Merging Geo and Segment Records for Census 2000 Summary Files Using PC Based SPSS

Naming Conventions

The suffix for all the SPSS data definition files is .sps. Since all the summary files will use the same format for the geographic header file one will work for all. This file is named cen2k.geo.sps. For all the other files/segments the name starts with cen2k followed by pl for the PL94-171 files, sf1 for summary file 1, sf2 for summary file 2, etc. This is followed by a period, an s followed by the segment number. I've used as many digits as would be required for the highest numbered segment. For example, in summary file 1 there are 39 segments, so segment 1 for would be named cen2ksf1.s01.sps. The last segment would be named cen2ksf1.s39.sps.

Launch SPSS. If you have already created the .sav files for the segment(s), go to step 7. If you need to create .sav files for multiple segments, repeat steps 1-6.

1) Click on file/open/syntax. Find the directory containing the .sps and data files.
2) Load the .sps file for the appropriate segment. A new SPSS window (the syntax editor window) will open with the contents of the .sps file displayed.
3) The second line of the .sps file for the segment starts with /FILE. If necessary, change the line so that it refers to the appropriate data file.
4) Click on Run/All.
5) When the processing of the .sps file is completed, the data editor window (the one containing the data view and variable view) will open.
6) Click on File/Save As. Enter the file name according to the conventions listed above. For example, to create a .sav file for the first segment of SF1 for Delaware, the file name should be cen2ksf1.des01. The .sav will be automatically added. Make sure the file will be saved in the appropriate directory.

7) If there is already a .sav file for the geographic header file, load it by clicking on file/open/data. Continue to step 13
8) In either the data editor or syntax editor window, load the .sps file for the geographic header – cen2k.geo.sps by clicking file/open/syntax. A new syntax editor window will open.
9) As with the segment file, the second line of the .sps file starts with /FILE. Change the line so that it refers to the appropriate geographic header file. To add value labels for county, do the following…
    a) Click on File/Open/Syntax. Open file .sps file named cen2k.county.sps
    b) "Cut" the lines from the word "county" through the last value definition associated with the counties for the state you're working on. Go back to the window containing cen2k.geo.sps and "Paste" the value labels into the value labels section of the .sps file. It can go anywhere, just make sure you don't split another definition, and don't leave any blank lines.
10) Click on Run/All.
11) When the processing of the .sps file is completed, the data editor window (the one containing the data view and variable view) will open.
12) Click on File/Save As. Enter the file name according to the conventions listed above. For example, to create a geographic header .sav file for SF1 for Delaware, the file name should be cen2ksf1.degeo. The .sav will be automatically added. Make sure the file will be saved in the appropriate directory. Okay, now you're ready to merge the geographic header file with the appropriate segment file(s).
13) The data editor window should be active. If not, make it the active window. Click on data/merge/add variables.
14) There will be a prompt for the .sav file for the segment that is to be merged with the geographic header .sav file that was created in step 12. Click on it.
15) An "Add variables from …" window will open. Click on the button that says "Match cases on key variables in sorted files. Make sure the option "Both files provide cases" is selected.
16) From the list of "Excluded variables" click on logrecno, and then the little arrow to the left of the key variables box.
17) Click on OK. A warning will come up telling you that both files need to be sorted in ascending order on the key variable. If the census bureau has done its stuff this is fine, so click on ok.
18) That's it – the file in the data window is now the product of the merging. Compute away.