

Class exercise 'Designing Weights'

Do the following exercise in pairs. The goal is to think as a class about how to design weighting models. For the purpose of this exercise, we will assume that we want to do a survey among the general population of the Netherlands. That is, it includes everyone living in the Netherlands on 1 January 2019, including babies and elderly people. The population register of the Netherlands is used as a sample frame, including information for individuals about:

- Their location (region, town, postal code)
- Age
- Gender
- Income of the household
- Born in the Netherlands or not

Imagine your survey is a general survey where all kinds of questions about behavior, attitudes and facts are asked. Some of the main outcome variables (Y) are listed below.

Question 1: Please pick just one of these outcome variables.

- Voting behavior in the parliamentary election if there was a election today
- Household spending per month (amount and type of products)
- Health issues (long-term illnesses)
- Computer literacy

Question 2: Imagine that the survey you are doing is a face-to-face survey, where interviewers approach every individuals at least 6 times on different days, and different times of the day. The budget for this survey is high, and the response rate is about 60%. Go to statline: <https://opendata.cbs.nl/statline/#/CBS/en/> and browse the data here. What variables would you like to include in your survey, so you can later weight for nonresponse?

Question 3: Now imagine, the survey actually has a very low budget. Paper invitations are sent, including a URL that people can type into their browser to get to the response. The response rate for this survey is about 5%.

Go to statline again. What other/different variables would you now like to include in your survey, so you can later weight for nonresponse?

Question 4 (if you have time left): Pick another outcome variable under question 1 and repeat the exercise