

SM 625: Week 5 Sampling Project Notes

For this week, we will consider the information available for stratified sampling of students. Eventually you are going to design a stratified cluster sample of students, where the clusters (or PSUs) are schools, but we aren't there yet.

Recall the regions of interest in the sampling project description:

Region	Counties (COUNTY_ID on the sampling frame)
1	7, 31, 42, 66
2	22, 27, 36, 55
3	2, 21, 52
4	17, 48, 49, 77
5	1, 4, 6, 16, 20, 26, 35, 60, 65, 68, 69, 71, 72
6	5, 10, 15, 18, 24, 28, 40, 43, 45, 51, 53, 57, 67, 83
7	3, 8, 11, 12, 13, 14, 34, 39, 41, 54, 59, 61, 62, 64, 70, 75, 80
8	9, 19, 23, 25, 29, 30, 33, 37, 38, 46, 47, 56, 73, 78, 81
9	32, 44, 50, 58, 63, 74, 76, 79, 82

As "State officials are interested in providing, if at all possible, separate estimates for each of nine education regions in the state, where the regions are defined by groups of counties", we will use these nine regions as strata.

Prepare a table that includes the overall population counts in each of these nine strata (the total count of students in the target population at each school is in the **tot_all** column on the sampling frame). Given these counts, once you have the *working* overall sample size (unknown for now and will be decided by your team next week), what is the *proportionate* allocation plan of that sample of students across these nine strata?