

Merging in TMG 5

by Ben Franklin

“Merging” can mean different things to different people. Merging can be used in TMG to eliminate duplicate sources, duplicate place names and duplicate people. It may also be employed to import several databases into one project or to combine multiple datasets into one dataset. This is a brief introduction to merging in The Master Genealogist v5

Caveat: *Back up all of your databases before you proceed. You will be making some rather drastic changes that may not result in what you are planning.*

Adding a Dataset to a Project

GEDCOM Data

During your research, you may encounter a fellow researcher who has entered data into some sort of genealogy program. You may find information on Rootsweb, Ancestry’s World Tree, Pedigree Resource Files, World Family Tree, FamilySearch.org, or any number of places. Many of these are able to download information in GEDCOM format. This data can be incorporated directly into your research without having to type it all over again.

In addition, with TMG there is the GenBridge^a interface that allows you to directly import data from various programs.

If you are only using the data as reference material, you might want to just leave the data in the form that the author sent it to you. In this case, you will essentially be using TMG as a viewer of this data. You would import this into a project - perhaps a project that includes a collection of similar GEDCOMs - and lock it as a read-only dataset.

The next section includes instructions are for those situations where you may want to incorporate the data you have found into another dataset.

Merging Datasets

See What You Have

Carefully inspect the new dataset. See if it includes lineages that are of no interest. Remember that TMG is already a very slow program. Having a few thousand unneeded people in your datasets will slow it down even more, every time you open a list, generate a report, etc. If this dataset includes such extraneous lineages, trim them off.

Look to see if the event types have been mapped properly. You will find that various people have entered data in unusual or unexpected ways into their genealogy programs. In my experience, this seems to most frequently afflict data stored in Family Tree Maker format. Users don’t seem to know how to properly input data, so they just put it anywhere they can. Because of this,

although TMG might translate the data properly, it could map into some strange event types in TMG. For instance, when you see the **Confirmation** tag with a place of “Such-and-Such Cemetery,” you can be pretty sure this was a mapping problem.

Unfortunately, because of the way that TMG groups event types, TMG will not allow you to merely change the event type. To correct this type of defect, you will have to re-add the tag copy the information and then delete the bad one. Remember to copy the source information.

Place Names

You may want to edit the place names. Be aware that many genealogy applications don't have the capability to manage place names, resulting in wildly inconsistent naming found within the same dataset.

Sources in Merged Data

Unfortunately, more frequently than not, your merged dataset will not have any sources listed. This of course casts doubt on every fact, name spelling, date, etc. You will want to make sure that you have marked everything in this dataset as coming from this GEDCOM before you proceed to use it or merge it into your other data.

Similarly, for those GEDCOMs that include source data, you may want to make a special notation in the sources of this new dataset to indicate who did the research. Perhaps you would want to include some indicator that you have not seen the source materials yourself, or do not have a copy of the source. Nothing can be more frustrating than searching through thousands of your source documents for something that is not there...

Caveat: *Before you merge datasets, mark all of the events with their source. It will be very difficult to do this afterward.*

In Summary

After you have imported a dataset, and before you merge it into another dataset:

- Trim off branches and other information, that don't interest you.
- Check the sources.
- Check the place names
- Label all of the events with the overall source of the data (the GEDCOM, etc.)

Merging Multiple Sources

It is easy to get duplicates of your various source references. You can accidentally add one after not finding the current source in the list of sources. Duplicate source references can also be introduced when you import a dataset.

When you have two differing copies of the same source, be sure to retain the more detailed version of the source. It is easy to lose valuable information by choosing the wrong one to merge.

Merging Multiple Places

Merging Multiple Places is the simplest type of merging in TMG. If you establish identical place names in identical field, TMG will automatically eliminate duplicates. The elimination of duplicates is one of the features of the **File->Maintenance->Optimize** function.

From TMG's help file (Optimize topic): *For two place records to be considered duplicates, all fields must be the same. This includes the Style, if one was used to create some of the entries. It is possible, for example, for two Master Place records to appear identical, but to have been generated using different Styles.*

Optimizing will not remove unused place records for which a start year, end year, comment, or short place name have been recorded.

Merging Multiple Copies of the Same Person

This is the most difficult and tedious of the types of merging.

Identify Candidates

In a small dataset of only a few hundred, or even a thousand individuals, or so, one of the easiest ways of identifying candidates for merger is to scan the Picklist. Remember that multiple name variants will all appear in the list. You can eliminate the non-primary names in the **Filter...** screen.

Another way to identify duplicates is to use the **Check for Duplicate People** function (**Tools->Check for Duplicate People**).

Actual Merging

Here is the screen:

The screenshot shows a 'Merge Two People - Split Display' window. It contains two tables side-by-side, both for 'John Drewry GOODWYN Jr. (2:1)**' and 'John Drewry GOODWYN Jr. (2:24)'. The left table has columns: Label, *, Date/Name/Place, X, and -. The right table has columns: Label, *, Date/Name/Place, X, and ^. Both tables have checkboxes for marking records. Below the tables are buttons for 'Mark All', 'Unmark All', and '>> Combined Display <<'. At the bottom are buttons for 'Flip', 'Merge', 'Close', and 'Help'.

Label	*	Date/Name/Place	X	-
Name-Var	*	GOODWYN, John Drewry Jr. (2:1) (1938-)	✓	
Dau-Bio	*	GOODWYN, Janet Deigh (2:3) (1962-)	✓	
Dau-Bio	*	GOODWYN, JoAnne Drew (2:4) (1969-)	✓	
Birth	*	22 Aug 1938, Durham County, North Carolina	✓	
Marriage	*	01 Apr 1961, Durham, Durham County, North Carolina	✓	

Label	*	Date/Name/Place	X	^
N	*	GOODWYN, John Drewry Jr. (2:24) (1938-)	✓	
Birth	*	22 Aug 1938, Durham County, North Carolina	✓	
Father-Bio	*	GOODWYN, John Drewry (2:25) (1901-1954)	✓	
Mother-Bio	*	LIFSEY, Hattie Davis (2:26) (1903-1991)	✓	

** ID #, primary and flags will override

One standard that I adhere to is to merge the person with the higher ID number into the person with the lower ID number. In the Merge screen, the lower numbered person should be on the left. This is the record that will remain when you are done. The ***Flip*** option allows you to swap the left and right person with each other.

Cleanup

After merging, inspect the various tags. You will note that if your people each had parents, the new person has two sets of parent tags, pointing to different parents. If you merge each of the parents together, the person will have two copies of the parent tags.

Caveat: *TMG does NOT merge Tags!!!*

Visually inspect the resulting person for duplicate tags. Delete exact duplicates.

Caveat: *Remember to copy source references from the tag to be deleted when you delete duplicate tags.*