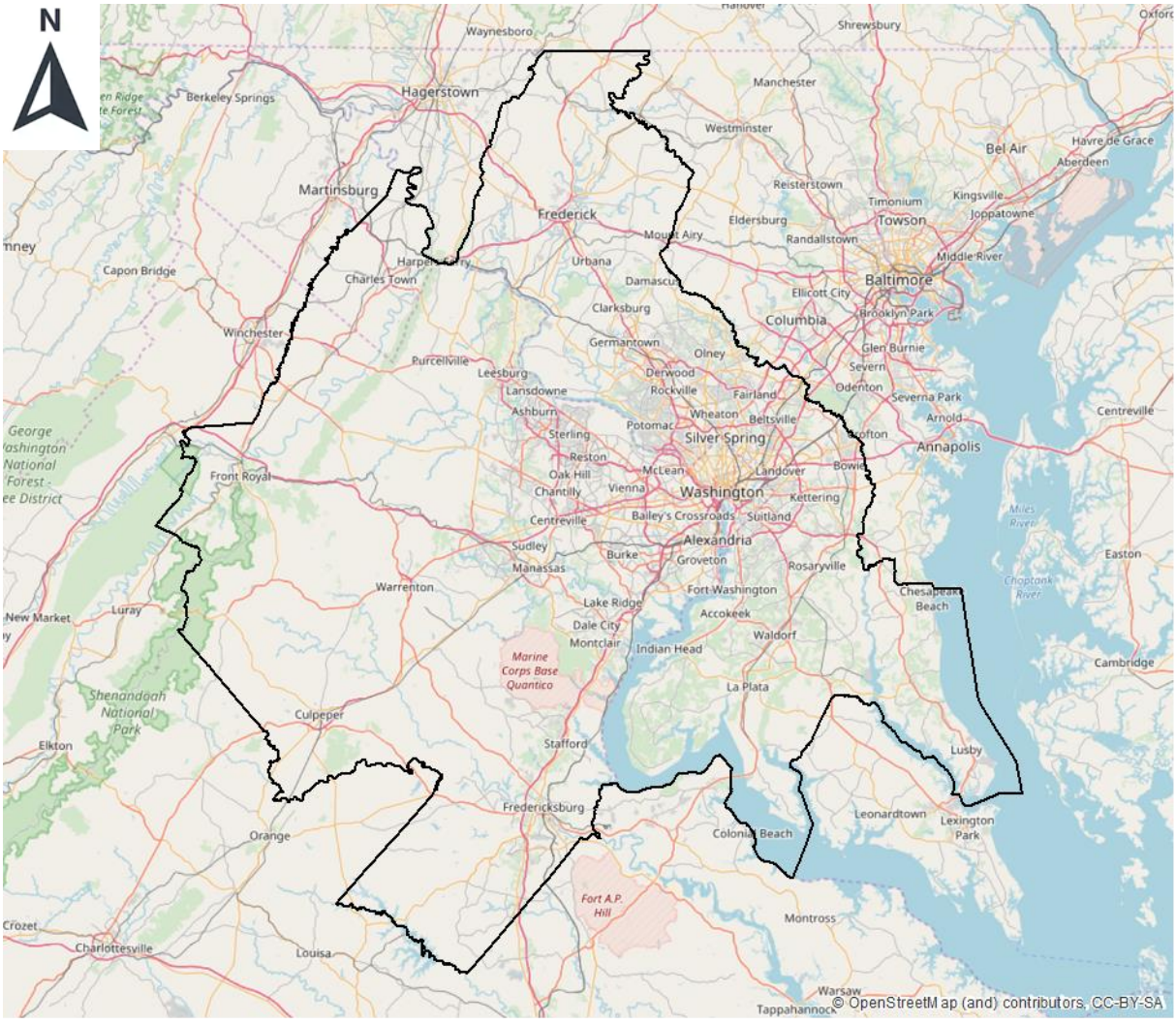
Lab 1: Exploring Income Inequality in the Washington D.C. Core Based Statistical Region

Reference Map: Washington D.C. CBSA



Projected Coordinate System: NAD 1983 StatePlane Maryland FIPS 1900 (Meters)

Data obtained from US Census TIGER Database and OpenStreetMap

This lab examines the Washington D.C. Core Based Statistical Area (Reference Map) using data sourced from the American Community Survey 2012-2016, 5 Year-Estimates on Social Explorer. Six variables are included in this exploration: GINI index, percentage of male population 25 years and older with a bachelor's degree or higher (e01), percentage of female population 25 years and older with a bachelor's degree or higher (e02), unemployment rate, percent foreign born, and percent minority. All six variables are measured by census tract.

Figure 1

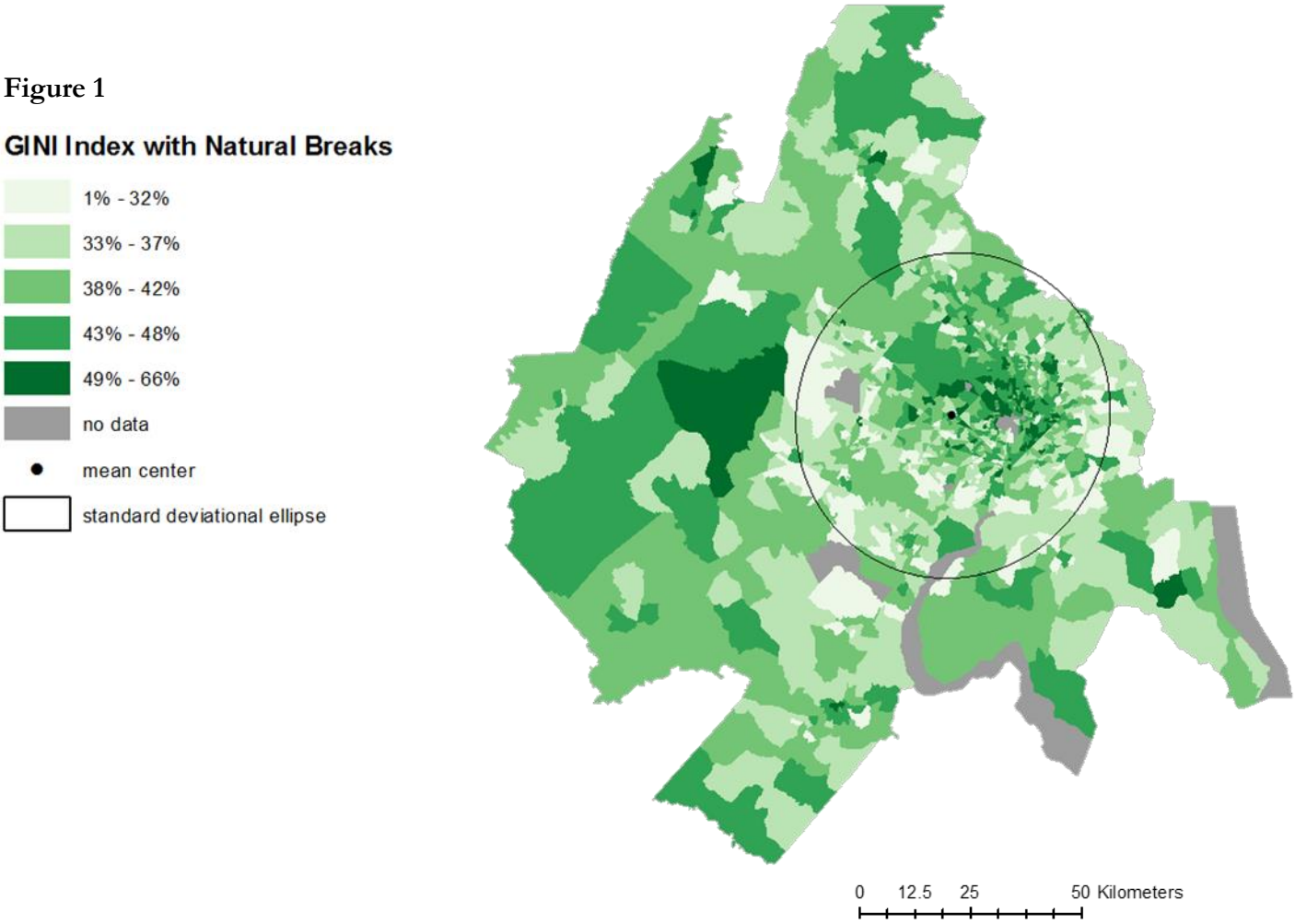


Figure 2

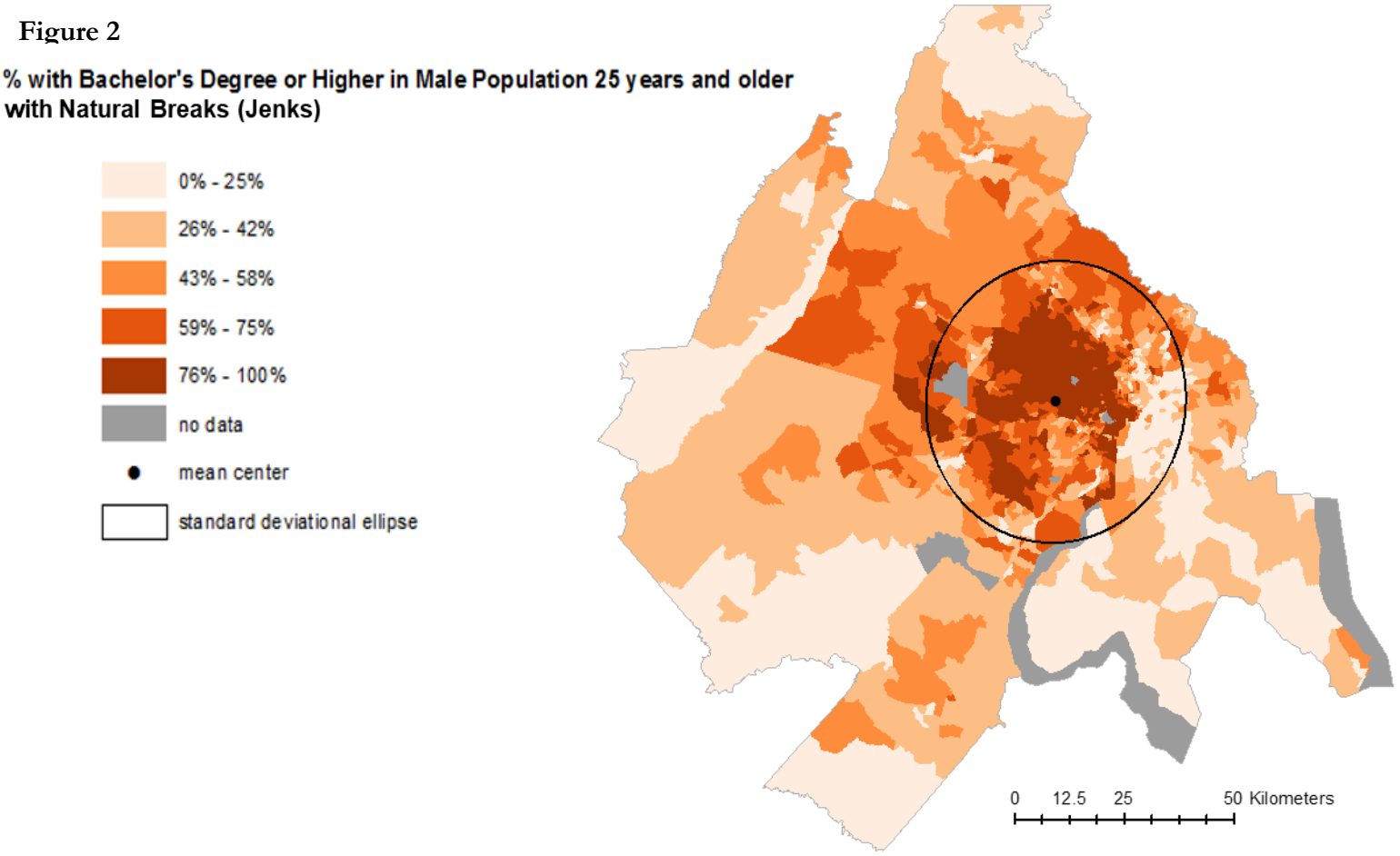


Figure 3

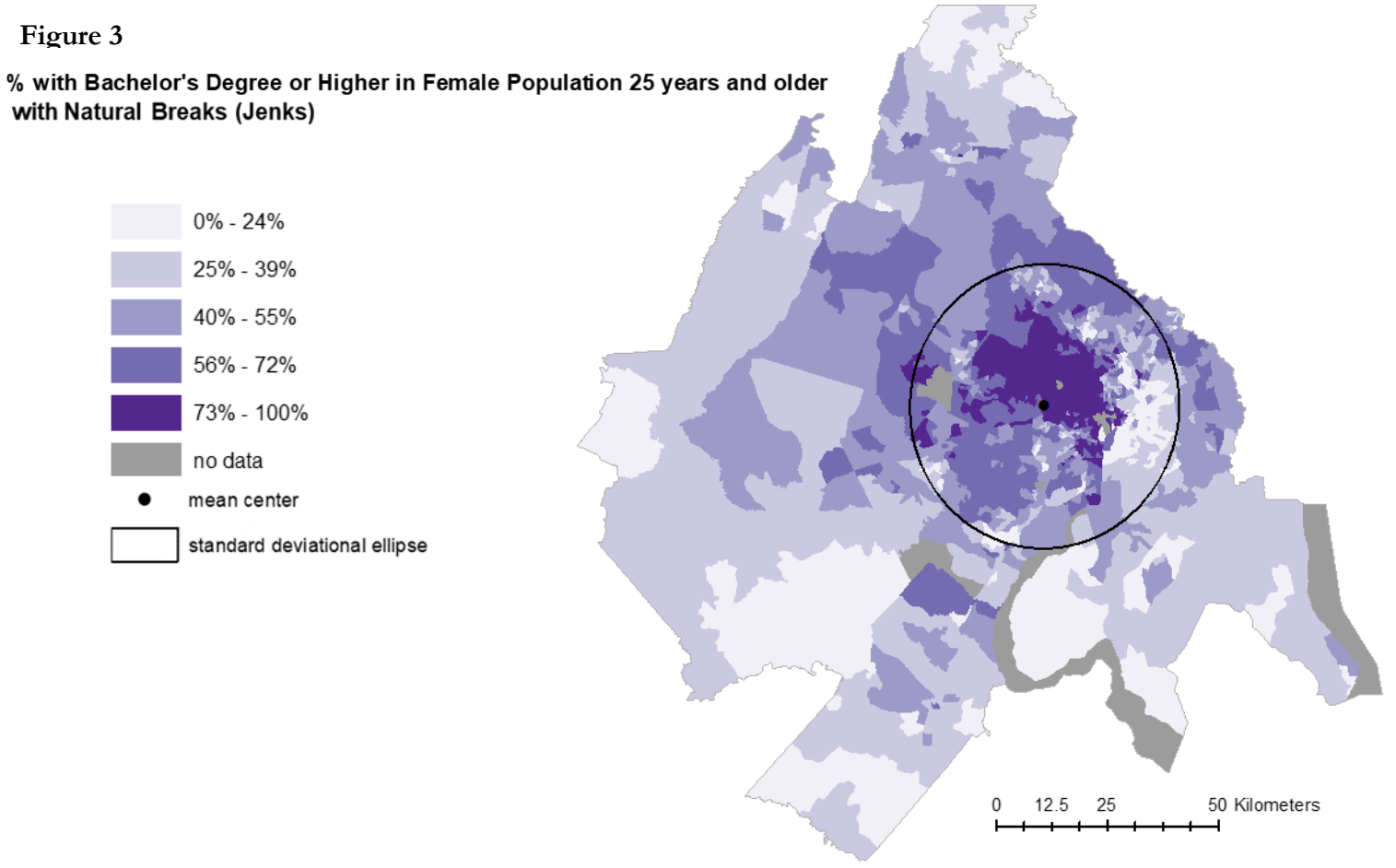


Table 1. Descriptive Spatial Statistics for the Washington D.C CBSA Metropolitan Region

Variable	Name	GINI Index	% with Bachelor's Degree or Higher in Male Population 25 years and older	% with Bachelor's Degree or Higher in Female Population 25 years and older
Mean	Mean Center - X Coord	383601.4	382838.9	383631.4
	Mean Center - Y Coord	136385.0	138048.1	138055.0
Directional Distribution	Angle Rotation	24.6	9.7	5.6
	X Axis Length (m)	34759.0	29391.5	30147.3
	Y Axis Length (m)	36625.3	32027.4	31887.8
	Area of the Ellipse (m^2)	3999225786.5	2957135961.7	3019951290.7
Standard Distance	Standard Distance (m)	35704.3	30737.7	31029.7
	Area (m^2)	4004697587.9	2968051298.3	3024710414.5

Source: American Community Survey 2012-2016, Table created by Emma Blackwood

Figures 1-3 are choropleth maps for three variables in the Washington D.C. CBSA. The projected coordinate system used for all the maps is the NAD 1983 StatePlane Maryland FIPS 1900 (Meters). Mean centers and standard deviational ellipses were derived using ArcMap and based on 1 standard deviation.

While the mean centers and standard deviational ellipses for all three variables share similar areas, the choropleth maps (Figure 1-3) reveal varying spatial patterns for GINI index and the education variables. All the mean centers are located near the core of the city and slightly towards the west. The GINI index by census tract is higher near the core of the city as well as in some of the tracts further away from the center. The suburbs around the city have a low GINI index, indicating most of these census tracts have a similar income levels within each tract. The highest percentages of men and women 25 years or older with a bachelor's degree or higher by census tract are closest to the city, meaning the highest percentages by tract for attaining a higher education are found in more urban areas. (Figures 2 & 3). The standard deviational ellipses are slightly smaller for these variables than the one for the GINI index.

Table 2. Summary Table of Moran's I for the Washington D.C. CBSA

	GINI Index	% with Bachelor's Degree or Higher in Male Population 25 years and older	% with Bachelor's Degree or Higher in Female Population 25 years and older	Unemployment Rate	Percent Foreign Born	Percent Minority
Moran's I	0.397023	0.767622	0.731096	0.524081	0.729942	0.505433
Z-Score	25.0976	49.5371	46.2343	33.5136	46.8805	33.2075
pseudo p-value	0.001	0.001	0.001	0.001	0.001	0.001

Source: Table Created by Emma Blackwood. Data is from the U.S. Census, American Community Survey, 2012-2016 Estimates

Table 2 was created using a queen contiguity weight matrix. Positive spacial autocorrelation is seen in all six variables, meaning there is clustering of similar values. This violates the assumption of independence of data. Using a 999-permutation prediction resulted in a pseudo p-value of 0.001 across all variables, indicating statistical significance. With a 5 percent significance level ($p<0.05$), the null hypothesis of random dispersion can be rejected. The education variables (% with bachelor's degree or higher in population 25 years and older for male and female) and the percent foreign born by tract have strong positive spatial autocorrelation. GINI index, unemployment rate, and percent minority by census tract have moderate positive autocorrelation. There is spatial dependence occurring with the variables chosen.

Table 3. Summary Table of Bi-Variate Moran's I

	e01-GINI	e02-GINI	u01-GINI
Bi-Variate Moran's I	0.0400387	0.0504639	0.174129
Z-score	3.616	4.4941	14.8572
pseudo p-value	0.001	0.001	0.001

Source: Table Created by Emma Blackwood. Data is from the U.S. Census, American Community Survey, 2012-2016 Estimates

Table 3 summarizes bi-variate Moran's I for three variables in relation to the GINI index. The queen contiguity weight matrix was used to calculate the results. There is minimal correlation between the percent with bachelor's degree or higher in population 25 years and older for male and female (e01 and e02) and GINI index. There is mild spatial correlation between the unemployment rate (u01) and GINI index. These results are statistically significant, but they do not reveal much spatial correlation between and dependent and independent variables on a global scale.

Table 4. Summmary Table of Moran's I for Census Tracts vs. Grids

	Census Tracts		Grids	
	Black	White	Black	White
Moran's I	0.723784	0.557939	0.65827	0.656367
Z-score	47.4728	34.6362	25.6518	25.512
pseudo p-value	0.001	0.001	0.001	0.001

Source: Table Created by Emma Blackwood. Data is from the U.S. Census, American Community Survey, 2012-2016 Estimates

Grids were created at a scale of 7 kilometers (height and width). Looking at Moran's I for the white and black populations, using census tracts and grids both yielded positive spatial autocorrelation. All of the Moran's I values are statistically significant and the null hypothesis of random dispersion of the white and black populations can be rejectd. Using a grid system reduced postive spatial autocorrelation for the black population and increased positive spatial autocorrelation for the white population. Interestingly, using grids resulted in nearly identical values for spatial autocorrelation in the black and white populations. When equal area grids are used in favor of census tracts, there is equal clustering of the black and white populations.

Figure 5 and 6 compare spatial patterns in the white and black populations using grids and census tracts. Using census tracts, the black population (Figure 5) is highest in the tracts on the eastern side of Washington D.C. When grids are used, the distribution of the black population moves west more towards the center of the city. Looking at the white population (Figure 6), there is a large difference between the population distribution using grids as compared to census tracts. While the census tracts show the white population highest throughout the majority of the CBSA (except for the eastern portion within the city), the grids show the white population highest in the city.

Figure 4. Black Population

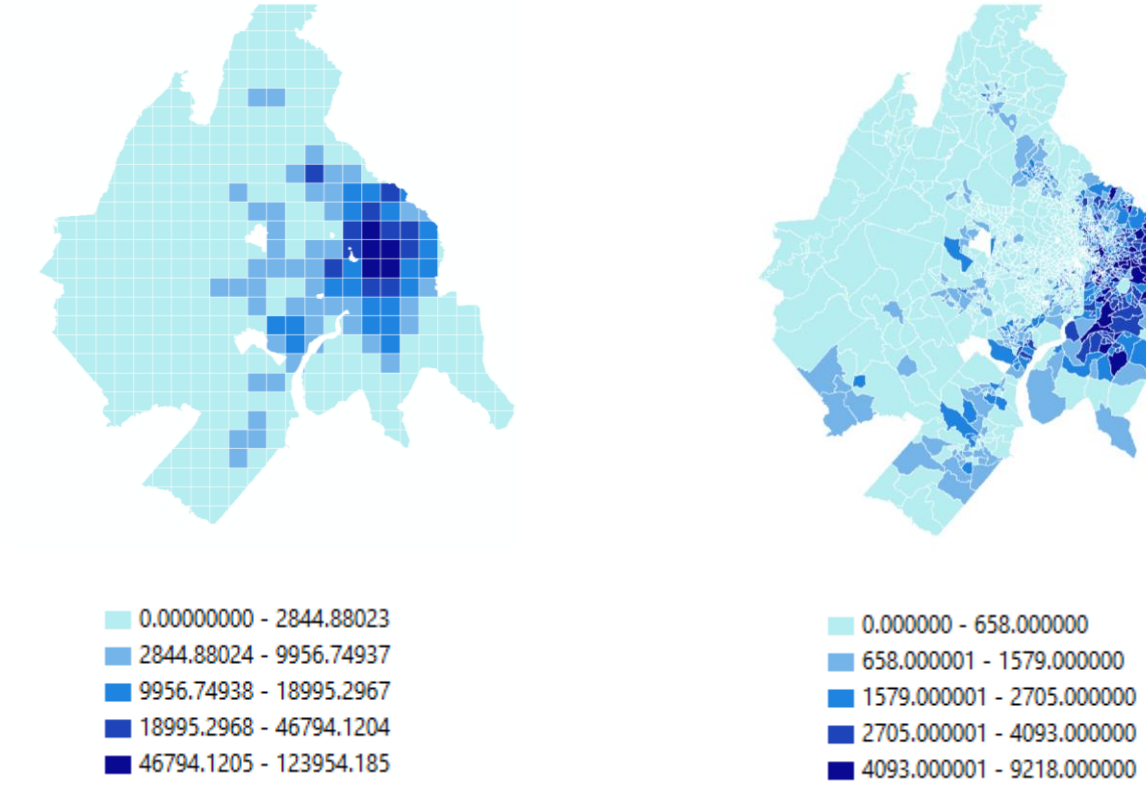


Figure 5. White Population

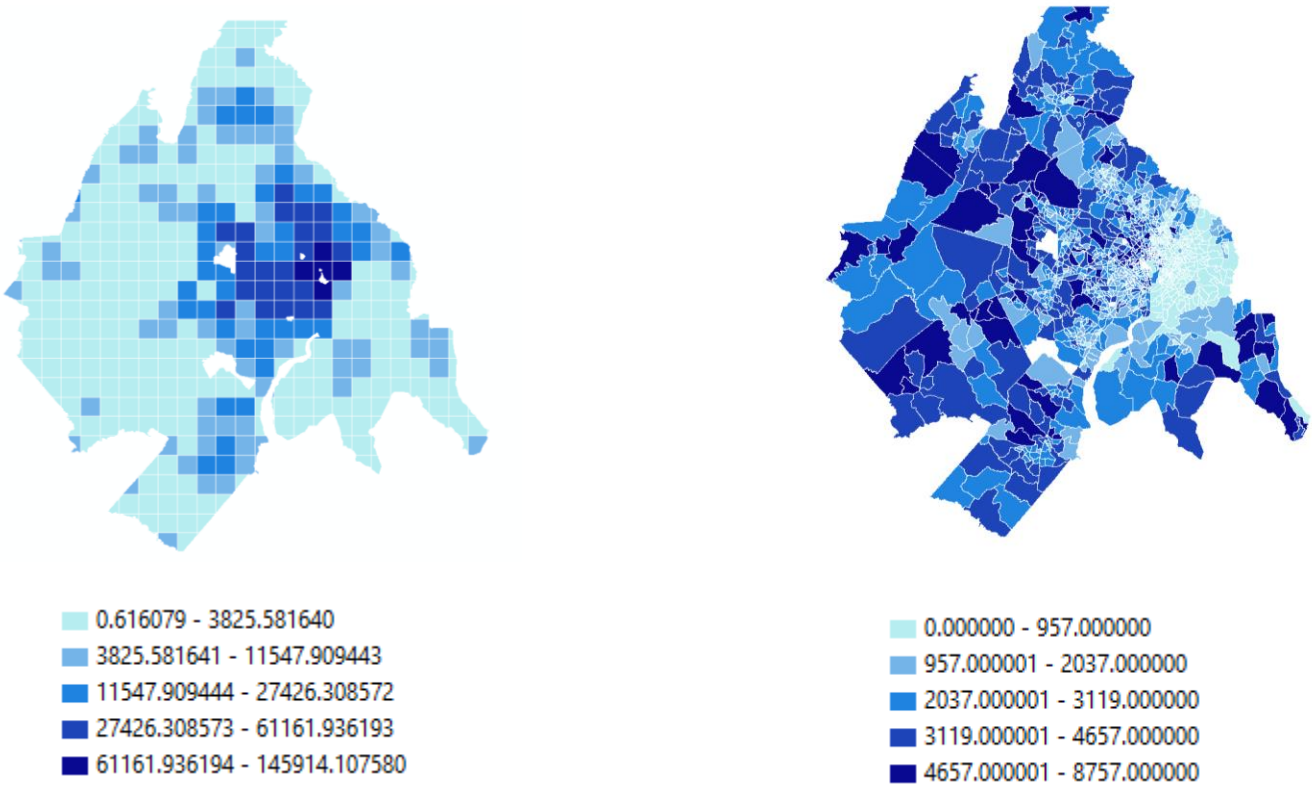


Figure 6. LISA map for GINI Index

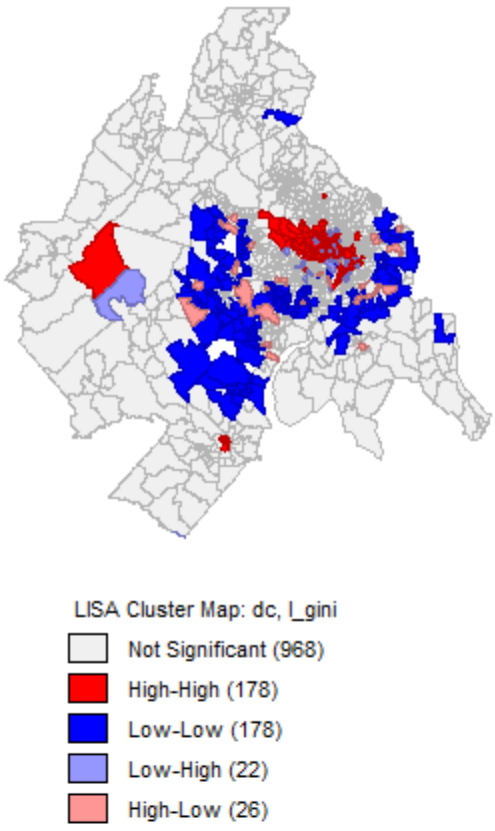


Figure 7. LISA map for e01

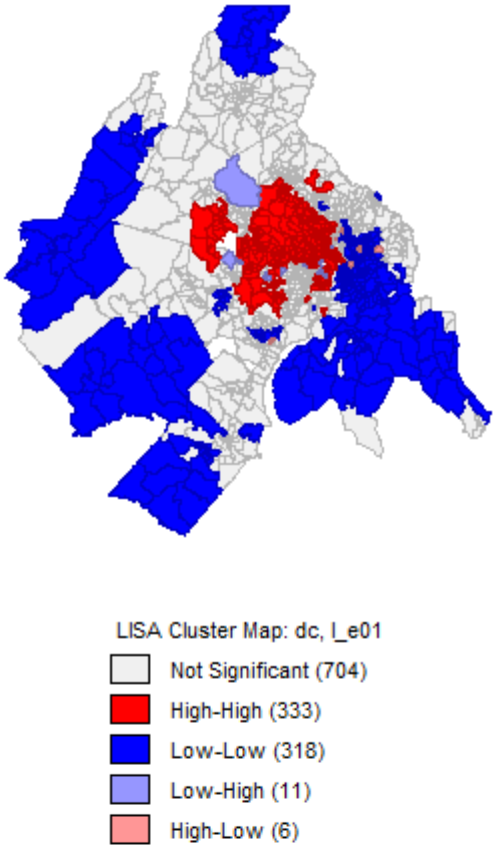
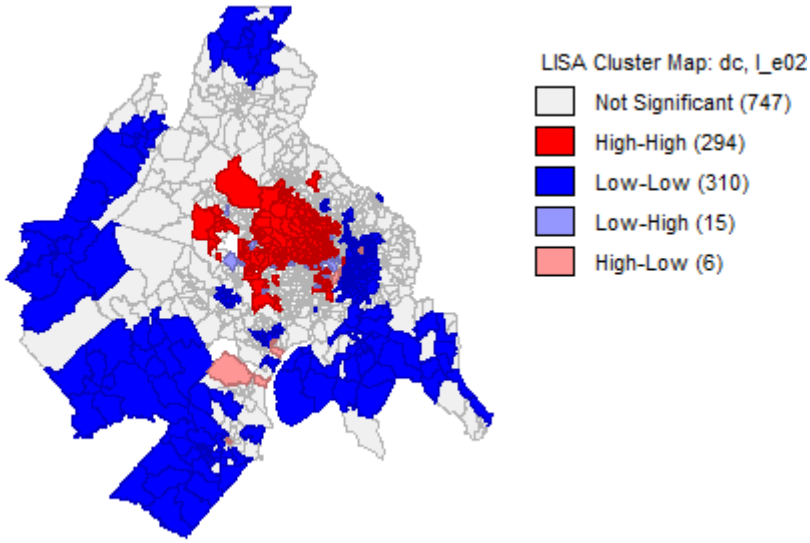


Figure 8. LISA map for e02



The LISA maps (Figures 6-8) were created using a queen contiguity weight matrix. The two education variables (e01 and e02) have large areas of significant local clustering. The census tracts closer to the city's core show High-High clustering, and tracts around the CBSA edges show Low-Low clustering. Overall, these cluster patterns contribute significantly to the positive global spatial autocorrelation values for the education variables. The lack of homogeneity suggests that different spatial models should be considered for the High-High areas compared to the Low-Low areas. The LISA map for the GINI Index also shows local clustering, but much fewer tracts have significant clustering compared to the LISA maps for e01 and e02.

Figure 9. BiLISA map for e01 with GINI Index

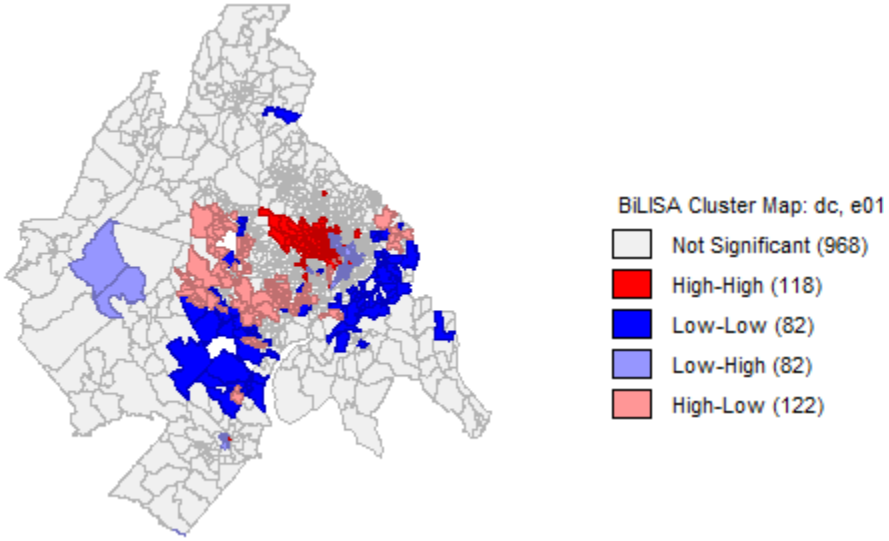
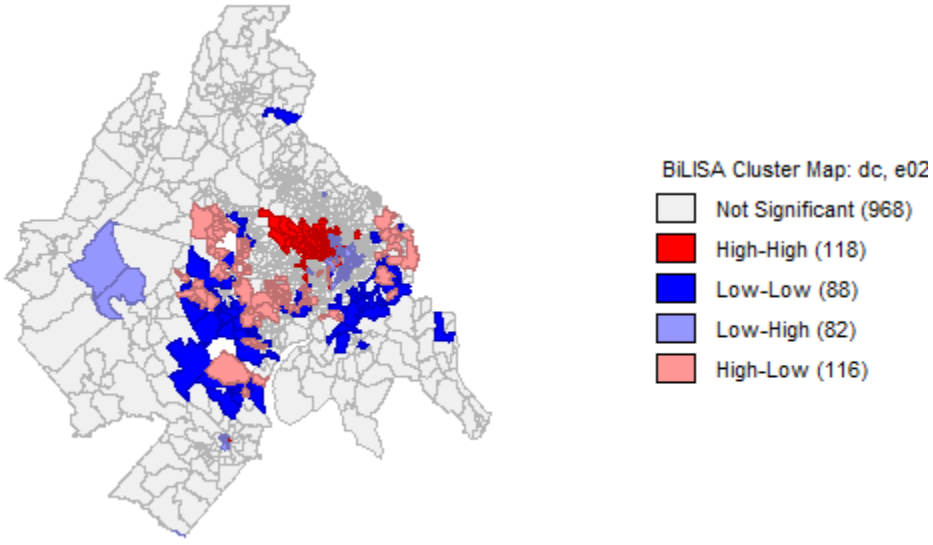


Figure 10. BiLISA map for e02 with GINI Index



The BiLISA maps were also created using the queen contiguity weight matrix. In some urban tracts, high values of the education variables are correlated with high GINI index values. In these tracts, obtaining a bachelor's degree or higher for males and females was correlated with higher levels of income inequality. There are also Low-Low clusters where lower values for bachelor's degree attainment are matched with low levels of income inequality. Many tracts had no significant clustering and others had outliers that would contribute to a lower value for bi-variate global Moran's I. There appear to be many processes influencing the dependent variable, the GINI index. While on a global level, bi-variate spatial autocorrelation suggested minimal autocorrelation between attainment of a bachelor's degree and GINI index, the BiLISA maps reveal heterogeneity and spatial correlation on the local scale.