

Merge Codes Resource Document

In merging restricted-use data files using the school control number (CNTLNUMS), both data files being merged must be sorted by the school control number prior to performing the merge. In the Stata code shown within the module (also shown here), words in *italics* are meant to be replaced by file or variable names that the user specifies.

```
use dataset1
sort CNTLNUMS
save dataset1, replace
use dataset2
sort CNTLNUMS
save dataset2, replace
merge CNTLNUMS using dataset1 /*merges the two files */
drop if _merge= =2 /*specifies dataset1 as unit of analysis*/
save newfilename, replace /*saves a new file keeping all dataset1 records and only
    matching dataset2 records*/
```

When merging any of the school, principal, teacher, or school library media center files together for a given school, the school's control number, CNTLNUMS, is used to merge data files. When using Stata, a default merge variable is created during the merging of data files. The default name of this variable is *merge*. The variable *merge* identifies the various categories of data in a one-to-one merge and can be used to specify a unit of analysis. To merge the Public School District Data File with other public sector data files, the district's control number (CNTLNUMD) should be used. This variable is included on the Public School District Data File as well as the public school, principal, teacher, and library data files. The sample code provided above is correct, except that the merging variable will be CNTLNUMD.

The codes for merge results, and an example of merging the public school district data file with the public school principal data file are provided below.

If users merge the public school district file ("using" data file) onto the principal file ("master" data file):

_merge= =1

Observations are only from principal data file, with no public school district data added (occurs when associated district is a nonresponding case)

_merge= =2

Observations are only from public school district data file (e.g., district responded, but there is no responding principal(s) from the associated district), with no principal data added

_merge= =3

Observations are from public school district and principal data files.