Homework 4: The Spatial Distribution of Farmers Markets in Philadelphia

CPLN 671/MUSA 500

"In some Philadelphia neighborhoods, it is easier to find orange drinks than an actual orange."

- Get Healthy Philly website

Ensuring access to healthy, locally grown food has been a challenge for many American cities. As one of the ways to get fresh and healthy food to residents of the different Philadelphia neighborhoods, the Philadelphia Food Trust has set up numerous farmers markets throughout the city. These markets have been shown to offer numerous benefits to consumers who shop there and to their communities, and Wikipedia (http://en.wikipedia.org/wiki/Farmers'_market) lists a few of them:

1. *Reduced overhead: driving, parking, etc.*
2. *Fresher foods*
3. *Seasonal foods*
4. *Healthier foods*
5. *A better variety of foods, e.g.: organic foods, pasture-raised meats, free-range eggs and poultry, handmade farmstead cheeses, heirloom produce heritage breeds of meat and many less transport-immune cultivars disfavored by large grocers*
6. *A place to meet neighbors, chat, etc.*
7. *A place to enjoy an outdoor walk while getting needed groceries*

That said, not all neighborhoods in Philadelphia have access to farmers markets; in fact, parts of South Philadelphia and North Philadelphia, and essentially all of Northeast Philadelphia have no farmers markets at all, depriving people who live there and their communities of the numerous benefits listed above.

This assignment asks you to examine the spatial distribution of farmers markets in Philadelphia, and see whether markets are randomly placed, dispersed or clustered throughout the city. It makes use of a shapefile containing locations of all farmers markets in Philadelphia for the year of 2013, obtained from the PA Spatial Data Access website (www.pasda.psu.edu). This shapefile, as well as a (zip code level) shapefile of Philadelphia, is available in the course folder.

You will be asked to carry out several types of point pattern analyses that were discussed in class using ArcGIS and Excel, and determine whether the farmers markets are randomly placed, clustered or dispersed. These include Nearest Neighbor Analysis and K-Function Analysis with the following settings in ArcGIS:

a. Use *10* as the 'Number of Distance Bands'
b. Select *99 Permutations* under 'Compute Confidence Envelope'
c. Check 'Display Results Graphically'
d. Leave the 'Weight Field' blank
e. Put in *0* for 'Beginning Distance'
f. Put in *2500* feet for 'Distance Increment'
g. Select *Simulate_Outer_Boundary_Values* for 'Boundary Correction Method'
h. Select *User_Provided_Study_Area_Feature_Class* under 'Study Area Method'
i. Select the *Philadelphia* shapefile under 'Study Area Feature Class')

* For extra credit, you may use R in addition to ArcGIS. If you do, please present your R code in an appendix.

Your deliverable should be a report that follows the outline below.

A. 1 paragraph introduction to the problem, which describes the importance of the issue, and briefly states what you will be doing in this report.
B. 1-3 page description of the methods:
    a. What hypotheses you will be testing
        i. What is CSR? Describe the conditions that need to hold in order for the point pattern to be completely spatially random.
        ii. What are your null/alternative hypotheses in point pattern analyses?
    b. Even though you will not be using the Quadrat method in this assignment, talk about it:
        i. What does this method entail? That is, describe the method and be sure to write about it in your own words.
        ii. Discuss the limitations of the Quadrat method and why it's generally not used in practice.
    c. Talk about the Nearest Neighbor Analysis method:
        i. What does this method entail? That is, describe the method and be sure to write about it in your own words.
        ii. What statistical test will you use here? What is the test statistic? That is, how do you determine whether you have significant clustering or dispersion (uniformity)? Include the relevant formulas and be sure to describe them and all the terms that they include.

      iii.   What are the limitations of the Nearest Neighbor Analysis? In particular, compare the problem at hand with the hospital example in the slides.

  d. Talk about the K-Functions Analysis method:

      i.   What does the method entail? That is, describe the method and be sure to write about it in your own words.

        1. Be sure to include what K(d) and L(d) functions are in this description, and how they are defined in the software package that you use.

      ii. Talk about the beginning and incremental distances, and specify how they are calculated.

      iii. Describe the testing procedure and the concept of confidence envelopes.

      iv. What happens to points next to the border? Talk about Ripley's Edge Correction and the Simulate Outer Boundary Values Edge Correction (and specify which one you use and why).

      v. Talk about the nonhomogeneous K-Functions – that is, situations when you need to take into account a reference measure such as population in your K-function analysis.

        1. Briefly describe when/where this would be appropriate

        2. In your own words, describe how this may be done

          a. Hint: this is what is done in the last 25 or so slides.

C. Results

  a. Present results of the Nearest Neighbor Analysis.

      i.   Describe the results. Do you reject the null hypothesis?

      ii. Re-run the analysis using the area of Philadelphia and not the minimum enclosing rectangle. Describe the results. Do you reject the null hypothesis?

  b. Present results of the K-function Analysis.

      i.   Describe the results, and be sure to specify what the beginning and incremental distances were, and why you set them at those levels.

        1. Do you reject the null hypothesis? At what distances?

      ii. Without actually doing the analyses, do you expect that the reason for not seeing any farmers markets in Northeast Philly and certain parts of North and South Philly is that the population in that part of the city is too low? That is, do you expect that you would get different results if you were to take population (say, at the zip code level) into consideration?

D. Discussion

  a. Are the results obtained with the Nearest Neighbor Analysis and K-function Analysis consistent with each other?

b. Are they consistent with your expectations, based on the visual examination of the point data, and given the limitations of each method?

c. * Are the results from R and ArcGIS consistent with each other?

d. Does it seem to be the case that areas with lower median incomes have fewer farmers markets? Without doing any statistical tests, present a map of the farmers markets overlaid on a map of median household income at the zip code level and discuss. The *Philadelphia_ZipCodes* shapefile has the MedIncome variable which is the median household income for the year 2000.

e. What conclusions can you make based on these findings? That is, can you conclude that farmers markets are clustered in Philly? Does it make sense to use the minimum enclosing rectangle as the study area for this problem?

f. What might the policy implications of these findings be?