

DATA QUALITY REPORT

1. DATA IDENTIFICATION	
Data File:	enseignant.dta
Dictionary:	enseignant.dtc
Data Cleaning Do-File:	enseignant.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 the questionnaire « enquête enseignant- Visite 2eme jour ».

3. Editing

The question 2 asking about the presence of the employee has been incorrectly answered in many cases. If the person is not present, the enumerator should go and carry out the questionnaire to the next employee. When all the questions have been answered and that the employee is reported to be absent, the assumption that it was present at the time of the interview is made. It seems a reasonable assumption as all these questions asking for some detailed information cannot be filled if the employee is not present. The employees for which such changes have been made are indicated below:

20505221	51901225	103072315	30455235
20505222	30813235	103072316	30406231
20505223	40611233	22029238	40811232
10923223	71631235	30406211	11419214
10923224	20509233	30406212	11423211
10931231	11131233	30406233	51402232
10931234	301992811	30455222	51455221
20401212	103072314	30455234	

Answer to question 8 has been replaced by missing as the employee is reported to be absent and this question is the only one to have been answered for employee 20509233.

Answer to questions 15 and 34 have been replaced by missing as the employee is reported to be absent and these questions are the only ones to have been answered for employee 20505234.

For the employees listed below, a certain number of questions have been answered by zero or no have thus be replaced by missing to respect the skip pattern induced by question 2. Note that for employees 20705221, 30199247 and 40912212, 1 or 2 questions concerning the salary have been filled but are also replaced by missing due to the assumed absence of the employees and the scarcity of the number of questions answered.

10923221	20601233	21343232
11405221	20607222	51119211
11405223	20607232	30199247
11405232	20705221	40912212
11405234	20705223	40912232
11405235	21329236	71631221

Answer to questions 28, 29 and 30 have been replaced by missing as the employee is reported to be absent and these questions are the only ones to have been answered for employee 11216232.

Answer to questions 33 to 36 have been replaced by missing as the employee is reported to be absent and these questions are the only ones to have been answered for employee 301992820, 301992823 and 30802211.

Answers to questions 34 to 36 have been replaced by missing as the employee is reported to be absent and these questions are the only ones to have been answered for employee 20509235 and 10609211.

Answers to questions 7a, b and c have been replaced by missing for 3 employees in order to respect the skip pattern induced by answering 2 or 3 to question 6.

Answer to question 29 has been replaced by missing for 11 employees in order to respect the skip pattern induced by answering 1 to question 28.

Answer to question 29 has been replaced by missing for 10 employees in order to respect the skip pattern induced by answering 2 to question 28.

Answer to question 29 has been replaced by missing for 56 employees in order to respect the skip pattern induced by answering 4 to question 28.

Answer to question 29 has been replaced by missing for 2 employees in order to respect the skip pattern induced by answering 5 to question 28.

Answers to question 32, 33 and 34 have been replaced by missing for 1 employee in order to respect the skip pattern induced by answering 2 to question 31.

Answers to question 33 have generally been misunderstood and thus they have been replaced by missing instead of zero for 565 employees in order to respect the skip pattern induced by answering 2 question 31.

Answers to question 34 have generally been misunderstood and thus they have been replaced by missing instead of nothing (or food for 8 employees) for 376 employees in order to respect the skip pattern induced by answering 2 question 31.