Access to low-cost food in the Washington, DC region

The problem – disparate access to healthy low-cost food

> The presence of supermarkets/grocery stores in a neighbourhood is associated with buying and consuming healthier food. Grocery stores typically sell healthier food items at affordable prices compared with convenience stores and fast-food outlets. Wide ranges of fresh food choices, with relatively lower price points, frequent availability and visually appealing presentation, are associated with choosing healthier food options and healthier diets.

Abeykoon, Stringer and Muhajarine, “Health-related outcomes of new grocery store interventions: a systematic review”

Can Foursquare data shed light on the issue?

The question at hand is whether higher-income areas have a comparatively higher preponderance of supermarkets and a comparatively lower preponderance of lower-quality food vendors than lower-income areas.

Project elements:

(i) Defining geographic limits – grid search over the Washington, DC area.

(ii) Extracting and cleaning Foursquare venue data – extracted supermarkets, convenience stores, fast food restaurants and dollar stores.

(iii) Geographic groupings – k-means clustering based on distance (approach needs improvement).

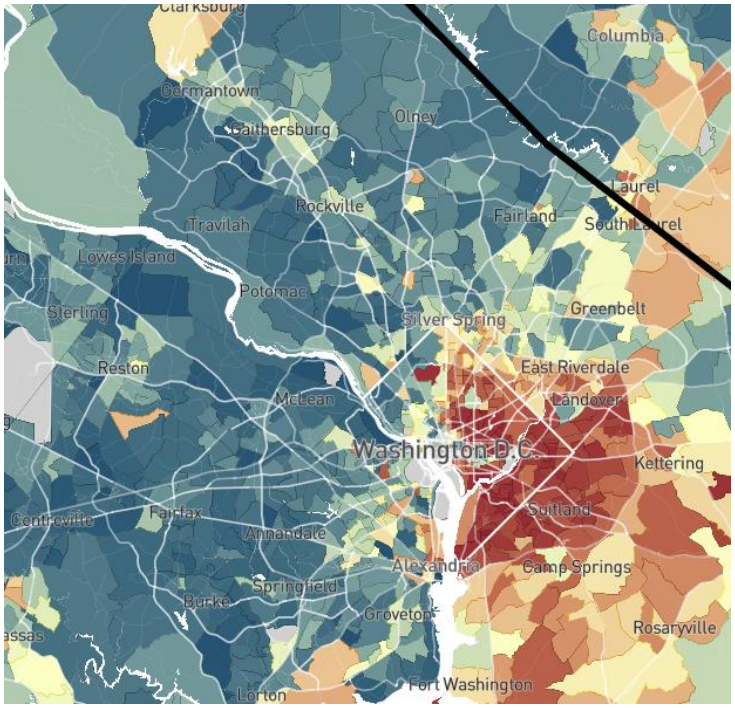
(iv) Food availability clustering – k-means clustering of the step (iii) geographic zones, based on the relative percentages of supermarkets, convenience stores, fast food restaurants and dollar stores within each zone.

Income distribution in the Washington, DC area

Source: The Opportunity Atlas (https://www.opportunityatlas.org/)

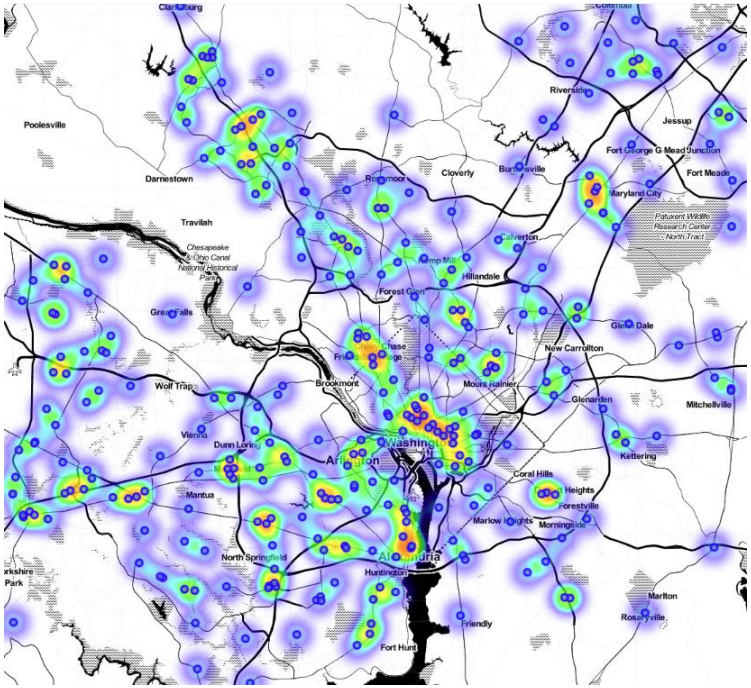
Higher income – blue-green

Lower income – red-maroon

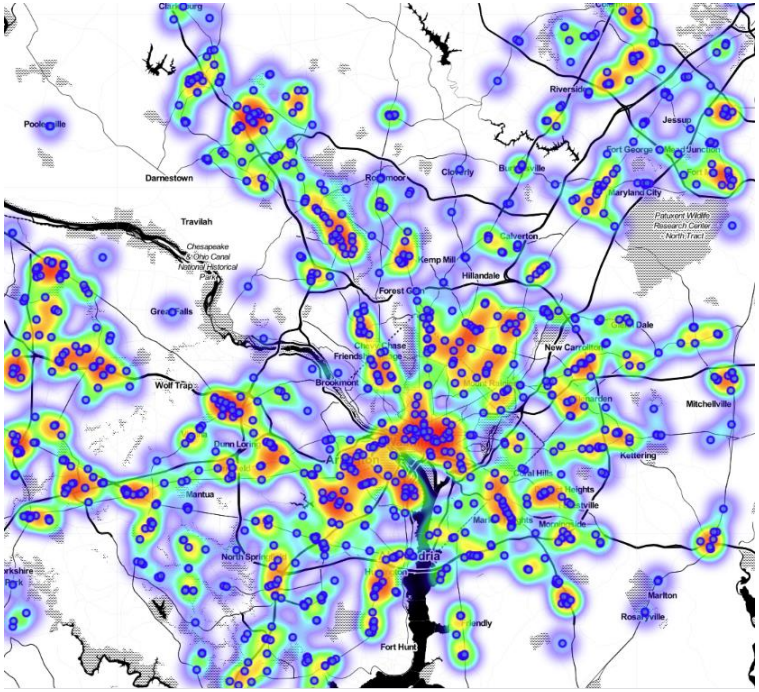


Initial results – patterns, but nothing conclusive

Supermarkets

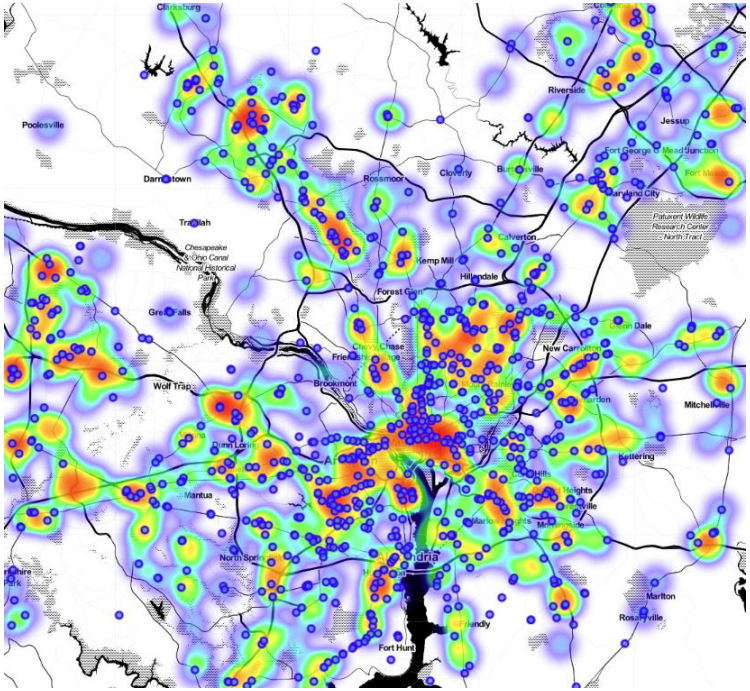


Fast food

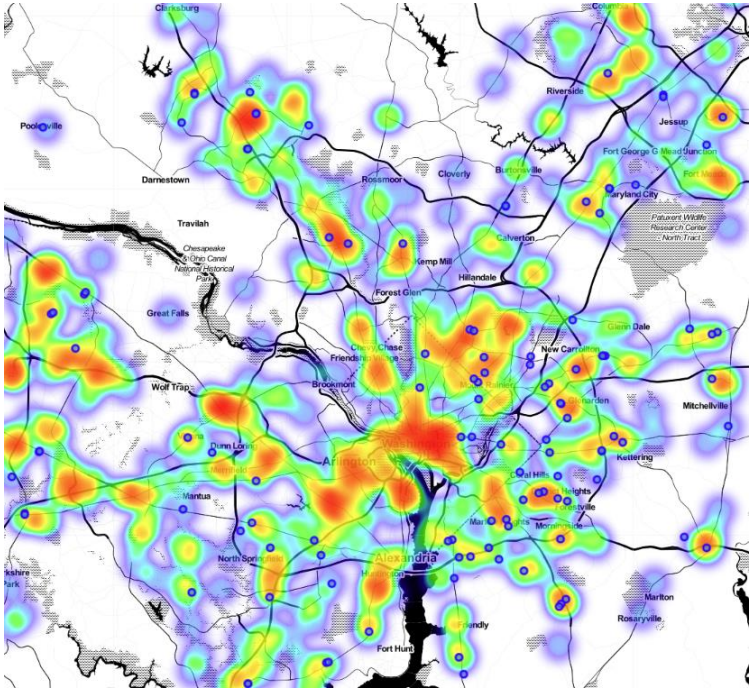


Initial results (cont'd)

Convenience stores



Dollar stores



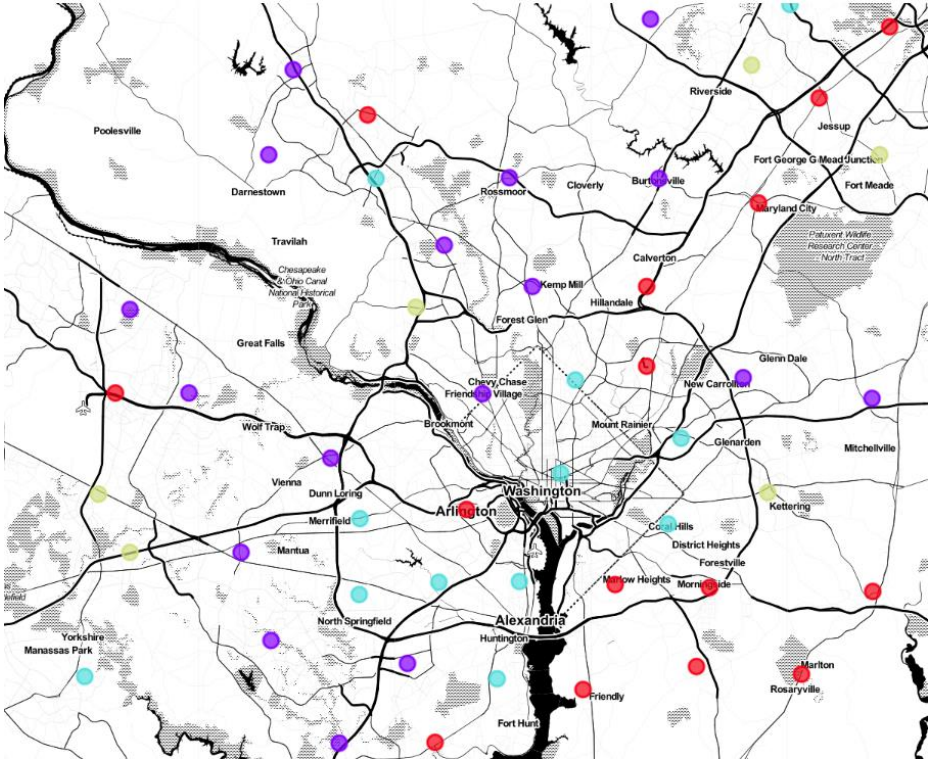
Clustering of geographic zones

Cluster 0 – red
Cluster 1 – purple
Cluster 2 – cyan
Cluster 3 – light green

Clusters 0 and 2 tend to correspond with lower-income areas

Clusters 1 and 3 tend to correspond with higher-income areas

There are significant exceptions in both directions



Cluster differences

Table 1 – Food availability clusters

	Cluster 0		Cluster 1		Cluster 2		Cluster 3	
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Mean	Std dev
Convenience	0.425	0.030	0.309	0.034	0.485	0.041	0.197	0.026
Dollar	0.054	0.043	0.044	0.038	0.041	0.041	0.049	0.042
Fast Food	0.441	0.029	0.503	0.030	0.350	0.030	0.602	0.043
Supermarket	0.080	0.028	0.144	0.039	0.124	0.049	0.152	0.053

The proportion of supermarkets in Cluster 0 has a statistically different mean from the other three clusters.

But we cannot reject the hypothesis that the proportion of supermarkets across Clusters 1, 2 and/or 3 have the same mean.

Bottom line

Foursquare venue data may serve as a useful tool for examining the availability of low-cost food vendors among various local areas in the Washington, DC region.

Potential enhancements:

- developing refined geographic groupings that are more precisely linked to locality data regarding income (as well as other relevant criteria, including health-related criteria)
- accounting for the fact that the distribution of fast food restaurants and convenience stores reflects the distribution of offices and tourist areas as well as the distribution of residences
- if practicable, accounting for other types of food vendors.