

Updated: April 24, 2020

COVID-19 Dataset and Visualization Collection

By [Data Collaborations Lab](#), Carnegie Mellon University Libraries.

Attribution: CC-BY-NC 4.0

Summary

During the COVID-19 pandemic, clinicians, epidemiologists, scientists and data scientists joined hands to share information related to the disease as open data and open source tools, including datasets, visualizations, simulations. We are collecting some of these together to share with the community who are interested in either exploring these data, or to inform themselves about the COVID-19 case, distribution, progression, or what is published in the literature. Due to the limited bandwidth we have, this is by no means a complete collection. We welcome any collaboration, contrition, questions, or feedback! Email us: dataCoLAB@andrew.cmu.edu

Legend:

[***] Highly recommended resources

Case Datasets:

- [***] The COVID Tracking Project: collects information from 50 US states, the District of Columbia, and 5 other US territories to provide the most comprehensive testing data we can collect for the novel coronavirus, SARS-CoV-2. We attempt to include positive and negative results, pending tests, and total people tested for each state or district currently reporting that data: <https://covidtracking.com/>
- [***] Live streaming COVID-19 data from John Hopkins University, initially in spreadsheet
<https://docs.google.com/spreadsheets/d/14quQPFErG-hlpsrNgYcX85vW7JMMK5X2vNZrafRcH8c/edit#gid=2087034741>, now update in GitHub repository:
<https://github.com/CSSEGISandData/COVID-19>
- A collection of data sources, as part of the [sars2pack project](https://seandavi.github.io/sars2pack/articles/datasets.html), from Sean Davis et al:
<https://seandavi.github.io/sars2pack/articles/datasets.html>
- [***] COVIDcast - a project led by [Delphi Research Group](#) at CMU. The dataset includes doctor visits data, [symptoms survey](#) (Facebook and Google), Google search trend, and

flu tests (Quidel): <https://covidcast.cmu.edu> An open API can be found here:
<https://cmu-delphi.github.io/delphi-epidata/api/>

Publications Datasets:

- The WHO's downloadable database:
<https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov>
- [***] [COVID-19 Open Research Data](#) (CORD-19): Machine readable publications dataset on COVID-19 related publications, curated by [Allen Institute for AI](#) and hosted by Semantic Scholar.
- Collection from Dimensions database, including publications, datasets, clinical trials.
https://www.dimensions.ai/news/dimensions-is-facilitating-access-to-covid-19-research/?utm_source=pardot&utm_medium=email&utm_campaign=newsletter20&utm_term=dimensions
and <https://covid-19.dimensions.ai/>
- [***] The Cushing/Whitney Medical Library's public collection of COVID-19 citations: The collection, accessible through [this public Zotero library](#) (no account required), includes daily updates from PubMed, ClinicalTrials.gov, Disaster Lit, and preprints* from bioRxiv and medRxiv.
- [***] [LitCovid](#): Curated downloadable collection of more than 1,200 journal articles hosted by the National Library of Medicine
- Resources collected by a Wikidata group:
https://www.wikidata.org/wiki/Wikidata:WikiProject_COVID-19

Misc Datasets:

- Collections of COVID-19 research datasets and publications deposited in public repositories including: figshare <https://covid19.figshare.com>, Zenodo <https://zenodo.org/communities/covid-19/>, and Harvard Dataverse <https://dataverse.harvard.edu/dataverse/2019ncov>
- A comprehensive versioned dataset of the repositories and related metadata about **public projects** hosted on GitHub related to COVID-19:
<https://github.com/github/covid-19-repo-data>

- Center for Open Science recently launched the [Coronavirus Outbreak Research Collection](#) in Open Science Framework, in response to the urgent need for aggregated and discoverable research on the virus.

Visualizations and Dashboards:

- ******* A really nice corona-data-simulator that's shows the positive effects of social distancing <https://www.washingtonpost.com/graphics/2020/world/corona-simulator/>
- ******* JHU real-time tracker dashboard: <https://coronavirus.jhu.edu/map.html>
- ******* COVID-19 case tracker (by John Burn-Murdoch @ Financial Times): <https://www.ft.com/coronavirus-latest>
- COVID-19 dashboards: <https://covid19dashboards.com>
- Infection Trajectory: See Which Countries are Flattening Their COVID-19 Curve: <https://www.visualcapitalist.com/infection-trajectory-flattening-the-covid19-curve/>
- World meter: <https://www.worldometers.info/coronavirus/>
- COVID-19 tracker: covid19.nguy.dev
- Covid-19: active cases and extrapolation: [Covid19-dash.github.io](https://covid19-dash.github.io)
- [COVID-19 Open Research Map](#): Interactively engage with COVID-19 research output and explore connections between publications
- ******* View Allegheny county COVID-19 data from the local health department: <https://www.alleghenycounty.us/Health-Department/Resources/COVID-19/COVID-19.aspx>

Other Collections of Resources

- A list of COVID-19 dashboards on GitHub: <https://github.com/CSSEGISandData/COVID-19/issues/576#issuecomment-609454667>
- An article that includes a comprehensive list of datasets and visualizations: <https://www.zdnet.com/article/as-covid-19-data-sets-become-more-accessible-novel-coronavirus-pandemic-may-be-most-visualized-ever/>
- Best COVID-19 resources: <https://www.visualcapitalist.com/7-best-covid-19-resources/>

- eLife community discussions:
<https://elife.discourse.group/t/fast-question-what-covid-tracker-dashboard-do-people-in-our-country-use/2514>

Slide Deck

- How to fight the coronavirus SARS-CoV-2 and its disease, COVID-19.
<https://drive.google.com/file/d/1ZaiDO87me4puBte-8VycSRtpQ3PVpkK/view?fbclid=IwAR2SPvfFb4zutunpCYzjRjSuyiX8sgJnPZiedV68rstEVk018rkgOPRBYSY>