
ESCUELA DE GOBIERNO

DOCUMENTOS DE TRABAJO 2020/09

COVID-19 DATABASE

EDUARDO LEVY YEYATI

LUCA SARTORIO

JULIO 2020

Documentos de trabajo: <https://bit.ly/2REorES>

UTDT: Av. Figueroa Alcorta 7350, C1428BCW Buenos Aires, Argentina

COVID-19 database

Eduardo Levy Yeyati & Luca Sartorio¹

This update: 07/24/2020

The COVID-19 database is a statistical package developed by the Center for the Evaluation of Evidence-Based Policies (CEPE). This dataset collects more than 50 indicators in more than 150 countries to consolidate them into a public database available for download that compiles a multiplicity of variables related to the health targets, the policies implemented and the contextual characteristics of each country that significantly affect the impact of the pandemic.

The goal of this dataset is to consolidate in a single, standardized and usable tool, a set of indicators that capture the multidimensionality of factors that influence the capacities of national states in dealing with the COVID-19 and affect the effectiveness of containment measures or the testing, tracing and isolating policies. This includes a wide set of variables that seek to take into account the incidence in the pandemic of social, economic, demographic, political and climatic characteristics (among others).

The base is structured at the country level and on a daily frequency. A large number of variables in the database are daily time series and track the health outcomes of the epidemic, the stringency of containment measures and mobility trends, variations in public policies associated with COVID-19, and climatic variables such as temperature or humidity. These variables will be updated weekly in updated versions of the dataset that will be available on the CEPE website. On the other hand, other variables are considered “fixed effects” during the course of the epidemic, such as demographic, economic, social, political and cultural characteristics of each country that have a relevant incidence in the main health outcomes or in the effectiveness of social distancing policies.

This document describes the detailed content of the database, its primary sources from which the information compiled was collected, and a codebook for its practical use. The document will be updated as new variables are added to the database.

1. Topics and primary sources

The COVID-19 database is structured by the 9 following topics:

- COVID-19 Data
- Mobility and Stringency
- Government Response
- Health System
- Demographics
- Economic and Social

¹ Eduardo Levy Yeyati and Luca Sartorio are with the School of Government of Universidad Torcuato Di Tella. They thank the invaluable assistance of Augusto Castro in the preparation of this document.

- Governance and State Capability
- Globalization and International Integration
- Temperature and Humidity

In this section we describe the rationale for the inclusion of each topic, its main components and their primary sources.

1.1. COVID-19 Data

The first set of variables are directly related to the fundamental health outcomes to analyze the effectiveness of the policy responses of the different countries. We use COVID-19 cases and deaths data from Our World in Data, both new daily cases and those accumulated up to each day. In addition, we also obtained from this dataset the daily tests for those countries that report them, both the new ones carried out each day, the accumulated ones and also a daily indicator of tests for each positive case. Finally, we consider three different estimates of the daily effective reproduction number (R_t) to characterize the COVID-19 spread rate, defined as the expected number of secondary infectious cases produced by a primary infectious case: the first one is developed by the Harvard Chan School of Public Health, the second is estimated by Arroyo et al. (2020)² and the last one by the Centre for the Mathematical Modelling of Infectious Diseases³.

1.2. Mobility and Stringency

Containment measures and social distancing policies are maybe the main policy intervention used by national governments to limit the spread of the virus. First, we report the Stringency Index estimated by the Oxford COVID-19 Government Response Tracker, a composite index with daily frequency that considers several response indicators like school closures, workplace closures, restrictions on gatherings, travel bans, among others to generate a measure of the “de jure” rigidity of containment measures. Moreover, to assess the “de facto” compliance of the social distancing policies, we add to the database the daily mobility data garnered from the movement of cell phones estimated in the Google COVID-19 Community Mobility Reports, considering its multiple variants (workplaces mobility, residential mobility, grocery and pharmacy mobility, etc.).

1.3. Government Response

In addition to social distancing policies, governments deploy multiple complementary interventions that can be tracked on a daily basis like:

- The “de jure” testing policy, namely, government policy on the scope to who has access to testing
- The “de jure” contact tracing policy, namely, government policy on the scope of contact tracing after a positive diagnosis
- Announced short term spending on healthcare system, e.g. hospitals, masks, etc.
- The presence of public information campaigns

² Arroyo Marioli, F., Bullano, F., Kučinskas, S., & Rondón-Moreno, C. (2020). Tracking R_t of COVID-19: A New Real-Time Estimation Using the Kalman Filter. *Available at SSRN 3581633*.

³ Abbott, S., Hellewell, J., Thompson, R. N., Sherratt, K., Gibbs, H. P., Bosse, N. I., ... & Chan, Y. W. D. (2020). Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Research*, 5(112), 112.

- Announced fiscal measures and economic stimulus spending

All these measures are globally tracked on a daily basis by the Oxford COVID-19 Government Response Tracker.

1.4. Health System

As the effectiveness of the policy response is strongly affected by the capacity of the health systems of each different country we incorporate some variables to measure these cross-country differences like the Healthcare Access and Quality Index Based on Amenable Mortality developed by the Institute for Health Metrics and Evaluation (IHME), the Global Health Security Index developed jointly by the Nuclear Threat Initiative (NTI), the Johns Hopkins Center for Health Security (JHU) and The Economist Intelligence Unit (EIU) and the number of hospital beds per every 1.000 inhabitants provided by the World Health Organization⁴. All these metrics are treated as fixed effects during the pandemic and they not vary on a daily basis.

1.5. Demographics

Two demographics aspects have a significant relevance in the impact of the pandemic. First, countries more densely populated may be exposed to a more accelerated spread of the virus limiting their ability to control the impact on main health outcomes. Second, aged populations have higher mortality rates, especially countries with high proportions of individuals older than 65 years. For these reasons, we add several demographic variables reported by the World Bank like the population density, the share of individuals older than 65 or 70 years, and the median age of the population. All these metrics are treated as fixed effects during the pandemic and they not vary on a daily basis.

1.6. Economic and social

As the levels of per capita income, labor informality and economic inequality of the countries strongly limit their ability to maintain economic lockdowns and stringent social distancing over time, it seems important to us to incorporate to the analysis indicators published by the World Bank such as the GDP per capita or the Gini Index as well as the informal employment share provided by the International Labour Organization.

Furthermore, given that poor habitat conditions also undermine the capacity to maintain strict social distancing and “stay at home” policies, we believe it is important to consider the share of the population living in slums estimated by the United Nations. Finally, the lack of basic infrastructure for handwashing at home can affect some fundamental prevention measures, which is why we also incorporate the share of the national population that has handwashing facilities at home, jointly estimated by the World Health Organization and UNICEF. All these metrics are treated as fixed effects during the pandemic and they not vary on a daily basis.

1.7. Governance and State Capability

The effectiveness of the policy response to the pandemic and the capacity to sustain stringent social distancing policies are naturally constrained by the national state capabilities and the political context of each country. In order to consider the impact of these factors, we consider the

⁴ This variable does not consider the additional beds that were incorporated especially for the COVID-19 epidemic, nor does it differentiate between standard beds and intensive care beds.

Government effectiveness, Political Stability and Absence of Violence/Terrorism indexes developed by the World Bank and the Democracy Index estimated by The Economist. All these metrics are treated as fixed effects during the pandemic and they not vary on a daily basis.

1.8. Globalization and International Integration

With the exception of China, the first cases of COVID-19 in each country were imported by international travelers. So, it is plausible to consider that more globalized countries or those with greater tourist traffic were more exposed to having stronger initial outbreaks, conditioning their response capacity and the time available to prepare the health system and the state capabilities required to deal with the pandemic. To empirically analyze this hypothetical constraint, we incorporate tourist traffic measures reported by the World Bank as well as the Globalization Index developed by the KOF Swiss Economic Institute. All these metrics are treated as fixed effects during the pandemic and they not vary on a daily basis.

1.9. Temperature and humidity

Finally, climatic factors can influence both the spread of the virus and the use of available sanitary infrastructure due to seasonal outbreaks of other diseases, and for this reason we report daily temperature and humidity levels for each country in the capital or the most populous city available in the Air Quality Open Data Platform.

2. Variable definitions and codebook

Topic	Variable	Description	Coding	Source
ID Variables	country	Country name	Country name	-
ID Variables	country_code	Country ISO 3 code	Country code	-
ID Variables	region	Country region	Country region	-
ID Variables	region_2	Country region (alternative version)	Country region (alternative version)	-
ID Variables	date	Date	Date since 01-01-2020	-
ID Variables	date_code	Unique ID. The concatenation of country ISO 3 code and date as number with origin in January 1, 2020	Country_code concatenated with date as number (origin: January 1, 2020)	-
COVID data	total_cases	Total confirmed cases of COVID-19	Total cases in units 0 - no cases until that day NA - no data	Our World in Data
COVID data	new_cases	New confirmed cases of COVID-19	New cases in units 0 - no cases that day NA - no data	Our World in Data
COVID data	total_deaths	Total deaths attributed to COVID-19	Total deaths in units 0 - no cases until that day NA - no data	Our World in Data
COVID data	new_deaths	New deaths attributed to COVID-19	New deaths in units 0 - no deaths that day NA - no data	Our World in Data
COVID data	total_tests	Total tests for COVID-19	Total deaths in units 0 - no tests until that day NA - no data	Our World in Data
COVID data	new_tests	New tests for COVID-19	New tests in units 0 - no tests that day NA - no data	Our World in Data
COVID data	tests_units	Units used to report its testing data	Tests units NA - no data	Our World in Data

COVID data	tests_per_case	Number of test per detected case of COVID-19	Tests per confirmed case NA - no data	Our World in Data
COVID data	r_t	R_t is the daily effective reproduction number (Rt) to characterize COVID-19 spread rate, defined as the expected number of secondary infectious cases produced by a primary infectious case.	The effective reproduction rate -88 - the number of total cases is above 50 but the average number of new cases within the previous 7 days is below 10 -888 - number of total cases below 50 NA - no data	Harvard Chan School of Public Health
COVID data	rt_arroyo	R_t is the daily effective reproduction number (Rt) to characterize COVID-19 spread rate, defined as the expected number of secondary infectious cases produced by a primary infectious case. This estimation provides a “smoothed” version of the growth rate of the number of infected individuals. We choose estimation for seven days of infectious realized by Arroyo et al. (2020).	The effective reproduction rate NA - no data	Arroyo et al (2020)
COVID data	rt_epi	R_t is the daily effective reproduction number (Rt) to characterize COVID-19 spread rate, defined as the expected number of secondary infectious cases produced by a primary infectious case. This estimation accounts for reporting delays, so that temporal variations in reproduction number estimates can be compared directly with the times at which interventions are implemented. We choose the nowcast estimation realized by Epi Forecasts.	The effective reproduction rate NA - no data	EpiForecasts

Mobility and Stringency	retail_and_recreation_percent_change_from_baseline	Mobility trends for places like restaurants, cafes, shopping centers, theme parks, museums, libraries, and movie theaters, compared to baseline. These reports show how visits and length of stay at different places change compared to a baseline. These changes are calculated using the same kind of aggregated and anonymized data used to show popular times for places in Google Maps. Changes for each day are compared to a baseline value for that day of the week and the baseline is the median value, for the corresponding day of the week, during the 5- week period Jan 3–Feb 6, 2020.	Percentage change from the median value, for the corresponding day of the week, during the 5-week period Jan 3–Feb 6, 2020. NA - no data	Google Mobility
Mobility and Stringency	grocery_and_pharmacy_percent_change_from_baseline	Mobility trends for places like grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies, compared to baseline.	Same as above.	Google Mobility
Mobility and Stringency	parks_percent_change_from_baseline	Mobility trends for places like national parks, public beaches, marinas, dog parks, plazas, and public gardens.	Same as above.	Google Mobility
Mobility and Stringency	transit_stations_percent_change_from_baseline	Mobility trends for places like like public transport hubs such as subway, bus, and train stations, compared to baseline.	Same as above.	Google Mobility
Mobility and Stringency	workplaces_percent_change_from_baseline	Mobility trends for places of work compared to baseline.	Same as above.	Google Mobility
Mobility and Stringency	residential_percent_change_from_baseline	Mobility trends for places for places of residence compared to baseline.	Same as above.	Google Mobility
Mobility and Stringency	stringency_index	The Government Response Stringency Index is a composite measure based on nine response indicators including school closures, workplace closures, and travel bans, rescaled to a value from 0 to 100 (100 = strictest response).	Stringency of containment measures (between 0 and 100) NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	stringency_legacy_index	The Government Response Stringency Index is a composite measure based on seven response indicators including school closures, workplace closures, and travel bans, rescaled to a value from 0 to 100 (100 = strictest response). It is generally not	Stringency of containment measures with alternative index (between 0 and 100) NA - no data	University of Oxford, Blavatnik School of Government

		recommended to use this legacy index, but it may be useful for continuity purposes.		
Mobility and Stringency	c1_school_closing	Record closings of schools and universities	0 - no measures 1 - recommend closing 2 - require closing (only some levels or categories, eg just high school, or just public schools) 3 - require closing all levels NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c1_flag	Binary flag for geographic scope. It indicates if the policy is general or geographically targeted.	0 - targeted 1 - general NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c2_workplace_closing	Record closings of workplaces	0 - no measures 1 - recommend closing (or recommend work from home) 2 - require closing (or work from home) for some sectors or categories of workers 3 - require closing (or work from home) for all-but-essential workplaces	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c2_flag	Binary flag for geographic scope. It indicates if the policy is general or geographically targeted.	0 - targeted 1 - general NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c3_cancel_public_events	Record cancelling public events	0 - no measures 1 - recommend cancelling 2 - require cancelling NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c3_flag	Binary flag for geographic scope. It indicates if the policy is general or geographically targeted.	0 - targeted 1 - general NA - no data	University of Oxford, Blavatnik School of Government

Mobility and Stringency	c4_restrictions_on_gatherings	Record limits on private gatherings	0 - no restrictions 1 - restrictions on very large gatherings (the limit is above 1000 people) 2 - restrictions on gatherings between 101-1000 people 3 - restrictions on gatherings between 11-100 people 4 - restrictions on gatherings of 10 people or less NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c4_flag	Binary flag for geographic scope. It indicates if the policy is general or geographically targeted.	0 - targeted 1 - general NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c5_close_public_transport	Record closing of public transport	0 - no measures 1 - recommend closing (or significantly reduce volume/route/means of transport available) 2 - require closing (or prohibit most citizens from using it) NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c5_flag	Binary flag for geographic scope. It indicates if the policy is general or geographically targeted.	0 - targeted 1 - general NA - no data	University of Oxford, Blavatnik School of Government

Mobility and Stringency	c6_stay_at_home_requirements	Record orders to "shelter-in-place" and otherwise confine to the home	0 - no measures 1 - recommend not leaving house 2 - require not leaving house with exceptions for daily exercise, grocery shopping, and 'essential' trips 3 - require not leaving house with minimal exceptions (eg allowed to leave once a week, or only one person can leave at a time, etc) NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c6_flag	Binary flag for geographic scope. It indicates if the policy is general or geographically targeted.	0 - targeted 1 - general NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c7_restrictions_on_internal_movement	Record restrictions on internal movement between cities/regions	0 - no measures 1 - recommend not to travel between regions/cities 2 - internal movement restrictions in place NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c7_flag	Binary flag for geographic scope. It indicates if the policy is general or geographically targeted.	0 - targeted 1 - general NA - no data	University of Oxford, Blavatnik School of Government
Mobility and Stringency	c8_international_travel_controls	Record restrictions on internal movement between cities/regions	0 - no restrictions 1 - screening arrivals 2 - quarantine arrivals from some or all regions 3 - ban arrivals from some regions 4 - ban on all regions or total border closure NA - no data	University of Oxford, Blavatnik School of Government

Government Response	fiscal_measures	Announced economic stimulus spending	Monetary value in USD of fiscal stimuli, includes any spending or tax cuts 0 - no new spending that day NA - no data	University of Oxford, Blavatnik School of Government
Government Response	h2_testing_policy	Record government policy on who has access to testing. This records policies about testing for current infection (PCR tests) not testing for immunity (antibody test).	0 - no testing policy 1 - only those who both (a) have symptoms and (b) meet specific criteria (e.g. key workers, admitted to hospital, came into contact with a known case, returned from overseas) 2 - testing of anyone showing Covid-19 symptoms 3 - open public testing (e.g. "drive through" testing available to asymptomatic people) NA - no data	University of Oxford, Blavatnik School of Government
Government Response	h3_contact_tracing	Record government policy on contact tracing after a positive diagnosis.	0 - no contact tracing 1 - limited contact tracing; not done for all cases 2 - comprehensive contact tracing; done for all identified cases	University of Oxford, Blavatnik School of Government
Government Response	h4_emergency_investment_in_healthcare	Announced short term spending on healthcare system, eg hospitals, masks, etc. Only record amount additional to previously announced spending.	Monetary value in USD 0 - no new spending that day NA - no data	University of Oxford, Blavatnik School of Government

Government Response	h1_public_information_campaigns	Record presence of public info campaign	0 - no Covid-19 public information campaign 1 - public officials urging caution about Covid-19 2- coordinated public information campaign (eg across traditional and social media) NA - no data	University of Oxford, Blavatnik School of Government
Government Response	h1_flag	Binary flag for geographic scope. It indicates if the policy is general or geographically targeted.	0 - targeted 1 - general NA - no data	University of Oxford, Blavatnik School of Government
Health System	haq_index	Healthcare Access and Quality Index. These estimates were used in an analysis of national levels of personal healthcare access and quality based on 32 causes considered amenable to healthcare over time.	HAQ Index (between 0 and 100) NA - no data	Institute for Health Metrics and Evaluation (IHME)
Health System	ghs_index	To create the Global Health Security Index (GHS), NTI, JHU, and the EIU project team worked with an international advisory panel of 21 experts from 13 countries to create a detailed and comprehensive framework of 140 questions, organized across 6 categories, 34 indicators, and 85 subindicators to assess a country's capability to prevent and mitigate epidemics and pandemics. The Index has 6 main dimensions: prevention, detection and reporting, rapid response, health system, compliance with international norms and risk environment.	GHS Index NA – no data	Nuclear Threat Initiative, John Hopkins Center for Health Security, The Economist Intelligence Unit
Health System	hospital_beds_per1k	The number of hospital beds (public or private) available for every 1000 inhabitants in a population.	Number of hospital beds per every 1000 inhabitants NA - no data	World Health Organization
Demographics	total_population	Total Population	Total population in units NA - no data	World Bank

Demographics	population_density	Population density is midyear population divided by land area in square kilometers. Land area is a country's total area, excluding area under inland water bodies, national claims to continental shelf, and exclusive economic zones. In most cases the definition of inland water bodies includes major rivers and lakes.	Population density NA - no data	World Bank
Demographics	median_age	Median age of the population	Median age NA - no data	Our World in Data
Demographics	aged_65_older	Share of the population that is 65 years and older. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship.	Share of the population that is 65 years and older NA - no data	World Bank
Demographics	aged_70_older	Share of the population that is 70 years and older. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship.	Share of the population that is 70 years and older NA - no data	World Bank
Economic and Social	real_gdp_pc	GDP per capita, PPP (constant 2017 international \$)	GDP per capita, PPP (constant 2017 international \$) NA - no data	World Bank
Economic and Social	informal_employment	Share of informal employment in total employment	Share of informal employment NA - no data	International Labour Organization
Economic and Social	urban_informal_employment	Share of informal employment in total employment, only in urban locations.	Share of informal employment (urban locations) NA - no data	International Labour Organization

Economic and Social	gini	Gini index measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution. A Lorenz curve plots the cumulative percentages of total income received against the cumulative number of recipients, starting with the poorest individual or household. The Gini index measures the area between the Lorenz curve and a hypothetical line of absolute equality, expressed as a percentage of the maximum area under the line. Thus a Gini index of 0 represents perfect equality, while an index of 100 implies perfect inequality.	Gini Index NA - no data	World Bank
Economic and Social	unemployment	Unemployment refers to the share of the labor force that is without work but available for and seeking employment.	Unemployment rate NA - no data	World Bank
Economic and Social	population_living_in_slums	Population living in slums is the proportion of the urban population living in slum households. A slum household is defined as a group of individuals living under the same roof lacking one or more of the following conditions: access to improved water, access to improved sanitation, sufficient living area, and durability of housing.	Population living in slums as percentage of total population NA - no data	United Nation's Millennium Development Goals database
Economic and Social	people_with_basic_handwashing_facilities	The percentage of people living in households that have a handwashing facility with soap and water available on the premises. Handwashing facilities may be fixed or mobile and include a sink with tap water, buckets with taps, tippy-taps, and jugs or basins designated for handwashing. Soap includes bar soap, liquid soap, powder detergent, and soapy water but does not include ash, soil, sand or other handwashing agents.	Population with basic handwashing facilities as percentage of total population NA - no data	WHO/UNICEF Joint Monitoring Programme (JMP)

Governance and State Capability	political_stability	Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.	Standard deviation of political stability NA - no data	World Bank
Governance and State Capability	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. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. ranging from approximately -2.5 to 2.5.	Standard deviation of government effectiveness NA - no data	World Bank
Governance and State Capability	democracy_index	Index ranging from 1 to 10 based on the ratings of five categories of indicators: whether national elections are free and fair, the security of voters, the influence of foreign powers on government, the capability of the civil service to implement policies.	Democracy index NA - no data	The Economist

<p>Globalization and international integration</p>	<p>inbound_tourism</p>	<p>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. When data on number of tourists are not available, the number of visitors, which includes tourists, same-day visitors, cruise passengers, and crew members, is shown instead. Sources and collection methods for arrivals differ across countries. In some cases data are from border statistics (police, immigration, and the like) and supplemented by border surveys. In other cases data are from tourism accommodation establishments. For some countries number of arrivals is limited to arrivals by air and for others to arrivals staying in hotels. Some countries include arrivals of nationals residing abroad while others do not. Caution should thus be used in comparing arrivals across countries. The data on inbound tourists refer to the number of arrivals, not to the number of people traveling. Thus a person who makes several trips to a country during a given period is counted each time as a new arrival. Year: 2018</p>	<p>Inbound tourists in thousands NA - no data</p>	<p>World Bank</p>
<p>Globalization and international integration</p>	<p>outbound_tourism</p>	<p>International outbound tourists are the number of departures that people make from their country of usual residence to any other country for any purpose other than a remunerated activity in the country visited. The data on outbound tourists refer to the number of departures, not to the number of people traveling. Thus a person who makes several trips from a country during a given period is counted each time as a new departure. Year: 2018</p>	<p>Outbound tourists in thousands NA - no data</p>	<p>World Bank</p>

Globalization and international integration	globalization_index	The KOF Globalization Index measures the economic, social and political dimensions of globalization. Each variable is transformed to an index with a scale from one to one hundred, where 100 is assigned to the maximum value of a specific variable over the whole sample.	Globalization index NA - no data	KOF Swiss Economic Institute
Temperature and Humidity	humidity	The data for each city is based on the median of several stations. The humidity reports the median for the most populous city available or its capital city. The complete list of cities used for each country is available in the appendix.	Median daily value of humidity in the most populous city NA - no data	Air Quality Open Data Platform
Temperature and Humidity	temperature	The data for each city is based on the median of several stations. The temperature reports the median for the most populous city available or its capital city. The complete list of cities used for each country is available in the appendix.	Median daily value of temperature in the most populous city NA - no data	Air Quality Open Data Platform

Appendix – Cities considered for climatic variables

Country	City
Afghanistan	Kabul
Algeria	Algiers
Argentina	Buenos Aires
Australia	Sydney
Austria	Vienna
Bahrain	Manama
Bangladesh	Dhaka
Belgium	Brussels
Bolivia (Plurinational State of)	Cochabamba
Bosnia and Herzegovina	Sarajevo
Brazil	SÃ£o Paulo
Bulgaria	Sofia
Canada	Toronto
Chile	Santiago
China	Shanghai
Colombia	BogotÃj
Costa Rica	San JosÃ©
Côte d'Ivoire	Abidjan
Croatia	Zagreb
Cyprus	Nicosia
Czechia	Prague
Denmark	Copenhagen
Ecuador	Quito
El Salvador	San Salvador
Estonia	Tallinn
Ethiopia	Addis Ababa
Finland	Helsinki
France	Paris
Georgia	Tbilisi
Germany	Berlin
Ghana	Accra
Greece	Athens
Guatemala	Guatemala City
Guinea	Conakry
Hong Kong	Hong Kong
Hungary	Budapest
Iceland	ReykjavÃk
India	Mumbai
Indonesia	Jakarta
Iran (Islamic Republic of)	Tehran

Iraq	Baghdad
Ireland	Dublin
Israel	Jerusalem
Italy	Rome
Japan	Tokyo
Jordan	Amman
Kazakhstan	Almaty
Korea (the Republic of)	Seoul
Kuwait	أَبُو ظَبْيَان
Kyrgyzstan	Bishkek
Lao People's Democratic Republic (the)	Vientiane
Lithuania	Kaunas
Macao	Macao
Malaysia	Kuala Lumpur
Mali	Bamako
Mexico	Mexico City
Mongolia	Ulan Bator
Myanmar	Yangon
Nepal	Kathmandu
Netherlands (the)	Amsterdam
New Zealand	Auckland
Norway	Oslo
Pakistan	Karachi
Peru	Lima
Philippines (the)	Manila
Poland	Warsaw
Portugal	Lisbon
Puerto Rico	San Juan
Republic of North Macedonia	Skopje
Romania	Bucharest
Russian Federation (the)	Moscow
Saudi Arabia	Riyadh
Serbia	Belgrade
Singapore	Singapore
Slovakia	Bratislava
South Africa	Cape Town
Spain	Madrid
Sri Lanka	Colombo
Sweden	Stockholm
Switzerland	Zurich
Taiwan (Province of China)	Taipei
Tajikistan	Dushanbe

Thailand	Bangkok
Turkey	Istanbul
Turkmenistan	Ashgabat
Uganda	Kampala
Ukraine	Kyiv
United Arab Emirates (the)	Dubai
United Kingdom	London
United States of America (the)	Washington D.C.
Uzbekistan	Tashkent
Viet Nam	Ho Chi Minh City