

Spatial Neighbors

Lecture #15 | GEOG 510
GIS & Spatial Analysis in Public Health
Varun Goel

Outline

- Neighborhoods and neighbors
- Distance and topology based
- Neighborhood weight matrix
- How to choose?

Tobler's First Law

- Tobler's first law of Geography
 - *Everything is related to everything else, but near things are more related than distant things*
 - Values at locations near each other tend to be similar, with similarity decreasing with distance
 - Implies that phenomena are not distributed randomly (throughout space)
 - Imagine how the world would appear if everything were randomly distributed!

Spatial Autocorrelation

- Spatial Autocorrelation
 - The degree of similarity between objects that are located near each other
 - Attribute similarity
 - Can be measured, quantitatively
 - Over an entire region (global)
 - In a smaller area within the region (local)
 - Use in health geography
 - Prediction, distance decay, cluster analysis

Neighborhood

- What is a neighborhood?
 - Neighborhood has many definitions
 - Zone of influence
 - Idea of nearness or connectedness
 - Things or objects that are near one another
 - Things or objects that affect one another
- Why important?
 - To describe or characterize spatial relationships among objects requires us to define a neighborhood

Neighbors

- Neighbors are features located within a neighborhood
 - To describe or characterize spatial relationships among objects requires us to define the neighbor relationships
 - Neighbors for each observation!

Neighbors

- Required for...
 - Spatial point pattern tests
 - Spatial autocorrelation tests/measures
 - Spatial regression

Neighbors

- Basic approaches to characterize neighbors
 - Binary (Y,N)
 - Either you are a neighbor, or not
 - Continuous
 - Amount of “neighborliness”
 - Generally, based on distance
 - On a conceptual level, some neighbors may be strong, while others are weaker

Neighbors

- Basic approaches to characterize neighbors
 - Absolute distance
 - Objects are considered neighbors based upon the actual distance separating them
 - Relative distance
 - Nearest feature
 - The nearest feature is considered a neighbor
 - Or, nearest k features
 - Topology-based
 - Connecting features are considered neighbors

Neighbors

- Absolute distance approach
 - Objects are considered neighbors based upon a predetermined threshold distance
 - For points
 - Distance between points
 - For polygons
 - Distance between polygon centroids

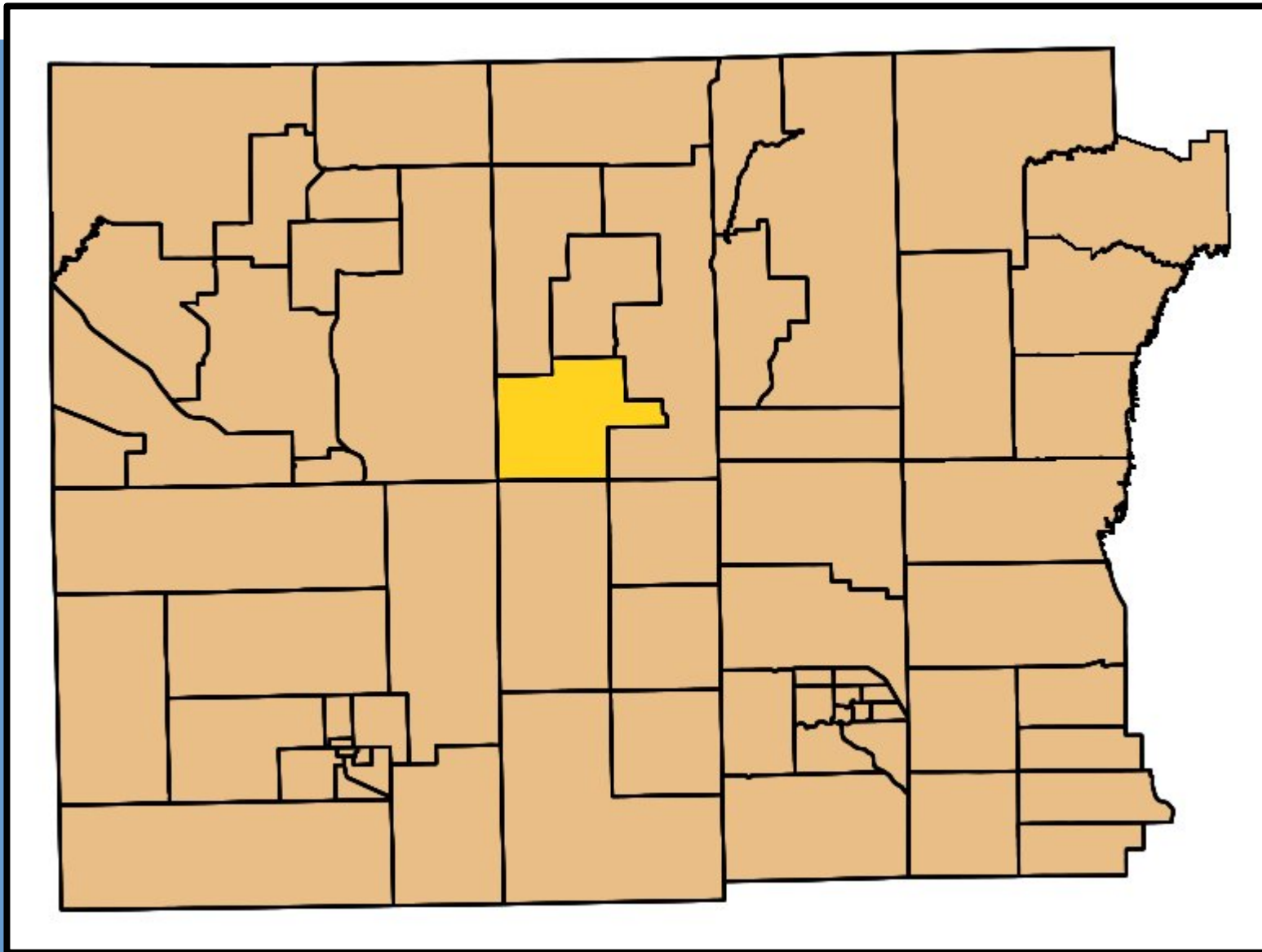
Neighbors

- Relative distance approach
 - Nearest feature
 - For points
 - Distance between points
 - For polygons
 - Distance between polygon centroids

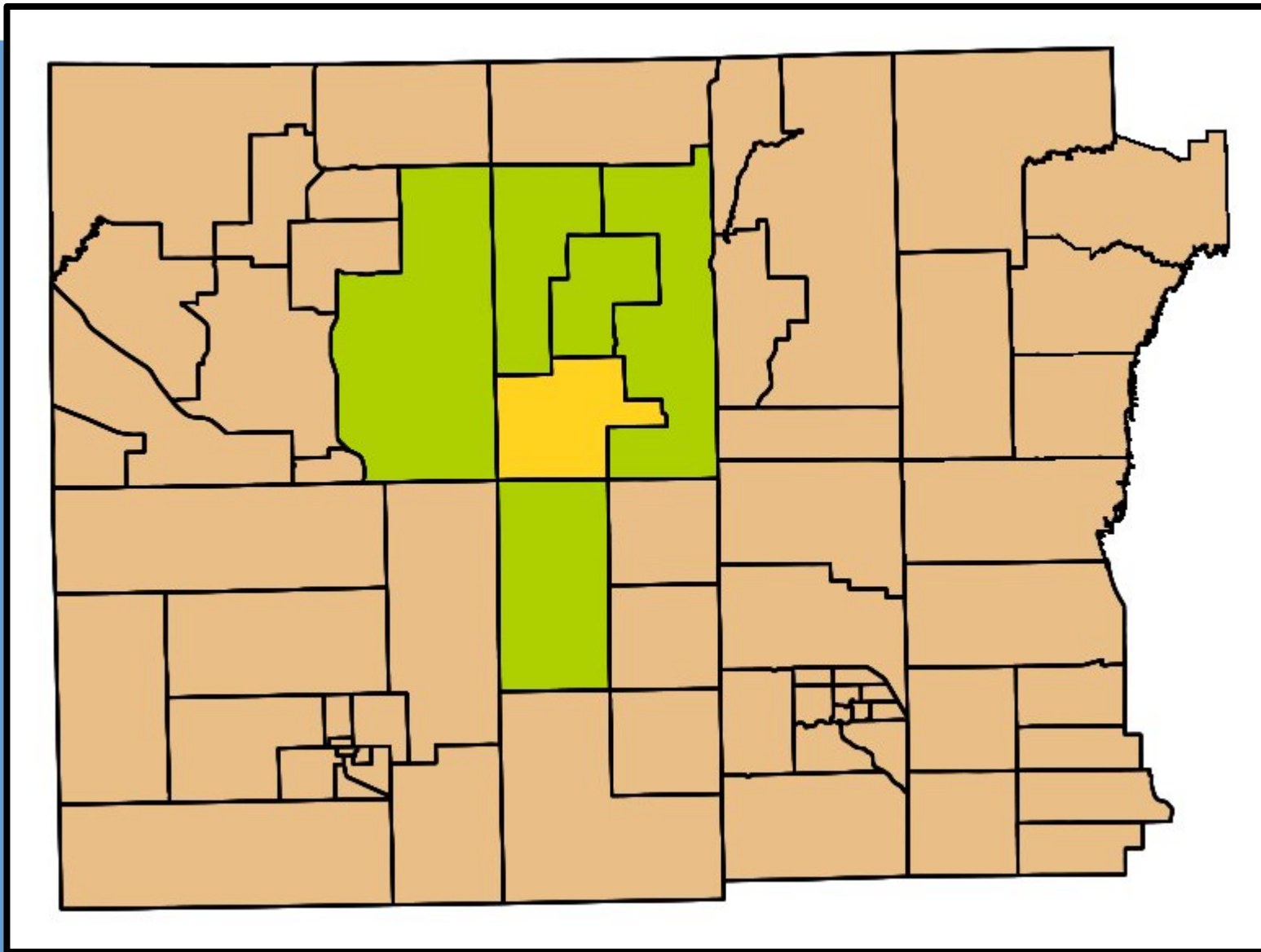
Neighbors

- Relative distance approach
 - Topology-based
 - For points
 - Not available
 - For polygons
 - Based on shared borders (points)

RELATIVE DISTANCE, TOPOLOGY

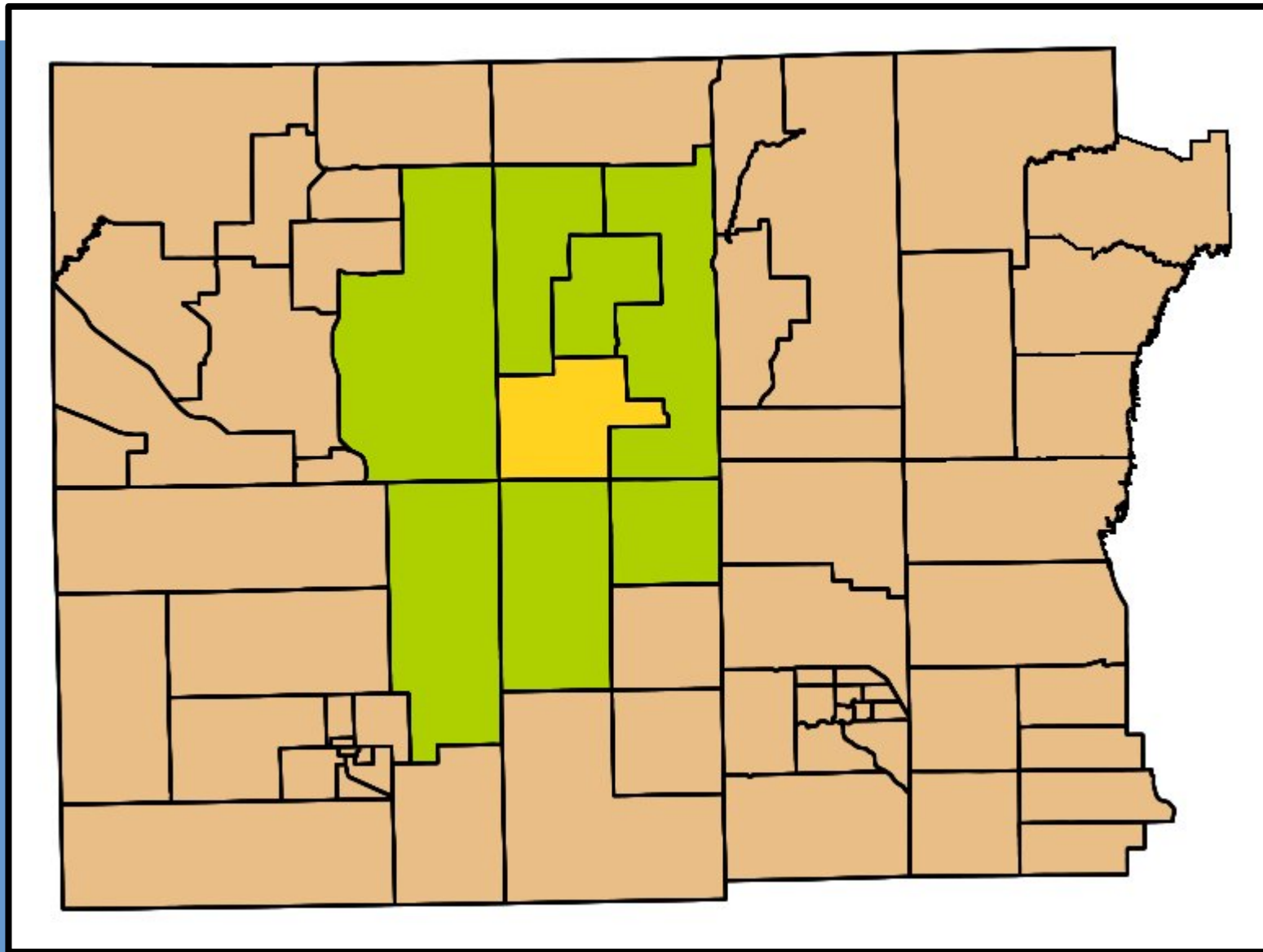


RELATIVE DISTANCE, TOPOLOGY



Rook's case, 1st order neighbors

RELATIVE DISTANCE, TOPOLOGY



Queen's case, 1st order neighbors

Neighbors

- Stored in a neighborhood weight matrix
 - ...or, a similar format
- Matrix is $n \times n$
 - The number of observations = n
 - Entries in this matrix describe the relationships between observations

Neighborhood Weight Matrix

	A	B	C	D	E	F	G	H	I
A	0	1	0	1	0	0	0	0	0
B	1	0	1	0	1	0	0	0	0
C	0	1	0	0	0	1	0	0	0
D	1	0	0	0	1	0	1	0	0
E	0	1	0	1	0	1	0	1	0
F	0	0	1	0	1	0	0	0	1
G	0	0	0	1	0	0	0	1	0
H	0	0	0	0	1	0	1	0	1
I	0	0	0	0	0	1	0	1	0

A	B	C
D	E	F
G	H	I

Rook
Contiguity

Neighborhood Weight Matrix

	A	B	C	D	E	F	G	H	I
A	0	1	0	1	0	0	0	0	0
B	1	0	1	0	1	0	0	0	0
C	0	1	0	0	0	1	0	0	0
D	1	0	0	0	1	0	1	0	0
E	0	1	0	1	0	1	0	1	0
F	0	0	1	0	1	0	0	0	1
G	0	0	0	1	0	0	0	1	0
H	0	0	0	0	1	0	1	0	1
I	0	0	0	0	0	1	0	1	0

A	B	C
D	E	F
G	H	I

Rook
Contiguity

Neighborhood Weight Matrix

	A	B	C	D	E	F	G	H	I
A	0	1	0	1	0	0	0	0	0
B	1	0	1	0	1	0	0	0	0
C	0	1	0	0	0	1	0	0	0
D	1	0	0	0	1	0	1	0	0
E	0	1	0	1	0	1	0	1	0
F	0	0	1	0	1	0	0	0	1
G	0	0	0	1	0	0	0	1	0
H	0	0	0	0	1	0	1	0	1
I	0	0	0	0	0	1	0	1	0

A	B	C
D	E	F
G	H	I

Rook
Contiguity

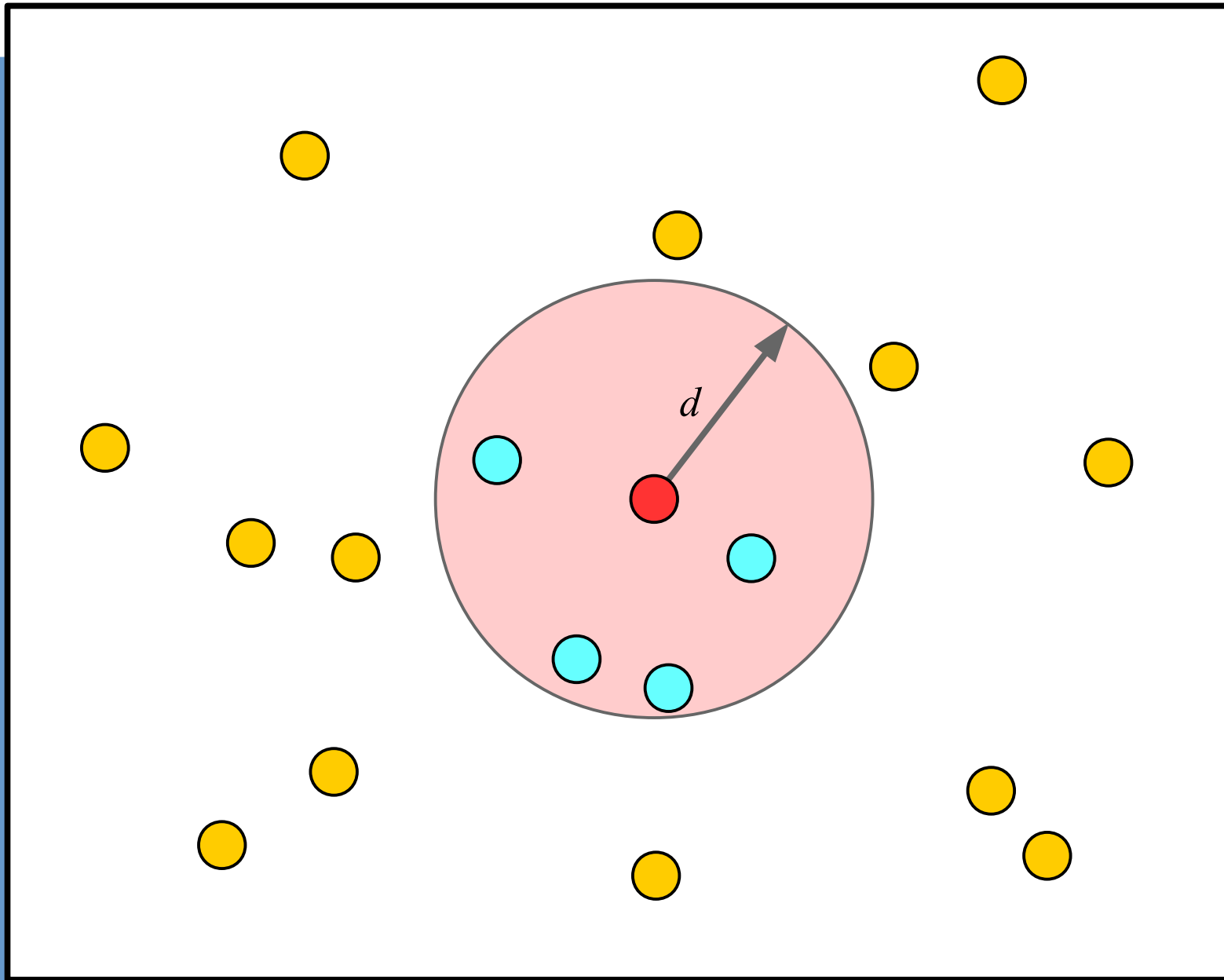
Neighborhood Weight Matrix

	A	B	C	D	E	F	G	H	I
A	0	1	0	1	1	0	0	0	0
B	1	0	1	1	1	1	0	0	0
C	0	1	0	0	1	1	0	0	0
D	1	1	0	0	1	0	1	1	0
E	1	1	1	1	0	1	1	1	1
F	0	1	1	0	1	0	0	1	1
G	0	0	0	1	1	0	0	1	0
H	0	0	0	1	1	1	1	0	1
I	0	0	0	0	1	1	0	1	0

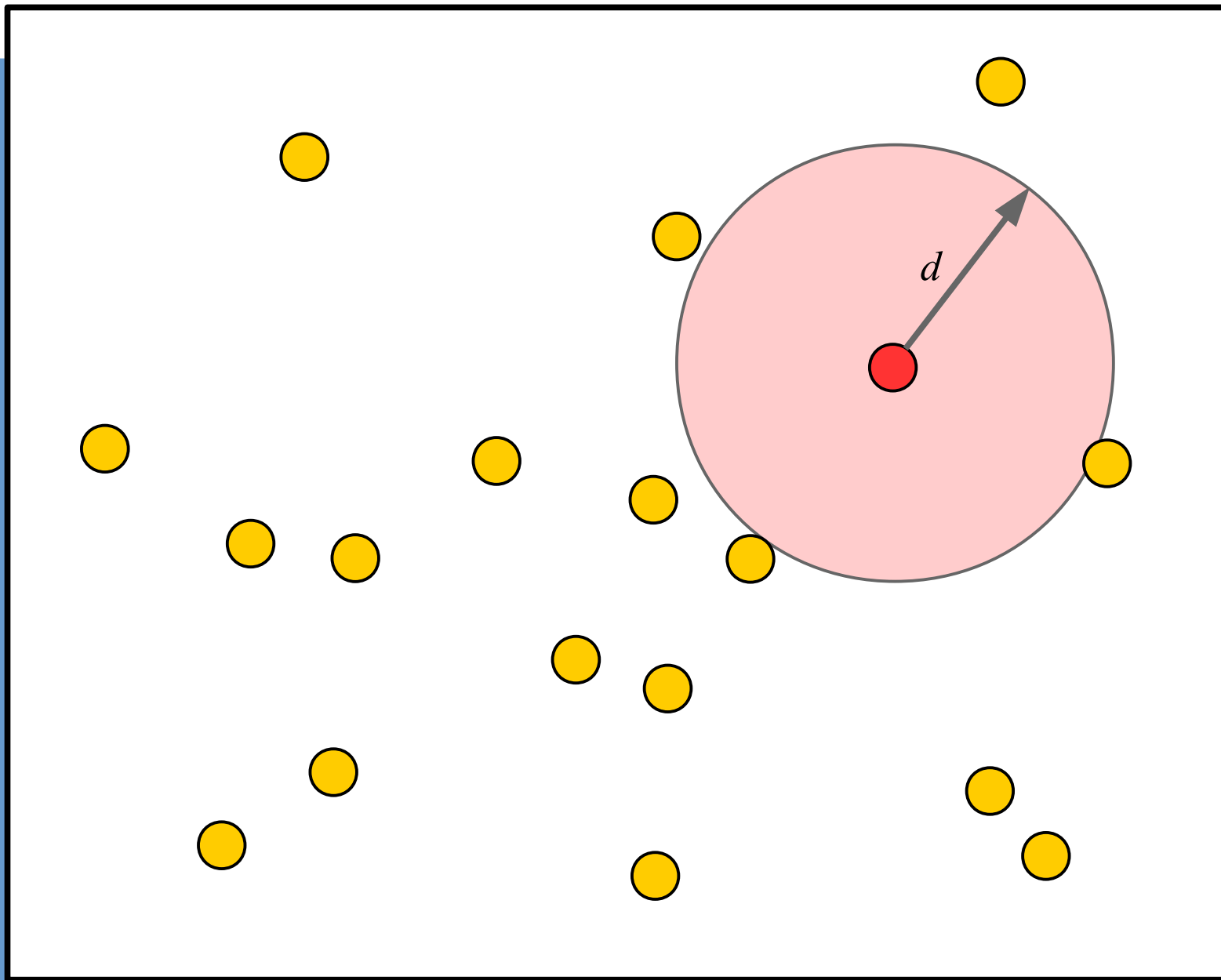
A	B	C
D	E	F
G	H	I

Queen
Contiguity

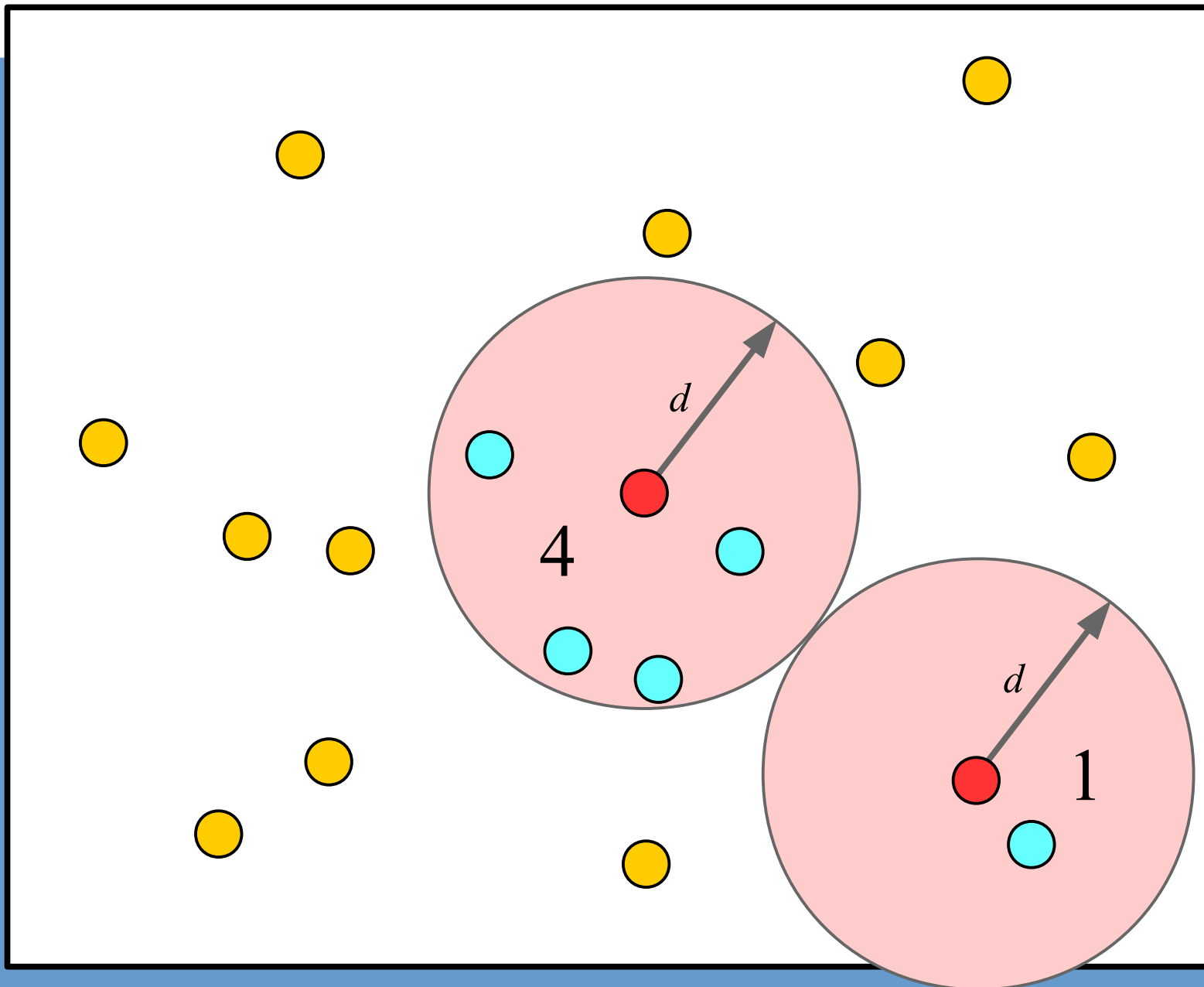
ABSOLUTE DISTANCE, BINARY



ABSOLUTE DISTANCE, BINARY



ABSOLUTE DISTANCE, BINARY



Neighborhood Weight Matrix

	A	B	C	D	E	F	G	H	I
A	0	1	0	1	0	0	0	0	0
B	1	0	1	0	1	0	0	0	0
C	0	1	0	0	0	1	0	0	0
D	1	0	0	0	1	0	1	0	0
E	0	1	0	1	0	1	0	1	0
F	0	0	1	0	1	0	0	0	1
G	0	0	0	1	0	0	0	1	0
H	0	0	0	0	1	0	1	0	1
I	0	0	0	0	0	1	0	1	0

A	B	C
D	E	F
G	H	I

Rook
Contiguity

*No Row
Standardization*

Neighborhood Weight Matrix

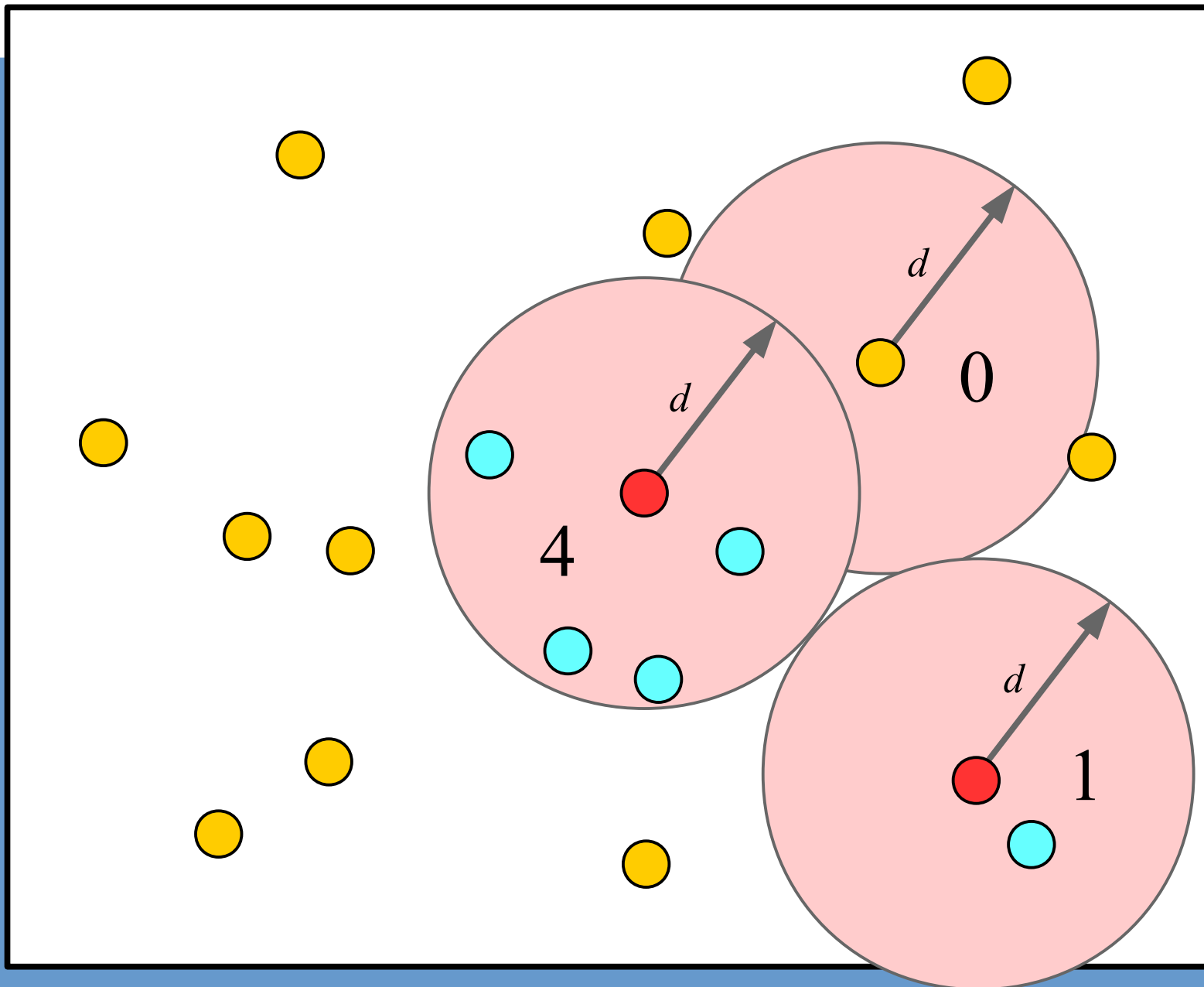
	A	B	C	D	E	F	G	H	I
A	0	0.50	0	0.50	0	0	0	0	0
B	0.33	0	0.33	0	0.33	0	0	0	0
C	0	0.50	0	0	0	0.50	0	0	0
D	0.33	0	0	0	0.33	0	0.33	0	0
E	0	0.25	0	0.25	0	0.25	0	0.25	0
F	0	0	0.33	0	0.33	0	0	0	0.33
G	0	0	0	0.50	0	0	0	0.50	0
H	0	0	0	0	0.33	0	0.33	0	0.33
I	0	0	0	0	0	0.50	0	0.50	0

A	B	C
D	E	F
G	H	I

Rook
Contiguity

Row
Standardized
Divide by Row Sum

ABSOLUTE DISTANCE, BINARY



ABSOLUTE DISTANCE, BINARY

GeoDa

Weights File Creation

Weights File ID Variable:

Contiguity Weight

☐ Queen contiguity Order of contiguity:

☐ Rook contiguity ☐ Include lower orders

☐ Precision threshold:

Distance Weight

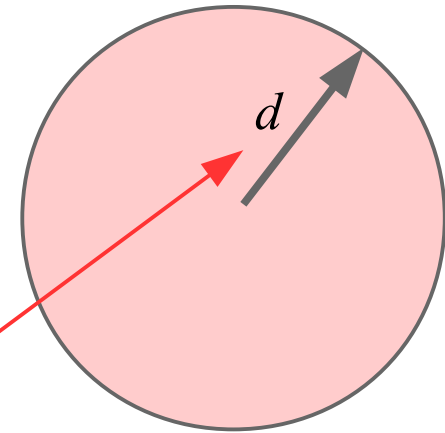
Distance metric:

X-coordinate variable:

Y-coordinate variable:

☒ Threshold distance:

☐ k-Nearest Neighbors Number of neighbors:

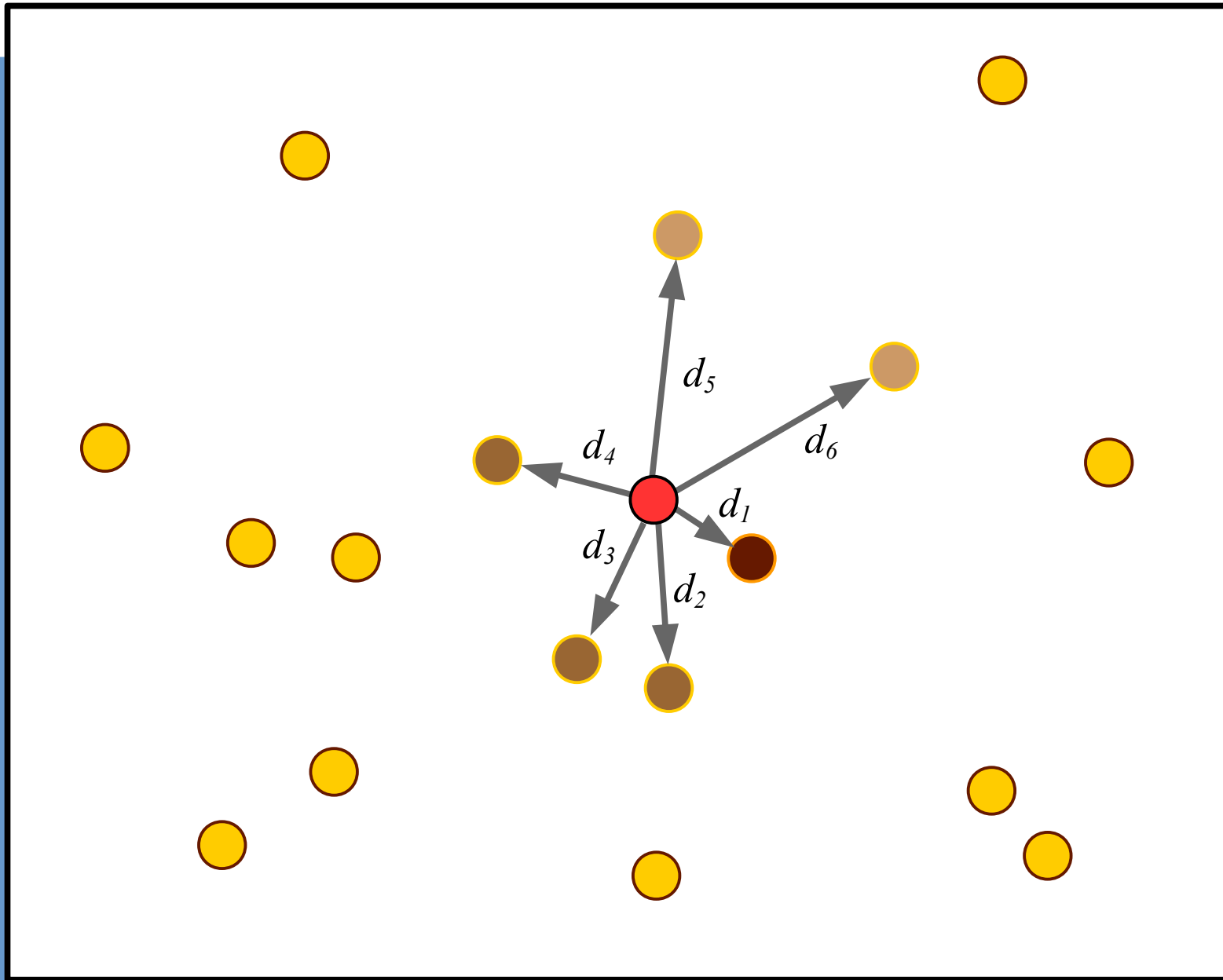


Default (auto)
Distance:

*Minimum distance for
each feature to have at
least one neighbor*

All neighborhood
weight matrices are
Row Standardized

ABSOLUTE DISTANCE, CONTINUOUS



Distance Parameter

- Weight of relationship determined by an “inverse” relationship with distance
 - Short distance = High weight
 - Long distance = Low weight

$$w_{i,j} = \frac{1}{d_{i,j}^x}$$

$w_{i,j}$ = Weight value in neighborhood weight matrix for observation i to observation j

$d_{i,j}$ = Distance from observation i to observation j

x = Distance effect parameter

Distance Parameter

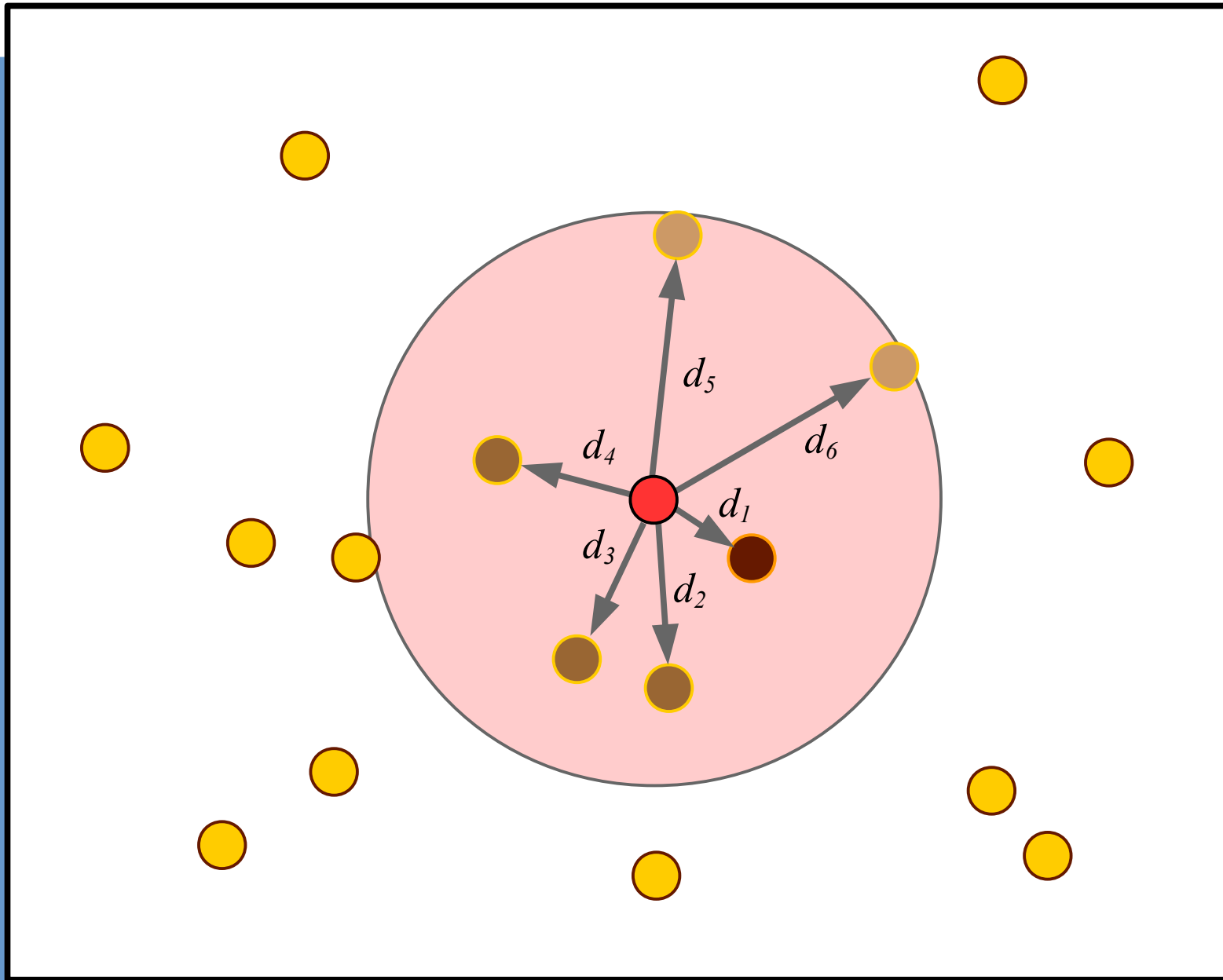
$$w_{i,j} = \frac{1}{d_{i,j}}$$

$$w_{i,j} = \frac{1}{d_{i,j}^2}$$

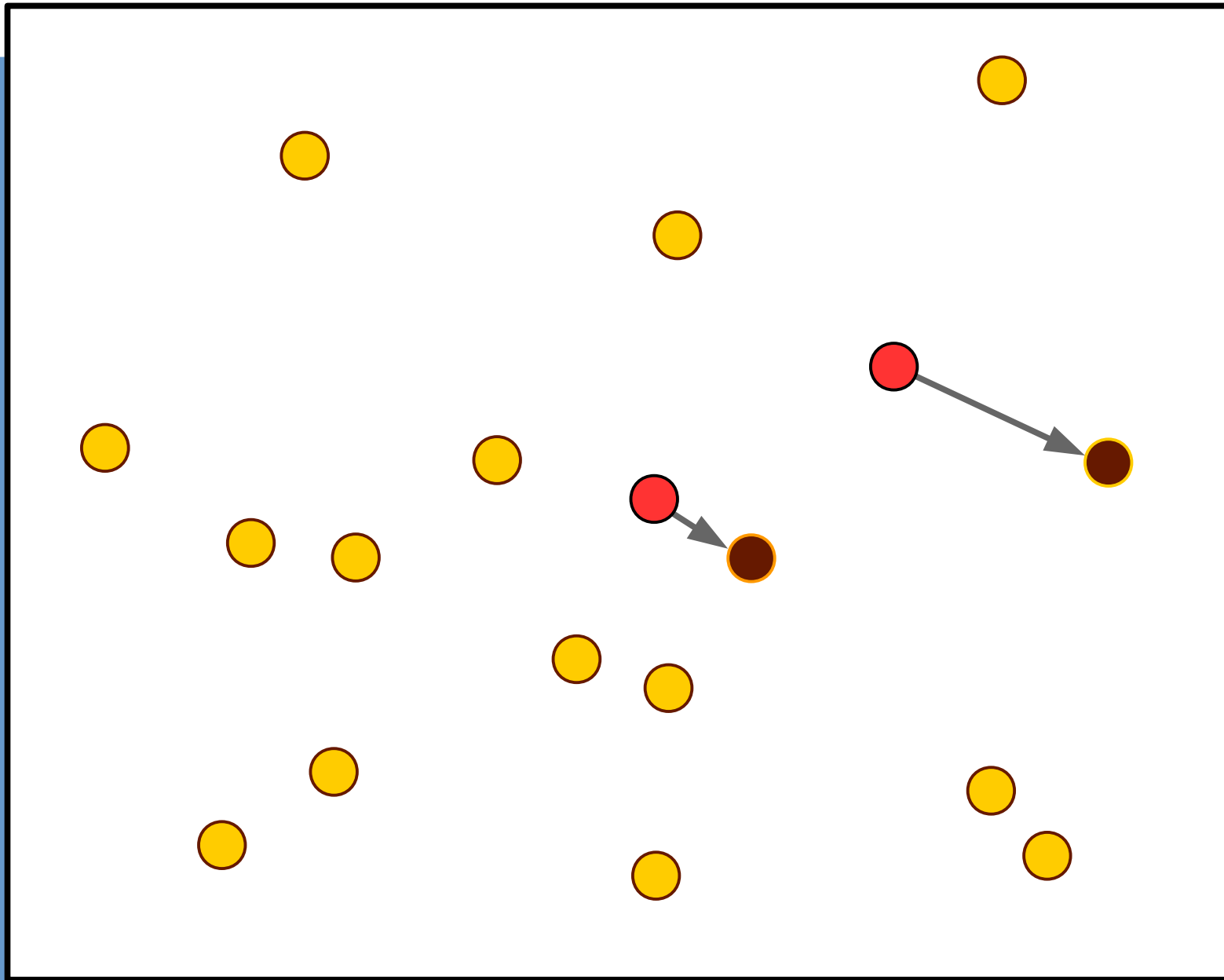
Distance (d)	$w(1)$	$w(2)$
10	0.1000	0.0100
12	0.0833	0.0069
20	0.0500	0.0025
42	0.0238	0.0006
46	0.0217	0.0005
58	0.0172	0.0003

*Important
to note the
relative
differences
among
weights!*

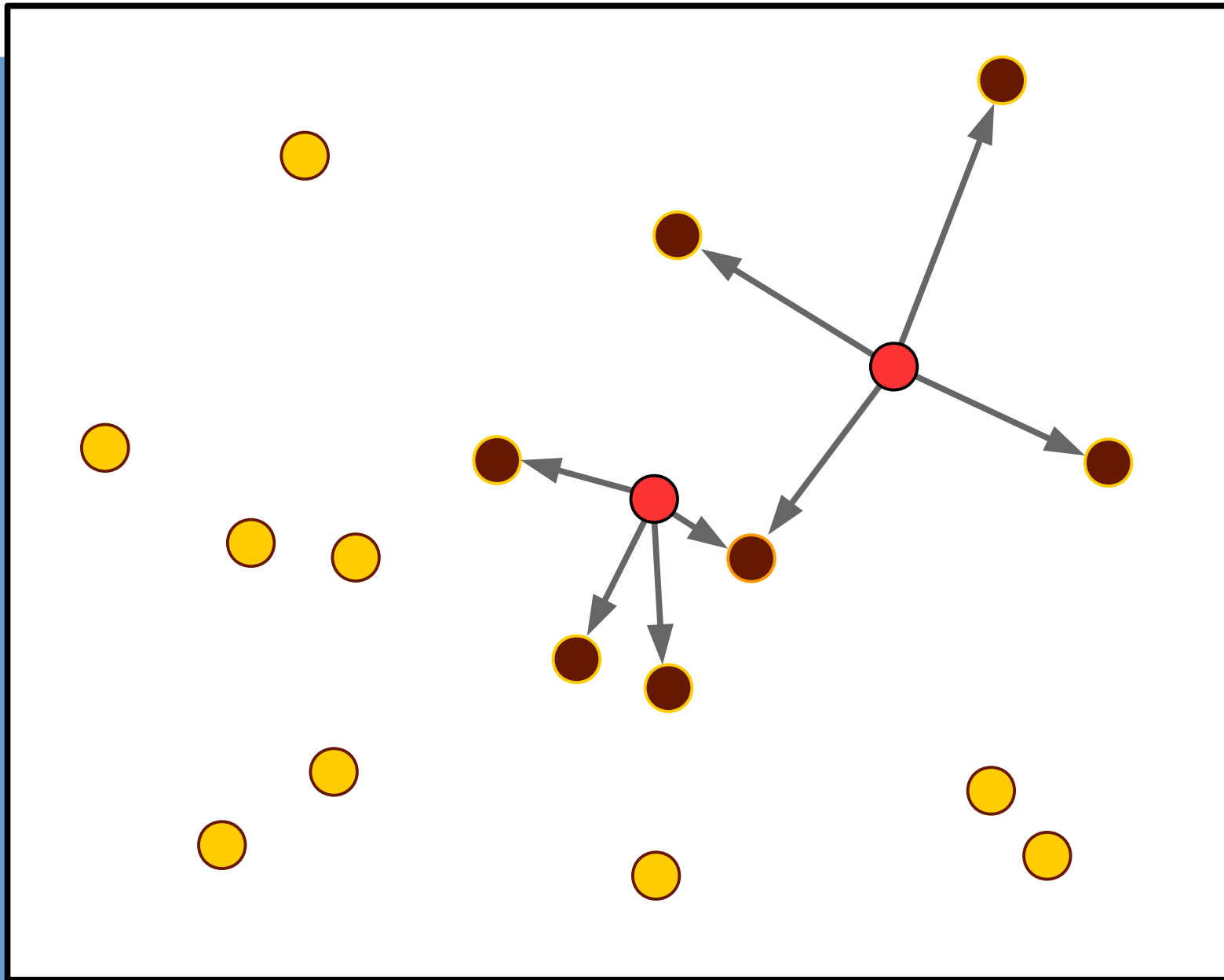
ABSOLUTE DISTANCE, CONTINUOUS



RELATIVE DISTANCE, $K(1)$ NEAREST



RELATIVE DISTANCE, $K(4)$ NEAREST



RELATIVE DISTANCE, $K(4)$ NEAREST

GeoDa

Weights File Creation

Weights File ID Variable:

Contiguity Weight

☐ Queen contiguity Order of contiguity

☐ Rook contiguity ☐ Include lower orders

☐ Precision threshold

Distance Weight

Distance metric:

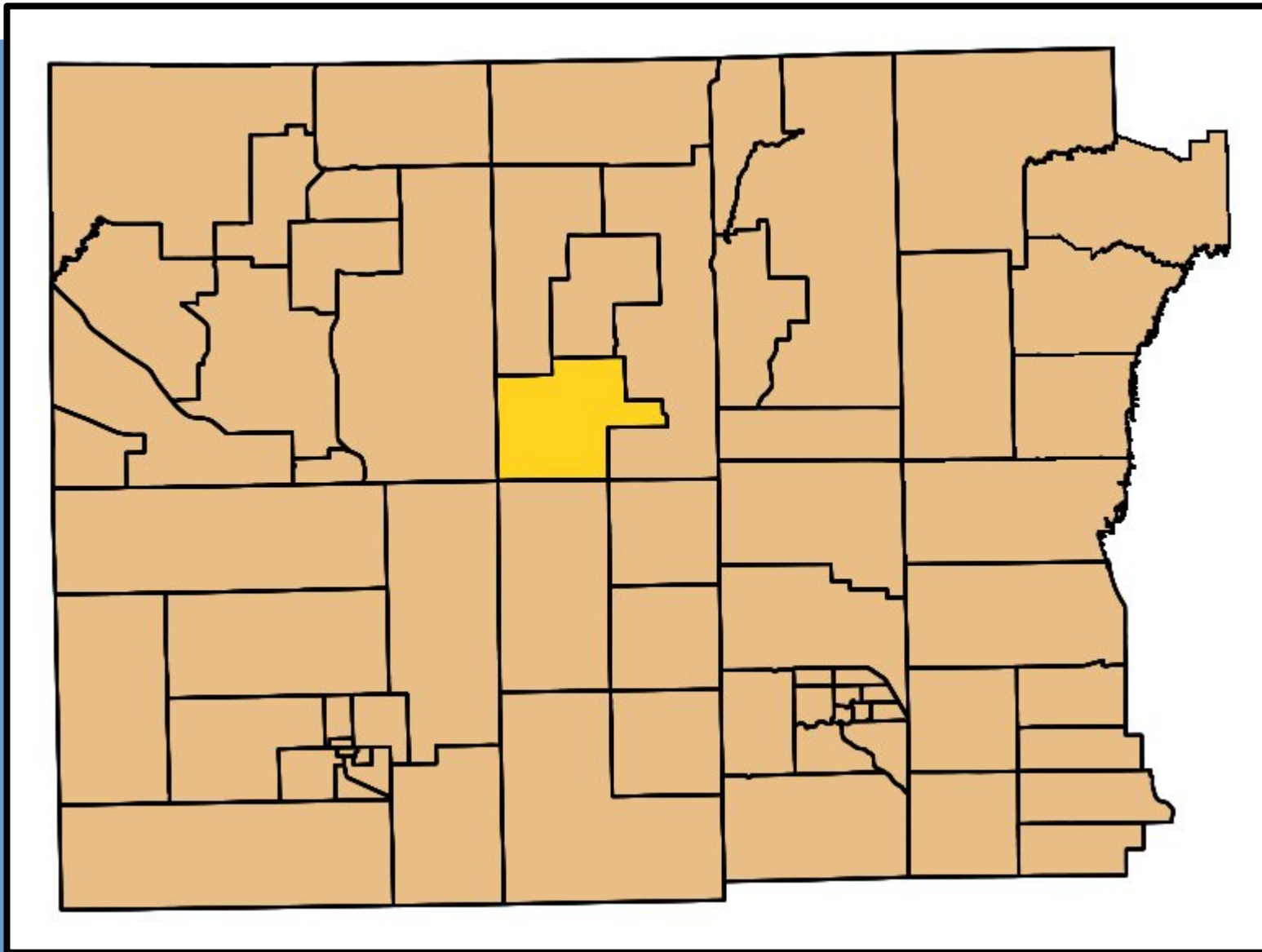
X-coordinate variable:

Y-coordinate variable:

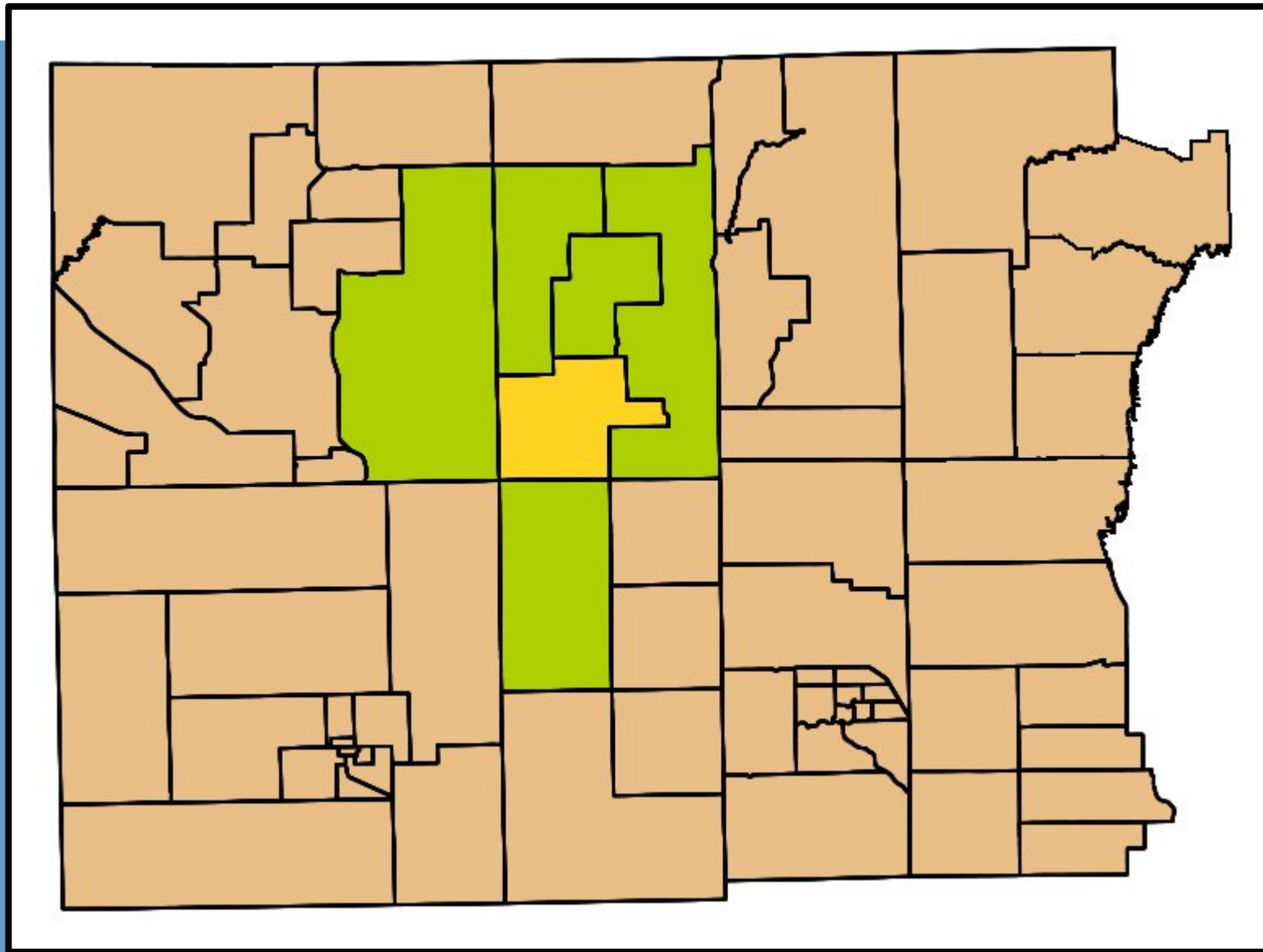
☐ Threshold distance

☒ k-Nearest Neighbors Number of neighbors

RELATIVE DISTANCE, TOPOLOGY



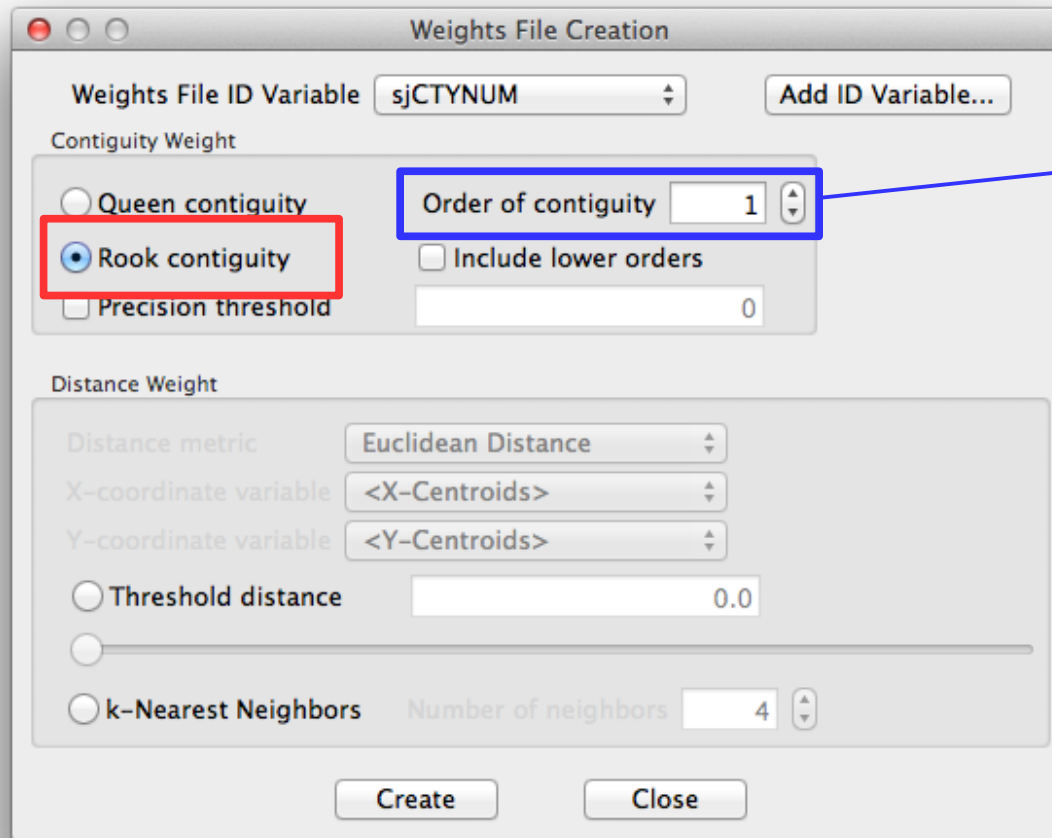
RELATIVE DISTANCE, TOPOLOGY



Rook's case, 1st order neighbors

RELATIVE DISTANCE, TOPOLOGY

GeoDa



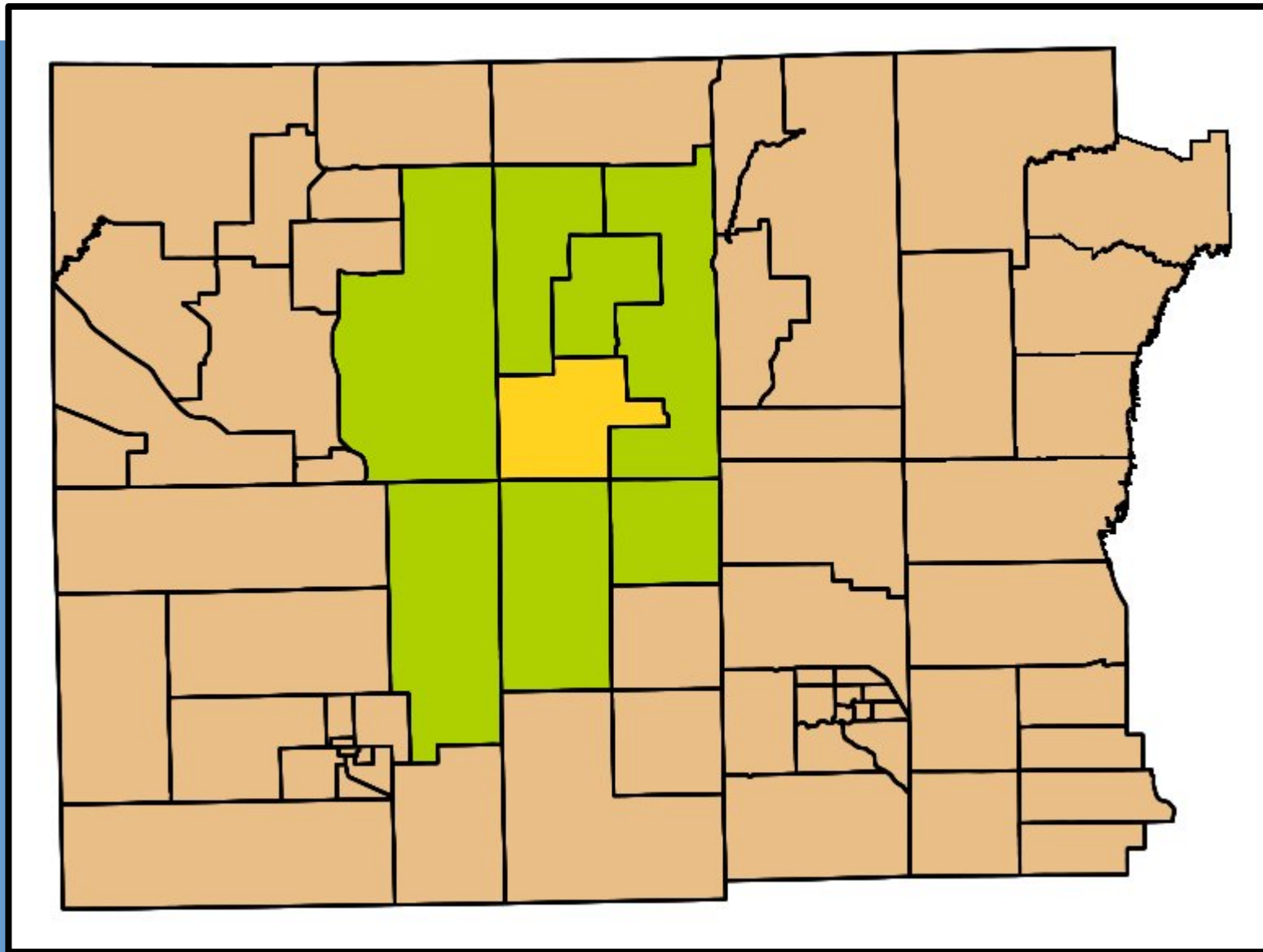
The image shows the 'Weights File Creation' dialog box in GeoDa. The 'Weights File ID Variable' is set to 'sjCTYNUM'. Under 'Contiguity Weight', 'Rook contiguity' is selected (highlighted with a red box), and 'Order of contiguity' is set to 1 (highlighted with a blue box and an arrow pointing to the text 'Order Number:'). The 'Distance Weight' section is disabled. At the bottom are 'Create' and 'Close' buttons.

Order Number:

All neighborhood weight matrices are Row Standardized

Rook's case, 1st order neighbors

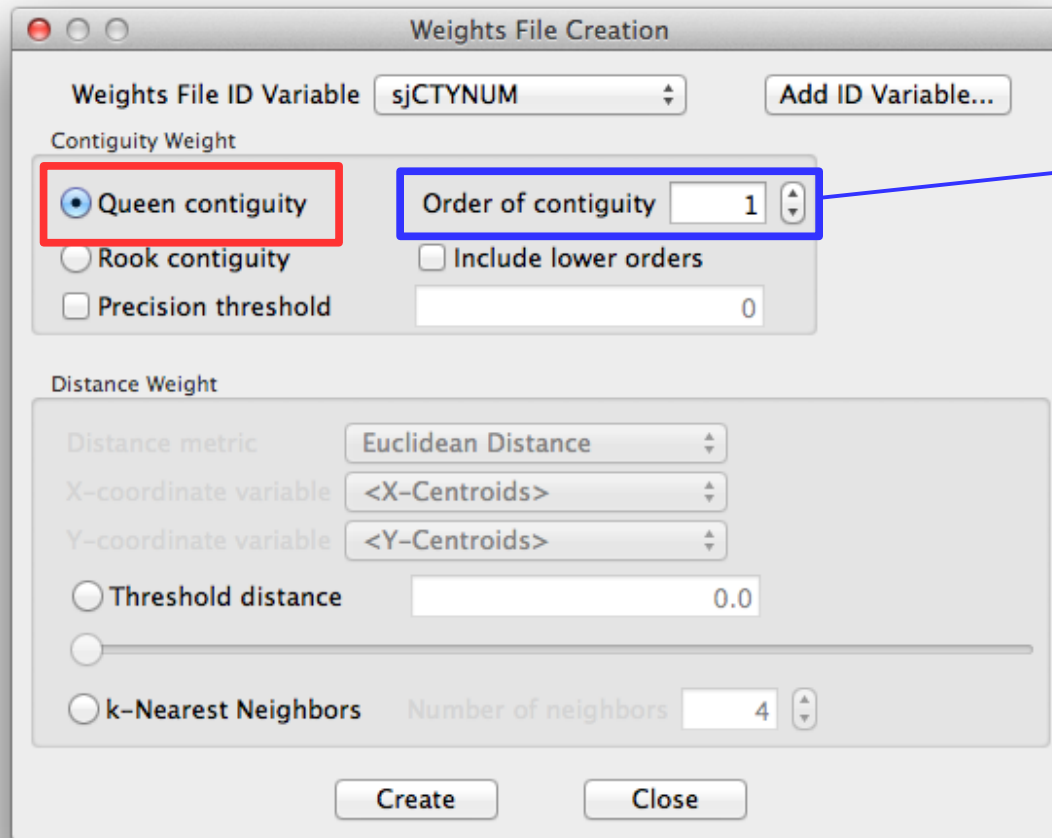
RELATIVE DISTANCE, TOPOLOGY



Queen's case, 1st order neighbors

RELATIVE DISTANCE, TOPOLOGY

GeoDa



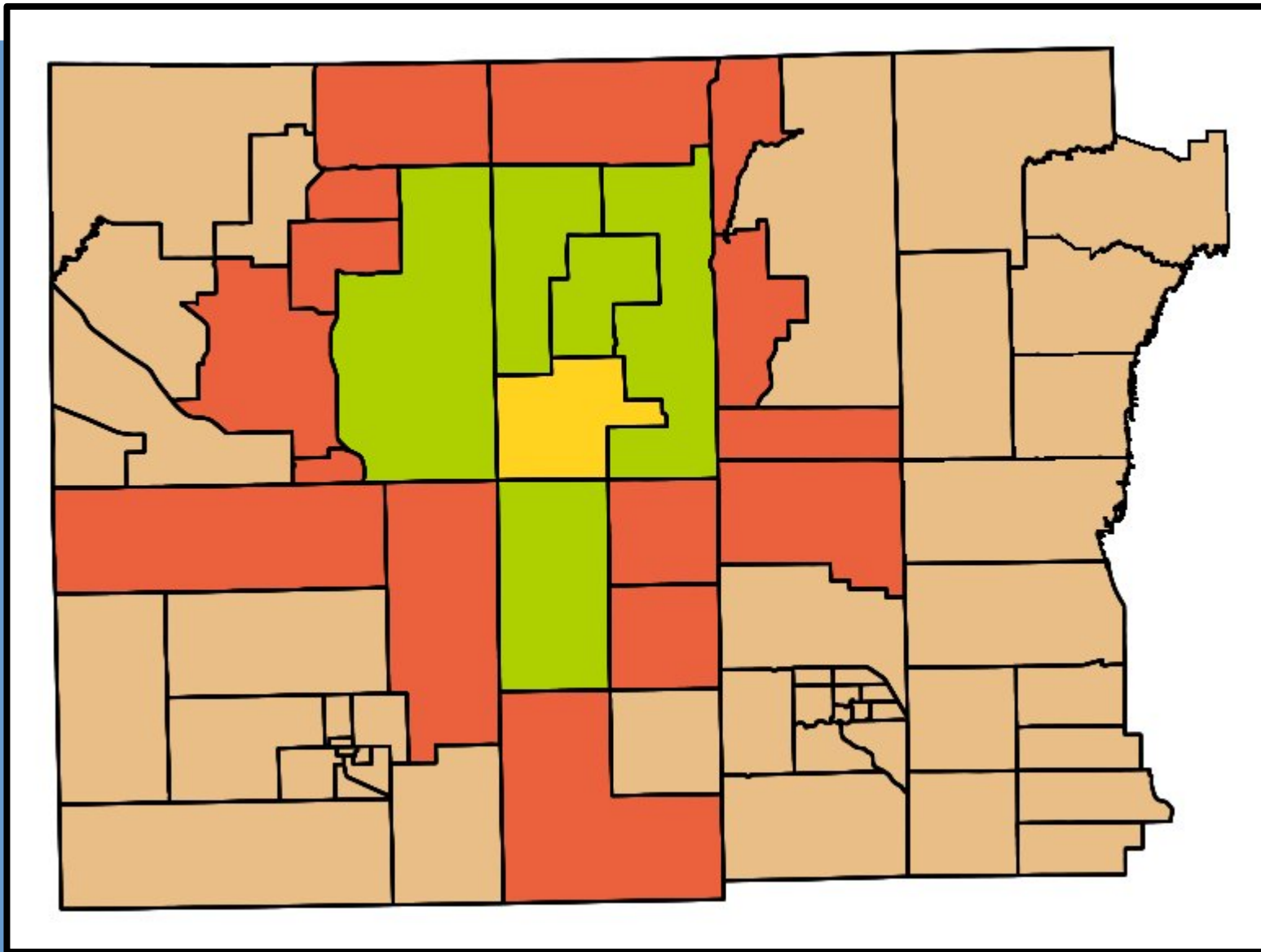
The image shows the 'Weights File Creation' dialog box in GeoDa. The 'Weights File ID Variable' is set to 'sjCTYNUM'. Under 'Contiguity Weight', 'Queen contiguity' is selected (highlighted with a red box), and 'Order of contiguity' is set to 1 (highlighted with a blue box). A blue arrow points from the 'Order of contiguity' box to the text 'Order Number:'. Other options include 'Rook contiguity', 'Precision threshold', 'Include lower orders', 'Distance Weight' (Euclidean Distance), 'X-coordinate variable' (<X-Centroids>), 'Y-coordinate variable' (<Y-Centroids>), 'Threshold distance' (0.0), and 'k-Nearest Neighbors' (Number of neighbors: 4). Buttons for 'Create' and 'Close' are at the bottom.

Order Number:

All neighborhood weight matrices are Row Standardized

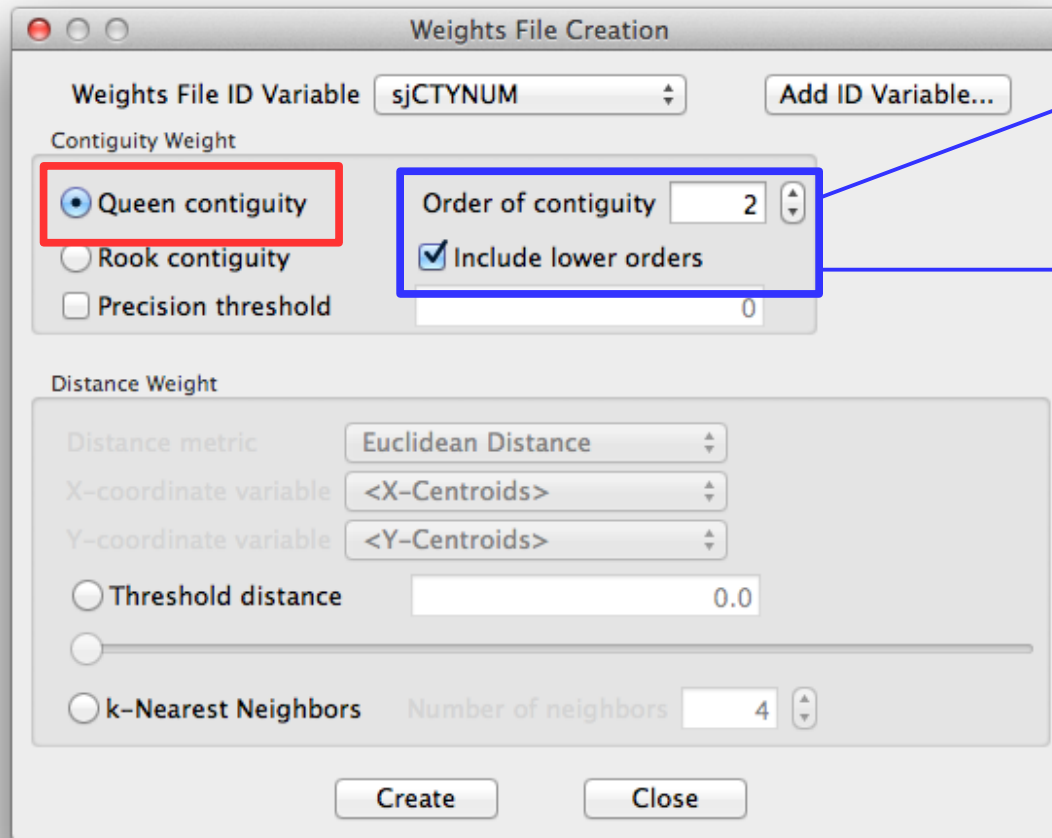
Queen's case, 1st order neighbors

RELATIVE DISTANCE, TOPOLOGY



Rook's case, 2nd order neighbors

RELATIVE DISTANCE, TOPOLOGY



The image shows a 'Weights File Creation' dialog box. It has a title bar with standard window controls. The 'Weights File ID Variable' is set to 'sjCTYNUM' with an 'Add ID Variable...' button. Under 'Contiguity Weight', 'Queen contiguity' is selected (highlighted with a red box), 'Rook contiguity' is unselected, and 'Precision threshold' is unselected. A blue box highlights the 'Order of contiguity' set to 2 and the 'Include lower orders' checkbox which is checked. Blue arrows point from these elements to text on the right. The 'Distance Weight' section is dimmed, showing 'Euclidean Distance' as the metric, '<X-Centroids>' for the X-coordinate variable, '<Y-Centroids>' for the Y-coordinate variable, a 'Threshold distance' of 0.0, and 'k-Nearest Neighbors' with 4 neighbors. 'Create' and 'Close' buttons are at the bottom.

Order Number:
Up to K orders

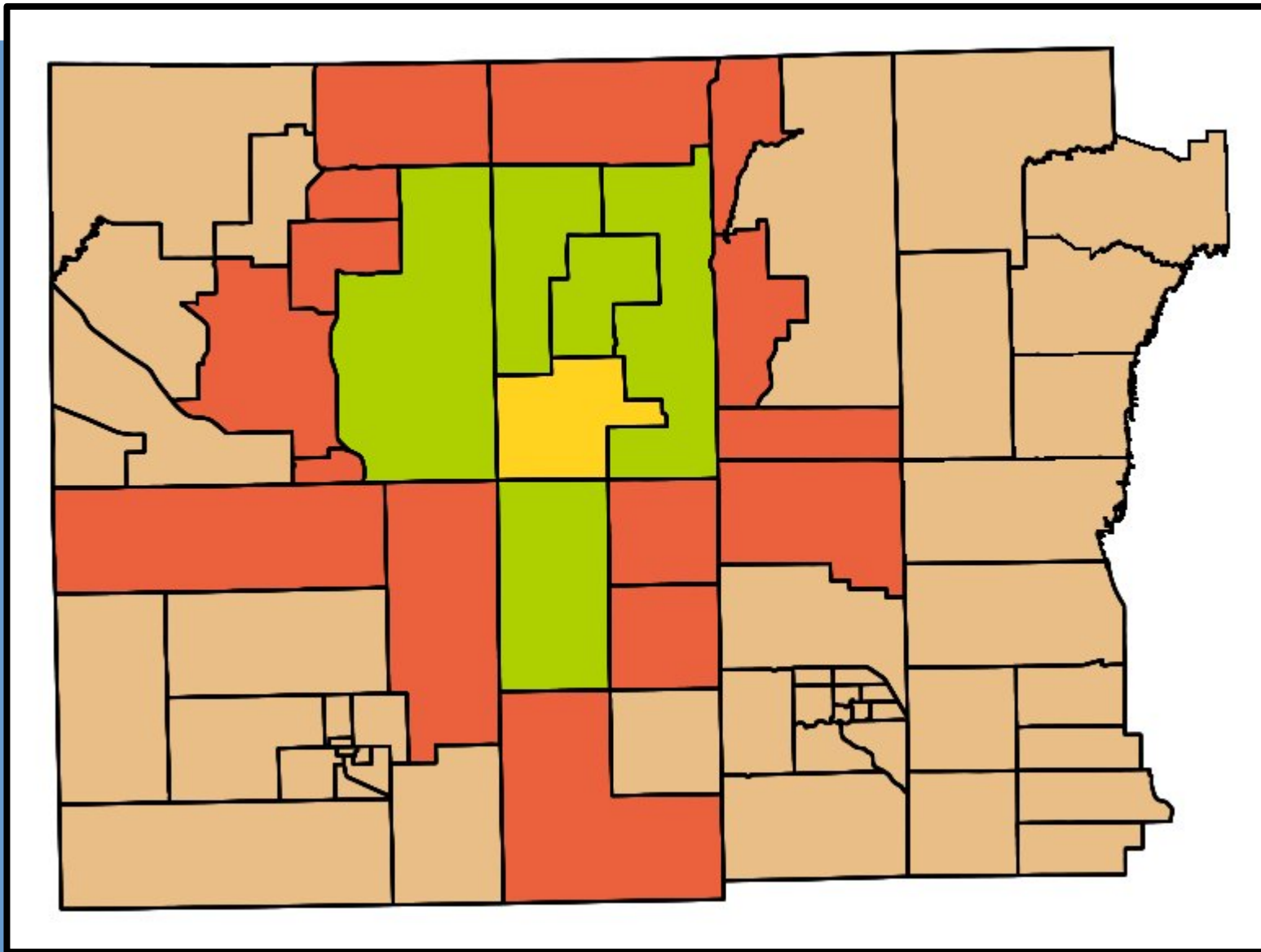
Includes 1st order
neighbors as well

All neighbors are
treated as equals!

All neighborhood
weight matrices are
Row Standardized

Rook's case, 2nd order neighbors

RELATIVE DISTANCE, TOPOLOGY



Queen's case, 2nd order neighbors

Neighborhoods

- How do I decide how to define my neighborhood?... and how to weigh my neighbors?
 - Unfortunately, no simple answer to this question
 - Theory-driven approach
 - Cite previous literature
 - Empirical approach
 - Rules of thumb: there are many, “each worse than the previous one”
 - ESDA: explore, optimize

Empirical Approaches

- Exploratory Spatial Data Analysis
 - Explore your data!
 - Choose a spatial autocorrelation metric
 - Test the metric over multiple neighborhood definitions
 - Test the metric over multiple neighborhood parameters
 - Chances are that your results will be semi-consistent
 - If not, eureka!... or oh no!

Empirical Approaches

- Exploratory Spatial Data Analysis
 - Example: test data for global spatial autocorrelation using multiple neighborhood definitions
 - Calculate Moran's I with multiple neighbor definitions
 - Inverse Distance, Rook Contiguity, Queen Contiguity
 - KNN, Inverse Distance, Inverse Distance sq
 - 1NN, 2NN, 5NN, 10NN, 20NN

Robustness Tests

- Run your analysis using multiple neighborhood definitions or neighbor relationships
 - Similar to testing over multiple scales
 - If your results are similar, they are robust with regard to neighbor(hood) definition!

Empirical Approaches

- Optimization
 - Find the neighborhood definition that produces the most “extreme” results
 - Beware!
 - Circular logic

Neighborhood Weight Matrix

- Entries in the neighborhood weight matrix describe the relationships between observations
 - These do not have to be based on geographic relationships!
 - e.g., network connectivity
 - e.g., sociodemographic similarity
 - Some software allows you to import your own matrix

Keywords

- Neighborhood, neighbors
- Binary, continuous
- Absolute, relative distance
- Neighborhood weight matrix
- Row standardization
- ESDA, robustness