

STATEICP

Obtain data

To access the data from IPUMS USA, we first selected “IPUMS USA” on the IPUMS website, then clicked “Get Data” and chose the “2022 ACS” under “SELECT SAMPLE.” For state-level information, we navigated to “HOUSEHOLD” > “GEOGRAPHIC” and added “STATEICP” to our cart. For individual-level data, we went to “PERSON” and added “EDUC.” After reviewing the selections in the cart, we clicked “CREATE DATA EXTRACT.” We then changed the “DATA FORMAT” to “.dta” and submitted the extract. After logging in or creating an account, we received an email notification once the extract was ready. We downloaded the file (e.g., “usa_00004.dta”) and saved it locally for use in R.

Overview of the ratio estimators approach

The ratio estimators approach is a statistical method used to estimate population totals or averages by leveraging known ratios from a sample. This technique involves calculating the ratio of a specific characteristic (e.g., the number of doctoral degree holders) to the total population in a known subset (such as California). The same ratio is then applied to other subsets to estimate their totals, under the assumption that similar relationships hold across the broader population. This method is especially helpful when the exact population size is unknown, but sample data provides proportional relationships that can be extended to the larger group.

Estimates and the actual number

Explanation

The estimates of total respondents in each state using the ratio estimators approach may differ from the actual numbers due to several factors:

- If the estimation is based on a sample rather than a full population census, random variability in the sample can impact the calculated ratio and the accuracy of the estimates.
- Educational attainment is not evenly distributed across the U.S. Factors like regional policies, cultural norms, and access to higher education differ, so the ratio derived from California may not apply to other states.
- The ratio estimator assumes that the proportion of doctoral degree holders in California reflects the proportion in other states. However, educational attainment varies widely due to differences in demographics, economic conditions, and educational resources across states. This variation can lead to discrepancies between estimated and actual figures.
- The Laplace ratio method assumes that the relationship between the characteristic of interest (doctoral degree holders) and the total population is consistent across all units. If California's ratio is not reflective of other states due to unobserved factors, the estimates will be biased.

These factors highlight that the assumption of uniformity in ratio estimators often leads to differences when applied to diverse populations, such as the varying states in the U.S.