

Missing Data

In the process of cleaning your data, you may have found that you have missing information. If this occurs, you have options.

The **Two** Most Common Ways of Dealing with Missing Data

Listwise Deletion (Complete Case Analysis)

In **listwise** deletion, the entire “case” (an entire participant’s data) containing the missing data is deleted from that particular analysis (even the data that is not missing for that participant). For example, if you have a participant who did not give you their age, all of that participant’s data will be excluded from any analysis of your dataset that includes age.

Pairwise Deletion (Available Case Analysis)

In **pairwise** deletion, only the missing data is excluded from the analysis, but the rest of the participants’ data is included. For example, if a participant’s age is missing from the dataset, other variables from that participant will still be used in analyses, even analyses including age as a variable.

Additional Way to Handle Data

Data Imputation

Listwise and pairwise deletion are ways of handling missing data if there is little data missing and you are sure that there is not a systematic reason underlying the missing data. Data imputation is a more complex way of handling missing data. This involves using information about the distribution of your data to estimate missing values. With this, you should consult a statistician about when to use this and which particular method is most sensible given the characteristics of your data.

(At Westcliff, you can contact a methodologist through your program for further support on complex statistical matters)

**Please note that the above processes must be done on clean data, as opposed to raw data.
Clean data is data that has been processed already.**