

13.3: Data and Assumptions

Each statistical test has some assumptions which must be met in order for the formula to function properly. In keeping, there are a few assumptions about the data which must be met before a simple linear regression is used. Note that these are the same assumptions that must be met for bivariate correlation. First, the scores for both variables must be matched. This means that the same participants have scores for each variable and that the two scores for each participant can be identified together. Second, data for each quantitative variable should be fairly normally distributed without notable impact due to either univariate or bivariate outliers. A bivariate outlier refers to a participant whose scores on X and Y together do not follow the general pattern of the other participants; if their data diverge markedly from the general pattern, they are considered a bivariate outlier. Finally, the relationship between the variables should be linear meaning they approximate a straight line rather than other shapes such as would be seen when graphing a quadratic equation, for example. See Chapter 12 for a review of how to assess a linear pattern visually using scatterplots.

This page titled [13.3: Data and Assumptions](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by .