Opioid Risk Environment Regionalization:

A Within Environment Approach

Ashlynn I. Wimer

Division of the Social Sciences, University of Chicago

GISC 28400: GIScience Practicum

Dr. Yue Lin

May 24, 2024

Author Note

All data and code can be found at the project GitHub repository:
https://github.com/bucketteOfIvy/chicago-places

Opioid Regionalization in Chicago Using Found Data

The opioid death rate has increased steadily since 2002, reaching 32.6 deaths per 100,000 standard population in 2022 (Spencer, 2024). The increasing rate of the epidemic and the amount of harm caused have spurred many computational researchers to undertake predictive studies of opioid overdose events (OOE) and opioid deaths. This task has proven to be a challenging, with many studies achieving low predictive values when using common models (Cuomo et al., 2023; Gavali et al., 2021; Schell et al. 2022). Studies using ensemble methods or novel data sources have managed to achieve larger predictability on known data (Bozorgi et al., 2021; Li et al., 2022), but the studies on the results of placing harm reduction resources based on the advice of predictive models are still forthcoming (Marshall et al., 2022). These prediction studies are vital contributions from computational scientists seeking to help alleviate the harms of the opioid epidemic and are potentially capable of addressing urgent opioid overdose event flareups. However, as the studies aim to nowcast opioid overdose events, they have limited ability to place resources and investment which can address more deeper causes of opioid risk.

Qualitative and theoretical studies offer multiple frameworks for understanding these base causes. One of the most common frameworks for approaching opioid risk is the opioid risk environment (ORE) framework, which conceives of opioid risk – defined as the chance of suffering opioid related harm and the amount of harm suffered – as being contingent on social and environmental factors (Rhodes, 2002). At its most basic, the ORE framework consists of four environments – the social, physical, economic, and political environments – operating at the micro- (individual or community) and macro- (systemic) scales (Rhodes, 2002). These environments and scales are understood to interact. Factors at the macropolitical level, such as restrictions on needle availability, influence social at the microsocial level, such as needle sharing tendencies and norms. The ORE framework thus offers a lens through which to understand the causes and potential interventions for opioid risk and harm.

Moreover, the ORE framework intentionally leaves room for modification and fuzziness – it's unclear, for example, whether the social norms of a local drug market is a macro or micro level factor – which has allowed for numerous modifications on the framework. Duff (2007), for instance, uses the framework to discuss the importance of context on drug risk and harms, while Collins et al. (2019) argues for adding an intentional intersectional lens to the framework. Modifications are often to better fit qualitative studies. Cooper et al. (2020), for instance, reintroduce a meso- scale to the framework (which is defined as the scale which mediates the influence of the macro- environments on the micro- environments) in order to understand the influence of the pharmaceutical legislation on buprenorphine access in Appalachian Kentucky.

Potentially owing to its limited use in computational studies, the ORE framework is rarely deployed or modified for deployment in computational and spatial studies. The primary and potentially only computational study conducted within an opioid risk environment paradigm was a rural regionalization analysis in Southern Illinois conducted by Kolak et al. (2020). Whereas predictive studies are vital for the ability to determine where to place stopgap resources, regionalization analyses provide potential pathways towards understanding the spatial distribution of root influences of opioid risk and could enable intentional investment and placement of longer-term harm reduction resources. Additionally, explicitly spatial computational studies using historical data have could enable researchers to identify historical influences on the modern opioid risk environment. However, a drought of computational studies limits our understanding of the best practices for and potential utility of computational studies on opioid risk environments.

Towards alleviating this drought, we undertake a study of the opioid risk environment in Chicago. Prior research by Kolak et al. (2020) utilized an *across* environment approach, finding regions for the overall risk environment instead of for each aspect of the risk environment. We instead utilize a *within* environment approach, which we believe may allow for deeper characterization of the risk environments' spatial distribution.

To do so, we use regionalization methods to find risk regions within the social, economic, and physical risk environments within Chicago. In order to explicitly validate the utility of our approach, we undertake two investigations on our resultant regions. First, we investigate the extent to which our identified high-risk regions align with areas that contain large amounts of narcotic related arrests, which we use as a proxy for opioid risk. We expect that, if our regions are reasonable reflections of opioid risk, then areas with factors that the opioid risk environment framework views as increasing risk should have larger amounts of opioid risk than areas that are indicated as safe.

Secondly, we investigate the extent to which our approach can illuminate historic influences on modern opioid risk through comparison to historic redlining grades. Redlining was a practice of discriminatory home loan lending and community investment instantiated in the first half of the 19th century, and has been linked to numerous modern health outcomes, including gunshot victims, preterm birth, late-stage lung cancer diagnosis, and infant mortality (Swope et al., 2022). Based on these prior correlations, we expect that areas of historic disinvestment, so red- and yellow-lined areas, will feature larger proportions of high risk opioid regions than blue- and green-lined areas.

Data

We acquire a large amount of health and environment related data in an attempt to quantitatively characterize the social, economic, and physical opioid risk environments (Table 1). While many of our variables, such as these data sources originate from traditional data sources such as the American Community Survey and CDC, others come from less traditional sources and require more extensive preprocessing, the process of which we discuss individually below.

Street Level Greenery

Everyday exposure to plants, trees, and other flora (“greenery”) has been shown to have positive mental health impacts, which hints that it could be preventative of opioid risk (CITE).

We thus construct a relative street-level relative greenery index for the city of Chicago. To create this dataset, we first uniformly select 25,000 points on Chicago's street network, and assign a random heading between 0° (North) and and 359° (North, but slightly east) to each point. We then attempt to acquire Streetview imagery for every point and every associated direction from the Google StreetView Static Imagery API, and successfully retrieve 24241 street level statics. Every pixel in each image is then assigned a semantic category through the MIT ADE20k semantic image segmentation algorithm (Zhou et al., 2017, 2018), allowing us to calculate an absolute greenery value G_{img} for each image img by summing up every pixel assigned to "tree", "grass", "field", "flower", or "hill." Lastly, we calculate a relative greenery index g_{img} through the following equation:

$$g_{img} = \frac{G_{img} - \min_{j \in Images} G_j}{\max_{j \in Images} G_j - \min_{j \in Images} G_j}.$$

As the minimum amount of greenery observed in any of our images was 0 greenery, a relative greenery index of 0 thus indicates that no greenery is in the image, while a relative greenery index of 1 means that the image contains the most greenery in any of the collected images.

311 Data

We download two datasets from the City of Chicago Data Portal: 311 requests and crime data from 2023. We use our 311 data to proxy social, economic, and physical disorder. To do so, we assign each request to one of three categories depending on whether the request reflects social concerns (e.g. a vehicle in bike lane request), economic concerns, (e.g. complaints about inaccurate retail sales), or physical concerns (e.g. a complaint about a dysfunctional traffic signal) (Appendix A). We then use a point-in-polygon analysis to separately count the number of physical, social, and economic 311 requests occurring in each Census Tract within Chicago. We then convert these to three rate variables – social 311 requests per 1,000 population, economic 311 requests per 1,000 population, and physical 311 requests per 1,000 population – which we use to proxy social, economic, and physical disorder, respectively.

Table 1

Anticipated categorizations of opioid risk environment variables

Environment	Concept	Construct	Data Source
Social	Social Disorder	Selected 311 requests per 1000 population	Chicago Data Portal 2024
	Educational Attainment	Percent of population with at most a high school degree	ACS 2018-2022
	Social Vulnerability	Social Vulnerability Index	CDC 2022
	Exposure to Violence	2022 Violent Crime Per 1000s Persons	Chicago Data Portal 2024
	Race Composition	% White Alone, Black Alone, and Asian Alone	ACS 2018-2022
	Age Composition	% 18-24 Years Old, %>65 Years Old	ACS 2018-2022
Economic	Economic Disorder	Selected 311 requests counts per 1000 population	Chicago Data Portal 2024
	Socioeconomic Status	Median Income, % Poverty, % Unemployment	ACS 2018-2022
	High Risk Jobs	Percent of working age population employed in agriculture, forestry, fishing, hunting, mining, quarrying, oil extraction, gas extraction, construction, manufacturing, or utilities	ACS 2018-2022
	Internet Access	% Households with Internet Access	ACS 2018-2022
Physical	Physical Disorder	Selected 311 request counts per 1000 population	Chicago Data Portal 2024
	Building Vacancy	Percent of vacant or unoccupied buildings	ACS 2018-2022
	Street Level Greenery	Relative greenery seen on Google Streetview Imagery	Google Streetview 2024

Crime Data

We use our crime data to construct a proxy for exposure to violence and to construct a proxy for high opioid risk. To proxy exposure to violence, we first classify crimes whose primary type is one of battery, homicide, assault, robbery, criminal sexual assault, sex offense,

kidnapping, and human trafficking as violent crimes, before conducting a point-in-polygon analysis to get the sum of violent crimes in each census tract which we then convert into rate of violent crimes per 1000 population. We similarly classify any crime whose primary offense is "narcotics" or "other narcotic violation" as a narcotic related crime and conduct another point-in-polygon analysis to find the number of narcotic crimes in each census tract.

Our analysis consists of three steps: a clustering step, a validation step, and an experimental step. For our clustering step, we conduct a regionalization analysis of opioid risk in Chicago. We specifically utilize an agglomerative clustering with queen contiguity in which we select the optimal number and type of clusters based on their interpretability and overall performance across three metrics – Calinski-Harabasz, Davies-Bouldin, and Silhouette – resulting in three sets of regions – one each for our social, economic, and physical risk environments. With our clusters in hand, we then validate the names and interpretations of our clusters by comparing the population-weighted occurrence of narcotic arrests between our different typologies. We argue that, if our cluster risk interpretations are sensible, then it should be the case that our higher risk clusters will feature greater proportions of narcotic arrests per 1000 population than our lower risk clusters. Lastly, we assign each census tract to a Home Owners Loan Corperation (HOLC) zone based on its centroid and conduct a chi-square test to investigate potential impact of historic redlining grades on modern opioid risk. To do so, we utilize the HOLC shapefile provided by Nelson and Winling (2023). Additionally, we aggregate the A and B grades and the C and D grades to ensure our categories have at least five members. All code and data used in our analysis can be found in the project's GitHub repository, linked on the cover page.

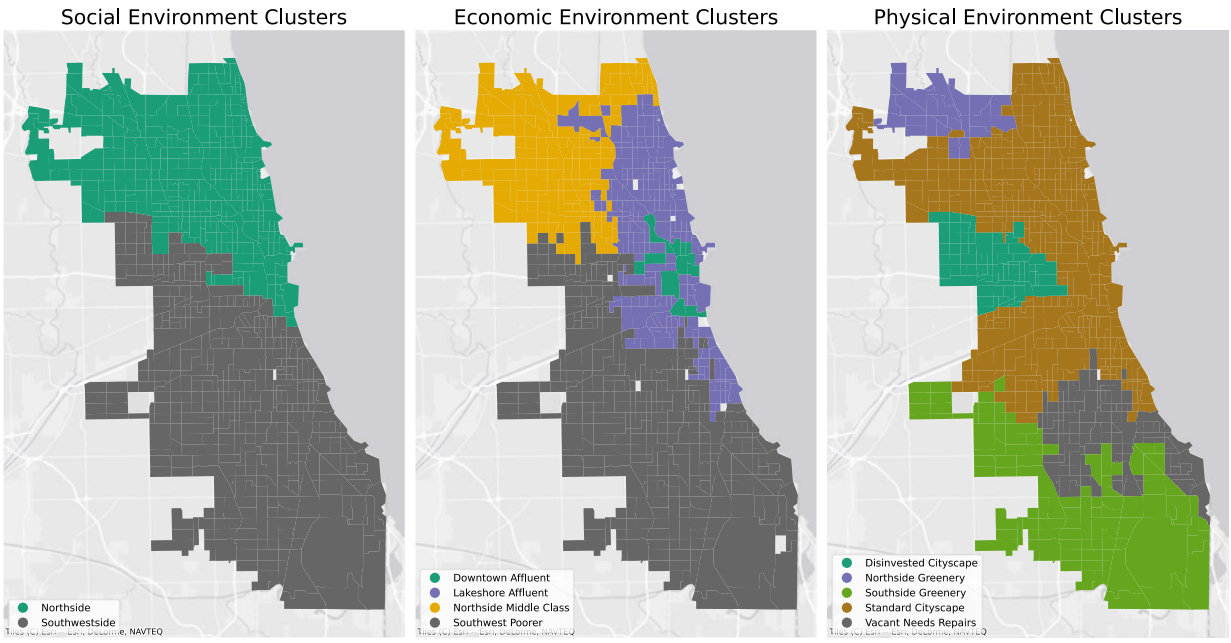
Results

Picking clusters based on those with the best performing metrics, we find two social environment clusters, four economic environment clusters (excluding a fifth singleton), and five physical environment clusters (Figure 1). We identify specific regions corresponding to higher

opioid risk within each environment, and successfully validate their higher risk through comparison to narcotic crime rates. Additionally, our chi-square test results indicate that historically C and D lined areas have more higher risk regions than A and B lined areas (Table 3).

Figure 1

Side by side maps of risk environment regions.



Social Environment Cluster

We identify a Northside social environment regions which features fewer 311s, a more highly educated population, a lower SVI, lower exposure to violent crime, and a population that is on average more white than the rest of the city, and a corresponding southwest side cluster which has a larger quantity of 311s, is one average less educated, has a higher average SVI, has larger minority populations, and has a larger exposure to violent crime (Table 2). The social disorder and exposure to violent crime ratios indicate that the Southwest side cluster would likely be a higher risk area, a prediction validated by its larger narcotic crime rate per 1000 population (Table 2). Additionally, this region makes up a greater portion of areas historically graded C and D (Table 3; $p = 0.15$).

Economic Environment Cluster

We identified four economic risk regions – a Downtown Affluent cluster, a Lakeshore Affluent cluster, a Northside Middle Class cluster, and a Southwest Lower Class cluster (Table 2). We identify two clusters as higher risk clusters: the Downtown Affluent Cluster as higher risk due to its larger economic disorder measure, and the Southwest Lower Class cluster due to its lower overall wealth. These trends are reflected in the narcotic crime rate per 1000 seen in Table 2. Additionally, we find that these higher risk clusters appear more often in historically C and D graded areas than they do in A or B graded areas (Table 3; $p < 0.001$).

Table 2

Cluster summary statistics and narcotic crime rate per 1000 population.

	Social Risk Environment Clusters			
Variable	**Northside**		**Southwest Side**	
	M	SD	M	SD
Social 311s	-0.121	0.862	0.102	1.095
Percent w/ No More Than a High School Diploma	-0.613	0.839	0.518	0.817
Social Vulnerability Index	-0.613	0.839	0.518	0.817
Violent Crime Rate	-0.614	0.994	0.519	0.651
Percent White Alone	0.803	0.603	-0.676	0.731
Percent Black Alone	-0.698	0.248	0.590	1.018
Percent Asian Alone	0.303	0.854	-0.257	1.044
Percent Hispanic	-0.015	0.792	0.012	1.148
Percent Between Ages 18 and 24	-0.0169	1.083	0.014	0.926
Percent Age >65	-0.319	0.994	0.269	0.925
Count	361		427	

Narcotic Crime Rate Per 1000	0.914	2.917

Economic Risk Environment Clusters

Variable	Downtown Affluent		Lakeshore Affluent		Northside Middle Class		Southwest Lower Class	
	M	SD	M	SD	M	SD	M	SD
Economic 311s Rate Per 1000	3.612	1.47	-0.063	0.848	-0.405	0.646	0.118	0.935
Median Income	0.063	<0.001	-0.069	1.453	0.062	<0.001	0.0618	<0.001
Percent Population in Poverty	-0.778	0.461	-0.561	0.576	-0.311	0.534	0.593	1.102
Percent Unemployed	-0.604	0.508	-0.617	0.472	-0.488	0.408	0.715	1.049
Percent Working High Risk Jobs	-0.934	0.365	-0.785	0.546	0.315	0.806	0.436	1.004
Percent Without Internet at Home	-0.717	0.974	-0.737	0.593	-0.009	0.668	0.601	0.994
Count	13		246		176		343	
Narcotic Crime Rate Per 1000	3.958		0.602		1.298		3.119	

Physical Risk Environment Clusters

Variable	Disinvested Cityscape		Northside Greenery		Southside Greenery		Standard Cityscape		Vacant Need Repairs	
	M	SD	M	SD	M	SD	M	SD	M	SD
Physical 311s	0.905	0.789	0.116	0.329	0.764	0.892	-0.524	0.585	0.899	1.188
Relative Greenery	-0.134	0.618	2.431	1.563	1.296	1.050	-0.380	0.582	-0.040	1.188
Percent Vacant Lots	0.533	0.945	-0.551	0.557	-0.177	1.053	-0.329	0.606	1.339	1.149
Count	74		25		103		474		111	
Narcotic Crime Rate Per 1000	11.644		0.357		1.050		0.906		2.738	

Note: With the exception of the Count and Narcotic Crime Rate Per 1000 Population variables, all values in this table are from standardized distributions (i.e. 0 is mean, positive is above mean, and -1 is one standard deviation below the mean).

Physical Environment Cluster

We identify five physical environment regions: a Disinvested Cityscape region, Northside Greenery, Southside Greenery, Standard Cityscape, and a region which consists of largely Vacant Buildings which may Need Repairs. *A priori* the Disinvested Cityscape region and the Vacant Needs Repairs region would be expected to have higher relatively higher narcotic crime rates per 1000 population due to their higher physical disorder, lower relative greenery, and higher vacant lot totals. While this is true of both regions, the Disinvested Cityscape region's narcotic crime rate is notably larger than that seen in any other region, suggesting that some either space is a vital factor or that differentiating features may be missing from the model (Table 2). Additionally, we note that these regions once again appear more often in historically C and D lined areas than they do in historically A or B lined areas (Table 3; $p < 0.001$).

Table 3

Risk environment cluster and counts by HOLC Grade.

Social Risk Environment Clusters

Grade	Northside		Southwest Side	
	N	%	N	%
A or B	41	62.1	25	37.9
C or D	265	45.5	317	54.5

Economic Risk Environment Clusters

Grade	Downtown Affluent		Lakeshore Affluent		Northside Middle Class		Southwest Lower Class	
	N	%	N	%	N	%	N	%
A or B	0	0.00	12	18.2	31	47.0	23	34.8
C or D	6	1.04	194	33.7	129	22.4	247	42.9

Physical Risk Environment Clusters

Grade	Disinvested Cityscape		Northside Greenery		Southside Greenery		Standard Cityscape		Vacant Needs Repairs	
	N	%	N	%	N	%	N	%	N	%
A or B	0	0.00	10	15.1	12	18.2	33	50	11	16.7
C or D	65	11.1	12	2.06	55	9.45	365	62.7	85	14.6

Discussion

Our regionalization attempt resulted in clusters which seem to both successfully identify aspects of opioid risk while also revealing longer standing influences of historic disinvestment.

Our discovered regions successfully discriminate regions of higher opioid risk from regions of lower opioid risk, demonstrating the potential of machine learning methods in charting out the boundaries of the opioid risk environment. While this is true of our social, economic, and physical environment regions, our physical environment regionalization is most reflective of our proxy for contemporary opioid risk, with the Disinvested Cityscape region – located in the West Side of the city – containing the largest narcotic crime rate per 1000 population of *any* region discovered in our analysis. Excluding the spatial location of the region, none of the regions summary statistics mark it as particularly different than the "Vacant Needs Repairs" region, indicating that variables outside of our current model may be at play. This opens the door for future research considering additional aspects of the physical environment, such as through traffic (i.e. perhaps this region sits along a larger travel corridor, allowing for easier access to opiates) or perceived aspects of the built environment (e.g. perhaps this region seems less lively, making it feel like a relatively safe space to inject).

Across all of our environments, the highest risk regions appear most frequently in C and D graded areas. These areas are those with the least historic investment from the Homeowners Loan Corporation, indicating that the echoes of historic disinvestment may still result in additional opioid risk to in modern times.

While our study has demonstrated spatially robust results, we are nonetheless limited by our data and methods. The largest of these is our proxy for opioid risk – narcotic crime rates – which persists of point data indicating that arrests were made due to opioid use or possession. Although this variable proxies some aspects of opioid risk, it is an imperfect tool for validation; if certain areas are more heavily policed than others, we would expect this proxy to overreport opioid risk in the heavily policed areas. Furthermore, narcotic arrest rates indicate are useful

proxies for aspects of the social risk environment; Cooper et al. (2020), for instance, found that increased DEA presence and involvement in their studied communities often lead to less buprenorphine availability. Future studies can improve on this limitation through data usage agreements with local government agencies in order to acquire opioid overdose events (OOE) data, which can serve as a proxy that sidesteps this issue. Other limitations to our study include our relatively limited data table, particularly in the physical risk environment (which features only three variables), and error inherent in variables derived from ACS sources. Future studies can improve on these limitations through the usage of larger varieties of data or through more robust data sources.

Despite our limitations, this study demonstrates the potential for understanding the distribution of the opioid risk environment through a *within environment* regionalization approach. Using this approach, we identify and validate regions of opioid risk within Chicago, and discover potential historical correlations with these regions. Given the critical need to address the ongoing opioid epidemic and to build infrastructure to prevent future opioid epidemics, these successes are hopeful, and may be fruitful bases for future research in different regions or using more robust data sources.

References

Bozorgi, P., Porter, D. E., Eberth, J. M., Eidson, J. P., & Karami, A. (2021). The leading neighborhood-level predictors of drug overdose: A mixed machine learning and spatial approach. *Drug and Alcohol Dependence*, *229*, 109143. https://doi.org/10.1016/j.drugalcdep.2021.109143

Cooper, H. L., Cloud, D. H., Freeman, P. R., Fadanelli, M., Green, T., Van Meter, C., Beane, S., Ibragimov, U., & Young, A. M. (2020). Buprenorphine dispensing in an epicenter of the U.S. opioid epidemic: A case study of the rural risk environment in Appalachian Kentucky. *International Journal of Drug Policy*, *8*. https://doi.org/10.1016/j.drugpo.2020.102701

Cuomo, R., Purushothaman, V., Calac, A. J., McMann, T., Li, Z., & Mackey, T. (2023). Estimating County-Level Overdose Rates Using Opioid-Related Twitter Data: Interdisciplinary Infodemiology Study. *JMIR Formative Research*, *7*. Scopus. https://doi.org/10.2196/42162

Duff, C. (2007). Towards a theory of drug use contexts: Space, embodiment and practice. *Addiction Research and Theory*, *15*(5), 503–5019. https://doi.org/10.1080/16066350601165448

Gavali, S., Chen, C., Cowart, J., Peng, X., Ding, S., Wu, C., & Anderson, T. (2021). Understanding the factors related to the opioid epidemic using machine learning. *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1309–1314. https://doi.org/10.1109/BIBM52615.2021.9669486

Kolak, M. A., Chen, Y.-T., Joyce, S., Ellis, K., Defever, K., McLuckie, C., Friedman, S., & Pho, M. T. (2020). Rural risk environments, opioid-related overdose, and infectious diseases: A multidimensional, spatial perspective. *International Journal of Drug Policy*, *85*. https://doi.org/10.1016/j.drugpo.2020.102727

Li, Y., Miller, H. J., Root, E. D., Hyder, A., & Lu, D. (2022). Understanding the role of urban social and physical environment in opioid overdose events using found geospatial data. *Health and Place*, *75*. https://doi.org/10.1016/j.healthplace.2022.102792

Marshall, B. D. L., Alexander-Scott, N., Yedinak, J. L., Hallowell, B. D., Goedel, W. C., Allen, B., Schell, R. C., Li, Y., Krieger, M. S., Pratty, C., Ahern, J., Neill, D. B., & Cerdá, M. (2022). Preventing Overdose Using Information and Data from the Environment (PROVIDENT): Protocol for a randomized, population-based, community intervention trial. *Addiction*, *117*(4), 1152–1162. https://doi.org/10.1111/add.15731

Nelson, R. K., & Winling, L. (2023). *Mapping Inequality: Redlining in New Deal America* [dataset]. Digital Scholarship Lab. https://dsl.richmond.edu/panorama/redlining/contactus

Rhodes, T. (2002). The "risk environment": A framework for understanding and reducing drug-related harm. *International Journal of Drug Policy*, *13*, 85–94. https://doi.org/10.1016/S0955-3959(02)00007-5

Schell, R. C., Allen, B., Goedel, W. C., Hallowell, B. D., Scagos, R., Li, Y., Krieger, M. S., Neill, D. B., Marshall, B. D. L., Cerda, M., & Ahern, J. (2022). Identify Predictors of Opioid Overdose Death at a Neighborhood Level with Machine Learning. *American Journal of Epidemiology*, *191*(3), 526–533. https://doi.org/10.1093/aje/kwab279

Spencer, M. R., Garnett, M. F., & Miniño, A. M. (2024, March 19). *Products—Data Briefs—Number 491—March 2024*. https://doi.org/10.15620/cdc:135849

Swope, C. B., Hernández, D., & Cushing, L. J. (2022). The Relationship of Historical Redlining with Present-Day Neighborhood Environmental and Health Outcomes: A Scoping Review and Conceptual Model. *Journal of Urban Health*, *99*(6), 959–983. https://doi.org/10.1007/s11524-022-00665-z

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). Scene Parsing through ADE20K Dataset. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., & Torralba, A. (2018). Semantic understanding of scenes through the ade20k dataset. *International Journal on Computer Vision*.