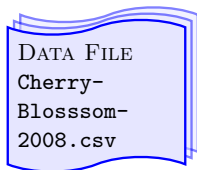


— be up-front and honest about it. In surveys, you should keep track of the non-response rate and include that in whatever report you make of your study.

Example 2.1: Struggling for a Random Sample Good researchers take great effort to secure a random sample. One evening I received a phone call at home from the state health department. They were conducting a survey of access to health care, in particular how often people have illnesses for which they don't get treatment. The person on the other end of the phone told me that they were dialing numbers randomly, checked to make sure that I live in the area of interest, and asked how many adults over age 18 live in the household. "Two," I replied, "Me and my wife." The researcher asked me to hold a minute while she generated a random number. Then the researcher asked to speak to my wife. "She isn't home right now, but I'll be happy to help you," I offered. No deal.

The sampling frame was adults over age 18 who live in a particular area. Once the researcher had made a random selection, as she did after asking how many adults are in my household, she wasn't going to accept any substitutes. It took three follow-up phone calls over a few days — at least that's how many I answered, who knows how many I wasn't home for — before the researcher was able to contact my wife. The researcher declined to interview me in order to avoid self-selection bias and worked hard to contact my wife — the randomly selected member of our household — in order to avoid non-response bias.

2.4 Longitudinal and Cross-Sectional Samples



Data are often collected to study the links between different traits. For example, the data in the following table are a small part of a larger data set of the speeds of runners in a ten-mile race held in Washington, D.C. in 2008. The variable **time** gives the time from the start gun to the finish line, in seconds.

Name	Age	Time	Sex
A. Renee Callahan	42	112.4	F
Aaron Alton	34	89.3	M
Cristine Doran	25	87.8	F
Ed Donahoe	58 D	77.3	M
Amy Lewis	30	94.0	F

... and so on.

Such data might be used to study the link between age and speed, for example to find out at what age people run the fastest and how much they slow down as they age beyond that.