Geo 578: GIS Applications 2019 Spring Semester

Green Urbanism: Exploring Spatial Correlation Between Access to Parks and Health

Tim Prestby - Joe Marks

Submission date: 4/30/19

Table of Contents

Capstone Statement

Recent scientific literature suggests a correlation between access to nature and overall health. As urbanization continues to skyrocket, most people's limited access to nature exists as urban greenspace, particularly parks. We seek to explore a possible spatial correlation between access to nature and human health for ten cities around 650,000 people spaced throughout the United States. We will perform a comparative spatial analysis by evaluating how close and far distances to parks link to the central tendency for physical health and mental health indices per census block group.

Background

Although the past century has brought huge advances of technology, medicine, and material wealth to most people, mental disorders have skyrocketed to an unprecedented rate; to the point where over one in four people have some sort of mental illness (Adi nd *et al*, np). In wake of this health epidemic, numerous studies have sought to uncover the root of this ailment. Many such studies have focused on examining the correlation between mental health and environmental factors in urban areas; since over half of the entire world's population today lives in cities and an estimated 68% of people will live in cities by 2050 (United Nations 2018, np). These studies point to artificial environments imposing a state of anxiety because of the lack of nature, which helps regulate our physiological functions tied to restoration, hormone regulation, and stress recovery (Hyvönen 2018, 2; Song 2016, 1-3).

Historically, humans have spent the majority of time with close access to nature until the industrial revolution, beginning in the 17th century, sparked urbanization (Boyle 2004, np). For urban residents, human exposure to nature has dwindled ever since then as trees grew into buildings and pristine paths turned into roads. The 'concrete jungle' of urban life lacks nature except for allocated green space, such as fields, gardens, and parks. These green spaces provide a calming, rejuvenating escape from the otherwise chaotic and stressful urban environment. Green spaces, such as parks, can alleviate a variety of mental health aspects including aggression, fatigue, stress, depression, restlessness, eating-disorders. (Hyvönen *et al* 2018, 2; Bjørnstad *et al* 2015, np; Song *et al* 2016, 1-3; Annerstedt, M., & Währborg, P. 2011 371-372,380).

Contact with nature in green spaces can elevate cognitive well-being such as elevated abilities to focus, maintaining a positive attitude, and relaxation (Hyvönen *et al* 2018, 2; Bjørnstad *et al* 2015, np; Song *et al* 2016, 1-3, 11, 13 ; Annerstedt and Währborg 2011, 380). Accordingly, nature soothes humans' psychological state by providing tranquil visuals, engaging yet calming environments, and feelings of harmony as well as being away (Hyvönen *et al* 2018, 2, Song *et al* 2016, 3-5; Annerstedt *et al* 2013, 1-2). In today's busy society, burnout and work inefficiency tend to run rampant while inducing stress. Green space deactivates the sympathetic nervous system which causes stress while turning on the parasympathetic nervous system which soothes the body with relaxation and restoration (Song *et al* 2016, 2).

The key factor to the mental-health epidemic is that today's urbanized society perpetuates long-term exposure to stressors (Annerstedt *et al* 2013, 1). Current initiatives to combat this epidemic focus on better landscape planning to promote green space access (USDA 2018, 10-11; Beyer *et al* 2014, 1-4). Above all, Grahn and Stigsdotter (2003,1-3) argue that the urban residents' "distance to public urban open green spaces seems to be of decisive importance" to the occurrences of mental-health related illnesses.

This paper seeks to examine the validity of these claims using geospatial data analysis techniques. We seek to determine if there is a correlation between access to green space/parks and mental health. Also, we will see if the results of our correlation demand further investigation by understanding socioeconomic groups. Park access may have ties to demography and socioeconomic status (Sister *et al* 2010, 231-232). As a scientific review done by the United States Agriculture Department explains (2018, 16), "A number of studies have concluded that the distribution of urban green space is related to measures of socioeconomic status, such as income, education, race/ethnicity, and occupation, and regularly report that neighborhoods with higher socioeconomic status enjoy greater access to nearby green space". Low-income families of diverse backgrounds especially benefit from greater access to green space (Jenning and Gaither 2015, 1953-1956). Specifically, these families have reduced psychological health conditions including depression and stress (Jenning and Gaither 2015, 1953-1956).

We will use 10 cities as our areas of study of similar population size to avoid ecological fallacy and its problems of scale (Mennis, 2019, np). These cities include Seattle Washington, Denver Colorado, Detroit Michigan, Washington D.C., Boston Massachusetts, Memphis Tennessee, Nashville Tennessee, Portland Oregon, and Oklahoma City Oklahoma. These cities were chosen because they have populations ranging from around 620,000 to 690,000 which helps avoid ecological fallacy (U.S. Census Bureau 2017, np).

In the following sections, we dive deeper into the methodology of our project. Specifically, we detail our workflow structure, geographic information system (GIS) techniques, and our decisions, as well as justification for our analysis. Then, we present the results of our GIS workflow and supplementary maps presenting the findings. The limitations of our analysis and the assumptions we made relating to GIS are discussed. Finally, we evaluate whether our research questions were met and the direction this points for future research.

Original Data Layer

One of our ten cities of interest, Memphis Tennessee, did not have a usable park/green space shapefile for our analysis. So, we georeferenced and digitized a park map provided by Tennessee's GIS department (Jackson, 2018, np). Originally, the park image was in the state plane projected coordinate system titled 'NAD 1983 StatePlane Tennessee FIPS 4100 Feet'. To georeference, we chose ten ground control points that were easily recognizable and spaced out throughout Memphis to adhere to standard georeferencing practices (UCONN, nd, np). After this, we rectified the map. Then, a geodatabase was created with line and polygon feature classes. Digitization is a process of tracing features on an analog map to make it available it digital form (Chang, 2015, 90-110). We digitized the 118 parks using ArcMap, projected it using Albers Conic Equidistant Conic projection to preserve distances across meridians, and exported it as a shapefile.

Conceptualization

In order to test the claim that there is a correlation between access to green space and human health, we will break down the research into key concepts for analysis. These concepts include an individual's access to green space and human health. For simplicity, we will be looking at access to parks to represent an individual's access to green space, as the data is more readily available for the cities included in our research. In each subsection below, we explain the variables selected to represent each concept, and the operationalization of those variables.

Access to Parks

Many people, because of lack of overall access or access to transportation, visit parks and green space rarely or not at all (Blanck et al 2012, 424-425). Most people point to distance and lack of time as the obstacle (Grahn and Stigsdotter 2003, 10). For these reasons, we chose the relative distance to parks to represent an individual's access. Although we acknowledge that other barriers to parks may exist, such as climate and terrain, safety fears such as high traffic speeds, and lack of sidewalks and lighting (Hansen et al 2015, 2, 4; Blanck et al 2012, 425), such factors are beyond the scope of this paper and certain simplifications had to be made.

Because this paper focuses on finding a correlation between access to parks and health, the distance will be calculated by creating buffers from each park. Centroids of census tracts falling within or outside of these buffers will be analyzed to see if a correlation exists between the distance to a park and human health.

In order to simplify the implementation and reduce the overhead of possibly creating 10 separate layers (one for each city) to represent parks, the existing park layers for each city will be used. We realize this may not include all available green space to the public, and that each city may have different definitions on what constitutes a park and thus may include different features within their corresponding layer, but we believe these layers provide a reasonable approximation of green space. All cities provided park shapefiles except for Memphis, Tennessee. Therefore, we digitize a park map of Memphis, Tennessee using a digital map provided by the Memphis GIS department.

When calculating the distance buffers from parks, we choose to use the euclidean distance from the the perimeter of a park. We also use a spatial query using the census tract centroid as it provides an approximate average distance for all residents within the census tract. We chose to calculate distance using the perimeter of the park because one does not need to walk to the center of a park to be considered "within" it, or even close to it. Our hope is also that by using the park perimeter instead of the park centroid, it will reduce distance error for very large parks. For example, if calculating distance based on a park's centroid, a census tract that is directly adjacent to a two mile long park would have the extra distance of one mile covered by he buffer. Our reasoning for using euclidean distance as our measure of distance is to simplify the workflow for the ten cities. While using network analysis on roads is typically more accurate (Nicholls 2001, 207, 216; Koohsari 2011, 69), euclidean distance is still used as a valid measure of approximation in GIS (Nicholls 2001, 207; Basnet et al 2001, 523).

To measure the access to parks, various zones of distances will be implemented according to the amount of walking time people are willing to walk. These walking times will be translated into distances that then can be analyzed. Three different zones/distances will be implemented as 'Excellent', 'Good', and 'Fair', 'Poor'. For the following class breaks, we recognize that differences of say 0.24 miles and 0.26 miles or 0.49 miles and 0.51 miles have minimal discrepancies in distances. However, by imposing these hard boundaries, a large difference in scoring will apply to values close to these class breaks. A general standard that the transit industry uses is that people are willing to indefinitely walk 0.25 miles (400 meters) and probably walk 0.5 miles (800 meters) which take around five and ten minutes respectively (Blanck et al 2012, 424; Martin 2004, 1; National Recreation and Park Association 2014, 4;

Bancroft et al 2015, 28; Nicholls 2001, 209). These standards have been enacted by over half of the largest cities in the U.S. (Harnik and Martin 2004, 1). A systematic literature review of access to parks included seven studies revealing that the maximum distance most people are willing to walk is one mile or 20 minutes (Bancroft et al 2015, 28; Moore 2008, 2). Anything above this amount is too much for most people to walk, so they would have to choose some sort of other mode of transportation. We exclusively focus on the mode of transportation of walking which almost all people can do (baring physical constraints). The 'Excellent' walking conditions are distance less than 0.25 miles. Good walking distances are distances equal to or above 0.25 miles and less than 0.5 miles. Fair walking distances are equal to or above 0.5 miles and less than 1 mile. Poor walking distances are equal to or above 1 mile. These walking distances will be implemented into buffers of their respected distances around the perimeter of the parks to determine which census tracts fall within each distance marker. Previous access to parks research has relied heavily on buffers to execute the analyses to minimize errors and computational demands (Nutsford et al 2013, 1007-1009; Bancroft et al 2015, 28; Nicholls 2001, 210). In addition, and other studies emphasize incorporating a distance decay metric to take into consideration spatial interaction processes across varying geographies (Zhang et al 2011, 4-5). However, Zhang et al (2011, 4-5) warns that for distance decay measurements, "the information or data needed to calibrate this parameter usually are not available". Therefore, we do not include a distance decay metric in our research.

Human Health

Recent demographic studies have found a positive association between exposure to urban green space and the perceived general health of residents (Song et al 2016, 5). In terms of physical health, access to green space can reduce many health conditions such as obesity, psychological health, and heat-related illness (Jenning and Johnson 2015, 1958). Living in areas with accessible green spaces for walking also increases the longevity of senior citizens, independent of age, sex, marital status, baseline functional status, and socioeconomic status (Takano et al 2002, 917), and a lack of recreational facilities and green space is associated with decreased physical activity, increased obesity, and other physical ailments. In fact, poor physical health tends to increase mortality rates and the occurrence of chronic ailments such as circulatory diseases. (Mitchell 2008, 1655-1656, 1658-1659). In terms of mental health, as mentioned before, green spaces, such as parks, can alleviate a variety of mental health aspects including aggression, fatigue, stress, depression, restlessness, eating-disorders. (Hyvönen *et al* 2018, 2; Bjørnstad *et al* 2015, np; Song *et al* 2016, 1-3; Annerstedt, M., & Währborg, P. 2011 371-372,380). Therefore, we choose to look at psychological health and physical health to define overall human health.

To measure psychological health for each census tract, we will be looking at survey data conducted by the Centers for Disease (CDC) in 2016 (CDC 2016, np). This survey was simplified to quantify the number of people over age 18 who reported negative mental health within the last fourteen days at the time of the survey (CDC 2016, np). To operationalize this variable, we chose to use a measure of central tendency as a statistical measure that identifies a single value as representative of the entire distribution (Manikandan 2011, 140). Then we classify census tracts into groups with "good" mental health and 'poor' mental health based on their relation to the mean of each respected city. We choose to classify the variable based on data distribution in order to create a layer which could easily be combined with our similarly classified physical health layer in order to create a layer representative of overall human health. Such practices are common in site suitability analysis', where classification was based on the data range, data type, and data distribution (Basnet et al 2001, 523). We decide to use the median as we have many outliers and skewed data that would otherwise distort results (Manikandan 2011, 214). It is also easy to calculate and will fall "between the mean and mode in a skewed distribution' (Manikandan 2011, 215).

Tracts whose median mental health is above the overall median for the city are those with "good" mental health, and tracts whose median mental health that are equal to or below the overall median for the city are those with "poor". We choose to sort tracts that have a median mental health equivalent to the overall median into the "worse" mental health category in order to give the benefit of the doubt to the opposing argument.

To measure the physical health for each census tract, we will be using a similar approach to that of mental health detailed in the paragraph above, using survey data conducted by the Centers for Disease (CDC) in 2016 (CDC 2016, np). This survey measures the number of people over age 18 who reported negative physical health within the last fourteen days at the time of the survey (CDC 2016, np). We use the median as a measure of central tendency for the same reasons stated above (Manikandan 2011, 140, 215) in order to sort each census tract into categories. Tracts with "good" physical health are those that have a tract median greater than the overall tract median for the city, while tracts with poor physical health are those that have a tract median less than or equal than the overall tract median for the city. Once again, we choose to sort the tracts equal to the overall physical health median in the city as poor, as we wish to better represent the opposing side to make our own claim more substantial.

We choose to sort the data into categories using medians for each city in order to let the data for each city define its own categories. We expect there to be variation in mental health and physical health data between cities, and want to avoid it as a possible source of error. In our analysis, we compare these values to the overall median for all cities to see what differences may arise. We also acknowledge that a disadvantage of using the median as a measure of central tendency is that it does not use each value of observations into the measurement which can harm the statistical precision (Manikandan 2011, 215). However, given the large size of our sample and many outliers that would otherwise distort the results, we believe it to be a good approximation. In our analysis, we examine the distribution of the data in the census tracts for possible error that may occur when using the median and report said errors in the result.

After categorizing the data for each census tract, we will combine the two into a single metric for human health using boolean combination, the result of which will be used for our analysis (Basnet et al 2001, 529).

We recognize that our data layers do not match exactly in terms of the dates the data was gathered, and that this may be further complicated by the fact that our park layers may also have varying dates since they were last updated. However, we believe that the dates are similar enough to provide an accurate estimate to show if a correlation exists.

Implementation

Our implementation was broken down into three stages: assigning a score for distance to parks for each census tract, assigning a score for health for each census tract, and then performing analysis using the resulting scores. All work was done in ESRI's ArcMap versions 10.5 and version 10.6, python version 3.7, and NumPy/SciPy version 1.16.3.

To begin, the census tract centroid locations were extracted from the 2018 release of the 500 cities data, which already had these coordinates defined under the "Geolocation" column. These centroids were added to the map using ArcMap's Make XY Event Layer (Data Management) tool with GCS_WGS_1984 set as the geographic coordinate spatial reference, as this was the datum the 500 cities data used in its 2018 release of the data. This data was then exported as a shapefile in order to add an 'Object ID' field, then added back to the map. The centroids were then projected to the World Equidistant Conic projection, which we use for all layers in our map. For each city, the parameters for the World Equidistant Projection were set using an online tool called 'Projection Wizard' to match the approximate boundaries of the city. We attempted to filter out extraneous features within the park layer by only including features with the land use type of park and any features which did not include "park" in their name. We then projected the parks layer obtained from each city's GIS records (excluding Memphis, which was digitized) and added it to the map. We used ArcMap's Buffer (Analysis) tool to create buffers around the park polygons with a distance of 0.25 miles, 0.5 miles, and 1 mile to represent excellent distance to parks, good distance to parks, and fair distance to parks respectively. Any centroids not within these buffers (equal to or above a 1 mile distance) would be considered to have poor distance to parks. We then added a field to our centroids layer to hold our assigned distance scores. Next, we used ArcMap's Select By Location (Data Management) tool in order to select centroids that fell within each respective buffer and assign them a score for the distance metric accordingly; a score of 3 for centroids that were within the excellent distance buffer, a score of 2 for centroids that fell within the good buffer, excluding those that were also within the excellent buffer, a score of 1 for centroids that fell within the fair buffer, excluding those that were also within the excellent and good buffers, and a score of 0 for all remaining centroids not within any buffer. Since we were performing these processes for 10 different cities, a tool was created using ArcMap's model builder in order to replicate these results. The details of the tool and its detailed workflow can be found in the Appendix D under Separate by Distance.

Next, we calculated the median mental health and physical health value from the city's 500 cities file which included columns for physical and mental health. We then added three more fields to our centroids layer. One field to hold our score given for mental health, one to hold our score for physical health, and one to hold the combined score. We used ArcMap's Select By Attribute (Data Management) to select centroids with a mental health value above the calculated mental health median and assigned them a mental health score of 1.5 to represent good mental health. Then, we selected centroids with a mental health value equal to or below the calculated mental health median and assigned them a mental health score of 0 to represent poor mental health. In the same manner, centroids with a physical health value above the physical health median were assigned a score of 1 to represent good physical health, and those equal to or below the median a score of 0 to represent poor physical health. Finally, these scores were combined in the third added column to represent each census block's overall health score. Note that we weighted mental health more heavily than physical health because the summary of our literature review pointed to green space having more of an impact on mental health than physical health, and we wanted our data to reflect that. Again, a tool was created in order to replicate these processes. The details of the tool and its detailed workflow can be found in the Appendix D under Separate by Health.

After using the tools created for this research, we were left with two datasets. One contained data for the health scores, the other contained data for the distance scores. In order to combine these metrics into a single score, ArcMap's Spatial Join (Analysis) tool was used on the two datasets, and a combined score was calculated from the result.

For analysis, we used ArcMap's Spatial Autocorrelation (Global Moran's I) (Spatial Statistics) tool to examine clustering on the global level, and ArcMap's Hot Spot Analysis (Getis-Ord Gi) (Spatial Statistics) tool to examine clustering on the local level (ESRI 2018a, np). We first performed some analysis using these tools on just the health scores we had assigned to the centroids, to see if there was any clustering of good and/or poor health within the city. If clusters existed at the health level, it was likely that there were factors influencing the health in different areas of the city. Then, we proceeded by combining our two scores and running the global and local cluster analysis on the combined metric to observe cluster patterns. As a final step, we used the NumPy/SciPy package for python and used the linregress function to find the Pearson product-moment correlation coefficients and P-Values for our distance scores and our health scores in order to determine if the two variables were correlated (SciPy Community 2018, np).

Results

What we discovered was that for every city we analysed, there was clustering for areas with good and poor health. The Spatial Autocorrelation showed that there was less than a 1% chance that the distribution of our health scores was random, as the Z-Score was well above 2.58 and the P-Value was below 0.01 for every city (ESRI, 2018b, np; Lentz 2009, 1-9). The Moran's Index showed some cities had more clustering than others. A positive Moran's index means that high value centroids attract each other, and the closer the index is to positive one, the stronger the clustering. See Appendix I for detailed Moran's I results for each city, including the calculated mean health score and mean distance score for reference. After performing the hotspot analysis, it was possible to visually estimate clustering of the hot spots and cold spots by overlaying them with the buffers created earlier. Further examination of the resulting hot spot analysis tables revealed that on a local level, the majority of the census tracts showed at least a 90% confidence interval that the clustering was not random. See Appendix E for detailed maps of the health metric, and Appendix F for detailed maps of health hot spots within each city. We then proceeded to run the same tests on our combined metric for distance and health and found similar results. Moran's I spatial autocorrelation determined a less than 1% chance that the distribution of our combined scores were random, and the local level hot spot analysis showed the majority of the census tracts had at least a 90% confidence interval that the clustering was not random. See Appendix I for the Moran's I results for the combined metric, Appendix G for detailed maps of the combined health and distance metrics, and Appendix H for detailed maps of the combined health and distance metrics' hot spot analysis. When calculating the Pearson product-moment correlation coefficients, we expected to see positive correlations coefficients, as our research pointed to our health score increasing as our distance to parks score increased (Laerd Statistics 2018, 1; Lentz 2009, 1-9). After running the tests, we found the correlation to be lower and much more inconsistent than we expected, with values for the Pearson Correlation Coefficient ranging from -0.254 in Washington DC to 0.3411 in Oklahoma City OK. Since the Pearson Correlation Coefficient ranges from one, representing perfect positive correlation, to negative one, representing perfect negative correlation, these results seem to indicate that while there may be some small correlation in some cities, our methods did not produce any correlation of statistical significance. Also, only Las Vegas, Oklahoma City, Nashville, and Washington had correlation coefficients that could actually be considered accurate, as they had a P-Value below 0.05 (Retherford, RD and Choe, MK 1993, 17-18). See the correlation table in Appendix J for the Pearson Correlation Coefficient in for each city.

While we did not detect correlation between distance to parks and health from our generated variables, we did show that there is definite clustering in terms of health and the combined metric of health and distance within some cities. This warrants further research into the subject, looking at different variables that may affect health, pre-existing conditions in each city such as historical health hazards from industrial cities that may influence clustering, and different scoring metrics then the ones used in this research. In future studies for this subject, we would recommend taking a closer look at demographics within each census tract as a possible contributing factor towards the clustering of health. Our research suggests examining income, race, education, and occupation (USDA 2018, 17-18). We would also recommend calculating the distance from each census tract to the closest park, rather than assigning a score based on a buffer zone for a more accurate variable representation, and ensuring that the selection of parks included in this analysis must include most properties of green space, and exclude features such as playground and industrial parks. Likewise, we suggest generating a new health metric from multiple variables and sources, as our health scores were mostly self-reported, which could have lead to inaccuracies. We hope the results of our research provide insight into this subject for future studies to come.

References

1. Rohit Adi, M.D. Mary Jo Pagel, M.D. Julian Whitaker, M.D. Anthony P. Urbanek, M.D. (2018). The Real Crisis in Mental Health Today. *Citizens Commission on Human Rights.* https://www.cchr.org/cchr-reports/the-real-crisis/introduction.html, Last accessed 7 March 2019.
2. United Nations (2018). 68% of the world population projected to live in urban areas by 2050, says UN. *Department of Economic and Social Affairs*. https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html, Last accessed 7 March 2019.
3. Hyvönen, K., Törnroos, K., Salonen, K., Korpela, K., Feldt, T., & Kinnunen, U. (2018). Profiles of Nature Exposure and Outdoor Activities Associated With Occupational Well-Being Among Employees. *Frontiers in psychology*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5968374/, Last accessed 7 March 2019.
4. Bjørnstad, S., Patil, G.G., & Raanaas, R.K. (2015). Nature contact and organizational support during office working hours: Benefits relating to stress reduction, subjective health complaints, and sick leave. https://www.ncbi.nlm.nih.gov/pubmed/26684700, Last accessed 7 March 2019.
5. Song, C., Ikei, H., & Miyazaki, Y. (2016). Physiological Effects of Nature Therapy: A Review of the Research in Japan. *International journal of environmental research and public health*, *13*(8). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4997467/, Last accessed 7 March 2019.
6. Annerstedt, M., & Währborg, P. (2011). Nature-assisted therapy: systematic review of controlled and observational studies. *Scandinavian journal of public health, 39 4*, 371-88. https://journals.sagepub.com/doi/abs/10.1177/1403494810396400, Last accessed 7 March 2019.
7. Matilda Annerstedt et. al (2013). Inducing physiological stress recovery with sounds of nature in a virtual reality forest. *Physiology and Behavior.* https://www.ncbi.nlm.nih.gov/pubmed/23688947, Last accessed 7 March 2019.
8. U.S. Department of Agriculture (USDA), Forest Service. (2018). Urban nature for human health and well-being: a research summary for communicating the health benefits of urban trees and green space. https://www.fs.fed.us/sites/default/files/fs_media/fs_document/urbannatureforhumanhealthandwellbeing_508_01_30_18.pdf, Last accessed 7 March 2019.
9. Patrik Grahn, Ulrika A. Stigsdotter (2003). Landscape planning and stress, *Urban Forestry & Urban Greening,* Volume 2 https://www.sciencedirect.com/science/article/pii/S1618866704700199, Last accessed 7 March 2019.
10. Beyer, Kirsten & Kaltenbach, Andrea & Szabo, Aniko & Bogar, Sandra & Javier Nieto, F & Malecki, Kristen. (2014). Exposure to Neighborhood Green Space and Mental Health: Evidence from the Survey of the Health of Wisconsin. International journal of environmental research and public health. https://www.mdpi.com/1660-4601/11/3/3453, Last accessed 7 March 2019.
11. D. Nutsford, A.L. Pearson, S. Kingham (2013) An ecological study investigating the association between access to urban green space and mental health, Journal of Public

Health. https://www.sciencedirect.com/science/article/pii/S0033350613002862, Last accessed 7 March 2019.

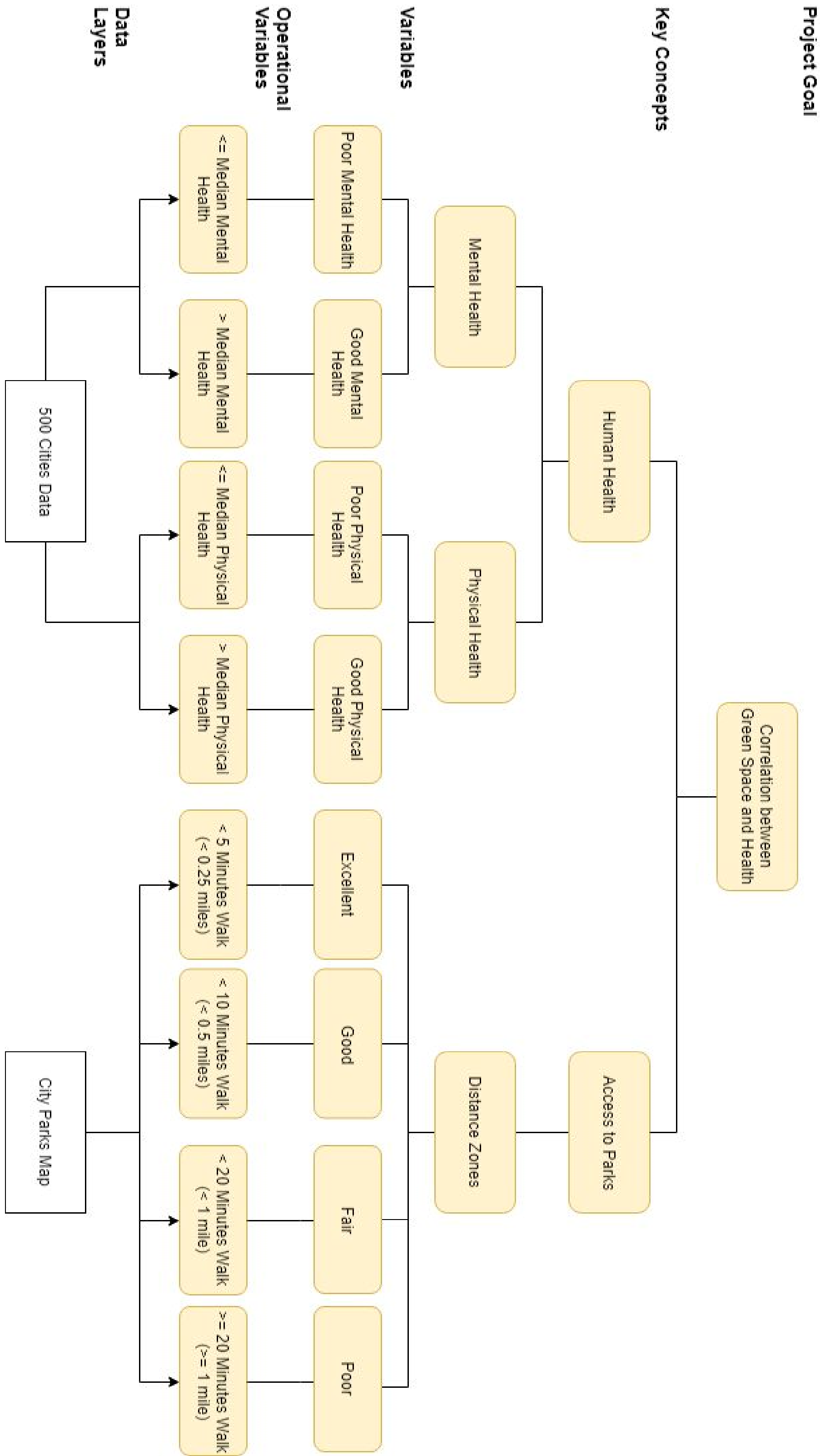
12. Blanck, H.M.; Allen, D.; Bashir, Z. [and others]. (2012). Let's go to the park today: the role of parks in obesity prevention and improving the public's health. Childhood Obesity. https://pdfs.semanticscholar.org/dac9/efa3531f7e6fbbf5617cdd227d1f0957f083.pdf, Last accessed 7 March 2019.
13. Yousefian Hansen, A.; Hartley, D. (2015). Promoting active living in rural communities. San Diego, CA: Active Living Research. https://activelivingresearch.org/promoting-active-living-rural-communities, Last accessed 7 March 2019.
14. Mitchell, R.; Popham, F. (2008). Effect of exposure to natural environment on health inequalities: an observational population study. https://www.sciencedirect.com/science/article/pii/S014067360861689X, Last accessed 8 March 2019.
15. Jenning, V.; Johnson Gaither, C. (2015). Approaching environmental health disparities and green space: an ecosystem services perspective. International Journal of Environmental Research and Public Health. https://www.mdpi.com/1660-4601/12/2/1952, Last accessed 8 March 2019.
16. Leah Binkovitz (2018). A study of city Park Scores reveals inequities tied to a city's racial and ethnic composition and other factors. *Rice Kinder Institute for Urban Region*. https://kinder.rice.edu/2018/06/24/how-city-demographics-affect-parks, Last accessed 8 March 2019.
17. Alessandro Rigolon, Matthew Browning, Viniece Jennings (2018). Inequities in the quality of urban park systems: An environmental justice investigation of cities in the United States. *Landscape and Urban Planning*. https://www.sciencedirect.com/science/article/pii/S0169204618304316, Last accessed 8 March 2019.
18. Lauren C. Abercrombie, James F. Sallis, Terry L. Conway, Lawrence D. Frank, Brian E. Saelens, James E. Chapman (2008). Income and Racial Disparities in Access to Public Parks and Private Recreation Facilities, *American Journal of Preventive Medicine.* https://www.sciencedirect.com/science/article/pii/S0749379707006502, Last accessed 8 March 2019.
19. McKenzie, T. L., Moody, J. S., Carlson, J. A., Lopez, N. V., & Elder, J. P. (2013). Neighborhood Income Matters: Disparities in Community Recreation Facilities, Amenities, and Programs. *Journal of park and recreation administration*, *31*(4) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4082954/, Last accessed 8 March 2019.
20. Sister, C.; Wolch, J.; Wilson, J. (2010). Got green? Addressing environmental justice in park provision. GeoJournal. https://link.springer.com/article/10.1007/s10708-009-9303-8, Last accessed 8 March 2019.
21. El-Geneidy, Ahmed & Levinson, David & Boisjoly, Geneviève & Verbich, David & Loong, Charis & Diab, Ehab. (2016). The cost of equity: Assessing transit accessibility and social disparity using total travel cost. Transportation Research Part A Policy and Practice. 91. https://www.sciencedirect.com/science/article/pii/S0965856416305924?via%3Dihub, Last accessed 8 March 2019.

22. Peter Harnik and Abby Martin (2004). Close-to-Home Parks: A Half-Mile or Less *The Center for City Park Excellence, The Trust for Public Land.* https://parkscore.tpl.org/Methodology/TPL_10MinWalk.pdf, Last accessed 8 March 2019.
23. National Recreation and Park Association (2014). Safe Routes to Parks: Improving Access to Parks through Walkability. *Safe Routes to Parks*. https://www.nrpa.org/contentassets/f768428a39aa4035ae55b2aaff372617/park-access-report.pdf, Last accessed 8 March 2019.
24. Carolyn Bancroft, Spruha Joshi, Andrew Rundle, Malo Hutson, Catherine Chong, Christopher C. Weiss, Jeanine Genkinger, Kathryn Neckerman, Gina Lovasi. (2015) Association of proximity and density of parks and objectively measured physical activity in the United States: A systematic review. Social Science & Medicine,Volume 138. https://www.sciencedirect.com/science/article/pii/S0277953615003160, Last accessed 8 March 2019.
25. Nicholls S. Measuring the accessibility and equity of public parks (2001). A case study using GIS. *Managing Leisure.* http://staff.washington.edu/kwolf/Archive/Classes/ESRM304_SocSci/304%20Soc%20Sci%20Lab%20Articles/Nicholls_2001.pdf, Last accessed 8 March 2019.
26. *Moore LV, Diez Roux AV, Evenson KR, McGinn AP, Brines SJ.* (2008). Availability of recreational resources in minority and low socioeconomic status areas. *Prev Med.* https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2254179/, Last accessed 8 March 2019.
27. Manikandan S. Measures of central tendency: (2011)The mean. J Pharmacol Pharmacother. http://www.jpharmacol.com/article.asp?issn=0976-500X;year=2011;volume=2;issue=2;spage=140;epage=142;aulast=Manikanda, Last accessed 8 March 2019.
28. Gravetter FJ, Wallnau LB. (2000). Statistics for the behavioral sciences. 5th ed. Belmont: Wadsworth – Thomson Learning. https://craftx.org/sites/all/themes/craft_blue/pdf/Exploring_College_Textbook_Sample_2_p3.pdf, Last accessed 8 March 2019.
29. Dawson B, Trapp RG. Basic and Clinical Biostatistics (2004). 4th ed. New York: Mc-Graw Hill; https://accessmedicine.mhmedical.com/book.aspx?bookID=356, Last accessed 8 March 2019.
30. BASNET, B., APAN, A. & RAINE, S. Environmental Management (2001). Last accessed https://doi.org/10.1007/s002670010241, 8 March 2019.
31. Takano T., Nakamura K., Watanabe M. (2002). Urban residential environments and senior citizens' longevity in megacity areas: The importance of walkable green spaces. J. Epidemiol. Community Health. 56 https://jech.bmj.com/content/56/12/913, Last accessed 8 March 2019.
32. Barbara Boyle Torrey (2004). Urbanization: An Environmental Force to be Reckoned With. *Population Reference Bureau.* https://www.prb.org/urbanization-an-environmental-force-to-be-reckoned-with/, Last accessed 8 March 2019.

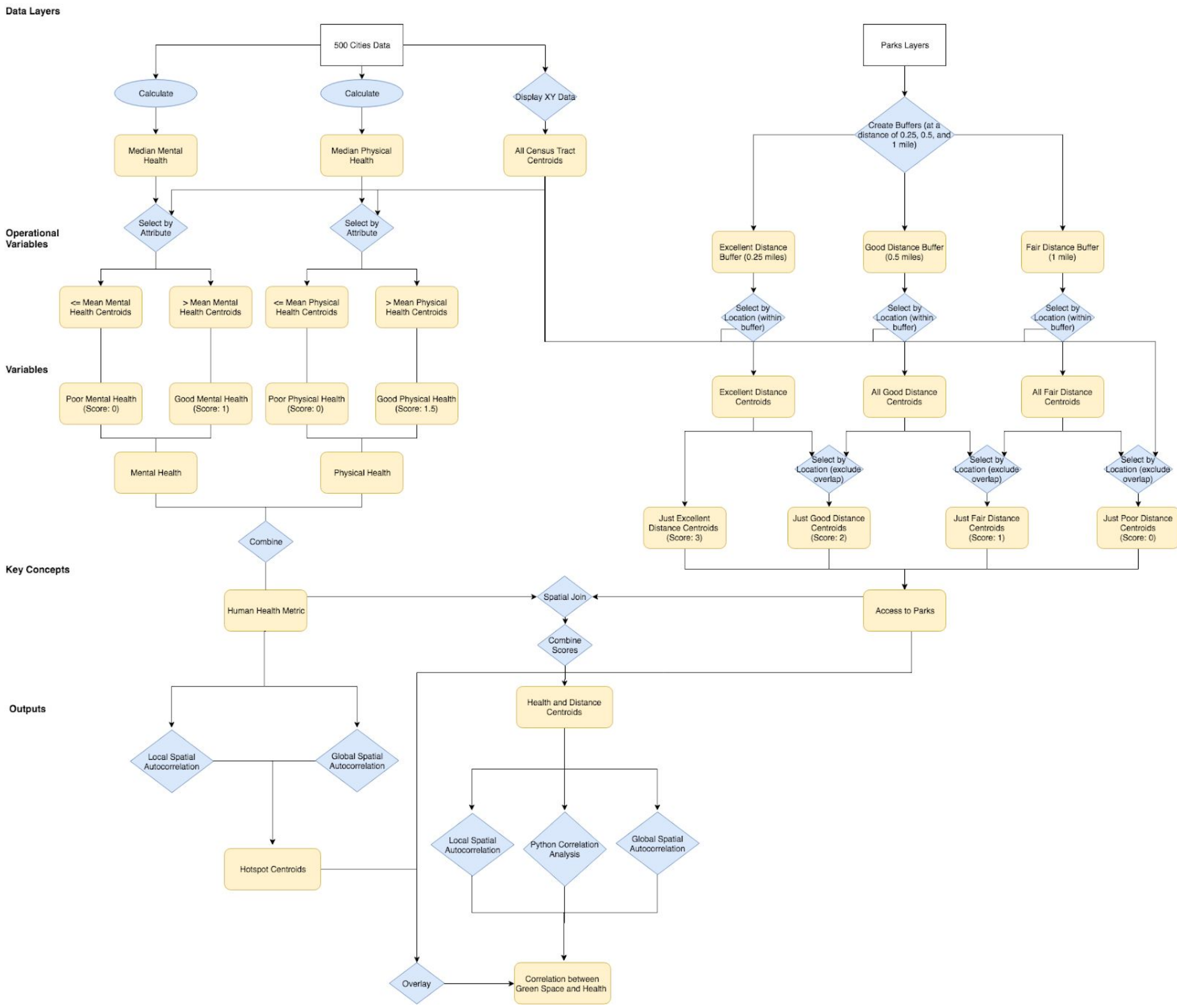
33. 500 Cities: Local Data for Better Health. Center for Disease Control and Prevention (2016). *U.S. Department of Health and Human Service.* https://www.cdc.gov/500cities/index.htm, Last accessed 8 March 2019
34. Koohsari, M. (2011). ACCESS TO PUBLIC OPEN SPACE: IS DISTRIBUTION EQUITABLE ACROSS DIFFERENT SOCIO-ECONOMIC AREAS. *Journal of Urban and Environmental Engineering, 5*(2). http://www.jstor.org/stable/26203358, Last accessed 7 March 2019.
35. University of Connecticut (nd). Digital Images and Georeferencing, *University of Connecticut Library* (np) http://magic.lib.uconn.edu/help/help_DigitalImagesandGeoreferencing.html Last accessed 21 April 2019.
36. Chang, K.T., (2015). Introduction to GIS (7th Edition).*McGraw Hill*, Ch5.
37. Mennis, J. (2019). Problems of Scale and Zoning. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2019 Edition), John P. Wilson (Ed.). (np) https://gistbok.ucgis.org/bok-topics/problems-scale-and-zoning Last accessed 24 April 2019.
38. Zhang, X., Lu, Hua L., Holt J. (2011). Modeling spatial accessibility to parks: a national study: Last accessed 24 April 2019. https://ij-healthgeographics.biomedcentral.com/articles/10.1186/1476-072X-10-31
39. U.S. Census Bureau: Annual Estimates of the Resident Population: April 1, 2010 to July 1, (2018) 2017, *Population Division.* (np) https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk, Last accessed 8 March 2019.
40. ESRI (2018a). “Spatial Autocorrelation (Global Moran's I).” *ArcGIS*, https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/spatial-autocorrelation.htm, Last accessed 28 April 2019.
41. ESRI (2018b). What is a z-score? What is a p-value? *ArcGIS*, https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/what-is-a-z-score-what-is-a-p-value.htm, Last accessed 28 April, 2019.
42. Laerd Statistics (2018). Pearson Product-Moment Correlation. *2018 Lund Research Ltd* https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php, Last accessed 28 April 2019.
43. SciPy (2018). Community Scipy.stats.linregress. *SciPy* https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.linregress.html, Last accessed 29 April 2019.
44. Jackson, J, (2018): Map Memphis Parks. *City of Memphis Enterprise GIS,* *http://memegis.maps.arcgis.com/home/item.html?id=9750a4dc8cf5462788a0d2c987f55682*, Last accessed 28 April 2019.

45. Retherford, RD, Choe, MK. (1993) Statistical methods for causal analysis. New York: John Wiley & Sons. https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118033135, Last accessed 29 April 2019.

Appendix A: Conceptualization Diagram

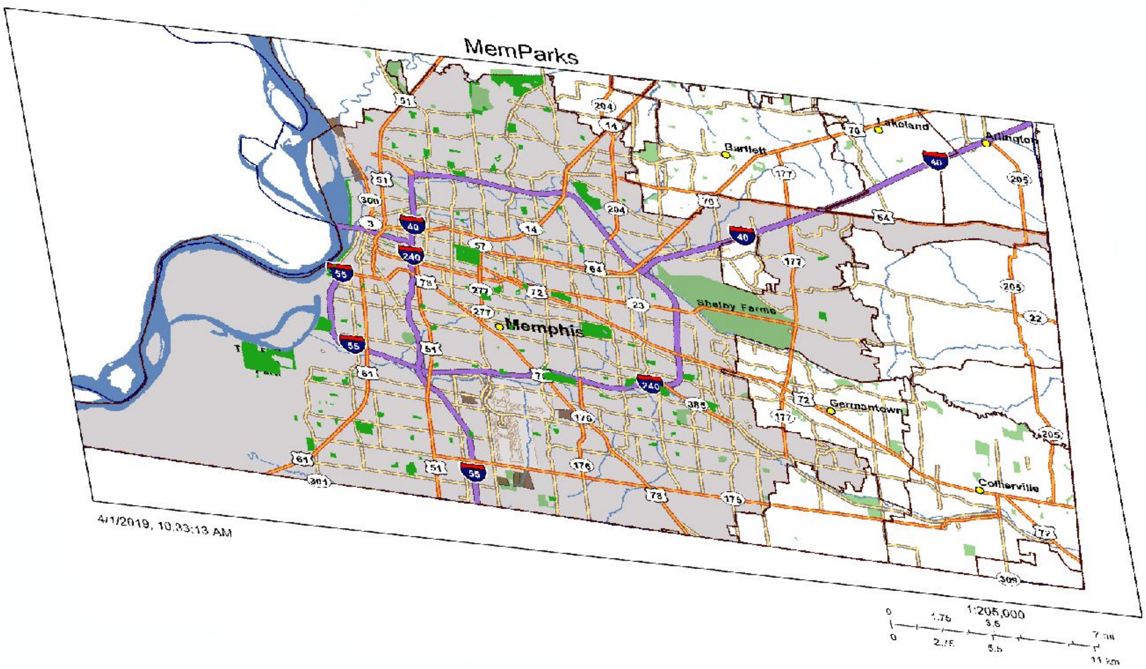


Appendix B: Implementation Diagram

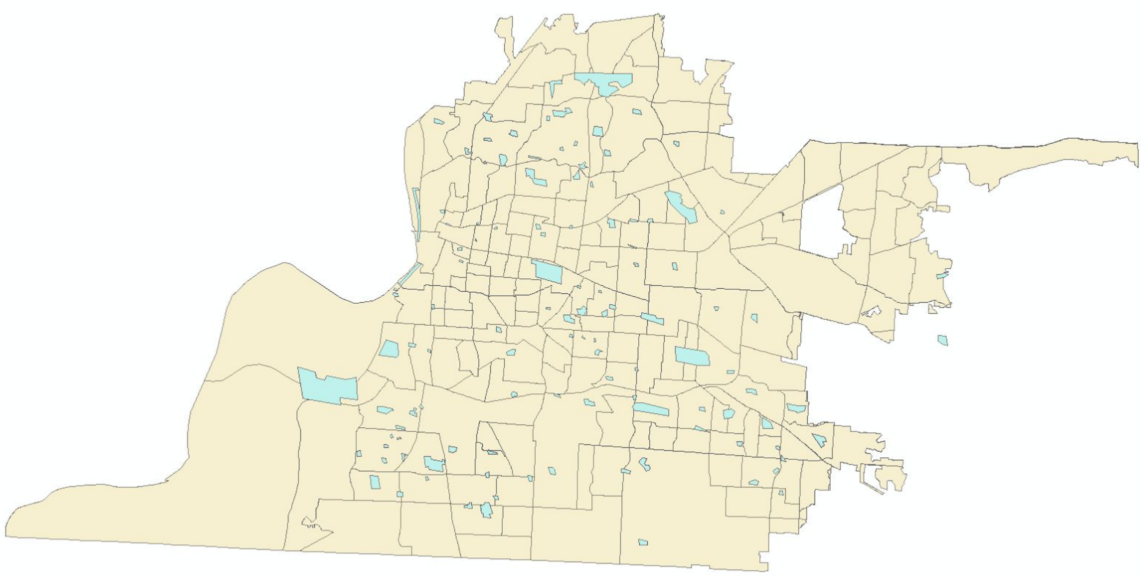


Appendix C: Original Data Layer (Memphis,TN)

Rectified Basemap

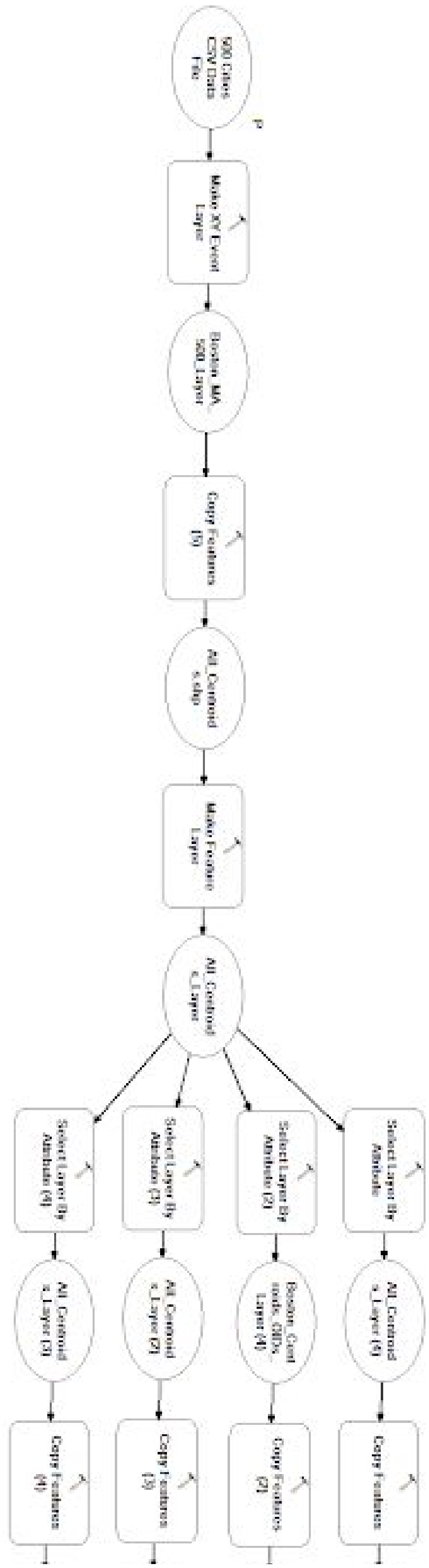


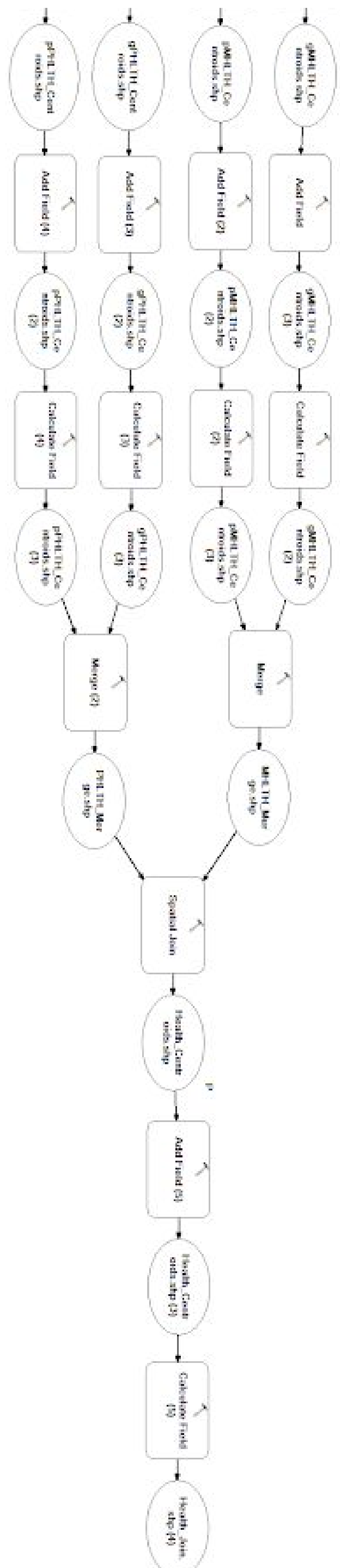
Digitized Parks Feature Class Overlayed with Census Tracts



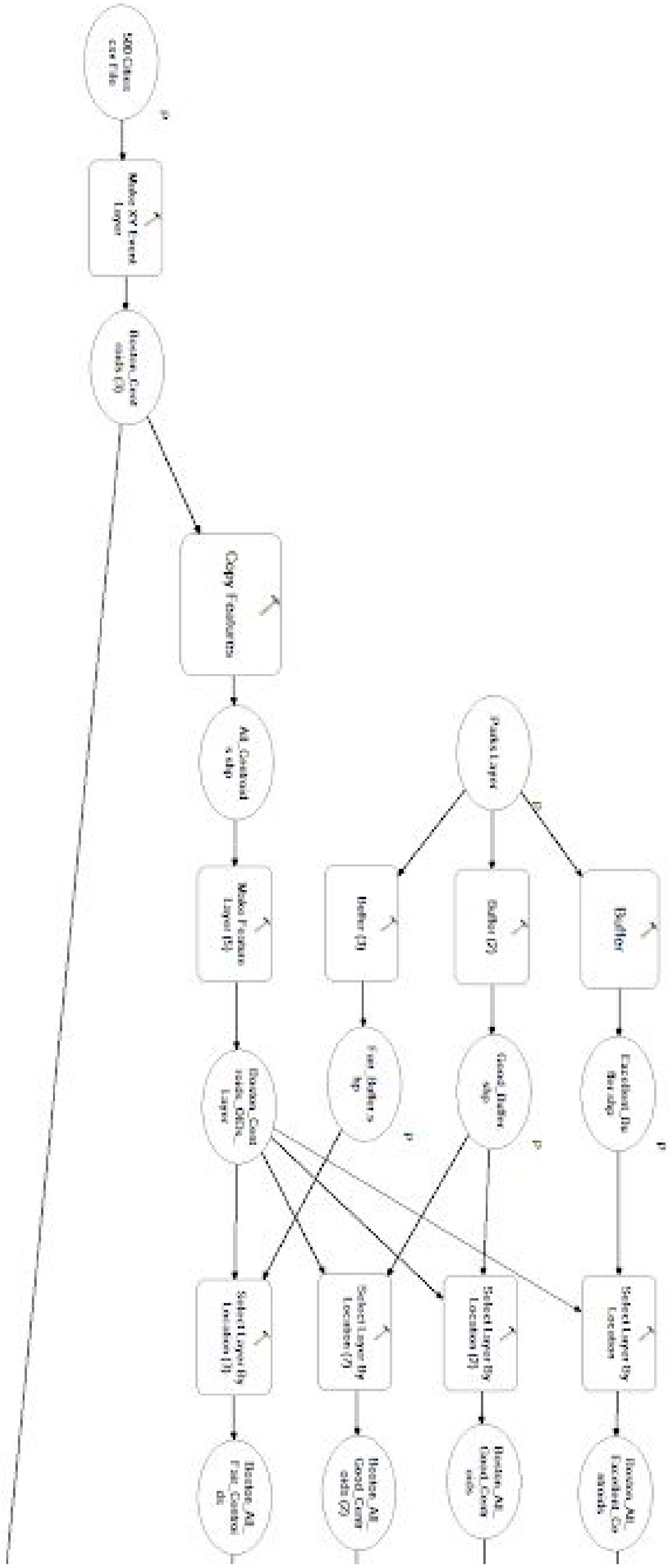
Appendix D: ArcMap Models and Detailed Workflow

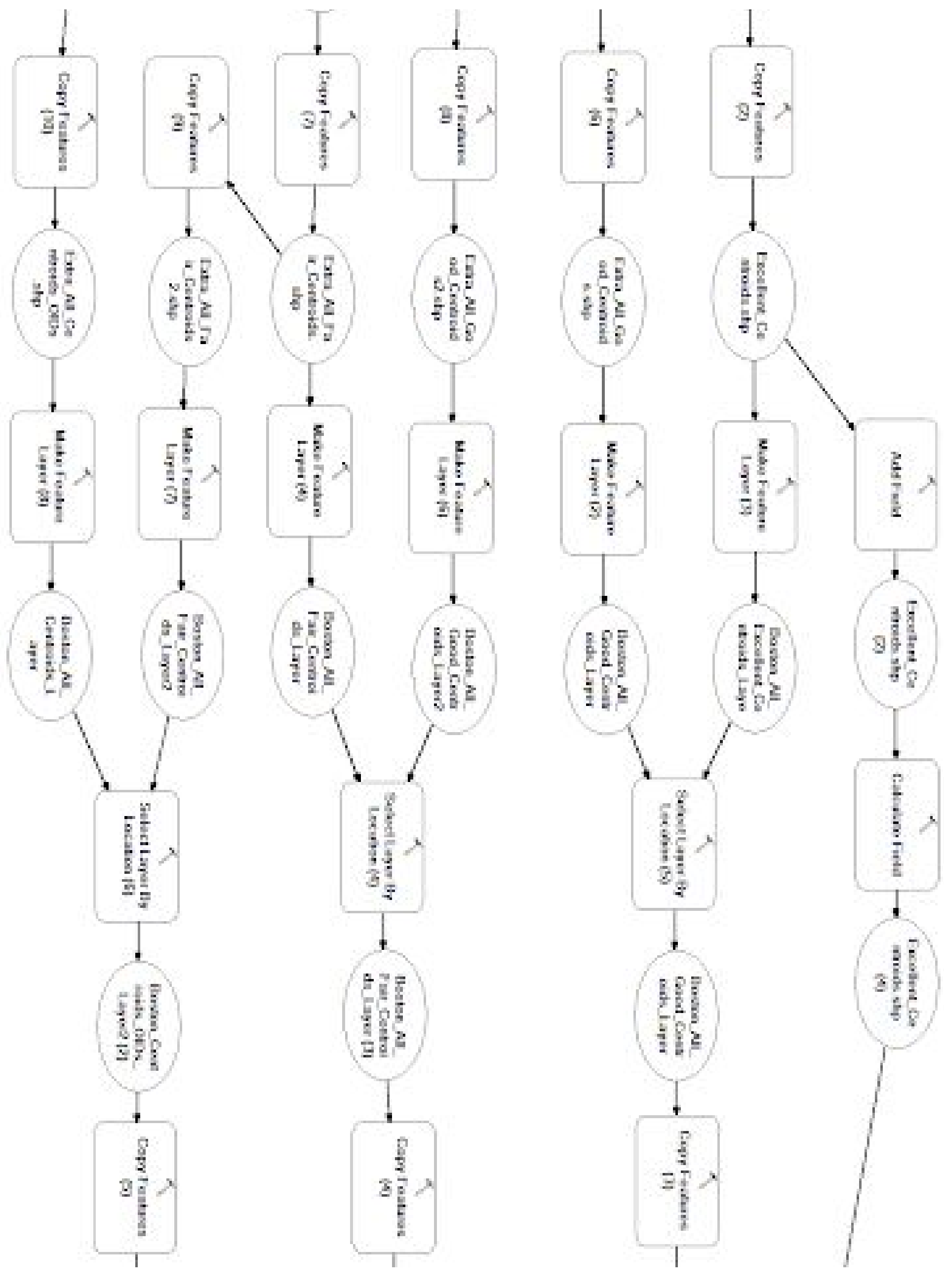
Separate by Health

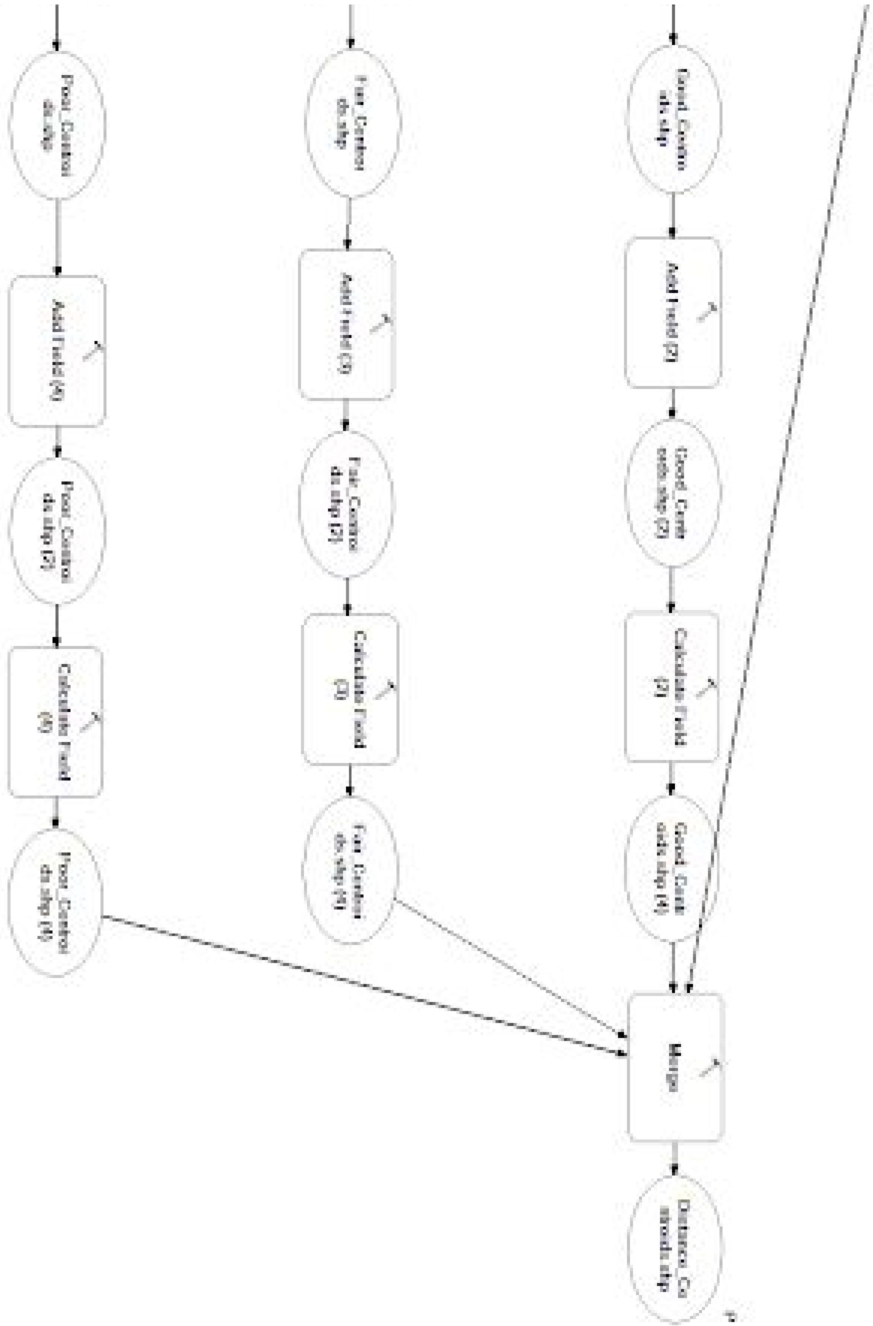



gMHLTH_Ce ntroids.shp
Add Field
gMHLTH_Ce ntroids.shp (3)
Calculate Field
gMHLTH_Ce ntroids.shp (2)
pMHLTH_Ce ntroids.shp
Add Field (2)
pMHLTH_Ce ntroids.shp (2)
Calculate Field (2)
pMHLTH_Ce ntroids.shp (3)
gPHLTH_Cent roids.shp
Add Field (3)
gPHLTH_Ce ntroids.shp (2)
Calculate Field (3)
gPHLTH_Ce ntroids.shp (3)
pPHLTH_Cent roids.shp
Add Field (4)
pPHLTH_Ce ntroids.shp (2)
Calculate Field (4)
pPHLTH_Ce ntroids.shp (3)
Merge
MHLTH_Mer ge.shp
Merge (2)
PHLTH_Mer ge.shp
Spatial Join
Health_Centr oids.shp
P
Add Field (5)
Health_Centr oids.shp (3)
Calculate Field (5)
Health_Join.shp (4)

Separate by Distance

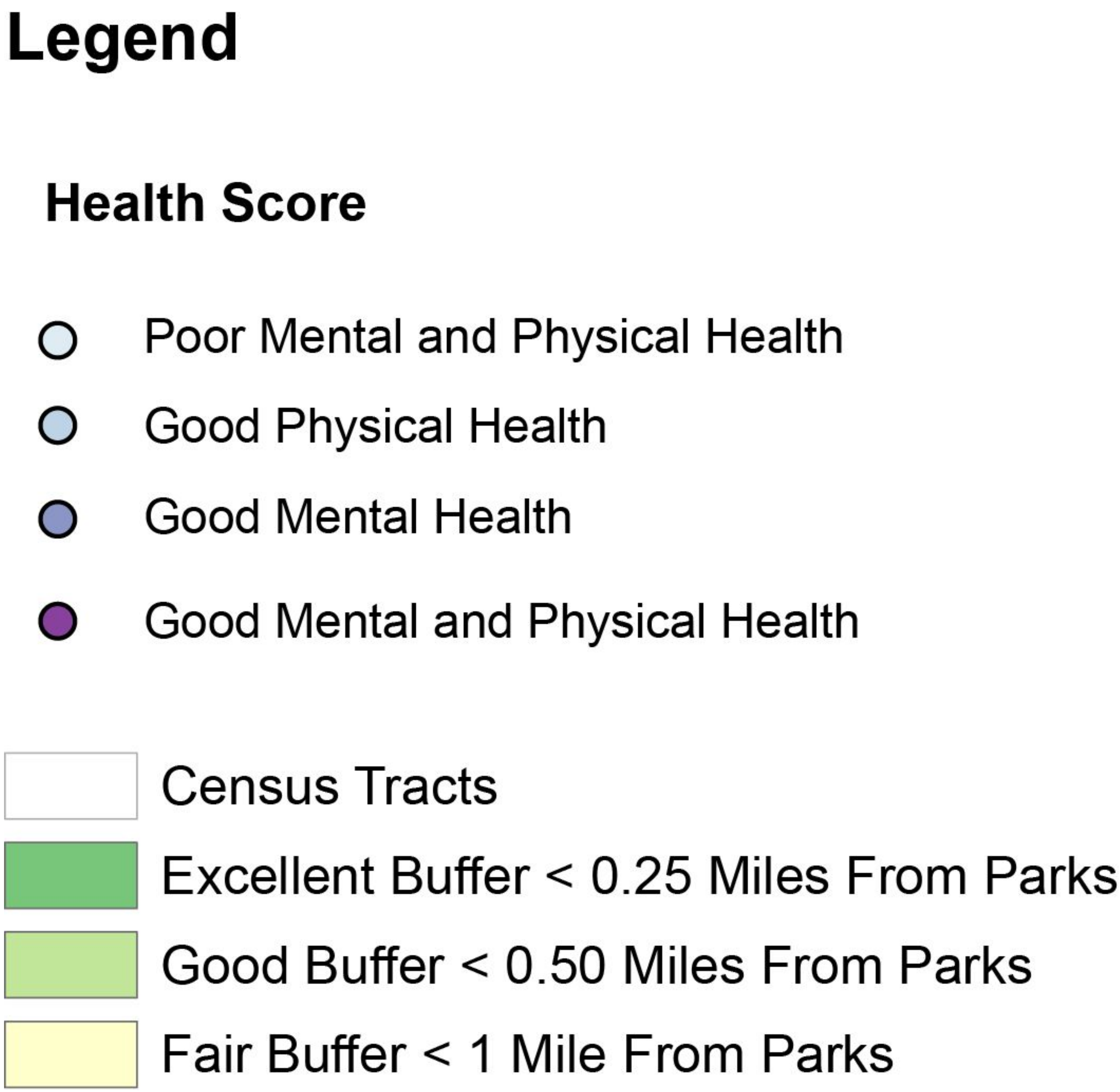


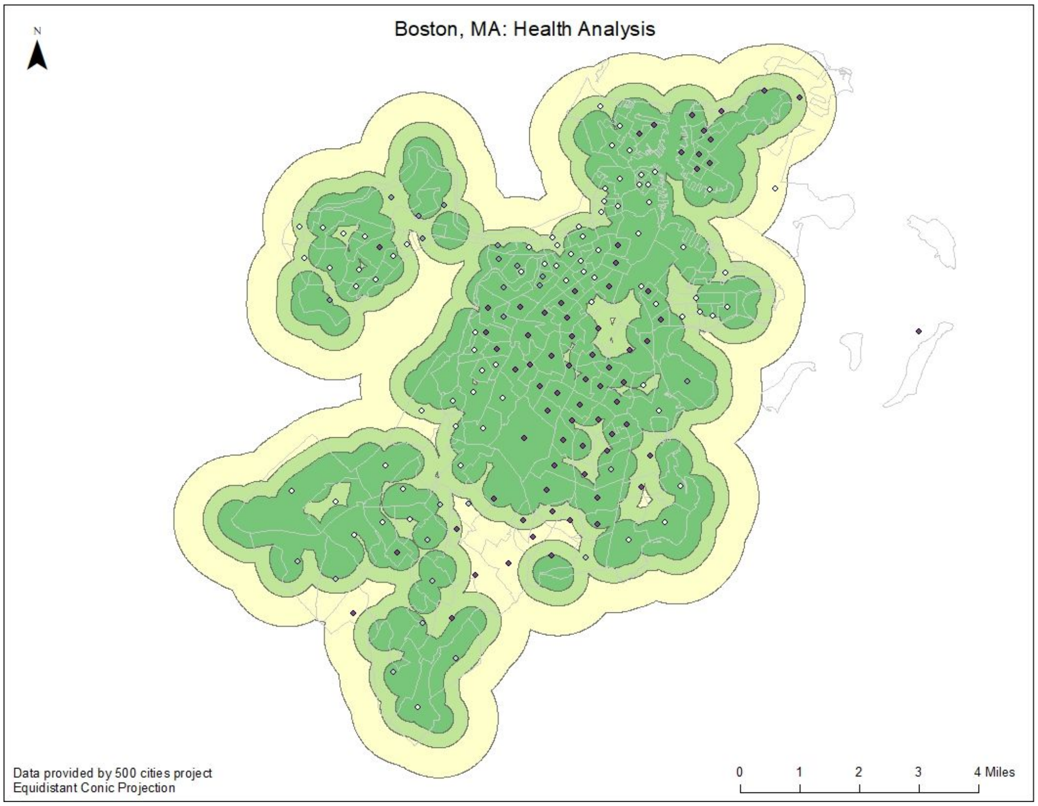


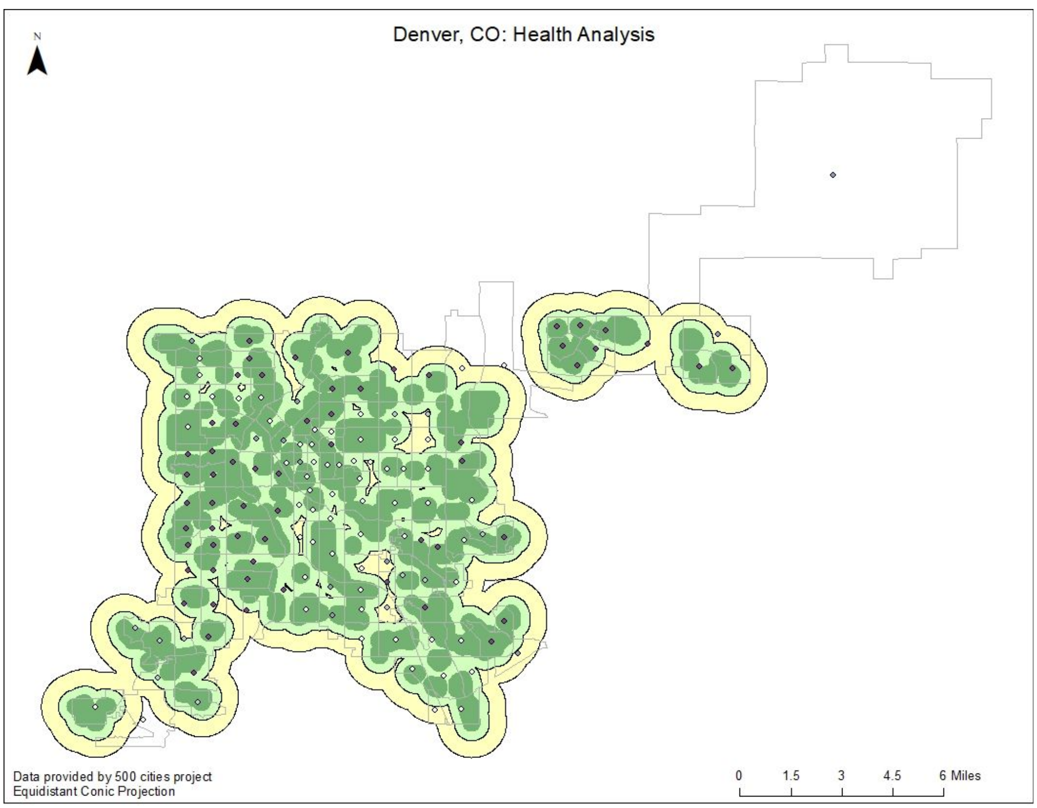

Good_Contro
ids.shp
Add Field (2)
Good_Centr
oids.shp (2)
Calculate Field
(2)
Good_Centr
oids.shp (4)
Merge
Distance_Co
ntroids.shp
Fair_Control
ds.shp
Add Field (3)
Fair_Control
ds.shp (2)
Calculate Field
(3)
Fair_Control
ds.shp (4)
Poor_Control
ds.shp
Add Field (4)
Poor_Control
ds.shp (2)
Calculate Field
(4)
Poor_Control
ds.shp (4)

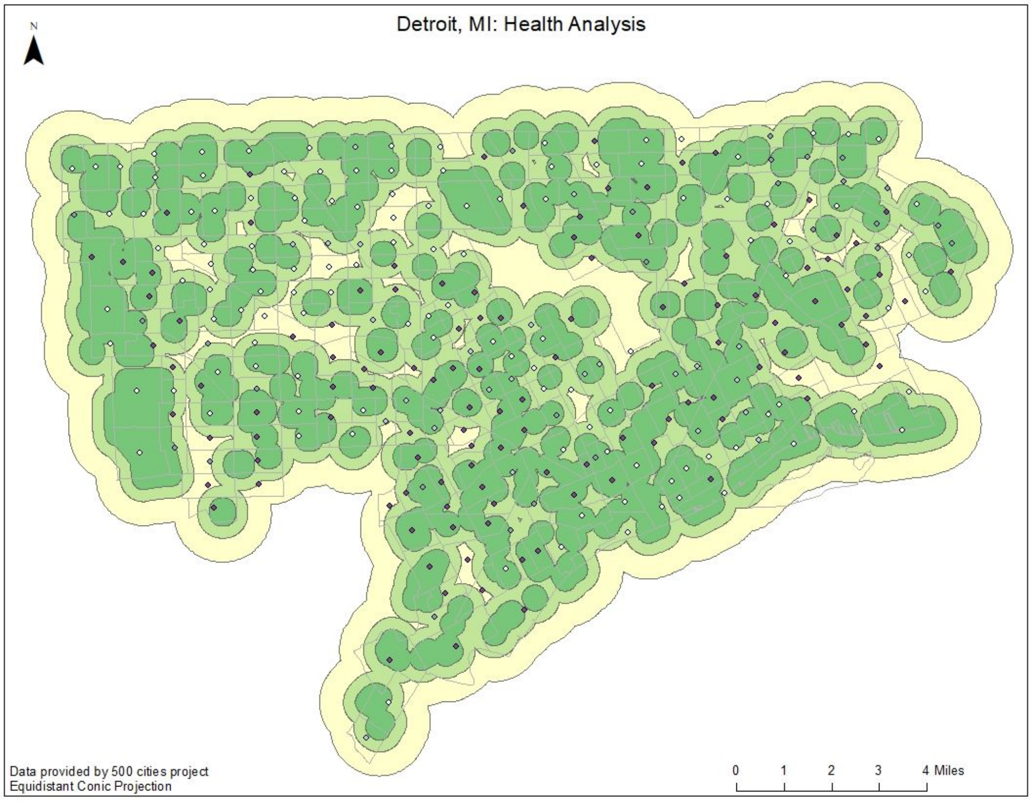
Appendix E: Health Analysis Maps

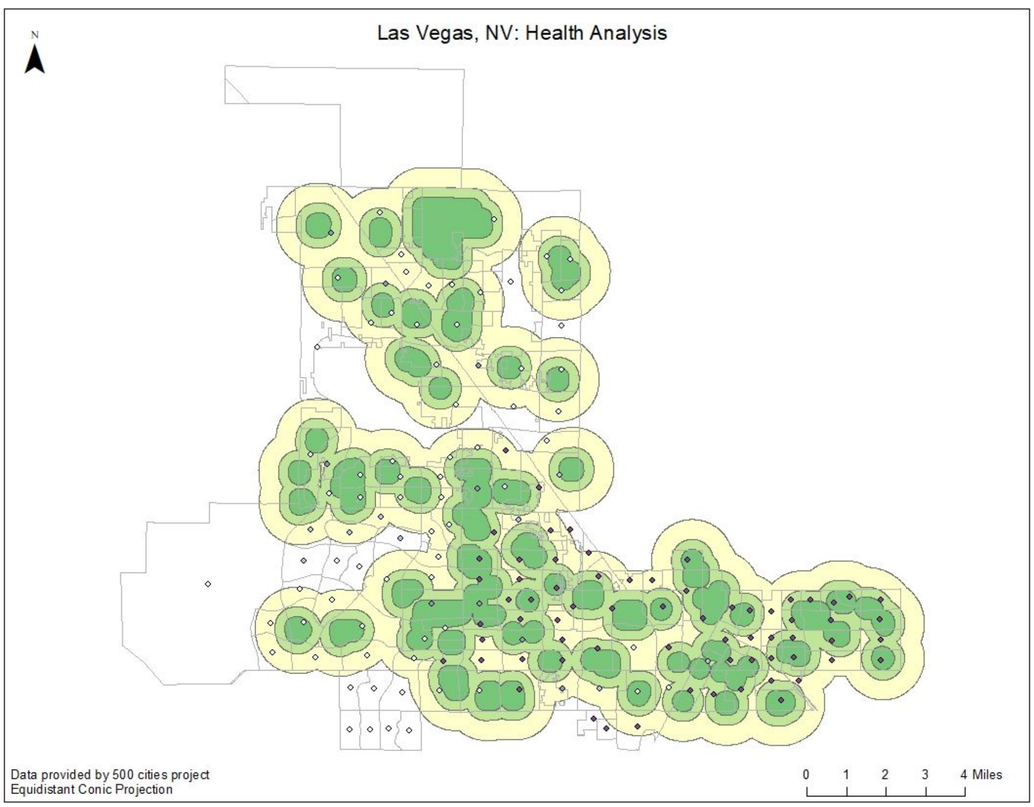
Legend applies to all maps in section

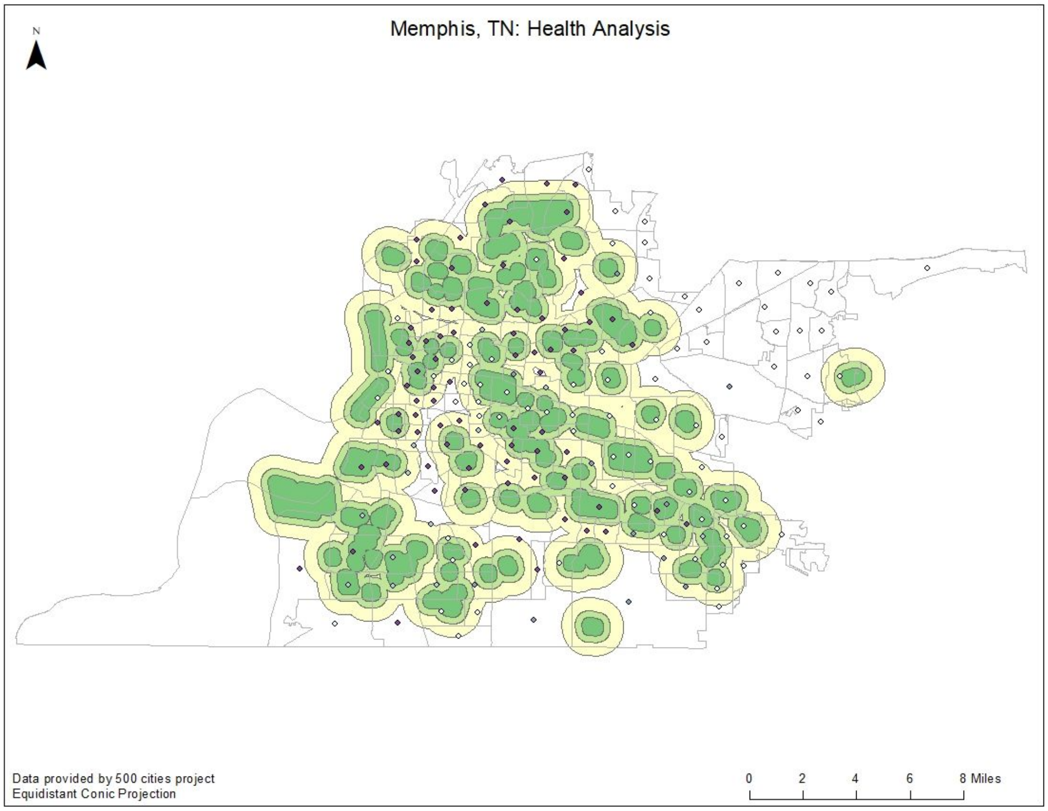


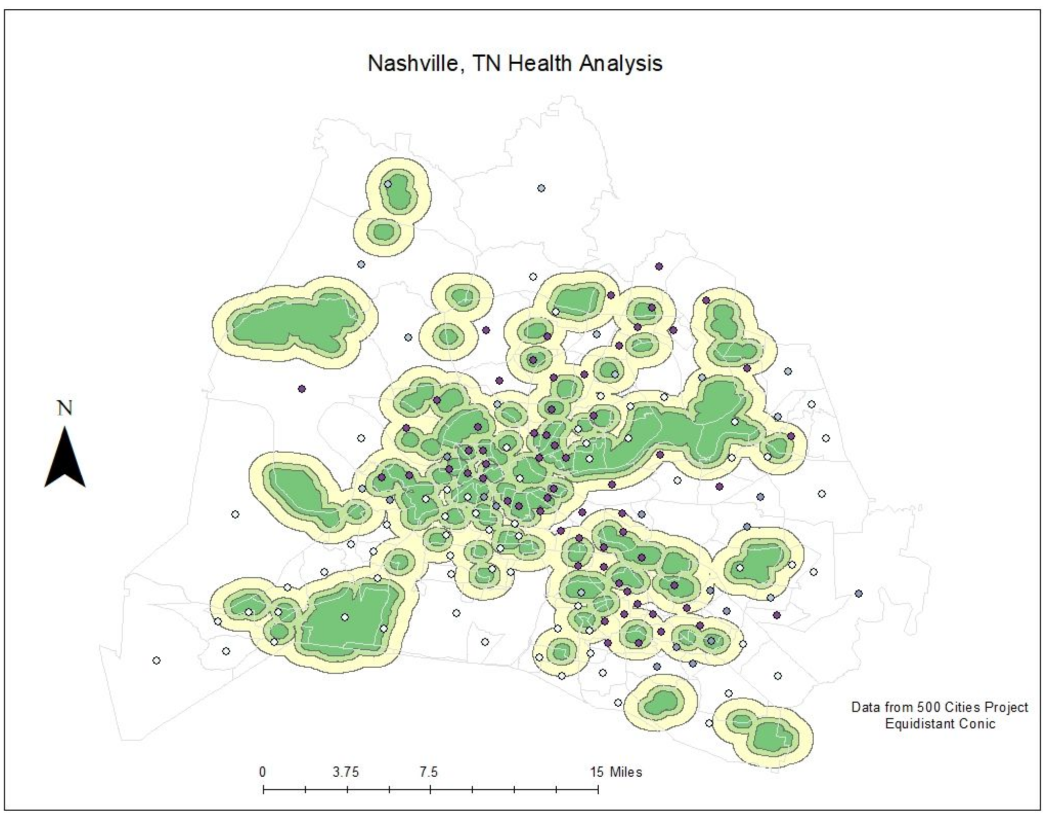

Boston, MA: Health Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0
1
2
3
4 Miles

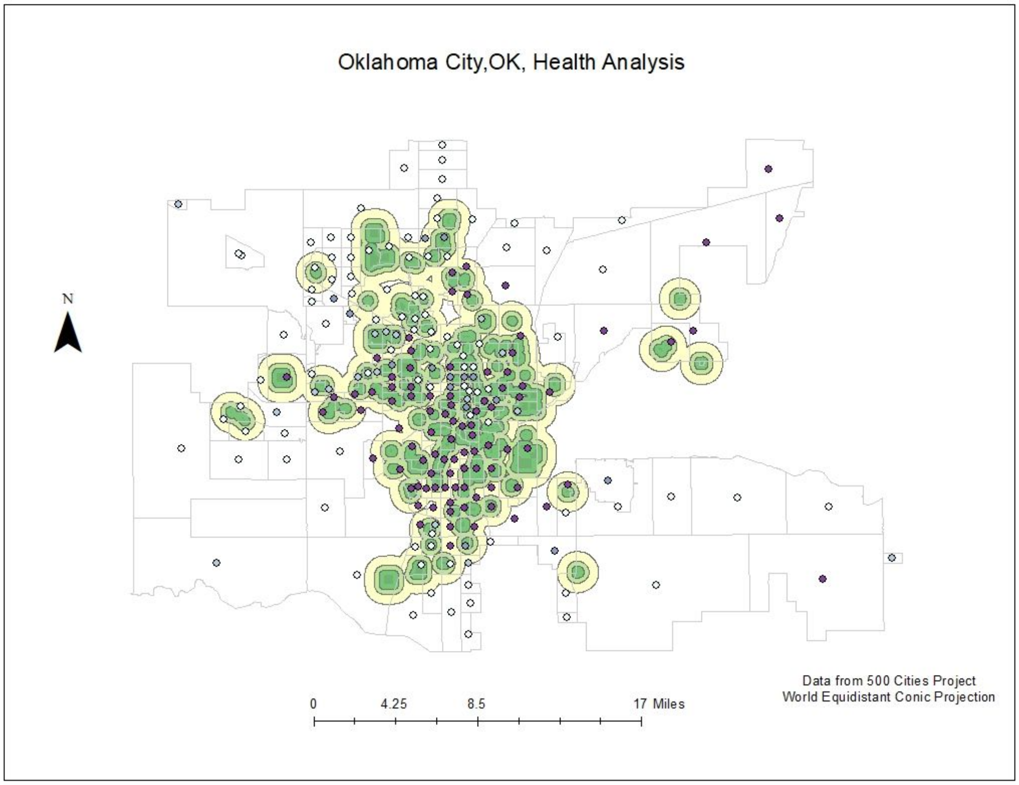

Denver, CO: Health Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0
1.5
3
4.5
6 Miles

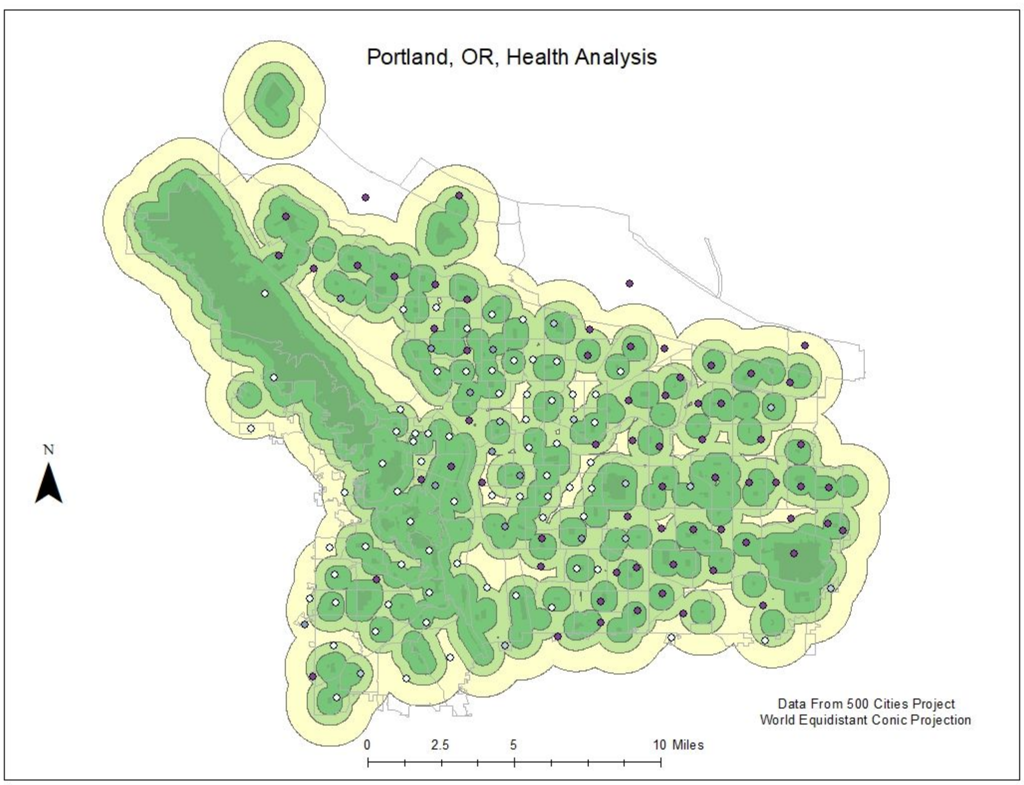

Detroit, MI: Health Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

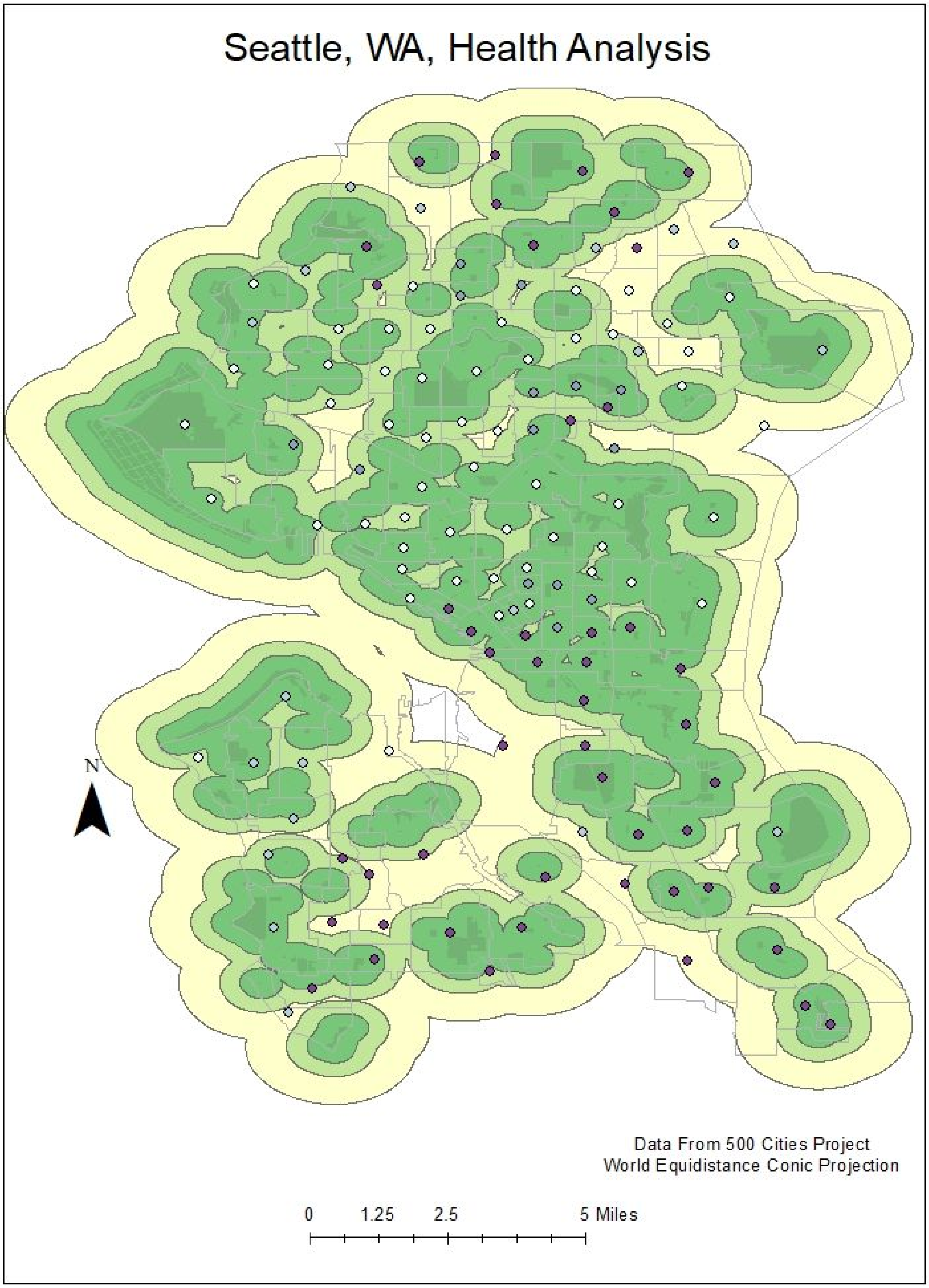

Las Vegas, NV: Health Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

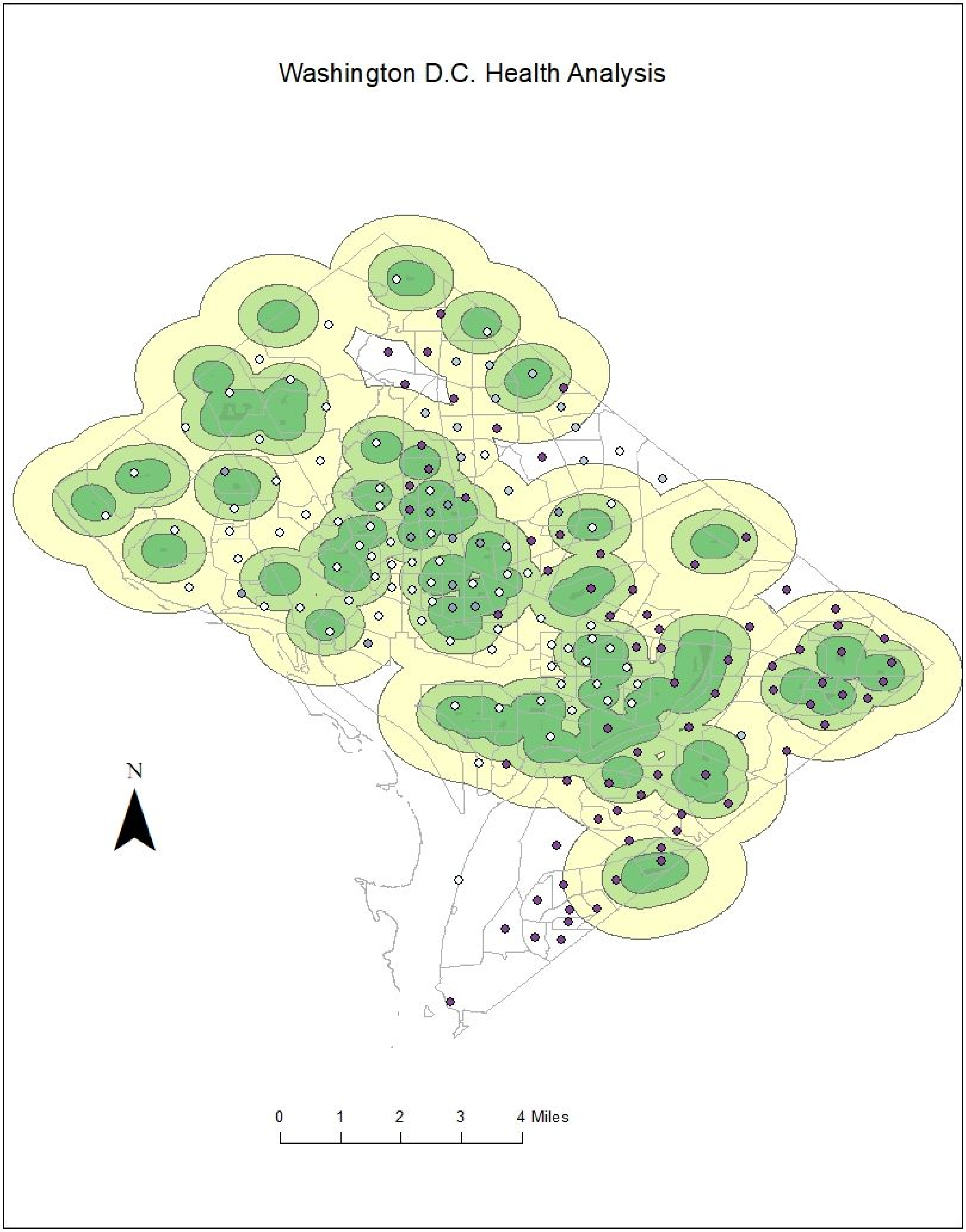

Memphis, TN: Health Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0
2
4
6
8 Miles


Nashville, TN Health Analysis
N
Data from 500 Cities Project
Equidistant Conic
0
3.75
7.5
15 Miles


Oklahoma City,OK, Health Analysis
N
Data from 500 Cities Project
World Equidistant Conic Projection
0
4.25
8.5
17 Miles


Portland, OR, Health Analysis
N
Data From 500 Cities Project
World Equidistant Conic Projection
0
2.5
5
10 Miles


Seattle, WA, Health Analysis
N
Data From 500 Cities Project
World Equidistance Conic Projection
0
1.25
2.5
5 Miles


Washington D.C. Health Analysis
N
0
1
2
3
4 Miles

Appendix F: Health Hot Spots Maps

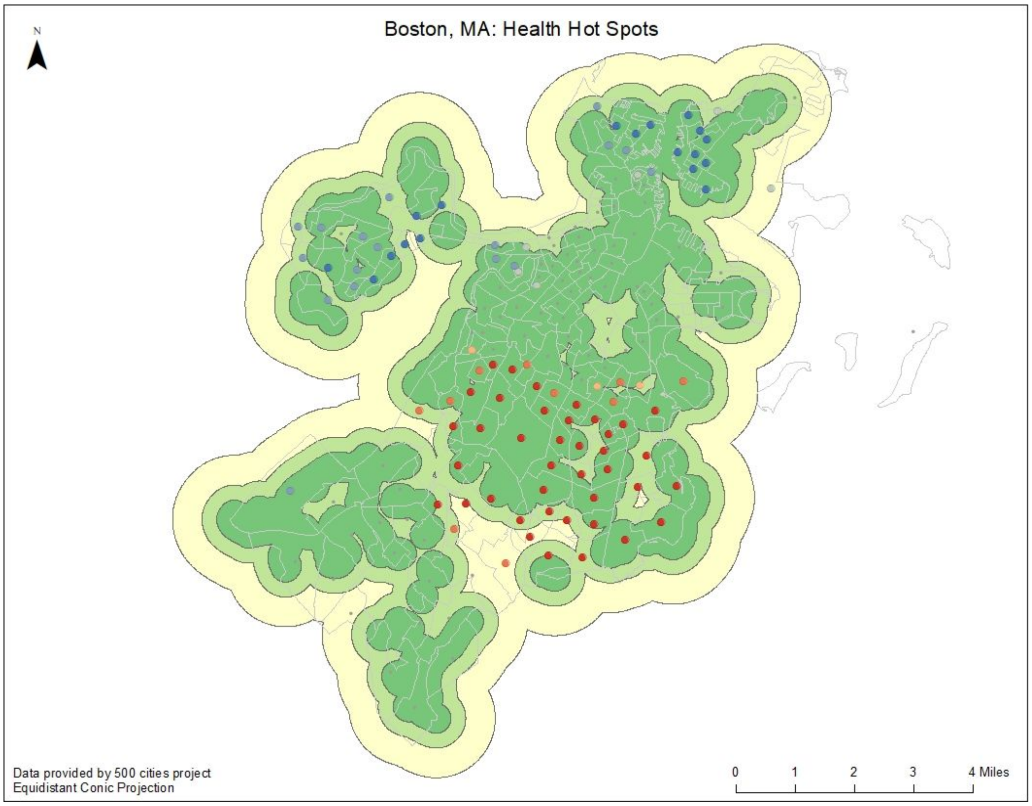
Legend applies to all maps in section

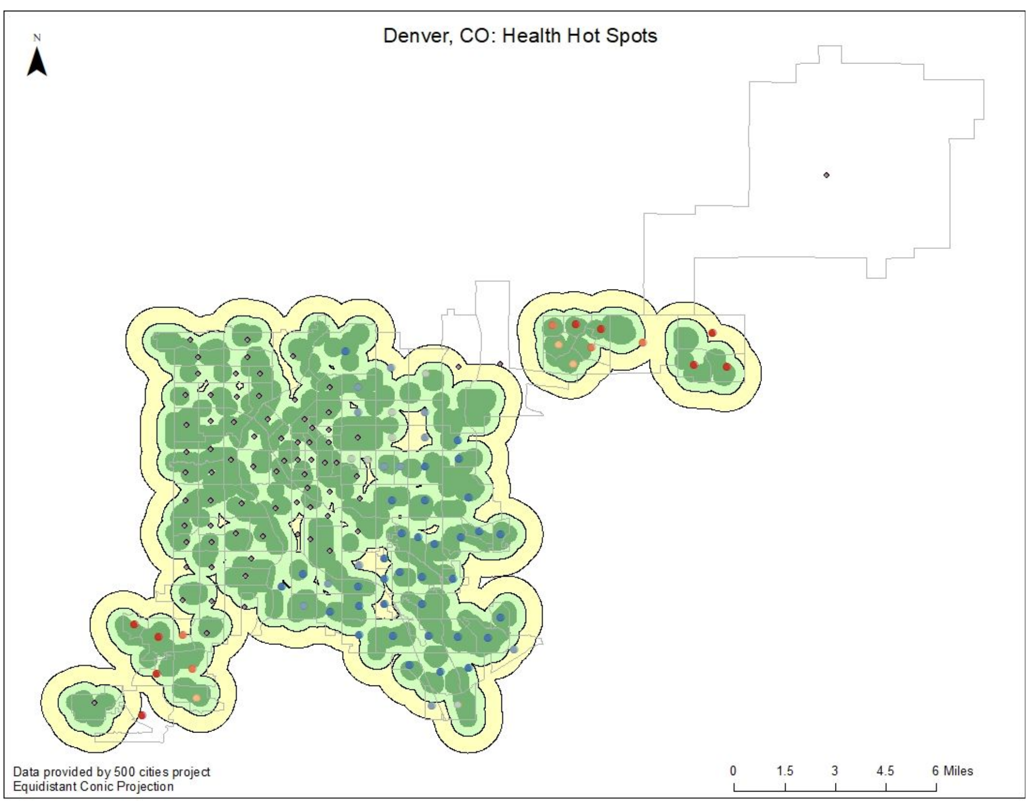
Legend

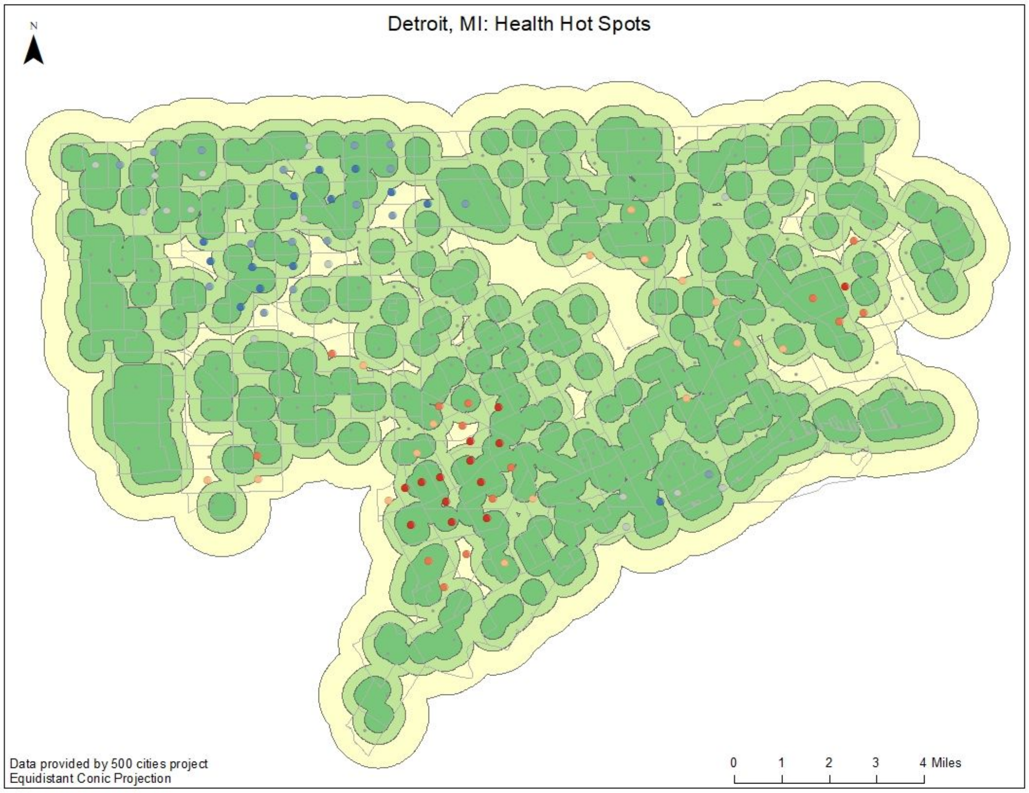
Hotspot Scores

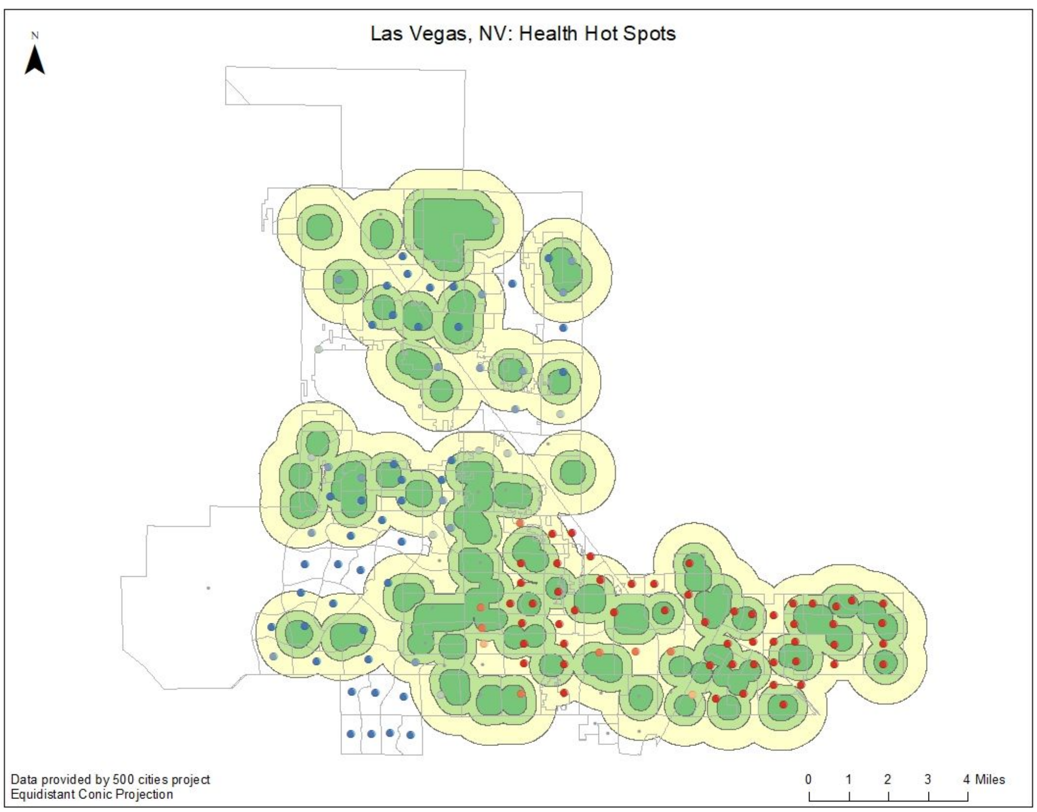
- Cold Spot - 99% Confidence
- Cold Spot - 95% Confidence
- Cold Spot - 90% Confidence
- Not Significant
- Hot Spot - 90% Confidence
- Hot Spot - 95% Confidence
- Hot Spot - 99% Confidence

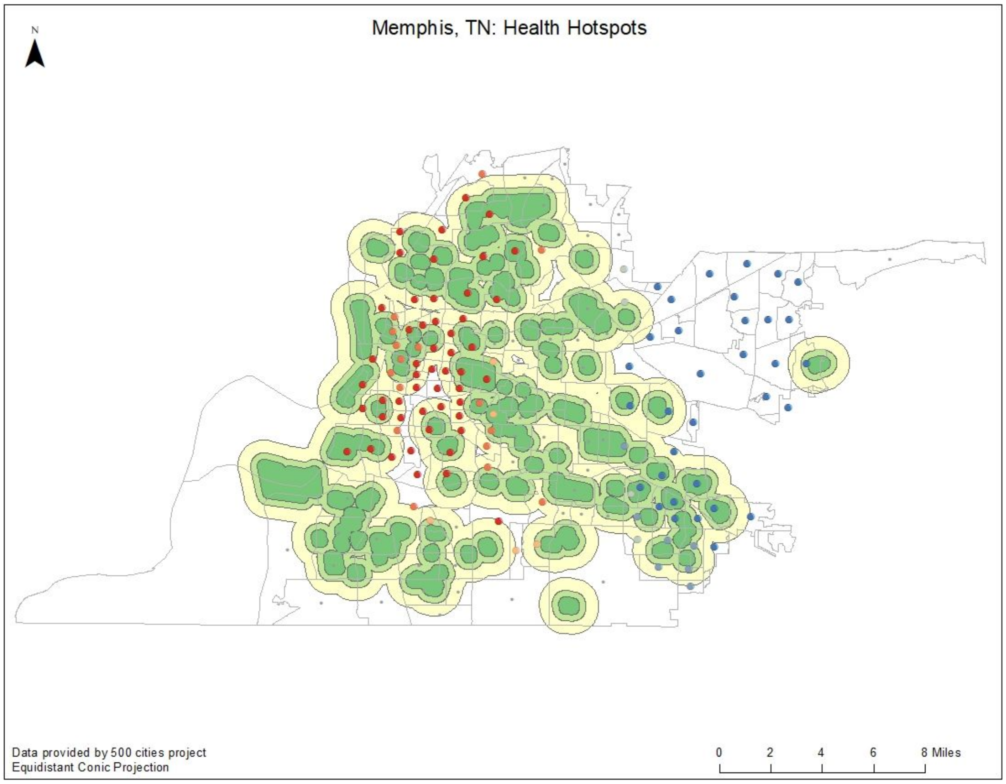
- Census Tracts
- Excellent Buffer < 0.25 Miles From Parks
- Good Buffer < 0.50 Miles From Parks
- Fair Buffer < 1 Mile From Parks

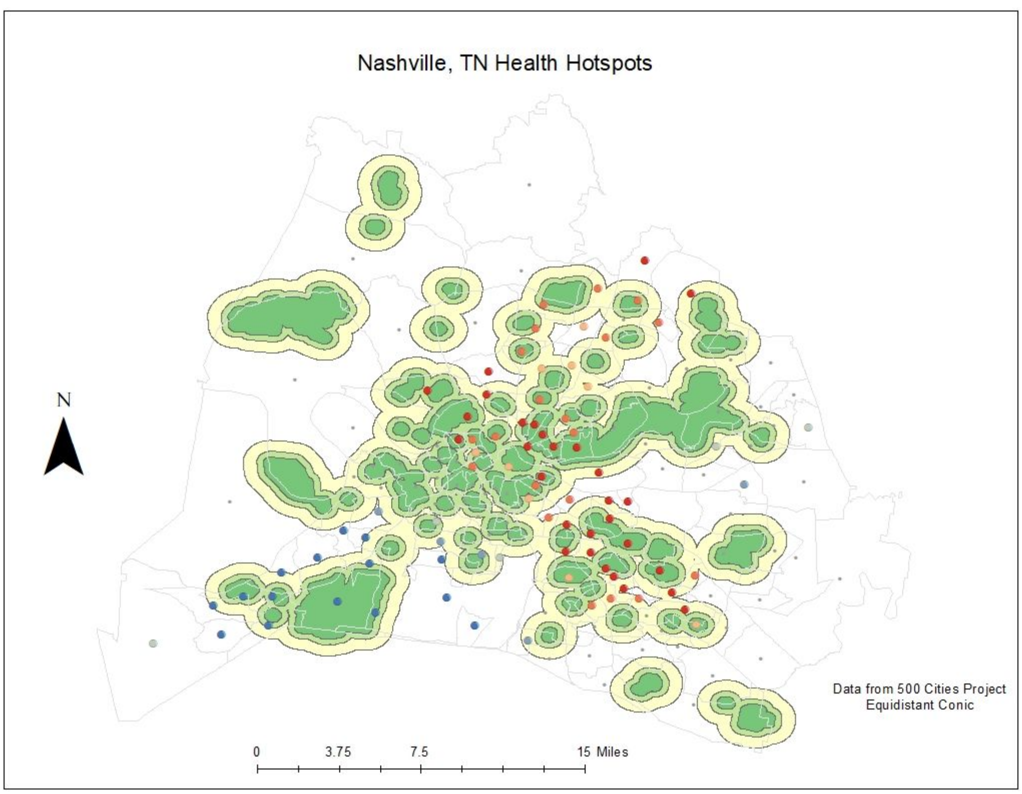

Boston, MA: Health Hot Spots
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

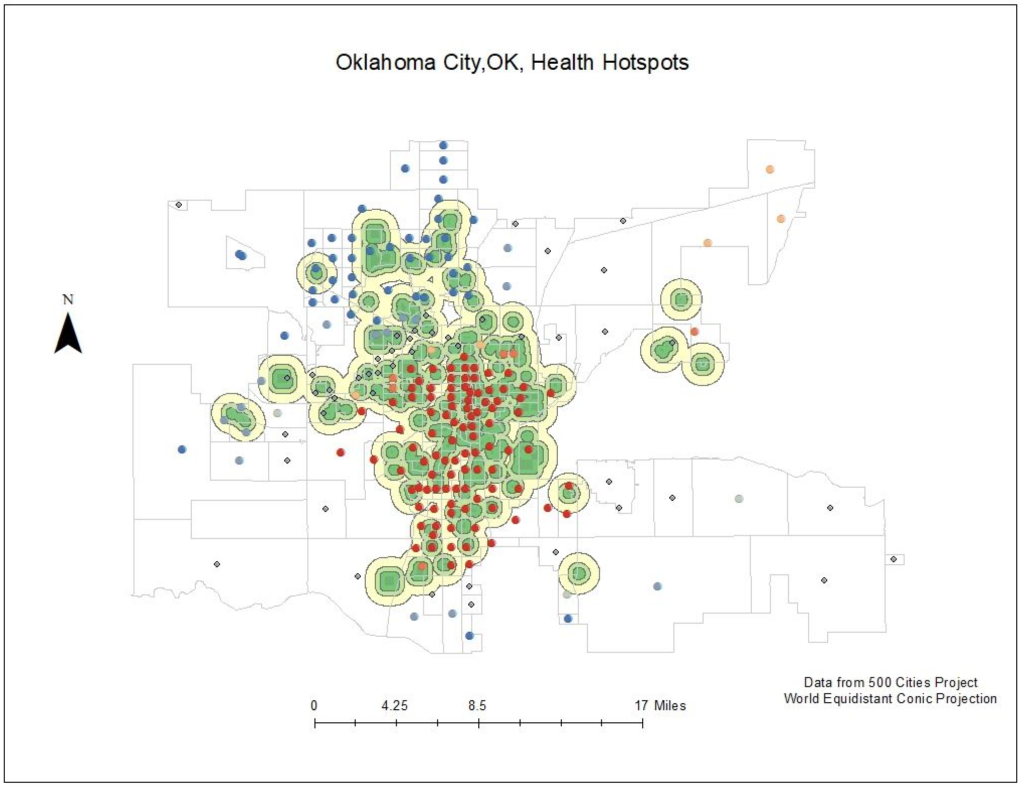

Denver, CO: Health Hot Spots
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1.5 3 4.5 6 Miles

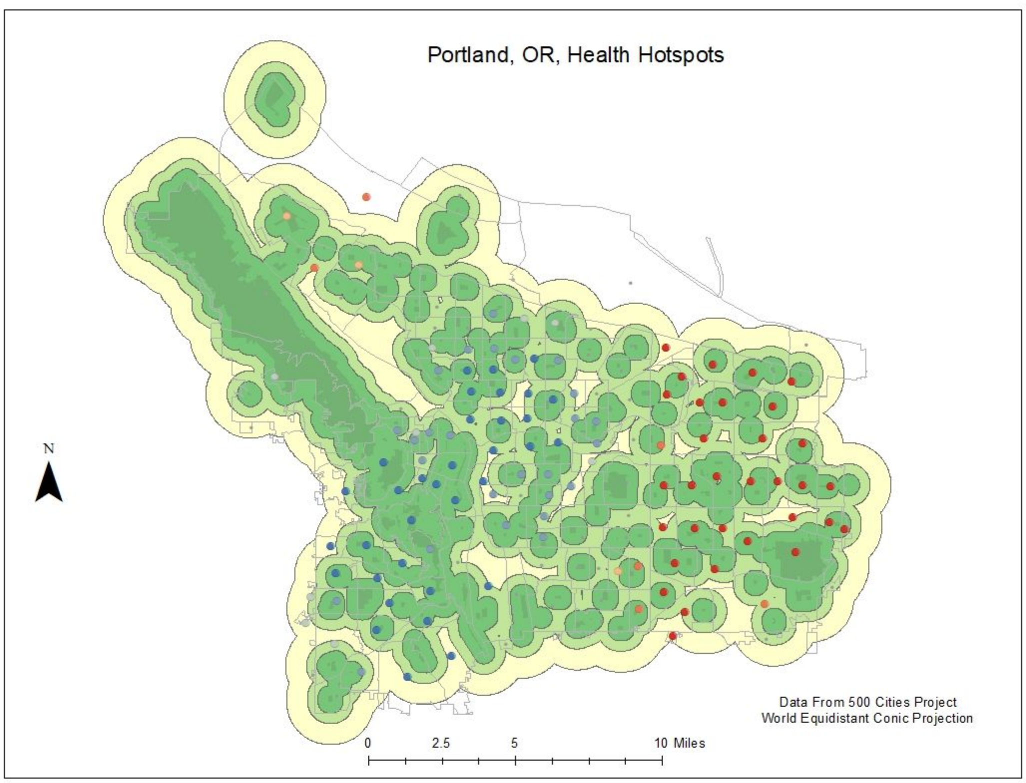

Detroit, MI: Health Hot Spots
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

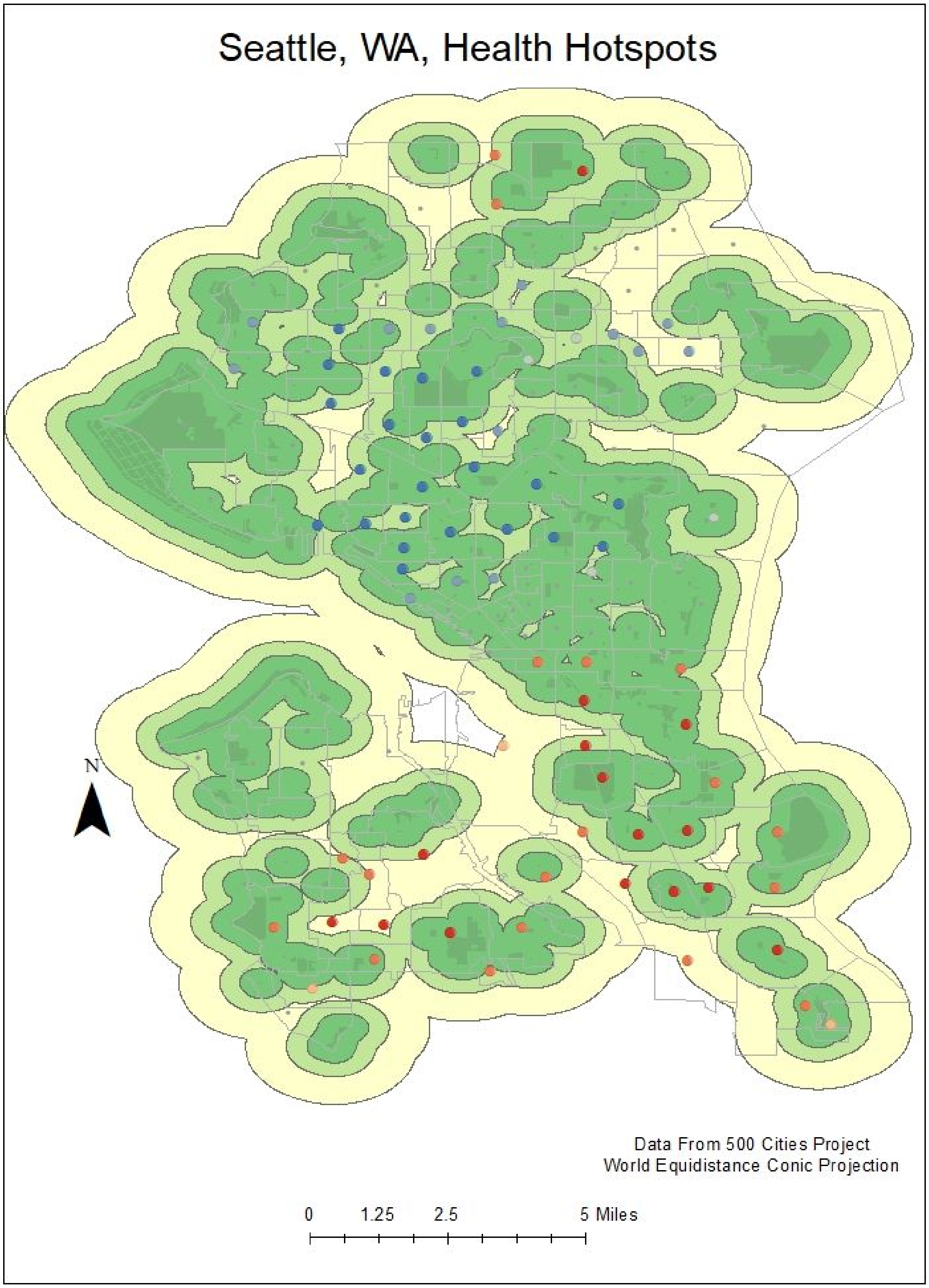

Las Vegas, NV: Health Hot Spots
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

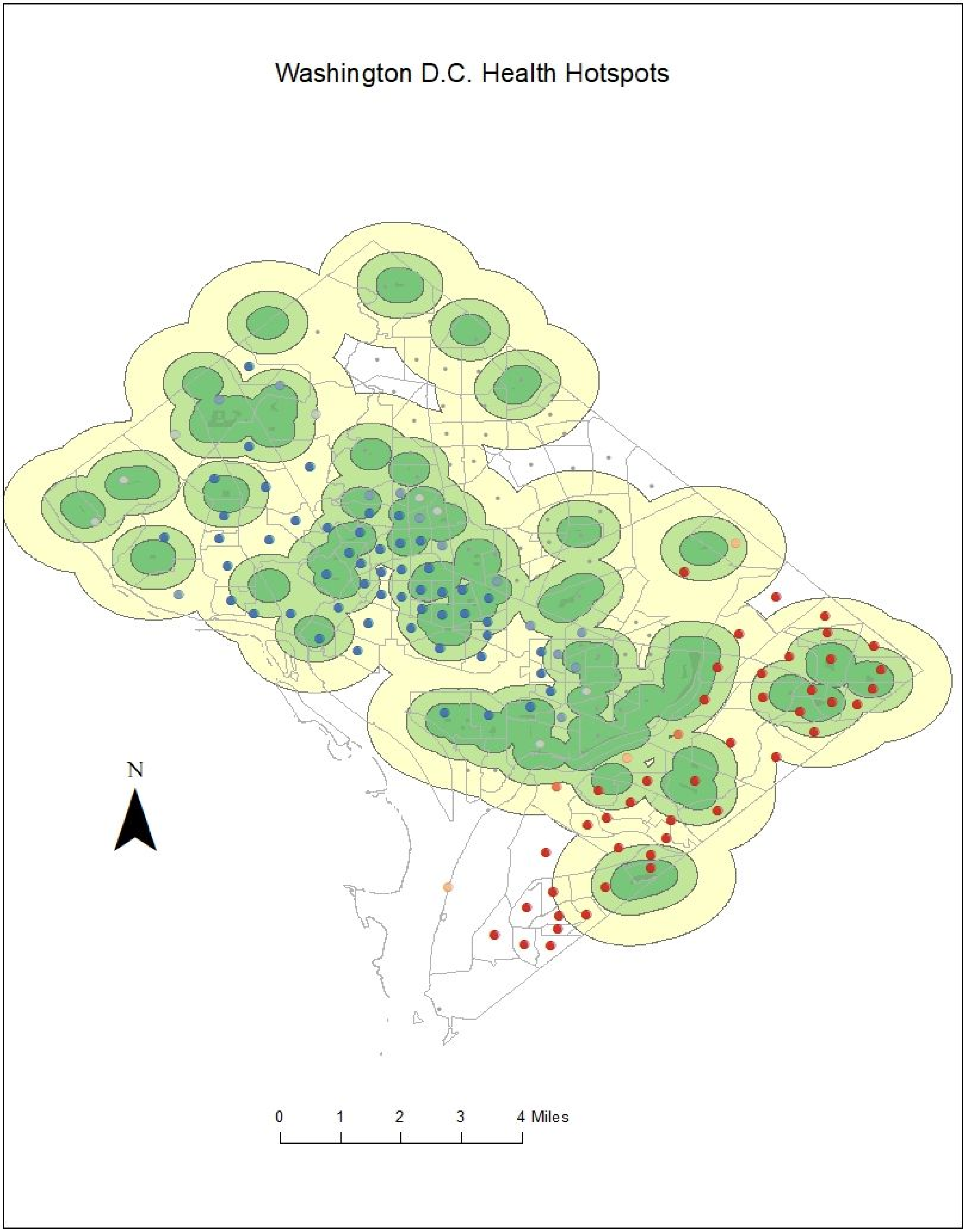

Memphis, TN: Health Hotspots
N
Data provided by 500 cities project
Equidistant Conic Projection
0 2 4 6 8 Miles


Nashville, TN Health Hotspots
N
Data from 500 Cities Project
Equidistant Conic
0 3.75 7.5 15 Miles


Oklahoma City,OK, Health Hotspots
N
Data from 500 Cities Project
World Equidistant Conic Projection
0
4.25
8.5
17 Miles

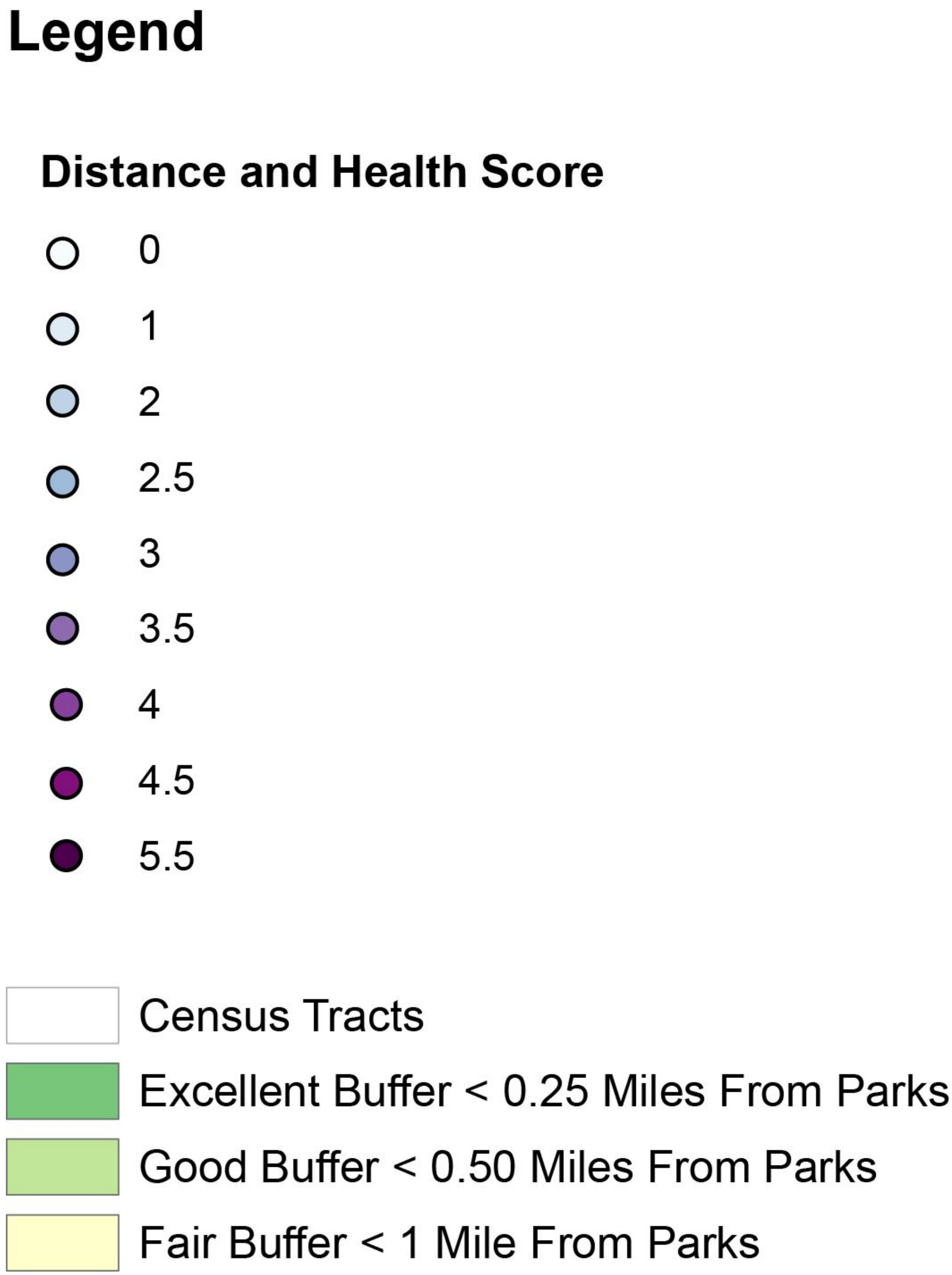

Portland, OR, Health Hotspots
N
Data From 500 Cities Project
World Equidistant Conic Projection
0
2.5
5
10 Miles

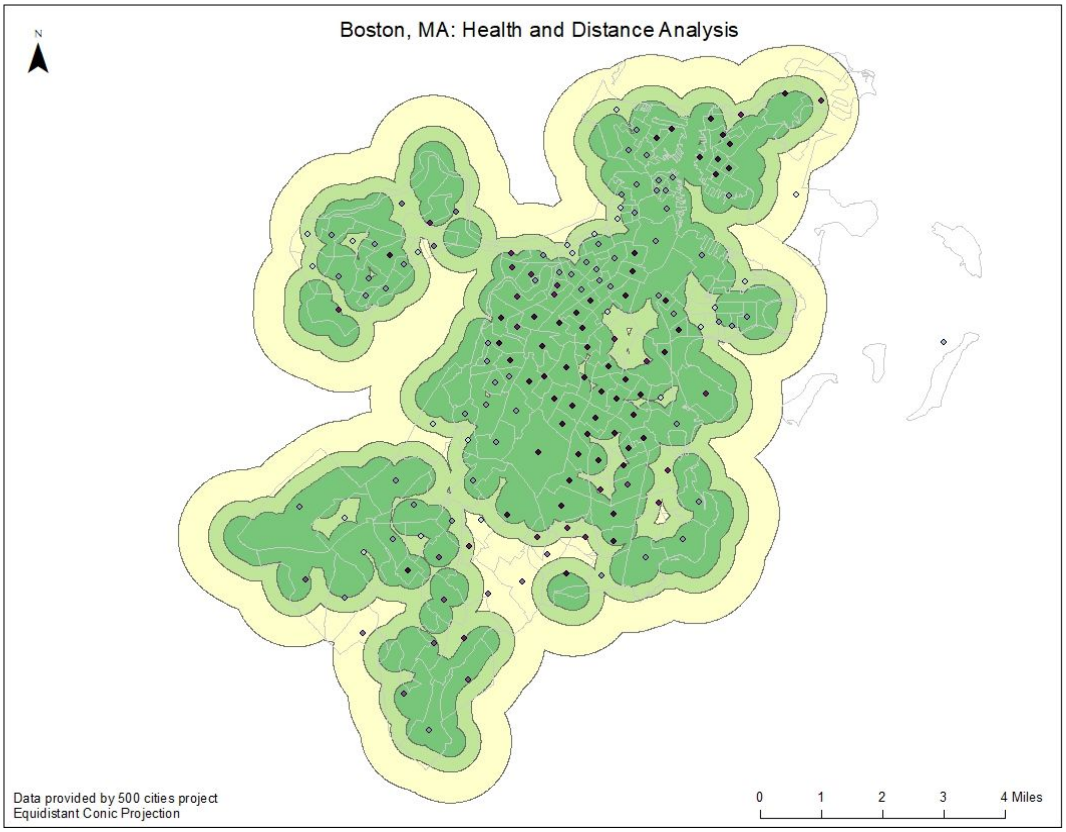

Seattle, WA, Health Hotspots
N
Data From 500 Cities Project
World Equidistance Conic Projection
0
1.25
2.5
5 Miles

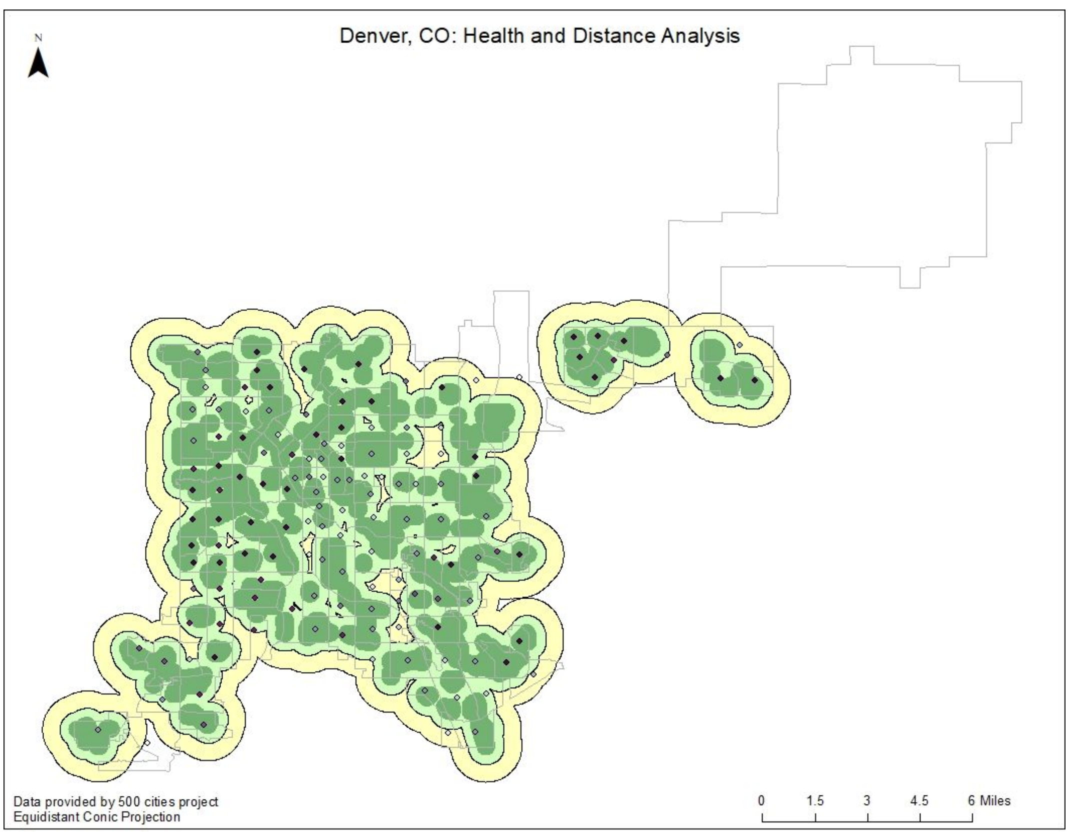

Washington D.C. Health Hotspots
N
0
1
2
3
4 Miles

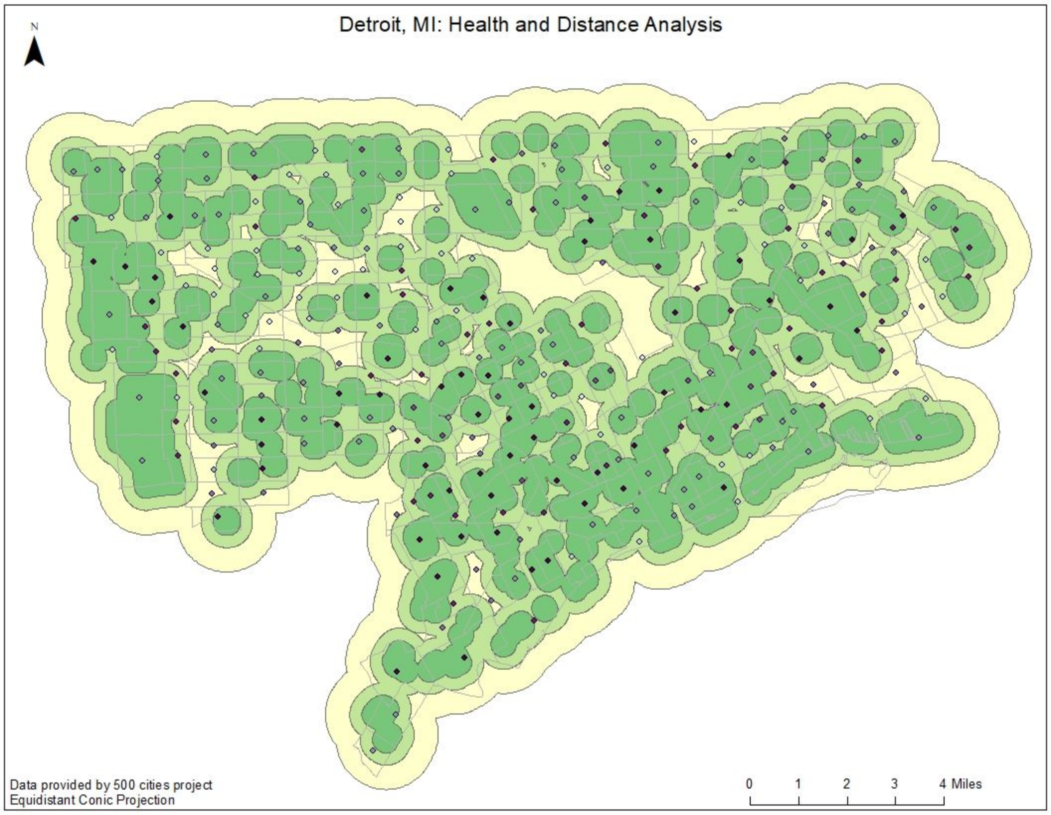
Appendix G: Health and Distance Analysis Maps

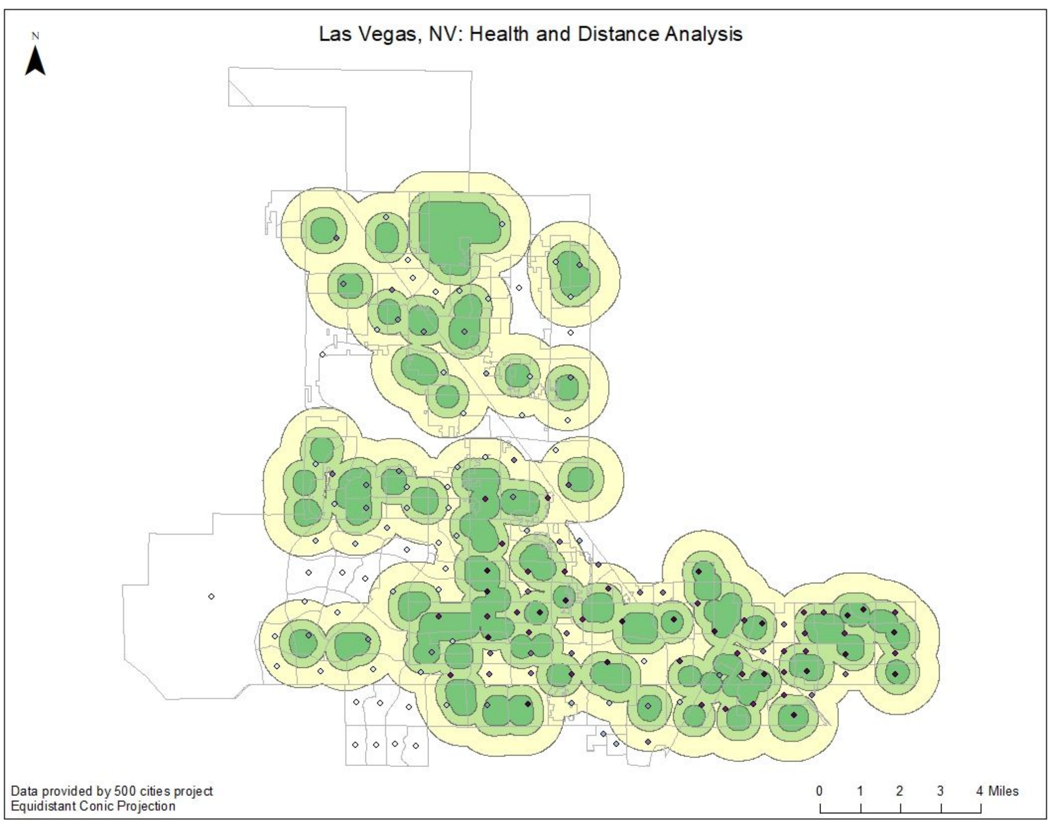
Legend applies to all maps in section

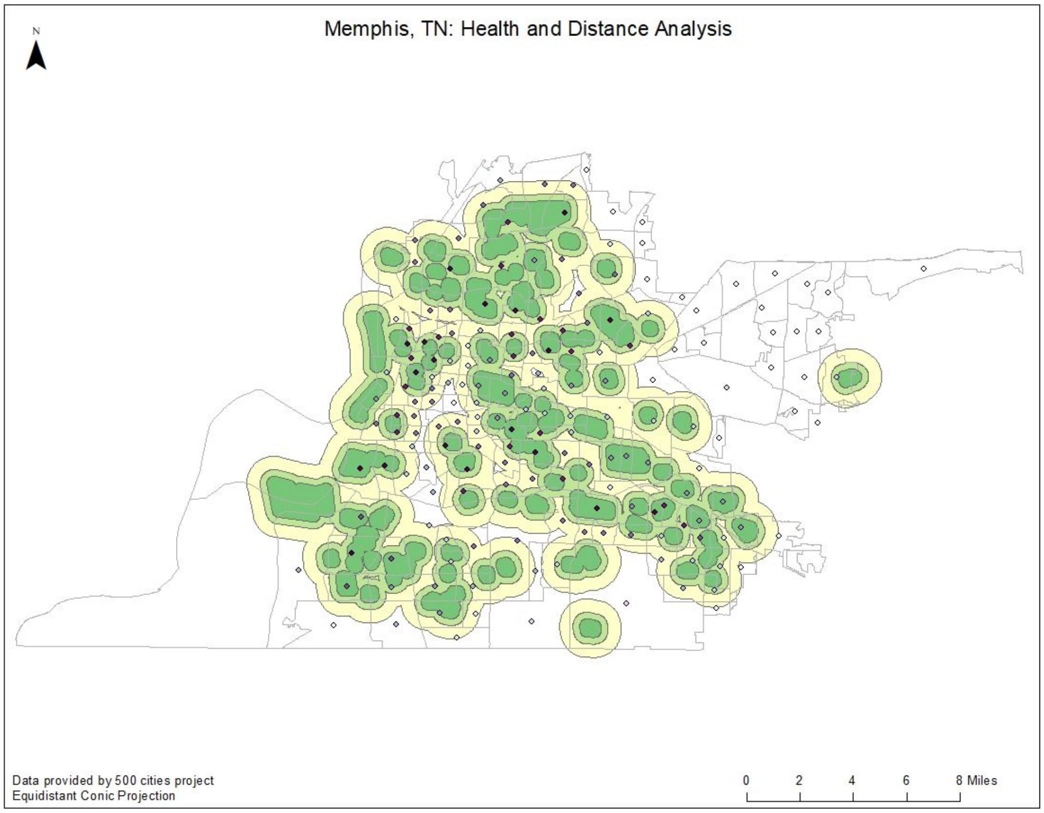


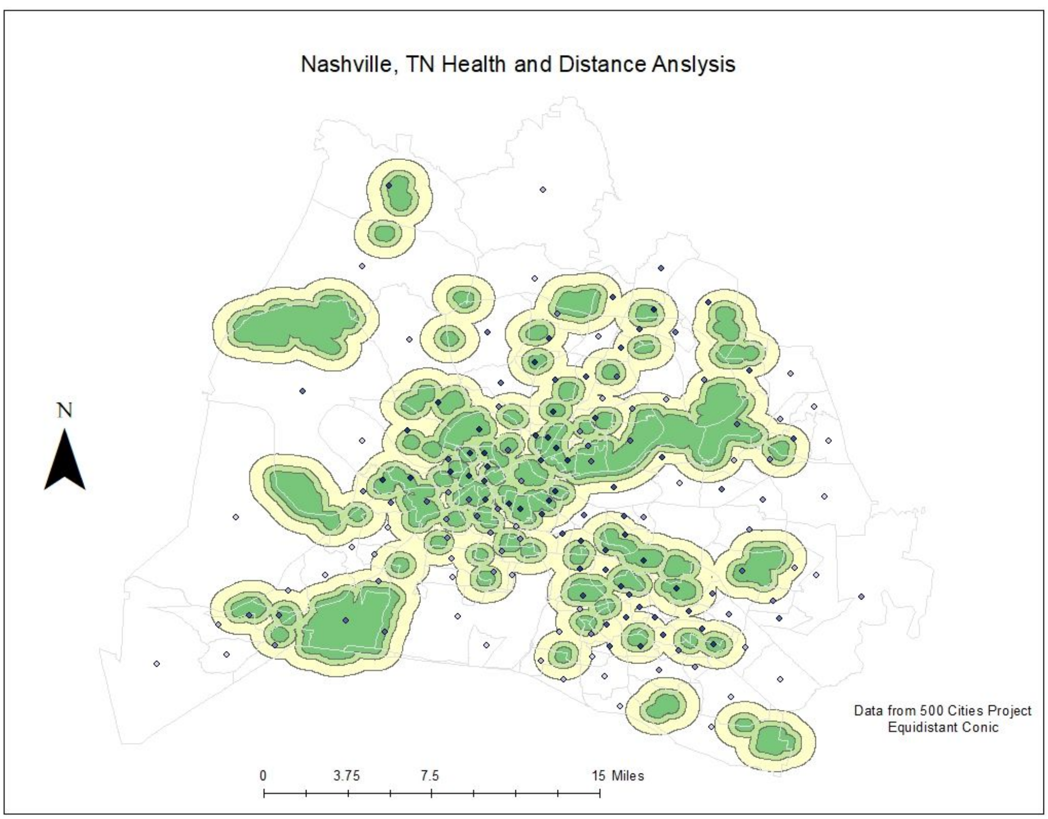

Boston, MA: Health and Distance Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

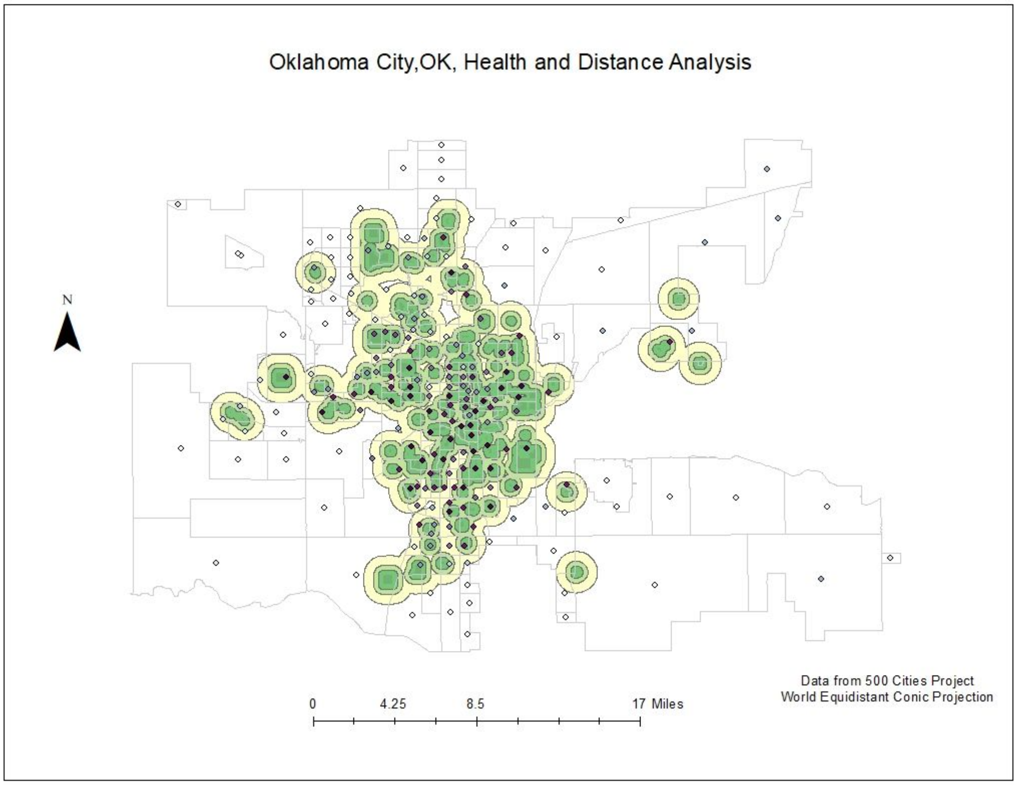

Denver, CO: Health and Distance Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1.5 3 4.5 6 Miles

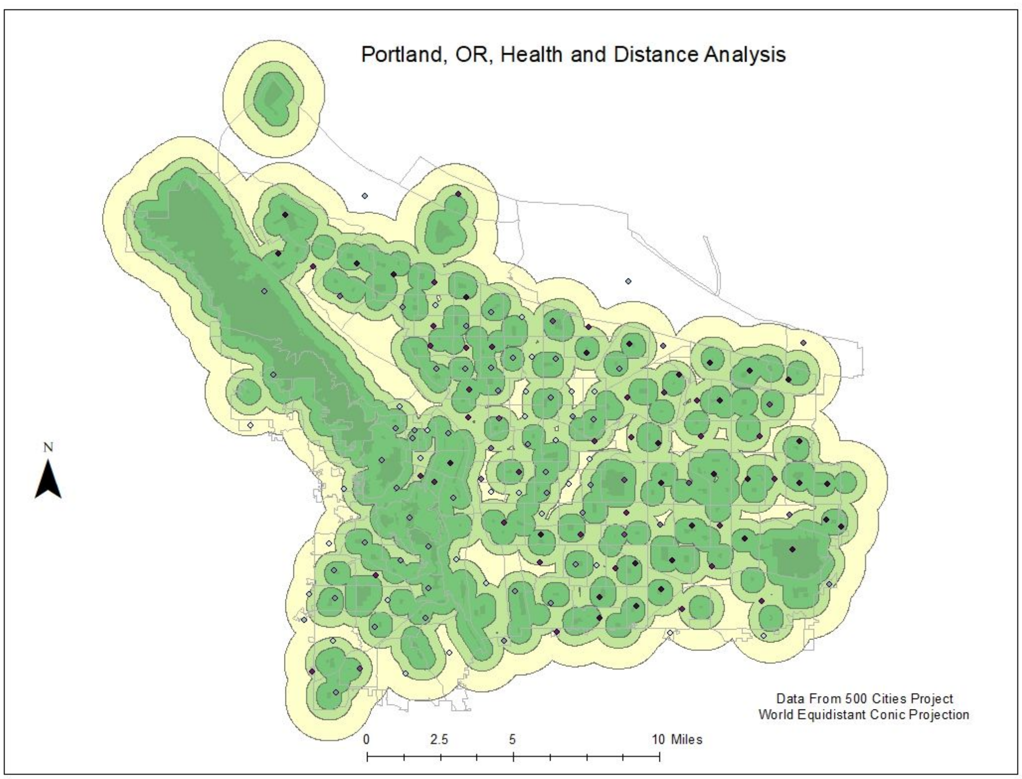

Detroit, MI: Health and Distance Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

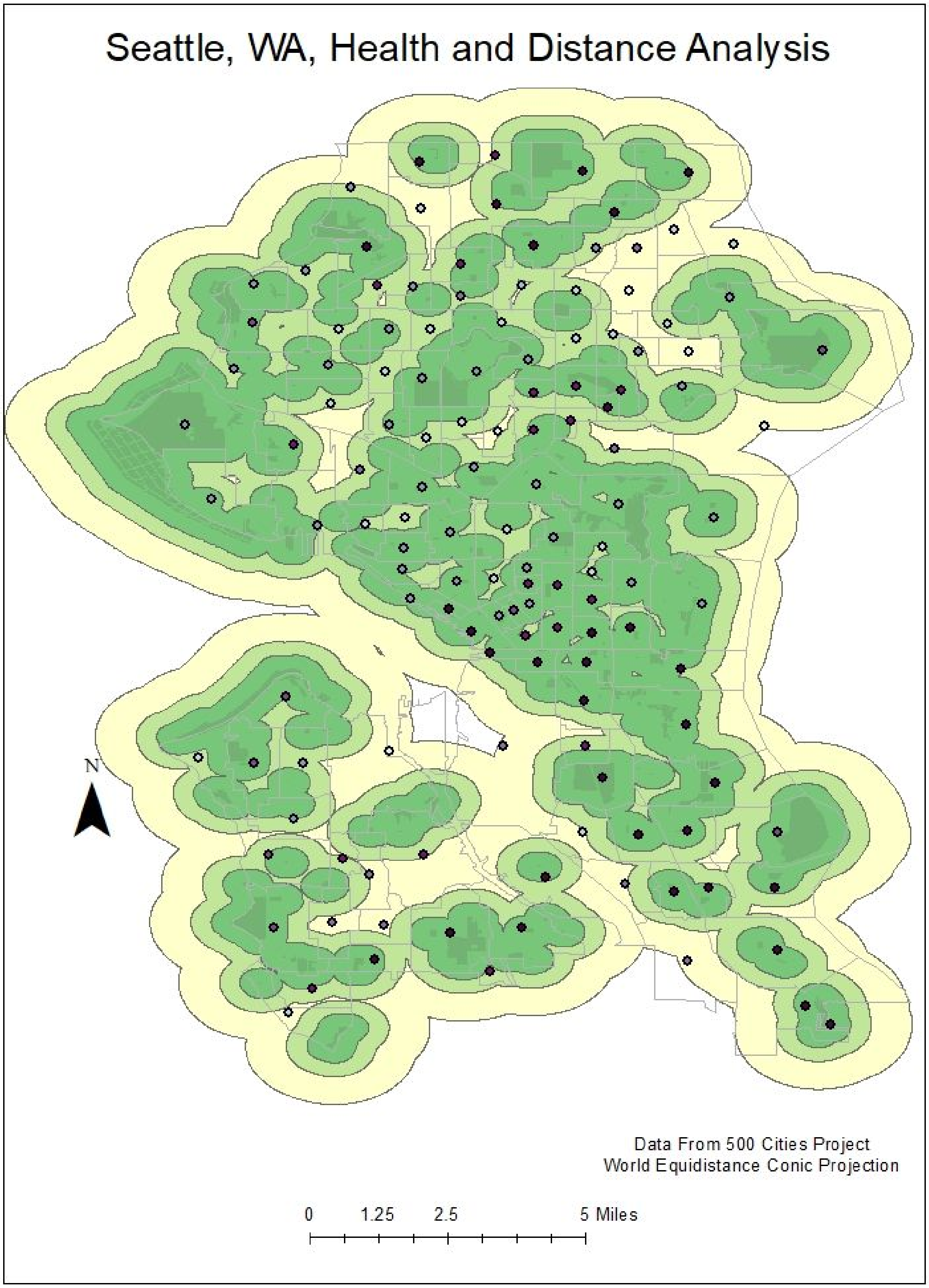

Las Vegas, NV: Health and Distance Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles


Memphis, TN: Health and Distance Analysis
N
Data provided by 500 cities project
Equidistant Conic Projection
0
2
4
6
8 Miles

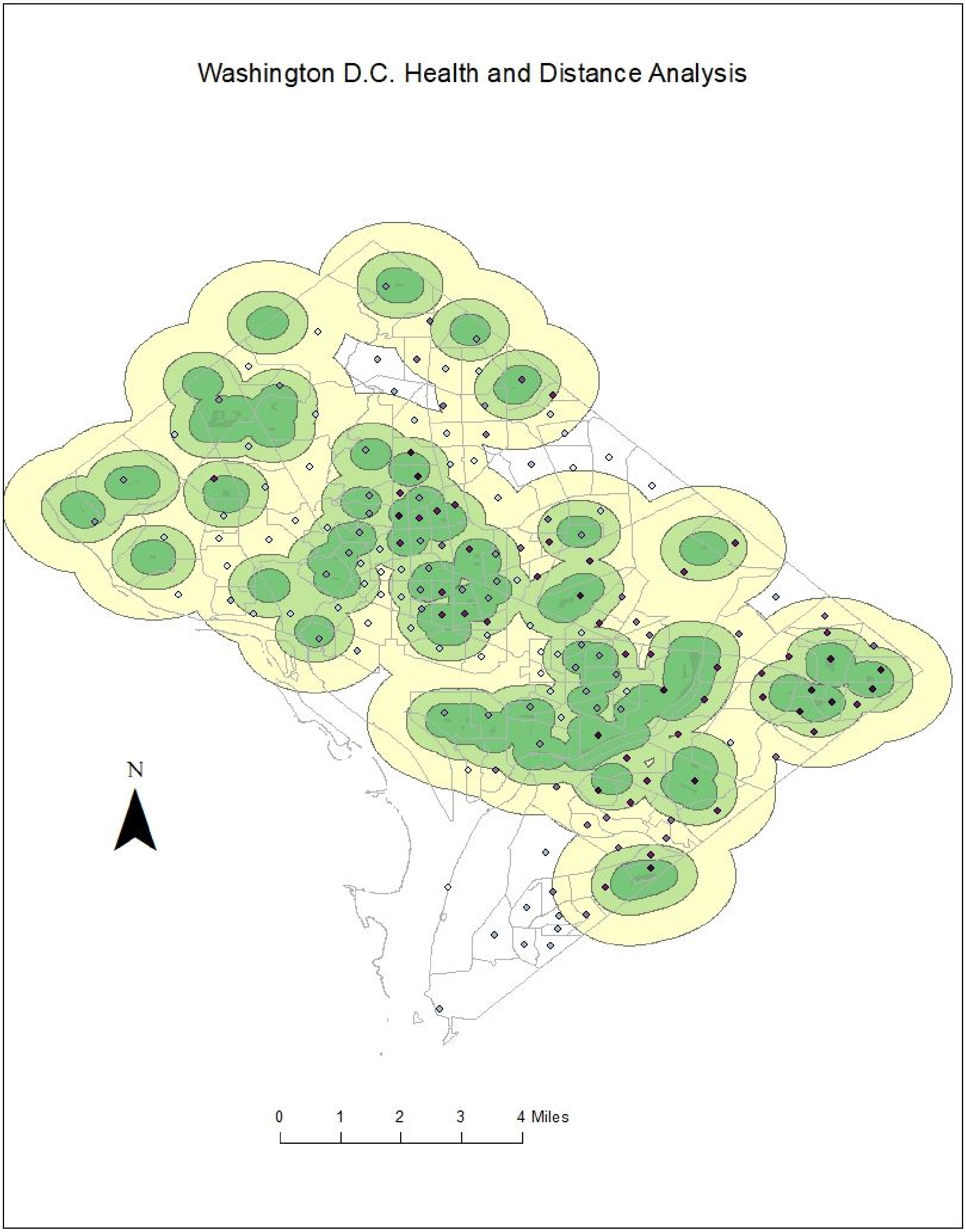

Nashville, TN Health and Distance Anslysis
N
Data from 500 Cities Project
Equidistant Conic
0
3.75
7.5
15 Miles


Oklahoma City,OK, Health and Distance Analysis
N
Data from 500 Cities Project
World Equidistant Conic Projection
0
4.25
8.5
17 Miles


Portland, OR, Health and Distance Analysis
N
Data From 500 Cities Project
World Equidistant Conic Projection
0
2.5
5
10 Miles



Seattle, WA, Health and Distance Analysis
N
Data From 500 Cities Project
World Equidistance Conic Projection
0
1.25
2.5
5 Miles


Washington D.C. Health and Distance Analysis
N
0 1 2 3 4 Miles

Appendix H: Health and Distance Hot Spots Maps

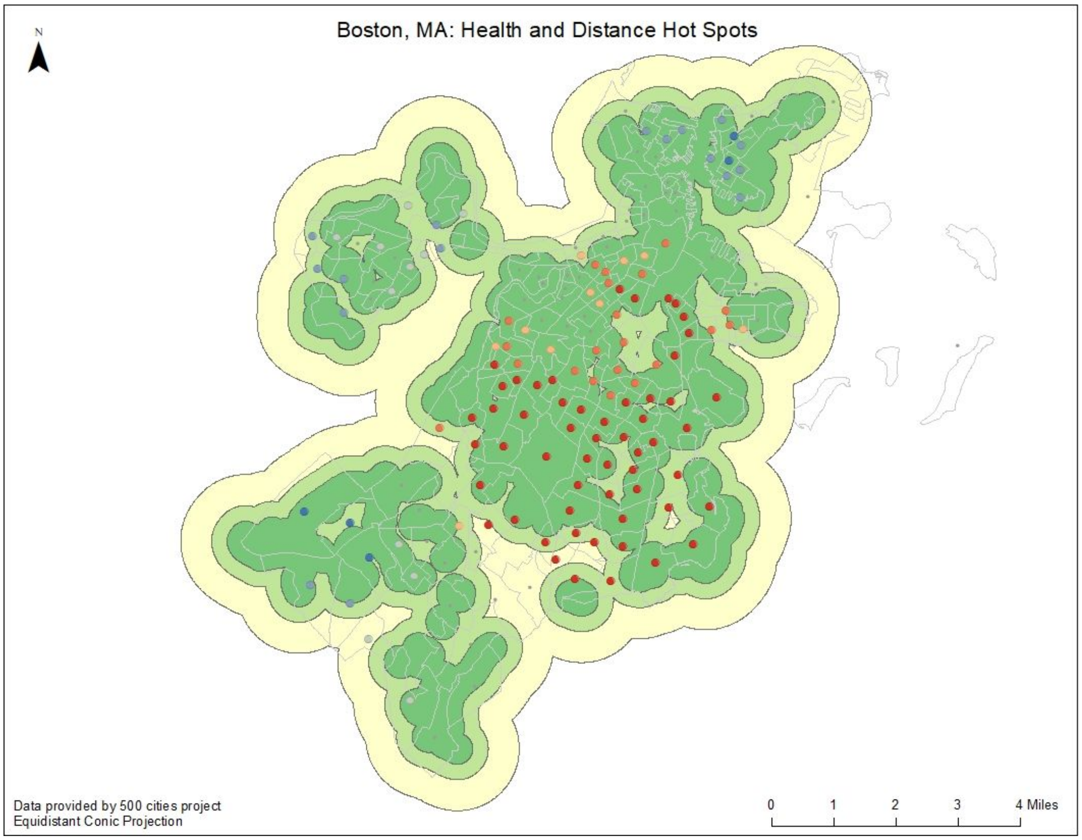
Legend applies to all maps in section

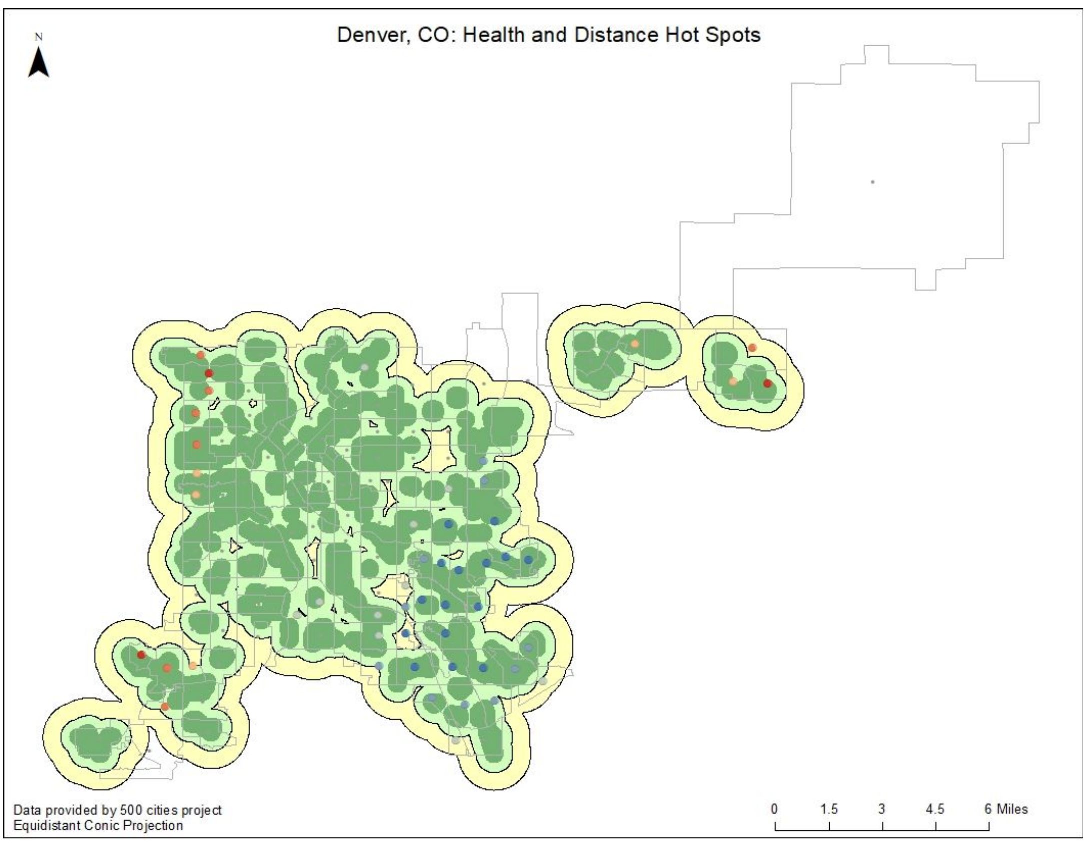
Legend

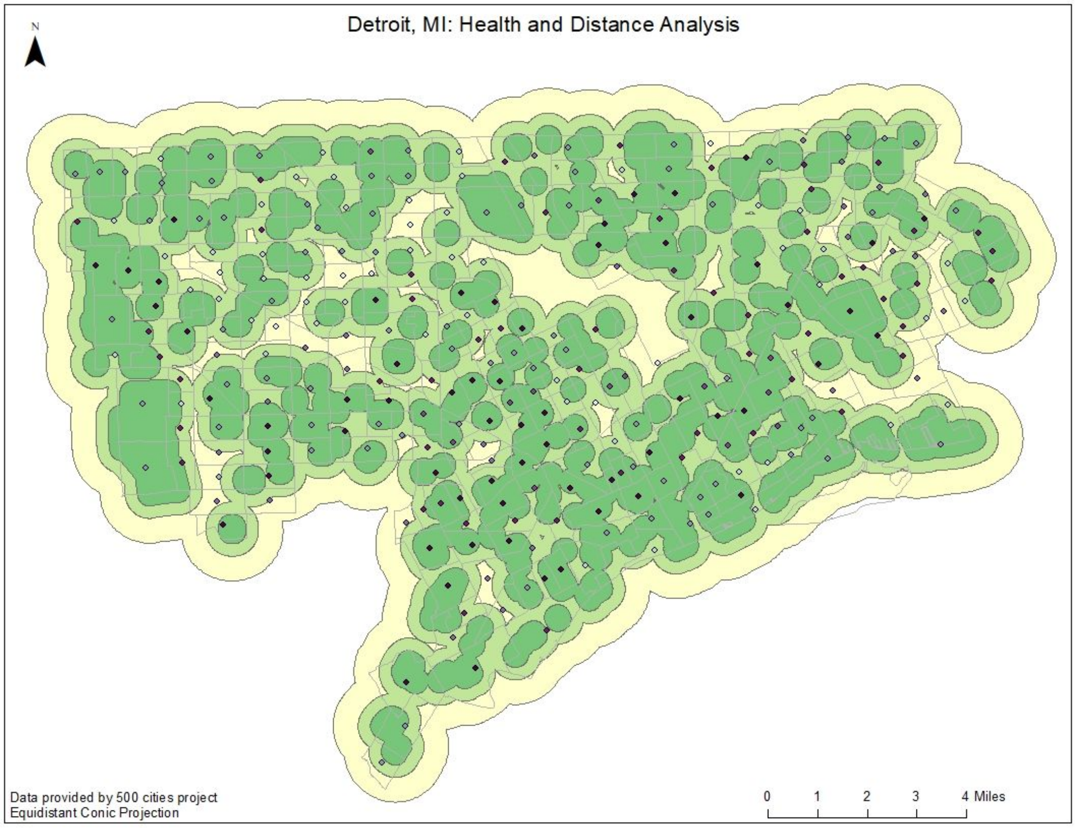
Hotspot Scores

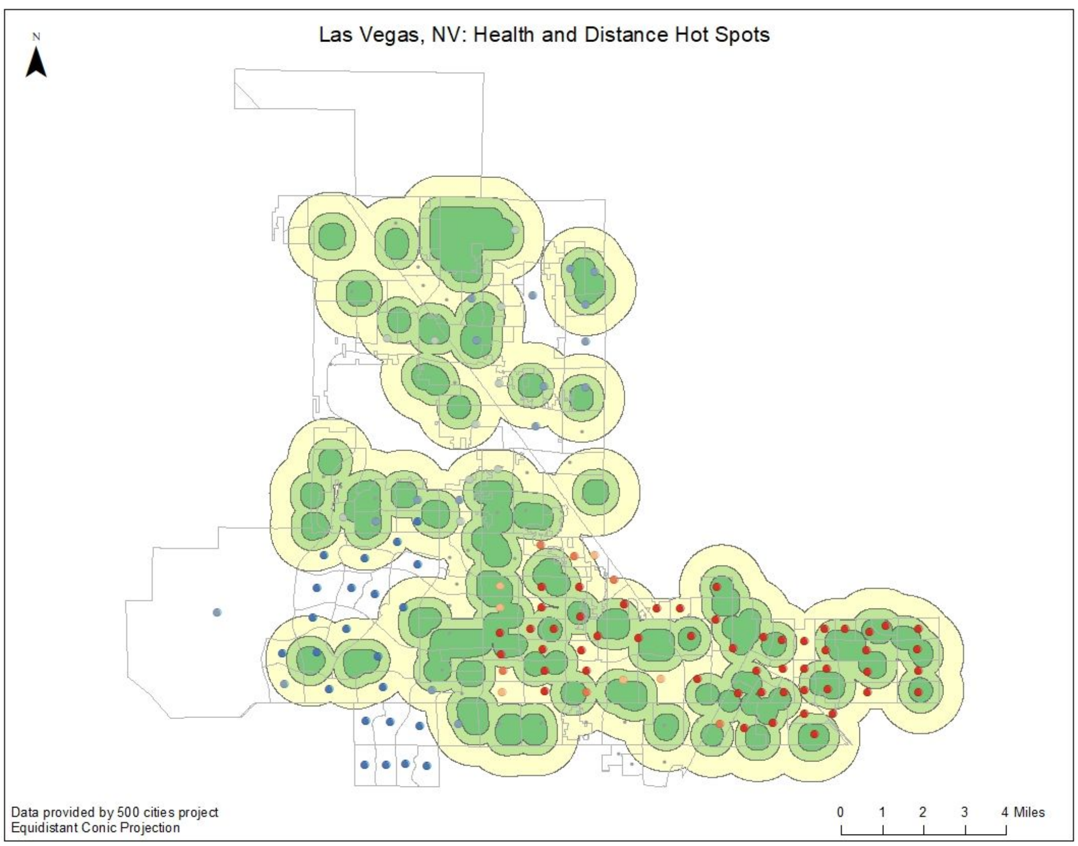
- Cold Spot - 99% Confidence
- Cold Spot - 95% Confidence
- Cold Spot - 90% Confidence
- Not Significant
- Hot Spot - 90% Confidence
- Hot Spot - 95% Confidence
- Hot Spot - 99% Confidence

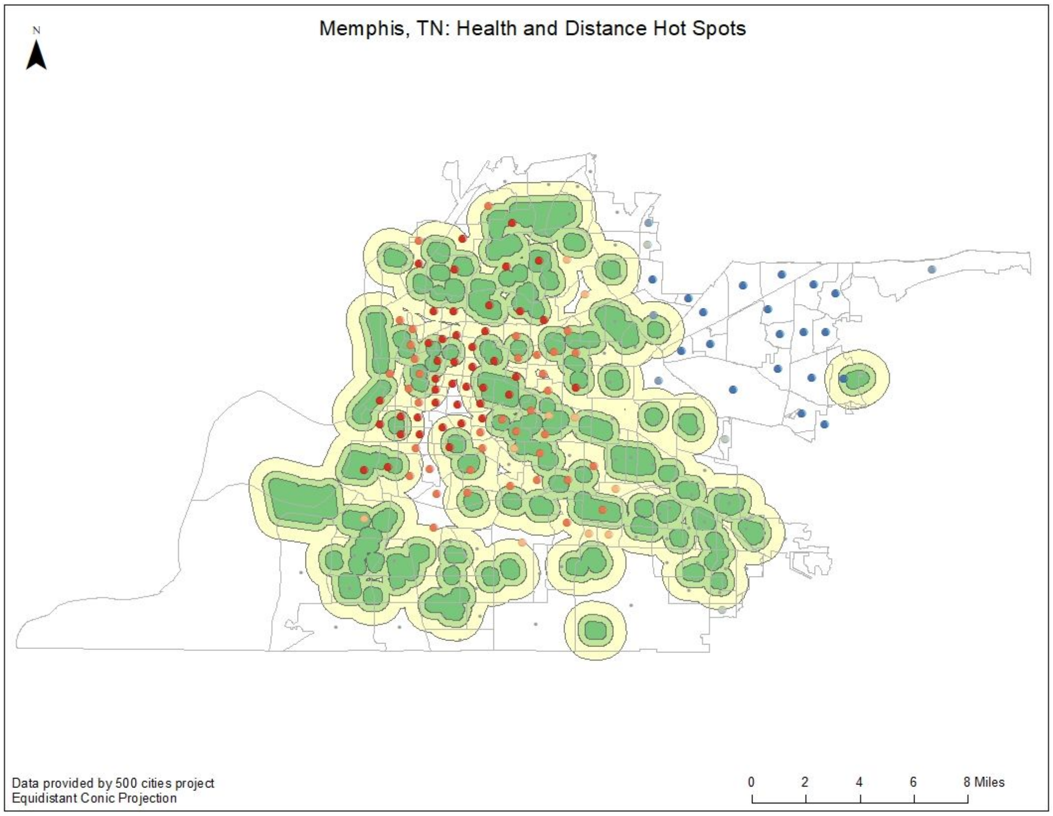
- Census Tracts
- Excellent Buffer < 0.25 Miles From Parks
- Good Buffer < 0.50 Miles From Parks
- Fair Buffer < 1 Mile From Parks

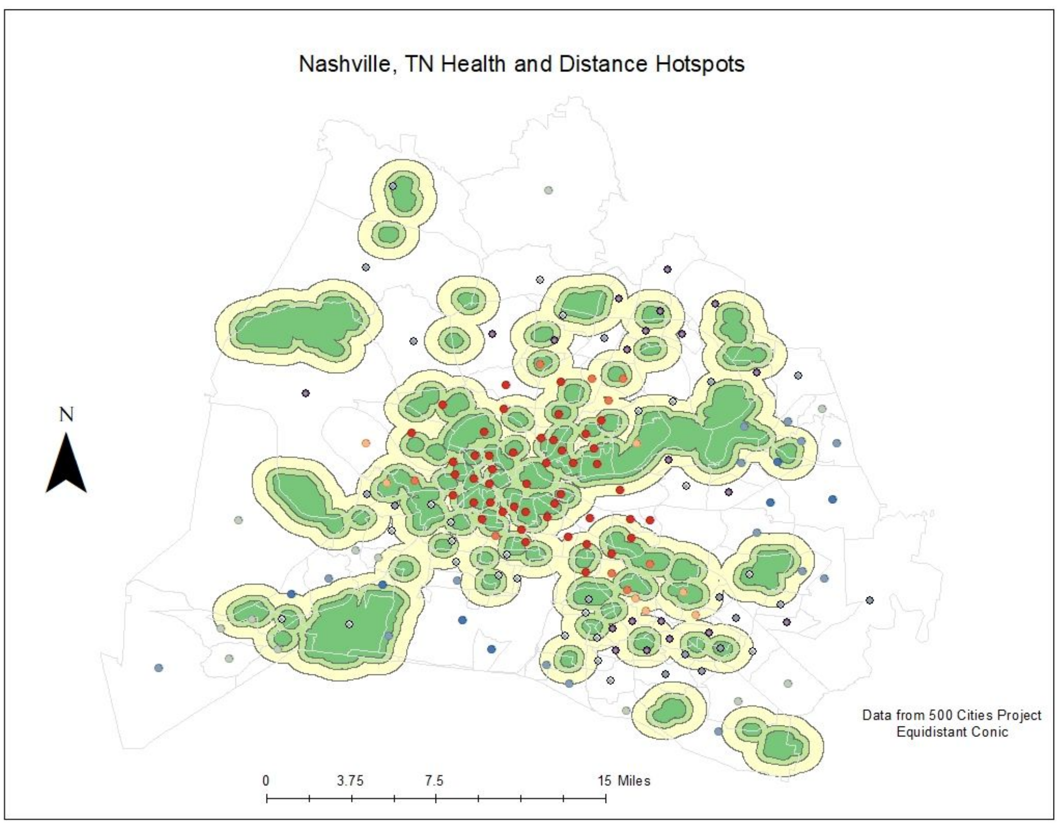

Boston, MA: Health and Distance Hot Spots
N
Data provided by 500 cities project
Equidistant Conic Projection
0
1
2
3
4 Miles

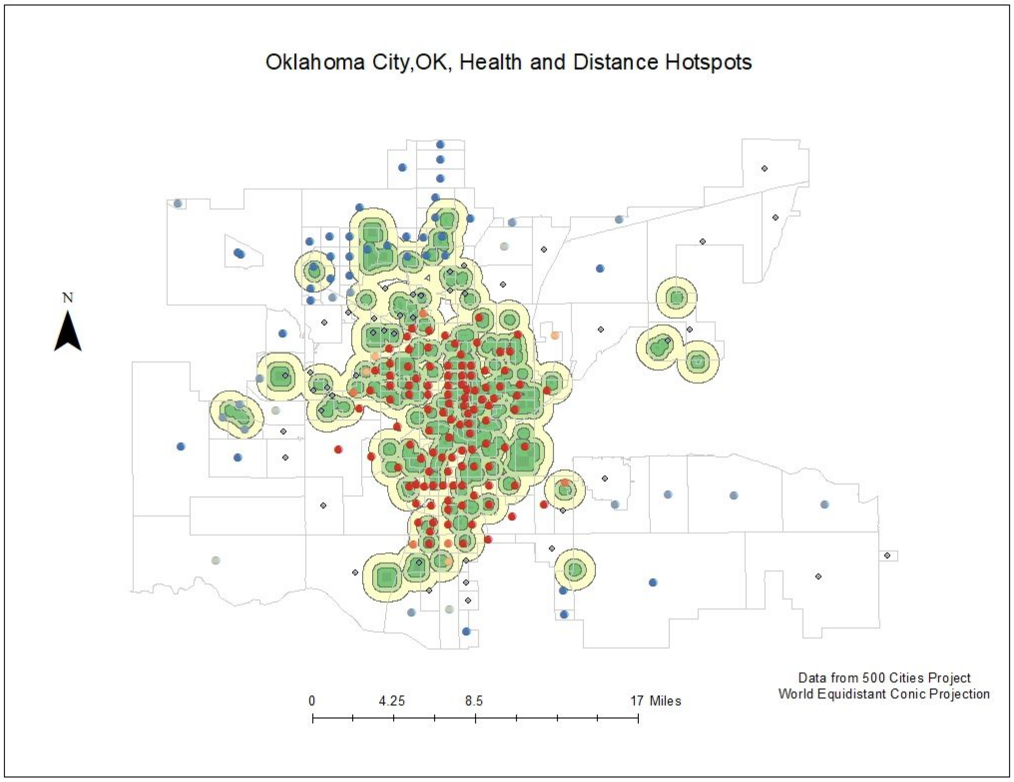

Denver, CO: Health and Distance Hot Spots
N
Data provided by 500 cities project
Equidistant Conic Projection
0
1.5
3
4.5
6 Miles

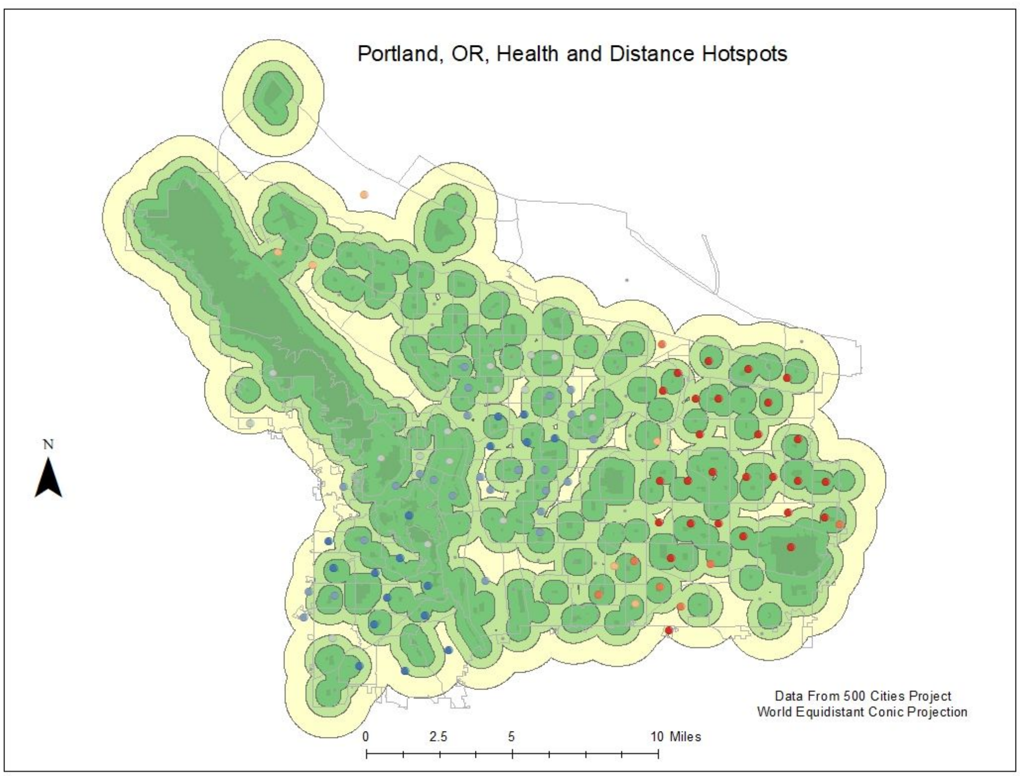

N
Detroit, MI: Health and Distance Analysis
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

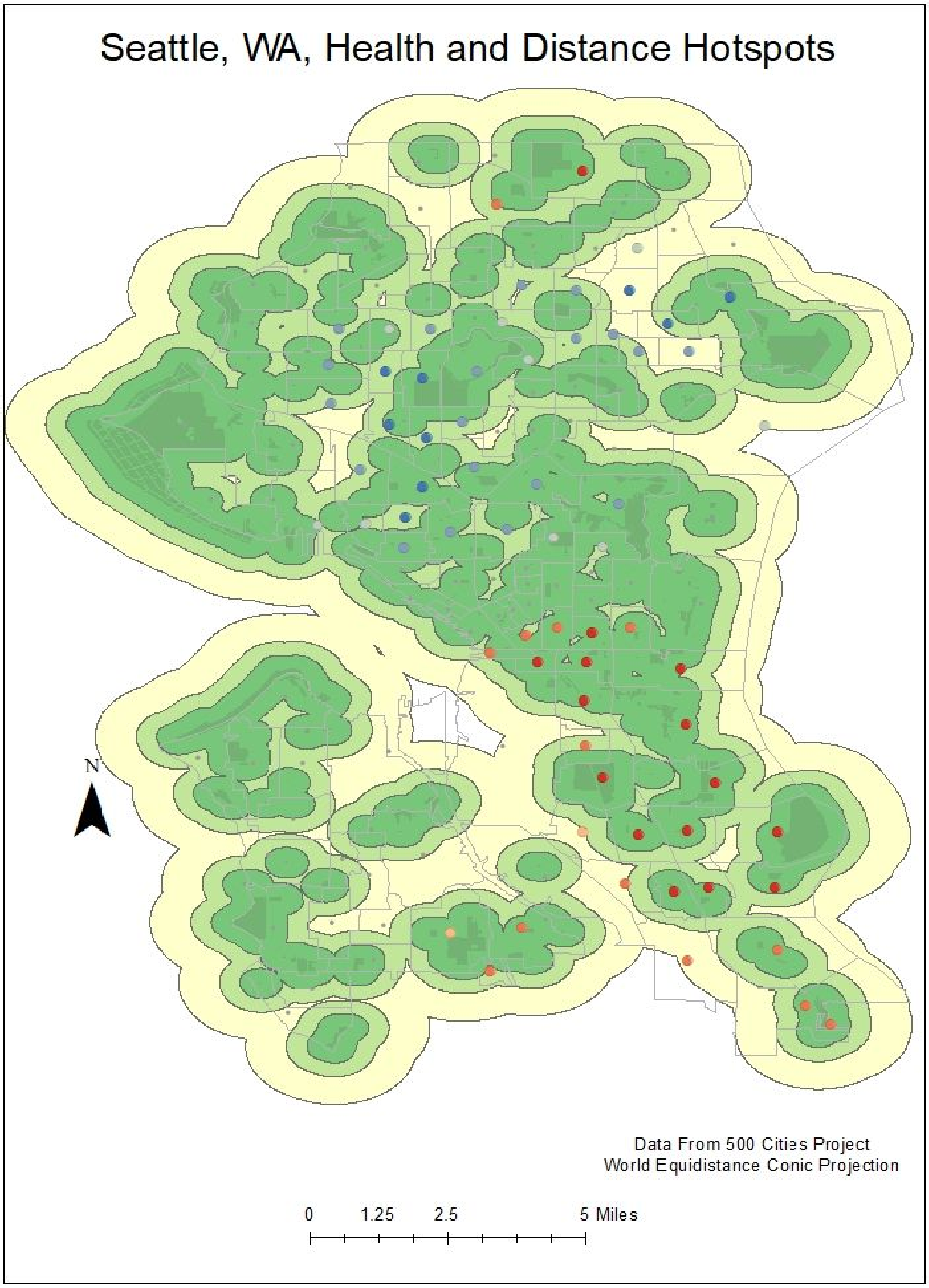

N
Las Vegas, NV: Health and Distance Hot Spots
Data provided by 500 cities project
Equidistant Conic Projection
0 1 2 3 4 Miles

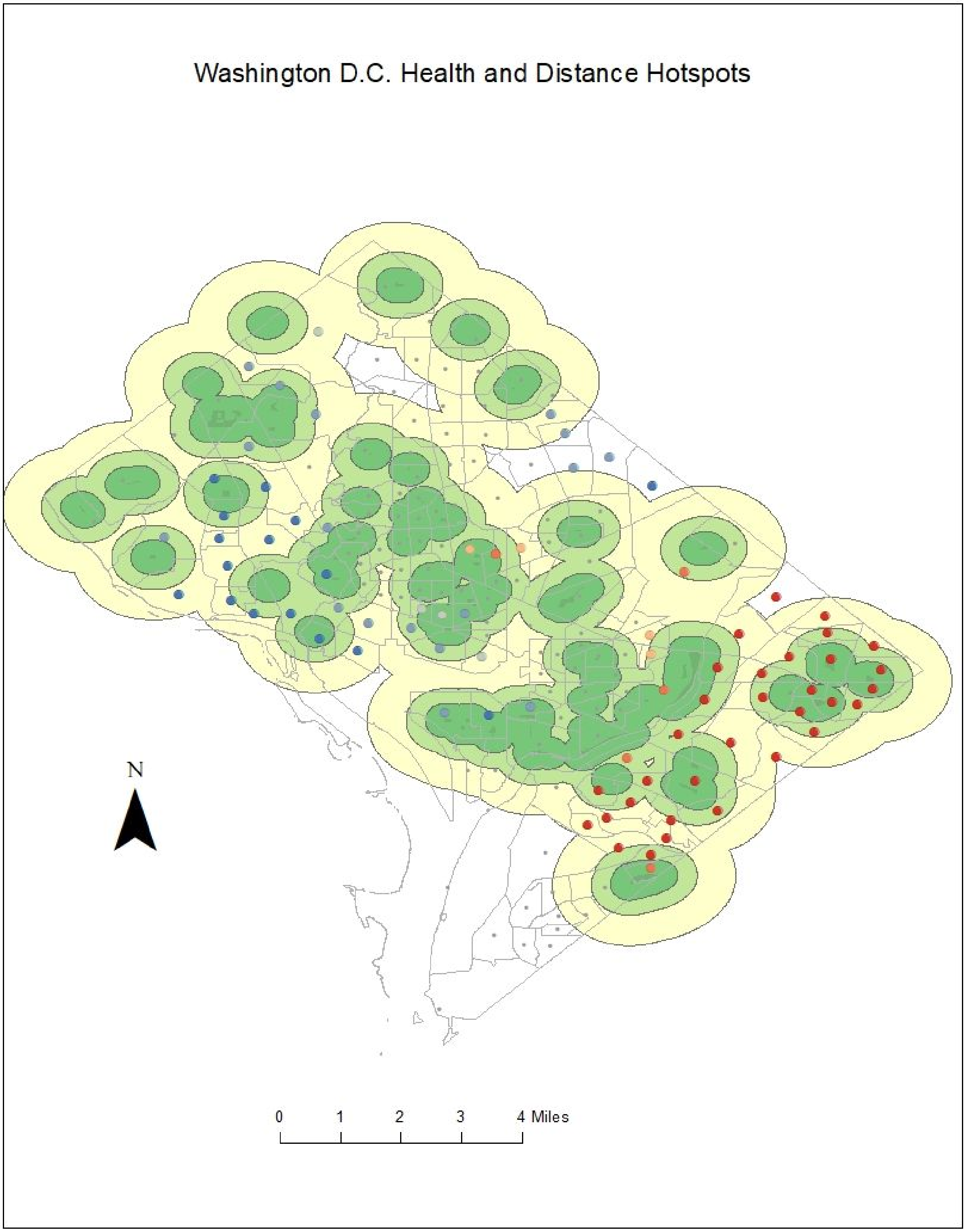

Memphis, TN: Health and Distance Hot Spots
N
Data provided by 500 cities project
Equidistant Conic Projection
0
2
4
6
8 Miles


Nashville, TN Health and Distance Hotspots
N
Data from 500 Cities Project
Equidistant Conic
0
3.75
7.5
15 Miles


Oklahoma City,OK, Health and Distance Hotspots
N
0
4.25
8.5
17 Miles
Data from 500 Cities Project
World Equidistant Conic Projection


Portland, OR, Health and Distance Hotspots
N
0
2.5
5
10 Miles
Data From 500 Cities Project
World Equidistant Conic Projection


Seattle, WA, Health and Distance Hotspots
N
Data From 500 Cities Project
World Equidistance Conic Projection
0
1.25
2.5
5 Miles


Washington D.C. Health and Distance Hotspots
N
0
1
2
3
4 Miles

Appendix I: Summary Statistics of Spatial Autocorrelation

	Health Index	Health Z-score	Health P-value	Combined Index	Combined Z-score	Combined P-value	Distance Mean Score	Health Mean Score
Boston, MA	0.190191	15.499239	0.000000	0.175964	14.390506	0.000000	2.693642	1.231214
Denver, CO	0.144718	12.434471	0.000000	0.089473	7.926696	0.000000	2.426573	1.241259
Detroit, MI	0.236931	6.75134	0.000000	0.259111	7.384318	0.000000	2.438356	1.243151
Las Vegas, NV	0.739285	21.739032	0.000000	0.628715	18.56551	0.000000	1.738562	1.24183
Memphis, TN	0.226052	11.481348	0.000000	0.237137	12.060738	0.000000	1.483146	1.241573
Nashville, TN	0.288679	11.975869	0.000000	0.379511	15.698555	0.000000	1.397436	1.24359
Oklahoma City, OK	0.331596	25.184377	0.000000	0.427536	32.403529	0.000000	1.523148	1.236111
Portland, OR	0.319746	11.701129	0.000000	0.246307	9.092587	0.000000	2.444444	1.22549
Seattle, WA	0.455027	11.475789	0.000000	0.343234	8.721072	0.000000	2.462687	1.19403
Washington, DC	0.455685	17.987204	0.000000	0.291424	11.619702	0.000000	1.876404	1.227528

Appendix J: Pearson Product-Moment Correlation Coefficients

City	Pearson Correlation-Coefficient	P-Value
Boston, MA	-0.015412605	0.840492551
Denver, CO	0.120142077	0.152928533
Detroit, MI	-0.053266119	0.364432325
Las Vegas, NV	0.218799784	0.006583404
Memphis, TN	0.134773889	0.072871908
Nashville, TN	0.169250508	0.034669323
Oklahoma City, OK	0.34111236	0.000000276
Portland, OR	-0.078142249	0.336990033
Seattle, WA	0.047276068	0.587517895
Washington, DC	-0.254220118	0.000616484