Description and disclaimer
The downloadable data file contains information about the volume of COVID-19 sequencing, the number and percentage distribution of variants of concern (VOC) by week and country. Each row contains the corresponding data for a country, variant and week (the data are in long format). The file is updated weekly. You may use the data in line with ECDC's copyright policy and with GISAID's data usage policy. We gratefully acknowledge both the originating and submitting laboratories for the sequence data in GISAID EpiCoV on which these outputs are partially based.

Source
Available data on the volume of COVID-19 sequencing, the number and percentage distribution of VOC for each country, week and variant submitted since 2020-W40 to the GISAID EpiCoV database (https://www.gisaid.org/) and TESSy (as either case-based or aggregate data) are displayed. Where countries have submitted data to both TESSy case-based and aggregate data record types in the same week preference is given to the record type with a valid denominator (see below); if both have a valid denominator the record type with the largest number of sequences reported is selected.

The number of weekly cases per used to estimate the proportion of cases sequenced per week is based on data collected by ECDC Epidemic Intelligence. The information sources are Ministries of Health or National Public Health Institutes (websites, twitter official accounts or Facebook official accounts), and the obtained data is systematically cross checked with data from WHO. More information is available at https://www.ecdc.europa.eu/en/covid-19/data-collection.

Interpretation of COVID-19 data
The 14-day notification rate of newly reported COVID-19 cases is based on data collected by the ECDC Epidemic Intelligence from various sources and are affected by the local testing strategy, laboratory capacity and the effectiveness of surveillance systems. Comparing the epidemiological situation regarding COVID-19 between countries and subnational regions should therefore not be based on these rates alone. However, at the individual country or regional level, this indicator may be useful for monitoring the national situation over time.

Testing policies and the number of tests performed per 100 000 persons, vary markedly across the EU/EEA. More extensive testing will inevitably lead to more cases being detected.

The 14-day notification rate of new COVID-19 cases should be used in combination with other factors including testing policies, number of tests performed, test positivity, excess mortality and rates of hospital and Intensive Care Unit (ICU) admissions, when analysing the epidemiological situation in a country. Most of these indicators are presented for EU/EEA Member States in the Country Overview report.

Even when using several indicators in combination, comparisons between countries should be done with caution and relevant epidemiological expertise.

Variable (as of 20210419)	Definition	Code
Country		String
country_code	2-letter ISO country code	String
year_week		yyyy-Www
Source	Data source, either GISAID EpiCoV database or TESSy.	String
new_cases	Weekly number of new confirmed cases. Set to zero in the event that countries have negative case counts due to retrospective correction of data.	Numeric

number_sequenced	Weekly number of sequences carried out	Numeric
percent_cases_sequenced	100 x new_cases/number_sequenced.	Numeric
valid_denominator	GISAID data: TRUE TESSY data: FALSE if there are discrepancies in the data reported for a given week, such as where the sum of number_detections_variant across all variants exceeds number_sequenced (aggregate data), or where no sequences have been reported that are coded as 'wild type' (case-based data).	Numeric
Variant	Each VOC, Other or UNK	Numeric
number_detections_variant	Number of detections reported of the variant	Numeric
percent_variant	100 x number_detections_variant/ number_sequenced. Np value given if valid_denominator == FALSE	Numeric