



THE  
**OPEN DATA INVENTORY**  
**2016** METHODS  
REPORT

---

# Acknowledgements

---

The Open Data Inventory is a team effort. We are pleased to acknowledge the help of all who contributed to this work.

## Open Data Watch

Shaïda Badiee, Misha Belkindas, Eric Swanson, Deirdre Appel, Jamison Crowell, Reza Farivari, Martin Getzendanner, Chandrika Kaul, Amelia Pittman, Tawheeda Wahabzada.

## ODIN Project Team

Eric Swanson, Director  
Jamison Crowell, Project Manager

## ODIN Assessors

Sarah Belford, Hannah Krentler-Brown, Dominic Scerbo, Deirdre Appel, Janeen Rashmawi, Pikria Saliashvili, Maryan Carbuccia, Caitlin Vejby, Sydney Taylor, Elizabeth Rogers, and Molly Geiser.

## ODIN Reviewers

Jamison Crowell, Chandrika Kaul, Tawheeda Wahabzada, Amelia Pittman, and Pikria Saliashvili.

## Website and Publication Design

District Design Group

## Website Development

Akron, Inc.

## Funding

Provided by the William and Flora Hewlett Foundation

## Backcover Photograph

Courtesy of Jonathan Torgovnik /Getty Images Reportage.  
Some rights reserved.

Copyright 2017 Open Data Watch – Licensed under a Creative Commons Attribution 4.0 license  
Please cite any uses of these data as Open Data Watch – Open Data Inventory

[www.odin.opendatawatch.com](http://www.odin.opendatawatch.com)



# THE OPEN DATA INVENTORY 2016 METHODS REPORT

---



OPEN DATA WATCH, INC

**OPENDATAWATCH.COM**

ACKNOWLEDGEMENTS

TABLE OF CONTENTS2

INTRODUCTION3

ODIN 2016 FAQs & Methodology Updates8

DATA CATEGORIES AND INDICATORS7

Data Sources7

Data Categories7

DATA COVERAGE14

Coverage Element 1: Coverage and Disaggregation14

Coverage Element 2: Data Availability for Last 5 Years20

Coverage Element 3: Data Availability for Last 10 Years20

Coverage Element 4: First Administrative Level21

Coverage Element 5: Second Administrative Level21

DATA OPENNESS22

Openness Element 1: Machine Readability22

Openness Element 2: Non-proprietary Format22

Openness Element 3: Download Options23

Openness Element 4: Metadata Availability23

Openness Element 5: Free/Unrestricted Terms of Use24

AGGREGATE SCORES25

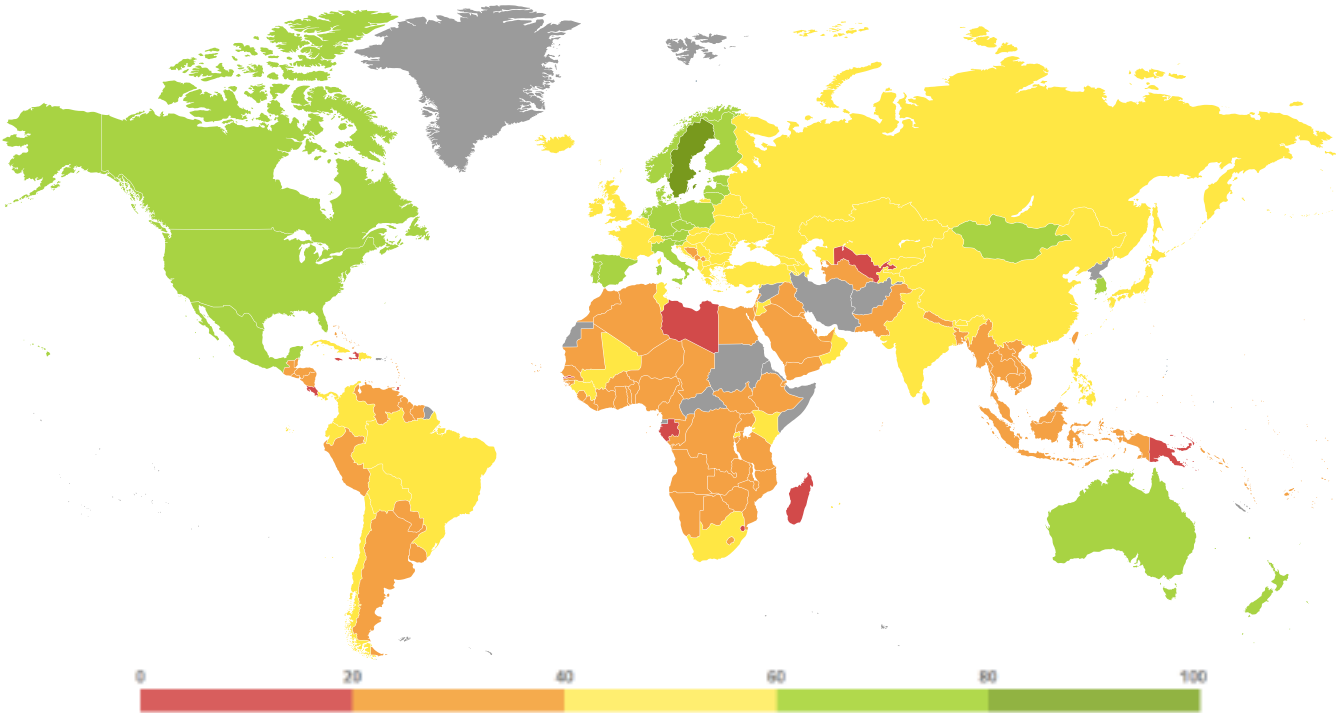
Standardized Scores25

Weighting25

ODIN 2016 is an assessment of the coverage and openness of data provided on the websites maintained by national statistical offices (NSOs) in 173 countries. Building on the previous year’s work, ODIN 2016 now includes most of the world’s high-income and OECD countries. The assessments analyze datasets across 20 data categories that are the most pertinent to managing and monitoring progress on the SDGs, as well as the social, economic, and environmental development of a country more generally.

The ODIN 2016 assessments were carried out between June and September 2016. Assessment results were reviewed and, if necessary, revised between October and December 2016. All results are based on the data available from the principal NSO website during the assessment period.

Figure 1. ODIN global scores, 2016



173 Complete Country Assessments

ODIN 2016 FAQs & Methodology Updates

What is ODIN?

The Open Data Inventory (ODIN) is an evaluation of the coverage and openness of data provided on the websites maintained by national statistical offices (NSOs). The overall ODIN score is an indicator of how complete and open an NSO's data offerings are. The summary scores for social, economic, and environmental statistics and summary scores for coverage and openness produce a picture of the national statistical systems' strengths and weaknesses. To see a video on ODIN 2016, visit [www.opendatawatch.com/odin2016video](http://www.opendatawatch.com/odin2016video).

What is ODIN's Purpose?

ODIN helps identify critical gaps, promote open data policies, improve data access, and encourage dialogue between NSOs and data users. NSOs and their development partners can use ODIN as part of a strategic planning process and as a measuring rod for the development of the statistical system. ODIN provides valuable information to data users within the government and private sectors and to the public about the availability of important statistical series. In addition to the ratings of coverage and openness in twenty statistical categories, ODIN assessments record the online location of key indicators in each data category, permitting quick access to hundreds of indicators.

How are Open Data Defined?

There is general agreement on the core meaning of open data. As summarized in the Open Definition, version 2.1, "Knowledge is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness." This definition has been operationalized in the International Open Data Charter. In practical terms, open data should be machine readable in non-proprietary formats, selectable by users, accompanied by descriptive

metadata, and free to be used and reused for any purpose without limitations other than acknowledgement of the original source. These requirements have been incorporated in the five elements of the ODIN openness assessment.

What Data Categories are Included?

ODIN assessments review published statistics in twenty categories, grouped as social statistics, economic and financial statistics, and environmental statistics. The ODIN overall score weights the three groups equally. In each category, principal or sentinel indicators were selected because they are frequently needed for public policies or private initiatives and because they provide evidence of underlying statistical processes for which statistical offices are responsible.

The twenty categories of ODIN are:

- 1. Population and Vital Statistics
- 2. Education Facilities
- 3. Education Outcomes
- 4. Health Facilities
- 5. Health Outcomes
- 6. Reproductive Health
- 7. Gender Statistics
- 8. Poverty Statistics
- 9. National Accounts
- 10. Labor Statistics
- 11. Price Indexes
- 12. Government Finance
- 13. Money and Banking
- 14. International Trade
- 15. Balance of Payments
- 16. Land Use
- 17. Resource Use
- 18. Energy Use
- 19. Pollution
- 20. Built Environment

Why Assess National Statistical Offices?

ODIN assessments begin with the websites maintained by national statistical offices because, in most countries, the NSO is the lead agency of the national statistical system, coordinating its work with other governmental bodies that produce official statistics. If an official national data source can be reached from the NSO's website, it is included in the ODIN assessment. NSOs, as producers and caretakers of official statistics, have a special obligation to maximize their public benefit. NSOs can and should be the leading advocate for and provider of high quality, official statistics to government, the public, and to the international community.

Methodology Changes Since ODIN 2015

Assessment process

In 2016 the ODIN research process recorded assessments using an online platform developed for exclusive use of ODIN. This platform allows assessors immediate access to methodology guidelines, examples of data they may encounter and FAQs. The platform has built-in functionality designed to catch common mistakes and automate some of the scoring process to ensure comparability of scores across countries. The new platform also provides greater ease of use for assessors and reviewers alike, making possible the inclusion of additional countries without extending the research period. Because of the benefits of the new platform, combined with a more extensive training program for country assessors, only one assessor was assigned per country. Following the initial assessment, each country assessment underwent two layers of review, as was the case for ODIN 2015.

Greater coverage

ODIN 2015 covered 125 countries in mostly low- and middle-income countries. In ODIN 2016, this scope was expanded to include many high-income and OECD countries; an addition of 51 countries.

Between 2015 and 2016, there are comparable data for 122 countries. (Three countries – Afghanistan, Iran, and Sudan – were removed from the 2016 assessment due to non-functioning websites).

Non-proprietary formats

In ODIN 2016 several file formats are scored as non-proprietary which were treated as proprietary in ODIN 2015. These changes have been made to align with international standards widely recognized by the open data community. The following formats now receive full scores for non-proprietary format: XLSX, DOCX, PPTX, and PDF. These changes had the effect of raising the openness scores for many countries.

Frequency criteria

Many countries currently participate in the IMF's Enhanced General Data Dissemination System (e-GDDS) or Special Data Dissemination Standard (SDDS). These standards provide countries guidance on, among other things, the coverage, periodicity and timeliness of data. During this edition of ODIN, we incorporated the IMF's e-GDDS periodicity requirements for three economic data categories: National Accounts, Price Indexes, and International Trade. For countries to score full points on the first coverage element for these categories, data published in the last 5 years must be available on at least a quarterly basis. In future versions of ODIN, we will extend this requirement to data published in the last 10 years and likely to additional data categories.

Indicator guidelines

To greater align our representative indicators with international guidelines and ensure comparability with the additional countries added in ODIN 2016, updates were made to the indicators and disaggregation of those indicators within certain data categories. The following is a summary of

those substantive changes.

- 1. Gender Statistics: The indicator, “specialized studies on the status and condition of women” was removed in favor of less subjective indicators.
- 2. Poverty Statistics: An exception was added, allowing for data on low-income or at-risk poverty rates to be sufficient if less than 1 percent of the country’s population fell below the international poverty line.
- 3. National Accounts: Indicators and disaggregation were updated to better align with recommendations from international organizations, not limited to, the UN Statistics Division.
- 4. Central Government Finance: Disaggregation of indicators were updated to better align with the

IMF’s Government Finance Statistics Manual.

- 5. Money and Banking: Indicators and disaggregation were updated to better align with guidelines set by international organizations, not limited to, the World Bank.
- 6. Balance of Payments: Indicators and disaggregation were updated to better align with the IMF’s Balance of Payments Manual.
- 7. Land Use: Indicators and disaggregation were updated to better align with the UN’s Guidelines of International Classifications for Agricultural Statistics.
- 8. Energy Use: Indicators and disaggregation were updated to better align with guidelines set by international organizations, not limited to, the International Energy Agency.

Data Categories and Indicators

The following sections explain the assessment methodology and the assumptions underlying the 2016 ODIN assessments.

Data Sources

The Open Data Inventory assesses the coverage and openness of statistics available from websites maintained by national statistical offices. Websites maintained by private or non-governmental agencies or international agencies are not included in the assessment. Websites maintained by other units of the national government or by sub-national governmental units are included if and only if they can be reached from the national statistical office website.

For example, if the national accounts are maintained by the central bank, then data would be included in the ODIN assessment only if the NSO’s website provides a link to the appropriate page on the central bank’s website or if the NSO reproduces the data on its own website. ODIN is premised on the belief that NSOs can and should take responsibility for providing access to all official statistics.

Data Categories

The Open Data Inventory assesses macrodata. By this we mean data from surveys or administrative sources that have been aggregated above the unit record level. We focus on these data because they are the final product released by the NSO or other official agencies They are used most frequently for policy making and for tracking policy outcomes. Microdata from censuses and surveys are very important, but require a different approach to assessing their openness.

Twenty categories of data are included in the ODIN assessment. [Tables 1-3](#) list the data categories, the representative indicators, required disaggregation, and explanatory notes for each indicator. For the construction of summary measures, the data categories are grouped as social statistics, economic statistics, and environmental statistics.

Table 1. ODIN social data categories, 2016

Social Statistics			
Data category	Representative Indicators	Disaggregation	Additional Notes
Population & Vital Statistics	(1.1) Population by 5 year age groups (see additional notes for exceptions);  (1.2) Crude birth rate;  (1.3) Crude death rate.	(1.1) Sex; marital status; (1.2) Sex; marital status;  (1.3) Sex.	(1.1) Estimations and projections are included.  (1.1) Points for first and second administrative levels can be given for population data not disaggregated by 5-year age groups as long as disaggregated data is also available at the national level.  (1.2, 1.3) Birth and death registration figures are not accepted, but number of births/deaths are accepted.  (1.2, 1.3) Births and deaths presented as sex ratios are not accepted.

Data category	Representative Indicators	Disaggregation	Additional Notes
Education Facilities	(2.1) Number of schools and/or classrooms;  (2.2) Number of teaching staff;  (2.3) Annual education budget.	(2.1) Student age groups; school stages; school types.  (2.2) Student age groups; school stages; school types;  (2.3) School stages; school/ calendar/fiscal year; functional categories (such as teachers, supplies, administration, etc.).	(2.1-2.3) School stages at a minimum refers to primary, secondary, and/or post-secondary/ tertiary. Some countries may use different terms.  (2.1, 2.2) School types may refer to technical schools, public, private, etc.  (2.3) Education budget must reference corresponding year. GDP contribution to education is not accepted.. Budgetary figures for the Ministry of Education are accepted.
Education Outcomes	(3.1) Enrollment rates;  (3.2) Completion and/or graduation rates;  (3.3) Competency exam results.	(3.1-3.3) Sex; school stages; age groups; school types.	(3.1-3.3) Attainment, attendance, and literacy rates are not accepted.  (3.1) Number of students is accepted.  (3.3) Matriculation exam results are accepted as disaggregated by school stage.
Health Facilities	(4.1) Number of health facilities;  (4.2) Hospital capacity (number of beds and/or budget);  (4.3) Number of health care staff.	(4.1) Facility/Department types;  (4.2) Facility/Department types; fiscal/calendar year (for budget data only)  (4.3) Facility/Department types; staff types.	(4.1-4.3) Facility types generally include clinics, hospitals, pharmacies, dental clinics, public/ private hospitals, etc. Department types generally include surgery, ophthalmology, internal medicine, pediatric care, dentistry, etc. Some countries may have different designations and this is analyzed on a case by case basis.
Health Outcomes	(5.1) Immunization rates (must indicate specific disease types).  (5.2) Number/rate of major communicable diseases prevalence (must indicate specific disease types).	(5.1) Age groups; sex; optional.  (5.2) Age groups; sex; optional.	(5.1) Immunization rates for children/infants is accepted as age disaggregation if the age group is defined.  (5.2) HIV incidence/prevalence is accepted even if other disease types are not listed.  (5.2) Mortality rates disaggregated by disease type is not accepted.

Data category	Representative Indicators	Disaggregation	Additional Notes
Reproductive Health	(6.1) Maternal mortality ratio;  (6.2) Infant mortality rate;  (6.3) Under-5 mortality rate;  (6.4) Fertility rate;  (6.5) Contraceptive prevalence rate;  (6.6) Adolescent birth rate.	(6.1) Optional.  (6.2, 6.3) Sex;  (6.4-6.6) Optional.	(6.1-6.6) Whole numbers cannot replace rates/ratios (as specified).  (6.3) Mortality rates disaggregated by age/age groups are accepted so long as rates for under 5 years of age are included.  (6.6) Birth rates disaggregated by age/age groups are accepted so long as rates for 15-19 years of age are included.
Gender Statistics	(7.1) Data on violence against women;  (7.2) Proportion of women in parliament and/or management/ senior positions.	(7.1, 7.2) Optional.	(7.1) Perception or opinion surveys on domestic violence or violence against women are not accepted. Data on victims of crimes disaggregated by sex are accepted. It should not be assumed that data on domestic violence or sexual assault include data only on women. Data may be presented as a rate or whole number.  (7.2) Share of women in employment in the non-agricultural sector is not accepted.  (7.2) Data must be presented as a rate. Whole numbers are only accepted if the data is disaggregated by sex.
Poverty & Income Statistics	(8.1) Number/ percentage of poor and/or poverty rate/incidence;  (8.2) Distribution of income by, at minimum, deciles.	(8.1, 8.2) Optional.	(8.1) This indicator can include data based on international or national poverty lines. If a country has less than 1% of the population that falls below the international poverty line, figures on at-risk poverty, low income rates, or similar are accepted.  (8.1) Data that are not accepted include: Gini coefficients, proportion of population with low disposable income.



Table 2. ODIN economic data categories, 2016

Economic Statistics			
Data category	Representative Indicators	Disaggregation	Additional Notes
National Accounts	(9.1) GDP (production approach) and/or gross value added;  (9.2) GDP (expenditure approach).	(9.1) Industrial classification; current prices; constant prices;  (9.2) Major expenditure categories; current prices; constant prices.	(9.1) Industrial classes include agriculture, mining, manufacturing, etc.  (9.2) Major expenditure categories include final consumption expenditure, government expenditure on goods and services, gross fixed capital formation, etc.
Labor Statistics	(10.1) Employment rate;  (10.2) Unemployment rate.	(10.1) Sex; age groups; industry; occupation type;  (10.2) Sex; age groups.	(10.1, 10.2) Active and idle figures are not accepted as employment or unemployment, nor are distribution figures by occupation type that do not show an aggregate rate for all occupations.  (10.1, 10.2) Youth unemployment alone is not accepted.  (10.1, 10.2) Data must be presented as a proportion of the entire population. Data disaggregated by age groups must include all the population age 15 or older.
Price Indexes	(11.1) Consumer price index (CPI);  (11.2) Producers price index (PPI).	(11.1, 11.2) Optional.	(11.1) CPI is only accepted if calculated for the entire country.  (11.2) This indicator may also be called Output Price Index. PPI disaggregated by industry or commodity qualifies only if an aggregate for all components is also presented. Industrial PPI is not accepted as it does not include construction and services.
Government Finance	(12.1) Actual revenues;  (12.2) Actual expenditures.	(12.1) Revenue source;  (12.2) Administrative, economic, or functional classes.	(12.1) Revenue sources may include : taxes, social contributions, grant, other, tax/non-tax.  (12.2) Administrative classes may include specific ministries, departments, etc.  Economic classes are typically presented as: expense, compensation of employees, use of goods and

Data category	Representative Indicators	Disaggregation	Additional Notes
(continued from previous page)			services, consumption of fixed capital, interest, subsidies, grant, social benefits, etc.  Functional classes are typically presented as: general public services, defense, public order and safety, economic affairs, environmental protection, health, education, etc.
Money & Banking	(13.1) Money supply;  (13.2) Domestic credit (provided by the entire financial sector).	(13.1) M0, M1, M2, M3, and/or M4.  (13.2) To private sector and/or public, partially state-owned enterprises.	(13.1) M1 and M2 (most common) are measures of money supply. M1 includes coins, currency, demand deposits, checking accounts, etc. M2 (also called quasi money) includes cash, checking deposits and near money.  (13.2) Credit data must be given by the entire financial sector, not just the central bank. Data on domestic claims are not accepted.
International trade	(14.1) Merchandise exports;  (14.2) Merchandise imports.	(14.1, 14.2) Major product categories (agricultural products, fuels/mining, manufactures, etc.).	(14.1, 14.2) Product categories may be called commodity groups.  Information available on a country's Balance of Payments is not accepted for any indicator in this category.
Balance of payments	(15.1) Current account;  (15.2) Capital and Financial account.	(15.1) Goods and services; income; current transfers;  (15.2) Direct investment and/or international investment position.	(15.1, 15.2) Disaggregation of current and capital and financial accounts are based on the IMF's Balance of Payments <a href="#">guide</a> . All disaggregations for a particular account must be present to accept data.

Table 3. ODIN environmental data categories, 2016

Environmental Statistics			
Data category	Representative Indicators	Disaggregation	Additional Notes
Land Use	(16.1) Land use and/or land cover.	16.1) Urban/rural; agricultural use (crop types); environmental zones.	<p>(16.1) Environmental zones refer to classifications as defined in the UN's <a href="#">Guidelines of International Classifications for Agricultural Statistics</a>.</p> <p>Environmental zones for land use may include: agriculture, forestry, aquaculture, etc.</p> <p>Environmental zones for land cover may include: crops, trees, mangroves, grassland, etc.</p> <p>Data should can presented in physical units (hectares, sq. km, etc.) or percentage of total land.</p>
Resource Use	(17.1) Data on fishery harvests;  (17.2) Data on deforestation and/or Forest coverage;  (17.3) Data on major mining activities;  (17.4) Water supply and/or consumption.  .	(17.1, 17.2) Optional.  (17.3) Type of mining activity;  (17.4) Optional.	<p>(17.1-17.4) Data should be presented in physical units and/or value.</p> <p>(17.1, 17.3) Contribution to GDP for given resources are not accepted.</p> <p>(17.1) Data may also be on fishery landings and catches.</p> <p>(17.2) Land use/cover data with a disaggregation category of forestry is not accepted, nor is data only on areas affected by forest fires.</p> <p>(17.3) Types of mining activities may include gas, petroleum, coal, etc.</p>
Energy Use	(18.1) Consumption of energy (must specify type).	(18.1) End-use sector; industrial sector.	<p>(18.1) Data should be in physical units, such as kilowatt hours of electricity or tons of oil equivalent for fossil fuels.</p> <p>End-use sector may include categories: construction, agriculture, transport, industry, residential, etc.</p>

Data category	Representative Indicators	Disaggregation	Additional Notes
Pollution	(19.1) CO2/Greenhouse gases (GHG) emissions (must specify type);  (19.2) Emissions of air and/or water pollutants (must specify types).	(19.1,19.2) Optional.	<p>(19.1, 19.2) Data should be in physical units.</p> <p>Production and consumption based accounting of emissions are both accepted.</p> <p>(19.1) Data for ozone depleting substances are accepted for this indicator.</p> <p>(19.2) Common air pollutants include: PM10, PM2.5, O3, NO2, CO, and SO2.</p>
Built Environment	(20.1) Proportion of people with access to water;  (20.2) Proportion of people with access to sanitation;  (20.3) Housing quality indicators.	(20.1) Optional  (20.2) Optional.  (20.3) Housing types; number of rooms; construction material, and other indicators on a case by case basis.	<p>(20.1) Data on access to clean/ drinking water or improved water sources is accepted.</p> <p>(20.2) Data on access to improved sanitation is accepted.</p> <p>(20.3) Housing type is defined broadly and is analyzed on a case by case basis.</p>



Data Coverage

The data categories are assessed against ten elements of coverage and openness shown in [Tables 1-15](#). Each element has a possible score of 1, 0.5, or 0, indicating that the data in a category satisfy the criteria for that element, partially satisfy them, or fail to satisfy them or the data are entirely missing. Thus a country has a maximum potential score of 200: 100 for data coverage and 100 for data openness.

The scoring scheme is deliberately coarse. A finer scoring grid (say from 1 to 10) would inevitably invite greater subjectivity on the part of assessors and create problems when comparing results produced by different assessors or at different times. The scoring guidelines for each element are summarized in the tables below.

Coverage Element 1: Coverage and Disaggregation

The first element requires assessors to locate representative indicators within each data category and determine whether important topical disaggregations are available. Guidelines for scoring each data category are shown in [Tables 4-6](#). The representative indicators and disaggregations are also listed in [Tables 4-6](#) for reference.

If the score for element 1 is less than 1, the remaining four elements of data coverage cannot exceed the score of element 1. However, the

elements of data openness (elements six through 10) are scored on the basis of available data, which may receive a full score for openness if they satisfy the guidelines for those elements. If no data are available for a category, all elements are scored 0.

Additionally, though ODIN assesses information available at the first or second administrative level, data must be first and foremost be available at the national level. If national level data for an indicator is not available, no points are awarded.

Table 4. Scoring guidelines for coverage element 1, social statistics

Social Statistics			
Data category	Representative Indicators	Disaggregation	Additional Notes
Population & Vital Statistics	(1.1) Population by 5 year age groups (see additional notes for exceptions);  (1.2) Crude birth rate;  (1.3) Crude death rate.	(1.1) Sex; marital status;  (1.2) Sex; marital status;  (1.3) Sex.	To receive full points: Must have all indicators disaggregated as follows: (1.1) population data by 5 year age groups by sex (0.5 points); and (1.2) crude birth and (1.3) death rates by sex (0.5 points).  To receive half points: Must have at least population by 5 year age groups with one disaggregation, or crude birth and death rates with one disaggregation each.

Data category	Representative Indicators	Disaggregation	Additional Notes
Education Facilities	(2.1) Number of schools and/or classrooms;  (2.2) Number of teaching staff;  (2.3) Annual education budget.	(2.1) Student age groups; school stages; school types.  (2.2) Student age groups; school stages; school types;  (2.3) School stages; school/ calendar/fiscal year; functional categories (such as teachers, supplies, administration, etc.).	To receive full points: Must have all three indicators disaggregated as follows: (2.1) number of schools and/or classrooms and (2.2) teaching staff by multiple school stages (at least 2) and one other disaggregation such as age groups and/or school types as well as (2.3) annual budget by year and one other disaggregation.  To receive half points: Must have at least one indicator with two disaggregations or two indicators with one disaggregation each.
Education Outcomes	(3.1) Enrollment rates;  (3.2) Completion and/or graduation rates;  (3.3) Competency exam results.	(3.1-3.3) Sex; school stages; age groups; school types.	To receive full points: Must have all indicators disaggregated as follows: (3.1) enrollment and (3.2) completion and/or graduation rates by sex and either school stage, school type, or age groups (0.5 points); as well as (3.3) exam results by sex and either school stage, school type, or age groups (0.5 points). If matriculation exam data is used as 3.3, disaggregation by school stage is implied.  To receive half points: Must have at least one indicator disaggregated by both sex and either school stage, school type, or age groups. Two indicators with only one disaggregation each is not enough.
Health Facilities	(4.1) Number of health facilities;  (4.2) Hospital capacity (number of beds and/or budget);  (4.3) Number of health care staff.	(4.1) Facility/Department types;  (4.2) Facility/Department types; fiscal/calendar year (for budget data only)  (4.3) Facility/Department types; staff types.	To receive full points: Must have two indicators with one disaggregation each.  To receive half points: Must have at least one indicator with one disaggregation.
Health Outcomes	(5.1) Immunization rates (must indicate specific disease types).  (5.2) Number/rate of major communicable diseases prevalence (must indicate specific disease types).	(5.1) Age groups; sex; optional.  (5.2) Age groups; sex; optional.	To receive full points: Must have all indicators disaggregated as follows: (5.1) immunization rates by sex and/or age groups (0.5 points); and (5.2) disease prevalence by sex and/or age (0.5 points).  To receive half points: Must have both indicators without disaggregation or one indicator with one disaggregation.

Data category	Representative Indicators	Disaggregation	Scoring Guidelines for First Coverage Element
Reproductive Health	(6.1) Maternal mortality ratio;  (6.2) Infant mortality rate;  (6.3) Under-5 mortality rate;  (6.4) Fertility rate;  (6.5) Contraceptive prevalence rate;  (6.6) Adolescent birth rate.	(6.1) Optional.  (6.2, 6.3) Sex;  (6.4-6.6) Optional.	To receive full points: Must have at least five indicators with one disaggregation each (unless disaggregation is optional).  To receive half points: Must have at least two disaggregated indicators (unless disaggregation is optional). One must be a mortality rate..
Gender Statistics	(7.1) Data on violence against women;  (7.2) Proportion of women in parliament and/or management/ senior positions;	(7.1, 7.2) Optional.	To receive full points: Must have (7.1) data on violence against women (0.5 points); and (7.2) data on women in management or political office (0.5 points). To receive half points: Must have at least one indicator.
Poverty & Income Statistics	(8.1) Number/ percentage of poor and/or poverty rate/incidence;  (8.2) Distribution of income by, at minimum, deciles.	(8.1, 8.2) Optional.	To receive full points: Must have (8.1) number/ percentage of poor and/or poverty rate/incidence (0.5 points); as well as (8.2) income distribution by, at least, deciles (0.5 points).  To receive half points: Must have at least one indicator.

Table 5. Scoring guidelines for coverage element 1, economic statistics

Economic Statistics			
Data category	Representative Indicators	Disaggregation	Scoring Guidelines for First Coverage Element
National Accounts	(9.1) GDP (production approach) and/or gross value added;  (9.2) GDP (expenditure approach).	(9.1) Industrial classification; current prices; constant prices;  (9.2) Major expenditure categories; current prices; constant prices.	To receive full points: Must have all indicators disaggregated as follows: (9.1) GDP (production approach) and/ or gross value added by industrial classification (0.5 points); and (9.2) GDP (expenditure approach) by major expenditure categories (0.5 points). Any data in the most recent 5 years (2011 onward) must be presented on at least quarterly basis to receive a full points.

Data category	Representative Indicators	Disaggregation	Scoring Guidelines for First Coverage Element
(continued from previous page)			To receive half points: Must have at least one indicator with one disaggregation. Data can be presented on a monthly, quarterly, or annual basis.
Labor Statistics	(10.1) Employment rate;  (10.2) Unemployment rate.	(10.1) Sex; age groups; industry; occupation type;  (10.2) Sex; age groups.	To receive full points: Must have all indicators disaggregated as follows: (10.1) employment by sex and age groups (0.5 points); and (10.2) unemployment by sex and age groups (0.5 points).  To receive half points: Must have at least one indicator with one disaggregation.
Price Indexes	(11.1) Consumer price index (CPI);  (11.2) Producers price index (PPI).	(11.1, 11.2) Optional.	To receive full points: Must have all indicators: (11.1) CPI (0.5 points); as well as (11.2) PPI (0.5 points). Any data in the most recent 5 years (2011 onward) must be presented on at least quarterly basis to receive a full points.  To receive half points: Must have at least one indicator. Data can be presented on a monthly, quarterly, or annual basis.
Government Finance	(12.1) Actual revenues;  (12.2) Actual expenditures.	(12.1) Revenue source;  (12.2) Administrative, economic, or functional classes.	(To receive full points: Must have all indicators disaggregated as follows: (12.1) actual revenues by revenue source (0.5 points); and (12.2) actual expenditures by at least one classification (0.5 points). If only tax revenues are disaggregated or the only disaggregation shows tax/ non-tax or oil/non-oil revenues, a score higher than .5 cannot be given.  To receive half points: Must have at least one disaggregated indicator.

Data category	Representative Indicators	Disaggregation	Scoring Guidelines for First Coverage Element
Money & Banking	(13.1) Money supply;  (13.2) Domestic credit (provided by the entire financial sector).	(13.1) M0, M1, M2, M3, and/or M4.  (13.2) To private sector and/or public, partially state-owned enterprises.	To receive full points: Must have all indicators disaggregated as follows: (13.1) money supply with one disaggregation (0.5 points); and (13.2) Domestic credit with one disaggregation.  To receive half points: Must have at least one disaggregated indicator.
International Trade	(14.1) Merchandise exports;  (14.2) Merchandise imports.	(14.1, 14.2) Major product categories (agricultural products, fuels/mining, manufactures, etc.).	To receive full points: Must have all indicators disaggregated as follows: (14.1) exports and (14.2) imports by major product categories. Any data in the most recent 5 years (2011 onward) must be presented on at least quarterly basis to receive a full points.  To receive half points: Must have one disaggregated indicator. Data can be presented on a monthly, quarterly, or annual basis.
Balance of Payments	(15.1) Current account;  (15.2) Capital and Financial account.	(15.1) Goods and services; income; current transfers;  (15.2) Direct investment and/or international investment position.	To receive full points: Must have (15.1) current account with all disaggregations and (15.2) with one disaggregation.  To receive half points: Must have all disaggregations under current or capital and financial account for 0.5 points.

Table 6. Scoring guidelines for coverage element 1, environmental statistics

Environmental Statistics			
Data category	Representative Indicators	Disaggregation	Scoring Guidelines for First Coverage Element
Land Use	(16.1) Land use and/or land cover.	16.1) Urban/rural; agricultural use (crop types); environmental zones.	To receive full points: Must have (16.1) land use and/or land cover with two disaggregations. Land use data with one disaggregation and land cover data with one disaggregation is accepted.  To receive half points: Must have one indicator with at least one disaggregation.

Data category	Representative Indicators	Disaggregation	Scoring Guidelines for First Coverage Element
Resource Use	(17.1) Data on fishery harvests;  (17.2) Data on deforestation and/or Forest coverage;  (17.3) Data on major mining activities;  (17.4) Water supply and/or consumption.	(17.1, 17.2) Optional.  (17.3) Type of mining activity;  (17.4) Optional.	To receive full points: Must have all four indicators, disaggregated if applicable.  To receive half points: Must have at least two indicators, disaggregated if applicable.
Energy Use	(18.1) Consumption of energy (must specify type).	(18.1) End-use sector; industrial sector.	To receive full points: Must have (18.1) consumption data on three energy types with one disaggregation each.  To receive half points: Must have consumption data on two energy types with one disaggregation each.
Pollution	(19.1) CO2/Greenhouse gases (GHG) emissions (must specify type);  (19.2) Emissions of air and/or water pollutants (must specify types).	(19.1,19.2) Optional.	To receive full points: Must have all indicators: (19.1) CO2 emissions and/or other GHG emissions (0.5 points); and (19.2) other air/water pollutant emissions (0.5 points). If data on CO2 emissions is not included, a score higher than .5 cannot be given.  To receive half points: Must have at least one indicator.
Built Environment	(20.1) Proportion of people with access to water;  (20.2) Proportion of people with access to sanitation;  (20.3) Housing quality indicators.	(20.1) Optional  (20.2) Optional.  (20.3) Housing types; number of rooms; construction material, and other indicators on a case by case basis.	To receive full points: Must have (20.1) proportion of people with access to water and (20.2) proportion of people with access to sanitation (0.5 points); as well as (20.3) one disaggregation of housing quality data (0.5 points).  To receive half points: Must have either (20.1) access to drinking water and (20.2) sanitation data, or (20.3) housing quality data with at least one disaggregation.

Scoring guidelines for the data coverage elements 2 through 5 are summarized in [Tables 7-10](#). Elements 2 and 3 assess the availability of annual data within each category over the 10-year period, 2006 – 2015. Elements 4 and 5 score the availability of subnational data at the level of first and second administrative units. Assessors are instructed to determine the administrative levels from official sources. Certain categories of economic statistics are not expected to be available for first or second administrative levels; no scores are recorded for those categories.

Coverage Element 2: Data Availability for Last 5 Years

Table 7. Scoring guidelines for coverage element 2

Scores for this element cannot be greater than the score for data coverage and disaggregation (element 1).	
For the categories National Accounts, Price Indexes, and International Trade, if data are presented on a quarterly basis, 3 out of 4 quarters for a particular year must be available to award credit. If data are presented on a monthly basis, 7 out of 12 months for a particular year must be present to award credit. Be aware if countries are using calendar years or other fiscal years prior to determining whether a full year of data is present. Other common fiscal years include 1 July – 30 June, 1 April-31 March. Less common fiscal years are also used in Afghanistan, Nepal, Iran, and the US.	
Scoring Options	Additional Notes
a. 1 point if data are available for 3 of the last 5 years.	The three years must occur within 2011-2015 (or 2010/2011-2014/2015). If the indicator shows data for a non-calendar year (i.e. 2012/2013), this should be counted as one year.
b. 0.5 points if data are available for 1-2 of the last 5 years.	The 1-2 years must occur within 2011-2015 (or 2010/2011-2014/2015). If the indicator shows data for a non-calendar year (i.e. 2012/2013), this should be counted as one year.
c. 0 points if data are unavailable for last 5 years.	A 0 is given for the categories National Accounts, Price Indexes, and International Trade if data are available within the last 5 years on a quarterly or monthly basis, but not for a majority of quarters/months in any particular year (see note above).

Coverage Element 3: Data Availability for Last 10 Years

Table 8. Scoring guidelines for coverage element 3

Scores for this element cannot be greater than the score for data coverage and disaggregation (element 1).	
For any data from 2011 onward in the categories National Accounts, Price Indexes, and International Trade, if data are presented on a quarterly basis, 3 out of 4 quarters for a particular year must be available to award credit. If data are presented on a monthly basis, 7 out of 12 months for a particular year must be present to award credit. Be aware if countries are using calendar years or other fiscal years prior to determining whether a full year of data is present. Other common fiscal years include 1 July – 30 June, 1 April-31 March. Less common fiscal years are also used in Afghanistan, Nepal, Iran, and the US.	
Scoring Options	Additional Notes
a. 1 point if data are available for 6 of the last 10 years.	The six years must occur within 2006-2015 (or 2005/2006-2014/2015). If the indicator shows data for a non-calendar year (i.e. 2012/2013), this should be counted as one year.
b. 0.5 points if data are available for 3-5 of the last 10 years.	The 3-5 years must occur within 2006-2015 (or 2005/2006-2014/2015). If the indicator shows data for a non-calendar year (i.e. 2012/2013), this should be counted as one year. If an indicator received 1 point for “data coverage for the last 5 years”, it should receive at least a 0.5 points in “data coverage for the last 10 years.”
c. 0 points if data are unavailable for 2 or fewer of last 10 years.	If data only exists prior to 2006, 0 points should be given.

Coverage Element 4: First Administrative Level

Table 9. Scoring guidelines for coverage element 4

Scores for this element cannot be greater than the score for data coverage and disaggregation (element 1).	
Additionally, data disaggregated at the first administrative level is only scored if national level data also exists.	
To identify the first administrative levels, reference the CIA World Factbook. The World Factbook largely draw on the ISO 3166-2 standard to identify first administrative levels ( <a href="http://www.iso.org/iso/home/standards/country_codes.htm">http://www.iso.org/iso/home/standards/country_codes.htm</a> ). In many countries, first administrative levels refer to governates, regions, or provinces. However, this is not always the case. For instance, in China, four municipalities are considered to be at the first administrative level (Beijing, Chongqing, Shanghai, and Tianjin).	
We accept population data at the first administrative level even if it is not disaggregated by 5-year age groups.	
Money & Banking, International Trade, and Balance of Payments are not scored for this element.	
Scoring Options	Additional Notes
a. 1 point if all data are available at first administrative level (state, province, and so forth).	
b. 0.5 points if data available at first administrative level for some indicators, or for a subset of one indicator.	
c. 0 points if no data only available at this level.	

Coverage Element 5: Second Administrative Level

Table 10. Scoring guidelines for coverage element 5

Scores for this element cannot be greater than the score for data coverage and disaggregation (element 1).	
Additionally, data disaggregated at the second administrative level is only scored if national level data also exists.	
If geographical disaggregation was included for locations not listed as first administrative level, assume that the data are disaggregated to the second administrative level as long as the classification is more numerous than the first.	
National Accounts, Government Finance, Money & Banking, International Trade, and Balance of Payments are not scored for this element. We accept population data at the second administrative level even if it is not disaggregated by 5-year age groups.	
Scoring Options	Additional Notes
a. 1 point if data available at second administrative level (municipality or other similar division)	
b. 0.5 points if data available at second administrative level for some indicators, or for a subset of one indicator.	
c. 0 points if no data only available at this level.	



Data Openness

Elements 6 through 10 assess the openness of data in a category using criteria derived from the Open Definition. (See <http://opendefinition.org/>.) Scores for coverage and openness were considered independently. If only one indicator for a certain category was published but that indicator was published in a fully open manner, it was given full points for openness. Scores for openness could, therefore, exceed the scores for coverage in the same category, but in practice this rarely happens. The scoring guidelines for the elements of openness are shown in Tables 11-15.

Elements 6 and 7 assess whether data are downloadable in machine readable, non-proprietary formats. Open data should be available to anyone in convenient and readily modifiable form. Element 8 asks whether users can select the data they are interested in and whether they are able to establish an API connection to the data, which would allow data to be linked to other applications. The alternative is often that data are only available in predetermined tables. The availability of metadata (element 9) is of importance in providing users with information on how the data were collected and compiled. Clear licensing terms/ Terms of Use (element 10) state what users may do with the data and permit for reuse of data with some restrictions; fully open data may be used and reused without restriction other than providing attribution to the original source.

Openness Element 1: Machine Readability

Table 11. Scoring guidelines for openness element 1

For the elements of data openness, scoring is calculated independent of the data coverage.	
Scoring Options	Additional Notes
a. 1 point if all data are downloadable in a machine-readable format (such as XLS, XLSX, CSV, Stata, SAS, SPSS, JSON and so forth).	CDF, RDF, XML, and TXT files are also classified as machine readable. PDFs, picture files, and HTML formats are not machine readable.
b. 0.5 points if some but not all the published data are downloadable in machine-readable format.	If published data are available in both a machine readable format and in a non-machine readable format, a full points should be awarded for machine readability. If, however, the data in the non-machine readable format is less complete, a half point should be awarded.
c. 0 points if published data are not available in machine-readable format.	

Openness Element 2: Non-proprietary Format

Table 12. Scoring guidelines for openness element 2

For the elements of data openness, scoring is calculated independent of the data coverage.	
If data files are compressed in RAR format (which is proprietary), data for that indicator should be considered proprietary even if the enclosing files are in a non-proprietary format.	
Scoring Options	Additional Notes
a. 1 point if published data are downloadable in non-proprietary format (such as XLSX, DOCX, CSV, XML, HTML, and JSON).	TXT files are also treated as non-proprietary. XLS, Stata, SAS, SPSS, DOC and PPT files are proprietary.

Scoring Options	Additional Notes
b. 0.5 points if some but not all published data are downloadable in a non-proprietary format.	If some published data are available in non-proprietary format but the same data are available in a proprietary format, full points should be awarded for non-proprietary format. If, however, the data in the non-machine readable format is less complete, a half points should be awarded.
c. 0 points if published data are not downloadable in a non-proprietary format.	If data files are compressed in RAR format, 0 points should be given. ZIP compression is non-proprietary and does not affect a country's score.

Openness Element 3: Download Options

Table 13. Scoring guidelines for openness element 3

For the elements of data openness, scoring is calculated independent of the data coverage.	
If all indicators in a category have an option for user selected downloads AND are available through bulk download/API, a full points should be given for this element. Zero points should be given for this element if no variables have bulk downloads or user selection. Any other combination would be given a half points.	
Scoring Options	Additional Notes
a. 0.5 points if user can select specific indicators in software for download.	For data to qualify as having an “user selection” option, users must be able to select an indicator and at least one other dimension to create a download or table. These dimensions could include time periods, geographic disaggregations, or other recommended disaggregations.
b. An additional 0.5 points is given if an Application Program Interface (API) is available that allows for bulk download or if a bulk download is available. APIs should be clearly visible on the website.	IBulk downloads are defined somewhat liberally as a feature that allows a user to easily download an entire dataset without being required to select customizations for that dataset. Bulk downloads should include multiple indicators, not just a time series for one indicator or multiple disaggregations of one indicator.

Openness Element 4: Metadata Availability

Table 14. Scoring guidelines for openness element 4

For the elements of data openness, scoring is calculated independent of the data coverage.	
Metadata are defined as information about how the data were collected, including the name of the survey/census/report, dates when data were compiled, and definitions of indicators included. In scoring this element, we distinguish metadata that provide specific information about an indicator from general information that applies to number of different indicators.	
Scoring Options	Additional Notes
a. 1 point if metadata are present that provide specific details about the definition of the indicator or the method of data collection and compilation for that indicator.	Metadata may be included in the same file as the data or it can be included elsewhere on the website.
b. 0.5 points if metadata are provided about a large survey or group of data of which the indicator is part.	If only some indicators have specific metadata or all/ some indicators have general metadata, a half point should be given.
c. 0 points if no metadata are available.	

Openness Element 5: Free/Unrestricted Terms of Use

Table 15. Scoring guidelines for openness element 5

For the elements of data openness, scoring is calculated independent of the data coverage.	
In order for Terms of Use policies to be scored, they need to clearly refer to the data found on the website. Terms of Use that refer to non-data content (such as pictures or news stories) of the website are not considered. A copyright symbol at the bottom of the page is not sufficient.	
If directed to an external website or portal, separate Terms of Use should exist unless that content is explicitly referenced in the Terms of Use on the NSO's website.	
Scoring Options	Additional Notes
a. 1 point if Terms of Use are the Creative Commons Attribution 4.0 (CC BY 4.0) or specifically reference CC BY 4.0 without adding restrictions.	Creative Commons Attribution 4.0 (CC BY 4.0) or similar means that data must be licensed to permit for free use and reuse for commercial and noncommercial use with, at most, an obligation to attribute data to original source. If the policy is unclear if restrictions exist, it cannot score higher than 0.5 points.
b. 0.5 points if Terms of Use allow for some reuse of data with some restrictions.	A typical restriction may limit reuse to non-commercial products, require the inclusion of original metadata with the attribution, or specify that material may not be reused in a misleading manner.
c. If no terms of use are found or if terms of use do not allow for use or reuse.	Restrictive terms of use include clauses that require users to seek permission, in any format, register or provide personal information about themselves, require payment, or other restrictions that exceed what is allowed as a "typical restriction." If no terms of use are provided other than a copyright symbol, 0 points are also given.

Aggregate Scores

ODIN scores are summarized along both dimensions of the ODIN assessment: by categories and by elements. In addition, subscores are computed for the combined categories of social statistics, economic statistics, and environmental statistics and for the combined elements of coverage and openness. The overall score aggregates all scores across both dimensions. For convenience, all aggregate scores are standardized by rescaling them to a range of 0 to 100.

Standardized Scores

The aggregate scores shown in ODIN tables and charts have been standardized. Scores are standardized by dividing by the maximum score achievable and multiplying by 100. For most subscores, the maximum score is the product of the number of data categories and the number of elements included. However, some of the elements of geographic disaggregation have been excluded a priori from the economic categories.

Specifically, it is assumed that the national accounts and government finance statistics will not be available at the second administrative level and that money and banking, international trade, and balance of payments statistics will not be available at the first or second administrative levels. Therefore, the maximum, unweighted score for five data coverage elements across all seven economic categories is 27 not 35 and the maximum achievable score over all data categories and elements is 192 not 200. Standardized scores involving any of these categories are reweighted to give them full weight.

Because of this discrepancy, subscores over data categories or across elements involving economic statistics will not “add up” consistently, but the treatment of each subscore is internally consistent.

Weighting

Because the three principal topical groupings (social, economic, and environmental) contain different numbers of data categories, aggregates computed over these categories would be implicitly weighted by the number of categories in each grouping. To neutralize this effect, the data categories are reweighted so that each group has equal weight in aggregates computed over all categories. The reweighting does not affect aggregates computed within each grouping. All elements have equal weights in all aggregates.

ODIN Online has an option for downloading both the raw and weighted scores for further analysis. In the future an option for user-specified weights for both categories and elements will be included in the online version of ODIN.



