Lecture 2: Spatial Models and Data Types

Spatial Data Mining

Instructor: Yiqun Xie

Attribution: Slides modified based on lecture notes from Dr. Xun Zhou (UI) and Dr. Shashi Shekhar (UMN)

Goals

- Learn two different paradigms to model spatial phenomenon
  - Field vs. object based models
- Understand what can and cannot be done using each paradigm
- Know major types of data within each paradigm
  - And models to describe their spatial relationships
- Understand common operations on spatial data

Acknowledgement: some materials are borrowed from the following sources

- *Slides of textbook "Spatial Database: A tour"*
- *Lecture notes from Dr. Xun Zhou (UI) and Dr. Shashi Shekhar (UMN)*

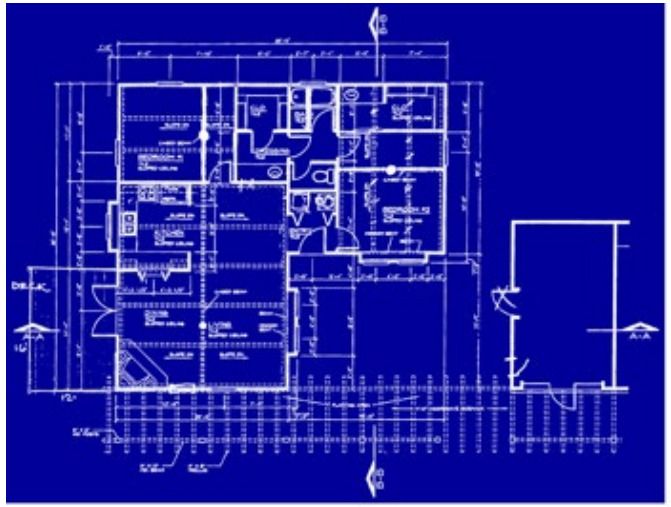
What is a Data Model?

- What is a model? (Dictionary meaning)
  - A set of plans (blueprint drawing) for a building
  - A miniature representation of a system to analyze properties of interest

Building

https://upload.wikimedia.org/wikipedia/commons/0/04/Garfield_Building_Detroit.jpg

Blueprint



http://causecapitalism.com/wp-content/uploads/2010/04/Blueprint.png

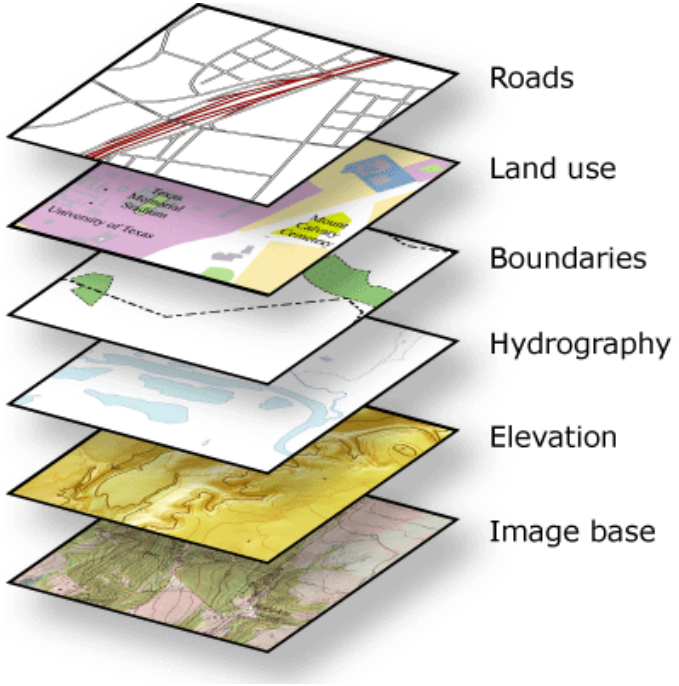
Why Data Models?

- Data models facilitate
  - Early analysis of properties, e.g. storage cost, querying ability, ...
  - Reuse of shared data among multiple applications
  - Exchange of data across organization
  - Conversion of data to new software / environment

- Different ways to think about data models
  - Fine-grained level: Spatial data representations: vectors, rasters, …
  - System level: Relational database, GIS…

Data Model Examples

- Geographic Information System (GIS) organizes spatial set as a set of layers
- Databases organize dataset as a collection of tables



http://desktop.arcgis.com/

Nontrivial tasks in map annotation & generalization

Roads

Road_ID	Length	Geometry
…	…	…

Landuse

Land_ID	Use	Geometry
…	…	…

Boundaries

Boundary_ID	Geometry
…	…

Hydrography

Location	Volume
…	…

Elevation

Location	Elevation
…	…

Image base

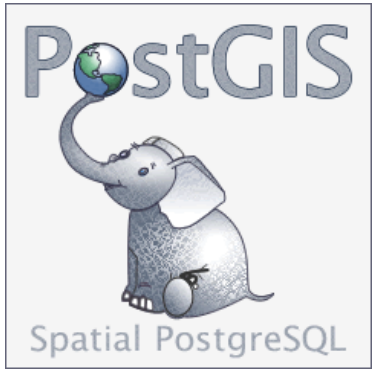
Pixel	Color
…	…

ORACLE SPATIAL

Object-relational databases

Tools for spatial data management

- ArcGIS, Quantum GIS, GeoData, …
- Spatial Databases: Oracle Spatial, PostGIS
- Spatial libraries and packages of programming languages
  - r-spatial for R
  - ArcGIS API, Shapely, geopandas, GDAL/OGR

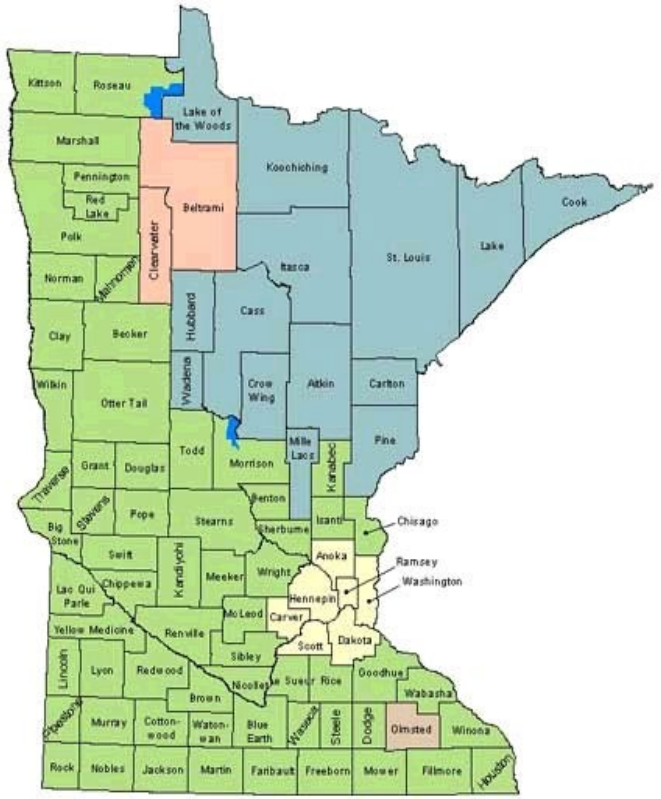


Spatial Data Models

- Field model
- Object model

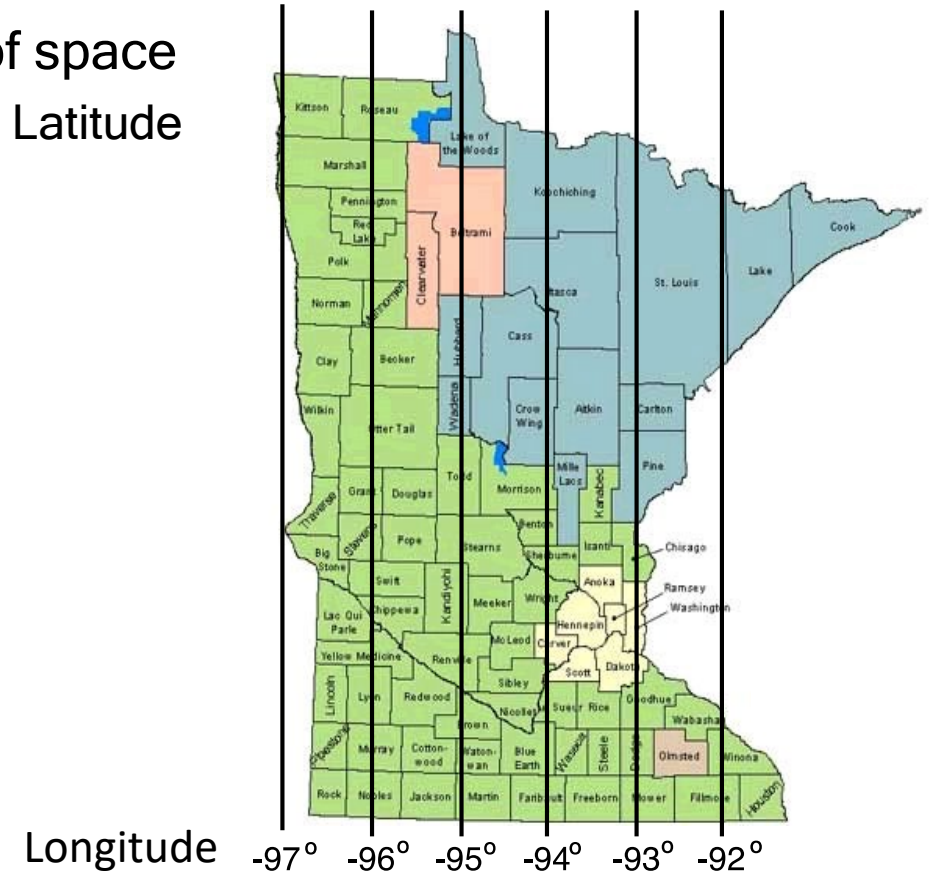
Field based Model

- Spatial Framework is a partitioning of space



Field based Model

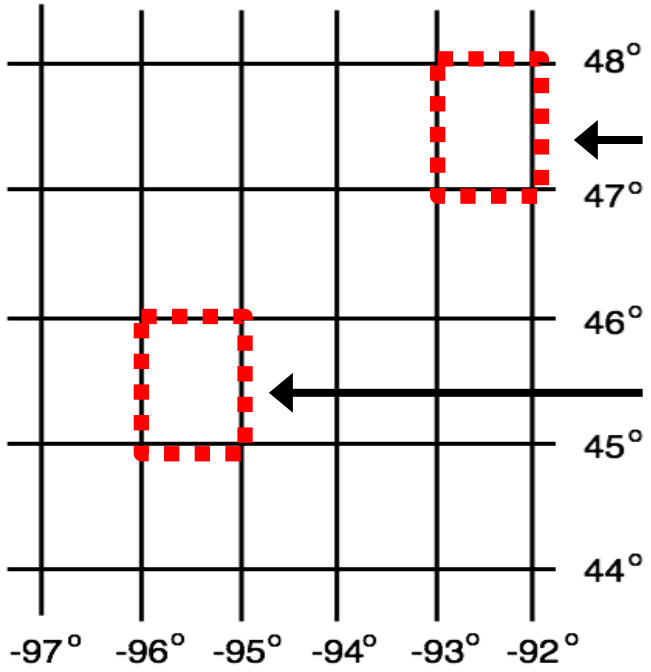
- Spatial Framework is a partitioning of space
  - e.g., Grid imposed by Longitude and Latitude



Field based Model

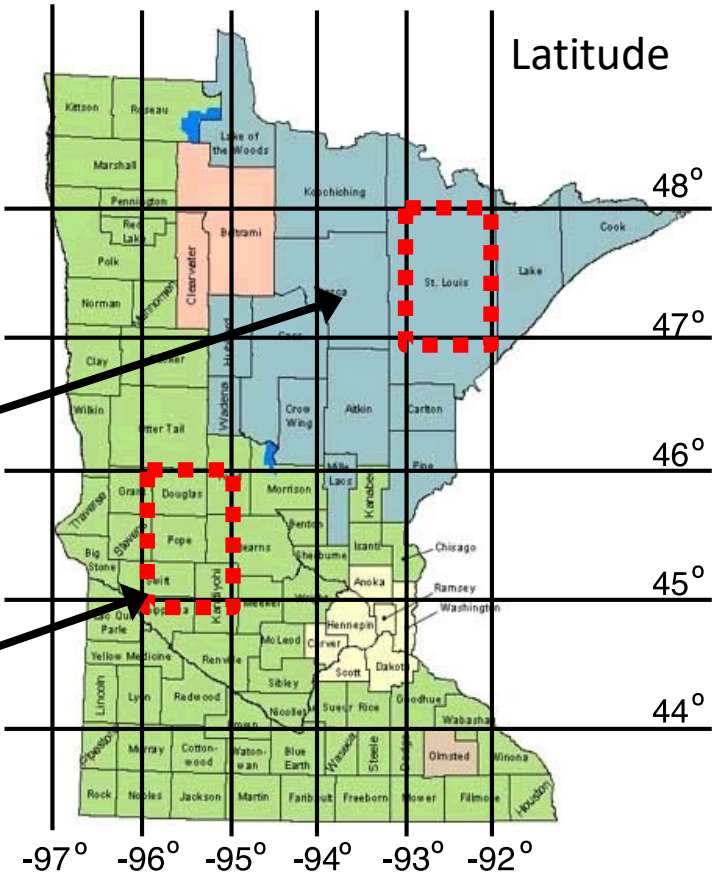
Different ways to partition the space

- Spatial Framework is a partitioning of space
  - e.g., Grid imposed by Longitude and Latitude



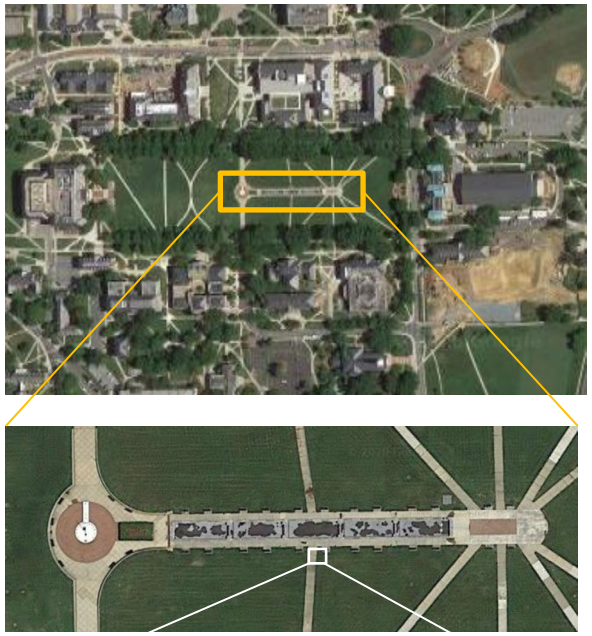
Forested Area

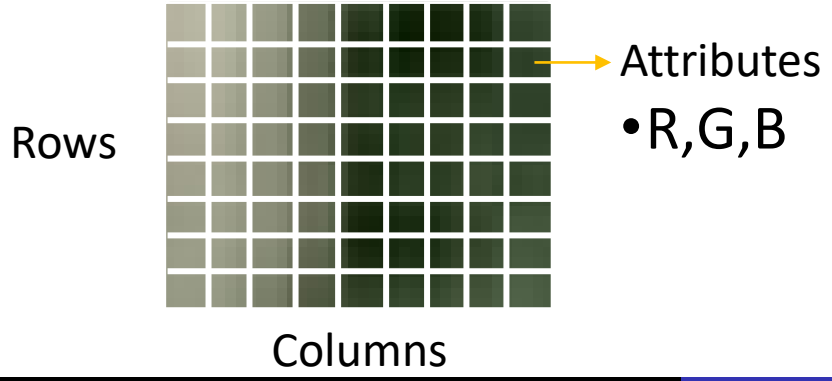
Agricultural



Field/Raster

- An “image”
  - Rows
  - Columns
  - Depth





- Get a value
  - Data[i,j,k]
  - Implicit spatial locations

Think about your excel spreadsheet

What it looks like for a single-attribute raster:

…

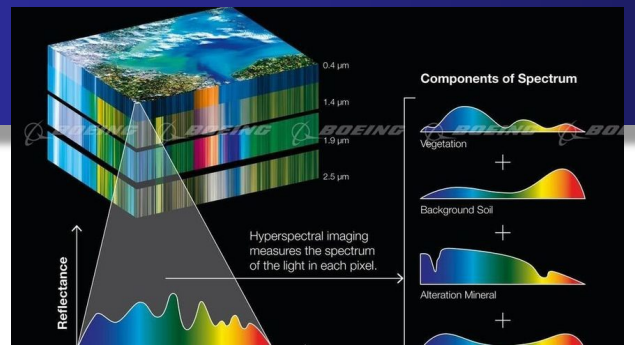
… 100, 201, 255, 255 …
110, 205, 255, 255
110, 205, 255, 255
115, 209, 255, 255

…

No need to store row_id & column_ids
Only need to know

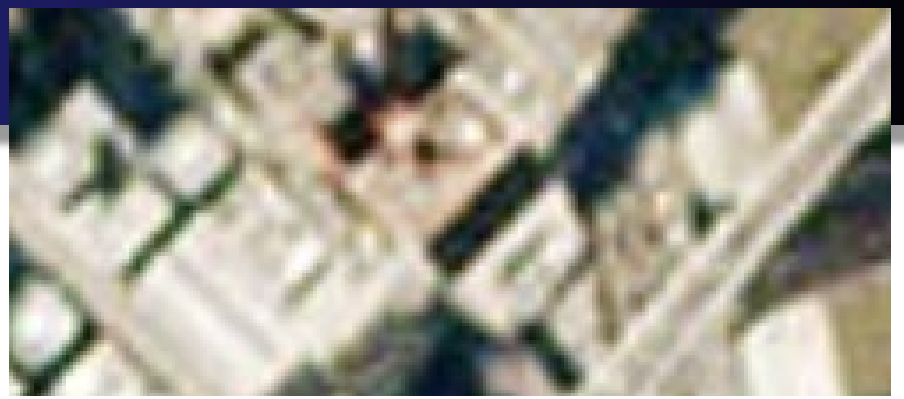
- #rows, #columns and cell sizes
- Location of the cell at [1,1]

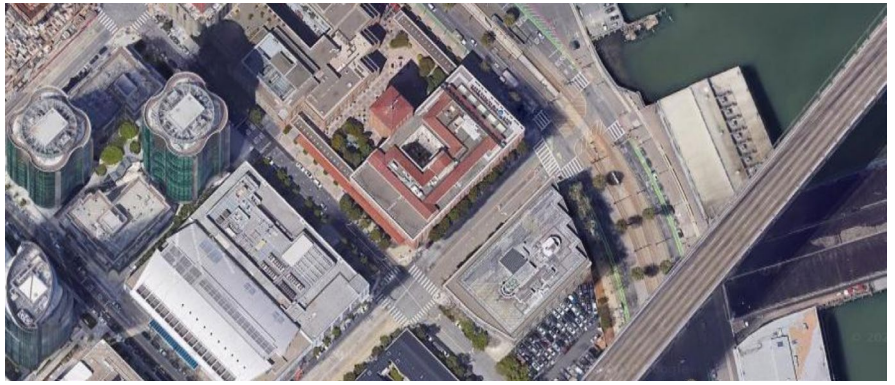
Field/Raster

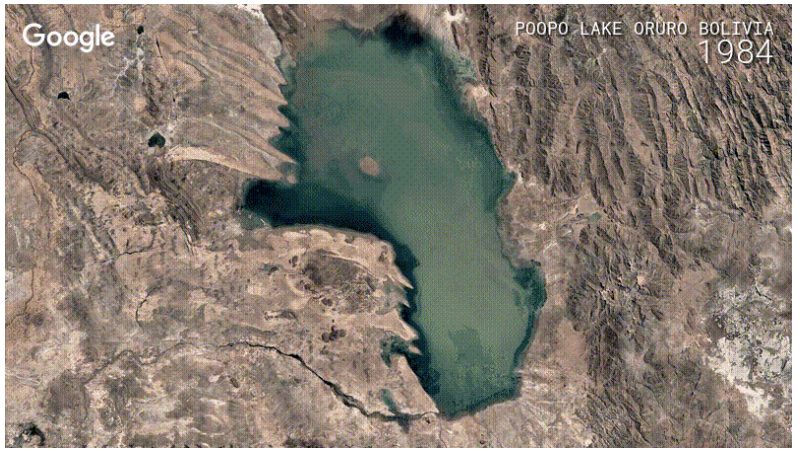


- Examples
  - Satellite imagery
    - Lower spatial resolution (e.g., 1m, 30m, 1km)
    - High temporal frequency
    - Multispectral, hyperspectral (high dimensional)
  - Aerial imagery (e.g., by Unmanned Aerial Vehicles)
    - Higher spatial resolution (e.g., 3-inch)
    - Lower frequency and fewer bands
  - A time-series of imagery

Traditional methods do not work well for very high-resolution datasets

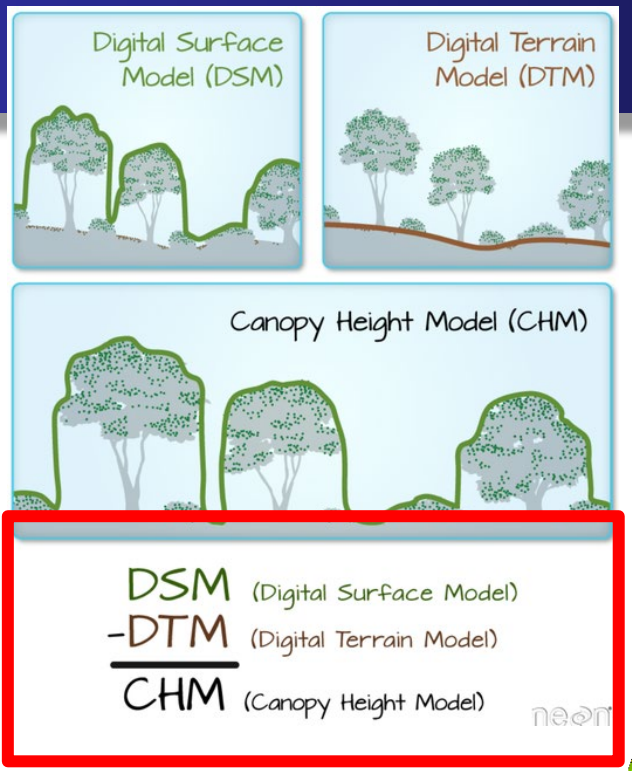


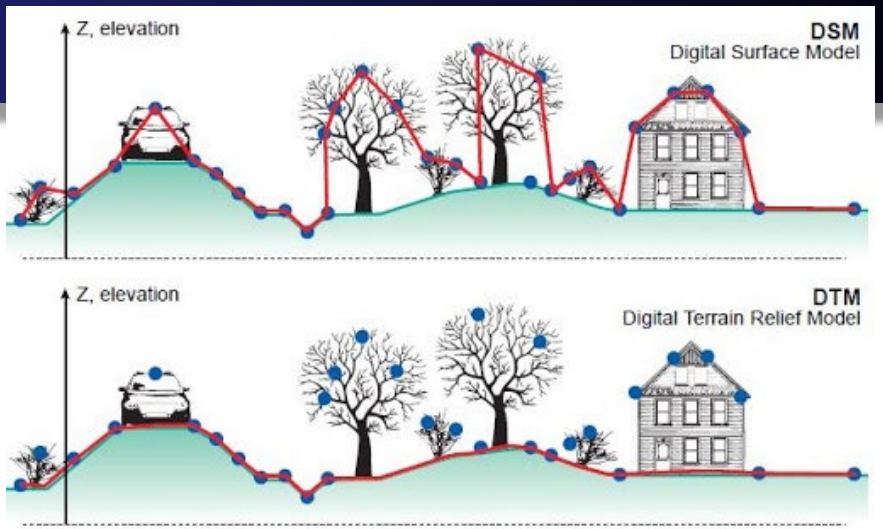


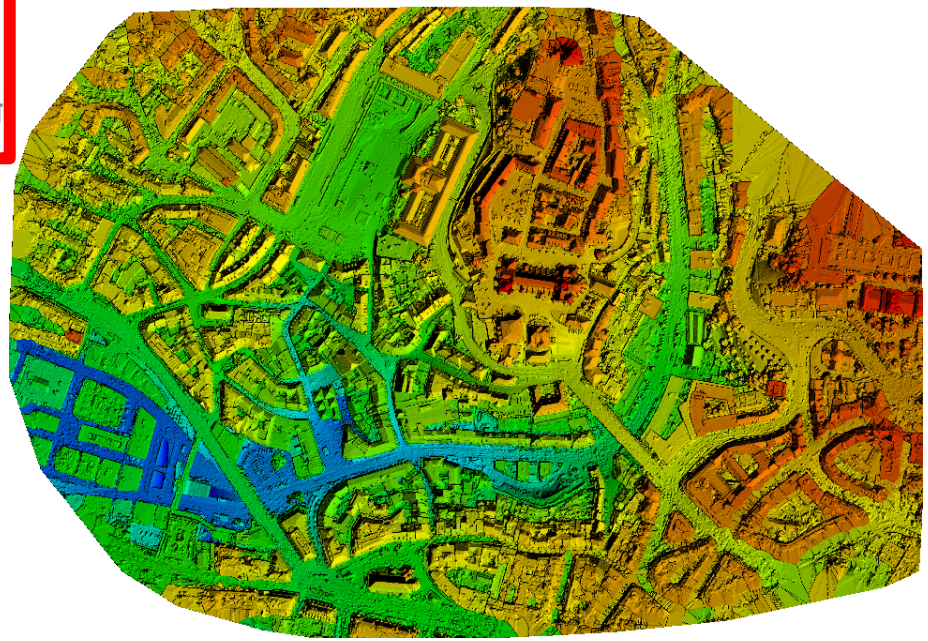


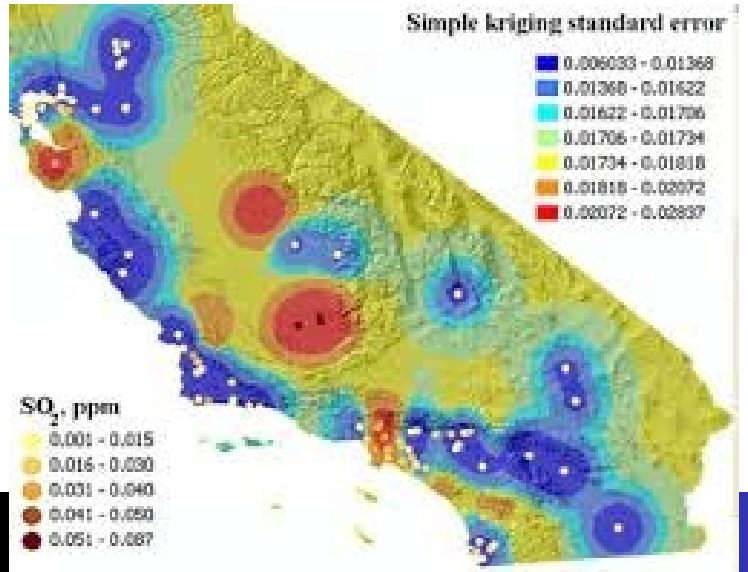
Field/Raster

- Examples continued
  - Height models
    - Digital terrain models
    - Digital surface models
    - Canopy height models
  - Interpolated air quality maps
  - Grid…







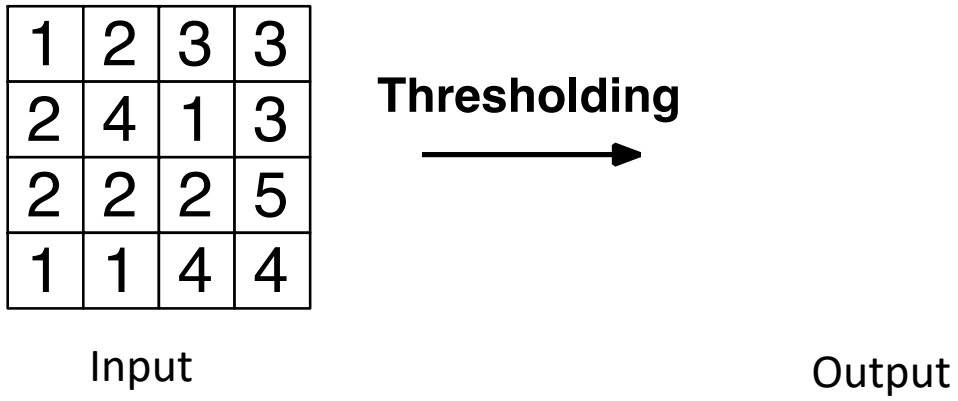


Field based Model – Field Operations

- Field Operations (operations between fields)
  - Can be classified into four groups:
    - Local
    - Focal
    - Zonal
    - Global

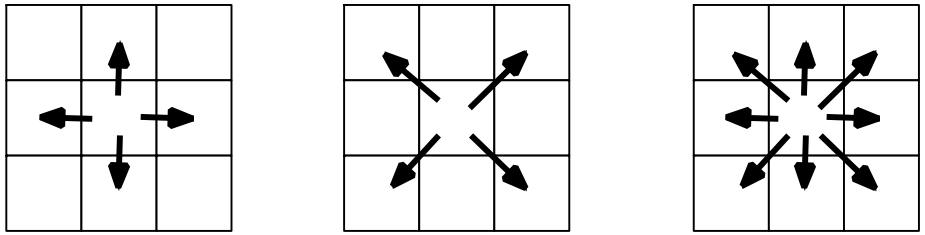
Local Operation

- Local operation: output at a given location depends only on the input at that location
- Example: thresholding
  - For each location (cell)
    if input > 2, then output = 1
    else, output = 0



Focal Operation

- Focal operation: output at a given location depends on the input in a small neighborhood of the location
- Example: FocalSum
  - Different neighborhoods (e.g., Rook, Bishop and Queen)



Rook, Bishop and Queen neighborhoods

Input Output

Zonal Operation

- Zocal operation: output at a given location depends on the input in a pre-defined zone around the location
- Example: ZonalSum

1	2	3	3
2	4	1	3
2	2	2	5
1	1	4	4

Input

Zonal map

Output

Global Operation

- Global operation: output at a given location depends on input in all locations
- Example: Does the cell has the largest value?

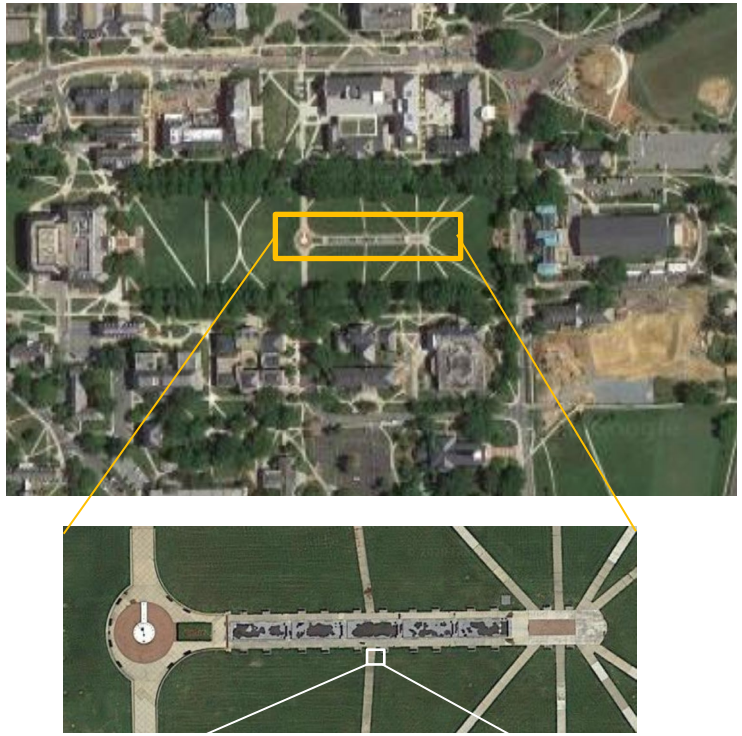
Quiz

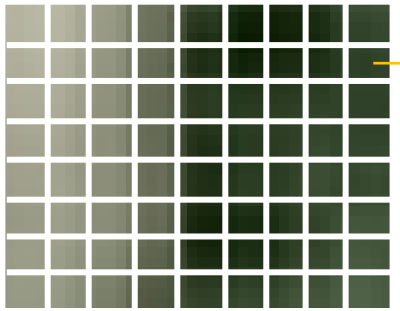
- Which type of operation is the following operations on elevation field classified?

(i) Identify peaks (points higher than its neighbors)
(ii) Identify mountain regions (elevation over 2000 feet)
(iii) Identify the highest point in the whole area
(iv) Determine average elevation within a set of river basins

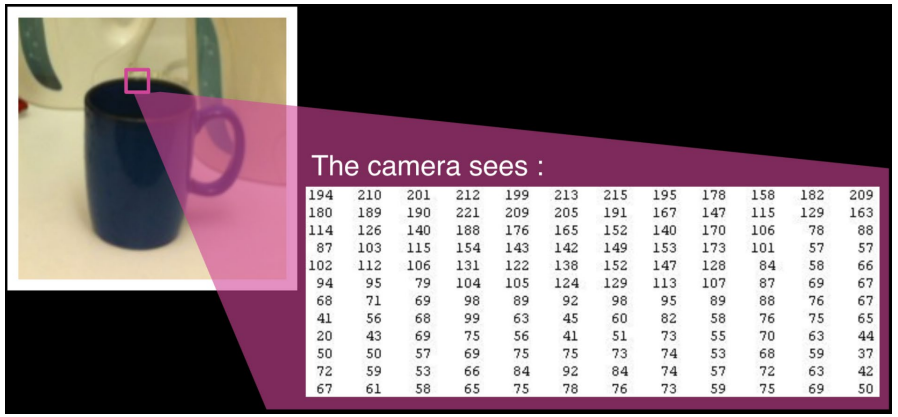
(a) Local
(b) Focal
(c) Zonal
(d) Global

Field vs. Object Models



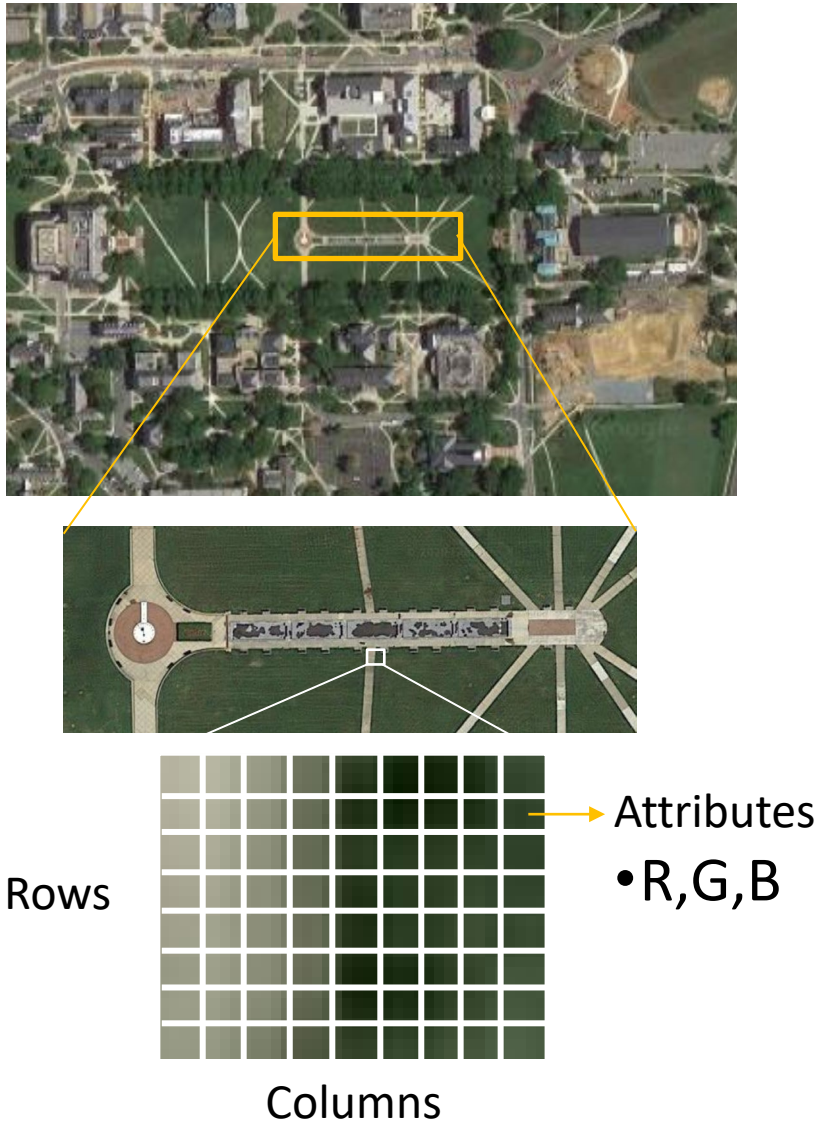


- What cannot be trivially done on field models?
  - A. Calculate the portion of pixels with a greenish color (e.g., vegetation)
  - B. Which building has the largest footprint?
  - C. Which pixel has the highest brightness?
  - D. What is the distance between LeFrak Hall and the Office of Registrar?



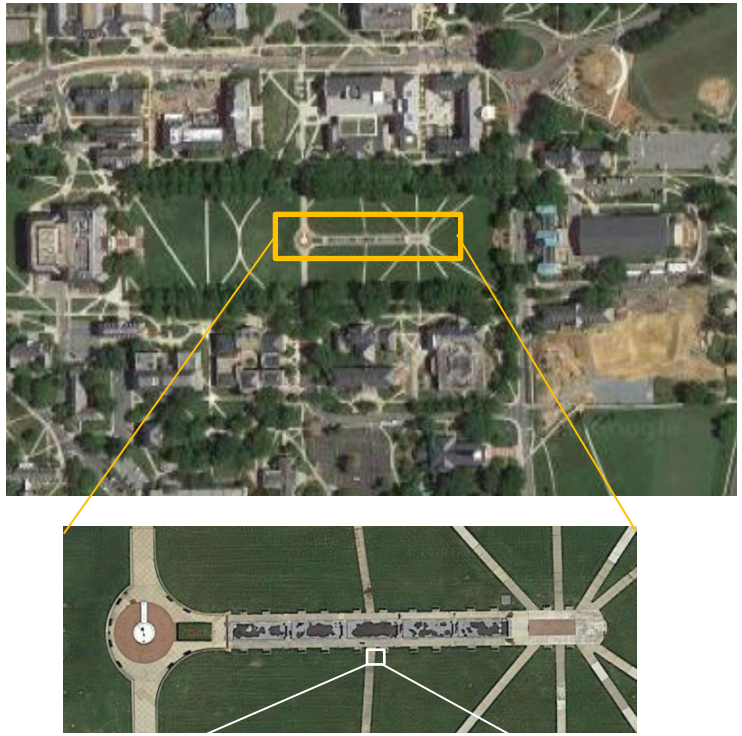
Source: Andrew Ng. http://cs229.stanford.edu/materials/CS229-DeepLearning.pdf

Field vs. Object Models



- What cannot be trivially done on field models?
  - A. Calculate the portion of pixels with a greenish color (e.g., vegetation)
  - B. Which building has the largest footprint?
  - C. Which pixel has the highest brightness?
  - D. What is the distance between LeFrak Hall and the Office of Registrar?

Field vs. Object Models

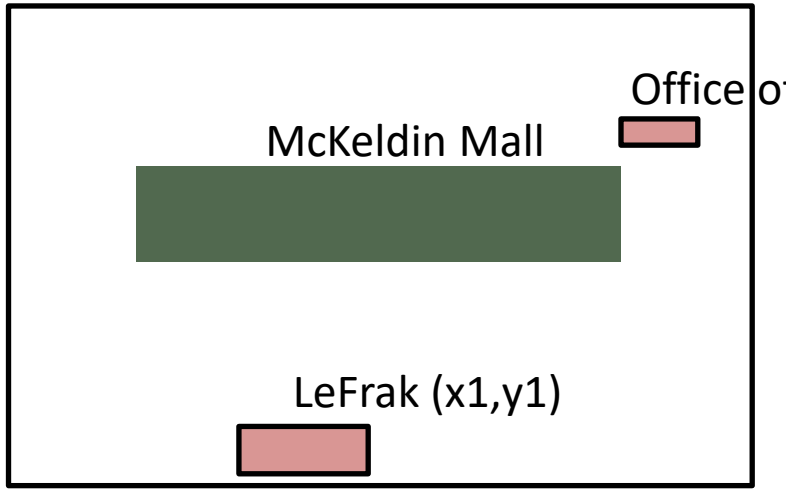


Rows

Columns

Attributes

- R,G,B



Object Model

- Object model concepts
  - Spatial information stored as vectors (e.g., points, lines, polygons…)
  - Objects have attributes and operations
  - Attribute: a simple (e.g., numeric, string) property of an object

Spatial attributes	**Non-spatial attributes**
Footprint	Build ID
Address	Year
Latitude	
Longitude	

Object Model – Example of Keller Hall

Object: Kenneth Keller Hall in UMN campus

Spatial attributes
Footprint
Latitude: 44.97
Longitude: -93.23

Non-spatial attributes
Build ID: 165
Year: 2010

Address? 200 Union St SE, Minneapolis, Minnesota, 55455

Classifying Spatial Objects

- Spatial objects are objects with spatial attributes (and non-spatial attributes)
- Spatial objects are of many types
  - Simple
    - 0 dimensional (points) , 1 dimensional (curves) , 2 dimensional (surface)

St. Paul (center of city)

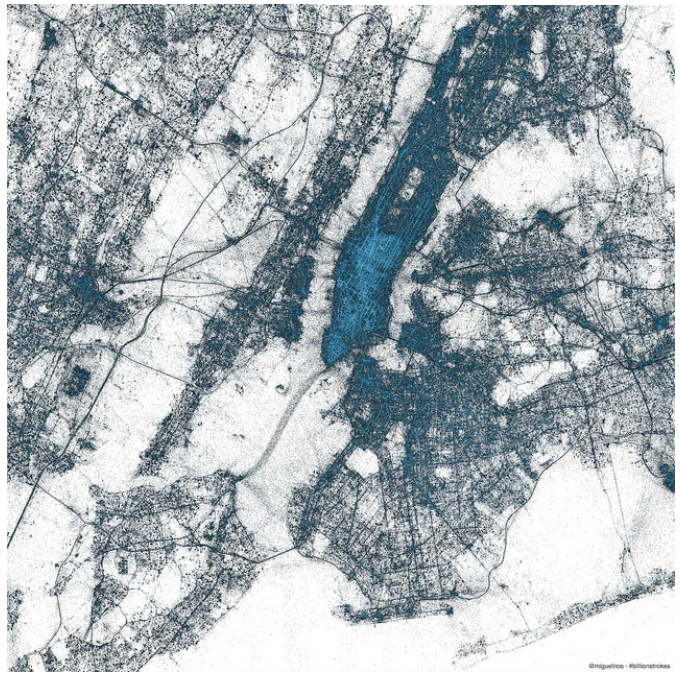
Mississippi River (center line)

Minnesota State

Point

- [**x, y**, z, t, attribute_1, …]
- Non-continuous phenomenon

- Examples
  - Geotagged tweets
  - Check-in data
  - Point-of-interest (POI)
  - Environment: Air quality sensor readings…
  - Public safety: crime cases…
  - Public health: not likely

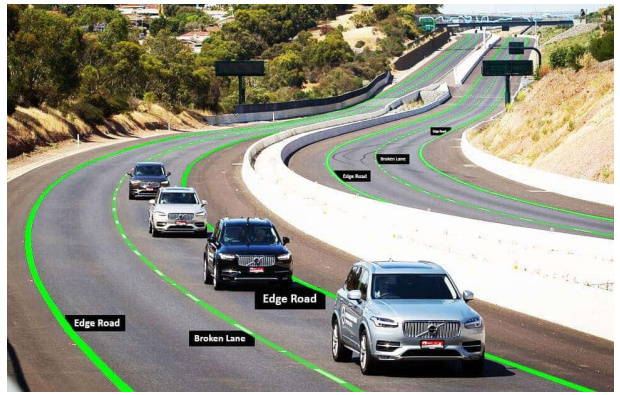


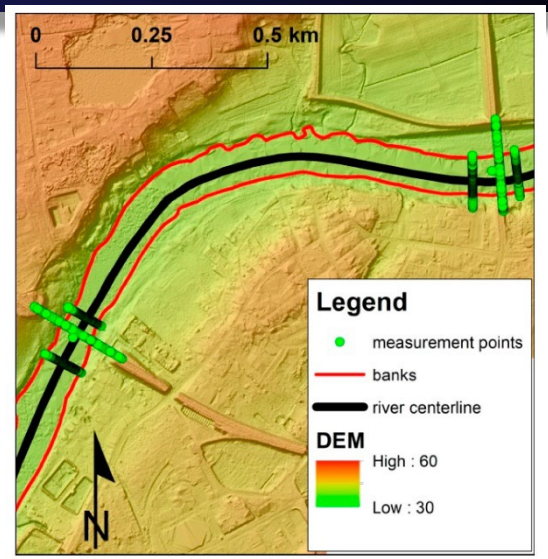
Billions…

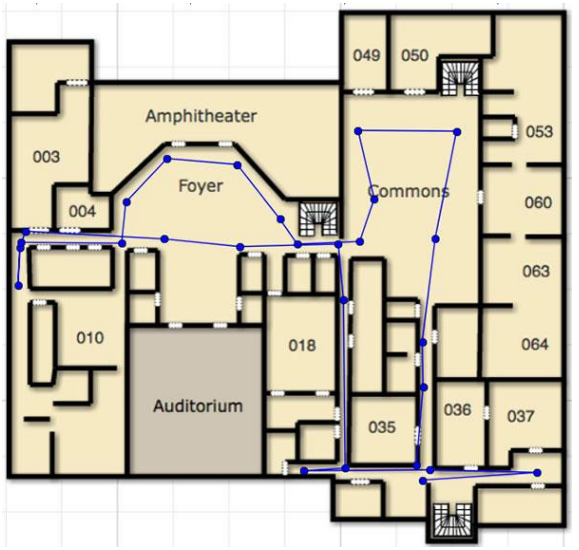
Polyline (Line String)

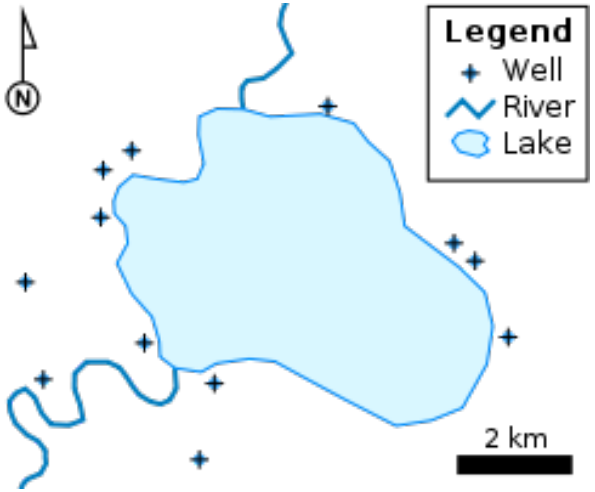
- [(x1,y1), (x2,y2), (x3,y3), …, attribute_1, …]
- Linear phenomenon

- Examples
  - River lines
  - Road lines
  - Walls/fences (especially indoor)
  - **More commonly used in visualization**





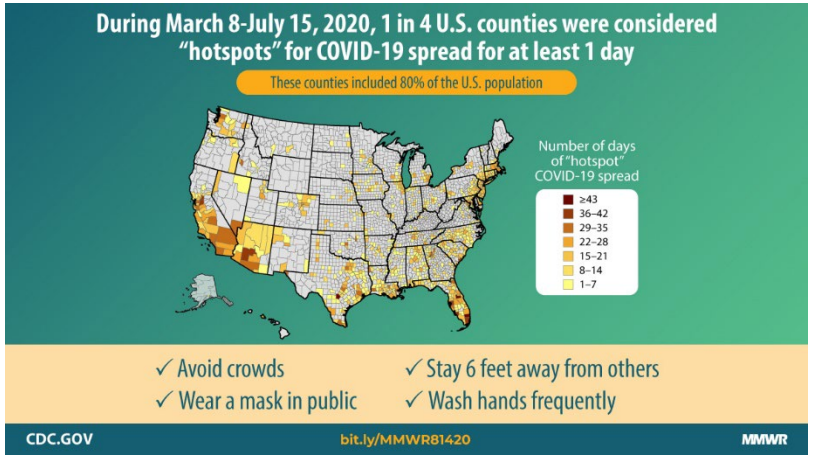




Polygon

- [(x1,y1), (x2,y2), …, (x1,y1), …, attribute_1, …]

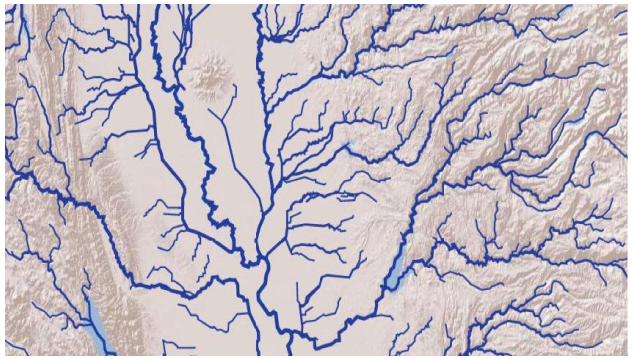
- Examples
  - Building footprints
  - Abstraction of real-world objects
    - Rectangles for buildings, vehicles…
    - Circles for trees…
  - Space partitions
    - City and county maps
    - Hydrological response units



Classifying Spatial Objects - Continued

- Collections
  - Multi point , Multi linestring , Multi polygon

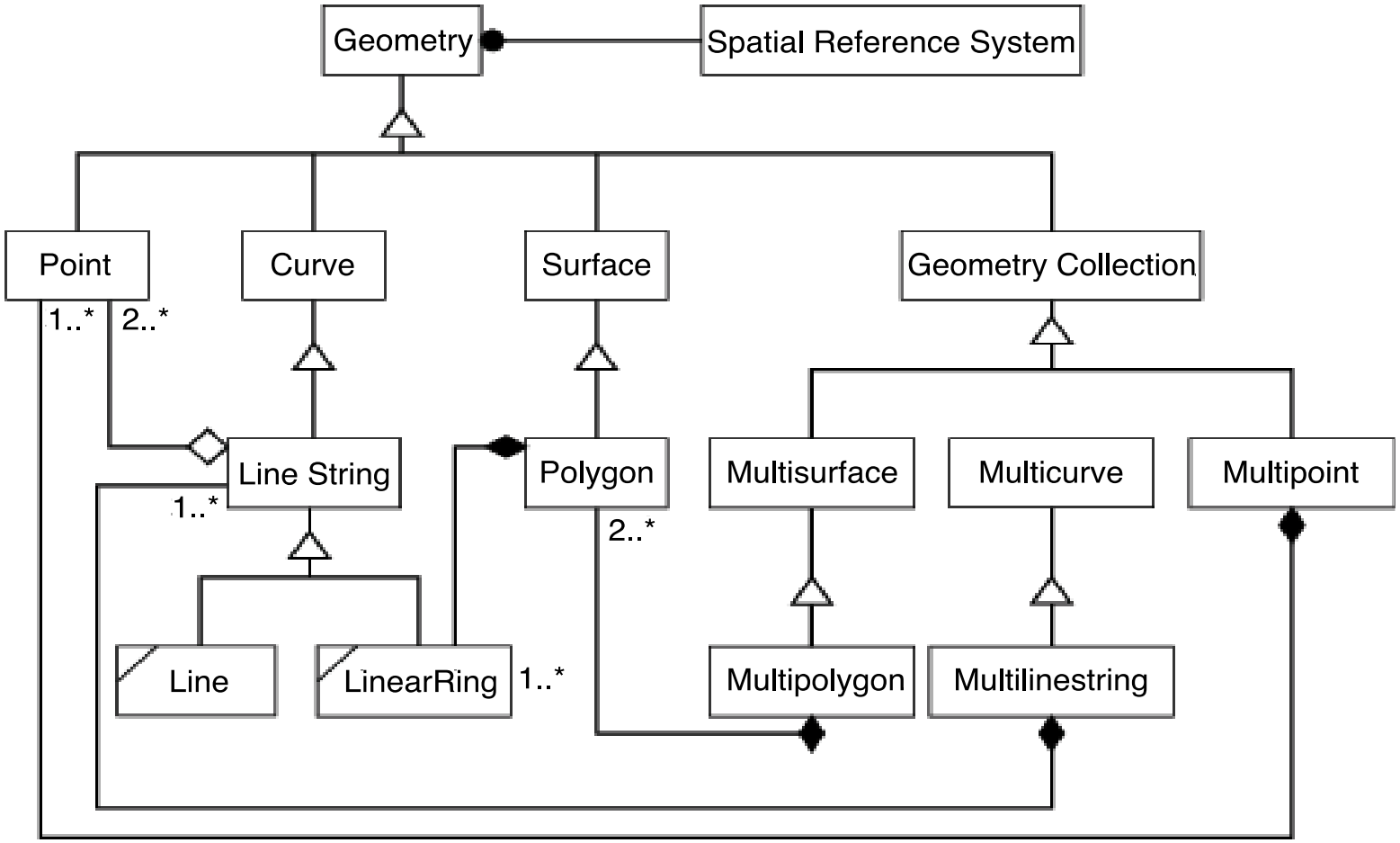
Wind farm



Group of rivers

Hawaii

Spatial Object Types in OGIS Data Model

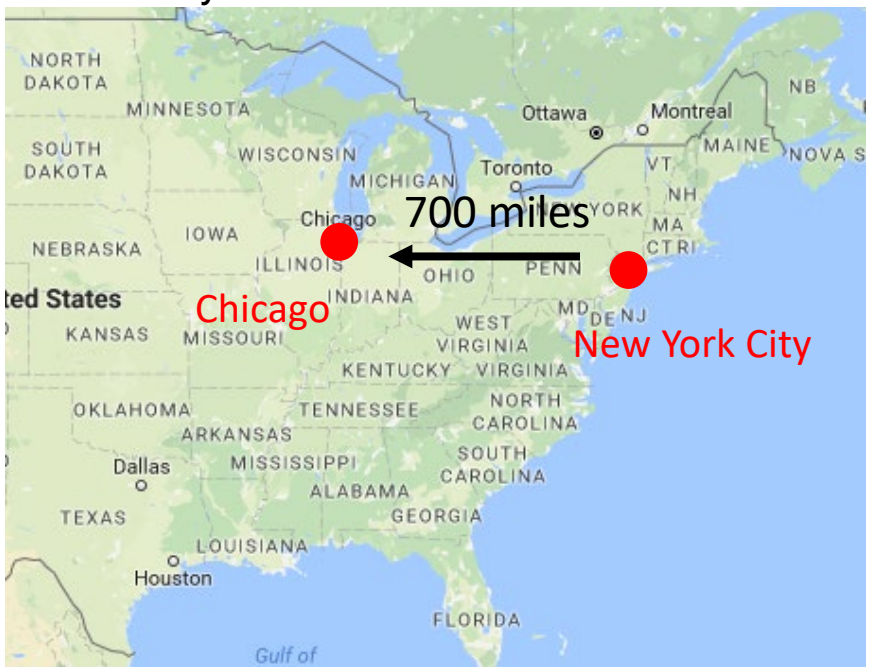


Operations on Spatial Objects

- Classifying operations
  - Set based: **Union** of East and West Germany
  - Topological operations (example on paper): Boundary of East Germany **touches** boundary of West Germany
  - Directional: New York city is to east of Chicago
  - Metric: Chicago is about 700 miles from New York city.

West Germany

East Germany



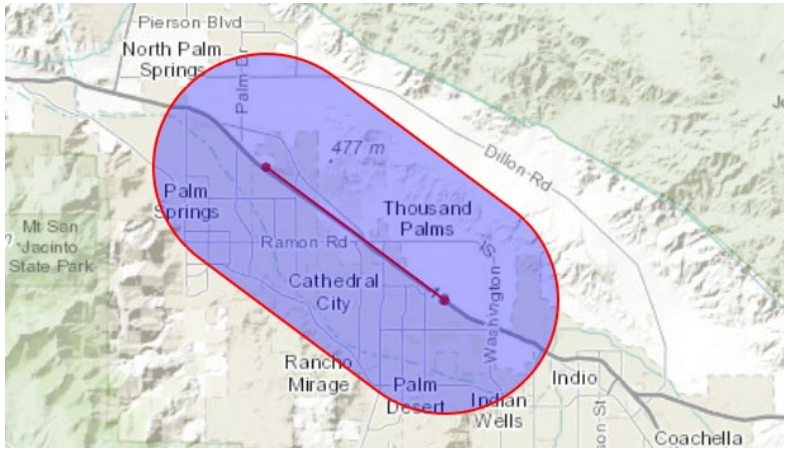
Other Object-Model Operations

A sample of Spatial Analysis operations listed in the Open Geodata Interchange Standard (OGIS) for SQL

Spatial Analysis	Distance	Returns the shortest distance between two geometries
	Buffer	Returns a geometry that consists of all points whose distance from the given geometry is less than or equal to the specified distance
	ConvexHull	Returns the smallest convex geometric set enclosing the geometry
	Intersection	Returns the geometric intersection of two geometries
	Union	Returns the geometric union of two geometries
	Difference	Returns the portion of a geometry that does not intersect with another given geometry
	DymmDiff	Returns the portions of two geometries that do not intersect with each other

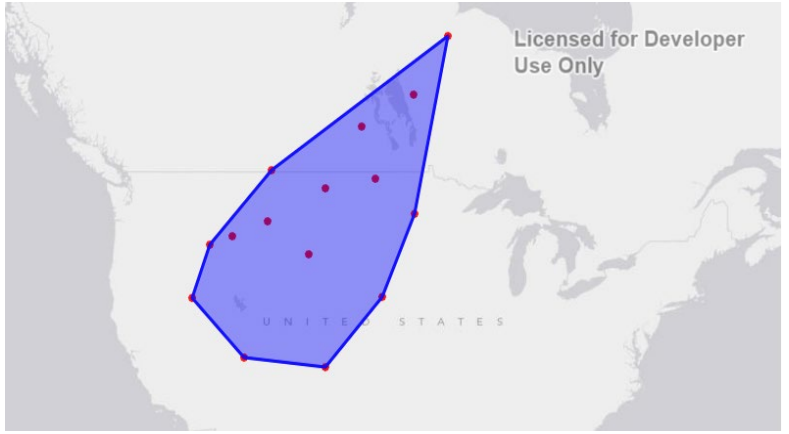
Operations on Spatial Objects

Buffer of a lineString



www.esri.com

ConvexHull of a set of points

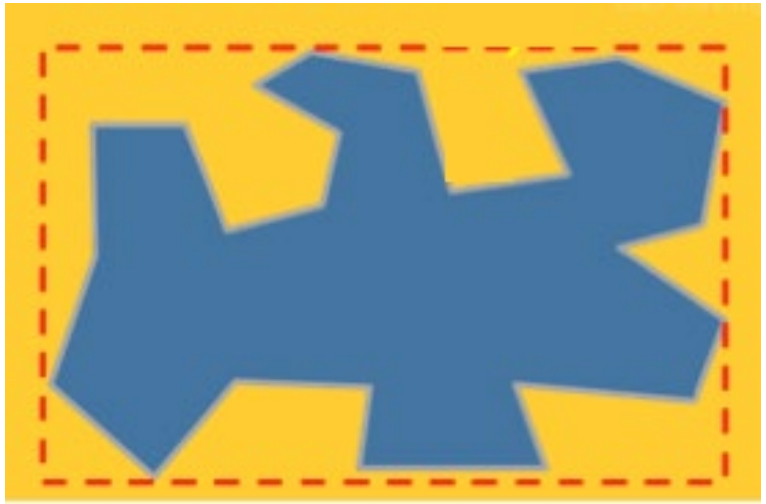


Other Object-Model Operations

A sample of Basic function operations listed in the OGIS for SQL

Basic Functions	SpatialReference ()	Returns the underlying coordinate system of the geometry
	Envelop ()	Returns the minimum orthogonal bounding rectangle of the geometry
	Export ()	Returns the geometry in a different representation
	IsEmpty ()	Returns true if the geometry is a null set
	IsSimple ()	Returns true if the geometry is a simple (no self-intersection)
	Boundary ()	Returns the boundary of the geometry

Envelop (Minimum orthogonal bounding rectangle) of a polygon



Topological Relationships

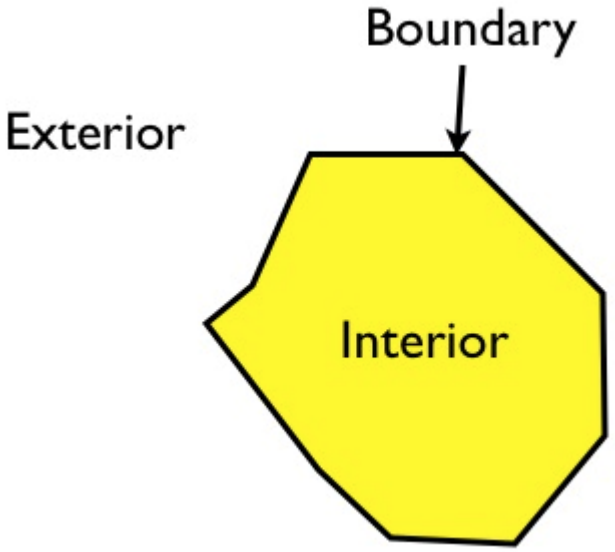
- Topological Relationships
  - Invariant under elastic deformation (without tear, merge).
  - Two countries which touch each other in a planar paper map will continue to do so in spherical globe maps.
- Topology is the study of topological relationships
- Example queries with topological operations
  - What is the topological relationship between two objects A and B ?
  - Find all objects which intersect with object A? (example in ArcGIS)

Planar map

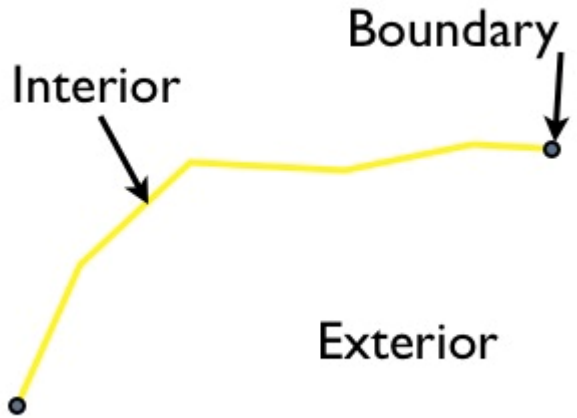
Spherical globe map

Topological Concepts

- Interior, boundary, exterior



Polygons

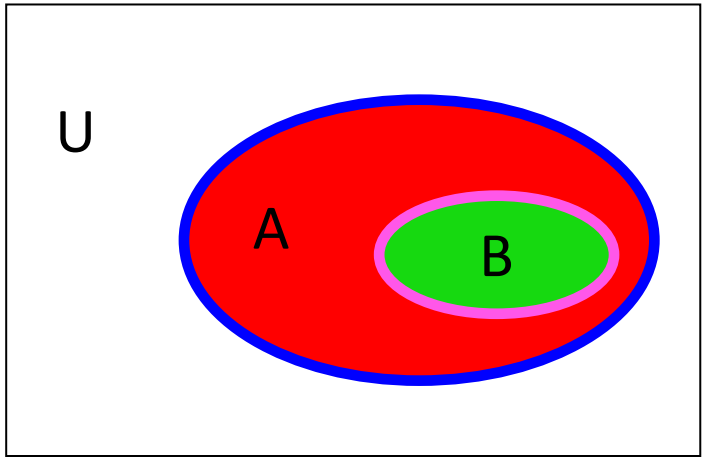


Polylines

https://postgis.net/workshops/postgis-intro/de9im.html

Topological Concepts

- Interior, boundary, exterior
  - Let A and B are two objects in a “Universe” U.



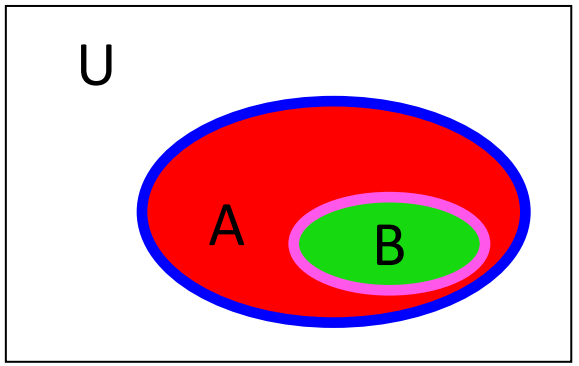
Red is interior of A: **Interior(A)**

Blue is boundary of A: **Boundary(A)**

White – Blue - Red is exterior of A: **Exterior(A)**

Nine-Intersection Model (9IM) of Topological Relationships

- Many topological relationship between A and B can be
  - Specified using 9 intersection model
  - A and B are spatial objects in a two dimensional plane.
  - Can be arranged as a 3 by 3 boolean matrix



$$\text{Matrix (A,B)} = \begin{pmatrix} In(A) \cap In(B) & In(A) \cap Bo(B) & In(A) \cap Ex(B) \\ Bo(A) \cap In(B) & Bo(A) \cap Bo(B) & Bo(A) \cap Ex(B) \\ Ex(A) \cap In(B) & Ex(A) \cap Bo(B) & Ex(A) \cap Ex(B) \end{pmatrix}$$

In = Interior, Bo = Boundary, Ex = Exterior

$$\text{Matrix (A,B)} = \begin{pmatrix} T & T & T \\ F & F & T \\ F & F & T \end{pmatrix}$$

T = True, F = False