

The POPGRID Data Collaborative

ADVANCING THE USE AND IMPACT OF POPULATION AND
INFRASTRUCTURE DATA IN SUPPORT OF THE SUSTAINABLE
DEVELOPMENT GOALS

Kytt MacManus, Linda Pistolesi, Robert S. Chen, Alex de Sherbinin,
Greg Yetman, Susana Adamo, Marc Levy, and Al Pinto

CIESIN, The Earth Institute, Columbia University
NASA Socioeconomic Data and Applications Center (SEDAC)



Determination of Global Fundamental Geospatial Data Themes

United Nations Committee of Experts on Global Geospatial Information Management (UN-GGIM)

- adopted proposed minimum list of 14 **global fundamental geospatial data themes** designed to facilitate the measurement, monitoring, and management of sustainable development in a consistent way over time to facilitate evidence-based decision- and policy-making.

Data Themes and Reference Frame

- Addresses
- Buildings and Settlements
- Elevation and depth
- Functional Areas
- Geographical Names
- Geology and Soils
- Land Cover and Land Use
- Land Parcels
- Orthoimagery
- Physical infrastructure
- Population distribution
- Transport Networks
- Water
- Reference Frame: Global Geodetic Reference Framework



UN-GGIM: EUROPE

UNITED NATIONS COMMITTEE OF EXPERTS ON
GLOBAL GEOSPATIAL INFORMATION MANAGEMENT



Theme title:	Population Distribution
Description	

“It’s vital to understand the spatial distribution of the population and its characteristics, as well as how

... population
... on already lives
... and economic
... tions. Some
... disasters or war.
... as well as

Possible sources of geospatial data

- United Nations Population Division, and individual country census agencies;
- National Statistical Institutes;
- Commercial providers include: Michael Bauer Research GmbH, Environics, Esri, and Facebook;
- Other potential providers include: The European Union’s Joint Research Commission; Center for International Earth Science Information Network (CIESIN), Columbia University; NASA Socioeconomic Data and Applications Center (SEDAC); Oak Ridge National Laboratory (Landscan products); and,
- KAPSARC.

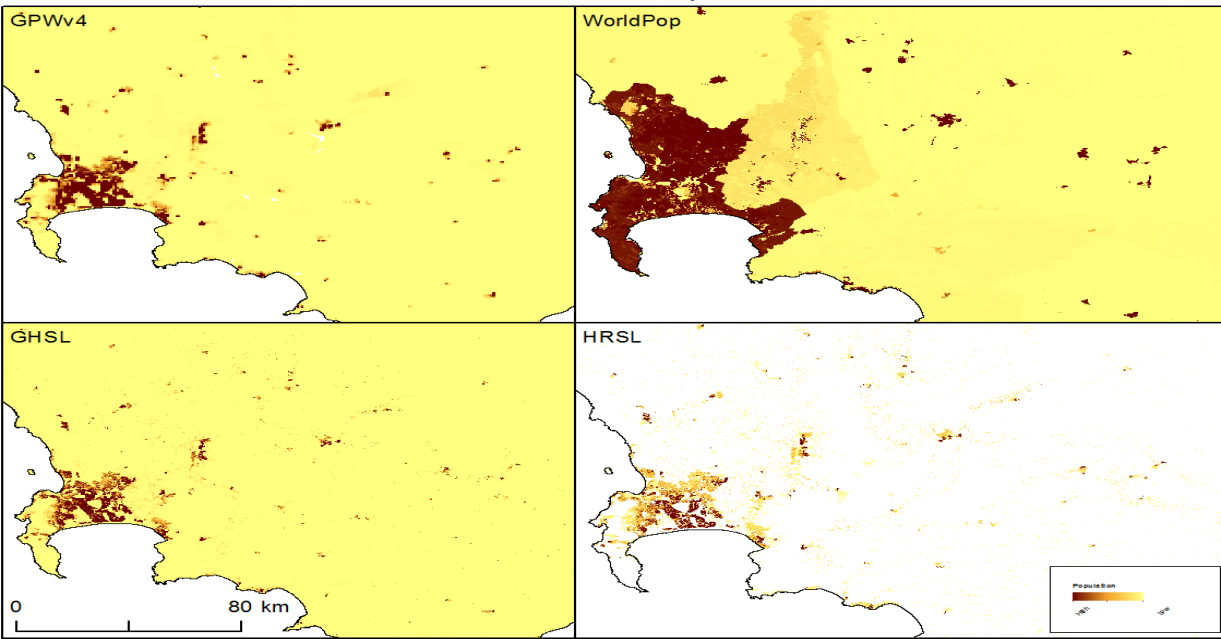
POPGRID

Possible sources of geospatial data

- United Nations Population Division, and individual country census agencies;
- National Statistical Institutes;
- Commercial providers include: Michael Bauer Research GmbH, Environics, Esri, and Facebook;
- Other potential providers include: The European Union’s Joint Research Commission; Center for International Earth Science Information Network (CIESIN), Columbia University; NASA Socioeconomic Data and Applications Center (SEDAC); Oak Ridge National Laboratory (Landscan products); and,
- KAPSARC.

Challenges to Using These Data for Sustainable Development Monitoring and Applications

- Access is scattered; not all open access
- Methods not clear; inconsistent documentation, metadata
- Some are 1-time, research-oriented products; not updated regularly
- Quality may vary by region, time period
- No rigorous validation or inter-comparison conducted
- Poor interoperability or integration with related data, e.g., topography, water bodies, administrative units, land tenure



Four population models for Cape Town and environs, South Africa:

- Gridded Population of the World, version 4 (GPWv4)
- WorldPop
- Global Human Settlement Layer (GHSL)
- High Resolution Settlement Layer (HRSL)

POPGRID: A “Data Collaborative” for Settlement, Infrastructure, and Population Data

Public-private data partnership involving intergovernmental organizations, national & academic research institutions, large and small companies, NGOs, foundations, universities, data stewards, etc.

Overall Goal: Accelerate the development and use of high quality, highly usable georeferenced data on population, human settlements, and infrastructure, drawing on an international, interdisciplinary community of data developers and users from both the public and private sectors.



Diversity of Products with Different Characteristics

Project	Prop. Allocation	Dasymmetric	Statistical / machine learning	Multiple Time Points	Imagery / spectral data	Radar	Nominal Spatial Resolution
GPW	✓			✓			1 km
Landscan		✓	✓?	✓	✓		1 km
WorldPop			✓	✓*			100m
GHSL		✓	✓	✓	✓	**	30m, 250m, 1km
GUF		✓	✓			✓	~12m for scientific research ~84m public
Esri		✓					250m
HRSL			✓		✓		30m
GMIS/HBASE			✓		✓		30 m

* Exists for some countries, planned for WorldPop Global

** Forthcoming GHSL based on GHS built-up R2018A derived from Sentinel-1

Objectives of the POPGRID Data Collective

Objectives

- Facilitate easy access to the data
- Share resources
- Improve data quality, consistency and documentation
- Clarify user needs and priorities
 - Offer variety of data formats and resolutions
- Address scientific and technical challenges
 - Encourage collaboration and innovation
- Facilitate appropriate data use and interpretation in a range of sustainable development application areas

Key Activities

Documentation & Publication

- Review paper
- Consistent metadata and data documentation

Contents

I.	Introduction	2
II.	Data and Methodology	3
III.	Data Set Description(s)	3
IV.	How to Use the Data	4
V.	Potential Use Cases	4
VI.	Limitations	5
VII.	Acknowledgments	5
VIII.	Disclaimer	5
IX.	Use Constraints	6
X.	Recommended Citation(s)	6
XI.	Source Code	6
XII.	References	6
XIII.	Documentation Copyright and License	7
Appendix 1.	Data Revision History	7
Appendix 2.	Contributing Authors & Documentation Revision History	7

Documentation for WorldPop Global Data Products

Documentation for the WorldPop Global Data Products

September 2018

University of Southampton, School of Geography and Environmental Science; University of Louisville, Department of Geography and Geosciences; Center for International Earth Science Information Network (CIESIN), Columbia University

Abstract
This document outlines the basic methodology and data sets used to construct the WorldPop Global data collection including gridded population distribution and age/sex structure datasets with a resolution of 2 arc seconds, along with use cases, limitations

European Commission, Joint Research Centre (JRC)
Documentation for the GHS Population Grid (GHS-POP)

Documentation for the GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015) (GHS-POP)

December 2017

European Commission, Joint Research Centre (JRC)
ghsl-data@jrc.ec.europa.eu

Abstract
This document outlines the basic methodology and data sets used to construct the Global Human Settlement Layer (GHSL) Population Grid (GHS-POP), referred to as the

NASA Socioeconomic Data and Applications Center (SEDAC)
Documentation for Gridded Population of the World (GPW), v4

Documentation for the Gridded Population of the World, Version 4 (GPWv4), Revision 10 Data Sets

November 2017

Center for International Earth Science Information Network (CIESIN)
Columbia University

Abstract
This document outlines the basic methodology used to construct the Gridded Population of the World, Version 4 (GPWv4) data collection and describes the data sets included in the collection, all of which have been updated as Revision 10 for this release. The Introduction briefly describes the input data, the purpose of the collection, the main characteristics of GPWv4, and lists the data sets in the collection. Details of the

Key Activities

Documentation & Publication

- Review paper
- Consistent metadata and data documentation

Framework Data, Standards, & Interoperability

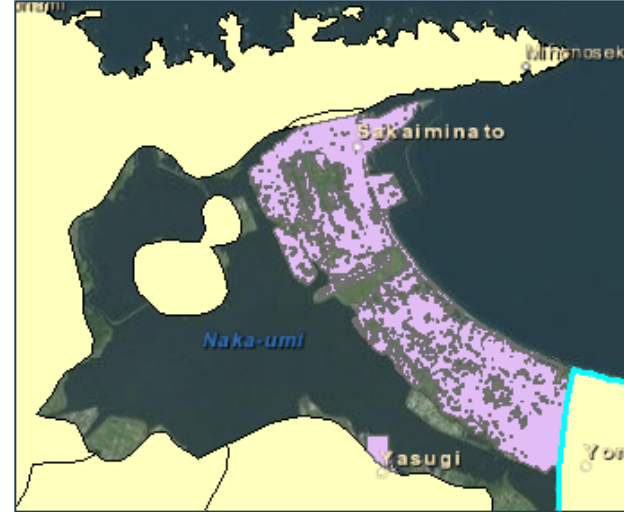
- Sharing input data, addressing coastline issues

Collaboration identified these issues and allowed them to be fixed for GPWv4.10 and forthcoming GHSL r2018.

Yellow = GPWv4

Purple = GHS Built Up

Land missing in GPWv4



GPWv4 edited



Coastline/country shift in GPWv4



GPWv4 edited



Key Activities

Documentation & Publication

- Review paper
- Consistent metadata and data documentation

Framework Data, Standards, & Interoperability

- Sharing input data, addressing coastline issues

Validation & Intercomparison

- Formulating framework/proposal for validation/intercomparison experiment(s)
- Identification of potential validation data sets

Tools, Portals, & Shared Resources

- POPGRID.ORG site
- Coordination of tools efforts

User Needs, Stakeholder Engagement, & Governance

- Working with Global Partnership for Sustainable Development Data and UN Sustainable Development Solutions on governance, staff support, engagement



GLOBAL PARTNERSHIP
ON SUSTAINABLE DEVELOPMENT DATA



SUSTAINABLE DEVELOPMENT
SOLUTIONS NETWORK
A GLOBAL INITIATIVE FOR THE UNITED NATIONS



- Information on available datasets provided by POPGRID participants

User

Username

Password

[Create](#)
[Request](#)

[Log in](#)

Table of Input Layers for Global Gridded Data Sets

Dataset	Input Variables									
	Population	Roads	Land Cover	Built structures	Cities or Urban areas	Night-time lights	Infrastructure	Environmental data (climate, topography, elevation)	Protected areas	Water bodies
Gridded Population of the World (GPW)	x									x
Global Rural-Urban Mapping Project (GRUMP)	x				x	x				x
LandScan Global Population database	x	x	x	x	x		x	x	x	x
Global Human Settlement Layer – Population (GHS-POP)	x			x						
World Population Estimate	x	x	x		x					x
WorldPop	x	x	x	x	x	x	x	x	x	x

to consistent
 a and
 ntation
 son tables
 son tools
 o user support
 AC)

Lightly Modeled Population Grids							
Global Human Settlement Layer – Population (GHS-POP)	European Commission Joint Research Centre (JRC) and Center for International Earth Science Information Network (CIESIN), Columbia University	Nighttime population (population counted at place of domicile)	Binary dasymmetric* using Landsat-derived built up areas and proportional allocation to distribute population data from subnational census data to the settlement extents. See table of input layers below. Details.	250 m (7.5 arc-minutes), 1 km (30 arc-seconds) (World Mollweide projection)	1975, 1990, 2000, 2015	United Nations Population Division (UNPD) estimates and projections	Open access
Global Rural	Center for International	Nighttime population	Binary dasymmetric* using	1 km (30 arc-	1990, 1995, 2000	United Nations	Open access



✓ GHS-Pop 2015 Count (ID: 0)

Choose a different map layer



✓ GPW4.10 UN Adjusted 2015 Count (ID: 1)

- Roads
- Satellite
- Terrain
- Gray Canvas



Downloadable Data Layers

Choose A Layer

POPGRID Data Comparison Viewer

+	GHS Pop 2015 Count (ID: 0)	Estimates of GHS Pop 2015 Count
+	Landscan 2015 Count (2)	Estimates of Landscan 2015 Count
+	World Population Estimate 2015 Count (3)	Estimates of World Population Estimate 2015 Count
+	Worldpop 2015 Count (4)	Estimates of Worldpop 2015 Count

Close

☑ GHS-Pop 2015 Count (ID: 0)

Leaflet | Powered by Esri | SEDAC

☑ Sync Panels | Repeat Map | Zoom Out | Legend | Overview | Download View

☑ GPW4.10 UN Adjusted 2015 Count (ID: 1)

Leaflet | Powered by Esri | SEDAC

Zoom Out | Legend | Overview | Download View

☑ Landscan 2015 Count (ID: 2)

Leaflet | Powered by Esri | SEDAC

Zoom Out | Legend | Overview | Download View

☑ World Population Estimate 2015 Count (ID: 3)

Leaflet | Powered by Esri | SEDAC

Zoom Out | Legend | Overview | Download View



St Helena,
Ascension
and Tristan
da Cunha

Angola

Zambia

Malawi

Comoros

Glorioso
Islands

Mayotte

Mozambique

Mozambique
Channel

Zimbabwe

Madagascar

Mauritius

Réunion

Namibia

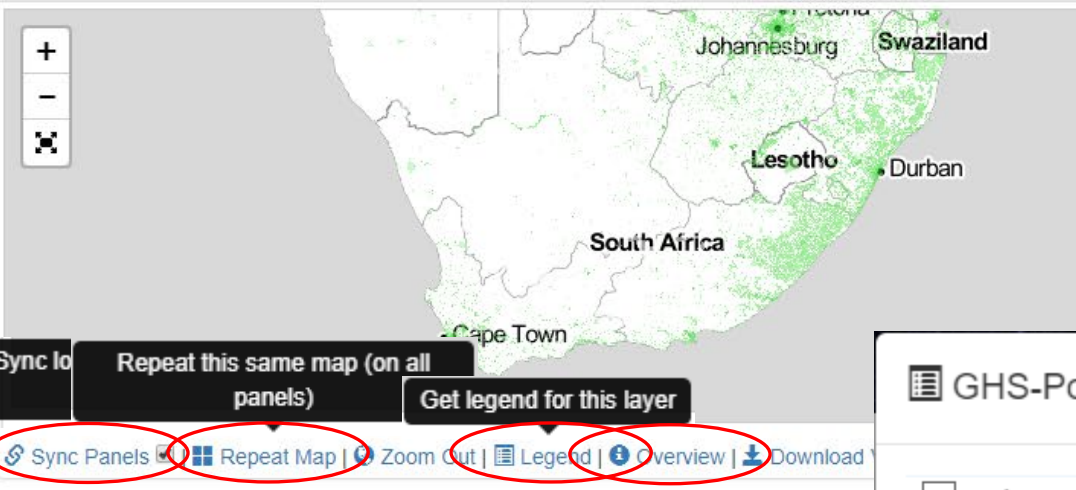
Botswana

Swaziland

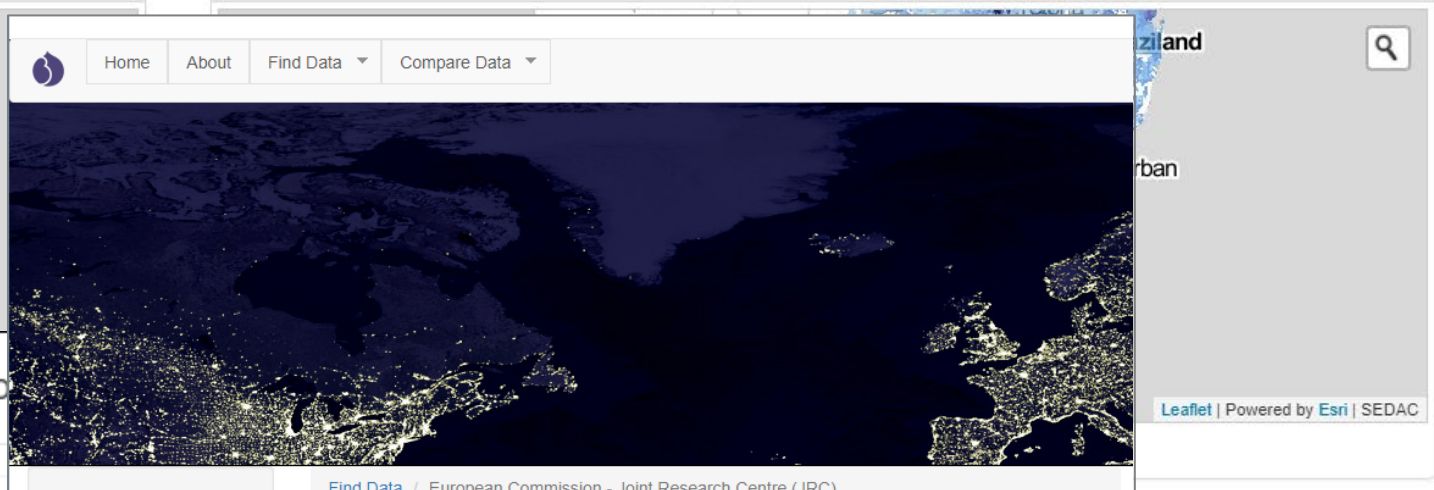
Lesotho

South Africa

GHS-Pop 2015 Count (ID: 0)



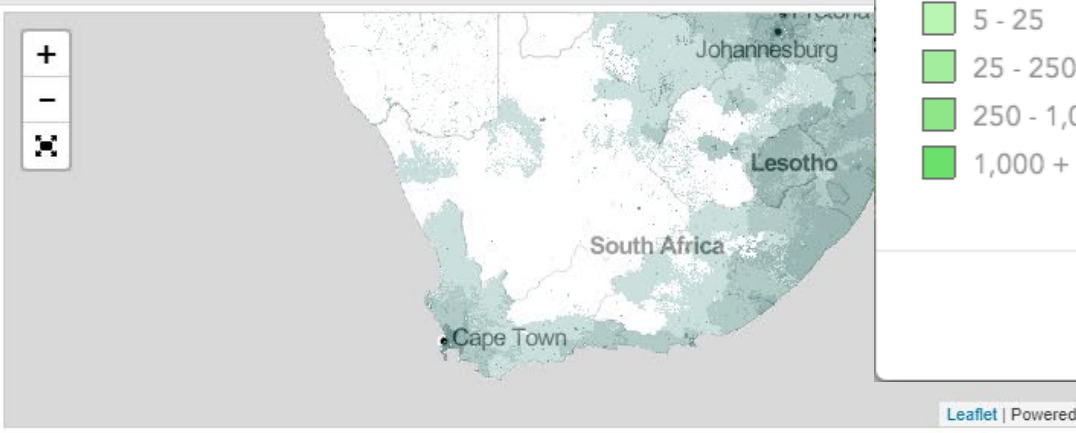
GPW4.10 UN Adjusted 2015 Count (ID: 1)



Sync Panels | Repeat this same map (on all panels) | Get legend for this layer

Sync Panels | Repeat Map | Zoom Out | Legend | Overview | Download

Worldpop 2015 Count (4)



User login

Username *

Password *

[Create new account](#)

[Request new password](#)

[Log in](#)

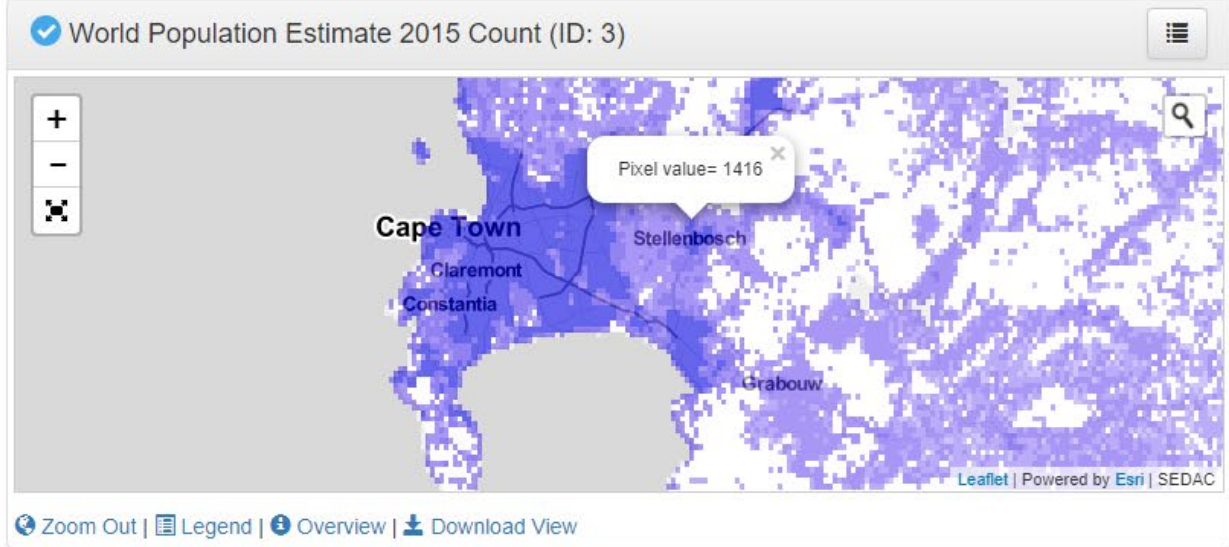
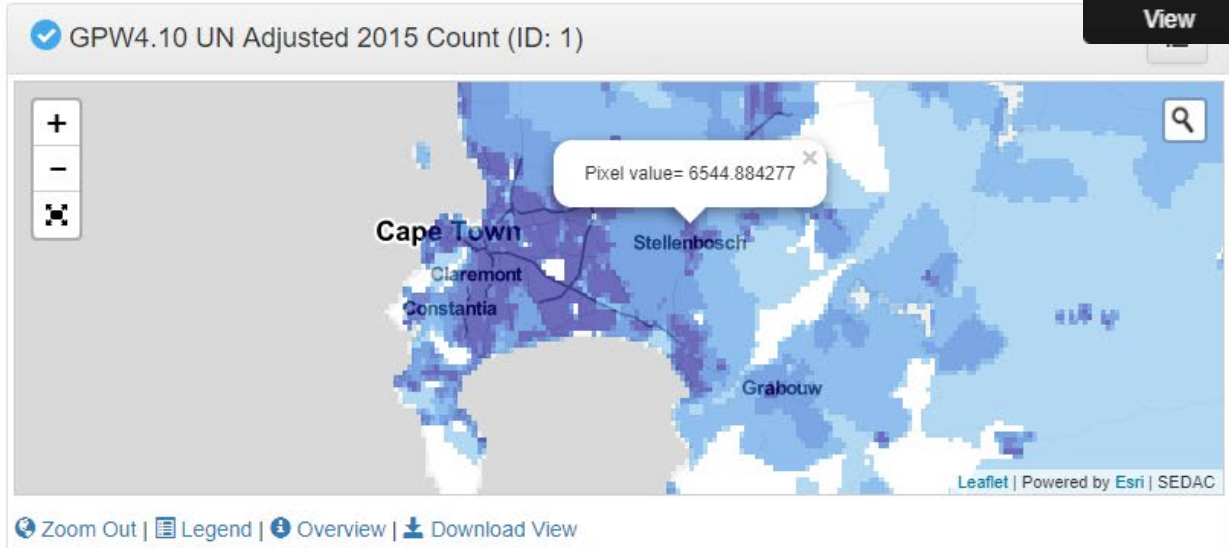
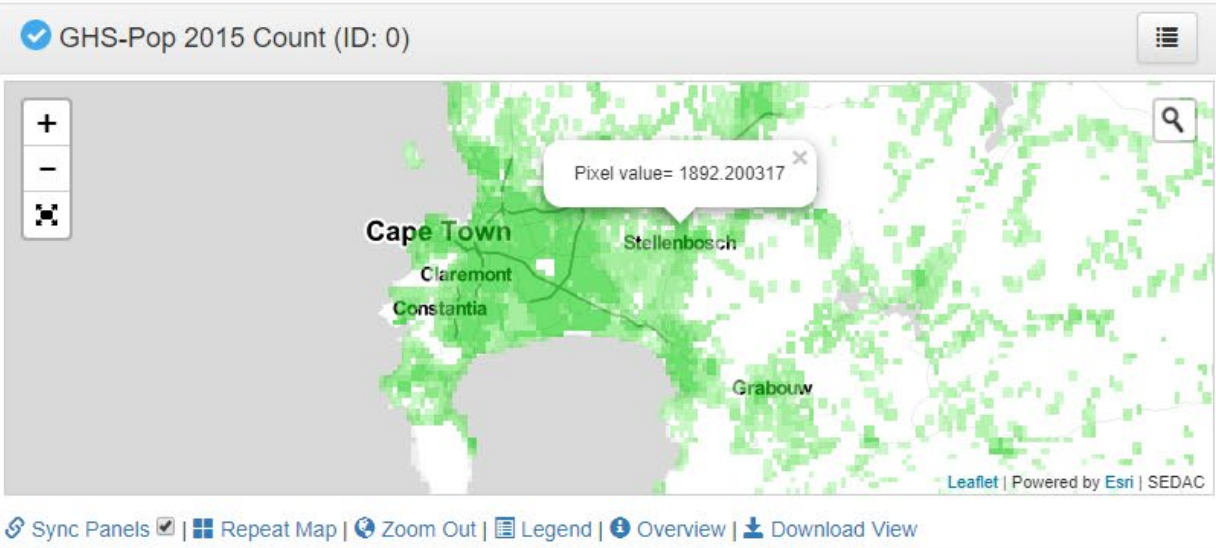
European Commission - Joint Research Centre (JRC)

The Joint Research Centre is the European Commission's science and knowledge service. The Global Human Settlement (GHS) framework produces new global spatial information, evidence-based analytics and knowledge describing the human presence on the planet. It operates in an open and free data and methods access policy (open input, open method, open output). It is supported by the Joint Research Centre (JRC) and the DG for Regional Development (DG REGIO) of the European Commission, together with the international partnership [GEO Human Planet Initiative](#).

- Global Built-up Grids:
- GHS Built-Up Grids P2016 (1975, 1990, 2000, 2013/14)
 - GHS Built-Up Quality P2016 (1975, 1990, 2000, 2013/14)
 - GHS Built Up Sentinel-1 Grid (2016)
- Global Population Grids:
- GHS Population Grids P2016 (1975, 1990, 2000, 2015)

Zoom Out | Legend | Overview | Download View

Zoom Out | Legend | Overview | Download View





Get Population Estimates

- Shapefile
- Rectangle / Polygon

Population Data

- GPW4.10 UN Adjusted 2015 Count (ID: 1)

Opacity

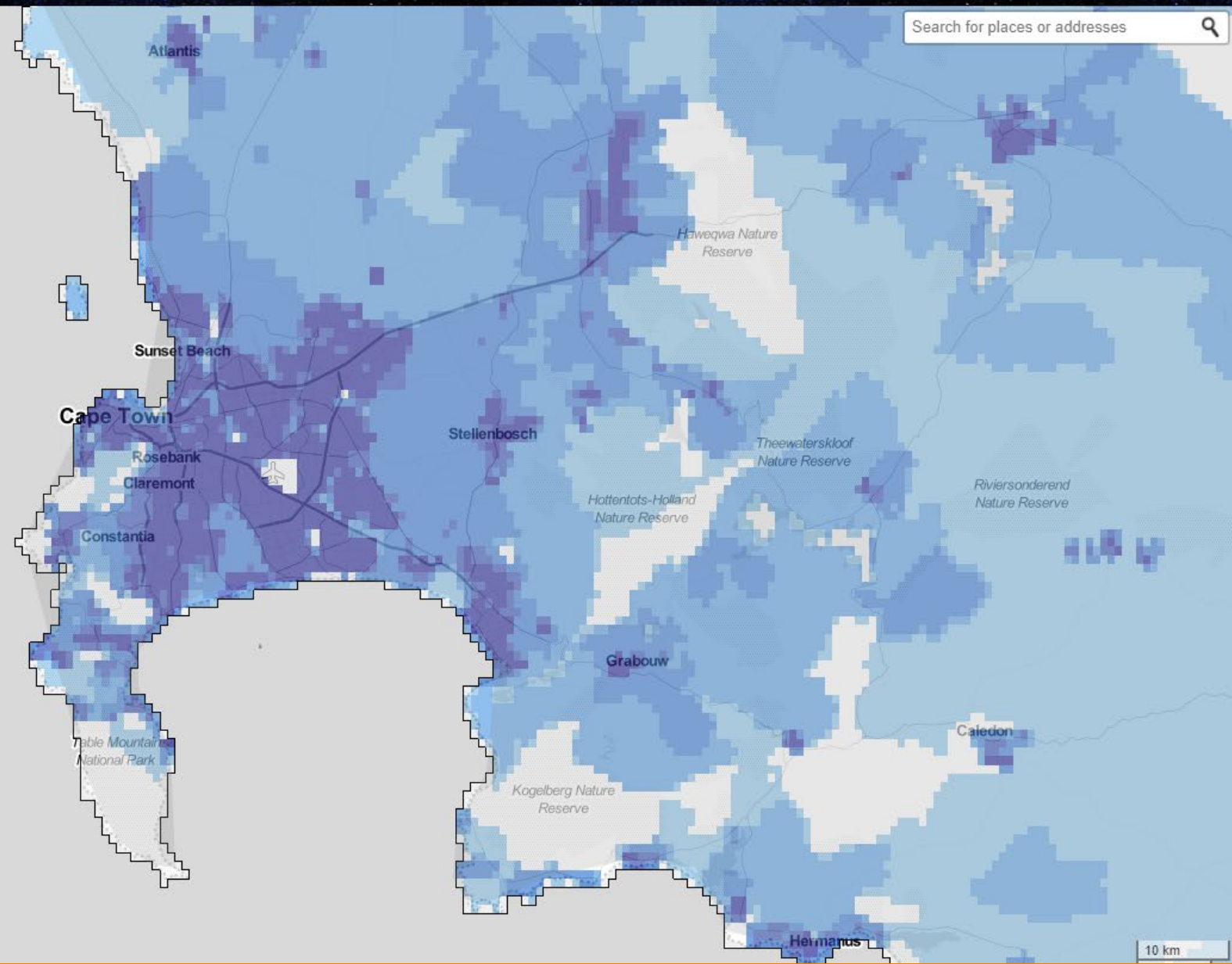
- < 1
- 1 - 5
- 5 - 25
- 25 - 250
- 250 - 1,000
- 1,000 +

- GHS-Pop 2015 Count
- Landscan 2015 Count
- World Population Estimate 2015 Count
- Worldpop 2015 Count

GPWv4.10 Metadata Layers

- Input Data Type (ID: 3)
- Year of Input Data (4)
- Admin Level of Input Data (5)

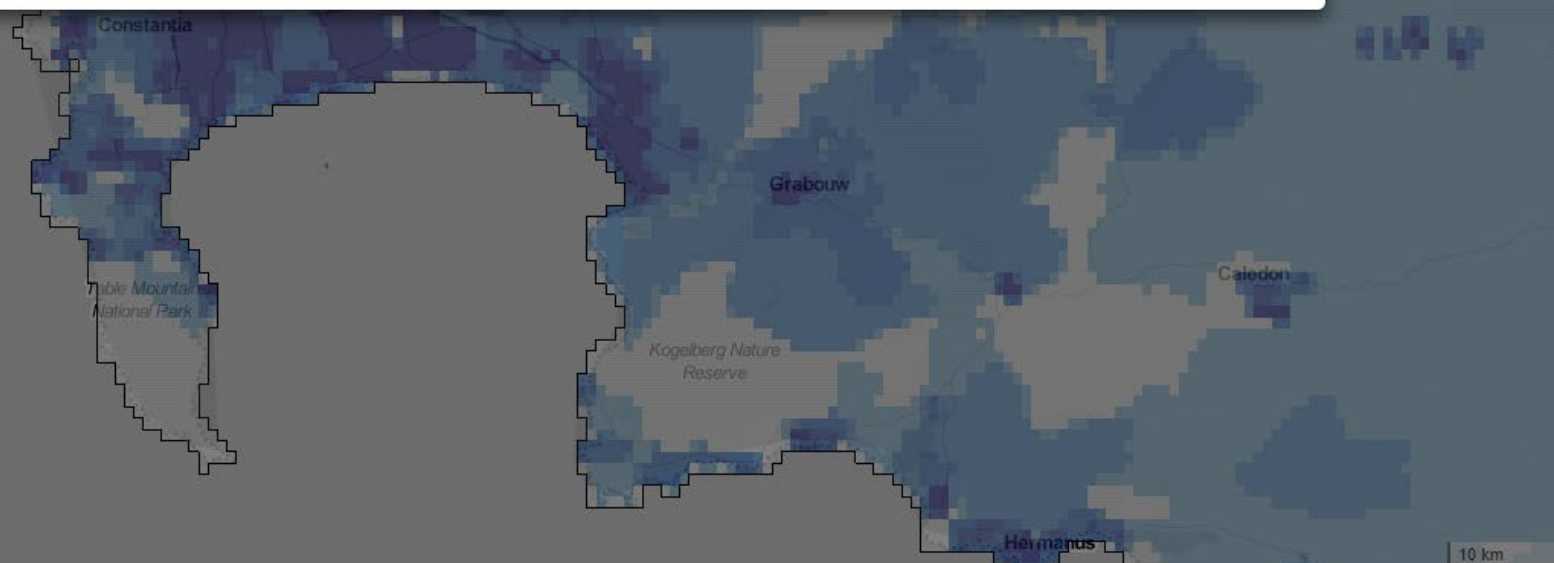
Search for places or addresses



South Africa - Detailed Source Comparison

	ESRI WPE	GPWv4	GHSL	Worldpop	Landscan
Data Source	Michael Bauer Research GmbH	Statistics South Africa, Census 2011, CD: Community Profiles in SuperCross, Table 1: Geography by Gender, generated using SuperCross software, Received 08/19/2013.	Statistics South Africa, Census 2011, CD: Community Profiles in SuperCross, Table 1: Geography by Gender, generated using SuperCross software, Received 08/19/2013.	N/A	N/A
Year of Input Data	2014	2011	2011	-99	-99
Input Feature Count	234	86983		-99	-99
Year of Last Census	2014-MBR-est	2011	2011	-99	-99

Done



Get Populat

Shapefile

Rectangle / P

Population D

GPW4.10 UN

Opacity

GHS-Pop 2015 Count

Landscan 2015 Count

World Population Estimate 2015 Count

Worldpop 2015 Count

GPWv4.10 Metadata Layers

Input Data Type (ID: 3)

Year of Input Data (4)

Admin Level of Input Data (5)

Get Population Estimates

Shapefile

Rectangle / Polygon

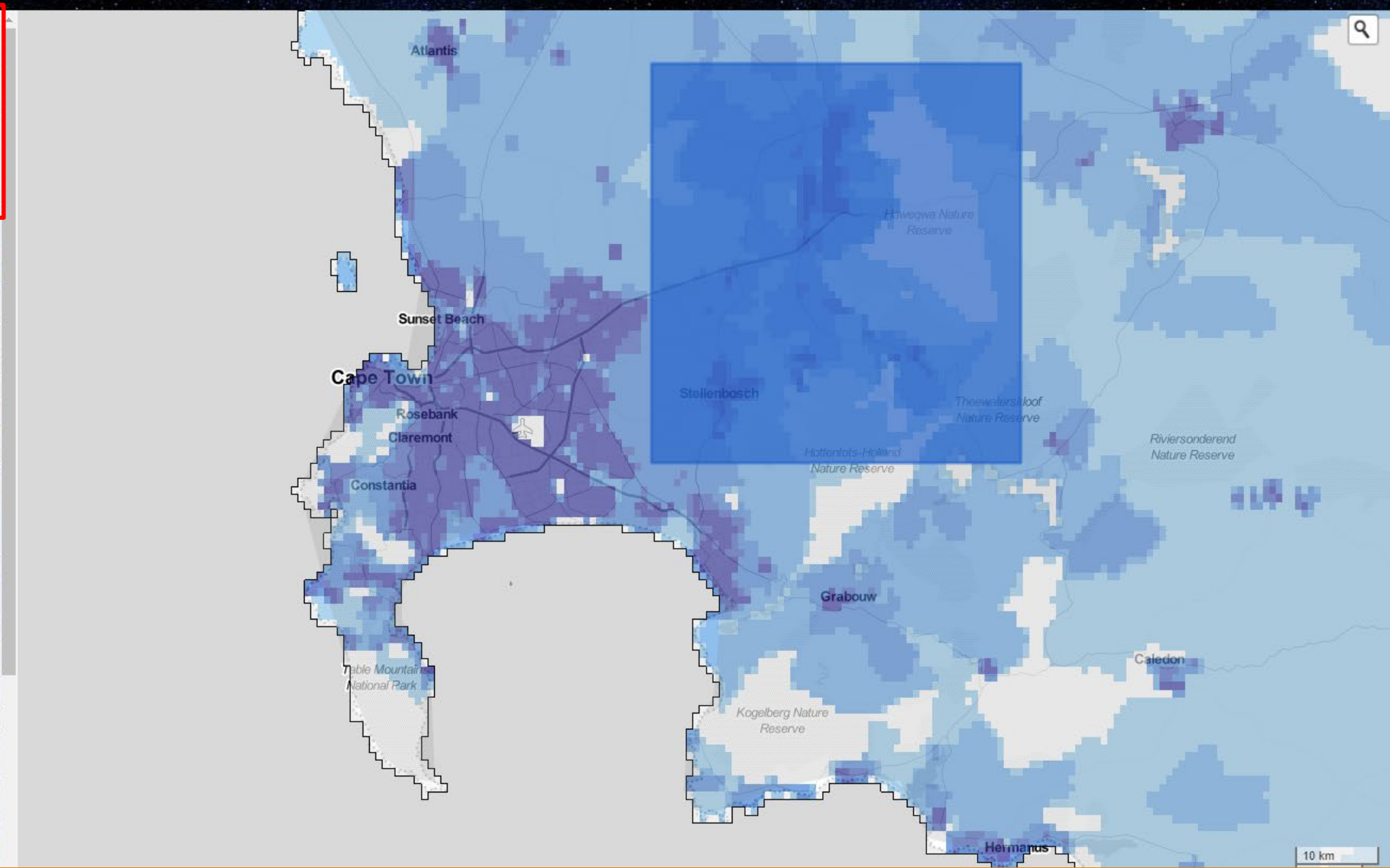
Population Data

GPW4.10 UN Adjusted 2015 Count (ID: 1)

Opacity

<input type="checkbox"/>	< 1
<input type="checkbox"/>	1 - 5
<input type="checkbox"/>	5 - 25
<input type="checkbox"/>	25 - 250
<input type="checkbox"/>	250 - 1,000
<input type="checkbox"/>	1,000 +

- GHS-Pop 2015 Count
 - Landscan 2015 Count
 - World Population Estimate 2015 Count
 - Worldpop 2015 Count
- GPWv4.10 Metadata Layers**
- Input Data Type (ID: 3)



Get Population Estimates

Shapefile

Rectangle / Polygon

Population Data

GPW4.10 UN Adjusted 2015 Count (ID: 1)

Opacity

- < 1
- 1 - 5
- 5 - 25
- 25 - 250
- 250 - 1,000
- 1,000 +

GHS-Pop 2015 Count

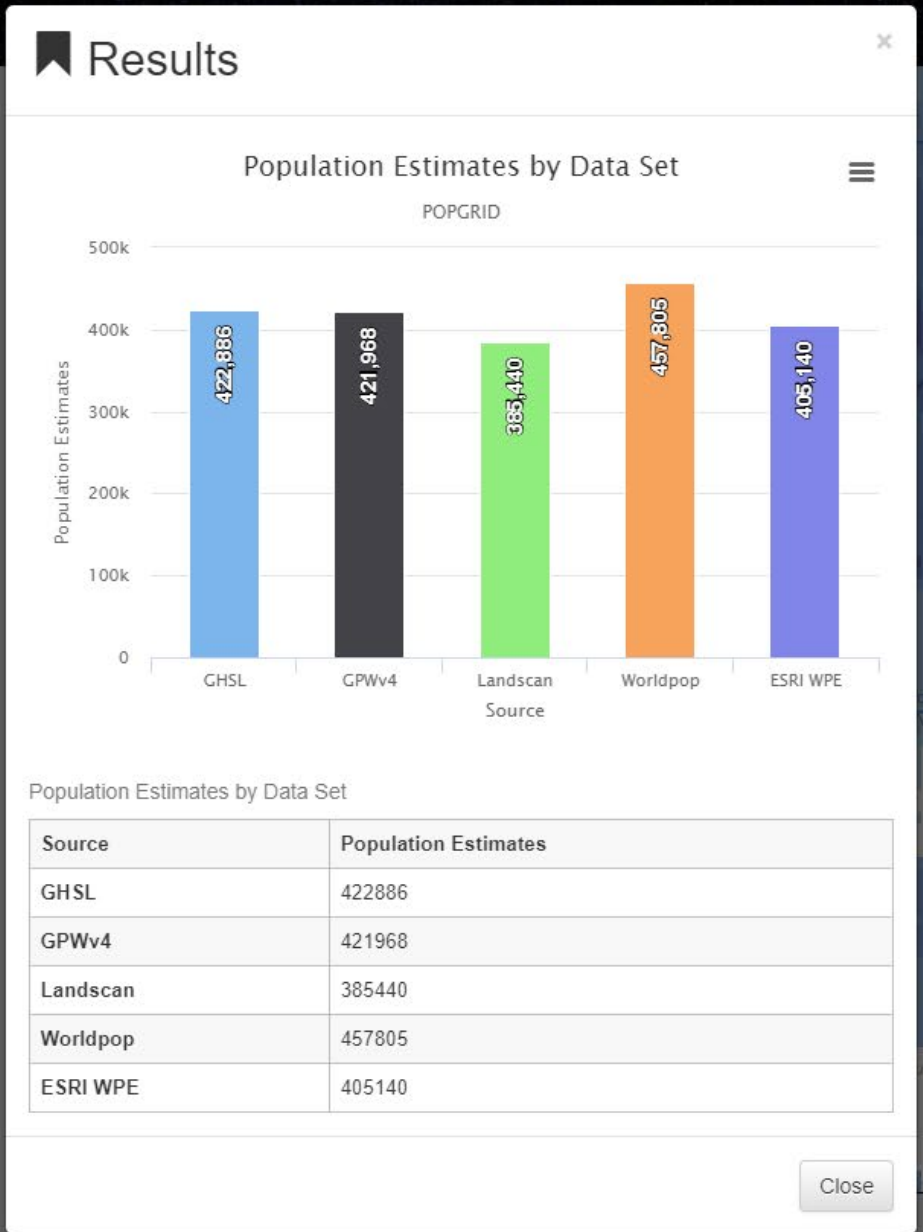
Landscan 2015 Count

World Population Estimate 2015 Count

Worldpop 2015 Count

GPWv4.10 Metadata Layers

Input Data Type (ID: 3)



Get Population Estimates

- Shapefile
- Rectangle / Polygon

Population Data

- GPW4.10 UN Adjusted 2015 Count (ID: 1)
- GHS-Pop 2015 Count
- Landscan 2015 Count
- World Population Estimate 2015 Count
- Worldpop 2015 Count
- GPWv4.10 Metadata Layers**

Input Data Type (ID: 3)

Opacity

- Preliminary/provisional census
- Population register
- Population estimate/projection
- Not applicable
- Final census

Year of Input Data (4)

Admin Level of Input Data (5)

GPWv4.10 Metadata Layers

Input Data Type (ID: 3)

Opacity

- Preliminary/provisional census
- Population register
- Population estimate/projection
- Not applicable
- Final census

Year of Input Data (4)

Admin Level of Input Data (5)

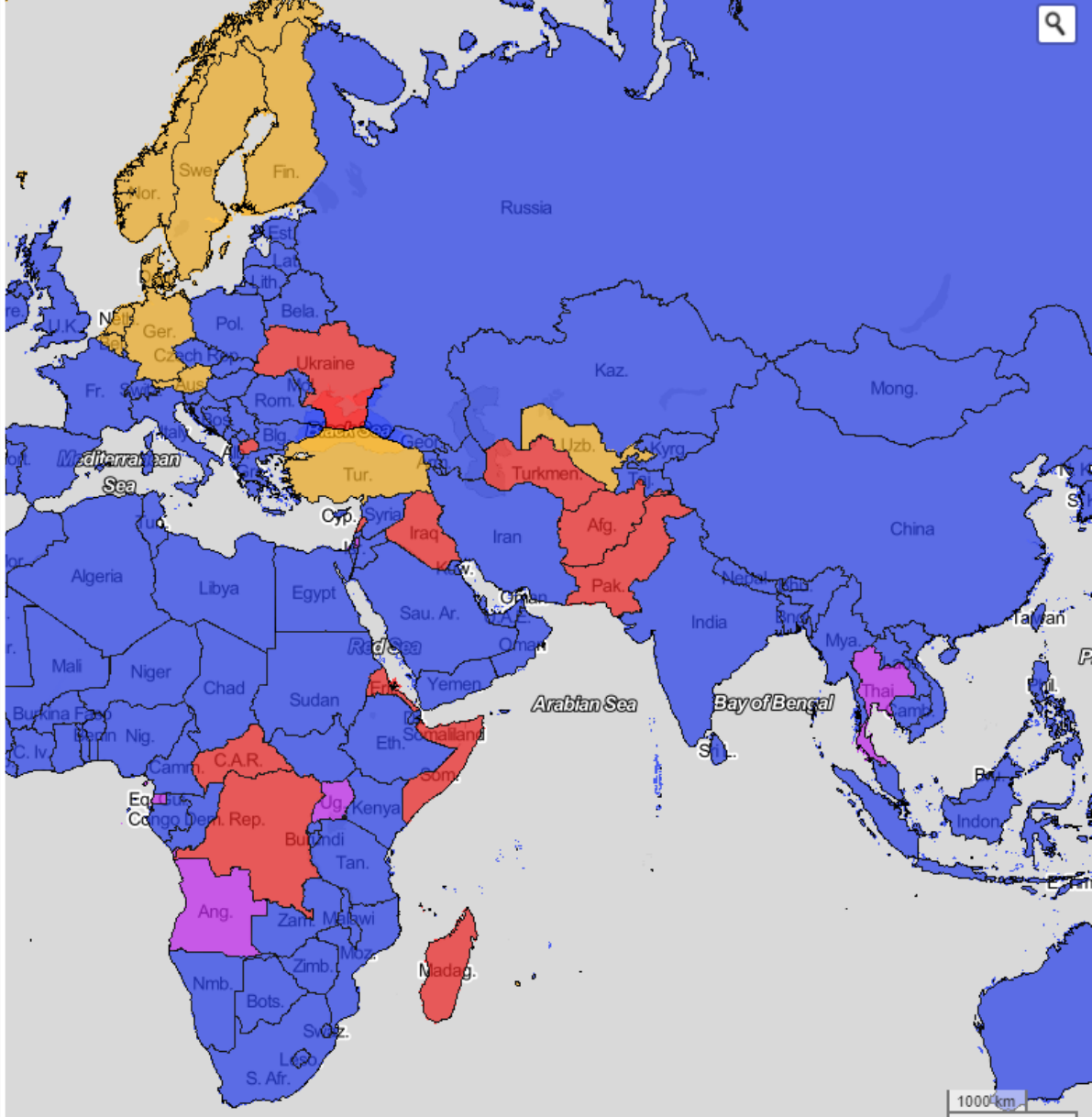
Growth Rate Start Year (6)

Growth Rate End Year (7)

Growth Rate Admin Level (8)

Year of Last Census (9)

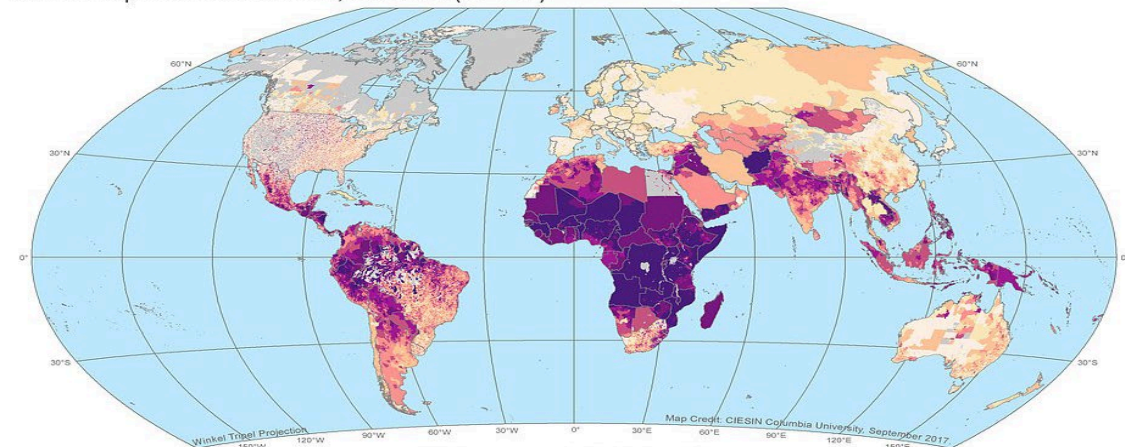
Average Size of Input Geography (10)



GPWv4.10, Population by Age Group & Sex, 2010

Basic Demographic Characteristics, v4.10, 2010: Children (Ages 0–14)

Gridded Population of the World, Version 4 (GPWv4)



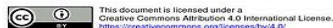
Children as a percent of the total population

<15 15–20 20–25 25–30 30–35 35–40 40–45 >45 No Data

Gridded Population of the World, Version 4 (GPWv4) Basic Demographic Characteristics, Revision 10 consists of estimates of human population by age and sex, consistent with national censuses and population registers, for the year 2010. The global distribution of children, ages 0–14, is represented here as a percent of the total population in the year 2010. It is calculated by summing the counts in the v4.10 5-year age group rasters from ages 0 to 14, then dividing the sum by the total population in the year 2010 and multiplying the quotient by 100.

Center for International Earth Science Information Network - CIESIN - Columbia University, 2017. Gridded Population of the World, Version 4 (GPWv4): Science Information Network Basic Demographic Characteristics, Revision 10, Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H4SH7D7F>.

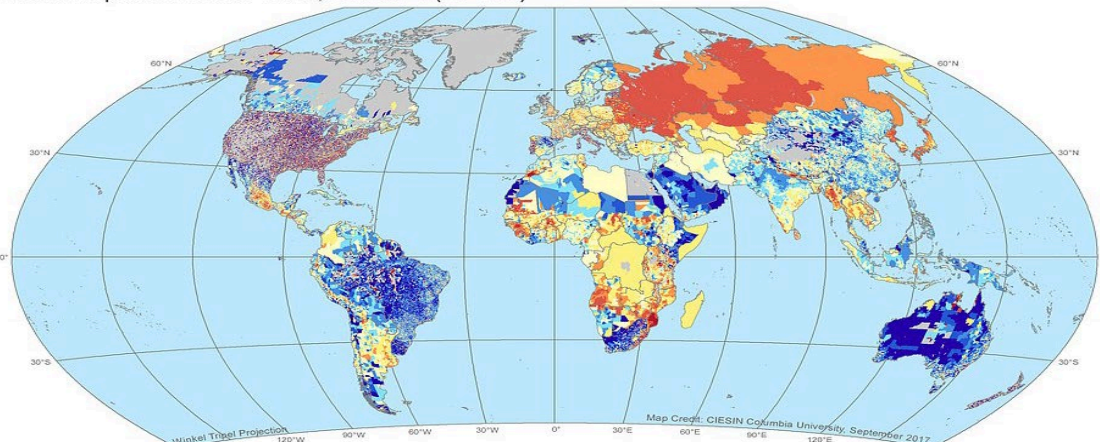
© 2017. The Trustees of Columbia University in the City of New York.



GPWv4.10 includes new dataset, Basic Demographic Characteristics, with gridded estimates of population by 5-year and broad age groups and sex for the year 2010.

Basic Demographic Characteristics, v4.10, 2010: Sex Ratio

Gridded Population of the World, Version 4 (GPWv4)



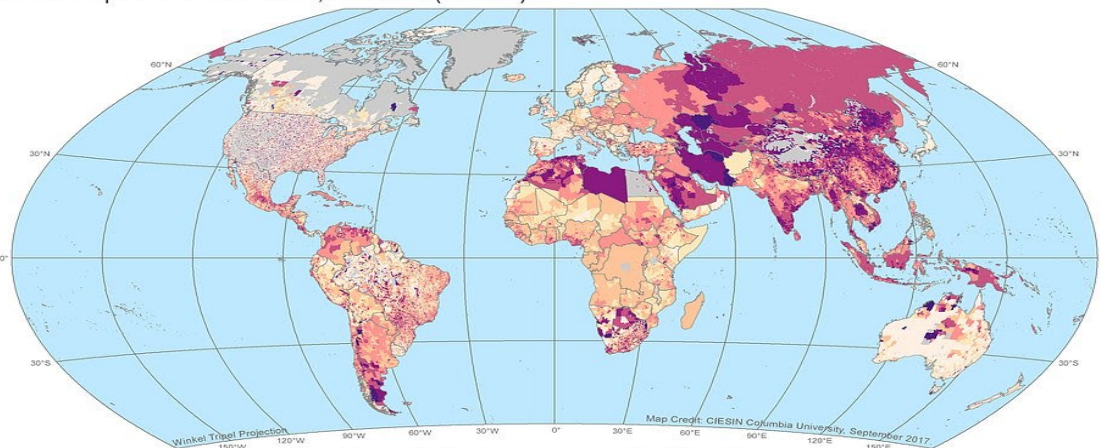
Number of males per 100 females

<80 80–90 90–95 95–100 100–105 105–110 110–120 120–130 >130 No Data

Gridded Population of the World, Version 4 (GPWv4) Basic Demographic Characteristics, Revision 10 consists of estimates of human population by age and sex, consistent with national censuses and population registers, for the year 2010. The sex ratio represents the number of males for every 100 females in the population. It is calculated by dividing the total male count raster by the total female count raster and multiplying the quotient by 100.

Basic Demographic Characteristics, v4.10, 2010: Women of Childbearing Age (Ages 15–49)

Gridded Population of the World, Version 4 (GPWv4)



Women of childbearing age as a percent of the total population

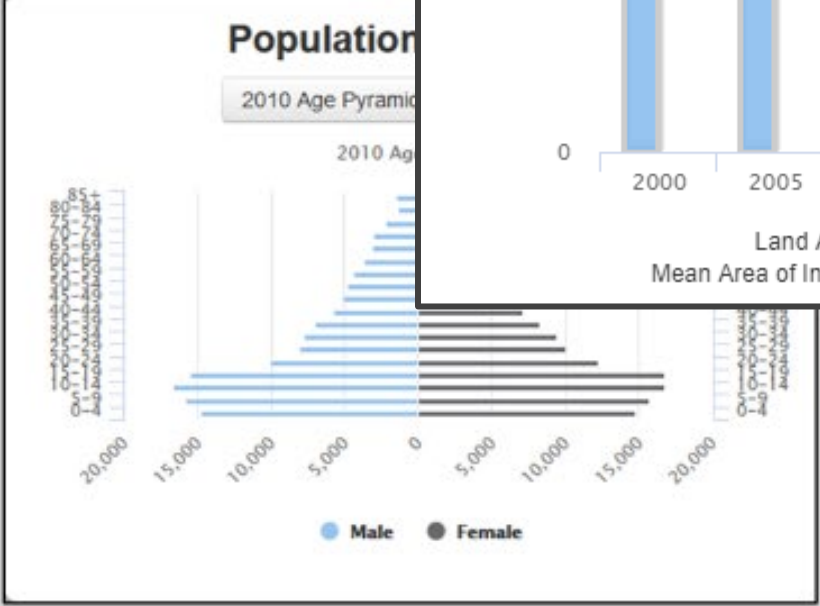
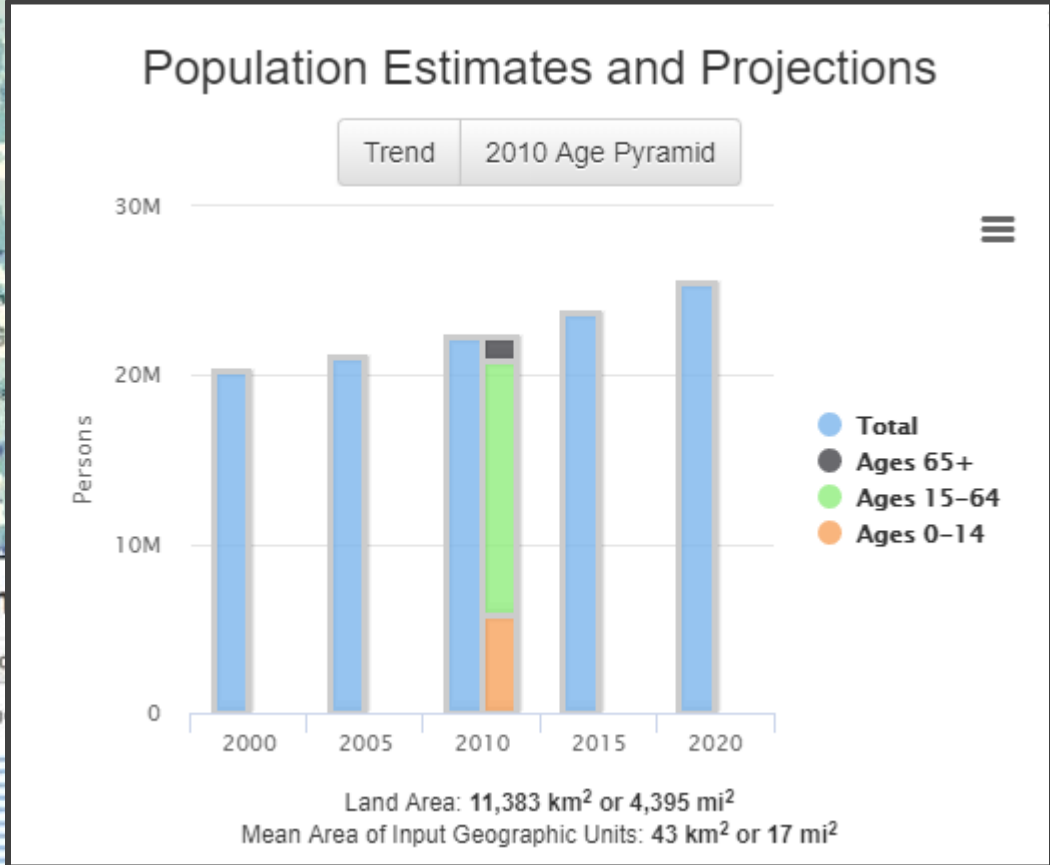
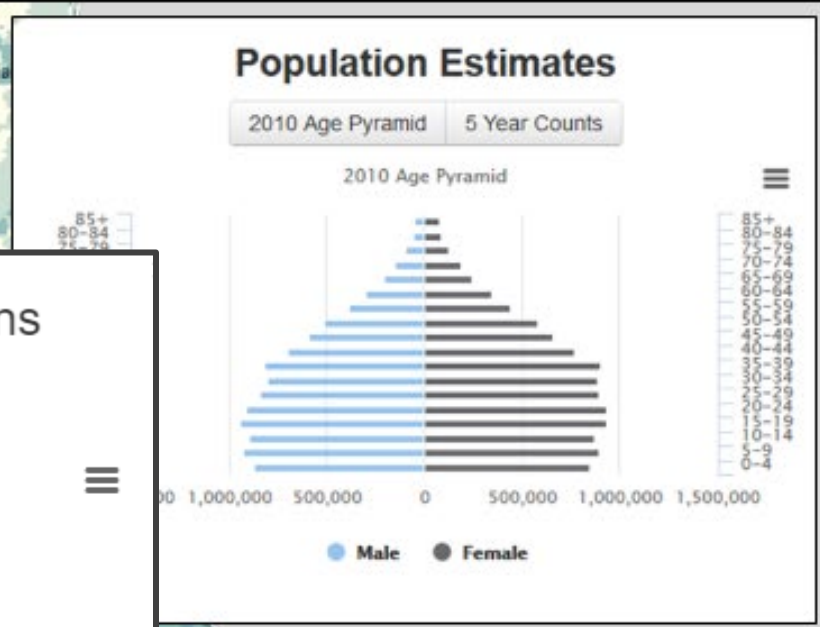
<20 20–22 22–24 24–26 26–28 28–30 >30 No Data

Gridded Population of the World, Version 4 (GPWv4) Basic Demographic Characteristics, Revision 10 consists of estimates of human population by age and sex, consistent with national censuses and population registers, for the year 2010. The global distribution of women of childbearing age, ages 15–49, is represented here as a percent of the total population in the year 2010. It is calculated by summing the counts in the v4.10 female 5-year age group rasters from ages 15 to 49, then dividing the sum by the total population in the year 2010 and multiplying the quotient by 100.

<http://sedac.ciesin.columbia.edu/data/collection/gpw-v4/sets/browse>

Mexico Input Data

Year	2010
Administrative Level	3
Source	Final Census, INEGI





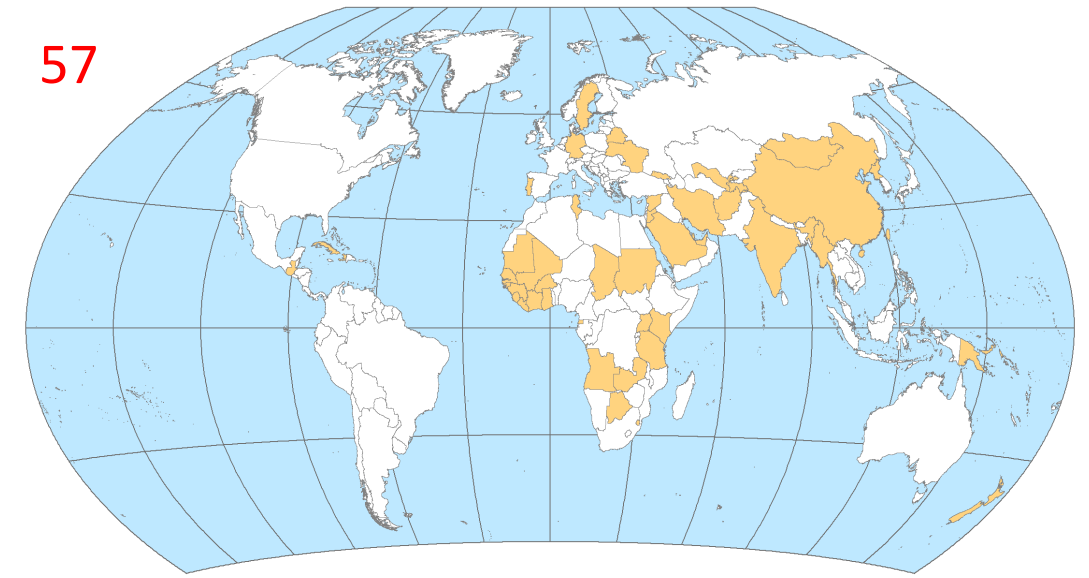
Age and/or sex input data at a coarser resolution than total population

Countries for which input data for age are at a coarser resolution than for sex.

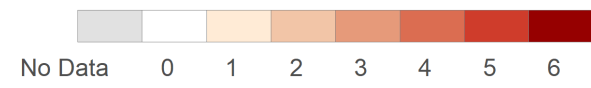
84



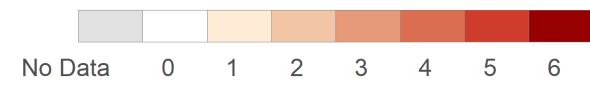
57



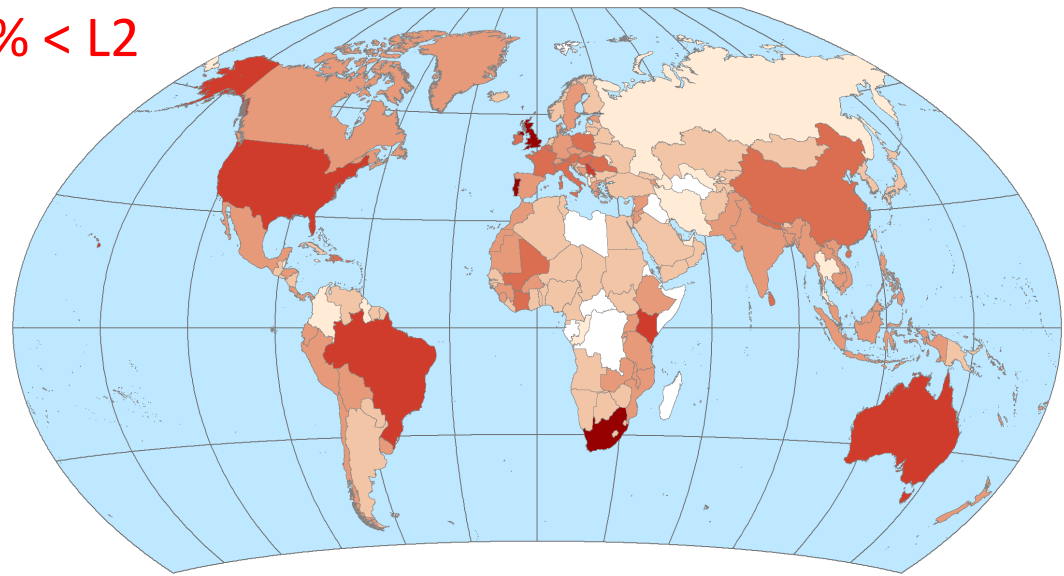
Administrative Level of Sex Input Data



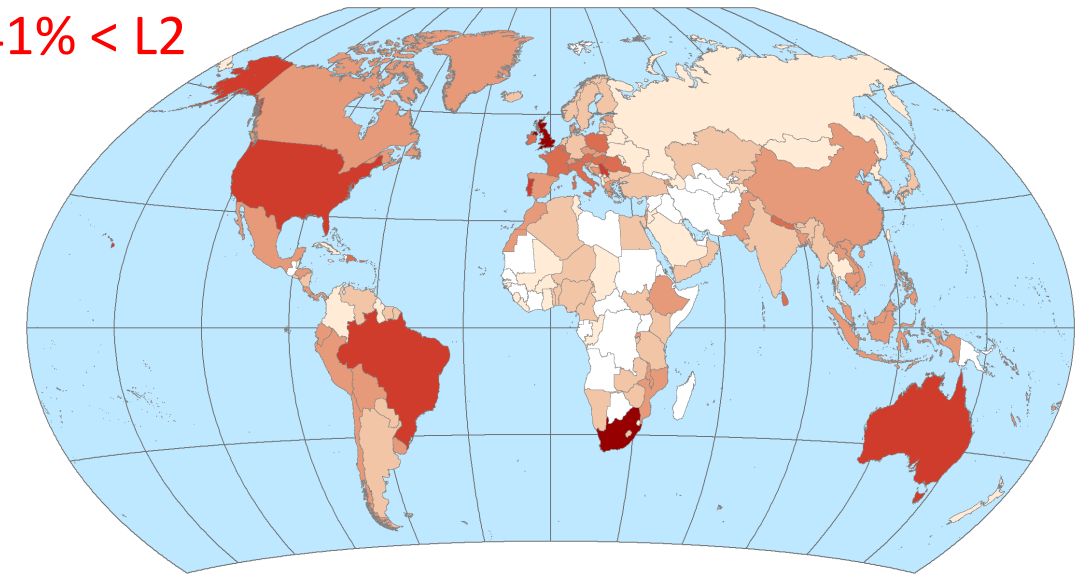
Administrative Level of Age Input Data



27% < L2



41% < L2





Ongoing Population Work

- JRC Global Human Settlement Layer (**GHSL**) – forthcoming R2018, currently in beta.
- DLR Global Urban Footprint (**GUF**) – forthcoming version 2
- University of Louisville/CIESIN collaboration to examine population covariates with detailed census data from Mexico and India.
- Facebook/CIESIN collaboration to bring the High Resolution Settlement Layer (**HRSL**) to (near) global coverage
- “**Human Planet**” projects funded by NASA to investigate population movements with night-time lights (VIIRS) data, integrate population/settlement data with selected SDG indicators, and map “missing” populations
- Geo-referenced Infrastructure and Demographic Data for Development (**GRID³**) initiative funded by BMGF and UK DFID, to build national capacity to collect, analyze, integrate, disseminate, and utilize high-resolution population, infrastructure, and other reference data in support of sustainable development and humanitarian response priorities

The work being done by the EFGS community to develop, document, and share fundamental data on settlements, population, infrastructure, and related themes is essential to the achievement of the SDGs.

Thank you!

Linda Pistolesi

LPISTOLE@CIESIN.COLUMBIA.EDU

