SDG indicator metadata

(Harmonized metadata template - format version 1.1)

0. Indicator information (SDG_INDICATOR_INFO)

0.a. Goal (SDG_GOAL)

Goal 17. Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development

0.b. Target (SDG_TARGET)

Target 17.18: By 2020, enhance capacity-building support to developing countries, including for least developed countries and small island developing States, to increase significantly the availability of high-quality, timely and reliable data disaggregated by income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts

0.c. Indicator (SDG_INDICATOR)

Indicator 17.18.1: Statistical capacity indicators

0.d. Series (SDG_SERIES_DESCR)

IQ_SPI_PIL4 - Data Sources performance index (Statistical Performance Indicators Pillar 4) [17.18.1]

IQ_SPI_PIL5 - Data Infrastructure performance index (Statistical Performance Indicators Pillar 5) [17.18.1]

0.e. Metadata update (META_LAST_UPDATE)

2024-03-28

0.f. Related indicators (SDG_RELATED_INDICATORS)

17.18.2, 17,18.3, 17.19.1, 17.19.2

0.g. International organisations(s) responsible for global monitoring (SDG_CUSTODIAN_AGENCIES)

World Bank (WB)

1. Data reporter (CONTACT)

1.a. Organisation (CONTACT_ORGANISATION)

World Bank (WB)

2. Definition, concepts, and classifications (IND_DEF_CON_CLASS)

2.a. Definition and concepts (STAT_CONC_DEF)

The new Statistical Performance Indicators (SPI) replaces the Statistical Capacity Index (SCI), which the World Bank has regularly published since 2004. Although the goals are the same, to offer a better tool to measure the statistical systems of countries, the new SPI framework has expanded into new areas including data use, administrative data, geospatial data, data services, and data infrastructure. The SPI provides a framework that can help countries measure where they stand in several dimensions and offers an ambitious measurement agenda for the international community.

The data sources overall score (Pillar 4 score) is a composite measure of whether countries have data available from the following sources: Censuses and surveys, administrative data and geospatial data. The data sources (input) pillar is segmented by the following sources generated by (i) the statistical office (censuses and surveys), and sources accessed from elsewhere such as (ii) administrative data, and (iii) geospatial data.

The data sources pillar (Pillar 4) covers the following dimensions:
- **Dimension 4.1: Censuses and Surveys:** *Availability of recent censuses and surveys covering broad areas. The following censuses and surveys are considered: Population & Housing census, Agriculture census, Business/establishment census, Household Survey on income/ consumption/ expenditure/ budget/ Integrated Survey, Agriculture survey, Labor Force Survey, Health/Demographic survey, Business/establishment survey. Source: NSO websites, World Bank microdata library, ILO microdata library, IHSN microdata library.*
- **Dimension 4.2: Administrative Data:** *Availability of Civil Registration and Vital Statistics (CRVS) indicator. A complete CRVS is one type of administrative data system that is well documented based on the United Nations Demographic Yearbook Vital Statistics Questionnaire.*
- **Dimension 4.3: Geospatial Data:** *Geospatial data available at 1st Admin Level. This data source from Open Data Watch focuses on data availability at the sub-national level and provides a partial understanding of a country's ability to produce geospatial data.*

The data infrastructure overall score (Pillar 5 score) measures standards and methods addressing compliance with recognized frameworks and concepts.

Standards and Methods: *This set of indicators is based on countries' use of internationally accepted and recommended methodologies, classifications and standards regarding data integration. These indicators help facilitate data exchange and provide the foundation for the preparation of relevant statistical indicators. The following methods and standards are considered: System of national accounts in use, National Accounts base year, Classification of national industry, CPI base year, Classification of household consumption, Classification of status of employment, Central government accounting status, Compilation of government finance statistics, Compilation of monetary and financial statistics, Business process. Further work could improve the validity of this indicator and reduce the risk that countries may be incentivized to adopt only traditional standards and methods and neglect innovative solutions that may be more valid in the current context.*

2.b. Unit of measure (UNIT_MEASURE)

0-100 scale

2.c. Classifications (CLASS_SYSTEM)

The data sources and data infrastructure index are produced by aggregating several underlying indicators. There is no standard international classification, but the approach has been peer reviewed in Dang, Pullinger, Serajuddin, & Stacy (2023).

3. Data source type and data collection method (SRC_TYPE_COLL_METHOD)
3.a. Data sources (SOURCE_TYPE)

For pillar 4, on censuses and surveys, two complementary approaches are taken to collecting data. The first is to make use of data submitted to the World Bank, International Household Survey Network (IHSN), International Labour Organization (ILO), and Food and Agriculture Organization (FAO) microdata libraries. Only surveys that are marked as nationally representative are used, and surveys are classified (as either health surveys, agriculture surveys, labor force surveys, etc.) using the classifications submitted to the microdata libraries. The contents of searches on these databases are available in the github repository. The second approach is a manual data collection effort, where national statistical office (NSO) websites have been visited to ensure no surveys were missed. To be included in this search, the survey or census must be publicly available and accessible. If surveys or censuses are missed in this search, the easiest way for a country to get it included would be to create an entry for the survey at one of the microdata libraries. Information on the completeness of the Civil Registry and Vital Statistics (CRVS) is sourced from the UN SDG global monitoring database. Information on whether data is available at the 1st administrative level, for the geospatial indicator, is sourced from Open Data Watch.

For pillar 5, indicators in the standards and methods pillar are sourced primarily through the International Monetary Fund (IMF). Information on the system of national accounts in use and national accounts base year are sourced through the World Bank's World Development Indicaotrs (WDI) metadata. Data for the business process indicator is sourced through the United Nations Industrial Development Organization (UNIDO) and the United Nations Economic Commission for Europe (UNECE).

3.b. Data collection method (COLL_METHOD)

The World Bank transparently shares and makes public the methodologies for all adjustments to original data (e.g., through its Statistical Performance Indicators (SPI) website, https://www.worldbank.org/spi and its various analytical documents). All code and data to generate the SPI scores is available on the SPI public github repository (https://github.com/worldbank/SPI).

3.c. Data collection calendar (FREQ_COLL)

Data is collected at the beginning of November each calendar year.

3.d. Data release calendar (REL_CAL_POLICY)

Data is released annually near the beginning of April.

3.e. Data providers (DATA_SOURCE)

Data come from the World Bank, International Monetary Fund (IMF), Open Data Watch, PARIS21, International Labour Organization (ILO), World Health Organization (WHO), United Nations Educational, Scientific and Cultural Organization (UNESCO), International Household Survey Network (IHSN), United Nations (UN), and Food and Agriculture Organization (FAO). Additionally, some data is collected directly from the websites of national statistical offices.

3.f. Data compilers (COMPILING_ORG)

World Bank (WB)

3.g. Institutional mandate (INST_MANDATE)

Within the World Bank, the Development Data Group (DECGG) is in charge of the collection, validation, and calculation of SPI indicators.

4. Other methodological considerations (OTHER_METHOD)

4.a. Rationale (RATIONALE)

The data sources pillar (Pillar 4) is segmented by three types of sources generated by (i) the statistical office (censuses and surveys), and sources accessed from elsewhere such as (ii) administrative data and (iii) geospatial data. The appropriate balance between these source types will vary depending on a country's institutional setting and the maturity of its statistical system. High scores should reflect the extent to which the sources being utilized enable the necessary statistical indicators to be generated. For example, a low score on environment statistics (in the data production pillar) may reflect a lack of use of (and low score for) geospatial data (in the data sources pillar). This type of linkage is inherent in the data cycle approach and can help highlight areas for investment required if country needs are to be met.

The data infrastructure pillar (Pillar 5) includes standards and methods addressing compliance with recognized frameworks and concepts.

4.b. Comment and limitations (REC_USE_LIM)

In the data sources pillar (Pillar 4), more information is needed in the areas of administrative data and geospatial data. On administrative data, the picture could be complemented with indicators on whether countries have administrative data systems in place to measure health, education, labor, and social protection program statistics. For the geospatial indicator, the information could be complemented by new measures of how countries are using geospatial information in other ways, for instance using satellite data. And while the world is increasingly awash with private and citizen generated data (e.g., on mobility, job search, or social networking), on a global scale there is no reliable source to measure how national statistical systems are incorporating this information, but in the future measures could be adopted that do so.

Several 'soft' components of the data infrastructure index could be added, but currently lack adequate data. This includes the areas of skills and of partnerships between entities in the national statistical system. The SPI dashboard makes use of the PARIS21 led SDG indicator on whether the statistical legislations in countries met the standards of the UN Fundamental Principles of Statistics, but this was not incorporated into the overall SPI score, because of inadequate country coverage. This is also true of the PARIS21 led SDG indicator on whether the national statistical system is fully funded.

4.c. Method of computation (DATA_COMP)

An overall score is produced by combining the Statistical Performance Indicators to yield one single index. The statistical performance indicators have a nested structure, and the SPI Pillar 4 and Pillar 5 scores are formed by sequentially aggregating each level. To begin we produce a score for each dimension, which, unless otherwise stated, is an unweighted average of the indicators within that dimension. After computing a score for each dimension, a score for the pillar is computed as an unweighted average of the dimensions in that pillar.

For Pillar 4 on data sources, censuses and surveys are given separate weights, so that censuses, surveys, admin data, and geospatial data each receives a weight of 1/4. While censuses and surveys are in the same pillar in the framework, and therefore each would typically only receive a weight of 1/6 (for a total weight of 1/3) in this dimension, because of their importance in producing many indicators, they are given extra weight such that they each gets a weight of 1/4 (for a total weight of 1/2).

For SPI Pillar 5, only the standards and methods dimension is included in the final score, although additional indicators that had lower levels of country coverage are available in the SPI database.

The SPI Pillar 4 and Pillar 5 scores have a maximum score of 100 and a minimum of 0. A score of 100 would indicate that a country has every single element that we measure. A score of 0 indicates that none are in available.

The nested structure of the index and the summation methods used to build an overall score meet the axiomatic properties outlined in Cameron et al. 2021. These properties include symmetry, monotonicity, and subgroup decomposability. Symmetry refers to property where if the values of two indicators in a nesting are switched, then the resulting index scores are unaffected. Monotonicity implies that if the value of an indicator improves, then the resulting index scores improve as well. Subgroup decomposability results from the fact that the scores are a weighted average of the subgroups (either indicators, dimensions, pillars) that make up that score and so can be written as a linear combination of those subgroups.

Scores for each dimension in Pillar 4 are discussed below:

Dimension 4.1: censuses and surveys
This indicator draws from data collected by the Statistical Performance Indicators team. The following censuses and surveys are considered:
- Population & Housing census
- Agriculture census
- Business/establishment census
- Household Survey on income/ consumption/ expenditure/ budget/ Integrated Survey
- Agriculture survey
- Labor Force Survey
- Health/Demographic survey
- Business/establishment survey

Population & Housing census
Population censuses collect data on the size, distribution and composition of population and information on a broad range of social and economic characteristics of the population. It also provides sampling frames for household and other surveys. Housing censuses provide information on the supply of housing units, the structural characteristics and facilities, and health and the development of normal family living conditions. Data obtained as part of the population census, including data on homeless persons, are often used in the presentation and analysis of the results of the housing census. It is recommended that population and housing censuses be conducted at least every 10 years.
1 Point. Population census done within last 10 years
0.5 Points. Population census done within last 20 years
0 Points. Otherwise

Agriculture census

Agriculture censuses collect information on agricultural activities, such as size of holding, land tenure, land use, employment and production, and provide basic structural data and sampling frames for agricultural surveys. Censuses of agriculture normally involves collecting key structural data by complete enumeration of all agricultural holdings, in combination with more detailed structural data using sampling methods. It is recommended that agricultural censuses be conducted at least every 10 years.

1 Point. census done within last 10 years
0.5 Points. census done within last 20 years
0 Points. Otherwise

Business/establishment census

Business/establishment censuses provide valuable information on all economic activities, number of employed and size of establishments in the economy. Business Register information is establishment-based and includes business location, organization type (e.g. subsidiary or parent), industry classification, and operating data (e.g., receipts and employment).

1 Point. census done within last 10 years
0.5 Points. census done within last 20 years
0 Points. Otherwise

Household Survey on income/consumption/expenditure/budget/Integrated Survey

These surveys collect data on household income (including income in kind), consumption and expenditure. They typically include income, expenditure, and consumption surveys, household budget surveys, integrated surveys. It is recommended that surveys on income and expenditure be conducted at least every 3 to 5 years.

1 Point. 3 or more surveys done within past 10 years
0.67 Points. 2 surveys done within past 10 years;
0.33 Points. 1 survey done within past 10 years;
0 Points. None within past 10 years

Agriculture survey

Agricultural surveys refer to surveys of agricultural holdings based on the sampling frames established by the agricultural census. These are surveys on agricultural land, production, crops and livestock, aquaculture, labor and cost, and time use. Some issues, such as gender and food security, are of interest to most agriculture surveys.

1 Point. 3 or more surveys done within past 10 years
0.67 Points. 2 surveys done within past 10 years;
0.33 Points. 1 survey done within past 10 years;
0 Points. None within past 10 years

Labor Force Survey

Labor force survey is a standard household-based survey of work-related statistics at the national and sub-national employment or unemployment levels, rates or trends. The surveys also provide the characteristics of the employed or unemployed, including labor force status by age or gender, breakdowns between employees and the self-employed, public versus private sector employment, multiple job-holding, hiring, job creation, and duration of unemployment.

1 Point. 3 or more surveys done within past 10 years

0.67 Points. 2 surveys done within past 10 years;

0.33 Points. 1 survey done within past 10 years;

0 Points. None within past 10 years

Health/Demographic survey

Health surveys collect information on various aspects of health of populations, such as health expenditure, access, utilization, and outcomes. They typically include Demographic and Health Surveys. It is recommended that health surveys be conducted at least every 3 to 5 years.

1 Point. 3 or more surveys done within past 10 years

0.67 Points. 2 surveys done within past 10 years;

0.33 Points. 1 survey done within past 10 years;

0 Points. None within past 10 years

Business/establishment survey

The business/establishment survey provides information on employment, hours, and earnings of employees from a sample of business establishments including private and public, entities that are classified based on an establishment's principal activity from the business or establishment census. Establishment surveys include surveys of businesses, farms, and institutions. They may ask for information about the establishment itself and/or employee characteristics and demographics.

1 Point. 3 or more surveys done within past 10 years

0.67 Points. 2 surveys done within past 10 years;

0.33 Points. 1 survey done within past 10 years;

0 Points. None within past 10 years

Dimension 4.2: CRVS

Civil registration is the act of recording and documenting of vital events in a person's life (including birth, marriage, divorce, adoption, and death and cause of death) and is a fundamental function of national governments. Birth registration establishes an individual's legal identity at birth. A legal identity, name, nationality, and proof of age, are important human rights. They enable individuals to be included in various government, social and private services, and include the right to vote, etc. Vital statistics are compiled using civil registration information on these vital events. The availability of reliable and up-to-date vital statistics depends on the level of development of civil registration programs. An effective civil registration and vital statistics (CRVS) system is critical for planning and monitoring programs across several sectors.

Data comes from the UNSD Global SDG monitoring database. Scoring is as follows:

1 Points. Both of at least 90% of births registered and at least 75% of deaths registered

0.5 Points. One of at least 90% of births registered or at least 75% of deaths registered

0 Points. Neither

Dimension 4.3: geospatial data

Our source for this indicator is Open Data Watch. From Open Data watch:

> The Open Data Inventory (ODIN) assesses the coverage and openness of official statistics to help identify gaps, promote open data policies, improve access, and encourage dialogue between national statistical offices (NSOs) and data users. ODIN 2018/19 includes 178 countries, including most all OECD countries. Two-year comparisons are for all countries with two years of data between 2015-2017. Scores can be compared across topics and countries.

We use their indicator on whether indicators are available at the first administrative level. To identify the first administrative levels, ODIN largely draws on the ISO 3166-2 standard. In many countries, first administrative levels refer to governorates, regions, or province.

Scoring for the ODIN indicators for geospatial information is below:

- 1 point if all published data in a data category are available at first administrative level.
- 0.5 points if some published data in a data category are available at first administrative level.
- 0 points if no data are available at this level

There are 21 data categories.

Social Statistics
1. Population and Vital Statistics
2. Education Facilities
3. Education Outcomes
4. Health Facilities
5. Health Outcomes
6. Reproductive Health
7. Gender Statistics
8. Crime and Justice Statistics
9. Poverty Statistics

Economic Statistics
10. National Accounts
11. Labor Statistics
12. Price Indexes
13. Government Finance
14. Money and Banking
15. International Trade
16. Balance of Payments

Environmental Statistics
17. Land Use
18. Resource Use
19. Energy Use
20. Pollution
21. Built Environment

Scores for each dimension in Pillar 5 are discussed below:

Dimension 5.2: standards and methods

System of National Accounts in use
The national accounts data are compiled using the concepts, definitions, framework, and methodology of the System of National Account 2008 (SNA2008) or European System of National and Regional Accounts (ESA 2010). The manual has evolved to meet the changing economic structure, to follow systematic accounting and ensure international compatibility.
Scoring:
1 point for using SNA2008 or ESA 2010,
0.5 points for using SNA 1993 or ESA 1995,
0 points otherwise

National Accounts base year
National accounts base year is the year used as the base period for constant price calculations in the country's national accounts. It is recommended that the base year of constant price estimates be changed periodically to reflect changes in economic structure and relative prices.
1 point for chained price,
0.5 for reference period within past 10 years,
0 points otherwise.

Classification of national industry
The industrial production data are compiled using the International Standard Industrial Classification of All Economic Activities (ISIC) Rev.4 and Statistical Classification of Economic Activities in the European Community (NACE) Rev.2. ISIC Rev.4 is a standard classification of economic activities arranged so that entities can be classified per the activity they carry out using criteria such as input, output and use of the products produced, more emphasis has been given to the character of the production process in defining and delineating ISIC classes for international comparability. The manual and classification have changed to cover the complete scope of industrial production, employment, and GDP and other statistical areas.
1 Point. Latest version is adopted (ISIC Rev 4, NACE Rev 2 or a compatible classification)
0.5 Points. Previous version is used (ISIC Rev 3, NACE Rev 1 or a compatible classification)
0 Points. Otherwise

CPI base year
Consumer Price Index serves as indicators of inflation and reflects changes in the cost of acquiring a fixed basket of goods and services by the average consumer.
Weights are usually derived from consumer expenditure surveys and the CPI base year refers to the year the weights were derived. It is recommended that the base year be changed periodically to reflect changes in expenditure structure.
1 Point. Annual chain linking.
0.5 Points. Base year in last 10 years.
0 points. Otherwise

Classification of household consumption
Classification of Individual Consumption According to Purpose (COICOP) is used in household budget surveys, consumer price indices and international comparisons of gross domestic product (GDP) and its component expenditures.
Although COICOP is not strictly linked to any particular model of consumer behavior, the classification is designed to broadly reflect differences in income elasticities. It is an integral part of the SNA1993 and more detailed subdivision of the classes provide comparability between countries and between statistics in these different areas.
1 Point. Follow Classification of Individual Consumption by Purpose (COICOP)
0 Points. Otherwise

Classification of status of employment
Classification of status of employment refers to employment data that are compiled using the current international standard International Classification of Status in Employment (ISCE-93). It classifies jobs with respect to the type of explicit or implicit contract of employment between the job holder and the economic unit in which he or she is employed. Therefore, it aims to provide the basis for production of

internationally comparable statistics on the employment relationship, including the distinction between salaried employment and self-employment.

1 Point. Follow International Labour Organization, International Classification of Status in Employment (ICSE-93) or 2012 North American Industry Classification System (NAICS).

0 Points Otherwise.

Central government accounting status

Government finance accounting status refers to the accounting basis for reporting central government financial data. For many countries' government finance data, have been consolidated into one set of accounts capturing all the central government's fiscal activities and following noncash recording basis. Budgetary central government accounts do not necessarily include all central government units, the picture they provide of central government activities is usually incomplete.

1 Point. Consolidated central government accounting follows noncash recording basis

0.5 Points. Consolidated central government accounting follows cash recording basis

0 Points. Otherwise

Compilation of government finance statistics

(GFSM) in use for compiling the data. It provides guidelines on the institutional structure of governments and the presentation of fiscal data in a format similar to business accounting with a balance sheet and income statement plus guidelines on the treatment of exchange rate and other valuation adjustments. The latest manual GFSM2014 is harmonized with the SNA2008.

1 Point. Follow the latest Government Finance Statistical Manual (2014)/ ESA2010

0.5 Points. Previous version is used (GFSM 2001)

0 Points. Otherwise

Compilation of monetary and financial statistics

Compilation of monetary and financial statistics refers to the Monetary and Financial Statistics Manual (MFSM) in use. It covers concepts, definitions, classifications of financial instruments and sectors, and accounting rules, and provides a comprehensive analytic framework for monetary and financial planning and policy determination. The Monetary and Finance Statistics: Compilation Guide (2008) provides detailed guidelines for the compilation of monetary and financial statistics in addition to MFSM.

1 Point. Follow the latest Monetary and Finance Statistics Manual (2000) or Monetary and Finance Statistics: Compilation Guide (2008/2016)

0 Points. Otherwise

Business process

The Generic Statistical Business Process Model (GSBPM) aims to describe statistics production in a general and process-oriented way. It is used both within and between statistical offices as a common basis for work with statistics production in different ways, such as quality, efficiency, standardization, and process-orientation. It is used for all types of surveys, and "business" is not related to "business statistics" but refers to the statistical office, simply expressed.

1 Point. GSBPM is in use

0 Points. Otherwise

4.d. Validation (DATA_VALIDATION)

The technical quality of the dataset is measured in a number of different ways. First, a set of automated data checks are performed, which involve comparing new values added to the dataset to the values of previous years to highlight unexpected changes or outliers and applying a set of data validation rules to check if any values of any columns are outside of expected bounds.

For new data added, the following validation checks are performed to highlight potential issues. First, the data set will be compared to the original dataset published on github. It is natural for some values to change from this vintage compared to last vintage, as the underlying data sources have been updated. However, this comparison helps to systematically monitor the differences and potentially highlight unexpected changes. The dataCompareR package is used for this purpose in R (Johnston et al. 2021).

Second, the data set is compared to a set of expected values for each column. The exact expected values are listed in the appendix (Supplementary Information), but the general idea is to check that the values adhere to rational min/max values and positive correlate in expected ways. For instance, the availability of SDG indicators columns is the fraction of SDG indicators in a goal with a value in the previous 5 years. Thus, they should be between 0 and 1. Additionally, the countries and dates included are matched to a pre-specified expected list to uncover the accidental inclusion/exclusion of dates or countries. The validate R package is used for this purpose (van der Loo & de Jonge (2021).

Third, we examine the index using different weighting methods, which does not change the results significantly. For example, using a weight of 1/6 for censuses and surveys (instead of ¼ in pillar 4 on data sources) provides very similar results. In particular, the correlation between the SPI overall score under the preferred approach and the alternative approach is 0.998.

Finally, we provide qualitative data checks by regularly consulting with colleagues who are experts at the World Bank and other international organizations (e.g., FAO, IMF, ILO) on different development subjects. In particular, we obtain data help through our long-running collaboration with colleagues at other international agencies (e.g., by asking them to reach out to their contacts in the NSOs for data verification or clarification). Similarly, wherever data for some indicators (countries) are in doubt, we also discuss with our country-based colleagues who are charged with daily monitoring of country development. This internal validation process is based on established data quality procedures that we implement for all the other databases produced by the World Bank such as the World Development Indicators.

4.e. Adjustments (ADJUSTMENT)

Not applicable

4.f. Treatment of missing values (i) at country level and (ii) at regional level (IMPUTATION)

In general, missing data for a particular country-year observation is not imputed. There is one exception. One is that data from Open Data Inventory (ODIN) on data openness and geospatial data is available in 2015, 2016, 2017, 2018, and 2020, but not 2019. For 2019 data, the 2018 value is imputed by carrying forward the 2018 value for 2019.

4.g. Regional aggregations (REG_AGG)

Regional aggregations are calculated by taking an unweighted average of country level SPI scores for that region.

4.h. Methods and guidance available to countries for the compilation of the data at the national level (DOC_METHOD)

Not applicable

4.i. Quality management (QUALITY_MGMNT)

Before publishing the numbers the data will follow a three-level review process. First, an internal team review will take place, where a number of data quality checks are completed. These data checks are described in greater detail below and in Dang et al. (2023). This involves for each country a comparison of the updated value to the value for the latest year from the previous vintage of data. It also includes a set of automated data checks built on the R package: validate. This package automatically checks that database adheres to a set of expectations including the number of countries, the time span covered by the indicators, and that the indicators themselves fall into predefined ranges (e.g. indicators scored between 0 and 1 all have values between 0 and 1). Second, a DECDG internal review takes place where indicators values for countries are shared within the DECDG team for review. At this opportunity, DECDG staff will have the opportunity to look at country scores for each indicator and comment on whether these reflect ground truth. Finally, country scores will be sent out the regional directors for information. In this case, regional reports will be produced for each World Bank region, which include the regional aggregate score and the score for other regions, the country scores for countries within each region, and time trends for that region.

4.j Quality assurance (QUALITY_ASSURE)

To create a layer of transparency and accountability, the SPI team maintains an SPI github repository, so users can track changes to the raw information collected in detail through the version control tools of Github. The github repository contains the version control history of the code and data, which documents every change in the data and code of the entire project dating back to July 2020 to build confidence and transparency. Any user can openly view how an indicator was constructed, and any change to the code or data that took place back to July 2020, before the project launch.

4.k Quality assessment (QUALITY_ASSMNT)

Prior to release, the SPI team produces a "what's new" publication detailing information about the data production process. The SPI "what's new" publication will detail the dates that data was collected for the indicators and show a table showing changes between the current release and the previous release for all countries.

5. Data availability and disaggregation (COVERAGE)

Data is available for 186 countries and areas as of the 2022 data release. These economies cover roughly 99.3% of the world population.

6. Comparability / deviation from international standards (COMPARABILITY)

Not applicable

7. References and Documentation (OTHER_DOC)

URL:
https://www.worldbank.org/spi

Github:
https://github.com/worldbank/SPI

SPI Working Paper:
https://documents.worldbank.org/en/publication/documents-reports/documentdetail/440191616164007723

SPI Technical Paper:
https://documents.worldbank.org/en/publication/documents-reports/documentdetail/815721616086786412

Dang, H., Pullinger, J., Serajuddin, U., Stacy, B. 2023. "Statistical performance indicators and index—a new tool to measure country statistical capacity". *Scientific Data (10), 146.* https://doi.org/10.1038/s41597-023-01971-0

Grant J. Cameron & Hai-Anh H. Dang & Mustafa Dinc & James Foster & Michael M. Lokshin, 2021. "**Measuring the Statistical Capacity of Nations**," Oxford Bulletin of Economics and Statistics, Department of Economics, University of Oxford, vol. 83(4), pages 870-896, August.

Johnston, S., Noble-Eddy, R., van Horssen, M., Bhasin, K., and Pollicott, S. *dataCompareR: Compare Two Data Frames and Summarise the Difference*. R package version 0.1.4. https://CRAN.R-project.org/package=dataCompareR (2021).

Loo, M. P. J. van der & Jonge, E. de. Data validation infrastructure for R. *Journal of Statistical Software* **97**, 1–31 (2021).