IPUMS Data Training Exercise:

An introduction to IPUMS DHS

(Online Exercise)

Learning goals

- Understand how IPUMS DHS dataset is structured
- Understand how to choose between related variables based on your research question
- Understand how IPUMS DHS employs composite coding to handle differences in response codes.
- Develop a research question and hypothesis using IPUMS DHS dataset

Summary

In this exercise, you will gain basic familiarity with the IPUMS DHS data exploration to answer research questions related to data on maternal health care. You will explore the website and investigate variables BIRTHSZ_ALL, BIRTHSZ_01, BIRTHWT_ALL, TOSMOKE, and ALDRINK, and explore research questions and develop hypotheses related to socioeconomic status and maternal health care and delivery. After completing this exercise, you will have experience navigating the IPUMS DHS website and should be able to leverage these data to explore your own research interests.

Part I. Exploring the IDHS website

- Go to https://www.idhsdata.org/idhs/, click on "Get Data."

Select Unit of Analysis

- Chose "Women" as the unit of analysis.
- Note that by selecting women as the unit of analysis, you will now only see variables available for women.

Select Variables

- The variable drop-down menu allows you to explore variables by topic. For example, you may expect to find woman-level variables related to antenatal and maternal health under the "Maternal and infant health" tab.
- The search tool allows you to search for variables. Observe the options for limiting your search results by variable characteristics.
- You may add a variable to your cart by clicking on the plus sign in the "Add to Cart" column of the topical variable list, or list of search results.
- You may view more information about the variable by clicking on the variable name, and navigating through the tabs that include a description of the variable, codes and value labels, the universe of persons asked the question, and information on the comparability of the variable among other pieces of information. If you are reviewing variable-specific information, you may click on the "Add to Cart" button near the top of the screen to add this variable to your data cart.
- Note that the _01 ending of variables in this exercise indicates that the material relates to the woman's most recent birth. Answers are at the end of the exercise (page 6).

Exercises

For each question below, use either the drop-down menu or the search feature to browse for these variables and search through the tabbed sections of the variable description to answer each question.

1. What variable would you use if you are interested in whether the woman is currently breastfeeding? _____

2. The variable BIRTHSZ_ALL reports the woman's description of the size of her child at birth. Click on that variable name to bring up the variable description.

    The Universe tab for BIRTHSZ_ALL reports which women were asked this question. The most common universe is "Women age 15-49 who gave birth in the five years before the survey." What samples have a different universe, and in what ways are the universes different?

    _____

    _____

3.   Click on the Codes tab and click the link to variable BIRTHSZ_01. Select the "case count view" option of the Codes tab for BIRTHSZ_01.
    A. How many women in the Egypt 1992 sample reported their child was "very large" or "very small"? _____

    _____

    B. What categories might you combine if you were concerned about small sample size for these categories?

    _____

    _____

    C. Which sample(s) include less detail in their responses than the 5-category responses of very large, larger than average, smaller than average, very small?

D. How does IDHS handle that difference in response codes?

Note. Review the integrated codes and the text in the Comparability tab.

_____

_____

_____

4. BIRTHWT_ALL provides an alternative measure of the last-born child's size at birth. Look at the Codes and Description pages of that variable. What are the advantages of this variable, compared to BIRTHSZ_ALL? What are the disadvantages?

_____

_____

_____

_____

5. In the United States, poorer birth outcomes, such as children with low birthweight or fetal alcohol syndrome, are often tied to health behaviors such as smoking cigarettes or consuming alcohol during pregnancy. Examine the case counts for the variables TOSMOKE (smokes cigarettes) and ALDRINK (drink alcohol) and indicate whether cigarette smoking and alcohol consumption are major issues for pregnant women in the countries included in IDHS (so far as you can tell for the samples included in those variables). Hint: Use the "Search" option to locate these variables quickly.

_____

_____

_____

Open-ended questions

6. Examine the variables available under the "Socioeconomic status" tab. These cover a variety of measures, including those relating land ownership, work and occupation, education, media exposure, and food security. Examine the variables available under the "Household characteristics" tab, including those relating to household demographics, personal wealth and possessions, and access to healthcare and disability. Which variable would you choose to identify the least privileged women, when studying maternal health, and why?

   _____

   _____

7. Maternal health care and delivery may be affected by or have effects on other areas of women's health and well-being. Explore other topic areas for Women and frame a hypothesis about some relationship between an aspect of maternal health and a variable from another topical area.

   For example: is the place of delivery (DELPL_ALL), such as outside a medical facility, associated with the likelihood that a woman has experienced a fistula (FSEVHAD)? How does intimate partner violence (DVPMSEVER, DVPLSEVER) relate to maternal healthcare or birth outcomes?

   _____

   _____

   _____

Answers

1. What variable would you use if you are interested in whether the woman is currently breastfeeding? <u>BFEEDANYNOW</u>

2. What samples have a different universe, and in what ways are the universes different? <u>Different reference period of three years before the survey: Bangladesh 2014, Benin 1996, Cameroon 1998, Cote d'Ivoire 1994, Ghana 1993, India 1998, Kenya 1998, Madagascar 1997, Mali 1995, Mozambique 1997, Nepal 1996, Niger 1998, Nigeria 1999, Uganda 1995, Zimbabwe 1994.</u>
<u>Different reference period of four years before the survey: India 1992, Uganda 1995.</u>
<u>Limiting sample to ever-married women: Afghanistan 2015, all Bangladesh samples, all Egypt samples, India 1992 and 1998, all Jordan samples, Nepal 1996 and 2001, all Pakistan samples, all Yemen samples.</u>
<u>Different age range for women of childbearing age: India 1992 includes women 13-49, Nigeria 1999 includes women 10-49, and Yemen 1991 includes women age 15-54.</u>

3. BIRTHSZ_01 variable:
   A. How many women in the Egypt 1992 sample reported their child was "very large" or "very small"? <u>27 very large and 39 very small last-born births in Egypt 1992.</u>

   B. What categories might you combine if you were concerned about small sample size for these categories? <u>Combine responses into Large, Average, and Small (first digit codes of 1,2, and 3) in the composite coding to handle small numbers of cases or incorporate samples with fewer responses, such as India 1992 and 1998.</u>

   C. Which sample(s) include less detail in their responses than the 5-category responses of very large, larger than average, smaller than average, very small? <u>India 1992 and 1998.</u>

D. How does IDHS handle that difference in response codes? Review the integrated codes and the text in the Comparability tab to find out. IDHS uses composite coding with "large," "average," and "small" responses code as the first digit, and further detail provided in the second digit.

4. What are the advantages of this variable, compared to BIRTHSZ_ALL? What are the disadvantages? BIRTHWT_ALL is a less subjective measure and allows the data to be recoded into standard categories such as "low birth weight" (under 2500 grams), "very low birthweight" (under 1500 grams), and "extremely low birth weight" (under 1000 grams). The disadvantage is that it is only available for infants who were weighed at birth. This variable may also be subject to recall error, and to limit this, the researcher may wish to limit the analysis to cases where the birthweight was written on the child's health card (code 1 in BIRTHWTREF_ALL).

5. Examine the case counts for the variables TOSMOKE (smokes cigarettes) and ALDRINK (drink alcohol) and indicate whether cigarette smoking and alcohol consumption are major issues for pregnant women in the countries included in IDHS (so far as you can tell for the samples included in those variables). Smoking and alcohol use seem very minor maternal health issues for African and Indian women, given the rarity of smoking by women and the large proportion of women who don't drink alcohol. Nepal and Turkey are the exception, with a larger proportion of women who smoke cigarettes.

Open-ended questions

6. Which variable would you choose to identify the least privileged women, when studying maternal health, and why? WEALTHQ would identify the least privileged women by the relative wealth of the household where she lives. This is a measure of the household's living standard, a better reflection of a woman's access to care and her socioeconomic status than her personal earnings, especially if her husband works and she does not.

7. Explore other topic areas on the IDHS Select Data page for Women and frame a hypothesis about some relationship between an aspect of maternal health and a variable from another topical area. <u>If a woman has the final say on how her earnings are spent (DECFEMEARN), she will be more likely to receive antenatal care (ANCARENOONE_01).</u>