

# Not All Crimes Are Created Equal

Analysis of How Criminal Indicators May Vary by Crime Type in Philadelphia County, Pennsylvania

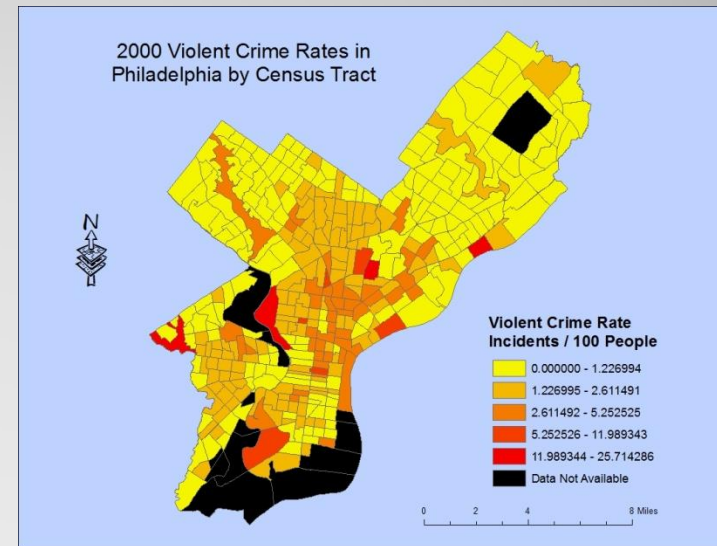
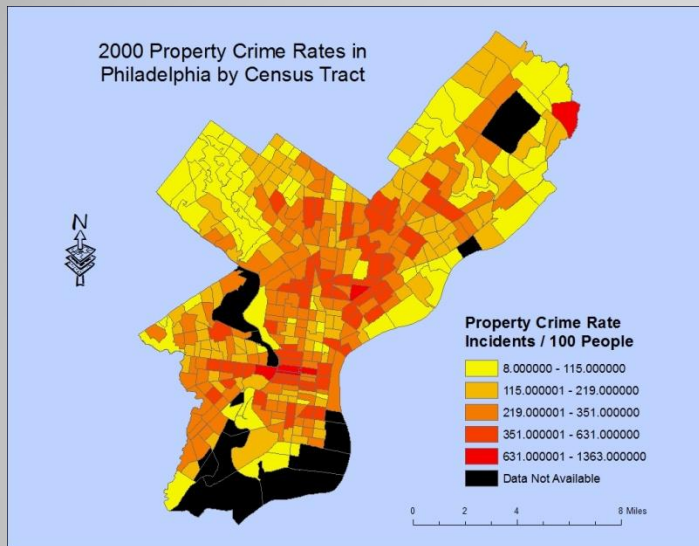
Jay Gregory  
GEOG 606  
Quantitative Spatial Analysis



- Crime in the United States
  - Crime rates contain two indices: violent crime and property crime
  - Larger cities tend to have higher rates of crime than smaller cities
  - More affluent cities and neighborhoods tend to have lower rates of crime than poorer areas
  - Little research available (online) investigating different predictors between violent crime and property crime, since both indices are often combined into one rate for reporting purposes
- Philadelphia
  - The 2000 U.S. Census ranked Philadelphia as the 6<sup>th</sup> largest city in population
  - The 2000 U.S. Census ranked Philadelphia as the 7<sup>th</sup> poorest cities among cities with populations over 250,000
  - In 2006, Philadelphia's homicide rate of people was the highest of the country's 10 most populated cities<sup>1</sup>
  - In 2005, Morgan Quitno ranked Philadelphia as the sixth most dangerous city out of 32 cities with populations over 500,000<sup>2</sup>

## Introduction

1. Determine what 2000 census variables best predict incidents of violent crime in Philadelphia County, PA
2. Investigate how some of these predictive variables differ spatially across Philadelphia county
3. Understand if similar or different variables also accurately predict incidents of property crime



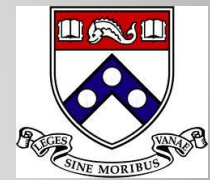
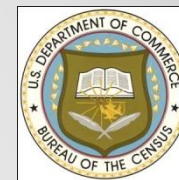
# Research Objectives

- Previous studies have found strong relationships between criminal activity and the following indicators
  - Education level
  - Socio-economic status
  - Race
  - Age
  - Gender
  - Population density
  - Land use / land cover and zoning
- A 1982 study found that socioeconomic differences between races and economic inequality in general causes increased rates of violent crime<sup>3</sup>
- A 2005 study found that population density is a significant negative predictor of crime in urban areas, but not necessarily suburban areas<sup>4</sup>

## Previous Research

- Crime and census data for the year 2000 obtained from the University of Pennsylvania's Cartographic Modeling Lab (CML)
- CML provides counts of different crimes aggregated by type of areal units such as census tract, ZIP code, police district
- Violent crime includes counts of robbery and aggravated assault, and property crime includes counts of burglary, personal theft, and motor vehicle theft
- Crime count data only represents categorized incidents at the time of the response, usually incorporating updates that have occurred within five days. However, categorizations that have changed outside the five-day window are not represented in the data
- Crime data was aggregated by census tract to facilitate tabular joins between crime counts and census data. The author recognizes possible effects from the modifiable areal unit problem
- Removed 11 census tracts from analysis because census data was unavailable or suspicious (e.g. population of 1)
- 2000 census data and shape files downloaded from the ESRI-hosted geography network
- Coordinate system: NAD 1983 State Plane projection for Pennsylvania's southern section

## Data Description



- Attempt to find best indicators of violent crime in Philadelphia through two different methods
  - ArcMap 9.3.1 – Trial and error OLS regression analysis with different explanatory variables
  - Excel – Stepwise regression algorithm to find significant variables for violent crime
- Compare to property crime
  - Used OLS to investigate if explanatory variables were similarly predictive for property crimes
  - Stepwise regression to find significant variables for property crime
- Spatial Autocorrelation for Violent and Property Crime
  - Perform Geographically Weighed Regression (GWR) analysis to determine how variables differ spatially across Philadelphia
  - Determine whether OLS results show signs of spatial autocorrelation using Moran's I and Local Moran's I functions in ArcMap
- Comparison Between Property and Violent Crime
  - Investigate difference in explanatory variable's spatial effects on violent crime and property crime

## Methodology

Trial and Error Regression Results for Violent Crime		Stepwise Regression Results for Violent Crime	
Variables	Confidence	Variables	Confidence
# White	> 99.9%	# White	> 99.9%
# Females	> 99.9%	# Asian	> 99%
# Age 30-39	> 99%	# Age 30-39	> 99.9%
# Income < 200% Poverty	> 99.9%	# Income < 100% Poverty Level	> 99.9%
Regression Diagnostics			
Adjusted R <sup>2</sup>	0.735	OLS Adjusted R <sup>2</sup>	0.764
ANOVA (F-Test)	257.497	ANOVA (F-Test)	300.106
F Confidence	> 99.9%	F Confidence	> 99.9%
		GWR Adjusted R <sup>2</sup>	0.832

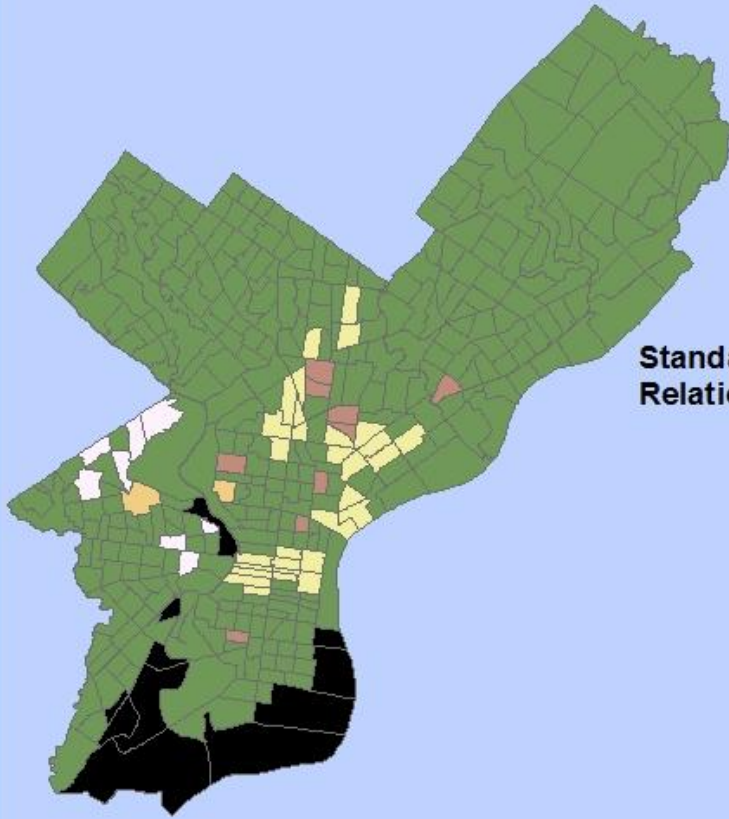
Regression Analysis for Property Crime Using Variables from Violent Crime Analysis		Stepwise Regression Results for Property Crime	
Variables	Confidence	Variables	Confidence
<b># White</b>	<b>&gt; 70%</b>	# Hispanic	> 95%
<b># Females</b>	<b>&gt; 90%</b>	# Age 18-21	> 99%
# Age 30-39	> 99%	# Households w/ 1 Male	> 99.9%
# Income < 200% Poverty	> 99.9%	# Income < 100% Poverty Level	> 99.9%
Regression Diagnostics			
Adjusted R <sup>2</sup>	0.309	OLS Adjusted R <sup>2</sup>	0.415
ANOVA (F-Test)	40.827	ANOVA (F-Test)	66.405
F Confidence	> 99.9%	F Confidence	> 99.9%
		GWR Adjusted R <sup>2</sup>	0.622

NOTE: All OLS and GWR regression analysis was done using the explanatory variables found from the stepwise algorithms

# Anselin Local Moran's I Statistic of Standardized Residuals from Ordinary Least Squares Regression Analysis

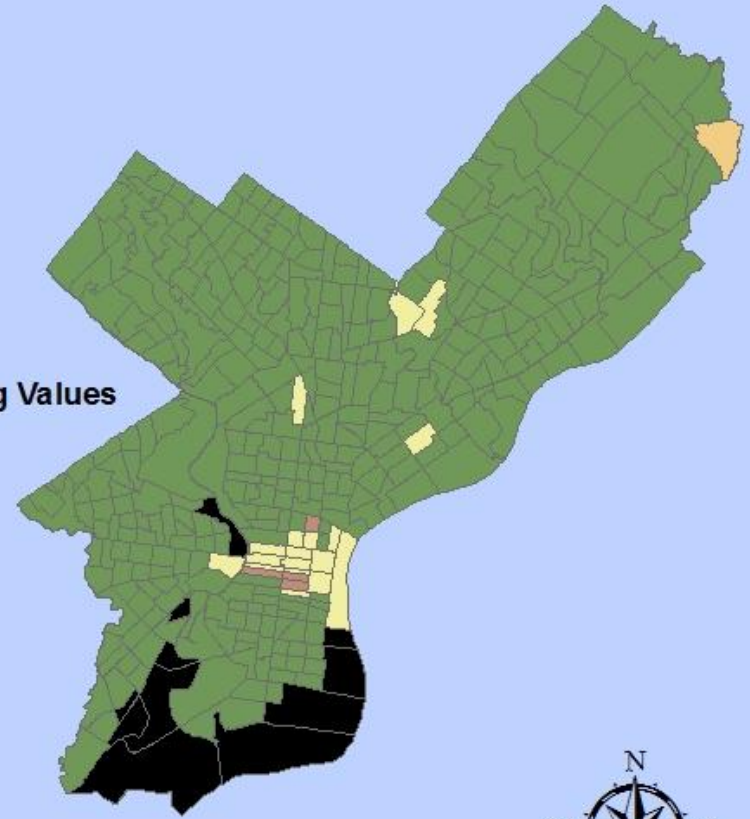
## Violent Crime

Global Moran's I Statistic 0.16  
Significance: 7.6 Standard Deviations



## Property Crime

Global Moran's I Statistic 0.17  
Significance: 8.28 Standard Deviations



### Standardized Residual / Relationship to Surrounding Values

- Black: Data Not Available
- Green: Not Significant
- Yellow: High/Similar
- Orange: High/Dissimilar
- Brown: Low/Dissimilar
- White: Low/Similar

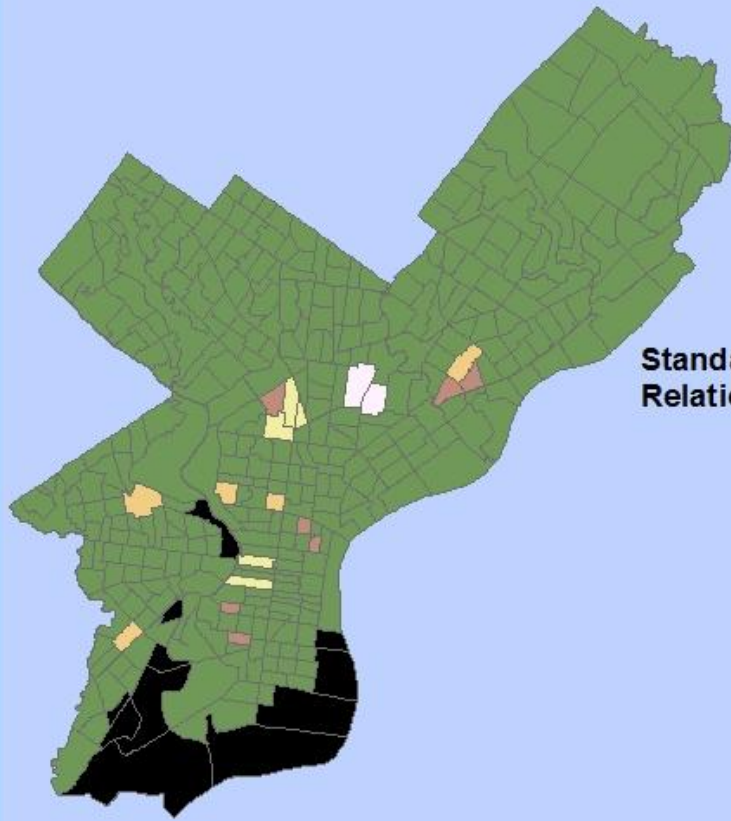
0 3 6 12 Miles



# Anselin Local Moran's I Statistic of Standardized Residuals from Geographically Weighted Regression Analysis

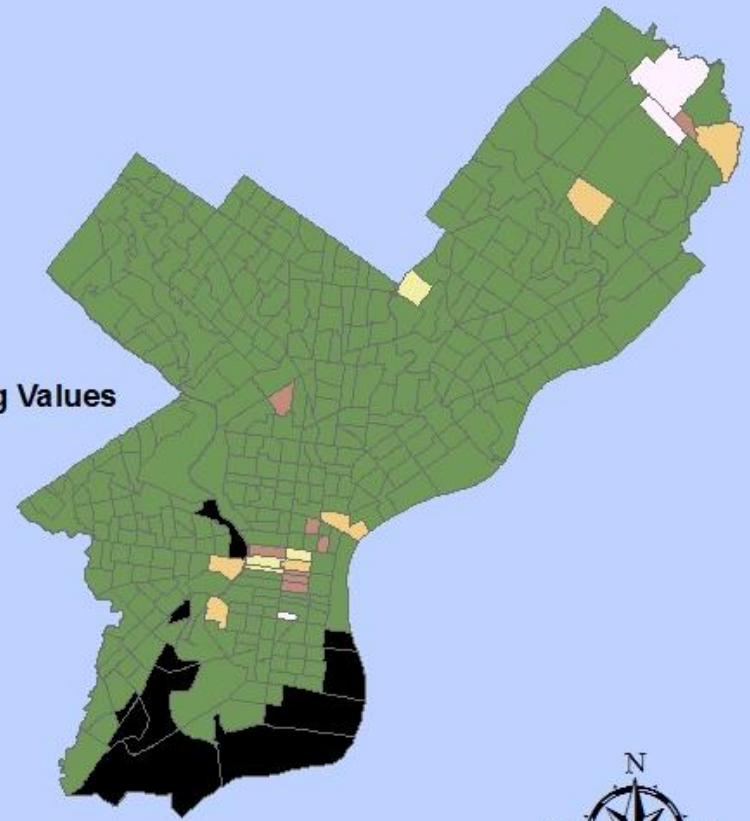
## Violent Crime

Global Moran's I Statistic -0.02  
Significance: -0.9 Standard Deviations



## Property Crime

Global Moran's I Statistic -0.02  
Significance: -0.78 Standard Deviations



### Standardized Residual / Relationship to Surrounding Values

- Black: Data Not Available
- Green: Not Significant
- Yellow: High/Similar
- Orange: High/Dissimilar
- Brown: Low/Dissimilar
- White: Low/Similar

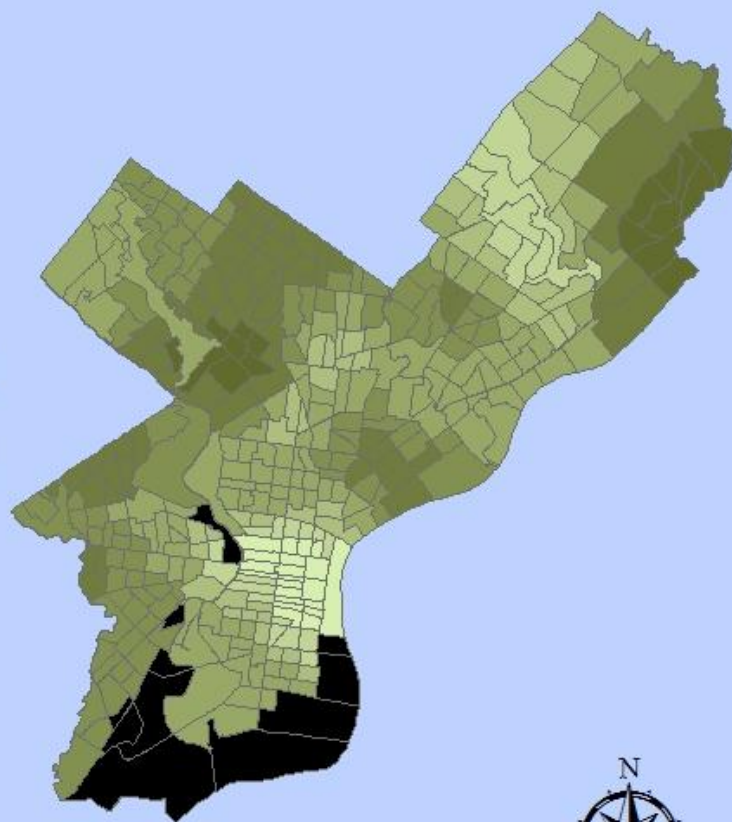
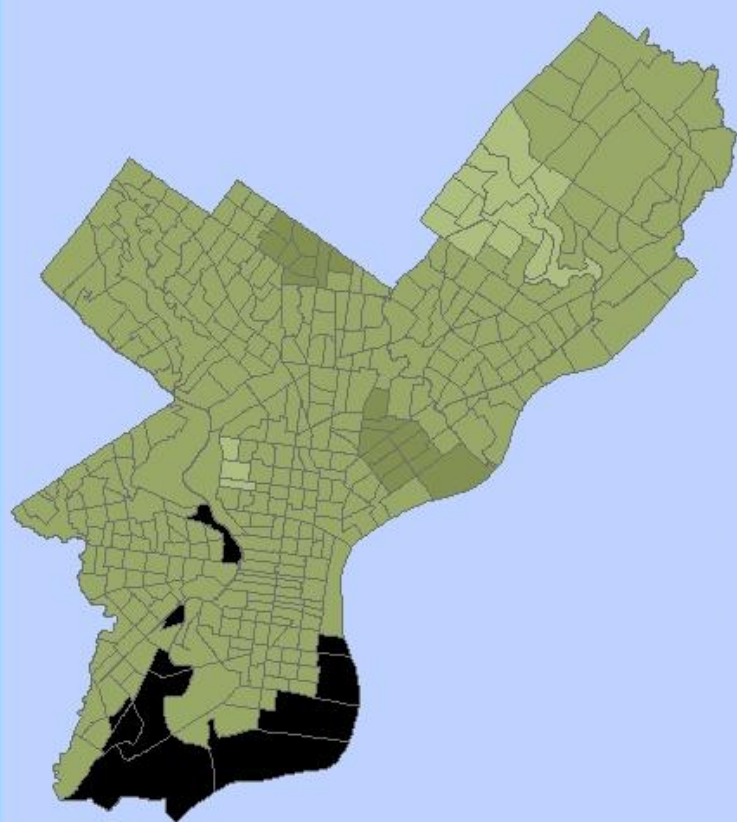
0 3 6 12 Miles



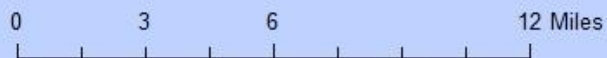
# Geographically Weighted Regression Coefficients for Population with Income Below 100% of Poverty Level

## Violent Crime

## Property Crime



### Coefficient Values



# Geographically Weighted Regression Coefficients for Race

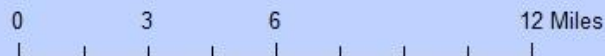
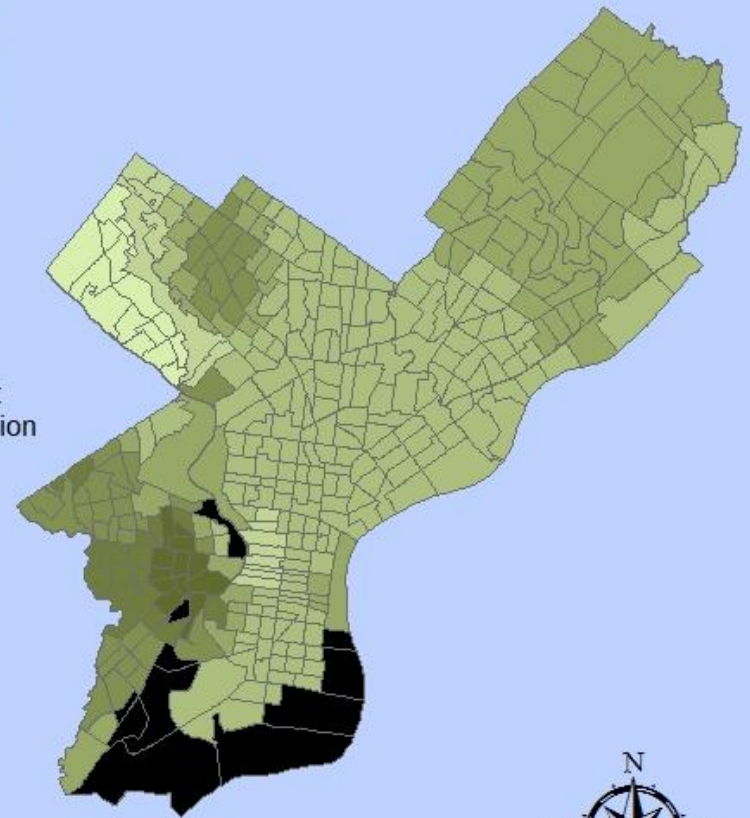
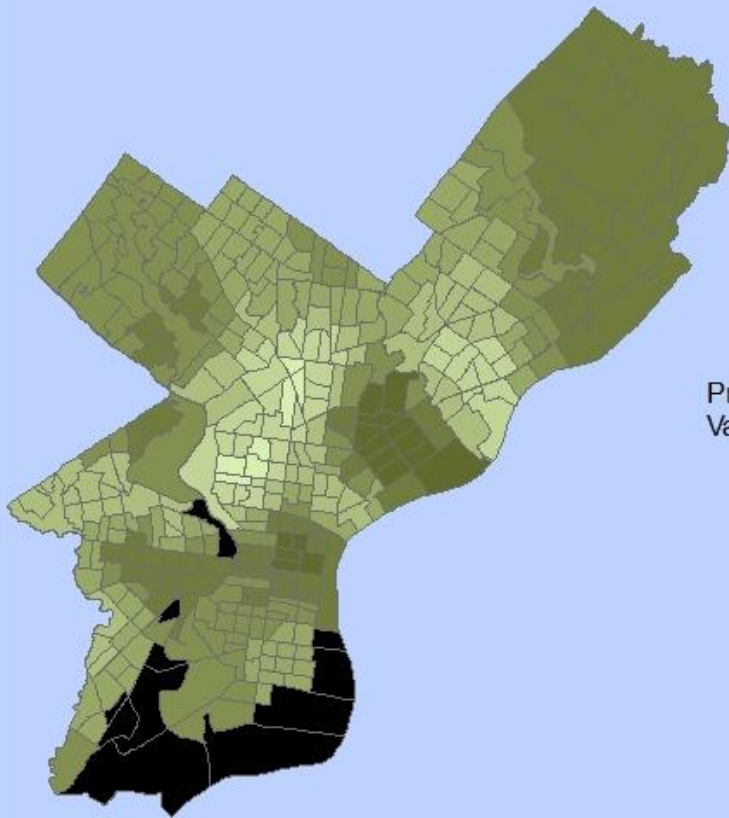
## Violent Crime

## Property Crime

Violent Crime - Coefficient Values for White Population



Property Crime - Coefficient Values for Hispanic Population



- Explanatory variables that perform well on violent crime prediction do not on property crime prediction
  - Although both regression models are significant according to F-Test, only 42% of property crime variation is explained by its independent variables compared with 76% for violent crime
  - Number of females and number of population that is white are significant variables for violent crime but not for property crime
  - Stepwise regression for property crime does not perform as well (lower  $R^2$ ) than for violent crime, suggesting that other variables than are available may better explain variations in property crimes
- The OLS residuals from the property crimes shows a higher degree of spatial autocorrelation than violent crimes
  - Probably because property crime's significant variables are a poorer fit in regression analysis
- The GWR residuals from property and violent crimes show no sign of significant spatial autocorrelation
  - Combined with the higher  $R^2$  values for GWR analysis, indicates that GWR is a good model for crime analysis in Philadelphia
- Population with income below 100% of poverty level has a more spatially varied and stronger effect for property crime than violent crime
  - Range of GWR coefficients is larger and vary more across Philadelphia for property crimes than for violent crimes
- Although race elements were significant explanatory variables for both types of crime, the Hispanic population accounts for greater variation in property crime (due to magnitude of the coefficients) than the white population does for violent crime
  - However, the effects (however minimal) of white population on violent crime vary more across Philadelphia than the effects of the Hispanic population on property crime

# Conclusions

- Additional Research
  - Investigate how independent variables correlate with one another, and those effects on the regression models
  - Perform additional analyses using different explanatory variables such as land use / land cover and a more specific set of explanatory variable based on existing research into what factors contribute most to violent and/or property crime
  - More rigorous statistical and visual analysis, including better understanding of regression model analysis, and comparison of two regression models with intersecting but not identical sets of explanatory variables
  - How to best account for missing data, especially in spatial analysis
  - Investigate effects of the modifiable areal unit problem and how it can affect results
  - Include measures of population density, another variable that has been linked to crime rates in previous research, in the analysis
  - Learn how to incorporate irregularly shaped areal units, and ones with low populations but high crime counts (presumably due to individuals committing crimes in a different census tract than which they live)
  - Use specific crime incident locations for point-pattern analysis
- Sources
  - Shapefiles: [Pennsylvania Geospatial Data Clearinghouse](#)
  - Crime Data: [CrimeBase](#) at UPENN's Cartographic Modeling Lab
  - Census 2000 Data: [ArcData from ESRI](#)
  - Stepwise Regression Add-In: [Boston University School of Management](#)
- References
  - 1: [Wikipedia](#) (via [The Philadelphia Inquirer](#))
  - 2: [Morgan Quitno](#)
  - 3: [The Cost of Inequality: Metropolitan Structure and Violent Crime](#)
  - 4: [Predicting Violent Crime using Urban and Suburban Densities](#)

# Additional Research and References