Deck 6: Midterm Review!

Intro to GIS – UMass Amherst – Michael F. Nelson

Overview

- Practical Exam Info
- Coordinate Systems
- Spatial Data
- Selections and Geoprocessing
- Cartography

Announcements

- Final poster analytical proposal (checkpoint 2) is due next Monday!
  - Time is flying by.
  - There are some very cool projects this term!
  - All on-time submissions were graded, and feedback provided.
  - I'll do a sweep to catch any late assignments today.

Midterm Logistics

- The midterm will automatically open on Friday June 16th at 12:01 AM, and close on Friday June 23rd at 11:59PM. This gives you about 8 days to complete it.
- It is formatted as a Moodle assignment; there is no downloadable pdf.
  - You'll see the questions within the Moodle assignment page.
  - This is not meant to be a public-facing assignment.
- You'll submit a single pdf midterm report document.

Midterm Resources

- You **can** use:
  - Your notes and old labs
  - Google/Internet/ESRI Help
  - Discussion with others in the class in public rooms during the lab periods.
  - Moodle Midterm Forum
- You **can't** use:
  - Myself or the TA
  - Private communication with others in the class

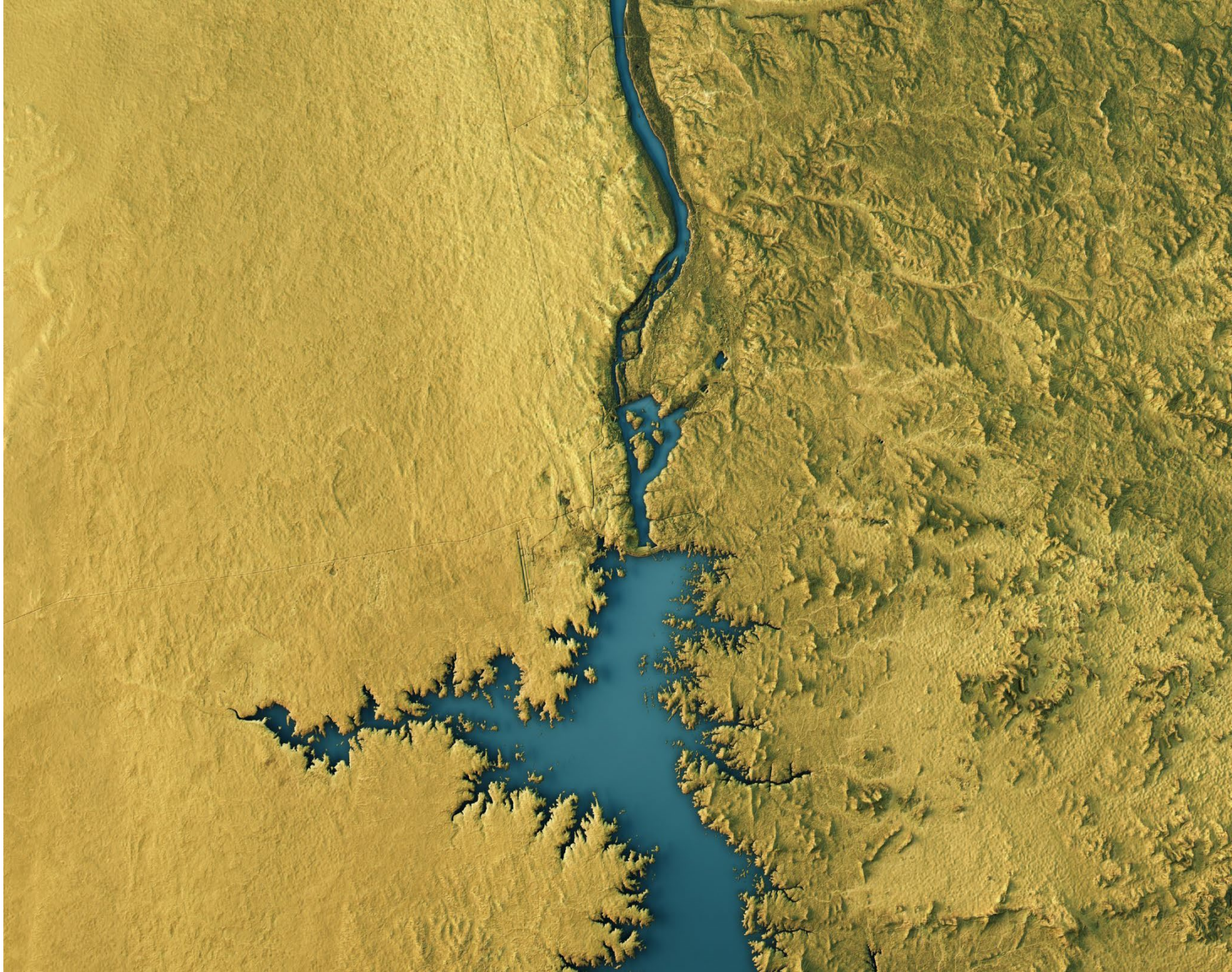
Midterm Resources

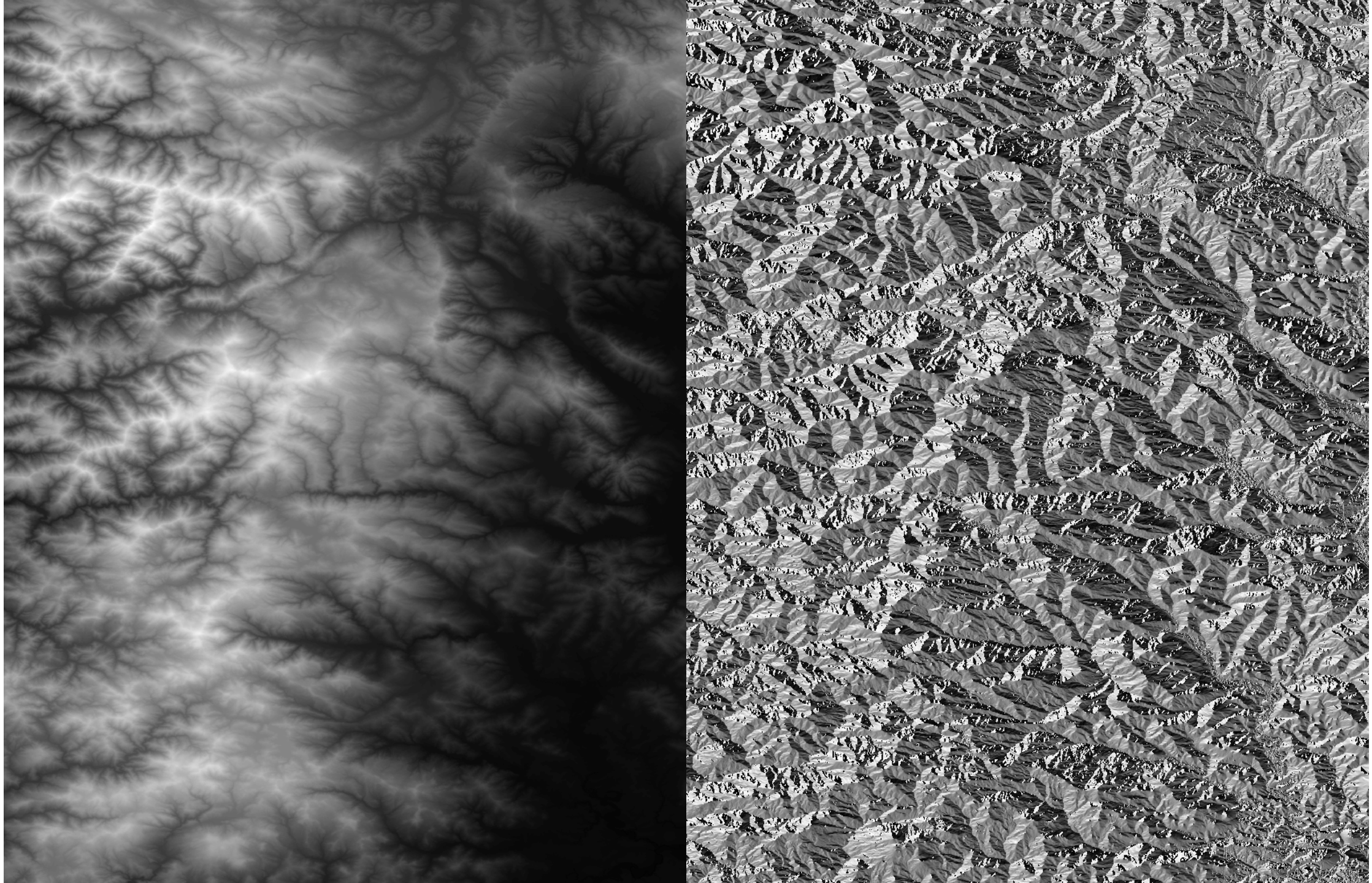
- All the material and practical skills you need to complete the midterm have been covered in the lectures and labs.
- Use the practice exam!
  - It contains an answer key you can use to check your work.
- Use the study guide.

Midterm Content

- There is no new material on the midterm!
- Practical skills:
  - Importing and manipulating data in ArcMap
  - Selections and geoprocessing tools
  - Mapmaking: remember the principles of good map design
- Concepts:
  - Everything we've covered, which is???
  - The lecture notes, practice materials, and labs are your best friends.
  - These review slides…!

Now, a map puzzler!

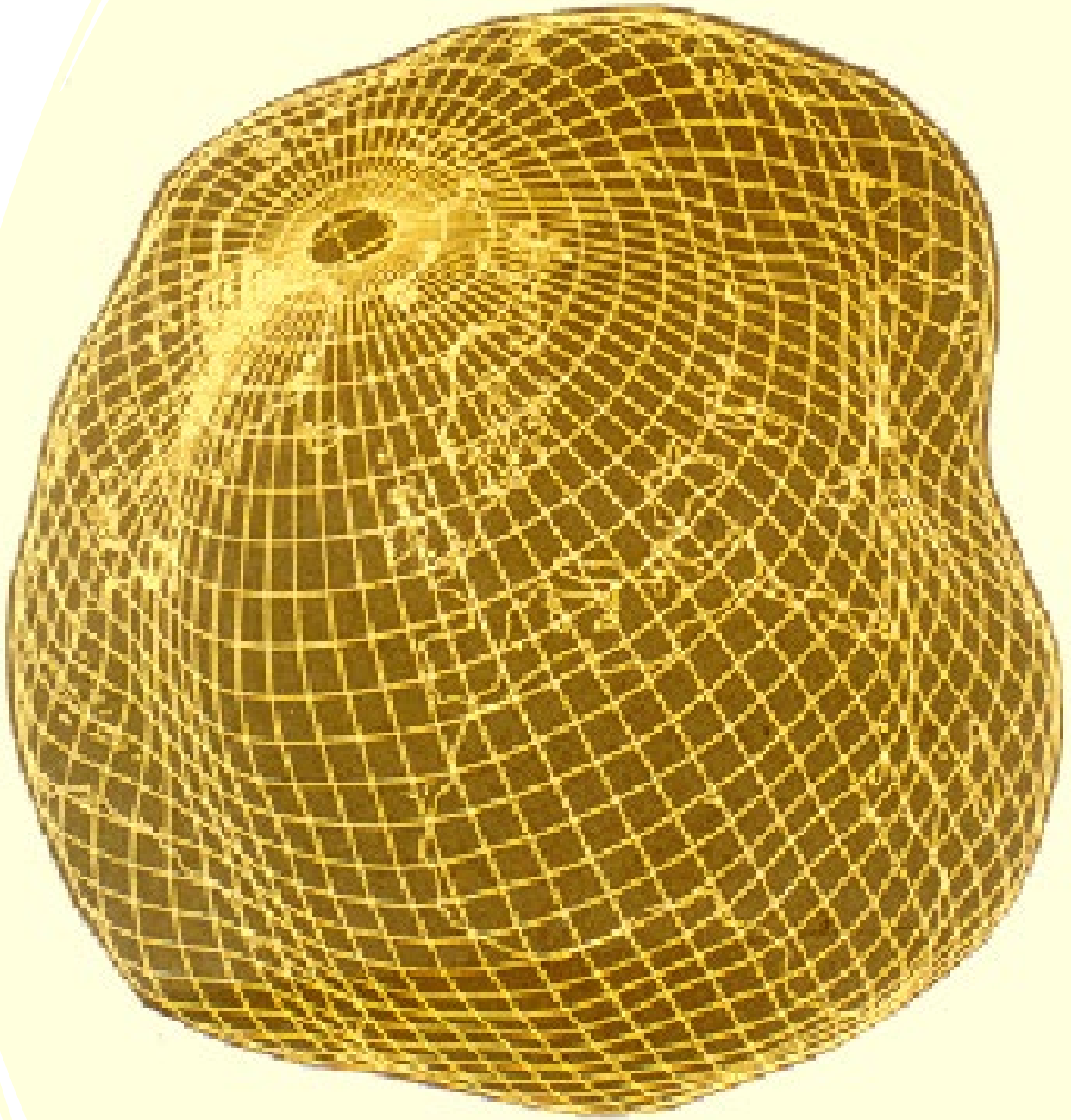




Review: Earth's Shape

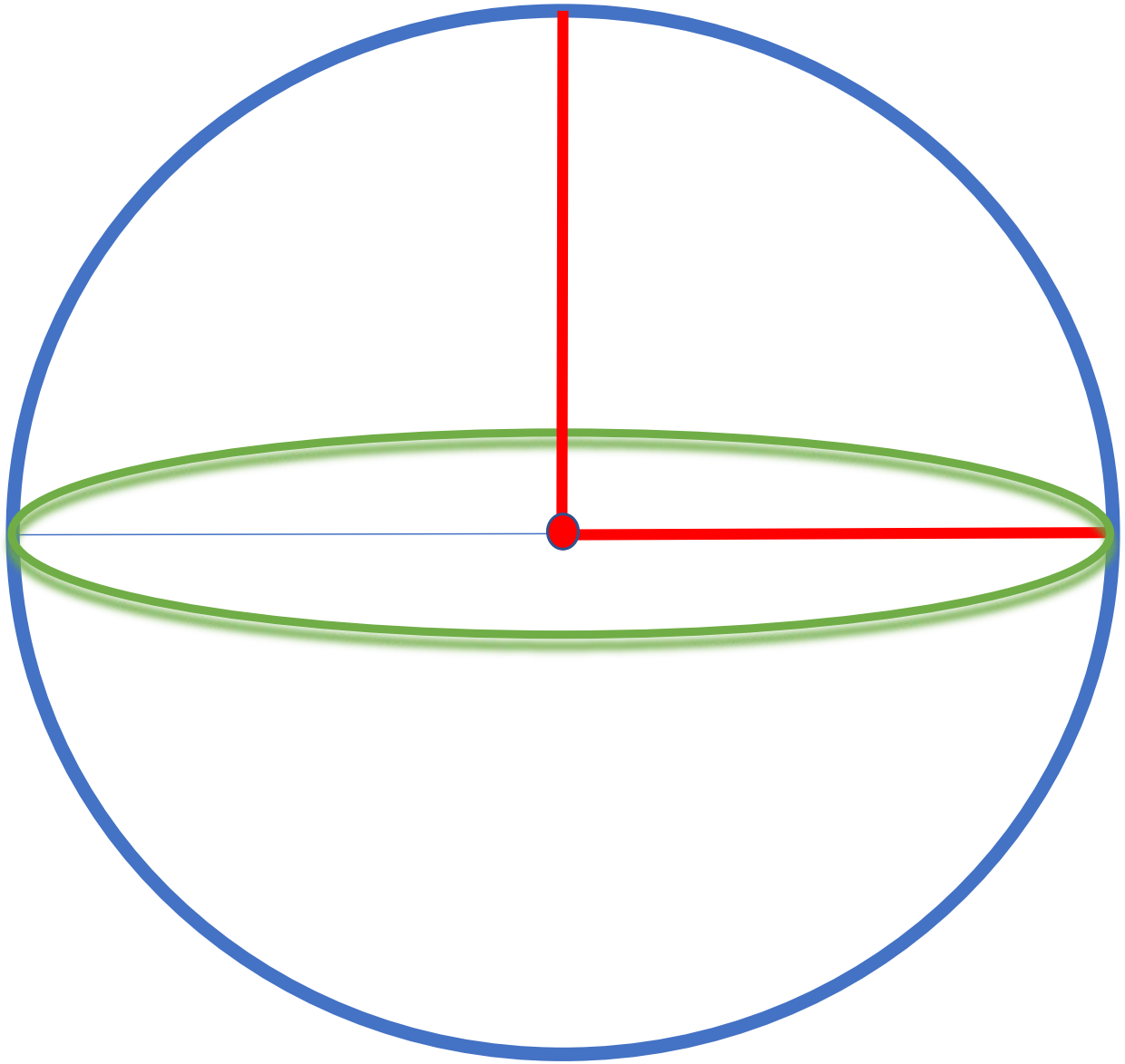
Coordinate Systems, Projections, and Maps

What is Earth's shape?



Model Thinking: A useful simplification of the earth's shape?

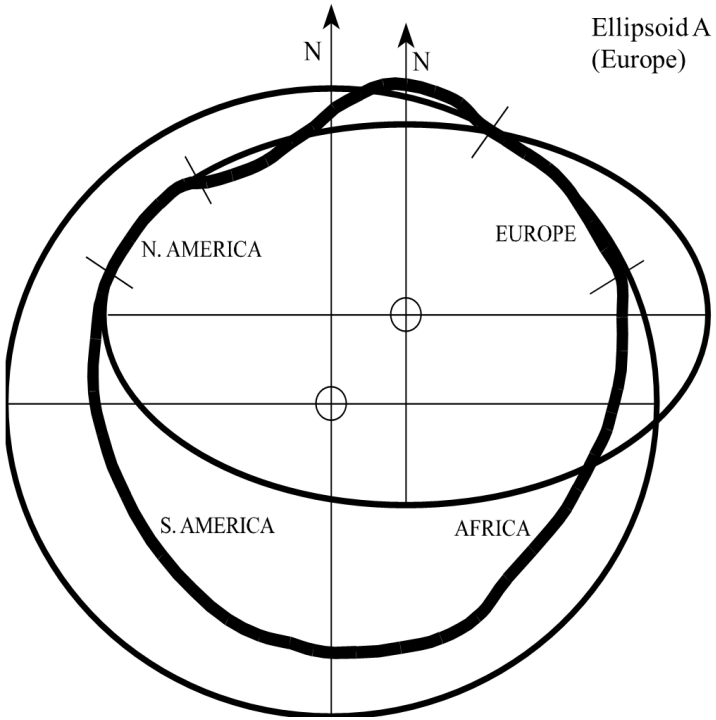
- Flat*?
- Sphere?
- Ellipsoid?
- Lumpy Space Potato?
- Geoid?

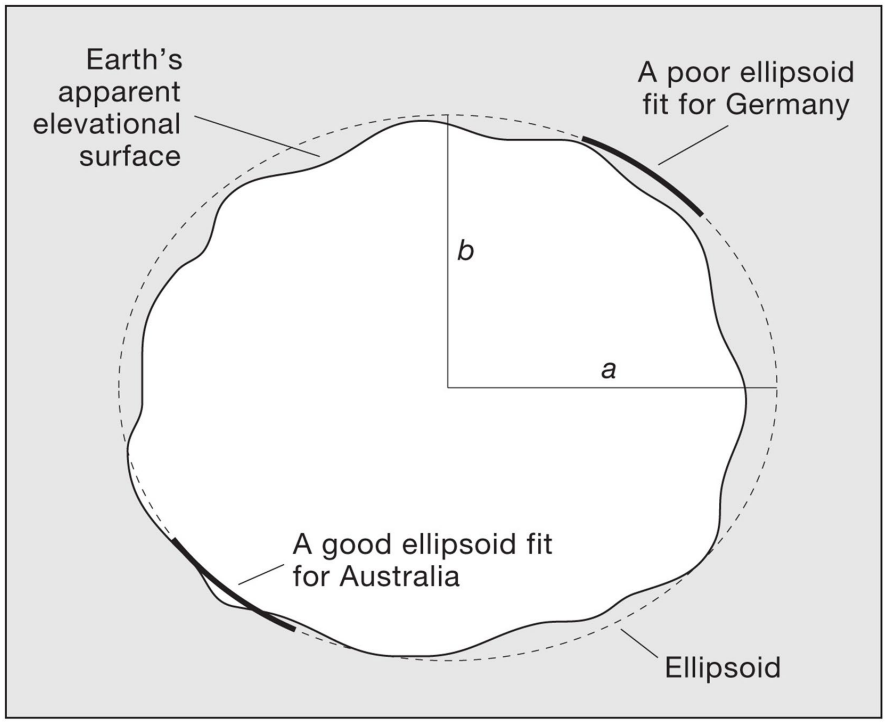


* The earth is not flat.

Local Ellipsoids

Different Ellipsoids are developed to fit the area of interest accurately over the area of interest

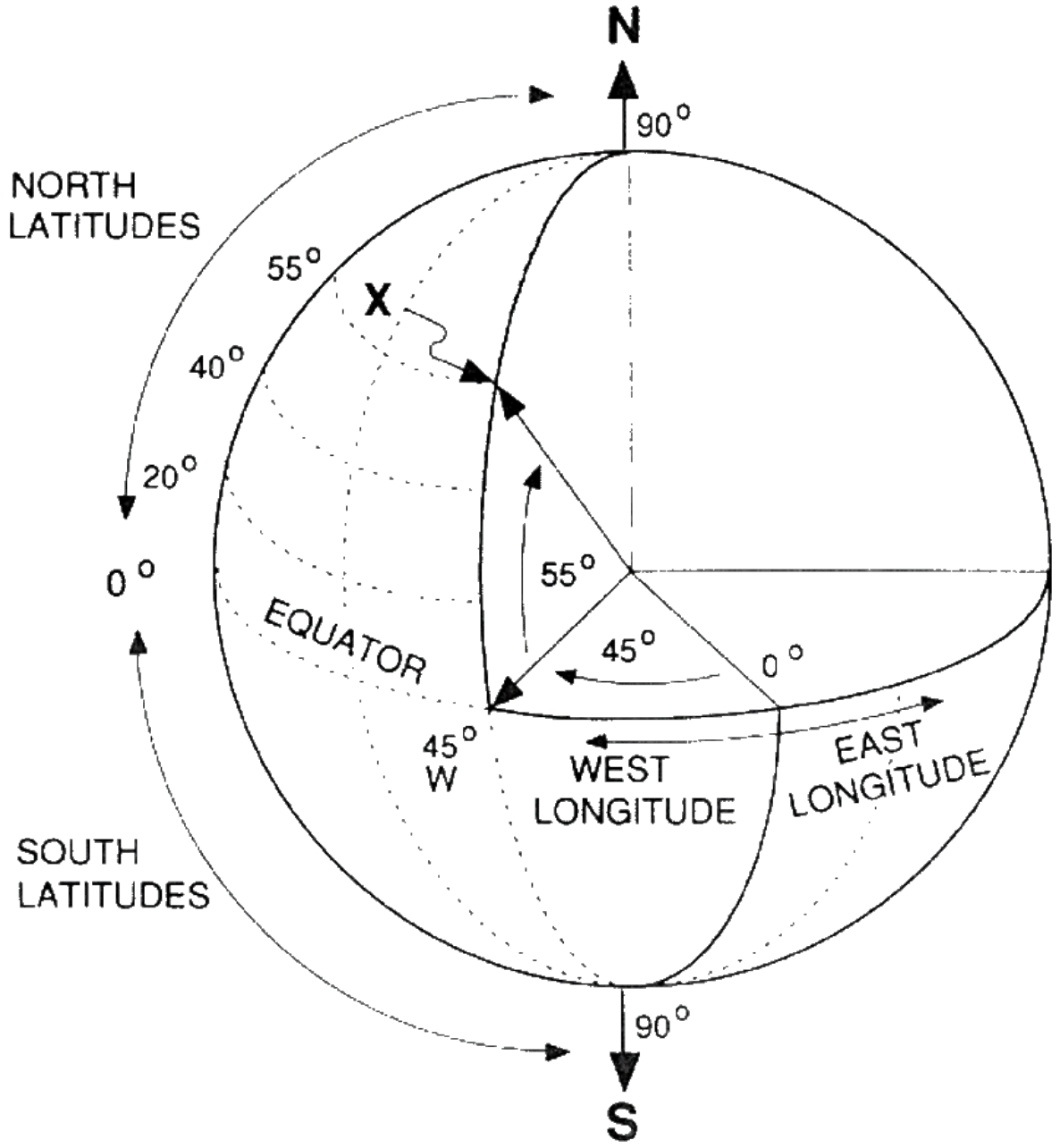




Spherical Coordinate System (2D)

Latitude: degrees (°)
North or South of the Equator

Longitude: degrees (°)
East or West of The Prime Meridian



What is a Datum?

*In surveying and geodesy, a **datum** is a reference point or surface against which position measurements are made, and an associated model of the shape of the earth for computing positions*

http://en.wikipedia.org/wiki/Geodetic_system

- A datum is a reference system with **two** components:
  - A specified **ellipsoid with a spherical coordinate system** and an **origin**
  - A set of surveyed **points** and lines to anchor the ellipsoid
- There are *Regional* and *Global* Datums.

Coordinate Systems, Projections

- Geographic Coordinate System (GCS): uses degrees longitude and latitude.
  - ‘Unprojected coordinates’
  - Not great for mapmaking. Hint: do not use a GCS for your final project posters!
- Projected Coordinate System (PCS): applies a mathematical function to the coordinates of a GCS.
  - ‘Projected coordinates’ or ‘projections’
  - A PCS attempts to minimize, or manage, distortion.
  - Three main types: cylindrical, conical, planar (azimuthal)

Lumpy Space Potato

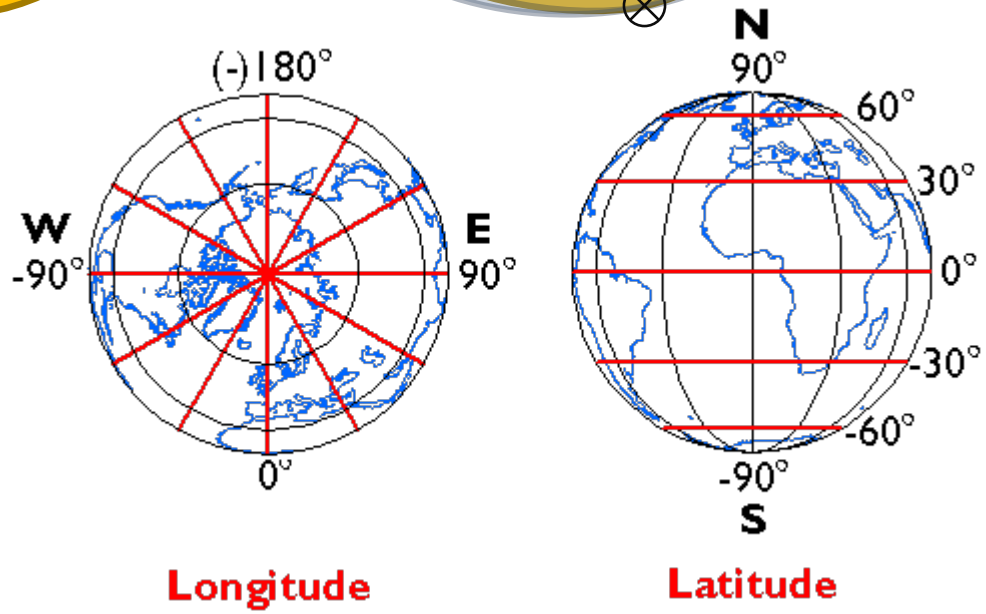
Geoid

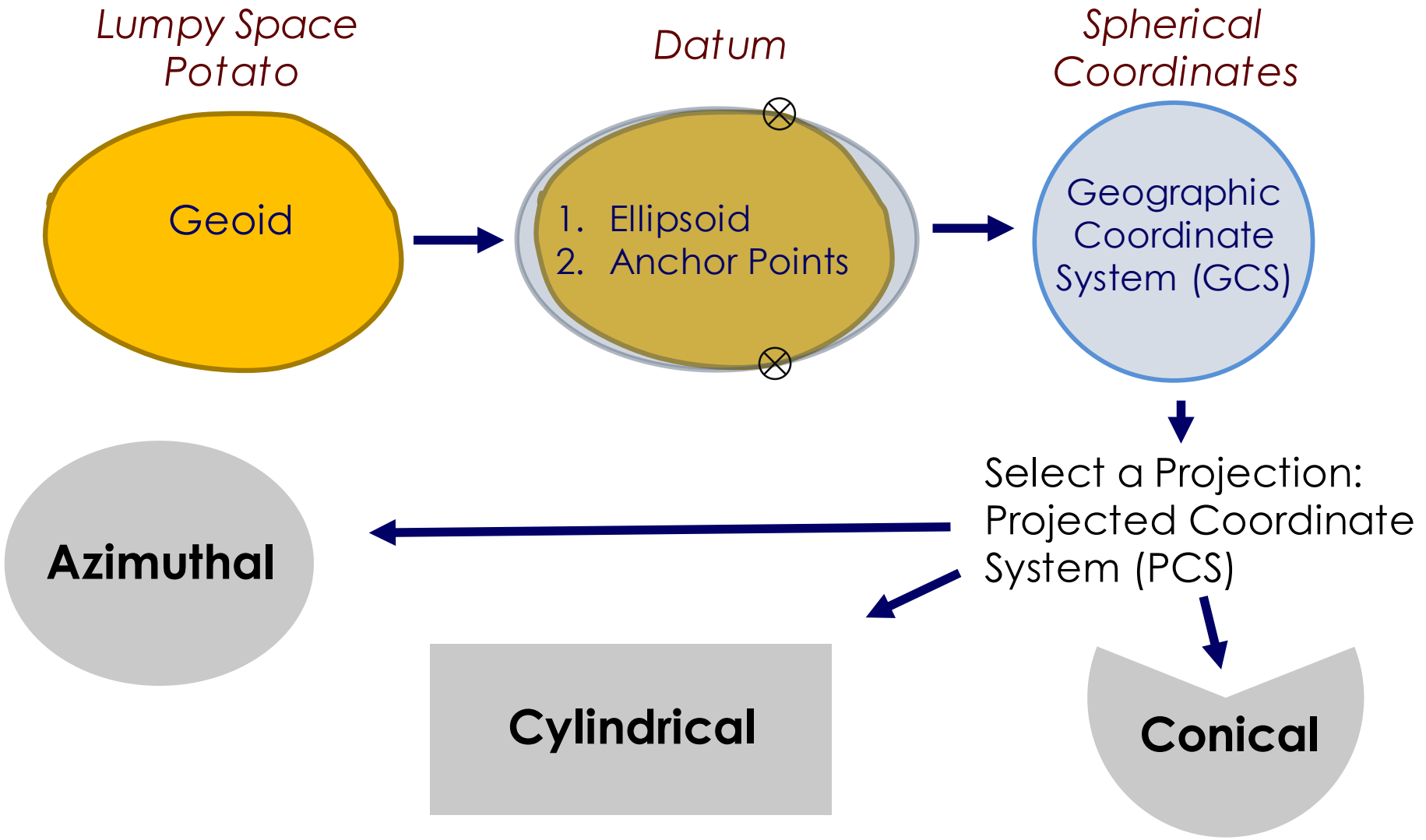
Datum

1. Ellipsoid
2. Anchor Points

Spherical Coordinates

Geographic Coordinate System (GCS)




Lumpy Space Potato
Geoid
Datum
1. Ellipsoid
2. Anchor Points
Spherical Coordinates
Geographic Coordinate System (GCS)
Select a Projection: Projected Coordinate System (PCS)
Azimuthal
Cylindrical
Conical

Map Classes: maps can preserve:

1

Size: equal-area

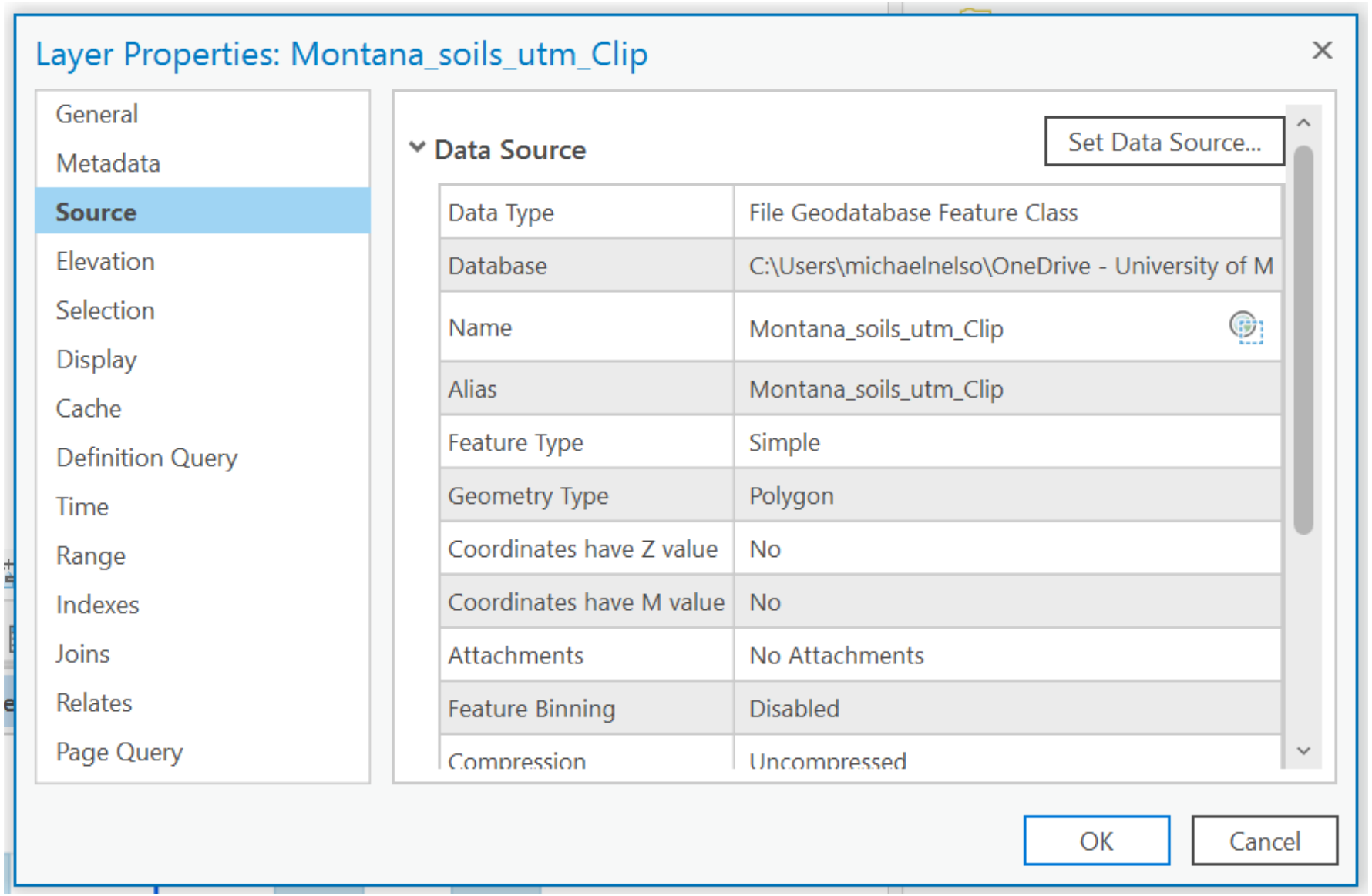
2

Distance: equidistant

3

Shape: conformal

Need projection info about your data?



Use the source (tab)

Coordinate System and Map Resources

Lab 5 is all about coordinates!

Use the coordinate system supplement in Slide Deck 5.

Review: Spatial Data

Vector and Raster Data Models

Vector (Feature) Data

- Vectors can represent:
  - Points
  - Lines
  - Polygons
- All vector data are built from points (vertices)
  - Each point has x- and y-coordinates
- The vector data model associates locations and attributes.

Vector Data Model

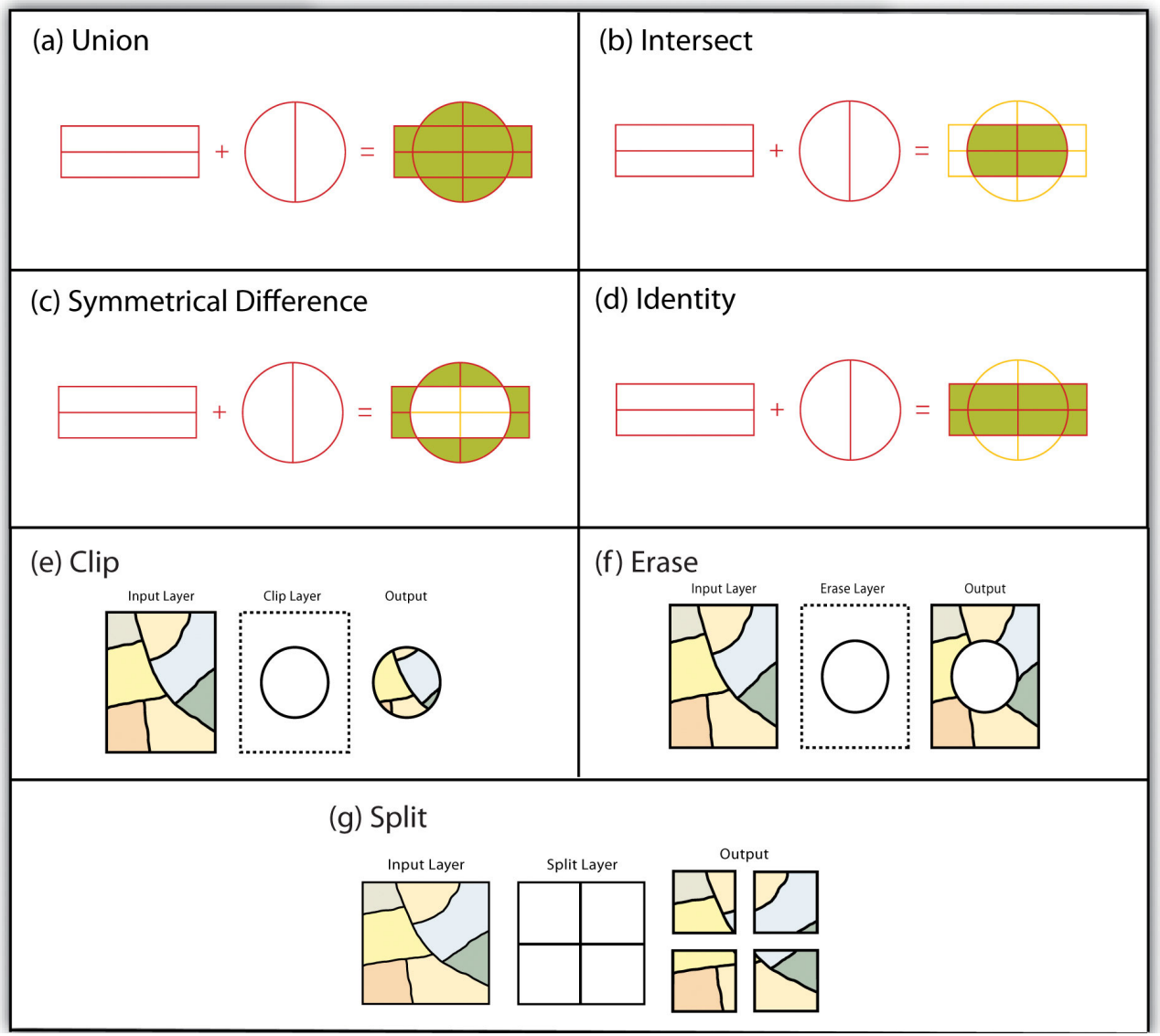
The vector data model associates location data to attributes.

Vector paradigm uses the row-data paradigm for attributes.

Location Data: stores the spatial information as vertices with explicit x- and y-coordinates.

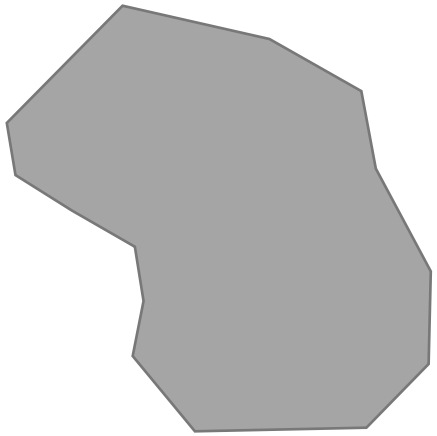
Attribute: stores the associated data values. Attributes by themselves have no spatial information.

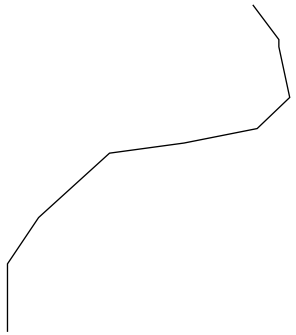
Feature: This is a spatial entity – combines location information with attributes. A feature may have zero or more attributes

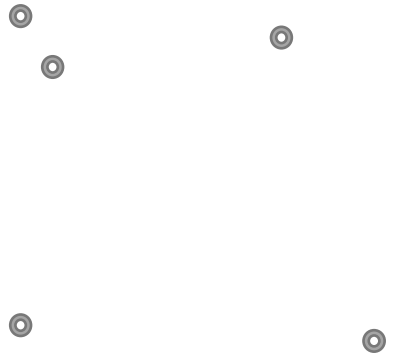


Vector Operations

How should you represent a feature?







As a point?

As a polygon?

As a line?

The choice depends on your goals.

How are raster and vector different?

Raster vs. Vector

Vector: vertices and edges

- Location with explicit x and y coordinates.
- Transforming coordinate systems is reversible.
- Location and attribute data are separate, but associated, entities.

Raster: grid

- Locations implicitly defined by corner coordinates and row/cell indices.
- Transforming is not reversible – it's a 'lossy' operation. Think about resizing or rotating a digital image.
- Arc can temporarily display rasters in a different coordinate system.
- The location and attribute data are the same.

Manipulating Spatial Data

Selections and Geoprocessing

Let’s take a quick poll!

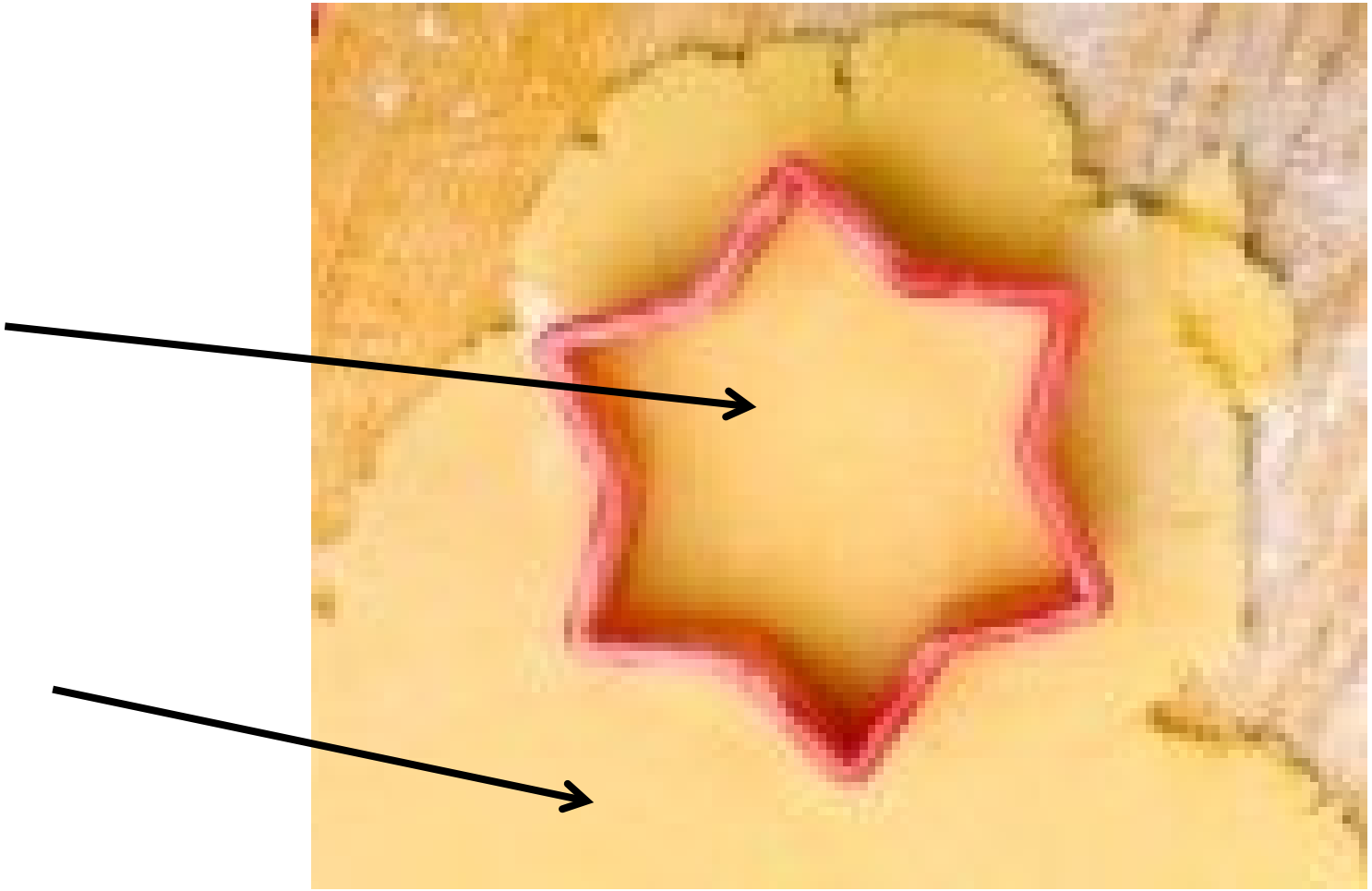
Table

roads_in_MA_towns

STREET_NAM	RT_NUMBER
RAMP-RT 140 NB TO RT 195 EB	
RAMP-RT 140 NB TO RT 195 WB	
RAMP-RT 140 NB TO RT 95 NB	
RAMP-RT 140 SB TO RT 195 EB	
RAMP-RT 140 SB TO RT 195 EB	
RAMP-RT 140 SB TO RT 195 EB	
RAMP-RT 140 SB TO RT 195 WB	
RAMP-RT 140 SB TO RT 495 NB	
RAMP-RT 140 SB TO RT 95 NB	
RAMP-RT 140 TO RT 2 WB	
RAMP-RT 140 TO RT 2 WB	
RAMP-RT 140 TO RT 2 WB	
RAMP-RT 140 TO RT 24 NB	
RAMP-RT 140 TO RT 24 NB	
RAMP-RT 140 TO RT 24 NB	
RAMP-RT 16 EB TO RT 1 NB	

0 (0 out of 18081 S

roads_in_MA_towns



- Two classes of tools
  Selections
  Geoprocessi
  ng

Selections

- What sorts of questions can we answer?

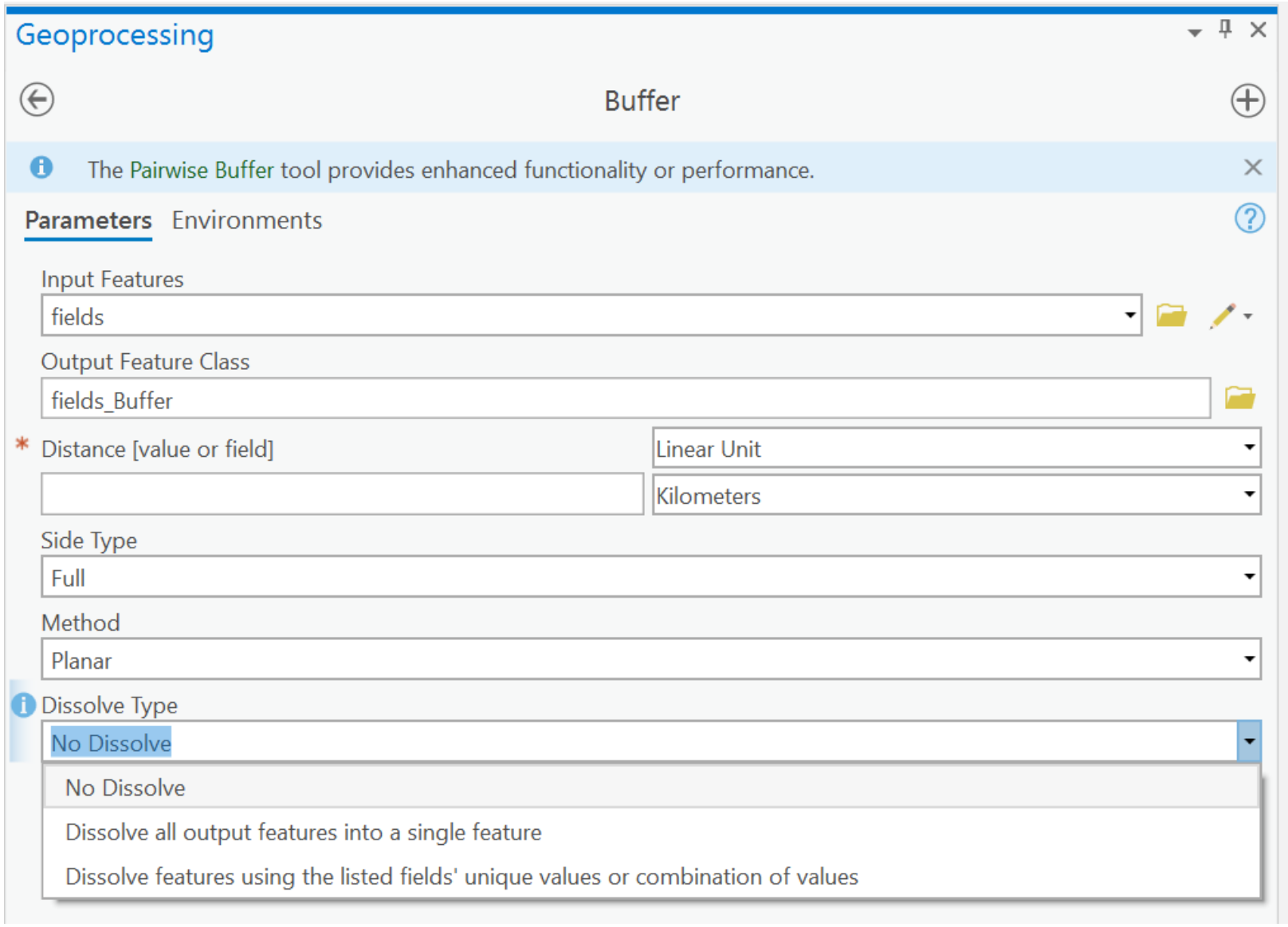
Geoprocessing

- What sorts of questions can we answer?

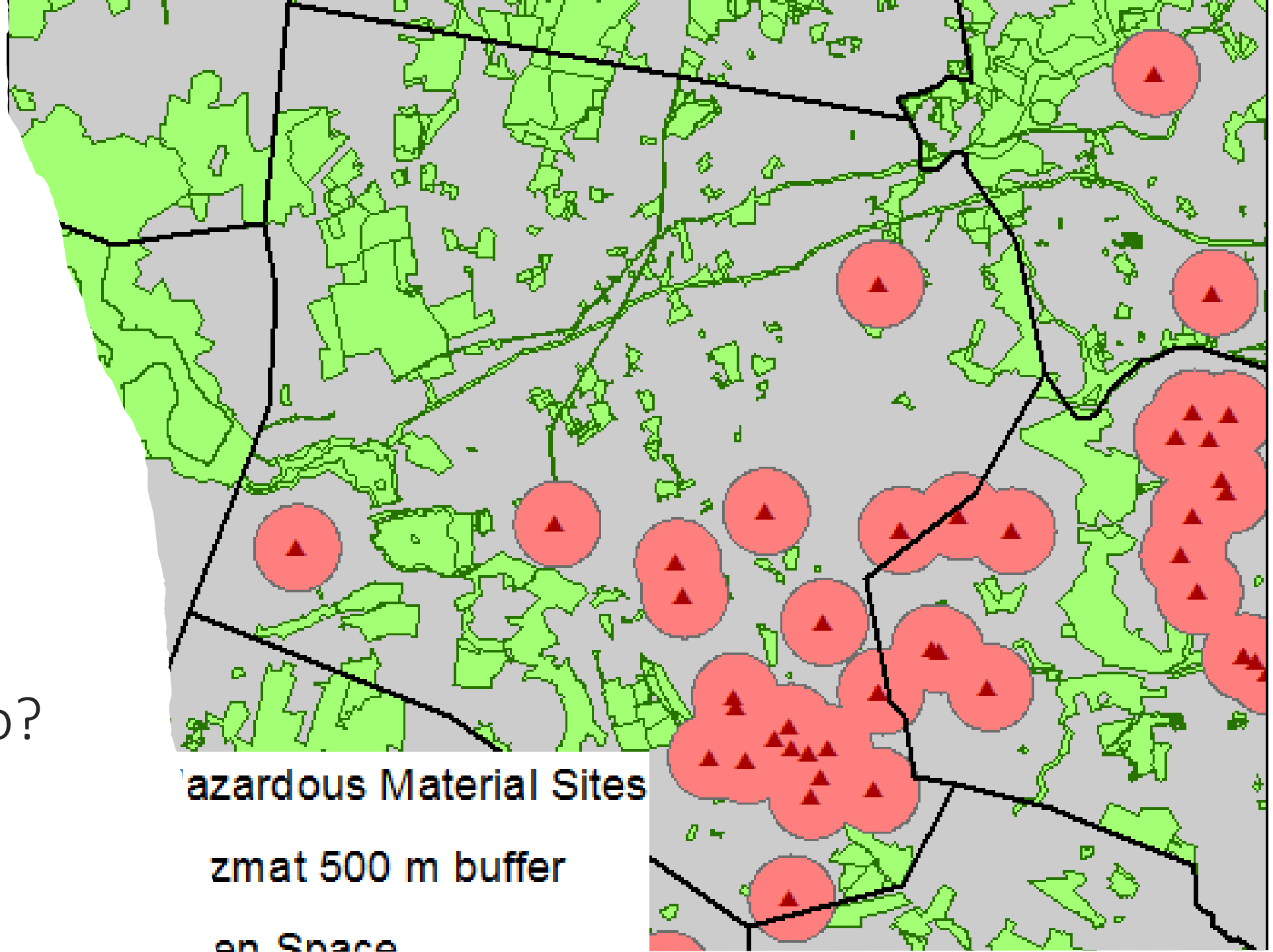
Which Geoprocessing Tools Have We Met So Far?

Buffer!

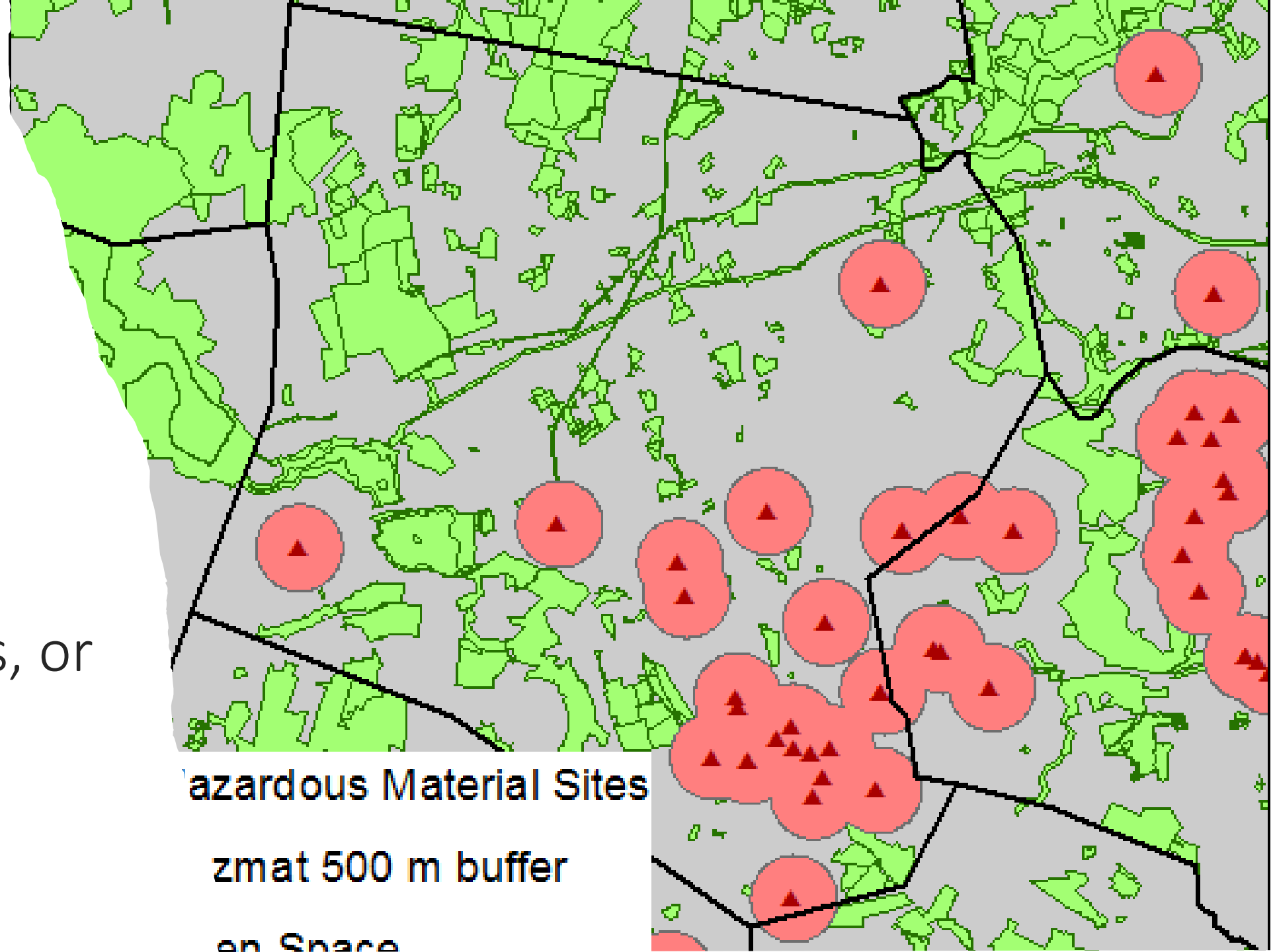
Dissolve options are important!



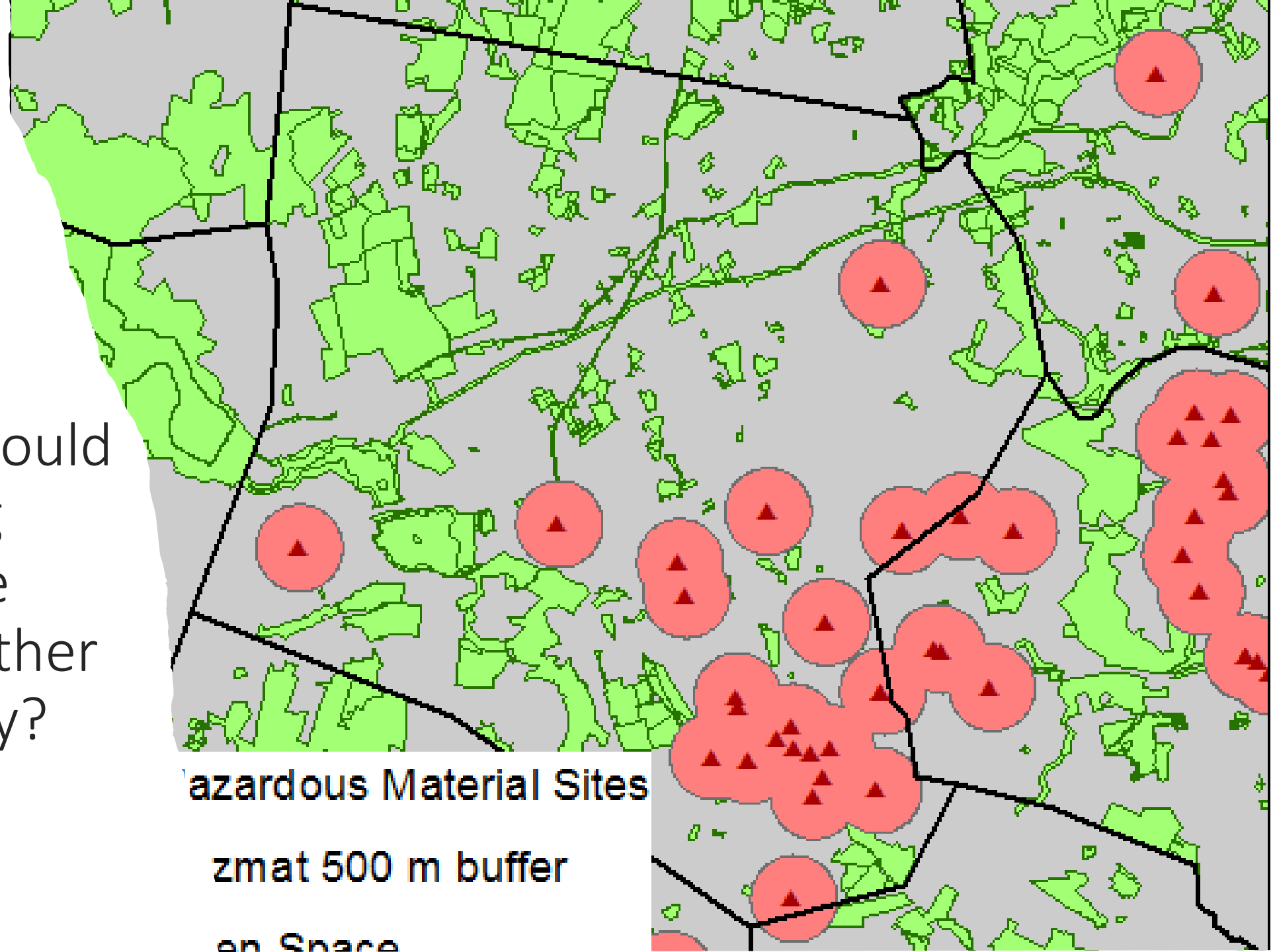
What does buffering do?



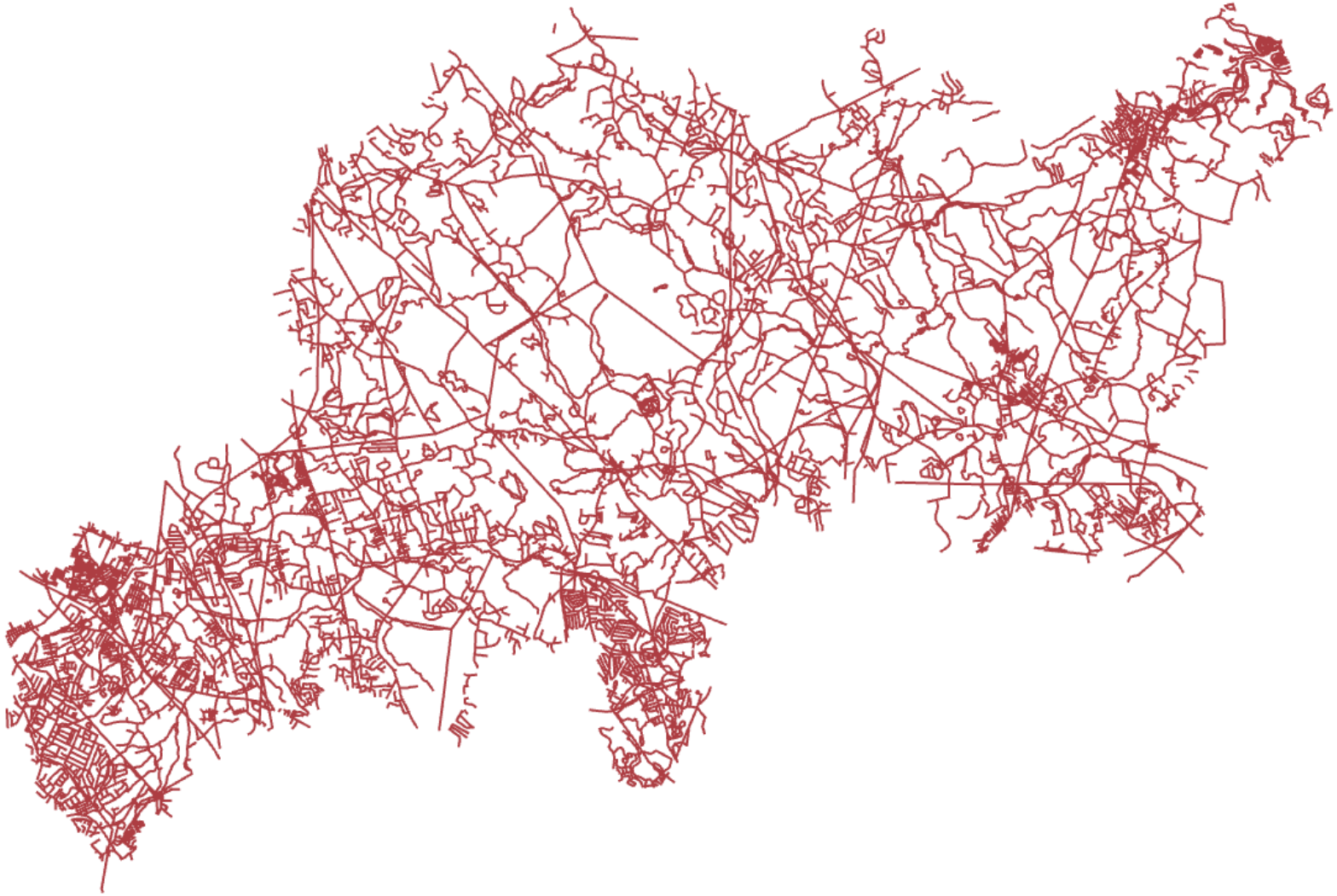
Is buffering faster for points, lines, or polygons?

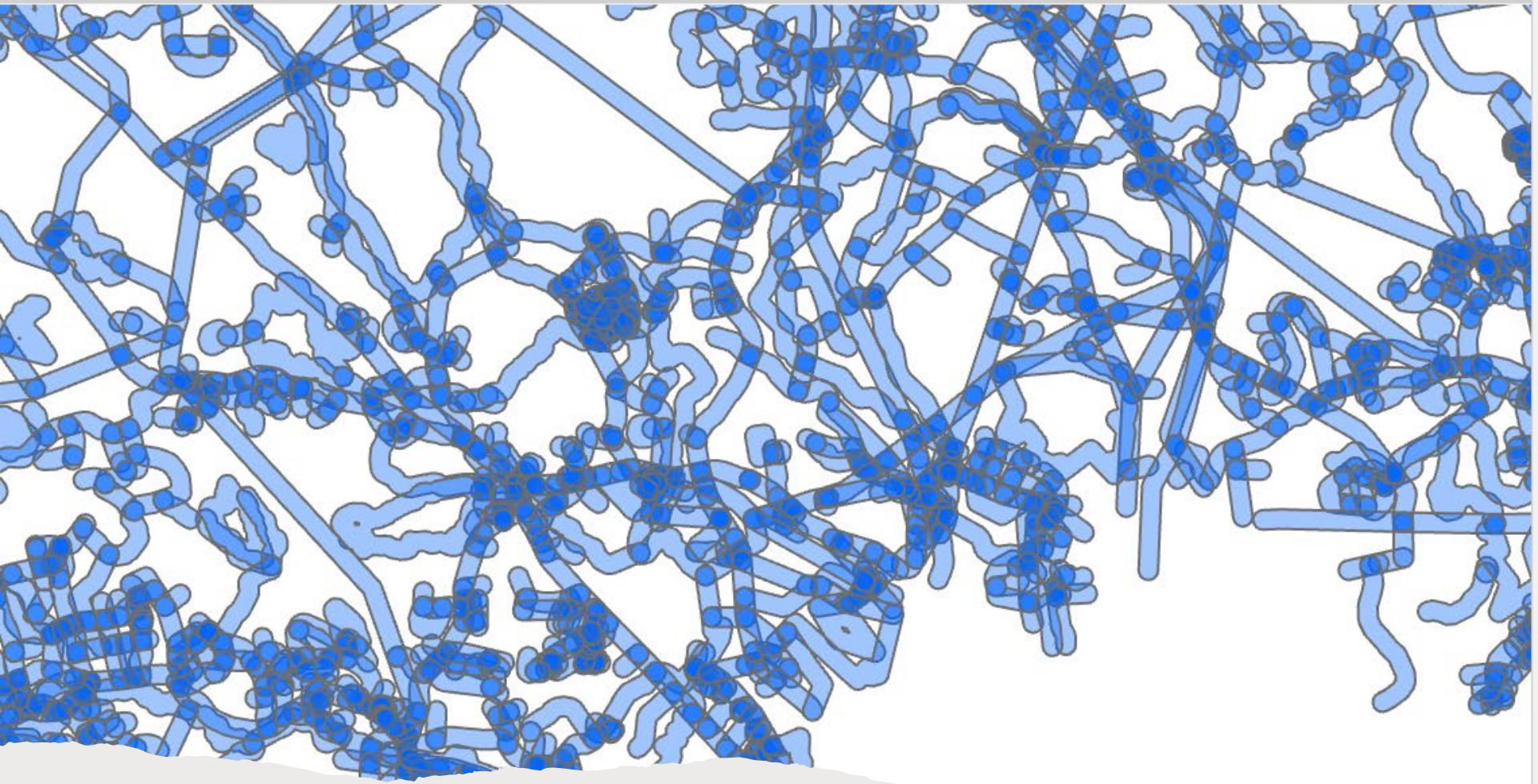


Dissolve: should overlapping polygons be joined together in some way?



Let’s buffer some roads!





Geoprocessing

Buffer

The Pairwise Buffer tool provides enhanced functionality or performance.

Parameters | Environments

Input Features
roads

Output Feature Class
roads_Buffer

Distance [value or field]
Linear Unit
100
Meters

Side Type
Full

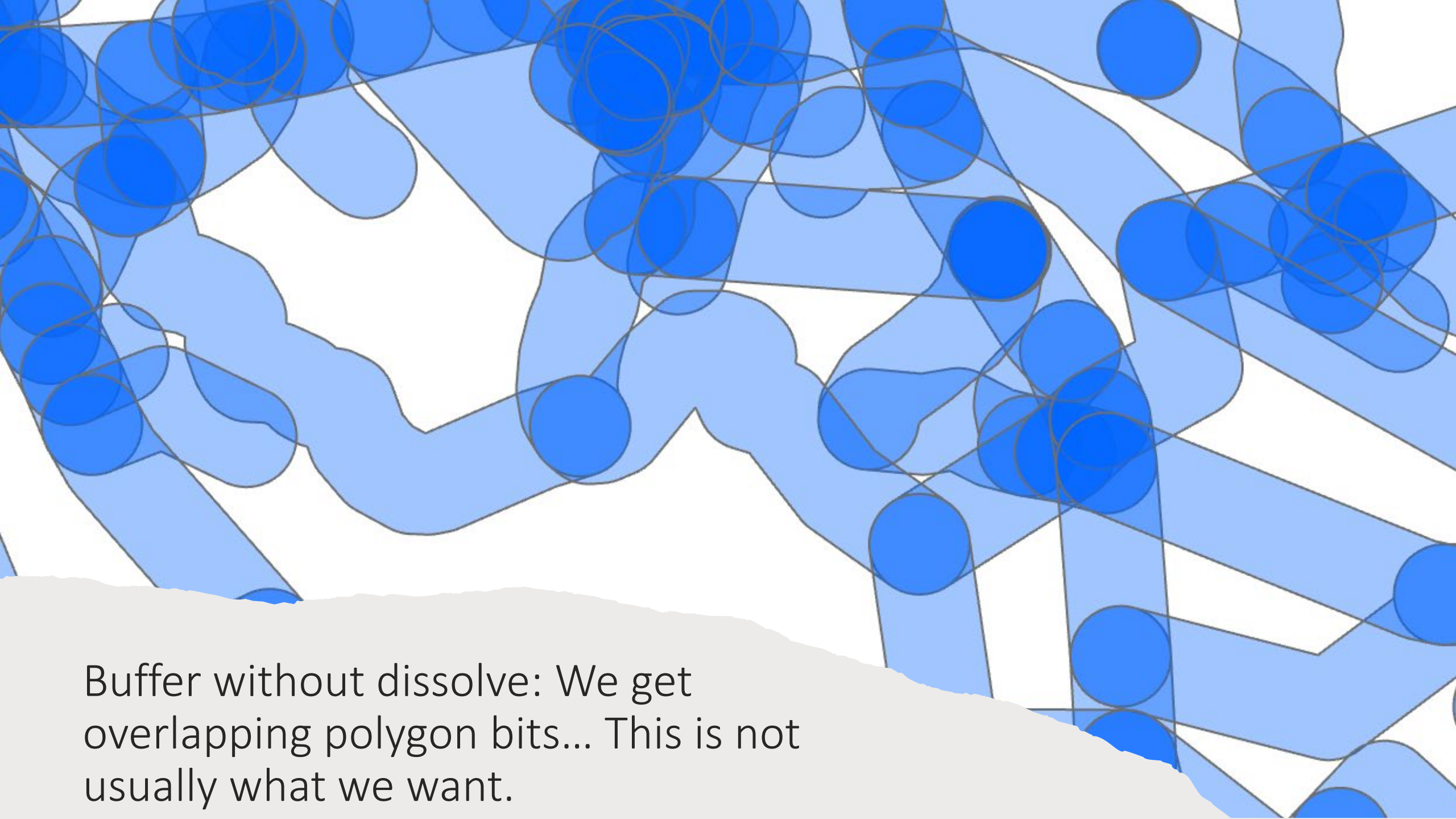
End Type
Round

Method
Planar

Dissolve Type
No Dissolve

Run

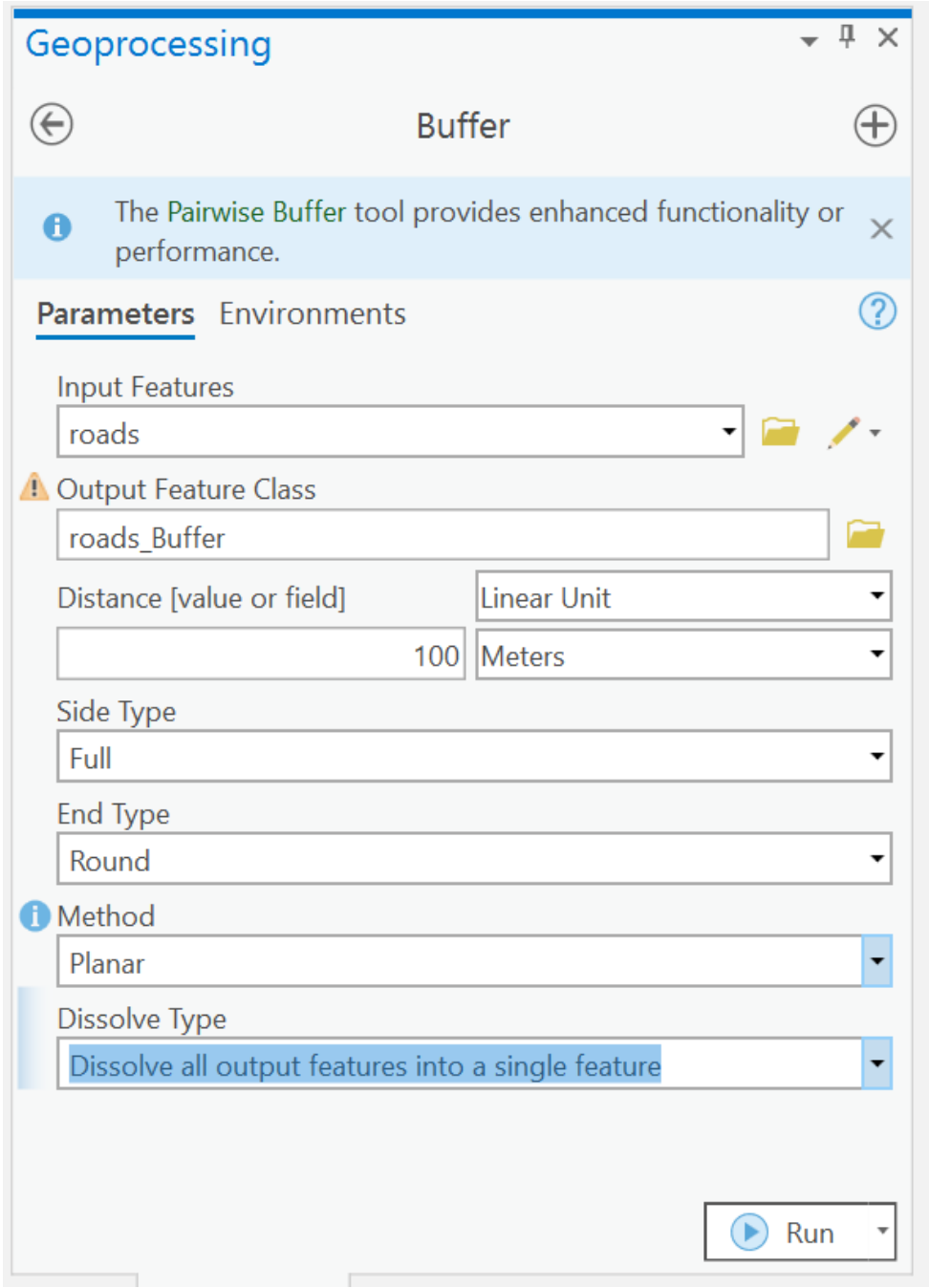
Buffer without dissolve



Buffer without dissolve: We get overlapping polygon bits… This is not usually what we want.

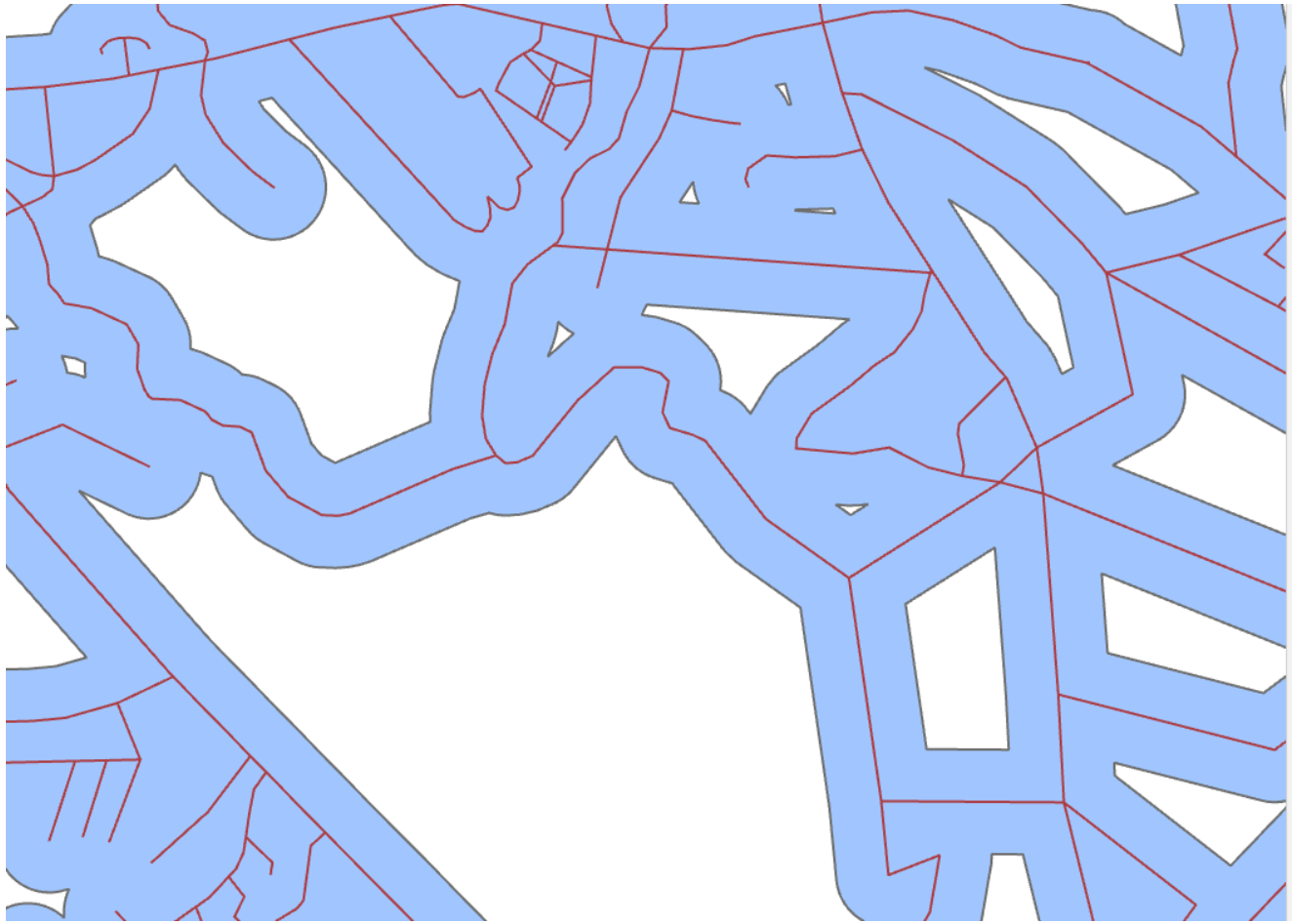
Buffer With Dissolve

This is typically what we want, we end up with a single feature



Look here!

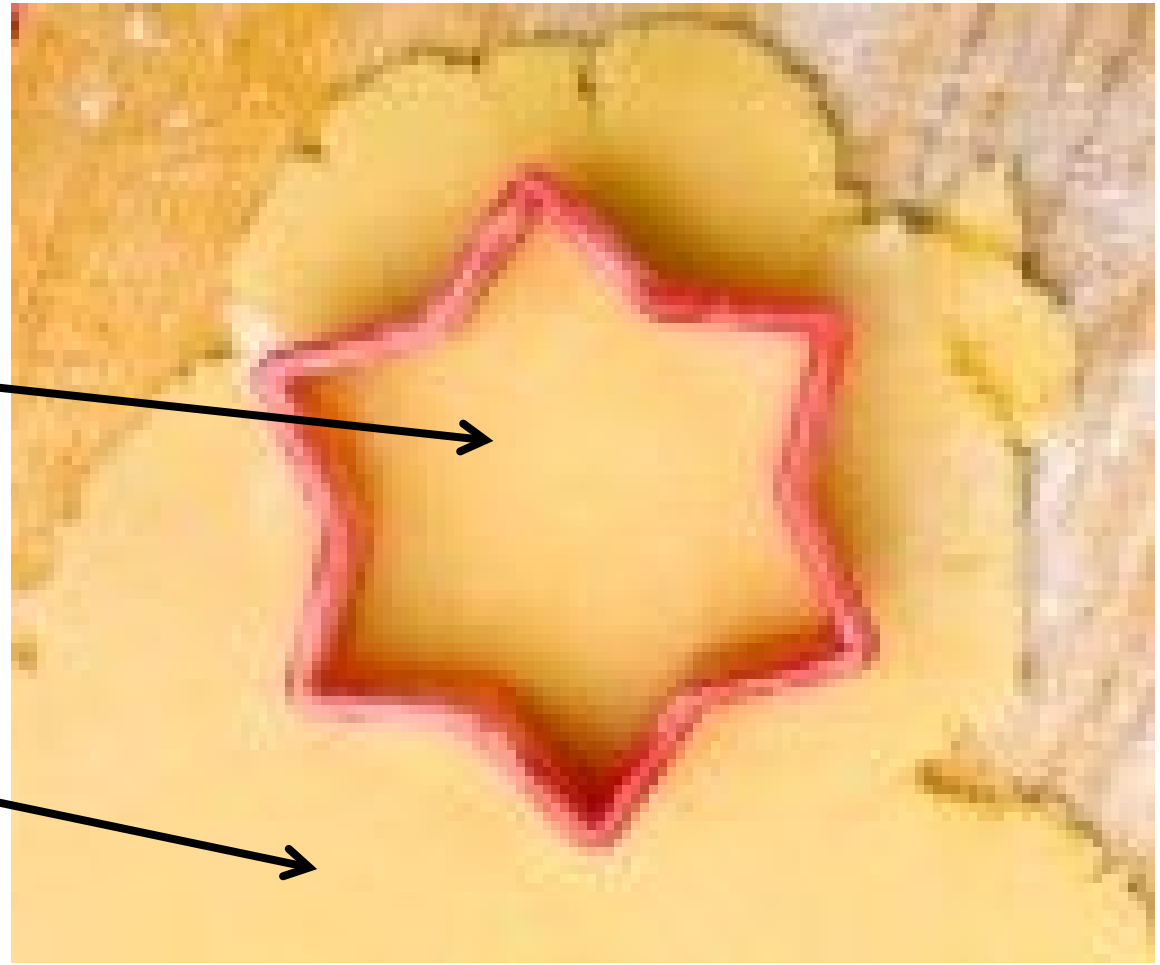
That’s better!



Clip and Erase

Clip keeps the info inside the shape

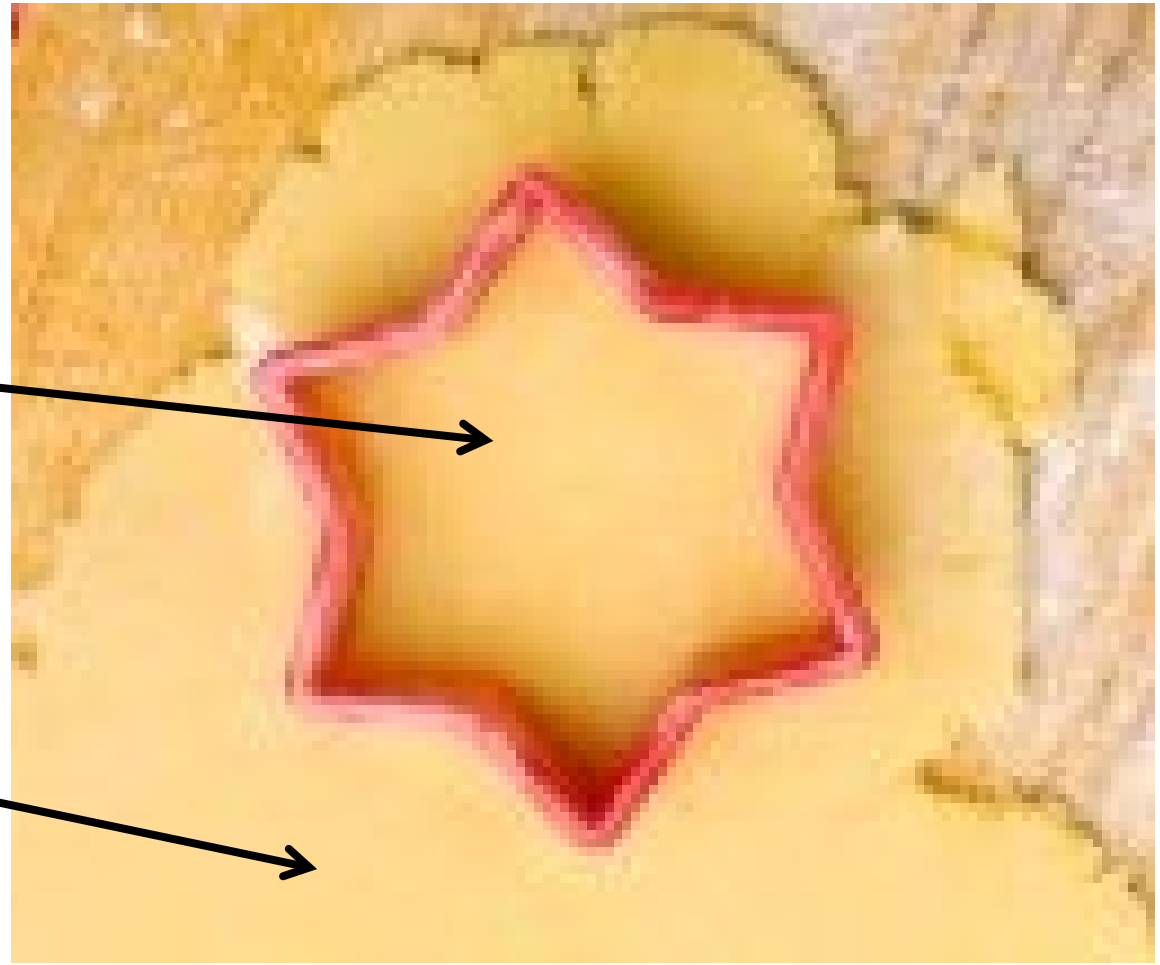
Erase keeps the info outside the shape



Clip and Erase

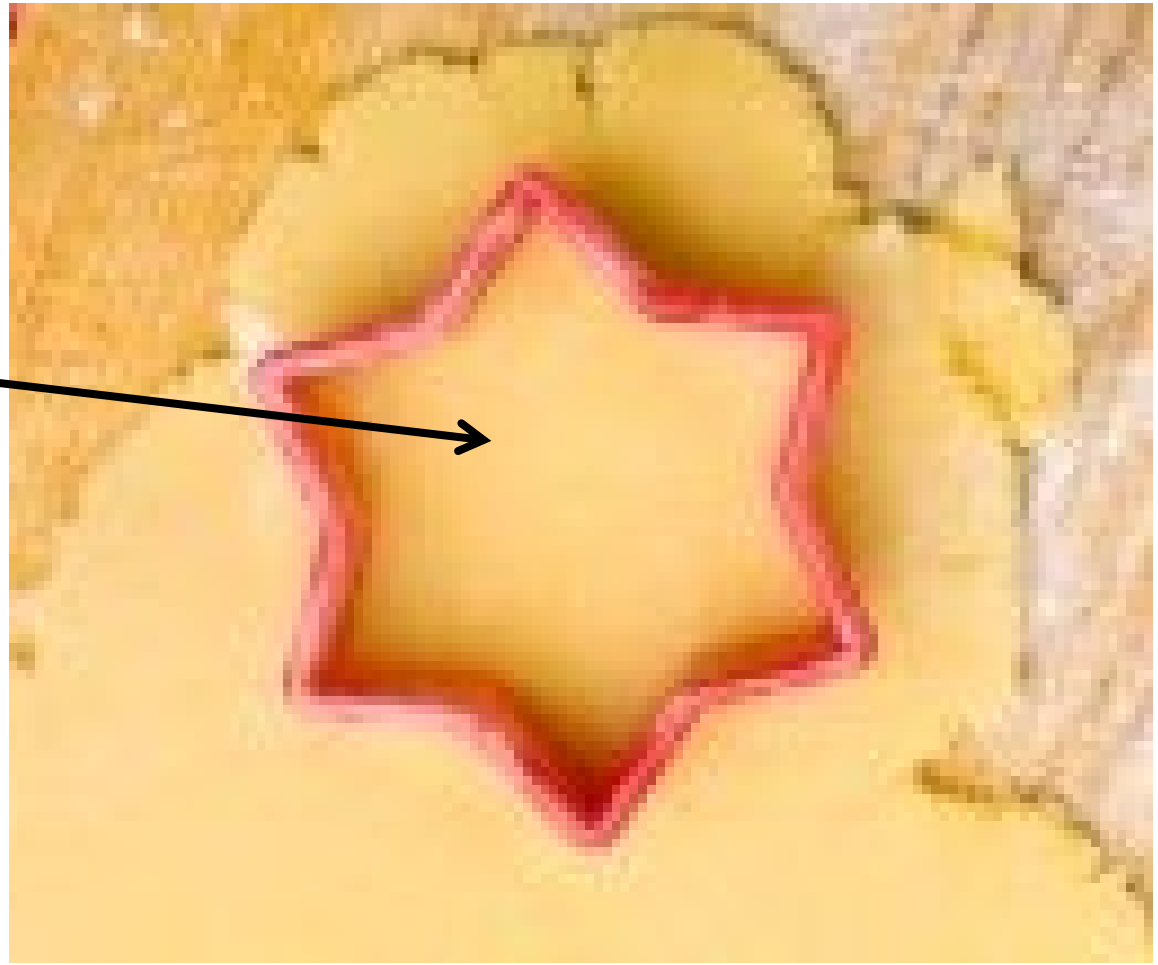
Clip keeps the info inside the shape

Erase keeps the info outside the shape



Intersect

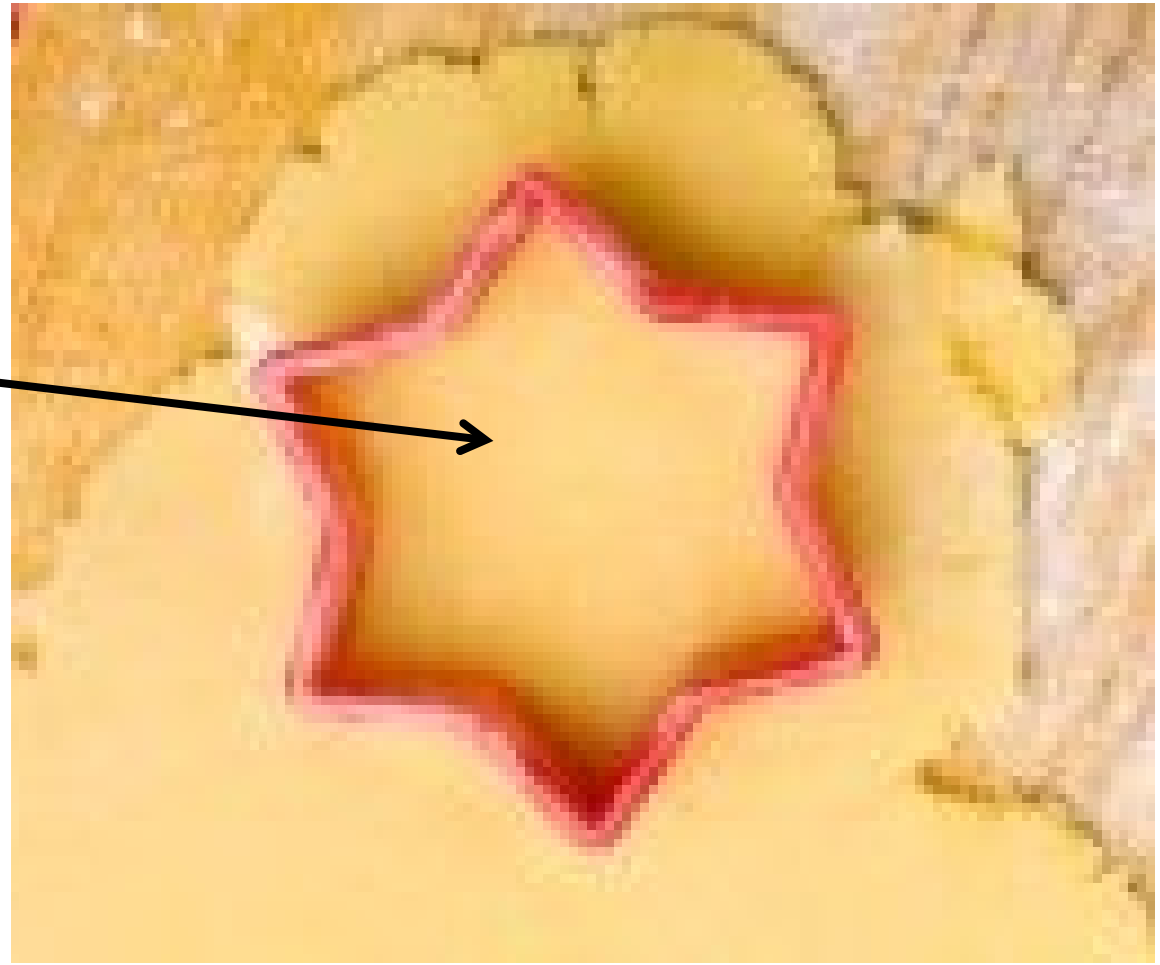
Intersect is like a clip (you end up with the inside), except you retain the attributes from BOTH shapefiles



Intersect

Intersect is like a clip (you end up with the inside), except you retain the attributes from BOTH shapefiles

Intersect will not be on the midterm



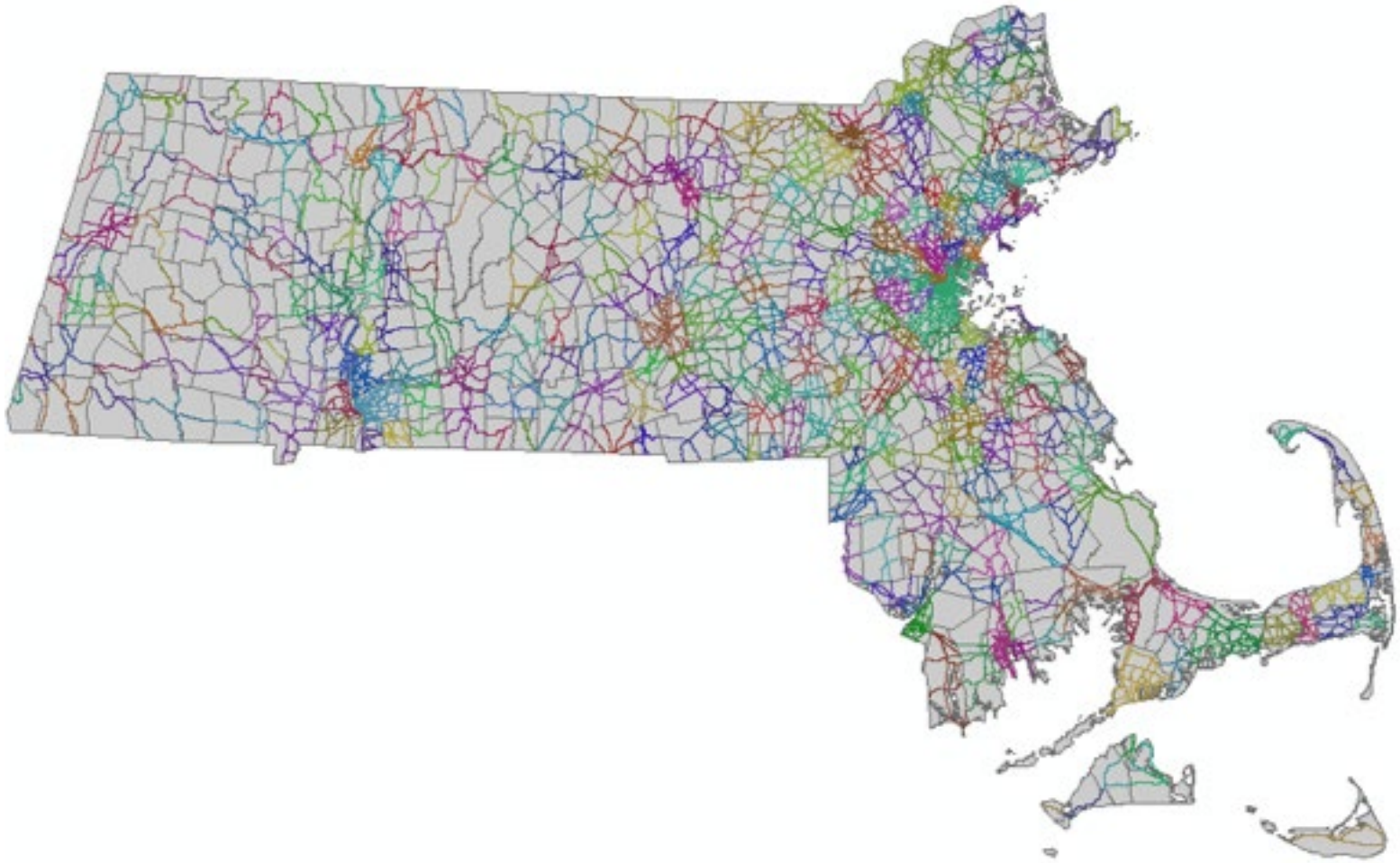
Intersect: Road segments within towns

Intersect combines attributes from two layers.

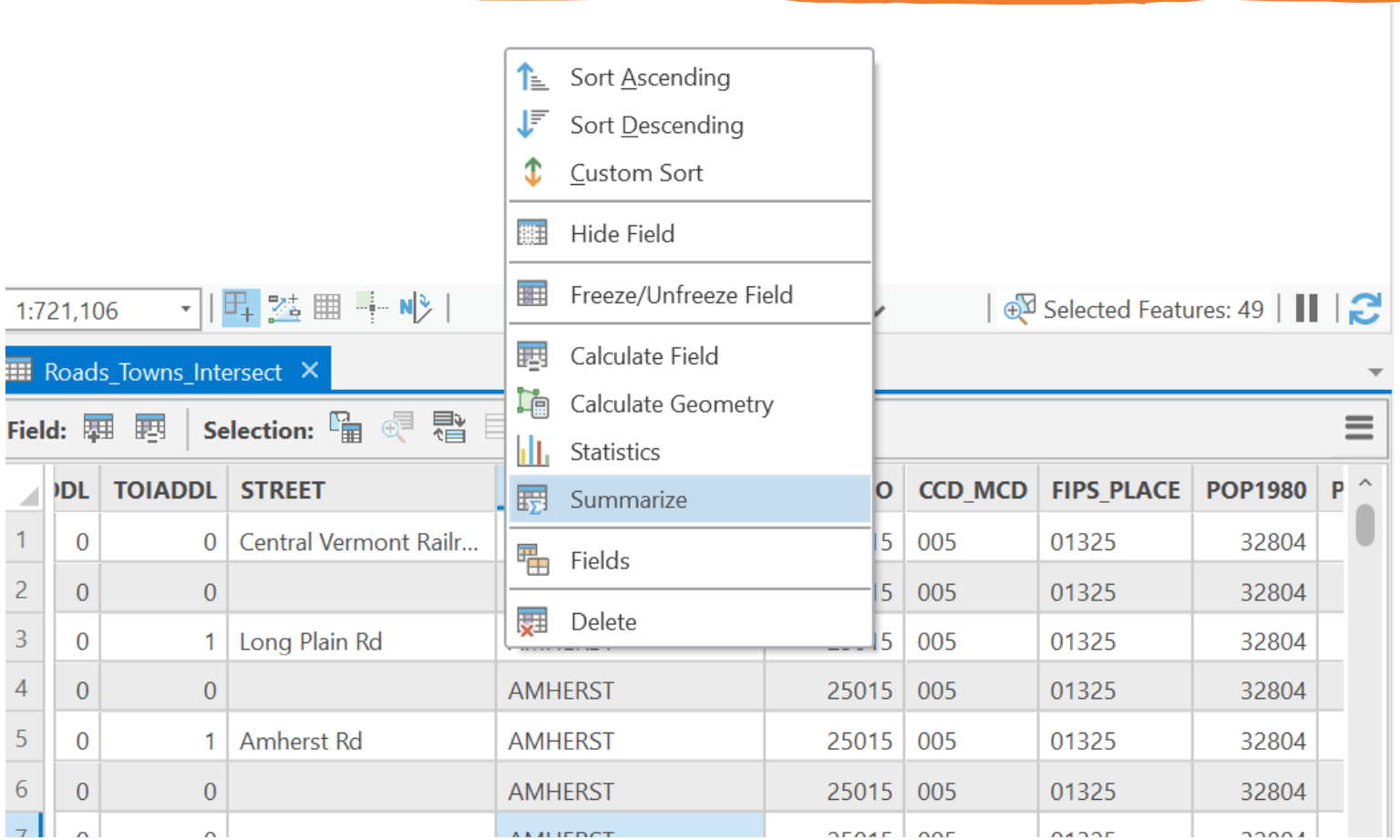
For example, you have two shapefiles:

1. Towns
2. Roads

More later, this isn't on the midterm.

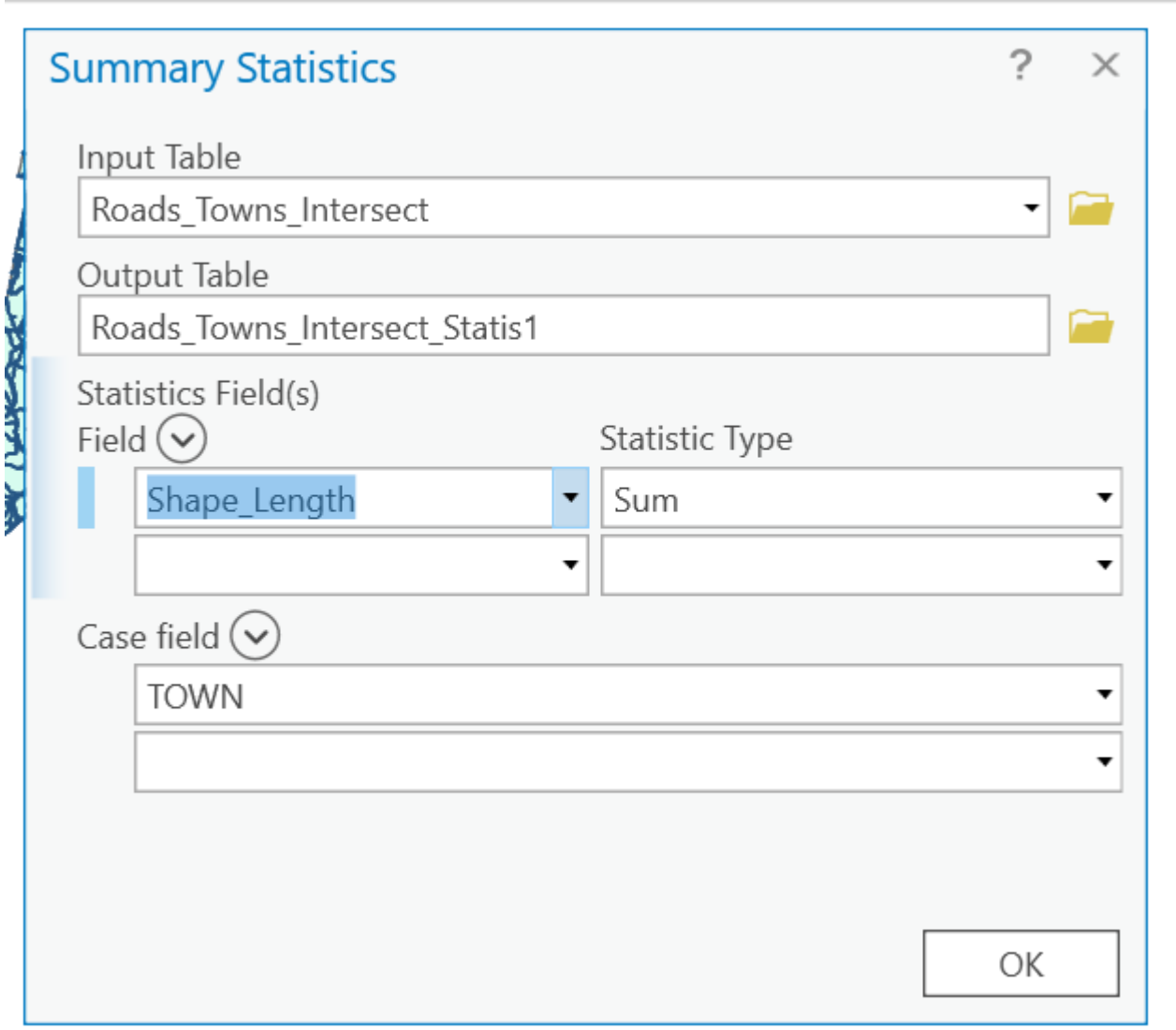


Why is the summarize tool awesome?



Aggregation is Awesome!

- Summarize will count features, aggregated by unique values
- It will also perform calculations like total length, area, etc!

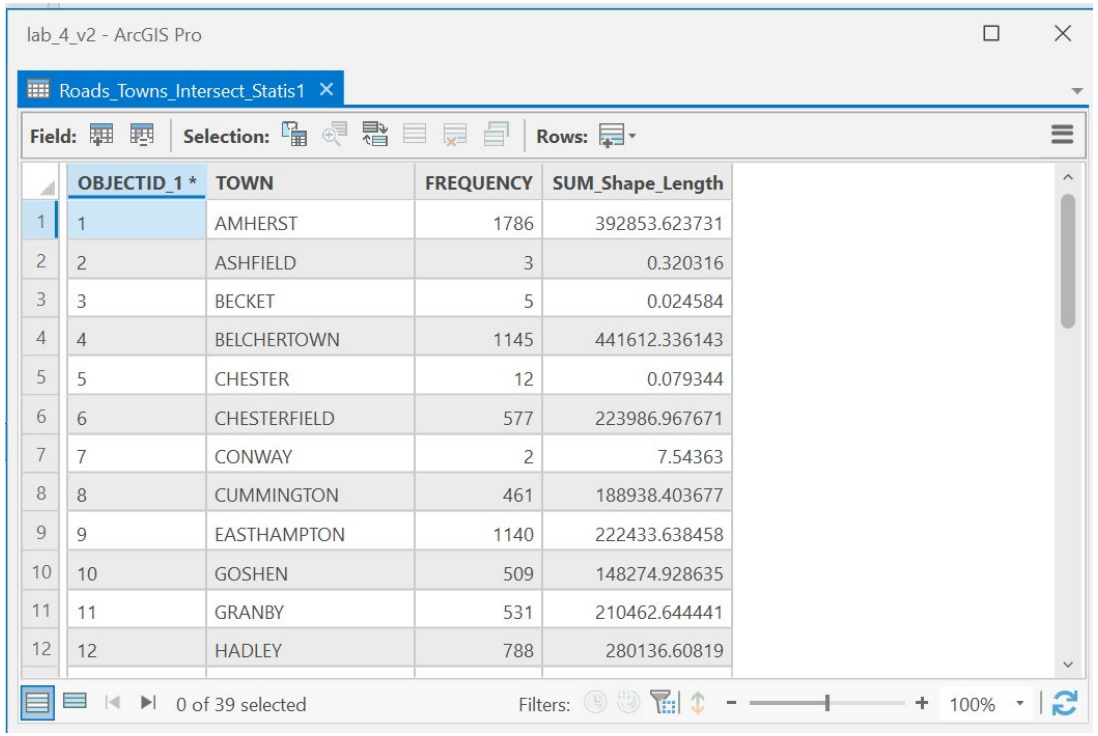


Aggregation

Summarize creates a table with:

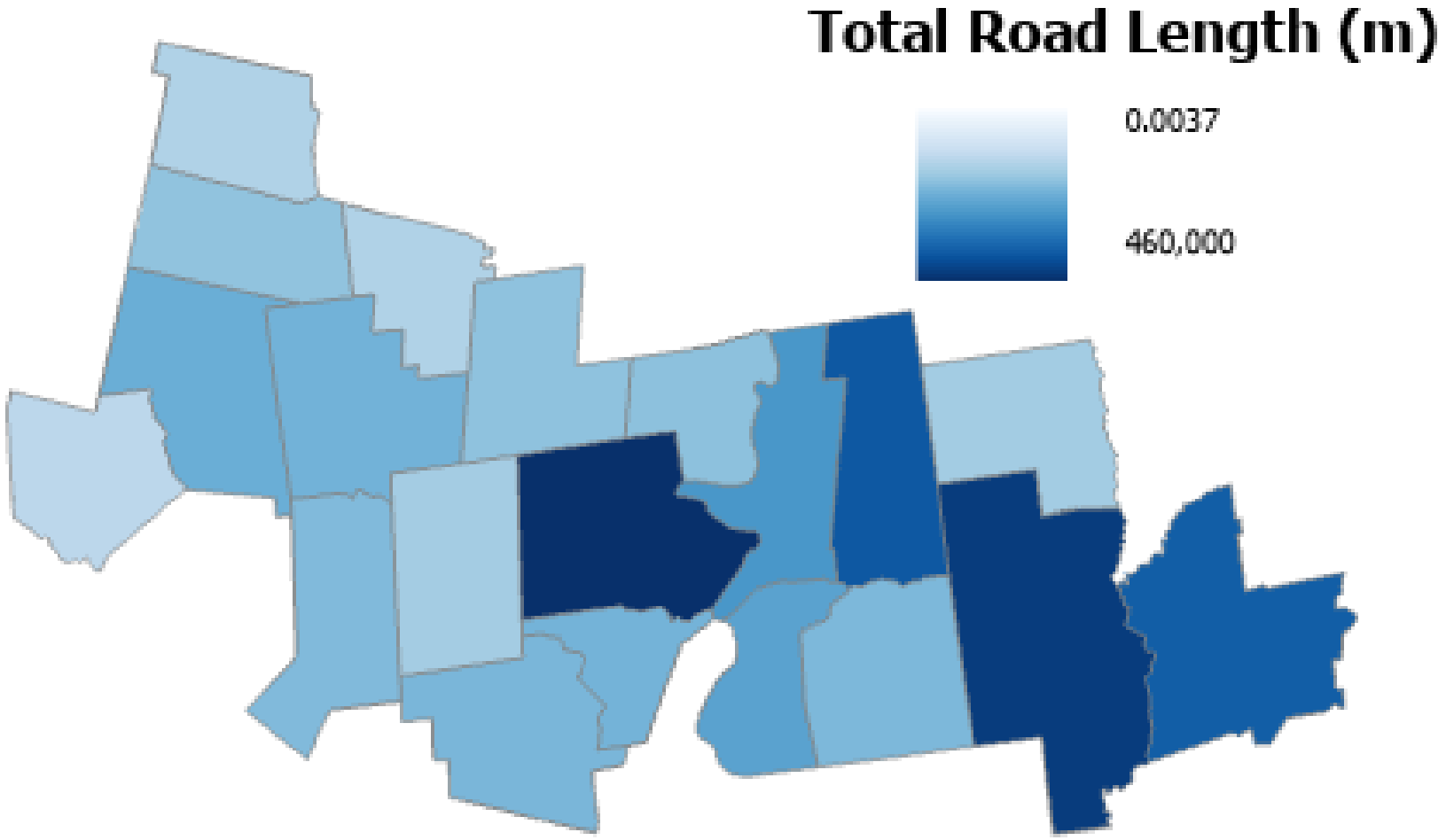
- **FREQUENCY:**
- Total number of major road segments in that town.
- **Sum_Shape_Length:**
- Sum of the lengths of all major roads in that town.

Now we can make maps using the new attributes!



	OBJECTID_1 *	TOWN	FREQUENCY	SUM_Shape_Length
1	1	AMHERST	1786	392853.623731
2	2	ASHFIELD	3	0.320316
3	3	BECKET	5	0.024584
4	4	BELCHERTOWN	1145	441612.336143
5	5	CHESTER	12	0.079344
6	6	CHESTERFIELD	577	223986.967671
7	7	CONWAY	2	7.54363
8	8	CUMMINGTON	461	188938.403677
9	9	EASTHAMPTON	1140	222433.638458
10	10	GOSHEN	509	148274.928635
11	11	GRANBY	531	210462.644441
12	12	HADLEY	788	280136.60819

What is the name for this kind of map?



Selections

It’s all about the options

Selection Options: Refining a Selection

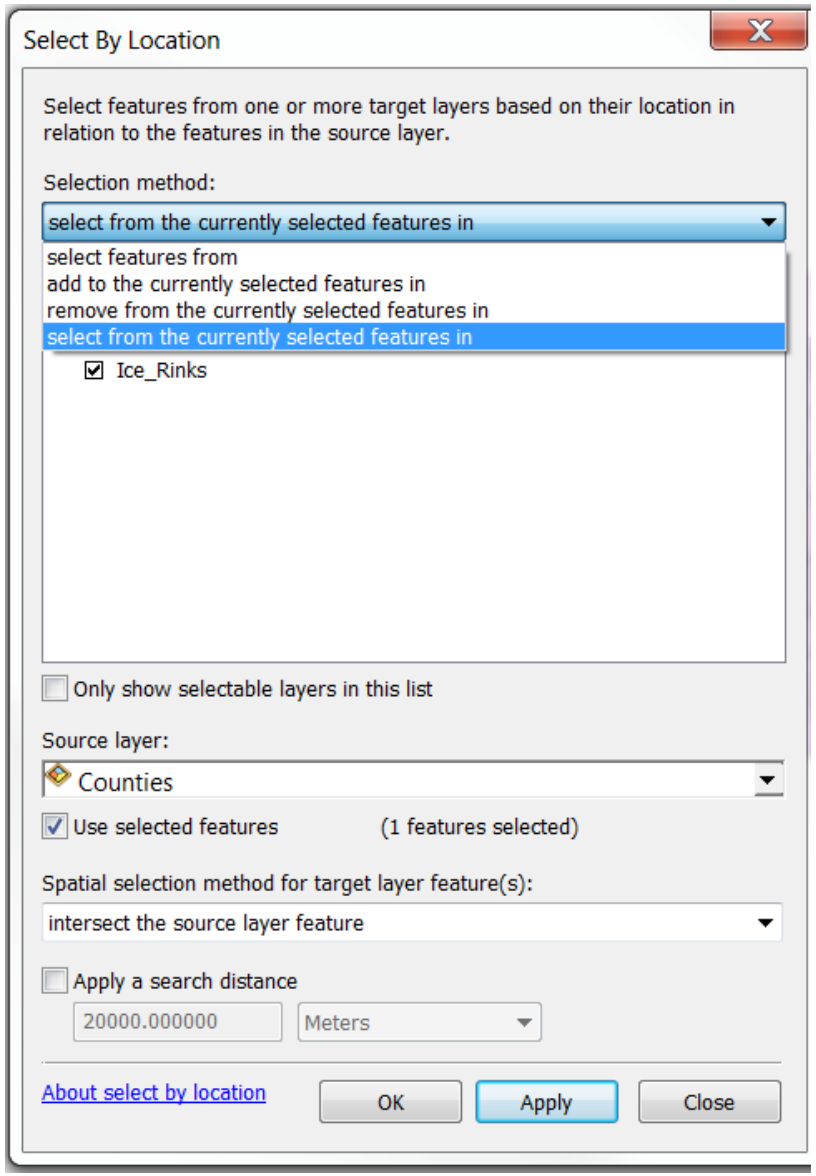
Select from currently selected features:

- Create a subset of the features you've already selected.
- Usually creates a smaller selection

Add to current selection

Remove from current selection

Make sure this box is checked to use the already selected features



Selection Options: Refining a Selection

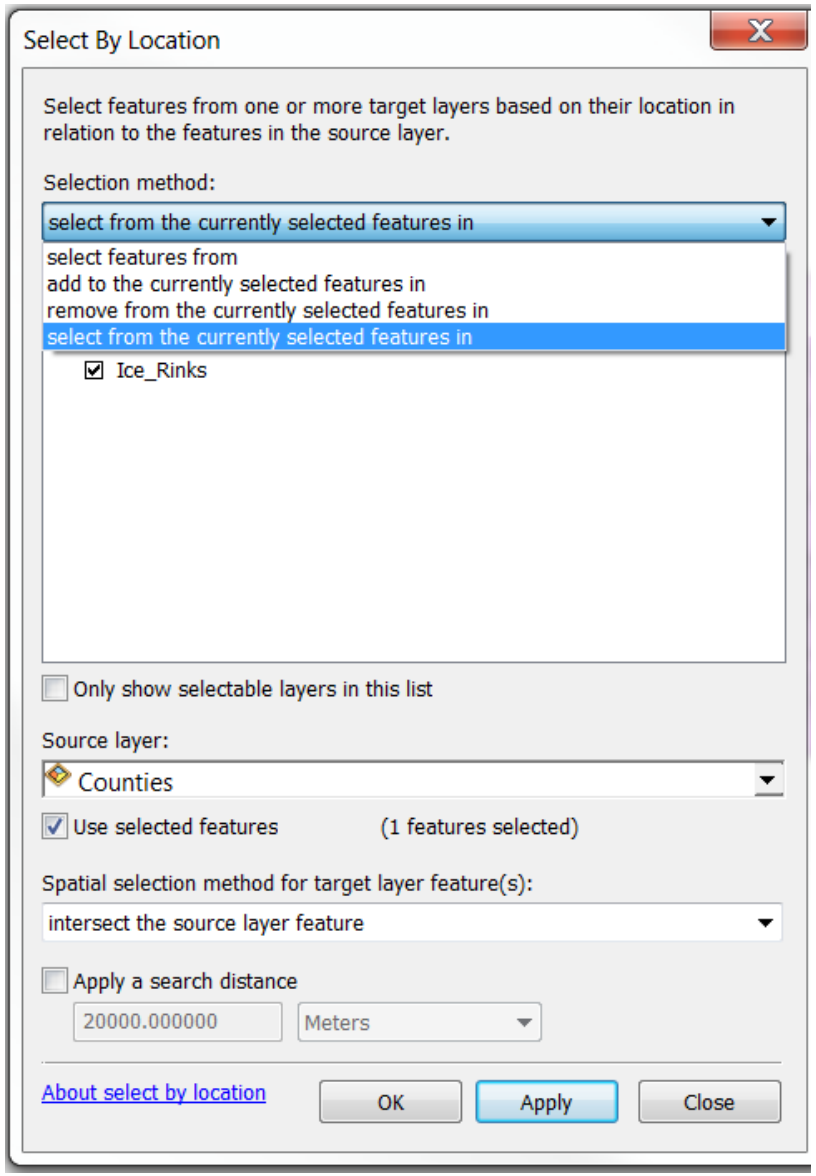
Select from currently selected features

Add to current selection

- Create a new selection criterion and add the new features to the current selection.
- May result in a bigger subset

Remove from current selection

Make sure this box is checked to use the already selected features



Selection Options: Refining a Selection

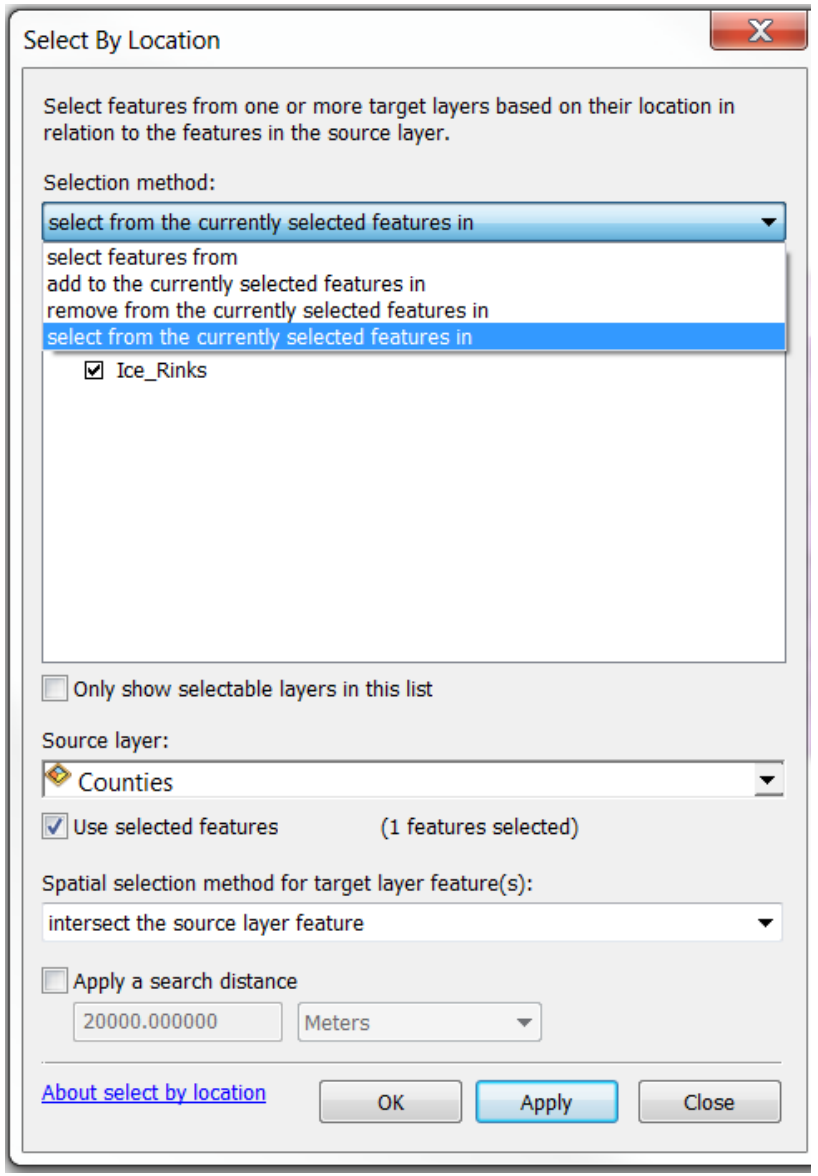
Select from currently selected features:

Add to current selection

Remove from current selection

- Create a new selection criterion and removes the newly selected features from the current selection.
- May result in a smaller subset.

Make sure this box is checked to use the already selected features



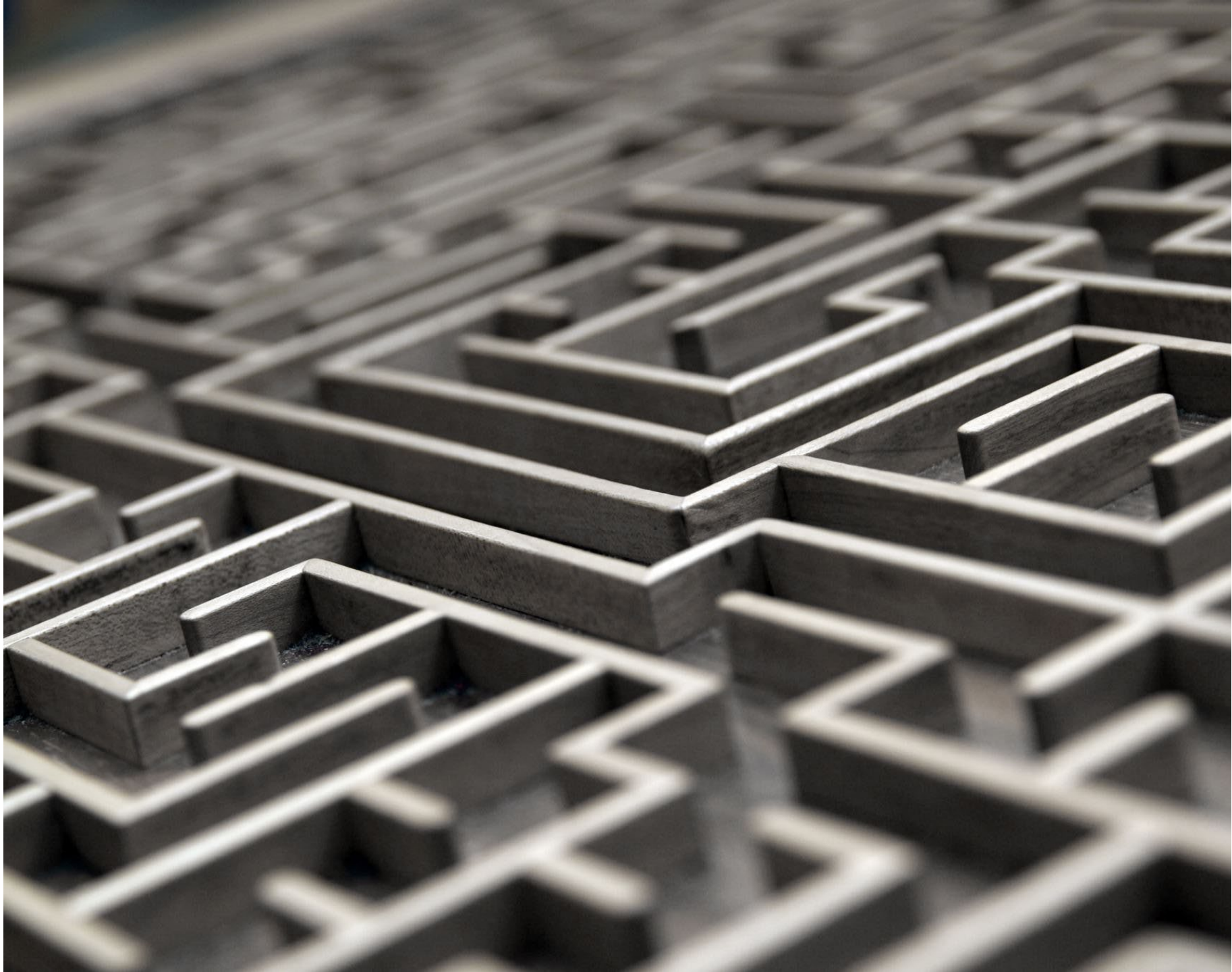
How could we use selection to find all public ice rinks in Middlesex county?

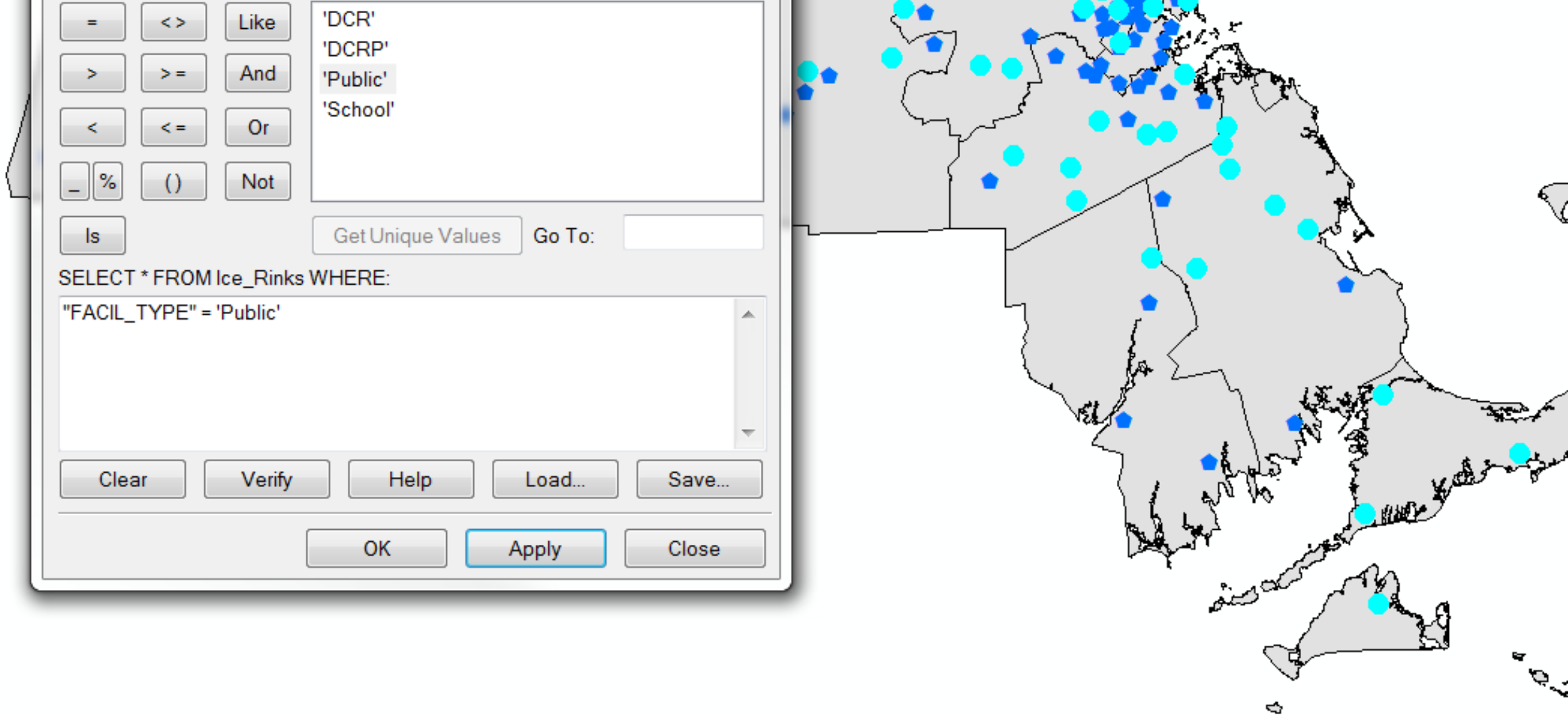
Why wolves?*

Silly AI!

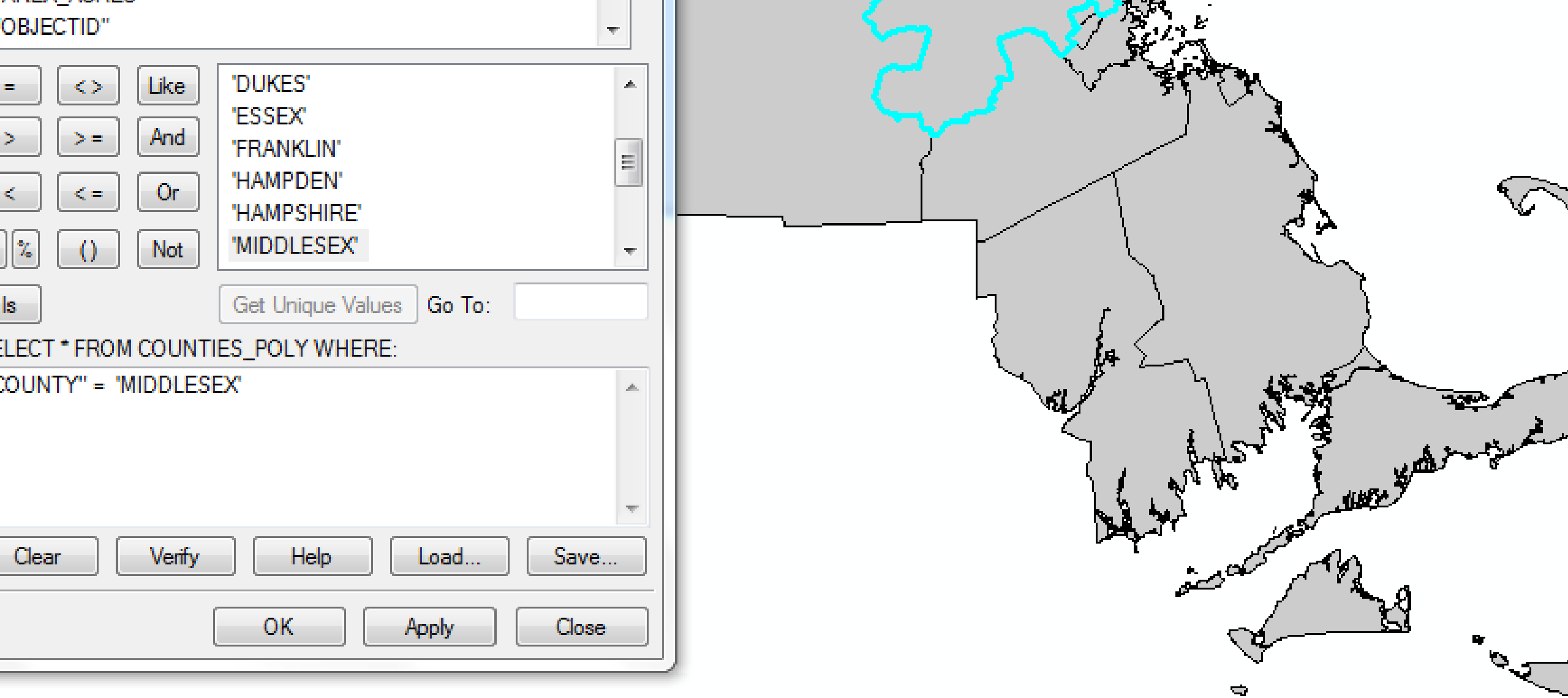
Selection by location and attribute

- We need to do 2 steps!
- Does the order matter?
- Do we use attribute or location selections?



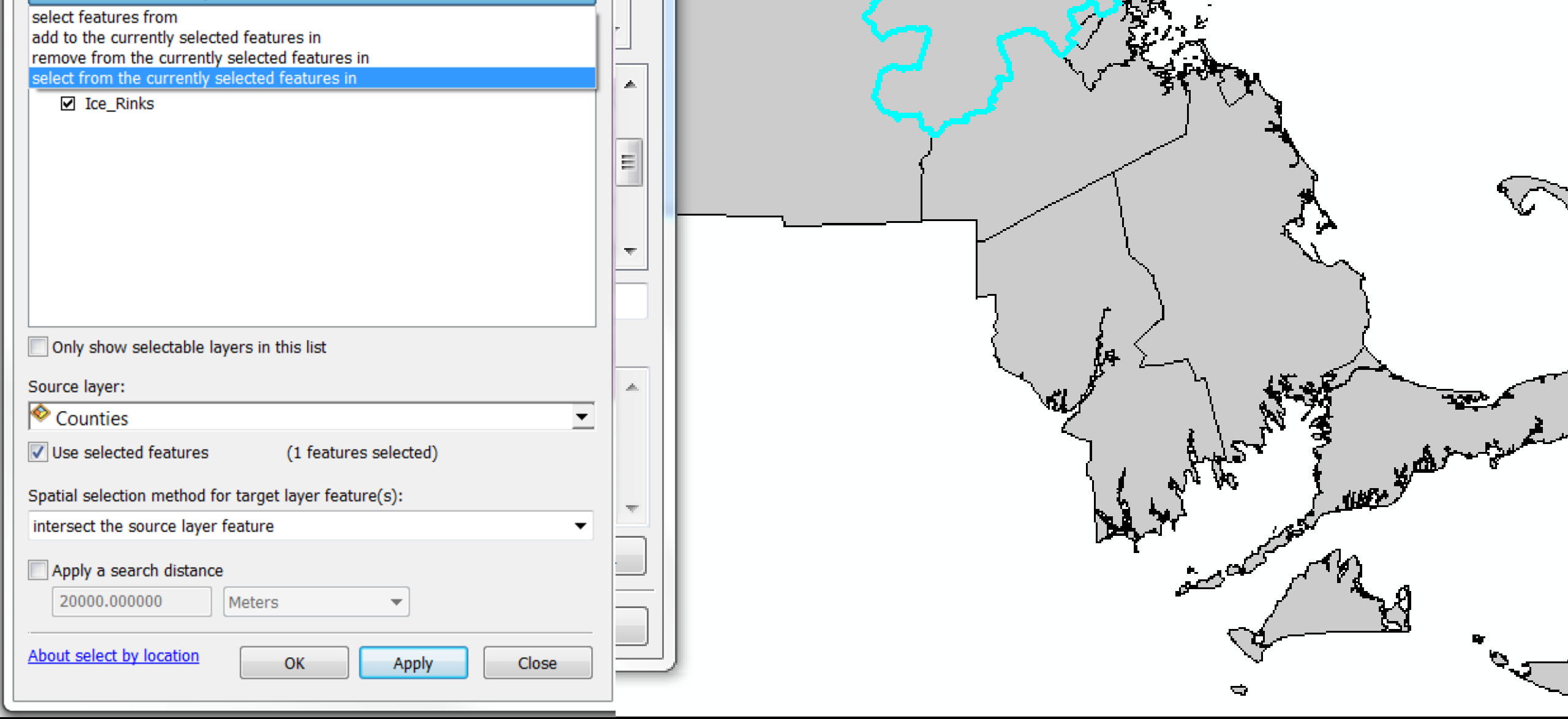


Select by attribute 1: SQL query for public rinks



Select by attribute 2: SQL query for Middlesex

* Note that this only works because the ice rinks have a county attribute



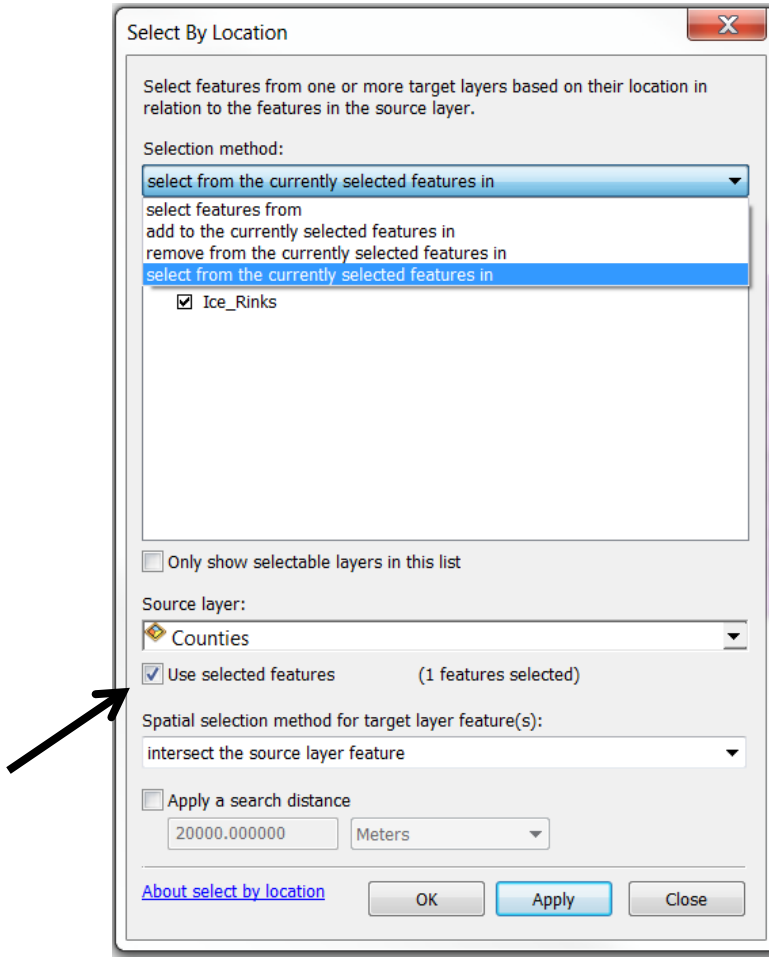
Select by Location: Middlesex County

Select by Location
Public ice rinks located in Middlesex County

Selection Options:

- New selection
- Add to selection
- Remove from selection
- Select from selection

Make sure this box is checked to use the selected features



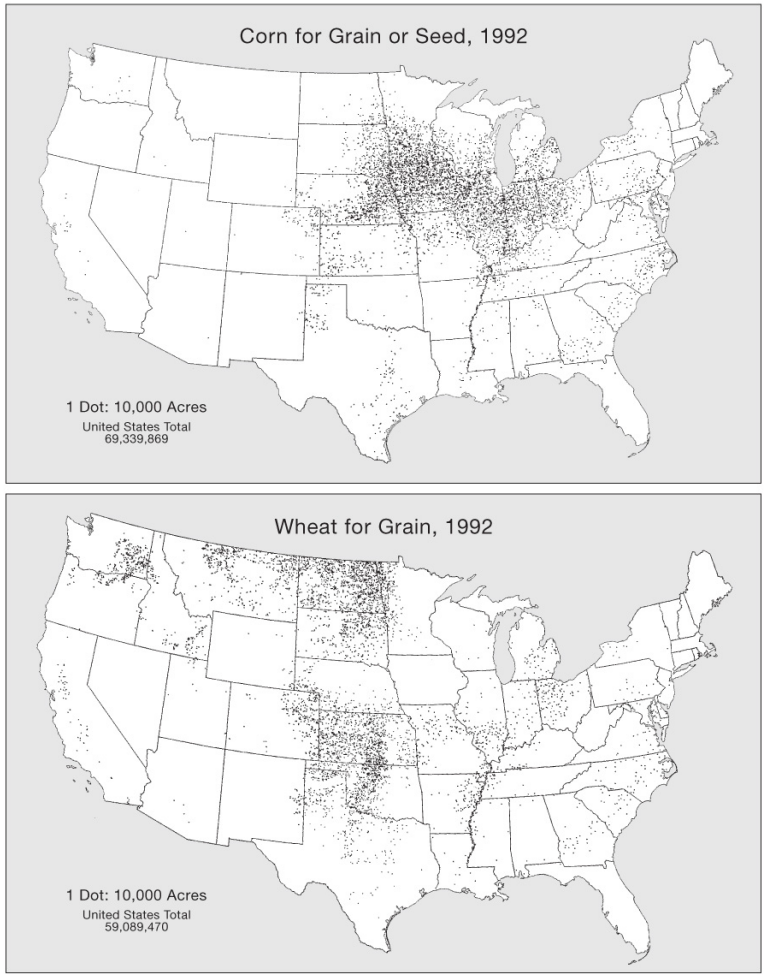
Cartography

Basic Best Practices

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



Let’s talk about dot density…

- Dot density can be difficult to understand and fussy to successfully implement, as you’ve seen in the grizzly bears lab!
- Dots do not represent actual points.
- A single dot does not usually represent a single individual.
  - That’s why we need the legend to tell us what a dot represents.

Let’s talk about dot density…

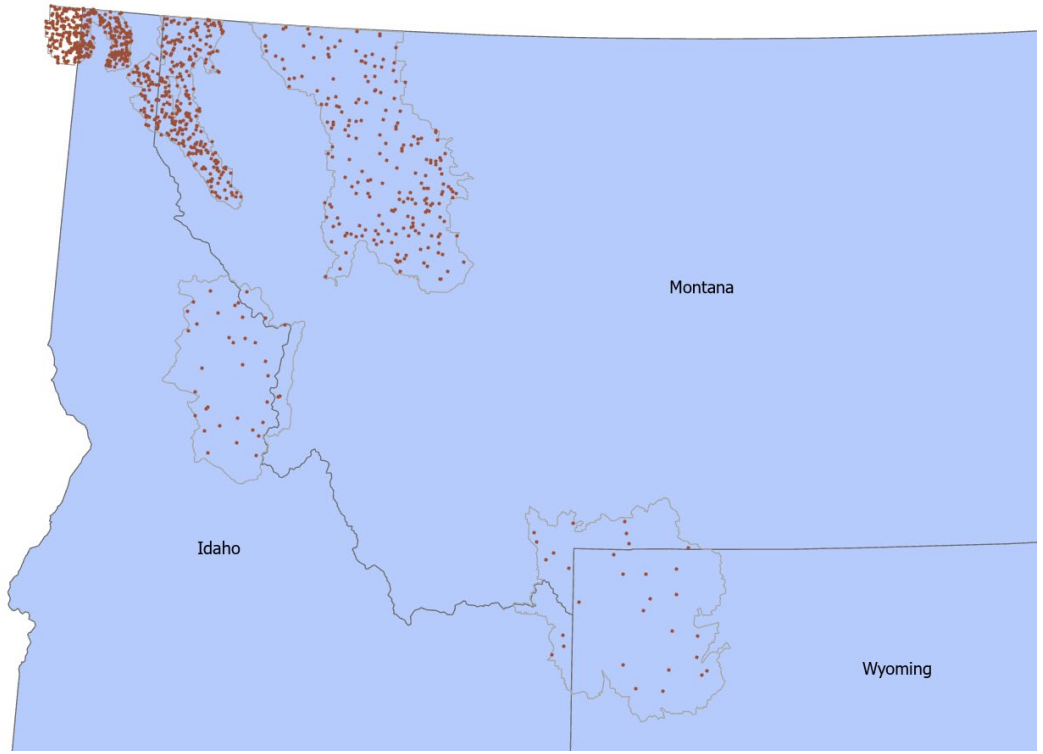
- The ‘density’ of dots = approximate number of dots in a unit of map area.
- The density is scaled by some numeric attribute.
  - Two polygons of different sizes with the same value for the attribute will have the same density but will have different numbers of dots.
- Dots are randomly placed within polygons. Arc uses a random number generator to place them. If you re-do the dot density symbology, dots will be in different locations.

Dots and Bears

- The lab asks for dot density scaled by bears per km2.
  - Not the bear population!
  - There is a terminology clash: ‘density’
    - Dot density: the symbology used on the map. The density of dots is proportional to the bear density.
    - Bear density: the number of bears per km2. This is a property of polygons within the recovery zones layer.
    - Is this confusing?!!? Yes!

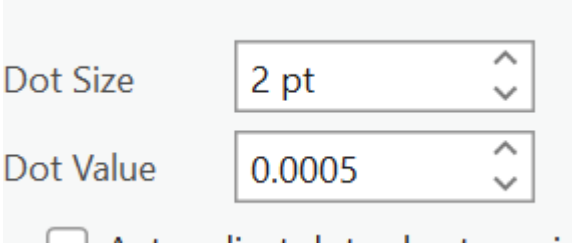
Density and Bears

	gnzone.ZONE_NAME	Area_km2	BEARS	Pop. Density
1	Selkirk Mountains	2889.29	515	0.178244
2	Cabinet-Yaak	6682.66	1125	0.168346
3	Northern Continental...	23130.4	2450	0.105921
4	Bitterroot Mountains	13809.8	250	0.018103
5	Greater Yellowstone	24026.7	360	0.014983



Notice the extreme differences in dot density.

- Selkirk density is more than 10 times the density of Yellowstone.
- Extreme difference in value makes dot density difficult to resolve.
- Small pop density values mean you have to use a small dot value.

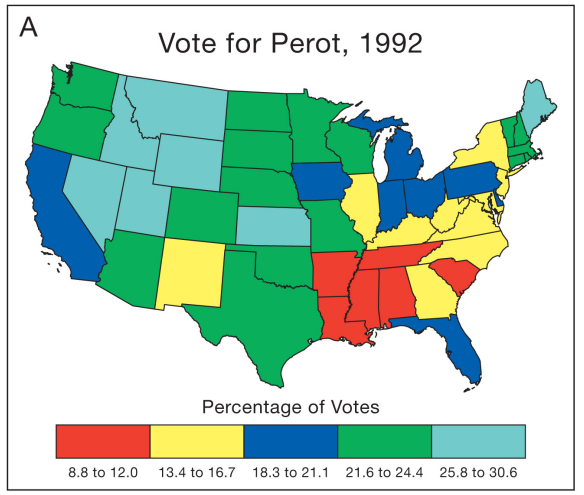


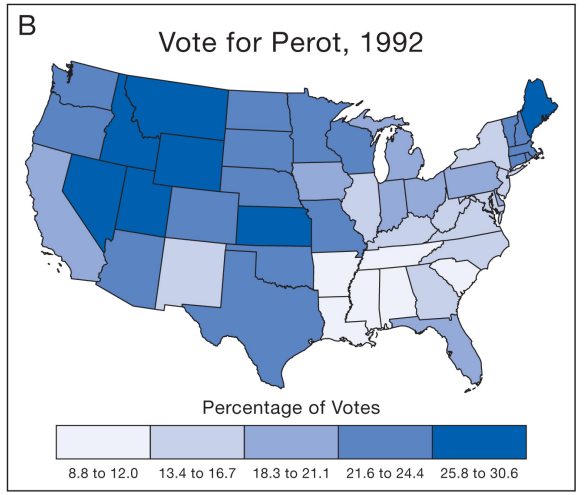
Thematic Cartography: Choropleths

Use color to emphasize theme: population density, family income, daily temperature maximums, etc.

Choropleth: color is proportional to a numerical value

- Value (from HSV) is proportional to Perot support.
- Color becomes **thematically informative**

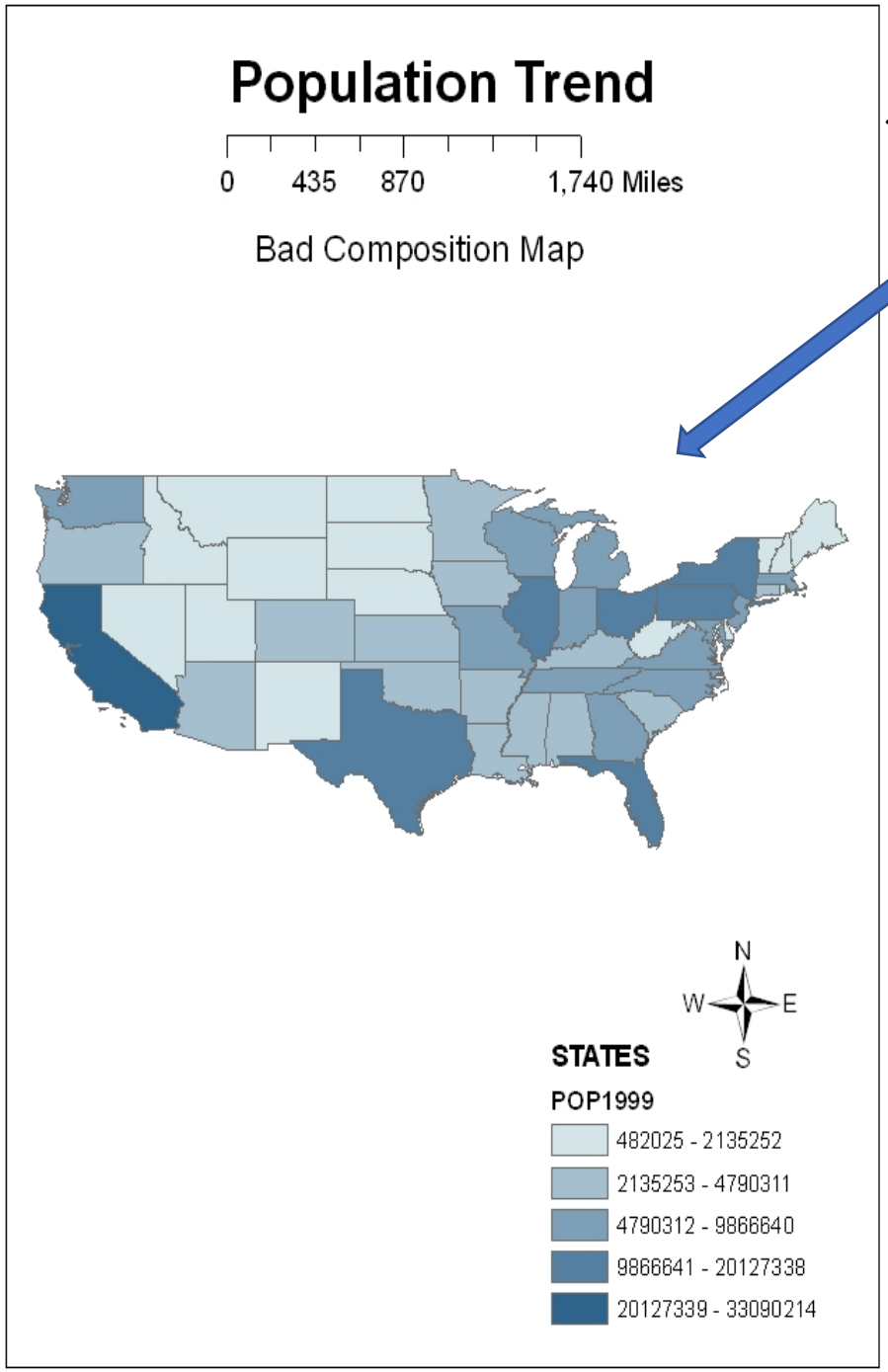




Composition refers to position of elements on a map

Composition: Avoid excess white space

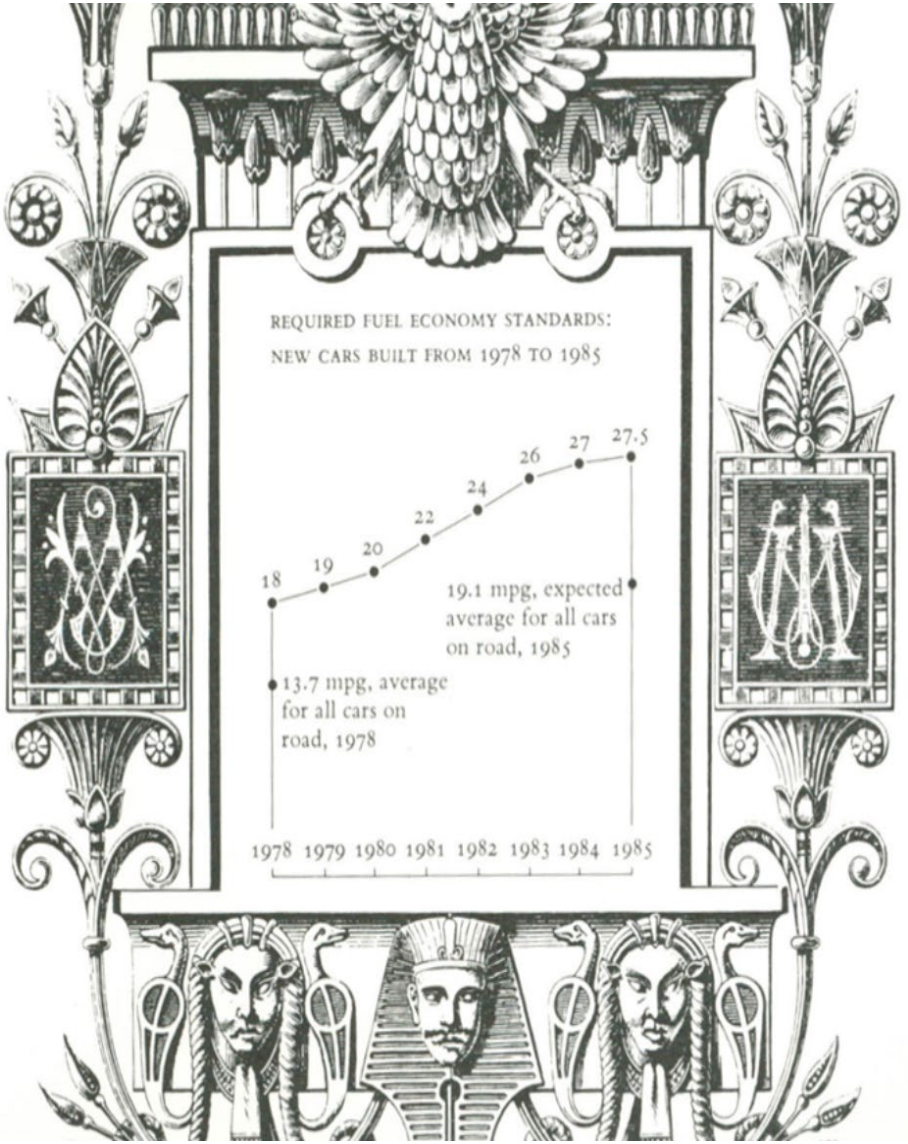
Like this…



…or this

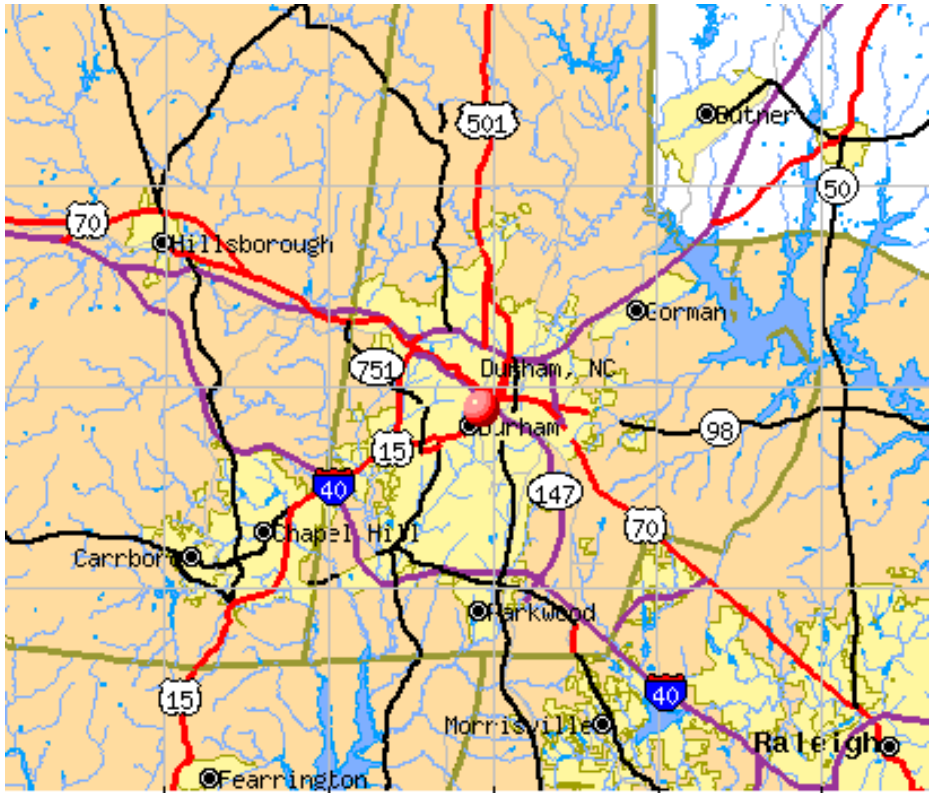
Composition and Clarity: Reduce Distractions

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



Map Design: Clarity and intuition

- Use intuitive colors schemes and appropriate cropping



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

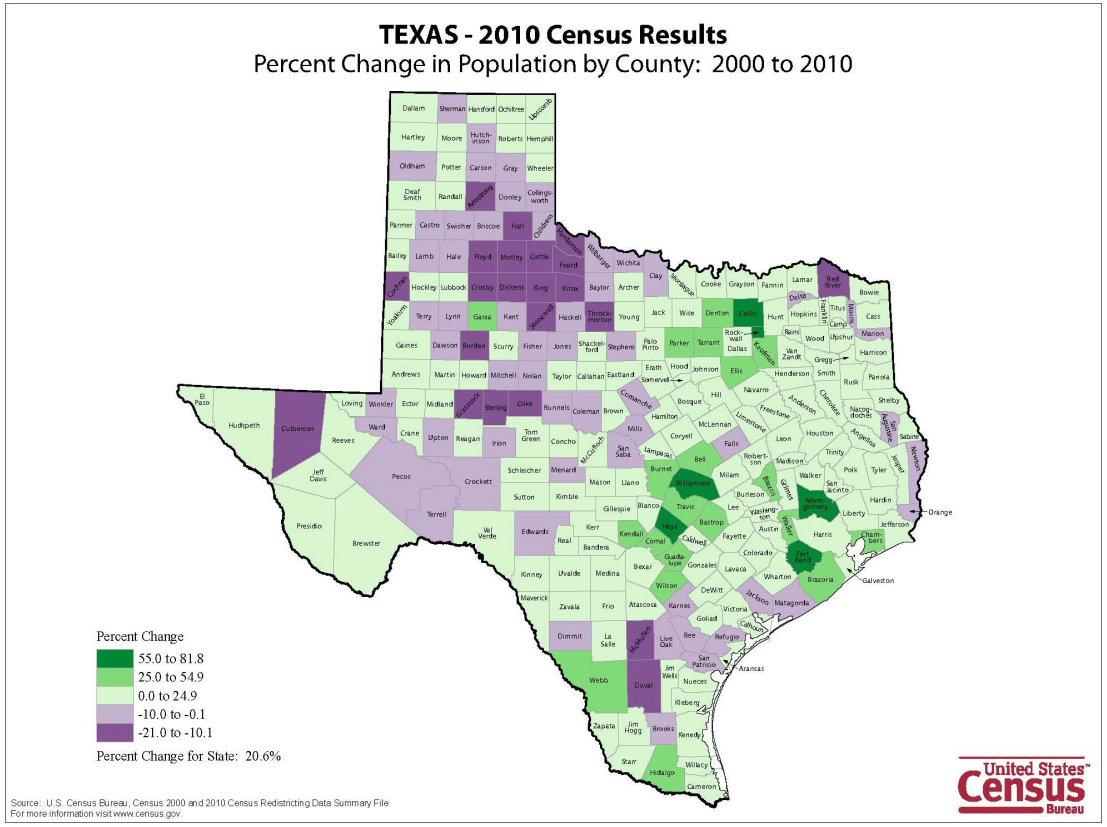
When making a map, you have:

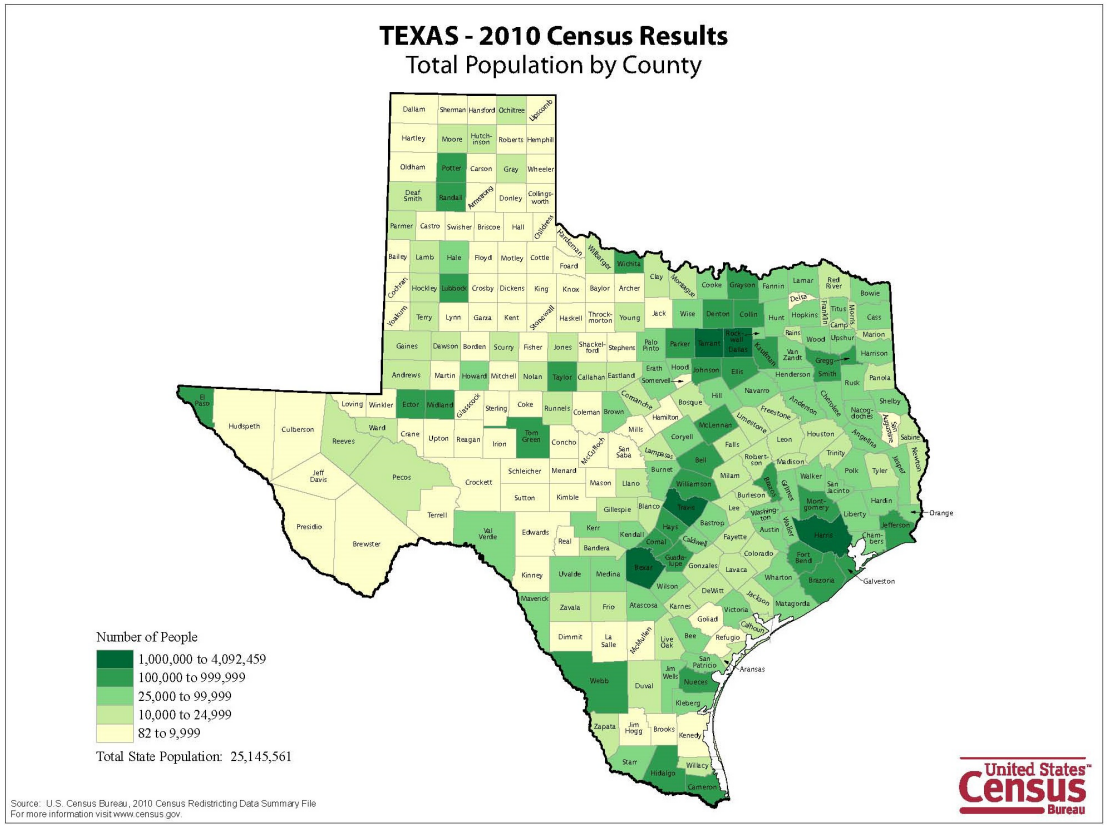
- Control over content
- Control over area
- Control over emphasis
- Topology

Control over Content

Control over Area

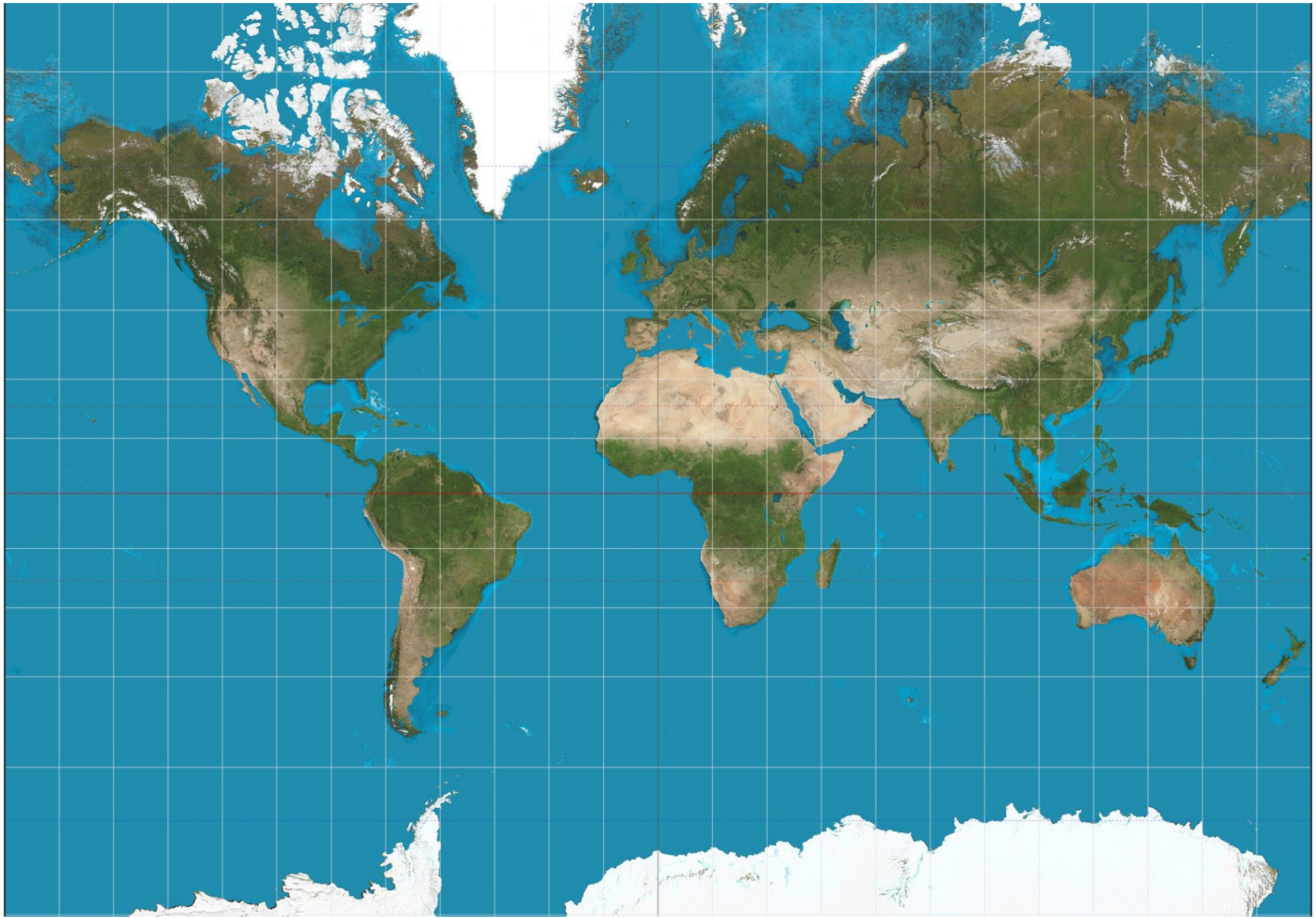
Control over Emphasis



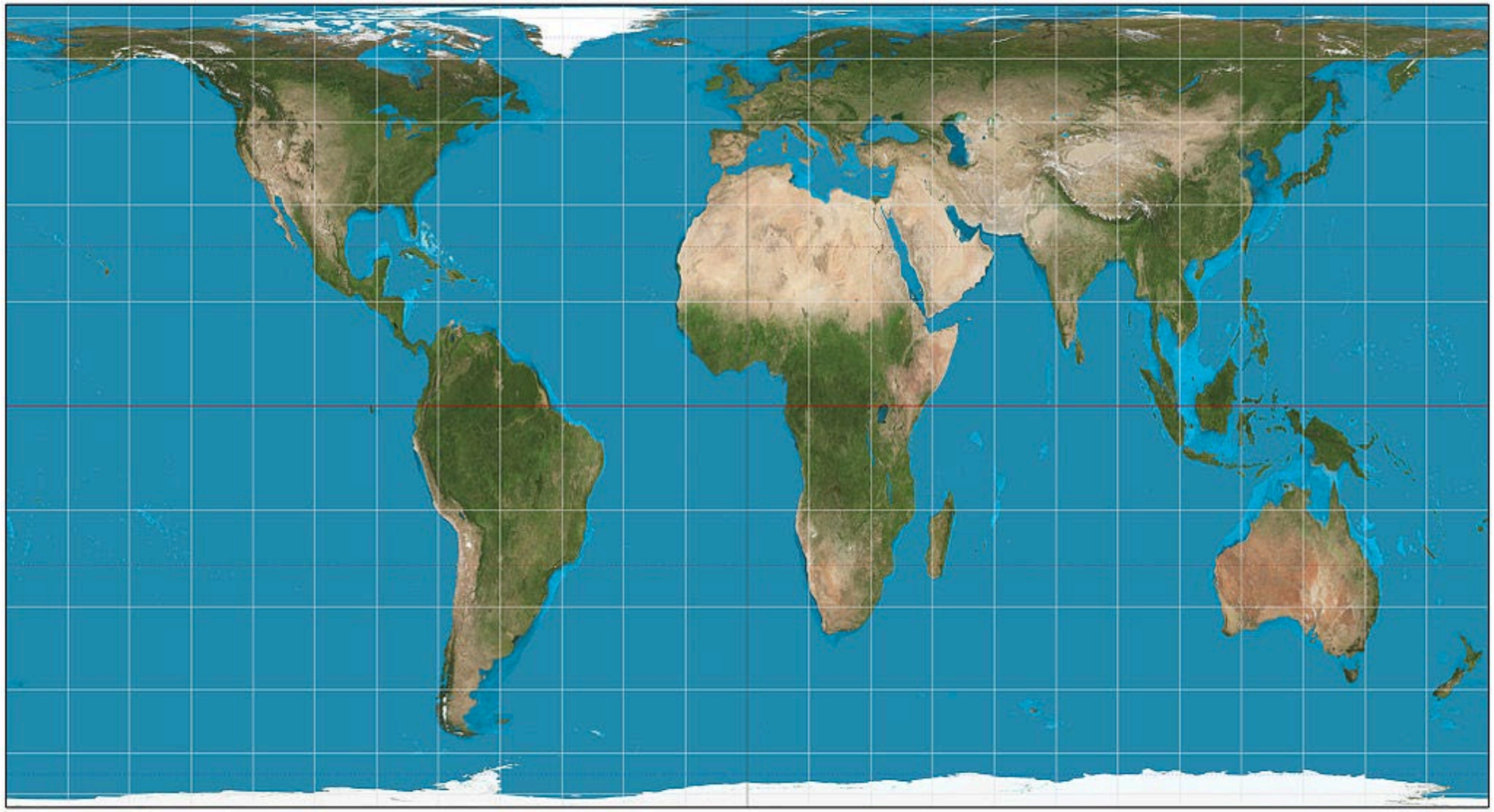


Analog of the real world (Topology)

Mercator Projection – reality?



Equal Area Mercator Projection – a different form of reality?



Final Projects

Methods Outline and Examples

Outline and Analytical Proposal

Use our feedback to refine your project idea.

- You'll need to narrow it down to a small number of *specific* questions or goals.

Find your data sources!

Intersections

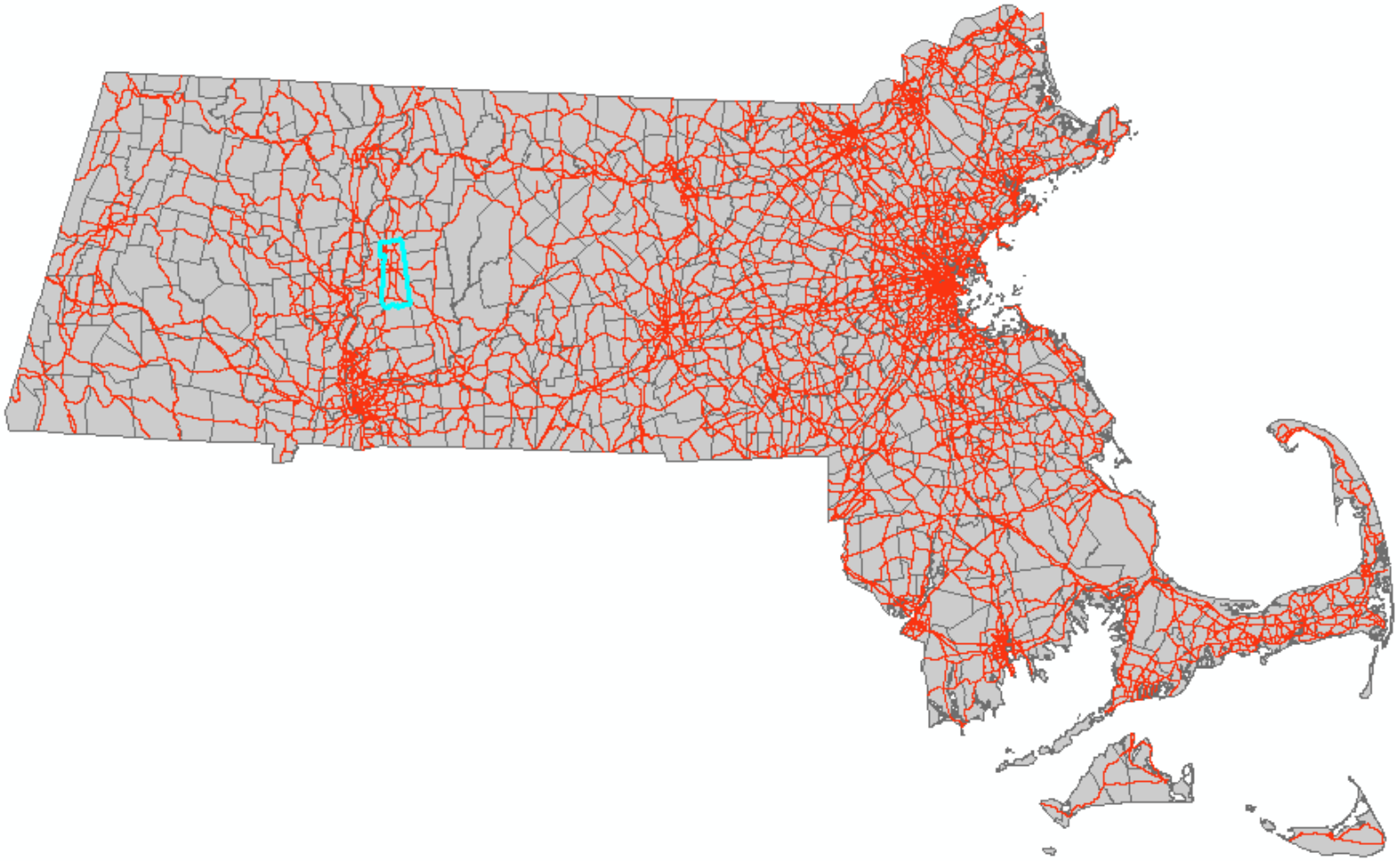
Intersections in further detail

Intersections aren’t on the midterm, so I glossed over them in the previous lecture.

The next few slides contain an optional, more detailed explanation of the motivation and an example.

What is the total length of roads in Amherst?

What steps could you use to find out?

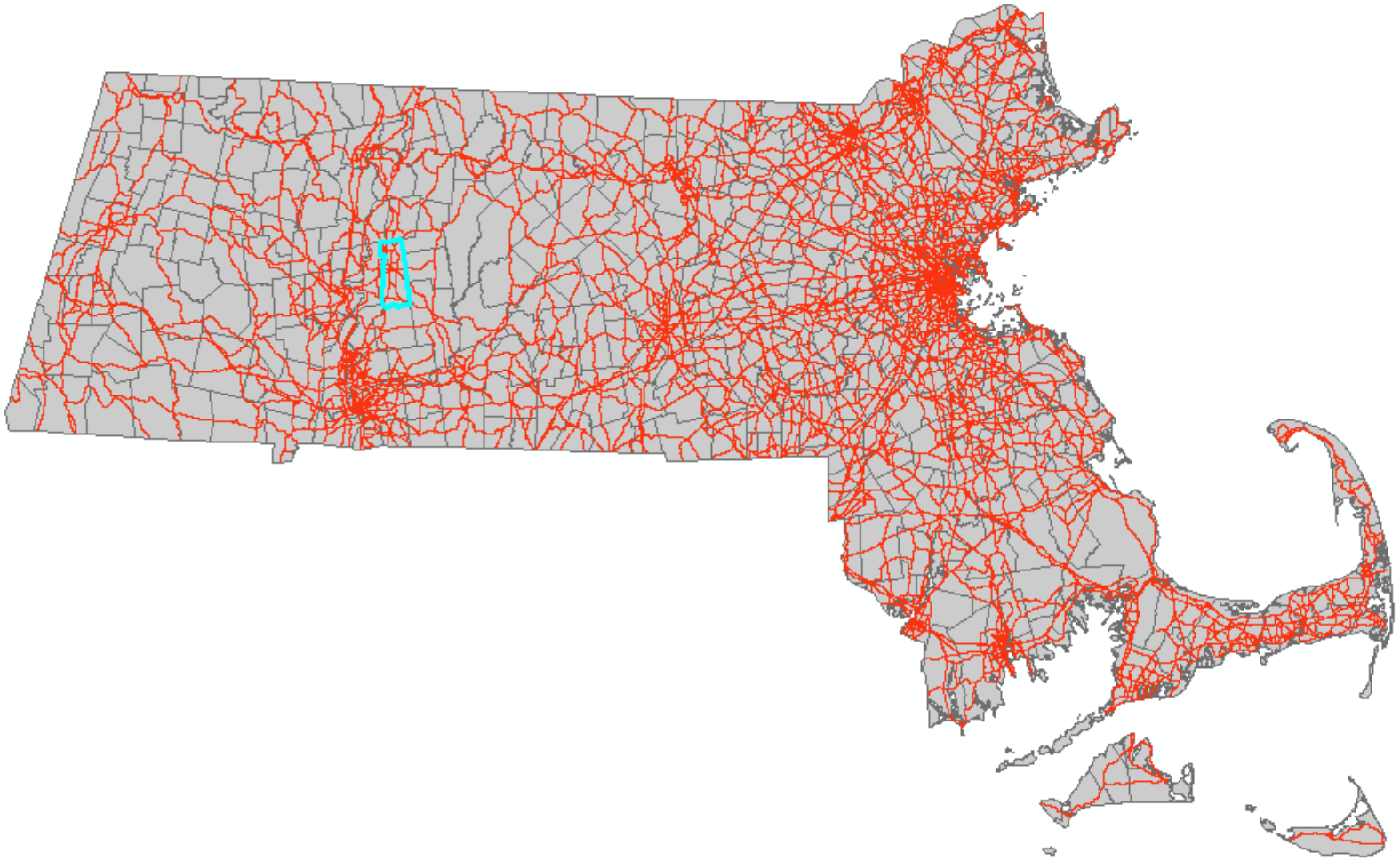


What is the total length of roads in Amherst?

What steps could you use to find out?

You have two layers:

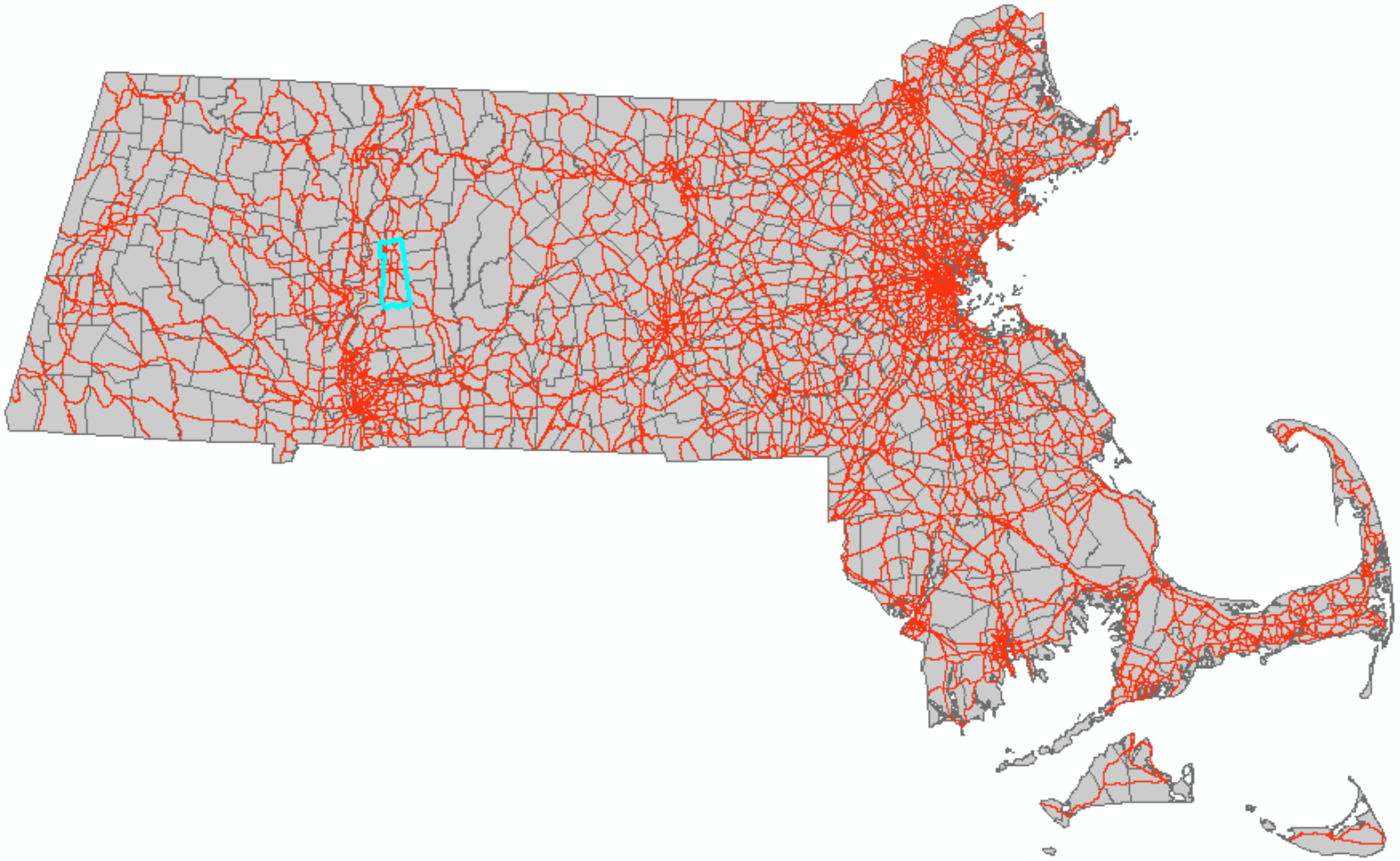
1. Towns
2. Roads



What is the total length of roads in Amherst?

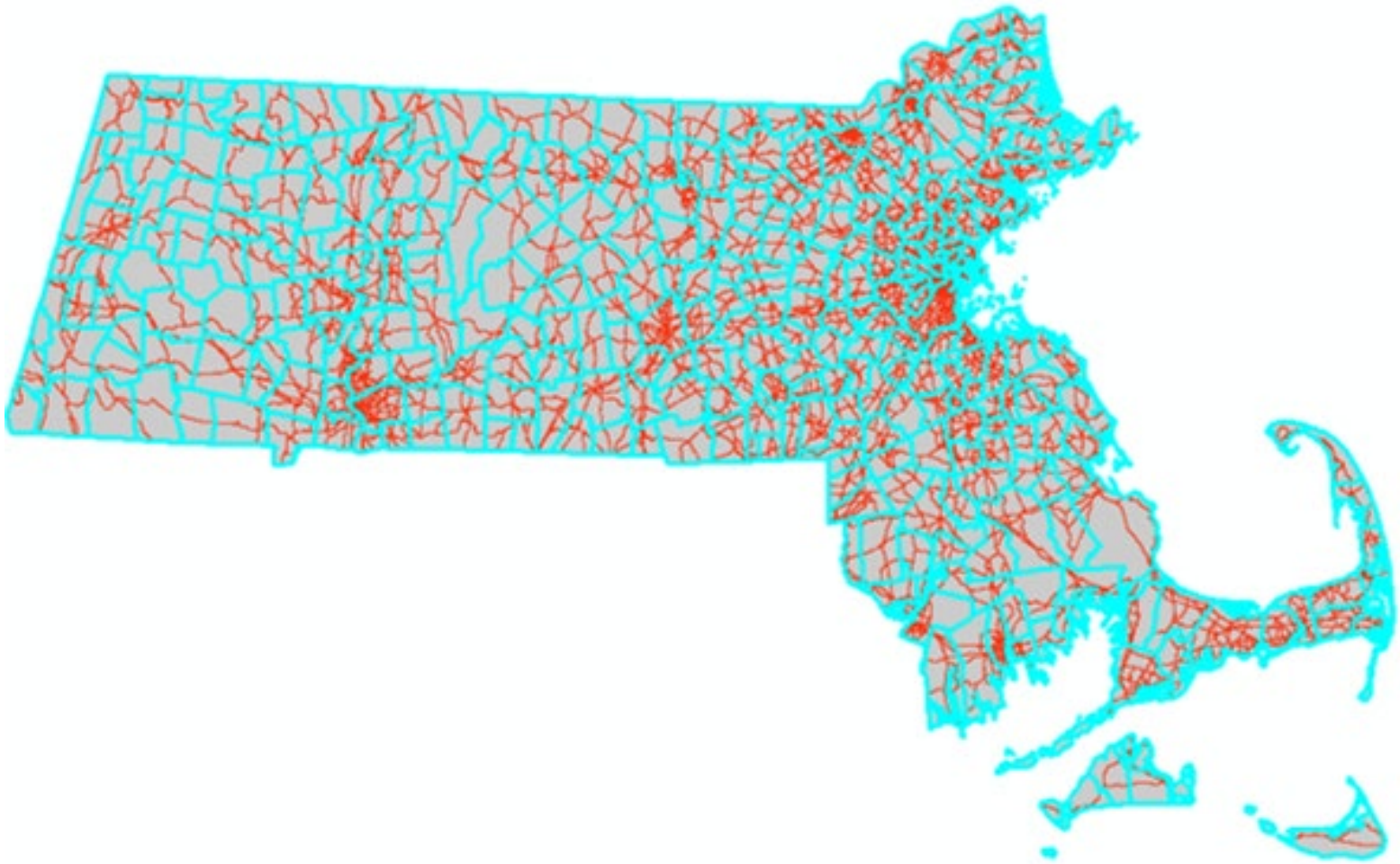
A possible workflow:

1. Select the Amherst polygon from the towns layer.
2. Clip roads by Amherst.
3. Calculate lengths of clipped road segments.
4. Calculate length sum with he ‘Statistics’ tool.



What is the total length of roads in each town in MA?

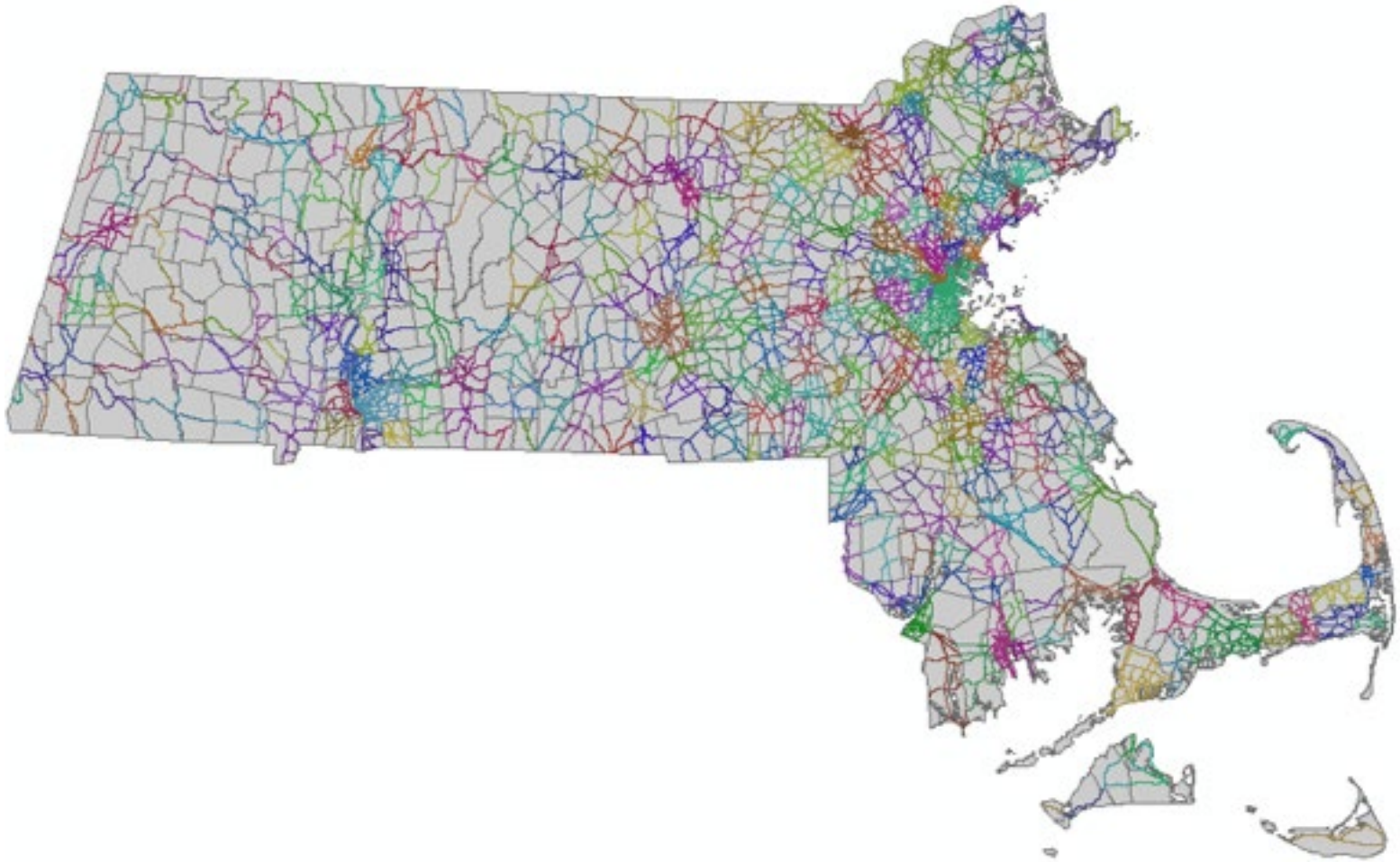
- More complicated question.
- How could you find out?



Road segments within towns

You have two shapefiles:

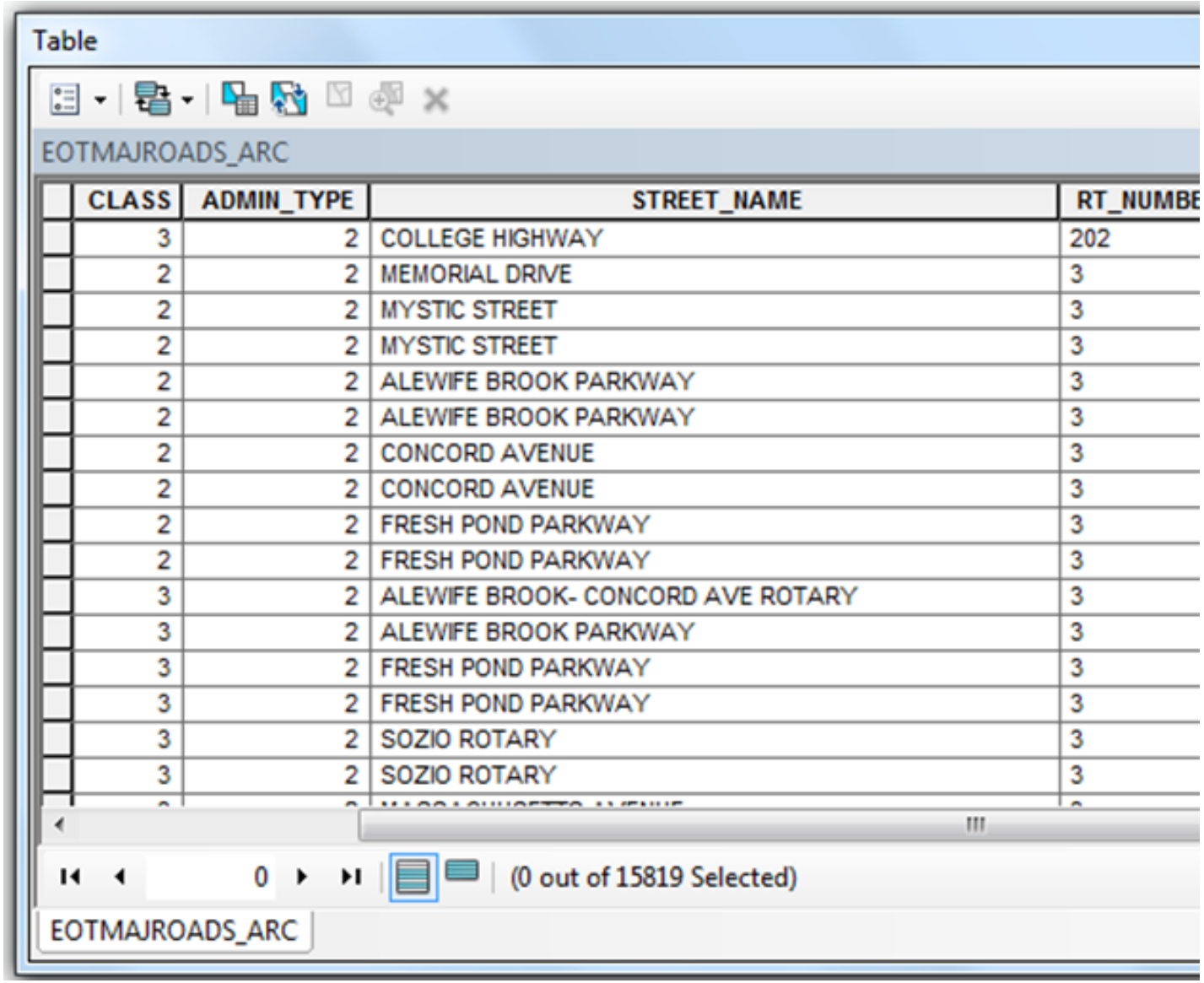
1. Towns
2. Roads



Major Roads in MA

Attribute table **before** intersection

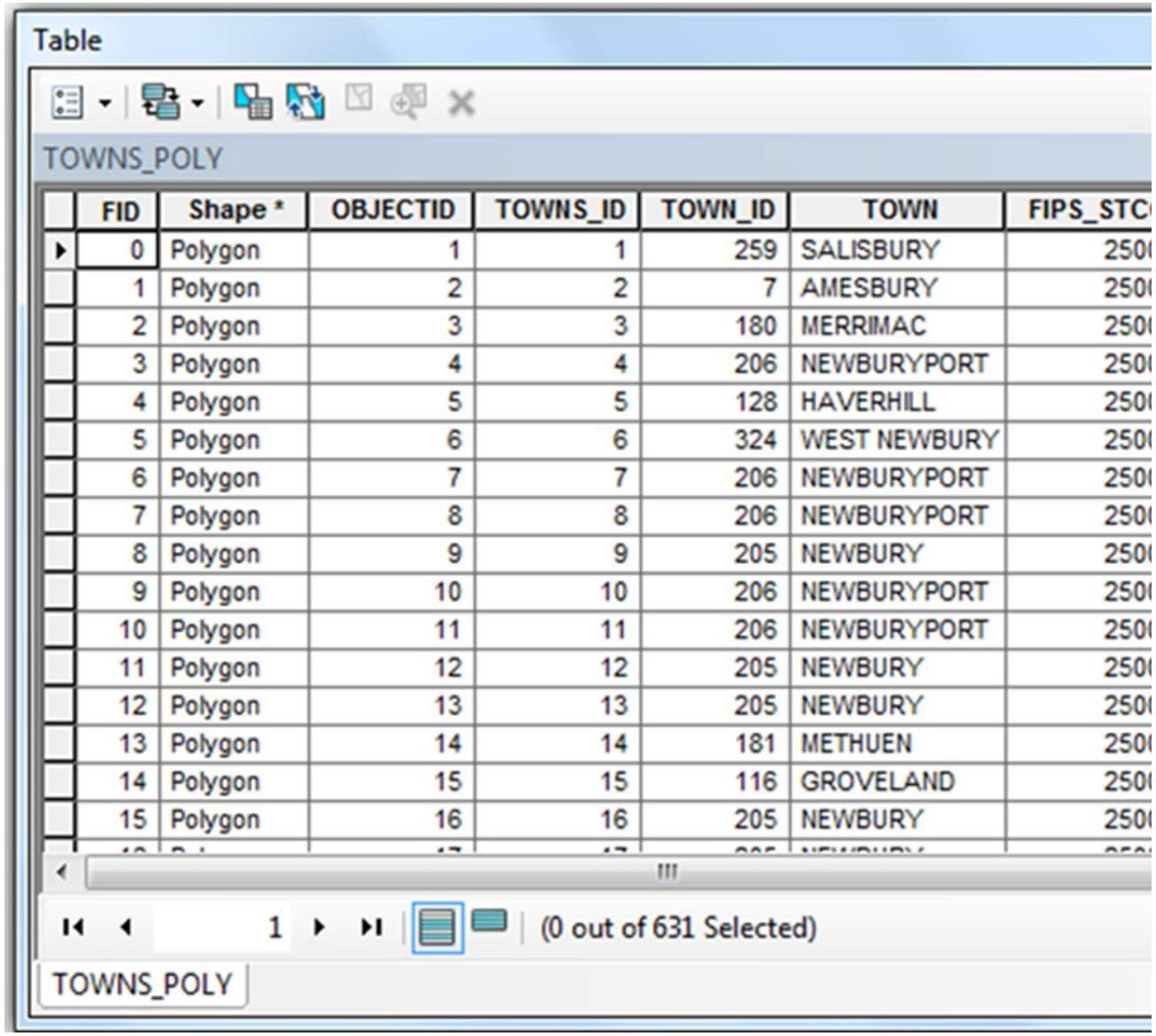
- There are 15819 major roads



Towns in MA

Attribute table **before** intersection

- There are 631 towns.

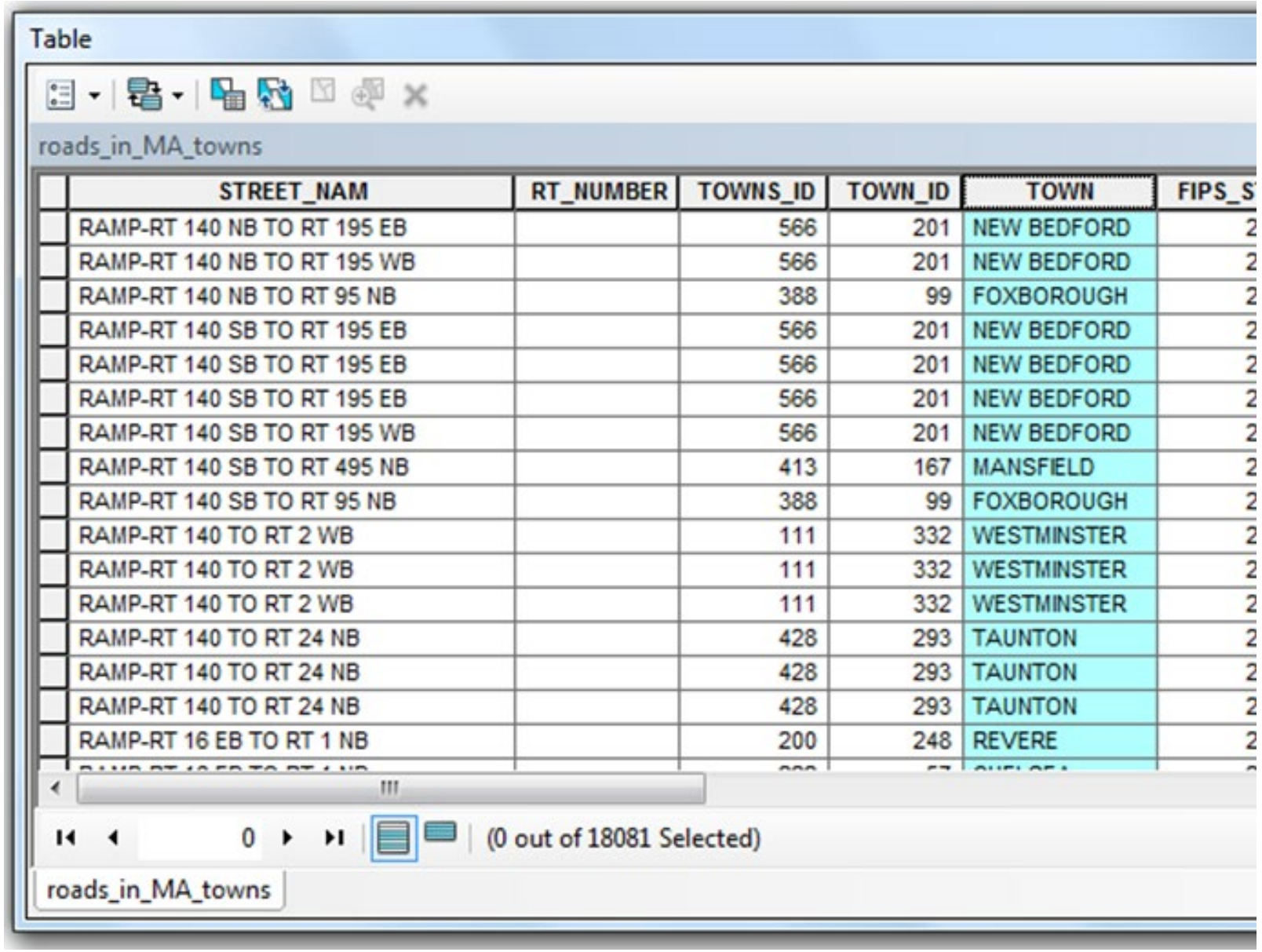


FID	Shape *	OBJECTID	TOWNS_ID	TOWN_ID	TOWN	FIPS_STC
0	Polygon	1	1	259	SALISBURY	250
1	Polygon	2	2	7	AMESBURY	250
2	Polygon	3	3	180	MERRIMAC	250
3	Polygon	4	4	206	NEWBURYPORT	250
4	Polygon	5	5	128	HAVERHILL	250
5	Polygon	6	6	324	WEST NEWBURY	250
6	Polygon	7	7	206	NEWBURYPORT	250
7	Polygon	8	8	206	NEWBURYPORT	250
8	Polygon	9	9	205	NEWBURY	250
9	Polygon	10	10	206	NEWBURYPORT	250
10	Polygon	11	11	206	NEWBURYPORT	250
11	Polygon	12	12	205	NEWBURY	250
12	Polygon	13	13	205	NEWBURY	250
13	Polygon	14	14	181	METHUEN	250
14	Polygon	15	15	116	GROVELAND	250
15	Polygon	16	16	205	NEWBURY	250

Road Segments in Towns

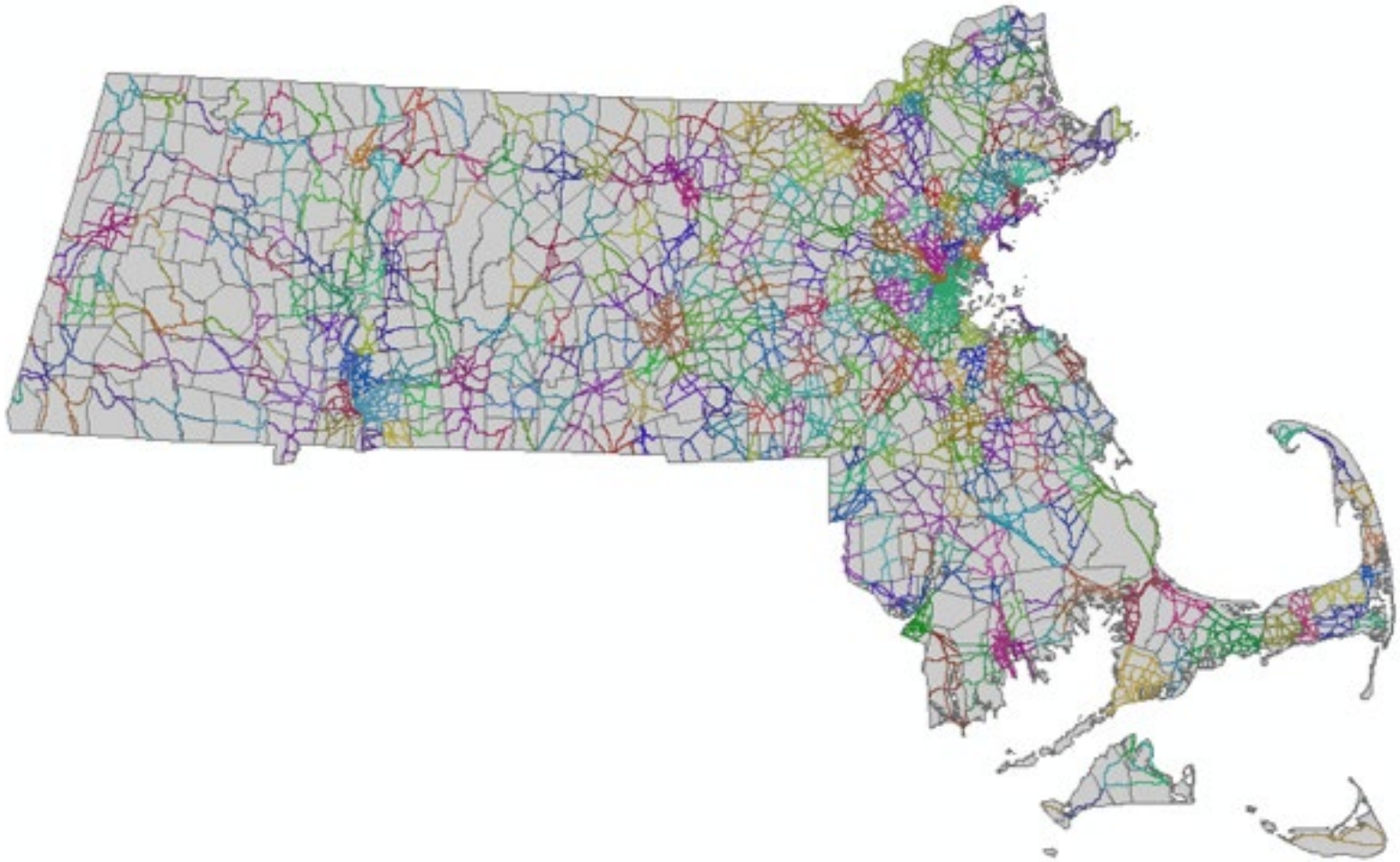
Attribute table **after** intersection

- There are 18081 features.
- What does a feature represent?



Road segments within towns

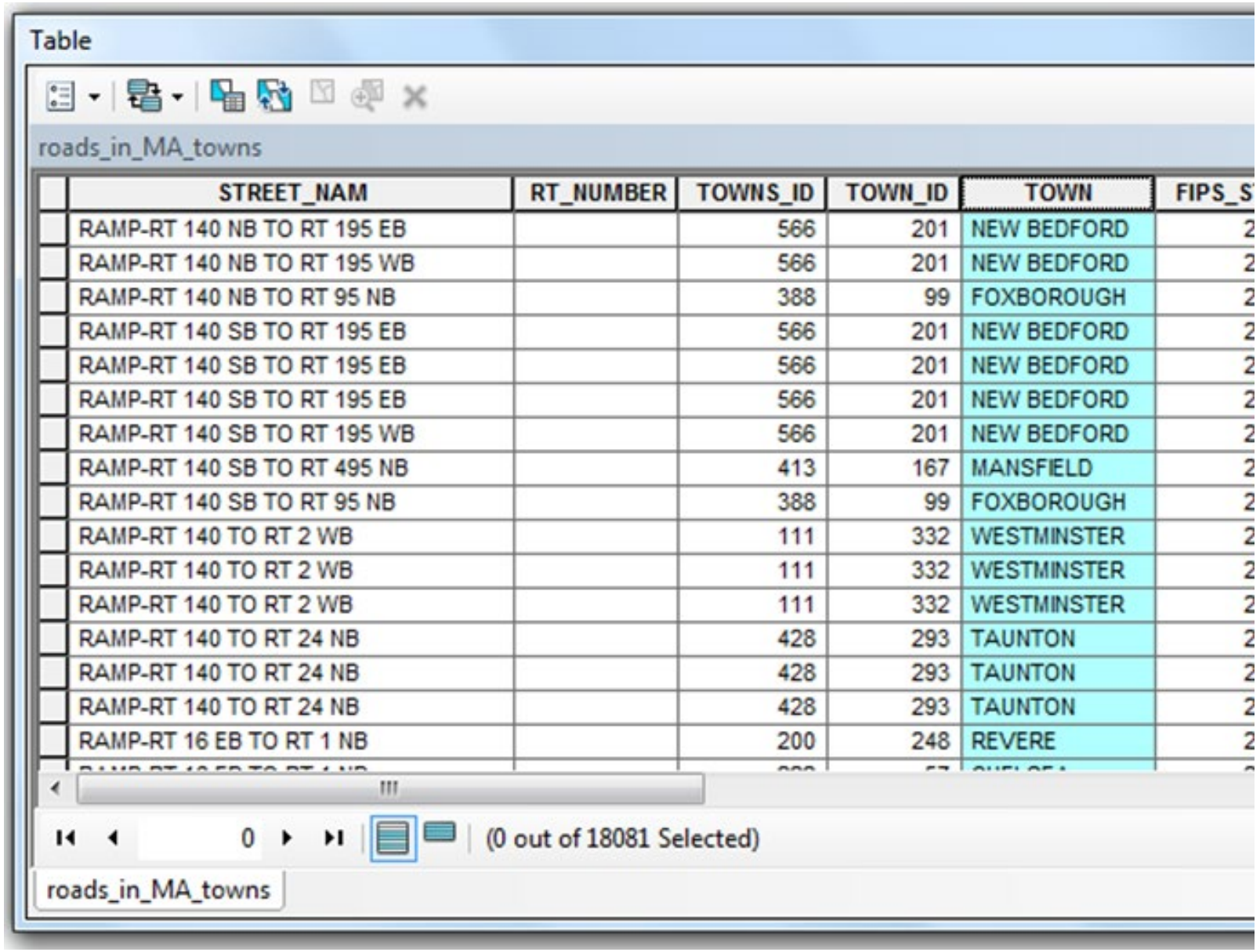
- Segments symbolized by town



What is the total length of roads in each town in MA?

Our original question

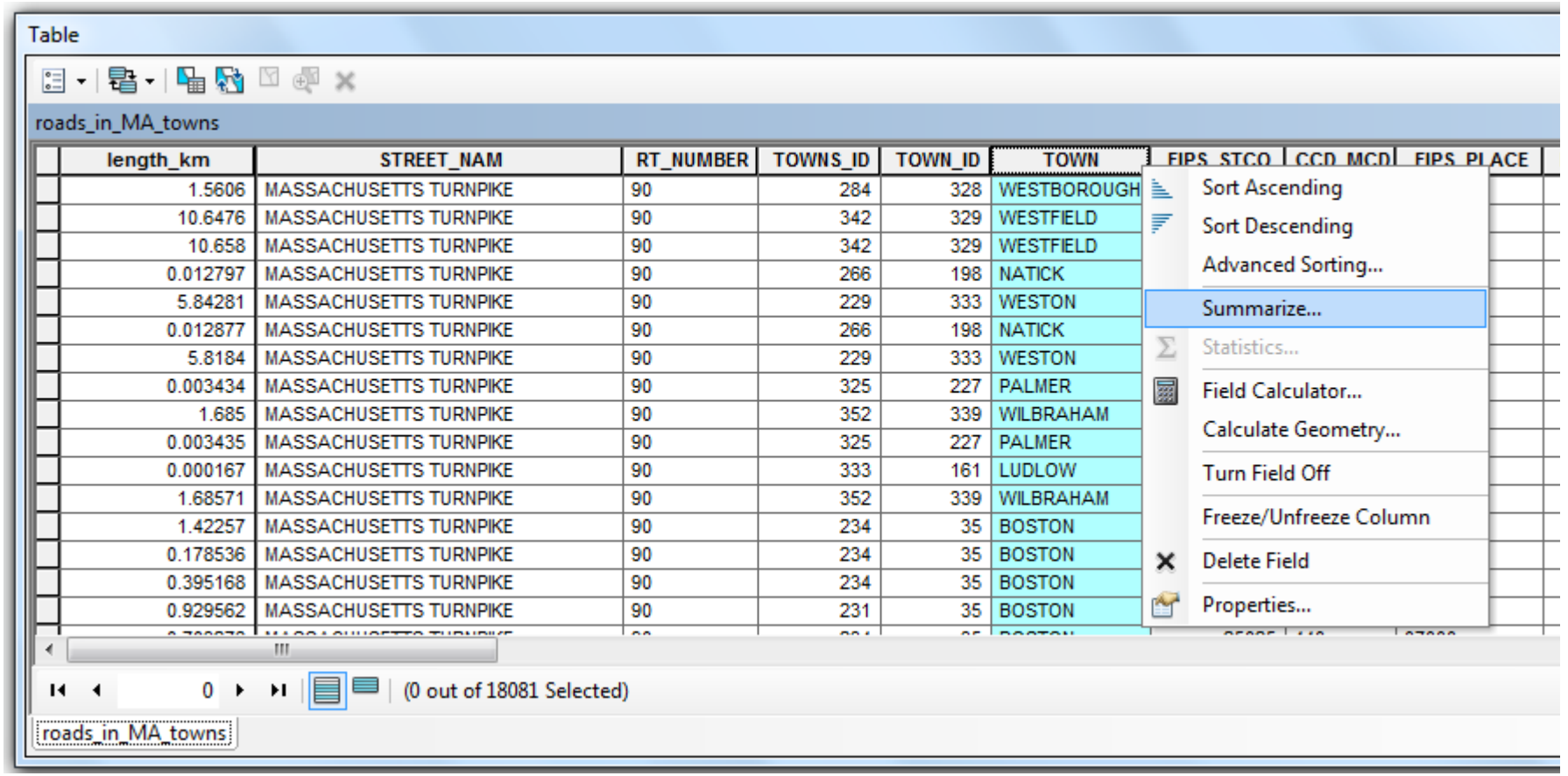
- We could use brute force to manually select features that match each of the 631 towns, or…



The Power of Summarize and Intersect!

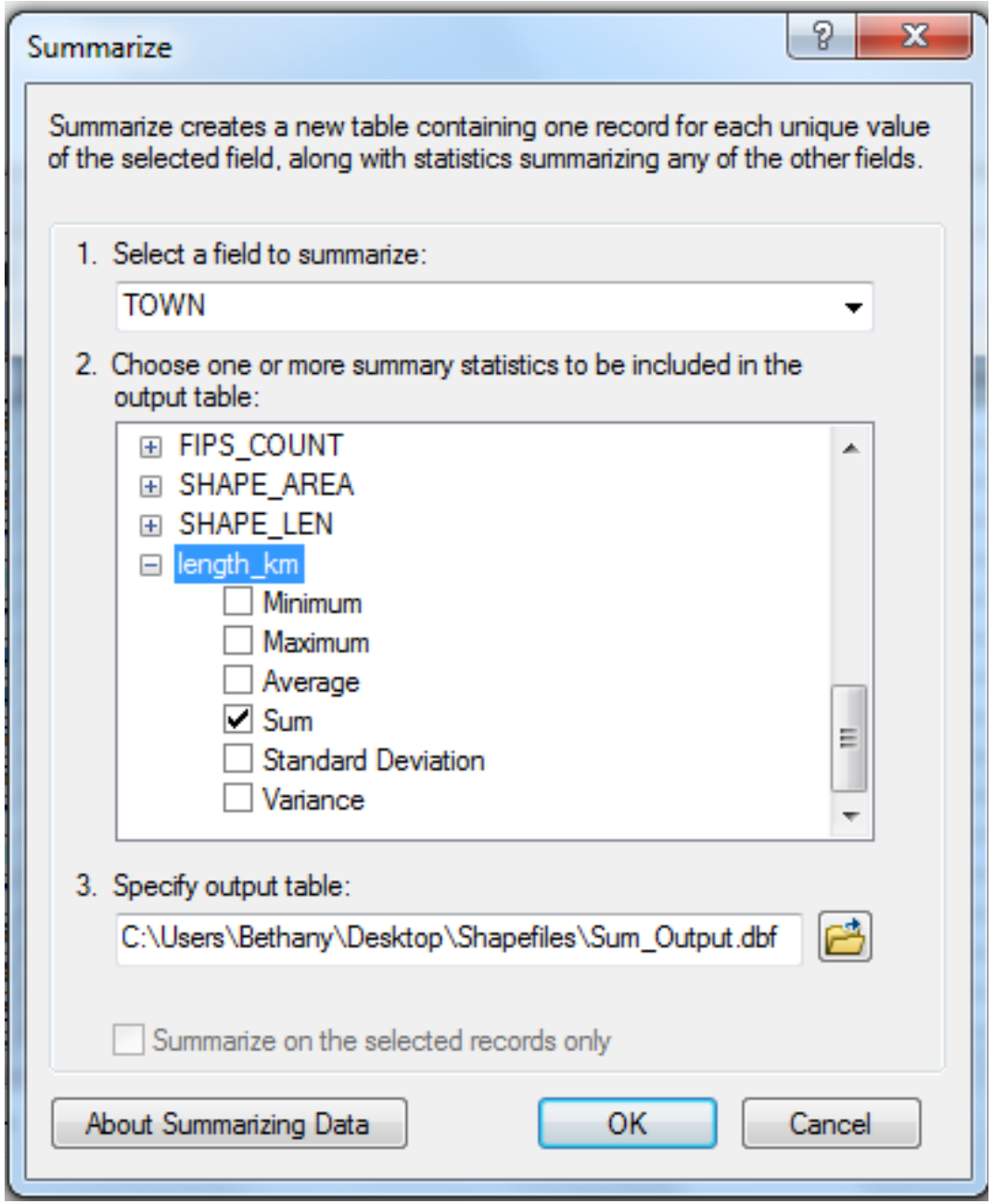
Two great tastes that taste great together.

Why is the summarize tool awesome?



Aggregation

- Summarize will count features, aggregated by unique values
- It will also perform calculations like total length, area, etc!



Aggregation

Summarize creates a table with:

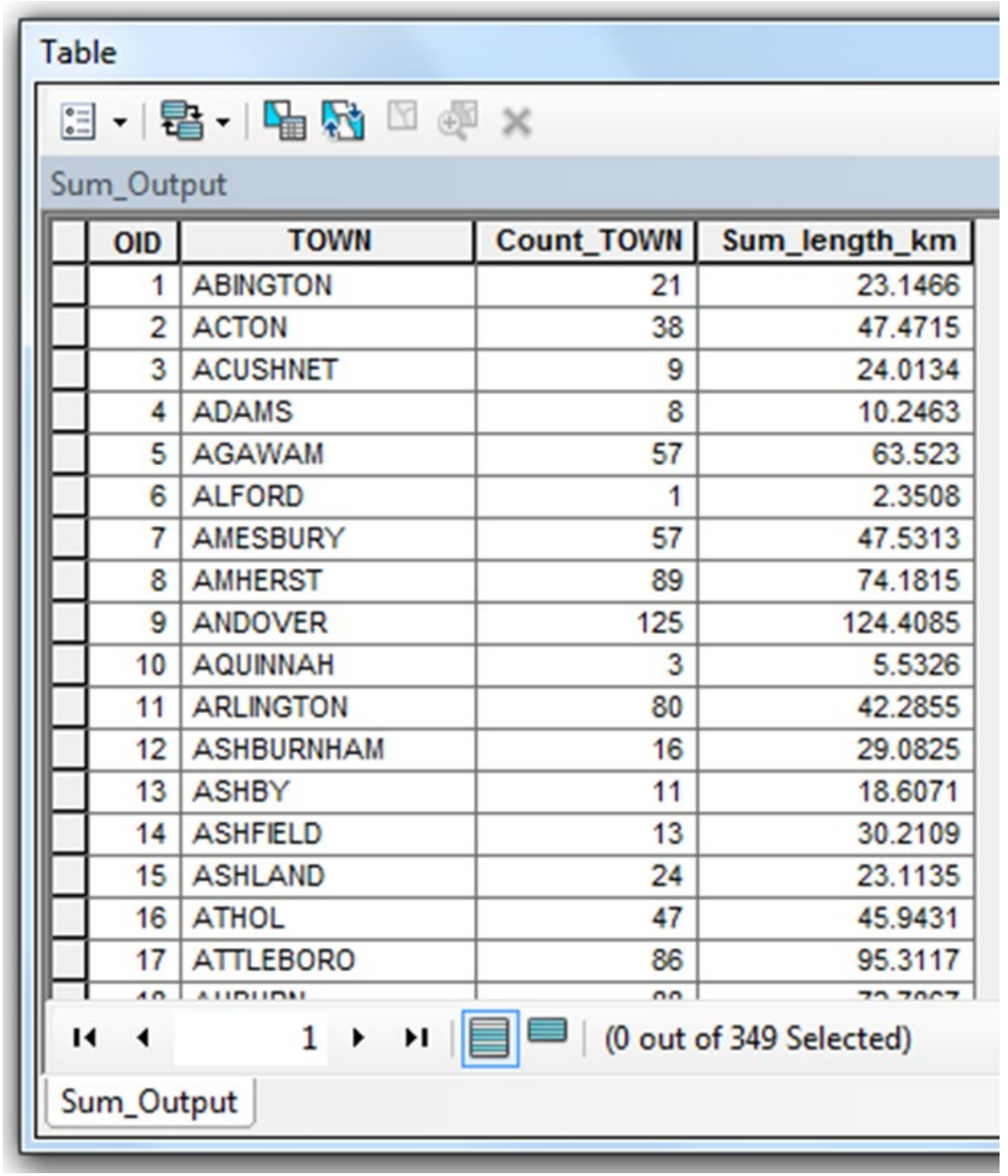
Count_TOWN:

- Total number of major road segments in that town.

Sum_length_km:

- Sum of the lengths of all major roads in that town.

Are we finished?

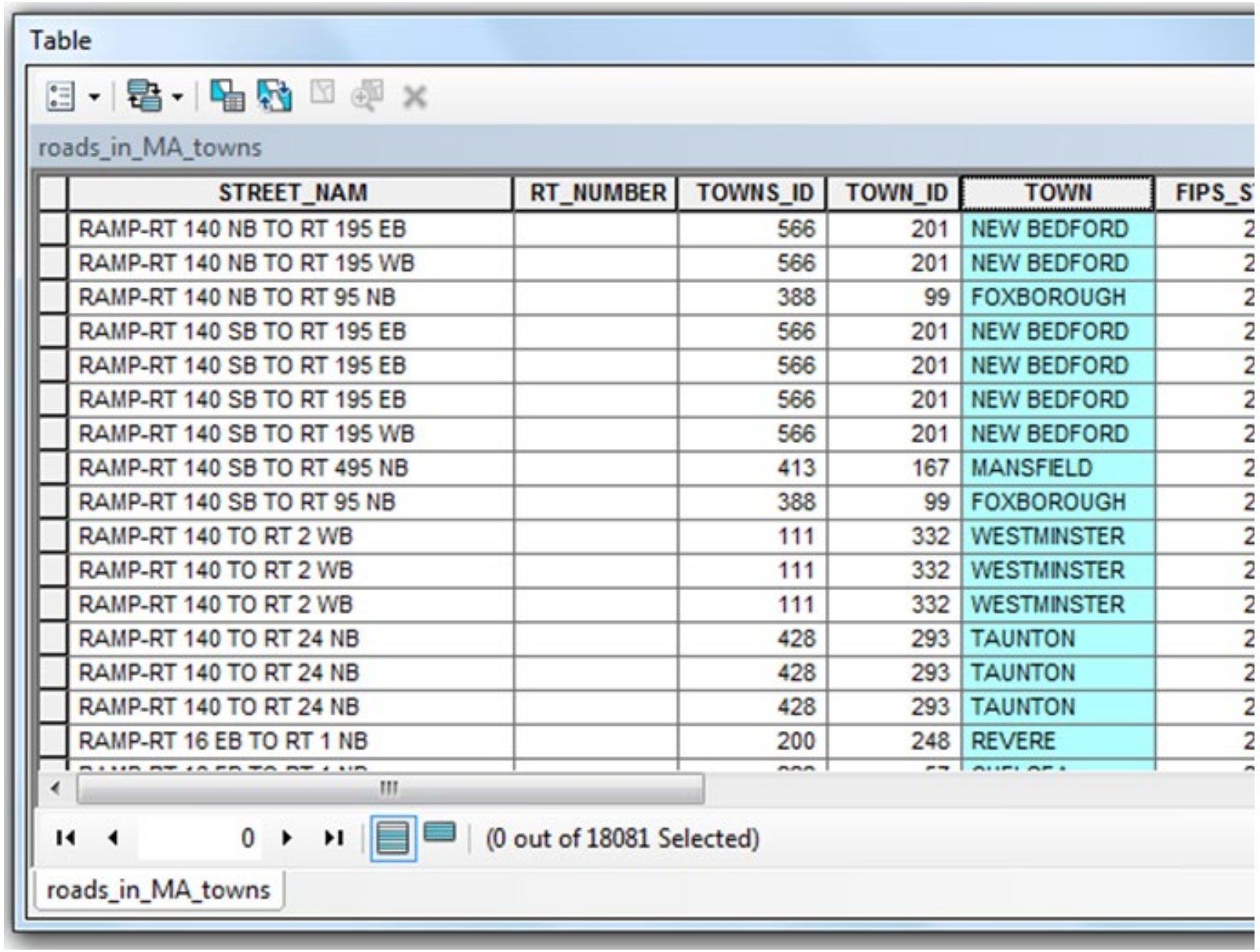


OID	TOWN	Count_TOWN	Sum_length_km
1	ABINGTON	21	23.1466
2	ACTON	38	47.4715
3	ACUSHNET	9	24.0134
4	ADAMS	8	10.2463
5	AGAWAM	57	63.523
6	ALFORD	1	2.3508
7	AMESBURY	57	47.5313
8	AMHERST	89	74.1815
9	ANDOVER	125	124.4085
10	AQUINNAH	3	5.5326
11	ARLINGTON	80	42.2855
12	ASHBURNHAM	16	29.0825
13	ASHBY	11	18.6071
14	ASHFIELD	13	30.2109
15	ASHLAND	24	23.1135
16	ATHOL	47	45.9431
17	ATTLEBORO	86	95.3117

What is the total length of roads in each town in MA?

We now have a total road length for each town…

how could we **join** these measurements to our towns attribute table?

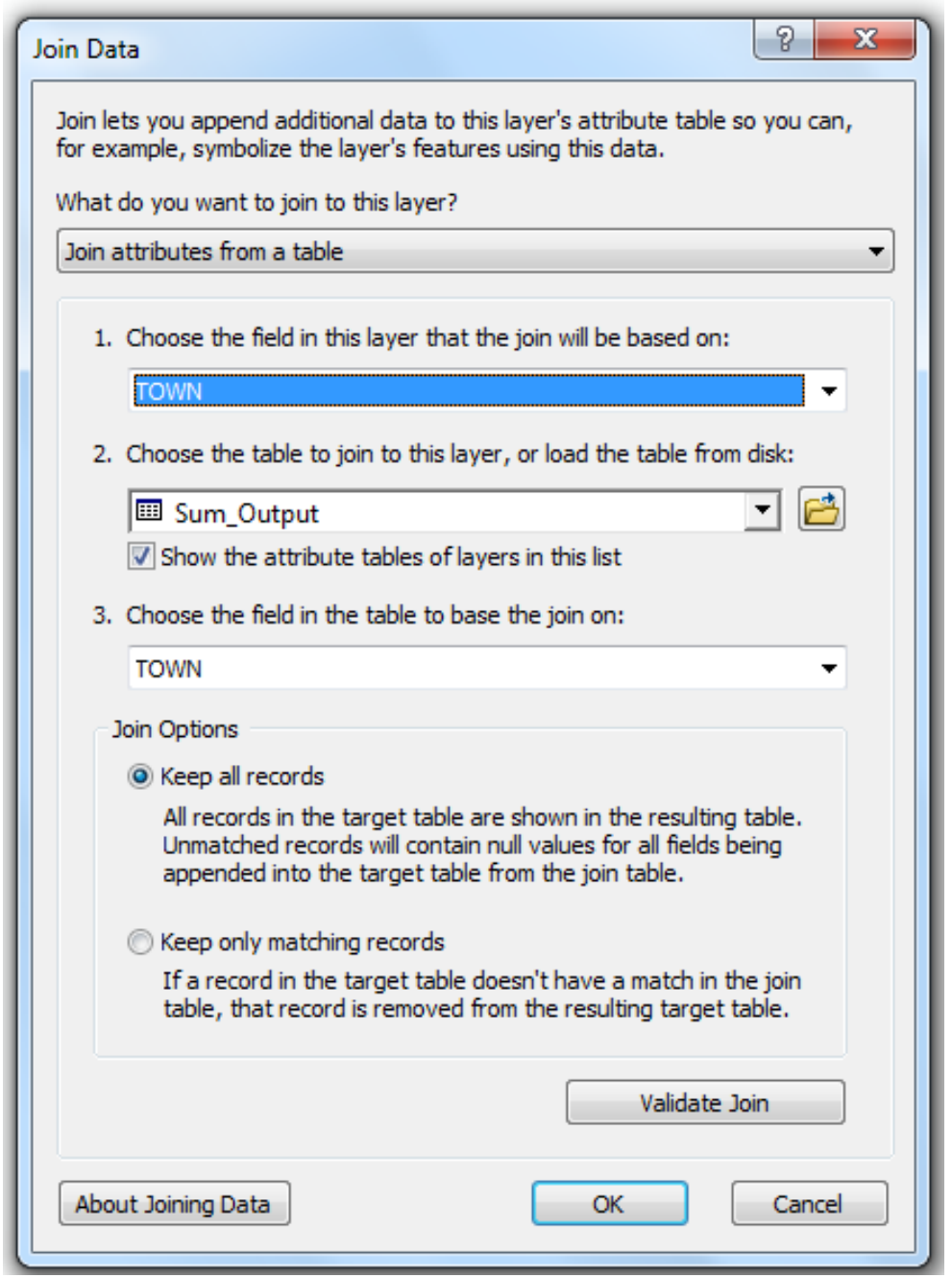


Join!

Join the summary table back to the ‘towns’ attribute table.

- Both have the common attribute of TOWN.

This will bring the attribute of total road length into your polygon shapefile of roads.



What is the name for this kind of map?