

Property Analytics – Our World in Data Project

Edward Yiu

What is Our World in Data (OWiD) Project?

- Max Roser is the founder of Our World in Data in 2011.
- The OWiD is a project of the [Global Change Data Lab](https://ourworldindata.org/#entries)
- All free: open access and open source
- <https://ourworldindata.org/#entries>

The screenshot shows the homepage of the Our World in Data website. The header is dark blue with the logo 'Our World in Data' on the left. To the right of the logo are navigation links: 'Articles by topic', a search bar, and 'Latest', 'About', and 'Donate'. Further right are logos for 'OXFORD MARTIN SCHOOL', 'UNIVERSITY OF OXFORD', and 'GC DL'. Below the header is a yellow banner with the text 'Coronavirus pandemic: daily updated research and data.' and a 'Read more' button. The main content area has a dark blue background with a world map pattern. It features the text 'Research and data to make progress against the world's largest problems' in yellow. Below this is a yellow button that says 'Scroll to all our articles' with a dropdown arrow. Underneath, it states '3149 charts across 297 topics' and 'All free: open access and open source'. The next section is white and titled 'TRUSTED IN RESEARCH AND MEDIA', listing logos for Science, nature, PNAS, ROYAL STATISTICAL SOCIETY, BBC, The New York Times, CNN, FT, theguardian, THE WALL STREET JOURNAL, CNBC, The Washington Post, and Vox. Below that is 'USED IN TEACHING' with logos for HARVARD UNIVERSITY, Stanford, Berkeley, UNIVERSITY OF CAMBRIDGE, OXFORD, and MIT. The final section is light blue and titled 'Latest publications', showing a grid of article thumbnails. One prominent thumbnail is for 'nature human behaviour'.

COVID Data Hub


- CHART (Time Series Data)
- MAP (Spatial Data)
- TABLE
- SOURCES
- DOWNLOAD
- SHARE
- <https://ourworldindata.org/coronavirus>

Statistics and Research






Coronavirus Pandemic (COVID-19)

Research and data: Hannah Ritchie, Esteban Ortiz-Ospina, Diana Beltekian, Edouard Mathieu, Joe Hasell, Bobbie Macdonald, Charlie Giattino, Cameron Appel, Lucas Rodés-Guirao and Max Roser

Web development: Ernst van Woerden, Daniel Gavrilov, Matthieu Bergel, Jason Crawford, and Marcel Gerber

 The data on the coronavirus pandemic is updated daily. Last update: 7 hours ago.  [Reuse our work freely](#)  [Cite this research](#)

Coronavirus > [By country](#) [Data explorer](#) [Deaths](#) [Cases](#) [Tests](#) [Hospitalizations](#) [Vaccinations](#) [Mortality risk](#) [Excess mortality](#) [Policy responses](#) [Exemplars](#) [All charts](#)

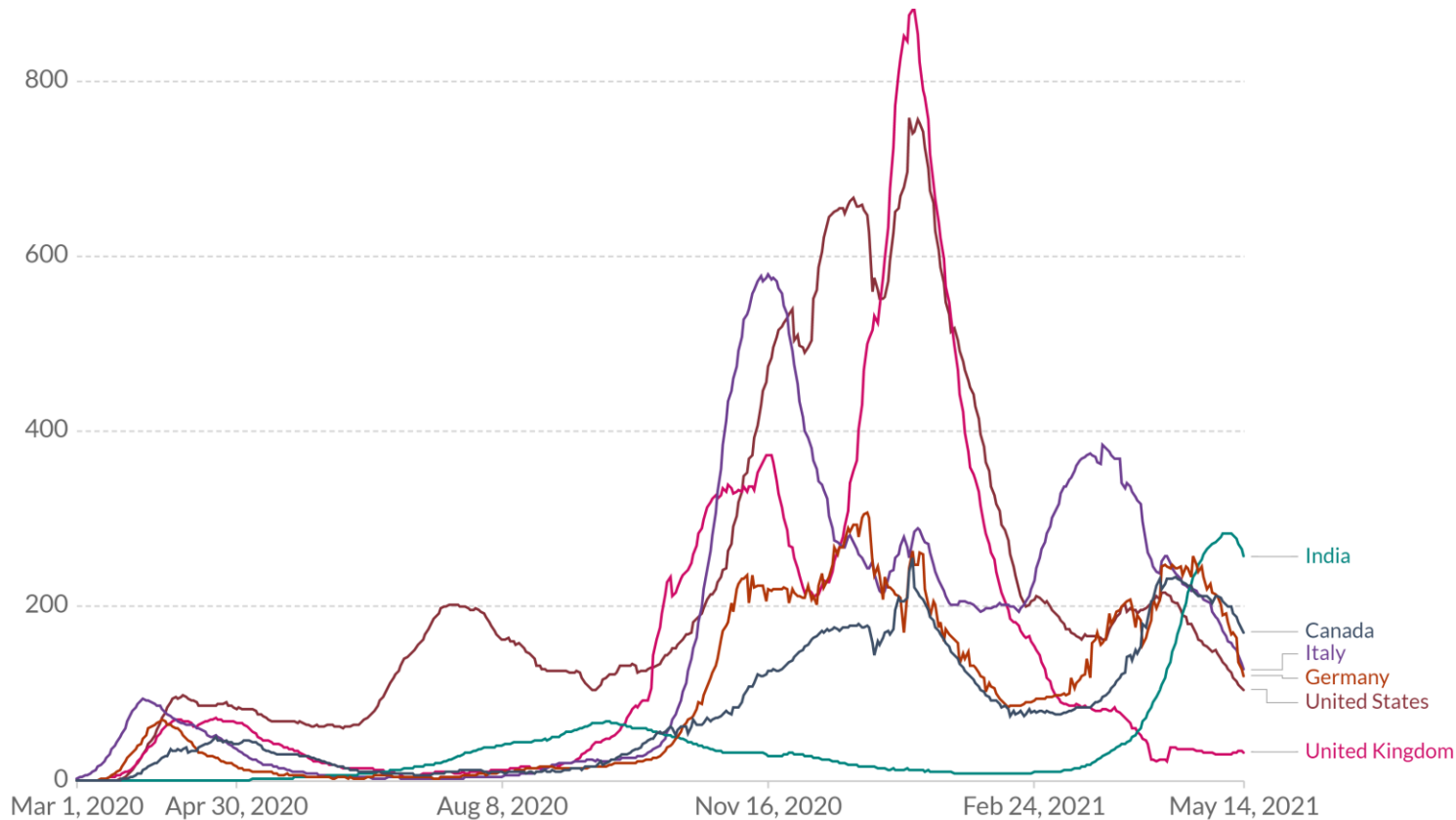
 <p>Data Explorer</p> <p>Explore all metrics - including cases, deaths, testing, and vaccinations - in one place.</p>	 <p>Country Profiles</p> <p>Get an overview of the pandemic for any country on a single page.</p>	 <p>Download Dataset</p> <p>Download our complete dataset of COVID-19 metrics on GitHub. It's open access and free for anyone to use.</p>	 <p>Vaccinations</p> <p>Explore our global dataset on COVID-19 vaccinations.</p>
 <p>US Vaccinations</p>	 <p>Cases</p>	 <p>Deaths</p>	 <p>Testing</p>

Panel Data

No. of Daily New
Confirmed Cases of
Different Countries at
Different Time

Daily new confirmed COVID-19 cases per million people

Shown is the rolling 7-day average. The number of confirmed cases is lower than the number of actual cases; the main reason for that is limited testing.



Source: Johns Hopkins University CSSE COVID-19 Data

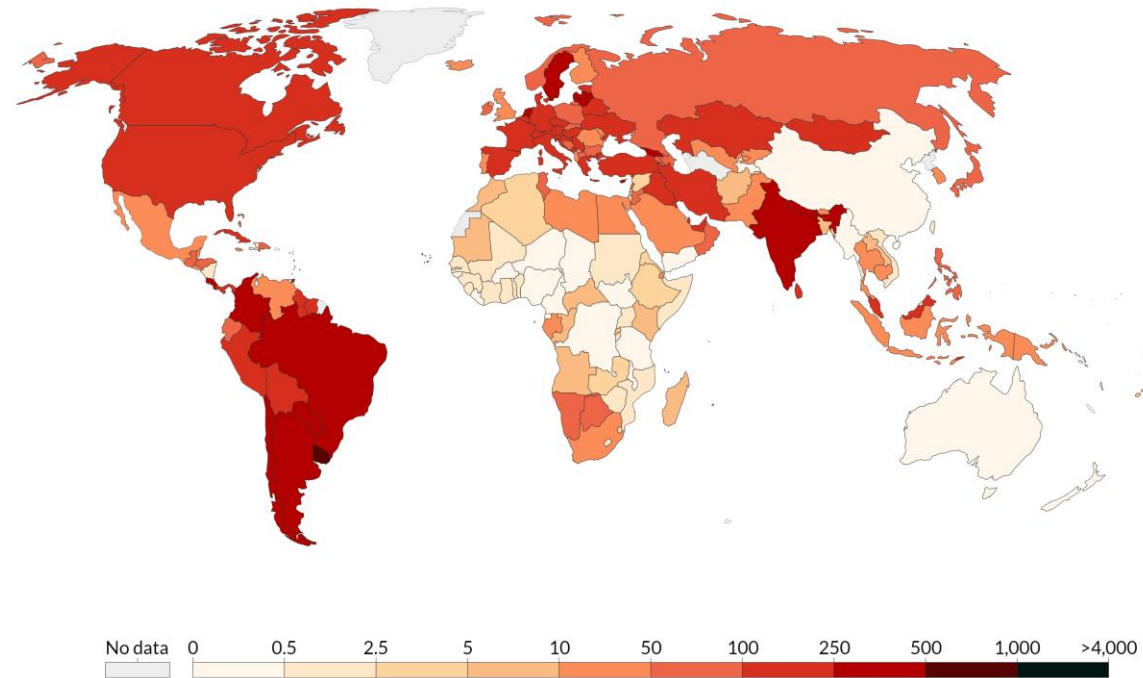
CC BY

Spatial Data by Coloured Map (a Snapshot)

- GIS shaded file
- Multiple maps can show changes over time

Daily new confirmed COVID-19 cases per million people, May 14, 2021
Shown is the rolling 7-day average. The number of confirmed cases is lower than the number of actual cases; the main reason for that is limited testing.

Our World
in Data



Source: Johns Hopkins University CSSE COVID-19 Data

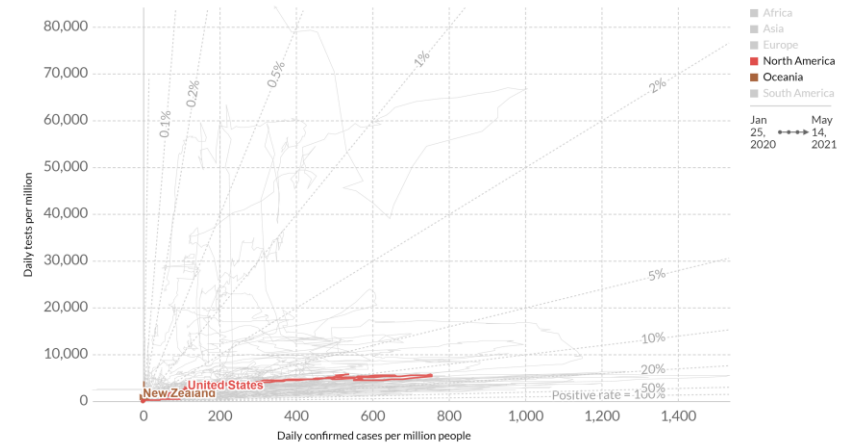
CC BY

Exercise - Choosing Metrics

- Absolute versus 7-day average, which one? why?
- Absolute number of cases versus Number of cases per 1 million population, which one? Why?
- Linear versus logarithm scale, which one? Why?
- New cases versus Cumulative cases, which one? Why?
- Number of deaths or Number of cases, which one? Why?
- Case fatality rate (CFR) versus Infection fatality rate? (See next slide)

COVID-19: Daily tests vs. Daily new confirmed cases per million

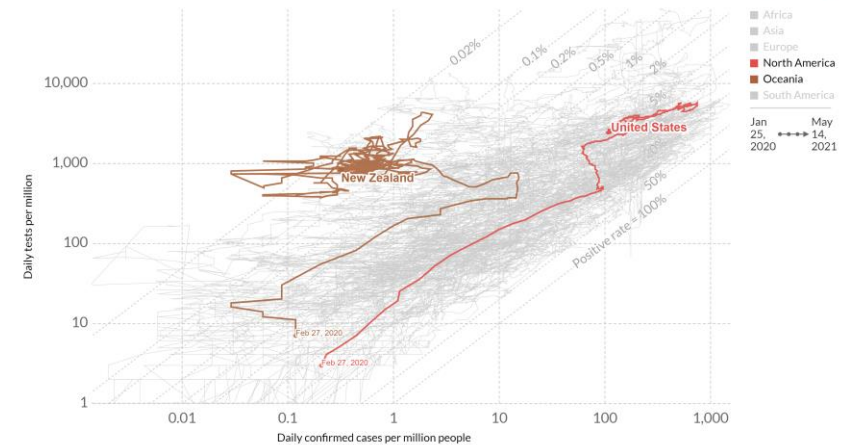
The figures are given as a rolling 7-day average.



Source: Testing data from official sources collated by Our World in Data, confirmed cases from Johns Hopkins University CSSE
Note: Comparisons of testing data across countries are affected by differences in the way the data are reported. Daily data is interpolated for countries not reporting testing data on a daily basis. Details can be found at our Testing Dataset page.
OurWorldInData.org/coronavirus • CC BY

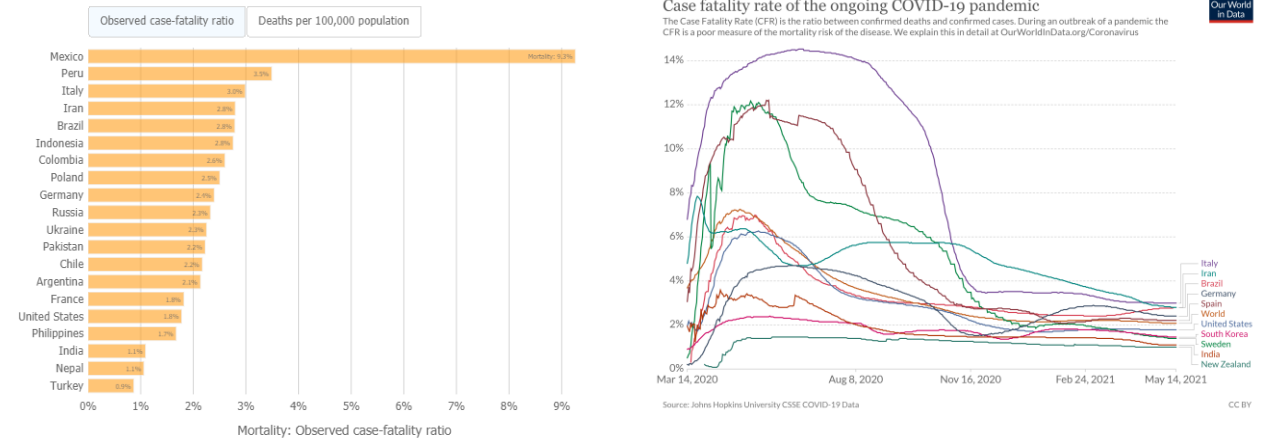
COVID-19: Daily tests vs. Daily new confirmed cases per million

The figures are given as a rolling 7-day average.

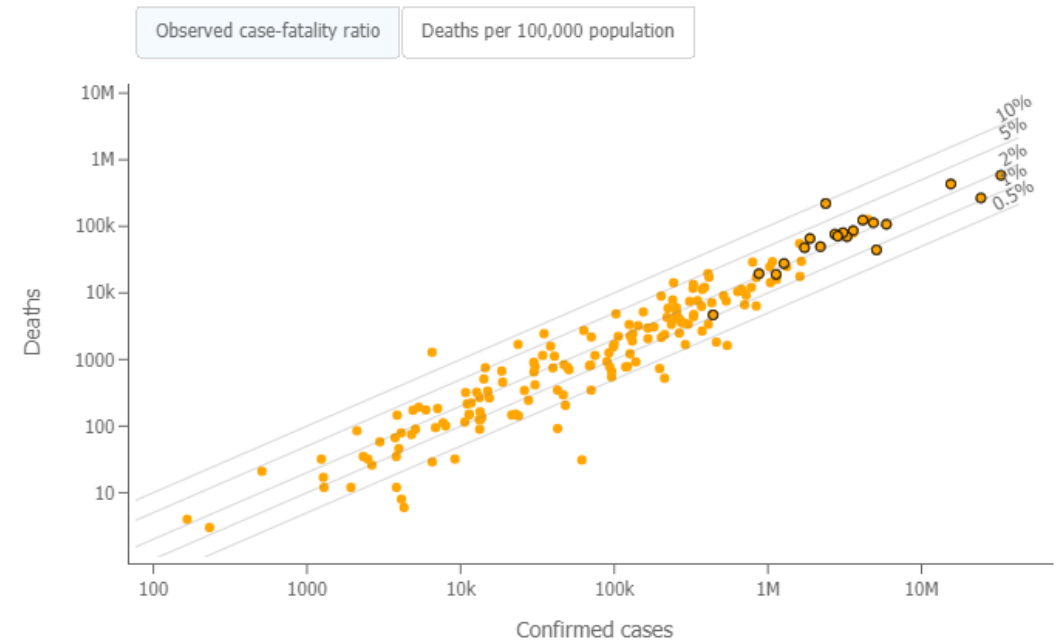


Source: Testing data from official sources collated by Our World in Data, confirmed cases from Johns Hopkins University CSSE
Note: Comparisons of testing data across countries are affected by differences in the way the data are reported. Daily data is interpolated for countries not reporting testing data on a daily basis. Details can be found at our Testing Dataset page.
OurWorldInData.org/coronavirus • CC BY

Bar chart, Line graph or Scatterplot?



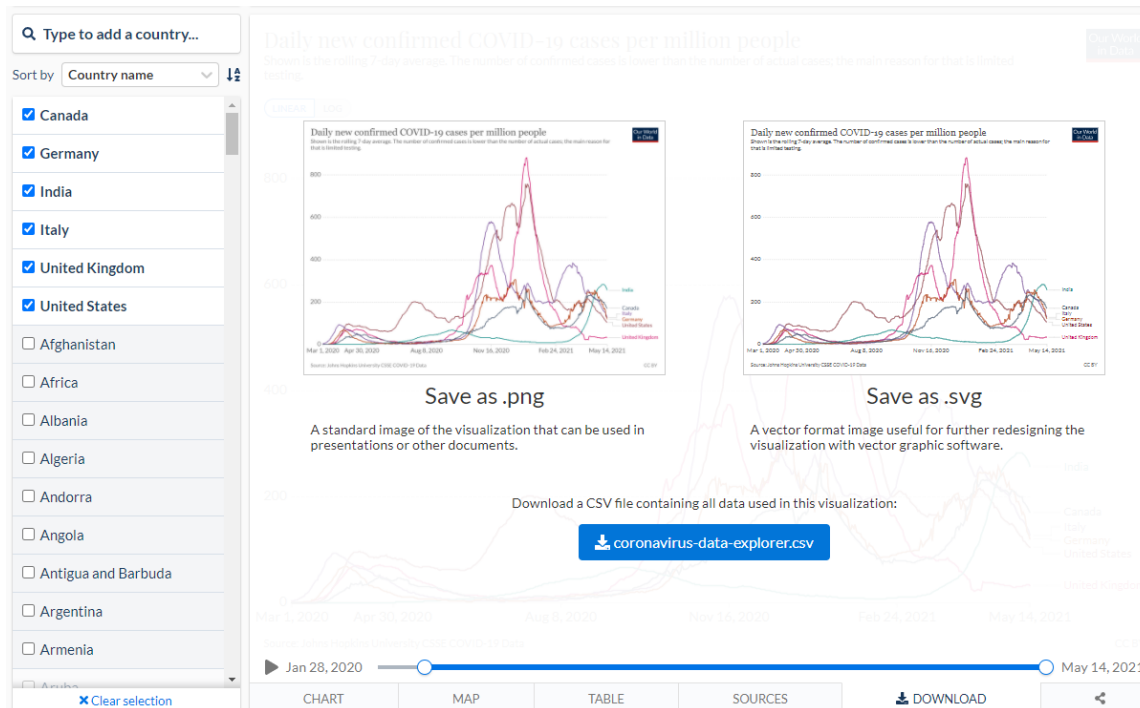
- See <https://ourworldindata.org/covid-mortality-risk>
- Line graph from OWiD, Bar chart and Scatterplot from John Hopkins Coronavirus Resource Center at <https://coronavirus.jhu.edu/data/mortality>
- Bar chart – a snapshot compare countries
- Line graph – multiple time series
- Scatterplot – a snapshot showing relationship



How to process panel data series?

Pivot Table

- Coronavirus-data-explorer.csv
- Panel Data – sequence by country and by date
- How to plot multiple time series of various countries by a line chart?



iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed
AFG	Asia	Afghanistan	24/02/2020	1	1	
AFG	Asia	Afghanistan	25/02/2020	1	0	
AFG	Asia	Afghanistan	26/02/2020	1	0	
AFG	Asia	Afghanistan	27/02/2020	1	0	
AFG	Asia	Afghanistan	28/02/2020	1	0	
AFG	Asia	Afghanistan	29/02/2020	1	0	0.143
AFG	Asia	Afghanistan	1/03/2020	1	0	0.143
AFG	Asia	Afghanistan	2/03/2020	1	0	0
AFG	Asia	Afghanistan	3/03/2020	2	1	0.143

Exercise – Plot a panel data line chart

- Convert the panel data series into a pivot table in excel (Excel can open a csv file, but remember to save it in a xls file)

Sum of new_cases_smoothed	Column Labels				
Row Labels	Asia	Europe	North America	Oceania	Grand Total
2020	19646660	23091771	22357719	31346	65127497
Jan	3865	2	3	1	3872
Feb	75184	670	23	11	75888
Mar	75579	361286	135564	3975	576404
Apr	327874	926600	947474	4227	2206176
May	572693	621873	867357	474	2062396
Jun	1110513	459221	994527	566	2564827
Jul	1887439	445520	2203358	8200	4544516
Aug	2754494	729149	1876146	10267	5370056
Sep	3513003	1361035	1500852	1719	6376609
Oct	3124531	4264924	2155045	691	9545192
Nov	2965360	7612451	4727698	496	15306005
Dec	3236125	6309040	6949672	719	16495556
2021	24340753	22862684	15782381	13954	62999772
Jan	2417965	6760267	7215427	669	16394328
Feb	1916784	4038655	3114828	571	9070838
Mar	3205540	5345759	2219908	4347	10775554
Apr	10009167	5204151	2401390	6377	17621085
May	6791297	1513850	830829	1990	9137966
Grand Total	43987414	45954455	38140100	45301	128127269

