

Community Population Survey

[ECON 5753] — *University of Arkansas*

The [CPS](#) contains survey data on workers collected by the US Government. I downloaded this from [IPUMS](#) using 2017, 2018, and 2019 data. Then I slightly cleaned it to create some useful variables.

Table 1: Codebook

variable	description
year	Survey year
statefip	State FIPS code of state where surveyed person lives
hhincome	Total household income
incwage	The surveyed person's annual income/wages
age	Person's age
race	String identifying a person's race/ethnicity
male	=1, if the surveyed person a male
marital_status	String identifying a person's marital status
veteran_status	=1, if the surveyed person is a veteran
citizen_status	=1, if the surveyed person is a U.S. citizen (including naturalized)
has_college_experience	=1, if the surveyed person has at least some years of college attendance
years_educ	The number of years a person was educated for

Preview of data:

Rows: 226,666

Columns: 12

\$ year	<dbl> 2017, 2017, 2017, 2017, 2017, 2017, 2017, 2017,~
\$ statefip	<dbl> 23, 23, 23, 23, 23, 23, 23, 23, 23, 23, 23, 23,~
\$ hhincome	<dbl> 53400, 53400, 53400, 106010, 106010, 70821, 708~
\$ age	<dbl> 55, 42, 21, 59, 60, 28, 29, 62, 51, 45, 17, 58,~
\$ race	<chr> "White", "White", "White", "White", "White", "W~
\$ incwage	<dbl> 25000, 20000, 8400, 55002, 50002, 37000, 32800,~
\$ male	<dbl> 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0,~
\$ marital_status	<chr> "Married, Spouse Present", "Married, Spouse Pre~
\$ veteran_status	<dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
\$ citizen_status	<dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
\$ has_college_experience	<dbl> 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1,~
\$ years_educ	<dbl> 12, 14, 12, 16, 14, 12, 16, 14, 14, 14, 10, 16,~