

1) Suppose that we want to find the average weight of cows in a region. A list of all farms is available for the region, and 50 farms are selected at random. Then the weight of each cow at the 50 selected farms is recorded.

Element or Observation Unit:

*Cow*

Measurement on the element:

*Weight*

Numerical quantity of interest:

*Mean weight*

Target population:

*All cows in the region*

Sampling Unit:

*Farm*

Sampling Frame:

*List of Farms*

Variable of interest:

*Weight of all cows in each sampling unit*

Parameter of interest:

*True mean weight of cows*

Sampled population: (Population that the sample is taken from.)

*All farms in the list*

The parameter that we would be interested in is the population or true mean cow weight for all the cows in the region. This would be easily estimated if we could randomly sample from a list of all the cows in the region and weigh them (measurement=weight, element=cow) and calculate the average of those measurements (numerical quantity of interest). The target population is all the cows in the region.

Unfortunately, the situation was not that simple and other aspects needed to be considered. Sampling had to occur at the level of the farms (the sampling units), which were randomly selected from a list of all the farms in the region (the sampling frame and also the sampled population since all farms were available for sampling). The variable of interest is the measurement on the sampling unit, which is best considered to be the weight of all the cows in each farm. We would also want to record how many cows were measured at each farm. Just recording the average at each farm might allow you to estimate an overall mean, but with just the average information, you would not be able to weight the estimate by the number of cows in each farm. Note also that the sample is 50 farms that information was collected about.