Evaluating Insufficiencies in Food Market Accessibility Across U.S Metropolitan Populations

Aayush Joshipura, Amanda Stark, Danny Heinrichs

DS2001 Science Programming Practicum Cocola

December 11th, 2020

Abstract

In this project, we explored the presence of "food deserts" across different metropolitan areas in the United States. We examined the role of low-income statistics, poverty, income levels, and geographic location in order to create connections to continuing food scarcities across America. The project was conducted through use of dataset cleaning, dataframe creation, geocoding, and other methods. Using professional studies as a basis we built on prior research by extracting census data from the USDA Food Access Research Atlas to draw comparisons between different U.S cities and evaluate statistics regarding food accessibility and income levels. We isolated Baltimore and mapped food access in the city to evaluate a single city. while discussing successful efforts in Boston as a standard for other cities. Our analysis let us compare incomes and food accessibility statistics between major U.S cities and identified the key factors that cause discrepancies. In addition, we described legislative and economic efforts in Boston to display how cities may improve the situation. Ultimately, we made three key conclusions: (1) Food deserts disproportionately affect low-income populations, (2) the severity of the issue varies among different urban areas across the US, and (3) Boston's success proves improving infrastructure and increasing grocery store access work only when there is a nutritional focus.

Introduction

Despite innovations in technology and endeavors to reduce hunger, convenient access to grocery stores and food sources continues to pose an economic challenge for legislators and city planners across the United States. About 23.5 million people in the United States today live in geographic areas where access to affordable, healthy food options are limited or nonexistent because grocery stores are too far away (DoSomething 2015). Otherwise known as "food deserts", these areas exist in metropolitan and suburban areas alike, afflicting populations with food scarcities and low quality service. The situation further worsens when income is taken into consideration as "food deserts" disproportionately affect low-income areas, with approximately 2.3 million people (2.2% of all US households) living in low-income, rural areas that are more than 10 miles from a supermarket (DoSomething 2015). Without proper sources of nutrition, large portions of these deprived populations are forced to turn to unhealthier choices, such as fast food and other cheap alternatives. These people continue to suffer under these "food deserts" as they fail to lift themselves out of their socio-economic struggles.

Various factors contribute to the propagation of the "food desert" problem across the country today, including the state of low-income areas, insufficient legislation, and promotion of inferior alternatives. For starters, the low-income areas in which many "food deserts" continue to exist, have poor infrastructure to support the growth of additional grocery stores, food banks, and other resources. Many afflicted areas showcase high poverty rates, disease rates, and crime rates, which discourages efforts to establish business in the area. These problems are only furthered by insufficient legislative efforts and lack of awareness from local and state governments regarding the issue. American cities have been negligent in attacking "food deserts", continuing the use of fast food restaurants as alternatives to healthy choices in low-income communities. These actions have influenced much of the low-income population to pursue cheaper, unhealthier alternatives, even when grocery stores are launched, due to developed dangerous eating habits that are difficult to reverse. Solely building grocery stores cannot solve the problem, as we need to shift dietary habits in these low income communities. Legislation is necessary to create change and current efforts in many U.S cities fall short of this goal.

Observing current research in the field, we were able to identify two professional studies that provide insights into the "food desert" problem. Shima Hamidi, a researcher in collaboration with John Hopkins University conducted a 2020 study regarding the urban sprawl of food deserts and possible solutions to encourage investment from grocery retailers. She cited statistics from the USDA database to support her claims regarding the initiatives urban areas can take in order to promote grocery store access in their communities (Hamidi 2020). A similar study was conducted in 2008 by the United States Department of Health and Human Services (HHS), where they evaluated the different characteristics of food deserts and what initiatives could be taken by local governments (Karpyn, Riser, Tracy, Wang & Shen).

While these studies serve as a baseline to operate from, our study fills some of the holes in current research, specifically the lack of geographical comparisons between cities across the United States and a lack of visuals between the relation between low-income populations and food deserts. With this vision in mind, our project goals sought to address these issues. Using data science, we sought to find how prevalent these food deserts are across the country and visualize them geographically. In addition, we hoped to illustrate a relationship between income and food deserts by analyzing different metropolitan areas and their low-income populations.

Most importantly, we sought to conduct this research using different metropolitan areas in order to cross analyze any factors or efforts between cities. With this research, we can identify possible solutions by observing successful cases and their endeavors in combating the problem. Are food deserts a wide ranging issue, are metropolitan cities across the US facing the problem differently, and can we use data science to highlight useful information to increase food accessibility?

Data

The dataset that we used for our research is provided by the Food Access Research Atlas, a database assembled by the United States Department of Agriculture. The Food Access Research Atlas was designed to "present a spatial overview of food access indicators for low-income and other census tracts using different measures of supermarket accessibility" and further "provides food access data for populations within census tracts" (USDA). Regarding the overall structure of the data, the database was available for use on DataWorld or as an Excel spreadsheet from the USDA website. The dataset provides descriptive statistics critical to identifying "food deserts" and their impacts on the populations, reporting poverty rates, food access, grocery store proximity, and other variables for hundreds of cities across the United States. The set contains the state, country, population, housing units, and census tracts of urban areas. For each area, it identifies areas of low income, low food access, high poverty, low median family income, food scarcities, and other key variables. We selected the Food Access Research Atlas due to the comprehensive information it provided about the variables we sought to examine. In particular, we wanted to analyze the role of income, poverty, and other demographic factors involved in the food desert issue and the Atlas provided all of these variables in great depth. It sorted proximity to stores by different miles, allowing us to see how many people were 10 or 20 or more miles away from a certain location. In addition to its valuable information, the data is trustworthy, provided by a trusted government source and frequently updated by the Department of Agriculture. In summary, USDA's Food Access Research Atlas served as our dataset due to its detailed analysis of key variables and trustworthy source.

Methods

For the data science component of the project, our methodology consisted of several steps that ensured we could generate comprehensive and descriptive visuals to present. Our code was

created utilizing information from our dataset and the use of Jupyter Notebook to run procedures. Our first step in creating our final product was to isolate the census data from our dataset and convert the file into a CSV file. We made this conversion to develop code after cleaning the dataset and reading the file into a dataframe. Next, to begin our analysis across different U.S cities, we created new dataframes for each urban area and a synthesized dataframe to show the differences between them. This step was necessary to illustrate the cross-analysis we aimed to complete in our project, comparing performance and statistics between the different cities we selected. Then, we added columns of calculated percentages of different populations in food deserts to the new dataframe to standardize the data. Our next task was to create graphs showing the food desert distribution in different cities, with additional graphics displaying the effect income has on a person's food access. These graphs were generated to conduct the analysis of metropolitan populations and the impact their income status has on their surrounding area. After these visuals were displayed, we began a procedure to isolate our selected case study in Baltimore. We utilized dataset cleaning after isolating the Baltimore grocery store location data. Then to create our detailed food access map, we utilized available census data and geocoding to create a unique dataframe for the city. From here, we mapped grocery store locations in Baltimore. Through these steps, we were able to display the graphs and food access map that was critical in presenting our research findings.

Results & Discussions

With the quantifications of the number of food deserts, populations in food deserts, and low income populations in food deserts, we were able to determine which cities had the most people affected by food deserts (Fig 1A/B). However, by finding the ratios of people in food deserts compared to the whole city population, we were able to standardize the data so that the cities could be compared on the same scale without the factor of size affecting the findings. From this, new cities emerged as having larger issues with food access than expected. For example, almost a third of the populations of Atlanta and Austin are people experiencing food

Figure 1

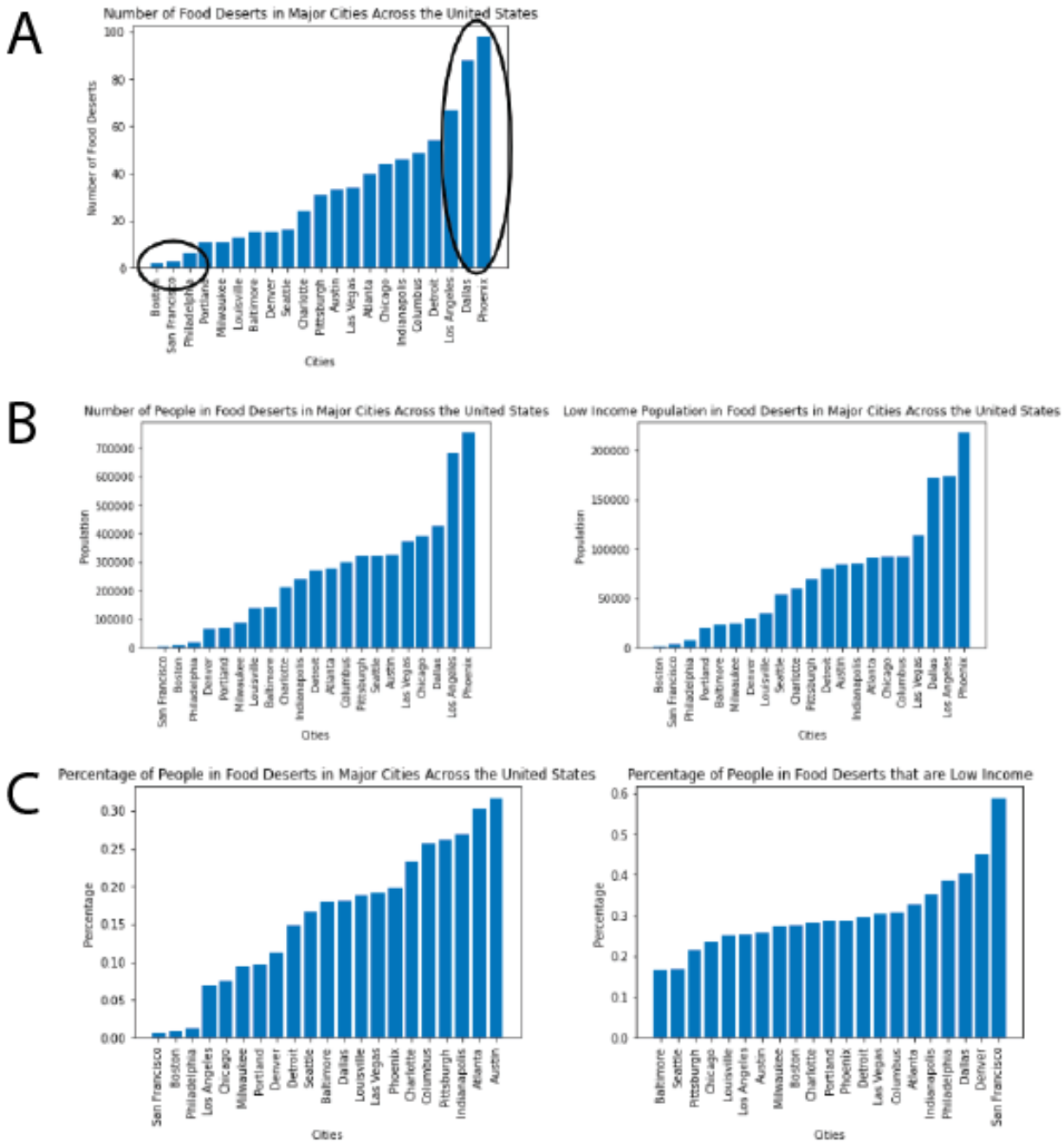
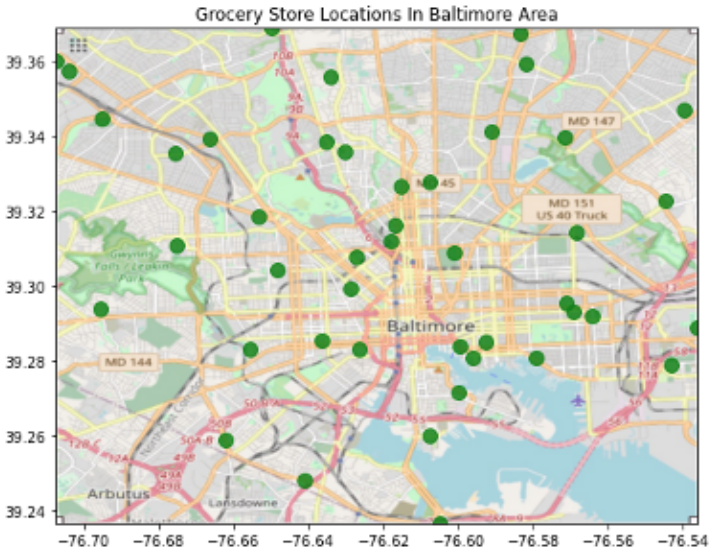


Fig 1. **These represent differential populations in food deserts throughout US Metropolitan Areas.** A) There is a wide range in the number of food deserts, as quantified by census tract areas, in major cities. Boston, San Francisco, and Philadelphia are seen to have the lowest numbers of food deserts and Los Angeles, Dallas, and Phoenix have the largest numbers. B) General populations as well as low income populations in food deserts across major cities show a similar trend to the number of food deserts (as shown in A). C). The percentage of people in food deserts was quantified by dividing the general and low income populations in food deserts by the total metropolitan area population from 2010 (same year as the other census data). Indianapolis, Atlanta, and Austin arise as cities with a large percentage of their populations in food deserts.

access difficulties by living in food deserts (Fig 1C). By isolating the data sets for metropolitan cities in the United States and displaying which cities have the largest issues regarding lack of food access, we can help nonprofits and policy makers understand the depth of the issue and find solutions by looking to the cities with low numbers of food deserts or low percentages of low income people in food deserts depending on the issue they want to tackle. Although there aren't any solutions found just by isolating this data set and highlighting the cities with the best and worst food access, we hope this can be a great resource for others. Food access is a fundamental and non-partisan issue that affects low income populations the most, so we hope that aggregating the data for easier use could help people and organizations with more power and influence tackle the source of the problem.



Using a provided street map of Baltimore, we isolated Baltimore's food distribution and represented access across the city (Baltimore Street Map) A geocoded map of Baltimore grocery store locations reveals gaps in store coverage. These are noted with blue rectangles and represent areas where residents have to travel a significant distance to buy food. This map adds an additional visualization and context to the trends the graphs above describe. City planners can use this data to plan new locations for grocery stores.

Throughout our analysis, Boston stood out as an extremely successful case in relation to the city's contemporaries across the United States. To examine the reason for their high ranking performance, we conducted a qualitative analysis inside Boston's legislative and economic strategies to counter the problem, with the hopes other cities could feasibly implement similar initiatives. Utilizing a study by Kimberly Etingoff and Kimberly Zeuli, we found Boston's success largely stems from their commitment to add a grocery store in every neighborhood, sustained legislation to continue enacting changes, and supporting programs to increase

affordability and availability of healthy food. Under the guidance of the mayor, Boston has added grocery stores in poorer neighborhoods by improving surrounding infrastructure. This action helped raise incomes, provide jobs, and lower crime rates in these cities (Etingoff & Zeuli). In addition to these new stores, Boston sponsored healthy habits by doubling farmer market purchases, starting healthy corner store programs, and expanding agriculture (Etingoff & Zeuli). These initiatives led to success in Boston and could easily carry over to other major U.S cities.

While we were able to achieve our research goals and contribute valuable information to the research field, our project contended with different challenges both logistically and with our programming that may limit our data. Regarding practical challenges with coding, we struggled with converting our dataset into our new data frame to fit the needs of our project. The Food Access Research Atlas was very informative and served as a relevant dataset for our research, but the sheer size of the database made the file difficult to parse through for the necessary information, which created challenges for us during coding. Processing large data sets throughout this project was also quite difficult and greatly slowed our early progress when we ran into errors. Learning and utilizing geopandas had a similar slowing impact, requiring us to troubleshoot code multiple times. Logistically, many potential datasets are locked behind paywalls or university libraries, which restricted us from useful information. As a result,our data is limited as we were unable to examine specific variables regarding transportation, grocery store concentration, and statistics regarding race. Acquiring information on these variables would have enhanced our understanding and provided more insight into why people are unable to access nutritional food sources. Despite these setbacks, we achieved our primary research goals and created detailed graphics that we believe represent original contributions to the field.

Summary & Conclusions

In summary, we sought to explore the prevalence of "food deserts" across different metropolitan areas in the United States. In this analysis, we aimed to isolate the role of variables such as percentage of low-income residents, poverty, income levels, and location in order to study their effects on the continuing food scarcities in these areas. Building on prior research, we synthesized data from the USDA Food Access Research Atlas to draw comparisons between different U.S cities and their ability to feed their populations. We extended this research further

by isolating Baltimore as a case study and mapping food access in the city, while discussing successful efforts in Boston as a standard for other cities. Through our analysis, we distinguished food accessibility statistics between different major U.S cities and worked towards finding solutions by identifying key factors that may cause discrepancies. In addition, we described legislative and economic efforts in Boston to show how cities may succeed in mitigating the problem. All in all, we were able to make three key conclusions: (1) Food deserts disproportionately affect low-income populations, (2) the severity of the issue varies among different urban areas across the US, and (3) making improvements in overall infrastructure and increasing grocery store access may mitigate the problem, as evidenced by Boston's success.

For future endeavors, our research can be expanded on with numerous different strategies that enhance our efforts. Researchers can expand by introducing more demographics into the data analysis to gain additional insights about what contributes to deserts, whether that be race, location, or another variable. They can also expand our mapped visuals by heat mapping census tracts to create a more advanced and detailed graphic given additional time. Examining nutrition and cultural appropriateness of foods was another topic we discussed briefly, but could certainly be expanded on greatly in a supporting study, as we focused more on relations to income and geography. Outside of building on research, we hope that someone who views our study may try to conduct a similar analysis in other countries to analyze food deserts internationally. We can spread awareness by sharing our findings with higher entities in local governments to examine data and advocate for solutions. Ultimately, we believe our project adds new insights to the field and displays valuable information to solve a pressing problem in our country today.

Project Contributions

Throughout this process, the team collaborated well and divided work evenly, despite the unique circumstances. Aayush Joshipura was heavily involved with the report, writing 4 sections of the paper. He also built the presentation outline and filled out slides for the Update and Presentation, alongside code contributions. Danny Heinrichs created the presentations and finished numerous slides. He took a larger role in coding, parsing data and creating graphs, also writing the Data and Methodology sections. Amanda Stark provided the project idea, with major coding contributions, using geocode and other techniques to create our graphs. She input all of

our graphs and explained our data in the Results section of our paper. Amanda also filled out slides for the update and final presentation. Overall, our group worked cohesively and efficiently.

References

1. "11 Facts About Food Deserts." *DoSomething.org*, 2015, www.dosomething.org/us/facts/11-facts-about-food-deserts.

2. (Dataset 1) https://www.ers.usda.gov/data-products/food-access-research-atlas/

3. "Baltimore Open Street Map." *Open Street Map*, 2020, www.openstreetmap.org/#map=12/39.3040/-76.6250.

4. Etingoff, Kimberly, and Kimberly Zeuli (2017). "A Grocery Store in Every Neighborhood: Boston's Quest to Eliminate Food Deserts." icic.org/blog/grocery-store-every-neighborhood/.

5. Hamidi, S. (2020). Urban sprawl and the emergence of food deserts in the USA. Urban Studies, 57(8), 1660–1675. https://doi.org/10.1177/0042098019841540

6. Karpyn AE, Riser D, Tracy T, Wang R, Shen YE. The changing landscape of food deserts. *UNSCN Nutr*. 2019;44:46-53.