Lesson 6: Spatial Weights and Applications

Dr. Kam Tin Seong
Assoc. Professor of Information Systems

School of Computing and Information Systems,
Singapore Management University

2020-5-1 (updated: 2021-06-06)

Content

- The concept of spatial autocorrelation and how it help us to understand real world phenomena
- Defining spatial Neighbourhoods and Weights
- Contiguity-Based Spatial Weights
    - Rook's
    - Queen's
- Distance-Band Spatial Weights
- Applications of Spatial Weights

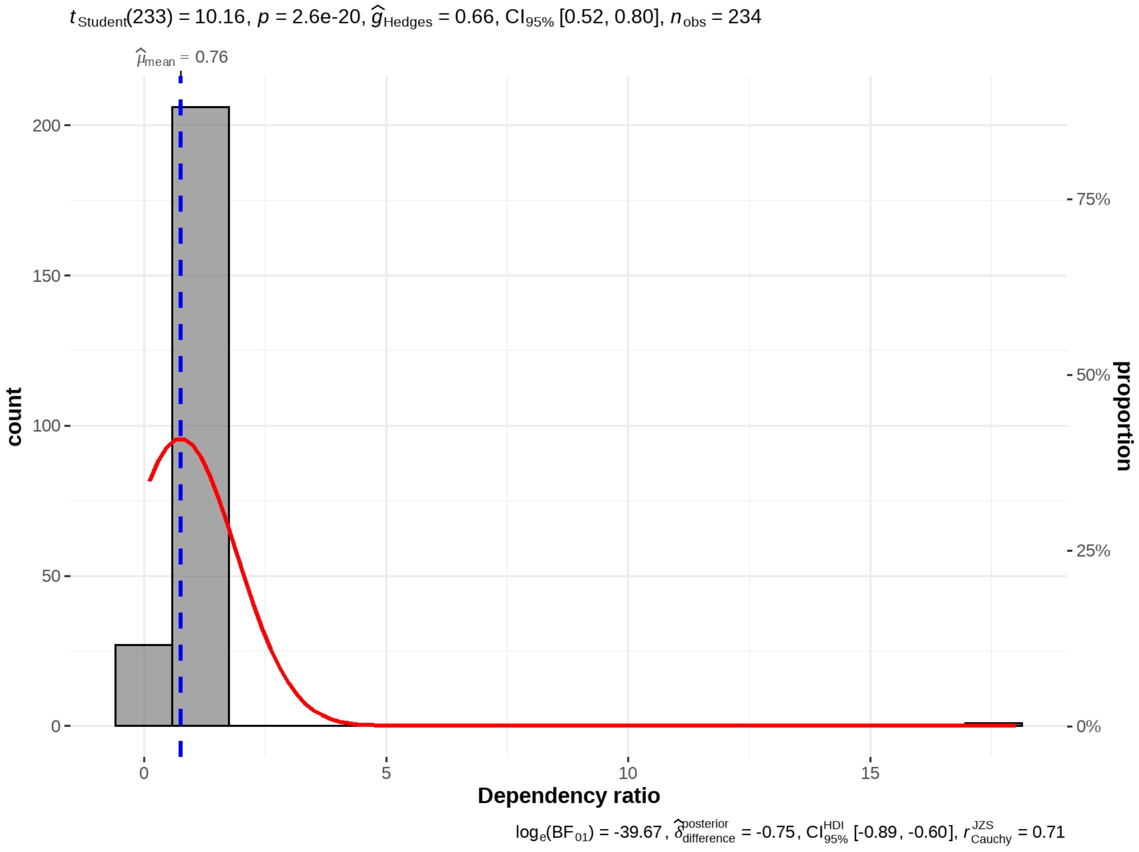
What is geographically referenced attribute?

A kind of data that is very similar to an ordinary data. The only difference is that each observation is associated with some form of geography such as numbers of aged population by planning zone.

```
## Rows: 323
## Columns: 12
## $ SUBZONE_N          <chr> "MARINA SOUTH", "PEARL'S HILL", "BOAT QUAY", "HENDERS~
## $ SUBZONE_C          <fct> MSSZ01, OTSZ01, SRSZ03, BMSZ08, BMSZ03, BMSZ07, BMSZ0~
## $ PLN_AREA_N         <fct> MARINA SOUTH, OUTRAM, SINGAPORE RIVER, BUKIT MERAH, B~
## $ PLN_AREA_C         <fct> MS, OT, SR, BM, BM, BM, BM, SR, QT, QT, QT, BM, ME, R~
## $ REGION_N           <fct> CENTRAL REGION, CENTRAL REGION, CENTRAL REGION, CENTR~
## $ REGION_C           <fct> CR, CR, CR, CR, CR, CR, CR, CR, CR, CR, CR, CR, CR, C~
## $ YOUNG              <dbl> NA, 1100, 0, 2620, 2840, 2910, 2850, 0, 1120, 30, NA,~
## $ `ECONOMY ACTIVE`   <dbl> NA, 3420, 50, 7500, 6260, 7560, 8340, 50, 2750, 210, ~
## $ AGED               <dbl> NA, 2110, 20, 3260, 1630, 3310, 3590, 10, 560, 50, NA~
## $ TOTAL              <dbl> NA, 6630, 70, 13380, 10730, 13780, 14780, 60, 4430, 2~
## $ DEPENDENCY         <dbl> NA, 0.9385965, 0.4000000, 0.7840000, 0.7140575, 0.822~
## $ geometry           <MULTIPOLYGON [m]> MULTIPOLYGON (((31495.56 30..., MULTIPOL~
```

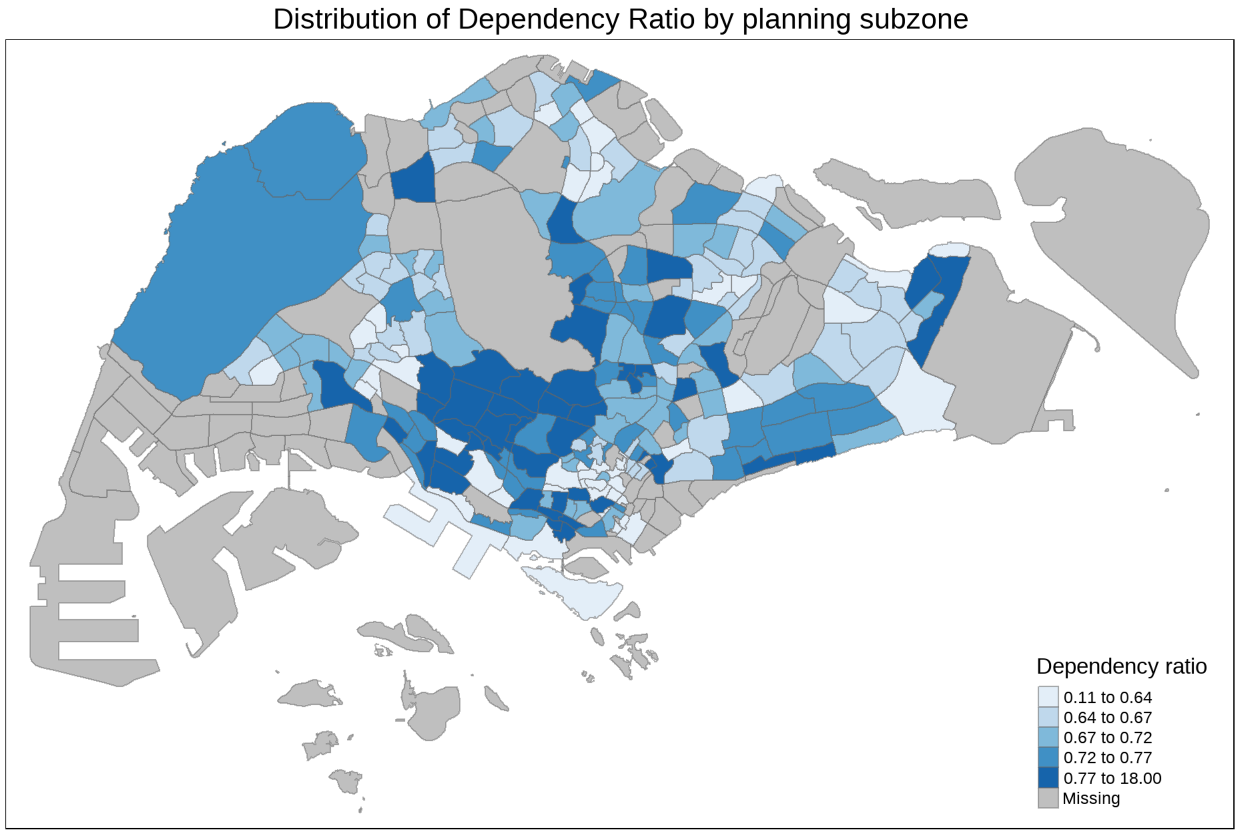
Describing attribute distribution

The dependency ratio values by planning subzone are normally distributed.



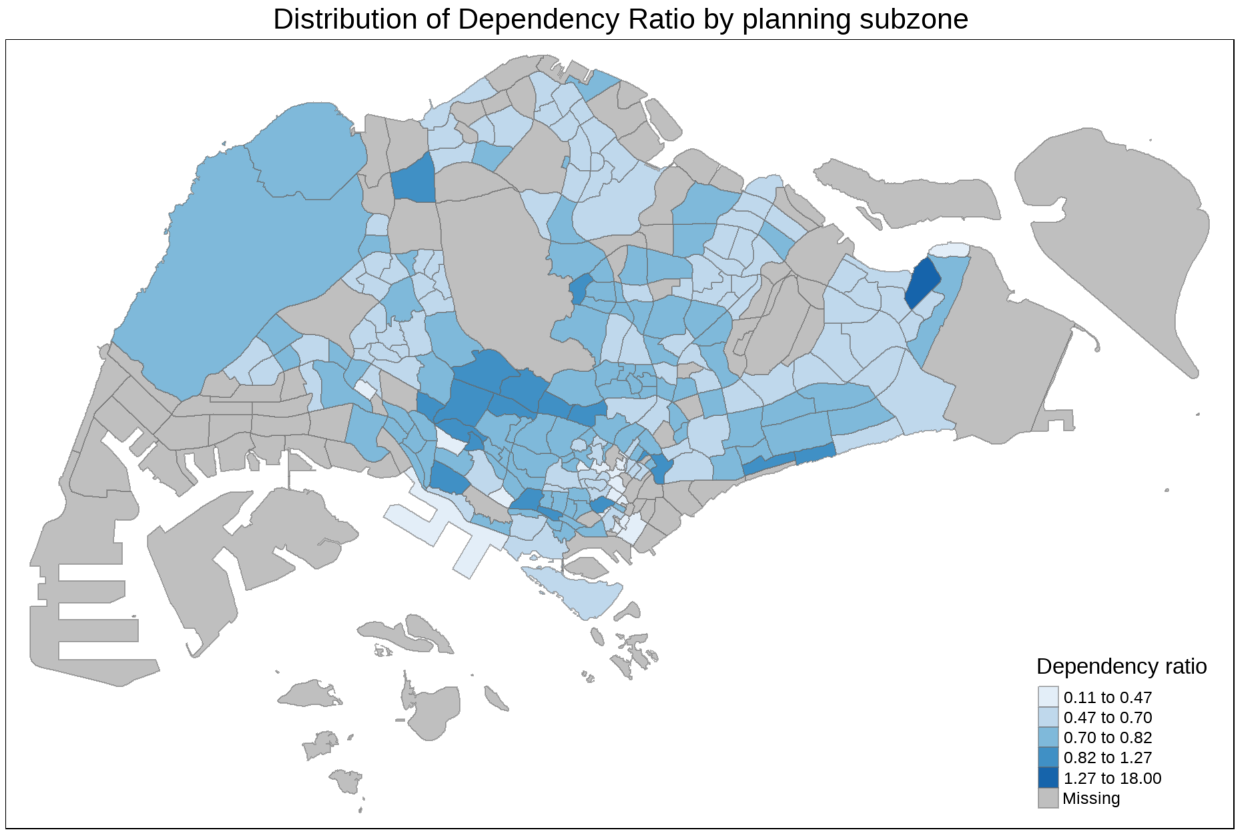
Geographical distribution question

Are the planning subzones with high proportion of dependency ratio randomly distributed over space?



Geographical distribution question

Are the planning subzones with high proportion of dependency ratio randomly distributed over space?



Tobler’s First law of Geography

Everything is related to everything else, but near things are more related than distant things.

The foundation of the fundamental concepts of:

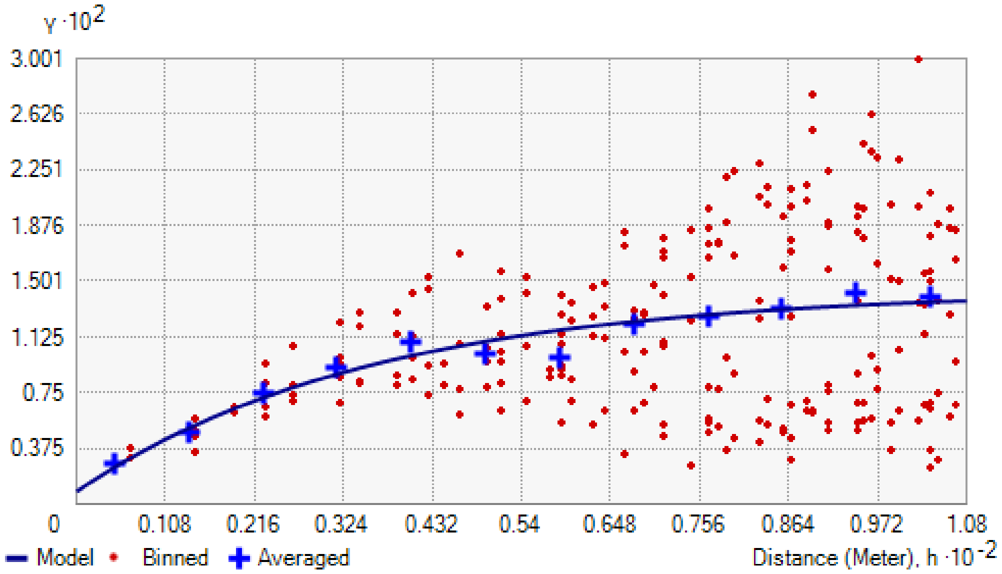
- spatial dependence, and
- spatial autocorrelation

Reference: A Computer Movie Simulating Urban Growth in the Detroit Region

Spatial Dependency

- Spatial dependence is the spatial relationship of variable values (for themes defined over space, such as rainfall) or locations (for themes defined as objects, such as cities).

- Spatial dependence is measured as the existence of statistical dependence in a collection of random variables, each of which is associated with a different geographical location.



Spatial Autocorrelation

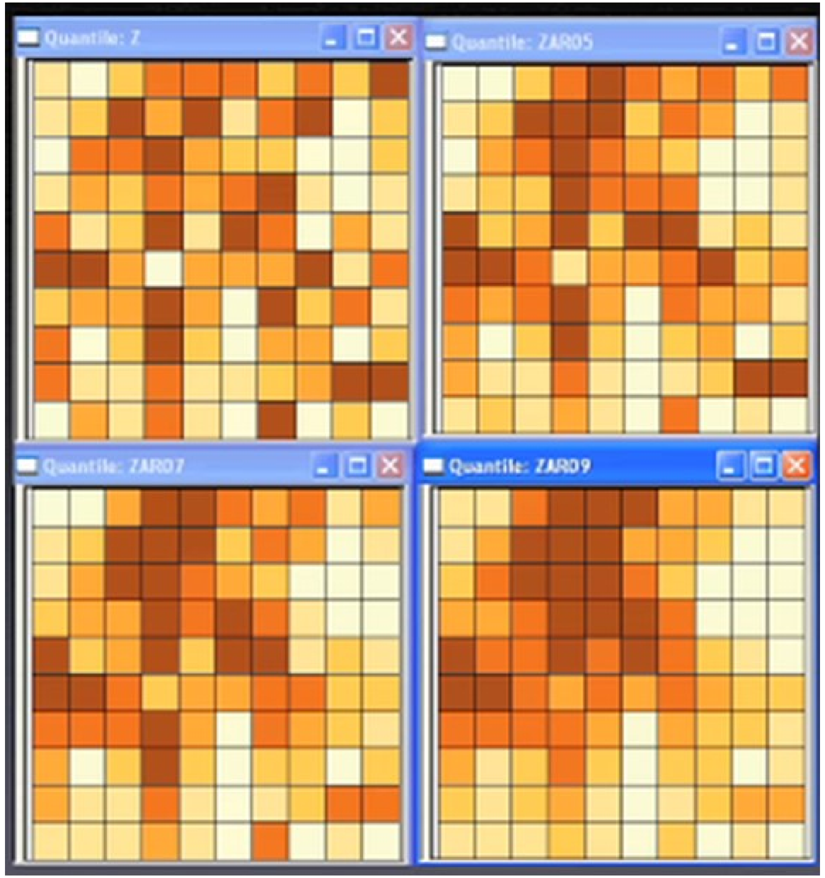
- Spatial autocorrelation is the term used to describe the presence of systematic spatial variation in a variable.

- The variable can assume values either:

    - at any point on a continuous surface (such as land use type or annual precipitation levels in a region);
    - at a set of fixed sites located within a region (such as prices at a set of retail outlets); or
    - across a set of areas that subdivide a region (such as the count or proportion of households with two or more cars in a set of Census tracts that divide an urban region).

Positive Spatial Autocorrelation

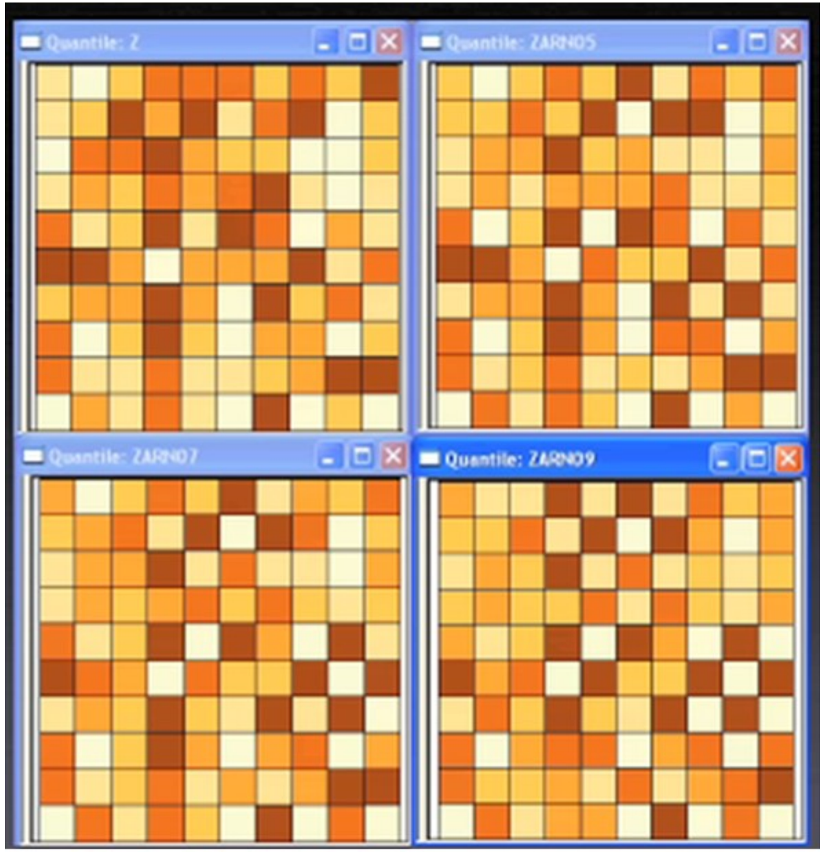
- Clustering
  - like values tend to be in similar locations.
- Neighbours are similar
  - more alike than they would be under spatial randomness.
- Compatible with diffusion
  - but not necessary caused by diffusion.



Legend: 0.1, 0.5, 0.7, 0.9

Negative Spatial Autocorrelation

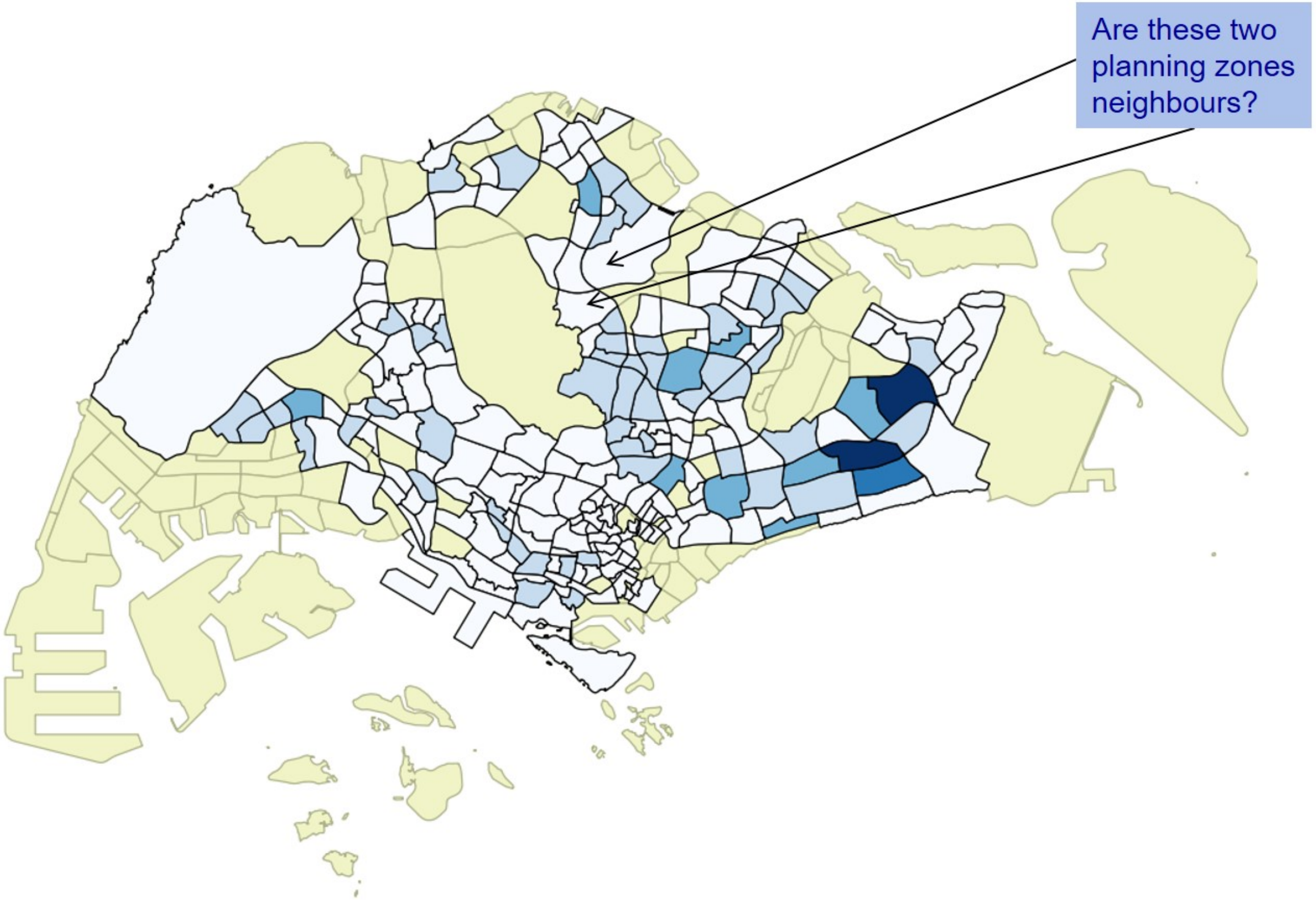
- Checkerboard patterns
    - “opposite” of clustering
- Neighbours are dissimilar
    - more dissimilar than they would be under spatial randomness
- Compatible to competition
    - but not necessary competition



Legend: -0.1,-0.5, -0.7, -0.9

What are Spatial Weights (*wij*)

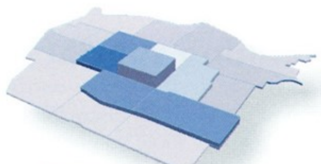
- A way to define spatial neighbourhood.



Defining Spatial Weight Matrices

Adjacency criterion:

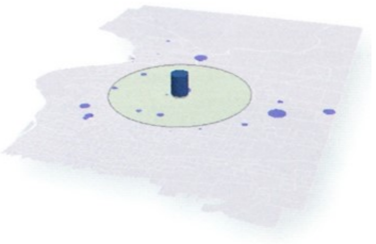
$$w_{ij} = \begin{cases} 1 & \text{if location j is adjacent to i,} \\ 0 & \text{if location j is not adjacent to i.} \end{cases}$$



Adjacency

Distance criterion:

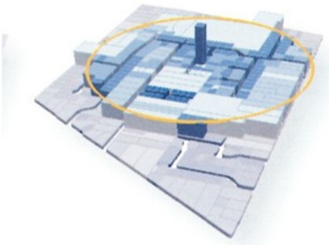
$$w_{ij}(d) = \begin{cases} 1 & \text{if location j is within distance d from i,} \\ 0 & \text{otherwise.} \end{cases}$$



Distance: discrete features

A general spatial distance weight matrices:

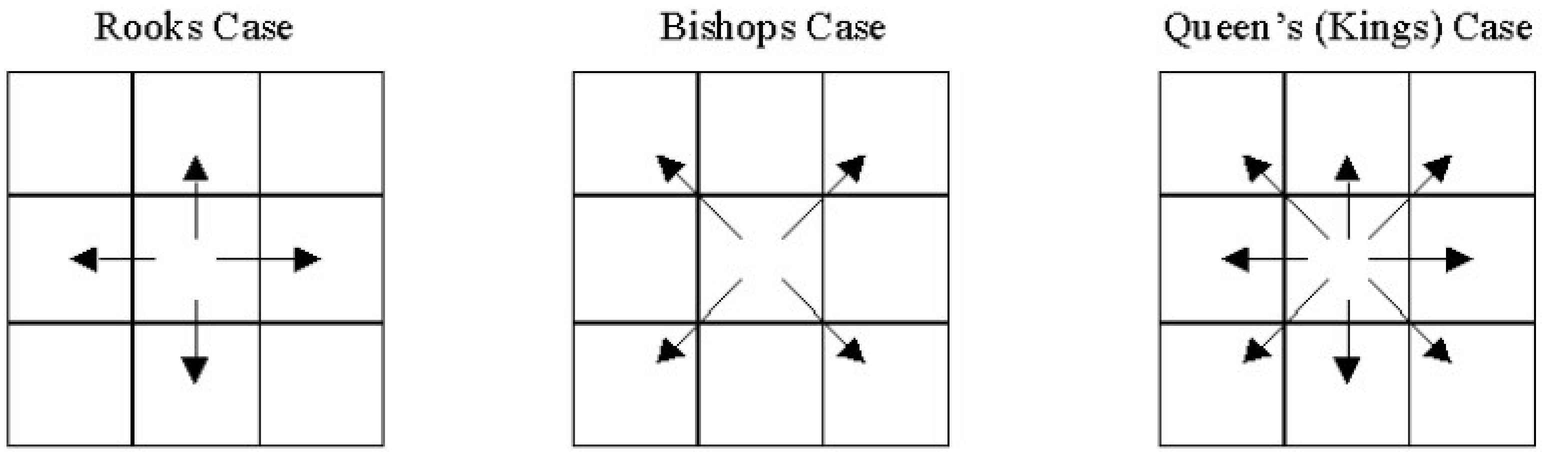
$$w_{ij}(d) = d_{ij}^{-a} \cdot \beta^{b}$$



Distance: contiguous features

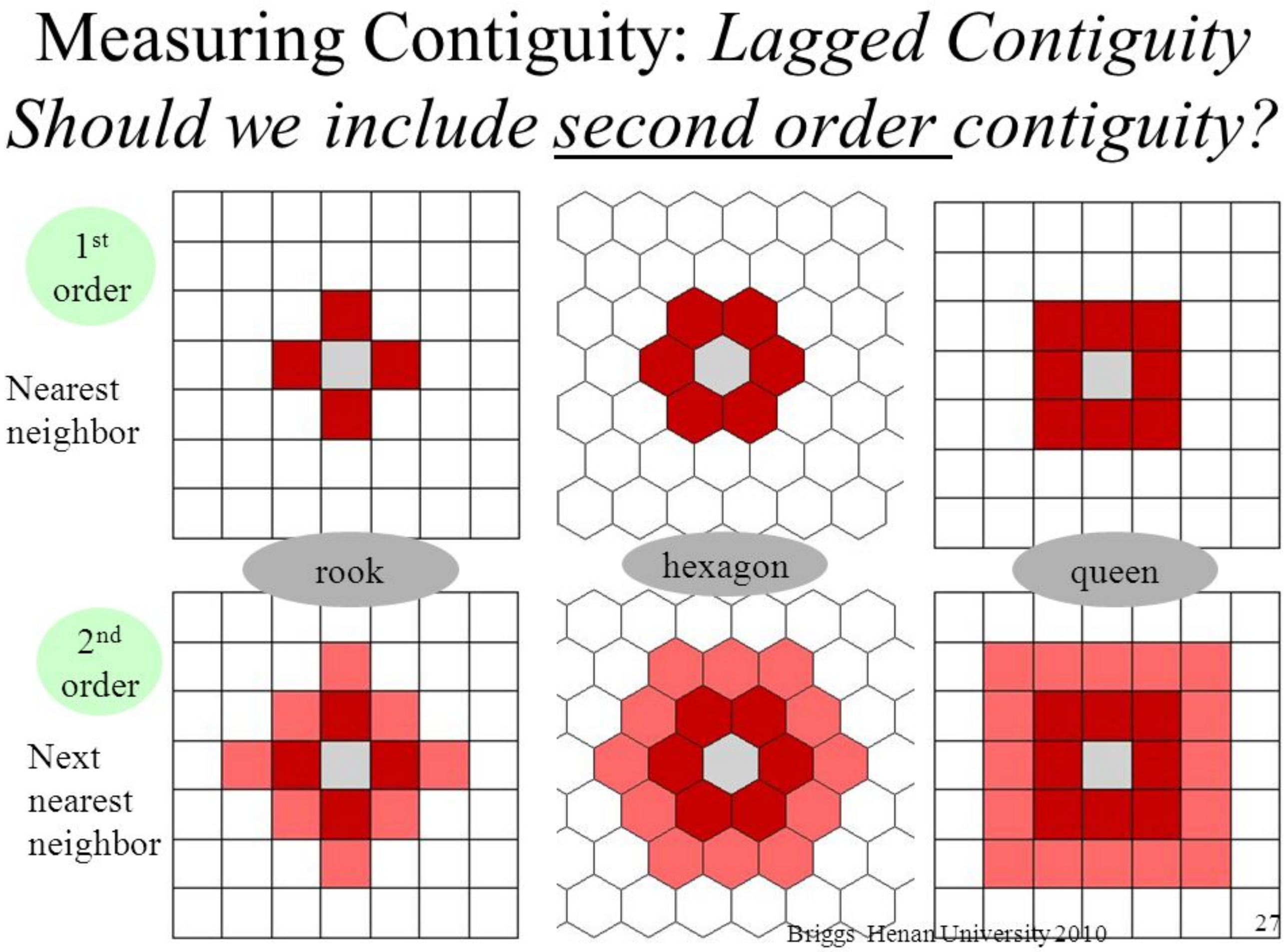
Contiguity Neighbours

- Contiguity (common boundary)
- What is a “shared” boundary?



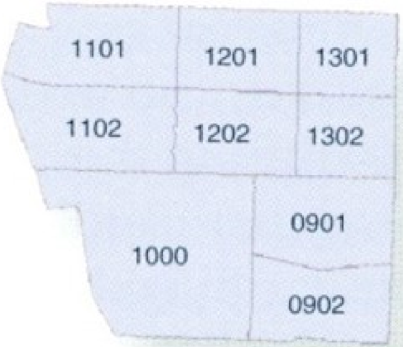
Beyond the basic contiguity neighbours

There are also second-order, third-order, forth-order, etc contiguity



Weights matrix: Adjacency-based neighbours

Quiz: With reference to the figure below, list down the neighbour(s) of area 1202 using Rook case

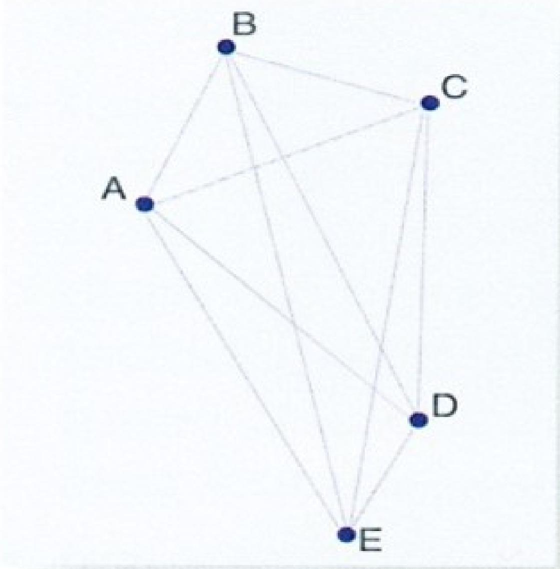


Weights matrix for an adjacency-based neighborhood

	1101	1201	1301	1102	1202	1302	1000	0901	0902
1101	0	1	0	1	1	0	0	0	0
1201	1	0	1	1	1	1	0	0	0
1301	0	1	0	0	1	1	0	0	0
1102	1	1	0	0	1	0	1	0	0
1202	1	1	1	1	0	1	1	1	0
1302	0	1	1	0	1	0	0	1	0
1000	0	0	0	1	1	0	0	1	1
0901	0	0	0	0	1	1	1	0	1
0902	0	0	0	0	0	0	1	1	0

Weights Matrix: Distance-based neighbours

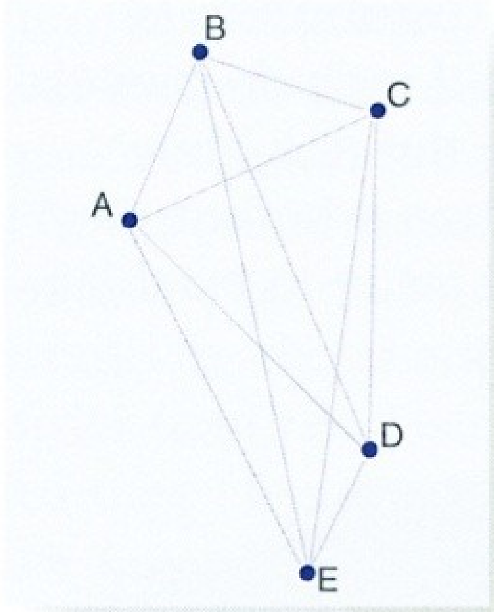
Quiz: With reference to the figure below, create a weights matrix for d = 650.



Weights matrix for a distance-based neighborhood

	A	B	C	D	E
A	0	353	516	641	757
B	353	0	357	837	1025
C	516	357	0	659	901
D	641	837	659	0	263
E	757	1025	901	263	0

Weights matrix: Measured distances



Weights matrix with measured distances

	A	B	C	D	E
A	0	353	516	641	757
B	353	0	357	837	1025
C	516	357	0	659	901
D	641				
E	757				

$$w_{ij} = \frac{1}{d_{ij}}$$

	A	B	C	D	E
A	0	0.00283	0.00194	0.00156	0.00132
B	0.00283	0	0.00280	0.00119	0.00098
C	0.00194	0.00280	0	0.00152	0.00111
D	0.00156	0.00119	0.00152	0	0.00380
E	0.00132	0.00098	0.00111	0.00380	0

Weights matrix with inverse distances

Row standardisation

- Row-standardised weights increase the influence of links from observations with few neighbours, which binary weights vary the influence of observations.
  - Those with many neighbours are up-weighted compared to those with few.

Binary W matrix:

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

Row standardized W matrix:

$$W = \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \end{bmatrix}$$

Spatially Lagged Variables

With a neighbor structure defined by the non-zero elements of the spatial weights matrix W, a **spatially lagged variable** is a weighted sum or a weighted average of the neighboring values for that variable. In most commonly used notation, the spatial lag of y is then expressed as Wy.

Formally, for observation i, the spatial lag of yi, referred to as [Wy]i (the variable Wy observed for location i) is:

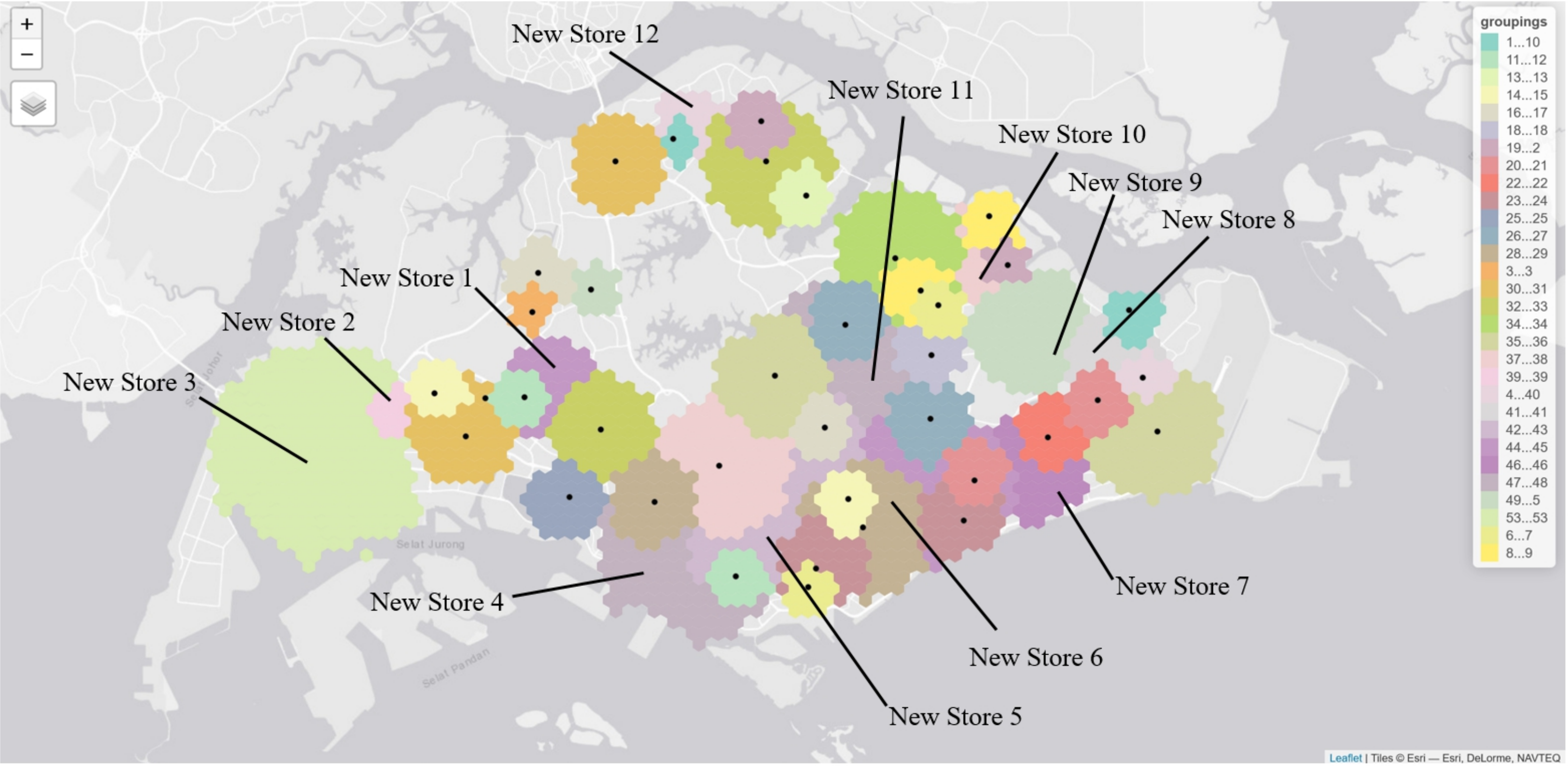
$$[Wy]_i = w_{i1}y_1 + w_{i2}y_2 + \cdots + w_{in}y_n,$$

$$[Wy]_i = \sum_{j=1}^{n} w_{ij}y_j,$$

where the weights wij consist of the elements of the i-th row of the matrix W, matched up with the corresponding elements of the vector y.

Application of Spatially Lagged Variables

In this project, spatially lagged variables approach is used to delineate potential locations for new DELCO stores of a Quick Service Restaurants (QSR) in Singapore.



References