EXERCISE 3. CREATING A DATA FILE TAILORED TO YOUR RESEARCH NEEDS

IPUMS-DHS Training, Nairobi

In this exercise, we'll return to some of the variables from the first exercise, to make a data file (called an "extract') and some properly weighted tables to examine change over time for Kenyan women.

Step 1: Log in
If you have not already done so, log in so you can start making a customized data file. To log in, click on the Login link at the top of the IPUMS-DHS home page.
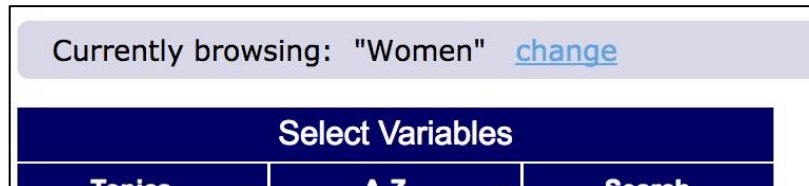
Log in using your own DHS email and password.

If you are not yet approved as a registered DHS user, see Exercise 2.

Step 2. Choose the unit of analysis
IPUMS-DHS allows you to pick between 3 units of analysis: women of childbearing age, children under age 5, and all births to women currently of childbearing age. When children or births are chosen as the unit of analysis, many of the variables relate to their mothers' characteristics.

Choose the unit of analysis that best fits your research topic. For example, if you want to know what percentage of young children recently had diarrheal disease, or were breastfed for less than a year, or were low birthweight, you should pick children as the unit. For this workshop exercise, we will work with variables relating to women.

Look at the gray box at the top of the topics menu on the "Select Data" page. If the text says, Currently browsing: "Women," then you are in the right place. If the gray bar says you are browsing "Children" or "Births," click on the link for "change" and choose "Women" from the drop-down box.

Now, each line of data in your customized extract will represent one woman, and the variables displayed will relate to women of childbearing age.

Step 3. Select samples
Once you have logged in and chosen the unit of analysis, go to "Sample selections."

Check the box next to Kenya only, and submit your sample selection. Now you'll see information on only the variables available for the 5 Kenyan samples (1989, 1993,1998, 2003, and 2008).

Step 4. Select variables of interest

Making a customized data file from IPUMS-DHS is somewhat like shopping online, except you are choosing samples and variables--and everything is free! Now we'll start adding variables to our "data cart," to make a data extract.

Please find and add the following variables: EDUCLVL, EDACHIEVER, LITBRIG, WEALTHQ, URBAN, and AGE5YEAR.

When you have added all these variables to your data cart, you should have 6 variables and 5 samples in your data cart. Click on the green "view cart" box.

Step 5. Review and/or modify your data cart

After you click the "view cart" box, you will see the list of variables you added to your data cart. You will also see other variables that the system automatically adds to your extract to handle weighting (e.g., PERWEIGHT), sample identification (e.g., SAMPLE), variables used for linking (e.g., HHID), and variables used for variance estimation (e.g., PSU).

When you have finished looking at the list of variables and are satisfied, click on the green "create data extract" box. If you had decided to add another variable at this point, such as "AGE5YEAR" for woman's age, you could click on "Add more variables" and modify your extract request.

Step 6. Submit data extract request

After you click on "Create data extract," you are brought to another page, which allows you to specify the type of file you want. For example, if you want a Stata file, click on "Change" next to "Data format" and then click on the box in front of Stata.

Next, click on "Submit." Then fill in a description in the box headed "Describe your extract," including YOUR NAME, so you can identify your own extract if you are using the workshop account for this exercise.

When finished, click on the green "Submit extract" box.

Step 7. Check for your data extract

Within a very few minutes, your data file will be available for downloading to your computer, and will show up on the Download or Revise Extracts page, accessible through the "My Data Extracts" link on the IPUMS-DHS home page.