

WIC (Women, Infants, and Children) project in New York State (NYS) GIS Data Resources Document (WICNYS)

Sally Findley's Notes - handwritten doc in right middle drawer of 735 desk

Data Resources Documentation (Version 0.3) - 2014.09.08



 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



Table of Contents

Data Resources Documentation (Version 0.3) - 2014.09.08	1
Purpose:	3
Note on Geocoding:	3
ZCTA vs ZIP CODE	3
Geographic ZIP/ZCTA Data	4
ZIP code (postal data)	4
ZIP code - postal codes, NYS and/or USA.....	4
ZCTA	4
WIC.....	5
Density of WIC vendors.	5
WIC Locations.	5
WIC catchment area – possible methods.....	5
Data from Table 2: Community Health Living Resources and Constraints	5
Neighborhood Census/ACS Variables.....	6
Neighborhood deprivation - percent unemployed.	6
Neighborhood deprivation - percent household overcrowded.....	6
Neighborhood mobility - upward economic mobility. - percent white-collar workers	6
Neighborhood mobility - linguistic isolation.....	7
Walkability:.....	7
Neighborhood Walkability Index.....	7
Neighborhood Fitness Resources:	8
Playground access.....	8
Density of recreational and fitness outlets.....	8
Neighborhood Health Eating Resources:.....	9
Farmers Market Access.....	9
Density of food stores and food service places.....	9
Neighborhood Programs:.....	10
Programs for Mothers of young Children - Early childhood education programs for mothers with young children.	10
Day care - Density of child care and HS.....	10
Programs for healthy day care-Number of Head Start, Pre-K programs with healthy diet and activity programs.....	10
Programs for school-age - Average FITNESSGRAM for grades K-5 at schools in catchment (NYC only).....	10
SPARKS/EWPH:	11
SPARKS (Sports, Play and Active Recreation for Kids)/EWPH (Eat Well Play Hard) - licensed daycare facilities.	11
Geoprocessing Methods Explained	11
Areal Weighting Interpolation	11
References	13

Purpose:

The purpose of this document is to list the GIS data resources 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. As 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 developed further during the more specific time estimation process.

All **RED TEXT** signifies an issue that requires the WICNYS Group's feedback or guidance or signifies an important note for the WICNYS Group.

Note on Geocoding:

(BEH-GIS geocoding estimates are ~3-5 minutes per record geocode depending on the data quality).

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.

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

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.

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/) [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>].

WIC

No - these are the WIC sites, not WIC vendors - you need to get # vendors by ZCTA from the VMA's (is VMA's the Vendor Management Associations)

Density of WIC vendors.

- [Women Infant and Children WIC Program Site Map](#)

Description from [<https://health.data.ny.gov/Health/Women-Infant-and-Children-WIC-Program-Site-Map/ns2x-9zqf>]:

This map shows the locations of sites across NYS where WIC services are delivered and the locations of agencies that oversee WIC vendors (Vendor Management Associations). Each type of site is represented by a different color on the map. Temporary sites may not be located in a permanent location and may have limited hours of operation. The Special Supplemental Nutrition Program for Women, Infants and Children (WIC) offers nutrition education, breastfeeding support, referrals and a variety of nutritious foods to low-income pregnant, breastfeeding or postpartum women, infants and children up to age five to promote and support good health. For more information, check out: <http://www.health.ny.gov/prevention/nutrition/wic/> or go to the "About" tab.

WICNYS Group: Please validate that this data matches the WIC Vendors data requirement.

WIC Locations.

BEH-GIS assumes this data will either be provide by WICNYS Group or will have to be geocoded from this following data 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.

The site map above is what you want to use, as it includes temp. sites

WICNYS Group: Please validate that this data matches the WIC locations data requirement or if you will be providing this data.

WIC catchment area – possible methods

- 1) Spider Diagram. (ZIPs to facility distance) - nearest facility to ZIP (likely centroid) for ZIP without WIC locations.

Yes. These will work

- 2) Or only ZIP's with WIC facilities in the ZIP bounding geometry.

WICNYS Group will need to provide clarification as to how to define the WIC catchment area.

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 - upward economic mobility. - percent white-collar workers

- 2008-2012 American Community Survey. SEX BY OCCUPATION FOR THE CIVILIAN EMPLOYED POPULATION 16 YEARS AND OVER - Universe: Civilian employed population 16 years and over.

Select occupations from the following list of Census variables for occupations to identify which occupations WICNYS Group would like included as White Collar occupations.

- C24010003 Management, business, science, and arts occupations:
- C24010004 Management, business, and financial occupations:
- C24010005 Management occupations
- C24010006 Business and financial operations occupations
- C24010007 Computer, engineering, and science occupations:
- C24010008 Computer and mathematical occupations
- C24010009 Architecture and engineering occupations
- C24010010 Life, physical, and social science occupations
- C24010011 Education, legal, community service, arts, and media occupations:
- C24010012 Community and social service occupations
- C24010013 Legal occupations
- C24010014 Education, training, and library occupations
- C24010015 Arts, design, entertainment, sports, and media occupations
- C24010016 Healthcare practitioners and technical occupations:
- C24010017 Health diagnosing and treating practitioners and other technical occupations
- C24010018 Health technologists and technicians
- C24010019 Service occupations:
- C24010020 Healthcare support occupations
- C24010021 Protective service occupations:
- C24010022 Fire fighting and prevention, and other protective service workers including supervisors
- C24010023 Law enforcement workers including supervisors
- C24010024 Food preparation and serving related occupations
- C24010025 Building and grounds cleaning and maintenance occupations
- C24010026 Personal care and service occupations
- C24010027 Sales and office occupations:
- C24010028 Sales and related occupations
- C24010029 Office and administrative support occupations
- C24010030 Natural resources, construction, and maintenance occupations:
- C24010031 Farming, fishing, and forestry occupations
- C24010032 Construction and extraction occupations
- C24010033 Installation, maintenance, and repair occupations

- C24010034 Production, transportation, and material moving occupations:
- C24010035 Production occupations
- C24010036 Transportation occupations
- C24010037 Material moving occupations

Neighborhood mobility - linguistic isolation.

- 2008-2012 American Community Survey.

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

- The count of English speaking households by linguistic isolation [P020002].
- The count of Spanish speaking households linguistic isolation [P020003].
- Percent of population who are linguistically isolated $[(P020004 + P020007 + P020010 + P020013) / P020001]$.
- The count of households that are linguistically isolated that speak Spanish [P020004].
- The percent of households that are linguistically isolated that speak Spanish $[P020004] / [P020003]$ (total Spanish households).
- The count of households that are linguistically isolated that speak Other Indo-European languages [P020007].
- The percent of households that are linguistically isolated that speak Other Indo-European languages $[P020007] / [P020006]$ (total Other Indo-European households).
- The count of households that are linguistically isolated that speak Asian and Pacific Island languages [P020010].
- The percent of households that are linguistically isolated that speak Asian and Pacific Island languages $[P020010] / [P020009]$ (total Asian and Pacific Island households).
- The count of households that are linguistically isolated that speak Other languages [P020013].
- The percent of households that are linguistically isolated that speak Other languages $[P020013] / [P020012]$ (total Other languages households).

Walkability:

Neighborhood Walkability Index.

- For NYS locations - WalkScore website – purchase from WalkScore for New York State WIC locations.

BEH-GIS cannot estimate a price for WalkScore data without knowing the number of addresses and then consulting with WalkScore.

- NYC Walkability index for New York City (BEH).

BEH-GIS creates and maintains this dataset.

Can you please do this for the WIC sites - per WIC site locations on p. 5

Neighborhood Fitness Resources:

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.

okay to use state parks

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.**

okay to use 2014

- New York City (NYC) subjects will be assigned NYC playgrounds GIS dataset from BEH-GIS NYC Playgrounds GIS file (Parks-derived).

BEH-GIS has a fairly up-to-date version of this file.

- 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 States and Canada at national, state and local levels.

This is a parks file not playground but may be included as a variable or not. WICNYS Group, please decide if Parks is a variable that would like to be included.

Can you include this also & combine as Park or State Park or Playground?

Density of recreational and fitness outlets.

- US Census Bureau County Business Patterns

Can you use what you already did w. SIC codes?

For the Neighborhood Fitness Resources data on commercial recreational and fitness outlets by ZIP code are likely to be included in the County Business Patterns dataset which categorizes businesses by the North American Industry Classification System. BEH-GIS would need to build an algorithm to select for NAICS codes that reflect recreational and fitness outlets. We have previously done this using SIC codes which is another business classification system.

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-cg8a>]:
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.

Okay to use 2014

This dataset was updated August 14, 2014. BEH-GIS has contacted NYS Department of Agriculture and Markets to determine whether there is a 2010 dataset. Contact: Jacqueline.Boyer@agriculture.ny.gov

- [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

This dataset was updated July 16, 2014. BEH-GIS has contacted USDA to determine whether there is a 2010 dataset. It may be possible that the JSON API has more detailed founding year information but will require some programming with the API. Contact: edward.ragland@ams.usda.gov

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.

This dataset was updated August 27, 2014. BEH-GIS has yet to contact NYS Agriculture and Markets to determine whether there is a 2010 dataset. Contact: currently no email, would have to create NYS Open Data account to contact owner.

okay to use 2014

Neighborhood Programs:

Programs for Mothers of young Children - Early childhood education programs for mothers with young children.

- Average number of programs per child < 5

Ok - will not have it.

BEH-GIS cannot find a datasource for Neighborhood Programs for Mothers of young Children. This dataset would have to be provided by WICNYS Group.

Day care - Density of child care and HS.

- Number of day care slots/ pop 3-5 years of age (2010 Census - perhaps 2008-2012 ACS for consistency).
- [Child Care Regulated Programs](#)
Description from [<https://data.ny.gov/Human-Services/Child-Care-Regulated-Programs-Map/s8uq-s4wqj>]:

The point map shows the locations of regulated child care providers. The initial view of the map is broken up into large geographic areas and displays the number of providers 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. The map can be filtered by number of school age capacity by changing this option under the Filter tab. Data is the most recently submitted and available data. This map is currently a snapshot in time. For more information check out http://www.ocfs.state.ny.us/main/childcare/ccfs_template.asp, or go to the "About" section.

Ok to use most current data

WICNYS Group needs to verify that this dataset is what they had in mind for Neighborhood Program day care slots. In addition the years of operation are not included in the forward-facing data table, they may have to be scraped from the individual DayCare page. For example, [<http://it.ocfs.ny.gov/DayCareFacilitySearch/Profile/Index/333056>]. A python script may be written by BEH-GIS to scrape this. That time would be necessary to include in the time estimate.

Programs for healthy day care-Number of Head Start, Pre-K programs with healthy diet and activity programs.

- HS and Pre-K slots with programs having healthy eating or activity options/ pop 3-5 years of age (2010 Census - perhaps 2008-2012 ACS for consistency).

use the list of sites w. EWPH or SPARKS

BEH-GIS has not been able to find publicly available data for listing which day care, Head Start, or Pre-K programs provide healthy diet and activity programs. These data would have to be provided by WICNYS Group.

Programs for school-age - Average FITNESSGRAM for grades K-5 at schools in catchment (NYC only).

- Average FITNESSGRAM for the WIC program catchment area.

Can you use most current date you have? Is this not avail. from DOHMH?

BEH-GIS does not have this data for 2010. These data would have to be provided by WICNYS Group. If the data require geocoding, there are two options: geocoding the data to the child's home ZIP code and estimation of FITNESSGRAM scores by ZIP code of residence; or linkage of the child to their school and then estimation of FITNESSGRAM scores for children attending schools in a ZIP code.

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
Cannot find a NYS SPARKS dataset.

SPARKS was supported by NYD DOHMH. Get the list from them. It will be by ZIP code.

WICNYS Group would 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>] :

yes!

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.

WICNYS Group: Please validate that this data matches the EWPH data requirement.

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:

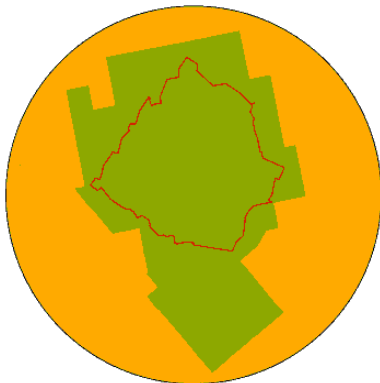
- 1) Include all full block groups that intersect your measurement geographies and their full variables.
- 2) Include all block groups with their centroid within the measurement geographies and their full variables.
- 3) Include all block groups completely contained within the measurement geographies 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.

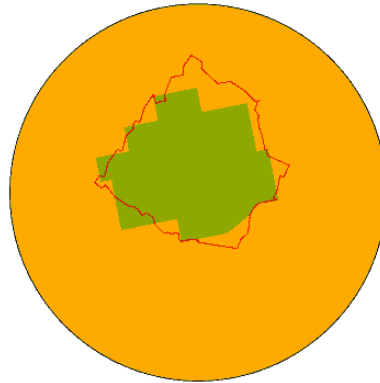
This analysis will likely utilize option 4. Where block groups (or any other spatial feature) are cut by the end of the buffer measurement geographies, the census variables are apportioned according to the percentage of land area falling inside and outside the buffers before calculating the results. 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.

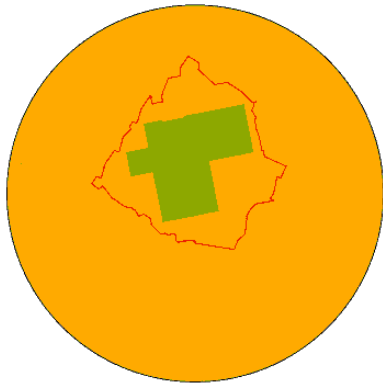
Option 1: full intersecting



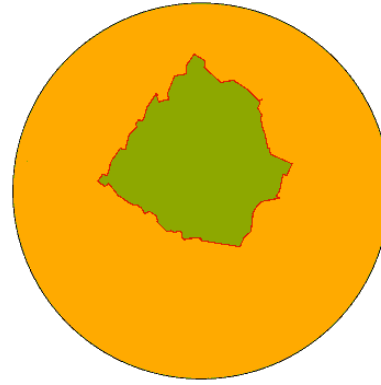
Option 2: centroid within block groups



Option 3: completely contained block groups

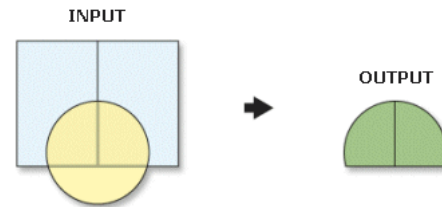


Option 4: apportioned intersecting block groups.



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.

$$\text{apportioned population} = (\text{new_area} / \text{original_area}) * \text{total_population}$$



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.