# Documentation for the Sustainable Development Goal Indicators (SDGI) Data Sets, 2023 Release

**July 2023**

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

## Abstract

This document introduces the Sustainable Development Goals Indicators (SDGI) collection, and describes the input data and methodologies used to produce the first four data sets in the 2023 Release. The data sets in the SDGI collection represent Sustainable Development Goal (SDG) Indicators 7.1.1, 9.1.1, 11.2.1, and 11.7.1. Using global open data, each indicator was tabulated using a series of Python scripts that apply zonal statistics to compute a population's access to a resource or amenity considered essential for people's prosperity and achievement of the United Nations (UN) SDGs. The Introduction describes the purpose of the SDGI collection and presents an overview of each indicator and their parent SDG. The Data and Methodology section provides details on the input data and the individual methodologies used to produce each indicator data set. The Data Set Descriptions section includes information about each data set including a code book of attributes and the file formats available. Potential use cases, limitations, and constraints of the 2023 Release data sets are briefly discussed. The development of the data sets included in this collection serves to demonstrate the potential of global open data in producing timely, transparent, and standardized data sets to help countries monitor their progress towards SDGs and to facilitate international comparisons.

**Suggested citation for this document:** Center for International Earth Science Information Network (CIESIN), Columbia University. 2023. Documentation for the Sustainable Development Goal Indicators (SDGI) Data Sets, 2023 Release. Palisades NY: NASA Socioeconomic Data and Applications Center. https://doi.org/10.7927/zc4h-hh18. Accessed Day MONTH YEAR.

We appreciate feedback regarding this data collection, such as suggestions, discovery of errors, difficulties in using the data, and format preferences. Please contact:

NASA Socioeconomic Data and Applications Center (SEDAC)
Center for International Earth Science Information Network (CIESIN)
Columbia University
Phone: 1 (845) 365-8920
Email: ciesin.info@ciesin.columbia.edu

## Contents



## I. Introduction

*The Sustainable Development Goals Report 2022* highlighted the persistence of data gaps and their effects on United Nations (UN) Sustainable Development Goal (SDG) monitoring (UN DESA, 2022). The Center for International Earth Science Information Network (CIESIN) has produced the Sustainable Development Goals Indicators (SDGI) collection to provide globally consistent, regularly updated, operational data sets of SDG indicators. The 2023 Release includes four data sets. These data sets, representing SDG indicators 7.1.1, 9.1.1, 11.2.1, and 11.7.1, exemplify the potential of non-traditional global open data sources in bridging data gaps to help monitor SDG progress. CIESIN plans to release regular updates of each indicator data set and is currently developing methodologies to compute additional indicator data sets to be included in subsequent releases. These data sets were developed in support of the Group on Earth Observations (GEO) Human Planet Initiative (HPI).

UN SDG 7 is “ensure access to affordable, reliable, sustainable and modern energy for all” (UN DESA, n.d.a). *Tracking SDG 7: The Energy Progress Report* estimated that in 2019, 759 million people around the world lacked access to electricity. Moreover, due to current policies and the detrimental effects of the COVID-19 crisis, it is predicted that by 2030, 660 million people will still not have access to electricity, with a majority of these people residing in Sub-Saharan Africa (United Nations, n.d.b). As one measure of progress, the UN agreed on SDG indicator 7.1.1: “Proportion of population with access to electricity” (United Nations, 2022). The SDG indicator 7.1.1 data set included in the SDGI collection provides estimates for the proportion of population with access to electricity for 206 countries and 45,979 level 2 subnational units. The data set is available at both national and level 2 subnational resolutions.

UN SDG 9 is "build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation" (UN DESA, n.d.c). Addressing inadequate access to roads, especially in rural areas, is critical to achieving SDG 9. Sustainable transportation helps to eliminate poverty, promote food security, improve access to key health services, increase trade competitiveness, and bolster human rights (United Nations, 2021a). As one measure of progress, the UN has established SDG indicator 9.1.1: "Proportion of the rural population who live within 2 km of an all-season road" (United Nations, 2020). This indicator is commonly known as the Rural Access Index (RAI) (Workman and McPherson, 2020; World Bank, n.d.). The SDG indicator 9.1.1 data set included in the SDGI collection provides estimates for the proportion of the rural population with access to all-season roads for 209 countries and 45,073 subnational units. The data set is available at both national and level 2 subnational resolutions.

UN SDG 11 is "make cities and human settlements inclusive, safe, resilient and sustainable" (UN DESA, n.d.d). Improving access to public transport services is integral to achieving the objectives of SDG 11. According to the UN *Sustainable Transport, Sustainable Development 2021 interagency report*, "only about half the world's urban population have convenient access to public transport" (United Nations, 2021a). The report highlights that access to sustainable transport can help reduce food insecurity, boost economies, empower women, and connect people to key health, education, and financial services (United Nations, 2021a). As one measure of progress, the UN has established SDG indicator 11.2.1: "Proportion of population that has convenient access to public transport, by sex, age and persons with disabilities" (United Nations, 2021b). The SDG indicator 11.2.1 data set included in the SDGI collection provides estimates for the proportion of population with convenient access to public transport for 5,749 urban centers across 178 countries.

SDG indicator 11.7.1 is also important in measuring progress towards SDG 11. SDG indicator 11.7.1 is defined as the "Average share of the built-up area of cities that is open space for public use for all, by sex, age and persons with disabilities" (United Nations, 2021c). The SDG indicator 11.7.1 data set included in the SDGI data set collection provides estimates for both the proportion of urban centers that are designated public space and the urban population's access to Open Public Space (OPS). SDG indicator 11.7.1 was computed for 8,873 urban centers across 180 countries.

All four SDGI data sets can be accessed as Web Feature Services (WFS), or downloaded in GIS compatible or tabular file formats. For the latest data releases, collection information, and other important news, follow SEDAC and/or CIESIN on any of the following:

- Twitter: http://www.twitter.com/ciesin
- Facebook: https://www.facebook.com/socioeconomicdataandappsctr
- SEDAC YouTube channel: https://www.youtube.com/@nasasocioeconomicdataandap4874

# II. Data and Methodology

An appropriate methodology for each indicator was developed based on the UN SDG Indicators Metadata repository and additional literature review. Each methodology was implemented through a series of multiprocessing Python scripts (Python Software Foundation, 2022a; Python Software Foundation, 2022b). SDG indicator 7.1.1 measures access to electricity and was computed as the proportion of population within illuminated areas defined by annual VIIRS Nighttime Lights Version 2 (VNL V2) data (Elvidge et al., 2021). SDG indicator 9.1.1 measures rural access to all-season roads and was computed as the population within rural areas defined by the Global Human Settlement Layer Settlement Model grid (GHS-SMOD) that were within 2 kilometers of a subset of OpenStreetMap (OSM) roads (Pesaresi et al., 2019; OSM, n.d.). SDG indicator 11.2.1 measures convenient access to public transport and was computed as the population of a given urban center within 0.5 kilometer of a low capacity OSM public transport point or within 1 kilometer of a high capacity public transport point (United Nations, 2021b). Lastly, SDG indicator 11.7.1 measures the average share of built-up areas within a defined urban center extent that are public space, as well as the urban population's access to Open Public Space (OPS). SDG indicator 11.7.1 was computed by estimating OSM road and public space area within urban center polygons and tabulating the proportion of the urban population within 400 meters to OPS (United Nations, 2021c). For both SDG indicators 11.2.1 and 11.7.1, urban center extents were obtained from the GHS Urban Center Database (GHS-UCDB) (Florczyk et al., 2019). SDG indicators 7.1.1 and 9.1.1 are reported at administrative level 0 (national) and administrative level 2 (subnational), while SDG indicators 11.2.1 and 11.7.1 are reported for urban centers. WorldPop's Unconstrained Individual Countries 2020 UN Adjusted (100 meter resolution) data set was the input gridded population data set for all four indicators (WorldPop and CIESIN, 2020). It was the most current global gridded population data available at the time of production, however, it is expected that future releases of gridded population data and our automated processing scripts will allow for regular updates to the SDGI collection.

## Input data

The SDGI collection was produced using multiple input data sets (Table 1), including administrative boundaries, urban center extent boundaries, gridded population data, nighttime lights data, roads, public spaces, public transport points, and United Nations population estimates. The sources and applications of the input data are described below.

### Boundaries

The Database of Global Administrative Areas (GADM) v3.6 was used for delineating national and level 2 subnational boundaries for indicators 7.1.1 and 9.1.1 (GADM, n.d.). The Global Human Settlement Layer's (GHSL) Settlement Model (GHS-SMOD) R2019 was used for classifying rural areas for SDG indicator 9.1.1 (Pesaresi, 2019). The GHSL Urban Center Database (GHS-UCDB) R2019A was used for defining the boundaries of urban centers used in computing SDG indicators 11.2.1 and 11.7.1 (Florczyk, 2019).

More specifically, the GHS-UCDB was selected because it applies the principles of the Degree of Urbanization (DEGURBA), a UN-endorsed method for characterizing the urban-rural continuum (United Nations Statistical Commission, 2020).

**Nighttime Lights**
The Visible Infrared Imaging Radiometer Suite (VIIRS) is a collection of scientific instruments integrated on the NASA/NOAA Suomi National Polar-orbiting Partnership (Suomi NPP) and NOAA-20 satellites, which provides global nighttime lights data (NASA, n.d.). The 2021 nighttime lights data used as input to SDG indicator 7.1.1 were provided by the Earth Observation Group at The Payne Institute for Public Policy, Colorado School of Mines. The annual VIIRS Nighttime Lights Version 2 (VNL V2) (15 arc-second resolution) product was used because it incorporates additional statistical processing that attempts to further remove light from fires and other unwanted pixels (Elvidge et al., 2021; Earth Observation Group, n.d.).

**Roads, Public Spaces, and Public Transport Points**
OpenStreetMap (OSM) is a crowd-sourced mapping project, in which people around the world contribute to digitizing the planet. OSM data used in this effort were downloaded on 3/31/2022 from https://download.osmdata.xyz. Data were downloaded by OSM key. GeoPackages downloaded include Highway, Landuse, Leisure, Natural, and Public Transport (OSM, n.d., osmdata.xyz, n.d.).

**United Nations Population Estimates**
For both SDG indicators 7.1.1 and 9.1.1, the United Nations *World Population Prospects: The 2019 Revision*, referred to as UNWPP 2019 in this document, was used in quality checks of the indicator data sets (UN DESA Population Division, 2019). Total country population values derived from zonal statistics using the underlying WorldPop data set were compared to total population values reported in UNWPP 2019. Countries with a percent error greater than 5% were given an error code in the data set.

**Gridded Population Data**
WorldPop is an innovative organization focused on developing and supporting spatial demographic data sets. For the purposes of this project, gridded population data were downloaded per country from WorldPop's Unconstrained Individual Countries 2020 UN Adjusted (100 meter resolution) data set. This data set implements a Random Forest-based dasymetric redistribution model, in order to map population distributions as pixels for each country. After applying this mapping approach, WorldPop then adjusts the results according to the official United Nations population estimates reported in the 2019 Revision of World Population Prospects. WorldPop raster data sets were used as the input gridded population data sets for the zonal statistics conducted in each SDG indicator methodology (WorldPop and CIESIN, 2020).

*Table 1. Input Data Sets*

| Data Set | Data Source | Time Frame | Format | Resolution |
|---|---|---|---|---|
| OpenStreetMap (OSM) | osmdata.xyz | Downloaded 03/31/22 | Vector | N/A |
| Unconstrained Individual Countries 2020 UN Adjusted | WorldPop Hub | 2020 | Raster | 100 m |
| The Database of Global Administrative Areas (GADM) Version 3.6 | gadm.org | 2020 | Vector | N/A |
| Annual VIIRS Nighttime Lights Version 2 (VNL V2) | eogdata.mines.edu | 2020 | Raster | 15 arc-second |
| GHS Urban Center Database (GHS-UCDB) R2019A | Global Human Settlement Layer (GHSL) | 2015 | Vector | 1 km |
| GHS Settlement Model Grid (GHS-SMOD) R2019A | Global Human Settlement Layer (GHSL) | 2015 | Raster | 1 km |

## Methods

Each indicator methodology was automated using a series of Python scripts (Python Software Foundation, 2022a; Python Software Foundation, 2022b), and Esri ArcPy site package modules (Esri, 2022). All input data sets used in the creation of the SDGI data set collection are open data and all scripts are available in a GitHub repository. All spatial data were projected to the WGS84 Geographic Coordinate System before being used in any processing.

1. *SDG Indicator 7.1.1*

   SDG indicator 7.1.1 measures access to electricity by computing the proportion of population within illuminated areas defined by VIIRS nighttime lights data. In 2020, Falchetta et al. published a paper in which they used nighttime lights data to assess electrification and progress towards SDG 7 in Sub-Saharan Africa (Falchetta et al., 2020). Falchetta et al.'s methodology was adapted and simplified in order to create the SDG indicator 7.1.1 data set.

   First, the Esri ArcGIS Select tool was used to extract GADM national and administrative level 2 subnational boundaries for a given country. Then the 2021

VNL V2 raster data set was clipped to the national boundary. Next, the clipped national VNL V2 data were reclassified to generate a binary raster model representing zones with or without light per subnational unit. The binary raster was converted into a polygon feature class and intersected with the GADM subnational units to generate level 2 subnational zones with and without lights. Next, the Esri ArcGIS Select Layer By Attribute tool was used to create separate feature classes for admin 2 zones with lights and without lights. Zones with lights were assumed to represent areas with access to electricity. Zonal Statistics were performed for both the feature class of illuminated admin 2 zones and the feature class of unilluminated admin 2 zones. Additional Zonal Statistics were run to obtain the total population per admin 2 unit. WorldPop gridded population data corresponding to the given country were used in each Zonal Statistics operation.

SDG indicator 7.1.1 was calculated by dividing the population with access to electricity by the total population of a given area and multiplying by 100 (Figure 1). The subnational results were summed to estimate SDG indicator 7.1.1 nationally. If a country's total population based on the WorldPop gridded population data set resulted in a percent error greater than 5 % when compared to the 2019 UNWPP, the country results were given an error code. This work was completed previously using 2015 and 2020 VNL V2 data and the results compared well with electrification data reported by the World Bank (IEA et al., n.d.). A comparison for the results obtained using 2021 VNL V2 data could not be made because the World Bank has not yet released 2021 electrification data.

*2. SDG Indicator 9.1.1*

SDG indicator 9.1.1 measures rural access to all-season roads. First, the Esri ArcGIS Select tool was used to extract GADM national and administrative level 2 subnational boundaries for a given country. The national boundary was then used to clip GHS-SMOD and to extract OSM roads data using the Esri ArcGIS Select Layer By Location tool. For each country, OSM highway tags were used as a proxy for road condition data in order to identify which road features were most likely all-season roads. Tags from the OSM Highway, Access, and Surface keys were all used to filter extracted OSM roads data to remove any road features that probably did not meet all-season road conditions (Table 2). Next, these roads were buffered, and then the buffers dissolved. The resulting feature class was intersected with level 2 subnational units and GHS-SMOD, to create rural zones that are within 2 kilometers of an all-season road per subnational unit. Zonal statistics were run using these zones and the 2020 WorldPop gridded population data set to compute the population per zone. Finally, the RAI for each subnational unit was calculated by dividing the total rural population within 2 kilometers of an all-season road by the total corresponding rural population and multiplying by 100. Subnational data were aggregated to compute indicator 9.1.1 at the national level. If a country's total population based on the WorldPop gridded population data set resulted in a percent error greater than 5% when compared to the 2019 UNWPP, the country results were given an error code.

*Table 2. OSM Keys and Tags Used to Identify All-Season Roads for SDG Indicator 9.1.1*

| Key | Condition | Tags |
|---|---|---|
| Highway | *Keep* | primary, primary link, secondary, secondary link, tertiary, tertiary link, trunk, trunk link, motorway, motorway link, road, unclassified, track (grade 1 only) |
| Access | *Remove* | no, private |
| Surface | *Remove* | unpaved, compacted, fine gravel, gravel, rock, pebblestone, ground, dirt, earth, grass, grass paver, mud, sand, woodchips, snow, ice, salt |

*3. SDG Indicator 11.2.1*

SDG indicator 11.2.1 monitors access to public transport in cities. In order to help simplify the scope of this indicator, the UN SDG indicator 11.2.1 metadata characterizes convenient access to public transport as either 0.5 kilometer walking distance to a low-capacity public transport point or 1 kilometer walking distance to a high-capacity public transport point (United Nations, 2021b).

First, the Esri ArcGIS Select tool was used to extract urban centers per country from the GHS-UCDB using the ISO field. Then the Esri ArcGIS Select Layer By Location tool was used to extract OSM public transport points using a country's urban centers. OSM public transportation points within each GHS-UCDB urban center extent were filtered and labeled as either low-capacity or high-capacity points based on OSM tags. OSM tags from the Amenity, Railway, Bus, and Highway keys were all used to categorize public transport points (Tables 3 and 4). Then these points were either buffered by 0.5 kilometer or 1 kilometer according to their categorization. The resulting buffers were then dissolved. The dissolved output was then intersected with the urban center extents to create zones of access to public transport per urban center. Zonal statistics were then performed using the 2020 WorldPop gridded population data to tabulate the total number of people with convenient access to public transport for a given city. This value was then divided by the total population of the urban center, previously computed with zonal statistics, and multiplied by 100 to estimate SDG indicator 11.2.1.

*Table 3. OSM Keys and Tags Used to Identify Acceptable Low and High Capacity Public Transport Points for SDG Indicator 11.2.1.*

| **OSM Key** | **OSM tag(s)** | **High or Low Capacity *Public Transport Points*** |
|---|---|---|
| Amenity | bus_station | Low Capacity |
| Amenity | ferry_terminal | High Capacity |
| Railway | tram_stop | Low Capacity |
| Railway | stop, stop_position, platform, station, subway_entrance | High Capacity |
| Bus | yes | Low Capacity |
| Highway | bus_stop | Low Capacity |

*4. SDG Indicator 11.7.1*

SDG indicator 11.7.1 measures both an urban center's share of public space and access to its Open Public Space (OPS). Public space includes both roads and OPS. Since SDG indicator 11.7.1 is concerned with area, input data should ideally be polygon features. While there are a variety of OSM polygon tags used to describe OPS, OSM highways data are packaged as line features. In order to estimate the spatial contribution of OSM roads to SDG indicator 11.7.1, road widths were required to approximate road area. Therefore, before any analysis was conducted, a Regional Median Road Widths (RMRW) database was created based on the OSM Highway key's width tag. Adhering to the geographic regions used by the UN Statistics Division, median regional road width values were tabulated from the OSM Highway GeoPackage and stored in a CSV file (UNSD, n.d.).

First, the Esri ArcGIS Select tool was used to extract urban centers per country from the GHS-UCDB using the ISO field. Then the Esri ArcGIS Select Layer By Location tool was used to extract OSM OPS and road data using a country's urban extents. OSM OPS data came from OSM Leisure, Landuse, and Natural GeoPackages and were filtered based on OSM tags (Table 4). After filtering, data extracted from the Leisure, Landuse, and Natural GeoPackages were merged together to form one feature class representing all OPS polygons in a country's urban centers. The OPS feature class was then clipped to the country's urban center extents.

Width values from the RMRW database were used to buffer the extracted roads. These buffered roads were then dissolved and clipped to the country's urban center extents. Next, road and OPS features were merged and dissolved into a single polygon

representing all public space in the country. The resulting public space feature class was then intersected with the country's urban center extents to get the area of public space within each urban center. To compute the urban center's share of public space, the total area of the intersected public space within each urban center was divided by the total area of the corresponding urban center and multiplied by 100.

In order to compute access to OPS, OSM OPS polygons were buffered by 400 meters, as specified by the UN SDG indicator 11.7.1 metadata, and dissolved into a single polygon (United Nations, 2021c). The resulting feature class was clipped to and intersected with the country's urban center extents to create zones of access to OPS. Lastly, zonal statistics were conducted using the country's 2020 WorldPop gridded population data to calculate the number of people with access to OPS per urban center. This value was then converted into a percentage, by dividing by the total population of a given urban center and multiplying by 100.

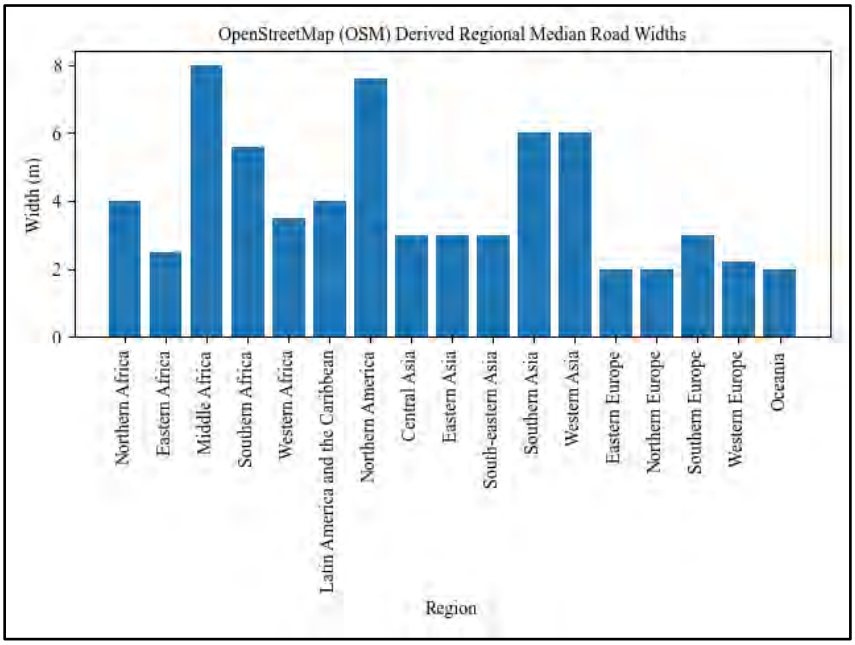


*Figure 1. Regional Median Road Widths*

*Table 4. OSM Keys and Tags Used to Identify Acceptable OPS and Road Features for SDG Indicator 11.7.1.*

| Key | Condition | Tags |
|---|---|---|
| Leisure | *Keep* | park, nature_reserve, playground, common, garden, allotments, recreation_ground, pitch, dog_park, fitness_station |
| Landuse | *Keep* | forest, village_green, recreation_ground, allotments, conservation |
| Natural | *Keep* | fell, grassland, heath, scrub, wood |
| Highway | *Keep* | motorway, trunk, primary, secondary, tertiary, unclassified, residential, motorway_link, trunk_link, pirmary_link, secondary_link, tertiary_link, living_street, pedestrian, road, busway, sidewalk |

## III. Data Set Description(s)

All data sets in the SDGI data collection are stored in a WGS84 Geographic Coordinate System, and are available as Web Feature Service (WFS) services. In addition, data sets are available in GIS (Geodatabase and Shapefile) and tabular (CSV) formats. SDG indicator 7.1.1 and 9.1.1 data sets are available at both national and level 2 subnational resolutions. SDG indicator 11.2.1 and 11.7.1 data sets are available at the urban center level.

Each data set includes error codes, described in the table below, to indicate units for which the indicator could not be computed.

**Codebook for SDG Indicator Error Codes**

| Error Code | Description |
|---|---|
| -9997 | Population Percent Error Greater than 5 %. The indicator is not reported for these units |
| -9998 | There is a population value of zero in the denominator therefore the indicator could not be computed. |
| -9999 | Indicator could not be computed due to missing input data or inaccurate numerator. The latter will be remedied in a future version of the data set. |

## 1. SDG Indicator 7.1.1: Access to Electricity, 2023 Release

**Data set description:**
The SDG Indicator 7.1.1: Access to Electricity, 2023 Release data set, part of the Sustainable Development Goal Indicators (SDGI) collection, measures the proportion of the population with access to electricity for a given statistical area. The indicator was computed as the proportion of WorldPop gridded population located within illuminated areas defined by annual VIIRS Nighttime Lights Version 2 (VNL V2) data. The SDG indicator 7.1.1 data set included in the SDGI collection provides estimates for the proportion of population with access to electricity for 206 countries and 45,979 level 2 subnational units

**Codebook for National Level SDG Indicator 7.1.1: Access to Electricity**

| Field Name | Alias | Description |
|---|---|---|
| ISO3 | ISO3 Code | International Standards Organization (ISO) unique three-letter country or area code |
| NAME_0 | Country/Area Name | Country/Area name |
| Total_Pop | Total Population of Country/Area | Total population of the country/area |
| Pop_Elec | Population with Access to Electricity | Population in the country/area with access to electricity |

| | | |
|---|---|---|
| SDG711pct | Percent of Population with Access to Electricity | Percent of population in the country/area with access to electricity |
| PCT_ERROR | Population Percent Error | Total population percent error when compared to UNWPP 2019 |

**Codebook for Subnational Level SDG Indicator 7.1.1: Access to Electricity**

| Field Name | Alias | Description |
|---|---|---|
| ISO3 | ISO3 Code | International Standards Organization (ISO) unique three-letter country or area code |
| NAME_0 | Country/Area Name | Country/Area name |
| NAME_1 | Admin 1 Name | Name of administrative level 1 |
| NAME_2 | Admin 2 Name | Name of administrative level 2 |
| Pop_A2 | Population in Admin 2 | Total population of administrative level 2 unit |
| Pop_A2_E | Population in A2 w/Electricity | Population in administrative level 2 unit with access to electricity |
| POP_A2_Err | Error in Admin 2 Population | Difference in administrative level 2 based on VNL V2 zones versus GADM |
| SDG711_Pct | Percent of A2 Pop w/Electricity | Percent of population in administrative level 2 unit with access to electricity |
| PCT_ERROR | Country/Area Population Percent Error | Total country/area population percent error when compared to UNWPP 2019 |

**Data set web page:**
SEDAC URL: https://sedac.ciesin.columbia.edu/data/set/sdgi-7-1-1-access-electricity-2023
Permanent URL: https://doi.org/10.7927/gxnr-sx57

**Data set format:**
The SDG Indicator 7.1.1: Access to Electricity, 2023 Release data set is available in Esri file Geodatabase (GDB), Shapefile (SHP), and CSV formats. Each downloadable zip file contains: 1) global layers for the national and subnational indicator in the selected file format, and 2) PDF Documentation. The data are also accessible as WFS services.

**Data set downloads:**
sdgi-7-1-1-access-electricity-2023-csv
sdgi-7-1-1-access-electricity-2023-gdb
sdgi-7-1-1-access-electricity-2023-shp

**2. SDG Indicator 9.1.1: The Rural Access Index (RAI), 2023 Release**

**Data set description:**
The SDG Indicator 9.1.1: The Rural Access Index (RAI), 2023 Release data set, part of the SDGI collection, measures the proportion of the rural population who live within 2 kilometers of an all-season road for a given statistical area. The indicator was computed as the proportion of WorldPop gridded population within 2 kilometers to an OpenStreetMap (OSM) all-season road. The SDG indicator 9.1.1 data set included in the SDGI collection provides estimates for the proportion of the rural population with access to all-season roads for 209 countries and 45,073 subnational units.

**Codebook for National Level SDG Indicator 9.1.1: The Rural Access Index (RAI)**

| Field | Alias | Description |
|---|---|---|
| ISO3 | ISO3 Code | International Standards Organization (ISO) unique three-letter country or area code |
| NAME_0 | Country/Area Name | Country/Area name |
| RAI_Pct | RAI (%) | Rural Access Index (%) for country/area |
| PCT_ERROR | Population Percent Error | Total population percent error when compared to UNWPP 2019 |

**Codebook for Subnational Level SDG Indicator 9.1.1: The Rural Access Index (RAI)**

| Field | Alias | Description |
|---|---|---|
| ISO3 | ISO3 Code | International Standards Organization (ISO) unique three-letter country or area code |
| NAME_0 | Country/Area Name | Country/Area name |

| | | |
|---|---|---|
| NAME_1 | Admin 1 Name | Name of administrative level 1 |
| NAME_2 | Admin 2 Name | Name of administrative level 2 |
| Pop_A2 | Population Admin 2 | Total population in administrative level 2 unit |
| Pop_A2_R | Rural Population Admin 2 | Rural population in administrative level 2 unit |
| Pop_A2_R2K | Rural Pop in Admin2 within 2 KM of Road | Rural population in administrative level 2 unit within 2 kilometers of an all-season road |
| RAI_Pct | RAI (%) | Rural Access Index (%) for administrative level 2 unit |
| PCT_ERROR | Population Percent Error | Total country/area population percent error when compared to UNWPP 2019 |

**Data set web page:**
SEDAC URL: https://sedac.ciesin.columbia.edu/data/set/sdgi-9-1-1-rai-2023
Permanent URL: https://doi.org/10.7927/fcre-m572

**Data set format:**
The SDG Indicator 9.1.1: The Rural Access Index (RAI), 2023 Release data set is available in Esri file Geodatabase (GDB), Shapefile (SHP), and CSV formats. Each downloadable zip file contains: 1) global layers for the national and subnational indicator in the selected file format, and 2) PDF Documentation. The data are also accessible as WFS services.

**Data set downloads:**
sdgi-9-1-1-rai-2023-csv
sdgi-9-1-1-rai-2023-gdb
sdgi-9-1-1-rai-2023-shp

**3. SDG Indicator 11.2.1: Urban Access to Public Transport, 2023 Release**

**Data set description:**
The SDG Indicator 11.2.1: Urban Access to Public Transport, 2023 Release, part of the SDGI collection, measures the proportion of the population in a city that has convenient access to public transport. The indicator was computed as the proportion of WorldPop gridded population within either 0.5 kilometer walking distance to a low-capacity OpenStreetMap (OSM) public transport point or 1 kilometer walking distance to a high-capacity OSM public transport point. Cities were delineated using the European

Commission Joint Research Centre (JRC) Urban Center Database (UCDB). The SDG indicator 11.2.1 data set included in the SDGI collection provides estimates for the proportion of population with convenient access to public transport for 5,749 urban centers across 178 countries.

**Codebook for SDG Indicator 11.2.1: Urban Access to Public Transport**

*Note that field names: CTR_MN_ISO, CTR_MN_NM, UC_NM_MN, and UC_NM_LST all come directly from the GHS Urban Center Database 2015 (UCDB) data set.

| Field | Alias | Description |
|---|---|---|
| CTR_MN_ISO | Main Country Identification: ISO3 | International Standards Organization (ISO) unique three-letter country or area code |
| CTR_MN_NM | Main Country Identification: Name | Country name |
| UC_NM_MN | Name of the Urban Center | Primary name of the urban center |
| UC_NM_LST | List of  Names | Other urban centers included in polygon |
| Total_POP | Total Population of Urban Center | Total population of urban center |
| Access_POP | Population with Convenient Access to Public Transport | Population within 0.5 kilometers of a low-capacity public transport point or 1 kilometer of a high-capacity point |
| SDG1121pct | Percent of Population with Convenient Access to Public Transport | Percent of population within 0.5 kilometers of a low-capacity public transport point or 1 kilometer of a high-capacity point |

**Data set web page:**
SEDAC URL: https://sedac.ciesin.columbia.edu/data/set/sdgi-11-2-1-urban-access-public-transport-2023
Permanent URL: https://doi.org/10.7927/1a5z-3h71

**Data set format:**
The SDG Indicator 11.2.1: Urban Access to Public Transport, 2023 Release data sets are available in Esri file Geodatabase (GDB), Shapefile (SHP), and CSV formats. Each downloadable zip file contains: 1) global layers for the national and subnational indicator

in the selected file format, and 2) PDF Documentation. The data are also accessible as WFS services.

**Data set downloads:**
sdgi-11-2-1-urban-access-public-transport-2023-csv
sdgi-11-2-1-urban-access-public-transport-2023-gdb
sdgi-11-2-1-urban-access-public-transport-2023-shp

**4. SDG Indicator 11.7.1: Urban Public Space, Availability and Access, 2023 Release**

**Data set description:**
The SDG Indicator 11.7.1: Urban Public Space, Availability and Access, 2023 Release, part of the SDGI collection, measures the average share of the built-up area of a city that is open space for public use for all. The indicator was computed by measuring both the proportion of OpenStreetMap (OSM) public space within a given urban center and the proportion of WorldPop gridded population within 400 meters to Open Public Space (OPS). Cities were delineated using the European Commission Joint Research Centre (JRC) Urban Center Database (GHS-UCDB). SDG indicator 11.7.1 was computed for 8,873 urban centers across 180 countries.

**Codebook for SDG Indicator 11.7.1: Urban Public Space, Availability and Access**

*Note that field names: CTR_MN_ISO, CTR_MN_NM, UC_NM_MN, and UC_NM_LST all come directly from the GHS Urban Center Database 2015 (UCDB) data set.

| Field | Alias | Description |
|---|---|---|
| CTR_MN_ISO | Main Country Identification: ISO3 | International Standards Organization (ISO) unique three-letter country or area code |
| CTR_MN_NM | Main Country Identification: Name | Country name |
| UC_NM_MN | Name of the Urban Center | Primary name of the urban center |
| UC_NM_LST | List of Names | Other urban centers included in polygon |
| AREA | Total Area of Urban Center | Total area of urban center in square kilometers |
| AREA_PS | Total Area of Public | Total combined area of |

| | | |
|---|---|---|
| | Space | OPS and roads in urban center in square kilometers |
| SDG1171pct | SDG Indicator 11.7.1 | Percent of urban center that is open space (OPS and roads) for public use |
| Total_POP | Total Population of Urban Center | Total population of urban center |
| Access_POP | Population with Access to OPS | Population within 400 meters of an OPS polygon |
| Pct_Access | Percent of Population with Access to OPS | Percent of Population within 400 meters of an OPS polygon |

**Data set web page:**
SEDAC URL: https://sedac.ciesin.columbia.edu/data/set/sdgi-11-7-1-urban-public-space-availability-access-2023
Permanent URL: https://doi.org/10.7927/eavc-4k45

**Data set format:**
The SDG Indicator 11.7.1: Urban Public Space, Availability and Access, 2023 Release data set is available in Esri file Geodatabase (GDB), Shapefile (SHP), and CSV formats. Each downloadable zip file contains: 1) global layers for the national and subnational indicator in the selected file format, and 2) PDF Documentation. The data are also accessible as WFS services.

**Data set downloads:**
sdgi-11-7-1-urban-public-space-availability-access-2023-csv
sdgi-11-7-1-urban-public-space-availability-access-2023-gdb
sdgi-11-7-1-urban-public-space-availability-access-2023-shp

## IV. How to Use the Data

Data is available in GIS (Geodatabase and Shapefile) and tabular (CSV) formats. The vector and tabular data can be used directly for statistical analysis.

## V. Potential Use Cases

It is expected that the data collection will be of interest to governments seeking to monitor their progress on SDGs and to other interested stakeholders. The SDGI collection will help countries: monitor indicators in a regular and consistent manner; validate, compute, and compare indicators developed using their own sources of data; and compare their progress with their peers. The SDGI collection may also be used by governments or civil society organizations to identify areas within countries that need prioritization. For example, the SDG indicator 9.1.1 data set can help policy makers identify level 2 subnational units that may need increased investment in road construction or rehabilitation. In addition, by demonstrating the feasibility of using global open data to help monitor progress towards SDGs, producers of global data are encouraged to release more frequent updates.

## VI. Limitations

It should be noted that GADM exhibits variable resolution per country. SDG indicator 7.1.1 subnational estimates may be less accurate due to the spatial resolution of the VIIRS Nighttime Lights Version 2 data (~500 meters at the equator) and the use of a simplistic binary raster model (i.e., presence/absence of lights) to identify populated pixels as having or lacking access to electricity. Future iterations of the indicator 7.1.1 data set will try to improve upon the accuracy of the results by advancing this baseline model. Lastly, it should be noted that OSM global digitization is incomplete. For example, in 2016, it was estimated that on a global-scale the OSM road data set was approximately 83% complete (Barrington-Leigh and Millard-Ball, 2017). The quality and completeness of classifications and tagging for roads and other features may also vary both between and within countries.

## VII. Acknowledgments

The Sustainable Development Goal Indicators (SDGI) Data Sets, 2023 Release was completed by CIESIN using data from various sources. The SDGI team included Senior Geographic Information Specialist Linda Pistolesi and Senior Research Staff Assistant James Gibson. Gibson was responsible for data acquisition and processing, and the production of map services and supporting materials, including Story Maps for each indicator. Pistolesi supervised the production of the data sets and supporting materials and collaborated on the data set documentation. John Scialdone coordinated the data set release and reviewed the data set documentation and maps. Merlie Hansen compiled and formatted the data set metadata. Joe Schumacher prepared the data set and documentation delivery mechanism on the SEDAC website

Funding for the development and dissemination of this data set was provided under the U.S. National Aeronautics and Space Administration (NASA) under Grant

80NSSC18K0328, Population and Infrastructure on Our Human Planet, and Contract 80GSFC18C0111 for the Socioeconomic Data and Applications Center (SEDAC), both led by Principal Investigator Robert S. Chen. SEDAC is operated by the Center for International Earth Science Information Network (CIESIN) of Columbia University, and is one of 12 Distributed Active Archive Centers (DAACs) in the NASA Earth Observing System Data and Information System (EOSDIS).

## VIII. Disclaimer

CIESIN follows procedures designed to ensure that data disseminated by CIESIN are of reasonable quality. If, despite these procedures, users encounter apparent errors or misstatements in the data, they should contact SEDAC User Services at ciesin.info@ciesin.columbia.edu. Neither CIESIN nor NASA verifies or guarantees the accuracy, reliability, or completeness of any data provided. CIESIN provides this data without warranty of any kind whatsoever, either expressed or implied. CIESIN shall not be liable for incidental, consequential, or special damages arising out of the use of any data provided by CIESIN.

## IX. Use Constraints



Users are free to use, copy, distribute, transmit, and adapt the work for commercial and non-commercial purposes, without restriction, as long as clear attribution of the source is provided.

## X. Recommended Citation(s)

### Data set(s):

Center for International Earth Science Information Network (CIESIN), Columbia University. 2023. SDG Indicator 7.1.1: Access to Electricity, 2023 Release. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/gxnr-sx57. Accessed DAY MONTH YEAR.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2023. SDG Indicator 9.1.1: Rural Access Index (RAI), 2023 Release. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/fcre-m572. Accessed DAY MONTH YEAR.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2023. SDG Indicator 11.2.1: Urban Access to Public Transport, 2023 Release. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/1a5z-3h71. Accessed DAY MONTH YEAR.

Center for International Earth Science Information Network (CIESIN), Columbia University. 2023. SDG Indicator 11.7.1: Urban Public Space, Availability and Access, 2023 Release. Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/eavc-4k45. Accessed DAY MONTH YEAR.

## XI. Source Code

The geoprocessing tools used for this work rely on the ArcPy Python module from Esri. The scripts used in the production of the SDGI are on Github and can be accessed using the links below:

- Indicator 7.1.1 Code: https://github.com/JIB777/SDG-Indicator-711
- Indicator 9.1.1 Code: https://github.com/JIB777/SDG-Indicator-911
- Indicator 11.2.1 Code: https://github.com/JIB777/SDG-Indicator-1121
- Indicator 11.7.1 Code: https://github.com/JIB777/SDG-Indicator-1171

## XII. References

Barrington-Leigh, C., and A. Millard-Ball. 2017. The world's user-generated road map is more than 80% complete. *PLoS ONE,* 12(8):e0180698. https://doi.org/10.1371/journal.pone.0180698.

Earth Observation Group. n.d. *VIIRS Nighttime Light.* https://eogdata.mines.edu/products/vnl/. Accessed 2022.

Elvidge, C. D., M. Zhizhin, T. Ghosh, F.-C., Hsu, and J. Taneja. 2021. Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sensing,* 13(5):922. https://doi.org/10.3390/rs13050922.

ESRI. 2022. ArcGIS Pro: Release 3.0. Redlands, CA: Environmental Systems Research Institute. http://www.esri.com.

Falchetta, G., S. Pachauri, E. Byers, O. Danylo, and S. C. Parkinson. 2020. Satellite observations reveal inequalities in the progress and effectiveness of recent electrification in Sub-Saharan Africa. *One Earth,* 2(4):364–379. https://doi.org/10.1016/j.oneear.2020.03.007.

Florczyk, A., C. Corbane, M. Schiavina, M. Pesaresi, L. Maffenini, M. Melchiorri, P. Politis, et al. 2019. GHS Urban Centre Database 2015, multitemporal and multidimensional attributes, R2019A. European Commission, Joint Research Centre (JRC). [Dataset]. https://ghsl.jrc.ec.europa.eu/ghs_stat_ucdb2015mt_r2019a.php.
GADM. https://gadm.org/. Accessed 2022.

International Energy Agency (IEA), International Renewable Energy Agency (IRENA), United Nations Statistics Division (UNSD), World Bank, and World Health Organization (WHO ). n.d. World Bank Global Electrification Database from "Tracking SDG 7: The Energy Progress Report". World Bank. https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS.

NASA. Nighttime Lights | Earthdata. https://www.earthdata.nasa.gov/learn/backgrounders/nighttime-lights. Accessed 2022.

OpenStreetMap (OSM). https://www.openstreetmap.org/#map=5/38.007/-95.844. Accessed 3/31/2022.

osmdata.xyz. https://github.com/michaelmgis/osmdata.xyz. Accessed 3/31/2022.

Pesaresi, M., A. Florczyk, M. Schiavina, M. Melchiorri, and L. Maffenini. 2019. GHS settlement grid, updated and refined REGIO model 2014 in application to GHS-BUILT R2018A and GHS-POP R2019A, multitemporal (1975-1990-2000-2015), R2019A. European Commission, Joint Research Centre (JRC) [Dataset]. 10.2905/42E8BE89-54FF-464E-BE7B-BF9E64DA5218 http://data.europa.eu/89h/42e8be89-54ff-464e-be7b-bf9e64da5218.

Python Software Foundation 2022a. IDLE (Python's Integrated DeveLopment Environment) version 3.9.1.1.

Python Software Foundation 2022b. Python programming language, version 3.9.1.1. https://www.python.org.

United Nations. 2020. Indicator 9.1.1, SDG indicator metadata. UN. https://unstats.un.org/sdgs/metadata/files/Metadata-09-01-01.pdf.

United Nations. 2021a. Sustainable Transport, Sustainable Development. UN. https://sdgs.un.org/sites/default/files/2021-10/Transportation%20Report%202021_FullReport_Digital.pdf.

United Nations. 2021b. Indicator 11.7.1, SDG indicator metadata. UN. https://unstats.un.org/sdgs/metadata/files/Metadata-11-07-01.pdf.

United Nations. 2022. Indicator 7.1.1, SDG indicator metadata. UN. https://unstats.un.org/sdgs/metadata/files/Metadata-07-01-01.pdf.

United Nations Department of Economic and Social Affairs (UN DESA). n.d.a. Goal 7 | Department of Economic and Social Affairs. https://sdgs.un.org/goals/goal7. Accessed 2022.

United Nations Department of Economic and Social Affairs (UN DESA). n.d.b. Universal access to sustainable energy will remain elusive without addressing inequalities | United Nations. https://www.un.org/en/desa/universal-access-sustainable-energy-will-remain-elusive-without-addressing-inequalities. Accessed 2022.

United Nations Department of Economic and Social Affairs (UN DESA). n.d.c. Goal 9 | Department of Economic and Social Affairs. https://sdgs.un.org/goals/goal9. Accessed 2022.

United Nations Department of Economic and Social Affairs (UN DESA). n.d.d. Goal 11 | Department of Economic and Social Affairs. https://sdgs.un.org/goals/goal11. Accessed 2022.

United Nations. 2021. Indicator 11.2.1, SDG indicator metadata. UN. https://unstats.un.org/sdgs/metadata/files/Metadata-11-02-01.pdf.

United Nations Department of Economic and Social Affairs (UN DESA). 2022. The Sustainable Development Goals Report 2022. New York, USA: UN DESA. https://unstats.un.org/sdgs/report/2022/.

United Nations Department of Economic and Social Affairs (UN DESA), Population Division. 2019. "World Population Prospects 2019, Online Edition. Rev. 1." https://population.un.org/wpp/Download/Archive/Standard/.

UNSD M49, United Nations Statistics Division (UNSD), Standard Country or Area Codes for Statistical Use (M49). https://unstats.un.org/unsd/methodology/m49/. Accessed 2022. (Direct data set url:https://unstats.un.org/unsd/methodology/m49/overview/).

United Nations Statistical Commission. 2020. A recommendation on the method to delineate cities, urban and rural areas for international statistical comparisons. New York, USA: UN Statistical Commission. https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Recommendation-E.pdf.

Workman, R., and K. McPherson. 2020. Measuring Rural Access for SDG 9.1.1. *Transactions in GIS*, 25(2):721–734. https://doi.org/10.1111/tgis.12721.

World Bank. Rural Access Index (RAI) | Data Catalog. https://datacatalog.worldbank.org/search/dataset/0038250. Accessed 2022.

WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton;, Department of Geography and Geosciences, University of Louisville, Departement de Geographie, Universite de Namur) and Center for International Earth Science Information Network (CIESIN), Columbia University. 2020. Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076) [Datasets]. https://hub.worldpop.org/geodata/listing?id=69 https://dx.doi.org/10.5258/SOTON/WP00660.

## XIII. Documentation Copyright and License



## Appendix 1. Data Revision History

This is the first release of these data sets and therefore, no revisions have been made.

## Appendix 2. Contributing Authors & Documentation Revision History

| Revision Date | ORCID | Contributors | Revisions |
|---|---|---|---|
| July 20, 2023 | 0009-0003-2580-366X<br>0000-0001-9484-1705 | James Gibson,<br>Linda Pistolesi | This document is the 1st instance of documentation. |