

A Surname Relationship Diagram

Joseph Kummer

Typical genealogy applications offer a small set of views for your data. They seem to fall into one of two camps. Either they present the lowest level data so that you can edit the details of individuals lives or they present lists of all people, places, or events. In the first case, it is hard to understand how the current item on which you are working fits into the overall story of your data. In the second case, many pieces of information are presented with no indication of relationship between the pieces. The goal is to create an intermediate view.

My initial impulse was to use the family as the simplifying concept, but with over 430 families, the goal of condensing them all down to a single view seemed unlikely. I then considered a grouping by surname; 192 surnames. That's still a lot, but maybe the list could be pruned. Upon closer look at the names on my list, I had to wonder how much certain names weighed on the overall understanding of what the data was representing. If we start filtering this set by how many persons were associated with each surname, the number of surnames drops significantly. Ignore surnames with only 1 member, there are only 94 remaining; ignore those with less than 7, 27 remain; ignore those with less than 16, there are only 14. A diagram with 14 objects is generally understandable. Moreover, these 14 surnames represent 446 individuals, 51% of all individuals with surnames in the database (68 individuals have no surname defined; typically wives listed in censuses by first name only).

The surname relationship diagram consists of just two primary features. Bubbles represent a surname and lines represent the relationship between surnames. To give a visual indication of the number of individuals with a given surname, the bubbles are sized relative to the number of individuals. If surname A has twice the number of individuals as surname B, it will be displayed twice as big. The bubbles also contain text of the surname and quantity of individuals.

Surnames are related to one another via the relationship event that joined them. The most common relationship event is a marriage, but a relationship can also be a surname spelling change or a legal name change. These relationships are each illustrated as a line connecting two surname bubbles. For marriages, an arrow head is attached to the end connected to the surname to which a name is being changed. In the relation A --> B, a woman from surname A has married a man from surname B and changed her name. The connecting line can also be adorned with the date of the

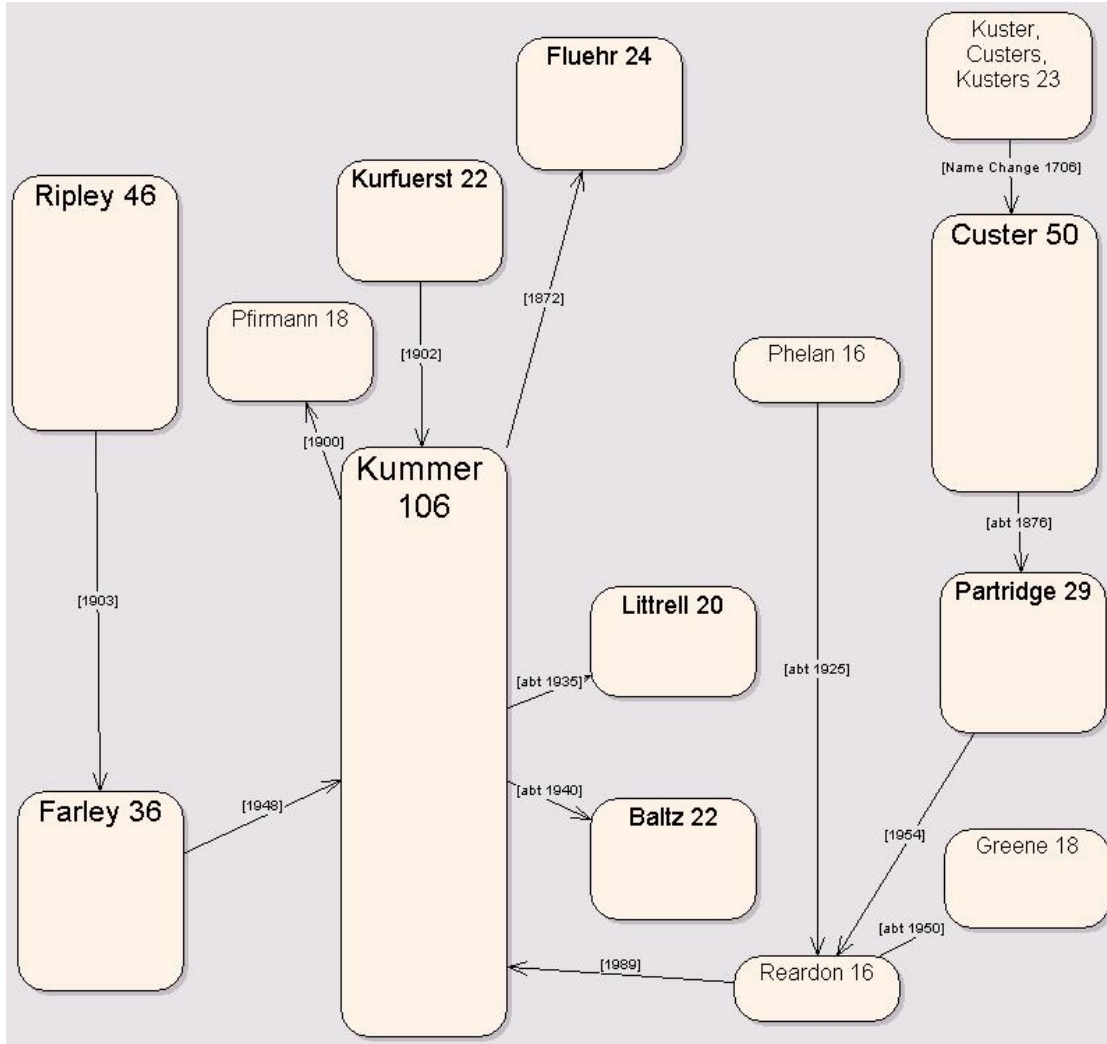
Copyright 2009 Joseph Kummer. Permission to reproduce is granted if not for profit or commercial advantage, and if the original work is properly cited.

Author's address: Mount Ephraim, NJ

Cite as: Joseph Kummer, "Surname Relationship Diagram", *Annals of Genealogical Research* Vol. 5, No. 1 (2009) at <http://www.genlit.org/agr/viewarticle.php?id=26>

event or other information.

Further context can be added to the diagram by arranging the events ascending chronologically from top to bottom. For layout purposes, this can only be done as an approximation.



By referring to the diagram included, important statements can be made and areas of possible inquiry are exposed. Typical observations are: the Kummer surname is the most represented at 106 individuals; In 1948 a Farley woman married a Kummer and took his name; the Custer surname had several variations that resolved to Custer for my database in 1706. Research possibilities include: Reardon is an important surname in my tree, yet there are only 16 members, I could delve deeper into the Reardon surname; there are several marriage dates that are approximated, they could be nailed down better; why was there confusion with the Custer name variations?

The diagram that appears in this article was generated using a tool similar to those

readily available in Microsoft Office or software freely downloaded from the Internet. The bubbles and lines were manually labeled and sized. Manually created diagrams provide a snapshot of the database, but having these diagrams automatically generated as part of an existing genealogy data management system would provide two significant advantages: dynamic behavior and an alternate navigation ability. As individuals are added to the database, the surname bubble could grow. Clicking on a surname bubble could bring the user to a list of all people with that surname.

A few final thoughts: My choice of filtering at a 16 person minimum per surname worked for my data. The diagram creation tool should allow the user to vary this value. It should be understood that this method says nothing about the personal importance of a particular surname. If a particularly large surname exists, that is actually a distant offshoot of the focus of the database, it might overwhelm the presentation. The diagram creation tool should have features to hide user-selected surnames. The tool should also allow positioning of all objects.

An old expression states: "A picture is worth a thousand words." I'd change this to "A *successful* picture is worth a thousand words." An unsuccessful picture might require a thousand words. The same goes for diagrams. I hope I've saved someone a thousand words.