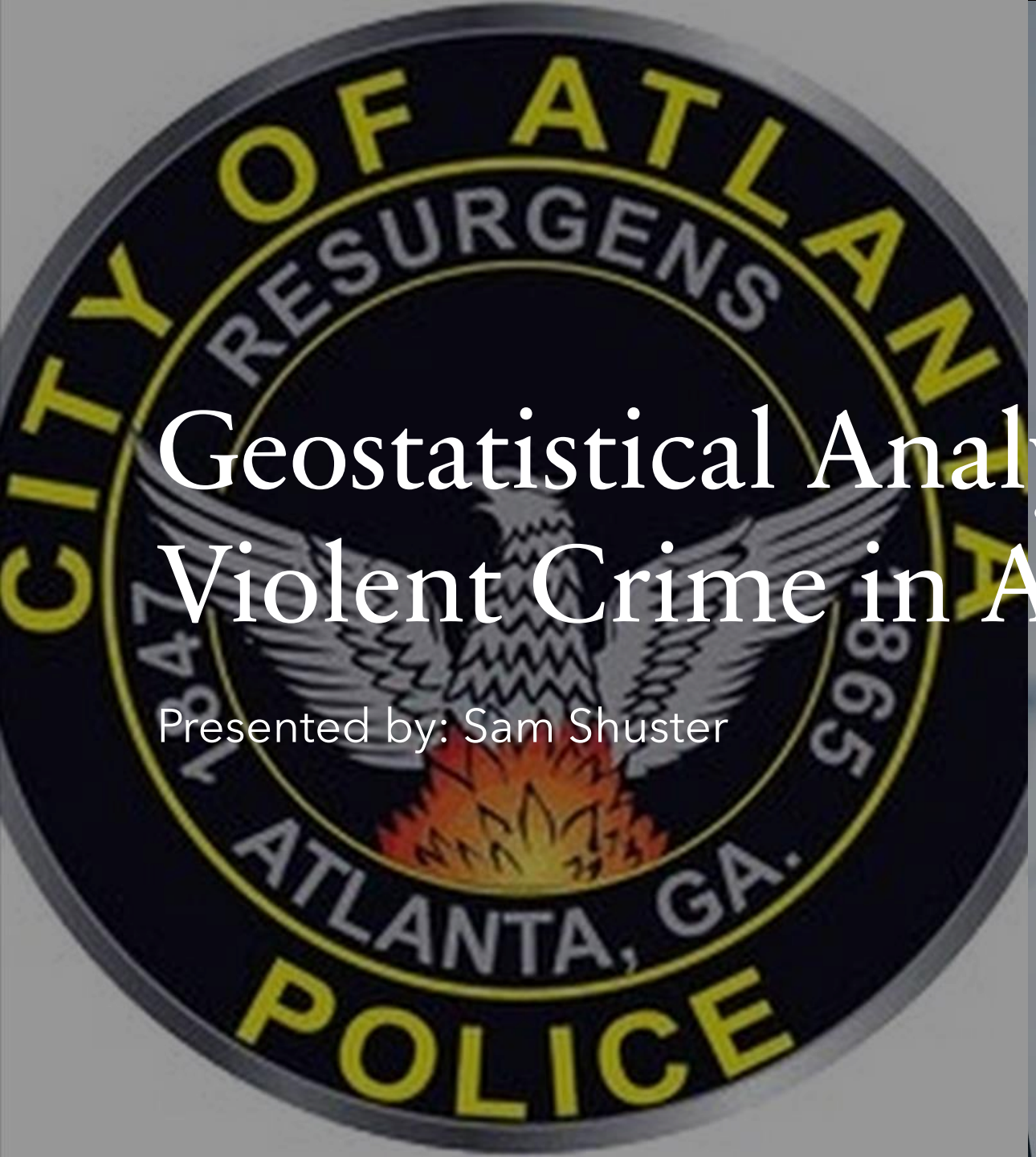


Geostatistical Analysis of Violent Crime in Atlanta

Presented by: Sam Shuster

Background:

- Home to fortune 500 company headquarters from companies like:
- Atlanta's Population is estimated to be approximately 488,800
  - (According to the 2019 American Community Survey: 5-Year Dataset)

Success in Law Enforcement:

In the last ~15 years, crime has decreased approximately 40%

Crime rates in Atlanta by year													
Type	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Murders (per 100,000)	110 (22.6)	129 (25.9)	105 (19.7)	80 (14.5)	93 (17.3)	88 (20.7)	83 (19.0)	84 (18.6)	93 (20.5)	94 (20.2)	111 (23.5)	79 (16.4)	88 (17.7)
Rapes (per 100,000)	171 (35.2)	148 (29.8)	126 (23.6)	135 (24.4)	89 (16.6)	148 (34.8)	113 (25.9)	105 (23.3)	151 (33.2)	170 (36.6)	136 (28.8)	282 (58.6)	245 (49.4)
Robberies (per 100,000)	2,959 (609.1)	3,577 (719.3)	3,308 (620.6)	2,725 (492.9)	2,162 (403.0)	2,343 (550.6)	2,276 (520.8)	2,363 (523.9)	2,329 (512.6)	1,995 (429.3)	2,070 (438.0)	1,413 (293.6)	1,099 (221.5)
Assaults (per 100,000)	4,308 (886.8)	4,221 (848.8)	3,864 (724.9)	3,419 (618.4)	3,405 (634.7)	3,518 (826.7)	3,555 (813.4)	2,965 (657.4)	3,004 (661.1)	2,944 (633.5)	2,804 (593.3)	2,730 (567.2)	2,382 (480.1)
Burglaries (per 100,000)	7,401 (1,523)	8,859 (1,781)	9,989 (1,874)	9,112 (1,648)	8,016 (1,494)	7,499 (1,762)	6,192 (1,417)	5,938 (1,317)	5,470 (1,204)	4,781 (1,029)	4,411 (933.4)	3,390 (704.3)	3,082 (621.2)
Thefts (per 100,000)	18,952 (3,901)	20,353 (4,093)	22,499 (4,221)	19,511 (3,529)	17,741 (3,307)	17,274 (4,059)	17,212 (3,938)	17,158 (3,804)	16,498 (3,631)	16,493 (3,549)	16,400 (3,470)	16,304 (3,387)	16,701 (3,366)
Auto thefts (per 100,000)	5,878 (1,210)	7,020 (1,412)	6,490 (1,218)	5,726 (1,036)	5,043 (940.0)	5,371 (1,262)	5,150 (1,178)	4,432 (982.7)	4,146 (912.5)	4,282 (921.4)	3,993 (844.9)	3,297 (685.0)	3,308 (666.8)
Arson (per 100,000)	175 (36.0)	155 (31.2)	147 (27.6)	157 (28.4)	95 (17.7)	121 (28.4)	103 (23.6)	71 (15.7)	75 (16.5)	50 (10.8)	50 (10.6)	74 (15.4)	90 (18.1)
City-Data.com crime index	793.0	862.6	786.9	659.4	607.9	786.4	728.6	667.5	656.0	612.9	591.8	520.8	469.0

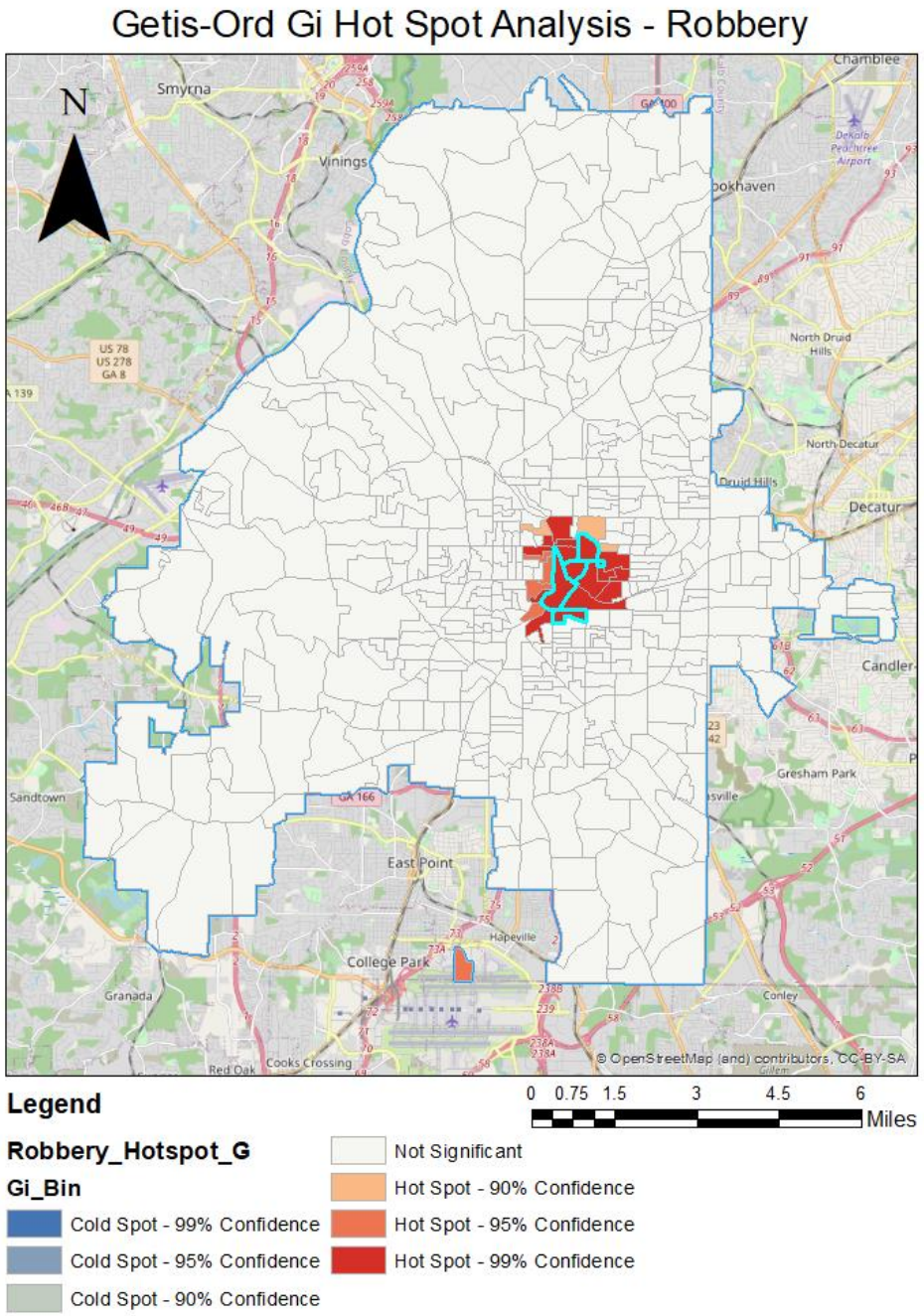
Objectives:

- Evaluate crime data to explore potential spatial relationships
- Determine whether there is any correlation between *location* and violent crimes like **robbery** and **rape**
- Determine whether there are any factors that increase the likelihood of criminality
- Roughly determine the safest and most dangerous regions to live within the city

Methodology:

1. Conducted Hot Spot Analysis using Getis-Ord Gi in ArcGIS
2. Conducted OLS for **Robbery** and **Rape** in ArcGIS
3. Conducted OLS test in GeoDa
4. Conducted Spatial Lag test in GeoDa
5. Conducted Geographically Weighted Regression in ArcGIS

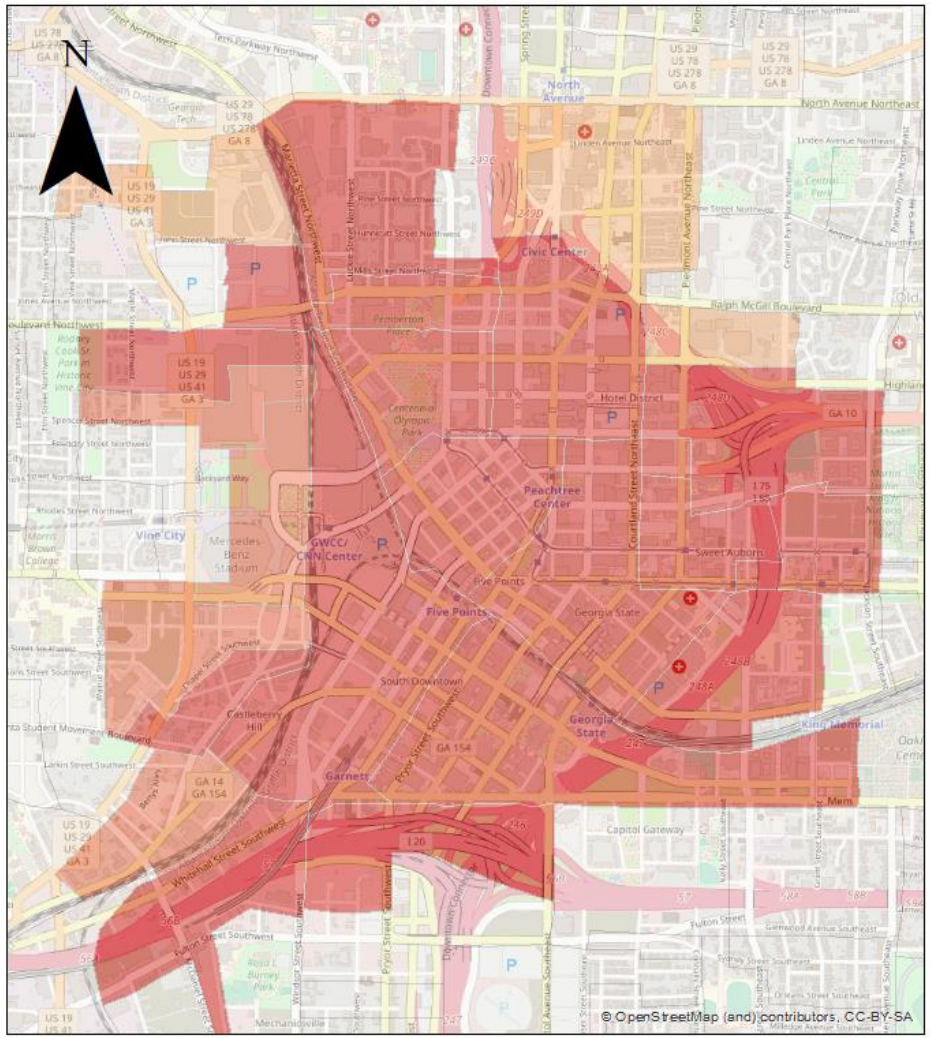
- Spatial resolution for all tests was conducted at the Census Block level



Location Analysis:

- Using Getis-Ord, a Hot Spot region was found bounded by GA Tech to the north, & I-20 to the South, US-19 to the West, and Jackson Street (SE) to the East.
- Queen continuity SWM used.

Getis-Ord Hot Spot Analysis - Robbery



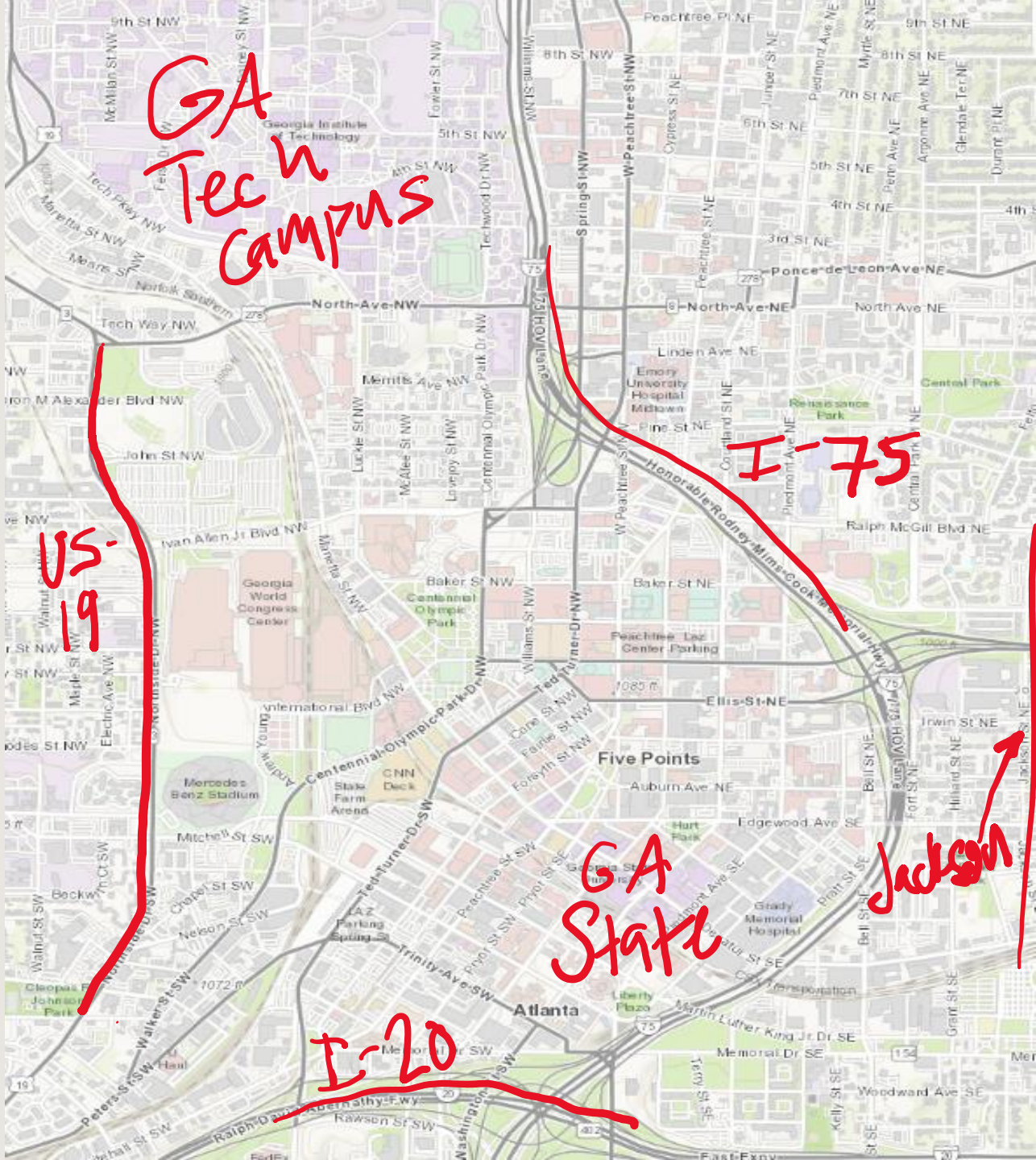
0 0.1 0.2 0.4 0.6 0.8 Miles

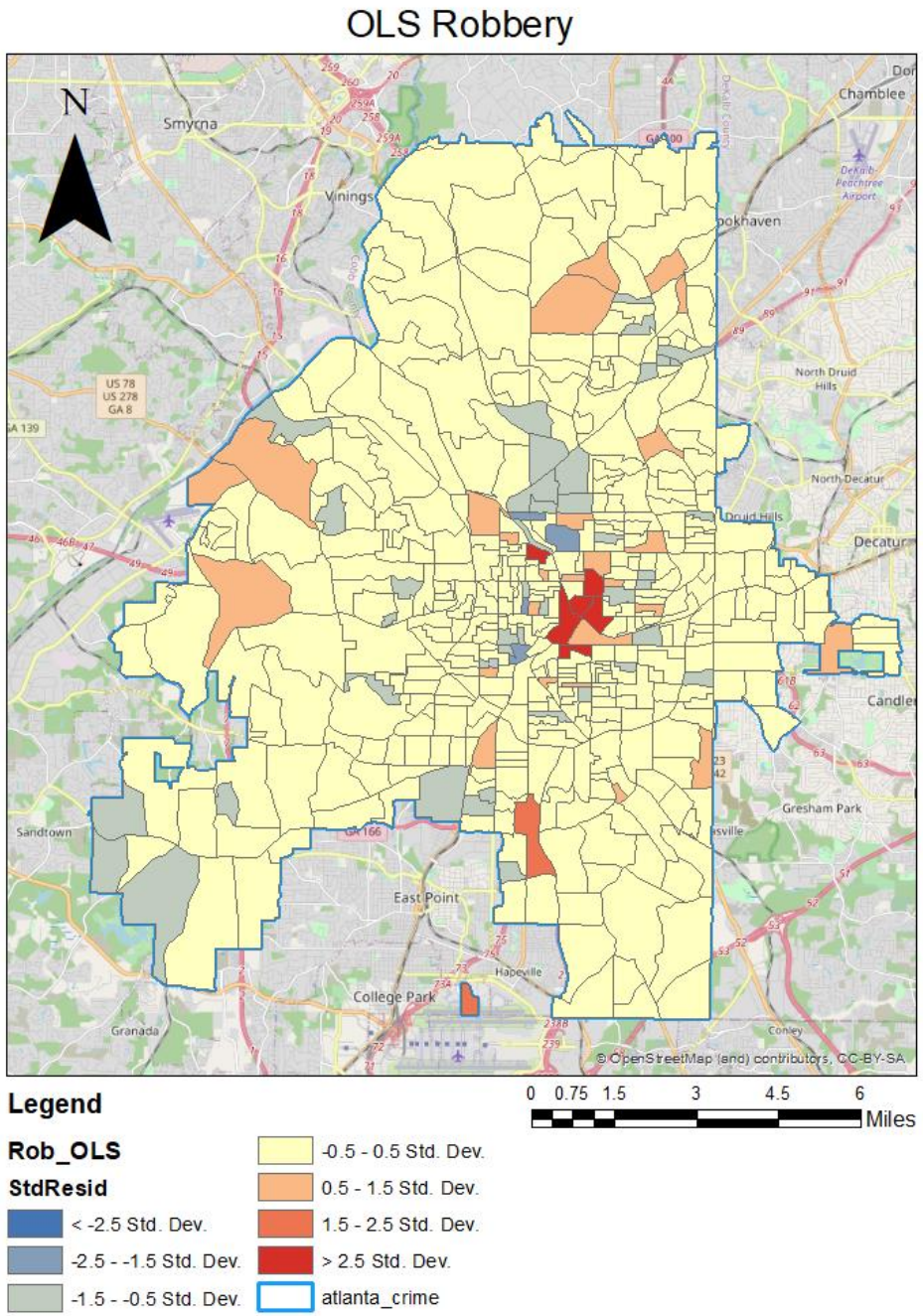
Legend

Robbery_Hotspot_G

Gi_Bin

- Cold Spot - 99% Confidence
- Cold Spot - 95% Confidence
- Cold Spot - 90% Confidence
- Not Significant
- Hot Spot - 90% Confidence
- Hot Spot - 95% Confidence
- Hot Spot - 99% Confidence





OLS Robbery

- Evaluated robbery based on:
  1. White (cnt)
  2. Black (cnt)
  3. Hispanic (cnt)
  4. Other Races (cnt)
  5. Housing Units (cnt)
  6. Vacant Housing Units (cnt)
  7. Median Family Income ($)
  8. Median Home Value ($)
  9. 1 Adult Households (%)
  10. 1 Parent Households (%)
  11. Fulltime Employment (%)
  12. HS Education (%)
  13. Poverty (%)
  14. Unemployment (%)
  15. Ages 15-24 (%)

Initial Findings (OLS Robbery):

1. White (cnt)
2. Black (cnt)
3. Hispanic (cnt)
4. Other Races (cnt)
5. Housing Units (cnt)
6. Vacant Housing Units (cnt)
7. **Median Family Income ($)**
8. **Median Home Value ($)**
9. 1 Adult Households (%)
10. **1 Parent Households (%)**
11. Fulltime Employment (%)
12. HS Education (%)
13. **Poverty (%)**
14. Unemployment (%)
15. **Ages 15-24 (%)**

Significant factors:

- Poverty is the most influential factor contributing to elevated robbery rates.
- Elevated Risk Factors:
  - Low Median Family Income (MFI)
  - Low Median Home Value (MHV)
  - High numbers of single Parent HH's
  - Areas w/ high numbers of teens – a large portion of robberies are committed by those in their late teens.

Initial Findings (OLS) cont.

Assessment of model fitness:

- **R^2=0.118, AIC=5102, *Significant heteroskedasticity***
- At glance, race appears to be a factor for both white and black demographics. Robust measures proved otherwise.
- With that in mind, associated VIFs for the White, Black, and Housing Unit variables are high. For that reason, Housing unit and race data is excluded in other models.
- Interpretation? **Insignificant finding.**

Findings (OLS Robbery) GeoDa:

1. **Median Family Income ($)**
2. **Median Home Value ($)**
3. 1 Adult Households (%)
4. **1 Parent Households (%)**
5. *Fulltime Employment (%)*
6. HS Education (%)
7. **Poverty (%)**
8. *Unemployment (%)*
9. **Ages 15-24 (%)**
10. **Population 1990 (cnt)**

Significant Factors:

- Variables with high VIF removed
- Similarly, to previous OLS:
  - Low MFI
  - Low MHV
  - Single Parent HH
  - High Poverty rate
  - High numbers of teens in area
- High Unemployment and low fulltime employment rates are factors of lesser influence.
- **R^2=0.114, AIC=5096, *Significant heteroskedasticity***
  - Slightly better fit than ArcGIS model (according to AIC)

Findings (Spatial Lag) GeoDa:

1. **Median Family Income ($)**
2. **Median Home Value ($)**
3. 1 Adult Households (%)
4. **1 Parent Households (%)**
5. **Fulltime Employment (%)**
6. **HS Education (%)**
7. **Poverty (%)**
8. Unemployment (%)
9. **Ages 15-24 (%)**
10. **Population 1990 (cnt)**

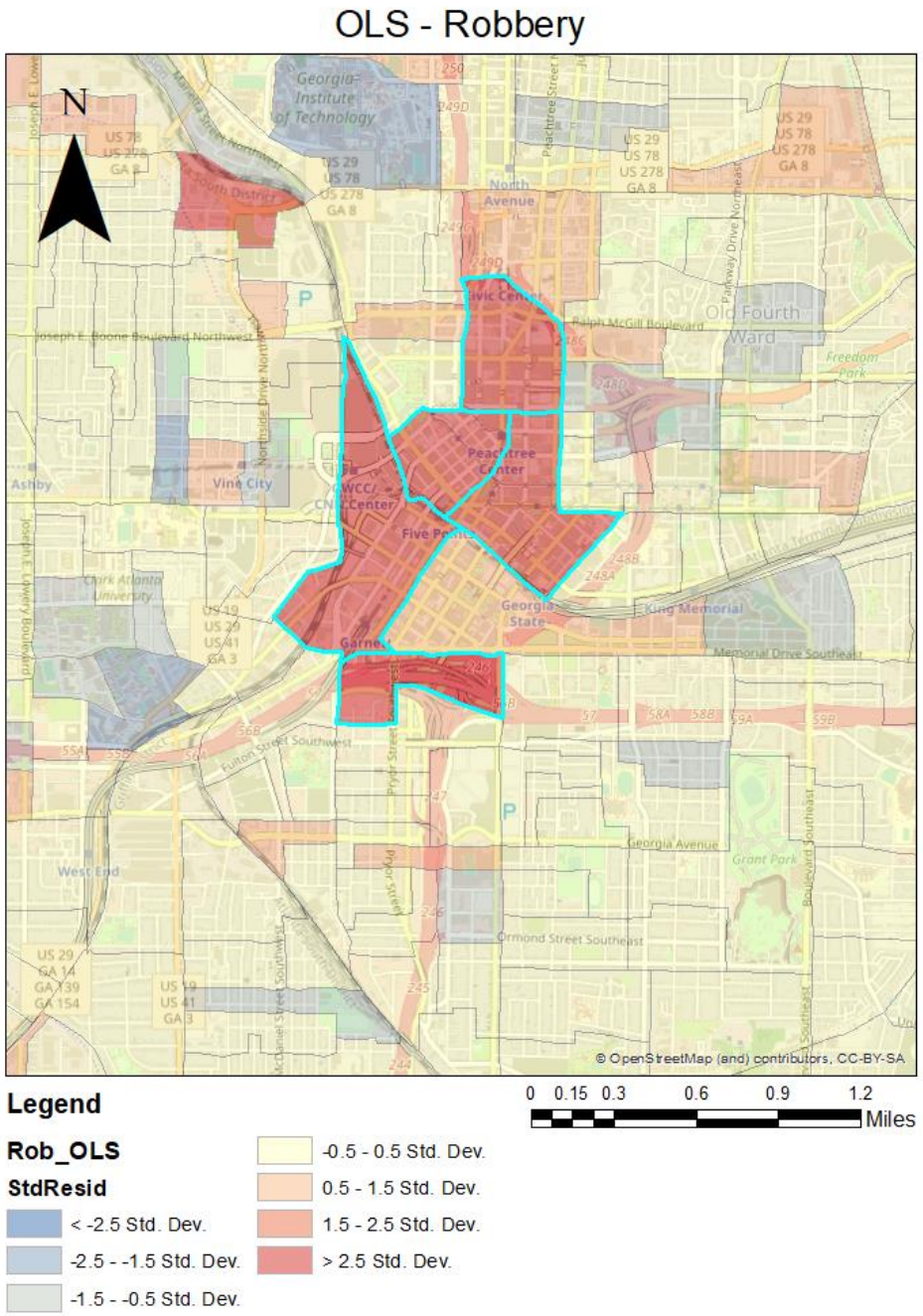
Significant Contributing Factors:

- Similarly, to previous OLS tests:
  - Low MFI
  - Low MHV
  - Single Parent HH
  - Low Fulltime Employment rate
  - Low rate of High School completion
  - High Poverty rate
  - High numbers of teens in area
  - **R^2=0.3707, AIC=4995, *Significant heteroskedasticity***
    - *Best fitting model so far.*

Findings (GWR Robbery):

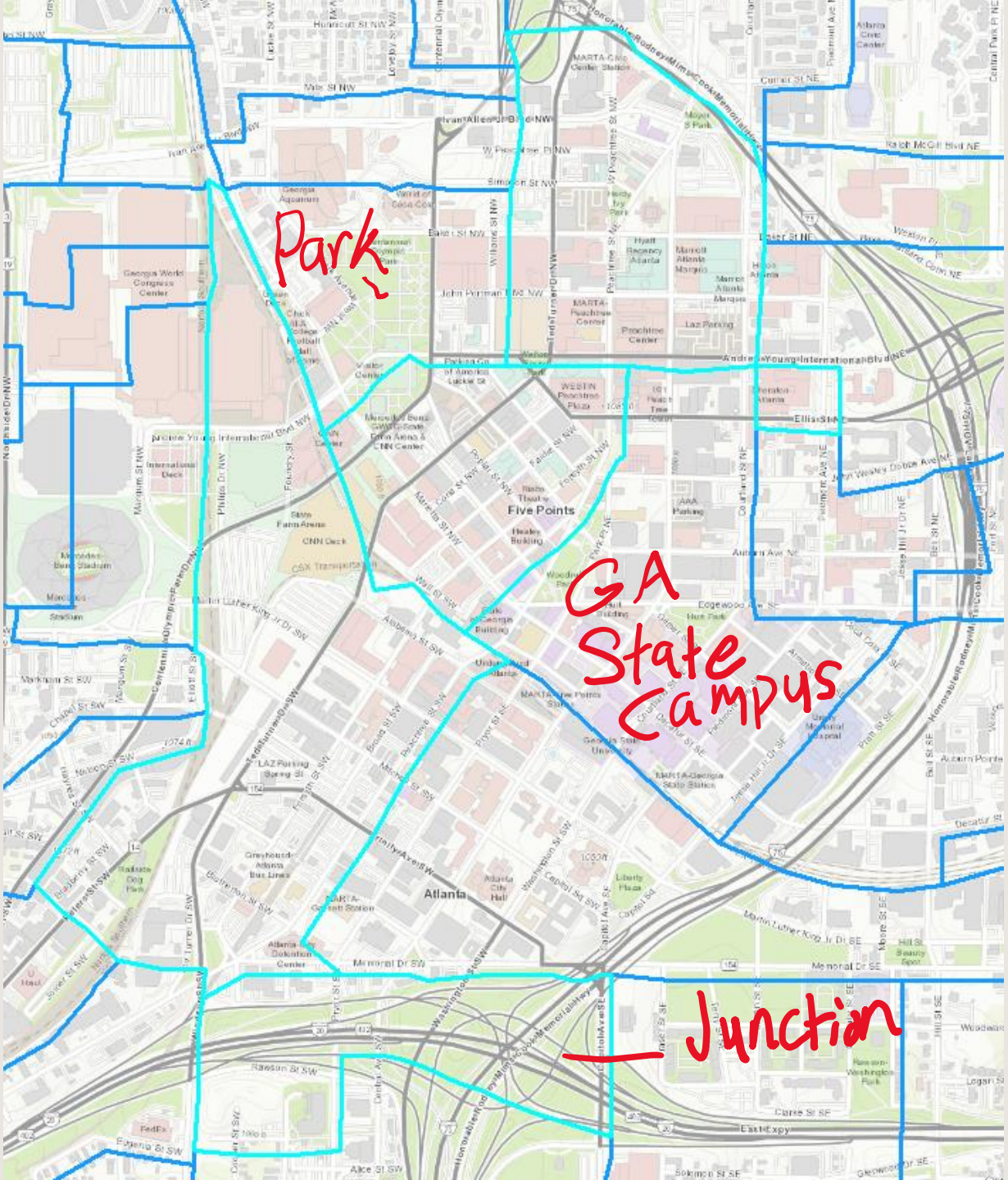
Assessment of model fitness:

- R^2=**0.243**
- **AIC=5044**
- ***Significant heteroskedasticity***



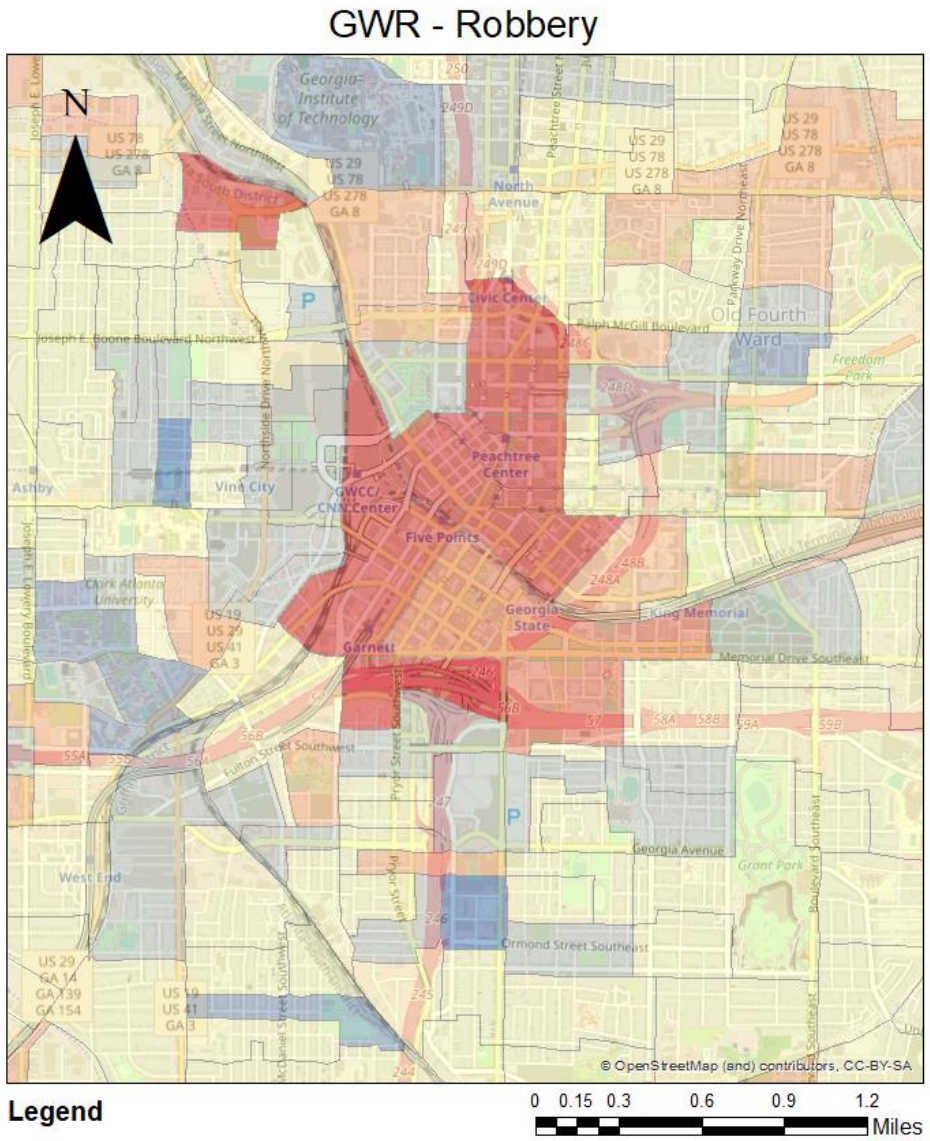
Location Analysis (Robbery):

- The highest concentration of robberies appears to occur in the heart of the city – roughly spanning the area from spaghetti Junction to the South, Georgia State to the East and Centennial Olympic Park to the Northwest.



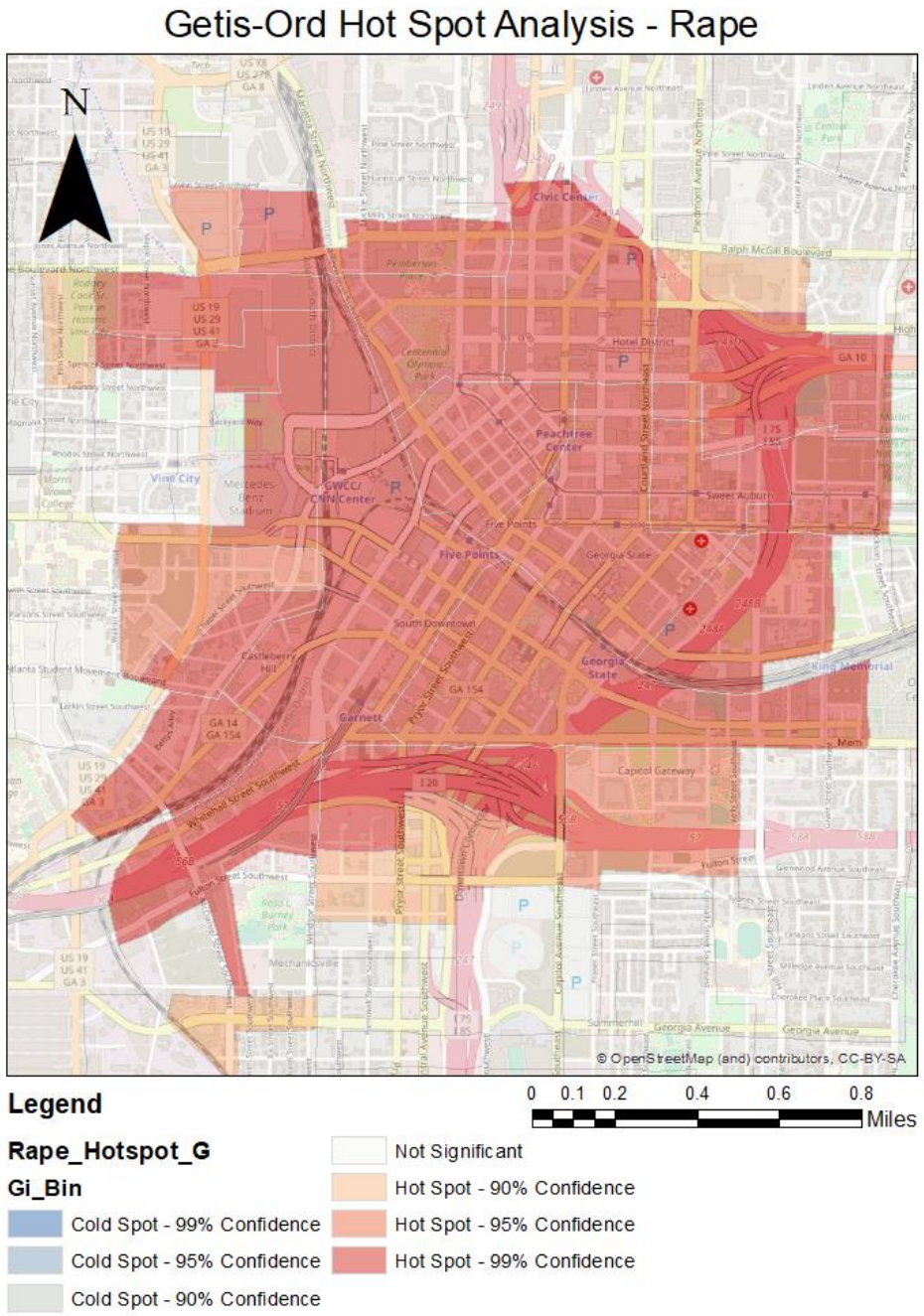
Location Analysis:

- The highest concentration of robberies appears to occur in the heart of the city – roughly spanning the area from spaghetti Junction to the South, Georgia State to the East and Centennial Olympic Park to the Northwest.



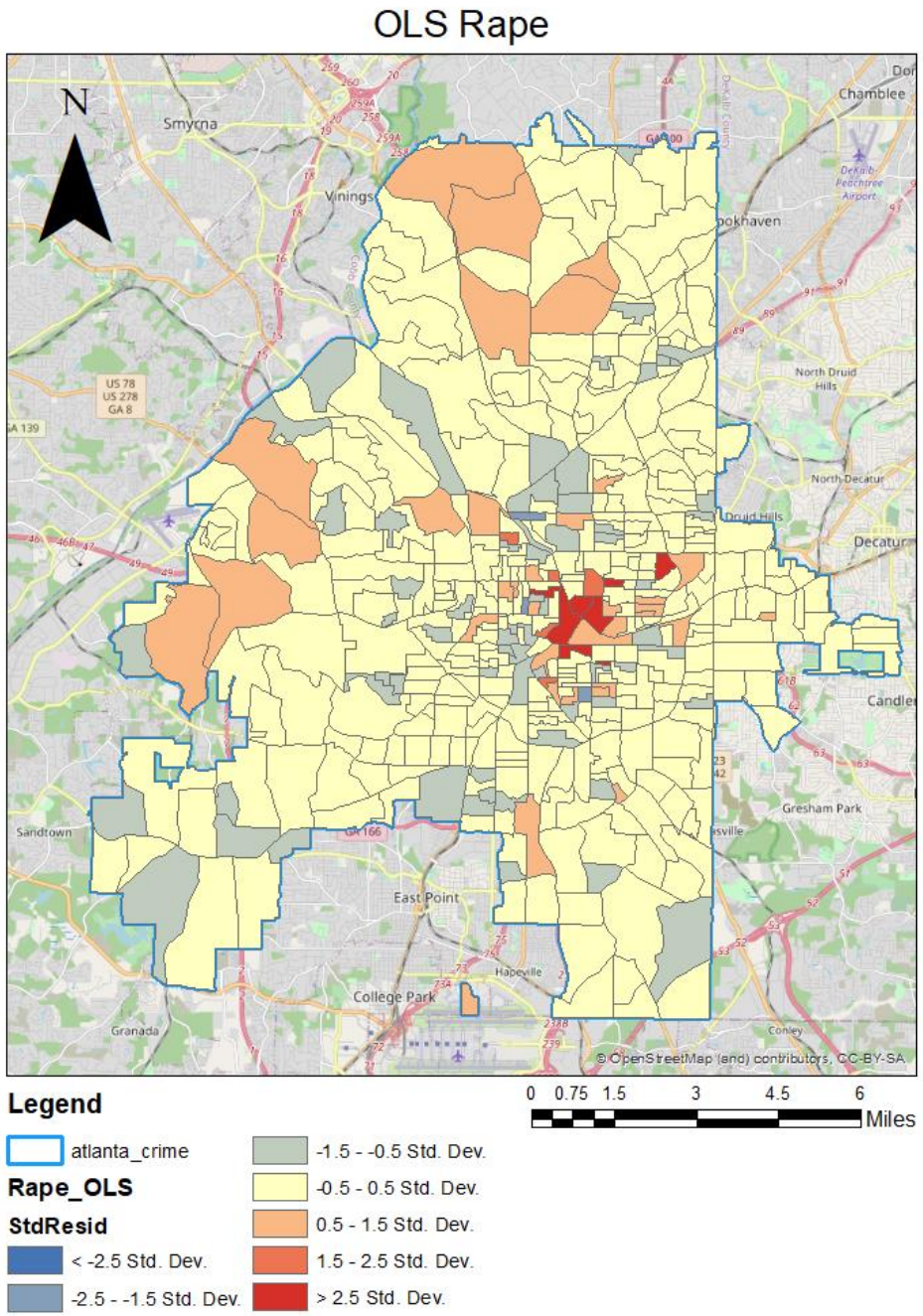
Location Analysis:

- The highest concentration of robberies appears to occur in the heart of the city – roughly spanning the area from spaghetti Junction to the South, Georgia State to the East and Centennial Olympic Park to the Northwest.



Location Analysis (Rape):

- Using Getis-Ord, a Hot Spot region was found bounded by GA Tech to the north, Spaghetti JCTN & I-20 to the South, US-19 to the West, and Jackson Street (SE) to the East.
- Queen continuity SWM used.



OLS Rape

- Evaluated rape based on:
  1. White (cnt)
  2. Black (cnt)
  3. Hispanic (cnt)
  4. Other Races (cnt)
  5. Housing Units (cnt)
  6. Vacant Housing Units (cnt)
  7. Median Family Income ($)
  8. Median Home Value ($)
  9. 1 Adult Households (%)
  10. 1 Parent Households (%)
  11. Fulltime Employment (%)
  12. HS Education (%)
  13. Poverty (%)
  14. Unemployment (%)
  15. Ages 15-24 (%)

Initial Findings (OLS Rape):

1. *White (cnt)*
2. **Black (cnt)**
3. Hispanic (cnt)
4. Other Races (cnt)
5. Housing Units (cnt)
6. Vacant Housing Units (cnt)
7. Median Family Income ($)
8. *Median Home Value ($)*
9. 1 Adult Households (%)
10. *1 Parent Households (%)*
11. *Fulltime Employment (%)*
12. **HS Education (%)**
13. **Poverty (%)**
14. Unemployment (%)
15. *Ages 15-24 (%)*

Significant factors:

- The most influential factor in analysis of rape rates appears to be related to education.
- Areas with low High School completion rates have higher rates of rape.
- Other elevated risk factors include:
  - Low Median Home Value (MHV)
  - High numbers of single Parent HH's
  - High Poverty rates
  - Areas w/ high numbers of teens – statutory rape is included in rape rate.
  - The Black community appears to have a significantly higher rate of rape than other races

Initial Findings (OLS) cont.

Assessment of model fitness:

- **R^2=0.129, AIC=2997, *Significant heteroskedasticity***
- At glance, race appears to be a factor for both white and black demographics. Robust measures however indicate that the Black community has a statistically significant propensity towards elevated rape rates.
- That said, the associated VIFs for the White, Black, and Housing Unit variables are high. For that reason, Housing unit and race data is excluded in other models.
- Interpretation? **The model may be biased.**

Findings (OLS Rape) GeoDa:

Significant Factors:

- **R^2=0.129, AIC=2990, *Significant heteroskedasticity***
- Variables with high VIF removed
- Similarly, to previous OLS:
  - Low Median Home Value (MHV)
  - High numbers of single Parent HH's
  - High Poverty rates
  - Areas w/ high numbers of teens – statutory rape is included in rape rate.
  - The Black community appears to have a significantly higher rate of rape than other races

Findings (Spatial Lag) GeoDa:

1. Median Family Income ($)
2. **Median Home Value ($)**
3. 1 Adult Households (%)
4. **1 Parent Households (%)**
5. **Fulltime Employment (%)**
6. **HS Education (%)**
7. **Poverty (%)**
8. **Unemployment (%)**
9. **Ages 15-24 (%)**
10. **Population 1990 (cnt)**

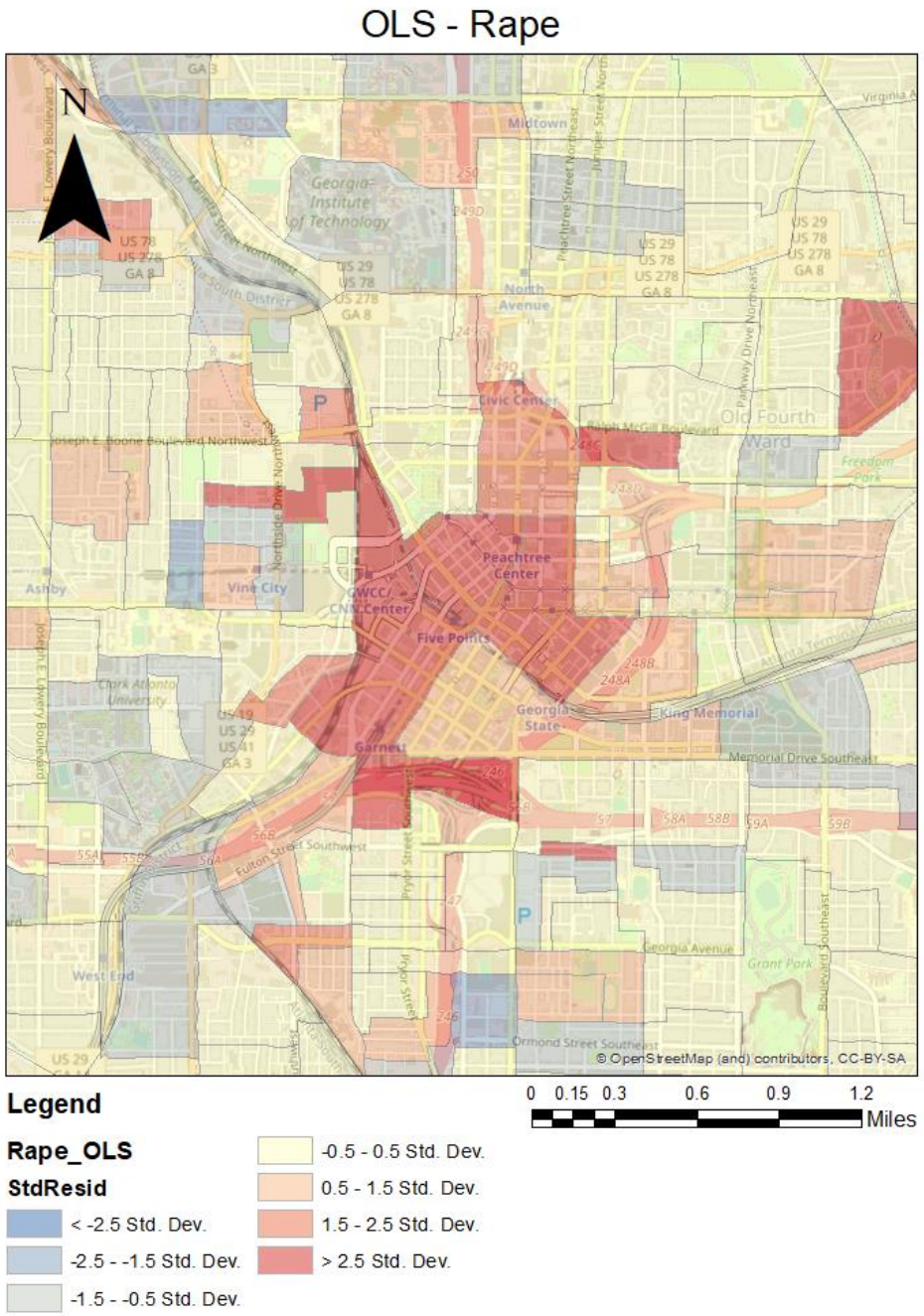
Significant Contributing Factors:

- Similarly, to previous OLS tests:
  - Low rate of High School completion
  - Low MHV
  - High Poverty rate
  - Low Fulltime Employment rate
  - Single Parent HH
  - High numbers of teens in area
  - **R^2=0.318, AIC=2920, *Significant heteroskedasticity***
    - *Best fitting model.*

Findings (GWR Rape):

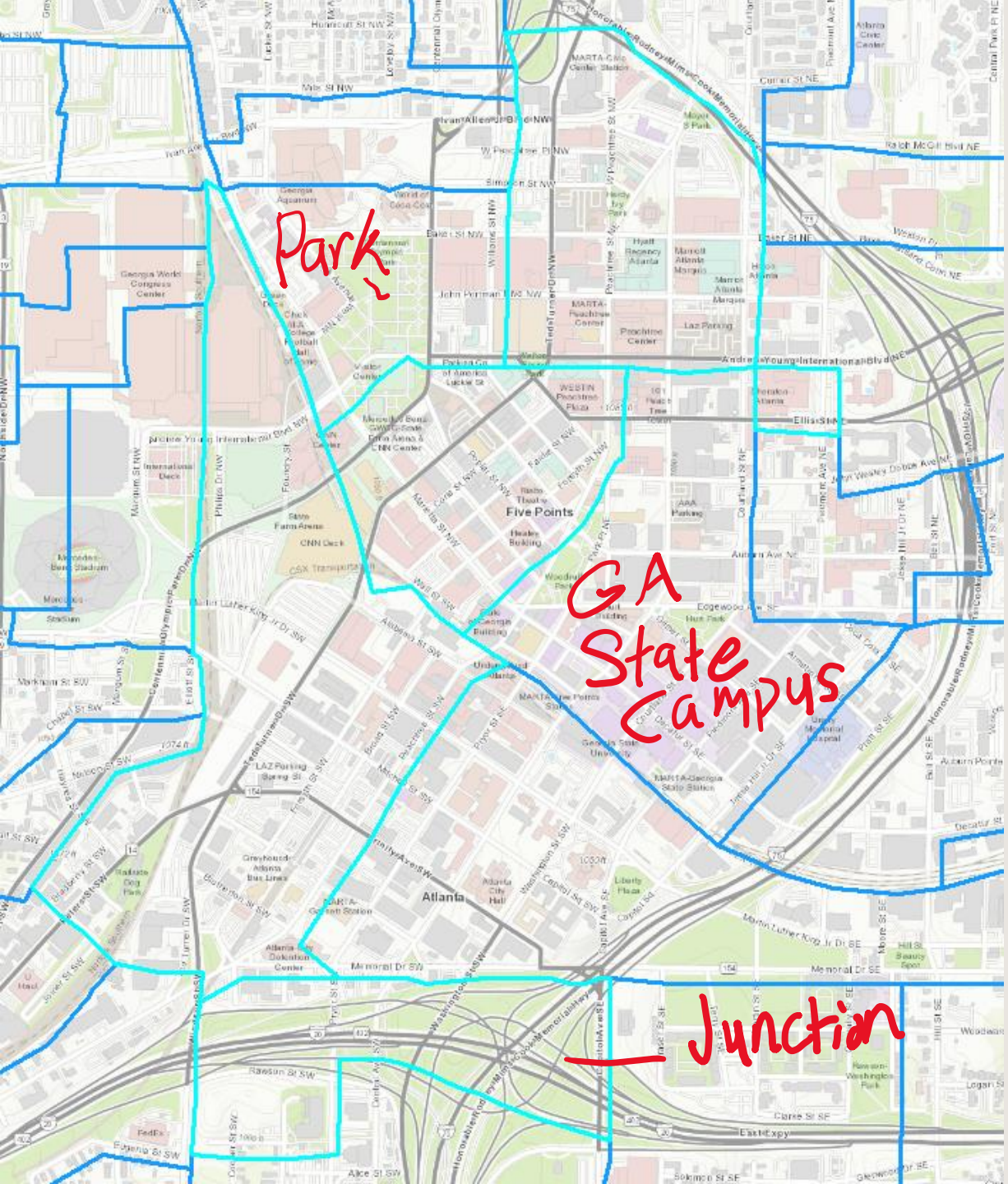
Assessment of model fitness:

- R^2=0.153
- AIC=2989
- ***Significant heteroskedasticity***



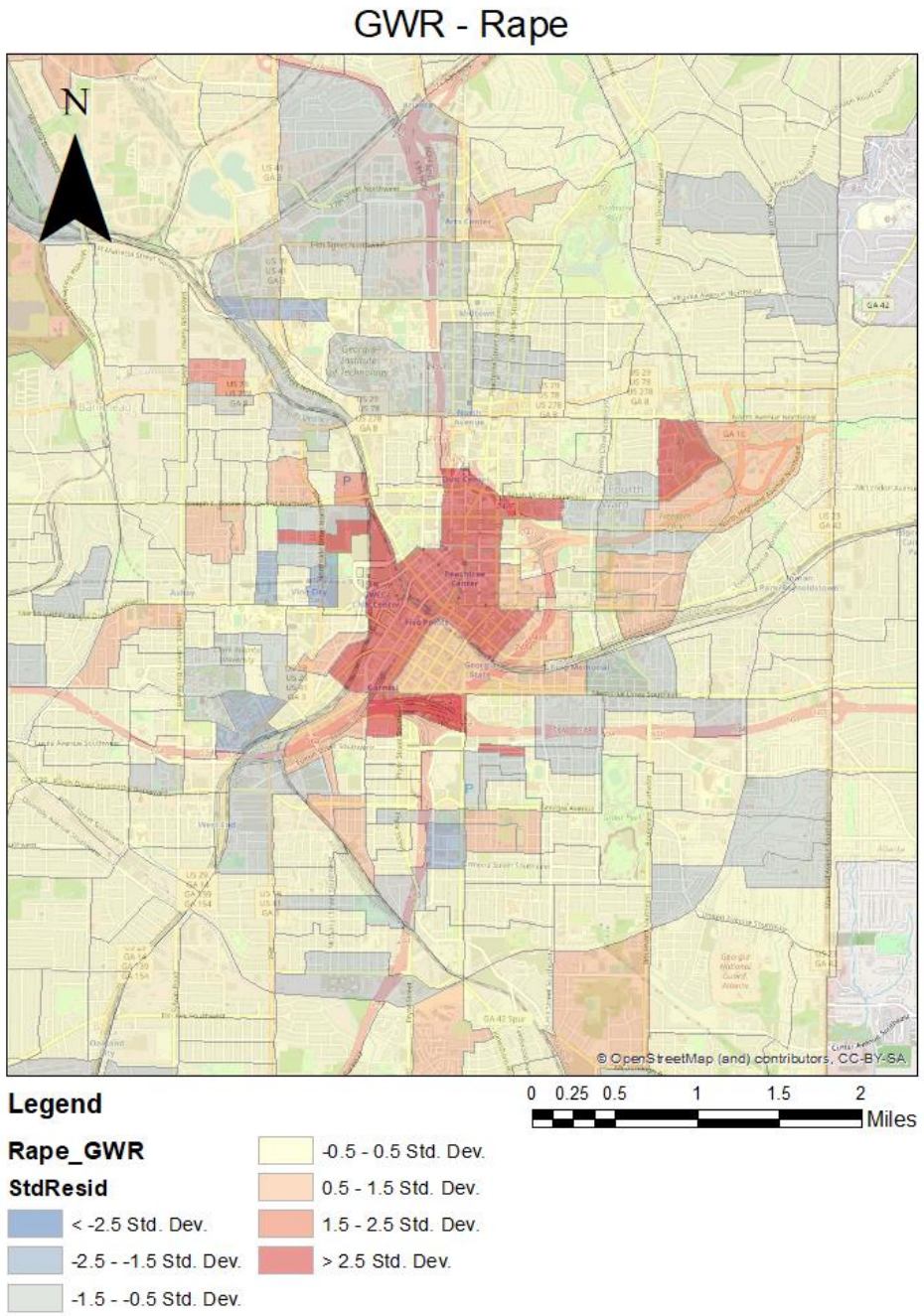
Location Analysis:

- The highest concentration of rapes appears to occur in the heart of the city – roughly spanning the area from spaghetti Junction to the South, Georgia State to the East and Centennial Olympic Park to the Northwest.



Location Analysis:

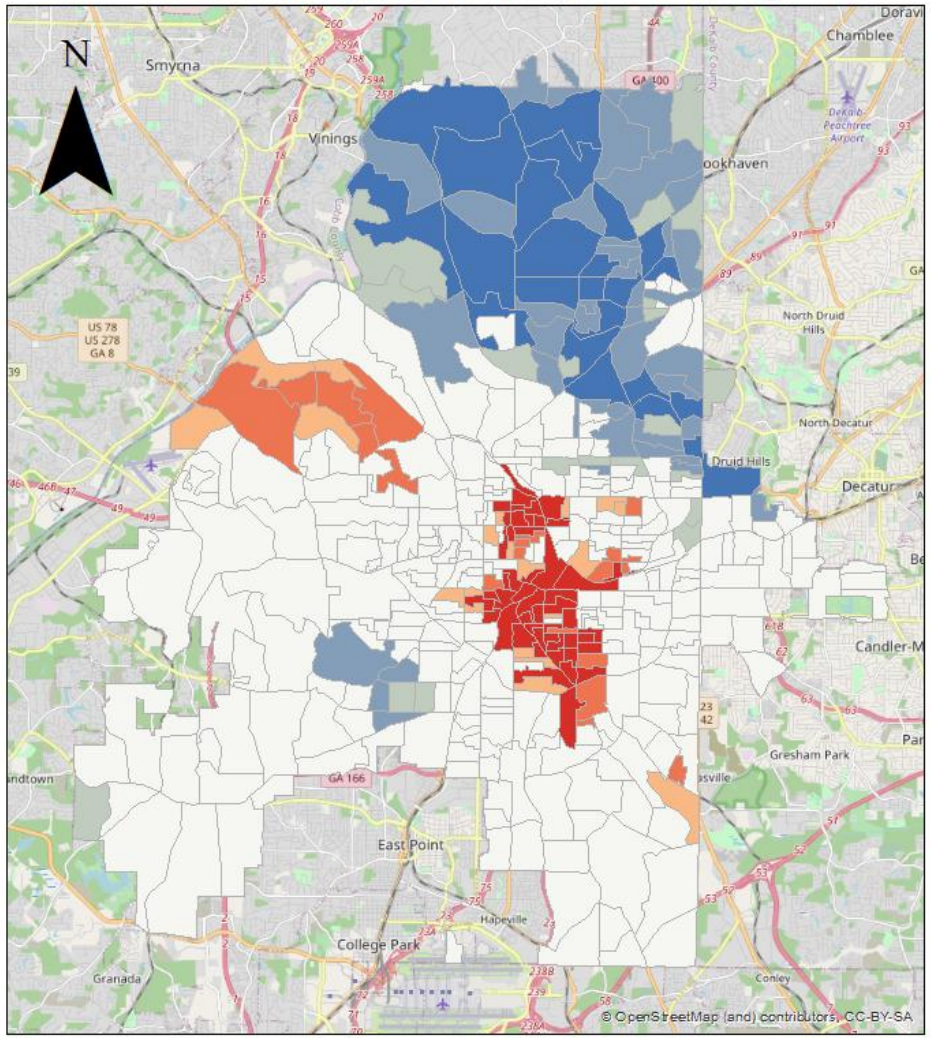
- The highest concentration of rapes appears to occur in the heart of the city – roughly spanning the area from spaghetti Junction to the South, Georgia State to the East and Centennial Olympic Park to the Northwest.



Location Analysis:

- The highest concentration of rapes appears to occur in the heart of the city – roughly spanning the area from spaghetti Junction to the South, Georgia State to the East and Centennial Olympic Park to the Northwest.

Getis-Ord Hot Spot Analysis - Poverty



0 0.75 1.5 3 4.5 6 Miles

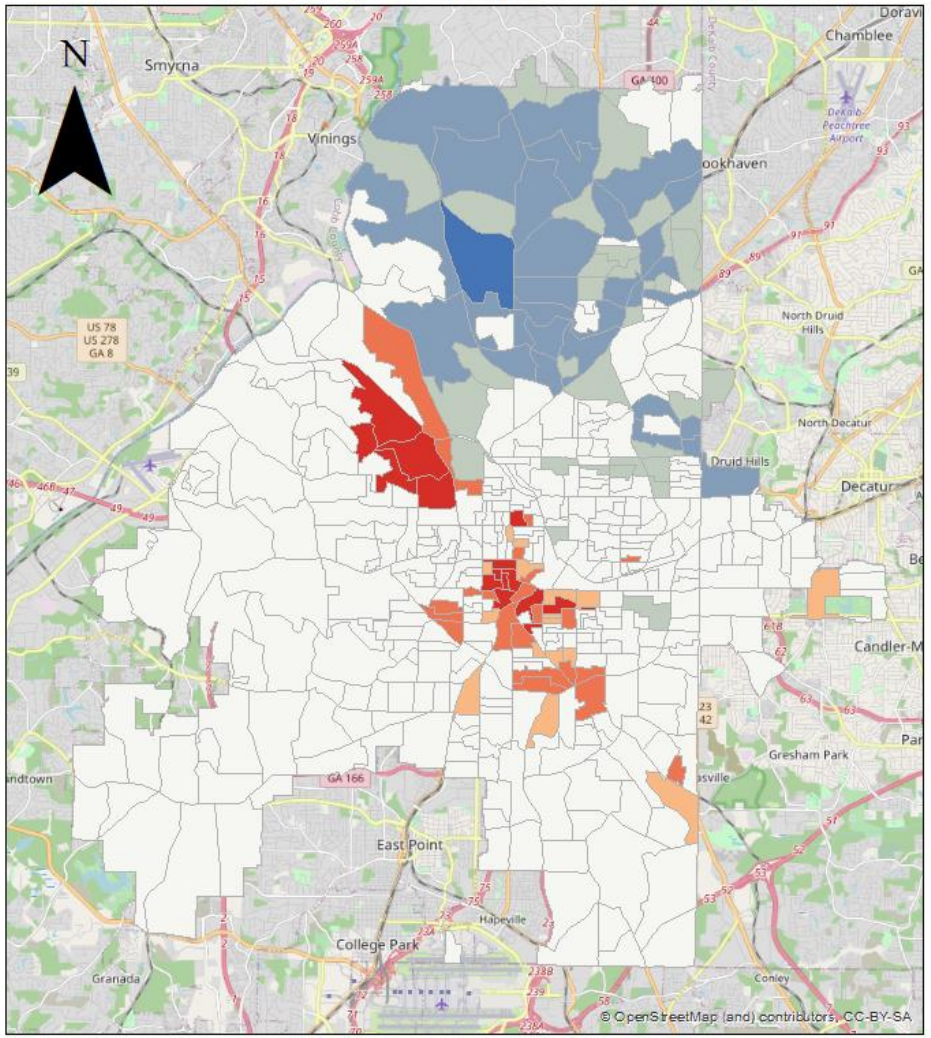
Legend

Poverty_Hotspot_G

Gi_Bin

- Cold Spot - 99% Confidence
- Cold Spot - 95% Confidence
- Cold Spot - 90% Confidence
- Not Significant
- Hot Spot - 90% Confidence
- Hot Spot - 95% Confidence
- Hot Spot - 99% Confidence

Getis-Ord Hot Spot Analysis - Unempolyment



0 0.75 1.5 3 4.5 6 Miles

Legend

unemp_Hotspot_G

Gi_Bin

- Cold Spot - 99% Confidence
- Cold Spot - 95% Confidence
- Cold Spot - 90% Confidence
- Not Significant
- Hot Spot - 90% Confidence
- Hot Spot - 95% Confidence
- Hot Spot - 99% Confidence

Conclusions:

- Poverty is a primary driver for robbery.
- Low-income communities with high numbers of single parent HHs are at high risk for violent crime.
- Lack of education is the best indicator for predicting rape.
- Northern part of the city is the safest; the central portion is the most dangerous.
- The Spatial lag models were the best fits for both Robbery & Rape.
- Model heteroskedasticity indicates further analysis must be conducted.

Resources:

- NHGIS
- Atlanta Police Department
- City-Data.com
- GeoDa
- ArcGIS

Questions?