TABLE OF CONTENTS

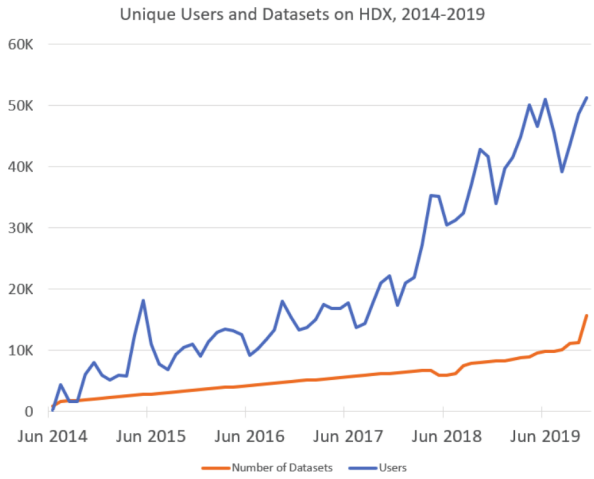
Acknowledgements

This report was produced by the United Nations Office for the Coordination of Humanitarian Affairs (OCHA) Centre for Humanitarian Data in The Hague, which manages the Humanitarian Data Exchange platform. OCHA thanks all of the organizations that have shared data through HDX, the donors who have supported this work over the years, and the HDX users who are committed to ensuring humanitarian response is data driven. For additional information, contact the Centre for Humanitarian Data at centrehumdata@un.org.

1. INTRODUCTION

The goal of this report is to increase awareness of the data available for humanitarian response activities and to highlight what is missing, as measured through OCHA's Humanitarian Data Exchange (HDX) platform. We want to recognize the valuable and long-standing contributions of data-sharing organizations. We also want to be more targeted in our outreach on what data is required to understand crises so that new actors might be compelled to join the platform. Data is not an end in itself but a critical ingredient to the analysis that informs decision making. With nearly 168 million people in need of humanitarian assistance in 2020[1] – the highest figure in decades – there is no time, or data, to lose.

When HDX was launched in 2014, it held around 800 datasets. Over the past five years, that number has skyrocketed to over 17,000 datasets. The data covers every active humanitarian crisis, from Afghanistan to Yemen, and has been shared by dozens of organizations, from ACLED to WFP.[2] In 2019, HDX was accessed by over 600,000 users.



This is a tremendous achievement for collective action in a sector that relies on cooperation. It also shows the value of an open data platform. OCHA's work to aggregate data from many sources in one place has undoubtedly created efficiency in the system. Humanitarians, donors, academics, and journalists no longer need to chase contacts to locate data; they can go to HDX and search for it. If the data is not there, the HDX team will help find it.

> "Accurate data is the lifeblood of good policy and decision-making. Obtaining it, and sharing it across hundreds of organizations, in the middle of a humanitarian emergency, is complicated and time-consuming – but it is absolutely crucial.[3]"
>
> - United Nations Secretary-General António Guterres at the opening of the OCHA Centre for Humanitarian Data in The Hague in December 2017

One downside to all of this data sharing is knowing what data is most relevant to understanding a crisis context. In May 2019, HDX added a new feature called the Data Grid to help people in their quest for good and relevant data. Based on extensive user research, the Data Grid places the most important crisis data into six categories: affected people; coordination and context; food security and nutrition; geography and infrastructure; health and education; and population and socio-economy.

[1] 2020 Global Humanitarian Overview: **https://www.unocha.org/sites/unocha/files/GHO-2020_v9.1.pdf**

[2] Humanitarian data is defined as: 1) data about the context in which a humanitarian crisis is occurring; 2) data about the people affected by the crisis and their needs; and 3) data about the response by organizations and people seeking to help those who need assistance.

[3] Taken from the opening remarks of the UN Secretary-General at the launch of the Centre for Humanitarian Data in The Hague in December 2017: **https://www.un.org/sg/en/content/sg/speeches/2017-12-22/opening-centre-humanitarian-data-remarks**

Within each category, there are several sub-categories. For example, within affected people, there are sub-categories for internally displaced people, refugees, returnees, humanitarian needs, and casualties. Within coordination and context, there are sub-categories for 'who is doing what where', affected areas, humanitarian access, and funding, among others. (See Annex A for all sub-category definitions).

There are three main criteria for whether relevant data is included in the Data Grid: 1) disaggregated beyond the national level; 2) commonly-used formats; and 3) timeliness. If at least one dataset meets all criteria, that sub-category is considered 'complete'. If at least one dataset meets some of these criteria, the sub-category is considered 'incomplete'. If a dataset does not meet the criteria or does not exist on HDX, the sub-category is considered empty or as having no data. (See Annex B for details on the Data Grid criteria).

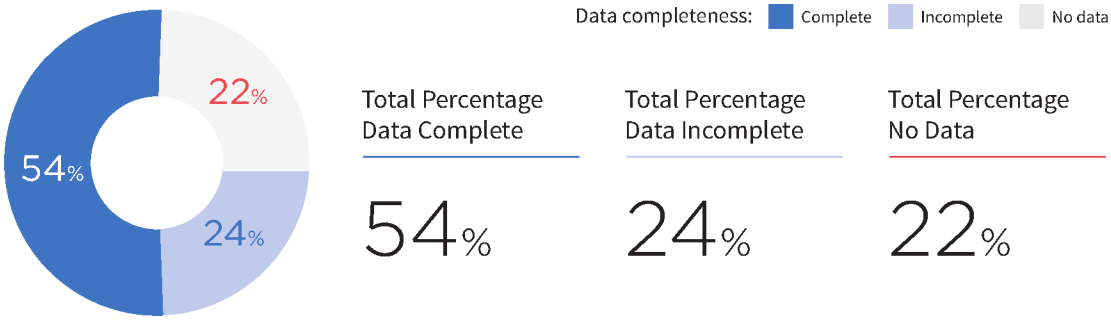
DATA GRID CRITERIA: 1) SUB-NATIONAL; 2) COMMONLY-USED FORMATS; AND 3) TIMELINESS.

Complete	Incomplete	No Data
At least one dataset in the sub-category meets all criteria.	At least one dataset in the sub-category meets some criteria.	Available data in the sub-category does not meet the criteria or does not exist on HDX.

Of course, relevant data will greatly depend on who is looking and what they are looking for. A dataset might have the right data, but not cover the part of the country needed for the analysis. Or it might cover the right geographic area but be in a format that is difficult to work with. For this reason, the HDX team reviews all relevant datasets and assesses them against the criteria. This careful curation process is undertaken daily on all newly shared or updated datasets. So far, some 700 datasets have been taken through this process.

By the end of December 2019, HDX included Data Grids for 14 locations.[4] These include: Afghanistan, Bangladesh, Central African Republic, Chad, Colombia, Democratic Republic of the Congo, Mozambique, Myanmar, the Philippines, State of Palestine, Somalia, Sudan, Venezuela and Yemen. We will expand the Data Grids to cover all locations with a Humanitarian Response Plan throughout 2020.[5] We may also expand the categories and sub-categories based on feedback.

As we start 2020, the completeness of all Data Grids combined is 54 percent. That is, 54 percent of relevant, comprehensive data is available across 14 locations. If we add the data that is relevant but incomplete, the total is 78 percent. This leaves 22 percent of categories with data that does not meet the criteria or with no data. The Data Grids include an average of 20-30 datasets per location.



It is important to note that not all humanitarian data can be shared openly. Data about the location of affected people and responders can put people at risk, especially in conflict environments. The HDX Terms of Service[6] prohibit the sharing of data that includes personally identifiable information. For sensitive, non-personal data

[4] A location refers to a country or territory.

[5] See all response plans here **https://www.hpc.tools**. A response plan may be for a country, region, or unique to a specific crisis.

[6] Organizations sharing data through the HDX platform should ensure that the data was collected in a legal, ethical and responsible manner. For more on the HDX Terms of Service, visit **https://data.humdata.org/about/terms**