

Why are countries the main DMU?

You can use other DMUs. Indeed, you can create a leader year dataset if you are ONLY using the Achigos or LEAD variable groups (variable groups 10a and 10b).

However, you have to work with a country DMU in order to "speak to" the other datasets (as they are all country based). For example, data from the COW Country Capabilities dataset cannot be merge to a leader, since the leader IDs are not in the COW Country Capabilities Variable Group (Group 02a).

But if, for example, you select the "ccode" variable in both the LEAD variable group and the COW Country Capabilities variable group, then you could create a leader-country-year dataset.

Specifically, you can follow these steps:

1. Under the "Select variables" tab, go to the "02a. COW Country Capabilities" variable group, and check the variables "ccode", "year", and "milex" (by placing your cursor over the box next to each variable name and then clicking your mouse button).
2. Still under the "Select variables" tab, go to the "10b. LEAD Leader Data" variable group, and check the variables "leadid", "ccode", and "year" (by placing your cursor over the box next to each variable name and then clicking your mouse button).
3. Still under the "Select variables" tab, go to the window near the middle top of the screen that reads "# COW Country cols:", and use the up or down arrows to set the number to 1.
4. Now go to the "Prepare run" tab.
5. Set the years to the desired range. For instance, if you wish to have data from 1960 to 1970, then set the date range to run from a "Starting date" of 1/1/1960 to an "Ending date" of 1/1/1971 (NOTICE: the dataset will end with the 1970 annual observation if the day and year are kept at January 1).
6. Click on "Choose location" to name the output data file and choose its location on your hard drive.
7. Click "Generate Output".
8. A window will now open asking you to "Select top-level variable group". This window appears because NewGene recognizes both "ccode" and "leadid" as DMUs.
 - o If you choose "02a. COW Country Capabilities", the output dataset will look as follows:

ccode	year	milex	leadid	year
2	1960	45380000	49	1960
2	1961	47808000	52	1961
2	1962	52381008	52	1962
2	1963	52295008	55	1963
2	1964	51213008	55	1964
2	1965	51827008	55	1965
2	1966	67572000	55	1966
2	1967	75448000	55	1967
2	1968	80732000	55	1968
2	1969	81446000	58	1969

- If you choose ``10b. LEAD Leader Data'', the output dataset will look as follows:

cocode	leadid	year	Year	milex
2	49	1960	1960	45380000
2	52	1961	1961	47808000
2	52	1962	1962	52381008
2	55	1963	1963	52295008
2	55	1964	1964	51213008
2	55	1965	1965	51827008
2	55	1966	1966	67572000
2	55	1967	1967	75448000
2	55	1968	1968	80732000
2	58	1969	1969	81446000