



## How to merge two data files when both files contain only one record per participant, using SPSS

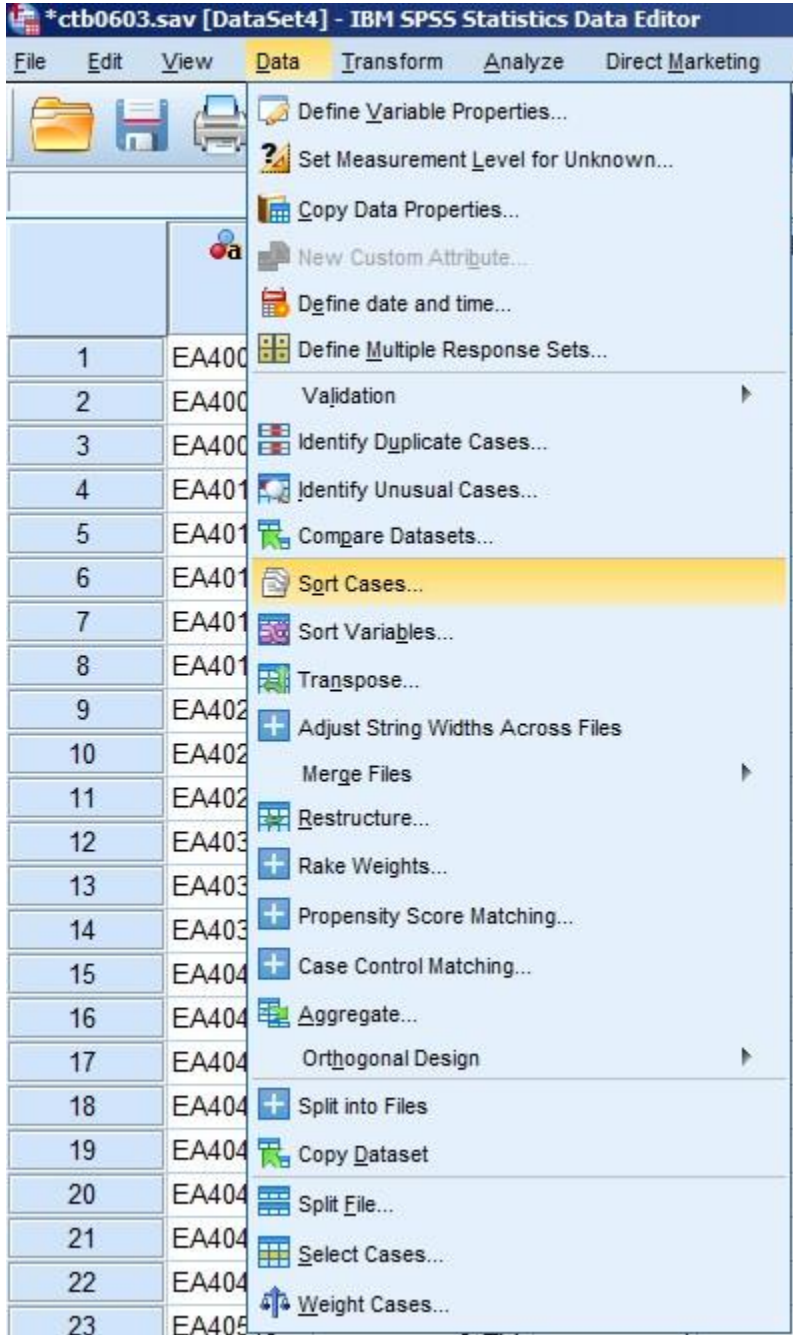
Based on SPSS version 24

This document is intended to assist data users with merging two data files which each contain only one record per participant. This file format is also known as a “wide” data file. This document will show data users how to navigate the SPSS menus to conduct a merge and also how to use syntax to conduct the merge.

Launch SPSS and open both of your two data files of interest that each have only one record per participant. For this instruction sheet, we will merge the Conflict Tactics Scale data files from the LONGSCAN dataset: Partner to respondent item level data (CTB0603.sav) with the scored data for the same measure (CTS0603.sav).

You must sort each data file by the key IDS which in the case of LONGSCAN is the LONGSCAN ID variable called "ID." Click on "Data" from the menu bar at the top of the main SPSS window and select "Sort cases..." from the Data menu options, as shown in Figure 1.

Figure 1. Screenshot of the "Data" menu options showing "Sort Cases..." highlighted.



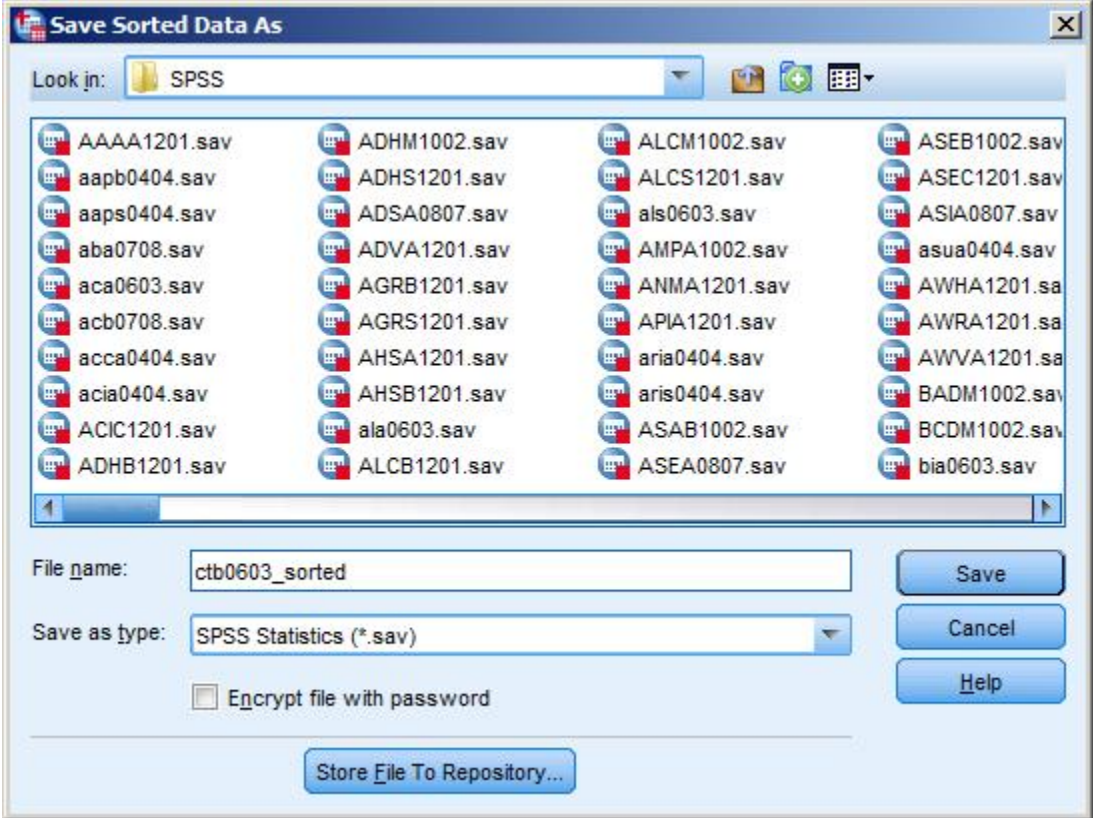
In the “Sort Cases” window that appears, select the “LONGSCAN SUBJECT ID (ID)” variable from the list on the left of the window and use the arrow button to click the variable into the “Sort by” box on the right side of the window. Be sure that the “Ascending” radio button is active to sort the file in ascending order by ID. Click in the checkbox next to the words “Save file with sorted data” as shown in Figure 2 to save the resulting sorted data file. Click on the “File...” button beneath the “Save file with sorted data” text.

Figure 2. Screenshot of the “Sort Cases” window.



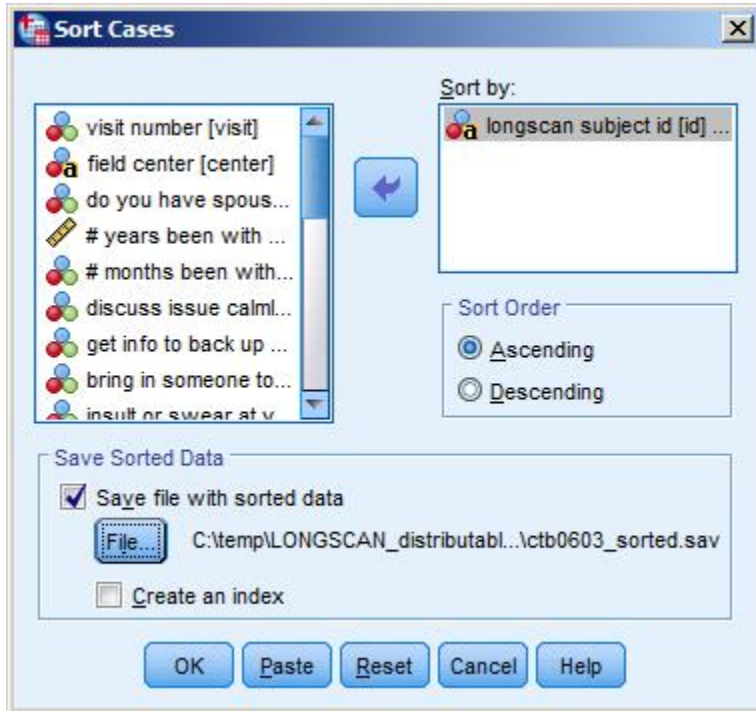
Once the “Save Sorted Data As” window is open, specify a file location and name to save the sorted data file (see Figure 3). Click the “Save” button on the lower right portion of the window to save the sorted data file once you are finish.

Figure 3. Screenshot of the “Save Sorted Data As” window.



After clicking “Save” in the “Save Sorted Data As” window, the “Sort Cases” window will reappear as shown in Figure 4. If the information appearing in the window looks correct, then click “OK” to sort the cases and save the file as specified.

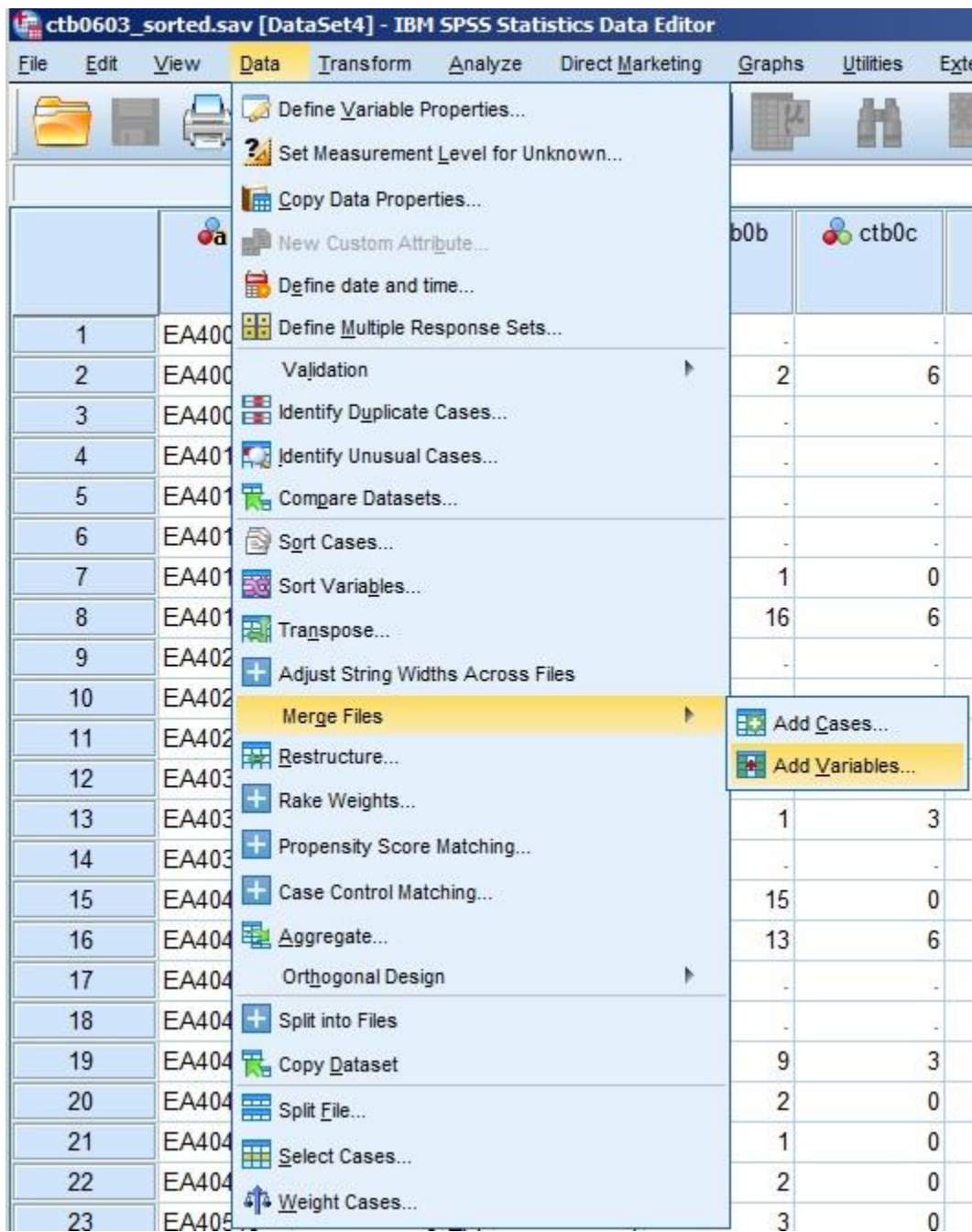
Figure 4. Screenshot of the “Sort Cases” window.



Repeat all instructions from the beginning of the document through this point for the other wide format data file you want to use in the merge. In the case of the example being used for this instruction sheet, we would repeat the steps for the CTS0603.sav data file. After both files have been sorted by ID and saved, proceed to the next step which is the merge. If you want variables from the item level data to appear first in the merged data file, be sure to open the item level sort data file and conduct the next steps from inside of that SPSS window.

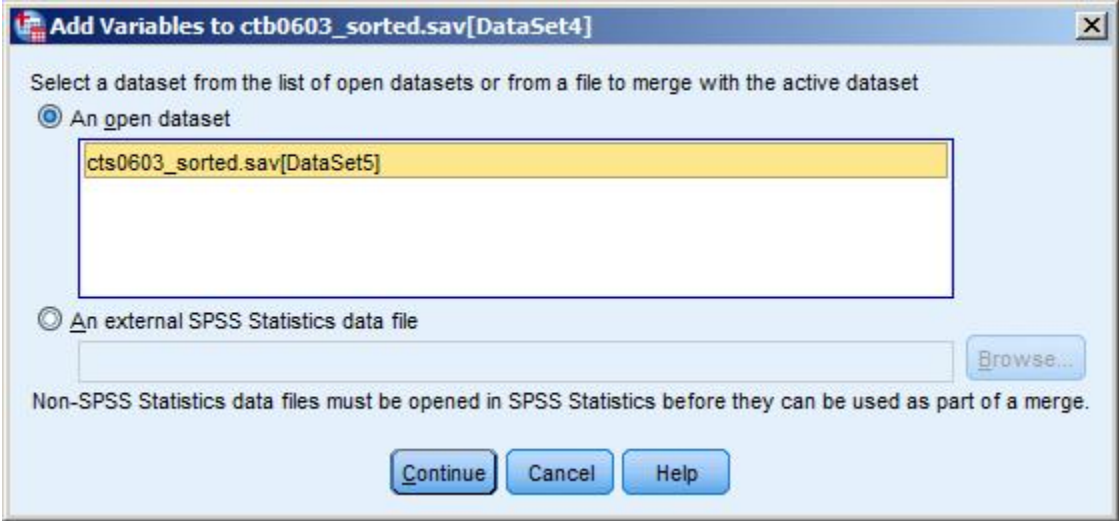
Click on the “Data” menu in the SPSS main window and select “Merge Files” from the list of options in the Data menu and then select “Add variables...” Figure 5 shows the “Data” menu list with the “Merge Files” option selected. We are adding variables because we know two things about these data files: First, they are related and the only way to produce scores in the scored file the participant must have a record in the item level data, so records will not be added during the merge. Second, each of the files only contain one time point of data collection, which is Visit 8.

Figure 5. Screenshot of the “Data” menu list showing “Merge Files” highlighted.



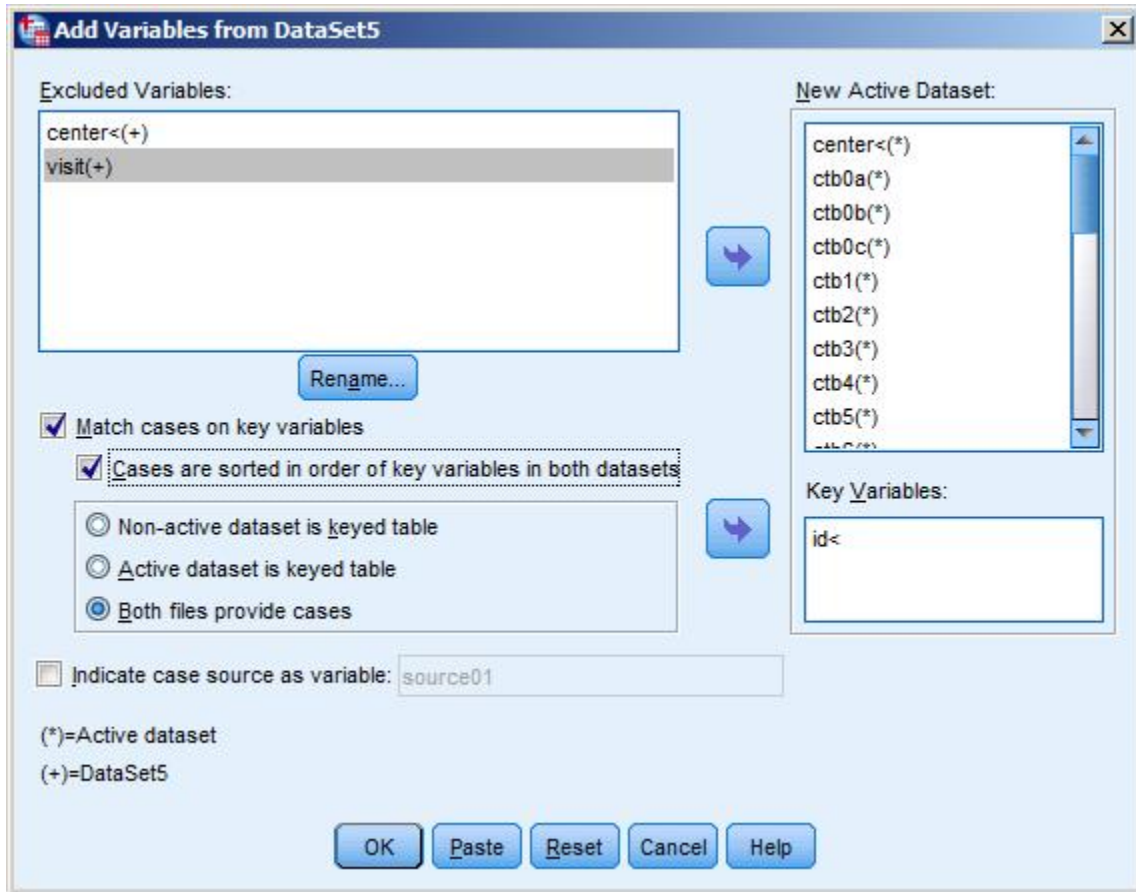
If both data files are already open, then you can activate the radio button in the “Add Variables to ctb\_sorted.sav” (or insert your item-level data file here) which says “An open dataset.” The open scored data file should appear in the list inside of the box beneath the “An open dataset” text. Click on the name of the data file to select it. Be sure that it is your sorted data file. Click the “Continue” button along the bottom of the window (see Figure 6).

Figure 6. Screenshot of the “Add Variables to ctb0603\_sorted.sav” window.



The next window that will open is called the “Add Variables from Datasets” window (see Figure 7). Check mark the “Match cases on key variables” check box. Check mark the “Cases are sorted in order of key variables in both datasets” box. From the “Excluded Variables” list, select the “id” variable and then use the arrow button to the left of the “Key Variables” box to click “id” into that box. We can leave “Center” and “Visit” variables in the “Excluded Variables” box since those variables are duplicated in the other data file. Click the “Ok” button.

Figure 7. Screenshot of the “add Variables from Datasets” window.



Once you click the “Ok” button, a warning will pop up like the one below in Figure 8. If you are certain that you sorted both data files in the same order by the key id, then click the “Ok” button.

Figure 8. Screenshot of the warning window.





After clicking the “Ok” button on the warning window, your files will be merged and the merged dataset will be open in front of you. Be sure to examine the data to make sure that the merge was performed correctly. One quick way to be sure that your merge was performed correctly is to add the count of variables from each data file together and then subtract the duplicate variables that were in the “excluded variables” section of Figure 7 which was the “Add Variables from Datasets” window and also subtract one Keyid variable, since that was used to match and therefore only one copy of it is included in the resulting merged data file.

Data users who are comfortable using syntax to perform statistical procedures can modify and use the syntax below to conduct the merge.

\*Open and sort the ctb0603 data file in ascending order by the ID variable.

```
GET FILE = 'C:\temp\LONGSCAN_distributable\Data\SPSS\ctb0603.sav'.  
EXECUTE.  
DATASET NAME ctb.  
EXECUTE.  
DATASET ACTIVATE ctb.  
EXECUTE.  
SORT CASES BY id (A).  
EXECUTE.  
SAVE OUTFILE= 'C:\temp\LONGSCAN_distributable\Data\SPSS\ctb0603_sorted.sav'.  
EXECUTE.
```

\*Open and sort the cts0603 data file in ascending order by the ID variable.

```
GET FILE = 'C:\temp\LONGSCAN_distributable\Data\SPSS\cts0603.sav'.  
EXECUTE.  
DATASET NAME cts.  
EXECUTE.  
DATASET ACTIVATE cts.  
EXECUTE.  
SORT CASES BY id (A).  
EXECUTE.  
SAVE OUTFILE= 'C:\temp\LONGSCAN_distributable\Data\SPSS\cts0603_sorted.sav'.  
EXECUTE.
```

\*perform the merge.

DATASET ACTIVATE ctb.

MATCH FILES FILE= 'C:\temp\LONGSCAN\_distributable\Data\SPSS\ctb0603\_sorted.sav'

/FILE= 'C:\temp\LONGSCAN\_distributable\Data\SPSS\cts0603\_sorted.sav'

/BY id.

EXECUTE.

SAVE OUTFILE = 'C:\temp\LONGSCAN\_distributable\Data\SPSS\ctb-cts-merged.sav'.

execute.

This concludes the instructions for merging two data files using SPSS. Please send questions to [NDACANsupport@cornell.edu](mailto:NDACANsupport@cornell.edu).