**Data Brief**

The data is made from merged datasets. This gives rise to a range of possibilities for analysis but requires us to be careful:

* A given country may not have data available for every field.
* The level of granularity of the COVID 19 data is daily whereas the rest of the data is for the most recent year of survey as a whole. This means that each row of data corresponds to a specific date, but because the non-COVID data doesn't change with the date, values are repeated across the rows for each country. Therefore make sure you don't aggregate values by summing them, instead take averages.

You are welcome to include other data sources to supplement this dataset, should you wish to explore an avenue of analysis further.

**Field Taxonomy**

|  |  |  |
| --- | --- | --- |
| **Field Name** | **Units** | **Description** |
| Date |  |  |
| Continent |  |  |
| Countries |  | *A set of countries named according to UN Convention* |
| New\_Cases | # | *A daily estimate of COVID-19 Cases per Country.* |
| New\_Deaths | # | *A daily estimate of COVID-19 Deaths per Country.* |
| Total\_Cases | # | *A running total of COVID-19 Cases per Country.* |
| Total\_Deaths | # | *A running total of COVID-19 Deaths per Country.* |
| Weekly\_Cases | # | *NB: This may show up as negative where a Country has made an adjustment of their Total Cases estimate.* |
| Weekly\_Deaths | # | *NB: This may show up as negative where a Country has made an adjustment of their Total Deaths estimate.* |
| Population\_Density | # | *The average number of people populating each square kilometre of land of given country as of 2020 estimates.* |
| Fertility\_Rate | # | *The number of children that would be born to each woman with prevailing age-specific fertility rates as of 2020 estimates.* |
| GDP\_PPP | PPP $ | *PPP GDP is GDP converted to international dollars using purchasing power parity rates. A PPP dollar has the same purchasing power over GDP as the US dollar has in the united states data are in constant 2005 international dollars. Calculated as of 2020 estimates.* |
| Corruption | In units of standard normal distribution (Roughly -2.5 to 2.5, where a strongly negative score corresponds to relatively higher levels of 'Corruption', and a strongly positive score corresponds to relatively lower levels of 'Corruption') | *Control of Corruption captures perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests. Calculated as of 2018 estimates.* |
| Government\_Effectiveness | In units of standard normal distribution (Roughly -2.5 to 2.5, where a strongly negative score corresponds to relatively lower 'Government Effectiveness', and a strongly positive score corresponds to relatively higher 'Government Effectiveness') | *Government Effectiveness captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. Calculated as of 2018 estimates.* |
| Political\_Stability | In units of standard normal distribution (Roughly -2.5 to 2.5, where a strongly negative score corresponds to relatively lower 'Political Stability', and a strongly positive score corresponds to relatively higher 'Political Stability') | *Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism. Calculated as of 2018 estimates.* |
| Rule\_of\_Law | In units of standard normal distribution (Roughly -2.5 to 2.5, where a strongly negative score corresponds to relatively lower 'Rule of Law', and a strongly positive score corresponds to relatively higher 'Rule of Law') | *Rule of Law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Calculated as of 2018 estimates.* |
| Government\_Healthcare\_Spend | % | *Domestic General Government Health Expenditure (GGHE-D) as % Current Health Expenditure (CHE) as of 2017 estimates.* |
| Urban\_Population | % | *Population in urban agglomerations of more than one million is the % of a country's population living in metropolitan areas that in 2000 had a population of more than one million people. Calculated as of 2019 estimates.* |
| Population | # | *Estimate of the countries' populations as of 2020.* |
| Smoking\_Prevalence | % | *Prevalence of smoking is the percentage of men and women ages 15 and over who currently smoke any tobacco product on a daily or non-daily basis. It excludes smokeless tobacco use. The rates are age-standardised. Calculated as of 2019 estimates.* |
| Tourism | # | *International inbound tourists (overnight visitors) are the number of tourists who travel to a country other than that in which they have their usual residence, but outside their usual environment, for a period not exceeding 12 months and whose main purpose in visiting is other than an activity remunerated from within the country visited. Calculated as of 2018 estimates.* |
| Women\_In\_Parliament | % | *The percentage of parliamentary seats in a single or lower chamber held by women as of 2019 estimates.* |
| Obesity\_Rate | % | *Adult prevalence rate gives the percent of a country's population considered to be obese. Calculated as of 2020 estimates.* |
| RunTot\_New\_Cases | # | *Running total of new coronavirus cases per country.* |
| RunTot\_New\_Deaths | # | *Running total of new coronavirus deaths per country.* |
| Diabetes Prevalence % of Population 20-79 | % | *The proportion of the population aged 20-79 that has been diagnosed with a form of Diabetes. As of 2019 estiamtes.* |
| Labour Force Total | # | *Size of the labour force of a country as of 2019 estimates.* |
| Population Aged 65 and above % of Total | % | *The proportion of a country's population aged 65 and above as a percentage of the total population (as of 2019 estimates).* |
| Urban Population | # | *The number of people living in urban areas of a country as of 2019 estimates.* |
| Death Rate Per 1000 (2017) | # | *Crude death rate indicates the number of deaths occurring during the year, per 1,000 population estimated at midyear. Subtracting the crude death rate from the crude birth rate provides the rate of natural increase, which is equal to the rate of population change in the absence of migration. As of 2017 estimates.* |
| PM2.5 air pollution, mean annual exposure (micrograms per cubic meter) (2017) | # | *Population-weighted exposure to ambient PM2.5 pollution is defined as the average level of exposure of a nation's population to concentrations of suspended particles measuring less than 2.5 microns in aerodynamic diameter, which are capable of penetrating deep into the respiratory tract and causing severe health damage. Exposure is calculated by weighting mean annual concentrations of PM2.5 by population in both urban and rural areas. As of 2017 estimates.* |
| PM2.5 air pollution, population exposed to levels exceeding WHO guideline value (% of total) (2017) | % | *Percent of population exposed to ambient concentrations of PM2.5 that exceed the WHO guideline value is defined as the portion of a country’s population living in places where mean annual concentrations of PM2.5 are greater than 10 micrograms per cubic meter, the guideline value recommended by the World Health Organization as the lower end of the range of concentrations over which adverse health effects due to PM2.5 exposure have been observed. As of 2017 estimates.* |
| Hospital Beds Year |  | *The year in which the hospital beds data is taken from. This is due to the sparsity in the data.* |
| Hospital Beds Per 1000 People | # | *The number of hospital beds available for a country per 1000 members of its population. As of 2019 estimates.* |

**Data Sources**

This data has been made by merging datasets from the following sources:

**COVID Data :** <https://ourworldindata.org/coronavirus-source-data> - Up to date resource on cases and fatalities

**Gapminder - Health and Economic Data :** <https://www.gapminder.org/data/> - Data on Population, GDP PPP, Fertility Rate and Urban Population was sourced here

**WHO - Health Spend Data :** <https://apps.who.int/nha/database/ViewData/Indicators/en> - Government Healthcare Spend

**World Bank - Political and Health-Related Data :**

* <https://data.worldbank.org/indicator/ST.INT.ARVL> - Tourism Arrivals
* <https://data.worldbank.org/indicator/SG.GEN.PARL.ZS?view=chart> - Women in Parliament
* [https://databank.worldbank.org/source/worldwide-governance-indicators#](https://databank.worldbank.org/source/worldwide-governance-indicators) - Government Effectiveness
* <https://data.worldbank.org/indicator/SH.PRV.SMOK> - Smoking Prevalence
* <https://databank.worldbank.org/source/world-development-indicators> - Diabetes Prevelance % of Population 20-79, Labour Force Total, Population Aged 65 and above % of Total, Urban Population - *(as a total rather than percentage),* Death Rate Per 1000 (2017), PM2.5 air pollution, mean annual exposure (micrograms per cubic meter) (2017), PM2.5 air pollution, population exposed to levels exceeding WHO guideline value (% of total) (2017), Hospital Beds Year, Hospital Beds Per 1000 People

**CIA - Obesity Data :** <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2228rank.html> - Obesity Rate

*Note - The country names used are purely based on the datasets availability and do not reflect any political views or affiliations of TrueCue.*

**Data Exploration ideas**

You are free to explore and analyse the data however you like but in case you get stuck or need a starting point, here are some ideas.

*How does Government Policy affect the number of COVID 19 fatalities?*

* *Did countries with higher public spending on healthcare have fewer fatalities?*
* *Do factors like corruption and political stability affect a country’s reaction to COVID 19?*

*Which countries are worst affected by COVID 19?*

* *Is the number of deaths proportional to population density in these countries?*
* *Are countries with high urban populations more severely affected?*
* *Are there any anomalies?*