

GIS: Geographic Information Systems

Module 3: The Vector Data Model

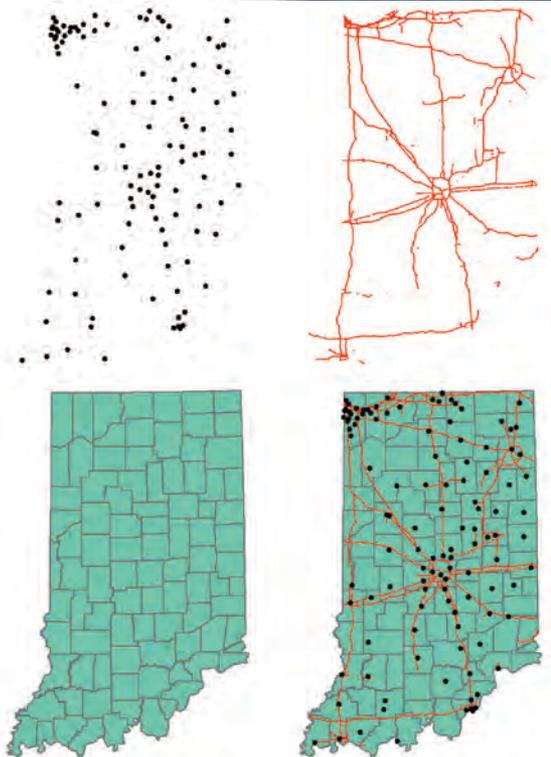
Matthew L. Sisk

Center for Digital Scholarship

Hesburgh Library, University of Notre Dame

library.nd.edu/cds/

Vectors:



Table

Indiana Counties

FID	Shape *	NAME_L	POP2000	MALES	FEMALES	AGE_UNDER5	AGE_5_17	AGE_18_21	AGE_22_29
0	Polygon	Steuben	33214	16771	16443	2199	6322	2241	3307
1	Polygon	Lagrange	34909	17681	17228	3432	8381	2199	3674
2	Polygon	Elkhart	182791	90848	91943	14800	37999	9881	20712
3	Polygon	St Joseph	265559	128133	137426	18673	49616	20658	28143
4	Polygon	Lake	484564	233367	251197	34639	95158	26621	48719
5	Polygon	Porter	146798	72046	74752	9488	28314	9093	14112
6	Polygon	La Porte	110106	56539	53567	7116	19888	5454	11258
7	Polygon	De Kalb	40285	20059	20226	3061	8238	1977	4210
8	Polygon	Noble	46275	23310	22965	3695	9729	2441	5074
9	Polygon	Marshall	45128	22415	22713	3290	9369	2335	4256
10	Polygon	Kosciusko	74057	36982	37075	5519	15043	3728	7655
11	Polygon	Starke	23556	11660	11896	1520	4792	1142	2169
12	Polygon	Whitley	30707	15238	15469	2101	6112	1469	2853
13	Polygon	Allen	331849	162425	169424	25440	66511	18022	36702
14	Polygon	Jasper	30043	14888	15155	2077	6157	1985	2856
15	Polygon	Newton	14566	7239	7327	902	2945	728	1265
16	Polygon	Fulton	20511	10139	10372	1348	3986	961	1835
17	Polygon	Pulaski	13755	6938	6817	845	2858	610	1171
18	Polygon	Wabash	34960	16957	18003	2073	6504	2377	3234
19	Polygon	Huntington	38075	18537	19538	2536	7412	2395	3600

(0 out of 92 Selected)

Indiana Counties

Three types of geometry

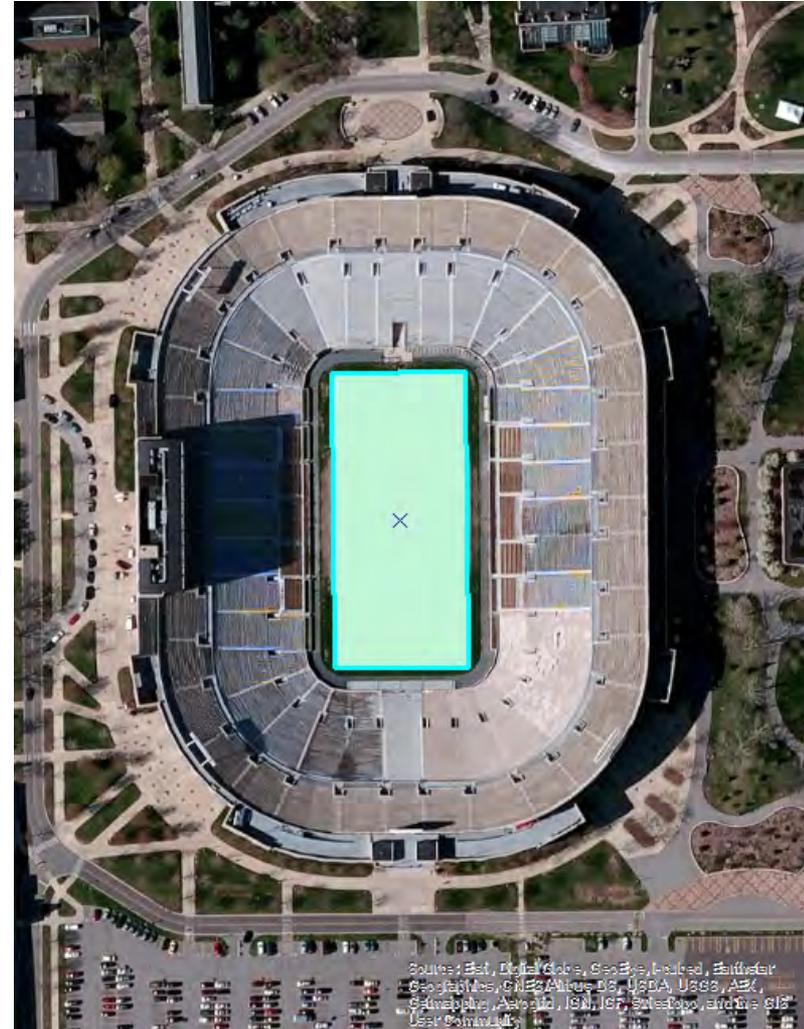
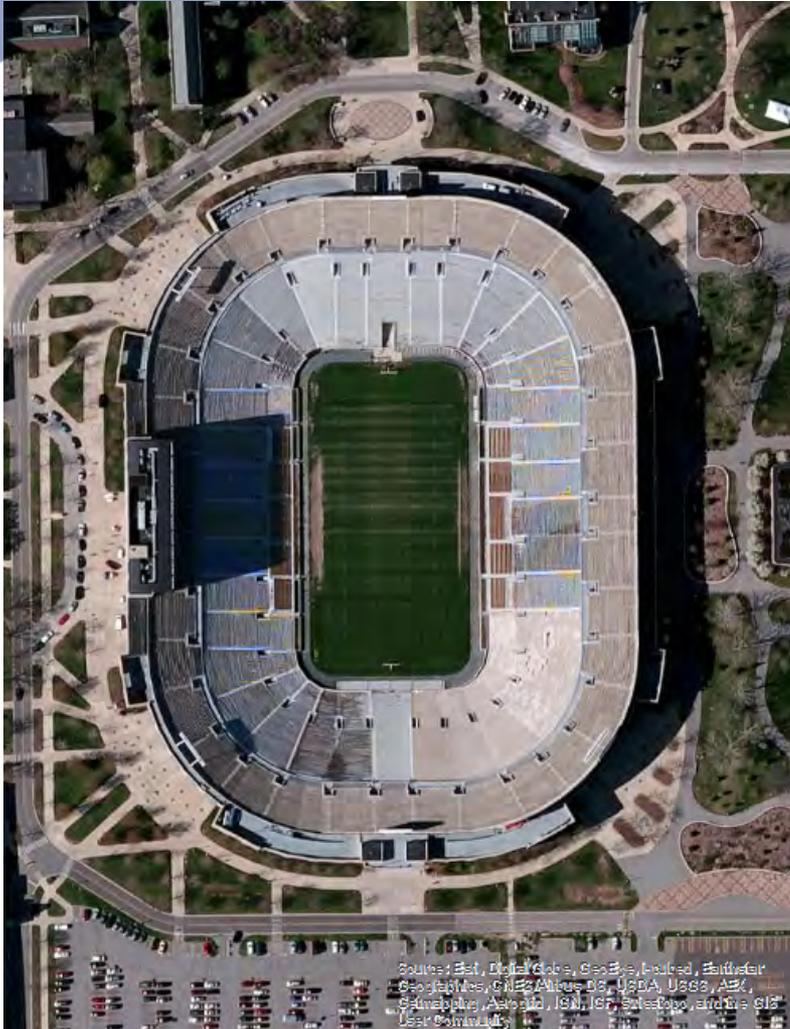
- Points
- Lines
- Areas

Data tables

- Geometric objects
 - Three basic types:
 - Points, lines and polygons
 - Other, more complicated, types can represent topography or movement
 - Basic unit is the point and coordinates

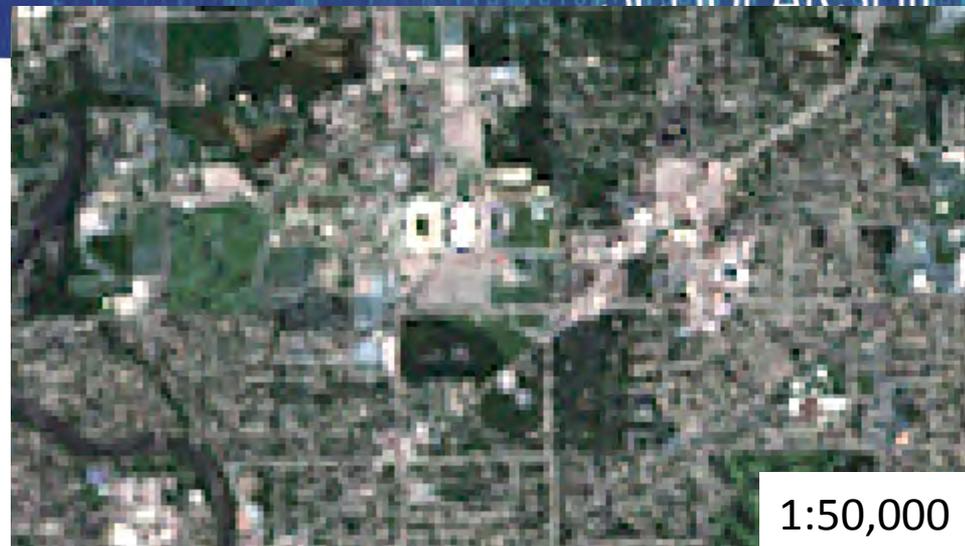
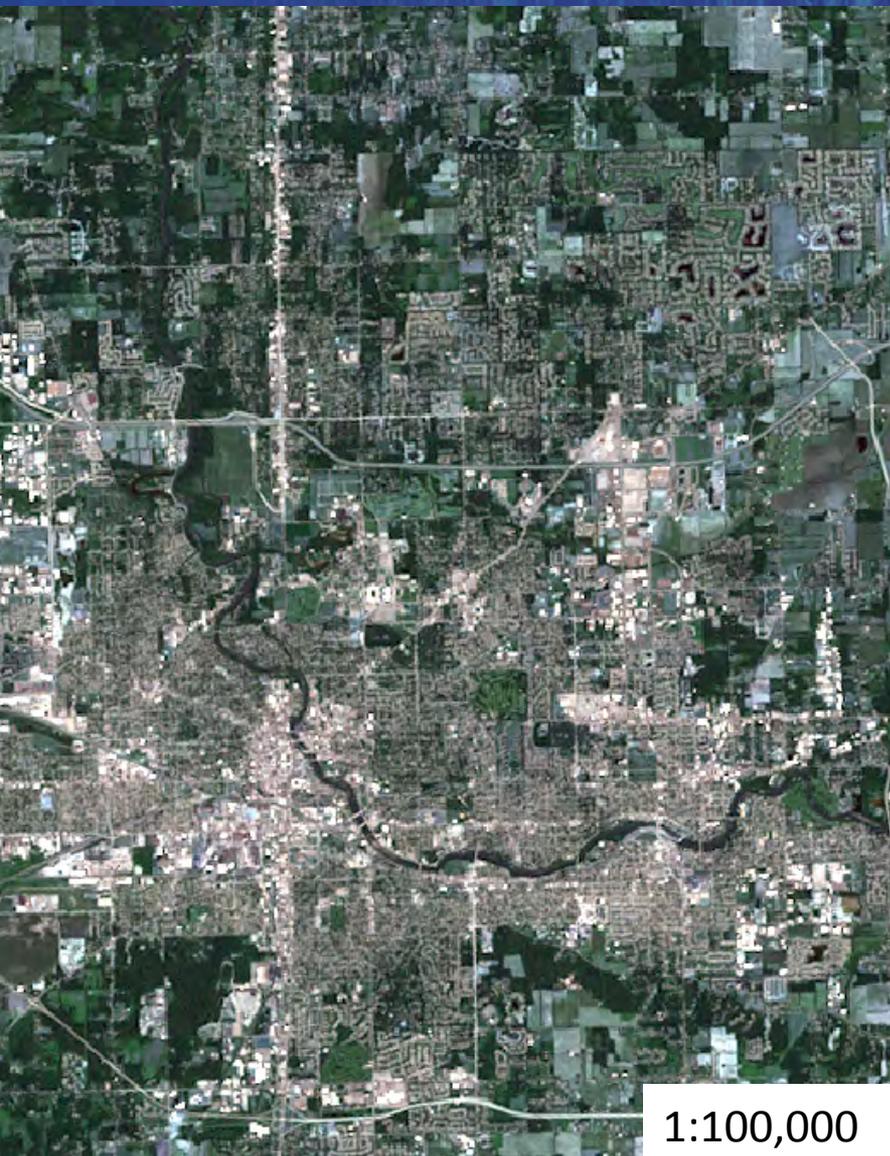
Raster and Vector

CENTER for
DIGITAL
SCHOLARSHIP



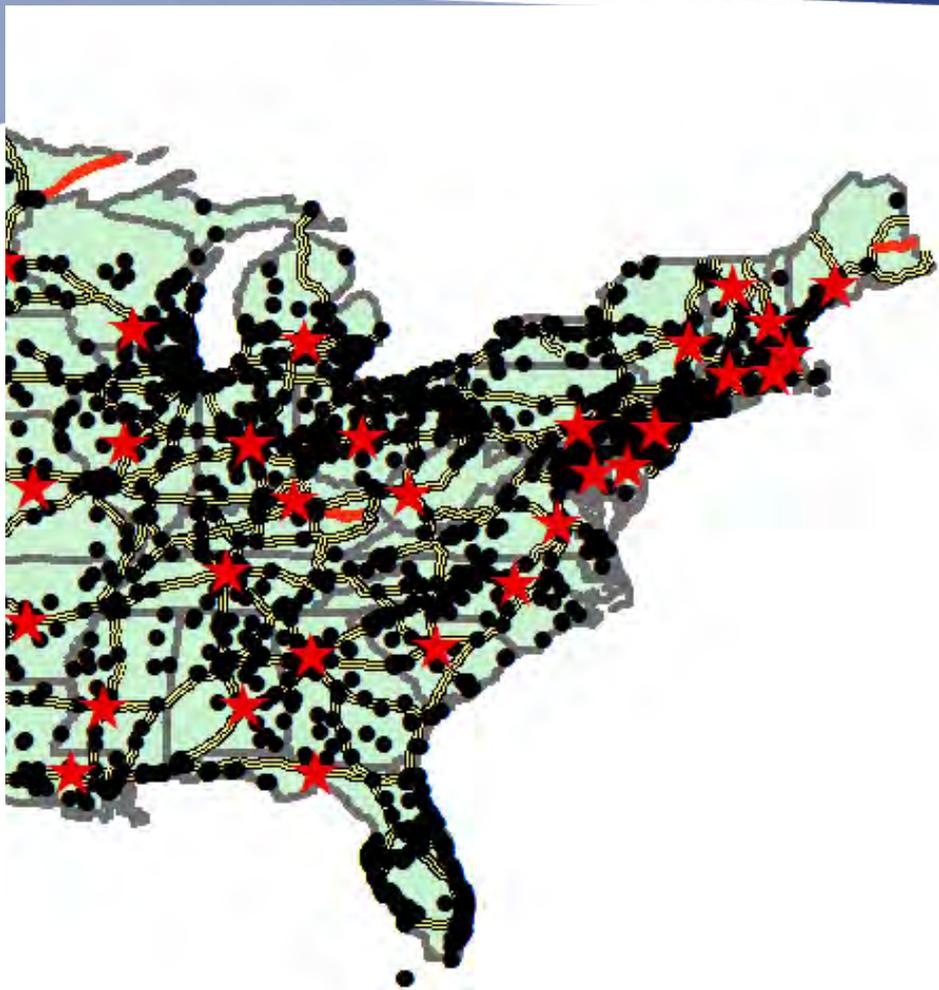
Raster resolution comparison

CENTER for
DIGITAL
SCHOLARSHIP

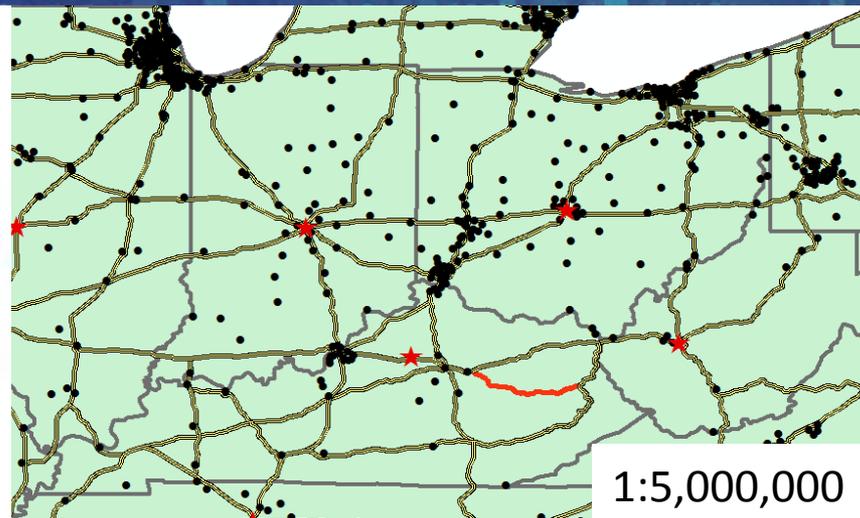


Vector Resolution Comparison

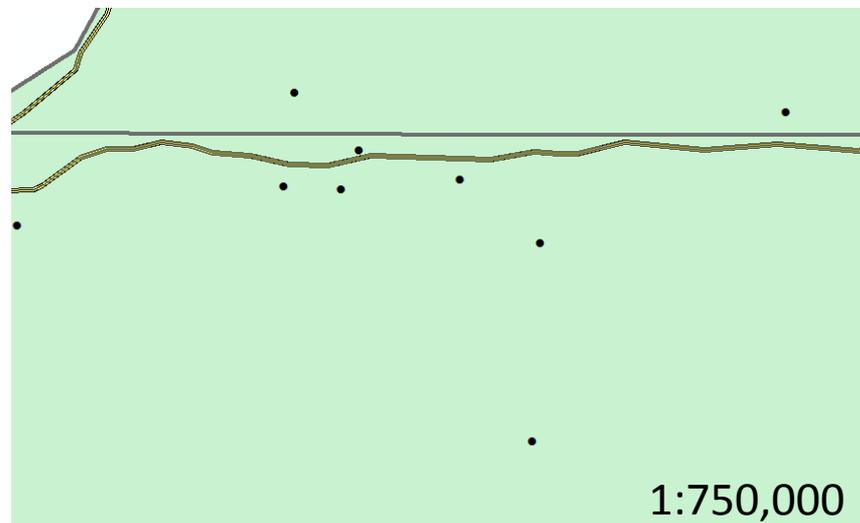
CENTER for
DIGITAL
SCHOLARSHIP



1:25,000,000

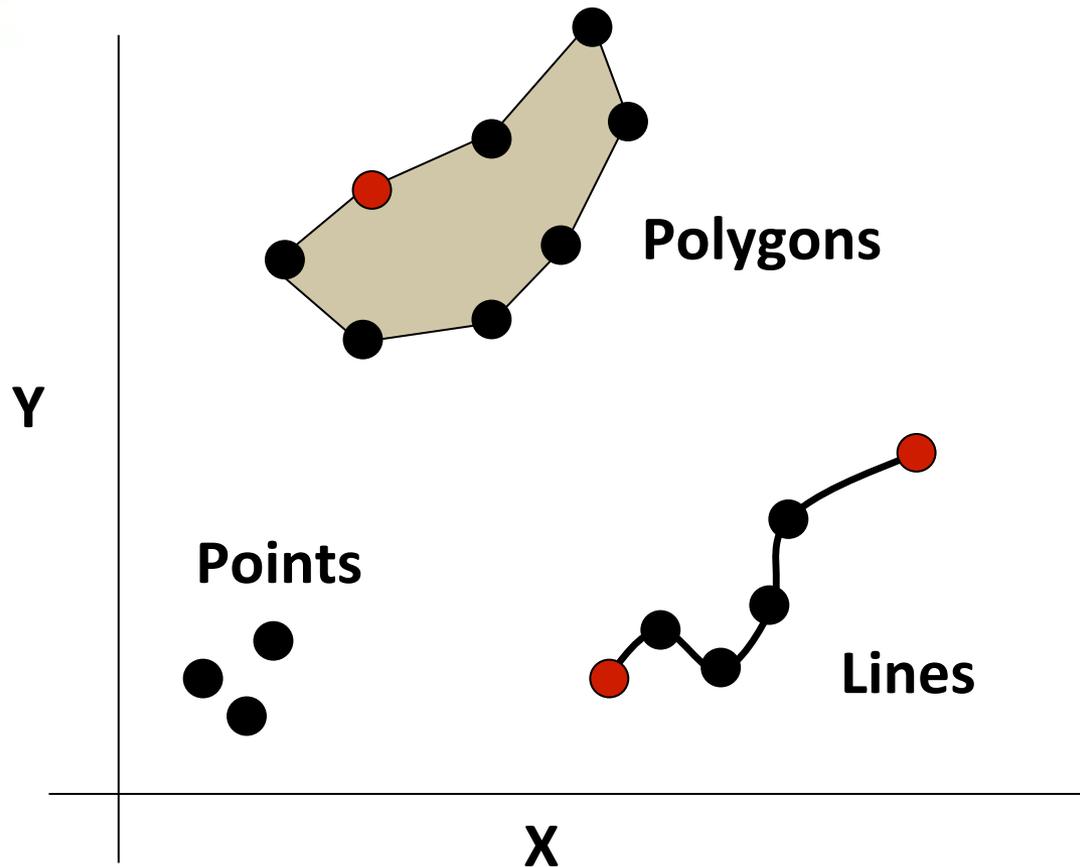


1:5,000,000

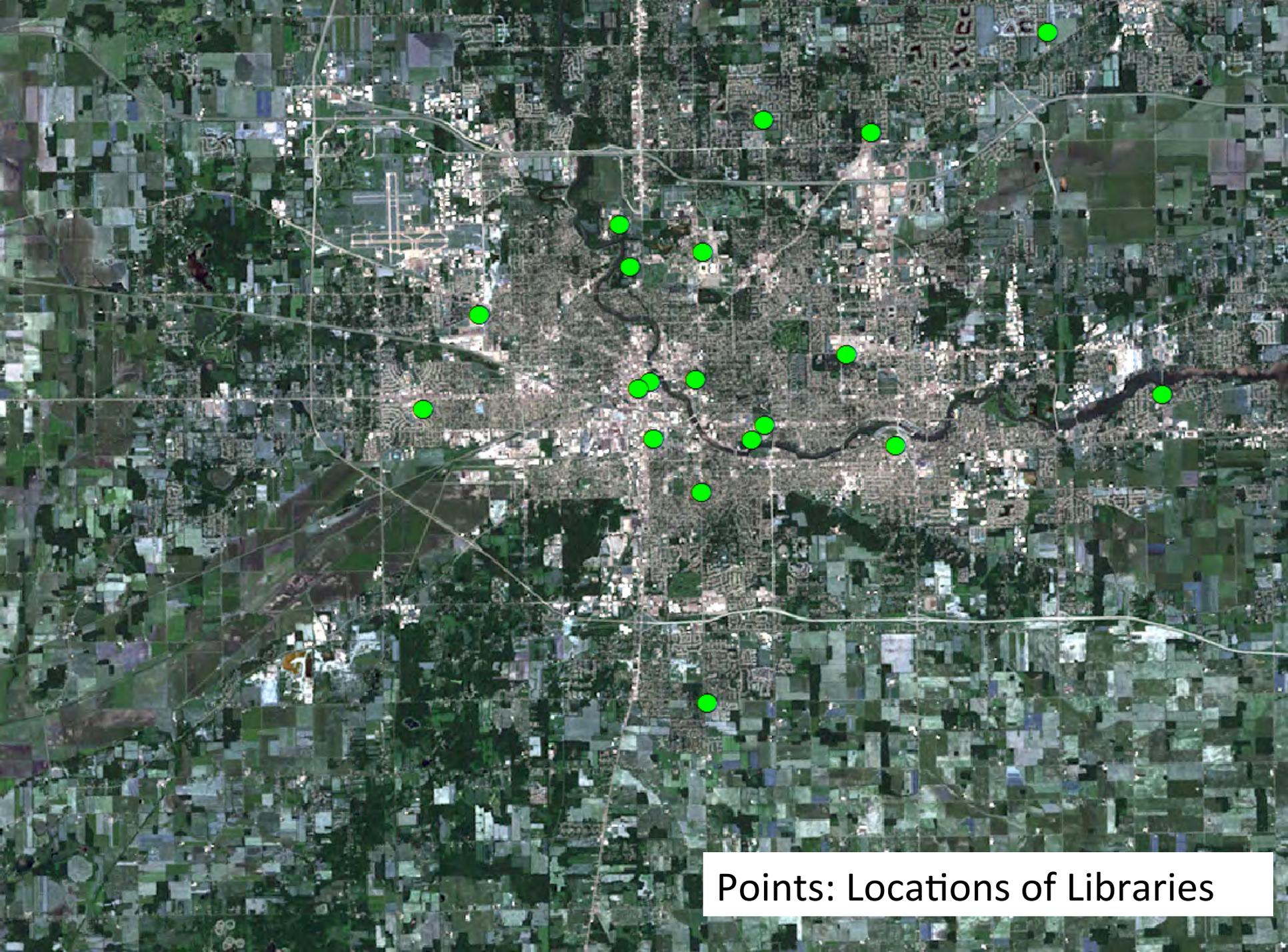


1:750,000

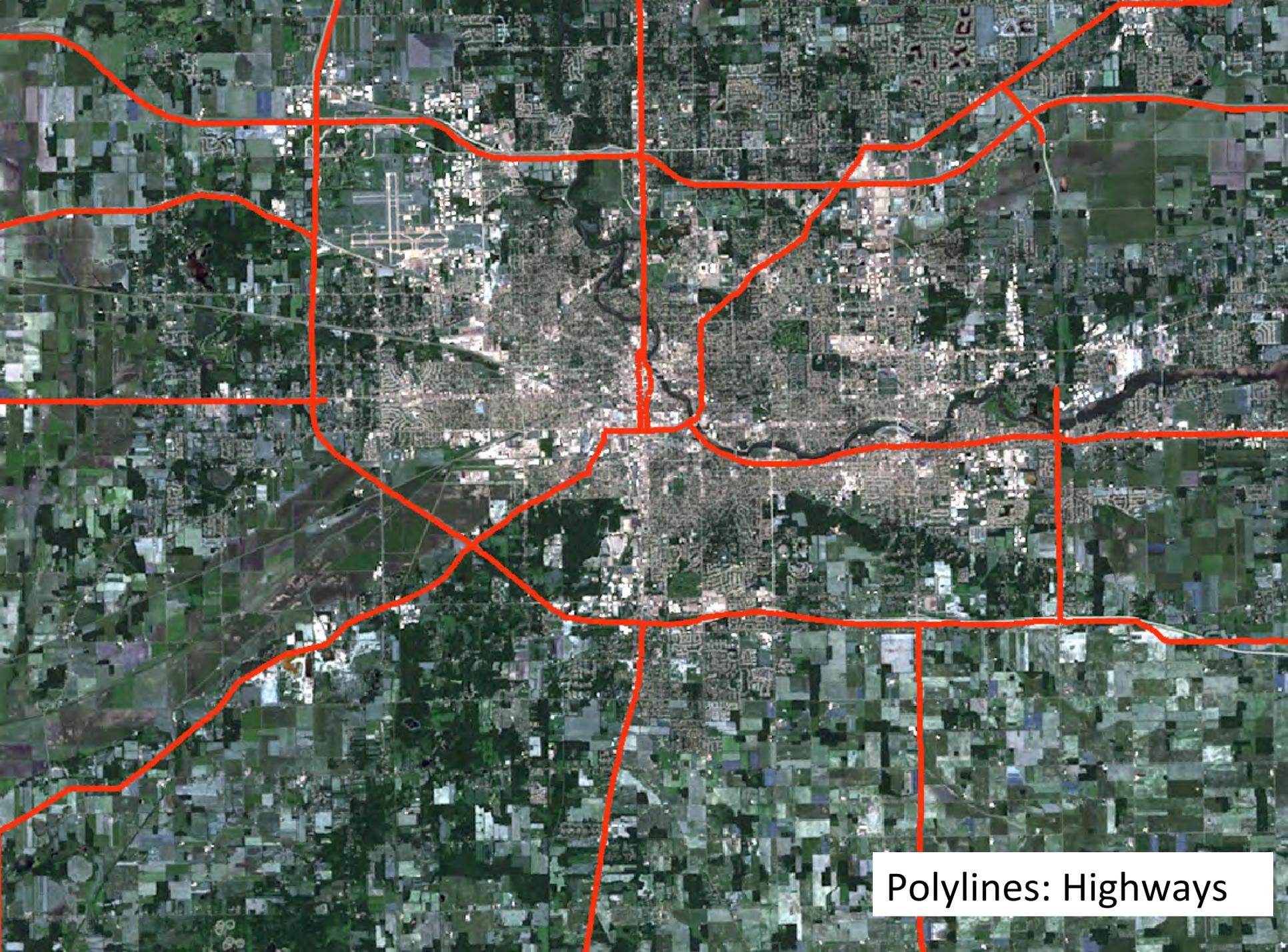
Vector model



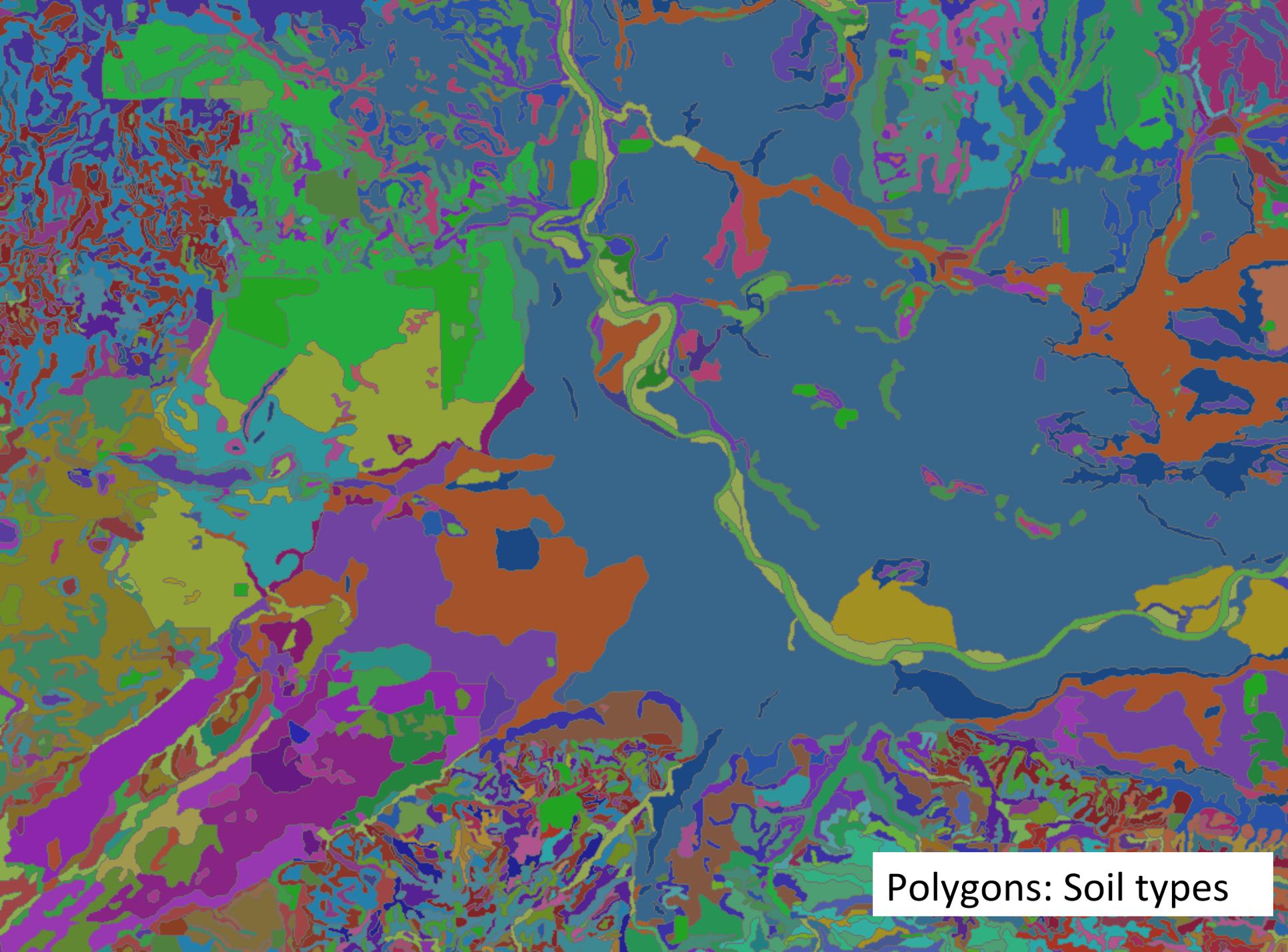
Features are stored as a series of x - y coordinates in a coordinate system.



Points: Locations of Libraries



Polylines: Highways

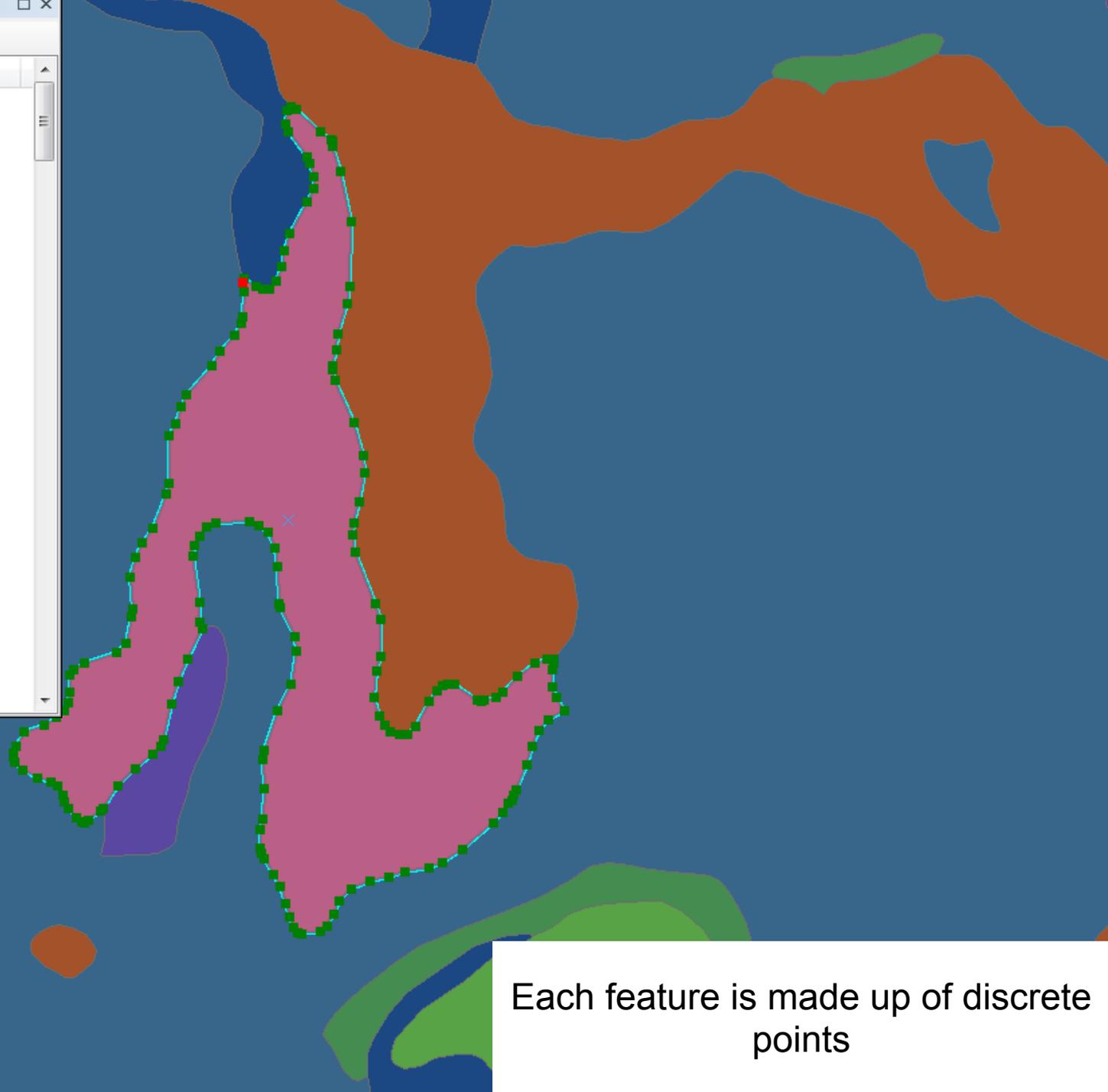


Polygons: Soil types

Edit Sketch Properties

Finish Sketch

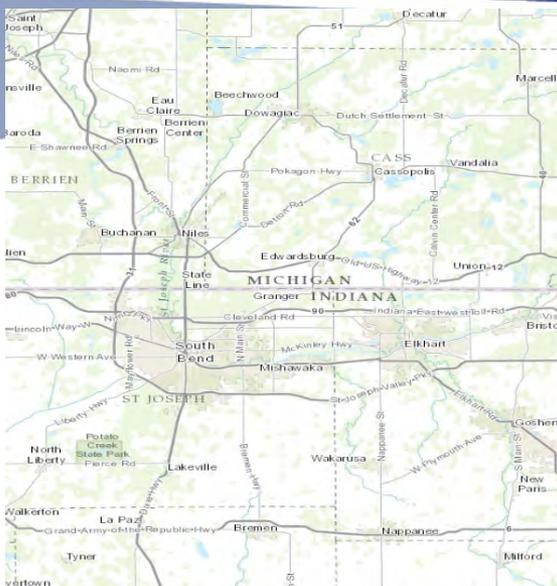
#	X	Y
0	562689.216	4618588.203
1	562708.227	4618573.590
2	562719.072	4618569.293
3	562730.268	4618571.289
4	562740.664	4618583.293
5	562748.765	4618607.282
6	562753.285	4618632.599
7	562761.107	4618658.279
8	562789.468	4618709.752
9	562799.034	4618732.060
10	562800.766	4618749.237
11	562794.590	4618771.277
12	562789.426	4618781.283
13	562759.963	4618822.197
14	562754.695	4618834.566
15	562757.079	4618855.341
16	562764.279	4618862.793
17	562772.009	4618857.631
18	562810.020	4618823.120
19	562828.484	4618810.327
20	562830.006	4618807.027
21	562829.959	4618799.581
22	562842.458	4618759.486
23	562859.771	4618677.683
24	562857.437	4618574.992



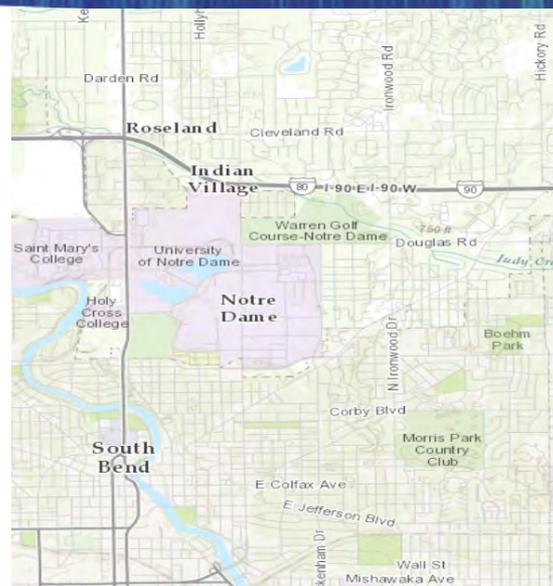
Each feature is made up of discrete points

- Very little in the real world is precise
 - The different types of vector data are choices we make to represent geographic features
 - Something is always lost
 - Boundaries are rarely clear, trees are not individual points, rivers have width
- Using a vector model makes analysis and visualization much easier
- But, we must always be aware of this schematization

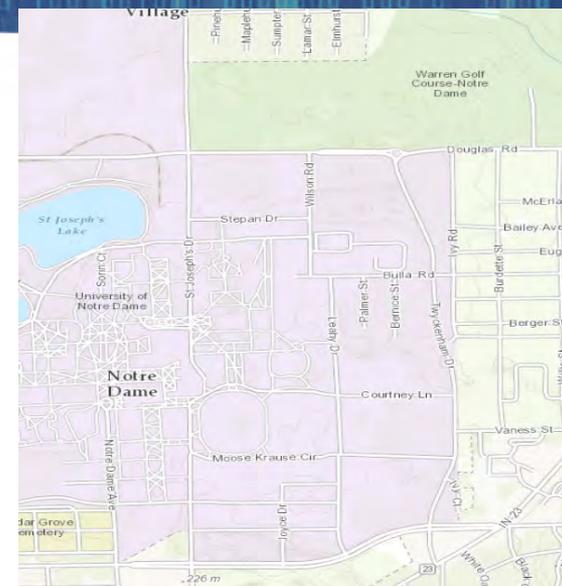
Different displays at different scales



1:500,000



1:50,000



1:15,000

- GIS data often need to be displayed differently at different scales
- This is much easier to do for vector data than for raster data
- Rendering points and polygons is no different at scale

Vector vs. Raster

Vector Advantages

- Economical in space
- Good for discrete features
- More flexible with regard to scale
- Can store many different attributes with each feature

Vector Disadvantages

- More schematized version of reality
- Poorly suited for continuous phenomena

The vector data model

ATTRIBUTE DATA

Attribute data

Table

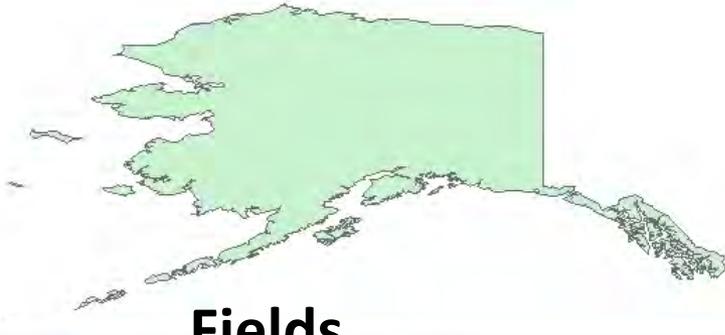
USCITIES

FID	Shape	STATE_CITY	CITY_NAME	STATE_NAME	CAPITAL	POP1990	HOUSEHOLDS	MALES	FEMALES
0	Point	0216750	College	Alaska	N	11249	3764	5936	5313
1	Point	0224230	Fairbanks	Alaska	N	30843	10885	16543	14300
2	Point	0203000	Anchorage	Alaska	N	226338	82702	116367	109971
3	Point	0236400	Juneau	Alaska	Y	26751	9902	13575	13176
4	Point	5305280	Bellingham	Washington	N	52179	21189	24838	27341
5	Point	3035050	Havre	Montana	N	10201	4027	4955	5246
6	Point	5301990	Anacortes	Washington	N	11451	4669	5506	5945
7	Point	5347560	Mount Vernon	Washington	N	17647	6885	8459	9188
8	Point	5350360	Oak Harbor	Washington	N	17176	5971	8532	8644
9	Point	3853380	Minot	North Dakota	N	34544	13965	16467	18077
10	Point	3040075	Kalispell	Montana	N	11917	5237	5474	6443
11	Point	3886220	Williston	North Dakota	N	13131	5133	6297	6834
12	Point	5355365	Port Angeles	Washington	N	17710	7360	8493	9217
13	Point	5349992	North Marysville	Washington	N	18711	6116	9335	9376
14	Point	5343955	Marysville	Washington	N	10328	4288	4860	5468
15	Point	5377542	West Lake Stevens	Washington	N	12453	4265	6236	6217
16	Point	5322640	Everett	Washington	N	69961	28679	34714	35247
17	Point	3832060	Grand Forks	North Dakota	N	49425	18531	24735	24690
18	Point	5352765	Paine Field-Lake Stickney	Washington	N	18670	7656	9376	9294
19	Point	5364452	Silver Lake-Fircrest	Washington	N	24474	7878	12365	12109
20	Point	5337705	Lake Serene-North Lynnwood	Washington	N	14290	5427	7194	7096
21	Point	5343815	Martha Lake	Washington	N	10155	3588	5124	5031
22	Point	5340840	Lynnwood	Washington	N	28695	11331	13974	14721
23	Point	5320750	Edmonds	Washington	N	30744	12628	14554	16190
24	Point	5340870	North Creek Canyon Park	Washington	N	32226	7744	11617	11610

1 (0 out of 3149 Selected)

STATES USCITIES

Feature Attributes



Fields



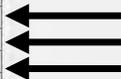
Table

STATES

FID	Shape	STATE_NAME	STATE_FIPS	STATE_ABBR
0	Polygon	Washington	53	WA
1	Polygon	Montana	30	MT
2	Polygon	Maine	23	ME
3	Polygon	North Dakota	38	ND
4	Polygon	South Dakota	46	SD
5	Polygon	Wyoming	56	WY
6	Polygon	Wisconsin	55	WI
7	Polygon	Idaho	16	ID
8	Polygon	Vermont	50	VT
9	Polygon	Minnesota	27	MN
10	Polygon	Oregon	41	OR
11	Polygon	New Hampshire	33	NH
12	Polygon	Iowa	19	IA
13	Polygon	Massachusetts	25	MA
14	Polygon	Nebraska	31	NE
15	Polygon	New York	36	NY
16	Polygon	Pennsylvania	42	PA
17	Polygon	Connecticut	09	CT
18	Polygon	Rhode Island	44	RI
19	Polygon	New Jersey	34	NJ
20	Polygon	Indiana	18	IN

STATES (0 out of 51 Selected)

Records



Spreadsheets vs. databases

- Spreadsheets
 - Can have more than one type of in each column
 - Best used for generating graphs and summaries
- Databases
 - Limited to one type of data in each field
 - Multiple tables in a database can be related on a common field and changes in one will be reflected in the others

GIS Attribute Tables:

- Fundamental to the strength of GIS
 - More accurately viewed as part of a larger database
 - Database structure is critical to the success of analyses
- Tables can be
 - Tied to features by spatial location
 - Tied to other tables by a common field

The vector data model

STORING VECTOR DATA

The Shapefile

Each shapefile is made up of several different files

- For example, a shapefile called “example” in windows explorer would have at least three different files
 - example.shp: shape format
 - the actual points for each feature
 - example.shx: shape index format
 - positional index of the feature geometry to allow seeking forwards and backwards quickly
 - example.dbf: attribute table
 - Columns with the attributes for each shape, in dBase format
 - And possibly up to 4 more

The vector data model

INTERACTING WITH VECTOR DATA

- In ArcGIS we have two main ways to query vector data
 - Select by Attributes
 - Within the attribute table
 - Using values of particular attributes
 - » e.g height > 10, name = “Steve”, percent_male = percent_female
 - Select by Location
 - Between different layers using geographic interactions
 - Using the principles of topology
 - » e.g. roads with 10 miles of a school, cities with the US, lakes touching Indiana

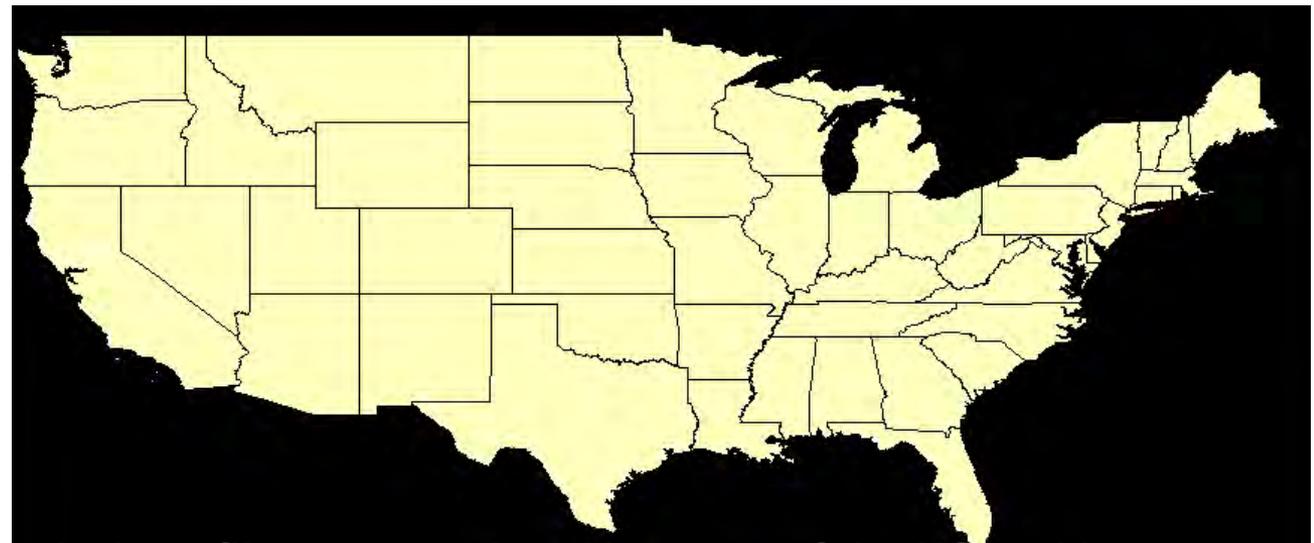
Select by Attribute: Example

Table

US_States

FID	Shape *	STATE_NAME	POP2000
0	Polygon	Hawaii	1184688
1	Polygon	Washington	5835089
2	Polygon	Montana	885795
3	Polygon	Maine	1257219
4	Polygon	North Dakota	631032
5	Polygon	South Dakota	734993
6	Polygon	Wyoming	479673
7	Polygon	Wisconsin	5277833
8	Polygon	Idaho	1273309
9	Polygon	Vermont	596714
10	Polygon	Minnesota	4820250
11	Polygon	Oregon	3356108
12	Polygon	New Hampshire	1215100
13	Polygon	Iowa	2877060
14	Polygon	Massachusetts	6206482
15	Polygon	Nebraska	1672199
16	Polygon	New York	18223519
17	Polygon	Pennsylvania	11986139
18	Polygon	Connecticut	3289062
19	Polygon	Rhode Island	992011
20	Polygon	New Jersey	8192386
21	Polygon	Indiana	5979311
22	Polygon	Nevada	1879204
23	Polygon	Utah	2164175
24	Polygon	California	33603430
25	Polygon	Ohio	11281851
26	Polygon	Illinois	12187552
27	Polygon	District of Columbia	513618
28	Polygon	Delaware	762227
29	Polygon	West Virginia	1804812
30	Polygon	Maryland	5212902
31	Polygon	Colorado	4139027
32	Polygon	Kentucky	3988695
33	Polygon	Kansas	2672387
34	Polygon	Virginia	6945067
35	Polygon	Missouri	5502243
36	Polygon	Arizona	4894006
37	Polygon	Oklahoma	3383158
38	Polygon	North Carolina	7762819
39	Polygon	Tennessee	5539577

US_States (0 out of 51 Selected)



A shapefile representing the states of the US and its attribute table

Select by Attribute: Example

Table

US_States

FID	Shape *	STATE_NAME	POP2000
46	Polygon	Arkansas	2566938

(1 out of 51 Selected)

US_States

Select by Attributes

Enter a WHERE clause to select records in the table window.

Method : Create a new selection

"FID"
"STATE_NAME"
"POP2000"

= <> Like
> >= And
< <= Or
_ % () Not
Is Get Unique Values Go To:

SELECT * FROM US_States WHERE:
"STATE_NAME" = 'Arkansas'

Clear Verify Help Load... Save...
Apply Close



Select by Attribute: Example

FID	Shape *	STATE_NAME	POP2000
1	Polygon	Washington	5835089
7	Polygon	Wisconsin	5277833
14	Polygon	Massachusetts	6206482
16	Polygon	New York	18223519
17	Polygon	Pennsylvania	11986139
20	Polygon	New Jersey	8192386
21	Polygon	Indiana	5979311
24	Polygon	California	33603430
25	Polygon	Ohio	11281851
26	Polygon	Illinois	12187552
30	Polygon	Maryland	5212902
34	Polygon	Virginia	6945067
35	Polygon	Missouri	5502243
38	Polygon	North Carolina	7762819
39	Polygon	Tennessee	5539577
40	Polygon	Texas	20398490
44	Polygon	Georgia	7950119
48	Polygon	Florida	15341185
49	Polygon	Michigan	9907530

Select by Attributes

Enter a WHERE clause to select records in the table window.

Method: Create a new selection

"FID"
"STATE_NAME"
"POP2000"

= <> Like 479673
> >= And 513618
< <= Or 596714
_ % () Not 624523
734993

Is Get Unique Values Go To:

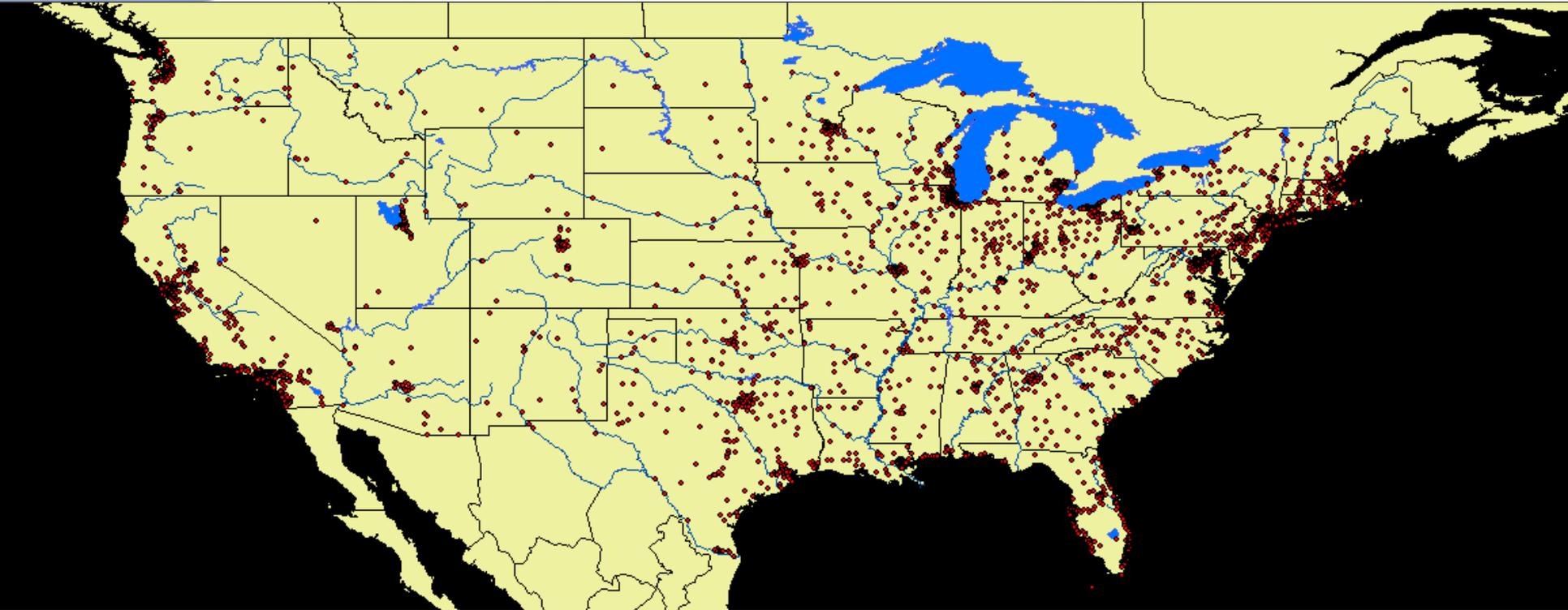
SELECT * FROM US_States WHERE:
"POP2000" >5000000

Clear Verify Help Load... Save...
Apply Close



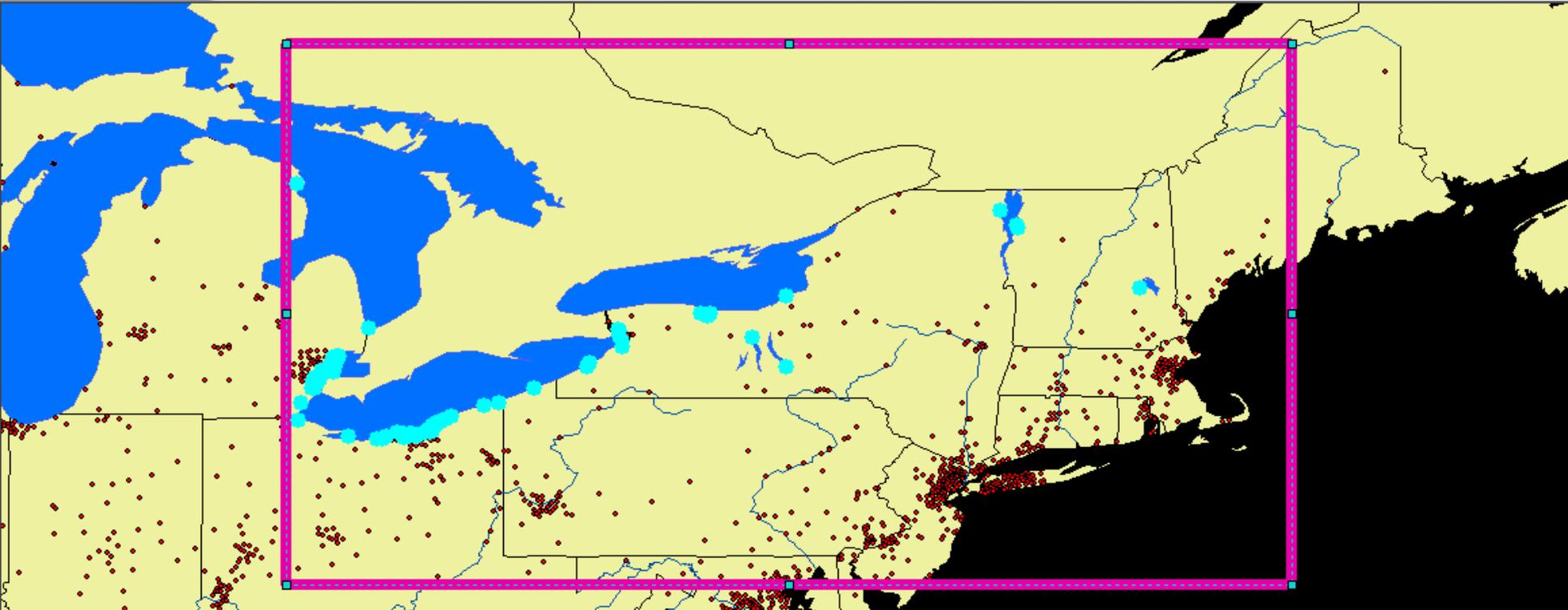
Select by Location: Example

CENTER for
DIGITAL
SCHOLARSHIP



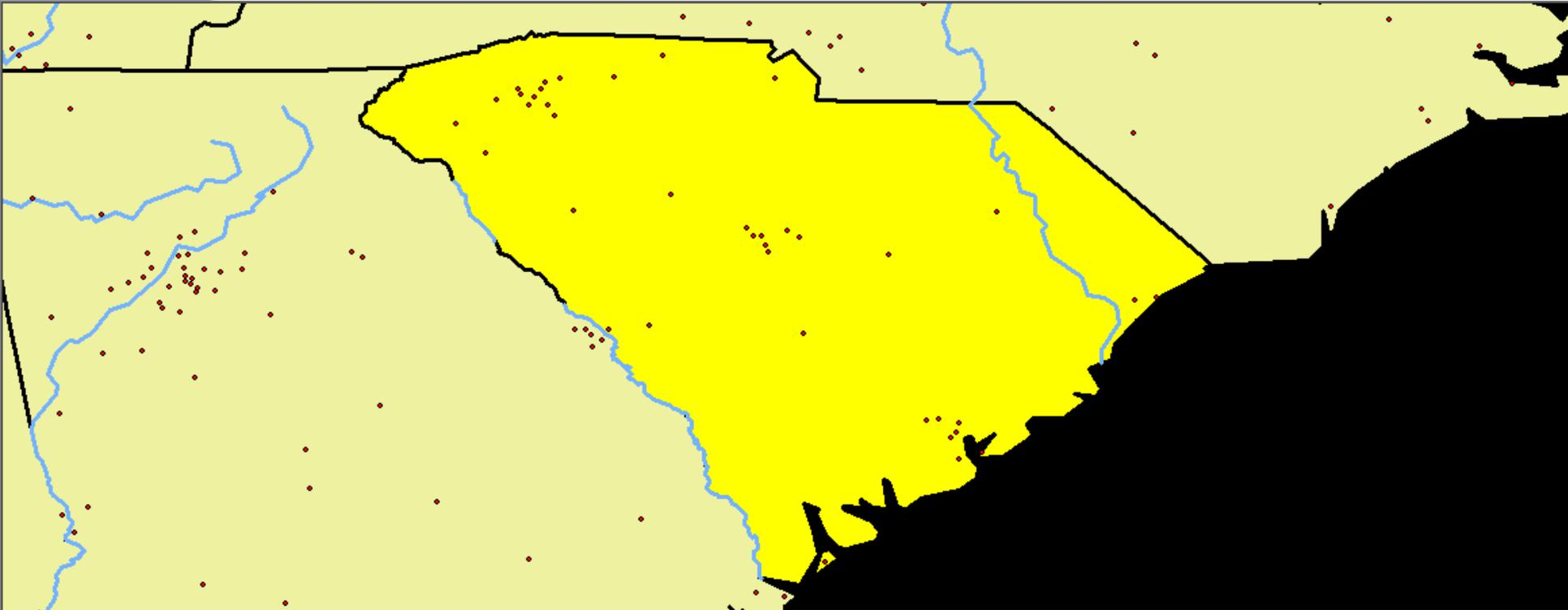
Shapefiles representing US cities, lakes, and rivers

Select by Location: Example



Select cities within 10 km of a lake using “Select by Location”

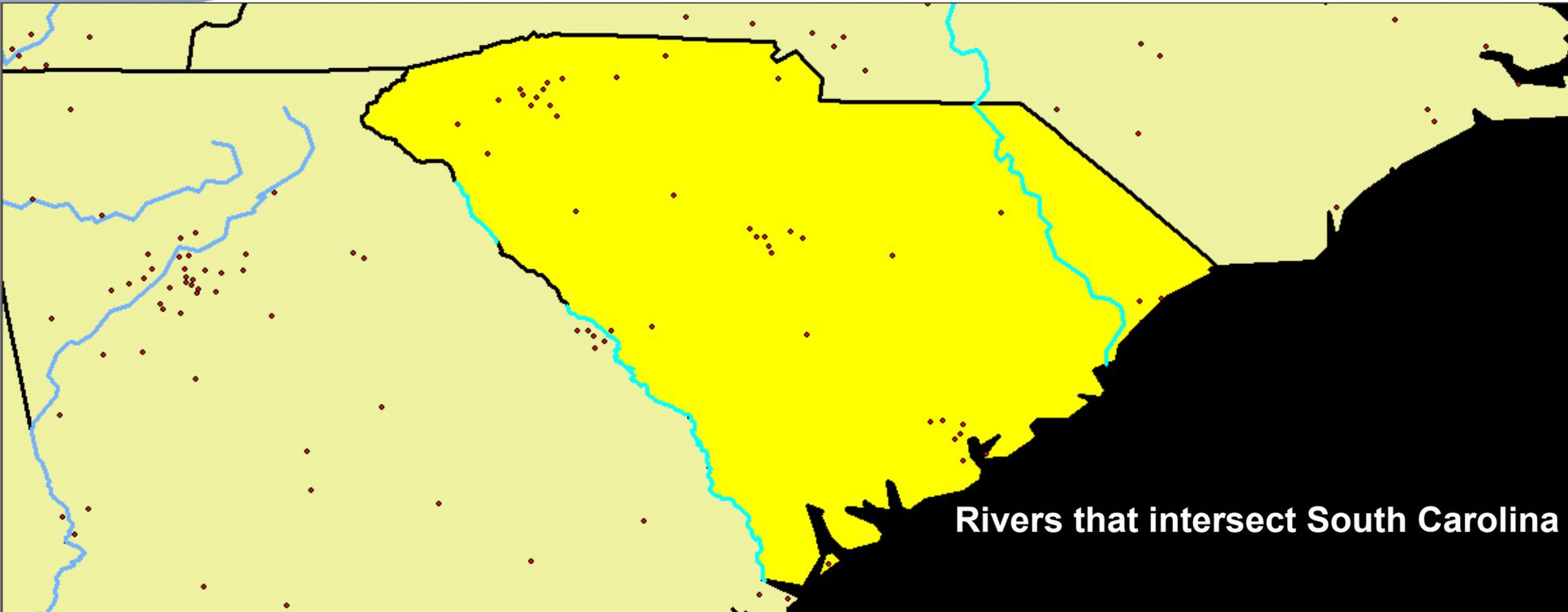
Select by Location: Example



1). Intersect: Polyline or Polygon features intersect on their edges.

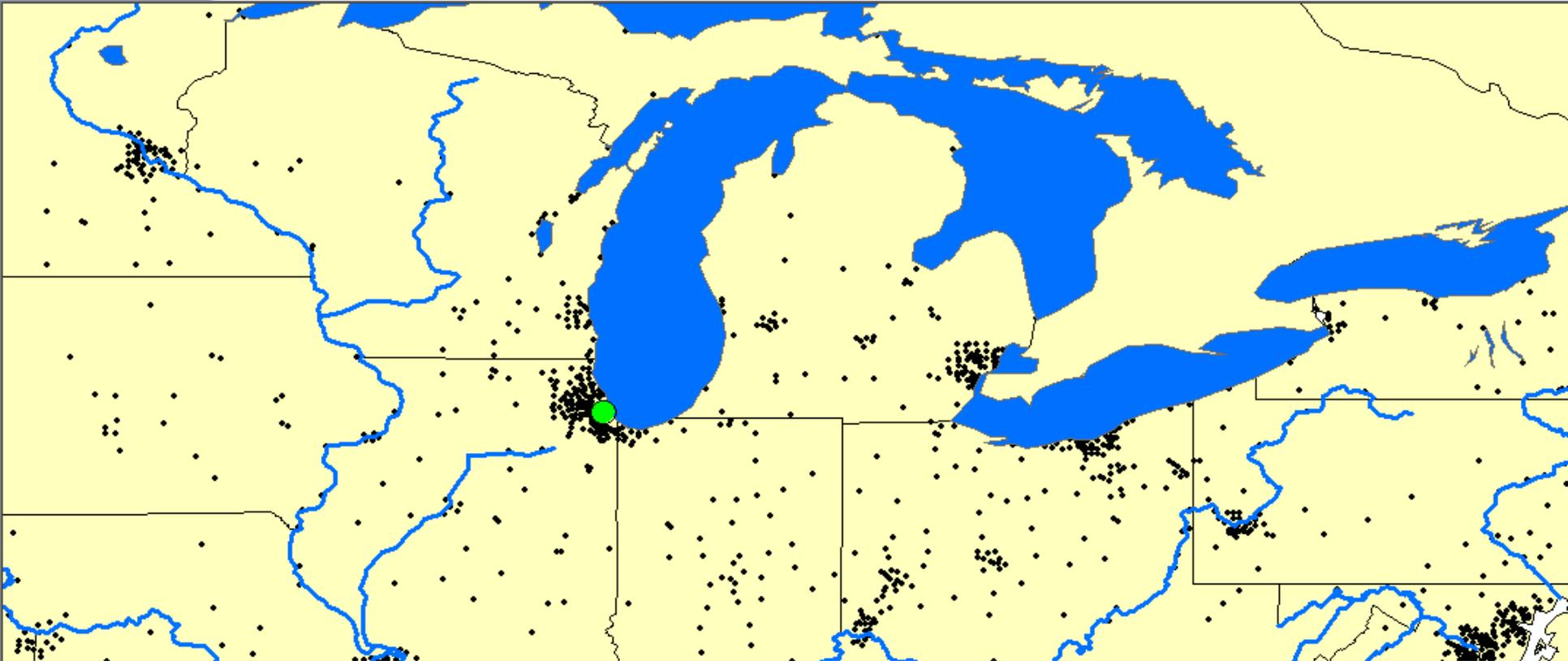
Select by Location: Other Types

CENTER for
DIGITAL
SCHOLARSHIP



1). Intersect: Polyline or Polygon features intersect on their edges.

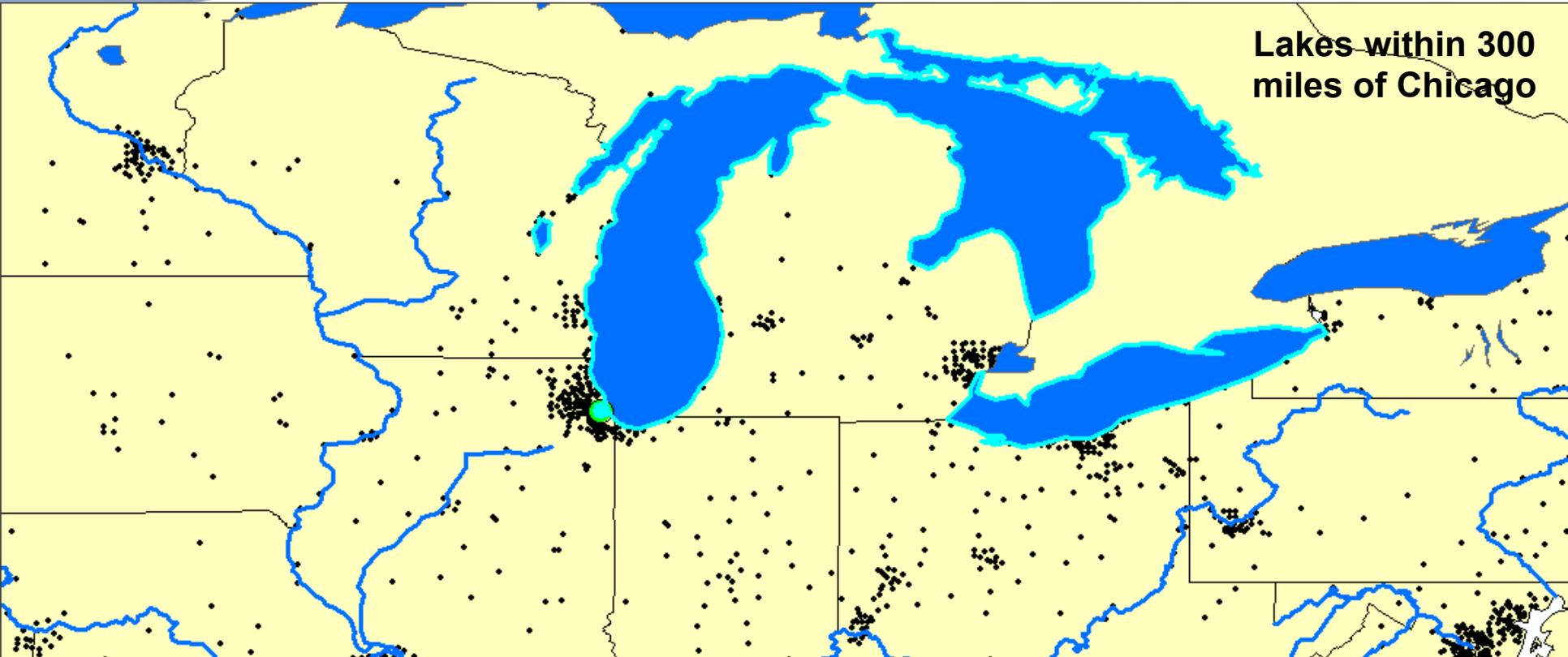
Select by Location: Other Types



2). Within Distance: Any features within given distance from edges

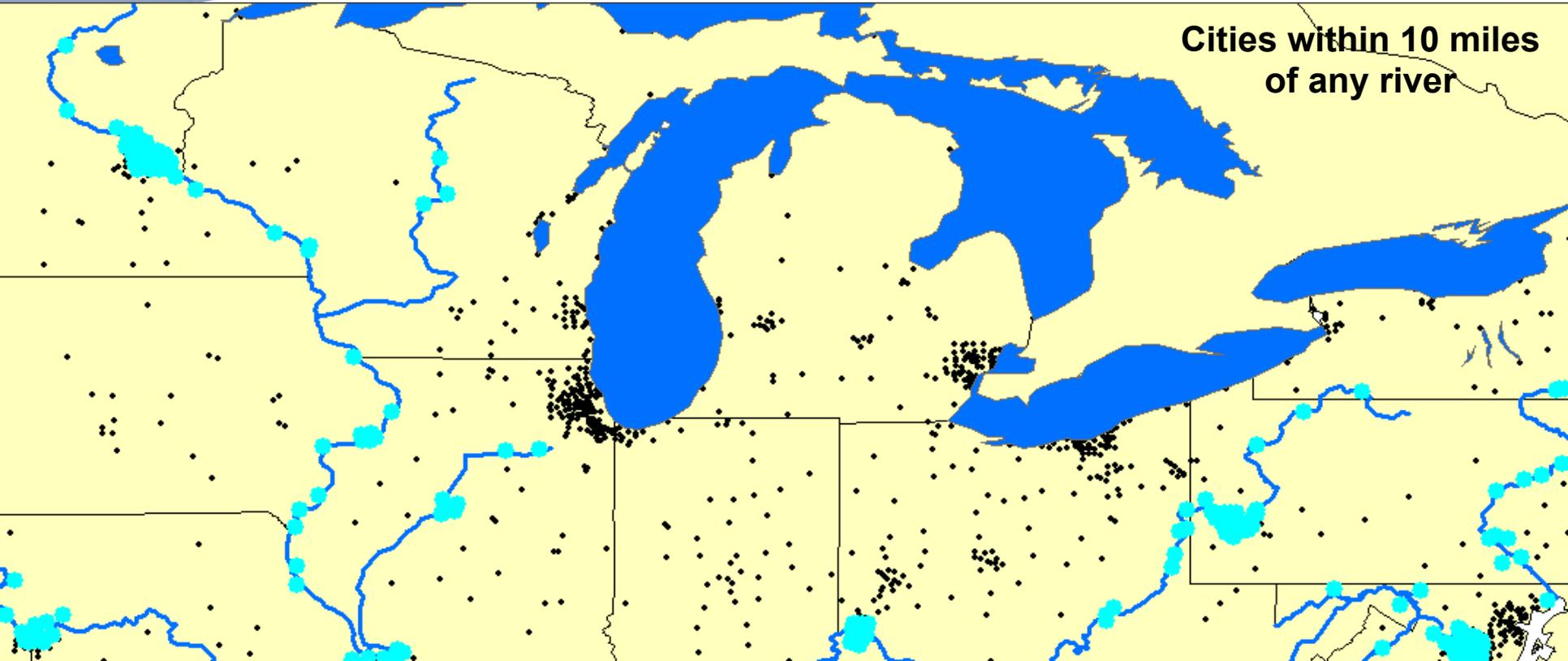
Select by Location: Other Types

CENTER for
DIGITAL
SCHOLARSHIP

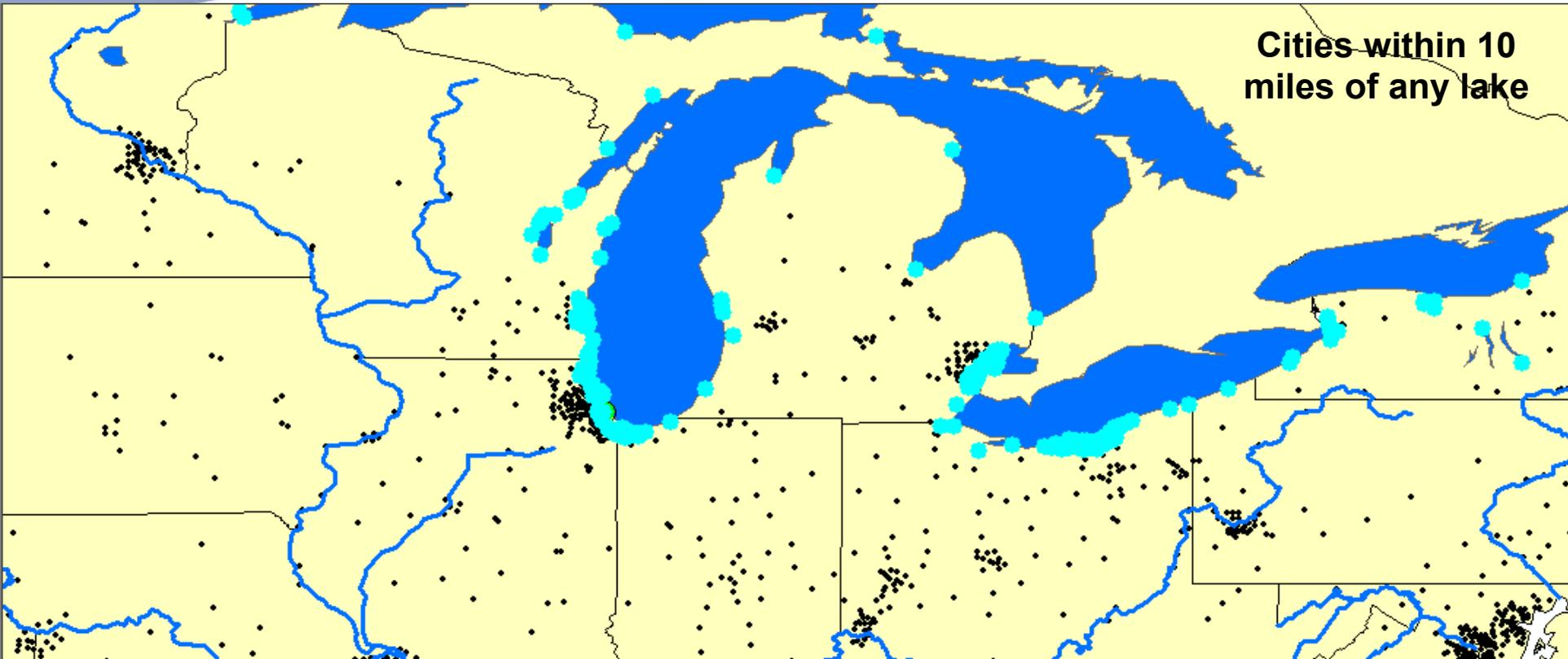


Select by Location: Other Types

CENTER for
DIGITAL
SCHOLARSHIP

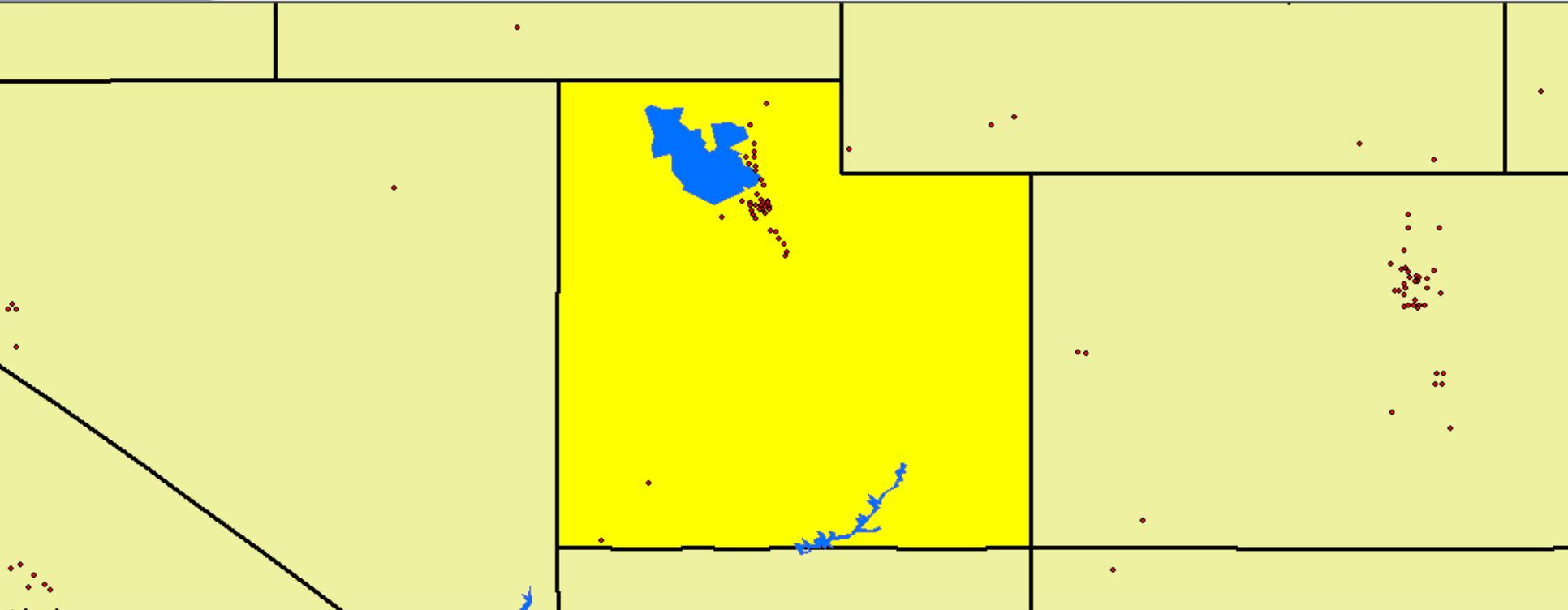


Select by Location: Other Types



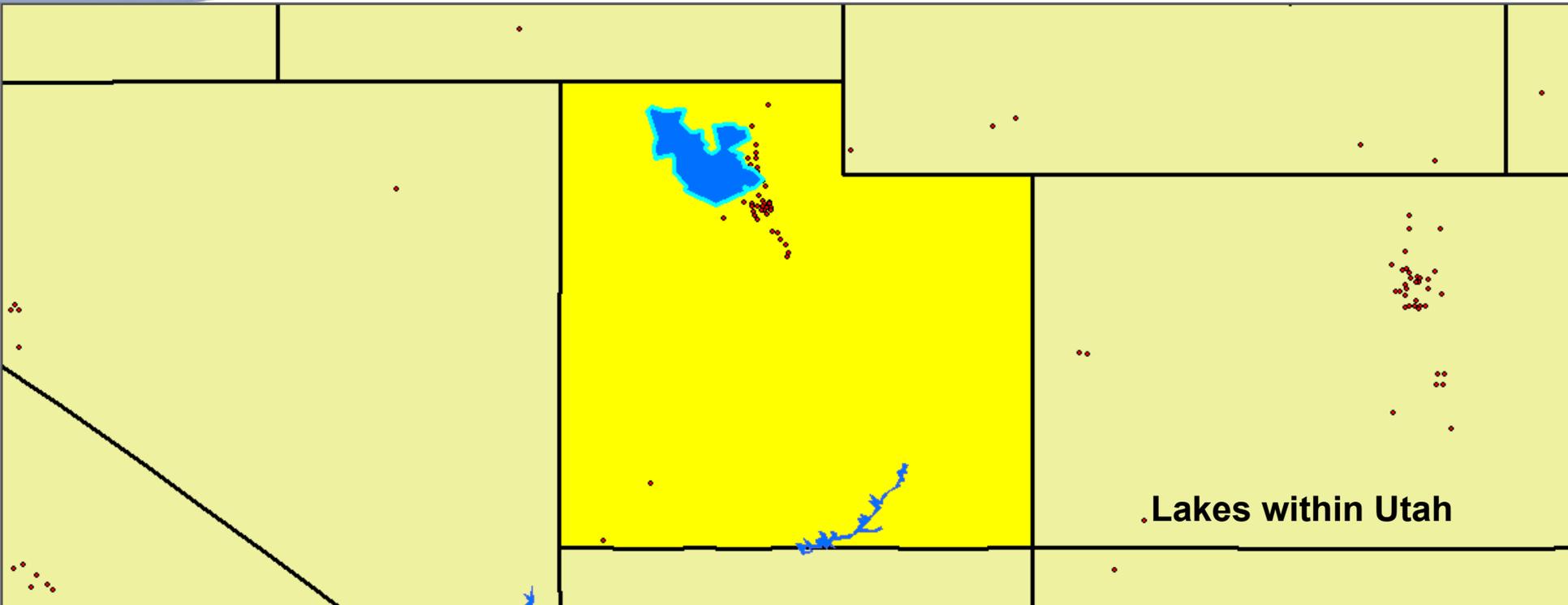
2). Within Distance: Any features within given distance from POLYGON

Select by Location: Other Types



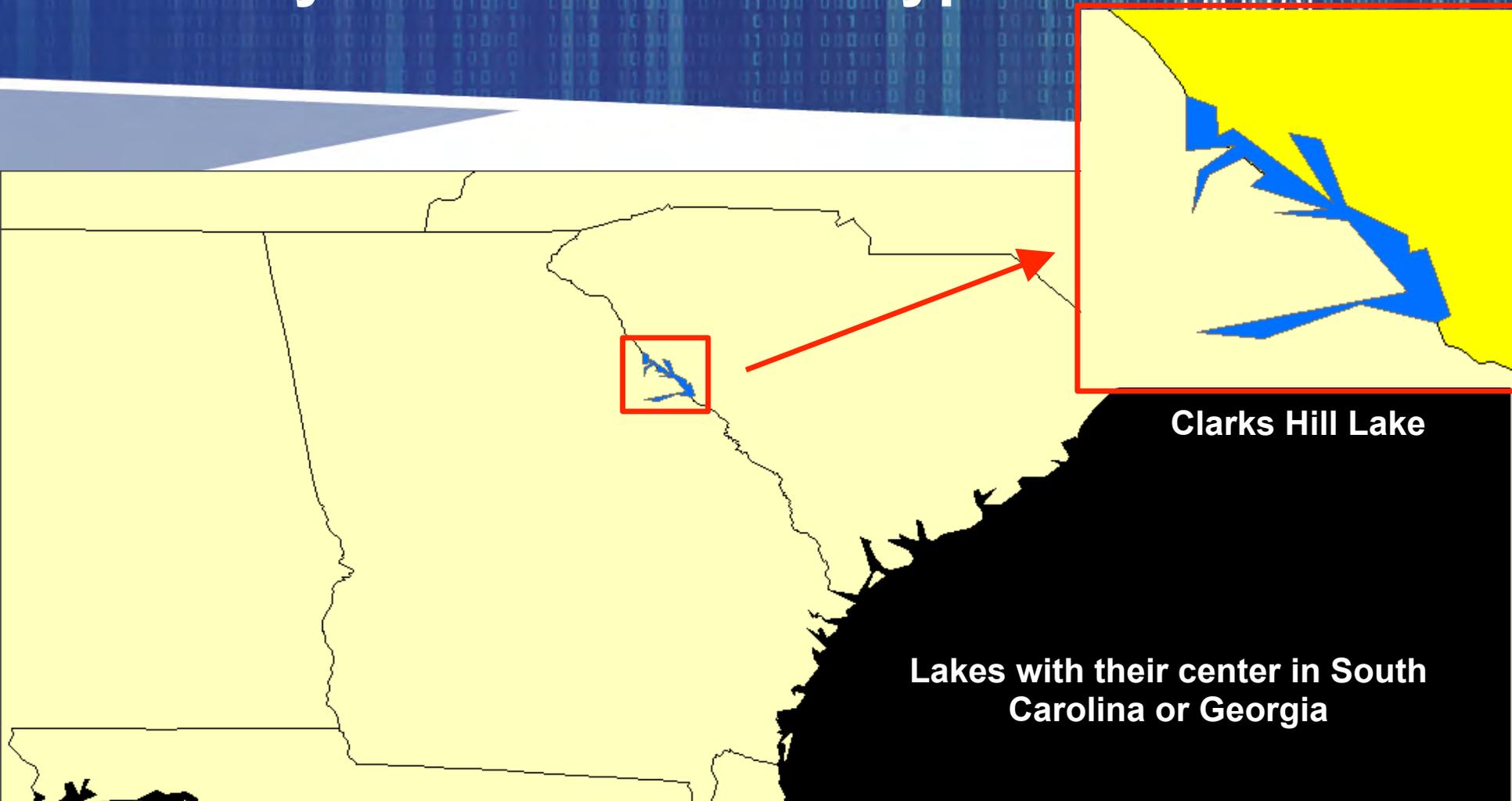
3). Completely Within: Any features completely inside polygon

Select by Location: Other Types



3). Completely Within: Any features completely inside polygon

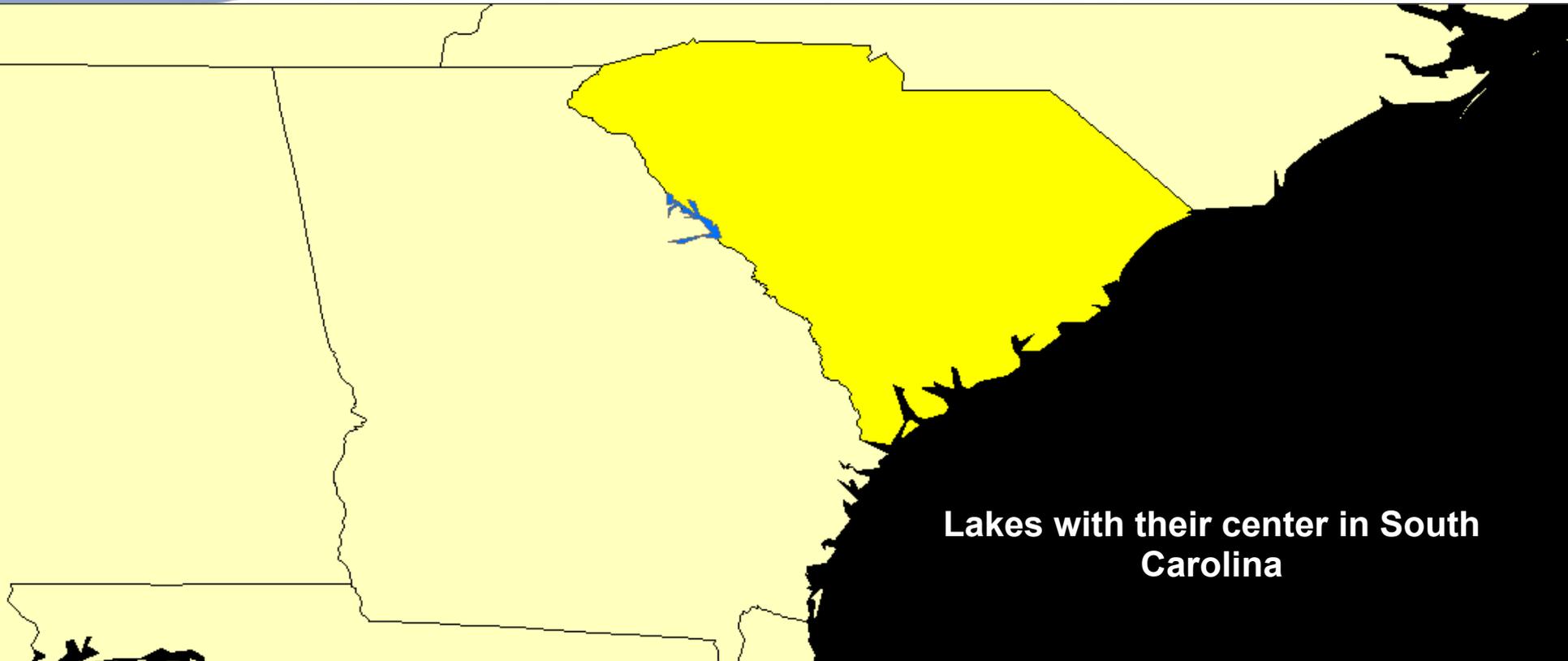
Select by Location: Other Types



5). Have Center Within: Features with a centroid inside polygon

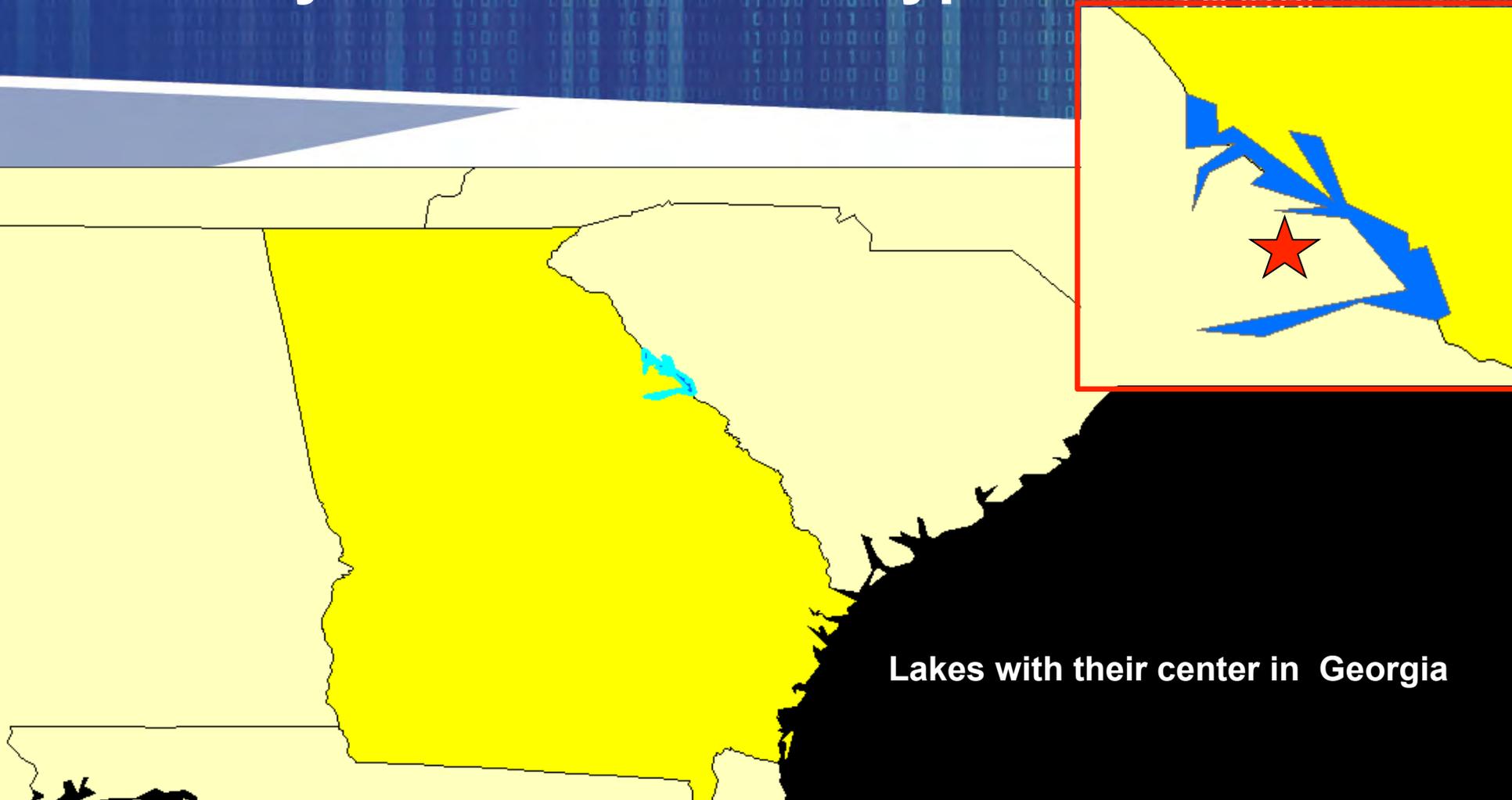
Select by Location: Other Types

CENTER for
DIGITAL
SCHOLARSHIP



5). Have Center Within: Features with a centroid inside polygon

Select by Location: Other Types



Lakes with their center in Georgia

5). Have Center Within: Features with a centroid inside polygon