

# WHAT'S IN A FRAME?

by Terry Holland

When I first started working for the Agency and before I had taken a sampling course, I kept hearing people talk about "the list frame" and "the area frame." "What are these things called frames?" I wondered. I imagined two gilded picture frames hanging on a wall - in one, a great list of farmers in tiny print; in the other, a huge detailed map of the United States. How they drew samples from these things I had no idea! Obviously I had the concept wrong, so I decided to do some digging. I started with Webster's dictionary:

frame (n) 1 a: something composed of parts fitting together and united ...  
2 a: the constructional system that gives shape or strength (as to a building) ...  
3 c: a structural unit in an automobile chassis supported on the axles and supporting the rest of the chassis and the body ...

None of the definitions seemed to fit, so I tried another dictionary. The same definitions were given: window frames, eyeglass frames, house frames, skeletons, baseball innings, bowling frames, film exposures, frame of mind, frame of reference, but no sampling frames! I tried a statistical dictionary. No luck! There was no clear definition of a frame. What should I do now?

With more digging, I finally discovered that the concept of the frame is relatively new. First stated by Frederick F. Stephan in 1936, the "frame" concept was adopted by the United Nations Subcommittee on Statistical Sampling in 1948. According to W. Edwards Deming, a frame is "a set of physical materials (census statistics, maps, lists, directories, records) that enables us to take hold of the universe piece by piece." A universe is simply "all the people or firms or material, conditions, concentrations, models, levels, etc., that one wishes to learn about ..."

More simply, a frame is a means by which a universe may be sampled. To select a sample, the universe must be divided into parts, called sampling units. The frame is a list or map of all sampling units, or a set of rules for identifying sample units.

When I was in Oregon, I worked on a Filbert Objective Yield Survey. We used a list of all filbert orchards in two States, identified by owner, operator, location, size, age, and variety. The list was updated every 4 years by a census of bearing and nonbearing orchards. A sample of orchards was selected and located. According to specific instructions, enumerators then selected three trees in each orchard to be in the sample. In this case, the universe was all bearing trees about which we wanted to know how many nuts they would produce. The frame was a list of orchards with a set of rules for further identifying trees.

This example illustrates the requirements of a frame. First, a frame must have enough information to define and locate any sampling unit that may be selected. If the sampling unit cannot be defined or found, there can be no frame. Even if the owner and operator of each

orchard changed during the 4 years between updates, the physical location of each orchard and tree did not. Second, the frame should cover the entire universe. If the frame covers only part of the universe, the results of the survey will apply only to that portion of the universe included in the frame. Since it takes 4 years for new filbert trees to come into production and the list was updated every 4 years, the frame meets this requirement. If instead the tree inventory was taken every 5 years, then every fifth year the results of the survey would reflect production from orchards that had been bearing for more than a year, not production from all bearing trees. Third, each element of the universe must be included in one and only one sampling unit. Obviously, each tree was found in one and only one orchard.

Over the years, NASS has developed a list frame and an area frame (those things I used to wonder about) to support many different survey needs. The list frame is simply a list of farm operators and agribusinesses. We maintain names, addresses, telephone numbers, and other items that enable us to locate the farm or business, if sampled. We also maintain control data that allow us to refine the frame for specific survey designs. Because the person or operation reporting may be different from the operator or business sampled, we have developed a set of rules to associate reported information uniquely with the name selected.

The list frame has a major problem - it does not cover the entire U.S. farm population! Sources of names for small and specialty farms are difficult to obtain, making construction of a complete list of farms nearly impossible. Even if a complete list could be built, it would be out-of-date before it could be used. About two-thirds of all U.S. farms are currently available for sampling from the list. Fortunately, we can usually compensate for this undercoverage by using samples from the area frame to supplement list samples.

The area frame is simply the entire U.S. land area divided into segments. Unique and identifiable boundaries of sampled segments are outlined on maps and photos to enable us to locate the segment. We have developed rules that allow us to associate farm operations, grain storage facilities, crops, animals, and other agricultural commodities with land inside each segment. Every possible item of interest can be uniquely associated with some segment of land, so we can say the area frame entirely covers any possible universe. Samples from the area frame may not estimate well for rare and specialty commodities, so it is important that all operations producing these types of items be maintained on the list.

You know, if you really think about it, Webster's definitions of frame aren't too far off the mark! A sampling frame is composed of parts (sampling units) that fit together and are united. In the constructional sense, NASS sampling frames give shape and strength to our sample designs. In the automotive sense, the list and area frames support the rest of the NASS survey and estimating programs (chassis and body), but without adequate coverage and rules of association (axles) the programs are of limited value.