2007/08/03 11:48 AM

Hi Lynn

At last. We've found the error. The questionnaire supplied (in word
format) was in fact the one for OHS 1997. Although the Flap referred to
OHS 1998 the other parts (Word files) were those of the OHS 1997
Questionnaire. The correct OHS 1998 questionnaire (in PDF format) can be
downloaded from:
http://www.statssa.gov.za/sending/thefiles/OHS_Questionnaires/OHS_Questionnaires.zip

Sorry for the Inconvenience.

Kind regards

Piet

>>> "Lynn Woolfrey" <Lynn.Woolfrey@uct.ac.za> 7/26/2007 11:29 AM >>>
Thanks, Piet - we've made a note in the dataset in the meantime.

Regards
Lynn

>>> "Piet Alberts" <PietA@statssa.gov.za> 2007/07/24 12:27 PM >>>
Hi Lynn

As mentioned before, something looks strange with marital status in
OHS
1999.
We are investigating the issue and will report back ASAP.

Piet

>>> "Lynn Woolfrey" <Lynn.Woolfrey@uct.ac.za> 7/13/2007 4:04:46 PM >>>
Dear Piet

It seems the OHS 1998 dataset used the marital status codes from the
OHS 1999 (see the query below), so that the data is incorrect. Could
you check if there is a corrected version of this?

Thanks
Lynn

It looks as if the marital status variable in the 1998 OHS was coded
according to the 1999 codes. In 1999 the codes were as follows:

1 Married - civil
2 Married - traditional (customary)
3 Living together
4 Widower/widow

5 Divorced/separated
6 Never married

>>> Martin Wittenberg 2007/07/13 02:57 PM >>>
Dear Lynn

My student Grace Kumchulesi has picked up what looks like a serious
error in the OHS 98 marital codes. The error is most obvious when
looking at the proportion of each age group in a particular marital
state, as shown in the graph that is attached to this e-mail
(generated
by Grace)

According to the questionnaire the codes are as follows:
1 Never married
2 Married - civil
3 Married - customary
4 Living together
5 Widowed
6 Divorced.

It is abundantly clear, however, that a code of 6 seems to have been
used for the "never married", a code of one looks like "married -
civil"
and what Grace coded as "married" (i.e. categories 2 to 4) cover the
widows.
The raw proportions bear this out: 65% of the sample is in category 6
and 17% in category 1.

I have tried to track down where this error could have come from. I've
gone back to my copy of the raw data and read it in taking care that I
read in the correct column into the "marital" variable (as per the
metadata) and got precisely the same results. So the error is
definitely
in the raw data.

Could you please help us to find out what could have gone wrong? It
would be useful to check whether the SADA copy of the data also has
this
error.

Thanks
Martin