AidData GeoQuery Request Documentation

Report Info

Request Name	Request 07-08-18 19:47
Request Id	5b424e3ac15e00a98d40d097
Email	ytolochko@gmail.com
Generated on	2018-07-08 13:48:10 (EDT)
Download Link	geo.aiddata.org/query/#!/status/5b424e3ac15e00a98d40d097

Processing Timeline

submitted	2018-07-08 13:47:38 (EDT)
prepared	2018-07-08 13:48:07 (EDT)
processed	2018-07-08 13:48:07 (EDT)
completed	2018-07-08 13:48:10 (EDT)

Citation

Please cite the following in any and all applications of the extracted datasets:

Goodman, S., BenYishay, A., Runfola, D., 2016. Overview of the geo Framework. AidData. Available online at geo.aiddata.org. DOI: 10.13140/RG.2.2.28363.59686

Contents of Request Zip

- request documentation (this pdf document)

- a comma separated value (CSV) file containing your data

- JSON file containing your request parameters

- "Introducing the AidData Geo Framework" paper (pdf)

For additional information, usage tips, guides and more please visit *geo.aiddata.org*.

To get in touch, please contact us via *geo@aiddata.org*.

Meta Information

Boundary

Title	Democratic Republic of the Congo ADM2 Boundary - GADM 2.8
Name	cod_adm2_gadm28
Version	2.8
Description	GADM Boundary File for ADM2 (Sub-region) in Democratic Republic of the Congo.
Details	(no additional details)
Bounding Box	[[[12.206628799438477, 5.386097908020076], [12.206628799438477, -13.455676078796273], [31.30572319030756, -13.455676078796273], [31.30572319030756, 5.386097908020076], [12.206628799438477, 5.386097908020076]]]
Date Added	2016-10-19
Date Updated	2017-09-27
Source Name	Global Administrative Areas (GADM)
Source Link	http://www.gadm.org
Citation	Global Administrative Areas (GADM) http://www.gadm.org.

Selection 1 - World Bank Geocoded Aid Data v1.4.2

Title	World Bank Geocoded Aid Data v1.4.2
Name	worldbank_geocodedresearchrelease_level1_v1_4_2
Version	1.4.2
Column Names	worldbank_geocodedresearchrelease_level1_v1_4_2. fa137b9. sum
Filters	hash: fa137b97af6f70839bffd55e8546defa39d31f36
ad_sector_names	All
Description	Aid data from World Bank Donor System, geocoded and published by AidData. Covers projects from 1995 to 2014. Version 1.4.2.
Details	(no additional details)
Bounding Box	[[[-175.6332, 72.0], [-175.6332, -54.666669999999996], [179.19981, -54.666669999999996], [179.19981, 72.0], [-175.6332, 72.0]]]

Date Added	2017-03-29
Date Updated	2017-03-29
Source Name	World Bank
Source Link	http://data.worldbank.org/
Citation	AidData. 2017. WorldBank_GeocodedResearchRelease_Level1_v1.4.2 geocoded dataset. Williamsburg, VA and Washington, DC: AidData. Accessed on [date]. http://aiddata.org/research-datasets.
Download Link	http://aiddata.org/datasets

Selection 2 - DMSP-OLS Nighttime Lights

Title	DMSP-OLS Nighttime Lights
Name	v4composites_calibrated_201709
Version	4
Column Names	Format: "v4composites_calibrated_201709.<temporal>.<method>" for all combinations of <temporal> and <method> which can be found in the "Temporal Selection" and "Extract Types Selected" fields below (88 columns total)
Temporal Selection	2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002, 2001, 2000, 1999, 1998, 1997, 1996, 1995, 1994, 1993, 1992
Extract Types Selected	max (maximum value measured within each unit of analysis), mean (average value measured within each unit of analysis), min (minimum value measured within each unit of analysis), sum (sum of values measured within each unit of analysis)
Description	Version 4 DMSP-OLS Nighttime Lights composites. The lights from cities, towns, and other sites with persistent lighting, including gas flares. Ephemeral events, such as fires have been discarded. Calibrated across sensors and years using Elvidge 2014 coeficients.
Details	Calibrated across sensors and years using Elvidge 2014 coeficients (coefficients for 2014 acquired via communications with Elvidge, 2016). Values greater than 63 after calibration were truncated to 63. Calibration formula fixed from previous processing of raw v4composites dataset. Calibration citation: Elvidge, Christopher D., Feng-Chi Hsu, Kimberly E. Baugh, and Tilottama Ghosh. 'National trends in satellite-observed lighting.' Global urban monitoring and assessment through earth observation 23 (2014): 97-118.

Bounding Box	[[[-180, 75.00416666665], [-180, -65.00416610665], [180, -65.00416610665], [180, 75.00416666665], [-180, 75.00416666665]]]
Date Added	2017-09-25
Date Updated	2017-10-03
Source Name	NOAA National Geophysical Data Center
Source Link	https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html
Citation	Image and Data processing by NOAA's National Geophysical Data Center. DMSP data collected by the US Air Force Weather Agency.
Variable Description	digital number 0-63.
Resolution	0.00833333333333
Factor	1.0

Selection 3 - UCDP Conflict Deaths

Title	UCDP Conflict Deaths
Name	ucdp_deaths_171
Version	17.1
Column Names	Format: "ucdp_deaths_171.<temporal>.<method>" for all combinations of <temporal> and <method> which can be found in the "Temporal Selection" and "Extract Types Selected" fields below (28 columns total)
Temporal Selection	2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002, 2001, 2000, 1999, 1998, 1997, 1996, 1995, 1994, 1993, 1992, 1991, 1990, 1989
Extract Types Selected	sum (sum of fatalities per unit of analysis)
Description	Number of total fatalities per 0.01 decimal degree grid cell resulting from conflict event using UCDP Georeferenced Event Dataset (GED) global version 17.1.
Details	
Bounding Box	[[[-117.3, 60.0], [-117.3, -34.61], [160.27999999999997, -34.61], [160.27999999999997, 60.0], [-117.3, 60.0]]]
Date Added	2017-07-27
Date Updated	2017-07-27
Source Name	UCDP

Source Link	http://ucdp.uu.se/downloads/
Citation	Sundberg, Ralph, and Erik Melander, 2013, 'Introducing the UCDP Georeferenced Event Dataset', Journal of Peace Research, vol.50, no.4, 523-532 Croicu, Mihai and Ralph Sundberg, 2017, "UCDP GED Codebook version 17.1", Department of Peace and Conflict Research, Uppsala University
Variable Description	the best estimate of total fatalities resulting from conflict events
Resolution	0.01
Factor	1.0

Selection 4 - Population Density (GPW V4, UN Adjusted)

Title	Population Density (GPW V4, UN Adjusted)
Name	gpw_v4_density
Version	4
Column Names	Format: "gpw_v4_density.<temporal>.<method>" for all combinations of <temporal> and <method> which can be found in the "Temporal Selection" and "Extract Types Selected" fields below (20 columns total)
Temporal Selection	2020, 2015, 2010, 2005, 2000
Extract Types Selected	count (total count of pixels per unit of analysis), mean (average population density per unit of analysis), max (maximum population density per unit of analysis), min (minimum population density per unit of analysis)
Description	Population density (UN Adjusted values) from Gridded Population of the World v4. GPWv4 depicts the density of human population across the globe. Source data provided in 30 arc-second (~1 km) grid cells.
Details	No additional processing of raw data.
Bounding Box	[[[-180.0, 85.0000000000092], [-180.0, -59.9999999999918], [180, -59.9999999999918], [180, 85.0000000000092], [-180.0, 85.0000000000092]]]
Date Added	2017-09-19
Date Updated	2017-10-19
Source Name	CIESIN
Source Link	http://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-adjusted-to-2015-unwpp-country-totals
Citation	Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. Gridded Population of the World, Version 4 (GPWv4): Population Density Adjusted to Match 2015 Revision UN WPP Country Totals. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4HX19NJ.

Variable Description	persons per square kilometer
Resolution	0.00833
Factor	1.0

Selection 5 - Physical Elevation

Title	Physical Elevation
Name	srtm_elevation_500m
Version	1
Column Names	Format: "srtm_elevation_500m.<temporal>.<method>" for all combinations of <temporal> and <method> which can be found in the "Temporal Selection" and "Extract Types Selected" fields below (3 columns total)
Temporal Selection	none
Extract Types Selected	mean (average elevation within each unit of analysis), max (maximum elevation within each unit of analysis), min (minimum elevation within each unit of analysis)
Description	Global elevation (in meters) from Shuttle Radar Topography Mission (SRTM) dataset (v4.1) at 500m resolution.
Details	(no additional details)
Bounding Box	[[[-180.0, 60.008333313696056], [-180.0, -60.00833333376795], [180.0, -60.00833333376795], [180.0, 60.008333313696056], [-180.0, 60.008333313696056]]]
Date Added	2016-10-28
Date Updated	2017-02-06
Source Name	CGIAR-CSI
Source Link	http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1
Citation	Jarvis A., H.I. Reuter, A. Nelson, E. Guevara, 2008, Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT), available from http://srtm.csi.cgiar.org.
Variable Description	meters
Resolution	0.004166
Factor	1.0

Interpreting CSV Column Names

Each CSV will contain a column labeled "asdf_id" which has values for each feature that are unique (within that boundary dataset), one or more columns for your extract data, followed by the original source attributes for the boundary file (e.g., from GADM)

The standard format for extract data column names is a three part string delimited by periods (.)

<dataset>.<filter>.<method>

where

<dataset> is the name of the dataset which was extracted

<filter> describes how the dataset was filtered. This is usually a temporal value (e.g., YYYY format for year such as "1999", "none" for temporally invariant data, or a unique hash describing more complex filters, such as for aid datasets)

<method> is the extract method used to aggregate dataset values to boundary features (e.g., "mean", "sum")

Notes - Aid data extracts

The <filter> component of aid data extracts is a unique hash that corresponds to the filter combination used to generate that particular aid data extract (e.g., donor, sector, year, status). For each aid data extract you request, you will see three columns in the CSV that have the same <dataset> and <filter> sections of the column name with the <methods> of the three being different.

These three <method> values are:

- "sum" is the total aid for each feature within the boundary based on the distribution of aid used when building the aid data

- "potential" is the maximum aid that could have been allocated to each feature regardless of the distribution of aid used

- "reliability" is a ratio of sum:potential representing a simplistic measure of how accurate the distribution and aggregation of aid was relative to the boundary features used during the extract process

Notes - Categorical extracts

Data extracted using the categorical method will have multiple columns with the same <dataset> and <filter> where the .<method> for each is "categorical_.<category>.

For a simple landcover dataset this might look like:

- landcover.2000.categorical_water

- landcover.2000.categorical_forest

- landcover.2000.categorical_desert

Usage Notes

- If you attempt to merge geo(query) results wth vector data (e.g., shapefiles) downloaded from GADM, the GADM data may not always contain a unique id field to merge on. In these cases, please feel free to contact us and we can provide you with a modified file that contains a unique field for merging ("asdf_id" field, found in all result csvs).

Notes About Aid Datasets

- When requesting aid data using a very specific filter (usually resulting in only a single project match), the location count shown in geo(query) may be inaccurate. This can result in aid filters which appear valid while building your request, but result in no aid data in your results csv. This is due to a slight reduction in the accuracy of location counts for the web page in order to make the responses fast enough for user interaction.

- The year filter for aid data is based on project start and end dates (determined by earliest and latest transactions). Because projects are represented by year ranges, multiple aid data selections for individual years may contain duplicate aid. This will result in an inflated total if you sum the aid from each individual year (compared to a single selection for all years). Limited source information on individual or even yearly transactions for a project prevent us from offering more granular temporal aid values for projects.

- All aid data selections result in commitment values, regardless of whether you filter by commitment values or disbursement values (or both). This is due to the notably better project coverage of commitments vs disbursements (e.g., World Bank aid dataset has 99% commitment coverage vs ~75% for disbursements).

Terms of Use

The database and derived products produced by this tool are governed by the licenses described at http://http://geoquery.org/toolsguides. By clicking submit you agree to the terms, which are summarized as:

As long as you:

Attribute: You must attribute any public use of the database, or works produced from the database, in the manner specified in the license. For any use or redistribution of the database, or works produced from it, you must make clear to others the license of the database and keep intact any notices on the original database.

You are free:

To Share: To copy, distribute and use the database.

To Create: To produce works from the database.

To Adapt: To modify, transform and build upon the database.

This is not a license. It is simply a handy reference for understanding the ODC-BY 1.0 — it is a human-readable expression of some of its key terms. This summary has no legal value, and its contents do not appear in the actual license. Read the full ODC-BY 1.0 license text at geo.aiddata.org/license for the exact terms that apply.

Acknowledgements

GeoQuery is an academic research project based out of AidData at William and Mary dedicated to enabling the use of spatial data in decision-making.

This work was performed in part using computational facilities at the College of William and Mary which were provided with assistance from the National Science Foundation, the Virginia Port Authority, and Virginia's Commonwealth Technology Research Fund.