

Estimating the Number of Respondents in U.S. States with a Doctoral Degree*

Kevin Cai

October 3, 2024

1 Instructions to Obtain the Data

To replicate the analysis, please follow these steps to obtain the 2022 ACS data from IPUMS (Ruggles et al. (2022)):

1. Go to [IPUMS USA](#).
2. Navigate to the “Select Data” page and select the 2022 ACS dataset.
3. Choose relevant variables, including EDUCD (to identify respondents with doctoral degrees) and STATEICP (to identify the state).
4. Download the dataset in `.dta` format.
5. Extract the data and clean it using R Core Team (2024), and R library Wickham et al. (2019).

2 Overview of Estimation Approach

To estimate the number of respondents with doctoral degrees as their highest education in each state, I used the ratio estimators approach of Laplace. The ratio estimators approach is used to estimate the total population size using a known ratio from a sample (Alexander (2023)). In this case, we first use the dataset we obtained and find the ratio of people with doctoral degree in California. The ratio can be obtained by dividing the number of people with doctoral degree in California by the total number of respondents in California. Then we can apply that ratio to other states to get an estimate of the number of doctoral respondents in each state.

*Code and data are available at: <https://github.com/kevicai/us-education-estimate>

3 Analysis

Table 1: Estimated vs. Actual Number of Doctoral Degree Holders by State

State	Actual Doctoral Count	Estimated Doctoral Count
alabama	460	835
alaska	51	113
arizona	896	1201
arkansas	251	507
california	6336	6336
colorado	1031	969
connecticut	600	605
delaware	152	156
district of columbia	311	109
florida	2731	3528
georgia	1451	1771
hawaii	214	243
idaho	175	322
illinois	1457	2074
indiana	620	1131
iowa	258	544
kansas	321	485
kentucky	448	755
louisiana	450	730
maine	165	235
maryland	1608	1011
massachusetts	2014	1184
michigan	991	1644
minnesota	572	955
mississippi	263	483
missouri	621	1046
montana	113	180
nebraska	153	324
nevada	282	498
new hampshire	244	228
new jersey	1438	1509
new mexico	350	328
new york	2829	3303
north carolina	1421	1769
north dakota	60	131
ohio	1213	1954

Table 1: Estimated vs. Actual Number of Doctoral Degree Holders by State

State	Actual Doctoral Count	Estimated Doctoral Count
oklahoma	281	639
oregon	647	708
pennsylvania	1620	2148
rhode island	177	168
south carolina	647	885
south dakota	71	151
tennessee	841	1172
texas	3216	4745
utah	428	576
vermont	131	111
virginia	1531	1438
washington	1195	1309
west virginia	159	294
wisconsin	513	1004
wyoming	72	97

From Table 1, we can see that there is a difference between the estimated and actual number of doctoral degree holders by state. I suspect this difference is caused by various factors affecting the quality of education that residents in each state can receive. Factors such as whether the state has good high schools to prepare residents for university, and whether it has research universities, will all influence the rate at which people attain doctoral degrees. This explains why using California as a reference for all states results in a noticeable difference among them. California is a relatively affluent state, leading to a higher quality of education. Consequently, California has a higher ratio of individuals with doctoral degrees as their highest level of education compared to most other U.S. states, as seen in the table.

References

- Alexander, Rohan. 2023. “Telling Stories with Data.” *Tellingstorieswithdata.com*, July. <https://tellingstorieswithdata.com/>.
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2022. “IPUMS USA: Version 11.0.” Minneapolis, MN: IPUMS. <https://doi.org/10.18128/d010.v11.0>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.