

---

Subject: Labeling mistakes in exposure to mass media variables in Zambia 2007 stata dataset

Posted by [ykang12](#) on Sun, 03 Aug 2014 16:00:43 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

I am Yunhee, graduate student of Johns Hopkins School of Public Health.

I have been doing data analysis with Zambia 2007 DHS dataset (stata format) and I wanted to know exposure to mass media among adolescent groups by tabulating such variables (v157,v158,v159). But I found labeling mistakes below and questions/answers seemed to be inconsistent with questions in DHS questionnaires. Could you help me clear about this?

< STATA dataset>

. tab v157

```
frequency of reading |
newspaper or magazine |   Freq.   Percent   Cum.
-----+-----
no education, preschool |   883    55.60    55.60
    primary |   233    14.67    70.28
    secondary |   261    16.44    86.71
    higher |   209    13.16    99.87
    9 |     2     0.13   100.00
-----+-----
Total |   1,588   100.00
```

: their answers were recorded as 0,1,2,3,9, which is not consistent with Zambia questionnaire itself and even compared with other countries' dataset.

< DHS questionnaires>

Do you read a newspaper or magazine almost every day, at least once a week, less than once a week or not at all?

ALMOST EVERY DAY .....1

AT LEAST ONCE A WEEK .....2

LESS THAN ONCE A WEEK .....3

NOT AT ALL .....4

---

Subject: Re: Labeling mistakes in exposure to mass media variables in Zambia 2007 stata dataset

Posted by [Bridgette-DHS](#) on Mon, 04 Aug 2014 12:45:03 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

I am looking at the frequencies in the file: zmir51fl.dta, and can't see what you have in your post. Please send me the complete name of the file you are using, and I will run some frequencies to compare with the result you are having.

---

---

Subject: Re: Labeling mistakes in exposure to mass media variables in Zambia 2007 stata dataset

Posted by [dab000@mail.harvard.edu](mailto:dab000@mail.harvard.edu) on Wed, 28 Jun 2017 13:57:47 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Hello Bridgette,

I am having similar problems to what is described above for V157-V159 of the Kenya DHS for 2003 and 2008 - the response options in the dataset range from 0 to 3, rather than the options in the questionnaires, which range from 1 to 4 in the surveys, and tabulations of these responses do not cleanly map onto published results. For example, the final report for the Kenya DHS 2003 survey says that 22.5% of women read the newspaper at least once a week, 28.8% watch television at least once a week, and 75.1% listen to the radio once a week. However, this does not match the results I have calculated below.

I am not having problems reproducing other results, and I am also not having a problem reproducing these statistics in KDHS 2014 (although there is still a discrepancy in the response options in the dataset vs. the response options in the questionnaire).

Any advice or insights into this issue?

#### File Attachments

---

1) [Screen Shot 2017-06-28 at 9.48.19 AM.png](#), downloaded 810 times

2) [Screen Shot 2017-06-28 at 9.48.13 AM.png](#), downloaded 844 times

---

---

Subject: Re: Labeling mistakes in exposure to mass media variables in Zambia 2007 stata dataset

Posted by [Bridgette-DHS](#) on Tue, 11 Jul 2017 14:35:32 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Here is the weighted distribution of the responses to radio listening in the Kenya 2003 survey:

"Listening to the radio once a week" means listening at least that often. It is the sum of codes 2 and 3, which is 75.18%. If the report says 75.1%, I would consider that to be a match. (The difference of a tenth of a percent is annoying but is probably due to the treatment of code 9.)

I don't know why codes 0 to 3 were used rather than codes 1 to 4, for this survey. We occasionally encounter survey-specific or country-specific inconsistencies of this sort. Your

problem in getting a match was that your consolidation of codes was not appropriate for the 0 to 3 coding.

I assume that this will explain the problem getting a match for newspaper and TV too, and for Kenya 2008 as well as Kenya 2003.

### File Attachments

---

1) [tab-v158.jpg](#), downloaded 828 times

---