

## **SDI Education datasets: Anonymization Protocol**

### **Objectives**

In order to release the SDI Education dataset as a Public Use File, it is necessary to ensure the privacy of its participants. To this end, and with the intention of avoiding the re-identification of the schools, teachers and pupils, the following measures are taken:

1. Deletion of variables that contain confidential information or lead to re-identification.
2. Deletion of value labels that contain confidential information or lead to re-identification.
3. Recoding of variables that could lead to re-identification of the observations into broader group categories.
4. Trimming and/or censoring of specific unique values and outliers that might allow re-identification.

### **Results**

The resulting datasets keep the usefulness of the data intact while greatly protecting the privacy of the respondents and reducing the identification risk. Most harmonized variables (>80 %) are unaffected by the anonymization process. In very few instances, some variables are relabeled to avoid the publication of the exact questions/answers of confidential assessment instruments.

All identifiers are being cleaned (deletion of value labels). Some variables are being recoded into categorical variables to ensure confidentiality.

## 1. Disclosure risk and confidentiality protection

Microdata often contain confidential or sensitive information, which makes release of these datasets in their original form impossible. Release of the data could reveal this confidential information and lead to a breach of privacy of the respondents. This has ethical and, in many cases, legal objections. Furthermore, when confidentiality is not guaranteed, current and potential future respondents are less likely to be willing to respond in future surveys.

The aim of this process is to create a Public Use File (PUF), which is a dataset that is freely accessible to the greater public taking in consideration the identified issue. This PUF must also minimize as much as possible unnecessary disturbances to the original to preserve the usability and quality as much as possible when possible.

Risk in the Statistical Disclosure Control context is the probability or likelihood that disclosure by an (hypothetical) intruder of a record occurs. Disclosure can be **identity disclosure**, when the identity of an individual or entity in the dataset is correctly revealed, or **attribute disclosure**, when the intruder gains new (confidential) information from the dataset. Identity disclosure can imply attribute disclosure. The risk is dependent on several factors, amongst others the frequency of **keys** (i.e. combinations of values of key variables), sample size and sampling weights as well as the availability of external information to intruders to use for re-identification. The disclosure scenarios for a particular dataset describe these parameters and the way an intruder can use a dataset to gain new information.

The acceptable level of risk depends on the release type, e.g. scientific use file, public use file, or other ways of release and the sensitivity of the data. This dataset was prepared to be released as **PUF and hence needs a higher level of protection**. Also, the potential harm caused by disclosure should be considered when determining the acceptable risk level.

In similar microdata releases with, for instance, business survey data, the geographical level is highly reduced, large companies are suppressed and the level of detail in the data is reduced to protect the records. Generally, the period between the survey and data release is also specified, e.g. 1 year. In case of the SDI education survey, the period between the survey and the data release introduces already uncertainty into several variables, such as number of pupils. It should be noted that a complete elimination of disclosure risk is not possible.

## 2. SDI Education data set

The SDI Education dataset consists of a series of country-year surveys, each containing information of schools, teachers and pupils. The main concern for re-identification and confidentiality are the teachers and pupils. However, since the data are hierarchical, i.e. teachers and pupils belong to schools, the re-identification of a school might lead to the re-identification of a teacher or pupil too.

The datasets consist of the following:

- Set 1: School level data
  - Set 1a: School management/finances (Module 3)<sup>1</sup>
- Set 2: Teacher level
- Set 3: Child level
- Set 4: Time on task

These datasets contain sensitive and confidential variables, especially on the level of the teacher but also on the school level.

### 3. Actions taken

The Anonymization process is done with the aid of the statistical software Stata. All anonymization steps are reproducible with the Stata script for each of the detailed datasets. The process starts from the harmonized dataset: before the start of the anonymization process, any final data quality corrections are made.

Each anonymization script covers the following steps:

1. Identification of ready-to-release variables.
2. Identification and removal/anonymization of variables due to the sensitivity of its data. Identification and removal/anonymization of variables due to the distribution of the data that could lead to high risks of re-identification.
3. Identification and recoding of variables into categories to deal with outliers (top recoding), sample unique and continuous variables whose values represent high risk of reidentification (special unique).

#### 3.1 Ready to release variables.

There is a subset of variables that do not imply considerable risks of disclosure. These variables were identified and revised. We proceeded to select them considering:

- General information: Country, survey year, urban, etc. This information is preserved and there is no risk associated with the release.
- Randomized id/keys: region, district, province, etc. This information is shared considering it doesn't contain descriptive information. In general terms, it preserves the variability of the data but doesn't provide an associated value label.

---

<sup>1</sup> The publication of new (or updated) PUF datasets includes the Education SDI's module 3 which compiles information of the school's management and financial activities (e.g. budget, expenditure, fundraising, monitoring and evaluation, etc.). This module has been published separately from the school level dataset as it was not harmonized to a common standard. Given distinct educational (and administrative) structures, module 3 is uniquely designed according to the country's interests and context, creating constraints for a harmonization process across SDIs. It is important to note that, despite this dataset not being harmonized, it is still subject to a thorough anonymization process like the rest of PUF datasets.

- Specific information: a subset of infrastructure variables, a subset of assessment variables, other school and teacher characteristics, etc. These variables gathered through SDI survey represent a high risk of breach of confidentiality. We analyze each of these and their distributions to check if unique values allow for reidentification.

### 3.2 Identification of variables to delete

- Sensitive information

The SDI surveys gather data that includes sensitive information: names, financial information and specific descriptions, among others, that should not be available to public.

To avoid its disclosure, some of them are directly removed from the database and others are just transformed to a “Confidential” or “.c” value, allowing only the knowledge of their existence (to be considered its release upon request under strict protocols) and the identification of missing values.

There are other variables that do not represent a risk of disclosure but include important information that could damage future waves of surveys. These variables contain detailed descriptions of the assessments and their correct answers. All of them are being removed, keeping only general information on the type of question, its result/score and the label associated to interpret the latter.

- Variables with distributions that represent risks of reidentification

We identify variables with distributions that represent high risks of reidentification and for which recoding and other perturbative methods cannot account for the mentioned risks. These variables contain sample unique (a unique combination of values for the selected categorical key variables in the dataset and is at high risk of re-identification) and are represented mainly by descriptions/details of “Other” categories. It is not possible to share the information contained in any of the respective variables due to the specificity of the answers.

Table A and B shows the deleted and anonymized variables.

Table A. Overview of removed variables by dataset

Deleted Variable	Description
<u>School level dataset</u>	
gps_lat, gps_lon, gps_miss_ind, gps_lat2, gps_lon2	GPS variables
hfid, _GIS_merge	Variables that come from internal analysis and checks done by SDI team
<u>Teacher level dataset</u>	
All variables ending in “_mark”	Information on teacher’s specific answer
All variables ending in “_correction”	Information on correct answer

lit\_4\_denom

Variables that come from internal analysis of SDI team

Child level dataset

lang\_nomiss-total\_theta\_se

Variables that come from internal analysis and checks done by SDI team

Note: a. Additional variables deleted for the 2016 Morocco's Education SDI.

Table B. Overview of anonymized variables

<b>Anonymized Variable</b>	<b>Description</b>
<u>School level database</u>	
refusal_reason1	If refused, reason for refusal (1)
fv_enum_name1	First Visit Enumerator (s) 1 Name
sv_enum_name1	Second Visit Enumerator (s) 1 Name
region_name	Region name
district_name	District name
district_code <sup>b</sup>	District code
county_code <sup>b</sup>	County code
ward_code	Ward
ward_name	Ward name
village_street_code	Village/street
village_street_name	Village/street name
pedag_admin_zone_c <sup>b</sup>	Pedagogical Administration Zone Code
massar_code <sup>b</sup>	School Massar Code
school_name	School Name
school_code	School Code (Other ID)
emis_code	EMIS Code
fv_date_day	Date of 1st Visit (day)
fv_date_month	Date of 1st visit (month)
fv_date_year	Date of 1st visit (year)
sv_date_day	Date of 2nd Visit (day)
sv_date_month	Date of 2nd visit (month)
sv_date_year	Date of 2nd visit (year)
team_lead_date_day	Team leader date (day)
team_lead_date_month	Team leader date (month)
team_lead_date_year	Team leader date (year)
super_date_day	Supervisor date (day)
super_date_month	Supervisor date (month)
super_date_year	Supervisor date (year)
fv_arrival_time_hr	At what time did you arrive at the school? (hour)
fv_arrival_time_min	At what time did you arrive at the school? (min)
fv_end_time_hr	At what time did you leave at the school? (hour)
fv_end_time_min	At what time did you leave at the school? (min)

sv_arrival_time_hr	At what time did you arrive at the school? (hour) (second visit)
sv_arrival_time_min	At what time did you arrive at the school? (min) (second visit)
sv_end_time_hr	At what time did you leave at the school? (hour) (second visit)
sv_end_time_min	At what time did you leave at the school? (min) (second visit)
fv_date_comb	Date of First Visit (combined)
sv_date_comb	Date of Second Visit (combined)
team_lead_date_comb	Team leader date (combined)
super_date_comb	Supervisor date (combined)
fv_arrival_time	At what time did you arrive at the school: First Visit
fv_end_time	At what time did you leave the school: First Visit
sv_arrival_time	At what time did you arrive at the school: Second Visit
sv_end_time	At what time did you leave the school: Second Visit
county_name	County name
controller_name	Controller Name
village_name	Village name
control_date	Control Date
town_name	Town name
fokontany_name	Fokontany name
pedag_admin_zone_n	Pedagogical Administration Zone Name
fv_enum_name2	First Visit Enumerator (s) 2 Name
sv_enum_name2	Second Visit Enumerator (s) 2 Name
refusal_reason2	If refused, reason for refusal (2)
subcounty_name	Sub-County/Division/Town Council
parish_name	Parish Code
operator_id	Data entry operator ID
state_school_code	State School Code
private_school_code	Private School Code
knec_code_new	School KNEC Code (new)
knec_code_old	School KNEC Code (old)
tsc_code_public	School TSC Code (public)
moe_code	School MoE Code
division_name	Division name
IDEN_name	IDEN name
m1_resp_name	M1 What is your name?
m1_resp_phone	M1 Please can we have your Mobile Phone number
m1_resp_post_other <sup>b</sup>	M1 Which Position do you occupy in this facility? (Other)
m1_sc_ownership	M1 What's this school's ownership type
m1_satelite_num	M1 If your school is a central or community school, how many satellites does it
m1_satelite_id	M1 If it is a satellite school, enter the identifier of its central school

m1\_toilet\_type5  
 m1\_toilet\_type6  
 m1\_toilet\_type\_other1  
 m1\_toilet\_type\_other2  
 m1\_water\_type\_other  
 m1\_road\_type\_other  
 m1\_satelite\_dist  
 m1\_comments1  
 m1\_qao\_feed\_discuss  
 m1\_school\_type\_other  
 m1\_school\_open[n]<sup>b</sup>  
 m1\_school\_close[n]<sup>b</sup>  
 m1\_comments2  
 m2\_comments1  
 m2\_comments2  
 m2\_comments3  
 m2\_n\_comp\_rooms  
 m4\_late\_enum  
 m4\_teachername  
 m4\_comments1  
 m4\_comments2  
 m4\_comments3  
 m4\_comments4  
 m4c\_teach\_callname  
 m4c\_teach\_callname\_n  
 m4c\_teach\_projector  
 m4c\_teach\_screen  
 m4c\_teach\_ed\_software  
 m4c\_teach\_videogame  
 m4c\_teach\_video  
 m4c\_teach\_ed\_website<sup>b</sup>  
 m4d\_start\_teach  
 m4d\_start\_teach\_this<sup>b</sup>  
 m4d\_post\_other  
 m4d\_educ\_training\_other<sup>b</sup>

M1 Toilet covered pit  
 M1 Toilet uncovered pit  
 M1 Other – toilet  
 M1 Other – latrine  
 M1 Other Source of water  
 M1 Other types of road  
 M1 For satellite schools, how far is the school from Central and satellite? (km)  
 M1 Comments 1  
 M1 If Yes, did you discuss the information with:  
 M1 What's the school type (Other)  
 M1 For each term *n* last year, when did the school open (day/month/year)  
 M1 For each term *n* last year, when did the school close (day/month/year)  
 M1 Comments 2  
 M2A Comments 1  
 M2B Comments 1  
 M2B Comments 2  
 M2 How many computer rooms?  
 M4 How many minutes late were you? (If not late, write 0).  
 M4 Teacher name  
 M4 Comments: 1  
 M4 Comments: 2  
 M4 Comments: 3  
 M4 Comments: 4  
 M4 Did the teacher call pupils by name while teaching?  
 M4 How many pupils did the teacher call by name?  
 M4 Did the teacher use a video projector?  
 M4 Did the teacher use a projector screen?  
 M4 Did the teacher use an educational software or CD?  
 M4 Did the teacher use a video game?  
 M4 Did the teacher use a video?  
 M4 Did the teacher use an educational website?  
 M4 What year did you begin teaching?  
 M4 What year did you begin teaching at this school?  
 M4 Other Position Specify  
 M4 Other Training Specify

#### School management/finances (M3)<sup>b</sup>

q3a\_1a  
 q3a\_3a  
 q3a\_13a

M3 respondent's position - Other (Specify)  
 M3 respondent's highest degree - Other (Specify)  
 M3 respondent's reason for not continuing in education, if had to choose profession again

q3a_17[x]/a	Main constraints to the proper functioning of this school – Other_[x] (Specify)
q3b_5a	In case of absence of teachers, specify other measures taken for students in these cases
q3b_8a	Parent Association: Number of contract teachers (pre-school)
q3b_8b	Parent Association: Number of contract teachers (primary school)
q3b_11aa_[x]	Who is mainly responsible for taking decision/action [x]? - Other (Specify)
q3c_3a	For what reasons the last inspection had taken place? - Other (Specify)
q3d_2a	What are the main components of the «projet d'établissement»? - Other (Specify)
q3d_7	The Council (CGE) is composed of how many members?
q3d_8	How many members are women?
q3d_9	How many members of the CGE are members of the community (parents, etc.)
q3d_13	«Projet d'établissement»: for a duration of how many years? - Other (Specify)
q3d_14a	Who participates in the development of this «projet d'établissement»? - Other (Specify)
q3d_15aa	What is the frequency of parents' meetings in your school? - Other (Specify)
q3d_19aa	How does the school share financial information with parents and the community? - Other (Specify)
q3d_21a	How does the school share information on purchase of educational materials and equipment with parents and the community? - Other (Specify)
q3d_23a	How often do your school place student information about Massar? - Other (Specify)
q3d_24a	Who takes care of placing the information about Massar within the school? - Other (Specify)
q3d_25a	How does M3 respondent use Massar in reforms and modifications in school practices? - Other (Specify)
q3e_2a	Financial resources (other than wages) received from the Government: Other government structure (to be specified)
q3e_6a0l_[n]	Name of non-governmental organization [n] that financially supported the school
q3e_6ba_[n]	Non-governmental organization [n]: type of organization - Other (Specify)
q3e_6c_[n]	Amount of the financial support provided by non-governmental organization [n]
q3e_6d_[n]	Value of the financial support (in-kind) provided by non-governmental organization [n]
q3e_6ea_[n]	What is the most important in-kind support provided by organization [n]? - Other (Specify)



### Teacher level dataset

m2a_name	M2a First and last names
m2a_post_other	M2a Specify - Position in the school
m2a_educ_training_other <sup>b</sup>	M2a Other - Highest level of educational training (certificate) attained
m2a_contract_other <sup>b</sup>	M2a Specify - Contract Status
m2a_comments	M2a Module comments
m2b_fv_absence_r_other <sup>b</sup>	M2b Specify - Reason for absence in first visit
m2b_name	M2b First and last names
m2b_born_here	M2b Were you born in this [geographical area]?
m2b_post_other <sup>b</sup>	M2b Specify - Position in the school
m2b_educ_level_other <sup>b</sup>	M2b Other specify
m2b_observed_class_other <sup>b</sup>	M2b Specify - How often does the admin/HT observe your classroom?
m2b_observed_class_last <sup>b</sup>	M2b In what year did the last visit of the inspector?
m2b_perf_area_c11 <sup>b</sup>	M2b Performance areas commented on other (specify)
m2b_unpaid_claims_1	M2b Do you have any other unpaid claims? (1)
m2b_unpaid_claims_2	M2b Do you have any other unpaid claims? (2)
m2b_unpaid_claims_3	M2b Do you have any other unpaid claims? (3)
m2b_unpaid_claims_4	M2b Do you have any other unpaid claims? (4)
m2b_unpaid_claims_other	M2b Other unpaid claims?
m2b_comment1	M2b Module 2B Comments: 1
m2b_civil_status	M2b What is your civil status?
m2b_spouse_employment	M2b What is the employment status of your spouse?
m2b_comment2	M2b Module 2B Comments: 2
m6_date_day	M6 Date (day)
m6_name	M6 Teacher name (Official first, middle and last names)
m6_date	M6 Date (Day/Month/Year)
m6_post_other	M6 Specify - What is your position at this school
m6_enum_name	M6 Enumerator(s) name
m6_comments	M6 Comments
teach_a1	3a) Tell the pupils what the aims of the lesson
teach_a2	3a) Tell the pupils what the aims of the lesson
teach_b1	3b) What specific learning outcomes do you want the pupils to achieve? (list 2 I
teach_b2	3b) What specific learning outcomes do you want the pupils to achieve? (list 2 I
teach_c1	3c) Write out two questions that you will ask to pupils to show that they have (...)
teach_c2	3c) Write out two questions that you will ask to pupils to show that they have (...)
teach_d1	3d) Write out a question that you will ask pupils to show that they can apply (...)
teach_e_i	3e) What can or should the government do about road accidents?

teach\_e\_ii

teach\_e\_iii

teach\_f\_i

teach\_f\_ii

teach\_f\_iii

teach\_d2

swot\_as1

swot\_as2

swot\_as3

swot\_aw1

swot\_aw2

swot\_aw3

swot\_bs1

swot\_bs2

swot\_bs3

swot\_bw1

swot\_bw2

swot\_bw3

swot\_as4

swot\_aw4

swot\_bs4

swot\_bw4

eval\_a

eval\_c1

eval\_c2

eval\_c3

c\_name

c\_age

c\_gender

c\_post

c\_contract

c\_fulltime

c\_educ\_level

c\_educ\_training

Id

3e) What can or should the government do about road accidents?

3e) What can or should the government do about road accidents?

3f) Why is it difficult?

3f) Why is it difficult?

3f) Why is it difficult?

3d) Write out a question that you will ask pupils to show that they can apply wh

4a. Strength1. short assessment of XXX's letter

4a. Strength2. Short assessment of XXX's letter

4a. Strength3. Short assessment of XXX's letter

4a. Weakness 1. Short assessment of XXX's letter

4a. Weakness 2. Short assessment of XXX's letter

4a. Weakness 3. Short assessment of XXX's letter

4b. Strength 1. Short assessment of YYY's letter

4b. Strength 2. Short assessment of YYY's letter

4b. Strength 3. Short assessment of YYY's letter

4b. Weakness 1. Short assessment of YYY's letter

4b. Weakness 2. Short assessment of YYY's letter

4b. Weakness 3. Short assessment of YYY's letter

4a. Strength4. Short assessment of XXX's letter

4a. Weakness 3. Short assessment of XXX's letter

4b. Strength 4. Short assessment of YYY's letter

4b. Weakness 4. Short assessment of YYY's letter

5a) What is the class average for English?

5c) Comment on these three learners (1)

5c) Comment on these three learners (2)

5c) Comment on these three learners (3)

Name (Consolidated)

Age (Consolidated)

Gender (Consolidated)

Position (Consolidated)

Contract (Consolidated)

Full-time/Part-time (Consolidated)

Education Level (Consolidated)

Education Training (Consolidated)

Id IRT

#### Child level

emis\_code

m5s0\_date\_d

m5s0\_date\_m

m5s0\_date\_y

m5s0\_teacher\_n

School EMIS Code

M5 Date (day)

M5 Date (month)

M5 Date (year)

M5 Teacher name

m5s0_1st_enumerator_n	M5 First Visit Enumerator (name)
m5s0_1st_enumerator_c	M5 First Visit Enumerator (code)
m5s0_2nd_enumerator_n	M5 Second Visit Enumerator (name)
m5s0_2nd_enumerator_c	M5 Second Visit Enumerator (code)
m5s0_region_c	M5 Region Code
m5s0_region_n	M5 Region Name
m5s0_district_c	M5 District code
m5s0_district_n	M5 District name
m5s0_ward_c	M5 Ward code
m5s0_ward_n	M5 Ward name
m5s0_village_c	M5 Village/street code
m5s0_village_n	M5 Village/street name
m5s0_school_n	M5 School name
m5s0_school_c	M5 School code
m5s0_teamleader_result	M5 Team leader result
m5s0_supervisor_result	M5 Supervisor result
m5s0_enumerator_c	M5 Enumerator Name
m5s0_county_n	M5 County name
m5s0_county_c	M5 County code
m5s0_division_n	M5 Division name
m5s0_division_c	M5 Division code
m5s0_parish_n	M5 Parish name
m5s0_parish_c	M5 Parish code
m5s0_date_dmy	M5 Date (Day/Month/Year)
m5s0_province_n	M5 Province name
m5s0_IDEN_n	M5 IDEN Name
m5s0_pupils_mod4_why <sup>b</sup>	M5 If only some or none students selected from classroom observed in mod 4, why?
m5s1_first_name	M5 Pupil's First name
m5s1_pupil_c <sup>b</sup>	M5 Pupil code
m5s1_age	M5 Age
m5s1_lang1_teacher_thisyr_n	M5 Name of your language 1 teacher this year
m5s1_lang1_teacher_thisyr_c <sup>b</sup>	M5 Code from module 2A
m5s1_lang2_teacher_thisyr_n <sup>b</sup>	M5 Name of your language 2 teacher this year
m5s1_lang2_teacher_thisyr_c <sup>b</sup>	M5 Code from module 2A
m5s1_math_teacher_thisyr_n	M5 Name of your Math teacher this year
m5s1_math_teacher_thisyr_c <sup>b</sup>	M5 Code from module 2A
m5s1_inschool_lastyr	M5 Were you in this school last year
m5s1_lang1_teacher_lastyr_n	M5 Name of your language 1 teacher last year
m5s1_math_teacher_lastyr_n	M5 Name of your Math teacher last year
m5s1_lang1_teacher_lastyr_c <sup>b</sup>	M5 Code from module 2A
m5s1_lang2_teacher_lastyr_c <sup>b</sup>	M5 Code from module 2A
m5s1_math_teacher_lastyr_c	M5 Code of your Math teacher last year
m5s1_agreetoparticipate	M5 Do you agree to participate in this exercise

m5s1_time_starttest_hr	M5 Time started the test (HR)
m5s1_time_starttest_min	M5 Time started the test (MN)
m5s1_time_endtest_hr	M5 Time ended the test (HR)
m5s1_time_endtest_min	M5 Time ended the test (MN)
m5s1_name	M5 Pupil's name (first and last)
m5s1_lang2_teacher_lastyr_n	M5 Name of your language 2 teacher last year
m5s1_master_lastyr_n	M5 Name of your Master last year
m5s1_master_thisyr_n	M5 Name of your Master this year? (name)
m5s1_school_n	M5 School name
m5s1_teacher_CE2_n	M5 Name of teacher (CE2)
m5s1_teacher_CE1_n	M5 Name of teacher (CE1)
m5q12a	M5 District name
m5q13a	M5 County name
agree_participate	Do you agree to participate in this exercise
uniqueid	

#### Time on task

date_day	Date (Day)
date_month	Date (Month)
date_year	Date (Year)
date_comb <sup>b</sup>	Date (Day/Month/Year)
observation_end	Time observed (Constructed)
comments	Module 4a Time on Task comments

Note: b. Additional variables anonymized for the 2016 Morocco's Education SDI.

### 3.3 Recoding of variables

Some variables represent a risk of reidentification because of its composition and distribution. In order to share its contents, it is necessary to recode them in categories. The recoded variables cover both continuous (e.g. age, number of employees, etc.) and specific categorical variables in which certain categories are too scarce (e.g. position in the establishment, etc.). A detailed list of the recoded variables and the changes applied is listed in Table C.

The recoding consists in trimming tails of distribution or top recoding (e.g. year the school opened: every school opened before 1950 is grouped), transforming continuous values into ranges (e.g. using decades instead of years), and/or broadening categories to group possible answers (e.g. "Owner/Director/Head teacher/Principal/Deputy head teacher" into one value). Lastly, rather than grouping the whole distribution of values into categories, few variables' values are censored or anonymized when their specificity/uniqueness (e.g. excessively large values, rare values, etc.) might allow re-identification.

As part of the recoding process, we ensure that all identifiers have no value labels that could

contain specific information that leads to the reidentification of schools, teachers or pupils.

Table C. Overview of recoded variables

Recoded Variable	Description	Recode
<u>School level</u>		
m1_resp_post	M1 Which Position do you occupy in this facility?	1-3 Owner/Director/Head teacher/Principal/Deputy head teacher as grouped in one category; others are kept the same
m1_school_type	M1 What's the school type	Suppress 3 schools with different classification to missing
m1_school_year	M1 When did this school begin operating	Years transformed into decades [1950-2010], all years below 1950 are grouped
m1_travelttime_govt	M1 Approximate traveling time from school to local govt education office (in min	Grouped in 30 min intervals, [0-360+], all above 360 are grouped
m1_days_in_session	M1 What was the actual number of days during which school was in session in the	Grouped in 10 days intervals, starting less than [100-250+], all below 100 are grouped and above 250 into another
m1_resp_age	M1 What is your age?	Transformed into decades [30s-60+], all below 30 are grouped and above 60 into another
m1_road_distance	M1 What is the distance between the school and the asphalt road?	Transformed in 3 categories: less than 100, 100-1000, more than 1000. Values like 99, 999, are considered missing values
m1_delegation_dist	M1 What is the distance in kilometers between school and delegation?	Transformed in 2 categories: 10 or less, more than 10
m1_water_time_dist	M1 Avg time to go and come back to main source of water (including avg waiting t	Transformed in 3 categories: less than 15, 16-60, 61+
m1_travelttime_govt_h	M1 Approximate traveling time from school to local govt education office (Hours)	Transformed in 3 categories: 0, 1-3, 4 or more
m1_travelttime_govt_m	M1 Approximate traveling time from school to local govt education office (Minute	Transformed in 4 categories: 0-15,16-30, 31-45, 46-60
m1_strike_days <sup>c</sup>	M1 Stikes (days)	10 or more days were grouped together
m1_days_cancelled <sup>c</sup>	M1 How many school days were cancelled last year?	10 or more days were grouped together
m2_n_teachers	M2 How many teachers work in this school?	Transformed in 5 categories: 0-10, 11-20, 21-30, 31-40, 41+
m2_n_non_teachers	M2 How many non-teaching staff work in this school?	Transformed into a dummy: 0, 1 or more
m4d_age	M4 Age	Transformed into decades [30s-60+], all below 30 are grouped and above 60 into another

School  
management/finances  
(M3)<sup>c</sup>

q3a_4	In what year did [M3's respondent] start working in the field of education (teaching and or direction)?	Years transformed into decades [1990-2010], all years below 1990 are grouped together
q3a_6	Since what year does [M3's respondent] head this school?	Years transformed into decades [1990-2010], all years below 1990 are grouped together
q3a_8	In what year did [M3's respondent] get his/her first director position?	Years transformed into decades [1990-2010], all years below 1990 are grouped together
q3b_6	In 2014/2015, what is the number of teachers with whom you held an individual meeting to assess their performance?	Transformed in 5 categories: 1-4, 5-9, 10-14, 15-19, 20+ teachers. Observations recorded with "0" (no teachers assessed) are maintained untouched.
q3c_2	How many inspection visits did the school receive in the last 5 years?	Transformed in 5 categories: 1-4, 5-9, 10-14, 15-19, 20+ visits. Observations recorded with "0" (no inspection visit) are maintained untouched.
q3c_7	During the 2014/15 school year, how many tours of teacher inspections did you receive? (ie, inspections during which teachers were observed in class)?	Transformed in 5 categories: 1-4, 5-9, 10-14, 15-19, 20+ visits. Observations recorded with "0" (no inspection visit) are maintained untouched.
q3c_8	During the 2014/15 school year, what is the number of teachers who received an inspection visit?	Transformed in 5 categories: 1-4, 5-9, 10-14, 15-19, 20+ teachers. Observations recorded with "0" (no teachers visited) are maintained untouched.
q3d_4	How many times has the CGE gathered during the 2015/2014 school year?	Transformed in 5 categories: 1-4, 5-9, 10-14, 15-19, 20+ meetings. Observations recorded with "0" (no meetings) are maintained untouched.
q3d_16	How many general meetings (open meetings) did the school have with parents/community?	Transformed in 5 categories: 1-4, 5-9, 10-14, 15-19, 20+ meetings. Observations recorded with "0" (no meetings) are maintained untouched.
q3e_2_1b	Amount of financial resources (other than wages) received by the provincial delegation	Grouped in five dh 10,000 intervals, [1-40,000+]; all above dh 40,000 are grouped. Observations recorded with "dh 0" (no transfers) are maintained untouched.

q3e_2_2b	Amount of financial resources (other than wages) received by other government structure	Grouped in five dh 10,000 intervals, [1-40,000+]; all above dh 40,000 are grouped. Observations recorded with "dh 0" (no transfers) are maintained untouched.
q3f_1c_[x]	Annual average amount (in DH) paid to the school per student as contribution for <i>[item X]</i>	Amounts grouped into categories following an incremental-pattern interval: starting with an upper (exclusive) limit of dh 10 for the first category, each successive category's upper limit is multiplied by 2; all equal or above dh 10,240 are grouped together. Observations recorded with "dh 0" (no contributions) are maintained untouched.
q3e_7	For the 2014-2015 budget year, what were the total expenses of the school?	Grouped in five dh 10,000 intervals, [1-40,000+]; all above dh 40,000 are grouped. Observations recorded with "dh 0" (no expenses) are maintained untouched.
q3e_8	What is the balance at the end of the 2014-2015 budget year?	Grouped in five dh 10,000 intervals, [1-40,000+]; all above dh 40,000 are grouped. Observations recorded with "dh 0" (no balance) are maintained untouched.

#### Teacher level

m2a_post	M2a Position in the school	1-3 Owner/Director/Head teacher/Principal/Deputy head teacher as grouped in one category; others are kept the same
m2a_age	M2a Age	Transformed into decades [30s-60+], all below 30 are grouped and above 60 into another
m2b_years_teach <sup>c</sup>	M2b Number of years spent in the education sector	less than 10, 10-19, 20-29, 30+
m2b_post	M2b Position in the school	1-3 Owner/Director/Head teacher/Principal/Deputy head teacher as grouped in one category; others are kept the same
m2b_start_teach	M2b Since what year have you been teaching?	Years transformed into decades [1980-2010], all years below 1980 are grouped together
m2b_age	M2b Age	Transformed into decades [30s-60+], all below 30 are grouped and above 60 into another
m2b_start_teach_this	M2b Since what year do you teach in this school?	Transformed into 2 categories: before 2010, 2010-2019
m6_age	M6 Age	Transformed into decades [30s-60+], all below 30 are grouped and above 60 into another



M6_educ_training_yr <sup>c</sup>	M6 When did you complete your highest level of teacher training?	Years transformed into decades [1960-2010], all years below 1980 are grouped together
m6_years_teach	M6 Number of years teaching	less than 10, 10-19,20-29, 30+
m6_post	M6 What is your position at this school	1-3 Owner/Director/Head teacher/Principal/Deputy head teacher as grouped in one category; others are kept the same
c_age	Age (Consolidated)	Transformed into decades [30s-60+], all below 30 are grouped and above 60 into another
c_post	Position (Consolidated)	1-3 Owner/Director/Head teacher/Principal/Deputy head teacher as grouped in one category; others are kept the same

Note: c. Additional variables recoded for the 2016 Morocco's Education SDI.