

DATA QUALITY REPORT

1. DATA IDENTIFICATION	
Data File:	headteacher1.dta
Dictionary:	headteacher1.smcl
Data Cleaning Do-File:	headteacher1.do

2. STRUCTURAL STABILITY			
	Yes		No
2.1 Do all variables in the questionnaire match those in the data set?			X
2.2 Do all value labels in the questionnaire match those in the data set?			X
2.3 What is the name of the dataset's unique identifier?	schoolco		

If 2.1 or 2.2 is false, please append additional sheets describing the discrepancies.

3. EDITING AND IMPUTATION			
	Yes		No
3.1 Did any variable exhibit errors in more than 5% of observations?	X		
3.2 Was any editing or imputation applied to the data set <i>due to illegal or incoherent responses?</i> (see attached do-file)	X		
3.3 Were any outliers rejected?			X

If either 3.1 or 3.3 are true, please append additional sheets describing the errors identified. 3.2 should appear in the attached data cleaning do-file.

This dataset corresponds to sections I, II and III of the questionnaire « Questionnaire for Head-teacher of school Expenditure and Service Delivery Survey ZAMBIA : Educational Sector The World Bank». The document was named Head-Teacher_Questionnaire.doc in the questionnaire file.

2.1 Inconsistence

- (1) Section 7 in the questionnaire have no data set.
- (2) Section VIB s6bq5 said “if 2 skip to 10” , there is no question 10 in this section.
- (3) s6cq13a should be labelled as s6cq13, s6cq14a and s6cq14b are created. The questionnaire only has s6cq14.

2.2 Matching

- (1) s1q14 in the questionnaire and s1q14 in the data set are not matched.

3.1 Variable exhibit errors in more than 5% of observation

The following variables exhibit errors in more than 5% of observations:

s1q12, s1q16, s3q3a-s3q3f, s3q4a-s3q4g, and s3q7a-s3q7e

3.2 Editing

The following editing have been corrected.

- (1) The skip pattern induced by missing data or conflict in s2q6 and s2q7, if s2q6 is yes, but no answer in s2q7, s2q6 was replaced by “no”.
- (2) s3q7e have been replace by zero if s3q7e is missing.
- (3) If s3q2a-g is no, the corresponding s3q3a-g is replaced to zero

3.2.1 Skip pattern that has not been corrected

- (1) There are a big block info can be removed. If s1q16 is no, s1q17-s1q21a,b should not be answered. While there are over 60 percent of the survey participates answered question s1q17-s1q21a,b. Which are incorrect responses. But we will not correct these inconsistence.
- (2) s3q7a-e and s3q8 questions, there are many people answered s3q8 questions though they confirmed no in s3q7. This skip pattern is not correct due to too many errors.

3.2.2 Other Editing

- (4) “-99” in the cell is replaced by “.r”
- (5) S2q7 initial value is 10 and is replaced by zero.
- (6) S1q18 initial vale is 10 and is replaced by zero.

3.3 Outliers

(1) No outliers were found.