

# SA-TIED Geospatial Analysis Workshop

## Overview



Dr Justin van Dijk



[j.t.vandijk@ucl.ac.uk](mailto:j.t.vandijk@ucl.ac.uk)



# Objectives

We will explore the following topics:

- Fundamentals of using R for data analysis.
- Creating thematic maps using R.
- Quantifying the degree of spatial dependence in a dataset.
- Incorporating space into statistical models.

# Schedule

Day 1 – Morning	R for Data Analysis
Day 1 – Afternoon	R for Spatial Analysis
Day 2 – Morning	Spatial Autocorrelation
Day 2 – Afternoon	Spatial Models



# SA-TIED Geospatial Analysis Workshop

## S03 – Spatial Autocorrelation



# This session

- Spatial dependence.
- Spatial autocorrelation.
- Formalising space.

# Spatial dependence

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

Walter Tobler 1970

# Spatial dependence

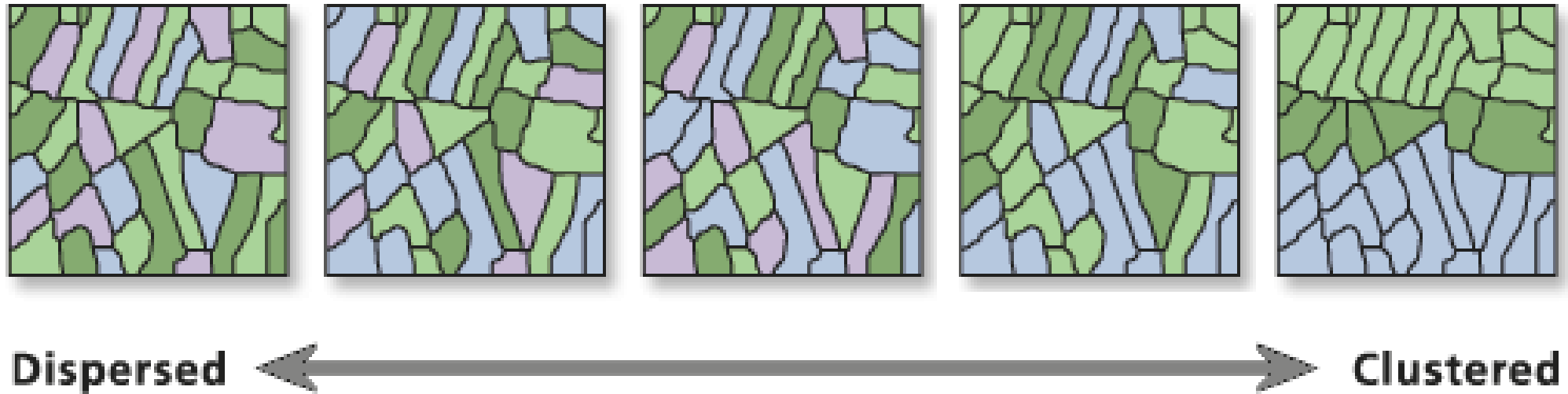
- Spatial dependence refers to the concept that the value of a variable at one location is influenced, to some extent, by the value of the same variable at nearby locations.
- This is often understood through the concept of distance decay, where the influence decreases as distance increases.
- Spatial dependence is a key principle in various geographical applications, such as spatial interpolation and spatial interaction modeling.

# Spatial autocorrelation

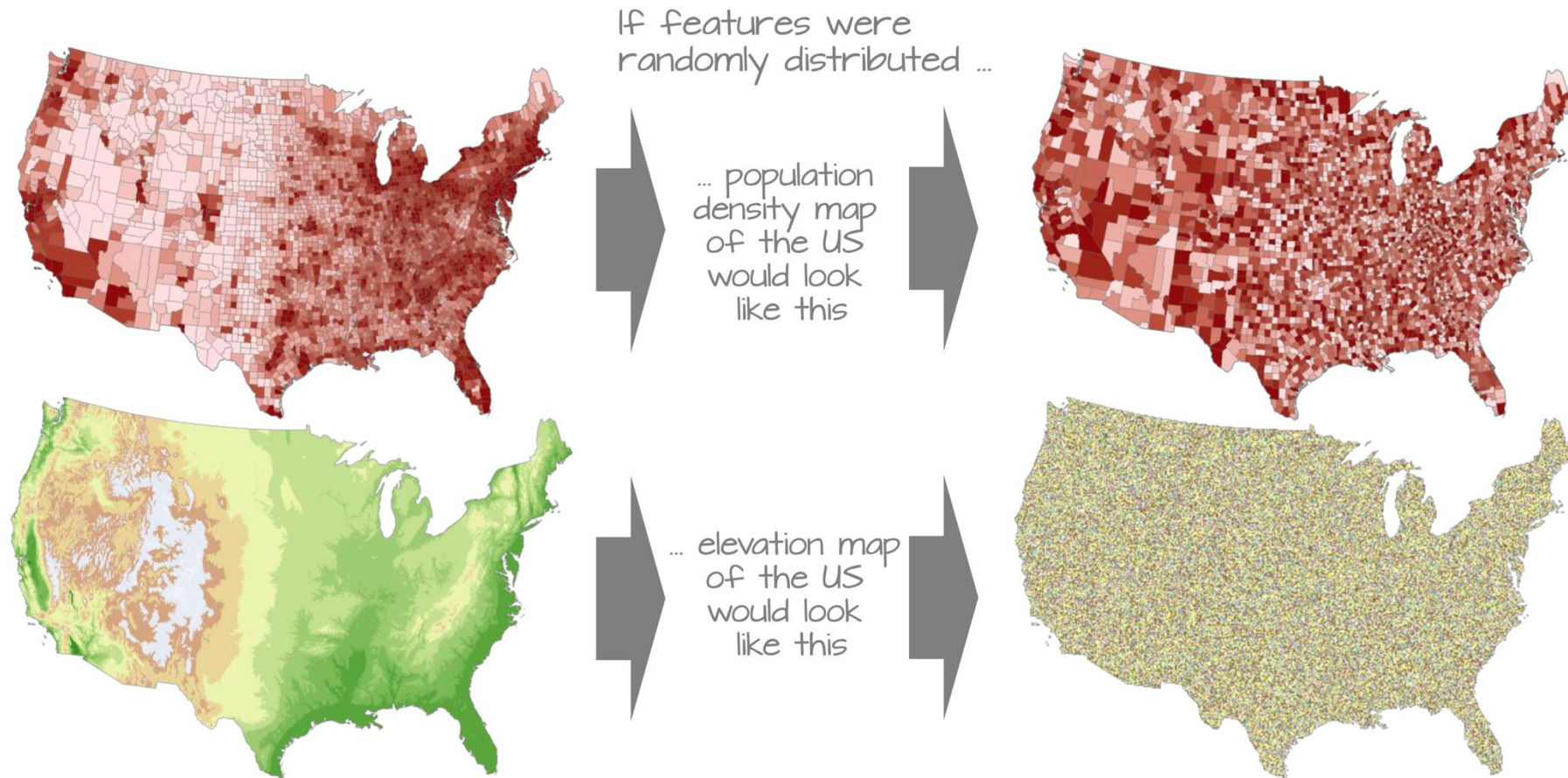
- Measurement of spatial autocorrelation is the idea of formalising spatial dependency: measuring the degree to which similar values cluster together in space.
- By measuring spatial autocorrelation, we try to identify hotspots where high values are concentrated versus areas where low values are concentrated.
- Spatial Autocorrelation indicates the absence of Complete Spatial Randomness (CSR).
- CSR suggests that a pattern is entirely the result of random chance, with no underlying spatial structure.



# Spatial autocorrelation



# Spatial autocorrelation



# Spatial autocorrelation

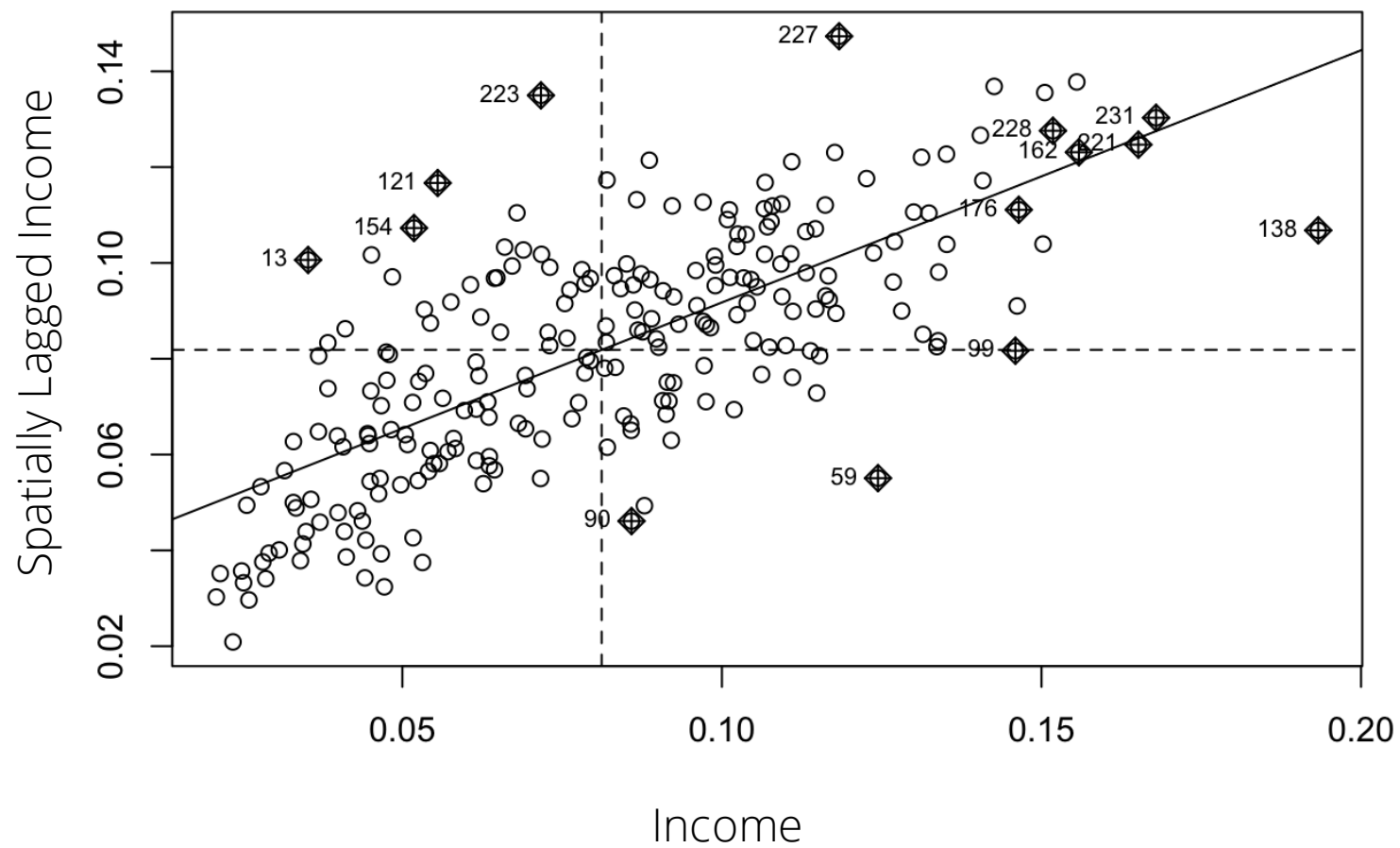
Spatial Autocorrelation can be measured in two ways:

- 1) Global Spatial Autocorrelation: This assesses the overall spatial dependence across the entire dataset.
- 2) Local Spatial Autocorrelation: This focuses on the differences between each unit of analysis and its neighbors.

# Spatial autocorrelation

- Moran's I: The most commonly used indicator of global spatial autocorrelation.
- Identifies neighbours for each target feature (e.g. polygon) and summarises their values by computing their means to create a **spatially lagged variable** value.
- Plots the target feature's value against its spatially lagged mean value and fits a linear model to the points.
- The slope of the fitted line ( $\beta$  estimate) is the Moran's I statistic.

# Spatial autocorrelation

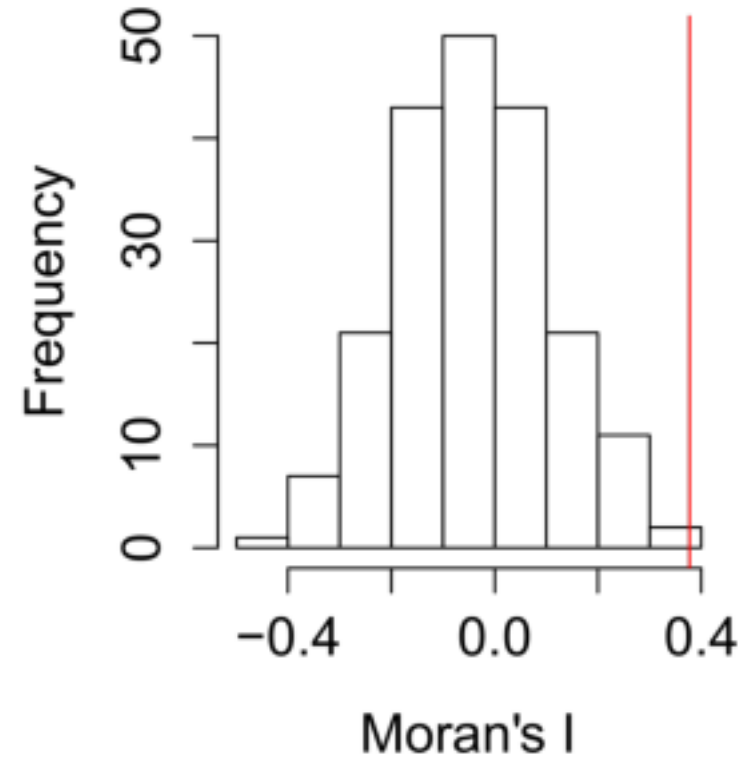
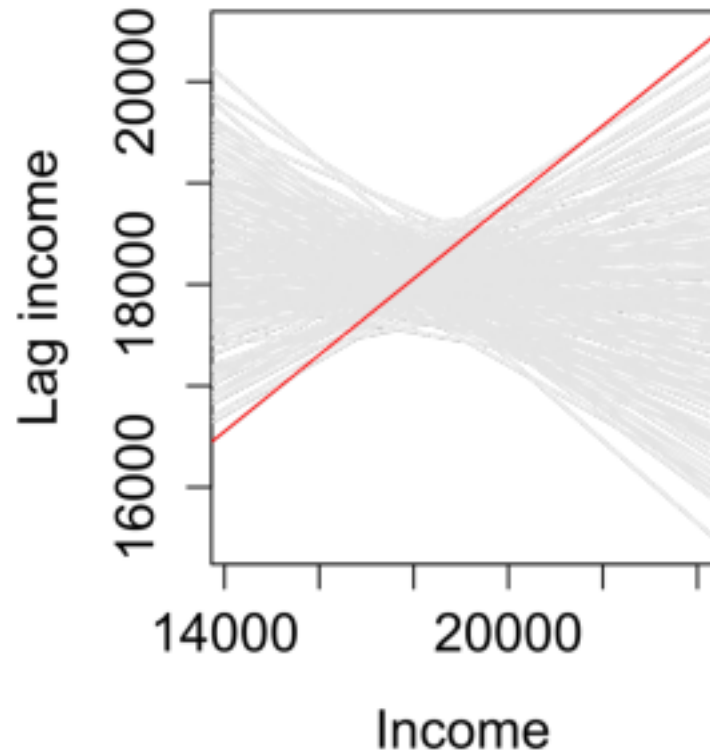




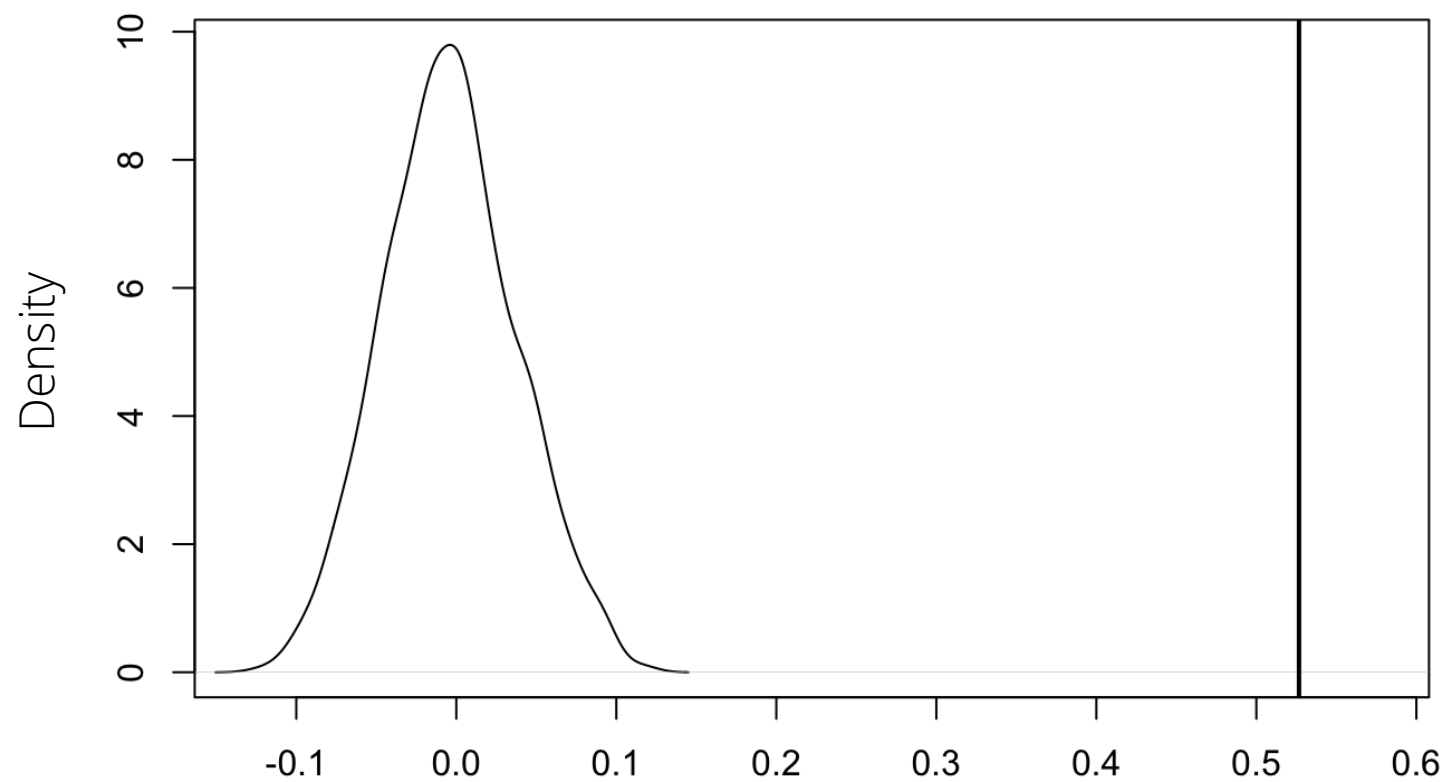
# Spatial autocorrelation

- The Moran's I statistic typically ranges from -1 to +1:
  - +1 Indicates perfect clustering (positive spatial autocorrelation).
  - 0 Suggests a random pattern (no spatial autocorrelation).
  - -1 Indicates perfect dispersion (negative spatial autocorrelation).
- How to assess the significance of the relationship?

# Spatial autocorrelation

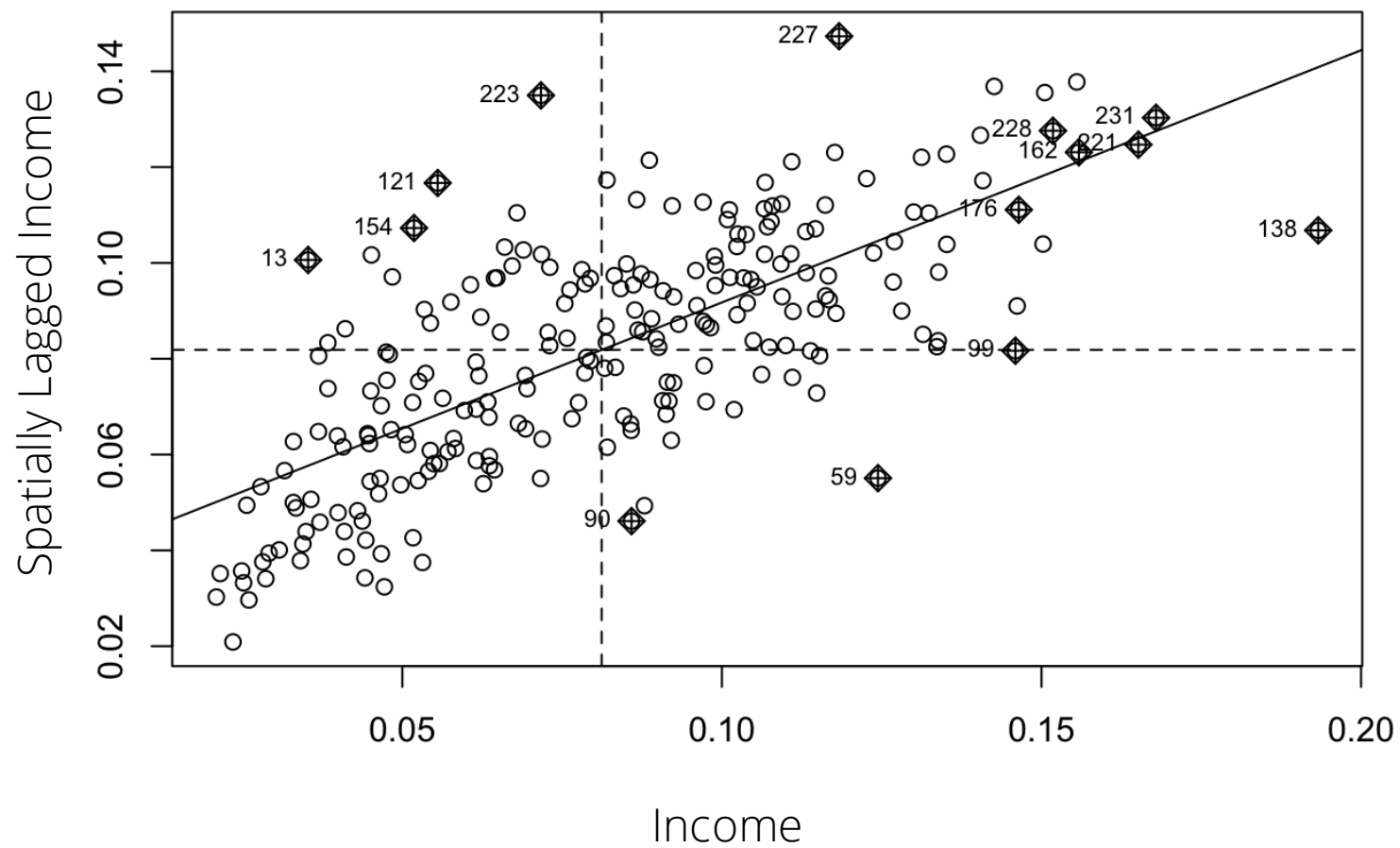


# Spatial autocorrelation

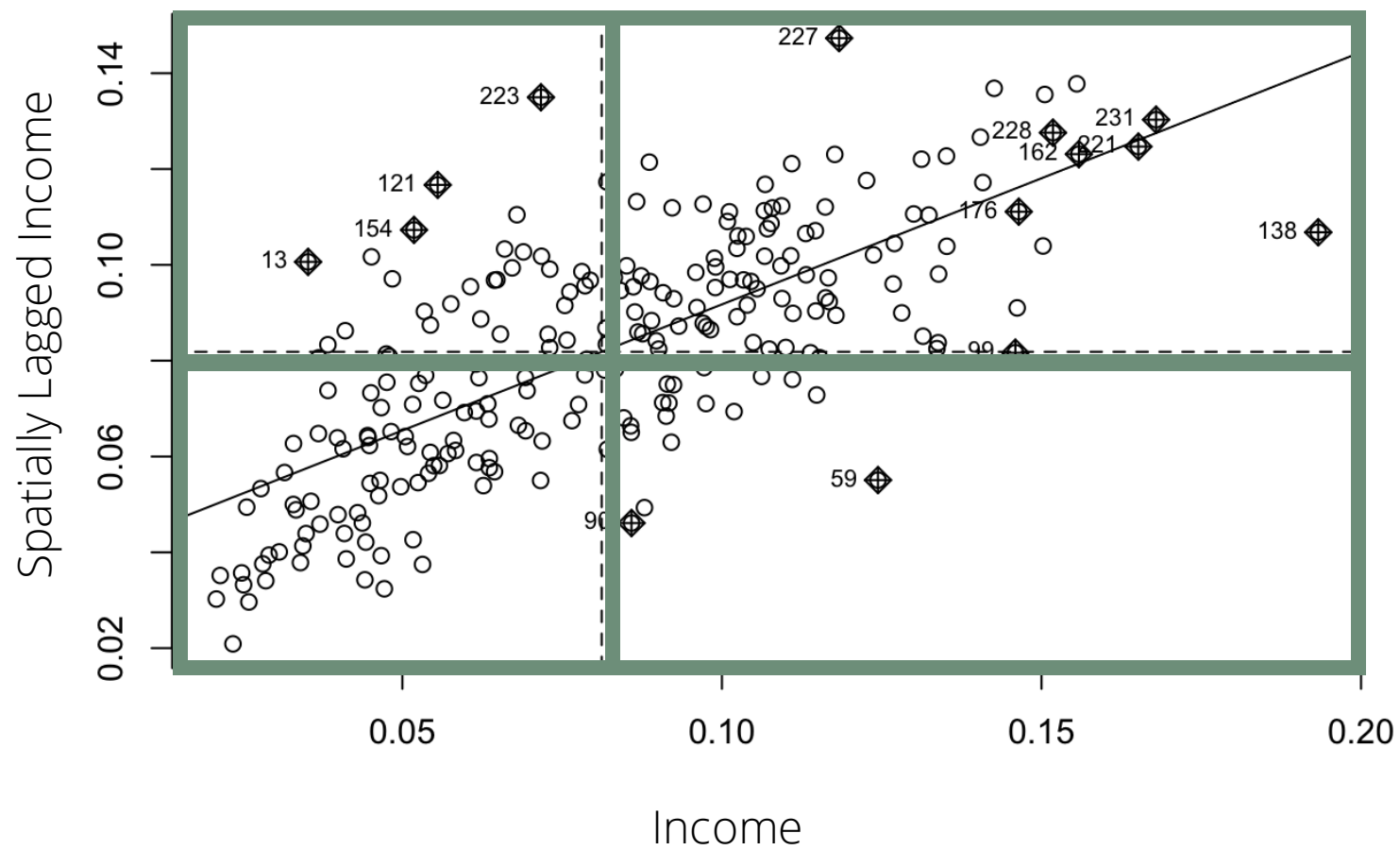


Monte Carlo Simulation of Moran's I

# Spatial autocorrelation

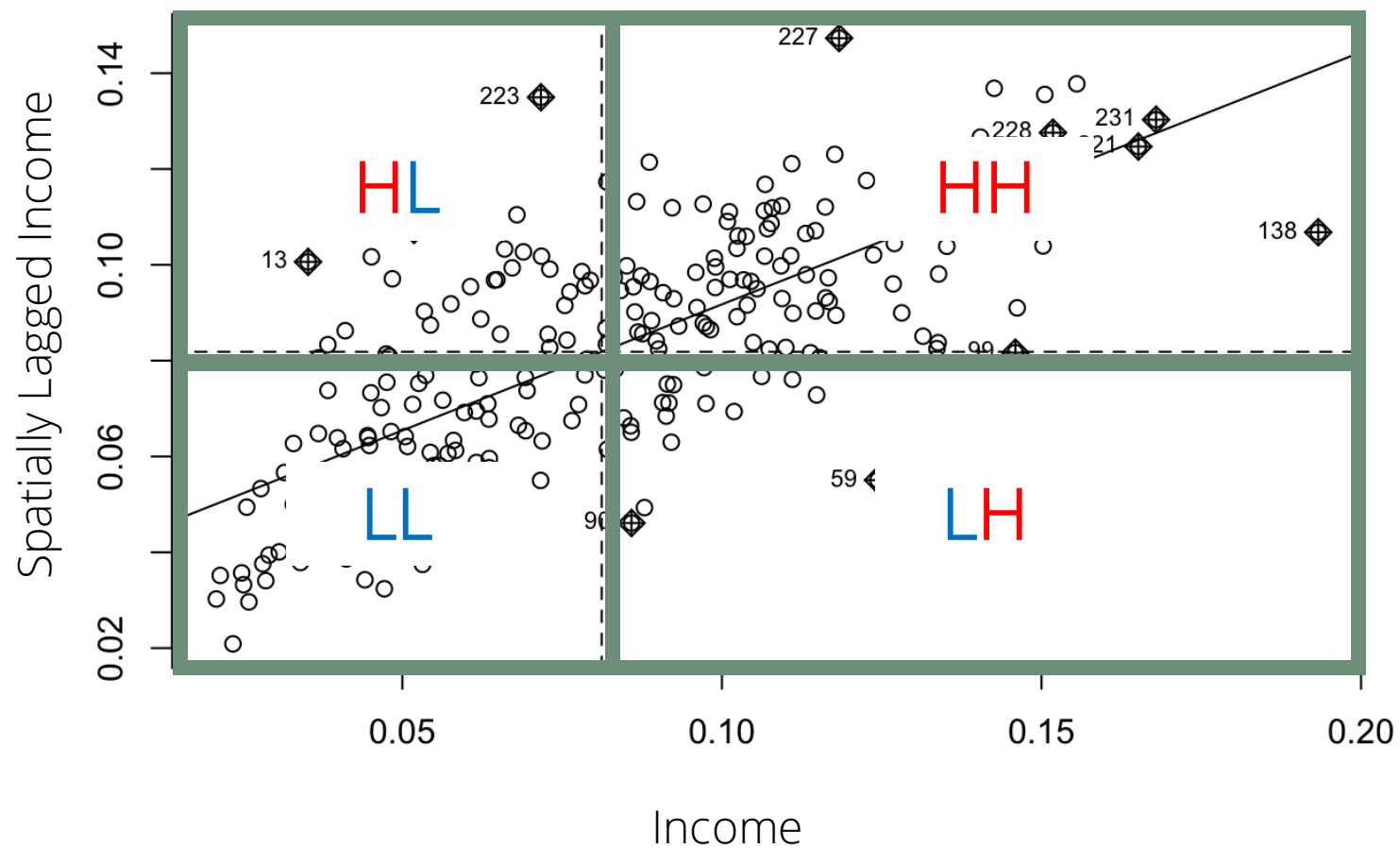


# Spatial autocorrelation





# Spatial autocorrelation



# Spatial autocorrelation

- Decomposing the Moran's I statistic: Local Moran's I
- Assesses spatial autocorrelation at the local level by evaluating each feature and its surrounding neighborhood.
- Four cluster types: high-high, low-low, but also outliers: high-low, low-high.
- Commonly known as cluster and outlier analysis.
- Monte Carlo simulation can be used to assess significance of these clusters.

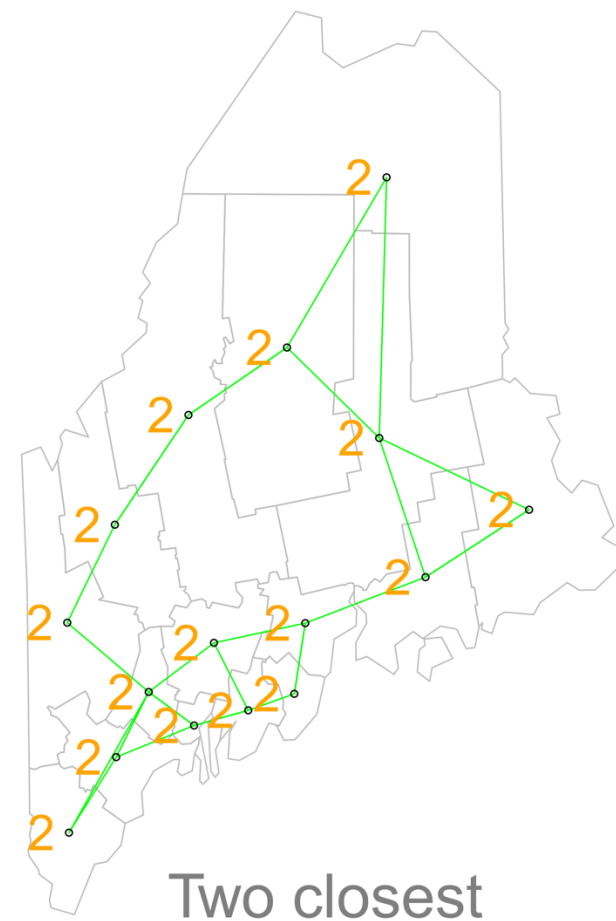
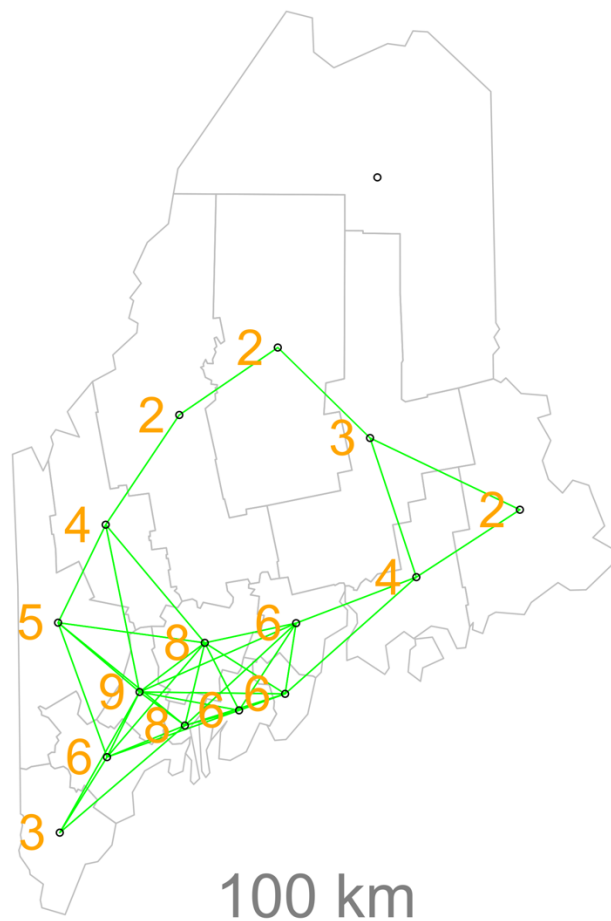
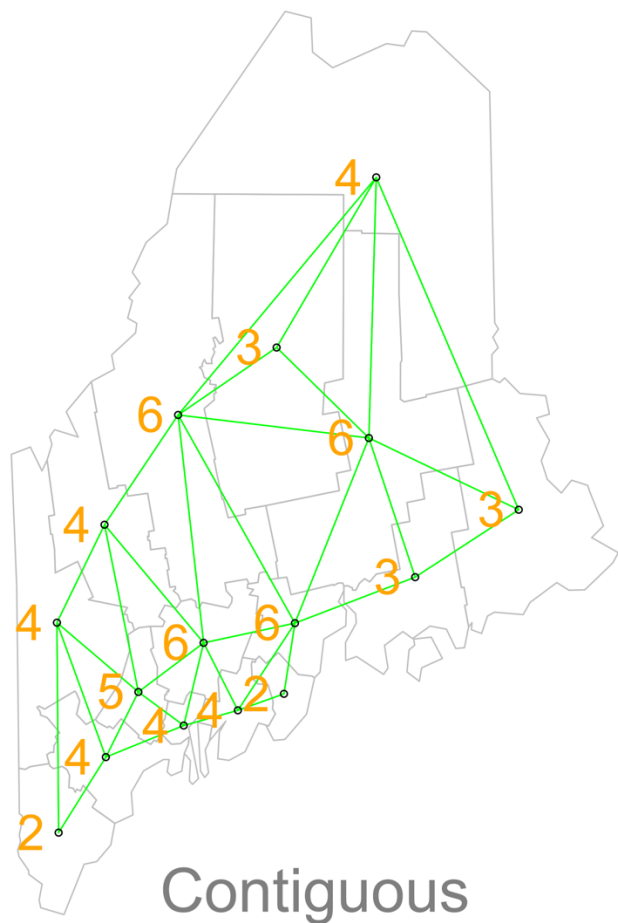
# Who is your neighbour?

- To calculate any measure of spatial autocorrelation, it's essential to understand how spatial units relate to each other as neighbours, specifically how we define their spatial relationships.
- There are two primary approaches to defining neighbours:
  - Proximity: Neighbours are determined based on the distance between spatial units.
  - Contiguity: Neighbours are defined by shared boundaries or edges.

# Who is your neighbour?

- To calculate any measure of spatial autocorrelation, it's essential to understand how spatial units relate to each other as neighbours, specifically how we define their spatial relationships.
- There are two primary approaches to defining neighbours:
  - Proximity: Neighbours are determined based on the distance between spatial units. A unit is considered a neighbour if it is within a certain distance from another unit.
  - Contiguity: Neighbours are defined by shared boundaries or edges. Spatial units are considered neighbours if they touch each other directly, either along edges (rook contiguity) or at corners (queen contiguity).

# Who is your neighbour?

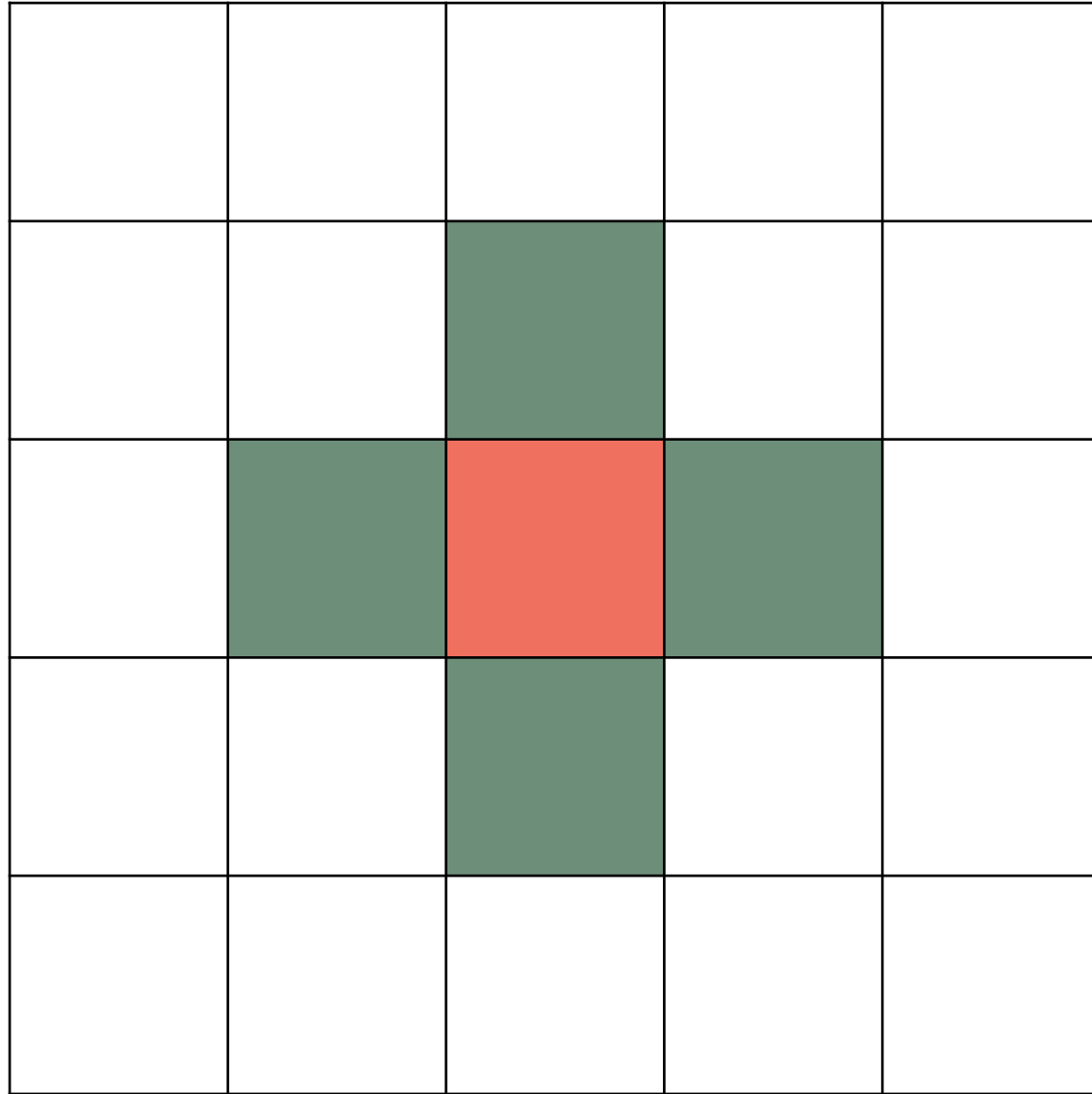




# Spatial weights matrix

- How to use the relationships between neighbouring features?
- Spatial weight matrix: an  $N \times N$  positive matrix ( $W$ ) that summarises all spatial relationships between the features in a dataset.
- It provides a formal expression of spatial dependency by quantifying the influence of each feature on every other feature based on their spatial proximity or other relevant criteria.

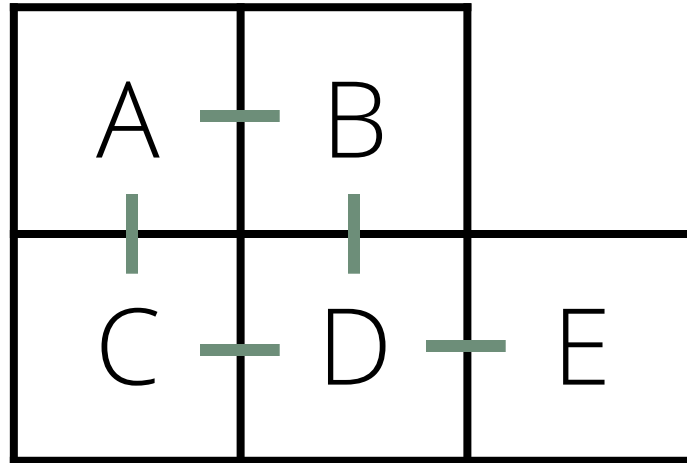
# Spatial weights matrix



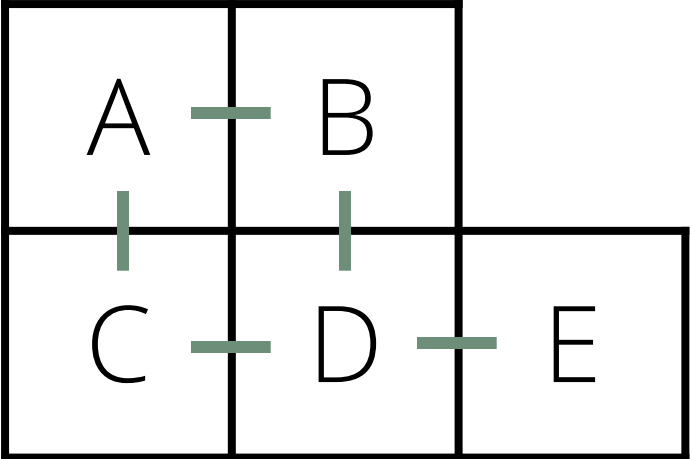
# Spatial weights matrix

A	B	
C	D	E

# Spatial weights matrix



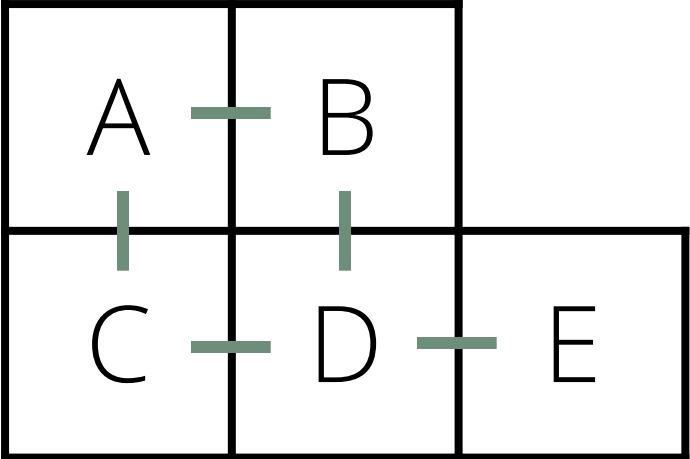
# Spatial weights matrix



	a	b	c	d	e
a	0	1	1	0	0
b	1	0	0	1	0
c	1	0	0	1	0
d	0	1	1	0	1
e	0	0	0	1	0



# Spatial weights matrix

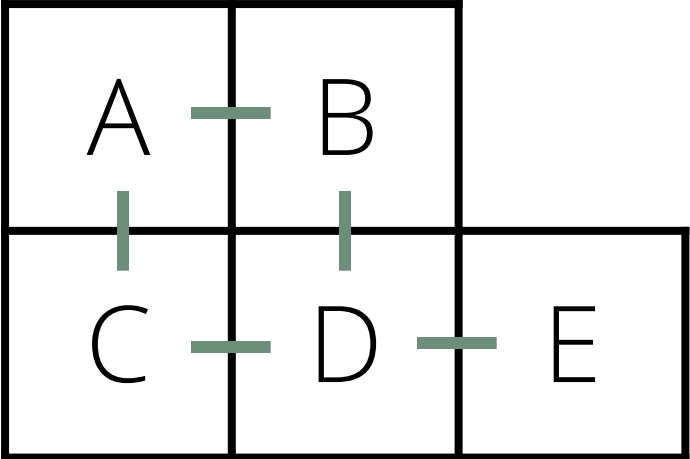


	a	b	c	d	e
a	0	1	1	0	0
b	1	0	0	1	0
c	1	0	0	1	0
d	0	1	1	0	1
e	0	0	0	1	0

=

w
2
2
2
3
1

# Spatial weights matrix

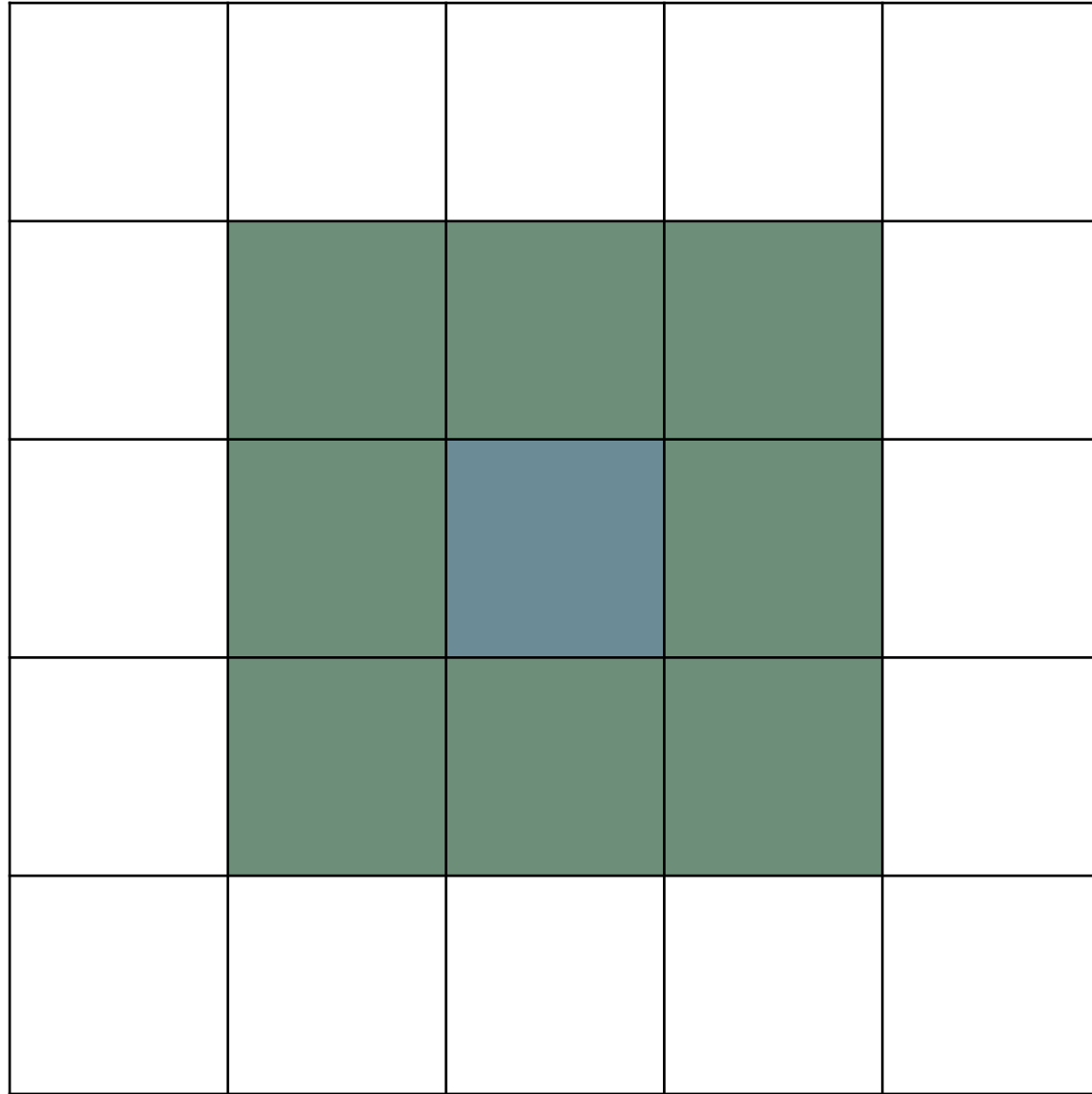


	a	b	c	d	e
a	0	.5	.5	0	0
b	.5	0	0	.5	0
c	.5	0	0	.5	0
d	0	.33	.33	0	.33
e	0	0	0	1	0

=

w
2
2
2
3
1

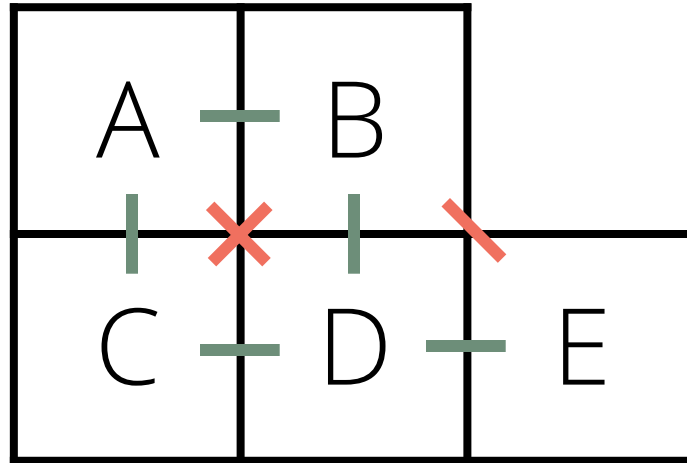
# Spatial weights matrix



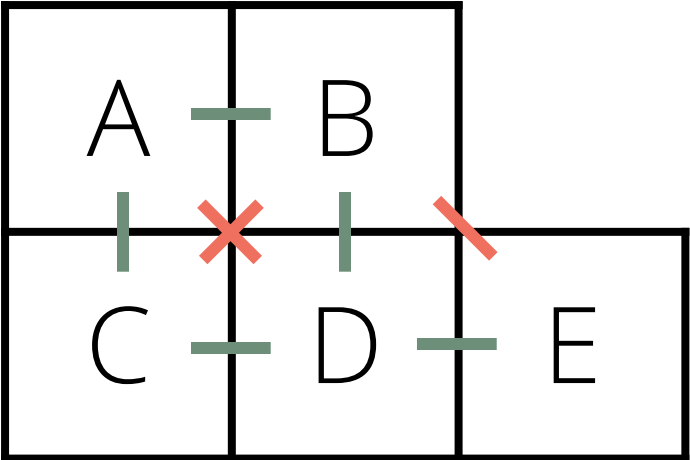
# Spatial weights matrix

A	B	
C	D	E

# Spatial weights matrix

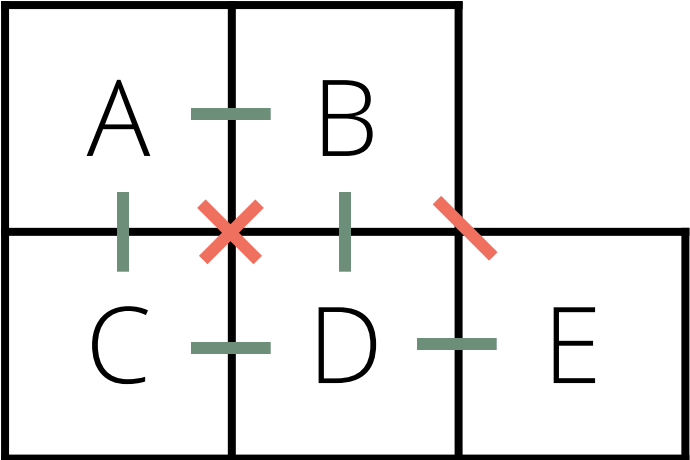


# Spatial weights matrix



	a	b	c	d	e
a	0	1	1	1	0
b	1	0	1	1	1
c	1	1	0	1	0
d	1	1	1	0	1
e	0	1	0	1	0

# Spatial weights matrix

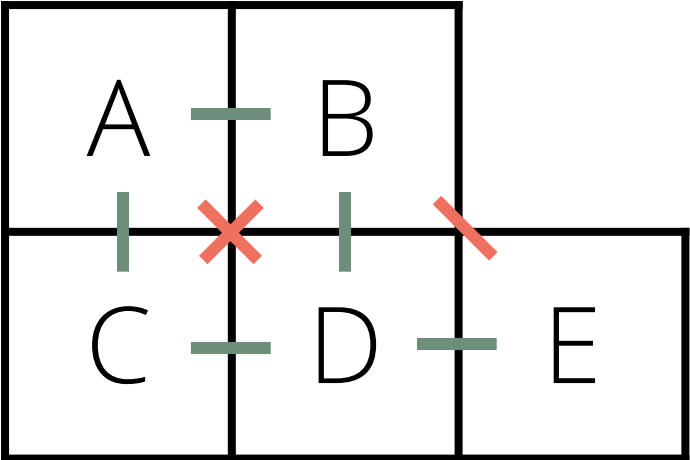


	a	b	c	d	e
a	0	1	1	1	0
b	1	0	1	1	1
c	1	1	0	1	0
d	1	1	1	0	1
e	0	1	0	1	0

=

w
3
4
3
4
2

# Spatial weights matrix



	a	b	c	d	e
a	0	.33	.33	.33	0
b	.25	0	.25	.25	.25
c	.33	.33	0	.33	0
d	.25	.25	.25	0	.25
e	0	.5	0	.5	0

=

w
3
4
3
4
2



# Topology



# Conclusion

- Measuring spatial autocorrelation is essential for understanding spatial relationships and patterns in data.
- We discussed two common measures of spatial autocorrelation, though other methods also exist.
- A spatial weights matrix is required to calculate the spatially lagged variable used in these measures.
- The way neighbours are defined (proximity or contiguity) can affect the results of spatial autocorrelation tests.

# Questions

Justin van Dijk  
[j.t.vandijk@ucl.ac.uk](mailto:j.t.vandijk@ucl.ac.uk)

