

## Treatment of Missing Values in Survey Analysis

Missing data in survey research are a reality and faced by researchers and analysts often. It may occur in the following situations –

- Complete nonresponse by a sample element i.e. when a respondent did not participate in the survey at all
- Item nonresponse by a responding sampled element i.e. when a participating respondent did not provide answers to some of the survey items or questions
- Often, in surveys designed in Likert scale where only one response can be chosen, cases of multiple responses for a single question can also be represented as invalid or missing data

Ignoring missing values leads to data loss. The simplest method to handle missing data is to omit the cases with missing values. This is called listwise deletion. If out of 100 survey respondents, there are 7 cases that had missing data, under listwise deletion we delete those 7 cases and take the rest of 93 complete cases for further analysis.

If we delete cases listwise, the sample size gets reduced. Therefore we need to check whether the number arrived at after listwise deletion is sufficient or not. Say, our target for sample is  $N$ . The filled-in survey size is  $N_1$  where  $N_1 > N$ . Normally, to handle missing data cases, we typically roll out the survey to a larger sample where  $N_1 > N$ . After deleting the missing data cases, the size becomes, say,  $N_2$ . If  $N_2 > N$ , then listwise deletion does not really reduce the sample size as we have adequate size to go ahead with analysis.

However, there is an important assumption here – data is missing completely at random (MCAR). This means that any piece of data is as likely to be missing as any other piece of data. This is ensured in the sampling plan by the way the sample is drawn.

In situations when sample size is getting reduced due to listwise deletion, we may have to go for replacement or substitution methods. We shall discuss these methods, and their advantages and disadvantages in future posts. However, it would be pertinent to note here that in some cases we are better off to go for listwise deletion than to go for substitution methods in order to keep the analysis and interpretations non-complicated and comprehensible.