**Week 4 Discussion**

*Why do you typically need to change data from wide to long when working with data in R or other programming languages? Your response should be one paragraph and you need to respond to three other student’s posts.*

Typically, one needs to change data from wide to long when working with data in R or other programming languages for a few main reasons. The first reason is because having data in long format makes graphing and statistical modeling easier [1]. For example, if I have a wide data set showing time series data and each day’s data is presented in a different column, modeling this in the form of a graph would be very difficult because the days are spread out across individual columns. In R, the use of a function such as ggplot would be nearly impossible because x and y aesthetics would be spread numerous columns. The below representation of wide data illustrates what I have described above.

id gender race trt day1 day2 day3

1 F 0 0 19.81310 18.05777 14.84996

2 M 0 0 17.91846 18.75825 15.30547

3 M 0 0 17.22526 19.79218 15.10622

However, if “day” is presented as a single column, and the various days are listed underneath this variable in a repeating fashion, it is much easier to represent the data in a visual format using a function such as ggplot. The same argument can be made for statistical modeling. If “day” is assigned to a single variable, it is much easier to manipulate all data within “day” as opposed to attempting to apply data transformations and modeling techniques across multiple columns. The transformed long dataset below again illustrates this point.

id gender race trt day amt

1 F 0 0 day1 19.81310

1 F 0 0 day2 18.05777

1 F 0 0 day3 14.84996

2 M 0 0 day1 17.91846

2 M 0 0 day2 18.75825

2 M 0 0 day3 15.30547

3 M 0 0 day1 17.22526

3 M 0 0 day2 19.79218

3 M 0 0 day3 15.10622

For these reasons, the conversion of data from wide to long is typically required. One additional note is that data can be transformed back to wide after the application of analysis and visualization has been completed to the long format for purposes of presentation and reporting.

Resources

1. *Reshaping data from wide to long*. Reshaping Data from Wide to Long | UVA Library. (n.d.). https://library.virginia.edu/data/articles/reshaping-data-from-wide-to-long