Identifying Geospatial Indicators of Poor Health in Chicago

Max Rodrigues
Lyubov Chebotareva

Abstract
Built environment features and their relationship to health outcomes is a highly researched field due to the many nefarious implications of unhealthiness. In this paper, we examine multiple built environment features that are known to be related to obesity and other health outcomes in a GIS setting. These features include grocery stores density, fast food density, and proximity to areas individuals can exercise. For food availability, we utilized approaches similar to other researchers but examined densities in terms of the closest 5 features to each census tract, and for exercise accessibility we viewed densities by converting point data to continuous raster data with a kernel density estimation function. Both aspects of the analysis found underserved areas mostly in the south and west sides of Chicago.

Introduction

Overview of Related Research

The topic of geospatial factors influencing health outcomes is a heavily researched field. In an urban environment, there are a multitude of human-designed elements such as land usage, transportation systems, and accessibility to certain neighborhood features that all influence health behaviors and have been studied for their relationship to obesity (Thornton et al. 2011). Other researchers have examined the relationship between built environment factors such as walkability, fitness facility density, and recreational center density and childhood obesity (Jia et al 2019). What is clear from the research is that while certain built environment features stand out as highly correlative with health outcomes, there are many features that have influence and a holistic approach may be best when conducting geospatial analysis on health outcomes.

Methodologies of Other Researchers

Other researchers (e.g. Jia et al. & Kolak et al. 2018) have looked at geospatial measures of accessibility to identify areas where people may be at higher risk of poor health. In the case of Kolak et al., they identified food deserts based on grocery store data by calculating centroids of census tracts then averaging the distance to all grocery stores for each census tract in the city with higher average distances indicate potential food deserts. Centroids are often used when identifying accessibility as shown by Jia et al., but so is density (i.e., the count of features within a certain area) and kernel density estimation (i.e., transforming point data to a continuous

density surface map). There are also several other measures of accessibility (connectivity, network distance, walkability) but these measures will not be explored in this analysis.

Problem Definition

Since some neighborhoods in a city are more economically advantaged than others, there are large disparities in accessibility to certain neighborhood features such as fast-food restaurants, convenience stores, and grocery stores between neighborhoods. Given the demonstrated strong relatedness between certain geospatial built environment features and health outcomes, we chose to investigate the distribution of some potential health-related geospatial indicators in Chicago. Specifically, we focused on two general built environment factors that could be analyzed in several ways: food availability as it relates to grocery stores, fast-food restaurants, and farmers markets; and accessibility to places where an individual can exercise. By identifying areas with inaccessibility to healthy food options and areas where exercise options are limited, we can target areas that could use assistance in addressing these inequities that result in poor health and fitness.

Data and Methods

Food Availability Maps: Data Collection and Initial Processing

For the food availability maps, we used grocery store locations (Kolak et al. 2018) as a variable that is indicative of healthy food availability as well as farmers market locations (Chicago Data Portal 2012 data) and we used fast food restaurant locations (obtained from fastfoodmaps.com) as an indicator of unhealthy food availability. Both of these data files are relatively old (grocery stores are from 2014 and fast-food restaurant locations are from 2007), so the general geospatial distribution of these variables have likely changed due to closings and openings of stores and restaurants, but general patterns obtained from this data are likely to hold true. Fast food locations contain ten of the largest national chains including McDonalds, Burger King, and Wendys. Census tract boundaries were made available by the US Census Bureau and were downloaded from the Chicago Data Portal. To eliminate census tracts without households, we used the 2019 ACS Community Survey average population values at the level of census tracts and joined this data with the census tract boundaries, excluding tracts with zero population. Income data was utilized for one thematic map after this join. When creating maps, we used the NAD 1983 2011 StatePlane Illinois East FIPS 1201 map projection since data is being viewed at the city level. For simple point maps, we either used shapefiles provided by researchers or used text files with WGS 1984 longitude and latitude attributes and converted these to appropriate map projections.

Food Availability Maps: Data Processing, Defining Latent Variables, and Map Creation

We developed five thematic maps looking at the distribution of grocery stores and fast-food restaurants, as well as one that includes income. The first two maps show the average distance to the five closest fast-food restaurants and grocery stores to each census tract centroid, respectively. The third and fourth maps utilizes a latent variable called "Healthy Food Availability" which is the average distance to fast food restaurants divided by the average distance to grocery stores & farmers markets (one map excluded farmers markets since these are mostly seasonal). For the latent variable, smaller values are indicative of areas where the density of fast-food restaurants is greater than the density of grocery stores which acts as an approximation for easier access to unhealthy food. Following the work of other researchers (e.g. Kolak et al.), we initially viewed choropleths that took the average distance from the centroid of each census tract to all grocery stores/fast food restaurants in Chicago. However, we found that the maps were essentially population density maps with the loop having the highest density and surrounding neighborhoods becoming less dense which isn't very informative for our analysis. We changed our approach programmatically by calculating the average distance to the "n" closest grocery stores and fast-food restaurants from each census tract centroid. We chose the 5 closest but other values could also potentially provide useful insights. The fifth thematic map utilizes map algebra in a raster analysis using these calculated distances by highlighting areas with a high density of fast-food restaurants, a low density of grocery stores, and are also in low-income areas. Quantiles were used for classification in both the choropleths and the raster map.

Exercise Accessibility Maps: Data Collection

To obtain fitness center locations, we used web scraping software to pull address information from gyms, recreational centers, fitness centers, and cycling studios into a csv file and then geocoded the addresses so locations could be marked on a map. Web scraping involves extracting data directly from websites. For example, the Chicago Park District website has addresses to all Chicago park district fitness centers, and we used web scraping to obtain all of these addresses that were then geocoded. We went through chains (Planet Fitness, Xsport, LA Fitness, etc) and various google searches including gyms, fitness centers, recreational centers and cycling classes pulling as many as we could find but we admit this is an imperfect process as there will likely be some erroneously omitted and the omissions may be biased. However, the densities were similar to what we expected based on other aspects of the analysis. In addition to centers a person could exercise, we also used park locations and bike paths taken from the Chicago Data Portal since these are places people can potentially get exercise.

Exercise Accessibility Maps: Data Processing, Defining Latent Variables, and Map Creation

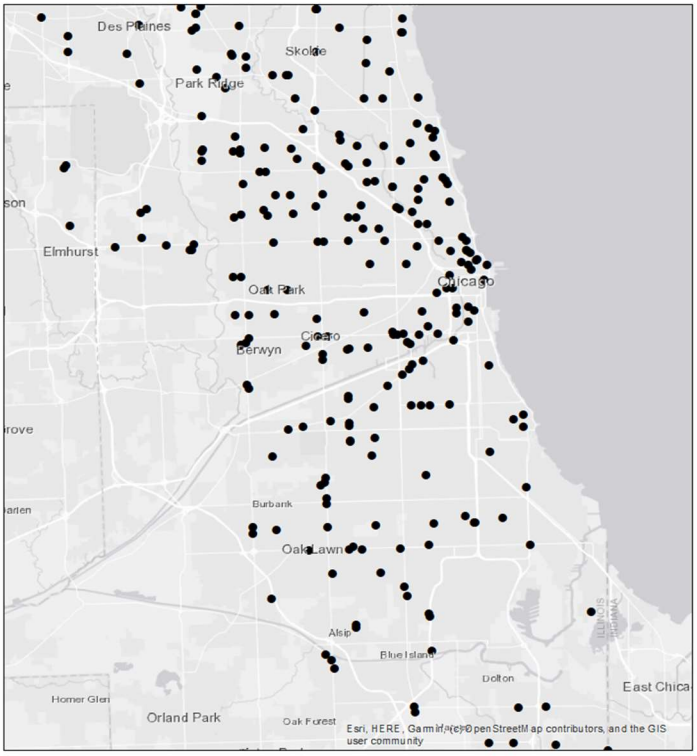
After creating a point map of fitness center locations and a map displaying park location, we found the centroids of each park to create a thematic map looking at the overall densities of fitness centers and parks throughout Chicago. Kernel density estimation converts point data to continuous raster data to highlight areas of higher density that are difficult to distinguish from simpler maps. Kernel density estimations were found for both fitness centers and park centroids with densities classified by quantiles. Quantiles were then reclassified to ordinal values (numbers 1-9) and map algebra was used to add values. For this map, darker values are indicative of higher densities of places to exercise (both parks and fitness centers).

Results

Food Availability Maps

In figure 1, the distribution of grocery stores appears denser near the loop and parts of the north and near west side, and sparser in the south side. In figure 2, fast food locations appear much denser near the loop, about equally dense in the northside neighborhoods as grocery stores, and slightly denser in the south/southwest side neighborhoods. Figure 3 shows that there are few farmers markets on the southwest side, but most farmers markets are not open year-round. The thematic maps viewing the average distance to the five closest grocery stores and fast-food restaurants (figures 4 & 5) supplement the first two point maps and convey a similar message: grocery store density is higher around the loop, north, northwest, and near southwest parts of Chicago and much sparser elsewhere, and fast food restaurants are dense around the north side but also have pockets of higher density in the southside. The thematic map viewing average distance to fast food restaurants divided by average distance to grocery stores/farmers markets (figures 6 & 7) gives a sense of the quality of food available to each census tract. In the south side, most census tracts have fast food restaurants as their major option for food, and some pockets on the west side also seem to mostly have fast food as their main food option. Figure 6 includes farmers markets and figure 7 excludes them.

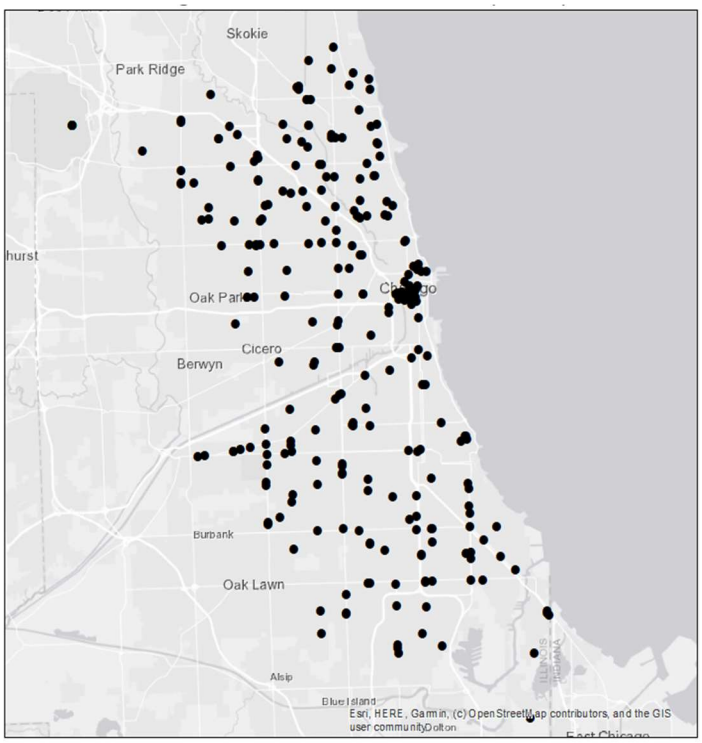
Figure 8 highlights areas with high densities of fast-food restaurants, low densities of grocery stores, and low income by reclassifying and utilizing map algebra to combine the three attributes into one. This map is similar to the “Healthy Food Availability” map, much of the southside, part of the west side, and a small pocket of land at the very Northern edge of Chicago face poverty, fewer grocery store options, and more fast-food options than most other parts of Chicago.



Legend

- Grocery Store Location

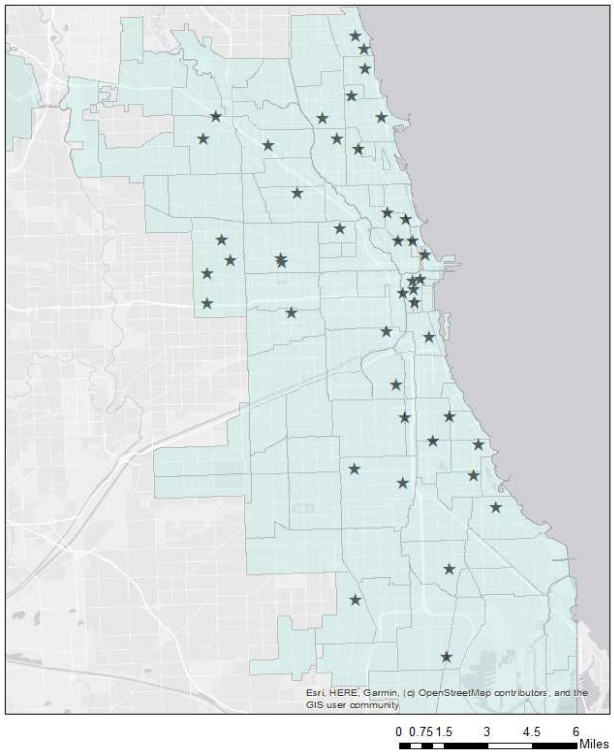
Figure 1: Grocery store locations (2014)



Legend

- Fast Food Restaurant Location

Figure 2: Fast-food restaurant locations (2007)



Source:Chicago Data Portal

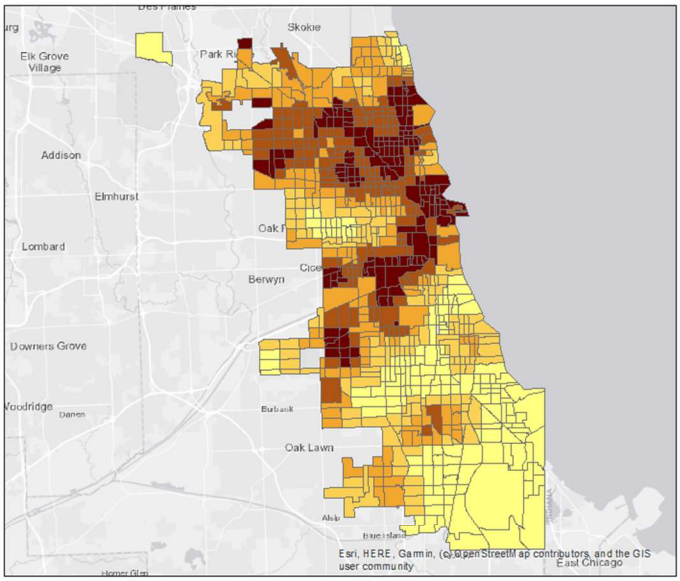
Legend

Neighborhoods boundaries

★ farmers markets 2012

Figure 3: Farmers market locations (2012)

Figure 4: Choropleth showing average distance to five closest grocery stores to each census tract centroid

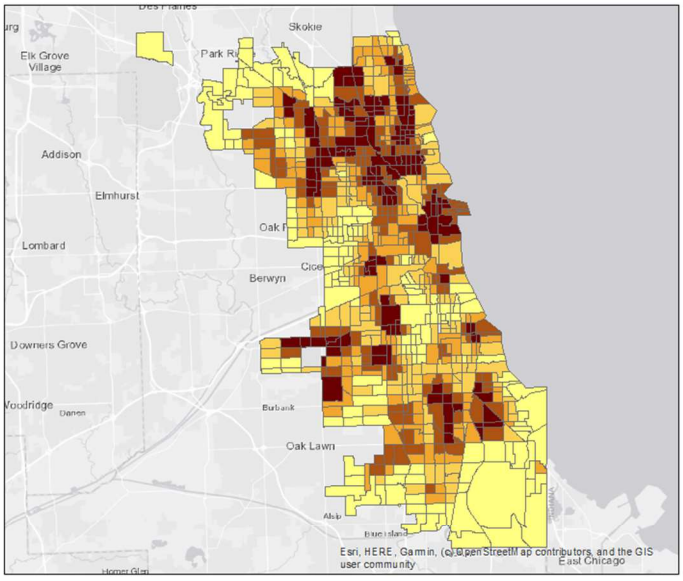


Legend

Avg dist to 5 closest grocery stores (meters)

- 409 - 1,157
- 1,158 - 1,458
- 1,459 - 1,884
- 1,885 - 2,485
- 2,486 - 6,397

Figure 5: Choropleth showing average distance to five closest fast-food restaurants to each census tract centroid



Legend

Avg dist to 5 closest fast food restaurants (meters)

- 246 - 983
- 984 - 1,190
- 1,191 - 1,427
- 1,428 - 1,708
- 1,709 - 4,837

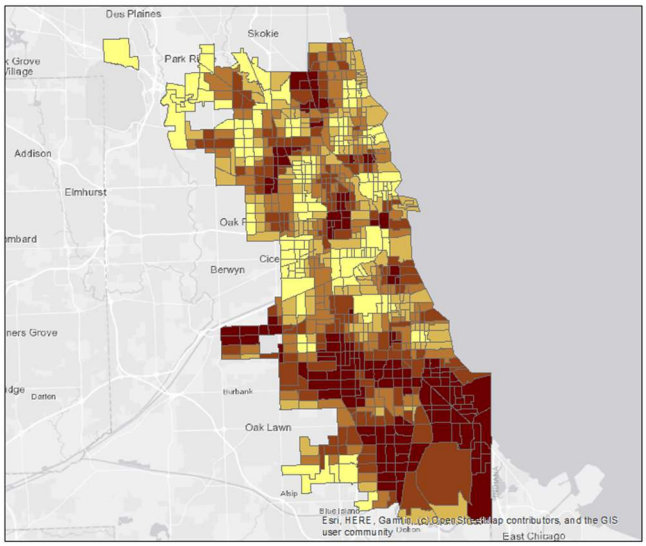


Figure 6: Food Availability Map showing average distance to five closest fast-food restaurants divided by average distance to grocery stores and farmers markets (smaller values indicate higher density of fast-food than grocery stores & farmers markets

Legend

Avg Dist to Fast Food/Avg Dist Grocery & Farmers Markets

- 0.24 - 0.65
- 0.66 - 0.83
- 0.84 - 1.00
- 1.01 - 1.20
- 1.21 - 4.30

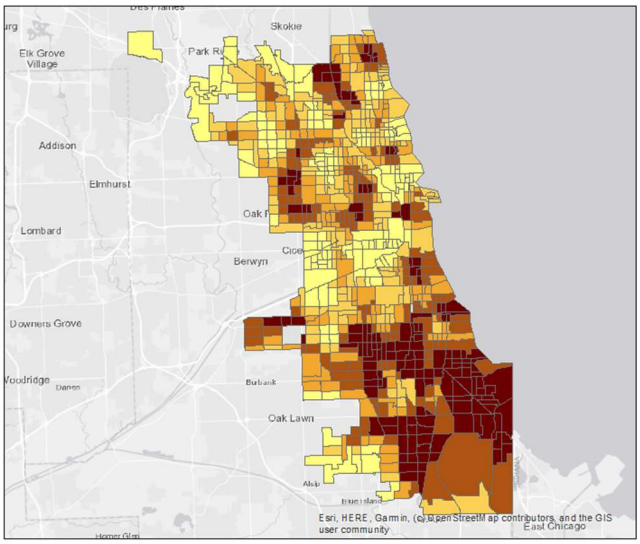


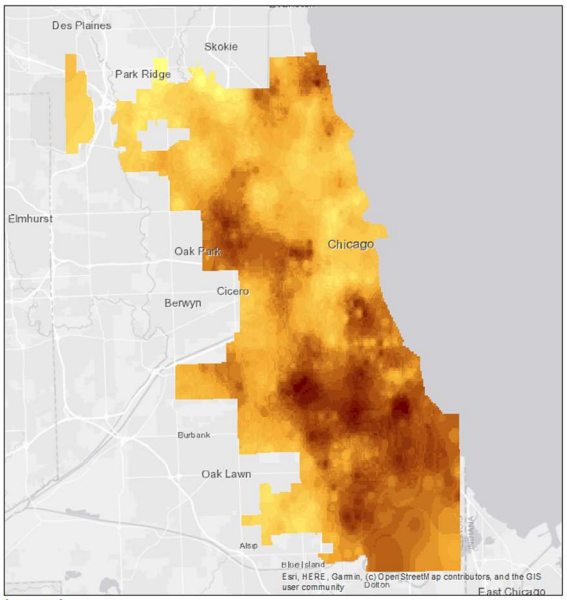
Figure 7: Food Availability Map showing average distance to five closest fast-food restaurants divided by average distance to grocery stores excluding farmers markets

Legend

Proportion of avg dist to fast food over avg dist to grocery stores (5 closest)*

- 0.21 - 0.56
- 0.57 - 0.70
- 0.71 - 0.85
- 0.86 - 1.07
- 1.08 - 4.30

*Darker areas indicate higher density of fast food restaurants than grocery stores



Legend

Summation of variable quantiles (higher is more vulnerable)

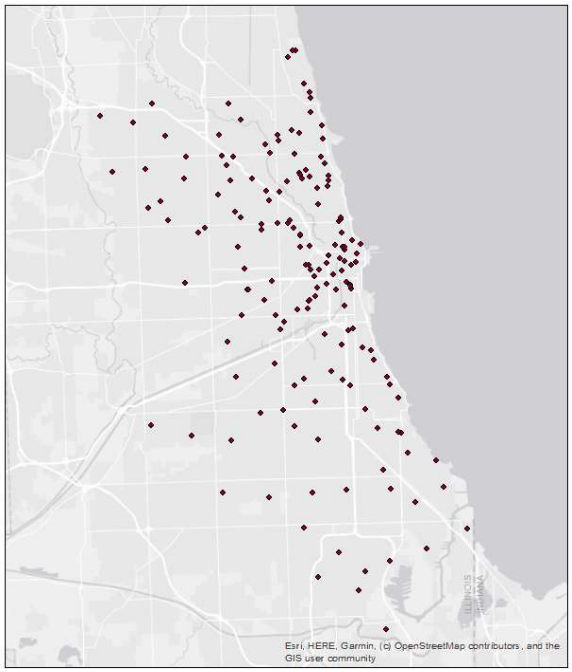
Value

High : 25

Low : 4

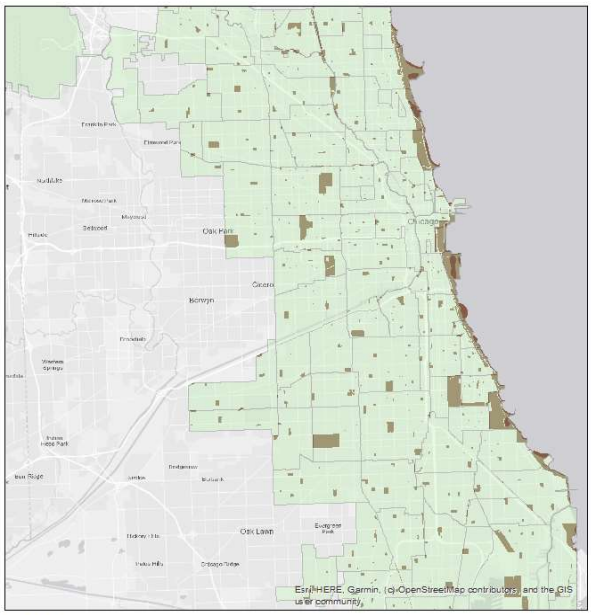
Figure 8: Chicago areas with high fast-food density, low grocery store density, and low income.

Exercise Accessibility Maps



Fitness centers

Figure 9: Fitness center point map (datascraped 2021)



Legend

Neighborhoods boundaries

Parks

Source :Chicago Data Portal

Figure 10: Parks in Chicago (2012)

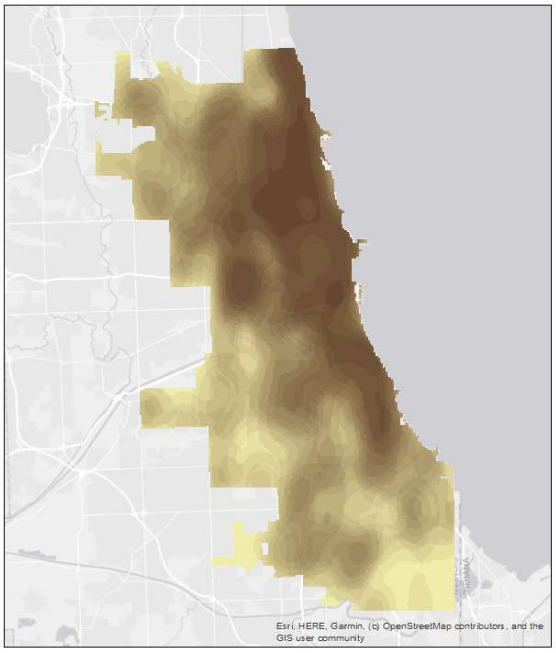
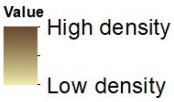


Figure 11: Kernel density estimation plot of fitness centers and parks in Chicago. Darker areas indicate higher densities of parks and fitness centers



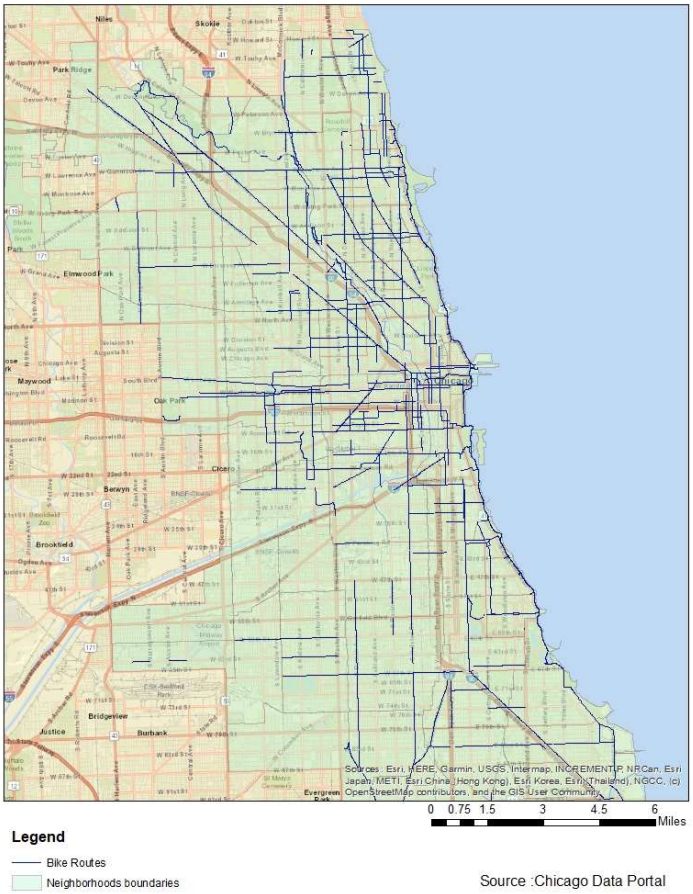


Figure 12: Plot of available bike routes in Chicago (2012)

Figure 9 shows a simple dot map of fitness centers. From this map, it is evident that there are fewer fitness centers located in the south and southwest parts of Chicago. Figure 10 highlights park locations, but this map is somewhat indeterminant as there seem to be parks evenly distributed around Chicago with some of the larger parks along the lake. One map was created to calculate the density of parks and fitness centers in Chicago in a thematic approach. This map shows that the south and southwest parts of Chicago have the lowest densities of these two attributes, but it seems clear from the simpler maps that fitness centers are the driving attribute in this since parks appear to be uniformly distributed.

To supplement this section, simple bike routes map was created to indicate which part of Chicago has access to biking as an alternative and healthier way of travelling. South and west part of Chicago again are the most underserved with fewer or no access to bike paths.

Limitations of Analysis

Several data sources in this analysis are out of date, so more recent data would give more relevant results and would allow for more informative policy suggestions. For food availability, there are likely many other sources of food not included such as convenience stores, non-fast food restaurants, and delivery from non-local merchants. There are also potentially many other places people can exercise such as apartments with gyms or even home gyms, and our webscraping technique may have missed some fitness centers. Additional built environment features could help supplement the analysis and as other researchers have done, additional measures of accessibility would improve the analysis.

Conclusion

Based on our analysis of food availability in Chicago, pure point plots make it difficult to gauge how neighborhoods are served but more complex thematic maps display a clearer story. Thematic choropleths indicated that the north side, loop area, and areas in the near southwest side of Chicago have many accessible grocery store options while the west side and far south side have much more limited accessibility. Fast food restaurant density is strongest around the northside which is likely because of higher overall economic activity, but there are also dense patches in the south side where there are few grocery stores. When viewing these attributes together in terms of overall healthy food availability, it is evident that the southside, parts of the west side, and the very northern edge of Chicago have more unhealthy food options than healthy food options. In regards to policy that could assist these areas, grocery store companies could receive subsidies to incentive opening in locations where healthy food availability is low.

The fitness center point map, kernel density estimation map, and bike path maps also found similar results: the southside, southwest, and west areas of Chicago are underserved in accessibility in areas to exercise. Despite a somewhat uniform distribution of parks throughout Chicago, parks are somewhat inaccessible for most of the year due to weather, making indoor

exercise facilities of huge importance. Many fitness center locations located in underserved communities are run by the Chicago Park District, but other equitable action could be taken to help bring private business owners to these areas.

References

1. Thornton L, Pearce J, Kavanagh A.: Using Geographic Information Systems (GIS) to assess the role of the built environment in influencing obesity: a glossary. International Journal of Behavioral Nutrition and Physical Activity 2011 8:71.
2. Peng Jia, Hong Xue, Xi Cheng, Yaogang Wang, Youfa Wang, Association of neighborhood built environments with childhood obesity: Evidence from a 9-year longitudinal, nationally representative survey in the US, Environment International, Volume 128, 2019, Pages 158-164, ISSN 0160-4120, https://doi.org/10.1016/j.envint.2019.03.067
3. Kolak M, Bradley M, Block DR, Pool L, Garg G, Toman CK, et al. Urban foodscape trends: Disparities in healthy food access in Chicago, 2007–2014. Health Place. 2018; 52: 231–239. pmid:30015180