

# Sociology 4670-5670

# Introduction to Spatial Demography

---

Ness Sandoval

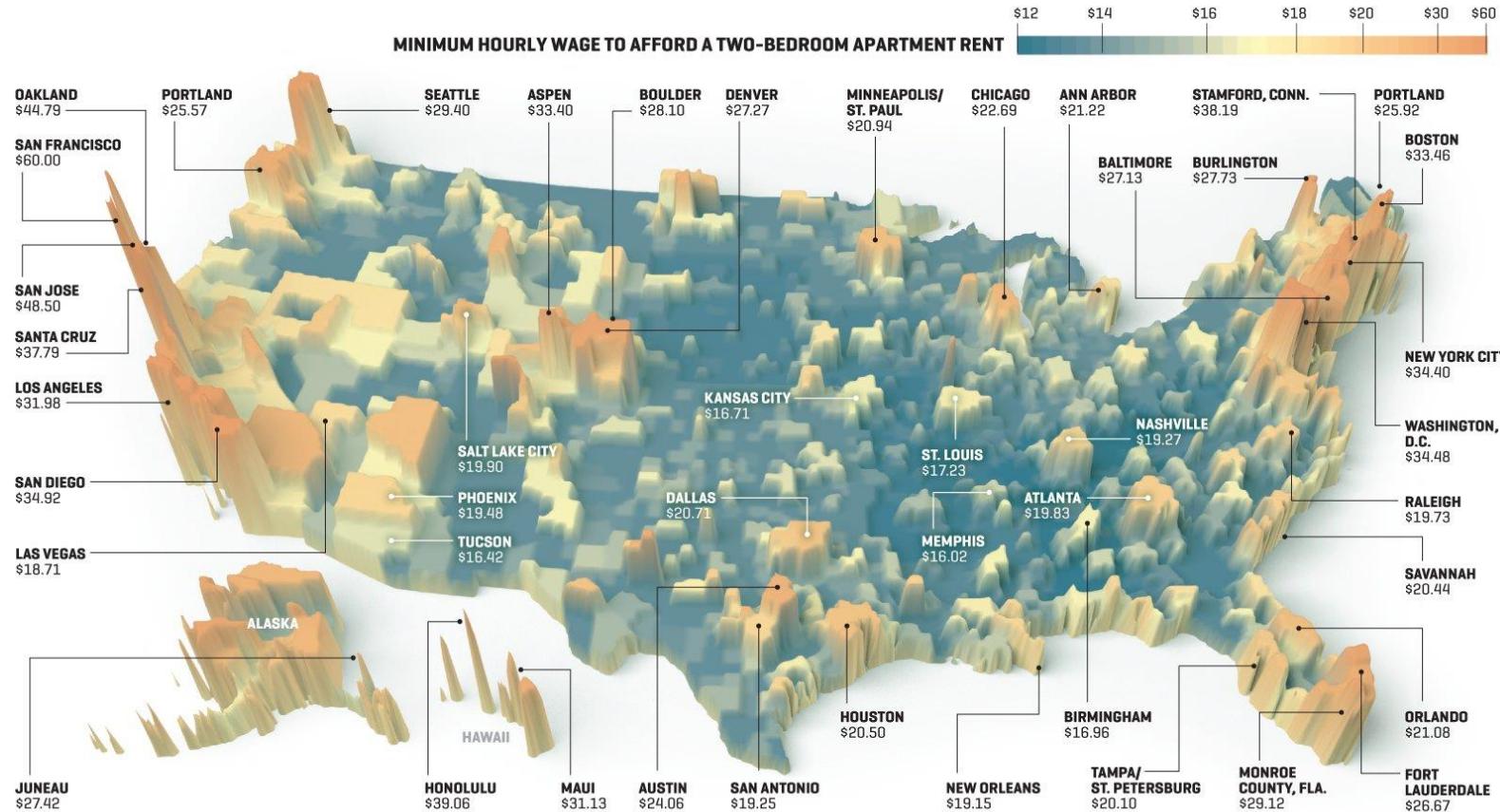
[ness.sandoval@slu.edu](mailto:ness.sandoval@slu.edu)

Sociology

Saint Louis University

# Outline

- Overview of the Course
- Basic Demographic Mapping
- Gravity Model
- Racial Diversity
- Spatial Descriptive Statistics
- Spatial Statistics
- Spatial Models
- Introductions of Students



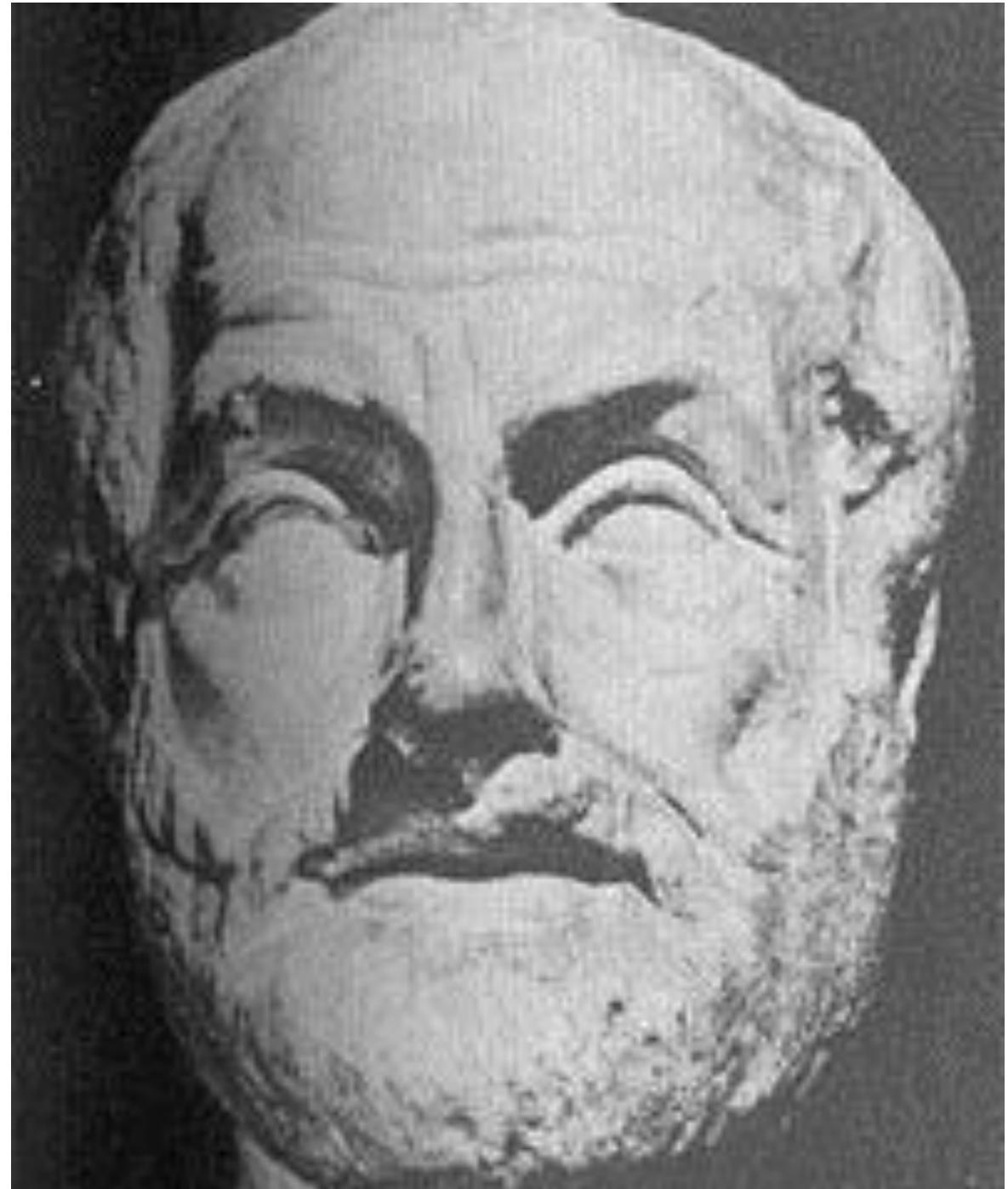
SOURCE: NATIONAL LOW INCOME HOUSING COALITION

# Overview of the Course

---

“The gods did not reveal, from the beginning, all things to us; but in the course of time, through seeking we may learn, and know things better. But as for certain truth, no man has known it, nor will he know it; neither of the gods, nor yet of all the things of which I speak. And even if by chance he were to utter the perfect truth, he would himself not know it; for all is but a woven web of guesses.”

*Xenophanes, 6<sup>th</sup> Century B.C.*



“Even to have expressed a false thought boldly and clearly is already to have gained a great deal.”

*Ludwig Wittgenstein  
(1948)*



# Expectations for class

- Pedagogy
  - Individual learning
  - Group learning
  - Technology
- Upper Lever Undergraduate Course and Graduate level course
- Two Big Projects - Don't procrastinate

# ASSIGNMENTS

Homework #1 – Spatial Introduction to a City	25%
Homework #2 – Class Presentation on Final Project	15%
Homework #3 – Peer Review of Class Presentations	10%
Homework #4 – Research Paper on Final Project	50%
<b>Total</b>	<b>100%</b>

# ASSIGNMENT DUE DATES

Homework #1 – Spatial Introduction to a City	March 6
Homework #2 – Class Presentation on Final Project	April 24 and May 01
Homework #3 – Peer Review of Class Presentation	April 24 and May 01
Homework #4 – Research Paper on Final Project	May 07

# GRADING CRITERIA

A                  100-94

A-                93-91

B+                90-88

B                87-85

B-                84-82

C+                81-79

C                78-76

C-                75-73

D                72-62

F                62 and below

## GRADING POLICY

Each assignment will be graded on scale of 0-100. There is no final exam, but there is a final in-class presentation and final research paper.

# Recommended Texts

- Anselin, Luc and Rey, Sergio. 2014. Modern Spatial Econometrics in Practice: A Guide to GeoDA, GeoDaSpace and PySAL. GeoDA Press LLC.
- Bivand, Roger S. Pebesma, Edzer, and Gómez-Rubio. 2013. Applied Spatial Data with R. Springer.
- Brunsdon, Chris and Comber, Lex. 2015. An Introduction to R for Spatial Analysis and Mapping. SAGE Publications.
- Haining, Robert. 2009. Spatial Data Analysis. Cambridge University.
- Fotheringham, A. Stewart, Brunsdon, Chris and Charlton, Martin. 2000 Quantitative Geography: Perspectives on Spatial Data Analysis. Sage Publishing.

# Outline of the class

## **Week 1 – January 16**

Overview and Goals of the Course and Seminar - Why Spatial Demography?

## **Week 2 – January 23**

Spatial Demography Concepts and Databases – Part 1

## **Week 3 – January 30**

Spatial Demography Concepts and Databases – Part 2

## **Week 4 – February 6**

Exploratory Spatial Data Analysis

# Outline of the class

**Week 5 – February 13**

Spatial Interpolation

**Week 6 – February 20**

Spatial Autocorrelation – Global

**Week 7 – February 27**

Spatial Autocorrelation – Local

**Week 8 – March 6**

Point Pattern Analysis and Kernel Density

**Week 9 - March 13**

No Class Spring Break

# Outline of the class

**Week 10 – March 20**

Spatial Temporal Analysis

**Week 11 – March 27**

Spatial Autoregressive Models

**Week 12 – April 3**

Geographically Weighted Regression

**Week 13 and 14 – April 10 and April 17**

Spatial Statistics with R

**Week 15 and 16 - April 24 and May 1**

Presentations on Paper

# Content of Class

- We need to speak the same language
  - Review of Statistics
  - Review of Demographic concepts
- Learn how to effectively use demography with maps and statistics
- Beginning, Intermediate, and Advanced Use of Spatial Statistics
- Plenty of time built into class schedule to work on your individual projects

# Demography as Spatial Social Science or Computational Spatial Science

- Macro Demography vs. Micro Demography
- Applied Demography
- The future of Demography is spatial
- The future of spatial demography is integrated science
- R and Python are the future of open science

# The Definition of Spatial Demography

- The formal demographic study of areal aggregates.
- Demography viewed from the spatial perspective. [...]  
Spatial, together with temporal, variations in mortality, fertility and migration are studied as preliminaries to the investigation of population structure in its entirety.
- Population geography vs. spatial demography

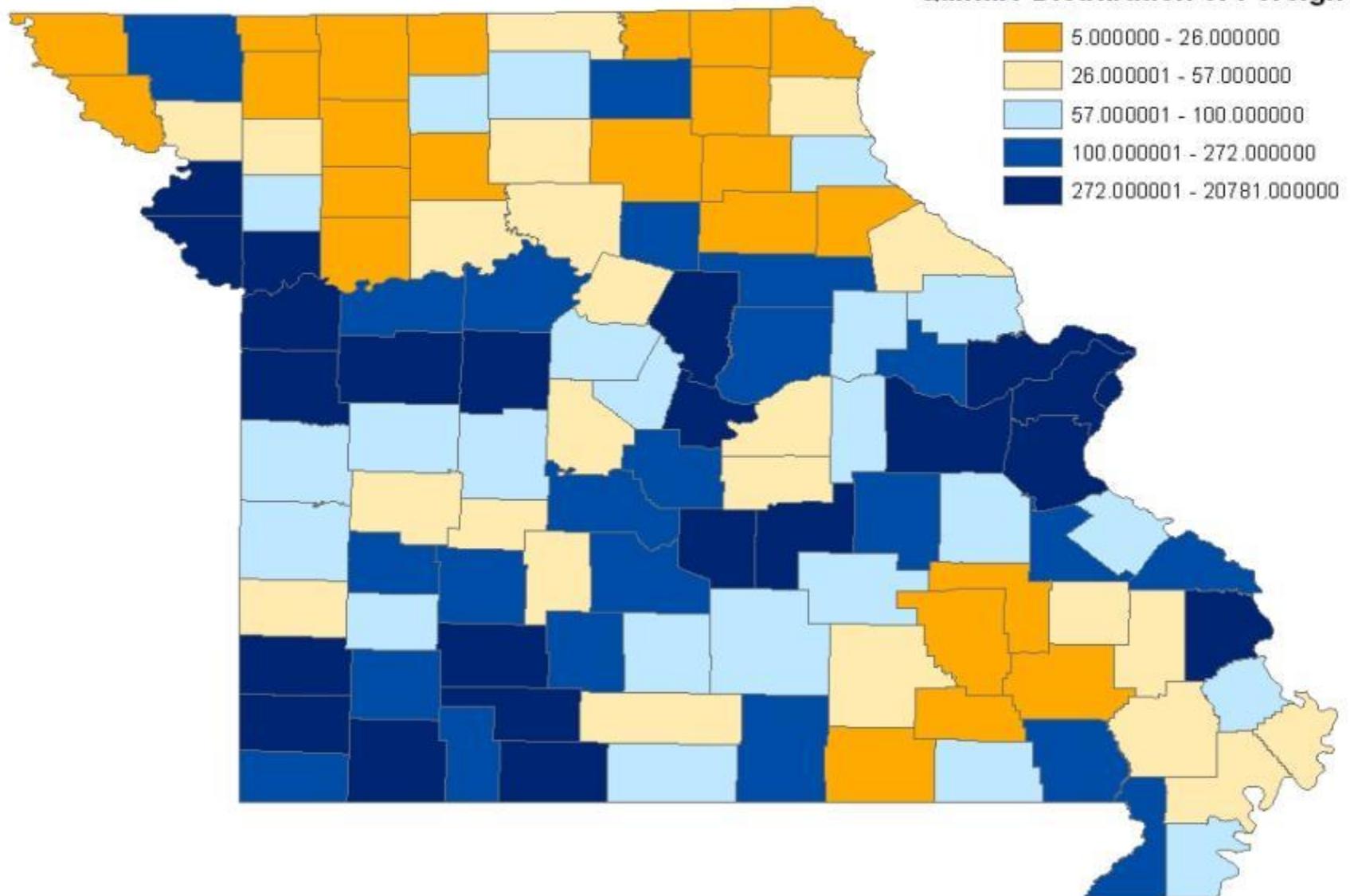
# The Definition of Spatial Demography

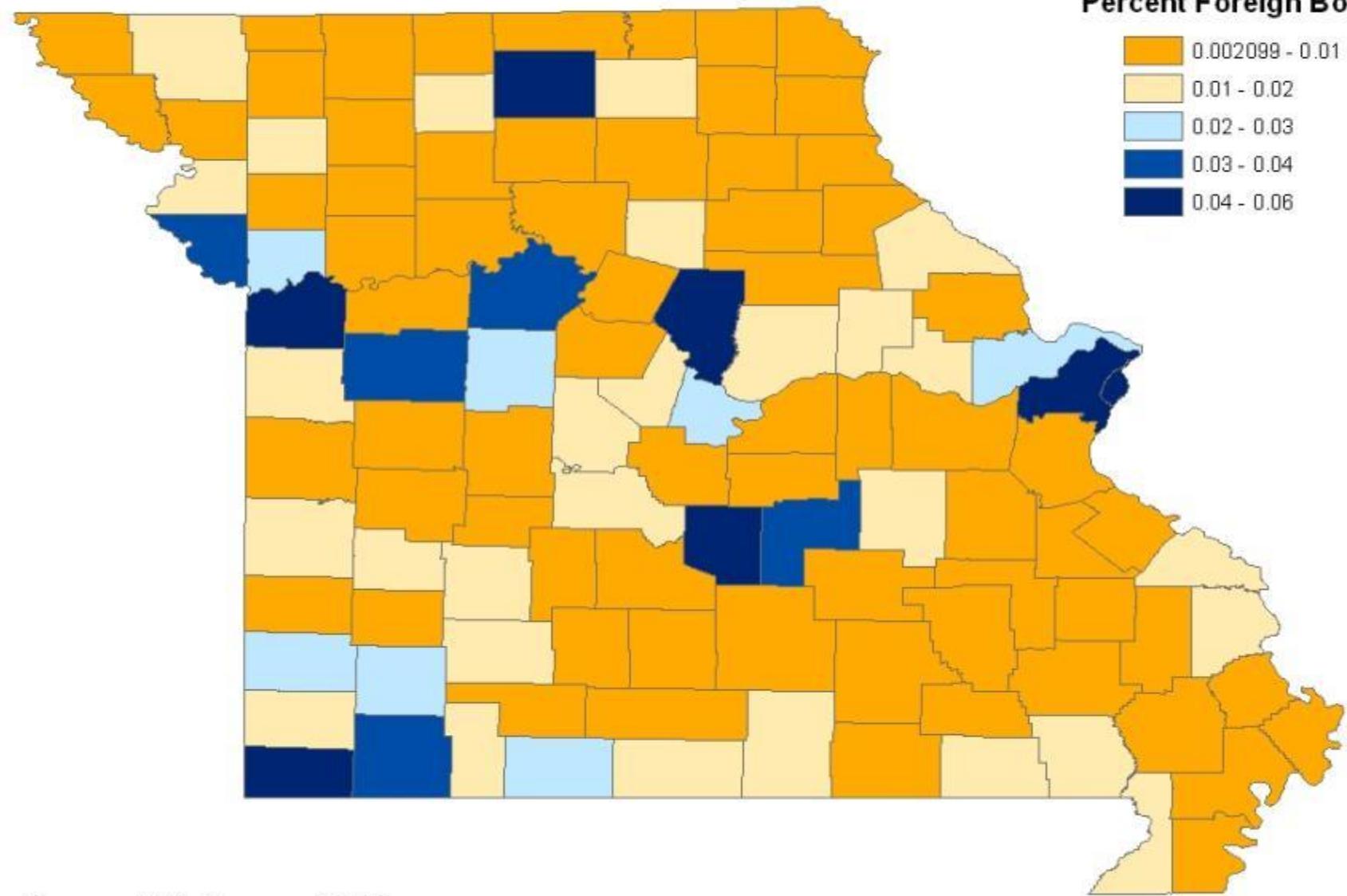
- Spatial Demography includes research on the core of demographic analysis, defined by the three demographic components—fertility, mortality, and migration—and by models on population size, change, and composition. Spatial Demography also includes research on population and environment, poverty, inequality, crime analysis, health outcomes, and land use. Second, spatial demography implies the use of formal statistical modeling, such as techniques in spatial statistical analysis, geostatistics, Bayesian modeling, and spatial econometrics, as well as the formulation of population models in three dimensions: age, time, and space.

# Basic Demographic Mapping

---

### Quintile Distribution of Foreign Born





Source: U.S. Census 2000.

# Gravity Model

---

# Job Accessibility Measure: Regional Index

$$A_{ik} = \sum_j [ \text{Jobs}_j * \exp(-v \text{ Time}_{ijk}) ]$$

$A_{ik}$  = Accessibility indicator

Jobs = # of non-professional jobs in tract

Time = Travel Times

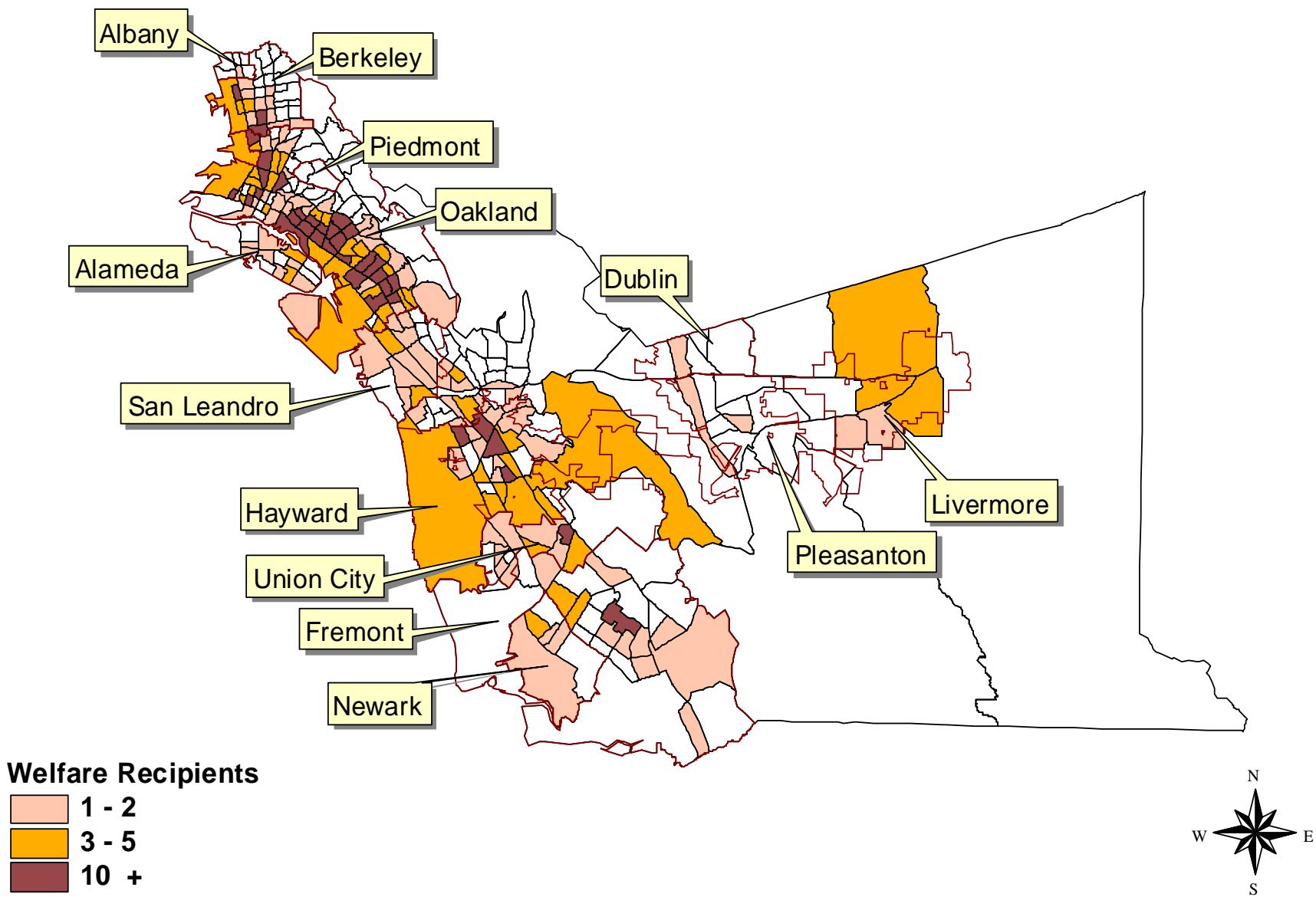
i = Residential Address

j = Centroid of employment zones

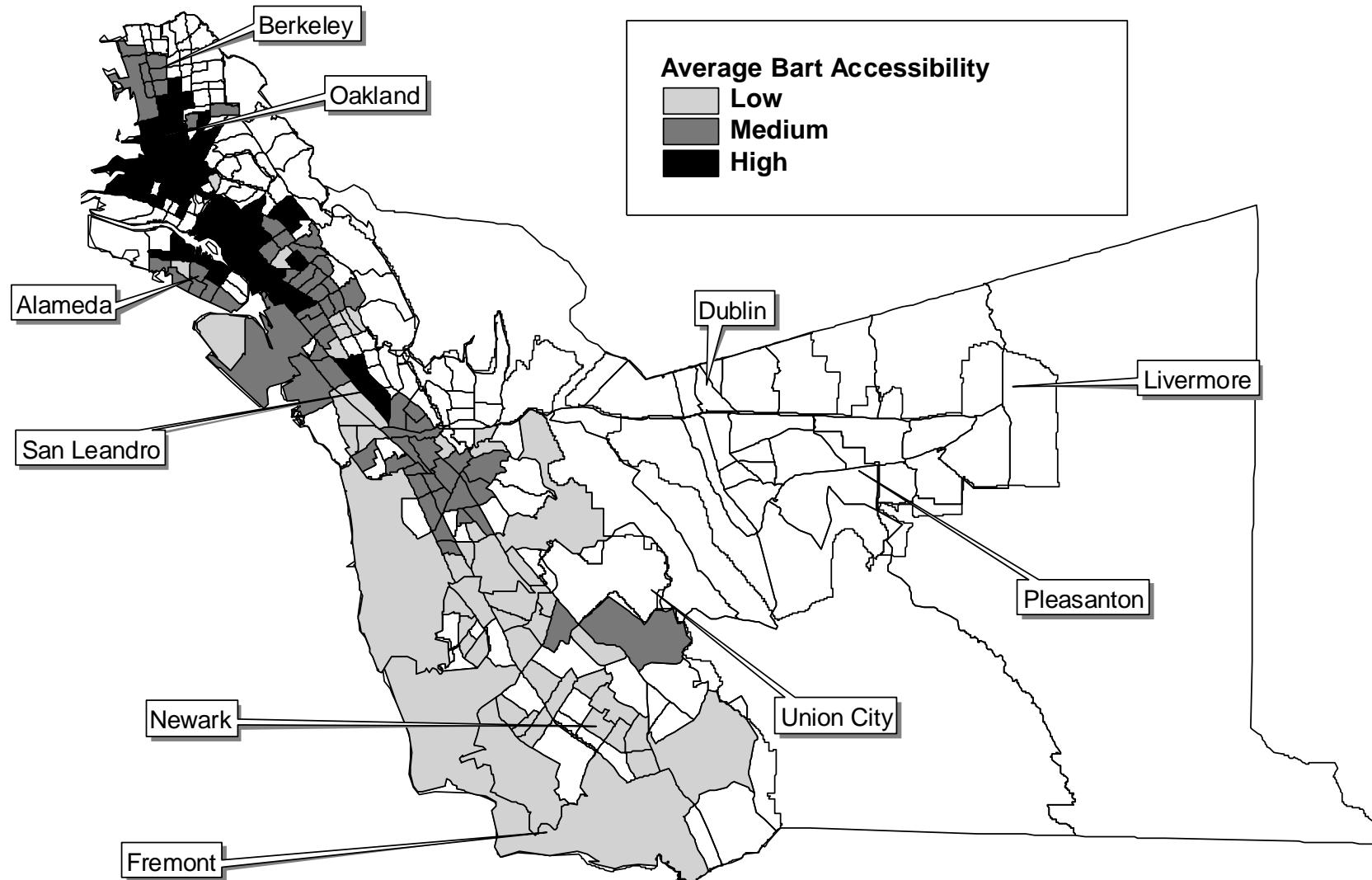
k = Mode (transit vs. highway)

# Alameda County

## Wave II – AFDC Recipients



**Map 3**  
**Job Accessibility Levels via Transit Network**  
**for Alameda County Census Tracts of Residence of Panel Sample**



# Racial Diversity

---

# Diversity Index

$$E_i = \left| \sum_{i=1}^n \frac{(\ln(P(i))) * P(i)}{(\ln k)} \right|$$

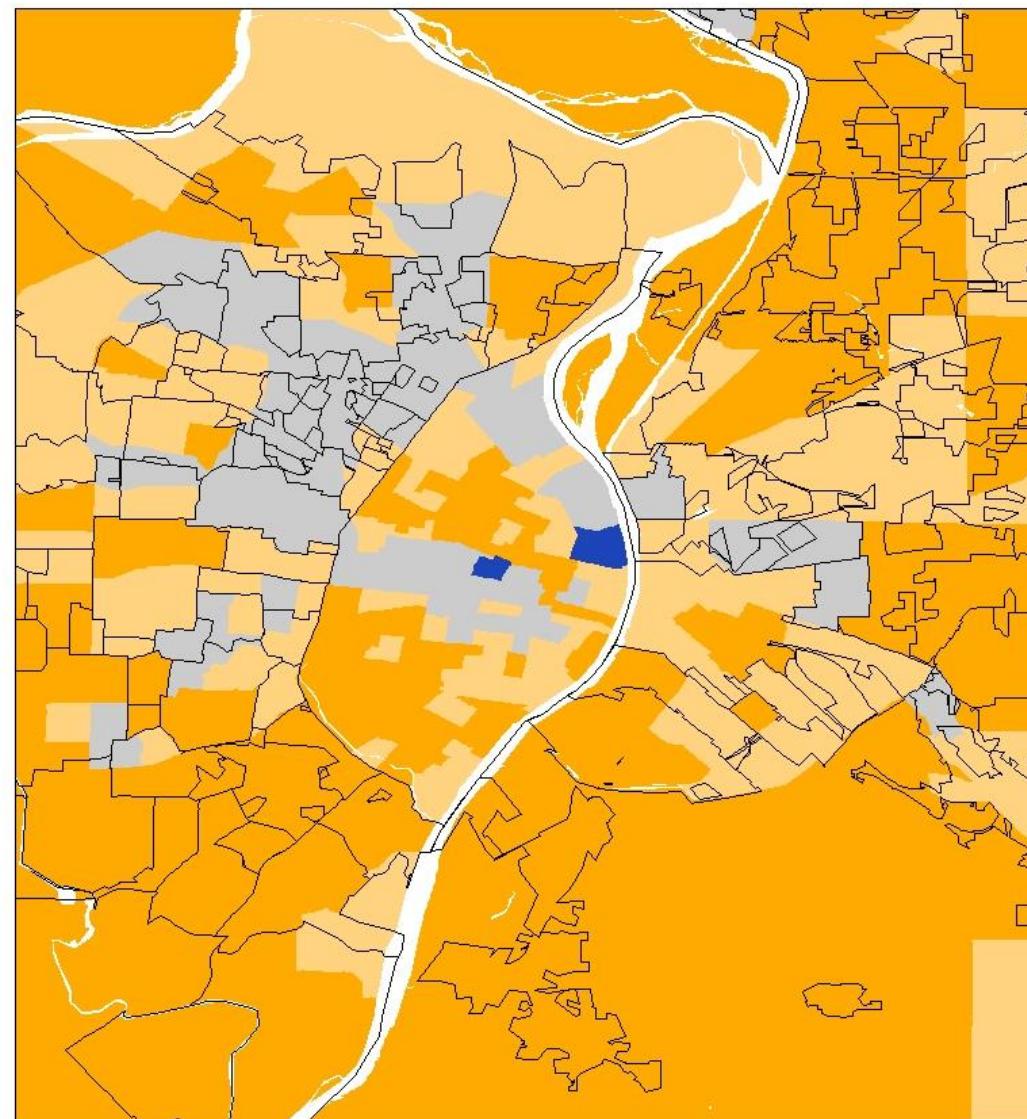
where:

$E_i$ = Diversity score for tract j

$P(i)$ =Proportion of the tract population in race/ethnic group

n=the total number of racial/ethnic categories

Diversity score will range from 0=(No Diversity) to 100=(Very Diverse)



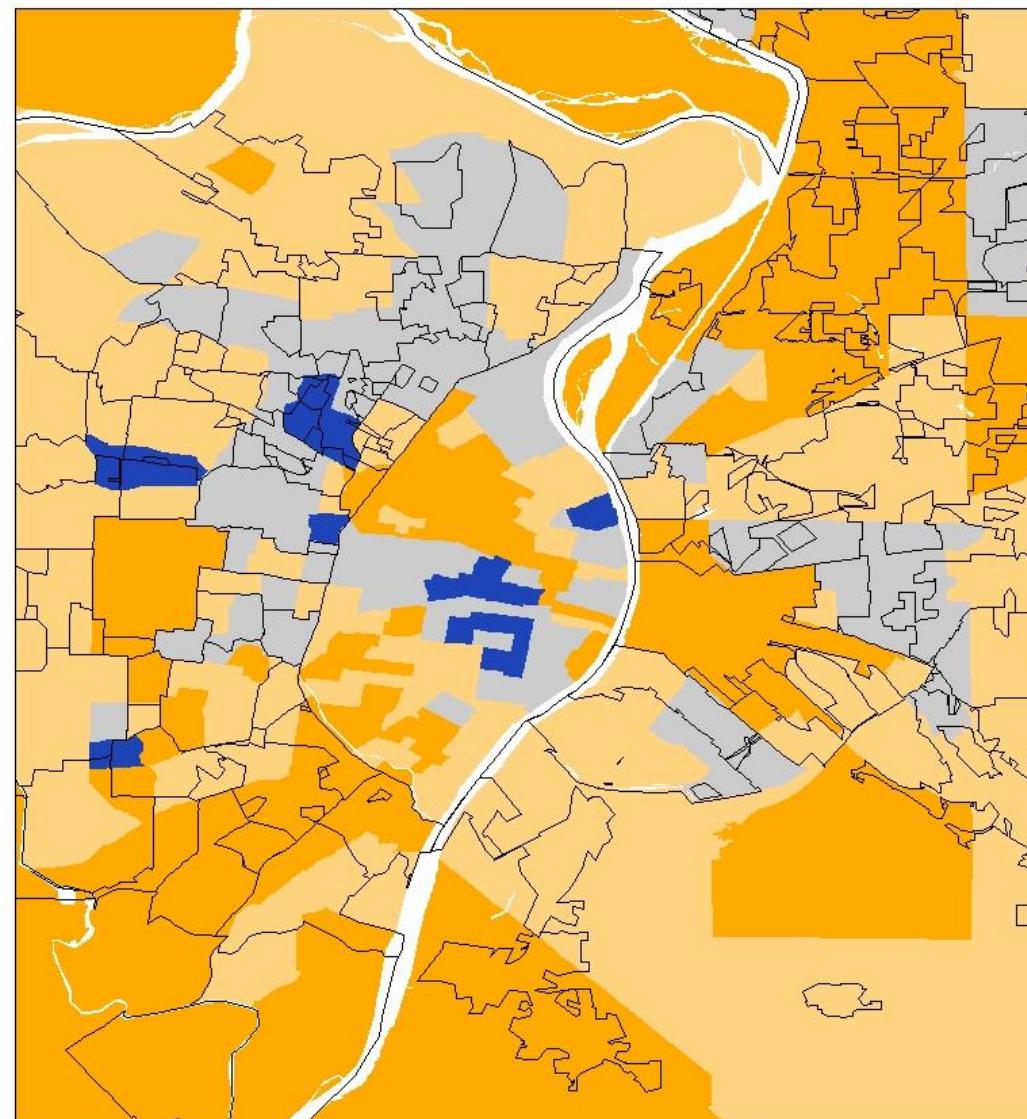
Neighborhood Diversity

mastertract

E80

0.000000 - 0.100000
0.100001 - 0.310000
0.310001 - 0.520000
0.520001 - 0.730000
0.730001 - 1.000000

**Neighborhood Racial Diversity  
Saint Louis 1980**



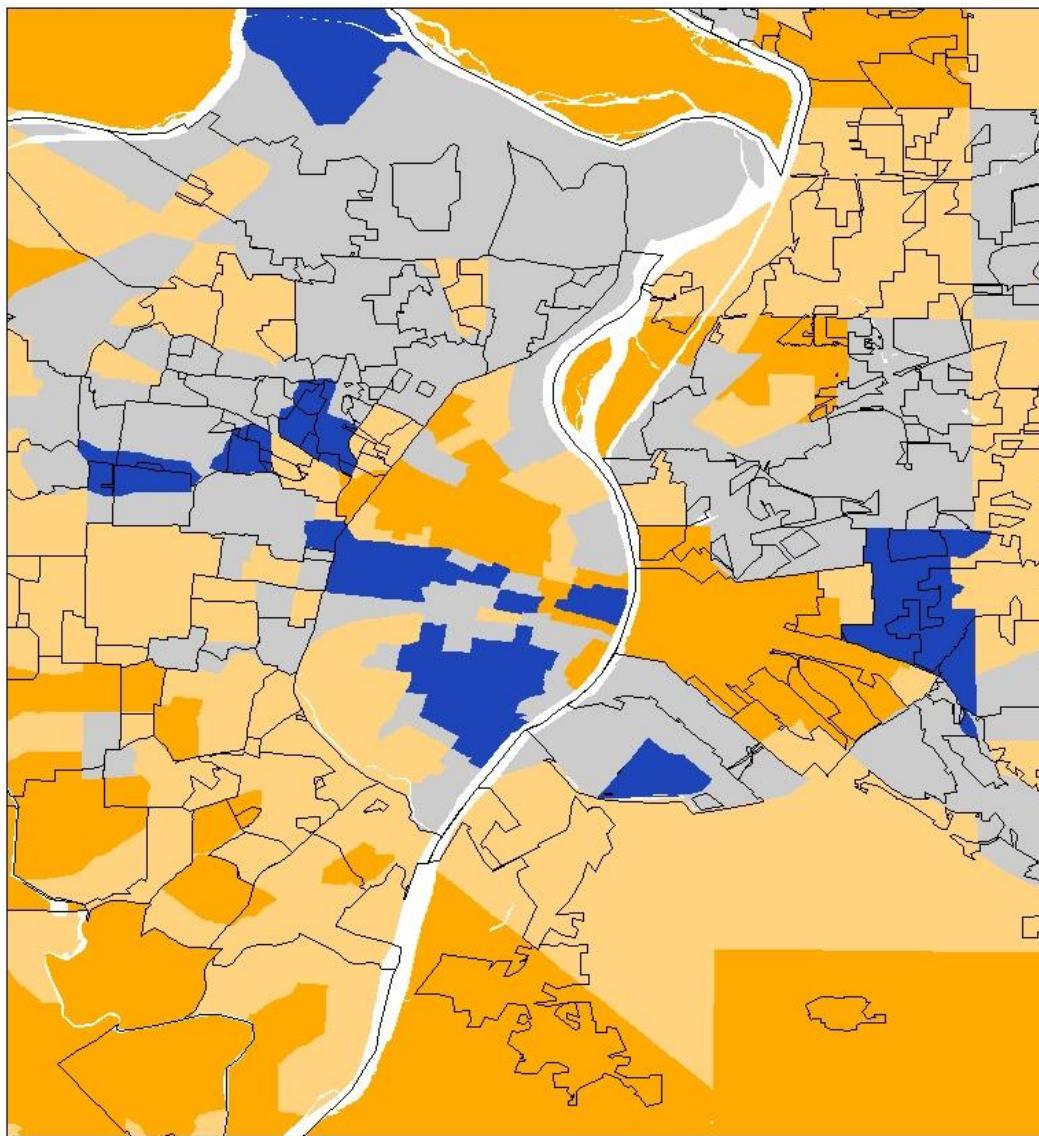
Neighborhood Diversity

mastertract

E90

0.000000 - 0.100000
0.100001 - 0.310000
0.310001 - 0.520000
0.520001 - 0.730000
0.730001 - 1.000000

## Neighborhood Racial Diversity Saint Louis 1990



Neighborhood Diversity

mastertract

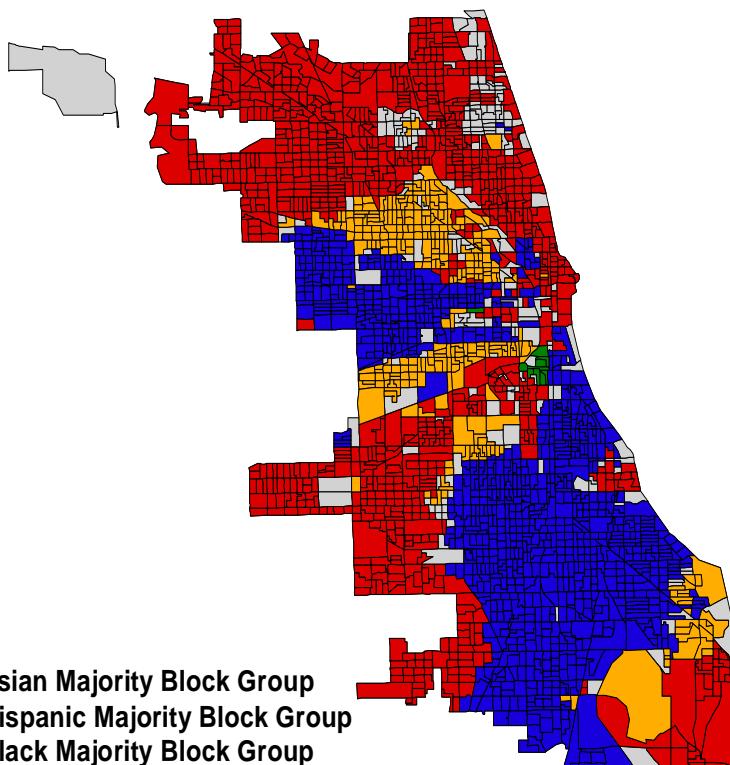
E00

0.000000 - 0.100000
0.100001 - 0.310000
0.310001 - 0.520000
0.520001 - 0.730000
0.730001 - 1.000000

## Neighborhood Racial Diversity Saint Louis 2000

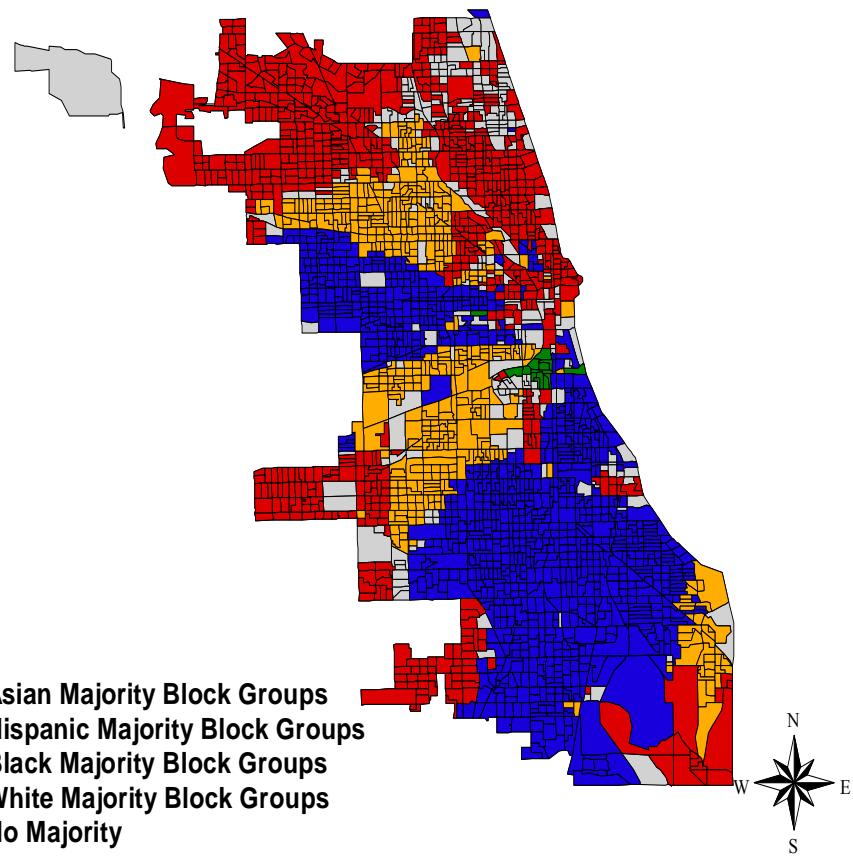
# Chicago Block Groups

1990



- █ Asian Majority Block Group
- █ Hispanic Majority Block Group
- █ Black Majority Block Group
- █ White Majority Block Group
- █ No Majority

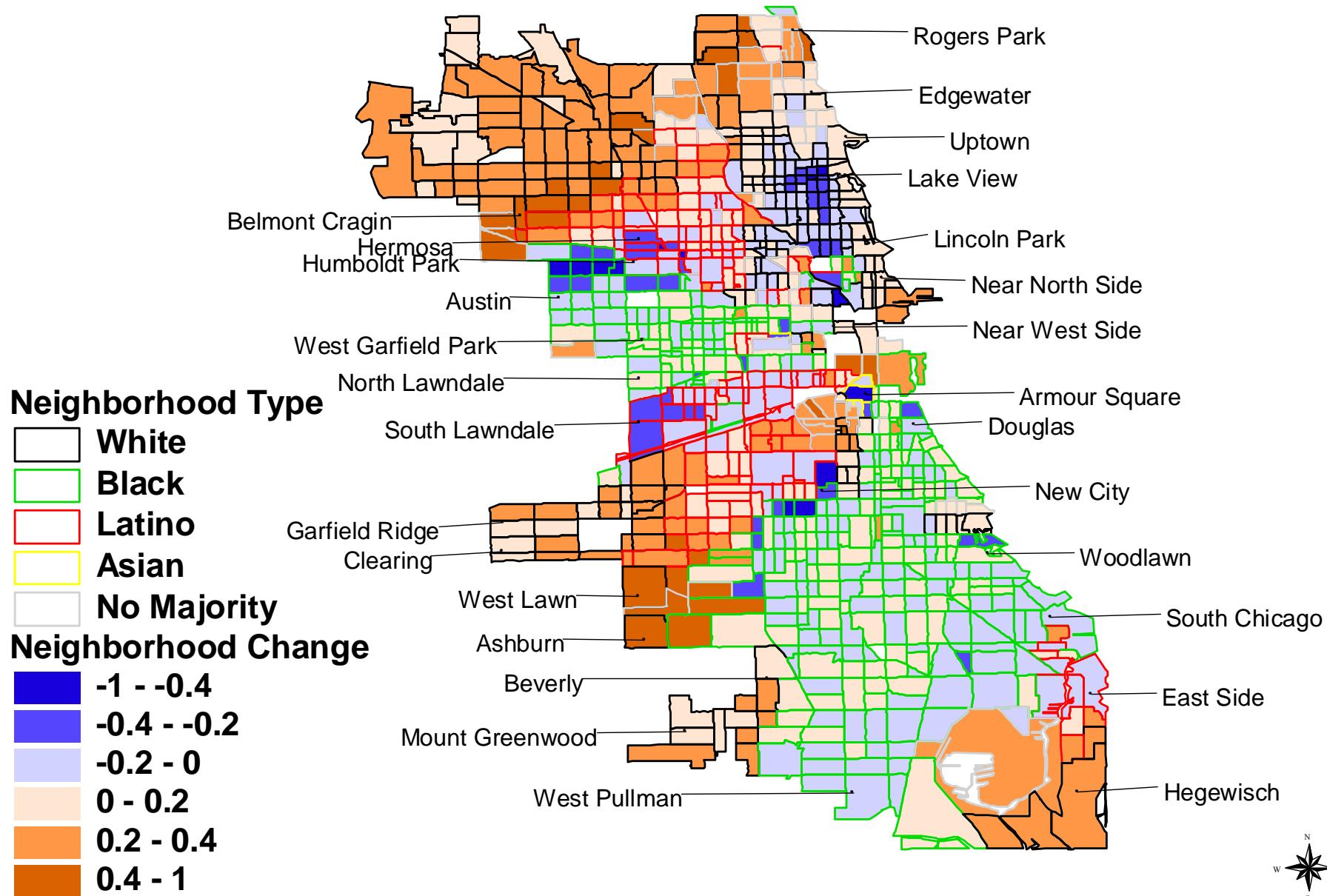
2000



- █ Asian Majority Block Groups
- █ Hispanic Majority Block Groups
- █ Black Majority Block Groups
- █ White Majority Block Groups
- █ No Majority

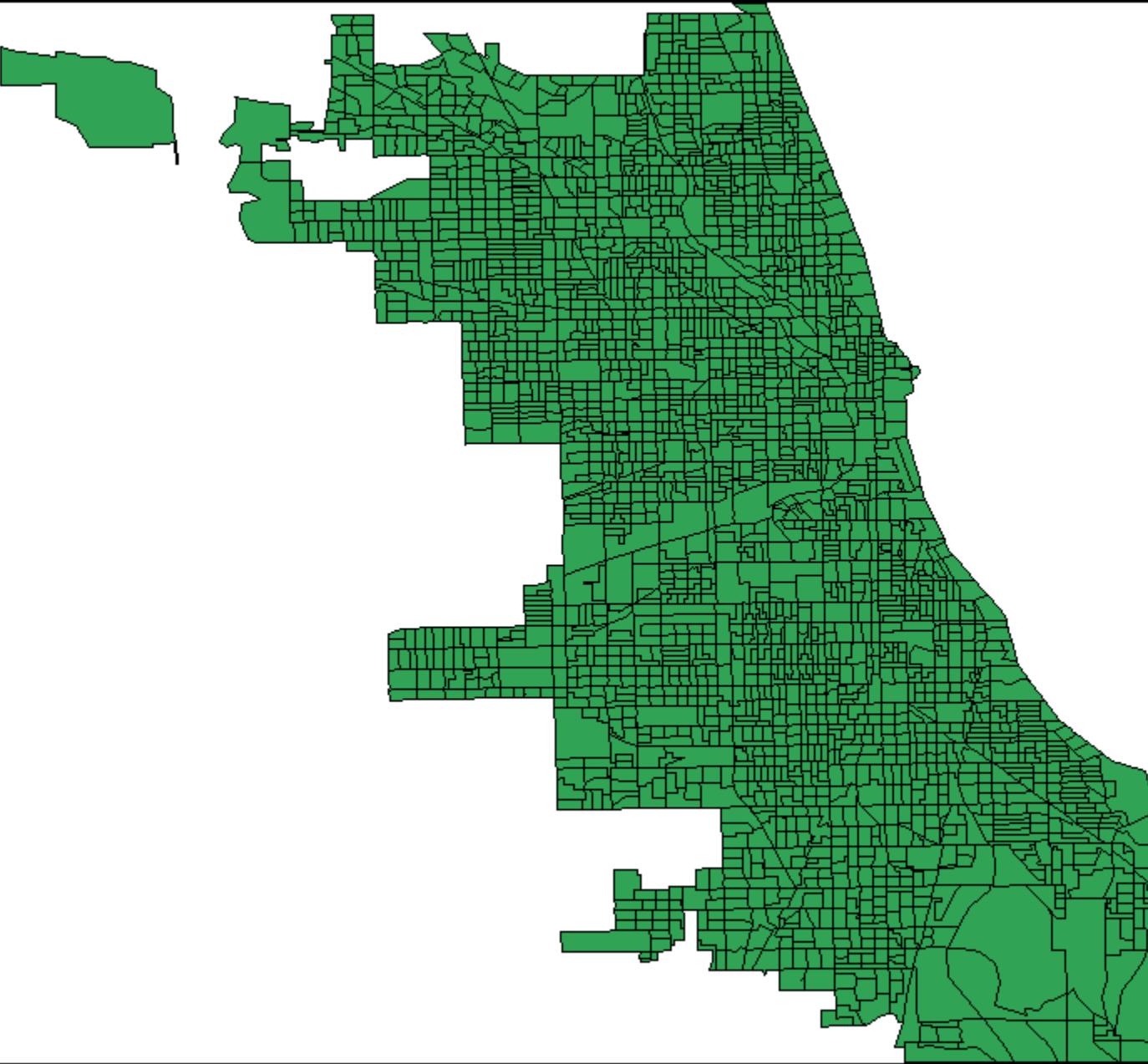
# City of Chicago

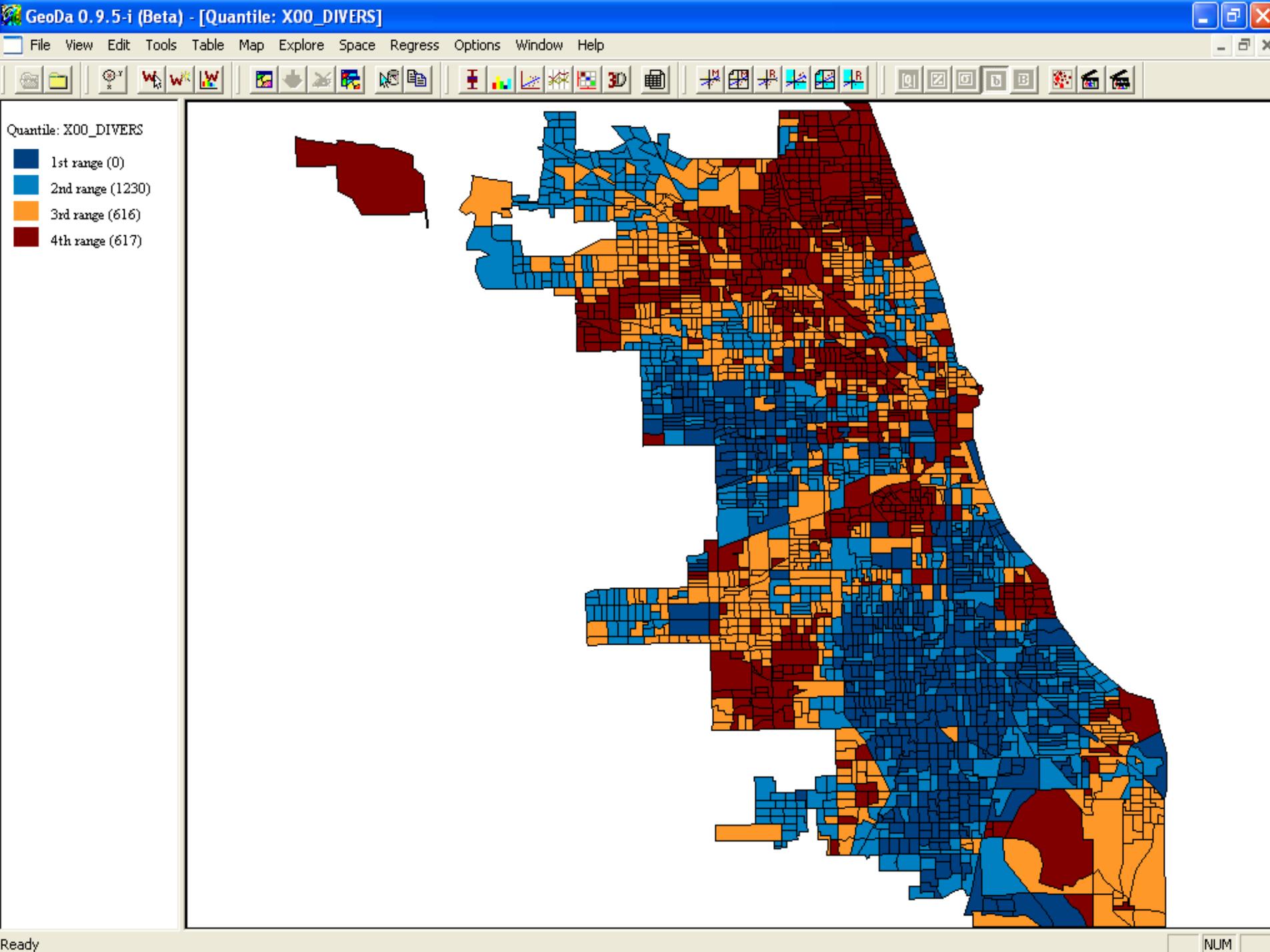
## Change in Diversity Score from 1980 to 2000

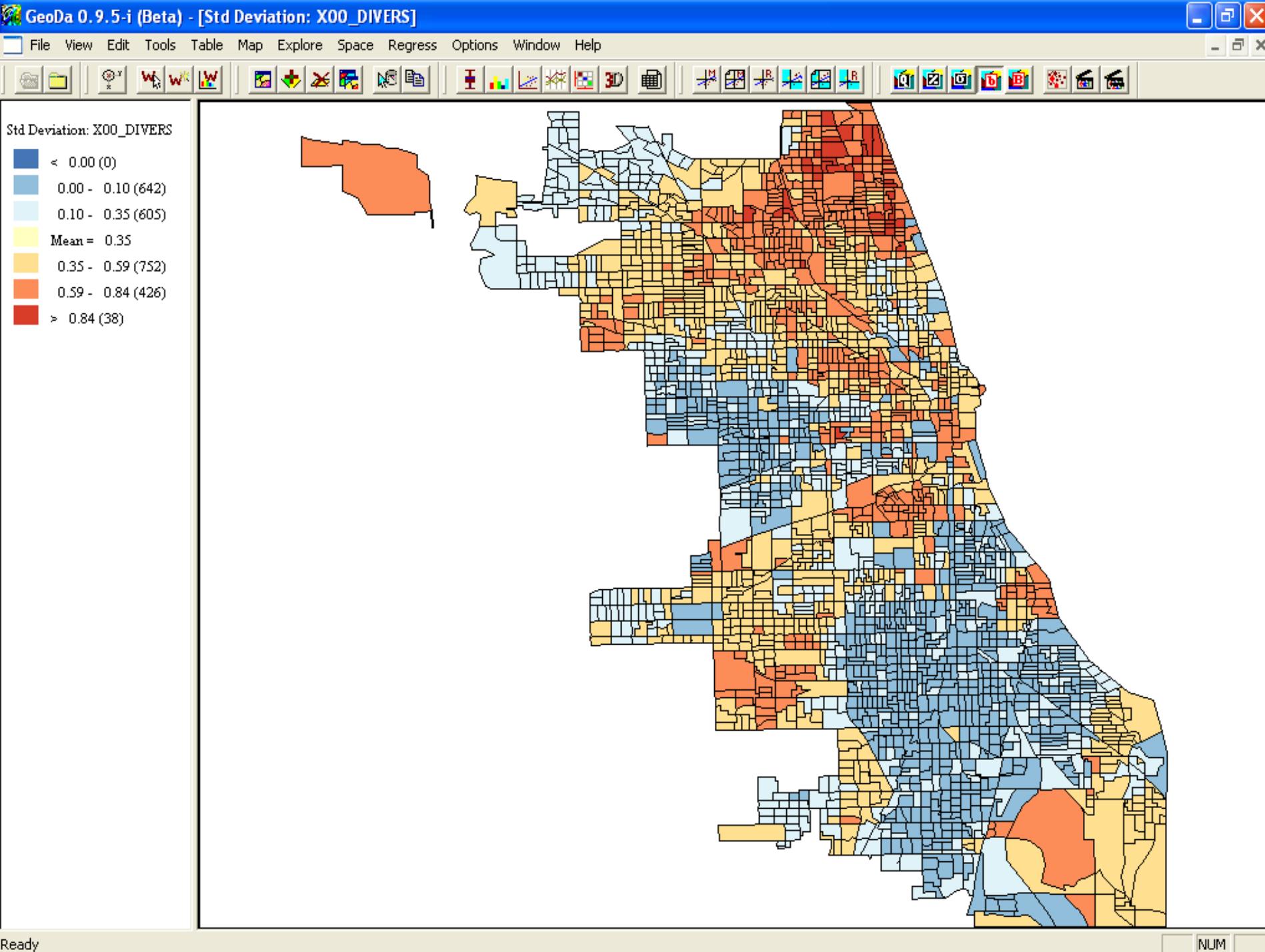


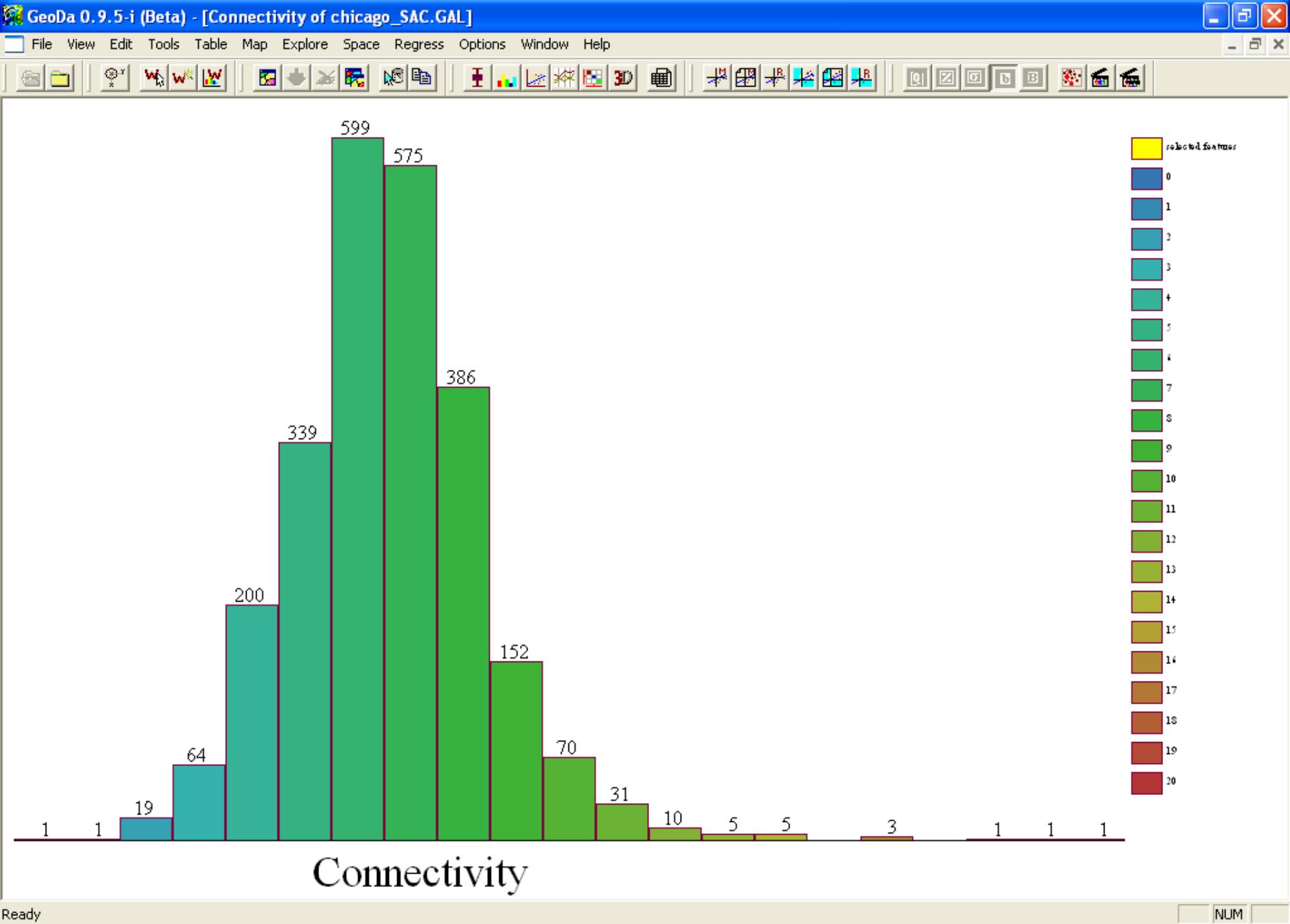


Map Legen







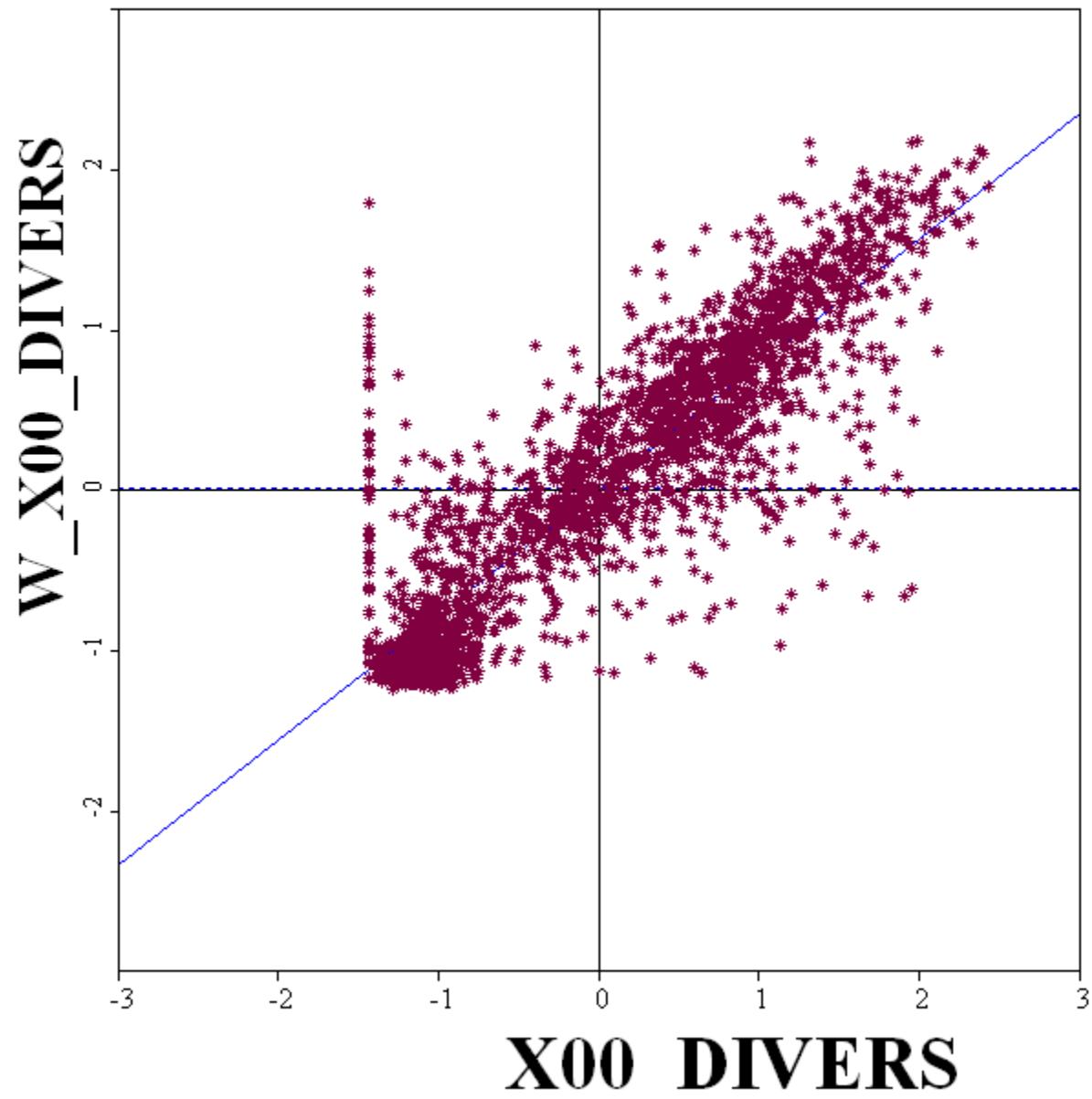


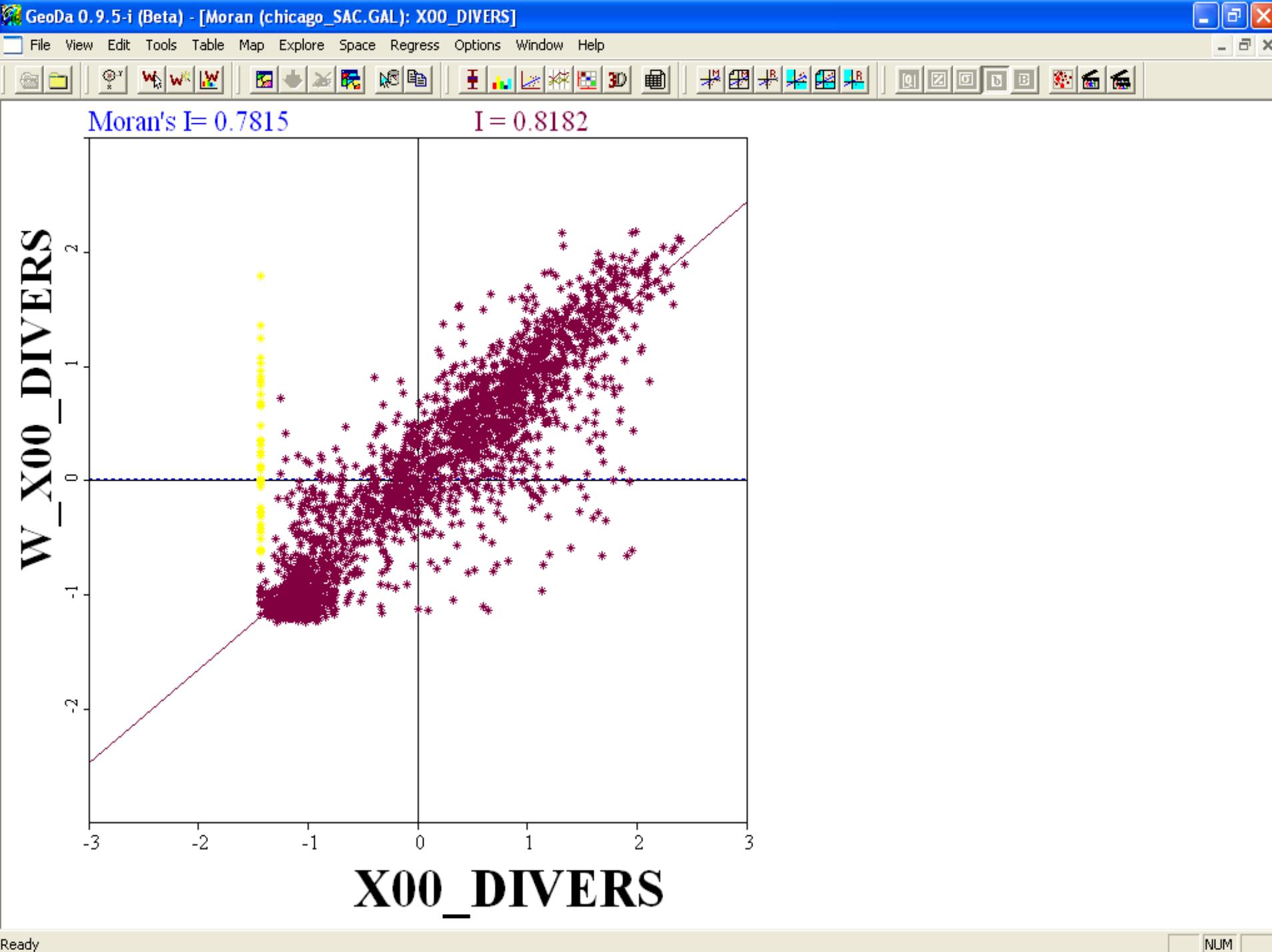
File View Edit Tools Table Map Explore Space Regress Options Window Help



Close

Moran's I= 0.7815





# Randomization



permutation: 499  
p-value : 0.0020

Close

Run

I:0.7815 E[I]:-0.0004 Mean:-0.0006 Sd:0.0113



(1) LISA Significance Map

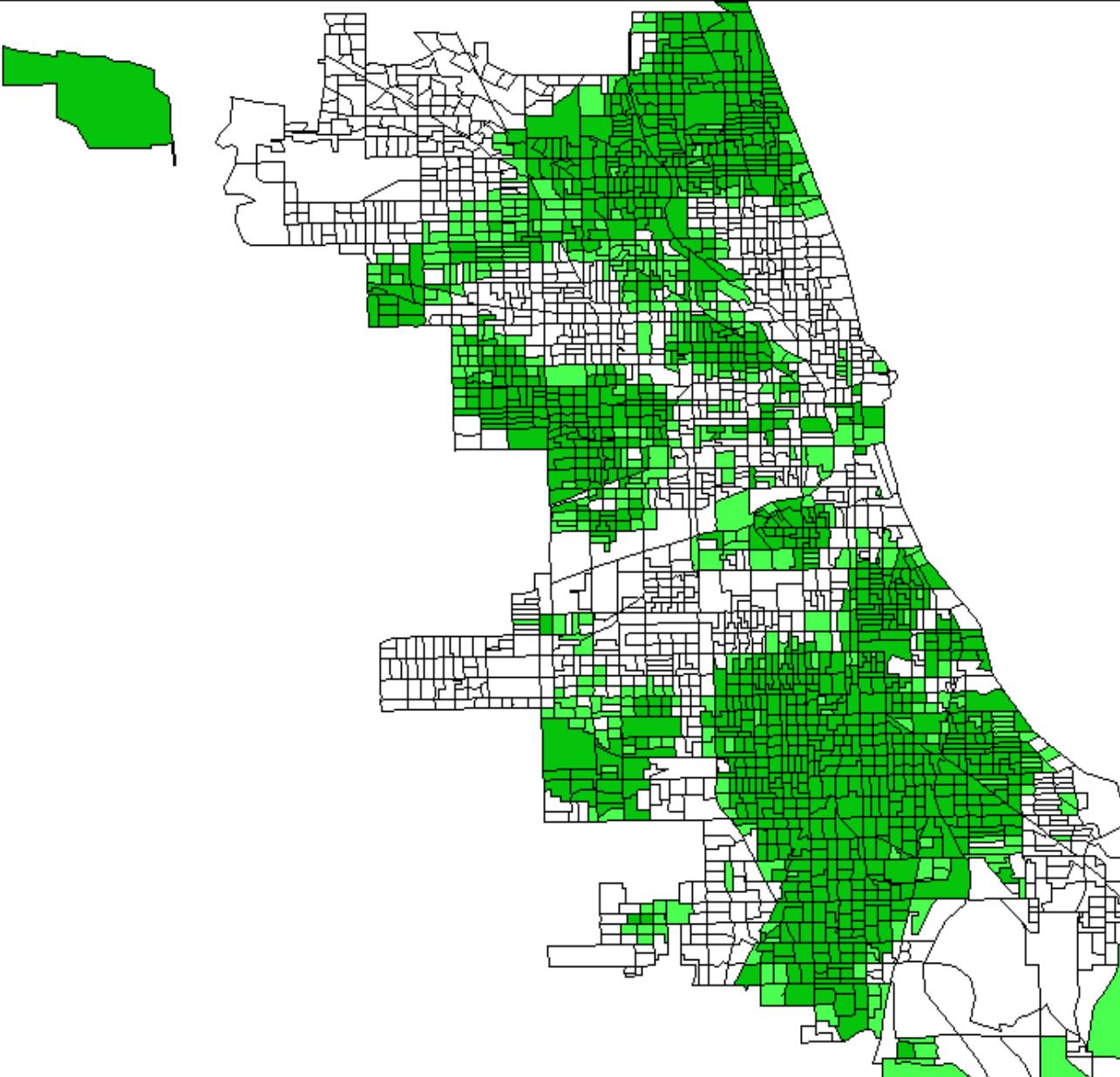
Not Significant

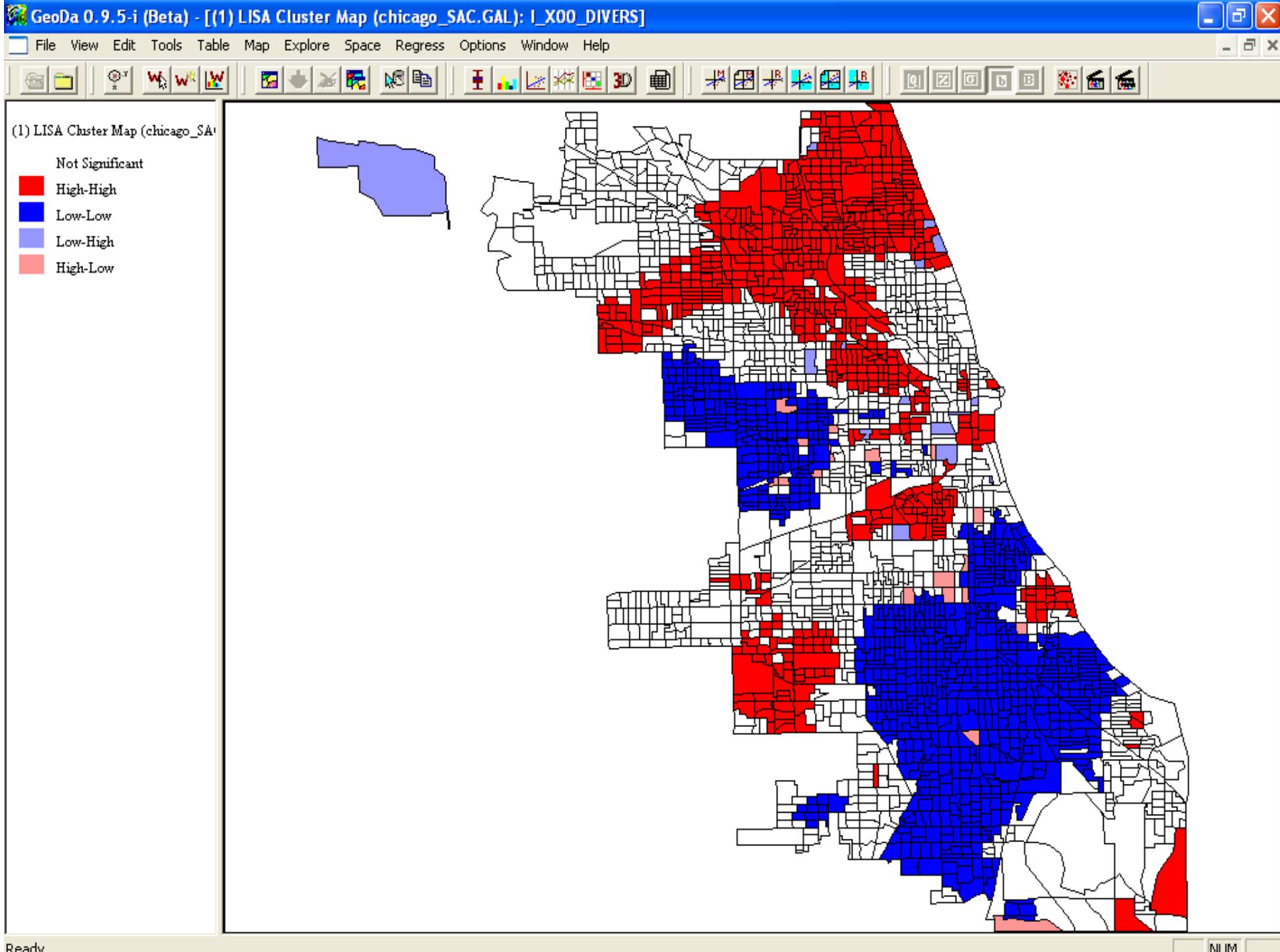
p = 0.05

p = 0.01

p = 0.001

p = 0.0001

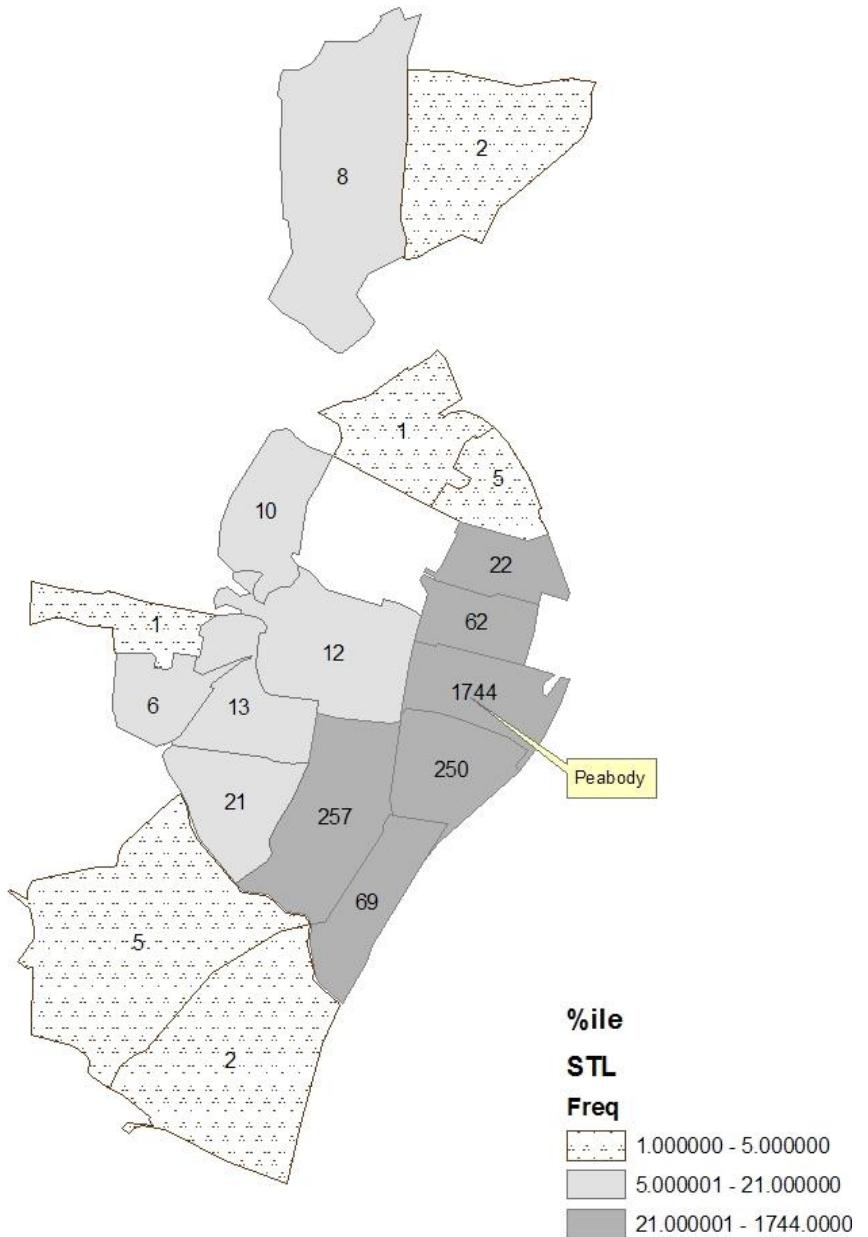




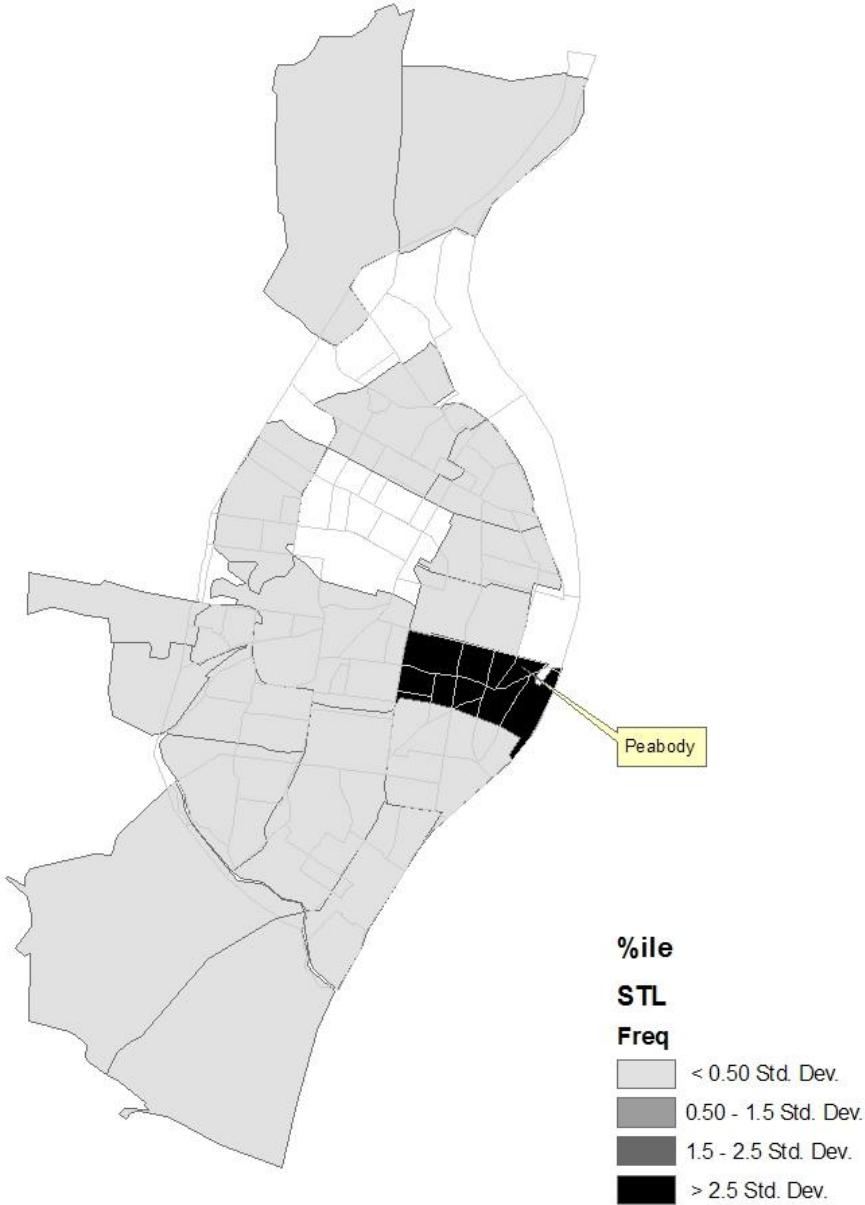
# Spatial Descriptive Statistics

---

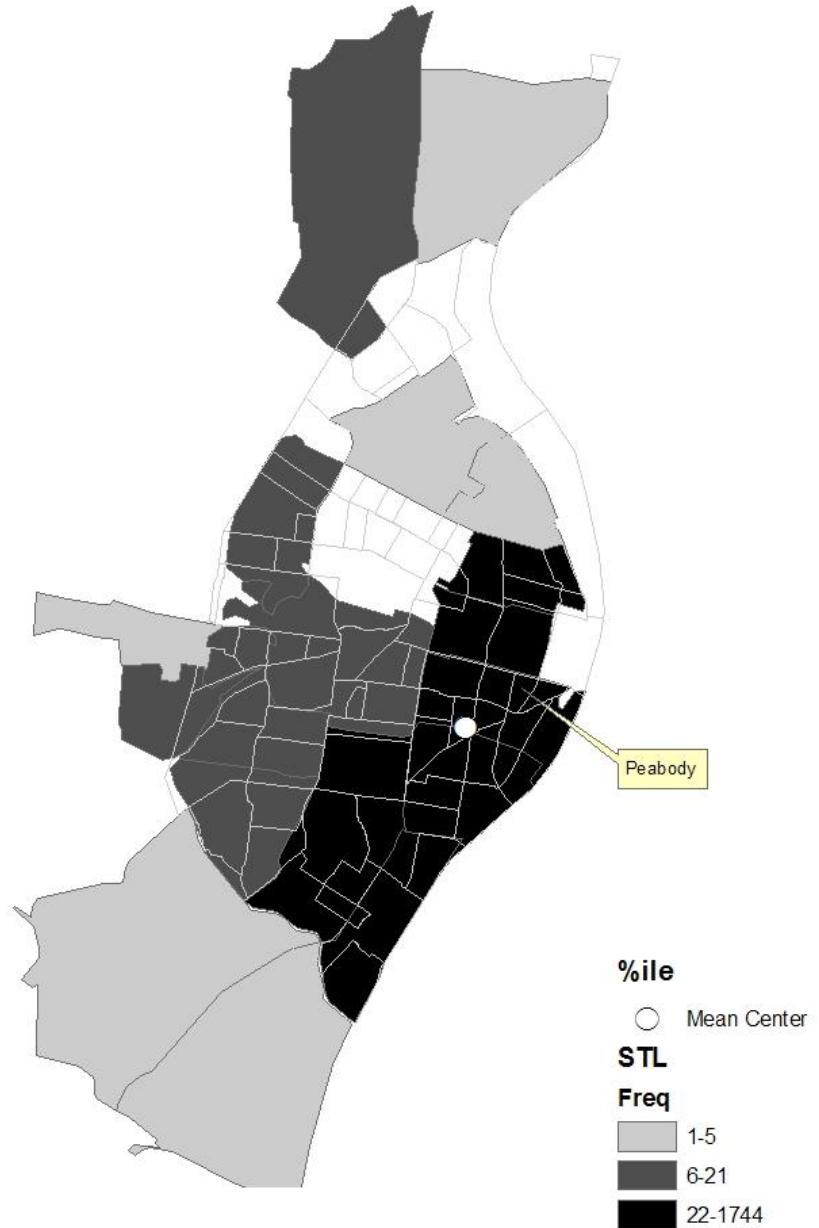
# Spatial Distribution of Kingdom House Clients (2009) Frequency



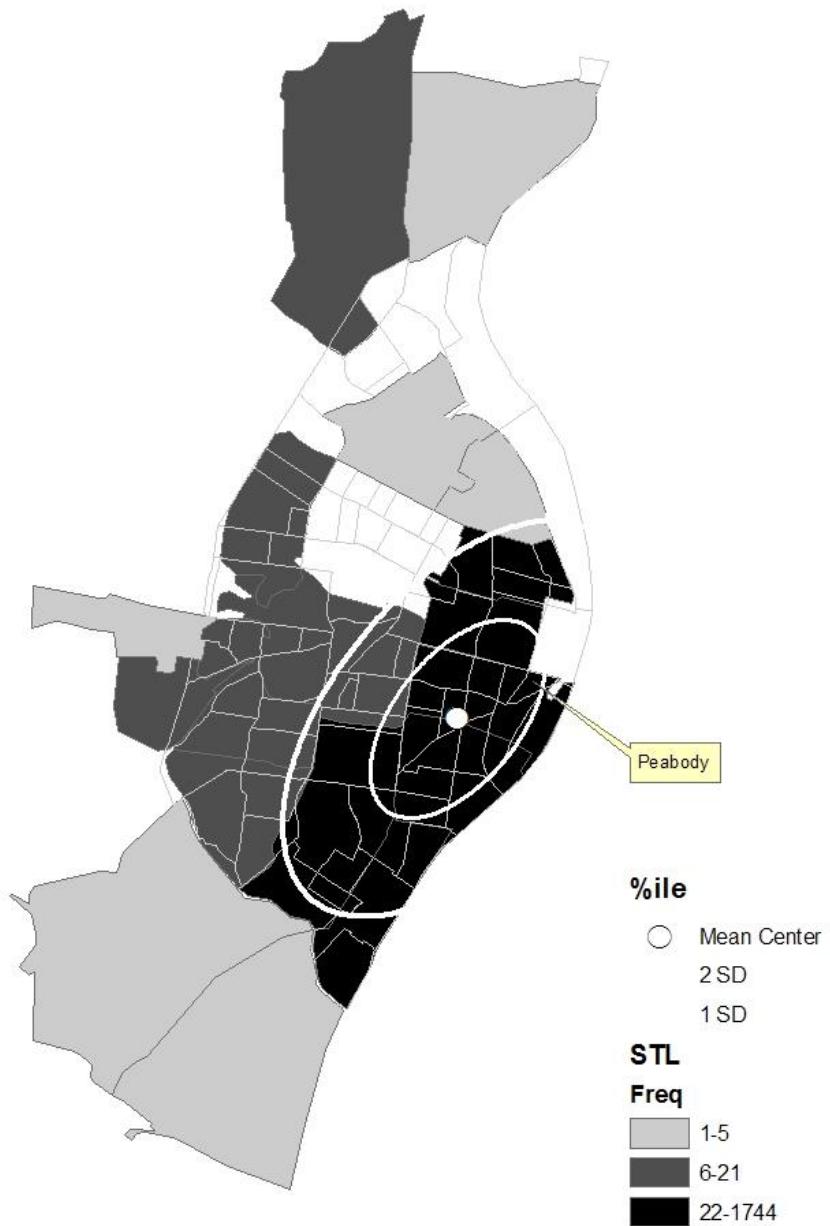
**Spatial Distribution of Kingdom House Clients (2009)**  
**Standard from the Mean**



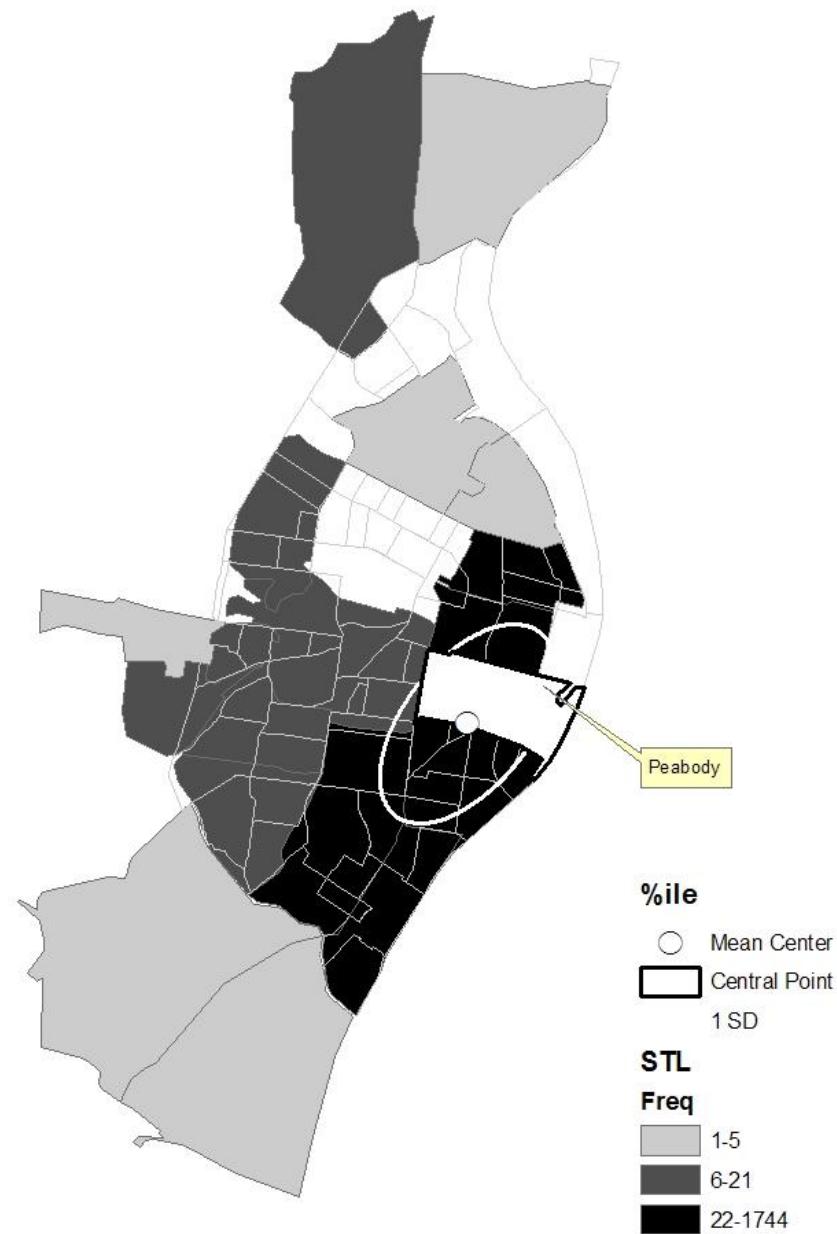
Spatial Distribution of Kingdom House Clients (2009)  
Mean Center



Spatial Distribution of Kingdom House Clients (2009)  
Spatial Direction (2SD)



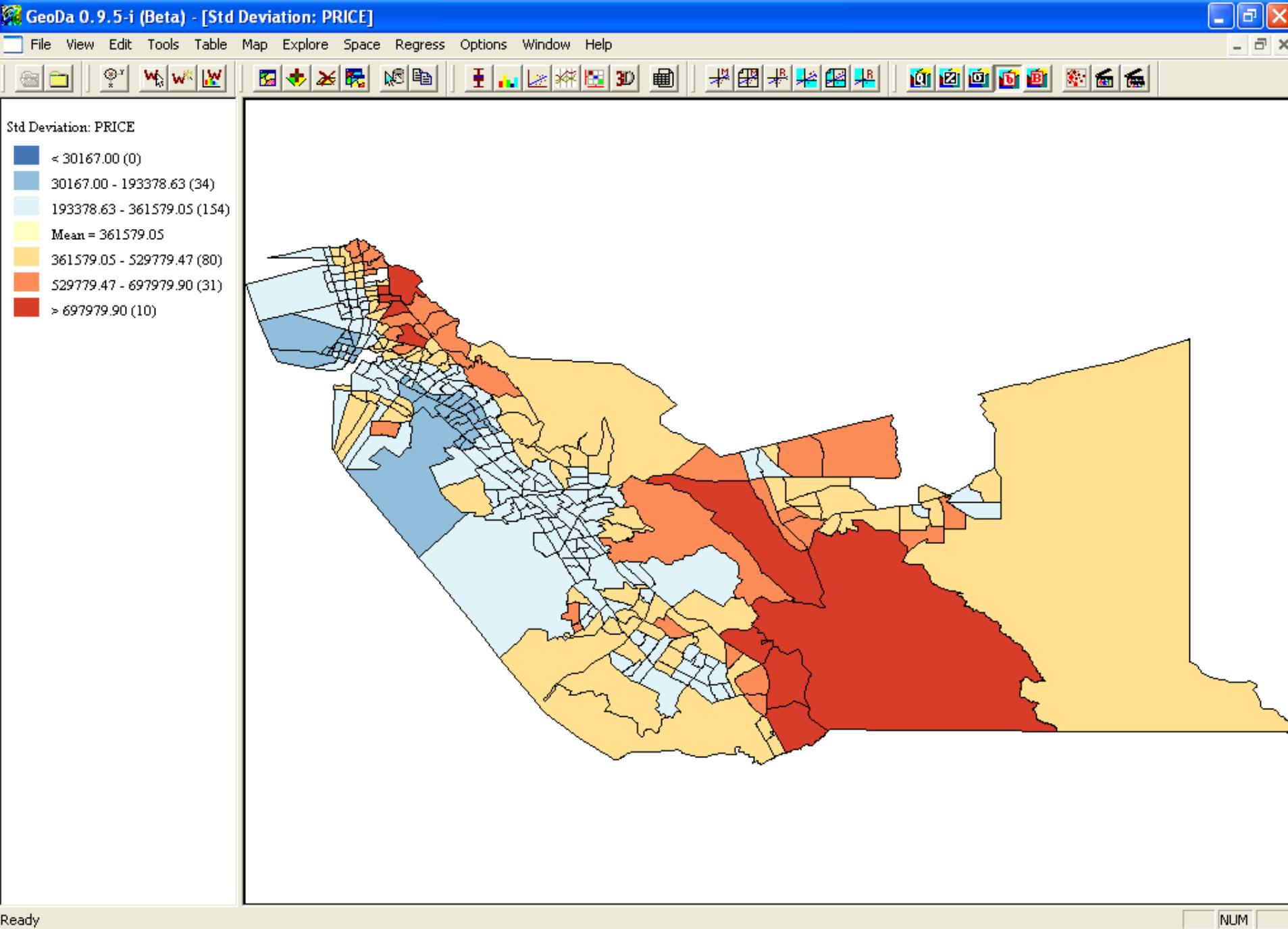
**Spatial Distribution of Kingdom House Clients (2009)**  
**Central Point**

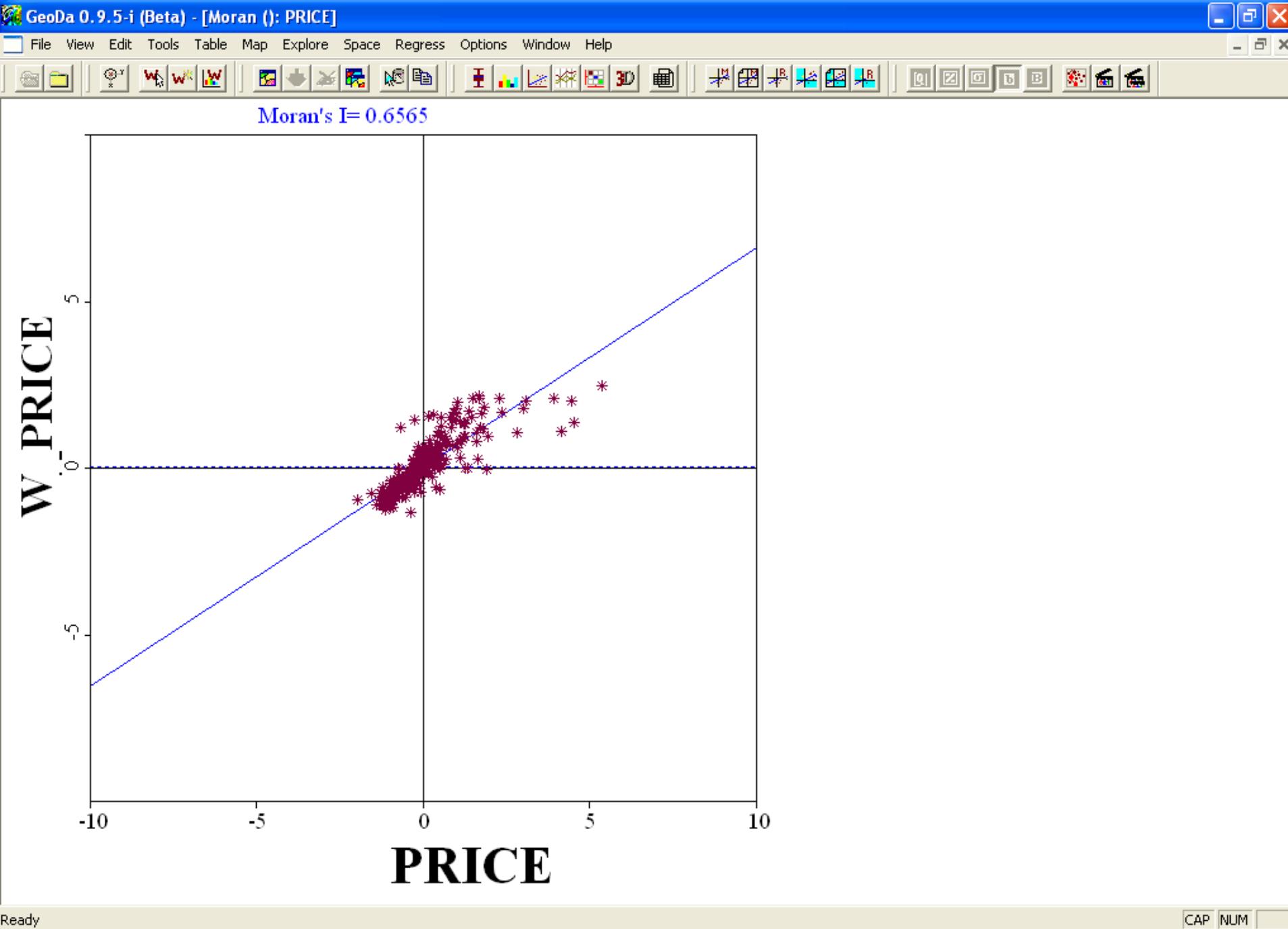


# Spatial Statistics

## Example 1

---





**GeoDa 0.9.5-i (Beta) - [Regression]**

File View Edit Tools Table Map Explore Space Regress Window Help

**REGRESSION**

**SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION**

Data set	:	HOMESALESFINAL
Dependent Variable	:	PRICE Number of Observations: 309
Mean dependent var	:	361579 Number of Variables : 2
S.D. dependent var	:	167928 Degrees of Freedom : 307
R-squared	:	0.209579 F-statistic : 81.4005
Adjusted R-squared	:	0.207004 Prob(F-statistic) : 2.05819e-017
Sum squared residual:	6.88753e+012	Log likelihood : -4119.78
Sigma-square	: 2.24349e+010	Akaike info criterion : 8243.57
S.E. of regression	: 149783	Schwarz criterion : 8251.03
Sigma-square ML	: 2.22897e+010	
S.E. of regression ML	: 149297	

---

Variable	Coefficient	Std. Error	t-Statistic	Probability
CONSTANT	742980.1	43123.72	17.22904	0.0000000
RACE_DIV	-580013.7	64287.23	-9.022223	0.0000000

---

**REGRESSION DIAGNOSTICS**

MULTICOLLINEARITY CONDITION NUMBER 10.02214

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	324.4905	0.0000000

**DIAGNOSTICS FOR HETEROSKEDASTICITY**

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	1	122.2451	0.0000000
Koenker-Bassett test	1	37.7615	0.0000000

SPECIFICATION ROBUST TEST

Ready NUM

**REGRESSION****SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION**

Data set : HOMESALESFINAL  
 Dependent Variable : PRICE Number of Observations: 309  
 Mean dependent var : 361579 Number of Variables : 2  
 S.D. dependent var : 167928 Degrees of Freedom : 307

R-squared : 0.209579 F-statistic : 81.4005  
 Adjusted R-squared : 0.207004 Prob(F-statistic) : 2.05819e-017  
 Sum squared residual: 6.88753e+012 Log likelihood : -4119.78  
 Sigma-square : 2.24349e+010 Akaike info criterion : 8243.57  
 S.E. of regression : 149783 Schwarz criterion : 8251.03  
 Sigma-square ML : 2.22897e+010  
 S.E of regression ML: 149297

---

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	742980.1	43123.72	17.22904	0.0000000
RACE_DIV	-580013.7	64287.23	-9.022223	0.0000000

---

**REGRESSION DIAGNOSTICS**

MULTICOLLINEARITY CONDITION NUMBER 10.02214

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	324.4905	0.0000000

**DIAGNOSTICS FOR HETEROSKEDASTICITY**

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	1	122.2451	0.0000000
Koenker-Bassett test	1	37.7615	0.0000000

SPECIFICATION ROBUST TEST

TEST	DF	VALUE	PROB
White	2	43.52151	0.0000000

**DIAGNOSTICS FOR SPATIAL DEPENDENCE**

FOR WEIGHT MATRIX : QUEENWGT.S.GAL (row-standardized weights)

TEST	MI/DF	VALUE	PROB
Moran's I (error)	0.574094	17.4888140	0.0000000
Lagrange Multiplier (lag)	1	289.1299549	0.0000000
Robust LM (lag)	1	10.8049103	0.0010123
Lagrange Multiplier (error)	1	291.6297134	0.0000000
Robust LM (error)	1	13.3046687	0.0002647
Lagrange Multiplier (SARMA)	2	302.4346237	0.0000000

---

# Spatial Statistics

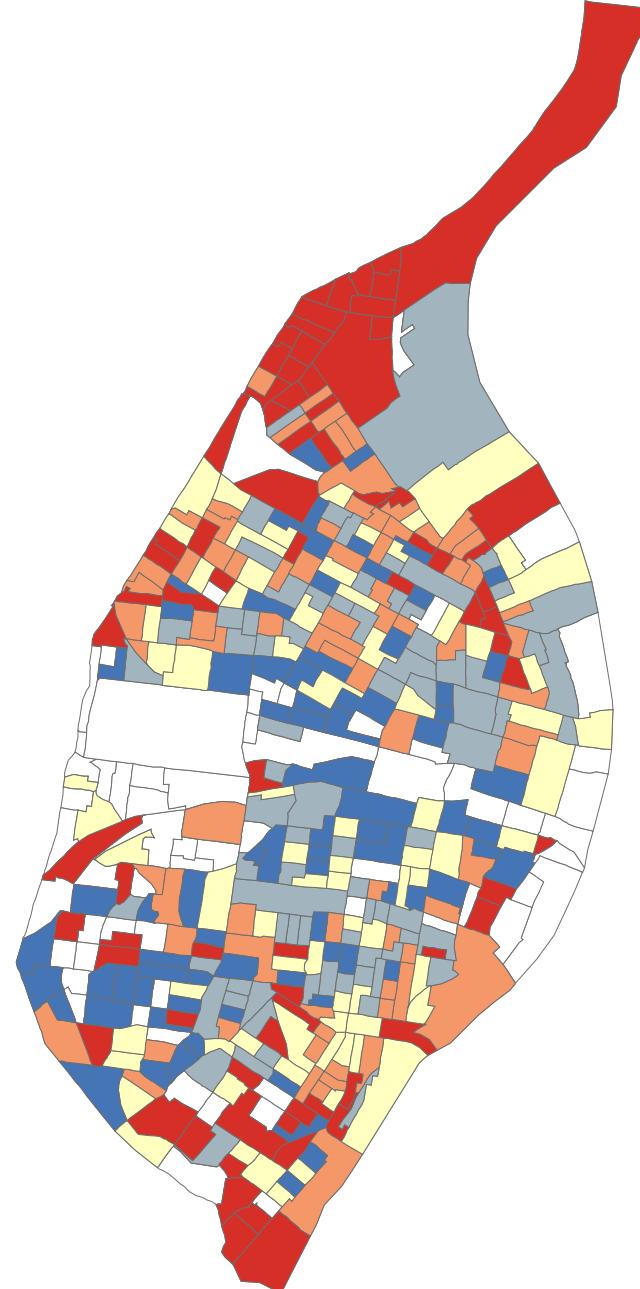
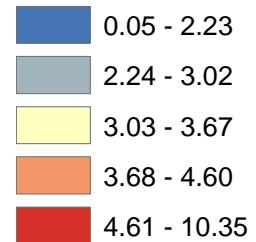
## Example 2

---

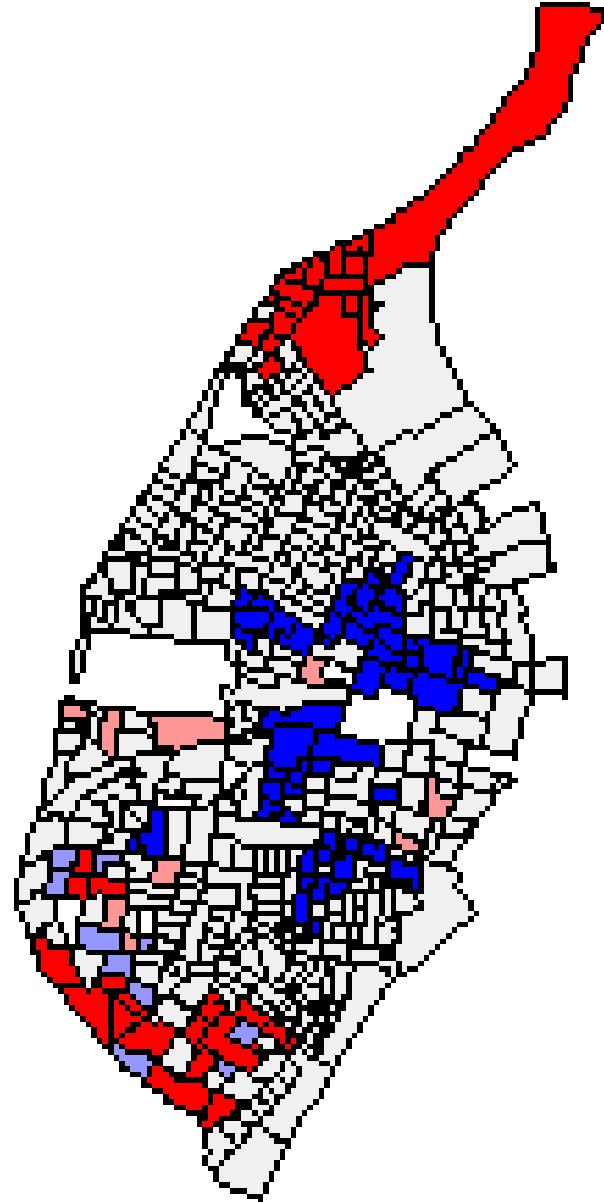
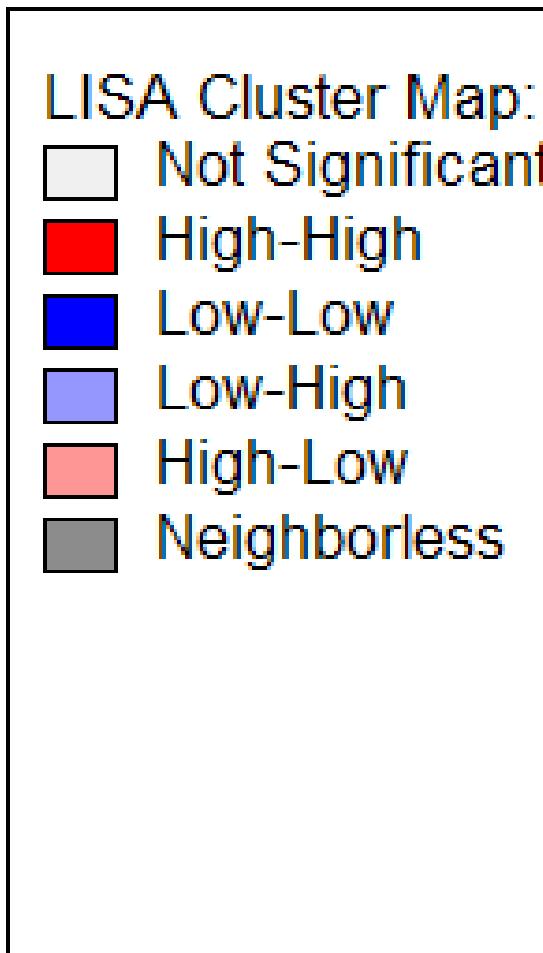
### Census Block Groups

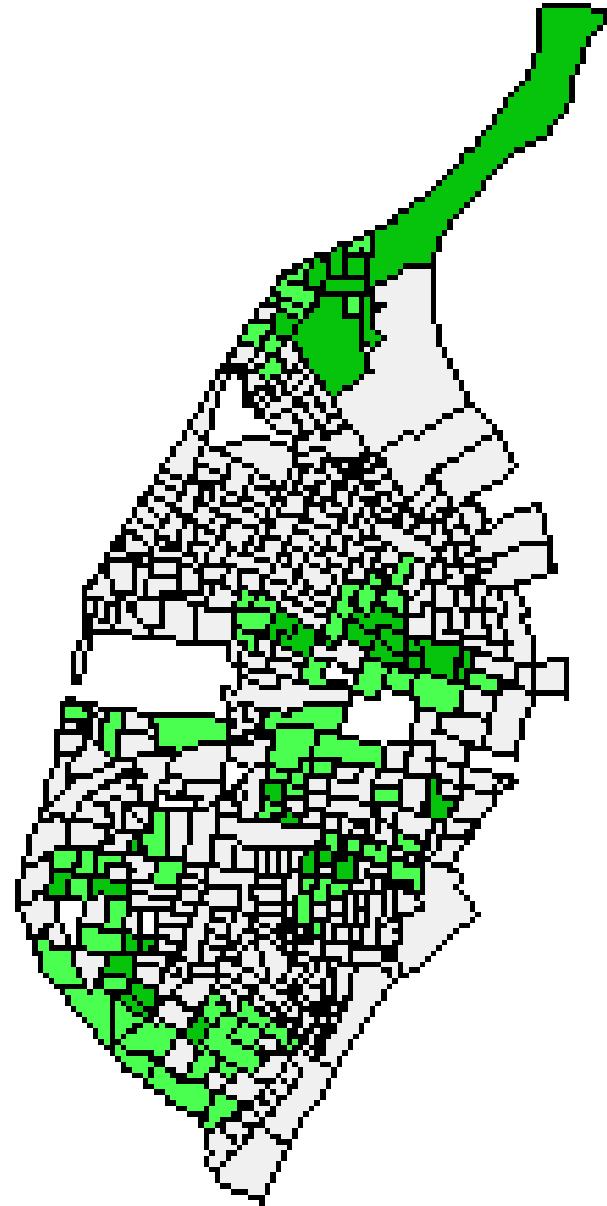
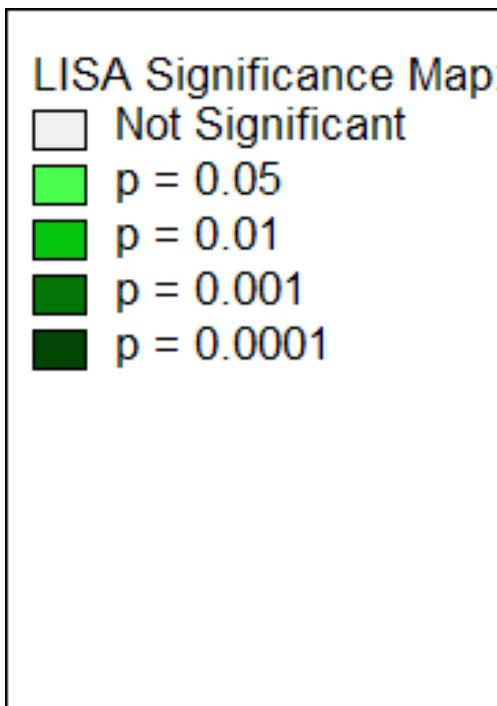
### Stealing - Misdemeanor

### Average Distance



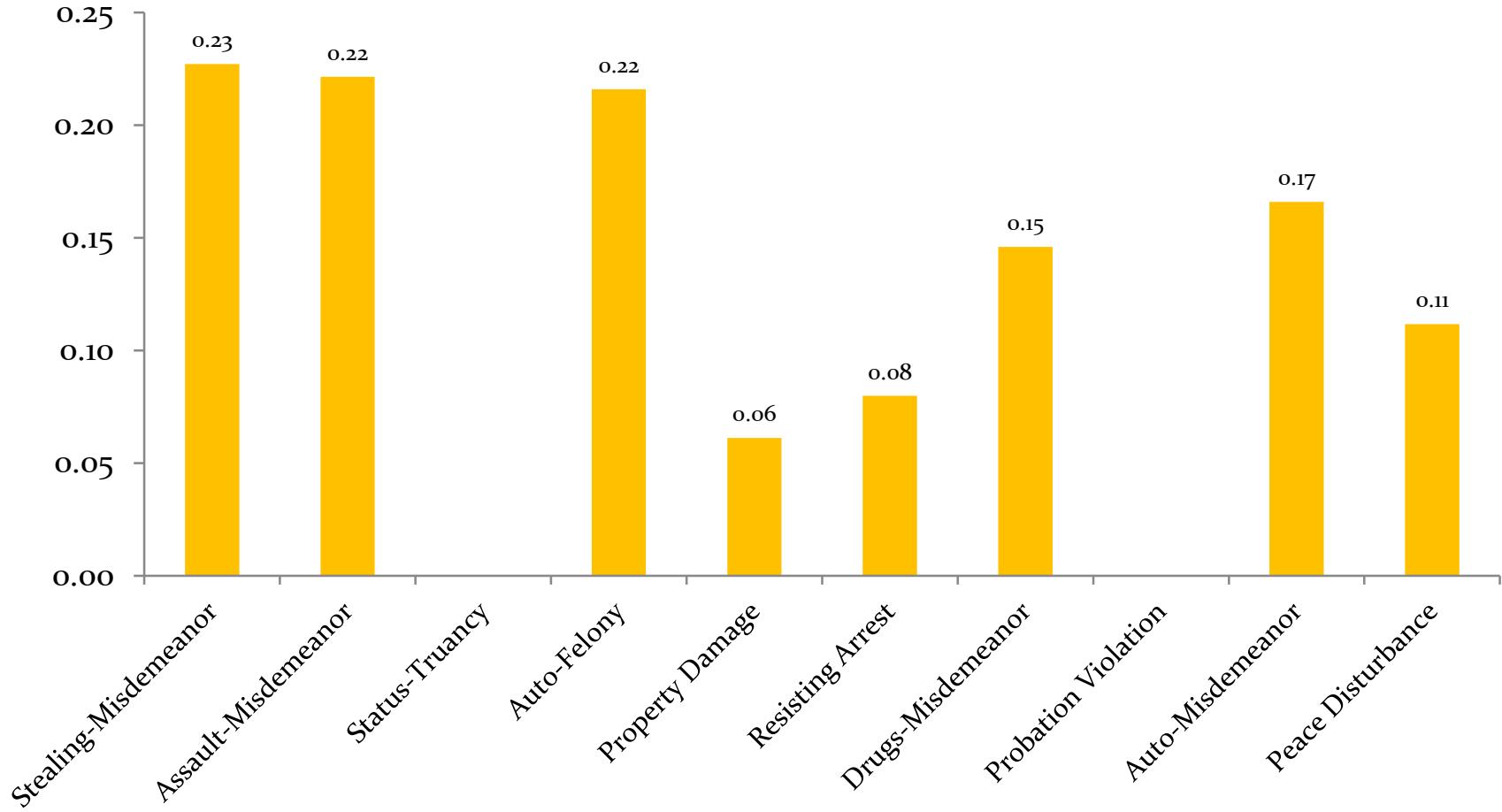
# Moran's I (.2271)





# Moran's I

## Average Distance by Crime Type



# Spatial Statistics

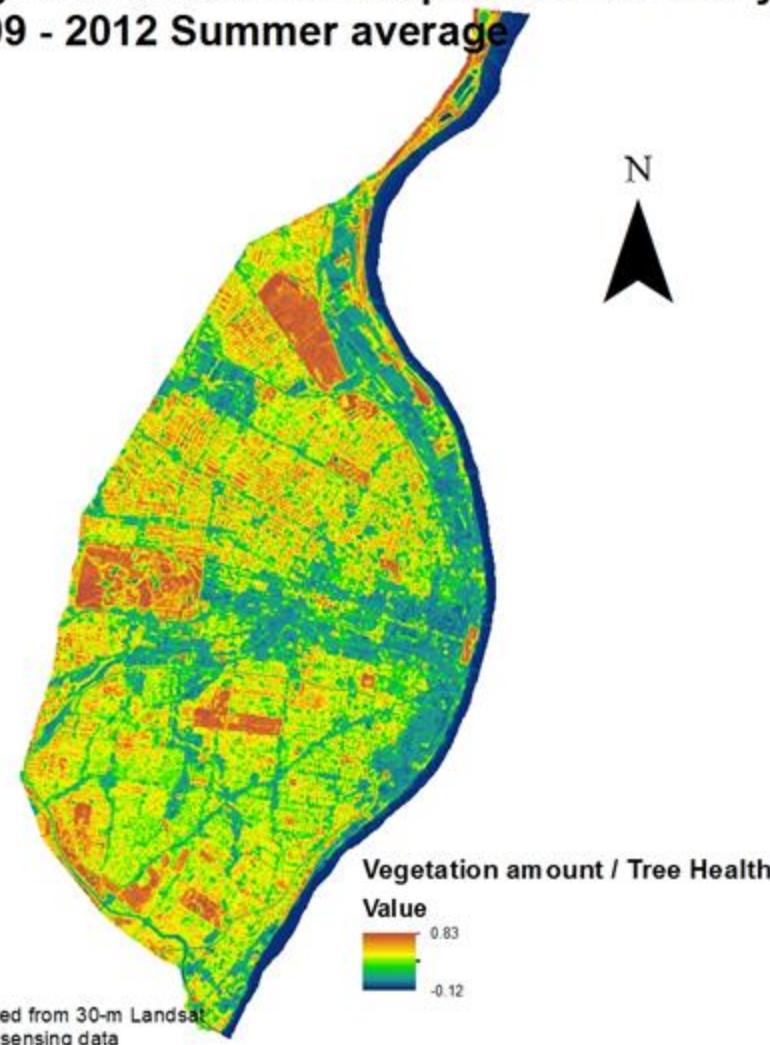
## Example 3

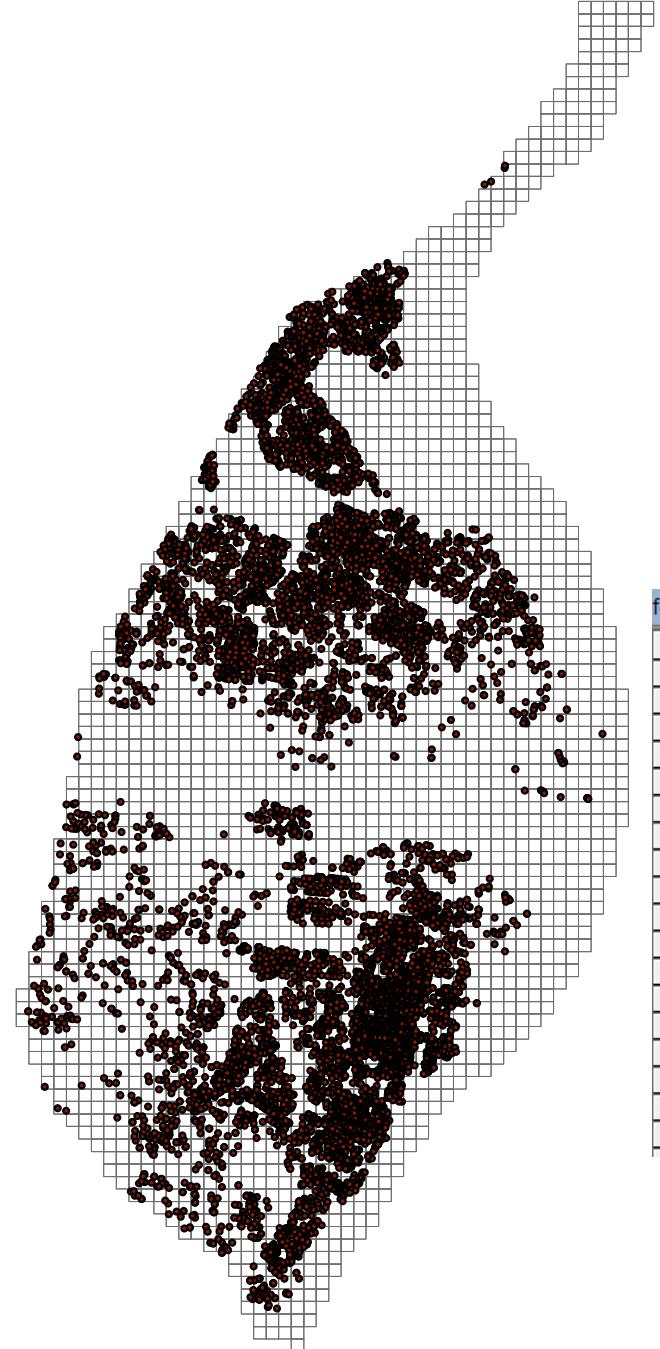
---



# A picture is worth 1000 data points

Vegetation condition map of St. Louis City  
2009 - 2012 Summer average





# Point Data

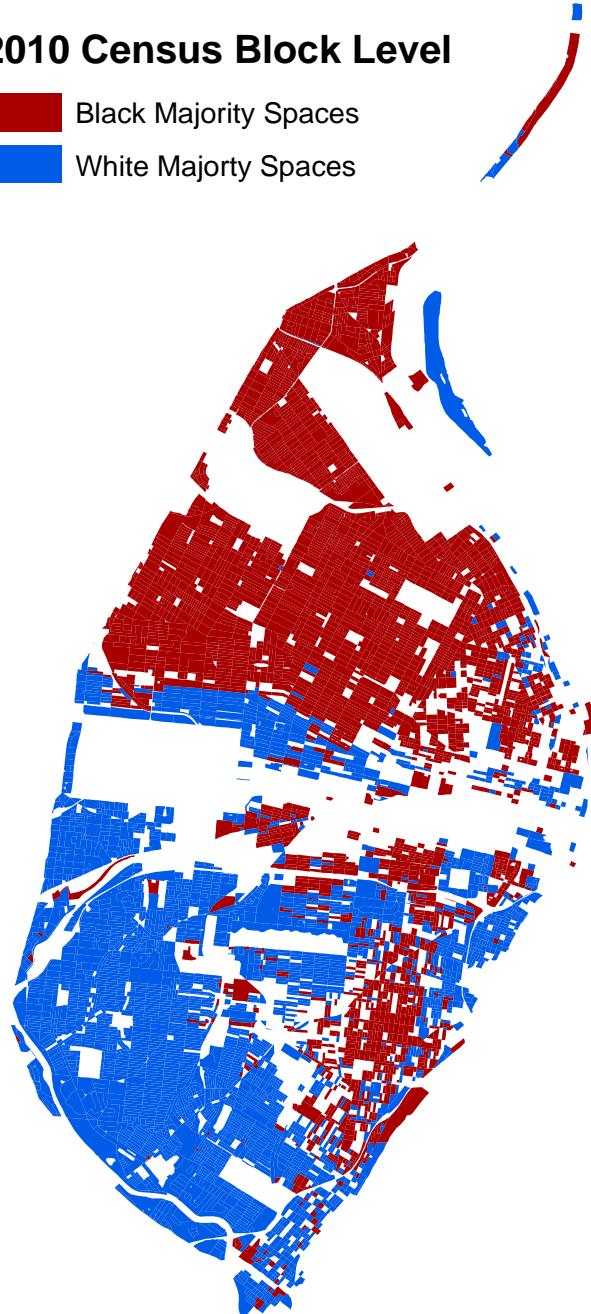
foreclosureexperiment

FID	Shape *	ParcelId	SalePrice	salesdate	X	Y
8558	Point	49070202400	790000	11/30/2007	566886.145311	1025296.37418
6266	Point	38800003300	760500	11/25/2008	567975.841347	1024350.43386
11468	Point	58490000100	658355	7/15/2008	555979.115773	1023121.88425
42	Point	07890000300	627230	4/30/2008	582891.971193	1009621.35558
4931	Point	31020000700	612774	8/20/2007	567323.728629	989409.337702
9298	Point	51870000500	576000	4/4/2005	578048.061043	1029732.75312
10961	Point	56390001500	571500	1/18/2007	555888.467177	1021851.87888
5578	Point	36670001100	558637	12/23/2005	576224.96913	1029049.16243
8170	Point	45800501400	555819	3/23/2009	570507.025897	1024321.43424
6337	Point	39900500200	500000	2/28/2006	570922.000981	1016179.85815
6338	Point	39900500300	500000	2/28/2006	570864.938914	1016249.90417
6339	Point	39900500701	500000	2/28/2006	570957.63763	1016354.26532
8171	Point	45800501600	495000	1/25/2005	570399.874148	1024381.09759
4381	Point	27850002400	489900	8/11/2006	574932.843928	995866.064826
8164	Point	45770000800	480345	7/28/2006	569199.25991	1024808.43307
8163	Point	45760002700	479006	11/22/2006	569180.521064	1025053.10013
2772	Point	18050001560	467500	9/11/2008	582268.184873	1012531.56814
184	Point	13100001600	465913	5/30/2008	577121.257494	1012030.68626

# Polygon Data

2010 Census Block Level

Black Majority Spaces  
White Majority Spaces



FID	Shape *	OBJECTID	AREAKEY	P0050001	pct_wht	pct_blk
977	Polygon	978	295101025002001	39	0.641	0.231
5905	Polygon	5906	295101174003010	39	0.641	0.231
9502	Polygon	9503	295101276003036	39	0.641	0.256
8138	Polygon	8139	295101268001023	25	0.64	0.04
4716	Polygon	4717	295101152003004	36	0.639	0.111
4983	Polygon	4984	295101155005003	36	0.639	0.278
6303	Polygon	6304	295101192001023	36	0.639	0.111
7119	Polygon	7120	295101243003043	36	0.639	0.333
1003	Polygon	1004	295101025002027	80	0.638	0.175
3746	Polygon	3747	295101121002007	343	0.638	0.227
4886	Polygon	4887	295101154002018	224	0.638	0.223
4590	Polygon	4591	295101151002010	91	0.637	0.297
261	Polygon	262	295101014001017	66	0.636	0.182
1302	Polygon	1303	295101038001012	22	0.636	0.136
1534	Polygon	1535	295101045002019	11	0.636	0
1629	Polygon	1630	295101051982013	33	0.636	0.303
4837	Polygon	4838	295101153005023	22	0.636	0.136
4941	Polygon	4942	295101155001009	88	0.636	0.33

## Social (N=22)

Dependency Ratio

Crime

Education

Cooling, Heating, and Food Centers

Grocery Stores

Vacant Land and Home

Infant Health

Abuse and Neglect

Single with Children

Condemnations

## Economic (N=8)

Under Banked

Distance to Banks

Poverty

Public Assistance

Family Income

Per Capita Income

Homeownership

Foreclosures

## Environment (N=3)

Lead Poisoning

Brownfields

Access to Parks

## Ecology (N=5)

Vegetation

Tree Cover

Tree Health

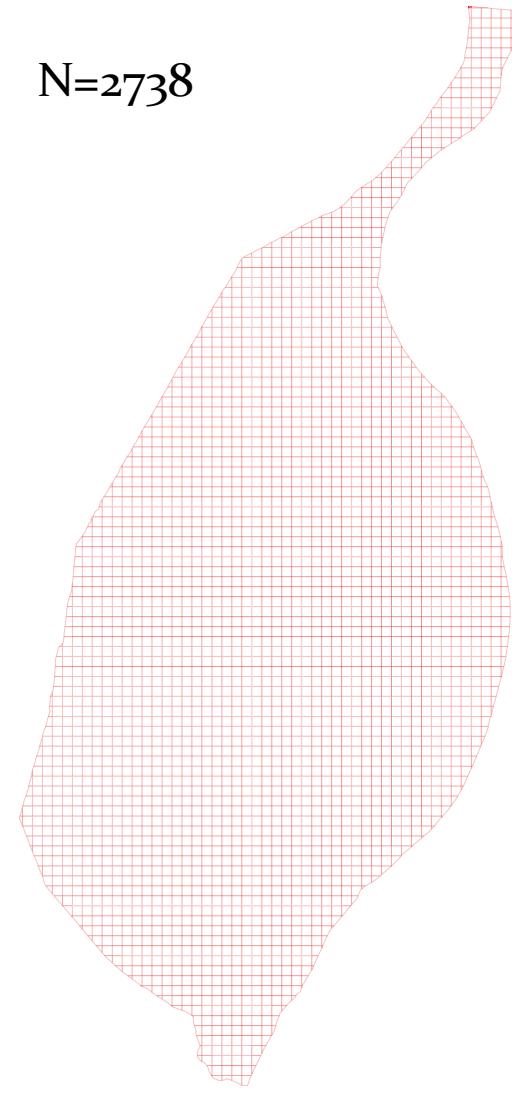
Geological Hazard

# Synthesis is Spatial

N=360



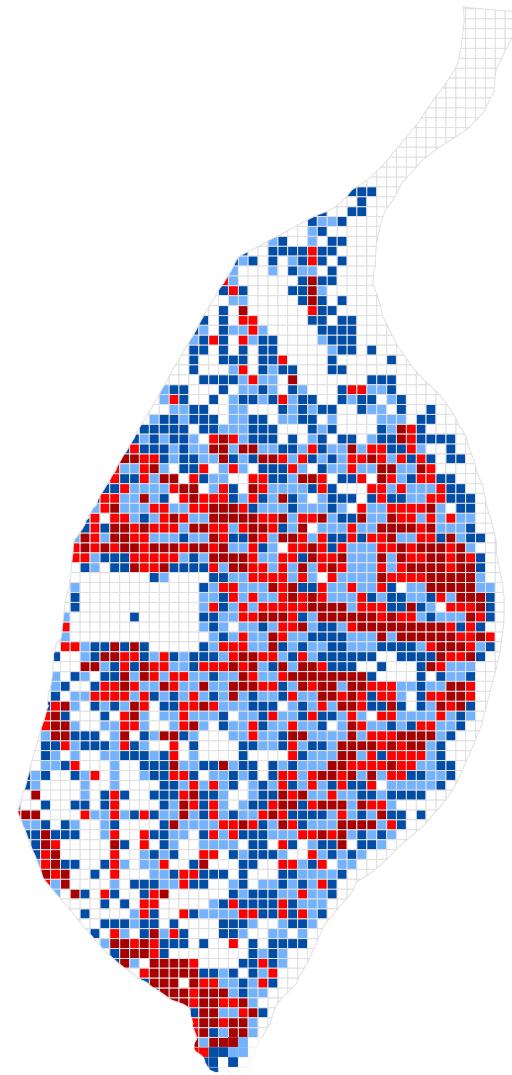
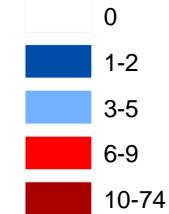
N=2738



# Brownfields



Saint Louis  
bronwfield

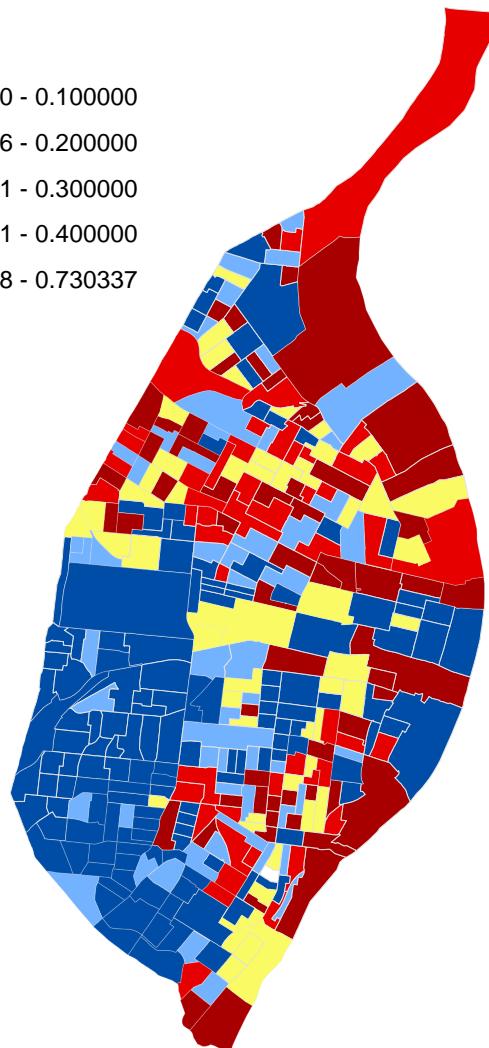


# Poverty Rate

Saint Louis

povrate

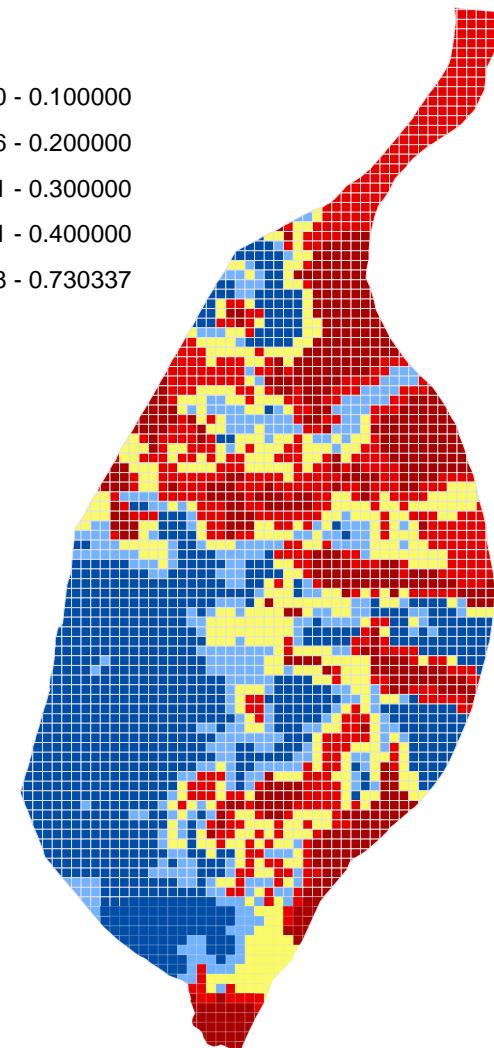
- 0.000000 - 0.100000
- 0.101356 - 0.200000
- 0.200381 - 0.300000
- 0.300041 - 0.400000
- 0.400788 - 0.730337



Saint Louis

poverty\_ra

- 0.000000 - 0.100000
- 0.101356 - 0.200000
- 0.200381 - 0.300000
- 0.300041 - 0.400000
- 0.400788 - 0.730337



# Index

Conceptual Formula for the Social-Environmental Synthesis Risk Index

$$(1) V_i = \left( \frac{X_j - Y_i}{Z_i - Y_i} \right)$$

$V_i$  is the standarized index score for each grid cell

$X_j$ = actual value for a grid cell

$Y_i$ =minimum value in the universe of grid cells

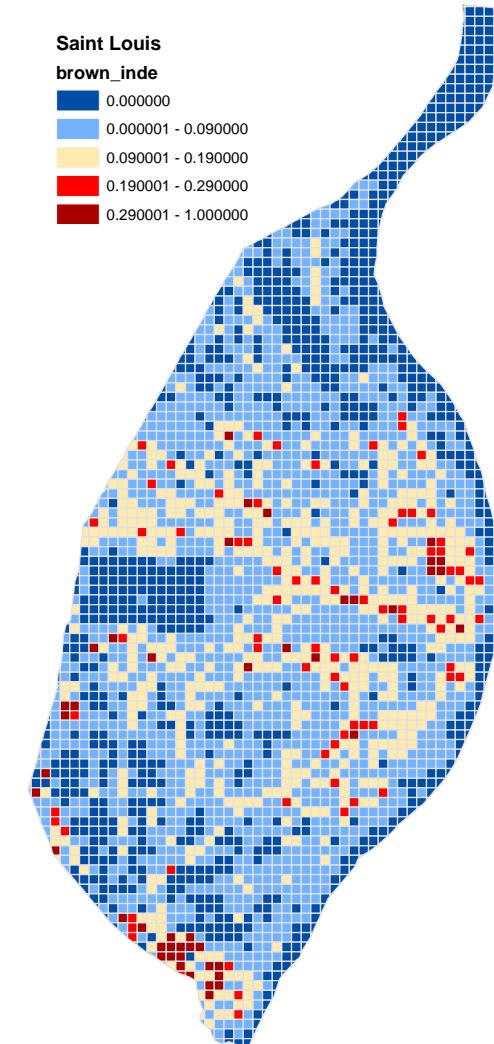
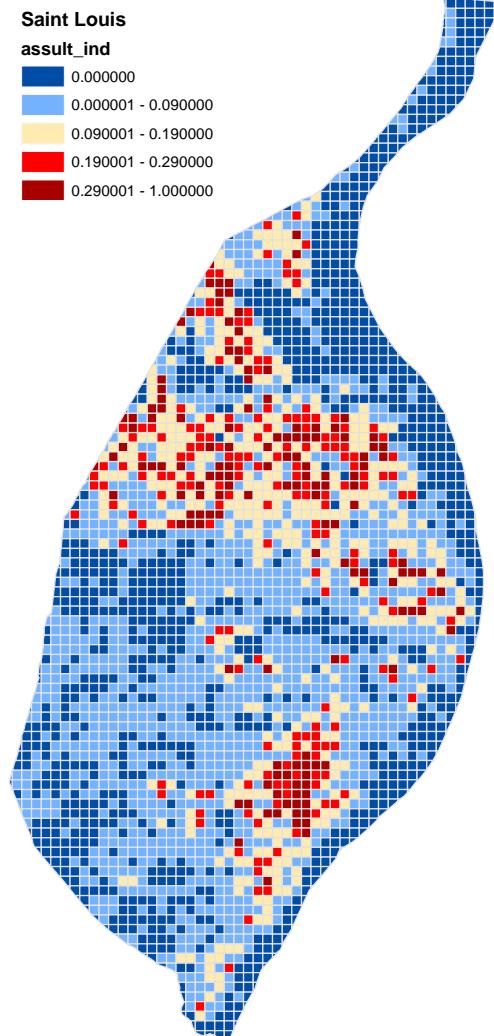
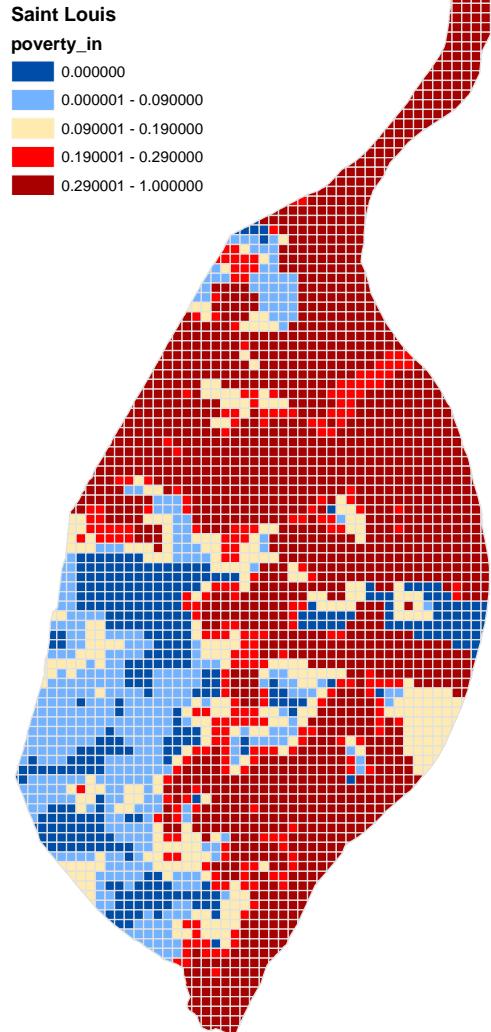
$Z_i$ =maximum value in the universe of grid cells

$$(2) S_i = \frac{\sum_{i=1}^N V_i}{N}$$

$S_i$  is the SER index sore for each grid cell

$N$  is the number of dimensions

# Three Dimensions

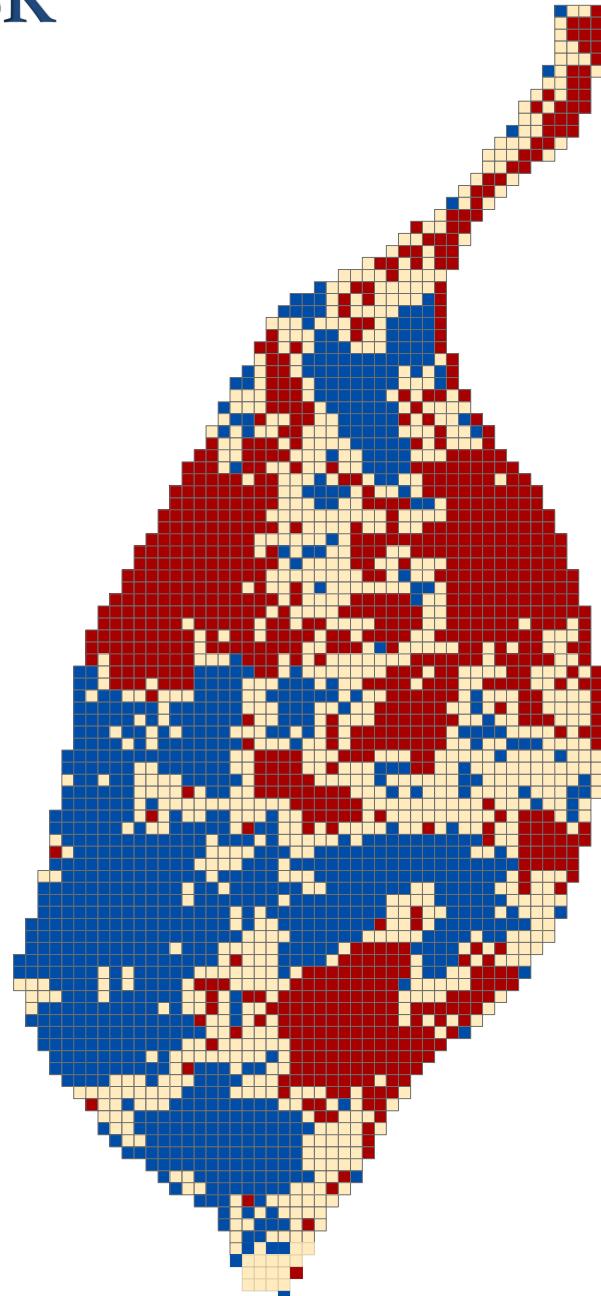


# Social Environmental Risk Index

## Social Environmental Risk Index



Moran's I  
.8127  
p value <.001



# Social Environmental Risk Index

## Social Environmental Risk Index

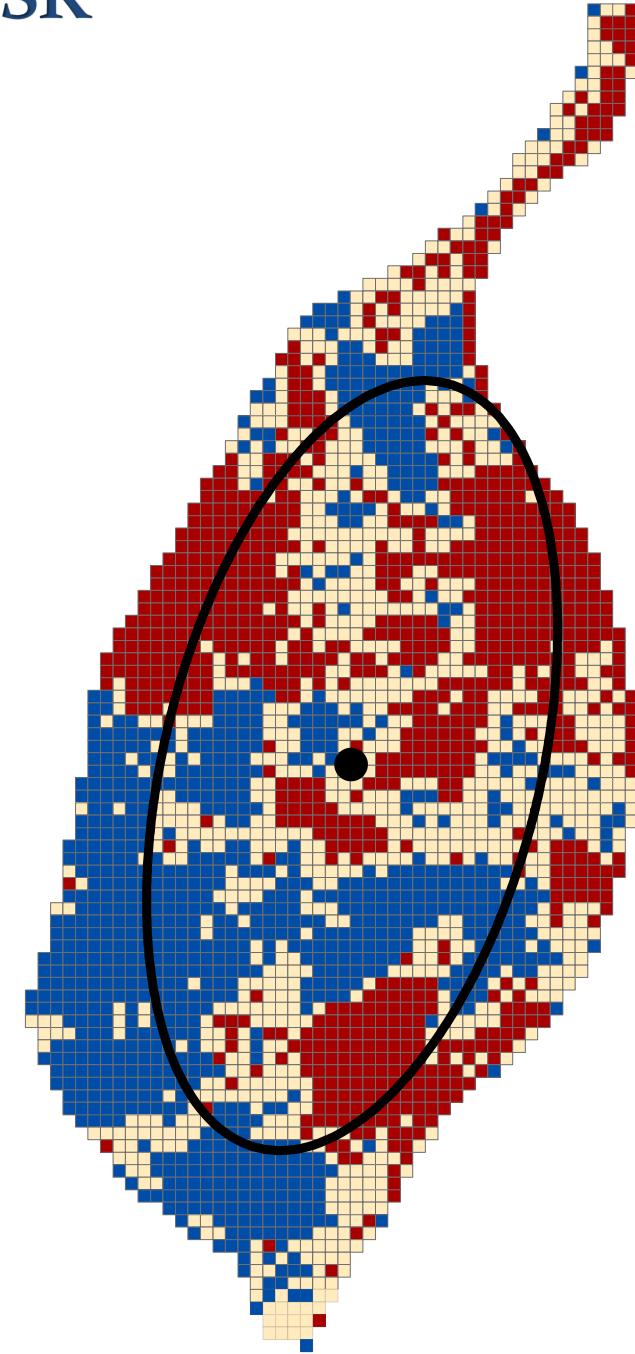
● Mean Center

◻ Standard Deviational Ellipse

■ 0.236 - 0.370

■ 0.371 - 0.409

■ 0.410 - 0.528



# Spatial Models

---

# Regression Models

$$(1) Y_{risk} = B_0 + B_1 X_{perwhite} + B_2 X_{perblack} + B_3 X_{perpoploss} + error$$

$$(2) Y_{risk} = B_0 + \rho w_i y_i + B_1 X_{perwhite} + B_2 X_{perblack} + B_3 X_{perpoploss} + error$$

$$(3) Y_{risk} = B_0 + B_1 X_{perwhite} + B_2 X_{perblack} + B_3 X_{perpoploss} + \lambda W\mathcal{E} + \xi$$

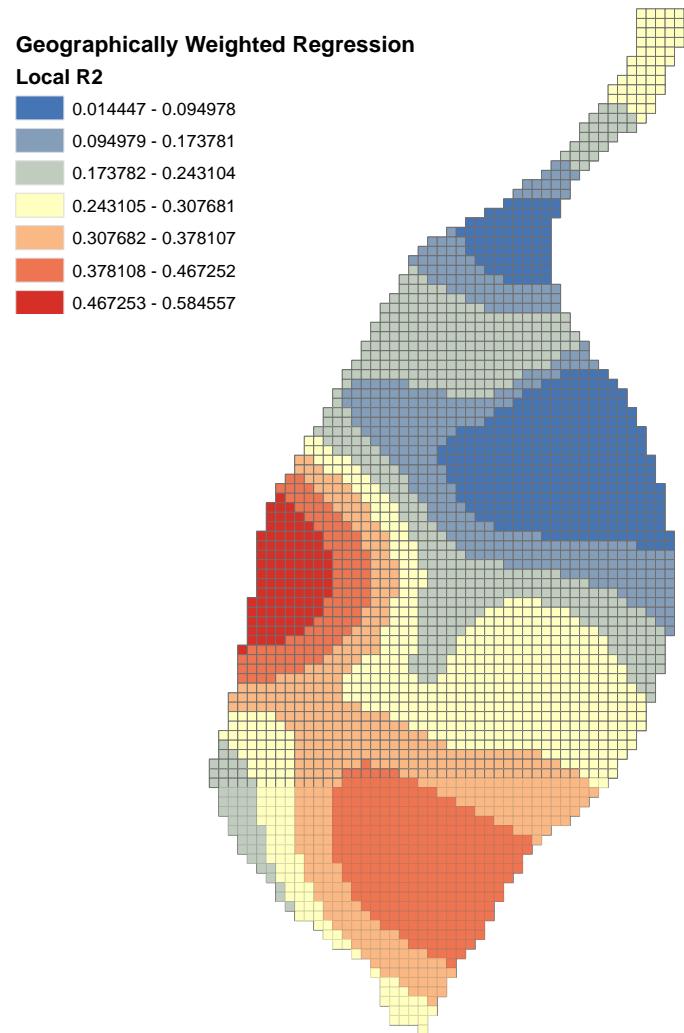
	OLS	SLM	SEM
Percent White	-0.05905254 (0.002552181)***	-0.006379363 (0.001311031) ***	-0.01383606 (0.003214683) ***
Percent Black	0.03196312 (0.002320322) ***	0.005242719 (0.001177692) ***	0.02557729 (0.003307808) ***
Percent Population Loss	-0.0002005318 (0.0002589599)	-7.434595e-005 (0.000122794)	-0.0003127578 (0.0002814473)
Spatial Lag $\rho$		0.9178052 (0.008517798) ***	
Spatial Error $\lambda$			0.9342123 (0.007792058) ***
Constant	0.3964256 (0.001279275)	0.03215798 (0.003386138)	0.3865446 (0.005246035)
Log likelihood	5113.64	6888.78	6896.02
AIC	-10219.3	-13767.6	-13784

\* ≤ .05, \*\* ≤ .01, \*\*\* ≤ .001

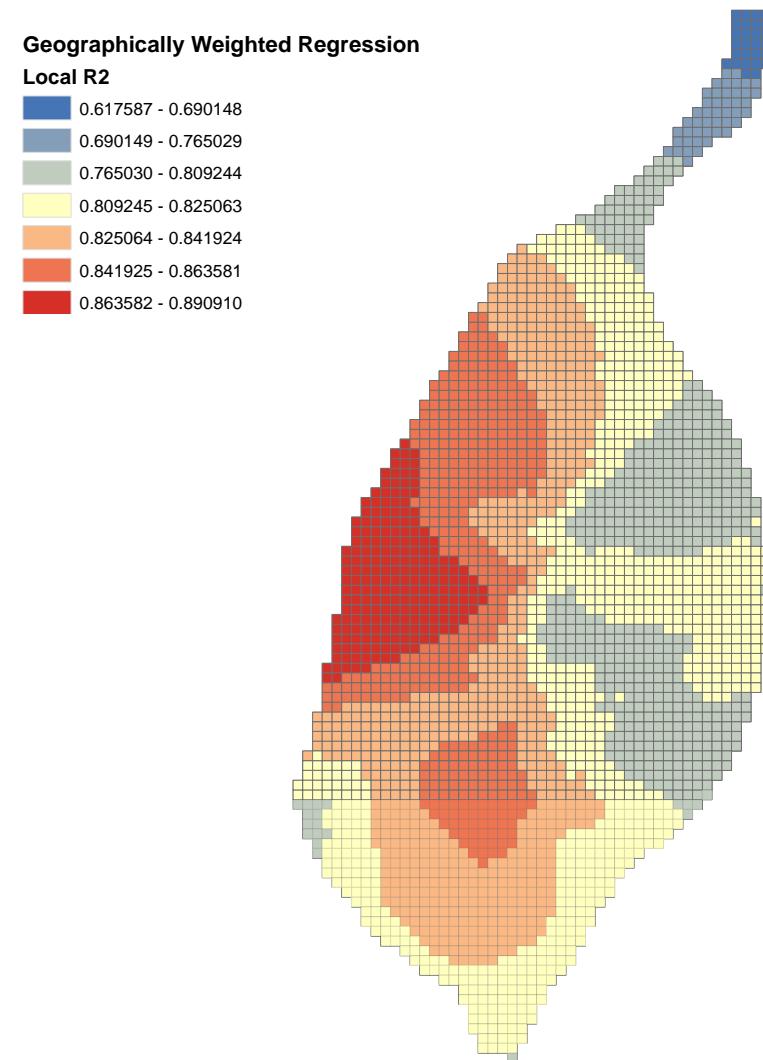
Standard Errors in Parentheses

# Geographically Weighted Regression

Race and Population Model



Spatial Lag Model



# Introductions of Students

---