Introduction to Geographic Information Systems

Peter Galante
Biodiversity Informatics Scientist
Center for Biodiversity and Conservation
American Museum of Natural History

Outline

How do we represent real world geographic information with a computer?

- What is a Geographic Information System?
- Data models – representations of geographic reality

- GPS
- Georeferencing, sampling bias
- Remote Sensing

Notable GIS software

Commercial:

- ArcGIS, including ArcMap (Esri)
- IDRISI (Clark Labs)
- ERDAS IMAGINE (ERDAS Inc.)

Open source:

- **Quantum GIS**
- GRASS GIS
- MANIS tool
- GPSvisualizer
- Biogeomancer

Free (but not open source):

- DIVA GIS

What is a Geographic Information System?

- A Geographic Information System (GIS)
  - software, hardware, geographic information, spatially explicit data

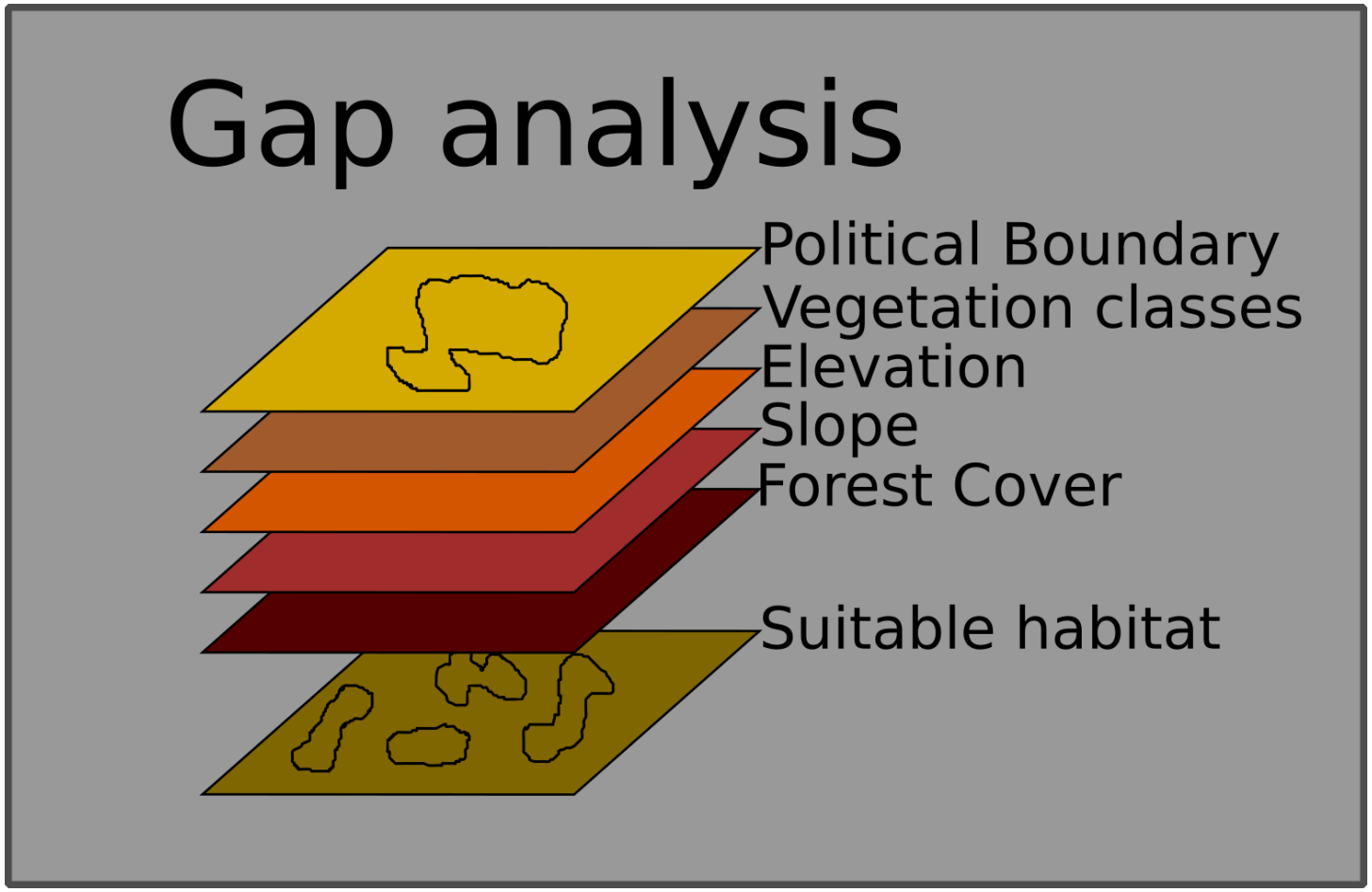
- Software can be used to store, query, manage, analyze and display spatial data.

- A GIS links cartographic, statistical, and data-basing technologies.

What is a Geographic information system?

- Data (layers)
  - Spatial – Images, shape data, geographic information
  - Tabular – Spreadsheet data with geographic properties

Layering


Gap analysis
Political Boundary
Vegetation classes
Elevation
Slope
Forest Cover
Suitable habitat

Outline

How do we represent real world geographic information with a computer?

- What is a Geographic Information System?
- Data models – representations of geographic reality

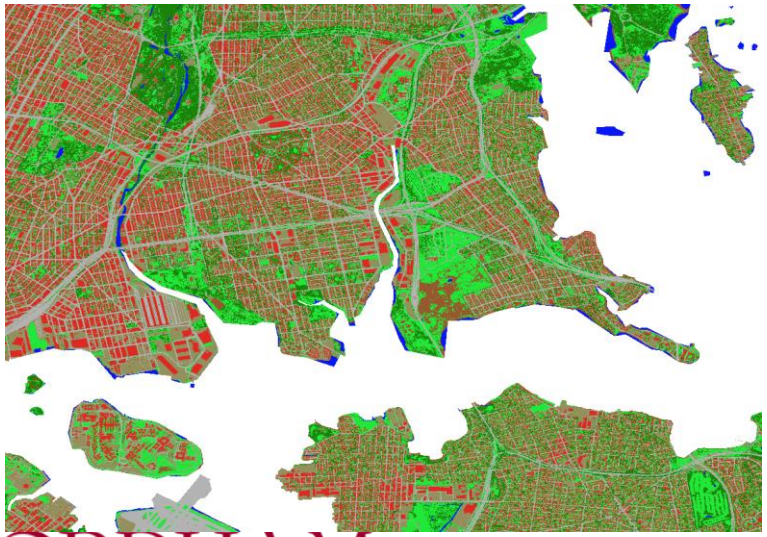
- Georeferencing, sampling bias
- Remote Sensing

Two methods of GIS representation

- There are two main ways a GIS can create a logical data model of the world:
  - Raster = continuous fields
  - Vector = discrete features

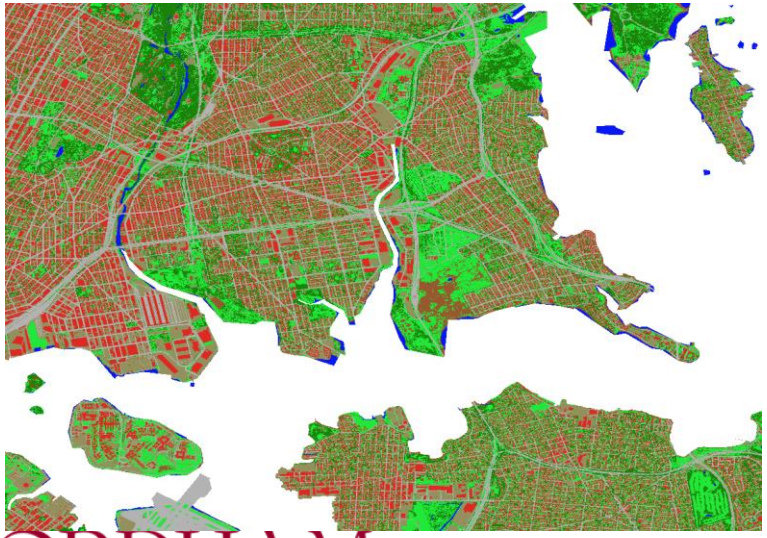
Two methods of GIS representation

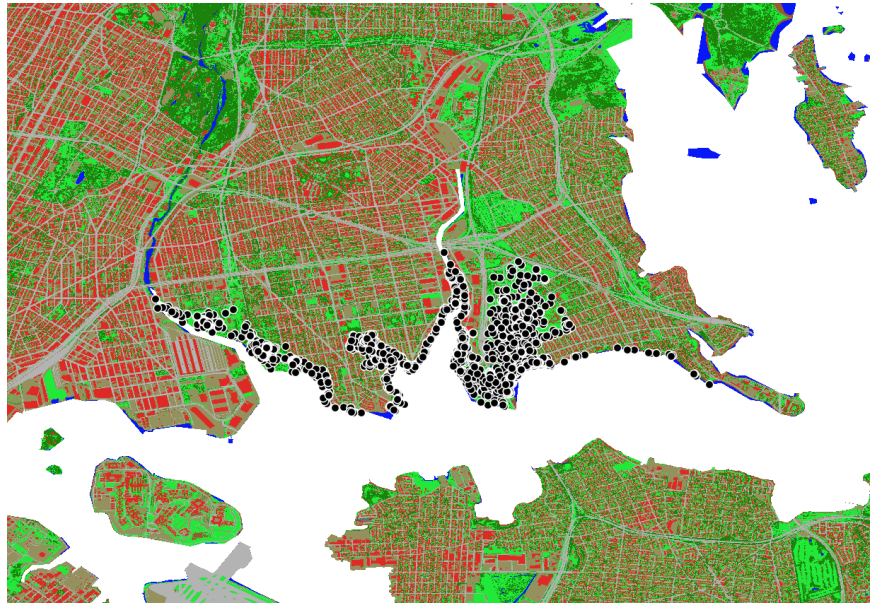
- There are two main ways a GIS can create a logical data model of the world:
  - Raster = continuous fields
  - Vector = discrete features



Two methods of GIS representation

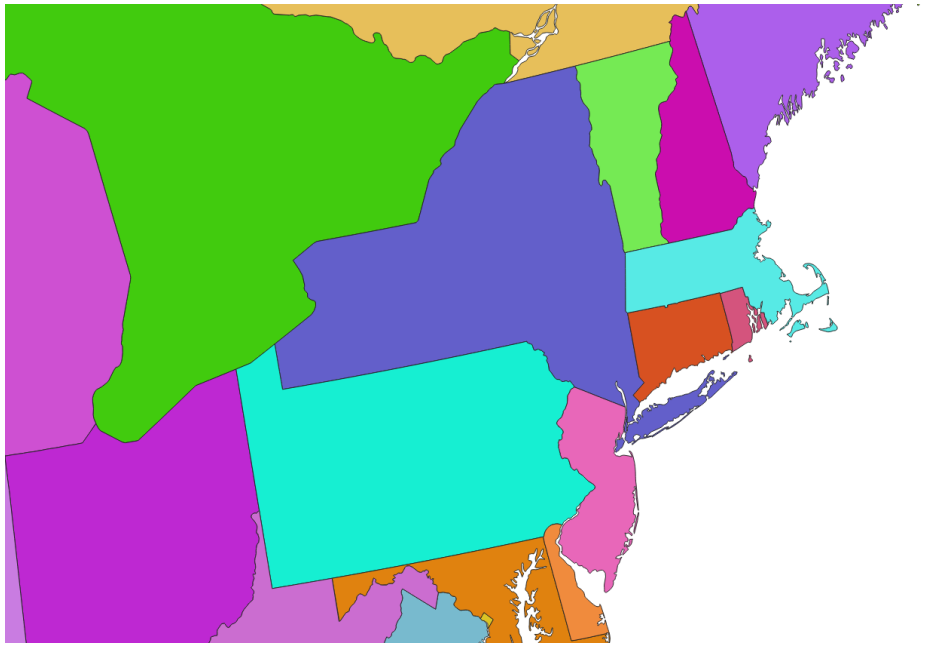
- There are two main ways a GIS can create a logical data model of the world:
  - Raster = continuous fields
  - Vector = discrete features





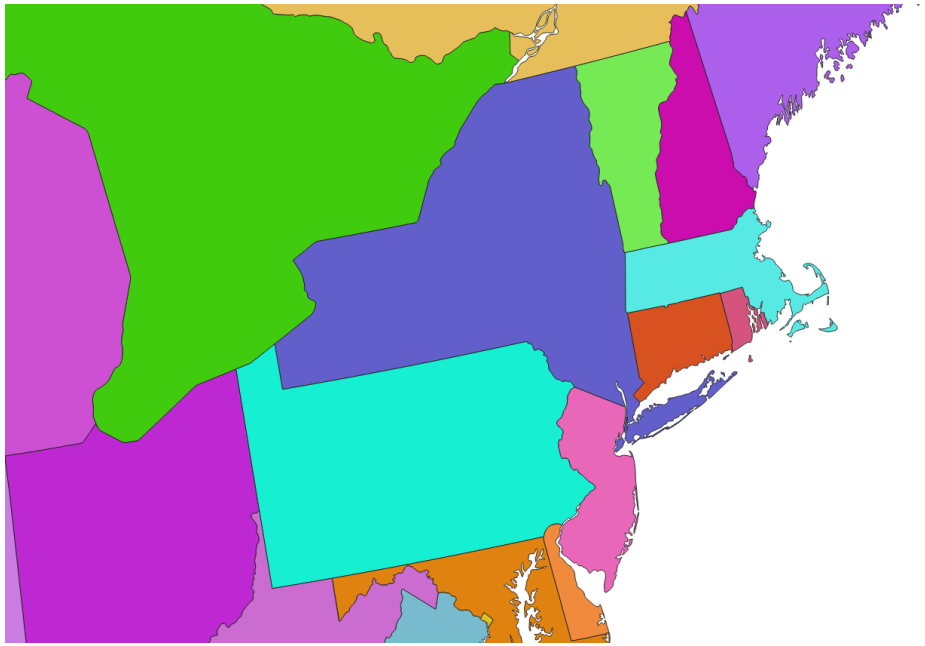
Vector

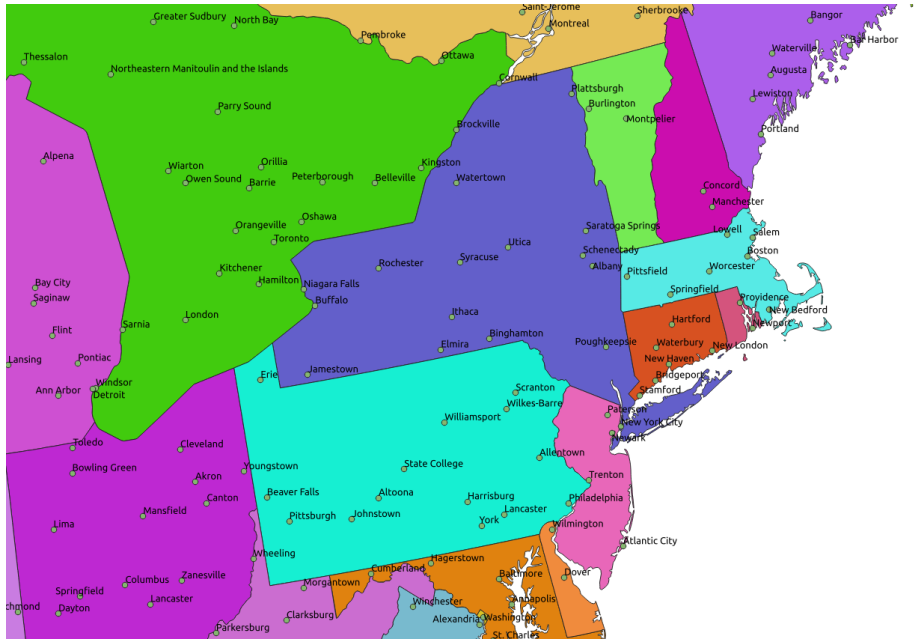
- Discrete objects
  - Represents the geographic world as objects with well-defined boundaries in otherwise empty space



Vector

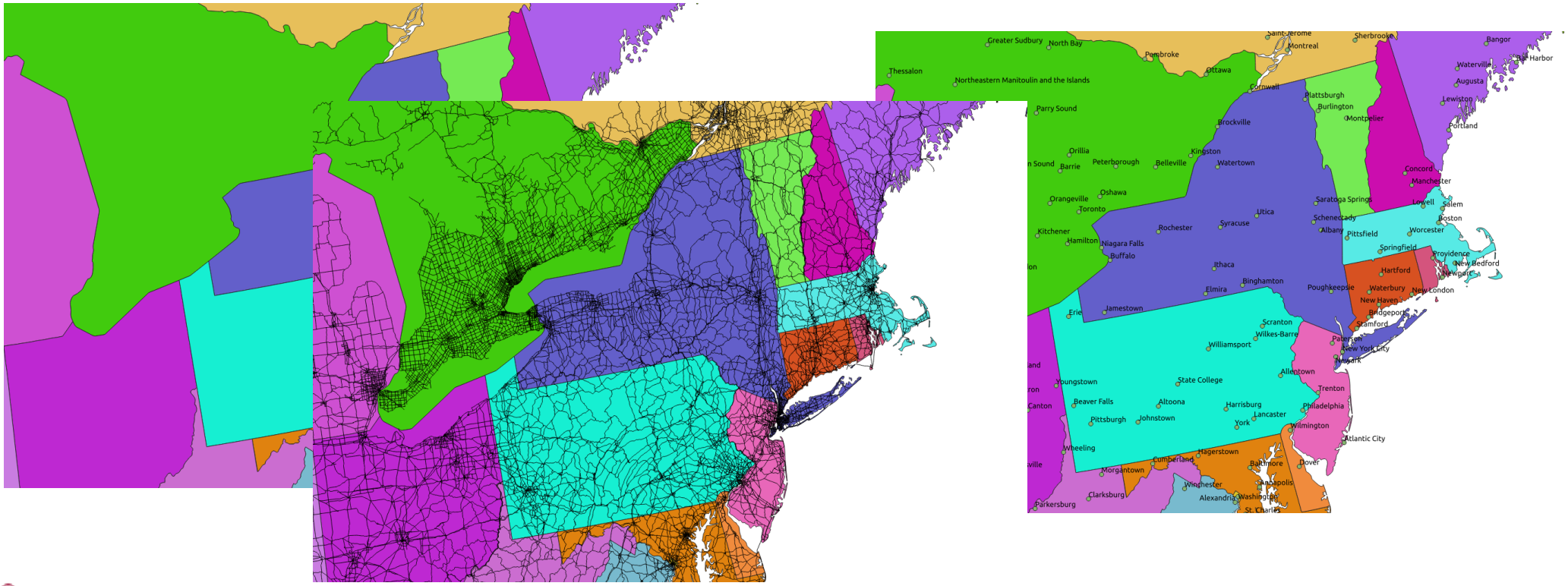
- Discrete objects
  - Represents the geographic world as objects with well-defined boundaries in otherwise empty space





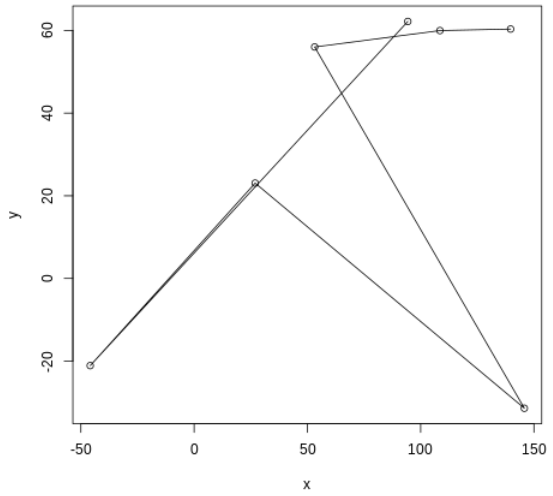
Vector

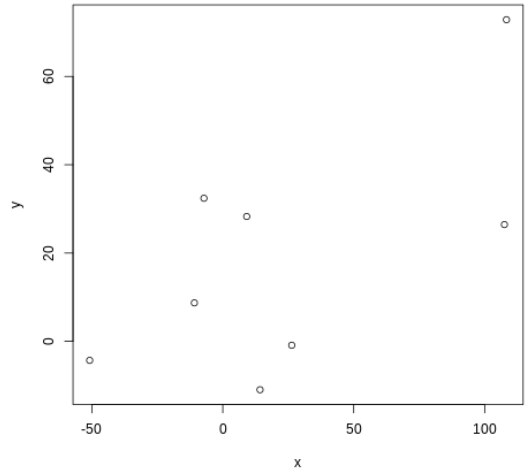
- Discrete objects
  - Represents the geographic world as objects with well-defined boundaries in otherwise empty space



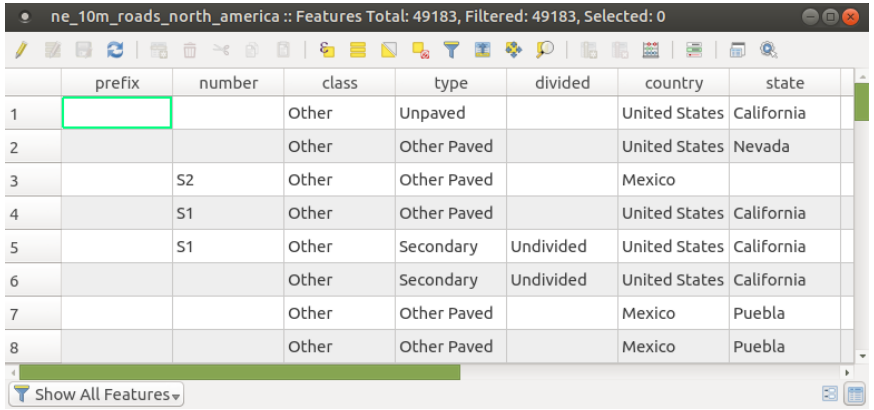
Vector Data

- Vector data represents features as *points*, *lines*, and *polygons*.
- It is best applied to discrete objects with defined shapes and boundaries.
- It is tied with tabular data





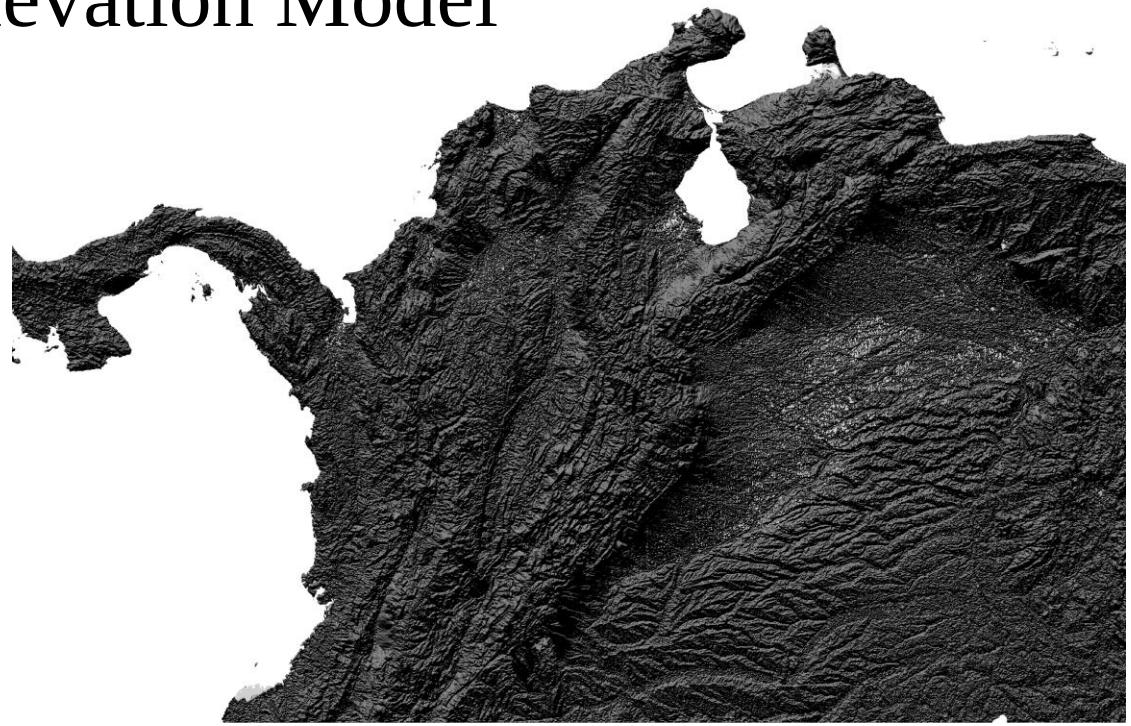
- Examples include:
  - GPS points
  - Roads
  - Protected areas



FORDHAM
THE JESUIT UNIVERSITY OF NEW YORK

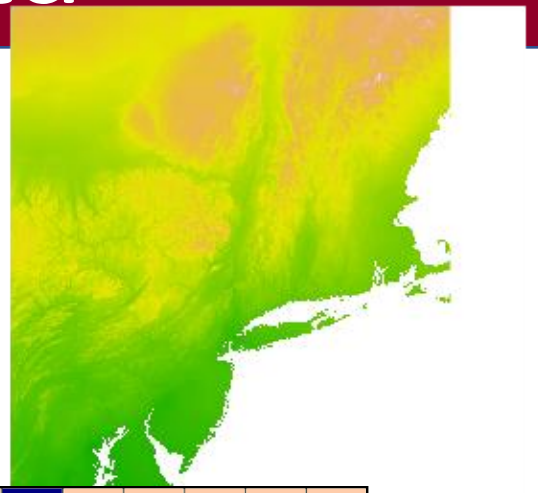
Raster

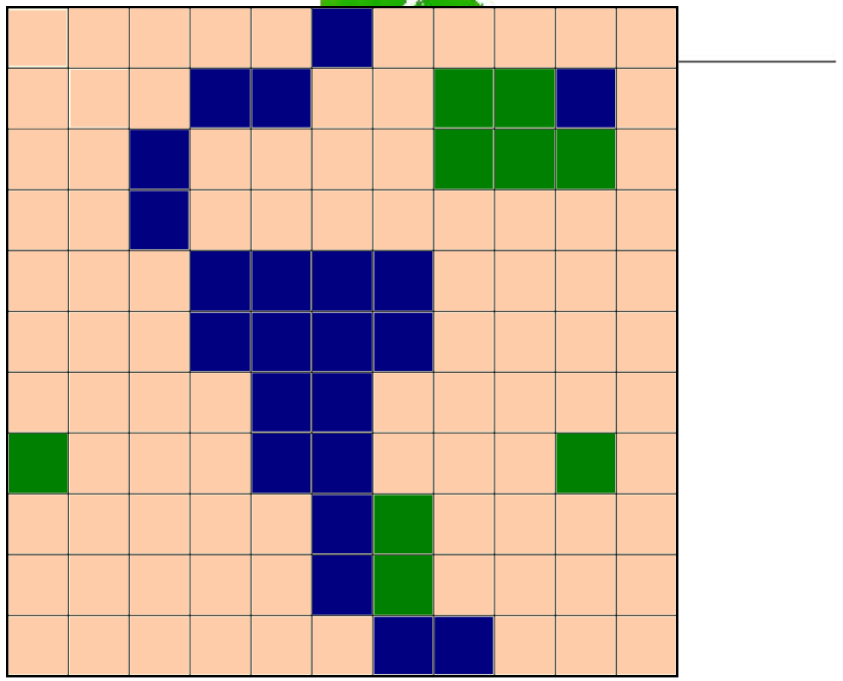
- Continuous fields
  - Represents the real world as a finite number of variables, each one defined at every possible position.
- E.g. a Digital Elevation Model



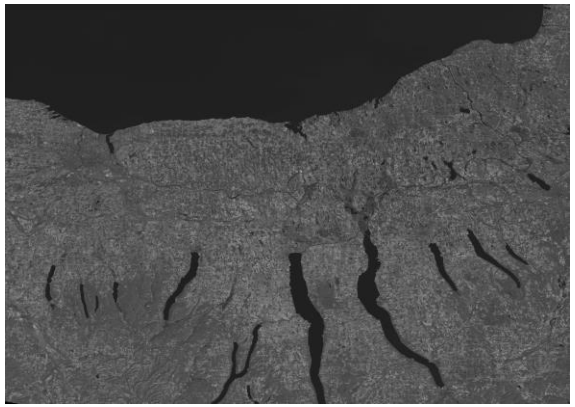
Raster Data

A raster model represents the world as a surface that is divided into a regular grid of cells, each of which are assigned a value.





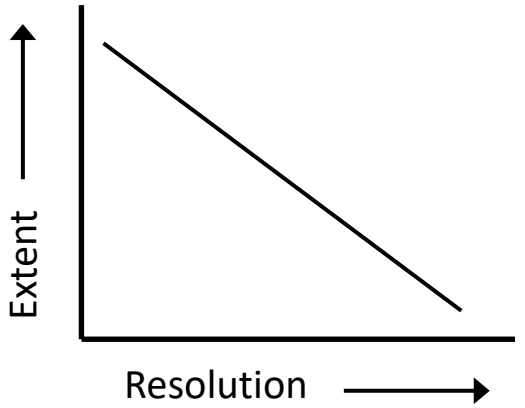
Raster data includes images and grids. aerial photograph, a satellite image, or a scanned map, are often used for generating raster datasets.



General data issues: spatial scale

Spatial scale has two elements: **resolution** and **extent**

Resolution and extent
tend to be
inversely related



Vector data

- Advantage
  - precise representation of points, boundaries, and linear features
    - defining spatial relationship (connectivity and adjacency) between coverage features
    - network analysis (for example to find an optimal path between two nodes in a complex transport network)
- Disadvantage
  - boundaries of the resulting map polygons are discrete/enclosed by well-defined boundary lines. In reality the map polygons may represent continuous gradation or gradual change, as in soil maps
  - Resolution is not clear – how many points were used to make up a line?

Raster data

- Advantage
  - Represent indistinct boundaries
    - thematic information on soil types, soil moisture, vegetation, ground temperatures
  - Reconnaissance satellites and aerial surveys use raster-based scanners
    - scanned images can be directly incorporated into GIS
- Potential disadvantage
  - The higher the grid resolution, the larger the data file is going to be
  - If converting from vector to raster can lose information:

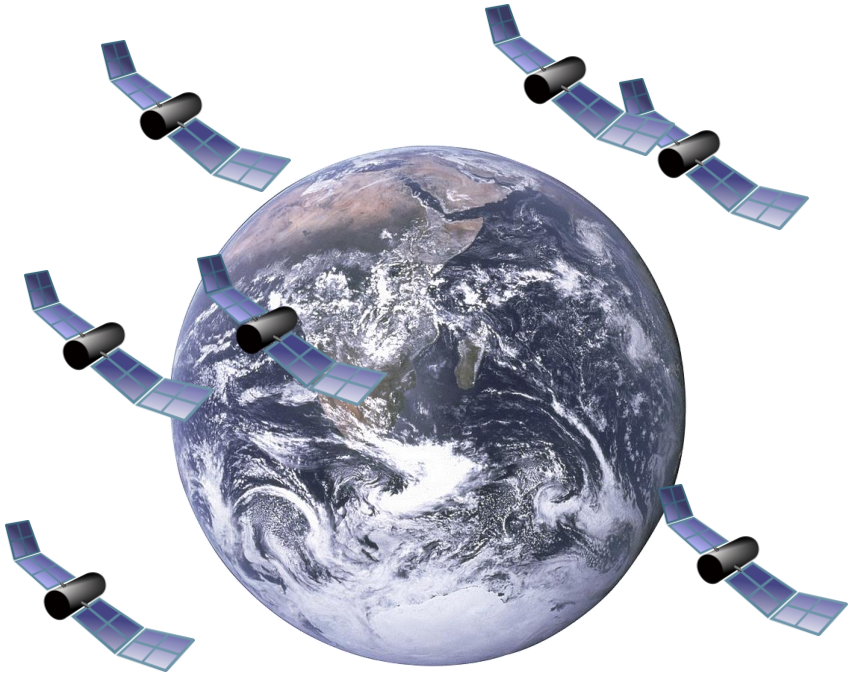
Outline

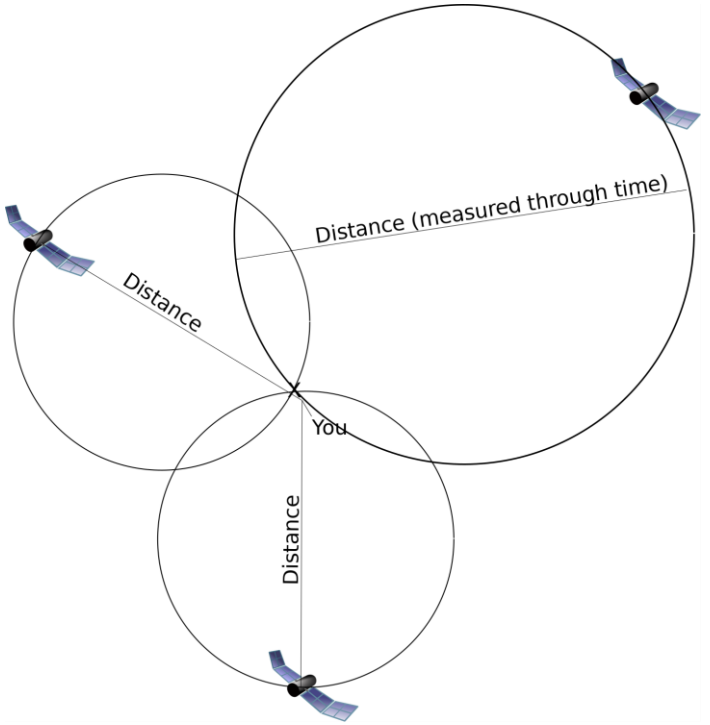
How do we represent real world geographic information with a computer?

- What is a Geographic Information System?
- Data models – representations of geographic reality

- GPS
- Georeferencing, sampling bias
- Remote Sensing

GPS





Outline

How do we represent real world geographic information with a computer?

- What is a Geographic Information System?
- Data models – representations of geographic reality

- GPS
- Georeferencing, sampling bias
- Remote Sensing

Species' distribution data: possible sources

- Personal collection during field surveys (e.g., using a GPS receiver or smartphone)
- Extracting data from large surveys (e.g., North American Breeding Bird Survey)
- Digitizing atlases (e.g. *The new atlas of breeding birds in Britain and Ireland: 1988-1991*)
- Collections in natural history museums
- On-line distributed databases (e.g. GBIF, HerpNET, FishNET, ORNIS)

Georeferencing

- Woodbury, NY
- Near Woodbury
- 5 miles from Woodbury
- 1 mile east of Woodbury
- 10 minutes along the path leading out of Woodbury

FORDHAM
THE JESUIT UNIVERSITY OF NEW YORK

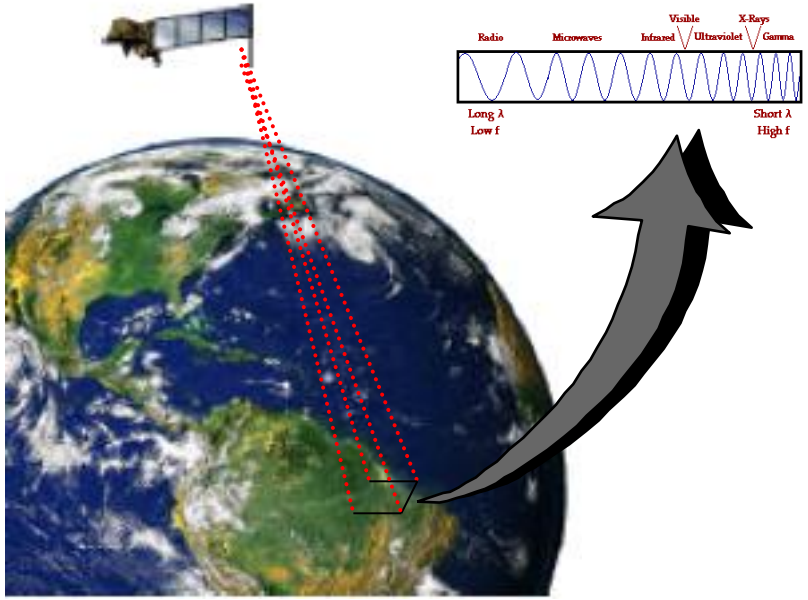
Outline

How do we represent real world geographic information with a computer?

- What is a Geographic Information System?
- Data models – representations of geographic reality

- GPS
- Georeferencing, sampling bias
- Remote Sensing

What is remote sensing?



Source: P. Ersts, July 2004

General definition: *Acquiring information about an object without physical contact.*

Definition in context of Earth observation: *A technology for sampling reflected and emitted electromagnetic (EM) radiation of features on the Earth's land surface, oceans, and atmosphere.*

GIS projects

The majority of your grade for this course will be based on a comprehensive final project. Ideally, your project will represent some aspect of your graduate research. It should be of sufficient quality to be used in a publication, whether as just a map (or series of maps), supporting documentation, or values generated as part of an analysis.

If you do not have a project, or an interesting spatial aspect of a project, we can chat and come up with one.