SPATIAL DATA MODELS

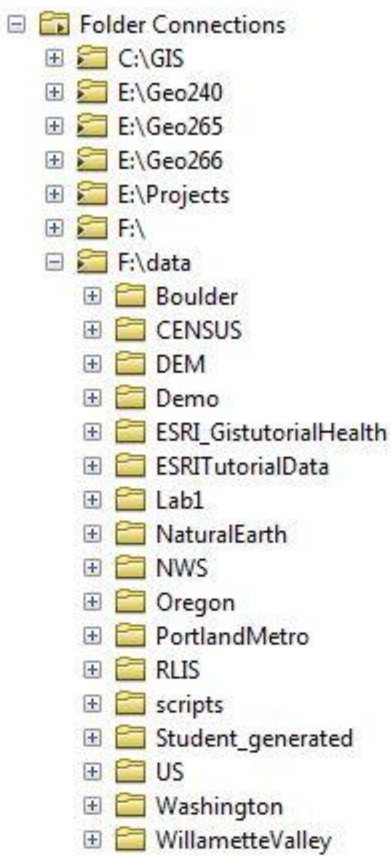
Introduction to GIS| Winter 2015

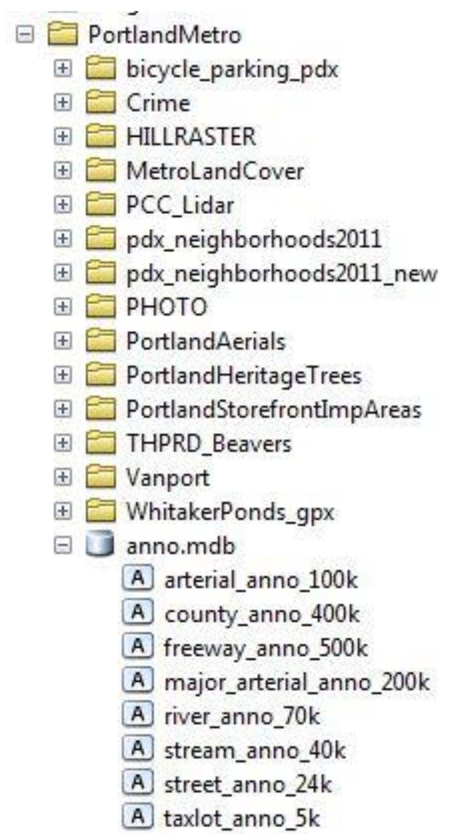
GIS Data Organization

The basics

- Data can be organized in a variety of ways
  - Spatial location, content (attributes), frequency of use
  - Come up with a system and stick to it
- GIS data becomes confusing quickly
- GIS files have multiple parts
- When working in GIS you create lots of new files

Data by folder or geodatabase





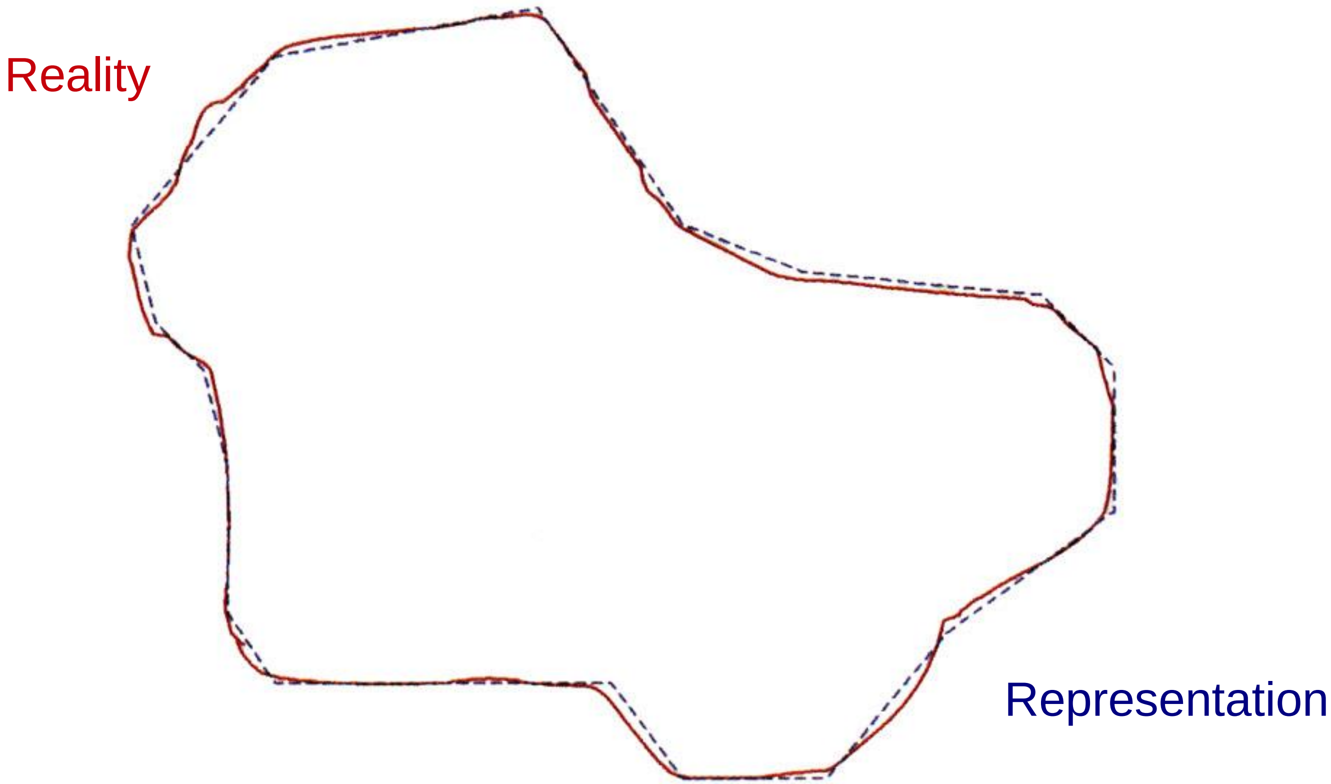
Essential organization tips

- Create a folder for every lab
- Create a ‘default’ geodatabase
- Title everything appropriately
  - Not: export_output
- Use successional number to *version* your data

GIS Data models

What is a data model?

- Digital *representation* of some *limited* part of the Earth's surface
- *Simplified* relationship between the digital model and the feature being represented

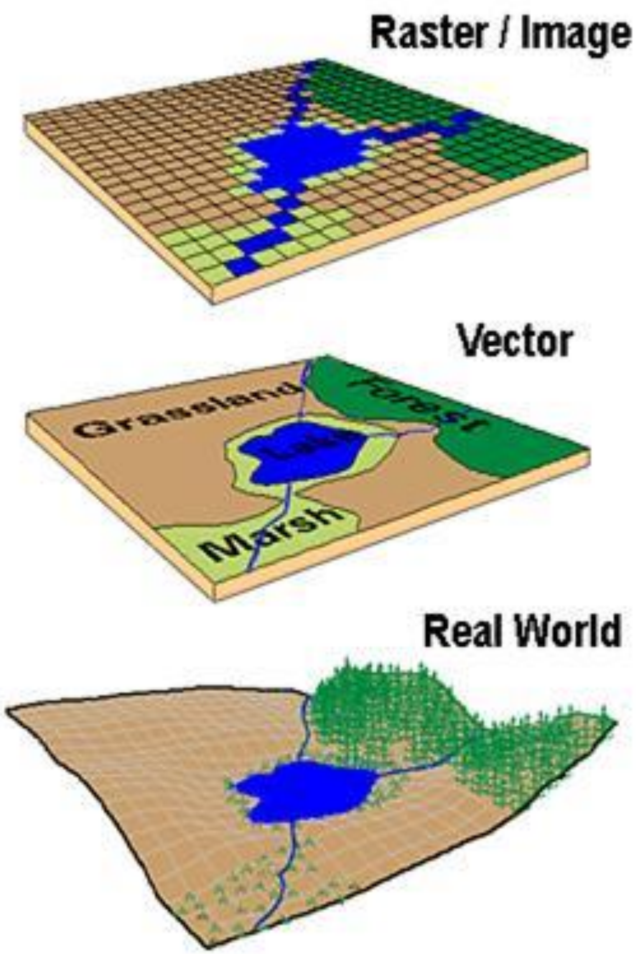

Reality
Representation

Why do we need them?

- The world is infinitely complex, but:
  - A computer is finite
  - A map is finite

GIS data models

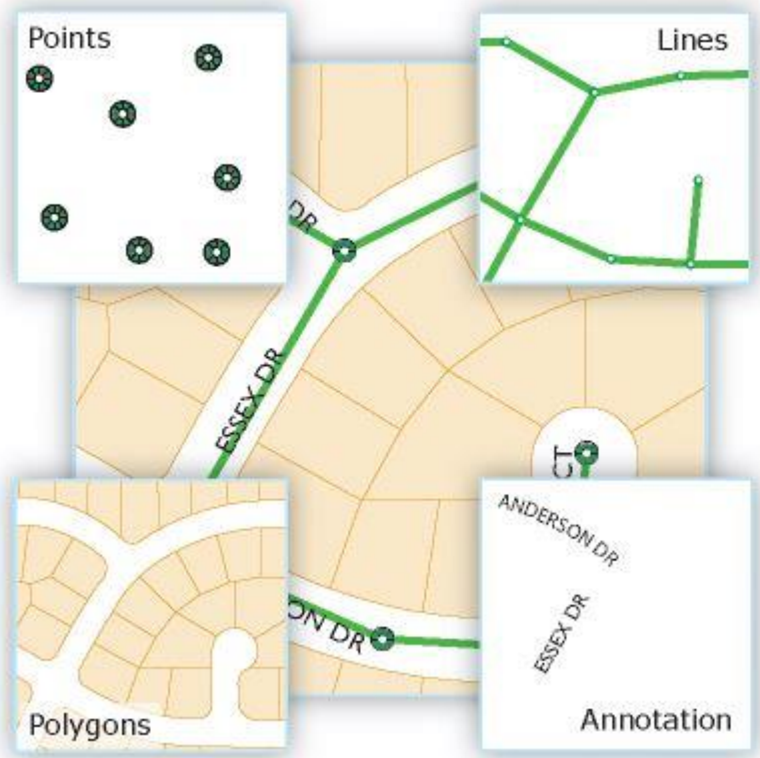
- Vector
  - Points
  - Lines
  - Polygons
- Raster
  - Cell based
- Network
  - Flows of data
- TIN
  - Points connected by lines to create traiangles



Vector data

Vector data models

- Objects are defined by:
  - Points
  - Lines
  - Polygons
- Point, line & polygon datasets are referred to as ‘feature classes’

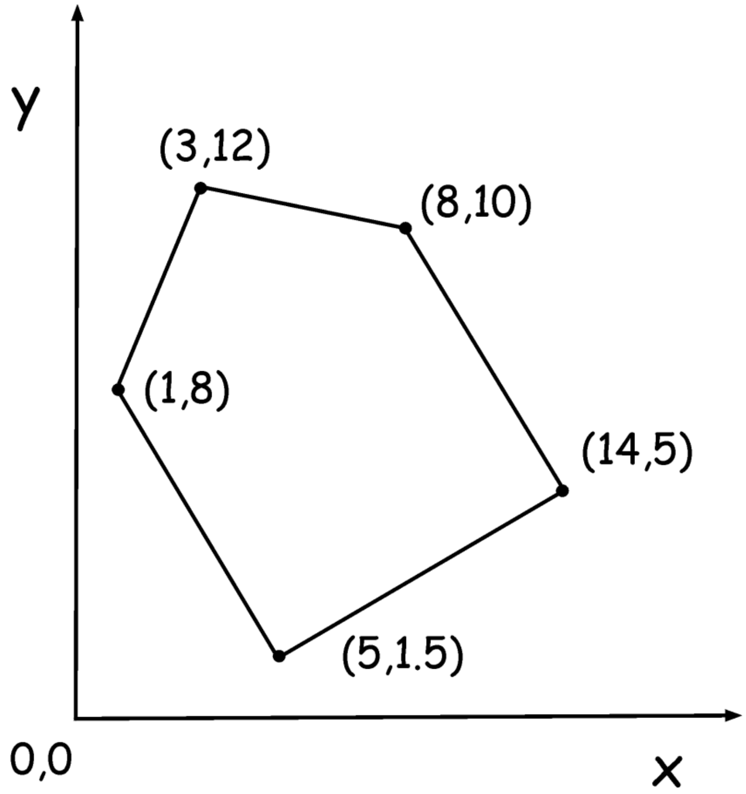


Vector data terminology

- **Points.** features defined by a single x/y coordinate
- **Vertices.** points defining a line or polygon feature
- **Nodes.** vertices representing the start and end of a line feature
- **Polyline.** feature comprised of a series of straight line segments or mathematical curves
- **Polygon.** feature represented by a ‘closed’ series of polylines (areas)

Feature Class

- Objects are defined by *points*, *lines* connecting the points, or *polygons* bounded by lines

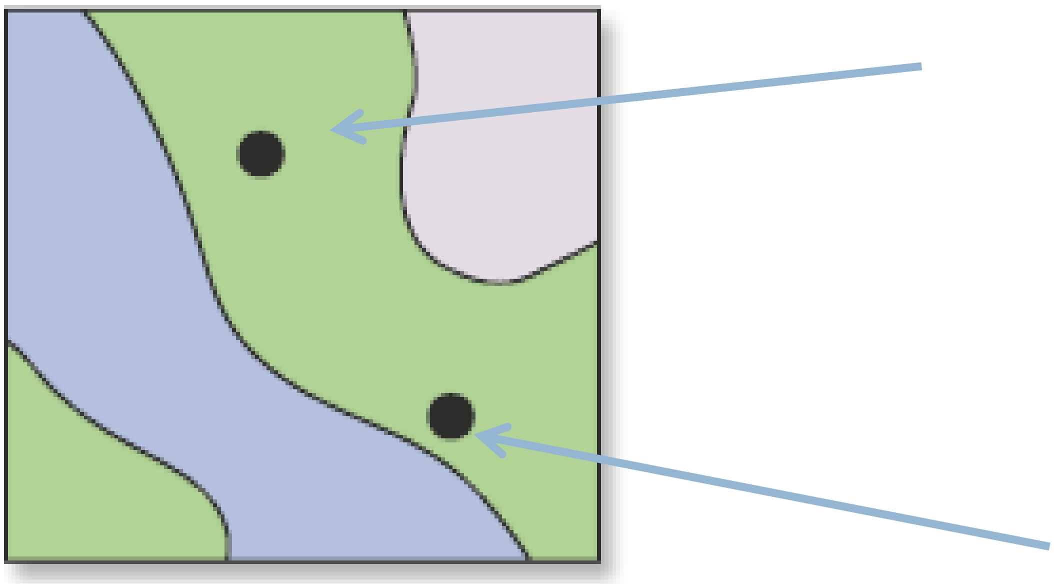


Coordinates: 1,8
3,12
8,10
14,5
5,1.5
1,8

Attributes:
Lot # :1347
Street: Willow Lane
Town: Hopkins

Feature Class

- Each point has x/y values (can also be a z component)



Feature Class

- Each object, or "*Feature*", is a single *row* in the GIS database

Name	Position	Altitu	Lat H	Lat Deg	Lat Min	Lat Sec
1	N30 17.824 W97 44.081	552	1	30	17	824
2	N30 17.824 W97 44.083	568	1	30	17	824
3	N30 17.824 W97 44.082	577	1	30	17	824
4	N30 17.823 W97 44.082	580	1	30	17	823
5	N30 17.824 W97 44.081	574	1	30	17	824

Feature class

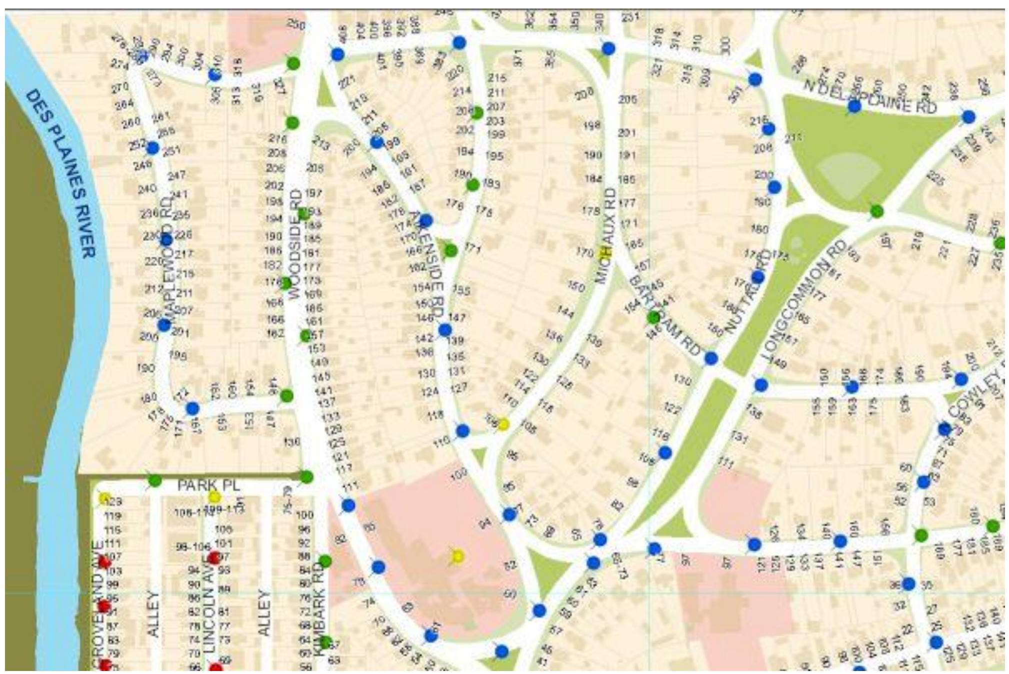
- "*Attributes*" are information associated with features; each attribute is a *column* in the database

Name	Position	Altitu	Lat H	Lat Deg	Lat Min	Lat Sec	Lat
1	N30 17.824 W97 44.081	552	1	30	17	824	30.51
2	N30 17.824 W97 44.083	568	1	30	17	824	30.51
3	N30 17.824 W97 44.082	577	1	30	17	824	30.51
4	N30 17.823 W97 44.082	580	1	30	17	823	30.51
5	N30 17.824 W97 44.081	574	1	30	17	824	30.51

Feature Class

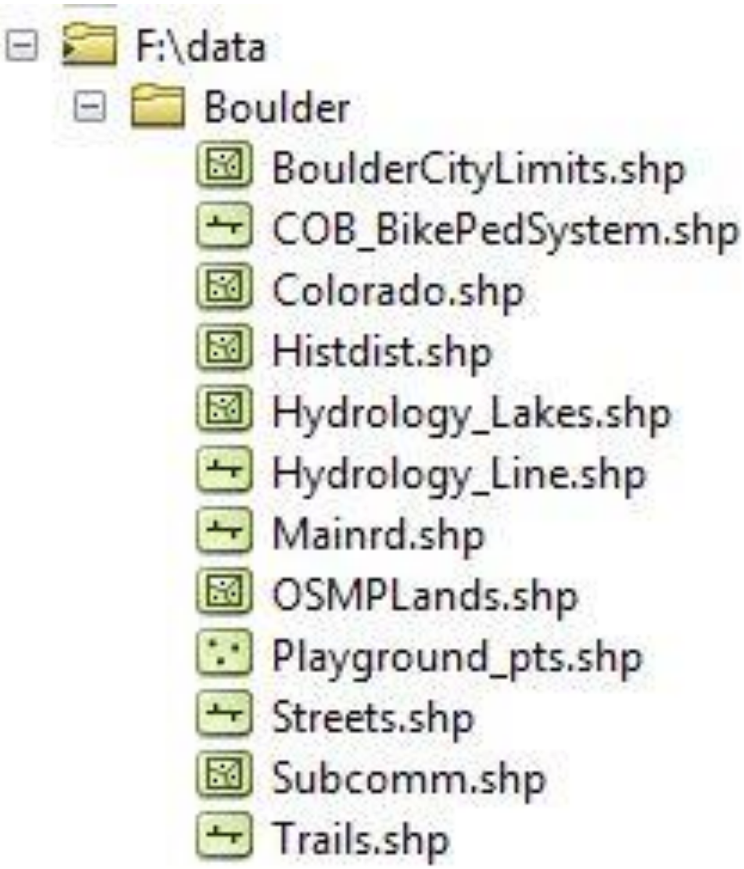
- Good way to represent *Discrete* objects, e.g. buildings

(“discrete” = constituting a separate, distinct entity)



Shapefiles

- *Simple* feature class
- Most common GIS vector data type

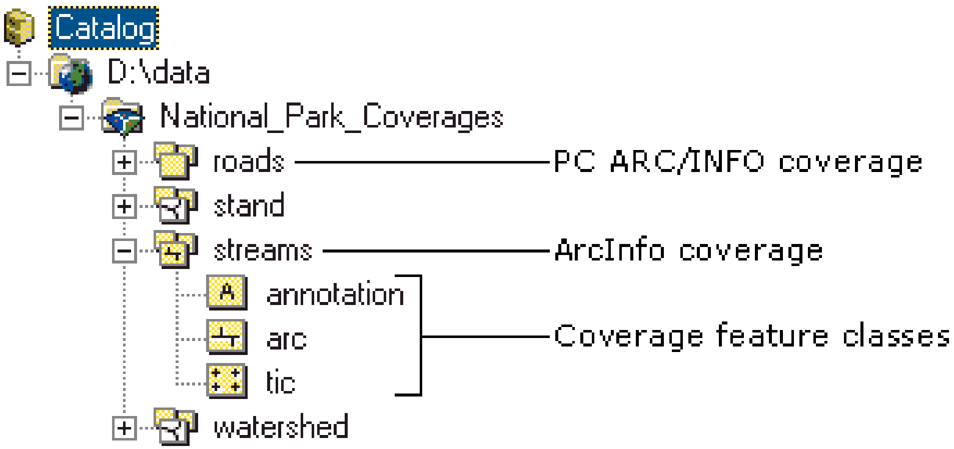


Shapefile

- **.shp – feature geometry (shape)**
- **.dbf – attribute table**
- **.shx – index file**
- .prj – map projection info
- .sbx and .sbn – spatial index
- .ain and aih – attribute index

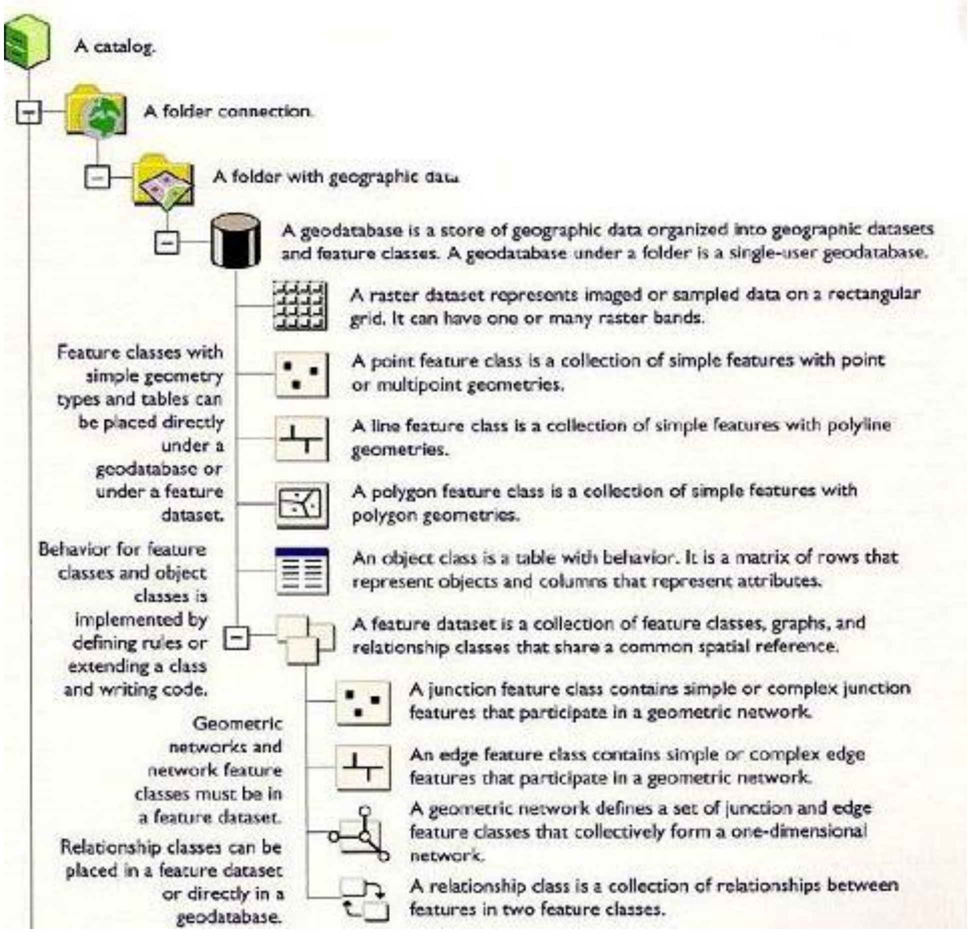
Coverage

- *Topological* vector features
- Stored in .e00 file format (aka interchange files or Export files)
- Older data model not used much anymore



Geodatabase

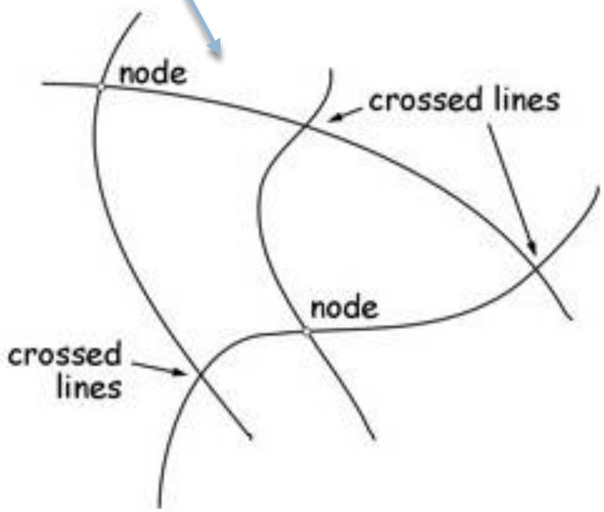
- Collection of various geographic datasets
- Primary data format used for editing and data management
- Build in additionally functionality
  - Behaviors
  - Rules for spatial integrity
  - Tools for working with spatial relationships within and between data sets

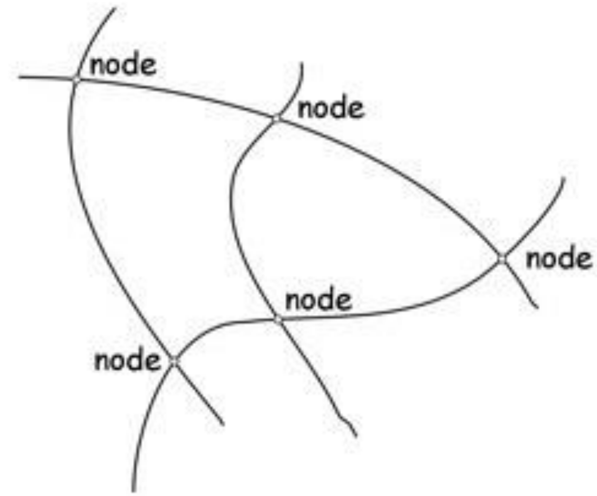


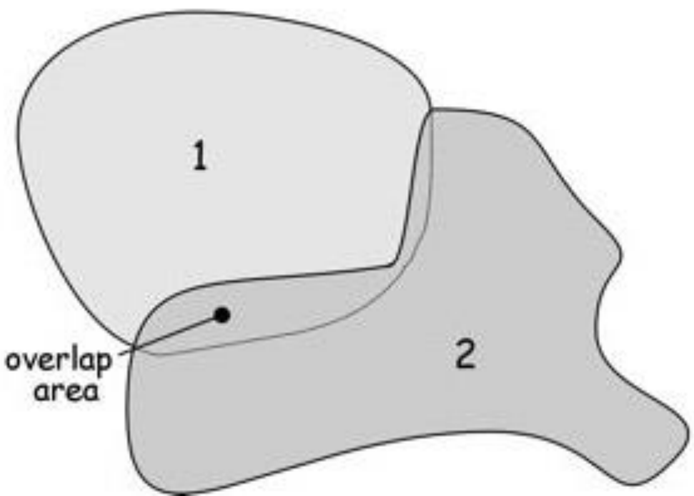
Topological data

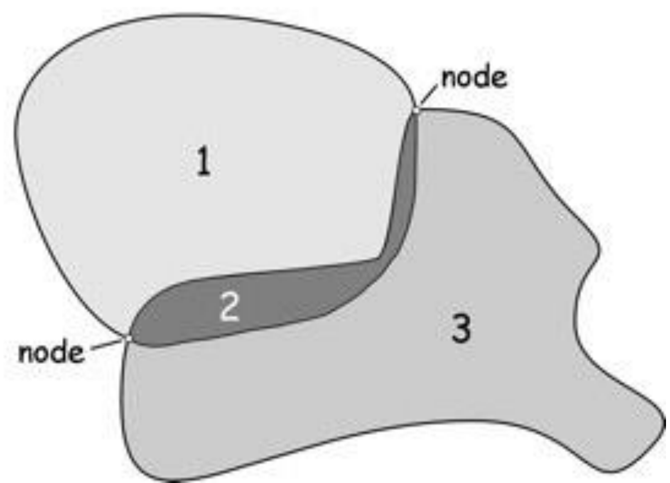
- Has explicit spatial relationships
  - **Connectivity**: Lines that cross or touch are connected to one another
  - **Planar**: Lines and polygons cannot overlap
  - **Contiguity**: Polygons that share a border are contiguous and adjacent – they don’t overlap and are next to each other

Simple v. Topological data









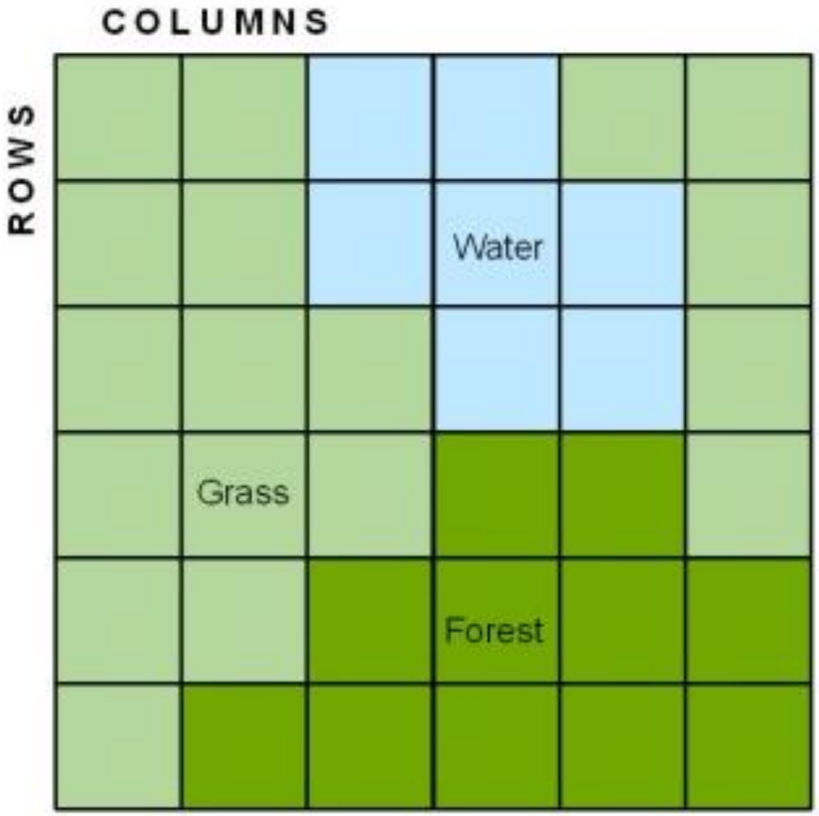
Vector Data Model

- Advantages
  - Vector is ‘correcter’ (more accurate)
  - Intuitive
  - Analytical capabilities
  - Easy to edit
  - Better cartography
  - Includes topology
- Disadvantages
  - Complex features / data structure
  - Time consuming to design (topological)
  - Software dependent

Raster data

Raster data model

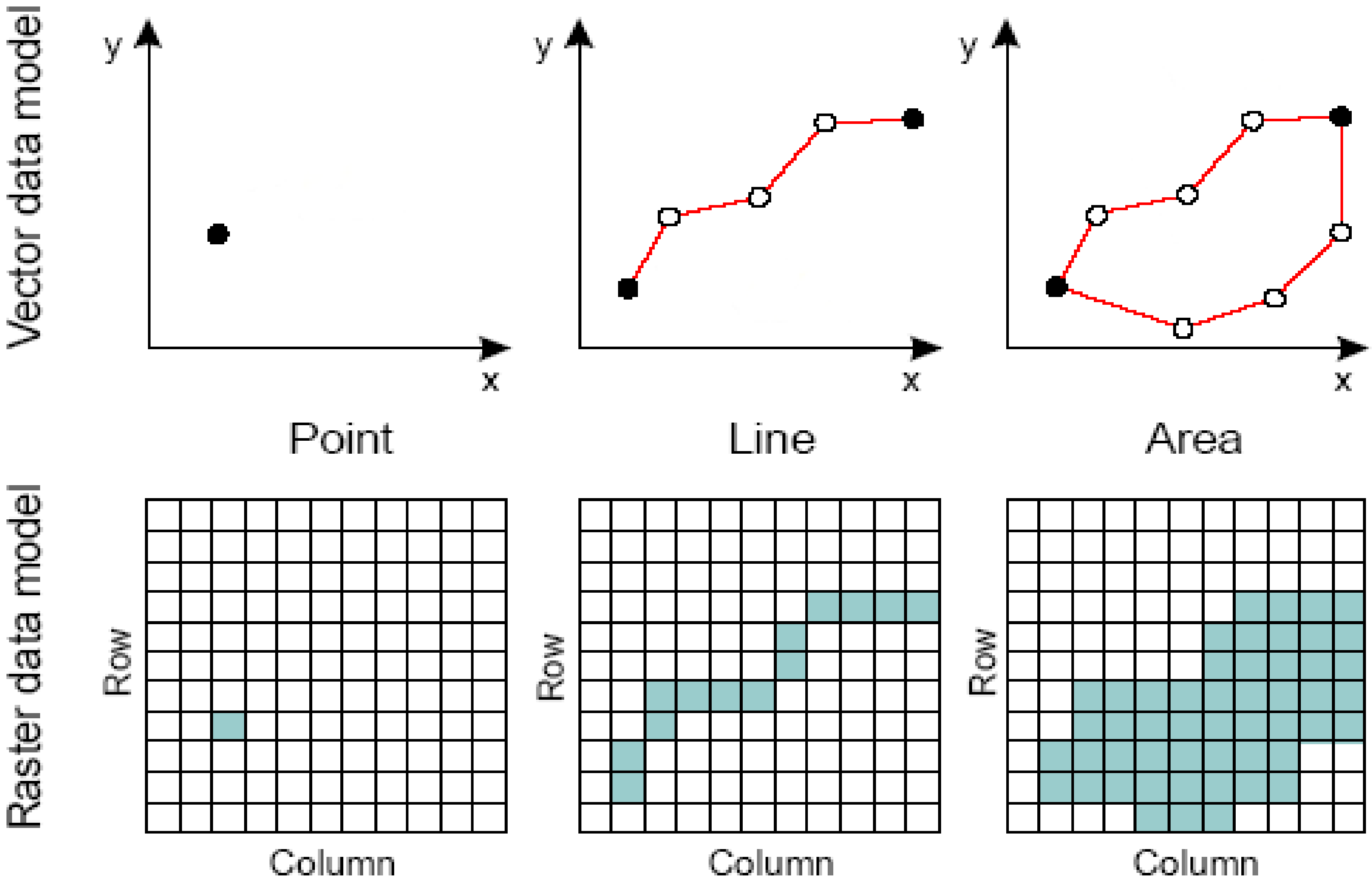
- Contains array of cells (pixels) that are organized in columns and rows



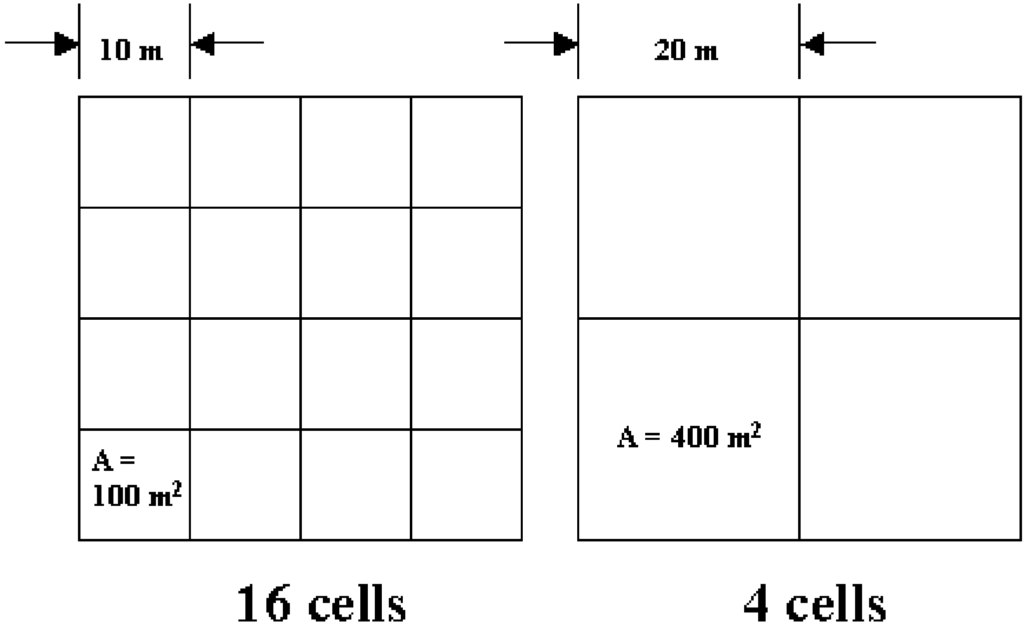
Raster data

- *Attributes* are associated with each cell
- Used mostly for *continuous* data (i.e. rainfall or elevation)
- Can also be used to represent discrete data

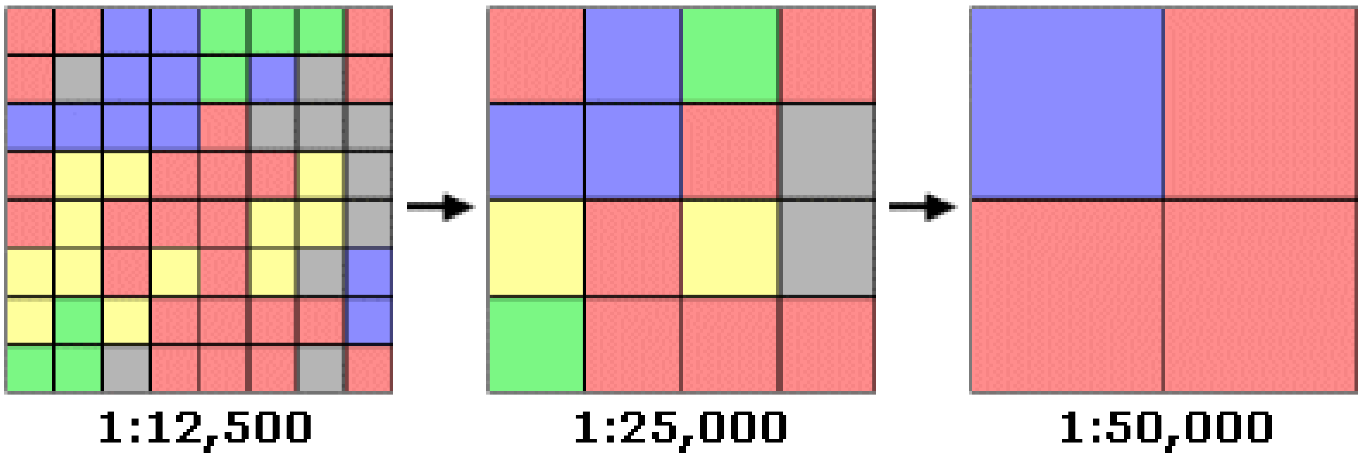
Raster v. Vector



Raster data resolution



Raster data resolution

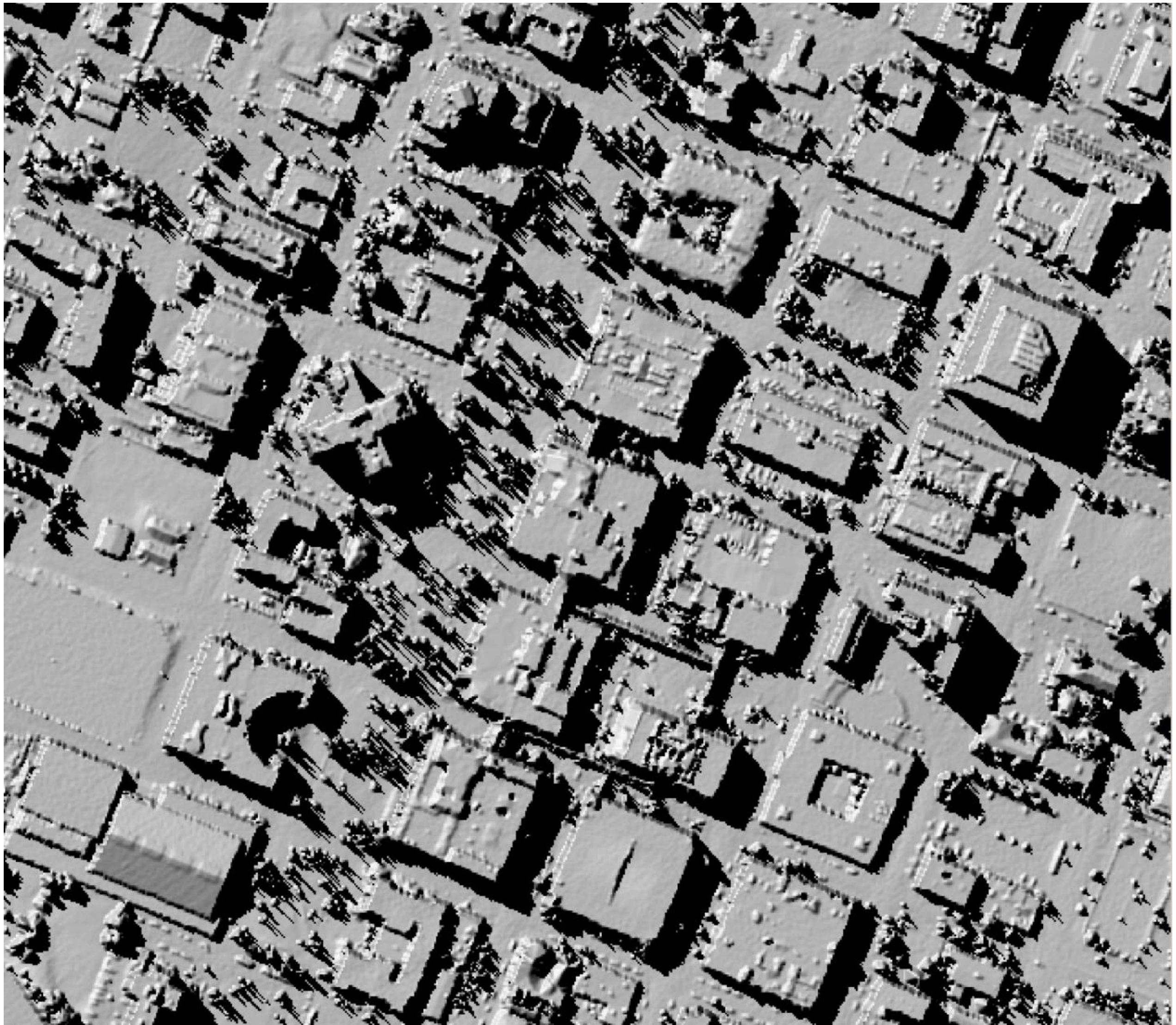


Examples of raster data types

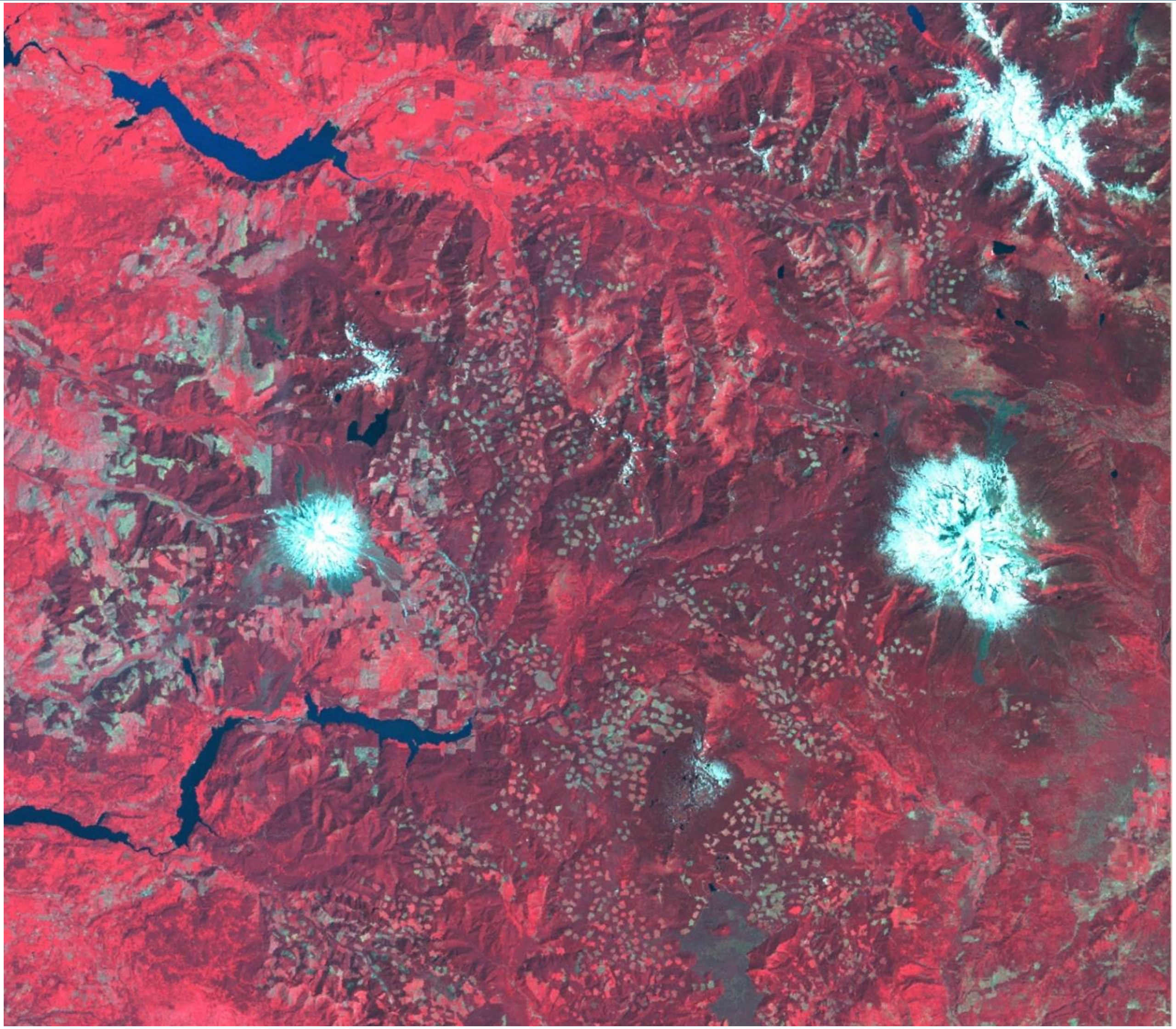
- ESRI Grid
- ESRI geodatabase
- ERDAS Imagine (.img)
- MrSID (.sid)
- Other image formats (.tif, .jpg)

Aerial photograph

Hillshade - LiDAR



Satellite image



Raster data model

- Advantages
  - Simple data structure
  - Large # of common, software independent formats
  - Faster processing time (generally)
- Disadvantages
  - Large datasets
  - Cartographic limitations
  - Limited analytical capability

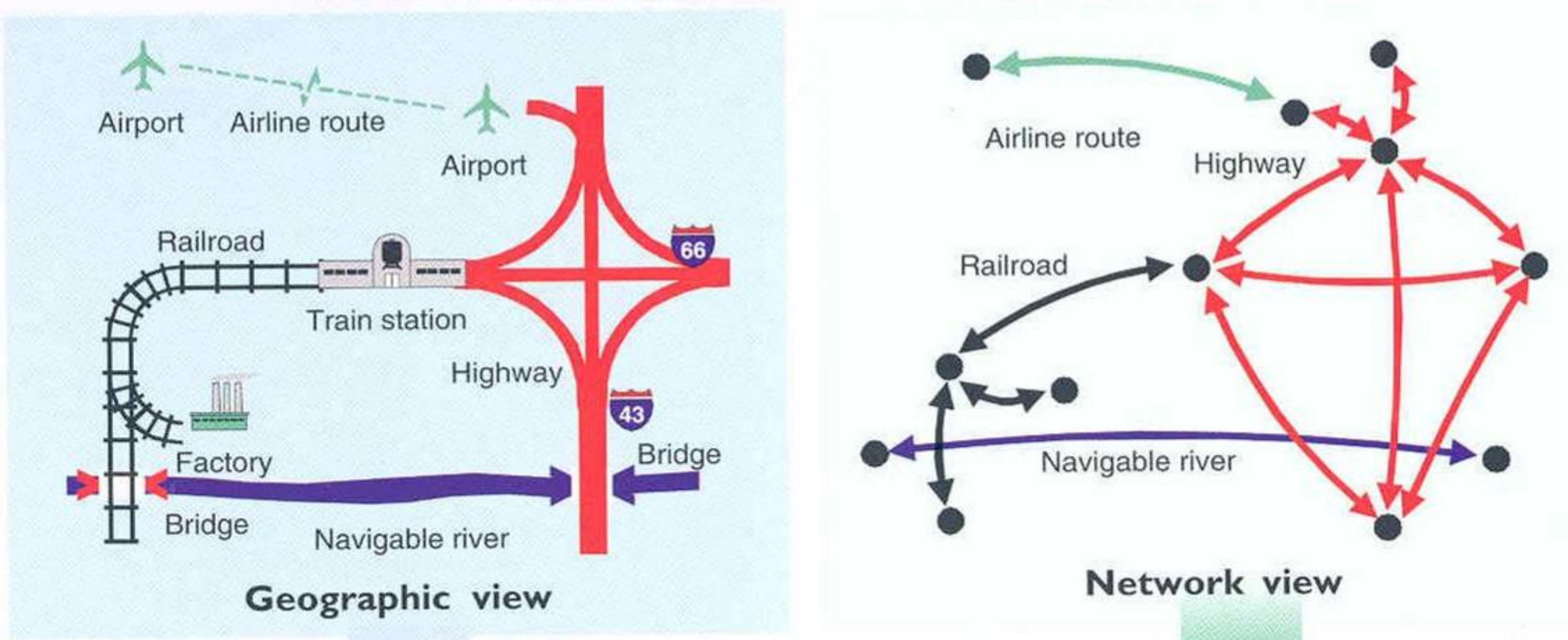
Network data model

Network data model

- A system of points and lines defining how flows move through a network (i.e. streams, streets)
- **Linear Referencing**: features are stored as distances along a network

Two views of a network

You can view a network as a collection of geographic objects such as rails, roads, stations, and bridges and also as a pure network of edges and junctions.



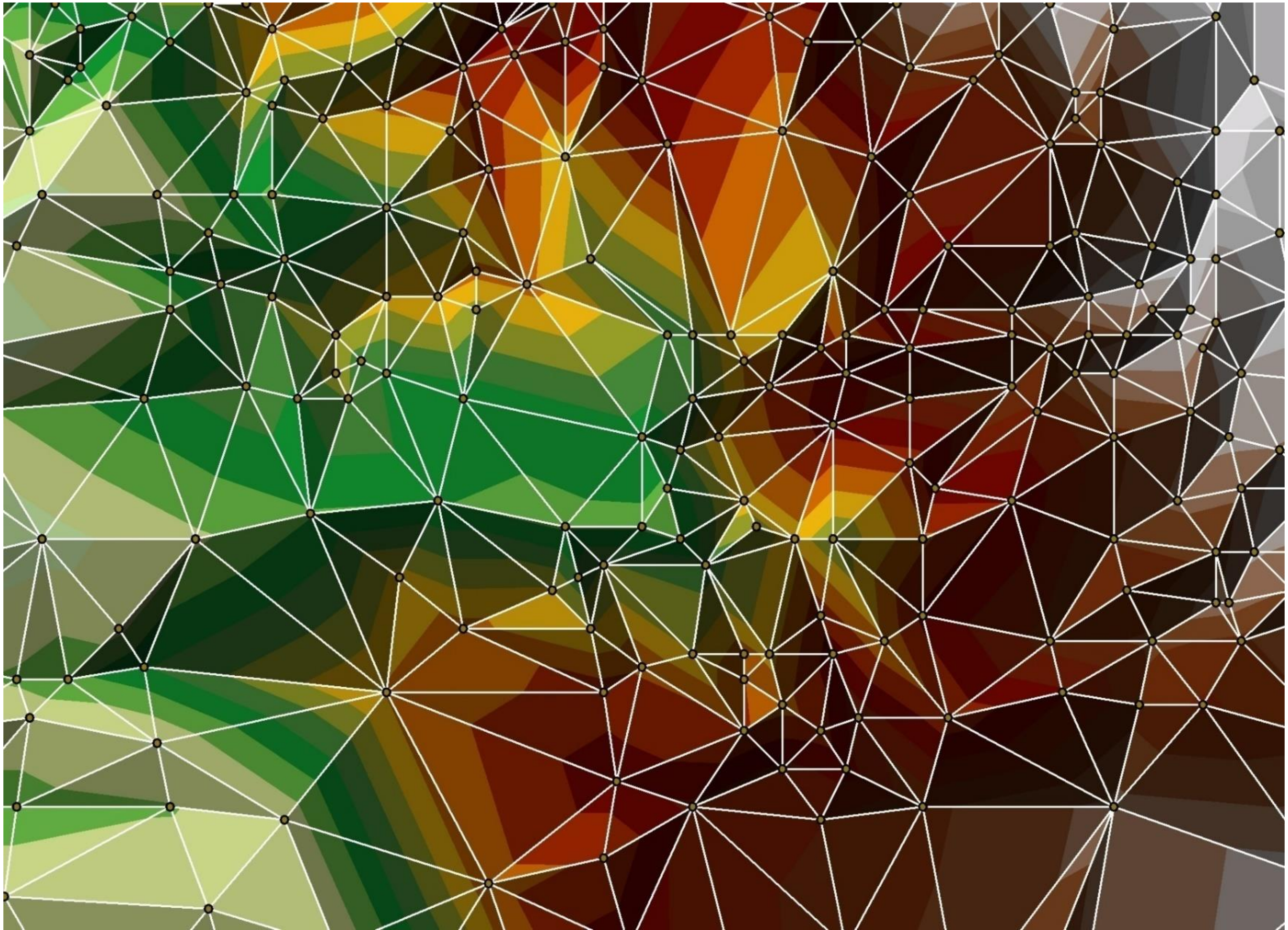
TIN data

Triangulated Irregular Network (TIN)

TIN data model

- Created from a series of points with *x/y/z coordinates* (3D)
- Points are connected by lines to form triangles

TIN data model



TIN data model

- Used to represent *continuous* surfaces (i.e. rainfall or elevation)
- Best used when the *density* of data points varies (irregular spacing of data points)