

SAMPLING, WEIGHTING AND ESTIMATION

EXERCISE 1

Stefan Zins, Matthias Sand
and Jan-Philipp Kolb

GESIS - Leibniz Institute
for the Social Sciences

January 22, 2016

- 1 Download the ESS dataset for Sweden (Sampling Data and Country File) of the 5th round

<http://www.europeansocialsurvey.org/data/country.html?c=sweden>

- 2 Setup your workspace and load the R-packages `foreign` and `survey`
- 3 Load the ESS dataset and the country file
- 4 Merge both data frames by their ID-variable, using the `merge()`-command

- 5 Determine the sampling strategy (Inspect the variables `SAMPPOIN`, `STRATFY` and `PROB`)
- 6 Add the variable `N` to your data frame that equals the total population size for the target population of the ESS Round 5 of Sweden

Advice: `dweight * pweight * 10000 * n`

- 7 Convert your data frame to a survey object (`survey`-Package)
- 8 Estimate the total and mean of the variable `tvttot`

- The survey package provides a large range of applications for complex survey samples
- Typically, the first step is to define a survey object with the `svydesign()` command

Simple survey object (Simple Random Sample)

```
data(api)

surv.obj <- svydesign(id=~1, fpc = ~fpc, data = apisrs)
```

```
surv.obj <- svydesign(id=~1,fpc = ~fpc, data = apisrs)
```

- `id` specifies the identifier of PSU and SSU; `id= ~0` or `id=~1` stipulates a single stage sampling
 - For multi-stage samples the `id` argument should always specify a formula with the cluster-identifier at each stage
 - `fpc` should be used for the finite population correction
- ⇒ Either as the total population size of each stratum or as a fraction of the total population that has been sampled
- `data` reflects the data set for which the design object should be defined

Important commands

<code>svytotal</code>	returns the estimated total of a variable and its standard error (+ <i>deff</i>)
<code>svymean</code>	returns the estimated mean of a variable and its standard error (+ <i>deff</i>)
<code>svyquantile</code>	Computes quantiles for data from complex surveys
<code>svyvar</code>	Computes variances for data from complex surveys
<code>weights</code>	Returns the (design) weights of a survey object
<code>calibrate</code>	Calibration of a data set (uses the GREG-Estimator)
...	...

```
svytotal(~api00,surv.obj)
```

```
##           total      SE  
## api00 4066888 57293
```