



Random Samples: How and Why

What steps do I take to draw a random sample of a target audience in my county to do a needs assessment or gather baseline data? And why is a random sample so important?

There are two steps in the sampling process. The first is to compile a comprehensive and accurate list of your target audience; for example, farmers in the dairy business, childcare providers, agency personnel or 4-H leaders in your county. This list is called the *sampling frame* or universe. The second step is to *draw a random sample* from that list. By far, the more challenging and often neglected task is the first.

Step One: A Sampling Frame: Your goal in constructing a county list is to make sure the list is comprehensive. For a bio-security needs assessment it should include every dairy farmer and for a child safety needs assessment, every childcare provider. The list should include all members of the target audience whether they are currently extension cooperators or not. If the list has other related people such as sanitarians or agribusinesses in dairy, or in the case of childcare providers, children's supply companies, you must be able to extract just the names of the dairy farmers and childcare providers that are currently in business.

Why does the list need to be comprehensive?

- Research often draws upon health and accident rates to demonstrate that problems exist, but the extent to which the target audience practices *particular* behaviors that contribute to say, poor bio-security on farms or child accidents, is *not* known from such general data. Thus, it is not known which behaviors to emphasize in instructional videos, slides, printed materials and educational meetings. Therefore, it is important to include a complete list of the target audience in your sampling frame in order to know the widest distribution of all the management practices that relate to the problem.
- Conducting a needs assessment or collecting baseline data to find the distribution of these practices and attitudes toward them will have an immediate and direct bearing on the content of the program.
- There are long-term effects of collecting such data. Having an up-to-date list of dairy farmers or childcare providers in a county means that a county agent will have a complete list of the principal target audience for marketing and evaluating the educational program.
- Sometimes a seminal list may exist at the state level but you might have to add to it. For dairy farmers, begin with the Milk Ring list, a list of all farmers shipping milk. You can obtain it for your county by calling the Bureau of Animal Industries at 717/783-5301. Ask for the Milk Ring results. This list is updated every three months. Next, compare this list with those of co-ops in the area. Check with farm realtors for farmers who have recently sold their business and for the names of the new owners. Check with other appropriate sources.
- At other times, such as compiling a list of daycare providers, you may have to start at the county level. Begin with local chapters of professional organizations (NAEYC) or support groups. Glean from phone books for centers. Be sure to include providers who have already participated in recent extension trainings.

The list you compile does not have to be in any order.



Step Two: Drawing a Sample: Since it is impossible to interview everyone in your target audience, methods have been established to select a fraction to accurately represent the whole. That fraction is known as the “sample.” What is important from a scientific point of view is to make certain everyone in the sampling frame gets a known chance of being selected. That means two things: no particular persons on the list should be selected, but they should not be selected “hit or miss” either.

There are four easy steps in drawing a sample:

1. Decide how many you need in the sample. Start with the number of people you want to interview (for example, 15), add to that number 5-7 people for outright rejection and another 5-8 people to cover a variety of reasons who withdraw at some point in the process. Thus, the sample here stands at 30. It is better to have more names in your sample than too few because you don't want to have to draw another sample. If, for some reason, you think individuals not in your target population have slipped into your list, add another 5-6 people or possibly more. Avoid this problem at all costs by constructing the type of sampling frame described above.
2. In order to find out which names to select from the list, do two things. First, divide the sample number into the total number of dairy farmers or childcare providers in your county.

Example 1.	
30 sample	$\frac{2.0}{60 \text{ list}}$
Example 2.	
30 sample	$\frac{4.6}{138}$



The first example tells you to select every 2nd number on your list.
The other example tells you to select every 4th number on your list.

3. Select names, but don't begin to select names from the top of the list. You want *to start* in a random place. To find out where on the list to start, put numbers in a container from 1 to whatever the number is in Step 2 above. For example 1, put #1 and #2 in a container. For example 2, put #s 1, 2, 3 and 4 in a container. Pull out a number. This is the place on the list you *start* to select the names in your sample. The next names on the list will be every nth (Step 2) name after your first selection.
4. When you finish selecting names, you should have a sample. The number will be somewhat larger than the number of interviews you need to complete. To set up dates for the interviews, you can begin with any name in the sample since they were all randomly chosen.

Nancy Ellen Kiernan, Ph.D., Program Evaluator, nekiernan@psu.edu

The reference citation for this Tipsheet is: Kiernan, Nancy Ellen (2001). Random Samples: How and Why: Tipsheet #57, University Park, PA: Penn State Cooperative Extension. Available at: <http://www.extension.psu.edu/evaluation/pdf/TS57.pdf>

This Web site is copyrighted by The Pennsylvania State University. The information may be used for educational purposes but not sold for profit.