WEIGHTING

What is Weighting?

Weighting is the adjusting of response data to ensure that your results reflect the population from which your sample was selected. This is often used where one or more particular groups have been over or under-sampled (i.e. – an age group) during a consultation.

When should Weighting be used?

Weighting should be carried out when you want the responses of your survey to reflect the whole population of your consultation area but the responses you have received do not allow you to do this, i.e. if a certain age group or ethnic group are under or over represented amongst your respondents.

Why Should Data be Weighted?

If data requiring weighting are not weighted the resulting estimates will be biased if they are interpreted as estimates for the wider population (as opposed to estimates relating to the achieved sample).

Types of Weight

1. Sample design or probability weights

Sample design or probability weights correct for cases having unequal probabilities of selection that result from sample design. It is important to note that non-equal selection probabilities can also occur due to differentials in non-response, which is corrected by non-response weights described below.

To illustrate how a sample design weight is calculated, consider a survey design that interviews one dwelling per address, one household per dwelling and one adult per household. Provided information concerning

dwellings per address, households per dwelling and adults per household is enumerated by the interviewer, one can subsequently calculate sample design weights that correct for the lower selection probabilities of adults in multi-adult (and household/dwelling) households. The general formula for a sample design weight is arithmetically very simple, it is 1 divided by the probability of selection due to the survey design. However, these are usually scaled, so we define the weight as proportional to this number. For example, if there are 3 adults in a given household the resulting sample design weight for the single interviewed adult will be proportional to 1/(1/3), i.e. proportional to 3. In a one adult household, the weight will be simple proportional to 1/1, i.e. proportional to one. In other words the influence of the former respondent is being increased threefold relative to the influence of the latter respondent to exactly compensate for the fact the former respondent was three times less likely to be included in the sample.

The weights are often scaled to have a mean of 1, which maintains an effective sample size when the data are weighted.

2. Non-Response Weights

Non-response weights compensate for differential response rates. Response rate in this sense refers to unit non-response, whereby someone refuses to take part in the survey at all, as opposed to item non-response, which relates to refusing to answer specific questions, which is addressed by missing data methods rather than weighting.

Non-response weights are typically obtained by defining weighting classes, which are based on information available for both responding and non-responding households. Such information typically relates to geographical location, primary sampling unit (PSU) characteristics (which are derived from other data sources, often the Census) and often household and dwelling type (which need to be recorded by the interviewer).

Respondents in each weighting class are weighted to compensate for the proportion of non-respondents in that class. More formally, the non-response rate weight is proportional to 1 divided by the response rate for the weighting class, i.e. directly analogous to sample design weights.

Sample design weights usually control exactly for differences in selection probability due to sample design, but non-response weights are seldom entirely accurate.

The utility of the non-response weights is governed by the amount of information available to define the weighting classes. By definition, information about non-respondents is limited. The assumption of non-response weights is that the characteristics of respondents and non-respondents within each weighting class are the same; only if they are will the non-response weight be entirely accurate. If you as an analyst are examining characteristics that do vary between respondents and non-respondents within weighting classes, then the weighted estimates you derive will be biased population estimates. This problem is most likely to occur when examining measures of social engagement such as voting/not voting, which are likely to be highly correlated (even within a weighting class) with whether a respondent agrees to take part in a survey or not.

3. Post-Stratification Weights

Post-stratification weights (also known as population or calibration weights) are constructed after the other types of weights have been constructed and applied to the data. They are applied to make the data even more representative of the population. As for probability weights, information on the population is usually derived from the decennial Census of Population.

These weights allow for more accurate population totals of estimates, they reduce non-response bias further (over and above non-response weights), and improve precision

Whereas sample design (probability) and non-response weights result from a very simple computation (1/selection probability), post-stratification weights are mathematically complex, requiring iterative algorithms that maximise the fit of the data to the population. This procedure is called 'raking', and requires specialist software. The Office for National Statistics currently use a SAS based macro called CALMAR to calculate post-stratification weights, but are switching to the Generalized Estimation System (GES) programme.

For example, calculation of the Labour Force Survey individual level post-stratification weights (see: Barton 2004) involves 'raking' to 3 controls (derived from the Census and population projections):

5-year age group by sex within region
Local Authority
Single years 16-24 by sex
Population projections

The raking procedure iterates until the data best match all three controls, and computes the post-stratification weight accordingly.

The above information on types of weighting is from the ESDS Government Service.