Greater DC Data Explorer: Methodology and Notes:

Last Revised April 5, 2018

Overview

The [Greater DC Data Explorer](#) on [GreaterDC.Urban.org](#) is part of Urban–Greater DC's effort to democratize data and aid community organizations, foundations, and local government in understanding their communities. The Greater DC Data Explorer uses data collected from a variety of sources and standardizes them to create indicators that are consistent over time and across different geographic areas. It captures indicators in several domains for the entire Washington, DC Metropolitan Area, which includes 23 counties or independent cities, plus the District of Columbia. As of March 2018, data for communities in Maryland, Virginia, and West Virginia come from the decennial U.S. Censuses and the American Community Survey. Data for the District of Columbia come from those sources, as well as other local administrative data sources.
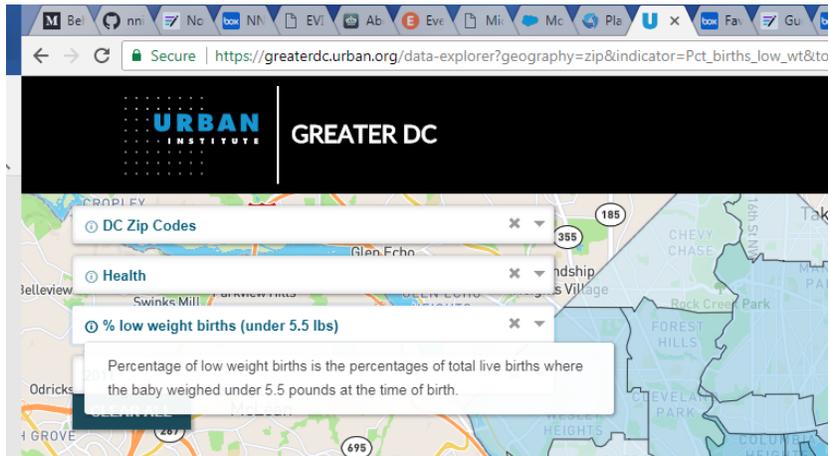
There is detailed information in the Data Explorer about geographies, data sources, and definitions for indicators wherever you see the ⓘ icon, such as on the left-hand side of the selector boxes in the top left corner of the Data Explorer (figure 1).

FIGURE 1:
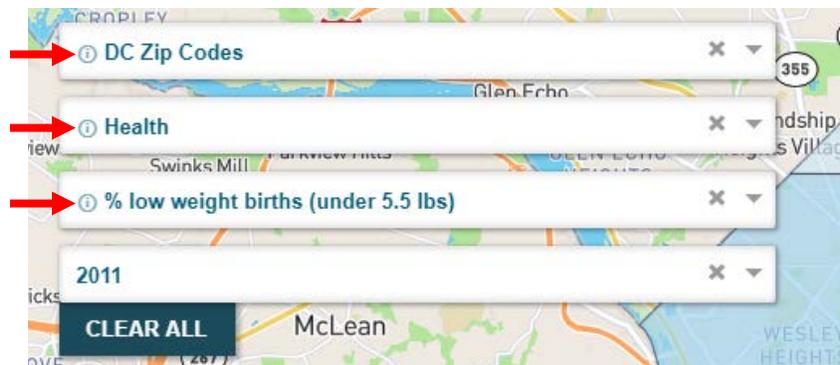
Clicking on the ⓘ will reveal the information available (figure 2).

FIGURE 2

In this document, we have provided additional details and notes on geography, data sources, and table elements, that provide more complete information about the data and methodologies used to produce the data on the Greater DC Data Explorer. If you have questions about data or methodology, please email us at greaterdc@urban.org.

Geography

In this section we define the geographic coverage of the Data Explorer and additional details on geographies used for the District of Columbia.

Coverage

The Greater DC Data Explorer includes data for the Washington-Arlington-Alexandria, DC-VA-MD-WV Metropolitan Statistical Area as defined by the Office of Management and Budget in 2015 as:

- District of Columbia

- In Maryland: Calvert County, Charles County, Frederick County, Montgomery County, and Prince George's County

- In Virginia: Arlington County, Clark County, Culpeper County, Fairfax County, Fauquier County, Loudoun County, Prince William County, Rappahannock County, Spotsylvania County, Stafford County, Warren County, Alexandria City, Fairfax City, Falls Church City, Fredericksburg City, Manassas City, and Manassas Park City

- In West Virginia: Jefferson County

Data for all these areas are provided at the county (or city) and census tract-level. Data for the District of Columbia are provided at additional levels of geographic detail.

Geographies Specific to the District of Columbia

In addition to census tracts, data for the District of Columbia are also provided for several types of small geographic areas including: Advisory Neighborhood Commissions (ANCs), Police Service Areas (PSAs), neighborhood clusters, city council Wards, and zip codes. Additional information about these geographies can be found by clicking on the ⓘ in the Data Explorer where the geography is selected.

Some of the source data are provided to Urban–Greater DC for geographic areas, such as census tracts and census block groups, which do not line up perfectly with other DC geographies like Wards and zip codes. The Urban–Greater DC team mapped these data to each set of geographical boundaries using population weights to split tracts or block groups, when necessary. Consequently, for some indicators at geographic levels such as, ANCs, neighborhood clusters, PSAs, and zip codes, the indicators represent close approximations of the values, rather than precise counts.

Data Sources

In this section we briefly describe some of the data sets used in the Greater DC Data Explorer and provide more details about their characteristics and sources.

American Community Survey

The American Community Survey (ACS) is a national survey of households and housing units, administered by the U.S. Census Bureau on an ongoing basis. Starting after the 2000 Census, the ACS replaced the Decennial Census long form, which was administered only once every ten years, collecting much of the same information on demographics, poverty, employment, housing, and other detailed characteristics.

While the ACS offers the advantage of more frequent data collection, it has a smaller sample size than the long form. To obtain reliable estimates for small areas, multiple years of ACS data must be combined to produce a single indicator. In the Greater DC Data Explorer, we must use the ACS 5-year estimates, which combine data for 60 consecutive months of surveys. The ACS data on the Data Explorer are updated based on the most recent available 5-year ACS data and are displayed with the start and end years of 5-year period, e.g. 2012-16. The estimates on indicators from ACS can be thought of as an average over this 5-year period.

One-year ACS estimates are available for geographic areas with more than 65,000 people; however, we use the 5-year estimates on the Data Explorer for two reasons: 1) the 1-year estimates are not available for all jurisdictions in the metropolitan area, and 2) to maintain consistency between the census tract-level and county-level indicators provided.

Because surveys are asked only of a sample or subset of the population, an estimate will be associated with a range of values above and below the estimate at which it is mostly likely the true value lies. If the same survey questions were asked of a different group of people, the survey may produce slightly different results. A "margin of error" reflects this variation of estimates. With each ACS estimate, the Census Bureau reports the margin of error (MOE). The MOE enables data users to measure the range of uncertainty around each estimate. The larger the MOE, the lower the accuracy of the estimate and the less confidence one

should have that the estimate is close to the true value. More information on the ACS can be found [on the Census Bureau's web site](#). We have included the MOEs for all ACS data in the tables in the Greater DC Data Explorer. Smaller geographic levels, such as census tracts, will have much larger MOEs, and therefore estimates that are less reliable, than higher geographic levels like cities or counties.

Decennial Census

Every decade the Census Bureau updates census tract boundaries, making it difficult to compare data over time for a consistent area. To transform data from previous geographic boundaries to new ones, the Neighborhood Change Database (NCDB) was created by GeoLytics and the Urban Institute with financial support from the Rockefeller Foundation. The NCDB is based on Decennial Census data from 1990, 2000, and 2010, and remapped to 2010 boundaries for the Data Explorer. More information about this process can be found in the [NCDB User's Guide](#).

Washington, DC Administrative Data

The Greater DC Data Explorer contains several sources of administrative data from the District of Columbia– that is data that are created through the operation of government programs or oversight functions. Urban–Greater DC obtains these data in one of several ways. For some data, we download data from [Open Data DC](#), the city's open data portal. Data we obtain through this method includes crime incidents and real property sales. In other cases, we have a Memorandum of Understanding (MOU) in place with the agency that houses the source data and specifies the use of the data as well as how sensitive or confidential information will be protected. For example, we have a MOU established with the DC Department of Human Services to receive data on TANF and SNAP recipients. Additional information about individual indicators can be found by clicking on the ⓘ in the Data Explorer.

Tables in the Data Explorer

After selecting a geography, topic in the Data Explorer, a table will appear at the bottom of the screen. The table will display all indicators and years of data available for that topic, even if there is no map displayed. Users can continue to filter the data and select an indicator and year. On the table, after the indicator label and year, the first numeric column shows the data value for a geography once it has been selected. The name of that geographic area should display in the table header, e.g. "ZIP 20010."

The next three columns display reference values (average, low, and high) to help users interpret the value for the selected geographic area. The column "average" displays the average values for the indicator across *all* geographic areas of the same type. For example, for total population at the census tract level, the average represents the average population in each census tract across the DC metropolitan area. In another example, the number of TANF recipients at the Ward level represents the average number of recipients in each Ward in the District of Columbia, as the data are only available for DC.

The low and high columns on the right side of the table display the value for the geography with the lowest and highest values. For example, for total population at the census tract-level, the low value

represents the census tract with the fewest people and the high value represents the census tract with the most people.

Recent Updates to the Explorer

As new data or new geographies get added to the site or any corrections are made to the data, we will note them here.

URBAN
INSTITUTE

2100 M Street NW
Washington, DC 20037

www.urban.org