Identifying Priority Areas for Poor Air Quality Mitigation

1. Introduction

The impact of global warming is a pressing issue to everyone at a global scale. However, at a more granular scale, within the urban context, it has been well established that socially vulnerable communities of color, lower education, and lower income are exposed to higher emissions and poorer air quality (Barnes et al., 2019; Lane et al., 2022; Tessum et al., 2019; Jbaily et al., 2022; Servadio et al., 2018; Martenies et al., 2017). Even when conditions improved, a study of PM2.5 concentrations over 36 years of research in the U.S. found that the most exposed subpopulations remained at a disadvantage (Colmer et al., 2020). Consequently, this disparity is evident in the health and mortality rates as well. Several studies, dedicated to studying exposure to air pollutants, NO2, PM10 and SO2, across socio-economic levels and demographics in Europe, found a positive association between less affluent and deprived communities, and significant health effects of pollutants: cardiovascular mortality, asthma attacks, and non-trauma mortality (Deguen and Zmirou-Navier 2010; Servadio et al., 2018; Martenies et al., 2017; Ouyang et al., 2018).

1.1 Detroit Background

This GIS project could be conducted in any large urban context, but I chose Detroit due to its standing in the top 20 most polluted cities (American Lung Association, 2022). Additionally, most of the previous academic research I have come across have chosen to explore a city with a smaller African American population than that of White. Detroit's racial makeup currently stands at 77.4% African American, 10.6% White (non-Latino), 1.75% Asian, and 4.09% Hispanic White. It will be intriguing to see the results of the study in a predominantly African American city (Data USA, 2022).

1.2 Problem Statement

Given the ill effects of air quality and the disparity faced by more vulnerable communities, it is inefficient for olicy makers to address the whole city at once. If anything, renovations may simply improve affluent communities once again. So rather policy makers and urban planners need a guide to identify the areas to address and which are in most need of help. Utilizing the existing correlations between air quality and its indicators, socioeconomic status, demographic, green space, and traffic density, I plan to spatially analyze vulnerability, with respect to air pollution, and determine the areas within Detroit, MI that require immediate improvement.

2. Relevant Work

The current academic research, about building indicators of air quality, examines at least 1 of four indicators: socioeconomics and demographics, urban environment structure (green infrastructure, permeable surfaces, Urban Heat Islands (UHI), etc.), and traffic density. There have been numerous attempts at creating a vulnerability index or tool through GIS spatial analysis to help identify urban locations that require intervention through the addition of green infrastructure. However, every study took into consideration different variables or utilized a different method of analysis and consequently produced a different index.

2.1 *Vulnerability Index with Greenspace*

One such study based in Chile, utilized air pollution, population density, socioeconomic status and vegetation data to first calculate the environmental stress indicator (ESI) and the social relevance indicator (SRI) through model analysis at 3 different scales. Then both indicators were combined into a comprehensive vulnerability index and spatially analyzed to determine the priority populations (Ignacio et al., 2018). Similarly, Sabrin et al. built an environmental risk impact index (EVII) to reflect the urban heat islands (UHI) and Ozone-PM2.5 pollution risk with respect to the urban environment, such as permeability of surfaces and vegetation, and a social vulnerability impact index (SVII), which builds on the variables included in EVII, to reflect risk with respect to socioeconomic status and demographics. The indices were created with regression, but both were mapped and spatially analyzed using GIS techniques to identify target neighborhoods in Camden, NJ, for policy makers and planners to focus on for future mitigations (Sabrin et al., 2020). In contrast, Ge et al. defined a new SVI, with respect to air pollution, while factoring in 20 indicators, such as age, gender, green space, and even the availability of medical and management services. The results of the project pursuit cluster (PPC) model revealed high SVI at the northern and southern edges of the study area, and lower SVI at the urban core (Ge et al., 2017). Another approach, an auto-regressive model (SAR), was utilized in the study of total suspended particulate (TSP) across socioeconomic levels in Hamilton, Canada over the course of 10 years at an older study period (1985 – 1994), yet they found similar results; there was an association between exposure to TSP and housing values, low income, and unemployment (Jerret et al., 2016). A more relevant study, in the context of Detroit, instead focuses on identifying areas that would specifically benefit from the addition of green infrastructure. Utilizing Pearson’s bivariate correlation, the researchers build the Green Infrastructure Spatial Planning (GISP) model which balances 6 factors: stormwater abatement, social vulnerability, green spaces, air quality, urban heat amelioration, and accessibility to the green spaces, to provide the most optimal solution at the census tract level.

2.2 Vulnerability w.r.t to Traffic Density Exposure

While the aforementioned studies focused on mitigating air pollution and UHIs through the addition of green space, others have solely examined disparities in exposure to traffic pollution across socioeconomic and demographic variables, in efforts to identify areas for urban planners and policy makers to improve through the alleviation of traffic density (Sun et al., 2020; Schindler et al., 2021). When high-resolution traffic data is available, researchers have relied on kernel density to create a traffic density surface (Pratt et al., 2014; Pratt et al., 2015) that captures the areas exposed to emissions or the inverse, where a 500ft. buffer, around each residence, captures all the traffic density that exposes the respective residence (English et al., 1999). Once the exposure levels are quantified, its correlation with air pollution, socio-demographic factors, and health effects are assessed through mixed methods, such as an interaction matrix (Mavroulidou et al., 2004), principal component analysis (Pratt et al. 2015) or Pearson/Gaussian models (English et al., 1999). The results analyses concluded that less affluent and deprived populations faced worse health effects, such as asthma (English et al., 1999), due to an increased exposure to traffic emissions (Pratt et al. 2015).

3. Analytic Process

Rather than aiming to confirm the correlation between air quality and its indicators once again, my goal is to utilize this established pre-existing relationship to assess and determine the area(s) (a small subset of CBGs) where policy makers and urban planners can focus their efforts of improvement, at least at the beginning. In addition, researchers in the past have focused on either traffic or green infrastructure, but none consider them together. I believe the consideration of both will fine tune the results to more specific CBG's. So, rather than use complex model or regression analysis for a simple goal, I utilized Moran's I clustering to identify the CBS's that were faced with the worst conditions (high-high or low-low clusters) according to each indicator.

3.1 Socioeconomics and Demographics

The socioeconomic and demographic variables are readily available at the CBG level on the U.S. Census tables website as CSV files (U.S. Census). Specifically, based on previous research studies (Colmer et al., 2020; Ge et al., 2017; Deguen and Zmirou-Navier, 2010; Martenies et al., 2017; Ouyang et al., 2018; Jbaily et al., 2022) the socioeconomic status and demographics of Detroit will be represented by their income, education, race and median age. Those are the variables that were found to be most indicative of air quality. So, I first downloaded the CSV files for each of the census variables: educational attainment for age 25 & over, household income in the past 12 months (2019 inflation adjusted), race, and Hispanic or Latino origin (US Census, 2019). These files don't have a spatial feature attached to them, instead just the CBG id ("GEOID") and the values for each of those variables. So, in parallel, I also downloaded the U.S. CBG boundary Shapefile (Census, 2017), which had a polygon for each CBG and its unique ID (GEOID). Although the "GEOID" from both files are meant to be the same, one has an extra id segment attached to it ("150000US"). This issue was resolved simply calculating a new field with the "GEOID" and the extra characters replaced with empty space. Once loading the variable tables and shapefile into ArcGIS, I joined each table by the attribute of "GEOID" to the CBG Shapefile. After joining each attribute to the CBG Shapefile, each polygon is attributed with race, income, education, and age. However, the age variable can easily be used for the purposes of clustering, but educational attainment, income, and race are provided as percentages of each bracket (Ex. 2% of CBG earns below 10K or 7% of CBG only has a high school diploma). So, to allow for clustering, I discretized the income and educational attainment variables. Firstly, I encoded each income bracket and educational level with a discrete representation (Ex. there are 16 income brackets meaning below 10k = 1 and above 200k = 16), such that the discrete variable represents the level appropriately. So, a lower educational attainment level and lower income bracket have lower discrete variables. Next, using "Calculate Field" on ArcGIS, I multiplied the percentage (of each tax bracket or of each educational level) with the corresponding discrete variable and summed the result up as a new field. Now, there is one number that represents the weighted mean of the educational attainment and another that represents income. It did not feel appropriate discretizing race, since the relative vulnerability has not been quantified, so instead it will be included in the overlap analysis through visual inspection.

Figure 1: Percent of Black or African American per CBG in Detroit, MI

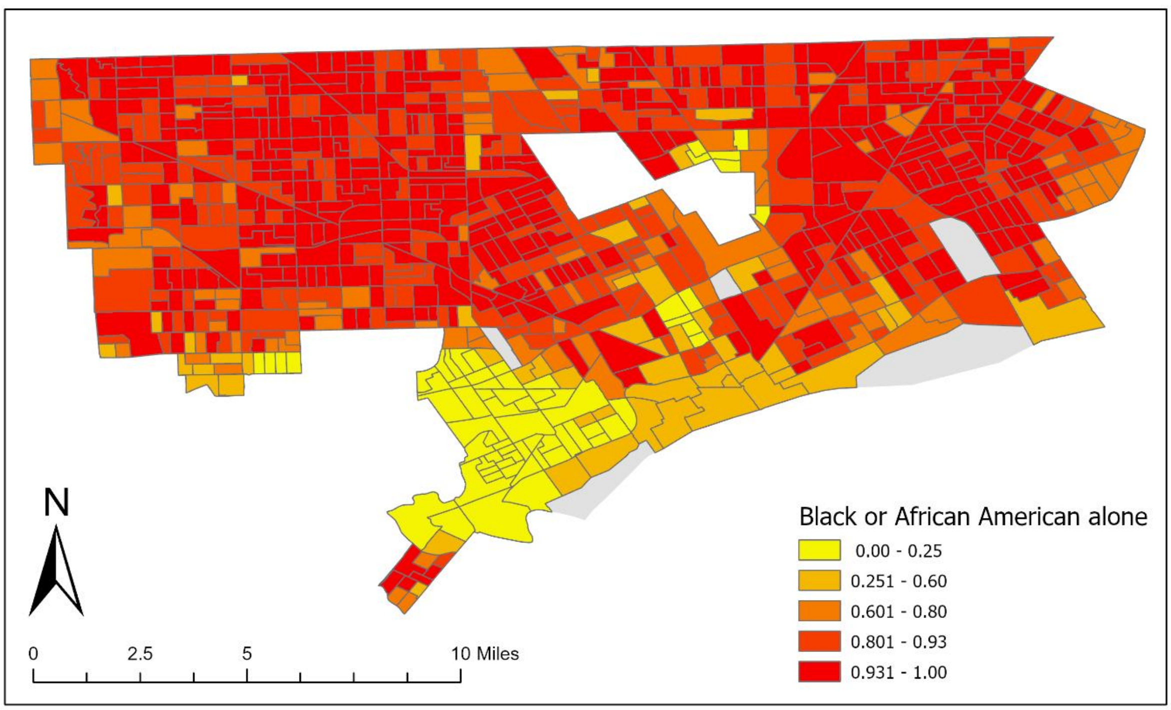


Figure 2: White Percentage per CBG in Detroit, MI

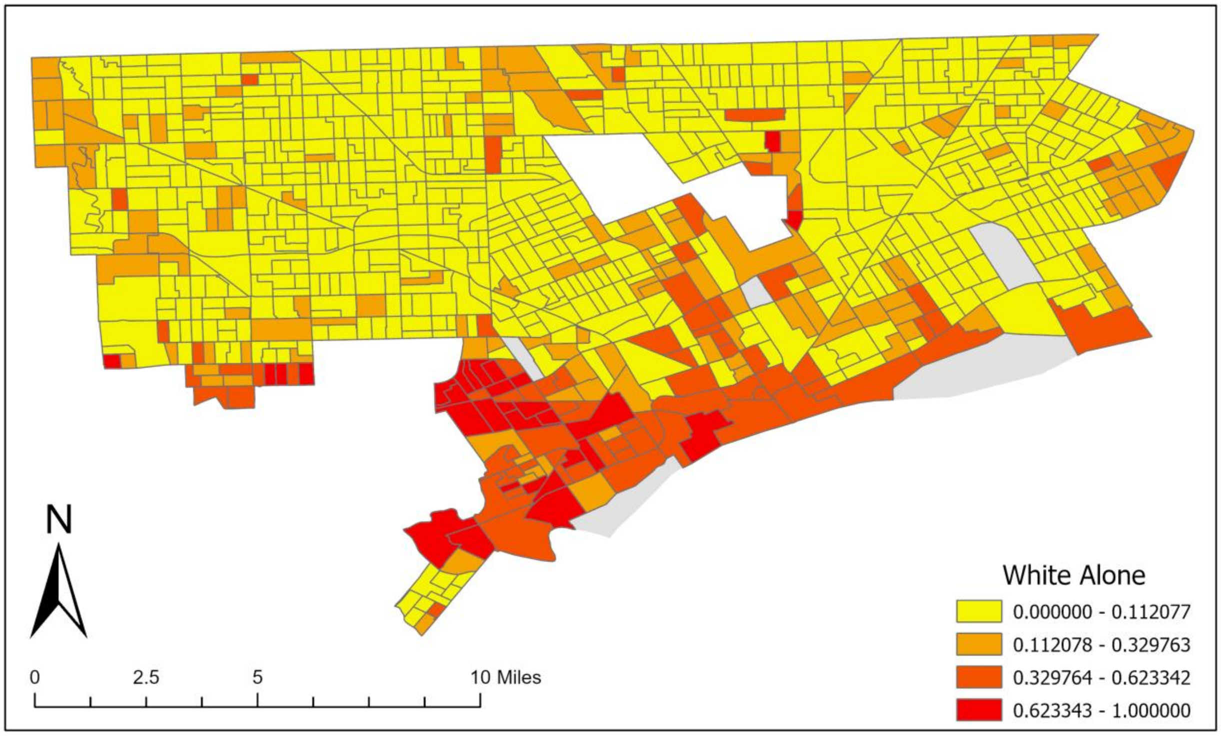


Figure 3: Hispanic or Latino Percentage per CBG in Detroit, MI

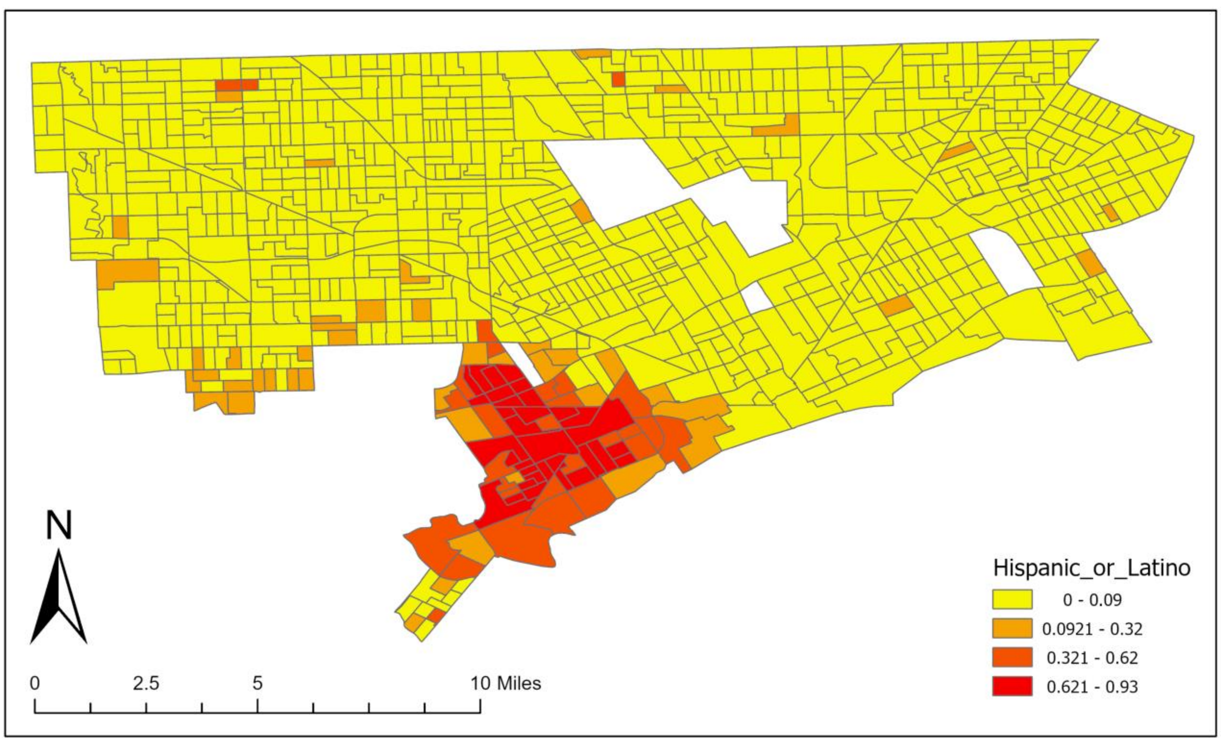


Figure 4: Weighted Mean of Educational Attainment per CBG in Detroit, MI

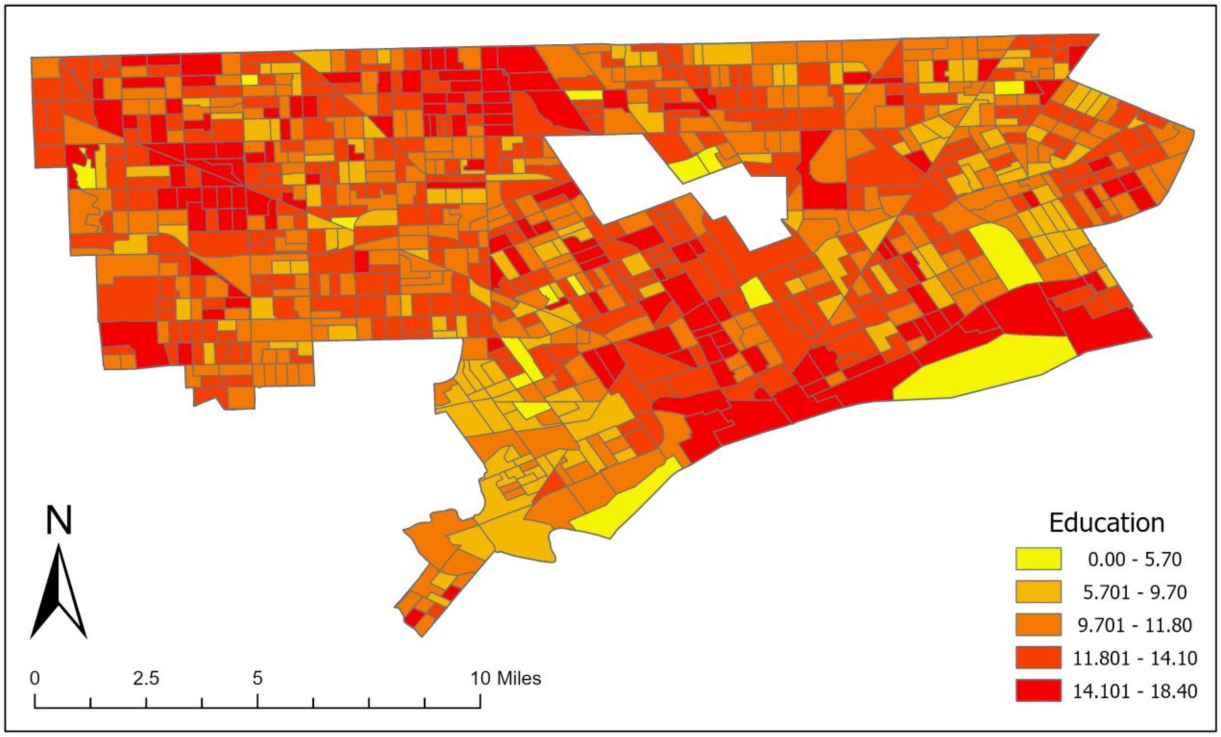
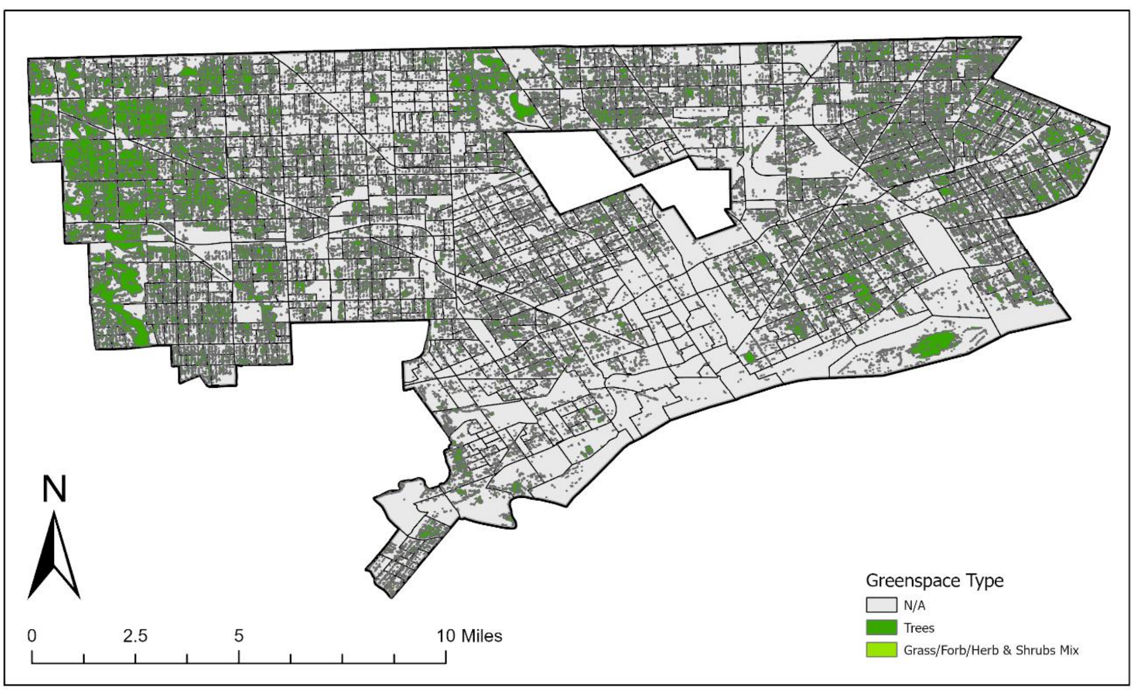


Figure 5: Weighted Mean of Income per CBG in Detroit, MI

3.2 Greenspace

Most of the necessary green space data is readily available on the Landscape Change Monitoring System (LCMS), developed by the USA Forest Service (Housman, 2022). The LCMS provided high-resolution “Land Cover” data, divided into 15 categories (trees, shrubs, grass, snow, etc.), for the contiguous U.S. as a tag image file (.tif) that can directly be loaded into ArcGIS. Although, the LCMS system has a resolution of yards, I aggregated to the percentage of trees, shrubbery, and grass for each CBG. To begin with, clipping a raster file with a polygon, of the Detroit city boundary (“Detroit City Boundary, 2022), resulted inaccurate results, so I first converted the raster file into a polygon (“Convert Raster to Polygon”), after which I was able to clip the polygons and preserve only the areas that encompasses Detroit. Any other polygon outside the boundary was spatially selected and deleted. Next, I removed all the polygons that represented anything other than “Trees” and “Grass/Forbs & Trees Mix” as they are irrelevant to our goal. Lastly, using the “Tabular Intersection” tool, I was able to ascertain the percentage of each land cover/greenspace type within each CBG zone. The resulting table was inconvenient to use, since there was row for each unique CBG and greenspace type pair. This issue was easily resolved by creating a copy of the table and then (self) joining with the original and removing the rows that contained the same land cover category twice. Lastly, for the purposes of clustering, I calculated the weighted mean of greenspace within each CBG, using a process identical to the one in used in section 3.1. Here, Trees received 2 and “Grass/Forbs & Trees mix” received a 1, because tree canopy has a more positive impact on greenspace.

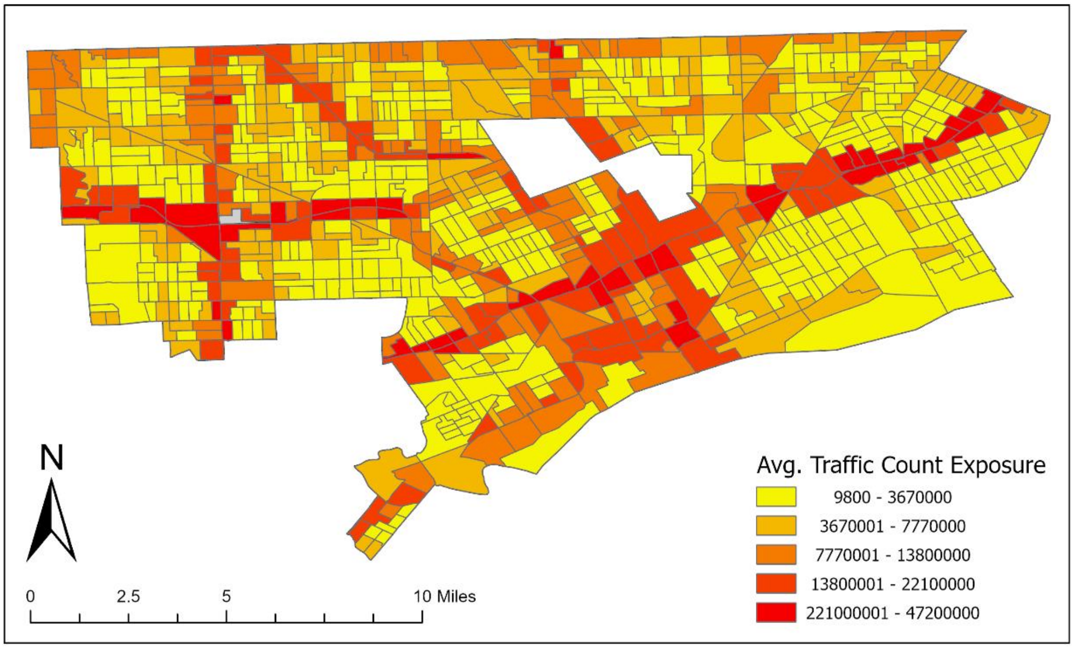
Figure 6: Greenspace per CBG in Detroit, MI



3.3 Traffic Density

Lastly, the Southeast Michigan council of governments (SEMCOG) built a detailed repository of the average annual daily traffic (AADT) counts for every street (east-bound and west-bound) in southeast Michigan (SEMCOG, 2019). The difficulty in analyzing this data lies in that the AADT is attributed individual road segments, and not every segment of a road. Similarly, segments belonging to the same road do not have the same traffic density, rather it depends on location. The cost of high resolution is the analytical difficulty. The SEMCOG dataset is provided as points that are attributed with the road's (the one that the segment belongs to) unique ID ("LRS") and the AADT. To accompany this dataset and build a full road network's traffic density, I downloaded as Shapefile of Michigan's road network from the ArcGIS online portal (Michigan DOT, 2022), which contained a polygon for every road segment and for each segment the corresponding ID ("PR") of the road that the segment belongs to. After a close manual inspection of the road network, I able to determine that a road or a highway with the same name doesn't necessarily mean that there is an equivalent unique ID to represent that entire road/highway. Rather every named road/high (like Route 10) may be broken up into multiple unique road ID's, which preserves locality. Given this revelation, I exported the table for the SEMCOG AADT feature class, and using *python*, I averaged the traffic density for every unique road ID ("LRS"), and imported the resulting table (containing average AADT for every road ID). From visual inspection, I was able to confirm that "LRS" represented the same value as "PR", so I joined the imported table, with the average AADT calculations, using the "LRS" field to the Michigan road network on the matching "PR" field. The resulting Shapfile

Figure 7: Avg. Traffic Count Exposure per CBG in Detroit, MI

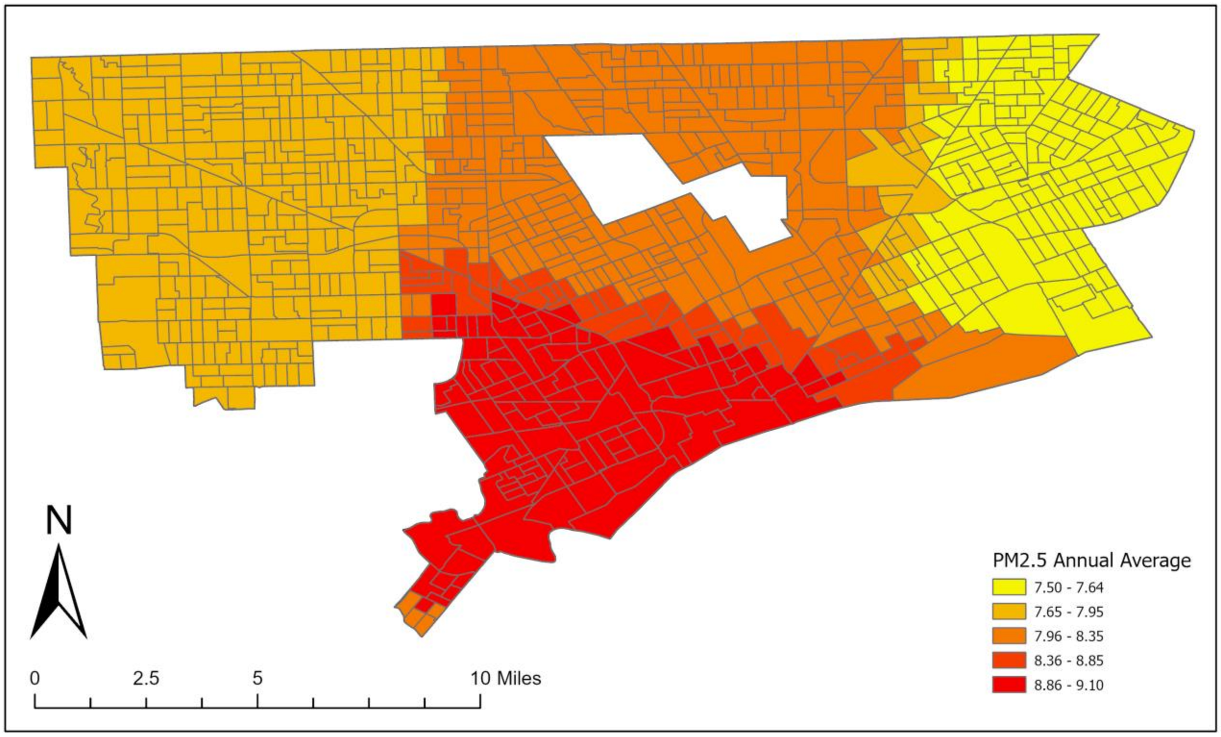


contained the AADT for every road, highway, ramp, arterial network, etc. that was listed in the SEMCOG data. Although this data is not complete, it has enough resolution to support effective analysis. Given the traffic density, the next step was to define exposure, which I did utilizing Kernel Density Estimation (KDE). (Pratt et al., 2014; Pratt et al., 2015). Based on review of academic research I determined that the resolution of the KDE would be 550ft. Using the "Kernel Density" tool on ArcGIS I was able to produce a raster of traffic exposure, meaning expected count represented as pixels. Using a process similar to the one described in section 3.3., I converted the raster to polygons, after which I used "Tabular Intersection" to calculate how much total traffic (expected) count each CBG (zone) is exposed to.

3.4 Air Quality

One of the biggest limitations to the study of air pollution is the lack of high-resolution data (Tessum et al. 2017), as urban concentrations significantly change at granular distances (<1km), but monitoring sites are relatively scarce (Apte et al., 2017). Even the Environmental Protection Agency (EPA) only has 4000 ambient monitoring stations across the U.S, which support at best county-level observations (EPA). So, for research purposes air pollution models are often utilized to simulate the concentrations at a high-resolution (Tessum et al., 2017; Martenies, 2017; Lane et al., 2022; Sabrin et al., 2020). For this GIS project, I used the Community Multi-Scale Air Quality System (CMAQ), which was developed by the EPA to use atmospheric science and air quality modeling to estimate concentrations of ozone, particulates (PM2.5 and PM10), and other toxins. Due to the complexity of the model, it would take a lot of time and computational power to run it from scratch. So, to make it easier for the end user, the EPA published their results for "daily average surface concentrations for 14 chemical species" from 2002 to 2017. To begin with I downloaded the files for 2017 and calculated the annual daily average for every unique coordinate using *python*, loaded the point data on to ArcGIS and then removed any points that were farther than 12km from the Detroit city boundary. The preprocessing results simply represent 12km by 12km grid cells that will be attributed with the coordinates and PM2.5 concentrations in micrograms per second. Next, to convert the coordinates to square grid cells, I first created a 12km circular buffer, then used the "Envelope Polygon" to create squares that covered the circular buffer and was centered at the coordinate. Once clipped with the Detroit city boundary, while only preserving what remains inside, I arrive at the end of the air quality layer. Note that the resolution is still fairly low, higher than raw data, but there are only 7 resulting grid cells to explain the gradient in Detroit. Although it is evident that the air quality alone could be used to identify a small subset of CBGs for policy planners to focus on. However, it does without our problem statement, since it does show a clear gradient amongst the grid cells. The urban core and the center section of the city have the worst PM2.5 concentrations. (Figure 1) Furthermore the results here could be fine-tuned by the other indicators to reduce or home in on the priority CBGs.

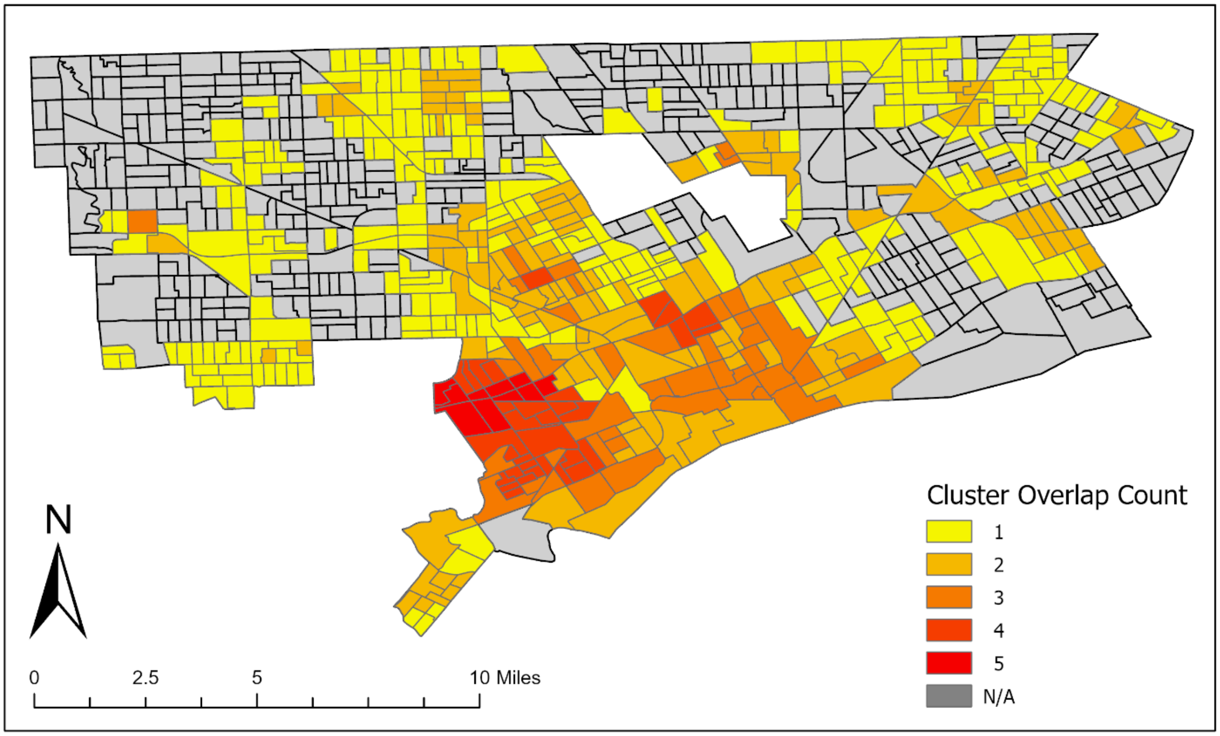
Figure 8: PM2.5 per CBG in Detroit, MI



3.5 Clustering Analysis with Moran's I

Clustering is an appropriate analysis tool in this scenario, because my goal is to help policy makers of Detroit identify areas of block groups that require the most attention for a policy enaction or funds. Ideally, it should have some gradient so as to indicate the level of priority as well. So, finding where greenspace, density, socioeconomic, demographic, and air quality variables cluster into low or high values would identify the troubled areas for each category. Then aggregating all the priority CBG's (clusters) will help create a single vulnerability map with the exclusion of race, which will be added through visual analysis. Race could not be included in the cluster analysis, since it can't be discretized relative to other races, as opposed to income and educational attainment. To begin the analysis, I simply used the "Cluster and Outlier Analysis (Anselin Local Moran's I)" to calculate the clusters and outliers of the feature class created for PM2.5 concentrations, greenspace, traffic density, income, age, and educational attainment. Each resulting feature class has high-high and low-low clusters, but the vulnerability concept for every variable is not the same. So, using queries I selected all the high clusters for air quality, low clusters for greenspace, high clusters for traffic density, low clusters for income, both high and low clusters for age, and low clusters for educational attainment, and remove the remaining clusters and all outliers. The resulting feature classes simply represent the most vulnerable CBGs with respect to each indicator (expect race). Lastly, all that was left to do was calculate the overlaps, meaning ask the question "how many indicators identified this CBG as vulnerable?" for each CBG. I answered the question with the "Count Overlapping Features" geoprocessing tool in ArcGIS. The results are displayed in Figure 9.

Figure 9: Vulnerability Cluster Overlap Count per CBG in Detroit, MI



4. **Findings and Conclusions**

4.1 Results

Examining the results of each individual layer, it is evident that Claytown, Springwells and the Eastern Market are the most vulnerable. For the most part, the outcomes were as expected. Starting with socioeconomics, the weighted mean of percent income and median age across the city was dispersed and varied; there was slight indication of lower income in Claytown, but it was not conclusive. (Figure 5) In contrast, it was clear from looking at the weighted mean of educational attainment (Figure 4) that Claytown and Springwells have the lowest levels of educational attainment compared to the rest of the city. It also happens to be the case that the areas to the south surrounding Springwells and Claytown have the highest percentage of White and Hispanic or Latino origin per CBG (Figures 2 & 3). Keep in mind however that White does include the Hispanic or Latino population in its counts. On the other hand, the remaining CBGs, which evidently have higher educational attainment levels, are heavily populated by the African American population (Figure 1). Moving on to greenspace, it is once again the case that the outer edges of the city (specifically to the east and west) have high concentrations of greenspace per CBG, but as we move closer to the urban core and southwards, the amount of greenspace diminishes and is sparse (Figure 6). In addition to a higher traffic density exposure in the Eastern Market and the downtown areas of Detroit, another pattern emerges; obviously the traffic exposure is much higher for those CBGs that reside next to a major highway (Figure 7). Detroit's highways spread evenly across the city. Lastly, even though there are very few divisions/buffers to describe the air quality (PM2.5) concentrations, the urban core and the surrounding CBGs have the poorest air quality (Figure 8). However, extending outward form the center there is an upward gradient and improvement in PM2.5 concentrations.

With the exception of income and median age, it is evident that the south of the city, not specifically the urban core, but the CBGs surrounding the Eastern Market, Springwells, and Claytown are dictated as the most vulnerable by the indicators of air quality and the air quality (PM2.5) itself (although general). Although valid, these conclusions are purely from visual inspection. Regardless, these conclusions were supported by overlapping cluster and outlier analysis with Moran's I. Once, I performed cluster analysis, and retrieved the low clusters for greenspace, income and education, high clusters for traffic density and air quality and both clusters for median age, I was able to count the number of overlaps, meaning asking "how many indicators classified this CBG as most vulnerable?" for each CBG. While the cluster analysis was singular in depth, the overlapping brough depth and consequently a gradient in vulnerability. Once again, this gradient confirms that the CBG's that face the most disparity with respect to poor air quality reside in Claytown. The vulnerability count diminishes, and in many cases there is not even an overlap, as we move to the outer edges of the city. Contrary to previous studies, the Black or African American population is not more vulnerable than the White and Hispanic population. Further research is required, but since there is a higher Hispanic or Latino population in the southernmost part of the city, similar to White percentages (a value which includes the Hispanic or Latino population as well), it can be reasonably said that a minority population is still at a disadvantage and faces worse air quality. This skew, w.r.t the better air quality, in predominantly African American CBGs is most likely caused by the significantly higher African American population in the city, in general. Since, we used clustering to identify the "most" vulnerable areas within the city, it doesn't not mean that the other CBGs have good air quality, rather they have improved air quality relative to the CBGs to the southern most areas of Detroit. So, I can confidently recommend that the CBG's in Claytown should be the focus of improvement from policy makers and urban planners.

4.2 Limitations

There were two significant limitations to this study. The resolution of both the air quality data and the traffic density data. While the resolution of the air quality data did not skew results, I would have been able to conclude with much more confidence given a higher resolution. Unlike, air quality the lower (relative) resolution of the traffic density data resulted in many smaller roads and streets not having a recorded AADT, and consequently a false exposure count. Lastly, discretization of race would've allowed for much more quantitative and concrete results, that would've provided more confidence in my results.

4.3 Conclusion

After gathering data on 4 different air quality indicators, as dictated by previous academic research: greenspace, traffic density, socioeconomics (income and education) and demographics (race and median age), I was able to establish vulnerability w.r.t each indicator using clustering and outlier analysis. An initial mapping of (relatively) high resolution air quality data provided a “base” map for poor air quality, the addition of each indicator fine-tuned the results and added a gradient of priority that would be useful for policy makers and urban planners. Once the final overlapping results were calculated and mapped, it was evident that the CBGs in and near the Claytown region of Detroit face the most disparity and require immediate attention with respect to the improvement of air quality.

5. References

English, Paul, Raymond Neutra, Russell Scalf, Moira Sullivan, Lance Waller, and Li Zhu. 1999. "Examining Associations between Childhood Asthma and Traffic Flow Using a Geographic Information System." Environmental Helath Perspectives 107, no. 9. https://doi.org/10.1289/ehp.99107761.

Apte, J. S., K. P. Messier, S. Gani, M. Brauer, T. W. Kirchstetter, M. M. Lunden, J. D. Marshall, C. J. Portier, R. C. H. Vermeulen, and S. P. Hamburg. 2017. "High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data." Environ Sci Technol 51 (12): 6999-7008. https://doi.org/10.1021/acs.est.7b00891. https://www.ncbi.nlm.nih.gov/pubmed/28578585.

Barnes, Joanna H., Tim J. Chatterton, and James W. S. Longhurst. 2019. "Emissions vs exposure: Increasing injustice from road traffic-related air pollution in the United Kingdom." Transportation Research Part D: Transport and Environment 73: 56-66. https://doi.org/10.1016/j.trd.2019.05.012.

Colmer, Jonathan, Ian Hardman, Jay Shimshack, and John Voorheis. 2020. "Disparities in PM_{2.5} air pollution in the United States." Science 369 (6503): 575-578. https://doi.org/doi:10.1126/science.aaz9353. Deguen, S., and D. Zmirou-Navier. 2010.

"Social inequalities resulting from health risks related to ambient air quality--A European review." Eur J Public Health 20 (1): 27-35. https://doi.org/10.1093/eurpub/ckp220. https://www.ncbi.nlm.nih.gov/pubmed/20081212.

English, Paul, Raymond Neutra, Russell Scalf, Moira Sullivan, Lance Waller, and Li Zhu. 1999. "Examining Associations between Childhood Asthma and Traffic Flow Using a Geographic Information System." Environmental Health Perspectives 107, no. 9. https://doi.org/10.1289/ehp.99107761.

Fernández, Ignacio C., and Jianguo Wu. 2018. "A GIS-based framework to identify priority areas for urban environmental inequity mitigation and its application in Santiago de Chile." Applied Geography 94: 213-222. https://doi.org/10.1016/j.apgeog.2018.03.019.

Ge, Yi, Haibo Zhang, Wen Dou, Wenfang Chen, Ning Liu, Yuan Wang, Yulin Shi, and Wenxin Rao. 2017. "Mapping Social Vulnerability to Air Pollution: A Case Study of the Yangtze River Delta Region, China." Sustainability 9 (1). https://doi.org/10.3390/su9010109.

Guo, L., J. Luo, M. Yuan, Y. Huang, H. Shen, and T. Li. 2019. "The influence of urban planning factors on PM2.5 pollution exposure and implications: A case study in China based on remote sensing, LBS, and GIS data." Sci Total Environ 659: 1585-1596. https://doi.org/10.1016/j.scitotenv.2018.12.448. https://www.ncbi.nlm.nih.gov/pubmed/31096368.

Housman, Ian. LCMS Data Explorer. U.S.A Forestry Service. Accessed April 6, 2022. https://apps.fs.usda.gov/lcms-viewer/.

Jbaily, A., X. Zhou, J. Liu, T. H. Lee, L. Kamareddine, S. Verguet, and F. Dominici. 2022. "Air pollution exposure disparities across US population and income groups." Nature 601 (7892): 228-233. https://doi.org/10.1038/s41586-021-04190-y. https://www.ncbi.nlm.nih.gov/pubmed/35022594.

Jerrett, Michael, Richard T. Burnett, Pavlos Kanaroglou, John Eyles, Norm Finkelstein, Chris Giovis, and Jeffrey R. Brook. 2016. "A GIS–Environmental Justice Analysis of Particulate Air Pollution in Hamilton, Canada." Environment and Planning A: Economy and Space 33 (6): 955-973. https://doi.org/10.1068/a33137.

Lane, Haley M., Rachel Morello-Frosch, Julian D. Marshall, and Joshua S. Apte. 2022. "Historical Redlining Is Associated with Present-Day Air Pollution Disparities in U.S. Cities." Environmental Science & Technology Letters. https://doi.org/10.1021/acs.estlett.1c01012.

Martenies, S. E., C. W. Milando, G. O. Williams, and S. A. Batterman. 2017. "Disease and Health Inequalities Attributable to Air Pollutant Exposure in Detroit, Michigan." Int J Environ Res Public Health 14 (10). https://doi.org/10.3390/ijerph14101243. https://www.ncbi.nlm.nih.gov/pubmed/29048385.

Mavroulidou, M., S. J. Hughes, and E. E. Hellawell. 2004. "A qualitative tool combining an interaction matrix and a GIS to map vulnerability to traffic induced air pollution." J Environ Manage 70 (4): 283-9. https://doi.org/10.1016/j.jenvman.2003.12.002. https://www.ncbi.nlm.nih.gov/pubmed/15016437.

Meerow, Sara, and Joshua P. Newell. 2017. "Spatial planning for multifunctional green infrastructure: Growing resilience in Detroit." Landscape and Urban Planning 159: 62-75. https://doi.org/10.1016/j.landurbplan.2016.10.005.

Ouyang, W., B. Gao, H. Cheng, Z. Hao, and N. Wu. 2018. "Exposure inequality assessment for PM2.5 and the potential association with environmental health in Beijing." Sci Total Environ 635: 769-778. https://doi.org/10.1016/j.scitotenv.2018.04.190. https://www.ncbi.nlm.nih.gov/pubmed/29710600.

Pratt, G. C., K. Parson, N. Shinoda, P. Lindgren, S. Dunlap, B. Yawn, P. Wollan, and J. Johnson. 2014. "Quantifying traffic exposure." J Expo Sci Environ Epidemiol 24 (3): 290-6. https://doi.org/10.1038/jes.2013.51. https://www.ncbi.nlm.nih.gov/pubmed/24045427.

Pratt, G. C., M. L. Vadali, D. L. Kvale, and K. M. Ellickson. 2015. "Traffic, air pollution, minority and socio-economic status: addressing inequities in exposure and risk." Int J Environ Res Public Health 12 (5): 5355-72. https://doi.org/10.3390/ijerph120505355. https://www.ncbi.nlm.nih.gov/pubmed/25996888.

Sabrin, Samain, Maryam Karimi, and Rouzbeh Nazari. 2020. "Developing Vulnerability Index to Quantify Urban Heat Islands Effects Coupled with Air Pollution: A Case Study of Camden, NJ." ISPRS International Journal of Geo-Information 9 (6). https://doi.org/10.3390/ijgi9060349.

Schindler, Mirjam, Judith Y. T. Wang, and Richard D. Connors. 2021. "A two-stage residential location and transport mode choice model with exposure to traffic-induced air pollution." Journal of Transport Geography 93. https://doi.org/10.1016/j.jtrangeo.2021.103044.

Servadio, J.L., Lawal, A.S., Davis, T. et al. Demographic Inequities in Health Outcomes and Air Pollution Exposure in the Atlanta Area and its Relationship to Urban Infrastructure. J Urban Health 96, 219–234 (2019). https://doi.org/10.1007/s11524-018-0318-7

Sun, Chuanwang, Shuhua Xu, Mian Yang, and Xu Gong. 2022. "Urban traffic regulation and air pollution: A case study of urban motor vehicle restriction policy." Energy Policy 163. https://doi.org/10.1016/j.enpol.2022.112819.

Tessum, C. W., J. S. Apte, A. L. Goodkind, N. Z. Muller, K. A. Mullins, D. A. Paolella, S. Polasky, N. P. Springer, S. K. Thakrar, J. D. Marshall, and J. D. Hill. 2019. "Inequity in consumption of goods and services adds to racial-ethnic disparities in air pollution exposure." Proc Natl Acad Sci U S A 116 (13): 6001-6006. https://doi.org/10.1073/pnas.1818859116. https://www.ncbi.nlm.nih.gov/pubmed/30858319.

Tessum, C. W., J. D. Hill, and J. D. Marshall. 2017. "InMAP: A model for air pollution interventions." PLoS One 12 (4): e0176131. https://doi.org/10.1371/journal.pone.0176131. https://www.ncbi.nlm.nih.gov/pubmed/28423049.

Xia, Liping, and Yaping Shao. 2005. "Modelling of traffic flow and air pollution emission with application to Hong Kong Island." Environmental Modelling & Software 20 (9): 1175-1188. https://doi.org/10.1016/j.envsoft.2004.08.003.

“Air Data Basic Information.” EPA. Environmental Protection Agency. Accessed April 6, 2022. https://www.epa.gov/outdoor air-quality-data/air-data-basic information#:~:text=Monitoring%20Data%20%2D%20Ambient%20(outdoor),mainly%20by%20state%20environmental%20agencies.

Explore census data for Michigan. U.S. Census. Accessed April 6, 2022. https://data.census.gov/cedsci/all?q=population&g=0400000US26%241000000.

SEMCOG. Traffic counts. Southeast Michigan Council of Governments. Accessed April 6, 2022. https://www.semcog.org/traffic-counts

US Census. 2017. “2017 Tiger/Line® Shapefiles: Block+Groups.” United States Census Bureau. Accessed May 6, 2022. https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2017&layergroup=Block%2BGroups.

Michigan Department of Transportation. GIS Open Data. Accessed May 6, 2022. https://gis-michigan.opendata.arcgis.com/datasets/5fddb53b929b4b729b8d4282b4d23ade.

American Lung Association “Most Polluted Cities: State of the Air.” State of the Air. Accessed May 6, 2022. https://www.lung.org/research/sota/city-rankings/most-polluted-cities.

US EPA Office of Research and Development. “CMAQ.” Zenodo. Enviornmental Protection Agency, August 17, 2021. https://zenodo.org/record/5213949#.YnVGhofMK3A.

“Detroit City Boundary.” Arcgis.com. Accessed May 6, 2022. https://www.arcgis.com/home/item.html?id=5ddcdd8528d349629d8b6bb4d0e27823.

Bureau, U.S. Census. “US Cenus Tables for Wayne County.” Explore census data. U.S Census. Accessed May 7, 2022. https://data.census.gov/cedsci/table?g=0500000US26163%241500000&y=2019.

“Detroit, MI.” Detroit, MI Data. Data USA. Accessed May 7, 2022. https://datausa.io/profile/geo/detroit-mi/.

6. Appendix

6.1 Data Description Table

Spatial Layer	Components	Source	Relevant Citations	Geography Covered	Scale	Projection	File Type
Greenspace	Greenspace	USDA	(Housman, 2022; LCMS)	United States	United States	WGS 1984	Raster (.tiff)
Air Pollutant Concentrations	Air Pollutant Concentrations	EPA’s CMAQ Model	(US EPA Office of Research and Development, Byun et al., 2006)	United States	12km x 12km grid	WGS 1984	CSV (.csv)
Traffic Density	Traffic Density Counts	Southeast Michigan Council of Governemnts (SEMCOG)	(SEMCOG)	Southeast Michigan Region	Exact Road Coordinates	WGS 1984	CSV (.csv)
	Road Network	Michigan Department of Transportation	(Michigan Department of Transportation)	Michigan	Every Road	WGS 1984	Shapefile (.shp)
Socioeconomics and Demographics	Socioeconomic and Demographic variables	2015 ACS Survey	(US Census)	Michigan	CBG	WGS 1984	CSV (.csv)
	Census Block Groups (CBG)	US Census	(US Census, 2017)	Michigan	CBG	NAD 1983	Shapefile (.shp)
	Detroit City Boundary	ESIRI	("Detroit City Boundary"	Detroit	Detroit	NAD 1983	Shapefile (.shp)
Base Map	World Topographic Map	ArcGIS Default	N/A	World	CBG	WGS 1984	Vector Tile Service

Table 1: Spatial Layer Descriptions

Table 2: Spatial Layer Descriptions

Spatial Layer	Components	Spatial Features	Attributes	Preprocessing Steps	Limitations
Greenspace	Land Cover	Pixels of Land Cover	Land Cover Type (Figure ___)	N/A	Slightly overestimates the amount of tree canopy. For example, it captures a house or two when capturing trees in between houses/backyards.
Air Pollutant Concentrations	Air Pollutant Concentrations	Point for each grid	Latitude, Longitude, Date, PM2.5 Concentration (micrograms per second)	Clean CSV file*. Average the concentrations over a year for each unique coordinate.	Resolution is low. Only 6 records for all of detroit.
Traffic Density	Traffic Density Counts	Point for each road segment (Not entirety of a road)	Road Unique ID (LRS ID), Annual Average Daily Traffic	Clean CSV file*	Although relatively high in resolution, a better dataset would have counts for every road segment of every road.
	Michigan Road Network	Polygon of each road segment (Not entierty of a road)	Road Unique ID (PR ID)	N/A	N/A
Socioeconomics and Demographics	Socioeconomic and Demographic variables	N/A	CBG ID, Race, Age, Income, Education	Project to WGS 1984	N/A
	Census Block Groups (CBG)	Polygon of CBG	CBG ID	Project to WGS 1984	N/A
Base Map	World Topographic Map	Raster	N/A	N/A	N/A