

DATA QUALITY REPORT

1. DATA IDENTIFICATION	
Data File:	pupilroster.dta
Dictionary:	pupilroster_codebook.smcl
Data Cleaning Do-File:	pupilroster.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?	ID		

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 « Head-teacher Matching Roster Expenditure and Service Delivery Survey ZAMBIA : Educational Sector The World Bank». This questionnaire was named pupil Matching Roster.doc in the file.

2.1 Extra variables

Question 1 has only three choices (yes[1], no[2], do not know[3]), but the survey data set has the fourth choice (not in the roster[4])

3.2 Editing

The following editing have been accomplished.

- (1) Corrected skip pattern errors in q1 and q2; q6 and q7; and q8 and q9. The values of q2 , q7, and q9 that are greater than 0 were replaced by missing values.

Other Editing

- (1) Generate unique ID (pupil_id) by concatenate school code and pupil code. And destring pupil_id from string variable to long variable.
- (2) "99" in the cell was replaced by ".r",
- (3) "77" in the cell was replaced by ".n"
- (4) Replace "roster not done" in question 1 choice 4 to "missing"