A study of Food Deserts in the Madison, WI Neighborhoods

Josh Jarvey

November 16th, 2019

1. Introduction

1.1 Study Description:

A food desert is an area that has limited access to affordable and nutritious food, in contrast with an area with higher access to supermarkets or vegetable shops with fresh foods, which is called a food oasis. The designation considers the type and quality of food available to the population, in addition to the accessibility of the food through the size and proximity of the food stores.

In 2010, the United States Department of Agriculture (USDA) reported that 23.5 million Americans live in "food deserts", meaning that they live more than one mile from a supermarket in urban or suburban areas, and more than 10 miles from a supermarket in rural areas.

Food deserts tend to be more populated by low-income residents, who are not as attractive a market for large supermarkets, and have reduced mobility. Food deserts lack suppliers of fresh protein sources such as poultry, fish and meat, along with fresh fruit and vegetables, instead, relying on convenience stores, which provide processed and sugar- and fat-laden foods, which are known contributors to the United States' obesity epidemic.[1]

1.2 Problem Definition & Stakeholder Identification:

In this capstone project, I will focus my study of food deserts within the various, and unique neighborhoods that make up the city of Madison, Wisconsin (a place I call home!).

Recently, a proposed project on the cities' south side caused a lot of community concern[2] when the project asserted to demolishing a Pick-n-Save supermarket to build a new health care clinic in its place, leaving this community without access to fresh produce within the guidelines posted above. While nothing to date has taken place with the development project, this also does raise the interesting question of not only determining current desert areas, but those also at risk of becoming a food desert.

Target stakeholders of this study are the various community groups & leaders, associations, alderman's, city officials, lawmakers, and even supermarket businesses that operate in the city.

2. Data acquisition, cleaning, and feature selection

2.1 Data Sources:

This study will require datasets that contain not only geographical information about the neighborhoods that make up the city of Madison, but also geographical information on grocery stores within the area. I have identified two key data sources as described below:

i) The city of Madison provides a repository of different datasets via their Open Data Portal[3]. This repository contains various datasets via the ARCGIS open data module. I will use the Neighborhood Plans[4] data set to collect information about the different neighborhoods within the Madison area.

ii) Using the Foursquare Places API[5], I will focus my queries on the venue name "Supermarket" & "Grocery Store", using the neighborhood data as a reference point. This will provide the necessary location data of the specified venue to conduct the analysis.

By combining these two datasets, it should provide an easy and intuitive visualization to identify areas of the cities and their associated neighborhoods of where food deserts exist within the city. I also expect to identify areas of risk for which a food desert could occur based on the density of fill from overlapping radii of the venues.

2.2 Data Cleaning:

As with every data project, an effort of exploration, cleansing, and aggregation must take place. An initial review of the Neighborhood plans dataset identified geographic data of neighborhood locations was in the format of ERSI standards and uses a "local" coordinate system described at this website[6].

9/7/2017 Feature Layer Custom License

Download APIs

Showing 1 to 10 of 27

Hint: Filter columns using

PlanName	Status	Year	SHAPE_Leng	Link	Shape.STArea()	Shape.STLength()
Schenk-Atwood-Starkweather-Wo...	Adopted	2000	23821.4443845	https://www.cityofmadison.com/...	28355961.548706055	23821.445050297672
Tenney-Lapham-Old Market Neig...	Adopted	2008	23041.9865548	https://www.cityofmadison.com/...	13296750.531860352	17193.245760875183
Marquette-Schenk-Atwood Neigh...	Adopted	1994	24224.7831786	https://www.cityofmadison.com/...	22165823.134490967	24224.783772685994
Brittingham-Vilas Neighborhood ...	Adopted	1989	21127.2789675	https://www.cityofmadison.com/...	19990114.138427734	21127.27945590609
Bay Creek Neighborhood Plan	Adopted	1991	18309.3863695	https://www.cityofmadison.com/	15800758.821472168	18309.38670984485

Since I am unfamiliar with GIS data and ERSI standards, I conducted a bit of research on GIS systems and Shape files. I learned that using this information, a GIS program can calculate Latitude and Longitude from this Shape information.

I contacted the city of Madison via the online question form, and was issued a response almost immediately that I should be able to download the .Shape files from the website, and then using an open source program such as QGIS[7]. I could calculate my required values to continue my progress. Ultimately, using the QGIS program allowed me to convert the Shape data into Latitude and Longitude, and I was able to load the resulting .CSV output into my project.

Focusing on my second data set for this project, I use the Foursquare API 'explore' method by passing in specific parameters to return venue information in a .JSON format. First, a URL needed to be crafted for each unique neighborhood in order to return the results. The following outlines the specific values used and the reason it was selected:

Parameter	Description	Value	Reason
Query	the key word passed to the API call in order to match on and return specified results	'Grocery Store' 'Supermarket'	I select two terms to query so I can conduct a comparison between the two
Latitude	the latitude value used to define the central point of the search query	[iterate each neighborhood's latitude]	Each neighborhood required this value to create its centroid
Longitude	the longitude value used to define the central point of the search query	[iterate each neighborhood's longitude]	Each neighborhood required this value to create its centroid
Radius	the measurement (in Meters) of which to complete the search based on the latitude and longitude center point	5,000 meters (~3 mi)	This size selection was inclusive enough to gather all the available stores near this neighborhood (with some overlap)
Version	the timestamped version of the API call	'20191115'	The date I completed the project

Each neighborhood's latitude and longitude were passed to the algorithm in order to create a list of unique URL calls that would ultimately be executed against the Foursquare API to return the information required.

2.3 Feature Selection:

After converting the shape files to latitude and longitude and importing into the notebook, I would only select the features required for the food desert analysis and rename them for greater intuition. The remaining columns would be dropped once loaded into a dataframe.

Secondly, when a call is made to the API, the .JSON file is returned containing a rich set of feature data as it relates to that search. This ultimately provided more information than what's needed for this analysis, so I selected only the values required. See the table below for more specific details.

Data Source	Feature Selected	Action	Result (dataframes)
Converted City's Open Data Portal	PlanName, Y, X	.rename('PlanName':'Neighborhood') .rename('Y':'Latitude') .rename('X':'Longitude')	Madison Neighborhood dataframe containing name and coordinates of the neighborhoods
Foursquare	'Category' name, 'venue' name, formatted address, latitude, longitude	Iterate through the returned .JSON object and extract each feature into a list. Appended list to an aggregated list. Dropped any duplicates due to overlapping search radius.	Grocery Store dataframe containing unique grocery store and coordinates within the studied area.

3. Conducting the Analysis

3.1 Approach:

In order to provide a clear intuition of the datasets, I decided to use visualization methods as my primary tool in conducting my analysis. Folium[8] is a great visualization package available for use in the Python programming language, and with just a few lines of code I would be able to plot out my neighborhoods and grocery stores onto a map of the city.

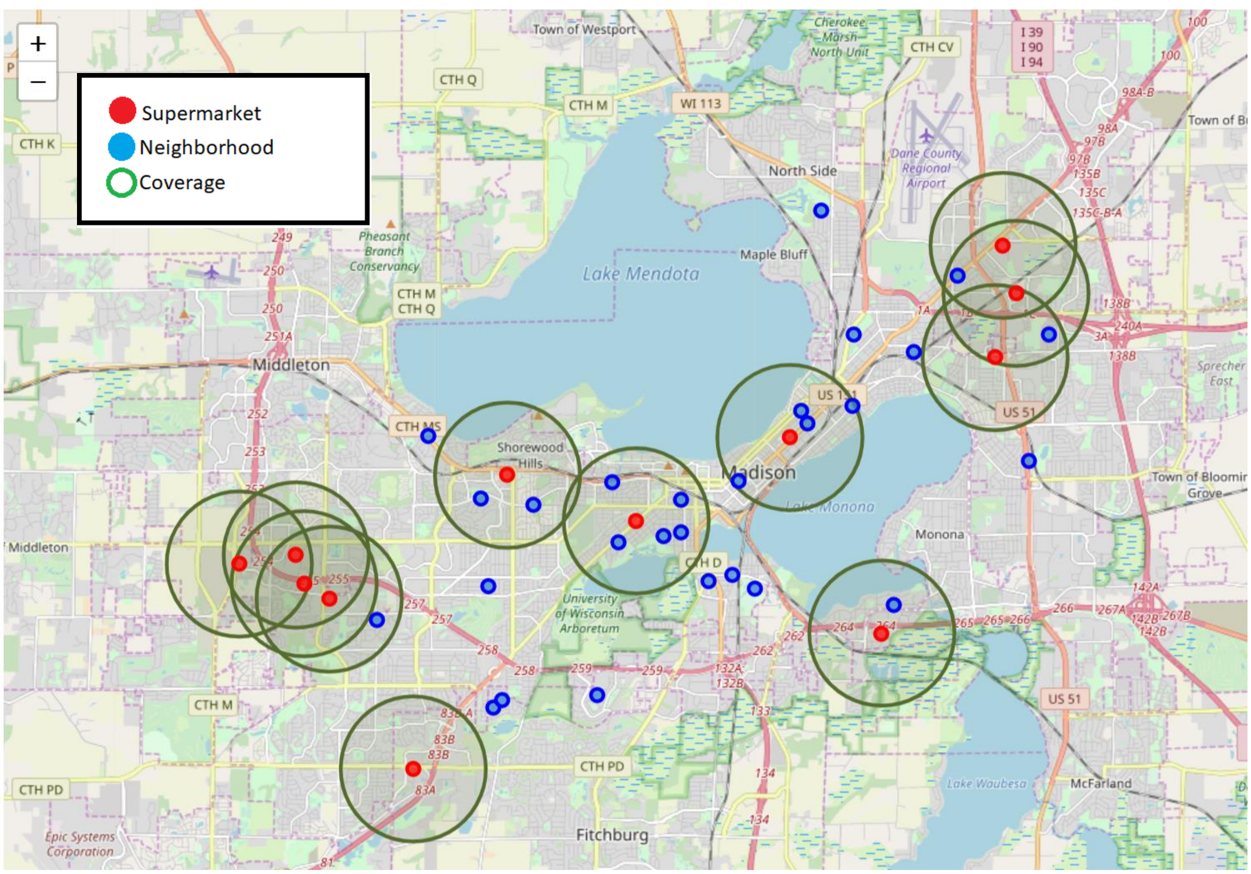
To build my Food Desert map using the Folium package, I execute the following steps:

1) I will first start out by iterating through the grocery store dataframe, plotting a circle object for each entry on the map. The size of the circle object will have exactly a 1610 meter (e.g. 1 mile) radius in order to indicate its “coverage” of that area. The opacity of the object will also be set to 0.1 (e.g. 10%) in order to also identify overlapping “coverage” (e.g. indicating those communities are at less of a risk).
2) I will iterate through the entire grocery store dataframe a second time, plotting a red circle marker object for each entry on the map on top of the existing map structure. When the icon is clicked by the user, it will pop up a window and provide the store name.
3) Finally, I will iterate through the neighborhood dataframe, plotting a blue circle market object for each entry on the map on top of the existing map structure. When the icon is clicked by the user, it will pop up a window and provide the neighborhood name.

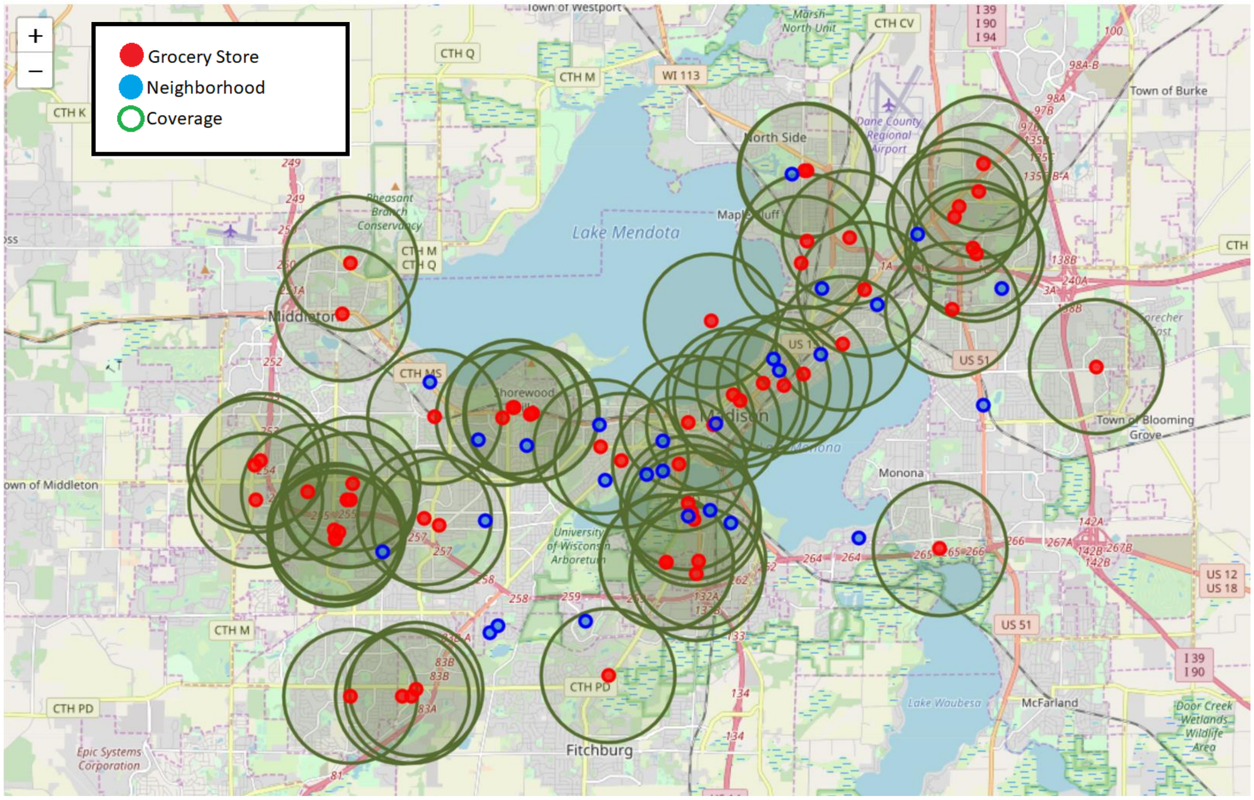
3.2 Maps:

In the following section I provide screen prints of the maps that were generated using the datasets created for this study.

<u>“Supermarket” vs. neighborhoods in the city of Madison, WI</u>



"Grocery Store" vs. neighborhoods in the city of Madison, WI



3.3 Analysis Results:

From a visual perspective, you can see there is a clear difference between the two maps, and how a slight change in terminology used when calling the Foursquare API can generate drastically different results.

Looking at the "supermarket" map, one may get the impression that there are multiple food deserts in the city of Madison. However, upon further investigation of the map, I notice a large supermarket chain is missing from the dataset – Roundy's (e.g. Pick-n-Save) – which is a popular grocery chain here in the upper Midwest that offers fresh fruits, vegetables, and meats.

Analyzing the "grocery store" map, one may get the impression that there is a food oasis (opposite of a desert), albeit a few small areas that suffer from a desert situation. Upon further investigation of the map, I notice a large representation of "mom-n-pop" & ethnic grocery stores, which don't all necessary meet the guidelines per the offering of fresh fruits, vegetables, and meats.

Although the maps contrast quite differently upon initial review, you will notice similar voids of coverage when overlaying the two. More specifically using the Foursquare data, members of these neighborhoods are either in, or at risk, of living in a food desert: Broadway-Simpson-Waunona, Allied-Dunn's Marsh, Southwest, Stoughton Road Revitalization Project, South Madison.

4. Conclusion

This study has provided an opportunity to take a bird's eye view, geographically, of Supermarkets and Grocery Stores in the various neighborhoods within the city of Madison, Wisconsin in order to identify areas of risk. I've taken city neighborhood location data, plotted against Foursquare venue location data, and analyzed "grocery store coverage" according to the parameters of the USDA definition of a food desert.

I discovered that Foursquare data should be considered unreliable for this type of study due to the integrity of the search results. There were missing results when using the "supermarket" search term (very large chain store – Roundy's), vs. using the "Grocery Store" search term returned too many results including Mom-n-Pop convince stores and ethnic stores that don't not tend to carry fresh fruits, vegetables, and meats that meat the USDA definition when evaluating a food desert.

5. Future Direction

While this study uncovered there are some areas of concern within the city, it would take further analysis, additional datasets & features, and geographic knowledge to further isolate and identify these areas with greater confidence. For example, adding the dimension of median household income or poverty rate within the neighborhoods has been strongly correlated to be at risk of a food desert.

Ultimately, this project provides some of the foundational analysis to continue the study of food deserts. Using publicly available datasets with a few lines of python code, the compiled maps provide "areas of further interest" for stakeholders to investigate further.

Tools & Sources:

[1] Food Desert description: https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/

[2] Park street resident's concerns over zoning proposals : https://isthmus.com/news/news/park-street-residents-worry-about-what-will-happen-if-pick-n-save-grocery-closes/

[3] City of Madison open data portal: https://data-cityofmadison.opendata.arcgis.com/

[4] Neighborhood plans dataset: https://data-cityofmadison.opendata.arcgis.com/datasets/neighborhood-plans

[5] Foursquare Places API: https://developer.foursquare.com/places

[6] Spatial Reference System: https://spatialreference.org/ref/sr-org/6675/

[7] Open GIS, QGIS website: https://qgis.org/en/site/

[8] Folium map package website: https://python-visualization.github.io/folium/