

COVID-19 Case Surveillance Public Use Data

Case Surveillance

View Data

Visualize

Export

API

...

Featured Content Using this Data

Demographic Trends of COVID-19 cases and death...

External Content

CDC is working with states to provide more information on race/ethnicity for reported cases. The percent of reported cases that in...

About this Dataset

Mute Dataset

Updated

January 14, 2021

Data Last Updated

December 31, 2020

Metadata Last Updated

January 14, 2021

Date Created

May 15, 2020

Views

577K

Downloads

29.5K

Data Provided by

CDC Case Surveillance Task Force

Dataset Owner

Data.CDC.gov Administrator

Contact Dataset Owner

Common Core

Publisher	CDC
Contact Name	Surveillance Review and Response Group
Contact Email	eocevent394@cdc.gov
Suggested Citation	Centers for Disease Control and Prevention, COVID-19 Response. COVID-19 Case Surveillance Public Data Access, Summary, and Limitations (version date: December 31, 2020).
Bureau Code	009:20
Program Code	009:020

Data Quality

Geographic Coverage	US
Temporal Applicability	2020-01-01/2020-12-16
Update Frequency	Monthly

Show More

Currently, CDC provides the public with COVID-19 case surveillance data in two ways: an 12 data element public use dataset of the line-listed dataset of all COVID-19 cases shared with CDC and a 32 data element [restricted access dataset](#) of the line-listed dataset of all COVID-19 cases shared with CDC.

What's in this Dataset?

Rows	Columns	Each row is a
13.4M	12	Deidentified Patient

The following are true for both the public use dataset 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.

Columns in this Dataset

	Column Name	Description	Type	
- 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.	cdc_case_earliest_dt	The earlier of the Clinical Date (date r...	Date & Time	⌵
- Data are suppressed to protect individual privacy.	cdc_report_dt	Initial case report date to CDC. Depre...	Date & Time	⌵
- The datasets include all cases with an initial report date of case to CDC at least 14 days prior to the creation of the previously updated datasets. This 14-day lag will allow case reporting to be stabilized and ensure that time-dependent outcome data are accurately captured.	pos_spec_dt	Date of first positive specimen collect...	Date & Time	⌵
	onset_dt	Symptom onset date, if symptomatic	Date & Time	⌵
- Datasets are updated monthly.	current_status	Case Status: Laboratory-confirmed ca...	Plain Text	T ⌵
- Both datasets are created using CDC's operational Policy on Public Health Research and Nonresearch Data Management and Access and include protections designed to protect individual privacy.	sex	Sex: Male; Female; Unknown; Other	Plain Text	T ⌵
	age_group	Age Group: 0 - 9 Years; 10 - 19 Years; ...	Plain Text	T ⌵
- For more information about data collection and reporting, please see https://wwwn.cdc.gov/nndss/data-collection.html				

Show All (12)

- For more information about the COVID-19 case surveillance, please see <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/fao-surveillance.html>

Table Preview

View Data

Create Visualization

<https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/08/05/>

The deidentified data in the public use dataset include demographic characteristics, exposure history, disease severity indicators and outcomes, clinical data, laboratory diagnostic test results, and comorbidities. All data elements can be found on the COVID-19 case report form located at www.cdc.gov/coronavirus/2019-ncov/downloads/pui-form.pdf.

The Case Surveillance Task Force and Surveillance Review and Response Group (SRRG) within CDC's COVID-19 Response provide stewardship for datasets that support the public's access to COVID-19 data while protecting individual privacy.

COVID-19 Case Reports

COVID-19 case reports have been routinely submitted using standardized case reporting forms. On April 5, 2020, CSTE released an Interim Position Statement with national surveillance case definitions for COVID-19 included. Current versions of these case definitions are available here: <https://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. Implementation of case reporting using this new form is ongoing among U.S. states and territories.

Data are Considered Provisional

- The 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 shared with CDC based on any updated information from jurisdictions.
- National case surveillance data are constantly changing. For instance, as new information is gathered about previously reported cases, health departments provide updated data to CDC. As more information and data become available, analyses might find changes in

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 re-classified 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 missing value is re-coded to Missing if the respondent did not answer 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 that is in the future is reported to CDC, this value is set to null until the reporting jurisdiction updates this information appropriately.
- The initial report date of the case to CDC is intended to be completed by the reporting jurisdiction when data are submitted. If blank, this variable is completed using the date the data file was first submitted to CDC.
- Additional data quality processing to recode free text data are ongoing. Data on

cdc_c... :	cdc_r... :	pos_s... :	onset... :	curre... :	sex :	age_g... :	race... :	hosp... :	icu_yn :	death... :	medc... :
2020/03/23	2020/03/31	2020/03/23	Related Content Using this Data	Laborator...	Female	0 - 9 Years	Black, No...	Unknown	Unknown	Unknown	Unknown
2020/03/22	2020/03/23	2020/03/23		Laborator...	Female	0 - 9 Years	Hispanic/L...	Yes	Unknown	Unknown	Unknown
<div><div><div><div> COVID-19 Case Surveillance</div><div>Public Use Data Profile</div></div><div><div>January 21, 2021</div><div>22.5K Views</div></div><div><div><div><div></div></div></div><div>Summary of variables.</div><div>The COVID-19 case surveillance system dat...</div></div></div></div>				Laborator...	Female	0 - 9 Years	Hispanic/L...	No	No	No	No
				Laborator...	Female	0 - 9 Years	Hispanic/L...	No	Missing	No	No
				Laborator...	Female	0 - 9 Years	Hispanic/L...	Unknown	Unknown	Unknown	Unknown
				Laborator...	Female	0 - 9 Years	Hispanic/L...	No	Missing	No	No
				Laborator...	Female	0 - 9 Years	Hispanic/L...	Yes	Unknown	Unknown	Unknown
				Laborator...	Male	0 - 9 Years	Missing	Missing	Missing	Missing	Missing
				Laborator...	Male	0 - 9 Years	Missing	Yes	No	No	Yes
				Laborator...	Male	0 - 9 Years	Black, No...	Yes	Unknown	Unknown	Unknown
				Laborator...	Male	0 - 9 Years	Unknown	Missing	Missing	Missing	Missing
2020/03/23	2020/03/23	2020/03/23	Laborator...	Male	0 - 9 Years	Unknown	Missing	Missing	Missing	Missing	
2020/03/23	2020/03/25	2020/03/23	Laborator...	Male	0 - 9 Years	Unknown	Unknown	Unknown	Unknown	Unknown	

<https://www.cdc.gov/covid-data-tracker/index.html>

<https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html>

<https://www.cdc.gov/coronavirus/2019-ncov/php/open-america/surveillance-data-analytics.html>

