Geospatial Analytics

NRCAAO

May 22, 2024

Patrick Alesandrini, CAE, MAI, RES, SRA, CFE

Geospatial Analytics Defined

- Geospatial analytics is a form of computational analysis that utilizes geographic information, spatial data, location data, and increasingly, high-resolution imagery, computer vision, and other forms of modeling and AI to extract structured data that can be used for specific applications and industries.
- Geospatial analytics is used to add timing and location to traditional types of data and to build data visualizations. These visualizations can include maps, graphs, statistics and cartograms that show historical changes and current shifts. (IBM)

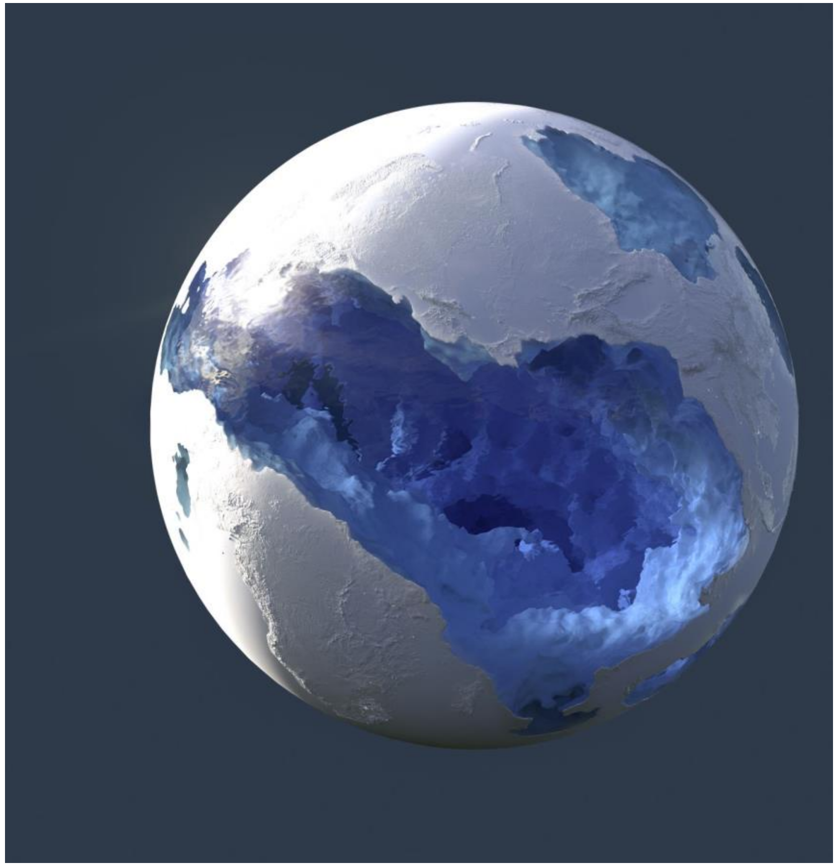
Geographic Data

Geographic Data:

Geographic(al) means 'pertaining to geography (the study of the surface of the earth)’

'referring to or characteristic of a certain locality, especially in reference to its location in relation to other places' (Macquarie Dictionary).

- Geographic data uses different feature types (raster, points, lines, or polygons) to uniquely identify the location and/or the geographical boundaries of spatial (location based) entities that exist on the earth surface.
- Geographic data are a significant subset of spatial data



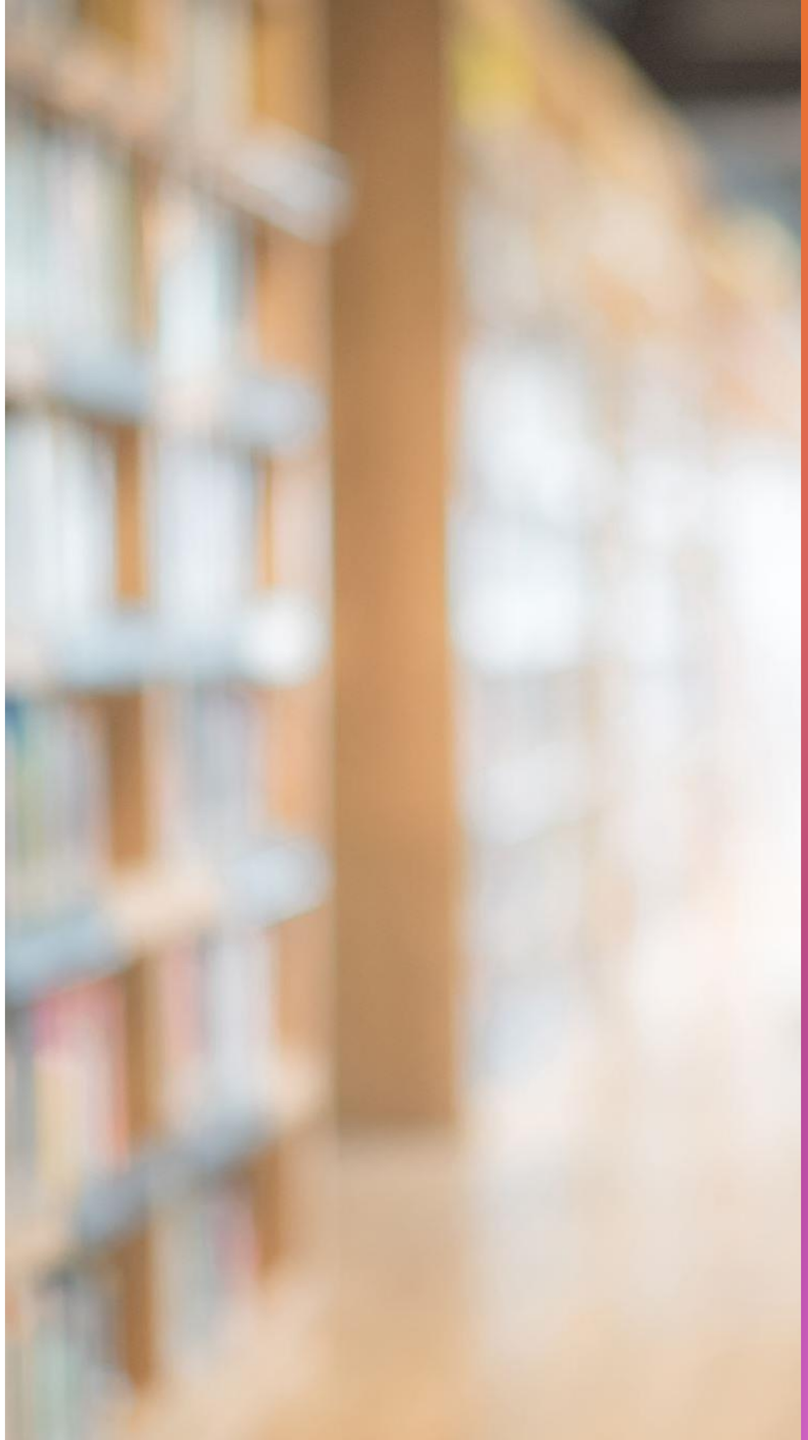
Spatial Data

- The word spatial originated from Latin 'spatium', which means space. Spatial means 'pertaining to space' or 'having to do with space, relating to space and the position, size, shape, etc.' (Oxford Dictionary),
- refers to features or phenomena distributed in three-dimensional space (any space, not only the Earth's surface) and, thus, having physical, measurable dimensions.
- In GIS, 'spatial' is also referred to as 'based on location on map'.
- The terms geographic, spatial, and geospatial are often used interchangeably, although technically incorrect.

Spatial Data

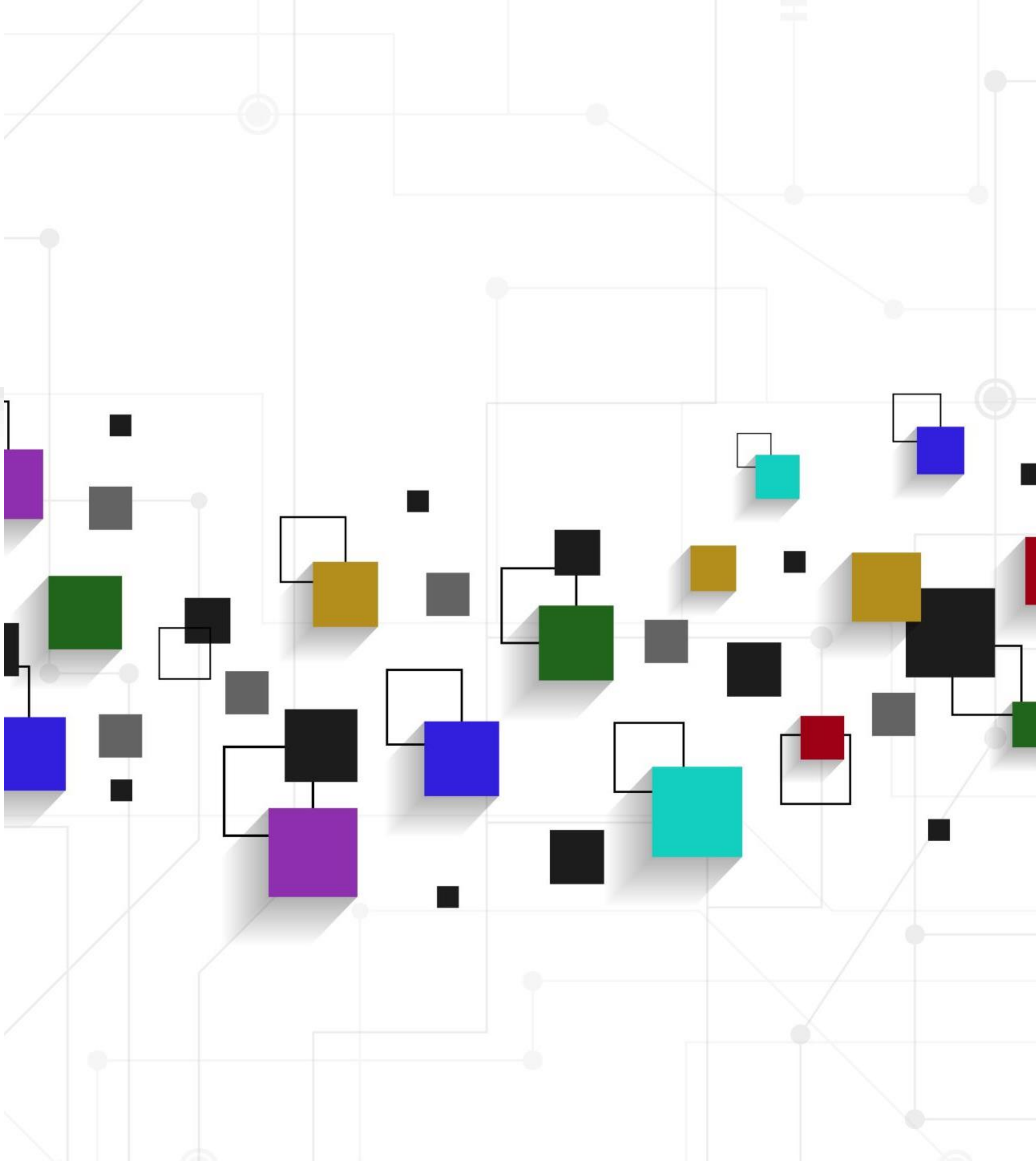
- Attribute domains are rules that describe the available values of a field type.
    - They are used to constrain the values allowed in any attribute for a table or feature class.
    - They provide a method for enforcing data integrity by limiting what can be placed on a field to a valid list or range of choices.

- Three types of spatial data are distinguished through the characteristics of the domain, namely, areal (or lattice) data, geostatistical data, and point patterns



Spatial Data (continued)

- Areal data
    - Areal data usually arise when the number of events corresponding to some variable of interest are aggregated in areas such as neighborhoods or counties.



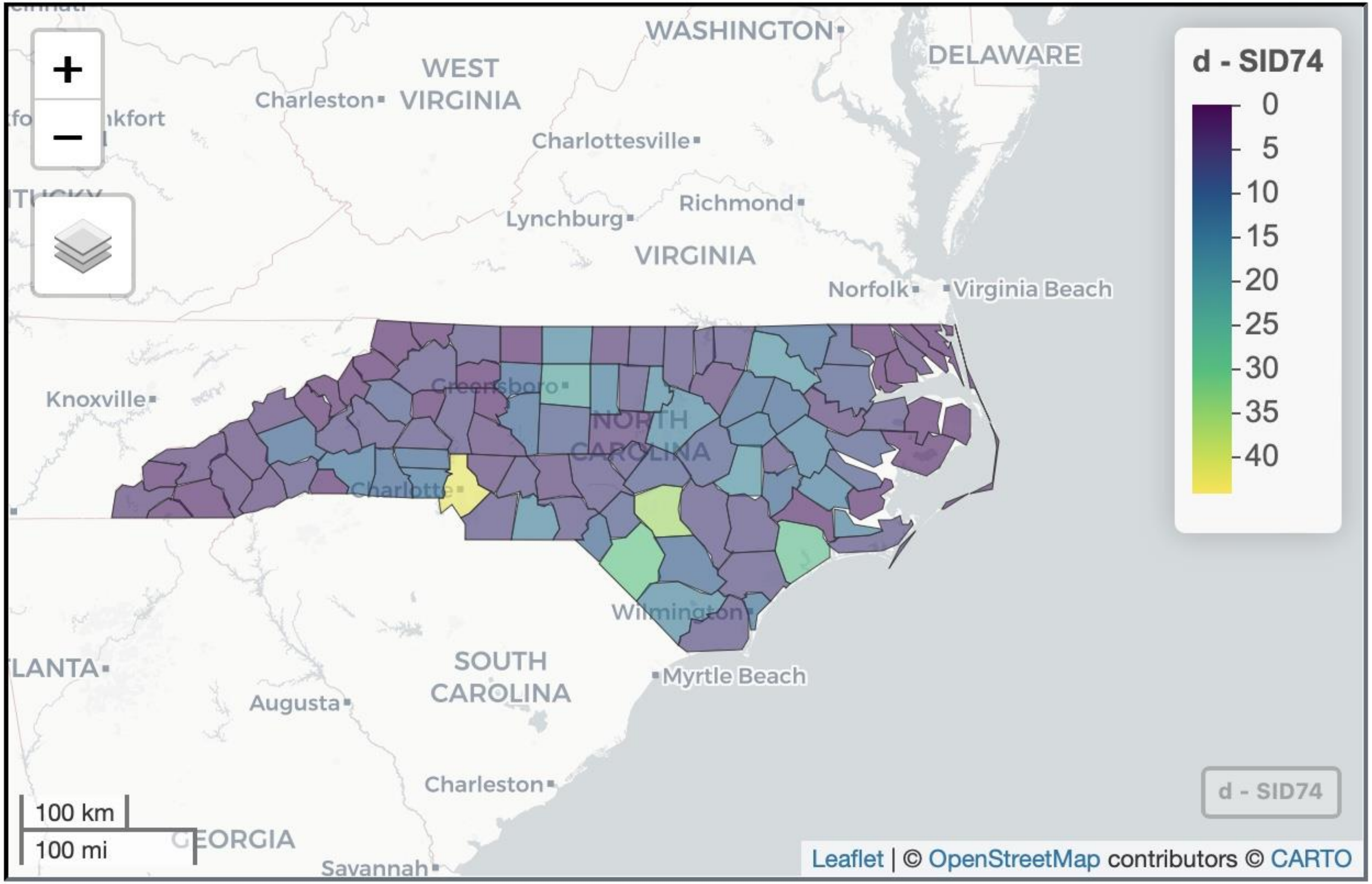


FIGURE 1.1: Example of areal data. Number of sudden infant deaths in counties of North Carolina, USA, in 1974.

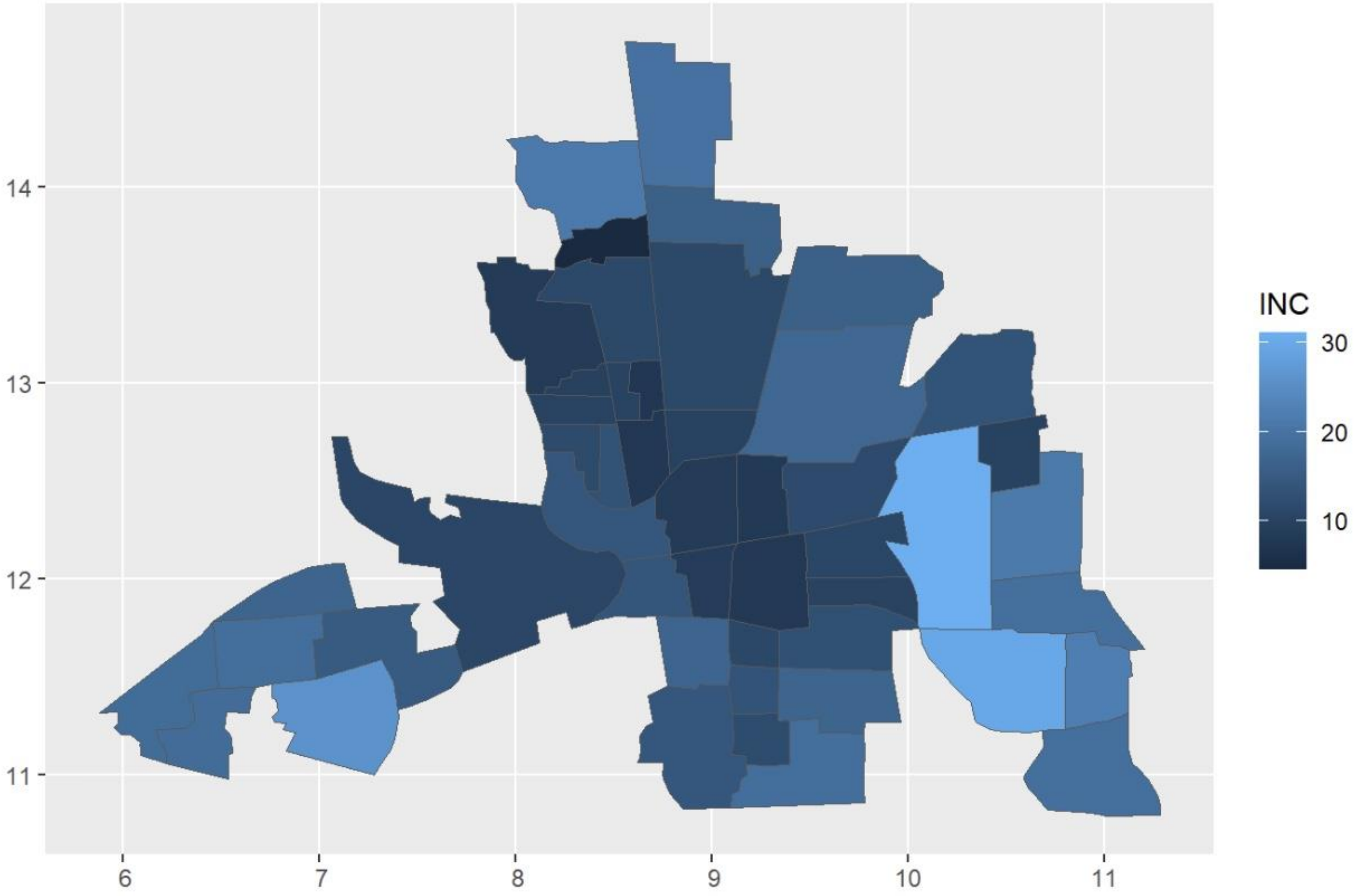


FIGURE 1.2: Example of areal data. Household income in $1000 USD in neighborhoods in Columbus, Ohio, in 1980.

Spatial Data (continued)

- Areal data
  - Areal data usually arise when the number of events corresponding to some variable of interest are aggregated in areas such as neighborhoods or counties.
- Geostatistical data
  - The use of data observed at known spatial locations to predict the values of the variable of interest at unsampled locations.
  - For example, we can use air pollution measurements at a number of monitoring stations to predict air pollution at other locations taking into account spatial autocorrelation and other factors that are known to predict the outcome of interest

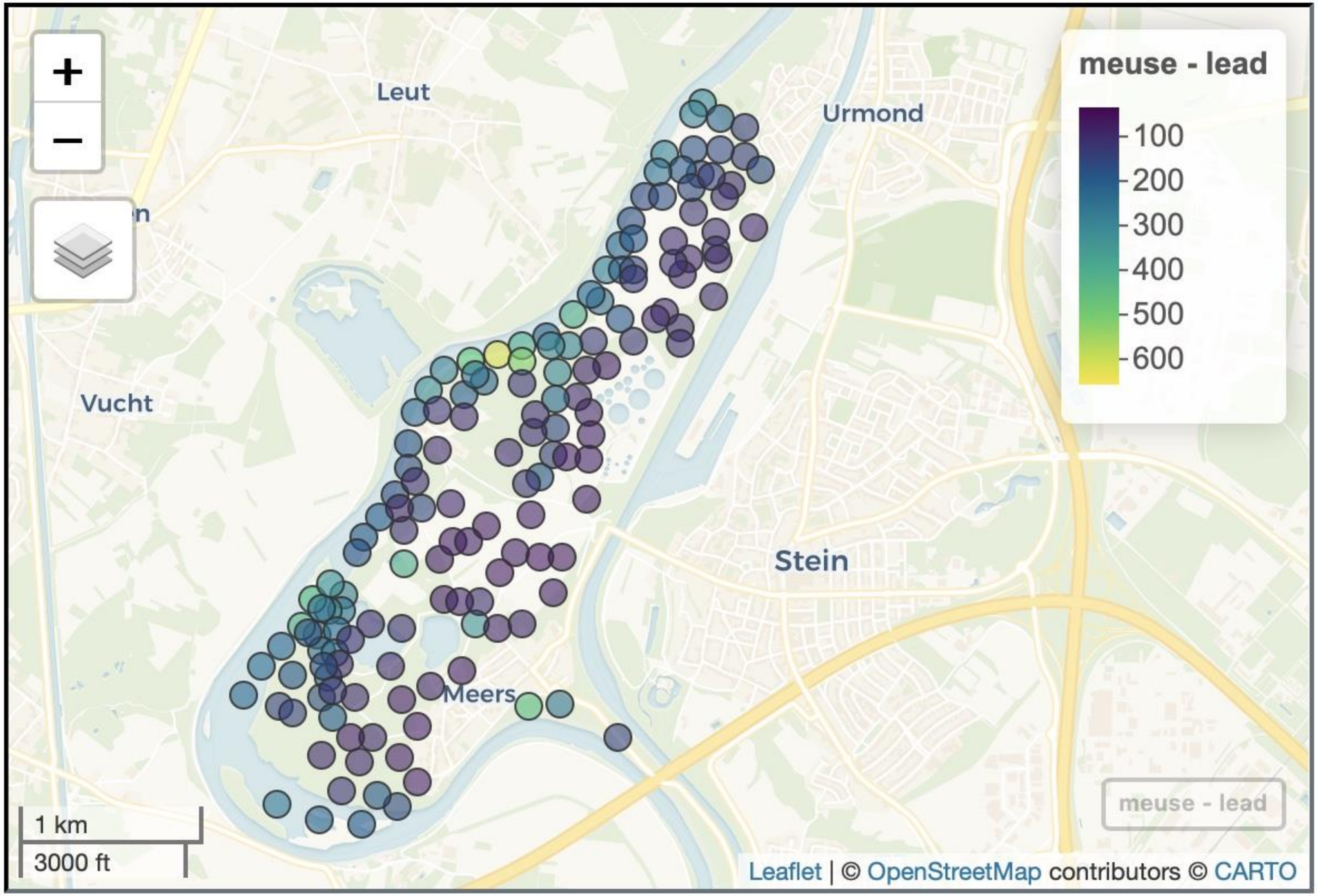


FIGURE 1.4: Example of geostatistical data. Topsoil lead concentrations at locations sampled in a flood plain of the river Meuse, The Netherlands.

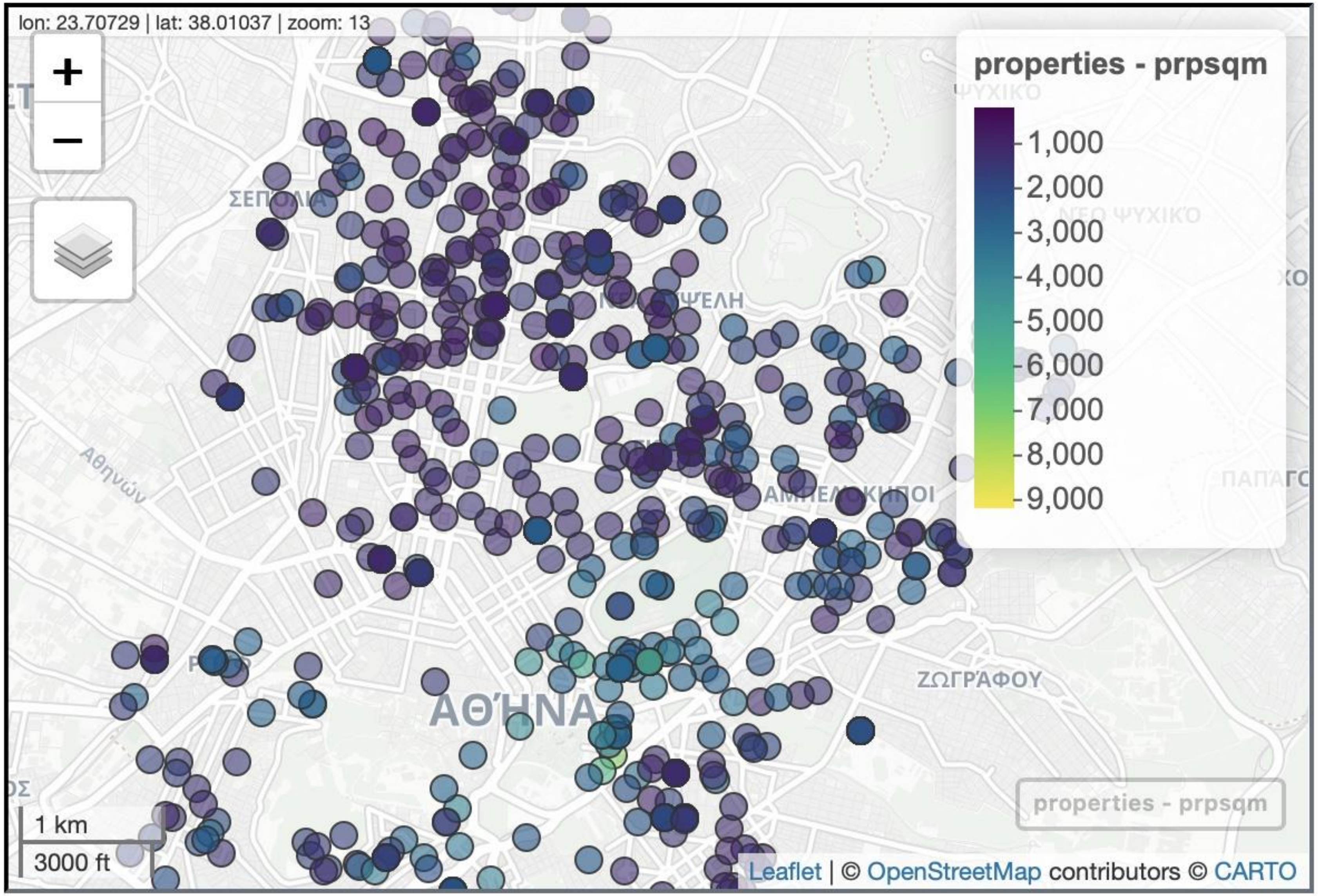


FIGURE 1.5: Example of geostatistical data. Price per square meter of a set of apartments in Athens, Greece, in 2017.

Spatial Data (continued)

- Areal data
  - Areal data usually arise when the number of events corresponding to some variable of interest are aggregated in areas such as neighborhoods or counties
- Geostatistical data
  - The use of data observed at known spatial locations to predict the values of the variable of interest at unsampled locations
  - For example, we can use air pollution measurements at a number of monitoring stations to predict air pollution at other locations taking into account spatial autocorrelation and other factors that are known to predict the outcome of interest
- Point patterns
  - In point patterns, an observation index set gives the locations of random events of the spatial point pattern indicating occurrence of the event

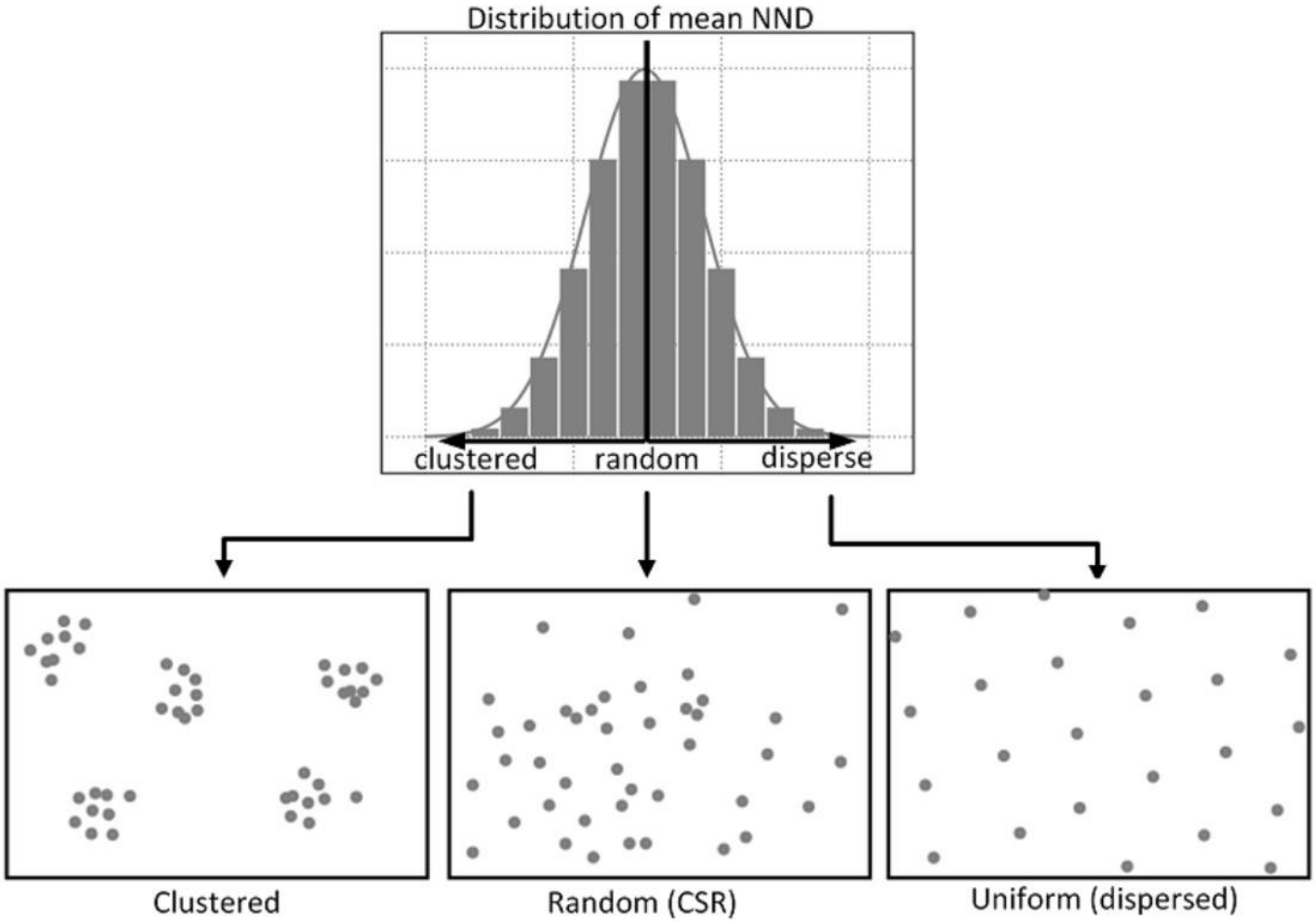


Figure 3: Relations between different point patterns and mean nearest neighbor distance (NND).

Spatial Data (Continued)

- Spatio-temporal data
  - Spatio-temporal data arise when information is both spatially and temporally referenced.
    - Temporal is data that represents an observation or pattern over time
    - These past patterns are used to predict future events

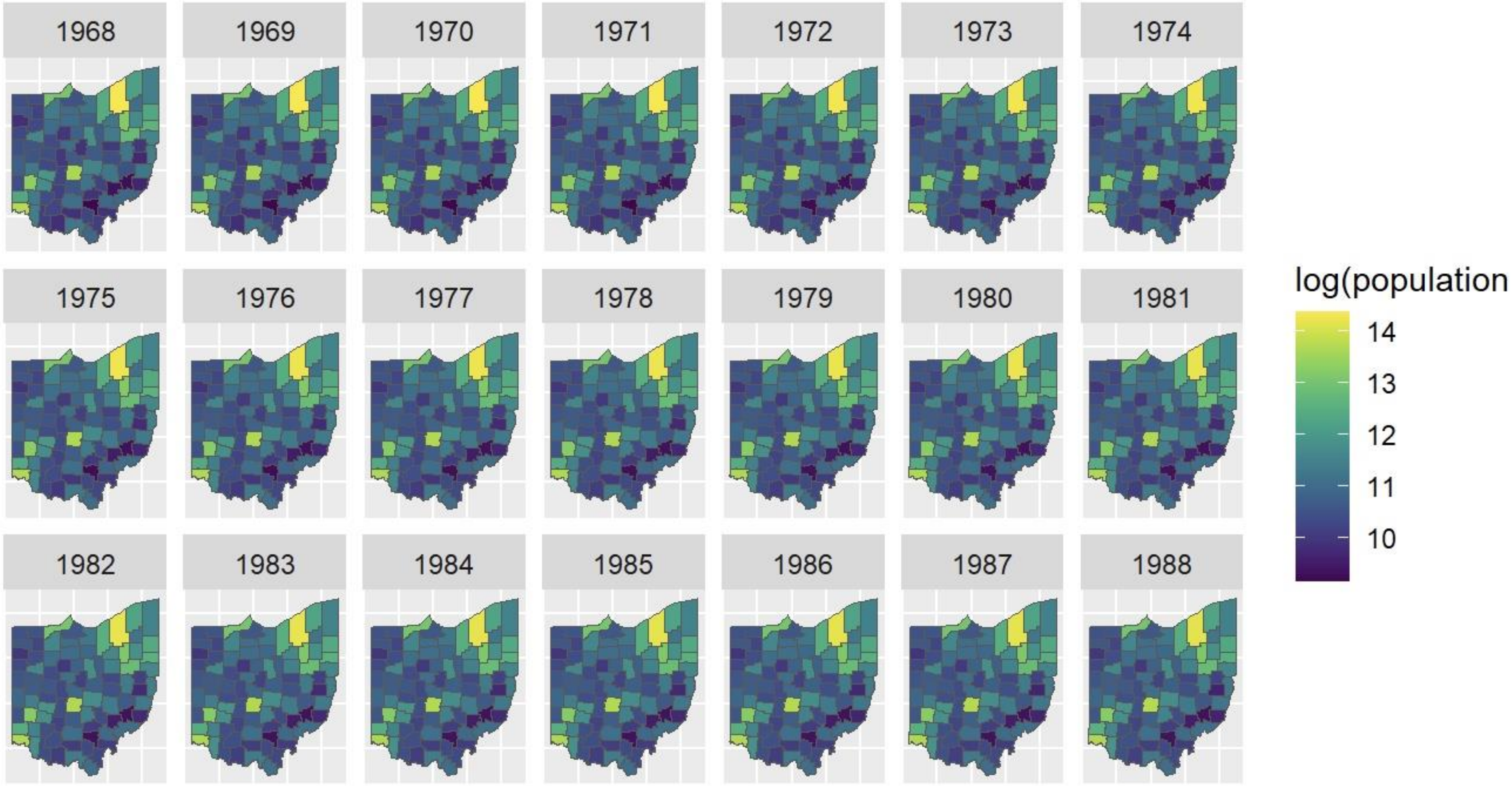
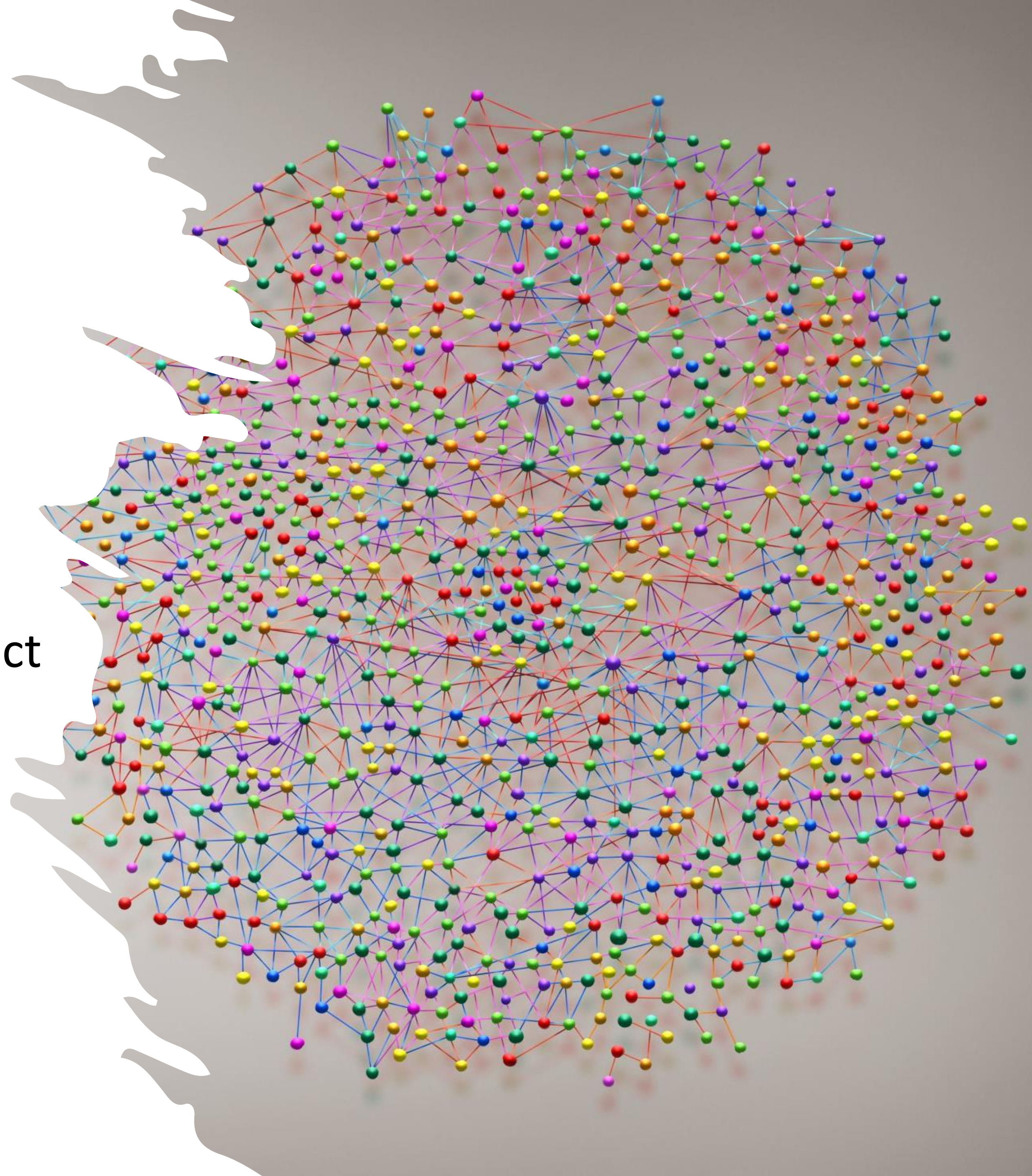


FIGURE 1.9: Example of spatio-temporal data. Population of the counties of Ohio, USA, from 1968 to 1988.

Spatial Data (Continued)

- Spatio-temporal data
  - Spatio-temporal data arise when information is both spatially and temporally referenced.
    - Temporal is data that represents an observation or pattern over time
    - These past patterns are used to predict future events
- Spatial functional data
  - Spatial functional data arise when the three types of spatial data (areal, geostatistical, and point patterns) are combined with random functions.



Geostatistical Data

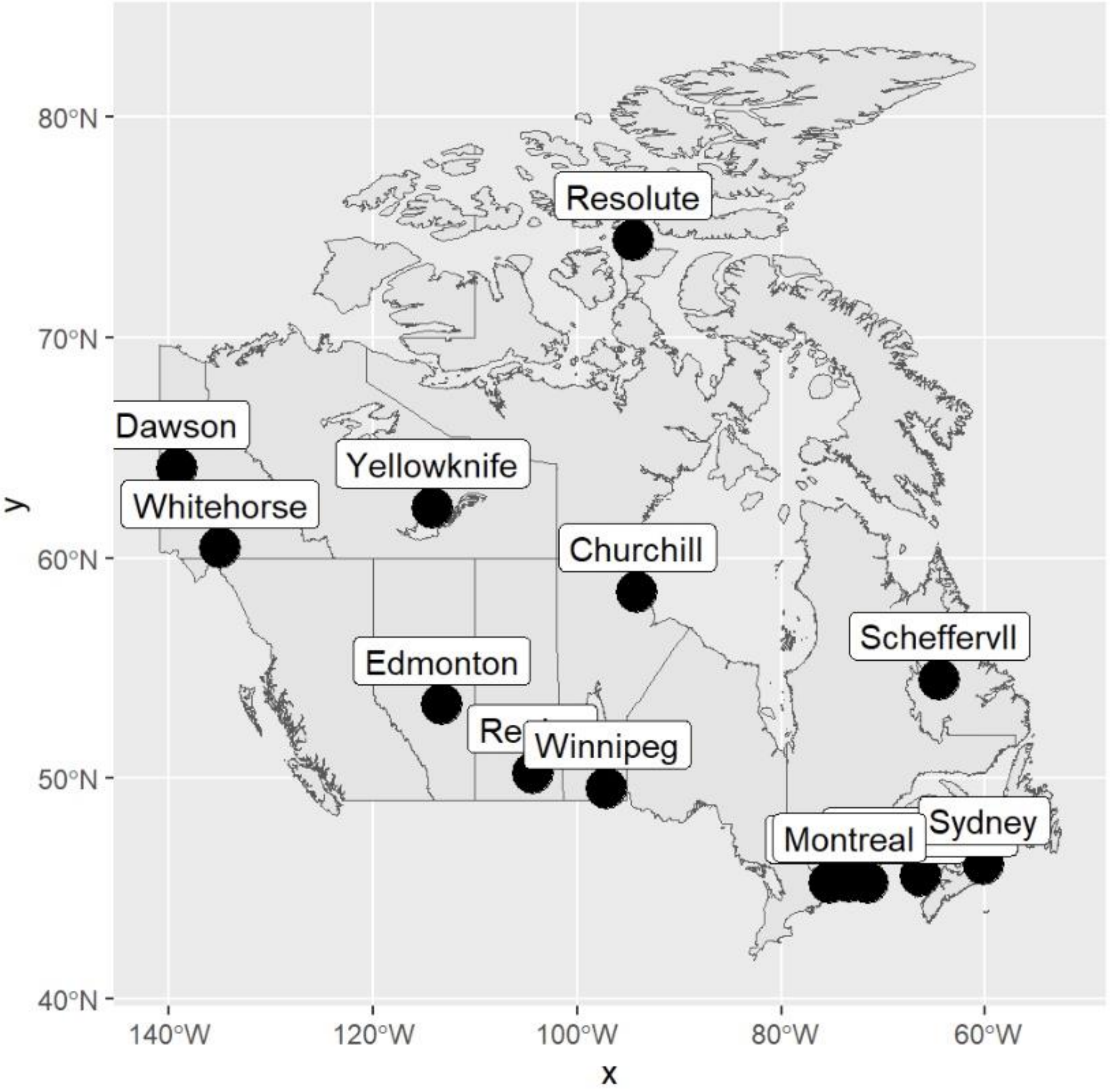


FIGURE 1.10: Locations of Canadian weather stations where daily temperature is measured.

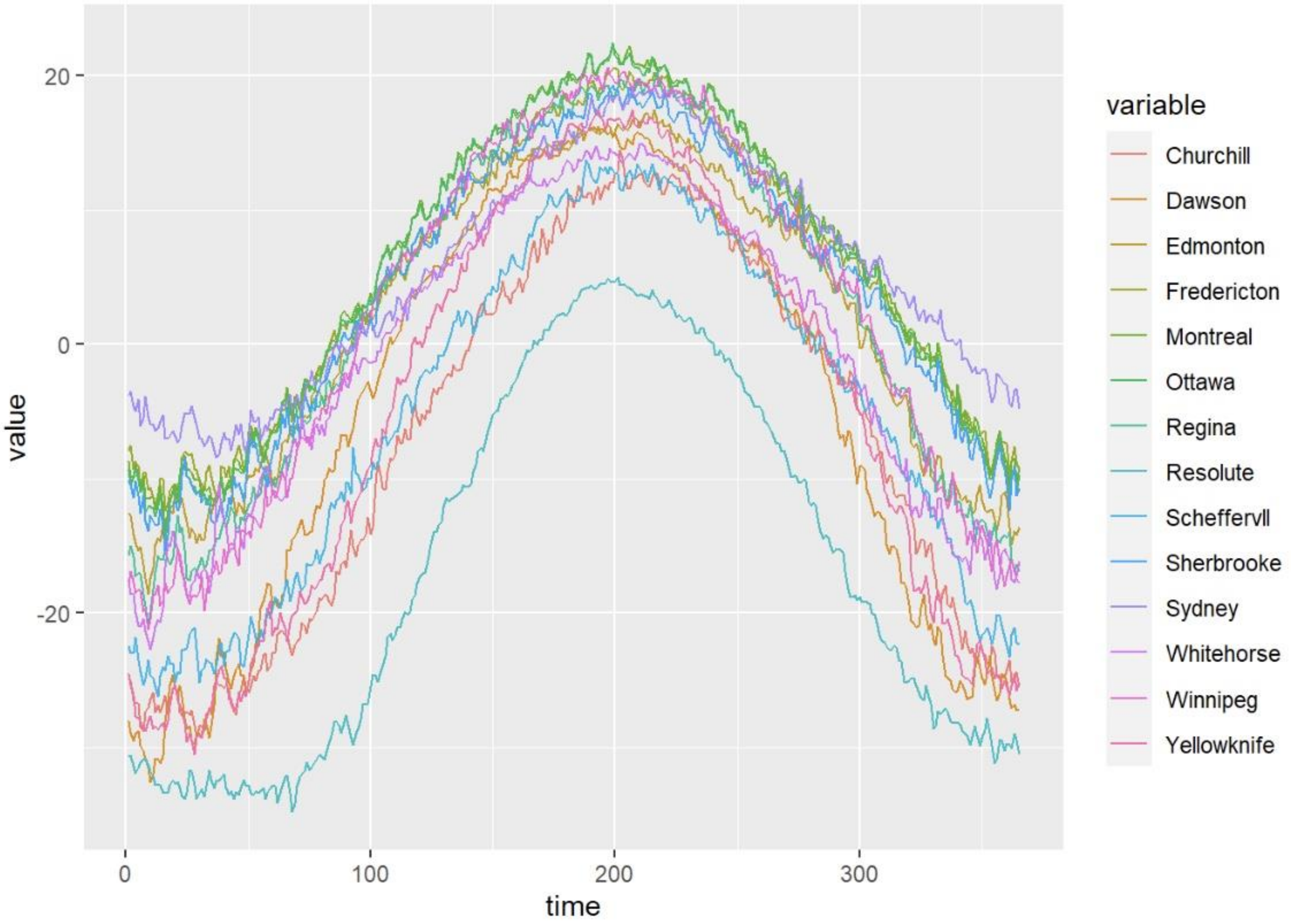


FIGURE 1.11: Example of spatial functional data. Daily temperature averaged over 30 years measured at 35 Canadian weather stations.

Geospatial

- Consisting of, derived from, or relating to data that is directly linked to specific geographical locations (Merriam-Webster)
- Geospatial data is used to develop information about features, objects, and classes on Earth's surface and/or near Earth's surface.
- Many organizations, including agencies of the United States consider Geospatial and Spatial as equivalent terms.

Cartogram: A cartogram is a transformation of a map that uses some variable instead of land area to expand or contract the area of the original polygons based on an attribute value.

Cartograms are often used for displaying population data.

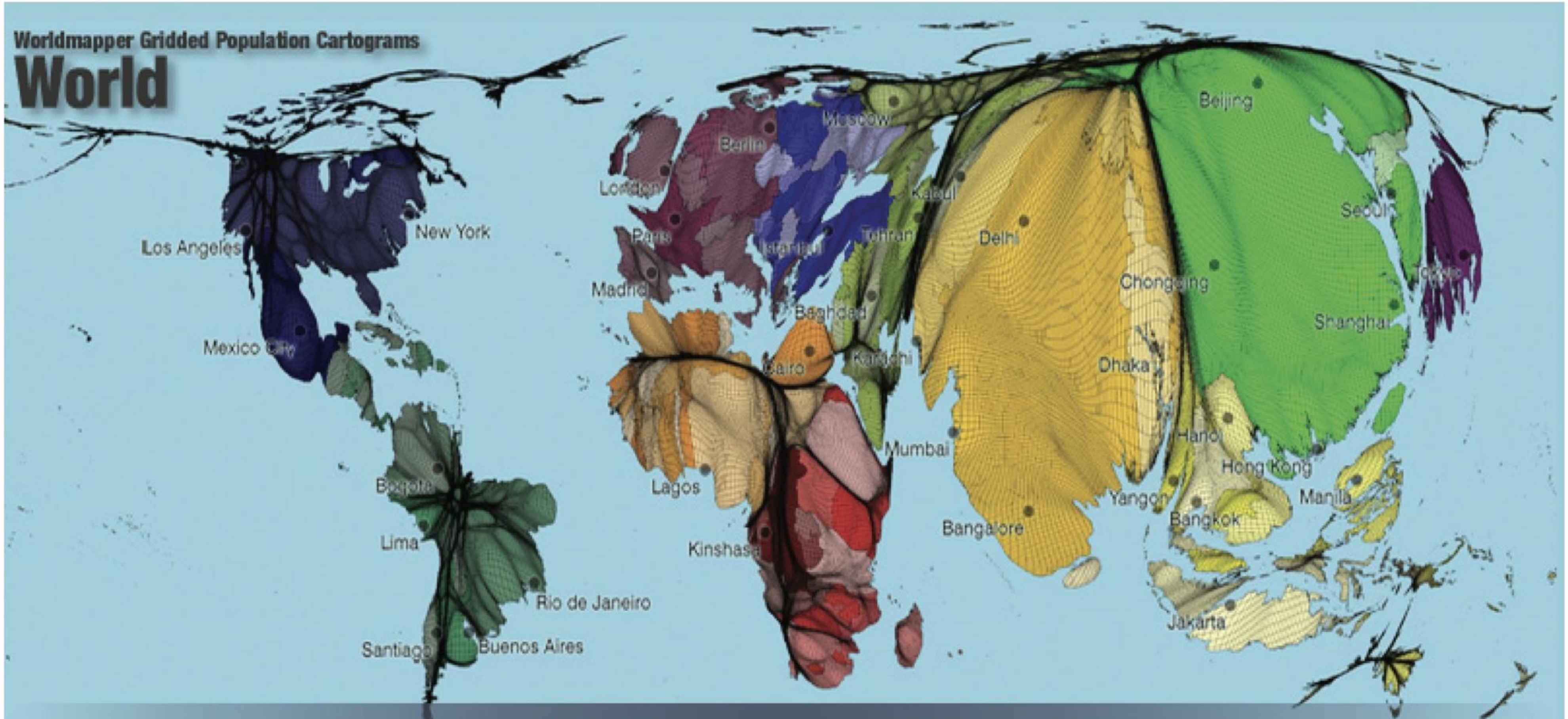
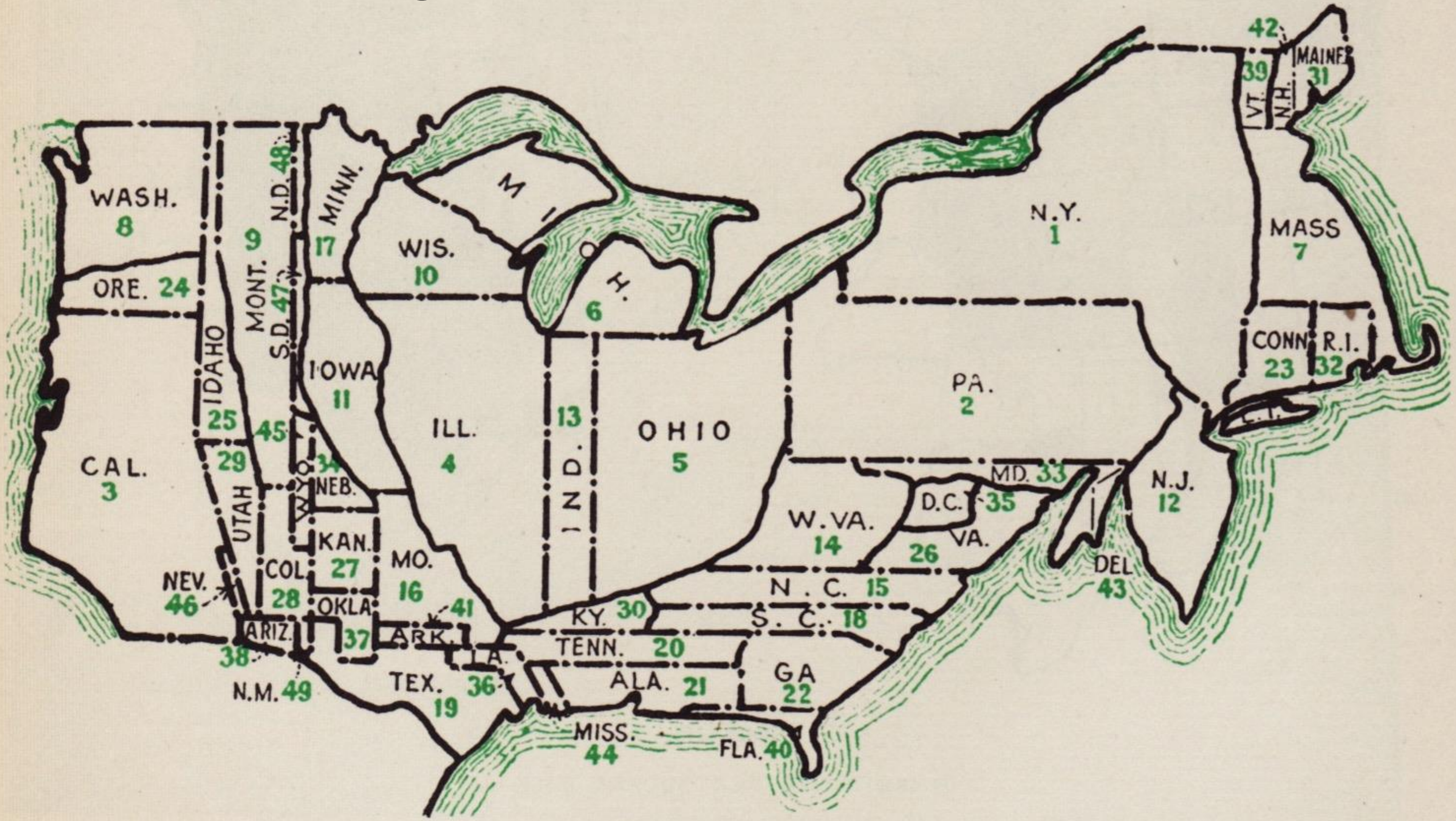


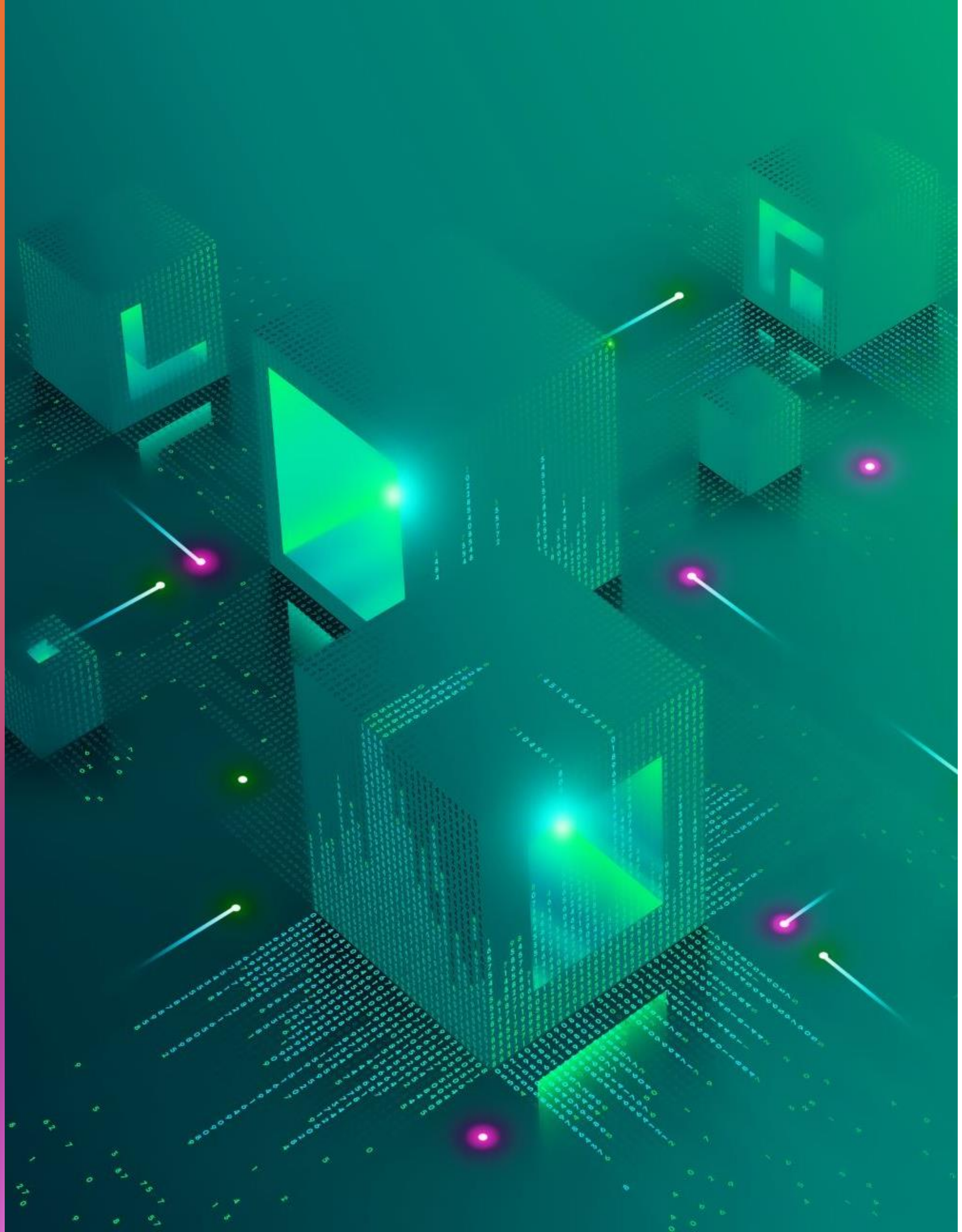
Figure 2: Grid-Based World Population Cartogram (2000)

Old School Cartogram



Literary Digest, April 23, 1921.

Relative Size of Each of the United States If Based on Electrical Energy Sold for Light and Power in 1921.

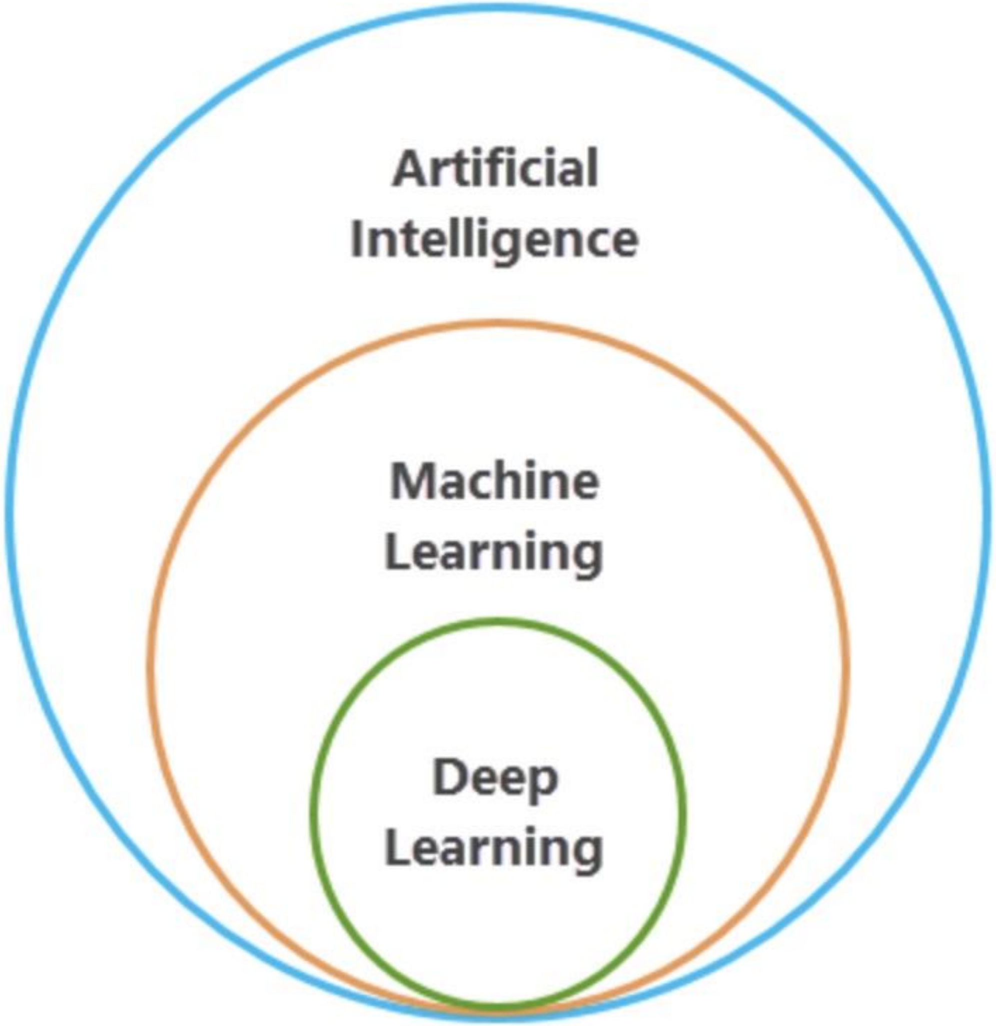


Computer vision

- Computer vision applications use input from sensing devices, artificial intelligence, machine learning, and deep learning to replicate the way the human vision system works.
  - Computer vision applications run on algorithms that are trained on massive amounts of visual data or images in the cloud.
  - They recognize patterns in this visual data and use those patterns to determine the content of other images. (Microsoft Azure)
  - A form of computer vision would be the change technology that is used in many offices today.

Machine learning is a branch of artificial intelligence in which structured data is processed with an algorithm to solve a problem.

Traditional structured data requires a person to label the data, such as a pictures of cats and dogs, so that specific features for each animal type can be understood within the algorithm and used to identify these animals in other pictures.



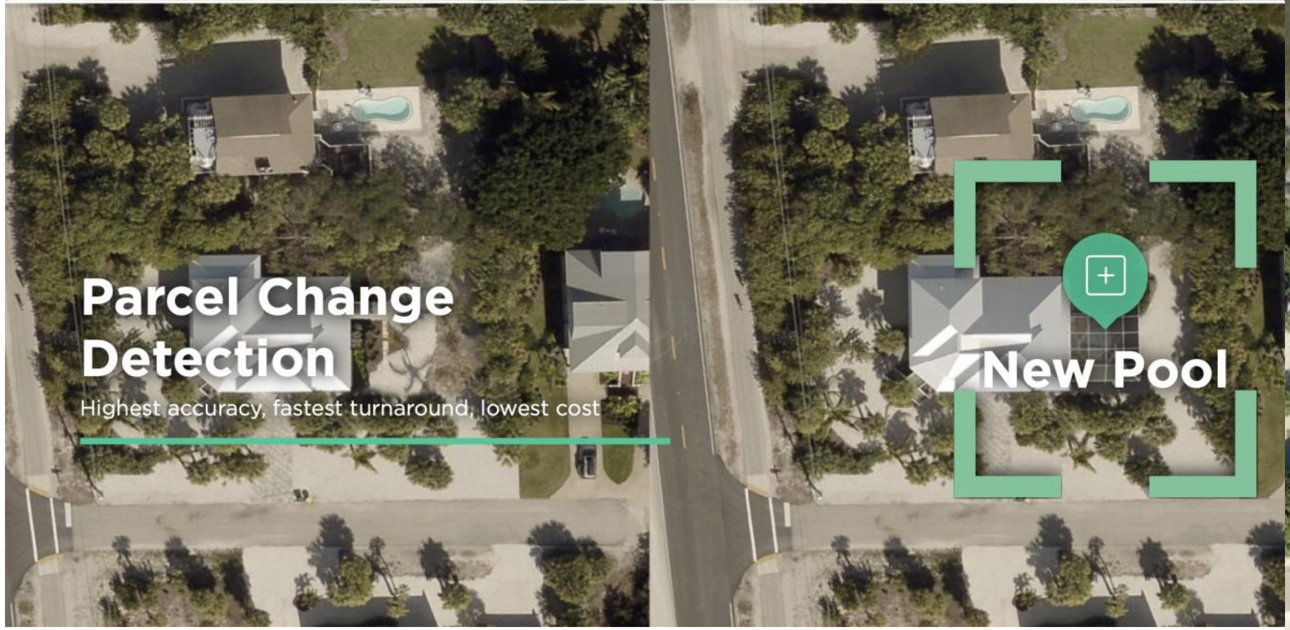
Deep learning is a subset of machine learning that uses several layers of algorithms in the form of neural networks.

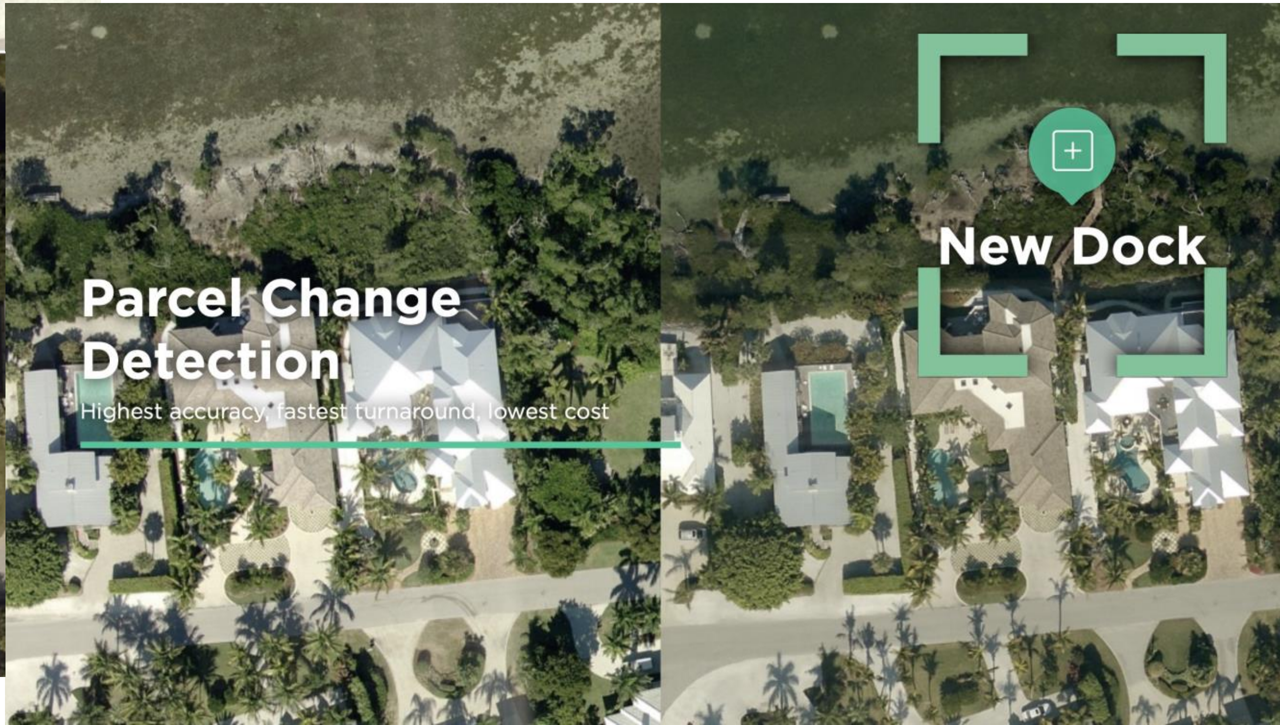
Input data is analyzed through different layers of the network, with each layer defining specific features and patterns in the data.

For example, if you want to identify features such as buildings and roads, the deep learning model can be trained with images of different buildings and roads, processing the images through layers within the neural network, and then finding the identifiers required to classify a building or road.

Change Technology - PushPin


Parcel Change Detection
Highest accuracy, fastest turnaround, lowest cost
Addition


Parcel Change Detection
Highest accuracy, fastest turnaround, lowest cost
New Pool


Parcel Change Detection
Highest accuracy, fastest turnaround, lowest cost
New Dock

Refining Neighborhoods with GIS & Statistics

In 2019, while working with an island neighborhood, the following statistical output were observed:

Median Ratio: 88%

PRD: 129

Mean Ratio: 72%

Weighted Mean: 55.81%

Staff wanted to put a positive multiplier on the entire neighborhood to correct the underassessment.

2024 Harbo
0 - 25
26 - 50
51 - 75
76 - 19

Refining Neighborhoods with GIS & Statistics

- Obviously my first concern was a PRD of 129!
- We created a map displaying the quartile of all sales by sales ratio with the sale price displayed
- Quickly it was found that the first quartile was underassessed and the fourth quartile was over assessed.
- We also noted that the culprit causing most of the trouble was one of nine mansions on the island that was significantly underassessed and it also greatly out valued the typical nice but less expensive condominiums
- The mansions were removed from the dataset and added to the adjacent island that is made up primarily of mansions

Refining Neighborhoods with GIS & Statistics

- From the initial quartile study, we found that the assessments not meeting IAAO Standards required a more refined analysis and so we moved from quartiles to percentiles.
- As will be seen on the following screen, the 30^{th} percentile captured those properties that are below the 90%.
- The group of properties below the 30^{th} percentile had an COD of 5.6471
- The median ratio for this group was 0.85
- By dividing the desired ratio of 100% by the median ratio of 0.85, a factor of 1.176471 is established.

COD of the 30th Percentile			Median	Absolute Dev	
03C001 04 00013	51	0.70	0.85	0.15	
03C001 04 00015	51	0.71	0.85	0.14	
03C001 04 00014	51	0.74	0.85	0.11	
03C001 04 00008	51	0.85	0.85	0	
03C001 04 00009	51	0.85 ***5th***	0.85	0	
03C001 04 00011	51	0.85 ***6th***	0.85	0	
03C008 00 00006B	51	0.86	0.85	0.01	
03C008 00 00004B	51	0.87	0.85	0.02	
03C008 00 00005B	51	0.87	0.85	0.02	
03C001 04 00010	51	0.88	0.85	0.03	
				0.48	Sum of Abs Dev
				÷ 10	N
n+1÷2	0.85 + 0.85 ÷	2 = 0.85		0.048	Ave Absolute Dev
				÷ 0.85	Median
				0.056471	AAD / Median
				x 100	
				5.647059	COD
Factor to adjust these properties					
	1.00 ÷ 0.85 =	1.176471 Factor			

Refining Neighborhoods with GIS & Statistics

- As will be seen on the next slide, those properties above the 110% level of assessment existed above the 64th percentile
- This group had a very low COD of 1.902748, showing that the median was an excellent representative of the population
- The ratios above the 64th percentile had a median of 1.295
- These properties could easily be lowered to the 100% level of assessment by utilizing a factor of 0.772201

COD of the 64th Percentile			Median	Absolute Deviation	
03C001 07 00013	51	1.25	1.295	0.045	
03C001 07 00018	51	1.28	1.295	0.015	
03C001 07 00024	51	1.28	1.295	0.015	
03C001 07 00012	51	1.29	1.295	0.005	
03C001 07 00014	51	1.29	1.295	0.005	
03C001 07 00019 **6th**	51	1.29	1.295	0.005	
03C001 07 00016 **7th**	51	1.30	1.295	0.005	
03C001 07 00017	51	1.31	1.295	0.015	
03C001 07 00010	51	1.33	1.295	0.035	
03C001 07 00009	51	1.36	1.295	0.065	
03C001 07 00025	51	1.36	1.295	0.065	
03C001 07 00022	51	1.37	1.295	0.075	
				0.35	Sum of Abs Dev
				÷ 12	n
n+1÷2 12 + 1 = 13 ÷ 2 = 6.5				0.0291667 ÷ 1.295	Ave Absolute Dev Median
1.29 + 1.30 ÷ 2 = 1.295				0.0225225	AAD / Median
				x 100	
				2.2522523	COD
Factor to adjust these properties					
1.00 ÷ 1.295 =	0.772201	Factor			

Refining Neighborhoods with GIS & Statistics

- Upon further investigation, it was found that those properties that were being undervalued were smaller waterfront townhomes.
- After this process, all of the properties in this area were adjusted accordingly

Refining Neighborhoods with GIS & Statistics

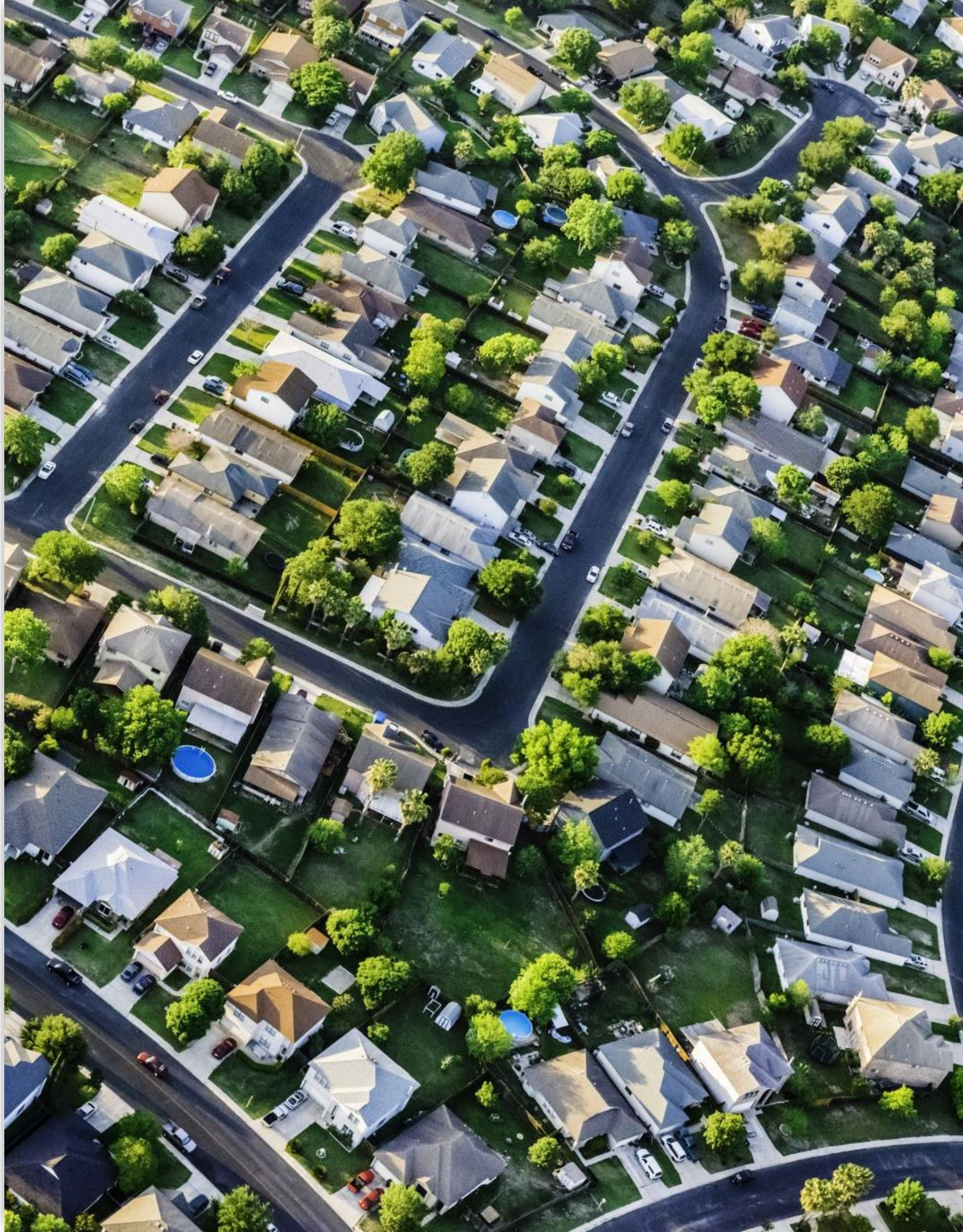
- The properties that were over-assessed were predominantly those on the east shore of the Island that had views of the commercial harbor

- After is process, all of the properties in this area were adjusted accordingly

Refining Neighborhoods with GIS & Statistics

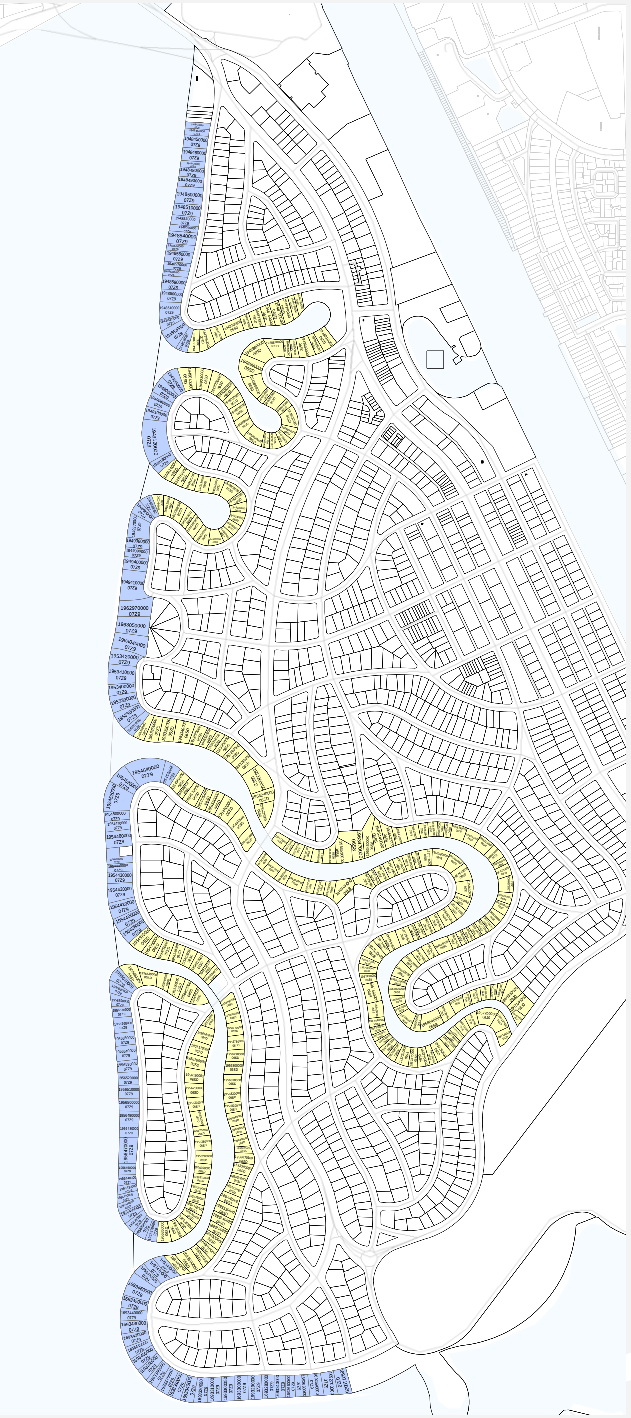
After adjusting the neighborhood for the several conditions, the 2024 descriptive statistics for this neighborhood are:

Median:	.9939
Mean:	.9897
COV:	15.76
COD:	11.71
PRD:	.9904



Refining Land Values with Geospatial Data

- On an adjacent Island, the land values were not keeping pace with the sale price of properties on the water or canals.
- We utilized water depth data from the Army Corp of Engineers, along with sonar readings taken by our office
- What we created was a map relevant to access to deep water
- Significant premiums were being paid for those properties with open blue water and there was a steady decrease in values at that access diminished
- Prior to this research, the assumption was based upon whether or not access to blue water was impeded by a bridge or not



Refining Land Values with Geospatial Data

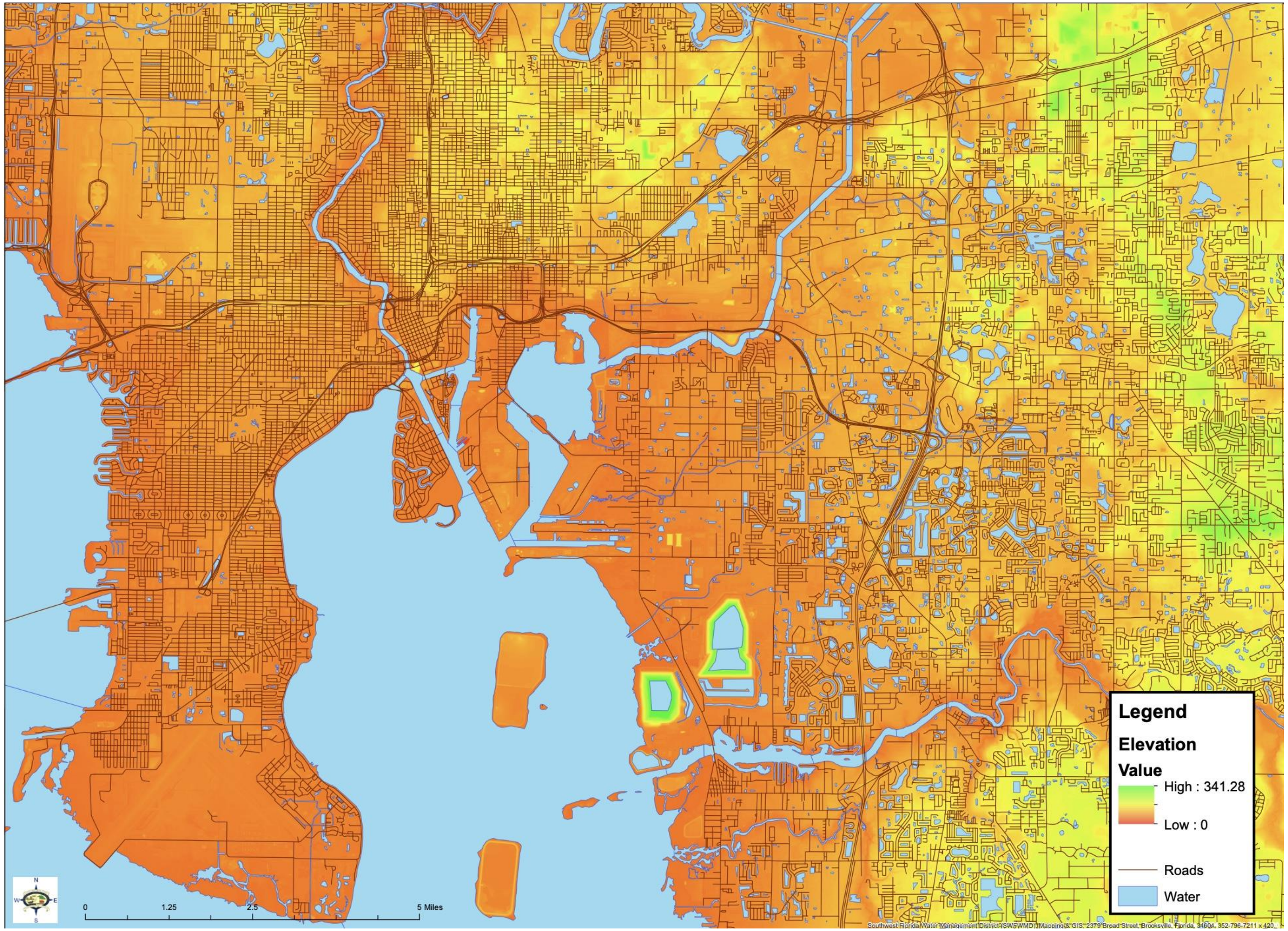
- As can be seen, the blue lots have direct deep water access.
  - Low tide depth of at least six (6) feet
- The yellow lots are considered to be canal lots
  - Low tide depth below six (6) feet
  - These lots gradually decrease in value as the water depth of low tide diminishes
  - Once the low water tide depth reaches below two (2) feet, the value is constant
  - The center canal is impacted by both a bridge crossing and diminishing depths

Refining Models with Geospatial Data

- Over 105,000 properties of Hillsborough County are in an A or B flood zone.
- After Hurricane Michael, FEMA readjusted their Hurricane flood maps
  - What was learned from Michael was the idea of “stacking”
  - Stacking occurs when the water can’t dissipate and simply stacks up higher and higher
- With these new maps and increasing insurance rates, we utilize lidar maps to find the base elevation of each home, to utilize as a characteristic in our regression models

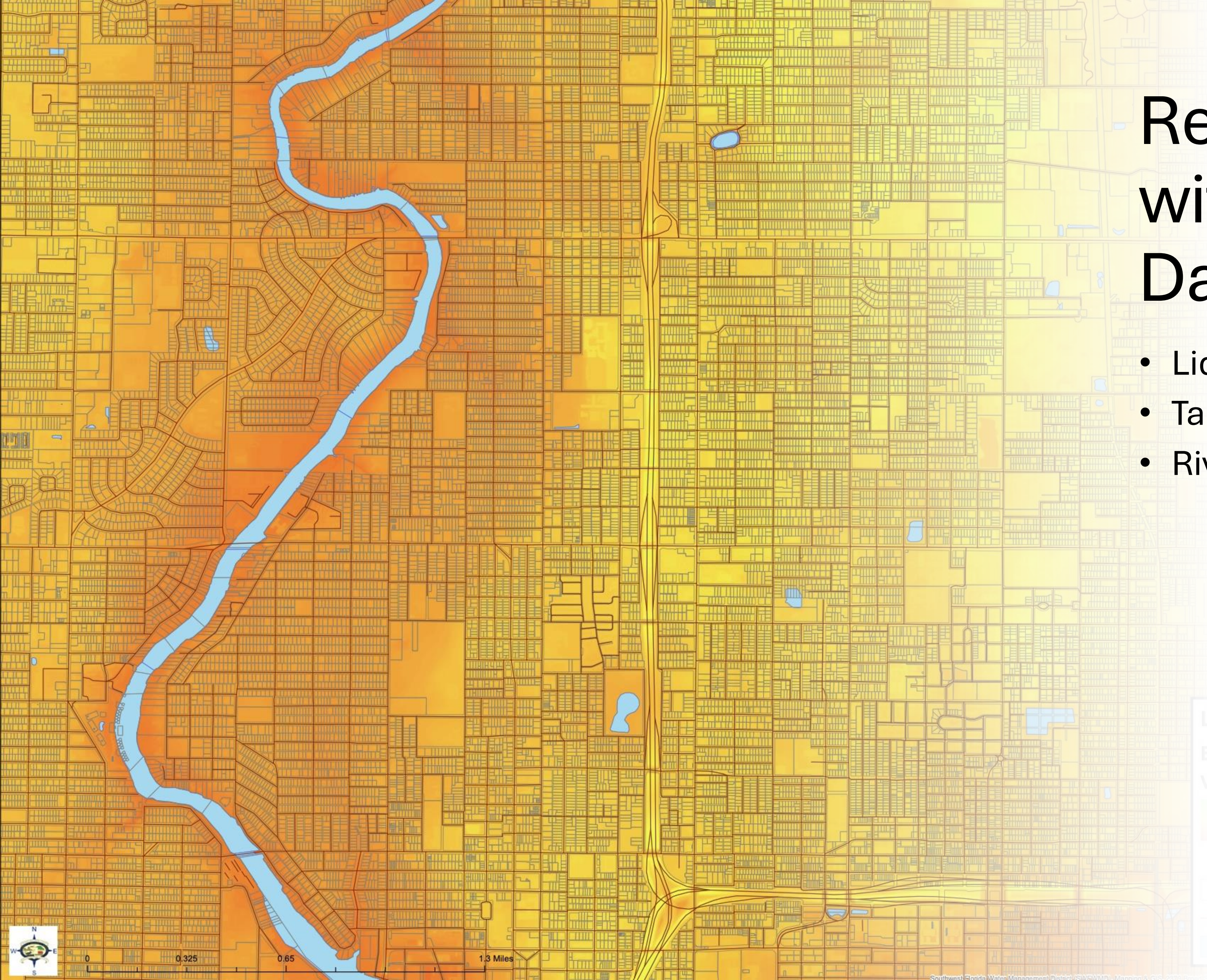
Refining Models with Geospatial Data

Lidar Elevation Model Map
Tampa, FL



Refining Models with Geospatial Data

- Lidar Elevation Model Map
- Tampa Heights, FL
- Riverside Heights, FL



Hillsborough Public Facing GIS

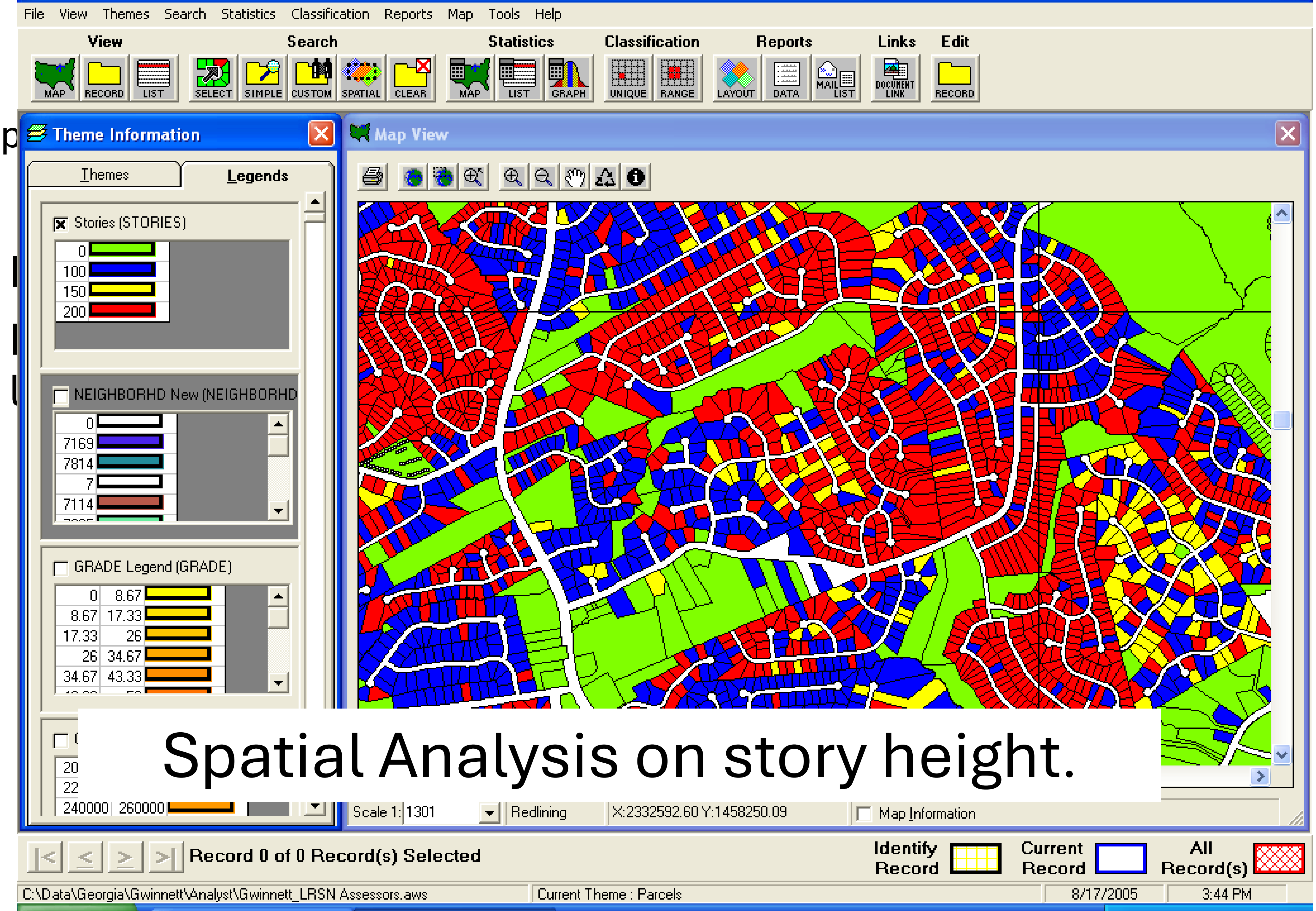
https://www.arcgis.com/apps/dashboards/dfa0789a1f264acb884745d029036117

Thank you

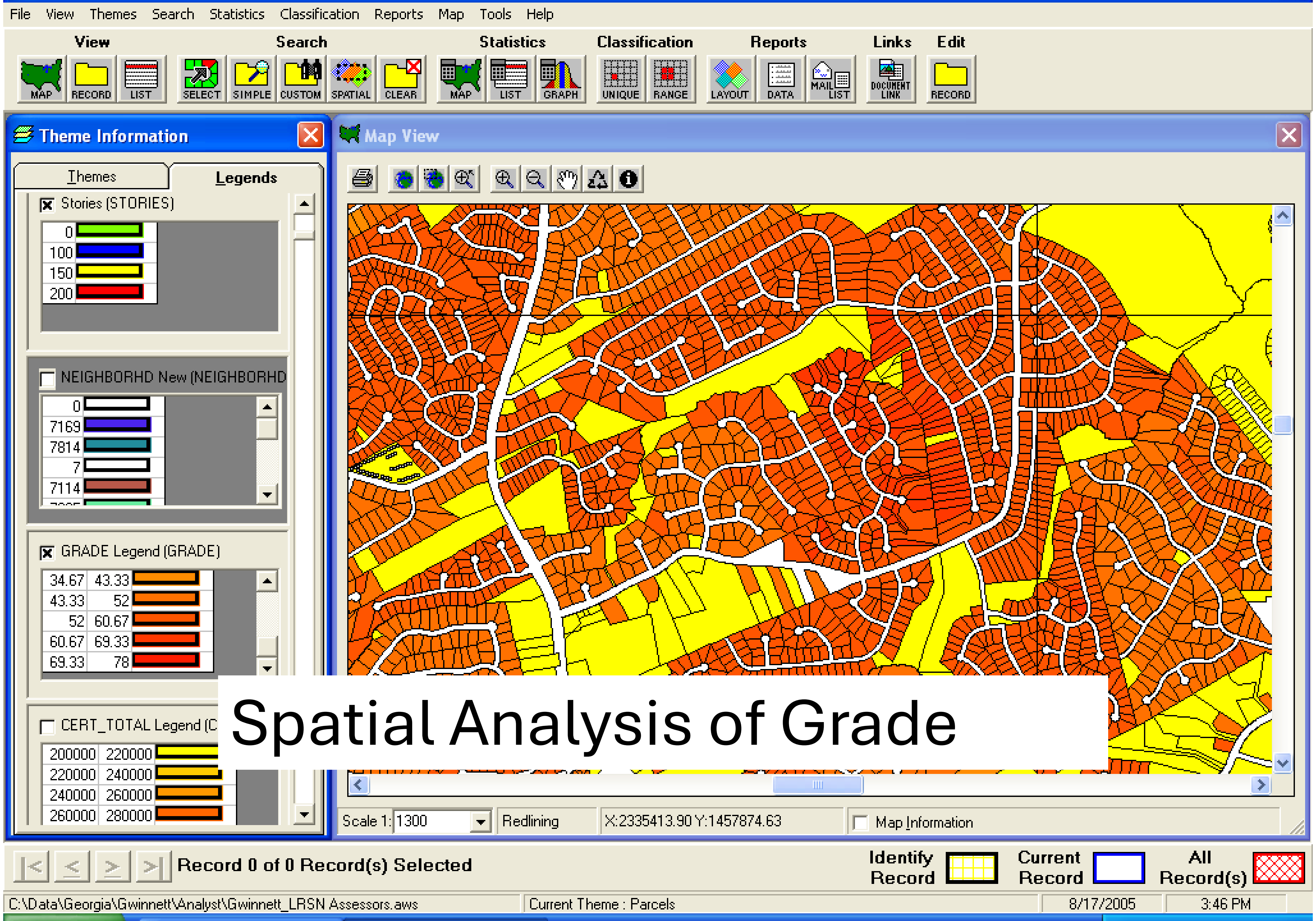
Patrick Alesandrini, CAE, MAI, RES, SRA, CFE

drini57@outlook.com

+1 309-453-6684

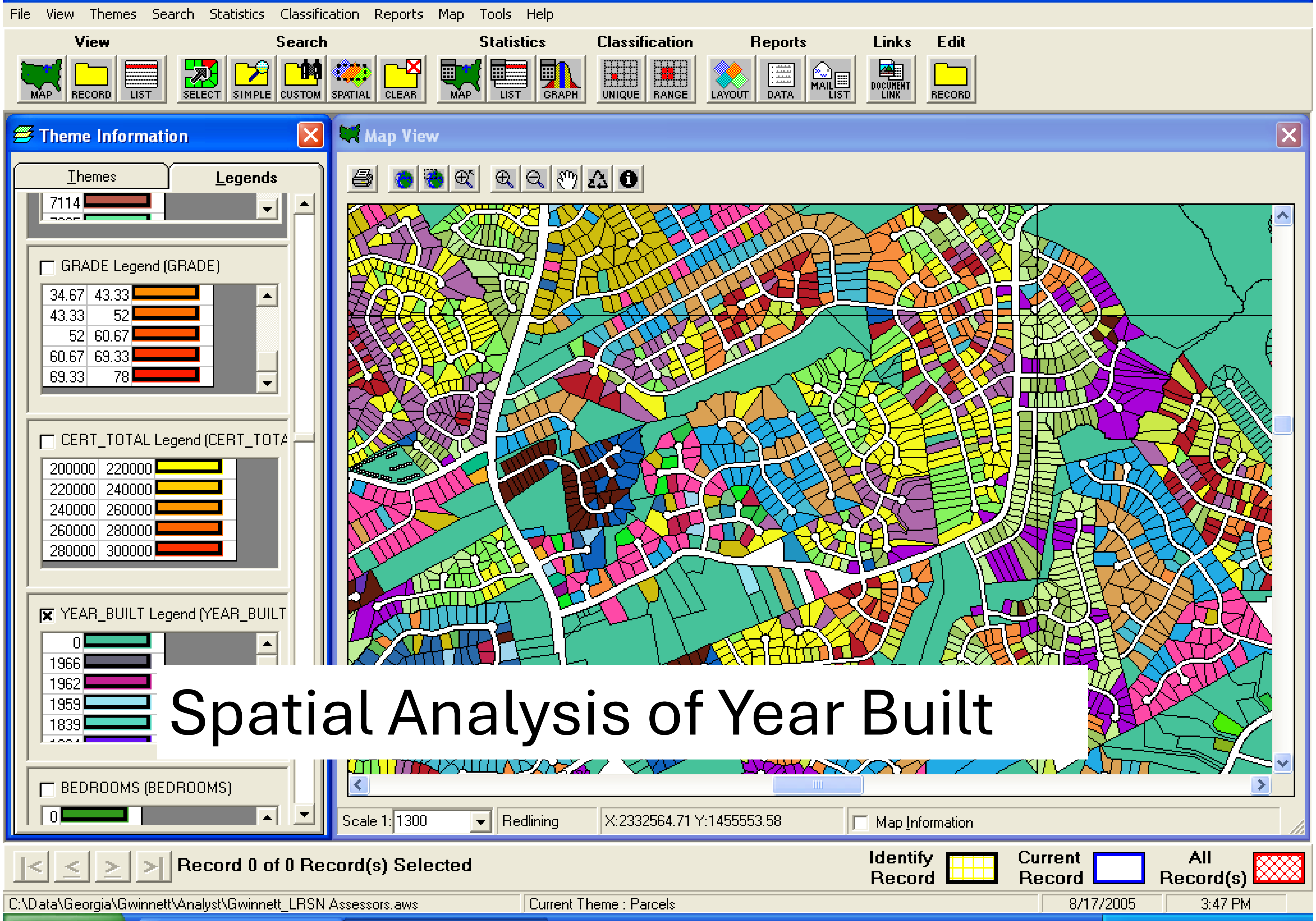


Spatial Analysis on story height.

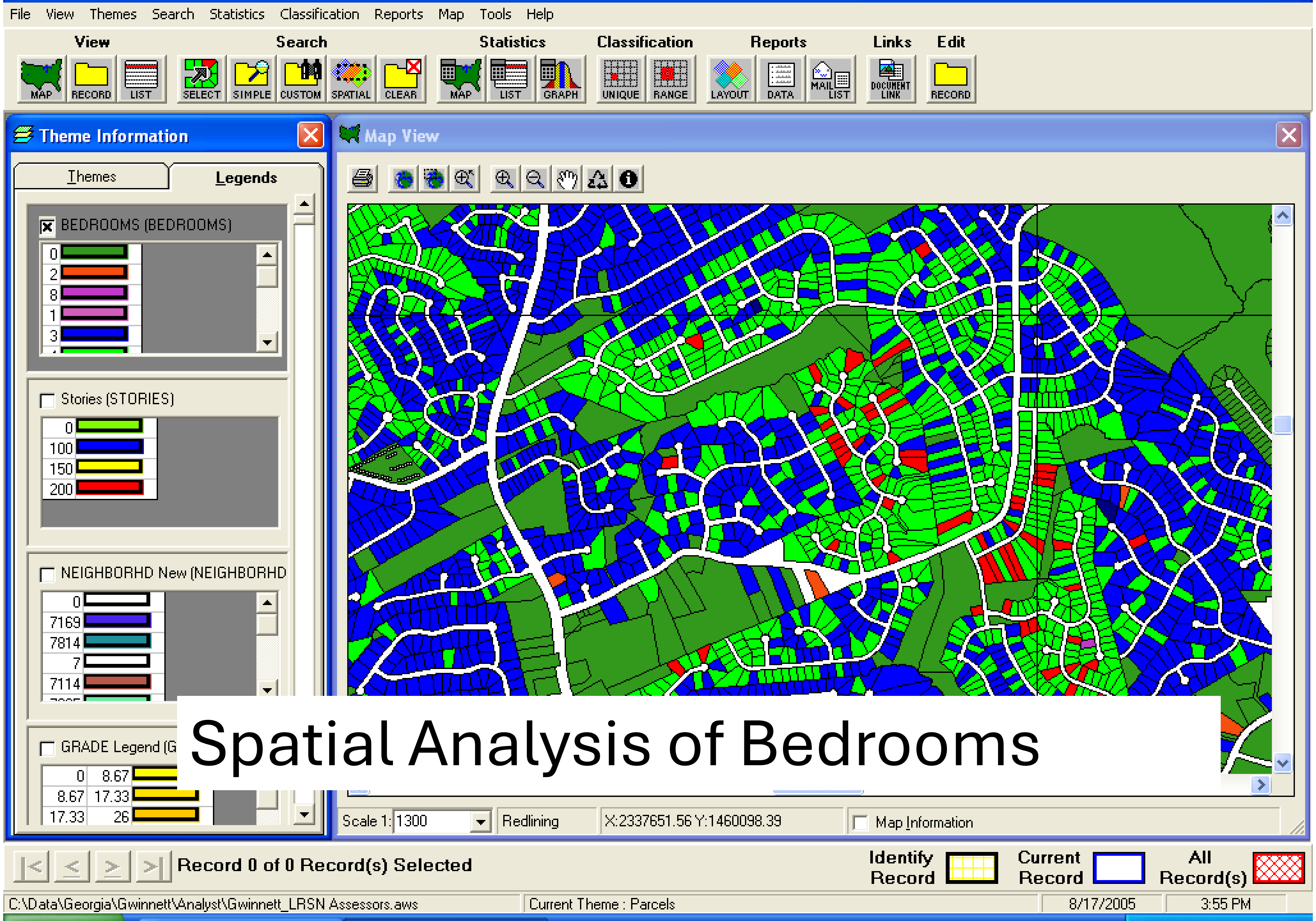


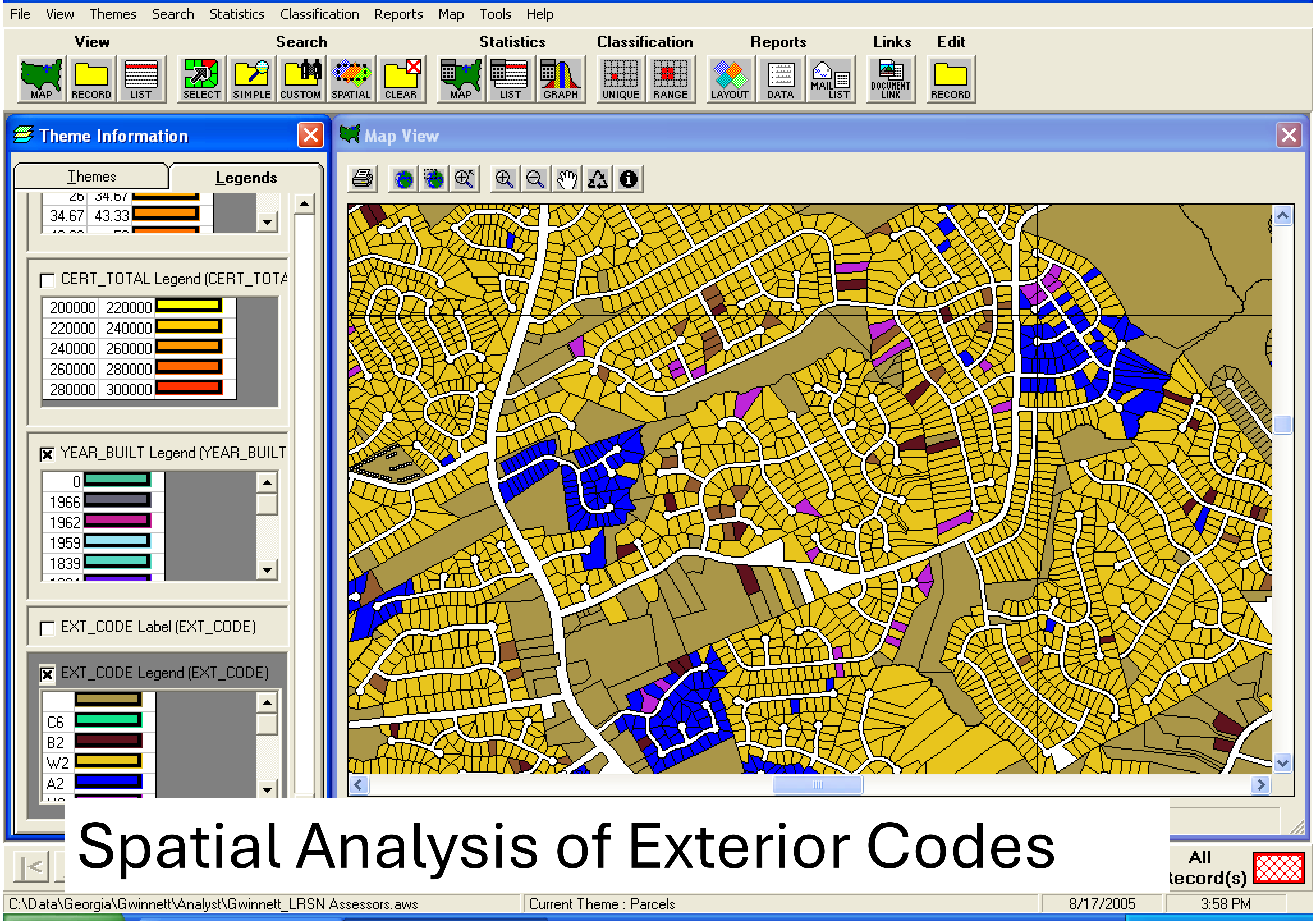
Spatial Analysis of Grade

Spatial Analysis of Year Built



Spatial Analysis of Bedrooms





Spatial Analysis of Exterior Codes