Translating Community Survey Data: A Pilot Study of Food Access Near Somerville Schools

Cartographer: Alice Tin
UEP 232: Introduction to GIS
December 10, 2009

Introduction

The Centers for Disease Control and Prevention estimate that over 9 million (16%) children ages 6-19 are overweight or obese; this figure has tripled since 1980. An additional 15% were considered at risk of becoming overweight.[1] Adolescence is a critical period in shaping eating habits and weight status later in life, which underlines the importance of helping children form healthy eating habits at this time. Schools are pivotal in this part of their education, as the children spend a large amount of time there.

Previous studies have shown that the built environment can influence energy intake based on food availability, which in turn demonstrates an association between higher convenience density with a higher probability of being overweight.[2] Therefore Shape Up Somerville, a city-wide campaign to promote physical activity and healthy eating, designed a survey to gather more information about food access within a walking distance of schools. Due to some constraints, this project is a pilot to demonstrate the ways that the survey data could be transcribed onto a computer database and analyzed and displayed through GIS. The ultimate goal of the project was to gain a better understanding of Somerville's food system and children's access to junk food, which will in turn lend insight to childhood obesity and may inform future actions in regulating types of food that can be sold around schools.

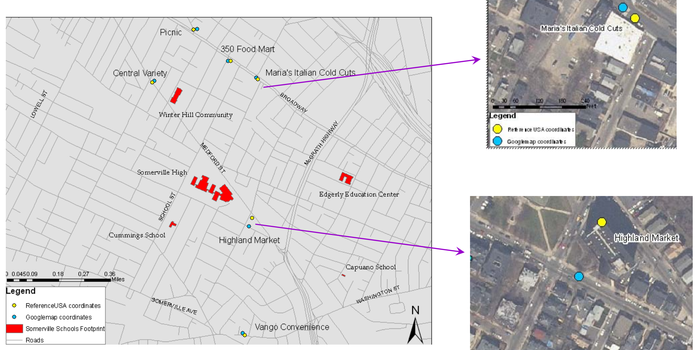
Methodology

The Shape Up Somerville survey was conducted in schools located within a "walking distance," which was defined as a 1/4 mile radius, of Somerville schools. We recorded the presence and price of the items listed on the survey. The categories of food items (health grocery, junk food, etc) were identified by Shape Up Somerville, and sit-down restaurants and fast food establishments were excluded from this survey at their request. The survey was not formatted in a way that would facilitate easy data entry; therefore I coded the survey responses numerically and created a database in Microsoft Access. Considering that this project was limited to one semester and including time allowance for data collection and data entry, I was not able to enter all the collected data. I simplified the coding for some parts of the survey and focused on recording health grocery and junk food items in a detailed manner.

Analysis

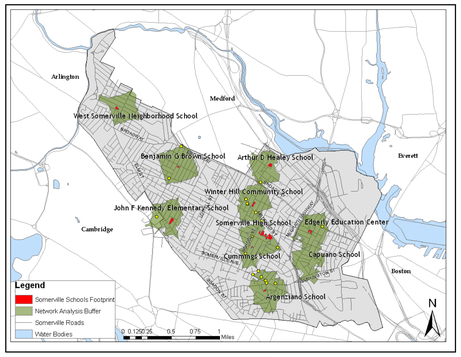
Geocoding store locations

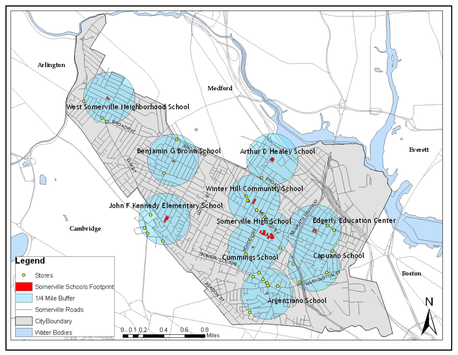
One of the Shape Up Somerville interns had compiled a database of food establishments in Somerville. She had used Google maps to obtain the longitude and latitude information for each store based on the given address. However, I realized that the coordinate data could also be obtained from Reference USA, which is a database of 14 million U.S. businesses containing verified data, updated monthly. Access was obtained through the Tisch Library at Tufts University. I geocoded these points and compared them to those that were provided by Google maps and orthophotos. I acknowledge that both datasets are estimates, and the most accurate data would be obtained by visiting the actual stores and collecting coordinate data using a hand-held GPS unit.

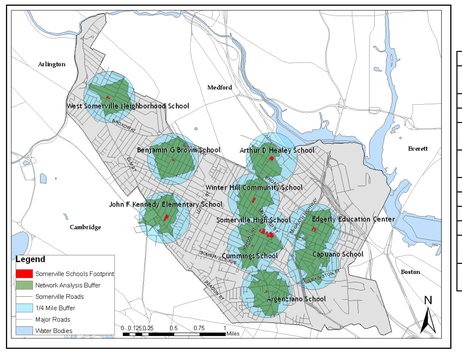


Creating the ¼ mile buffer

As discussed before, "walking access" was defined as a ¼ mile radius around schools. ArcMap was first used to create buffers that were drawn by simply extending a ¼ mile radius from the data points of Somerville schools. However, this project is considering walking access to stores, and it would be more accurate to consider a ¼ mile buffer comprised of streets. Therefore I also used the Service Area function in ArcMap's Network Analysis to create such a buffer. The difference in area is shown in the table below. This comparison also highlights the distinction of the number of stores that would be taken into consideration.







School name	Buffer area (sq mi) "As the crow flies"	Buffer area (sq mi) Network Analysis Service Area
Argenziano School	1.57	0.112916
Arthur D Healey School	1.57	0.079418
Benjamin G Brown Elementary School	1.57	0.115689
Edgerly Early Education Center	1.57	0.16191
Capuano School	1.57	0.08836
Cummings School	1.57	0.065938
John F Kennedy School	1.57	0.09306
Somerville High School	1.57	0.097701
West Somerville Neighborhood School	1.57	0.086643
Winter Hill Community School	1.57	0.102484

Proximity Analysis

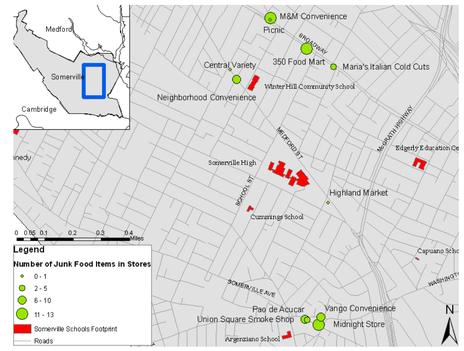
The point distance tool was used to calculate the distance between the schools and stores within a ¼ mile radius based both on the network analysis buffer and the coordinates from Google map.

School Name	Store Name	Distance (mile)
Argenziano School	Capone Foods	0.14
	Casa de Carnes	0.21
	Bombay Market	0.21
	La Internacional Foods Corps	0.07
	Reliable Market	0.10
	Vango Convenience	0.13
Arthur D Healey School	M&M Convenience	0.21
Benjamin G Brown Elementary School	White Hen Pantry	0.20
	Y-Not Variety	0.16
Edgerly Education Center	Minute Market	0.13
Capuano School		
Cummings School	Lil Peach Market	0.11
	Summer Convenience & Pizzeria	0.10

John F Kennedy School	Shaws Porter Square	0.14
Somerville High School	Thurston Spa	0.23
	Lil Peach Market	0.11
	Highland Market	0.22
	Mobil Mart	0.14
West Somerville Neighborhood School	/	
Winter Hill Community School	Thurston Spa	0.10
	M&M Convenience	0.23
	Corner Market	0.06
	Mobil Mart	0.21
	Central Variety	0.08
	Hispana American	0.06
Total number of stores	19	

Visual Representation of Stores with Healthy Food or Junk Food

Since the survey data was coded with numerical values, I was able to select and aggregate parts of the survey to display how many junk food items are available at each store and represent it with a graduated symbol.



Challenges

The main challenges of this project were related to data quality and data entry. The survey was not structured for data entry and therefore a lot of effort was put into the creation of a database that would also be formatted to be compatible with ArcMap. Additionally, the data collected varied in quality. Multiple surveys were un-useable because they did not have basic identifying information, such as the store name. A few stores refused to participate in the survey. Therefore, the survey is not a comprehensive list of stores within the ¼ mile buffer. Multiple surveys were conducted at the same store, which resulted in a small sample size. Additionally, most surveys were not filled out completely. Many fields were left blank, which could either be interpreted as an absence of said item, or that the information could not be found. In this project, I assumed the latter. These deficiencies in data quality could be rectified if project would benefit from the provision of training for surveyors on a standardized format for data collection.

Other challenges in data quality included the validity of established data layers, such as the Somerville schools layer, which I acquired from the City of Somerville. The data was not up-to-date. For example, the East Somerville school was burned down in 2008, and this change was not reflected in the data layer, and as noted in an earlier section, the currency of the data is unknown. However, I was able to correct these errors using the Editor function in ArcMap.

Future Direction

Based on the coding that I did for these two categories of foods, someone who will be working on this project in the future will be able to codify the rest of the survey. As discussed before, data quality could be improved. When this data set is complete, Shape Up Somerville has expressed interest in formatting the map so that it can be an interactive feature on a website. For example, users would be able to click on a store on the map and information about the types and prices of foods sold would be displayed.

As suggested by the literature referenced, further analyses with regards to advertisements displayed at the stores, evaluation of the sidewalks or recreational facilities could be pursued. These factors, such as effect of advertising on consumption behavior and level of physical activity are also contributing factors to childhood obesity. An analysis of these contributing elements would further our understanding of the childhood obesity phenomenon and fuel intervention strategies.

It is also important to consider that children spend a significant amount of time in other places outside of school. A separate survey of the school children could reveal these other locales (for example, an after-school activity center or local park that serves as a gathering spot). Junk food access close to these other locations should be included in the analysis. Shape Up Somerville and Cambridge Health Alliance staff have also mentioned mobile sources of junk food, such as ice cream trucks or mobile fast food stands, which are sources of unhealthy food access, which present a unique challenge for representing these elements on a map.

The analysis could also be expanded to the residential population in these areas. Since these stores would be the closest source of food, residents may also be obtaining their food from these same locales. I know that other students in the class are working on analyzing food access in Somerville, and this would be an opportunity to share data layers.

Data Sources: MassGIS, City of Somerville, Original Survey Designed by Shape Up Somerville

Map Projection: StatePlane Massachusetts Mainland FIPS 2001, NAD 1983

Works Cited:

[1] Centers for Disease Control. (2004) Prevalence of Overweight and Obesity Among Children and Adolescents: United States, 1999-2002. Accessed on December 11, 2009 from http://www.childrenshospitals.net/AM/Template.cfm?Section=Site_Map3&CONTENTID=47972&TEMPLATE=/CM/ContentDisplay.cfm

[2] Grafova, I. Overweight children: Assessing the contribution of the built environment. *Preventive Medicine.* 2008. 47:3 pp304-308