

Group Dictionaries

County by Population

Name	Definition	Data Type	Possible Values	Required
countyFIPS	Unique Identifying Integer for the County	Integer	0, 1011, 1012, 1035	Yes
County Name	Colloquial English name of the county	String	Calhoun County, Henry County, Etowah County	Yes
State	Two letter unique identifier for the state	String	AL, NC, SC, VA	Yes
Population	Integer that identifies county population	Integer	0, 55869, 19448	Yes

Number of Deaths Per Date

Name	Definition	Data Type	Possible Values	Required
countyFIPS	Unique Identifying Integer for the County	Integer	0, 1011, 1012, 1035	Yes
County name	Colloquial English name of the county	String	Calhoun County, Henry County, Etowah County	Yes
State	Two letter unique identifier for the state	String	AL, NC, SC, VA	Yes
StateFIPS	Integer identifier for the state	Integer	1, 20, 35	Yes
Deaths Until Date	Amount of deaths recorded	Integer	0, 101, 200	Yes

	up to this day			
--	----------------	--	--	--

Number of New Cases Per Date

Name	Definition	Data Type	Possible Values	Required
countyFIPS	Unique Identifying Integer for the County	Integer	0, 1011, 1012, 1035	Yes
County name	Colloquial English name of the county	String	Calhoun County, Henry County, Etowah County	Yes
State	Two letter unique identifier for the state	String	AL, NC, SC, VA	Yes
StateFIPS	Integer identifier for the state	Integer	1, 20, 35	Yes
New Cases Up To Date	Amount of cases up to this date	Integer	0, 101, 200	Yes

Preliminary Observations

Taking a look at the new cases and deaths data, there appears to be a decline in new cases and deaths. The peak is somewhere in the middle, which is what we would expect considering the peak of Covid activity is past.

Individual Dictionaries

President By County - Cole Yonkers

Name	Definition	Data Type	Possible Values	Required
State	State the county is in	String	Florida, District of Columbia, North Carolina	Yes
County	Name of the county the election took place in	String	Kent County, Bradford County	Yes
Candidate	Name of the presidential candidate	String	Joe Biden, Donald Trump, Jo Jorgenson	Yes
Party	Political party the presidential candidate is affiliated with	String	DEM, REP, GRN, LIB, PSL	Yes
Total_vote	How many total votes a candidate received in the county	Integer	420, 437, 46, 39397	Yes
won	Boolean value that displays whether or not the candidate won (true) lost (false)	Boolean	True, False	Yes

The data can be merged with the main datasets by joining based on county and state name, however the data requires a few alterations. County must become County Name and a two letter abbreviation for the state must be included, luckily this can be done in excel.

The data can help determine how political leanings, propaganda, and biases might have affected the spread, lethality, and influence of Covid-19 during the election cycle.

Employment Dataset - Calvin Ng

Name	Definition	Data Type	Possible Values	Required
Year	The year this data was recorded	Integer	2021	Yes
Qtr	The yearly quarter the data was recorded. Because this is the overall data, this column is irrelevant	Char	A	Yes
Area Type	This specifies whether the row is a county, state, or nationwide	String	Nation, State, County	Yes
St Name	This is the name of the state	String	Alabama, North Carolina, NULL	No
Area	This is the name of the area	String	Alabama – Statewide, U.S. Total, Rockland County, New York	Yes
Ownership	Refers to who/what owns the industry	String	Federal Government, Total Covered, Private	Yes
Industry	This is the industry the row is reporting data for; the format is the industry code followed by its name	String	“10 Total, all industries”, “1012 Construction”, “1024 Professional and Business Services”	Yes

Status Code	This marks whether the information is disclosed or not	Char/String	N, Null	No
Annual Average Establishment Count	This is the number of establishments in a given area	Int	142,212, 1,251, 28	Yes
Annual Average Employment	The number of positions employed	Int	2,883,380, 193,660	Yes
Annual Total Wages	The total sum of wages earned by all employees in the year	Int	2,622,652,420,011, 1,312,978	Yes
Annual Average Weekly Wage	The average wage of an employee in a week	Int	883, 539, 913	Yes
Annual Average Pay	The average wage of an employee in a year	Int	67,610, 73,020	Yes
Employment Location Quotient Relative to U.S.	A number representing the local share of an industry in comparison to the national share as a whole. A number larger than one indicates a greater local share than nationwide, and a number smaller than one indicates a smaller local share than	Float	0.43, 1.54, 0.91	Yes

	nationwide.			
Total Wage Location Quotient Relative to U.S.	A number representing the relative wages paid in a certain industry, quarter, and area in comparison nationwide. A number higher than 1 indicates higher wages than the average for that industry nationwide, and a number lower than 1 indicates lower wages than the average for that industry nationwide.	Float	0.84, 1.62, 0.44	Yes

The way to join this dataset with the main data would be to join on account of county name, statename, and date. This dataset only covers data from the first half of 2022, so we would only want to use a subset of the main dataset. We would sum up the cases/deaths for all the dates in the first quarter per county and per state and would map those to county and state information of this dataset.

This enrichment dataset is valuable because we can analyze how the spread of covid impacted the economy on a county and state level. One of the things that often comes up when discussing covid is its economical impact, and this gives us a way to measure it. Some questions we can answer with this data are which industries were impacted the most, which geographical areas were impacted the most, etc.

Demographic and Housing Dataset - Kachif Adefalou

Name	Definition	Data Type	Possible Values	Required
State	State in which data were reported	String	New York, Texas, Georgia	Yes

Population	Total population in the state where census is reported	Int	19835913	Yes
Sex	Sexual orientation of individual based on population report	String	Male, Female, etc..	Yes
Age	Age group of individual based on population report	Int	25 to 34 years	Yes
Race	Race of individual based on population report	String	Black, White, etc.	Yes
Hispanic or Latino and Race	Boolean values that indicate whether or not individual reporting is of hispanic origin	Boolean	True, False	Yes
Citizen, Voting Age Population	Citizen of voting age in the specific state	Int	14,109,037	No

We could possibly merge the primary COVID-19 dataset with this data by joining either the population or the state name. We might have to change State name to the abbreviated version of each state if using state name as the mapping variable otherwise population should match without any issue.

This data could help determine the spread of Covid-19 based on age group, as well as the level of infection in regards to Race or Sex