Alex Fielder

A02367780

April 16th, 2024

WATS 4930

“The Journey to Freshness: A Geospatial Analysis on Travel Distances to Farmers Markets Using ZIP Codes and Transaction Data”

Abstract

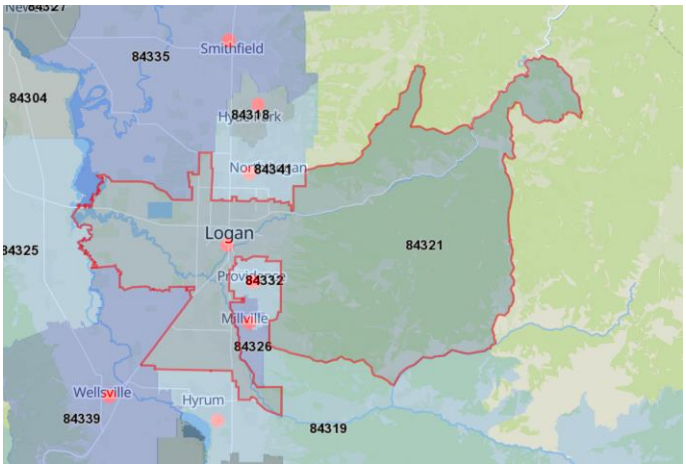
The double up food bucks program is an incentive to get SNAP (food stamp) users to eat more fruits and vegetables grown by Utah farmers. Each farmers market location and stand has collected zip code data for each transaction. This report will go over the steps taken to translate the zip code data to estimates about distances traveled to the participating farm stands and the transactions per zip code.

Study Area/Methodology

The study area is primarily the state of Utah, with markets and farm stands scattered around, but includes any zip code in the United States for transaction data. Some transactions can be found to originate all the way from hawaii to the florida keys. The main goal of this project was to give market runners across the state participating in the double up food bucks program a visual representation of where people are coming from and an estimate on how far people are traveling to get to the fresh produce. The first goal was easy and rather straightforward once zip code areas were figured out, as outlined in the next section.

The second goal of finding statistics on distances traveled to farmers markets posed a different challenge due to the spatial resolution of zip code areas. For example, the zip code “84321” in cache county covers a huge area including the summer homes up Logan Canyon

and a large portion of the national forest. If the zip code center, it would probably land somewhere on the bench of the mountains or in the center of the largest area of the polygon. This would throw off the estimated distances by a large sum, and while not all zip codes are this large and strangely shaped, a good amount are in the mountain west. The population weighted centroids, offered by the Department of Housing and Urban Development, provided a better estimate to where people are living and takes out some of that spatial uncertainty with such large areas. The population centroids are the light red dots on the map and do match very well to population density, following east of main street for much of the valley in this example. With this data, distances can be estimated in a much more confident way.



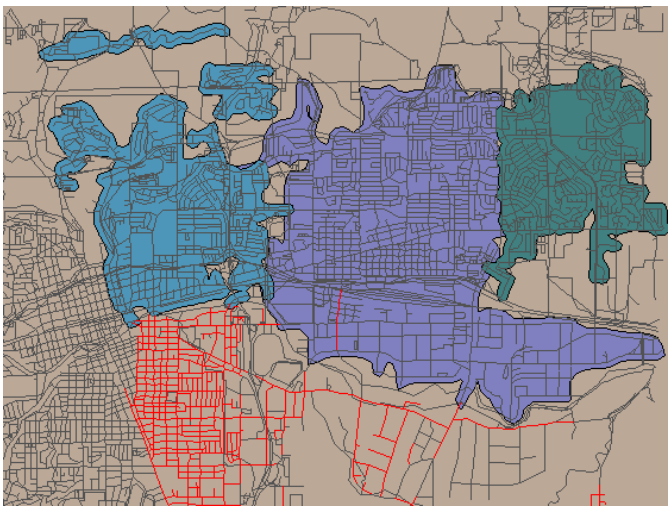
Data report/error

The data was received from the Utah Department of Health and covers the data for the 2021 to 2022 season. This covers 27 different markets and stands across the state participating in the double up food bucks program. Each market collects data in different ways (directly to a computer, self reporting, writing on paper and transcribing later, ect…) and as such have different amounts of error. Most of the unused data is either missing or invalid and was excluded from analysis due to inability to spatially locate it.

Initially it was imagined that we would be able to analyze distance traveled to market by transaction and customer ID, but this proved to be unreliable. While many customers only went once, there are several customers that went 30+ times and reported several different zip codes over that year's time. This may be due to card sharing within a family, or people moving around.

Because there is no one “true” way of determining the location that should be used for each customer, it was decided to go only based on transaction.

Zip code polygons proved to be a challenge when geolocating the zip code data due to uncertainty and misunderstanding of what Zip codes are. It was previously believed that zip codes were areas that could be fully constrained by polygons, but they are actually “routes” and points which can not be reliably boxed around in an exclusive polygon (Grubesic). Despite this complication, zip code polygons are still useful for public facing reports and comprehension, just not analysis. The map on the left is an example of the zip code polygon issue from Reno, Nevada. The red lines are covered by the zip code, but would be impossible to exclusively constraint using a polygon due to size and overlap.



While not a super common issue, the outliers in the zip code locations threw off some of the averages in markets like the Cache Valley Farmers Market. There were transactions associated with zip codes in Hawaii, which can be seen in the large max distances and the unexpected large average travel distance. There is no way to know for sure if one of these zip codes was the result of a mistake, moving, or where food stamps were applied from, so they could not be reasonably excluded.

Analysis

The first step in processing the data was geolocating it for spatial analysis. This was originally going to be done using ESRIs Geocoding service which required credits, but after calculating the number of credits needed to just geolocate zip codes, it was decided that making a geolocator would be easier. This was originally done using the Zip Code polygons (ESRI,

“United States ZIP Code Boundaries”) but was redone using zip code points (ESRI, “USA ZIP Code Points”) due to the error already discussed.

From here the counts were created using the overlapping features tool and recombined with the zip code and population data (found in the zip code layer) using a spatial join. All this was cross referenced using the search by attribute tool in the table to ensure accurate counts. A second geolocator was created using population weighted centroids (Department of Housing and Urban Development) to better represent where people in these zip codes might live, rather than representing location using the geographic center of the zip code area. All checks were redone to ensure the tools ran correctly. The farmers market locations were geocoded using the world geocoding service provided by ESRI and were checked with the project contact for accuracy.

After the locating of the zip codes, the transaction dataset was split by market name using the split by attribute tool and was once again checked by crossref renting the original data table. The near tool was then run to measure the straight line geodesic distance from each transaction to its corresponding market. This was checked using the measure tool for each market. One everything was double and triple checked, the data was exported to excel using the aptly named table to excel tool.

The rest of the analysis was done in excel (finding the mean, median, average, max, count, and standard deviation) and was added to the farmers market location points. All values were rounded to the nearest 1/10th of a mile, but could be rounded further due to the uncertainty associated with the resolution of varied zip code sizes. In a normal application the units would be in metric, but for audience comprehension imperial units work best. The buffer and merge tools were used to check the calculated averages to see if they seemed reasonable, and further geographic representation using the calculated statistics might be done at a later date.

Tools ran

- Create locator
- Geocode addresses
- Cout overlapping features
- Near
- Spatial join
- Table to Excel
- Split by attribute
- Buffer
- Merge

Results

OnlineMap-

https://www.arcgis.com/apps/mapviewer/index.html?webmap=fbecaf539532460aafa518fb8d876c8f&extent=-113.41,39.9284,-110.0372,41.2238

Story map- https://arcg.is/1mCeuT

Paper maps- https://drive.google.com/drive/folders/1nkauWDnOqD1FpEEwwBI7vPLCgLkFDvNj

Table-

https://docs.google.com/spreadsheets/d/1NNk2aoOtAtc6EKFDZN_cMFx1eY2I7p5IOlZvXCQu5_I/edit?usp=sharing

To note, the Nature Hills farm stand did not collect any zip code data.

Works Cited

CDA International Ltd. “Zip Codes Are Not Area Polygons.” *Www.nku.edu*, 9 Feb. 2005, www.nku.edu/~longa/modules/av/lab/zips/zip_codes.html. Accessed 16 Apr. 2024.

Department of Housing and Urban Development. “ZIP Code Population Weighted Centroids.” *Hudgis-Hud.opendata.arcgis.com*, 4 Aug. 2023, hudgis-hud.opendata.arcgis.com/datasets/HUD::zip-code-population-weighted-centroids/about.

ESRI. “USA ZIP Code Boundaries.” *Hub.arcgis.com*, 19 June 2015, hub.arcgis.com/datasets/d6f7ee6129e241cc9b6f75978e47128b_0/about. Accessed 16 Apr. 2024.

---. “USA ZIP Code Points.” *Hub.arcgis.com*, 19 Mar. 2012, hub.arcgis.com/content/1eeaf4bb41314febb990e2e96f7178df/about.

Grubesic, Tony H. “Zip Codes and Spatial Analysis: Problems and Prospects.” *Socio-Economic Planning Sciences*, vol. 42, no. 2, June 2008, pp. 129–49, https://doi.org/10.1016/j.seps.2006.09.001.