Data for Environmental Health

Last Updated: 3/22/2019

Data about Environmental Health is actually somewhat of an anomaly. There are few accurate and reliable sources of data, because, quite simply, the rule that correlation is not causation dictates many limitations for the use. What isn't ruled out of usefulness from that scenario is removed from use because of personally identifiable information.That being said, this assemblage of a variety of health-related datasets is an expansive view of what may be health-related as many of these datasets fall in the environmental health area. The mission at Go Code Colorado is to make public data more accessible and useful for informed business decision making. We do that primarily through publishing state-level data to the state's open data platform, the Colorado Information Marketplace. But we know data in combination is key so we look at all kinds of data sources when considering analysis. The datasets highlighted in **RED** are our featured datasets--that we have assisted in publishing. The blue highlight is for community resources we are aware of.

Denver Resources
- DRCOG is the regional planning council for the Denver Metro area, and they have superb GIS data. http://gis.drcog.org/datacatalog/
- Denver Public Schools has a really strong data analysis and GIS group. They have great data they might share.

Colorado Resources
- Colorado Information Marketplace https://data.colorado.gov/ is the state's central repository for machine-readable data. Check out our resources on Go Code Colorado for tips and tricks on how to maximize your experience and featured datasets on Current Datasets. Also we have some good info on the previous years' apps built with public data, and Additional Data Resources has links to data portals with downloadable data.

Colorado Resources with Environmental Health Relevance

- **Colorado Department of State**: Durable Medical Equipment Suppliers in Colorado. Durable medical equipment (DME) and disposable medical supplies are medically necessary, in general, means that the DME/supply has a medical purpose that is appropriate for the intended use, it is the most cost effective item available, and it provides for a safe environment. Some examples of equipment that are considered DME include: oxygen, wheelchairs, walkers and bathroom/bedroom safety equipment. This dataset **Durable Medical Equipment Suppliers in Colorado** provides the location of DME suppliers.
- **Colorado Department Of Public Safety**: Crime Statistics By Police District. Crime data is collected by local Police Districts (sheriff, police and state agencies) and then reported up to the State Department of Criminal Justice to provide these aggregate statistics for the state. Arrests are incidents in which an individual was actually booked and charged. The annual report represents data on crimes (**Offenses** and **Arrests**) reported by law enforcement agencies through the Colorado Uniform Crime Reporting Program (UCR) and are stored in the Colorado Crime Information Center (CCIC) computer located at the Colorado Bureau of Investigation. During 2015, 245 law enforcement agencies were responsible for reporting crimes and arrests to the Colorado Bureau of Investigation (CBI). This data is now available dating back to 2001.
- **Boulder County**: Septic Systems In Boulder County Colorado. Septic system data including type, status and location for dwellings within Boulder County, Colorado provided by Boulder County can be found in the **Septic Systems in Boulder County Colorado** dataset. Additionally, there is information on the duration of permitting process, the engineer and installer, the lot size and number of bedrooms (for residential sites), and also the depth to bedrock and site soil type. Septic systems are the second most frequently cited source of groundwater contamination in the US. Unapproved, aging, and failing septic systems have a large impact on the quality and safety of the water supply.
- **Colorado Department Of Public Health And Environment**: Healthcare Facilities. There are five themed data sets pertaining to the location and function of healthcare facilities for the State of Colorado. The primary data set, **Health Facilities**, includes known regulated health facilities in Colorado, and is developed directly from address information obtained directly from the Health Facilities and Emergency Medical Services Division of CDPHE. Also in this suite are **Community Behavioral Health Centers**, **Drug Treatment Programs and Resources** and a wide variety of other socio-demographic data sets relevant to healthcare. For more on Behavioral Statistics and Emergency Management related data, such as Durable Medical Equipment, see the **Health and Environment Available Dataset Summary for Go Code Colorado** on GitHub. For more information about GitHub, see the **Resources** page in the **Compete** section of this website.
- **Colorado Department Of Local Affairs Housing Division**: Foreclosures. The Department of Local Affairs Housing Division tracks both the number of home **Foreclosure Filings in Colorado** and **Foreclosure Sales in Colorado** by county for each quarter. As a complement to the existing Rents and Vacancies data, the Housing Division has released Foreclosures in Colorado, a new data set this year with 2015-2016 data. Foreclosure filings provide a good guide to foreclosure activity in a given county, and whereas a property may be withdrawn from the foreclosure process after a filing is made, the filings statistics nevertheless indicate where borrowers are delinquent and in default.

- **Colorado Department Of Local Affairs Demography Division**: Census. This year the DOLA Census division has published 10 geospatial data sets, five for **2009-2013** and five for **2011-2015**. For those who have tried and struggled, you know why it is so valuable that Go Code Colorado and the Demography Division work together to make census data updates easily available each year. Quite simply, it can be very difficult to get data from the American Community Service website that is specific to the region you are looking for and the fields you want.
- **Colorado Department Of Human Services**: Licensed Child Care Facilities. The Colorado Department of Human Services (CDHS) updates the **Licensed Child Care Facilities** data set monthly with any changes or additions to non-24 hour licensed child care facilities in the State of Colorado. Fields include CDHS-issued child care license numbers and legal business names of the child care facility programs as they appear in the licensing application as well as the physical location addresses (including coordinates). For each location, the total licensed capacity and the types of services and programs are provided. Also included in this data set is the Colorado Shines quality rating level (if applicable), and whether the location utilizes the Colorado Child Care Assistance Program (CCCAP).
- **Colorado Department Of Labor And Employment**: Open/Closed Petroleum Release Events In Colorado. All the **Open Petroleum Release Events** that are currently active petroleum releases from regulated storage tank systems are reported to the State of Colorado's Division of Oil & Public Safety. These petroleum releases (or "events") are currently being investigated, assessed or remediated, or are obtaining closure. For active and closed events, navigate to the **Active and Closed Petroleum Release Events in Colorado** data set. The Petroleum Program within the division is responsible for protecting human health and the environment from petroleum releases caused by regulated storage tank systems. This includes oversight of assessment and cleanup of regulated substances from both underground storage tank (UST) and above ground storage tank (AST) systems. This data set is updated monthly from the Colorado Storage Tank Information System (COSTIS) Database, and the locations are generated from point address based geocoding.
- **Colorado Department of Health and Environment** Health and Environmental Data
- **Colorado Department of Health and Environmen**t Community Level Health Estimates
- Colorado School View Data Center - View profiles of individual schools and districts as well as state education data, including district and school performance frameworks, performance results, enrollment, disciplinary actions and more.

Federal Resources with Relevance to Environmental Health
- CDC - (Socrata Portal) - Center for Disease Control Behavioral Risk Factor Surveillance System. Has datasets related to Rates of Obesity, Rates of Cardiovascular Disease, Tobacco Use, Binge Drinking among Adults, Chronic Health Indicators, Rates of Physical Activity, Number of Fruits and Vegetables Per Week, Oral Health, WOmen's Health, Immunization, Injury, Arthritis, Blood Pressure, Hypertension, Flu Shots, Prostate Cancer, {*anyone notice some other dataset titles in this catalog worth listing here?*}
- Bureau of Transportation Statistics - Provides a wealth of data that encomposes many years of granular data on: Airlines and Airports, Highway, Maritime, Multimodal Sources, Pipeline and Hazardous Materials, Rail, Transit

- [The National Map](#) has countless endpoints available for datasets modeling the earth and primed for environmental analysis.
- [HIFLD](#) - Homeland Infrastructure data catalog houses 228 datasets that include, National Area Code Boundaries, DHL Facilities, Landfills, Federally Insured Bank Locations, Major Sport Venues, Overhead Cables, Critical Habitats, Watershed Boundaries, Dam Areas, Bridge Areas, Landslide Areas, and from Geographic Names Information System (GNIS): Structures, Historical Features, Cultural Features, Landform Features, Community Features.
- [Map of Community Initiatives](#) - Federal Government data on initiatives across 15 different agencies, including AmeriCorps, Promise Zones, Sustainable Communities, and TIGER Grants
- [DataUSA.io](#) - This portal brings together a selected set of Federatl Data Resources that allow the user to click through the data visualization tools to get a view into the data, or to ping their dite to access that same dat athrough an API. Data sources include the Census' American Community Survey, The American Community Survey (ACS) Public Use Microdata Sample (PUMS) files are a set of untabulated records about individual people or housing units, Bureau of Economic Analysis (BEA) publishes data on Input-Output also called Make-Use for industries in the United States. This Dataset is provided by the US Department of Commerce, The Bureau of Labor Statistics (BLS) is a public dataset made available by the United States Department of Labor, The University of Wisconsin County Health Rankings (CHR) dataset is a collaboration between the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute.The Integrated Postsecondary Education Data System (IPEDS) is the core postsecondary education data collection program for the National Center for Education Statistics, a part of the Institute for Education Sciences within the United States Department of Education.The O*Net Skills is a dataset containing detailed descriptions of the required and used skills for specific occupations. The O*Net dataset is sponsored by the United States Department of Labor. Data USA currently uses version 19.0 (July 2014) of the O*NET Database.

General Resources with Relevance to Environmental Health

- [Google Public Data Directory](#) - Google Public Data viewer is a great resource to view various large open datasets to better understand things like Labor Statistics, Broadband performance, in this scenario there are a handful of visualizations that are powered by downloadable data.
- [Google Tables](#) - helps you search thousands of public Fusion Tables, or millions of public tables from around the web you can import to Fusion Tables (make sure you click Fusion Table instead of Web Table in the options bar on the left if you want to get raw data). For example - here is a geocoded table of all the [Colorado Breweries](#) - public and free to use!
- **[OpenStreetMap](#)** has millions of points of interest which will serve you well for this. These are "nodes" (points) each tagged with a set of keywords that indicate what each point is. For example, you might have a node with latitude 51.398, longitude -1.082, and tags amenity=place_of_worship, name=St Mary's Church. This indicates that there's a church called St Mary's Church at that lat and long.
  - So you need to get this data out of OSM. The first step is to identify which tags you're interested in. You might want, for example, everything tagged as a church:

in other words, everything with amenity=place_of_worship. (The first part of the tag is the 'key', the second the 'value'. You can see a long list of common OSM tags [on our wiki](#).). There are two ways you can get this data out of the OSM database. One is by using XAPI, the Extended API. XAPI allows you to query the server for data matching certain tags. It can be very slow, so certainly isn't suitable for live use (you shouldn't use OSM servers for your live app anyway). But it's an easy way to download a lot of data matching certain criteria. The ideal way, however, is to download the OpenStreetMap data dump and filter it according to your needs. The data dump is called planet.osm. You can either download the whole world (really huge), or just an extract covering one country or state.Once you've downloaded it, a wonderful Java program called Osmosis will do all the hard work of extracting the data for you, and either loading the result into a database, or outputting it as a file.

- [Yelp LIVES](#) - Local Inspector Value-Entry Specification, Restaurant Inspection Data integrated with Yelp in [~20 cities/locations](#)
- [Urban Observatory](#) - Uses lots of great data, good site for inspiration.