GIS Level 1: Introduction to GIS & Mapping

Outline

- Introduction – What is GIS?
- Software options
- Applications
- Understanding Maps & Data
  - Data Layers
  - Spatial Data Types
  - Characteristics of Spatial Data
  - Metadata
- Making Great Maps – Data Visualization Principles

INTRODUCTION

Geographic Information System

"A system for

capturing, storing, checking, integrating, manipulating, analyzing and displaying

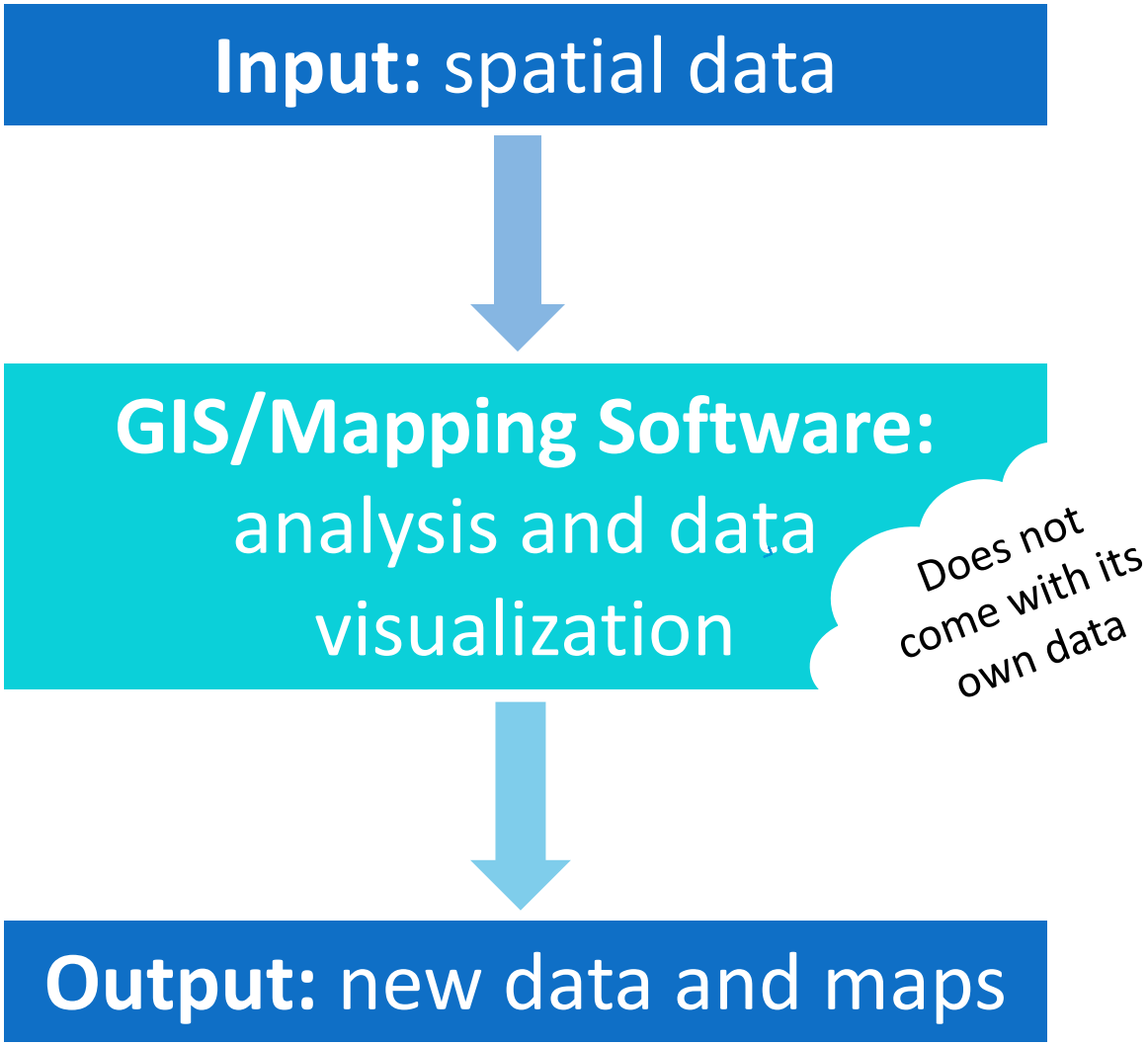
spatial data"

Geographic Information System

"A system for

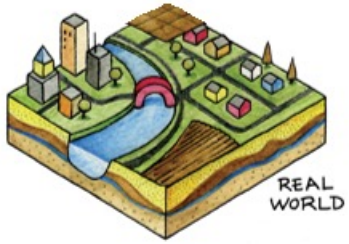
capturing, storing, checking, integrating, manipulating, analyzing and displaying

spatial data"


Input: spatial data
GIS/Mapping Software: analysis and data visualization
Does not come with its own data
Output: new data and maps

Theoretical Overview

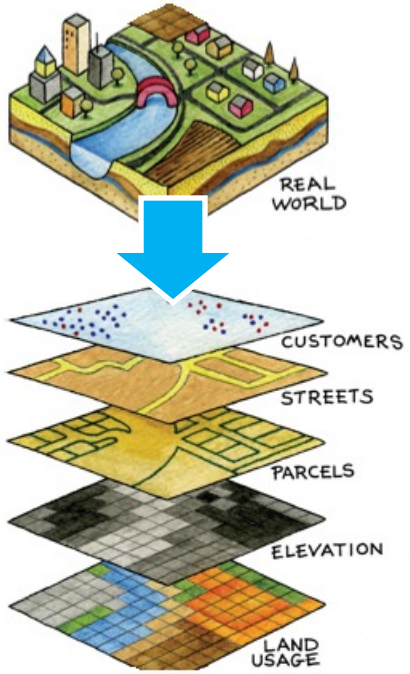
GIS recreates real world spatial data



Theoretical Overview

GIS recreates real world spatial data

as digitized themed data “layers”
(e.g. locations, boundaries, infrastructure, socioeconomic hydrology, land use/cover)

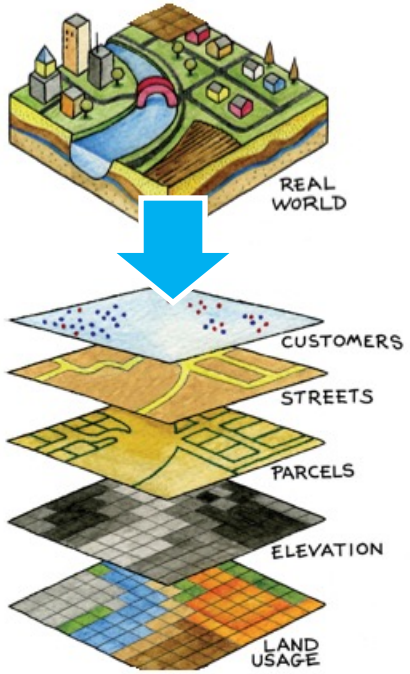


Theoretical Overview

GIS recreates real world spatial data

as digitized themed data “layers”
(e.g. locations, boundaries, infrastructure, socioeconomic hydrology, land use/cover)

assembled in any combination



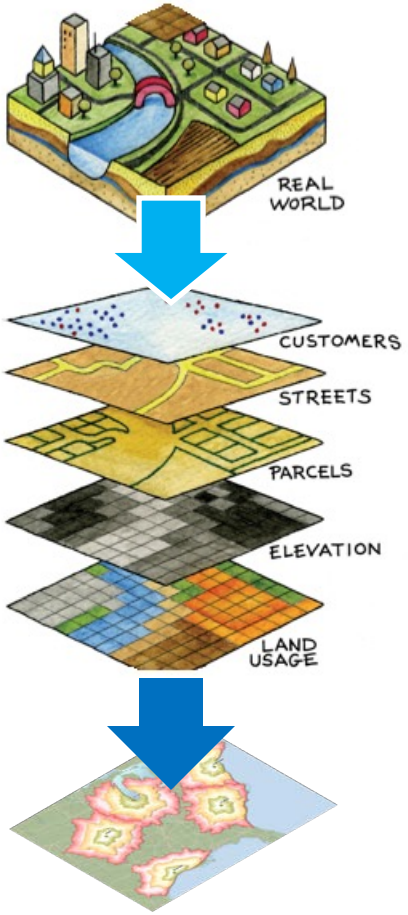
Theoretical Overview

GIS recreates real world spatial data

as digitized themed data “layers”
(e.g. locations, boundaries, infrastructure, socioeconomic hydrology, land use/cover)

assembled in any combination

and overlaid for analysis



SOFTWARE

Types of GIS & Mapping Software

Type	Analysis Power	Example(s)
Geobrowser	**Weak** (mainly only to display data)	Google Maps, Google Earth, Apple Maps, Waze, etc.
Web-based	**Medium** (able to upload additional data, customize display, and perform basic analyses)	Carto, ArcGIS Online, Mapbox, Google MyMaps, etc.
Desktop	**Strong** (installed locally, provides full control of map creation, and perform advanced analyses)	ArcGIS Pro QGIS

Which desktop software should you use today?

ArcGIS Pro (by ESRI)

- Commercial software (expensive to purchase)
- Only runs on Windows
- Larger program – can run slowly on some computers
- Full set of GIS functions and tools
- Integration with ArcGIS Online
- Fully developed training program (online modules, written tutorials, MOOCs)
- Comprehensive support (direct support from ESRI, documentation for every tool)

QGIS

- Free, open-source tool
- Runs on any operating system
- Smaller program that will not affect performance of your computer
- Many available tools, but lacking some for specific functions, such as network analysis (i.e. routing) and spatial statistics
- Basic tutorials by QGIS developers and users
- Tools can be developed by anyone so performance and documentation is inconsistent.
- Support via forums

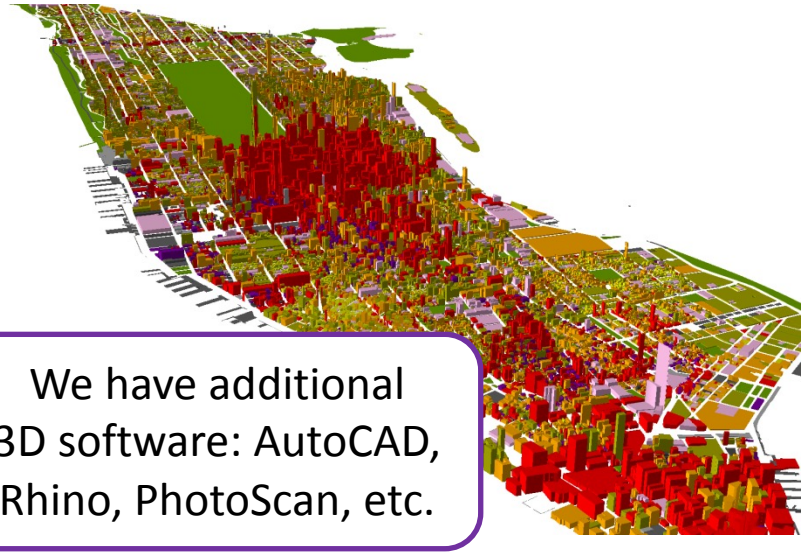
GIS APPLICATIONS

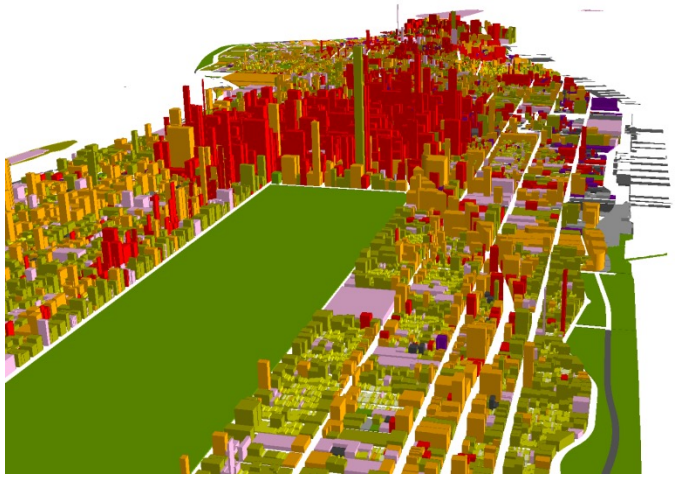
View Imagery

Check out our workshops on **Remote Sensing & Imagery**

City of Cambridge Aerial Photograph, April 2010

Create 3D models

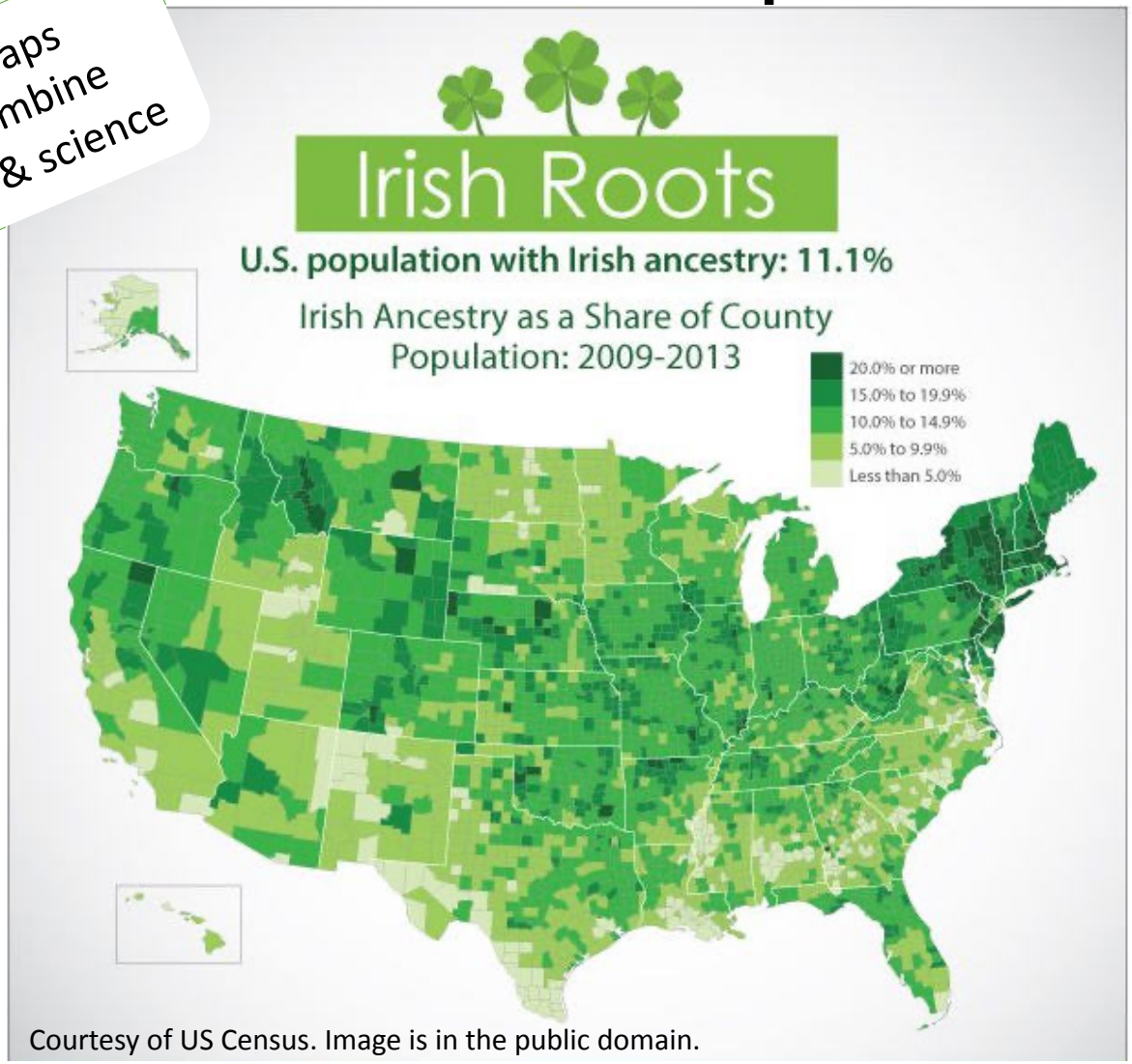




We have additional 3D software: AutoCAD, Rhino, PhotoScan, etc.

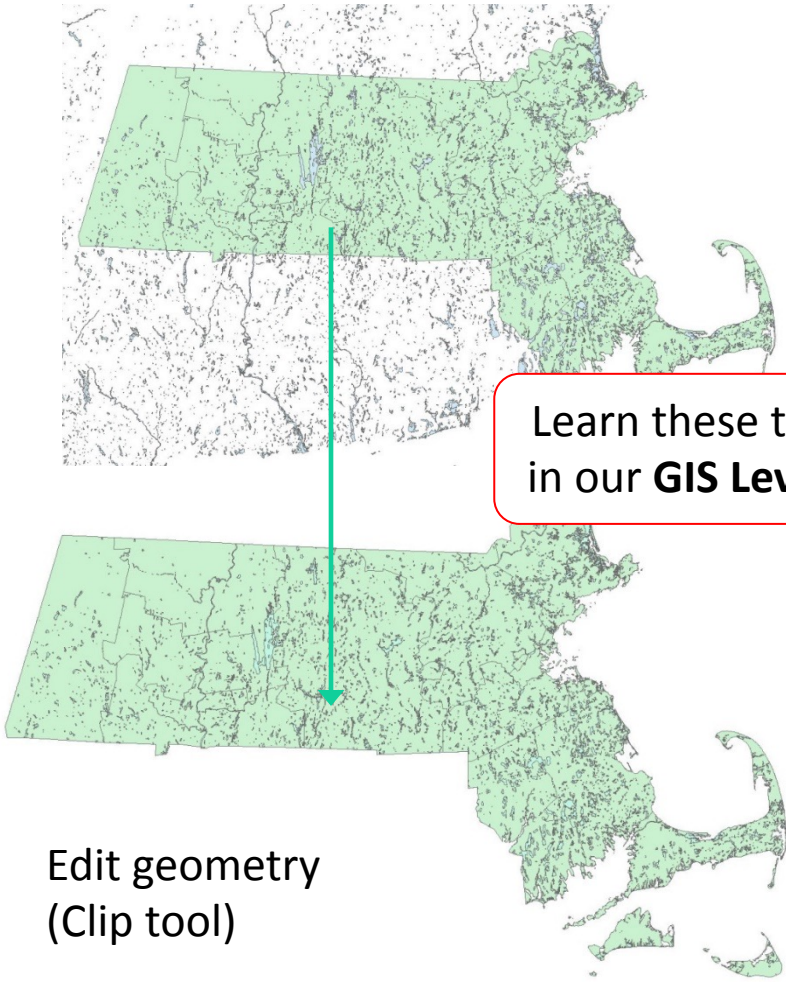
Create Maps

Maps combine art & science

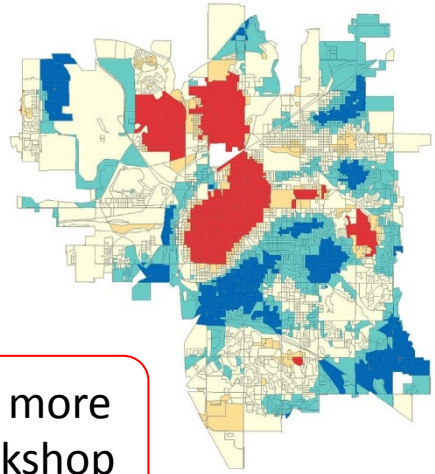


Courtesy of US Census. Image is in the public domain.

Conduct Analyses

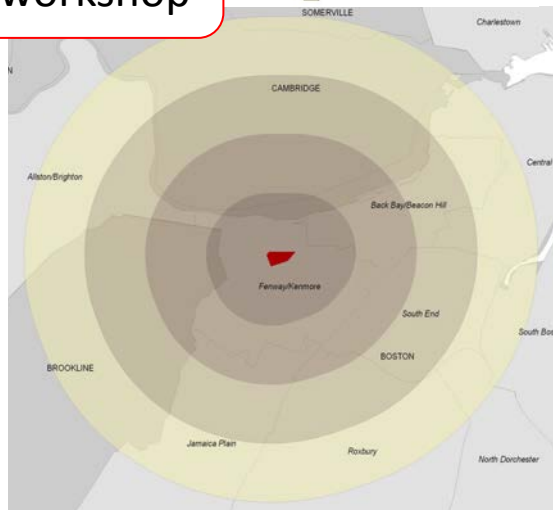


Edit geometry
(Clip tool)



Analyze values
(Spatial Statistics)

Learn these tools and more in our **GIS Level 2** workshop



Create data
(Buffer tool)

UNDERSTANDING MAPS & DATA

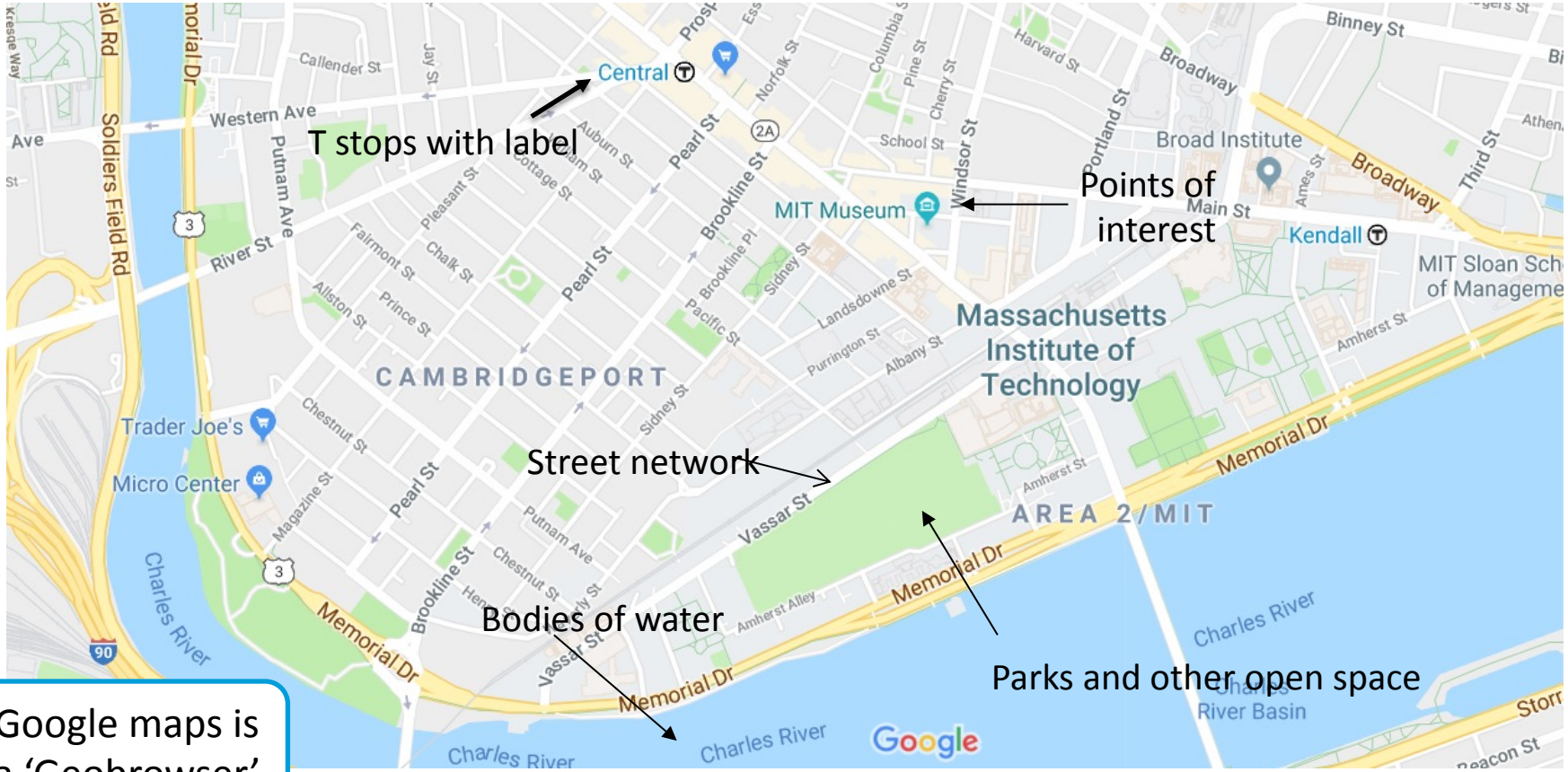
You may have been looking at geospatial data for a long time

Understanding data ‘layers’

What individual data layers were used to create this map?

Understanding data ‘layers’

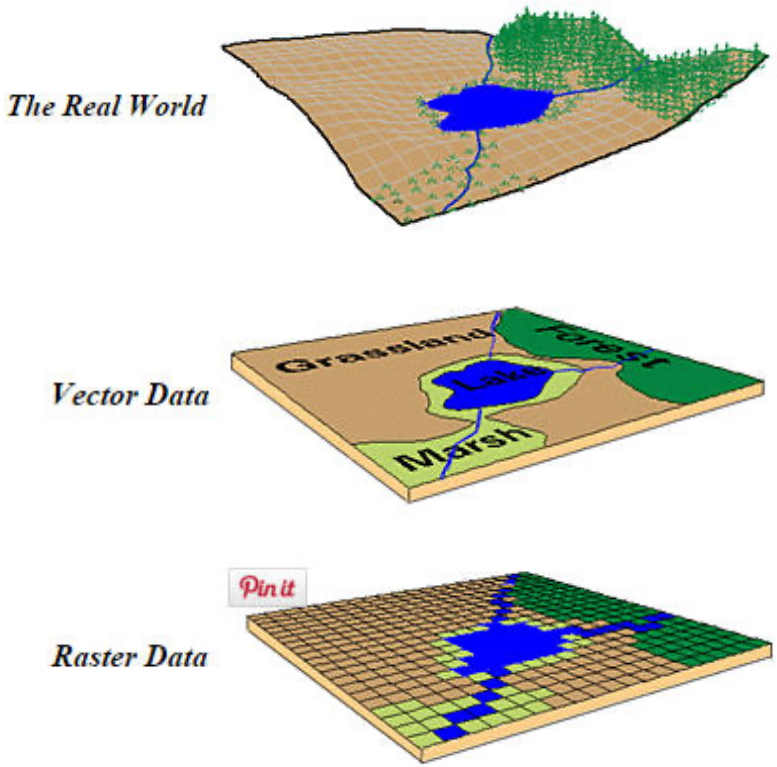
What individual data layers were used to create this map?



Google maps is a ‘Geobrowser’

MAPS & DATA:
SPATIAL DATA TYPES

Geospatial Data Types



Geospatial or coordinate data can be represented in two different data formats:

Vector:

e.g. points, lines, and polygons

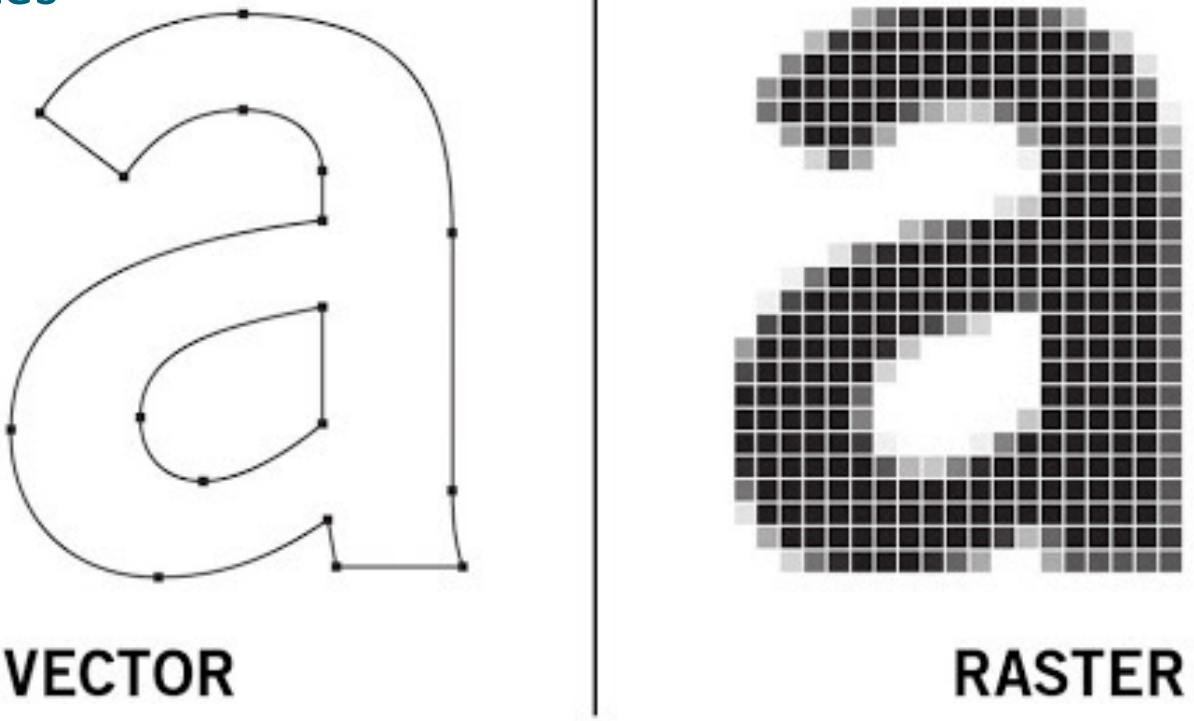
Raster:

e.g. row and column matrix

Data Types: Vector versus Raster

Vectors are composed of coordinates

Raster's are composed of pixels



These are often used for variables with:

defined borders, e.g. manmade

continuous surface, e.g. environmental

Image courtesy of Zina Yonten. Used under CC BY-NC.

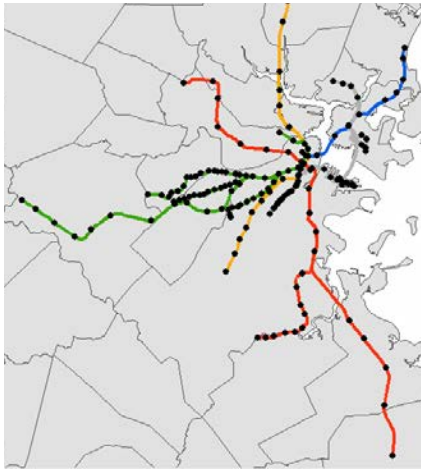
Data Types: Vector examples

Points

Lines

Polygons

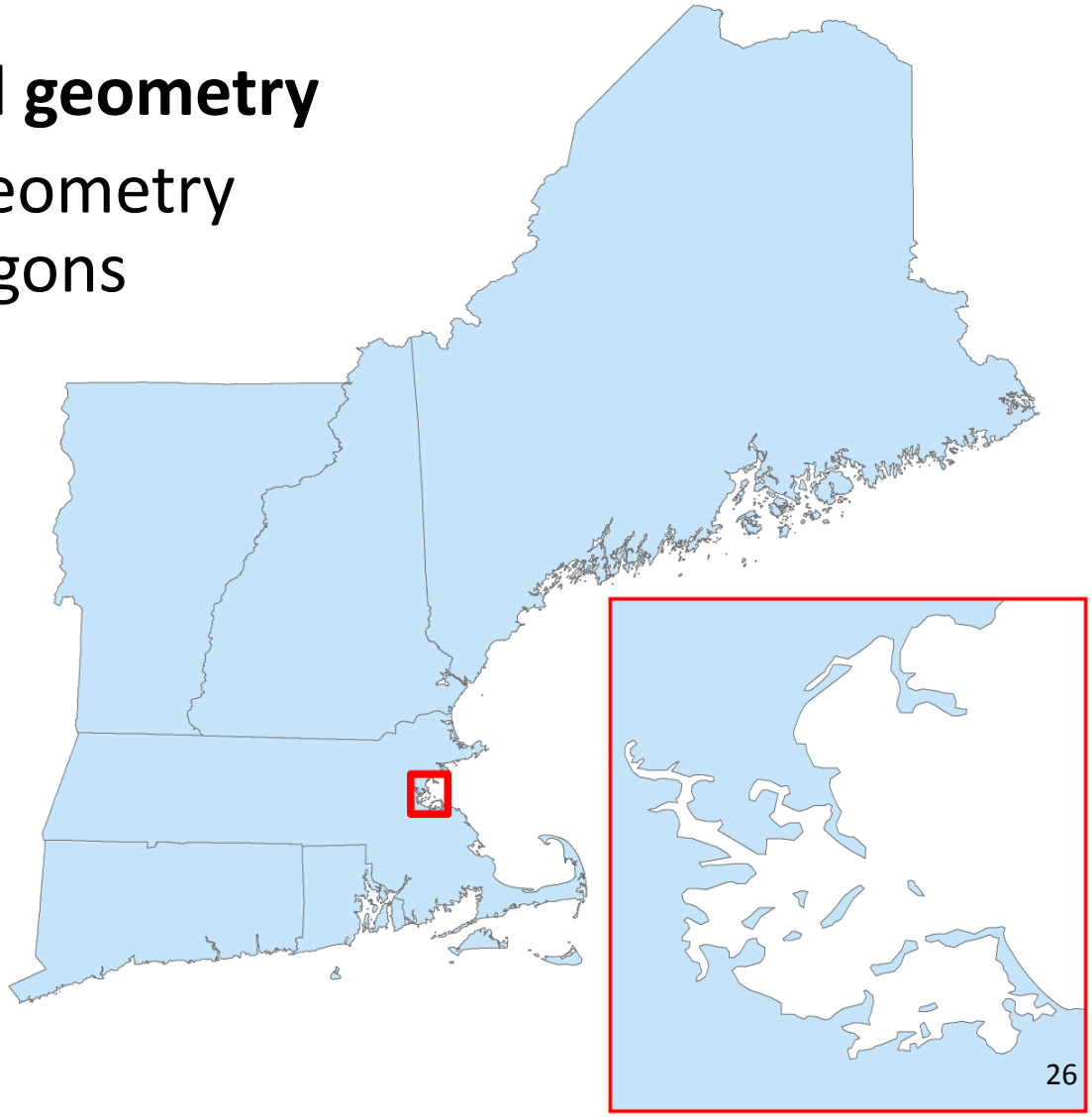
(Combined)



Data Types: Vector mapping

Vectors have a **frontend geometry**

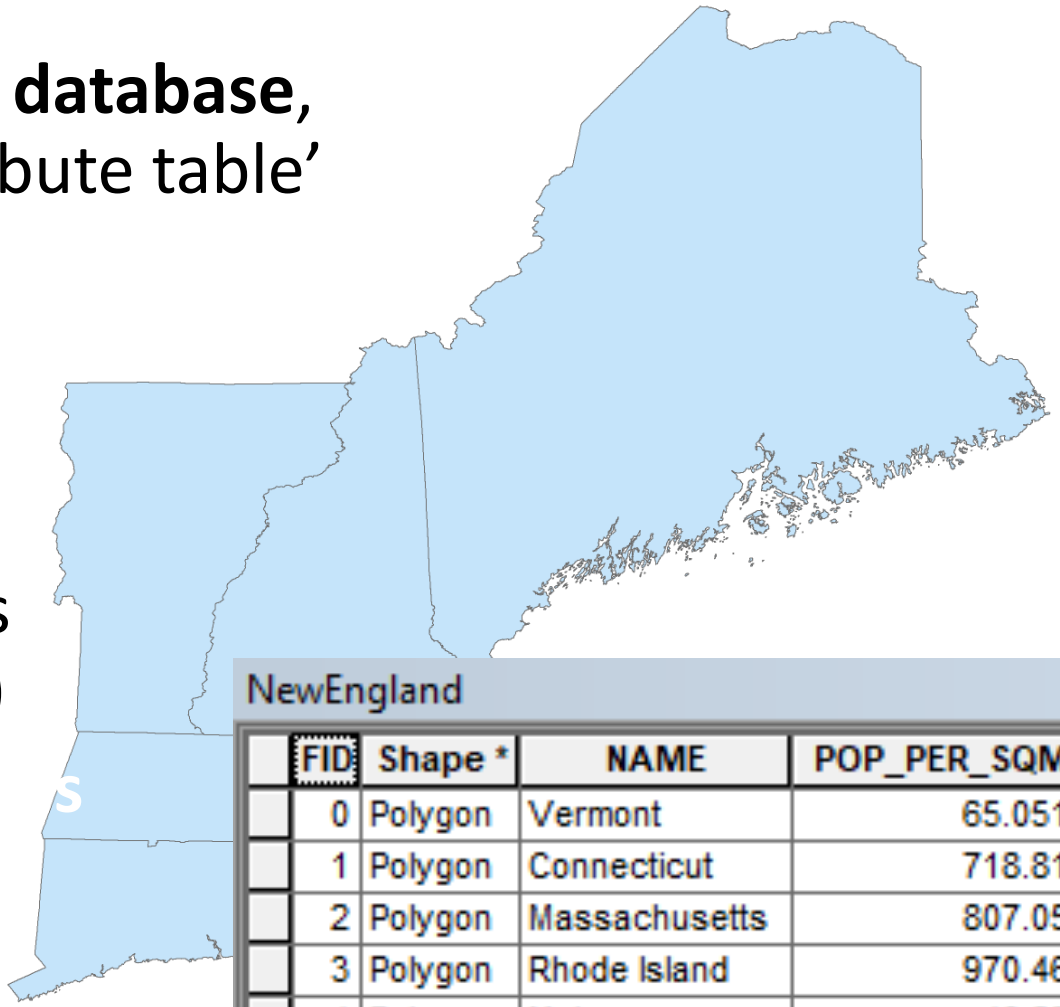
- In this example the geometry represents state polygons



Data Types: Vector mapping

Vectors have a **backend database**, normally called an 'attribute table'

- **rows** represent unique geometries (e.g. state polygons)
- **columns** represent a number of variables (theoretically infinite)



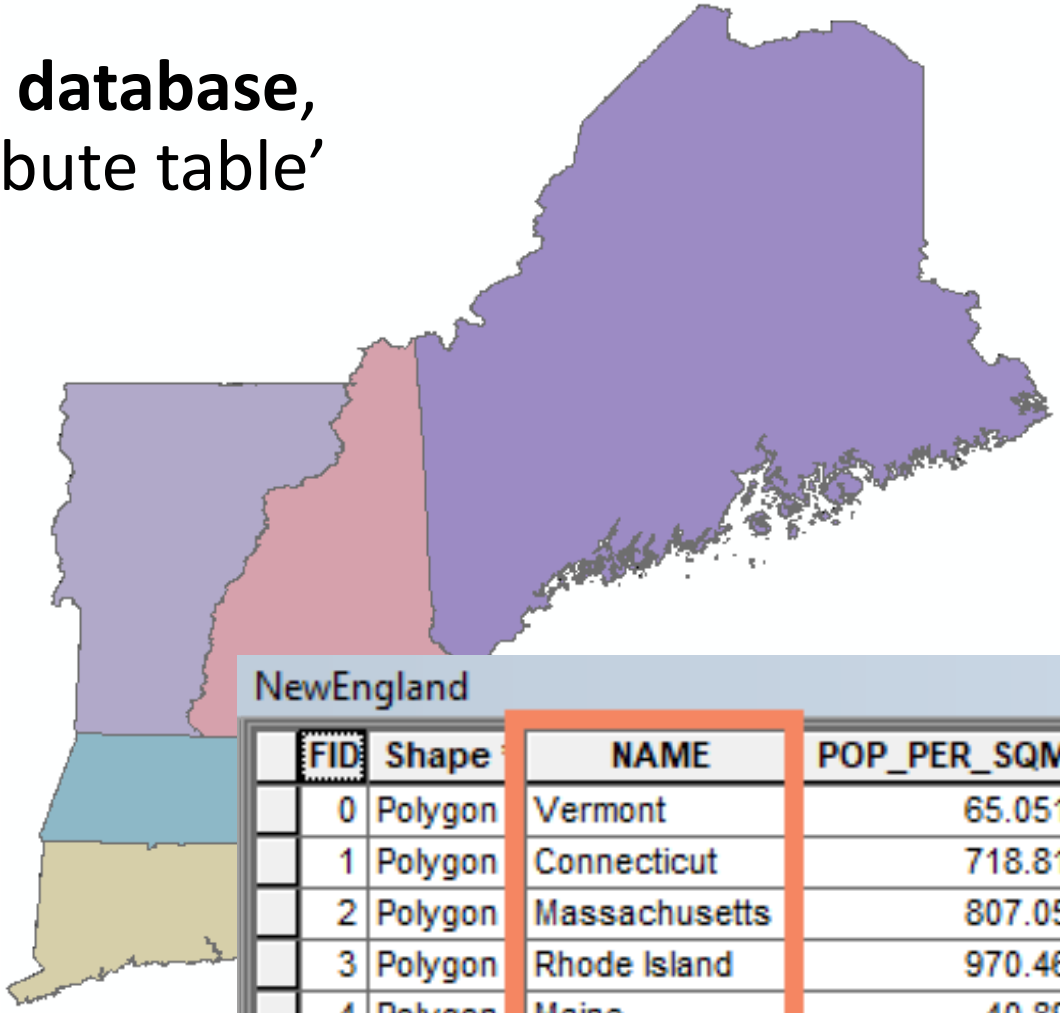
NewEngland

FID	Shape *	NAME	POP_PER_SQMI
0	Polygon	Vermont	65.0519
1	Polygon	Connecticut	718.816
2	Polygon	Massachusetts	807.054
3	Polygon	Rhode Island	970.466
4	Polygon	Maine	40.891
5	Polygon	New Hampshire	142.186

Data Types: Vector mapping

Vectors have a **backend database**, normally called an ‘attribute table’

- **Here each state is being symbolized by ‘NAME’ (qualitative variable)**



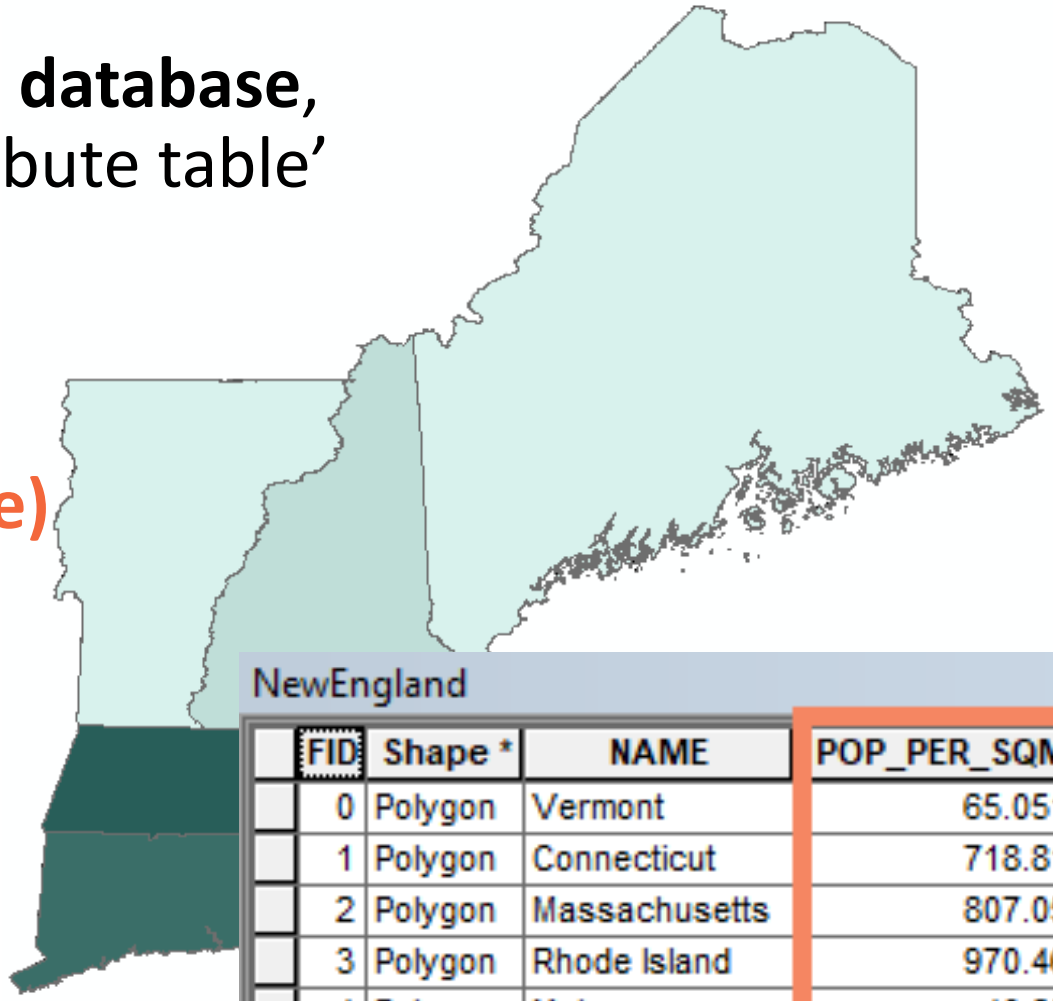
NewEngland

FID	Shape	NAME	POP_PER_SQMI
0	Polygon	Vermont	65.0519
1	Polygon	Connecticut	718.816
2	Polygon	Massachusetts	807.054
3	Polygon	Rhode Island	970.466
4	Polygon	Maine	40.891
5	Polygon	New Hampshire	142.186

Data Types: Vector mapping

Vectors have a **backend database**, normally called an ‘attribute table’

- **Here each state is being symbolized by ‘POP_PER_SQMI’ (quantitative variable)**



NewEngland

FID	Shape *	NAME	POP_PER_SQMI
0	Polygon	Vermont	65.0519
1	Polygon	Connecticut	718.816
2	Polygon	Massachusetts	807.054
3	Polygon	Rhode Island	970.466
4	Polygon	Maine	40.891
5	Polygon	New Hampshire	142.186

Data Types: Vector file formats

- The shapefile is the most common vector file format.
- “A” shapefile is actually a collection of several different files with different extensions.

When adding files to ArcGIS Pro, you will only see one file, not every extension.

Shapefile =

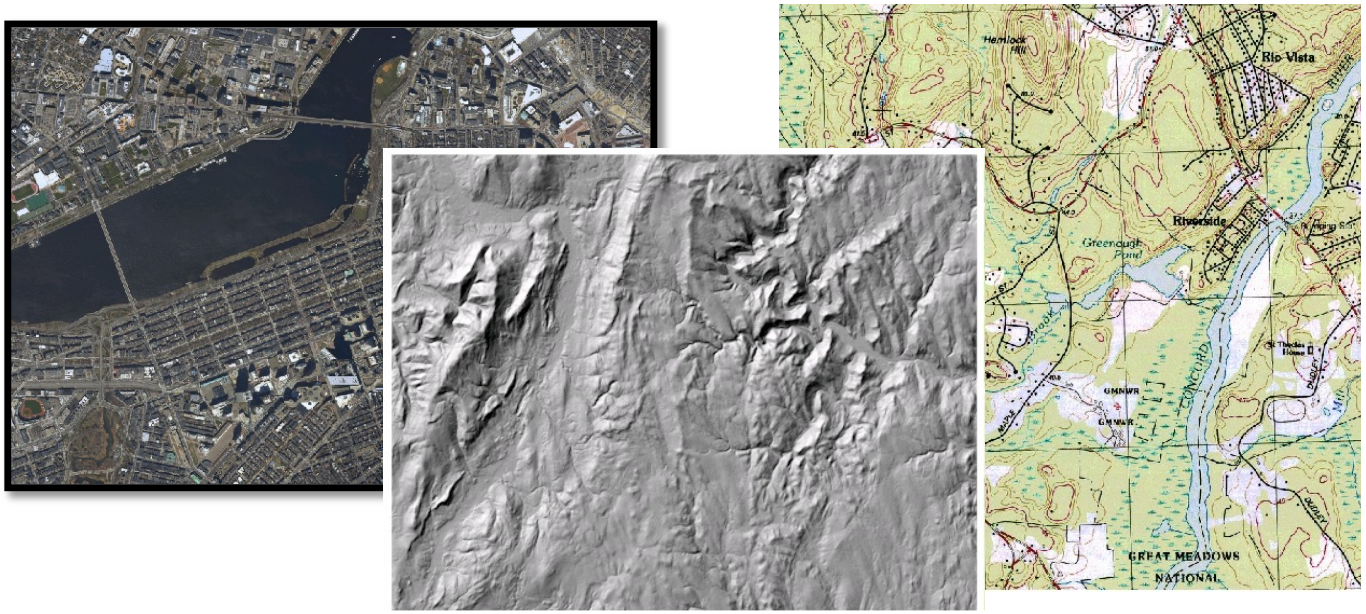
.shp .shx .sbx

.dbf .prj

Make sure to keep all files together when moving.

Data Types: Raster

Raster data includes aerial photographs, digital elevation models, and scanned maps.

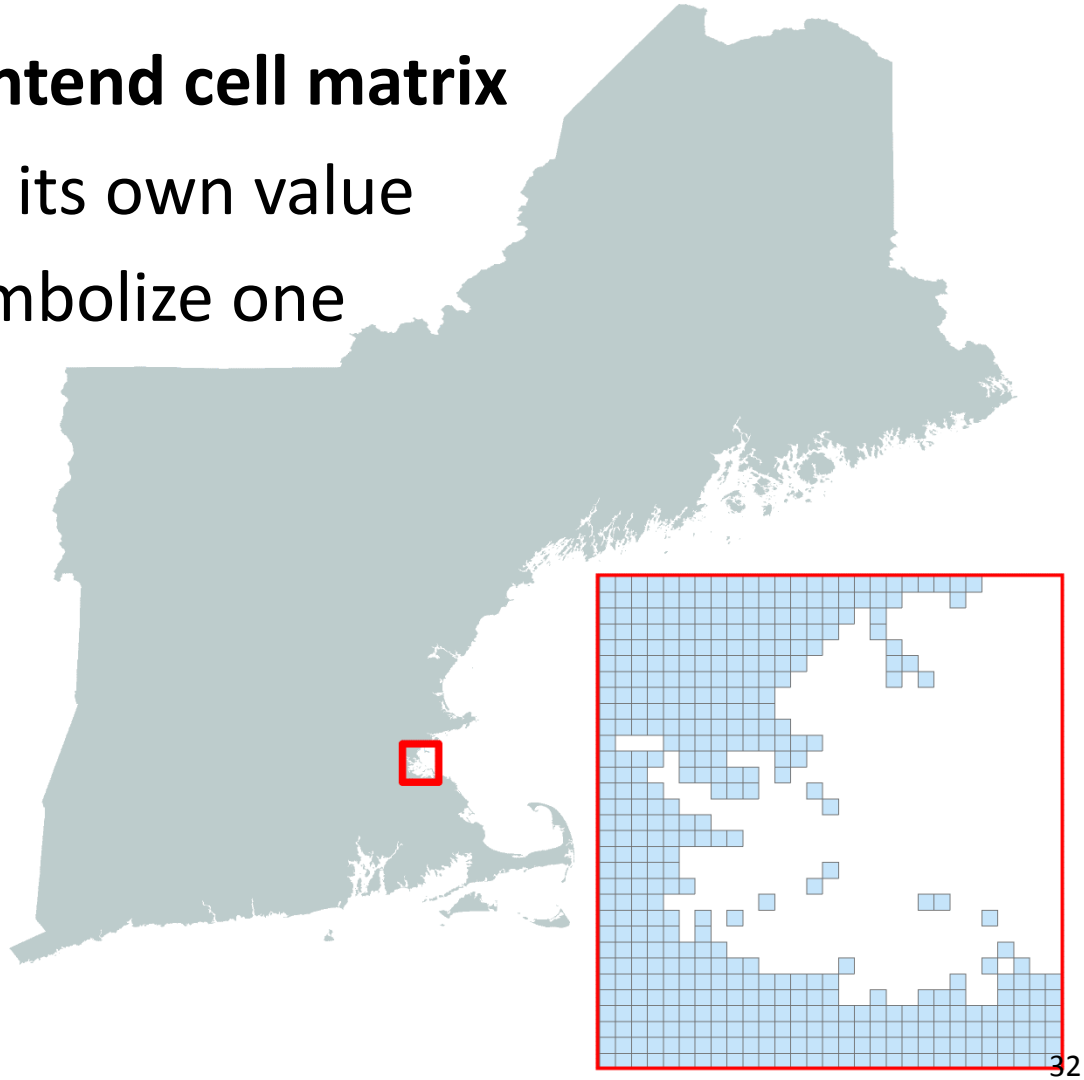


(Remember these are constructed from pixels)

Data Types: Raster mapping

Raster data have a **frontend cell matrix**

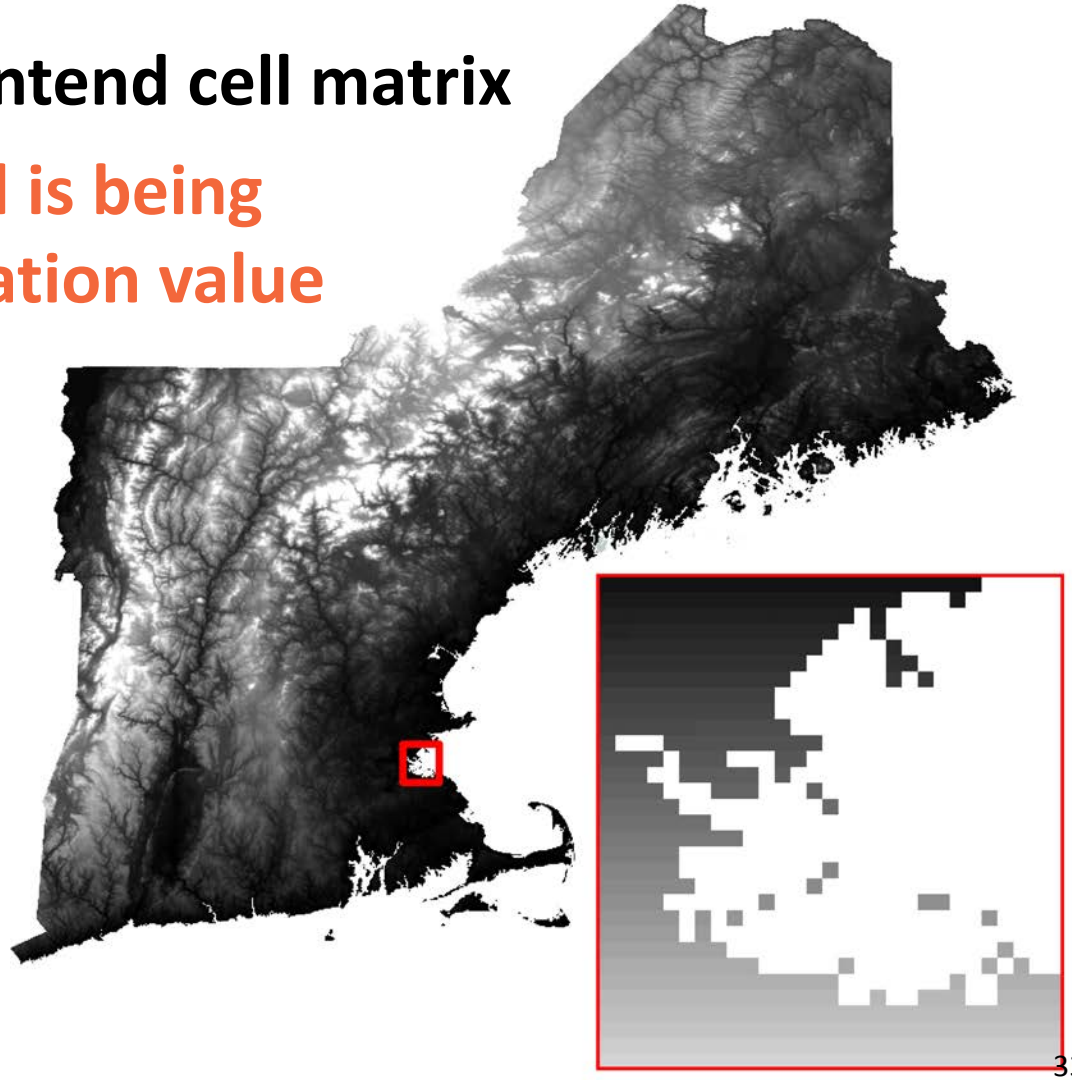
- Where each cell has its own value
- A raster can only symbolize one variable at a time



Data Types: Raster mapping

Raster data have a **frontend cell matrix**

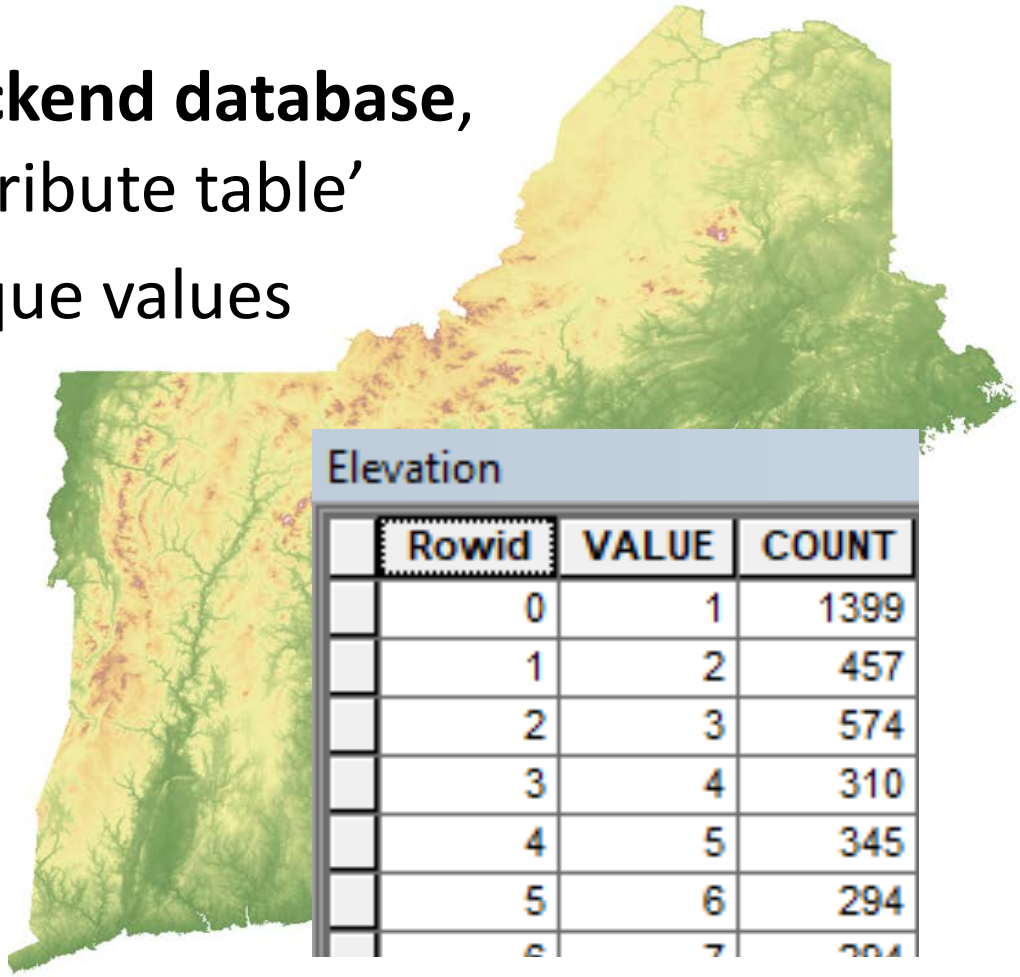
- **Here each cell/pixel is being symbolized by elevation value**



Data Types: Raster mapping

Raster data have a **backend database**, normally called an 'attribute table'

- **rows** represent unique values (1m, 2m, 3m, etc.)
- **columns** have specific variables
  1) unique 'ROW ID'
  2) unique 'VALUE'
  3) 'COUNT' of pixels with that 'VALUE'



Elevation

Rowid	VALUE	COUNT
0	1	1399
1	2	457
2	3	574
3	4	310
4	5	345
5	6	294

Data Types: Raster file formats

There are many different raster file extensions, including common image formats.

.tiff

.asc

.img

.jpg

Learn more about raster formats in this ArcGIS Pro documentation. QGIS supports similar formats.

Some formats may include a collection of files with different extensions, similar to a shapefile.

Data Types: Tabular

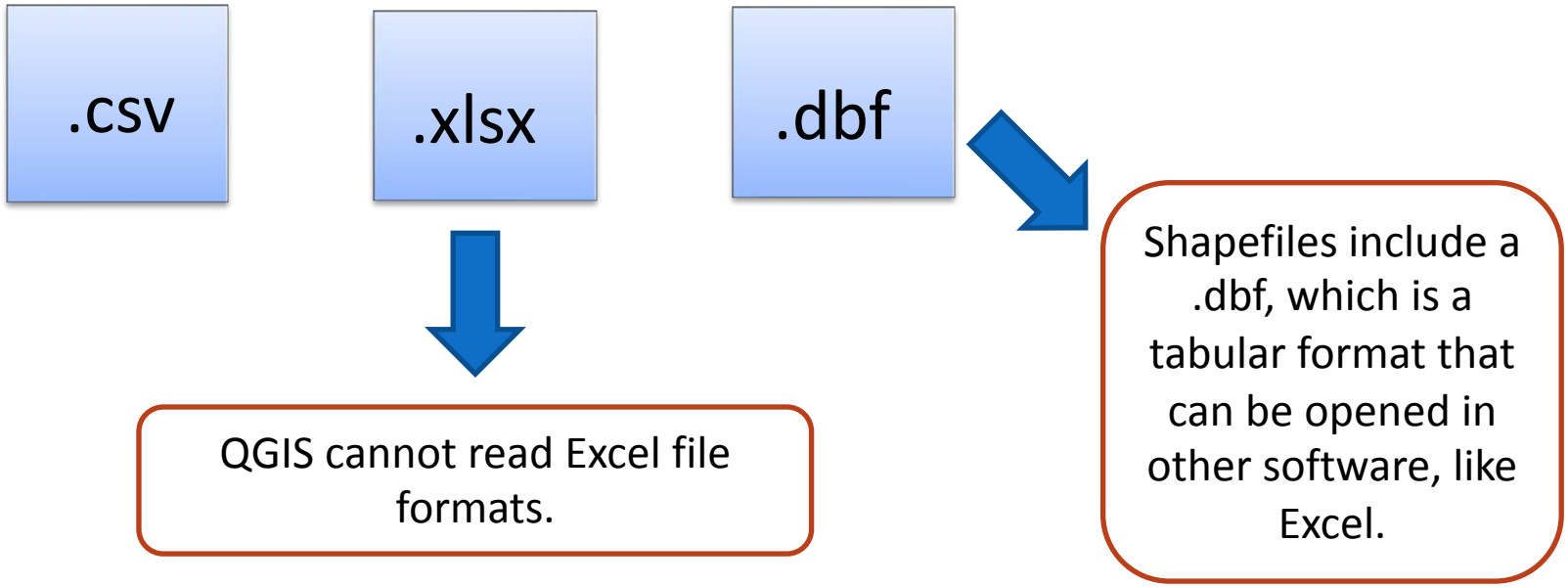
Tabular data can be transformed into spatial data in two ways:

Created by Musavvir Ahmed from Noun Project

1. Joining
   - **Use a shared unique identifier** (GEOID, name, etc.) to match up tabular data to the spatial data's attribute table.
2. Geocoding
   - **Use lat/lon** coordinates in table to plot as points on map
   - **Use addresses** to plot locations based on a street network

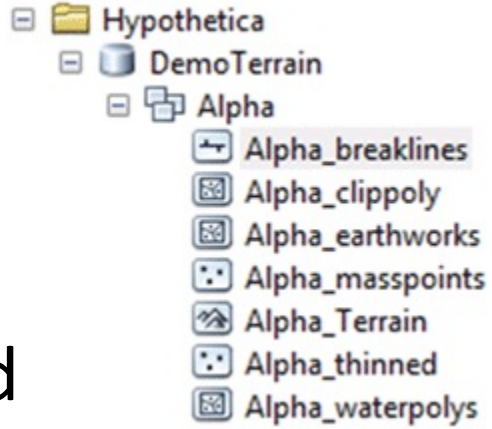
Data Types: Tabular file formats

GIS software can read commonly used tabular formats in order to transform them into spatial data.



Geodatabases

- ESRI/ArcGIS storage system
- a collection of geographic datasets of various types held in a common file system folder
- **Advantages:** larger files size limits, faster processing time when using analysis tools
- **Disadvantages:** can only be opened in ESRI software
- Learn more about using geodatabases in Pro.



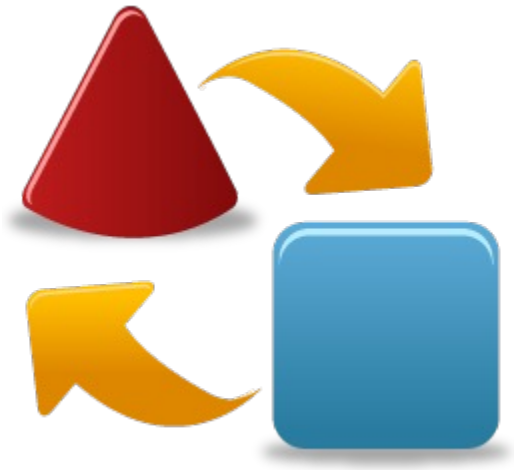
Other data formats

GIS can import and convert data produced in other formats:

- KML / KMZ files (Google Earth)
- DXF / DWG (CAD)
- NetCDF (scientific data)
- LAS (Lidar)
- GPX (GPS units)
- Geojson

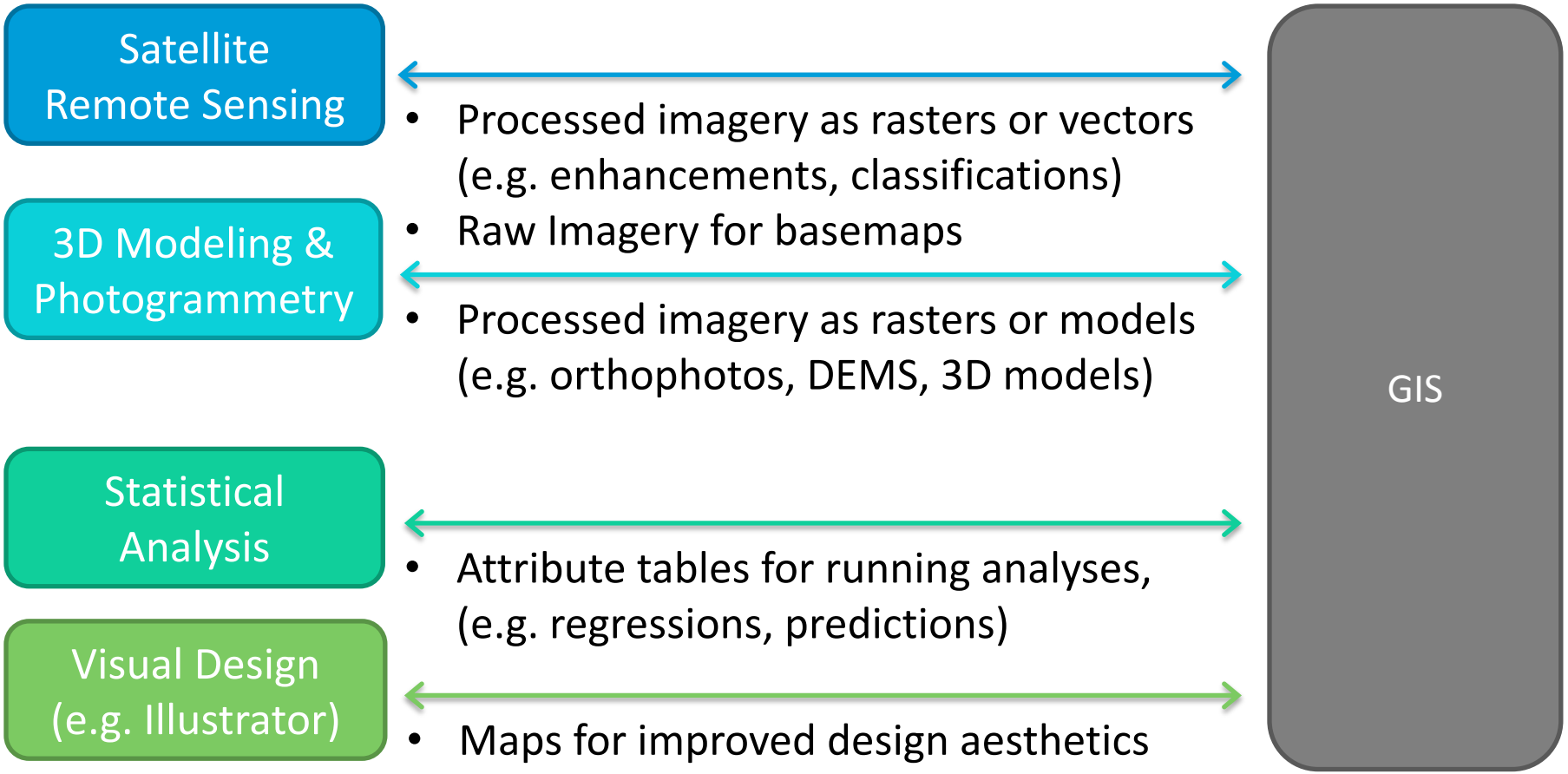
GIS software can export many formats:

- Adobe Illustrator
- KML
- CAD
- TIF
- JPG



The GIS & Data Lab has many types of data visualization software.

Common Associated Workflows

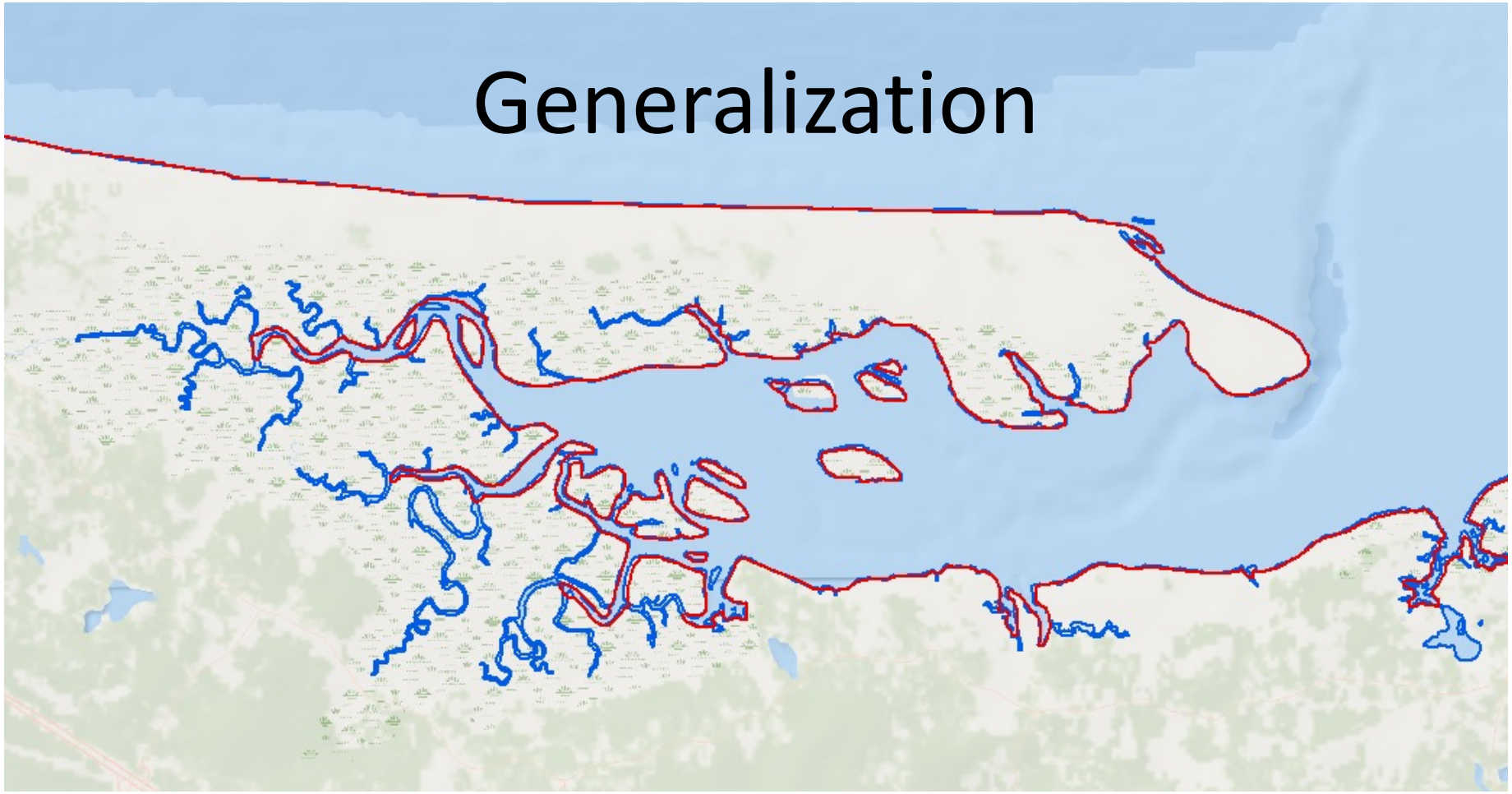


Exercise 1

- Goals:
  - Become familiar with the GIS interface
  - Learn how to add data
  - Explore data types & attributes
- Complete either the QGIS or ArcGIS Pro exercise from your workshop folder.

MAPS & DATA:
CHARACTERISTICS OF SPATIAL DATA

Generalization



- The most detailed data available is not suitable for all purposes (or often a manageable file size)
- e.g. resolution of coastline data for this map is scale dependent
  - **Red**: county map
  - **Blue**: town map

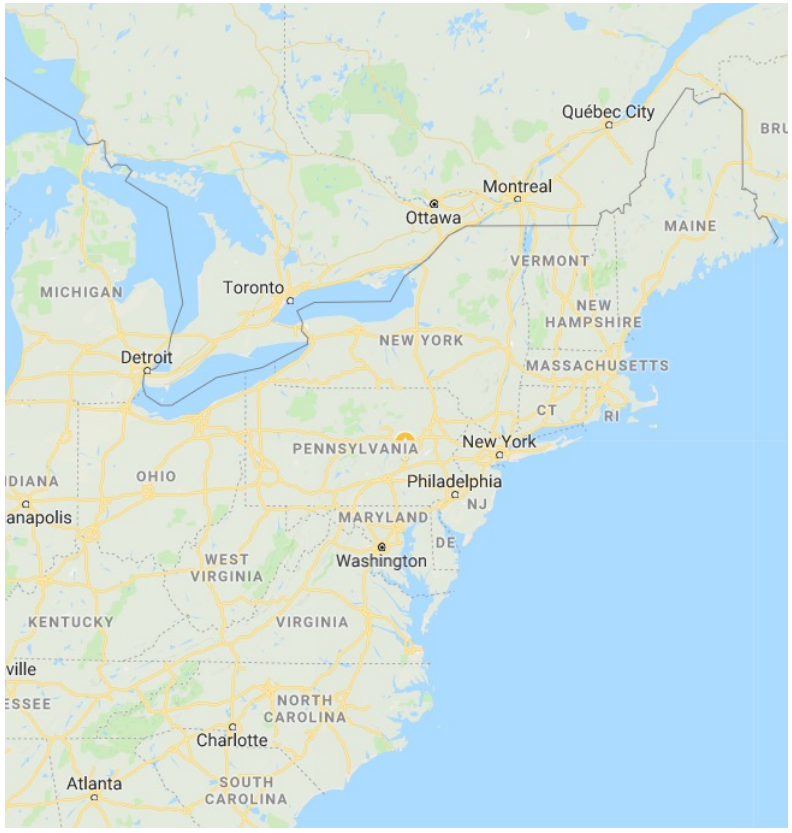
Abstraction

The process of reducing data from its complete state to what is necessary for use and presentation

Quiz: Which data symbology (pictured above) would you select for each of the following maps?

- Land use study of adjacent property
- Development map of the airport
- National map of airports

Spatial Resolution/Scale

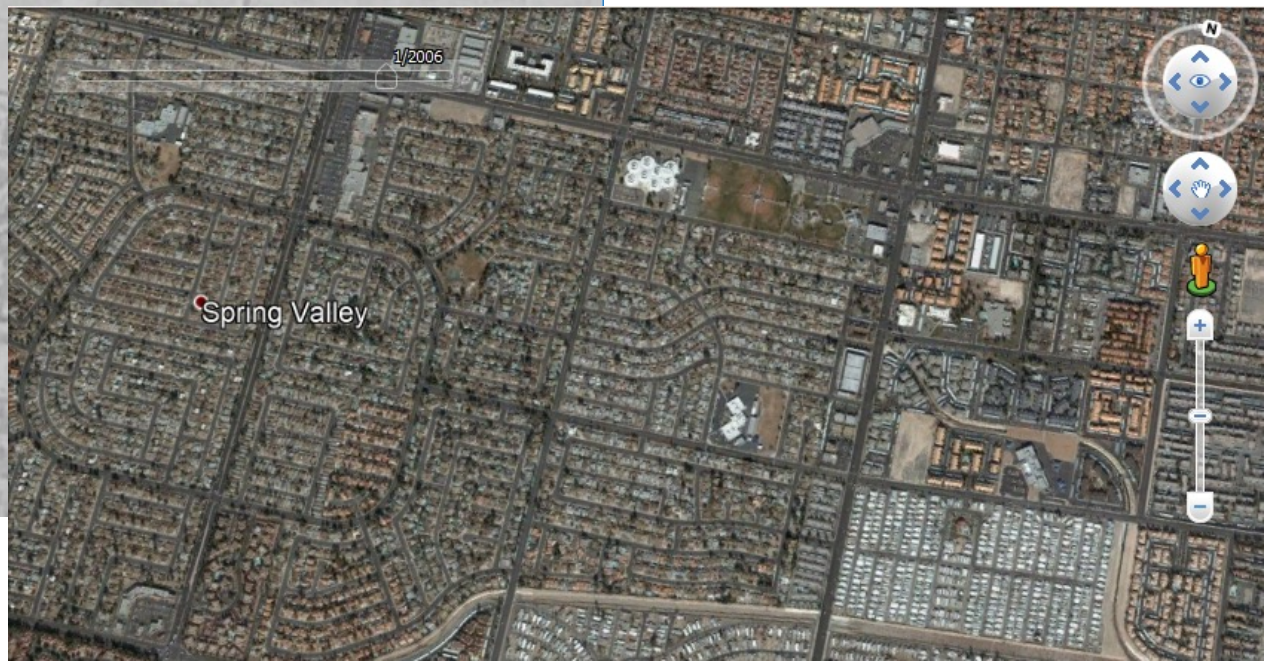


Suitable data geometry is dependent on scale:

e.g. roads are polygons at local scale but lines at national scale

Temporal Resolution

1977 to 2006



Keep in mind temporal resolution when obtaining data

Searching for Spatial Data

- Look in general GIS data repositories
- Search the internet
  - Include “gis”, or “data” in the search terms
  - Search by location and/or topic
- Search for country statistical agencies or open data sites (large cities often have their own open data portals as well)
- Contact GIS departments, universities, or researchers in your area of interest.
- Search for articles on your topic and look for the sources of the data.

Repositories and Websites

Find many more on our website.

Libguides.mit.edu/gis

- Can also find by googling 'MIT GIS', first result
- Click on Find Data Tab for a list of resources, including an assembled links of common data sources per topic.

Geodata.mit.edu (Geoweb)

- Includes data licensed freely or restricted to MIT and other institutions, plus CDs and DVDs in the GIS lab.
- MIT instance is mainly historical-local or purchased data.

OpenStreetMap.org

- Crowd-sourced maps; content will vary by location
- Download as a shapefile via http://www.geofabrik.de/
- Best source to start for rural international data.

MAPS & DATA: METADATA

What is Metadata?

Use metadata to learn how and why the data were created, access restrictions, columns in the attribute table, and much more!

Citation Information

Originator	MassGIS
Publication Date	20030301
Title	Massachusetts (Major Drainage Basins, 2003)
Geospatial Data Presentation Form	vector digital data
Publication Place	Massachusetts
Publisher	Massachusetts Office of Geographic and Environmental Information (MassGIS)
Online Linkage	Server=arrowsmith.mit.edu; Service=5150; Database=oracle

Description

Abstract	This layer is a statewide digital datalayer of the 28 major drainage basins of Massachusetts as defined by the USGS Water Resources Division and the MA Water Resources Commission. It contains 885 arcs and 82,000 vertices. It is stored as a single statewide coverage. MassGIS created this data.MassGIS is maintaining this datalayer. In February 2002 the DEP GIS Group modified some linework within the USGS Woronoco (#33) Quadrangle to synchronize the major basin divide with changes in the DEP Surface Water Supply Protection Area datalayers (Zone A, B, C) and Sub-basin tiles. In May 2002 the DEP GIS Group revised the N. Coastal/Ipswich basin boundary in Wenham. In March 2003 linework was updated along the Westfield/Connecticut and N. Coastal/Ipswich boundaries.
Purpose	This layer is intended for researchers, students, and policy makers for reference and mapping purposes, and may be used for basic applications such as viewing, querying, and map output production, or to provide a basemap to support graphical overlays and analysis with other spatial data. This layer will provide a basemap for layers related to real property and public utilities More advanced user applications may focus on urban planning, public works, or utilities purposes (including defining boundaries, managing assets and facilities, integrating attribute databases with geographic features, spatial analysis, and presentation output.)

Time Period of Content

Calendar Date	20030301

Metadata Examples

1. MassGIS:

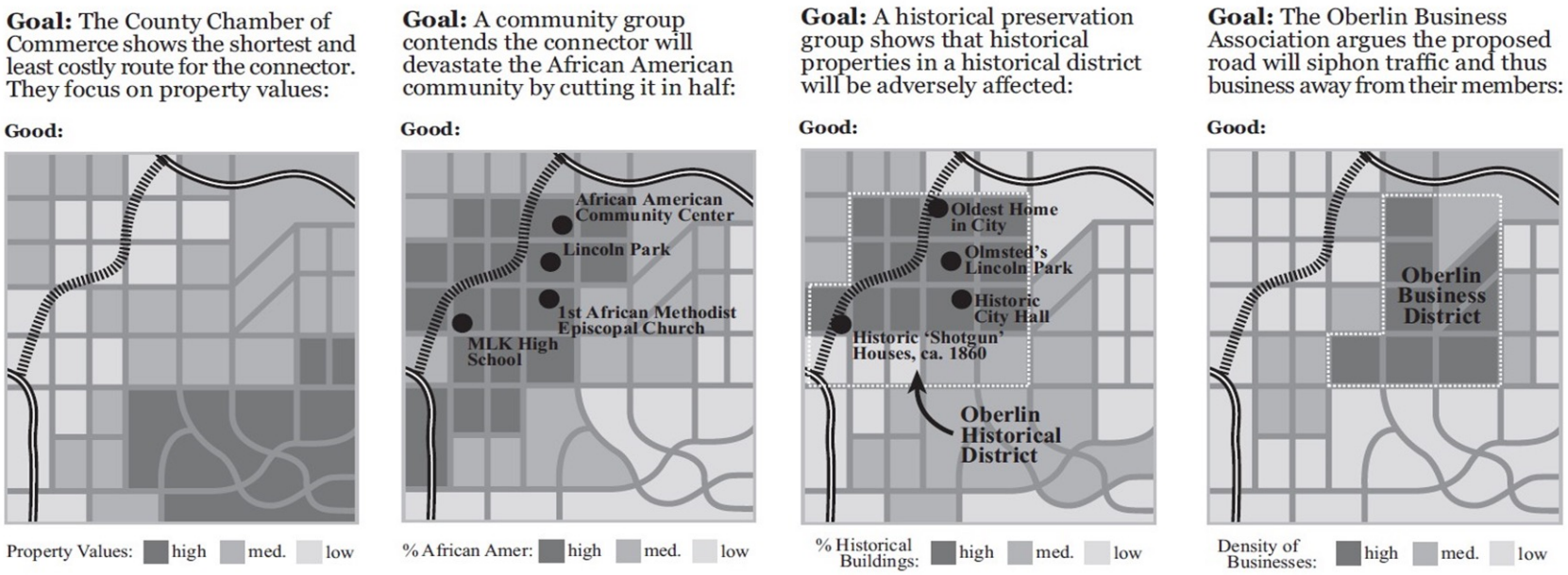
   https://www.mass.gov/info-details/massgis-data-marine-beaches
2. GeoWeb: geodata.mit.edu/catalog/mit-w37ehgh6nvl4w
3. City of Boston: https://data.boston.gov/dataset/traffic-signals

MAKING GREAT MAPS:
DATA VISUALIZATION PRINCIPLES

Making Great Maps

- Cartography is the art and science of making maps
- Maps are always simplifications of reality, which makes them helpful when making decisions or explaining patterns
- Maps are designed by people (who have intentions), so we have to create them responsibly

Making Great Maps



From: ***Making Maps: A Visual Guide to Map Design for GIS*** *by John Krygier and Denis Wood*

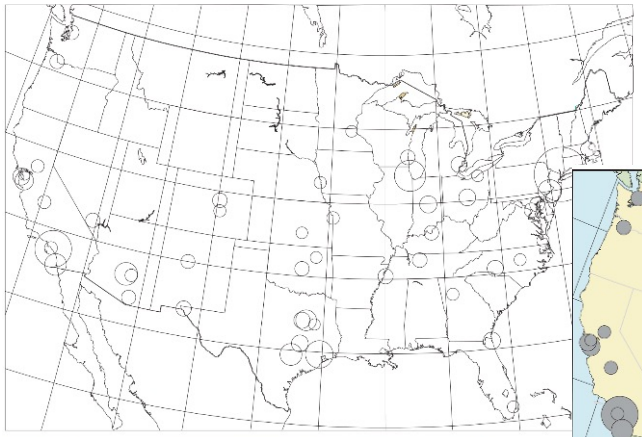
Example of how a map can be used to prove many different points.
However in the past only those in power had the software and data to do so.

Three Key Questions

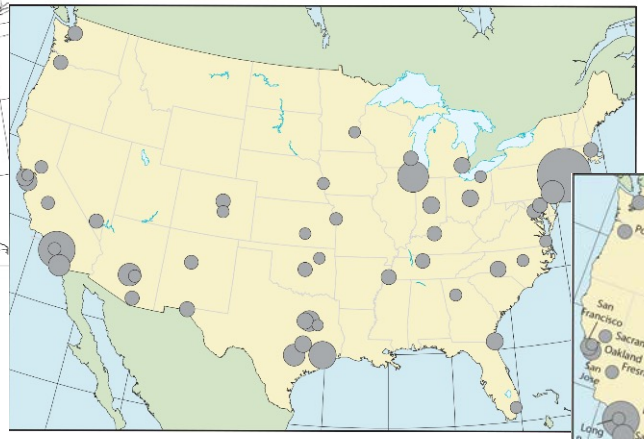
1. **Who wants the map?**
   - e.g. experts (detailed), students (contextual), the community (interactive)
2. **Where will it be seen?**
   - e.g. 8x11 paper (static small, room for main points)
   - e.g. 30x40 poster board (static large, room for detail)
   - e.g. web map (interactive, users control navigation of map)
3. **What is it's purpose?**
   - e.g. to show a variable through time (time series)
   - e.g. to show change over time (change detection)
   - e.g. to combine multiple variable into an index to pick best/worst (sustainability/risk/vulnerability mapping, site selection)

Each question deserves a well-thought answer before mapping

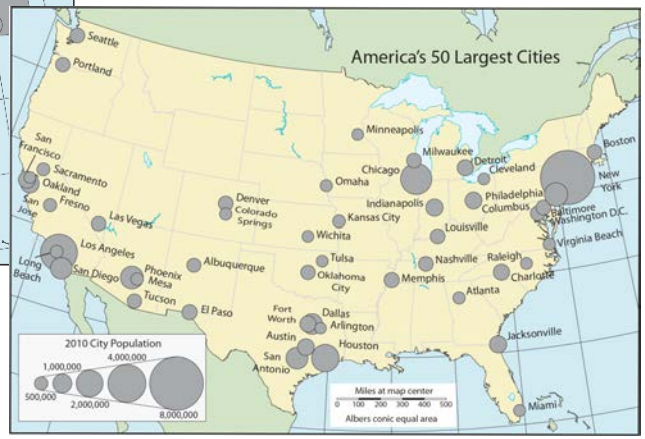
Map Design Process



Start with assembling the data from multiple sources

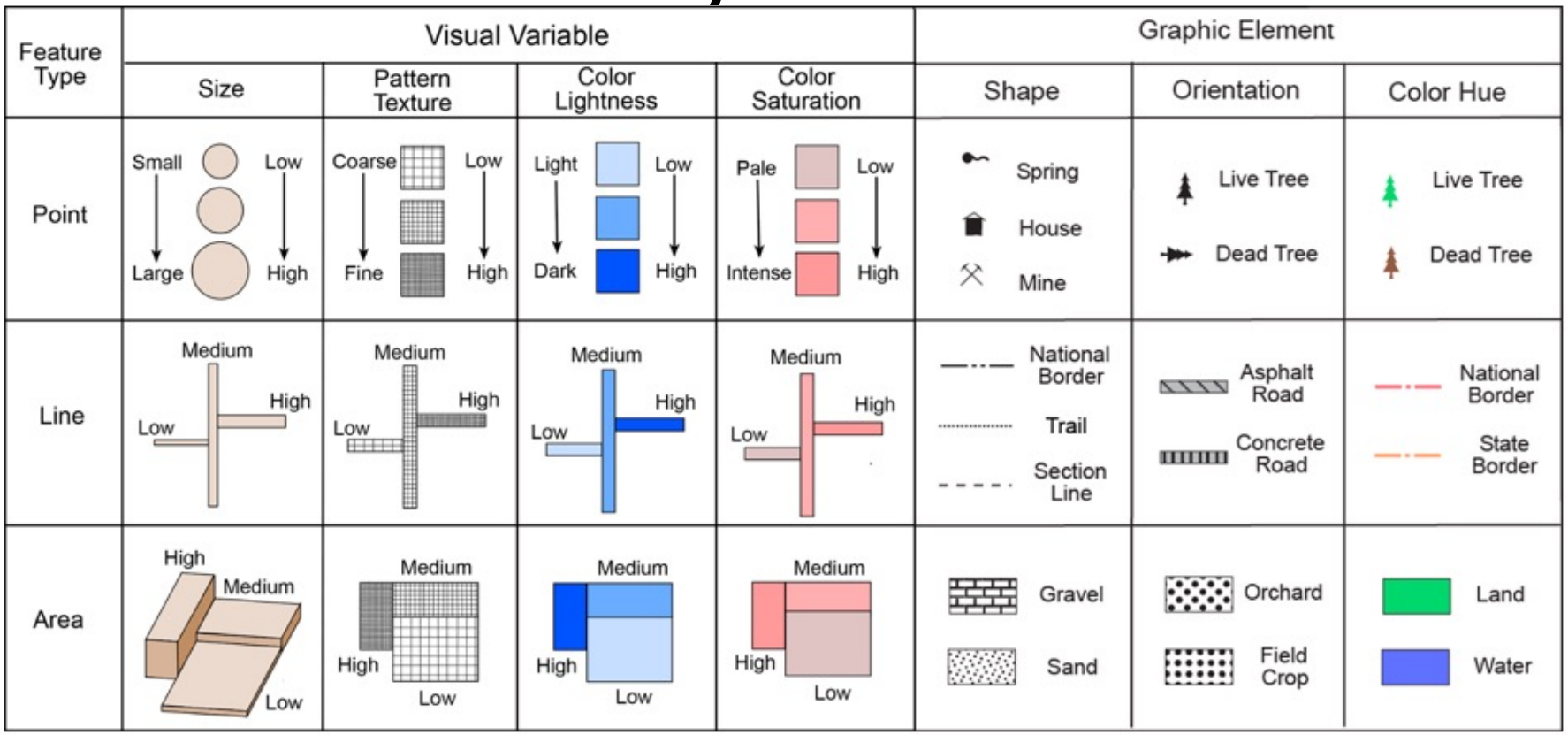


Next choose the data, analyses, & symbolization



Lastly insert the title, legend, north arrow, scale bar, & labels

Vector Symbolization

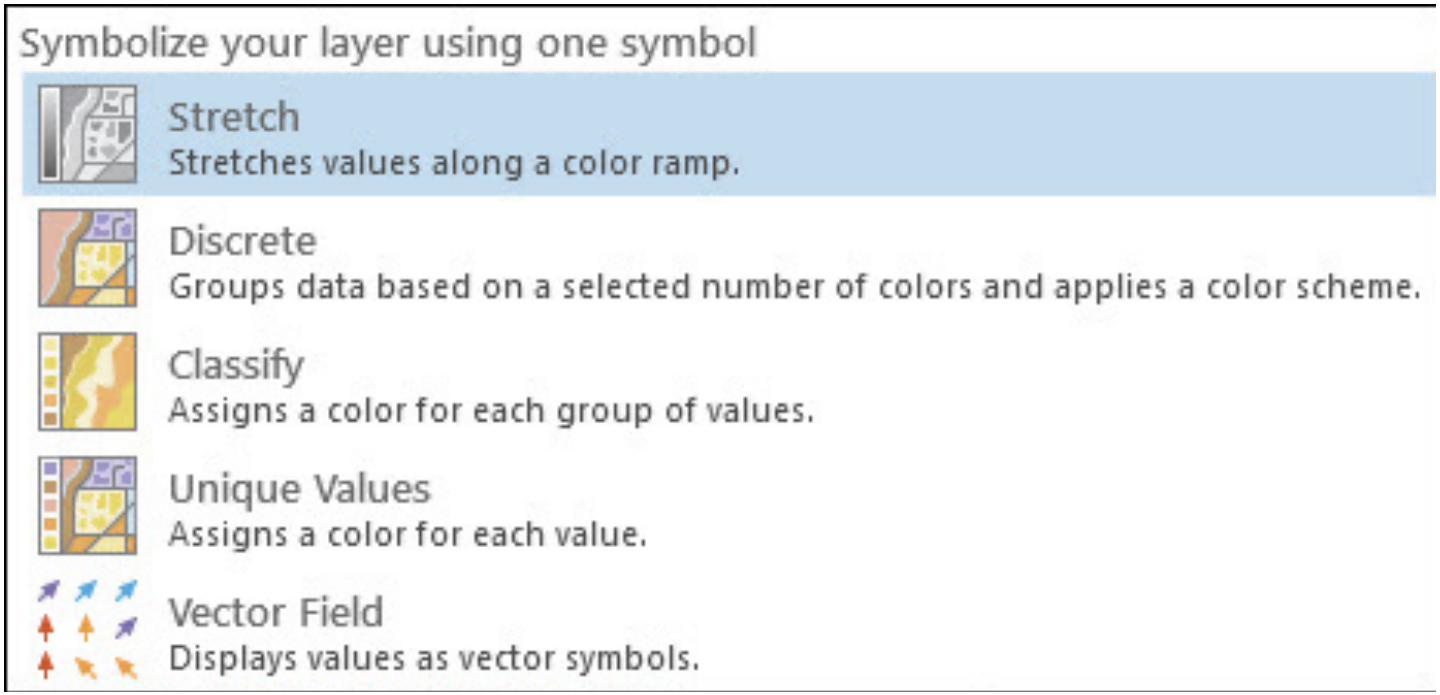


From: *Making Maps: A Visual Guide to Map Design for GIS* by John Krygier and Denis Wood - *makingmaps.owu.edu*

See our tutorial for additional "Cartography Tips".

Colorbrewer provides accessible color options.

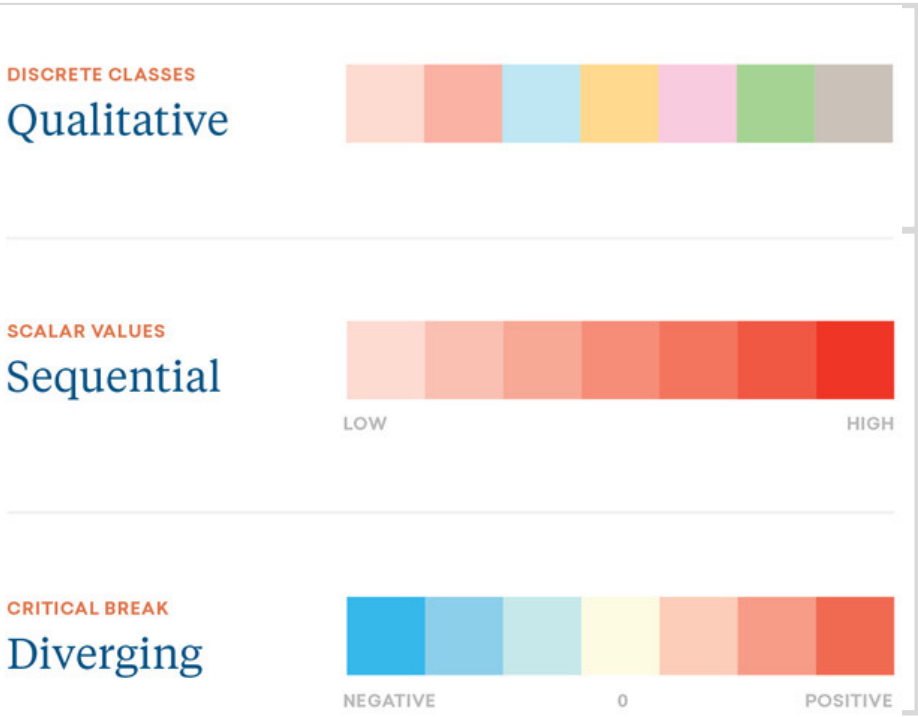
Raster Symbolization



Choosing Color Tips

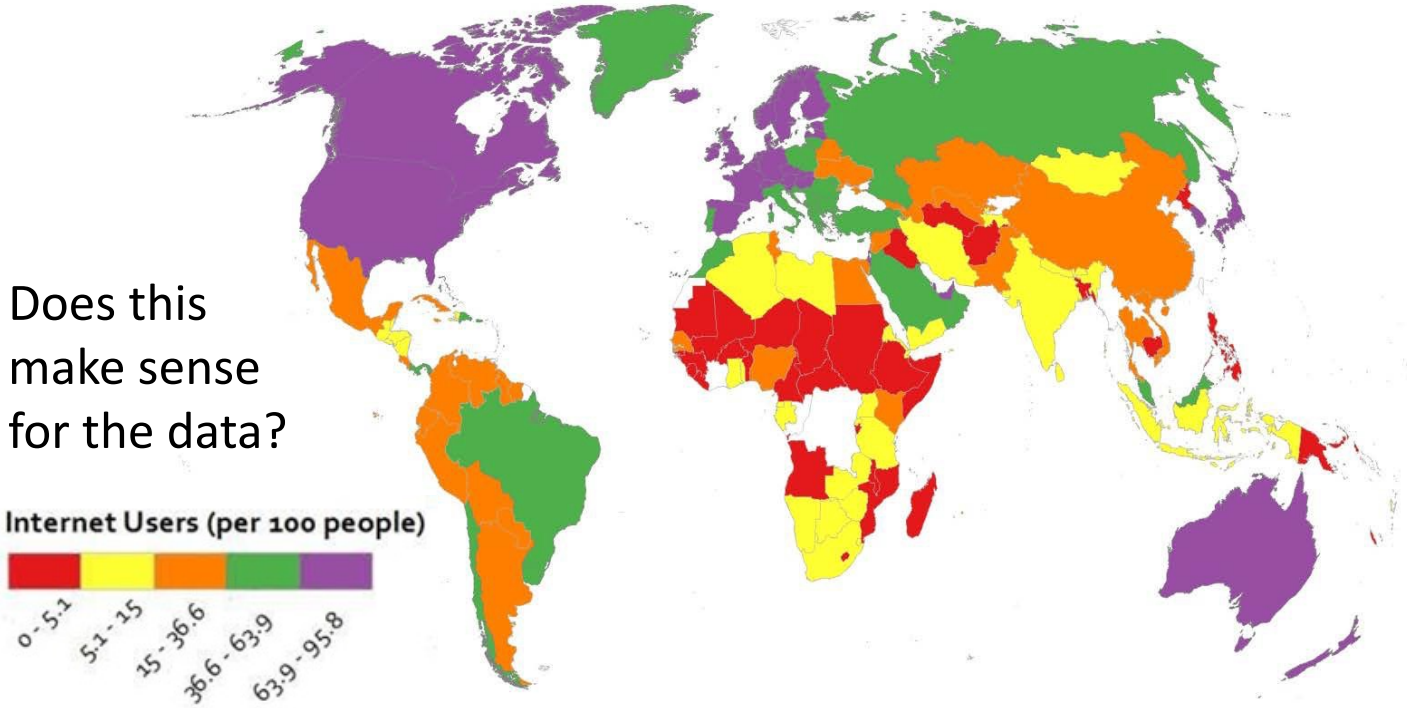
Match the type of data to the type of color scheme:

- **Qualitative** (categories)
- **Quantitative** (numbers)



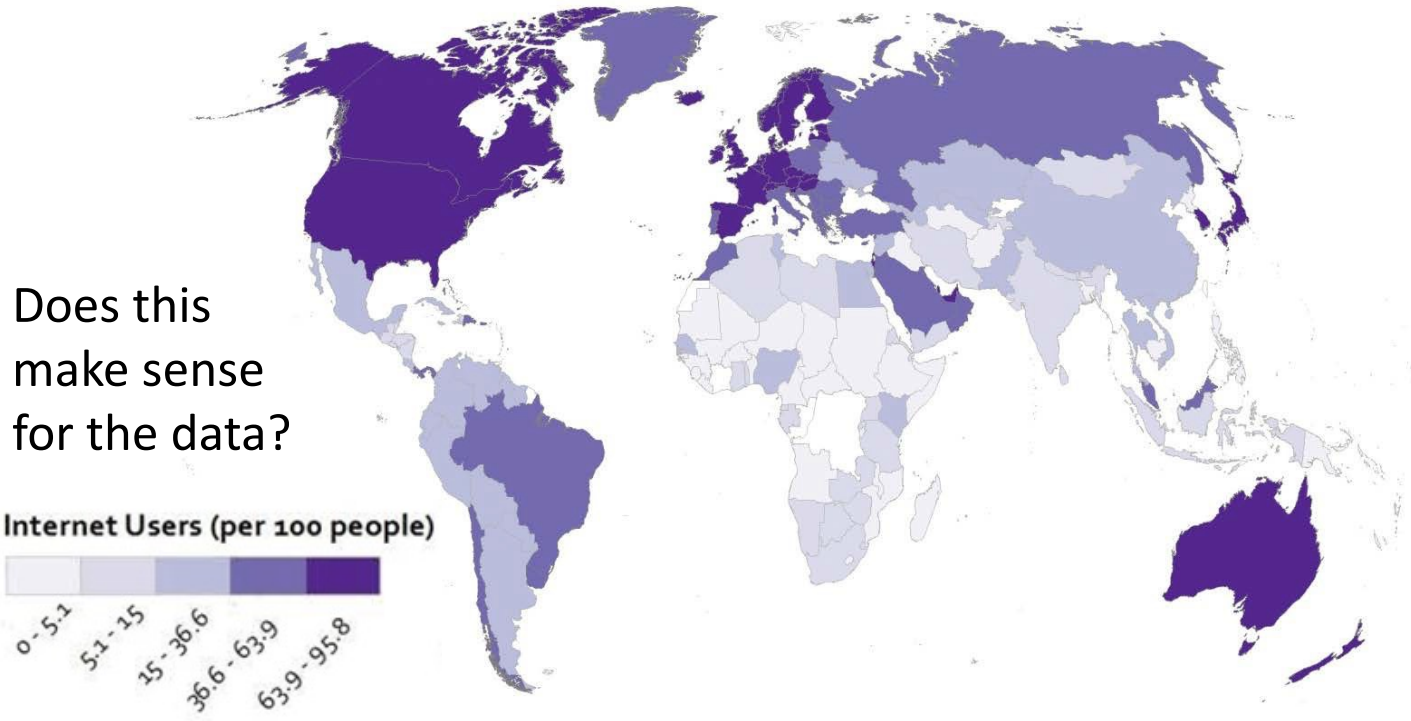
Qualitative Color Example

Does this make sense for the data?



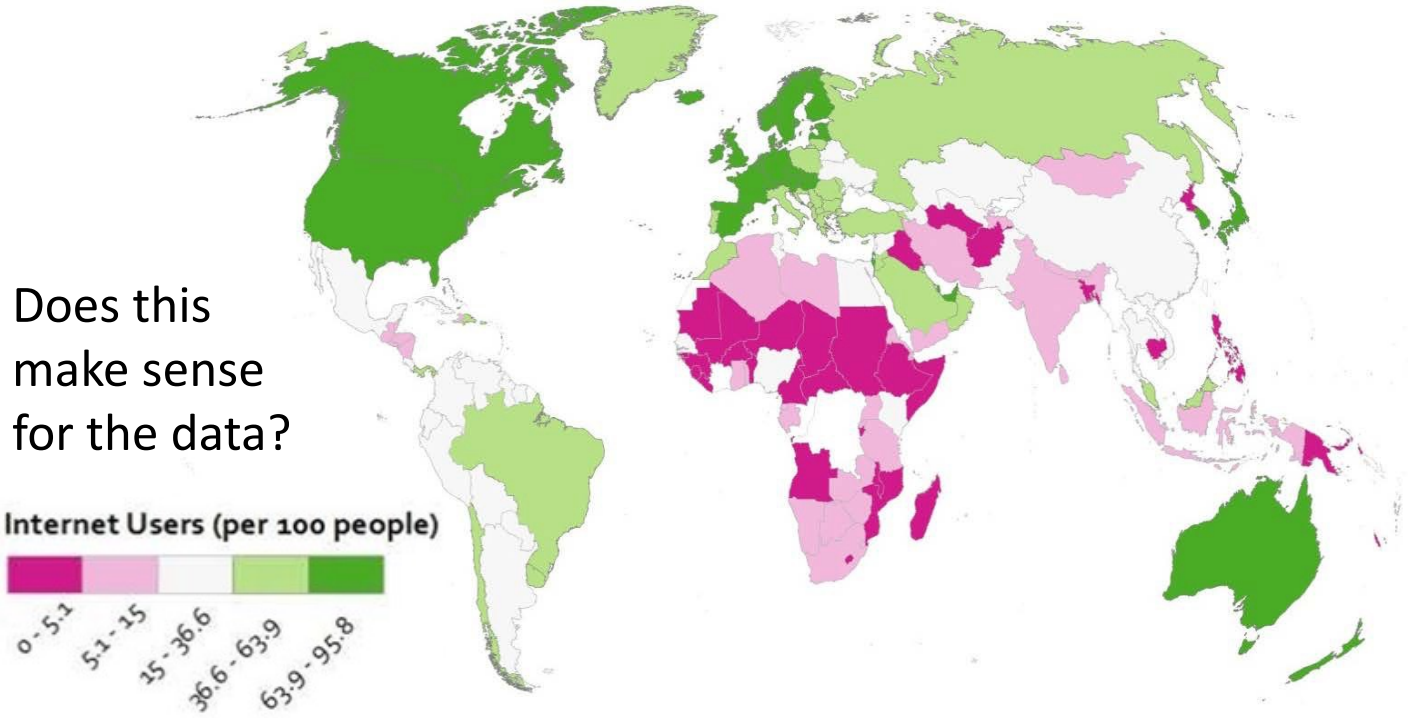
Sequential Color Example

Does this make sense for the data?



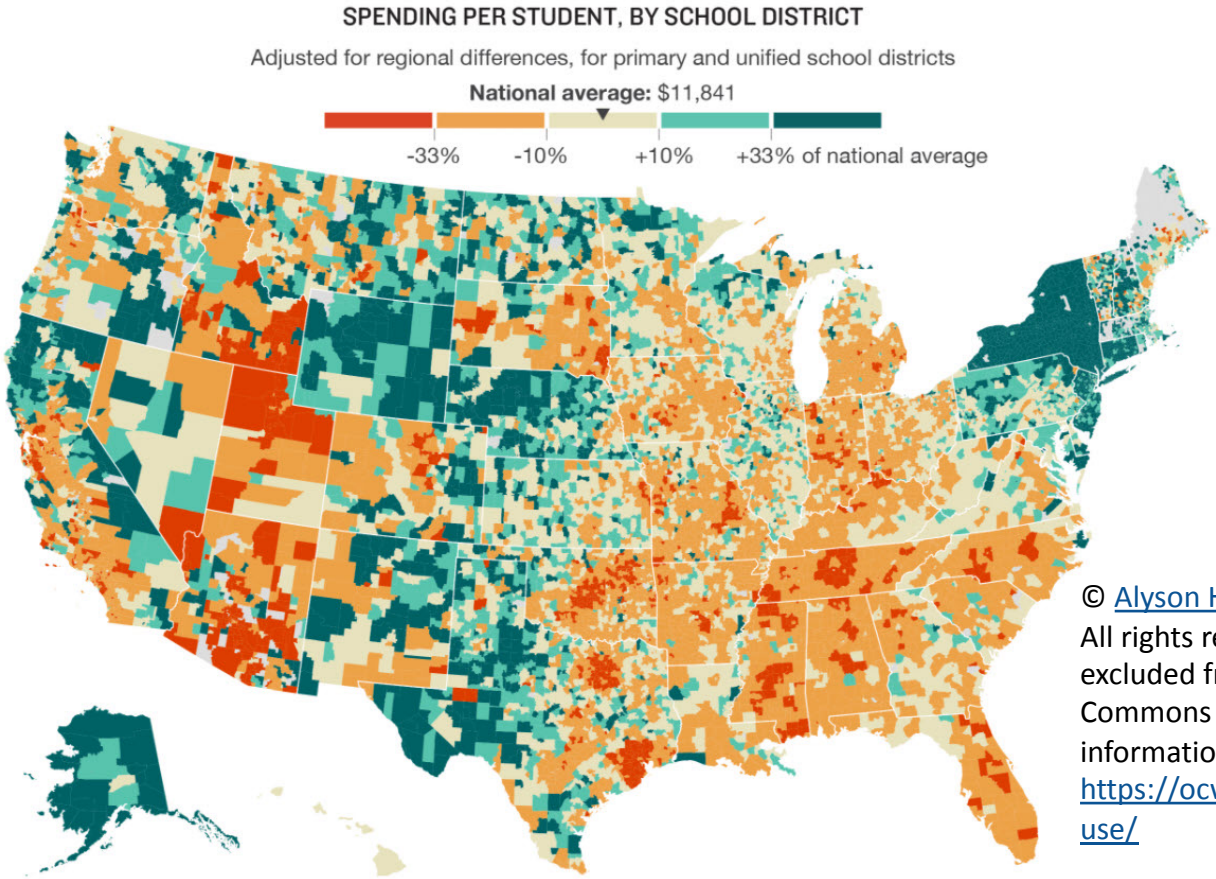
Diverging Color Example

Does this make sense for the data?



Commonly used map type: Choropleth

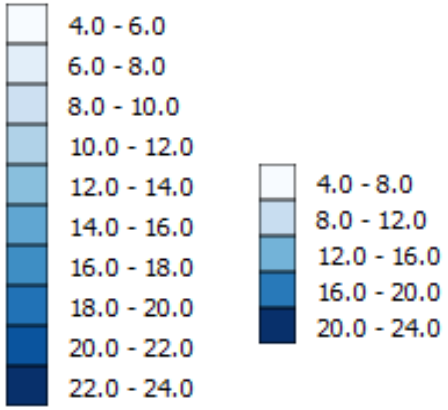
These use different shading and coloring to display the quantity or value in defined areas.



Choropleth map choices

1. Number of Classes

- Aggregates data for display
- More classes = more variation (best to have no more than 7)



2. Classification Method

- Data classification is how data is arranged into separate classes.
- Major types
  - Equal Intervals
  - Quantile (Equal Count)
  - Natural Breaks
  - Defined Intervals

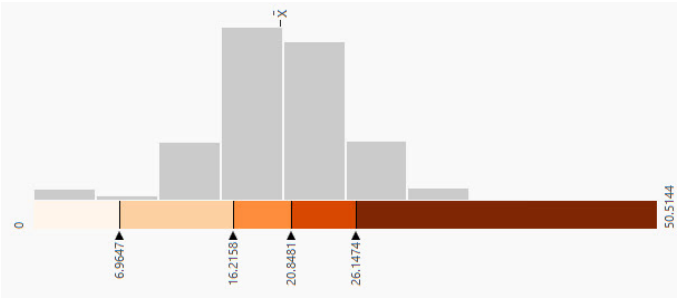
Classification Methods

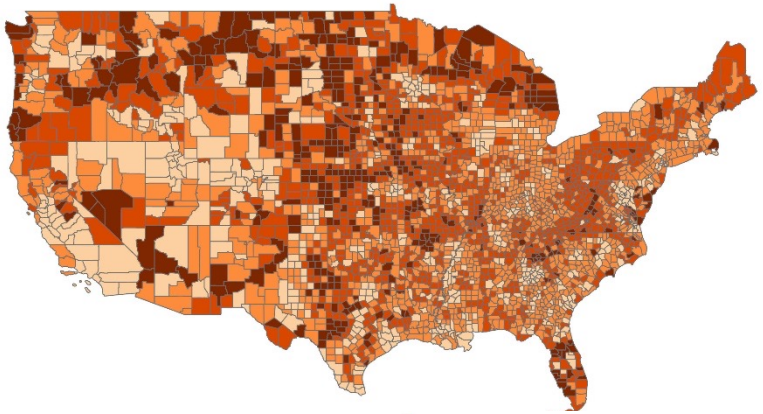
- **Equal Interval** = classes have equal ranges
- **Quantile** = classes have equal counts
- **Natural Breaks** = optimizes class variation
- **Manual** = you define classes

Note: each has pros/cons to their usage,
for "Choropleth Classification Methods" use this link:
https://libguides.mit.edu/gis/tutorials#s-lg-box-wrapper-4119325

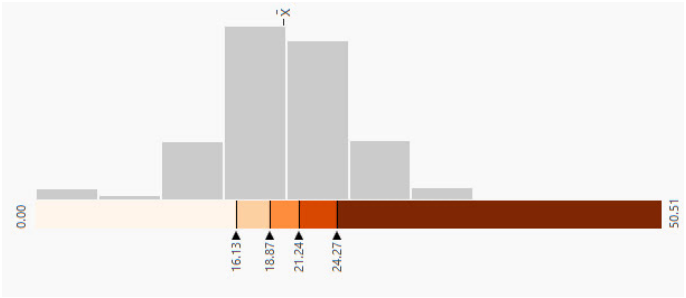
2020 % population over 65

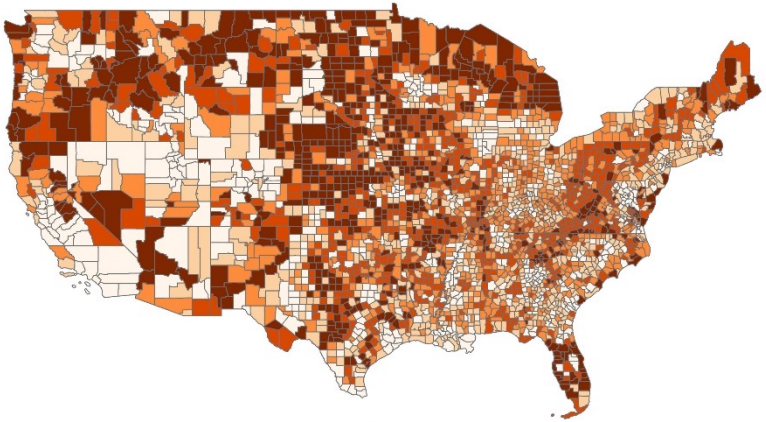
Natural breaks



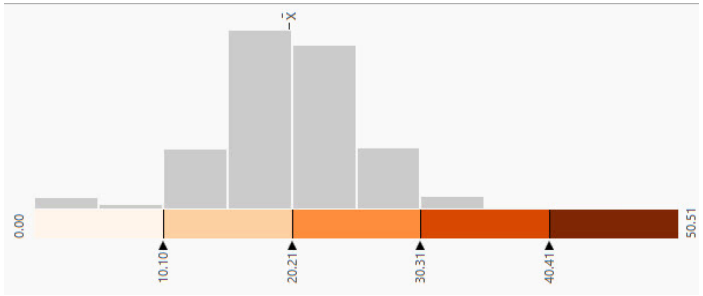


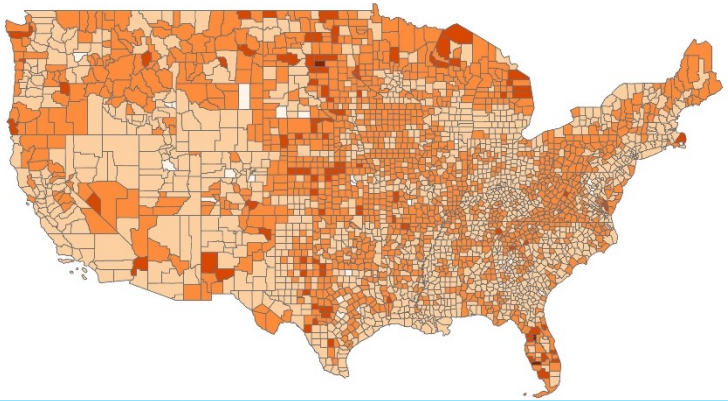
Quantile





Equal interval





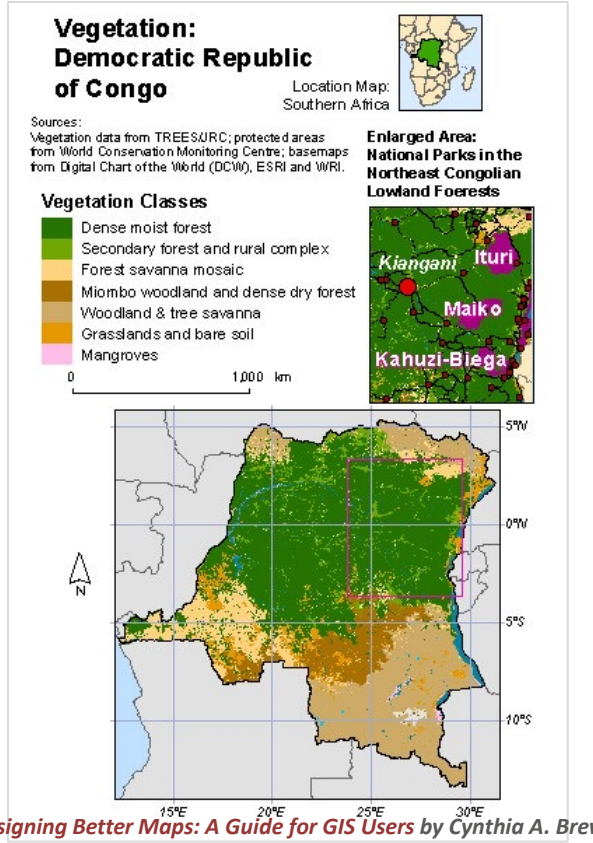
Exercise 2

- Goal:
  - Learn how to symbolize different types of data
- Complete Exercise 2 for either QGIS or ArcGIS Pro.

Map Layout Design Example

Overview:

- Map layout design is about developing a balanced arrangement
- Maps, title, legend, scale bar, labels, etc. all need relative positioning & sizing
- Goal is to design the map layout to support your design questions
  - Who wants the map
  - Where will it be seen
  - What is its purpose

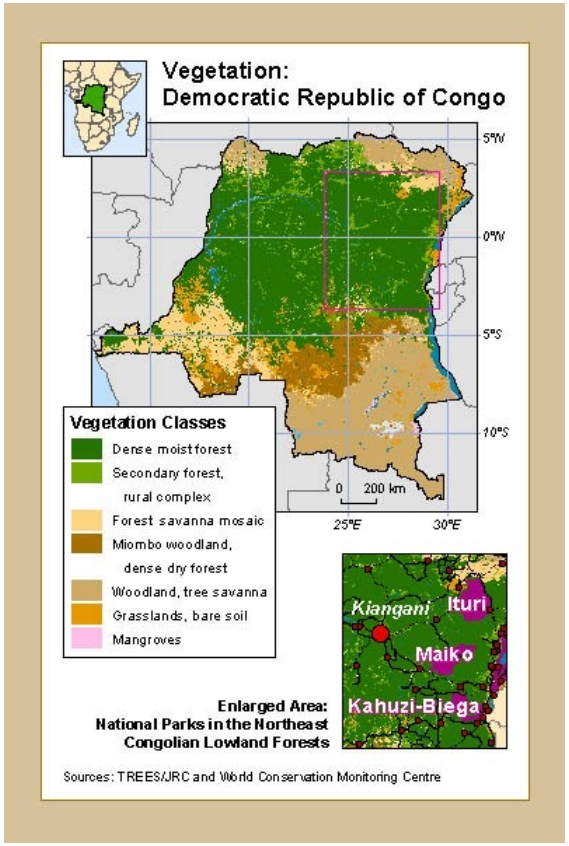


From: *Designing Better Maps: A Guide for GIS Users* by Cynthia A. Brewer

Map Layout Design Example

Tips:

- Inset/locator maps are often placed in the top/bottom corners (e.g. continent view top left and zoomed view in bottom right).
- Main map often placed in center (usually largest & most detailed).
- Legend is tucked into the main map for easy comparison with the data.
- Scale bars and north arrows shouldn't be a distraction from the main map.
- Sources should run along the bottom.



From: ***Designing Better Maps: A Guide for GIS Users*** *by Cynthia A. Brewer*

Complete the take-home exercise to learn more.

TAKE-HOME EXERCISE

Exercise Overview

Query and use unemployment and transportation data to create a map that helps you decide where to build a mixed use facility.

1. Navigate the software interface
2. Find and add data, including basemaps
3. Access and explore attribute information
4. Symbolize data layers, for vector and raster
5. Select data by attributes and spatial location
6. Design a simple map for export

Boburg, S. (2017, May 31). How Jared Kushner built a luxury skyscraper using loans meant for job-starved areas. *Washington Post*. Retrieved from

MIT OpenCourseWare
https://ocw.mit.edu

RES.STR-001 Geographic Information System (GIS) Tutorial
IAP 2022

For information about citing these materials or our Terms of Use, visit:
https://ocw.mit.edu/terms