

## Longer-Run State-by-Year Data Set Documentation

### Overview

The longer-run state-by-year data set includes a limited number of state-level economic variables data for all 50 states for the years 2008 to 2022 (the District of Columbia is not included). 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 state and year, so rows can be uniquely identified using a combination of `statefip` or `state_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)	
statefip	State FIPS code – numeric ID for states	See <a href="#">link here</a> for list of states and codes (you can also use the <code>state_name</code> variable to ID states)
state_name	State name (string format)	
total_population	Total population	Reported by the Census Bureau.
income_per_capita	Real per capita personal income (in 2017 dollars)	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	Fraction of the labor force that is currently employed. Calculated as the average of the monthly unemployment rate within each state and year reported by the Census Bureau.