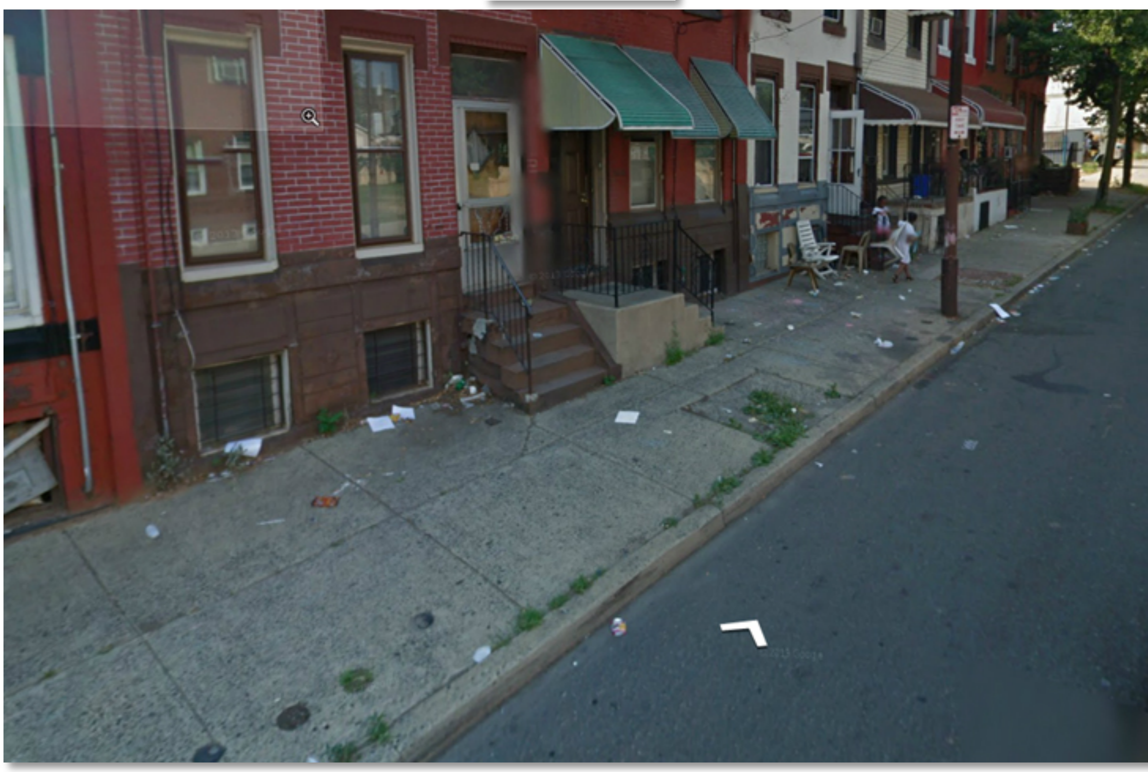
Fragile Families - 5 and 9 year (waves 4 & 5) Neighborhood Variables (250m and 1km)

American Community Survey 2008-2012, Crime (Esri 2010 & 2012), Street View (BEH) Neighborhood Disorder & Pedestrian Safety and NETS (1990-2010) Business data (by overall flag).



Prepared By:
Daniel M. Sheehan
Geographer, GIS Analyst/Programmer
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. Gina Lovasi
Assistant Professor of Epidemiology,
Co-Director Epidemiology and Population
Health Summer Institute at Columbia
University (EPIC) The Built Environment & Health Project
Columbia University, Mailman School of Public Health
Department of Epidemiology
722 West 168th St. – R804
New York, NY 10032

Table of Contents

Introduction

This data dictionary includes data that was originally developed for the fragile families project. Each record represents fragile families identifier, year and parent ID. For a single family with data from the mother and father for both waves will be represented with four (4) years of data. Two neighborhood geographies were created; a 250 meter radial buffer and a 1000 meter radial buffer. Each row includes variables from the following categories; US Census American Community Survey (ACS) 2008-2012, Crime Esri by Block Group for 2010 and 2012, business environment data NETS and Street View neighborhood disorder and pedestrian safety.

Important Note on variables:

Please note that Street View related variables are only for those living in the Street View audited cities and NETS data is only for the 23 County region around New York City. Census and Crime related variables should be available for all subjects.

Unique Identifier

geoid
Fragile Families family identifier for geoid.

year
Year from Fragile Families interview wave.

parent
Parent ID of Fragile Families
M = mother
F = father

Neighborhood Buffer Distances – variable prefix *****

The two neighborhood units of geography used for this project were a 250 meter radial (as the crow flies) buffer and a 1000 meter (1 kilometer) radial buffer.

In order to cut down on the size of the data dictionary and reduce redundancy for all variables, they will be preceded with a wildcard prefix; *****.

*****
[b0250m, b1000m] - prefix for buffer distance

b0250m
buffer 250 meter neighborhood geography unit.

b1000m
buffer 1000 meter neighborhood geography unit. (1 kilometer)

*******areasqmeter**
area of neighborhood geography in square meters.

American Community Survey (2008-2012) Variables.

Census Tract ACS 2008-2012 variables were acquired from the US Census Bureau API [https://www.census.gov/developers/] at the 2010 Census Tract level of geography. These variables were created using areal weighting interpolation. The original census variable names were slightly modified in order to provide a systematic naming convention across all project neighborhood definitions and to provide an indicator of the type of variable being provided. The ACS census variable names use the following naming convention:

acs = American Community Survey 2008-2012 5 Year

****
[b0250m, b1000m] - prefix for buffer distance in meters

> **b0250m**
> buffer 250 meter neighborhood geography unit.
>
> **b1000m**
> buffer 1000 meter neighborhood geography unit. (1 kilometer)

*******countrows**
Count of rows in Census-related variables (not useful for analysis).

Total Population Variables

*******acstotpop**
Area weighted total population derived from American Community Survey (ACS) 2008-2012 5 Year (Census Tract Level variables) in Neighborhood Geography (*****).

Age-related Variables

*******acspctage34nunder**
Percent population 34 years of age and younger in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001003E'] + df[geo+'B01001004E'] + df[geo+'B01001005E'] +
df[geo+'B01001006E'] + df[geo+'B01001007E'] + df[geo+'B01001008E'] +
df[geo+'B01001009E'] + df[geo+'B01001010E'] + df[geo+'B01001011E'] +
df[geo+'B01001012E'] + df[geo+'B01001027E'] + df[geo+'B01001028E'] +
df[geo+'B01001029E'] + df[geo+'B01001030E'] + df[geo+'B01001031E'] +
df[geo+'B01001032E'] + df[geo+'B01001033E'] + df[geo+'B01001034E'] +
df[geo+'B01001035E'] + df[geo+'B01001036E'] ) / df[geo+'B01001001E']
```

*******acspctage35up**
Percent population 35 years of age and older in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001013E'] + df[geo+'B01001014E'] + df[geo+'B01001015E'] +
df[geo+'B01001016E'] + df[geo+'B01001017E'] + df[geo+'B01001018E'] +
df[geo+'B01001019E'] + df[geo+'B01001020E'] + df[geo+'B01001021E'] +
df[geo+'B01001022E'] + df[geo+'B01001023E'] + df[geo+'B01001024E'] +
df[geo+'B01001025E'] + df[geo+'B01001037E'] + df[geo+'B01001038E'] +
df[geo+'B01001039E'] + df[geo+'B01001040E'] + df[geo+'B01001041E'] +
df[geo+'B01001042E'] + df[geo+'B01001043E'] + df[geo+'B01001044E'] +
df[geo+'B01001045E'] + df[geo+'B01001046E'] + df[geo+'B01001047E'] +
df[geo+'B01001048E'] + df[geo+'B01001049E'] ) / df[geo+'B01001001E']
```

*******acspctage60up**
Percent population 60 years of age and older in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001018E'] + df[geo+'B01001019E'] + df[geo+'B01001020E'] +
df[geo+'B01001021E'] + df[geo+'B01001022E'] + df[geo+'B01001023E'] +
df[geo+'B01001024E'] + df[geo+'B01001025E'] + df[geo+'B01001042E'] +
df[geo+'B01001043E'] + df[geo+'B01001044E'] + df[geo+'B01001045E'] +
df[geo+'B01001046E'] + df[geo+'B01001047E'] + df[geo+'B01001048E'] +
df[geo+'B01001049E'] ) / df[geo+'B01001001E']
```

*******acspctage18_24**
Percent population 18-24 years of age in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001007E'] + df[geo+'B01001008E'] + df[geo+'B01001009E'] +
df[geo+'B01001010E'] + df[geo+'B01001031E'] + df[geo+'B01001032E'] +
df[geo+'B01001033E'] + df[geo+'B01001034E']  ) /  df[geo+'B01001001E']
```

*******acspctage25_34**
Percent population 25-34 years of age in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001011E'] + df[geo+'B01001012E'] + df[geo+'B01001035E'] +
df[geo+'B01001036E'] ) / df[geo+'B01001001E']
```

*******acspctage35_44**
Percent population 35-44 years of age in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001013E'] + df[geo+'B01001014E'] + df[geo+'B01001037E'] +
df[geo+'B01001038E'] ) / df[geo+'B01001001E']
```

*******acspctage45_54**
Percent population 45-54 years of age in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001015E'] + df[geo+'B01001016E'] + df[geo+'B01001039E'] +
df[geo+'B01001040E'] ) / df[geo+'B01001001E']
```

*******acspctage55_64**
Percent population 55-64 years of age in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001017E'] + df[geo+'B01001018E'] + df[geo+'B01001019E'] +
df[geo+'B01001041E'] + df[geo+'B01001042E'] + df[geo+'B01001043E']) /
df[geo+'B01001001E']
```

*******acspctage65up**
Percent population 65 years of age and older in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B01001020E'] + df[geo+'B01001021E'] + df[geo+'B01001022E'] +
df[geo+'B01001023E'] + df[geo+'B01001024E'] + df[geo+'B01001025E'] +
df[geo+'B01001044E'] + df[geo+'B01001045E'] + df[geo+'B01001046E'] +
df[geo+'B01001047E'] + df[geo+'B01001048E'] + df[geo+'B01001049E'] ) /
df[geo+'B01001001E']
```

Sex, Race, Economic, Etc. Variables

*******acspctmale**
Percent population Male in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
df[geo+'B01001002E'] / df[geo+'B01001001E']
```

*******acspctwhite**
Percent population white in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
df[geo+'B02001002E'] / df[geo+'B02001001E']
```

*******acspcthisp**
Percent population Hispanic in Neighborhood Geography (*****).

American Community Survey 2008-2012 (5-year)
```
df[geo+'B03002012E'] / df[geo+'B03002001E']
```

*******acspctblack**
Percent population Black in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
df[geo+'B02001003E'] / df[geo+'B02001001E']
```

*******acspctasian**
Percent population Asian in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
df[geo+'B02001005E'] / df[geo+'B02001001E']
```

*******acspctother**
Percent population Other (than Asian, Black or White) in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
( df[geo+'B02001004E'] + df[geo+'B02001006E'] + df[geo+'B02001007E'] +
df[geo+'B02001008E']) / df[geo+'B02001001E']
```

*******acspctforborn**
Percent population foreign born in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
( df[geo+'B05001005E'] + df[geo+'B05001006E'] ) / df[geo+'B05001001E']
```

*******acspctlingiso**
Percent population linguistic isolation in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
( df[geo+'B16002004E'] + df[geo+'B16002007E'] + df[geo+'B16002010E'] +
df[geo+'B16002013E']) / df[geo+'B16002001E']
```

*******acspcthhownocc**
Percent household owner in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
df[geo+'B25003002E'] / df[geo+'B25003001E']
```

*******acspctsameh1y**
Percent population in same house 1 year ago in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
df[geo+'B07001017E'] / df[geo+'B07001001E']
```

*******acspctpov**
Percent population in poverty in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
( df[geo+'C17002002E'] + df[geo+'C17002003E']  ) / df[geo+'C17002001E']
```

*******acspctpub**
Percent population with public assistance in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
df[geo+'B19057002E'] / df[geo+'B19057001E']
```

*******acspctfemheadhh**
Percent households with Female householder, no husband present in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
df[geo+'B11001006E'] / df[geo+'B11001001E']
```

*******acspctearn50kup**
Percent population 15 years and over who work full-time in past 12 months that make $50,000 and over in income in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)
```
( df[geo+'B19325021E'] + df[geo+'B19325022E'] + df[geo+'B19325023E'] +
df[geo+'B19325024E'] +df[geo+'B19325025E'] + df[geo+'B19325044E'] +
```

```
df[geo+'B19325045E'] + df[geo+'B19325046E'] + df[geo+'B19325047E'] +
df[geo+'B19325048E'] ) / df[geo+'B19325001E']
```

*******acspctunemploy**
Percent population 16 years and over who are civilians in the labor force that are unemployed in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B23001008E'] + df[geo+'B23001015E'] + df[geo+'B23001022E'] +
df[geo+'B23001029E'] + df[geo+'B23001036E'] + df[geo+'B23001043E'] +
df[geo+'B23001050E'] + df[geo+'B23001057E'] + df[geo+'B23001064E'] +
df[geo+'B23001071E'] + df[geo+'B23001076E'] + df[geo+'B23001081E'] +
df[geo+'B23001086E'] + df[geo+'B23001094E'] + df[geo+'B23001101E'] +
df[geo+'B23001108E'] + df[geo+'B23001115E'] + df[geo+'B23001122E'] +
df[geo+'B23001129E'] + df[geo+'B23001136E'] + df[geo+'B23001143E'] +
df[geo+'B23001150E'] + df[geo+'B23001157E'] + df[geo+'B23001162E'] +
df[geo+'B23001167E'] + df[geo+'B23001172E'] )  / df[geo+'B23001001E']
```

*******acspctjobmanagr**
Percent population 16 years and over in Management, business, science, and arts occupations in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'C24010003E'] + df[geo+'C24010039E'] ) / df[geo+'C24010001E']
```

*******acspctednohisch**
Percent population 25 years and over with no high school diploma or GED or alternative in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B15002003E'] + df[geo+'B15002004E'] + df[geo+'B15002005E'] +
df[geo+'B15002006E'] +  df[geo+'B15002007E'] + df[geo+'B15002008E'] +
df[geo+'B15002009E'] +  df[geo+'B15002010E'] + df[geo+'B15002020E'] +
df[geo+'B15002021E'] + df[geo+'B15002022E'] + df[geo+'B15002023E'] +
df[geo+'B15002024E'] + df[geo+'B15002025E'] + df[geo+'B15002026E'] +
df[geo+'B15002027E'] ) / df[geo+'B15002001E']
```

*******acspctedyeshisch**
Percent population 25 years and over with at least a high school diploma or GED or alternative in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B15002011E'] + df[geo+'B15002012E'] + df[geo+'B15002013E'] +
df[geo+'B15002014E'] + df[geo+'B15002015E'] + df[geo+'B15002016E'] +
df[geo+'B15002017E'] + df[geo+'B15002018E'] + df[geo+'B15002028E'] +
df[geo+'B15002029E'] + df[geo+'B15002030E'] + df[geo+'B15002031E'] +
df[geo+'B15002032E'] + df[geo+'B15002033E'] + df[geo+'B15002034E'] +
df[geo+'B15002035E'] ) / df[geo+'B15002001E']
```

*******acspctedcolgeup**
Percent population 25 years and over with at least Associate;s degree, Bachelor's degree, Master's degree, Professional school degree or Doctorate degree in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
( df[geo+'B15002014E'] + df[geo+'B15002015E'] + df[geo+'B15002016E'] +
df[geo+'B15002017E'] + df[geo+'B15002018E'] + df[geo+'B15002031E'] +
df[geo+'B15002032E'] + df[geo+'B15002033E'] + df[geo+'B15002034E'] +
df[geo+'B15002035E'] ) / df[geo+'B15002001E']
```

*******acsmedhhinc**
Median Household Income in Neighborhood Geography (*****).
American Community Survey 2008-2012 (5-year)

```
df[geo+'B19013001E'] / df[geo+'countrows']
```

Crime – Esri Crime Indexes 2010 & 2012

Esri Crime Indexes from 2010 & 2012
Please see **Appendix A** for further information regarding the database methodology of the CrimeRisk Index data.

Please note that a crime index value of '100' is considered the national average. The CrimeRisk Index variables provided for this project are the original index values and **have not** been scaled in any way.

In order to cut down on the size of the data dictionary and reduce redundancy for all variables, they will be preceded with a wildcard; $$.

****
[b0250m, b1000m] - prefix for buffer distance

$$ - [10, 12]
wildcard for year of Esri Crime Data

> **10**
> Esri Crime Indexes for 2010
>
> **12**
> Esri Crime Indexes for 2012

[awi, avg]
awi = areal weighted interpolation
avg = average index per intersected feature (no spatial weighting by area of census block group intersected by neighborhood geography)

*******cr$$---totc**
Average Total Crime Index in Neighborhood Geography (*****).

*******cr$$---perc**
Average Personal Crime Index in Neighborhood Geography (*****).

*******cr$$---murd**
Average Murder Index in Neighborhood Geography (*****).

*******cr$$---rape**
Average Rape Index in Neighborhood Geography (*****).

*******cr$$---robb**
Average Robbery Index in Neighborhood Geography (*****).

*******cr$$---asst**
Average Assault Index in Neighborhood Geography (*****).

*******cr$$---proc**
Average Property Crime Index in Neighborhood Geography (*****).

*******cr$$---burg**
Average Burglary Index in Neighborhood Geography (*****).

*******cr$$---larc**
Average Larceny Index in Neighborhood Geography (*****).

*******cr$$---mveh**
Average Motor Vehicle Theft Index in Neighborhood Geography (*****).

NETS business variables (1990-2010) by overall flag

Raw versus Collapsed

The NETS data was also split by both raw and collapsed counts of business flags. All NAD 83 UTM Zone 18 coordinates were rounded to the nearest 10 meters, thus imposing a grid and then given a rounded grid index value and then that value is used for the collapse. Meaning if two records were geocoded as gyms and both snapped to the same 10 meter grid index then it would show up as 2 records in the raw counts and only 1 record in the collapsed counts.

Retail Environment

The National Establishment Time-Series (NETS) database is a commercially purchased longitudinal file derived from Dun & Bradstreet's (D&B) register of business information from January of each year. NETS is considered one of the most comprehensive establishment sources available, serving as a census of American businesses. Businesses that we considered health-relevant and pertinent to current research interests were categorized into one of 25 mutually exclusive researcher-defined categories. An eight-digit Standard Industrial Classification (SIC) code [BEH_SIC] and, where applicable, the most recently reported company name [COMPANYHERE], trade name[TRADENAMEHERE], sales volume [SALESHERE], or employee count [EMPLOYEESHERE], were used to assign businesses to one researcher-defined category, uniform across all years. Definitions coded in SAS 9.4.

> *Note: BEH_SIC is an SIC value used to represent the business type across all years. If a business reported one primary SIC code for at least 75% of the years it was in operation, that majority SIC code was used to represent the business type across all years. The most recently reported primary SIC code was used to characterize the overall nature of businesses whose most frequently reported SIC was reported less than 75% of the time*

****

[b0250m, b1000m] - prefix for buffer distance in meters

b0250m
buffer 250 meter neighborhood geography unit.

b1000m
buffer 1000 meter neighborhood geography unit. (1 kilometer)

YYYY is a placeholder for year [1990-2010]
for year in range(1990, 2010) – wildcard for year of NETS data

^^^ is a placeholder for count type [raw, col], where:

raw - count of all businesses in specified category in a given year

col - count after collapsing businesses that belonged to the same researcher-defined category[1], in the same year, at the same location[2]

[1] We did not collapse businesses categorized as "other potential walking destinations" due to the heterogeneity of industries included in this category.

[2] Because address strings for the same location can vary greatly and because our three geocoders return matched addresses in different formats, the final prioritized decimal degrees coordinates were projected to UTM NAD 83 ZONE 18N (meters) and rounded to the nearest 10 meters. Local projected coordinate systems ensure accurate distance calculations.

Food Outlet Categories

*******nets^^^YYYYwar–** count of warehouse and discount department stores selling food

```
if INDEX(TRADENAMEHERE, "SAMS CLUB")>0  OR INDEX(COMPANYHERE, "SAMS CLUB") >0 or
INDEX(TRADENAMEHERE, "COSTCO WHOLESALE")>0  OR INDEX(COMPANYHERE, "COSTCO
WHOLESALE") >0
or TRADENAMEHERE="PRICE CLUB"  OR COMPANYHERE="PRICE CLUB"
or INDEX(TRADENAMEHERE, "BJS WHOLESALE")>0  OR INDEX(COMPANYHERE, "BJS WHOLESALE")
>0 OR COMPANYHERE="BJS"
```

Note: that 18 (warehouse) does not show up in the any of the 250m buffers.

*******nets^^^YYYYaff–** count of all fast food restaurants

```
if BEH_SIC in (58120300, 58120307, 58120308) OR (BEH_SIC in (58120000:58129999)
AND (INDEX(TRADENAMEHERE, "ARBYS")>0  OR INDEX(COMPANYHERE, "ARBYS") >0
OR INDEX(TRADENAMEHERE, "BASKIN ROBBINS")>0  OR INDEX(COMPANYHERE, "BASKIN
ROBBINS") >0 OR INDEX(TRADENAMEHERE, "BLIMPIE")>0  OR INDEX(COMPANYHERE,
"BLIMPIE") >0 OR INDEX(TRADENAMEHERE, "BOJANGLES")>0  OR INDEX(COMPANYHERE,
"BOJANGLES") >0 OR INDEX(TRADENAMEHERE, "BOSTON MARKET")>0  OR INDEX(COMPANYHERE,
"BOSTON MARKET") >0 OR INDEX(COMPANYHERE, "BURGER KING")>0 OR
INDEX(TRADENAMEHERE,"BURGER KING")>0 OR INDEX(TRADENAMEHERE, "CHECKERS")>0  OR
INDEX(COMPANYHERE, "CHECKERS") >0 OR INDEX(TRADENAMEHERE, "CHIPOTLE")>0  OR
INDEX(COMPANYHERE, "CHIPOTLE") >0
OR INDEX(TRADENAMEHERE, "CHURCHS CHICKEN")>0  OR INDEX(COMPANYHERE, "CHURCHS
CHICKEN") >0 OR INDEX(TRADENAMEHERE, "DAIRY QUEEN")>0  OR INDEX(COMPANYHERE,
"DAIRY QUEEN") >0 OR INDEX(TRADENAMEHERE, "DOMINOS PIZZA")>0  OR
INDEX(COMPANYHERE, "DOMINOS PIZZA") >0 OR INDEX(TRADENAMEHERE, "DUNKIN DONUTS")>0
OR INDEX(COMPANYHERE, "DUNKIN DONUTS") >0 OR INDEX(TRADENAMEHERE, "EL POLLO
LOCO")>0  OR INDEX(COMPANYHERE, "EL POLLO LOCO") >0 OR INDEX(TRADENAMEHERE,
"FUDDRUCKERS")>0  OR INDEX(COMPANYHERE, "FUDDRUCKERS") >0 OR
TRADENAMEHERE="HARDEES" OR COMPANYHERE="HARDEES" OR INDEX(TRADENAMEHERE, "JAMBA
JUICE")>0  OR INDEX(COMPANYHERE, "JAMBA JUICE") >0 OR INDEX(TRADENAMEHERE,
"KENTUCKY FRIED CHICKEN")>0 OR INDEX(COMPANYHERE, "KENTUCKY FRIED CHICKEN") >0 OR
TRADENAMEHERE="KFC" OR COMPANYHERE="KFC" OR INDEX(TRADENAMEHERE, "KRISPY KREME")>0
OR INDEX(COMPANYHERE, "KRISPY KREME") >0 OR INDEX(TRADENAMEHERE, "LONG JOHN
SILVERS")>0  OR INDEX(COMPANYHERE, "LONG JOHN SILVERS") >0 OR INDEX(TRADENAMEHERE,
"MCDONALDS")>0  OR INDEX(COMPANYHERE, "MCDONALDS") >0 OR INDEX(TRADENAMEHERE,
"PANDA EXPRESS")>0  OR INDEX(COMPANYHERE, "PANDA EXPRESS") >0 OR
INDEX(TRADENAMEHERE, "PAPA JOHN")>0  OR INDEX(COMPANYHERE, "PAPA JOHN") >0 OR
INDEX(TRADENAMEHERE, "PIZZA HUT")>0  OR INDEX(COMPANYHERE, "PIZZA HUT") >0 OR
INDEX(TRADENAMEHERE, "POPEYES")>0  OR INDEX(COMPANYHERE, "POPEYES") >0 OR
INDEX(TRADENAMEHERE, "QUIZNOS")>0  OR INDEX(COMPANYHERE, "QUIZNOS") >0 OR
INDEX(TRADENAMEHERE, "ROUND TABLE")>0  OR INDEX(COMPANYHERE, "ROUND TABLE") >0 OR
TRADENAMEHERE="SBARRO" OR COMPANYHERE="SBARRO" OR TRADENAMEHERE="SONIC" OR
COMPANYHERE="SONIC" OR INDEX(TRADENAMEHERE, "STARBUCKS")>0 OR INDEX(COMPANYHERE,
"STARBUCKS") >0  OR TRADENAMEHERE="SUBWAY"  OR COMPANYHERE="SUBWAY" OR
TRADENAMEHERE="SUBWAY SANDWICHES"  OR COMPANYHERE="SUBWAY SANDWICHES"
OR TRADENAMEHERE="SUBWAY RESTAURANT"  OR COMPANYHERE="SUBWAY RESTAURANT" OR
INDEX(COMPANYHERE, "TACO BELL")>0 OR INDEX(TRADENAMEHERE,"TACO BELL")>0 OR
INDEX(TRADENAMEHERE, "TCBY")>0  OR INDEX(COMPANYHERE, "TCBY") >0 OR
INDEX(TRADENAMEHERE, "WENDYS")>0  OR INDEX(COMPANYHERE, "WENDYS OLD") >0 OR
INDEX(TRADENAMEHERE, "WHITE CASTLE")>0  OR INDEX(COMPANYHERE, "WHITE CASTLE") >0
or INDEX(COMPANYHERE, "BEN & JERRYS")>0  OR INDEX(TRADENAMEHERE, "BEN & JERRYS")>0
or INDEX(COMPANYHERE, "BEN AND JERRYS")>0  OR INDEX(TRADENAMEHERE, "BEN AND
JERRYS")>0 OR INDEX(COMPANYHERE, " CARVEL ICE CREAM CAKES")>0 OR
INDEX(TRADENAMEHERE," CARVEL ICE CREAM CAKES")>0 OR INDEX(COMPANYHERE, "COLD STONE
CREAMERY")>0  OR INDEX(TRADENAMEHERE, "COLD STONE CREAMERY")>0 OR
INDEX(COMPANYHERE, "HAAGEN-DAZS")>0  OR  INDEX(TRADENAMEHERE, "HAAGEN-DAZS") >0 OR
INDEX(COMPANYHERE, "HAAGENDAZS")>0  OR INDEX(TRADENAMEHERE, "HAAGENDAZS") >0 OR
INDEX(COMPANYHERE, "HAAGEN DAZS")>0  OR INDEX(TRADENAMEHERE, "HAAGEN DAZS") >0 OR
INDEX(COMPANYHERE, "I CANT BELIEVE ITS YOGURT")>0  OR INDEX(TRADENAMEHERE, "I CANT
BELIEVE ITS YOGURT")>0 OR INDEX(COMPANYHERE, "LITTLE CAESARS PIZZA")>0  OR
INDEX(TRADENAMEHERE, "LITTLE CAESARS PIZZA") >0
OR INDEX(COMPANYHERE, "SCHLOTZSKYS DELI")>0  OR INDEX(TRADENAMEHERE, "SCHLOTZSKYS
DELI") >0 OR INDEX(TRADENAMEHERE,"AU BON PAIN")>0 OR INDEX(COMPANYHERE, "A B P
CORPORATION")>0 OR INDEX(COMPANYHERE, "AU BON PAIN")>0 OR INDEX(TRADENAMEHERE, "A
B P CORPORATION")>0 OR INDEX(TRADENAMEHERE,"AUNTIE ANNES")>0 OR INDEX(COMPANYHERE,
"AUNTIE ANNES")>0 OR INDEX(TRADENAMEHERE,"CHICKEN HOLIDAY")>0 OR
INDEX(COMPANYHERE, "CHICKEN HOLIDAY")>0 OR INDEX(TRADENAMEHERE,"COFFEE SHOP NORTH
CENTRAL HOSP")>0 OR INDEX(COMPANYHERE, "DIRECTORS METRO FOOD SERVICE")>0 OR
INDEX(TRADENAMEHERE,"COSI SANDWICH")>0 OR INDEX(COMPANYHERE, "COSI SANDWICH")>0 OR
INDEX(TRADENAMEHERE, "CROWN FRIED CHICKEN")>0 OR INDEX(COMPANYHERE, "CROWN FRIED
CHICKEN")>0 OR INDEX(TRADENAMEHERE,"EVERYTHING YOGURT & SALAD")>0 OR
INDEX(COMPANYHERE, "EVERYTHING YOGURT & SALAD")>0 OR INDEX(TRADENAMEHERE,"GRAYS
PAPAYA")>0 OR INDEX(COMPANYHERE, "GRAYS PAPAYA")>0 OR INDEX(TRADENAMEHERE,
"KENNEDY FRIED CHICKEN")>0 OR INDEX(COMPANYHERE, "KENNEDY FRIED CHICKEN")>0 OR
INDEX(TRADENAMEHERE, "KOSHER DELIGHT CORP")>0 OR INDEX(COMPANYHERE, "KOSHER
DELIGHT CORP")>0
OR INDEX(TRADENAMEHERE,"MANHATTAN BAGEL")>0 OR INDEX(COMPANYHERE, "MANHATTAN
BAGEL")>0 OR INDEX(TRADENAMEHERE,"METROPOLITAN DELI")>0 OR INDEX(COMPANYHERE,
"METROPOLITAN DELI")>0 OR INDEX(TRADENAMEHERE, "MIAMI SUBS & GRILL")>0 OR
INDEX(COMPANYHERE, "MIAMI SUBS CORPORATION")>0 OR INDEX(TRADENAMEHERE,"NATHANS
FAMOUS")>0 OR INDEX(COMPANYHERE, "NATHANS FAMOUS")>0 OR
```

```
INDEX(TRADENAMEHERE,"PAPAYA KING")>0 OR INDEX(COMPANYHERE, "PAPAYA KING")>0 OR
INDEX(TRADENAMEHERE,"PUDGIES FAMOUS CHICKEN")>0 OR INDEX(COMPANYHERE, "PUDGIES
FAMOUS CHICKEN")>0 OR INDEX(TRADENAMEHERE,"ROY ROGERS")>0 OR INDEX(COMPANYHERE,
"ROY ROGERS")>0 OR INDEX(TRADENAMEHERE, "EL POLLO SUPREMO")>0 OR
INDEX(COMPANYHERE, "SUPREME CHICKEN OF NEW JERSEY")>0 OR
INDEX(TRADENAMEHERE,"TOGOS")>0 OR INDEX(COMPANYHERE, "TOGOS")>0 OR
INDEX(TRADENAMEHERE,"XANDO")>0 OR INDEX(COMPANYHERE, "XANDO")>0 OR
INDEX(TRADENAMEHERE,"ZORN FAMOUS CHICKEN & RIBS")>0 OR INDEX(COMPANYHERE, "ZORN
FAMOUS CHICKEN & RIBS")>0 OR INDEX(TRADENAMEHERE,"ZORNS FAMOUS CHICKEN")>0 OR
INDEX(COMPANYHERE, "ZORNS FAMOUS CHICKEN")>0))
```

*******nets^^^YYYYpiz –** count of pizza restaurants

```
if (*****AFFYYYY_xxx =0 and (BEH_SIC in (58120600, 58120601, 58120602))) or
(*****AFFYYYY_xxx =0 and index(CompanyHere, "PIZZA")>0) or (*****AFFYYYY_xxx =0
and index(TradeNameHere, "PIZZA")>0) or (*****AFFYYYY_xxx =0 and
index(CompanyHere, "PIZZERIA")>0) or (*****AFFYYYY_xxx =0 and index(TradeNameHere,
"PIZZERIA")>0) and *****URGYYYY_xxx=0 and *****HPYYYY_xxx=0 and *****RESYYYY_xxx=0
and *****RXYYYY_xxx=0 and *****MHYYYY_xxx=0 and *****DDSYYYY_xxx=0
```

*******nets^^^YYYYeat –** count of other restaurants

```
if BEH_SIC in (58120000:58129999) and *****AFFYYYY_xxx=0 and *****PIZYYYY_xxx=0
```

*******nets^^^YYYYbak –** count of bakeries and candy/confectionary stores

```
if BEH_SIC in (54610000:54619999, 54419901, 54419902, 54419903, 54419905) and
*****AFFYYYY_xxx=0
```

*******nets^^^YYYYmet –** count of meat markets

```
if BEH_SIC in (54210200: 54210299)
```

*******nets^^^YYYYfvm –** count of fruit and vegetable markets

```
if BEH_SIC in (54310000:54319999)
```

*******nets^^^YYYYnat –** count of natural food markets and nut stores

```
if BEH_SIC in (54990100: 54990199, 54990900:54990999, 54993500:54993599, 54419904)
```

*******nets^^^YYYYfsh –** count of fish markets

```
if BEH_SIC in (54210100:54210199)
```

*******nets^^^YYYYcon –** count of convenience stores and small grocery stores

```
if BEH_SIC in (54110200:54110299, 54110300:54110399) or (BEH_SIC in
(54110000:54119999) and (EmpHere<5 and EmpHere ne .))
```

*******nets^^^YYYYsmk –** count of large supermarkets

```
if BEH_SIC in (54110000:54119999) and ((SalesHere>=2000000 and  SalesHere ne .) or
(EmpHere>=18 and EmpHere ne .))
```

Alcohol Outlet Categories

*******nets^^^YYYYbar –** count of bars and other public drinking places

```
if BEH_SIC in (58130000:58139999, 86410401)
```

*******nets^^^YYYYliq –** count of liquor stores

```
if BEH_SIC in (59200000:59299999) or (BEH_SIC in (51800000:51899999) and
index(CompanyHere,'LIQUOR') ge 1))
```

Medical Facility Categories

*******nets^^^YYYYurg –** count of urgent care and hospital facilities

```
if BEH_SIC in (80110200, 80110201, 80110204, 80620000:80629999, 80690000,
80690200, 80690201, 80690300, 80690301, 80699901, 80699902, 80699903, 80699904,
80699905) and *****WAREYYYY_xxx=0
```

*******nets^^^YYYYhpc –** count of offices or clinics of health practitioners

```
if BEH_SIC in  (80110000: 80110110, 80110202, 80110205, 80110500: 80119905,
80310000:80490201, 80499900, 80499902, 80499903, 80499904, 80499906, 80499908,
80499909, 80920000, 80930000, 80930200:80939901, 80939903, 80939905, 80990103,
80990104, 80990200, 80990201, 80990203, 80999905, 80999906, 80999907)
```

*****nets^^^YYYYres** – count of residential facilities with health care (e.g., nursing homes)

```
if BEH_SIC in (80510000, 80519900, 80519901, 80519902, 80520000, 80529900,
80529902, 80590000, 80599901, 80599904, 80599905, 80599906, 83610000, 83610300,
83610400:83610499, 83619900, 83619901, 83619904)
```

*****nets^^^YYYYrxp** – count of pharmacies

```
if BEH_SIC in (51220000, 51220300, 51229900, 59120000:59129902, 80110203) and
*****WAREYYYY_xxx=0
```

*****nets^^^YYYYmhc** – count of mental health care facilities

```
if BEH_SIC in (80110400, 80110401, 80110402, 80110403, 80490400, 80490401,
80490402, 80490403, 80490404, 80519903, 80529901, 80599903, 80630000, 80639900,
80639901, 80690100, 80690101, 80690102, 80930100, 80930101, 80930102, 80930103,
80939902, 83610302, 83610304, 83619902, 83619903, 83619905)
```

*****nets^^^YYYYdds** – count of dental care facilities

```
if BEH_SIC in (80210000, 80210100, 80210101, 80210102, 80210103, 80210104,
80210105, 80210106, 80210107, 80210108, 80210200, 80210201, 80210202, 80219901,
80219902, 80490501)
```

Physical Activity Facility Categories

*****nets^^^YYYYmul** – count of multi-use physical activity venues

```
if BEH_SIC in (70110306, 70110307, 79910000, 79910100, 79910101, 79910102,
79910300, 79910302, 79970000, 79991127, 79999910) or INDEX(COMPANYHERE,'JCC OF')>0
or INDEX(COMPANYHERE,'JEWISH CMMN CTR')>0 or INDEX(COMPANYHERE,'JEWISH CMMNTY
CTR')>0 or INDEX(COMPANYHERE,'JEWISH CMNTY CNTRE')>0 or INDEX(COMPANYHERE,'JEWISH
CMNTY CTR')>0 or INDEX(COMPANYHERE,'JEWISH CMNTY HSE')>0 or
INDEX(COMPANYHERE,'JEWISH CMTY CENTER')>0 or INDEX(COMPANYHERE,'JEWISH COMMNTY
CTR')>0 or INDEX(COMPANYHERE,'JEWISH COMMUNITY CENTER')>0 or
INDEX(COMPANYHERE,'JEWISH COMMUNITY CENTRE')>0 or
INDEX(COMPANYHERE,'JEWISH COMMUNITY CTR')>0 or INDEX(COMPANYHERE,'JEWISH CMNTY
CTR')>0 or INDEX(COMPANYHERE,'JEWISH COMMUNTY CENTER')>0 or
INDEX(COMPANYHERE,'Y M & Y M H A')>0 or INDEX(COMPANYHERE,'Y M C A')>0 or
INDEX(COMPANYHERE,'Y W C A')>0 or INDEX(COMPANYHERE,'YM C A')>0 or
INDEX(COMPANYHERE,'YMHA')>0 or INDEX(COMPANYHERE,'YWHA')>0 or
INDEX(COMPANYHERE,'YMCA')>0 or INDEX(COMPANYHERE,'YMWCA')>0 or
INDEX(COMPANYHERE,'YWCA')>0 or INDEX(COMPANYHERE,'YNG MENS & YNG WOMENS HEBREW')>0
or INDEX(COMPANYHERE,'YOUNG MEN CHRISTIAN ASSOCIATON')>0 or
INDEX(COMPANYHERE,'YOUNG MEN YUNG WNS HBREW ASSN')>0 or INDEX(COMPANYHERE,'YOUNG
MEN YUNG WNS HEBREW ASSN')>0 or INDEX(COMPANYHERE,'YOUNG MENS & WOMENS ASSC')>0 or
INDEX(COMPANYHERE,'YOUNG MENS & WOMENS CHRISTIA')>0 or INDEX(COMPANYHERE,'YOUNG
MENS & YOUNG WOMENS')>0 or INDEX(COMPANYHERE,'YOUNG MENS AND YOUNG WOMENS')>0 or
INDEX(COMPANYHERE,'YOUNG MENS CHRISTIAN')>0 or INDEX(COMPANYHERE,'YOUNG MENS
CHRSTN ASSOC')>0 or INDEX(COMPANYHERE,'YOUNG MENS HEBREW ASSOCIATION')>0 or
INDEX(COMPANYHERE,'YOUNG MENS YNG WMNS HBRW ASSN')>0 or
INDEX(COMPANYHERE,'YOUNG MENS/YOUNG WOMENS HEBREW')>0 or INDEX(COMPANYHERE,'YOUNG
MNS CHRISTN ASSN')>0 or INDEX(COMPANYHERE,'YOUNG MNS CHRSTN ASSN')>0 or
INDEX(COMPANYHERE,'YOUNG MNS YUNG WNS HEBREW ASSN')>0 or INDEX(COMPANYHERE,'YOUNG
WNS CHRISTN ASSN')>0 or INDEX(COMPANYHERE,'YOUNG WOMANS CHRISTIAN ASSOC')>0 or
INDEX(COMPANYHERE,'YOUNG WOMENS CHRISTIAN')>0 or INDEX(COMPANYHERE,'YWC ASSOC')>0
or index(TradeNameHERE,'J C C')>0 or index(TradeNameHERE,'JCC')>0 or
index(TradeNameHERE,'JEWISH CMNTY CNTR')>0 or index(TradeNameHERE,'JEWISH CMNTY
CTR')>0 or index(TradeNameHERE,'JEWISH COMMUNITY CENTER')>0 or
index(TradeNameHERE,'JEWISH COMMUNITY CNTR')>0 or index(TradeNameHERE,'Y M C A')>0
or
index(TradeNameHERE,'Y W C A')>0 or index(TradeNameHERE,'YM & YW H A')>0 or
index(TradeNameHERE,'YWHA')>0 or index(TradeNameHERE,'YMCA')>0 or
index(TradeNameHERE,'YMWCA')>0 or index(TradeNameHERE,'YOUNG MENS CHRISTIAN
ASSOCIAT')>0 or index(TradeNameHERE,'YOUNG MNS YUNG WNS HEBREW ASSN')>0 or
index(TradeNameHERE,'YOUNG WOMENS CHRISTIAN ASSN')>0 or index(TradeNameHERE,'YOUNG
WOMENS CHRISTIAN ASSOC')>0 or index(TradeNameHERE,'YWCA')>0  and
*****URGYYYY_xxx=0 and *****HPYYYY_xxx=0 and *****RESYYYY_xxx=0 and
*****RXYYYY_xxx=0 and *****MHYYYY_xxx=0 and *****DDSYYYY_xxx=0
```

*****nets^^^YYYYlmp** – count of light/moderate physical activity venues

```
if (*****MULYYYY_xxx=0 and (BEH_SIC in (39490103, 70320301, 79110000, 79110100,
79110101, 79110200, 79110201, 79110202, 79110203, 79110204, 79330000, 79339903,
79920000, 79970101, 79970200, 79970201, 79970202, 79970204, 79970300, 79970301,
79970302, 79979904, 79979906, 79979908, 79990200, 79990202, 79990203, 79990204,
```

```
79990205, 79990601, 79990700, 79990701, 79990702, 79990703, 79990704, 79990705,
79991102, 79991104, 79991109, 79991115, 79991121, 79991123, 79991200, 79991201,
79991202, 79991205, 79991400, 79991402, 79991409, 79991411, 79991512, 79991601,
79991602, 79999903, 82999903)))
```

*******nets^^^YYYYvpa –** count of vigorous physical activity venues

```
if (*****MULYYYY_xxx=0 and (BEH_SIC in (70110201, 79410104, 79410201, 79910301,
79970100, 79970102,  79970203, 79970402, 79970403, 79970500, 79970502, 79970503,
79970504, 79990100, 79990101, 79990102, 79990300, 79990301, 79990302, 79990303,
79990501, 79990600, 79990602, 79990603, 79991103, 79991107, 79991110, 79991111,
79991112, 79991113, 79991116, 79991118, 79991119, 79991120, 79991122, 79991412,
82999914)))
```

Other Retail Categories

*******nets^^^YYYYbnk –** count of banks

```
if BEH_SIC in (60210000, 60219900, 60219901, 60220000, 60229900, 60229901,
60290000, 60350000, 60359900, 60359901, 60359902, 60360000, 60369900, 60369901,
60369902)
```

*******nets^^^YYYYcrd –** count of credit unions

```
if BEH_SIC in  (60610000,60620000,60629900,60629901)
```

*******nets^^^YYYYdes –** count of other potential destinations

```
if (BEH_SIC in (53000000:53999999) and (EmpHere<250 and EmpHere ne .))
or (BEH_SIC in (54000000:54999999) and (EmpHere<250 and EmpHere ne .))
or (BEH_SIC in (58000000:58999999) and (EmpHere<250 and EmpHere ne .))
or (BEH_SIC in (59000000:59999999) and (EmpHere<250 and EmpHere ne .))
or (BEH_SIC in (78000000:78999999) and (EmpHere<250 and EmpHere ne .))
or (BEH_SIC in (83000000:83999999) and (EmpHere<250 and EmpHere ne .))
or (BEH_SIC in (84000000:84999999) and (EmpHere<250 and EmpHere ne .))
and *****BARYYYY_xxx=0 and *****LIQYYYY_xxx=0 and *****FSHYYYY_xxx=0 and
*****FVMYYYY_xxx=0 and *****NATYYYY_xxx=0 and *****METYYYY_xxx=0 and
*****SMKYYYY_xxx=0 and *****EATYYYY_xxx=0 and *****CONYYYY_xxx=0 and
*****AFFYYYY_xxx=0 and *****PIZYYYY_xxx=0 and *****BAKYYYY_xxx=0 and
*****BNKYYYY_xxx=0 and *****CRDYYYY_xxx=0 and *****MULYYYY_xxx=0 and
*****LMPAYYYY_xxx=0 and *****VPAYYYY_xxx=0 and *****WAREYYYY_xxx=0 and
*****URGYYYY_xxx=0 and *****HPYYYY_xxx=0 and *****RESYYYY_xxx=0 and
*****RXYYYY_xxx=0 and *****MHYYYY_xxx=0 and *****DDSYYYY_xxx=0
```

*******nets^^^YYYYnot –** count of businesses that do not belong to any of the researcher-defined categories

```
if *****BARYYYY_xxx=0 and *****LIQYYYY_xxx=0 and *****FSHYYYY_xxx=0 and
*****FVMYYYY_xxx=0 and *****NATYYYY_xxx=0 and *****METYYYY_xxx=0 and
*****SMKYYYY_xxx=0 and *****EATYYYY_xxx=0 and *****CONYYYY_xxx=0 and
*****AFFYYYY_xxx=0 and *****PIZYYYY_xxx=0 and *****BAKYYYY_xxx=0 and
*****BNKYYYY_xxx=0 and *****CRDYYYY_xxx=0 and *****MULYYYY_xxx=0 and
*****LMPAYYYY_xxx=0 and *****VPAYYYY_xxx=0 and *****WAREYYYY_xxx=0 and
*****URGYYYY_xxx=0 and *****HPYYYY_xxx=0 and *****RESYYYY_xxx=0 and
*****RXYYYY_xxx=0 and *****MHYYYY_xxx=0 and *****DDSYYYY_xxx=0 and
*****DESYYYY_xxx=0
```

☐ Checked each variable and fixed overlaps using the following prioritizations:

- Medical Facility Categories [URG, HP, RES, RX, MH, DDS] > Pizza [PIZ]
- All Fast Food [AFF] > Pizza [PIZ]
- All Fast Food [AFF] > Other restaurants [EAT]
- All Fast Food [AFF] > Other restaurants [EAT]
- All Fast Food [AFF] > Bakeries and candy/confectionary stores [BAK]
- Warehouse and discount department stores selling food [WARE] > Urgent care and hospital facilities [URG]
- Warehouse and discount department stores selling food [WARE] > Pharmacies [RX]
- Multi-use physical activity venues [MUL] > Light/moderate physical activity venues [LMPA]
- Multi-use physical activity venues [MUL] > Vigourous physical activity venues [VPA]
- Alcohol Outlet Categories [BAR, LIQ] + Food Outlet Categories [+ Medical Facility Categories [URG, HP, RES, RX, MH, DDS] + Physical Activity Facility Categories [WARE, AFF, PIZ, EAT, BAK, MET, FVM, NAT, FSH, CON, SMK] + Banks [BNK] + Credit Unions [CRD] > Other potential destinations [DES]

Street View Variables
The following Street View variables were generated using a 100 meter by 100 meter grid created using kriging from the Street View online audit tool developed by Steve Mooney of BEH.

All values were intersected with the neighborhood geometry and averaged (divided by count of grid points) similarly to Zonal Statistics on a raster grid.

Street View Disorder
Kriging parameters for disorder are:

City	Nugget	Range	Partial Sill
NYC	0.14	5000	0.19
Philadelphia	0.35	10000	0.30
Detroit	0.18	3000	0.52
San Jose	0.09	1000	0.27

Neighborhood physical disorder, or the deterioration of an urban landscape, is reflected in indicators suchas litter, graffiti, and abandoned buildings. A physical disorder surface was constructed from observations of nine indicators of disorder (i.e., bottles, litter, graffiti, poorly maintained buildings, burned out buildings, abandoned buildings, abandoned cars, vacant lots, and bars on windows) on 532 block faces in NYC using Google Street View imagery that dated from August 2007 to October 2011. The nine indicators were combined into a scale using Item Response Theory, and ordinary kriging with an exponential variogram function with nugget = 0.19, range= 5000 and partial sill = 0.14 was used to estimate physical disorder levels at any point throughout the city.

For each subject, 11 neighborhood disorder variables were constructed for each buffer type; one from each of 10 kriged conditional realizations and one from the kriged point estimate ('value'). Because kriging presents over-smoothed estimates, when estimating associations with disorder, rather than use the point estimate, the 10 conditional realizations should be used as imputations within a multiple imputation framework. This will mitigate the over-smoothing issue and result in associational estimates that account appropriately for uncertainty and measurement error both in the point disorder measure and the interpolation. The simulations are conditional realizations based on the kriging parameters, observed values, and variance estimates – essentially, the conditional realizations use estimates that sample from the posterior distribution rather than select the peak of that distribution.

****
[b0250m, b1000m] - prefix for buffer distance in meters

> **b0250m**
> buffer 250 meter neighborhood geography unit.
>
> **b1000m**
> buffer 1000 meter neighborhood geography unit. (1 kilometer)

*******svdisvalue**
Street View Disorder Mean Value

*******svdisvariance**
Street View Disorder Mean Variance

*******svdissim1**
Street View Disorder Mean Simulation 1

*****svdissim2**
Street View Disorder Mean Simulation 2

*****svdissim3**
Street View Disorder Mean Simulation 3

*****svdissim4**
Street View Disorder Mean Simulation 4

*****svdissim5**
Street View Disorder Mean Simulation 5

*****svdissim6**
Street View Disorder Mean Simulation 6

*****svdissim7**
Street View Disorder Mean Simulation 7

*****svdissim8**
Street View Disorder Mean Simulation 8

*****svdissim9**
Street View Disorder Mean Simulation 9

*****svdissim10**
Street View Disorder Mean Simulation 10

*****svdiscount**
Street View Disorder Count of grid points

Street View Pedestrian Safety

The pedestrian safety measure was constructed from 6 indicators of investment to keep pedestrians safe from traffic. The four major indicators are (1) presence of crosswalks, (2) presence of curb cuts, (3) completeness of sidewalk and (4) presence of street trees between the sidewalk and the street. Road and sidewalk condition measures also contribute slightly to the scale. We assessed this measure on 532 block faces in NYC, 503 block faces in Philadelphia, 502 block faces in Detroit and 289 block faces in San Jose using Google Street View imagery that dated from July 2007 to October 2011. The six indicators were combined into a scale using Item Response Theory, and ordinary kriging with an exponential variogram function with was used to estimate pedestrian safety from traffic levels at any point throughout each city. Kriging parameters were nugget = 0.20, range= 8000 and partial sill = 0.08 for NYC; nugget = 0.22, range= 5000 and partial sill = 0.02 for Philadelphia; nugget = 0.22, range= 5000 and partial sill = 0.09 for Detroit; and nugget = 0.19, range= 1000 and partial sill = 0.08 for San Jose.

For each subject, 11 pedestrian safety from traffic variables were constructed for each buffer type; one from each of 10 kriged conditional realizations and one from the kriged point estimate ('value'). Because kriging presents over-smoothed estimates, when estimating associations with pedestrian safety, rather than use the point estimate, the 10 conditional realizations should be used as imputations within a multiple imputation framework. This will mitigate the over-smoothing issue and result in associational estimates that account appropriately for uncertainty and measurement error both in the point pedestrian safety measure and the interpolation. The simulations are conditional realizations based on the kriging parameters, observed values, and variance estimates –essentially, the conditional realizations use estimates that sample from the posterior distribution rather than select the peak of that distribution.

Street View Pedestrian Safety Measure Caution

The pedestrian safety infrastructure scale was created using item response theory, which assumes a uni-dimensional latent characteristic is indicated by observed items. This may not be appropriate for pedestrian safety infrastructure, which is usually developed as a result of a complex interaction of factors, including prevalence of walking destinations, municipal priorities, concepts of desired neighborhood character, neighborhood efficacy, etc. Additionally, because infrastructure investment decisions are multifactorial and tems assessed by this measure are micro-scale, they may be strongly affected by street type randomly selected for assessment. However, street type was not a component of the kriging process used to spatially interpret the measure.

*******svpedvalue**
Street View Pedestrian Safety Mean Value

*******svpedvariance**
Street View Pedestrian Safety Mean Variance

*******svpedsim1**
Street View Pedestrian Safety Mean Simulation 1

*******svpedsim2**
Street View Pedestrian Safety Mean Simulation 2

*******svpedsim3**
Street View Pedestrian Safety Mean Simulation 3

*******svpedsim4**
Street View Pedestrian Safety Mean Simulation 4

*******svpedsim5**
Street View Pedestrian Safety Mean Simulation 5

*******svpedsim6**
Street View Pedestrian Safety Mean Simulation 6

*******svpedsim7**
Street View Pedestrian Safety Mean Simulation 7

*******svpedsim8**
Street View Pedestrian Safety Mean Simulation 8

*******svpedsim9**
Street View Pedestrian Safety Mean Simulation 9

*******svpedsim10**
Street View Pedestrian Safety Mean Simulation 10

*******svpedcount**
Street View Pedestrian Safety Count of grid points

Appendix A: CrimeRisk Database Methodology Guide

Content

CrimeRisk is a block group and higher level geographic database consisting of a series of standardized indexes for a range of serious crimes against both persons and property. It is derived from an extensive analysis of several years of crime reports from the vast majority of law enforcement jurisdictions nationwide. The crimes included in the database are the "Part 1" crimes and include murder, rape, robbery, assault, burglary, theft, and motor vehicle theft. These categories are the primary reporting categories used by the FBI in its Uniform Crime Report (UCR), with the exception of Arson, for which data is very inconsistently reported at the jurisdictional level. Part II crimes are not reported in the detail databases and are generally available only for selected areas or at high levels of geography.

In accordance with the reporting procedures using in the UCR reports, aggregate indexes have been prepared for personal and property crimes separately, as well as a total index. While this provides a useful measure of the relative "overall" crime rate in an area, it must be recognized that these are unweighted indexes, in that a murder is weighted no more heavily than a purse snatching in the computation. For this reason, caution is advised when using any of the aggregate index values.

Methodology

The primary source of CrimeRisk was a careful compilation and analysis of the FBI Uniform Crime Report databases. On an annual basis, the FBI collects data from each of about 16,000 separate law enforcement jurisdictions at the city, county, and state levels and compiles these into its annual Uniform Crime Report (UCR). The latest national crime report can be obtained either from the FBI web site in Adobe Portable Document (PDF) format or can be ordered directly from the FBI. While useful, the UCR provides detailed data only for the largest cities, counties, and metropolitan areas.

The original analysis was undertaken by obtaining detailed jurisdictional level data for the years 1990 through 1996, which were supplemented with 1999 preliminary UCR statistics at the State level and for cities and metropolitan areas where those have been released. We are now using UCR data from 1998-2006. The preliminary 2007 release data was used to balance the models to the latest available data.

A considerable effort was made to correct a number of problems that are prevalent within the FBI databases, including:

- The standardization of jurisdictional names: the FBI does not employ Census bureau codes in its databases and the jurisdictional names contain numerous typographical errors and format discrepancies which needed to be manually corrected

- Reporting by individual jurisdictions can be inconsistent from year to year, in that data for some jurisdictions is missing for one or more years and required handling

- Reporting for some crime types is inconsistent between jurisdictions. The FBI handles this by simply suppressing the statistics entirely for those areas. This primarily affects the rape category for Illinois, where statistics are suppressed for all but the largest jurisdictions. These missing values were handled via the modeling process, in which rape estimates were prepared for these jurisdictions by using a model which related rape incidence to other crime types

- The standardization of the database to account for jurisdictional overlaps. For example, the California Highway

Patrol has jurisdiction over only state and interstate highways in urban areas.

- Crime rates in general have been declining over the past several years, so it was necessary to adjust the historical data to reflect current crime rates.

Once this correction and standardization effort was completed, the database consisted of a time series of six years of data covering:

- All cities and towns which have their own police agency
- All cities and towns where policing for the local jurisdiction is contracted to a higher level agency but which tracks statistics separately (e.g. the city of Thousand Oaks, California contracts with the Ventura County Sheriff's Department for police services, but the incident reports are separately compiled)
- A record for each county which covers the population not covered by either of the two cases above. This is normally either a County Sheriff (or equivalent) or a State level jurisdiction which reports incidence of crime by county (e.g. in New York, the State Trooper).

For a very limited number of areas, such as New York City, the local jurisdiction spans several counties.

The initial models were undertaken using a subset of this database. In the smallest cities, a single murder will have a profound effect on the crime rate per 100,000 population that would severely distort the resulting models. Cities with less than 2,500 people were reassigned to their parent counties for the purpose of the analysis. A wide range of 1990 Census and current year demographic attributes was extracted from AGS' databases for the remaining areas (approximately 8,500 separate "jurisdictions"). This database was then used as the primary modeling database and was used later for scaling purposes.

Each of the seven crime types was modeled separately, using an initial range of about 65 socio-economic characteristics taken from the 2000 Census and AGS' current year estimates. Separate models were constructed for each of the nine Census regions (e.g. New England, East North Central, Pacific) in order to account for regional differences in crime rates and the demographic characteristics which underlay them. The models constructed typically accounted for over 85% of the variance in crime rates at this "jurisdiction" level, although it should be noted that the results for property crimes were generally more reliable than for personal crimes.

The results of these models were then applied to the block group level using the same demographic attributes compiled at the block group level. The resulting estimates were then scaled to match the master database of 8,500 jurisdictions. For cities, the block groups within each city were scaled to match the city total. For areas outside of these cities (or for smaller centers), results were scaled to match the county total after adjusting for those cities scaled separately.

The final crime rate estimates were then weighted by population and aggregated to the national totals. The results were then scaled to match the 2007 preliminary estimates (at a state level) and converted to indexes relative to the national total.

Source: Applied Geographic Solutions, Inc. v2, May 2010

Reference:

Mooney, Stephen J., Michael DM Bader, Gina S. Lovasi, Kathryn M. Neckerman, Julien O. Teitler, and Andrew G. Rundle. "Validity of an Ecometric Neighborhood Physical Disorder Measure Constructed by Virtual Street Audit." *American Journal of Epidemiology* 180, no. 6 (2014): 626-635.