

12.3: 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 bivariate correlation is used. First, the scores for both variables must be matched. This means that participants have scores for each variable which are identified together. Second, data for each quantitative variable should be fairly normally distributed without notable impact due to outliers, including bivariate outliers. A bivariate outlier refers to a participant whose scores on X and Y together do not follow the pattern of the other participants. In order to be identified as an outlier, this divergence should be fairly dramatic and rare relative to the rest of the data set rather than mildly divergent or common. Finally, the relationship between the variables should be linear meaning they approximate a line rather than other shapes such as what would be seen when graphing a quadratic equation. We will focus on how to look for the shape visually to see if it looks linear later in this chapter.

This page titled [12.3: Data and Assumptions](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by .