COVID-19 and Food Insecurity in Chicago

Authors:
Valeria Balza (vbalza)
Mariel Wiechers (marielwb)
Sophia Mlawer (smlawer)
Gabriela Palacios (gabrielapg)

Overview

COVID-19 has revealed many underlying inequities in society. As of March 2021, Black residents accounted for 39% and White (non-Latinx) residents accounted for 22% of the total COVID-19 deaths in Chicago–despite comprising 29% and 33% of the population, respectively.

This trend inspired our group to explore health disparities related to COVID-19 and food insecurity in Chicago.[1] First, we identify zip codes that face restricted access to nutritious foods and whose demographic and socioeconomic characteristics predict higher food insecurity. Second, we identify COVID-19 mortality rates by zip codes. We hypothesize that residents in "food swamps" may be disproportionately affected by risk factors such as diabetes, hypertension, heart disease, and obesity–which consequently result in higher COVID-19 mortality rates.[2] Finally, we explore whether the locations of food banks serve the most food insecure areas.

Our team produced an interactive map of Chicago that visualizes (i) COVID-19 death rates; (ii) a "food swamp" indicator; a (iii) predicted "food swamp" indicator based on demographic and socioeconomic variables; and (iv) the location of food banks. The database leverages the American Community Survey as well as data from the City of Chicago's Data Portal to predict via regression analysis which zip codes are the most food insecure. The final product is a Django web app featuring an interactive map and a data search engine.

---

[1] We define 'food insecurity' as the state of being without reliable access to a sufficient quantity of affordable, nutritious food.

[2] We define 'food swamps' as areas with a high density of establishments selling high-calorie foods relative to healthier options (e.g., fast-food and discount variety stores). While obesity is a complex disease with many contributing factors, food insecurity is a big contributor. See Pan, L., Sherry, B., Njai, R., & Blanck, H. M. (2012). Food insecurity is associated with obesity among US adults in 12 states.

Software Structure

See Annex 1 for a diagram of our structure.

1. Data Sources

1a. American Community Survey

(acs_data) Data from the American Community Survey (ACS, 2019) for Illinois was used to produce food insecurity estimates by zip code for Chicago. The specific demographic and socioeconomic variables extracted are unemployment, poverty, homeownership, and minority population. We model the relationship between these variables and the relative density of high-calorie establishments (food swamp) to establish a predicted food swamp indicator by zip code.[3] Data was retrieved via downloaded CSV files: (i) demographic data to extract race and gender information; (ii) employment data which includes income observations; and (iii) housing data.

1b. COVID-19 data

(covid.py) From the City of Chicago's Data Portal, we used an API to download information on the COVID-19 death rate (per 100,000) people by zip code. Data was processed using pandas.The map at the bottom shows the distribution of COVID-19 death rates by Chicago zip code.

1c. Food Swamp Measure

(grocery_stores.py) From the Chicago Data Portal, we use an API to download information on all the grocery stores in Chicago by zip code.

(fast_food.py) From Wikipedia, we use web scraping to download the list of major fast food restaurants in the United States.

(business_license.py) From the City of Chicago Data Portal, we downloaded business licenses for food retail, generated a list of chain discount variety stores, and used **fast_food.py to** generate the names and locations of all the fast food restaurants in Chicago using the fuzzywuzzy package.[4]

(food_swamp.py) This combines the grocery store and fast food restaurant counts by zip code and generates the ratio of unhealthy to healthy food for each zip code:

$$Food\ Ratio\ =\ \frac{Fast\ Food\ Restaurants + Chain\ Discount\ Stores + 1}{Grocery\ Stores + 1}$$

1d. Food Banks

---

[3] Our model for predicted food swamp indicators was based on *Feeding America*'s Food Insecurity Rate methodology for the Map the Meal Gap 2020 study (see pp. 7). The intuition behind this model is that higher unemployment and poverty rates are associated with higher access to low quality food options. These vulnerable groups likely qualify for federal nutrition assistance programs, but their surroundings are still environments that lack access to affordable highly nutritious food.

[4] This is probably not fully accurate because it does not account for small convenience stores or local fast food restaurants.

(food_banks.py). Using the beautifulsoup package, we web scrape the Greater Chicago Food Depository to identify the addresses of food resources (food banks, pantries, etc.) and then convert them to latitude and longitude coordinates.

2. Predictive Component

Based on Feeding America's regression model, we examined the relationship between the unhealthy-to-healthy food ratio and socioeconomic factors for each zip code.[5] We then generated predictions of an expected food swamp ratio for each zip code. The map of actual and predicted food ratio can be seen below.

(3) Visualization

We then combined our data to produce a Django web application (**CS_covid_food**) featuring (i) a map of (i) COVID-19 death rates; (ii) a "food swamp" indicator; and a (iii) predicted "food swamp" indicator based on demographic and socioeconomic variables. The application also features a search engine that queries the database based on the different variables. This application includes the files **map_query/views.py/models.py** and **obs.py**, which build the interactive map and the search engine.[6]

Results, Shortcomings, Future Work

Our original hypothesis was that areas with a higher "food swamp" ratio would coincide with the areas most affected by COVID-19–as there are a lot of similar risk factors for both. However, as can be observed in our visualizations, both static and interactive, this trend does not appear to be captured.

There are many potential reasons for this shortcoming. First of all, in our process of determining unhealthy and healthy sources of food, we only included national fast food chains and variety stores. Other studies were able to include local options (and see if they were healthy or not) by actually travelling to the different Chicago neighborhoods.[7] Also, it appears the map shows the highest food swamp ratio in tourist-dense areas so it is possible that our regression could be refined with the addition of a tourist area control variable.

Furthermore, it could be the case that actual location-based access to food sources is not influential on health outcomes. When thinking about policies to improve health and access to nutrition for underserved communities, introducing healthy food options might not be an efficient approach. Additional research suggests that the mere introduction of healthy options would not

---

[5] See footnote (3).

[6] Please note the search query component modifies code used for Programming Assignment 3. The following tutorials were used to develop the Django project, which uses GeoDjango: https://www.youtube.com/watch?v=L3YoX9wrGDc and https://docs.djangoproject.com/en/3.1/intro/tutorial01/.

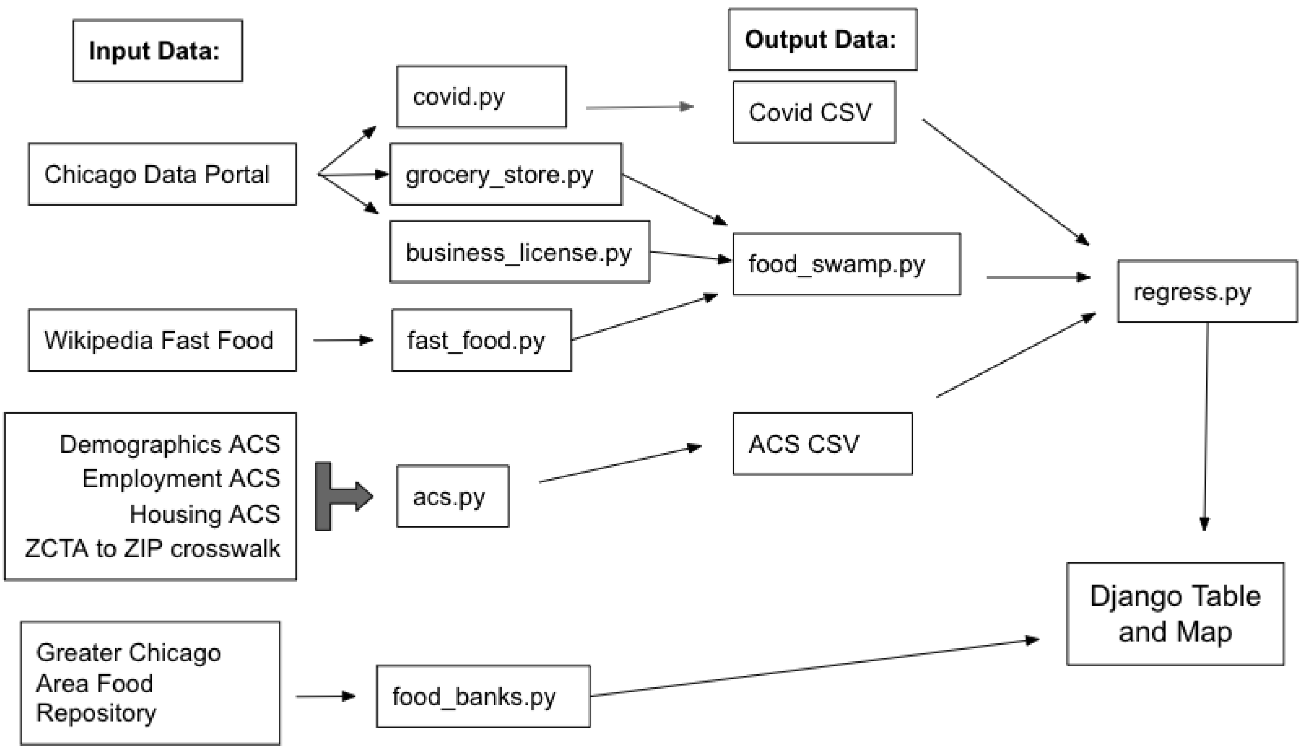
[7] Kolak M, Bradley M, Block DR, Pool L, Garg G, Toman CK, Boatright K, Lipiszko D, Koschinsky J, Kershaw K, Carnethon M, Isakova T, Wolf M. Urban foodscape trends: Disparities in healthy food access in Chicago, 2007-2014. Health Place. 2018.

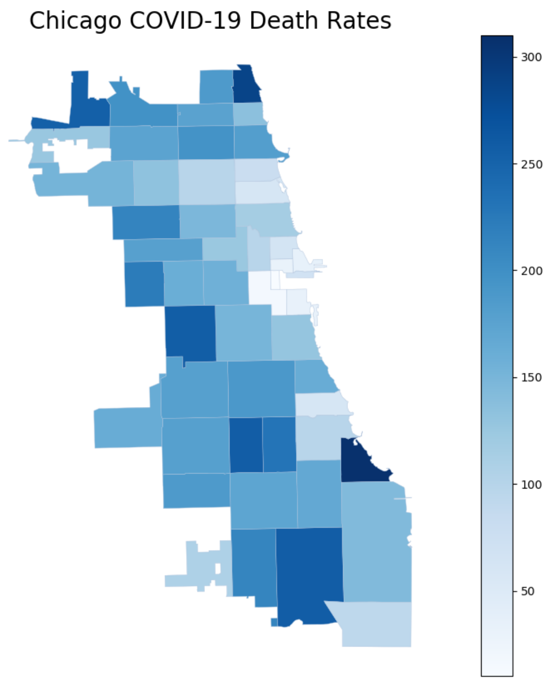
have a significant impact if not paired with efforts to influence eating habits and offset the cost of the healthier food.[8] It is most likely that an interplay of socioeconomic factors, themselves products of institutionalized inequalities, are risk factors of these communities. More than identifying the causes of health inequality, further research could aid in guiding action around improving quality access to resources that have a bigger impact on community health.

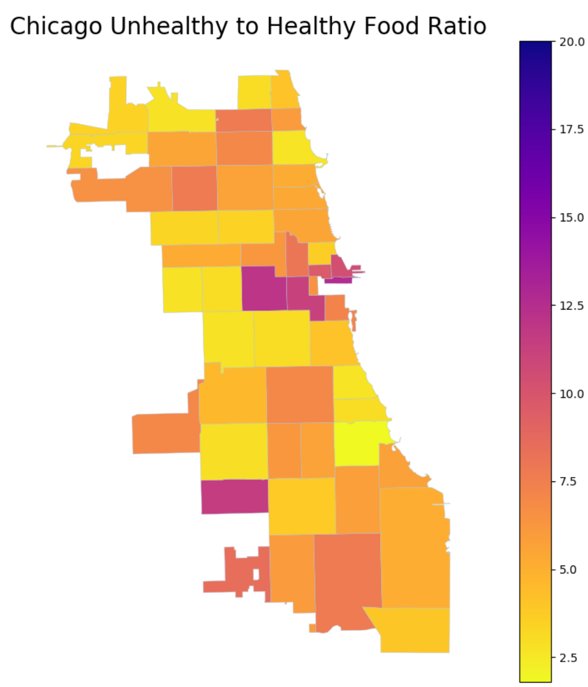
Finally, in relation to the resources related to food insecurity, it appears that locations are spread out throughout Chicago though more analysis would need to be done in order to determine any potential gaps for underserved populations.

[8] Ibid.

Annex 1.


Input Data:
Output Data:
covid.py
Covid CSV
Chicago Data Portal
grocery_store.py
business_license.py
food_swamp.py
regress.py
Wikipedia Fast Food
fast_food.py
ACS CSV
Demographics ACS
Employment ACS
Housing ACS
ZCTA to ZIP crosswalk
acs.py
Django Table and Map
Greater Chicago Area Food Repository
food_banks.py


Chicago COVID-19 Death Rates
300
250
200
150
100
50


Chicago Unhealthy to Healthy Food Ratio
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5

Chicago Predicted Unhealthy to Healthy Food Ratio