

2.1: Missing Values

Any large collection of data is probably incomplete. That is, it is likely that there will be cells without values in your data table. These missing values may be the result of an error, such as the experimenter simply forgetting to fill in a particular entry. They also could be missing because that particular system configuration did not have that parameter available. For example, not every processor tested in our example data had an L2 cache. Fortunately, R is designed to gracefully handle missing values. R uses the notation NA to indicate that the corresponding value is not available.

Most of the functions in R have been written to appropriately ignore NA values and still compute the desired result. Sometimes, however, you must explicitly tell the function to ignore the NA values. For example, calling the `mean()` function with an input vector that contains NA values causes it to return NA as the result. To compute the mean of the input vector while ignoring the NA values, you must explicitly tell the function to remove the NA values using `mean(x, na.rm=TRUE)`.

This page titled [2.1: Missing Values](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [David Lilja](#) ([University of Minnesota Libraries Publishing](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.