County-by-Year California Data Set Documentation

Overview

The longer-run county-by-year data set includes a limited number of county-level economic variables data for all 58 counties in the state of California for the years 2008 to 2022. Because this data set contains relatively few variables, it works best as a starting point for merging with other data sets, particularly if you want to use a longer sample of data.

Data set is `.csv` formatted – to load into R, use the `read_csv()` function from `tidyverse`.

- Each row of the data set corresponds to a county and year, so rows can be uniquely identified using a combination of `countyfip` or `county_name` and `year`.
- Documentation below lists all variables and includes definitions and coding notes.
- All variables for this data set are downloaded from the [St. Louis Fed's FRED website](#).

Variable Name	Definition	Coding and Notes
`year`	Year (covers 2008-2022)	
`countyfip`	3-digit county FIPS code – numeric ID for counties	See [link here](#) for list of counties and codes (you can also use the county_name variable to ID counties).
`county_name`	County name (string format)	
`state_name`	State name (string format)	
`total_population`	Total population	Reported by the Census Bureau.
`per_capita_personal_income`	*Nominal* per capita personal income	Note that dollar values here are **nominal**. To make comparisons in income levels over time, you should calculate real values. Personal income is income that is received by individuals from all sources including wages and salaries, rental income, transfer receipts, etc.
`unemp_rate`	Unemployment rate	Percentage of the labor force that is currently employed. Calculated as the average of the monthly unemployment rate within each county and year reported by the Census Bureau. Note that values here range from 0 to 29.39 – values are **not** fractions bounded between 0 and 1.