Week 7: 29th April 2025

Spatial Autocorrelation

Global Moran’s I
Local Moran’s I
Getis Ord (Gi*)

HYESOP SHIN

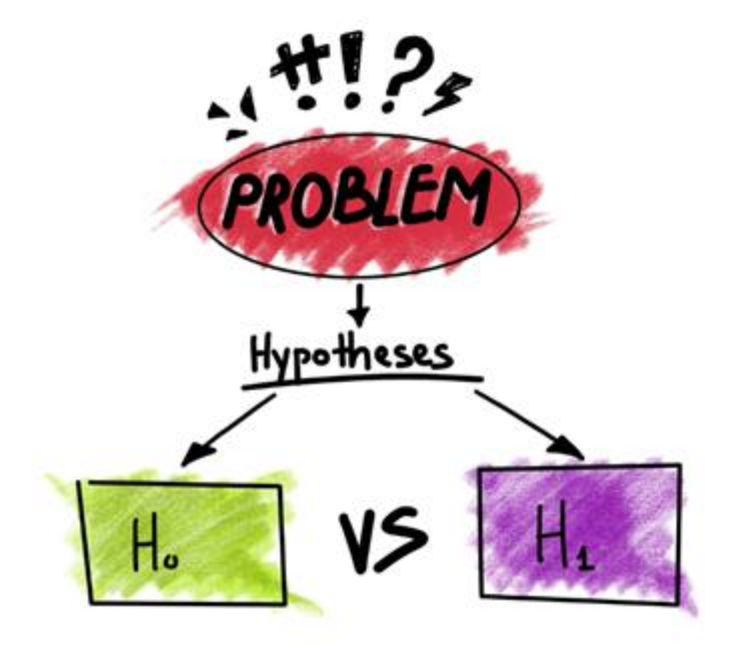
UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

Learning Objectives

- Key concepts: Correlation, spatial autocorrelation, hypotheses
- Know what are spatial weights
- Global Moran's I and Local Moran's I
- Four types of outcomes: HH, HL, LH, LL

Hypothesis?

- Explain what you expect to happen
- Be clear and understandable
- Be testable and measurable



Example:
Areas with more green spaces have lower air pollution levels than areas with fewer green spaces.

Null & Alternative Hypothesis

Null Hypothesis (H_0)

There is *no relationship* between the amount of green space and air pollution levels in an area

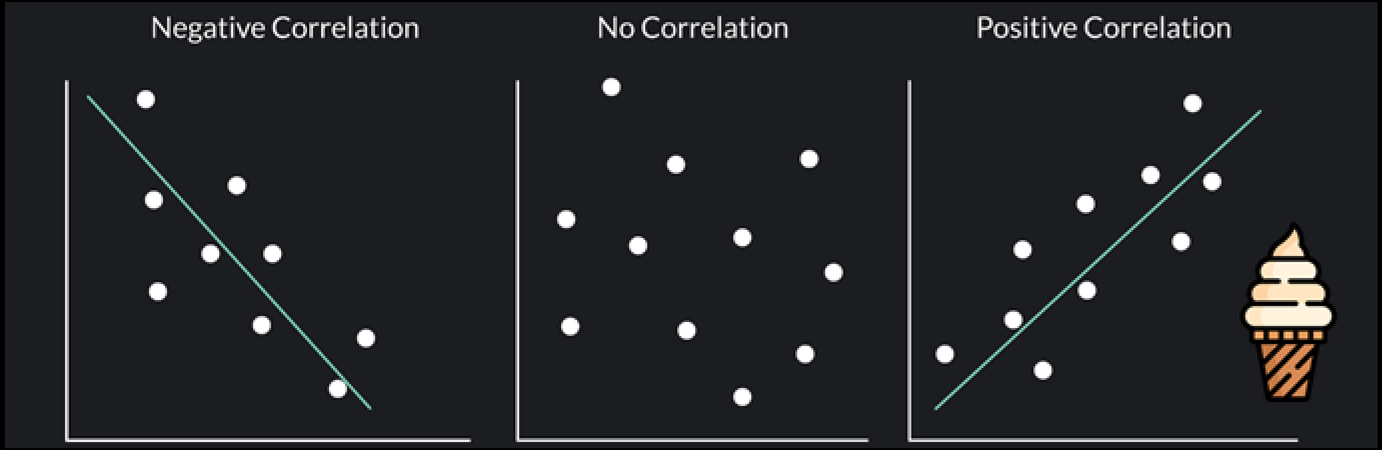
Alternative Hypothesis (H_1)

Areas with *more green spaces have lower air pollution* levels than areas with fewer green spaces.

Correlation

Correlation

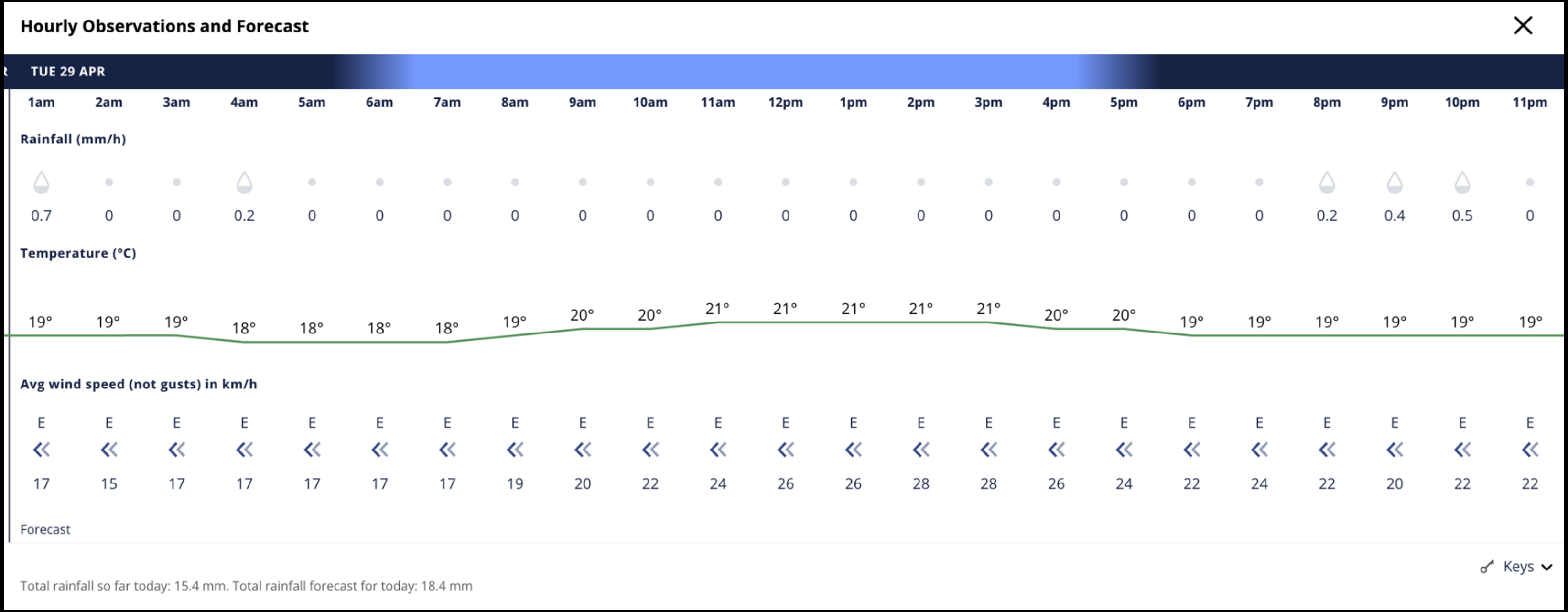
Correlation is a statistical measure that describes the relationship between two variables. It tells us whether an increase in one variable is associated with an increase or decrease in another.



Interest Rates x Housing prices

Temperature x Ice Cream Sales

Temporal Autocorrelation?



Hourly Observations and Forecast

TUE 29 APR

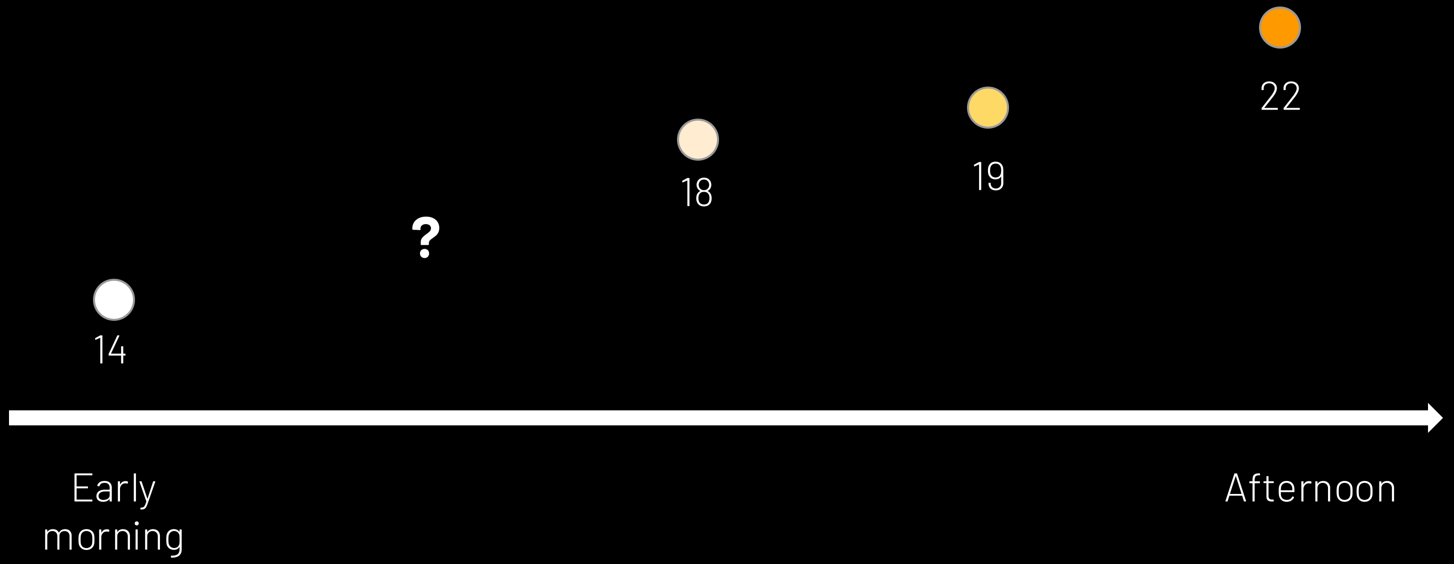
	1am	2am	3am	4am	5am	6am	7am	8am	9am	10am	11am	12pm	1pm	2pm	3pm	4pm	5pm	6pm	7pm	8pm	9pm	10pm	11pm
Rainfall (mm/h)	0.7	0	0	0.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.4	0.5	0
Temperature (°C)	19°	19°	19°	18°	18°	18°	18°	19°	20°	20°	21°	21°	21°	21°	21°	20°	20°	19°	19°	19°	19°	19°	19°
Avg wind speed (not gusts) in km/h	E 17	E 15	E 17	E 17	E 17	E 17	E 17	E 19	E 20	E 22	E 24	E 26	E 26	E 28	E 28	E 26	E 24	E 22	E 24	E 22	E 20	E 22	E 22

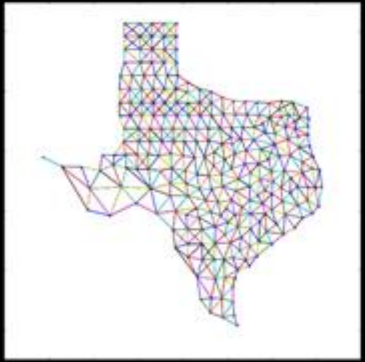
Forecast

Keys

Total rainfall so far today: 15.4 mm. Total rainfall forecast for today: 18.4 mm

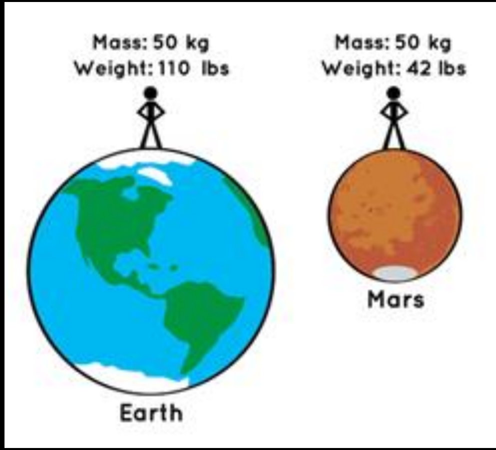
Temporal Autocorrelation?





Spatial Weights

How to include “space” in statistics?



UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

Spatial Weights

We need a mathematical representation of space
Geometries are spatial information but does not provide the spatial relationships

- Connectedness
- Proximity

But graphs do have neighbours and distances

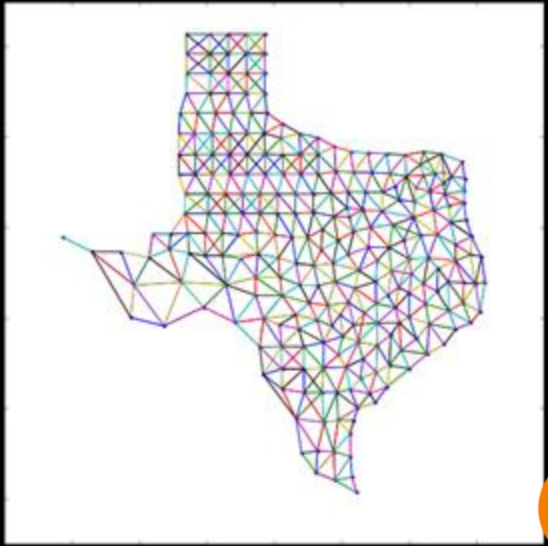
UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

What is a graph?

In this context a graph is:

- A data structure that consists of a set of objects called nodes and
- A set of connections between them called edges

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$



What is a neighbour?

It depends

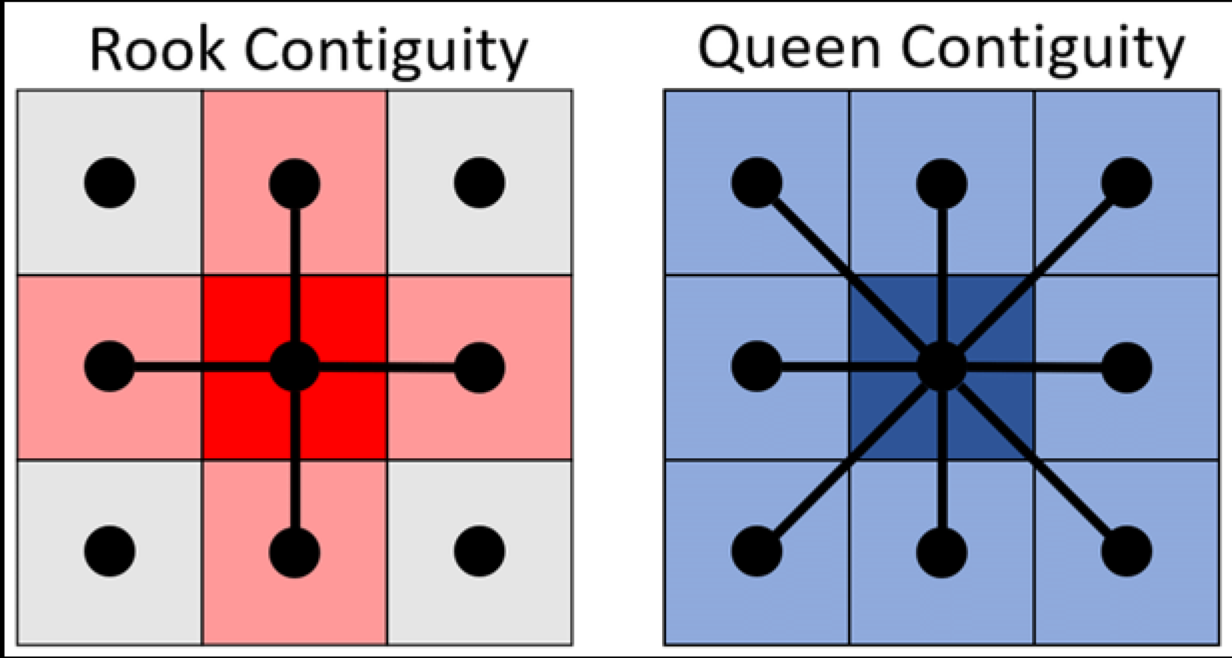
- What neighbourhoods are you surrounded by?
- How many primary schools within 5km radius from my home?

We are building a topology that we can use to examine the data

Topology: a mathematical structure that expresses the connectivity between observations

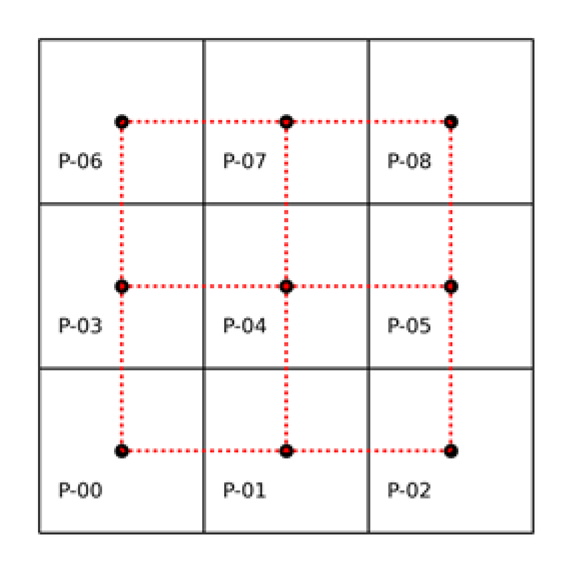
Contiguity

Contiguity means that two spatial units share a common border of non-zero length

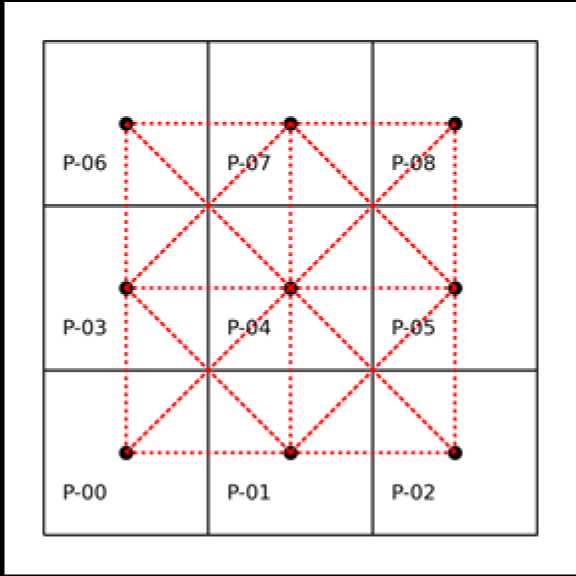


Graph Visualisation

Rook

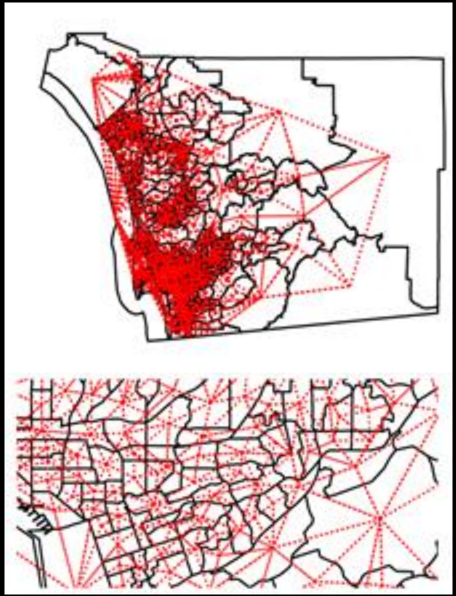


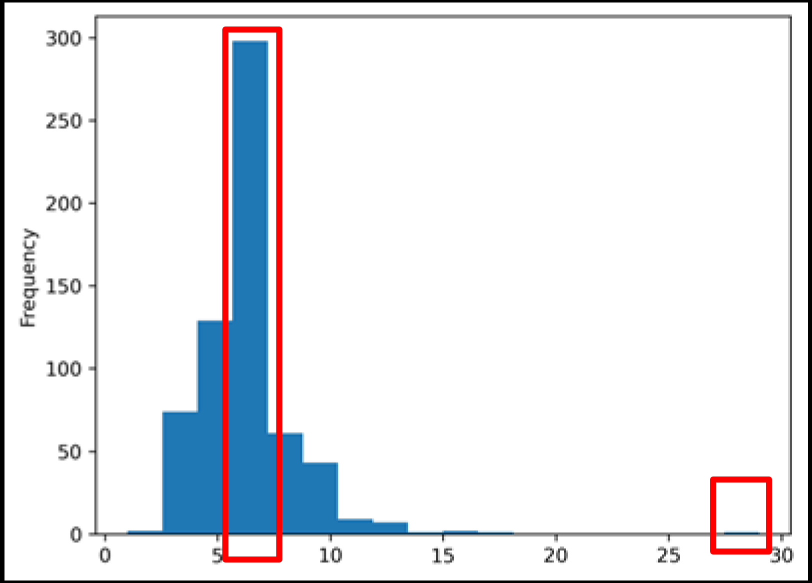
Queen



Spatial weights from real-world geographic tables

Queen Contiguity

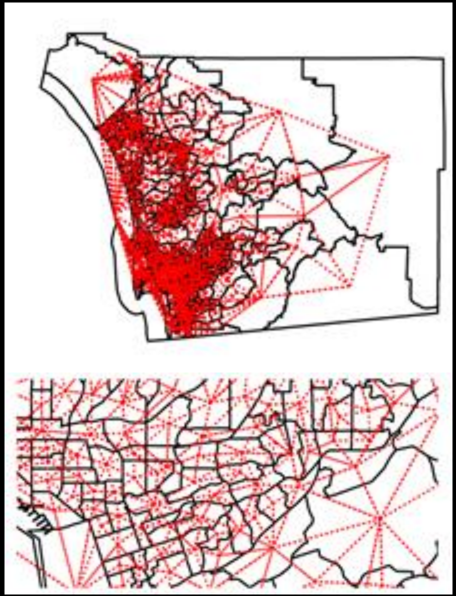


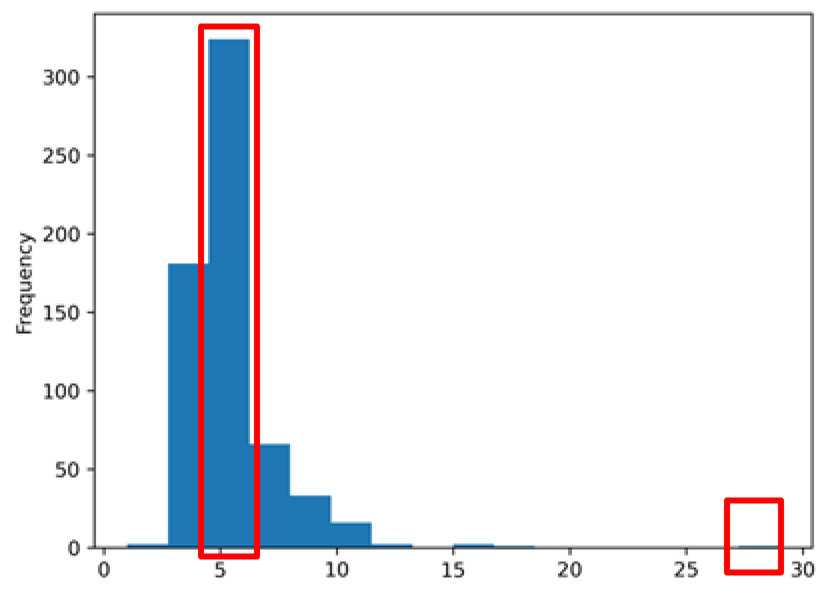


San Diego census tract

Spatial weights from real-world geographic tables

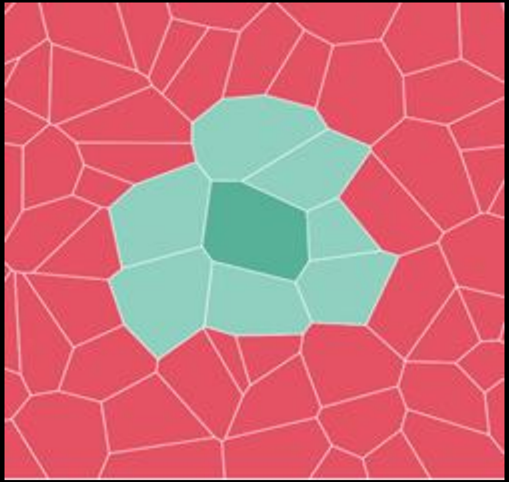
Rook Contiguity



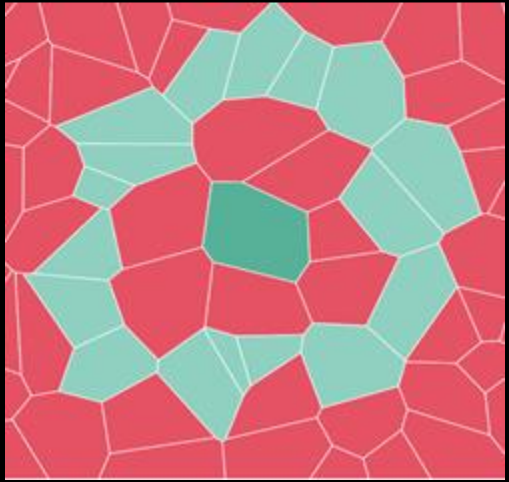


San Diego census tract

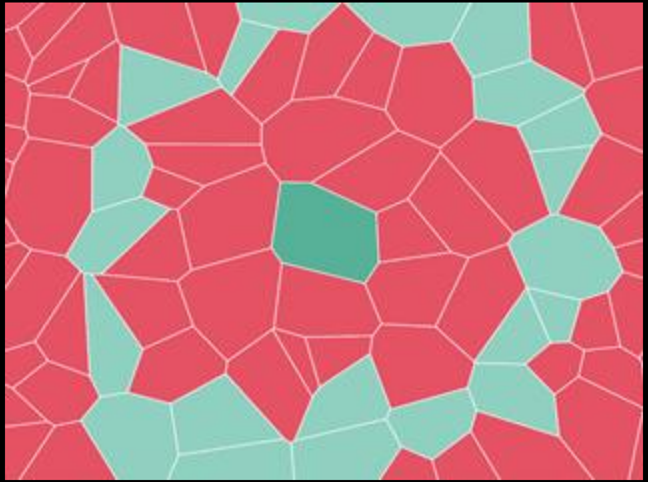
Order of Contiguity



1st order contiguity



2nd order contiguity



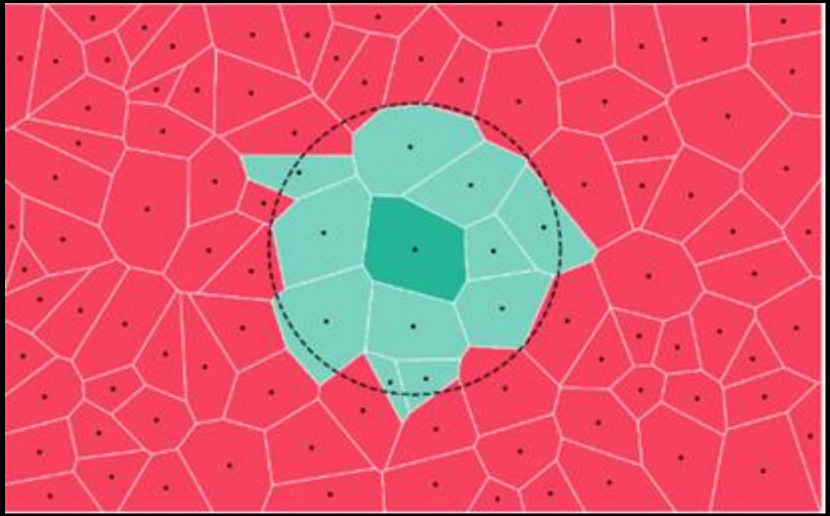
3rd order contiguity

UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

Distance-based

Neighbours can be defined by distances

Two geometries are considered neighbours if they lie within a set threshold from each other

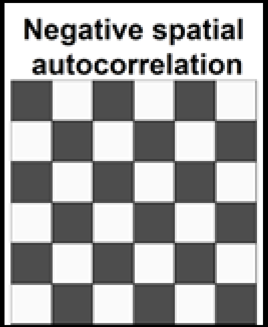


Spatial Autocorrelation

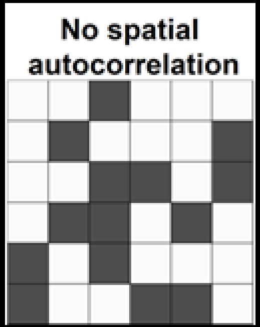
Spatial “auto”correlation

Everything is related to everything else, but near things are more related than distant things (Waldo Tobler, 1970).

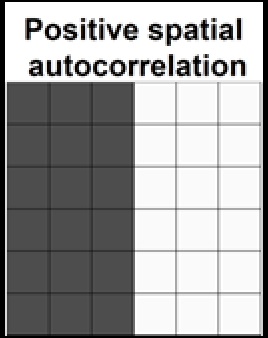
- Degree to which similar values are located in similar locations



Similar values → dissimilar location (further apart)



Random



Similar values → similar location (closeby)

Example

Positive SA

Income
Poverty
Vegetation
Temperature

Negative SA

Supermarkets
Police stations
Fire stations
Hospitals...

Scales

[Global]
Clustering: do values tend to be close to other (dis)similar values?

[Local]
Clusters: are there any specific parts of a map with an extraordinary concentration of (dis)similar values?

Global Spatial Autocorrelation

"Clustering"

Overall trend where the distribution of values follows a particular pattern over space

The premise is to understand how similar are you compared to your neighbourhood

The null hypothesis assumes spatial randomness – that there is no spatial autocorrelation in the data.

[Positive] Similar values close to each other (high-high, low-low)

[Negative] Similar values far from each other (high-low)

How do we measure it?

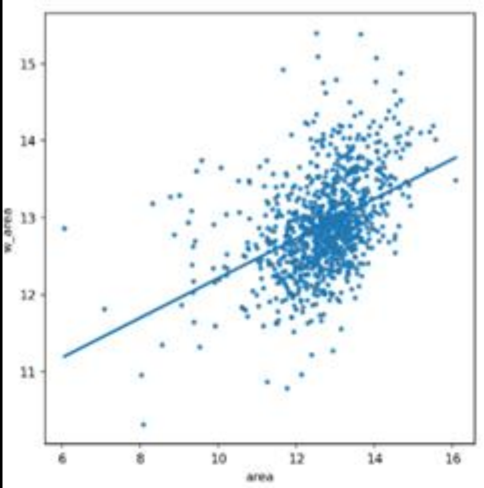
UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

Moran Plot

- Graphical device that displays a variable on the horizontal axis against its spatial lag* on the vertical one

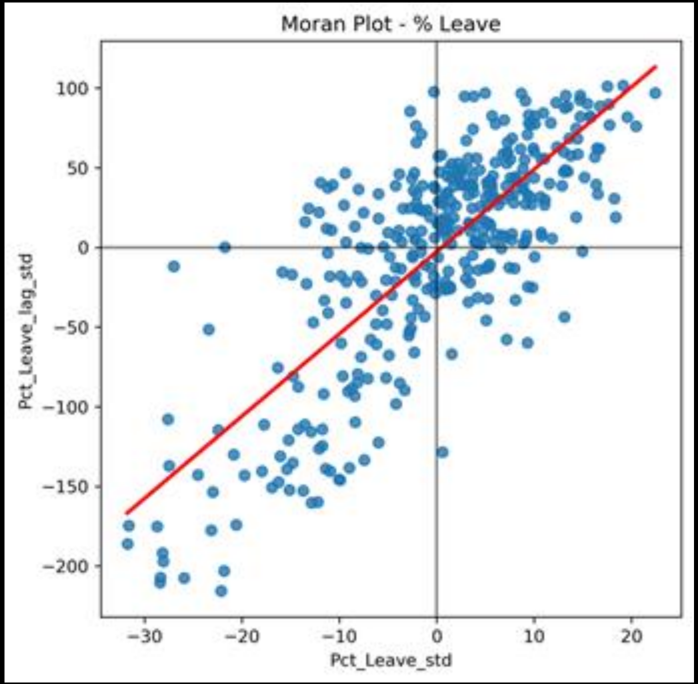
- Assessment of the overall association between a variable in a given location and in *its neighbourhood*

*Spatial lag: the average value of the variable at neighbouring locations)

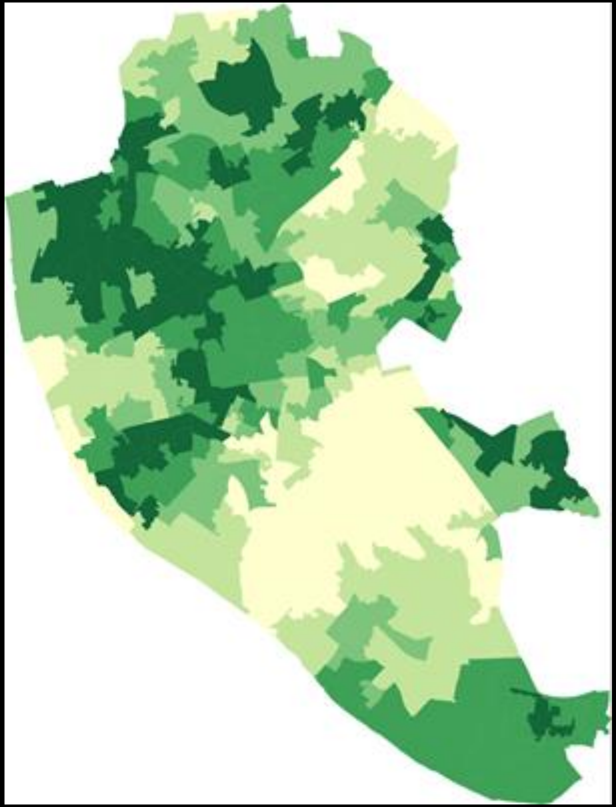


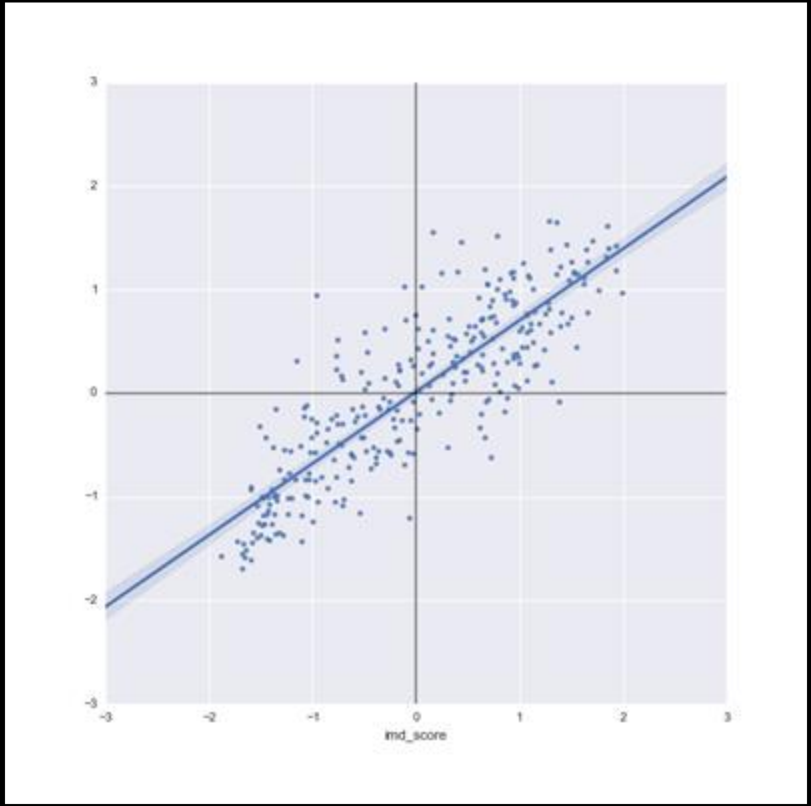
Moran's I

- Formal test of global spatial autocorrelation
- Statistically identify the presence of clustering in a variable
- Slope of the Moran plot
- Inference based on how likely it is to obtain a map like observed from a purely random pattern
- Value Range: 0 and 1



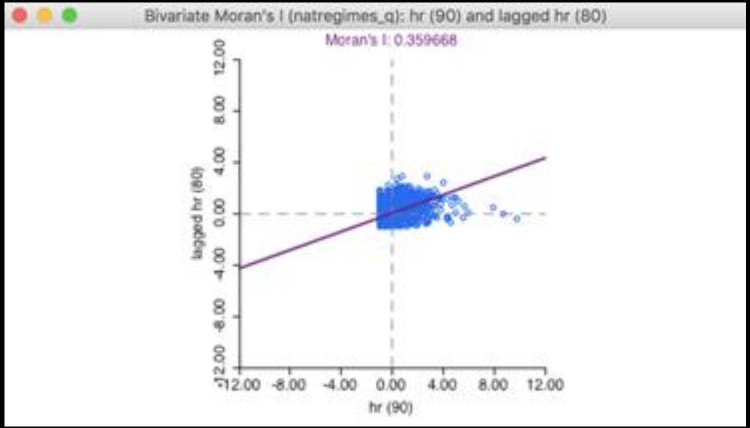
Graphic illustration

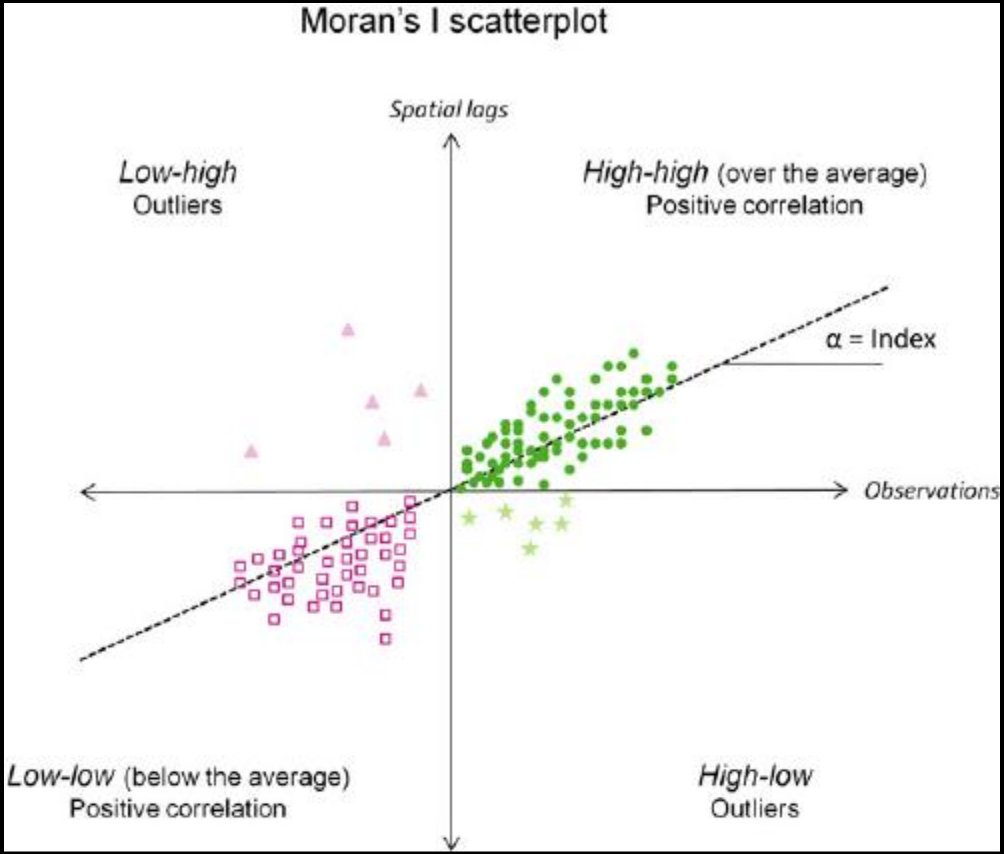




I statistics

I values between 0 and 1





How do we compute Moran's I?

Get your Choropleth map ready

Moran's I

Steps

1. State the null and alternative hypotheses

2. Calculate Queen contiguity

3. Extract NZDep variable

4. Standardise weights

5. Calculate Moran's I (and look at the p-value)

If p-value < α, we reject the null hypothesis.
If p-value $\geq \alpha$, we fail to reject the null hypothesis

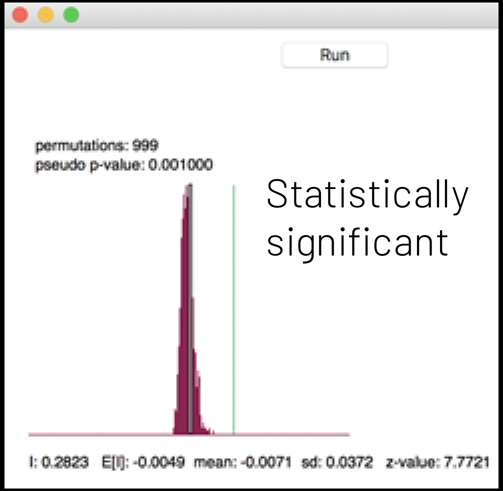
Check hypothesis: Permutation

Permutation (randomly distributing)

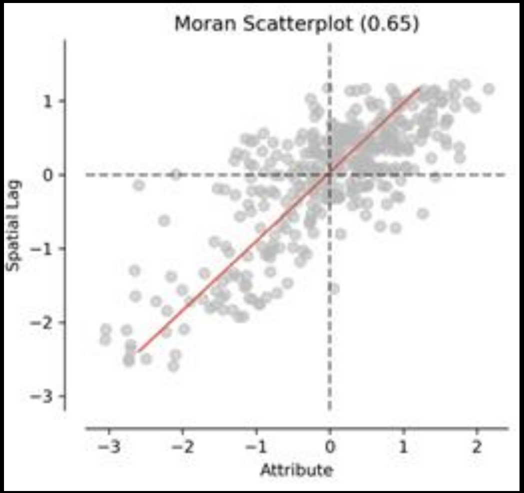
Permutations can be 9, 99, and 999

Checks how much the analysed I are different to the permuted results

Permuted: Histogram
Green: your Global moran's I



Statistically significant



UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

Limitations of Global Spatial Autocorrelation (e.g., Moran's I)

Global statistics like Moran's I summarise spatial patterns across the entire map.

Moran's I captures overall clustering (positive values) or dispersion (negative values) in your dataset. However, it does not tell you where clustering or dispersion occurs.

🧭 Think of it like getting an average temperature for a country – useful, but it doesn't tell you where the heatwaves or cold spots are!

To identify specific local patterns (e.g., clusters of high or low values), we need local spatial statistics (e.g., Local Moran's I or Getis-Ord Gi*).

Local Moran's I

Local measures of spatial autocorrelation focus on the relationships between each observation and its surroundings

Also known as _Local Indicators of Spatial Association (LISA)_ (Anselin 1995)

Clusters

High-High

Positive SA of high values (hotspots)

Low-Low

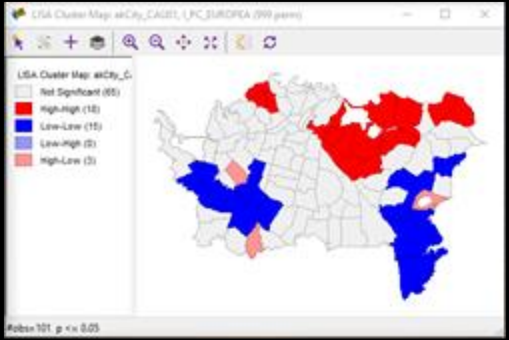
Positive SA of low values (coldspots)

High-Low

Negative SA (spatial outliers)

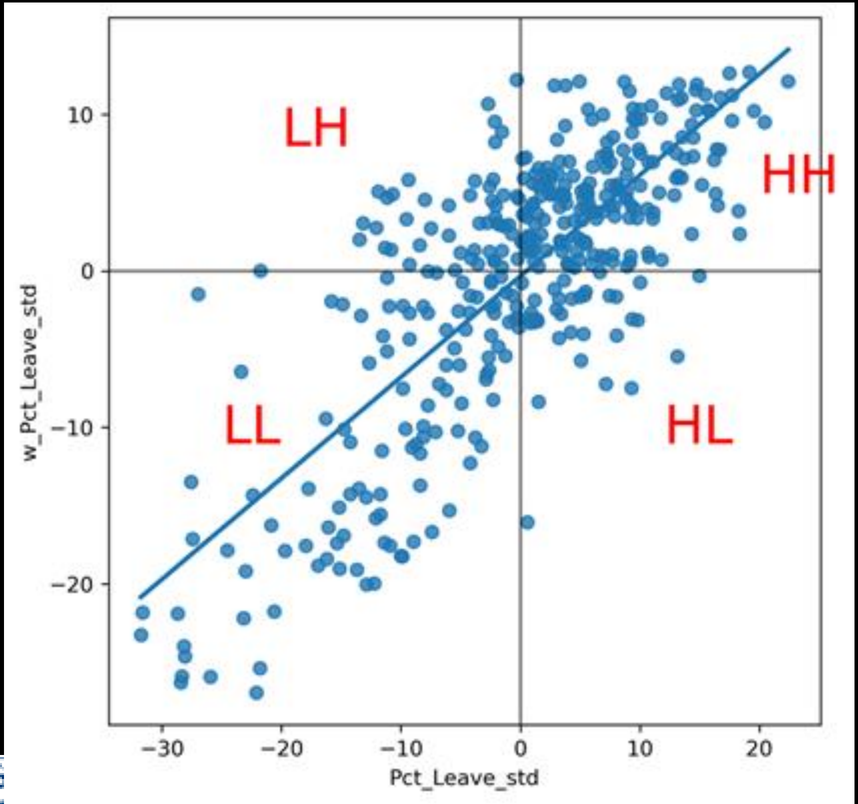
Low-High

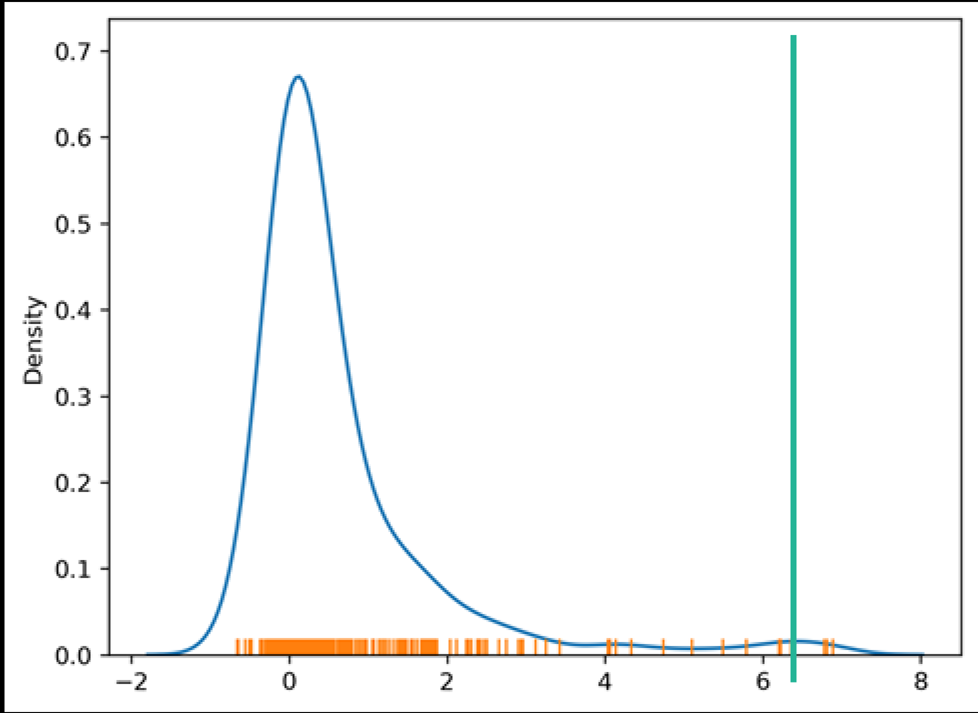
Negative SA (spatial outliers)



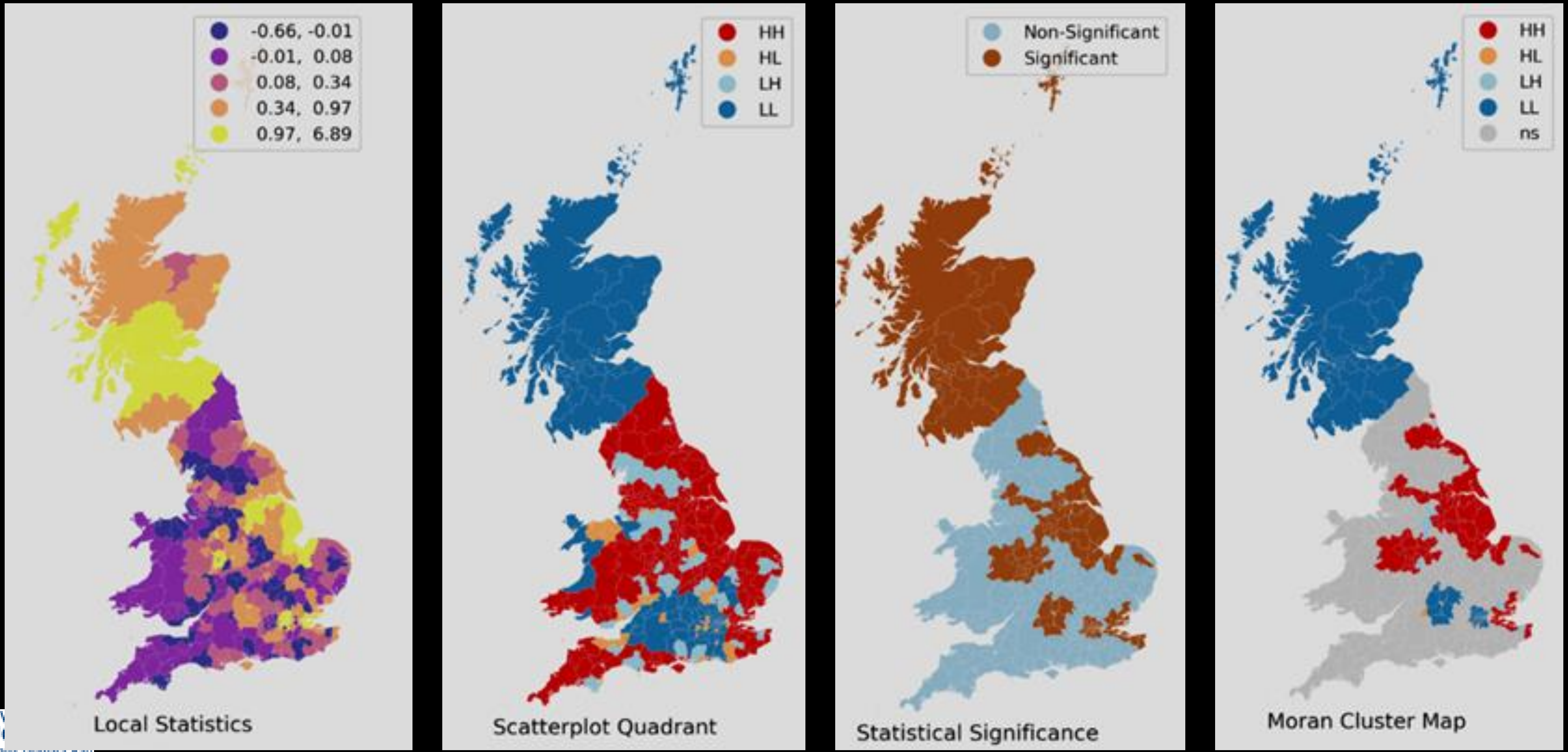
UNIVERSITY OF AUCKLAND Waipapa Taumata Rau NEW ZEALAND

Example: UK Referendum





Interpreting Results



Frequency Table

Category	Count	Percentage
Non-Significant	226	59.47%
HH	75	19.74%
LL	69	18.16%
LH	6	1.58%
HL	4	1.05%

Your turn

Getis Ord

Getis-Ord Gi* is a local spatial statistic used to identify hot spots (high values surrounded by high values) and cold spots (low values surrounded by low values) in geospatial data.

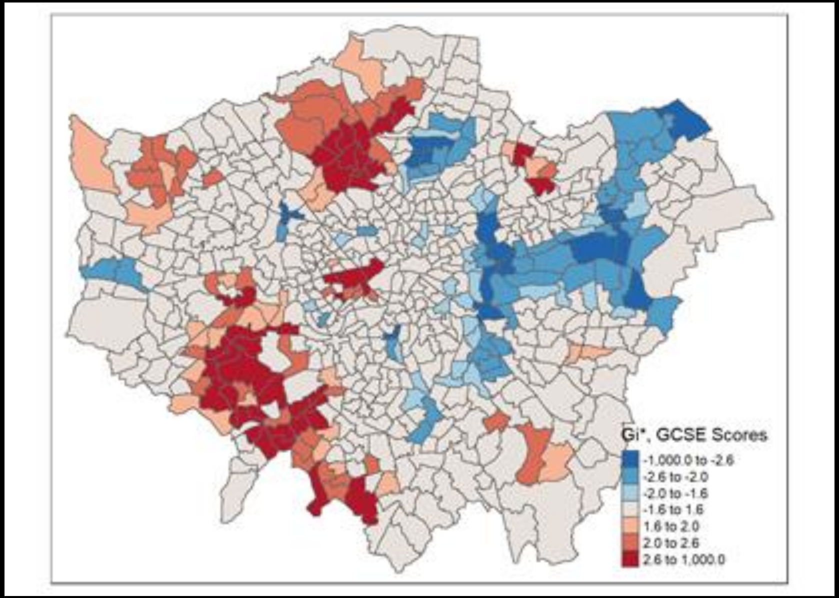
✅ High Gi* → Hot spots (clusters of statistically high values).
✅ Low Gi* → Cold spots (clusters of statistically low values).

You can get both Gi and Gi*

The * includes whether you add yourself into the computation

How it works

- For each feature/location i, the Gi* statistic sums the attribute values of i and its neighbours, weighted by spatial proximity.
- It then compares this local sum to the global sum of all attribute values.
- The result is a Z-score, which tells you whether high (or low) values are spatially clustered.



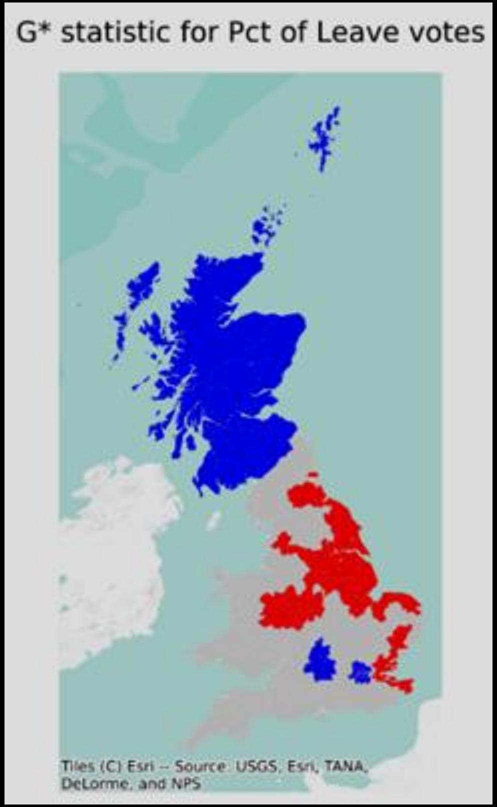
UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

LISA vs Getis Ord

Feature	Local Moran's I	Getis-Ord Gi*
Type	Local measure	Local measure
Measure	Similarity/dissimilarity between a location and its neighbours	Intensity of clustering of high or low values
Purpose	Detects overall spatial autocorrelation	Identifies hot/cold spots
Detects	Clusters and spatial outliers	Only clusters (hot/cold spots)
Based on	Deviation from mean	Sum of weighted values in neighbourhood

UNIVERSITY OF AUCKLAND
Waipapa Taumata Rau
NEW ZEALAND

Gi* Example


G* statistic for Pct of Leave votes
Tiles (C) Esri -- Source: USGS, Esri, TANA, DeLorme, and NPS

Example Use Cases for Getis-Ord Gi*

Crime Analysis – Identifying crime hot spots for targeted policing.

Urban Planning – Locating areas needing infrastructure investment.

Public Health – Detecting disease outbreak clusters.

Retail Analysis – Finding high-density consumer areas.

Your turn

Lecture Summary

Spatial Weights: Define proximity (contiguity, distance-based)

Spatial Autocorrelation: Measures how similar values cluster in space

Global Moran's I: Detects overall clustering (HH, LL = positive; HL, LH = negative)

Local Moran's I (LISA): Finds local clusters/outliers (HH, LL, HL, LH)

Getis-Ord Gi*: Identifies hot and cold spots

Use: Map patterns, test hypotheses, guide spatial decisions

Questions?

If not, I will all see you tomorrow with some example data

Reference

- Ray et al (2023), Geographic Data Science with Python, CRC Press
- Fleischmann (2024), Spatial Data Science in Python
- Dorman, M., Graser, A., Nowosad, J., & Lovelace, R. (2025). Geocomputation with Python. CRC Press