School of Geography and Earth Sciences McMaster University

Applied Spatial Statistics

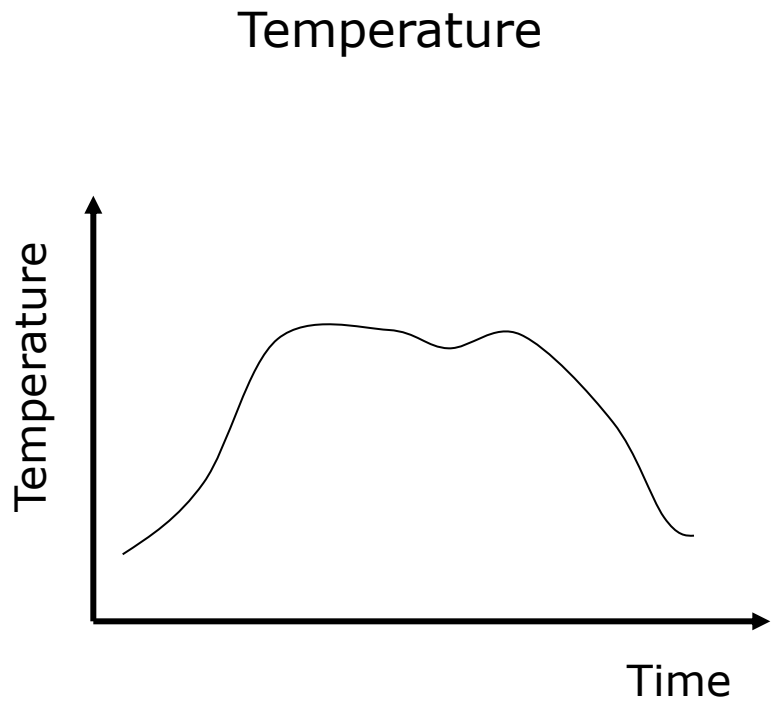
Area Data
I & II

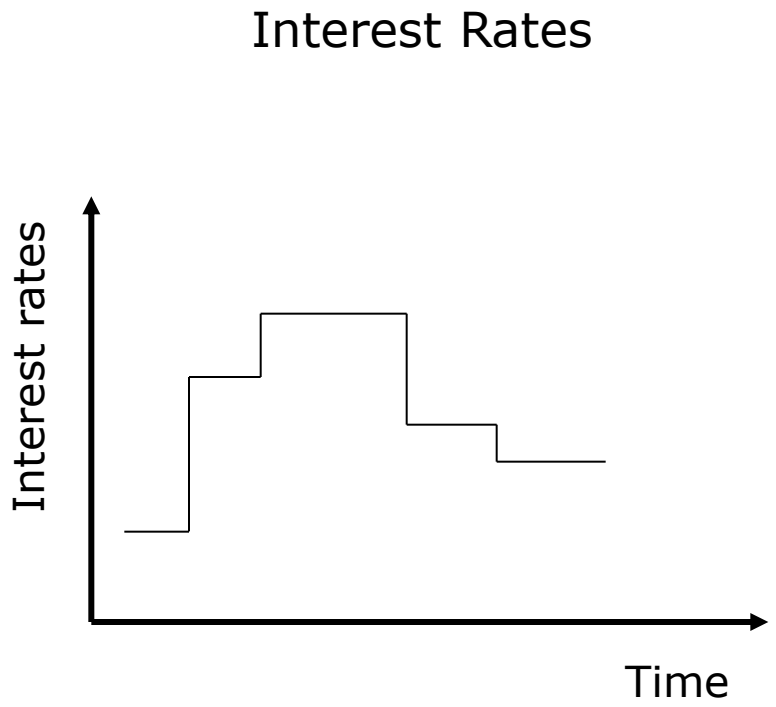
This session:

- **Area Data**
  - Definitions
  - Visualization
  - Spatial proximity matrices

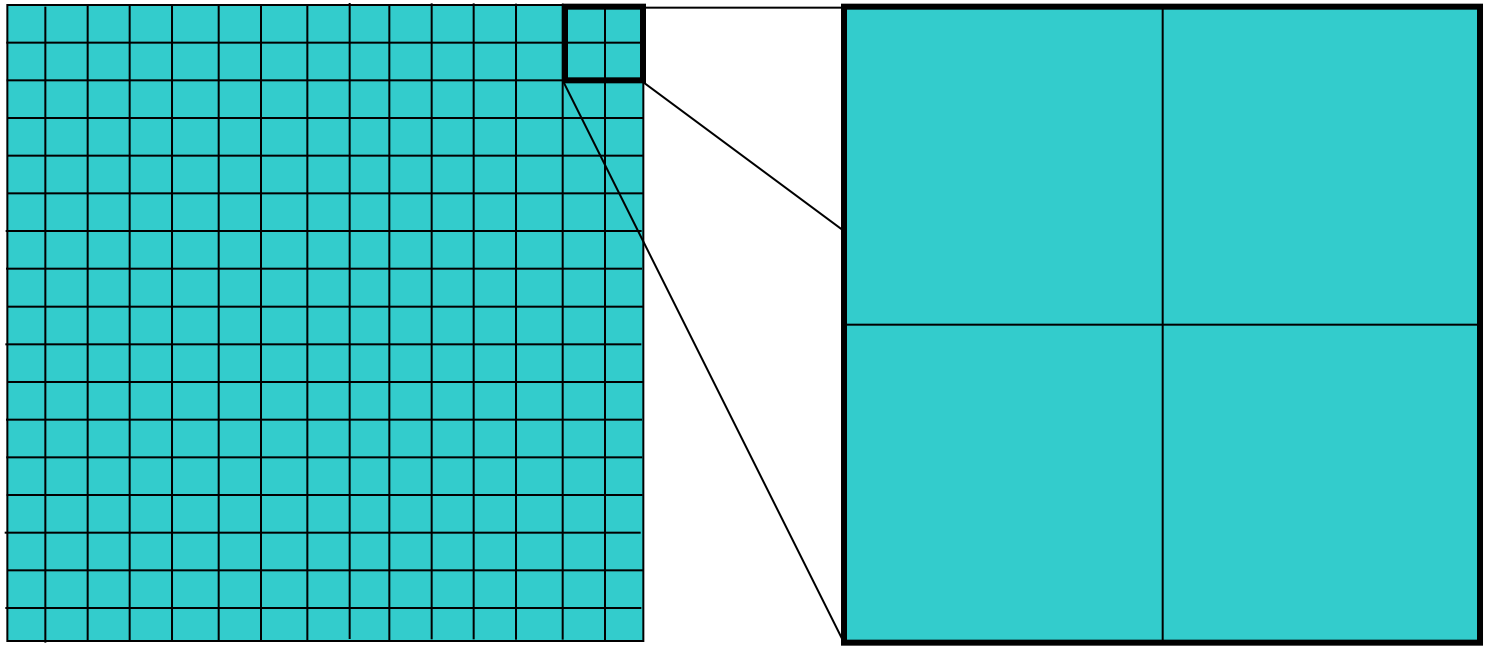
The difference between spatially continuous data and area data

- Time series





Scale of Analysis



Area Data

- Examples of data in the social sciences
  - Urban population density
    - What is the distribution of population in a city?
    - Links to locational theory
  - Residential insurance coverage
    - Do insurance companies redline neighborhoods?

Area Data

- Examples of data in the social sciences
  - Urban crime
    - Does criminal behavior follow spatial patterns?
    - Are there spillover effects?
  - Geographic distribution of minorities
    - Is there spatial segregation of minorities?

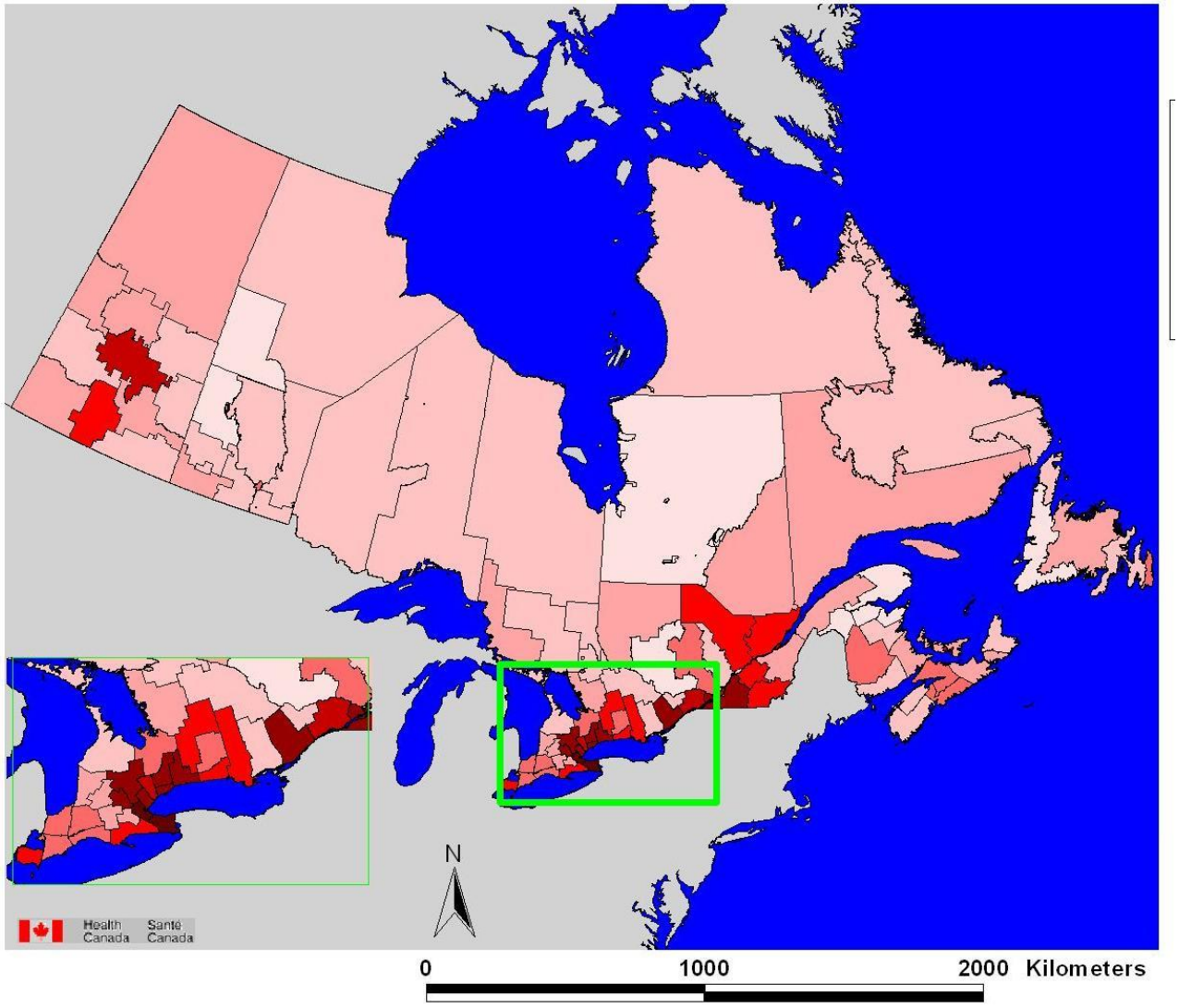
Area Data

- Examples of data in the social sciences
  - Voter turnout
    - What explains voter turnout in elections?

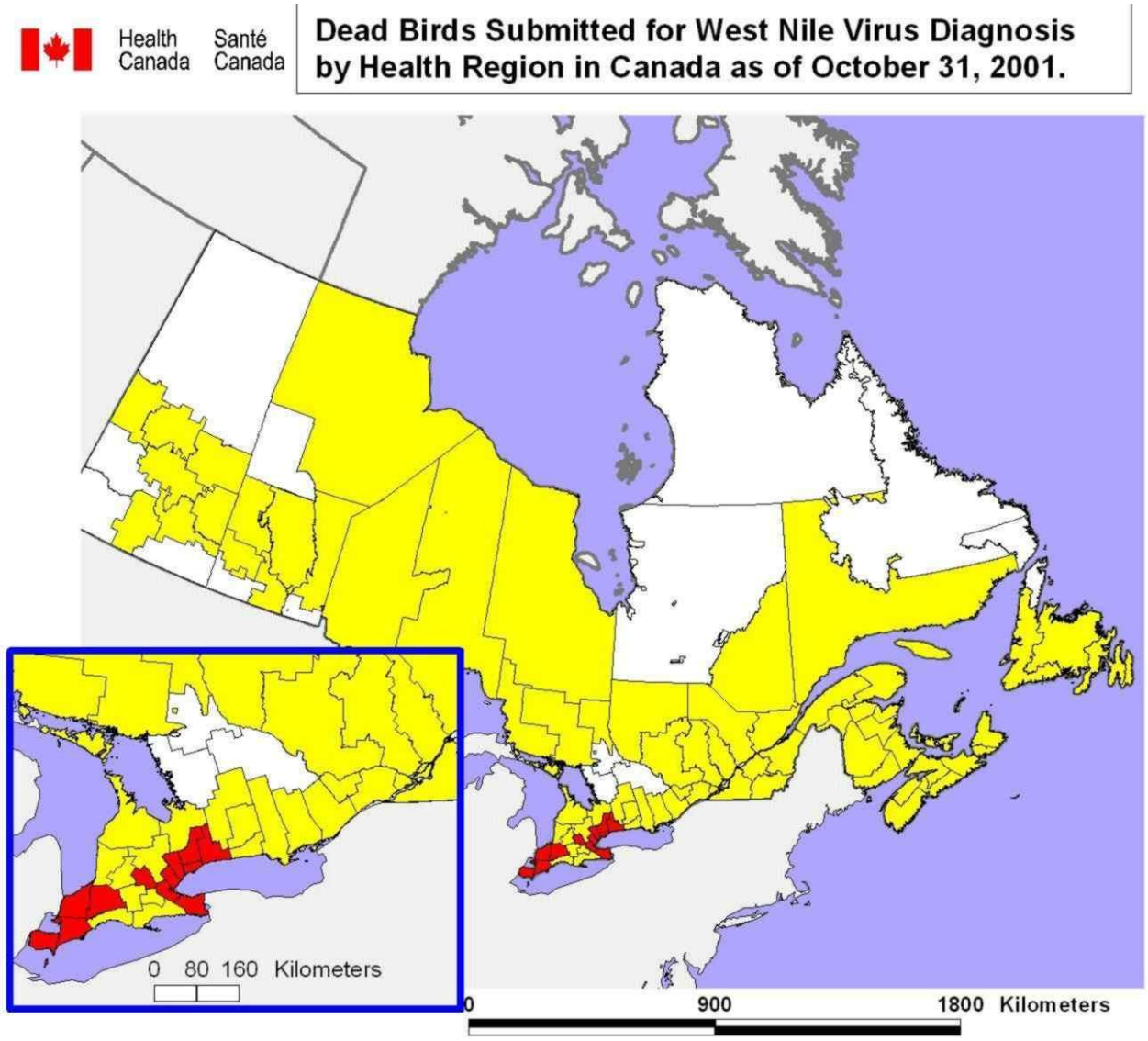
Area Data

- Area data in the environmental sciences
  - Agricultural experiments in regularly tilled fields
  - Species habitat
  - Intersection between social and environmental sciences
    - NOx spillovers in Europe

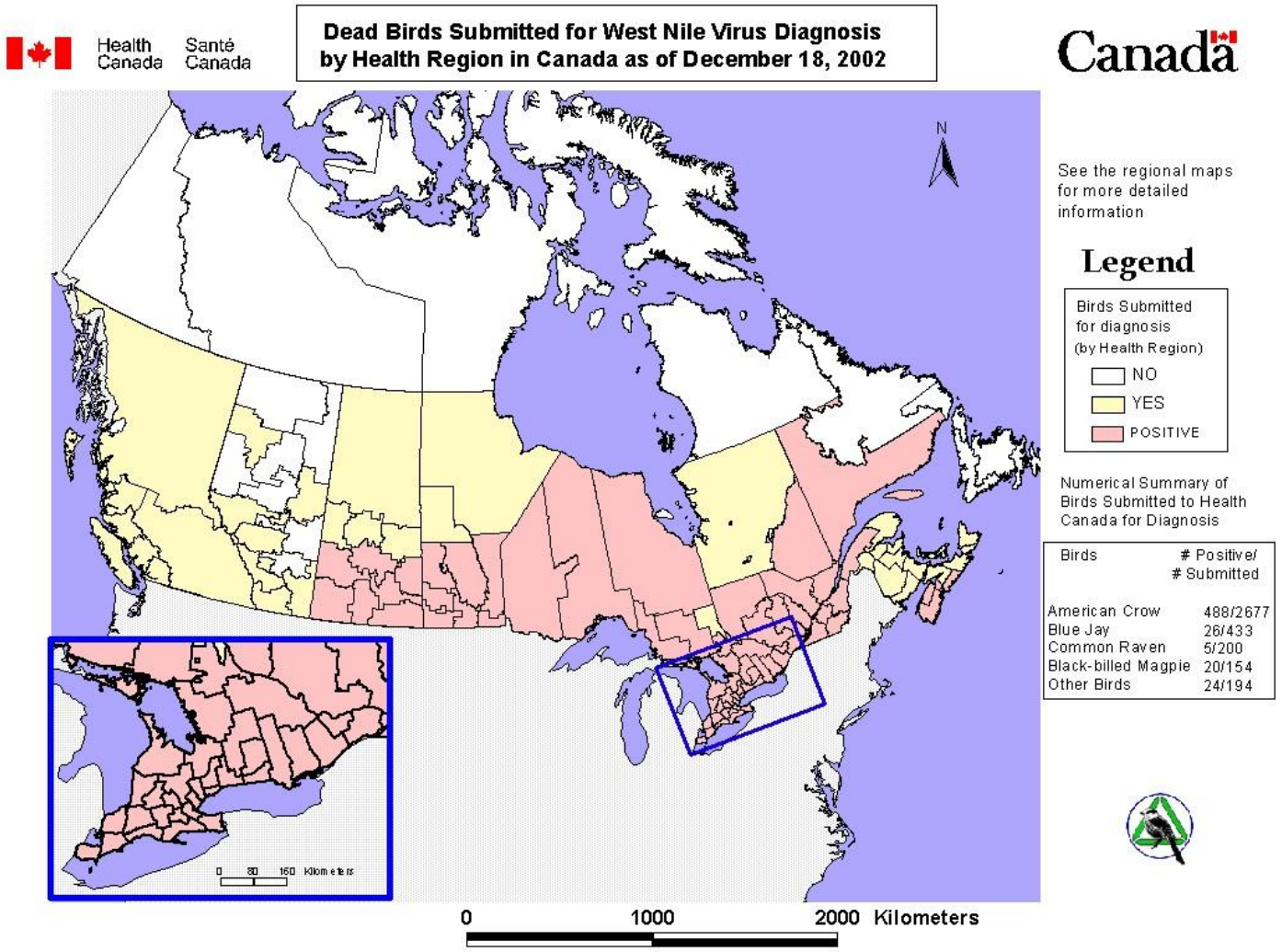
Health Issues:
Number of dead birds reported (2000)



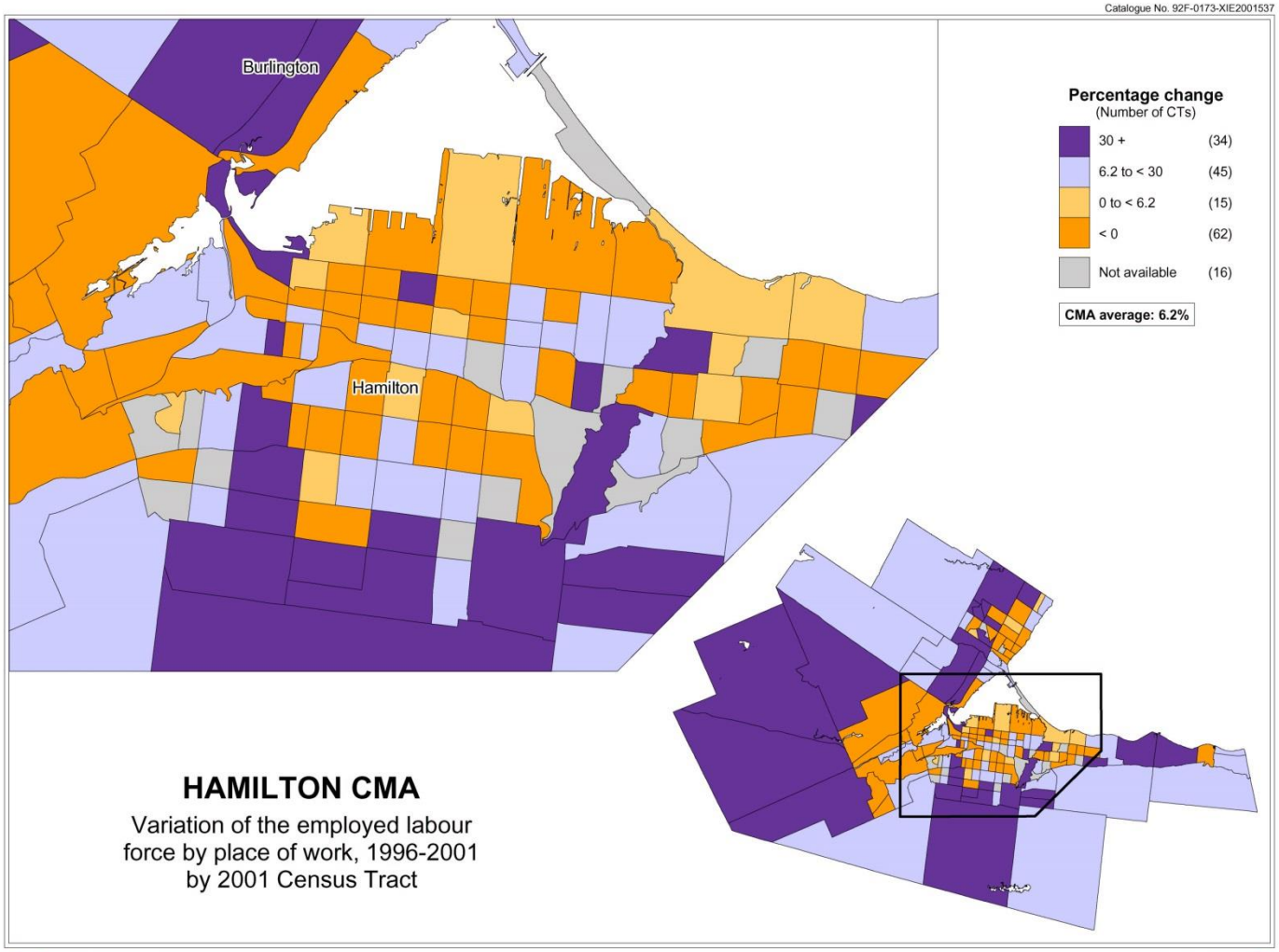
Health Issues:
Dead birds reports



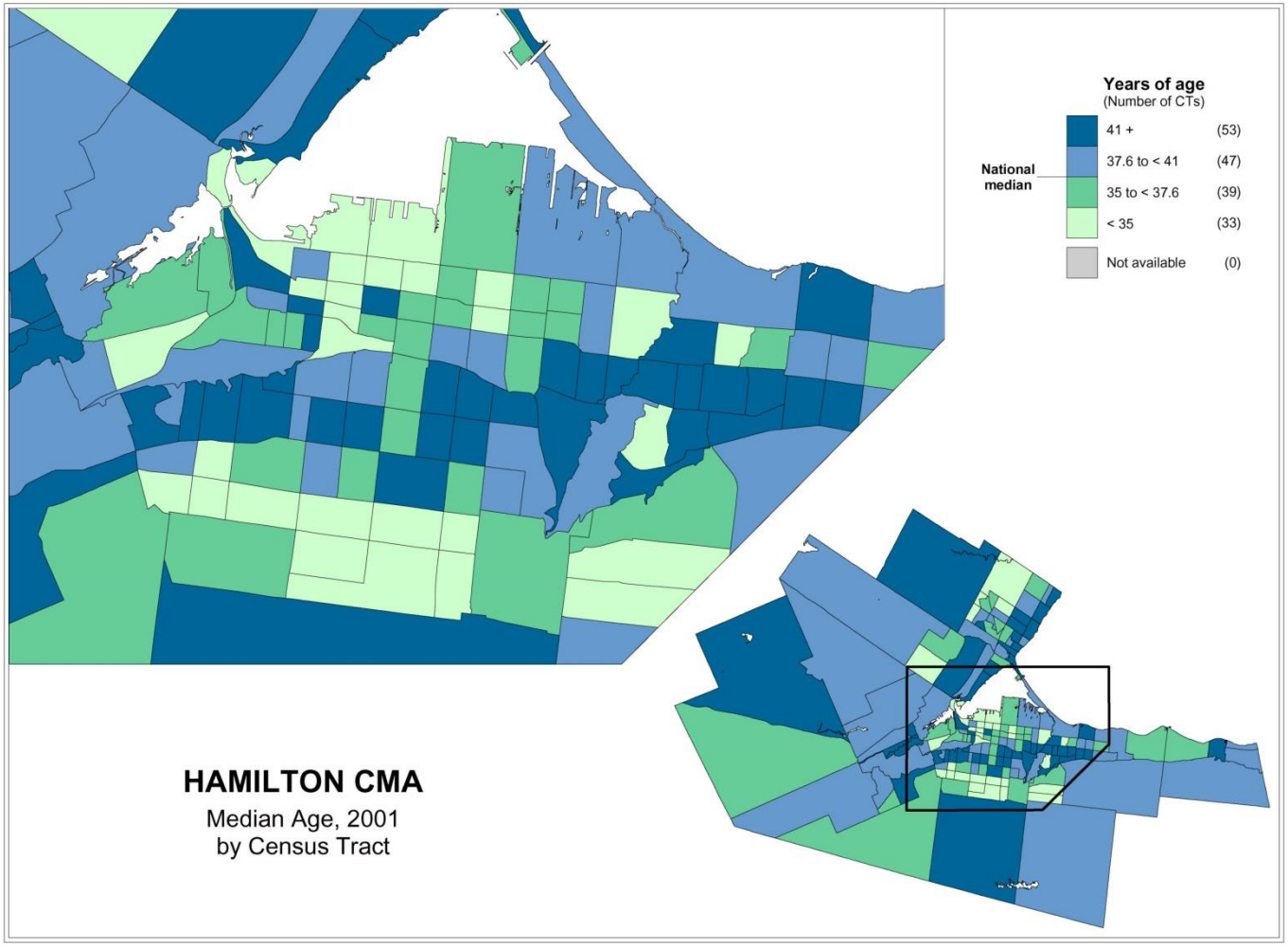
Health Issues:
Dead birds reports



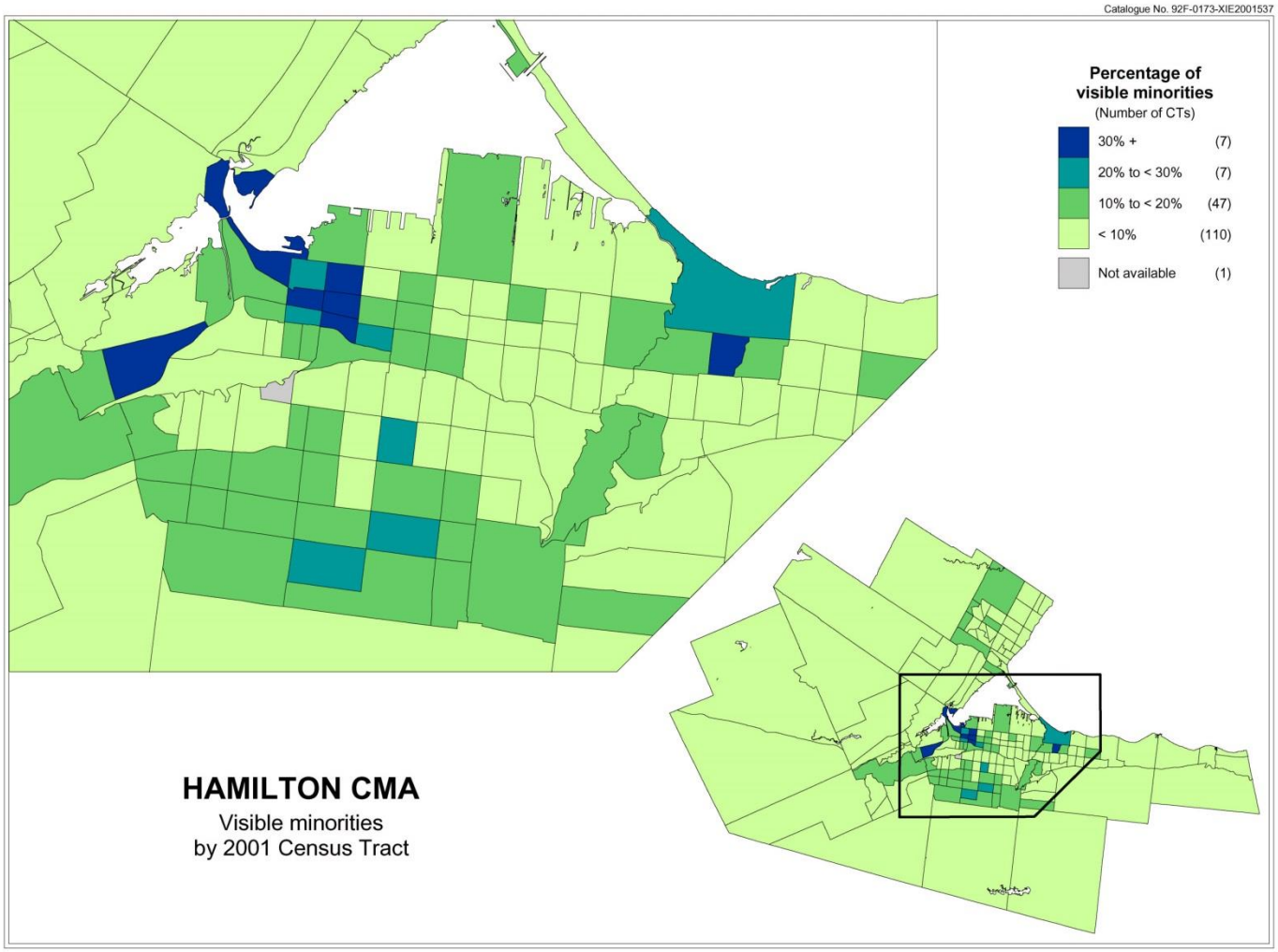
Socio-economic Issues: Employment



Demographic Issues: Aging



Demographic Issues: Ethnicity



Goals of Applications: Area data

- The description of important features
- Estimation of an average value over large areas
- The estimation of an average value over small areas
- Explanation of patterns and trends

(first order and second order effects)

Analysis of Area Data

- Main objective is to explain and possibly to forecast
- Analysis follows the usual sequence:
  - Visualization and exploration
  - Modeling
  - First order and second order effects

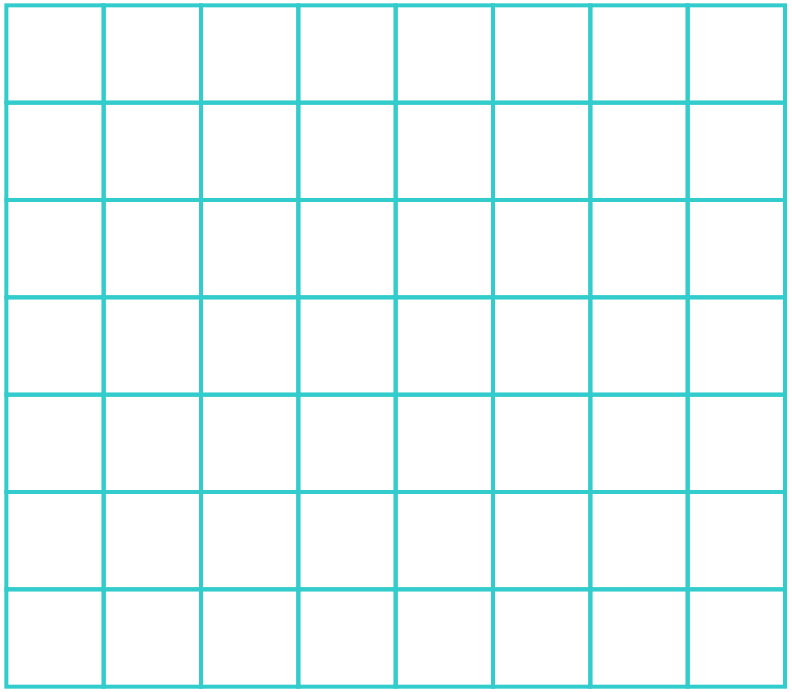
Definitions

- Region
- Regular lattice
- Irregular lattice
- Zones (Areas)
- Attributes
- First order effects
- Second order effects

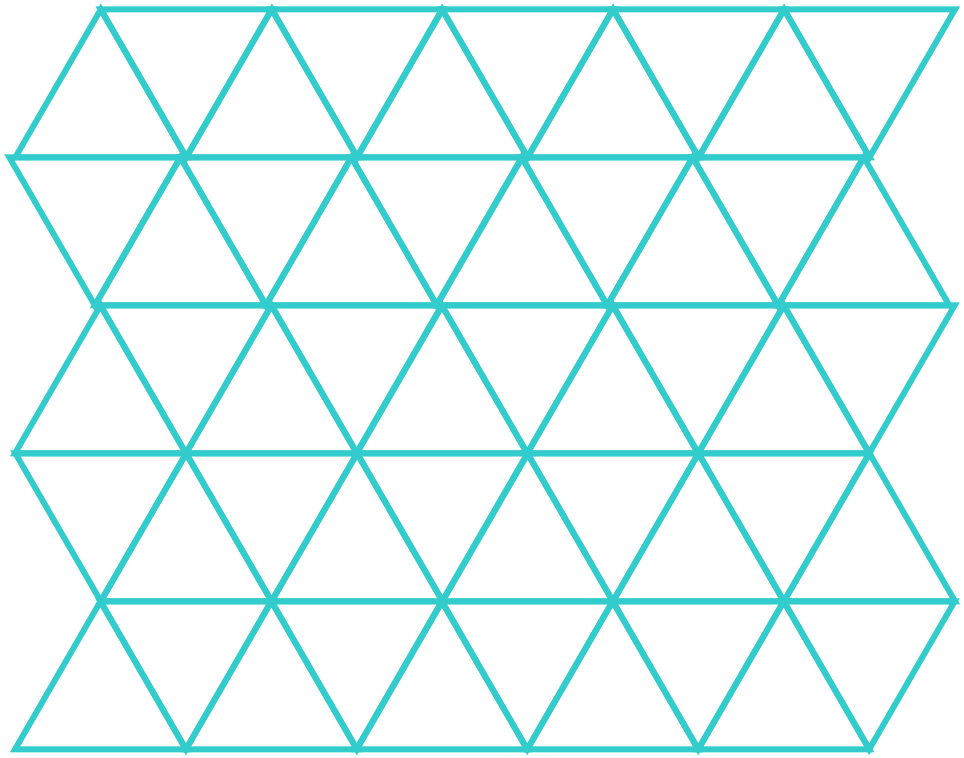
Definitions: Region

- Region (R) – Specific area over the surface of the earth that is of interest

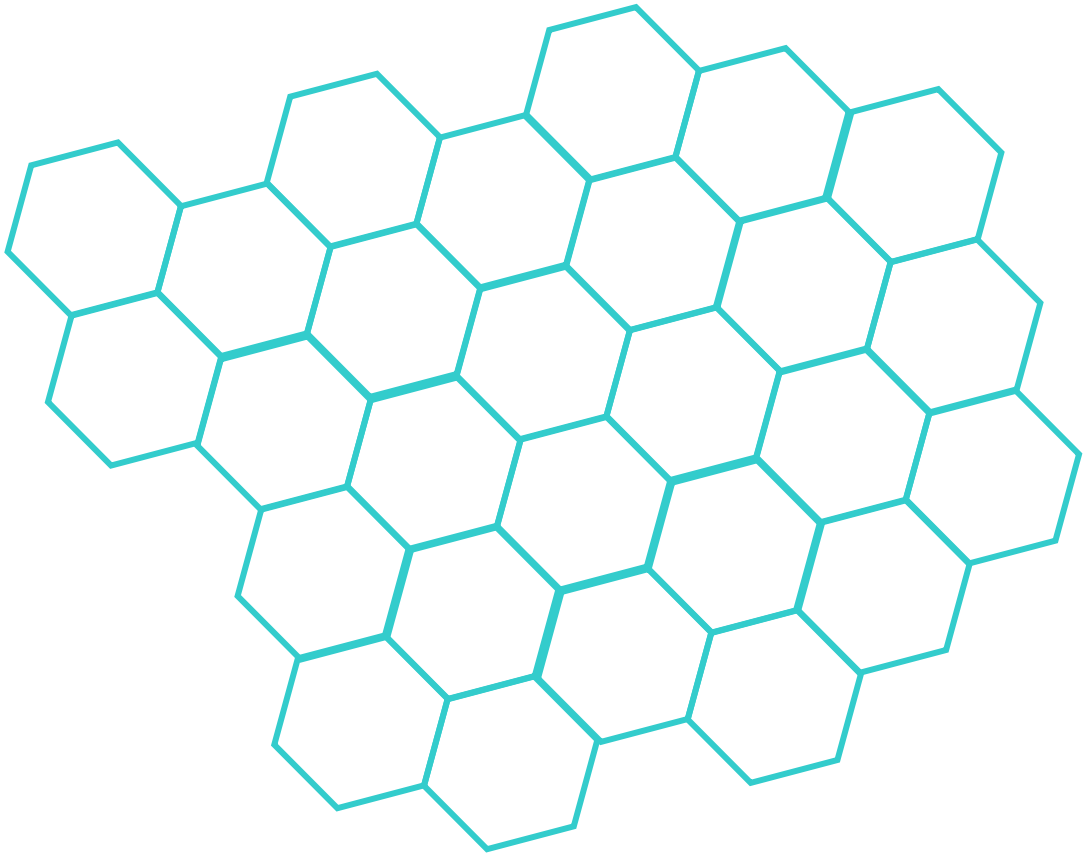
Definitions: Regular Lattice



Definitions: Regular Lattice

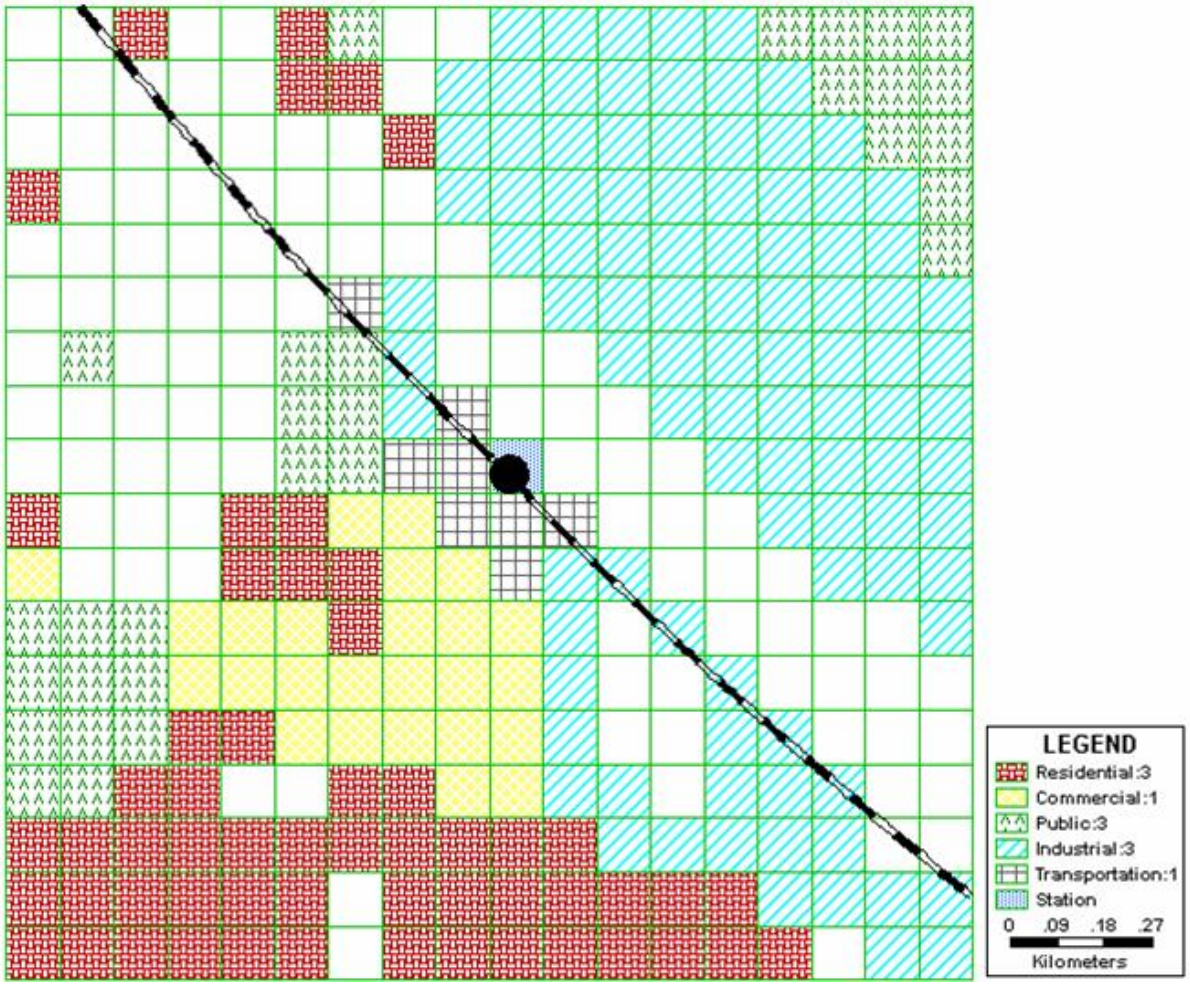


Definitions: Regular Lattice

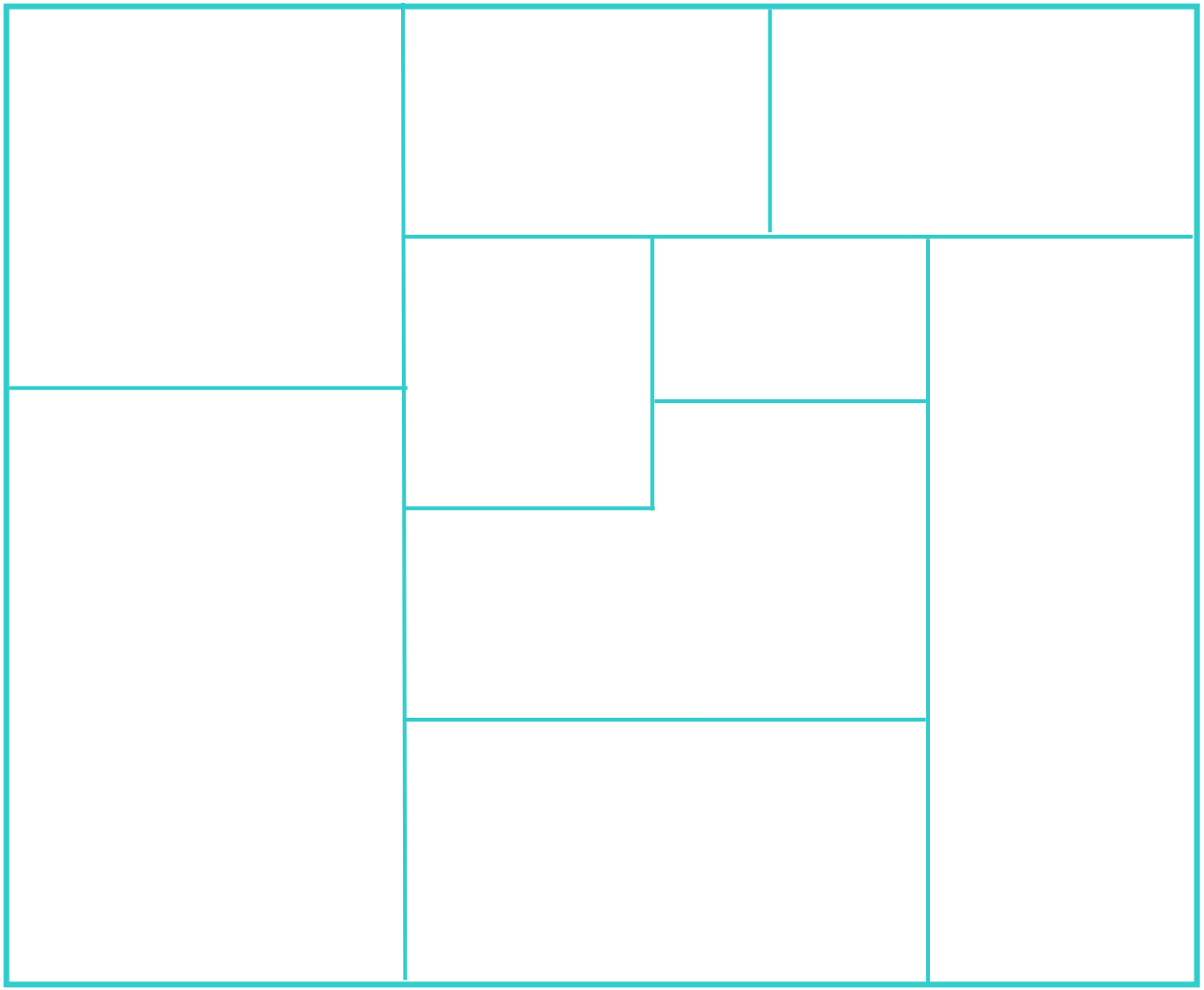


Definitions: Regular Lattice

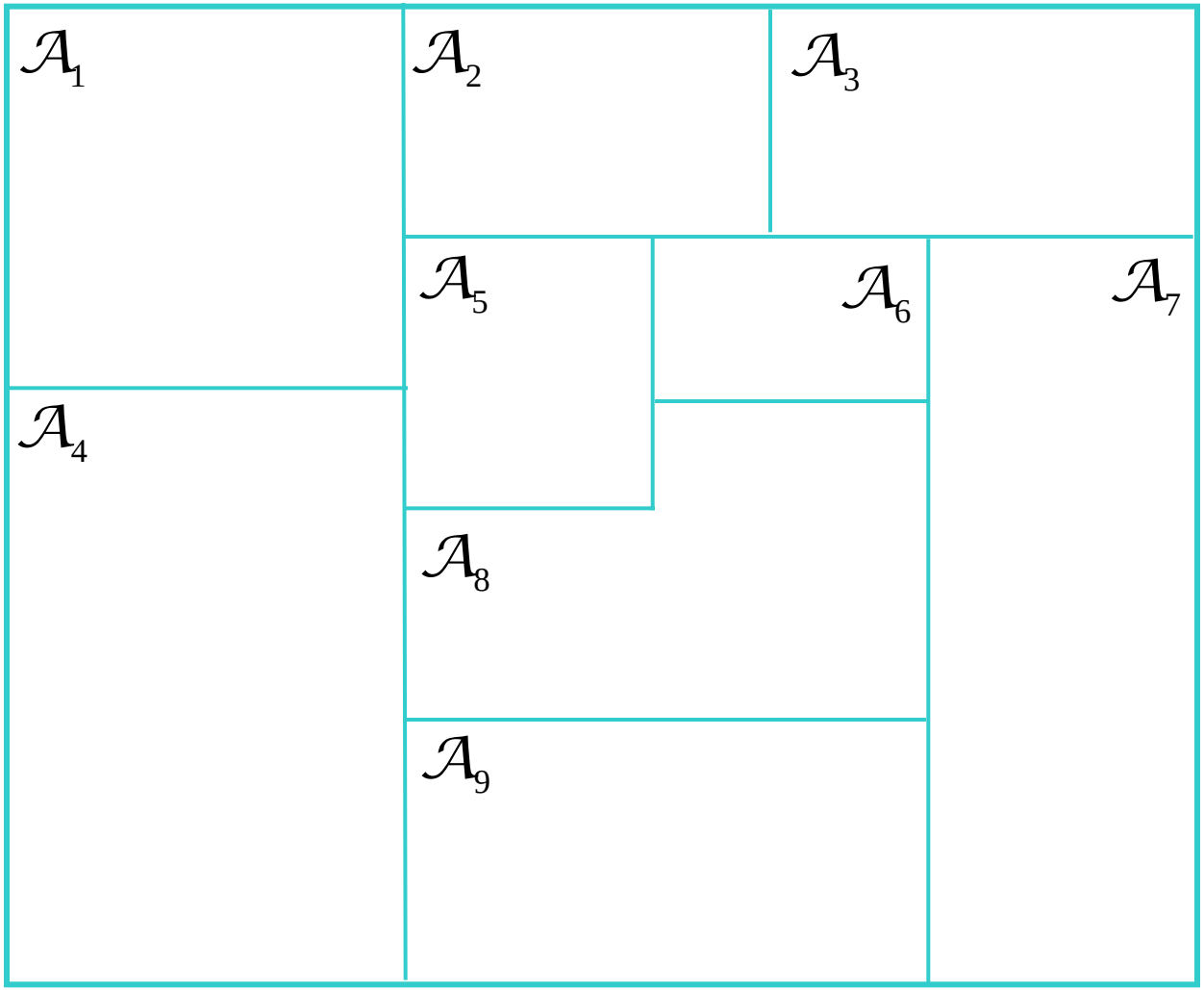
- Example: Land use change



Definitions: Irregular Lattice



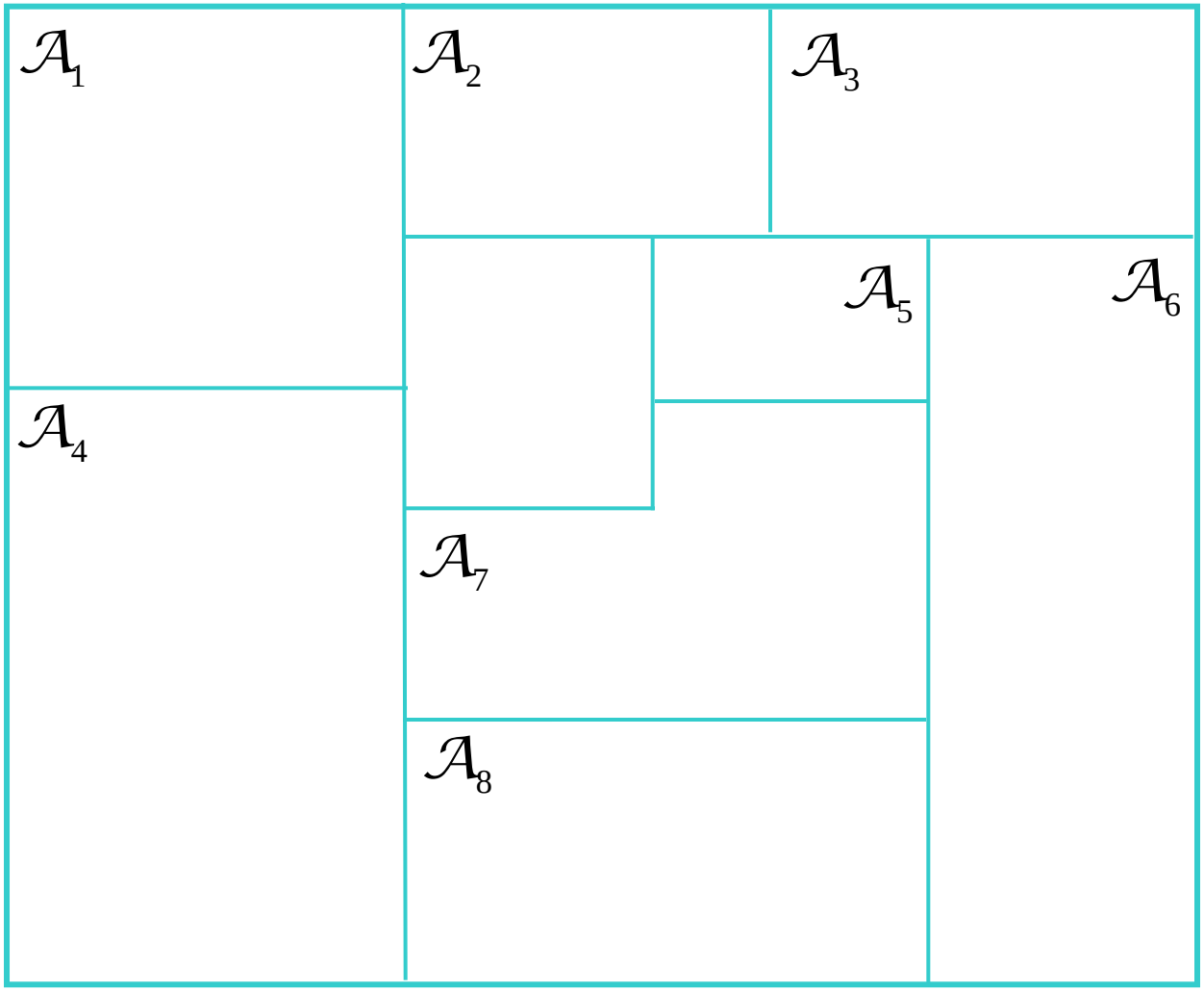
Definitions: Zones (Areas)



Definitions: Zones (Areas)

$$\mathcal{A}_1 \cup \mathcal{A}_2 \cup \cdots \cup \mathcal{A}_3 = \mathcal{R}$$

Definitions: Zones (Areas)



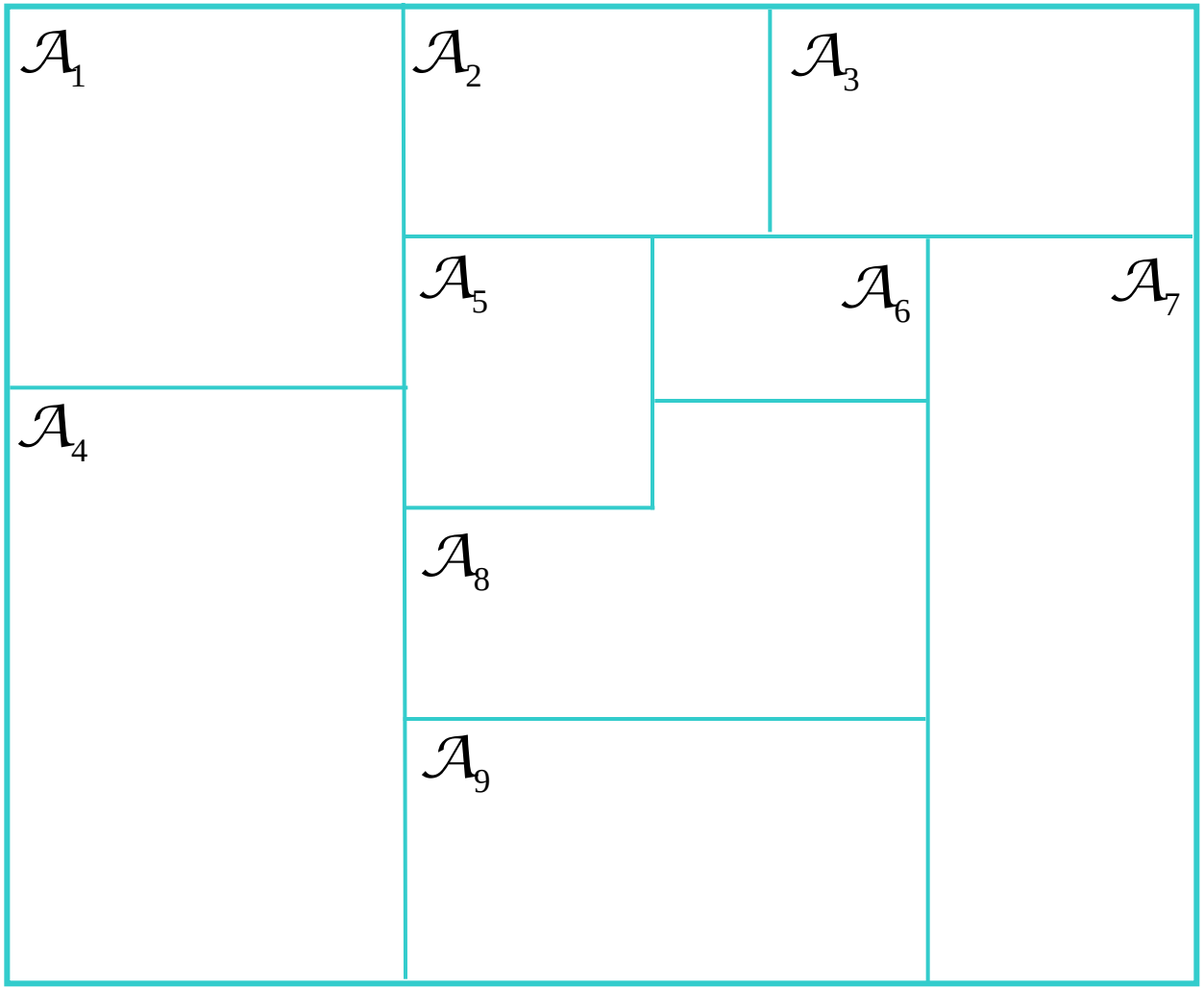
Definitions: Attributes (Area Data)

$Y(\mathcal{A}_i)$

$X_1(\mathcal{A}_i)$

$\vdots$

$X_k(\mathcal{A}_i)$



Definitions: Attributes (Area Data)

$$Y(\mathcal{A}_i) = Y_i$$

$$X_1(\mathcal{A}_i) = X_{1i}$$

$$\vdots$$

$$X_k(\mathcal{A}_i) = X_{2i}$$

First Order Effects

- Expected value of Y for area $\mathcal{A}$

$$\mu(\mathcal{A}_i) = \mu_i = E[Y(\mathcal{A}_i)]$$

(systematic, deterministic)

Second Order Effects

- Covariance between of $Y(\mathcal{A}_i)$ and $Y(\mathcal{A}_j)$

$$COV(Y_i, Y_j)$$

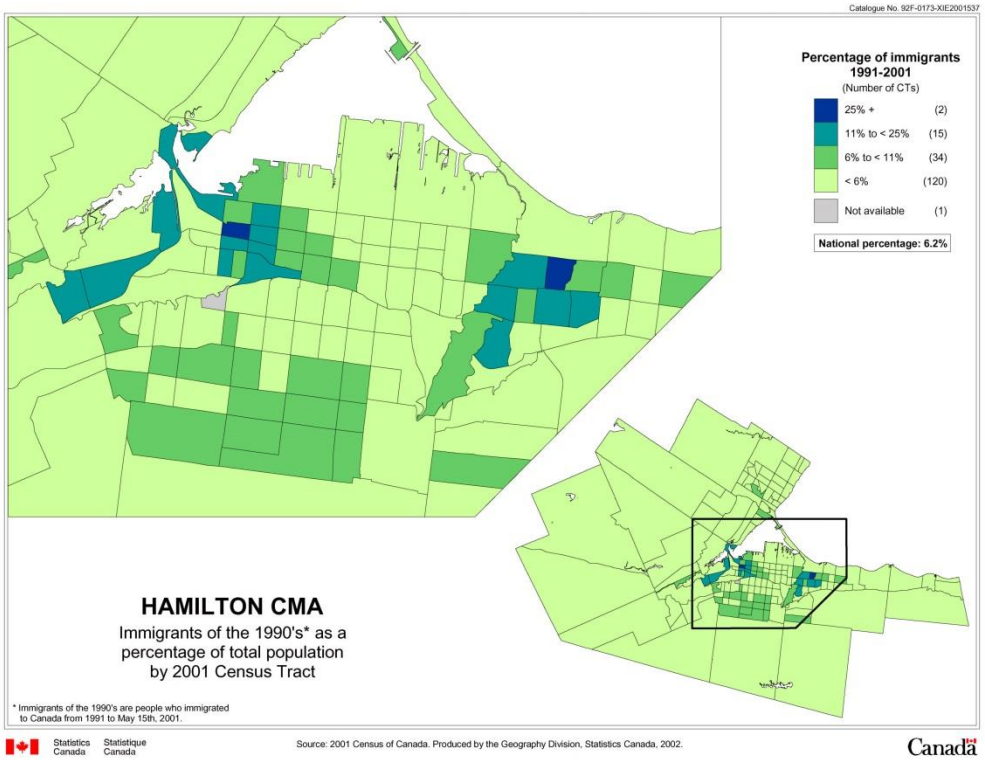
(systematic, deterministic? random, unpredictable?)

Visualization of Area Data

- Proportional symbols
- Choropleth maps
- Density equalized maps

Visualization: Choropleth Maps

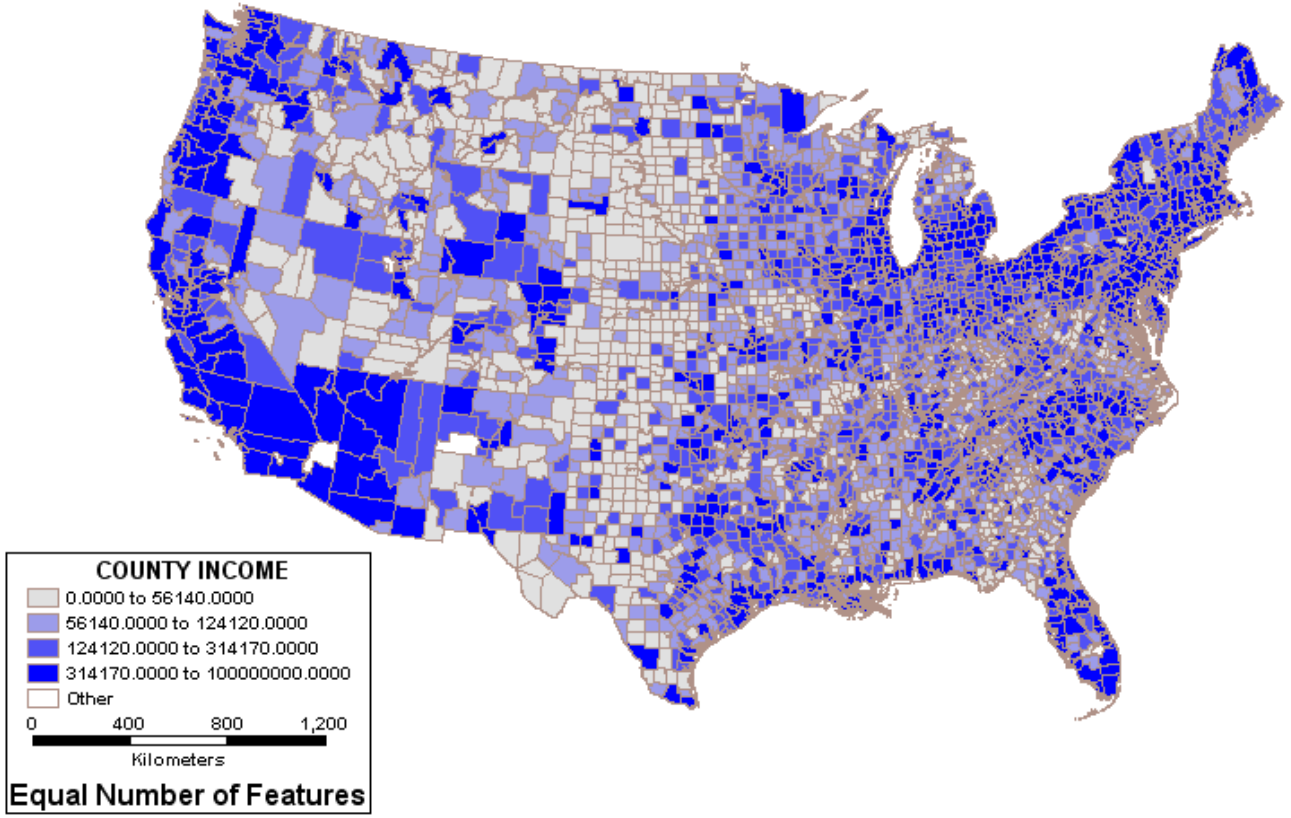
- Area colors depend on the value of the attribute



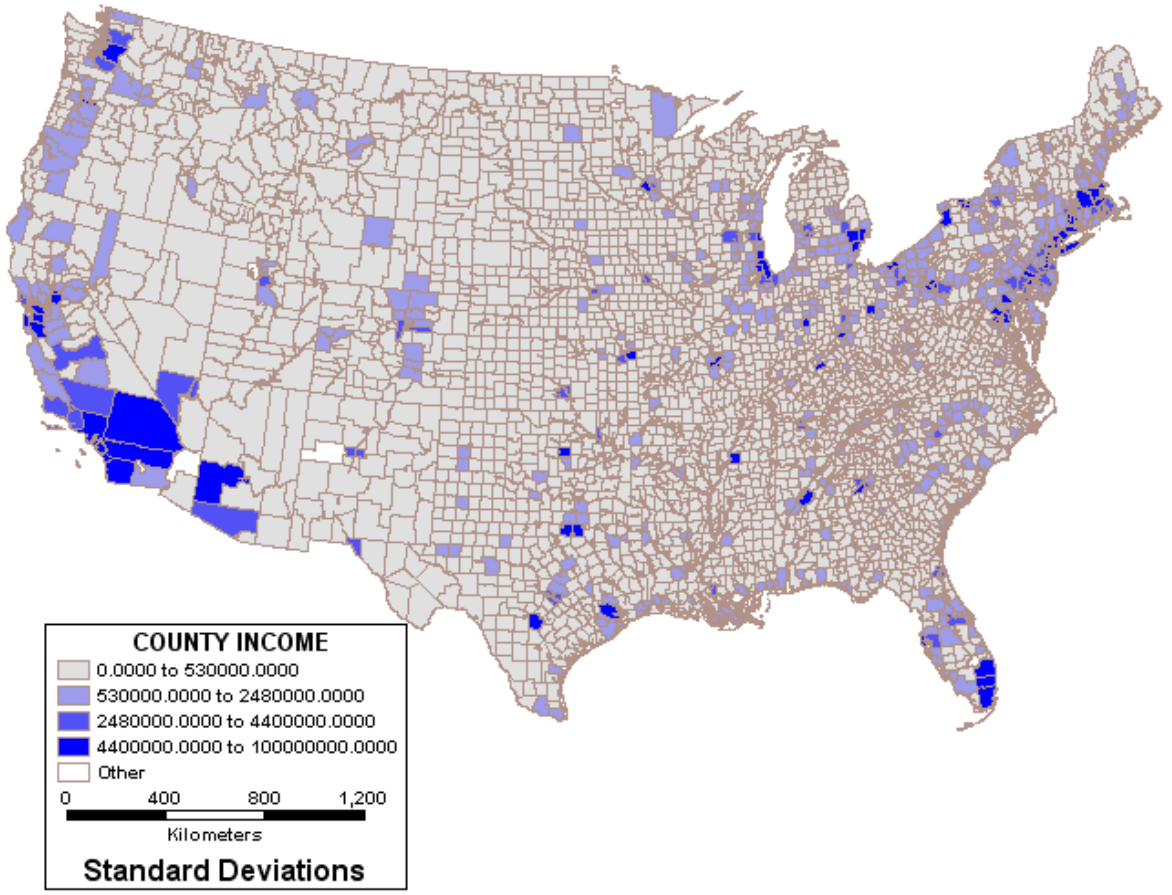
Visualization: Choropleth Maps

- Some issues
  - Selection of class intervals
  - Large areas – small areas
  - Interpretability

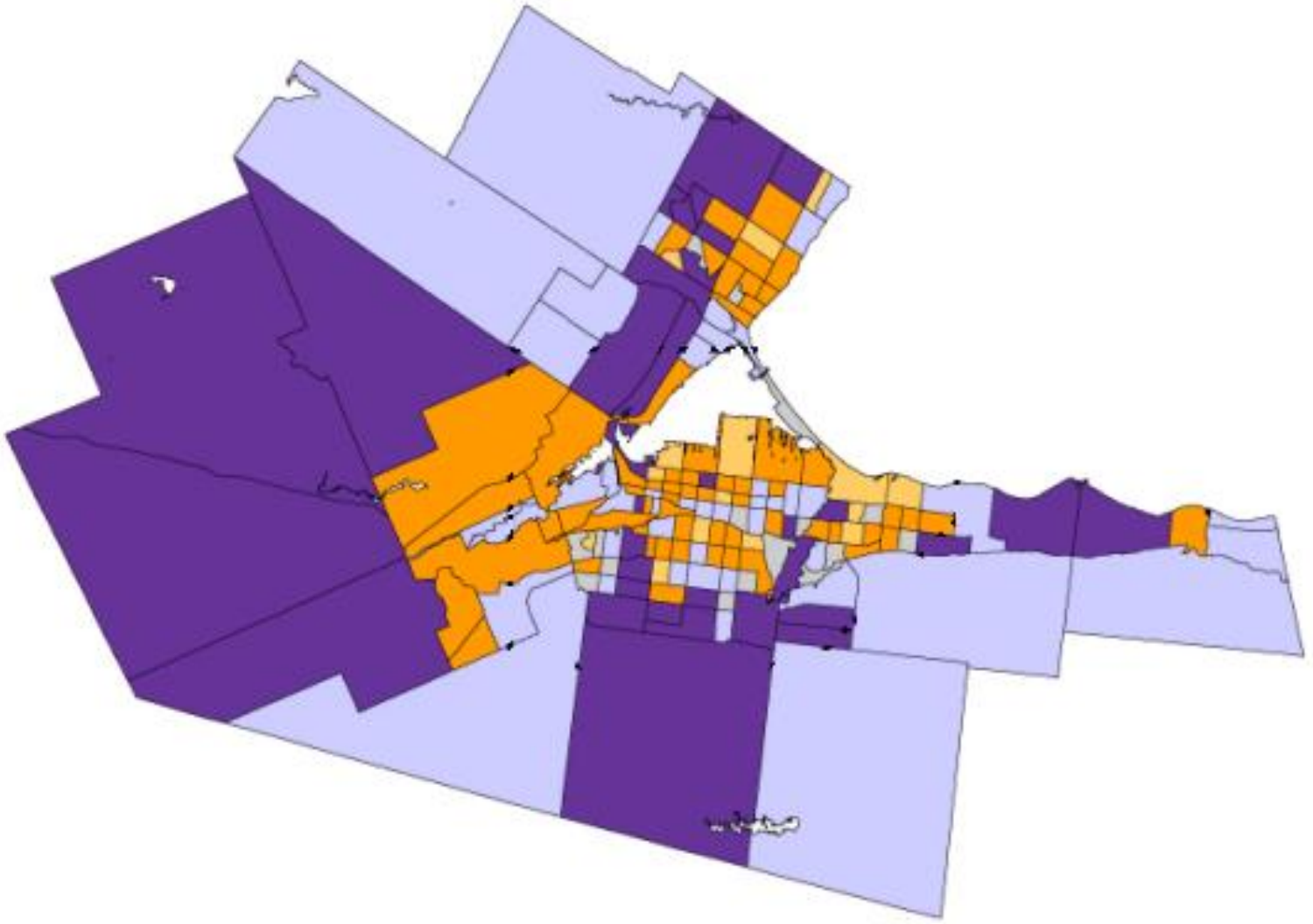
Choropleth Maps: Selection of Class Intervals



Choropleth Maps: Selection of Class Intervals



Choropleth Maps:
Large Zones – Small Zones

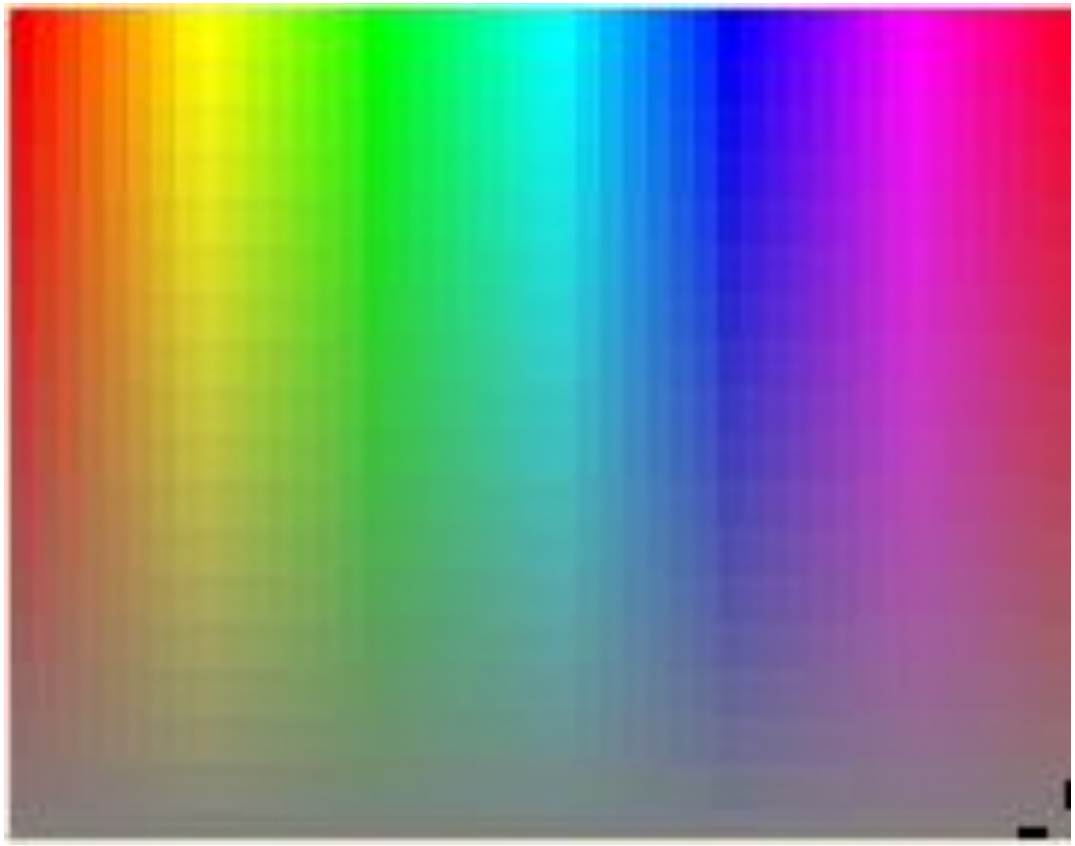


Choropleth Maps: Interpretability

- Number of classes
- Positive – negative attribute values

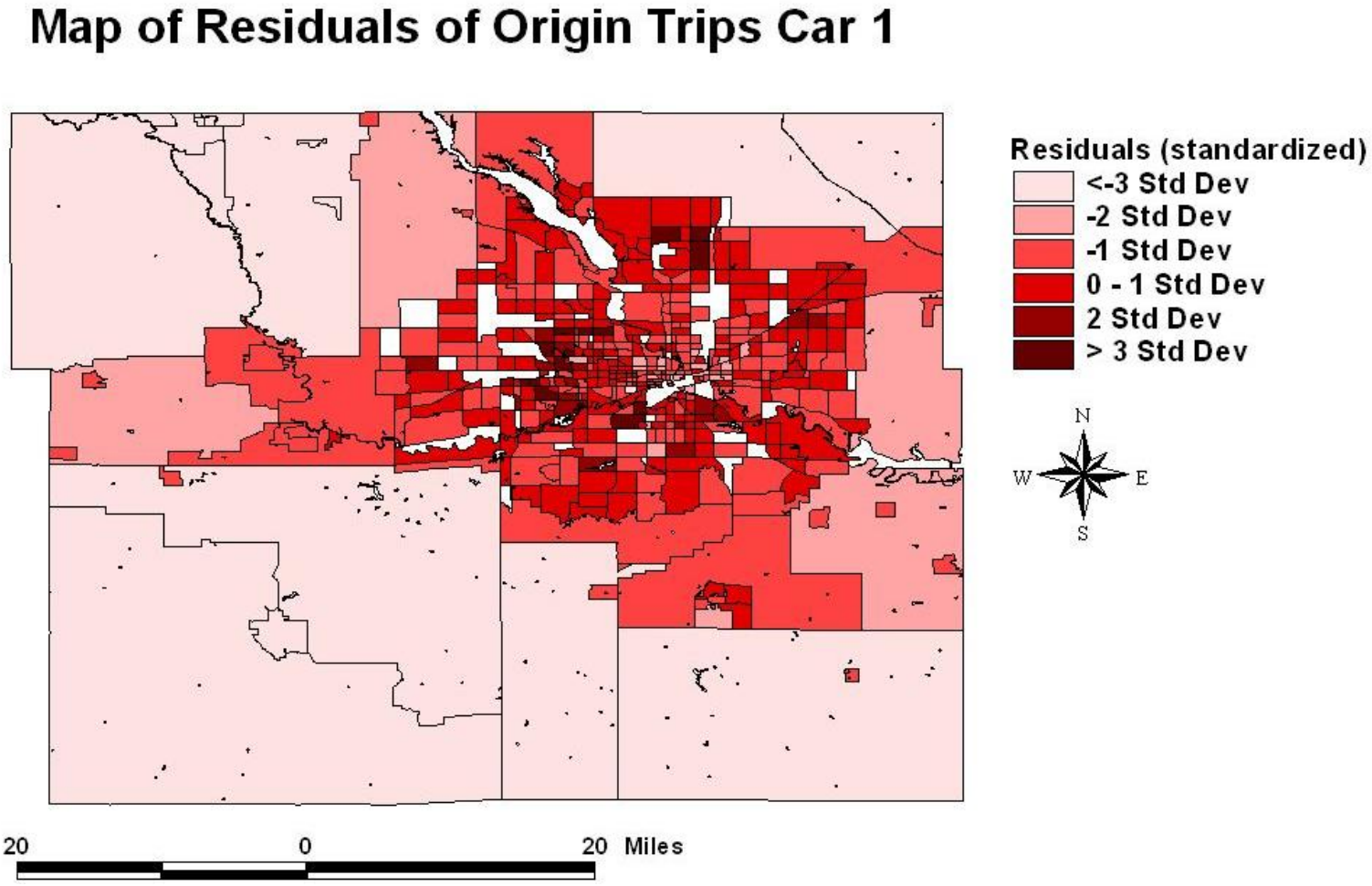
Choropleth Maps: Interpretability

- Number of classes



Choropleth Maps: Interpretability

- Positive – Negative attribute values



Visualization:
Density Equalized Maps

- Example

Fig. 7.2 Density equalised map of unemployment in Britain, 1988

Exploration of Area Data

- Spatial Moving Averages
- Kernel Estimation

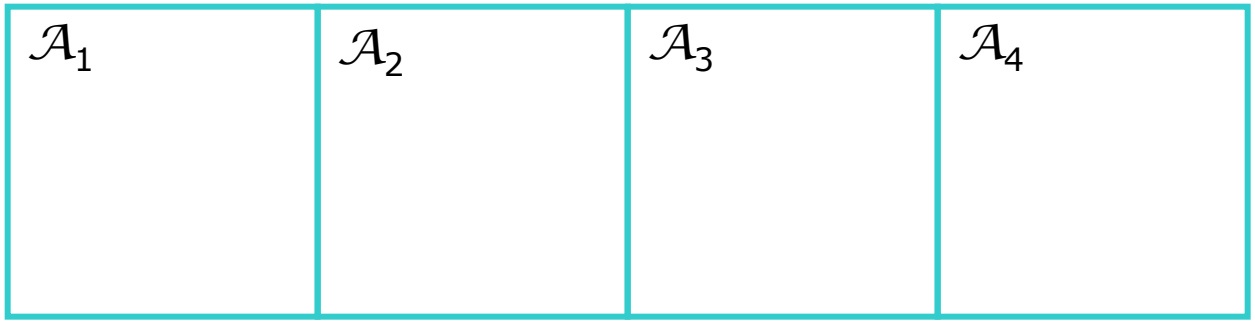
Spatial Proximity

- How is proximity defined for area data?

- Spatial proximity matrix $\mathbf{W}$

Spatial Proximity Matrix

- Contiguity



Spatial Proximity Matrix

	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$
$\mathcal{A}_1$	0			
$\mathcal{A}_2$		0		
$\mathcal{A}_3$			0	
$\mathcal{A}_4$				0

Spatial Proximity Matrix

$$\mathbf{W} = \{w_{ij}\} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ w_{n1} & \cdots & \cdots & w_{nn} \end{bmatrix}$$

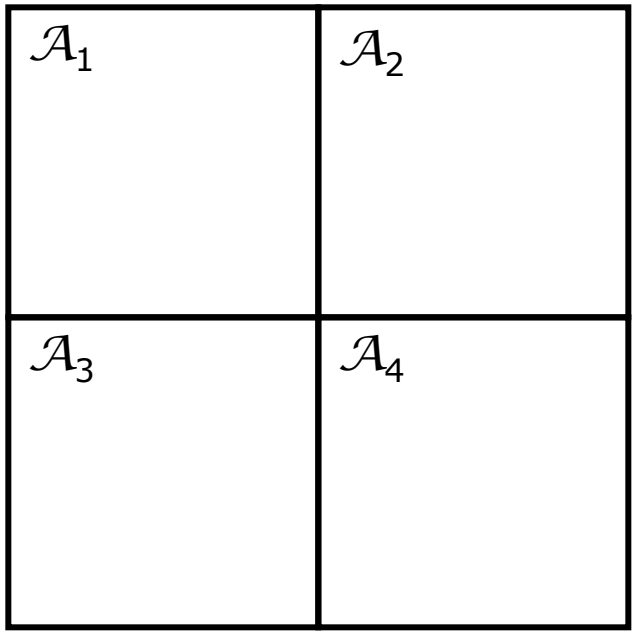
Spatial Proximity Matrix

- Definition of weights w_{ij}

$$w_{ij} = \begin{cases} 1 & \text{if } \mathcal{A}_j \text{ shares a border with } \mathcal{A}_i \\ 0 & \text{otherwise} \end{cases}$$

Spatial Proximity Matrix

- Contiguity

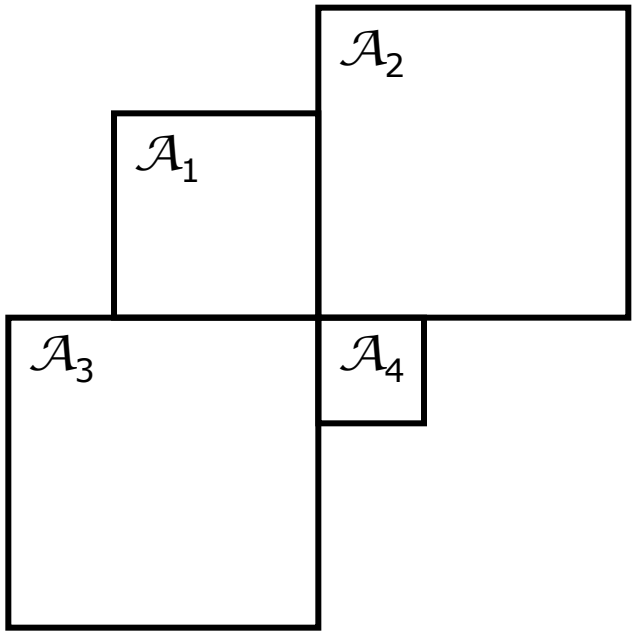


Spatial Proximity Matrix

	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$
$\mathcal{A}_1$	0			
$\mathcal{A}_2$		0		
$\mathcal{A}_3$			0	
$\mathcal{A}_4$				0

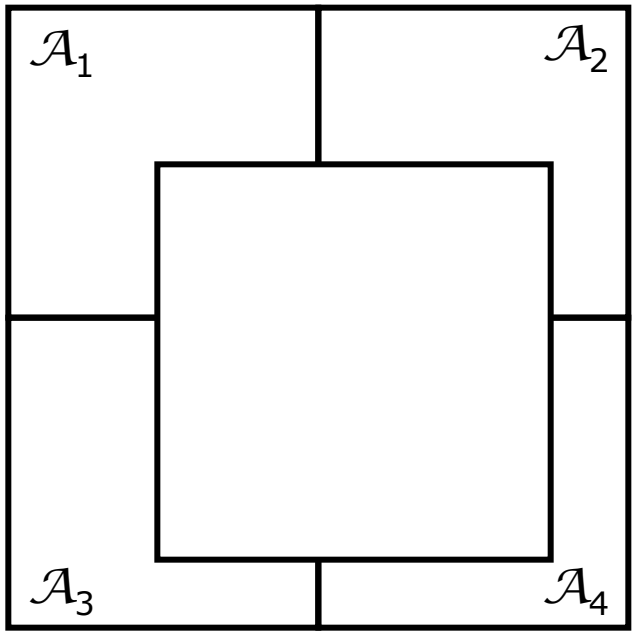
Spatial Proximity Matrix

- Topology?



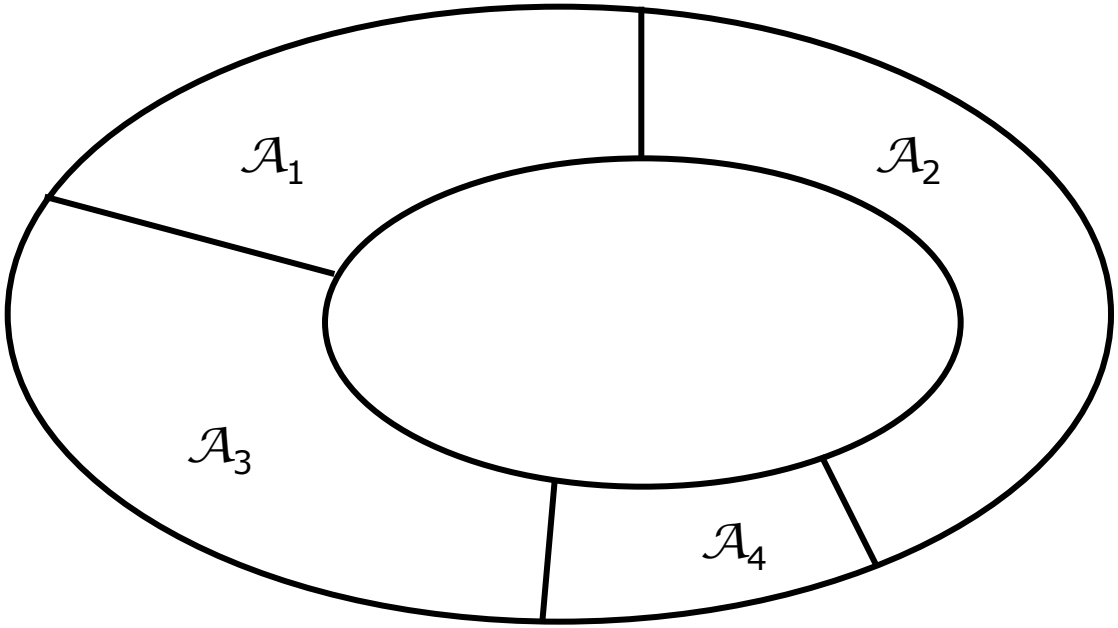
Spatial Proximity Matrix

- Topology?



Spatial Proximity Matrix

- Topology?



Spatial Proximity Matrix

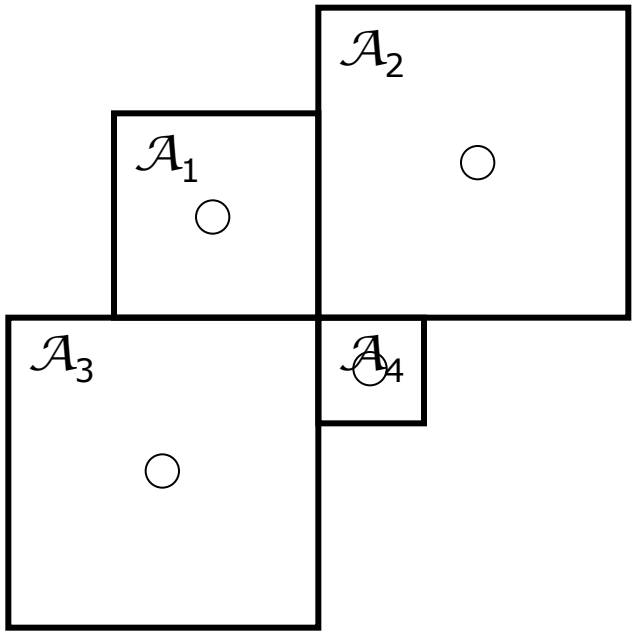
- Alternative definition of w_{ij}
  - Length of shared boundary

$$w_{ij} = \frac{l_{ij}}{l_i}$$

  - l_{ij}:length of common boundary between $\mathcal{A}_j$ and $\mathcal{A}_j$
  - l_i:perimeter of $\mathcal{A}_j$

Spatial Proximity Matrix

- Zone centroids: transforming area data into point data



Spatial Proximity Matrix

- Alternative definition of w_{ij}

$$w_{ij} = \begin{cases} 1 & \textit{if centroid of } \mathcal{A}_j \textit{ is one of} \\ & \textit{k nearest neighbors to that of } \mathcal{A}_i \\ \\ 0 & \textit{otherwise} \end{cases}$$

Spatial Proximity Matrix

- Alternative definition of w_{ij}

$$w_{ij} = \begin{cases} 1 & \textit{if centroid of } \mathcal{A}_j \textit{ is within specified distance to that of } \mathcal{A}_i \\ 0 & \textit{otherwise} \end{cases}$$

Spatial Proximity Matrix

- Alternative definition of w_{ij}

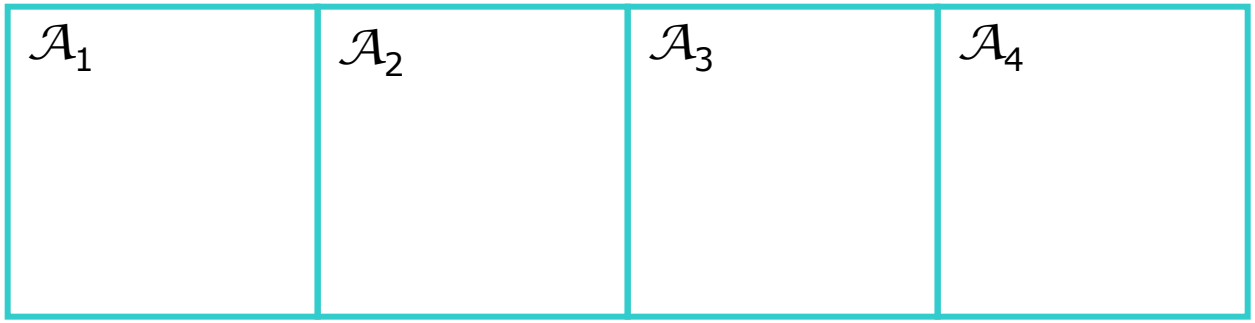
$$w_{ij} = \begin{cases} d_{ij}^{\gamma} & \text{if inter-centroid distance } d_{ij} < \delta \\ & (\delta > 0;\ \gamma < 0) \\ 0 & \text{otherwise} \end{cases}$$

Spatial Proximity Matrix

- Higher order matrices $\mathbf{W}$
  - First order “neighbors”
  - Second order “neighbors”

Spatial Proximity Matrix

- Higher order matrices

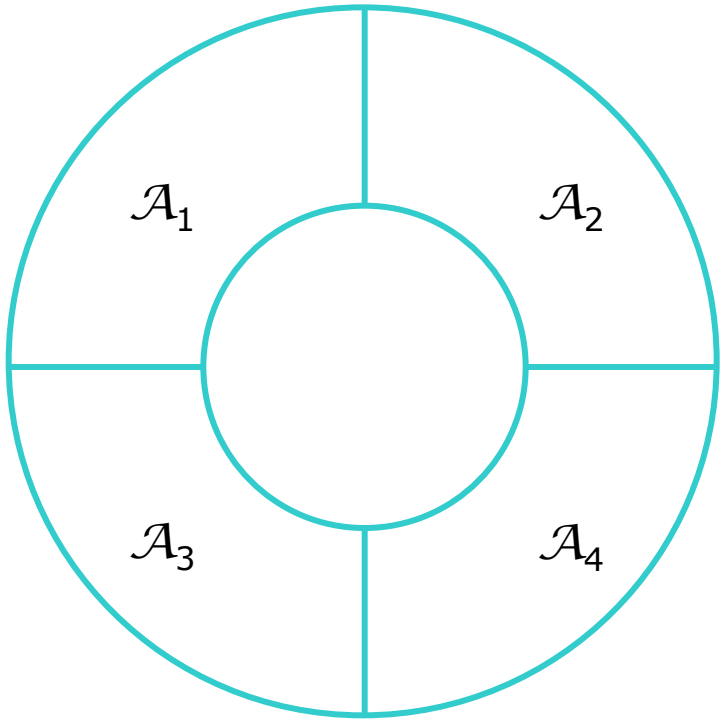


Spatial Proximity Matrix

	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$	$\mathcal{A}_4$
$\mathcal{A}_1$	0	0	1	0
$\mathcal{A}_2$	0	0	0	1
$\mathcal{A}_3$	1	0	0	0
$\mathcal{A}_4$	0	1	0	0

Spatial Proximity Matrix

- Row-standardization of $\mathbf{W}$



Spatial Proximity Matrix

- Row-standardization of W

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Spatial Proximity Matrix

- Row-standardization of W

$$\mathbf{W}_{st} = \left[\quad \right]$$

Spatial Proximity Matrix

- Network data
  - The units of analysis are nodes in a network
  - For example, cities in a transportation network
  - Network connectivity gives the proximity matrix

Next…

- Exploring area data
- First order effects
- Second order effects