



OPEN DATA INVENTORY 2018/19 ANNUAL REPORT

REPORTING ADVANCEMENTS IN OPEN DATA

1. Executive Summary

The 2018/19 Open Data Inventory (ODIN) is the fourth edition of the index compiled by Open Data Watch and provides an assessment of the coverage and openness of official statistics in 178 countries. The purpose of ODIN is to provide an objective and reproducible measure of the public availability of national statistics, and their adherence to open data standards. ODIN evaluates openness on a spectrum, measuring the degree of openness rather than classifying datasets as open or closed.

Results from 2018 indicate that national statistical systems are becoming more open. In 2016 and 2017 the median overall score remained unchanged at 38, but in 2018 the median score went up by 3 points. Some countries have made large improvements, resulting in new additions to the top ten highest scoring countries. In 2018, Singapore claimed the number one spot, bumping Denmark to second place by just over one percentage point. Germany entered the top ten ranked ODIN countries in 7th place - a grouping which has fluctuated over the years but remains relatively homogenous.

Although there remains a strong relationship between income level and overall ODIN scores, every year some of the largest improvements are found in developing regions. In 2018 the top 20 most improved countries include seven that are low- or lower-middle income and twelve that are upper-middle or high income. Furthermore, the geographic diversity has also increased. Central America makes the most frequent showing, followed by Eastern Asia and the Caribbean. Over half of all global regions have at least one country among the twenty most improved countries.

Most of the countries that made the greatest progress over the last year did so by improving the openness of existing data. Between 2017 and 2018, the openness elements with the highest average improvement were terms of use (or data licenses), with an average increase of 20 points, and metadata availability, with an average increase of 10 points. Many of the countries that made the largest positive changes amended or adopted an open terms of use or developed a new data portal that connects datasets to newly developed or existing metadata.

Though progress has been made, some challenges are prevalent:

1. Data availability remains a challenge in many data categories. While the availability of data, or data coverage, saw slight improvements, much of it can be explained by methodology changes in ODIN. The fact remains, data availability remains a challenge in many data categories with persistent data gaps. Data on crime and justice, pollution, and energy use remain severely underreported. However, the countries that do publish data in these categories, do so relatively comprehensively.

2. Coverage gaps differ by type of data. For instance, economic and financial data categories score significantly higher on 5- and 10-year data availability, while social indicators score better by having subnational data at the first and second administrative level.

Overall, the results from ODIN 2018/19 are promising. If countries take their open data commitments seriously, the open data goals of the [Cape Town Action Plan for Sustainable Development Data](#), which calls on countries to fully incorporate open data into national statistical systems, can be achieved. But there are no open data without data. Coverage gaps must be filled by utilizing new sources of information, collecting new data, or releasing previously non-public datasets, or their open data efforts will stall. Addressing these gaps requires a system-wide approach while filling other gaps may require specific, sectoral strategies.

2. Introducing the Open Data Inventory (ODIN)

2.1 How to Use This Report (not for PDF version)

The ODIN 2018/19 annual report gives an in-depth look at the results from the most recent assessments, as well as comparisons with previous years.

Navigate: To navigate through the report, use the menu on the top left of this page. Hover over each section title to reveal subsections underneath. Click any section or subsection title to move to the beginning of that section. To navigate to another section, choose a different title from the menu. Most figures and charts are interactive and have different functionality depending on their content. A description of how to interact with each figure or table appears below each one.

Read offline: To download a printable copy of this report as a PDF, scroll to the top of the page. Before the Executive Summary there is a link on the right side that provides a non-interactive version of the report for offline reading and printing.

Download data: To download the data used in this report, click “Data” in the menu to be redirected to the ODIN website download page.

2.2 ODIN’s Four Year Evolution

The Open Data Inventory, the flagship project of [Open Data Watch \(ODW\)](#), began its first assessment in the summer of 2015. Over time, the number of countries included in ODIN has grown, the methodology has improved, and the assessment process itself has been refined. Open data is an emerging field whose principles and best practices have evolved over the years. To better reflect this evolution and the changing landscape of countries’ priorities to align data with the [Sustainable Development Goals \(SDGs\)](#), ODIN has evolved as well. Below is a brief summary of ODIN’s evolution.

In 2015, ODIN included 125 low- and middle-income countries and assessed open data across 20 data categories. In 2016, ODIN added high-income and Organization for Economic Co-operation and Development (OECD) countries for a total of 173 countries. Quarterly publication requirements were introduced for many economic indicators. And to improve the quality and consistency of the assessments, an internal system was created to collect, validate, and review assessments. In 2017 and 2018, ODW made minor updates to the definitions for indicator, categorical, and geographic disaggregation. (See Updates in 2018 for more information.) The most significant change in last two years was the addition of a new category for crime and justice statistics.

Because of these changes, data from 2015 are not included in the time-series comparisons of this report. ODIN 2015 remains a record of the state of open data in 2015, but comparisons with the following years may be misleading.

ODIN is designed as a tool that will grow over time to respond to the open data priorities of the international community. In future years, additional data categories and new datasets are likely to be added as we work with countries to better understand the barriers to and opportunities for implementing open data.

2.3 Updates in 2018

There are a few methodological changes in the 2018/19 Open Data Inventory (ODIN). The main changes are described below. For further details, please read the [ODIN 2018/19 Methodology Report](#).

Country Coverage: The total number of countries assessed in 2018 was 178, down two from last year. Chad and Kiribati were removed from this year’s assessment because their websites experienced significant technical issues during the assessment period that made it impossible to complete their assessments.

Geographic Disaggregation: ODIN data coverage criteria requires geographic disaggregation of data, but some

data categories are excluded. For example, international trade data is excluded from disaggregation at the first or second administrative level. This year, additional categories were added to the exclusion list, including pollution, energy use, price indexes, and resource, which are now excluded from second administrative level disaggregation.

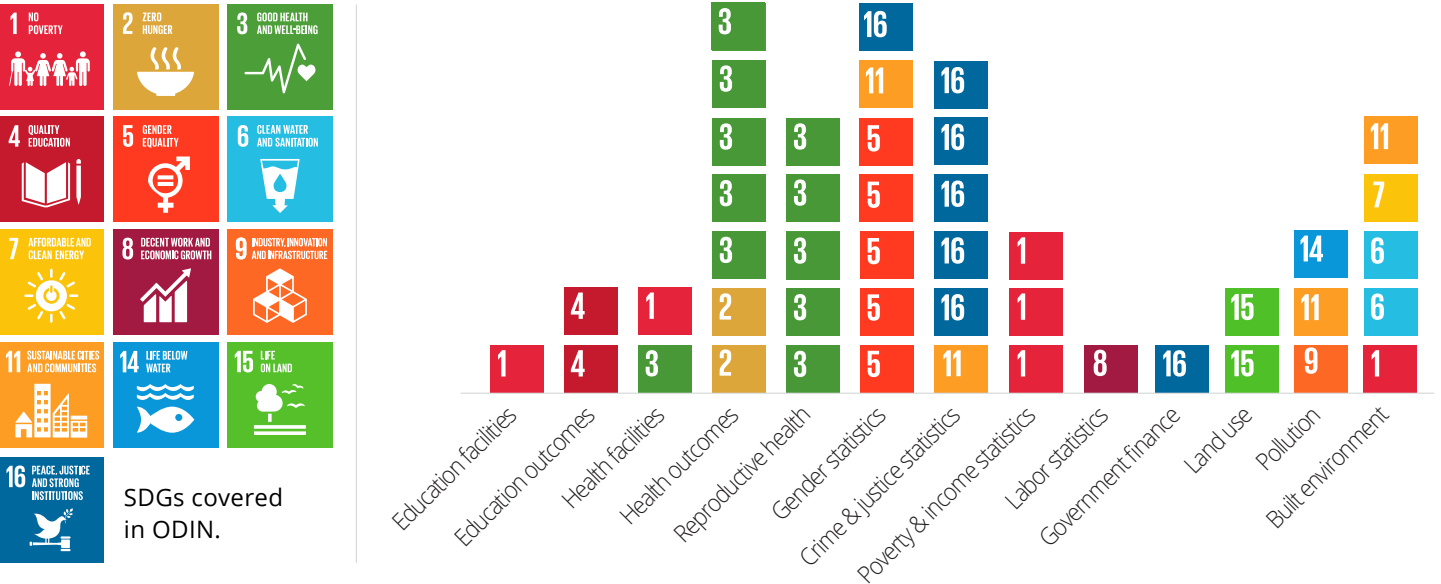
After numerous consultations, it became clear that even countries with excellent reporting practices often did not produce these datasets at the second administrative level. Therefore, these criteria were not appropriate as representative measures of data coverage. Furthermore, other individual indicators were also excluded from second administrative level disaggregation, including education and health budget data, maternal mortality ratio, infant mortality rate, proportion of women in government, persons in prison, distribution of income, and data on land use.

In line with these changes, we also accepted several alternative classifications for first administrative level disaggregation credit, as appropriate for the data type. For example, most environmental data can receive first administrative disaggregation credit for disaggregation by urban and rural areas. By default, first administrative disaggregations for each country are based on the [International Organization for Standards \(ISO\) 3166](#).

Increased Country Engagement: In 2018, Open Data Watch invited every country evaluated in ODIN to participate in the assessment process. During this process, countries’ national statistical offices were invited to suggest additional datasets for the ODIN team to consider to be counted in their assessment and to provide comments on their coverage gaps. In 2017, 65 countries agreed to participate and in 2018, 99 countries agreed to participate. The number of countries that provided comments increased as well, from 48 to 69 countries. This increased engagement with countries has led to more countries seeking direct consultation with Open Data Watch as they implement open data initiatives and an increased use of ODIN by countries as an internal tool to measure progress.

Sustainable Development Goals (SDG): Many SDG indicators are now accepted in place of certain ODIN indicators if the corresponding ODIN indicator is not published. For example, ODIN indicator 4.2 Number of beds or data on healthcare staff, can be substituted by SDG indicator 3.c.1 Health worker density and distribution. ODIN indicators are meant to be representative of the primary information published by countries and are crucial to the framework of many of the SDG indicators. It is rare for a country to publish data on SDG indicators, while not providing data on the more primary underlying ODIN indicators, but this methodology change allows countries to receive credit for either. Figure 1 shows the SDG indicators that are accepted in ODIN.

Figure 1. SDG indicators and corresponding ODIN categories



Please see [online version](#) to access interactive elements of all visualizations.

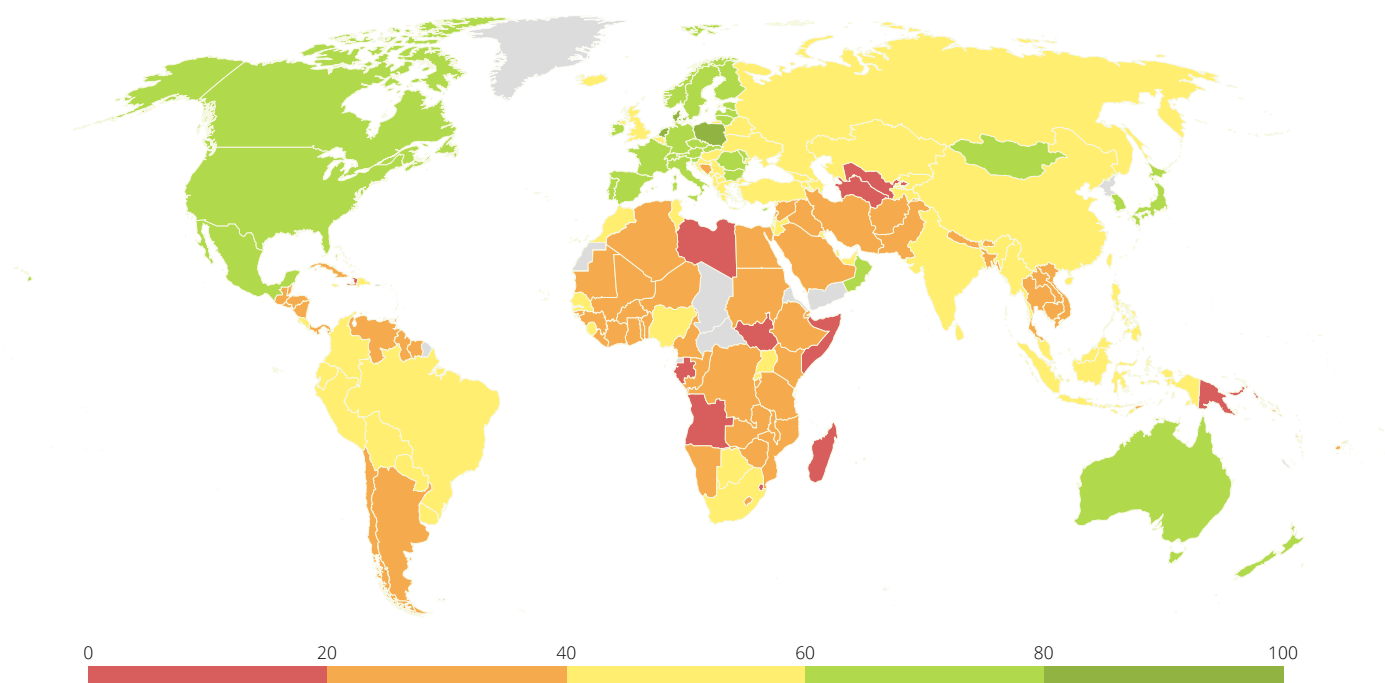
3. Findings

3.1 What Happened in 2018?

The Open Data Inventory (ODIN) 2018/19 assessed the coverage and openness of official statistical data in 178 countries. Median scores increased in coverage and openness as did average scores in nearly all categories. Many countries showed considerable improvements, resulting in new additions to the ODIN's highest scoring countries. Progress has been made as governments become more knowledgeable of open data [principles](#). At the same time, not many governments are publishing new datasets and, instead, have focused on opening existing datasets. As a result, openness scores have driven aggregate scores upward over time, while coverage scores have remained mostly static.

Figure 2 shows the scores of countries included in ODIN 2018/19 grouped by quintiles. Countries not included are shown in gray. Chad and Kiribati were removed from the index this year because their websites were non-functioning during the assessment period.

Figure 2. ODIN overall scores, 2018



ODIN scores are calculated as a percentage of the maximum score obtainable. The median ODIN country score for 2018 is 41.1, so fewer than half the countries satisfy more than 41 percent of the ODIN criteria for data coverage and openness across all data categories. The global median scores for coverage and openness rose by 2 and 3 points to 41.5 and 41.3 respectively. Overall, most countries are making steady progress in both openness and data coverage, but openness improvements are more common and have had greater impacts on overall scores.

2018 saw new open data champions emerge, as represented in ODIN's top ten highest scoring countries in Table 1. This year, Singapore and Germany have reached the top ten, in first and seventh place. Compared to last year, the gap between 1st and 10th place is smaller this year with only a 9-point difference in comparison to the 12-point difference in 2017. Within the top ten, only a few points separate many of the countries' rankings, including the highly competitive top ranked position where Singapore has overtaken two-year champion, Denmark, by 1.3 points.

Table 1. ODIN 2018/19 highest scoring countries

| Rank | Country Name | ODIN 2018 Score |
|------|--------------|-----------------|
| 1 | Singapore | 85 |
| 2 | Denmark | 84 |
| 3 | Netherlands | 82 |
| 4 | Poland | 81 |
| 5 | Slovenia | 78 |
| 6 | Finland | 78 |
| 7 | Germany | 77 |
| 7 | Sweden | 77 |
| 9 | Canada | 77 |
| 10 | Norway | 76 |

Singapore's Ascension

Between 2017 and 2018, [Singapore's](#) overall score increased by 21 points since 2017, a result of increased availability and openness of data. With a nearly perfect openness score of 99 out of 100, Singapore's national statistical office's ([SingStat](#)) website serves as a great example for other countries. In the last year, Singapore launched a completely redesigned website with many features that have increased accessibility and openness, including a multi-table download function (bulk download), an application programming interface (API), and an easy-to-use navigation to find datasets. In addition, Singapore has updated their [terms of use](#) for data to conform to open data standards, resulting in a fully open [data use policy](#).

The highest scoring countries in ODIN 2018/19 have many similarities, most notably their income level; each of them are a high-income country. However, moving farther down the ranks, but still among the 40 highest scoring countries, there is more variety. Seventy-five percent of the top 40 countries are classified as high income, 17.5 percent as upper-middle income, 2.9 percent as lower-middle income, and 2.5 percent as low income. Table 2 shows the income distribution of all countries in 2018 by score deciles.

Table 2. Countries (%) by income group and score decile, 2018

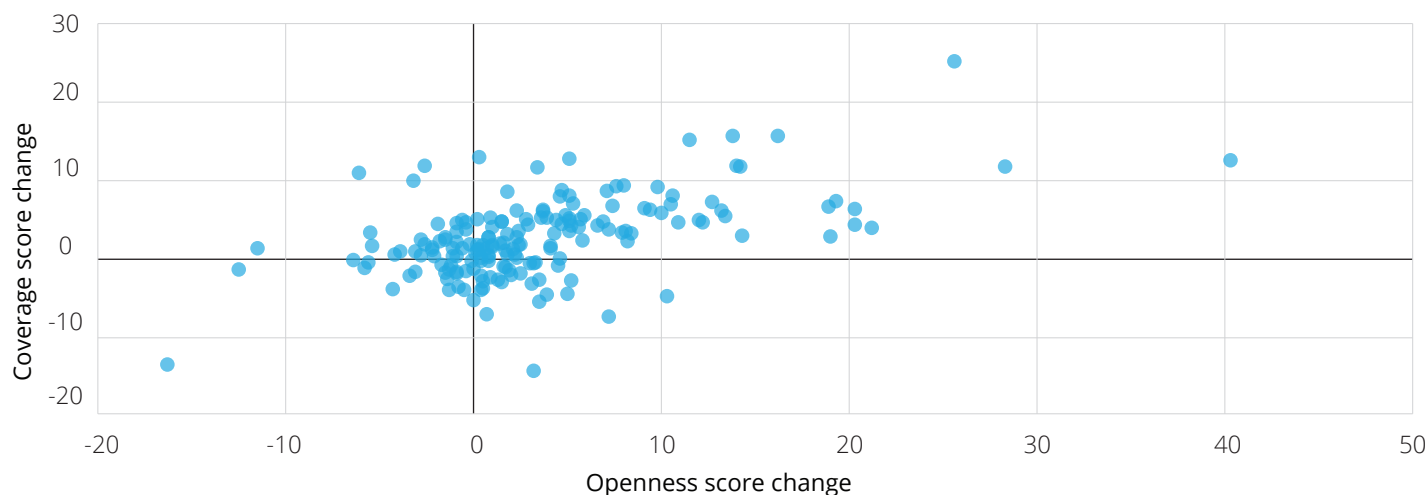
| Income Group | 0-20 | 21-40 | 41-60 | 61-80 | 81-100 |
|---------------------|------|-------|-------|-------|--------|
| High income | 2 | 19 | 26 | 46 | 7 |
| Upper-middle income | 10 | 31 | 53 | 6 | 0 |
| Lower-middle income | 9 | 61 | 27 | 2 | 0 |
| Low income | 14 | 71 | 14 | 0 | 0 |

Though there remains a strong relationship between income level and overall ODIN scores, every year some of the largest improvements are made by countries in low- and middle-income groups. You can read more about these countries in the section Improvements Since 2016.

3.2 Openness Drives Progress

As the ODIN 2018/19 scores illustrate, it is easier and more cost effective for countries to increase their score by making improvements to the openness of existing data while working to collect and publish new data. Therefore, results show the most dramatic improvements within countries occurred in openness scores. Of the 51 countries that made an overall increase in score of five or more points since 2017, the median coverage score increase was 6.7 and the median openness increase was 9.4. Figure 3 shows each countries' openness and coverage score change between 2017 and 2018. The greatest positive changes in openness were also accompanied by coverage score increases. However, coverage score increases were typically more modest changes in comparison.

Figure 3. Changes in coverage and openness scores, 2017-2018



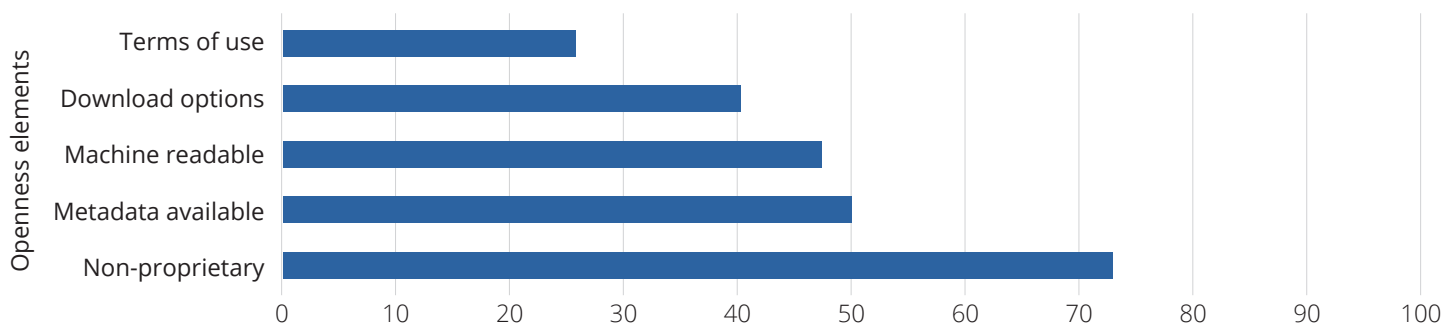
The countries with the greatest positive change in openness can be seen in the top right quadrant of Figure 3 (Morocco +40, Singapore +28, Romania +26, Hong Kong +21, Cyprus and Tunisia +20). Romania had the highest coverage change of +25 points, followed by Morocco +13, Singapore +12, Tunisia +6, and Cyprus and Hong Kong +4.

Openness scores and coverage scores are not independent because data availability and comprehensiveness are key factors in openness. The five openness elements cannot be assessed when data are not publicly available. Therefore, increases in coverage due to the release of new datasets nearly always lead to some increase in openness. For example, Singapore, by publishing additional datasets on land use and energy use, not only boosted their coverage score by 8 points, but also boosted their openness score by 13 points. Out of the 113 countries that increased their coverage scores by 1-point or more, 71% also increased their openness score by at least 1-point. In comparison, of the 37 countries that decreased their coverage scores by 1-point or more, only 38% increased their openness score by at least 1-point. The first step to making data open is to make datasets publicly available without any barriers to access.

3.3 Openness is a Spectrum

The five elements of openness used in ODIN mirror the principles of the [Open Definition](#) and the [Open Data Charter](#). Figure 4 shows average scores for each element of openness for 2018. In ODIN, openness is considered on a spectrum. Rarely are datasets fully open or closed. According to Figure 4, the highest scoring openness element is non-proprietary formats. An average of 73 percent of all data assessed in ODIN are made available in non-proprietary formats. Non-proprietary formats are important because they allow users to access data without requiring the use of a costly, proprietary software that may prevent some users from accessing the data. This number has increased steadily since ODIN 2016 from 67 to 73.

Figure 4. Average scores by openness element, 2018



All ODIN openness elements have seen improvements over the years, but certain elements have made greater strides than others. Between 2017 and 2018, the openness elements with the highest average increase were terms of use, with an average increase of 20 points since 2017, and metadata availability, with an average increase of 10 points. Download options had the lowest average increase of 8 points, likely because they are the costliest change to implement.

Many of the countries that have seen the largest positive changes in openness amended or adopted an open terms of use policy or data license. Others, redesigned or built a new data portal with improved capabilities and user features. Below is a summary of some countries' notable efforts.

Notable efforts to increase openness

Serbia: In the last year, the [Statistical Office of the Republic of Serbia](#) redesigned its website and published terms of use for their data. Although their [terms of use](#) could be further improved, their combined efforts resulted in a 10-point score increase overall and a 13-point increase in openness. On the newly designed website, data and metadata are not only easier to find, but more comprehensive.

Sierra Leone: Like Serbia, [Statistics Sierra Leone \(SSL\)](#) launched a new website and adopted [terms of use](#) with similar results. Between 2017 and 2018, Sierra Leone's overall score increased from 33 to 43 and their openness score increased from 33 to 46. Statistics Sierra Leone is a great example of how small but effective website updates can lead to improved accessibility without the expensive infrastructure required to build complex data portals.

Tunisia: In 2018, Tunisia saw a 20-point increase in openness. Like others, [Tunisia's National Institute of Statistics \(INS\)](#) also adopted an [open data license](#) that impacted openness scores across every category. Additionally, more information was published in their pre-existing data portal, resulting in higher scores in three additional categories that received zero scores in previous years.

Romania: In last year's assessment, Romania's score decreased drastically, partly due to their portal's mandatory registration requirement before users could access the data, which created a barrier to access. This year, Romania's [National Institute of Statistics \(INS\)](#) launched a [new portal](#), removing this barrier and including additional datasets. In 2018, Romania's coverage score increased from 34 to 59 and its openness score increased from 36 to 62.

Morocco: In 2018, most datasets assessed in ODIN were published in Morocco's new database created by the [High Commission for Planning \(HCP\)](#). The database features newly published data, increased openness of existing datasets, and an open [terms of use](#). Between 2017 and 2018, Morocco's openness score increased from 25 to 65. All openness elements increased in score except metadata availability, although HCP has stated that work in this area is on their future agenda.

Moldova: [Statistics Moldova](#) decided to amend its [terms of use](#) and omit a clause that would forbid "misleading use" of their data. After consultation with Open Data Watch. "Misleading clauses" are considered inimical to openness because they are unclear and may be enforced arbitrarily to discriminate against users. With the removal of this clause, Statistics Moldova's terms of use are now in full compliance with recommended open data standards.

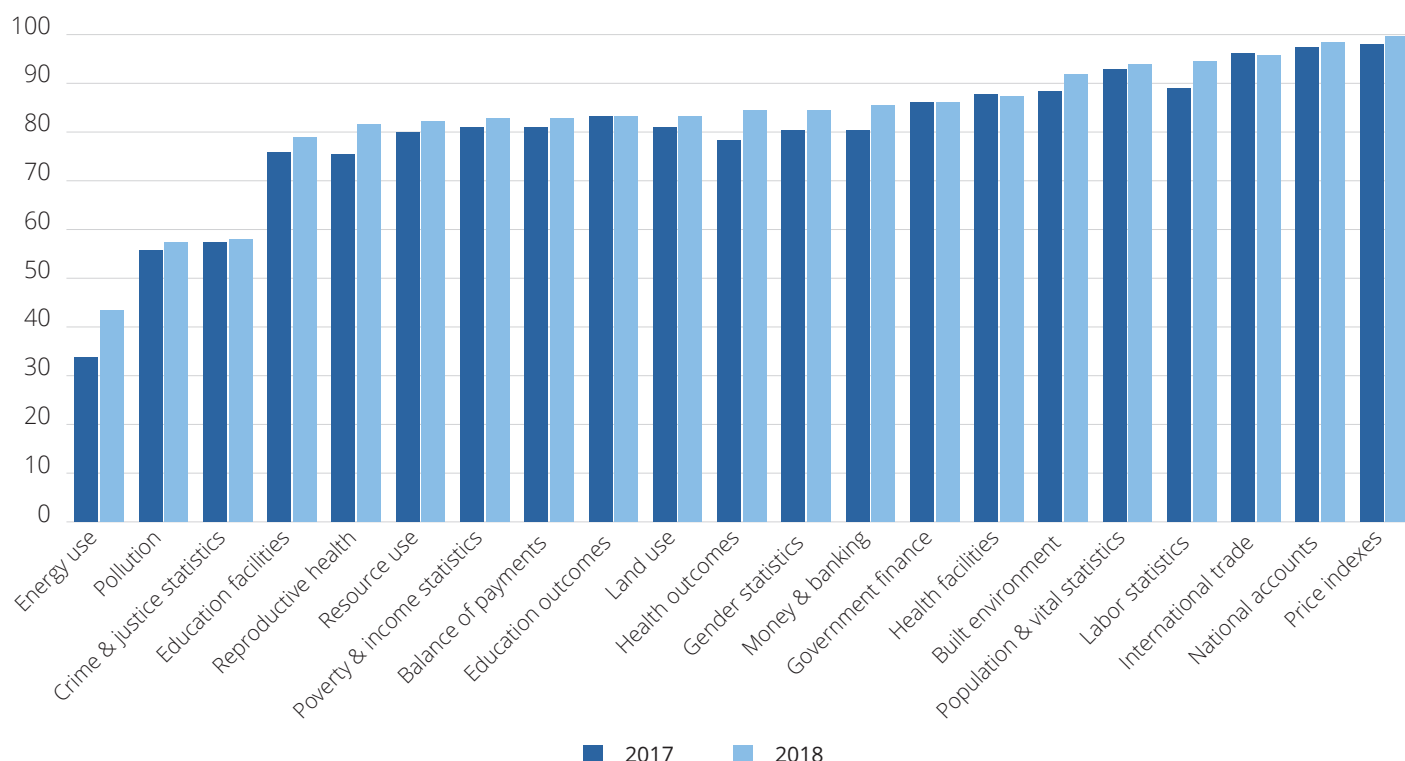
Palestine: Since the 2015 ODIN, Palestine has made significant progress in the openness of official statistics and between 2017 and 2018, [Palestine Central Bureau of Statistics \(PCBS\)](#) increased their overall score by 11 points and their openness score by 19 points. This jump is due to the release of more data in machine-readable and non-proprietary formats, more available metadata, and the creation of new [terms of use](#) that are fully open.

Oman: In 2018, the [National Center for Statistics \(NCSI\)](#) in Oman invited Open Data Watch to conduct a training with their staff on open data principles and the ODIN methodology to help inform their open data initiatives. Subsequently, NCSI made several changes that increased both their coverage and openness scores and launched them into the number one ranked position in Western Asia. Across all 21 categories, their openness scores are nearly perfect except for the two categories, energy and pollution, where no data are published.

3.4 The Missing Data

While the availability of data, or data coverage, saw slight improvements in 2018, much of it can be explained by methodology changes. The fact remains, data availability remains a challenge in many categories with persistent data gaps. Several countries continue to not publish datasets despite their importance for monitoring Sustainable Development Goals (SDGs) and other measures of progress. Figure 5 shows the percentage of countries in 2018 that published at least some data in each of the 21 categories assessed.

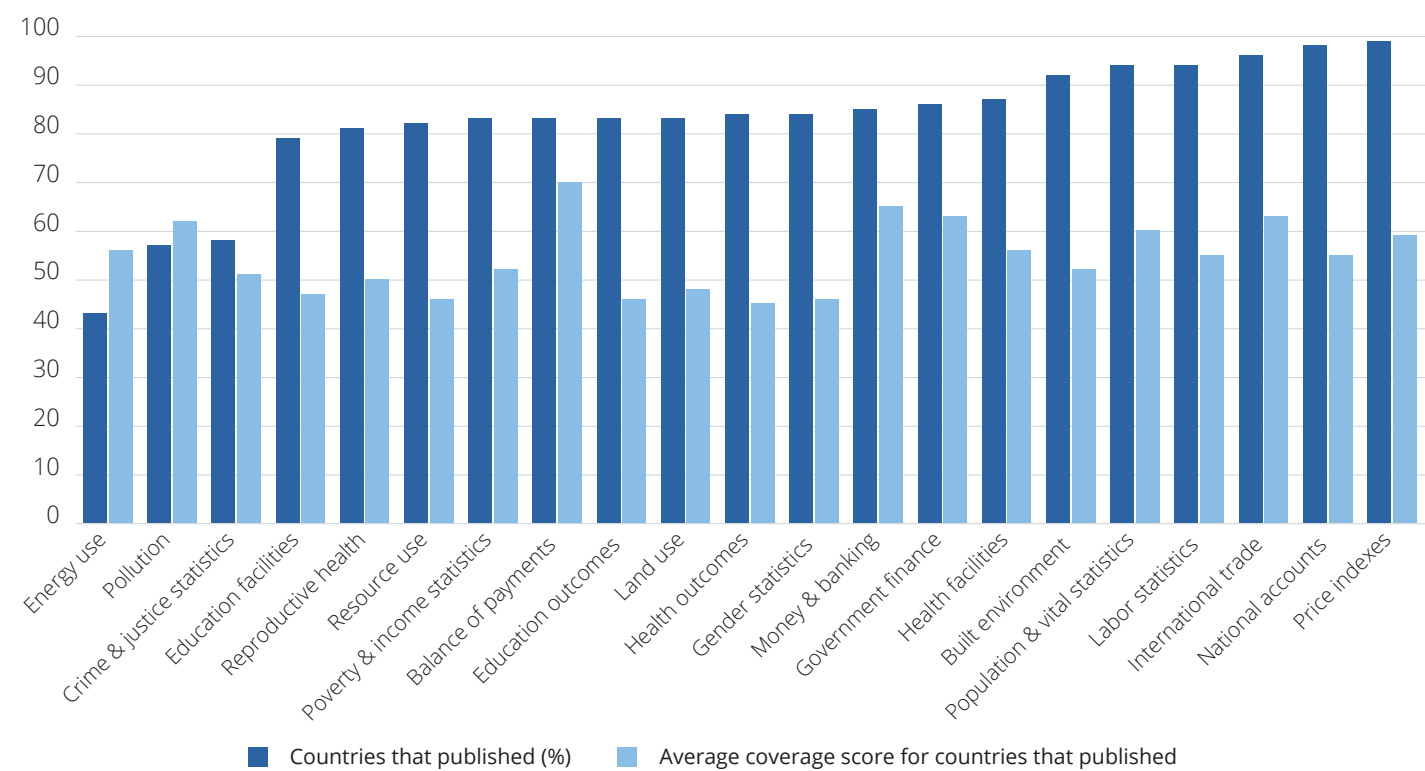
Figure 5. Countries that publish (%) by category, 2017-2018



Data on crime and justice, pollution, and energy use remain severely underreported, meaning a small percentage of countries publish any data in these categories. However, the countries that do publish data in these categories, do so relatively comprehensively.

For instance, average coverage scores for the category pollution are the 5th highest among all publishing countries, energy use is 8th, and crime and justice is 14th (See Figure 6). Data on health outcomes, resource use and education outcomes, though reported by over 80 percent of countries, perform at least four percentage points lower on average. These results demonstrate that the reasons for missing data and underreporting vary by category. Data least reported by countries (Figure 6, dark blue) are simply not prioritized in data collection, while data with low coverage (Figure 6, light blue) are not collected often enough or with the necessary disaggregation.

Figure 6. Countries that publish (%) by category and corresponding coverage scores, 2018



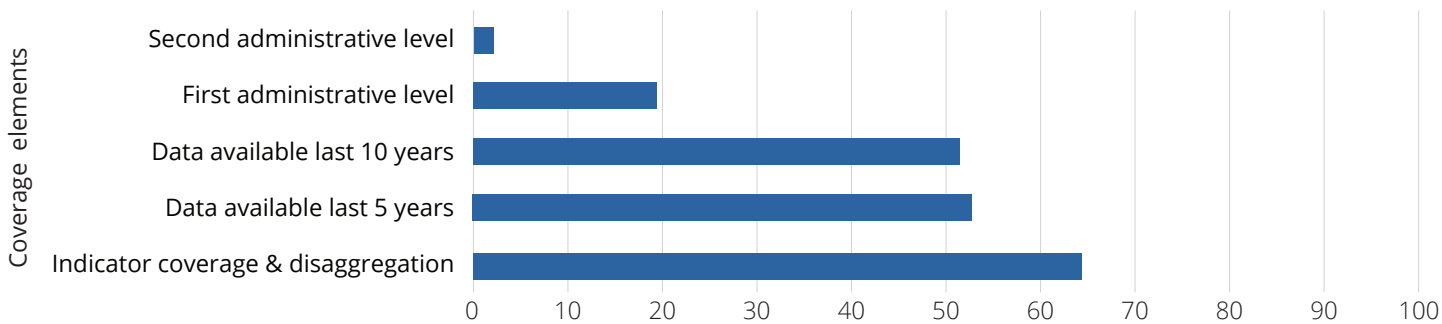
3.5 Different data need different strategies

Data coverage scores are based on five elements:

- (1) representative indicators are available and are disaggregated appropriately;
- (2) data are available for the preceding five years;
- (3) data are available for the preceding ten years;
- (4) data are disaggregated at the first administrative level;
- (5) data are disaggregated at the second administrative level.

Looking at the specific elements of coverage gives insight into which elements countries struggle with most and whether this differs between data categories. Figure 7 shows the average scores for the five elements of data coverage for 178 countries in 2018. Scores reflect the percentage share of possible scores on the five elements averaged over the 21 data categories.

Figure 7. Average scores by coverage element, 2018



The pattern of scores in Figure 7 have remained largely unchanged since the first ODIN assessment in 2015. Geographic disaggregation below the national level has been a predominant issue for countries across regions and income levels. Because some types of data are not commonly compiled at the subnational level, ODIN excludes eight of the 21 data categories from disaggregation at the second administrative level and three categories from disaggregation at the first administrative level. This means that countries are not penalized for not publishing data with geographic disaggregation for those categories. Nonetheless, very few datasets in the remaining categories are available at the first or second administrative levels.

Perhaps the most interesting findings show how coverage element scores differ between data categories. Figure 8 shows the average scores for the five coverage elements for each of the 21 data categories.

Figure 8. Countries that publish (%) by category and corresponding coverage scores, 2018

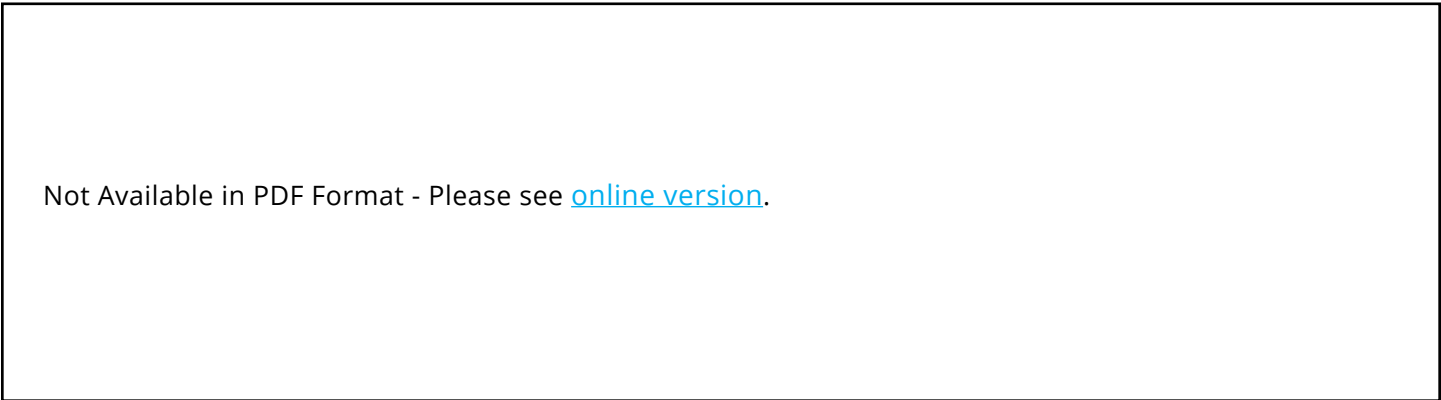


Figure 8 shows where efforts of improvement should be focused in countries. In general, economic and financial data categories score significantly higher on 5- and 10-year data availability, while social indicators are more likely to have subnational data at the first and second administrative level. Other results show that data such as poverty and income, gender, and built environment statistics are more likely to be published in the last 5 years rather than the last 10, suggesting that data are not collected often enough or older datasets are being removed from websites.

The first coverage element, indicator coverage and disaggregation, has the highest score in the category labor statistics, meaning countries are most likely to publish all representative indicators assessed in ODIN and their recommended disaggregations in this category. Government finance statistics is the second highest scoring category but lacks data available at the first administrative level. Although most governments have adopted regular publishing practices for national budget data, state or provincial budget data are rarely found.

Addressing gaps in coverage requires a system-wide approach to modernize national statistical data systems to address common issues, such as lack of subnational disaggregated data. At the same time, ODIN results show that different types of data may need different approaches since many gaps are not constant across data categories.

3.6 Improvements Since 2016

2018 is ODIN's fourth year running. There are now three comparable years to analyze countries' progress (read about why 2015 is excluded in section, Updates in 2018). Between 2016 and 2018, there are 169 countries that have been assessed every year. The following section looks at the countries that have made considerable and continued progress since 2016. Though these countries are not the always the highest scoring, their efforts should be acknowledged.

Table 3 shows the 20 countries that made the greatest improvement between 2016 and 2018. These countries have had the highest percentage increase in overall scores since 2016 and have improved by at least a 5 percent each year.

Table 3. Twenty most improved since 2016

| Country | Region | Income Group | ODIN 2018/19 Score | Relative Increase Since 2016 |
|--------------------------|--------------------|---------------------|--------------------|------------------------------|
| Jamaica | Caribbean | Upper-middle income | 58.5 | 228.9 |
| The Gambia | Western Africa | Low income | 32.4 | 113.3 |
| Hong Kong SAR, China | Eastern Asia | High income | 71.2 | 96.3 |
| Singapore | South-Eastern Asia | High income | 84.9 | 53.5 |
| El Salvador | Central America | Lower-middle income | 39.8 | 49.4 |
| St. Vincent & Grenadines | Caribbean | Upper-middle income | 36.7 | 46.8 |
| Oman | Western Asia | High income | 64.3 | 36.7 |
| Israel | Western Asia | High income | 52.9 | 36.5 |
| Myanmar | South-Eastern Asia | Lower-middle income | 45.8 | 36.2 |
| Ghana | Western Africa | Lower-middle income | 39.2 | 36.2 |
| Slovakia | Eastern Europe | High income | 74.3 | 30.0 |
| Honduras | Central America | Lower-middle income | 38.4 | 25.5 |
| Cyprus | Western Asia | High income | 63.6 | 36.2 |
| Guatemala | Central America | Lower-middle income | 35 | 30.0 |
| France | Western Europe | High income | 62.5 | 25.4 |
| Malaysia | South-Eastern Asia | Upper-middle income | 48.7 | 25.1 |
| Qatar | Western Asia | High income | 40.1 | 24.2 |
| Belize | Central America | Upper-middle income | 37.7 | 23.9 |
| Luxembourg | Western Europe | High income | 55.7 | 21.2 |
| Japan | Eastern Asia | High income | 65.1 | 19.4 |

Unlike the top ten highest scoring countries, which all are high-income, the 20 most improved are from a variety of regions and income groups. Six countries are low- or lower-middle income and 14 are upper-middle or high income. Their regional diversity is notable as well. Central America and Western Asia makes the most frequent showing, followed by South-Eastern Asia. In total, at least one country from half of all global regions is represented among the top 20 most improved countries.

The diversity among these countries proves that open data is not only an aspiration for high-income countries, but a feasible aspiration for all countries around the globe, regardless of income level. Jamaica is a great example of how a country under the strong leadership of its national statistical office can accomplish a great deal in a short time.

Jamaica: The Making of an Open Data Champion

In 2016, Jamaica received an overall score of 18, ranking 165th out of 173 countries. At that time, Jamaica published data on only half of ODIN's data categories, mostly in non-machine-readable formats. The country had serious coverage and openness gaps to address.

The [Statistical Institute of Jamaica \(STATIN\)](#) were concerned about their state of open data and knew there was a disconnect between their actions and recommended open data standards. By the summer of 2017, Jamaica was already making plans to improve. At the invitation of STATIN, Open Data Watch and [PARIS21](#) conducted a three-day training for all government statistical producers to not only learn more about open data, but how to design a National Strategy for the Development of Statistics (NSDS) that would help address their coverage and openness gaps. By the end of 2017, some of their initial efforts had come to fruition and

their ODIN score increased by 18 points, largely driven by the adoption of open data standards. Between ODIN 2017 and 2018, their work continued, this time focusing on increasing the coverage and availability of data. In ODIN 2018/19, Jamaica made another jump in score from 44 to 59 by publishing more datasets not previously publicly available. They did so by increasing their coordination with other statistical producers in the country. With minor adjustments to their existing data portal, the creation of an [open terms of use](#), and increased coordination between domestic statistical producers, Jamaica rightfully earned the number one spot of top 20 most improved countries with a score increase of over 200% in three years.

As evident, large improvements have been made by countries in all parts of the world. Figure 9 shows the progress made by countries within each region by comparing their global ranks since ODIN 2016.

ODIN classifies countries in 20 regions, based on the United Nations [Standard Country or Area Codes for Statistical Use](#) (M49). Seventeen out of 20 regions had at least one country that made an improvement in rank between 2016 and 2018. In Figure 9, you can view the most improved countries in each region, including the 20 most improved listed in Table 3.

Figure 9. Country ranks over time by region, 2016-2018

Not Available in PDF Format - Please see [online version](#).

In Central America, Eastern Asia, South America, South-Eastern Asia, Western Asia, and Western Europe, over 50 percent of countries made an improvement in rank between 2016 and 2018. To see the most improved countries from each region that didn't make the top 20 most improved list, select the region from the drop-down menu in Figure 9.

Progress measured by rank doesn't always indicate that a country made any commitments towards open data since ranks can change based on the actions of other countries. However, looking at rank changes over time is one way to measure whether open data practices in a country are improving, staying static, or falling behind

4. Special Topics

4.1 Gender Data in ODIN 2018

ODIN has seven data categories that include indicators of importance to measuring the status and well-being of women. The coverage and openness scores for each country can be seen in the [Country Profile](#) view, and a composite score can be constructed using ODIN's custom weight function.

Six of the seven gender data categories are from the social statistics domain. Not included are the data categories for health facilities, educational facilities, and poverty, which cannot be disaggregated by sex. The seventh category is labor statistics from the economic and financial statistics domain, which includes data on the employment and unemployment rates for women and men. Average overall scores for each category and their rank among the 21 ODIN data categories are shown by income group in Table 4.

Table 4. Average scores for gender-related Indicators by income group, 2018

| Income Group | Low income | | Lower-middle Income | | Upper-middle income | | High income | |
|-------------------------------|------------|------|---------------------|------|---------------------|------|-------------|------|
| | Score | Rank | Score | Rank | Score | Rank | Score | Rank |
| Population & vital statistics | 36.1 | 11 | 42.1 | 9 | 56.9 | 2 | 76.9 | 1 |
| Education outcomes | 31.6 | 14 | 34.4 | 14 | 31.7 | 17 | 49.3 | 15 |
| Health outcomes | 37.5 | 10 | 38.3 | 12 | 31.4 | 18 | 43.8 | 19 |
| Reproductive health | 38.0 | 8 | 40.8 | 11 | 39.2 | 13 | 43.4 | 20 |
| Gender statistics | 30.5 | 15 | 34.2 | 15 | 35.3 | 16 | 48.7 | 17 |
| Crime & justice | 15.4 | 19 | 21.0 | 19 | 29.0 | 20 | 43.9 | 18 |
| Labor | 35.9 | 12 | 42.1 | 10 | 51.3 | 7 | 67.2 | 8 |

Scores in each data category generally rise with average income. This is particularly true of population and vital statistics, which is ranked 11th among 21 categories in low-income countries but doubles in score and rises to 1st in the high-income group. Many low- and lower-middle-income countries rely on decennial censuses and infrequent surveys to update their demographic data and lack complete civil registration systems to record births and death. The categories of gender statistics and crime and justice statistics rank even lower. Both include statistics on assaults and gender-based violence that frequently go unreported. The coverage and openness of these categories is ranked low even in high-income countries. Two categories that display a contrary trend are health outcomes and reproductive health, which are ranked higher and, in some cases, have higher scores in low-income than in middle-income countries. This presumably reflects the support provided by donor-funded survey programs such as the Demographic and Health Surveys and Multiple Indicators Cluster Surveys.

The evidence from ODIN 2018/19 is that gender statistics and the larger set of data needed to measure women's well-being are still underreported and, when published, fall short of the standards for open data.

CRVS Systems and Gender Data

In addition to surveys supplying important gender data within countries, strong civil registration and vital statistics (CRVS) systems can improve the coverage of gender data. Most obvious perhaps are the links of CRVS systems to population and vital statistics. However, additional categories such as health outcomes, crime and justice, and reproductive health stand to reap many important benefits from well-functioning CRVS systems. A [new series](#) from the Centre of Excellence at the International Development Research Centre explores the links between CRVS systems and improving data on women and girls.

4.2 Terms of Use and Open Data Licensing

Open terms of use are one of the five criteria of openness in ODIN and often include disclaimers of warranty, limitations of liability and, most importantly, data licenses. Adopting an open license is a core component of the definition of open data. Data cannot be open unless it is licensed for reuse.

Nonetheless, many countries neglect adopting them. This may be due to the incorrect assumption that publishing data online implies authorization of public use. However, there can be widely different interpretations of what authorized use looks like to governments and users. An open data license is an opportunity for governments to encourage public use of their data by specifically addressing how people can use data, how they should attribute data, and what types of use, if any, are prohibited. These specifics encourage use and reuse of data by alleviating user concerns about legal ramifications of unapproved use.

Open Knowledge International's [Open Definition](#) presents nine criteria that a license must have to be considered fully open. It also lists certain conditions that can be added to a license without affecting its openness. A summary of these can be found in Table 5. In general, when referring to open statistical data, these criteria can be summarized as "[data that can be freely used, re-used and redistributed by anyone- subject only, at most, to the requirement to attribute and sharealike.](#)"

Table 5. The Open Definition guidelines on open licenses

| Open Data Licenses Must Allow: | Open Data Licenses Can Require: |
|--|---|
| 1. Free use | 1. Attribution |
| 2. Redistribution (including sale) | 2. Integrity (modified versions must carry a different name) |
| 3. Modification | 3. Share-alike (distributions must remain under the same license) |
| 4. Separation (any part of the work can be used, distributed or modified separately) | 4. Notice (retention of copyright notices) |
| 5. Compilation (can be distributed along with other works) | 5. Source (provide recipients with access to the source data) |
| 6. Non-discrimination against any persons or group | 6. Technical Restriction Prohibition (measures cannot be added to restrict access) |
| 7. Propagation without the need to agree to additional legal terms | 7. Non-aggression (in some cases, a license may require modifiers to grant the public additional permissions) |
| 8. Application to any purpose | |
| 9. No charge | |

The [Open Data Charter](#), currently adopted by 64 national and local governments, also mentions the importance of data licensing under Principle 3 and uses the Open Definition. It is generally recommended that governments adopt a Creative Commons (CC) license, or something similar, that meets these criteria, such as [CC0](#) (public domain), [CC BY 4.0](#) (with attribution), or [Public Domain Dedication and License](#) (PDDL).

ODIN 2018/19 results show that over half of the countries assessed have no published terms of use or they adopt restrictive data licenses for their statistical data. Only 14 countries have a fully open terms of use for all statistical data (see Table 6). However, none of these country received a perfect score in 2018 because none published data in every category. In many cases, countries have multiple data licenses or terms of use for different datasets on the same website or for datasets on different government agency websites. This approach is not recommended unless they are part of a broader government licensing framework and are clearly and coherently connected.

ODIN penalizes countries that adopt licenses or terms of use that do not grant all nine authorizations specified in the Open Definition (or have extra requirements beyond what is considered acceptable (see Table 5). ODIN classifies these as: [fully] open, semi-open, or restrictive/absent. Table 6 shows the percentage of countries in ODIN with various combinations of data licenses/terms of use and their openness level.

Table 6. Country data licenses/terms of use by openness level, 2018

| Openness Level | Percentage of Countries | Number of Countries |
|--|-------------------------|---------------------|
| All license/terms of use fully open | 8 | 14 |
| Combination of fully open and semi-open | 15 | 26 |
| Combination of fully open, semi-open and restrictive or absent | 2 | 4 |
| All licenses/terms of use semi-open | 8 | 14 |
| Combination of semi-open and restrictive or absent | 15 | 26 |
| All licenses/terms of use restrictive or absent | 53 | 94 |

Fourteen countries publish all data assessed in ODIN under fully open licenses/terms of use and 94 countries publish under restrictive ones or without one. Fifty-six countries publish data under multiple licenses/terms of use with different levels of openness, 44 of which have at least one that is fully open.

In 2017, 40 countries had at least one fully open data license/terms of use. In 2018, that number increased to 44. Some of the new additions include Oman and Moldova. In Moldova, the National Bureau of Statistics removed a condition prohibiting “misleading use” from their license. In Oman, they removed several restrictive conditions from their [terms of use](#) that forbid any type of use without receiving prior permission. This language is problematic because it is unclear and may be enforced arbitrarily to discriminate against users.

Another new addition, Singapore, improved their [terms of use](#) after removing two conditions that were not considered acceptable by the [Open Definition](#). The two conditions prohibited “presenting the Contents in a misleading or incorrect manner or misrepresenting the Contents;” and any use that would “damage our [SingStat’s] reputation.” The “misleading use” clause is problematic for the same reason as Moldova’s as it may be enforced arbitrarily to discriminate against users. The prohibition of use that would “damage” their reputation is restrictive for that reason as well. Also, by forbidding use that would “[prejudicial to the author’s honour or reputation](#),” or similar, Singapore would be claiming their moral right to the integrity of their work. However, in the view of Open Data Watch, the application of the moral right of integrity to statistical data is inconsistent with the rationale behind the Open Definition, which is based on the premise that open data should be reused freely without any civil or criminal liability attaching. With these clauses excluded, Singapore’s data license is now considered to be fully open, helping them to become the number one ranked statistical system in this year’s index.

Further elaborating on these discussions, here are some commonly asked questions by governments and their answers:

Does statistical data need a license to be open data?

Yes. For data to be considered open, it must be licensed as such. If the data is not licensed for reuse, any reuse would be unlawful.

In many countries, statistical data are not copyrightable since they do not meet the evidence of originality required by their copyright laws. In these cases, these data are considered part of the public domain. However, even in these instances, adopting an open data license such as CC0 or using a public domain mark (PDM) is still beneficial from a user perspective because it gives legal clarity and confirms that the data producer will not make a later claim of intellectual property rights or copyright infringement.

In the case where a government wants to license data separately from other types of work, they should adopt an open data license and specify in their terms of use what types of work the license applies to.

What is the best license to adopt?

There are several licenses countries can adopt for their data that would be fully compliant with the [Open Definition](#). As already mentioned, a [CC0](#) (public domain), [CC BY 4.0](#) (with attribution), or [Public Domain Dedication and License](#) (PDDL) are commonly used. However, a government can create a custom license if they meet the open license criteria and, ideally, explicitly state that the license is compatible with the Open Definition. Such examples include the [United Kingdom's Open Government License Version 3, Data License Germany- Version 2](#), [Canada's Open Government License \(Version 2\)](#), and [Taiwan's Open Government License Version 1](#). Previous versions of these licenses are not considered compliant with open data standards.

Many of these licenses include provisions where a country agrees to waive their moral rights, such as the right to integrity discussed previously. However, in some cases national law forbids the waiving of moral rights. In these cases, licenses such as CC0 and PDDL are ideal, as they include language for these circumstances.

Data

[Explore ODIN data](#)

Methodology Report

[ODIN 2018/19 Methodology Report](#)