Geographical Information Systems (GIS) for Data Science Introduction

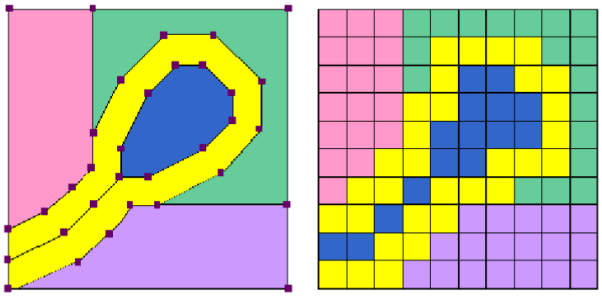
MIE223
Winter 2025

1 GIS

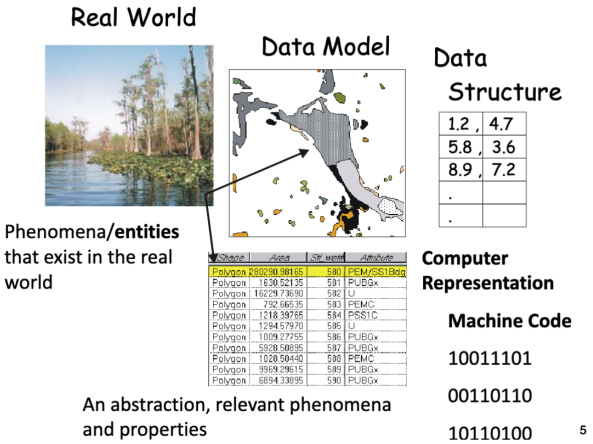
1.1 London Cholera: Epidemiological GIS in 1854

2 Geographical Information Systems (GIS)

2.1 Data Models



hard to go from raster to vector, but easy to go from vector to raster.



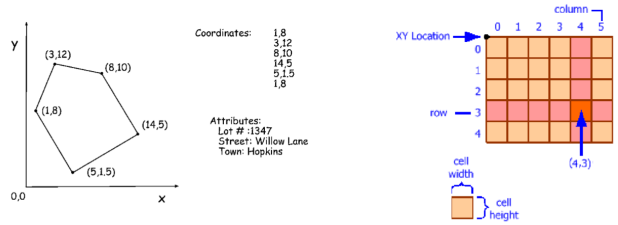
GIS and knowledge graphs are difficult to test, expect coding questions rather than conceptual.

2.2 Data Model

Objects in a spatial database.
The spatial data model provides a formal means of representing and manipulating spatially-referenced information.

2.3 Coordinates

Coordinates are used to define the location and extent of spatial objects.

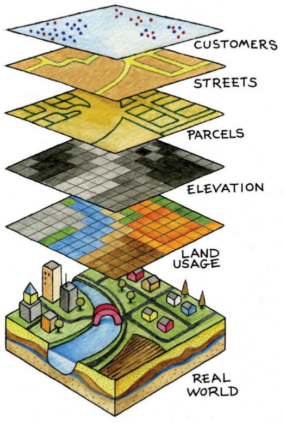


2.4 Attributes

Attributes describe the spatial object. Attribute data complements the coordinate data to define the spatial object.

2.5 Thematic Layers

- A logical separation of data according to theme.
- Each layer reflects a particular use or characteristic.
- Overlays.

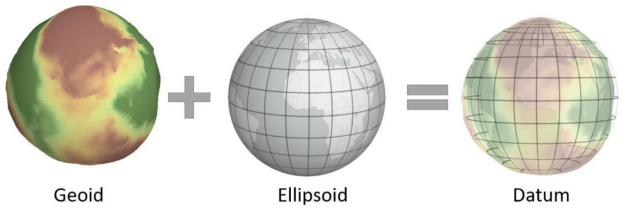


2.6 Coordinate Data

- Latitude & Longitude
  - Origin (intersection of the Equator and Greenwich meridian)
- Spherical Coordinates
  - Deg., min., sec. (DMS)
  - Decimal degrees (DD)

2.7 The Earth is not a Sphere

- It's close to an Oblate Spheroid. Why?
- But there are height variations
  - Not just mountains, also density anomalies

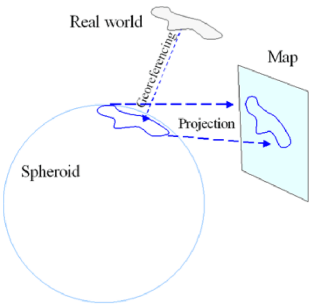


2.8 Ellipsoid

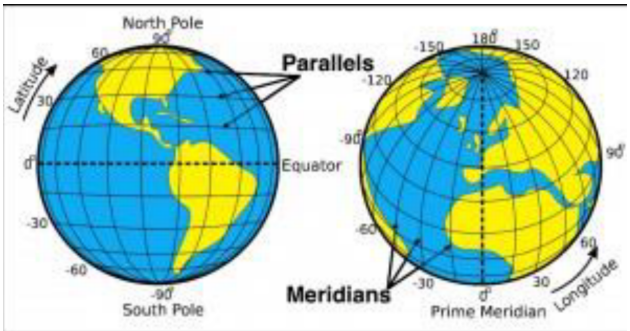
- The Earth is best approximated by an ellipsoid
- Ellipsoid is defined by:
  - Semi-major axis (a)
  - Semi-minor axis (b)
  - Flattening (f) = (a-b)/a

2.9 Geographic vs. Projected Coordinate System and Raster Coordinate Reference Systems

- Geographic Coordinate System (GCS)
  - Spheroidal approximation of Earth's surface
- Projected Coordinate System (PCS)
  - Projection of GCS onto rectangular (Euclidean) map coordinates for viewing
- Raster Coordinate Reference System
  - for warping 2D imaging to a CRS



2.10 Nomenclature of Geographic Coordinate System



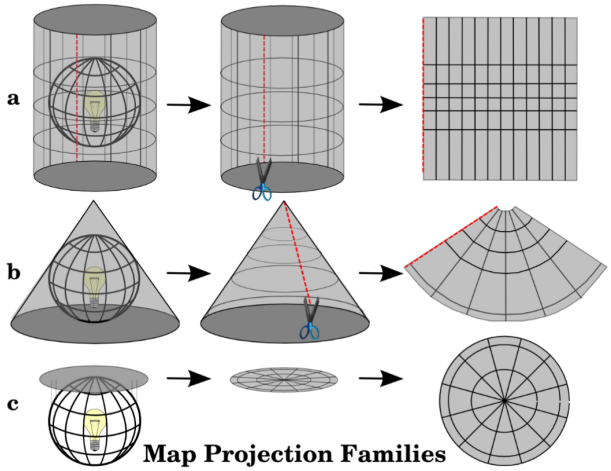
2.11 Properties of Projected Coordinate Systems (PCSs)

Three main types of projections:

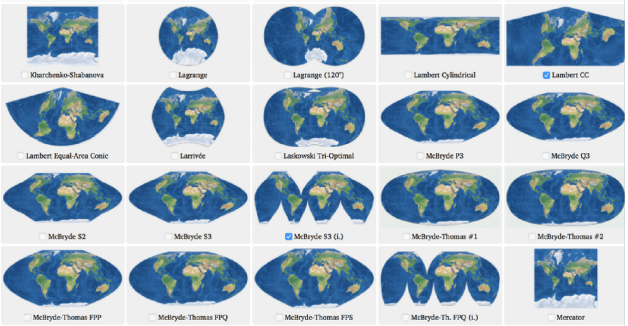
a Cylindrical

b Conical

c Planar



2.12 Projected Coordinate Systems

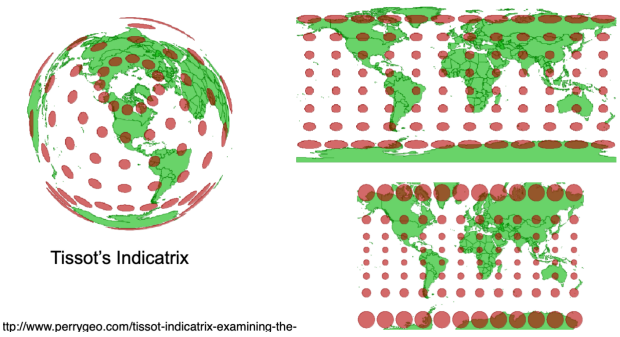


2.13 Properties of PCSs

- Why so many PCSs?

- Different PCSs maintain different spatial properties
  - NESW Cardinal Directions
  - Angles
  - Surface Area of Different Regions
  - Distances between points

2.14 Conversion from Spherical to Cartesian



Tissot's Indicatrix

ttp://www.perrygeo.com/tissot-indicatrix-examining-the-

2.15 Converting Arc to Surface Distance

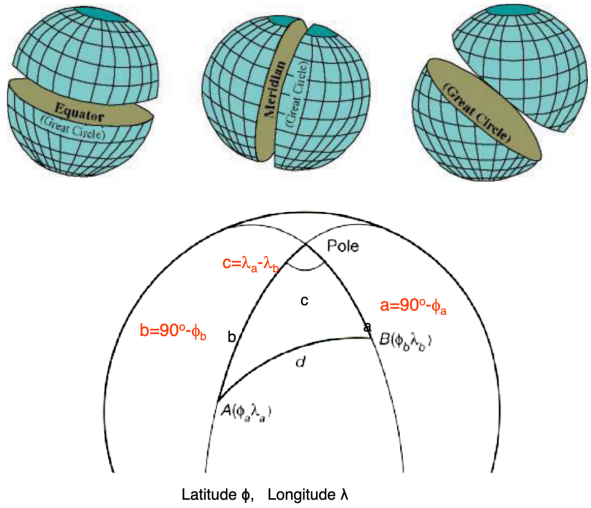
At the equator, one degree of longitude is about 111km
not tested on euclidean to spherical conversion

$$d = r \cdot \theta \tag{1}$$

Where d is the ground distance, r is the radius of the sphere and the angle is specified in radians, 360 degrees = 2 π radians.

2.16 Great Circle

A great circle is defined as any line resulting from a plane passing through the center of the globe



$$d = r\cos^{-1}[\sin\phi_a \sin\phi_b + \cos\phi_a \cos\phi_b \cos(\lambda_a - \lambda_b)]$$

2.17 Conversion from Geographic to Cartesian Coordinates

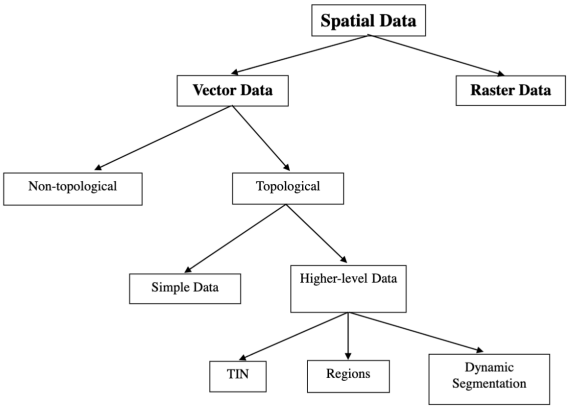
$$x = r \cdot \cos(\phi) \cdot \cos(\lambda) \tag{2}$$

$$y = r \cdot \cos(\phi) \cdot \sin(\lambda) \tag{3}$$

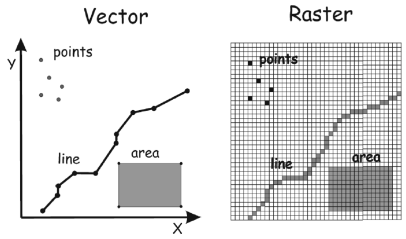
$$z = r \cdot \sin(\phi) \tag{4}$$

This formula requires latitude and longitude to be in radians.

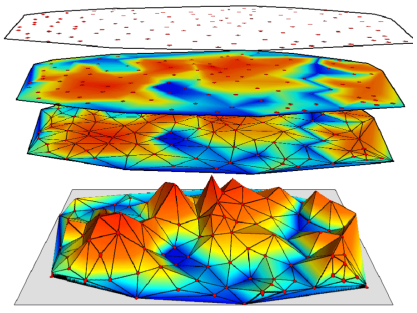
2.18 Common GIS Data Models



2.19 Two Most Common Spatial Data Models

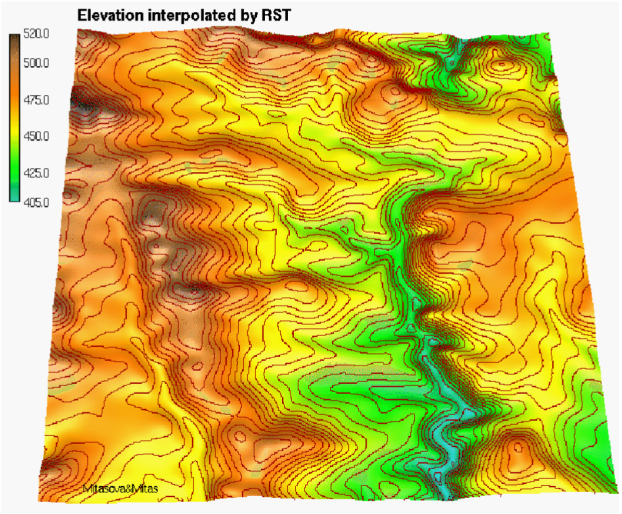


2.20 Triangulated Irregular Network (TIN)



2.21 Regions

2.22 Contour Lines over a Raster

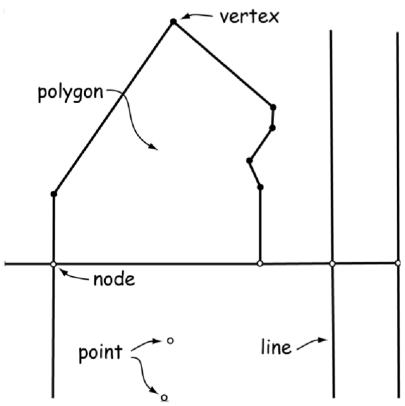


2.23 Vector & Raster

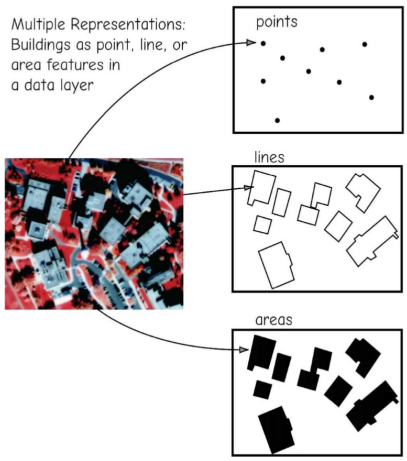
- Vector is better at representing discrete features. (objects, roads, buildings)
- Raster is better at representing continuous features (imaging, elevation, temperature)
- A project may contain both vector and raster layers.
- Spatial operations can only be performed on one type of layer.
- The best data model for a given layer depends upon the operations, the experience and the views of the user.
- No decision is final, as one can be converted to the other.

2.24 Vector Terminology

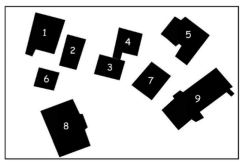
- The terms polygon / area / shape will be used interchangeably.
- A polyline is one or more connected lines.
- A polygon is a polyline that starts at ends at same point.



2.25 Multiple Representations

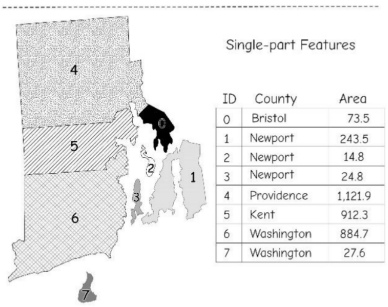


2.26 Vector Model



ID	Building Name	Floors	Roof Type
1	Hodson Hall	6.0	flat, sealed tar
2	Borlaug Hall	5.5	pitched 9/12, tile
3	Guilford Technology Bldg.	4.0	flat, gasket
4	Shop Annex	2.5	flat, sealed tar
5	Animal Sciences Bldg.	1.0	pitched 12/12, tile
6	Administration Bldg.	14.0	pitched 6/12, metal
7	Climate Sciences Center	6.0	flat, sealed tar
8	Grantham Tower	1.0	pitched, 9/12, tile
9	Biological Sciences Bldg.	9.0	pitched 12/12, tile

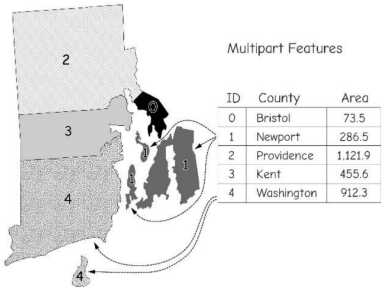
2.27 Single Part Features



ID	County	Area
0	Bristol	73.5
1	Newport	243.5
2	Newport	14.8
3	Newport	24.8
4	Providence	1,121.9
5	Kent	912.3
6	Washington	884.7
7	Washington	27.6

1 to 1 Relationship

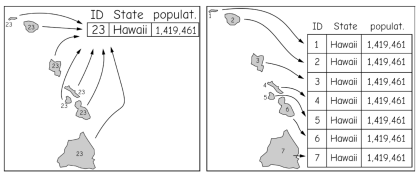
2.28 Multipart Features



ID	County	Area
0	Bristol	73.5
1	Newport	286.5
2	Providence	1,121.9
3	Kent	455.6
4	Washington	912.3

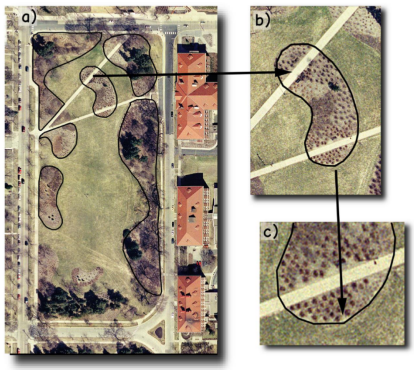
Many to 1 Relationship

2.29 Multipart to Single-Part Problem



ID	State	popular.
1	Hawaii	1,419,461
2	Hawaii	1,419,461
3	Hawaii	1,419,461
4	Hawaii	1,419,461
5	Hawaii	1,419,461
6	Hawaii	1,419,461
7	Hawaii	1,419,461

2.30 Polygon Inclusion



2.31 Polygon Inclusions

- Areas in polygons that are part of the polygon, but different from the rest of the polygon: e.g. Islands in a lake
- Solutions:
  - Create separate polygons for each inclusion.
  - Create an attribute column for coding inclusions.

2.32 Key Ideas to Understand

- GIS data consists of layers
- Layers can either be Raster or Vector
- Each layer has a coordinate reference system (CRS)
  - Geographic Coordinate System (GCS)
    - Spherical coordinates on approximation of Earth's surface
  - Projected Coordinate System (PCS)
    - Euclidean coordinates for projection onto a rectangular map for viewing
    - Different PCS's maintain different properties

exam: likely 1 high level question maybe 2