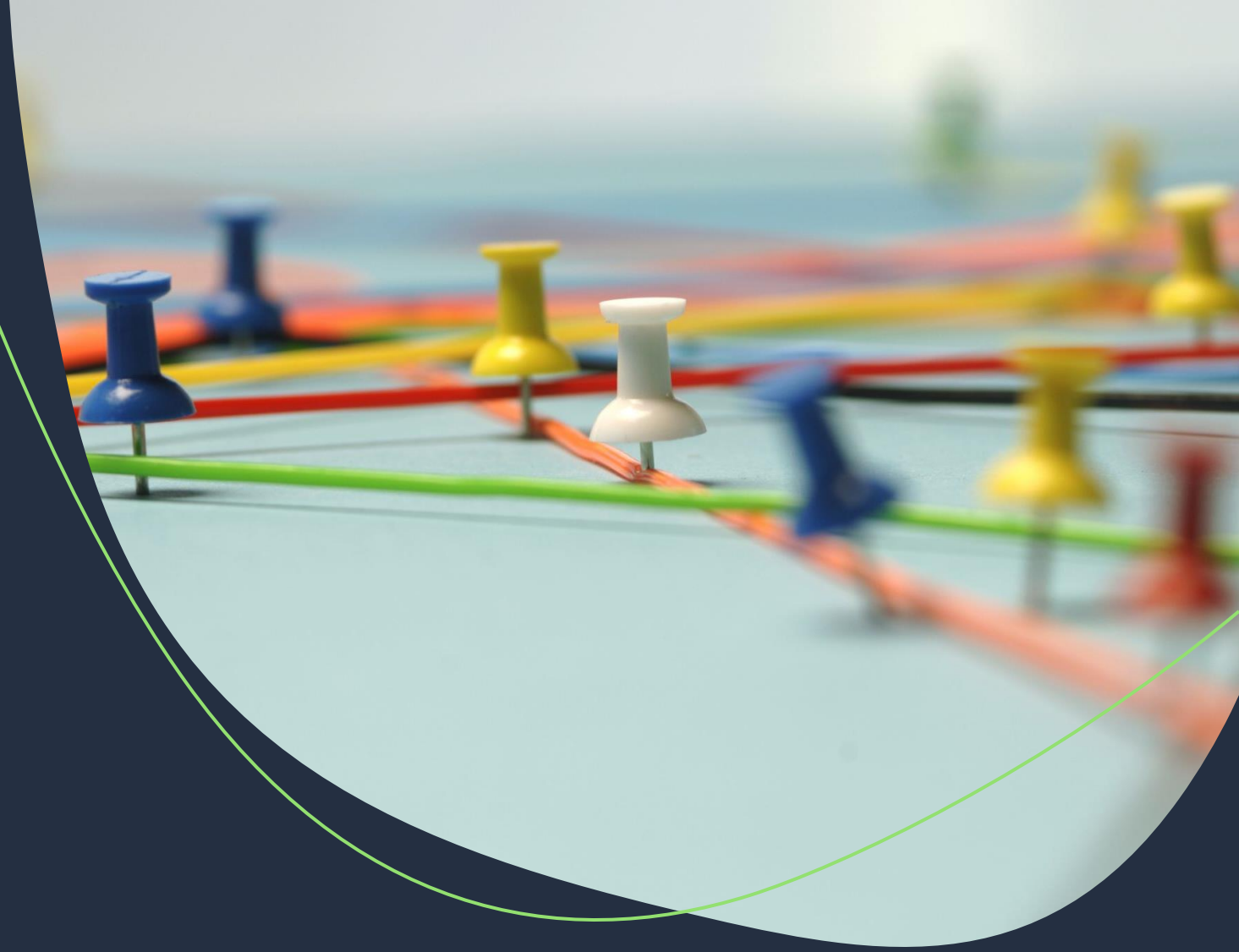
Mapeo de las violencias y los delitos

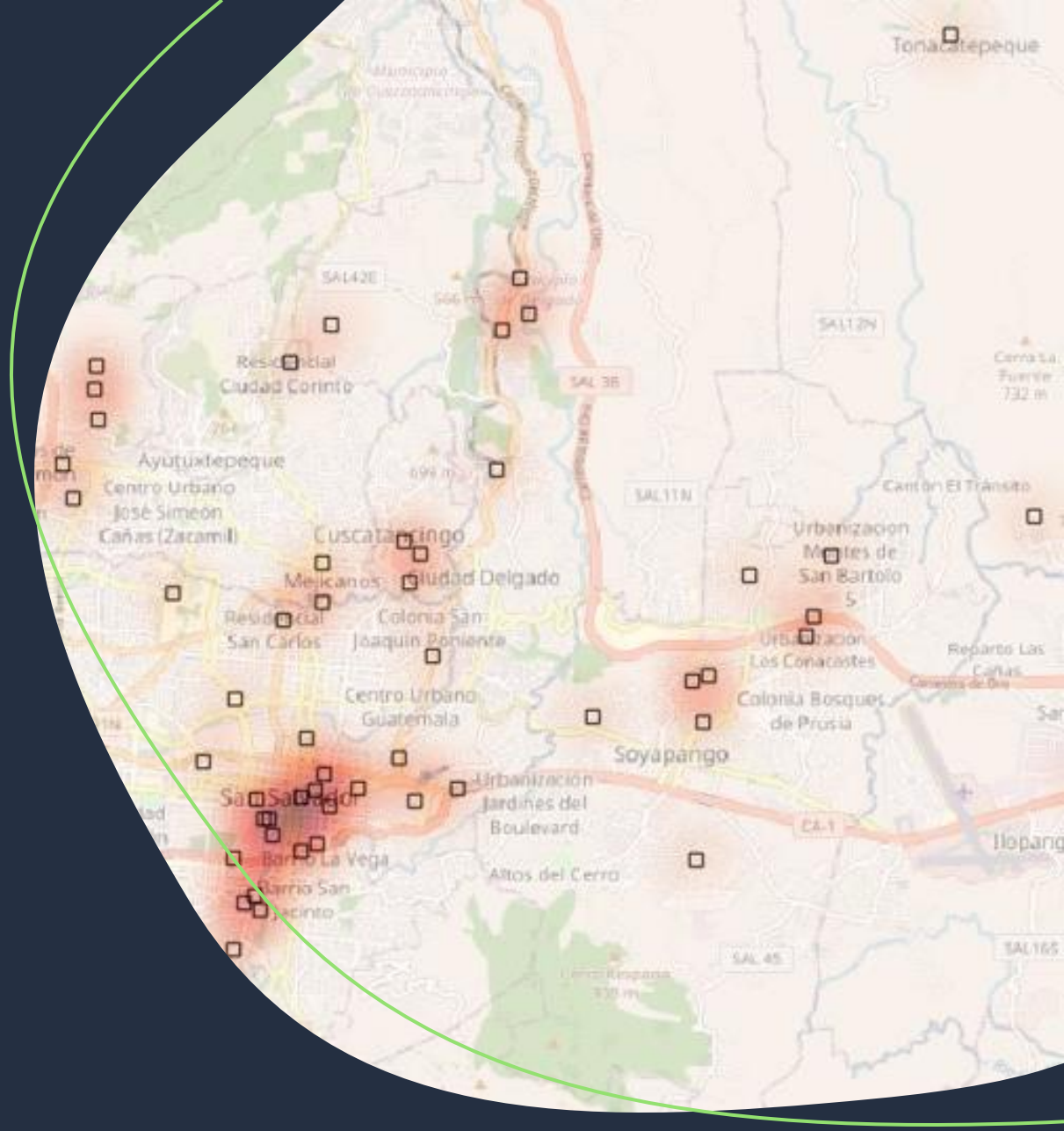


Clase abierta - Maestría en criminología y políticas de prevención

Metzi Aguilar

Temario

- Mapas de coropletas
- Estadística espacial
  - Inspección visual
  - Índices globales
  - Índices locales



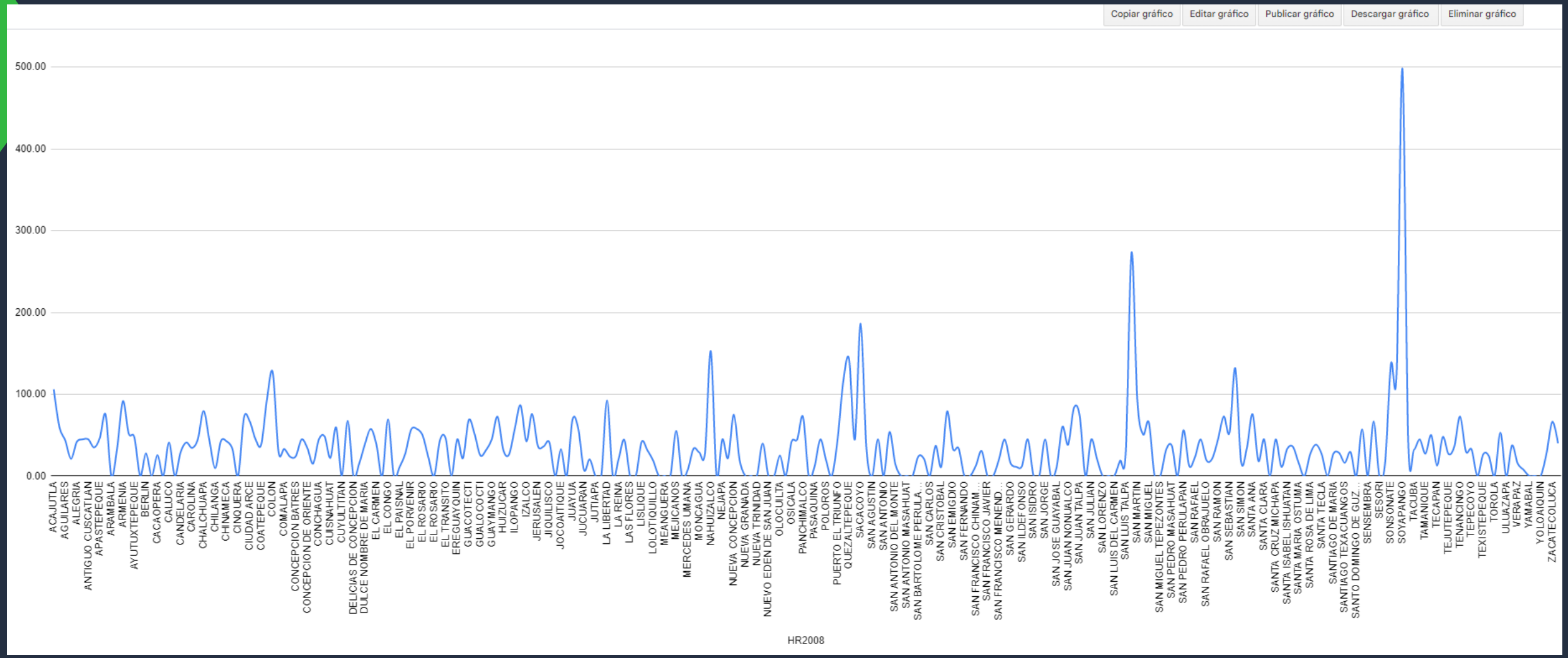
Base de datos de crimen (AAS – JJI 2008-2018)

B	D	G	CR
NOM_DPTO	NOM_MUN	COD_MUN4	HR2008
SONSONATE	ACAJUTLA	0301	105.99
CHALATENANG	AGUA CALIENTE	0401	59.40
SAN SALVADOF	AGUILARES	0601	43.52
AHUACHAPAN	AHUACHAPAN	0101	20.85
USULUTAN	ALEGRIA	1101	41.77
LA UNION	ANAMOROS	1401	44.80
LA LIBERTAD	ANTIGUO CUSCATLAN	0501	44.82
AHUACHAPAN	APANECA	0102	34.89
SAN VICENTE	APASTEPEQUE	1001	46.54
SAN SALVADOF	APOPA	0602	75.01
MORAZAN	ARAMBALA	1301	0.00
CHALATENANG	ARCATAO	0402	33.99
SONSONATE	ARMENIA	0302	91.49
AHUACHAPAN	ATIQUIZAYA	0103	52.02
SAN SALVADOF	AYUTUXTEPEQUE	0603	47.06
CHALATENANG	AZACUALPA	0403	0.00
USULUTAN	BERLIN	1102	27.87
LA UNION	BOLIVAR	1402	0.00
MORAZAN	CACAOPERA	1302	25.71
USULUTAN	CALIFORNIA	1103	0.00
SONSONATE	CALUCO	0303	40.98
CHALATENANG	CANCASQUE	0427	0.00
CUSCATLAN	CANDELARIA	0701	28.81
SANTA ANA	CANDELARIA DE LA FRONTER	0201	41.06
SAN MIGUEL	CAROLINA	1201	34.27
CHALATENANG	CHALATENANGO	0407	44.77

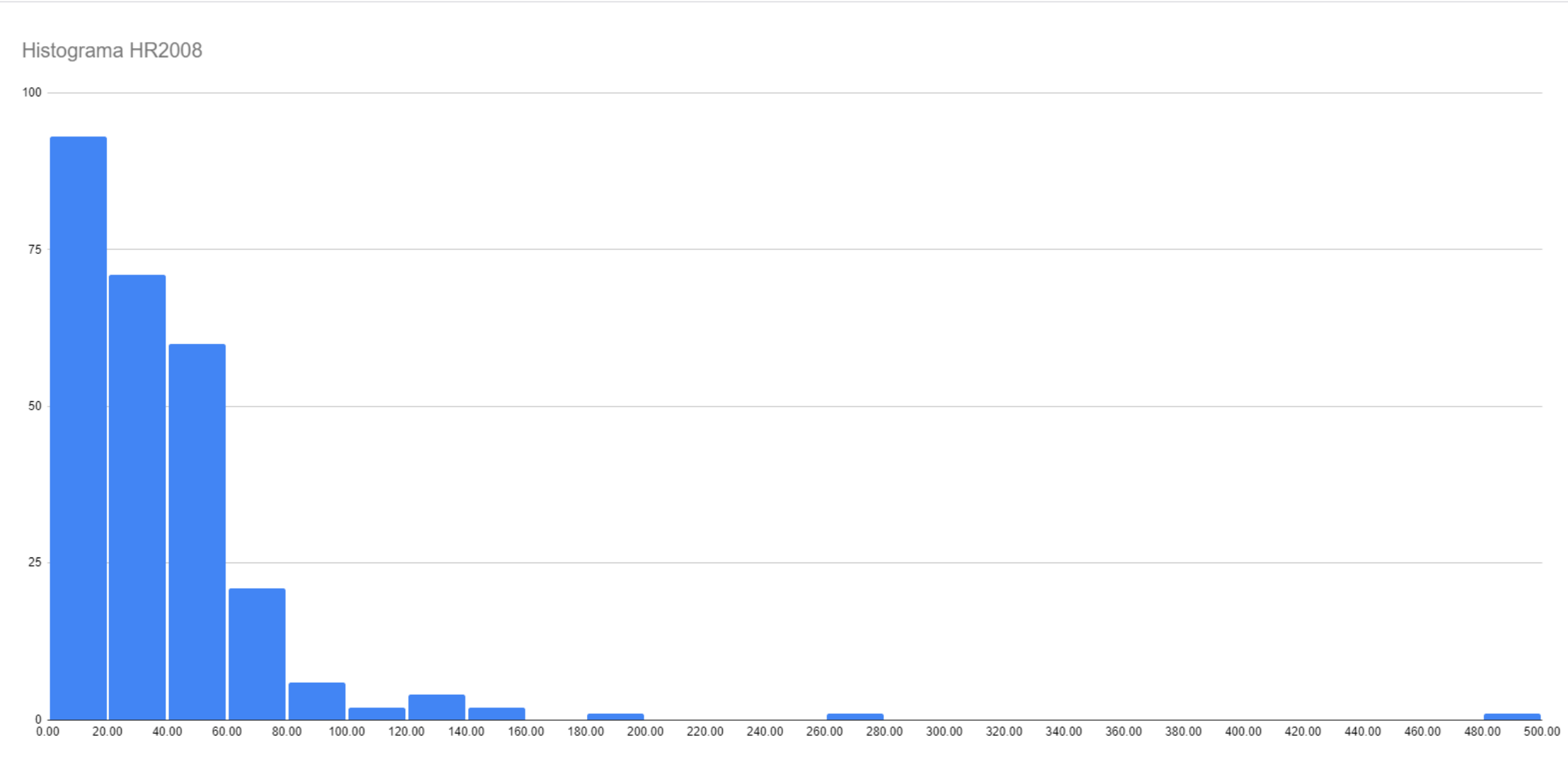
B	D	G	CR
NOM_DPTO	NOM_MUN	COD_MUN4	HR2008
SAN SALVADOF	SOYAPANGO	0617	497.78
SAN SALVADOF	SAN MARCOS	0612	273.11
LA LIBERTAD	SACACOYO	0513	186.08
SONSONATE	NAHUIZALCO	0308	152.62
LA LIBERTAD	QUEZALTEPEQUE	0512	142.77
SONSONATE	SONSONATE	0315	136.54
SANTA ANA	SAN SEBASTIAN SALITRILLO	0209	131.73
SONSONATE	SONZACATE	0316	127.37
LA LIBERTAD	COLON	0503	126.57
SAN MIGUEL	QUELEPA	1212	116.44
SONSONATE	ACAJUTLA	0301	105.99
CUSCATLAN	COJUTEPEQUE	0702	94.06
SAN SALVADOF	SAN MARTIN	0613	93.56
LA LIBERTAD	LA LIBERTAD	0509	92.51
SONSONATE	ARMENIA	0302	91.49
LA UNION	INTIPUCA	1407	86.45
LA LIBERTAD	SAN JUAN OPICO	0515	83.08
SANTA ANA	CHALCHUAPA	0203	79.51
USULUTAN	SAN DIONISIO	1117	78.90
LA LIBERTAD	JAYAQUE	0507	75.83
SANTA ANA	SANTA ANA	0210	75.59
CHALATENANG	NUEVA CONCEPCION	0416	75.07
SAN SALVADOF	APOPA	0602	75.01
LA PAZ	SAN JUAN TALPA	0811	73.83
SAN SALVADOF	SAN SALVADOR	0614	72.86
CUSCATLAN	TENANCINGO	0716	72.66

Estadísticas de columna

Valores únicos	200
Suma	9,408.40
Media	35.91
Mediana	31.01
Valor mínimo	0
Valor máximo	497.78


Copiar gráfico
Editar gráfico
Publicar gráfico
Descargar gráfico
Eliminar gráfico
500.00
400.00
300.00
200.00
100.00
0.00
ACAJUTLA
AGUILARES
ALEGRIA
ANTIGUO CUSCATLAN
APASTEPEQUE
ARAMBALA
ARMENIA
AYUTUXTEPEQUE
BERLIN
CACAOPERA
CALUCO
CANDELARIA
CAROLINA
CHALCHUAPA
CHILANGA
CHINAMECA
CINQUERA
CIUDAD ARCE
COATEPEQUE
COLON
COMALAPA
CONCEPCION BATRES
CONCEPCION DE ORIENTE
CONCHAGUA
CUISNAHUAT
CUYULTITAN
DELICIAS DE CONCEPCION
DULCE NOMBRE DE MARIA
EL CARMEN
EL CONGO
EL PAISNAL
EL PORVENIR
EL ROSARIO
EL ROSARIO
EL TRANSITO
EREGUAYQUIN
GUACOTECTI
GUALOCOCTI
GUAYMANGO
HUIZUCAR
ILOPANGO
IZALCO
JERUSALEN
JIQUILISCO
JOCOAITIQUE
JUAYUA
JUCUARAN
JUTIAPA
LA LIBERTAD
LA REINA
LAS FLORES
LISLIQUE
LOLOTIQUILLO
MEANGUERA
MEJICANOS
MERCEDES UMANA
MONCAGUA
NAHUIZALCO
NEJAPA
NUEVA CONCEPCION
NUEVA GRANADA
NUEVA TRINIDAD
NUEVO EDEN DE SAN JUAN
OLOCUILTA
OSICALA
PANCHIMALCO
PASAQUINA
POLOROS
PUERTO EL TRIUNFO
QUEZALTEPEQUE
SACACOYO
SAN AGUSTIN
SAN ANTONIO
SAN ANTONIO DEL MONTE
SAN ANTONIO MASAHUAT
SAN BARTOLOME PERULA...
SAN CARLOS
SAN CRISTOBAL
SAN EMIGDIO
SAN FERNANDO
SAN FRANCISCO CHINAM...
SAN FRANCISCO JAVIER
SAN FRANCISCO MENEND...
SAN GERARDO
SAN ILDEFONSO
SAN ISIDRO
SAN JORGE
SAN JOSE GUAYABAL
SAN JUAN NONUALCO
SAN JUAN TALPA
SAN JULIAN
SAN LORENZO
SAN LUIS DEL CARMEN
SAN LUIS TALPA
SAN MARTIN
SAN MIGUEL
SAN MIGUEL TEPEZONTES
SAN PEDRO MASAHUAT
SAN PEDRO PERULAPAN
SAN RAFAEL
SAN RAFAEL OBRAJUELO
SAN RAMON
SAN SEBASTIAN
SAN SIMON
SANTA ANA
SANTA CLARA
SANTA CRUZ MICHAPA
SANTA ISABEL ISHUATAN
SANTA MARIA OSTUMA
SANTA ROSA DE LIMA
SANTA TECLA
SANTIAGO DE MARIA
SANTIAGO TEXACUANGOS
SANTO DOMINGO DE GUZ...
SENSEMBRA
SESORI
SONSONATE
SOYAPANGO
TACUBA
TAMANIQUE
TECAPAN
TEJUTEPEQUE
TENANCINGO
TEPECOYO
TEXISTEPEQUE
TOROLA
ULUAZAPA
VERAPAZ
YAMABAL
YOLOAIQUIN
ZACATECOLUCA
HR2008

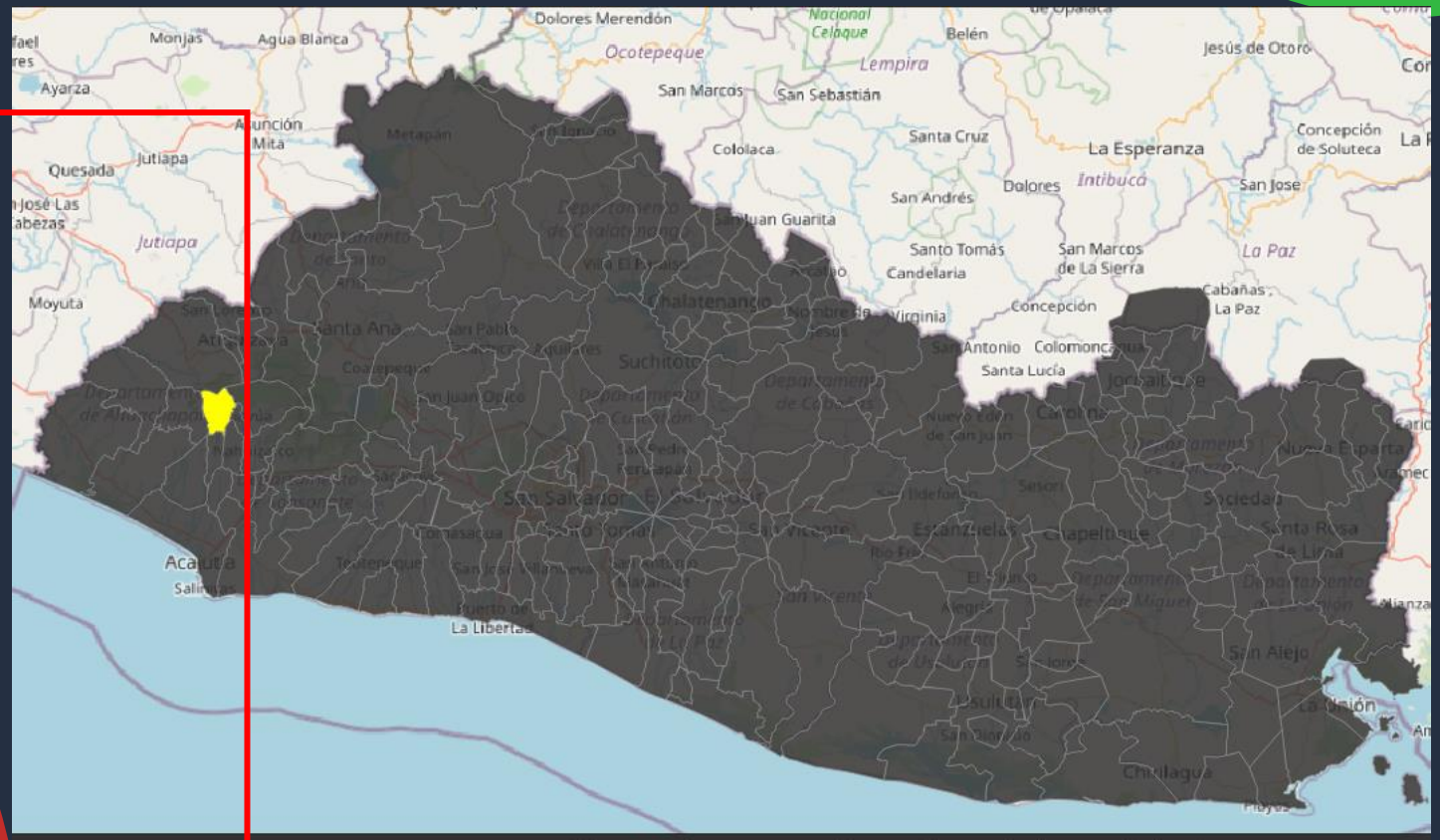
Histograma


Histograma HR2008
100
75
50
25
0
0.00
20.00
40.00
60.00
80.00
100.00
120.00
140.00
160.00
180.00
200.00
220.00
240.00
260.00
280.00
300.00
320.00
340.00
360.00
380.00
400.00
420.00
440.00
460.00
480.00
500.00

¿Con esta base de datos podemos identificar patrones del crimen distribuidos espacialmente?

Unión de datos con la geometría

B	D	G	CR
NOM_DPTO	NOM_MUN	COD_MUN4	HR2008
AHUACHAPAN	AHUACHAPAN	0101	20.85
AHUACHAPAN	APANECA	0102	34.89
AHUACHAPAN	ATIQUIZAYA	0103	52.02
AHUACHAPAN	CONCEPCION DE ATACO	0104	44.71
AHUACHAPAN	EL REFUGIO	0105	57.58
AHUACHAPAN	GUAYMANGO	0106	44.75
AHUACHAPAN	JUJUTLA	0107	20.26
AHUACHAPAN	SAN FRANCISCO MENENDEZ	0108	20.34
AHUACHAPAN	SAN LORENZO	0109	20.95
AHUACHAPAN	SAN PEDRO PUXTLA	0110	12.44
AHUACHAPAN	TACUBA	0111	32.25
AHUACHAPAN	TURIN	0112	53.07
SANTA ANA	CANDELARIA DE LA FRONTER	0201	41.06
SANTA ANA	COATEPEQUE	0202	38.43
SANTA ANA	CHALCHUAPA	0203	79.51
SANTA ANA	EL CONGO	0204	69.29
SANTA ANA	EL PORVENIR	0205	56.60
SANTA ANA	MASAHUAT	0206	0.00
SANTA ANA	METAPAN	0207	33.39
SANTA ANA	SAN ANTONIO PAJONAL	0208	0.00
SANTA ANA	SAN SEBASTIAN SALITRILLO	0209	131.73
SANTA ANA	SANTA ANA	0210	75.59
SANTA ANA	SANTA ROSA GUACHIPILIN	0211	38.42
SANTA ANA	SANTIAGO DE LA FRONTERA	0212	0.00
SANTA ANA	TEXISTEPEQUE	0213	26.36
SONSONATE	ACAJUTLA	0301	105.99



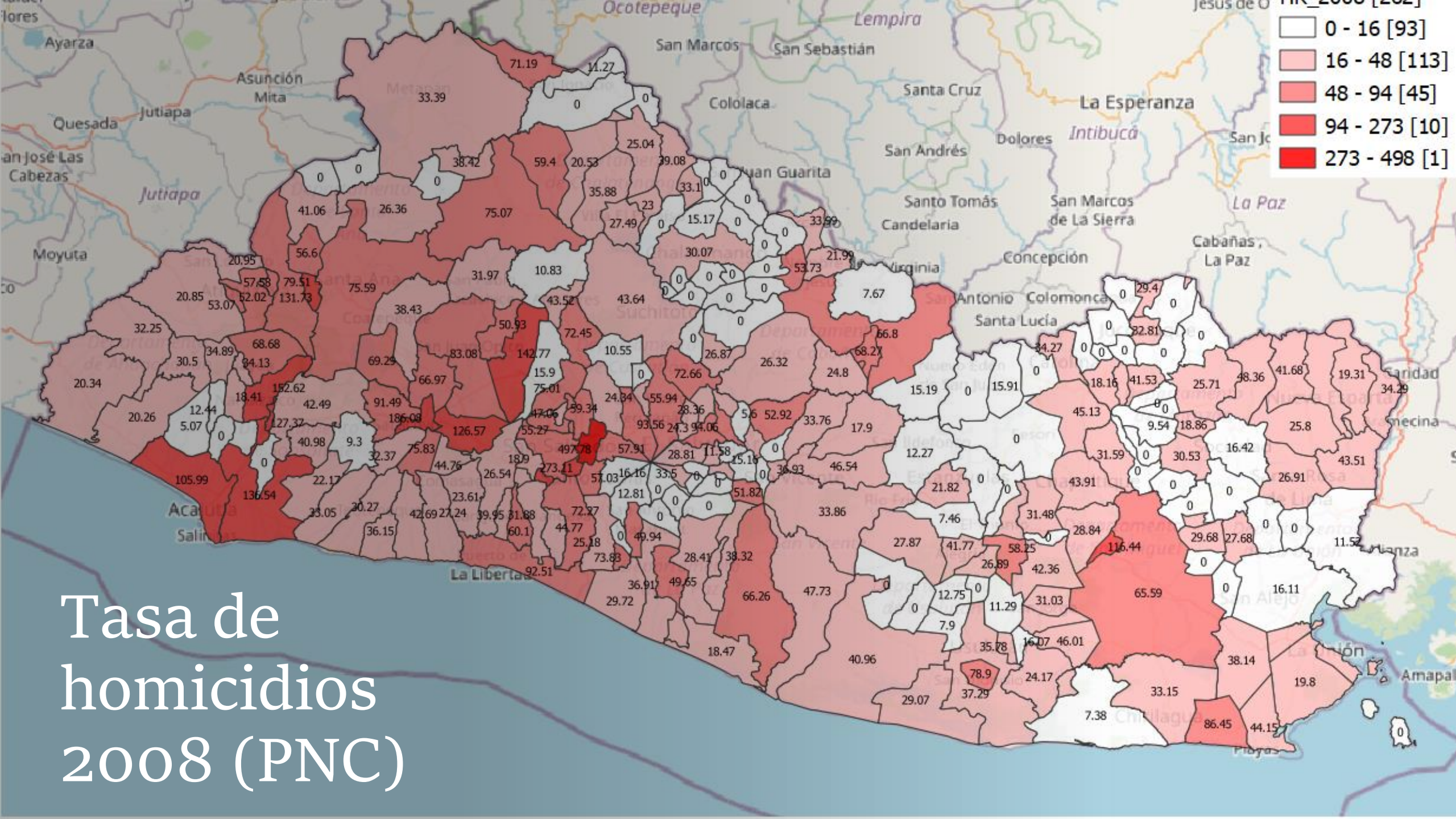
municipios_proyeccionesDIGESTYC— Objetos Totales: 271, Filtrados: 271, Seleccionados: 1

	fid	NOM_DPTO	COD_DPTO	NOM_MUN	COD_MUN	COD_MUN4	TASA_POBR	POBR_FISDL	PERIM_KM	AREA_KM2
10	237	AHUACHAPAN	01	AHUACHAPAN	01	0101	47.7	Pobreza Mod...	83.7769501109	251.1917483
11	230	AHUACHAPAN	01	APANECA	02	0102	52	Pobreza Alta	32.4664330927	40.6770800884
12	236	AHUACHAPAN	01	ATIQUIZAYA	03	0103	41.2	Pobreza Mod...	54.9150923877	65.1695533019
13	231	AHUACHAPAN	01	CONCEPCIO...	04	0104	48.7	Pobreza Mod...	47.5784226153	68.9204660813

Datos extendidos con geometría

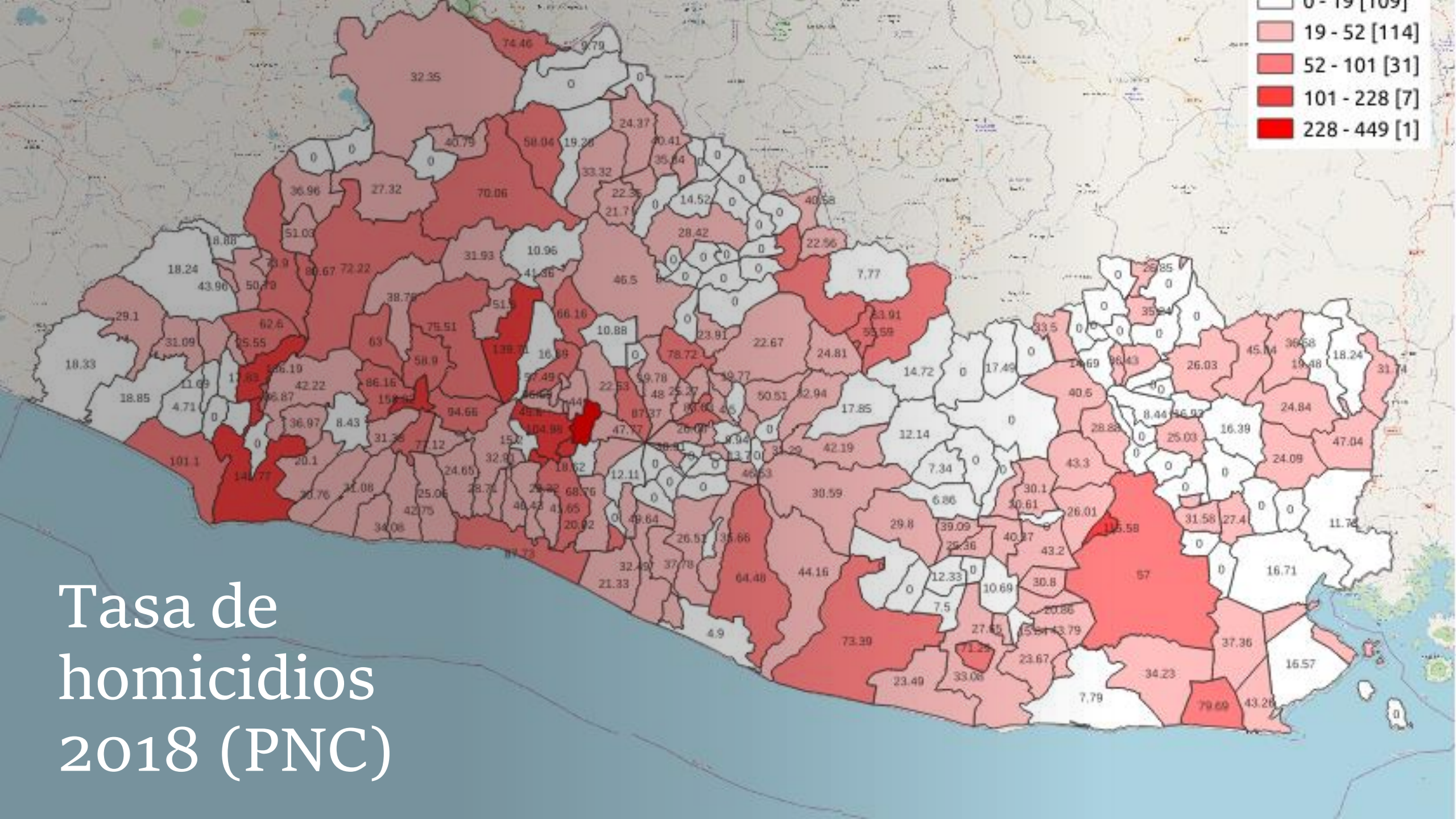
Homicide rate 2008— Objetos Totales: 273, Filtrados: 273, Seleccionados: 0

09	_dig_POB_2010	_dig_POB_2011	_dig_POB_2012	_dig_POB_2013	_dig_POB_2014	_dig_POB_2015	_dig_POB_2016	_dig_POB_2017	_dig_POB_2018	HR2008
58	53252	53881	54531	55197	55870	56551	57231	57915	58588	94.06
78	69202	69664	70154	70662	71182	71705	72227	72744	73249	93.56
10	39329	39557	39790	40022	40251	40472	40681	40867	41034	92.51
46	37551	37772	38007	38251	38498	38744	38989	39229	39463	91.49
93	8114	8156	8215	8287	8371	8463	8564	8671	8784	86.45
08	82218	83051	83899	84746	85592	86425	87238	88002	88729	83.08



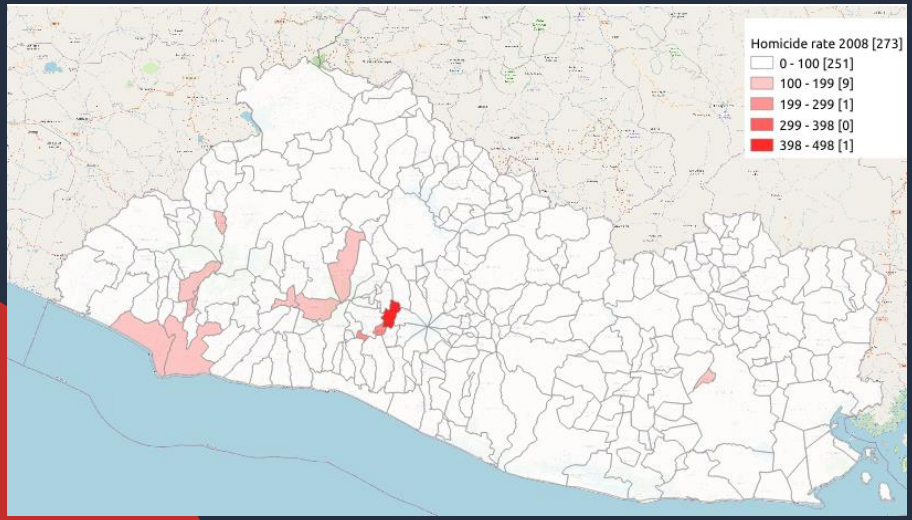
Tasa de homicidios 2008 (PNC)

#obs=271 #selected=271

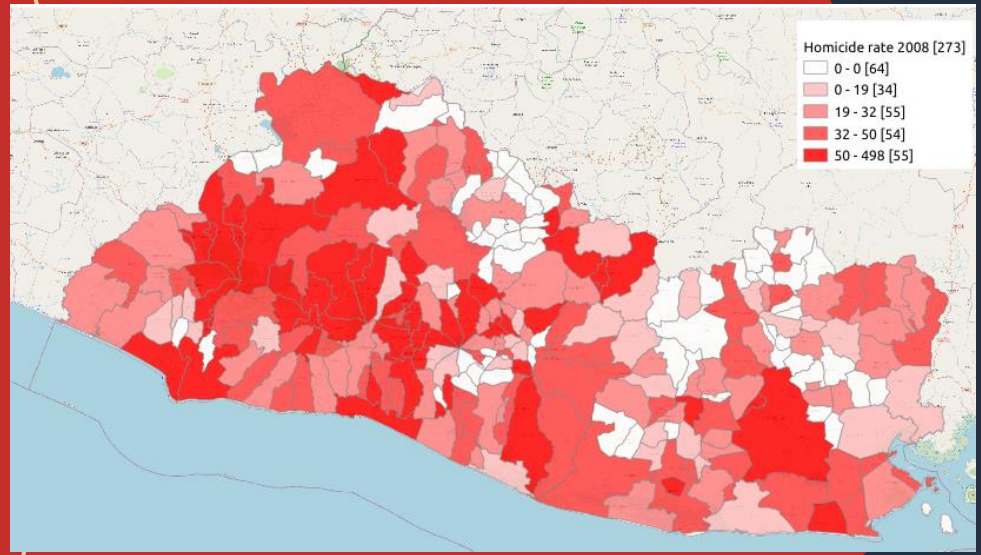


Tasa de homicidios 2018 (PNC)

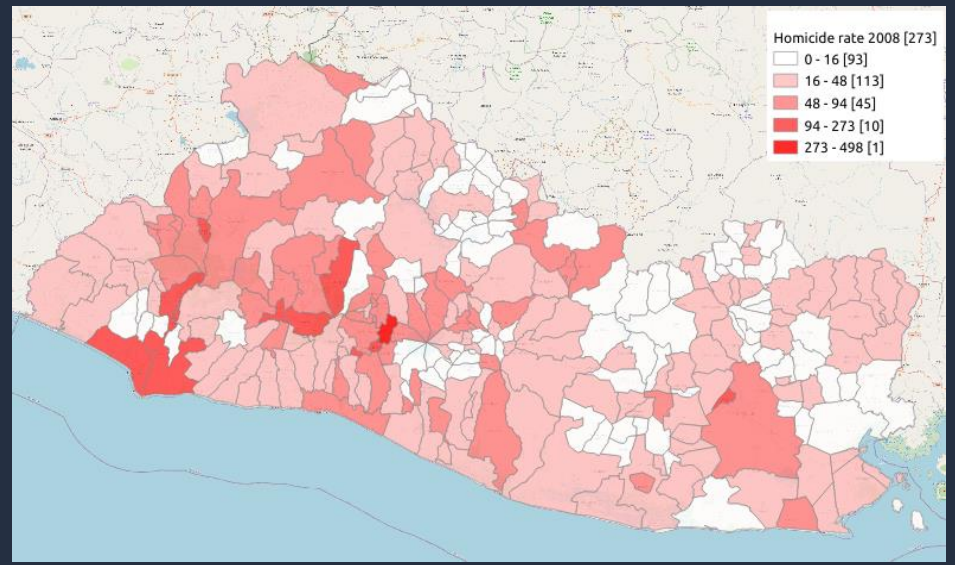
Mapas de coropletas
¿Cómo clasificar?



Intervalos iguales

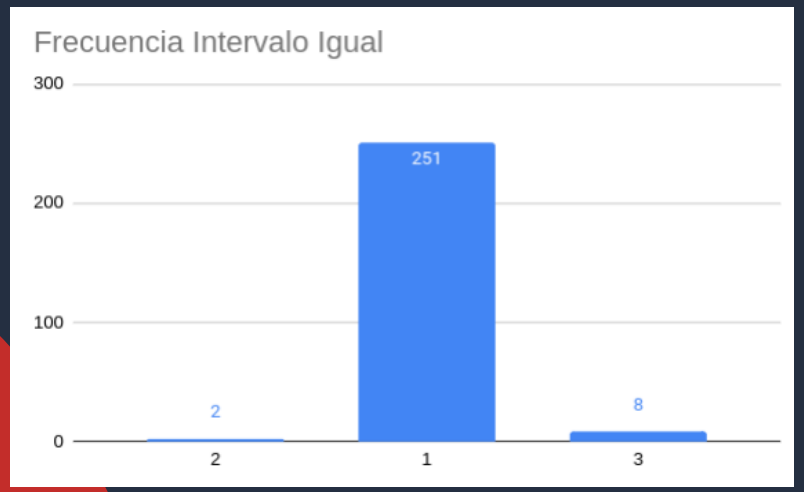


Quintiles

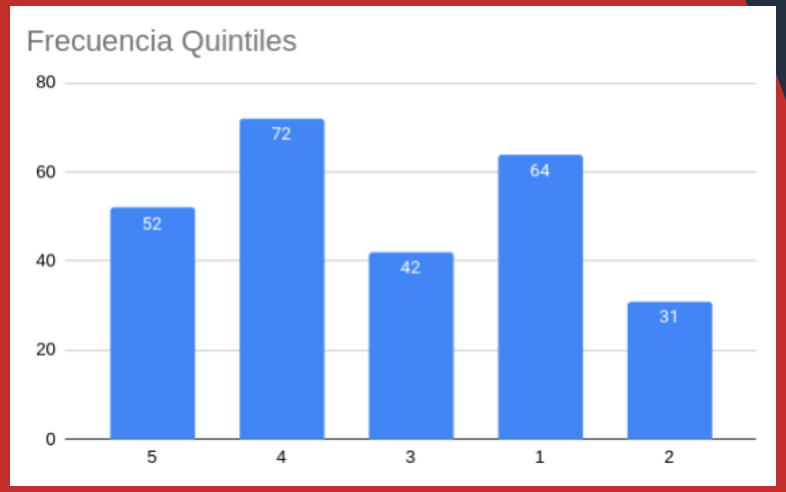


Jenks

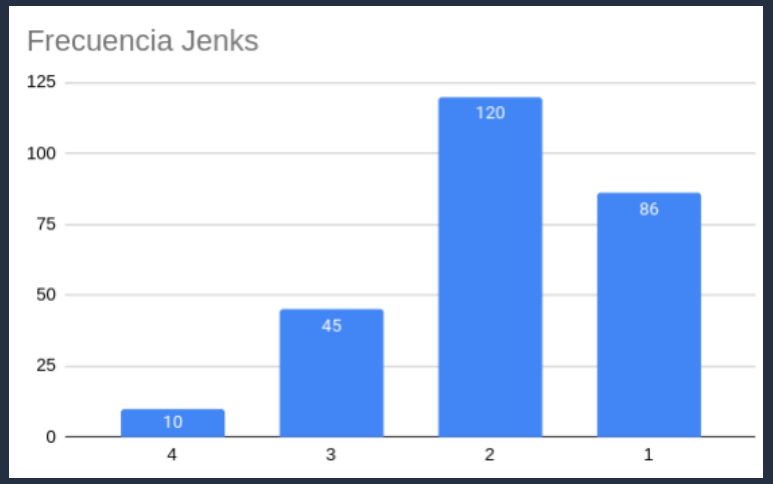
Mapas de coropletas
¿Cómo clasificar?



Intervalos iguales



Quintiles



Jenks

Mapas de coropletas
¿cómo clasificar? Criterio MEDA (ACDM)

2. La mediana de las desviaciones absolutas (MEDA)

La misma idea que nos llevó a definir la varianza como medida de dispersión a partir de la media, nos permite ahora introducir la **mediana de las desviaciones absolutas (MEDA)** para expresar la variabilidad de las observaciones alrededor de la mediana. Una vez construidas las desviaciones de los datos respecto de la mediana,

$$x_1 - med_x,\ x_2 - med_x, \ldots,\ x_{N-1} - med_x,\ x_N - med_x,$$

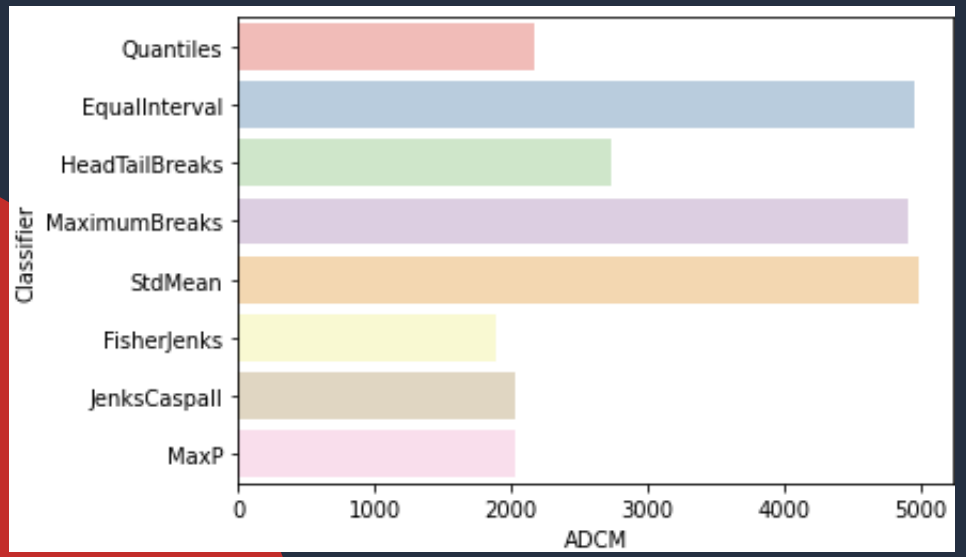
se toman valores absolutos para eliminar los signos,

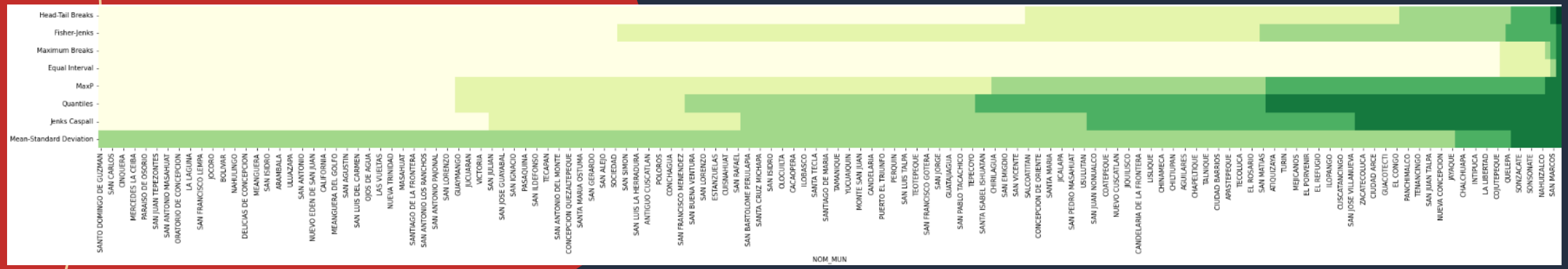
$$|x_1 - med_x|,\ |x_2 - med_x|, \ldots,\ |x_{N-1} - med_x|,\ |x_N - med_x|$$

y la mediana de estos valores es la MEDA, es decir,

$$MEDA_x = med(|x_1 - med_x|,\ |x_2 - med_x|, \ldots,\ |x_{N-1} - med_x|,\ |x_N - med_x|).$$

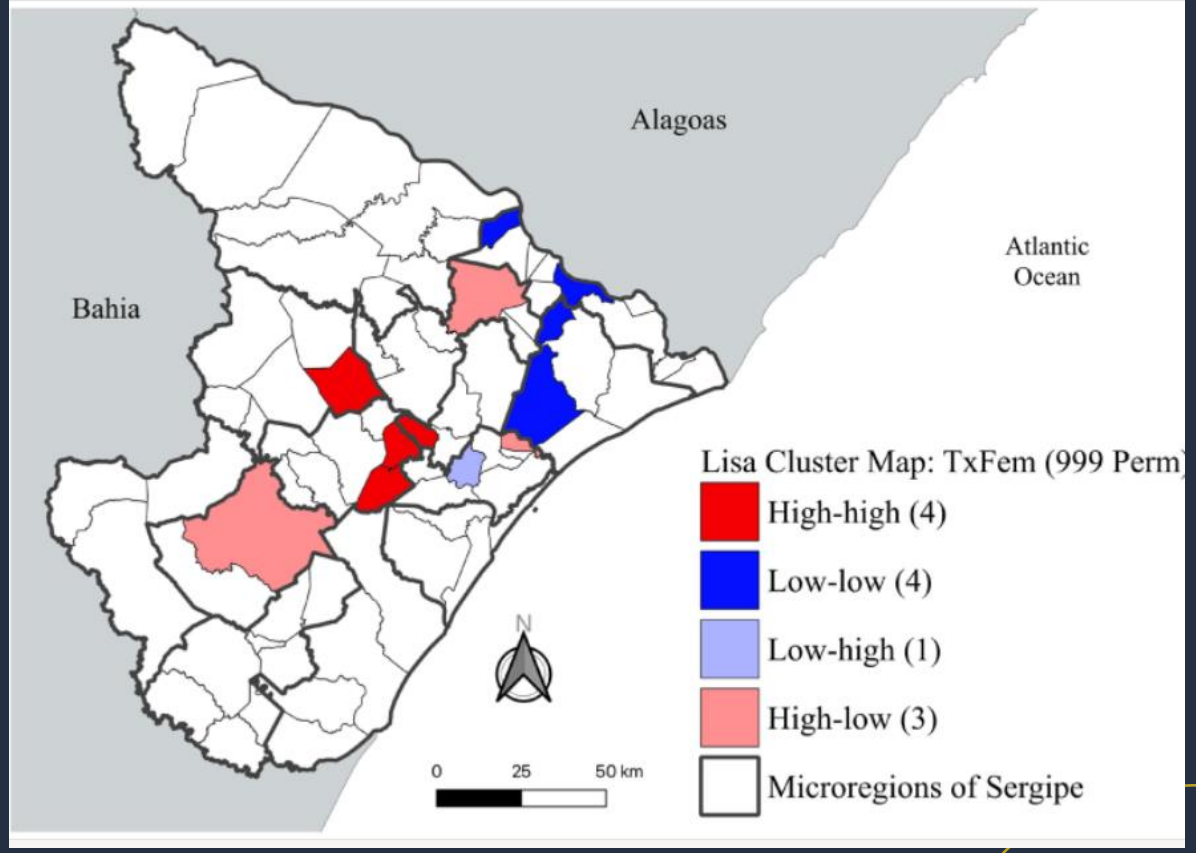
Mapas de coropletas
¿Cómo clasificar?

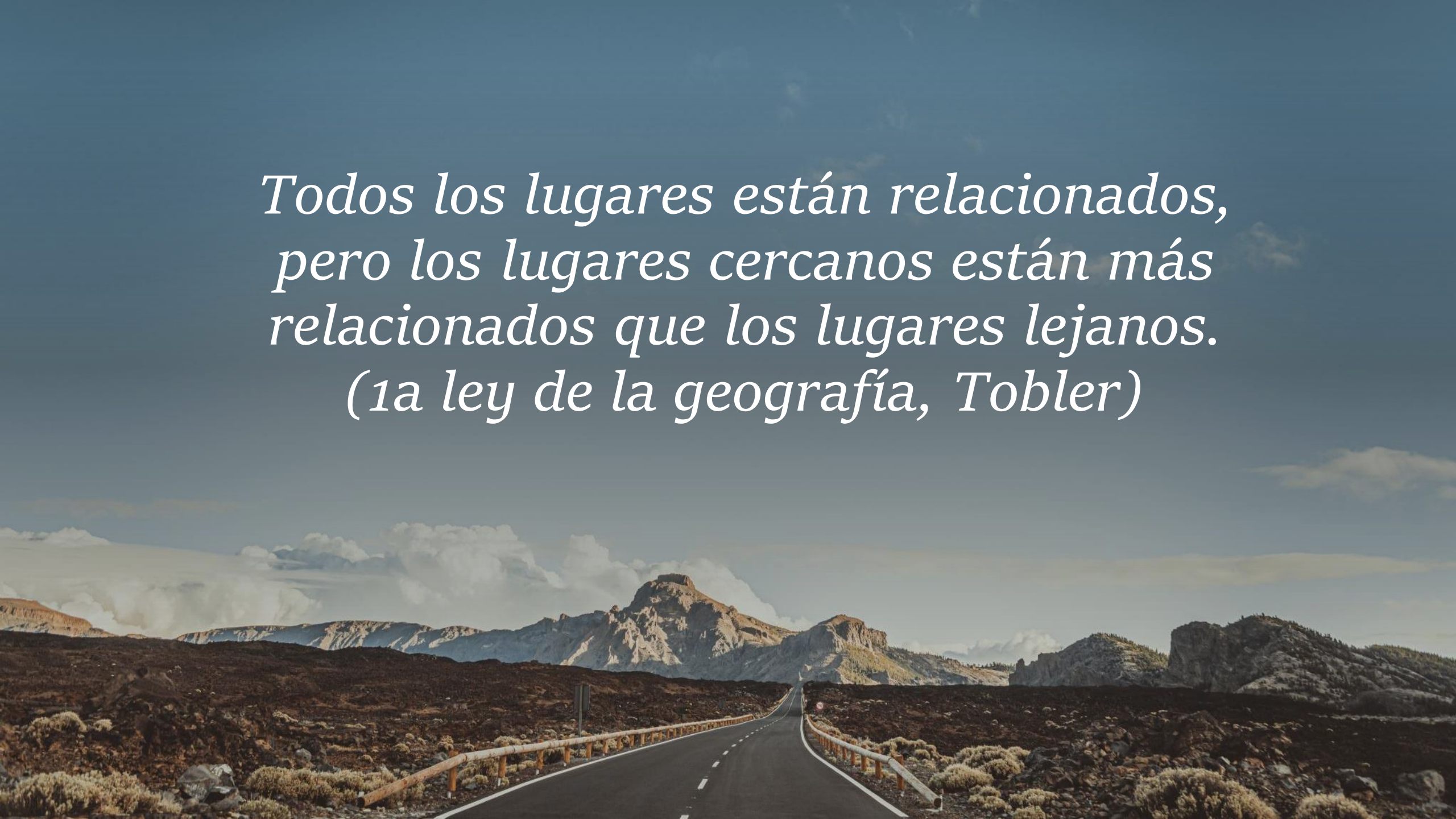




Importancia de la escala

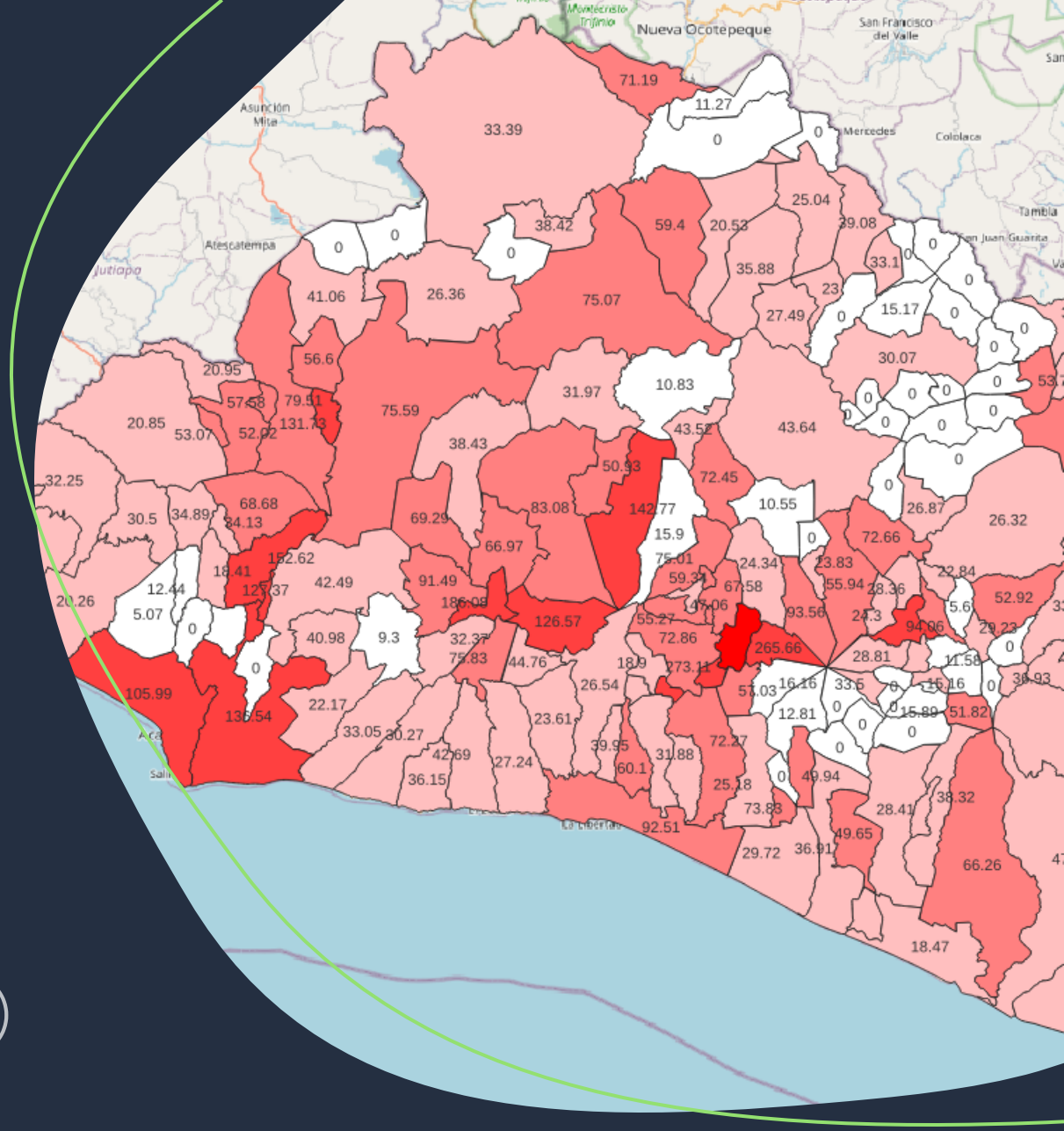
Figure 2 Moran's dispersion map for the femicide rates per 100 thousand inhabitants, in Sergipe's municipalities, in the period 2013–2017.




Todos los lugares están relacionados, pero los lugares cercanos están más relacionados que los lugares lejanos.
(1a ley de la geografía, Tobler)

Estadística espacial

Métodos para analizar la distribución espacial, patrones, procesos y relaciones entre los fenómenos. Se desarrolló específicamente para ser trabajada con datos geográficos a diferencia de los métodos tradicionales no espaciales, incorpora el espacio (proximidad, área, conectividad y otras relaciones espaciales) en su matemática.



Estadística espacial

Se basa en la **primera ley de la geografía**. Es la **correlación de una variable consigo misma** a través del **espacio** (sólo una variable está involucrada). Los patrones podrían indicar que los datos no son independientes unos de otros, rechazando la hipótesis de independencia de algunos test estadísticos.

Estadística espacial

Los test de autocorrelación espacial permiten contestar:

¿Cómo se distribuye el fenómeno?

¿Cuál es el patrón espacial del fenómeno?

¿Dónde se concentra el fenómeno) los aglomerados/cluster)?

¿Cómo se comparan los cluster de diferentes variables entre sí?

Vecinos espaciales y matriz de pesos

Vecindario = área en la que el SIG compara el valor objetivo con el valor de los vecinos.

Los vecindarios están definidos casi siempre por adyacencia o distancia, pero pueden ser definidos con base en tiempo de viaje, costo de viaje, etc.

Una matriz de pesos se usa para incorporar estos conceptos en el análisis estadístico.

Modelos de distancia

Distancia inversa: los fenómenos cercanos tienen más influencia.

Banda de distancia: objetos o fenómenos fuera de un nivel de tolerancia no tienen influencia.

Zona de indiferencia: Combina los dos criterios anteriores.

Modelos de adyacencia

K Nearest Neighbors – se especifica un número de vecinos a incluir en el cálculo.

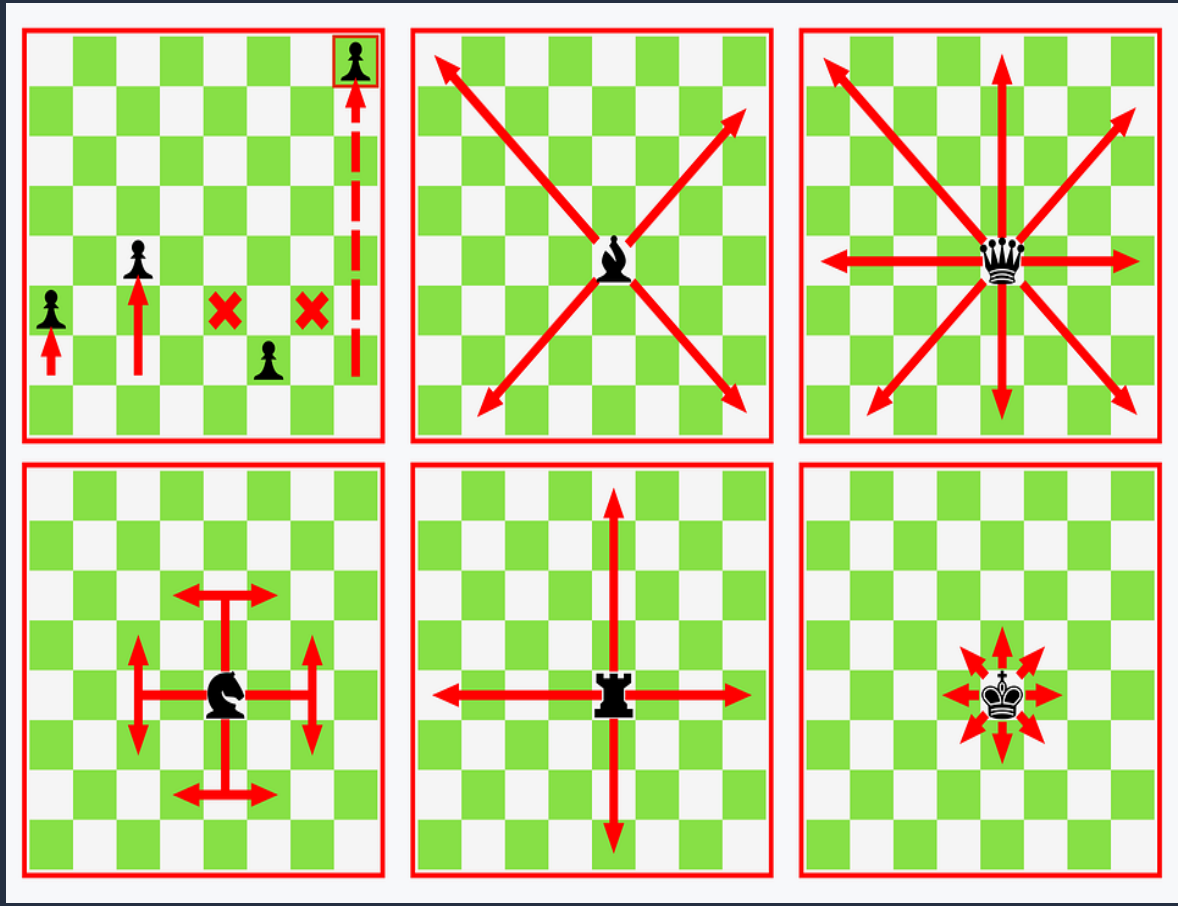
Polygon Contiguity – polígonos que comparten lado o nodo se influencian entre sí.

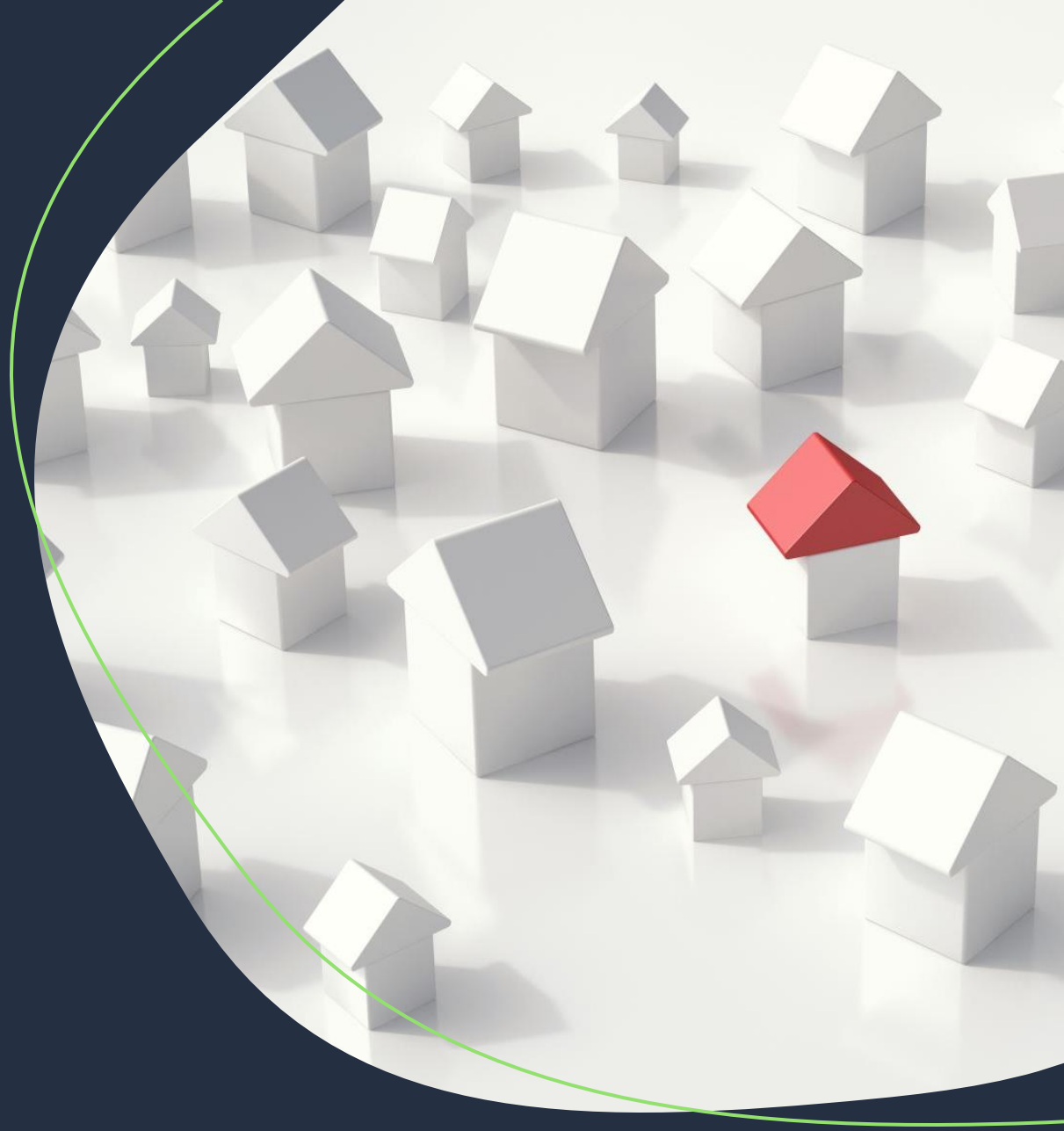
Spatial weights – peso de influencia se especifica por el usuario (ej: tiempos de viaje, etc.)

Modelos de adyacencia

- Torre: comparte lados
- Alfíl: comparte vértices
- Reina: comparte lados y vértices

Tipos de contigüidad de los vecinos





Estadística espacial

- Estadísticas globales – Identifican y miden el patrón de toda el área de estudio, no indica dónde sucede dicho patrón.

- Estadísticas locales – identifican variación a través de un área de estudio, enfocándose en objetos individuales y su relación con los vecinos (ej. áreas específicas de aglomerados o clusters)

Hay distintas herramientas:

- Inspección visual con mapeo
- I de Moran global
- Índices locales de correlación espacial : LISA (por sus siglas en inglés Local Indicator of Spatial Autocorrelation)

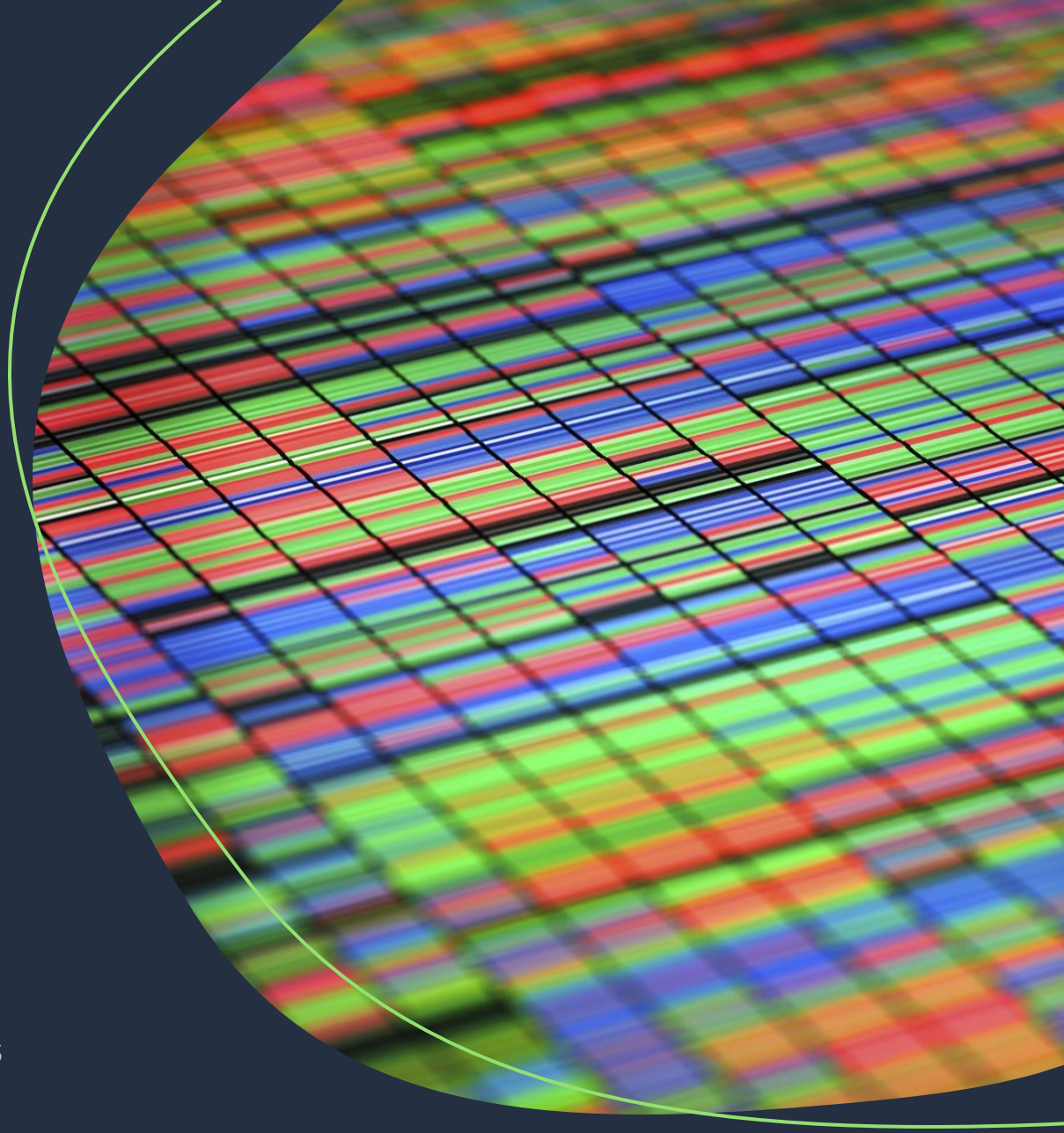
Índice de Morán global (Moran's I statistic)

A partir de la inspección visual, se observa que el homicidio se agrupa espacialmente.

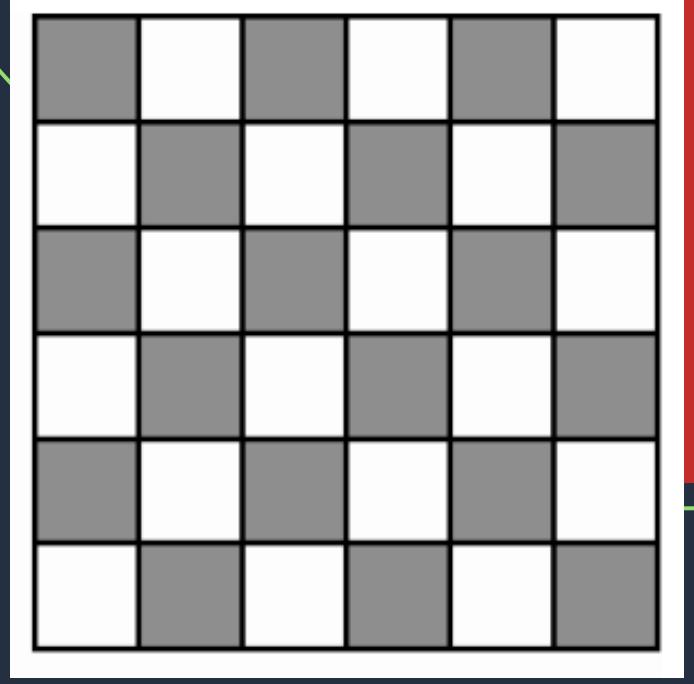
- Municipios que tienen una alta tasa de homicidio tienden a estar rodeados de otros municipios con alta tasa de homicidios.
- Ahora, ¿Es este agrupamiento espacial estadísticamente significativos?
-

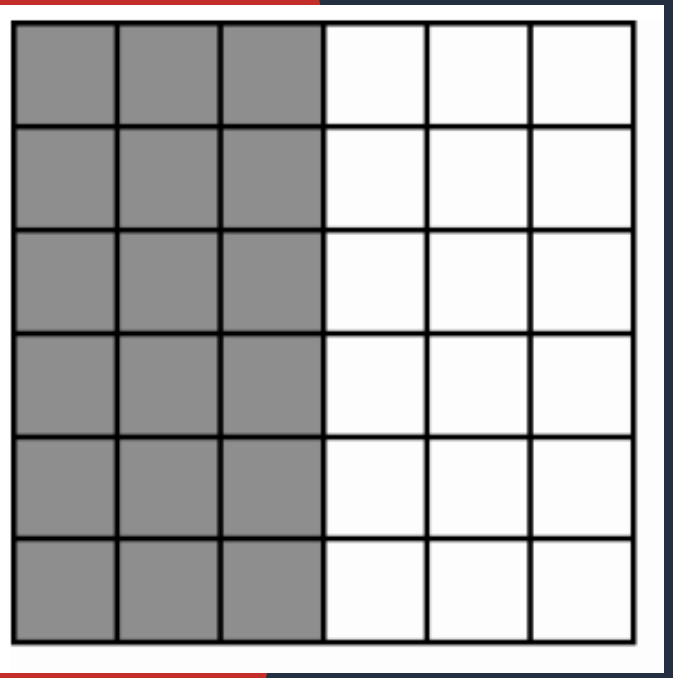
I de Moran global

- Es una medida global de autocorrelación espacial.
- Nos indica si valores altos están rodeado por valores altos (datos agrupados, dispersos o aleatorios?)



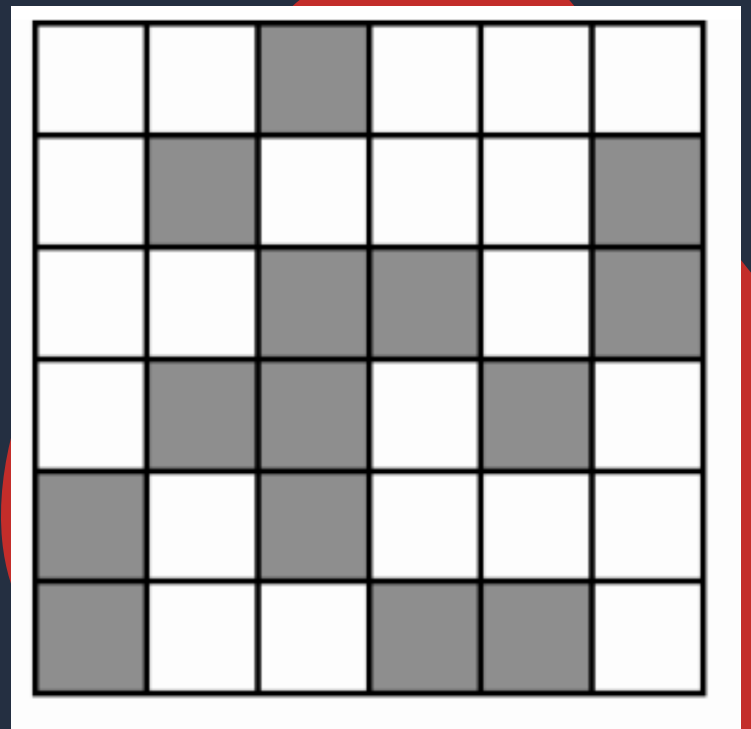
Correlación espacial negativa





Correlación espacial positiva

No correlación espacial



I de Morán global (Moran's statistic)

Queen (First Order)

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	1	1	1	0	0
0	0	1	i	1	0	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

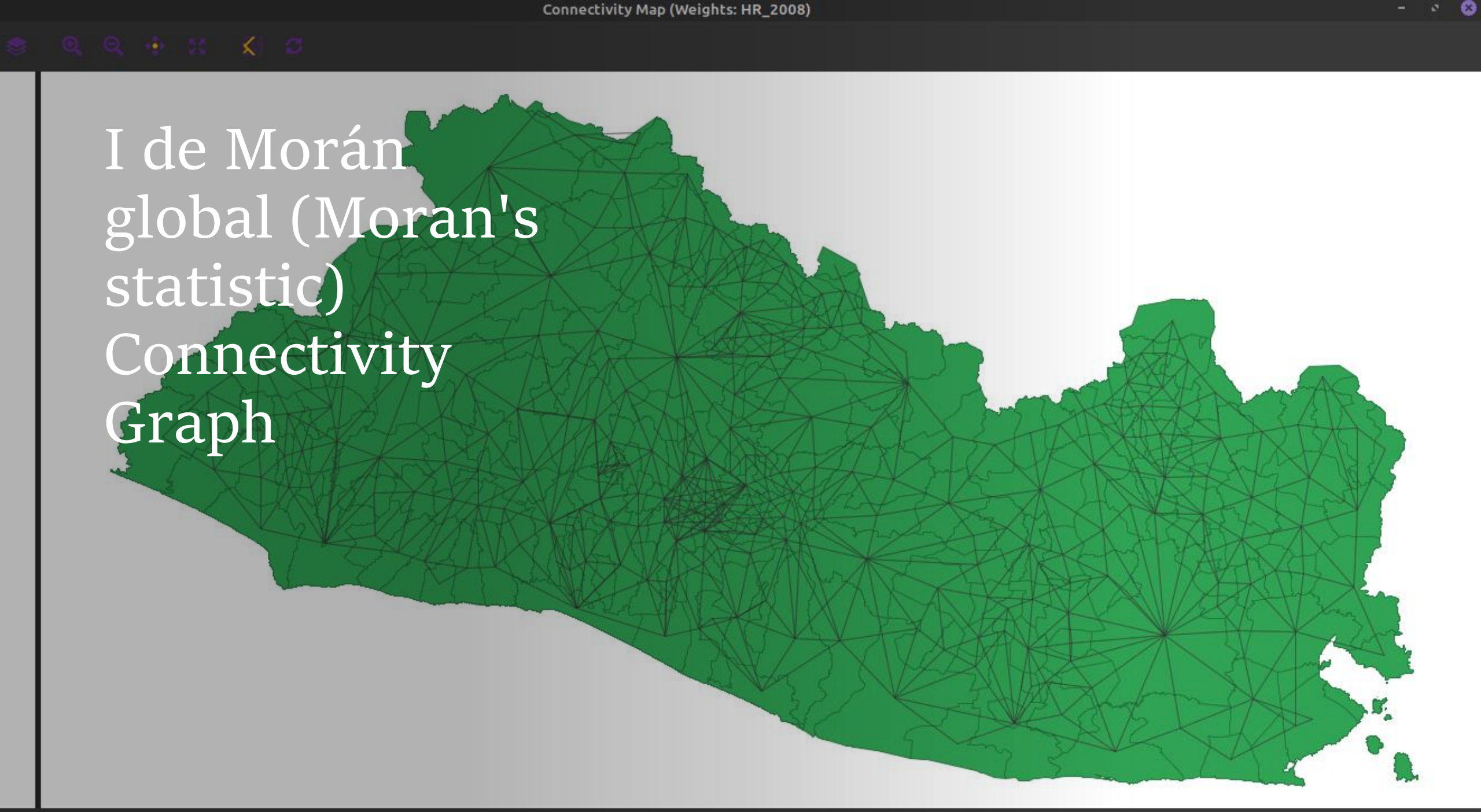
Queen (Second Order)

0	0	0	0	0	0	0
0	0.5	0.5	0.5	0.5	0.5	0
0	0.5	1	1	1	0.5	0
0	0.5	1	i	1	0.5	0
0	0.5	1	1	1	0.5	0
0	0.5	0.5	0.5	0.5	0.5	0
0	0	0	0	0	0	0

Antes de ejecutar la estadística de I de Moran Global, se tiene que crear una matriz de pesos espaciales.

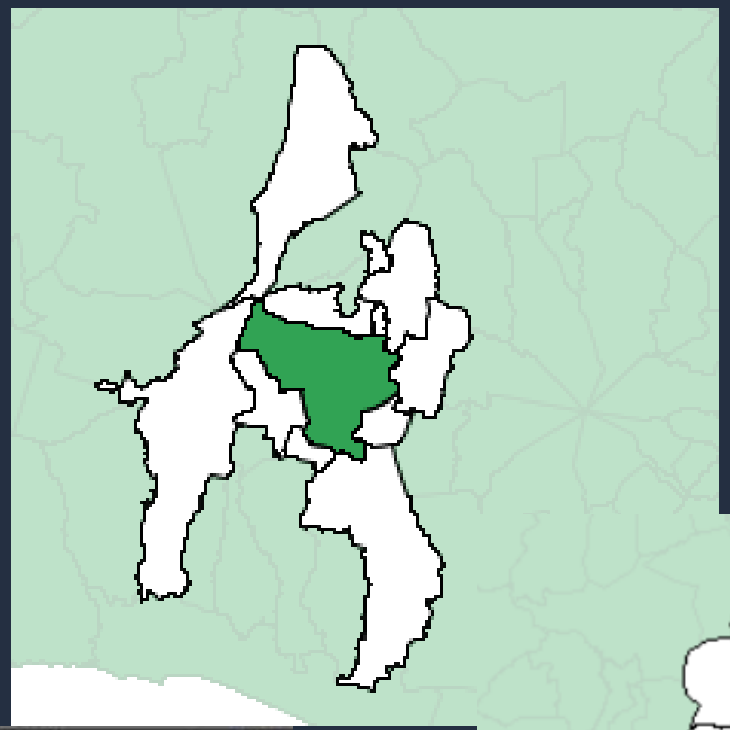
En esta hay que determinar quienes son los vecinos que cuentan.

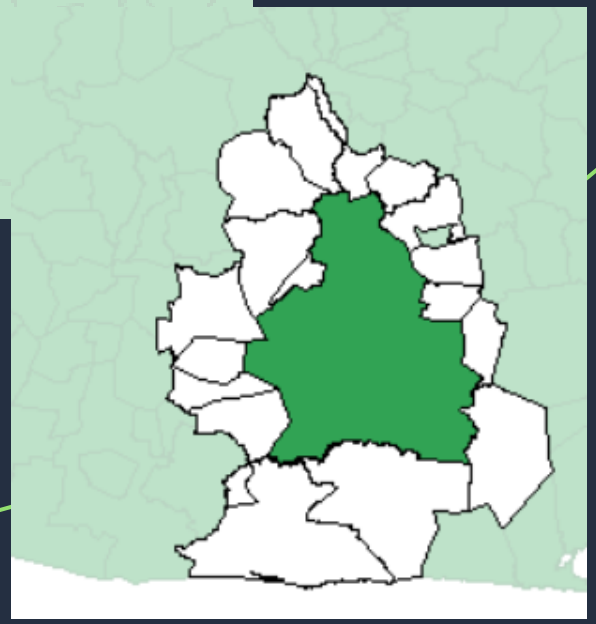
I de Morán global (Moran's statistic) Connectivity Graph

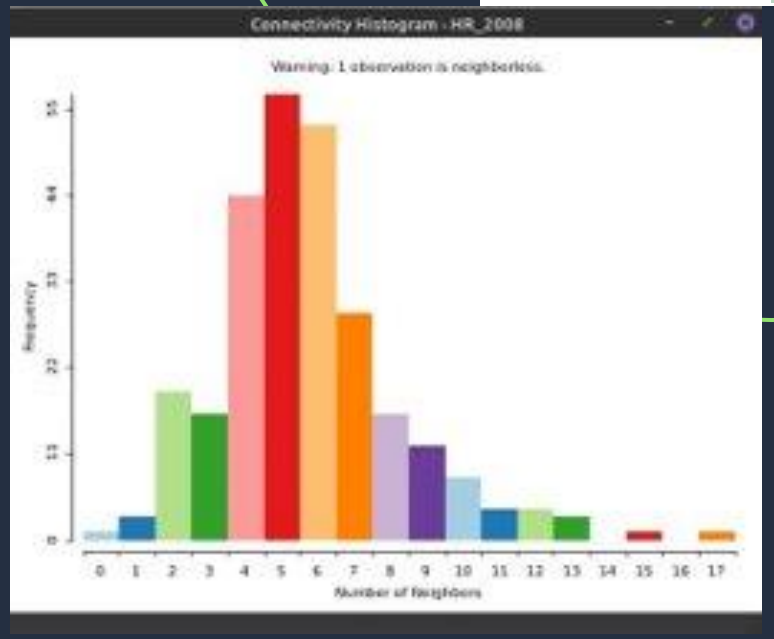


I de Morán global (Moran's statistic) Connectivity Graph

El “Connectivity map” muestra el municipio objetivo y los vecinos relacionados.

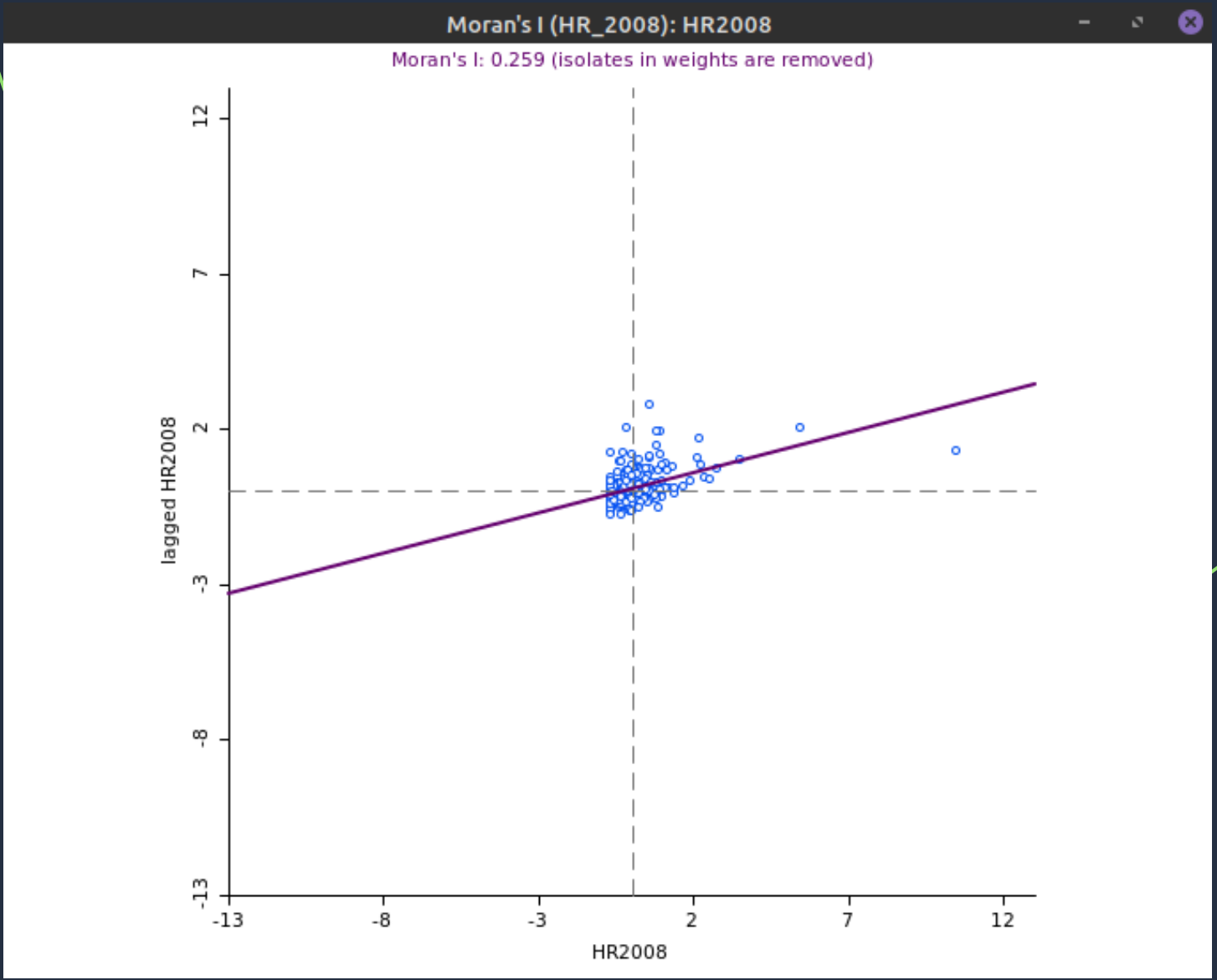






I de Morán global (Moran's statistic)

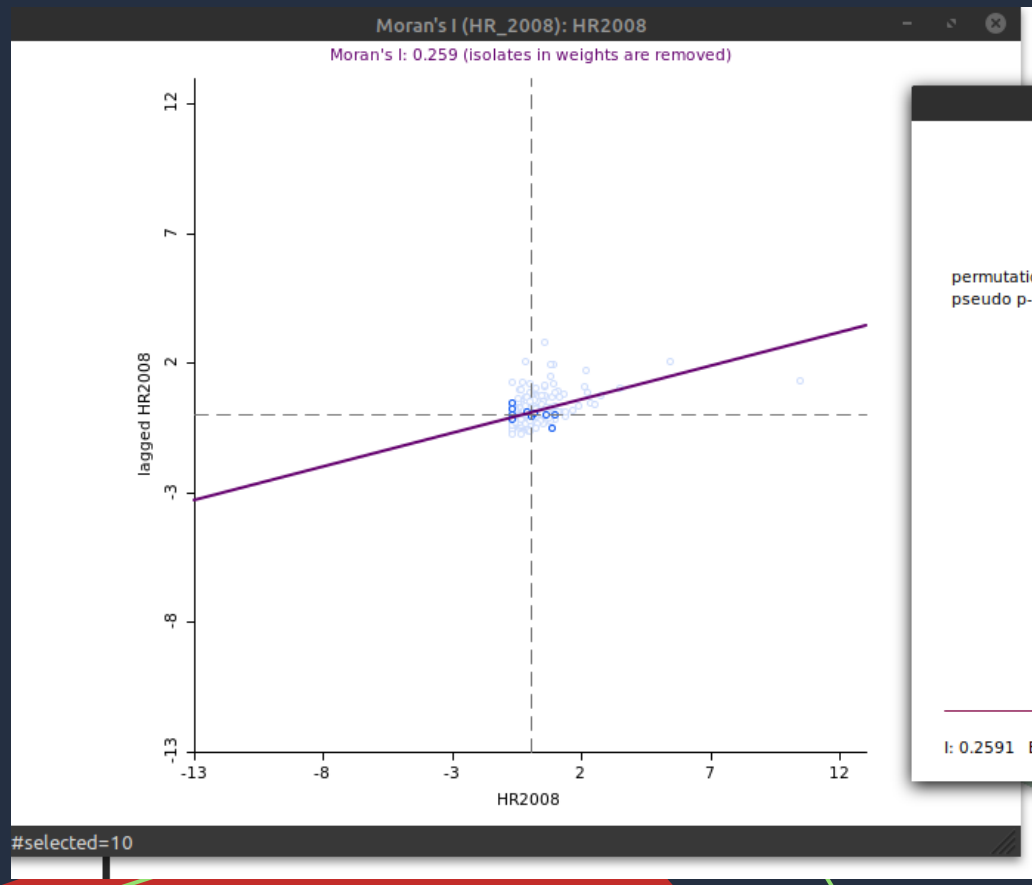
La I de Moran Global es positiva (0.26). Esto significa que municipios con alta tasa de homicidios están rodeados de otros municipios con alta tasa de homicidio o municipios de baja tasa están rodeados de otros municipios con baja tasa.

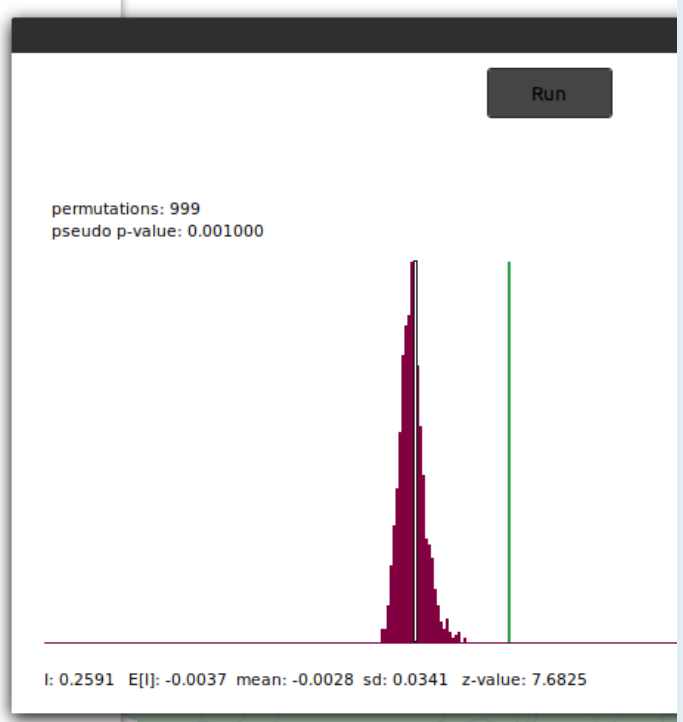


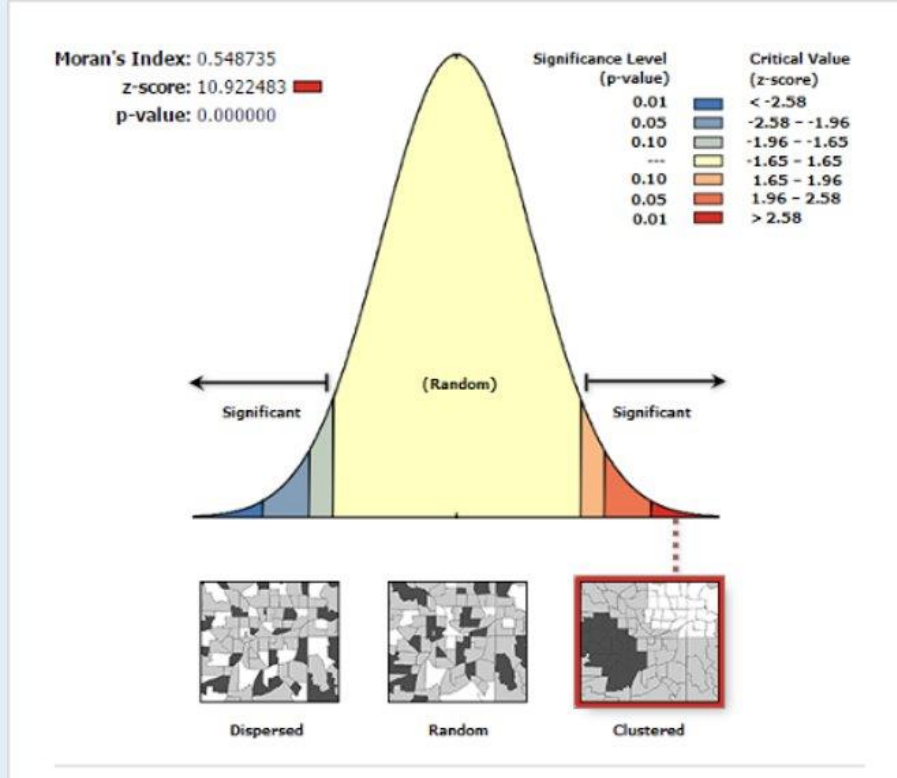
I de Morán global (Moran's statistic)

Randomize 999 permutaciones

Evaluación de hipótesis nula: fenómeno aleatorio







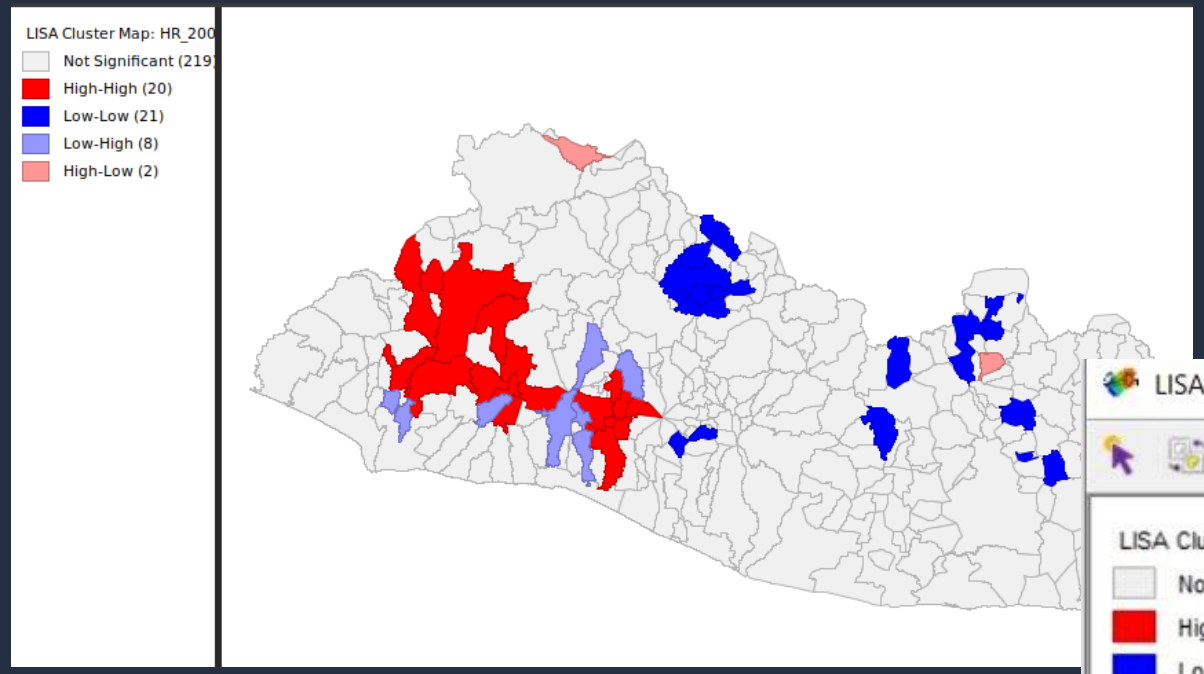
Indicadores locales de correlación espacial.LISA (Local Indicator of Spatial Autocorrelation)

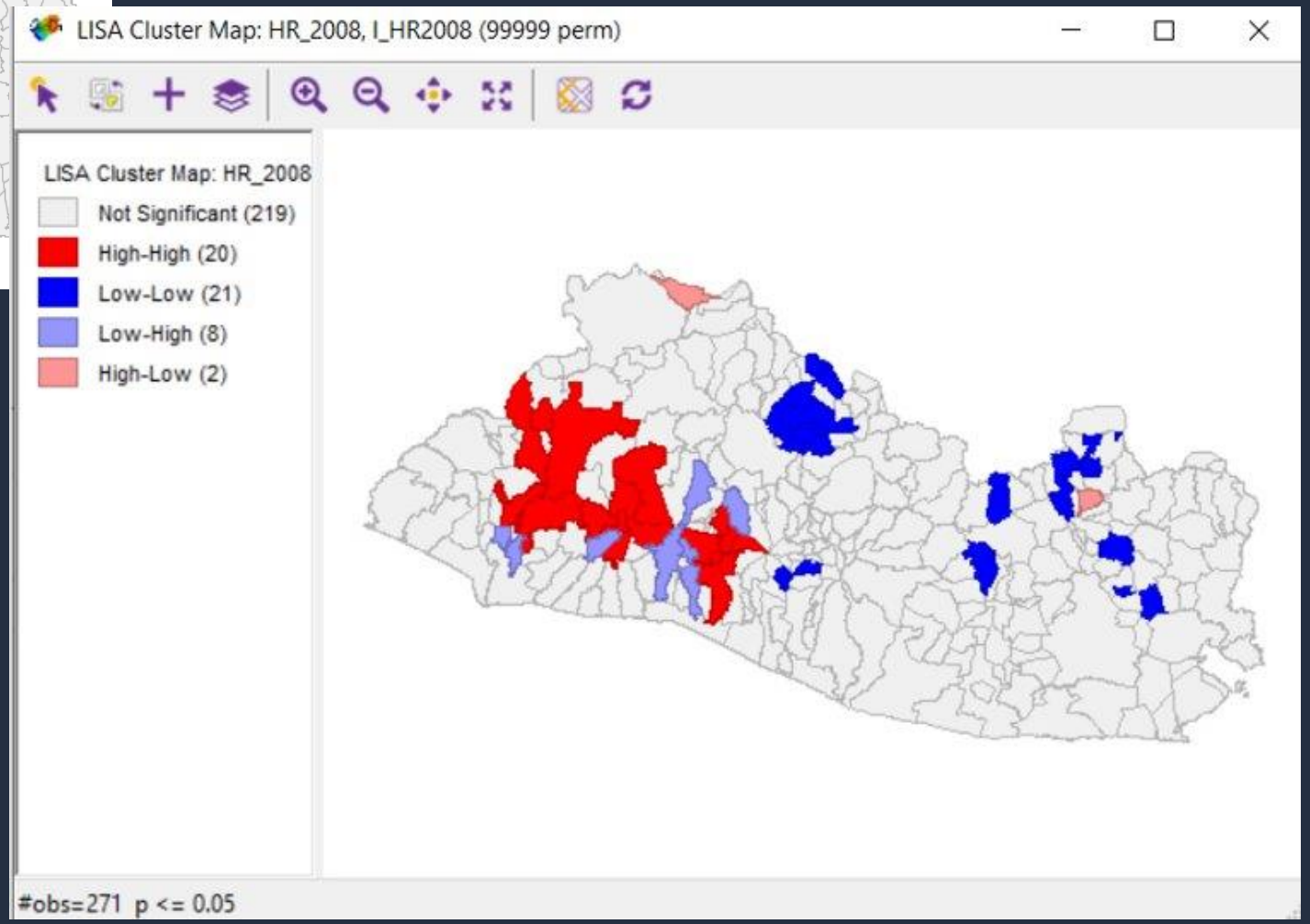
Pero ¿donde estan las concentraciones significativas?

I de Moran es una medida global.

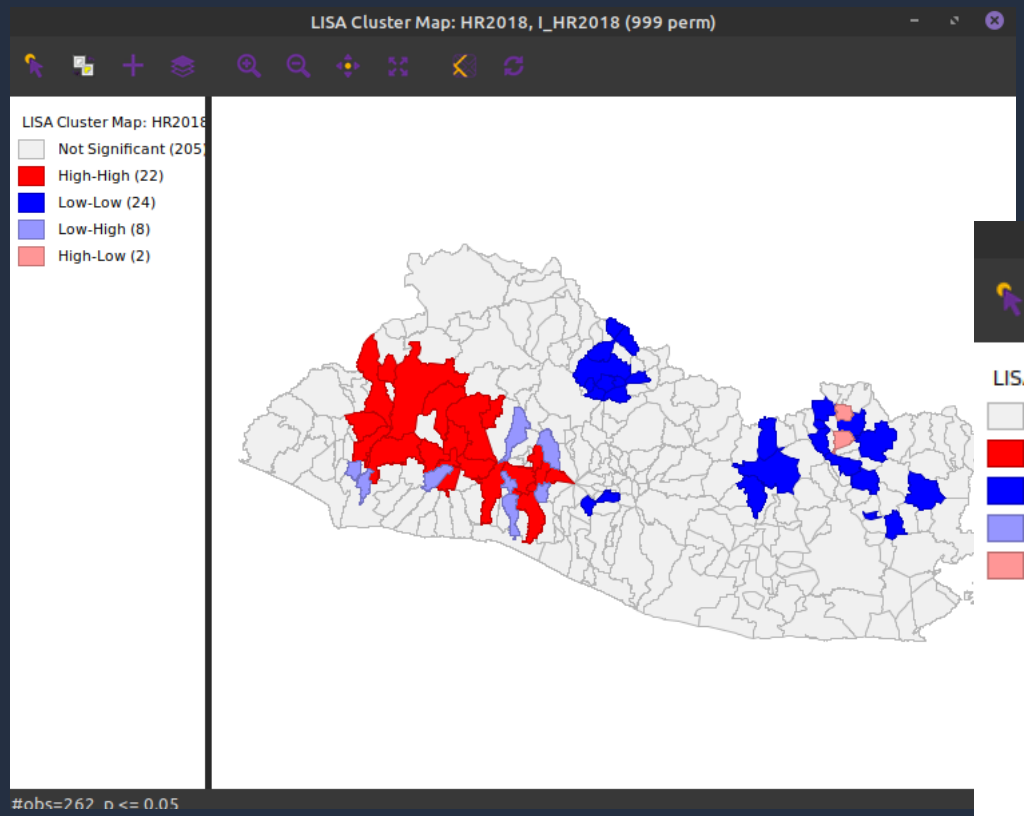
LISA

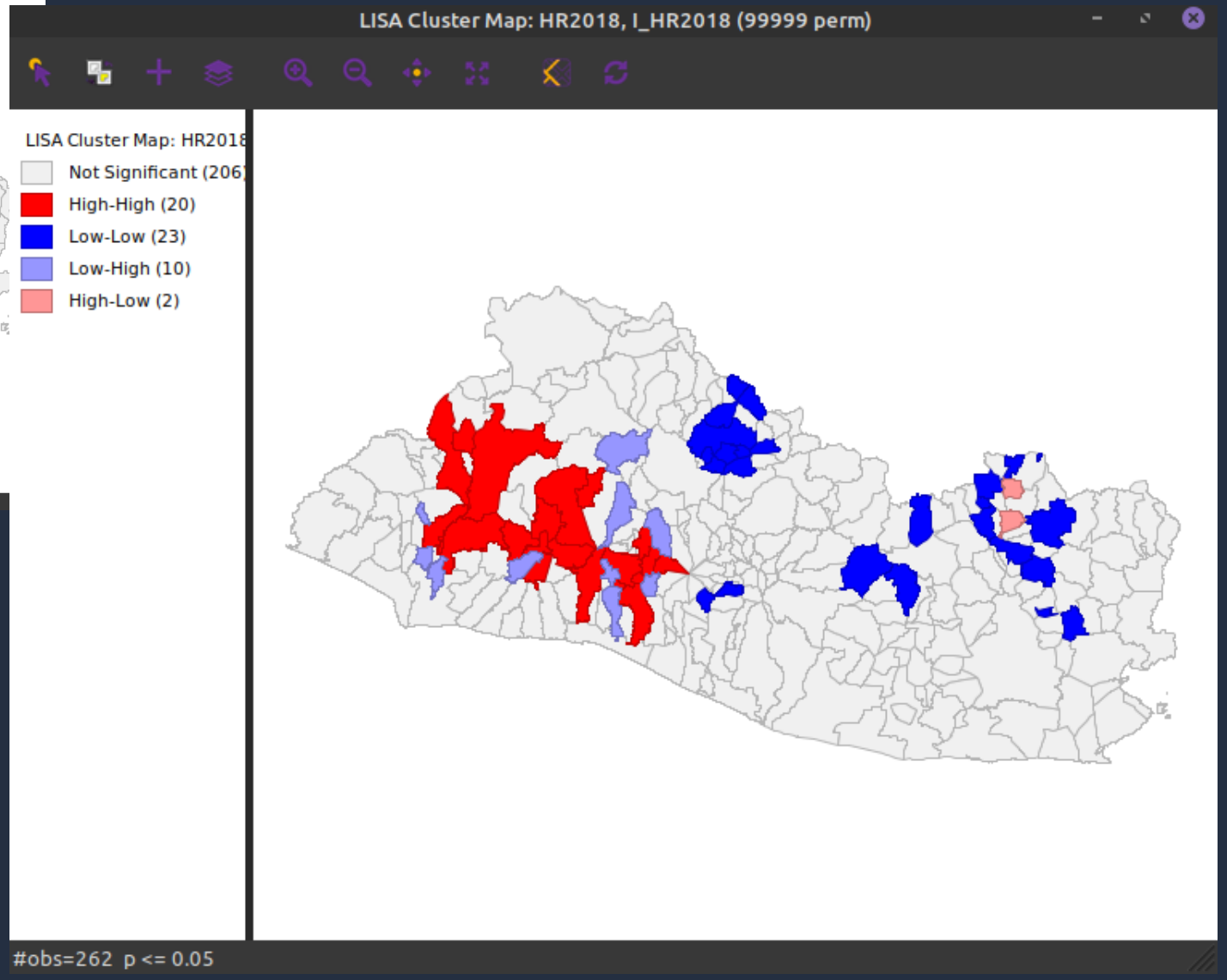
Nos muestra donde están la concentraciones significativas





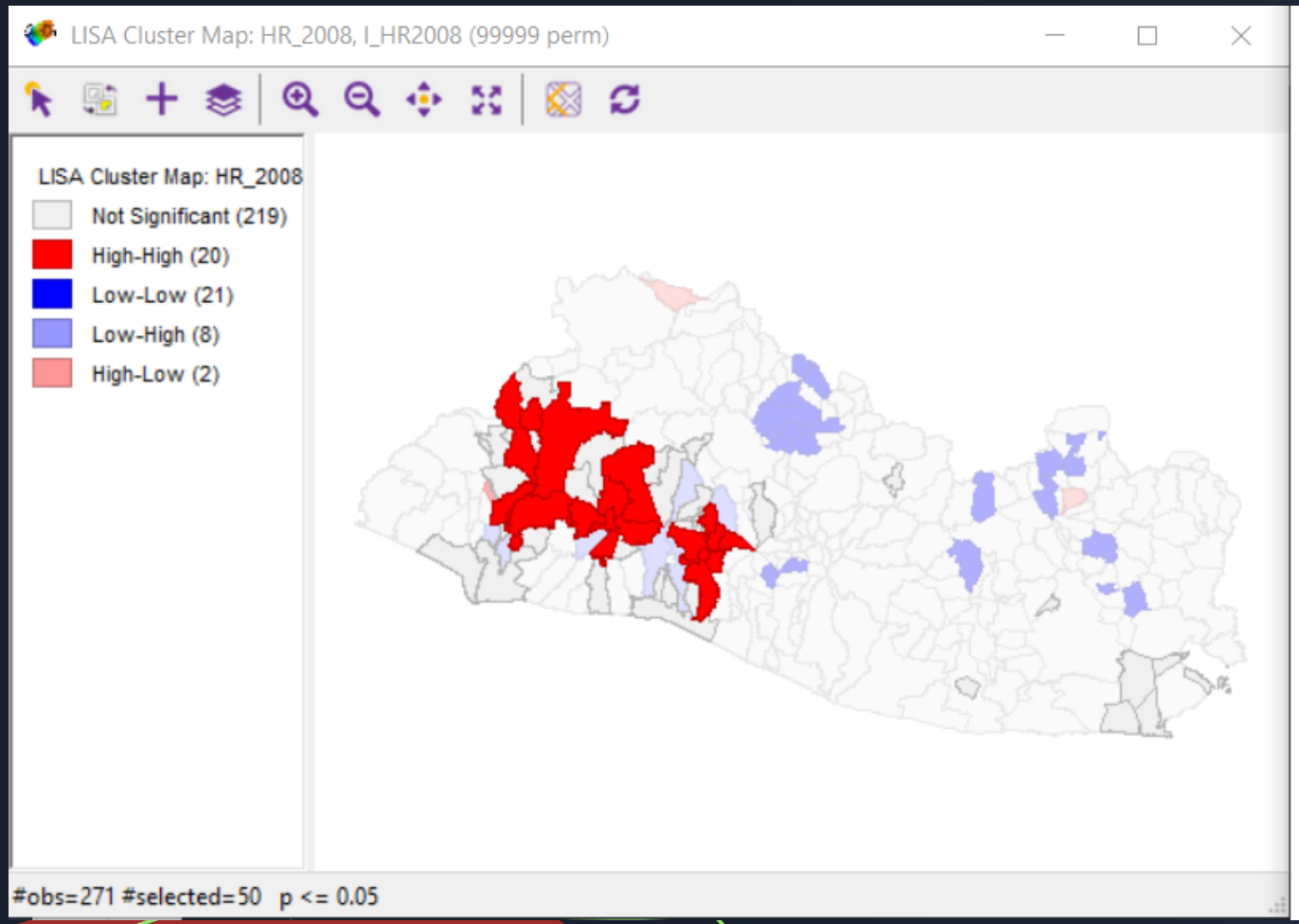
LISA Queen 1 2008

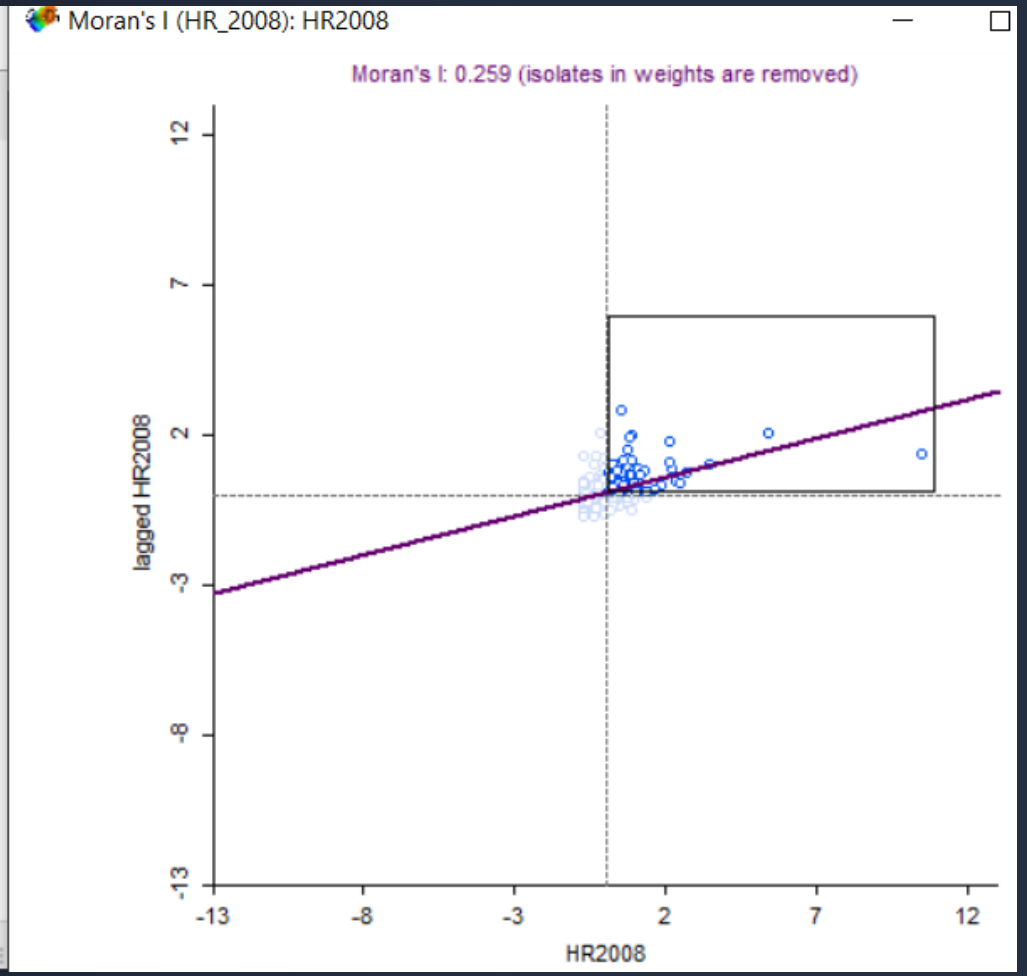




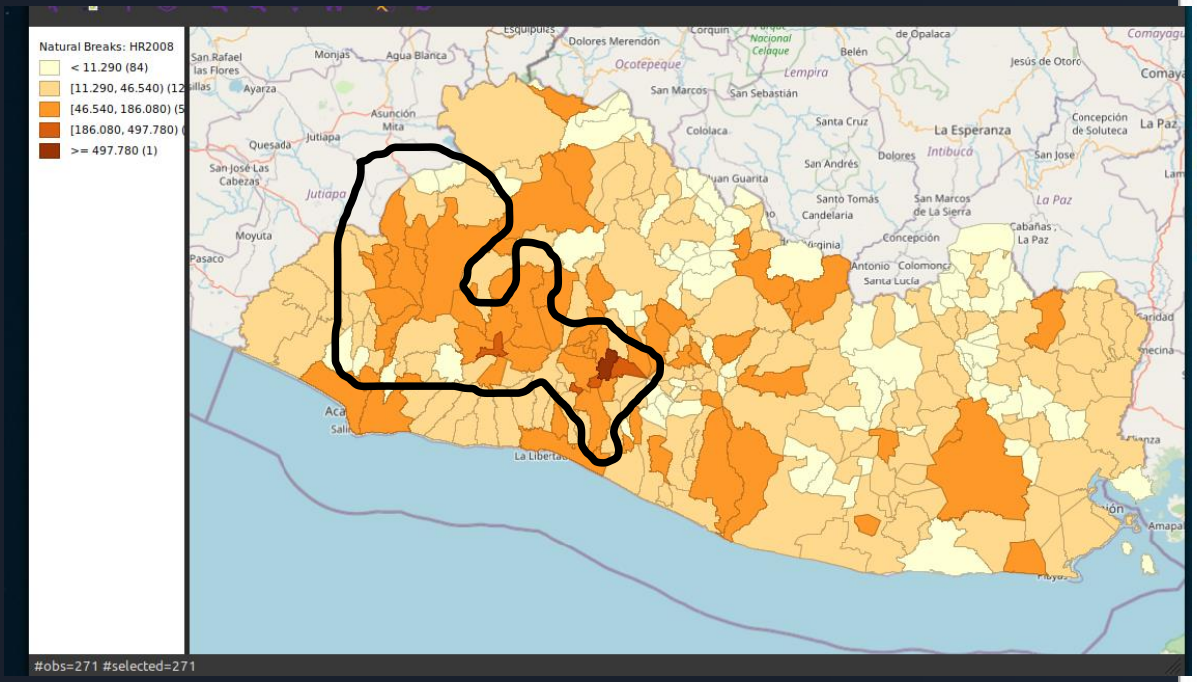
LISA Queen 1 2018

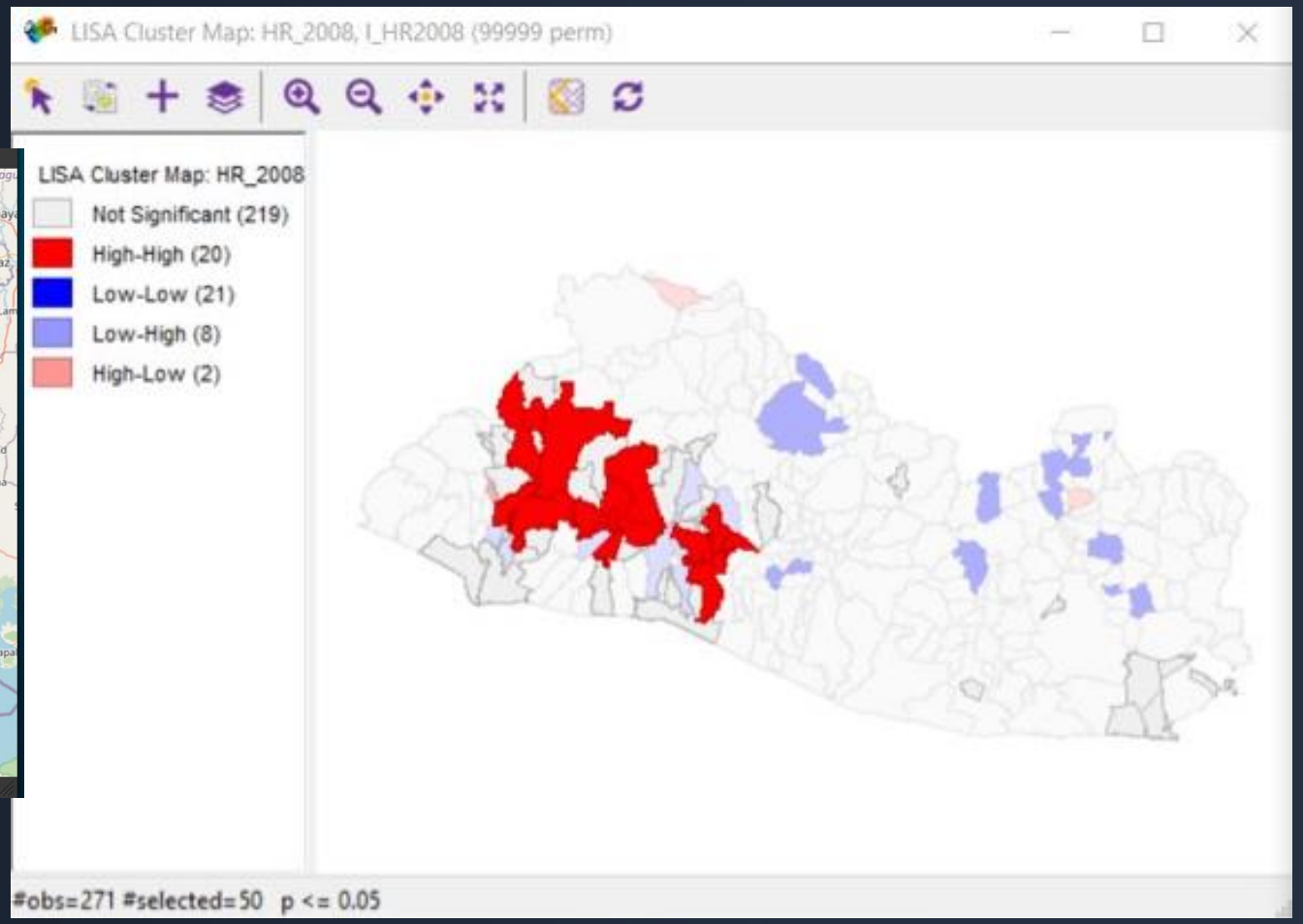
LISA Queen 1


LISA Cluster Map: HR_2008, I_HR2008 (99999 perm)
LISA Cluster Map: HR_2008
Not Significant (219)
High-High (20)
Low-Low (21)
Low-High (8)
High-Low (2)
#obs=271 #selected=50 p <= 0.05

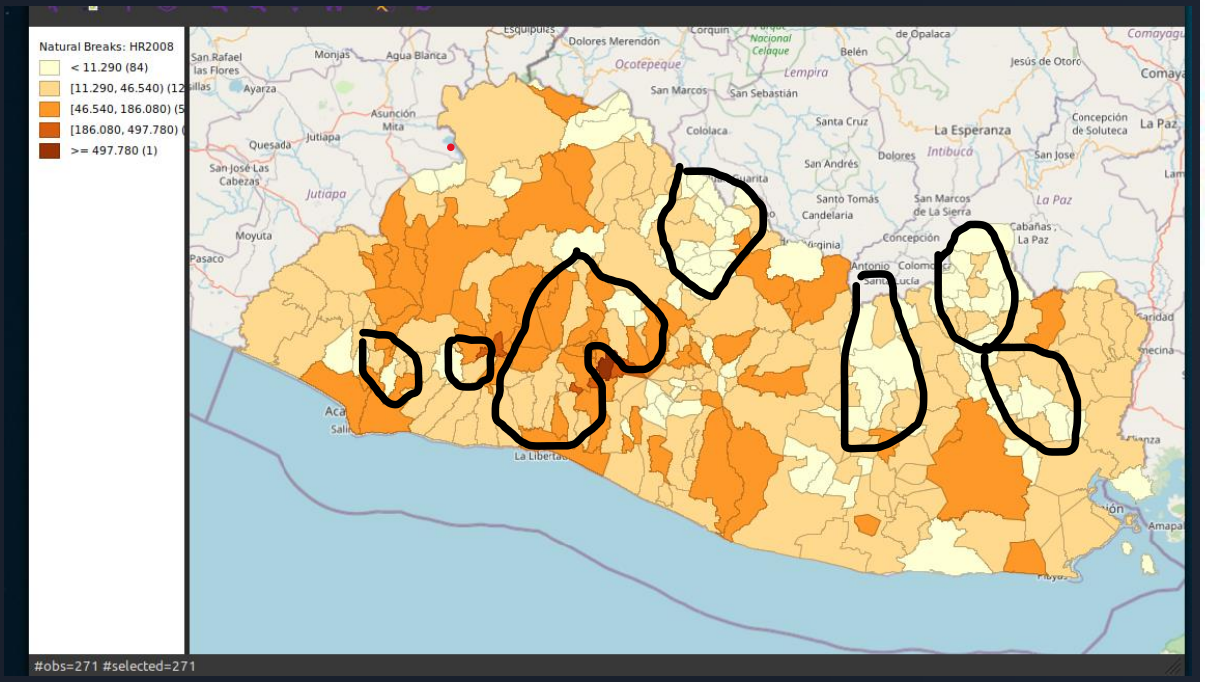

Moran's I (HR_2008): HR2008
Moran's I: 0.259 (isolates in weights are removed)
lagged HR2008
HR2008
12
7
2
-3
-8
-13
-13
-8
-3
2
7
12

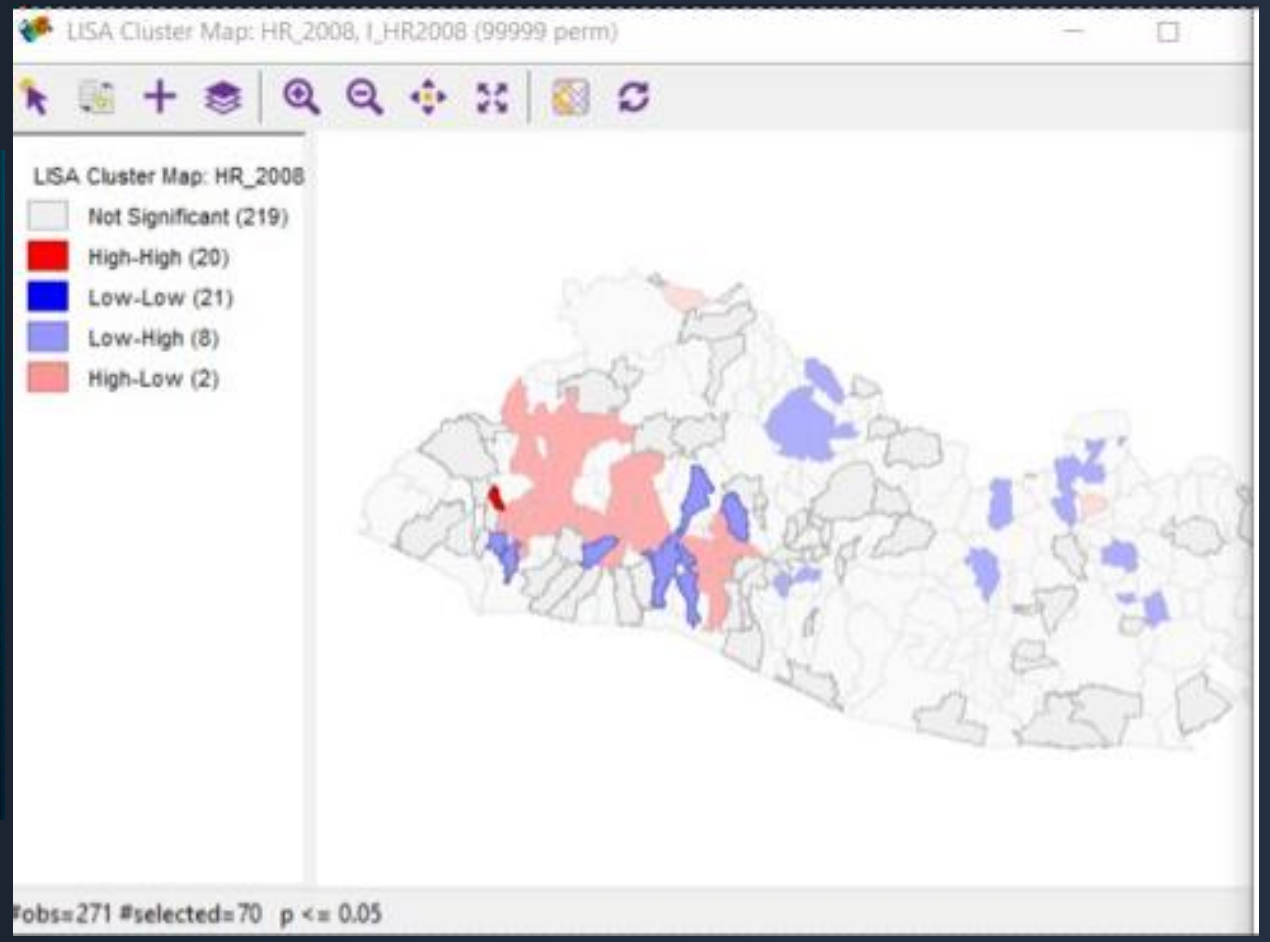
LISA Queen 1: Alto-alto



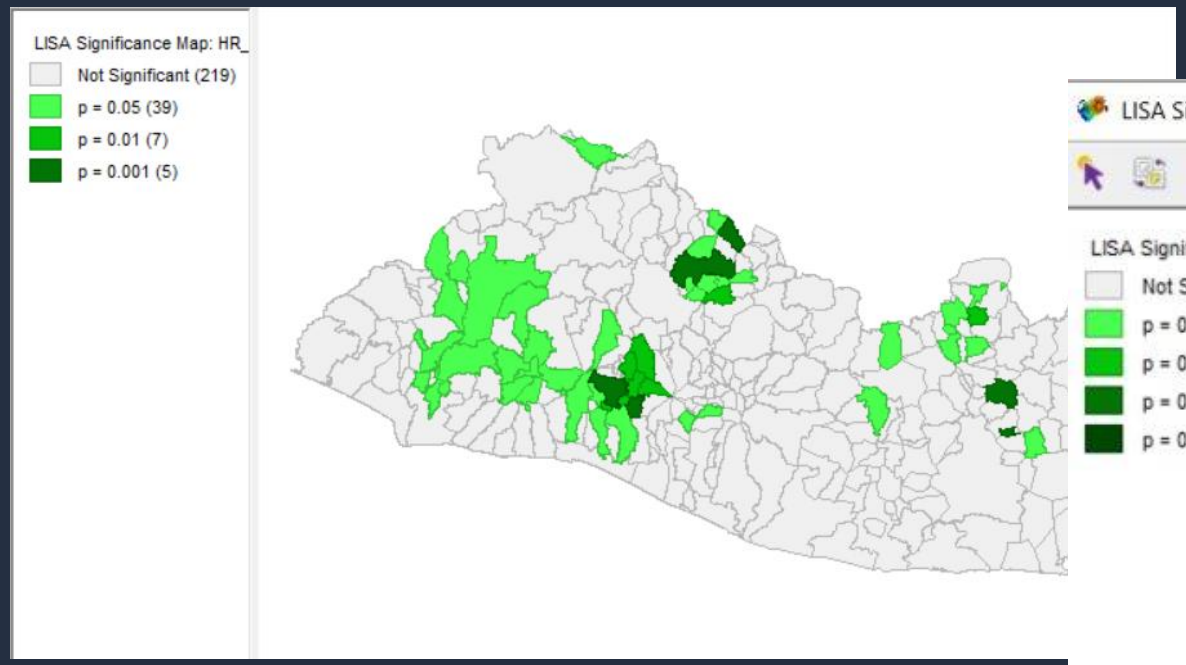


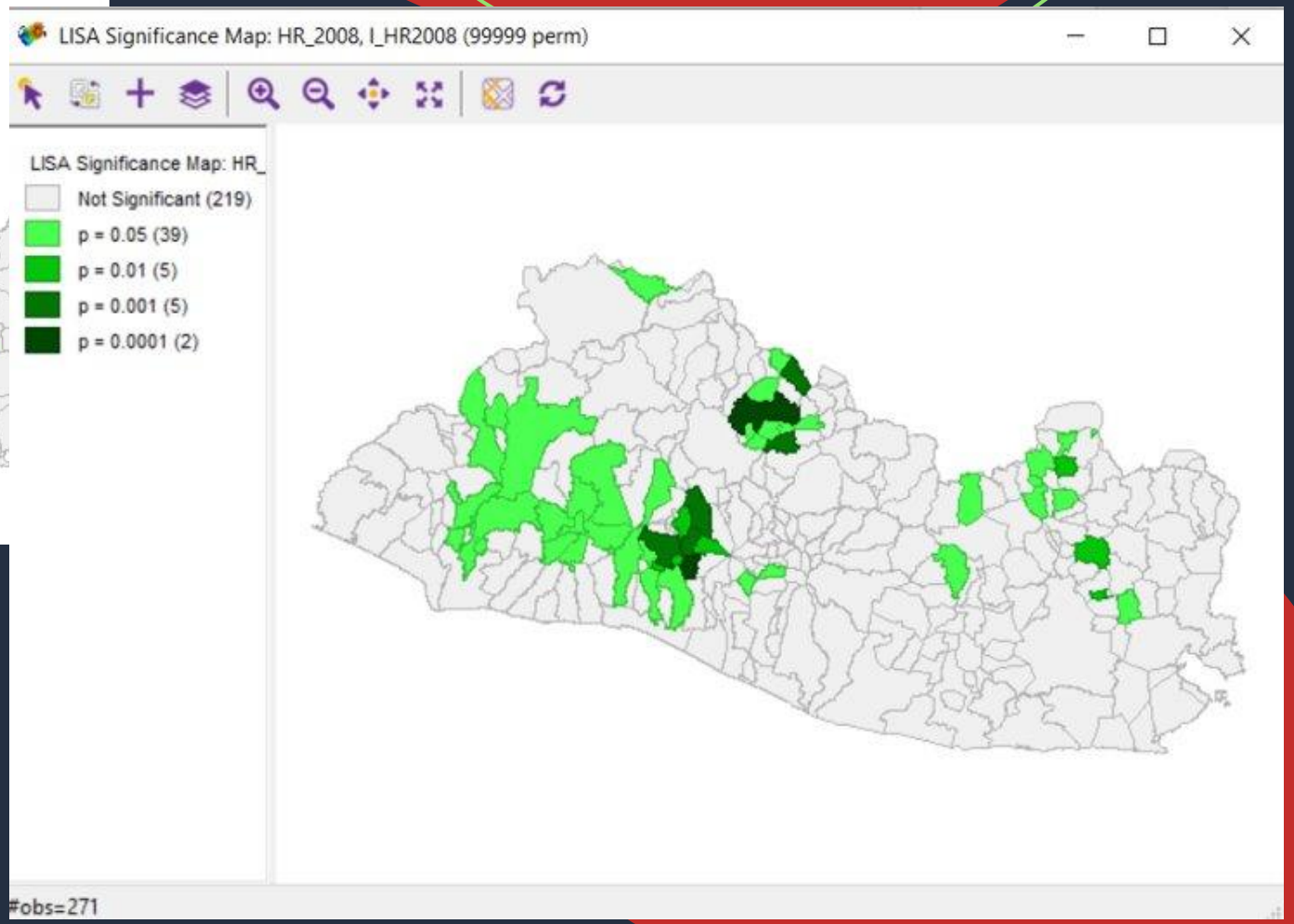
LISA Queen 1:outliers (valores atípicos)

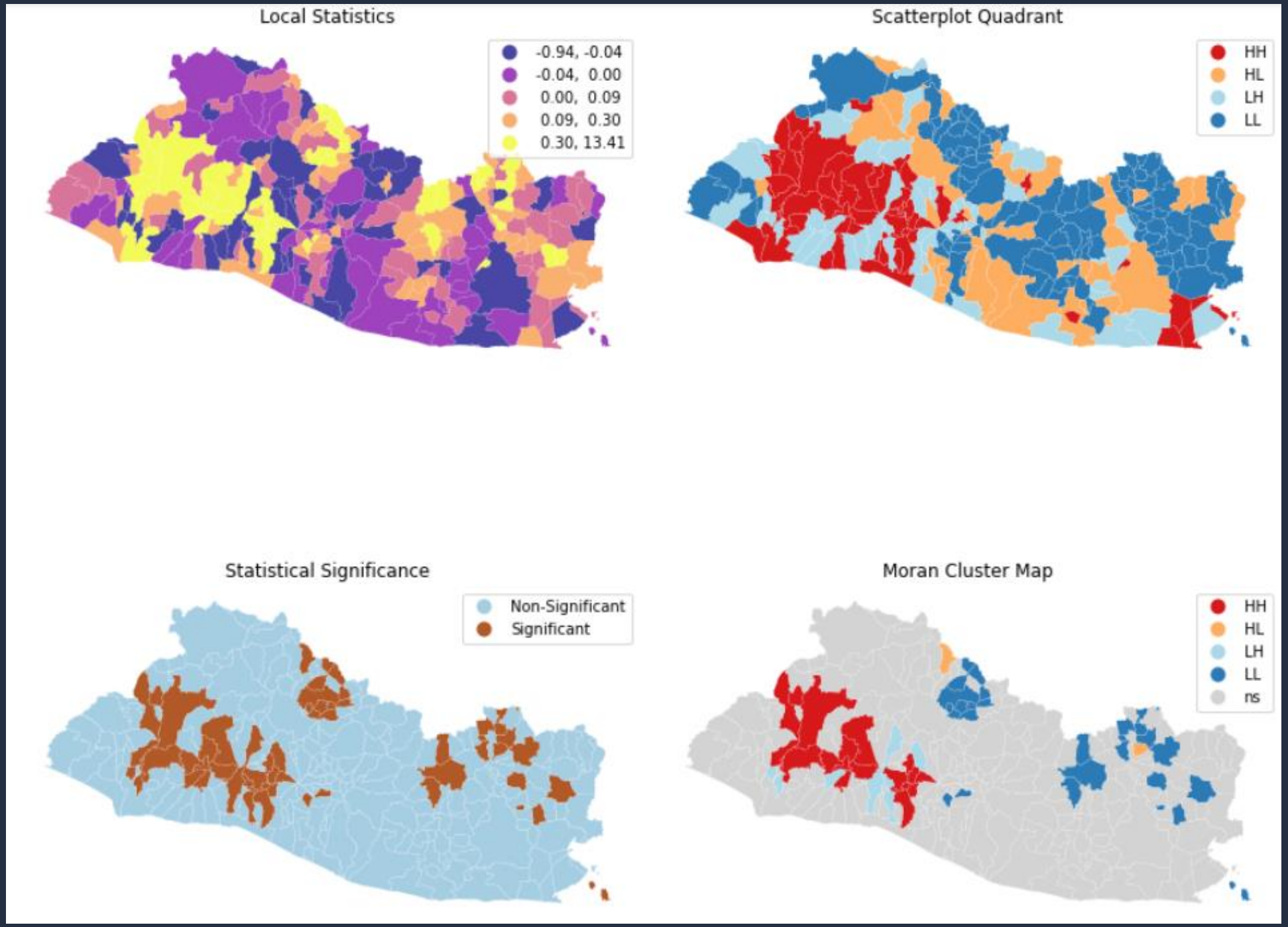




LISA Queen 1:Mapa de significancia





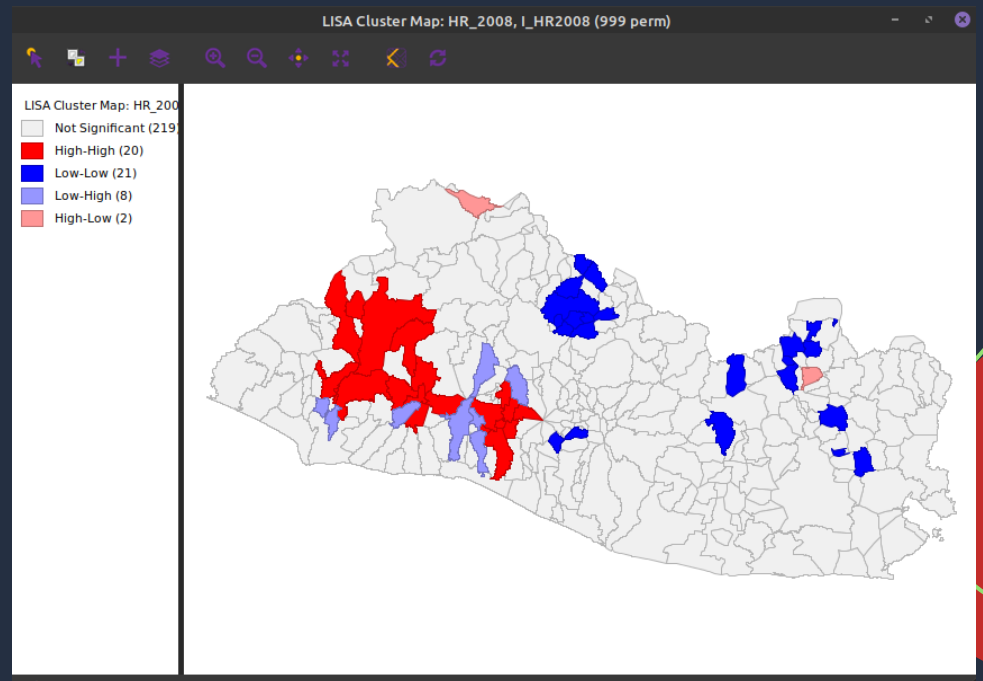

Local Statistics
-0.94, -0.04
-0.04, 0.00
0.00, 0.09
0.09, 0.30
0.30, 13.41
Scatterplot Quadrant
HH
HL
LH
LL
Statistical Significance
Non-Significant
Significant
Moran Cluster Map
HH
HL
LH
LL
ns

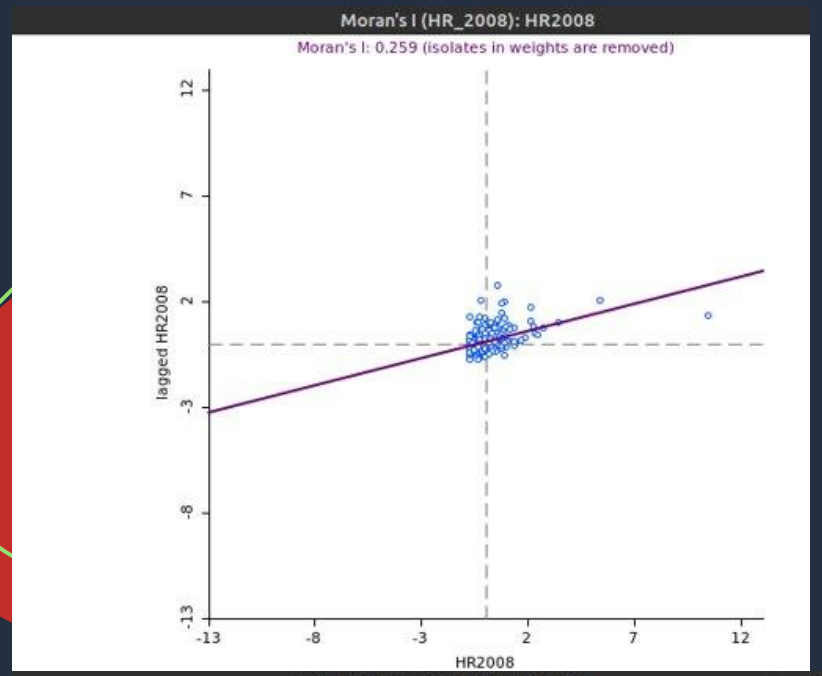
LISA Queen 1

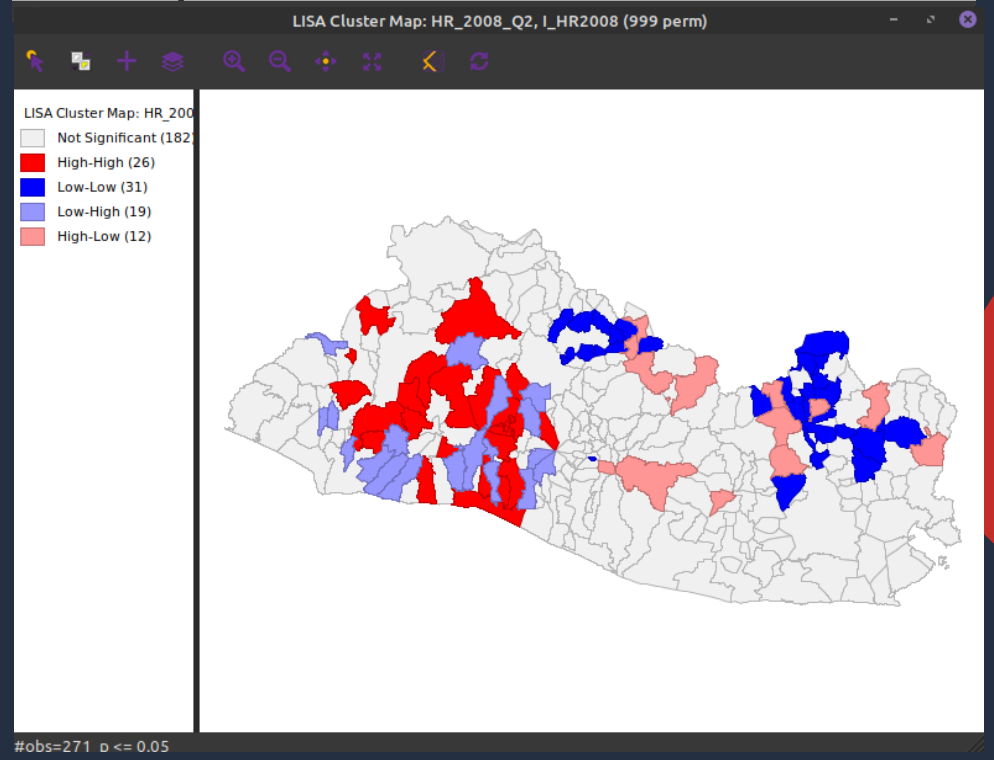
¿Cual resultado utilizaría usted?

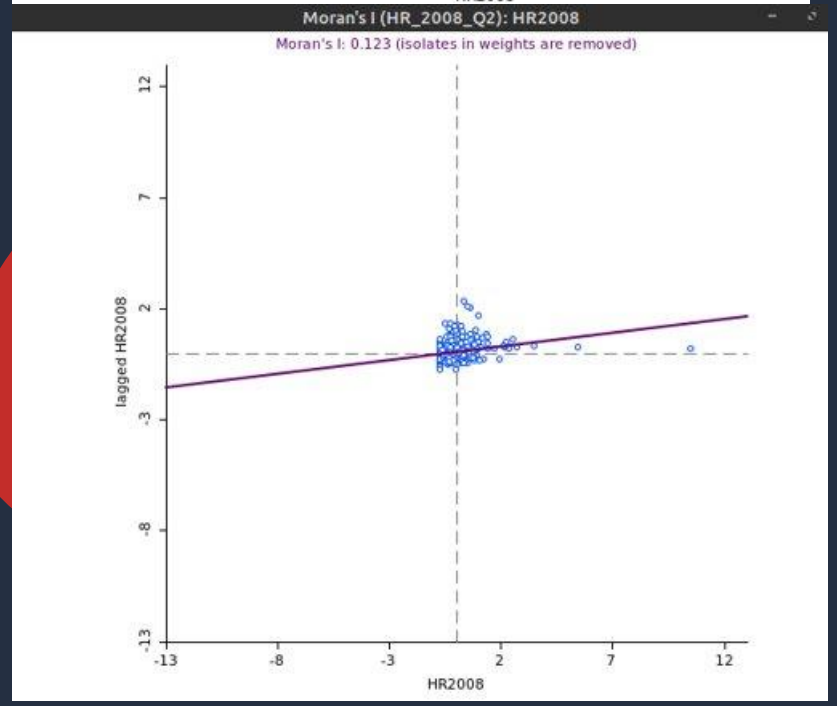
LISA Queen 2

Siempre elija la matriz que de el I de Moran's más alto.



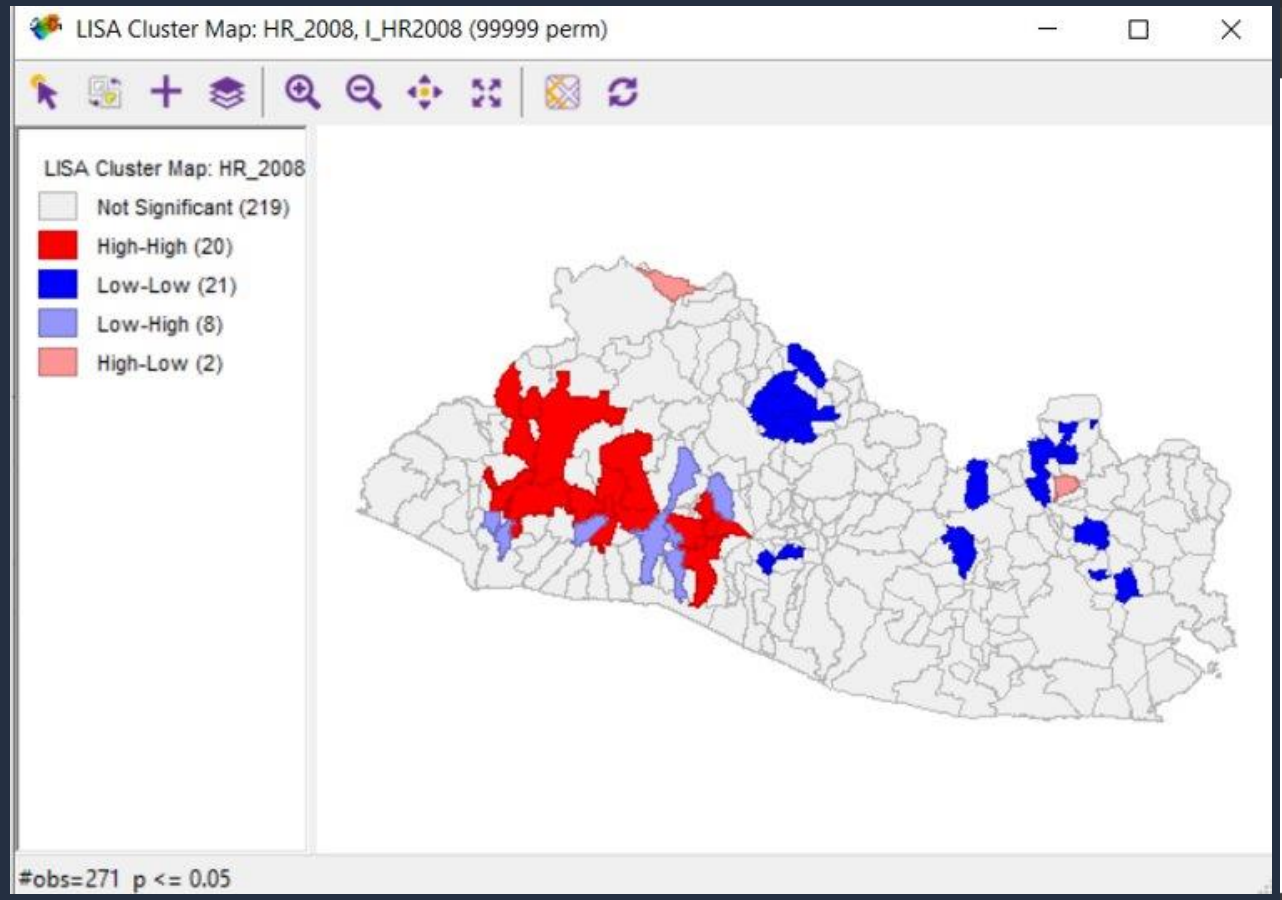




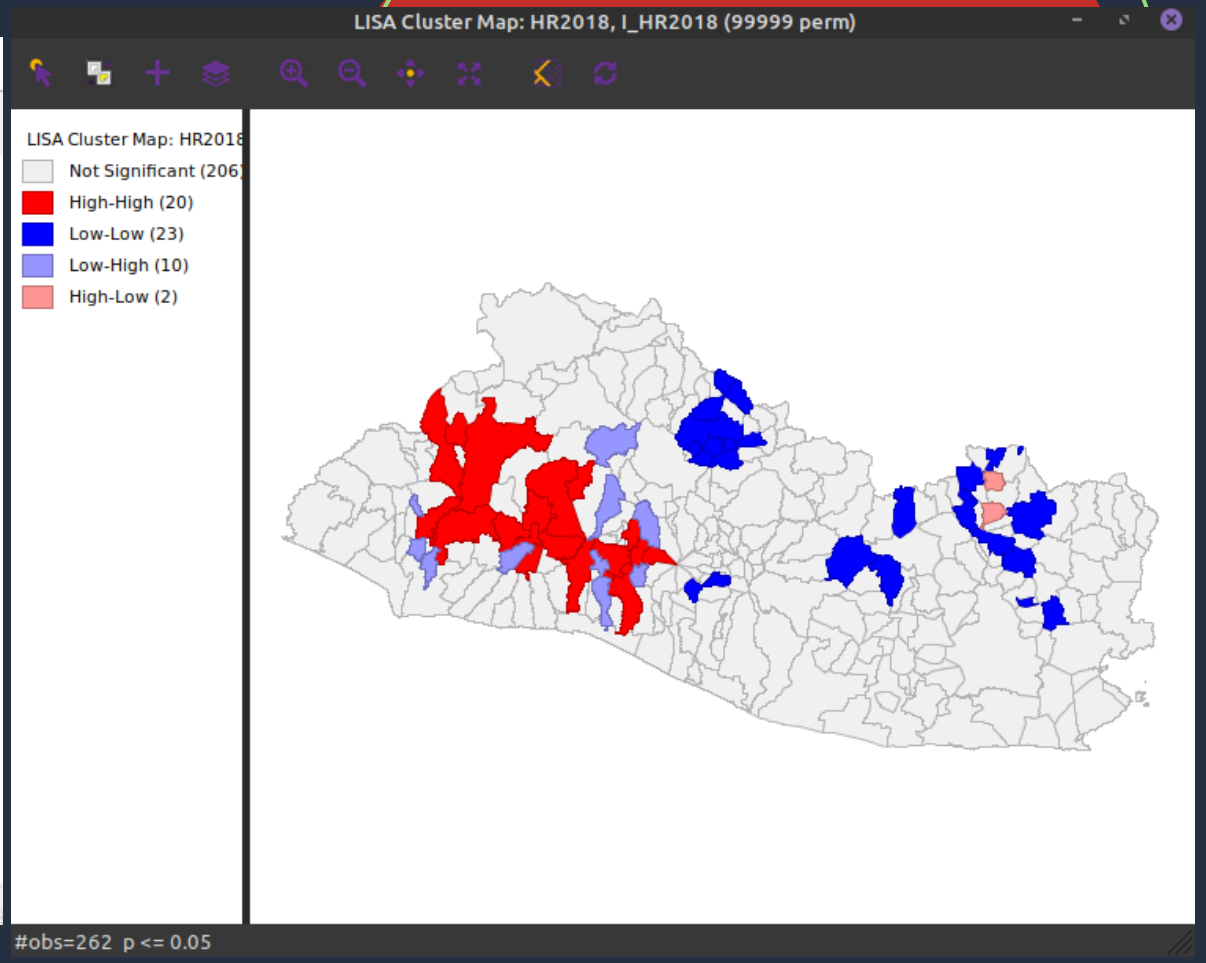


¿Cómo evolucionó el patrón en 10 años?

2008



2018



¿Cómo evolucionó el patrón en 10 años?

Puntos calientes históricos

1. SA: Chalchuapa
2. SA: Santa Ana
3. SA: El Porvenir
4. SO: Izalco
5. SO: Nahuizalco
6. SO: Armenia
7. LI: Sacacoyo
8. LI: Jayaque
9. LI: Ciudad Arce
10. LI: Opico
11. LI: Lourdes Colón
12. SS: Ilopango
13. SS: Soyapango
14. SS: Ciudad Delgado
15. SS: Panchimalco
16. SS: San Marcos
17. SS: San Salvador

Puntos calientes nuevos

1. LI: Santa Tecla
2. LI: San Matías

* SO: Salcoatitán pasó a ser valor atípico BAJO-ALTO

Puntos fríos históricos

1. CH: El Carrizal
2. CH: Ojos de agua
3. CH: Concepción Quezaltepeque
4. CH: Chalatenango
5. CH: San Luis del Carmen
6. CH: San Miguel de Mercedes
7. CH: Azacualpa
8. CH: Potonico
9. CH: San Isidro Labrador
10. PA: San Juan Tepezontes
11. PA: Santa María Ostuma
12. MO: Perquín
13. MO: Torola
14. MO: San Isidro
15. MO: San Simón
16. MO: San Francisco Gotera
17. UN: Yucuaiquín

* Jocoaitique pasó a ser valor atípico ALTO-BAJO

Puntos fríos nuevos

1. San Ildefonso
2. Cacaopera
3. Chilanga