Nancy Bowne
URSP 688Y Final Project Narrative
12 May 2025

Problem Statement & Research Motivation

This project's motivation is to measure changes in food distribution in Washington, D.C. and assess improvements in mitigating food deserts. The D.C. Department of Human Services would have hypothetical interest in this project, because it initiates important research on food insecurity and access in D.C. in identifying grocery store access and proximity mapping. This project is motivated by aligning the challenges of food deserts, food waste, and food insecurity, resulting in creating a better food distribution equilibrium. The first research question explores if there is an equal distribution and access between population growth, grocery store types, and food insecurity. The second research question focuses on if the quantity and duration of D.C. grocery stores have changed over time.

Methodology & Approach

For the first research question, I created graphics on available, relevant data from the USDA Food Environment Atlas. I developed graphics to compare changes in quantity and type of grocery store from 2011 and 2016. I compared percentage change over time of the number of grocery stores to the D.C. population and percentage of food insecure population, to detect whether there is a general equilibrium of food access per population size. I created linear regressions to compare the relationships between different types of grocery stores, for instance whether there is any relationship between convenience stores, farmers markets, and general grocery stores. I used regression analysis to detect any relationship between the D.C. population size and D.C. food insecurity rates.

For the second research question, I created a choropleth map to assess grocery store prevalence over time, by using the Open Data DC grocery store locations and DC tracts datasets. I plotted all the grocery store locations from the dataset, resulting in 91 locations. The dataset had years from 1990-2025, skipping a few years in the 90s, ultimately resulting in 19 years of data. The dataset featured a 'present [insert year]' column, that listed a yes, no, or 'blank' to whether that store location existed in that year. I created a counter that would create a new column for each of the yes, no, and blank responses, and to find the total sum of these counts among the 19 years of data. The yes column reflects the total number of years that the particular grocery store location was present. A grocery store location could collect as many as 19 yes responses in the yes count column, if it was open for the 19 years of data collected. The no column reflects how many years that a grocery store location did not exist, and 'blank' is unknown. I created three subplots of the choropleth map for each response. Some challenges included how to account for blank responses, classifications including geoid in index, and general discrepancies between converting data type objects.

Results

The results for the first research question found that there has generally been an increase in population growth, grocery stores from 2011 to 2016, and types of grocery stores (ex. convenience, specialty, supercenter), creating a general equal distribution. Convenience stores increased at a higher percentage of change, at 17.54%, grocery stores increased at 6.8%, and specialty stores increased at 16.98%, potentially suggesting that it is easier to open convenience stores than grocery stores (a common reason for food deserts in urban settings). The data includes five new supercenters and club stores that did not exist in 2011, suggesting larger real estate properties for urban grocery stores. However, club stores are not necessarily equitable for low-income communities, due to membership fees. The values and distribution between population, food insecurity, and types of grocery stores per 1,000 population maintained similar patterns in 2011 and 2016, with food insecurity and high food insecurity decreasing slightly, and other variables increasing.

The regression analysis reveals a positive relationship between farmers markets and convenience stores per population, as well as grocery stores and convenience stores. There was a negative relationship between food insecurity and the population, and grocery store and convenience store per population. While these regressions identify the variables' relationships, a lack of long-term data creates challenges to claiming a strong positive or negative relationship. Despite the numerous variables to explore food insecurity, access, and sales in the USDA atlas, a limitation was a lack of long-term, annual data to track the access of D.C. food distribution. These findings still share some insight on the distribution and equilibrium of food and population in D.C., further remarking on the importance of growth in grocery store options as it relates to population growth.

The grocery store prevalence and proximity choropleth map for the second question resulted in some mixed findings. The map organized findings through three subplots of the yes, no, and blank counts by region. There was a higher rate of grocery stores staying open for more years ('yes') in west D.C., but there were still scattered frequencies of stores that have been open. The no count mainly had a darker prevalence density in the center of D.C. The unknown count color density was also scattered all around D.C., especially helpful to reflect missing gaps in data of grocery store prevalence.

In addition to the choropleth map, the right-skewed histogram visualizes the common number of years for a grocery store to be present, absent, or unknown, demonstrating a higher frequency of values of a region of grocery store having 0-5 years of being present, not present, or unknown. There was a slightly higher quantity of higher values for present. This finding demonstrates general turnover of grocery store locations, perhaps predictable given a market system. These stores are generally replaced over time, but not resulting in a higher absence of stores, or a lack of distribution for the population or in that location. If all 91 grocery stores were open for the selected 19 years, the highest possible value for the yes count would be 1729, for 19 years of data per 91 grocery stores. I separately counted 863 occurrences of the yes response, 463

occurrences of the no response, and 403 occurrences of the unknown response. These values per count reflect the total number of stores that were present within the 19 years of data.

The choropleth map represents areas of a higher frequency of a known open grocery store in 19 years. It may help city officials, like D.C. Human Services, to pinpoint areas that may have longer-lasting grocery stores. While there is weak evidence of higher prevalence of a specific store being open or closed in a specific area, this map illustrates grocery store prevalence and a lack of grocery stores in south and east D.C.

Since the grocery store location data only lists grocery store names, a research limitation is the uncertainty of whether a new grocery store takes over the same location. These store locations represent the past and present operation of stores, and some of the grocery stores are not open in 2025. Missing grocery store data is another limitation, as well as challenges in the choropleth map's count by region, not by the estimating the total number of yes, no, and unknown counts (calculated separately above).

Discussion

Smart city practices shape a role in these results, reflecting the need for proper tracking and publicity of public services, like grocery stores, and user experiences of accessing data. Continuous tracking of these variables may encourage a better understanding of food distribution in D.C., and motivate policies to alleviate food deserts and food insecurity. This data enables grocery stores and local government partnerships to examine food supply and demand in D.C.

Smart city challenges emerge in the estimation of food insecurity, especially with the need for household reporting. While food insecurity seems to be slightly decreasing in D.C., food deserts are still a prevalent issue, especially in handling barriers to city grocery store zoning and fresh produce shipments. This research data encourages the continued use of how to assess food distribution, and track food access in urban settings. An ethical consequence of this research could be the publicity of localized food insecurity data, however, this data is already collected and open to the public. Another concern of this research application is that grocery store types may distort a city's understanding of valid food options, further believing that convenience stores count as the 'same type of food,' further creating food swamps. Better visibility of grocery store prevalence may rebalance the market by correcting shortfalls in grocery store supply and addressing demand by location. A concern is the potential erosion of small business convenience stores and local management. Supercenters and club stores might not be financially accessible for low-income communities.

Future research could investigate specific D.C. tracts and develop more interactive choropleth maps and dashboards to highlight the relationships between key features of food distribution in D.C. Future policy recommendations include promoting food waste redistribution, as well as methods to quantify food waste. Incentives to open new grocery stores include expanding land zoning contracts, tax credits, and better transit access. By mapping grocery store locations and food distribution, cities identify areas of need and long-term changes in neighborhoods' access and proximity to food.