Visualizing Food Accessibility in Massachusetts

Monisha Kapadia Preethi Gandluru Aadya Chaturvedi

Northeastern University

Abstract

The ability to find both accessible and affordable food options is a struggle affecting virtually every person. While Massachusetts has actively tried to prevent food accessibility issues from occurring, residents still face the issue of finding local food retailers that suit their needs. With this in mind, we created an interactive visualization tool made for everyday Massachusetts consumers to determine the location of food retailers based on their location as well as highlighting the number of each establishment type in the area. This tool will enable Massachusetts residents to be better equipped to make smart purchasing decisions when it comes to food and groceries.

1 Introduction

The concept of a food desert can be defined as a region where residents have limited or virtually no cheap and healthy food options. The presence of food deserts can have a multitude of detrimental effects. Food deserts often coincide with low-income and minority communities which exacerbates health issues within these demographics. These affected communities lack the means to seek out alternatives that are perhaps more expensive or are farther away with limited transportation methods. In the place of suitable food retailers, unhealthy and cheap options are relied on such as fast food or convenience stores. It is reported that over 2.8 million people in Massachusetts live in low-income regions where grocery stores are scarce. Additionally, Massachusetts possesses a pervasive wealth gap and gentrification problem that could only further isolate these affected communities as new establishments will be built to accommodate the influx of wealthy residents.

To inform Massachusetts residents of what grocery establishments are accessible and suited to their priorities, we created our visualization tool considering the prevalence of food deserts in Massachusetts. End users will be able to get a visual representation of the geographical locations of food retailers across the state along with the ability to focus on local establishments based on their town. Users will be able to visually perceive the clustering of food establishments or noticeably empty areas. This can aid them in finding new establishments or areas to frequent to purchase food. The visualization will also support informing the target end user to better understand their options by categorizing them by establishment type. They will be able to compare the totals of each establishment type given a specific area. Understanding the distribution of food options will empower Massachusetts residents to better advocate for themselves whether they are concerned for their health, budget, or ease of access of the food they need. These tasks are crucial to support as food is a daily necessity for every person and improving the process of obtaining it can help fight food-related health issues, financial struggles, or even food insecurity.

2 Related works

The issue of food accessibility and food deserts is something of increasing prevalence and has received attention for other visualization tools, like the one we created. One related work that created a visualization tool to bring awareness to this issue would be in the paper "Neighborhood Food Environment and Walkability Predict Obesity in New York City" [1]. The authors visualize the relationship between the neighborhood food environment, walkability, and obesity in New York City. The authors used a variety of visualization techniques to represent the data, including choropleth maps and spatial regression analysis. This visualization tool focuses more on the health impacts of food deserts while ours is more about informing residents and improving the consumer experience. However, these visualization techniques could be useful in informing our visualization of food accessibility based on location. For example, the possible use of choropleth maps to represent the distribution of food retailers or areas with limited access to healthy food options is something that could be helpful in our own tool.

Another related paper in a similar domain would be "Data-Driven Food Desert Metric to Understand Access to Grocery Stores Using Chi-Square Automatic Interaction Detector Decision Tree Analysis" [2]. The authors present a data-driven approach to identify areas with limited access to healthy and affordable food options. The food deserts are visualized using a choropleth map showing the distribution of food desert status across different neighborhoods. This would help to identify areas with the greatest need for interventions to improve food access. This type of map can also be implemented into our visualization of where food retailers are located. However, this tool is powered by statistical analysis while our tool works with simple geographical data for the purpose of informing.

3 Use Case

A specific use case for this visualization tool is to improve the accessibility of food retailers to the everyday consumers that frequent grocery stores, specific to the Massachusetts region. The typical family or even college student needs to purchase groceries as a part of daily living. Individuals on a budget or looking to adjust their diet could utilize this tool as it will outline what type of establishments are in their specific region. For instance, convenience stores may better appeal to certain consumers that are looking for quick access to cheap options while bigger groceries stores have greater appeal to large families. The visualization tool will enlist data on all local and established food retailers in the Massachusetts area.

Users will be able to filter the represented retailers based on which municipality they live in to get a better understanding of their local options. Within the set filters, consumers will be able to gauge based on a tooltip to gather the necessary information to better inform their choice including the store address and establishment type. Additionally, consumers will be able to determine the number of each establishment type in a selected location in the form of a bar graph to gauge if the location in mind has the options that suit their needs or if finding better establishments will require further travel. The tool's purpose is to overall improve the consumer experience of grocery shopping while trying to alleviate the stress of urban food deserts that do exist within Massachusetts by informing the average resident.

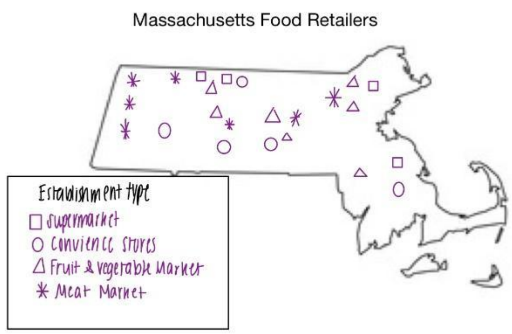
4 Data

The dataset used is a compilation of food retailers in Massachusetts that is from the Metropolitan Area Planning Council's compilation of food system datasets. The data was collected by Tufts University. The original dataset is found at this link https://datacommon.mapc.org/browser/datasets/416. The dataset originally has 9575 items and 16 attributes.

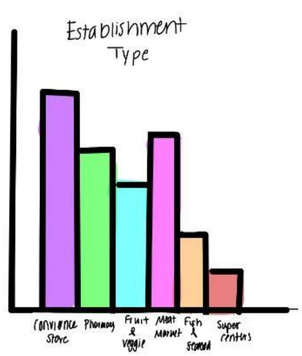
This dataset includes straightforward data regarding the geographical locations of different food retailers along with a basic classification of the retailer type**.** As for evident data collection biases or errors, there is no clear definition as to what qualified as a food retailer to be included in the dataset. Additionally, perhaps the data collection was not robust enough to fully explore all areas of Massachusetts with equal precision. As such some areas of Massachusetts could be well accounted for as others do not have a detailed breakdown of the food establishments in the area. The consequences of this bias in terms of our visualization would be not fully informing residents of a specific town or region in Massachusetts of their full options.

During the data exploration process, it was discovered that this dataset includes attributes that are not necessary for our visual encoding. For instance, the state column was the same for each entry as each retailer is located in Massachusetts, so the column was rendered useless. Additionally, many columns had ids and codes that did not have any clear meaning or significance to our visual encoding. To clean this dataset, the record type, state, year, source, farmers market id, EBT, and coupons were removed. The data in this dataset was nominal so explorations did not depict any outliers or other statistical anomalies. Lastly, the dataset had no null values which concluded our data exploration and processing.

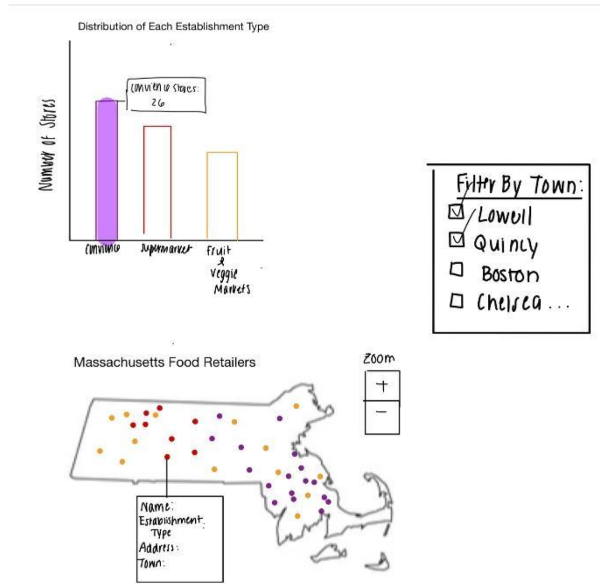
5 Design Process



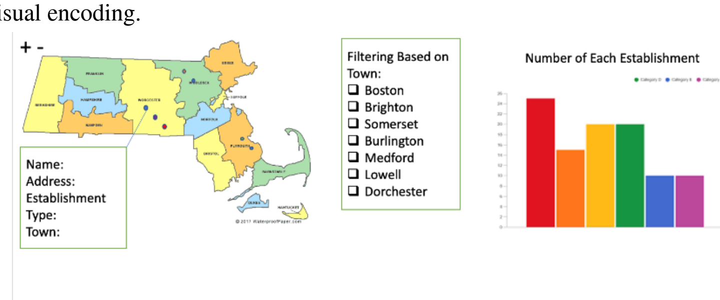
The first initial sketch of the visualization focused on representing the spatial data given by each food retailer in the dataset. Our original intention was to utilize points as marks and the channels of shapes to differentiate each location by establishment type.



The next partial design was to display the distribution of establishment type of each food retailer in the form of a bar graph with distinct colors differentiating the bars.



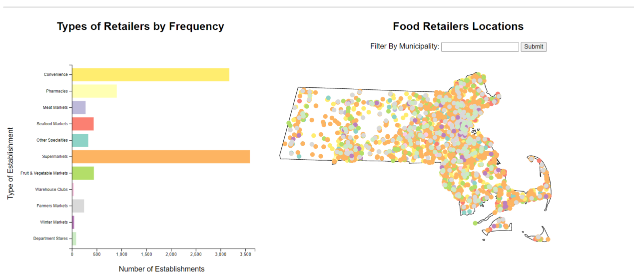
For a rough sketch of how to link the two visual encodings we had designed, we decided to implement a consistent use of color as a channel. The color of the dot on the map would be indicative of the color of the establishment type on the bar graph. To better suit the local aspect of the tool for end users, we envisioned filtering by the municipality that would control the outputs of each visual encoding.



The final sketch of our visualization compromises the two initial visualizations. The aggregation will be required to tally the total of each establishment type in the dataset for the bar graph. Each visual encoding will have a tooltip. On the map, a tooltip with the name of the retailer, the address, the establishment type, and the municipality where the retailer is located will appear. In the bar graph, a tooltip will be utilized to display the total number when hovered. The graphs will be linked by hovering capabilities on the map which will cause the highlighting of the bar represents the establishment type that the point on the map belongs too. Filtering will still occur with a checkbox that will limit or expand the view of each visual encoding based on user selection. The map will also have zooming capabilities as we expect there to be a significant number of marks representing each location. The marks in the map encoding will be points and utilize channels of position (both vertical and horizontal) and color with a categorical colormap. The bar graph utilizes marks of shapes with each bar and the channels of position (horizontal – indicating a different category of each establishment type), size (the height of the shape indicates the value of each bar), and the color with same categorical colormap to better keep the encodings consistent.

Throughout the implementation process, the design evolved to reach its final design based on usability testing and programming limitations. For instance, when testing the filtering abilities based on location in the visualization, users had a difficult time navigating through the long list of checkboxes as the number of municipalities was extensive. Additionally, there was obvious lag time trying to aggregate the data on multiple municipalities that were experienced in testing as well. These results encouraged us to implement a dropdown instead that could concisely list the many municipalities and select one at a time to aggregate. Additionally, without an explicit legend for the map, user testing also demonstrated that it was not overtly clear that the colormaps for the encodings were linked. As such, we decided to implement another linking method that would allow borders for a point on the map to appear when clicked and this border would also appear on the corresponding bar. We found using clicking rather than hovering made the visualization clearer and caused less of an overwhelming experience for the users when pursuing the map. For purely aesthetic and cohesion purposes, the bar chart was made horizontally.

6 Final Design



The final design of the visualization tool includes two visual encodings. On the left, a bar chart is implemented that totals the number of each establishment type. The x-axis represents the number of establishments while the y-axis labels the type of establishment type. There is a tooltip that conveys each bar's establishment type label and the exact number represented. The right visualization plots the geographical location of each food retailer on the map of Massachusetts using the data's longitude and latitude. The map has a tooltip that conveys the name of the store, the address, the establishment type, and the municipality the store is located in. The map also has zooming capabilities in and out to get a more specific view of the points in smaller regions. colormap utilized in both encodings is categorical based on the establishment type of the data depicted. The visualizations are linked in two methods. The dropdown above the map lists all the municipalities in the dataset. When the dropdown selection is submitted, the locations belonging to that municipality are plotted on the map and used to aggregate new totals in the bar chart. The second linking component allows a specific point on the map to be bordered and the corresponding bar belonging to the same establishment type to also become bordered.

The intended end user would be able to accommodate both domain tasks with the features of the visualization tool. To determine their local food retailers, the users can manually explore and zoom in on the retailer map based on where they are in Massachusetts. Additionally, the user can also use the filtering dropdown by selecting the municipality they live in to see the specific food retailer options near them. Utilizing the tooltip on the map, the user will be able to learn more about the location of interest to them. For the next domain task, the user will be able to see how the municipality they live in serves their specific grocery and food needs accordingly with the bar chart. Once again, the filtering dropdown would allow updated aggregation of the bar chart for the user to be more informed of how a certain municipality fails or meets its needs with the distribution of establishment types. Additionally, the point selection linking enhances the user's experience for this specific domain task to see if a certain location on the map is one of many of the same establishment types.

7 Discussion

The final implementation of our visualization tool, while addressing the majority of its intended domain tasks and its use case, there is room for improvements and future work. In its current state, the tool can aggregate retailer types and plot the retail locations on the map of Massachusetts with effective filtering. Addressing our ideal use case, users will be able to better understand their local options and determine if they suit their needs or not. However, the map of Massachusetts utilized lacks the detail and geographical information that a typical map interface would have that users are used to with common navigation tools. Using a more detailed map of Massachusetts with town lines and satellite imagery would be a better way to inform users of where these retailers are located relative to the users' specific geographic location. Additionally, while the use of municipality dropdown does effectively filter the locations that are plotted on the map, the points are clustered together and hard to view without utilizing the zoom feature. Ideally, filtering should also cause a zoom-in on the specific municipality on the map for a better user experience.

For future work, we would update to more recent data as this dataset was from 2016 and pre-COVID so many old retailers could close while new ones may have emerged. This would be to ensure users are better informed about their current options. Additionally, we would tweak the tool by enabling users to enter their specific location and determine the retailers closest to them using a certain mile radius. This would also enhance the user experience to get more personalized results. A possible extension for this tool is to allow for this visualization to be based on data from multiple states in the U.S. or perhaps the entire country to expand the intended users and apply the exploration of food deserts to more areas in the U.S.

8 Conclusion

We have created an interactive visualization tool in hopes of better informing Massachusetts consumers about their food purchasing options and their ability to advocate for their needs. The tool works off a dataset of geographical locations of food retailers in Massachusetts. One visual encoding consisted of a map that plotted each location while the other aggregated the data to create a bar graph based on the location's establishment type. The tool could lead to valuable insights into the presence of food deserts or a lack of affordable and nutritious options in specific regions of the state. These insights could be applied in many different domains: consumer experience, retail expansions, or even political advocacy. Food deserts will only become exacerbated with moving trends toward larger wealth gaps and gentrification, especially in states with large cities. Being informed is the first step for Massachusetts residents to combat the pervasive issue which is hopefully what our visualization tool contributes.

References

[1] Rundle, A., Neckerman, K. M., Freeman, L., Lovasi, G. S., Purciel, M., Quinn, J., Richards, C., Sircar, N., & Weiss, C. (2009). Neighborhood Food Environment and Walkability Predict Obesity in New York City. Environmental Health Perspectives, 117(3), 442–447. https://doi.org/10.1289/ehp.11590

[2] Chavis, C., & Bhuyan, I. A. (2022). Data-Driven Food Desert Metric to Understand Access to Grocery Stores Using Chi-Square Automatic Interaction Detector Decision Tree Analysis. Transportation Research Record: Journal of the Transportation Research Board, 036119812210893. https://doi.org/10.1177/03611981221089308

Appendix A: Data Abstraction

In the dataset, each row represents an item that indicates a discrete food retailer located in Massachusetts. There are several attributes that describe each item. The first is the name of the retailer, represented as categorical data. The second attribute is the address of the retailer, which includes the street and number where the store is located, also represented as a categorical attribute. The third attribute is the name of the municipality where the retailer is located and is a categorical attribute. The fourth attribute is the zip code associated with the retailer's location, which is a discrete quantitative attribute. The last attribute is the primary type of store that the retailer represents. This information is represented as a categorical variable, indicating the type of store. The fifth and sixth columns are positions that represent the latitude and longitude values of the retailer's location, respectively. Both values are represented as continuous quantitative attributes.

Appendix B: Task Abstraction

Using Tamara Munzner's visualization task taxonomy, the domain tasks of our visualization can be properly abstracted into the following visualization tasks:

- Domain Task 1: Determine the food retailers local to users based on the municipality they live in and convey specific details about these locations.
    - High-Level: Discover (Consume) – users are discovering new information about the food retailers in their vicinity
    - Medium-Level: Browse – users will know what locations to focus on but will be exploring new establishments with no specific target in mind.
    - Low-Level: Summarize – the users will get an overview of the retailers in each location
- Domain Task 2: Compare the totals of each establishment type to see how a region is equipped to fit differing consumer needs.
    - High-Level: Discover (Consume) – the user will be able to find new information on the frequency of each establishment type in each region
    - Medium-Level: Explore – the user does not have a specific target or location to look for when analyzing the totals of each retailer type.
    - Low-Level: Compare – the user will be able to take in multiple targets (establishment types) to see how their totals differ