

Data sources for Data Visualization projects

Data search, collection, examination, transformation, and exploration will account for 70% or 80% of the effort and time that you will dedicate to the Data Visualization project. To facilitate this process, this supplement provides some tips, clarifications, and suggested sources to find and collect data.

1) Do I need to use datasets that contain data up to 2020 or 2021 for COVID-related projects?

It depends on the story that you want to tell. For instance, if you want to explore the impact of COVID on other contexts as education, economy, etc. Or make projections. Then, yes, you need historical and actual data from all datasets that you are going to use. On the other side if you, want to explore correlations that may provide clues of why effects of COVID are different from country to country, you may not really need data from 2020-21 (except COVID cases or deaths of course). For instance, say you want to know whether countries that better manage the safety in roads (<https://roadskillmap.com/>), or health infrastructure, are also the ones that better managed the COVID crises (although if it exists, this just may be a spurious correlation). For this scenario, you can easily work with accidents data from 2019 and before.

2) What are good sources of data?

There are some well-known sources. In some cases you may need to download the data year by year:

<https://www.kaggle.com/datasets>. Here you can find datasets in all kind of topics, and COVID of course, just search by topic and sort them by last update.

<https://ec.europa.eu/eurostat/web/main/data/database>. All kind of European statistics. Some of them updated up to December 2020. Also includes georeferenced data

<https://www.openstreetmap.org/>. Geo located data as point of interests or areas (always updated).

<https://www.globalclimatemonitor.org/> Climate data including 2020.

https://www.ine.pt/xportal/xmain?xpgid=ine_main&xpid=INE. All kind of statistics about Portugal, some datasets updated to 2021.

<https://trends.google.com/> trends about any topic.

<https://www.google.com/covid19/mobility/>. Community mobility reports.

<http://insideairbnb.com/get-the-data.html>. Data about Airbnb listings.

There are also plenty of other websites that may offer very rich and open data but are very specific about the topic or about the country. These are harder to find and you should be familiar with the country, language, and topic for which you are looking the data for. For instance, COVID deaths may not be very well documented in developing countries, so you may need to explore the total number of deaths from other sources and try to find inconsistencies.