Description of each dataset

Vaccinations_ByCountry.csv

- **Country**- this is the country for which the vaccination information is provided

- **Country ISO Code** - ISO code for the country

- **Date** - date for the data entry; for some of the dates we have only the daily vaccinations, for others, only the (cumulative) total

- **Total number of vaccinations** - this is the absolute number of total immunizations in the country

- **Total number of people vaccinated** - a person, depending on the immunization scheme, will receive one or more (typically 2) vaccines; at a certain moment, the number of vaccination might be larger than the number of people

- **Total number of people fully vaccinated** - this is the number of people that received the entire set of immunization according to the immunization scheme (typically 2); at a certain moment in time, there might be a certain number of people that received one vaccine and another number (smaller) of people that received all vaccines in the scheme

- **Daily vaccinations (raw)** - for a certain data entry, the number of vaccination for that date/country

- **Daily vaccinations** = for a certain data entry, the number of vaccination for that date/country

- **Total vaccinations per hundred**- ratio (in percent) between vaccination number and total population up to the date in the country

- **Total number of people vaccinated per hundred**- ratio (in percent) between population immunized and total population up to the date in the country

- **Total number of people fully vaccinated per hundred**- ratio (in percent) between population fully immunized and total population up to the date in the country

- **Number of vaccinations per day**- number of daily vaccination for that day and country

- **Daily vaccinations per million**- ratio (in ppm) between vaccination number and total population for the current date in the country

- **Vaccines used in the country**- total number of vaccines used in the country (up to date)

- **Source name**- source of the information (national authority, international organization, local organization etc.)

- **Source website**- website of the source of information

<u>Vaccinations_ByCountry_ByManufacturer.csv</u>

- **Location**- country;

- **Date** - date;

- **Vaccine** - vaccine type;

- **Total number of vaccinations** - total number of vaccinations / current time and vaccine type.

<u>Covid19-TestingRecord.csv</u>

- **Entity**- this is the country for which the vaccination information is provided as well as the type of counting - such as people tested, tests performed or unclear unit;

- **ISO Code** - ISO code for the country;

- **Date** - date for the data entry;

- **Source URL** - this is the source for daily data in the country; usually, this will be a public site;

- **Source Label** - human readable title of the entity providing the information; ex: Ministry of Health;

- **Notes** - explanatory notes related to the collected data;

- **Daily change in cumulative total** - for a certain data entry, the number of tests for that date/country;

- **Cumulative total** - for a certain data entry, the total (cumulative, up to date) number of tests for that date/country;

- **Cumulative total per thousands** - for a certain data entry, the total (cumulative, up to date)number of tests for that date/country per thousands of people in the population;

- **Daily change in cumulative total per thousands** - for a certain data entry, the number of tests for that date/country per thousands of people in the population;

- **7-day smoothed daily change** - averaged value of daily change in number of tests for the current date in the country;

- **7-day smoothed daily change per thousands** - averaged value of daily change in number of tests per thousands people for the current date in the country;

- **Short-term positive rate** - Averaged on short period positive rate for the tests - for date/country;

- **Short-term tests per case** - Averaged on short period tests/confirmed case - for date/country;

<u>Covid19-VariantsFound.csv</u>

- location- this is the country for which the variants information is provided;

- date - date for the data entry;

- variant - this is the variant corresponding to this data entry;

- num_sequences - the number of sequences processed (for the country, variant and date);

- perc_sequences - percentage of sequences from the total number of sequences (for the country, variant and date);

- num_sequences_total - total number of sequences (for the country, variant and date);

<u>Comprehensive Global COVID-19 Dataset.csv</u>

- **Country Name** - The name of the country for which the data is provided

- **Cases** - The total number of confirmed COVID-19 cases in each country.

- **Deaths** - The total number of deaths attributed to COVID-19 in each country.

- **Recovered** - The total number of recovered COVID-19 patients in each country.