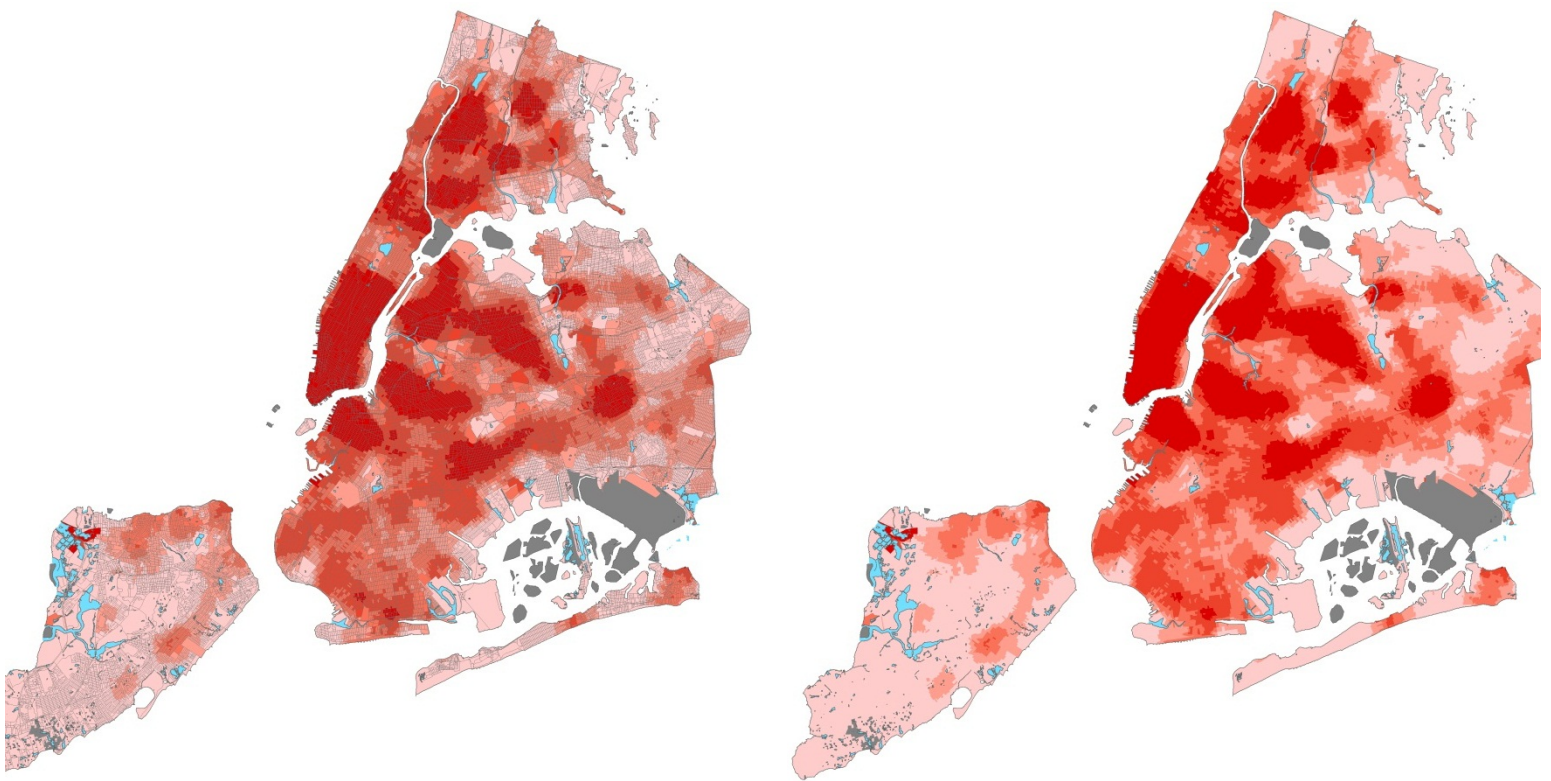# *GIS Code Book: New York City 2010 Census Tract Walkability Index*

Last Updated: 12-January-2015

Edited by:
Daniel M. Sheehan
Geographer, GIS Analyst

Prepared by:
James W. Quinn, MA
Geographer, GIS Analyst
Built Environment and Health Working Group
Department of Epidemiology
Mailman School of Public Health
Columbia University
e: jq2145@columbia.edu | p: (319) 332-1236

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

Walkability Index Scale

A number of researchers have constructed walkability indices which summarize built environment features believed to promote walking. Although specification details vary, these indices typically include measures of population density, land use, and street network. Our walkability measure was adapted from that employed in recent papers by Frank and colleagues (2005 and 2006), which includes four components: residential population density (density of population per total residential land area), intersection density, an entropy measure of land use based on the distribution of building floor area among six land use types (education, entertainment, single-family residential, multi-family residential, retail, and office), and the retail floor area ratio, or the ratio of retail building floor area to retail land area. All of the Frank components were z-scored and summed, with intersection density receiving a double weight for the *Frank Scale*, but not for our scale. Our "BEH walkability scale" is documented in a paper by Neckerman and colleagues (2009).

Frank et al. Linking objectively measured physical activity with objectively measured urban form: findings from SMARTRAQ. American Journal of Preventive Medicine. 2005;28(2 Suppl 2):117-125.

Frank et al. Many Pathways from Land Use to Health: Associations between Neighborhood Walkability and Active Transportation, Body Mass Index, and Air Quality. Journal of the American Planning Association. 2006;72(1):75-87.

Neckerman et al. Disparities in urban neighborhood conditions: Evidence from GIS measures and field observation in New York City. Journal of Public Health Policy. 2009;1(1 Suppl): S264-S285.

Important Note

Many of the figures in this document refer to 2010 Census Blocks. This data dictionary is for the 2010 Census Tracts. Figures are only for illustrative purposes.
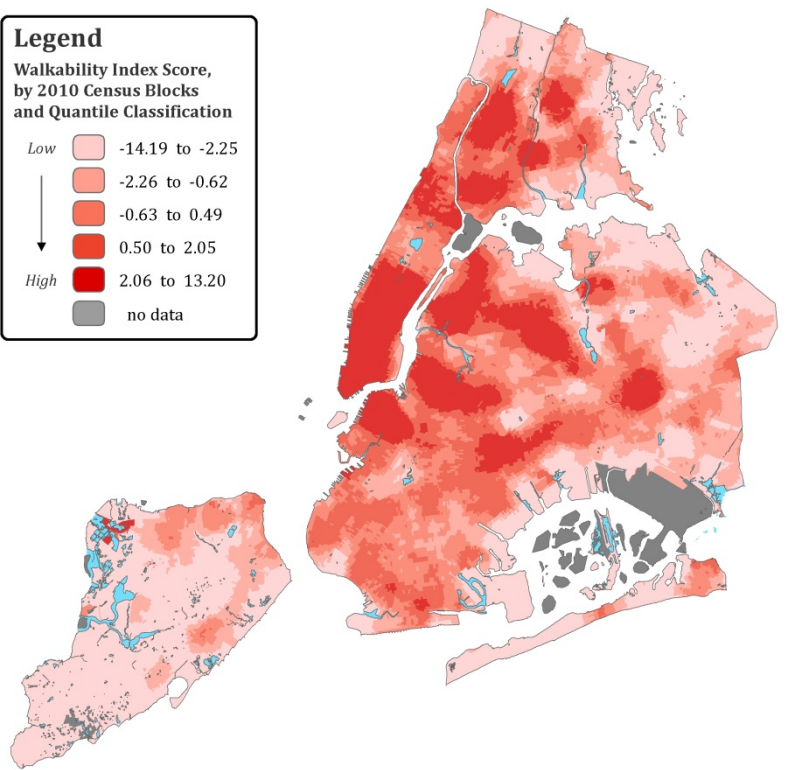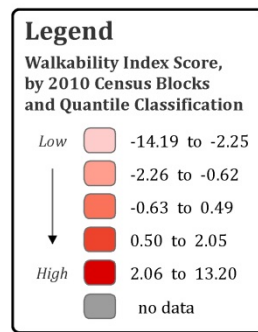
Walkability Index Scales: City-Wide Buffers

As an alternative to creating a 'neighborhood walkability index scale' based solely on the z-scored components within buffers specific to a study cohort, the BEH Working Group has developed a walkability index scale that considers the z-scored components across New York City as a whole.

The issue with creating a walkability index scale using buffer definitions specific to a study cohort is that the cohort may be spatially biased or clustered in a particular neighborhood(s) or parts of the City (e.g., northern Manhattan or the south Bronx). This could dramatically influence the range of z-scored components, and thus not provide a representative scale of walkability for that cohort compared to the walkability of the rest of the City.

The BEH Working Group has tackled this problem by buffering 2010 Census Block centroids *(n=38,526)* by 1-km radial buffers and then deriving the walkability components that are later z-scored. In doing so, a more appropriate apples-to-apple comparison of walkability can be made between, say, an address in Staten Island to the rest of the City as whole rather than a cluster of addresses all within a similar built environment.

New York City-wide Walkability Index Scale using Census Blocks

In order to assign each 2010 Census Tract with the city-wide walkability index scale, the walkability index scale values for each Block centroid were simply averaged up to the Tract level. The following variables are the result of that process:

t10_key – Unique 11-digit US 2010 Census Tract ID.

t10_km2 – Total area of US 2010 Census Tract geography in km2.

t10_lndkm2 – Total **land** area of US 2010 Census Tract geography in km2 *(inland water bodies removed)*.

t10_cnt – Count of unique 2010 Census Blocks nested within each Tract and whose walkability index scale values participated in averaging-up process for each Tract.

t10_walk – BEH Walkability Scale.

Walkability Index Scale Variables

To date, BEH has created and used two different versions of the Walkability Index Scales, which we refer to here as the *"Frank et al. 2006"* and BEH scales. The *"Frank 2006"* includes z-scored variables: residential density, land use mix using 5 land use types, intersection density * 2, and retail area ratio. The BEH scale includes z-scored variables: residential density, land use mix using 5 land use types, intersection density, retail area ratio, and subway stop density. Note that the BEH scale does not multiply intersection density by 2 and does include subway density.

t10_intden_z – Density of unique streets intersections per km2: z-scored.
t10_sub07d_z – Density of subway stations per km2: z-scored.
t10_rtlfar_z – Retail floor area ratio: z-scored.
t10_resdn1_z – Density of residential units: z-scored.
t10_entrpy_z – Entropy land use mix: z-scored.

t10_walk_cat – Quintiles of BEH Walkability Scale.

Please note that the variables in grey text below have not been provided in this deliverable and are only described here to provide context for the variables that are being provided.

Walkability Index Scale "Component" Variables (these variables are not included in the dataset)

b1_intden – Density of unique streets intersections per km2.
b1_sub07d – Density of subway stations per km2.
b1_rtlfar – Retail floor area ratio – Retail building floor area divided by retail land area in km2.
b1_resdn1 – Density of res units – Number of residential units divided by total residential building floor area in km2.
b1_entrpy – Land Use Mix – An entropy measure using the five of the six land use types employed in Frank et al. (2006). Single- and multi-family residential areas were combined because most housing in New York City is multi-family. Parcel-level measures of residential, office, and retail floor area were available from the MapPLUTO (version 04c; October 2004-October 2005) database. We used the MapPLUTO building class codes to identify buildings associated with education (schools) or entertainment (theaters, recreational facilities), and attributed the entire floor area of the identified building to education or entertainment. The entropy formula used was adapted from Frank et al. (2005), which yielded more plausible results: Land Use Mix = A/ln(N) where: A= –((b1/a)*ln(b1/a)+(b2/a)*ln(b2/a)+…) and b1 is the building floor area covered by the first land use, b2 is the building floor area covered by the second land use, etc., a is the total floor area across the five land uses, and N is the total number of land uses represented in the census tract. Zero values for b1…b5 were set to .000001 to avoid zero or undefined terms.

Components of the Entropy Measure:
b1_b1 – Total building area for Education uses in square feet (set to .000001 if 0)
b1_b2 – Total building area for Entertainment uses in square feet (set to .000001 if 0)
b1_b3 – Total building area Residential uses in square feet (set to .000001 if 0)
b1_b4 – Total building area for Retail uses in square feet (set to .000001 if 0)
b1_b5 – Total building area for Office uses in square feet (set to .000001 if 0)
b1_a – Total floor area across the five land uses in square feet (set to .000001 if 0)
b1_n – The total number of land uses represented (between 1 and 5)

ArcMap Entropy Field Calculation Expression:
$\text{entropy} = -(((([b1] / [a]) * \log([b1] / [a])) + (([b2] / [a]) * \log([b2] / [a])) + (([b3] / [a]) * \log([b3] / [a])) + (([b4] / [a]) * \log([b4] / [a])) + (([b5] / [a]) * \log([b5] / [a]))) / \log([n])$

Walkability Index Scale by 2010 Census Blocks

Legend

**Walkability Index Score,
by 2010 Census Blocks
and Quantile Classification**

Low

	-14.19 to -2.25
	-2.26 to -0.62
	-0.63 to 0.49
	0.50 to 2.05
	2.06 to 13.20
	no data

High