

Tips on sampling frame design

In designing a sampling frame for a survey (ie. an approach to sample selection), it is important to consider:

- The target groups of interest – to ensure that there is sufficient data to permit reliable analysis of data following data collection, the sampling frame should include all groups or segments of interest
- The use of random sampling – random sampling is the best technique to avoid sampling bias – if there are groups of interest, it is also possible to sample randomly within lists of interest (called stratified sampling) – eg. having a list of accountants and randomly selecting businesses from the list for interview
- The volume of sample required – when people are surveyed, it is important to ensure that a sufficiently large sample is included to ensure that the selected sample provides an accurate estimate of the entire “population” (ie. All the respondents of interest) – if too few respondents are interviewed, “random sampling error” may be high and imply a more biased estimate
- The following table may be used to work out appropriate sample sizes for surveys. While many other more complex factors can influence sampling design/confidence intervals, the table provides a “ballpark measure” of sample confidence intervals achieved for different sample sizes. This assumes an arbitrary population size of 5,000. It should be noted that population sizes in excess of thousands will not have a large effect on sample sizes required. So the following can be considered as indicative for population sizes even of 100,000 or more. This implies that there is a 95% chance that the given result is within +/- X% (confidence interval) of the obtained result. Practically, this means that if a result of 50% was obtained, a 5% confidence interval means the real result for the population is likely to be between 45% and 55%.

Confidence intervals for different sized samples assuming a population size of thousands of people (assuming 95% confidence level)				
N=94	N=357	N=536	N=880	N=1622
+/- 10%	+/- 5%	+/- 4%	+/- 3%	+/- 2%

- Working out confidence intervals can also assist with research budgeting – for example, if there is only a small improvement in sampling confidence intervals, it is arguably not worth spending additional budget on a larger sample – in this respect, confidence intervals can be useful for understanding when or when not to spend money on a larger budget
- If there are important sub-groups in the population of interest, remember that the above sample sizes should also be applied to the segments of interest – for instance, for roughly +/-5% confidence, this may include designing a sample of 800 accountants including 400 involved in consumer tax returns and 400 involved in business tax returns (this will help to ensure that there is reasonable confidence levels each at the subgroup and overall population level)