

CIC Challenge Data Set Variable Dictionary

The purpose of this document is to list and explain the different variables within the `cic_challenge_data_set` and the `cic_challenge_data_set_lat_long`.

Regional/Time Variables

- `state_ID` – dummy numeric variable created. Numbers 1-52 (IE: 1 = Alabama etc.)
- `State_` - 50 sovereign states. Territories were not included.
- `Year` – 2021, 2022, 2023

Population Distribution Variables – are percentages transformed into whole numbers unless otherwise noted.

KFF Notes: 1-year estimates are based off of the 2008-2023 American Community Survey and may differ from those published by the yearly by the Census Bureau. Population data used by KFF is restricted to the civilian, non-institutionalized population for whom ACS collects and reports poverty information. Persons of Hispanic origin may be of any race; all other racial/ethnic groups are non-Hispanic. Population numbers are rounded to the nearest 100.

- `Tot_Pop` – Measures the total population reported by State. *Whole number
- `Pop_Dist_White` – Total White population reported by State.
- `Pop_Dist_Black` – Total Black population reported by State.
- `Pop_Dist_Hispanic` – Total Hispanic population reported by State.
- `Pop_Dist_Asian` – Total Asian population reported by State.
- `Pop_Dist_Am_In` – Total Native American population reported by State.
- `Pop_Dist_Nat_Haw_PacIsl` – total Native Hawaiian/Pacific Islander population by State.
- `Pop_Dist_Mult_race` – Total Multiracial population by State.

Economic Variables are percentages transformed into whole numbers unless otherwise noted as whole number.

KFF Notes: 1-year estimates are based off of the 2008-2023 American Community Survey.

Pov and Povbr Variables: Definition of poverty is based on the U.S Census poverty threshold for a family with two adults and one child (the official measurement that is used by the Federal Government). You can see current and historical thresholds at <https://www.census.gov/data/tables/timeseries/demo/incomepoverty/historicalpovertythresholds.html>. Data may not sum do totals due to rounding.

Unemployment Variable: Unemployment rate represents the percentage of unemployed individuals within the civilian, non-institutional population ages 16 and older. Data are seasonally adjusted and based on place of residence. State level data are sourced from the LASST series via the BLS API. State-level figures may not sum to the national total because BLS independently develops the national employment series and does not force state estimates to align. Area definitions align with the Office of Management and Budget (OMB) Bulletin No. 15-01 (July 15, 2015) and are available of the [Bureau of Labor Statistics website](#).

GSP Variable: GSP = Gross State Product and is a measurement of a state's output. It is the sum of value added from all industries in the state. GSP is the state counterpart to the nation's Gross Domestic Product (GDP). GDP is the market value of goods and services produced by labor and property in the United States, regardless of nationality. Current dollar GAP components are compensation of employees, taxes on production and imports, and gross operating surplus.

SNAP Variables: KFF pulls national and state level monthly data from the United States [Department of Agriculture](#) (USDA). All data are subject to revision.

- Povbr_White – Population of White participants who fall under poverty threshold.
- Povbr_Black - Population of Black participants who fall under poverty threshold.
- Povbr_Hispanic - Population of Hispanic participants who fall under poverty threshold.
- Povbr_As_Pi - Population of Hawaiian/Pacific Islander participants who fall under poverty threshold.
- Povbr_Am_Na - Population of Native American participants who fall under poverty threshold.
- Povbr_Multiple_Races - Population of Multi-racial participants who fall under poverty threshold.

CIC Challenge Data Set Variable Dictionary

- Mean_Unempl – Mean of participants who reported being unemployed. *Whole number
- Tot_GSP_Mil – Gross State Product in millions of US dollars. *Whole number
- snap_tot_ben_Mil – Total annual spend on SNAP benefits. *Whole number
- SNAP_no_rep -Average Number of SNAP monthly participants over a year. *Whole number
- Pov_Ad_Ch – Adults with children living below the poverty threshold.
- Pov_Ad_No – Adults with no children living below the poverty threshold.
- Pov_Ad_Tot – Total Adults living below the poverty threshold.

Mental Health Variables - percentages transformed to whole numbers.

KFF Notes: This variable measures the percentage of survey participants who experienced a certain number of self-reported “bad mental health days” over a 30-day period. The KFF Custom State Report tool pulled data from the CDC’s Behavioral Risk Factor Surveillance System (BFRSS) 2013-2023. Which is a random digit dialed telephone survey of noninstitutionalized civilian adults aged 18 years and older.

Null Values- Florida was unable to collect data over enough months to meet the minimum requirements for inclusion in 2021. Kentucky and Pennsylvania did not meet the requirements for inclusion in 2023.

- Ad_Poor_MH_zero – participant experienced 0 bad mental health days
- Ad_Poor_MH_four – participant experienced 1-4 bad mental health days
- Ad_Poor_MH_thirteen – participant experienced 5-13 bad mental health days •
Ad_Poor_MH_fourteen – participant experienced ≥ 14 bad mental health days.

Covid19 Vaccination Rates –

CDC Socrata Official Variable Code Book is available in excel form within the raw downloads/CDC folder.

CIC Challenge Data Set Variable Dictionary

- Cov_Vacc_one – is the annual sum of all 7-day (MMWR_week) moving average counts of people with at least one dose of any vaccine by date of administration. *Whole number

Centroids - cic_challenge_data_set_long_lat - Google Data Set Publishing Language

These coordinates represent the geographic central locations within the 50 sovereign states. (Needed for JSON folium maps.)

- Latitude
- Longitude