**Predictions**

* Number of Cases
* Number of Deaths

**Train Set**

* World data

**Test Set**

* 50 US States

**Possible Features**

* ~~Average life expectancy~~
* ~~Median age (% >65)~~
* ~~Pre-existing conditions~~
  + ~~% diabetes, heart disease~~
* ~~Hospital beds per capita~~
* Hospitals per capita
* ~~Doctors per capita~~
* ~~Population density~~
* ~~Median income~~
* ~~GDP~~
* ~~Gini coefficient~~
* ~~Governance rating?~~
* ~~Ventilators?~~
* ~~Dummy variable for continent~~

Days since 100 confirmed cases

<https://www.bloomberg.com/opinion/articles/2020-04-11/coronavirus-this-pandemic-will-lead-to-social-revolutions>

**Sources**

<https://data-worldbank-org.proxy.library.upenn.edu/>

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.

GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars.

Life expectancy at birth, total (years)

( 1 ) United Nations Population Division. World Population Prospects: 2019 Revision, or derived from male and female life expectancy at birth from sources such as: ( 2 ) Census reports and other statistical publications from national statistical offices, ( 3 ) Eurostat: Demographic Statistics, ( 4 ) United Nations Statistical Division. Population and Vital Statistics Reprot ( various years ), ( 5 ) U.S. Census Bureau: International Database, and ( 6 ) Secretariat of the Pacific Community: Statistics and Demography Programme.

Population ages 65 and above as a percentage of the total population. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship.

Population density is midyear population divided by land area in square kilometers. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship--except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin. 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.

Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. The values shown are midyear estimates. (1) United Nations Population Division. World Population Prospects: 2019 Revision. (2) Census reports and other statistical publications from national statistical offices, (3) Eurostat: Demographic Statistics, (4) United Nations Statistical Division. Population and Vital Statistics Reprot (various years), (5) U.S. Census Bureau: International Database, and (6) Secretariat of the Pacific Community: Statistics and Demography Programme.

Population in largest city is the urban population living in the country's largest metropolitan area.

Population in largest city is the percentage of a country's urban population living in that country's largest metropolitan area.

Hospital beds include inpatient beds available in public, private, general, and specialized hospitals and rehabilitation centers. In most cases beds for both acute and chronic care are included. Data are from the World Health Organization, supplemented by country data.

Physicians include generalist and specialist medical practitioners. World Health Organization's Global Health Workforce Statistics, OECD, supplemented by country data.

Diabetes prevalence refers to the percentage of people ages 20-79 who have type 1 or type 2 diabetes. International Diabetes Federation, Diabetes Atlas.

Cause of death refers to the share of all deaths for all ages by underlying causes. Non-communicable diseases include cancer, diabetes mellitus, cardiovascular diseases, digestive diseases, skin diseases, musculoskeletal diseases, and congenital anomalies. Derived based on the data from WHO's Global Health Estimates.

Cause of death refers to the share of all deaths for all ages by underlying causes. Communicable diseases and maternal, prenatal and nutrition conditions include infectious and parasitic diseases, respiratory infections, and nutritional deficiencies such as underweight and stunting.

Adjusted net national income is GNI minus consumption of fixed capital and natural resources depletion. World Bank staff estimates based on sources and methods in World Bank's "The Changing Wealth of Nations: Measuring Sustainable Development in the New Millennium" (2011).

Current expenditures on health per capita in current US dollars. Estimates of current health expenditures include healthcare goods and services consumed during each year. World Health Organization Global Health Expenditure database (http://apps.who.int/nha/database).

|  |
| --- |
| 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. Percentile rank indicates the country's rank among all countries covered by the aggregate indicator |
| 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. Percentile rank indicates the country's rank among all countries covered by the aggregate indicator, with 0 corresponding to lowest rank, and 100 to highest rank. Percentile ranks have been adjusted to correct for changes over time in the composition of the countries covered by the WGI. |
| Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism. Percentile rank indicates the country's rank among all countries covered by the aggregate indicator, with 0 corresponding to lowest rank, and 100 to highest rank. Percentile ranks have been adjusted to correct for changes over time in the composition of the countries covered by the WGI. |
| Regulatory Quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development. Percentile rank indicates the country's rank among all countries covered by the aggregate indicator, with 0 corresponding to lowest rank, and 100 to highest rank. Percentile ranks have been adjusted to correct for changes over time in the composition of the countries covered by the WGI. |
| 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. Percentile rank indicates the country's rank among all countries covered by the aggregate indicator, with 0 corresponding to lowest rank, and 100 to highest rank. Percentile ranks have been adjusted to correct for changes over time in the composition of the countries covered by the WGI. |
| Voice and Accountability captures perceptions of the extent to which a country's citizens are able to participate in selecting their government, as well as freedom of expression, freedom of association, and a free media. Percentile rank indicates the country's rank among all countries covered by the aggregate indicator, with 0 corresponding to lowest rank, and 100 to highest rank. Percentile ranks have been adjusted to correct for changes over time in the composition of the countries covered by the WGI. |