Subject: Need clarification on Kenya DHS 2022 Dataset
Posted by Lubana on Tue, 23 Jan 2024 15:32:31 GMT
View Forum Message <> Reply to Message

Hello,

I hope this message finds you well.

I am working in a Project where I am going to create a comprehensive map of social vulnerability to climate change hazards in Kenya.

In pursuit of this goal, I have chosen to leverage the DHS-2022 dataset for Kenya. Having recently downloaded the Household data file (KEHR8BSD.zip) from the DHS website, I converted it to CSV format for streamlined integration into my research workflow. Unfortunately, I have found some challenges related to identifying the variable associated with the 'major source of drinking water.' I have looked at the DHS guidelines, such as the Microdata Library for Kenya and the USAID Guide to DHS Statistics, I have encountered difficulty in decoding the information contained in the Excel Sheet. The Microdata Library indicates that the code 'V113' corresponds to the major source of drinking water. Yet, upon examining the Excel file, I discovered 24 columns labeled HV113_01, HV113_02, and so forth. Unfortunately, the documentation does not shed light on the significance of these extension numbers, leaving me uncertain about the distinct sources represented by HV113_01 and HV113_02.

Moreover, I have encountered another code, 'V87 hv201 Source of drinking water' in in Microdata Library document, in Page 18, Regrettably, I am grappling with understanding the nuances of this code. I need help to figure this out. I not sure if I am using the right dataset since I am working with DHS data for the very first time. My apologies for bothering you. But I genuinely appreciate any guidance or insights you can provide.

Attached herewith is the Excel sheet containing the a small part of downloaded data for your convenience.

Thank you for your time and consideration.

Kind regards

File Attachments
1) DHS-Kenya.xlsx, downloaded 373 times

Subject: Re: Need clarification on Kenya DHS 2022 Dataset
Posted by Janet-DHS on Thu, 25 Jan 2024 21:24:10 GMT
View Forum Message <> Reply to Message

Following is a response from DHS staff member, Tom Pullum:

In the HR file, which you have opened, the variables begin with "hv". In that file, source of water is hv201. The HR file has one record per household. The multiple copies of hv113 that you found refer to different members of the household (hv113 is survival status of the father for children age 0-17). The PR file has the same data but is organized with one record per household member.

The variable v113 begins with "v", not "hv", and is in the IR file. The IR file has one record per woman age 15-49. v113 is the same as hv201 for all women in a given household.

Some variables that appear in the household files and the individual files have the same number, and differ only in terms of the prefix v or hv. However, most do not.

DHS data are not easy to use with Excel. If you can possibly switch to a package such as Stata or even SPSS, your analysis may be easier.

Thanks for using DHS data. Let us know if you have other questions.