CDC Centers for Disease Control and Prevention
CDC 24/7: Saving Lives. Protecting People.™

Data.CDC.gov

🔍 Search

Home   Data Catalog   Developers   Video Guides   🅕 🅞 🆈 ▶ 🅘   Sign In

COVID-19 Case Surveillance Public Use Data with Geography   Case Surveillance

View Data   Visualize ⌄   Export   API   •••

This case surveillance public use dataset has 19 elements for all COVID-19 cases shared with CDC and includes demographics, geography (county and state of residence), any exposure history, disease severity indicators and outcomes, and presence of any underlying medical conditions and risk behaviors.

Currently, CDC provides the public with three versions of COVID-19 case surveillance line-listed data: this 19 data element dataset with geography, a 12 data element public use dataset, and a 32 data element restricted access dataset.

The following apply to the public use datasets and the restricted access dataset:
- Data elements can be found on the COVID-19 case report form located at www.cdc.gov/coronavirus/2019-ncov/downloads/pui-form.pdf.
- Data are considered provisional by CDC and are subject to change until the data are reconciled and verified with the state and territorial data providers.
- Some data are suppressed to protect individual privacy.

- Datasets will include all cases with the earliest date available in each record (date received by CDC or date related to illness/specimen collection) at least 14 days prior to the creation of the previously updated datasets. This 14-day lag allows case reporting to be stabilized and ensure that time-dependent outcome data are accurately captured.
- Datasets are updated monthly.
- Datasets are created using CDC's Policy on Public Health Research and Nonresearch Data Management and Access and include protections designed to protect individual privacy.
- For more information about data collection and reporting, please see wwwn.cdc.gov/nndss/data-collection.html.
- For more information about the COVID-19 case surveillance data, please see www.cdc.gov/coronavirus/2019-ncov/covid-data/faq-surveillance.html.

Overview

The COVID-19 case surveillance database includes patient-level data reported by U.S. states and autonomous reporting entities, including New York City and the District of Columbia (D.C.), as well as U.S. territories and affiliates. On April 5, 2020, COVID-19 was added to the Nationally Notifiable Condition List and classified as "immediately notifiable, urgent (within 24 hours)" by a Council of State and Territorial Epidemiologists (CSTE) Interim Position Statement (Interim-20-ID-01). CSTE updated the position statement on August 5, 2020 to clarify the interpretation of antigen detection tests and serologic test results within the case classification (Interim-20-ID-02). The statement also recommended that all states and territories enact laws to make COVID-19 reportable in their jurisdiction, and that jurisdictions conducting surveillance should submit case notifications to CDC. COVID-19 case surveillance data collected by jurisdictions are shared voluntarily with CDC. For more information, visit:

wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-

covid-19/case-definition/2020/08/05/.

COVID-19 Case Reports

COVID-19 case reports are routinely submitted to CDC by public health jurisdictions using nationally standardized case reporting forms. On April 5, 2020, CSTE released an Interim Position Statement with national surveillance case definitions for COVID-19. Current versions of these case definitions are available at wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/. All cases reported on or after were requested to be shared by public health departments to CDC using the standardized case definitions for lab-confirmed or probable cases. On May 5, 2020, the standardized case reporting form was revised. States and territories continue to use this form.

Data are Considered Provisional

- COVID-19 case surveillance data are dynamic; case reports can be modified at any time by the jurisdictions sharing COVID-19 data with CDC. CDC may update prior cases based on any updated information from jurisdictions. For instance, health departments provide updated data to CDC as new information is gathered about previously reported cases. As more information and data become available, analyses might find changes in surveillance data and trends during a previously reported time window. Data may also be shared late with CDC due to the volume of COVID-19 cases.
- Annual finalized data: To create the final NNDSS data used in the annual tables, CDC works carefully with the reporting jurisdictions to reconcile data received during the year until each state or territorial epidemiologist confirms that the data from their area are correct.

Data Limitations

To learn more about the limitations in using case surveillance data, visit FAQ: COVID-19 Data and Surveillance.

Data Quality Assurance Procedures

CDC's Case Surveillance Section routinely performs data quality assurance procedures (i.e., ongoing corrections and logic checks to address data errors). To date, the following data cleaning steps have been implemented:
- Questions that have been left unanswered (blank) on the case report form are reclassified to a Missing value, if applicable to the question. For example, in the question "Was the individual hospitalized?" where the possible answer choices include "Yes," "No," or "Unknown," the blank value is recoded to "Missing" because the case report form did not include a response to the question.
- Logic checks are performed for date data. If an illogical date has been provided, CDC reviews the data with the reporting jurisdiction. For example, if a symptom onset date in the future is reported to CDC, this value is set to null until the reporting jurisdiction updates the date appropriately.
- Additional data quality processing to recode free text data are ongoing. Data on symptoms, race and ethnicity, and healthcare worker status have been prioritized.

Data Suppression

To prevent release of data that could be used to identify people, data cells are suppressed for low frequency (<11 COVID-19 case records with a given values). Suppression includes low frequency combinations of case month, geographic characteristics (county and state of residence), and demographic characteristics (sex, age group, race, and ethnicity). Suppressed values are re-coded to the NA answer option; records with data suppression are never removed.

option, records with data suppression are never removed.

Additional COVID-19 Data

COVID-19 data are available to the public as summary or aggregate count files, including total counts of cases and deaths by state and by county. These and other COVID-19 data are available from multiple public locations: COVID Data Tracker; United States COVID-19 Cases and Deaths by State; COVID Data Tracker Weekly Review; COVID-19 Surveillance and Data Analytics; COVID-19 Vaccination Reporting Data Systems; and COVID-19 Death Data and Resources.
Less

	Updated March 23, 2021	Data Provided by CDC Case Surveillance Task Force

About this Dataset

Mute Dataset

Updated
March 23, 2021

Data Last Updated
March 22, 2021

Metadata Last Updated
March 23, 2021

Date Created
February 3, 2021

Views	Downloads
7,402	**1,465**

Common Core

Publisher	CDC
Contact Name	Surveillance Review and Response Group
Contact Email	eocevent394@cdc.gov
References	https://data.cdc.gov/api/views/n8mc-b4w4/files/e5006bd3-c0d6-40e2-9eae-167d30fe1183?download=true&filename=faq_covid_cases_public_geo.pdf
Suggested Citation	Centers for Disease Control and Prevention, COVID-19 Response. COVID-19 Case Surveillance Public Use Data with Geography (version date: March 23, 2021).
Bureau Code	009:20
Program Code	009:020

Data Quality

Data Provided by	Dataset Owner
CDC Case Surveillance Task Force	Data.CDC.gov Administrator

Contact Dataset Owner

Geographic Coverage	US
Geospatial Resolution	State, County
Temporal Applicability	2020-01-01/2021-03-07
Update Frequency	Monthly
Analytical Methods Reference	https://github.com/CDCgov/covid_case_privacy_review
Data Dictionary	https://data.cdc.gov/api/views/n8mc-b4w4/files/5245e8f8-d328-4bad-8dc3-4c3a80d73e36?download=true&filename=data_dictionary_covid_cases_public_geo.xlsx

Attachments

faq_covid_cases_public_geo.pdf

data_dictionary_covid_cases_public_geo.xlsx

utility_summary_covid_cases_public_geo.pdf

Topics

Category	Case Surveillance
Tags	covid-19, covid19, coronavirus, cases, surveillance, county, state, microdata

Licensing and Attribution

License	Public Domain U.S. Government
Source Link	https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/index.html

Show Less

What's in this Dataset?

Rows	Columns	Each row is a
21.9M	19	Deidentified Patient Case

Columns in this Dataset

Column Name	Description	Type		
case_month	The earlier of month the Clinical Date (dat...	Plain Text	T	⌄
res_state	State of residence	Plain Text	T	⌄
state_fips_code	State FIPS code	Plain Text	T	⌄
res_county	County of residence	Plain Text	T	⌄
county_fips_code	County FIPS code	Plain Text	T	⌄
age_group	Age group [0 - 17 years; 18 - 49 years; 50 - ...	Plain Text	T	⌄
sex	Sex [Female; Male; Other; Unknown; Missi...	Plain Text	T	⌄

Show All (19)

Table Preview

View Data | Create Visualization

case_...	res_s...	state...	res_c...	coun...	age_...	sex	race	ethni...
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA
2020-01	NA	NA	NA	NA	NA	NA	NA	NA

< Previous   Next >   Showing Deidentified Patient Cases 1 to 14 out of 21,950,114

Related Content Using this Data

Public

COVID-19 Case Surveillance
Public Use Data with Geog…

March 24,	3 Views
2021	

About CDC   Jobs   Funding   Policies   Privacy   FOIA   No Fear Act   OIG

Podcasts/RSS   Contact Us

Powered By