AMERICAN MUSEUM OF NATURAL HISTORY
CENTER FOR BIODIVERSITY AND CONSERVATION

Introduction to Geographic Information Systems

Peter Galante
Biodiversity Informatics Specialist
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
- Location information: coordinate systems

- 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

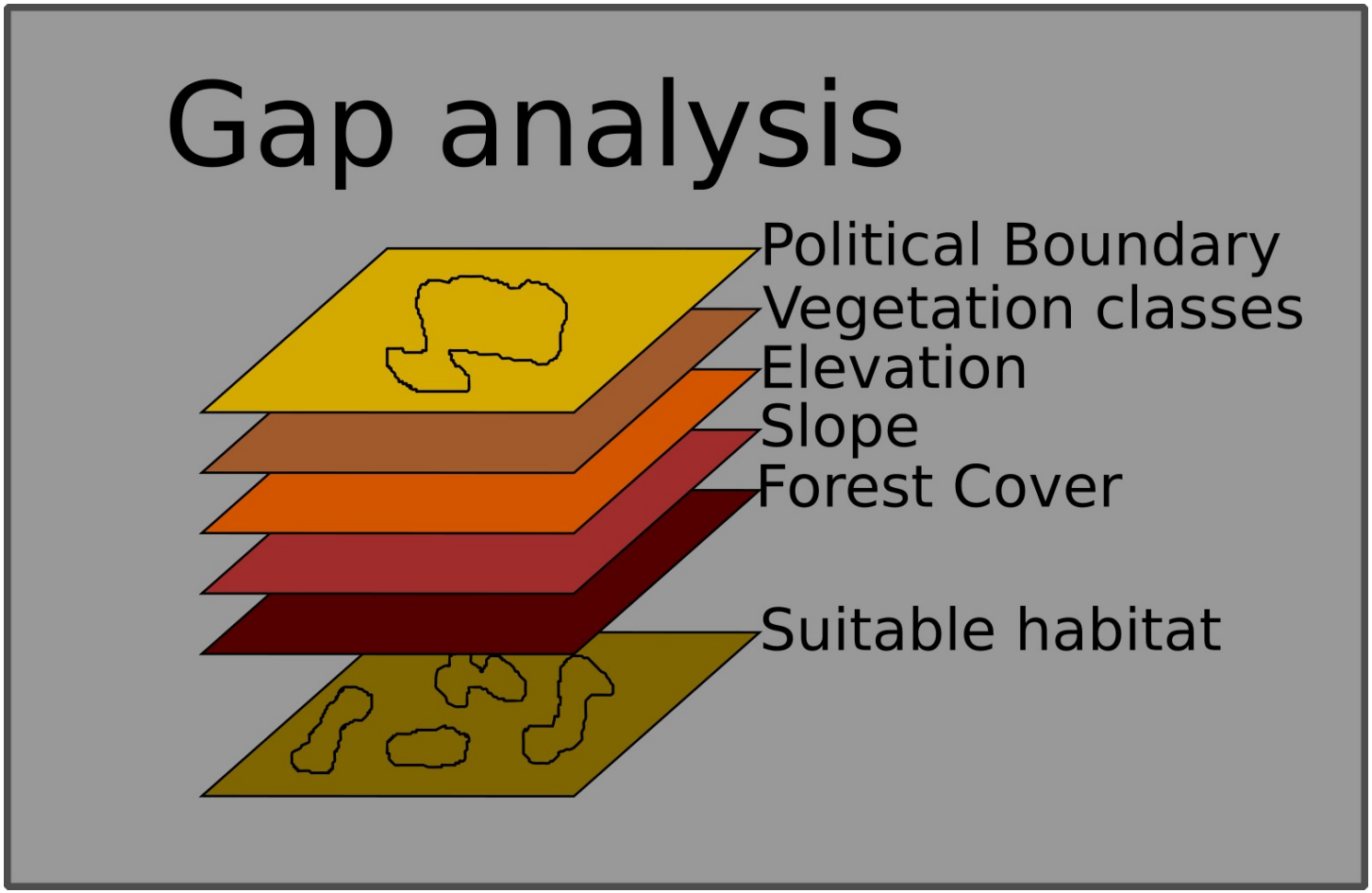
- Software can be used to store, query, manage, analyze and display spatial data.

- 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
- Location information: coordinate systems

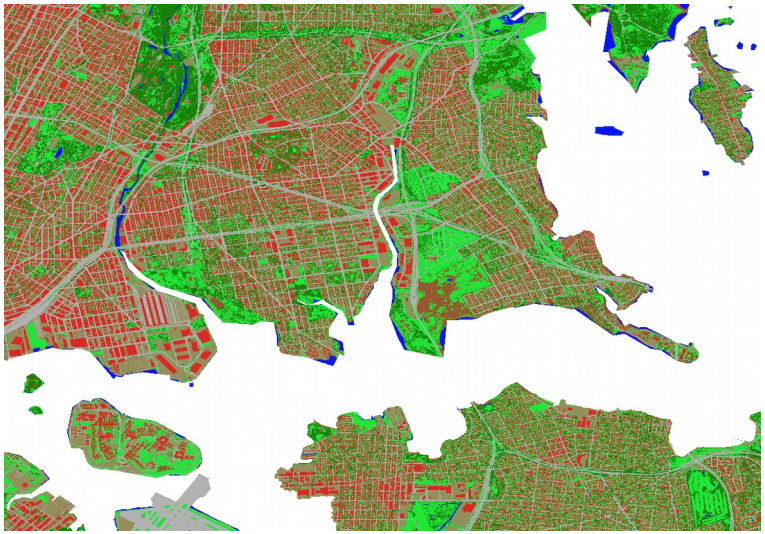
- Georeferencing, sampling bias
- Remote Sensing

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

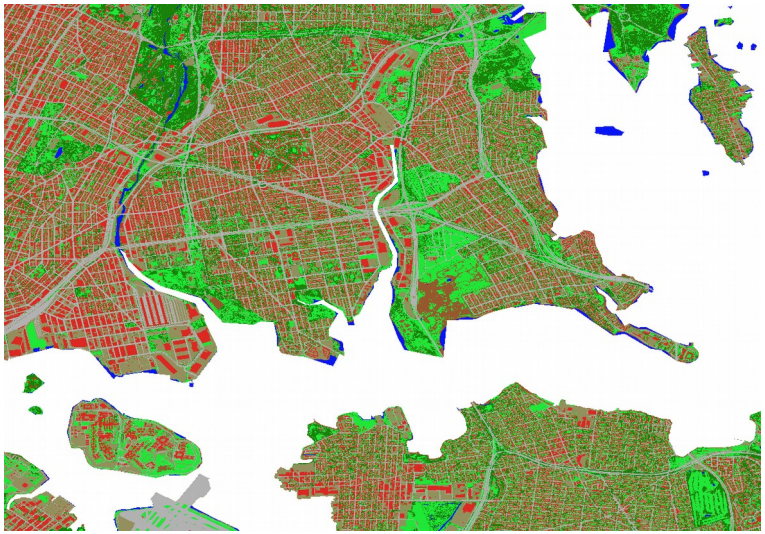
- There are two main ways a GIS can create a logical data model of the world:
  - Raster = continuous fields
  - Vector = discrete features



New York City landcover
classified at 3 foot resolution

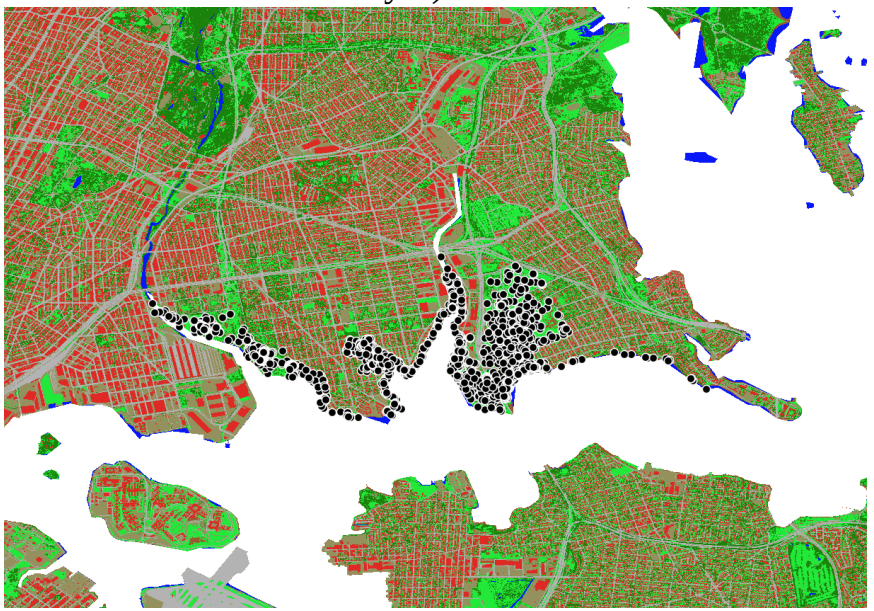
Two methods of GIS representation

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



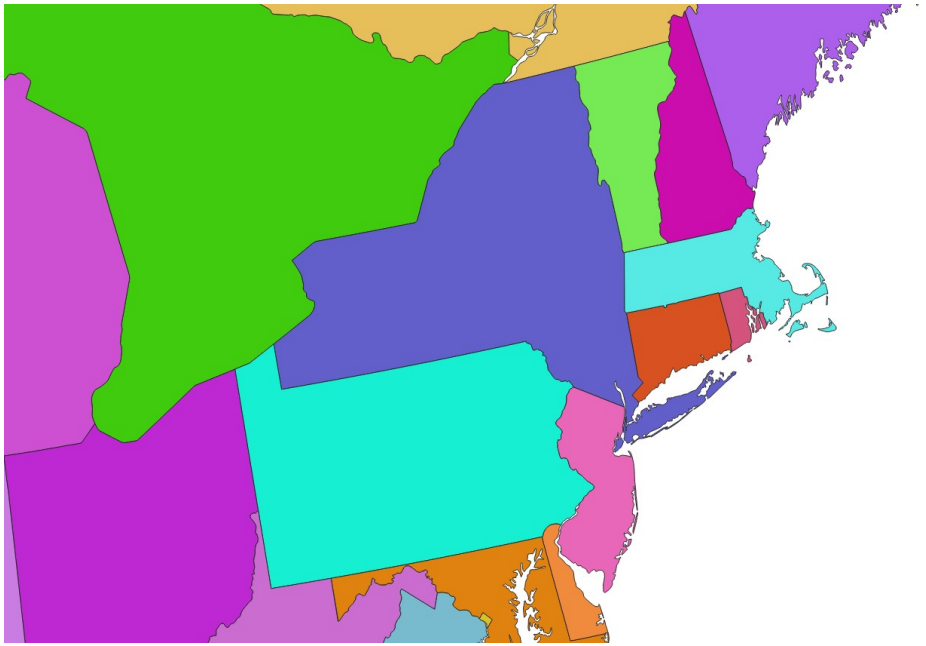
New York City landcover classified at 3 foot resolution

Coyote occurrence records (as vector layer)



Vector

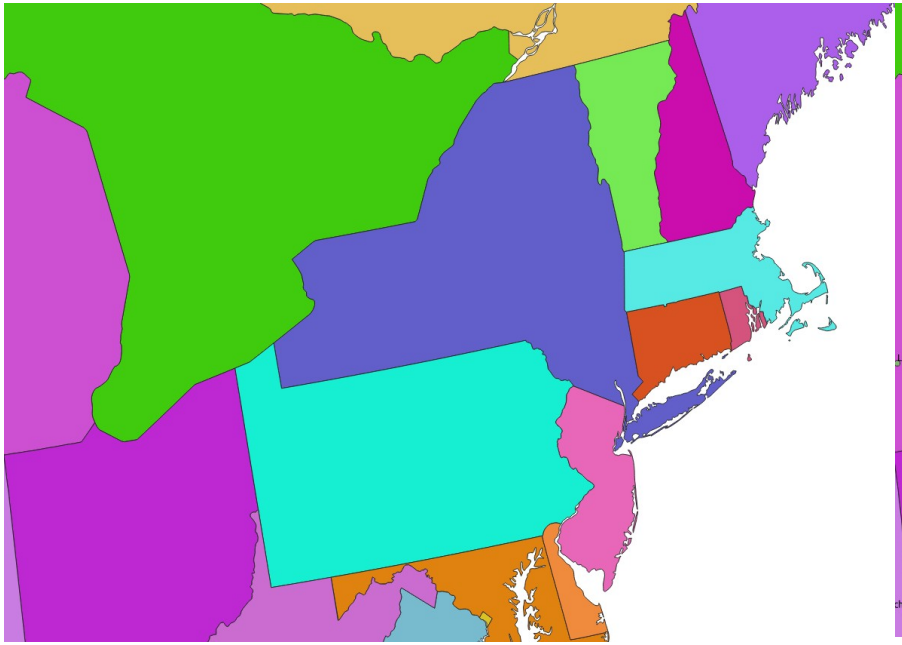
- Discrete objects
  - Represents the geographic world as objects with well-defined boundaries in otherwise empty space

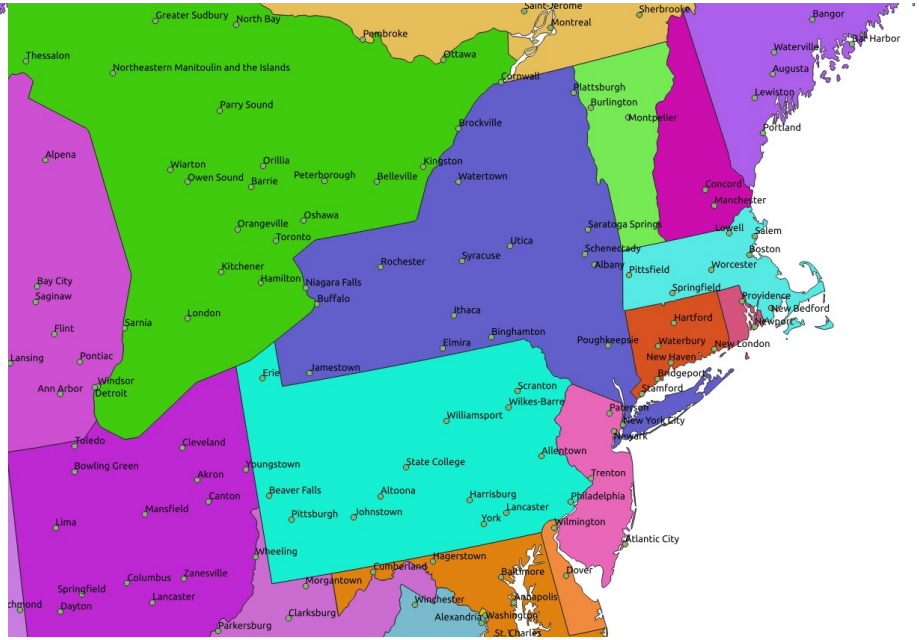


Data from
naturalearthdata.com

Vector

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

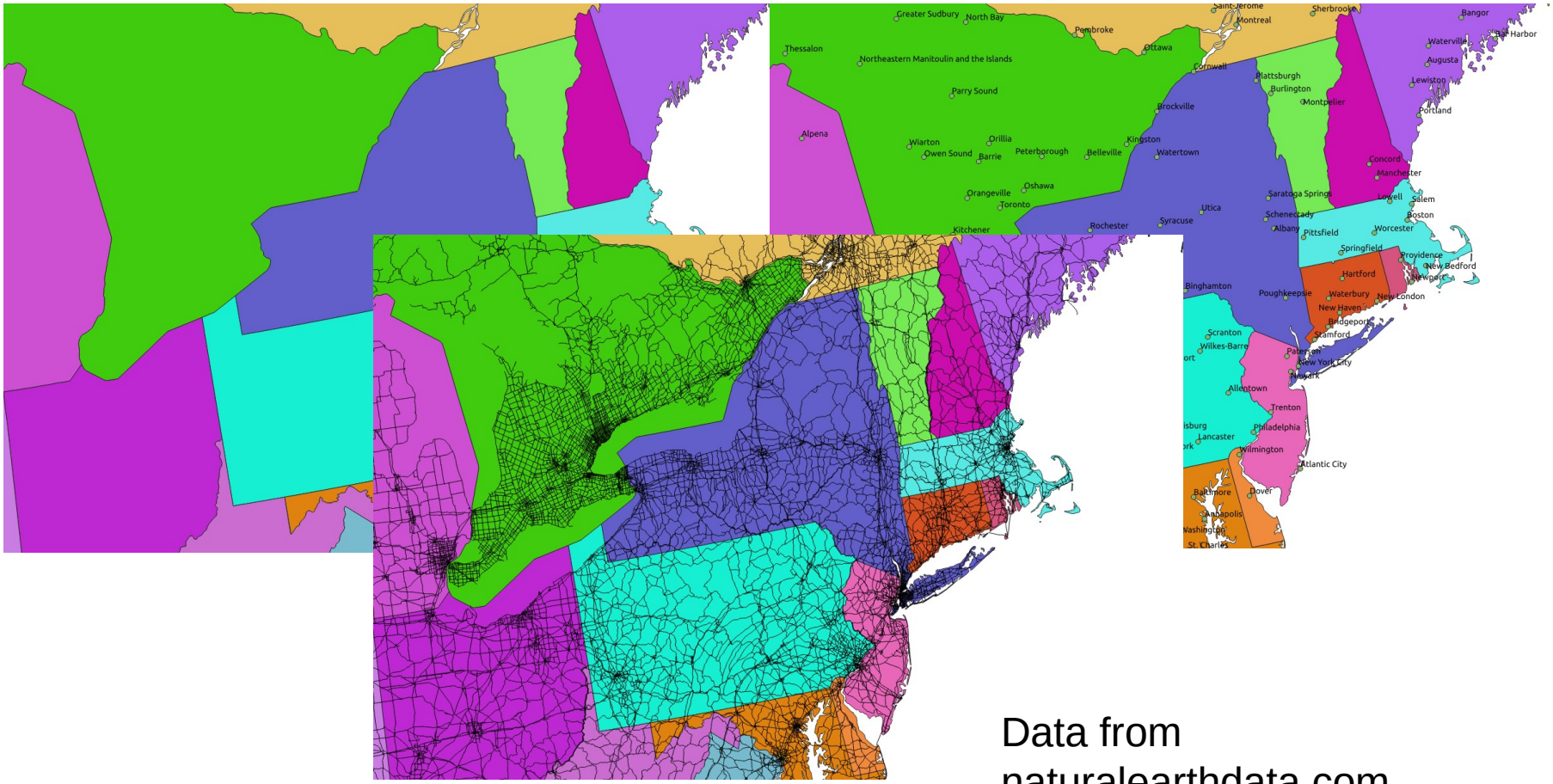




Data from naturalearthdata.com

Vector

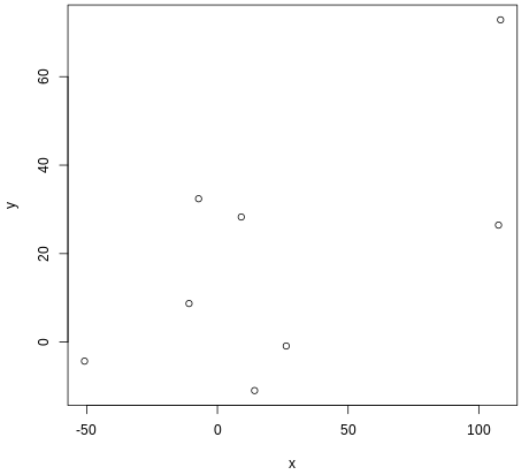
- Discrete objects
  - Represents the geographic world as objects with well-defined boundaries in otherwise empty space

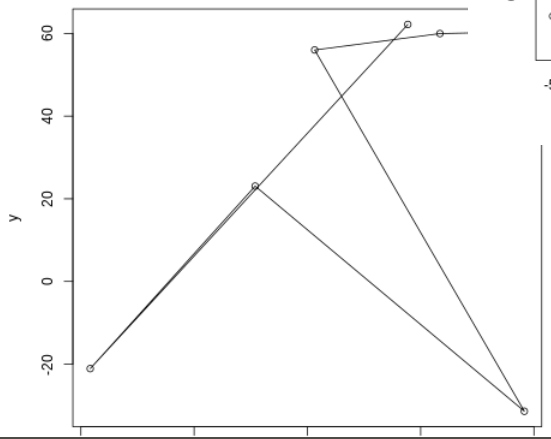


Data from naturalearthdata.com

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

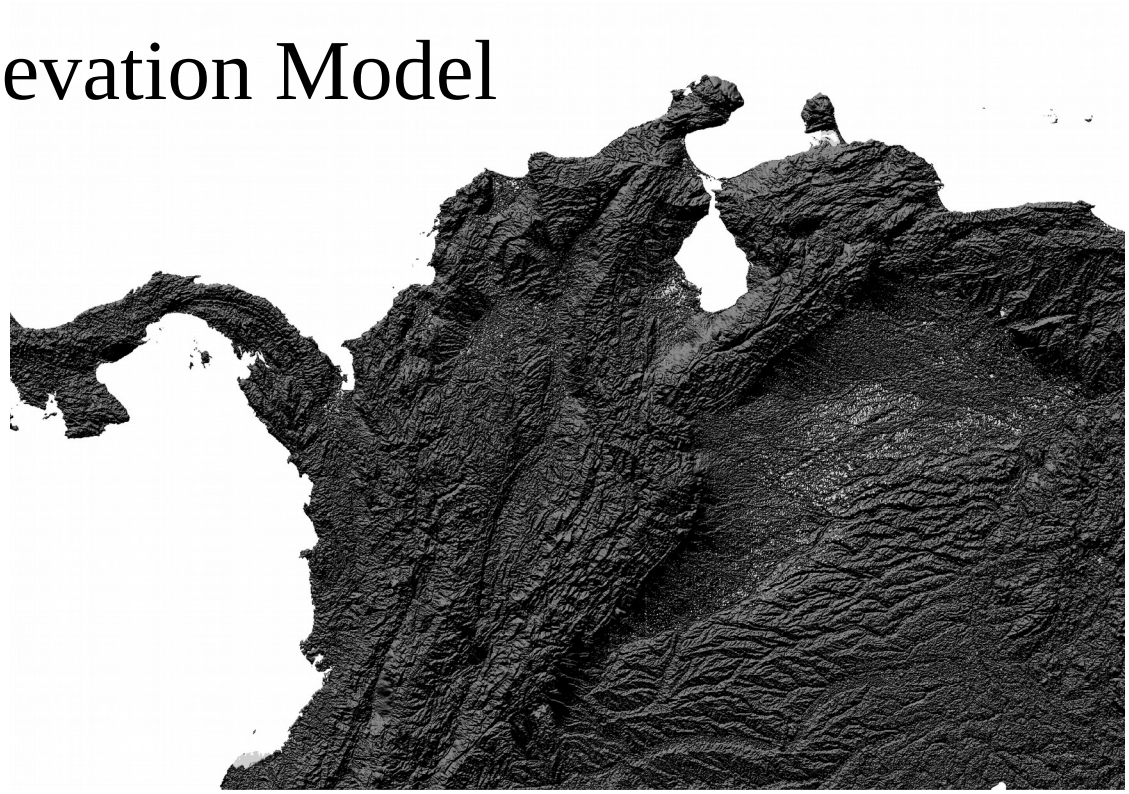
ne_10m_roads_north_america :: Features Total: 49183, Filtered: 49183, Selected: 0

	prefix	number	class	type	divided	country	state
1			Other	Unpaved		United States	California
2			Other	Other Paved		United States	Nevada
3		S2	Other	Other Paved		Mexico	
4		S1	Other	Other Paved		United States	California
5		S1	Other	Secondary	Undivided	United States	California
6			Other	Secondary	Undivided	United States	California
7			Other	Other Paved		Mexico	Puebla
8			Other	Other Paved		Mexico	Puebla

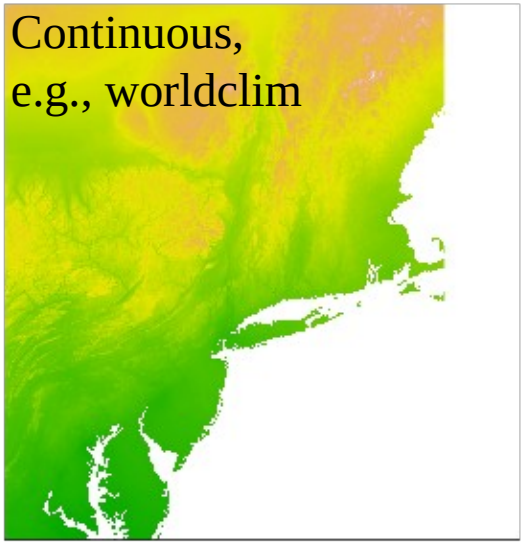
Show All Features

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

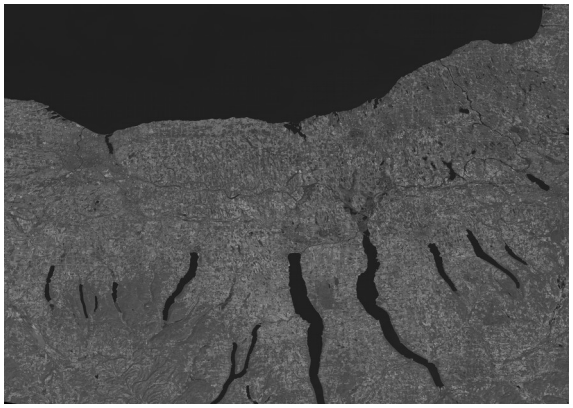


Raster Data

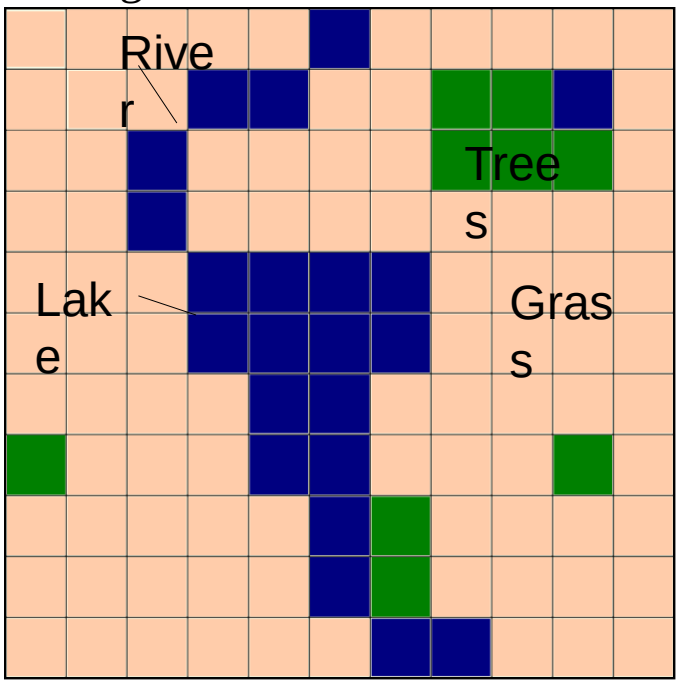


A raster model represents the world as a surface that is divided into a regular grid of cells, each of which are assigned a value.

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



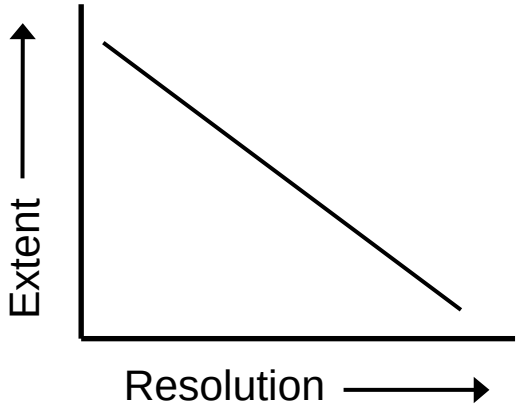
Categorical



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

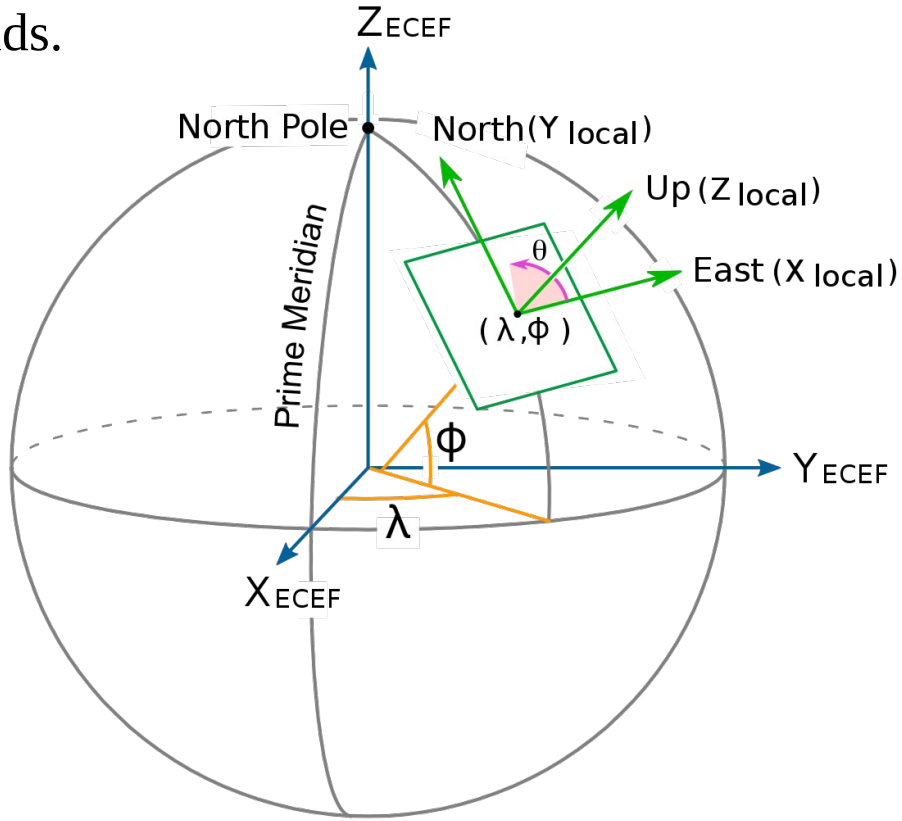
How do we represent real world geographic information with a computer?

- What is a Geographic Information System?
- Data models – representations of geographic reality
- Location information: coordinate systems

- Georeferencing, sampling bias
- Remote Sensing

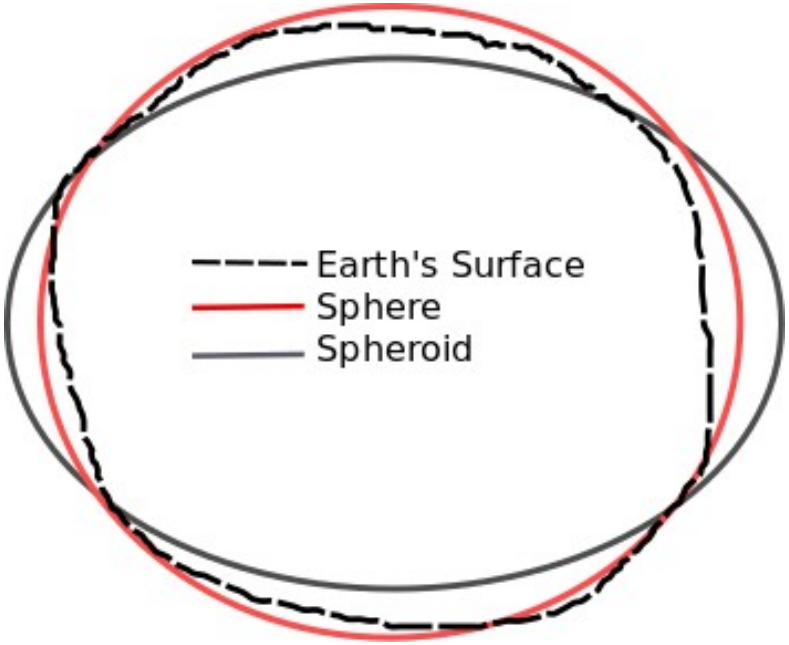
GIS fundamentals: Geographic coordinate systems (GCS)

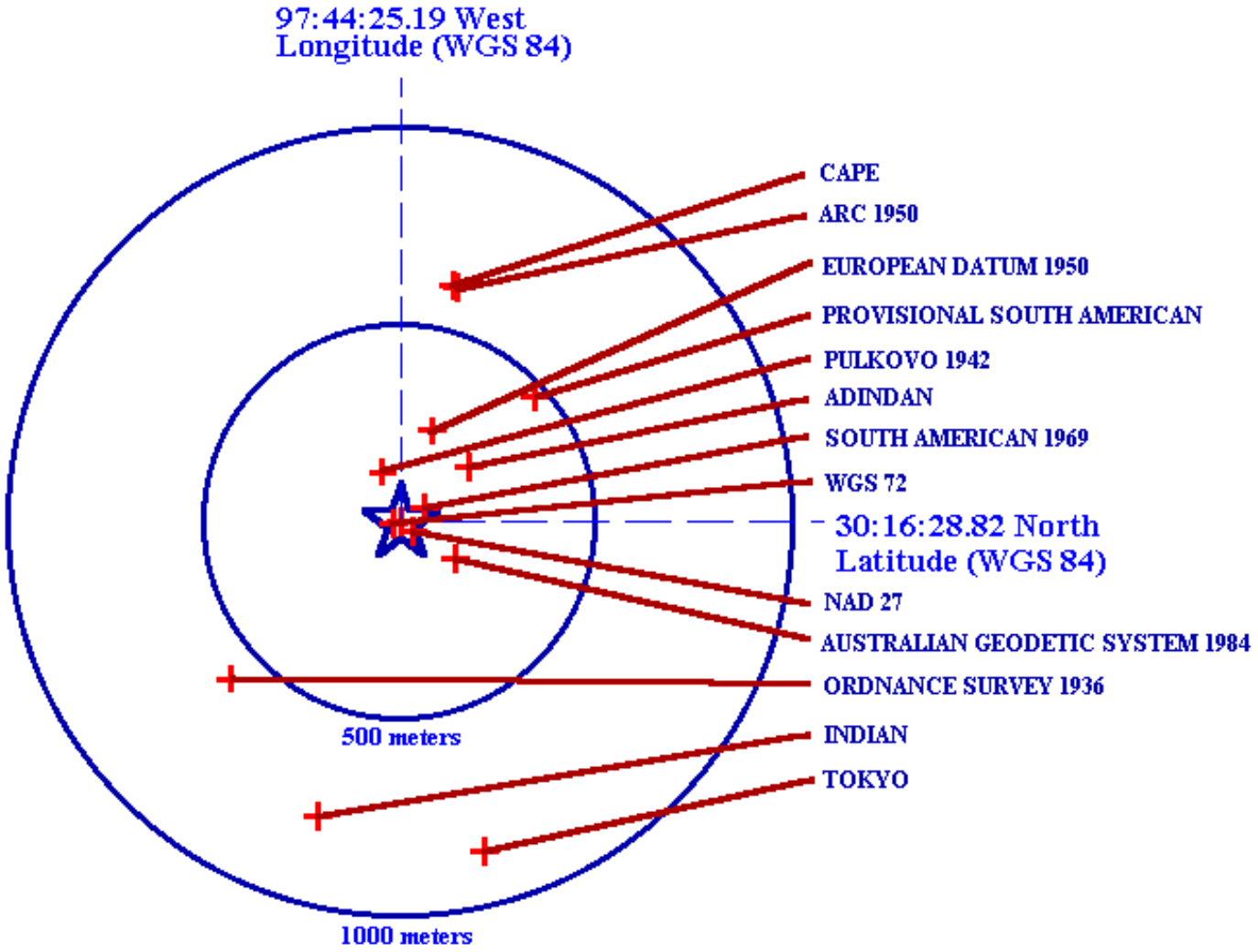
- A point is referenced by its *longitude* and *latitude*, which are angles measured from the Earth's center to a point on its surface.
  - Latitude is also the angle from Polaris to horizon (in N. hemisphere)
- Latitude and longitude are measured either in decimal degrees or in degrees, minutes and seconds.



GIS fundamentals: Geographic coordinate systems (GCS)

- The shape and size of a GCS's surface is defined by a sphere or a spheroid
- In fact, the Earth is neither a perfect sphere nor spheroid (the South Pole is closer to the equator than the North pole!)
- A *datum* defines the origin and orientation of latitude and longitude lines.
- Changing a GCS's spheroid or datum changes all values!
- The standard global system is called World Geodetic System 1984 (WGS1984)





Position Shifts from Datum Differences

Texas Capitol Dome Horizontal Benchmark

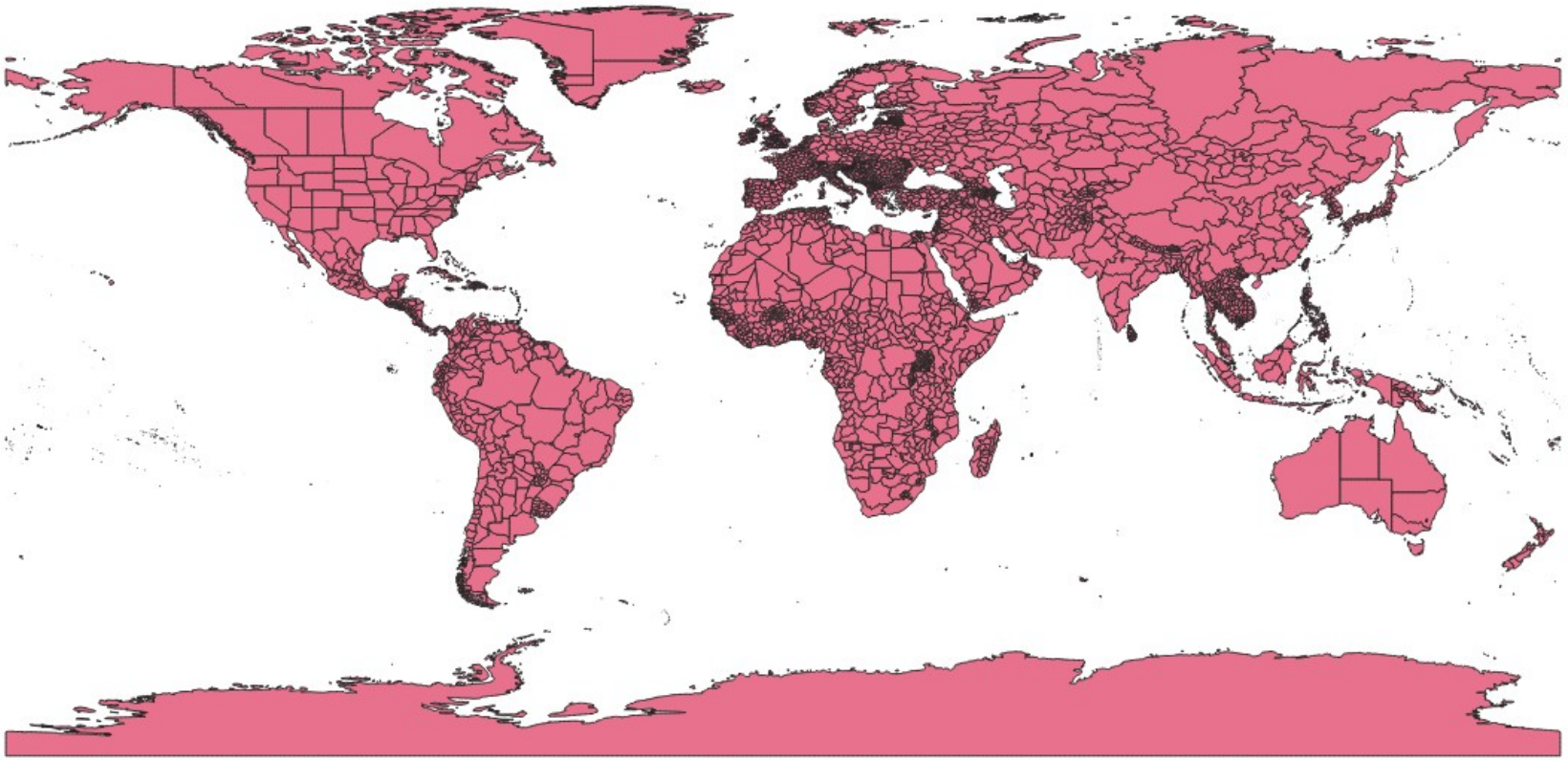
Peter H. Dana 9/1/94

GIS fundamentals: Geographic coordinate systems (GCS)

- WGS84
  - Default Datum for GPS units
  - Pretty good for global measurements
  - Small changes in landmass movements can be significant at global scale
    - Melting glaciers cause rising elevation
    - Tectonic movements
  - Local datums are better for local maps (e.g., North American Datum; NAD83) and reduce significance of landmass movements.

GIS fundamentals: Geographic coordinate systems (GCS)

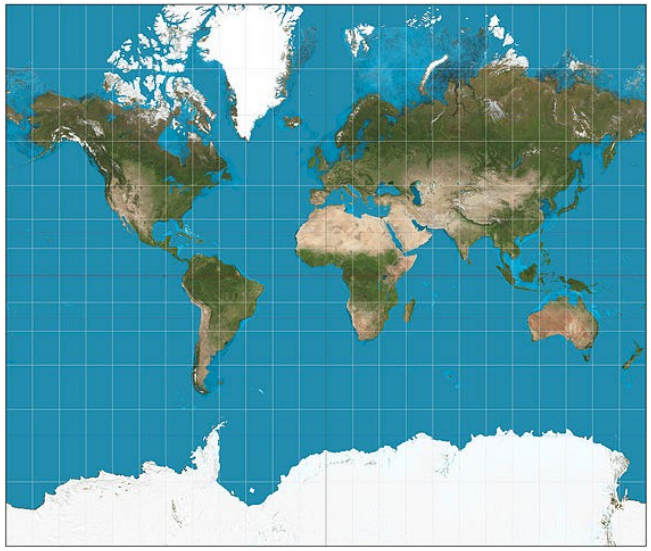
- WGS84 – Unprojected: Distortion near extents of map



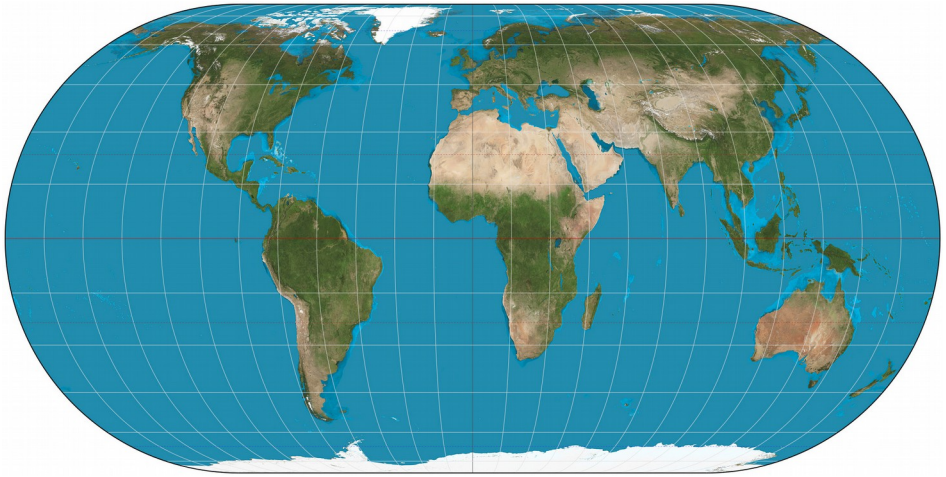
e

GIS fundamentals: Projected coordinate systems (PCS)

- Representing the earth's surface in two dimensions causes distortion in the shape, area, distance, or direction of the data.
  - Impossible to show spherical object in 2 dimensions without distortion
- A map projection uses mathematical formulas to relate spherical coordinates on the globe to flat, planar coordinates.
- Different projections are designed to minimize different distortions.

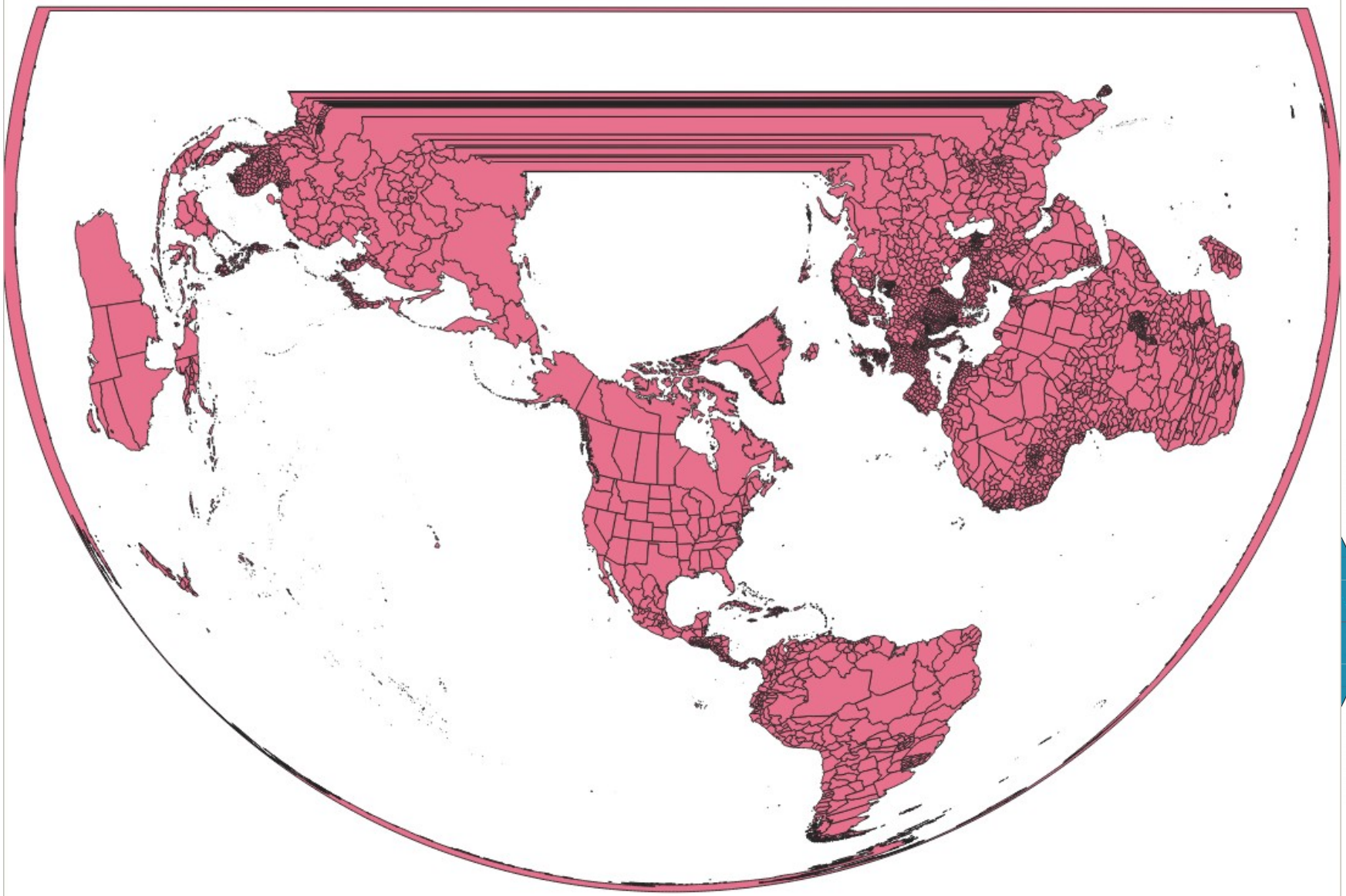


Mercator: maintains direction

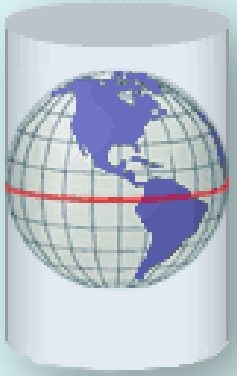


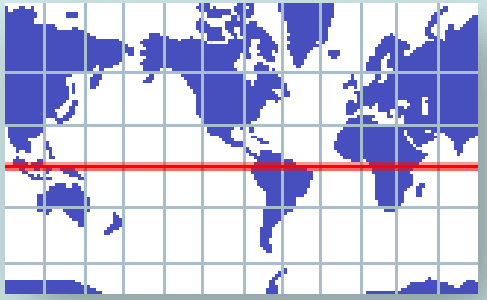
Eckert IV: maintains area

GIS fundamentals: Projected coordinate systems (PCS)



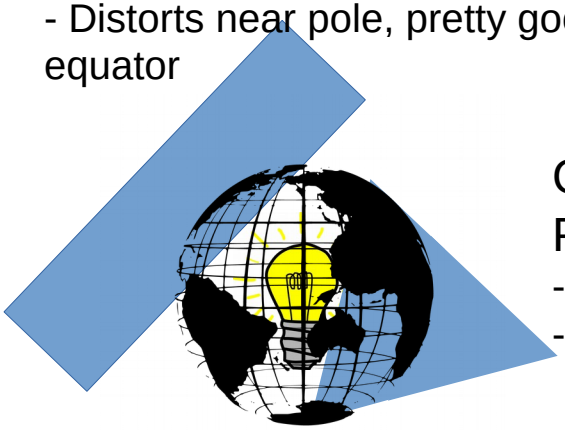
GIS fundamentals: Projected coordinate systems (PCS)





Mercator Projection

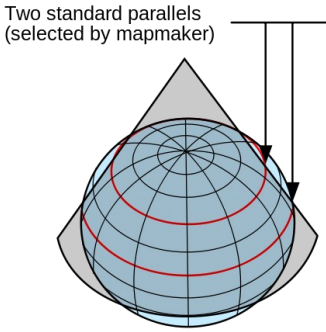
- Maintains direction
- Good for navigating
- Distorts near pole, pretty good near equator

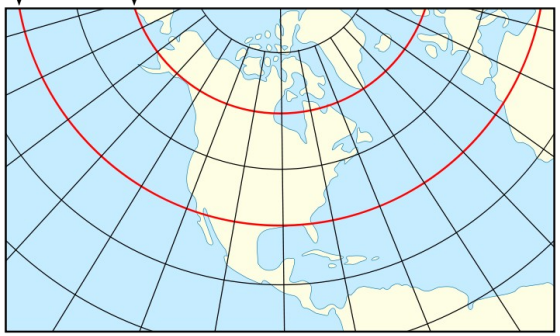


Gall-Peters Projection

- Maintains relative areas
- High distortion

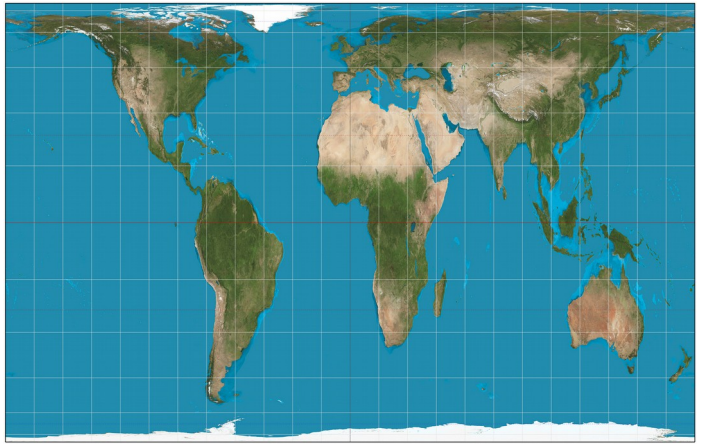
https://thetruesize.com/





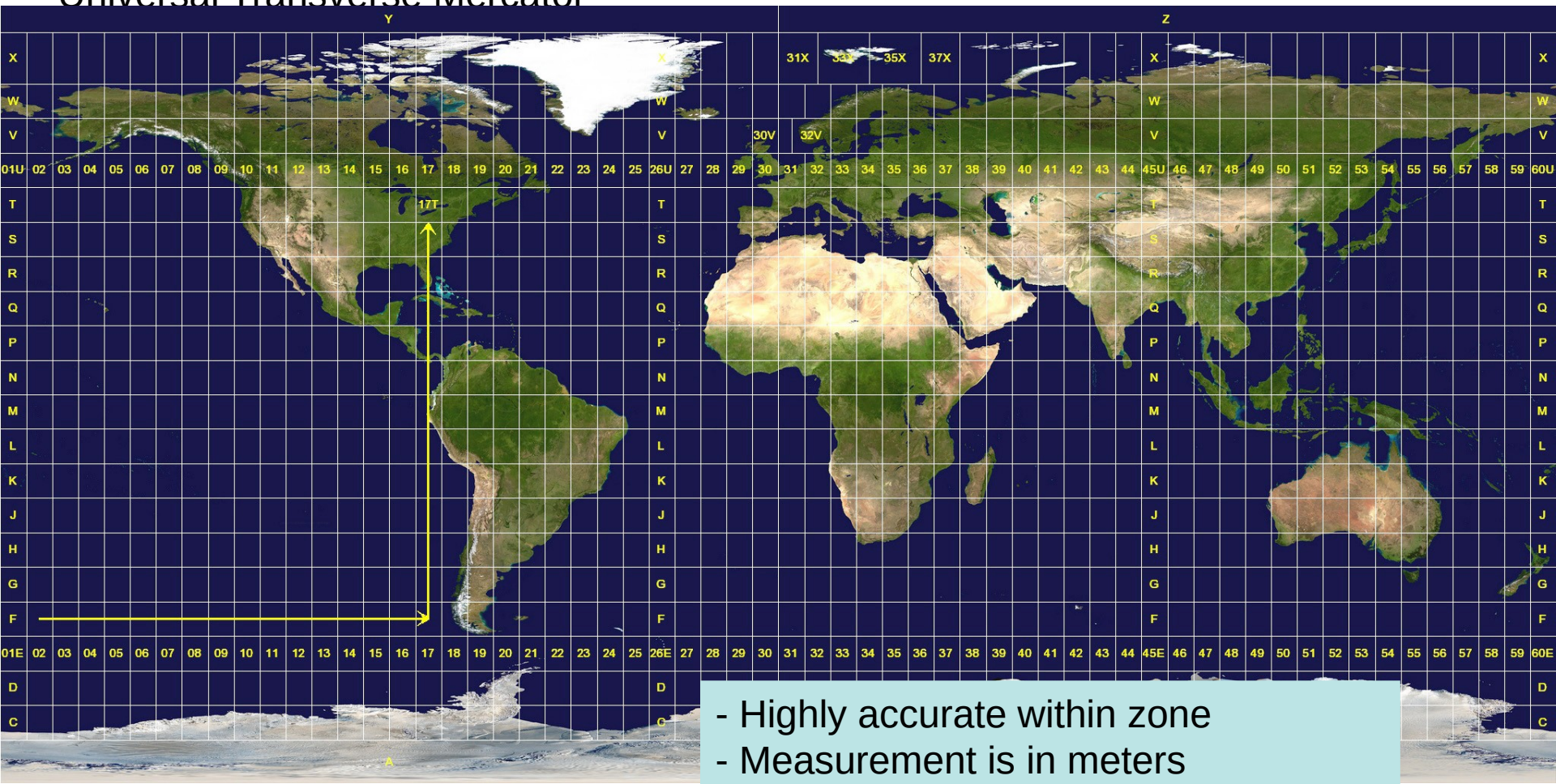
Conic Projection

- Maintains area near center
- Distortion near edges
- Good for east/west because distortion constant



GIS fundamentals: Projected coordinate systems (PCS)

Universal Transverse Mercator -



- Highly accurate within zone
- Measurement is in meters
  - intuitive
- Not great for large areas (across zones)

GIS fundamentals: Geographic and Projected coordinate systems

- Modern GIS make dealing with coordinate systems relatively straightforward; however:

IT IS ESSENTIAL TO KNOW THE COORDINATE SYSTEM OF ANY DATA WITH WHICH YOU ARE WORKING!

Outline

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

- What is a Geographic Information System?
- Data models – representations of geographic reality
- Location information: coordinate systems

- 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

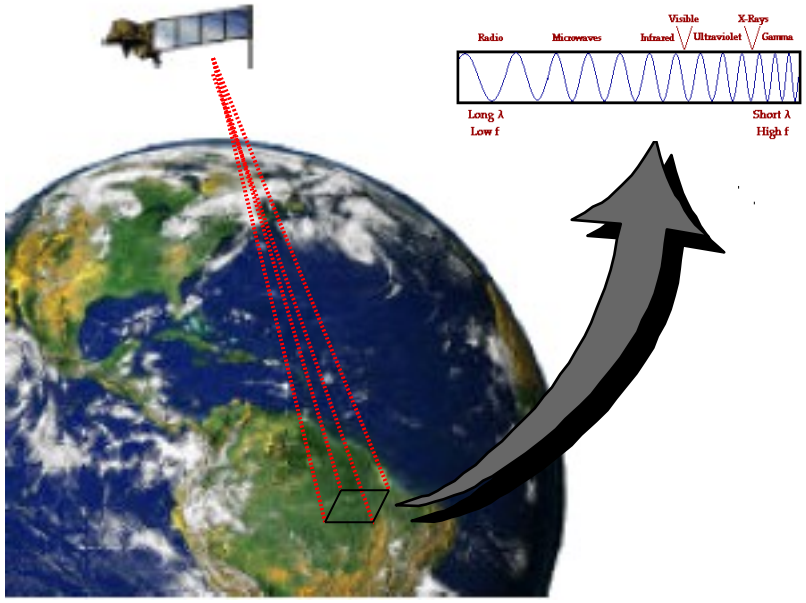
Outline

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

- What is a Geographic Information System?
- Data models – representations of geographic reality
- Location information: coordinate systems

- 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.*