Introduction to Geospatial Concepts

April 9, 2021

Public Policy Leaders Program – GIS Workshop - I

Parmanand Sinha

Computational Scientist, Research Computing Center

About RCC-GIS

Support for Geospatial Information Science Research

Assistance for all UChicago students, faculty, and staff

- Training
- Consultation
- Events

https://gis.rcc.uchicago.edu

Topics

- GIS, mapping, and spatial data
- The GIS data types and file formats
- Geocoding
- Basic GIS operations

ArcGIS Desktop is only available on Windows

UChicago Virtual Lab (vLab)

https://academictech.uchicago.edu/vlab/

Start **Microsoft Remote Desktop**

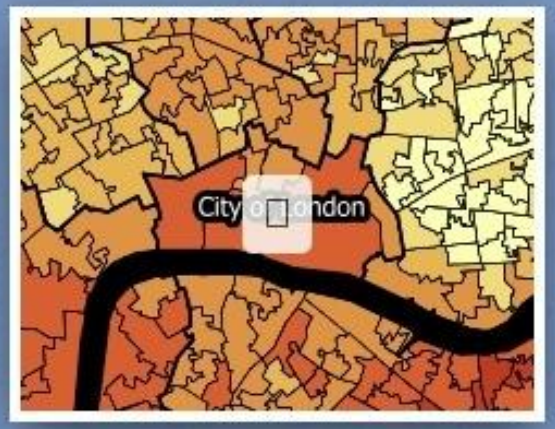
And connect to **vlab.uchicago.edu**

GIS

A geographic information system (GIS) lets us visualize, question, analyze, interpret, and **understand data to reveal relationships, patterns, and trends**.

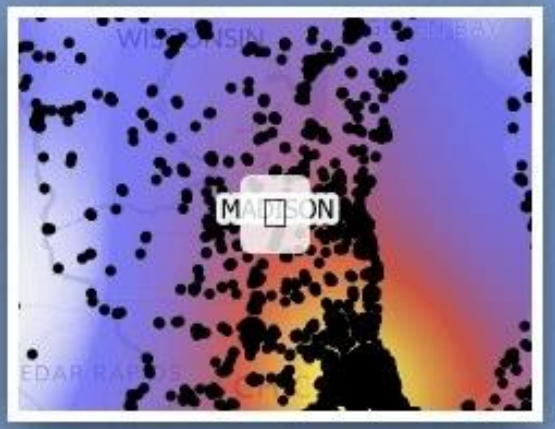
- esri

LOCAL GOVERNMENT



DEMOGRAPHICS

BUSINESS ANALYSIS

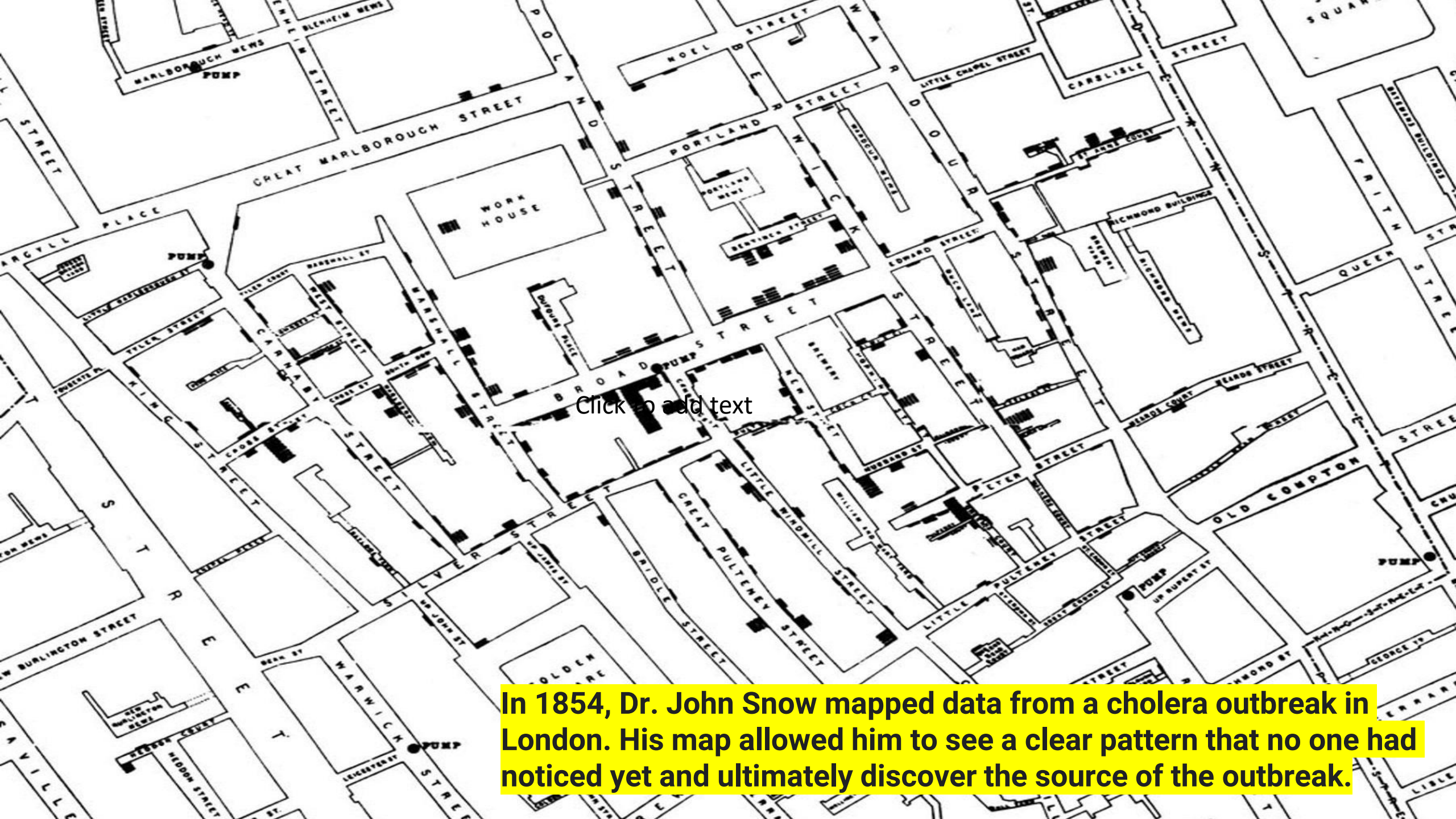


LAND USE AND PLANNING

GIS

A geographic information system (GIS) lets us visualize, question, analyze, interpret, and **understand data to reveal relationships, patterns, and trends**.

- esri


Click to add text
In 1854, Dr. John Snow mapped data from a cholera outbreak in London. His map allowed him to see a clear pattern that no one had noticed yet and ultimately discover the source of the outbreak.

GIS

Tobler's 1st Law of Geography

Everything is related to everything else, but near things are more related than distant things

GIS Themes: Layers

- Pull apart themes of your map to make layers
- Layers sit on top of one another
- Spatial relationships define how these layers interact with one another

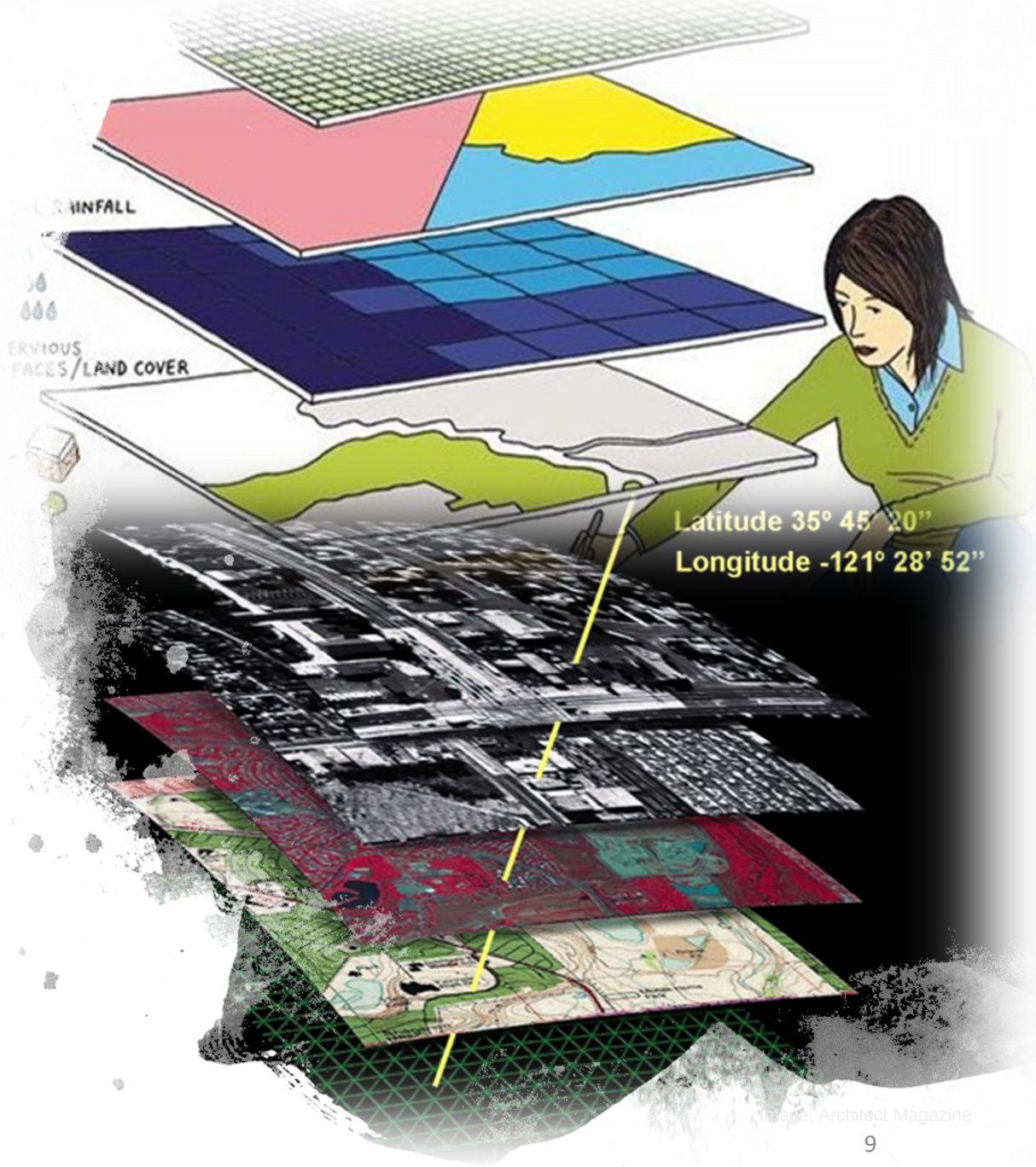


Image: Architect Magazine

What can GIS do?

As an integrated system for geographic data

- Capturing data
- Storing data
- Querying data
- Analyzing data
- Displaying data

What can GIS do?

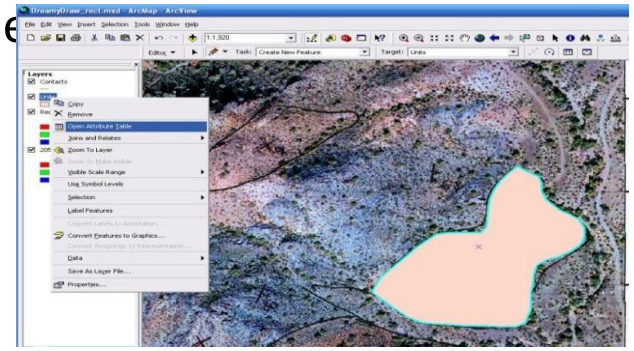
- Geographic Information System (GIS):

  An integrated system for geographic data

  - Capturing data
  - Storing data
  - Querying data
  - Analyzing data
  - Displaying data

Capturing data

- Digitizing
  - Creating a digital copy of existing data
  - Paper maps, Aerial imagery, Topographic maps
  - Output is a GIS friendly vector format



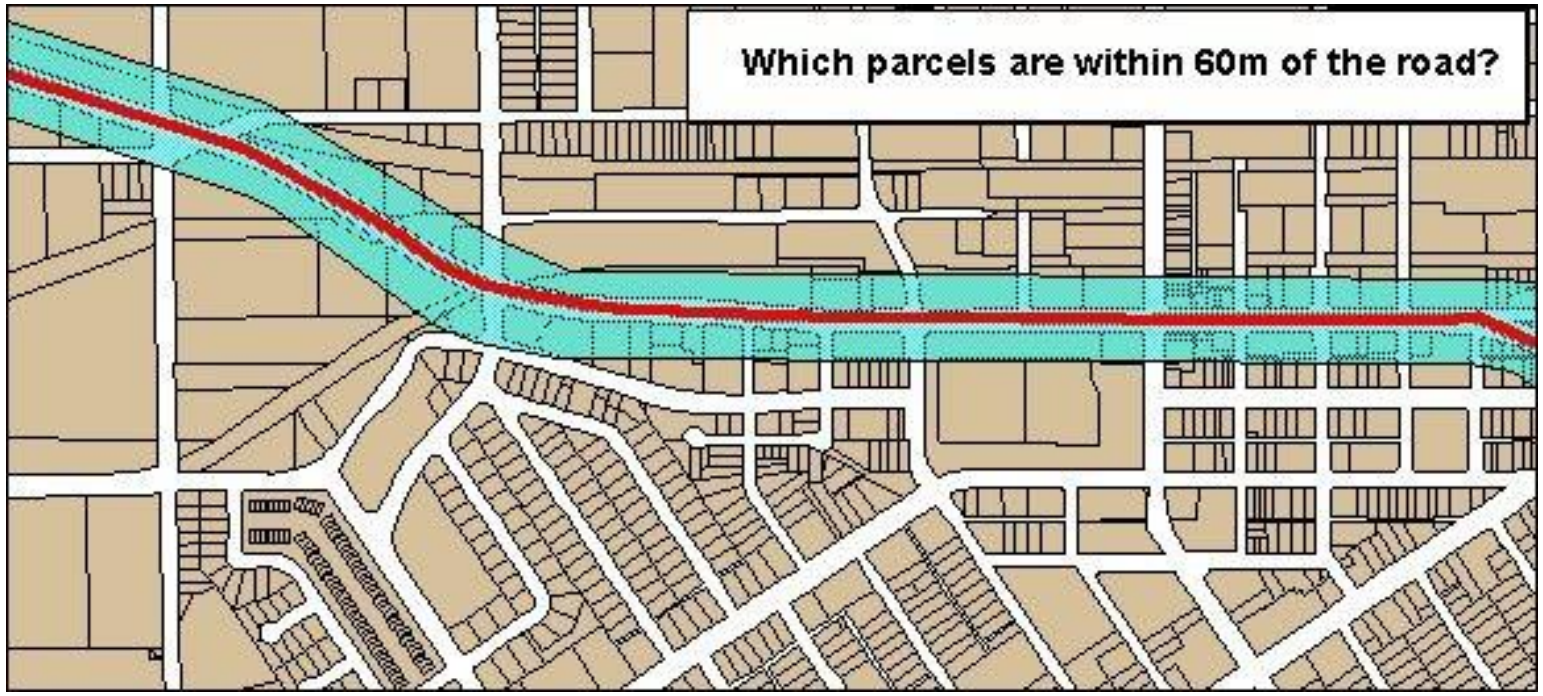
- GPS
  - Global Positioning System
  - Coverage of entire planet
  - Data Formats:
    - Temporal – accurate to about 14 nanoseconds
    - Spatial – sub-meter accuracy

Storing data

- Files
  - Shapefile
  - KML (Google)
  - Spreadsheet
  - Esri Geodatabase
- Spatial Database
  - PostGIS, Oracle, SQL Server
  - Traditional database management system that also stores vector feature geometry and location data

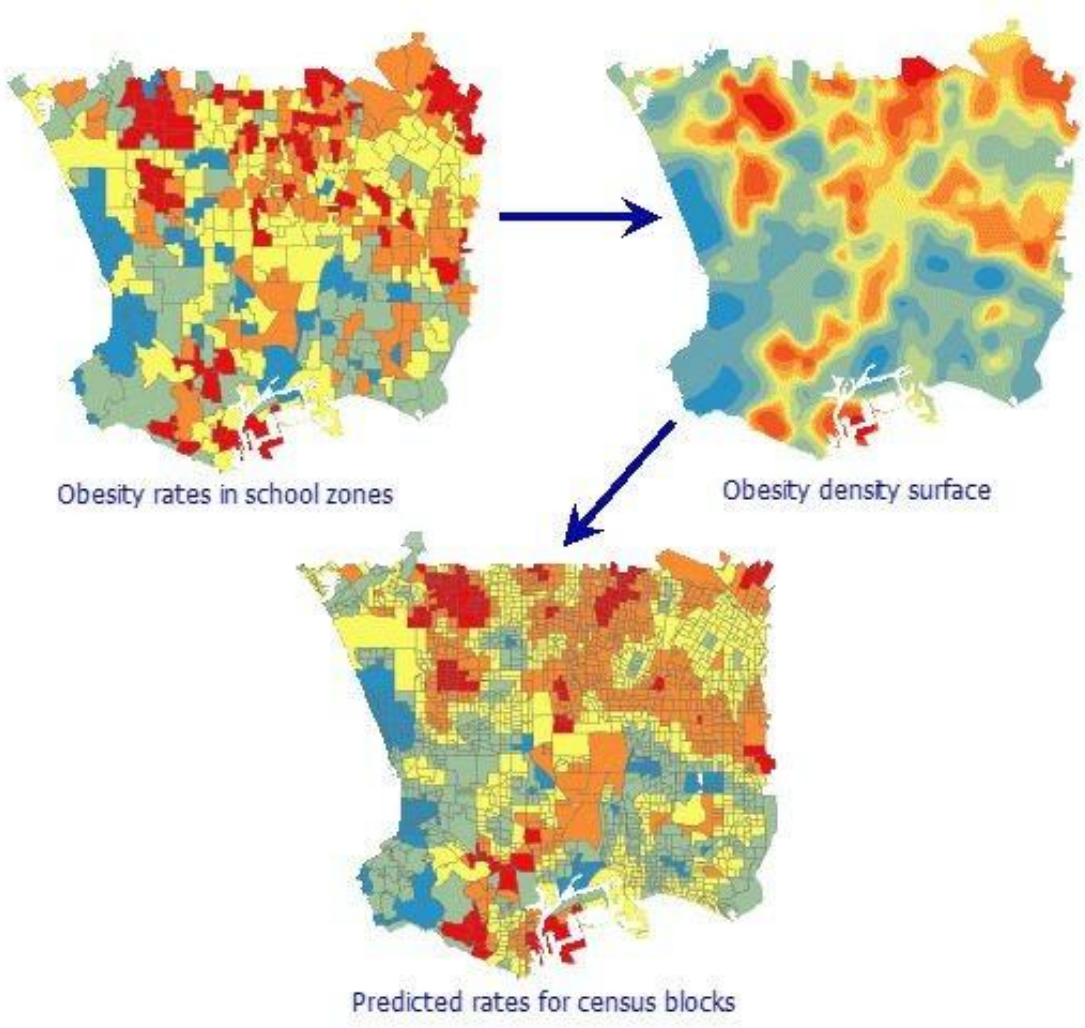
Querying data

Ask a question of the data based on location or other attribute



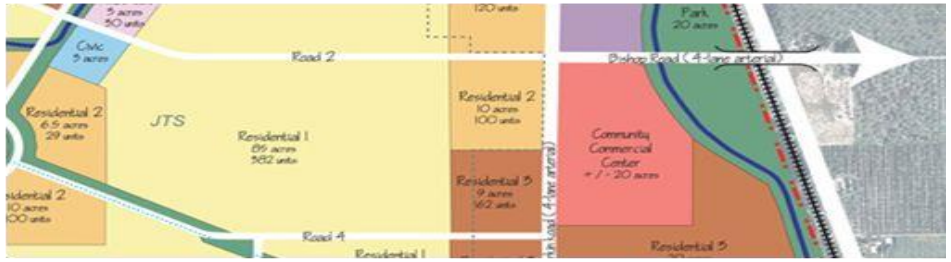
Analyzing data

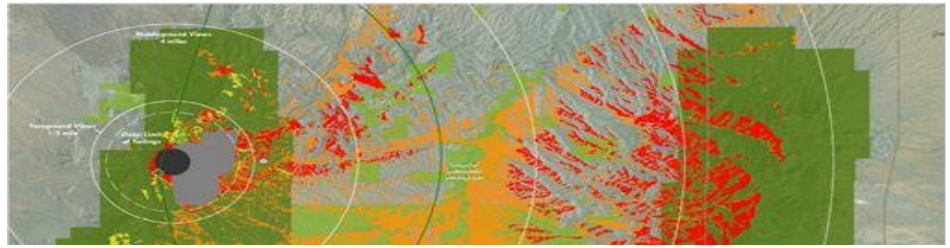
- Performing operations on the spatial data
  - “How far am I from the nearest park?”
  - Overlays
  - Extracts
  - Proximity
  - Spatial Statistics



Displaying data

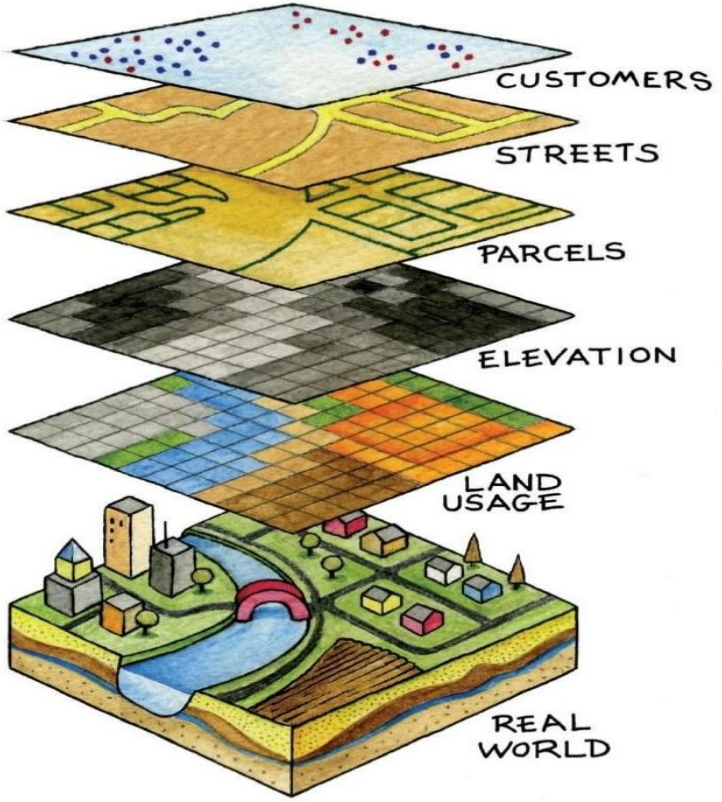
- Mapping
  - The display of spatial data for a specific purpose
  - Visualization
    - Re-emerging as a popular means to artistically portray traditional data
    - Not necessarily spatial data
- Output Format
  - Print or Web



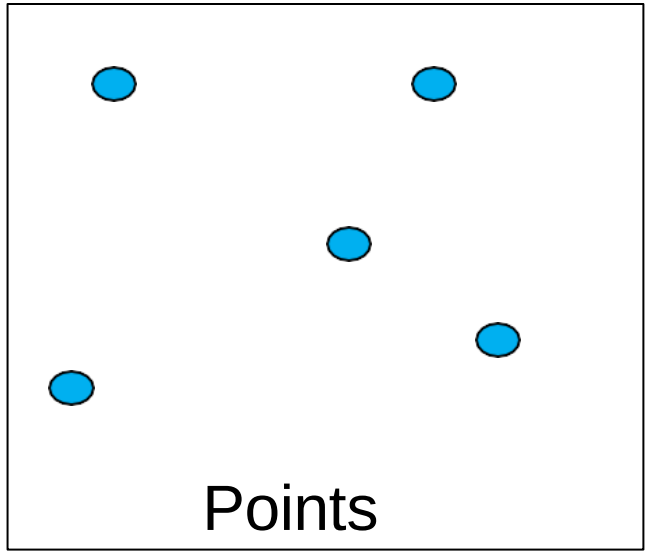


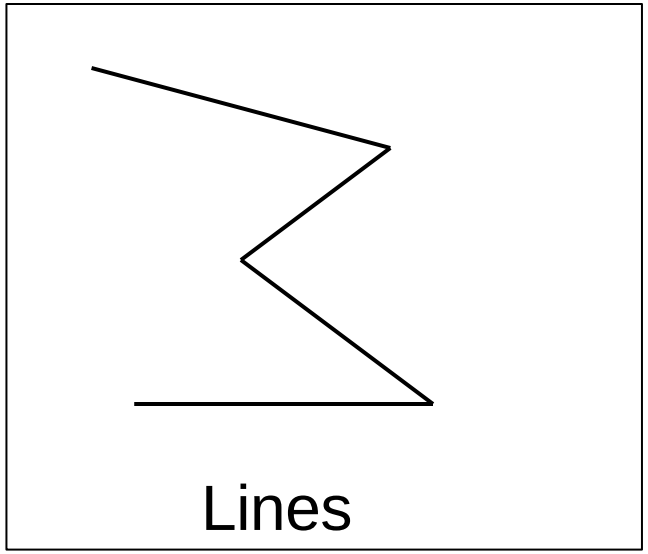
Spatial Data Models

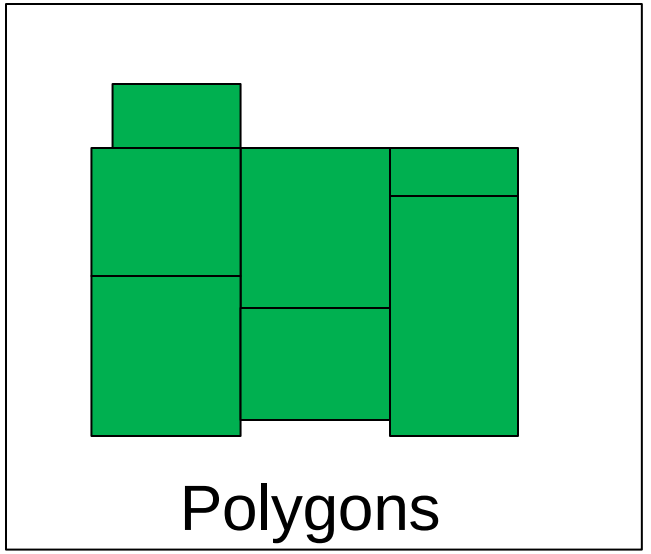
- Vector
- Raster



Vector Data Model





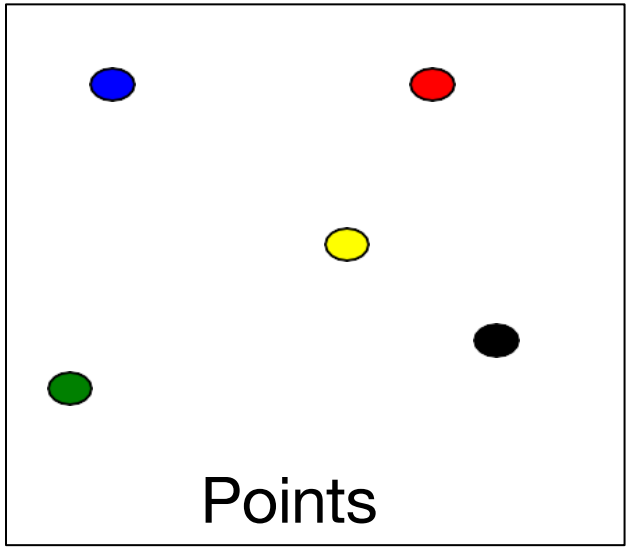


Spatial and Tabular Data

- Adds geometry to traditional data
  - Geometry: location and geometric characteristics of geographic (real-world) features
  - Attributes: data describing the characteristics of geographic features

Spatial and Tabular Data

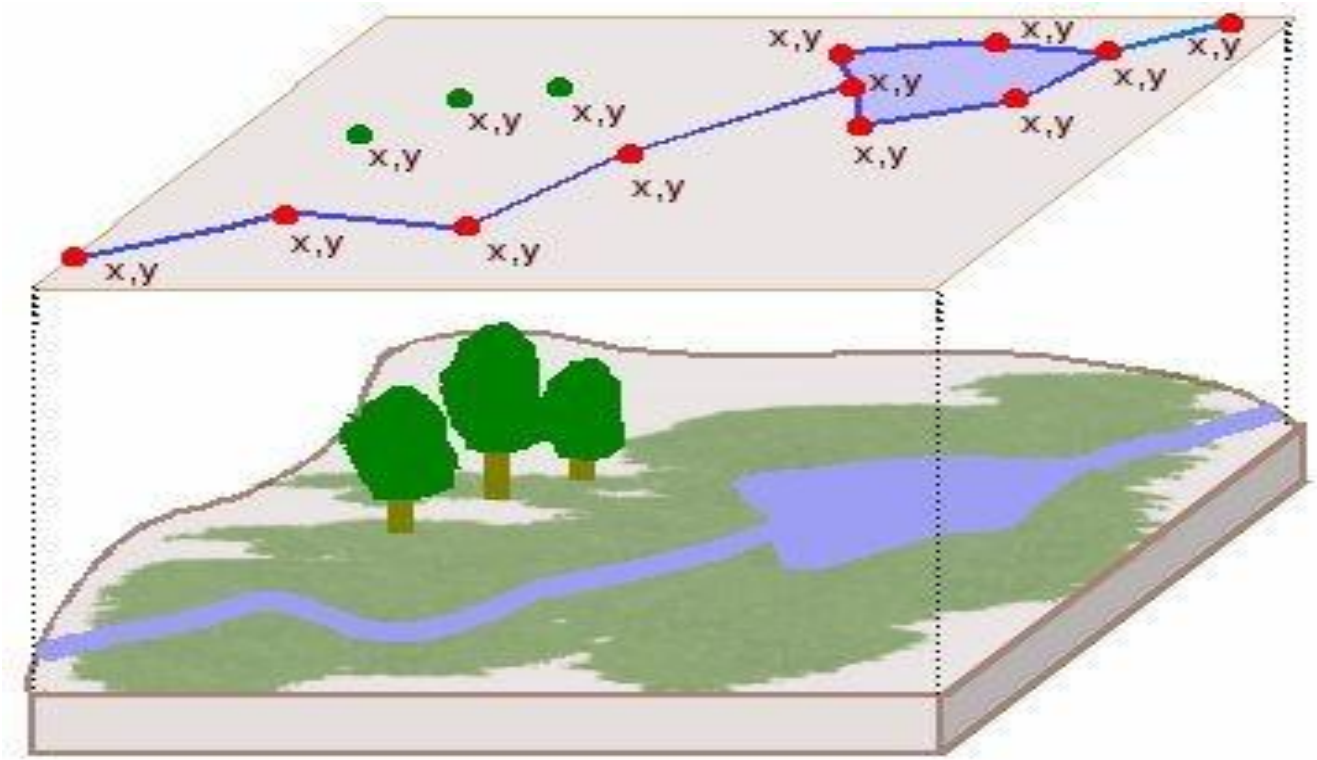
Geometry + Attributes



ID	**Color**	**Use**
1	Blue	Hospital
2	Red	Fire Dept.
3	Green	Office
4	Yellow	University
5	Black	Retail

Vector Model

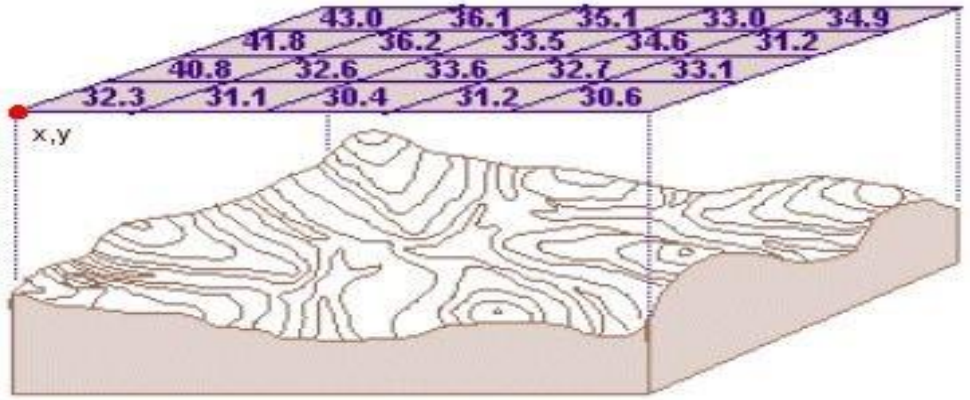
- Vector
  - Discrete entities defined by coordinate points
  - Three types of vector data
    - Point
    - Line
    - Polygon



Source: ESRI Virtual Campus

Raster Model

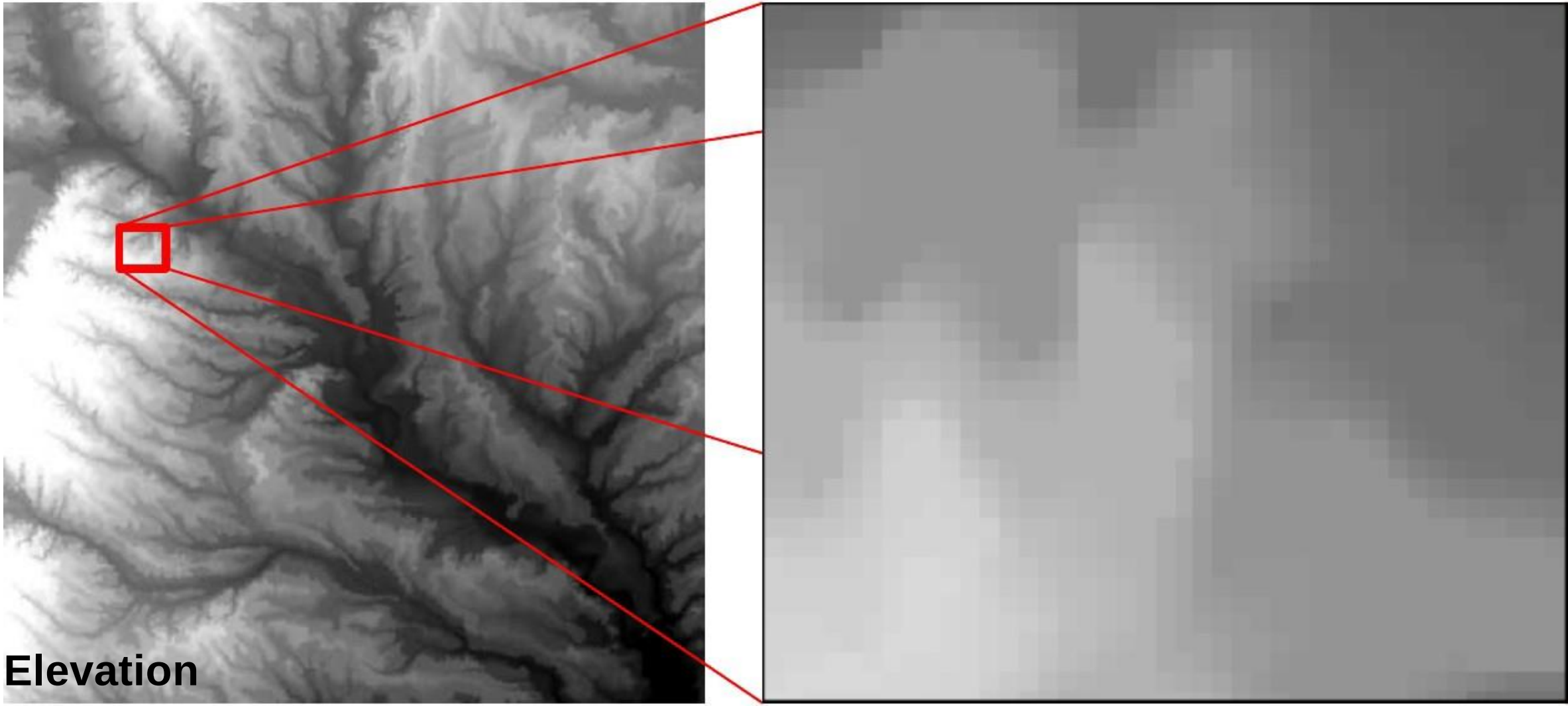
- Composed of a regular grid of cells
- Every grid cell has a value
- Every point on ground belongs to a grid cell
- Examples
  - Elevation
  - Crime hotspot
  - Temperature
  - Rainfall



AVERAGE ELEVATION

Raster Data Model

- Conventionally, stored row by row from the top left corner
- Attributes are recorded by assigning each cell a single value: e.g., landuse type
- Simple data structure
  - Directly store each layer as a single table
  - each layer is analogous to a ''spreadsheet'' or ''matrix''



- High : 262

- Low: 73

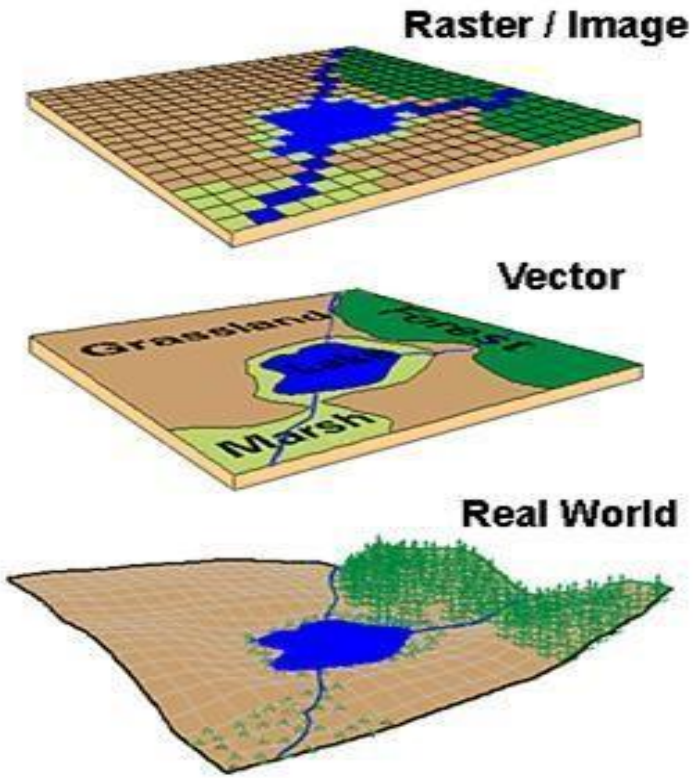
Elevation in Dallas county

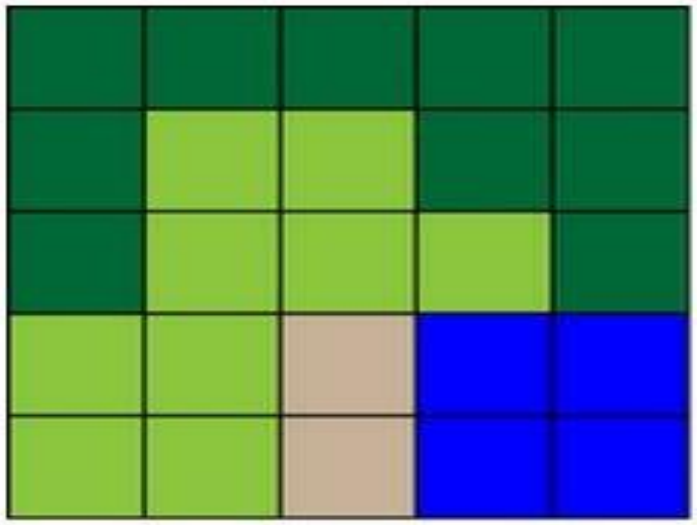
The whole county and a small area in thecounty

(Data Source: USGS)

Raster vs. Vector

- Different ways of displaying the “real world”





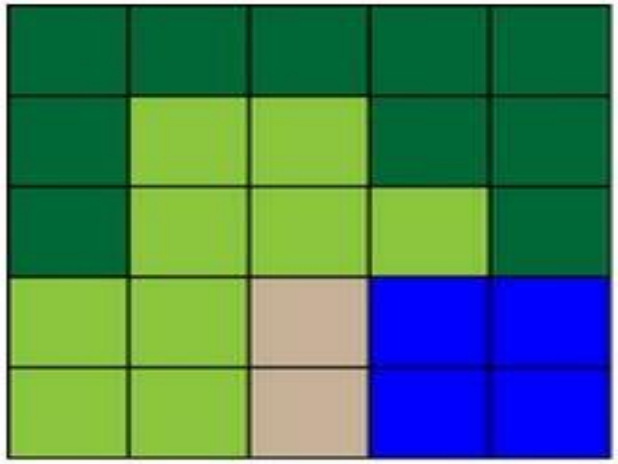
Raster

Vector

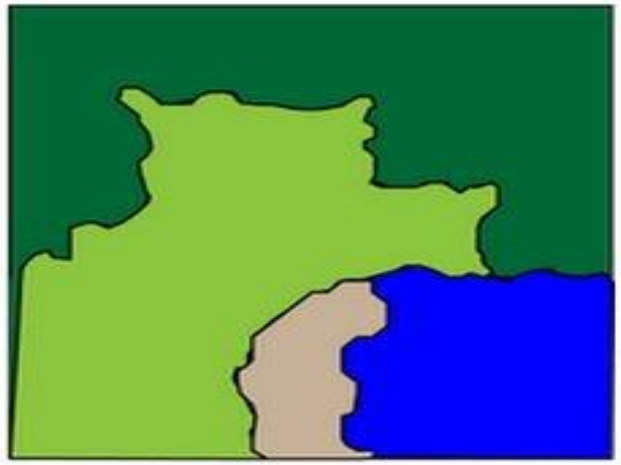
Source: University of Connecticut

Generalization

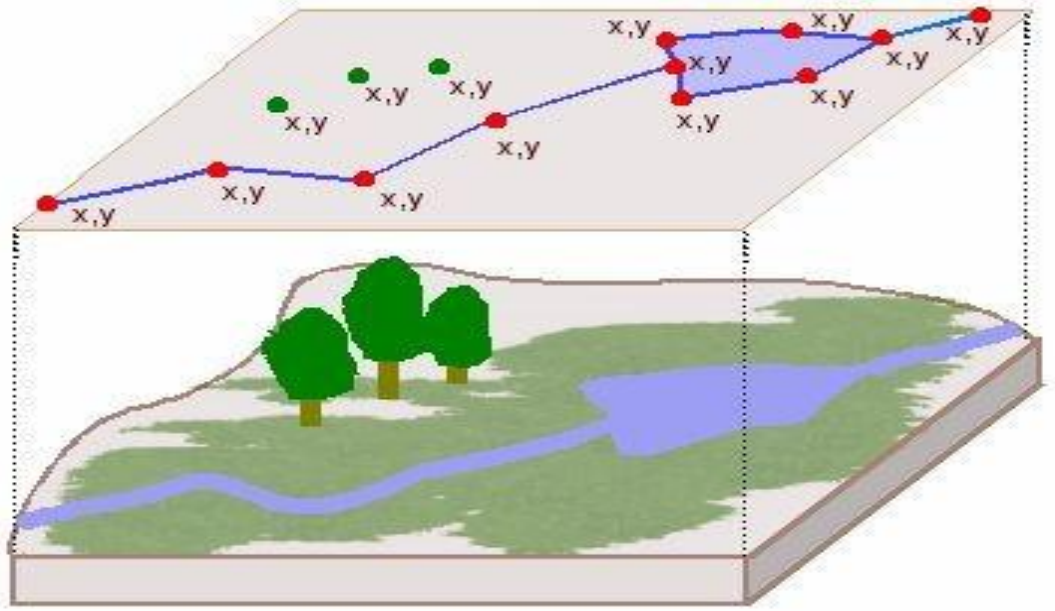
- The amount of detail given to an object



Raster



Vector

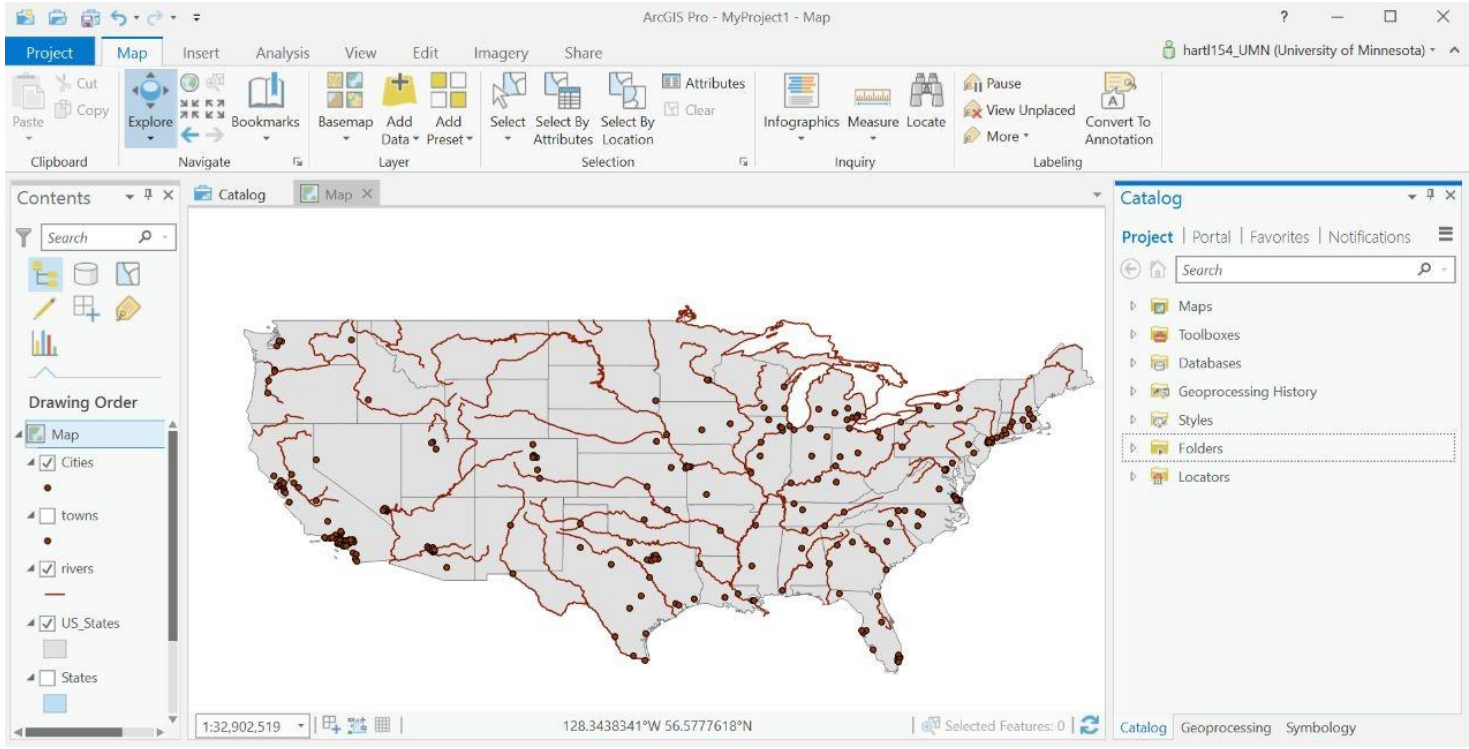


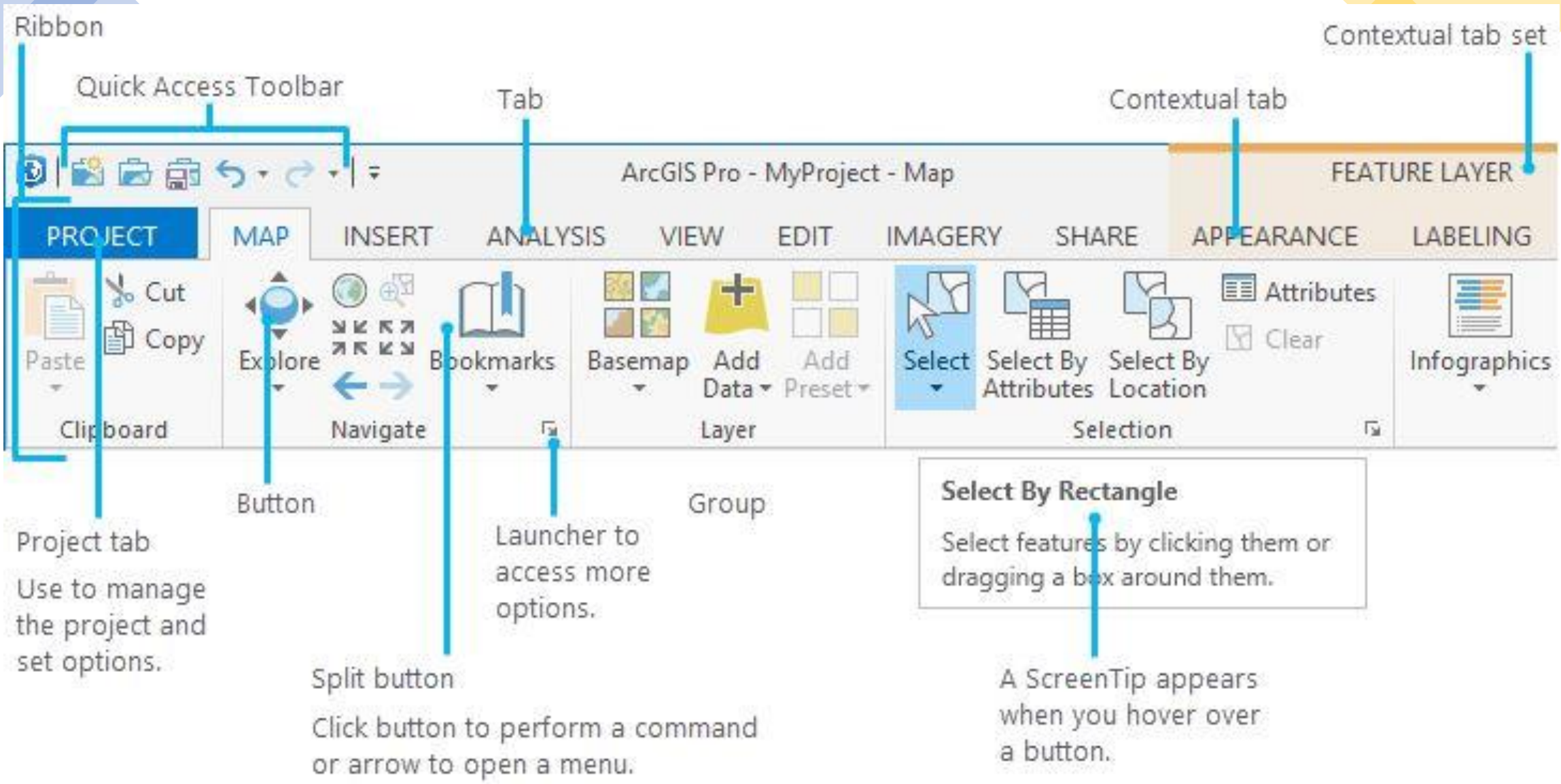
esri ArcGIS Pro

- Newest GIS software package offered by esri
- Currently included with ArcGIS Desktop
  - Licensing is different, uses ArcGIS Online for authentication
- 64-bit program
- Allows you to visualize, edit, and analyze geographic data in 2D and 3D
- More heavily integrated with ArcGIS Online
  - More options for sharing projects, access to ArcGIS Online geoprocessing tools

ArcGIS Pro

- Display GIS data
- Create maps
- Query data
- Analyze data
- Edit GIS data




Ribbon
Quick Access Toolbar
Tab
Contextual tab
Contextual tab set
ArcGIS Pro - MyProject - Map
FEATURE LAYER
PROJECT
MAP
INSERT
ANALYSIS
VIEW
EDIT
IMAGERY
SHARE
APPEARANCE
LABELING
Cut
Copy
Paste
Explore
Bookmarks
Basemap
Add Data
Add Preset
Select
Select By Attributes
Select By Location
Attributes
Clear
Infographics
Clipboard
Navigate
Layer
Selection
Project tab
Use to manage the project and set options.
Button
Launcher to access more options.
Group
Split button
Click button to perform a command or arrow to open a menu.
Select By Rectangle
Select features by clicking them or dragging a box around them.
A ScreenTip appears when you hover over a button.

Spatial data management

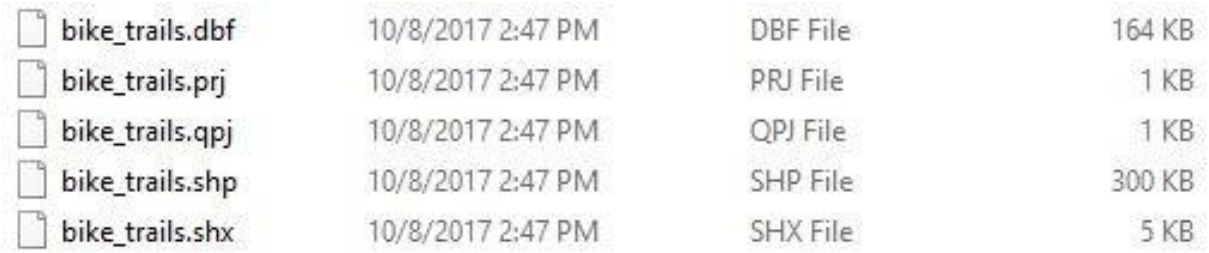
- Shapefiles
- Geodatabases
- Catalog Pane
- Metadata

Spatial Data: Shapefile

- Most common spatial data format in GIS
- Has been around for 40 years!
- Made up of points, lines, or polygons (**vector**)
- All GIS software will read shapefiles
- Used across all disciplines

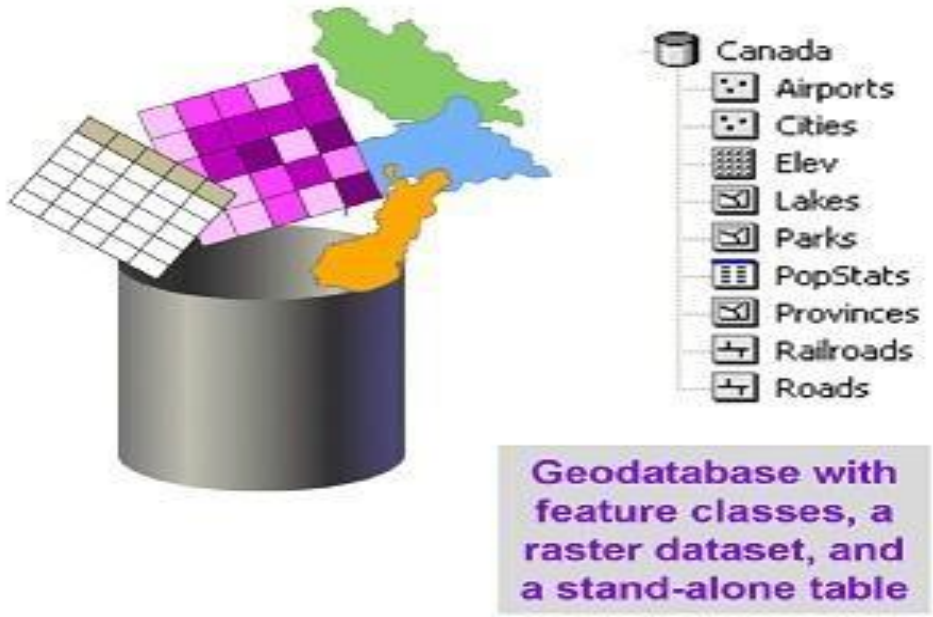
Spatial Data: Shapefile

- Single shapefile actually consists of multiple files
  - .shp – stores geometry
  - .dbf – stores attributes
  - .shx – index file
  - .prj – projection file
  - .xml – metadata file



Geodatabase

- Stores a set of files
- Also allows for data query, data management



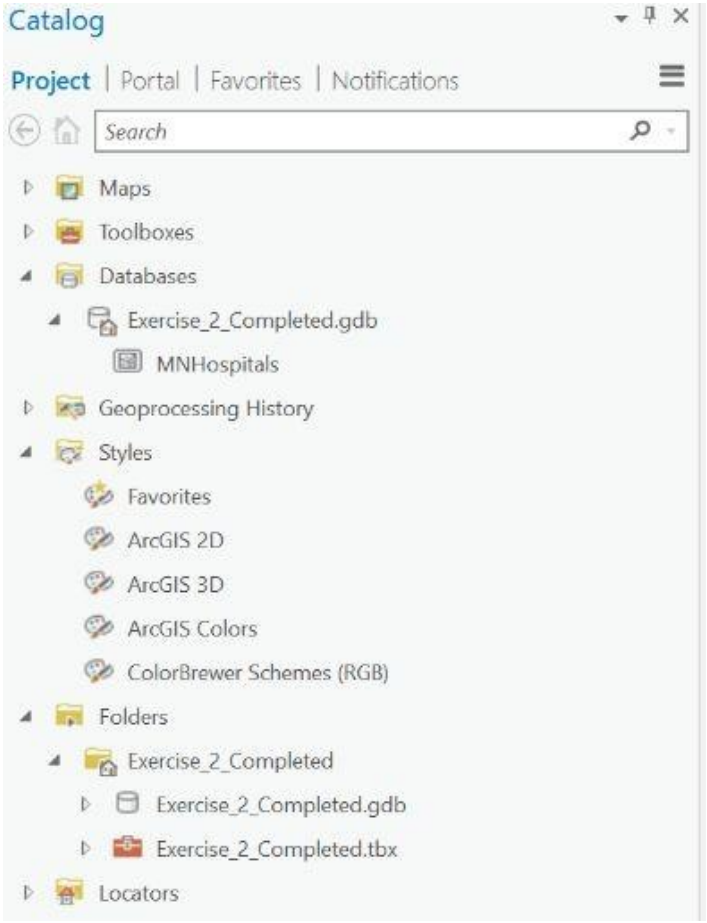
Feature Class

- Layer
- Grouping of one type of feature

  (i.e. points, lines, polygons)
- With spatial and attribute information for each feature
- A “shapefile” stored within a Geodatabase

Catalog Pane

- Primarily for data management
  - Browse GIS data
    - View geography, attributes, and metadata
  - Import/export GIS data



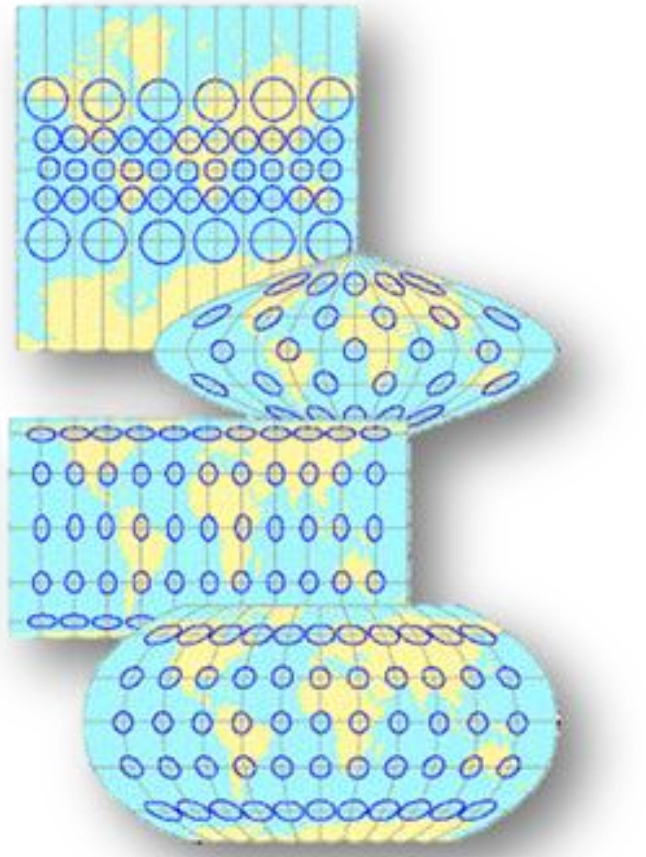
Metadata

- Data about data
- Describes the content, lineage, creator, distributor, processing steps, and spatial reference of the spatial data
- Helps users determine the availability and access requirements for data
- Helps users judge the quality and “fitness-for-use” of the data for their particular application
- Results only as good as input data!
- “**View Metadata**” in ArcGIS Pro

Projection

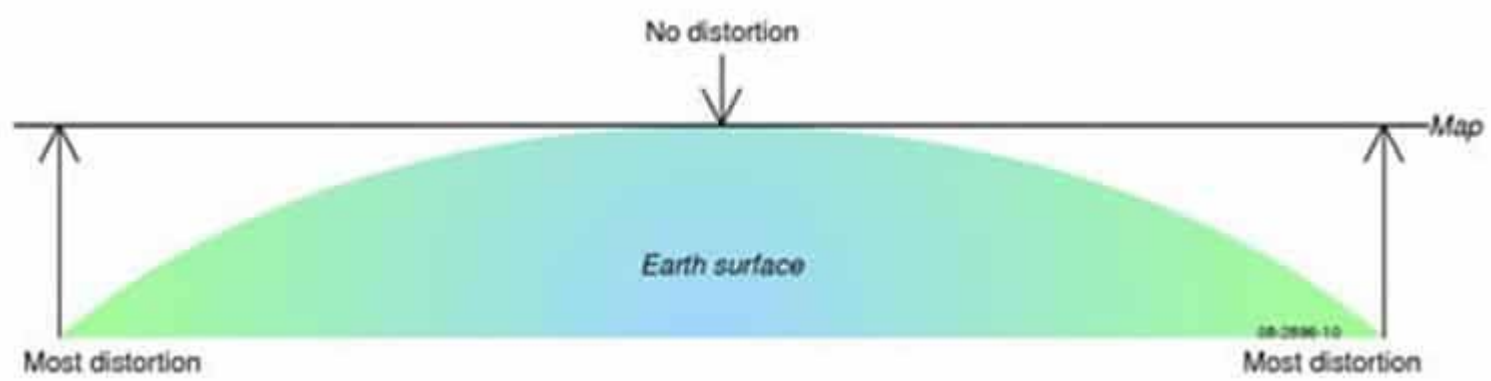
- Earth is spherical, maps are not
- Projections allow for “best” views of specific areas
- Will not be covered in detail but
  - Different Scales use different projections
  - States have their own projections
  - Counties have their own projections

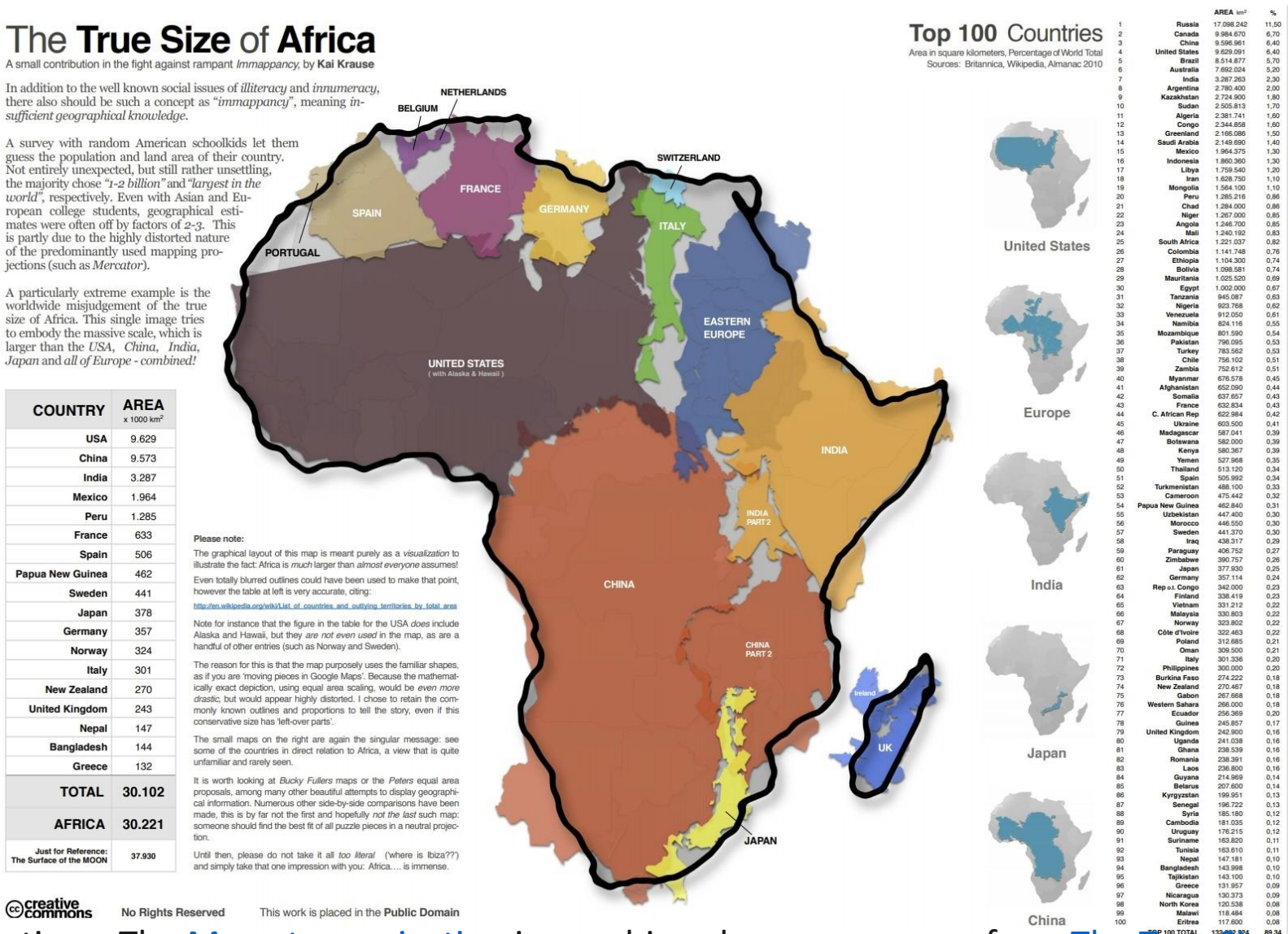
Projections



projection distortion

- The Earth is an ellipsoid whose larger radius is located at the Equator. If the jurisdiction of the owner is small enough, the planar assumption is valid.





COUNTRY	AREA x 1000 km²
USA	9.629
China	9.573
India	3.287
Mexico	1.964
Peru	1.285
France	633
Spain	506
Papua New Guinea	462
Sweden	441
Japan	378
Germany	357
Norway	324
Italy	301
New Zealand	270
United Kingdom	243
Nepal	147
Bangladesh	144
Greece	132
TOTAL	30.102
AFRICA	30.221
Just for Reference: The Surface of the MOON	37.930

Top 100 Countries

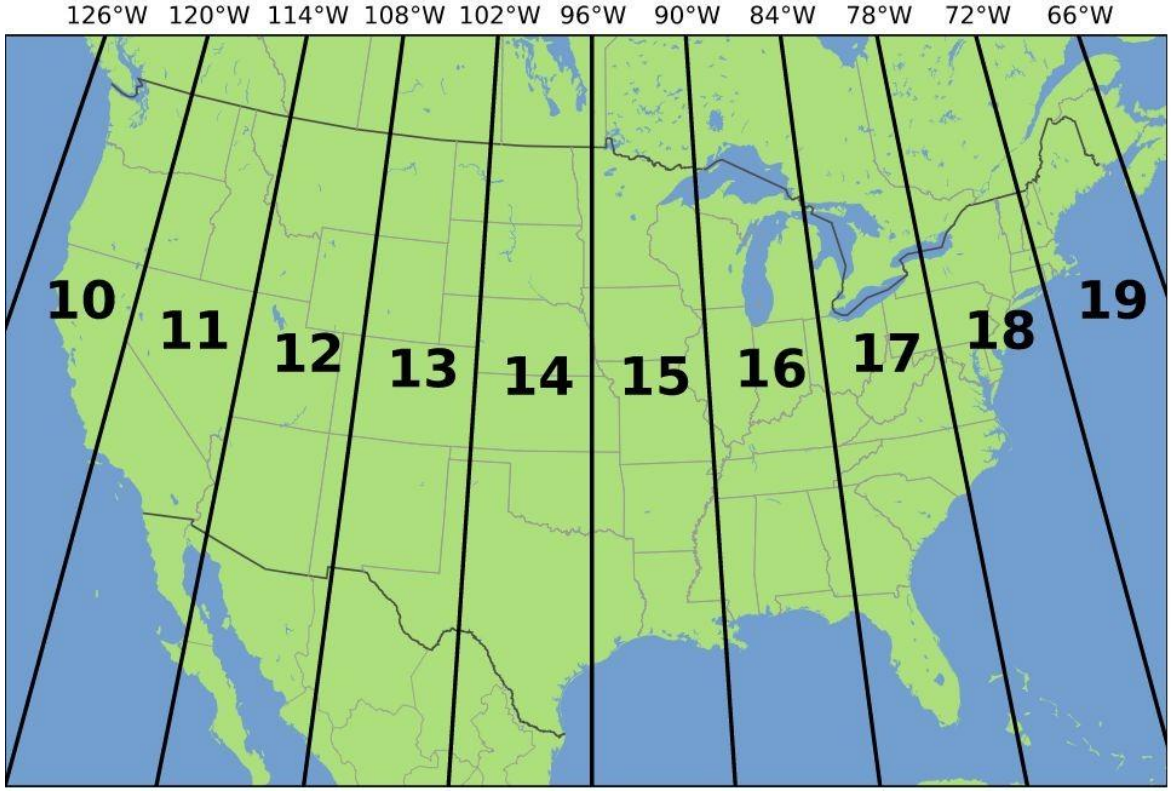
Area in square kilometers, Percentage of World Total
Sources: Britannica, Wikipedia, Almanac 2010

		AREA km²	%
1	Russia	17.098.242	11,50
2	Canada	9.984.670	6,70
3	China	9.596.961	6,40
4	United States	9.629.091	6,40
5	Brazil	8.514.877	5,70
6	Australia	7.692.024	5,20
7	India	3.287.263	2,30
8	Argentina	2.780.400	2,00
9	Kazakhstan	2.724.900	1,80
10	Sudan	2.505.813	1,70
11	Algeria	2.381.741	1,60
12	Congo	2.344.858	1,60
13	Greenland	2.166.086	1,50
14	Saudi Arabia	2.149.690	1,40
15	Mexico	1.964.375	1,30
16	Indonesia	1.860.360	1,30
17	Libya	1.759.540	1,20
18	Iran	1.628.750	1,10
19	Mongolia	1.564.100	1,10
20	Peru	1.285.216	0,86
21	Chad	1.284.000	0,86
22	Niger	1.267.000	0,85
23	Angola	1.246.700	0,85
24	Mali	1.240.192	0,83
25	South Africa	1.221.037	0,82
26	Colombia	1.141.748	0,76
27	Ethiopia	1.104.300	0,74
28	Bolivia	1.098.581	0,74
29	Mauritania	1.025.520	0,69
30	Egypt	1.002.000	0,67
31	Tanzania	945.087	0,63
32	Nigeria	923.768	0,62
33	Venezuela	912.050	0,61
34	Namibia	824.116	0,55
35	Mozambique	801.590	0,54
36	Pakistan	796.095	0,53
37	Turkey	783.562	0,53
38	Chile	756.102	0,51
39	Zambia	752.612	0,51
40	Myanmar	676.578	0,45
41	Afghanistan	652.090	0,44
42	Somalia	637.657	0,43
43	France	632.834	0,43
44	C. African Rep	622.984	0,42
45	Ukraine	603.500	0,41
46	Madagascar	587.041	0,39
47	Botswana	582.000	0,39
48	Kenya	580.367	0,39
49	Yemen	527.968	0,35
50	Thailand	513.120	0,34
51	Spain	505.992	0,34
52	Turkmenistan	488.100	0,33
53	Cameroon	475.442	0,32
54	Papua New Guinea	462.840	0,31
55	Uzbekistan	447.400	0,30
56	Morocco	446.550	0,30
57	Sweden	441.370	0,30
58	Iraq	438.317	0,29
59	Paraguay	406.752	0,27
60	Zimbabwe	390.757	0,26
61	Japan	377.930	0,25
62	Germany	357.114	0,24
63	Rep o.t. Congo	342.000	0,23
64	Finland	338.419	0,23
65	Vietnam	331.212	0,22
66	Malaysia	330.803	0,22
67	Norway	323.802	0,22
68	Côte d'Ivoire	322.463	0,22
69	Poland	312.685	0,21
70	Oman	309.500	0,21
71	Italy	301.336	0,20
72	Philippines	300.000	0,20
73	Burkina Faso	274.222	0,18
74	New Zealand	270.467	0,18
75	Gabon	267.668	0,18
76	Western Sahara	266.000	0,18
77	Ecuador	256.369	0,20
78	Guinea	245.857	0,17
79	United Kingdom	242.900	0,16
80	Uganda	241.038	0,16
81	Ghana	238.539	0,16
82	Romania	238.391	0,16
83	Laos	236.800	0,16
84	Guyana	214.969	0,14
85	Belarus	207.600	0,14
86	Kyrgyzstan	199.951	0,13
87	Senegal	196.722	0,13
88	Syria	185.180	0,12
89	Cambodia	181.035	0,12
90	Uruguay	176.215	0,12
91	Suriname	163.820	0,11
92	Tunisia	163.610	0,11
93	Nepal	147.181	0,10
94	Bangladesh	143.998	0,10
95	Tajikistan	143.100	0,10
96	Greece	131.957	0,09
97	Nicaragua	130.373	0,09
98	North Korea	120.538	0,08
99	Malawi	118.484	0,08
100	Eritrea	117.600	0,08
	TOP 100 TOTAL	132.082.924	89,34

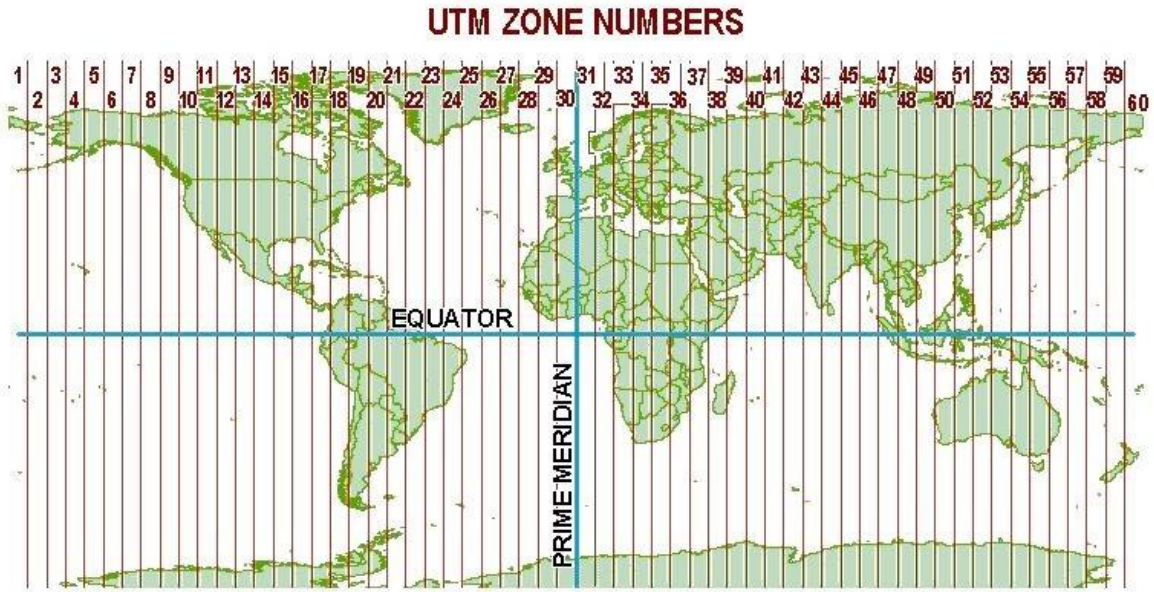
Different projections: The [Mercator projection](#) is very biased, as you can see from [The True Size of Africa](#).

UTM Projections

UTM Zones for the United States



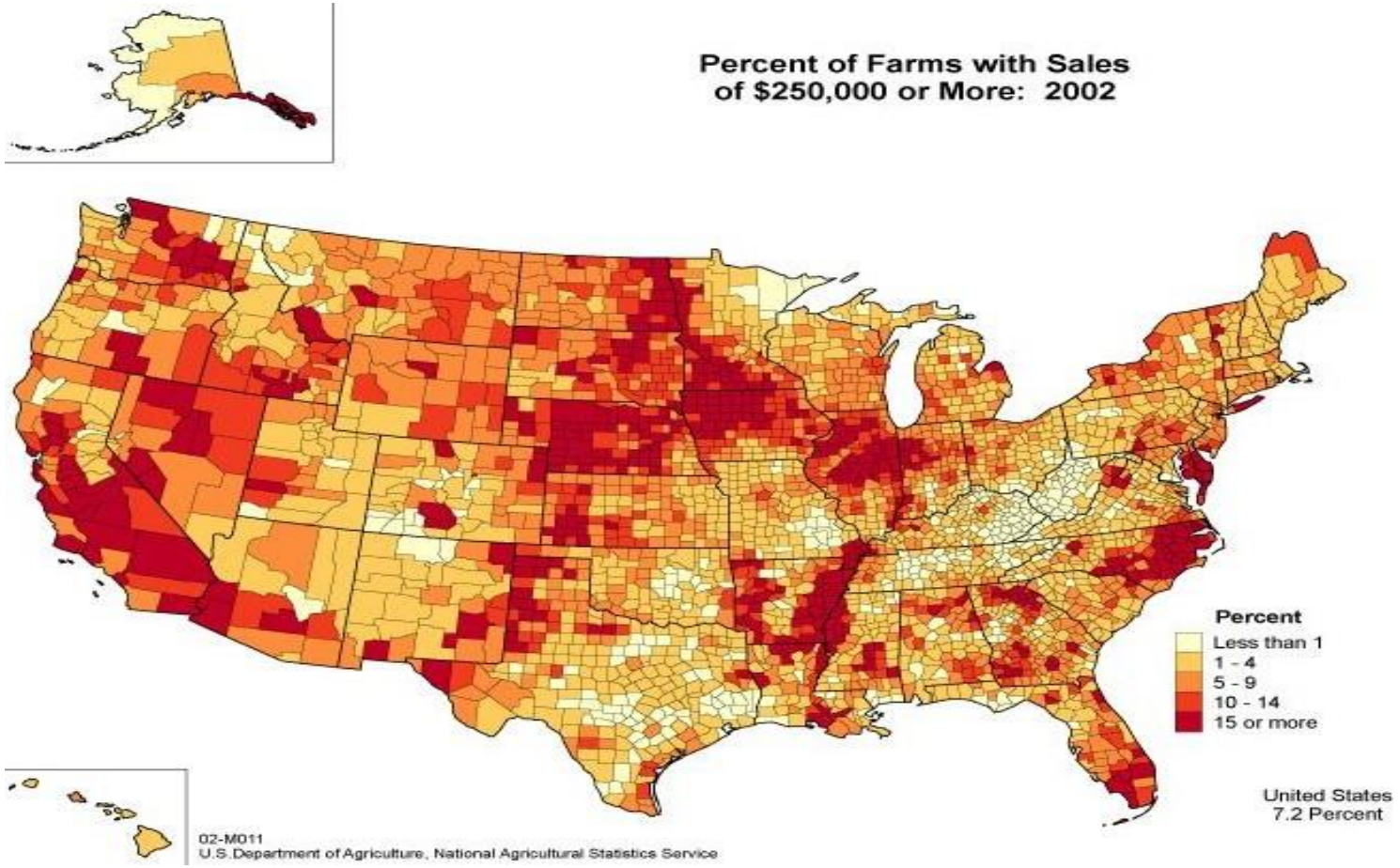
UTM Zones for the World



Types of Thematic Maps

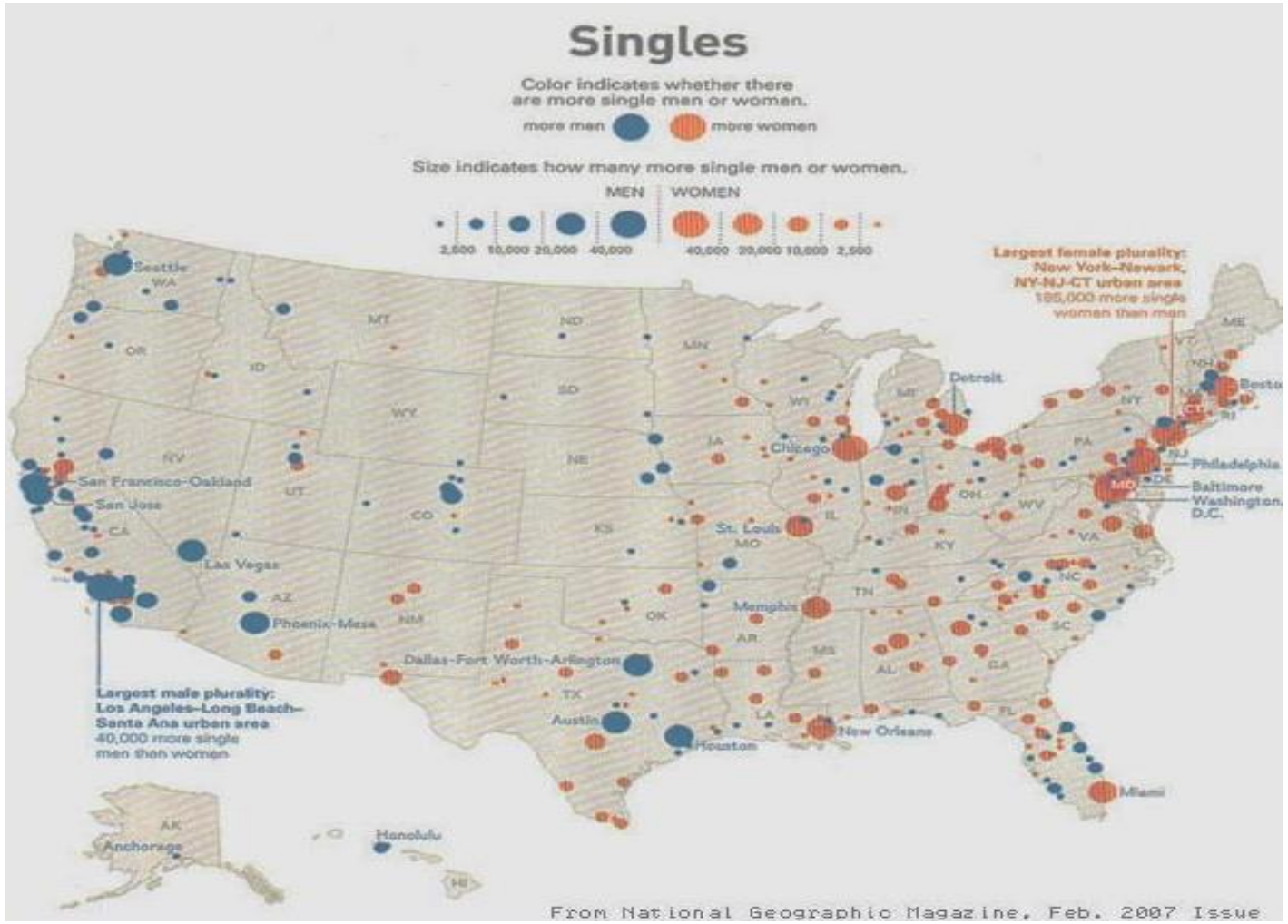
- Choropleth
- Graduated Symbol
- Hot Spot
- Isarithmic

Choropleth Map



Shows value per unit using colors. Use these for comparing relative numbers across space.

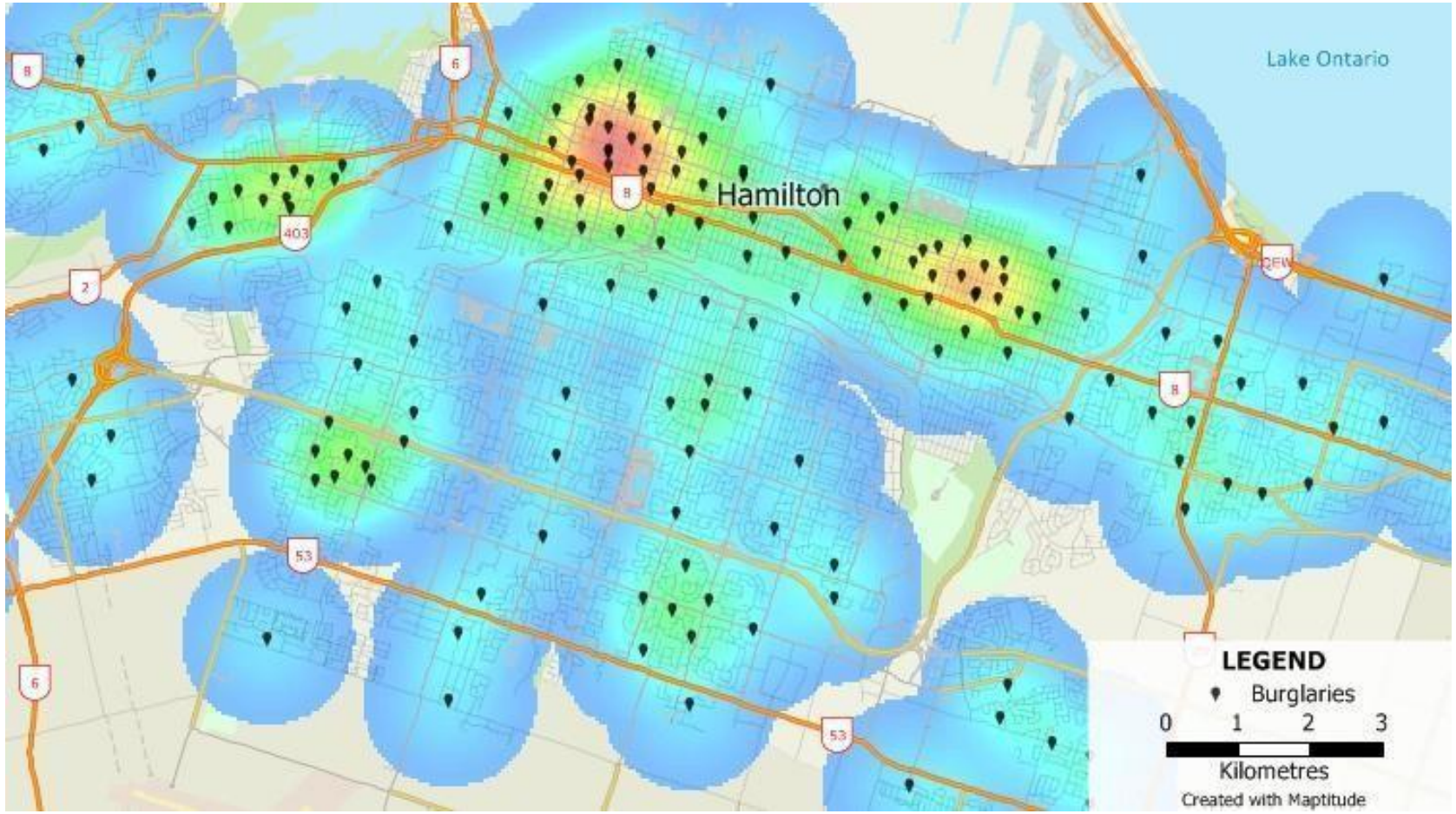
Proportional Symbol Map



The size of the symbol is representative of the value of the variable.

Difficult use well.

Hot Spot Map

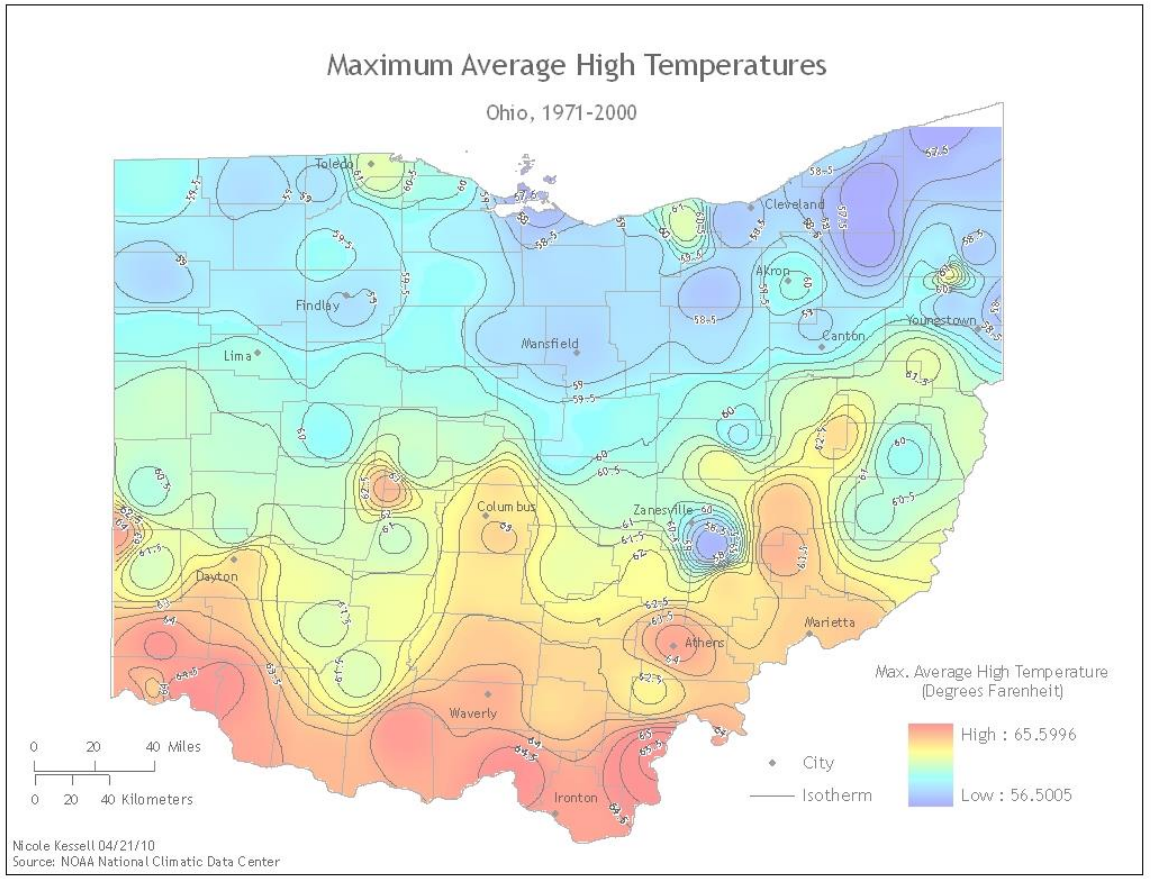


Shows density or clustering of phenomenon using color.

Statistically significant areas are displayed, excluding areas with no data.

Isarithmic Map

- Change in value is shown through intervals across map.

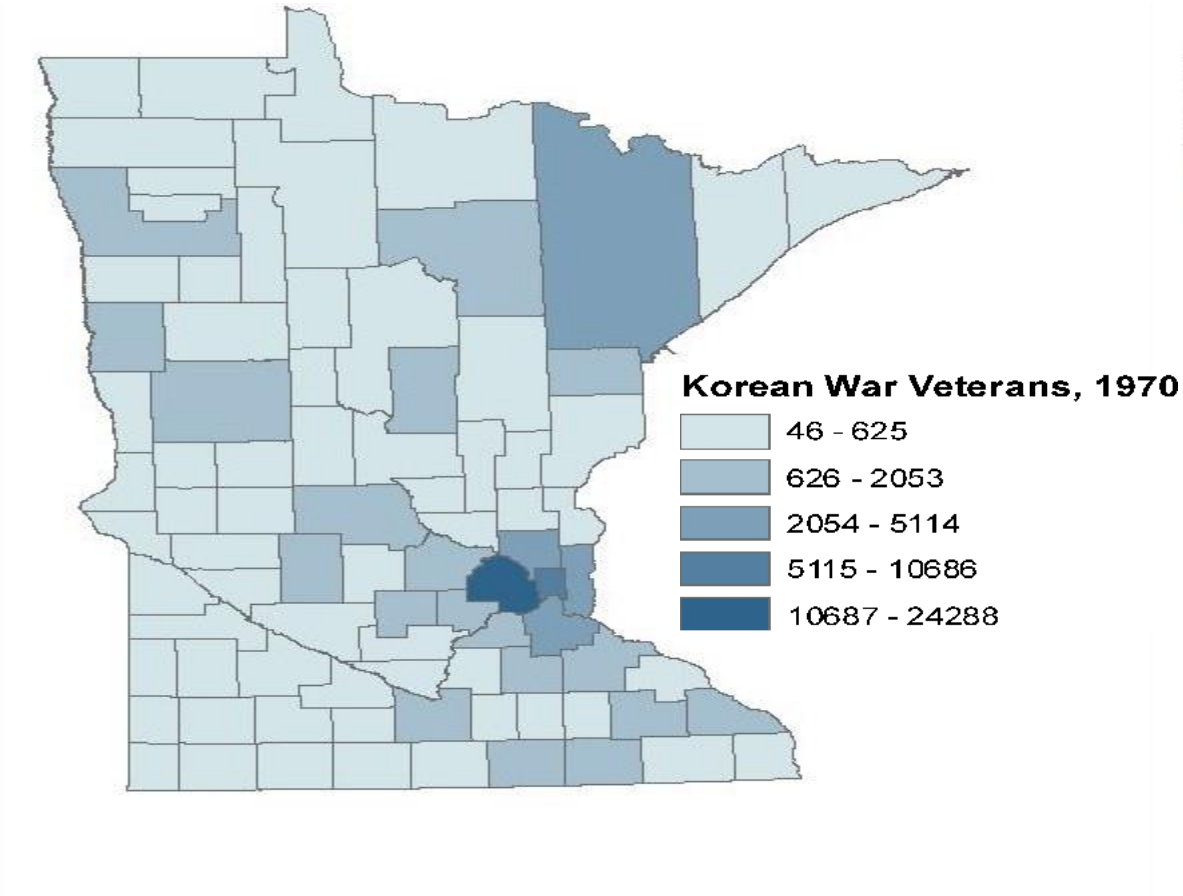


Data Standardization

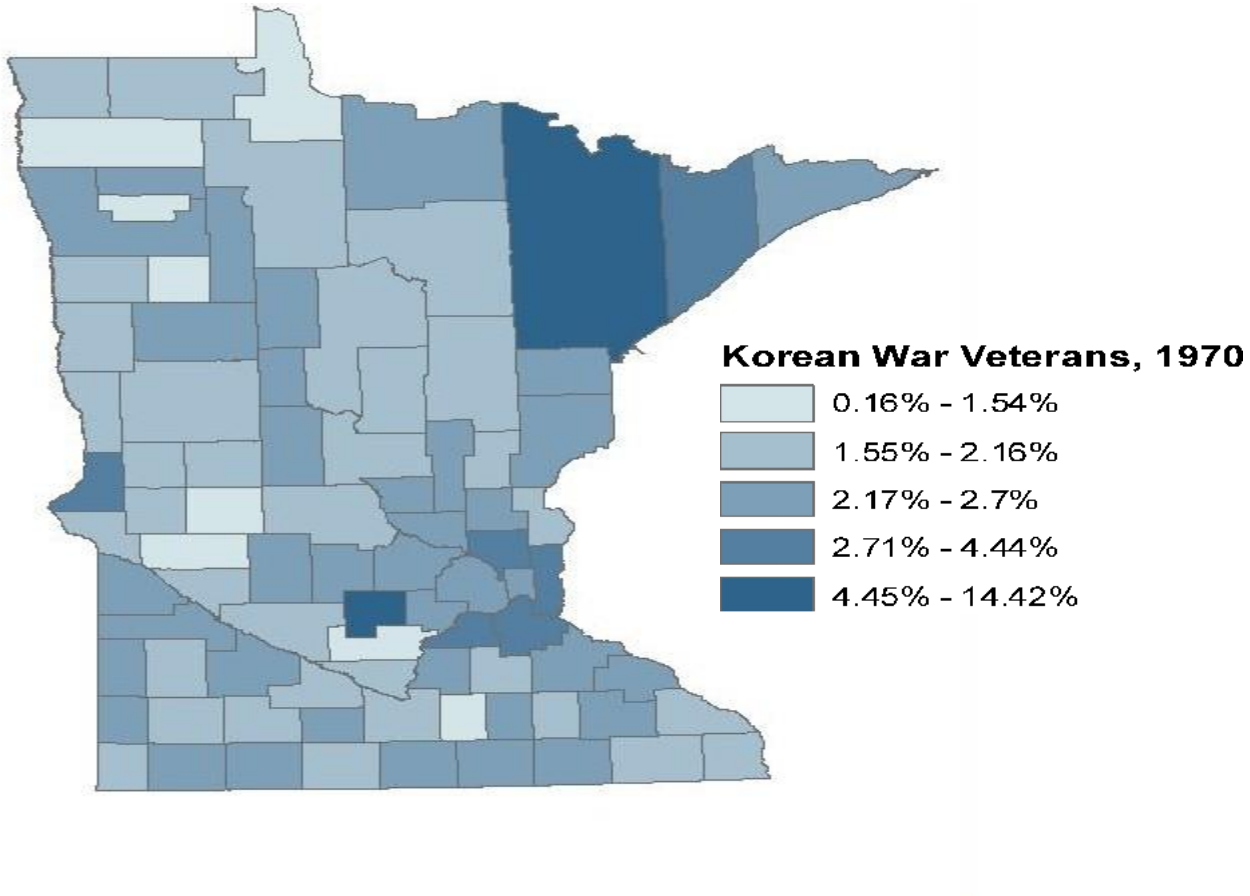
- Data should usually be standardized
    - Ability to compare areas
- By Population
    - Per person, Percentages
- By Area
    - Density

Data Standardization by Pop

Unstandardized

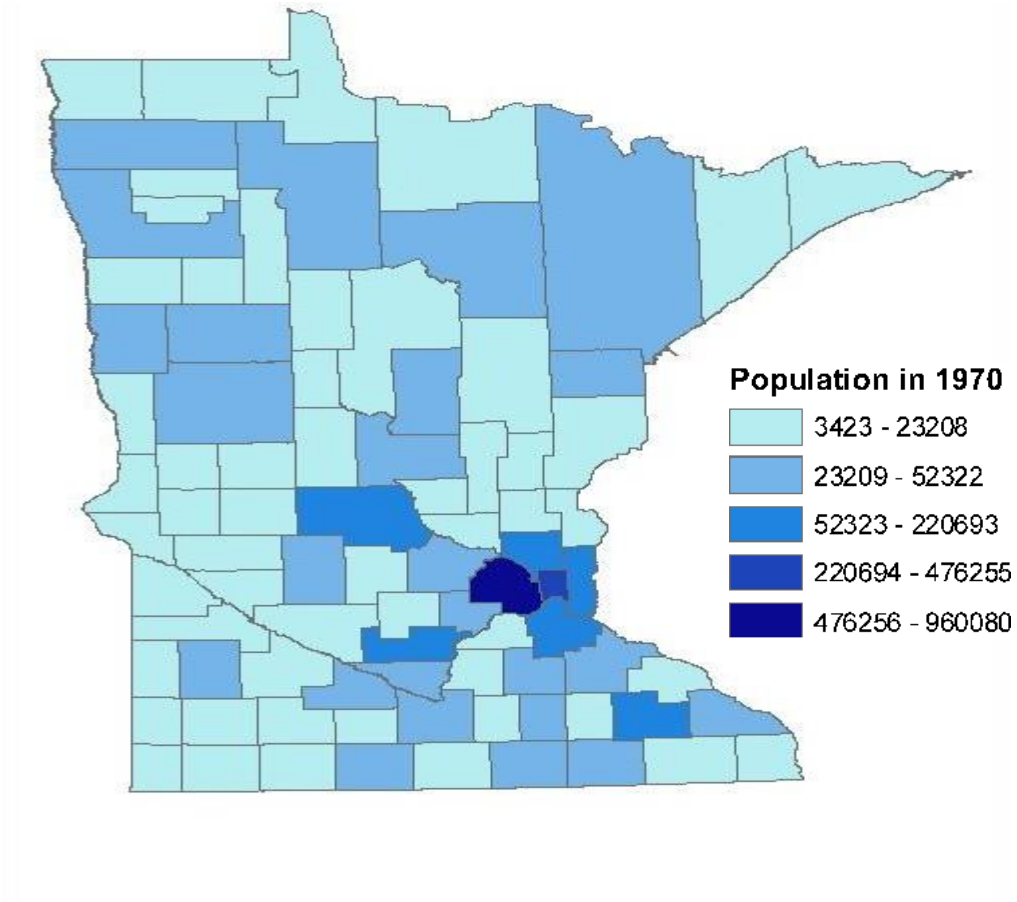


Standardized/Normalized

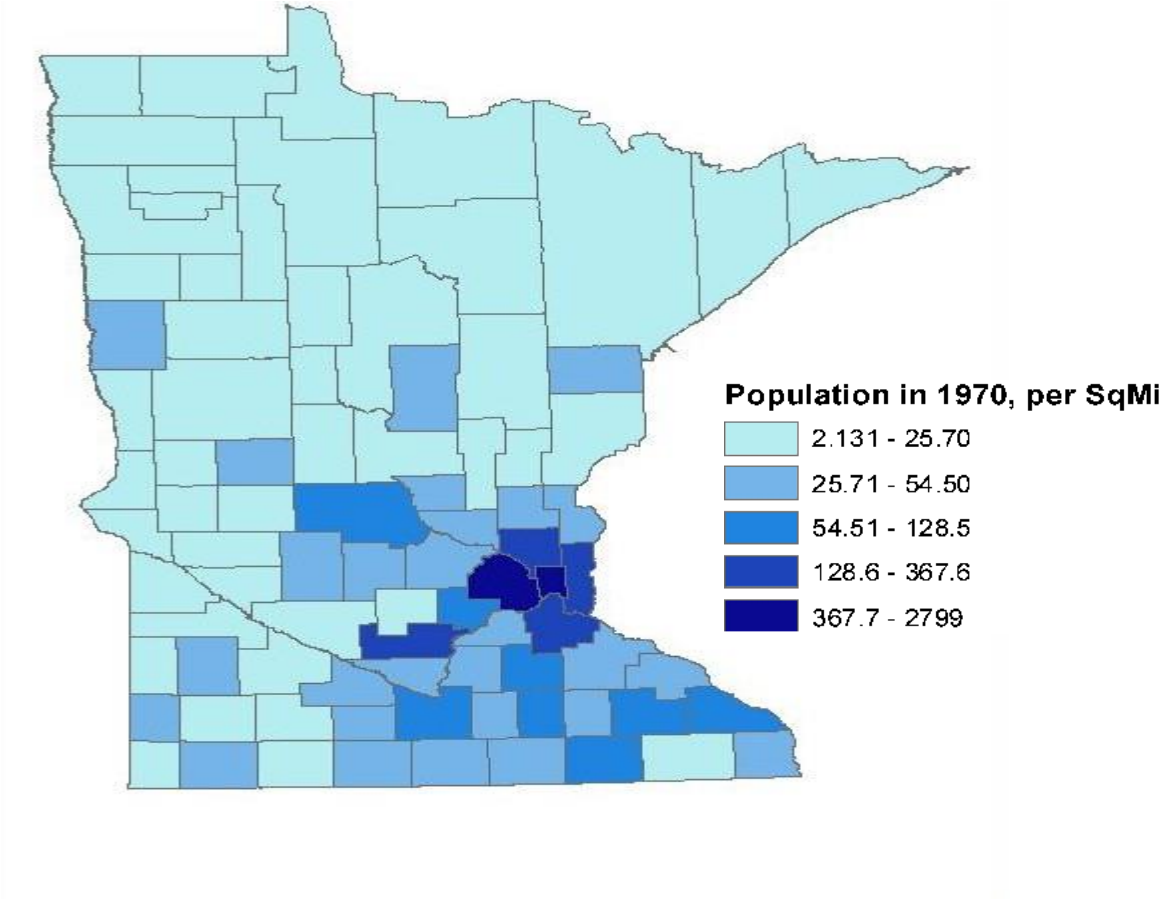


Data Standardization by Area

Unstandardized



Standardized/Normalized

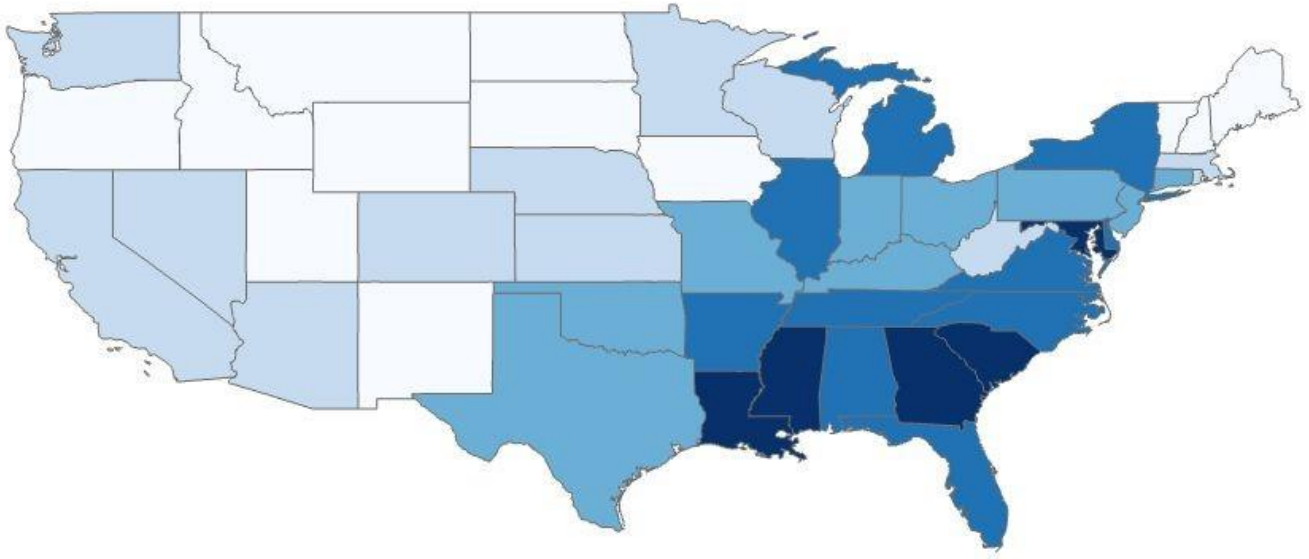


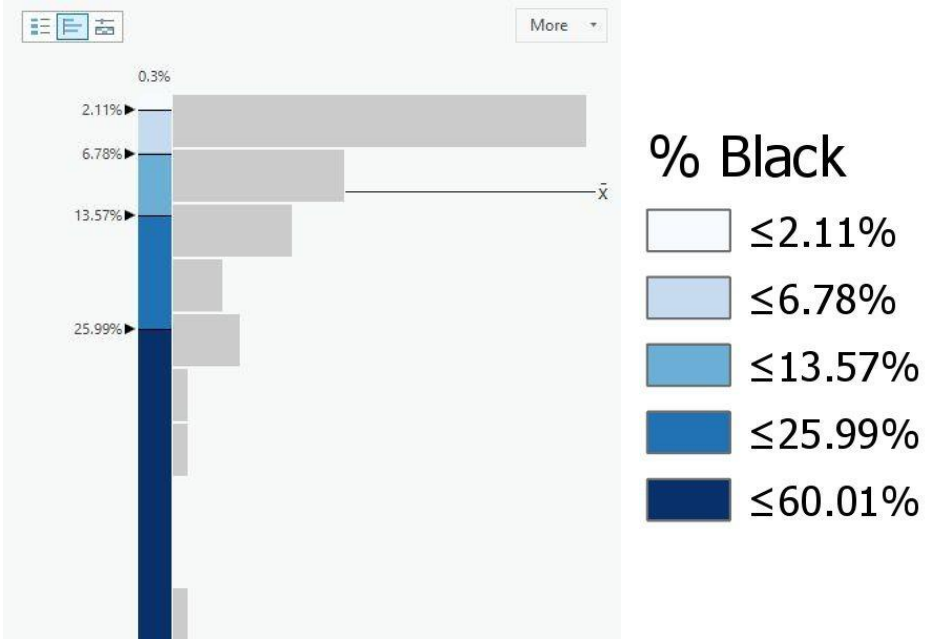
Classification

- Natural Breaks/Jenks
- Quantile
- Equal Interval Based on Range
- Equal Interval not Based on Range/Defined Interval

Natural Breaks (Jenks)

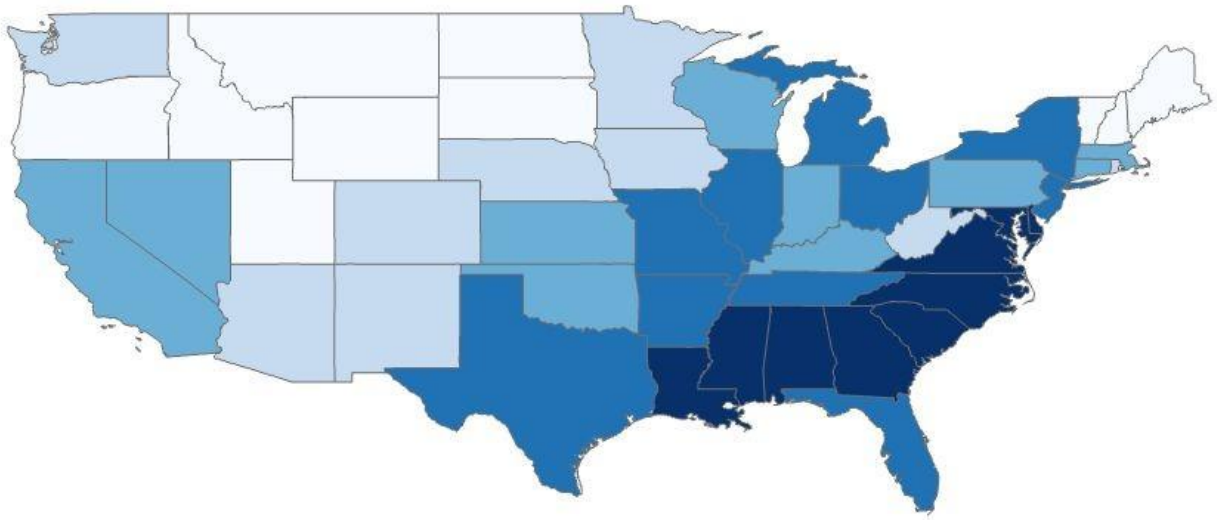
Natural Breaks has intervals that are created using natural clustering of the data. It maximizes variance between groups and minimizes variation within groups.

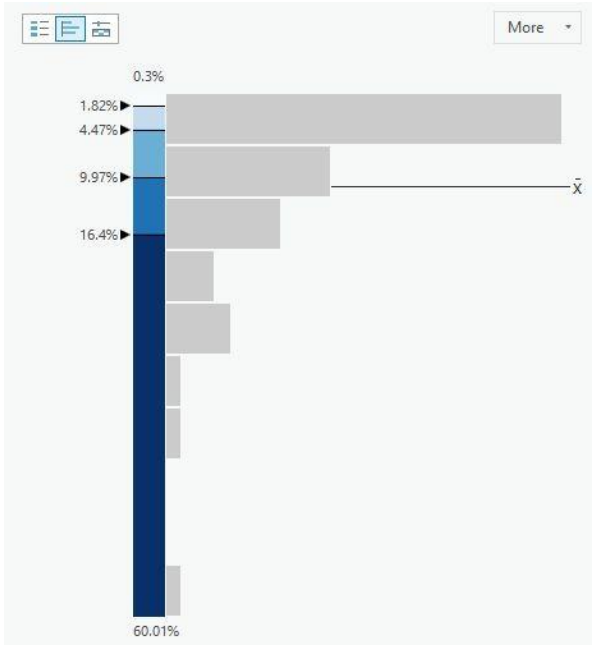




Quantile

Quantile has equal numbers of data in each class– sometimes called Quintile for 5 classes

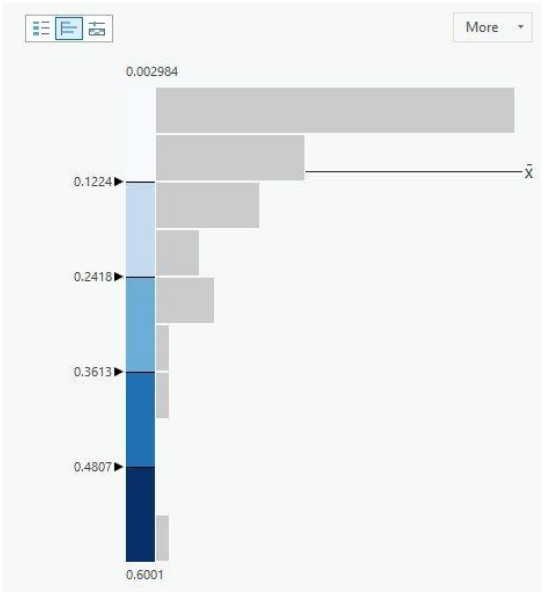




Equal Interval (Based on Range)

Equal-sized subranges

$$\frac{\text{Value of Highest Observation - Value of Lowest Observation}}{\text{Number of Classes}}$$

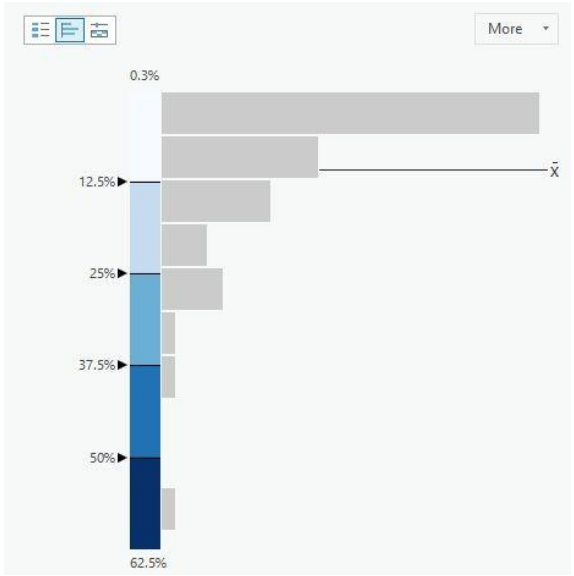


% Black

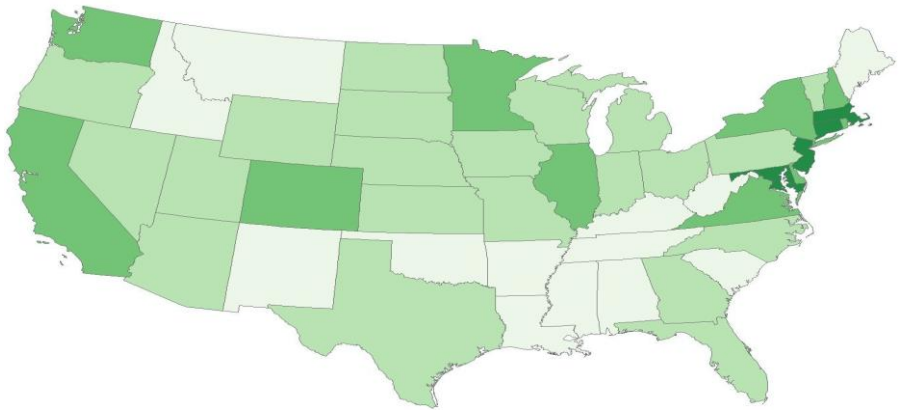
- ≤12.24%
- ≤24.18%
- ≤36.13%
- ≤48.07%
- ≤60.01%

Defined Interval

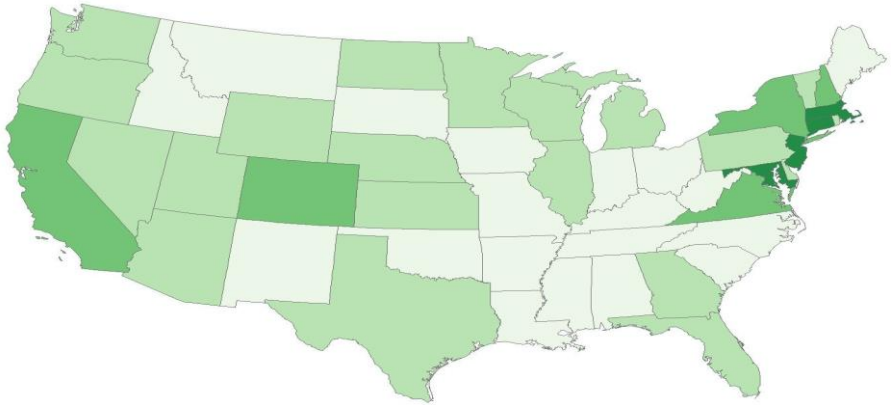
- Equal Interval not based on Range
- Good for comparing values across time
- Legends may be easier for audience to read



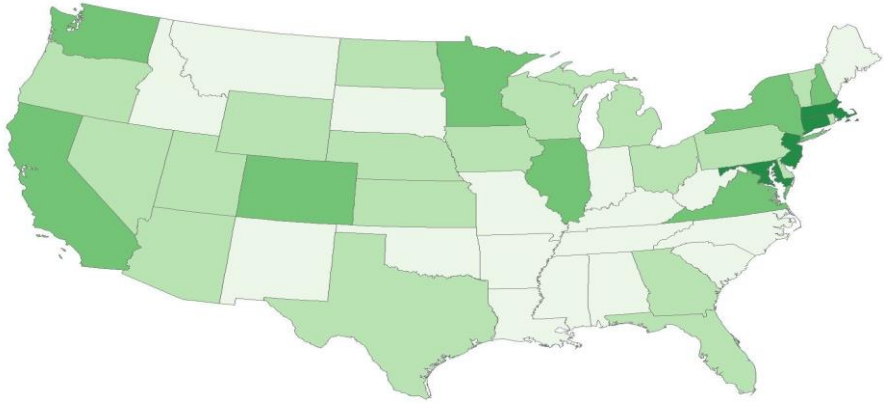
Classification: Method Comparison



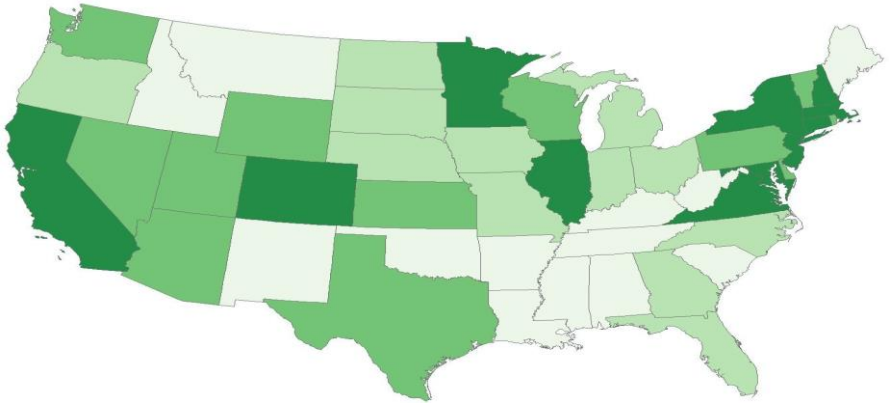
Natural Breaks



Equal Interval



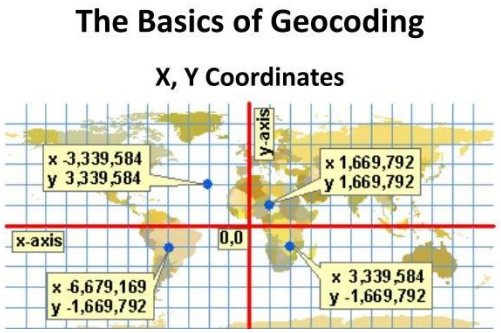
Standard Deviation



Quantile

Geocoding

- The process of finding the location of an address on a map.
- The location can be a pair of (X, Y) coordinate or a street address, postal delivery location, or building.
  - X – longitude (+ for E hemisphere; - for W hemisphere)
  - Y – latitude (+ for N hemisphere; - for S hemisphere)
- In GIS, geocoding requires a reference dataset that contains address attributes for the geographic features in the area of interest.



Address Matching

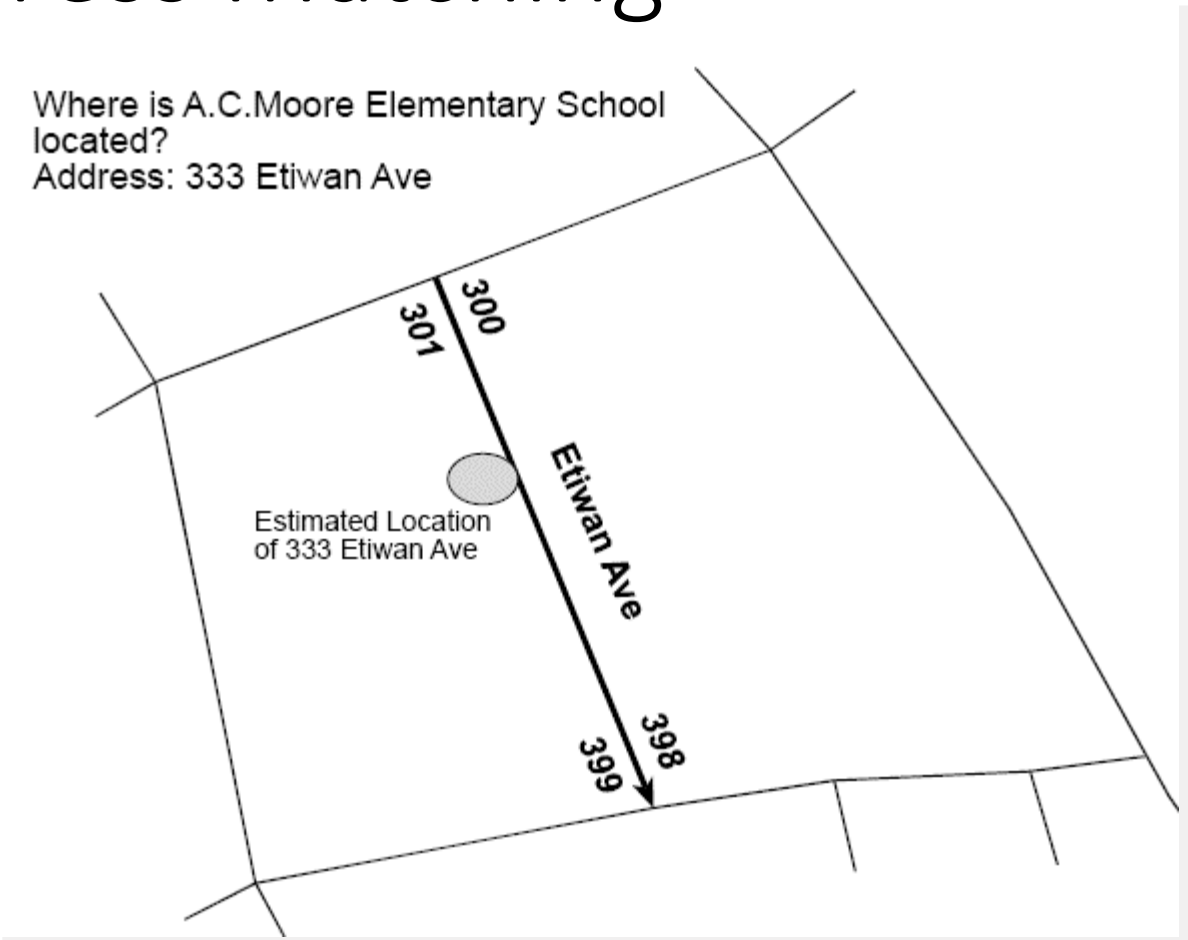
Address matching is the process of assigning an actual address to a geographic location on some reference files.

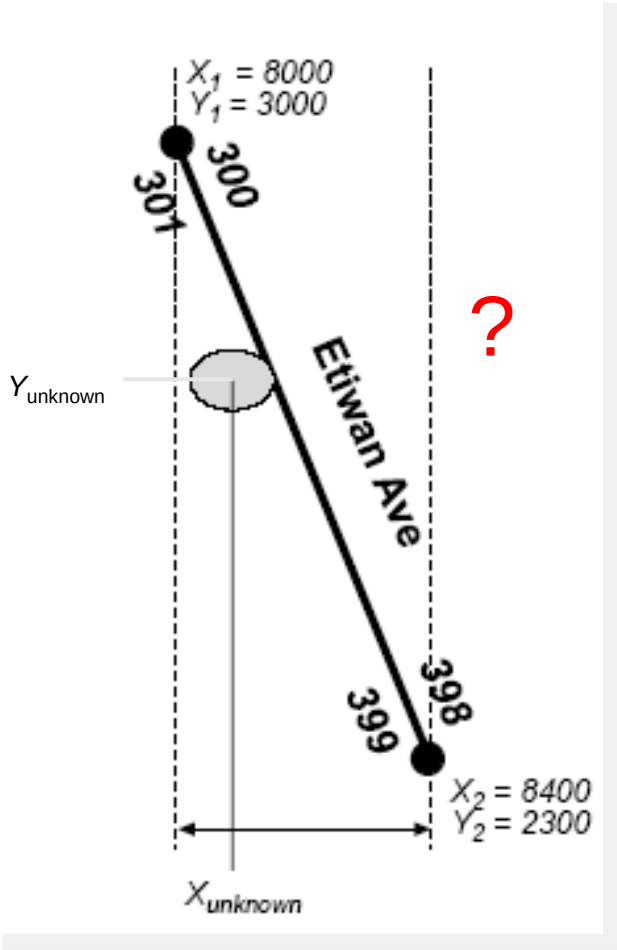
If an address falls within a feature's address range, it is considered a match and a location can be returned.

Address database

Reference map

Address Matching





- --- Both X and Y need to be interpolated allocate the address.

RCC-GIS Geocoding Service: Formatting Data for Processing

https://gis.rcc.uchicago.edu/content/rcc-gis-geocoding-service

Based on ESRI world Geocoder

Acceptable headers:

ID
ADDRESS
NEIGHBORHOOD
CITY
SUBREGION
REGION or STATE or ST
POSTAL or ZIP or ZIP CODE
COUNTRYCODE

Geocoding Review

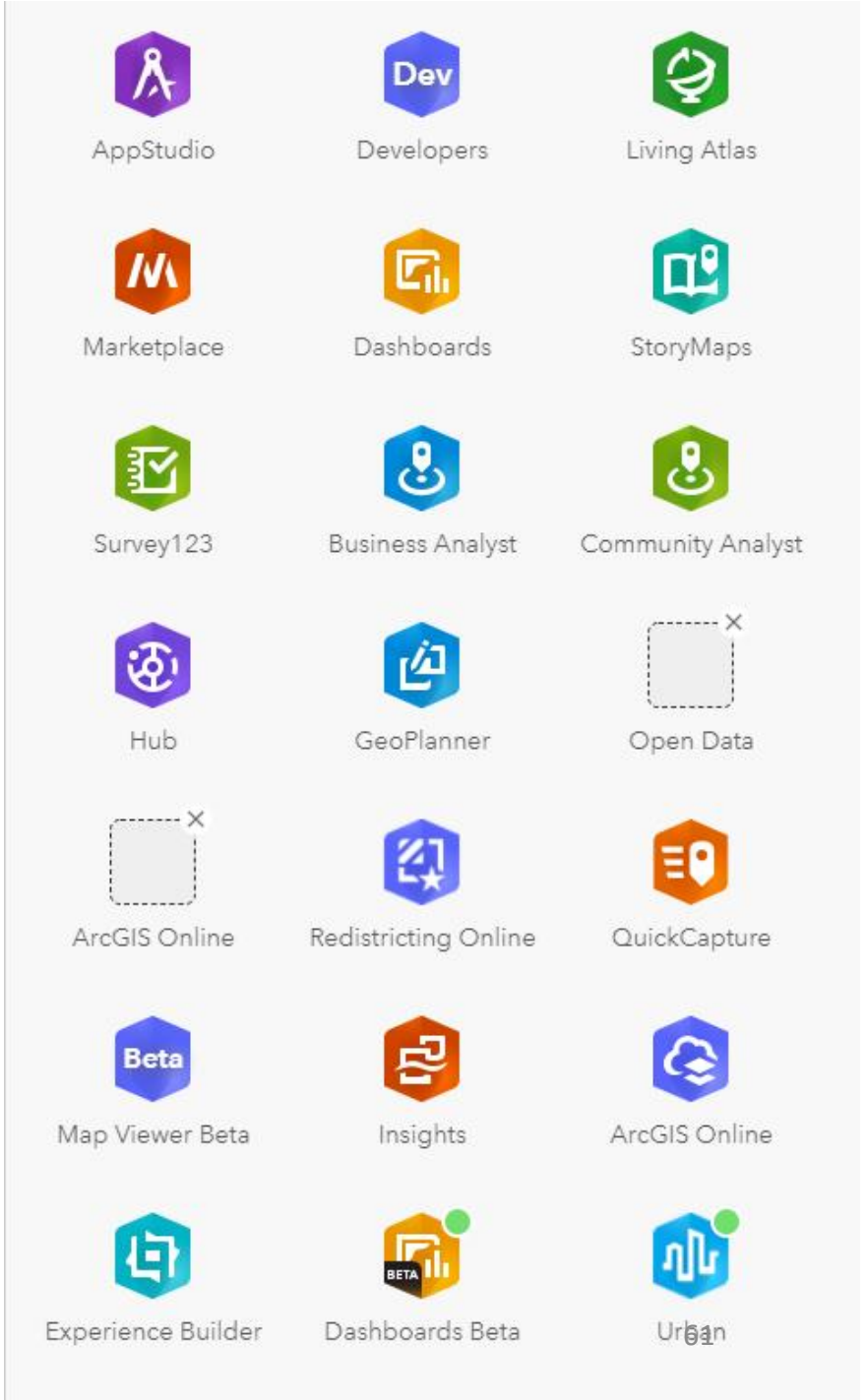
- Be careful which locator service you use, online geocoders are not HIPAA compliant
- Geocode to the appropriate geographic scale, can take care of confidentiality issues
- A high match score does not mean the point is accurate, best practice is to choose a small percentage of results to review
- Valid address does not necessarily mean correct location!

Spatial Analysis Examples

- Density of population around social services offices
- Demography around areas of high crime
- Site selection of interest
- Proposal for new bike path – how many people live within 1 mile of proposed path

ArcGIS Online

- https://uchicago.maps.arcgis.com



Using ArcGIS Online

- Wayback App
- https://livingatlas.arcgis.com/wayback

- Navigate to https://livingatlas.arcgis.com
- Sign in using CNETid
- Search “Global Imagery Browse Services”

esri ArcGIS Industries About Support

Parmanand Sinha
pnsinha_UChicago

ArcGIS Living Atlas of the World

Home | Browse | Apps | Blog | Contribute | My Contributions | My Favorites

Global Imagery Browse Services

Search examples

All | Trending | Basemaps | Imagery | Boundaries | People | Infrastructure | Environment

Filters: All content types | All time | All regions | Esri-only content

Sort by: Relevance

21 Results

MODIS True Color - Terra Surface Reflectance

Imagery Layer By: esri

This layer provides access to NASA Global Imagery Browse Services, which delivers global, full-resolution satellite imagery. This band composition (Bands 1 4 3) most accurately represents how we see the earth's surface with our own eyes.

Authoritative

MODIS True Color - Aqua Corrected Reflectance

Imagery Layer By: esri

This layer provides access to NASA Global Imagery Browse Services, which delivers global, full-resolution satellite imagery. This band composition (Bands 1 4 3) most accurately represents how we see the earth's surface with our own eyes.

Authoritative

MODIS True Color - Terra Corrected Reflectance

Imagery Layer By: esri

This layer provides access to NASA Global Imagery Browse Services, which delivers global, full-resolution satellite imagery. This band composition (Bands 1 4 3) most accurately represents how we see the earth's surface with our own eyes.

Daily Planet Imagery

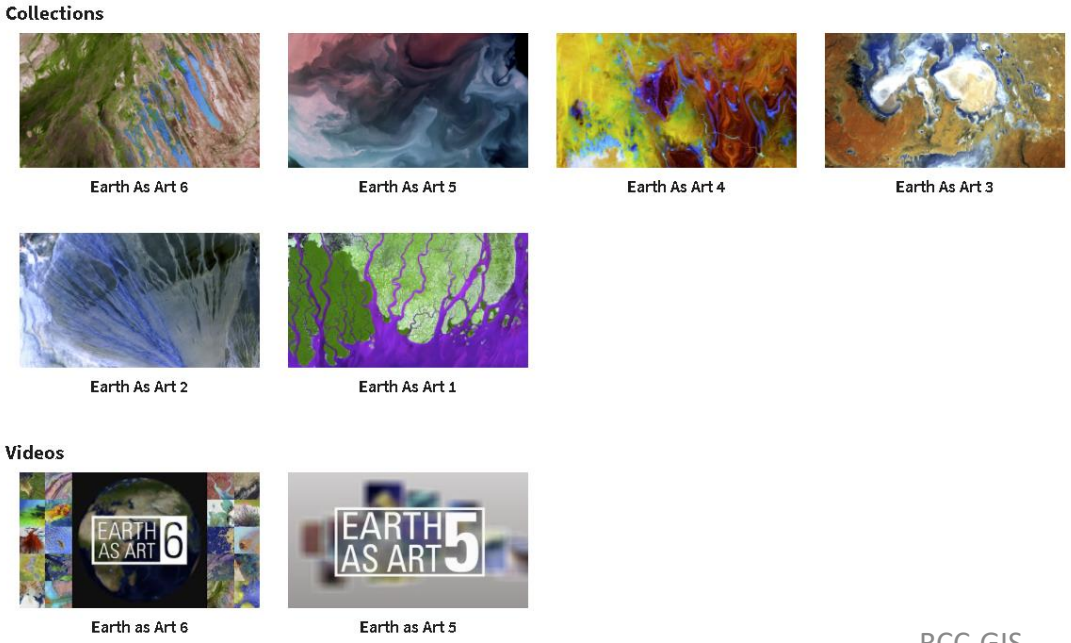
Web Map By: esri

This map shows imagery for the planet that is updated on a daily basis. It features the NASA MODIS imagery True Color band composition (Bands 1 4 3 | Red, Green, Blue) which most accurately shows how we see the earth's surface with our own eyes.

Earth As Art

- “In addition to their scientific value, many satellite images are simply intriguing to look at. Satellites capture an incredible variety of views of Earth. See the mesmerizing beauty of river deltas, mountains, and other sandy, salty, and icy landscapes. Some might even remind you of actual famous works of art!”

USGS - **Earth As Art**



Map Puzzle

https://www.indianaview.org/image_puzzle.html

Questions

Acknowledgement

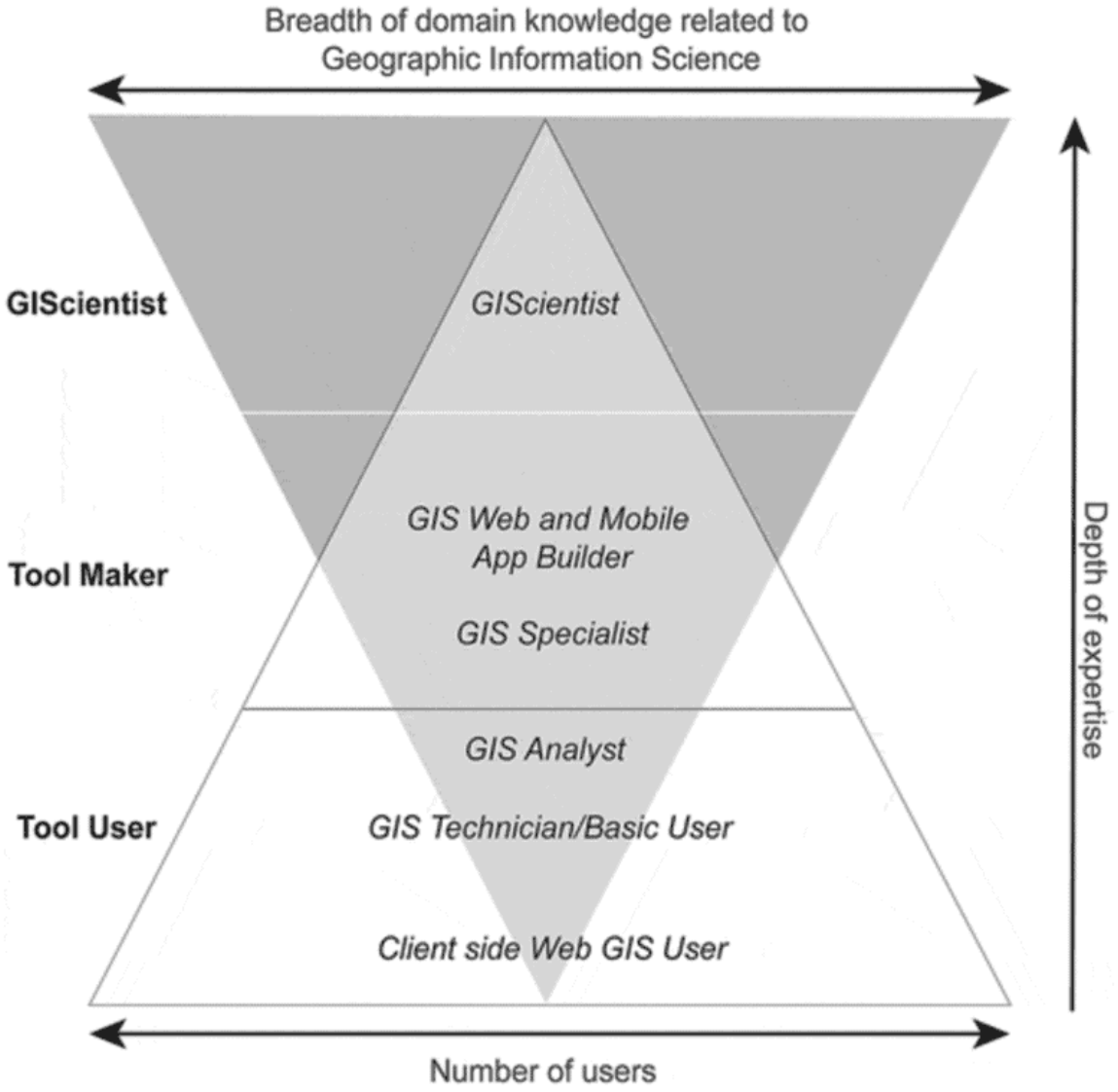
- U-Spatial Training: Desktop GIS 101: Analyzing Data and Creating Maps

- RCCGIS Training: Introduction to ArcGIS, Exploring Geospatial Raster Images

- https://www.indianaview.org

- https://www.esri.com/en-us/what-is-gis/overview

Skill-set



Breadth and depth of knowledge for GIS. Source: Ricker et al., 2020.