interview procedures usually require more than one visit to each household.[191] For example if a diary is kept, there will be an initial visit, a second visit to check that it is being kept correctly, and a collection visit after seven or fourteen days. In rural areas, where transport is a major cost of the survey, it therefore makes sense for the survey to group households into survey "clusters", often villages, with typically six to fifteen households in each cluster. The optimal number of households in the cluster involves a trade-off between the low marginal cost of drawing another household in a village that is already being visited, and the relatively low contribution to precision of such a household, given that it is likely to look rather similar to other households in the same village. The survey team remains in the village for a week or two, surveying all households in the cluster, and then moves on to a new cluster. Such surveys frequently attempt to give each household an equal chance of inclusion by using a two stage design, in which clusters are selected first, with a probability of inclusion proportional to size (i.e. the number of households in the cluster), while individual households are randomly selected at the second stage. The random selection of clusters and households is made from a "sampling frame", often a census. However, censuses are often badly outdated, and in some countries are not reliable, either because of political interference — census returns are typically used to make voter rolls and sometimes to allocate resources or because of difficulties of collection. Problems with censuses can be avoided as in India, where the National Sample Survey (NSS) selects villages from a village frame, and then lists all households in the village at a preliminary stage. The final drawing selects a stratified random sample from their own list, with stratification based on a few variables collected at the list stage, [see Murthy (1977, Chapter 15)] who also describes many of the other features of the design of the NSS, or (more generally) Casley and Lury (1981) for further description.

   The relatively simple — and sensible — designs of the previous paragraphs can be complicated ad infinitum. Adjustments can be made for non-response, and for the consequences of replacing non-responders by "look-alike" households, although it should be noted that, unless households are approached at obviously inappropriate times, like harvests, non-response is typically much less of a problem in LDCs than in the United States (US). Probabilities of selection can also be linked to any ancillary information in the sampling frame or listing, such as occupation, housing status, or landholdings. As a result, survey tapes usually report for each household a sampling probability, or its reciprocal, an "inflation factor", which is the number of households in the country for which the household stands proxy. In complex designs, the inflation factors will be different for every household in the survey. Although designs are often "self-weighting", whereby in spite of the many strata and levels of strata each household has an equal chance of being included, refusals or other