5-2014

An Analysis of the Patterns of Crime and Socioeconomic Status Visualized Through Self-Organized Maps

Jason Carlin Kaufman

University of Tennessee - Knoxville, jkaufma7@utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Jason Carlin Kaufman entitled "An Analysis of the Patterns of Crime and Socioeconomic Status Visualized Through Self-Organized Maps." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Geography.

Liem Tran, Major Professor

We have read this thesis and recommend its acceptance:

Nicholas Nagle, Bruce Ralson

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

An Analysis of the Patterns of Crime and Socioeconomic Status Visualized Through Self-Organized Maps

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Jason Carlin Kaufman
May 2014

Dedication

I owe a great deal of thanks for all the support I have had from my friends and family over the years. Thank you for listening and all you have done to help me. I must also thank Dr. Liem Tran, my primary advisor, for his invaluable help in completing this work; thanks as well as Drs. Nagle and Ralston for their assistance. It could not have been done without you. And lastly to Lauren Stachowiak who helped me see how far I needed to go, and Maria Martinez for helping me get there.

Abstract

This work is research to explore the association of spatial patterns between crime and socioeconomic status (SES) through the use of self-organized maps (SOM). It had been found that the spatial patterns of crime could be associated with those of socioeconomic, and this work sought to further these analyses in order to better understand how crime patterns and SES were related. To explore this association, patterns of crime and SES were examined in three cities: Nashville, TN; Portland, OR; and Tucson, AZ. Three SOMs were used in each city: one to analyze the patterns of crime, a second to analyze the patterns of SES, and a third to analyze the patterns of crime and SES. Nodes from each of these SOMs were also mapped to analyze the geographic distribution of their associated tracts. The results found an association between the patterns of crime and SES. In the Nashville Case Study, the patterns of high crime and low SES were not clearly associated in the combined Crime-SES SOM, but a stronger association was found in the geographic analysis. In the Portland Case Study, high crime and low SES patterns were found to be associated in the SOM. In the Tucson Case Study, high crime was found to be associated with low SES, but low SES was not always found to be associated with high crime. In each case study, the spatial patterns of low crime and high SES were found to be strongly associated. The spatial patterns of high crime were found to be associated with those of low SES, but the spatial patterns of low SES were not always found to be the same as those of high crime.

Table of Contents

List of Figures

1. Introduction

Criminologists, sociologists, law enforcement agencies and others interested in the occurrence of criminal behavior are continually seeking better ways to understand crime. Some criminological theories accept the idea that socioeconomic factors are a root cause of crime. This research focused on analyzing the similarities between spatial patterns of socioeconomic status (SES) and crime.

This study drew on three main research paths. The first was the self-organized map (SOM), created by Kohonen (1981). The second was research by Hagenauer, Helbich and Leitner (2011), who explored the relationship between crime and SES on a temporal scale in Houston, TX using an SOM. In their conclusion, they noted that future research might benefit from the inclusion of more crime and SES factors. This suggestion provided the genesis for this paper. The third path, the relationship between SES and crime, had been previously explored by the criminological theory of social disorganization, which was used to determine the measurements of SES to be included in this research.

The current work examined the association of spatial patterns between SES and crime using SOMs for three cities: Nashville, Tennessee; Tucson, Arizona; and Portland, Oregon. Each city was presented as a case study. In each of these case studies, the spatial patterns were analyzed using an SOM. Three research questions directed the analysis: what patterns in the component planes could be identified within the SOMs, how did the SOM help visualize these patterns, and how were the tracts from those patterns grouped within the city?

The results of this research suggest that there was an association of spatial patterns between crime and SES, but the results were mixed. The spatial patterns of high crime were found to be associated with those of low SES, but the spatial patterns of low SES were not always found to be the same as those of high crime. These results do suggest further research should be done with refinements to this method.

2. Literature Review

This work drew upon the clustering model of the SOM, crime research that has used the SOM, and the sociological theory of social disorganization. The SOM is a tool that re-projects multi-dimensional space onto a more easy to visualize two-dimensional map of nodes. In doing so, similar data points were located near each other based on their dimensional components. The work of Hagenauer, Helbich and Leitner (2011) was another key source; among their recommendations for future work was to add more sociological factors into their analysis. The sociological factors chosen for this analysis were consistent with ones presented in the criminological theory of social disorganization. The theory of social disorganization posits a causal relationship between SES and crime. This work does not present a causal relationship but for an association of spatial patterns between SES and crime.

2.1. Self-Organized Maps (SOM)

The SOM was originally proposed by Kohonen (1981) as an unsupervised classification system and is an analytical tool useful for modeling the organization of data points (Kohonen, 2001). The SOM visualizes a multi-dimensional dataset onto a created topology by organizing the dataset using unsupervised learning.

The creation of the SOM can begin with the user choosing the starting data point or by random selection (Kohonen, 2001). Each value of the data set is then added one by one and assigned to a node, based on the Euclidean distance of this data point to previously iterated data points; this is called the “learning process.” During the learning process, nodes that are close to each other will “learn” from one another. As the learning process continues, the Euclidean distances between data points are minimized, causing a smoothing of the weighted nodes in the SOM as data points are repositioned as necessary to create an optimal redistribution of the data set.

2.2. The U-Matrix for the SOM

The SOM geometry can be displayed in several ways. One method is the U-Matrix, developed by Ultsch and Siemon (1989), and Kraaijveld et al. (1992), and expanded upon by Ultsch (1993). In the U-Matrix, each node is mapped as a shape; this study uses a hexagon. In each node where data points were assigned, the node displays the scalar weight of the vectors assigned to that node. Each node is also surrounded by hexagons which display the scalar weight between nodes. The use of these hexagons in the U-Matrix displays how similar nodes are to each other.

2.3. Clustering Analysis Using the SOM

Various tools are capable of the analysis done in this work, such as *k*-means clustering, but the SOM was chosen for several reasons. The SOM is similar to *k*-means in that it produces a topology onto which it maps and uses a defined number of clusters for the model. The resultant geometry within the SOM displays scalar distances between points using the U-

Matrix. The SOM results also include the "component planes," which are visualizations of each dimensional component. Each component has its own geometry, allowing for comparisons of each component's influence on the SOM. As the purpose of this study was to compare how these components relate to each other, the SOM was an ideal model.

2.4. Crime and the Self-Organized Map

Hagenauer, Helbich, and Leitner (2011) used an SOM in a study of crime and SES. Their study conducted a temporal analysis of crime trends in Houston, Texas from August-September 2005. Using a spatio-temporal scan statistic, they mapped the change in crime rates over time and identified crime hotspots using a U-Matrix SOM. The crime hotspots compared the African-American population, the Hispanic population, the percentage of all persons below the poverty line, and the distance to the nearest police station. Their study found strong correlations of crime rates to minority populations and temporally to the landfall of Hurricane Rita. Included in their recommendations for future work was a suggestion to include more types of crime and socio-economic variables. It was this recommendation that provided the basis for this work.

2.5. Other Studies Using Self-Organized Maps

The SOM has been used in numerous fields to evaluate multivariate datasets. Friedel (2011) used a SOM to visualize potential hazards in a post-fire landscape. Chon et. al (1996) and Aguilera et. al. (2001) used SOMs to evaluate pollution and water quality. Chon et. al. (1996) also proposed the use of their results as a training module, allowing for easier classification of pollution responses in invertebrates. Tran et. al. (2003) used a SOM along with principal component analysis to evaluate environmental degradation in the Mid-Atlantic region of the US. These four works used smaller data sets, but others such has Lin, Chen and Nunamaker (1999) and Merdun (2011) used a SOM to visualize much larger data sets relating to document analysis and soil samples, respectively. This is a very small subset of the works that have used a SOM; the preface to Kohonen (2001) noted that as of that date more than 4000 different publications had used a SOM.

2.6. Social Disorganization

While there are a number of criminological theories, this study draws on social disorganization theory, which argued that crime was endemic to particular areas of a city. The work of Park and Burgess (1925) was instrumental among early theories; they proposed that cities are made of concentric zones. These zones were, in expanding order from the city center, the Central Business District (CBD), the Zone in Transition, the Zone of Working Men's Homes, the Residential Zone, the Commuters' Zone, the Agricultural District and the Hinterland (Figure 1). Of these zones, the Zone in Transition was the focus of social disorganization theory. Crime in this zone was often higher than in the rest of the city, and social disorganization theory attempted to explain why these areas were more criminogenic.

The Zone in Transition was a result of city growth (Park & Burgess, 1925). As the outer zones saw new investment and property value growth, the older and more financially

established inhabitants moved outward from the city center. As these citizens moved out, property values fell in their former neighborhoods, and the new inhabitants tended to be less financially successful. Park and Burgess developed this model at a time when these new inhabitants were often newly arrived poor European immigrants or African-Americans from the American South. This process, which Park and Burgess compared to biological secession, created fractured neighborhoods, resulting in social problems such as crime, poverty, and low education. Social disorganization theory argued that the conditions in these neighborhoods created a feedback loop of low SES, keeping property values low and thus encouraging the prosperous to leave; the result was a cycle of decay.

This theory was expanded upon by the research of Shaw and McKay (1942), which used data collected from Chicago over a 30-year period. To test the theory, Shaw and McKay mapped male juvenile delinquents who had been brought before a criminal court and sentenced. This criterion was more restrictive than the crime data used in this study, which was based on the location of the crime rather than where the criminal lived. In limiting their inquiry to convicted delinquents, Shaw and McKay estimated that they had removed 85% of the cases that occurred during the study period. The cases that were removed were ones where the incidents were resolved by the arresting officer without going to a trial. The remaining crimes brought to trial were therefore assumed to be more serious crimes (Shaw & McKay, 1942).

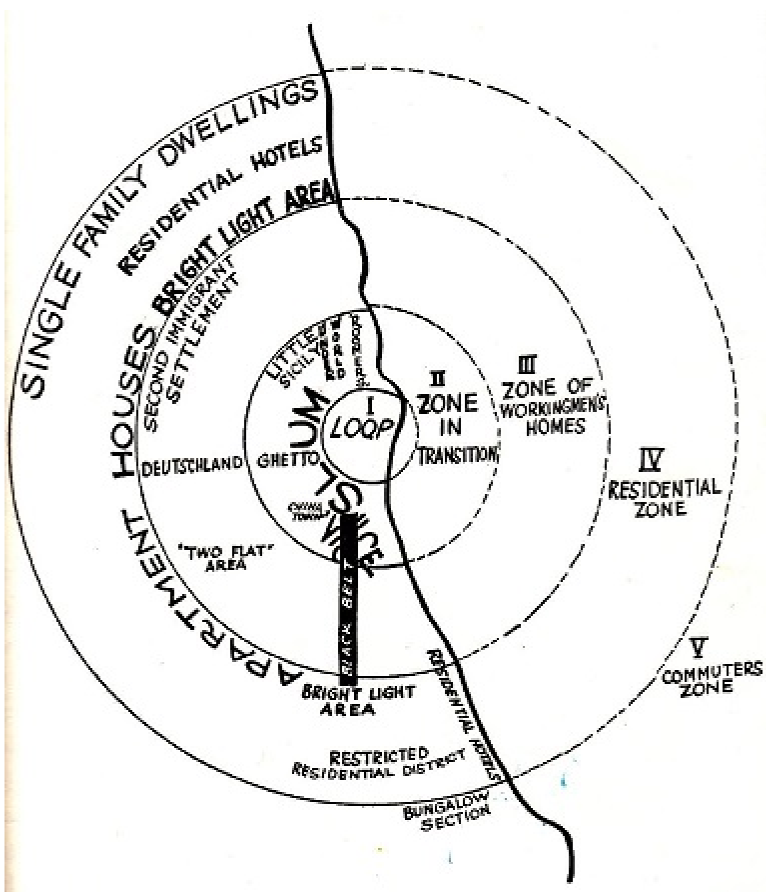

Figure 1 - Zones of the City (Park & Burgess, 1925)

The length of the study period of Shaw and McKay's work was instrumental in arguing that crime was endemic to areas of the city, rather than the persons living there. Early 20th century Chicago experienced a rapid overturn of population driven by immigration. New migrants moved into neighborhoods in the Zone in Transition, leading to large scale population replacement. Shaw and McKay showed that the highest crime areas remained in the same part of the city over the study period despite population replacement. The static nature of criminogenic areas in Chicago argued that the location and characteristics of these areas were responsible for criminal activity.

Shaw and McKay compared indicators of low SES to the crime rates of Chicago during their study period. These indicators included population change, number of structures condemned, percentage of families on relief, median rental costs, home ownership percentages, percentage of African-American and foreign born family heads, school truancy, infant mortality, rates of mental disorder and rates of tuberculosis. When these were mapped, neighborhoods in the Zone in Transition were found to have the lowest levels of SES.

Shaw and McKay's analysis of Chicago was used in this work primarily as a tool to direct the expanded research from Hagenauer, Helbich, and Leitner (2011). Drawing upon Shaw and McKay's method of analysis for low SES, measures similar to those of Shaw and McKay (1942) were included where possible when it was not possible to directly match their measures of SES.

3. Methods

This study analyzed crime reports and measures of SES from three cities (Nashville, TN; Portland, OR; and Tucson, AZ) to determine whether there was an association of spatial patterns between crime and SES. The analysis used tract-level data from the 2000 U.S. Census and crime data collected from the Metro Nashville Police Department, Portland Police Bureau, Tucson and South Tucson Police Departments and was compiled by Cahill (2004). The various categories of crime and measures of socioeconomic status that were used in the analysis were normalized on a 0-100 scale (Table 1). Data histograms for each variable and maps of selected variables are seen in Appendices I and II. Three SOMs were created for each city: a crime only SOM, an SES only SOM, and a third that combines both crime and SES into a single SOM. The SOMs were created using a software package created by Kohonen for use in MatLab.

The Nashville, Portland, and Tucson SOMs were structured into 5x7 matrices (35 nodes), 5x8 matrices (40 nodes) and 5x6 matrices (30 nodes) respectively. These sizes were chosen so that an average of four tracts per node was assigned in each SOM.

Two evaluations of distance were used: Tract Distance and Node Distance. Each tract has a calculated distance based on the value of each relevant variable assigned to that tract. Within the multivariate space, those values are the distance of that variable to the zero point in the center. The distance of the tract from the center is the Euclidean distance calculated from each variable within that tract, referred to in this study as the Tract Distance. Node Distance is the average Tract Distance for all tracts assigned to that node.

Table 1 - List of all the categories of crime and SES used in this study, and the identifier used in the Component Plane figures in the following case studies

Violent Crime	HOM: Homicide SEX: Sexual Assault ROB: Robbery AGG: Aggravated Assault
Non-Violent Crime	BURG: Burglary LARC: Larceny MVT: Motor Vehicle Theft JUV: Juvenile Crime*
Socioeconomic Status	ICE: Index of Concentration at the Extremes MHI: Median Household Income PCI: Per Capita Income PIP: Percentage of Total Population Living Below the Poverty Line PU: Percentage of All Adults in the Workforce Who Are Unemployed
Demographic Status	PSP: Percentage of All Parents Who Are Single Parents PAA: Percentage of All Persons Who Identify as African-American PL: Percentage of All Persons Who Identify as Hispanic PU18: Percentage of All Persons Who Are Under the Age of 18.

** Note: Juvenile Crime was not used in the Nashville case study due to a lack of data*

3.1. Research Questions

In order to determine how the patterns of crime and SES relate to each other within each case study, three questions were used to analyze each SOM.

1. What patterns in the component planes could be identified within the SOMs?

The analysis of each SOM began with an analysis of the global patterns of the component planes. While local patterns were also considered, of particular focus was the maximum and minimum of each component plane, as well as identifying how the values in each component plane compared to the others at these points. Patterns outside the extremes, in particular where several component planes had shared values, were also noted to for further analysis.

2. How did the SOM help visualize these patterns?

Three methods of evaluating the SOM were used to help evaluate the patterns identified in the previous question. These methods were an analysis of the Node Distance, analysis of how the tracts were distributed in each node of the SOM, and an analysis of the Standard Deviation of the Node Distance.

The Node Distance helped evaluate global, neighborhood, and local patterns. Node Distance was calculated by first finding the Euclidean distance of each tract in the dataset to the zero point (0, 0, … , 0). It was then calculated by averaging the calculated Euclidean distance of each tract assigned to that node. The standard deviation of the Node Distance was calculated using the Euclidean distance of all tracts assigned to that node.

The Node Distance was used with the tract distribution per node and clusters of similar nodes from the U-Matrix and component planes to assist in evaluating the boundaries of a cluster and determining common patterns, as well as evaluating the global patterns in the component planes. An analysis of the Node Distance determined if nodes were clustered; nodes with similar values of Node Distance were more likely to be part of the same cluster.

The distribution of tracts per node in the SOM identified over- or under-filled nodes. Analysis of over- or under-filled nodes identified how often a particular pattern was found in the dataset. Under-filled nodes suggest uncommon patterns; over-filled nodes grouped together show that that pattern was very common. Together with the Node Distance and the analysis of the component planes, the tract distribution helped identify which patterns were most common in the SOM.

The standard deviation of the Node Distance was also examined for each node to evaluate local patterns. The standard deviation within a node can help determine the similarity of assigned tracts: low values indicate that the assigned tracts were similar, high values indicate dissimilarity.

Neighborhood analysis was aided by identifying groups of nodes with similar values in each SOM. Nodes were identified as similar by their component plane values, their Node Distance values, how those nodes were related to adjacent nodes on the U-Matrix, and by tract distribution. These node groups were referred to as Node Types in this analysis and given a letter from A-E, depending on the SOM. Node Type was used to refer to these similar nodes within that SOM only.

3. How were the tracts grouped in SOM nodes distributed geographically the city?

This question was investigated by using nodes sampled from each SOM to determine whether clustering found in the SOM was also found in the corresponding city. The sampled nodes were chosen for having a high number of associated tracts and for an association with a local maximum of a component plane. These nodes were then mapped onto each city according to their associated tracts. The analysis here was both qualitative and quantitative based on the selection of nodes and an analysis of clustering within each model on the geographic space using Moran's I on the Tract Distance to determine spatial autocorrelation. Moran's I can be calculated using several different distance metrics for autocorrelation and, because the case studies were comprised of tracts of variable size, polygon continuity was chosen as the method with which to calculate Moran's I.

4. Case Study I: Nashville, TN

Nashville is characterized by a small central business district surrounded by decaying older neighborhoods, with more modern and stable neighborhoods moving outward from the city center (Figure 2). The city is separated into two parts by the Cumberland River. The central business district is located south of the river and is bounded by it on the north and east. Nodes in the Nashville Case Study SOMs were numbered as in Figure 3.

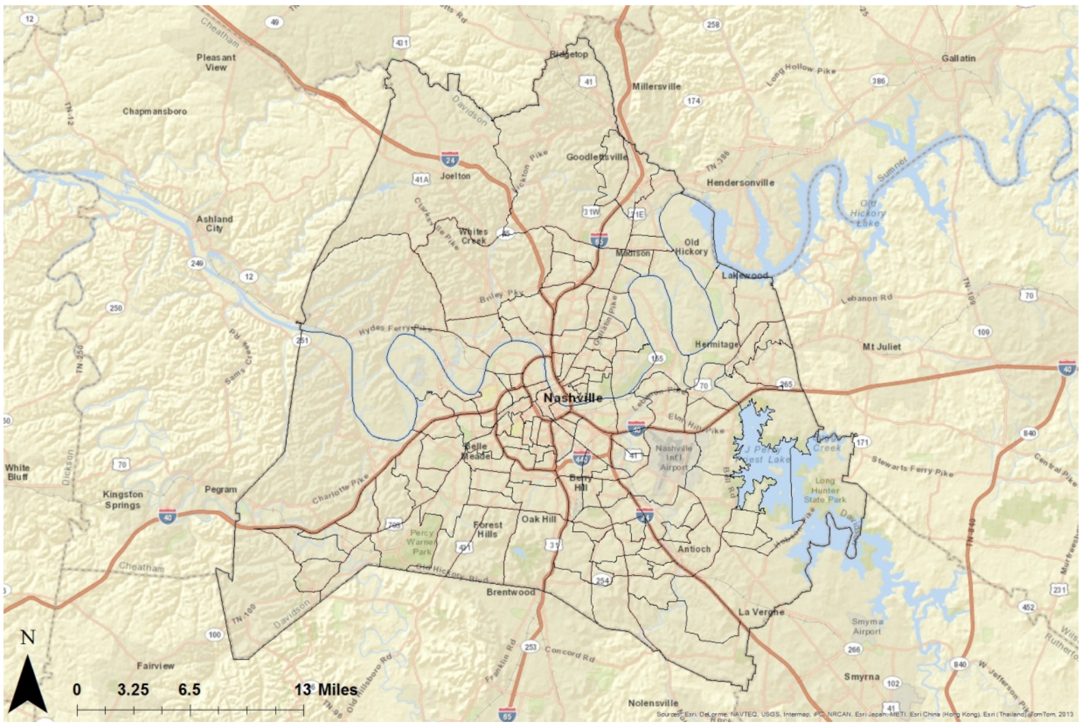

Figure 2 – Nashville, TN study area, with tracts

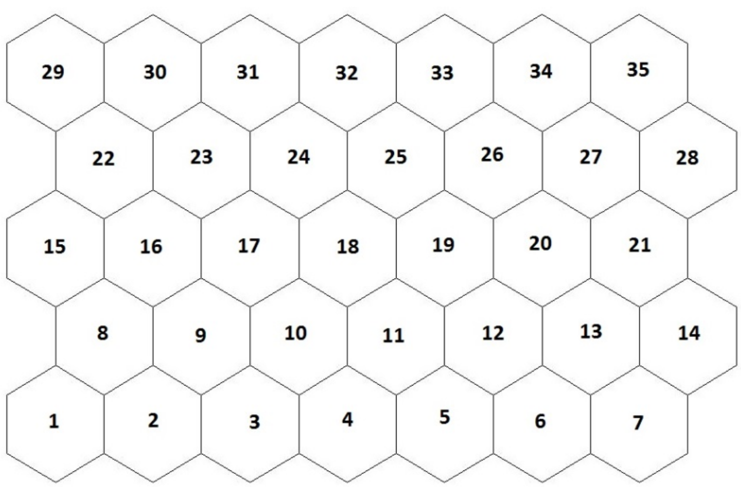

Figure 3 – SOM Node Numbers for Nashville, TN

4.1. Nashville Crime SOM

This case study found four types of nodes: Type A in the top left, Type B in the lower left, Type C in the lower right, and Type D nodes in the rest of the SOM. Violent crime was shown as highest in the lower left, and non-violent crime has a bipolar maximum in the bottom left and right. Crime values appear at their lowest in the top left. The analysis did not find any high tract density nodes associated with moderate crime values. These patterns were analyzed using the component planes (Figure 4), analysis of the Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance (Figure 5), and geographically (Figure 6).

4.1.1. What patterns in the component planes could be identified within the SOMs?

The patterns of the component planes fall into two main groups: violent and non-violent crimes (Figure 4). The four violent crime planes shared a general pattern of high crime in the lower left and low values of crime throughout the rest of the SOM. There was some variation to this pattern: robbery and aggravated assault was more common in Nodes 3-7, but the crime maximum was shared among the four planes.

Non-violent crime had a different pattern. In all three non-violent crime planes, there was a split distribution of the highest values of crime in the lower left and lower right, moderate crime in between these maximums and in the upper right, and low crime in the top left.

On the whole, crime patterns in this SOM have similar spatial patterns. Across all seven planes the highest and lowest values of crime were shared in the lower and upper left respectively. The spatial patterns of violent crimes were similar to each other, as were the spatial patterns of non-violent crimes.

4.1.2. How did the SOM help visualize these patterns?

Further analysis of the SOM was done using the Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance (Figure 5). Node Distance values were highest in the lower left and lowest in the top left. A secondary maximum was also found in the lower right. These maximums and minimums corresponded to the overall crime maximum, the crime minimum, and the non-violent crime maximum. As the global maximum was found in the lower left and the global minimum was found in the upper left, the primary axis of change across the SOM did not encompass the global minimum. The change in Node Distance across the two primary axes was roughly equal. Among the component planes, the majority were identified with a primary change axis from the bottom left to the top right.

Changes between nodes were largest along the left edge of the SOM, as was also seen in the U-Matrix (Figure 4). This is consistent with the global maximum and minimum located on the left side. Tract distribution was also closely associated with the minimum and maximum values of Node Distance, with the highest number of tracts at those nodes and few tracts assigned to other nodes. There were no high tract density nodes with moderate crime values. There is not a clear pattern to the Standard deviation of the Tract Distance, but nodes with higher internal variation are more likely to be high tract density nodes.

Moving from the global pattern to neighborhood analysis, four groups of nodes were identified in this SOM. These four groups were mapped in Figures 4 and 5 and labeled Type A,

B, C and D. Type A nodes were in the top left and were the crime minimum node in every component plane. These nodes had low Node Distance values, some internal variation, and many were high tract density nodes. Type B nodes were in the lower left, and were characterized by very high crime values, high Node Distance, and high internal variation. There was only one high tract density node in this Node Type. Type C nodes were in the lower right, and were characterized by very high non-violent crime values and high Node Distance. Most of the tracts associated with this Node Type were assigned to Node 7. Type D nodes were the most common, characterized by moderate crime values, moderate Node Distance values, low internal variation, and very low tract density nodes. These were nodes in the middle and upper right of the SOM. Nodes from Types A, B, and C were mapped in the next section.

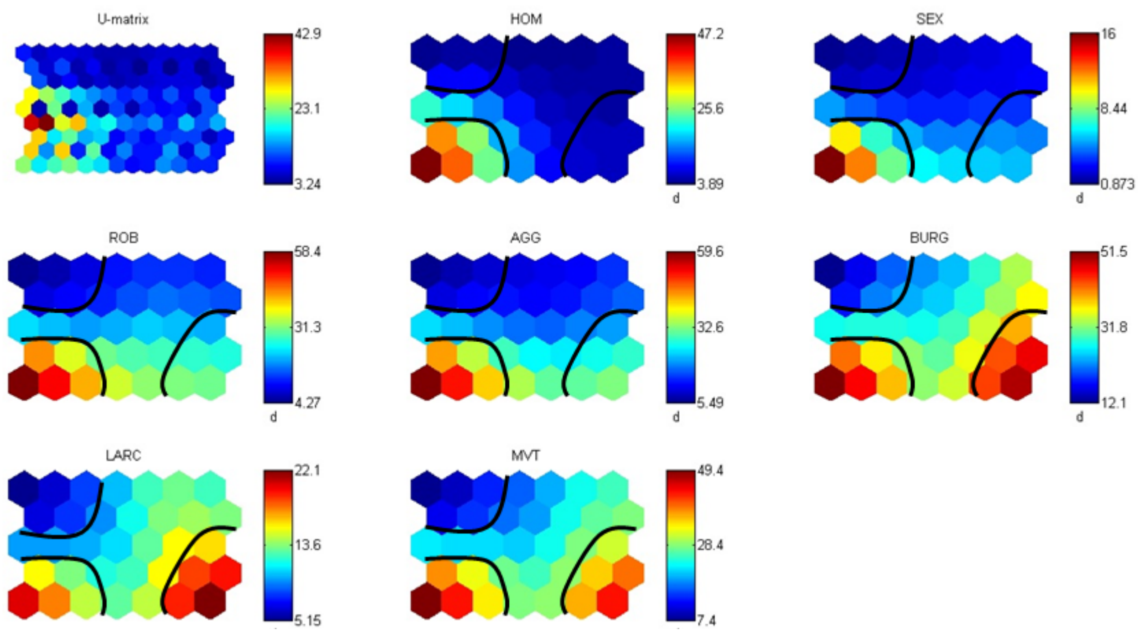

Figure 4 – Component planes and U-Matrix for the Nashville Crime SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the top left were Type A nodes, nodes in the bottom left were Type B nodes, nodes in the bottom right were Type C nodes, and nodes in the middle were Type D nodes. See Section 4.1.2 for further discussion.

4.1.3. How were the tracts grouped in SOM nodes distributed geographically the city?

Three Nodes (1, 7, and 29) were mapped in Nashville (Figure 6). These three nodes were associated with Node Types A, B and C respectively. Tracts associated with Node 1 were mostly found in two clusters near downtown. Despite the geographic separation, created by the Cumberland River, the tracts on both sides of the river were likely related to one another. Node 29 also showed very strong clustering to the south of the city. While some of the tracts associated with this node were scattered, the scattered tracts tended to be found at the edge of the city. There was no clustering to the tracts assigned to Node 7. Using the Tract Distance for each node, Moran's I was calculated to test for spatial autocorrelation. That result suggests

that crime was geographically clustered in Nashville, returning a Z-score of 6.07 and an associated p-value of 0.

4.1.4. Summary of the Nashville Crime SOM

These results clearly showed two crime patterns in Nashville: a general crime pattern and a non-violent crime pattern. These two patterns were noted as Node Types A and B. Nodes associated with moderate crime could be identified in between the high crime patterns, but there was no clear pattern of moderate crime as none of those nodes have a large number of tracts associated with them; many of the one-tract nodes were associated with moderate crime values in this model. When high tract density nodes associated with very high and low crime were mapped, two of them also showed strong clustering.

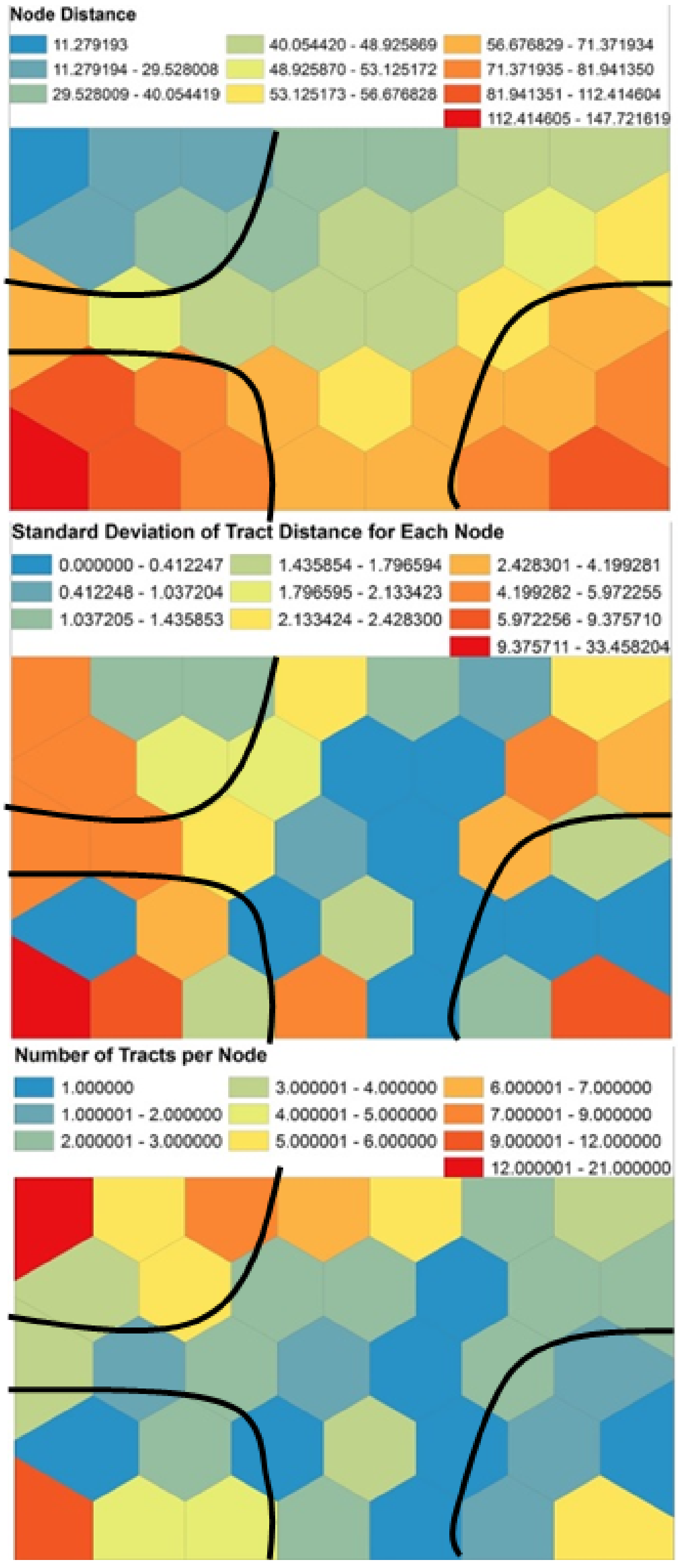

Figure 5 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Nashville Crime SOM. The drawn boundaries were the same as those in Figure 4 to denote Node Types A, B, C and D.

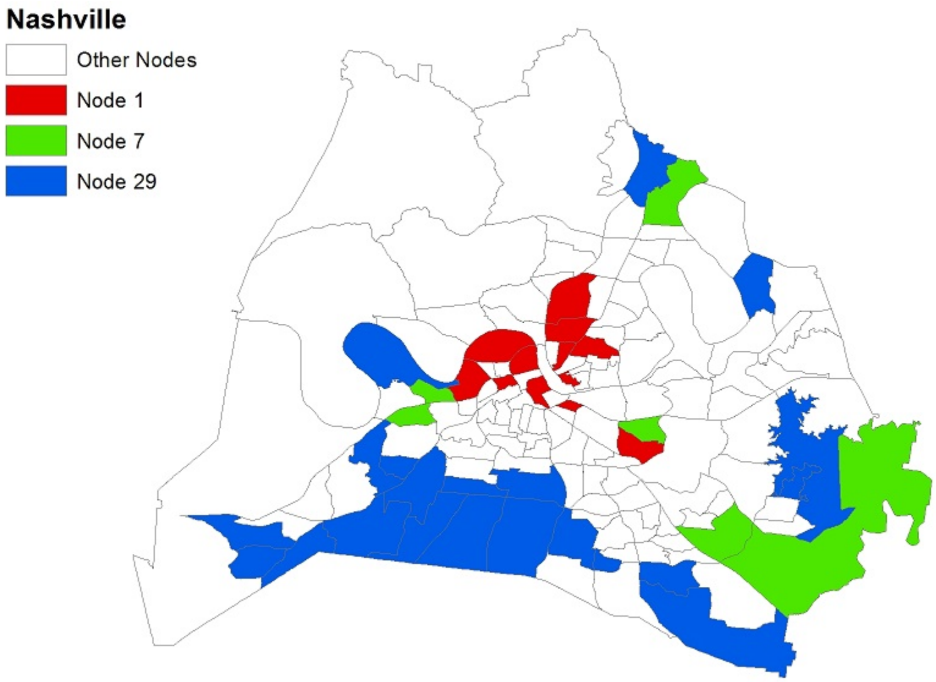

Figure 6 – Tracts associated with Node Types A, B and C in the Nashville Crime SOM. Node Type A was represented by Node 29, Node Type B was represented by Node 1, and Node Type C was represented by Node 7.

4.2. Nashville SES SOM

The analysis of the Nashville SES Case Study found four types of nodes: Type A in the top right, Type B in the lower right, Type C in the lower left, and Type D nodes in the rest of the SOM. While each of these was distinguishable by changes in their component planes, the upper right and middle patterns were not distinguished by changes in node values or tract distribution. These patterns were analyzed using the component planes (Figure 7), analysis of the Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance (Figure 8), and geographically (Figure 9).

4.2.1. What patterns in the component planes could be identified within the SOMs?

There were two patterns comprised of multiple SES component planes in this SOM (Figure 7). The income related component planes – ICE, median household income, and per capita income – shared a similar pattern. Four component planes made up a second pattern: single parents, African-American heritage, poverty, and unemployment. In general, where these four components were high, the income planes were low, and vice versa. This was expected for income, poverty, and unemployment; as income falls, poverty rises. Based on the values of the component planes, it appears that people in poverty were more likely to be single parents and African-Americans in Nashville.

The remaining two component planes had their own pattern. The percentage of persons who were children had a pattern similar to that of the low income pattern, but the main axis of this pattern ran from the bottom right to the top left. The component plane for Hispanic heritage was different from all of the above and did not appear to influence any of the others.

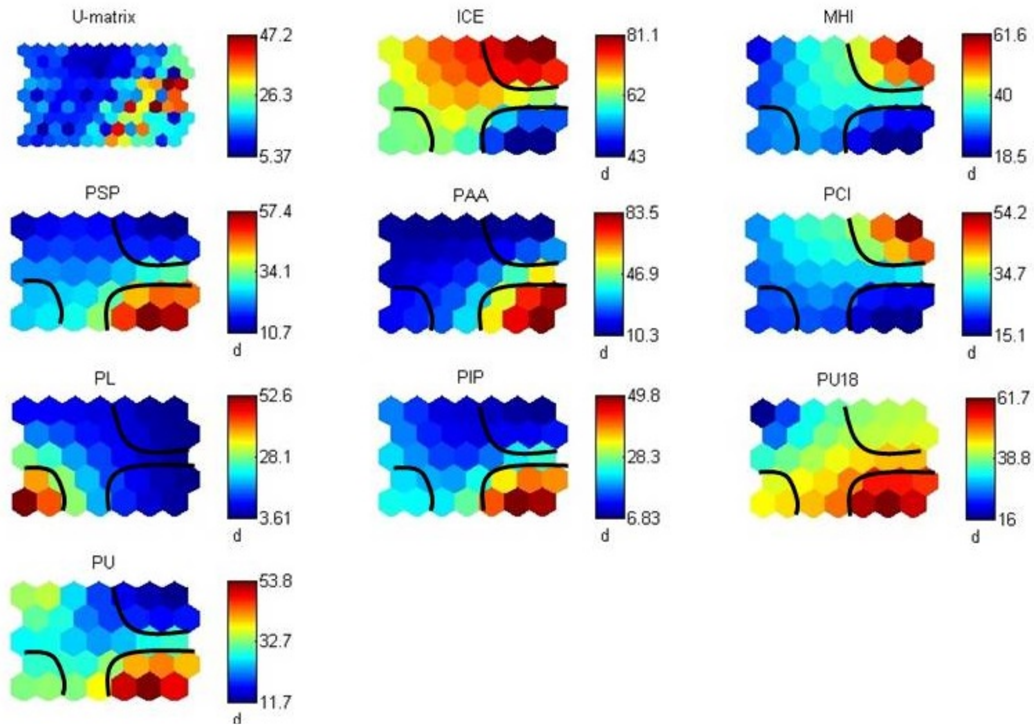

Figure 7 – Component planes and U-Matrix for the Nashville SES SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the top right were Type A nodes, nodes in the bottom right were Type B nodes, nodes in the bottom left were Type C nodes, and nodes in the middle were Type D nodes. See Section 4.2.2 for further discussion.

4.2.2. How did the SOM help visualize these patterns?

Node Distance showed an overall pattern of decreased node distance from right to left in the SOM, with three null-value nodes on the right (Figure 8). The highest values were found in the upper and lower right. The decrease across the SOM was gradual, and this could be seen in the large cluster on the left side of the U-Matrix (Figure 7). High tract density nodes were more common at the edge of the SOM (Figure 8), and were associated with patterns in the SOM. The Standard Deviation of the Tract Distance was again most associated with Tract Distribution.

The primary change axis was from the lower right of the SOM to the top left. This primary axis explained the change across five of the component planes, but did not encompass any significant change in income values across the SOM. Node Distance did not help with this evaluation due to how income planes were displayed. Very high SES and very low SES were both mapped as high values in these models, which reduced the use of Node Distance in evaluating the primary axis in SES SOMs. This was the reason why Figure 8 shows multiple high value nodes for Node Distance.

On the neighborhood scale, four types of nodes were identified and were mapped in Figures 7 and 8. Type A nodes were located in the top right, characterized by very high to moderate values of income and multiple high tract density nodes. Type B nodes were in the bottom right and were the SES minimum nodes in this model. There were also several high tract density Type B nodes. Type C nodes were in the lower right, and were notable as the only set of nodes where persons of Hispanic heritage were most common. Type D nodes comprised the rest of the set. Many of these nodes were low tract density nodes, and the values in all component planes except the percentage of children and ICE were low. Two Type A nodes, one high and one moderate income, were mapped in the next section, as were one Type B and Type C node each.

4.2.3. How were the tracts grouped in SOM nodes distributed geographically the city?

The four nodes were analyzed for geographic clustering (Nodes 1, 6, 32 and 35). These nodes were associated with Node Types C, B, and the last two with Type A nodes respectively. Tracts in Node 6 were clustered, but separate from any other tract. Nodes 32 and 35 were geographically separate, but not by a large distance. What was not expected was that tracts from Node 1 were adjacent to those in Node 32, as these nodes were very distant from one another in the SOM. Tracts in Nodes 6 and 35 show the greatest degree of clustering, while Nodes 1 and 32 show some clustering. Tract distance was again used to calculate Moran's I to test for spatial autocorrelation. The test indicates that SES was geographically clustered in Nashville, returning a Z-score of 6.46 and an associated *p*-value of 0.

4.2.4. Summary of the Nashville SES SOM

In this SOM, four types of nodes were identified. Low SES/Type B nodes were identified in only the lower right corner of the SOM, and the high income/Type A nodes were found in the top right corner. Outside of these nodes there was little change in the SOM, which was seen in the large number of Type D nodes with similar component plane values.

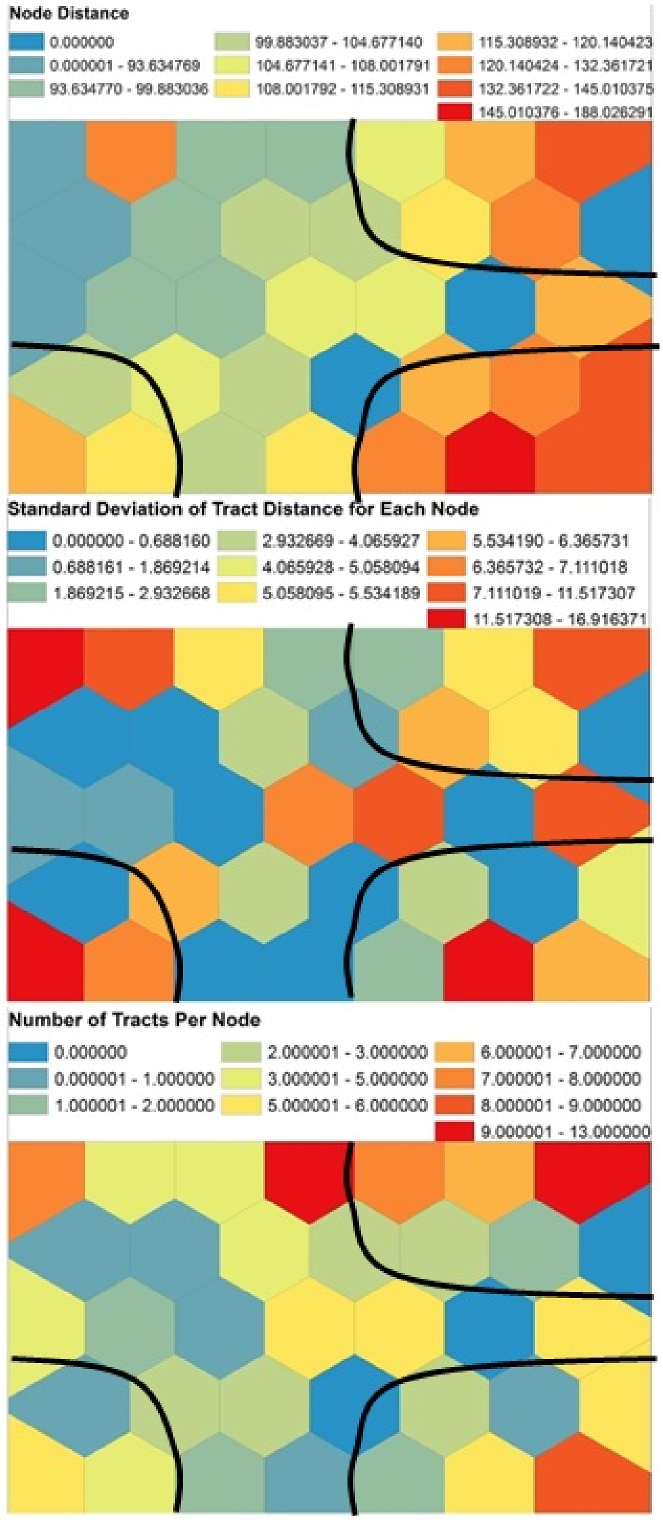

Figure 8 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Nashville SES SOM. The drawn boundaries were the same as those in Figure 7 to denote Node Types A, B, C and D.

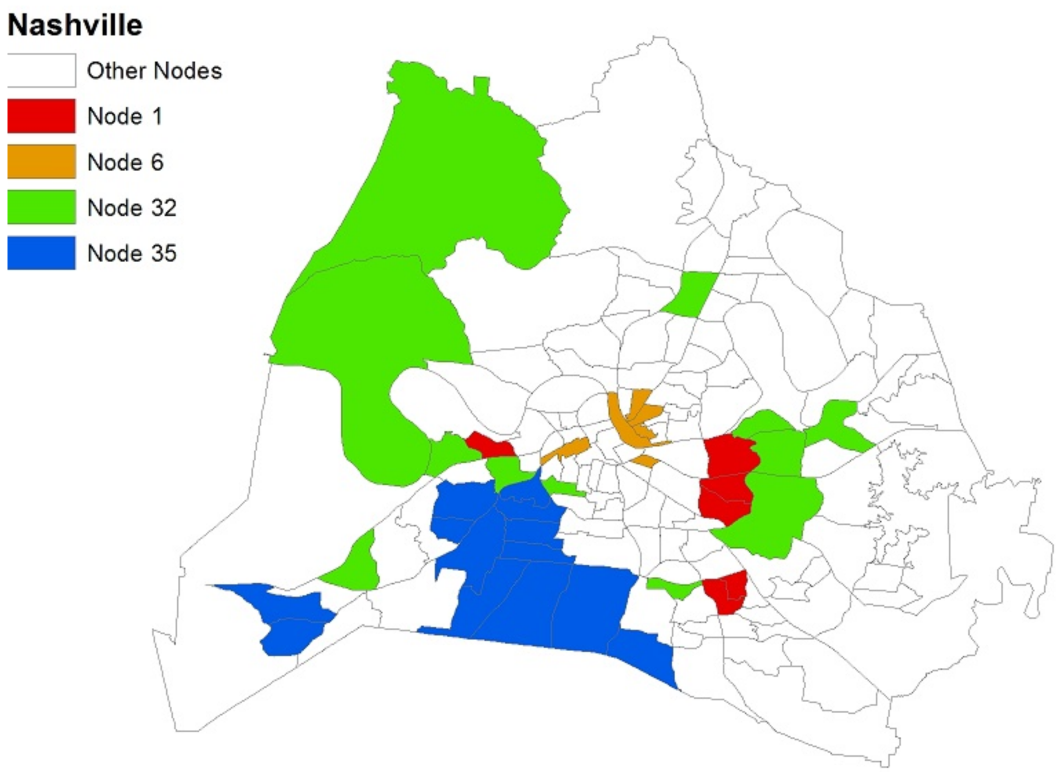

Figure 9 – Tracts associated with Node Types A, B and C in the Nashville SES SOM. Node Type A was represented by Nodes 32 and 35, Node Type B was represented by Node 6, and Node Type C was represented by Node 1.

4.3. Nashville Crime and SES SOM

Analysis of the combined SOM identified four types of nodes. These were Type A nodes in the lower left, Type B nodes in the bottom right, Type C nodes in the top right, and the remainder of the SOM was composed of Type D nodes. Node Types A and B were identified with different SES and crime values. This suggested there was no single SES pattern associated with crime. Component plane values associated with Type C nodes were mostly identical to similar nodes previously identified in the SES SOM. The crime minimum was also associated with Type C nodes. These results were analyzed using the component planes (Figure 10), analysis of the Node Distance, the number of Tracts per Node and Standard Deviation of the Tract Distance (Figure 11), and geographically (Figure 12).

4.3.1. What patterns in the component planes could be identified within the SOMs?

The patterns of crime and SES were largely unchanged from their respective SOMs. The main crime pattern was on the left of the SOM (Figure 10), with the highest values shared between Nodes 1 and 2. This pattern was found across all the component planes. A second pattern in the lower right was not as continuous, varying in strength across the seven crime planes.

The patterns of SES were largely identical to those noted previously. While household and per capita incomes were nearly identical, the pattern for ICE decreased much more slowly than the income planes. Of the remaining SES component planes, African-American heritage and percentages of single parents were most similar to one another, as were unemployment and poverty. The percentage of people who are children was also similar to this pattern. Separate from each of these was the pattern for Hispanic heritage.

4.3.2. How did the SOM help visualize these patterns?

Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance are seen in Figure 11. Node Distance values were highest in the lower left and right and lowest in the middle of the SOM. Five null-tract nodes helped separate the higher values in the bottom left and right from the lower values in the top middle. A number of high tract density nodes were located around the edges of the SOM. Standard Deviation of the Tract Distance was minimal and, again, best associated with Tract Distribution.

The primary change axis in this SOM was the lower left-upper right axis. This axis encompassed most high crime and the crime minimum in each component plane. The primary axis also explained most income variation, but the minimum in each of those planes was in the secondary axis. All of the remaining components were explained by the secondary axis with the exception of Hispanic heritage.

There were four types of nodes at the neighborhood scale, and these were identified in Figures 10 and 11. Within Type A nodes were the primary crime nodes in this SOM. There was very high internal variation within these nodes. Type B nodes were the SES minimum in this SOM, and these nodes were moderately criminogenic. Despite four nodes with similar values in this Node Type, most tracts were assigned to the lower right node only. Type C nodes were the high and moderate income nodes. These nodes were also associated with the crime minimum in this SOM, and a very large number of tracts were assigned to these nodes. Lastly, Type D nodes comprised the remainder of the SOM with relatively uniform values in all component planes. Most of these nodes had low tract densities.

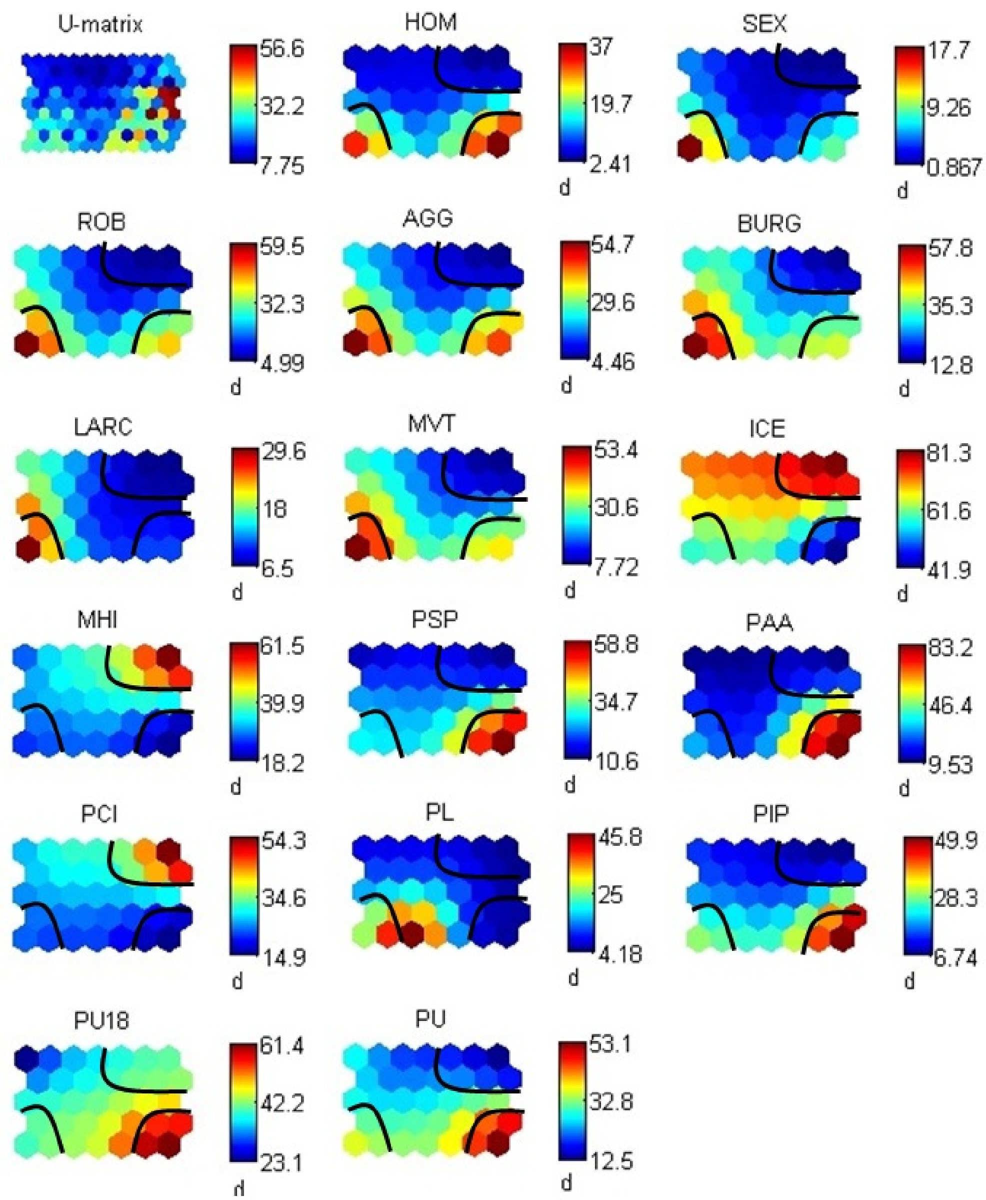

Figure 10 – Component planes and U-Matrix for the Nashville Crime-SES SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the bottom left were Type A nodes, nodes in the bottom right were Type B nodes, nodes in the top right were Type C nodes, and nodes in the middle were Type D nodes. See Section 4.3.2 for further discussion.

4.3.3. How were the tracts grouped in SOM nodes distributed geographically the city?

Six nodes (1, 3, 7, 15, 32, and 35) were mapped to analyze for geographic clustering (Figure 12). One of each of these nodes was associated with Node Types A and B and two each were associated with Node Types C and D. These nodes were split between geographic clustering

and scattering. Nodes 7 and 35 showed the strongest clustering, Nodes 1 and 3 showed some clustering, and Nodes 15 and 32 were scattered.

The tracts assigned to Node 35 were unchanged from their selection in the SES-only SOM, which suggests that the inclusion of crime played no role in that node or the related pattern. Similarly, tracts assigned to Node 7 were primarily from the high crime Node 1 in the crime SOM and the low income Node 6 in the SES SOM. Node 1 contained the remnant of the previous node combination and, as the remnant, it was not surprising that this node had less geographic clustering. Node 3 was very similar to Node 1 in the SES SOM in component planes and tract association. Node 15 was also similar to Node 7 in the Crime SOM. The only significant change could be seen in Node 32, which was almost entirely dissimilar to its equivalent SES node. This change may be why there was also no clustering to this node, though the reason for the change was unclear.

As there was some clustering in four of the six analyzed nodes, it was not surprising that when Tract Distance was analyzed in a Moran's I test, the returned Z-score is 7.20 with an associated *p*-value of 0.

4.3.4. Summary of the Nashville Crime-SES SOM

The results of this SOM suggest that there were four node groups, three of which are associated with the highs and lows of SES and crime. The bottom left was a crime maximum for six of the seven crime planes, and it was associated with slightly higher levels of poverty and unemployment. The lower right node group was partitioned from the rest of the SOM by two null value nodes and other low tract nodes. The values of the component planes associated with the equivalent SES-SOM pattern were nearly identical to those in this pattern. The addition of moderate values of crime was the only change here, and the difference previously noted between violent and non-violent crime patterns was not present. The top right and middle node group was also largely unchanged from the related patterns in the SES SOM; the addition of crime did not change these patterns.

These results show that the combined SOM was largely comprised of unchanged SES patterns from the SES-only SOM. The lowest crime values were associated with high and moderate income, and the highest crime values with the lowest levels of income. While moderate crime was found in the lower right, it did not cause the SES patterns seen here to change from their previous structure. The changes in crime patterns were therefore likely the result of the inclusion of the SES components.

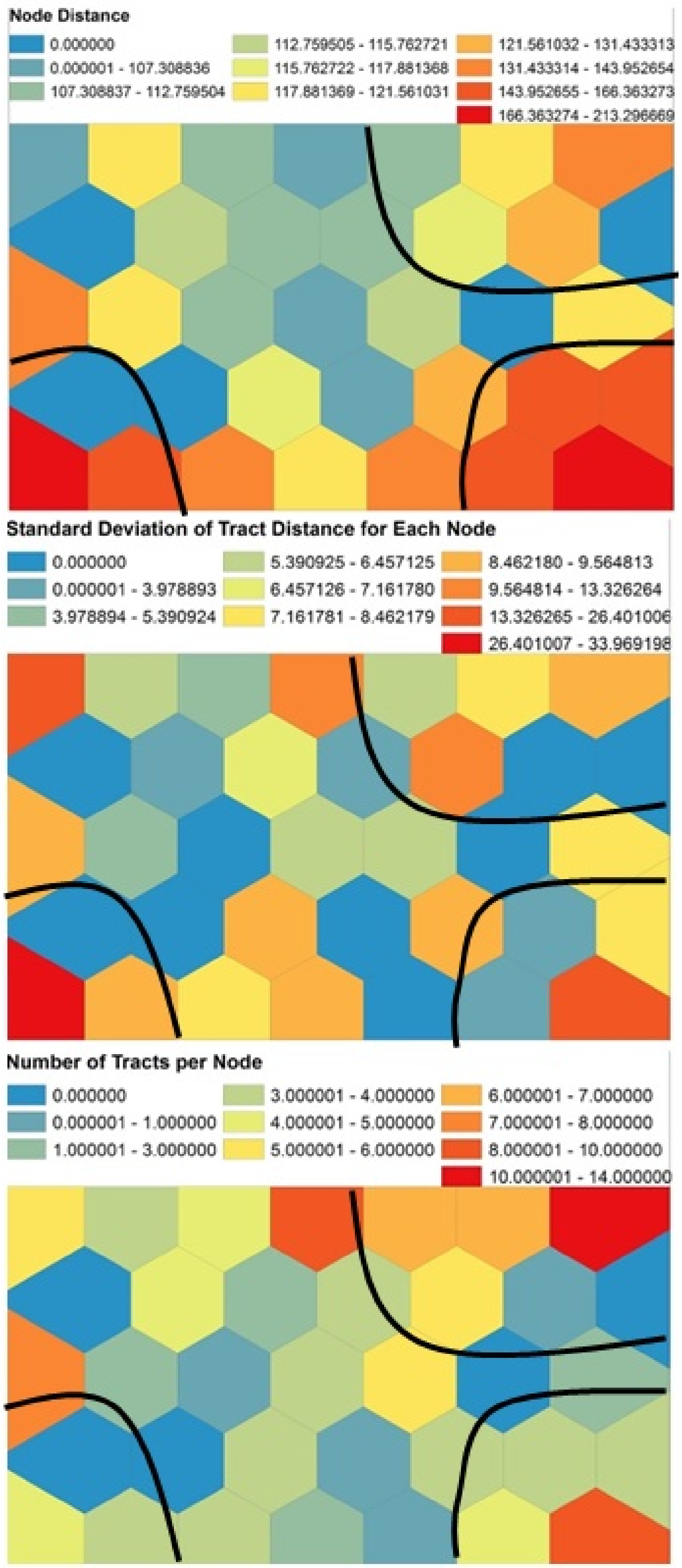

Figure 11 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Nashville Crime-SES SOM. The drawn boundaries were the same as those in Figure 10 to denote Node Types A, B, C, and D.

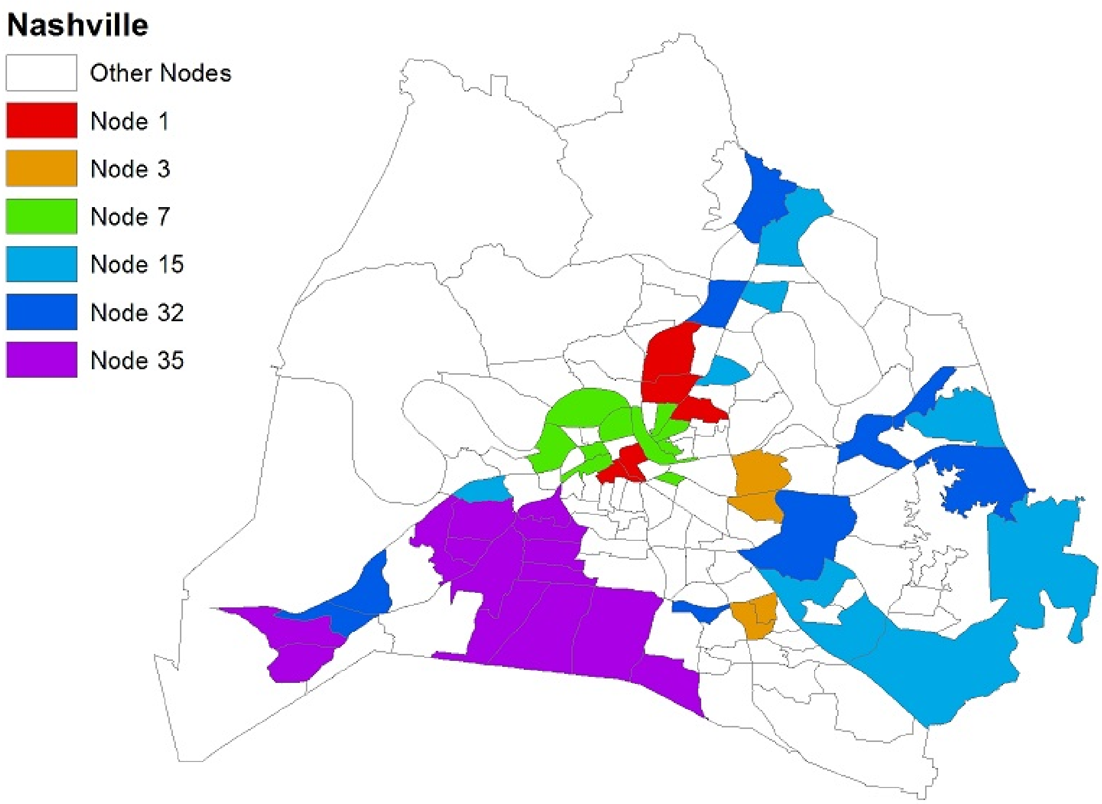

Figure 12 – Tracts associated with Node Types A, B, C, and D in the Nashville Crime-SES SOM. Node Type A was represented by Node 1, Node Type B was represented by Node 7, Node Type C was represented by Nodes 32 and 35, and Node Type D was represented by Nodes 3 and 15.

5. Case Study II: Portland, OR

Portland is defined by the flow of two rivers: one creating a northern border, the other splitting the downtown (Figure 13). To the east and west of downtown are residential areas, with the western the most affluent. Of the three cities in this case study, Portland has the smallest minority population, with no single minority comprising more than 10% of the population. Nodes in the Portland Case Study SOMs were numbered as in Figure 14.

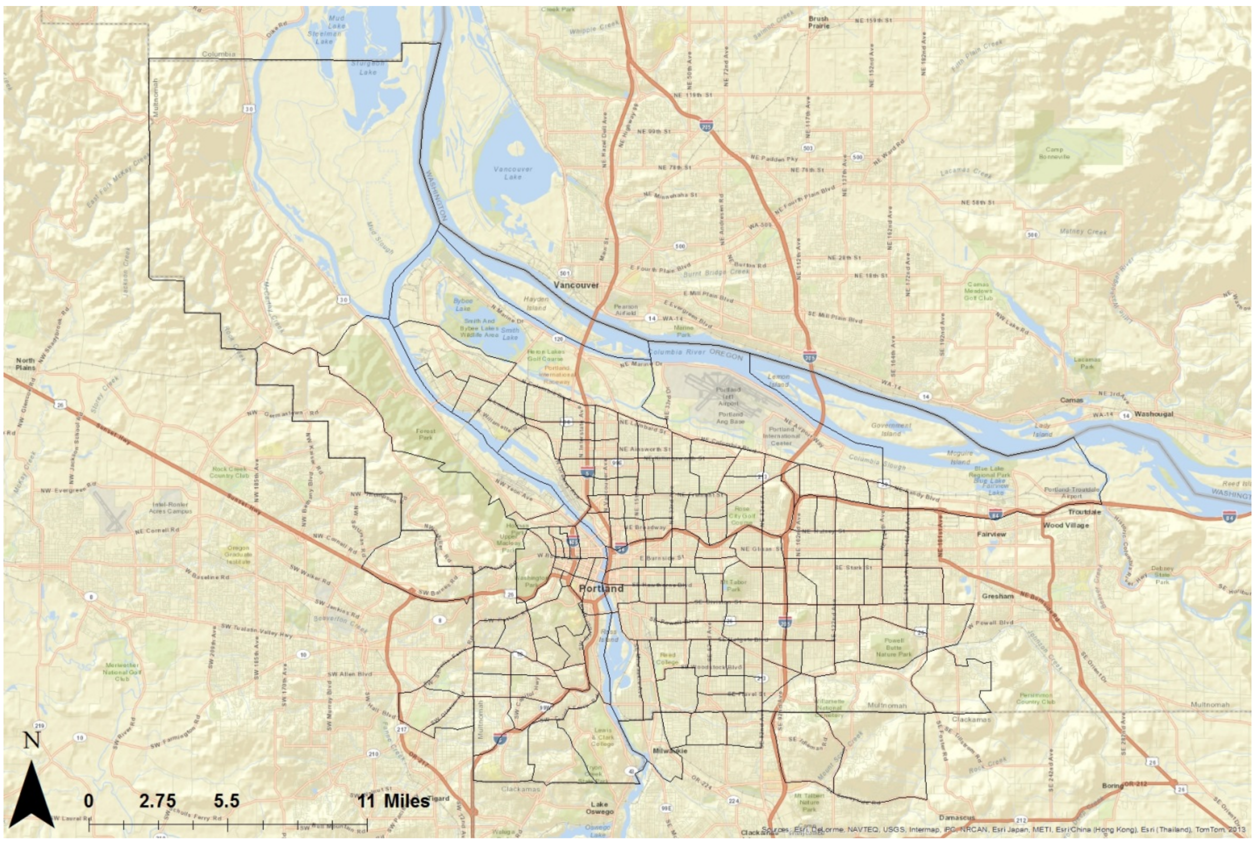

Figure 13 – Portland, OR study area, with tracts

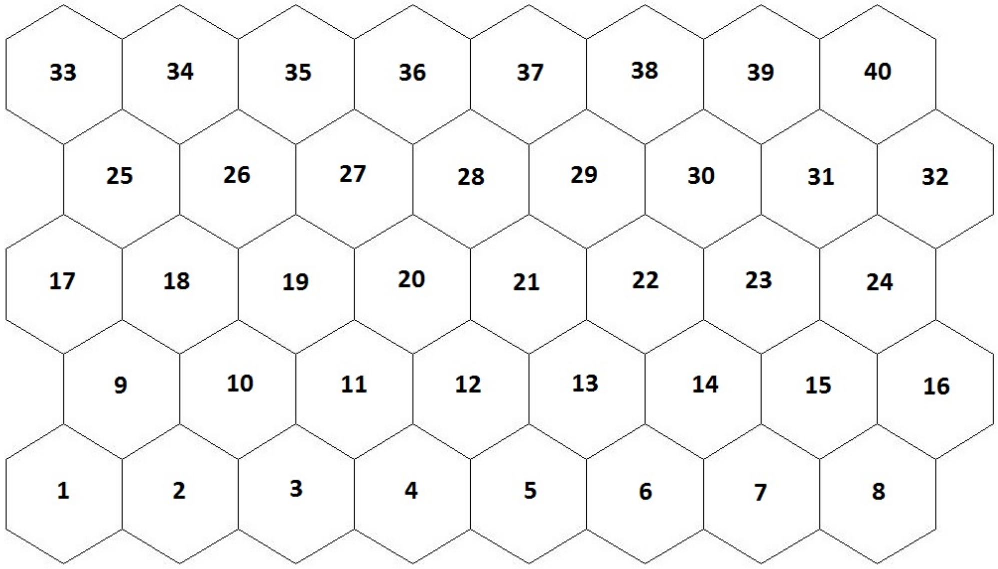

Figure 14 – SOM Node Numbers for Portland, OR

5.1. Portland Crime SOM

There were four main types of crime nodes in the Portland Crime SOM. The lower left (centered on Node 1) and right (Node 8) both showed mixed patterns of high crime. There were a large number of nodes grouped on moderate values of crime. These nodes were analyzed using the component planes (Figure 15), analysis of the Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance (Figure 16), and geographically (Figure 17).

5.1.1. What patterns in the component planes could be identified within the SOMs?

The component planes fall into two main patterns (Figure 15). In the first pattern, component planes on the left showed that homicide, aggravated assault and juvenile crime were highest, sexual assault and burglary were moderately high, motor vehicle theft and robbery were common, and larceny was rare. This was unlike the Nashville SOM, where there was a clear split between violent and non-violent crime patterns. On the whole, violent crimes were more common in more nodes than non-violent crimes in Portland.

The second pattern was simpler, located in the lower right: all crime was high except homicide, which was rare. In both of these patterns, crime values decrease quickly from their maximum and were lowest in the top middle left nodes. The component planes with split maximum nodes also showed a much slower decrease in crime values than the single maximum node component planes.

5.1.2. How did the SOM help visualize these patterns?

In analysis of Node Distance, the number of Tracts per Node and Standard Deviation of the Tract Distance, the presence of crime minimum was ambiguous due to eight null-tract nodes (Figure 16). Crime maximums were split to the lower left and right. Crime decreased

moving from the bottom to the top of the SOM. Nodes with four or more assigned tracts were in a minority, and were more commonly located on the right of the SOM. Node 35 may have been responsible for the fractured pattern, as there were 25 tracts assigned to this node. This was the crime minimum. The Standard Deviation of the Tract Distance was generally similar to the Tract Distribution, but there were several low tract nodes with high standard deviation.

Globally, the primary axis of change across the SOM was the lower right to top left axis. This axis explained a significant percentage of crime in every component plane except homicide. Homicide alone was aligned along the second axis, as well as secondary maximums in six other component planes.

Four neighborhood node groups were identified in Figures 15 and 16. The top left nodes were identified as Type A nodes in this analysis. These nodes were the crime minimum in every component plane except homicide, but the node with the highest tract density was offset in the top middle. The reasons for this were not clear. Type B nodes were in the lower left and were identified as a secondary crime maximum. Despite that crime values were generally lower in Type B nodes, the Node Distance values here were very high. Type C nodes were the crime maximum in the lower right. While crime values were uniform in the component planes, the internal variation in these nodes was very high. Type D nodes comprised the remainder of the SOM. While many of these nodes were underpopulated, several did show a concentration of tracts with moderate values of crime.

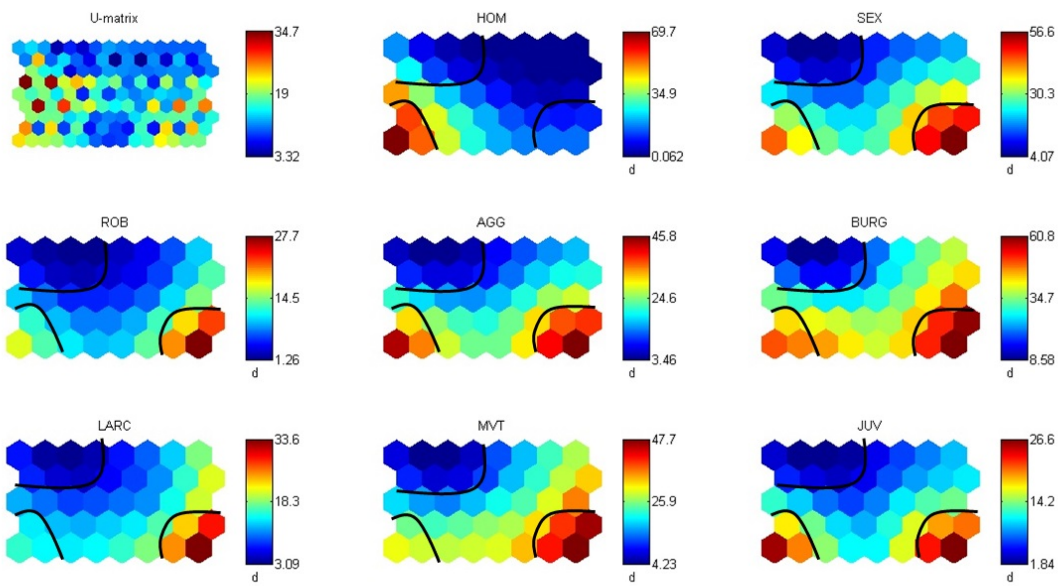

Figure 15 – Component planes and U-Matrix for the Portland Crime SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the top left were Type A nodes, nodes in the bottom left were Type B nodes, nodes in the bottom right were Type C nodes, and nodes in the middle were Type D nodes. See Section 5.1.2 for further discussion.

5.1.3. How were the tracts grouped in SOM nodes distributed geographically the city?

Six nodes were mapped, chosen for high numbers of assigned tracts (Figure 17). Nodes 1 and 8 were associated with crime maximums, Node 35 the crime minimum, and 3, 12, and 19 were associated with moderate crime values. Node 1 was representative of Type B nodes, Node 8 of Type C nodes, Nodes 3, 12 and 19 of Type D nodes, and Node 35 of Type A nodes. All of these nodes were geographically scattered. The clearest pattern was for Node 35, which was found around the edge of the city. The high and moderate crime tracts were found closer to the city center. Despite the scattered nature of this clustering, the Moran's I analysis of spatial autocorrelation returned a Z-score of 7.41, indicating that there was spatial autocorrelation of crime in Portland. This was likely because crime was lowest towards the edge of the city and higher towards the city center.

5.1.4. Summary of the Portland Crime SOM

This analysis shows that there were four types of crime nodes in Portland: two associated with high crime, one associated with moderate crime, and one associated with low crime. The two sets of high crime nodes were very different in their associated crime values and neither was comprised of solely violent or non-violent crime. There was little geographic correlation of crime values in this SOM.

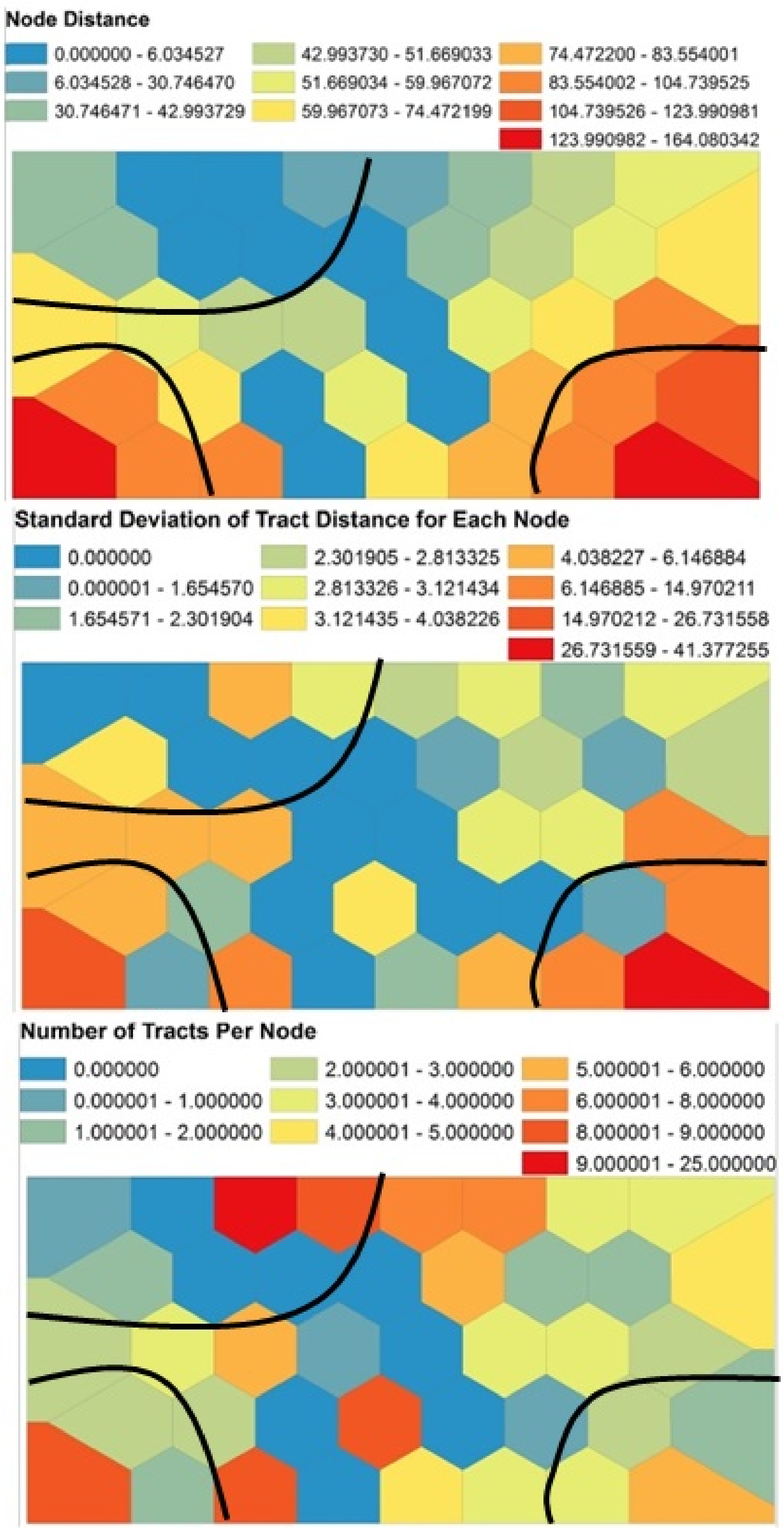

Figure 16 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Portland Crime SOM. The drawn boundaries were the same as those in Figure 15 to denote Node Types A, B, C and D.

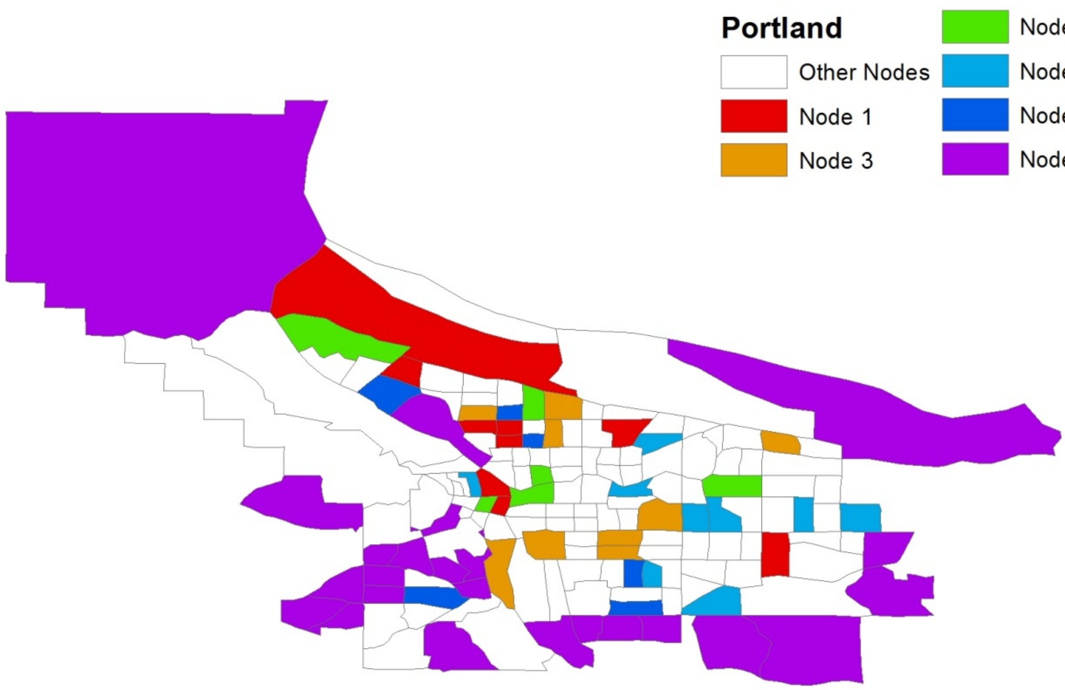

Figure 17 – Tracts associated with Node Types A, B, C, and D in the Portland Crime SOM. Node Type A was represented by Node 35, Node Type B was represented by Node 1, Node Type C was represented by Node 8, and Node Type D was represented by Nodes 3, 12 and 19.

5.2. Portland SES SOM

Four main types of nodes could be found in the analysis of the SES SOM. In addition, several high tract density nodes were found to be scattered in the largest of those Node Types. These patterns were variously defined by the component planes (Figure 18), the Node Distance, number of tracts per node, and Standard Deviation of the Node Distance (Figure 19), and geographically (Figure 20).

5.2.1. What patterns in the component planes could be identified within the SOMs?

There were three main patterns to the SES component planes (Figure 18). One was the change in income from the top left, where ICE and income were highest, to their minimum in the bottom right. Where income was lowest, unemployment and poverty was the most common; where income was highest, poverty, unemployment, and single parenthood was least common.

The top right-bottom left diagonal was similar, with ICE and income moderately high in the top right, and poverty and unemployment moderately low in the bottom left. Single parents were most common in the lower left, unlike the diagonal found in the Nashville results. African-American and Hispanic families were most common in the bottom left nodes.

The last component plane for the percentage of children did not align well with either of the previous patterns. High percentages of children were found where single parents were most common, but the percentage of children was still high in the top left where single parents were least common. Comparing directly to income, children were more common in high income families and less in low income families, which reverses the results of the Nashville case study.

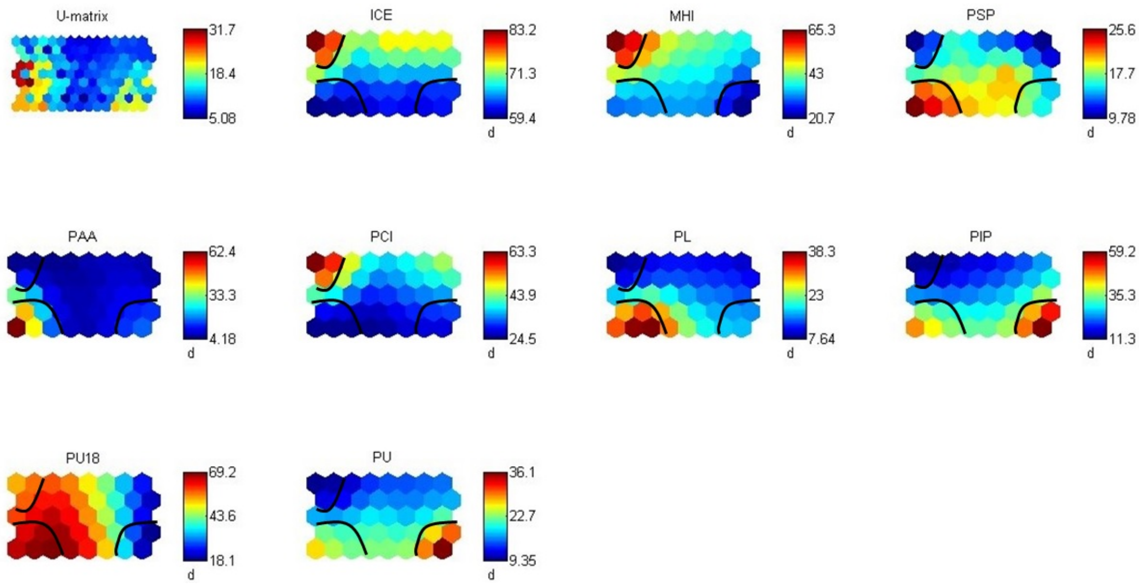

Figure 18 – Component planes and U-Matrix for the Portland SES SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the top left were Type A nodes, nodes in the bottom left were Type B nodes, nodes in the bottom right were Type C nodes, and nodes in the middle were Type D nodes. See Section 5.2.2 for further discussion.

5.2.2. How did the SOM help visualize these patterns?

The overall pattern for Node Distance was a general decrease from left to right as shown in the Node Distance, number of Tracts per Node, and Standard Deviation of the Tract Distance for this SOM (Figure 19). The highest values for Node Distance were found at the top and bottom left, and the lowest values were in the middle right. There were no null-value nodes. High tract density nodes were almost exclusively located around the edge of the SOM, with the three highest tract nodes in the corners. The Standard Deviation of the Tract Distance was again related to the number of tracts per node.

The primary axis of change for this SOM was from the bottom right to the top left. This primary axis encompassed the majority of the income, poverty and unemployment component planes, but did not include any of the other SES planes. A minor percentage of the income, poverty and unemployment planes were explained by the second axis, as well as both racial component planes and some of the variation in the single parent and percentage of children component planes.

Four Node Types were identified in this SOM. Type A nodes were in the top left. These nodes were the income maximum, as well as the minimum in nearly every other component

plane. Type B nodes were in the lower left. The four component planes that did not directly measure income - single parents, the two racial component planes, and the percentage of children - were at a local maximum in this Node Type. Type C nodes were the income and SES minimum in the lower right. Each of these three Node Types had a single high tract density node in the corner of the SOM, relatively high internal variation in that node, and each was a local or global Node Distance maximum. Type D nodes comprised the remainder of the SOM. While there was little variation across this Node Type, there were multiple high tract density nodes interspersed among the more common low tract density nodes. These were mapped in the following section, as well as the corner nodes from the first three Node Types.

5.2.3. How were the tracts grouped in SOM nodes distributed geographically the city?

In contrast to the widely scattered crime nodes, the SES nodes were very strongly geographically clustered (Figure 18). The low income Nodes 1 and 8 were adjacent, resulting in a single, large low SES cluster. Node 4, representing tracts where Hispanic families were common, also showed geographic clustering on the east side of the city. The high income tracts were a single large geographic area on the west side of the city. This area was also geographically separated from the city center by both a river and a set of hills at the right edge of this cluster of tracts.

Two of the nodes were less clustered. Node 37, representing average income homes, was scattered throughout the city. Node 22 was also somewhat scattered, though more commonly located near the high income tracts. The overall result was of geographic clustering of the SES nodes. Given then this geographic correlation, it was not surprising that the Moran's I test for spatial autocorrelation returned a Z-score of 7.74.

5.2.4. Summary of the Portland SES SOM

This SOM, like the Nashville SOM, showed strong geographic clustering of tracts from high and low SES nodes, resulting in a socially stratified landscape. This may have been due to the tract distribution in this SOM, as half of the tracts were assigned to a quarter of the nodes (Figure 19). The high tract density nodes that were found with Type D Nodes were interesting, as the adjacent nodes to those also had very similar component plane values. This result also suggested neighborhood stratification in Portland.

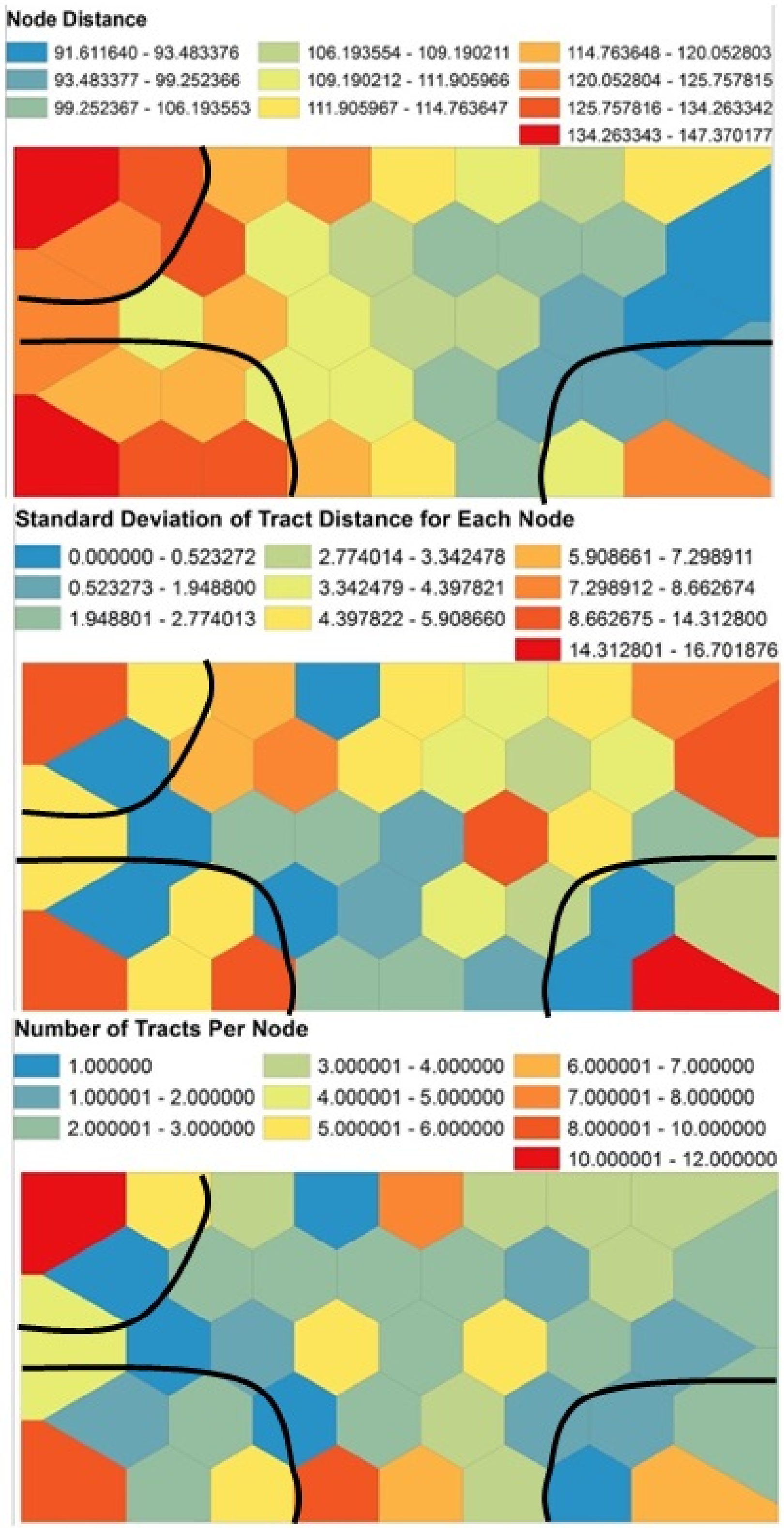

Figure 19 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Portland SES SOM. The drawn boundaries were the same as those in Figure 18 to denote Node Types A, B, C and D.

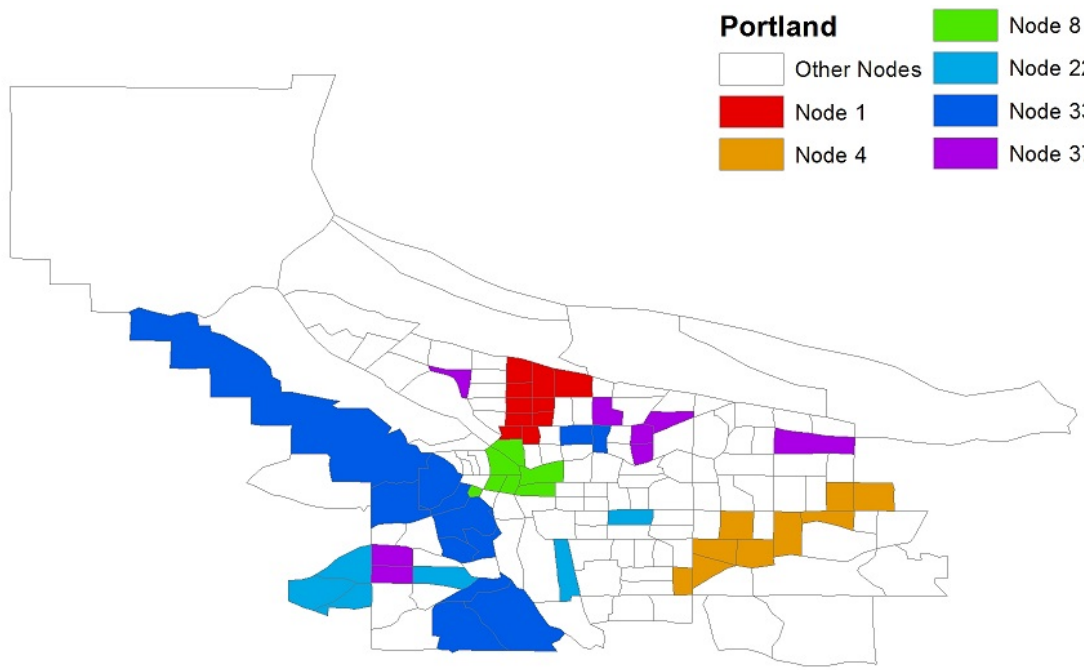

Figure 20 – Tracts associated with Node Types A, B and C in the Portland SES SOM. Node Type A was represented by Node 33, Node Type B was represented by Node 1, Node Type C was represented by Node 8, and Node Type D was represented by Nodes 4, 22 and 37.

5.3. Portland Crime and SES SOM

As with both the Crime and SES SOMs, the combined SOM had four identifiable node groups. Of these four node groups, two were similar low SES/high crime nodes, a single set of high income nodes that were not directly aligned with either of the two diagonal axes, and a large general set of nodes. As with other SOMs, the analysis was directed using the component planes (Figure 21), analysis of the Node Distance, the number of Tracts per Node and Standard Deviation of the Tract Distance (Figure 22), and geographically (Figure 23).

5.3.1. What patterns in the component planes could be identified within the SOMs?

Crime and low SES were both split to the lower left and right in this SOM (Figure 21). Six of the crime planes were highest in the lower right, with the exception of homicide and aggravated assault. The highest values of poverty were in the lower left, whereas unemployment was highest in the lower right. The lowest values of income were similarly split to the lower left (per capita) and lower right (household). The income maximum was in the top right; ICE, household, and per capita incomes all shared the same maximum value node, associated with minimums in poverty, unemployment, and crime. There was a pronounced overlap in African-American and Hispanic heritage patterns. Single parents were most common in the top left, and children were again common throughout, with low values only found in the far right.

5.3.2. How did the SOM help visualize these patterns?

In examining Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance for this SOM, the only large concentration of high tract density nodes was slightly offset in the top right (Figure 22). The majority of high tract density nodes were scattered around the edge of the SOM and separated from the others by at least one under-filled node. Only the bottom left node was not within two nodes of another high tract density node. Nodes with less than two tracts were also scattered throughout the SOM, largely in between high tract nodes. The Node Distance was tri-polar: maximums were found offset from the top right, the bottom left, and the bottom right. There was little decrease between the bottom row maximums, but values otherwise decline outside the maximums. The minimum was offset in the top left, and the single null-value node was in the bottom right. Standard Deviation of the Tract Distance was again largely correlated to the number of tracts, though with several exceptions. Several of the high tract density nodes in the middle of the SOM and one on the bottom row had very low values of standard deviation.

Globally, it was unclear which of the two primary axes the axis of largest change was. The offset income maximum from the primary axes was a further complication. As low SES and high crime was split to the lower left and right, there was no evidence to suggest one axes explained more of the change across the SOM than the other.

Four sets of nodes were identified at the neighborhood level. The bottom left nodes were labeled Type A nodes. This was one of two low SES/high crime Node Types. Both racial component plane maximums were associated with this Node Type. Type B nodes were located on the bottom right. Income was equally low in both of these Node Types, along with high poverty and unemployment. Crime was high in both Node Types, but there was no obvious split based on violent and non-violent crime. The high tract density nodes in both types also showed very high internal variation. The only significant difference between these two Node Types could be attributed to the racial component planes and the percentage of people under the age of 18, all of which were high in Node Type A and low in Type B.

Type C nodes were the income maximum in the offset top right. These nodes were also associated with very low crime, but not the crime minimum in most crime component planes. Type D nodes comprised the remainder of the SOM. Component plane values were less uniform across this Node Type, particularly in the top left of the SOM. There was also multiple high tract density nodes scattered across this Node Type.

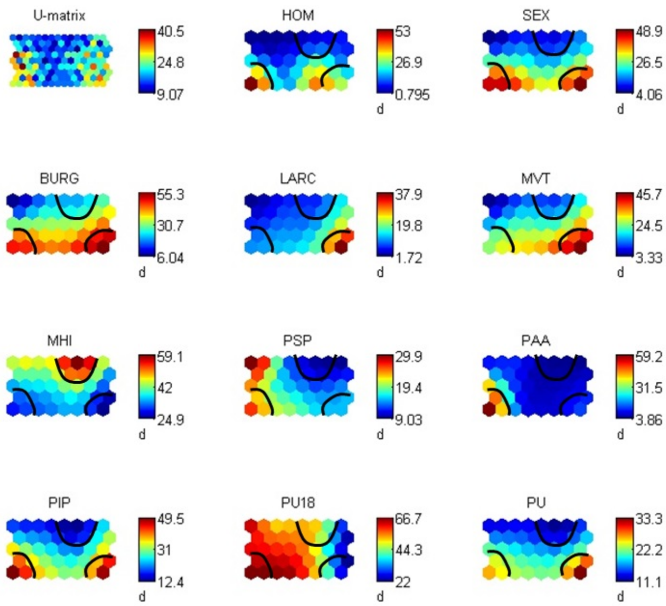

Figure 21 – Component planes and U-Matrix for the Portland Crime-SES SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the bottom left were Type A nodes, nodes in the bottom right were Type B nodes, nodes in the top right were Type C nodes, and nodes in the middle were Type D nodes. See Section 5.3.2 for further discussion.

5.3.3. How were the tracts grouped in SOM nodes distributed geographically the city?

Nodes from each Node Type were mapped: one each from Types A, B and C, and three from Type D. Six nodes demonstrated the expected results from the combination of the crime and SES SOMs: general clustering, but the clustering was less pronounced than in the SES-only SOM (Figure 23). This was best seen in Nodes 1, 4, 8 and 38, each of which had tracts similar to nodes identified previously (Figure 20). Five of the six nodes showed geographic clustering, but each of these had several tracts separate from the main cluster. Node 20 identified a collection of tracts separate from those noted in either previous analysis, but it was well clustered. Node 33 was quite scattered at the edge of the map. As with the previous maps of tracts, the Moran's I test returned a value indicative of autocorrelation: a Z-score of 6.11.

5.3.4. Summary of the Portland Crime-SES SOM

Nodes Types A and B were both high crime/low SES nodes, but with distinct patterns of crime and SES. While Node Type A was associated with high percentages of both African-Americans and Hispanics, there was also a high tract density node (Node 4) of high percentages of Hispanics with much lower crime values. Node Type B was associated with high crime and low SES, but not with either racial component plane. The crime minimum (Node 33) was also disconnected from the income maximum at Node Type D.

5.4. Case Study Two Summary

These results were markedly different from those of the Nashville case study. Crime was broadly scattered in the geographic analysis, despite not appearing as such in the analysis of the SOM, while SES showed very strong geographic clustering. When combined, the SOM was very scattered, but the tracts showed strong geographic clustering. While it was difficult to draw conclusions from the Crime SOM, the combined SOM does show that there was an association of patterns between those of crime and measures of SES. However, it appears that there was only a partial association between crime and SES, as demonstrated by the split high crime/low SES patterns (Figure 21).

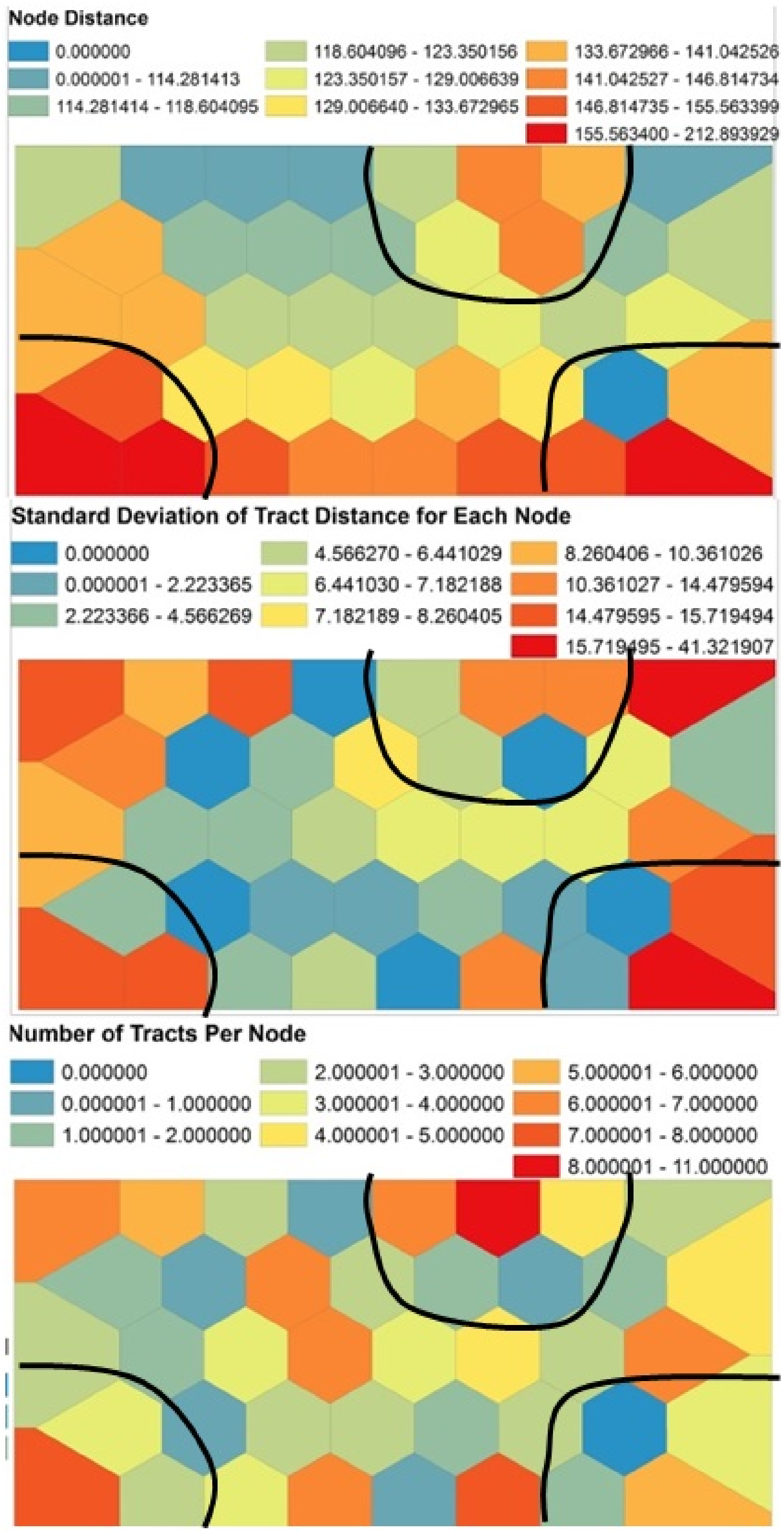

Figure 22 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Portland Crime-SES SOM. The drawn boundaries were the same as those in Figure 21 to denote Node Types A, B, C, D.

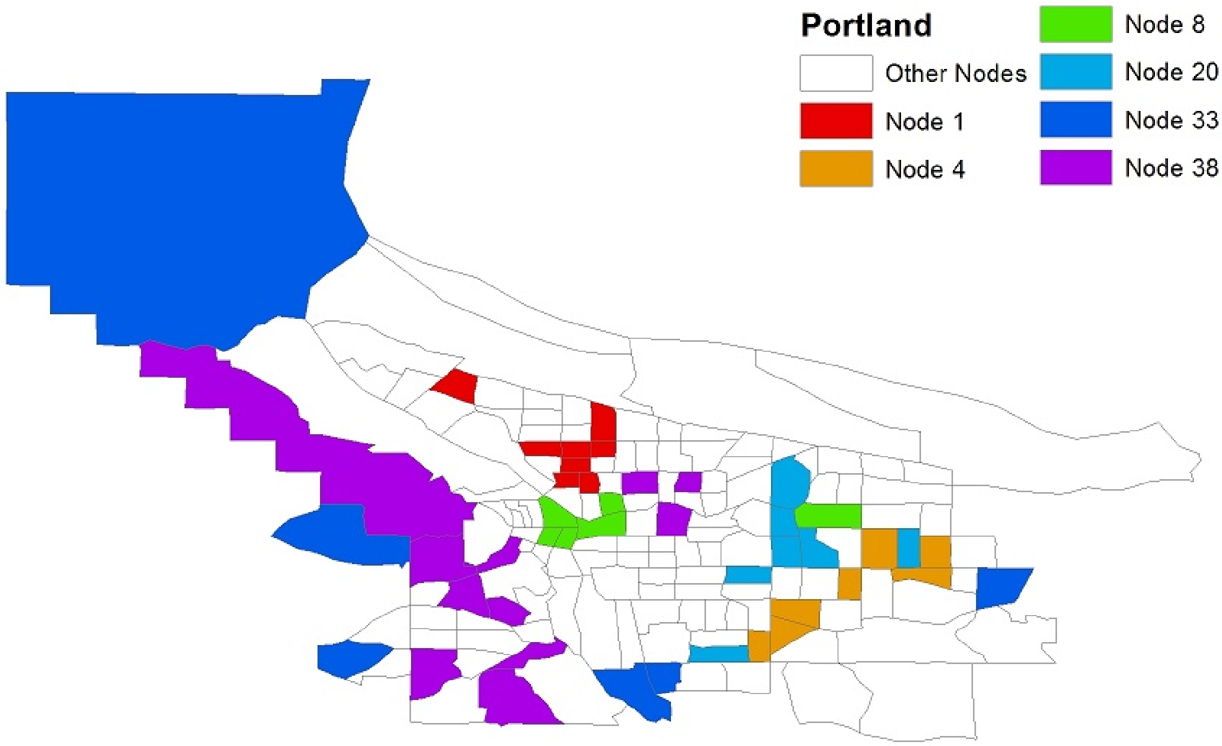

Figure 23 – Tracts associated with Node Types A, B, C, D and E in the Portland Crime-SES SOM. Node Type A was represented by Node 1, Node Type B was represented by Node 8, Node Type C was represented by Node 38, and Node Type D was represented by Nodes 4, 20, and 33.

6. Case Study III: Tucson, AZ

Tucson has few defining geographic features within the city, as the few rivers in the city are small and normally dry. The edges of the city are defined by mountains to the west, northeast and east. The city does not have a large downtown or clearly separated districts, and more resembles a suburb writ large. (Figure 24) Two main residential areas exist: along a north-south line paralleling the Interstate and a long stretch of residential area east of the city center. Demographically, the city has the highest percentage of persons of Hispanic heritage, with many immigrants arriving in the city as they arrive from Central America. Nodes in the Tucson Case Study SOMs were numbered as in Figure 25.

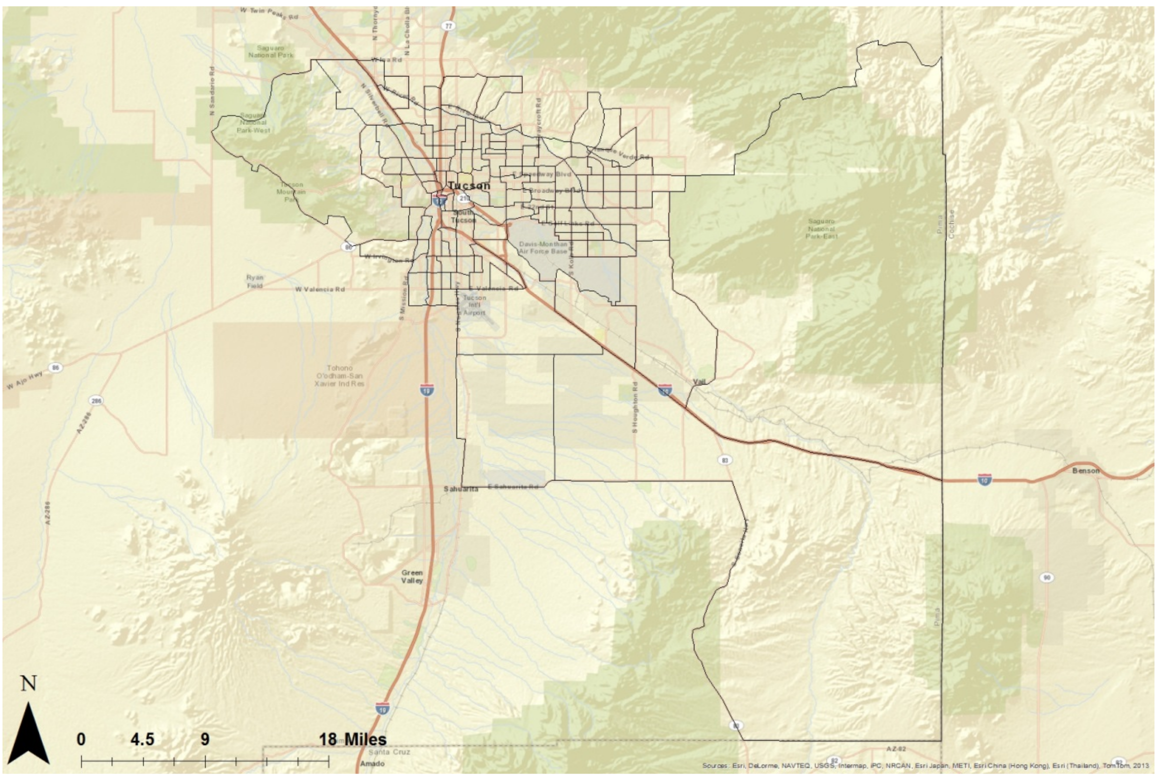

Figure 24 – Tucson, AZ study area, with tracts

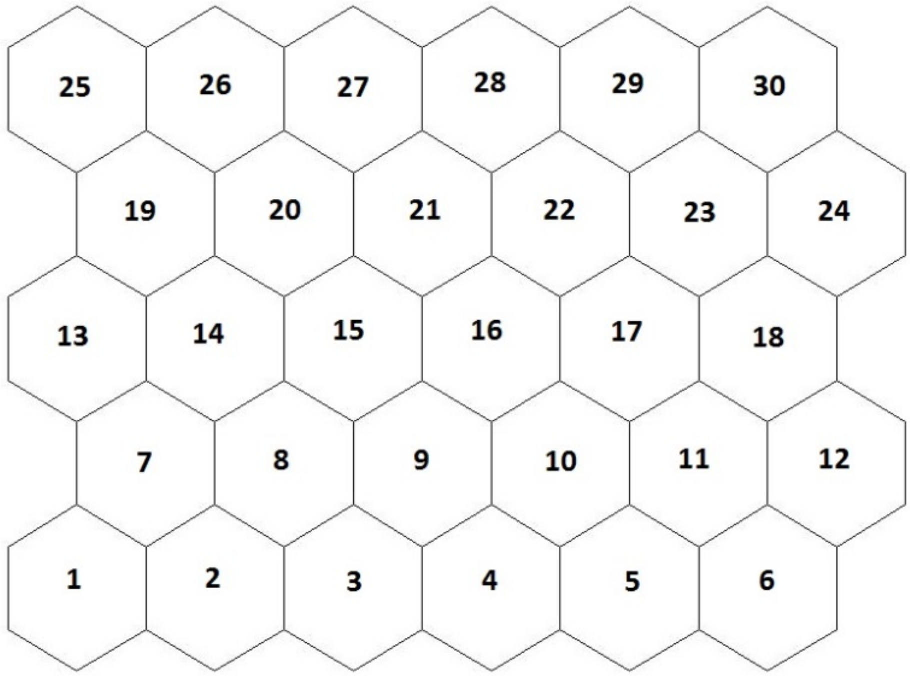

Figure 25 – SOM Node Numbers for Tucson, AZ

6.1. Tucson Crime SOM

There were four sets of node groups in the Tucson Crime SOM. The Type A nodes in the lower right were the primary crime pattern, with maximum values of crime in six of the seven categories. These patterns were analyzed using the component planes (Figure 26), analysis of the Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance (Figure 27), and geographically (Figure 28).

6.1.1. What patterns in the component planes could be identified within the SOMs?

Analysis of the component planes showed that the lower right was the clear crime maximum for five of the seven component planes, with larceny in a dual maximum at Nodes 1 and 7, and juvenile crime highest at Node 25 (Figure 26). While crime was uniformly high in the lower right of the SOM, the node of the maximum value of homicide was different, peaking at Node 12 instead of Node 6. The difference in values was minimal, as crime values were very high in both nodes in all planes except for juvenile crime. There was no clear pattern of moderate crime.

6.1.2. How did the SOM help visualize these patterns?

Tract distribution showed high tract density nodes in the top middle, bottom left, and bottom right of the SOM, each surrounded by under-filled nodes in analysis of Node Distance, the Standard Deviation of the Tract Distance, and the Number of Tracts per Node (Figure 27). The crime minimum could be identified as Node 28 from the Node Distance, and that node had 23 tracts assigned to it. There was also a single null-tract node on the left side. Node Distance

indicated a slow change of crime values across the SOM. The lower right could be identified as the crime maximum, with two secondary maximums in the upper and lower left. The Standard Deviation of the Tract Distance was again roughly correlated to the number of tracts per node.

The primary axis of change in the SOM was identified as the lower right-top left diagonal, due to the organization of the component planes and Node Distance. While crime in both diagonals decreased from the lower corner to the upper corner, the global minimum was identified in the top middle of the SOM.

Four node groups were identified in this SOM and mapped in Figures 26 and 27. Type A nodes were in the lower right, and were the clear crime maximum. Type B nodes were in the lower left and were a secondary crime maximum associated with largely non-violent crime. Type C nodes in the top center were the crime minimum. In between these three Node Types were the Type D nodes. As was seen from the tract distribution, these nodes were almost entirely underpopulated. Crime values were generally low in these nodes, with the exception of the top left node where Juvenile Crime was high.

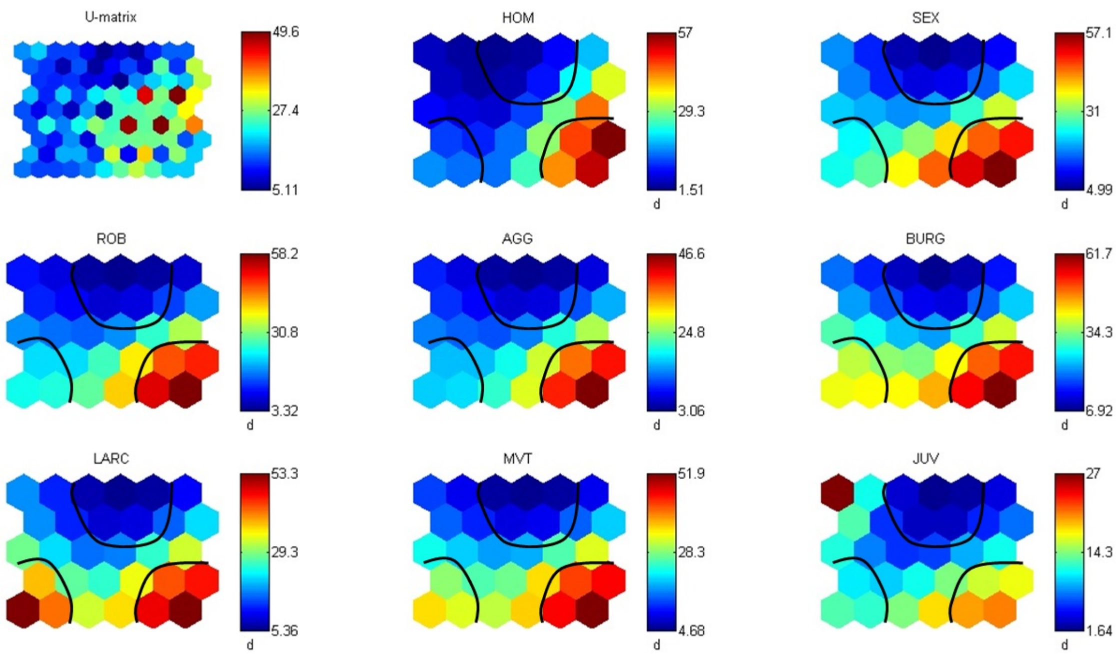

Figure 26 – Component planes and U-Matrix for the Tucson Crime SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the bottom right were Type A nodes, nodes in the bottom left were Type B nodes, nodes in the top middle were Type C nodes, and nodes in the middle were Type D nodes. See Section 6.1.2 for further discussion.

6.1.3. How were the tracts grouped in SOM nodes distributed geographically the city?

Four nodes (Nodes 1, 6, 25 and 28) were mapped based on the previous analysis (Figure 28). Node 1 was associated with Type B nodes, Node 6 with Type A nodes, Node 25 with Type D nodes, and Node 28 with Type C nodes. Each node was geographically scattered. Nodes 1 and 6 were generally located around the city center. Tracts assigned to Node 25 were found

near tracts in Nodes 1 and 6, but further outward from the city center. Node 28 was scattered at the edge of the map. While no geographic clustering of the tracts was present, the overall pattern was of high crime near the city center, decreasing outward. In testing for spatial autocorrelation, Moran's I returned a Z-score of 9.48, with a *p*-value of 0, despite the lack of clustering (Figure 28).

6.1.4. Summary of the Tucson Crime SOM

The analysis of SOM showed one high crime Node Type and one secondary crime Node Type associated with non-violent crime. The Type A nodes were the main crime nodes as it contained the maximum value of all crimes except Juvenile Crime (Figure 26). The lower left nodes were based on non-violent crime and where larceny was the most common in this SOM. To the top left was a single node of three tracts where only juvenile crime was high, and which was associated with an increase in Node Distance (Figure 27). There was also a very high standard deviation of the tract distance in this node. The number of low crime tracts was also significantly higher than the number of high crime tracts.

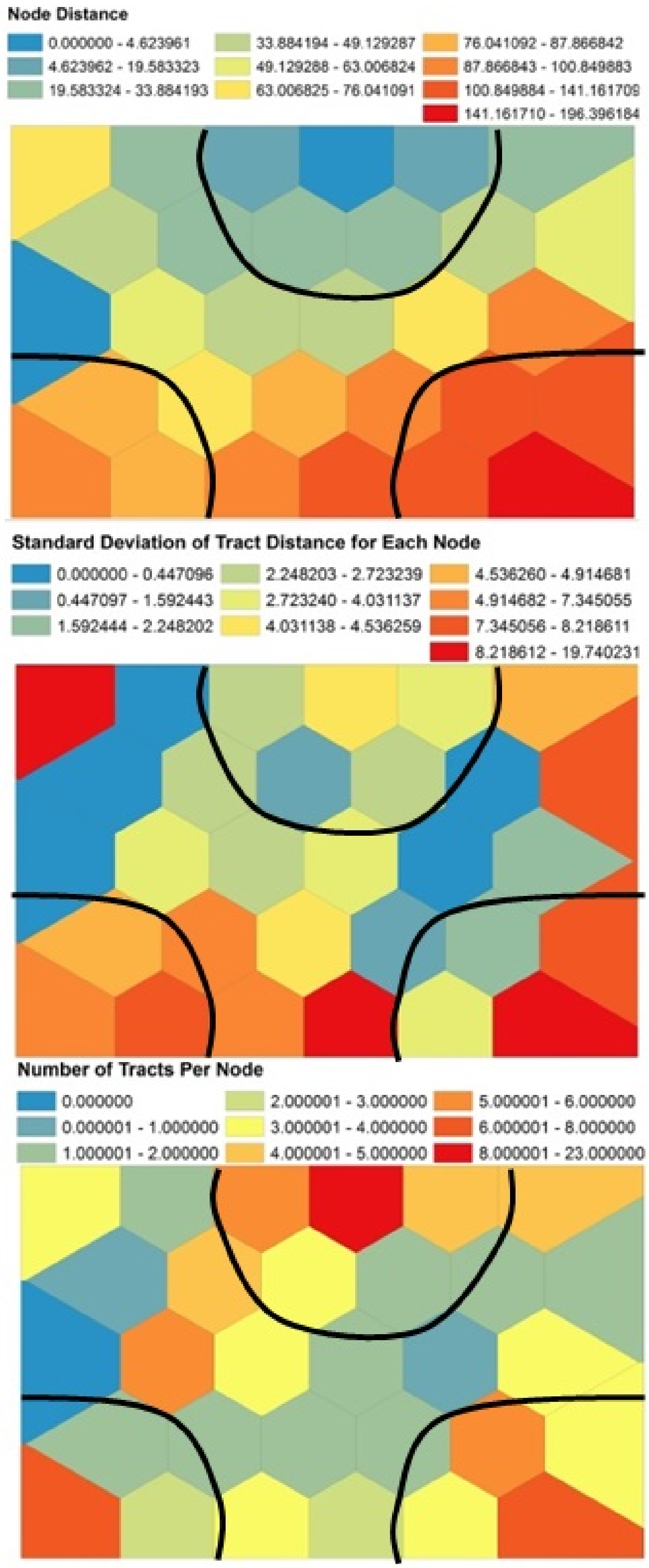

Figure 27 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Tucson Crime SOM. The drawn boundaries were the same as those in Figure 26 to denote Node Types A, B, C and D.

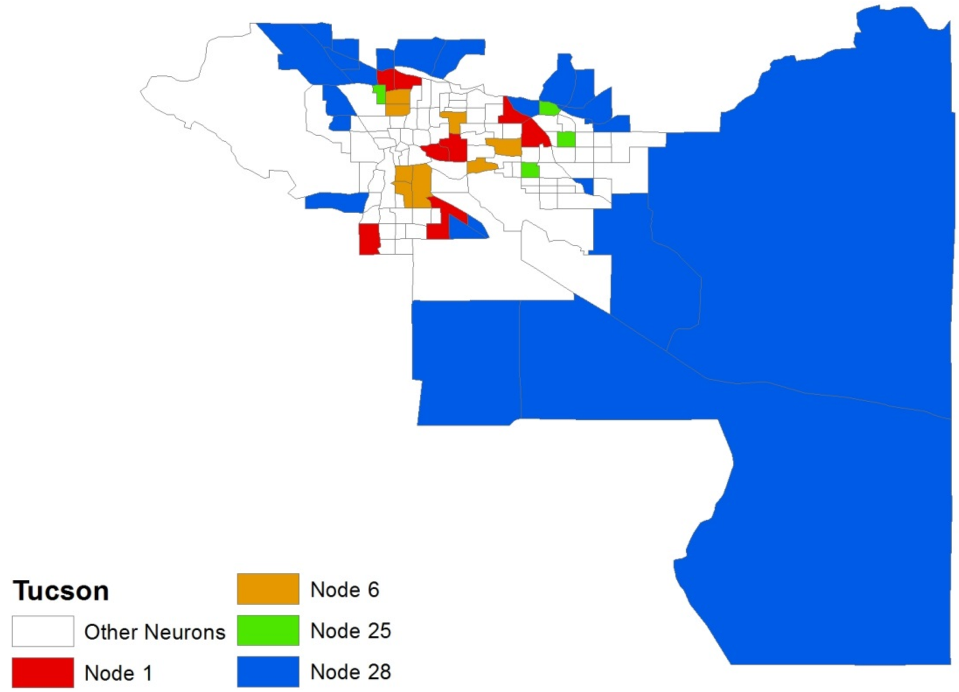

Figure 28 – Tracts associated with Node Types A, B, C, and D in the Tucson Crime SOM. Node Type A was represented by Node 6, Node Type B was represented by Node 1, Node Type C was represented by Node 28, and Node Type D was represented by Node 25.

6.2. Tucson SES SOM

Five sets of node groups were found from the SOM: one in each corner and a fifth in the middle. The upper two sets of nodes were both related to high levels of SES and the lower two sets of nodes were related to low values of SES. These patterns were analyzed using the component planes (Figure 29), analysis of the Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance (Figure 30), and geographically (Figure 31).

6.2.1. What patterns in the component planes could be identified within the SOMs?

Analysis of the component planes showed that SES and income were again well aligned in this SOM (Figure 29). The highest values of per capita income aligned with the lowest values of poverty and unemployment, and vice versa. High income was split into two maximums, likely due to differences in the percentage of children in the households, which may have indicated single person households versus family households. The top left node was likely represented by single person households: per capita income was roughly equal in the top left and right, but median household income was only average in the top left. Low SES was also split, with two maximums for poverty and unemployment. This appears related to differences in the racial makeup, the percentage of single parents, and the percentage of children in the nodes contributing to the difference. The component plane for African-American heritage was unique, centered in the lower middle of the SOM. African-American heritage was not closely

associated with high poverty and unemployment, unlike previous models that showed a close association.

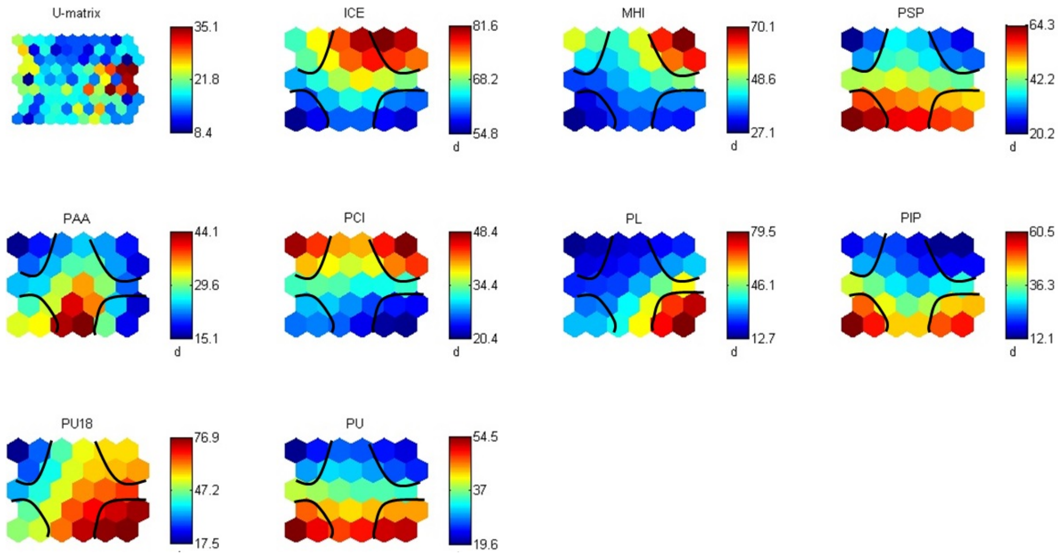

Figure 29 – Component planes and U-Matrix for the Tucson SES SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the top left were Type A nodes, nodes in the top right were Type B nodes, nodes in the bottom left were Type C nodes, nodes on the bottom right were Type D nodes, and nodes in the middle were Type E nodes. See Section 6.2.2 for further discussion.

6.2.2. How did the SOM help visualize these patterns?

The Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance analysis showed that Node Distance was highest in the lower right nodes and decreased to a minimum in the top left nodes, with two secondary maximums in the top right and lower left (Figure 30). There were no null tract nodes. Tracts were distributed with greatest frequency at the edges of the SOM, and most nodes in the center were under-populated. The Standard Deviation of the Tract Distance was again reflective of the tract distribution, where high tract density nodes had the largest standard deviation.

There was no clear primary axis in this SOM. Examining only Node Distance, the lower right-top left axis was clearly the primary axis, but the component planes were equally split; as many component planes showed a large change from the bottom left to the top right as showed a large change from the bottom right to the top left. This was reflective of the split in both high and low SES, as seen in Figure 29.

The split in high and low SES created five sets of node groups in this SOM. Type A and B nodes were identified in the top left and top right respectively. Both of these Node Types were sets of high SES nodes, with the split likely due to single-person households in Type A nodes versus family households in Type B nodes. Per Capita Income was equal in both Type A and B nodes, but Household Income was much higher in the Type B nodes, as well as a far higher percentage of children. There were also many more tracts associated with Type B nodes.

Type C and D nodes were the low SES Node Types, split to the lower left and right respectively. Income, poverty and unemployment were roughly equal in these two node types, but single parents were slightly more common in Type C nodes, and children and persons of Hispanic heritage were more common in Type D nodes. Type E nodes comprised the middle set of nodes. There was significant change in component plane values for income and race in Type E nodes, but there nodes were on the average underpopulated.

6.2.3. How were the tracts grouped in SOM nodes distributed geographically the city?

Five nodes were mapped from this SOM, one from each Node Type discussed above. Nodes 1 and 6 were associated with Node Types C and D, Nodes 25 and 30 with Node Types A and B, and Node 15 with Node Type E. (Figure 31). Geographic clustering was evident in Nodes 1, 6, and 16. Node 6 showed the strongest clustering, with most tracts concentrated in a single N-S line that continued in Node 1. Node 16 was associated with an area of suburban sprawl to the east of downtown. The two high income Nodes 25 and 30 were primarily to the east and north of the city center, with some clustering in the east. The Moran's test for spatial autocorrelation again returned a high Z-score of 7.428, confirming spatial autocorrelation.

6.2.4. Summary of the Tucson SES SOM

The five SES node groups here were split into two low patterns, two high patterns, and one of average income (Figure 31). The two low SES node groups were split due to racial component planes, one associated with Hispanic heritage and the other not strongly associated with any racial component. The high SES node groups were split by household income. The fifth node group represents neither high nor low values of any of the component planes. These results indicate that low SES was not always connected to racial heritage.

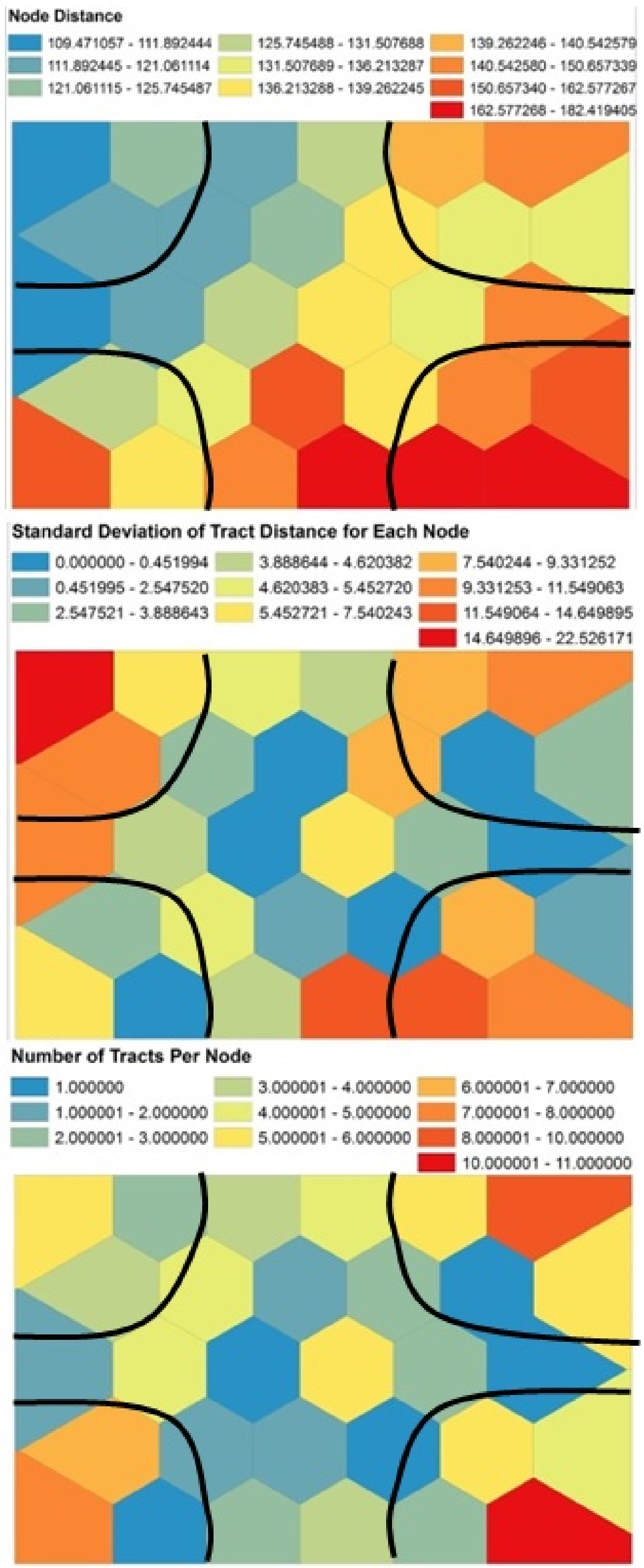

Figure 30 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Tucson SES SOM. The drawn boundaries were the same as those in Figure 29 to denote Node Types A, B, C and D.

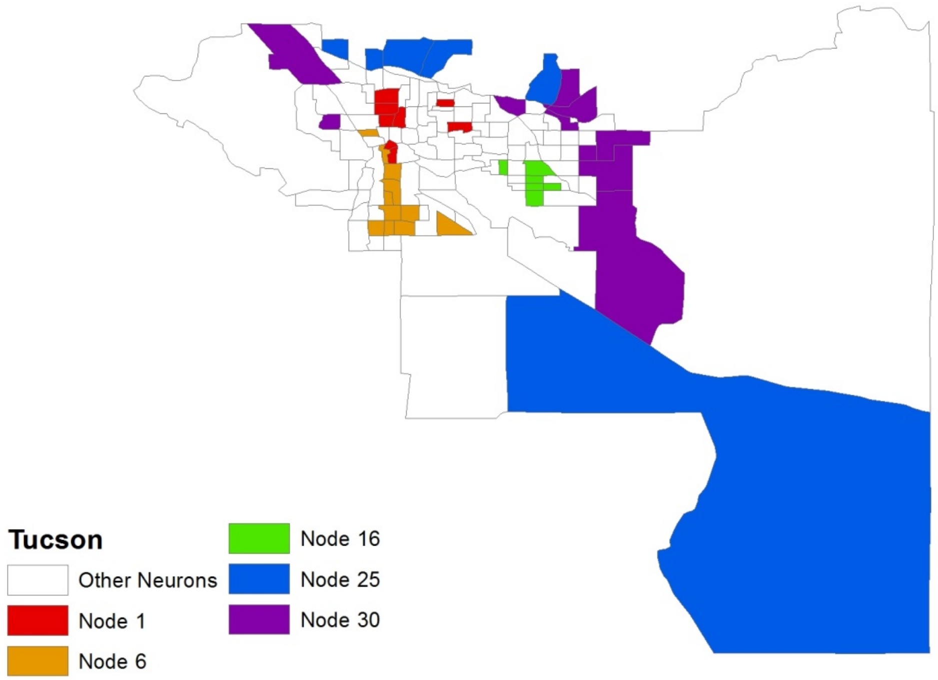

Figure 31 – Tracts associated with Node Types A, B, C, D, and E in the Tucson SES SOM. Node Type A was represented by Node 25, Node Type B was represented by Node 30, Node Type C was represented by Node 1, Node Type D was represented by Nodes 6, and Node Type E was represented by Node 16.

6.3. Tucson Crime and SES SOM

The lack of overlap in local maximums in the crime and SES component planes made analysis difficult, but five node groups were discerned. The lower two rows were split evenly into three node groups, the top right was identified as a fourth group, and the remainder of the SOM was identified as a fifth group. These patterns were analyzed using the component planes (Figure 32), analysis of the Node Distance, the number of Tracts per Node, and Standard Deviation of the Tract Distance (Figure 33), and geographically (Figure 34).

6.3.1. What patterns in the component planes could be identified within the SOMs?

Neither the crime nor the SES component planes were well organized in this SOM (Figure 32). The crime maximum was scattered across the bottom row for the 8 crime component planes, though they could be roughly grouped into violent and non-violent crime patterns. High values of violent crimes were shared between Nodes 2, 3, and 4, and high values of non-violent crime were shared between Nodes 4, 5, and 6. Minimum crime values were also shared between Nodes 25 and 30. As high crime values were previously aligned in a single node (Figure 26), this suggests that the crime patterns have been influenced by the inclusion of SES values.

Similarly, patterns of SES were not strongly organized. There were multiple nodes in the lower left associated with the highest values of poverty and unemployment, and the income patterns were aligned into a single node in the top right. The result of this model shows that while high crime was again found with low SES, the patterns did not overlap.

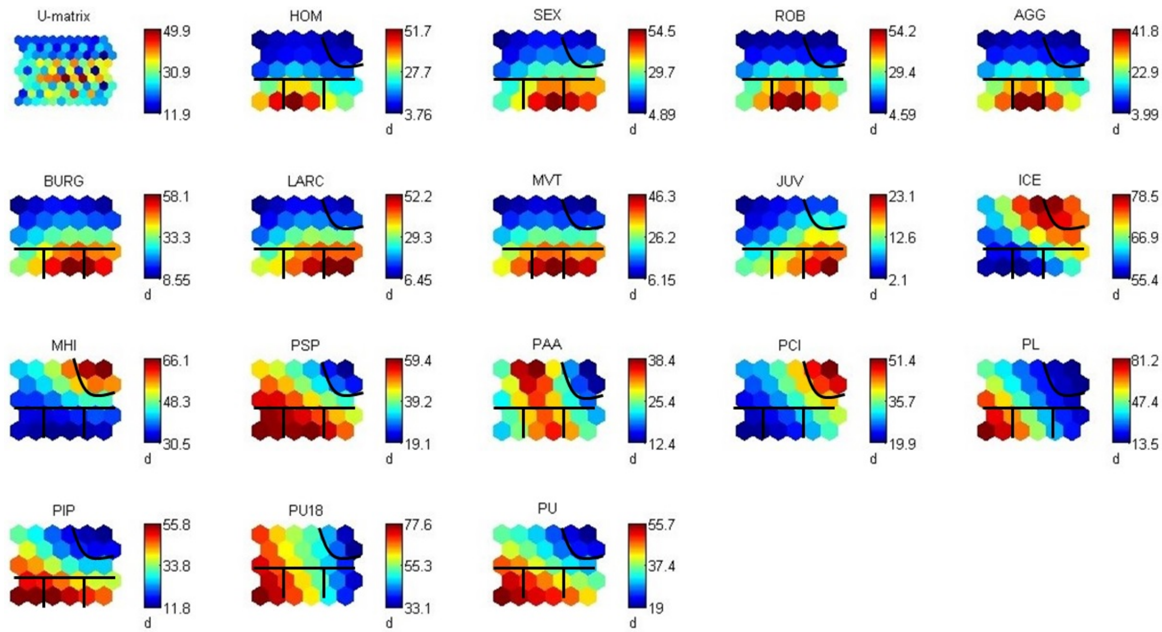

Figure 32 – Component planes and U-Matrix for the Tucson Crime-SES SOM. Refer to Table 1 for the reference names for the component planes. Nodes bracketed by the curve in the bottom left were Type A nodes, nodes in the bottom middle were Type B nodes, nodes in the bottom right were Type C nodes, nodes in the top right were Type D nodes, and nodes in the middle were Type E nodes. See Section 6.3.2 for further discussion.

6.3.2. How did the SOM help visualize these patterns?

In analysis of the Node Distance, Standard Deviation of Tract Distance for Each Node, and the Number of Tracts per Node, Node Distance was maximized in the bottom center at Nodes 3 and 4 and gradually decreased outwards from this maximum (Figure 33). High Node Distance value tracts were concentrated around the edge of the SOM and the majority of interior nodes under-filled. There were no null tract nodes. The Standard Deviation of the Tract Distance reflected the tract distribution and the Node Distance: standard deviation values were high where Node Distance was high or where there were a high number of tracts assigned to that node.

Global analysis was difficult due to the offset global maximum in the bottom middle of the SOM (Figure 33). The SES planes were generally identified with local axes from the lower left to the top right, and crime planes were identified with axes from the bottom right to the top left. Neither axis was clearly identifiable as the primary global axis.

Node Groups were difficult to identify, due to the lack of overlap in the crime and SES component planes. The bottom two rows were identified as Node Types A, B and C. Type A nodes were the least criminogenic of the three, and the nodes where SES was lowest. There

were also very few tracts associated with these nodes. Type B nodes tended to have slightly more violent than non-violent crime, and Type C nodes had slightly more non-violent crime than violent crime. SES was also slightly lower in Type B nodes. These classifications were not as clear as others have been in previous analysis, and there was little difference between the three in Node Distance.

Type D nodes in the top right were the income maximum of this SOM. Crime values were also at local minimums or very low values in these nodes. Type E nodes comprised the remaining nodes, which generally had low crime rates and moderate income values. One node in this Node Type, Node 13, was notable as one of the few that could be identified as a high tract density node with low crime and low SES.

6.3.3. How were the tracts grouped in SOM nodes distributed geographically the city?

Six nodes were mapped: 3, 6, 13, 17, 28, and 29 (Figure 34). Of these, Node 3 was associated with Type B nodes, Node 6 with Type C nodes, Nodes 13 and 17 with Type E nodes, and Nodes 28 and 29 with Type D nodes. Nodes 3 and 6 were associated with the violent and non-violent crime patterns, Node 13 with low SES, Node 17 with average SES, and Nodes 28 and 29 with moderate and high income. Nodes 3 and 29 showed strong geographic clustering, Node 13 showed some clustering, and the other three were scattered. This was reflective of the SOM and the maps of nodes in the Tucson Crime and SES SOMs. Node 3 was the only node with both high crime and low SES and was associated with tracts previously identified with those factors south of the city. The scattered nature of crime values was seen with Node 6. Nodes 17, 28, and 29 represented progressively increasing SES, which was seen in tracts in Node 17 due to being closer to the city center than those of Nodes 28 and 29 to the east of downtown. This was similar to the SES pattern (Figure 31). The test for spatial autocorrelation returned a Z-score of 8.3398, strongly suggesting spatial autocorrelation that could be seen in several nodes.

6.3.4. Summary of the Tucson Crime-SES SOM

Unlike previous SOMs, this SOM did not have discrete Node Types, as many of the component planes were slightly offset from each other. The crime component planes did not align well into one or two nodes as was observed in the crime-only Tucson SOM. Violent and non-violent crime was loosely organized around Nodes 3 and 6, and many of the SES component planes had similar values associated with these two patterns. High and moderate income nodes were identifiable around nodes 28 and 30, with similar values in all other component planes. High crime patterns were found with low SES patterns, but low SES was not always found with high crime.

6.4. Case Study Three Summary

The Tucson results differed from both previous case studies. Clear patterns of SES in the city could be established, but crime patterns could not be clearly established geographically or in the combined SOM. The result was that there was an incomplete association between SES and crime. Low SES and high violent crime were strongly associated with one another, but there was no clear association for non-violent crime.

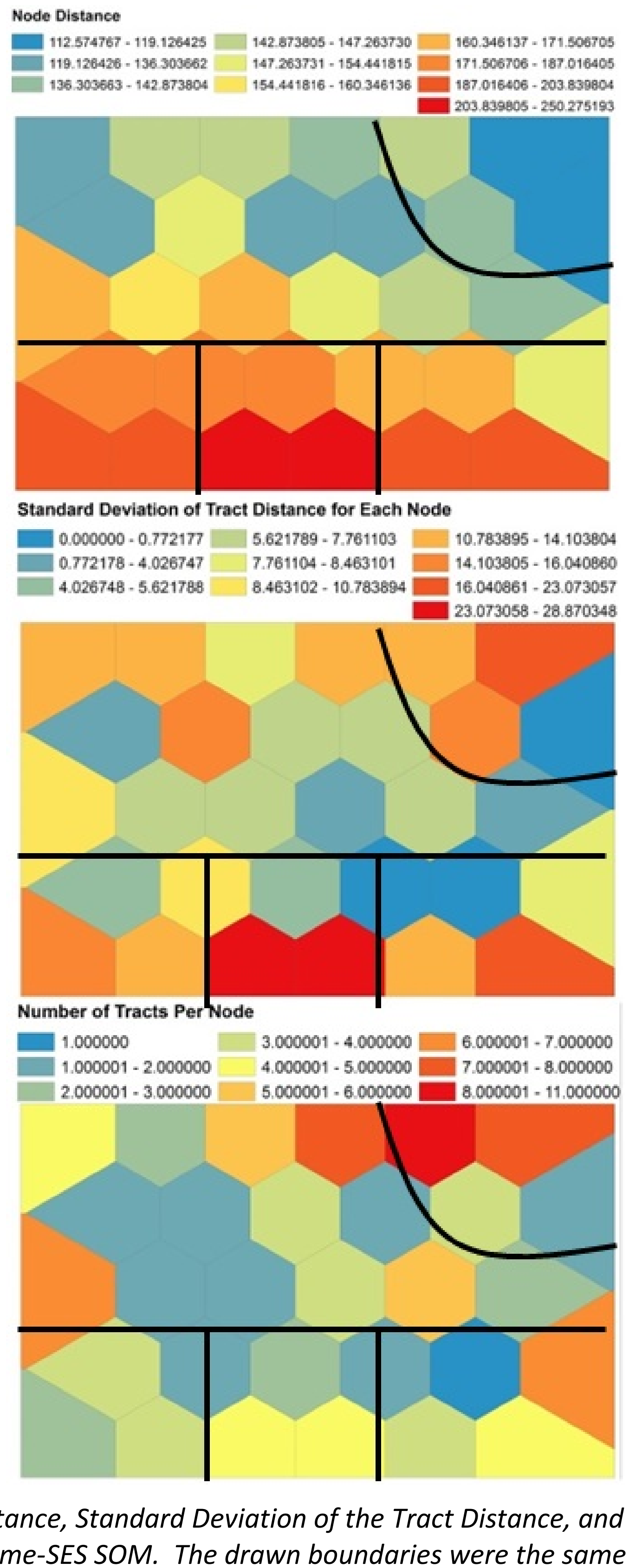

Figure 33 – Node Distance, Standard Deviation of the Tract Distance, and the Number of Tracts per Node, Tucson Crime-SES SOM. The drawn boundaries were the same as those in Figure 32 to denote Node Types A, B, C, D.

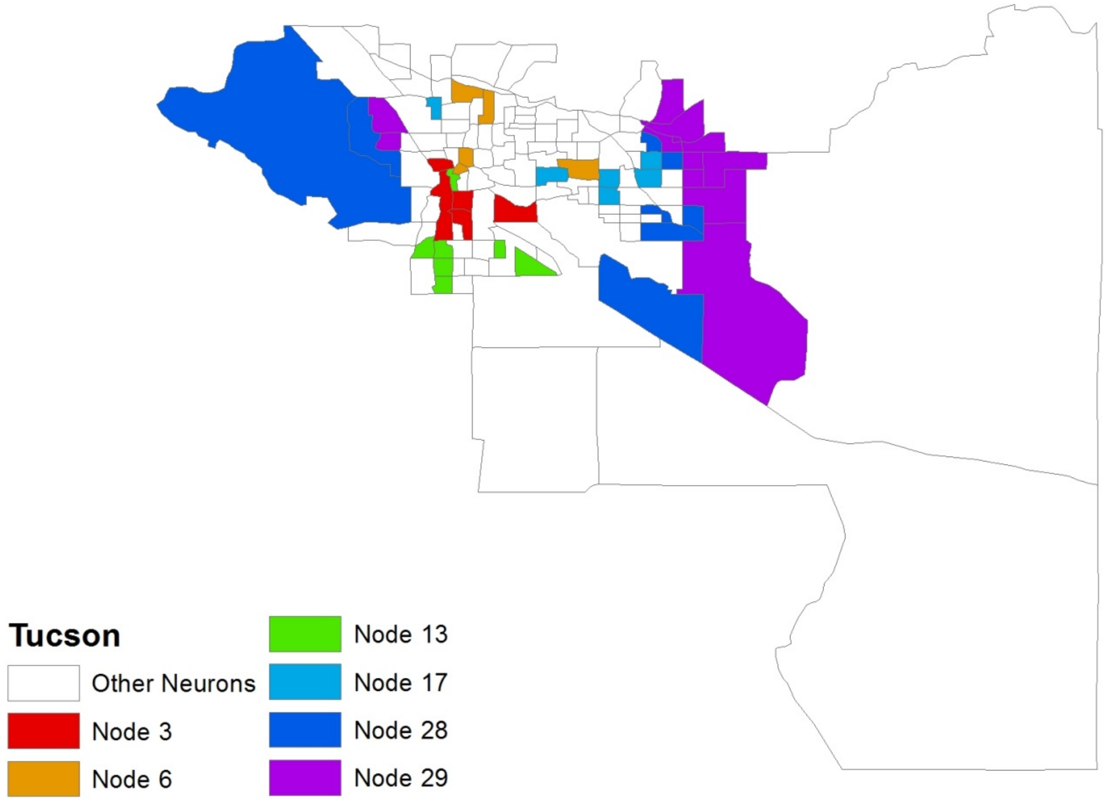

Figure 34 – Tracts associated with Node Types B, C, D and E in the Tucson Crime-SES SOM. Node Type B was represented by Node 3, Node Type C was represented by Node 6, Node Type D was represented by Nodes 13 and 17, and Node Type E was represented by Nodes 28 and 29.

7. Discussion

The results of these case studies showed a partial association of the spatial patterns of crime and SES. The spatial patterns of high crime were strongly associated with those of low income, high poverty and high unemployment. The converse of this was also true; the spatial patterns of low crime were strongly associated with those of high income, low poverty and low unemployment. While the spatial patterns of high crime were associated with low income, it could not be said that the spatial patterns of low income were strongly associated with high crime. The spatial patterns of other measures of SES were not consistently associated with any measure of crime or income across the three case studies, nor where they consistently associated with each other.

Each case study presented evidence of an association between crime and SES, but the evidence was not the same in each case study. The Nashville case study showed an association of the patterns of crime and SES in both the SOM and the geographic analysis of SOM nodes. In Portland, crime and SES were associated, but the highest crime was not associated with the lowest SES. Crime and SES had a weak association of patterns in the Tucson SOM but a much stronger association in the geographic mapping of nodes. The lowest levels of crime were associated with high SES and income in each case study. This section will compare the three case studies.

7.1 Comparison of Crime Patterns

Each of the three case studies found a limited number of crime patterns. In the Nashville case study, the spatial patterns of high violent and non-violent crime were separate. Two separate spatial high crime patterns were identified in the Portland case study, but neither pattern included solely violent or non-violent crime. The Tucson case study showed a single primary crime pattern, with two minor separate high crime patterns. In all three case studies the spatial patterns of low crime were strongly associated. While there was clustering of tracts at the crime maximums and minimums in each case study, there was little evidence of clustering for any level of crime in between the extremes despite a large number of tracts with moderate crime levels in all three. Geographically, the nodes with high crime and low crime were clustered only in the Nashville case study. Each study did suggest that high crime tracts were more common closer to the city center and low crime tracts were more common further away from the city center.

7.2 Comparison of SES Patterns

The analysis of SES patterns in each city found some consistent results across the three case studies. The highest percentages of people in poverty or unemployed were always associated with the lowest levels of per capita and household income. The reverse of this was also found in each case study. Both of these outcomes were expected and were the only consistent results in the analysis of the SES patterns. The Portland and Tucson case studies found two distinct patterns of low SES, but the Nashville case study only found one. This may be attributed to the racial component planes, which were inconsistently associated with low

SES patterns. The percentage of all people under the age of 18 and the percentage of mothers who were single mothers were also inconsistently associated with low SES patterns. When the SES patterns were mapped in each city, there were consistent patterns of geographic concentration of the tracts. Low SES tracts were near the city centers, and high income tracts were distant from the city centers.

7.3 Comparison of Crime and SES Patterns

The combined crime and SES analyses were analyzed and compared against the spatial patterns of crime and SES that had previously been analyzed separately. In each study, there was an incomplete association of high crime and low SES. The Nashville and Tucson Crime-SES SOMs showed the patterns of high crime separate from those of the lowest values of SES, though there were moderate crime values where SES was lowest. The Portland combined SOM showed two similar patterns of high crime and low SES, one of which was also associated with racial component planes. The geographic analysis of tracts associated with those patterns provided stronger evidence that high crime patterns and low SES patterns were associated in Nashville and Tucson. In both cities, tracts that had been previously identified as being associated with high crime and low SES patterns were found to be associated with tracts displaying low SES and moderate or higher crime levels in their respective SOMs. Therefore, it can be said that the patterns of high crime and low SES were found to be associated in each case study.

In each case study, the patterns of crime in the crime-only SOMs were different from those of the combined Crime-SES SOM. In the Nashville case study, the previous split in violent and non-violent crime patterns was not present in the combined SOM, but the combined SOM retained two separate crime patterns. The crime patterns in Portland were nearly identical, but a second homicide pattern did appear in the combined SOM. The Tucson crime patterns showed the largest change from a single crime pattern to a scattered set of crime patterns, with an approximate violent/non-violent crime split.

The SES patterns did not show significant changes between their separate SOMs and their combined SOM. The greatest changes occurred in the Portland SOM, where the locations of the high and low SES Node Types on the SOM were changed, but there was no change in how the components were associated with one another. While there were no changes in the spatial association of the SES component, there was a change in how crime patterns were associated with each other from the crime only SOM to the combined SOM. This outcome suggested that the patterns of crime were changed by the inclusion of the SES component planes. This may in part be due to the fact that there were more SES component planes than crime component planes, thus the combined SES patterns have a greater weight in the SOM result.

In each of the three case studies, the lowest levels of crime were associated with the highest levels of income. Where crime was highest, income was consistently low, but there was not high crime in all nodes where income was low. None of the other SES planes showed a consistent association of patterns with crime across the case studies. The two racial component planes showed an association with crime patterns in at least one case study, but

the pattern was not consistent. In addition, there was not any clear association of patterns with moderate values of crime to any SES level.

8. Conclusion

These three case studies offered evidence that moderate to high crime patterns were more likely than not to be associated with low SES patterns. In only one case study, Tucson, was there a low SES pattern associated with low levels of crime, but each case study did find at least one pattern where low SES was associated with high crime. The only consistent result across all three case studies was that high income was associated with low crime.

Three questions were used to direct this analysis. First, what patterns in the component planes could be identified within the SOMs? Second, how did the SOM help visualize these patterns? Third, how were the tracts grouped in SOM nodes distributed geographically the city?

The first question was useful in identifying the highs and lows of each component plane for further analysis. One problem that consistently came up in the analysis was how to characterize values in between these extremes, particularly in regards to crime values. There was not a precise definition of what a node associated with moderate crime should be, which left analysis outside of the extremes vague. The second question was used to build on the analysis in the first question and identify the global, neighborhood and local patterns in each SOM. The evaluation of the Standard Deviation of the Tract Distance was not as useful as hoped in analyzing each SOM, but Node Distance and Tract Distribution were both useful tools. Node Distance was less valuable than it could have been due to the inconsistent definition of a “high” value in the SES nodes. A low income value should have been normalized to a high value in the income planes, which would not have caused the offsetting high values in Node Distance. The mapping of certain nodes from each SOM was greatly beneficial to the analysis and in confirming that there was an association between high crime and low income.

The data of this study were insufficient to conclude whether there were other factors that could have been considered that would have resulted in a more consistent outcome across all three case studies. As these results did consistently find that economic factors are the most likely to be associated with levels of crime, further studies should include other measurements of the local economy. A study comparing the homes of the criminals, rather than the location of their victims, to SES would also be useful in future work.

This study could also have been improved by use of fewer dimensions in each SOM. The combined SOMs were too complicated, potentially hiding spatial associations between only one or two variables. An evaluation of one or two types of crime against SES variables, or of multiple crime types against a single SES variable may have resulted in a clearer analysis. The identification of correlated variables, or use of a principle component analysis, also could have assisted in filtering the data to a smaller number of components. More precise definitions, such as a precise definition of levels of crime, would have also been beneficial.

References

Aguilera, P. A., Frenich, A. G., Torres, J. A., Castro, H., Vidal, J. L. M., & Canton, M. 2001. Application of the Kohonen Neural Network in coastal water management: Methodological development for the assessment and prediction of water quality. *Water Resources, 35*(17), 4053-4062.

Andrienko, G., Andrienko, N., Bremm, S., Schreck, T., von Landseberger, T., Bak, P., & Keim, D. 2010. Space-in-time and time-in-space self-organizing maps for exploring spatiotemporal patterns. *Computer Graphics Forum, 29*(3), 913-922.

Bazarghan, M. 2012. Application of self-organizing maps to stellar spectral classifications. *Astrophysics and Space Science* (337), 93-98.

Brockett, P., Xia, X., & Derrig, R. A. 1998. Using Kohonen's self-organizing feature map to uncover automobile bodily injury claims fraud. *The Journal of Risk and Insurance*, *65*(2), 245-274.

Cahill, M. Geographies of urban crime: An intraurban study of crime in Nashville, TN; Portland, OR; and Tucson, AZ." (Ph. D. diss.). University of Arizona, 2004.

Chon, T., Park, Y. S., Moon, K. H., & Cha, E. Y. 1996. Patternizing communities by using an artificial neural network. *Ecological Modelling, 90*, 69-78.

Cooper, C., & Burns, A. 2007. Kohonen self-organizing feature maps as a means to benchmark college and university websites. *Journal of Science Education and Technology*, *16*(3), 203-211.

Friedel, M. J. 2011. Modeling hydrologic and geomorphic hazards across post-fire landscapes using a self-organizing map approach. *Environmental Modelling & Software, 26*, 166-1674.

Fritzke, B. 1995. Growing grid – a self-organizing network with constant neighborhood range and adaptation strength. *Neural Processing Letters*, *2*(5), 9-13.

Hagenauer, J., Helbich, M., & Leitner, M. 2011. *Visualization of crime trajectories with self-organizing maps: A case study on evaluating the impact of hurricanes on spatio-temporal crime hotspots.* 25th International Cartographic Conference, Paris, France.

Kohonen, T. 1981. *Construction of similarity diagrams for phonemes by a self-organizing algorithm.* Report TKK-F-A463, Helsinki University of Technology, Espoo, Finland.

---. 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*, 59-69.

---. 1990. The self-organizing map. *Proceedings of the IEEE*, 78, 1464-1480.

---. 2001. *Self-organizing maps* (3rd ed.). Berlin, Germany: Springer-Verlag.

Kaaijveld, M. A., Mao, J., & Jain, A. K. 1992. *Proceedings of the 11th ICPR, International Conference on Pattern Recognition.* Los Alamitos, CA: IEEE Computer Society Press.

Kulldorff, M. 1997. A spatial scan statistic. *Communications in Statistics: Theory and Methods*, *26*(6), 1481-1496.

Leitner, M., & Helbich, M. 2011. The impact of hurricanes on crime: A spatio-temporal analysis in the city of Houston, Texas. *Cartography and Geographic Information Science, 38*(2), 213-222.

Li, S., Kuo, S., & Tsai, F. 2010. An intelligent decision-support model using FSOM and rule extraction for crime prevention. *Expert Systems with Applications, 37,* 7108-7119.

Lin, C., Chen, H., & Nunamaker, J. F. 2000. Verifying the proximity and size hypothesis for self-organizing maps. *Journal of Management Information Systems, 16*(3), 57-70.

McQueen, J. B. 1967. Some methods for classification and analysis of multivariate observations. *Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, California: University of California Press.

Merdun, H. 2011. Self-organizing map artificial neural network application in multidimensional soil data analysis. *Neural Computing & Applications, 20*, 1295-1303.

Muruzabal, J. & Munoz, A., 1997. On the visualization of outliers via self-organizing maps. *Journal of Computational and Graphical Statistics, 6*(4), 355-382.

Park, R. E., & Burgess, E. W. 1925. *The city*. University of Chicago Press.

Rose, D. R., & Clear, T. R. 1998. Incarceration, social capital, and crime: Implications for social disorganization theory. *Criminology*, *36*, 441-480.

Sadeghi, A. 1998. Asymptotic behavior of self-organizing maps with nonuniform stimuli distribution. *The Annals of Applied Probability*, *8*(1), 281-299.

Sebestyen, G. S. 1962. *Decision making in pattern recognition*. New York: Macmillian.

Shaw, C. R., & McKay, H. D. 1942. *Juvenile delinquency in urban areas*. Chicago: University of Chicago Press.

Tran, L. T., Knight, C., O'Neill, R., Smith, E.& O'Connell, M. 2003. Self-organizing Maps for Integrated Environmental Assessment of the Mid-Atlantic Region. *Environmental Management*, *31*, 822-835.

Ultsch, A. & Siemon, H. 1989. *Technical report 329*. Univ. of Dortmund, Dortmund, Germany.

Ultsch, A. 1993. Self-organizing neural networks for visualization and classification. In O. Opitz, B. Lausen, & R. Klar (Eds.), *Information and Classification* (307-313). London: Springer.

Versanto, J., & Alhoniemi, E. 2000. Clustering of the self-organized map. *IEEE Transactions on Neural Networks, 11*, 586-597.

Appendices

Appendix I: Data Histograms

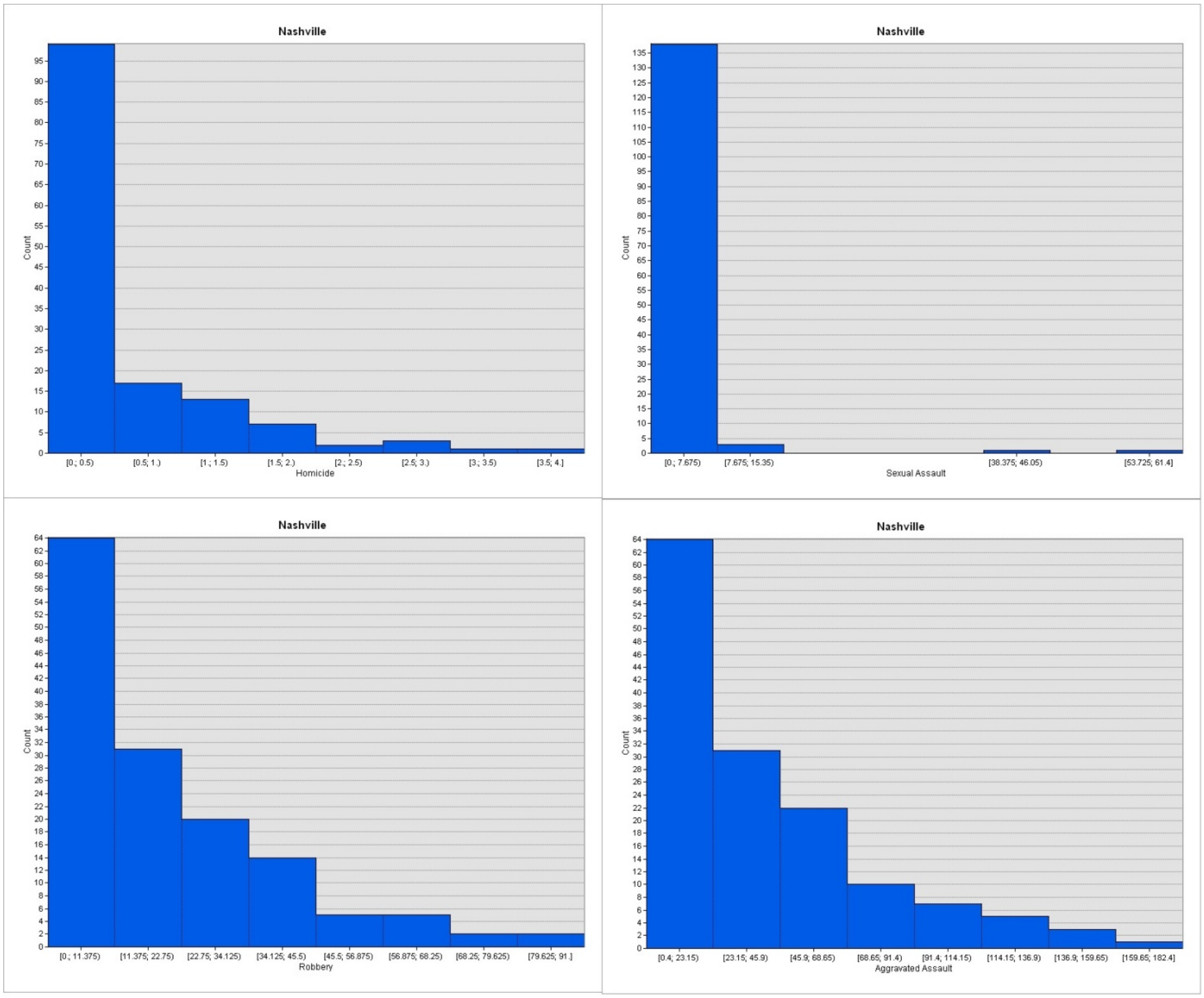

Figure 35 - Frequency histograms for four crimes in Nashville. From the top left, clockwise: homicide, sexual assault, aggravated assault, and robbery.

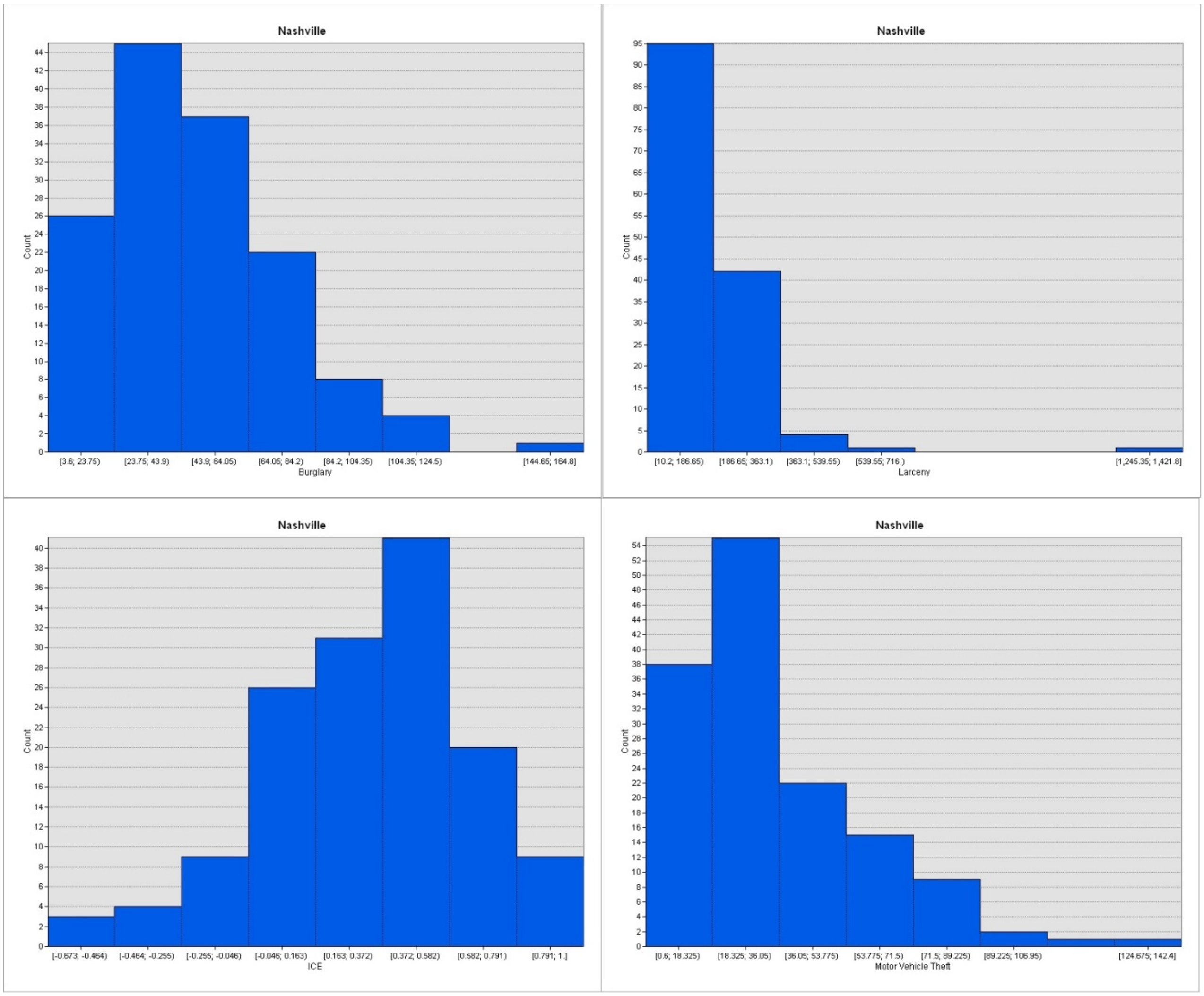

Figure 36 - Frequency histograms for three crimes and one measure of SES in Nashville. From the top left, clockwise: burglary, larceny, motor vehicle theft, and ICE.

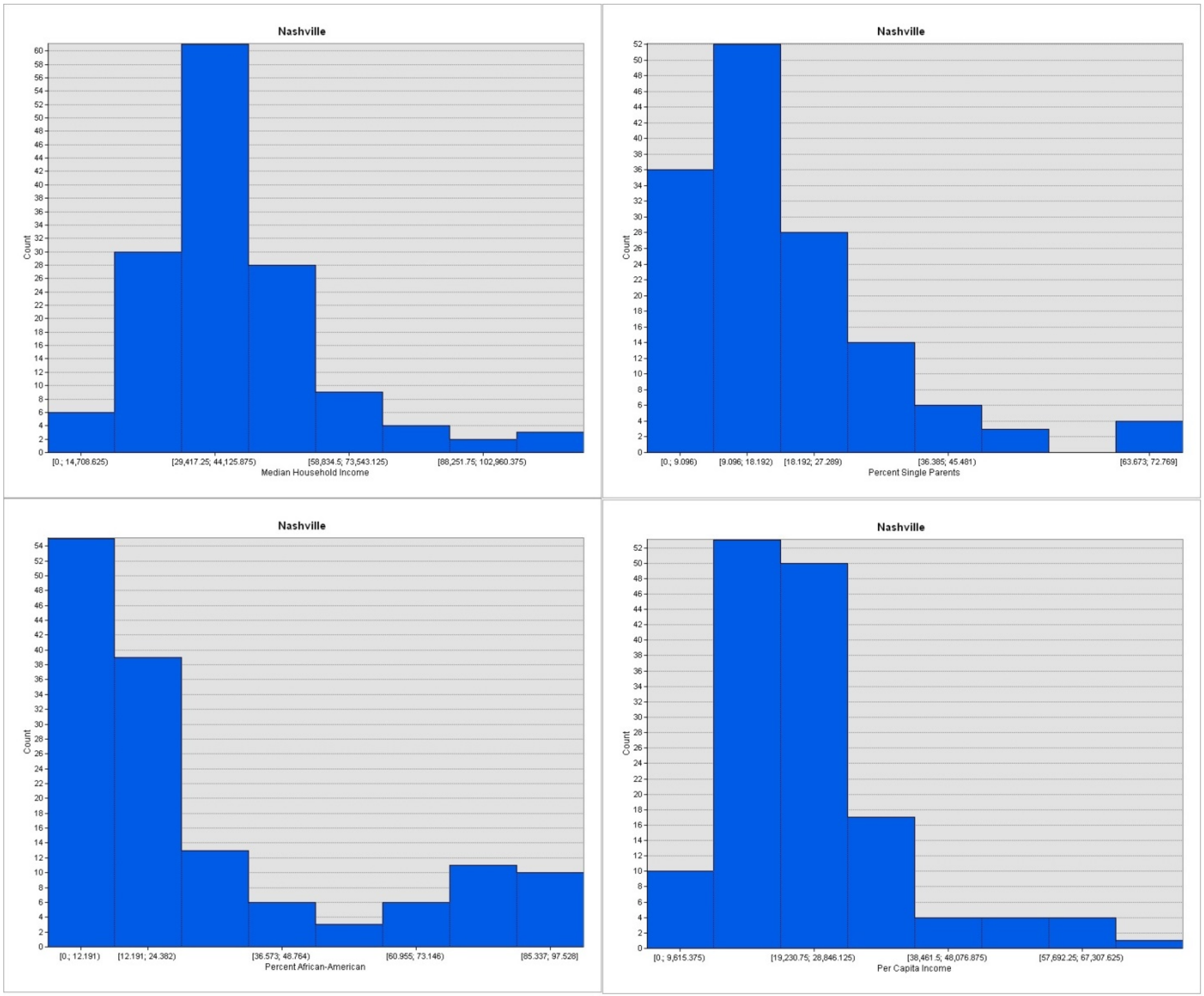

Figure 37 - Frequency histograms for four measures of SES in Nashville. From the top left, clockwise: median household income, percent of all parents that are single parents, per capita income, and percent of all people with African-American heritage.

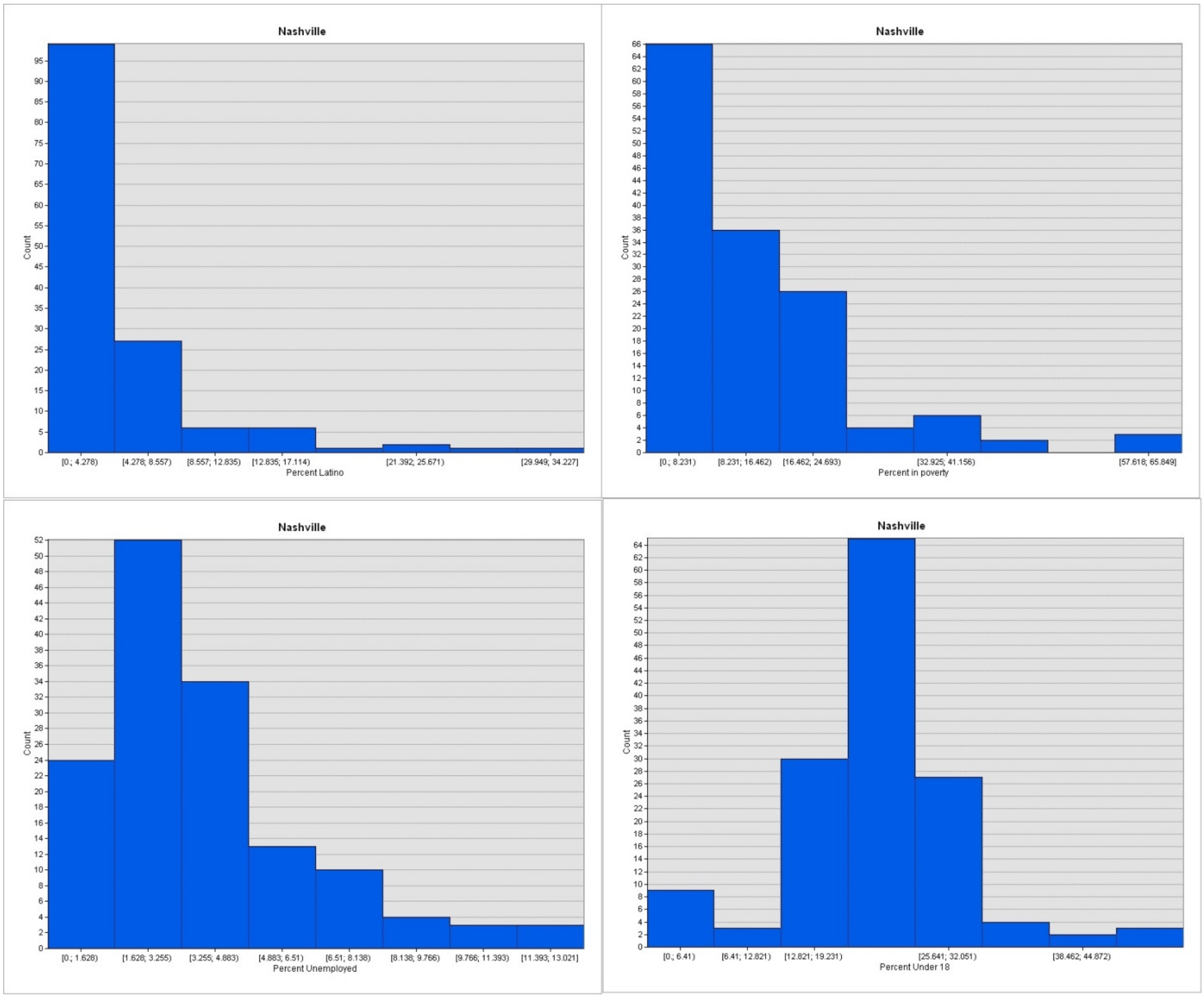

Figure 38 - Frequency histograms for four measures of SES in Nashville. From the top left, clockwise: percent of all people with Hispanic heritage, percent of people in poverty, percent of people under the age of 18, and percent of people unemployed.

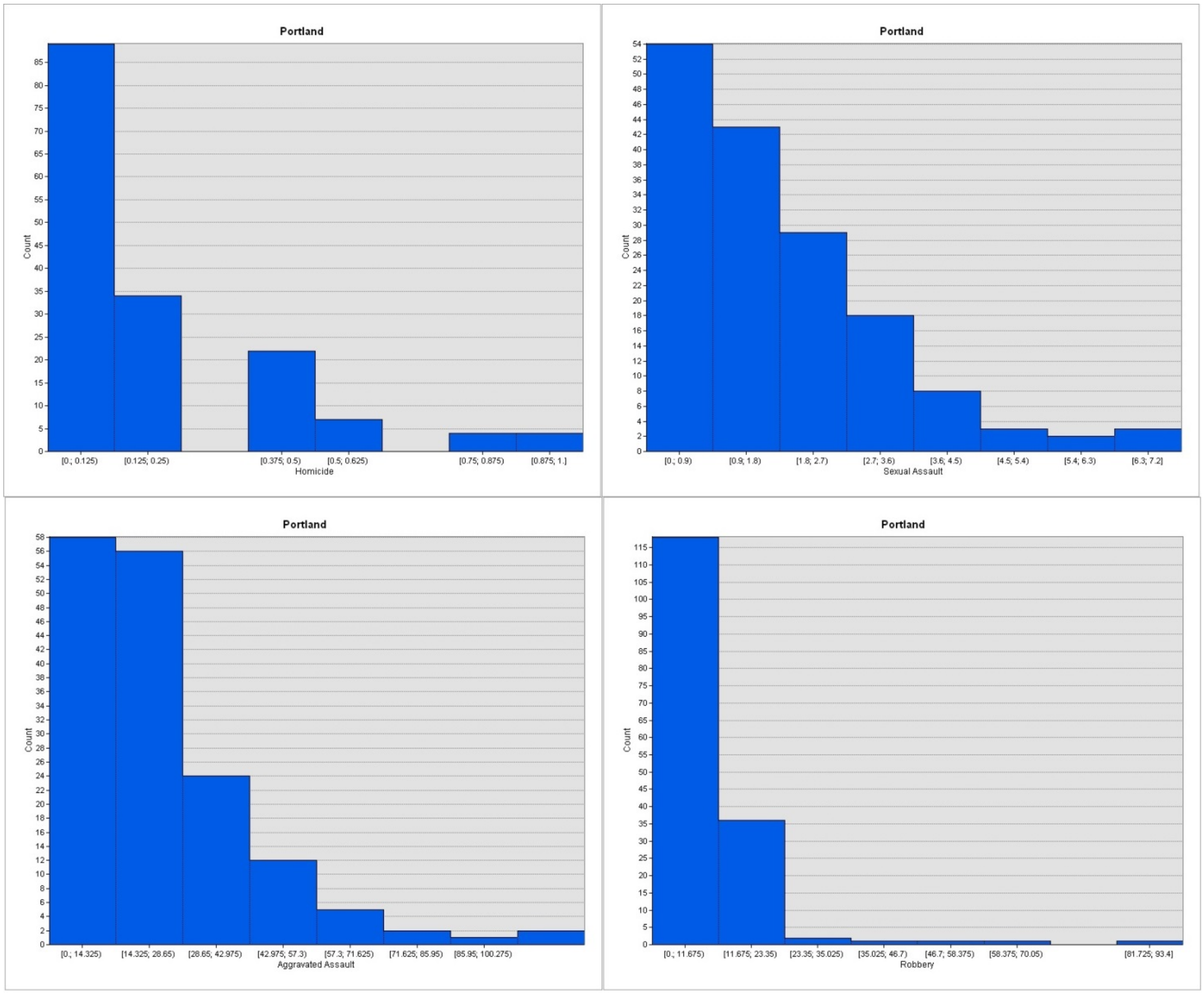

Figure 39 - Frequency histograms for four crimes in Portland. From the top left, clockwise: homicide, sexual assault, robbery, and aggravated assault.

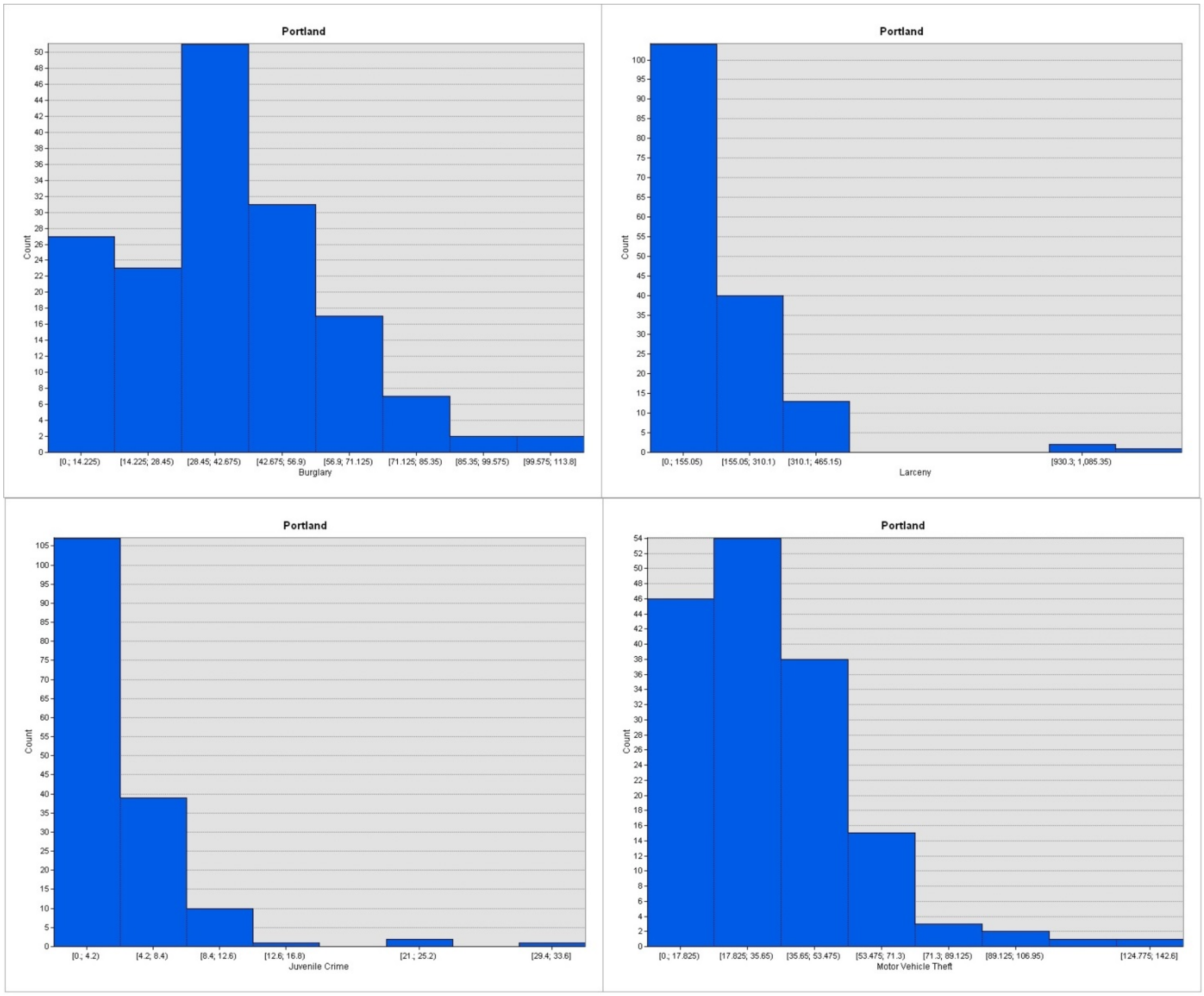

Figure 40 - Frequency histograms for four crimes in Portland. From the top left, clockwise: burglary, larceny, motor vehicle theft, and juvenile crime.

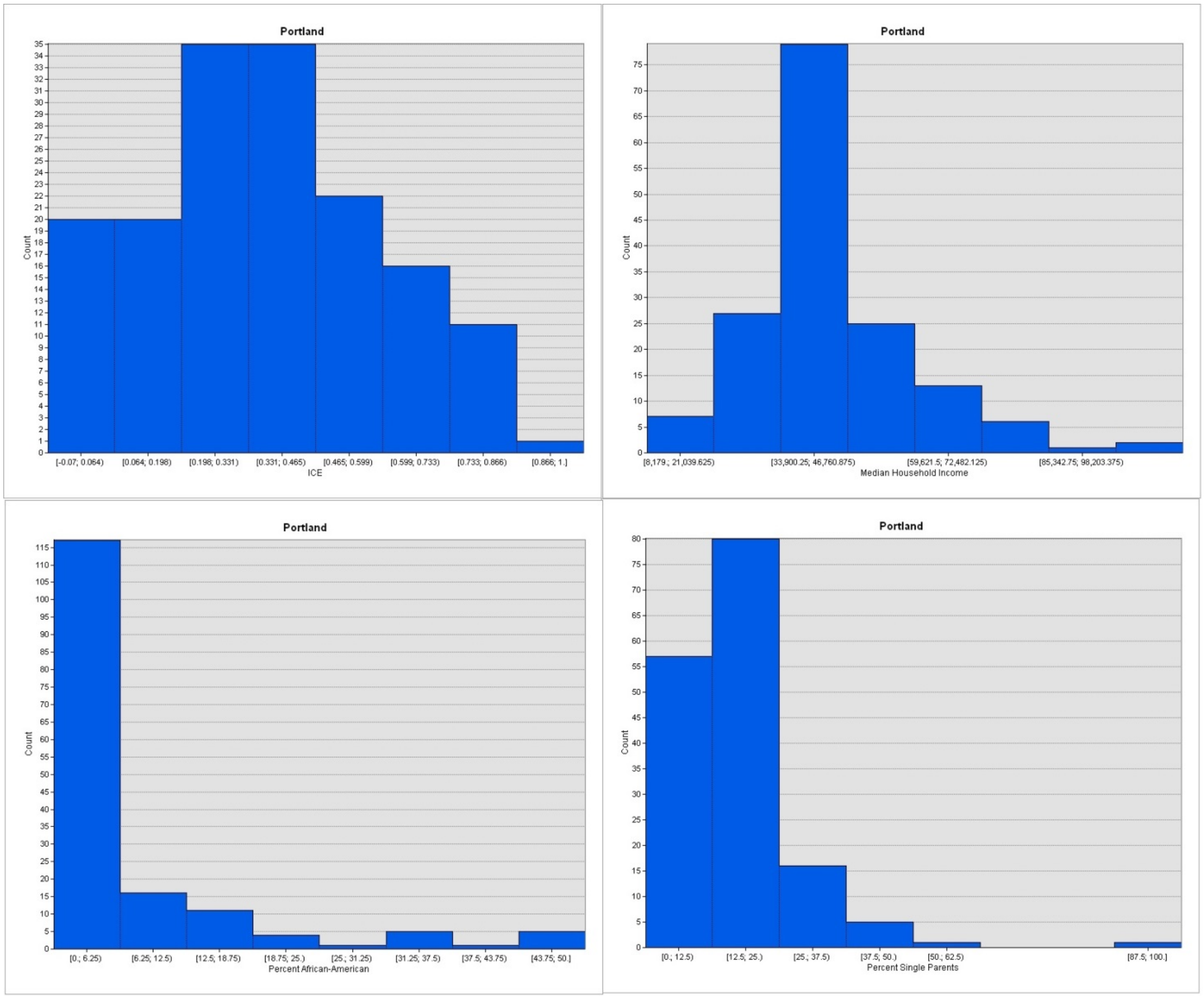

Figure 41 - Frequency histograms for four measures of SES in Portland. From the top left, clockwise: ICE, median household income, percent of all parents that are single parents, and percent of all people with African-American heritage.

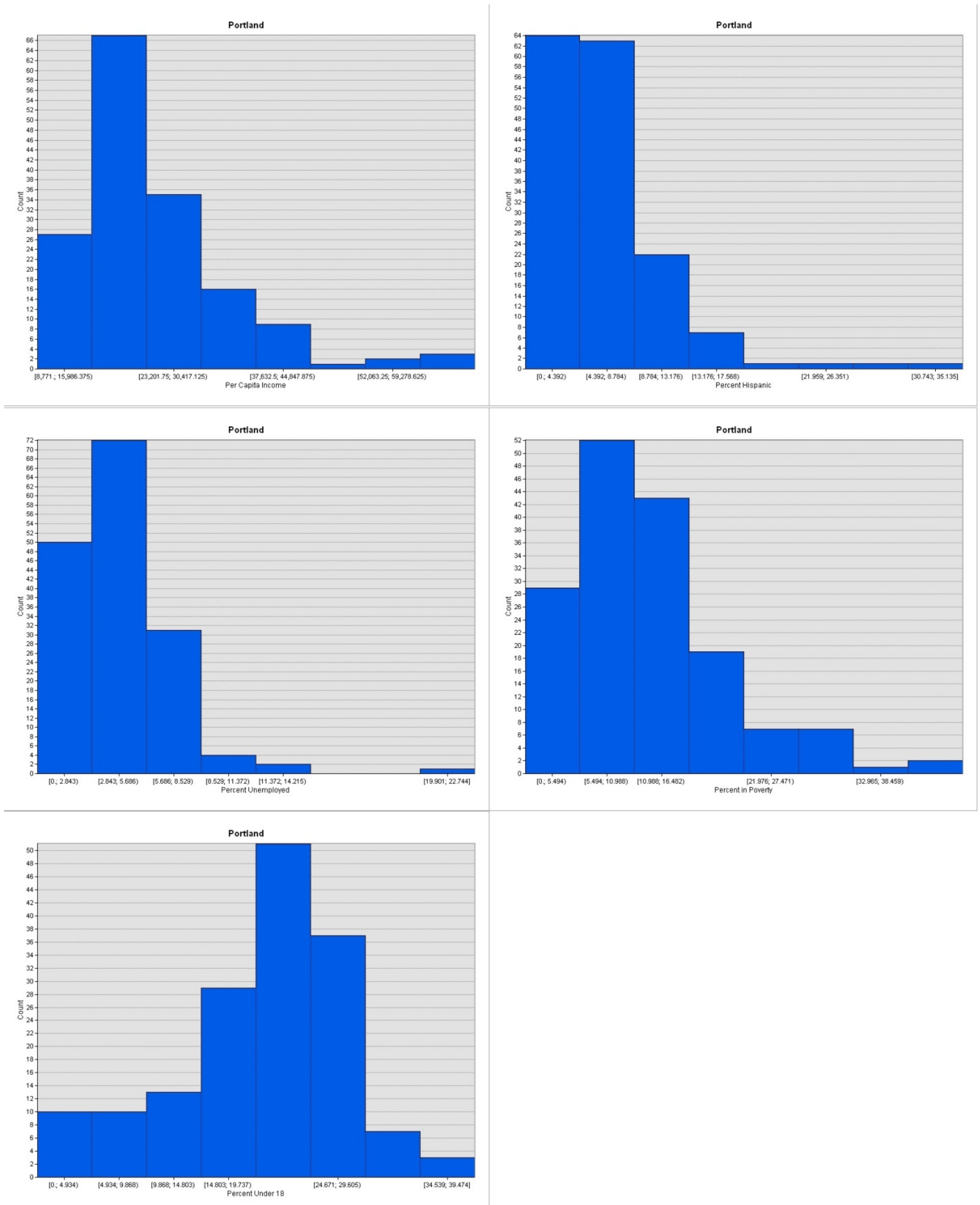

Figure 42 - Frequency histograms for five measures of SES in Portland. From the top left, clockwise: per capita income, percent of all people with Hispanic heritage, percent of people in poverty, percent of people under the age of 18, and percent of people unemployed.

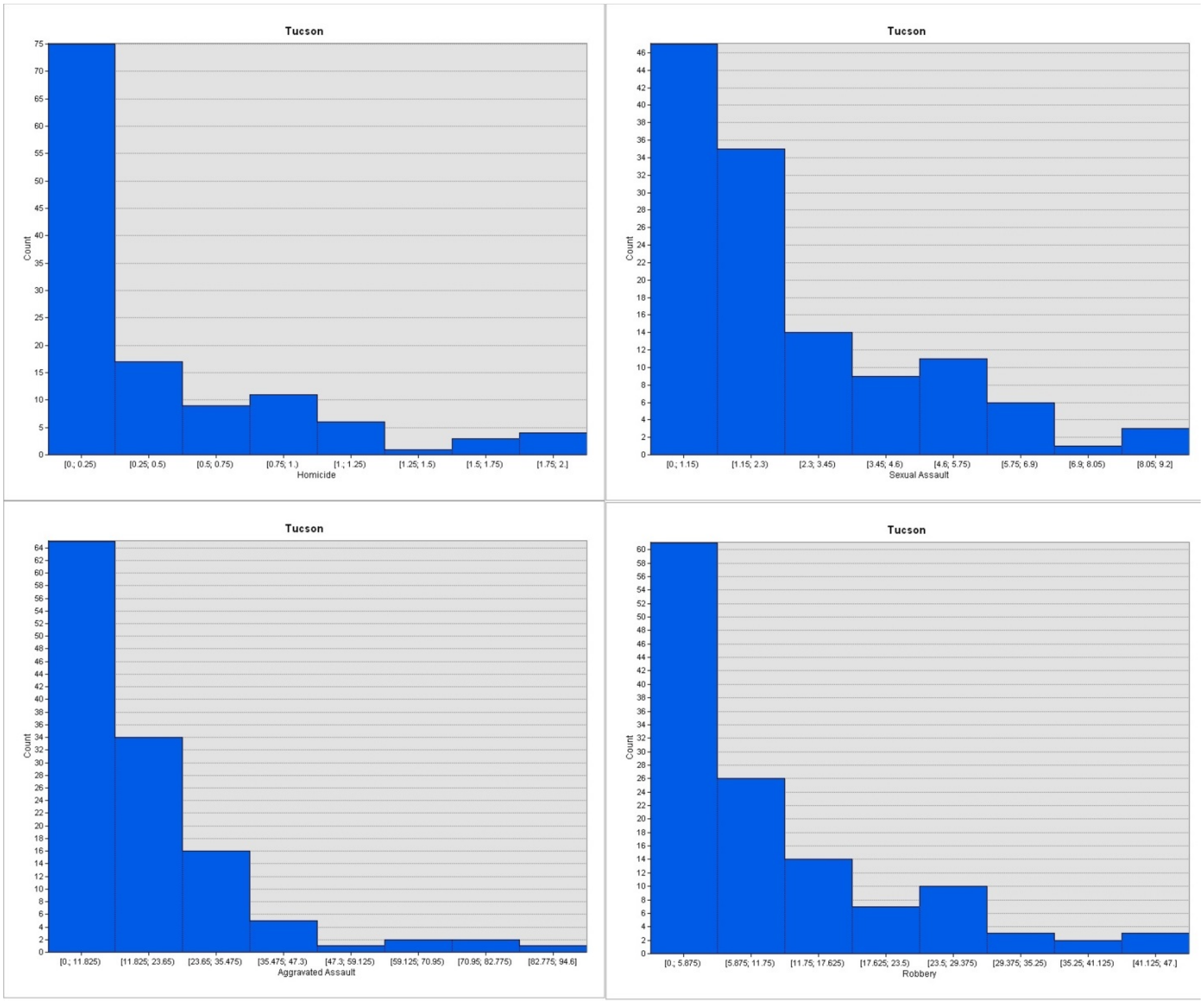

Figure 43 - Frequency histograms for four crimes in Tucson. From the top left, clockwise: homicide, sexual assault, robbery, and aggravated assault.

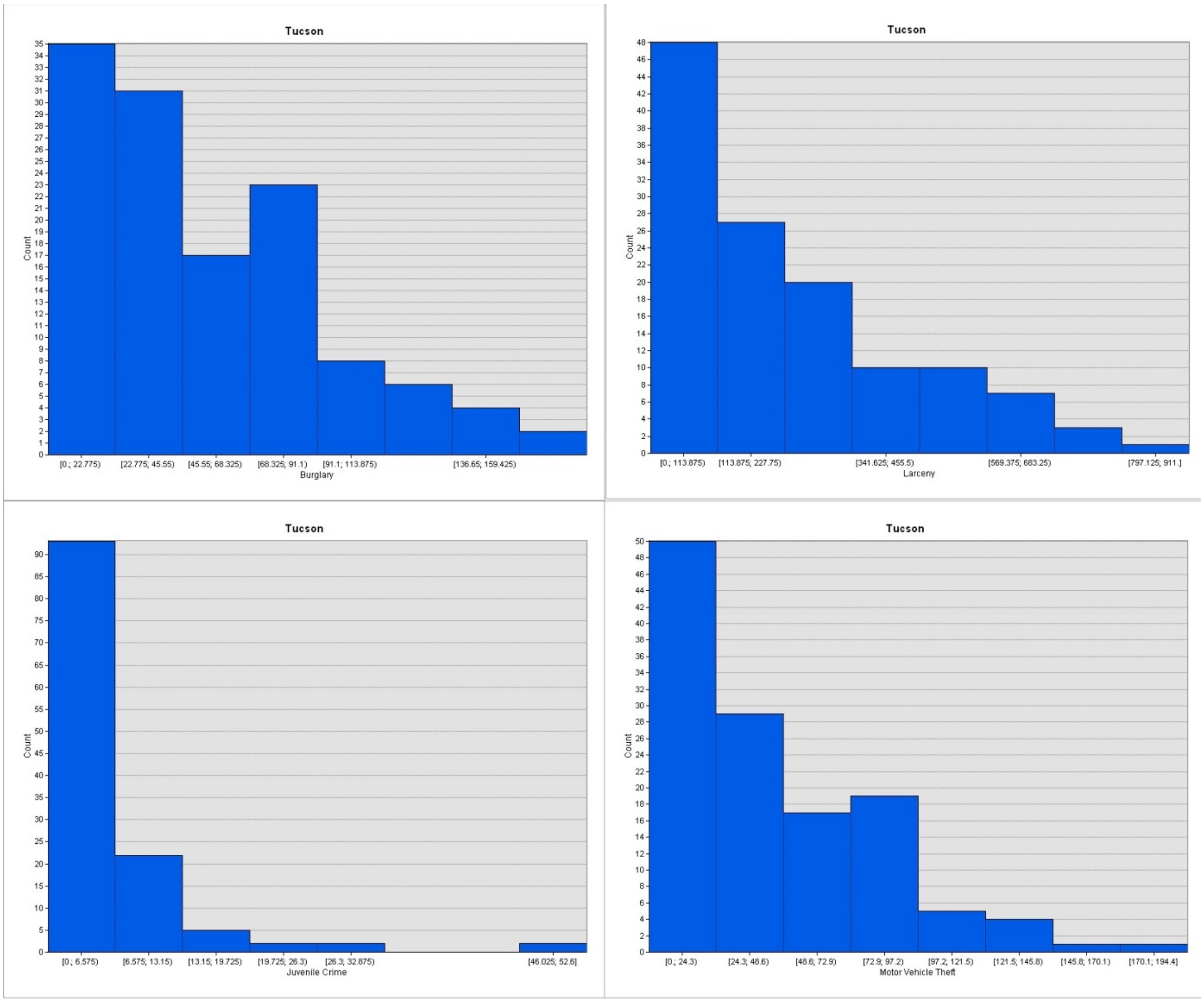

Figure 44 - Frequency histograms for four crimes in Tucson. From the top left, clockwise: burglary, larceny, motor vehicle theft, and juvenile crime.

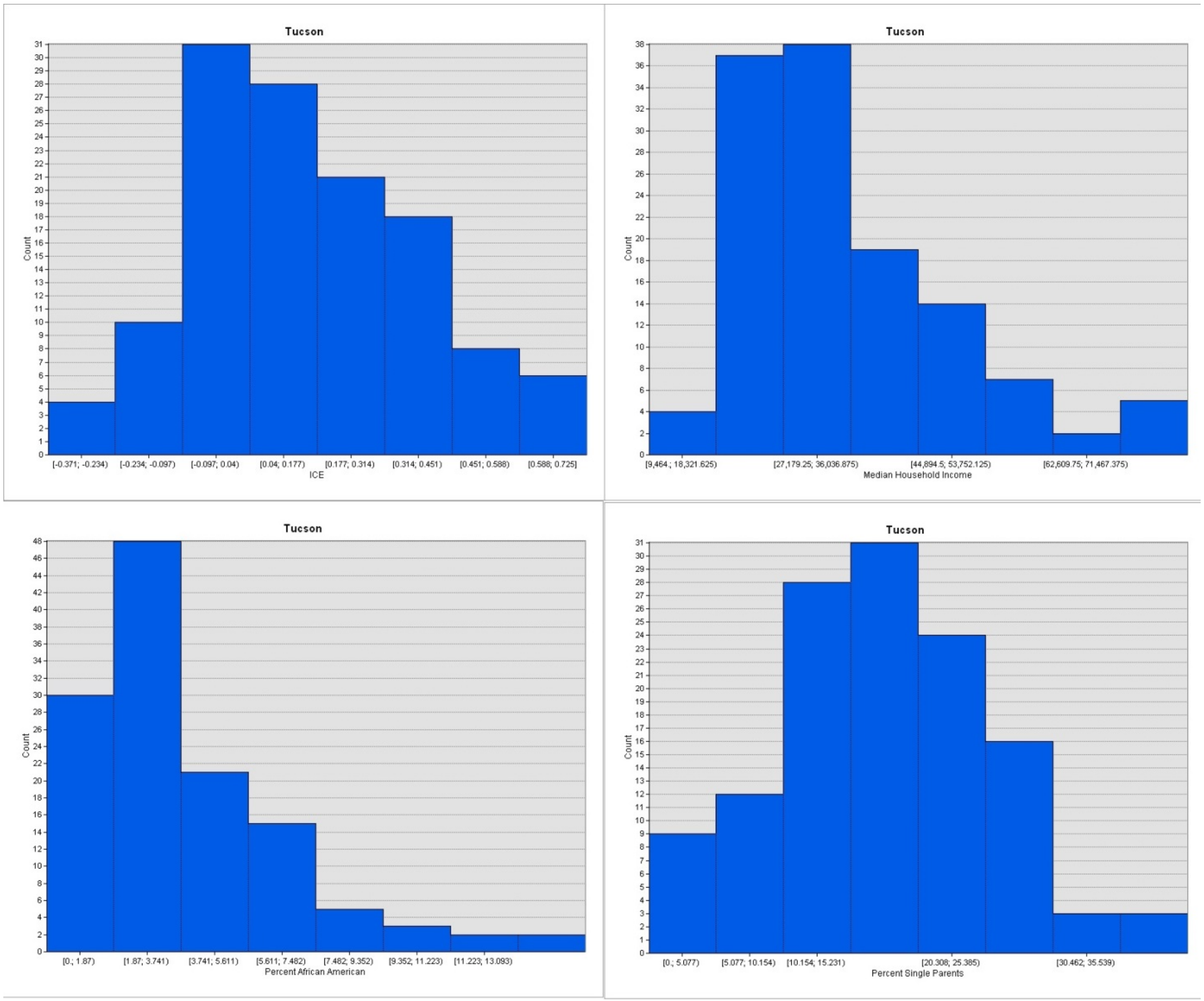

Figure 45 - Frequency histograms for four measures of SES in Tucson. From the top left, clockwise: ICE, median household income, percent of all parents that are single parents, and percent of all people with African-American heritage.

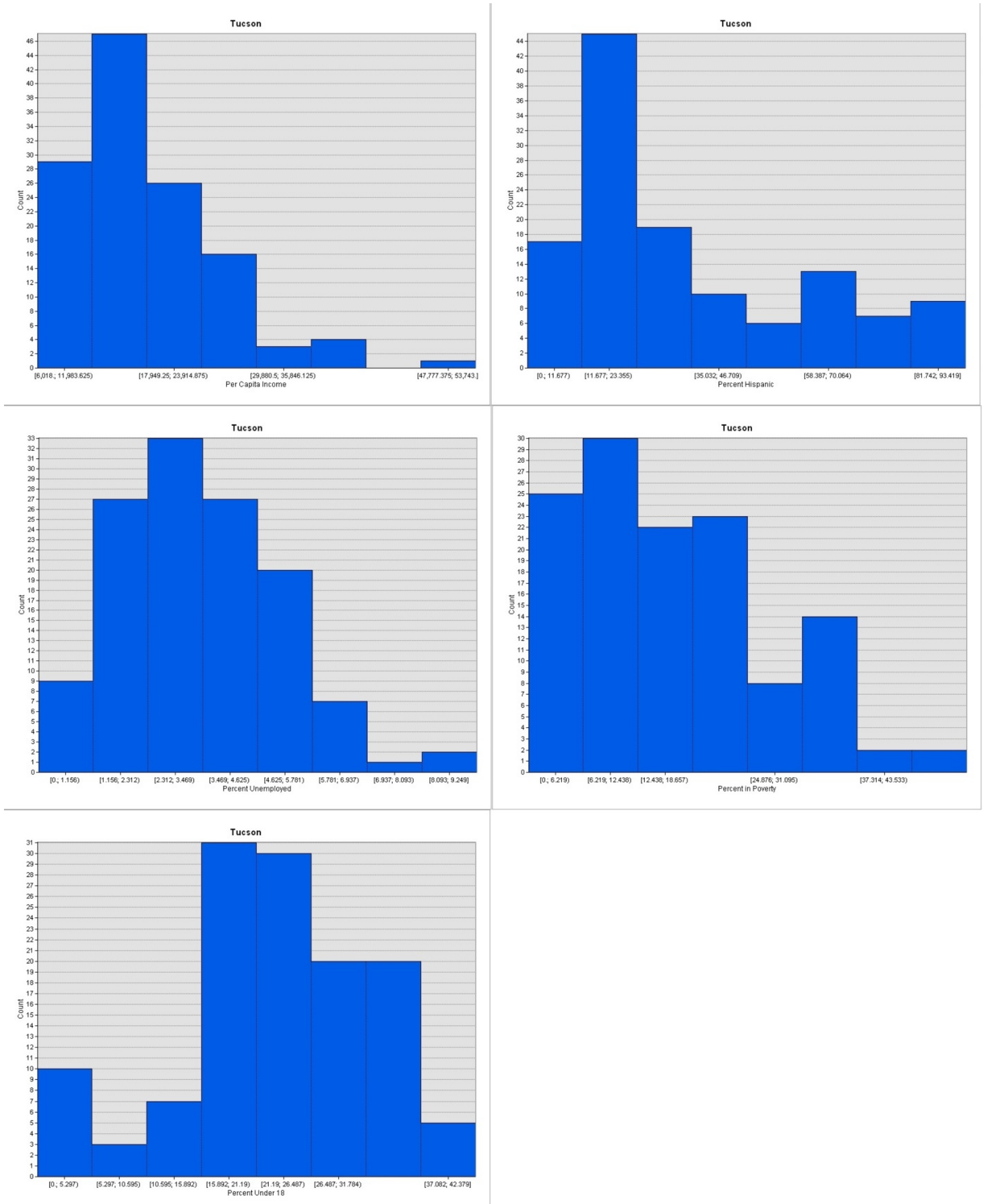

Figure 46 - Frequency histograms for five measures of SES in Tucson. From the top left, clockwise: per capita income, percent of all people with Hispanic heritage, percent of people in poverty, percent of people under the age of 18, and percent of people who are unemployed.

Appendix II: Data Visualizations

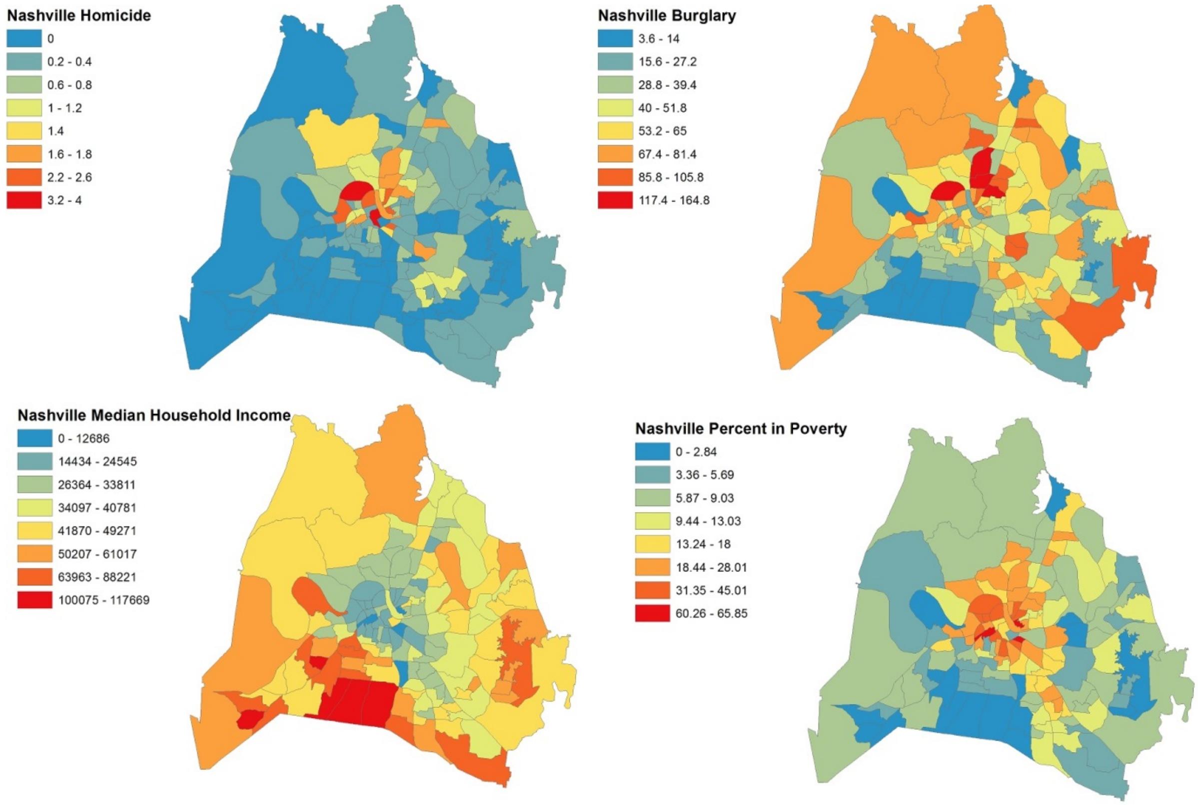

Figure 47 – Maps of two measures of crime and two measures of SES in Nashville. From the top left, clockwise: homicide, burglary, percent of people in poverty and median household income.

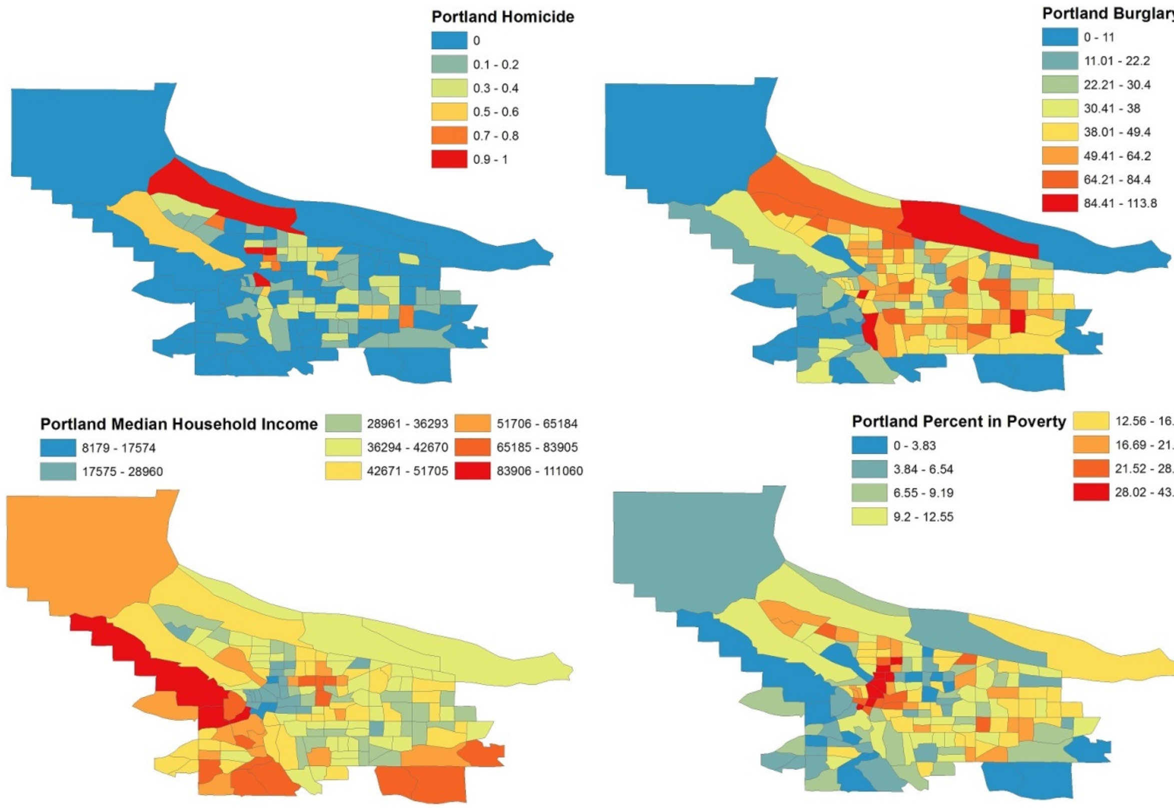

Figure 48 – Maps of two measures of crime and two measures of SES in Portland. From the top left, clockwise: homicide, burglary, percent of people in poverty and median household income.

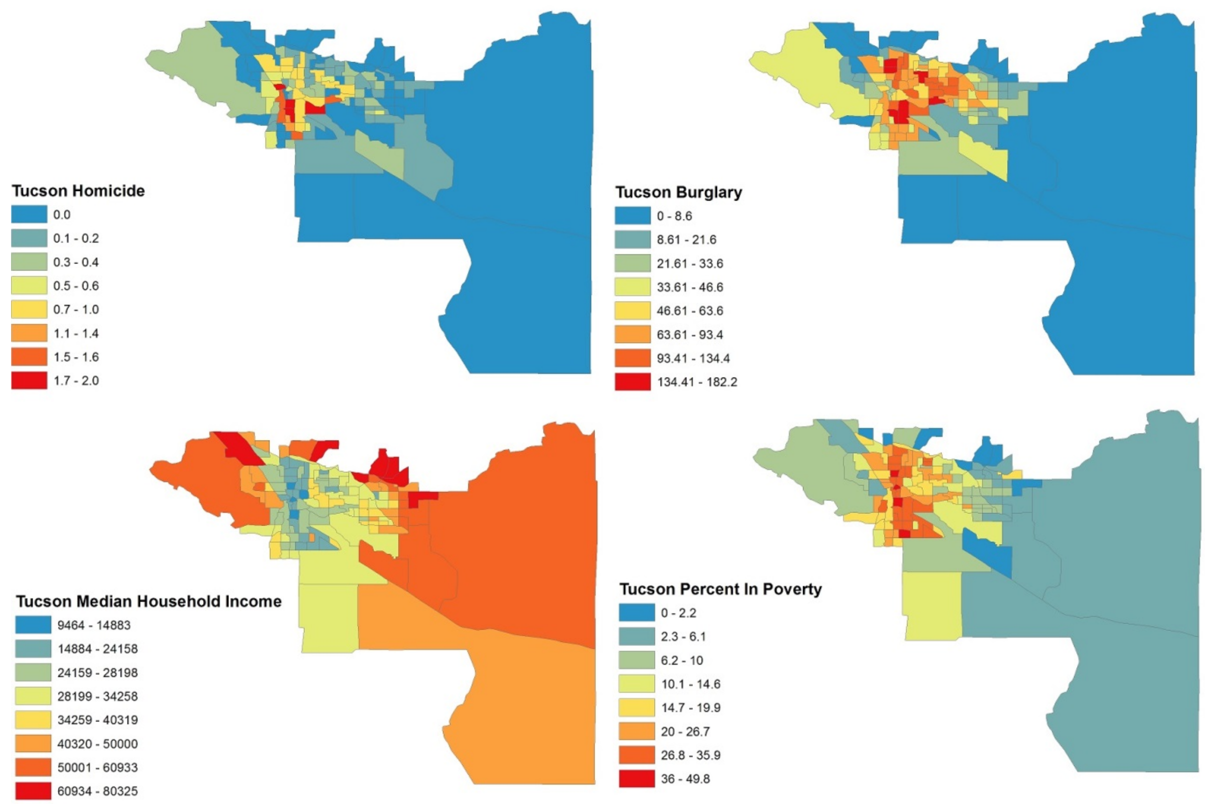

Figure 49 – Maps of two measures of crime and two measures of SES in Tucson. From the top left, clockwise: homicide, burglary, percent of people in poverty and median household income.

Vita

Jason Kaufman was born in State College, Pennsylvania to the parents of Joyce and Keith Kaufman. He has never lived long in one place, moving from Pennsylvania shortly after birth to Connecticut, Ohio and then Tennessee, where he graduated high school. From there, he moved to Huntsville, Alabama to attend the University of Alabama in Huntsville where he received his Bachelor's Degree in Earth System Science at the end of 2009. Moving afterwards to Knoxville, he eventually began attending the University of Tennessee in pursuit of a Master's Degree in the field of Geography. While completing that degree he has worked first as a graduate teaching assistant to the department, a graduate research assistant, and then finally working at the Oak Ridge National Lab, where he will continue to work after completing his Master's.