Homework Assignment I: Survey Sampling

Enter your student ID

Enter your group name (Niklas, Jonas)

Enter submission date

Please write up your answers, code, and results using this Markdown file as a template. Save the file as `HA1_yourstudentID.Rmd`. When everything works properly, knit the final version to HTML format and submit both Rmd and html files to the submission folder Homework 1 which you can find on ILIAS.

The deadline for submission is 16.12.2024, 23:55.

In our application using the `minaret`data, stratified sampling with equal sample sizes for each of the 26 cantons led to an efficiency loss relative to SRSWOR. Even when we allocated the sample proportional to cantonal population sizes, this only led to minimal efficiency gains. Other than bad luck, one reason for this might be the relatively large number of strata (and thus, large sampling variances within strata, particularly those with small samples). Your task is to improve the sampling design following the bullet points below.

You can earn one credit per bullet point. The maximum number of credits is 6. Grading key: 6 credits $\rightarrow$ grade 1, 5 credits $\rightarrow$ 2, 4 credits $\rightarrow$ 3, 3 credits $\rightarrow$ 4, 2 or 1 credits $\rightarrow$ 5. You will get one extra credit for correct Markdown syntax, package, and data handling.

1. Think about ways to group the cantons into fewer (say, 2 to 7) politico-culturally homogeneous regions (note that there is no *one* correct solution here). Explain your choice.
2. Recode `canton_id` into a new variable `region_id` according to your scheme. Sort minaret frame by `region_id` (this is a necessary for the sampling functions to work properly!)
3. Use `region_id` as a stratification variable. Draw a sample of size $n = 2000$, allocate the sample to the strata proportional to their size (look up the code in our previous Rmd file on Sampling).
4. Specify a `svydesign` object
5. Estimate the population share of No votes in the initiative.
6. Estimate the design effect, interpret the design effect in terms of the effective sample size.