

Michigan Center on the Demography of Aging  
(MiCDA) Geographic Linkages Repository (GLR)

Current Holdings

August 22, 2023

## About the Geographic Linkages Repository

MiCDA is partnering with the Health and Retirement Study (HRS), Panel Study of Income Dynamics (PSID) and National Health and Aging Trends Study/National Study of Caregiving (NHATS/NSOC) to make contextual files and related documentation available in the MiCDA Enclave to approved researchers upon request.

The Geographic Linkages Repository (GLR) facilitates linkages at multiple levels of geography (e.g. state, county, tract, zip code) so that researchers can more fully investigate the impact of the environments in which people live and work on later life health, well-being and related inequities.

## Citing MiCDA Geographic Linkages Repository Data Sources

By requesting to use data deposited with GLR, you agree to cite any GLR data sources used in your publications and presentations. A list of citations for current holdings in the GLR can be found at [micda.isr.umich.edu/micda/linkages](http://micda.isr.umich.edu/micda/linkages).

## Datasets in the GLR

**Contextual Data Resource (CDR).** The Contextual Data Resource (CDR) is a collection of datasets designed to facilitate research on the role of place in shaping health and well-being. CDR includes measures of health care resources, demographic and socioeconomic data, air quality, crime and the food environment. Measures are available at various levels of geography (e.g., census tract, county, metro area, state). More detailed information can be found on the [CDR webpage](#). Subfolders include:

**The Area Health Resources Files (AHRF)** - The Area Health Resources Files (AHRF) dataset includes information on the number of health care providers by profession and various types of health care facilities derived from medical associations and professional groups. The health professions data provide information on the supply and demographics of health profession such as primary care physicians, total MDs, doctors, dentists, and nurses. The health facilities data include information on types of health facilities and resources such as services providers, personnel, and beds. Data are longitudinal design, and available by county and year. Years covered include 2000-2019 and vary by dataset.

**Decennial Census and ACS** - The Decennial Census and American Community Survey (ACS) contain demographic and socioeconomic data on the US population and housing. The Census and ACS are collected by the U.S. Census Bureau. The HRS-CDR provides Census data for 1990, 2000, and 2010, and ACS data for 2008-12 and 2012-16 at the State, Metropolitan Statistical Area (MSA), County, and Census Tract levels of analysis.

**EPA FAQSD O3 Fused Air Quality** - The NO<sub>2</sub> land-use regression model estimate data includes national-scale estimates of NO<sub>2</sub> in the United States. It provides predictions for annual average NO<sub>2</sub> concentrations (ppb) using the land-use regression models. In order to predict NO<sub>2</sub> estimates at the national level, the following input data were used: Fixed-site regulatory monitors, Satellite-derived NO<sub>2</sub> estimates and GIS-derived land-use data.

**EPA FAQSD PM 2.5 Fused Air Quality** - The Fused Air Quality Surface Using Downscaling (FAQSD) Files is derived from data on Ozone (O<sub>3</sub>) and Particulate Matter 2.5 (PM<sub>2.5</sub>). The data provides predictions for national-scale annual average concentration of

O<sub>3</sub> (µg/m<sup>3</sup>) and PM<sub>2.5</sub> (µg/m<sup>3</sup>) in the United States using monitoring data and Community Multiscale Air Quality (CMAQ) output.

**Nitrogen Dioxide LUR Model Estimates** - The nitrogen dioxide (NO<sub>2</sub>) land-use regression model estimate data includes national-scale estimates of NO<sub>2</sub> in the United States. It provides predictions for annual average NO<sub>2</sub> concentrations (ppb) using the land-use regression models. In order to predict NO<sub>2</sub> estimates at the national level, the following input data were used: fixed-site regulatory monitors, satellite-derived NO<sub>2</sub> estimates, and GIS-derived land-use data. Data are available at census tract level and organized by time and cover the years 2000-2010.

**Uniform Crime Reporting** - The Uniform Crime Reports (UCR) data includes annual counts of offenses including murder, rape, robbery, aggravated assault, burglary, larceny, auto theft, and arson. The crime counts are available at the county level and organized by year, from 1994 to 2014 and 2016. The UCR data is collected from local law enforcement agencies through the Federal Bureau of Investigation (FBI) Uniform Crime Reporting Program and then aggregated, imputed, and distributed by the National Archive of Criminal Justice Data (NACJD).

**USDA Food Environment** - The USDA Food Environment and Access dataset provides information on geographic variation in access and proximity to grocery stores, restaurant availability and expenditures, food insecurity and assistance, food prices and taxes, and other food-related measures, as well as fine-grained data on food access for various demographic groups. The USDA dataset is derived from two data files: the Food Environment Atlas, which covers states and counties and spans the years from 2000 through 2016, and the Food Access Research Atlas, with measures from 2010 for census tracts.

### **LTC focus database**

The Long Term Care Focus (LTCFocus) datasets include data regarding the health and functional status of nursing home residents, characteristics of care facilities, and data characterizing the markets in which facilities exist. Compiled data are gathered from a variety of primary and secondary sources, including MDS, OSCAR, and other sources that characterize the policy environment and local market forces affecting nursing home providers at the state and county level. Most measures are available annually from 2000 to 2020. Visit [lctfocus.org](http://lctfocus.org) for additional information.

### **National Neighborhood Data Archive (NaNDA)**

NaNDA is a nation-wide collection of data with measures of the demographic, economic, social and physical environments, generally at the level of Zip Code Tabulation Areas (ZCTA) and Census Tract. A Zip Code to ZCTA crosswalk is included. More detailed information can be found on the [NaNDA webpage](#) and at [OpenICPSR](#). The individual datasets are listed below.

**Arts, Entertainment, and Recreation Organizations** - These datasets contain measures of the number and per capita density of select types of arts, entertainment, and recreation organizations—such as museums, libraries, spectator sports organizations, amusement parks, fitness centers, bowling alleys, and casinos per United States census tract or ZCTA from 2003 through 2017.

**Broadband Internet Availability, Speed, and Adoption** - These datasets contain measures of broadband internet access and usage per United census tract or ZCTA in 2014 through 2020. The data is derived primarily from internet service providers' Form 477

reports to the Federal Communications Commission. Key variables include the average upload and download speed of fixed broadband connections, the number of internet service providers, and the number of households with broadband.

**Crimes** - This dataset contains county-level totals for the years 2002-2014 for eight types of crime: murder, rape, robbery, aggravated assault, burglary, larceny, motor vehicle theft, and arson. These crimes are classed as Part I criminal offenses by the United States Federal Bureau of Investigations (FBI) in their Uniform Crime Reporting (UCR) program. Each record in the dataset represents the total of each type of criminal offense reported in (or, in the case of missing data, attributed to) the county in a given year.

**Dollar Stores** - These datasets contain measures of the number and density of dollar stores per United States census tract or ZCTA from 2003 through 2017.

**Eating and Drinking Places** - These datasets contain measures of the number and density of eating and drinking places – including fast food restaurants, coffee shops, and bars – per census tract or ZCTA in the United States from 2003 through 2017.

**Education and Training Services** - These datasets contain measures of the number and per capita density of education and training services per United States census tract or ZCTA from 2003 through 2017. This includes traditional education establishments such as elementary schools, secondary schools, and colleges, as well as businesses offering specialized training such as art classes, driving instruction, computer training, and standardized test preparation.

**Grocery Stores** - These datasets contain measures of the number and density of grocery stores – including supermarkets, specialty food stores, and warehouse clubs – per United States census tract or ZCTA from 2003 through 2017. These types of businesses represent places where neighborhood residents can obtain fresh and healthy foods.

**Health Care Services** - These datasets describe the number and density of health care services in each census tract or ZCTA in the United States from 2003 to 2017. The data includes counts, per capita densities, and area densities for many types of businesses in the health care sector, including doctors, dentists, mental health providers, nursing homes, and pharmacies.

**Internet Access** - These datasets contain measures of internet access per United States census tract or ZCTA from the 2015-2019 American Community Survey five-year estimate. Key variables include the number and percent of households with any type of internet subscription, with broadband internet, and with a computer or smartphone.

**Land Cover** - These datasets contain measures of land cover (e.g., low-, medium-, or high-density development, forest, wetland, open water) derived from the National Land Cover Database (NLCD) and aggregated by US census tract or ZCTA. Land cover is measured both in total square meters and as a proportion of all land within the tract or ZCTA.

**Law Enforcement Organizations** - These datasets contain measures of the number and per capita density of law enforcement and safety organizations—such as police and fire departments, courts, jails, and lawyers—per United States census tract or ZCTA from 2003 through 2017.

**Liquor, Tobacco, and Convenience Stores** - These datasets contain measures of the number and density of liquor, tobacco, and convenience stores per United States census tract or ZCTA from 2003 through 2017.

**Neighborhood-School Gap** - These datasets contain measures of neighborhood-school gap for 2009-2010 and 2015-2016. Neighborhood-school gap (NS gap) refers to the discrepancy between the demographics of a public school and its surrounding community.

For example, if 60% of a school's student body is Black, but 30% of the neighborhood population is Black, the school has a positive Black neighborhood-school gap. The datasets measure gaps in race and poverty between elementary school student populations and the census tracts or ZCTAs that those elementary schools serve. Supplemental data containing component variables used to calculate NS gap at the school and block group level is also available.

**Parks** - These datasets describe the number and area of parks in each census tract or ZCTA in the United States. Measures include the total number of parks, park area, and proportion of park area within each census tract or ZCTA.

**Personal Care Services and Laundromats** - These datasets contain measures of the number and per capita density of personal care services (such as barber shops, hair and nail salons, and spas) and laundromats per United States census tract or ZCTA from 2003 through 2017.

**Polluting Sites** - These datasets contain counts of polluting sites in each United States census tract or ZCTA and within a 0.5-mile buffer to capture spillover effects. Polluting sites are taken from the US Environmental Protection Agency's (EPA) Toxics Release Inventory. These facilities are typically larger and involved in manufacturing, metal mining, electric power generation, chemical manufacturing, and hazardous waste treatment.

**Post Offices and Banks** – These datasets contain measures of the number and per capita density of post offices and banks per United States census tract or ZCTA from 2003 through 2017.

**Primary and Secondary Roads** - These datasets contain measures of primary and secondary roads (highways and main arteries) per United States census tract or ZCTA in 2010. These measures may be used as a proxy for heavy traffic, high traffic speeds, and impediments to walking or biking. Variables include counts of primary, secondary, and all streets per tract or ZCTA; total length of primary, secondary, and all streets per tract or ZCTA; ratio of primary and/or secondary road counts to all roads; and ratio of length of primary/secondary roads to all streets.

**Public Transit Stops** - These datasets list the number of public transit stops per United States census tract or ZCTA based on data from the National Transit Map (NTM). Each observation represents the count and density (per capita and square mile) of transit stops within a census tract or ZCTA, as voluntarily reported to NTM between 2016 and 2018 by one of 270 regional transit agencies choosing to participate.

**Religious, Civic, and Social Organizations** - These datasets contain measures of the number and per capita density of select types of religious, civic, and social organizations – such as churches, mosques, synagogues, ethnic associations, and veterans' associations – per United States census tract or ZCTA from 2003 through 2017.

**Retail Establishments** - These datasets contain measures of the number and per capita density of select types of retail establishments—such as clothing, department, building and garden, furniture, and thrift stores—per United States census tract or ZCTA from 2003 through 2017.

**School Counts and Characteristics** - These datasets contain data on schools and school districts by district, census tract or ZCTA within the United States from 2000 through 2018. Key variables include district-level enrollment by race and ethnicity; numbers of teachers and counselors; teacher-student ratios; counts of public, private, and charter schools within districts; and expenditures and revenue, including per-pupil revenue.

**SES and Demography** - These datasets contain measures of socioeconomic and demographic characteristics by US census tract or ZCTA for the years 2008-2017. Example

measures include population density; population distribution by race, ethnicity, age, and income; and proportion of population living below the poverty level, receiving public assistance, and female-headed families. The dataset also contains a set of index variables to represent neighborhood disadvantage and affluence.

**Social Services** - These datasets contain measures of the number and per capita density of social services—such as senior centers, youth centers, food banks, job training programs, and day care centers—per United States census tract or ZCTA from 2003 through 2017.

**Street Connectivity** - These datasets contain measures of street connectivity (how well streets connect with one another) within all United States census tracts or ZCTAs. This includes measures of the number of street segments (links) and intersections (nodes) per tract, street length within tracts, and indices representing overall connectivity within the tract or ZCTA.

**Traffic Volume** - These datasets contain measures of traffic volume per census tract or ZCTA in the United States from 1963 to 2019 (primarily 1997 to 2019). High traffic volume may be used as a proxy for heavy traffic, high traffic speeds, and impediments to walking or biking. The dataset contains measures of the average, maximum, and minimum traffic volume per tract or ZCTA per year. These figures are available for all streets, highways, and non-highways.

**Urbanicity** - This dataset contains measures of the urban/rural characteristics of each census tract in the United States. These include proportions of urban and rural population, population density, rural/urban commuting area (RUCA) codes, and RUCA-based four- and seven- category urbanicity scales.

**Voter Registration, Turnout, and Partisanship** - This dataset contains counts of voter registration and voter turnout for all counties in the United States for the years 2004-2018. It also contains measures of each county's Democratic and Republican partisanship, including six-year longitudinal partisan indices for 2006-2016.

### **State Policy & Politics Database (SPPD)**

SPPD is a compilation of annual data on state policies and politics that are particularly relevant for population health. The SPPD includes several categories of policies, including labor and economic policies (e.g., minimum wage levels, right to work laws), social safety net policies (e.g., earned income tax credits, SNAP), behavior-related policies (e.g., tobacco taxes, opioid prescribing), as well as the political ideology of the states' government and citizens. Most measures are available annually from 1980 to 2021. More detailed information can be found on the [SPPD site](#).

## Merging GLR Contextual Files to HRS, PSID and NHATS/NSOC

The table below is an overview of the geographic level at which contextual data can be merged with HRS, PSID and NHATS/NSOC. For specific variable names, formats, and temporal availability please consult the *GLR Geographic Variables Crosswalk*, available in the enclave.

<b>MiCDA GLR Geographic Levels Available for Merging</b>				
	<b>HRS</b>	<b>PSID</b>	<b>NHATS</b>	<b>NSOC</b>
<b>COVID-19 US State Policy Database (CUSP)</b>	State	State	State	State
<b>CDR: Area Health Resources Files</b>	State, County	State, County	State, County	State, County
<b>CDR: Decennial Census and ACS</b>	State, County, Census Tract	State, County, Census Tract	State, County, Census Tract	State, County, Census Tract
<b>CDR: NO2 Land-Use Estimate Data</b>	State, Census Tract	State, Census Tract	State, Census Tract	State, Census Tract
<b>CDR: EPA FAQSD O3 Fused Air Quality*</b>	State, County, Census Tract	State, County, Census Tract	State, County, Census Tract	State, County, Census Tract
<b>CDR: EPA FAQSD PM 2.5 Fused Air Quality*</b>	State, County, Census Tract	State, County, Census Tract	State, County, Census Tract	State, County, Census Tract
<b>CDR: Uniform Crime Reporting</b>	State, County	State, County	State, County	State, County
<b>CDR: USDA Food Environment</b>	State, County, Census Tract	State, County, Census Tract	State, County, Census Tract	State, County, Census Tract
<b>NaNDA</b>	State, County, Zip Code, Census Tract	State, County, Zip Code, Census Tract	State, County, Zip Code, Census Tract	State, County, Zip Code, Census Tract
<b>State Policy &amp; Politics Database</b>	State	State	State	State
<b>LTCFocus Database</b>	State, County	State, County	State, County	State, County

\*State, county, and census tract are contained in one variable. Some data manipulation may be required to link to other data sets.

### Questions

If you have questions about the MiCDA Enclave or Geographic Linkages Repository, please contact Jana Deatrck at [jlbruce@umich.edu](mailto:jlbruce@umich.edu).