

Data Carpentry Learning Journal

Jeremy Ami

August 2019

1 Messy Data Journal

For the Mozambique table:

- The variables are ambiguously filled. It is not easy to know what exactly is being recorded in each table
- There is a colored cell with a number in it. It is vague whether or not the number in the cell refers to the variable plus the yellow highlight or if it is not inclusive
- There are three different tables, two of them are stacked. This is confusing and would be done better if there is one single table for all of the variables spread across each row
- It would be simpler, albeit a bit more tedious, if each variable was binary in nature. For example rather than having the names of the variable in each cell, a binary input of Y/N or 1/0 would make the table uniform and have no room for ambiguity
- There is a -99 in one of the tables. This is clearly a formatting and input error, and one which the binary input option I proposed would avoid

For the Tanzania Sheet

- There are blank cells. Computers do not like blank cells. A 'NULL' or 'Blank' input would be preferable.
- There is meta-data in this spreadsheet. This not only makes the sheet look messy but it can also confuse the computer when it comes to converting this to other kinds of files. Better to have the meta-data on the outside of this document as a legend or some other kind of external hermeneutic file

2 Dates as Data

I understand what is required to convert the cells with DD/MM/YYYY into cells with a single value for DD or MM or YYYY. Using the example from the Data Carpentry website, I inputted '41822' as a test and it automatically converted into a date '02/07/2014 '. When I then typed the formula '=DAY()' and selected the date it successfully converted to a 2. When I did the same thing for the already existing dates however it read as an error. 'VALUE!' came up as the input to the cell rather than the day. This may be because I am using LibreOffice as opposed to Excel for this function.

3 Quality assurance

Intention: Do the Quality Assurance lesson in Data Carpentry

Action: Open the tab for the lesson. Open the clean version of the SAFI dataset.

Error: I am using LibreOffice so I could not follow the instructions on in the lesson

4 Note

I did not successfully complete all of the Data Carpentry lessons for Data Organization in Spreadsheets for Social Scientists. I shall do so over the weekend.