Plan a multivariate analysis using Kenyan anthropometric data on children

For this exercise, participants should form groups of 3-4 people, working together to reach consensus on which variables to include in a multivariate analysis of predictors of growth in children, using data from Kenya.

Each group should choose one person to be the note-taker, to write down the answers you agree on for the questions below.

We'll be working with data on children. Click on "Select Data" on the internal website's home page, and choose "Children" as the unit of analysis. (This is a change from the earlier exercise). The drop-down menu will show variables relating to children under 5 (the top groups of variables) or the children's mothers. Use the "select samples" option to display only samples from Kenya.

You are going to select variables suitable for inclusion in a multivariate analysis of predictors of child growth and/or nutritional status.

The dependent variable will be an anthropometric variable related to the child's height and/or weight, and comparative standing relative to a model population.

The independent variables will be characteristics of the child (e.g., sex) and of the child's mother (e.g., education) that you think might influence children's physical development.

Step 1. Explore the Topics drop-down menu and choose the dependent variable for your model.
   a. Some of the anthropometric variables are expressed in Z scores. Read the variable descriptions and then briefly explain what that means. What is the advantage of using that measurement approach? _____ _____ _____
   b. Some of the variables use a standard distribution of values from the U.S. National Center for Health Statistics (NCHS), while others use a standard distribution of values from the World Health Organization (WHO). Which is a more appropriate standard for use with Kenyan data, and why? _____ _____ _____
   c. What variables are suitable for measuring recent acute undernutrition? What variables are suitable for measuring chronic undernutrition?
   d. If you were interested in using a dependent variable for measuring chronic undernutrition, you could recode the variable's values into a smaller number of categories (e.g., child is malnourished or is not malnourished) or use the full range of values. What are you implicitly testing if you make the dependent variable a) dichotomous, or b) use the full range of values?
   e. What variable did you choose as your dependent variable. Why?

   _____
   _____
   _____

f. Add the dependent variable you chose to your data cart, after you are logged in using the workshop email address and password.

Step 2. Decide which year(s) to analyze

a. You are probably interested in the latest results from Kenya 2014. But perhaps you also want to study change over time. How would you handle analyzing multiple years—running separate analyses for each year or include YEAR as a variable in a single regression?
b. Which year(s) did your group decide to include?
c. Modify your data cart to include the relevant year(s).

Step 3. Choose some variables related to context and child's characteristics.

a. Choose at least 3 variables related to the characteristics of the child and/or the surrounding context (e.g., URBAN). How do you think each of these variables would be related (positively or negatively) to your dependent variable?
b. Examine the documentation for your variables. If you are looking at multiple years, are there any comparability problems, and, if so, how will you handle them? Do any of the variable chosen greatly restrict the sample size? (For example, variables related to diarrheal treatment or fever treatment will limit the sample to children who recently had diarrhea or fever.) Modify your choices as needed after reviewing the documentation.
c. Add the variables you agree on to your data cart.

Step 4. Choose at least 3 variables related to the mother's characteristics. Repeat step 3, focusing on variables related to the mother's characteristics.

Step 5. Make the data extract, adding your team name to the "Description" box of your extract. Be prepared to report on your group's decisions in a discussion period.

9. Explore IPUMS-DHS on your own

Select the samples that interest you and explore which variables are available for them. Develop a research question that involves one or more dependent variables and several independent variables. Write out your ideas and explain why your question is important. You will use this material in the next exercise. _____

_____
_____
_____
_____
_____
_____