

World Development Report 2021: Data for Better Lives

Technical Note on Data Sources and Reproducing Data Analysis

Nikkie Pacheco and Elizabeth Salazar

The World Development Report 2021 (WDR 2021) Data Catalog collection includes all available source data and R/STATA replication files (for specific figures) necessary to create the figures and maps that appear in [WDR 2021: Data for Better Lives](#). These files have been posted on the World Bank Data Catalog under the [World Development Report 2021 Dataset Collection](#). The link¹ in the source note for each figure/map of the Report leads to their respective data catalog entry. Each dataset corresponding to a figure is a dataset in the collection and the dataset name is associated with the chapter number². Each data catalog entry also provides further details regarding the dataset including the full sources (in some cases, data are used from internal databases) and the license to use (in cases where some are restricted data and further permission is required from the original owner). Some of the figures and maps are also presented as data visualizations in the [WDR 2021 Data Stories website](#).

Users can reproduce figures individually after downloading the files. All data sources used for the creation of figures and maps are available in comma-separated values (CSV) file format. Variables that appear in original datasets but were not used in the analysis for the WDR 2021, have been removed from the data sets. At the same time, new variables and data labels have been added to facilitate further analysis. In cases where country-level or aggregate data were taken from published tables or figures, only those values are provided. Otherwise, the data files provide original data, as pulled from online databases, statistical annexes, published tables or figures, and unpublished datasets shared by scholars and researchers.

In most cases, all observations in the original data set were retained, even if the final figure does not include them. This is especially true for the figures in Chapters 2, 6 and 8. Three new World Bank databases are launching at the same time of the report and the chapters use the data behind them. These databases are as follows: (1) [Statistical Performance Indicators](#) (Chapter 2), (2) [Global Data Regulation Survey](#) (Chapter 6), and (3) [Data Governance Indicators](#) (Chapter 8).

Finally, maps were generated by the World Bank Cartography Unit using the [World Bank Official Boundaries](#).

¹ Full list of figures and maps with their respective links:

<https://www.worldbank.org/en/publication/wdr2021/brief/world-development-report-2021-maps-and-figures>

² In most cases, the dataset's name in the collection is the same as the figure's name in the report. For example, Figure 1.1 is named "World Development Report 2021 - Figure 1.1". On the other hand, this varies for duplicate figures and figures that share the same source within a chapter.