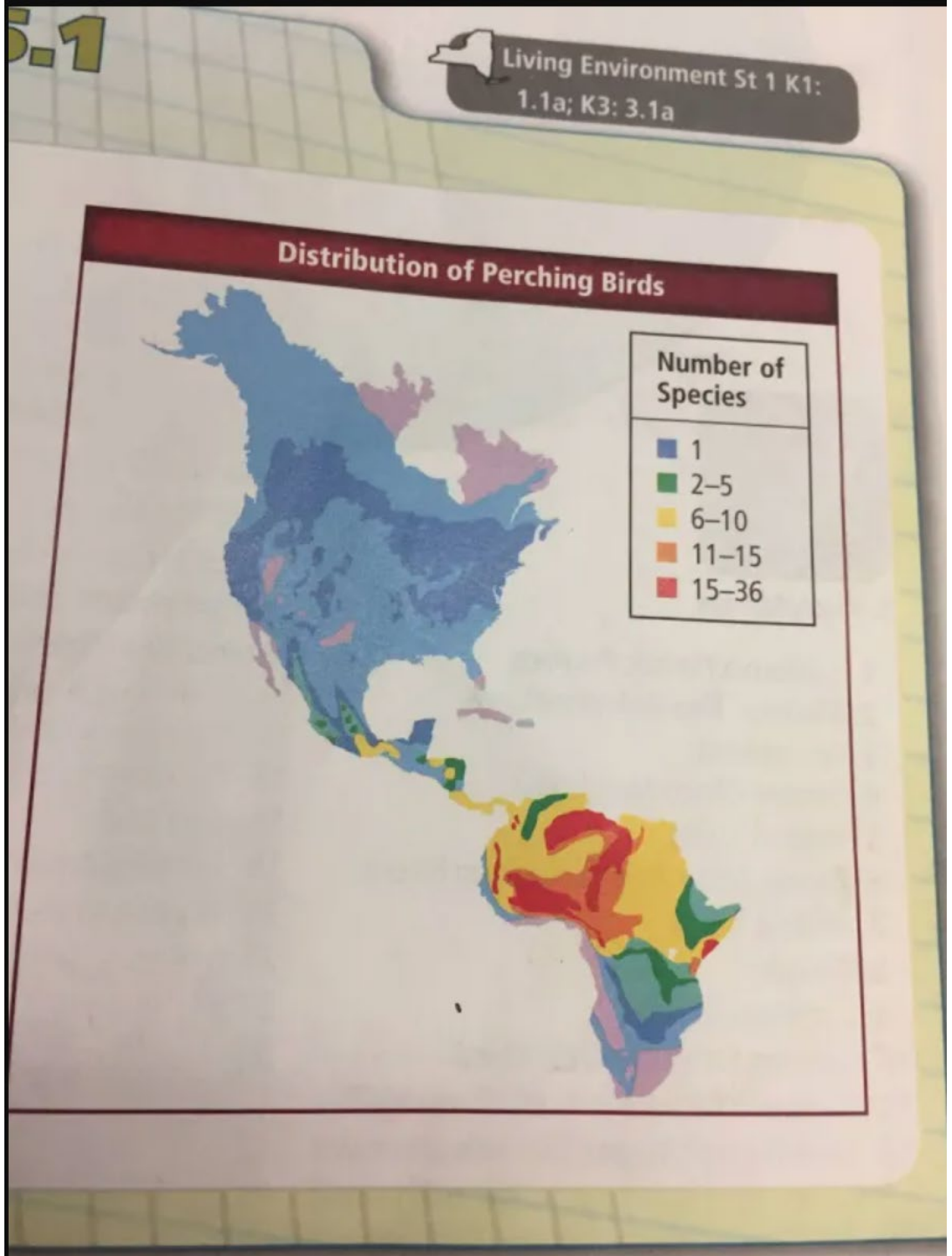
Deck 2: What is GIS, anyway?

Cartography, Data Formats, Querying, and Shape file Basics

Intro to GIS – UMass Amherst – Michael F. Nelson

Gallery of the Absurd



Overview

GIS's roots in Cartography

- What is a Map?
- What is Cartography?
- Basic Map Design
- Thematic Maps and Data Scales

Final Projects: Ideas

Thinking Spatially: How GIS Works

- Input
- Database Management
- Analysis
- Output

Cartography

Maps: Relevance to GIS

- **GIS output is often a map**
  - People produce more maps than ever now that GISs are available and accessible.

Maps: Relevance to GIS

- **Representational issues are very important for that reason**
  - Need for cartographic principles and understanding is increased by the availability of GIS.
  - GIS is helping to broaden the forms of maps we can feasibly produce.

What is a Map?

'A map is a graphical representation of the milieu'

Milieu means **environment,** and in this context broadly includes all aspects of **the cultural and physical environment**.

Here, milieu also includes ***mental abstractions*** that are not physically present on the geographic landscape (e.g. people's attitudes) yet can clearly be mapped.

A. H. Robinson and B. Bartz Petchenik. 1976. The Nature of Maps: Essays Toward Understanding Maps and Meaning. Chicago: University of Chicago Press. pp. 16-17.

Mapping is Model Thinking

A Model is a simplified version of reality. It is an ***abstraction****!*

All models are wrong, some models are useful.

Einstein's famous quote: "*Everything should be made as simple as possible, but not simpler*"

Cartography: The art and science of making maps

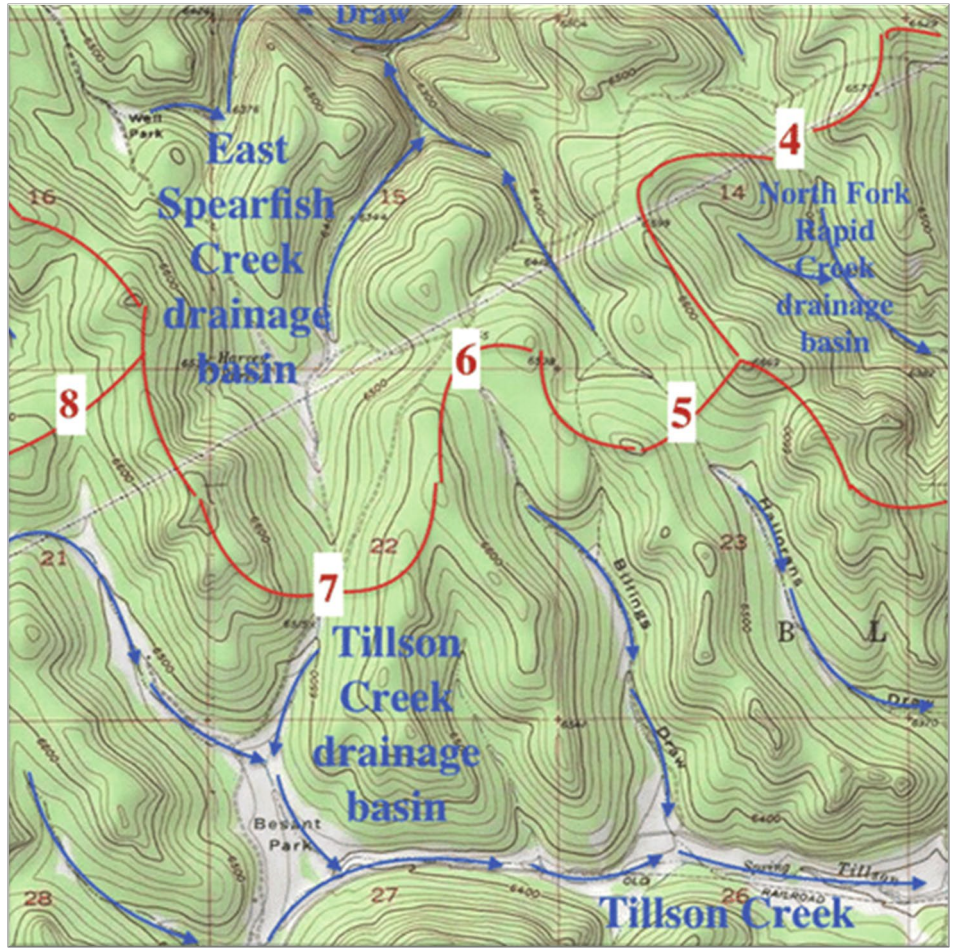
Goal: Communicate geographical information graphically

Conflict: There is an inherent conflict between maximizing information content and ease of understanding

Map Design: Purpose

General-Reference Cartography

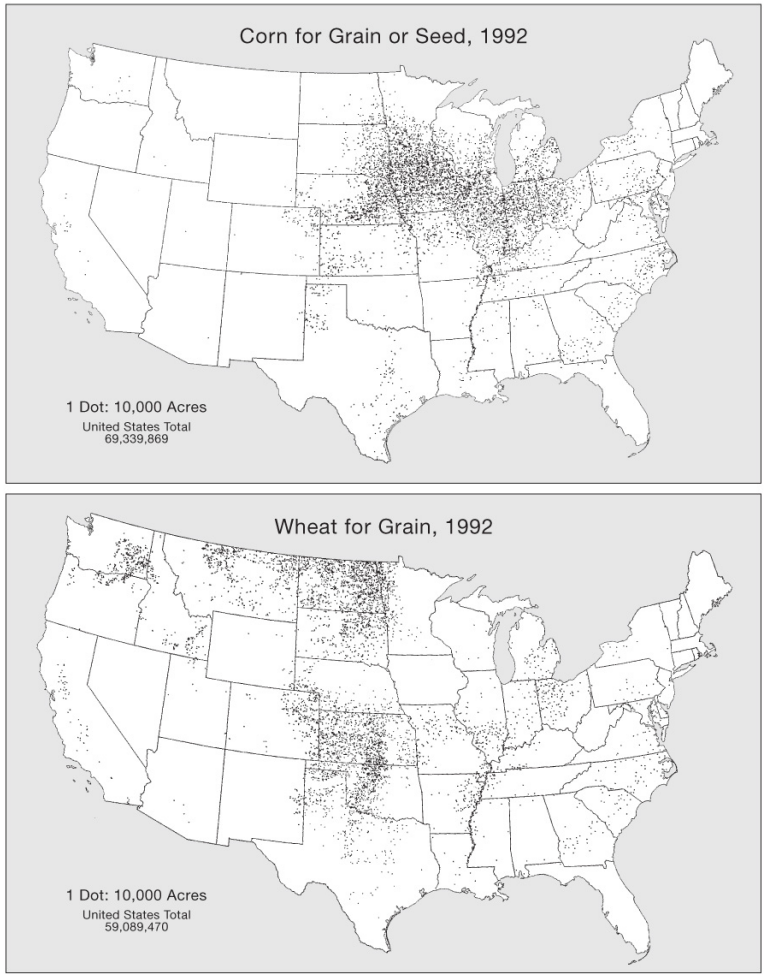
- Focus is usually on geographic location, not explicitly displaying data.
- Emphasize the relative location of spatial phenomena.
- E.g. : USGS Topographic Map
  - But these *do* show elevation!



Thematic Cartography

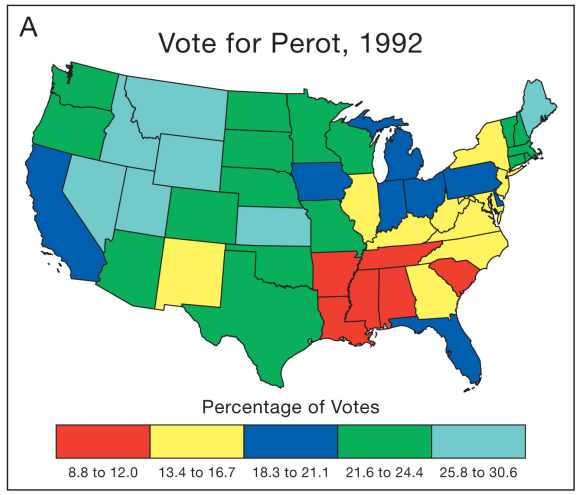
- Emphasize the spatial pattern of **one or more** geographic attributes.
- Example: Dot density
  - Higher density of dots = greater agricultural production
  - Note: dots do not represent farms!

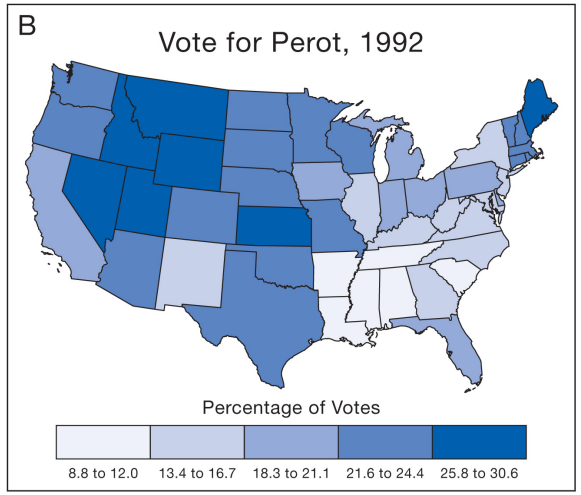
A dot-distribution map



Thematic Cartography: Choropleths

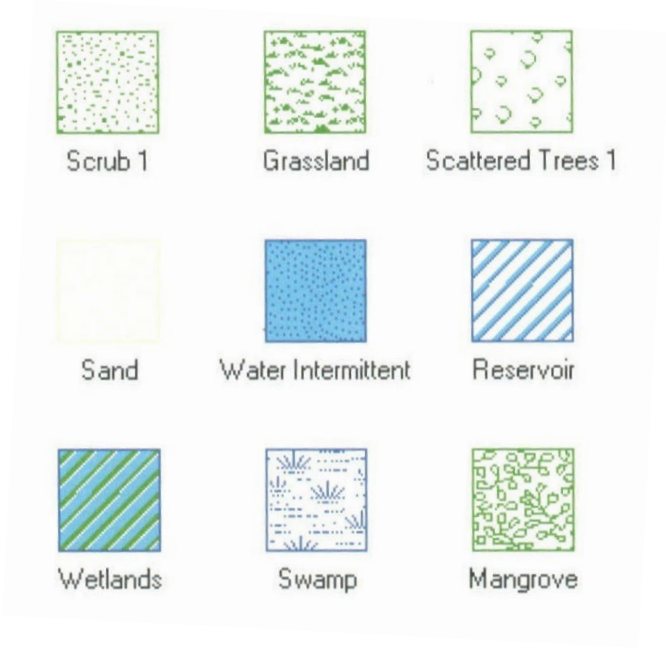
- Use color to emphasize theme: population density, family income, daily temperature maximums, etc.
- Choropleth: color is proportional to a numerical value
  - Value (from HSV – more info later) is proportional to Perot support.
  - Color becomes **thematically informative**

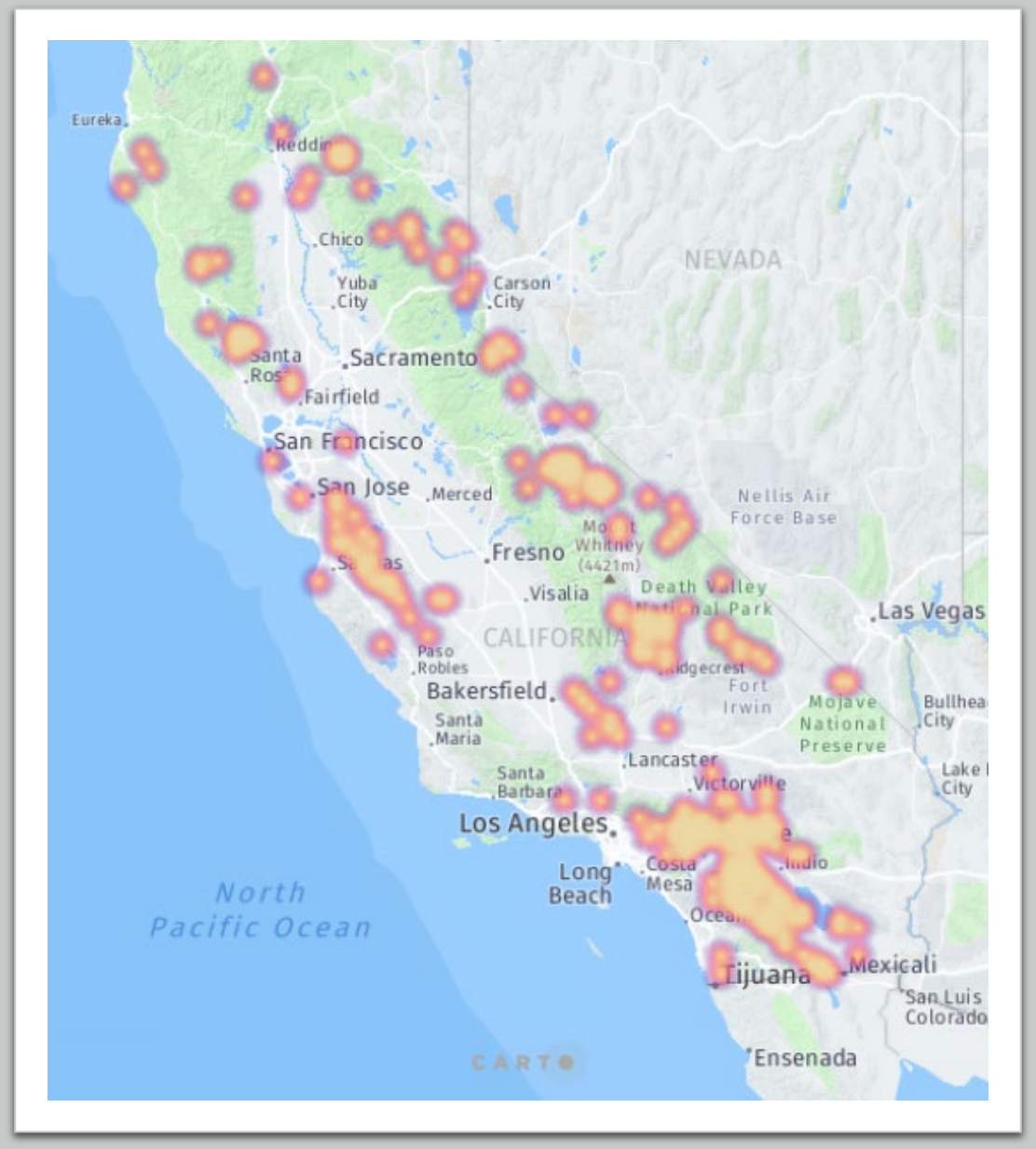




Map Design: Color Harmony

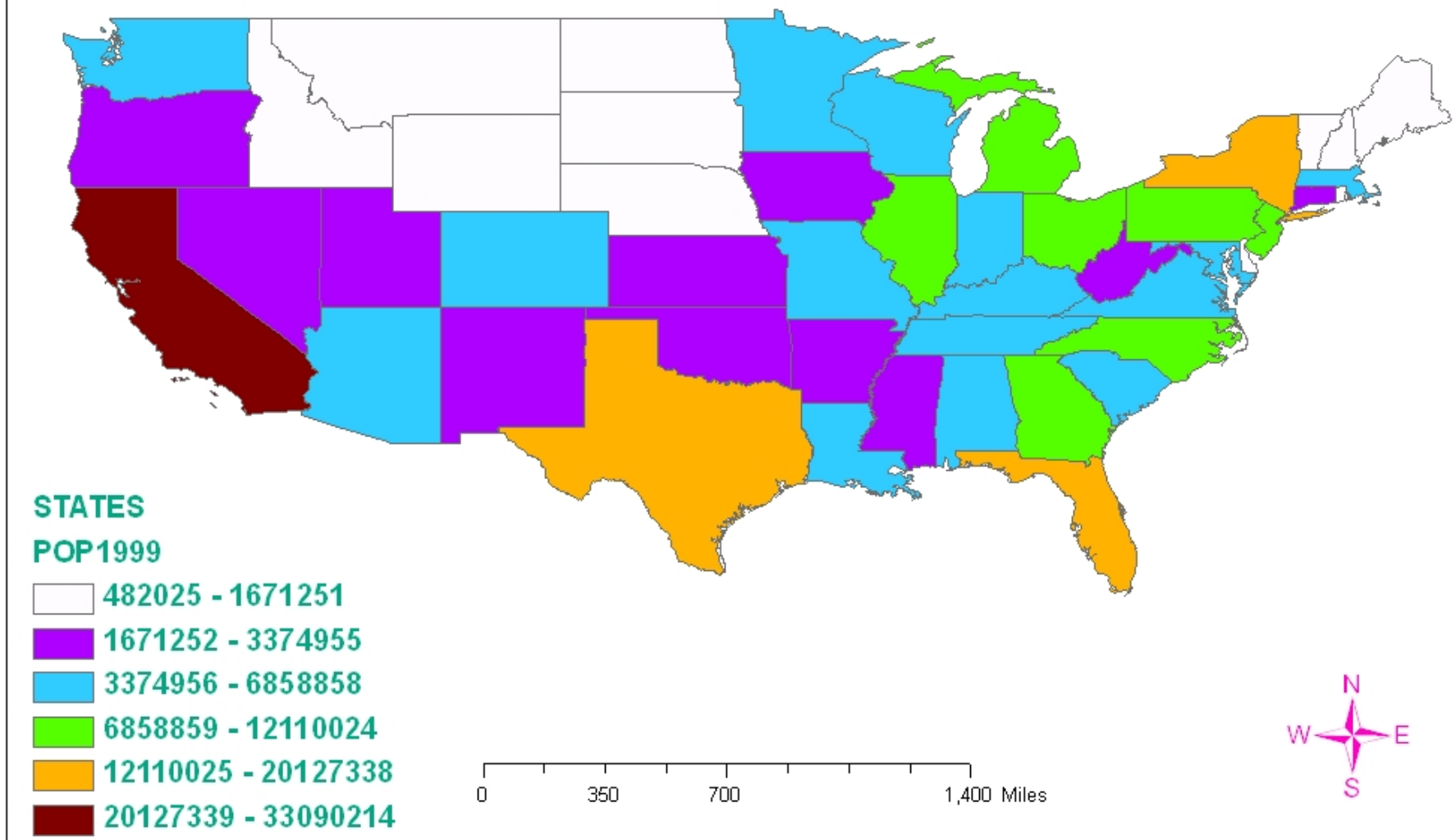
- Harmony of color/texture
- Arc can help!





Population Trend

Bad Harmony Map



Map Design: Color Models

Color Models

RGB

- Three color channels
  - Additive effect
  - Red, Green, Blue
- Simulates our eye physiology
  - Cone cells are differentially stimulated by different wavelengths.
- Computer screen pixels have R, G, and B light emitters.

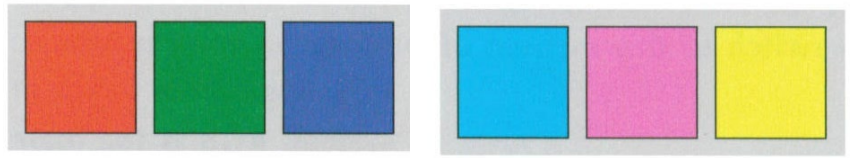
Hue Saturation Value (HSV)

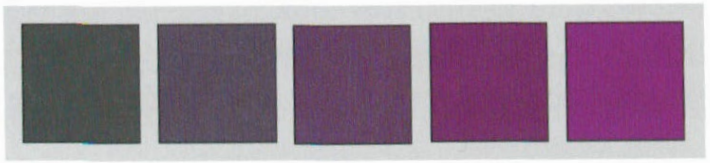
- **H**ue: color quality
  - Red, yellow, brown, etc.
- **S**aturation: amount of gray in the hue
- **V**alue: brightness

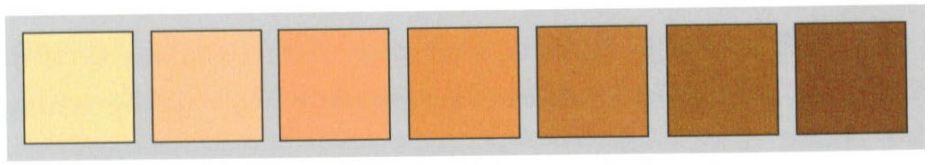
Harmony: Color Choices

HSV

- Hue
  - Color ‘quality’
- Saturation
  - aka ‘vividness’
- Value
  - aka Light/Dark





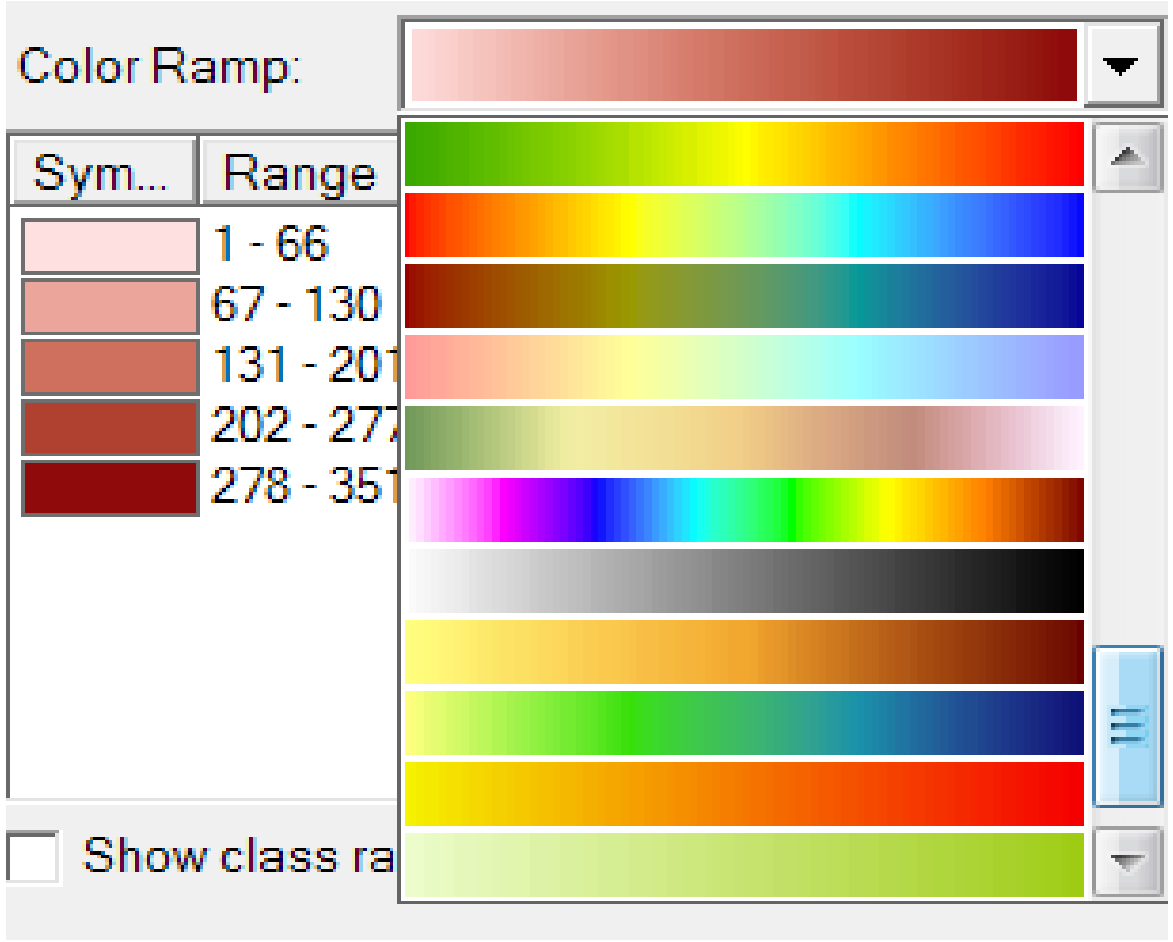


Harmony: Color Ramps

A color ramp maps values to colors (hues).

Arc has many built-in color spectra, a.k.a. color ramps.

Arc Symbology:

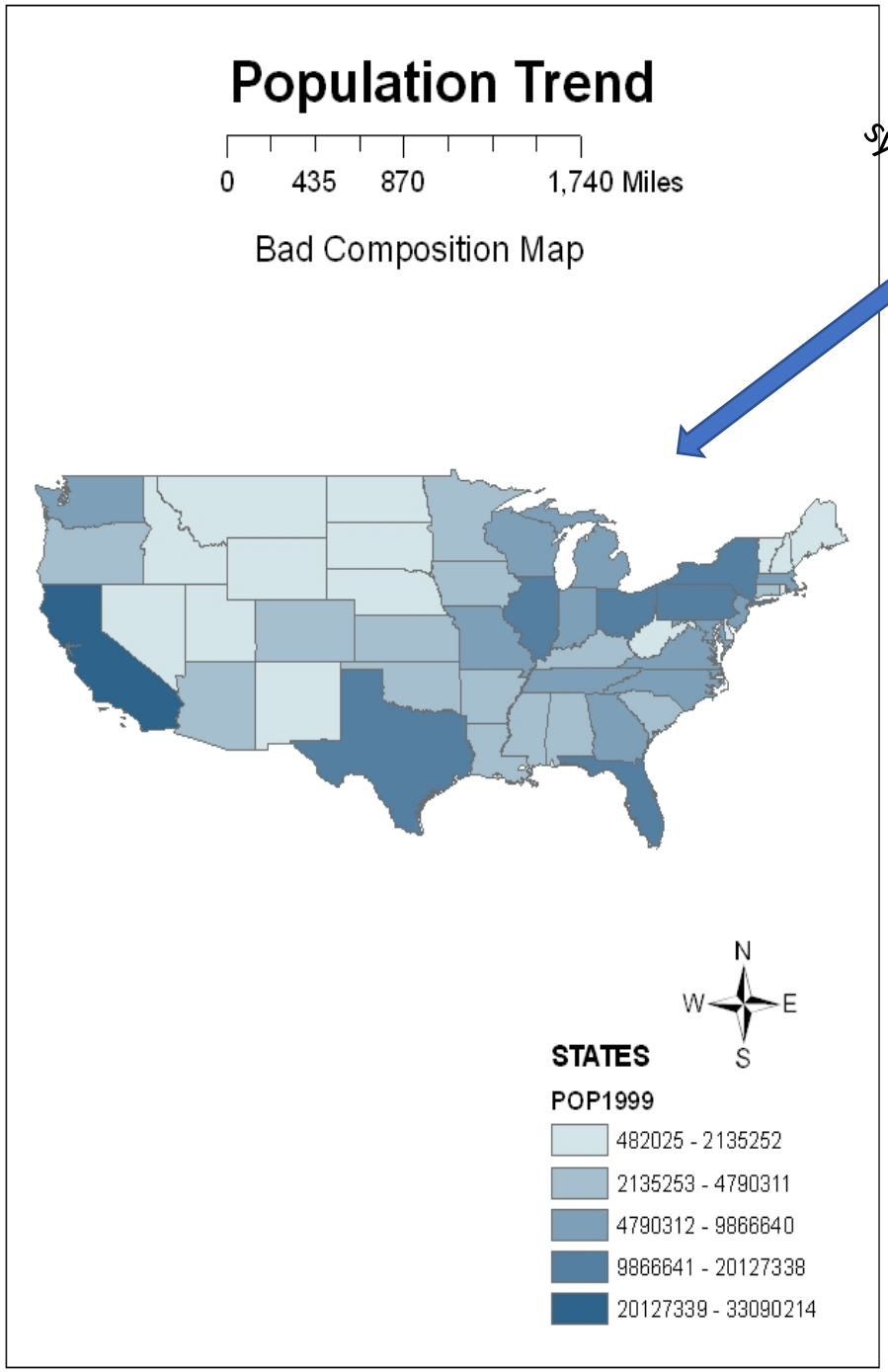


Map Design: Composition and Clarity

Composition refers to the relative positions of elements on a map

Composition: Avoid excess white space

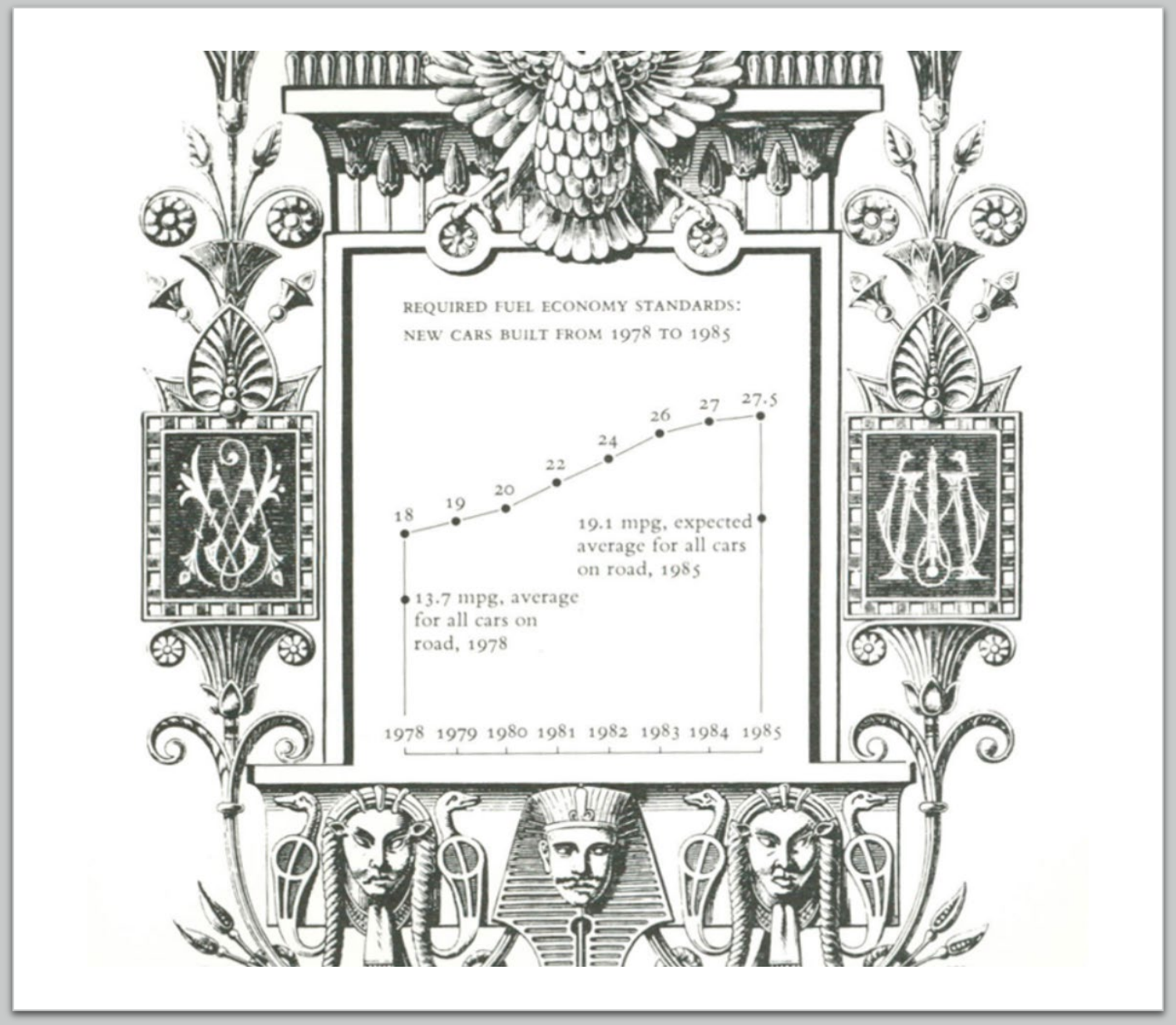
Like this…



…or this

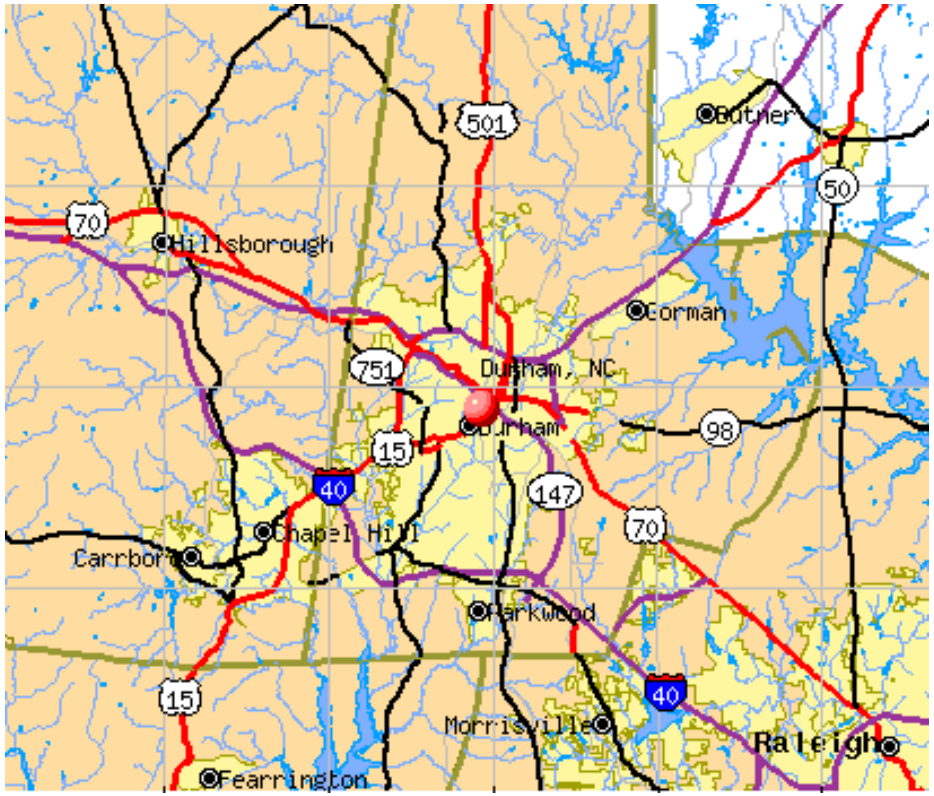
Composition and Clarity: Reduce Distractions

- Do not decorate your results!
- Make your results the focus.



Map Design: Clarity and intuition

- Use intuitive colors schemes and appropriate cropping/zoom level.



Clarity: Font Choices

- Serif (e.g., Times New Roman)
  - The quick brown fox jumped over the lazy dogs
- Sans-serif (e.g., Arial, Calibri)
  - The quick brown fox jumped over the lazy dogs
- Display: Impact, Forte, Comic Sans (gasp)
  - *The quick brown fox jumped over the lazy dogs*

Thematic Maps and Data Scales

Four Important Data Scales

Categorical: Nominal

Non-numeric, discrete, unordered

Categorical: Ordinal

Non-numeric, discrete, has an intrinsic order

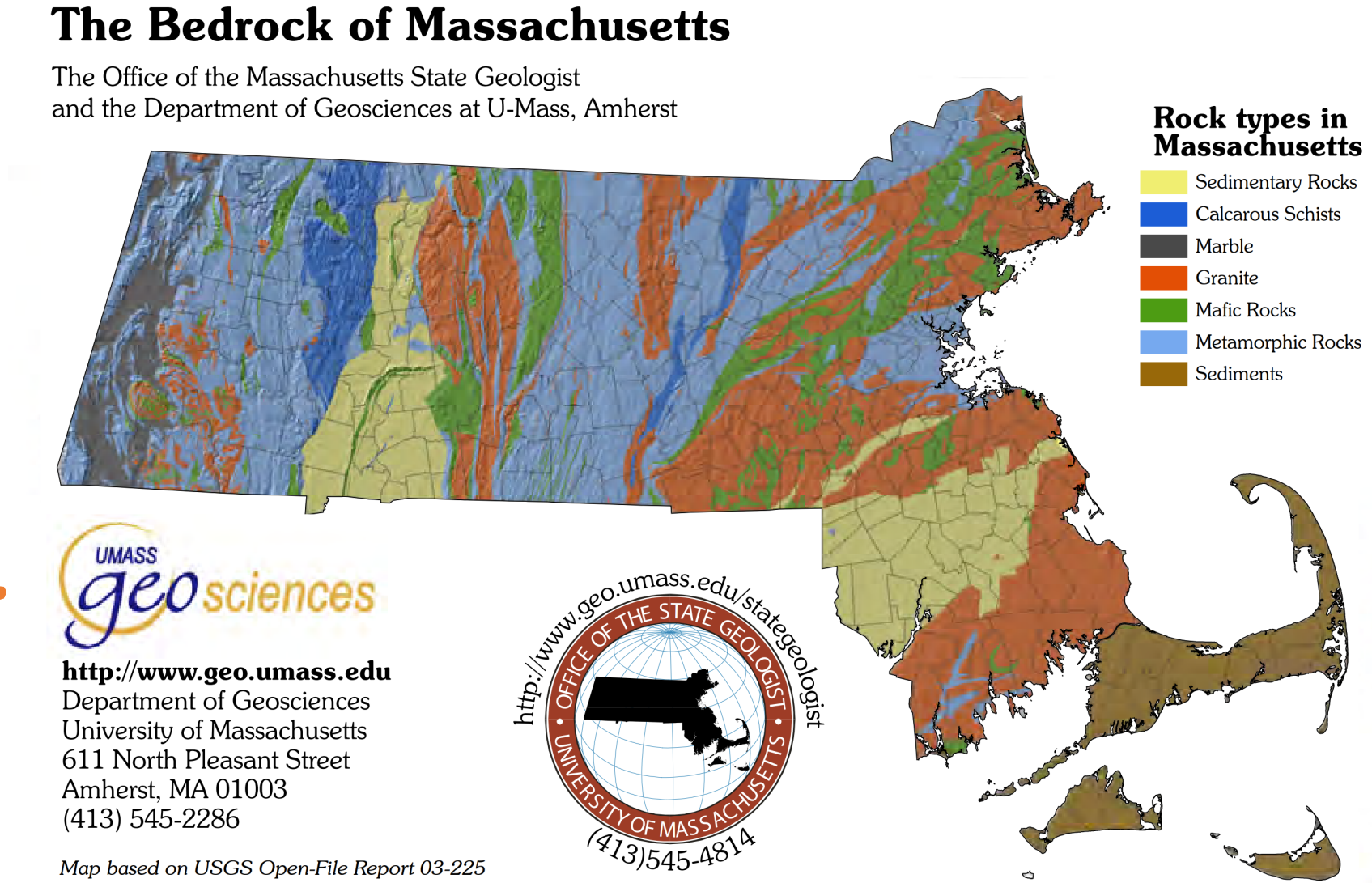
Numeric: Interval

Has a moveable zero

Numeric: Ratio

Has a meaningful zero

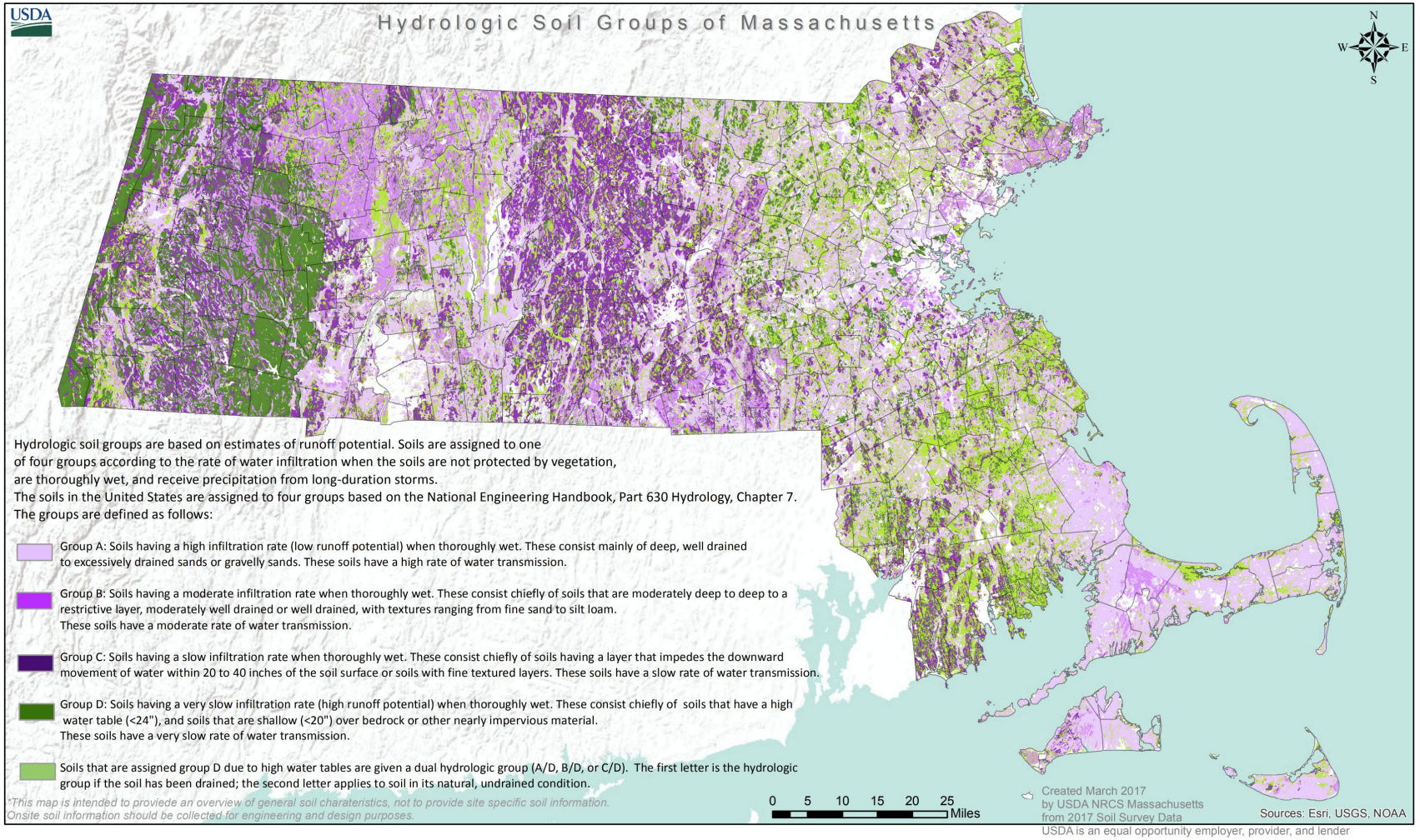
Nominal Data: Bedrock Map



Ordinal Data: Hydrographic Map

Groups are intrinsically ordered by infiltration rate.

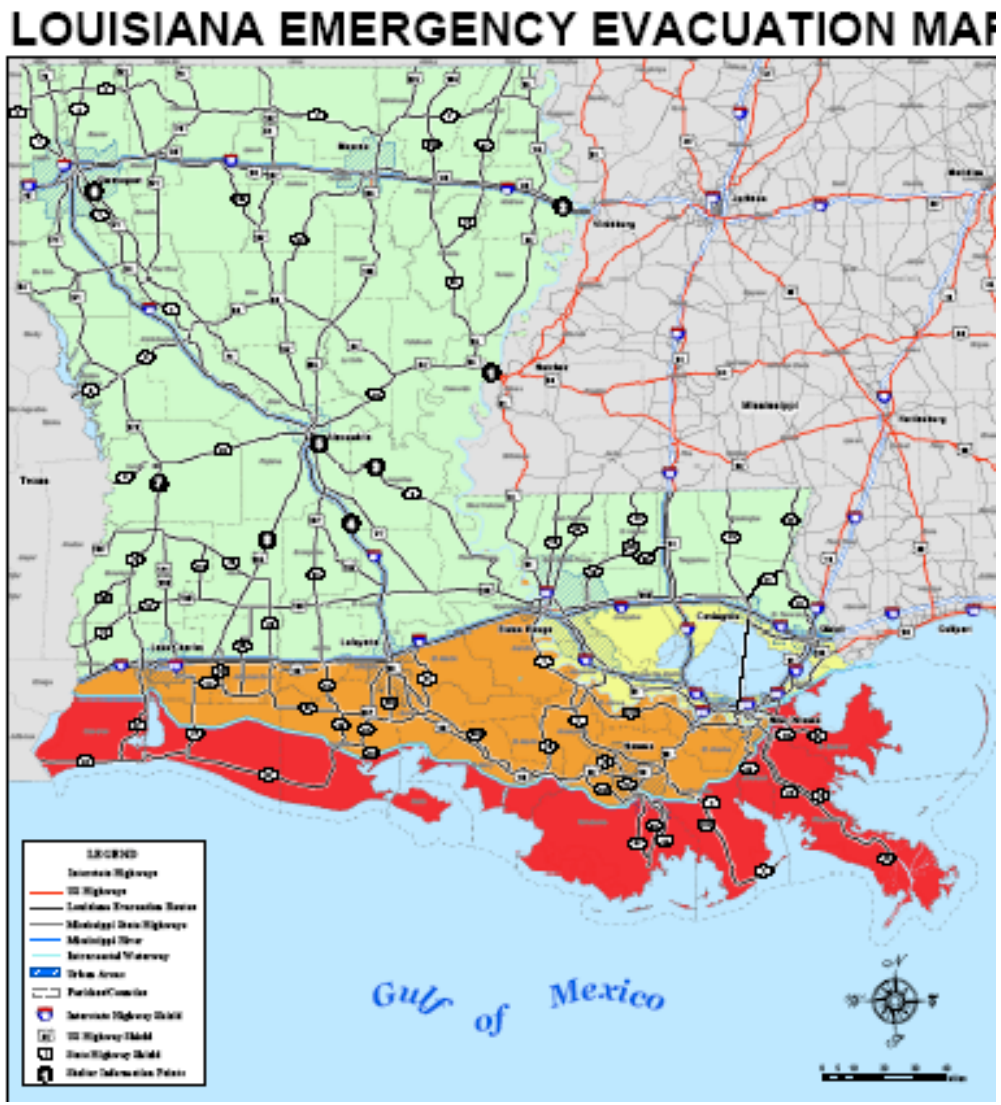
Color qualities are points along a gradient.



Ordinal Data: Hazard Map

Categories, but with an intrinsic ordering.

- Categories are not numeric. Why not?

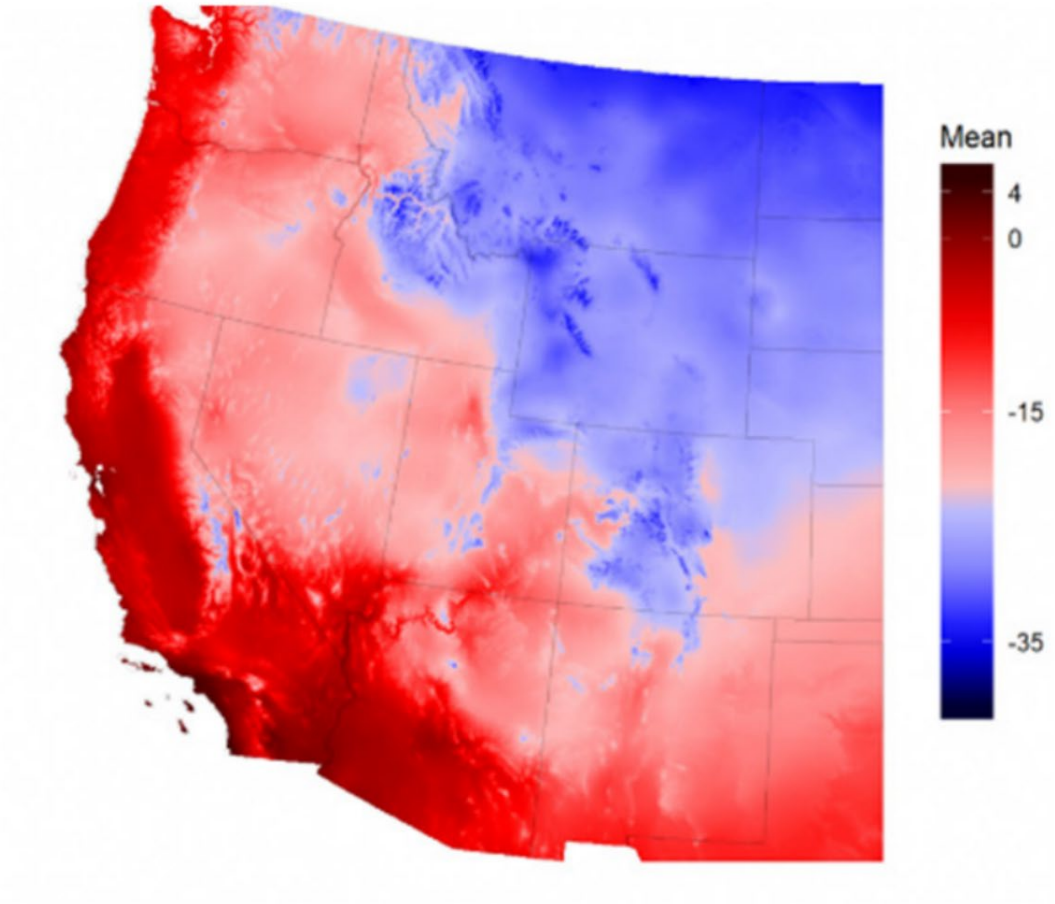


Interval Data: Choropleth

Color quality (hue) or intensity is scaled to a numeric quantity

Average minimum winter temperature, Celsius scale.

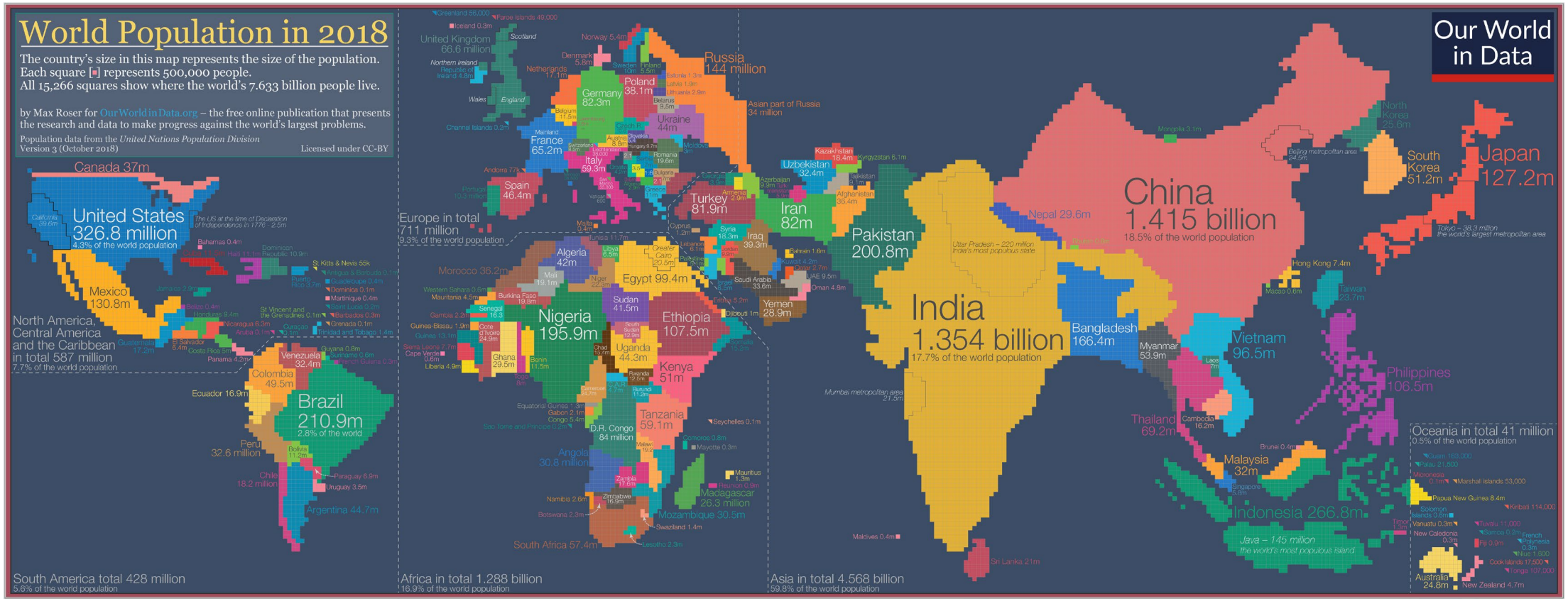
How could we convert this to a ratio scale?



Bone and Nelson, 2019. Forests

Ratio Data: Cartogram

Cartogram: Area is proportional to a quantity like population.



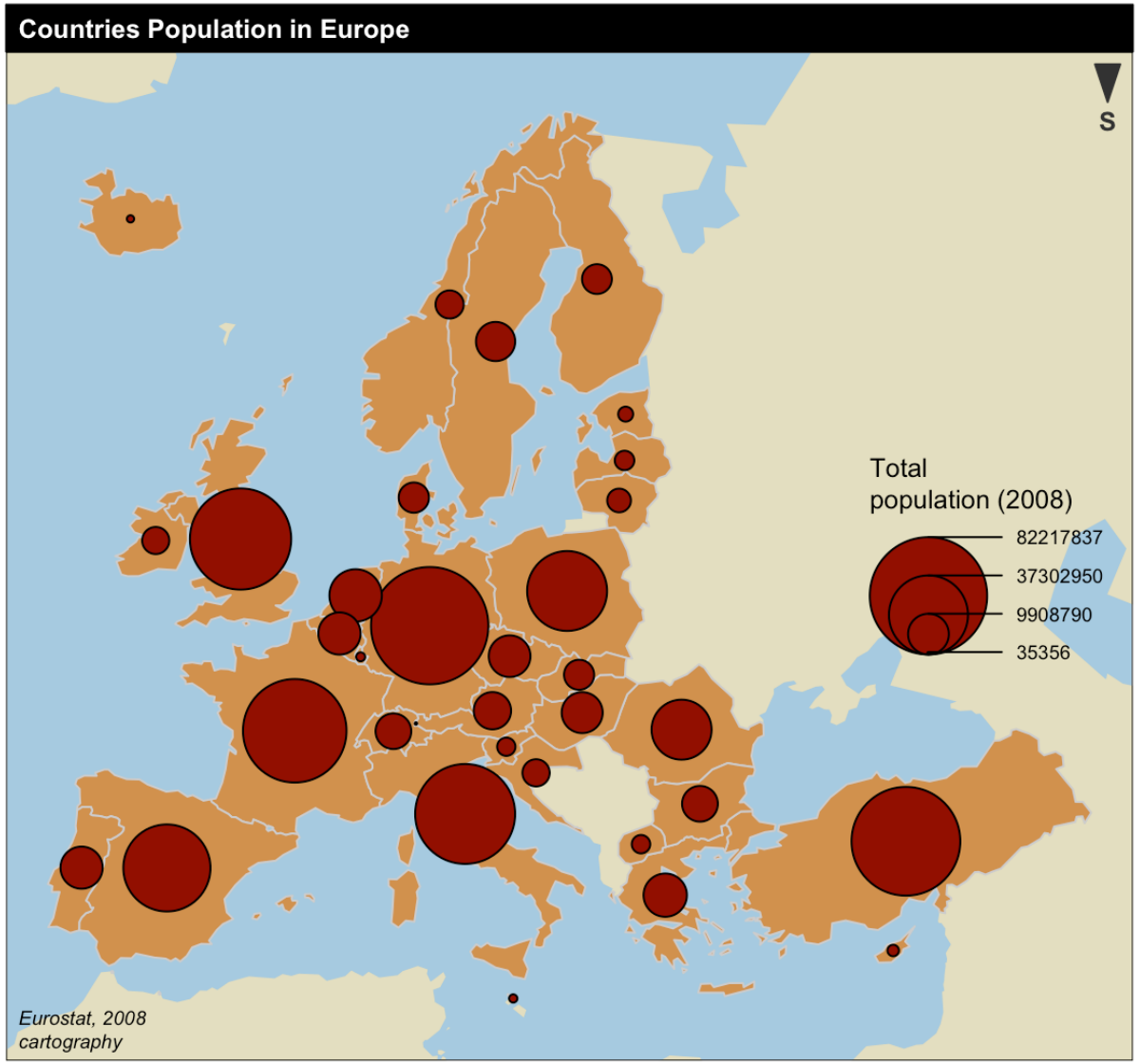
By Max Roser - https://ourworldindata.org/world-population-cartogram, CC BY 4.0, https://commons.wikimedia.org/w/index.php?curid=83545890

Ratio Data: Proportional Symbols

Circle area is proportional to a numeric quantity.

How useful is this representation for interval data?

- Hint: what about negative values?



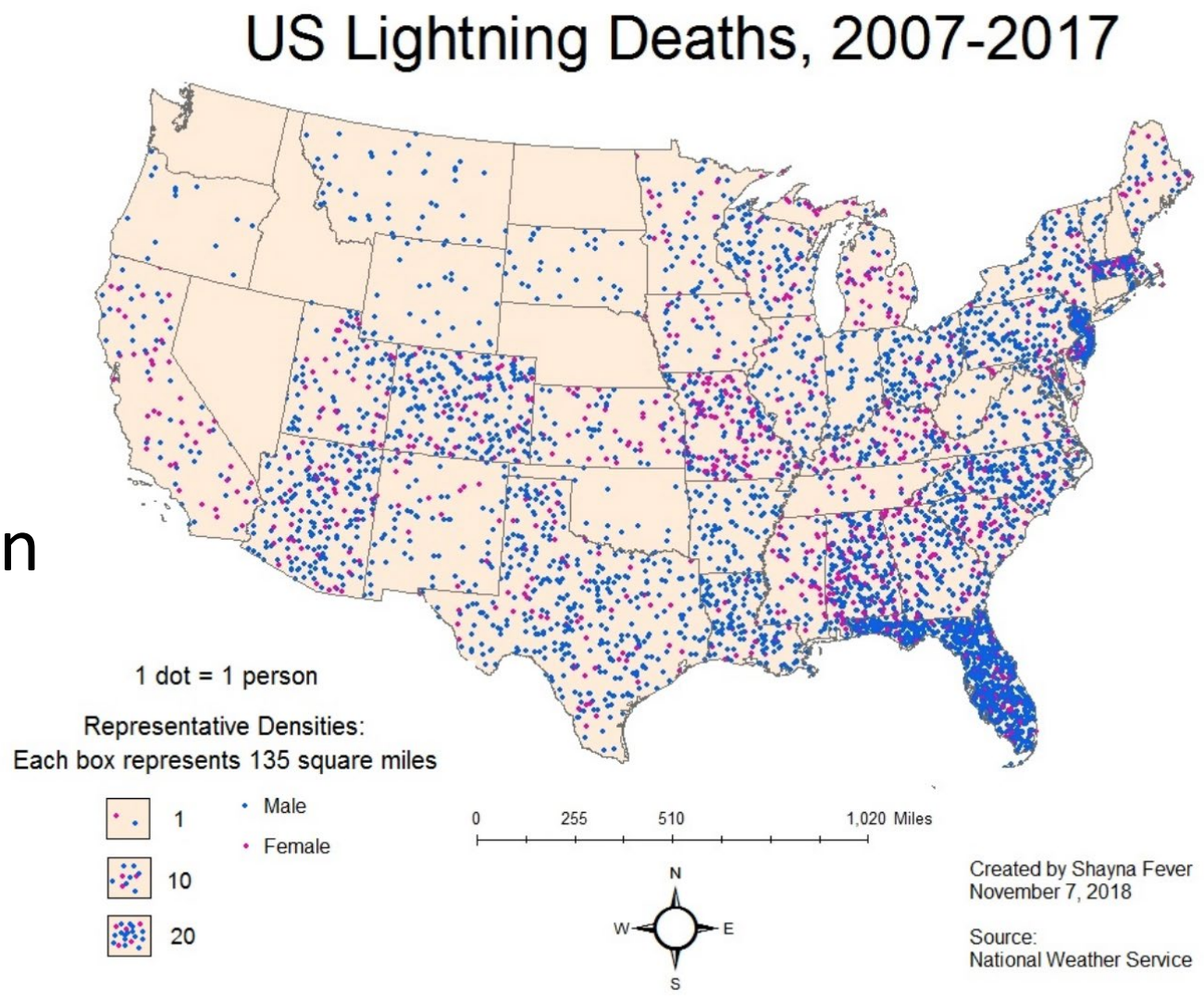
https://r-graph-gallery.com/177-map-with-proportional-symbols.html

Ratio Data: Dot Density

Number of dots/area is proportional to a numeric quantity.

Any opportunities for confusion with this map type?

More male deaths? What's up in Missouri?



Numeric Data and Map Types

Choropleth

- Maps quantities to colors
- Good for **interval** data

Cartogram

- Maps quantities to sizes
- Good for **ratio** data

Dot Density

- Density of dots corresponds to value
- Good for **ratio** data

Proportional Symbols

- Size corresponds to value
- Good for **ratio** data

Final Projects!

Already???

Final Project Components

Project Idea Proposal

- What is your general idea?
- Check out our list of project ideas or come up with your own!

Analytical Proposal + Map

- What is my question, idea, goal?
- What are my data sources, how will I use them?
- Map of study area.

Poster

- Presentation of your work
- 4 sections: Intro, Methods, Results, Discussion

Project Ideas: Help, what do I do next?

- Check out the example data sources and project posters on Moodle.
- Explore MassGIS and the US Census Bureau websites.
- Check out the Umass Library GIS Hub.
- Think about your research interests. Do you have some data you'd like to use?
- Chat with me and your TAs.

Thinking Spatially

How does geographic information science work?

Thinking Spatially: How GIS Works

GIS is not just a digital mapmaker! It is a way of thinking and solving problems.

Start thinking about the scope of problems for which GIS is useful.

Where did GIS come from?

Background in geography, cartography, computer science and mathematics

Fusion of information systems and imaging/positioning technologies

Geographic Information Science is a new interdisciplinary field built out of the use and theory of GIS

Abstraction of Reality

Maps and GIS abstract reality, and how that reality is recomposed determines how it is represented.

Your choices as a GIS user will determine how people see reality.

Thinking Spatially: Concepts vs. Implementation

You will learn foundational and theoretical concepts.

These help us solve any GIS problem.

Arc GIS is a specific implementation… but there are lots of others.

Don’t become tethered to a particular implementation!

Geographic Information Science

GIS as a Toolbox :

"…a powerful set of tools for storing and retrieving at will, transforming and displaying spatial data from the real world for a particular set of purpose"

Peter Burrough (1998)

Geographic Information Science

GIS as a System:

"An information system that is designed to work with data referenced by spatial or geographic coordinates. In other words, a GIS is both a database system with specific capabilities for spatially-referenced data, as well as a set of operations for working with the data"

Jeff Estes and Jeffrey Star (1990)

Geographic Information Science

GIS as a Science:

"The generic issues that surround the use of GIS technology, impede its successful implementation, or emerge from an understanding of its potential capabilities."

Goodchild (1992)

Geographic Information Science

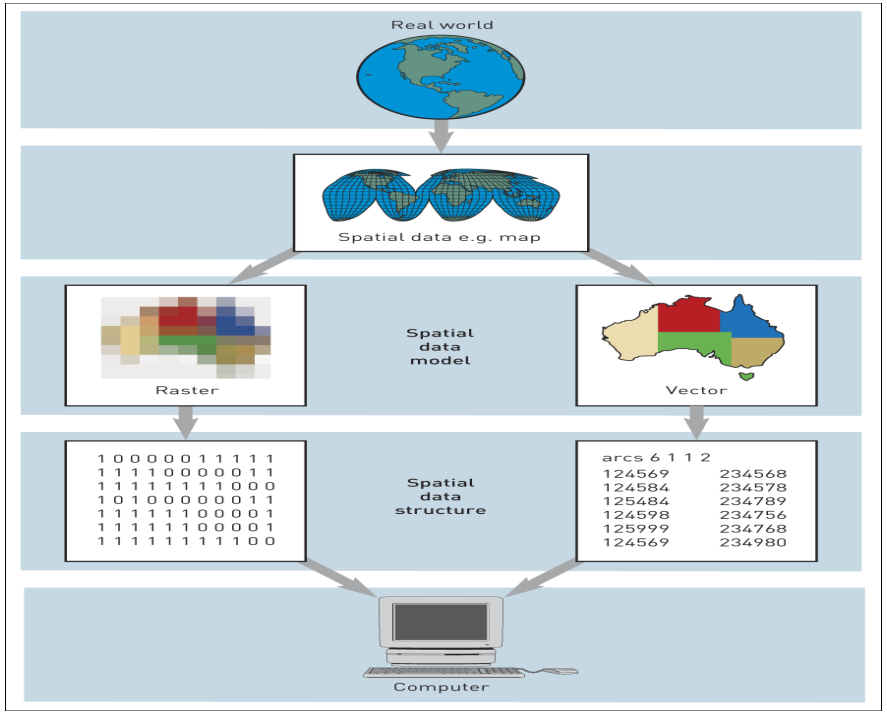
Don’t let this workflow diagram fool you!
A GIS process is iterative.

Geographic Information Science

Input → Management → Analysis → Output

Types of Spatial Data Available:

- Vector Data
- Raster Data
- Triangular Irregular Networks



Geographic Information Science

Spatial Data Terminology:

- **Vector Data** *also known as a* ***"shapefile"***
- **Raster Data** *also known as a* ***"grid"***

General Terminology

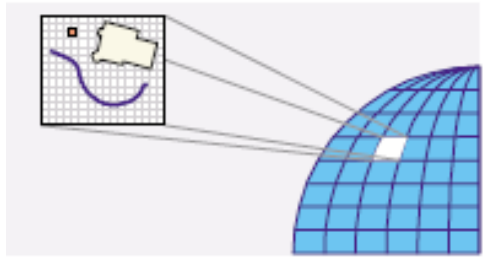
ArcGIS specific terminology

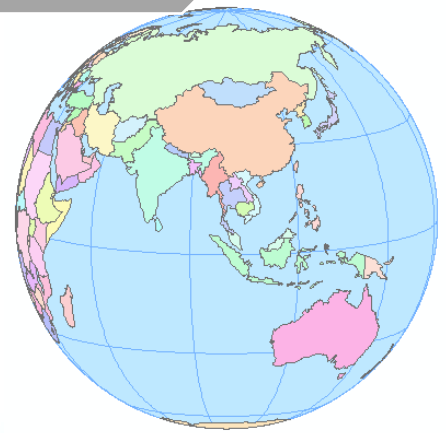
Geographic Information Science

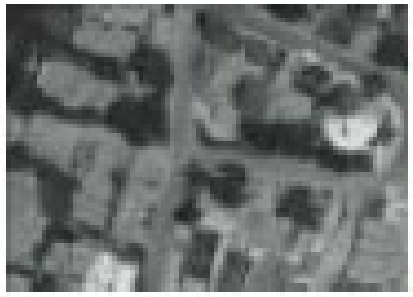
Input → Management → Analysis → Output

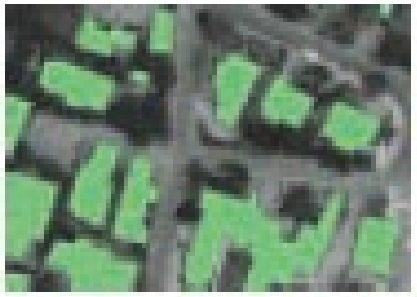
Types of Spatial Data:

- Vector data has 2 components
  - Attributes
  - Locations
- Raster data: just a grid!







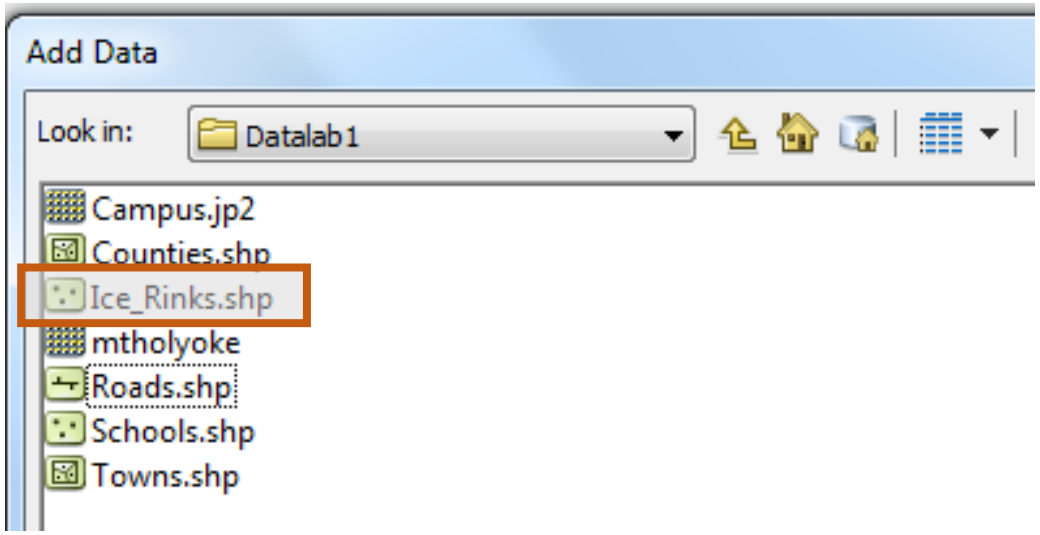


Geographic Information Science

Types of Spatial Data:

- Vector Data
  - Icons look like points, lines, or polygons in Arc

In ArcMap:



Geographic Information Science

Vector data: “Shapefiles”

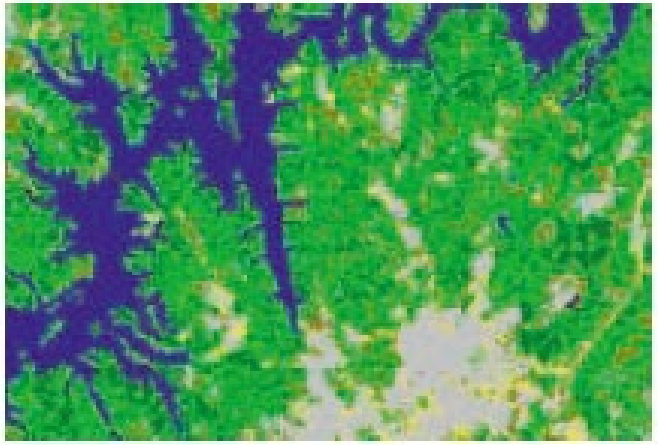
A “Shapefile” is actually a collection of related files with similar filenames, but different extensions.

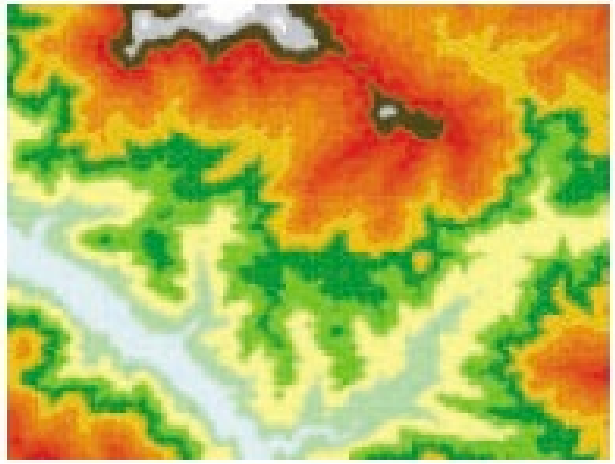
In Windows explorer:

Geographic Information Science

Types of Spatial Data:

- Raster Data
  - Grids are just like digital images
  - Continuous or categorical



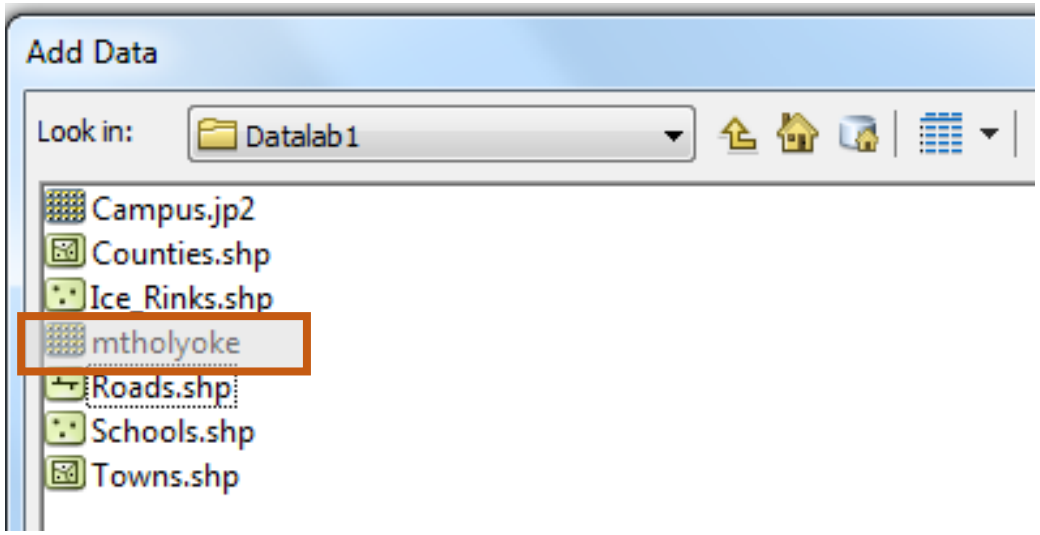


Geographic Information Science

Types of Spatial Data:

- Raster Data
  - Icon looks like a grid in Arc

In ArcMap:

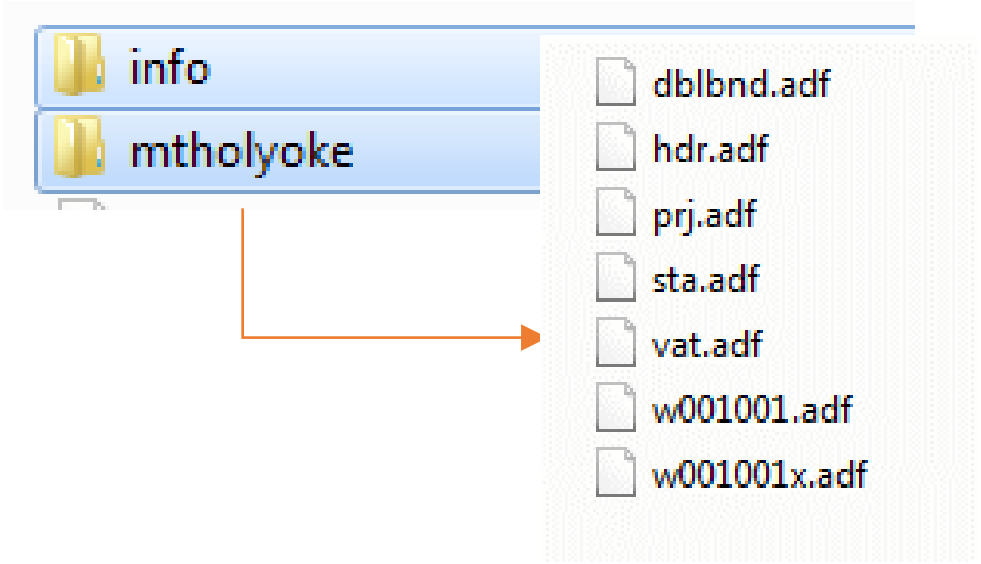


Geographic Information Science

Types of Spatial Data Available:

- Vector Data
- Raster Data
  - Images
  - Continuous and categorical

In Windows explorer:



Geographic Information Science

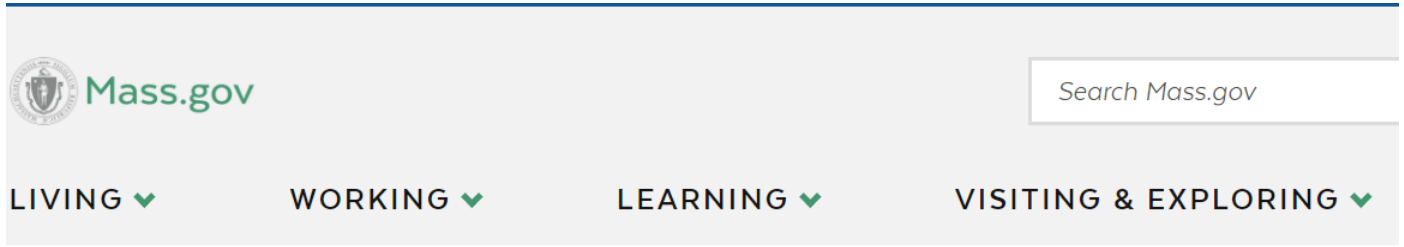
Input → Management → Analysis → Output

Geographic Information Science

Sources of Spatial Data:

- Web: WorldClim, Databasin, Federal and state government, so many more…
- Purchased data (e.g., remotely sensed imagery)
- Faculty & graduate students
- Create your own! We'll do this in lab 5.

Google ‘Mass GIS Data Layers’



MassGIS Data Layers

Each digital dataset name below links to a complete datalayer description. On each page you will find metadata and links to free data download.

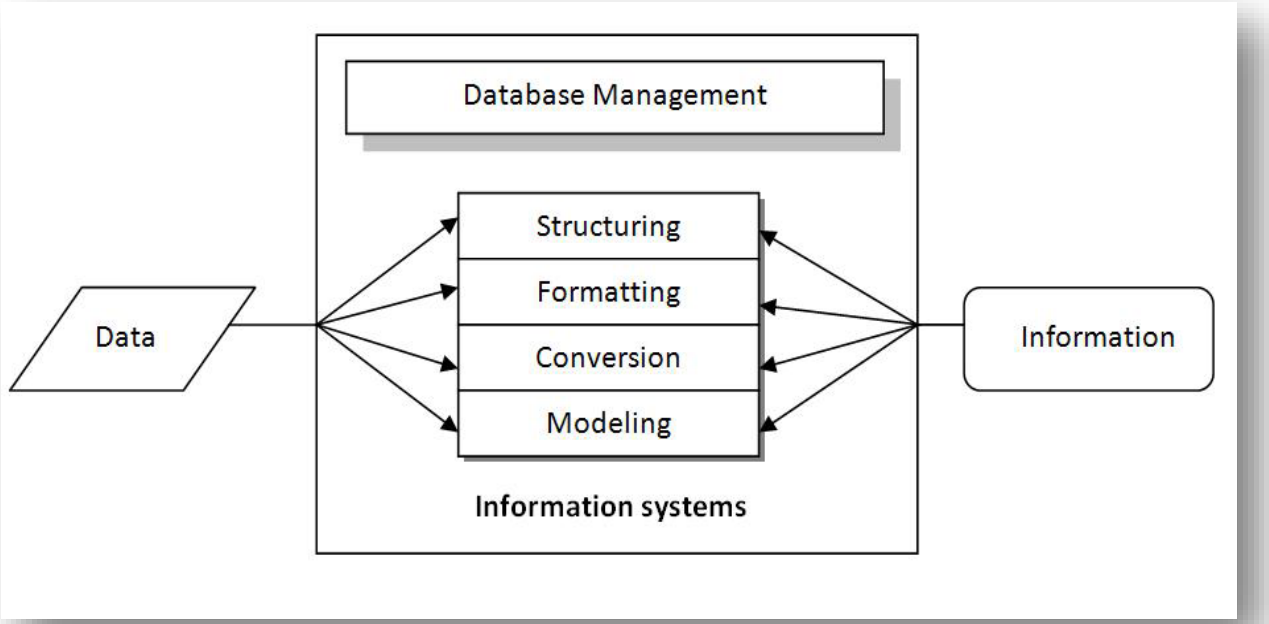
The date below the datalayer name on each page represents the month and year of the most recent update (or when the data first appeared in MassGIS' database). If you need more details, especially at the feature level, you will find contact information in the Maintenance section of each page.

See the Data Overview for data attribution and citation language and the spatial reference of

Some Data Sources

- U.S. Department of Labor: emerging fields
  - http://www.doleta.gov/brg/jobtraininitiative/
- GIS Certificate Program
  - http://www.gisci.org/
- U.S. Geological Survey
  - http://www.usgs.gov
- U.S. Census Bureau
  - http://www.census.gov/
- National Weather Service: GIS portal
  - http://www.weather.gov/gis/
- Hurricane Center
  - http://www.nhc.noaa.gov/
- U.S. Department of Housing and Urban Development: GIS portal
  - http://egis.hud.gov/
- U.S. Department of Health and Human Services: data warehouse
  - http://datawarehouse.hrsa.gov/
- Federal Highway Administration: GIS in transportation
  - http://www.gis.fhwa.dot.gov/apps.asp
- Forest Service: geospatial service and technology center
  - http://www.fs.fed.us/
- U.S. Department of Agriculture: program on precision, geospatial & sensor technologies
  - http://www.nifa.usda.gov/nea/technology/technology.cfm

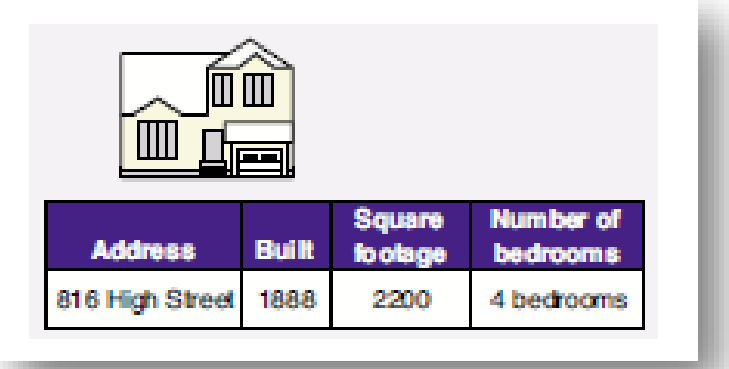
Geographic Information Science



Geographic Information Science

Data Management:

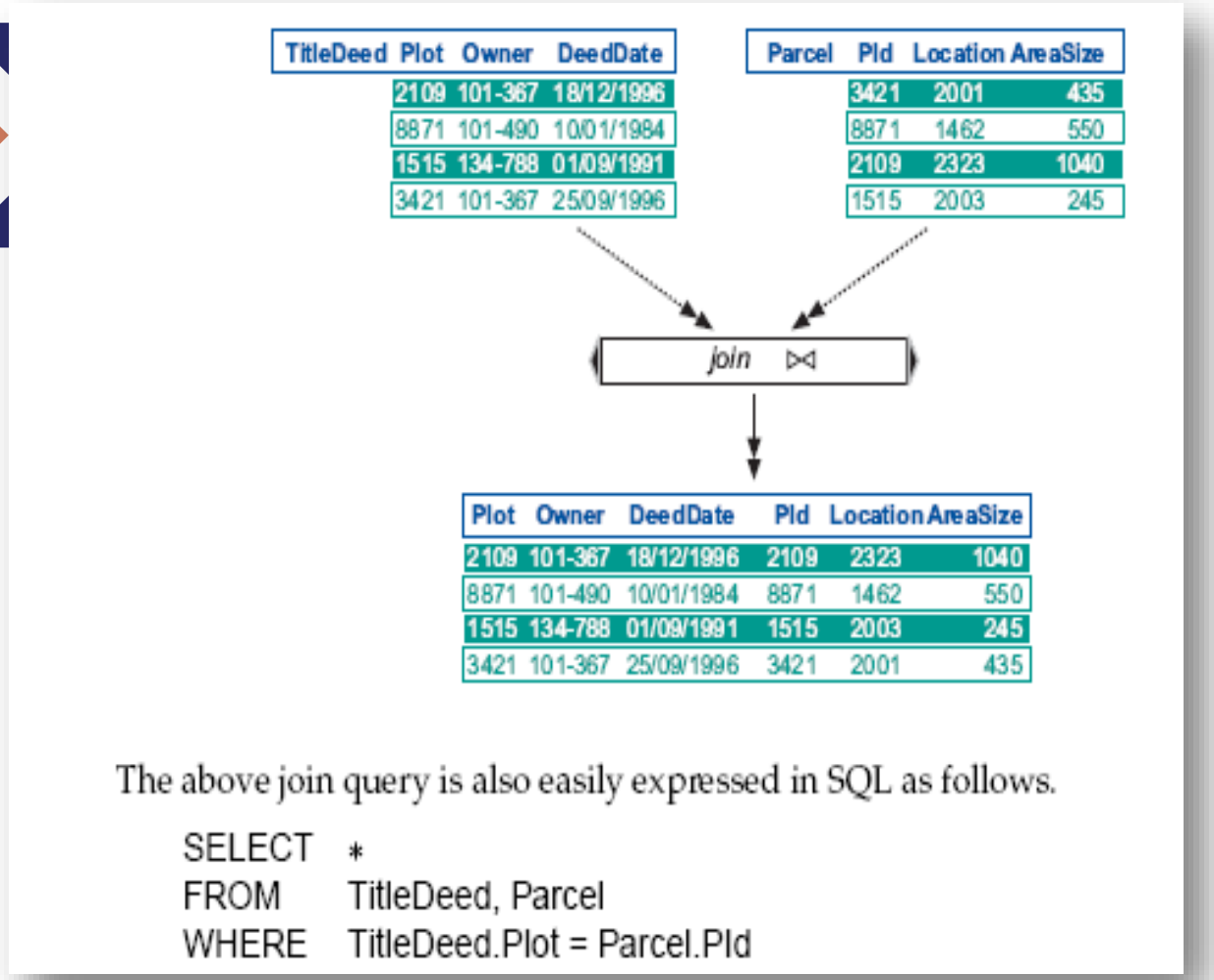
- [Vector] spatial data rely on databases, so they require database management
- Currently lots of movement towards online and multi-user geodatabases



Geographic Information Science

The Row Data Paradigm is implemented in vector data

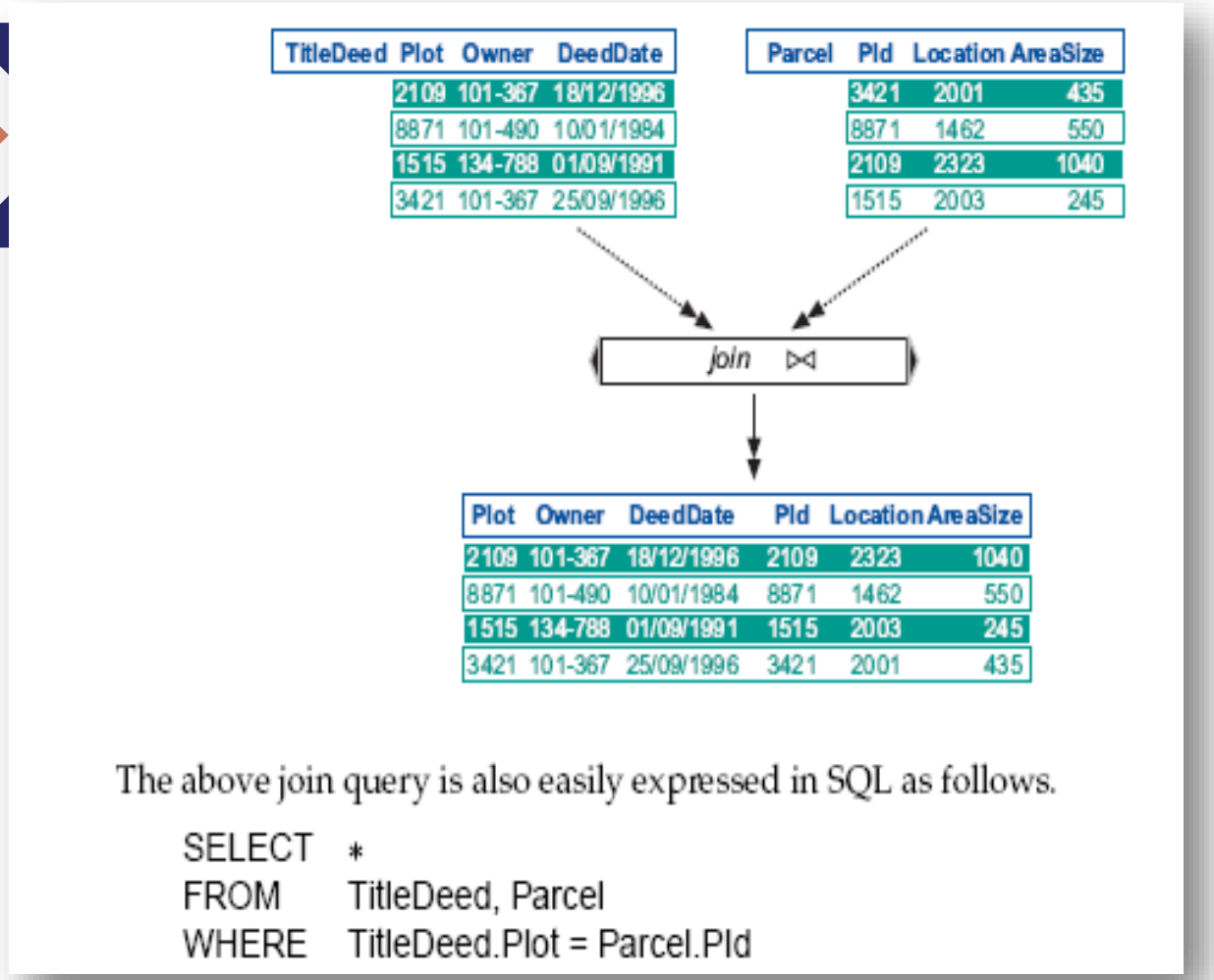
- Data arranged in 2D tables
- Rows are entities/observations
- Columns are attributes
- Row Data Paradigm allows us to relate different tables.
- Also known as ‘long-format’



Geographic Information Science

Structured **Q**uery **L**anguage

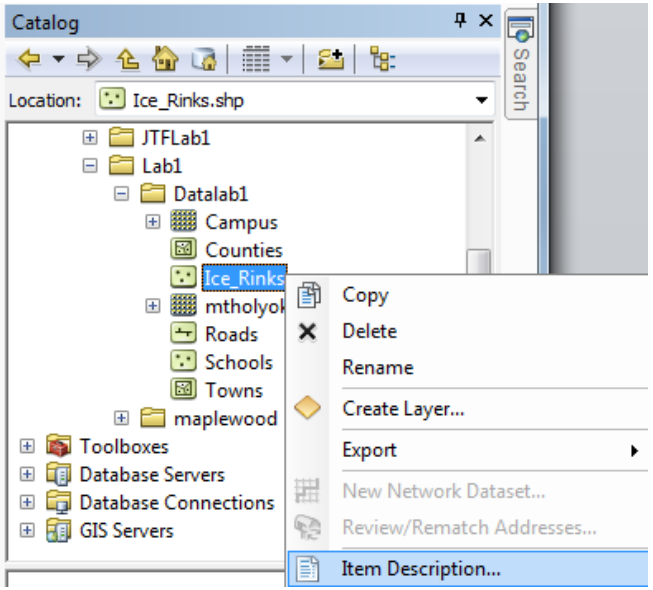
SELECT - extracts data (columns to be displayed) from a database
UPDATE - updates data in a database
DELETE - deletes data from a database
INSERT INTO - inserts new data into a database



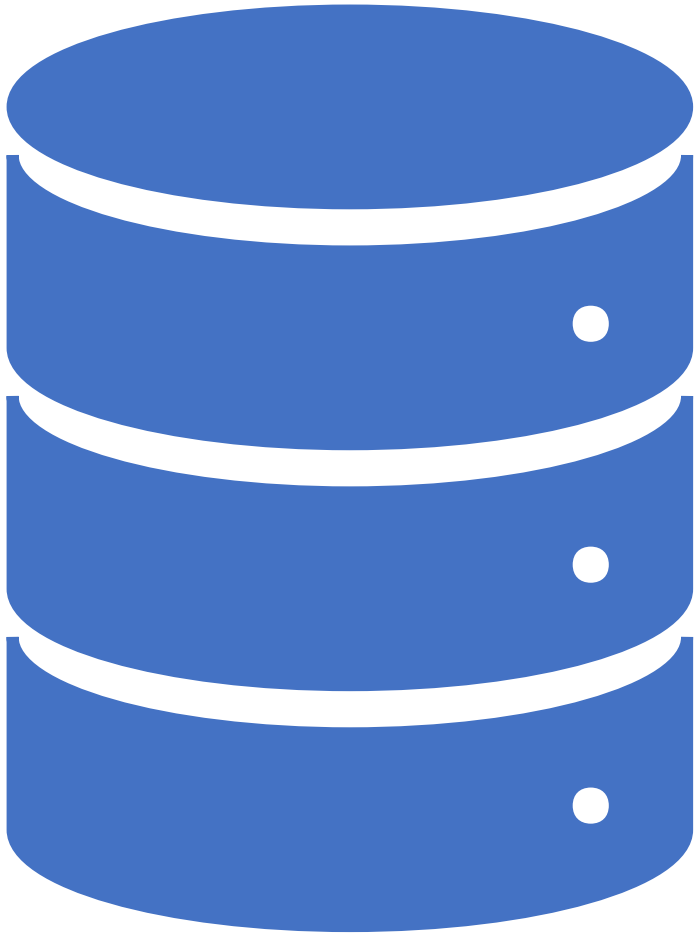
Geographic Information Science

What you need to know about data management:

- Spatial data often have metadata.
- But what is metadata?



Data and Metadata



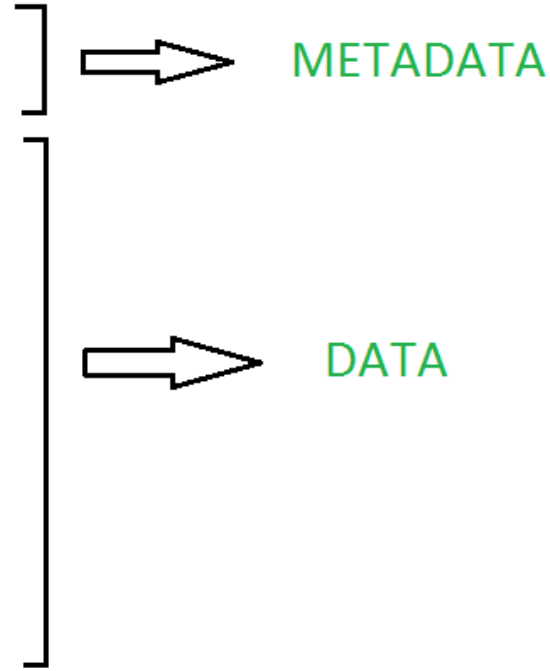
Data vs. Metadata

What does geeks 4 geeks have to say?

“1. DATA: The term data is derived from Latin word ‘Datum’ which refers to ‘something given’. Data is raw and unorganized facts…”

“2. METADATA: Metadata is a data about data. Metadata shows basic information about data, which can make finding and working with specific instances of data easier.”

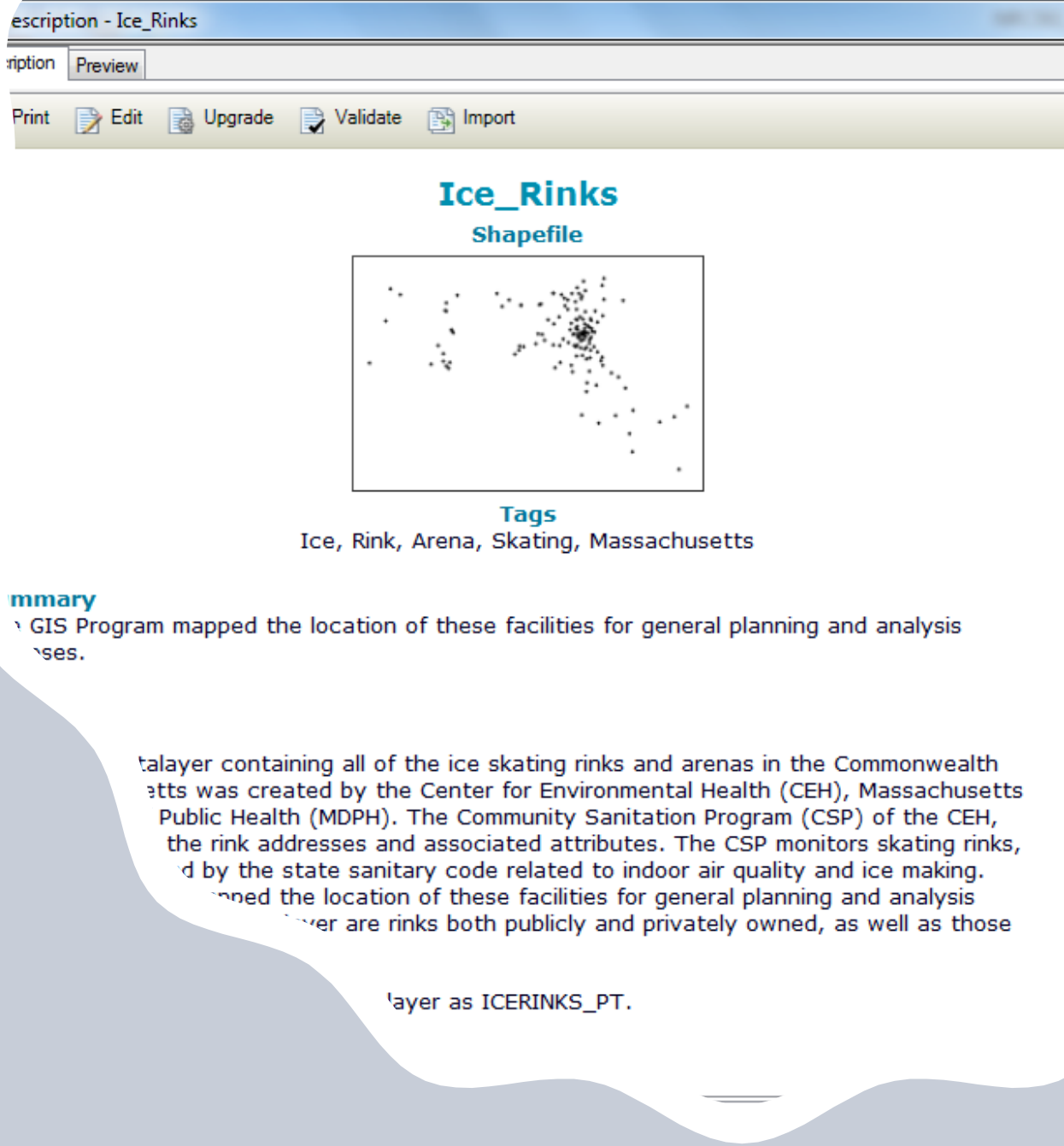
Species	Age	Body Mass
Adelie	5	3500
Gentoo	2	3550
Adelie	3	3567
Chinstrap	2	4724



Adapted from: https://www.geeksforgeeks.org/difference-between-data-and-metadata/

Metadata

Data without metadata is useless*

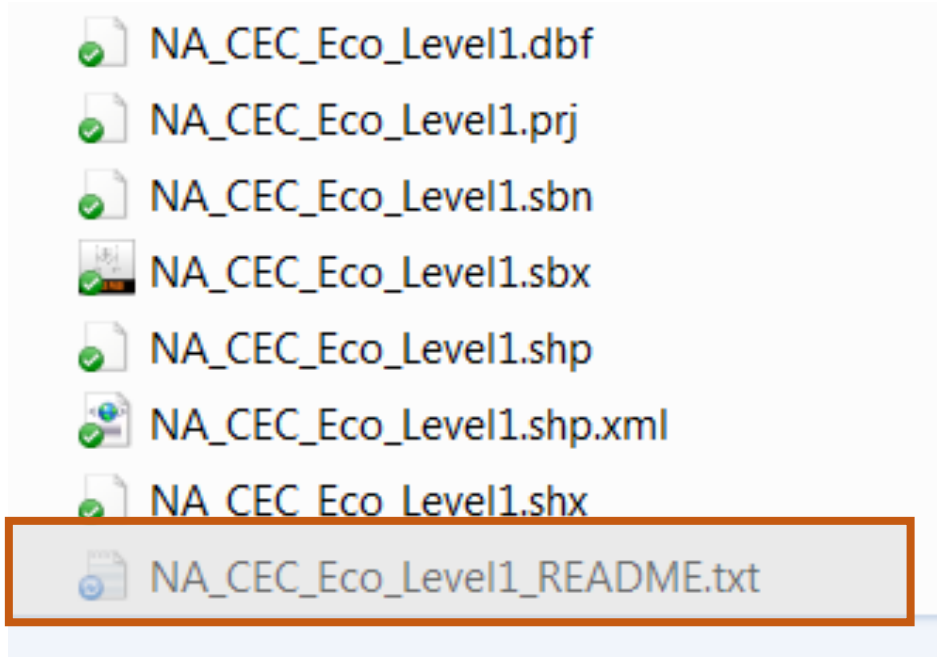


Geographic Information Science

What you need to know about data management:

- Spatial data often have metadata

(not always in obvious places)



Geographic Information Science

What you need to know about data management:

- ALWAYS use ArcCatalog to move spatial data
  - This will save lots of headaches and help avoid those red exclamation marks!
  - Drag+Drop is convenient, but it can cause unexpected errors later (especially for raster data or with Virtual Desktop)

Geographic Information Science

Analysis of Spatial Data is key in GIS

- Querying of data layers (Lab 3)
- Links between data layers (Lab 3; Lab 8)
- Spatial Modeling (Lab 4; Lab 7)
- Statistical analysis (Lab 8)

Example: Selection Tools
Select by Attribute

How many public ice rinks are located in Boston?

Ice Rinks

Field: Add Calculate | Selection: Select By Attributes Zoom To Switch Clear Delete Copy

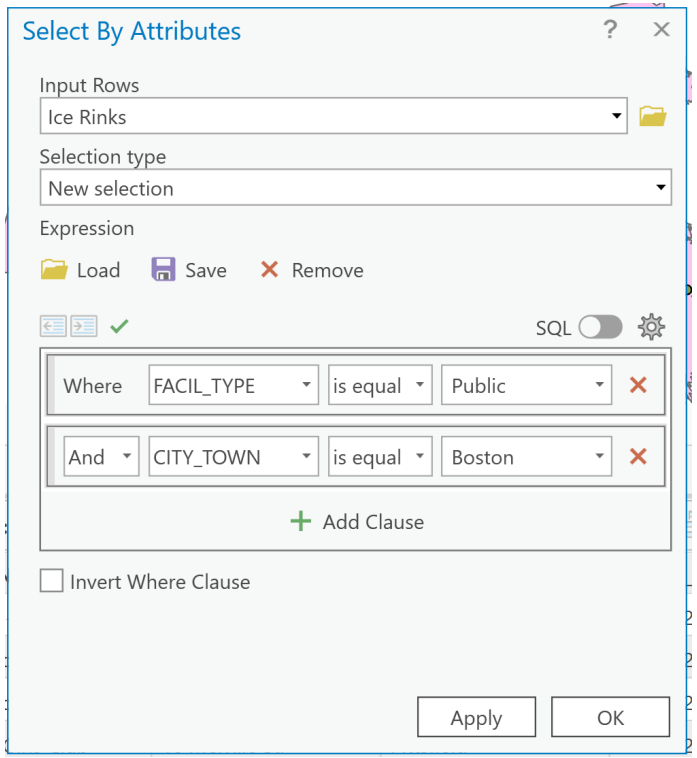
	FID	Shape *	OBJECTID	FACIL_NAME	ADDRESS	CITY_TOWN	TOWN_ID	PHONE	FACIL_TYPE
1	0	Point	584	McVann-O'Keefe Me...	511 Lowell St.	Peabody	229	(978) 535-2110	DCRP
2	1	Point	585	Hobomock Arena Rin...	132 Hobomock St.	Pembroke	231	(781) 294-0260	Public
3	2	Point	586	Hobomock Arena Rin...	132 Hobomock St.	Pembroke	231	(781) 294-0260	Public
4	3	Point	587	Boys & Girls Club	16 Melville St.	Pittsfield	236	(413) 448-8258	Public
5	4	Point	588	John A. Armstrong M...	103 Long Pond Rd.	Plymouth	239	(508) 746-8825	DCRP
6	5	Point	589	Shea Memorial Rink	651 Willard St.	Quincy	243	(617) 472-9325	DCR
7	6	Point	590	Quincy Youth Arena	60 Murphy Memorial...	Quincy	243	(617) 479-8371	Public
8	7	Point	591	Joseph Zapustas Arena	240 North St.	Randolph	244	(781) 961-0938	Public
9	8	Point	592	CDL Arena	1568 Broadway St.	Raynham	245	(508) 880-3311	Public
10	9	Point	593	Burbank Area	51 Symonds Way	Reading	246	(781) 942-2271	Public
11	10	Point	594	Cronin Memorial Rink	850 Revere Beach Par...	Revere	248		DCRP
12	11	Point	595	Rockland Rink	599 Summer St.	Rockland	251	(781) 878-5591	Public
13	12	Point	596	Massports Club, Bavis...	180 VFW Dr.	Rockland	251	(781) 982-7070	Public
14	13	Point	597	Massports Club, Mini...	180 VFW Dr.	Rockland	251	(781) 982-7070	Public

Did you notice that attribute tables follow the row data paradigm?

Example: Selection Tools

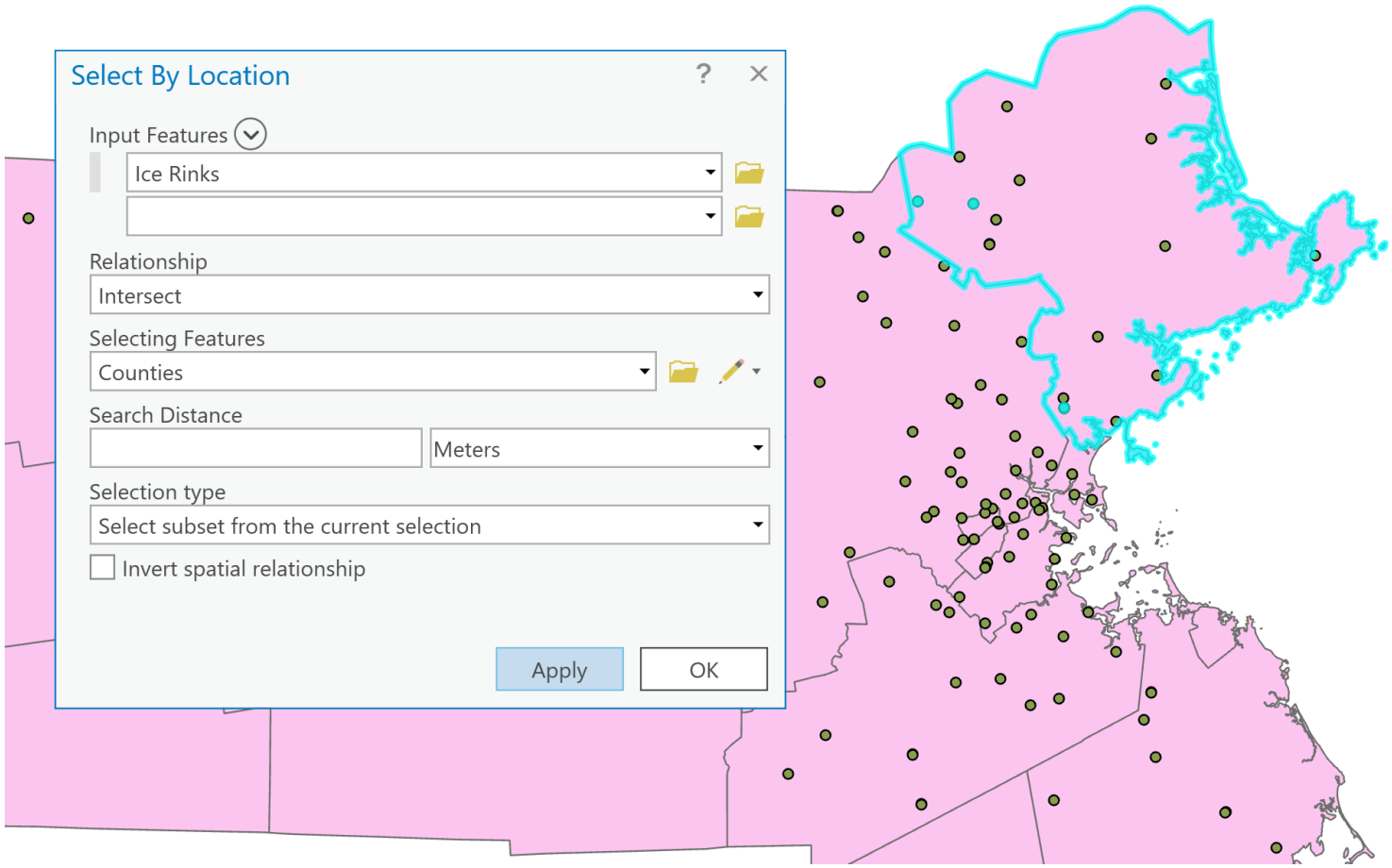
Select by Attribute: How many public ice rinks are located in Boston?

SELECT
*
FROM
Ice_Rinks
WHERE
“FACIL_TYPE”=‘Public’
AND
“CITY_TOWN”=‘Boston’



Example: Selection Tools

Select by Location: How many public ice rinks are located in Essex County?



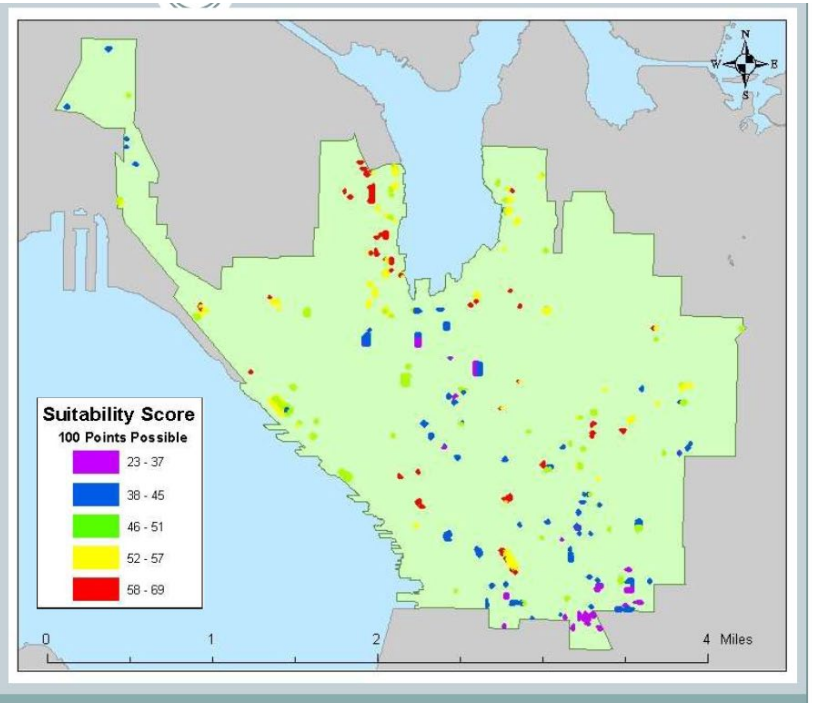
Geographic Information Science

Analysis of Spatial Data is key in GIS

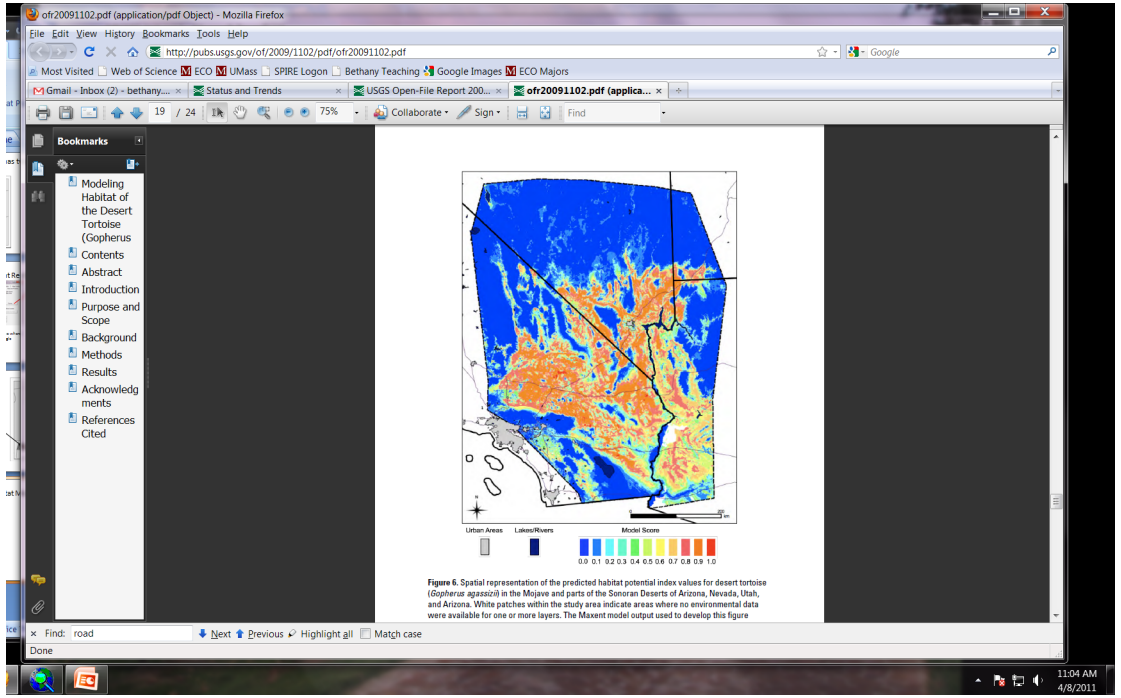
- Querying of data layers (Lab 2)
- Links between data layers (Lab 2; Lab 7)
- Spatial Modeling (Lab 3; Lab 6)
- Statistical analysis (Lab 7)

Doggie Daycare Suitability

After combining Customer Suitability, Distance Suitability and parcel criteria, you end up with a map of potential properties that meet all of your requirements.



Example: Suitability Analysis

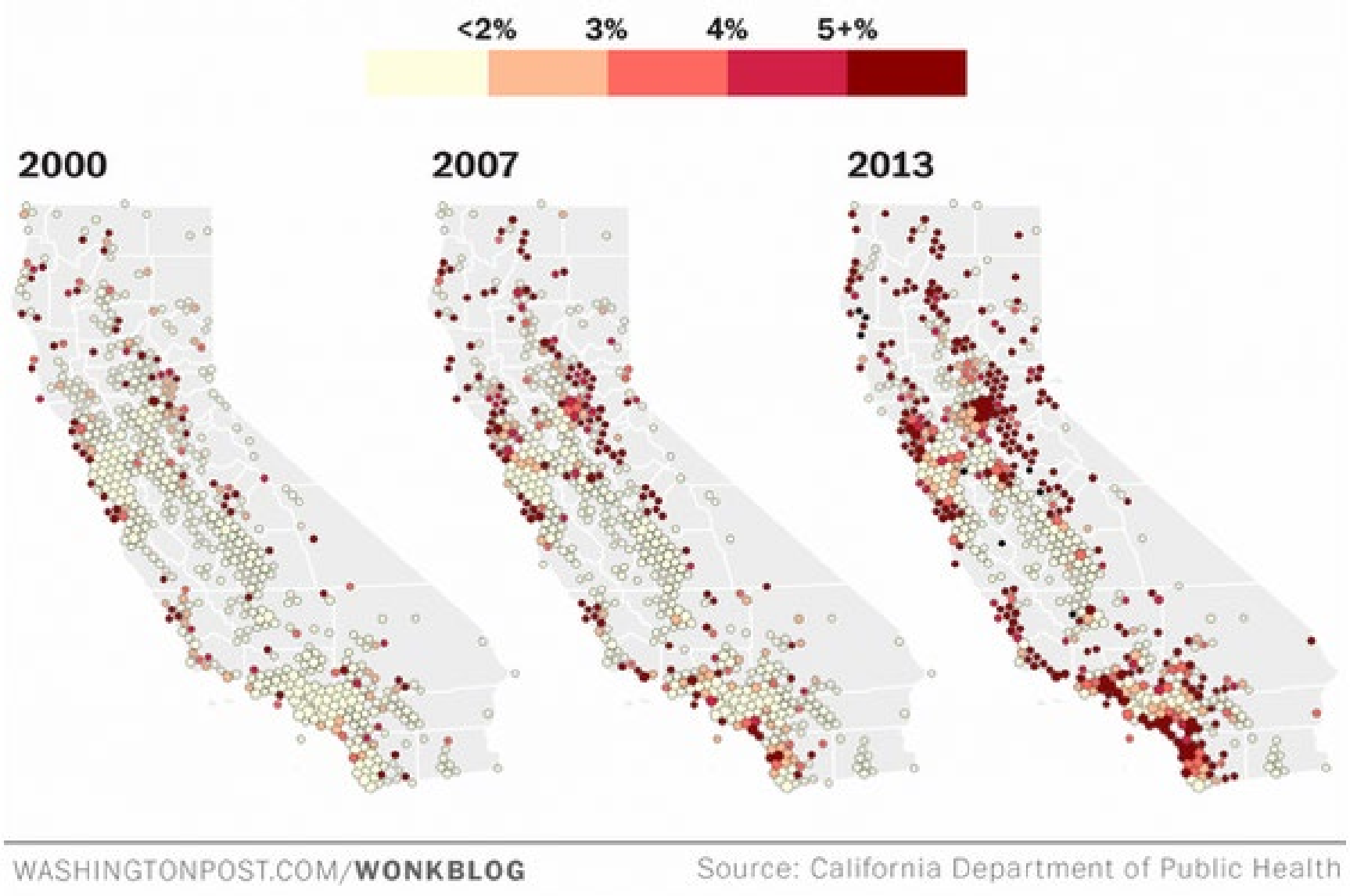


Geographic Information Science

GIS Output is often a map! GIS Output at its best is:

- Information rich
- Understandable (Intuitive!)

But first, a map puzzler!



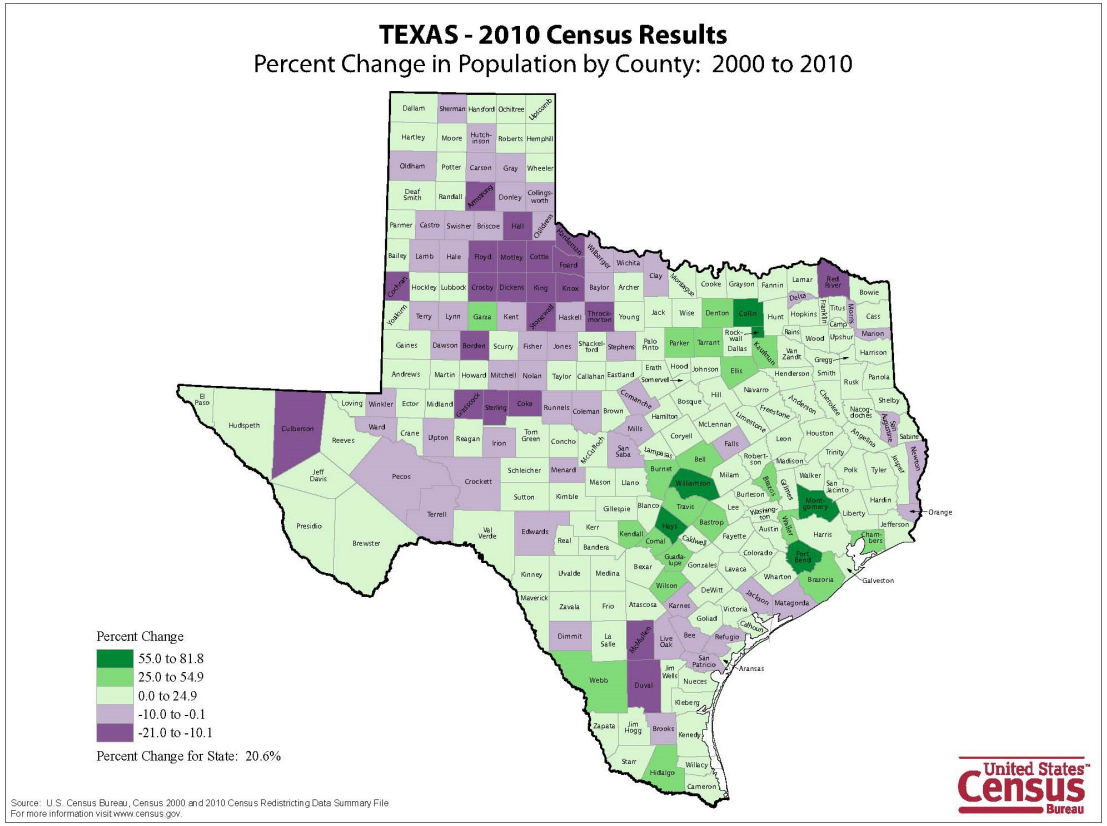
Maps are representations, i.e. abstractions, of reality

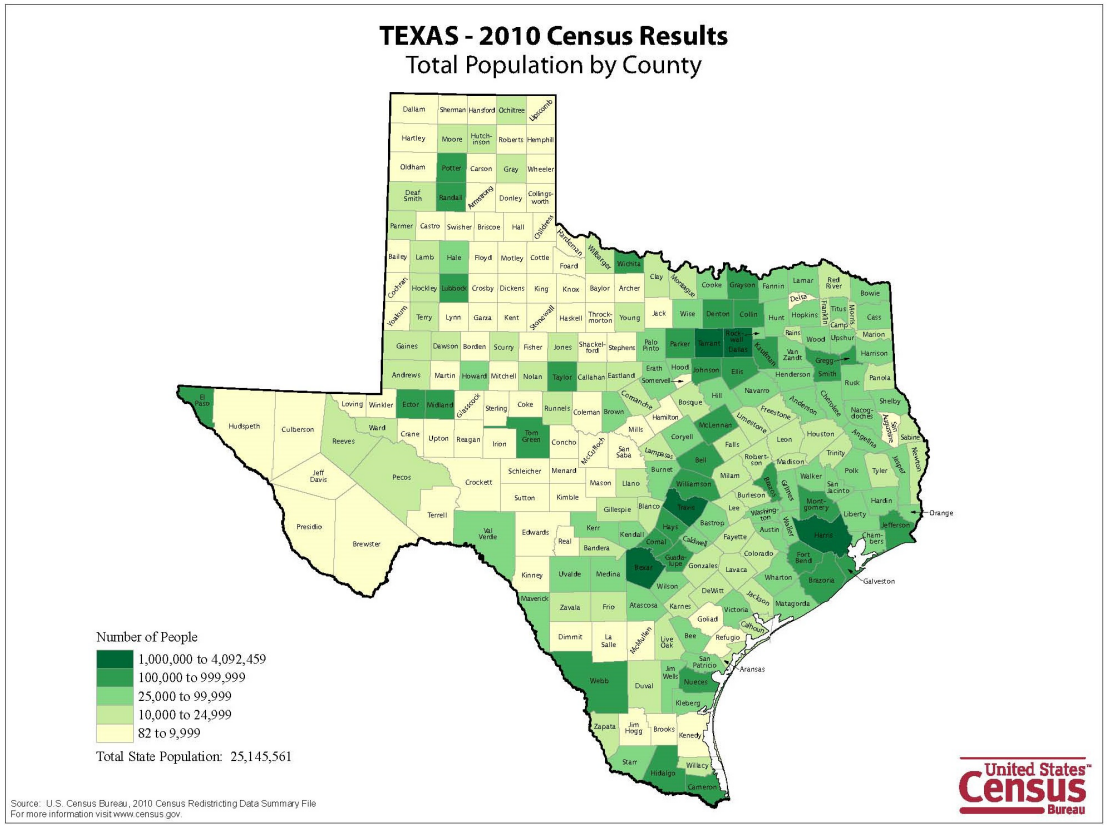
- When making a map, you have:
  - Control over content
  - Control over area
  - Control over emphasis
  - Analog of the real world

Control over Content

Control over Area

Control over Emphasis

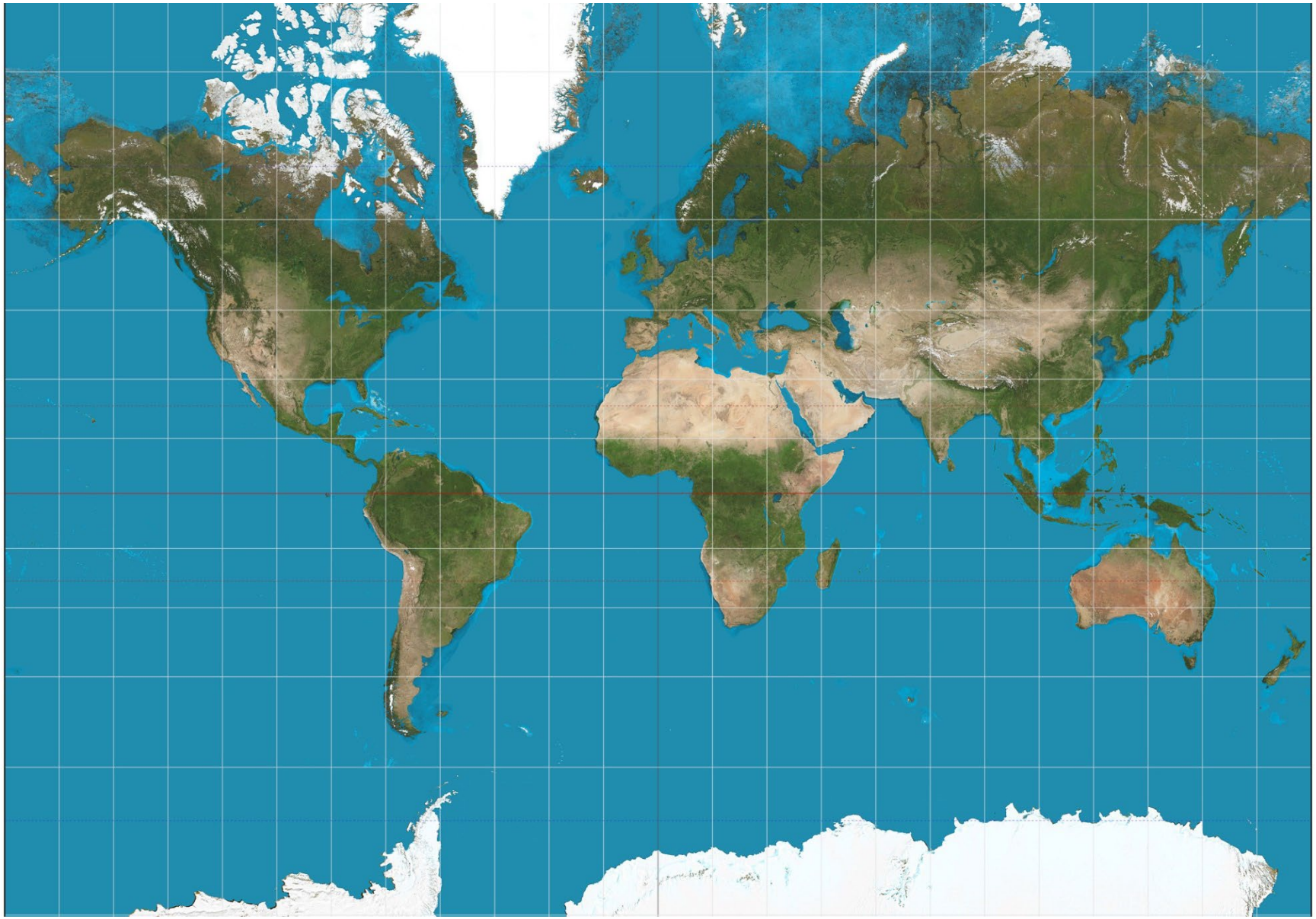




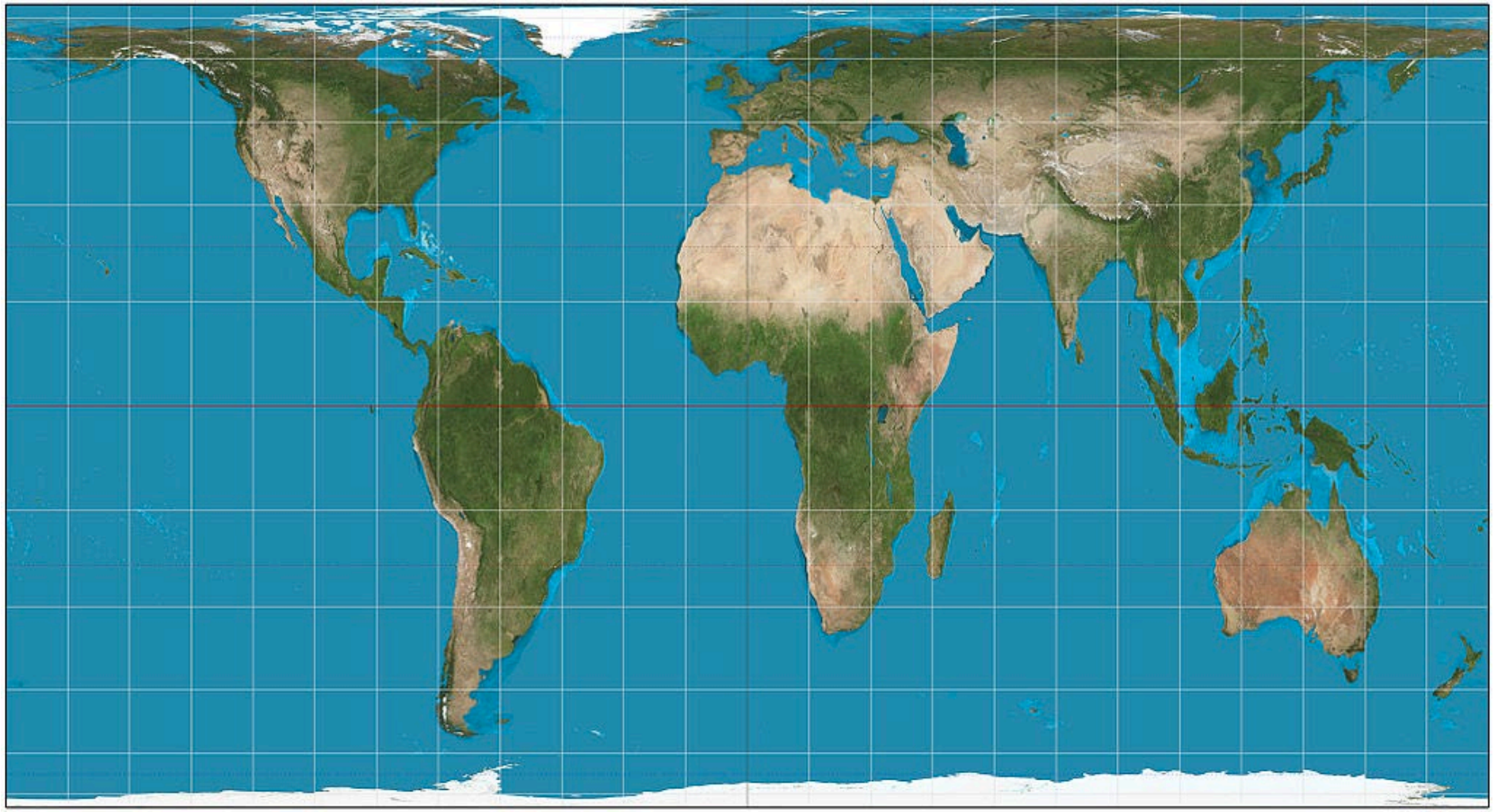
Analog of the real world (Topology)

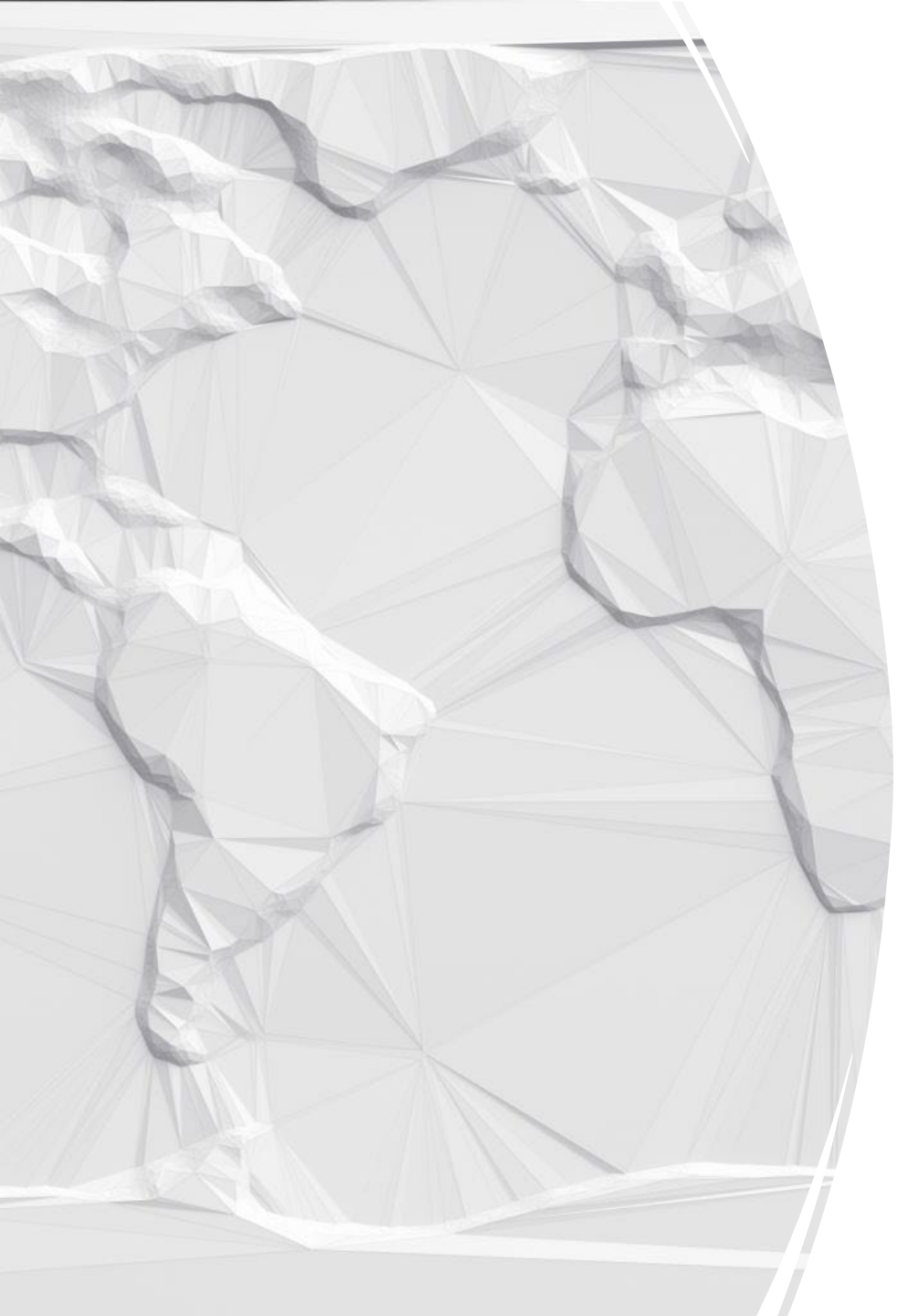
Analog of the real world (Topology)

Mercator Projection – reality?



Equal Area Projection – an alternate reality?





Supplement: More Map Design

Map Quality

“The quality of the map is also in part an aesthetic matter. Maps should have harmony within themselves. An ugly map, with crude colors, careless line work, and disagreeable poorly arranged lettering may be intrinsically as accurate as a beautiful map but is less likely to inspire confidence.”

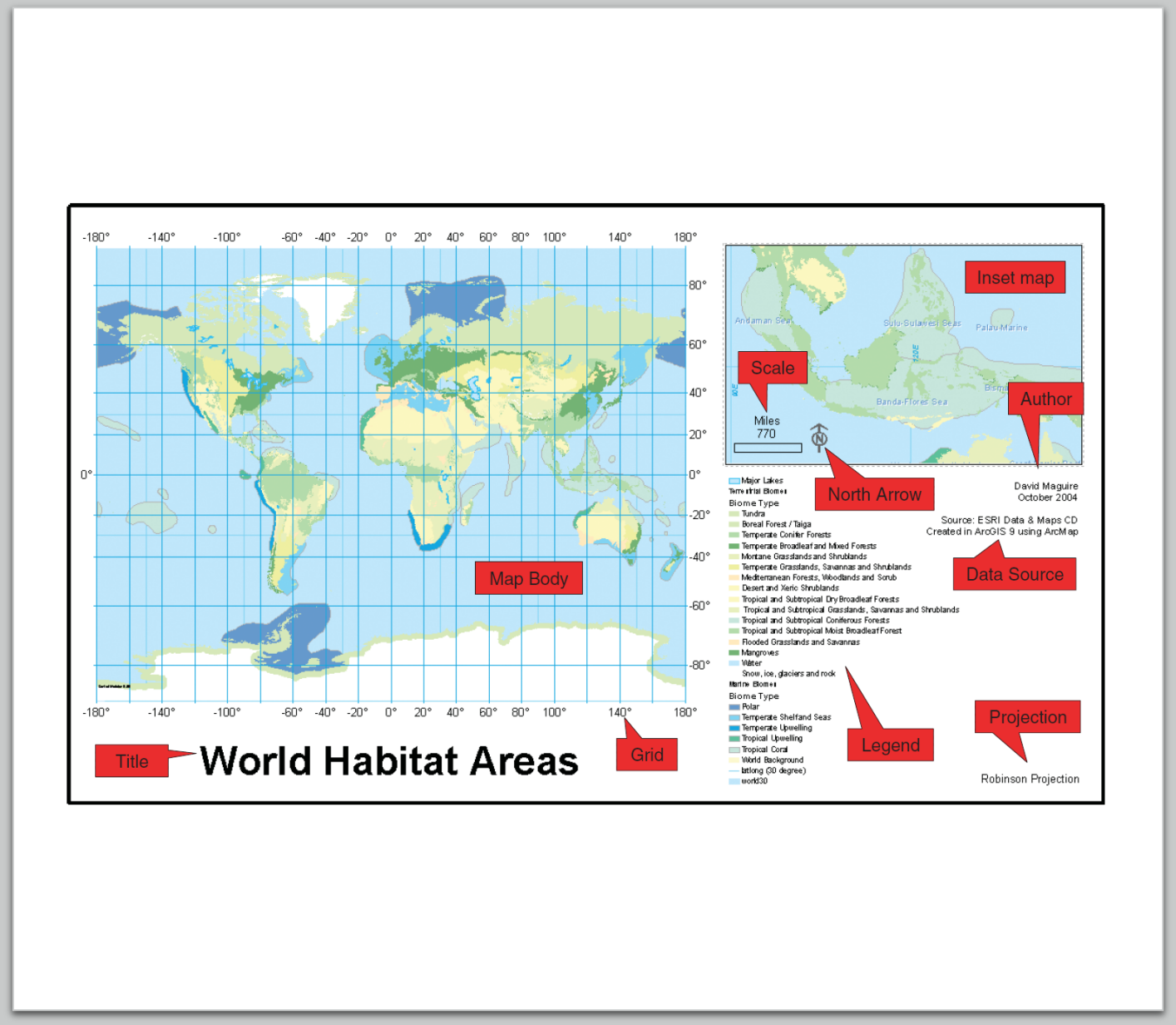
John K. Wright, cartographer and historian, and former director, American Geographical Society

Maps are Representations of Reality

- When making a map, you have:
  - Control over content
  - Control over area
  - Control over emphasis
  - Analog of the real world

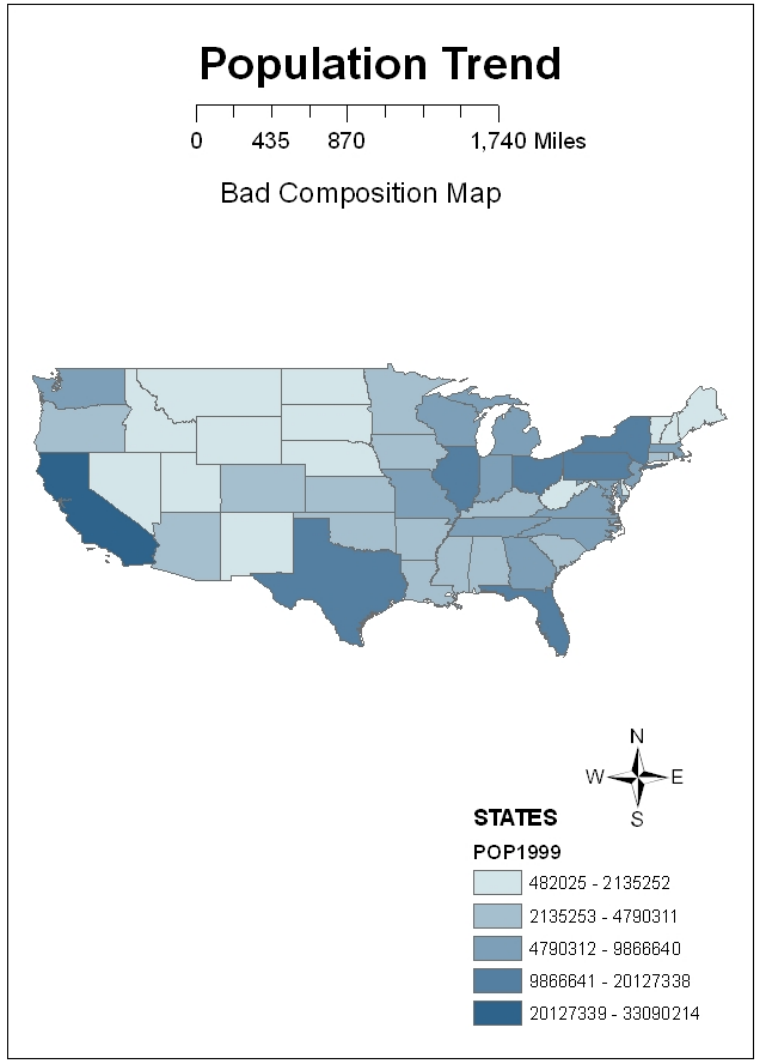
The Parts of the Map

- Frame and neat line
- Mapped area (figure)
- Inset
- Title and subtitle
- Legend
- Data source
- Scale
- Orientation



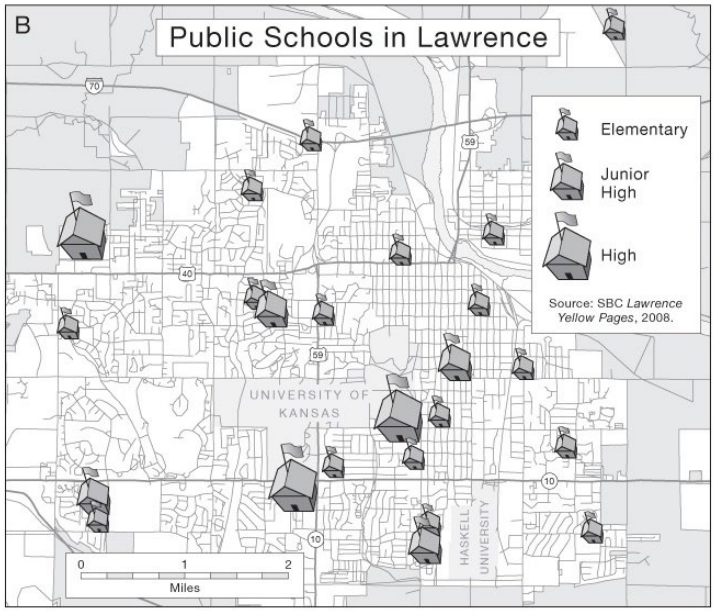
Designing the Map

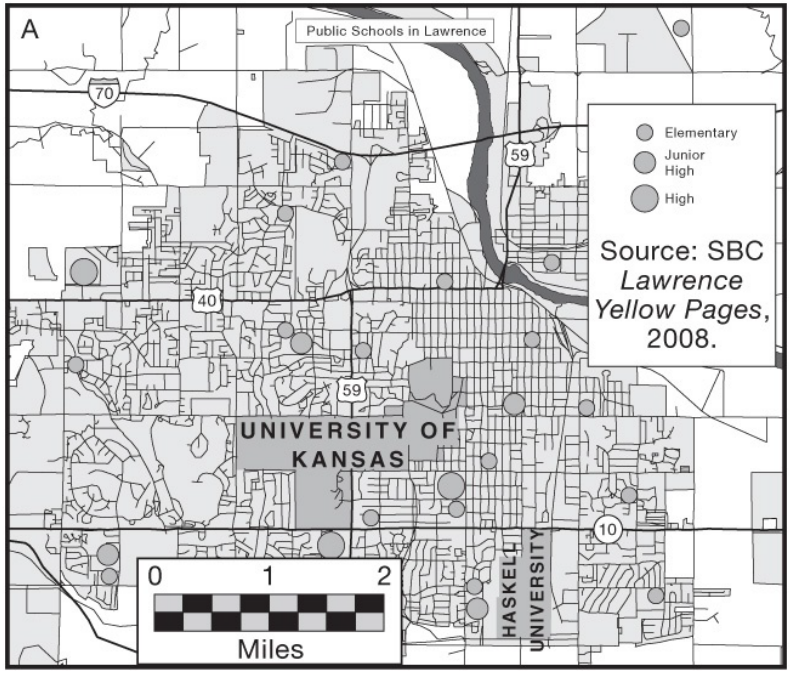
- Basics of the Map Design
  - Visual Balance
  - Visual Hierarchy
  - Simplicity the best!
- Pattern and Color
  - Hue
  - Intensity (Light/Dark)
  - Saturation (vividness)



Visual Hierarchy

- Graphical representation of the intellectual hierarchy , in which symbols and map elements are ranked according to their relative importance



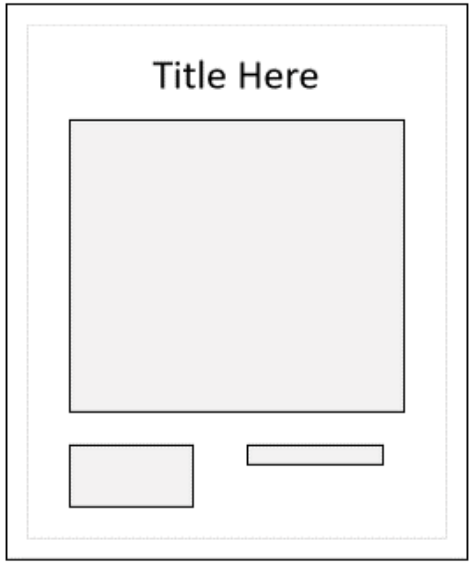


Visual Balance

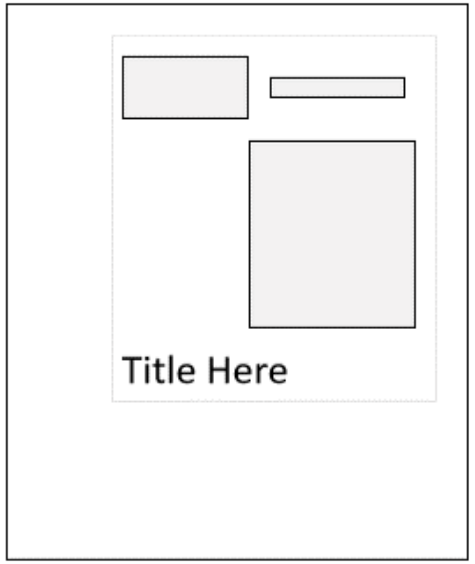
- The size of the symbols
- The pattern of the symbols
- The color of the symbols
- The visual hierarchy of the symbols and elements
- The location of the elements with respect to each other and the visual center of the map

Visual Hierarchy

- Graphical representation of the intellectual hierarchy.
- Symbols and map elements are ranked according to their relative importance.

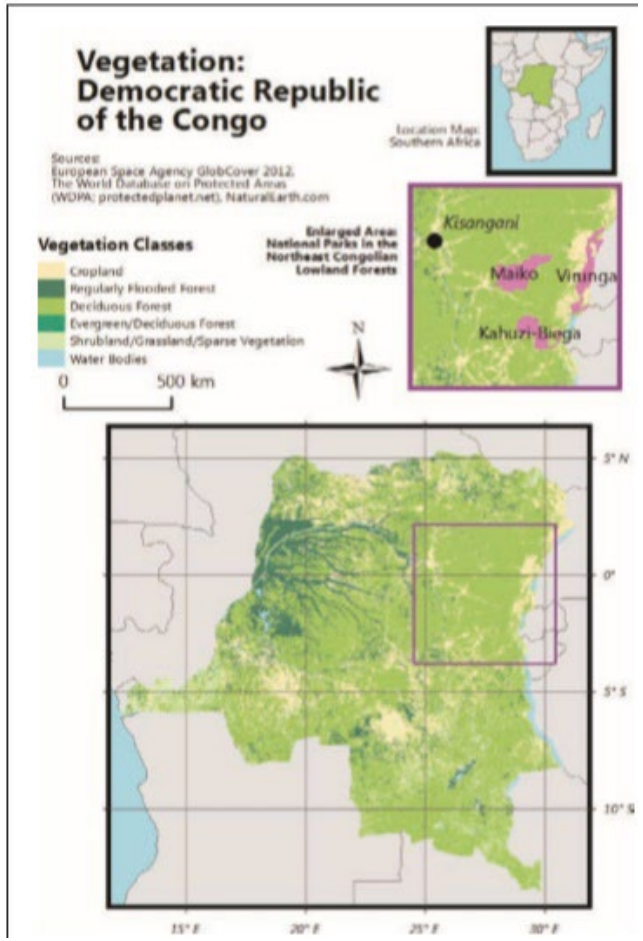


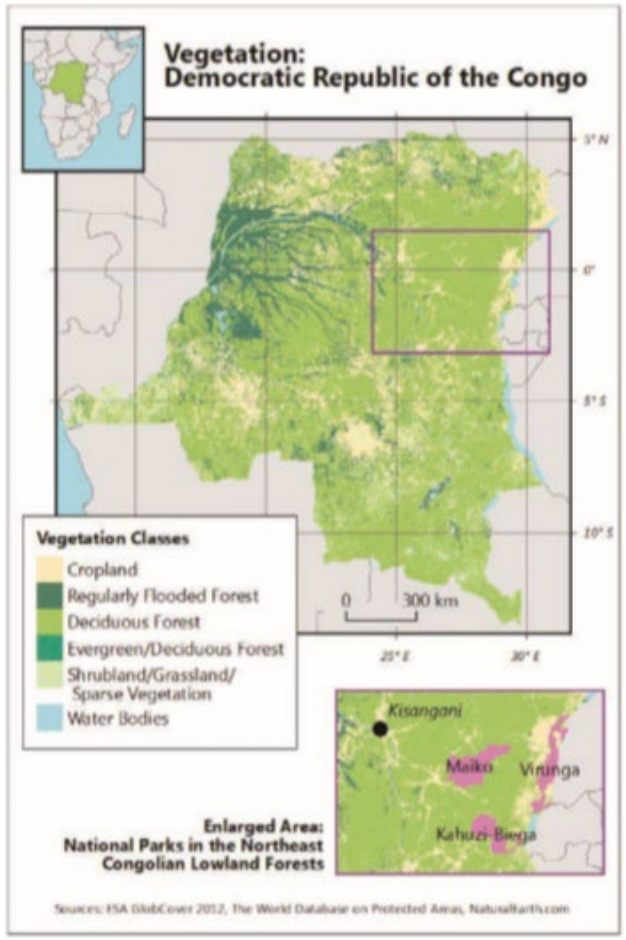
Good

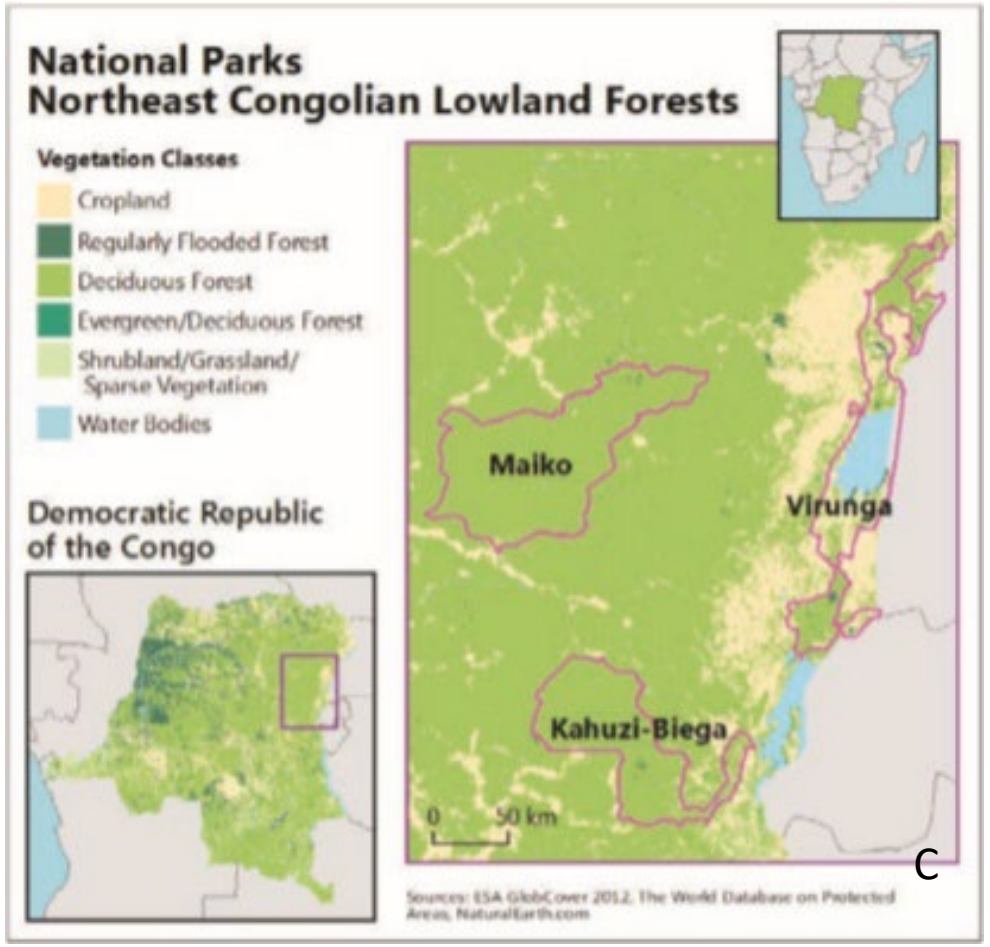


Bad

Good and bad design: Can you tell the difference?



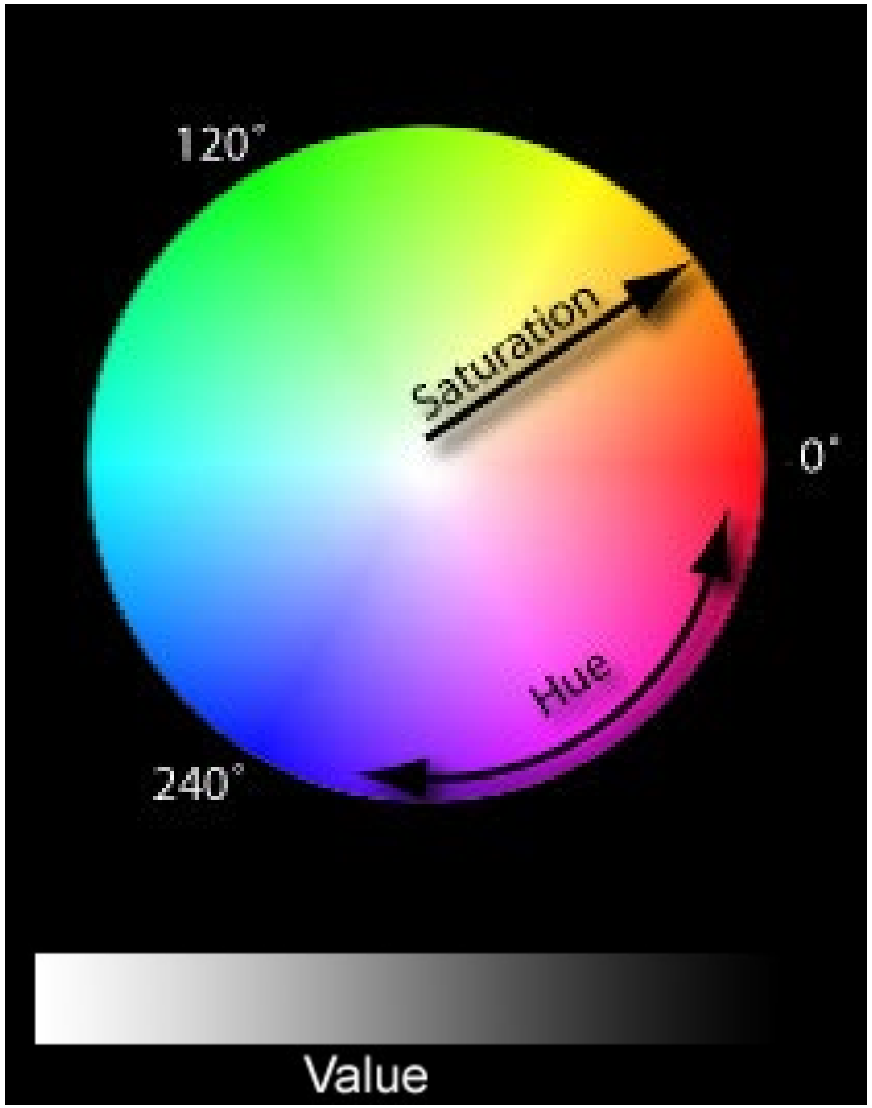




c

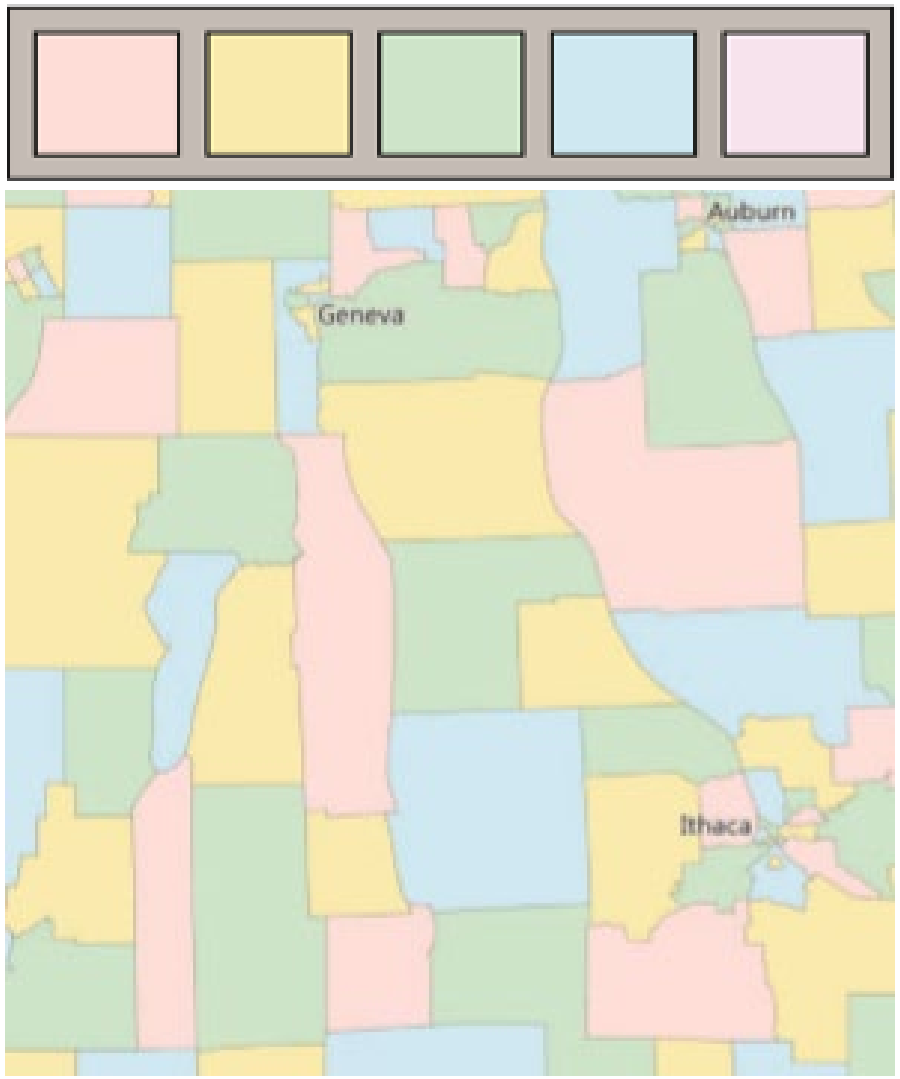
Designing the Map

- Basics of the Map Design
  - Visual Balance
  - Visual Hierarchy
- Pattern and Color
  - Hue
  - Saturation (vividness)
  - Intensity (Light/Dark)



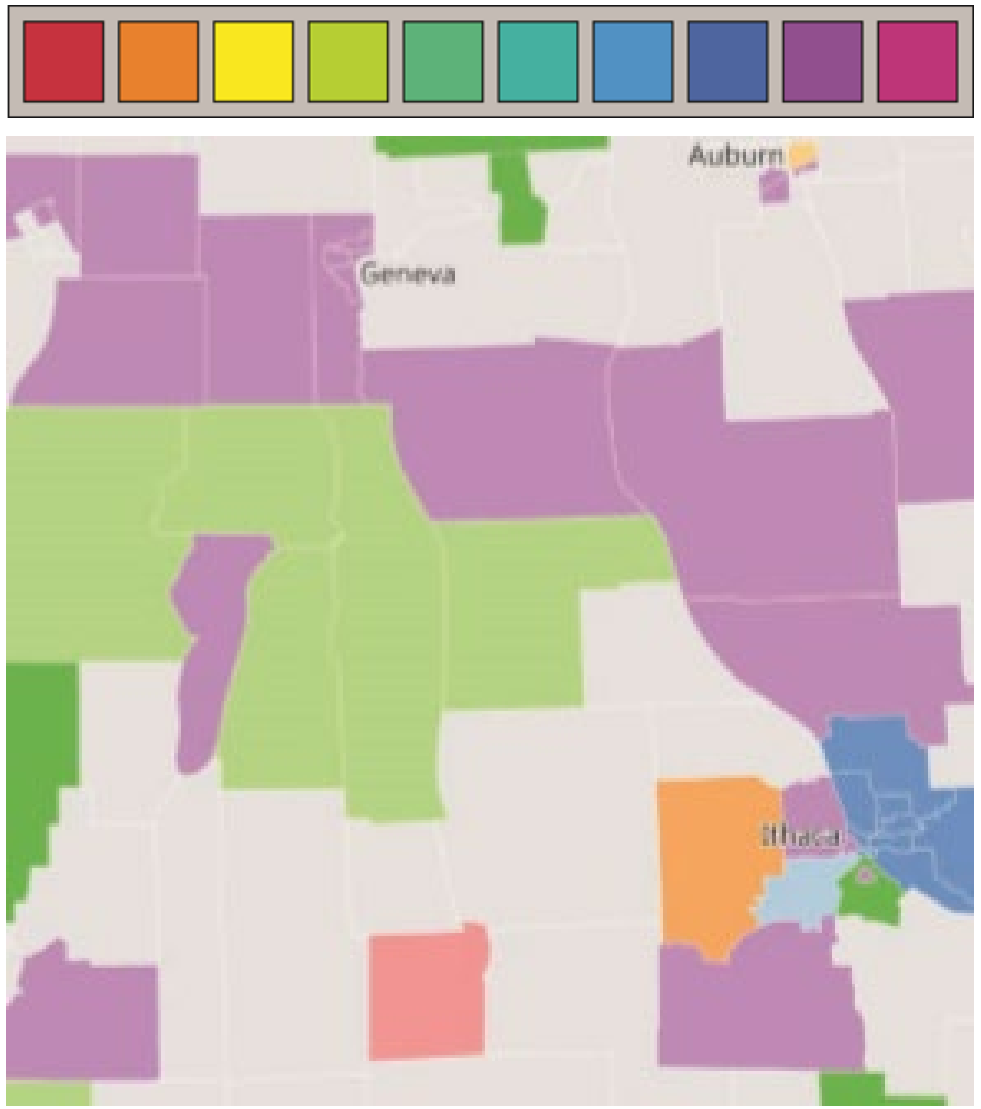
The Color Wheel: Hue

- Hues are used to differentiate census tracts in central New York State



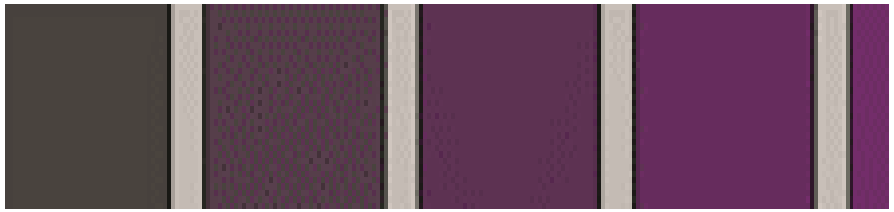
The Color Wheel: Hue

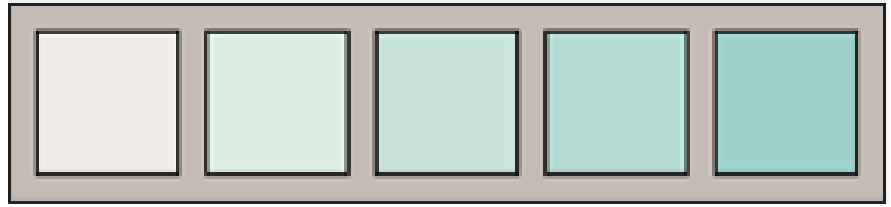
- Hues identify the majority non-English languages in each tract in central NY State.



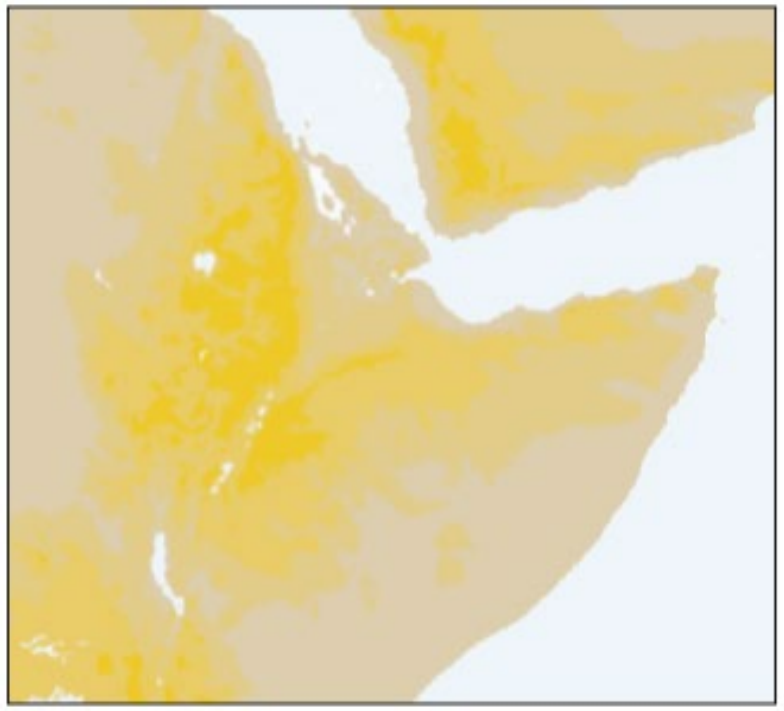
The Color Wheel : Saturation

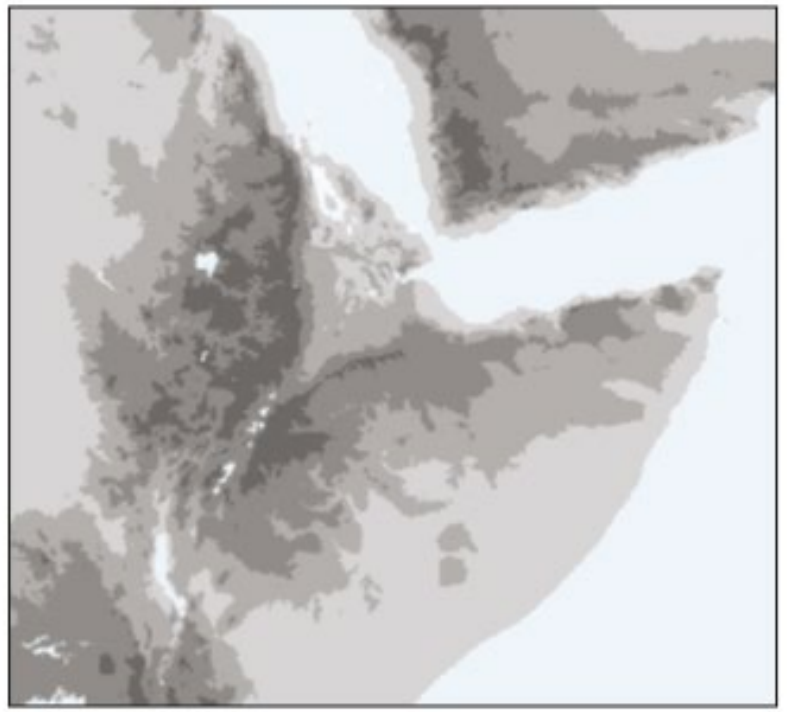
- Saturation: dark purples - near gray to vivid purple, same hue and lightness.
- Lightness: gray through desaturated cyans to the saturated hue.





The Color Wheel : Saturation

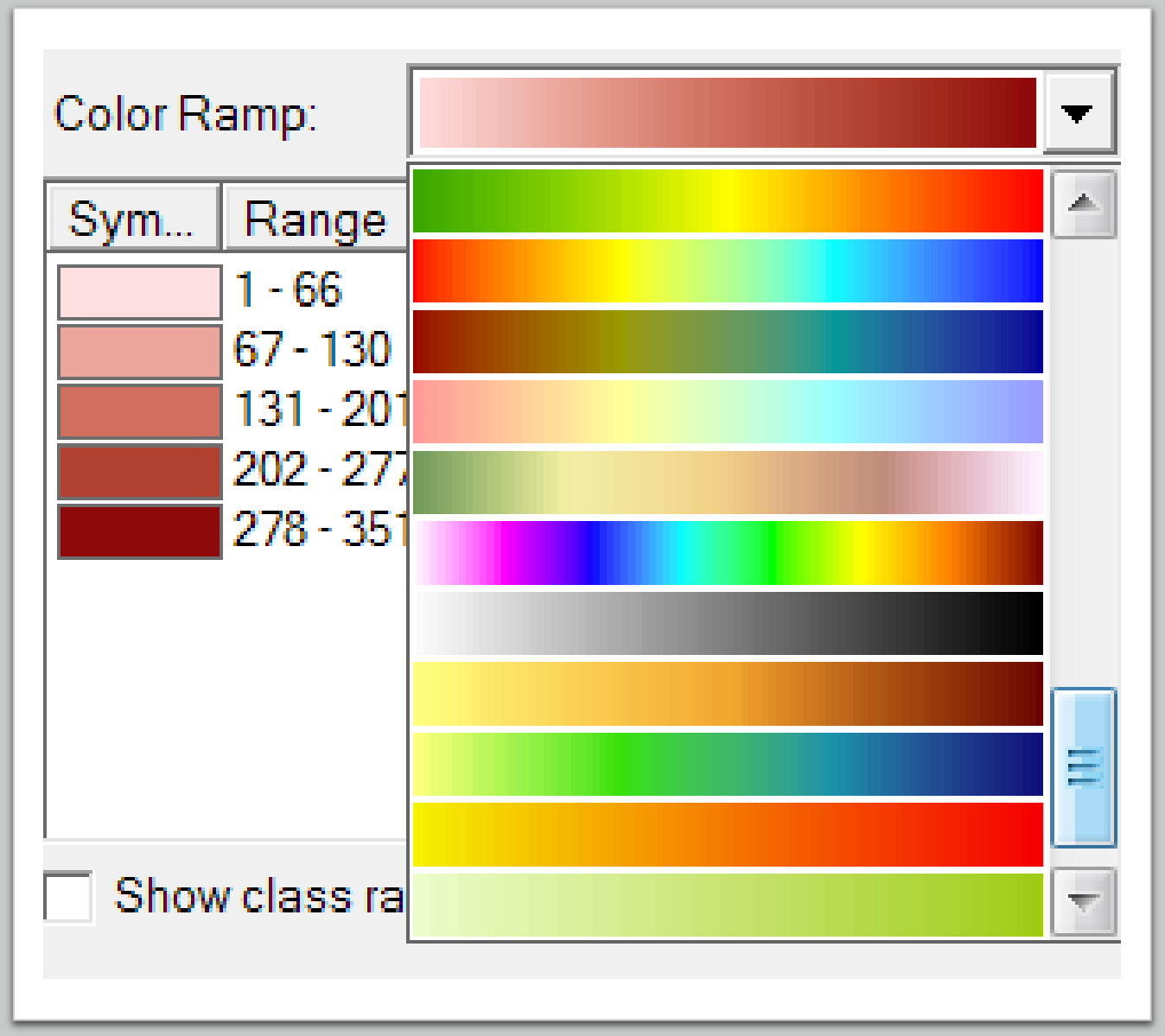




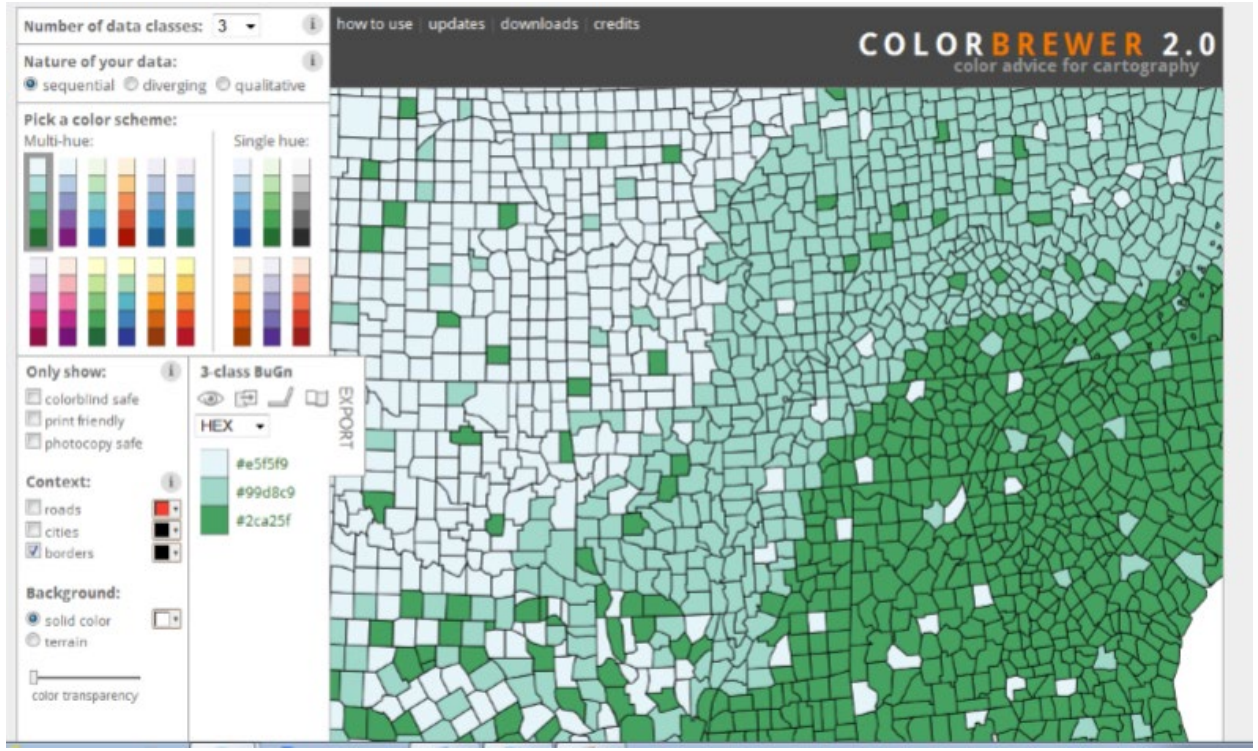
The Color Wheel : Value

- "the quality by which we distinguish a light color from a dark one."
- Albert Henry Munsell
- *A Colour Notation* 1905
-
- **Value** represents the luminescent contrast value between black and white

ArcMap Symbology Tab



Color Brewer



Summary

GIS stores both the map and the attributes associated with the features on the map

In order to understand GIS, we need to know a little cartography

GIS goes beyond just mapping by ability to manage and analyze spatial information