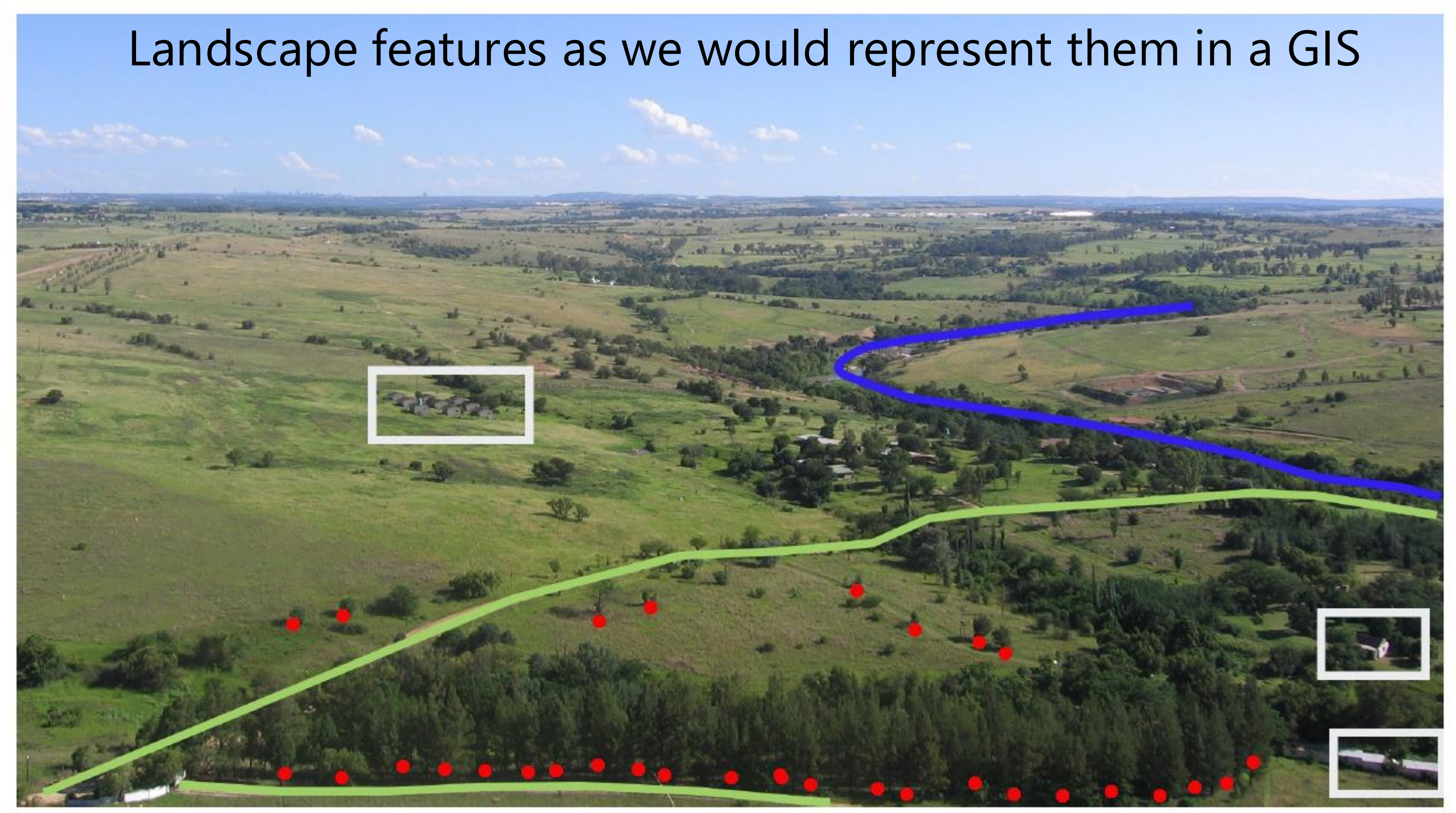
☐ Vector data model

Uses sets of coordinates and associated attribute data to define discrete objects. There are three basic types of vector objects: points, lines, and polygons.



Digital representation of spatial data

- Vector data model



Line

River

Road

Point

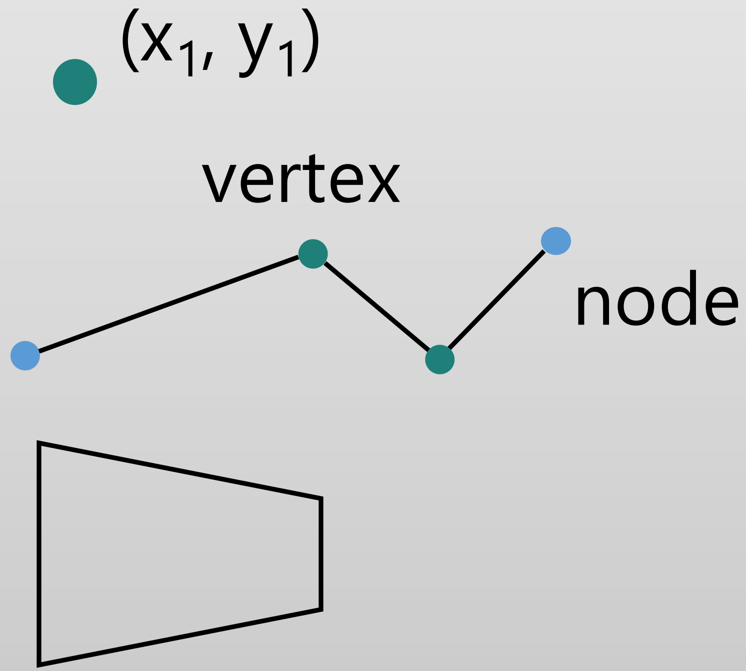
Trees

Polygons

Houses

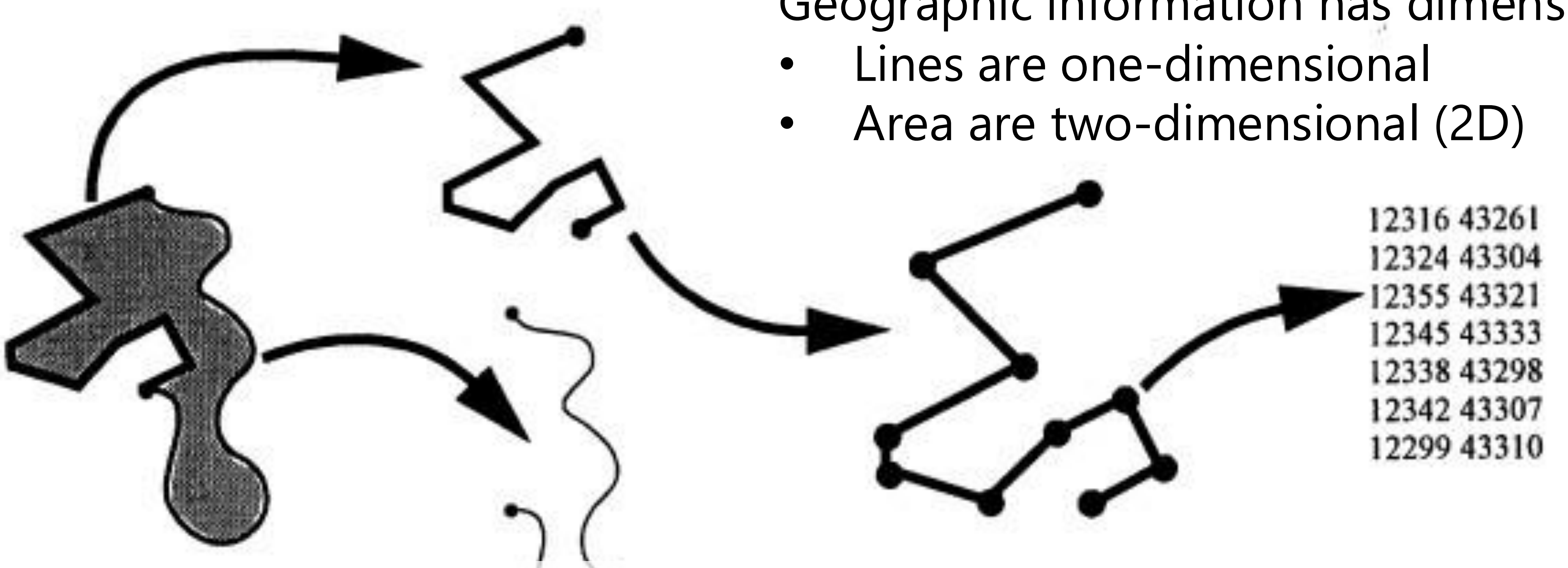
Digital representation of spatial data

- **Vector data are defined spatially as either:**
  - Point: a pair of x and y coordinates
  - Line (polyline): a sequence of points
  - Polygon: a closed set of lines



Digital representation of spatial data

- Vector data are defined spatially as either:



Geographic information has dimension:

- Lines are one-dimensional
- Area are two-dimensional (2D)

An **AREA** consist of... **LINES**, which consist of... **POINTS**, which consist of... **COORDINATES PAIR**

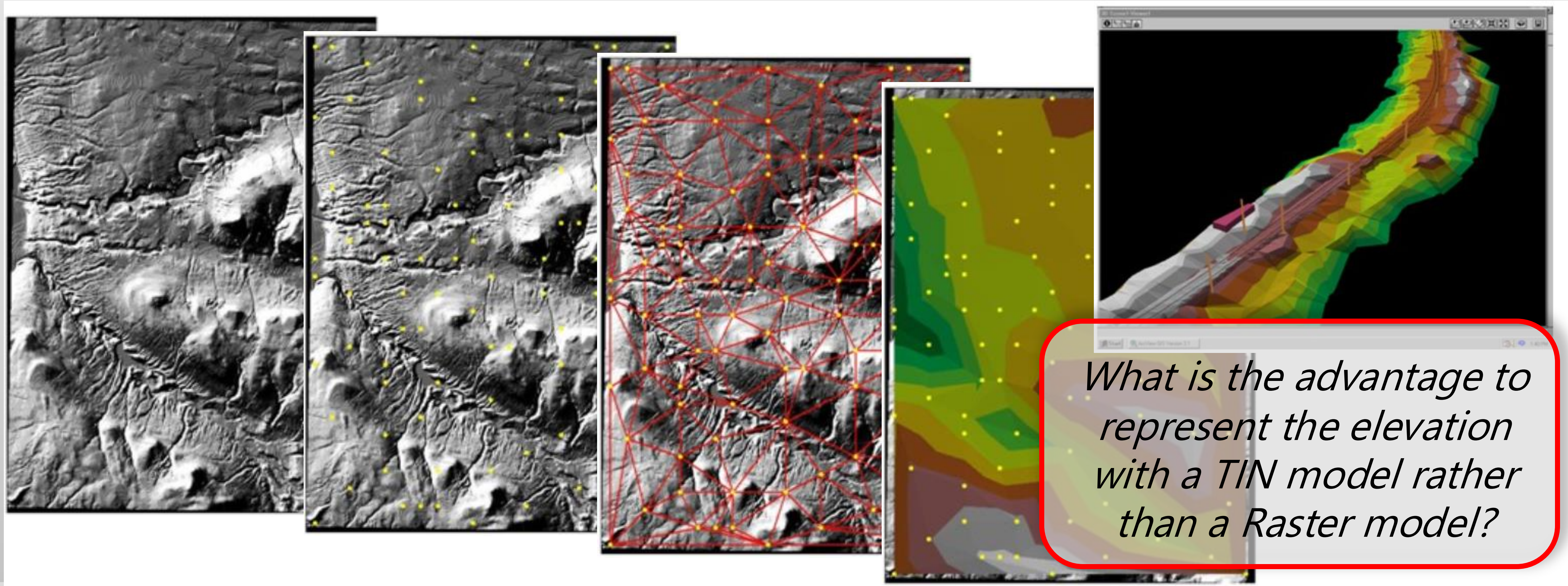
Digital representation of spatial data

TIN: Triangulated irregular network

- TIN is a form of vector-based digital geographic data
- A TIN is a data model that is used to represent three dimensional objects.
- TIN is composed of points, line and polygons and is constructed by triangulating a set of vertices (points).
- TIN is quite useful for representing elevations from points (x,y,z points where z is the elevation)
  - The points are connected into what is called a triangulation, forming a network of triangles
  - The lines of the triangles are called edges
  - The interior area is called a face

Digital representation of spatial data

- TIN: Triangulated irregular network



What is the advantage to represent the elevation with a TIN model rather than a Raster model?

Digital representation of spatial data

☐ Vector data

Advantages

- Good representation of reality
- Compact data structure
- Accurate graphics
- Small size of the files
- Topology can be described in a network

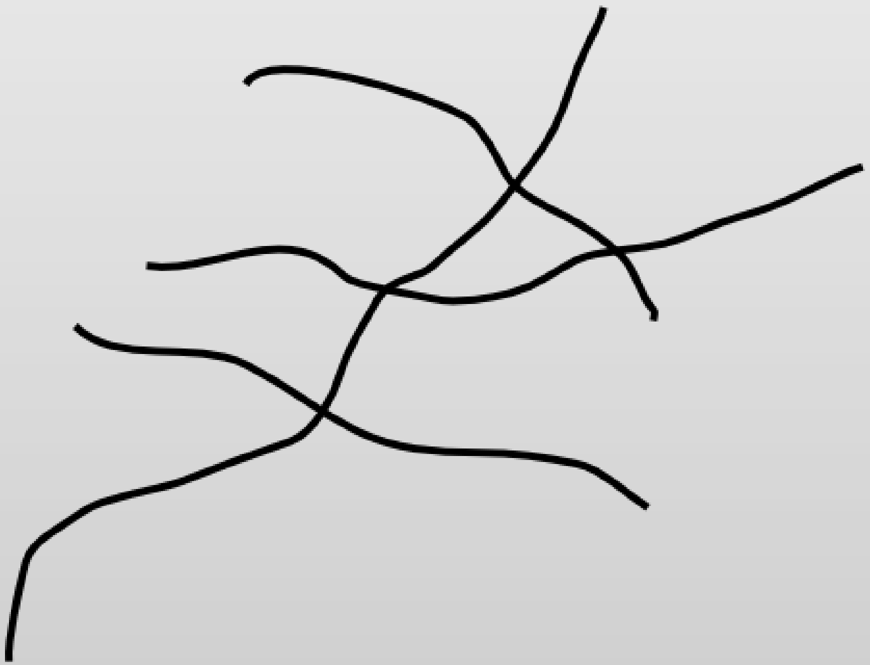
Disadvantages

- Complex data structures
- Simulation may be difficult
- Some spatial analysis is difficult or impossible to perform

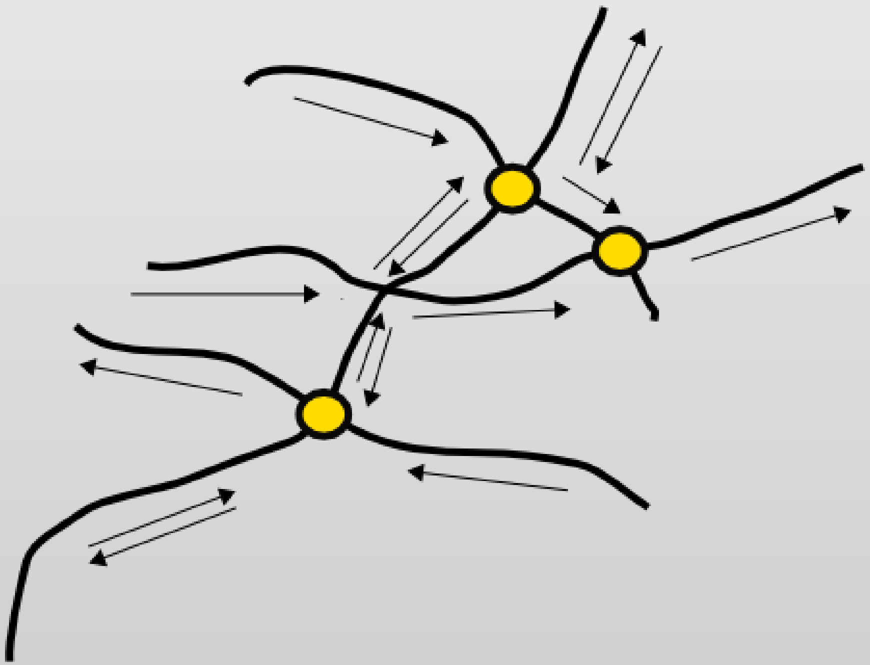
Topology: what is it?

- **in GIS, Topology is the science and mathematics of spatial relationships of vectors.**
  - A vector layer is said to contain topology if it contains the spatial relations between its features.
  - Assumption of a two-dimensional world:
    - all features are on the same plane
  - Raster files do not have topology

Topology: what is it?



A roads layer without topology



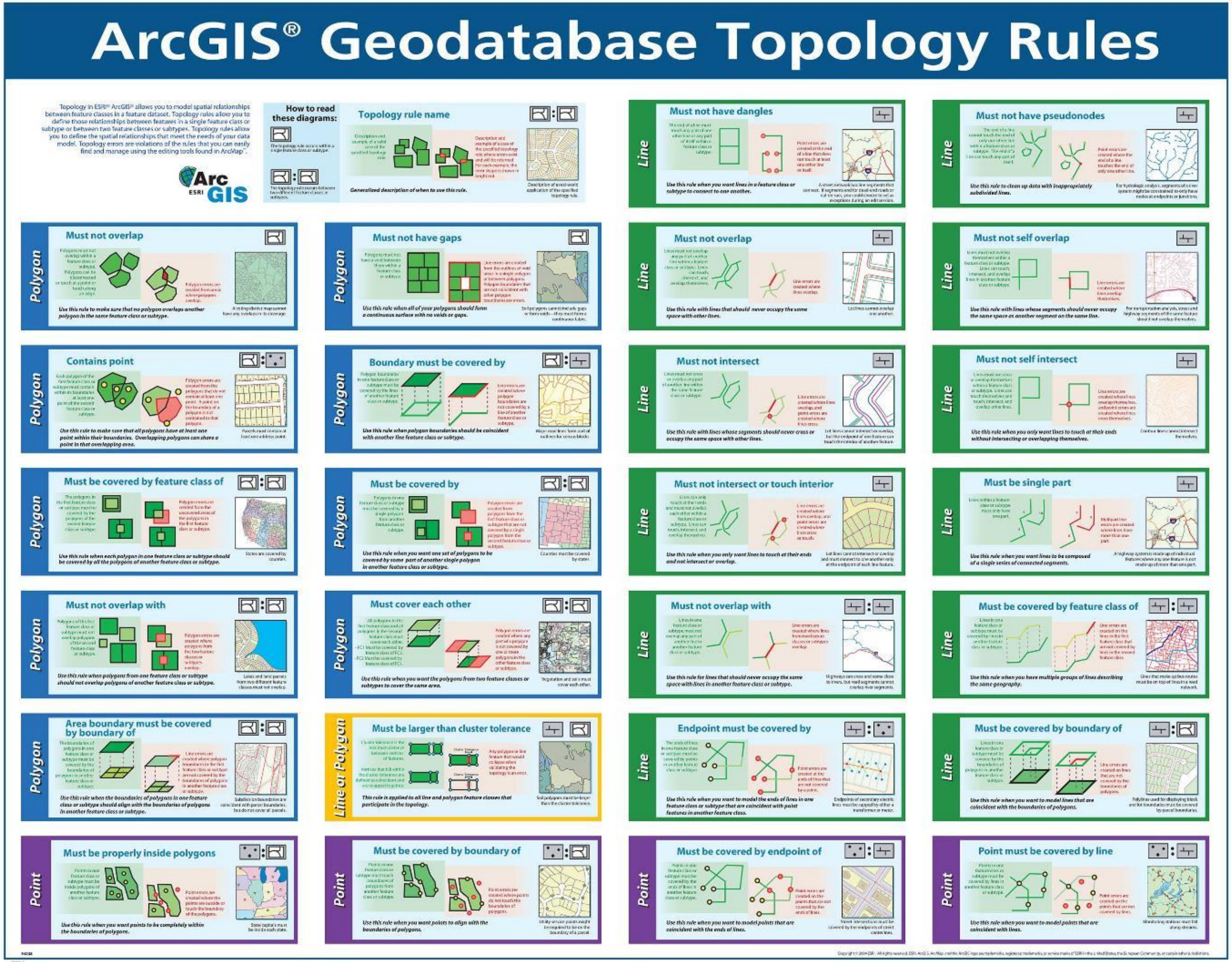
A roads layer with topology.
Circles indicate connections between roads.

Topology

- **in GIS, Topology is the science and mathematics of spatial relationships of vectors.**
  - 4 conditions of topology (i.e. spatial relationships with vector features):
    - Adjacency/Neighborhood: object with common boundary
    - Containment: object that encloses other objects ex. island in lake, house on property
    - Connectivity: links between object ex. network of roads and rivers
    - Coincidence: Objects occupying the same space

Topology

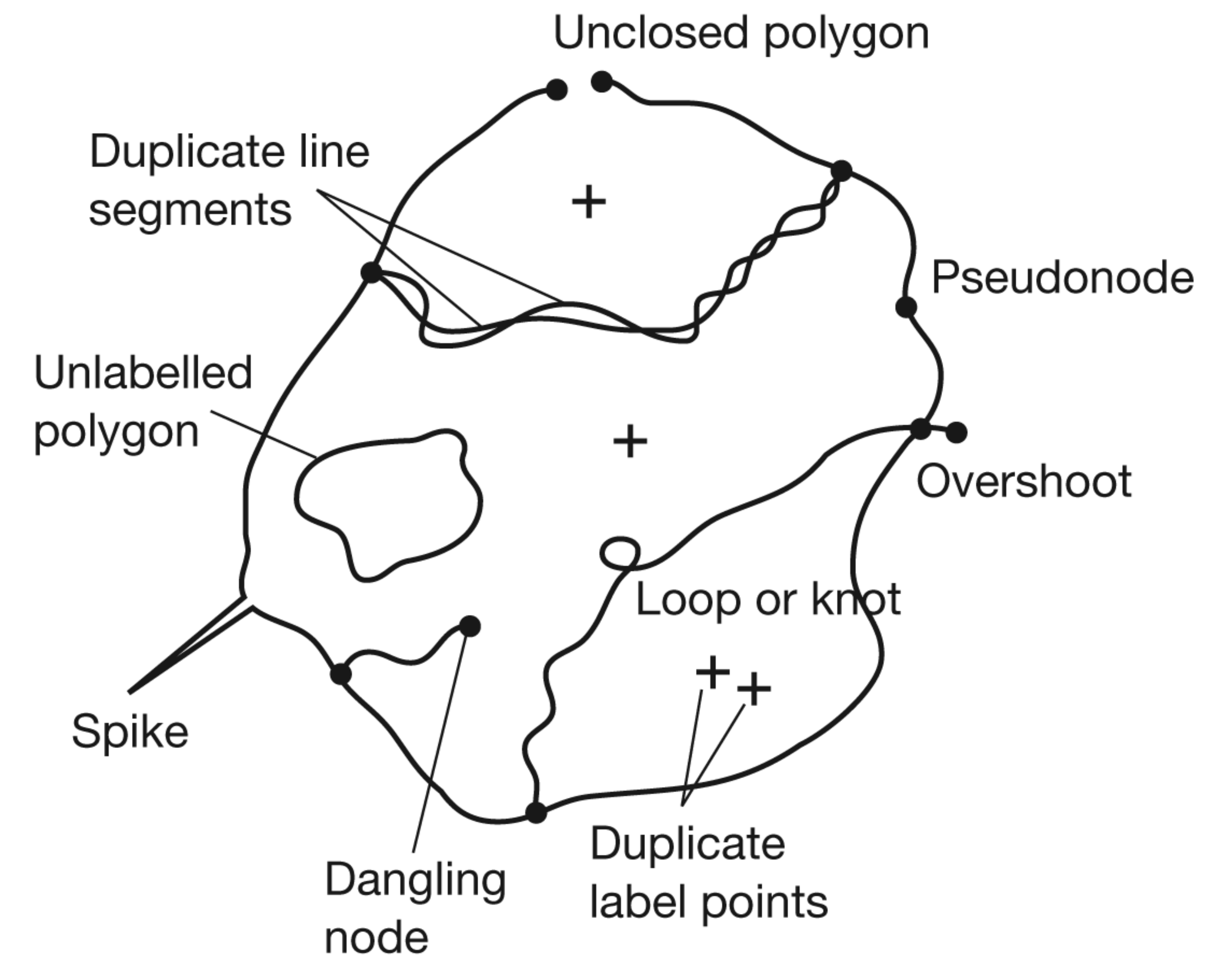
- in GIS, Topol
  relationships



https://apollomapping.com/blog/g-faq-spatial-topology-gis-part/topology_rules_poster-2

Topology: when do you need it?

- Defines rules for the geospatial data
- Support topological relationship queries.
- Prevent issues with gaps between polygons or overlaps
- When you create urban topography

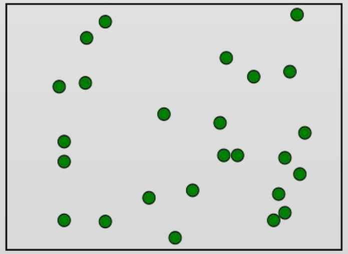
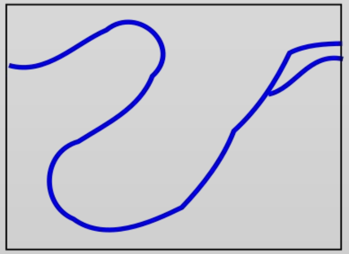
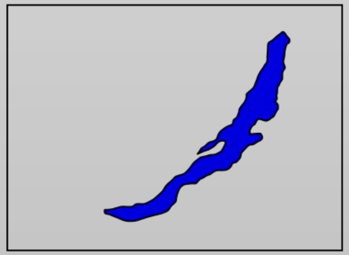


Basic characteristics of geographic data

- Spatial location
- Continuity (continuous/discrete)
- Attributes (level of measurement)

Attributes (level of measurement)

- **Attribute data** contain information about **what, where, and why of spatial data**
- It is appended in tabular format to spatial features

<table>
<tr><th>Primitive</th><th>Feature</th><th>Representation</th><th>Attributes</th></tr>
<tr><td>Points</td><td></td><td></td><td><table><tr><th>ID</th><th>HEIGHT</th><th>DIAMETER</th></tr><tr><td>1</td><td>17.5</td><td>35</td></tr><tr><td>2</td><td>22</td><td>45.6</td></tr><tr><td>3</td><td>15</td><td>27.2</td></tr><tr><td>4</td><td>19.7</td><td>36.1</td></tr><tr><td>.
.
.</td><td></td><td></td></tr></table></td></tr>
<tr><td>Lines</td><td></td><td></td><td><table><tr><th>WIDTH</th><th>DEPTH</th><th>LENGTH</th></tr><tr><td>15</td><td>4.3</td><td>35</td></tr><tr><td>6.3</td><td>3.9</td><td>5.2</td></tr></table></td></tr>
<tr><td>Polygons</td><td></td><td></td><td><table><tr><th>AREA</th><th>DEPTH</th></tr><tr><td>31494</td><td>1637</td></tr></table></td></tr>
</table>

Some examples of Attribute table for points, lines and polygons

Attributes (level of measurement)

Measurement values can be divided into four types:

- Nominal
- Ordinal
- Interval
- Ratio

Attributes (level of measurement)

- **Nominal to identify or distinguish one entity from another**
  - These values are qualities, not quantities
  - Identification / labelling of data (e.g. ID number in the attribute table)
  - These data have no particular order or hierarchy
  - Establish group, class, member, or category with which the object is associated

- *Examples: ID numbers, zip codes*

Attributes (level of measurement)

- **Ordinal**
  - Data ranked based on particular characteristics
  - Gives us insight into logical comparison of spatial objects (determine position)
  - Examples:
    - very hot, hot, cold, very cold, warm

Attributes (level of measurement)

- **Interval**
  - Number assigned to items measured
  - Measured on a **relative** scale rather than absolute scale
  - Values are **not** relative to a **true zero point** in time or space
  - Data can be compared with more precise estimates of the differences than nominal or ordinal levels
  - The difference between values makes sense but not ratio of interval data
  - Examples:
    - Time of day, calendar years, Fahrenheit temperature scale, pH values

Attributes (level of measurement)

- **Ratio**
  - Number assigned to items measured
  - Measured on a **absolute** scale rather than relative scale (use true 0 point in scaling)
  - Data can be compared with more precise estimates of the differences than nominal or ordinal levels
  - Examples:
    - Measurements of length, volume, density, total precipitation, population density, area of countries

Attributes (level of measurement): summary

- **Nominal**
  - A comparison is not possible
- **Ordinal**
  - Comparison in terms of greater than, less than, equal to
- **Interval / ratio**
  - Mathematical operations
    - Interval: addition, subtraction
    - Ratio: add, sub, multiply, divide

File format

1

GIS

- Spatial-temporal
- What is a GIS
- GIS applications

2

Digital representation of spatial data

- Geographic data
- Continuous vs discrete data
- Raster, Vector & Topology
- Attributes

3

File format

Digital representation of vector data

- **Shape Files**
  - Consisting of at least 3 files (.shp .shx .dbf)
    - Main file: objects with ID codes and geometry (* .shp)
    - Index file: indexes for each object -> navigation (* .shx)
    - Table (dBASE) with attributes (* .dbf)
  - The geometry of each object is stored as a "shape" consisting of a set of vector coordinates
  - Geometry types: Non topological data

Digital representation of vector data

- Shape Files

Three of the eight file are required

Extension	Description	Required?
.shp	The main file that stores the feature geometry. No attributes are stored in this file—only geometry	Yes
.shx	A companion file to the .shp that stores the position of individual feature IDs in the .shp file.	Yes
.dbf	The dBASE table that stores the attribute information of features.	Yes
.sbn and .sbx	Files that store the spatial index of the features.	No
.atx	Created for each dBASE attribute index created in ArcCatalog.	No
.ixs and .mxs	Geocoding index for read-write shapefiles.	No
.prj	The file that stores the coordinate system information.	No
.xml	Metadata for ArcGIS - stores information about the shapefile.	No

Digital representation of vector data

- **Many other file formats**
  - GeoJSON
  - GML
  - KML (KMZ)
  - SOSI (Norway 🇳🇴)
  - Spatialite
  - WKT
  - DXF
  - .....

GIS data formats (files)

Note: Georeferencing refers to "pinning" a raster to the earth. This is required because the pixels do not have coordinates stored within them as vector data does.

Raster data file format

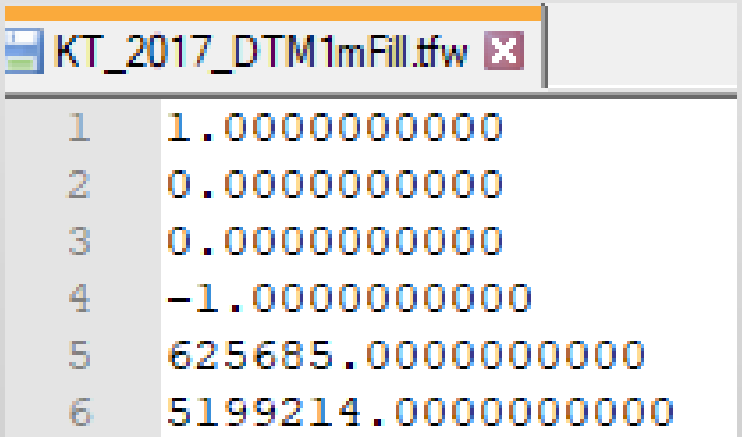
- Imagine (.img)
  - Originally created by an image processing software company called ERDAS. It is sometimes accompanied by an .xml file which usually stores metadata information about the raster layer.
- GeoTiff (.tiff)
  - A popular public domain raster data format. If maximum portability and platform independence is important, this file format may be a good choice.
  - Often is together with the **.tfw file (world file)**
- JPEG 2000 (.jpeg)
  - New JPEG format but not well supported
  - Often is together with the **. jgw file (world file)**
- "ASCII" or "GRID ASCII" (asc)
  - Older text format where you can "see" the raster data in a text editor but very slow
- File Geodatabase
  - Geodatabases have the benefit of defining image mosaic structures thus allowing the user to create "stitched" images from multiple image files stored in the geodatabase.

GIS data formats (files)

Raster data file format: World Files

- A "World" file contains the georeferencing for a raster file.
- The file is a simple text file with six floating point values, one on each line.

- The file has the following format:
  1. x-component of the pixel width (in map units) : X_{east}
  2. y-component of the pixel width (typically 0) : Y_{east}
  3. x-component of the pixel height (typically 0) : X_{north}
  4. y-component of the pixel height (typically negative) : Y_{north}
  5. X-coordinate of upper-left pixel (X_{origin})
  6. Y-coordinate of upper-left pixel (Y_{origin})


KT_2017_DTM1mFill.tfw

```
1 1.0000000000
2 0.0000000000
3 0.0000000000
4 -1.0000000000
5 625685.0000000000
6 5199214.0000000000
```

→ $Coord(i,j) = (X_{origin} + i*X_{east} + j*Y_{east}\ ,\ Y_{origin} + i*X_{north} + j*Y_{north})$

- You can georeference a JPEG file by adding a text file → extension of "jgw".
- For TIFF the extension would be "tfw".

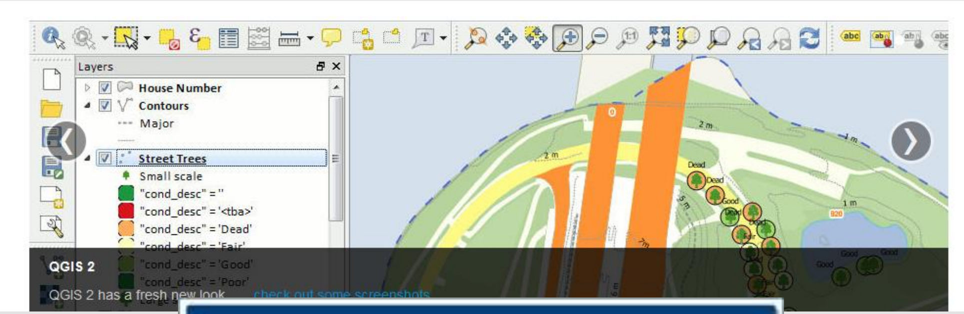
https://en.wikipedia.org/wiki/World_file

GIS Software

Open source:

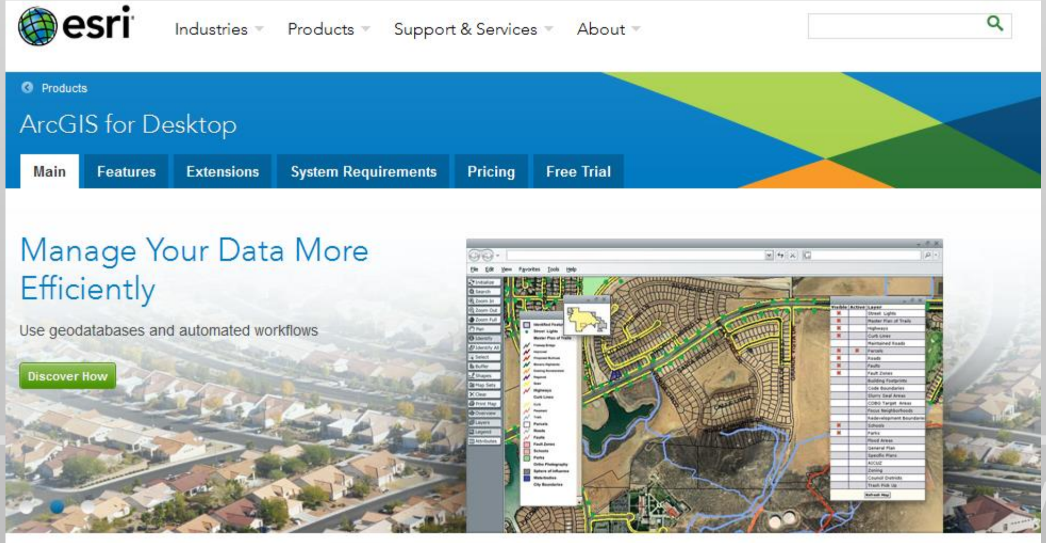
- Open Source Geospatial Foundation OSGeo (distribution of a broad set of open source geospatial software for Windows environments) http://trac.osgeo.org/osgeo4w/

  - **QGIS** (full desktop GIS)
  - gvSIG (full desktop GIS)
  - GRASS GIS (Hydrology packages)
  - SAGA GIS



Commercial:

- **ArcGIS** from ESRI
- **ENVI** from NV5 Geospatial



Summary of Today's topics

Summary: what is GIS

- **Data collection**

Import from other sources

- **Data storage**

Storage, organization and management of information

- **Data query**

Complex query options regarding the data properties

- **Data analysis**

Integration of multiple data sets, problem-based information retrieval

- **Data display**

Visualization of the data according to their respective characteristics

- **Data presentation**

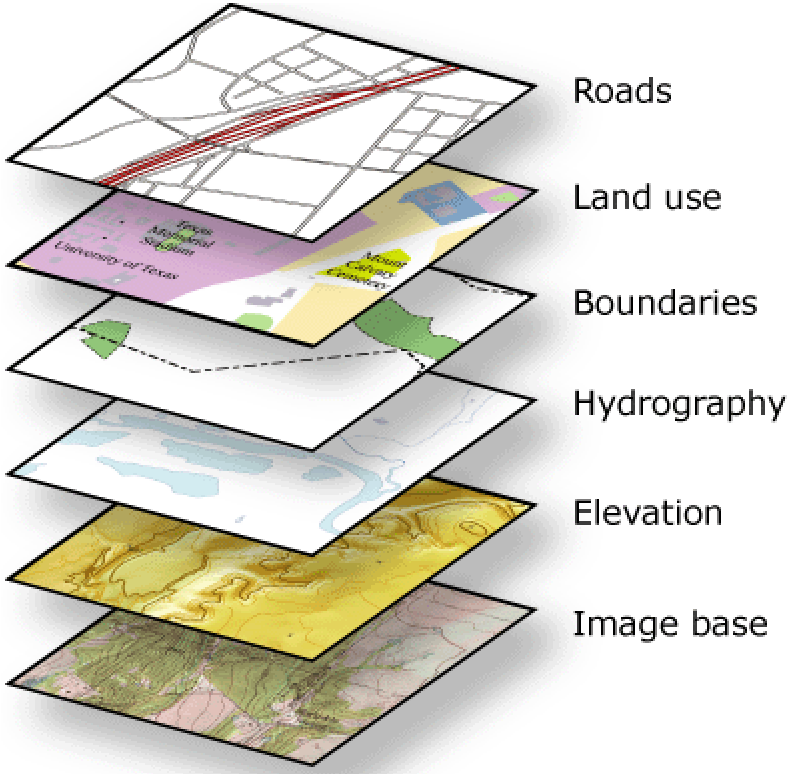
Creation and dissemination of high quality maps, graphs, reports, results of analysis

Summary: what is GIS

Answer questions by comparing different layer of data

<u>Map layers concept</u>

- each Geodata layer stores thematic spatial data
- analyse and visualize spatial dependencies by overlapping
- combine different datasets (hydrology, agriculture,...)
- spatial functions for data analysis (e.g. intersect, clip)
- perform map algebra and data modelling



ArcGIS for Desktop

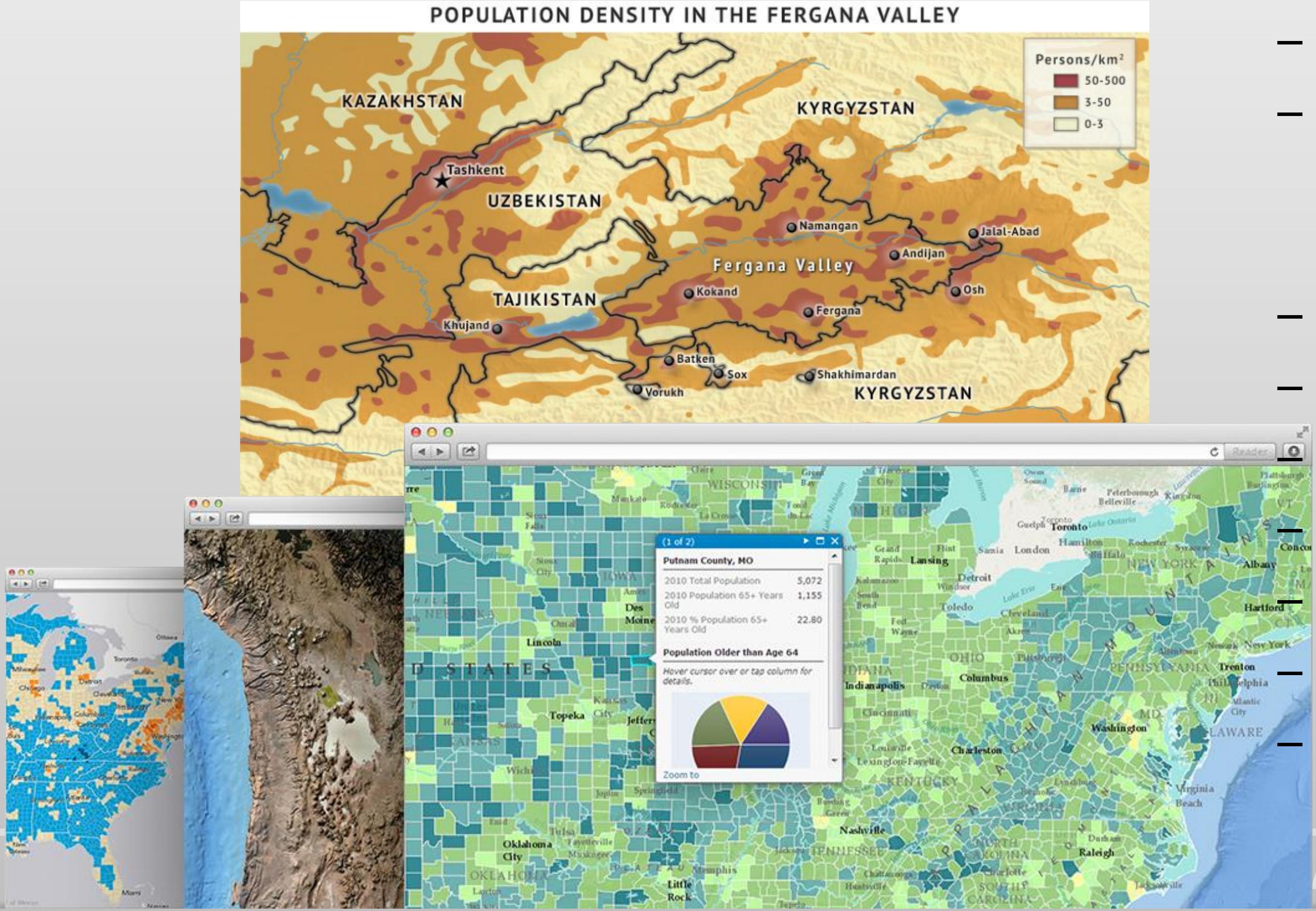
Summary: what is GIS

Answer questions by comparing different layer of data

Map layers concept

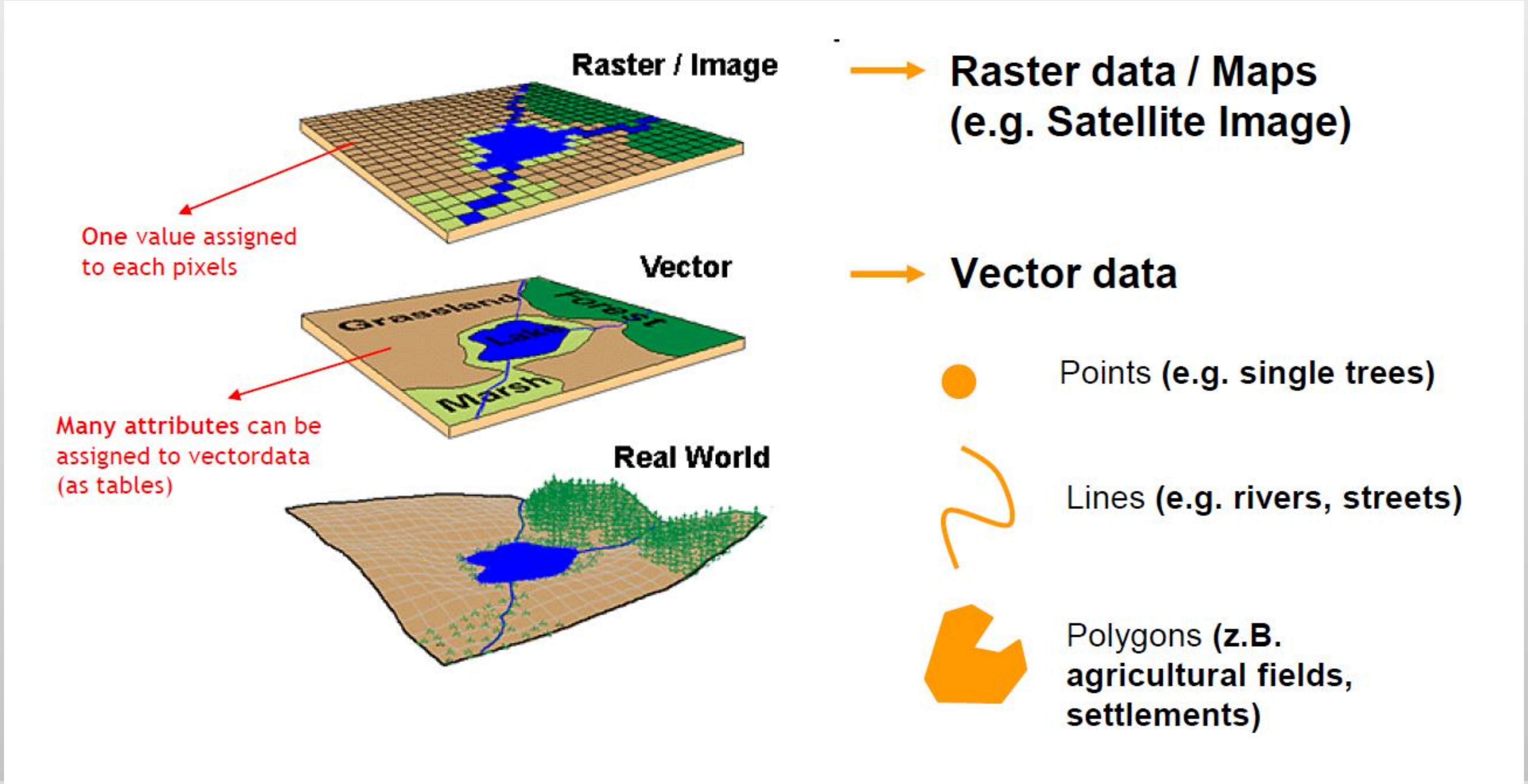
However: ensure data consistency and unitary reference system

Summary: GIS Applications



- engineering
- environment
  - water and land resources management
- planning
- health
- management
- transport/logistics
- insurance
- telecommunications
- business

Summary: Raster & Vector



Summary: Raster & Vector

Dimensions of the spatial component of geographical data