WIC Project in New York State

GIS Time Estimate

(Version 0.3) - 2015.04.09

Columbia University
Built Environment & Health Project
www.beh.columbia.edu

Prepared By:
Daniel M. Sheehan - Geographer, GIS Analyst
The Built Environment & Health Project Columbia University
Mailman School of Public Health - Department of Epidemiology
722 West 168th St. – R735
New York, NY 10032
e: dms2203 (at) columbia (dot) edu

Prepared For:
Dr. Sally Findley
Mailman School of Public Health
60 Haven Avenue, B2
New York, New York 10032
e: sef5 (at) columbia (dot) edu

Purpose

The purpose of this document is to list the GIS data resources and time estimate for the project; **Profiles of participation in WIC and other healthy living programs for preschoolers in New York.** The project will involve a WIC dataset (Women, Infants, and Children) provided by the investigators in the project study area of New York State (NYS). For the remainder of this GIS Data Resources Document the project will be referred to in this document as **WICNYS**.

The Co-Principal Investigators are Sally E. Findley, Jackson Sekhobo and Mary Ann Chiasson. This group of investigators will be referred to as the **WICNYS Group**. The Built Environment and Health Project (**BEH**) Geographic Information Systems group (**BEH-GIS**) constructed this document to begin a conversation about data resources for this WICNYS project. Some of these resources have not been used by BEH-GIS. There may be some additional data cleaning (data munging & carpentry) or pivots to alternative data resources. Some of these may be lead to additional conversation and consultation with the WICNYS Group . All **RED TEXT** signifies an important note in the document.

Tasks

This project will be broken down into several tasks that aid in organizing the GIS processing tasks. Each task (and for some subtasks) there will be a minimum and a maximum time estimate. The tightness of the time estimate minimum and maximum values is reliant upon previous experience with the particular task and data involved and, in the case of geocoding, a range that allows for some level of dirtiness in the initial address data of the study subjects.

1. **Geocoding**
2. **Neighborhood level variable creation**
3. **Data wrangling and QA/QC**
4. **Data Dictionary**
5. **Data Delivery and Mapping (optional)**

1. Geocoding

Note on Geocoding:

BEH-GIS geocoding estimates are ~3-5 minutes per address record geocode depending on the data quality. This estimate can be greatly reduced with very high address data quality. ZIP level geocoding has much lower estimate as it is more akin to a table join (or merge i.e. Stata) that is dependent on quality and usually take < 8 hours rectifying bad ZIPs so long as that number of initially unmatched ZIPs is < ~100.

A first pass in geocoding is running the data as-is through a geocoder (BEH maintains several secure/local geocoding services). This will determine how many records will require the additional 3-5 minutes per record estimate. It is difficult to say how many will successfully geocode in the first attempt. A very clean address dataset will generally geocode at 90-95% success. A dirty address table 80-85%. Anything below that will require a serious cleaning effort that does not scale very well and is time consuming OR it is a sign that the geocoder needs to be rebuilt.

ZCTA vs ZIP CODE

ZCTA (ZIP Code Tabulation Areas) and ZIP Codes have differing geographies. Quite simply, ZCTA's are "generalized areal representations of United States Postal Service (USPS) ZIP Code service areas." [1] These generalizations provide for nesting and topological relationships with existing Census geographies.

ZCTA's have Census/American Community Survey (ACS)-related variables so it will be determined whether this project will use:

1) **ZIP code to ZCTA crosswalk** or
2) **will use ZIP code polygons** (with some ZIP nesting rules for small ZIPs or point-level ZIPs) **and areal-weighting interpolation from data at the Census Tract geographic unit**.

At the moment BEH recommends choosing method 2) **ZIP code polygons** (with some ZIP nesting rules for small ZIPs or point-level ZIPs) **and areal-weighting interpolation from data at the Census Tract geographic unit.**

The following text is from [https://www.census.gov/geo/reference/zctas.html]:

> The Census Bureau first examined all of the addresses within each census block to define the list of ZIP Codes by block. Next, the most frequently occurring ZIP Code within each block was assigned to the entire census block as a preliminary ZCTA code. After all of the census blocks with addresses were assigned a preliminary ZCTA code, blocks were aggregated by code to create larger areas.
>
> The Census Bureau assigned blocks that contained addresses, but did not have a single most frequently occurring ZIP Code to the ZCTA with which the blocks had the longest shared boundary.
>
> If the area of an unassigned enclave was less than two square miles, it was assigned to the surrounding ZCTA. The Census Bureau used block group boundaries to identify and group unassigned blocks. These unassigned blocks were merged into an adjacent ZCTA based on the length of shared boundary.
>
> For the Census 2000 ZCTAs the Census Bureau created ZCTAs that ended in "XX" to represent large areas of land without ZIP Codes or "HH" to represent large areas of water without ZIP Codes. For the 2010 Census, large water bodies and large unpopulated land areas do not have ZCTAs.

---

1 ZIP Code™ Tabulation Areas (ZCTAs™) [https://www.census.gov/geo/reference/zctas.html]

ZCTAs were created using residential and nonresidential ZIP Codes that are available in the Census Bureau's MAF/TIGER database. ZIP Codes assigned to businesses only or single delivery point address will not necessarily appear as ZCTAs.

In most instances the ZCTA code is the same as the ZIP Code for an area.

In creating ZCTAs, the Census Bureau took the most frequently occurring ZIP Code in an area for the ZCTA code. Some addresses will end up with a ZCTA code different from their ZIP Code.

Some ZIP Codes represent very few addresses (sometimes only one) and therefore will not appear in the ZCTA universe.

For more information on ZCTA see:

1) ZCTA FAQ [https://www.census.gov/geo/reference/zctafaq.html]
2) ZCTA Delineation Animation [https://www.census.gov/geo/reference/zcta/zcta_delin_anim.html]
3) What is the difference between ZIP code "boundaries" and ZCTA areas? - gis.washington.edu [http://gis.washington.edu/phurvitz/zip_or_zcta/]

Geographic ZIP/ZCTA Data

ZIP code (postal data)

ZIP code - postal codes, NYS and/or USA

- They NYS GIS Clearinghouse state ZIP code polygon file is from 2008. [http://gis.ny.gov/gisdata/inventories/details.cfm?DSID=934]
- We also have a ZIP code polygon file from circa 2010 from Esri Data and Maps.

It is assumed by BEH-GIS that regardless of which methodology is selected to represent ZIP codes, we will take ZIP code input data for WICNYS Group and tabulate the data into a crosswalk table so that nested ZIP codes, point-level and building level (or very small) ZIP codes will be incorporated into a larger relevant ZIP code or ZCTA.

ZCTA

If ZCTA is used and needed we can download the TIGER files from the US Census Bureau [https://www.census.gov/geo/maps-data/data/tiger-line.html].

Geocoding counts

The following are hypothetical counts for geocoding pass/fails assuming ~85% success on first pass.

Geocoding counts	Count	Percent
Study Subject Count	**0**	
Count addresses to be geocoded (including WIC locations, etc. that may have to be geocoded)	**119**	100%
Hypothetical Count Passed on 1st Address Geocoding attempt	102	85.71%
Hypothetical Count Failed on 1st Geocoding attempt	17	14.29%

Geocoding Time Estimate

Address Geocoding

Below is the estimate if we geocode the study subjects to point addresses.

Address Geocoding	Minimum estimate	Maximum estimate
1st attempt in hours	4	12
Minutes per record failed at 1st attempt (# of failed records x 3/5 mins)	3	5
Total in Hours:	4.85	13.42

ZIP Code Geocoding

Below is the estimate if we geocode the study subjects to ZIP code.

ZIP Code Geocoding	Minimum estimate	Maximum estimate
1st attempt in hours	4	12
Minutes per record failed after 1st attempt (# of failed records x 1/2 mins)	1	2
Total in Hours:	4.28	12.57

If new geocoding service-street network or address layer-is needed

It might be possible that an existing geocoding service locator within BEH-GIS is not the best one to use. In this case we would have to build a new geocoding service on a local machine. This estimate is only included in the Maximum Geocoding estimate.

New Locator	Minimum estimate (hrs)	Maximum estimate (hrs)
Building a new secure geocoding service (address locator)	4	8

2. Neighborhood level variable creation

Creating Neighborhoods (buffers, ZIP, or ZCTA)

This step is taking point information or ZIP or ZCTA information and creating the analytical GIS layer to be used for the following geoprocessing analysis steps. Points and buffers are usually geoprocessed so it only includes land area (depending on the variable) and, if used in this project, network buffers are more time consuming to create.

Neighborhood GIS Layer	Minimum estimate (hrs)	Maximum estimate (hrs)
Total in Hours:	2	6

WIC

Density of WIC vendors.

This data layer does not exist in NY State Open Data but I have reached out to NY State WIC via email [NYSWIC@health.state.ny.us]. An estimate of will be updated and added to the table below if this data can be obtained.

Density of WIC vendors	Minimum estimate (hrs)	Maximum estimate (hrs)
Estimate (missing)	0	0

WIC Locations.

BEH will likely have to geocode this data and it will be added to the overall geocoding list:

WIC Locations geocoding	This count is sent to the Geocoding counts table (p. 6)
Count WIC Locations	119

Source:

- NY State WIC Local Agencies List
  Description from [https://www.health.ny.gov/prevention/nutrition/wic/local_agencies.htm]:
  You can find the WIC Program nearest you by locating your county (listed alphabetically), and finding an agency that serves participants in your area. Many programs have multiple sites available for your convenience, and you can get contact information for them by calling the main site number listed. You can also call the Growing Up Healthy Hotline at 1-800-522-5006 for further assistance.
  Click on a blue agency name to go to its WIC web page. Not all agencies currently have web pages available for the WIC Program.

WIC Locations Geoprocessing	Minimum estimate (hrs)	Maximum estimate (hrs)
Count of WIC Locations in Neighborhood Geography	2	4
Total in Hours:	2	4

Data from Table 2: Community Health Living Resources and Constraints

Data list is from Table 2: Community Health Living Resources and Constraints (by WIC catchment ZIP code) from document provided by Dr. Andrew Rundle to BEH-GIS.

Neighborhood Census/ACS Variables

Neighborhood variables using Census Tract (for area-weighted interpolation to ZIP, if ZIP's are crosswalked to ZIP) or ZCTA (if ZIP's are crosswalked to ZCTA's). Note: for percent-based variables raw counts will also be provided.

Neighborhood deprivation - percent unemployed.

- 2008-2012 American Community Survey.

Neighborhood deprivation - percent household overcrowded.

- 2008-2012 American Community Survey.

Neighborhood mobility - linguistic isolation.

- 2008-2012 American Community Survey.

Linguistic Isolation Variables (comparable ACS variables to the following Census 2000 variables)

- Percent of population who are linguistically isolated [(P020004 + P020007 + P020010 + P020013) / P020001)].

Census/ACS Variables	Minimum estimate (hrs)	Maximum estimate (hrs)
ZCTA/Census Tract Data to Neighborhood Geography	2	8
Census API data mining	6	12
Percent Unemployed	2	4
Percent Household Overcrowded	2	4
Percent Linguistic Isolated	2	4
Total in Hours:	14	32

Neighborhood Fitness Resources:

Park/Playground access.

- For New York State - Only State-owned/managed (not including town parks, etc) State Park Facility Points with Playgrounds Map - including town and other municipal level administrated parks will take an inordinate amount of time.
  Description from [https://data.ny.gov/Recreation/State-Park-Facility-Points-with-Playgrounds-Map/2esg-8ipp]: The New York State Office of Parks, Recreation and Historic Preservation oversees 179 state parks and 35 historic sites, which are visited by 60 million people annually. The point locations of these facilities are contained within this file along with the website address for the park or historic site along with any associated golf course or nature center. Availability of a playground or swimmable beach is also noted but not complete. Some attribute information such as availability of a playground or swimmable beach is incomplete. Unless noted, it should be assumed that there is not a playground or beach at the facility. The agency is currently working to update and complete that information. This dataset was updated July 18, 2014. BEH-GIS has yet to contact New York State Office of Parks, Recreation and Historic Preservation to determine whether there is a 2010 dataset. Contact: notes@parks.ny.gov Please note from Description: Availability of a playground or swimmable beach is also noted but not complete. Some attribute information such as availability of a playground or swimmable beach is incomplete. Also Note: this dataset only includes State-managed sites. Including town and other municipal level administrated parks will take an inordinate amount of time. WICNYS says 2014 data is okay to use.
- New York City (NYC) subjects will be assigned NYC playgrounds GIS dataset from BEH-GIS NYC Playgrounds GIS file (Parks-derived).
- Another source of data that is not specifically Playground data but is Parks. The Park data is the Esri 10-1 (or possibly earlier dataset, ie. from 10-0 or 9-3-1) Data and Maps-Street Map North America Detailed parks layer:
  Metadata:
  TomTom North America, Inc./Esri 2012 - StreetMap™ North America
  Description:
  U.S. and Canada Parks represents parks and forests within the United Stats and Canada at national, state and local levels.

Playground/Park Access	Minimum estimate (hrs)	Maximum estimate (hrs)
Area of NYS Parks in Neighborhood Geography	3	6
Count of NYS Parks in Neighborhood Geography	2	4
Area of NYC Parks in Neighborhood Geography	3	6
Count of NYC Parks in Neighborhood Geography	2	4
Area of NYC Playgrounds in Neighborhood Geography	3	6
Count of NYC Playgrounds in Neighborhood Geography	2	4
Area of Esri Parks in Neighborhood Geography	3	6
Count of Esri Parks in Neighborhood Geography	2	4
Total in Hours:	20	40

Neighborhood Health Eating Resources:

Farmers Market Access.

- NYS farmers markets has two possible datasets:
  - Farmers Markets in New York State Map
    Description from [https://data.ny.gov/Economic-Development/Farmers-Markets-in-New-York-State-Map/gfni-eg8a]:
    In the past decade the number of farmers markets in New York State has grown at a rapid rate. The dataset published on the Department website contains information detailing the time and location of community farmers markets as well as the name and phone number of the market manager.
  - Data.gov Farmers Markets - National
    Description from [http://catalog.data.gov/dataset/farmers-markets-geographic-data]:
    longitude and latitude, state, address, name, and zip code of Farmers Markets in the United States
    WICNYS Group stated it is okay to use the latest version of this data.

Retail Food Stores	Minimum estimate (hrs)	Maximum estimate (hrs)
Count of NYS Farmers Markets in Neighborhood Geography	2	4
Count of Data.gov Farmers Markets in Neighborhood Geography	2	4
Total in Hours:	4	8

Density of food stores and food service places.

- Density of supermarkets –
  NYS Ag and Market Retail Food Stores Map
  Description from [https://data.ny.gov/Economic-Development/Retail-Food-Stores-Map/p2dn-xhaw]:
  The point map shows the locations of all retail food stores which are licensed by the Department of Agriculture and Markets. The initial view of the map is broken up into large geographic areas and displays the number of retail food stores in each area. To drill down to a smaller geographic area, click directly on the area of the map or click the plus sign to zoom in on the map. Data is the most recently submitted and available data. This map is currently a snapshot in time. For more information check out http://www.agriculture.ny.gov/FS/FSHome.html, or go to the "About" section.
  WICNYS Group stated it is okay to use the most current version of this data

Retail Food Stores	Minimum estimate (hrs)	Maximum estimate (hrs)
Count of Retail Food Stores Locations in Neighborhood Geography	2	4
Total in Hours:	2	4

SPARKS/EWPH:

SPARKS (Sports, Play and Active Recreation for Kids)/EWPH (Eat Well Play Hard) - licensed daycare facilities.

BEH-GIS has not previously used SPARKS/EWPH data.

- SPARKS
  WICNYS Group informed BEH-GIS that this data is available from NYC DOHMH (only NYC). Is there a state-wide dataset and/or who is the primary contact for this data from NYC DOHMH?

  WICNYS Group may have to provide a SPARKS dataset in geographic form (with coordinate geometry) or in aspatial format that BEH-GIS could geocode.
- EWPH
  - Child And Adult Care Food Program Participation Map

  Description from [https://health.data.ny.gov/Health/Child-And-Adult-Care-Food-Program-Participation-Ma/izpu-8t68] :
  This map displays the names and locations of Child and Adult Care Food Program (CACFP) participating day care sites and whether or not the site is Breastfeeding Friendly Certified with CACFP, participating in the Eat Well Play Hard in Child Care Settings (EWPHCCS) project, or participating in the Eat Well Play Hard in Day Care Homes (EWPHDCH) project. This dataset excludes Child and Adult Care Food Program participation provided at homeless shelters and legally-exempt day care home providers. Not all counties in NYS are serviced by the grantees implementing the project EWPHCSS. The EWPHDCH project is currently limited to the areas served by the contracted agencies. The Child and Adult Care Food Program dataset is related to the Child Care Related Programs dataset on the Open NY website, but includes additional nutrition information. The Office of Children and Family Services is currently working to update the Child Care Related Programs dataset on a more frequent schedule than the Child and Adult Care Food Program dataset. The agencies are working to synchronize the update schedule in the near future. We appreciate your patience in the interim. Temporarily, we have omitted addresses for regulated child care providers that provide home care since this information is available on Open.ny.gov by using this link: https://data.ny.gov/Human-Services/Child-Care-Regulated-Programs/cb42-qumz. For more information please visit http://www.health.ny.gov/prevention/nutrition/cacfp/ or go to the "About" tab.

SPARKS/EWPH:	Minimum estimate (hrs)	Maximum estimate (hrs)
Count of SPARKS Locations in Neighborhood Geography	2	4
Count of EWPH Locations in Neighborhood Geography	2	4
Total in Hours:	4	8

Neighborhood level variable creation Task Time Estimate Table

Below is table including all the Minimum and Maximum estimates for all the neighborhood level variables.

Task	Minimum Estimate (hrs)	Maximum Estimate (hrs)
Neighborhood Geographies	2	6
WIC Vendors	0	0
WIC Locations	2	4
Census/ACS Variables	14	32
Playground/Park Access	20	40
Farmers Markets (NYS/ data.gov)	4	8
Retail Food Stores	2	4
SPARKS/EWPH	4	8
Total in Hours:	48	102

3. Data Wrangling and QA/QC

After creating all the geoprocessed data variables. Some data munging (carpentry) will occur, joining all the variables, standardizing variable names and then QA/QC-ing the data.

Data Wrangling and QA/QC	Minimum Estimate (hrs)	Maximum Estimate (hrs)
Data Wrangling (tables joins, etc.)	4	12
QA/QC Review	4	16
Total in Hours:	8	28

4. Data Dictionary

At the completion of the data processing BEH creates a document of all of the data variables and methods used in constructing those variables as well as the source of the original data used for the project. In addition a full documentation of the geocoding results and methods is also created.

Data Dictionary	Minimum Estimate (hrs)	Maximum Estimate (hrs)
Data Dictionary	8	16
Total in Hours:	8	16

5. Data Delivery and Mapping (optional)

At the completion of the project and data dictionary the data will be delivered as an encrypted zip file (AES 256 encryption or equivalent). It can be delivered as .csv, in a Stata file, in an SPSS file, in a SAS file or any format desired. If the WICNYS Group wishes to map out some of the variables either state-wide or at a local level BEH-GIS could develop a mapping template and send out first drafts to the WICNYS Group. Maps may be created specifically for a report, group exploratory data analysis, a research publication or for web.

Data Delivery and Mapping	Minimum Estimate (hrs)	Maximum Estimate (hrs)
Data Delivery	0.25	1
Mapping	*8*	*24*
Total in Hours (without Mapping option):	0.25	1

Total Time Estimate of All Tasks

Below is a table summing all of the individual tasks into a Project total Time Minimum Estimate and Maximum Estimate. The project hours will fall into this range and will be recorded so that the research group will only pay for the actual hours worked by BEH-GIS.

Task	Minimum Estimate (hrs)	Maximum Estimate (hrs)
1. Geocoding (Min at ZIP, Max at Address 5 min worst case scenario with new geocoder built) for study subjects and other data layers that require geocoding	4.28	21.42
2. Neighborhood level variable creation	48	102
3. Data munging and QA/QC	8	28
4. Data Dictionary	8	16
5. Data Delivery and Mapping (optional and not included in the total time estimate)	0.25	1
Total:	**68.53**	**168.42**
Weeks:	**1.83**	**4.49**

Geoprocessing Methods Explained

The process of characterizing neighborhoods with social and built environment variables is an ideal job for a Geographical Information System (GIS). A GIS is unique in that it harnesses the power of both relational databases and geographic space and place. Combined you have an efficient means of statistically aggregating and describing what lies within a specific measurement geography (e.g., state, county, community district, zip code, custom buffer). By overlaying spatial features from multiple layers (e.g., streets, census block groups, land-use, and crime) which are attached to descriptive variables (e.g., length, area, speed, total population, name, category) that task is achieved.

Areal Weighting Interpolation

Areal Weighting Interpolation is a data transfer procedure between incompatible zonal systems. There are diverse zonal systems used for aggregating and reporting spatial data, say, census tracts, administrative districts, school districts, and so forth. Since they are often geographically incompatible, integration of spatial data requires data transfer between zonal systems. This process is called areal interpolation, and the areal weighting interpolation method is one of the most popular interpolation methods in GIS (Markoff and Shapiro, 1973; Lam, 1983; Flowerdew and Green, 1991). Assuming a uniform distribution of spatial objects, the areal weighting interpolation divides the count of spatial objects according to area in each zone, and sums up the counts in another incompatible zone.

Take for example Census block groups. Say you want to calculate census related variables for custom measurement geographies such as the 0.50-mile network buffers created for this project. Some census geographies (e.g., block groups) will fall completely within your buffers, while others, only portions will fall within your buffers. You therefore need to decide how to deal with those census geographies that do not fall completely within your buffers. You really only have four options:

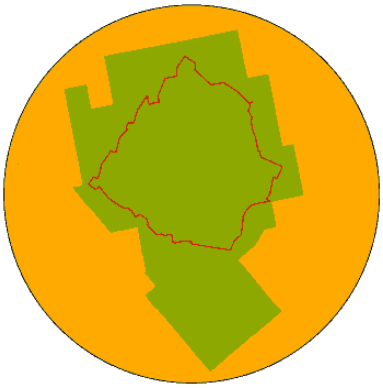
1) Include all full block groups that intersect your measurement geographies and their full variables.
2) Include all block groups with their centroid within the measurement geographies and their full variables.
3) Include all block groups completely contained within the measurement geographies and their full variables.
4) Include only those portions of block groups that intersect the measurement geographies and their apportioned
variables.

The four figures on the following demonstrate examples of these four options using a single network buffer and block group boundaries.

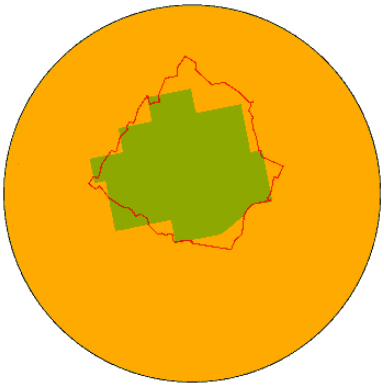
This analysis will likely utilize option 4. Where block groups (or any other spatial feature) are cut by the end of the buffer measurement geographies, the census variables are apportioned according to the percentage of land area falling inside and outside the buffers before calculating the results. Take for example the total population variable. Since we do not know where the population actually lives within each block group, all we can do is assume a normal distribution of the population. Where census blocks are cut by the end of the buffer measurement geographies, the population needs to be apportioned according to the percentage of block group falling inside the buffer. So we should know or you can easily calculate the area of the original block in whatever unit of measure you wish. Next, INTERSECT or compute the geometric intersection of the input features (i.e., census block groups) where features or portions of features which overlap the intersect features (i.e., custom network buffers) will be written to the final output feature. Then, recalculate the area of the output, and find the ratio of each block group falling inside the buffer. Finally, take that ratio times your total population value of each block to get your apportioned value.

Figures. – Examples of how to deal with non-contiguous data overlapping your study area.

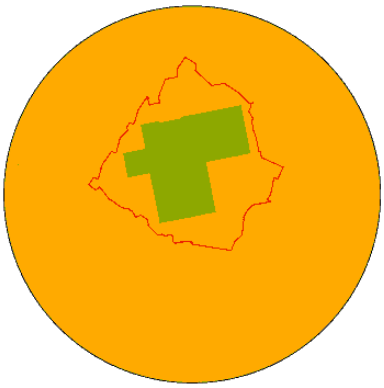
Option 1: full intersecting



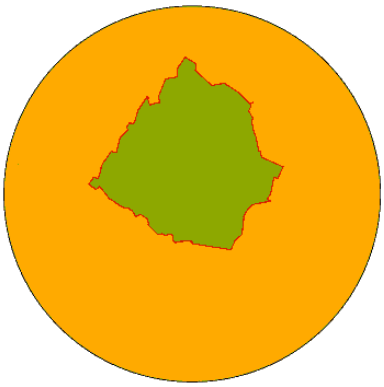
Option 2: centroid within block groups block groups



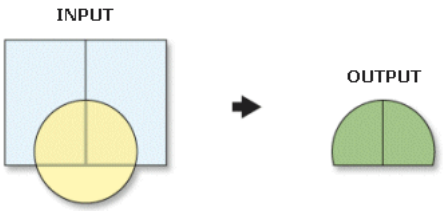
Option 3: completely contained block groups



Option 4: apportioned intersecting block groups.



For example, where a block group falls completely within the buffer the ratio will be 1.00 which in turns means the total population will be accounted for. But say you have a block group with 1,300 people living in it that is 250-km2 in total area but only 115-km2 of the block falls within the buffer. You would take 115 / 250 to get 0.46, which you then multiple against the total population for the block 0.46 * 1,300 to get 598 people. That is your apportioned population.



*apportioned population = (new_area / original_area) * total_population*

References

Flowerdew, R. and Green, M. (1991). Data Integration: Statistical Methods for Transferring Data between Zonal Systems. In Handling Geographical Information: Methodology and Potential Applications, edited by I. Masser and M. Blakemore (New York: Longman), pp. 38-54.

Gattrell, A.C. et al. (1996). Spatial point pattern analysis and its application in geographical epidemiology, Transactions of the Institute of British Geographers 21: pp. 256–274.

Guagliardo, M.F. et al. (2004). Physician accessibility: an urban case study of pediatric providers, Health & Place 10: pp. 273– 283.

Lam, N. N-S. (1983). Spatial Interpolation Methods: a Review. American Cartographer, 10: pp. 129-149.
Longley, P.A. et al. (2001) Geographic Information Systems and Science. Chichester, NY: John Wiley and Sons, Inc.

Markoff, J. and Shapiro, G. (1973). The Linkage of Data Describing Overlapping Geographical Units. Historical Methods Newsletter, 7: pp. 34-46.

O'Sullivan, D. and D. Unwin. (2002). Geographic information analysis. Hoboken, NJ: John Wiley and Sons, Inc.
Scott, D.W. (1998). Multivariate Density Estimation. John Wiley and Sons, Inc.