Eliot Padzensky
12/4/2017
EnSt 380

Nutrition Access Inequality in Chicago, Illinois

Introduction

In the United States, there exists vast disparities in access to healthy, fresh food. This is often closely linked to socioeconomic divides; areas with higher low-income or minority populations tend to have the least access to nutritious food. The USDA labels areas food deserts where this issue is most severe. In their 2009 report to congress on food deserts and access to nutrition, they found that 23.5 million Americans living in low-income areas live more than a mile from a supermarket or large grocery store. This represents more than a simple inconvenience to these populations. For many, their most convenient sources of food are small corner stores which often lack produce or other nutritious foods. The same study notes that there is a correlation between the lack of access to nutritious foods and nutrition related health issues such as obesity (however it also mentions that a causal link has not yet been fully proven).

In recent years the inequality in access to nutrition has become a much better understood and more salient issue and consequently policymakers from the federal to the local level have begun to try to address it. For instance, this was a core focus of Michele Obama's Let's Move program. In Chicago, Mayor Rahm Emanuel made the issue part of his platform and has worked to address it throughout his tenure. Nonetheless, he has been criticized for falling short on his goals (Ruthhart).

Fig. 1: Google searches for "food desert" (Google Trends)

To most effectively combat the issue, it is imperative that policymakers be equipped with research and data about its causes, scope and consequences. This paper focuses on Chicago, specifically, and uses both quantitative and spatial analysis to seek to better identify the extent to which the city is affected by nutritional access inequality, the extent to which it is tied to socioeconomic factors and how it correlates with negative health outcomes.

Chicago was an excellent subject for this research due to its size and extensive and pervasive racial and economic divides which facilitated the identification of correlations between food deserts and demographics. In 2015, the city was estimated to have a population of 2,717,534. Chicago is 32.2% percent white, 30.9% African American or black, 29.1% Hispanic or Latino and 5.9% Asian. (U.S. Census bureau) In 2012, the city's per-capita income was $28,202 and 19.6% of households were below the poverty line.

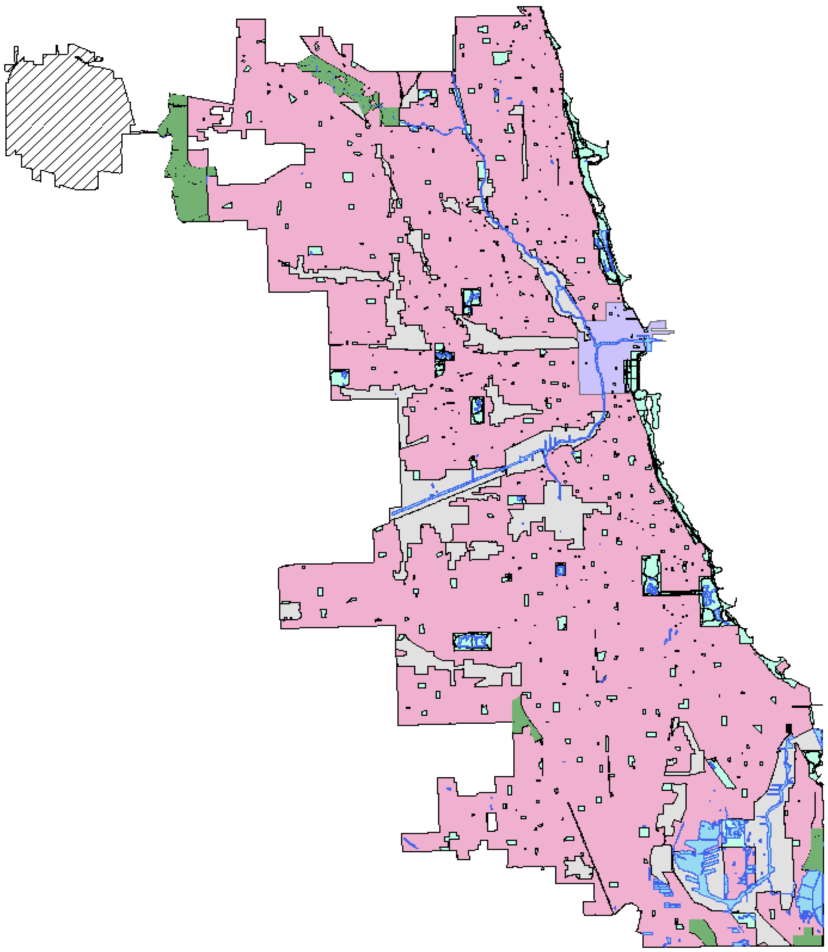
Scope of Analysis

For administrative and planning purposes, the city of Chicago is divided into 77 sectors called community areas. Community areas are static, precisely defined and unrelated to census blocks, wards or neighborhoods. Because the available relevant data is divided by community area, the analysis in this paper is centered around them. The scope of the analysis was limited to residential areas. Parks, industrial corridors, lakes, rivers, forestry, the central business district (the Loop) and O'Hare Airport were thus excluded in spatial analysis. Using these parameters, 172 square miles, or 72% of the total area of the city was designated as residential (Figure 2). Although the O'Hare community area does have a population, it was excluded because it mainly consists of the airport and this would likely affect spatial analysis. The Loop was excluded because the number of stores would likely be affected by the daily influx of commuters and the sizes of the stores by the high price of real estate. It should also be noted that there was missing data for the Loop and Montclare community areas, however, the Loop community area is entirely contained within the borders of the already-excluded central business district and Montclare is relatively affluent and, as such, not of concern in this analysis. This missing data is thus arguably not an issue.

This paper will only consider brick-and-mortar grocery stores. Although there exists data about farmer's markets, because they are seasonal, there are relatively few throughout the city and their prices tend to be higher than those in stores, the decision was made to exclude them from the analysis. The grocery stores data used for this research lists 506 stores in Chicago, 226 of which are larger than 10,000 square feet. When taking into account all grocery stores, no food deserts were clearly identified. However, when limited to only supermarkets (defined as a grocery store larger than 10,000 square feet), this was no longer the case.

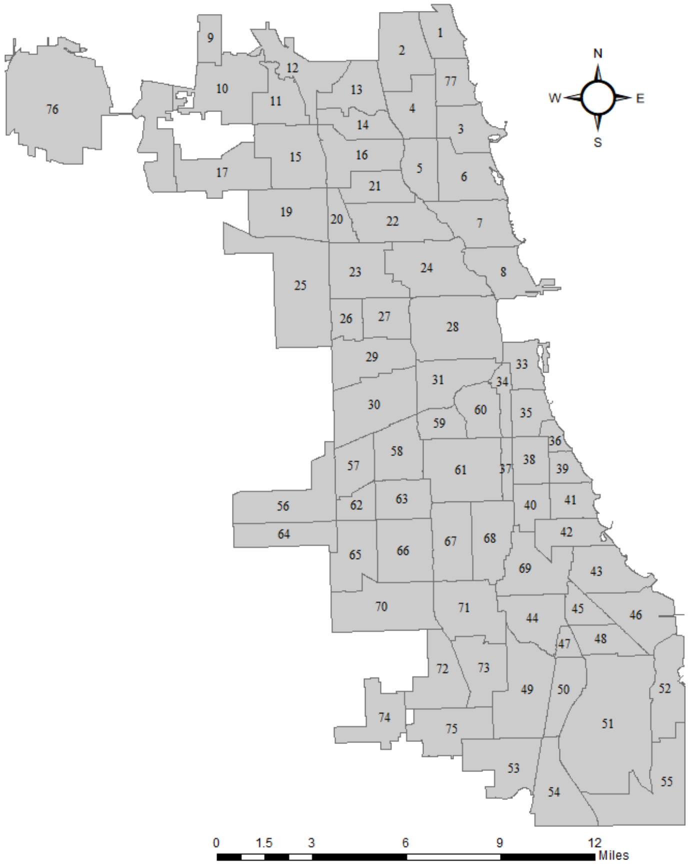
There were various dimensions to the analysis. First, the extent of the socioeconomic divisions within the city was identified. Next, a cursory quantitative analysis was done to identify the disparity in the number of grocery stores between community areas. Then, spatial analysis was used to find areas that had the farthest distances from the stores as well as their accessibility using public transit. Ultimately, several areas of the city were identified as food deserts. Their locations correlated with the distributions of African America, Latino and low-income populations. Likewise, they were a predictor of negative health outcomes. Ultimately, this analysis makes it clear that Chicago does indeed suffer from disparities in access to nutrition and this is closely tied to socioeconomic factors.

Fig. 2: Scope of Analysis



Category	Area (Miles^2)	Color
Residential Area	172	
Rivers and Lakes	5	
Forestry	5	
Park	13	
Central Business District	4	
Industrial Corridor	27	
O'Hare Airport	10	
All of Chicago	236	

Fig. 3: Community Areas in Chicago[*]



1	Rogers Park
2	West Ridge
3	Uptown
4	Lincoln Square
5	North Center
6	Lake View
7	Lincoln Park
8	Near North Side
9	Edison Park
10	Norwood Park
11	Jefferson Park
12	Forest Glen
13	North Park
14	Albany Park
15	Portage Park
16	Irving Park
17	Dunning
18	Montclare
19	Belmont Cragin
20	Hermosa
21	Avondale
22	Logan Square
23	Humboldt Park
24	West Town
25	Austin
26	West Garfield Park

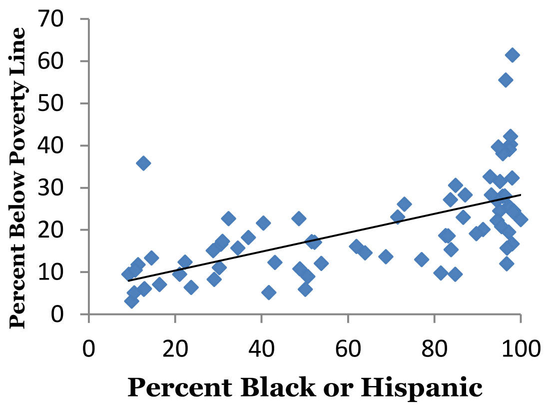
27	East Garfield Park	42	Woodlawn	57	Archer Heights	72	Beverly
28	Near West Side	43	South Shore	58	Brighton Park	73	Washington Heights
29	North Lawndale	44	Chatham	59	McKinley Park	74	Mount Greenwood
30	South Lawndale	45	Avalon Park	60	Bridgeport	75	Morgan Park
32	Loop	46	South Chicago	61	New City	76	O'Hare
31	Lower West Side	47	Burnside	62	West Elsdon	77	Edgewater
33	Near South Side	48	Calumet Heights	63	Gage Park	72	Beverly
34	Armour Square	49	Roseland	64	Clearing	73	Washington Heights
35	Douglas	50	Pullman	65	West Lawn	74	Mount Greenwood
36	Oakland	51	South Deering	66	Chicago Lawn	75	Morgan Park
37	Fuller Park	52	East Side	67	West Englewood	76	O'Hare
38	Grand Boulevard	53	West Pullman	68	Englewood	77	Edgewater
39	Kenwood	54	Riverdale	69	Greater Grand Crossing		
40	Washington Park	55	Hegewisch	70	Ashburn		
41	Hyde Park	56	Garfield Ridge	71	Auburn Gresham		

[*] *Unless otherwise indicated, all maps in this paper utilize the same scale and have the same orientation*

Methods

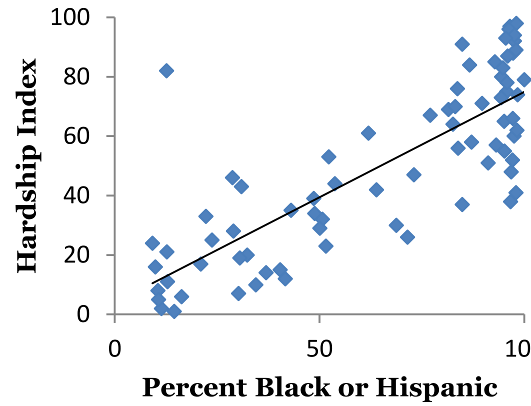
The first dimension that was examined was the degree to which the city of Chicago is socioeconomically divided. This analysis was based on two datasets: *Census Selected Socioeconomic Indicators in Chicago*, *2008 – 2012* and *Chicago Community Area (CCA) CDS Data*. The former lists data such as per-capita income, percent of people under the poverty line, percent unemployment and a metric called *hardship index* which takes all of this into account to determine a ranking for the community area. The higher rankings correspond to more economic hardship. The second dataset has demographic data for each community area as well as other data pertaining to income, employment, transportation, home values and population trends. Because this only listed the number of people from each race for the community areas, I had to create a set of new fields and use the field calculator to produce a breakdown by percentage. I also created a field to indicate if more than 50% of a community area's population was the same race. To do this, I again used the field calculator but with a python script that returned a number to correspond to the majority race or 0 if there was no majority. I also used Microsoft Excel to perform a linear regression on the combined data to find the correlation between race and percent of the population below the poverty line (fig. 4), and race and hardship index (fig. 5) for each community area. These variables were found to be strongly correlated and statistically significant. Likewise, the vast majority of community areas had a population consisting of at least 50% of the same race. Ultimately, it is clear that Chicago is deeply socioeconomically divided.

Fig. 4: Race and Poverty



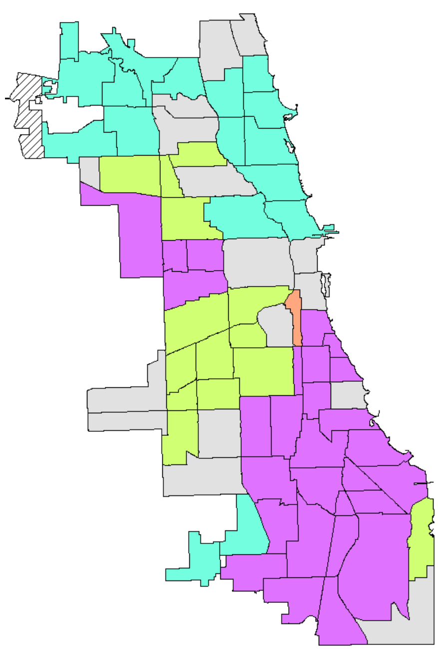
	Coefficients	*P-value*
Intercept	5.946053	0.014527
% Black or Hispanic	0.22378	2.07E-09

Fig. 5: Race and Hardship Index



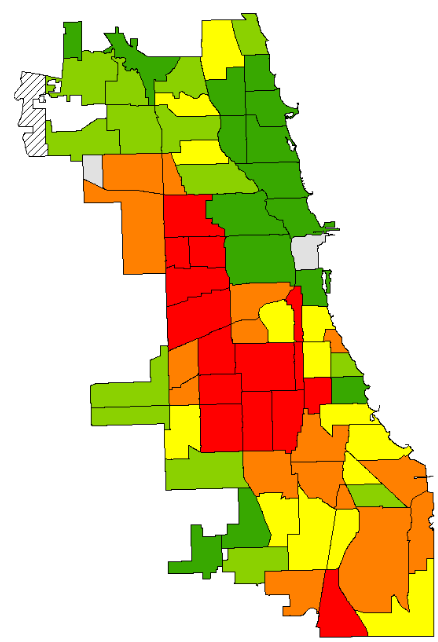
	Coefficients	*P-value*
Intercept	3.899012	0.386702
% Black or Hispanic	0.709974	4.55E-18

Fig. 6: Racial Makeup of Community Areas



Color	**Majority**	**Count**
	None	14
	White	18
	Black	28
	Hispanic	14
	Asian	1

Fig. 7: Hardship Index by Community Area



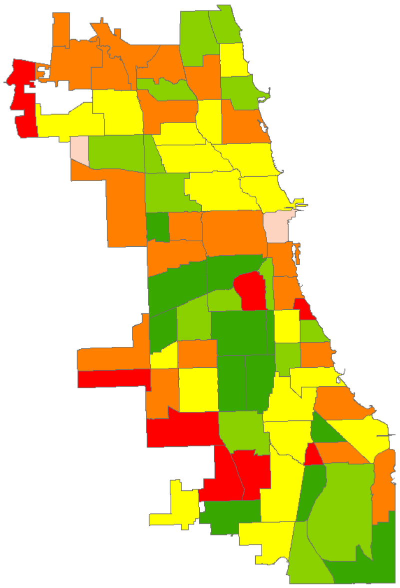
Color	**Range**	**Count**
	1 - 20	15
	21 - 39	15
	40 - 58	14
	59 - 79	16
	80 - 98	15

Finding Food Deserts through Quantitative Analysis

To identify food deserts in Chicago, I first did a quantitative analysis using ArcMap. The grocery stores data had a community area field for each store. I summarized this data to find the number of grocery stores for each community area. I then joined this to the combined census and demographics data. Next, I created a new field and used the field calculator to find the number of grocery stores per 10,000 people for each community area. I then exported the data and used linear regression in Excel to determine whether correlations existed between this and factors such as race, hardship index or poverty. Surprisingly, not only were the correlations small and not statistically significant, but they were positive i.e. poorer community areas had more stores.

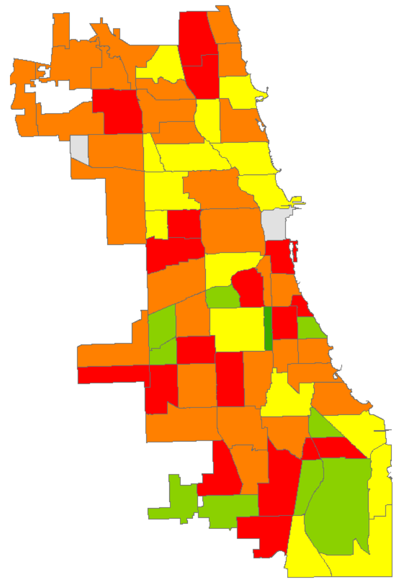
It was then that I realized that the grocery stores data would need to be further refined to make it useful. Many of the stores were small corner stores – stores where high quality, nutritious food is often not available. To create a subset of the data with only supermarkets, I selected by attribute to find the stores larger than 10,000 square feet and exported this. The new "supermarkets" dataset contained 226 of the 506 stores from the original dataset. After repeating the above steps with the refined data, the results much more closely reflected expectations. Many areas with the fewest stores were indeed primarily black or Hispanic. Furthermore, eight community areas had no supermarkets. The average percentage of individuals living below the poverty line in the eight exceeded that of the city by three percentage points. Of these eight community areas, 75% were at least 90% black or Hispanic. Overall, although cursory, this quantitative analysis clearly indicated the presence of a trend, however, further analysis would be necessary to obtain more meaningful results.

Fig. 8: Grocery Stores per 10,000 People



Color	Range	Count
	0.000000 - 0.737572	8
	0.737573 - 1.457070	21
	1.457071 - 2.164100	19
	2.164101 - 2.908510	15
	2.908511 - 4.146640	12

Fig. 9: Supermarkets per 10,000 People



Color	Range	Count
	0.000000 - 0.462670	18
	0.462671 - 1.010620	31
	1.010621 - 1.463520	16
	1.463521 - 2.261420	9
	2.261421 - 3.932360	1

Figures 8 and 9 demonstrate the difference between the two analyses. For the maps, I divided the data using natural breaks.

Fig. 10: Community Areas with no Supermarkets

	Community Area	Percent Black or Hispanic	Hardship Index	Percent Below Poverty Line	Supermarkets per 10,000 People
75	Oakland	95.84	78	38.10	0.0
74	Grand Boulevard	93.18	57	28.30	0.0
73	East Garfield Park	94.80	83	39.70	0.0
72	Burnside	100.00	79	22.50	0.0
71	Calumet Heights	96.69	38	12.00	0.0
70	Roseland	97.16	52	19.50	0.0
69	Bridgeport	30.98	43	17.30	0.0
68	Beverly	41.66	12	5.20	0.0
Averages		**81.29**	**55.25**	**22.83**	**0.0**
City of Chicago		**59**	**N/A**	**19.60**	**0.83**

Identifying Food Deserts through Spatial Analysis

Although the quantitative analysis indeed indicated the existence of nutrition access inequality in Chicago, precisely identifying the locations of food deserts required more extensive spatial analysis. The spatial analysis focused on two variables – the location of the stores and their proximity to public transit. This was primarily accomplished with Euclidian distance. Before beginning, however, I produced a “residential areas” layer (see figure 2). Starting with the city boundaries layer, I first used the feature editor to manually remove O’Hare airport and the I-190 spur (this was not arbitrary – there is a clear point where the city boundary juts out to follow the interstate to the airport). Next, I used the erase tool to remove bodies of water, forestry, parks, industrial corridors and the central business district (the loop). The end result was a polygon that roughly maps to residential areas, the area of concern in my analysis. I used the summarize and calculate geometry tools to find the areas listed in figure 2.

Next, I used Euclidian distance with a 250-foot cell size to begin the spatial analysis of the geographic distribution of the grocery stores. I then extracted by mask using the residential areas layer. I reclassified this raster to values representing half mile increments up to three miles and finally extracted by attribute where the distance was greater than one half mile, the distance used by the USDA in part of its definition of urban food deserts (the distance for non-urban food deserts is larger). The resulting map consisted of several large areas and clusters of smaller areas whose distance from the nearest grocery store was greater than one half mile. These results aligned with the results of the quantitative analysis in that they provided further evidence of disparities in access to supermarkets.

Fig. 11a: Euclidian Distance

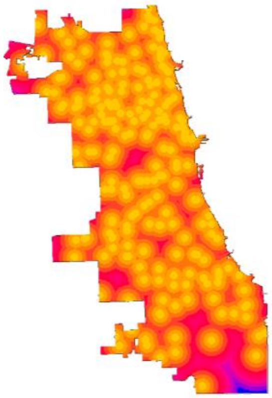


Fig. 11b: Extract by Mask

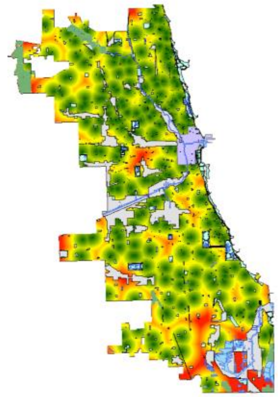


Fig. 11c: Reclassify

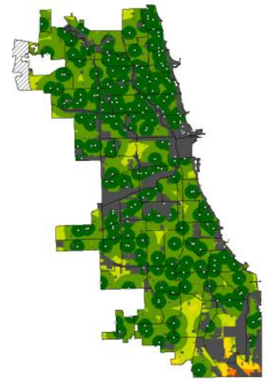


Fig. 11d: Extract by Value where Distance is Greater than 0.5 Miles

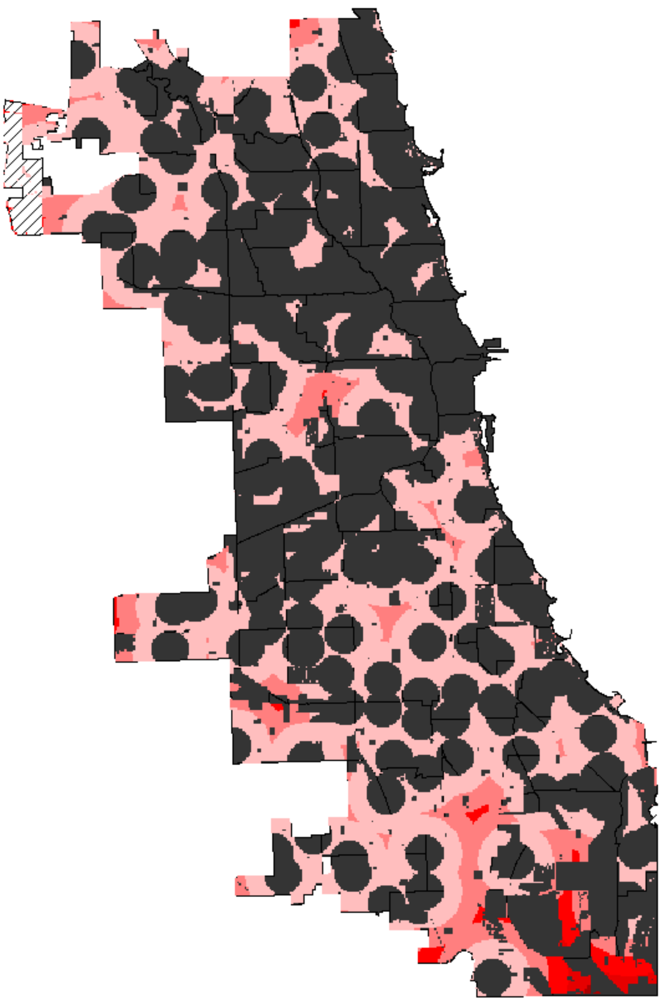


Fig. 11e: Reclassification Values

Feet	Miles	Value	Color
2540	.5	1	
5280	1	2	
7920	1.5	3	
10560	2	4	
13100	2.5	5	
15840	3	6	

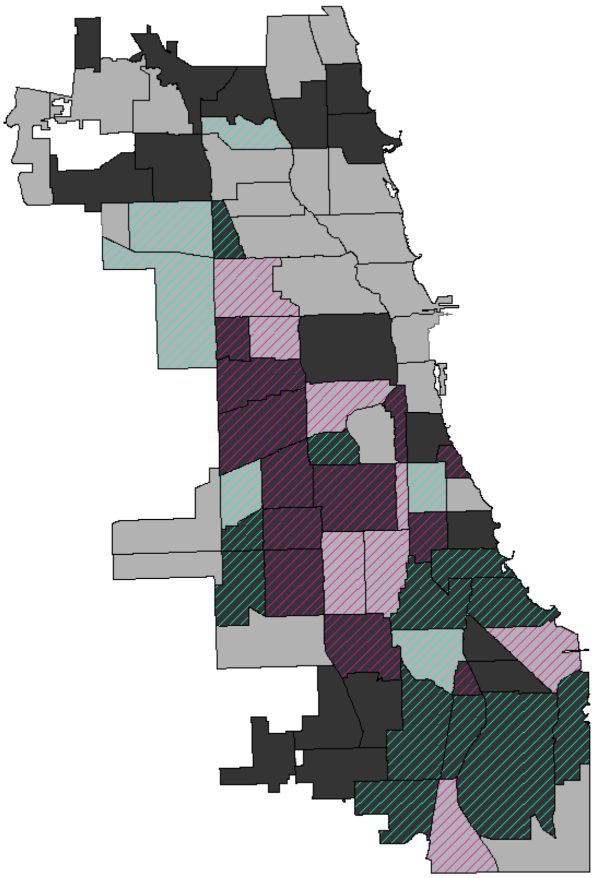
Averaging Euclidian Distances

While the previous section identified the areas farthest from supermarkets, the demographic, socioeconomic and public health data is broken down by community area. To be able to compare the results from the previous section to this data, I needed to find the average distance for each community area. To accomplish this, I used the zonal statistics tool with the community areas layer as my zones and the masked, pre-reclassification Euclidian distances layer as my values. I reclassified the output to half mile increments labeled one through four (the highest average value was just under two miles). I repeated this using the zonal statistics as table tool and joined the results to the combined community areas, demographics and public health data. I then selected by attribute to find the overlap between black and Hispanic majority community areas and the community areas where the average distance to supermarkets in residential areas was greater than a half mile. To find the average distance to supermarkets for the entire city (0.55 miles), I again used the zonal statistics as table tool but set the zone data to the city boundaries layer.

Fig. 12: Access to Supermarkets vs. Socioeconomic Factors

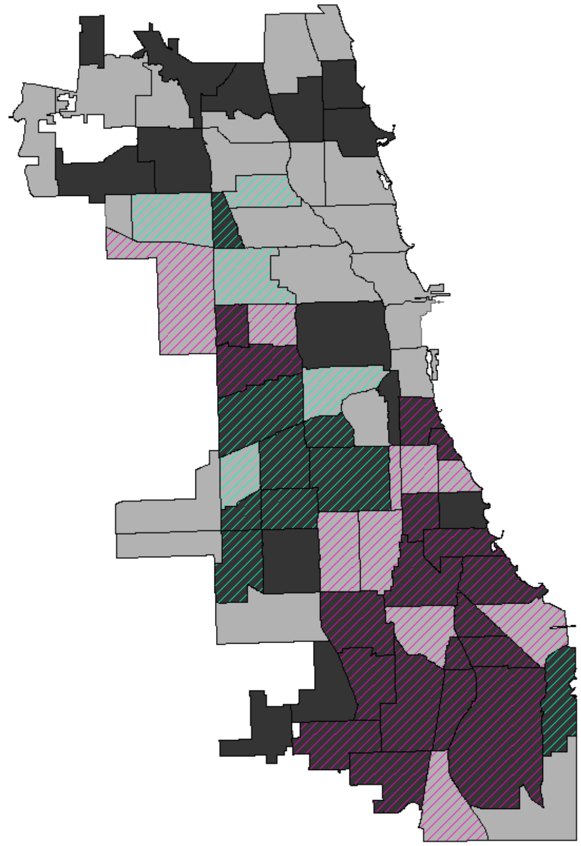
Community Area Category	Total Number	Number with Low Access to Supermarkets in Category	Percentage of Category with Low Access to Supermarkets	Percentage of Community Areas with Low Access to Supermarkets from that Category
Only Majority Black	28	18	64.3%	43.9%
Only Majority Hispanic	14	9	64.3%	22.0%
Only in Top 50% of Hardship Indices	38	24	63.2%	58.5%
Only in Top 25% of Hardship Indices	19	12	63.2%	29.3%
All	**75**	**41**	**54.6%**	

Fig. 13: Hardship Index vs. Access to Grocery Stores



	Average Distance to Supermarkets Less Than ½ Mile
	Average Distance to Supermarkets Greater Than ½ Mile
	Highest 50%
	Highest 25%

Fig. 14: Race vs. Access to Grocery Stores



	Average Distance to Supermarkets Less Than ½ Mile
	Average Distance to Supermarkets Greater Than ½ Mile
	Majority Hispanic
	Majority Black

Identifying Accessibility of Supermarkets

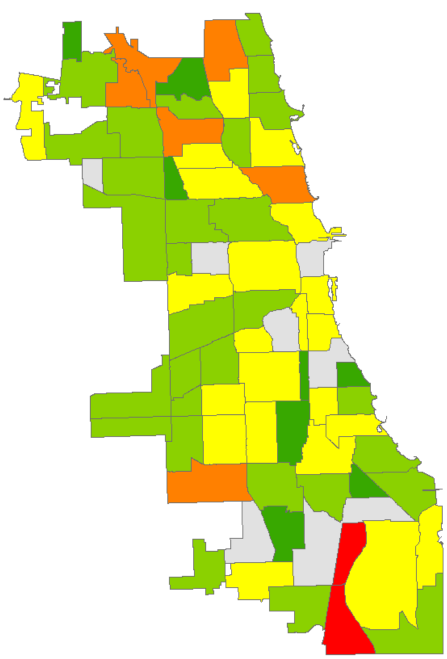
The USDA definition of a food desert does not only take into account location, but accessibility too. This is especially relevant in an urban center such as Chicago where many, especially low-income people rely on public transportation. Thus, my next step was to identify the accessibility of Chicago's supermarkets by finding their distances from the nearest L train or bus stop. To do this, I performed a spatial join on the supermarkets data with a layer of points representing the bus stops in Chicago, giving each supermarket the distance to the nearest one. I repeated this with the output of that and the layer of points representing the locations of the L train stations. In the resulting table I created a new field and used the field calculator with python to set it to the minimum of the two distance to get the distance to the nearest public transportation stop. Finally, I used the summarize tool to find the average distance to public transportation for each community area and joined it to the community areas / population / demographics data by community area.

Although I could not find data about rates of vehicle ownership by community area, the census data included fields for the number of people who commute by transit as well as the total number of commuters. I created a new field and used the field calculator to set it to the percent of people who commute by public transit. Figure 15 shows that many community areas with hardship indices greater than 60 whose population relies most on public transit either contain no supermarkets or have high average distances between supermarkets and the nearest transit stops.

Fig. 15: Ten Community Areas with Hardship Indices Greater than 60 with Highest Usage of Public Transportation

Community Area	Percent Black or Hispanic	Hardship Index	Supermarkets per 10,000 People	Average Distance to a Supermarket (Miles)	Average Stop Distance (Feet)	Percent Using Transit
Englewood	97.59	94	0.7657	0.50	18.086	41.98275862
Washington Park	97.29	88	0.8277	0.47	390.496	40.9219191
Burnside	100.00	79	0.0000	0.64	N/A	39.47368421
Oakland	95.84	78	0.0000	0.41	N/A	38.1613183
Greater Grand Crossing	97.18	66	1.2366	0.41	244.912	38.089406
East Garfield Park	94.80	83	0.0000	0.89	N/A	37.08165997
West Englewood	98.02	89	0.3110	0.63	263.775	34.44857497
West Garfield Park	97.60	92	1.1273	0.36	157.120	34.40166013
Riverdale	98.08	98	1.4104	0.76	1381.458	34.140625
North Lawndale	96.01	87	0.2835	0.70	244.802	33.38681948
Chicago	**59**	**N/A**	**0.83**	**0.55**	**276.44†**	**29.04**

Fig. 16: Community Areas by Average Distances Between Supermarkets and Public Transit



Color	Range (Feet)	Count
	18.0857095 - 74.6805441	8
	74.6805442 - 200.419103	29
	200.419104 - 459.897487	21
	459.897488 - 876.820623	7
	876.820624 - 1704.51391	2
	NO SUPERMARKETS	8

Divided by natural breaks

† Excludes community areas with no supermarkets

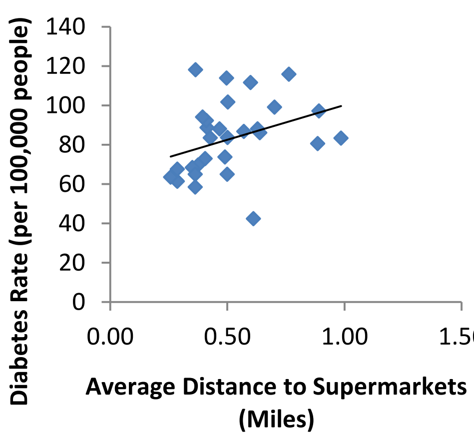
Public Health Data

The final component of my analysis was to determine the consequences of the disparities in access to nutrition. Specifically, I wanted to look at how low access corresponds to diet related health issues. Unfortunately, there was no available data about rates of obesity or cardiovascular illnesses by community area. For my analysis, I instead used data about rates of diabetes and life expectancy. To find the relation between access to nutrition and diabetes and life expectancy, I used the select by attribute and statistics tools. I also performed a regression analysis in Excel.

Fig. 17: Relationship Between Access to Nutrition and Negative Health Outcomes

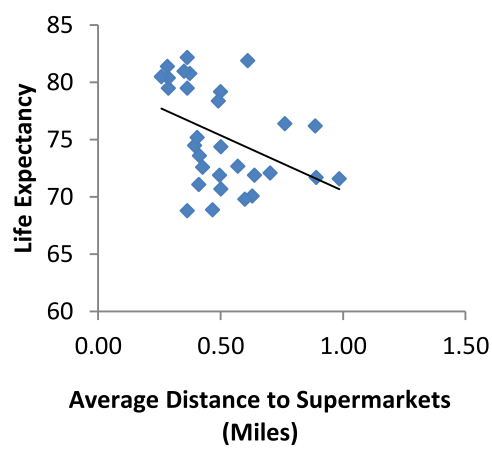
Community Areas	Average Distance to Supermarkets (Miles)	Supermarkets per 10,000 People	Average Distance Between Supermarkets and Transit (Ft)	Average Hardship Index
Bottom 50% of Diabetes Rates	0.49	0.88	258.68	38.87
Top 50% of Diabetes Rates	0.57	0.95	297.10	61.68
Bottom 50% Life Expectancies	0.59	0.91	308.71	64.05
Top 50% Life Expectancies	0.46	0.92	246.94	35.81

Fig. 18: Access vs. Diabetes



	Coefficients	*P-value*
Intercept	62.37542	3.46E-15
Average Distance	18.9375	0.09155

Fig. 19: Access Vs. Life Expectancy



	Coefficients	*P-value*
Intercept	79.58281	6.37E-67
Average Distance	-4.02357	0.061612

Results

As predicted, my analysis indeed indicated the existence of disparities in access to nutrition in Chicago. These differences were correlated with socioeconomic factors and negative health outcomes. This conclusion aligns with the existing research on the subject of food deserts including that of the USDA. Ultimately, although correlation does not prove causation, the results of this analysis nonetheless indicate that these disparities do exist and potentially pose an enormous risk to those affected by them.

While finding correlations with socioeconomic factors and health data was relatively straightforward, the identification of food deserts proved to be somewhat difficult. The initial quantitative analysis with all grocery stores did not yield any valuable information. Upon reexamining the USDA's definition of a food desert, I realized that my data was not specific enough and once I refined it to only include large grocery stores, socioeconomic-based nutritional disparities between community areas began to emerge. Nonetheless, this basic quantitative analysis had multiple shortcomings. For instance, South Deering was found to have 1.96 supermarkets per 10,000 people, well above average. These, however, were clustered in the north and thus the community's south and central portions were not identified as food deserts despite being far from the nearest store.

The spatial analysis yielded much more nuanced results. It showed several areas more than a half mile from the nearest supermarket, thus meeting the USDA's definition for an urban food desert. Furthermore, this analysis made it possible to identify areas with limited access to supermarkets which spanned multiple community areas. Using zonal statistics to find the average distance to a supermarket for each community area produced stronger correlations with socioeconomic factors and public health data.

Because the USDA definition of a food desert takes into account transportation in addition to distance, I elected to do the further spatial analysis utilizing the public transportation data. This identified the community areas with the supermarkets most and least accessible by bus or train. When looking at the city as a whole, these results were not particularly insightful. In more affluent areas where vehicle ownership is higher, it is not as important for supermarkets to be accessible by public transit and thus the spatial analysis indicated that many such community areas were relatively inaccessible. Nonetheless, when only looking at the community areas with the highest hardship indices and dependence on public transit, it became clear that there was indeed a link between poverty and accessibility; many of these community areas had increased average distances to super markets and between supermarkets and public transportation.

The final component of the analysis was to incorporate public health data. Using quantitative analysis, I found strong, statistically significant correlations between the average distance to a supermarket and diabetes and life expectancy (Figs. 17-19).

Discussion

Although my analysis did find correlations between socioeconomic factors and food deserts, and food deserts and negative health outcomes, there are several important limitations of which to be aware. First, correlation does not prove causation. As the USDA notes in its national study on food deserts, no causal link has been established between the aforementioned variables. While it is quite possible, for instance, that limited access to healthy food does result in a decreased life expectancy, it is also possible that both are function of another variable such as living near an industrial area. Another potential issue is that there may be multiple causes of food deserts. A supermarket's profitability is dependent on the size and affluence of its customer base. Thus, low population densities and poverty might both explain food deserts.

Another shortcoming of my analysis was the limitations of the available data. First, the grocery stores data did not contain more detailed information about the variety of products offered. Thus, when I refined it to only include large stores, I may have excluded places that do sell healthy food and included places that do not. Nonetheless, in the absence of this information, filtering by size was the most logical way to refine the data. Another potential issue was that the spatial analysis indicated that many of the affluent neighborhoods along Chicago's northern border had high distances to supermarkets. While this would seem to contradict the hypothesis that distance is correlated with income, the grocery store data only included locations within the city of Chicago and thus stores in nearby municipalities were not taken into account. A final limitation of the data was that it was only broken down my community area. Chicago's community areas were originally established in the 1920s and have remained unchanged since then, with the exception of additions when the city has grown. (Chicago Historical Society) Combined with their large sizes, it is possible that these demarcations do not adequately reflect the city's demographic and socioeconomic divisions. Future studies should take advantage of more localized data, should it become available.

Despite these limitations or potential issues, the link between socioeconomic factors, limited access to nutrition and negative health outcomes is clear. Regardless of the existence of a causal relationship, it is imperative that these issues be addressed. Analyses such as this project are crucial because they provide policymakers with the tools to understand the scope of the issues and thus more effectively solve them.

Literature Cited

"Chicago's Community Areas." Encyclopedia of Chicago, Chicago Historical Society, 2005, www.encyclopedia.chicagohistory.org/pages/1760.html.

Google Trends "Google Searches for 'Food Desert'" www.google.com/trends

Ruthhart, Bill. "Emanuel Still Short on Food Desert Goals, Even as Englewood Whole Foods Opens." The Chicago Tribune, 28 Sept. 2016, www.chicagotribune.com/news/local/politics/ct-rahm-emanuel-food-deserts-met-0928-20160927-story.html.

United States, Census Bureau, "2011-2015 American Community Survey 5-Year Estimates." 2011-2015 American Community Survey 5-Year Estimates, 2015. www.census.gov/programs-surveys/acs/.

United States, Department of Agriculture, Economic Research Service. "Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences." Access to Affordable and Nutritious Food: Measuring and Understanding Food Deserts and Their Consequences, June 2009. www.ers.usda.gov/webdocs/publications/42711/12698_ap036fm_1_.pdf?v=41055.