ODIN

OPEN DATA INVENTORY 2020/21

by OPEN DATA WATCH

EXECUTIVE SUMMARY

OPEN DATA WATCH

Acknowledgements

EXECUTIVE SUMMARY

The 2020/21 Open Data Inventory (ODIN) is the fifth edition of the index compiled by Open Data Watch. ODIN 2020/21 provides an assessment of the coverage and openness of official statistics in 187 countries, an increase of 9 countries compared to ODIN 2018/19. The year 2020 was a challenging year for the world as countries grappled with the COVID-19 pandemic. Nonetheless, and despite the pandemic's negative impact on the capacity of statistics producers, 2020 saw great progress in open data.

However, the news on data this year isn't all good. Countries in every region still struggle to publish gender data and many of the same countries are unable to provide sex-disaggregated data on the COVID-19 pandemic. In addition, low-income countries continue to need more support with capacity building and financial resources to overcome the barriers to publishing open data.

ODIN is an evaluation of the coverage and openness of data provided on the websites maintained by national statistical offices (NSOs) and any official government website that is accessible from the NSO site. The overall ODIN score is an indicator of how complete and open an NSO's data offerings are. It is comprised of both a coverage and openness subscore. Openness is measured against standards set by the **Open Definition** and **Open Data Charter.** ODIN 2020/21 includes 22 data categories, grouped under social, economic and financial, and environmental statistics. ODIN scores are represented on a range between 0 and 100, with 100 representing the best performance on open data.

Below is a summary of the main findings from ODIN 2020/21. The full report will be released in February 2021.

Figure 1. ODIN Global Scores, 2020/21
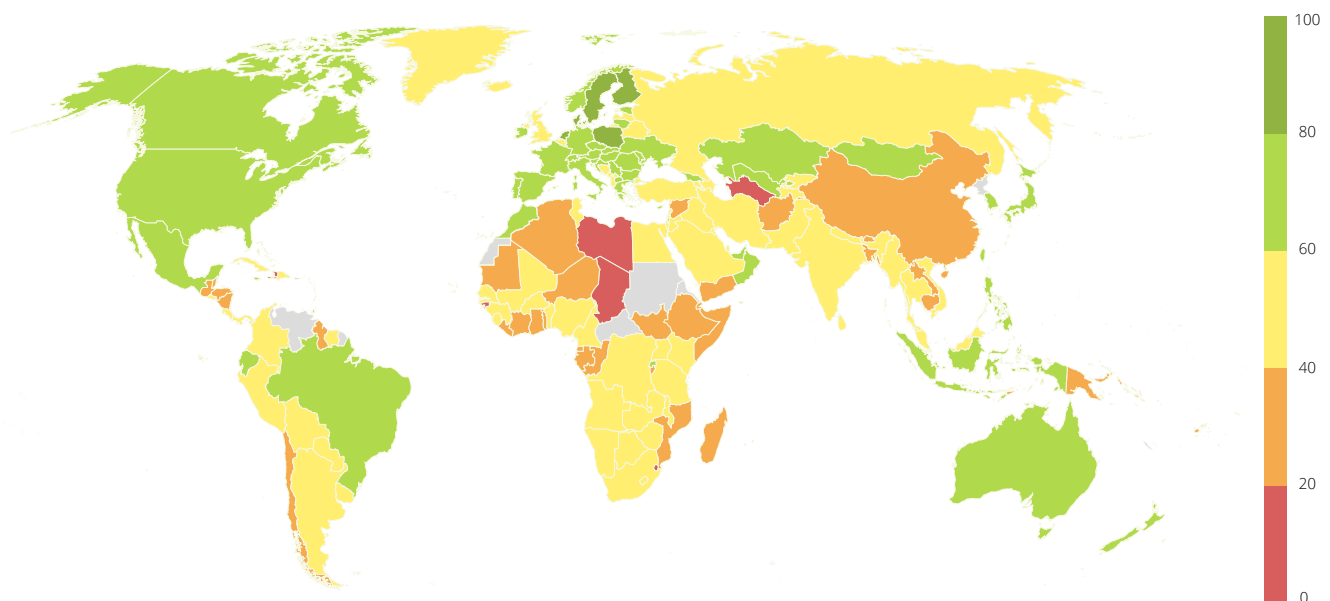
Open statistical data is on the rise, with countries demonstrating the greatest progress to date

The median ODIN score **increased by 6.4 points to 48.8** in 2020. This is the largest jump in any year since ODIN began. There was also a strong upward trend in the subscores for coverage and openness. In 2020, median coverage scores increased by 4.2 points and median openness scores increased by 10.5 points, tripling progress seen in the previous year. In 2020 there were more publicly available data from national statistical offices, and they were easier to access and share, than ever before.

Figure 2. Median ODIN Scores, 2016-2020/21
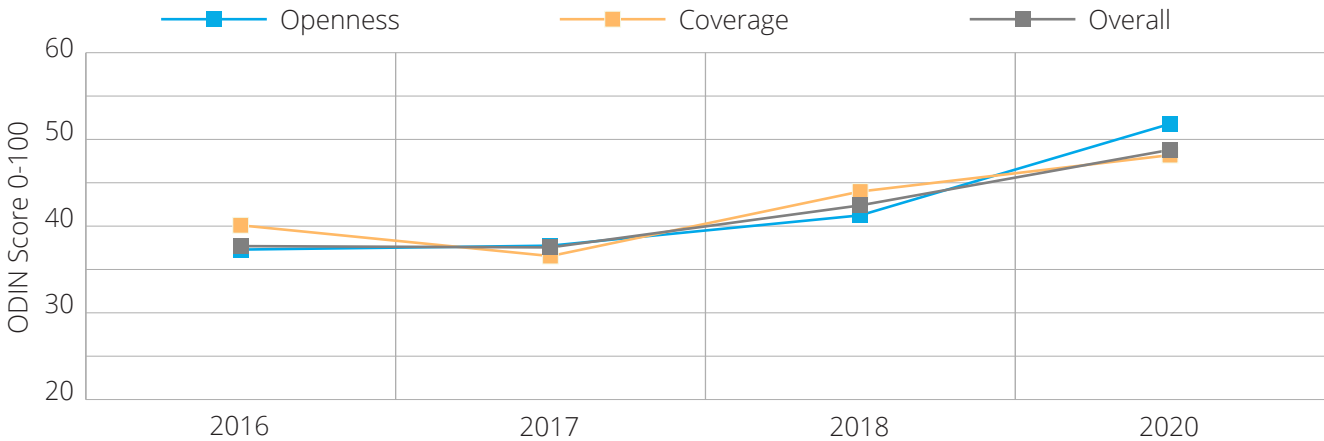
The impressive global progress on open data is thanks to the increasing number of data champions; 75 percent of countries increased their score since 2018. Although 24 percent of countries' scores decreased (and a few remained unchanged), they decreased by smaller amounts: Countries that increased their score had a median increase of 5.7 points, while countries whose score decreased saw only slight drops with a median decrease of -2.1.

**Figure 3.** Change in ODIN Scores, 2018/19-2020/21
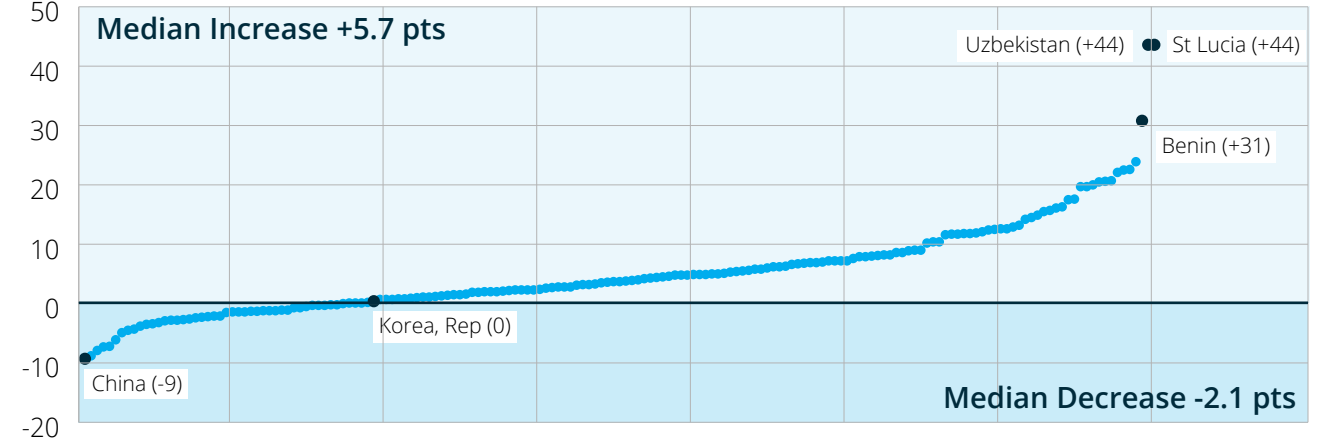
The top 30 countries are diversifying

The top performing 30 countries in ODIN have mostly remained the same since 2016, but that's starting to change. Twenty-three countries have appeared in the top 30 every year since 2016. They were joined by 5 new countries in 2018 and two more countries reached the top 30 for the first time in 2020. The newcomers are Palestine and the United Arab Emirates, both of which are countries that committed to making open data a priority over the last year.

0   20   40   60   80   100

Palestine
In 2020, Palestine built upon the ten-point increase they earned in 2018 and gained nearly 20 additional points for an overall score of 72 this year. Such an achievement was made possible by both publishing more data and opening up the data already published, as improvements were seen equally in data coverage and openness. Their coverage score increased because they published the indicators that were missing in 2018, as well as more recent data and more data at the subnational level. Openness scores were driven largely by making more data available in machine-readable formats through an option to download HTML tables from the website in Excel format. Palestine's Central Bureau of Statistics' **website** has continued to improve to meet users' data needs. In 2019, they launched a **Data Science Initiative** with the Arab American University of Palestine to build the capacity of students to use their data. They also participated in an open data workshop led by Open Data Watch in late 2019 to coordinate their open data strategy.

0   20   40   60   80   100
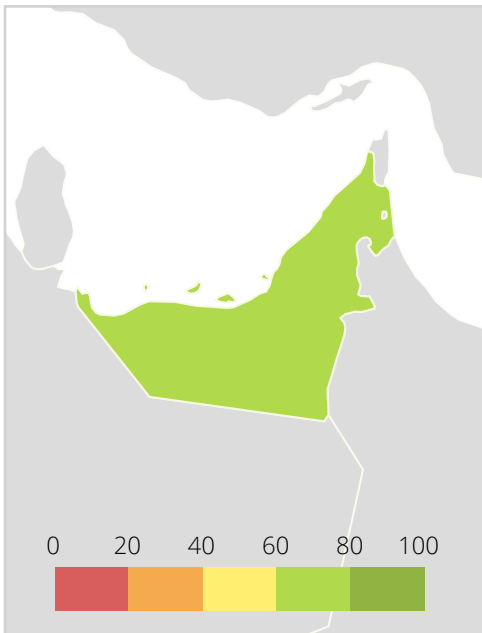
United Arab Emirates (UAE)
The UAE also took a dual approach of making more data available and improving elements of data openness. In 2018, their ODIN assessment showed they did not publish over 30 percent of the indicators reviewed by ODIN, which severely impacted their coverage and openness score as you cannot have open data until you have published data. In 2020, the Federal Competitiveness and Statistics Authority launched the Open Data Race — a competition among statistics-producing government agencies to see who could publish more open data through their **Bayanet data portal**. As a result, they reduced the number of missing indicators to 13 percent and their coverage score increased by over 20 points. Publishing data through the portal also resulted in 100 percent of ODIN indicators being made available in machine-readable format and made it easier to standardize the metadata that was made available for each dataset, doubling their metadata availability score.

Countries in the Africa and the Pacific Islands made the most significant improvements

Between 2018 and 2020, median scores increased in nearly every region, but countries in Africa and the Pacific Island saw some of the largest improvements as a region. The table below shows the median regional score changes, as well as the most improved countries in each region. Though countries in Africa and the Pacific Islands improved most collectively, the most improved individual countries were St. Lucia (Caribbean) and Uzbekistan (Asia) with increases of 44 points each. The only region that has no countries increase score was North America, where both Canada and the United States decreased scores slightly, largely due to changes in the assessment methodology, not necessarily changes in the countries' practices.

Table 1. Changes in Regional Median ODIN Scores, 2018/19-2020/21

Region	Number of Countries	Median Change 2018-2020	Most Improved Countries
Caribbean	9	2.4	St. Lucia (+44)
Africa	48	5.2	Benin (+31), Tanzania (+23), Angola (+22)
Pacific Islands	6	8.05	Marshall Islands (+12), Fiji (+10)
South and Central America	19	.7	Suriname (+21)
Asia	47	5	Uzbekistan (+44), United Arab Emirates (+24), Iraq (+23)
Europe	42	2.05	Ukraine (+21), Serbia (+16)
Australia and New Zealand	2	-.05	New Zealand (+5)
North America	2	-3	No countries improved

**Figure 4.** Regional Median ODIN Scores Increases, 2018/19-2020/21
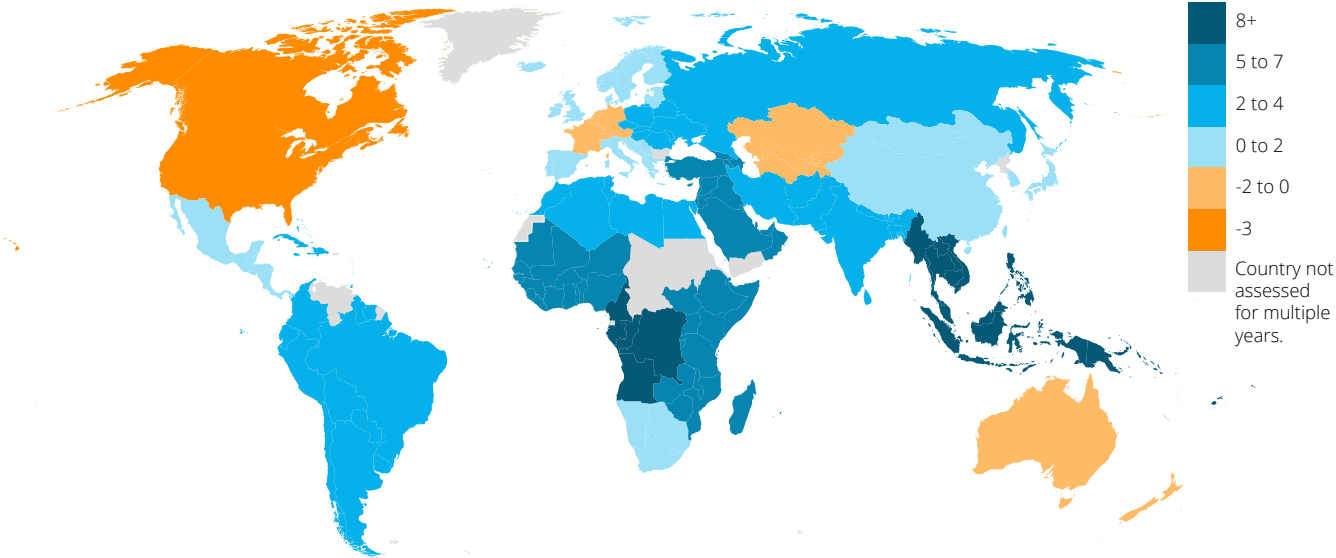
# Three countries that made great progress toward open data in 2020
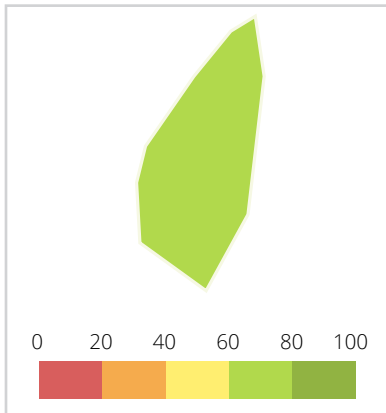
St. Lucia
St. Lucia has made the most progress of any country in any year with an increase of 44 points. In 2018, St. Lucia ranked in the bottom ten of all countries and has climbed to the 51st highest rank. In late 2018, the Central Statistical Office of Saint Lucia launched a new website and has since been dedicated to ensuring that not only are more data published, but that the published data are more open. In 2020, St. Lucia published 66 percent of the ODIN indicators, 16 percent more than what was published in 2018. In addition, they adopted a new open terms of use, published their data in XLSX and PDF files throughout the website, and made more metadata available.
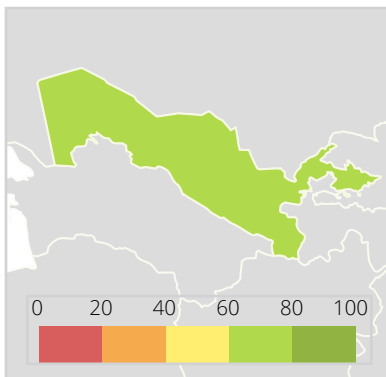
Uzbekistan
Uzbekistan matched St. Lucia's progress with an increase of 44 points. In 2018, they only published 39 percent of indicators, but in 2020 this increased to 73 percent. Like St. Lucia, they focused their openness efforts on adopting an open data license and ensuring data was published in machine-readable and nonproprietary formats throughout the website. In May 2020, the State Committee of the Republic of Uzbekistan on Statistics attended an open data workshop hosted by Open Data Watch in collaboration with the Organization for Security and Co-operation in Europe.
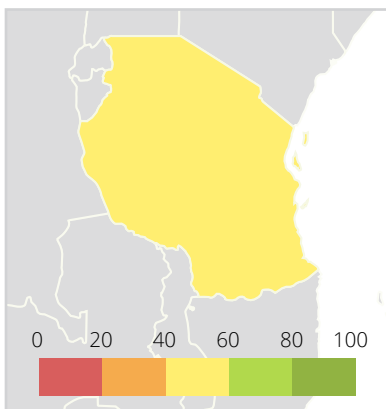
Tanzania
This year, Tanzania increased its score by 23 points. Their improvements were driven by efforts to increase the openness of data previously published. However, some new datasets were published as well, specifically those that included subnational data. Openness efforts focused on the publication of an open data license and making more data available through the Tanzania Social-Economic Database, which makes it easier to publish data in machine-readable formats. Hopefully these actions set a new precedent, in sharp contrast to the 2018 amendments added to their Statistical Law that made it a crime to publish statistics without prior approval. Since 2019, those amendments have been removed.

Lower-income countries struggle most with making data open, demonstrating the need for increased financial resources and capacity-building

Improvements in openness scores have driven most countries' progress. However, implementing data openness remains a problem for many low-income countries. When comparing openness and coverage scores across income groups, the difference between median openness scores of high- and low-income countries is nearly twice as high as the difference in median coverage scores.

Table 2. Median ODIN Scores by Income Group, 2020/21

Income Group	Overall Median	Coverage Median	Openness Median
High income	62.3	55.6	68.0
Upper-middle income	52.7	49.7	54.1
Lower-middle income	43.1	47.4	42.5
Low income	40.0	39.0	37.8
Difference between high- and low-income	**22.3**	**16.6**	**30.2**

Openness scores are measured against five openness elements (based on the Open Definition and Open Data Charter) and are shown in Figure 6. The biggest difference between high- and low-income countries is their ability to make data available in machine-readable formats and to adopt open data licenses. Many low-income countries only publish data in PDF format, which is nonproprietary but not machine readable. When data are made available in formats that are not machine readable, users cannot easily access and work with the data, which severely restricts the scope of the data's use. Depending on how a country manages their data, converting data from PDF files to machine-readable formats can be a time-consuming, manual task. More resources to help improve data management systems could greatly improve a country's ability to make data available in various formats with little effort.

Terms of Use is another openness element that has a large gap between high- and low-income countries. This issue, unlike machine readability, is not related to a need for financial support, but rather capacity building. Through Open Data Watch's engagements with countries, it has become clear that the main reason most countries lack an open terms of use or license is because of the lack of knowledge about open data and lack of technical and legal capacity to create the license.

Figure 5. Median Openness Element Scores by Income Group, 2020/21
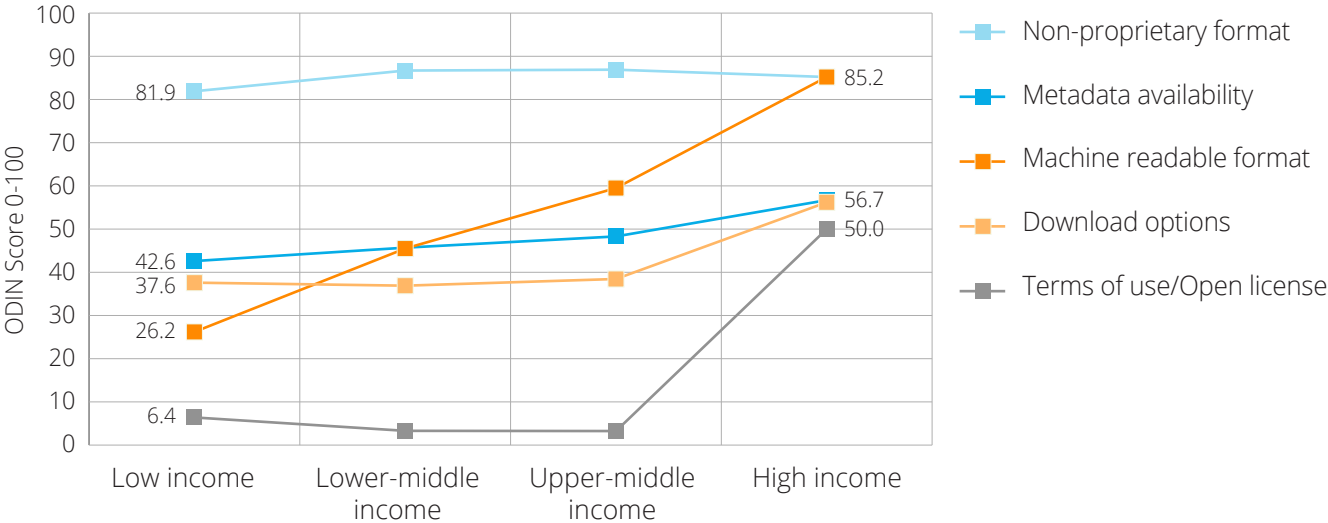
OPEN GENDER DATA

Gender data refers to data that are disaggregated by sex or that measure conditions and events that have a bearing on the welfare of women and their children. These data are used to identify specific needs, formulate policies to address shortcomings, and monitor their impact on women and their families. The ODIN – Open Gender Data Index (ODIN-OGDI) is a separate score based on the availability of 20 ODIN indicators in 8 statistical categories that require sex-disaggregated data or apply only to women. We include these 8 categories in the ODIN-OGDI along with two more categories whose data are not sex-disaggregated but have important consequences for women.

The ten data categories in the ODIN-OGDI are:

Sex-disaggregated
1. Population and vital statistics
2. Education outcomes
3. Health outcomes
4. Reproductive health
5. Food security and nutrition
6. Gender statistics

7. Crime statistics
8. Labor statistics

Not sex-disaggregated
9. Poverty statistics
10. Built environment

Combining the scores on these data categories yields a measure of the coverage and openness of gender statistics.

Scores on the ODIN-OGDI have been rising, but not as fast as non-gender data categories

Since 2016, the median score on the ODIN-OGDI has risen by 21 percent, but scores on the remaining non-gender-related categories have risen by 40 percent.

Figure 6. Median ODIN-OGDI and Non-Gender Scores, 2020/21
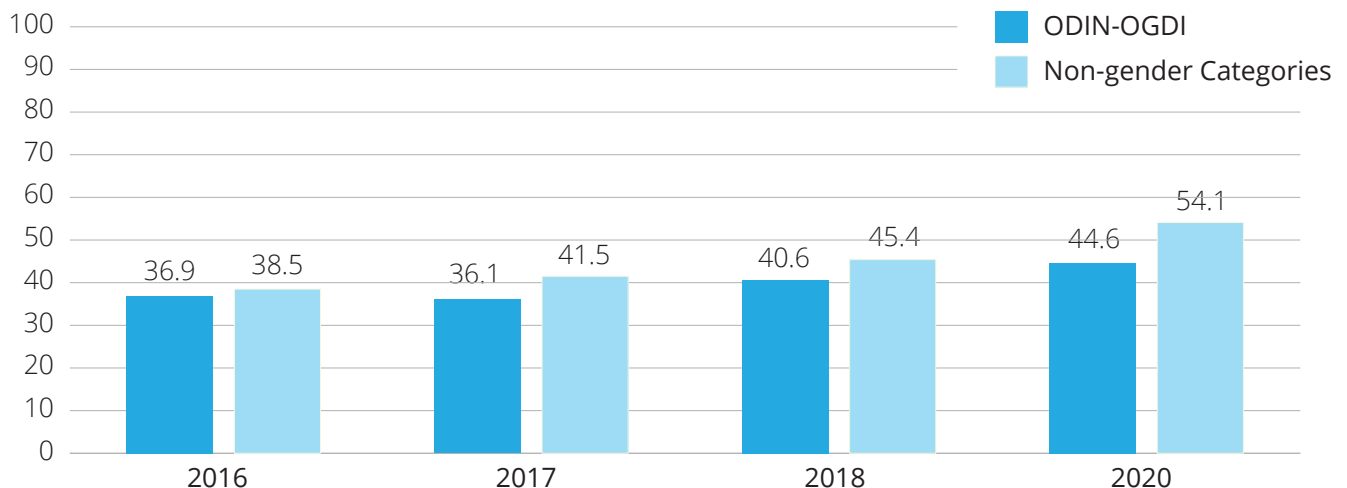
Countries that score low on the ODIN gender index are less able to provide sex-disaggregated data on the COVID-19 pandemic

Figure 7. ODIN-OGDI Scores by Countries Reporting Sex-Disaggregated COVID-19 Data, 2020/21
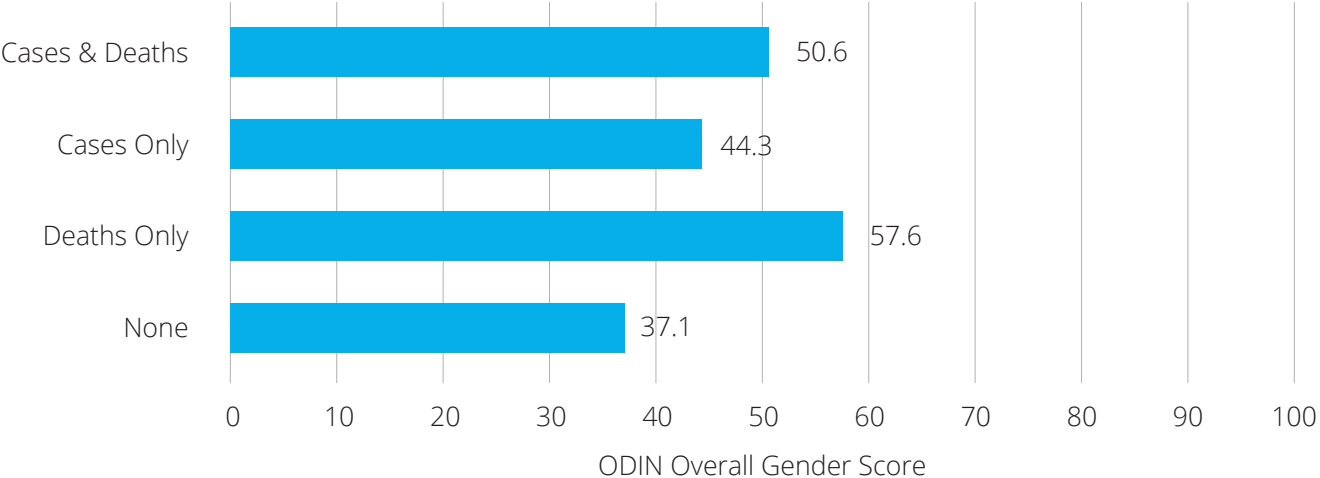
There is a 13.5 point difference in the ODIN-OGDI scores of the 86 countries that provide sex-disaggregated data on both the case and death rates from the COVID-19 pandemic and 63 that provide neither. There were 119 countries in the ODIN 2020/21 assessment with sex-disaggregated data on COVID-19 cases available in the **Global Health 5050 COVID-19 Sex-Disaggregated Data Tracker** and 68 countries without sex-disaggregated data as of 16 November, 2020. Death rates were less well-reported. Only 91 countries have data on male and female deaths from COVID-19 (of which 5 report deaths only), and 96 report no sex-disaggregated data on COVID-19 deaths.

Low scores on the ODIN-OGDI are not found only in poor countries

The lowest scoring countries on the ODIN-OGDI include countries across the income spectrum and of large and small size. Haiti, Turkmenistan, Anguilla, and Eswatini are among the lowest-ranked countries in ODIN 2020/21. China ranks 23 places higher on the overall ODIN index than it does on the ODIN-OGDI, but like other low-scoring countries, it lacks data for sex-disaggregated or gender-relevant indicators.

Table 3. Ten Lowest Scoring Countries on ODIN-OGDI, 2020/21

Country	Income Group	ODIN-OGDI (0-100)
Haiti	Low	0.0
Turkmenistan	Upper-middle	0.0
Anguilla	High	8.3
San Marino	High	8.5
Eswatini	Lower-middle	10.8
Libya	Upper-middle	11.5

Country	Income Group	ODIN-OGDI (0-100)
Gabon	Upper-middle	13.5
St. Kitts and Nevis	High	15.0
China	Upper-middle	16.0
Equatorial Guinea	Upper-middle	17.5

Crime and justice statistics are the least available data in the gender index

Sixty-five countries received no score in the Crime & justice category. Crime statistics that are assessed for sex-disaggregated data include the homicide rate, rate of other crimes, and data on the prison population. Crimes of violence against women are not specifically included in the ODIN assessment, but the Sustainable Development Goals specify five indicators that report on forms of violence including two specifically concerned with violence against women. If countries reported these indicators with sex-disaggregated data, they would be included in their ODIN scores. Frequent and accurate reporting of crime statistics is needed to halt the epidemic of femicide and violence against women.

Table 4. Data Categories Least Available in ODIN-OGDI, 2020/21

ODIN gender categories	Number of countries with no score
Population & vital statistics	8
Education outcomes	16
Health outcomes	25
Reproductive health	32
Food security & nutrition	48
Gender statistics	21
Crime & justice	65
Labor	9
Poverty & income	30
Built environment	35

IN CLOSING

ODIN 2020/21 strives to be an effective tool to assist countries in making progress toward open data. With each assessment, we continue to expand data sets included and countries covered. Through our program of country engagement, we encourage communication to addresses countries' practical concerns, while remaining independent. The new developments of the ODIN website highlight the future frontiers of open data, complementing our work in promoting data use and better data governance, both essential components of a sustainable approach to open official statistics. The ODIN 2020/21 Annual Report explores these themes further, as well as going deeper into the topics already discussed in this summary. Visit our website to read the Annual Report in late February, prior to the 52nd UN Statistical Commission in 2021.

ODIN

OPEN DATA INVENTORY

by OPEN DATA WATCH

OPEN DATA WATCH