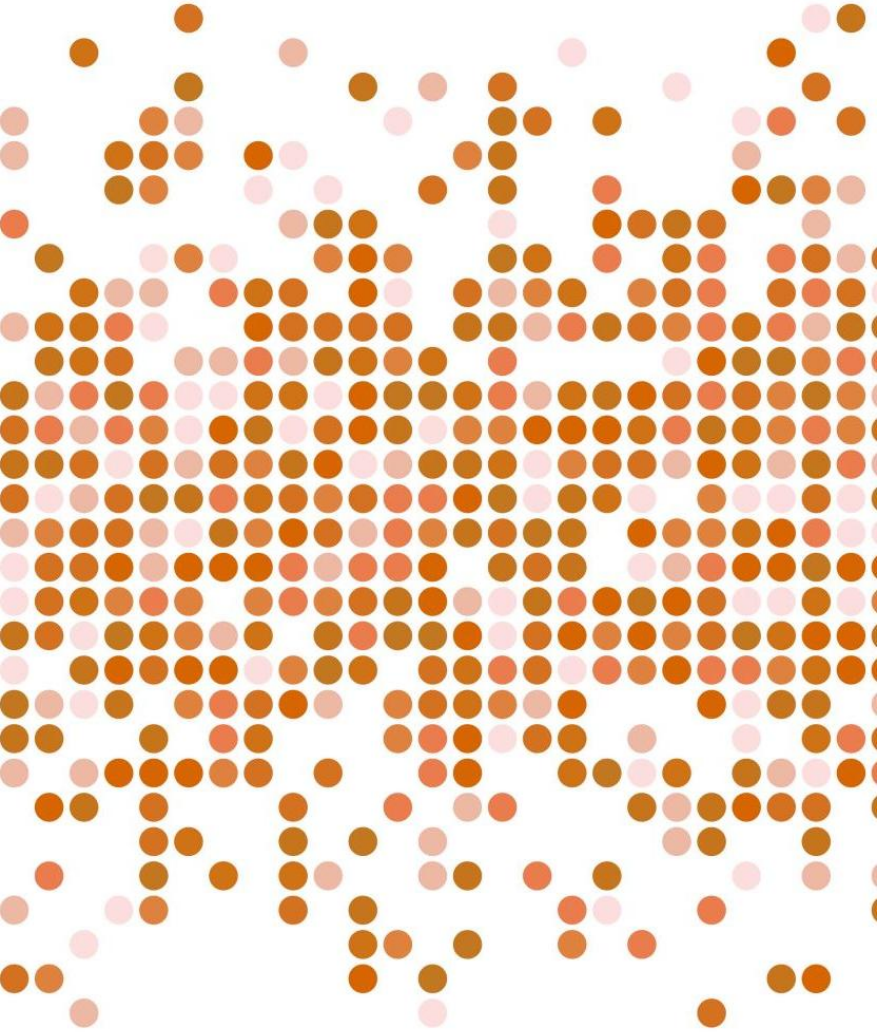
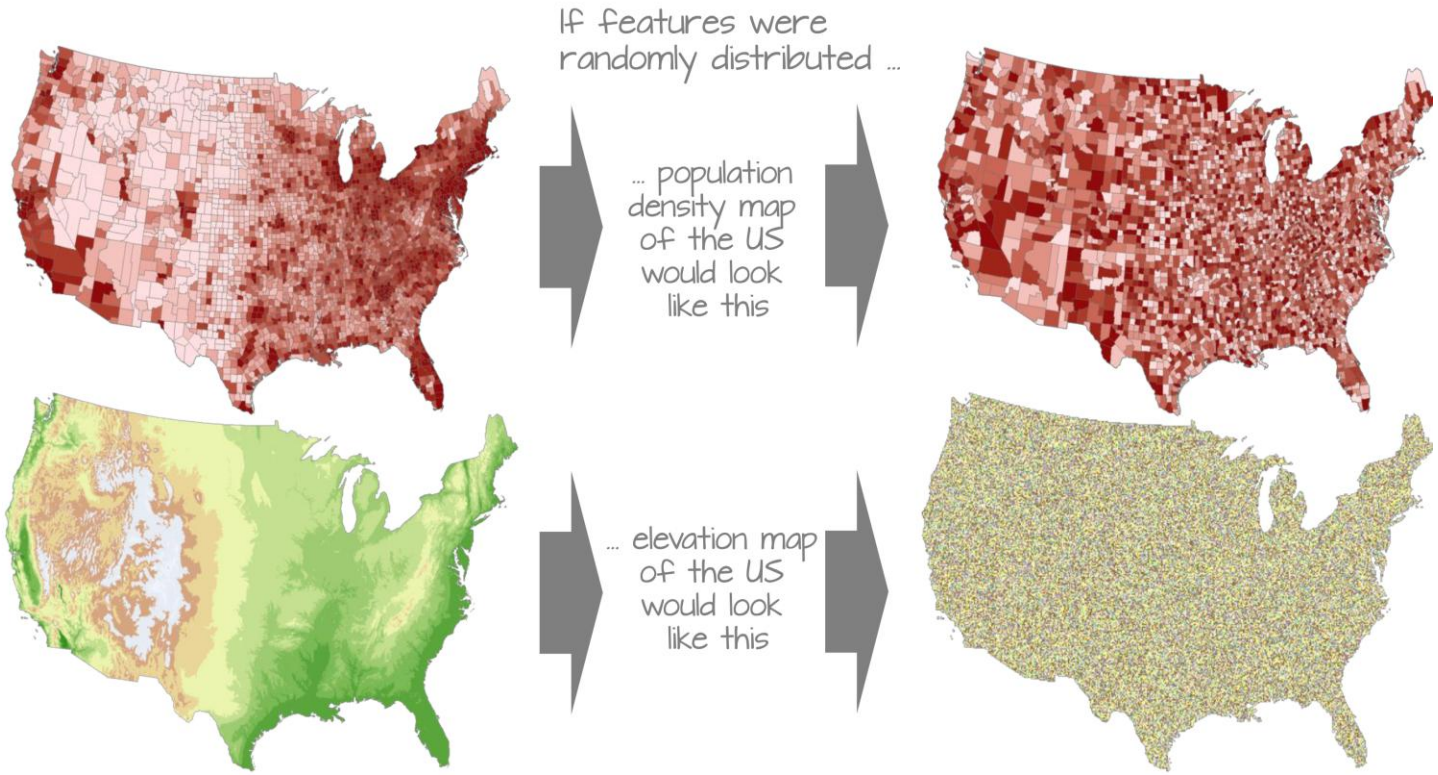


Spatial Autocorrelation and Local Moran's I

GGS 563

What is Spatial Autocorrelation?

- Measures the degree to which similar values occur near each other in space.
- Indicates whether spatial patterns are clustered, dispersed, or random.
- Considers location and attribute values together.
- Aligns closely with Tobler's First Law of Geography: "*Everything is related to everything else, but near things are more related than distant things*."



Spatial data normally includes non-spatial information (attributes). When mapped, these values often demonstrates some degree of spatial relatedness at some scale.

Importance of Spatial Autocorrelation

Detecting spatial patterns

- Identifies clusters and outliers

Improving spatial analysis, modeling and interpretation

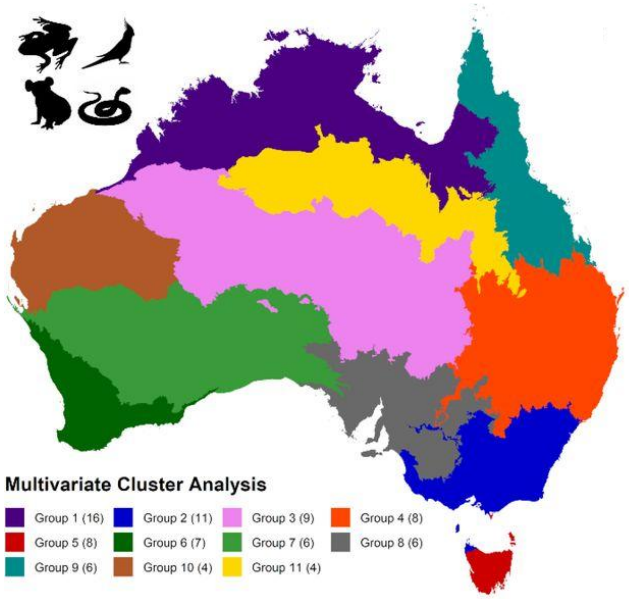
- Highlights underlying processes

Avoiding statistical assumptions

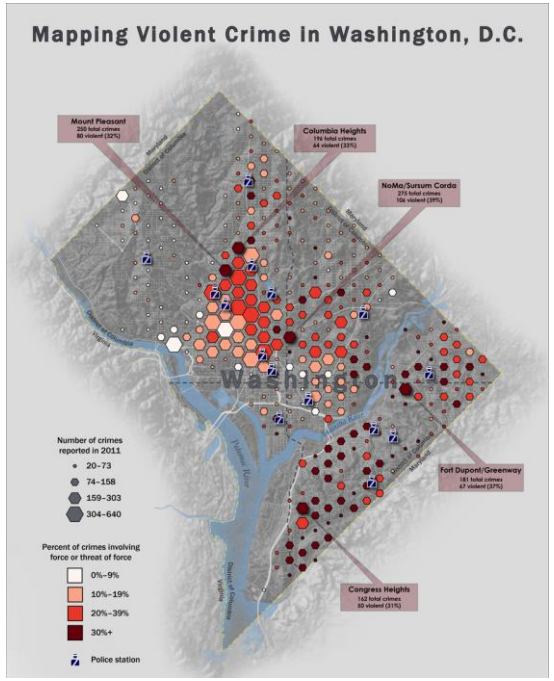
- Many statistical methods assume data independence → spatial autocorrelation violates this

Spatial Autocorrelation Applications

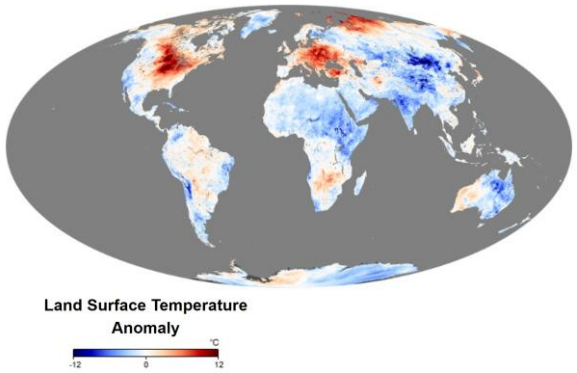
Discipline	Applications
Public Health	• Detecting disease clusters • Mapping low vaccination rates • Identifying environmental health risk areas
Urban Planning	• Analyzing housing price clusters • Detecting concentrated poverty or wealth • Locating traffic accident hot spots
Environmental Science	• Mapping deforestation patterns • Studying species distribution • Identifying high soil erosion risk areas
Crime Analysis	• Finding burglary or violent crime hot spots • Guiding patrol allocation
Agriculture	• Identifying yield clusters • Mapping pest or crop disease spread
Climate & Meteorology	• Tracking temperature anomalies • Identifying clusters of extreme weather events



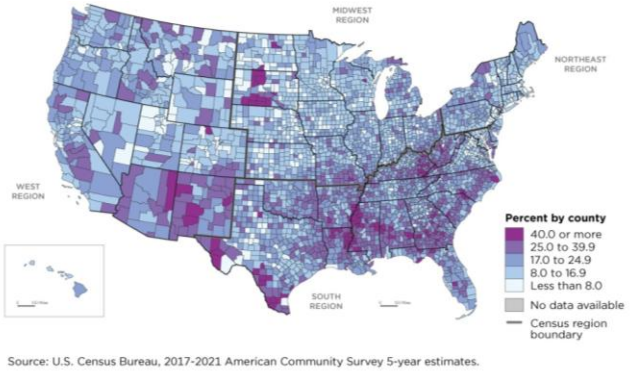
Spatial configuration of the optimal clusters for species diversity of vertebrates in Australia



Violent crime clustering in Washington D.C.



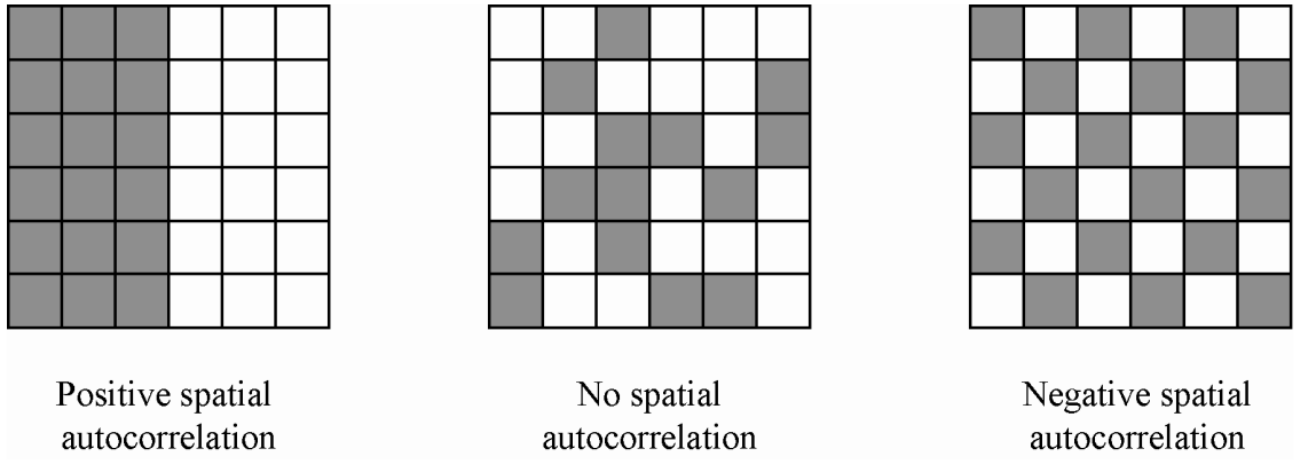
Global LST anomalies in February 2024



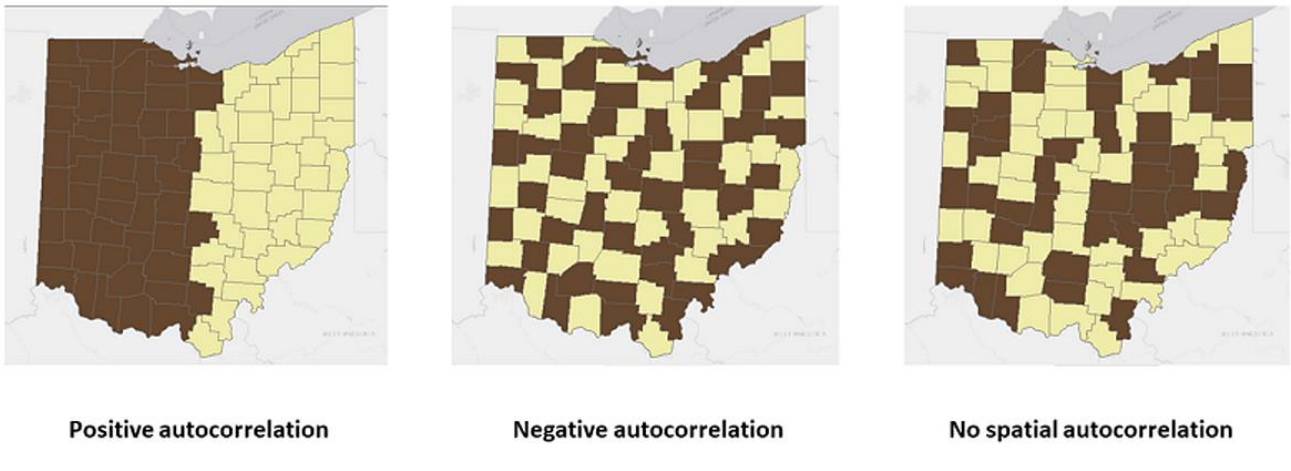
Spatial patterns of child poverty rates

Types of Spatial Autocorrelation

- **Positive spatial autocorrelation:** similar values are near each other
- **Negative spatial autocorrelation:** dissimilar values are near each other
- **No spatial autocorrelation:** values are randomly distributed

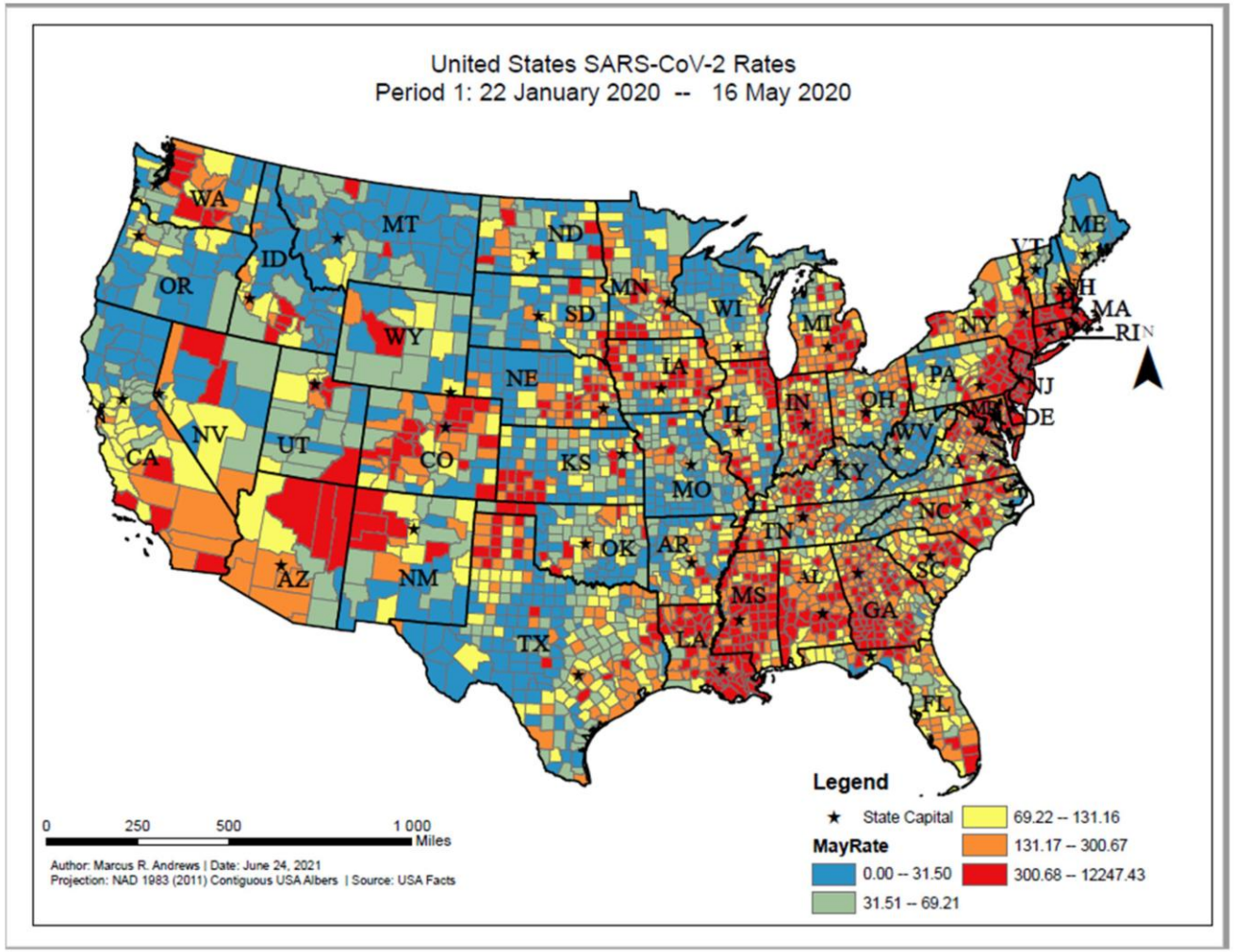


Different types of spatial autocorrelation with grid representation .



Different types of spatial autocorrelation with map representation .

Identifying Spatial Autocorrelation

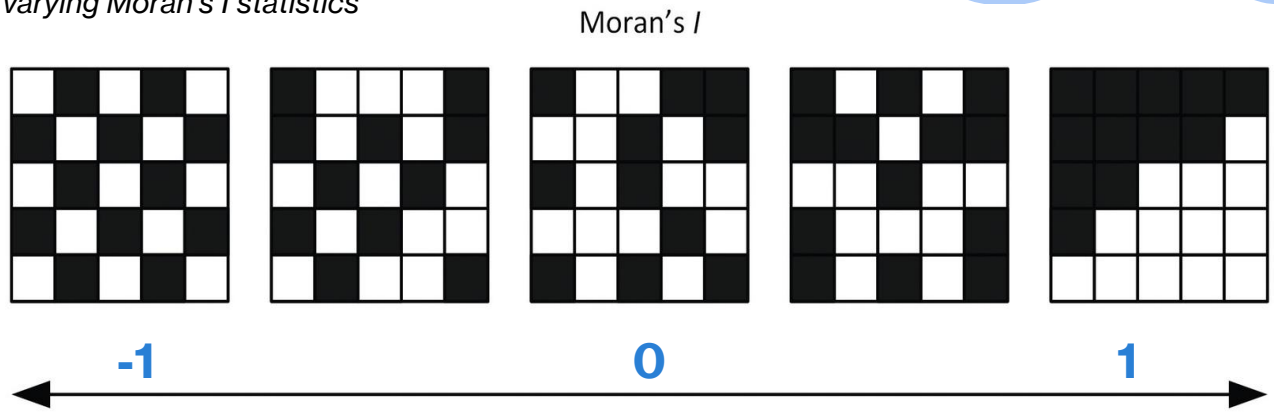


Spatial clustering of county-level COVID-19 rates in the U.S.

Moran's I

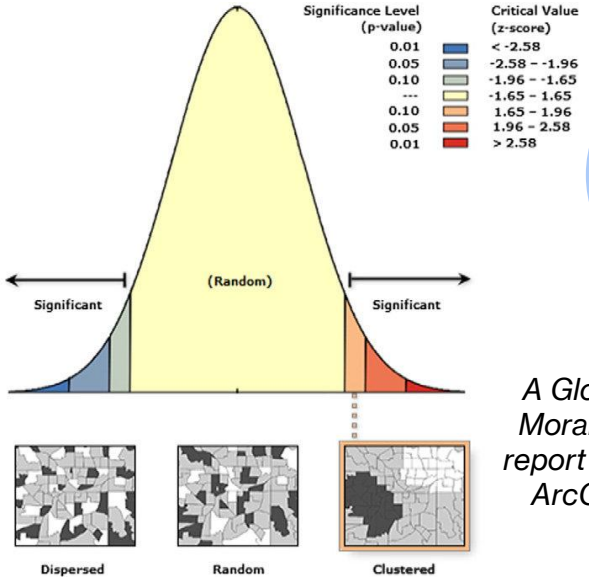
- A widely used statistical measure to assess spatial autocorrelation
  - Calculates a statistic between -1 and 1
    - *Positive values* = clustering
    - *Negative values* = dispersion
    - *Values close to 0* = random pattern
- Part of the **Local Indicators of Spatial Association (LISA)** family
- **Global Moran's I:** One statistic summarizing the entire study area
- **Local Moran's I**: A local measure of spatial autocorrelation that identifies **clusters** and spatial **outliers** for each feature in the study area → one statistic per feature in the study area

Example of patterns based on varying Moran's I statistics



$$I_i = z_i \sum_j w_{ij} z_j$$

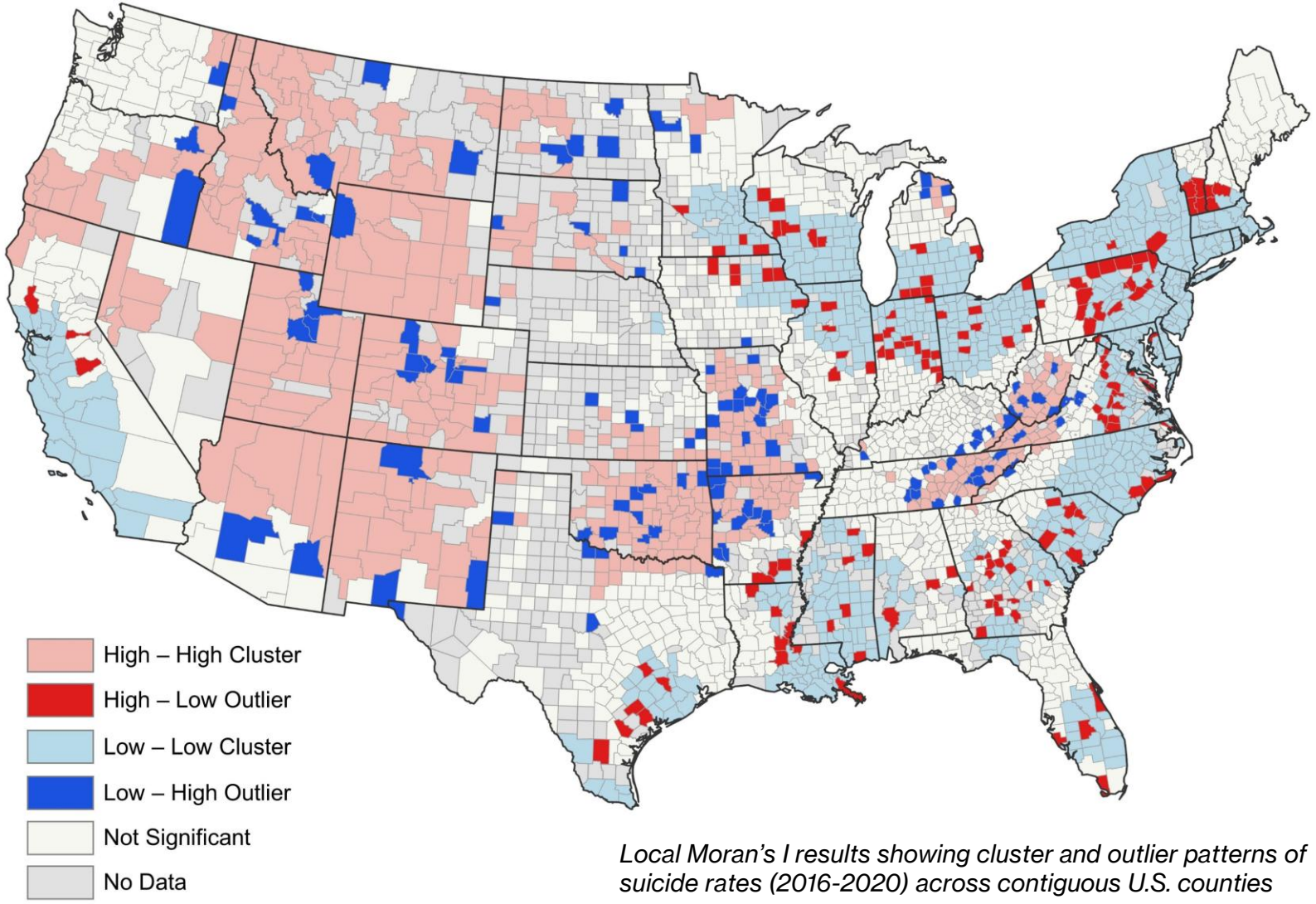
Equation for computing Local Moran's I statistic



A Global Moran's I report from ArcGIS

Visualizing Local Moran's I

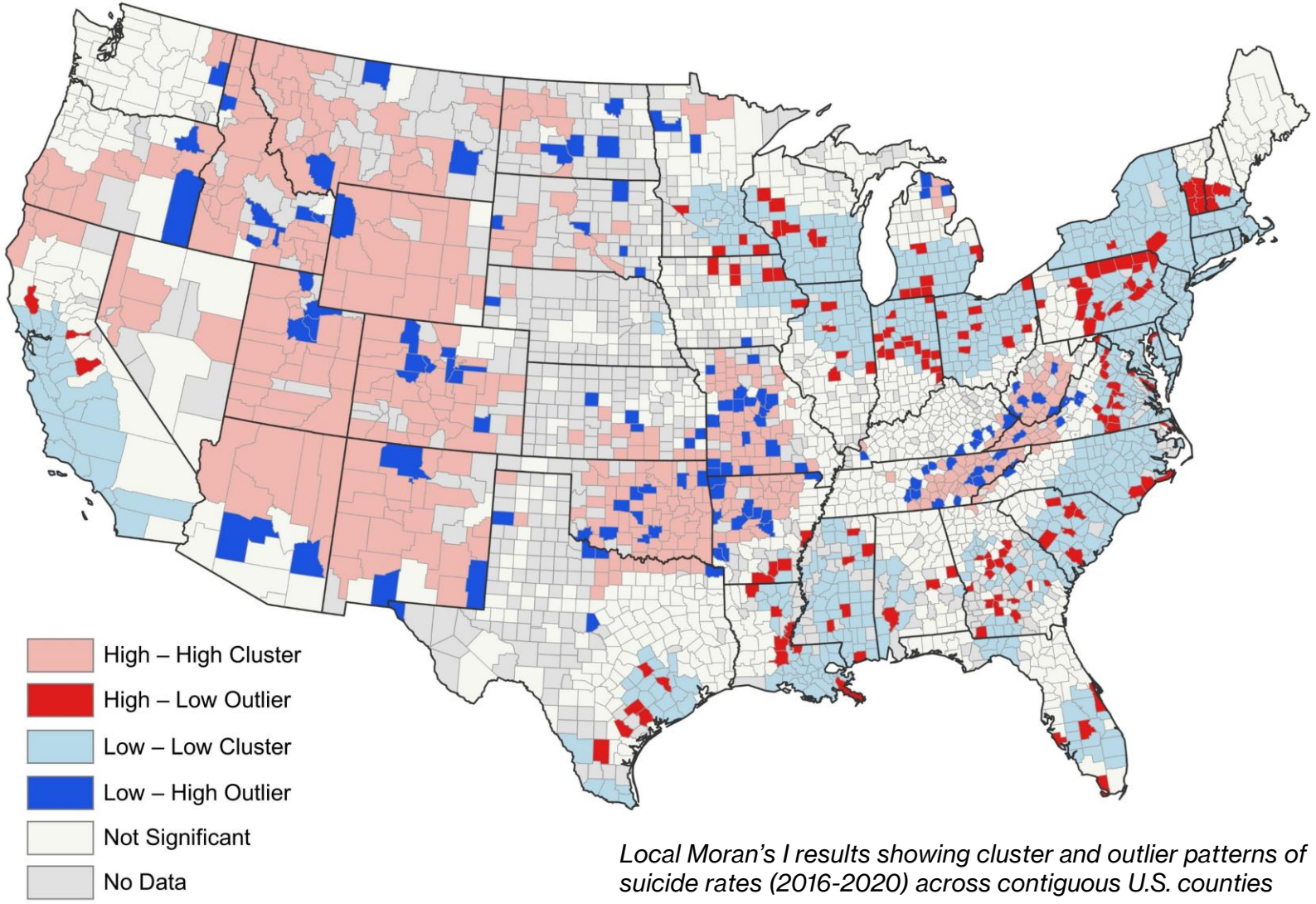
High – High cluster:
High values surrounded by high values



Local Moran's I results showing cluster and outlier patterns of suicide rates (2016-2020) across contiguous U.S. counties

Visualizing Local Moran's I

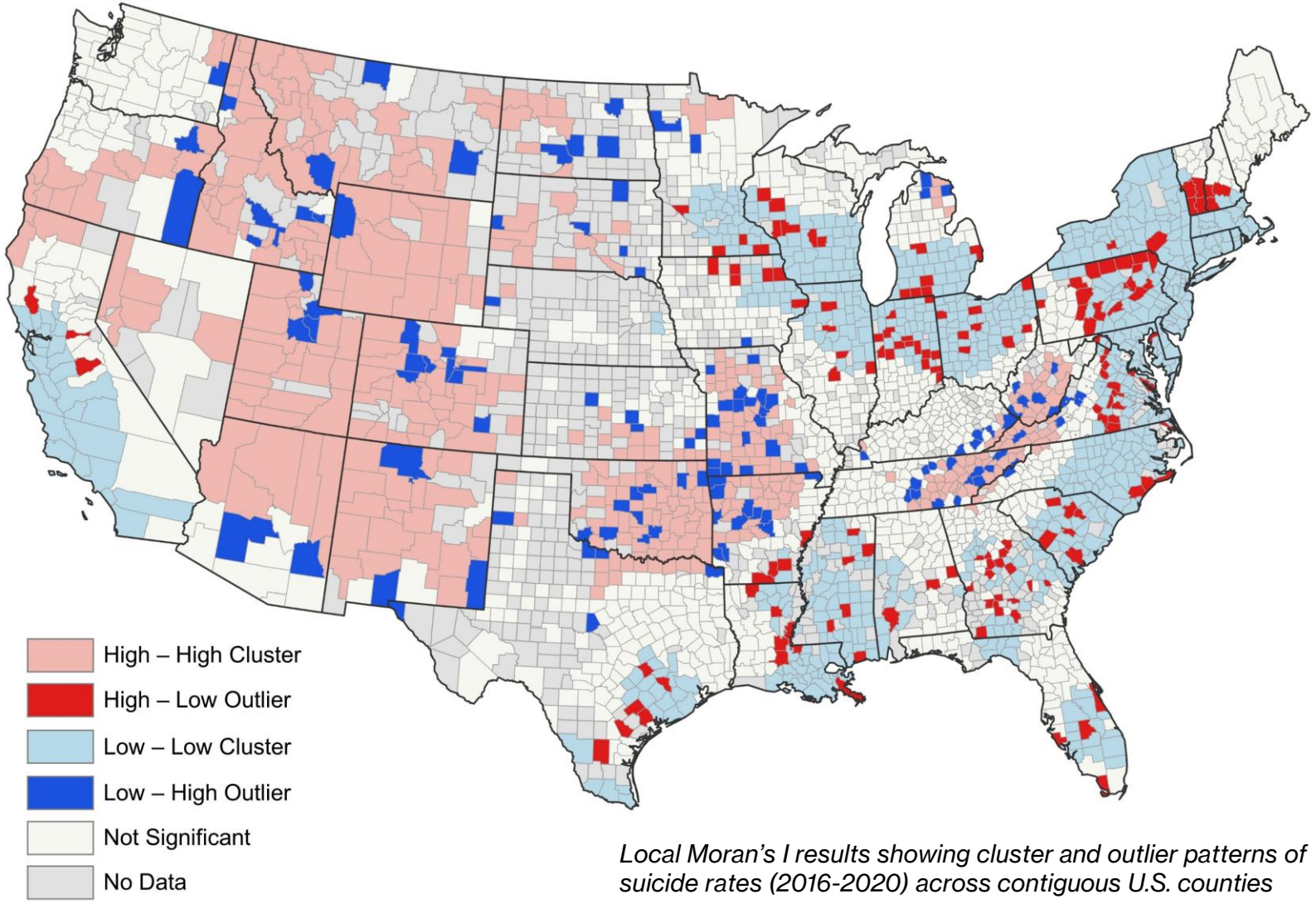
Low – Low cluster:
Low values surrounded by low values



Local Moran's I results showing cluster and outlier patterns of suicide rates (2016-2020) across contiguous U.S. counties

Visualizing Local Moran's I

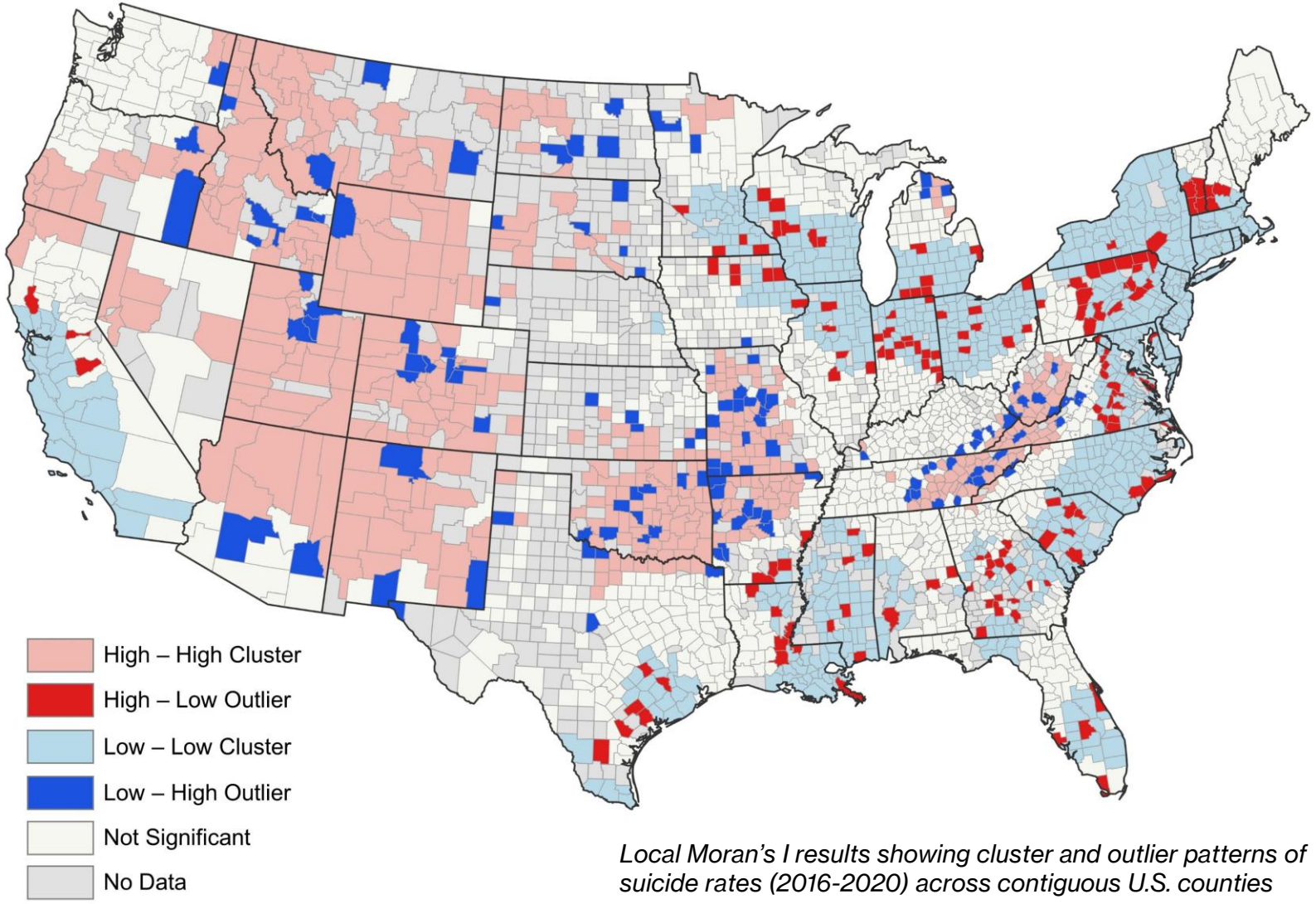
High – Low outlier:
High value surrounded by low values



Local Moran's I results showing cluster and outlier patterns of suicide rates (2016-2020) across contiguous U.S. counties

Visualizing Local Moran's I

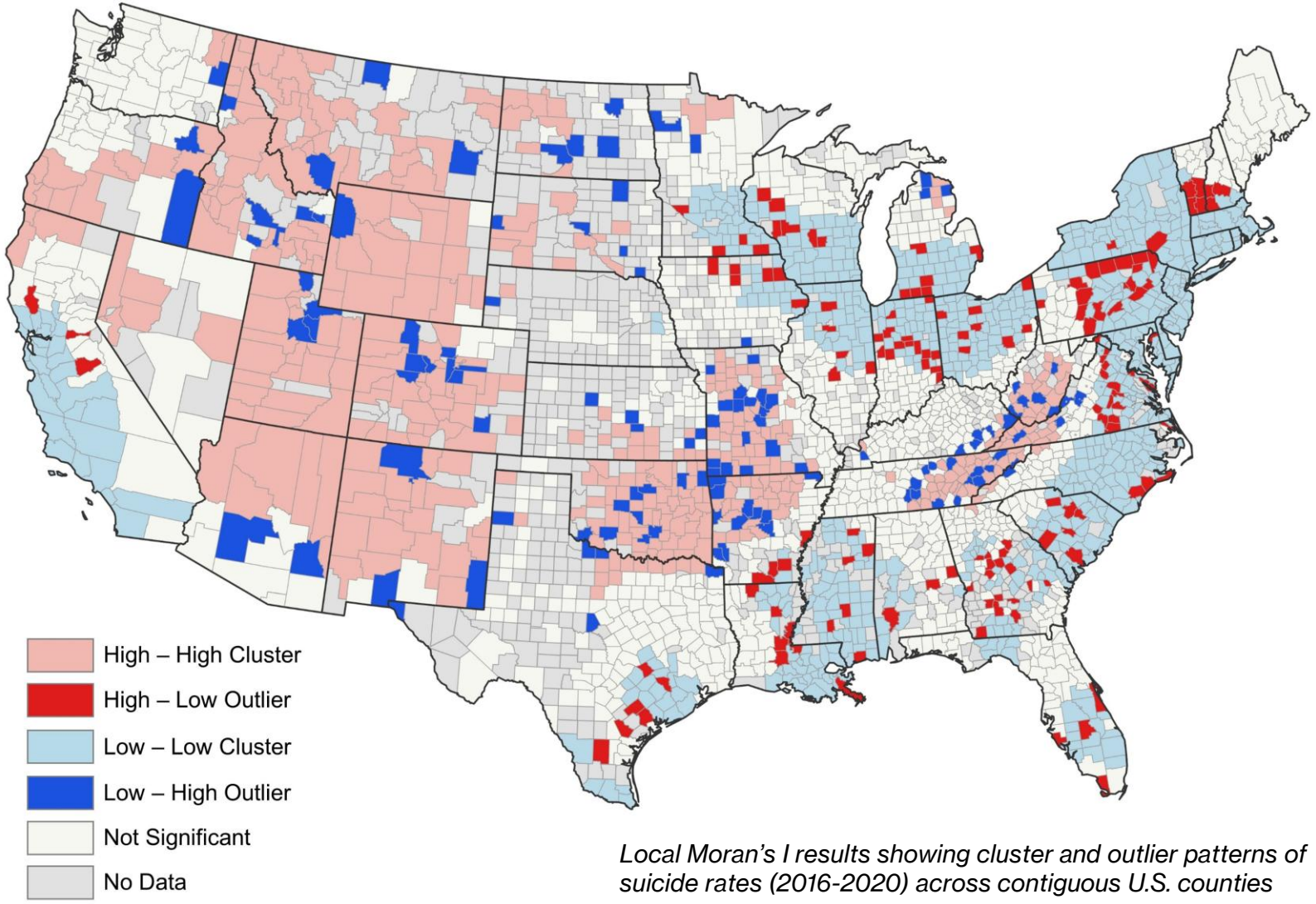
Low – High outlier:
Low value surrounded by high values



Local Moran's I results showing cluster and outlier patterns of suicide rates (2016-2020) across contiguous U.S. counties

Visualizing Local Moran's I

Not Significant:
Not considered a statistically significant cluster or outlier



Local Moran's I results showing cluster and outlier patterns of suicide rates (2016-2020) across contiguous U.S. counties

Interpreting Local Moran's I

Column name	Meaning
Ii	The **Local Moran's I statistic** for each county → how similar that county's value is to its neighbors • Positive = similar values (clustering) • Negative = dissimilar values (dispersion)
E.Ii	Expected value of ***Ii*** under spatial randomness
Var.Ii	Variance of ***Ii***
Z.Ii	Standardized z-score for ***Ii*** (used for significance testing)
P.Ii	p-value for the test of local spatial autocorrelation ($p < 0.05$ = significant)

A subset of Local Moran's I results for patterns of suicide rates (2016-2020) across Delaware, USA counties

suicide_z	Ii	E.Ii	Var.Ii	Z.Ii	P.Ii	LISA_cluster
-0.9351838	0.7190032	-0.0003644036	0.1747040	1.721074	0.08523749	Not significant
-1.3494274	1.1322124	-0.0007587308	0.2595051	2.224057	0.02614462	Low-Low
-1.0312252	0.9923223	-0.0004430938	0.2124133	2.154050	0.03123626	Low-Low

Interpreting Local Moran's I

1.Look at p-values

- **p < 0.05** = statistically significant local spatial autocorrelation
- p > 0.05 = no evidence of significant clustering or outlier behavior for that county

2.Check the sign of Ii

- **Positive Ii**: County has a similar value to its neighbors (part of a cluster)
- **Negative Ii**: County has a very different value from its neighbors (outlier)

3. Find the county's own value *

- High value: above the mean of the dataset
- Low value: below the mean of the dataset

4. Determine cluster or outlier type

- **High value + positive Ii** → High–High cluster
- **Low value + positive Ii** → Low–Low cluster
- **High value + negative Ii** → High–Low outlier
- **Low value + negative Ii** → Low–High outlier

A subset of Local Moran's I results for patterns of suicide rates (2016-2020) across Delaware, USA counties

suicide_z	Ii	E.Ii	Var.Ii	Z.Ii	P.Ii	LISA_cluster
All	All	All	All	All	All	All
-0.9351838	0.7190032	-0.0003644036	0.1747040	1.721074	0.08523749	Not significant
-1.3494274	1.1322124	-0.0007587308	0.2595051	2.224057	0.02614462	Low-Low
-1.0312252	0.9923223	-0.0004430938	0.2124133	2.154050	0.03123626	Low-Low

Is this a high or low value? Is it above or below the mean?

Is this a positive or negative value?

Is this greater or less than 0.05?

What cluster/outlier type is it?

* Remember, if the variable is standardized, then above the mean will be greater than 0 (positive value), while below the mean will be less than 0 (negative value)

Interpreting Local Moran's I: Practice!

Type: **Ii =** 1.45 **p-value =** 0.04 **Standardized County Suicide Rate =** -1.9	**Type:** **Ii =** 0.20 **p-value =** 0.55 **Standardized County Suicide Rate =** 0.21	**Type:** **Ii =** 1.82 **p-value =** 0.01 **Standardized County Suicide Rate =** 2.1
Type: **Ii = -**1.69 **p-value =** 0.03 **Standardized County Suicide Rate =** 3.4	**Type:** **Ii =** 1.01 **p-value =** 0.06 **Standardized County Suicide Rate =** 0.99	**Type:** **Ii =** -1.35 **p-value =** 0.00 **Standardized County Suicide Rate =** -1.7

Interpreting Local Moran's I: Practice!

Type: Low-Low **Ii =** 1.45 **p-value =** 0.04 **Standardized County Suicide Rate =** -1.9	**Type:** Not Significant **Ii =** 0.20 **p-value =** 0.55 **Standardized County Suicide Rate =** 0.21	**Type:** High-High **Ii =** 1.82 **p-value =** 0.01 **Standardized County Suicide Rate =** 2.1
Type: High-Low **Ii =** -1.69 **p-value =** 0.03 **Standardized County Suicide Rate =** 3.4	**Type:** Not Significant **Ii =** 1.01 **p-value =** 0.06 **Standardized County Suicide Rate =** 0.99	**Type:** Low-High **Ii =** -1.35 **p-value =** 0.00 **Standardized County Suicide Rate =** -1.7