Week 4 Reflection Exercise

Quarto

In this paper we explore the US 2022 Census Data, sourced from @ipums2024 We use @citeR, @readr, @dplyr, @xie2023knitr

Running Code

The data is obtained from IPSUMS USA. Once on their website, navigate to Get Data. This brings up a search function. First, click select samples, deselect the "Default sample from each year" tick, and manually tick 2022 ACS. This is data taken in the US, not quite at the scale of a census, but involves much more thorough questioning. Click on 'Submit Summary' to add this dataset. Under Harmonized Variables, select the following through the dropdown menus: Household -> State -> STATEICP, Person -> Demographic -> SEX, Person -> Education -> EDUC. Finally, click on 'View Cart' and follow the steps required to download data (an IPSUMS account is required when checking out).

```
Rows: 3373378 Columns: 14
-- Column specification -----
Delimiter: ","
dbl (14): YEAR, SAMPLE, SERIAL, CBSERIAL, HHWT, CLUSTER, STATEICP, STRATA, G...
i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
 [1] "YEAR"
                "SAMPLE"
                           "SERIAL"
                                       "CBSERIAL" "HHWT"
                                                             "CLUSTER"
 [7] "STATEICP" "STRATA"
                           "GQ"
                                       "PERNUM"
                                                  "PERWT"
                                                             "SEX"
[13] "EDUC"
                "EDUCD"
```

The ratio estimator approach involves the ratio of two random variables ($R = a_x/a_y$. It is used to estimate the population given a ratio and a preexisting population value ($a_y = R*a_x$). In this case, we are using California's ratio of doctorates (R), and the number of correspondents in California (a_x) to estimate the number of correspondents in other states (a_y).

Table 1: Comparison of Actual and Estimated Total Respondents in Each State Based on Doctoral Degree Ratio

$\begin{array}{cccccccccccccccccccccccccccccccccccc$	STATEICP	$actual_total_respondents$	$doctoral_count$	$estimated_total_respondents$
3 73077 2014 124340.024 4 14077 244 15064.035 5 10401 177 10927.596 6 6860 131 8087.658 11 9641 152 9384.153 12 93166 1438 88779.024 13 203891 2829 174656.370 14 132605 1620 100015.312 21 128046 1457 89952.043 22 69843 620 38277.465 23 101512 991 61182.207 24 120666 1213 7488.009 25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 <	1	37369	600	37042.708
4 14077 244 15064.035 5 10401 177 10927.595 6 6860 131 8087.658 11 9641 152 9384.153 12 93166 1438 88779.024 13 203891 2829 17456.377 14 132605 1620 100015.312 21 128046 1457 89952.043 22 69843 620 38277.465 23 101512 991 61182.207 24 120666 1213 7488.009 25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 1989 153 9445.891 36 8107 60 3704.271 37 9296 71<	2	14523	165	10186.745
5 10401 177 10927.598 6 6860 131 8087.658 11 9641 152 9384.153 12 93166 1438 88779.024 13 203891 2829 174656.370 14 132605 1620 100015.312 21 128046 1457 89952.043 22 69843 620 38277.465 23 101512 991 61182.207 24 120666 1213 74888.009 25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 5894 572 35314.049 34 64551 621 3839.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 153	3	73077	2014	124340.024
6 6860 131 8087.658 11 9641 152 9384.153 12 93166 1438 88779.024 13 203891 2829 174656.370 14 132605 1620 100015.312 21 128046 1457 89952.043 22 69843 620 38277.465 23 101512 991 61182.207 24 120666 1213 7488.009 25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28394.410 42 31288 251 15496.200	4	14077	244	15064.035
11 9641 152 9384.153 12 93166 1438 88779.024 13 203891 2829 174656.370 14 132605 1620 100015.312 21 128046 1457 89952.043 22 69843 620 38277.465 23 101512 991 61182.207 24 120666 1213 74888.009 25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28394.153 42 31288 251 15496.200 43 217799 2731 16806.060 <td>5</td> <td>10401</td> <td>177</td> <td>10927.599</td>	5	10401	177	10927.599
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	6	6860	131	8087.658
13 203891 2829 174656.370 14 132605 1620 100015.312 21 128046 1457 89952.043 22 69843 620 38277.465 23 101512 991 61182.207 24 120666 1213 74888.009 25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031	11	9641	152	9384.153
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	12	93166	1438	88779.024
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	13	203891	2829	174656.370
22 69843 620 38277.465 23 101512 991 61182.207 24 120666 1213 74888.009 25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	14	132605	1620	100015.312
23 101512 991 61182.207 24 120666 1213 74888.009 25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	21	128046	1457	89952.043
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	22	69843	620	38277.465
25 61967 513 31671.516 31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	23	101512	991	61182.207
31 33586 258 15928.365 32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	24	120666	1213	74888.009
32 29940 321 19817.849 33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	25	61967	513	31671.516
33 58984 572 35314.049 34 64551 621 38339.203 35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	31	33586	258	15928.365
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	32	29940	321	19817.849
35 19989 153 9445.891 36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	33	58984	572	35314.049
36 8107 60 3704.271 37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	34	64551	621	38339.203
37 9296 71 4383.387 40 88761 1531 94520.644 41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	35	19989	153	9445.891
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	36	8107	60	3704.271
41 51580 460 28399.410 42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	37	9296	71	4383.387
42 31288 251 15496.200 43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	40	88761	1531	94520.644
43 217799 2731 168606.061 44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	41	51580	460	28399.410
44 109349 1451 89581.616 45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	42	31288	251	15496.200
45 45040 450 27782.031 46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	43	217799	2731	168606.061
46 29796 263 16237.054 47 109230 1421 87729.481 48 54651 647 39944.387	44	109349	1451	89581.616
47 109230 1421 87729.481 48 54651 647 39944.387	45	45040	450	27782.031
48 54651 647 39944.387	46	29796	263	16237.054
	47	109230	1421	87729.481
49 292919 3216 198548.917	48	54651	647	39944.387
	49	292919	3216	198548.917

STATEICP	$actual_total_respondents$	$doctoral_count$	$estimated_total_respondents$
51	46605	448	27658.556
52	62442	1608	99274.458
53	39445	281	17348.335
54	72374	841	51921.530
56	18135	159	9816.318
61	74153	896	55317.111
62	59841	1031	63651.720
63	19884	175	10804.123
64	11116	113	6976.377
65	30749	282	17410.073
66	20243	350	21608.247
67	35537	428	26423.799
68	5962	72	4445.125
71	391171	6336	391171.000
72	43708	647	39944.387
73	80818	1195	73776.727
81	6972	51	3148.630
82	14995	214	13211.899
98	6718	311	19200.470

Our estimated number of respondents were much higher than the actual number of respondents for most states that are not California. One reason could be that California's count of doctorates is much larger than most other states, which would inflate the correspondent estimator. If we had used another state for our estimator, one closer to the average number across all states, the estimator would get closer to the actual number.