08 - Spatial Autocorrelation – Local

J.S. Onésimo Sandoval
Sociology
Saint Louis University

Outline

- Global vs. Local Spatial Autocorrelation
- Local Indicator of Spatial Association (LISA)
- Lab on LISA

Global vs. Local Spatial Autocorrelation

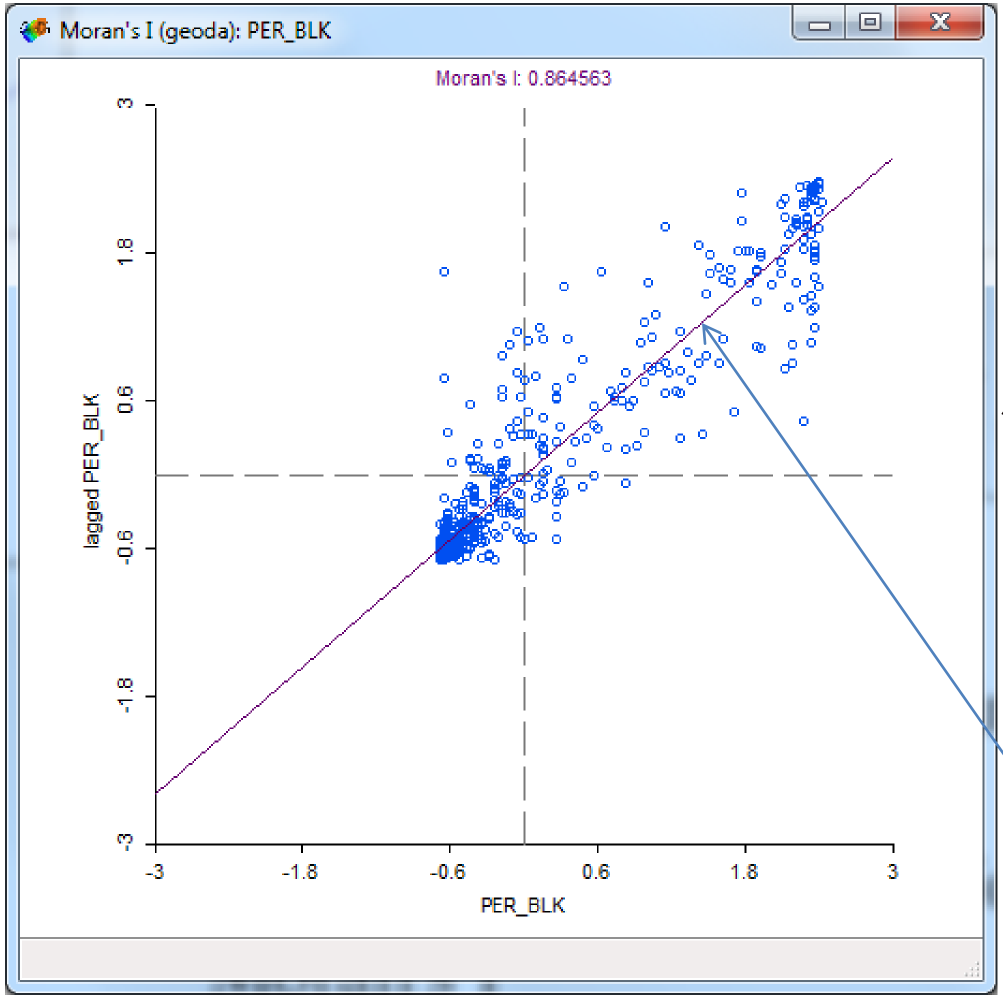
Global	Local
Summarize Data for the whole region	Local disaggregation of global statistics
Single-valued statistic	Multi-valued statistics – different values can occur in different locations – unique values
Non-mappable	Mappable
GIS-unfriendly	GIS-friendly – show how the relationship vary over space – our goal is to map these relationships
Aspatial or spatially limited	Spatial
Emphasize similarities across space	Emphasize differences across space
Search for regularities or "laws" – can be represented by one statistics – testing hypothesis – positivist school of thought	Search for exceptions or local "hot spots" – developing hypothesis from the data – grounded theory – exploratory spatial analyses.
Classic Regression	Spatial Regression or Geographically Weighted Regression

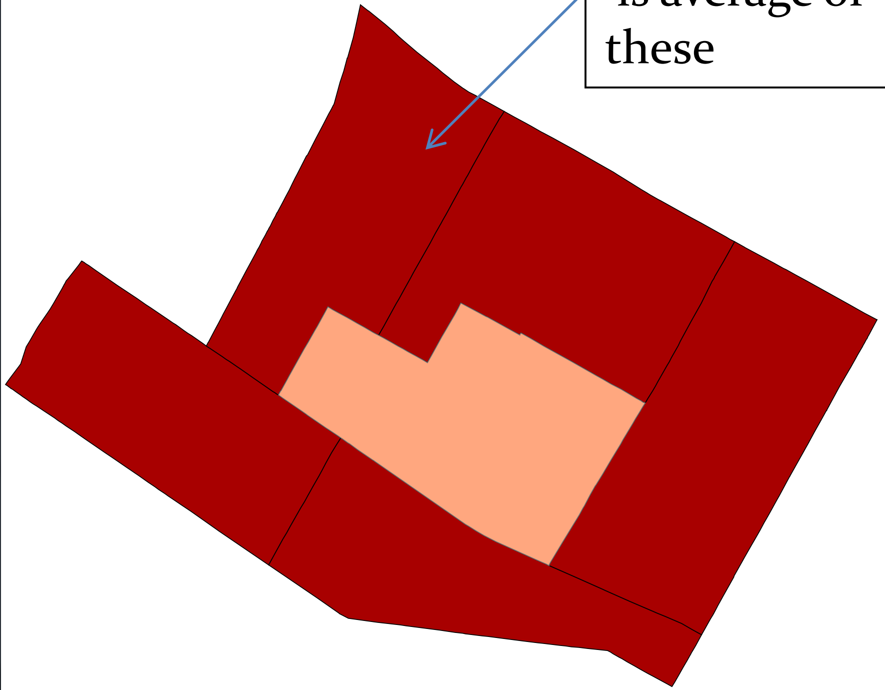
Moran Scatter Plots

Moran's I can be interpreted as the correlation between variable, X, and the "spatial lag" of X formed by averaging all the values of X for the neighboring polygons

We can then draw a scatter diagram between these two variables (in standardized form): **X** and **lag-X** (or W_X)

Example – Percent Black

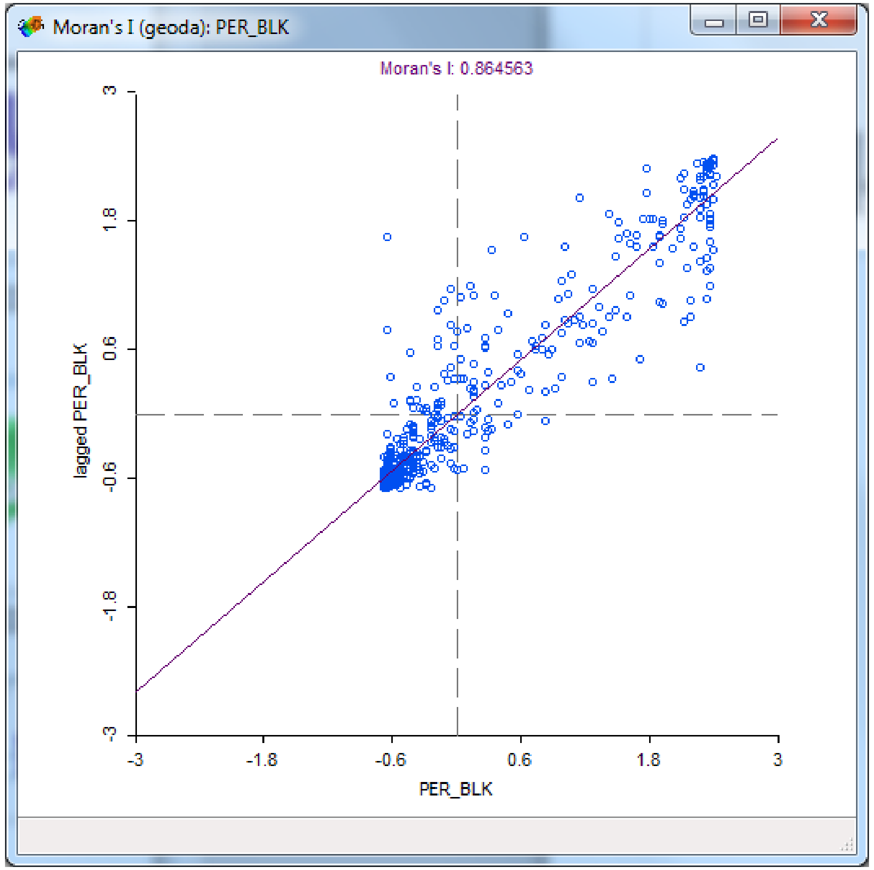


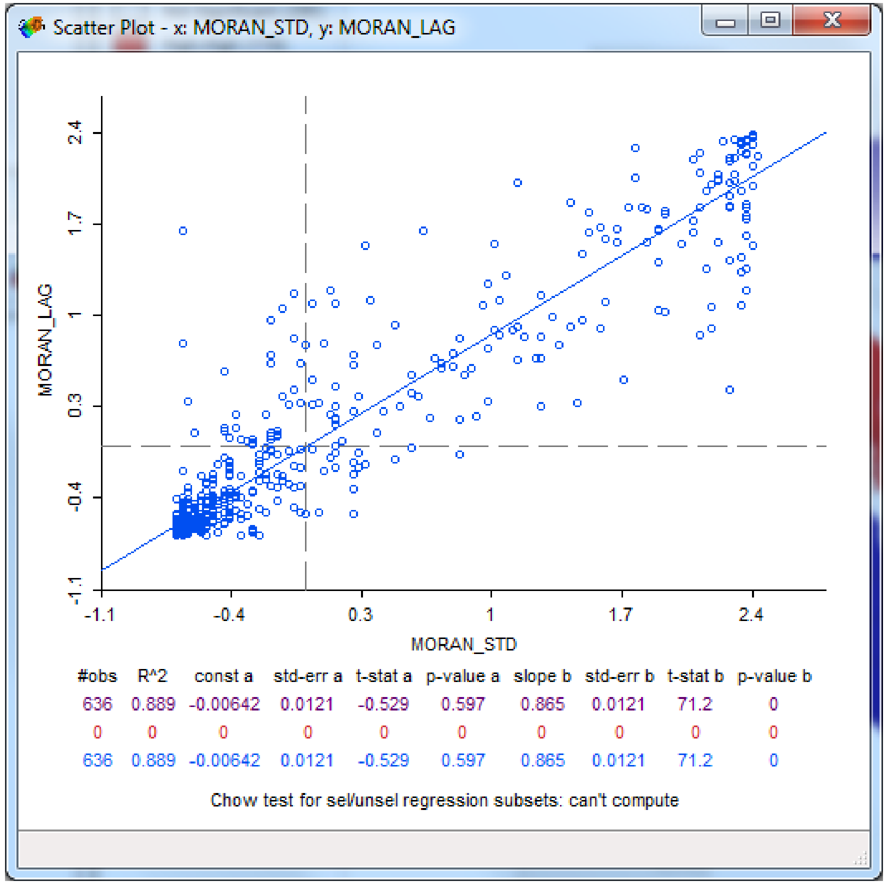


Lag X_i is average of these

Least squares "best fit" line to the points.

The <u>slope</u> of this *regression line* is Moran's I





The Moran's I scatter plot regresses a spatially lagged transformation of a variable (y-axis) on the original standardized variable (x-axis). The values of X are standardized in standard deviation units with a mean of zero

Moran's I for rate-based data

- Moran's I is often calculated for rates, such as crime rates (e.g. number of crimes per 1,000 population) or infant mortality rates (e.g. number of deaths per 1,000 births)

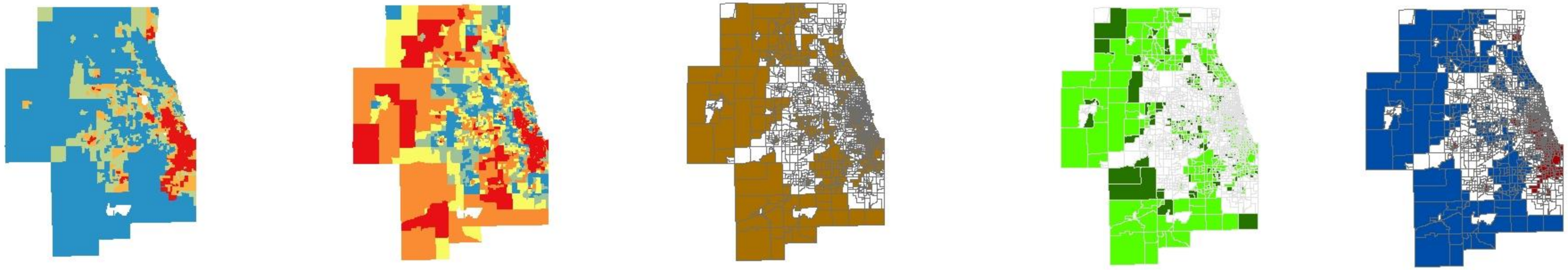
- An adjustment should be made, especially if the denominator in the rate (population or number of births) varies greatly (as it usually does)

- Adjustment is know as the *EB adjust*ment:
  - see Assuncao-Reis *Empirical Bayes Standardization* Statistics in Medicine, 1999

- *GeoDA* software includes an option for this adjustment

Local Indicator of Spatial Association (LISA)

Local Indicator of Spatial Association (LISA) or Cluster or Outlier Analysis



Calculating Anselin's LISA

- The Local Moran statistic for areal unit *i* is:

$$I_i = z_i \sum_{j=1}^{J_i} W_{ij} Z_j$$

Where:

z_i, z_j are standardized values

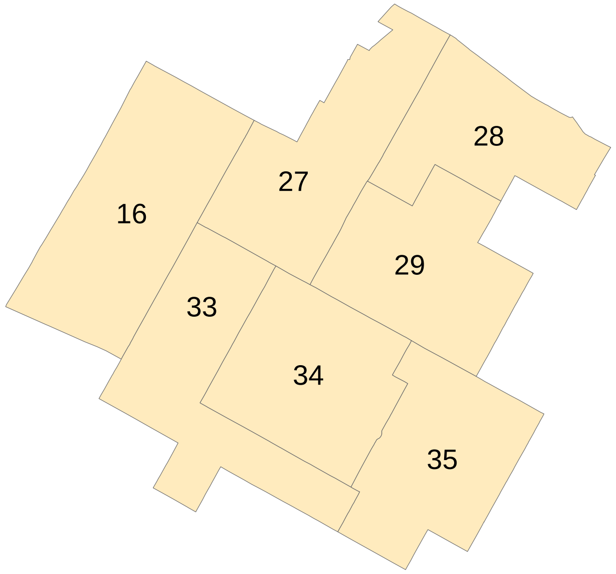
w_{ij} is generally the row-standardized spatial weight matrix

The summation $\sum_j$ is across each row i of the spatial weights matrix.

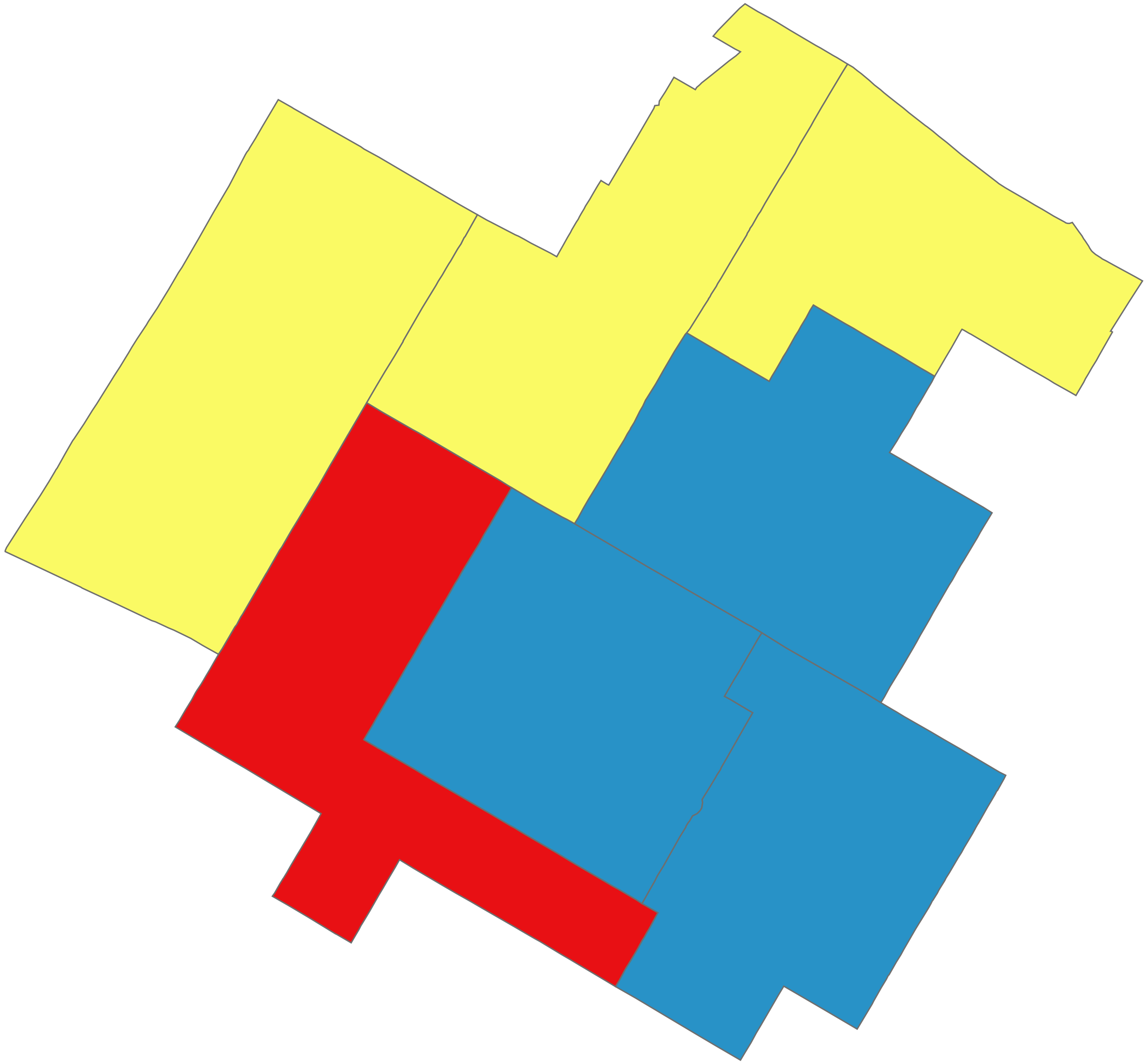
Local Indicators of Spatial Association (LISA)

- We will get a spatial auto association statistic for each polygon

- This statistics is calculated based on neighboring polygons with which it shares a border



Racial Diversity



Example

Diversity

0.119 - 0.147

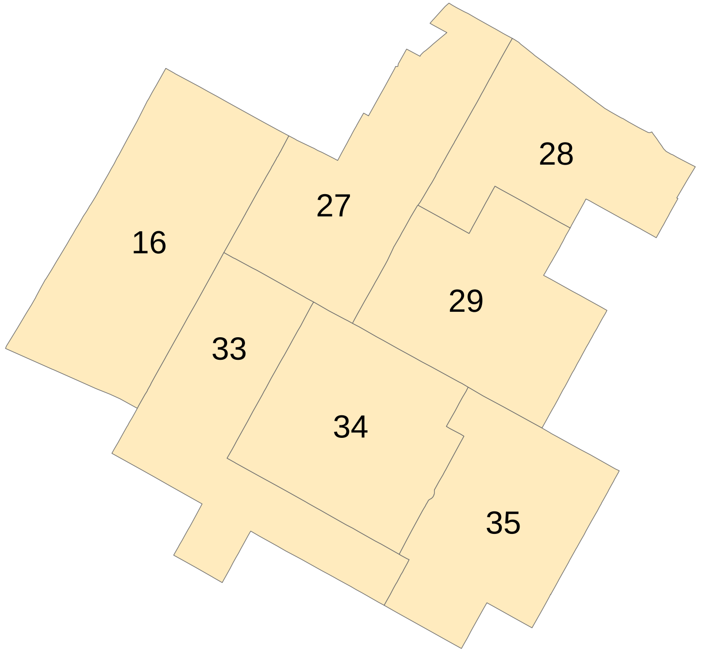
0.148 - 0.199

0.200 - 0.266

Contiguity Matrix

	16	27	28	29	33	34	35	Sum	Neighbors	Diversity
16	**0**	**1**	**0**	**0**	**1**	**0**	**0**	**2**	**27,33**	**0.17**
27	1	0	1	1	1	1	0	5	16,28,29,33,34	0.17
28	0	1	0	1	0	0	0	2	27,29	0.20
29	0	1	1	0	0	1	1	4	27,28,34,35	0.15
33	1	1	0	0	0	1	1	4	16,27,34,35	0.27
34	0	1	0	1	1	0	1	4	27,29,33,35	0.13
35	0	0	0	1	1	1	0	3	29,33,34	0.12

Each row in the contiguity matrix describes the **neighborhood** for that location



Contiguity Matrix and Row Standardized Spatial Weights Matrix

Contiguity Matrix

	16	27	28	29	33	34	35	Sum
16	0	1	0	0	1	0	0	2
27	1	0	1	1	1	1	0	5
28	0	1	0	1	0	0	0	2
29	0	1	1	0	0	1	1	4
33	1	1	0	0	0	1	1	4
34	0	1	0	1	1	0	1	4
35	0	0	0	1	1	1	0	3

Row Standardized Contiguity Matrix

	16	27	28	29	33	34	35	Sum
16	0.00	0.50 $\frac{1}{2}$	0.00	0.00	0.50 $\frac{1}{2}$	0.00	0.00	1
27	0.20	0.00	0.20	0.20	0.20	0.20	0.00	1
28	0.00	0.50	0.00	0.50	0.00	0.00	0.00	1
29	0.00	0.25	0.25	0.00	0.00	0.25	0.25	1
33	0.25	0.25	0.00	0.00	0.00	0.25	0.25	1
34	0.00	0.25	0.00	0.25	0.25	0.00	0.25	1
35	0.00	0.00	0.00	0.33	0.33	0.33	0.00	1

Calculating standardized z scores

Deviations from Mean and z scores. Mean and Standard Deviation

	X	$X - \overline{X}$	z
16	0.174	0.001	0.016
27	0.174	0.001	0.020
28	0.199	0.026	0.532
29	0.147	-0.026	-0.532
33	0.266	0.093	1.884
34	0.133	-0.040	-0.818
35	0.119	-0.054	-1.102
Sum	1.213		
Mean	0.173		
Variance	0.015		
SD	0.049		

$$z_i = \frac{x_i - \overline{x}}{SD_x}$$

$$z_i = \frac{.001}{.049} = .016$$

$$I_i = z_i \sum_{j=1}^{J_i} W_{ij} Z_j$$

Row Standardized Contiguity Matrix

	16	27	28	29	33	34	35	Sum
16	0.00	0.50	0.00	0.00	0.50	0.00	0.00	1
27	0.20	0.00	0.20	0.20	0.20	0.20	0.00	1
28	0.00	0.50	0.00	0.50	0.00	0.00	0.00	1
29	0.00	0.25	0.25	0.00	0.00	0.25	0.25	1
33	0.25	0.25	0.00	0.00	0.00	0.25	0.25	1
34	0.00	0.25	0.00	0.25	0.25	0.00	0.25	1
35	0.00	0.00	0.00	0.33	0.33	0.33	0.00	1

$$w_{ij}$$

Z_j is the standardized z score

z scores for row county and its potential neighbors

		16	27	28	29	33	34	35
Zi		0.016	0.020	0.532	(0.532)	1.884	(0.818)	(1.102)
16	0.016	0.016	0.020	0.532	(0.532)	1.884	(0.818)	(1.102)
27	0.020	0.016	0.020	0.532	(0.532)	1.884	(0.818)	(1.102)
28	0.532	0.016	0.020	0.532	(0.532)	1.884	(0.818)	(1.102)
29	(0.532)	0.016	0.020	0.532	(0.532)	1.884	(0.818)	(1.102)
33	1.884	0.016	0.020	0.532	(0.532)	1.884	(0.818)	(1.102)
34	(0.818)	0.016	0.020	0.532	(0.532)	1.884	(0.818)	(1.102)
35	(1.102)	0.016	0.020	0.532	(0.532)	1.884	(0.818)	(1.102)

$$Z_j$$

$$I_i = z_i \sum_{j=1}^{J_i} W_{ij} Z_j$$

Row Standardized Contiguity Matrix

	16	27	28	29	33	34	35	Sum
16	0.00	0.50	0.00	0.00	0.50	0.00	0.00	1
27	0.20	0.00	0.20	0.20	0.20	0.20	0.00	1
28	0.00	0.50	0.00	0.50	0.00	0.00	0.00	1
29	0.00	0.25	0.25	0.00	0.00	0.25	0.25	1
33	0.25	0.25	0.00	0.00	0.00	0.25	0.25	1
34	0.00	0.25	0.00	0.25	0.25	0.00	0.25	1
35	0.00	0.00	0.00	0.33	0.33	0.33	0.00	1

$$w_{ij}$$

Spatial Weight Matrix Multiplied by Z-Score Matrix (cell by cell multiplication)

		16	27	28	29	33	34	35	$\sum_j w_{ij} z_j$	LISA
		0.016	0.020	0.532	-0.532	1.884	-0.818	-1.102		
16	0.016	-	0.010	-	-	0.942	-	-	0.952	**0.015**
27	0.020	0.003	-	0.106	(0.106)	0.377	(0.164)	-	0.216	**0.004**
28	0.532	-	0.010	-	(0.266)	-	-	-	-0.256	**-0.136**
29	-0.532	-	0.005	0.133	-	-	(0.204)	(0.276)	-0.342	**0.182**
33	1.884	0.004	0.005	-	-	-	(0.204)	(0.276)	-0.471	**-0.887**
34	-0.818	-	0.005	-	(0.133)	0.471	-	(0.276)	0.068	**-0.055**
35	-1.102	-	-	-	(0.177)	0.628	(0.273)	-	0.178	**-0.196**

$$Z_j * W_{ij}$$

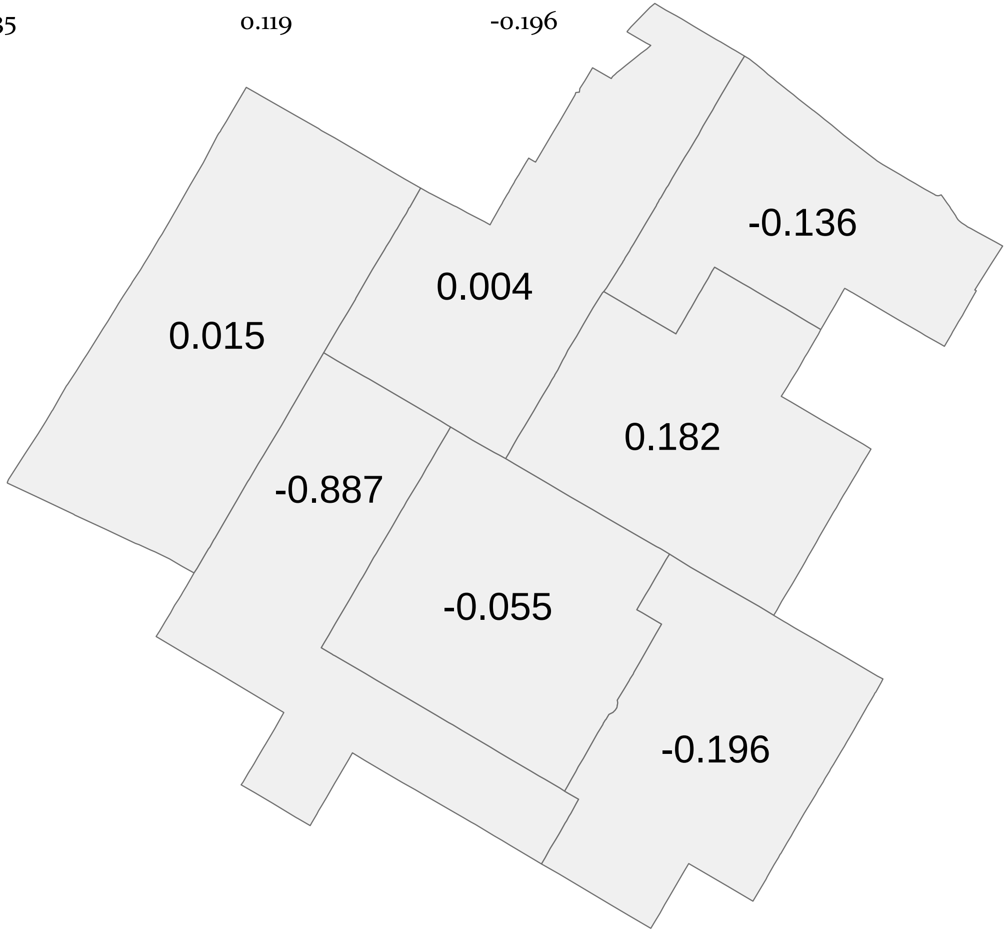
$$.020 * .50 = .010$$

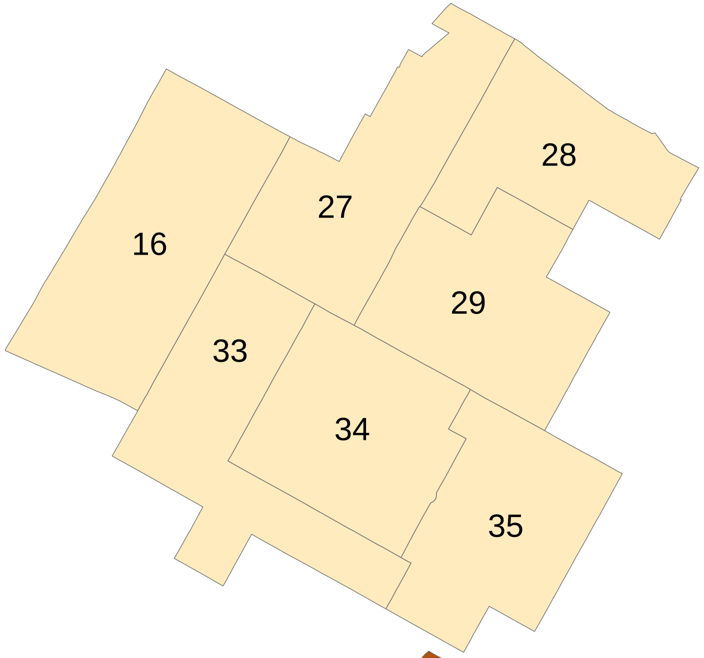
$$I_i = z_i \sum_{j=1}^{J_i} W_{ij} Z_j$$

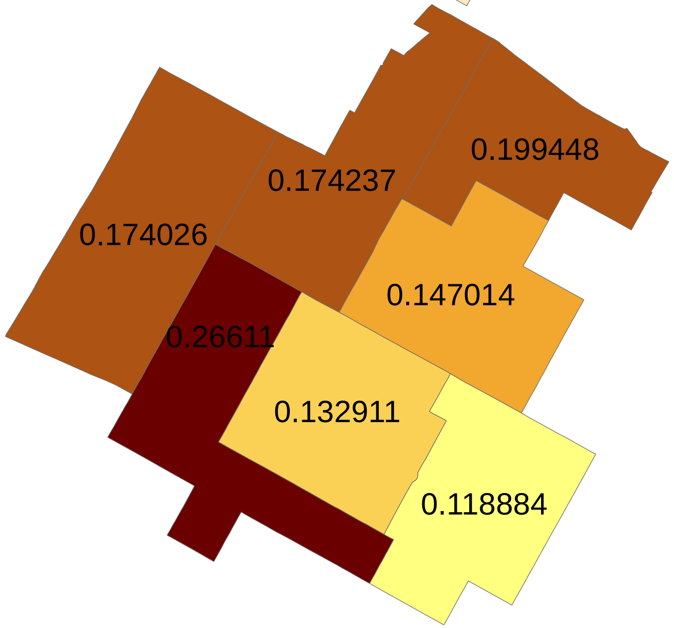
$$.015 = .016 * .952$$

Moran's I = -.178864

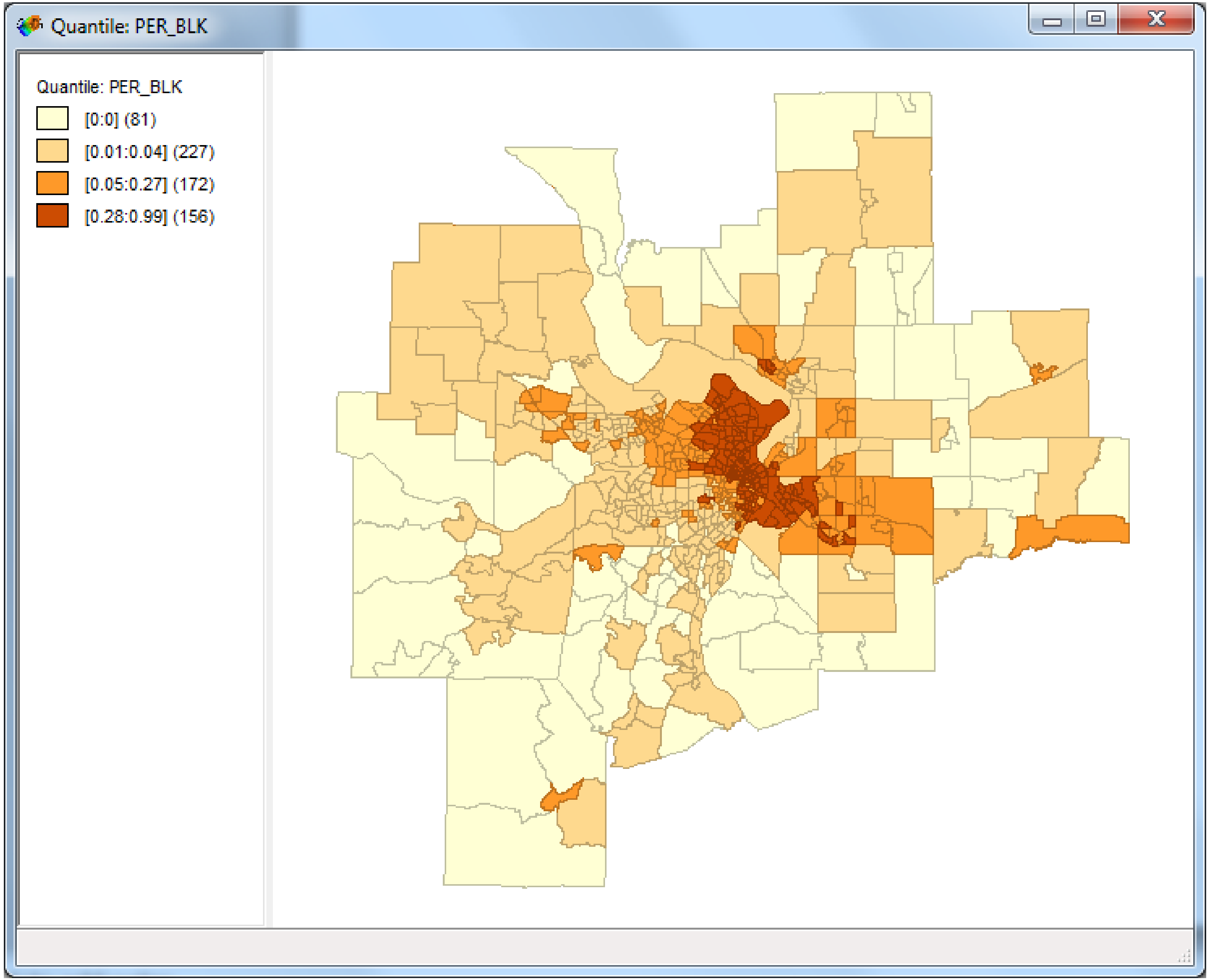
ID	Diversity	LISA
16	0.174	0.015
27	0.174	0.004
28	0.199	-0.136
29	0.147	0.182
33	0.266	-0.887
34	0.133	-0.055
35	0.119	-0.196



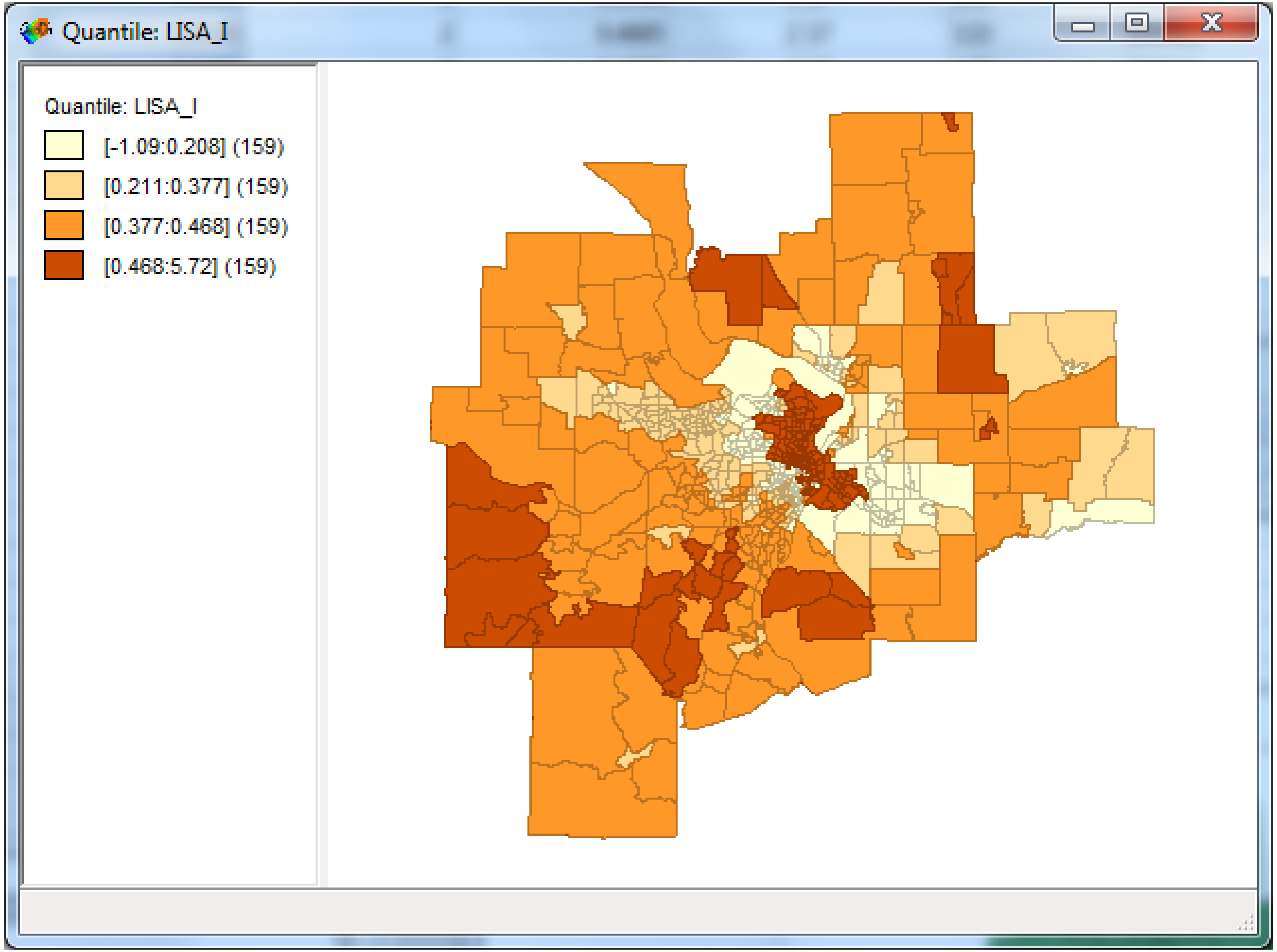




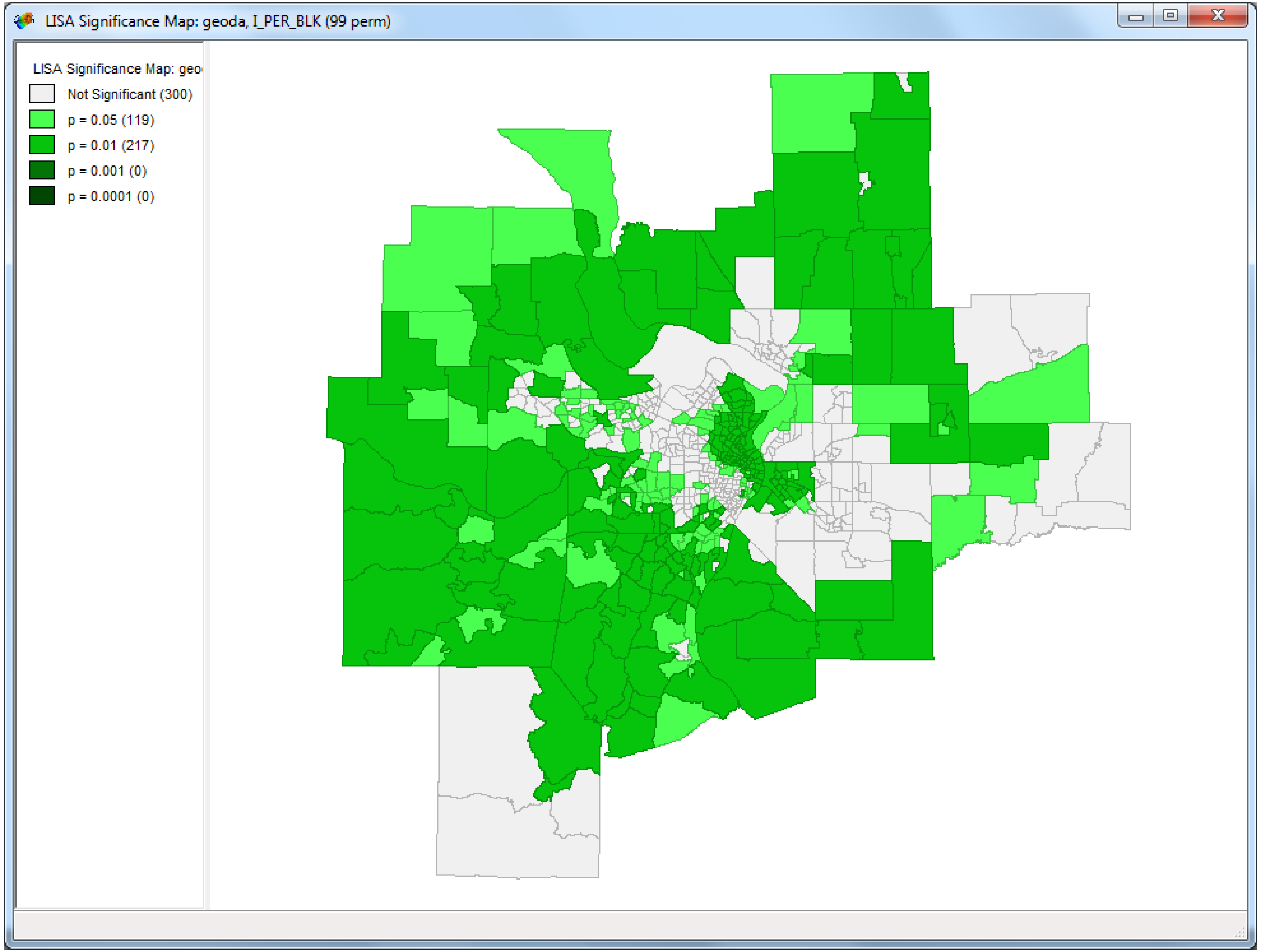
Percent Black



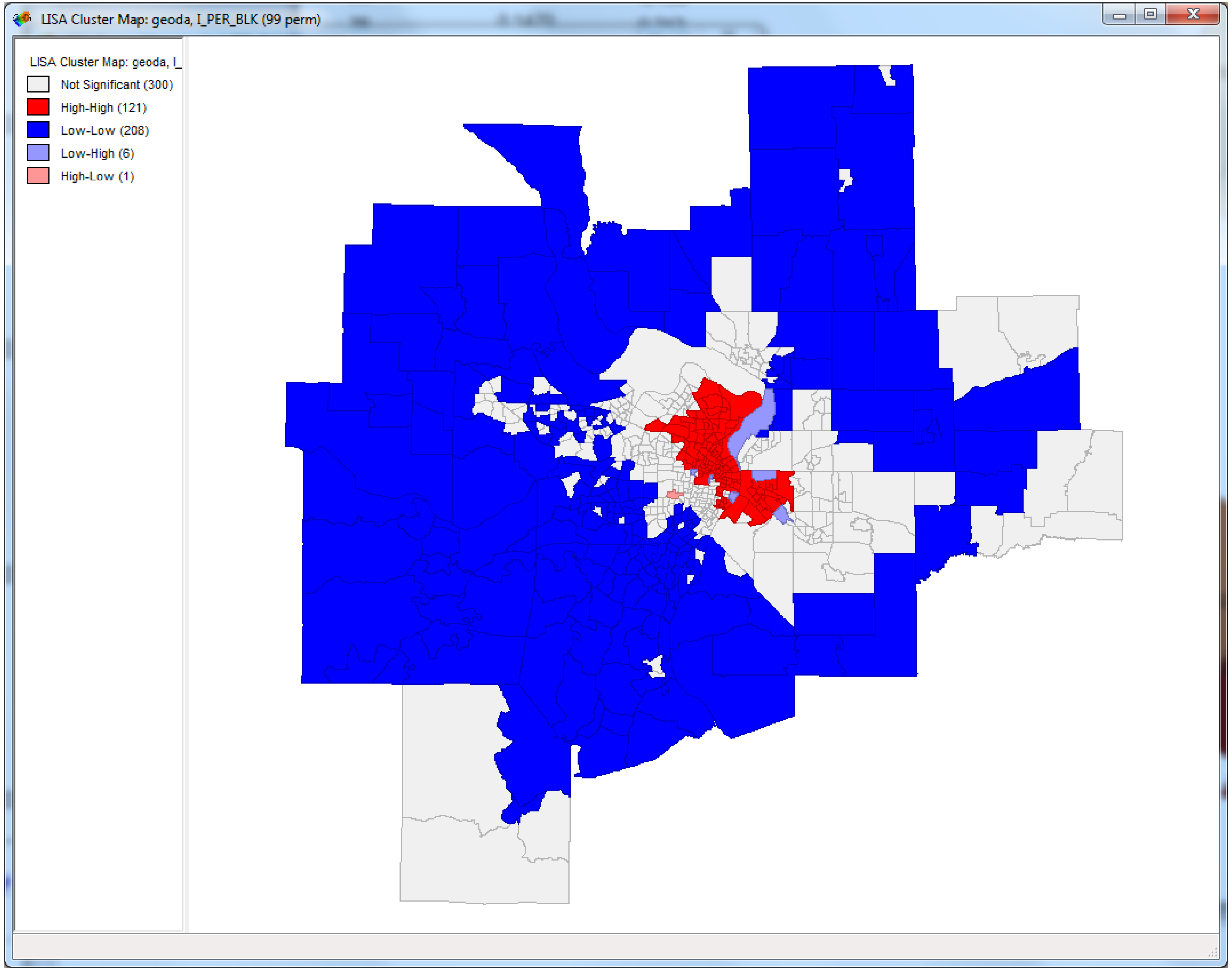
LISA Statistic Map for Percent Black



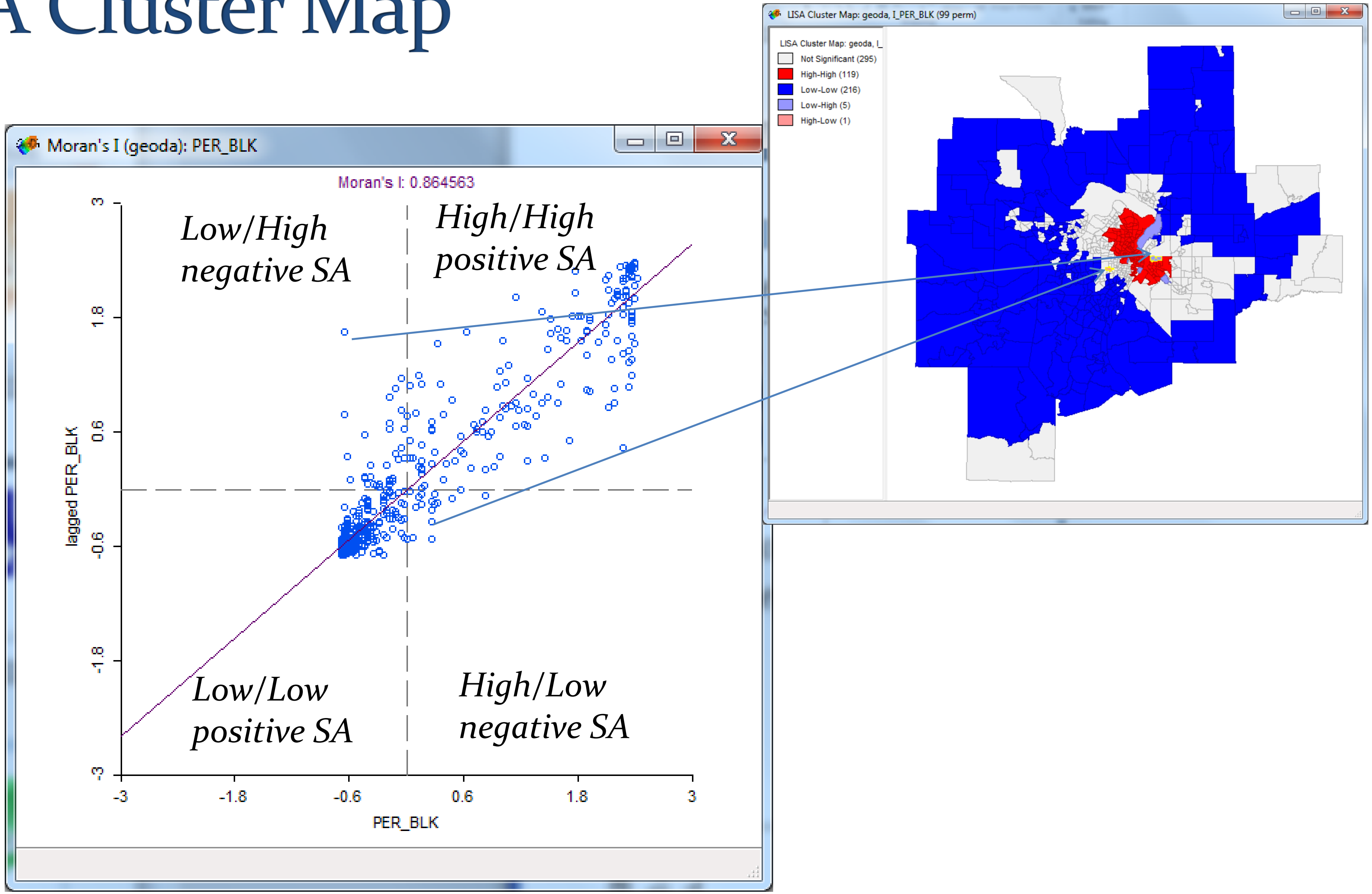
LISA Significance Map for Percent Black



LISA Cluster Map for Percent Black

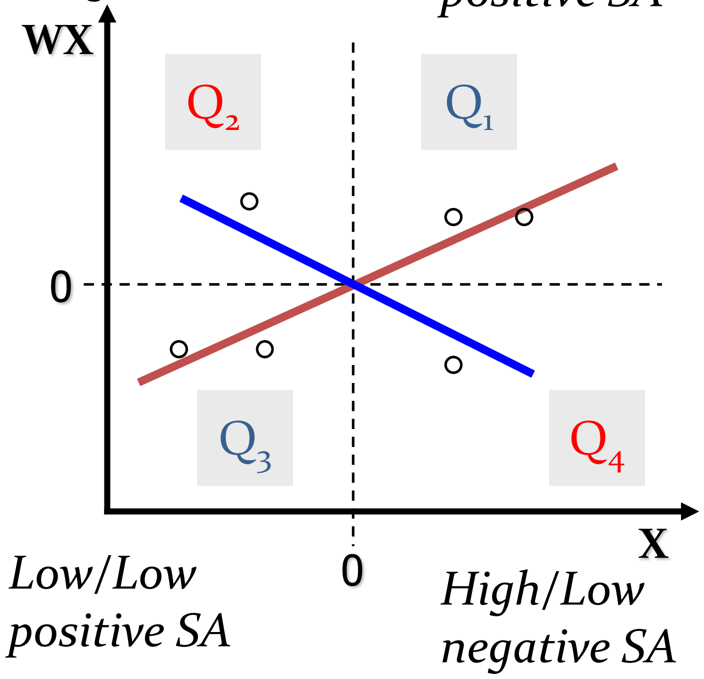


LISA Cluster Map



Quadrants of Moran Scatterplot

Each quadrant corresponds to one of the four different types of spatial association (SA)



Locations of positive spatial association ("I'm similar to my neighbors").

Q_1 (values [+], nearby values [+]): **H-H**

Q_3 (values [-], nearby values [-]): **L-L**

Locations of negative spatial association ("I'm different from my neighbors").

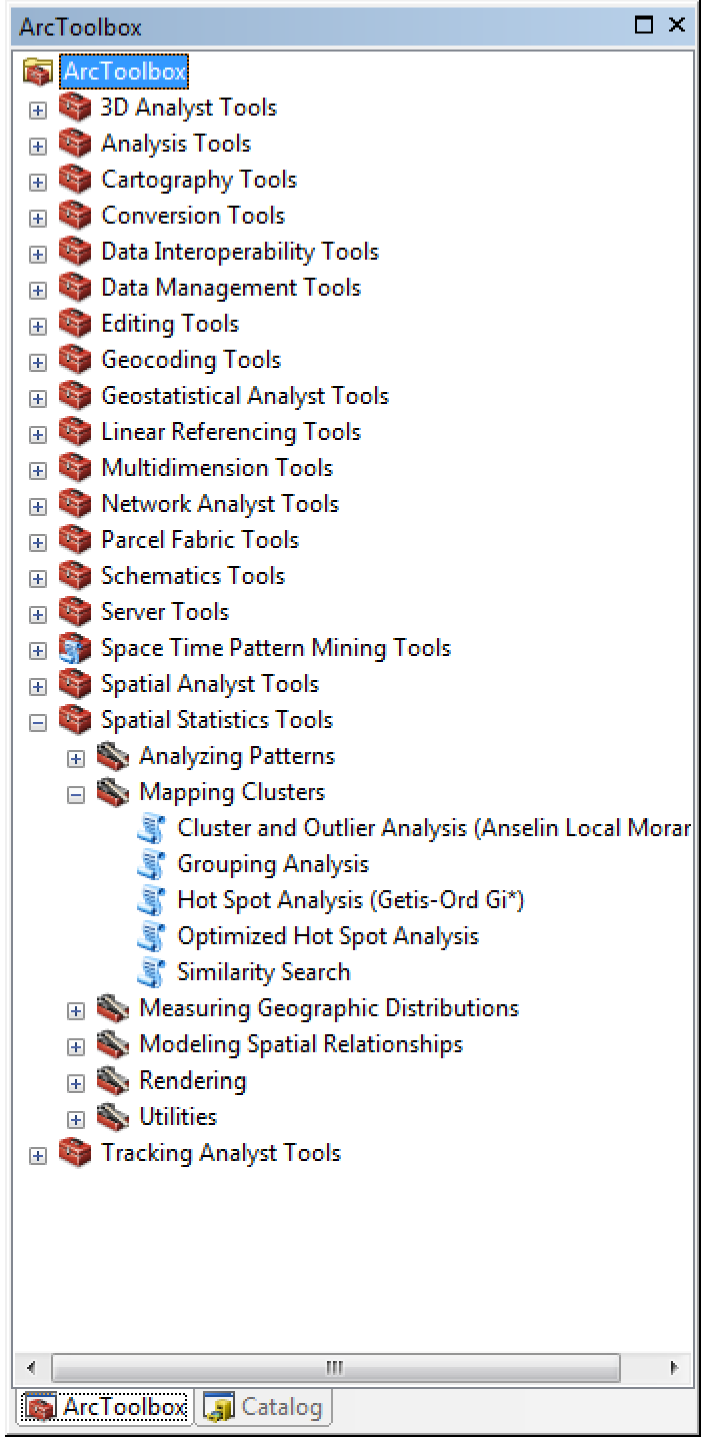
Q_2 (values [-], nearby values [+]): **L-H**

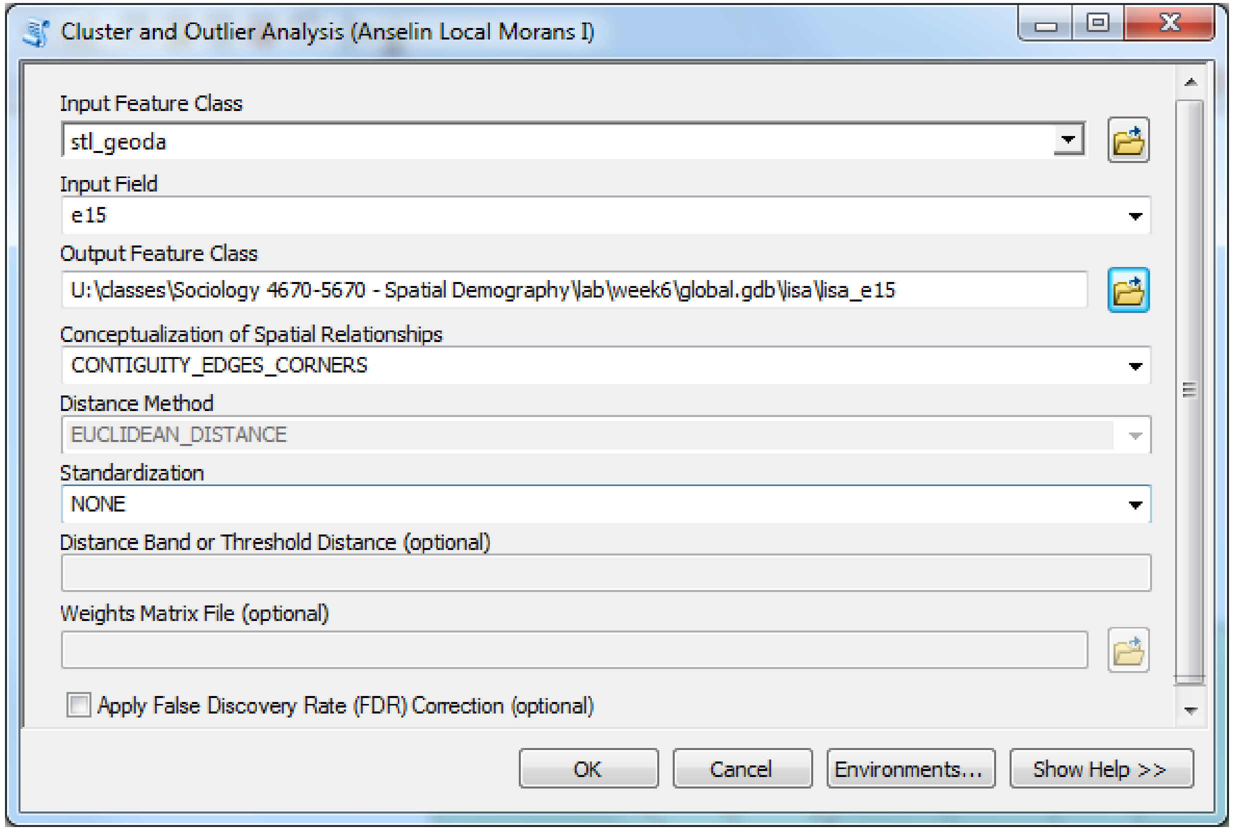
Q_4 (values [+], nearby values [-]): **H-L**

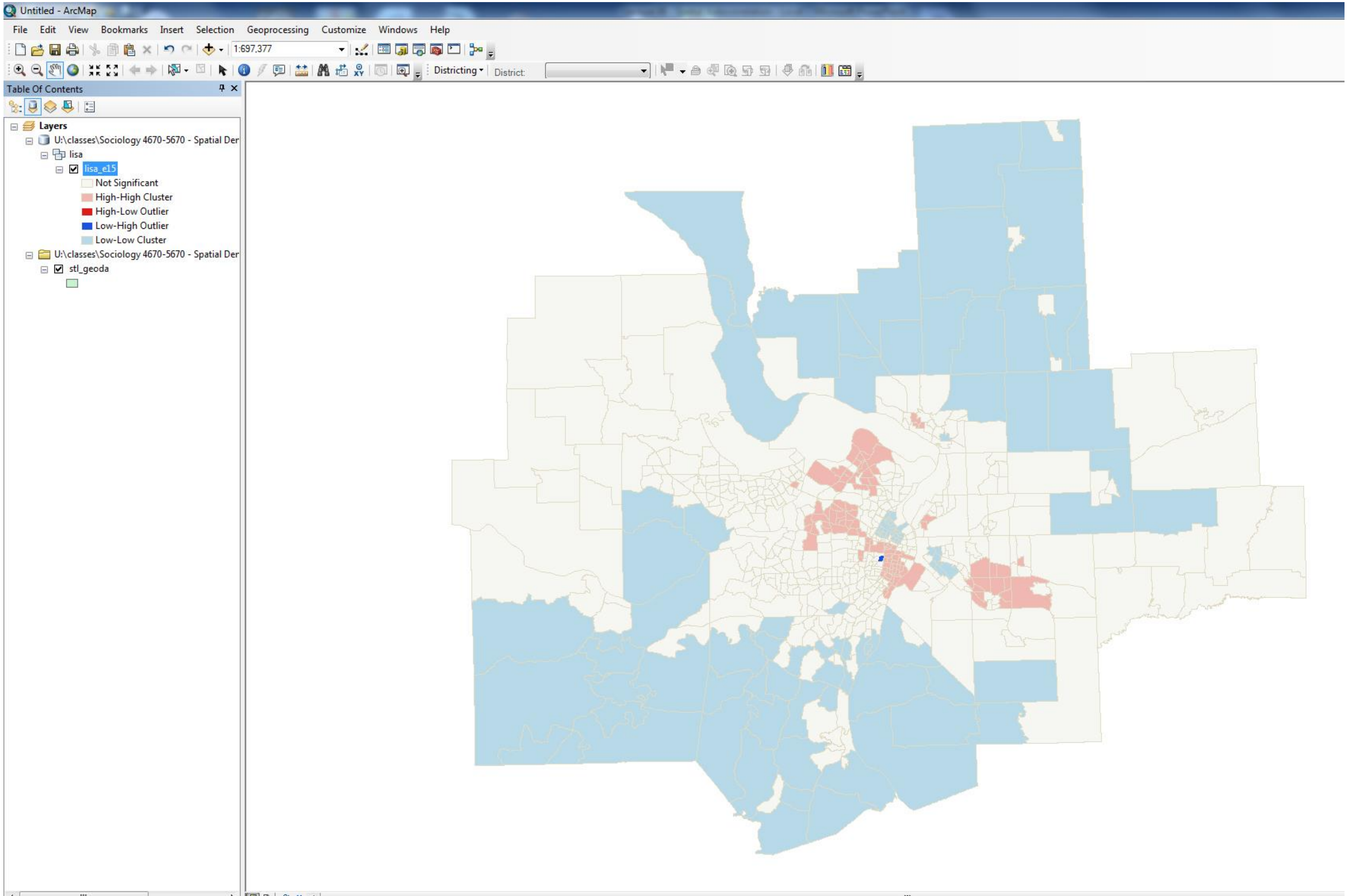
Interpreting the Legend

Color	X	Y	Correlation	Quadrant
	High	High	Positive	Q1
	Low	Low	Positive	Q3
	Low	High	Negative	Q2
	High	Low	Negative	Q4

Lab - ArcMap


ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Space Time Pattern Mining Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Cluster and Outlier Analysis (Anselin Local Morar
Grouping Analysis
Hot Spot Analysis (Getis-Ord Gi*)
Optimized Hot Spot Analysis
Similarity Search
Measuring Geographic Distributions
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools
ArcToolbox
Catalog


Cluster and Outlier Analysis (Anselin Local Morans I)
Input Feature Class
stl_geoda
Input Field
e15
Output Feature Class
U:\classes\Sociology 4670-5670 - Spatial Demography\lab\week6\global.gdb\lisa\lisa_e15
Conceptualization of Spatial Relationships
CONTIGUITY_EDGES_CORNERS
Distance Method
EUCLIDEAN_DISTANCE
Standardization
NONE
Distance Band or Threshold Distance (optional)
Weights Matrix File (optional)
Apply False Discovery Rate (FDR) Correction (optional)
OK
Cancel
Environments...
Show Help >>


Untitled - ArcMap
File
Edit
View
Bookmarks
Insert
Selection
Geoprocessing
Customize
Windows
Help
1:697,377
Districting
District:
Table Of Contents
Layers
U:\classes\Sociology 4670-5670 - Spatial Der
lisa
lisa_e15
Not Significant
High-High Cluster
High-Low Outlier
Low-High Outlier
Low-Low Cluster
U:\classes\Sociology 4670-5670 - Spatial Der
stl_geoda

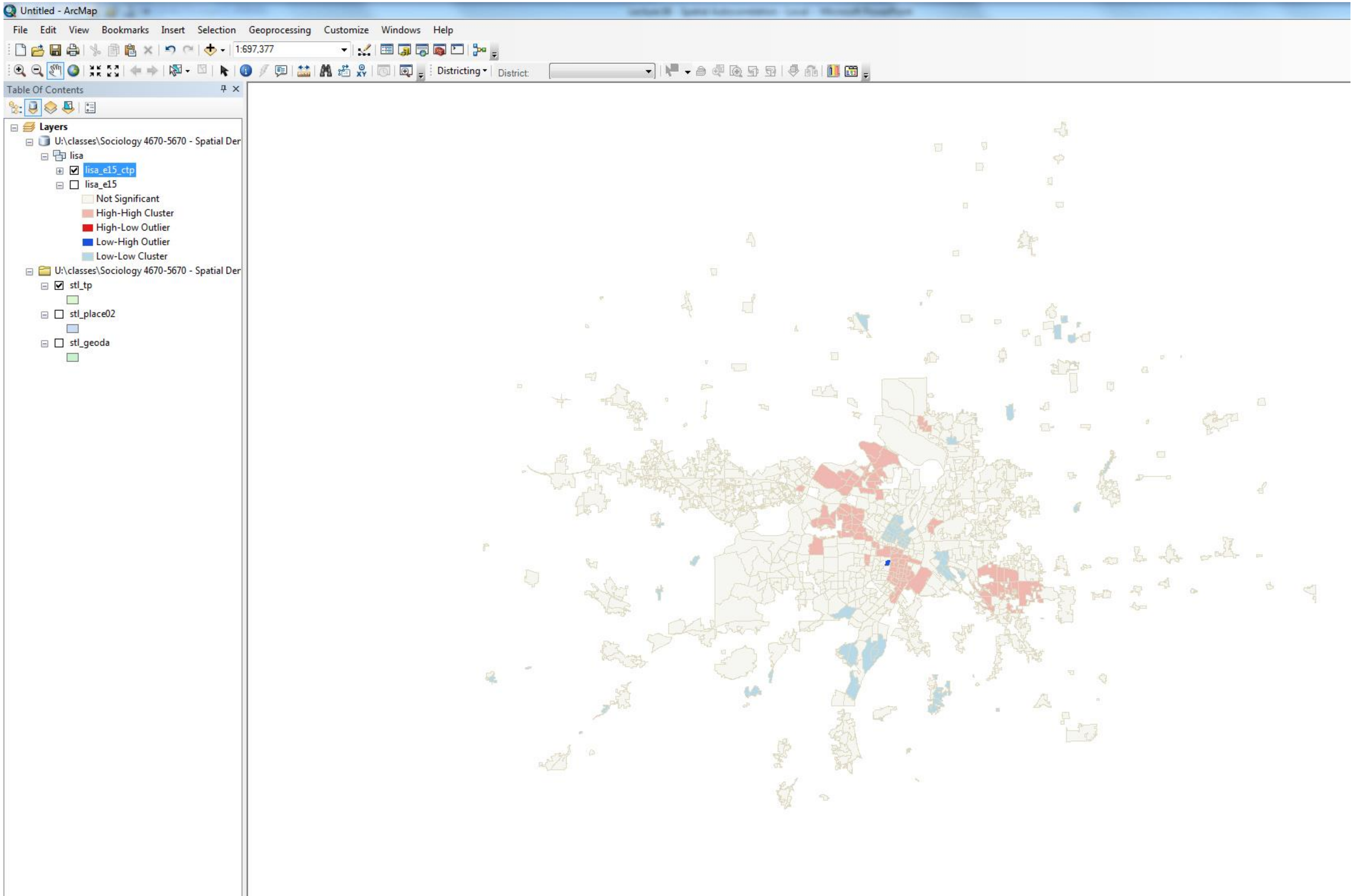
Table

lisa_e15

OBJECTID *	Shape *	SOURCE_ID	e15	Shape_Length	Shape_Area	LMiIndex	LMiZScore	LMiPValue	COType
1	Polygon	0	0.174953	8761.829111	4272870.555874	3.088376	1.271282	0.203628	
2	Polygon	1	0.108673	74670.041987	302972299.093549	9.902957	3.529558	0.000416	LL
3	Polygon	2	0.179487	24141.052946	12239863.036699	4.051972	1.446922	0.147919	
4	Polygon	3	0.245181	32447.102859	64237947.601333	-0.658008	-0.229578	0.81842	
5	Polygon	4	0.461556	24477.702542	23484806.847737	0.487526	0.168667	0.866059	
6	Polygon	5	0.207814	24627.48159	26085522.261326	2.499317	1.257261	0.208659	
7	Polygon	6	0.627683	17820.58705	8756896.988813	3.775546	1.273	0.203018	
8	Polygon	7	0.523876	31616.148359	43644601.078179	-0.865606	-0.303471	0.761531	
9	Polygon	8	0.139083	11693.467599	5361487.659518	0.776926	0.299729	0.764384	
10	Polygon	9	0.144203	23696.75189	18544701.199047	0.570597	0.259931	0.794917	
11	Polygon	10	0.087612	12303.791565	7768646.902752	3.618169	1.628694	0.103378	
12	Polygon	11	0.689025	16118.31168	6507532.898065	3.600576	1.373308	0.169657	
13	Polygon	12	0.321377	32872.735687	32319247.957447	-0.134725	-0.046889	0.962602	
14	Polygon	13	0.211058	31580.115661	55755835.904803	2.040455	0.780135	0.435311	
15	Polygon	14	0.186542	34244.099006	70283882.130466	3.623461	1.22192	0.221738	
16	Polygon	15	0.416704	15449.220454	10192525.952367	-0.336525	-0.134079	0.89334	
17	Polygon	16	0.243886	86178.744847	189612148.34056	2.239576	0.922989	0.356013	
18	Polygon	17	0.316824	89695.424095	233029641.088772	0.617587	0.281036	0.778683	
19	Polygon	18	0.064575	47756.872188	80113855.703524	1.046672	0.608725	0.542707	
20	Polygon	19	0.193177	78408.528966	209719578.945189	-0.306623	-0.214905	0.829842	
21	Polygon	20	0.487294	23544.144178	32942157.874046	4.027731	1.357701	0.174559	
22	Polygon	21	0.68924	12314.771481	9319333.888368	9.653406	3.965151	0.000073	HH
23	Polygon	22	0.52571	14970.173631	7640363.668159	1.738428	1.233812	0.217273	
24	Polygon	23	0.289171	66109.363948	123697516.702869	0.977571	0.376017	0.706905	
25	Polygon	24	0.49822	11409.036008	7424118.259007	2.765105	1.245555	0.212928	
26	Polygon	25	0.255239	28651.064246	41020510.676969	0.988225	0.336834	0.736242	
27	Polygon	26	0.579642	27333.52388	24644214.538218	3.800773	1.281474	0.200027	
28	Polygon	27	0.236817	23834.414319	31950926.933579	-1.534848	-0.541686	0.588035	
29	Polygon	28	0.207356	12394.017398	7468588.225768	2.070169	1.201206	0.229671	
30	Polygon	29	0.504297	14897.118212	10037997.064972	6.077651	2.315116	0.020607	HH
31	Polygon	30	0.193144	17709.929138	15493579.250651	1.166097	0.828372	0.40746	
32	Polygon	31	0.142983	61047.6532	109077422.039058	8.129512	2.597643	0.009387	LL
33	Polygon	32	0.775969	11099.899399	4348858.221391	6.529715	2.936363	0.003321	HH
34	Polygon	33	0.78033	9174.094142	4635101.658403	10.477149	4.709283	0.000002	HH
35	Polygon	34	0.579359	13890.986825	5807641.526327	5.540215	2.27736	0.022765	HH
36	Polygon	35	0.142614	91216.352379	288845706.646605	4.828585	2.17233	0.029831	LL
37	Polygon	36	0.115214	42481.20886	69313941.650369	3.568577	2.068604	0.038583	LL
38	Polygon	37	0.352831	18175.537564	10248292.516186	-0.027728	-0.017335	0.986169	
39	Polygon	38	0.042372	37431.546521	61974707.567582	9.864435	4.051743	0.000051	LL
40	Polygon	39	0.577095	18429.606293	9988372.628714	3.363702	1.384258	0.166279	
41	Polygon	40	0.075856	8284.968534	3123045.83976	4.934181	2.219757	0.026435	LL
42	Polygon	41	0.542497	14311.158489	12331252.667634	8.303984	2.960409	0.003072	HH
43	Polygon	42	0.601658	10440.791256	6151777.086163	6.751335	3.0359	0.002398	HH
44	Polygon	43	0.583403	10069.861854	5529545.076874	5.677631	2.553664	0.01066	HH
45	Polygon	44	0.503729	5981.411372	1946783.614727	2.859298	1.28786	0.197795	
46	Polygon	45	0.491212	11378.586478	5208019.523133	5.046962	1.923238	0.05445	

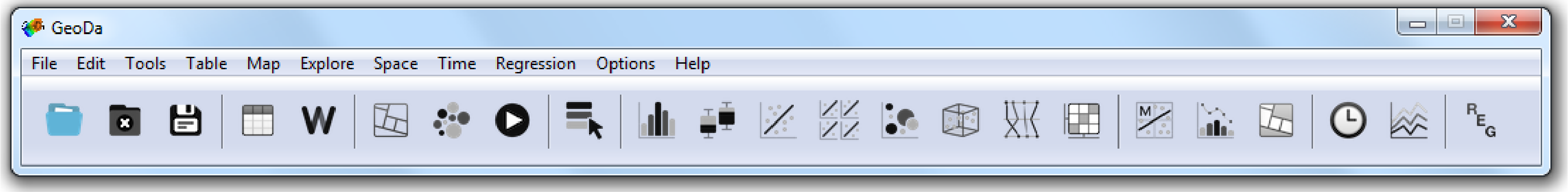
1 (0 out of 615 Selected)

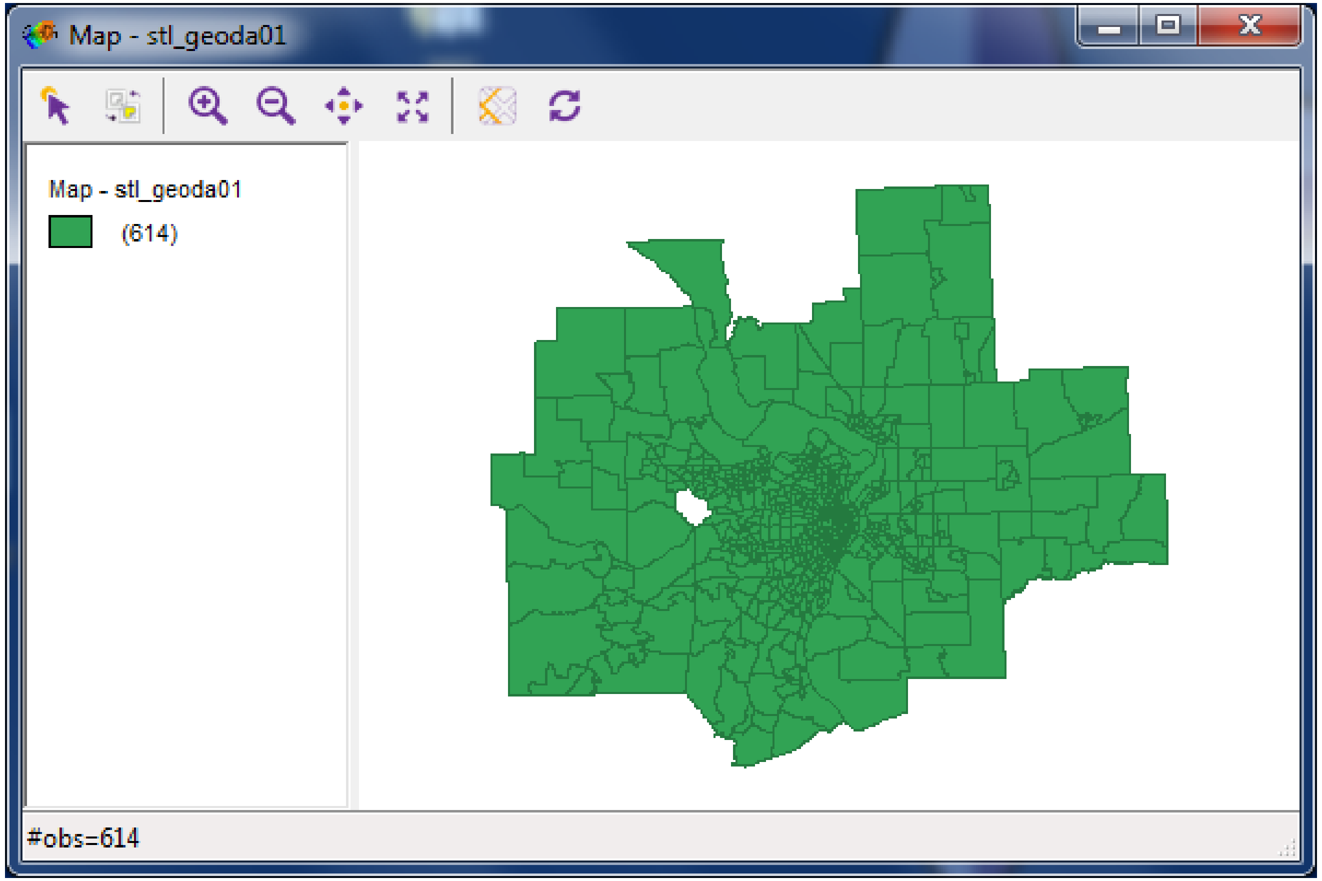
lisa_e15


Untitled - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
1:697,377
Districting
District:
Table Of Contents
Layers
U:\classes\Sociology 4670-5670 - Spatial Der
lisa
lisa_e15_ctp
lisa_e15
Not Significant
High-High Cluster
High-Low Outlier
Low-High Outlier
Low-Low Cluster
U:\classes\Sociology 4670-5670 - Spatial Der
stl_tp
stl_place02
stl_geoda

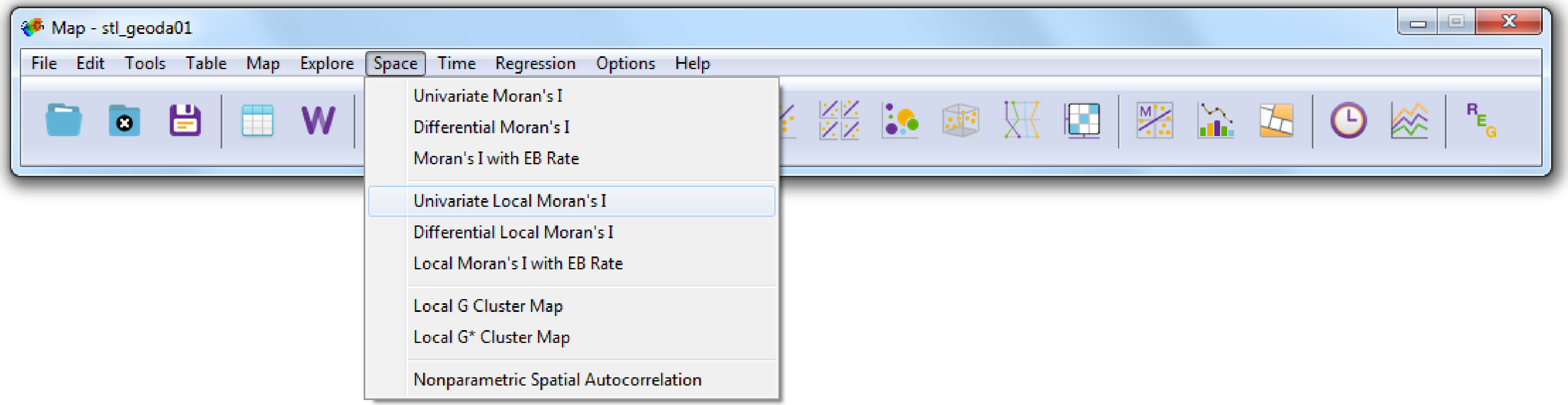
Lab on LISA

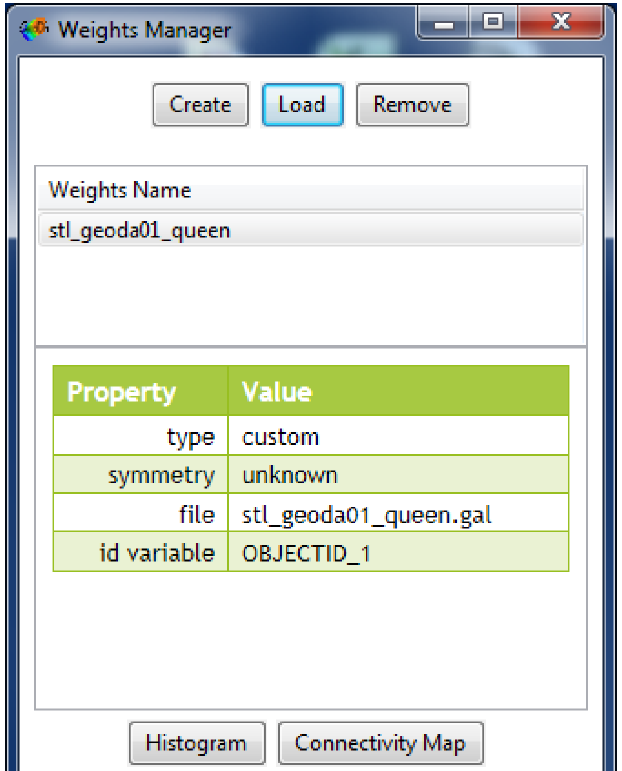
Step 1 – Open up shapefile

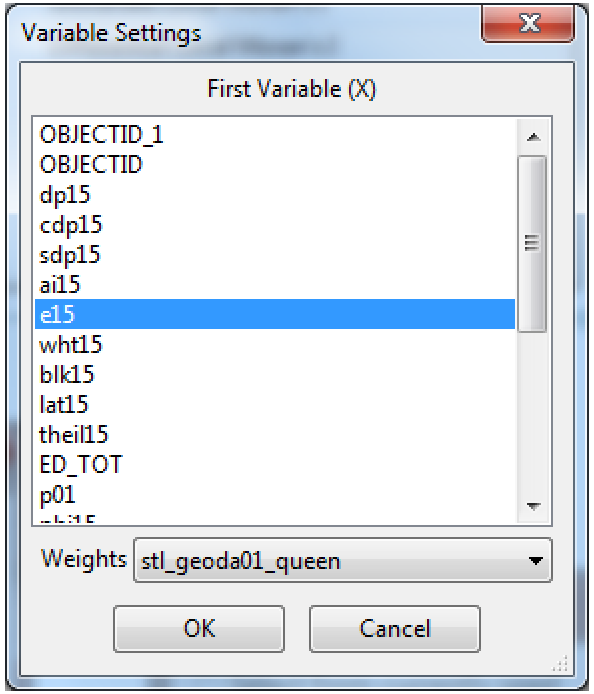


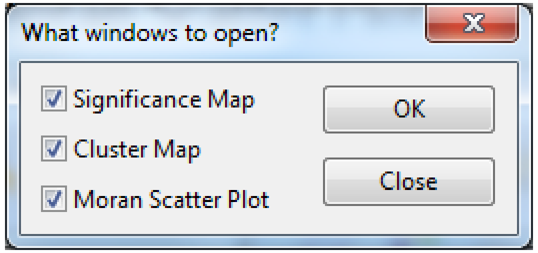


Step 2 – Select the Space Tab Menu and select the variable, select the weights and select significant map, cluster map, and Moran Scatter Plot

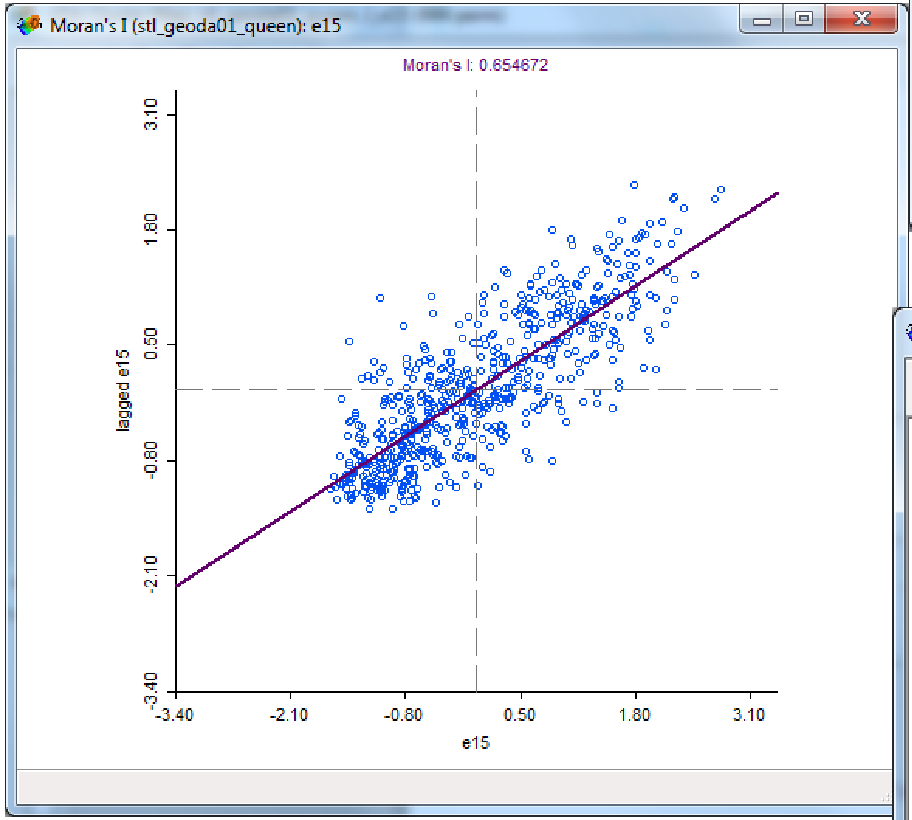


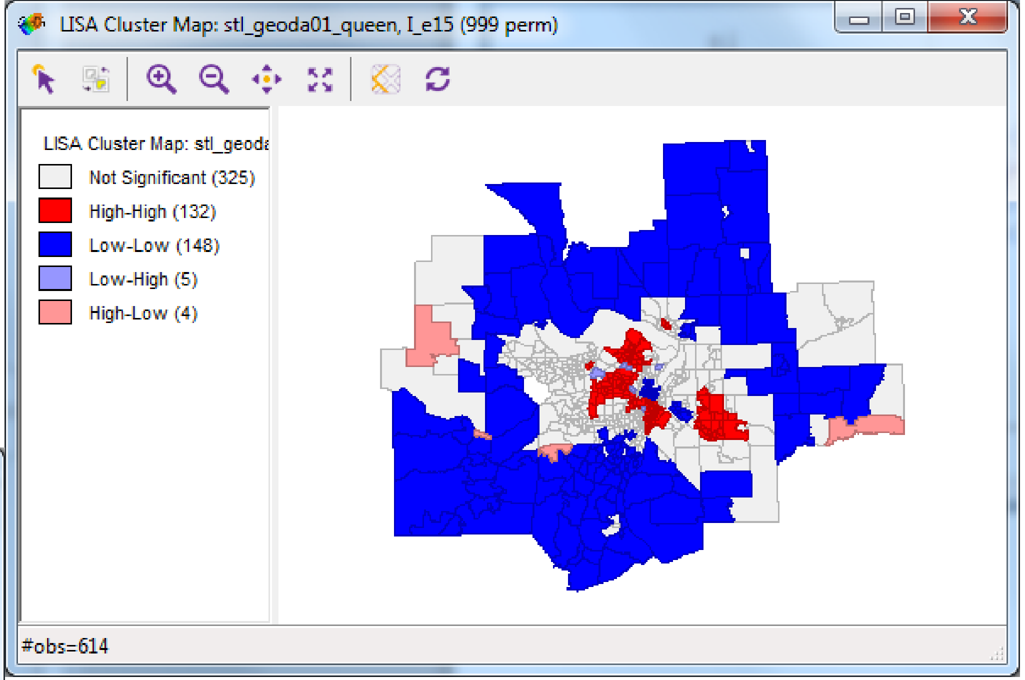


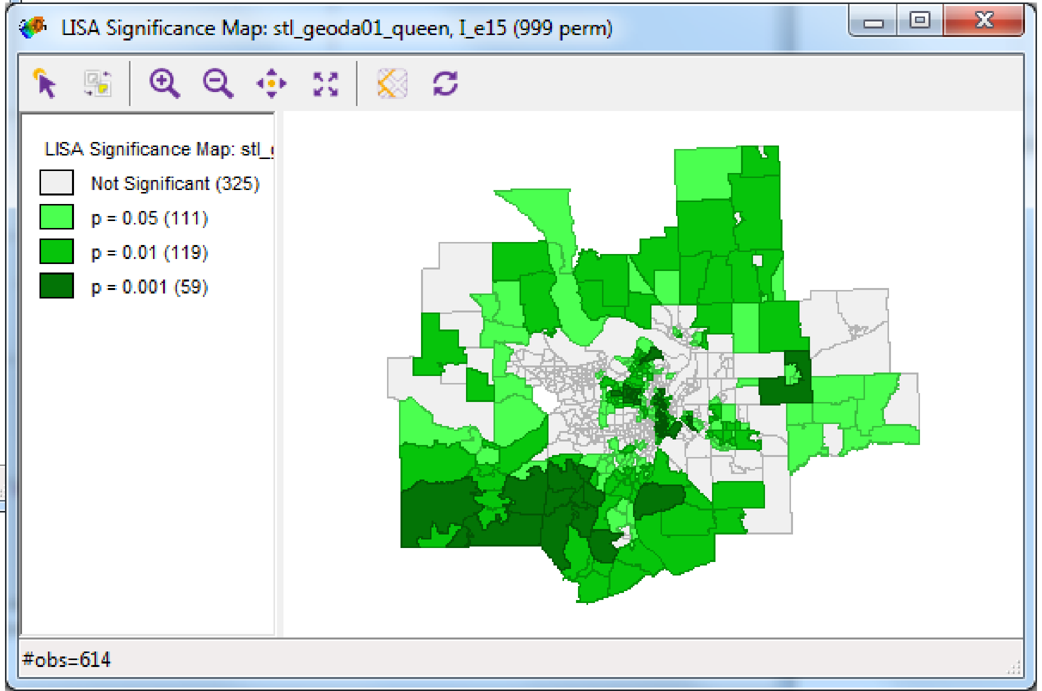




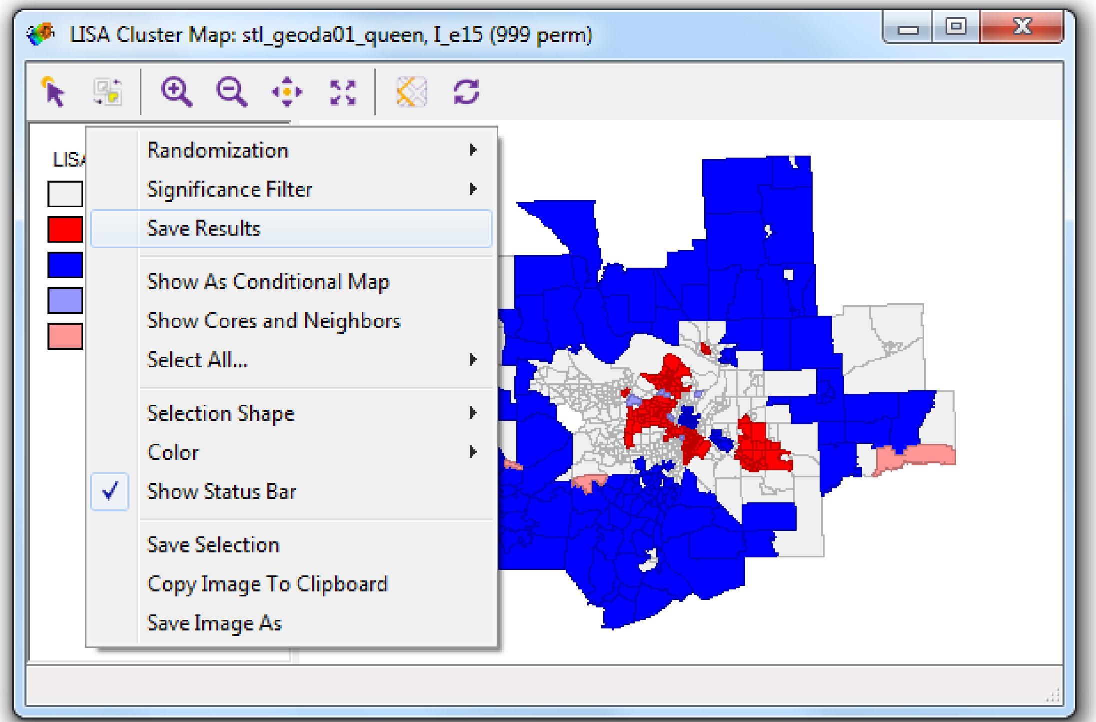
Step 3 – Review your output.

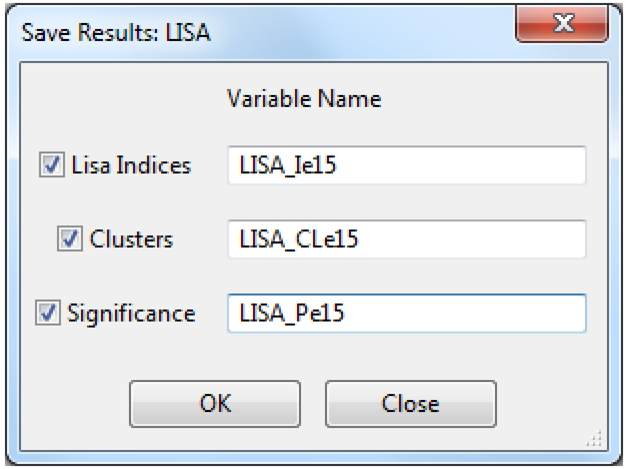


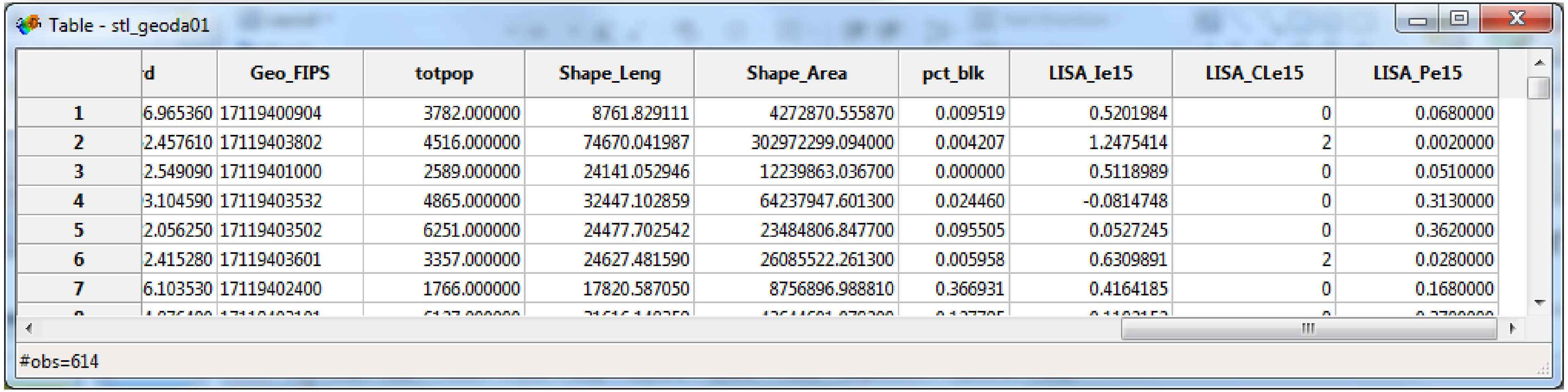




Step 4 – Save your results ->Put the cursor in the cluster map->right click->select Save Results

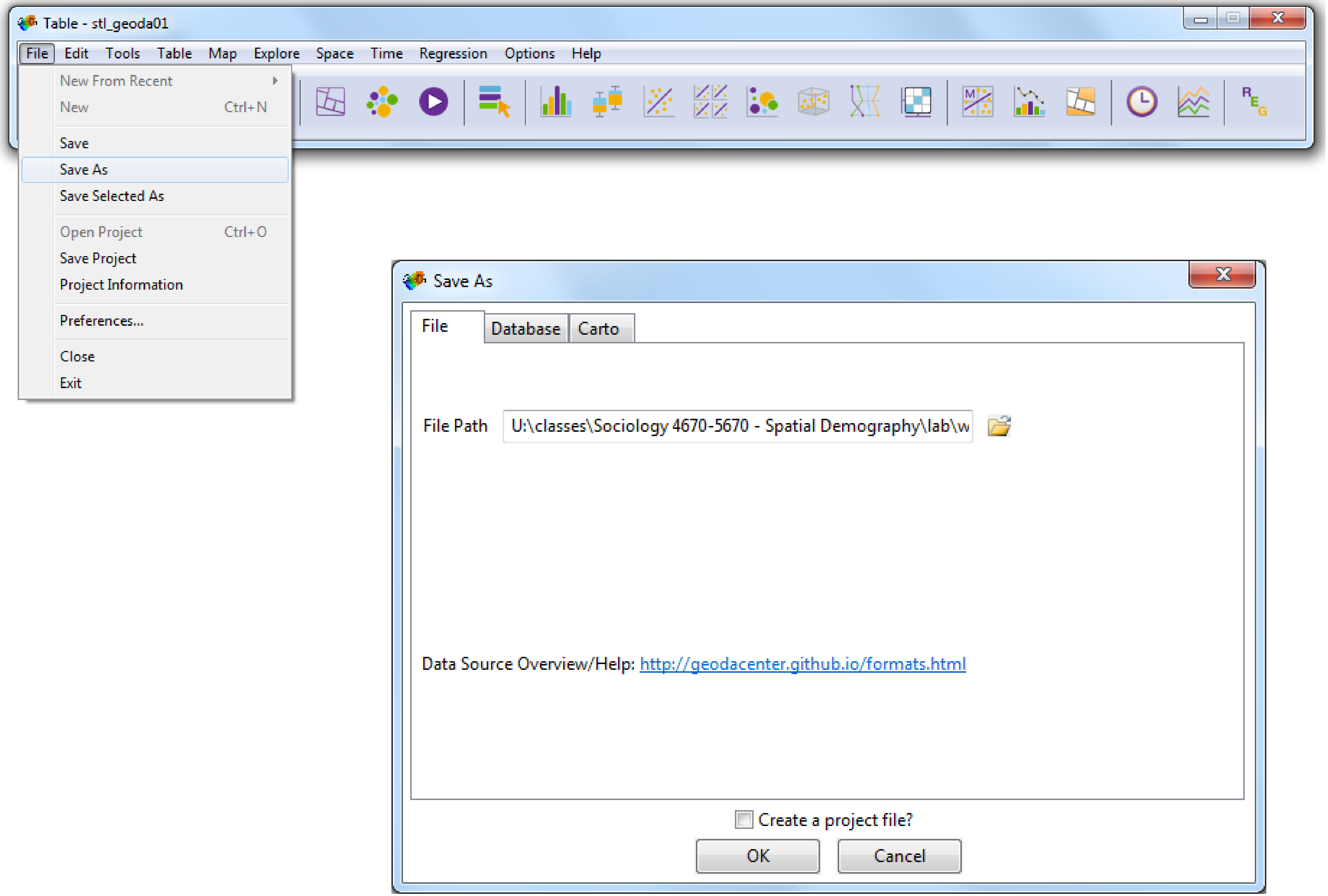


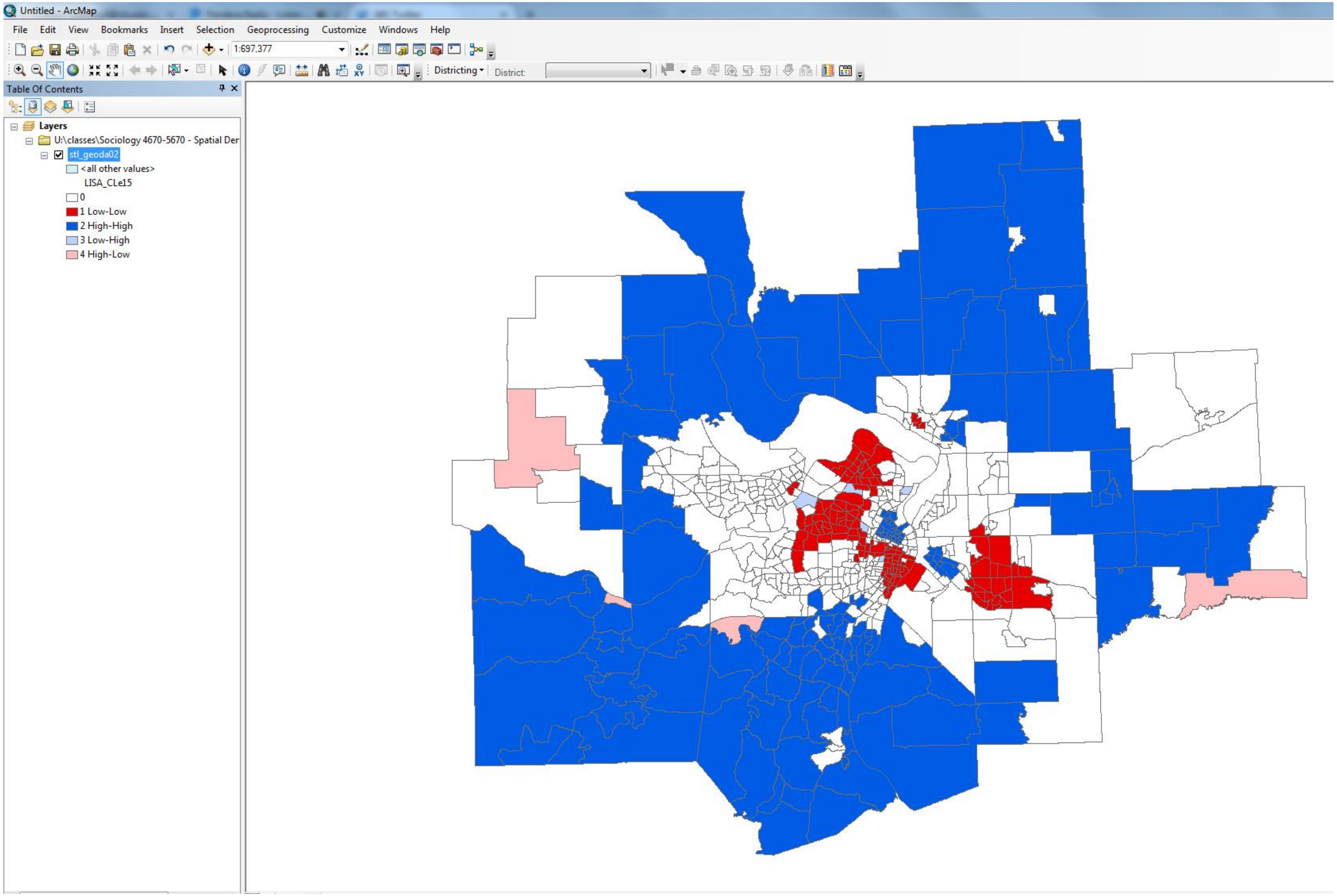




	d	Geo_FIPS	totpop	Shape_Leng	Shape_Area	pct_blk	LISA_Ie15	LISA_CLe15	LISA_Pe15
1	6.965360	17119400904	3782.000000	8761.829111	4272870.555870	0.009519	0.5201984	0	0.0680000
2	2.457610	17119403802	4516.000000	74670.041987	302972299.094000	0.004207	1.2475414	2	0.0020000
3	2.549090	17119401000	2589.000000	24141.052946	12239863.036700	0.000000	0.5118989	0	0.0510000
4	3.104590	17119403532	4865.000000	32447.102859	64237947.601300	0.024460	-0.0814748	0	0.3130000
5	2.056250	17119403502	6251.000000	24477.702542	23484806.847700	0.095505	0.0527245	0	0.3620000
6	2.415280	17119403601	3357.000000	24627.481590	26085522.261300	0.005958	0.6309891	2	0.0280000
7	6.103530	17119402400	1766.000000	17820.587050	8756896.988810	0.366931	0.4164185	0	0.1680000

Step 5 – Save your results as new shapefile ->Put cursor in the cluster map->right click->select Save Results




Untitled - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
1:697,377
Districting
District:
Table Of Contents
Layers
U:\classes\Sociology 4670-5670 - Spatial Der
stl_geoda02
<all other values>
LISA_CLe15
0
1 Low-Low
2 High-High
3 Low-High
4 High-Low

Technical Notes

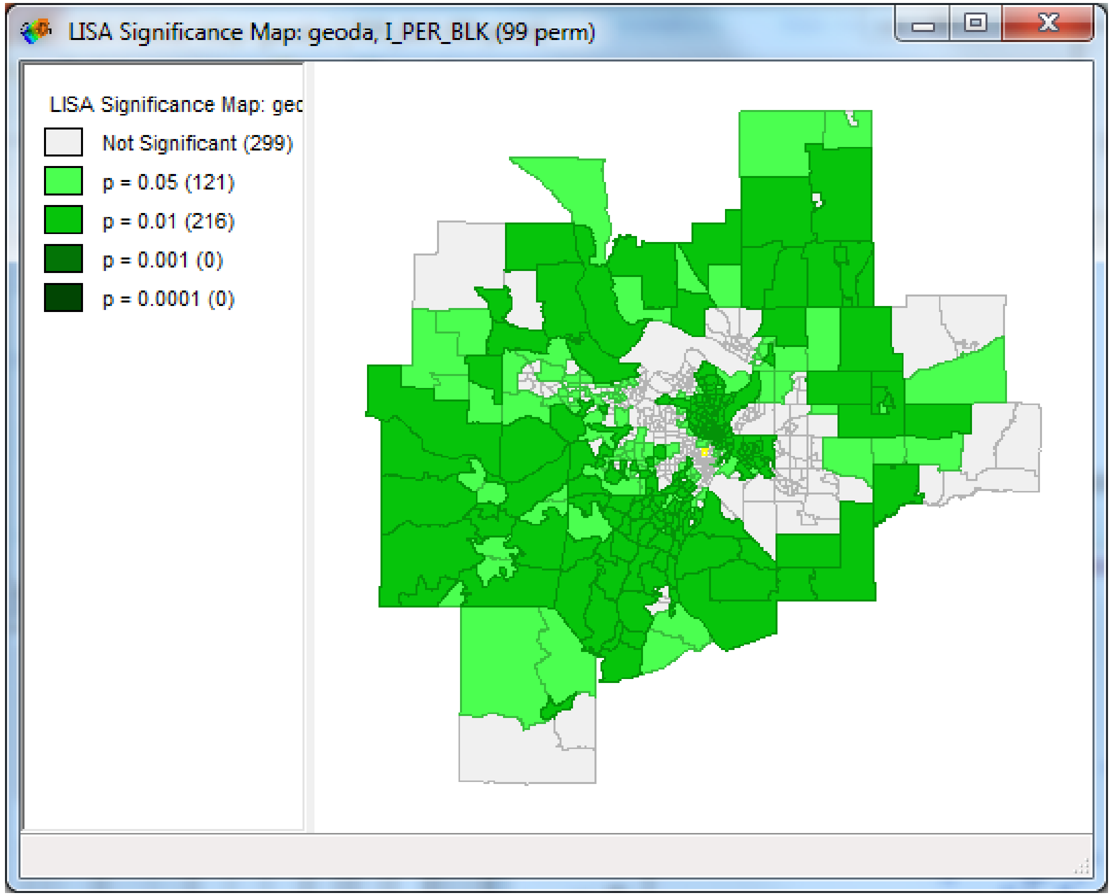
Technical Notes

- The LISA value for each location is determined from its individual contribution to the global Moran's I calculation.

- The **statistically significant values** are assessed by comparing the **actual value** to the **value calculated for the same location** by randomly reassigning the data among all the areal units and recalculating the values each time.

- Actual LISA values are ranked relative to the set of values produced by this randomization process.

- If an actual LISA score fall within the criteria (top 0.1%, 1% or 5%) of scores associated with that location under randomization, then it is judged statistically significant at the (0.001 , 0.01, or 0.05) level.



Technical Notes

- The combination of the Cluster Map and the Significance Map allows you to see which locations are contributing most strongly to the global outcome and in which direction.

- By adjusting the Significance Filter in the Cluster Map, you can see only those areas of highest significance.

- By selecting the Randomization right-click menu option and choosing a larger number of permutations (the default is 99), you can test just how strongly significant are the high-high and low-low outcomes seen in the Cluster Map.