Online Tool for Visualization of the MLSS Data

Proposal / Inception Report

Andres Arau

November 2021

Contents

[Summary 3](#_Toc88579612)

[Data Management System of the MLSS Survey 5](#_Toc88579613)

[1. Assessment of the Raw Data 5](#_Toc88579614)

[2. Harmonization System 7](#_Toc88579615)

[3. Harmonized Data 10](#_Toc88579616)

[4. Derivation System 11](#_Toc88579617)

[5. Derived Data 13](#_Toc88579618)

[Section 2: Suggested Data Management for the Online Tool 15](#_Toc88579619)

[Section 3: Elements and Content of the Online Tool 18](#_Toc88579620)

[Section 4: Software and tools 21](#_Toc88579621)

[Section 5: Work Plan 22](#_Toc88579622)

[Annex 1: Variable names and labels for the derived datasets 23](#_Toc88579623)

**List of Figures**

[Figure 1. Description of the MLSS Data Management System. 5](#_Toc88579624)

[Figure 2. Number of Raw Databases by Folder and Sub-Round. 6](#_Toc88579625)

[Figure 3. Consistency of dataset names across rounds 7](file:///C:\repositaries\1.work\MLS_concept\report\concept_note_AA.docx#_Toc88579626)

[Figure 4. Number of Harmonized Databases by Folder and Sub-Round 10](#_Toc88579627)

[Figure 5. Consistency of Harmonized Databases Across Rounds. 11](#_Toc88579628)

[Figure 6. Suggested Data Flow for the Online Tool. 15](#_Toc88579629)

[Figure 7. Data Architecture and Relationship of Tables for the Online Tool. 17](#_Toc88579630)

[Figure 8. Proposed Design of the Homepage. 18](#_Toc88579631)

[Figure 9. Proposed Design of the Visualization of Indicators. 19](#_Toc88579632)

[Figure 10. Proposed Design of the Background Information. 19](#_Toc88579633)

[Figure 11. Suggested Design for the Administrators Page. 20](#_Toc88579634)

[Figure 12. Workplan. 22](#_Toc88579635)

**List of Tables**

[Table 1. Software and Tools 21](#_Toc88579636)

[Table 2. Variable Names, Labels, and Fomat of the Derived Data at the School Level. 23](#_Toc88579637)

[Table 3. Variable Names, Labels, and Format of the Derived Data at the Student Level. 24](#_Toc88579638)

[Table 4. Variable Names, Labels, and Format of the Derived Data at the Teacher Level. 34](#_Toc88579639)

# Summary

The objective of this Inception Report is to provide a clear and detailed plan to set up an online dashboard to enable a friendly visualization of key MLSS indicators over time, and to enable browsing of these indicators at the school, teacher, and student level across regions and districts. The report is divided into 5 sections: **Section 1** describes the data management system of all the rounds of the longitudinal survey. Based on the findings of this section, **Section 2** proposes a data management system for an efficient implementation of the dashboard; **Section3** describes the elements of the dashboard to enable visualisation of indicators, present background information of the MLSS, and to facilitate administrators to add data for future rounds of the survey; **Section 4** Outlines the software that is recommended for the setup of the online tool; and **Section 5** presents a workplan for the development of the dashboard.

As it is understood from the ToRs developed by the World Bank, the online tool should enable a friendly visualization of key indicators of the MLSS but should also facilitate administrators to add additional data for existing indicators from future rounds.

The findings of the Data Management System show that although the current system achieves a good degree of harmonization and creates indicator for each round, it is still a system made of a complex interrelation of scripts that depend on human inspection of the data, that uses costumed functions, and that relies a great deal on manual coding to achieve its goals. This system, that is described in **Section 1**, applies over 40 scripts to more than 600 datasets to clean the data as the indicators are created. The main conclusions of this section are: 1) it is unlikely that the cleaning system could be separated from the derivation one, and most importantly 2) it would be very risky and complex for an online tool to replicate the whole process in an automatic and systematic process with out detailed human inspection.

For the reasons mentioned above, it is suggested that the initial input of the dashboard should be the derived data that contains the estimated indicators at the school, teacher, and student level. The advantages of using the derived data as the online tool’s input is that it increases the chances of future data being added into the tool and would reduce the amount of time spent in reproducing a complex data management system. Therefore, the recommendation is for the current data management system to continue happening off-line and for the online tool to execute the process of formatting and visualising the data. This process is detailed in **Section 2**.

The online tool will manipulate the derived data and will transform its format into a data architecture of relational tables that will enable the tool to perform an efficient analysis, visualisation, and integration of future rounds of the MLSS. This architecture, presented in **Section 2** has multiple purposes. First, it will allow to effectively organize the data, will enable a faster interaction with secondary data sources (like shapefiles), and more importantly, it will allow to have control over the harmonization of the data over the different rounds of the MLSS.

The suggested design of the tool, presented in **Section** **3**, is a simple but efficient user interface that enables an easy navigation of the dashboard. Basically, the proposed design is made of 6 web pages: a home page that introduces the MLSS and the dashboard, three pages to enable the browsing of indicators by region/district and level (school, teacher, student), a background section that has links to relevant sources, and a password protected page for the administration to enter data of future rounds into the system.

Finally, **Section 4** presents the software that will be used for the development of the dashboard and **Section 5** shows the workplan for a successful deliverable of the milestones.

# Data Management System of the MLSS Survey

The objective of this section is to map and assess every process of the MLSS data management system and to identify the likelihood of an online dashboard to atomate these processes. To do so, the MLSS data management system is analysed from its starting point, which is the raw data of the Survey, until the creation of indicators at the school, teacher, and student level. The following subsections describe each of these processes and identify possibilities and challenges for the online tool to replicate this complex system. After having conducted this analysis, a division of steps of process between off-line and on-line process is suggested in Section 2.

Because the MLSS is a Longitudinal Survey (there are multiple rounds of the data), the aim of this system is to clean the raw data, conduct necessary manipulations, and create indicators that inform about the school conditions, practices, and learning outcomes that can be traced over time. This complex system that is replicated for each round of the MLSS is summarised by Figure 1 below.

Figure 1. Description of the MLSS Data Management System.

Diagram

Description automatically generated

## Assessment of the Raw Data

This section explores the raw data, or the initial stage of the data management system. This data is stored in the 01\_Firm directory directory (throughout the document, the words folder and directory are use indistinctly). Each sub-round 01\_1AB\_2016 and 02\_1C\_2018, 01\_2A\_2018, and 01\_2B\_2019 has its own dedicated directory where the raw data is stored: 05\_Data/{round}/ {sub-round /01\_Firm.

To try to understand the volume of data and its consistency across sub-rounds, this assessment mapped all the directories within each sub-round’s raw data and counted the number of data files (\*.dta) stored in each. Apart from counting the number of data files, a comparison between folder architecture between rounds was conducted to try to have an estimate of the level of consistency between rounds. Thus, as can be seen in the Figure 2 below, that displays all the folders that exist in the raw data directory and counts the number of files that exist in each sub-round, there are 662 Stata datasets (\*.dta) across all sub-roundsof Baseline and Midline, 183 in sub-round 1AB\_2016 (1B), 194 in sub-round 1C\_2018 (1C), 54 (2A\_2018), and 324 (2B\_2019).

From that Figure, it is noticeable that there are some folders that do exist in one sub-round but do not exist in the rest. For example, the folders HT2\_NONC1, TEACHER\_NONC1, COMMM\_TRACK only exist in round 2B\_2019.

Another example is the folder HT2 that exists in all the sub-rounds but in the 2B\_2019. In summary, Figure 2 shows that there is a great number of databases saved in the raw data (01\_Firm) and that there is an important level of discrepancy between the structure of the directories between rounds.

Figure 2. Number of Raw Databases by Folder and Sub-Round.

A picture containing chart

Description automatically generated

Moreover, to assess the level of data heterogenicity between rounds, Figure 3 below displays all the databases stored in the folder 01\_Firm of the Baseline and Midline sub-rounds and colour codes if the file exists in each round. The main objective of this Figure is to map which datasets exist in sub-round 1B, sub-round 1C, sub-round 2A, sub-round 2B and in all the rounds.

Chart

Description automatically generatedAs part of this assessment, Figure 3 attempts to show the volume of data contained by each round and the difference in terms naming across directories.

Figure 3. Consistency of dataset names across rounds

In summary, Figure 3 shows that from the 324 datasets that exist in Baseline’s round 1B and round 1C, only 53 (16.4%) exist in both sub-rounds. Moreover, from the 378 data files stored in Midline’s raw data directory, **none** exist in both sub-rounds.

This finding shows that there’s is a big inconsistency in file names between sub-rounds. And that this is even more notorious for the midline raw data.

The original chart in its best format to inspect it can be seen by clicking this [link](https://github.com/araupontones/MLS_concept/blob/main/report/plots/Harmonization/databasesAcrossRounds.png).

The conclusion of the analysis of the raw data is that there is a high degree of discordance in terms of folder architecture and names of the datasets across rounds. Although the analysis does not inspect the name of the variables, it is likely that the discordance is even deeper at that level. Therefore, to make it possible to analyse indicators over time and to have a friendlier system, an harmonization process is required. Therefore, the data management system of the MLSS includes a process to harmonize the folder architecture and naming protocols between rounds. This harmonization process is explained in the next section of this report.

## Harmonization System

The main objective of this system is to solve the differences in terms of folder architecture, file and variable naming across rounds which was partially described in the previous section.

The harmonization system is a very important step in the Data Management System because without it, it would not be possible to conduct a longitudinal analysis of the data. The main input for this process is the data stored in 01\_Firm (described in the previous section) and the output for each round is stored in the 03\_Harmonized folder. The scripts/dofiles (throughout the document, the words dofile and script are use indistinctly) that run the harmonization process are stored in the 06\_Coding/{round}/{sub-round}/02\_Harmonization directory.

Although there are some small differences between rounds, the harmonization approach is very similar for all, below is a detailed description of each:

1. All the paths to key folders are defied in a dofile called 0\_00\_00\_Master\_NEW.do. Since this dofile defines the paths, it must be ran before starting or running any other script in the system. Below is an example of how the paths are defined:

|  |
| --- |
| Box 1: Extract of the definition of paths in the Master dofile of the harmonization system    global bl16 "${root}/01\_Baseline/01\_1AB\_2016"  global bl18 "${root}/01\_Baseline/02\_1C\_2018"  global ml2a "${root}/2. Midline/01\_2A\_2018"  global ml2b "${root}/2. Midline/01\_2B\_2019"  global el\_a "${root}/03\_Endline/01\_3AB\_2021"  global el\_b "${root}/03\_Endline/02\_3B\_2021" |

Thus, as it can see above, there is a dedicated and predefined directory for each round. Thus far, in the system, there is a folder for 01\_Baseline, 02\_Midline, 03\_Endline, and 04\_Longitudinal. Each of these rounds is subdivided into more directories that are based on sub-rounds (e.g., 01\_1AB\_2016, 02\_1C\_2018, 01\_2A\_2018, and 01\_2B\_2019).

In the harmonization system, the directory that contains the raw data is defined as dta\_raw and points to 01/\_Firm; the path to the intermediate data is defined as dta\_inter and points to 02\_Cleaned, and the path for the harmonized data is dta\_har and points to 03\_Harmonized. Thus, for example, if the raw data from baseline is to be called, the user should point to: $bl16/$dta\_raw.

1. Within each sub-round, there is a specific dofile to harmonize each module of the dataset. This system is composed of 13 dofiles:

* Append\_16BLto18BL.do
* 01. CFO\_recode.do
* 02. CLO\_recode.do
* 03. HT1\_recode.do
* 04. HT2\_recode.do
* 04. HT2\_recode\_other.do
* 05. HT3\_recode.do
* 06. MG\_PTA\_SMC\_recode.do
* 07. LAT\_recode.do
* 08. LATAK\_recode.do
* 09. student\_recode.do
* 10. teacher\_recode.do
* 11. SFO\_recode.do

The most important file of this system is 00.Append because all the other scripts are ran from this *“master”* script. In simpler words, the 00.Append runs all the other scripts listed above to replicate the complete analysis flow. Apart from running all the scripts needed for the harmonization, this dofile creates two functions (or commands) that are used for all the other dofiles: mpss\_scvalues that harmonises the single select value labels, and mpss\_kickout that removes old options and recode them accordingly.

After defining these functions, the 00.Append dofile runs all the other dofiles to mainly: assign harmomized labels to the variables, and , export the data to the 03\_Harmonized directory.

Below is an example of the code used during harmonization process for the CLO data.

|  |
| --- |
| Box 2: Extract of the the harmonization process for the CLO data    use "$path2/CLO/CLO\_main.dta", clear  ren qnum qq1  ren moecode MOEcode  ren class\_id classid  ren q1\_name school\_name  ren q3 enum\_first\_visit1  ren q5 enumerator\_name2  ren q13 teachername  ren q14 teacherid  \* CB  gen sup\_date = date(q15a3, "DMY")  drop q15a3  ren q15a4 reasons  replace q7=. if q7==777 | q7==999  gen q7new=1 if q7>=10  replace q7new=0 if q7<10  replace q7new=. if q7==. |

As can be seen in the example above, even though there is some level of atomation, each dataset requires a manual (very specific) manipulation of the variable names for the harmonization to be effective. Apart from the manual manipulation, in the harmonization system there are some other customed functions that help with the consistency of the naming protocols between rounds. For example, in the dofile 02.CLO\_recode, there is a note that explains what the fre command does:

|  |
| --- |
| Box 3: Example of the use of customed commands to clean the data    \*Task 1:  \*i) the variable names in BL18 align with the BL18 instrument and  \*ii) value labels align.  \*Using fre commandreplace q7new=. if q7==. |

Thus, at it was shown in the description of the harmonization process, this system relies on a complex combination of customed functions and manual coding for the harmonization to be successful. The next section describes in detail the output of the harmonization process (the harmonized data).

## Harmonized Data

The output data from the Harmoization system is stored, for each sub-round, in 05\_Data/{round}/{sub-round}/03\_Harmonized. Based on an analysis of this folder, it was found that there is “harmonized data” for the two sub-rounds of baseline but there is no data for the midline sub-rounds[[1]](#footnote-1). There are, in total, 60 datasets between sub-rounds 1AB and 1C. From all these datasets, 43 (71.7%) exist in both rounds.

Figure 4. Number of Harmonized Databases by Folder and Sub-Round

Diagram

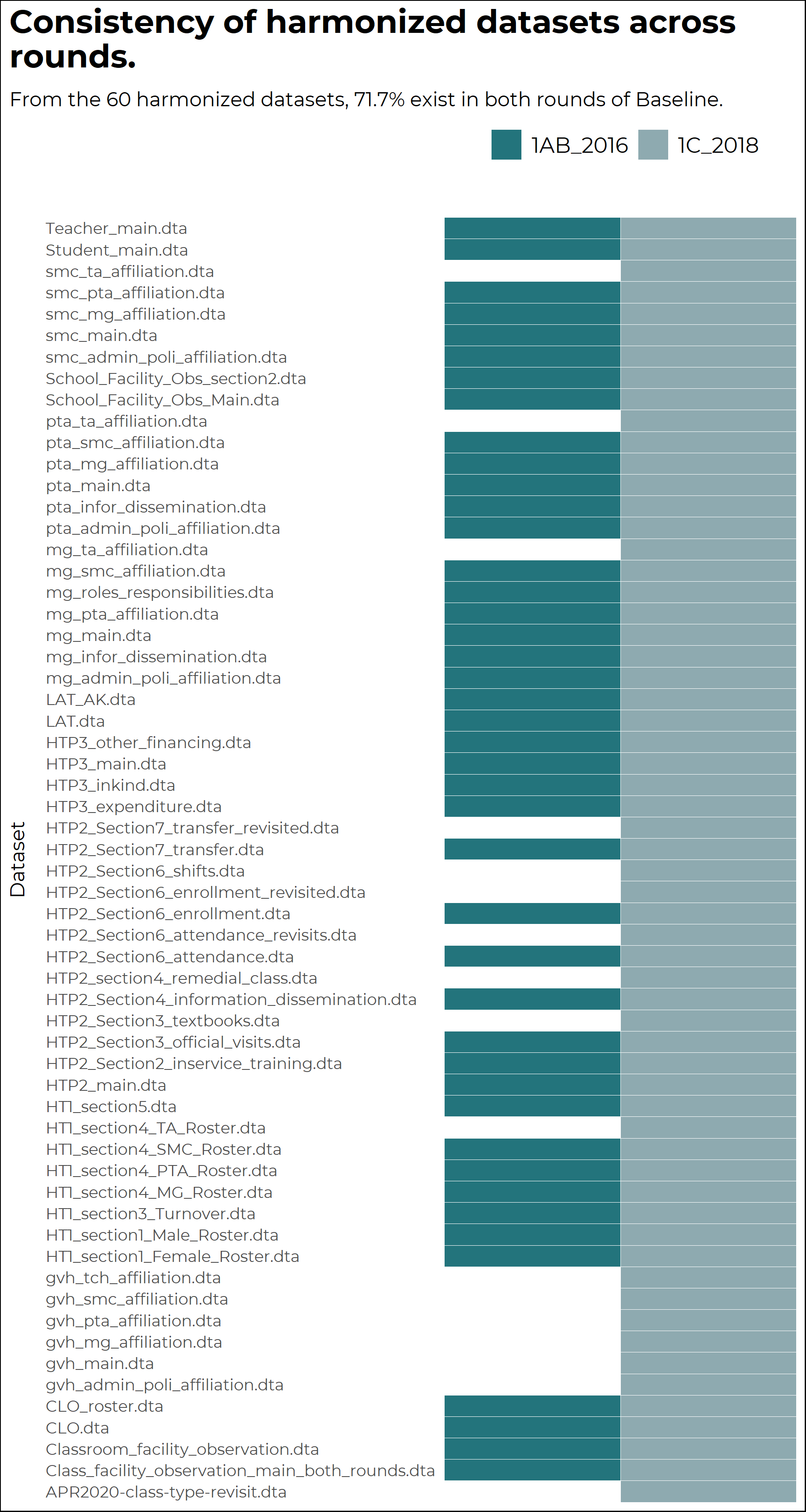
Description automatically generated

Figure 4 above shows the number of datasets that exist for each folder by each round in the harmonized directories. As it can be seen in the Figure, the module with the largest number of datasets is HT2 (20 stata data files), followed by HT1 (15 datasets). Although the Figure makes evident that the harmonization process does increased the homogenization of the folder architecture and the name of the files, there are still some inconsistencies and gaps between rounds: The number of datasets by module differs between rounds (see the case of HT2), not all the folders exist for both rounds (see the case of the GVH folder), and there is no harmonized data for any of the sub-rounds of the Midline round.

Figure 5 below shows that though the harmonization process considerably reduced the inconsistencies in terms of folder architecture and file naming, there are still some gaps between rounds. For example, the file smc\_ta\_affilitation.dta exists in sub-round 1C but does not exist in 1AB. In total there are 16 files present in round 1C that are missing in round 1AB. As it was described above, none of the “gvh” files exist in round 1AB. The analysis does not include the harmonized data from Midline because at the time of writing this report that data was not present in the project’s directory.

The original chart in an efficient format to be inspected can be access by accessing this [link](https://github.com/araupontones/MLS_concept/blob/main/report/plots/Harmonization/databasesAcrossRoundsHarm.png).

Figure 5. Consistency of Harmonized Databases Across Rounds.



Once the data is harmonized, the next step of the system is to create indicators at the school, teacher, and student level for each round. The findings of the inspection of this process are described in the next section.

## Derivation System

The derivation system is done at the round level (not at the sub-round). This system that is made of 27 dofiles, takes the data from the clean folder[[2]](#footnote-2), performs some cleaning like dropping duplicates and recoding variables, merges different datasets, creates indicators, defines a consistent naming and labelling, and exports the data to the Derived folder. In summary, the derivation cleans and prepares the data for analysis, creates intermediate datasets, estimates statistical weights, and finally creates indicators at the school, student, and teacher levels.

The main or master dofile for this system is 001\_master\_all\_indicators. This script reproduces the data flow of the derivation system by runninng in a sequence manner all the dofiles that create all the indicators:

* 000\_CFO-CLO-common-variables.do
* 000\_Records\_Available.do
* 001\_master\_all\_indicators.do

**These dofiles create intermediate datasets (data is exported to clean/Output\_dump):**

* 002\_distance\_quintile.do[[3]](#footnote-3)
* 003\_std\_offered.do
* 004\_enrollment\_by\_standard.do
* 005\_teacher\_roster\_list.do
* 006\_class\_size\_enrol\_unrestricted.do
* 007\_class\_size\_hc\_unrestricted.do
* 008\_PTR\_hc\_unrestricted.do
* 009\_school\_characterisitcs\_1.do
* 010\_enrollment\_calculation\_3\_periods.do
* 011\_attendance\_calculation\_3\_periods.do
* 012\_absence\_rate\_calculation\_3\_periods.do
* 013\_dr\_rr.do

**These dofiles create indicators at the school, student, and teacher level:**

* 014\_school\_characteristics\_2.do
* 015\_school\_characteristics\_3.do
* 016\_school\_characteristics\_1+2+3.do
* 017\_MLSS\_weights\_generation.do
* 018\_LAT\_IRT\_pct.do
* 019\_LAT\_IRT\_fitting\_raven.do
* 020\_student\_charcacteristics.do
* 021\_school\_characteristics\_final.do
* 022\_LAT\_AK\_score.do
* 023\_teacher\_characteristics.do

**Cleans the indicator names**

* 024\_label\_final\_indicators.do: Has more than 1,300 lines of code

Within each of these dofiles there is a considerable manual process to clean, recode, and merge the data. See example below from the 017\_MLSS\_weights\_generation.do:

|  |
| --- |
| Box 4: Example of code to generate the sample weights    \*population number of schools  gen tot\_sch\_strata=.  replace tot\_sch\_strata=248 if strata==7  replace tot\_sch\_strata=990 if strata==1  replace tot\_sch\_strata=1108 if strata==2  replace tot\_sch\_strata=1266 if strata==3  replace tot\_sch\_strata=503 if strata==4  replace tot\_sch\_strata=750 if strata==5  replace tot\_sch\_strata=550 if strata==6  bysort strata: gen num\_sch\_strata=\_N  gen wtg\_sch=tot\_sch\_strata/num\_sch\_strata  \*\*\* Student/teacher level weight  merge 1:1 MOEcode using "$bl\_dta\_clean/Output\_dump/school\_level\_indicators\_merged",keep(3)  drop \_m |

The above example shows the level of detailed and manual work necessary to create only one output of the derivation system which is, in this case, the sample weights of the baseline.

Moreover, as it is pointed out in the list of dofiles that are part of the system, there are some aspects that make this process very difficult to automate: **1**) Some scripts rely on customed reference files (like 02\_Data/962\_IE\_IDENTIFICATION\_COMPONENT\_101118\_complete\_infor") to produce outputs; **2)** The system is very intense for cleaning the data before it is ready to estimate indicators (the derivation process is both a cleaning and an indicator generator); **3)** Most, if not all, the dofiles require a manual inspection of the data before making cleaning decisions.

## Derived Data

The derivation system exports datasets that contain indicators at the school, teacher, and student level for each round (not sub-round) of the survey. At the moment of writing this report, there were derived datasets for the Baseline round only. Within the Baseline derived folder, there are 5 files - school\_level\_indicators\_final, school\_level\_indicators\_final\_c1, student\_level\_indicators\_final, student\_level\_indicators\_final\_c1, and teacher\_level\_indicators -.

The fact that there is more than one version of the indicators’ datasets serves as an example of the complexity of the system to automate the process. Moreover, it makes it confusing for the users of the data. A recommendation would be to only keep the latest version of the datasets in this directory.

The amount of data for each derived dataset is described below (this report is using the versions \_c1 for the school, and teacher level indicators).

:

* School derived dataset has 750 observations and 57 variables
* Students dataset has 18,512 observations and 501 variables
* Teachers dataset has 6,766 observations and 82 varibles

As can be seen in the tables that contain information about the detailed name of each variable and its label for each derived data set presented in Annex 1 **none of the derived datasets include geographic identifiers like region or district. Including these variables into these datasets is crucial for the online tool to aggregate the data at those levels**.

The findings of this section show that the steps to transform the data from its raw format to a clean version of estimated indicators at the school, teacher, and student level is a complex system that relies on some customed functions, a human inspection of the data, and in associations of different scripts that are heavily dependent on each other. Based on this findings, Section 2 proposes a feasible data management system of the online tool that allows an effective visualization of key MLSS indicators over time but that also a facilitates administrators to add additional data for existing indicators from future rounds of MLSS/other sources.

# Section 2: Suggested Data Management for the Online Tool

The understanding from the ToRs is that the dashboard should enable the visualization of key MLSS indicators over time, enabling the browsing of key indicators by different levels (school, teacher, and student). But also, that it must facilitate administrators to add additional data from future rounds of the survey. The later is very important for the decision of which data management approach the dashboard should follow.

As it was described in Section 1, the current data management system that transforms the raw data into indicators is a very complex system of over 40 scripts that although is replicable for the current sub-rounds, it is very unlikely that it could be transformed into an automatic system that allows adding raw or even harmonized data into it. Moreover, it was also described that the process to clean the data and to create the indicators is not separated but it is rather embedded into one (i.e., the data is cleaned as the indicators are created). Thus, in the current state of the system, it is not straightforward to separate the clean data from the derived one.

Based on this analysis, it is suggested that the input of the dashboard should be the derived version of the data. The version in which the indicators have been estimated at the different levels. If this advice is not approved, it is very likely that the online tool will fail and crash when new data from future rounds is added to it.

Figure 6. Suggested Data Flow for the Online Tool.

Diagram

Description automatically generated

Therefore, as it is displayed by Figure 6 that shows the suggested data management system of the online tool, it is advised to continue with the current off-line process and to input the data into the dashboard only when the indicators have been created. The suggested steps for an efficient and agile system are the following:

* 1. The World Bank’s team will transform the data from its raw version into indicators.
  2. The dashboard will have a system to allow inputting that data into the online system.
  3. Once that the data is inputted, the online tool will check for the consistency of the file and variable names. This step will be conducted by a specific script that will confirm that the inputted round of the data does not exist in the system yet and that the variable names are consistent with the previous rounds (this will assure that visualizations over time can be conducted). The list of indicators and their names will be defined by the Baseline derived data that already exists in the system. In the case of a future round having new variables, the online tool should identify that these were not part of previous rounds.

**3.a** In the case of the data not passing the test (if the names are not consistent with previous rounds), the online tool will inform the user about the problem, and this should be rectified by the off-line team.

**3.b** If the quality test is positive (if the names are consistent with previous rounds), the online system will use a script to transform the data into a format for effective visualisation and will append the data with the previous rounds (See Figure 7).

* 1. The online tool will integrate secondary data sources like shapefiles[[4]](#footnote-4) or population estimates (if necessary) to create customed visualization. This process will require a cleaning process to assure that the geographic variables of the data (region and district) match with the secondary data sources.
  2. A series of scripts will be developed for the user to interact with the data and for the dashboard to create customed charts. The creation of these charts will require an intense coding period where functions will be defined so they can run and respond based on user events or requests.

Figure 7 below shows the data architecture and the relationship of tables for the online tool to perform an efficient analysis, visualisation, and integration of future rounds of the MLSS.

Once the data is imported into the dashboard, the data will be transformed into different tables that will allow an effective interaction the components of the system. This architecture has multiple purposes. First, it will allow to effectively organize the data. Moreover, it will enable a faster interaction with secondary data sources (like shapefiles), and more importantly, it will allow to have control over the harmonization of the data over the different rounds of the MLSS.

As can be seen in Figure 7 below, the consistency of the data across rounds is very important for the correct interaction of the tables. This means, that to visualise key indicators across time it is necessary that the Parent Keys (PK or ID) of the datasets are consistent over time. For example, the IDs of the districts, regions, and schools should be consistent across rounds for the system to be able to aggregate the data and create a longitudinal format. The same applies for the name of the indicators, any change in the system will affect the capacity to trace that indicator over time, or to identify that the indicator is the same across rounds.

Thus, the main objective of introducing the architecture of the data in this report, is for the World Bank to be conscious about the importance of assuring the structure and naming consistency of the data between rounds of the MLSS. Not complying with this would make the system to fail to enable administrators to add new rounds of data and or to visualise indicators over time.

Figure 7. Data Architecture and Relationship of Tables for the Online Tool.

Diagram

Description automatically generated

It is suggested that during the first phases of the development of the online tool, the consultant will meet with the World Bank to agree on a naming protocol for the files and the names of the variables. A guide document with clear information about these will be produced so future rounds of the survey can follow the naming protocols.

The next section describes the suggested user interface and content of the online tool.

# Section 3: Elements and Content of the Online Tool

This section presents a sketch for the conceptual design of the online tool. This preliminary design intends to respond to the objectives of the tool as they are expressed in the ToRs:

* Enable visualization of key MLSS indicators over time (across rounds) and across space (region/district).
* Enable browsing of key indicators by level (school/teacher/student), topic (e.g. infrastructure, facilities, procedures, student profiles, learning outcomes),
* Additionally present background information on the MLSS and associated impact evaluations (provided by Task Team) including instruments, sampling, and interventions, Facility for administrators to add additional data for existing indicators from future rounds of MLSS/other sources,
* Facility for higher-level administrators to add new indicators in future, and
* Ability for administrators/’power users’ selected by administrators to extract visualizations for use in analytical outputs.

In terms of content, it suggested for the World Bank to define the final set of indicators that will be part of the online tool. The tables presented in Annex 1 can be used to facilitate the prioritization process. The ideal scenario, for a friendlier user interface would be to define the subset of indicators that are more likely to be of interest for the audience of the dashboard.

Thus, the content of the dashboard is expected to be formalized during the initial stages of the development and programming process (see work plan).

To improve the user experience of the online tool, it is suggested to design a basic but effective interface that enables an easy navigation of the dashboard. All the elements of the dashboard are presented below:

* + 1. **A home page** that briefly introduces the MLSS and the objectives of the online dashboard, this page will include the logos of the participating organizations and two buttons to direct the user to the dashboard and to its background information:

Figure 8. Proposed Design of the Homepage.

Graphical user interface, text

Description automatically generated

1. **Three separate pages to visualize indicators over time**, enabling the analysis to be conducted at the region and district level, and allowing the user to browse key indicators. These pages will present the data at the school, student, and teacher level respectively. Another possibility is to switch the hierarchical order of the attributes. For example, for the menu to have District and Region and for the dropdowns to allow filtering by level (school, teacher, student). This is suggested to be decided based on tests that will aim to understand which order is more effective for the experience of the users.

Figure 9. Proposed Design of the Visualization of Indicators.

A picture containing graphical user interface

Description automatically generated

1. **A page to present background information** of the MLSS. This information, that will be provided by the Task Team, can include things like Instruments, Sample, Relevant Links, etc. Apart from static text, this page can include links to external websites or other resources that can help to contextualize the importance and scope of the MLSS. The content of this page will be discussed at the beginning of the design phase.

Figure 10. Proposed Design of the Background Information.

Graphical user interface, application

Description automatically generated

1. Finally, a very important requirement of the online tool is that it should facilitate administrations to add data for existing indicators from future rounds of the MLSS. The suggested approach is to create **a specific page that is password protected**. Only users with the right set of credentials will be able to access this page. Once the password is validated, the administrators will be able to load more data into the system. To reduce the risk of the data not being in the right format, the administrator will be able to download a guide that will include the detail information about the necessary conditions that the data needs to meet to be accepted by the dashboard. These conditions, as it was explained in the section above, will be things like format of the file, variable names, etc.

Figure 11. Suggested Design for the Administrators Page.

Graphical user interface

Description automatically generated

# Section 4: Software and tools

This section introduces the software and tools expected to be used for the development and maintenance of the online tool:

Table 1. Software and Tools

|  |  |  |  |
| --- | --- | --- | --- |
| Software | Programming language | Process | Cost |
| Atom | HTML | Create the User Interface of the application | - |
| Atom | CSS | Create custom styling of the app | - |
| Atom | JavaScript | Create reactive events for the User Interface | - |
| RStudio | R | Manipulation and visualisation of the data | - |
| GitHub | Git | Storing and management of the data | - |
| Digital Ocean | Linux | Set up an online server to execute the back end of the app | USD 20 / monthly |
| GoDaddy | URL | Set up domain for the dashboard | UDF 10 /monthly |

As it can be seen in Table 1, most of the tools and software are open sources except for the configuration of the online server and the creation of a custom domain. Both costs are estimates but they could vary depending on the size of the application and the availability of the desired domain. It is expected that during the first phase of the programming, the World Bank will suggest a list of domains that could be used as the URL address of the online tool.

It is advised, to assure future sustainability and ownership of the application, that the payments of the software and tools will be made directly from the World Bank.

# Section 5: Work Plan

Based on the ToRs, it has been agreed that the consultant will work remotely and that the estimated duration is eight months. The appointment is on part-time basis for 60 contracted days to June 30, 2021.

The suggested workplan is presented in Figure 12 below.

Figure 12. Workplan.



# Annex 1: Variable names and labels for the derived datasets

Table 2. Variable Names, Labels, and Fomat of the Derived Data at the School Level.

| var\_name | Label | format |
| --- | --- | --- |
| school\_id | School ID | %10.0g |
| enrol\_tot\_bl | Baseline Total Enrolment | %9.0g |
| enrol\_lower\_tot\_bl | Baseline Lower Primary Enrolment | %9.0g |
| enrol\_PTR\_std1\_bl | Baseline enrolment in std1 / number of teachers in std1 | %9.0g |
| enrol\_PTR\_std1\_f\_bl | Baseline female enrolment in std1 / number of female teachers in std1 | %9.0g |
| enrol\_PTR\_std1\_m\_bl | Baseline male enrolment in std1 / number of male teachers in std1 | %9.0g |
| enrol\_PTR\_std2\_bl | Baseline enrolment in std2 / number of teachers in std2 | %9.0g |
| enrol\_PTR\_std2\_f\_bl | Baseline female enrolment in std2 / number of female teachers in std2 | %9.0g |
| enrol\_PTR\_std2\_m\_bl | Baseline male enrolment in std2 / number of male teachers in std2 | %9.0g |
| enrol\_PTR\_std3\_bl | Baseline enrolment in std3 / number of teachers in std3 | %9.0g |
| enrol\_PTR\_std3\_f\_bl | Baseline female enrolment in std1 / number of female teachers in std3 | %9.0g |
| enrol\_PTR\_std3\_m\_bl | Baseline male enrolment in std3 / number of male teachers in std3 | %9.0g |
| enrol\_PTR\_std4\_bl | Baseline enrolment in std4 / number of teachers in std4 | %9.0g |
| enrol\_PTR\_std4\_f\_bl | Baseline female enrolment in std4 / number of female teachers in std4 | %9.0g |
| enrol\_PTR\_std4\_m\_bl | Baseline male enrolment in std4 / number of male teachers in std4 | %9.0g |
| enrol\_PTR\_std5\_bl | Baseline enrolment in std5 / number of teachers in std5 | %9.0g |
| enrol\_PTR\_std5\_f\_bl | Baseline female enrolment in std5 / number of female teachers in std5 | %9.0g |
| enrol\_PTR\_std5\_m\_bl | Baseline male enrolment in std5 / number of male teachers in std5 | %9.0g |
| enrol\_PTR\_std6\_bl | Baseline enrolment in std6 / number of teachers in std6 | %9.0g |
| enrol\_PTR\_std6\_f\_bl | Baseline female enrolment in std6 / number of female teachers in std6 | %9.0g |
| enrol\_PTR\_std6\_m\_bl | Baseline male enrolment in std6 / number of male teachers in std6 | %9.0g |
| enrol\_PTR\_std7\_bl | Baseline enrolment in std7 / number of teachers in std7 | %9.0g |
| enrol\_PTR\_std7\_f\_bl | Baseline female enrolment in std7 / number of female teachers in std7 | %9.0g |
| enrol\_PTR\_std7\_m\_bl | Baseline male enrolment in std7 / number of male teachers in std7 | %9.0g |
| enrol\_PTR\_std8\_bl | Baseline enrolment in std8 / number of teachers in std8 | %9.0g |
| enrol\_PTR\_std8\_f\_bl | Baseline female enrolment in std8 / number of female teachers in std8 | %9.0g |
| enrol\_PTR\_std8\_m\_bl | Baseline male enrolment in std8 / number of male teachers in std8 | %9.0g |
| enrol\_PTR\_tot\_bl | Baseline Overall Pupil-Teacher Ratio | %9.0g |
| enrol\_PTR\_tot\_f\_bl | Baseline total female enrolment / total number of female teachers | %9.0g |
| enrol\_PTR\_tot\_m\_bl | Baseline total male enrolment / total number of male teachers | %9.0g |
| enrol\_PTR\_tot\_lower\_bl | Baseline Lower Primary Pupil-Teacher Ratio | %9.0g |
| enrol\_PTR\_tot\_upper\_bl | Baseline total upper class (std5-std8) enrolment / total number of teachers in u | %9.0g |
| enrol\_PTR\_tot\_lower\_f\_bl | Baseline total lower class (std1-std4) female enrolment / total number of female | %9.0g |
| enrol\_PTR\_tot\_upper\_f\_bl | Baseline total upper class (std5-std8) female enrolment / total number of female | %9.0g |
| PCR\_std1\_bl | Baseline | %9.0g |
| PCR\_std2\_bl | Baseline | %9.0g |
| PCR\_std3\_bl | Baseline | %9.0g |
| PCR\_std4\_bl | Baseline | %9.0g |
| PCR\_std5\_bl | Baseline | %9.0g |
| PCR\_std6\_bl | Baseline | %9.0g |
| PCR\_std7\_bl | Baseline | %9.0g |
| PCR\_std8\_bl | Baseline | %9.0g |
| PCR\_tot\_bl | Baseline Overall Pupil-Classroom Ratio | %9.0g |
| PCR\_lower\_bl | Baseline Lower Primary Pupil-Classroom Ratio | %9.0g |
| PCR\_upper\_bl | Baseline | %9.0g |
| hc\_PTR\_std1\_bl | Baseline PTR std 1 (HC) | %9.0g |
| hc\_PTR\_std2\_bl | Baseline PTR std 2 (HC) | %9.0g |
| hc\_PTR\_std3\_bl | Baseline PTR std 3 (HC) | %9.0g |
| hc\_PTR\_std4\_bl | Baseline PTR std 4 (HC) | %9.0g |
| hc\_PTR\_std5\_bl | Baseline PTR std 5 (HC) | %9.0g |
| hc\_PTR\_std6\_bl | Baseline PTR std 6 (HC) | %9.0g |
| hc\_PTR\_std7\_bl | Baseline PTR std 7 (HC) | %9.0g |
| hc\_PTR\_tot\_bl | Baseline PTR total (HC) | %9.0g |
| rr\_lower\_pri\_bl | Baseline Lower Primary Repetition Rate | %9.0g |
| dr\_lower\_pri\_bl | Baseline Lower Primary Dropout Rate | %9.0g |
| PTR\_top\_decile\_bl | Baseline school’s PTR (enrollment approach) is at top decile | %9.0g |
| progression\_rate\_lower\_pri\_bl | Baseline | %9.0g |

Table 3. Variable Names, Labels, and Format of the Derived Data at the Student Level.

| var\_name | label | format |
| --- | --- | --- |
| school\_id | School ID | %8.0g |
| student\_id | Student ID: | %8.0g |
| English\_IRT\_score\_bl | Baseline empirical Bayes means for Theta | %9.0g |
| Math\_IRT\_score\_bl | Baseline empirical Bayes means for Theta | %9.0g |
| Chi\_IRT\_score\_bl | Baseline empirical Bayes means for Theta | %9.0g |
| English\_klg\_score\_bl | Baseline English Knowledge score | %9.0g |
| Math\_klg\_score\_bl | Baseline Math Knowledge score | %9.0g |
| Chi\_klg\_score\_bl | Baseline Chichewa Knowledge score | %9.0g |
| MOEcode\_bl | Baseline School MoE code | %10.0g |
| school\_na\_bl | Baseline School Name | %30s |
| female\_bl | Baseline Female child (=1, 0 otherwise) | %9.0g |
| stu\_age\_bl | Baseline stu\_age | %9.0g |
| stu\_age\_above\_m\_bl | Baseline age is above the median age | %9.0g |
| Chewa\_speaking\_bl | Baseline Which language (s) do you usually speak at home?:Chewa | %9.0g |
| Tumbuka\_speaking\_bl | Baseline | %9.0g |
| less\_than\_15min\_dist\_bl | Baseline Distance to school <15 minutes | %9.0g |
| asset\_index\_bl | Baseline Scores for component 1 | %9.0g |
| above\_SES\_40th\_bl | Baseline asset\_index>SES\_40th | %9.0g |
| above\_SES\_10th\_bl | Baseline asset\_index>SES\_10th | %9.0g |
| above\_SES\_20th\_bl | Baseline asset\_index>SES\_20th | %9.0g |
| above\_SES\_40th\_sch\_bl | Baseline asset\_index>SES\_40th\_sch | %9.0g |
| above\_SES\_60th\_bl | Baseline asset\_index>SES\_60th | %9.0g |
| above\_SES\_75th\_bl | Baseline asset\_index>SES\_75th | %9.0g |
| live\_with\_parents\_bl | Baseline live with both mom and dad | %9.0g |
| live\_with\_parent\_bl | Baseline live with mom, dad, or both | %9.0g |
| both\_parent\_lit\_bl | Baseline both mom and dad literate (y or n) | %9.0g |
| parent\_lit\_bl | Baseline either mom, dad or both literate (y or n) | %9.0g |
| total\_raven\_bl | Baseline overall percentage raven score | %9.0g |
| total\_raven2\_bl | Baseline total raven score (out of 6) | %9.0g |
| absent\_last\_week\_yes\_bl | Baseline absent from school on any day last week (y or n) | %9.0g |
| hw\_help\_yes\_bl | Baseline student got homework last week (y or n) | %9.0g |
| stu\_repeat\_before\_bl | Baseline whether students repeated grades before (y or n) | %9.0g |
| exp\_score\_english\_bl | Baseline Expected score in English | %9.0g |
| exp\_score\_math\_bl | Baseline Expected score in Math | %9.0g |
| exp\_score\_chichewa\_bl | Baseline Expected score in Chichewa | %9.0g |
| stu\_self\_perception\_bl | Baseline student self-perception | %9.0g |
| a71a\_bl | Baseline 71a. You can score higher marks on tests/exams if you work hard) | %8.0g |
| a71b\_bl | Baseline 71b. You are interested in talking to new kids in school) | %8.0g |
| a71c\_bl | Baseline 71c. When you start your homework, you tend to finish it) | %8.0g |
| a71d\_bl | Baseline 71d. If your class fellows scores higher by putting in more effort, it | %8.0g |
| a71e\_bl | Baseline 71e. Failing in a test frustrates you) | %8.0g |
| a71f\_bl | Baseline 71f. You would prefer getting 2 sweets now instead of 5 tomorrow) | %8.0g |
| a71g\_bl | Baseline 71g. Teachers often tell you that you are not performing as well as you | %8.0g |
| a71h\_bl | Baseline 71h. Your parents often tell you that you are not performing as well as | %8.0g |
| a71i\_bl | Baseline 71i. School is boring for you) | %8.0g |
| a71j\_bl | Baseline 71j. Sometimes your friends lie to you and you lie to them) | %8.0g |
| think\_above\_ave\_bl | Baseline student self-evaluation of performance | %9.0g |
| general\_positive\_index\_bl | Baseline general positive mindset index (0-3) | %9.0g |
| growth\_mindset\_index\_bl | Baseline growth mindset index (0-5) | %9.0g |
| negative\_fb\_index\_bl | Baseline negative feedback: number of people given negative feedback (0-2) | %9.0g |
| stu\_tch\_perception\_bl | Baseline Student teacher perception | %9.0g |
| q65a\_bl | Baseline 65. Does the teacher encourage you to ask questions in the class? - En | %8.0g |
| q65b\_bl | Baseline 65. Does the teacher encourage you to ask questions in the class? - Ma | %8.0g |
| q65c\_bl | Baseline 65. Does the teacher encourage you to ask questions in the class? - Ch | %8.0g |
| q66a\_bl | Baseline 66. Does the teacher encourage you to answer questions in the class? - | %8.0g |
| q66b\_bl | Baseline 66. Does the teacher encourage you to answer questions in the class? - | %8.0g |
| q66c\_bl | Baseline 66. Does the teacher encourage you to answer questions in the class? - | %8.0g |
| ask\_the\_same\_anoth\_stu\_a\_bl | Baseline English teacher asks another student/the group the same question if stu | %9.0g |
| ask\_the\_same\_anoth\_stu\_b\_bl | Baseline Math teacher asks another student/the group the same question if studen | %9.0g |
| ask\_the\_same\_anoth\_stu\_c\_bl | Baseline Chichewa teacher asks another student/the group the same question if st | %9.0g |
| angry\_punish\_a\_bl | Baseline English teacher gets angry/upset and punishes me | %9.0g |
| angry\_punish\_b\_bl | Baseline Math teacher gets angry/upset and punishes me | %9.0g |
| angry\_punish\_c\_bl | Baseline Chichewa teacher gets angry/upset and punishes me | %9.0g |
| explain\_wrong\_anw\_a\_bl | Baseline English teacher explains what is wrong with the answer | %9.0g |
| explain\_wrong\_anw\_b\_bl | Baseline Math teacher explains what is wrong with the answer | %9.0g |
| explain\_wrong\_anw\_c\_bl | Baseline Chichewa teacher explains what is wrong with the answer | %9.0g |
| q68a\_bl | Baseline 68. Does the teacher mark your homework? - English Teacher | %8.0g |
| q68b\_bl | Baseline 68. Does the teacher mark your homework? - Maths Teacher | %8.0g |
| q68c\_bl | Baseline 68. Does the teacher mark your homework? - Chichewa Teacher | %8.0g |
| q69a\_bl | Baseline 69. Does the teacher explain the marks given on your homework? - Engli | %8.0g |
| q69b\_bl | Baseline 69. Does the teacher explain the marks given on your homework? - Maths | %8.0g |
| q69c\_bl | Baseline 69. Does the teacher explain the marks given on your homework? - Chich | %8.0g |
| q70a\_bl | Baseline 70. Is the teacher available after class if you need help with the less | %8.0g |
| q70b\_bl | Baseline 70. Is the teacher available after class if you need help with the less | %8.0g |
| q70c\_bl | Baseline 70. Is the teacher available after class if you need help with the less | %8.0g |
| asset\_index\_sum\_bl | Baseline the simple sum of all pca items on asset index | %9.0g |
| stu\_behav\_index\_bl | Baseline student part behavior indicators, general\_positive\_index + growth\_minds | %9.0g |
| greater\_than\_exp\_eng\_bl | Baseline whether greater than expected score english (y or n) | %9.0g |
| greater\_than\_exp\_math\_bl | Baseline whether greater than expected score Math (y or n) | %9.0g |
| greater\_than\_exp\_chi\_bl | Baseline whether greater than expected score Chichewa (y or n) | %9.0g |
| pca\_stu\_behav\_bl | Baseline Scores for component 1 | %9.0g |
| toilet\_bl | Baseline Is there a toilet at/ near your house that you use? Y or N | %9.0g |
| electricity\_bl | Baseline Do you have electricity in your house? Y or N | %9.0g |
| piped\_water\_bl | Baseline Do you have running water supply (pipe water) at your house? Y or N | %9.0g |
| tv\_bl | Baseline Does your family have … TV? Y or N | %9.0g |
| bed\_bl | Baseline Does your family have … Bed(s)? Y or N | %9.0g |
| radio\_bl | Baseline Does your family have … Radio? Y or N | %9.0g |
| phone\_bl | Baseline Does your family have … Mobile phone? Y or N | %9.0g |
| stove\_bl | Baseline Does your family have … Stove / Mafuwa / Mbaula? Y or N | %9.0g |
| bike\_bl | Baseline Does your family have … Bicycle? Y or N | %9.0g |
| motor\_bl | Baseline Does your family have … Motor cycle/ scooter? Y or N | %9.0g |
| truck\_bl | Baseline Does your family have … Tractor/ Truck/ Car? Y or N | %9.0g |
| chicken\_bl | Baseline Does your family have … Chickens? Y or N | %9.0g |
| goats\_bl | Baseline Does your family have … Goats/ Sheep? Y or N | %9.0g |
| pig\_bl | Baseline Does your family have … Pigs? Y or N | %9.0g |
| cattle\_bl | Baseline Does your family have … Cattle? Y or N | %9.0g |
| live\_with\_single\_parent\_bl | Baseline live with only mom or only dad (single parent) | %9.0g |
| hw\_time\_bl | Baseline How much time do you usually spend at home every day doing homework? | %9.0g |
| district\_na\_bl | Baseline District Name | %9s |
| school\_loc\_bl | Baseline Rural/urban location | %10.0g |
| visit\_filter\_bl | Baseline Is this a first (unannounced) visit or a second (announced) visit? | %23.0g |
| enum\_first\_visit1\_bl | Baseline Enumerator on first Visit | %24.0g |
| time\_first\_visit\_bl | Baseline Time Started | %9.0g |
| enum\_second\_visit1\_bl | Baseline Enumerator on second Visit | %10.0g |
| time\_second\_visit\_bl | Baseline Time Started | %9.0g |
| consent\_bl | Baseline Consent | %8.0g |
| a1\_bl | Baseline 1.First, middle and last name | %9s |
| a2\_bl | Baseline 2.Student Learner Identification Number (LIN) | %20s |
| a3\_bl | Baseline 3. Sex [Observed] | %8.0g |
| a4\_bl | Baseline 4. How old are you? | %8.0g |
| a5\_bl | Baseline 5. What is your father<U+FFFD>s full name? | %9s |
| a6\_bl | Baseline 6. Did you find any physical disability in this child? | %8.0g |
| a7a\_bl | Baseline 7. Which language (s) do you usually speak at home? | %14.0g |
| a7b\_bl | Baseline 7. Which language (s) do you usually speak at home? | %14.0g |
| a7c\_bl | Baseline 7. Which language (s) do you usually speak at home? | %14.0g |
| a7d\_bl | Baseline 7. Which language (s) do you usually speak at home? | %14.0g |
| a8\_bl | Baseline 8. Where do you live? | %30s |
| a9\_bl | Baseline 9. Do you live with both your mom and dad? | %8.0g |
| a10\_bl | Baseline 10. Who do you live with? | %21.0g |
| a10\_spec\_bl | Baseline 10. Who do you live with? - Other specify | %29s |
| a11\_bl | Baseline 11. Have you seen your mom reading the Bible, a book or a newspaper at | %8.0g |
| a12\_bl | Baseline 12. What does your mom do while you are in school? If child does no | %33.0g |
| a12\_spec\_bl | Baseline 12. What does your mom do while you are in school? If child does no | %74s |
| a13\_bl | Baseline 13. Have you seen your dad reading the Bible, a book or a newspaper at | %8.0g |
| a14\_bl | Baseline 14. What does your dad do while you are in school? | %33.0g |
| a14\_spec\_bl | Baseline 14. What does your dad do while you are in school? [Other (specify)] | %88s |
| a15\_bl | Baseline 15. How many brothers do you have? | %8.0g |
| a15\_m\_bl | Baseline 15\_m. How many of your brothers are older than you? | %8.0g |
| a16\_bl | Baseline 16. How many sisters do you have? | %8.0g |
| a16\_m\_bl | Baseline 16\_m. How many of your sisters are older than you? | %8.0g |
| a17m\_a\_bl | Baseline 17a. I get nervous when I don<U+FFFD>t know how to solve a task at school) 17\_ | %8.0g |
| a17m\_b\_bl | Baseline 17b. I perform poorly at school whether or not I study for my exams) 17 | %8.0g |
| a17m\_c\_bl | Baseline 17c.<U+FFFD>If I put in enough effort, I can succeed in school) 17\_m. Do you a | %8.0g |
| a17m\_d\_bl | Baseline 17d. I learn most when I work with other students) 17\_m. Do you agree w | %8.0g |
| a17m\_e\_bl | Baseline 17e.<U+FFFD>Many things I learn at school will help me in my life later on.) 1 | %8.0g |
| a18a\_bl | Baseline 18. On a typical day, how do you come to school from your home? Do you<U+FFFD> | %21.0g |
| a18b\_bl | Baseline 18. On a typical day, how do you come to school from your home? Do you<U+FFFD> | %21.0g |
| a21\_bl | Baseline 21. On a typical day, how long does it take for you to get to school fr | %18.0g |
| a22\_bl | Baseline 22. Do you usually reach school on time, before the start of first peri | %8.0g |
| a23\_bl | Baseline 23. How long would you like to stay in school? | %36.0g |
| q24\_bl | Baseline 24. What type of house do you live in? | %80.0g |
| q25\_bl | Baseline 25. What is the type of roof of your house? | %10.0g |
| q25\_spec\_bl | Baseline 25. What is the type of roof of your house? [Other (specify)] | %30s |
| q26\_bl | Baseline 26. What is the type of walls of your house? | %39.0g |
| q26\_spec\_bl | Baseline 26. What is the type of walls of your house? [Other (specify)] | %61s |
| q27\_bl | Baseline 27. What is the type of floor in your house? | %31.0g |
| q27\_spec\_bl | Baseline 27. What is the type of floor in your house? [Other (specify)] | %30s |
| q28\_bl | Baseline 28. Is there a toilet at/ near your house that you use? | %8.0g |
| q29\_bl | Baseline 29. Do you have electricity in your house? | %8.0g |
| q30\_bl | Baseline 30. Do you have running water supply (pipe water) at your house? | %8.0g |
| q31\_bl | Baseline 31. Does your family have farmland (cultivable land)? | %8.0g |
| a32a\_bl | Baseline 32. Does your family have -? (a. TV ) | %8.0g |
| a32b\_bl | Baseline 32. Does your family have -? (b. Bed(s) ) | %8.0g |
| a32c\_bl | Baseline 32. Does your family have -? (c.<U+FFFD>Radio ) | %8.0g |
| a32d\_bl | Baseline 32. Does your family have -? (d. Mobile phone ) | %8.0g |
| a32f\_bl | Baseline 32. Does your family have -? (f.<U+FFFD>Stove / Mafuwa / Mbaula) | %8.0g |
| a32g\_bl | Baseline 32. Does your family have -? (g.Bicycle ) | %8.0g |
| a32h\_bl | Baseline 32. Does your family have -? (h.Motor cycle/ scooter ) | %8.0g |
| a32i\_bl | Baseline 32. Does your family have -? (i.Tractor/ Truck/ Car) | %8.0g |
| a32k\_bl | Baseline 32. Does your family have -? (k.Chickens ) | %8.0g |
| a32l\_bl | Baseline 32. Does your family have -? (l.Goats/ Sheep) | %8.0g |
| a32n\_bl | Baseline 32. Does your family have -? (n.Pigs ) | %8.0g |
| a32o\_bl | Baseline 32. Does your family have -? (o.Cattle ) | %8.0g |
| a32o\_m\_bl | Baseline 32. Does your family have -? (o\_m.Other large items (Specify)) | %8.0g |
| a32o\_m\_spec\_bl | Baseline 32o\_m. Specify Other large items | %46s |
| a32m\_a\_bl | Baseline 32ma.<U+FFFD>It is completely my choice whether or not I do well at school) | %8.0g |
| a32m\_b\_bl | Baseline 32mb.<U+FFFD>I do certain things that are bad for me, if they are fun) | %8.0g |
| a32m\_c\_bl | Baseline 32mc.<U+FFFD>I often set a goal but later choose to pursue a different one.) | %8.0g |
| a32m\_d\_bl | Baseline 32md. I often worry that it will be difficult for me taking a test.) | %8.0g |
| a32m\_e\_bl | Baseline 32me. I prefer doing exercises in a group to doing them alone.) | %8.0g |
| sec1\_comments\_bl | Baseline Interviewer comments for this section | %215s |
| a33\_bl | Baseline 33. In which standard did you join this school? | %10.0g |
| a34\_bl | Baseline 34. Which standard are you currently in? | %10.0g |
| a35a\_bl | Baseline 35. How many years were you in<U+FFFD>Standard 4 | %8.0g |
| a36a\_bl | Baseline 36.Why did you repeat this grade? | %32.0g |
| a36a\_spec\_bl | Baseline 36.Why did you repeat this grade?…Standard 4.Other Specified | %9s |
| a35b\_bl | Baseline 35. How many years were you in<U+FFFD>Standard 3 | %8.0g |
| a36b\_bl | Baseline 36.Why did you repeat this grade? | %32.0g |
| a36b\_spec\_bl | Baseline 36.Why did you repeat this grade?…Standard 3.Other Specified | %120s |
| a35c\_bl | Baseline 35. How many years were you in<U+FFFD>Standard 2 | %8.0g |
| a36c\_bl | Baseline 36.Why did you repeat this grade? | %32.0g |
| a36c\_spec\_bl | Baseline 36.Why did you repeat this grade?…Standard 2.Other Specified | %9s |
| a35d\_bl | Baseline 35. How many years were you in<U+FFFD>Standard 1 | %8.0g |
| a36d\_bl | Baseline 36.Why did you repeat this grade? | %32.0g |
| a36d\_spec\_bl | Baseline 36.Why did you repeat this grade?…Standard 1.Other Specified | %9s |
| a35e\_bl | Baseline 35. How many years were you in<U+FFFD>Pre-school | %8.0g |
| a36e\_bl | Baseline 36.Why did you repeat this grade? | %32.0g |
| a36e\_spec\_bl | Baseline 36.Why did you repeat this grade?…Pre-school.Other Specified | %9s |
| a47\_m1\_bl | Baseline 47\_m1. Do you have a school bag? | %8.0g |
| a47\_m2\_bl | Baseline 47\_m2. Do you have a uniform? | %8.0g |
| a47\_m3\_bl | Baseline 47\_m3. Do you have a notebook/ exercise book/ workbook? | %8.0g |
| a47\_m4\_bl | Baseline 47\_m4. Did you get the notebook/ exercise book/ workbook from the schoo | %8.0g |
| q48\_bl | Baseline 48. Do you do your homework? | %23.0g |
| q49\_bl | Baseline 49. How much time do you usually spend at home every day doing homework | %8.0g |
| q50\_bl | Baseline 50. Who helps you the most with homework at home? | %14.0g |
| q50\_spec\_bl | Baseline 50. Who helps you the most with homework at home? [Other (specify)] | %36s |
| a51\_bl | Baseline 51. Last week, were you absent from school on any day? | %8.0g |
| a52\_bl | Baseline 52. Why were you absent? | %31.0g |
| a52\_spec\_bl | Baseline 52. Why were you absent? [Other (specify)] | %9s |
| q53\_bl | Baseline 53. Last week, were you absent from school on all days? | %8.0g |
| q54a\_bl | Baseline 54. During the last school week did your - teacher give you a written t | %8.0g |
| q54b\_bl | Baseline 54. During the last school week did your - teacher give you a written t | %8.0g |
| q54c\_bl | Baseline 54. During the last school week did your - teacher give you a written t | %8.0g |
| q55a\_bl | Baseline 55. During the last school week did your - teacher give you homework? ( | %8.0g |
| q55b\_bl | Baseline 55. During the last school week did your - teacher give you homework? ( | %8.0g |
| q55c\_bl | Baseline 55. During the last school week did your - teacher give you homework? ( | %8.0g |
| q56a\_bl | Baseline 56. Did your- teacher end class early on some days last week? (a. Engl | %8.0g |
| q56b\_bl | Baseline 56. Did your- teacher end class early on some days last week? (b. Math | %8.0g |
| q56c\_bl | Baseline 56. Did your- teacher end class early on some days last week? (c. Chic | %8.0g |
| q57a\_bl | Baseline 57. How many days last week was your - teacher absent from school? (a. | %8.0g |
| q57b\_bl | Baseline 57. How many days last week was your - teacher absent from school? (b. | %8.0g |
| q57c\_bl | Baseline 57. How many days last week was your - teacher absent from school? (c. | %8.0g |
| q58\_bl | Baseline 58. In your opinion how are you performing in school tests compared to | %15.0g |
| q59a\_bl | Baseline 59. If you are given 100-\_\_\_ questions from the topics in the textbook | %8.0g |
| q59b\_bl | Baseline 59. If you are given 100-\_\_\_ questions from the topics in the textbook | %8.0g |
| q59c\_bl | Baseline 59. If you are given 100-\_\_\_ questions from the topics in the textbook | %8.0g |
| q59\_m\_a\_bl | Baseline 59\_m\_a. I finish whatever I begin.) | %8.0g |
| q59\_m\_b\_bl | Baseline 59\_m\_b. I do the right thing, even if others do not like it.) | %8.0g |
| q59\_m\_c\_bl | Baseline 59\_m\_c.<U+FFFD>I often stay mad at people even when they apologize.) | %8.0g |
| q59\_m\_d\_bl | Baseline 59\_m\_d. If I get good grades at school it will not make a difference in | %8.0g |
| qsec2\_comments\_bl | Baseline Interviewer comments for this section | %271s |
| q60a\_bl | Baseline 60a.What is the name of your English teacher | %30s |
| q60b\_bl | Baseline 60b.What is the name of your Maths teacher? | %30s |
| q60c\_bl | Baseline 60c.What is the name of your Chichewa teacher? | %31s |
| q61a\_bl | Baseline 61a.Teacher ID - English Teacher | %12.0g |
| q61b\_bl | Baseline 61b.Teacher ID - Maths Teacher | %8.0g |
| q61c\_bl | Baseline 61c.Teacher ID - Chichewa Teacher | %8.0g |
| q67a\_bl | Baseline 67. What does the teacher do if you give wrong answer to the question? | %58.0g |
| q67b\_bl | Baseline 67. What does the teacher do if you give wrong answer to the question? | %58.0g |
| q67c\_bl | Baseline 67. What does the teacher do if you give wrong answer to the question? | %58.0g |
| q67\_m1\_aa\_bl | Baseline a.The teacher has to wait a long time for students to quiet down | %20.0g |
| q67\_m1\_ba\_bl | Baseline b.The teacher gives extra help when students need it | %20.0g |
| q67\_m1\_ca\_bl | Baseline c.The teacher explains things in different ways until I understand | %20.0g |
| q67\_m1\_da\_bl | Baseline d.The teacher tells me how I can improve my performance | %20.0g |
| q67\_m1\_ab\_bl | Baseline a.The teacher has to wait a long time for students to quiet down | %20.0g |
| q67\_m1\_bb\_bl | Baseline b.The teacher gives extra help when students need it | %20.0g |
| q67\_m1\_cb\_bl | Baseline c.The teacher explains things in different ways until I understand | %20.0g |
| q67\_m1\_db\_bl | Baseline d.The teacher tells me how I can improve my performance | %20.0g |
| q67\_m1\_ac\_bl | Baseline a.The teacher has to wait a long time for students to quiet down | %20.0g |
| q67\_m1\_bc\_bl | Baseline b.The teacher gives extra help when students need it | %20.0g |
| q67\_m1\_cc\_bl | Baseline c.The teacher explains things in different ways until I understand | %20.0g |
| q67\_m1\_dc\_bl | Baseline d.The teacher tells me how I can improve my performance | %20.0g |
| q70\_m1\_bl | Baseline 70m1.Do you agree with this statement: ’If I had different teachers, I | %17.0g |
| q70\_m2\_a\_bl | Baseline a) Teachers said something insulting to me in front of others | %30.0g |
| q70\_m2\_b\_bl | Baseline b) Teachers called on me less often than they called on other students | %30.0g |
| q70\_m2\_c\_bl | Baseline c) I got called names by other students | %30.0g |
| q70\_m2\_d\_bl | Baseline d) Other students took away or destroyed things that belonged to me | %30.0g |
| a72\_bl | Baseline 72. Which figure follows? Fc 1 | %8.0g |
| a73\_bl | Baseline 73. Which figure follows? Fc 2 | %8.0g |
| a74\_bl | Baseline 74. Which figure follows? Fc 3 | %8.0g |
| a75\_bl | Baseline 75. Which figure follows? Fc 4 | %8.0g |
| a76\_bl | Baseline 76. Which figure follows? Fc 5 | %8.0g |
| a77\_bl | Baseline 77. Which figure follows? Fc 6 | %8.0g |
| sec5\_comments\_bl | Baseline Interviewer comments for this section | %103s |
| activity\_6am\_7am\_bl | Baseline Activity (6am-7am) | %26.0g |
| activity\_6am\_7am\_spec\_bl | Baseline Specify activity | %88s |
| housework\_6am\_7am\_bl | Baseline Type of House Work (6am-7am) | %30.0g |
| housework\_6am\_7am\_spec\_bl | Baseline Specify Work | %39s |
| paidwork\_6am\_7am\_bl | Baseline Type of Paid Work (6am-7am) | %25.0g |
| paidwork\_6am\_7am\_spec\_bl | Baseline Specify Work | %9s |
| activity\_7am\_8am\_bl | Baseline Activity (7am\_8am) | %26.0g |
| activity\_7am\_8am\_spec\_bl | Baseline Specify activity | %62s |
| housework\_7am\_8am\_bl | Baseline Type of House Work (7am\_8am) | %30.0g |
| housework\_7am\_8am\_spec\_bl | Baseline Specify Work | %32s |
| paidwork\_7am\_8am\_bl | Baseline Type of Paid Work (7am\_8am) | %25.0g |
| paidwork\_7am\_8am\_spec\_bl | Baseline Specify Work | %9s |
| activity\_8am\_9am\_bl | Baseline Activity (8am\_9am) | %26.0g |
| activity\_8am\_9am\_spec\_bl | Baseline Specify activity | %127s |
| housework\_8am\_9am\_bl | Baseline Type of House Work (8am\_9am) | %30.0g |
| housework\_8am\_9am\_spec\_bl | Baseline Specify Work | %9s |
| paidwork\_8am\_9am\_bl | Baseline Type of Paid Work (8am\_9am) | %25.0g |
| paidwork\_8am\_9am\_spec\_bl | Baseline Specify Work | %9s |
| activity\_9am\_10am\_bl | Baseline Activity (9am\_10am) | %26.0g |
| activity\_9am\_10am\_spec\_bl | Baseline Specify activity | %47s |
| housework\_9am\_10am\_bl | Baseline Type of House Work (9am\_10am) | %30.0g |
| housework\_9am\_10am\_spec\_bl | Baseline Specify Work | %28s |
| paidwork\_9am\_10am\_bl | Baseline Type of Paid Work (9am\_10am) | %25.0g |
| paidwork\_9am\_10am\_spec\_bl | Baseline Specify Work | %9s |
| activity\_10am\_11am\_bl | Baseline Activity (10am\_11am) | %26.0g |
| activity\_10am\_11am\_spec\_bl | Baseline Specify activity | %36s |
| housework\_10am\_11am\_bl | Baseline Type of House Work (10am\_11am) | %30.0g |
| housework\_10am\_11am\_spec\_bl | Baseline Specify Work | %9s |
| paidwork\_10am\_11am\_bl | Baseline Type of Paid Work (10am\_11am) | %25.0g |
| paidwork\_10am\_11am\_spec\_bl | Baseline Specify Work | %9s |
| activity\_11am\_12pm\_bl | Baseline Activity (11am\_12am) | %26.0g |
| activity\_11am\_12pm\_spec\_bl | Baseline Specify activity | %34s |
| housework\_11am\_12pm\_bl | Baseline Type of House Work (11am\_12am) | %30.0g |
| housework\_11am\_12pm\_spec\_bl | Baseline Specify Work | %16s |
| paidwork\_11am\_12pm\_bl | Baseline Type of Paid Work (11am\_12am) | %25.0g |
| paidwork\_11am\_12pm\_spec\_bl | Baseline Specify Work | %9s |
| activity\_12pm\_1pm\_bl | Baseline Activity (12Pm\_1Pm) | %26.0g |
| activity\_12pm\_1pm\_spec\_bl | Baseline Specify activity | %37s |
| housework\_12pm\_1pm\_bl | Baseline Type of House Work (12Pm\_1Pm) | %30.0g |
| housework\_12pm\_1pm\_spec\_bl | Baseline Specify Work | %17s |
| paidwork\_12pm\_1pm\_bl | Baseline Type of Paid Work (12Pm\_1Pm) | %25.0g |
| paidwork\_12pm\_1pm\_spec\_bl | Baseline Specify Work | %9s |
| activity\_1pm\_2pm\_bl | Baseline Activity (1Pm\_2Pm) | %26.0g |
| activity\_1pm\_2pm\_spec\_bl | Baseline Specify activity | %55s |
| housework\_1pm\_2pm\_bl | Baseline Type of House Work (1Pm\_2Pm) | %30.0g |
| housework\_1pm\_2pm\_spec\_bl | Baseline Specify Work | %29s |
| paidwork\_1pm\_2pm\_bl | Baseline Type of Paid Work (1Pm\_2Pm) | %25.0g |
| paidwork\_1pm\_2pm\_spec\_bl | Baseline Specify Work | %9s |
| activity\_2pm\_3pm\_bl | Baseline Activity (2Pm\_3Pm) | %26.0g |
| activity\_2pm\_3pm\_spec\_bl | Baseline Specify activity | %101s |
| housework\_2pm\_3pm\_bl | Baseline Type of House Work (2Pm\_3Pm) | %30.0g |
| housework\_2pm\_3pm\_spec\_bl | Baseline Specify Work | %36s |
| paidwork\_2pm\_3pm\_bl | Baseline Type of Paid Work (2Pm\_3Pm) | %25.0g |
| paidwork\_2pm\_3pm\_spec\_bl | Baseline Specify Work | %32s |
| activity\_3pm\_4pm\_bl | Baseline Activity (3Pm\_4Pm) | %26.0g |
| activity\_3pm\_4pm\_spec\_bl | Baseline Specify activity | %42s |
| housework\_3pm\_4pm\_bl | Baseline Type of House Work (3Pm\_4Pm) | %30.0g |
| housework\_3pm\_4pm\_spec\_bl | Baseline Specify Work | %42s |
| paidwork\_3pm\_4pm\_bl | Baseline Type of Paid Work (3Pm\_4Pm) | %25.0g |
| paidwork\_3pm\_4pm\_spec\_bl | Baseline Specify Work | %33s |
| activity\_4pm\_5pm\_bl | Baseline Activity (4Pm\_5Pm) | %26.0g |
| activity\_4pm\_5pm\_spec\_bl | Baseline Specify activity | %45s |
| housework\_4pm\_5pm\_bl | Baseline Type of House Work (4Pm\_5Pm) | %30.0g |
| housework\_4pm\_5pm\_spec\_bl | Baseline Specify Work | %33s |
| paidwork\_4pm\_5pm\_bl | Baseline Type of Paid Work (4Pm\_5Pm) | %25.0g |
| paidwork\_4pm\_5pm\_spec\_bl | Baseline Specify Work | %107s |
| activity\_5pm\_6pm\_bl | Baseline Activity (5Pm\_6Pm) | %26.0g |
| activity\_5pm\_6pm\_spec\_bl | Baseline Specify activity | %71s |
| housework\_5pm\_6pm\_bl | Baseline Type of House Work (5Pm\_6Pm) | %30.0g |
| housework\_5pm\_6pm\_spec\_bl | Baseline Specify Work | %101s |
| paidwork\_5pm\_6pm\_bl | Baseline Type of Paid Work (5Pm\_6Pm) | %25.0g |
| paidwork\_5pm\_6pm\_spec\_bl | Baseline Specify Work | %9s |
| activity\_6pm\_7pm\_bl | Baseline Activity (6Pm\_7Pm) | %26.0g |
| activity\_6pm\_7pm\_spec\_bl | Baseline Specify activity | %39s |
| housework\_6pm\_7pm\_bl | Baseline Type of House Work (6Pm\_7Pm) | %30.0g |
| housework\_6pm\_7pm\_spec\_bl | Baseline Specify Work | %66s |
| paidwork\_6pm\_7pm\_bl | Baseline Type of Paid Work (6Pm\_7Pm) | %25.0g |
| paidwork\_6pm\_7pm\_spec\_bl | Baseline Specify Work | %9s |
| activity\_7pm\_8pm\_bl | Baseline Activity (7Pm\_8m) | %26.0g |
| activity\_7pm\_8m\_spec\_bl | Baseline Specify activity | %36s |
| housework\_7pm\_8pm\_bl | Baseline Type of House Work (7Pm\_8m) | %30.0g |
| housework\_7pm\_8pm\_spec\_bl | Baseline Specify Work | %22s |
| paidwork\_7pm\_8pm\_bl | Baseline Type of Paid Work (7Pm\_8m) | %25.0g |
| paidwork\_7pm\_8pm\_spec\_bl | Baseline Specify Work | %9s |
| activity\_8pm\_9pm\_bl | Baseline Activity (8Pm\_9Pm) | %26.0g |
| activity\_8pm\_9pm\_spec\_bl | Baseline Specify activity | %28s |
| housework\_8pm\_9pm\_bl | Baseline Type of House Work (8Pm\_9Pm) | %30.0g |
| housework\_8pm\_9pm\_spec\_bl | Baseline Specify Work | %9s |
| paidwork\_8pm\_9pm\_bl | Baseline Type of Paid Work (8Pm\_9Pm) | %25.0g |
| paidwork\_8pm\_9pm\_spec\_bl | Baseline Specify Work | %9s |
| activity\_9pm\_10pm\_bl | Baseline Activity (9Pm\_10Pm) | %26.0g |
| activity\_9pm\_10pm\_spec\_bl | Baseline Specify activity | %35s |
| housework\_9pm\_10pm\_bl | Baseline Type of House Work (9Pm\_10Pm) | %30.0g |
| housework\_9pm\_10pm\_spec\_bl | Baseline Specify Work | %9s |
| paidwork\_9pm\_10pm\_bl | Baseline Type of Paid Work (9Pm\_10Pm) | %25.0g |
| paidwork\_9pm\_10pm\_spec\_bl | Baseline Specify Work | %9s |
| sec6\_comments\_bl | Baseline INTERVIEWER COMMENTS | %276s |
| visit4\_bl | Baseline Is this still your first Visit ? | %20.0g |
| enum\_second\_visit2\_bl | Baseline Enumerator on second Visit | %9s |
| time\_second\_visit2\_bl | Baseline | %8.0g |
| refusal\_reason\_bl | Baseline Reasons of refusal | %9s |
| supervisor\_name\_bl | Baseline Select the name of the supervisor who accompanied you. | %24s |
| time\_review\_bl | Baseline Time | %9.0g |
| review\_result\_bl | Baseline Review Result | %24.0g |
| incomplete\_reason\_bl | Baseline Reason to why the questionnaire is incomplete | %9s |
| formdef\_version\_bl | Baseline Form version used on device | %12.0g |
| key\_bl | Baseline Unique submission ID | %41s |
| column1\_bl | Baseline Column1 | %16s |
| submissiondate\_bl | Baseline Date/time submitted | %tc |
| start\_bl | Baseline | %tc |
| date\_first\_visit\_bl | Baseline Date of Interview | %td |
| date\_sec\_visit1\_bl | Baseline Date of Interview | %td |
| date\_sec\_visit2\_bl | Baseline Date of Interview | %td |
| date\_review\_bl | Baseline Date | %td |
| BL16\_bl | Baseline | %8.0g |
| qnum\_bl | Baseline questionnaire number | %8.0g |
| q4\_bl | Baseline 4. date [first unannounced visit] | %10s |
| q5\_bl | Baseline 5. enumerator’s name & code [second unannounced visit if applicable] | %8.0g |
| q6\_bl | Baseline 6. date [second unannounced visit if applicable] | %10s |
| q8\_bl | Baseline 8. at what time did you end the survey? (24 hts format) first visit | %6s |
| q9\_bl | Baseline 9. at what time did you start the survey? (24 hts format) second visit | %6s |
| q10\_bl | Baseline 10. at what time did you end the survey? (24 hts format) second visit | %6s |
| q11a1\_bl | Baseline a. supervisor | %8.0g |
| q11a2\_bl | Baseline name and signature | %8.0g |
| q11a3\_bl | Baseline date | %10s |
| q11a4\_bl | Baseline reason | %9s |
| q11b1\_bl | Baseline a. supervisor | %8.0g |
| q11b2\_bl | Baseline name and signature | %8.0g |
| q11b3\_bl | Baseline date | %10s |
| q11b4\_bl | Baseline reason | %1s |
| q11c1\_bl | Baseline c. data entry clerk 2 | %8.0g |
| q11c2\_bl | Baseline name and signature | %1s |
| q11c3\_bl | Baseline date | %10s |
| q11c4\_bl | Baseline reason | %1s |
| q11d1\_bl | Baseline d. m&e supervisor (manager) | %8.0g |
| q11d2\_bl | Baseline name and signature | %8.0g |
| q11d3\_bl | Baseline date | %10s |
| q11d4\_bl | Baseline reason | %1s |
| int\_start\_bl | Baseline may i now beging the interview | %8.0g |
| A2\_bl | Baseline new student code | %10s |
| a7\_spec\_bl | Baseline Q1.7. Which language (s) do you usually speak at home? (Other (Specify | %27s |
| a17\_bl | Baseline 17. are there any brothers/sisters older than you? | %8.0g |
| a18c\_bl | Baseline 18. how do you come to school from your home? do you? | %21.0g |
| a18\_spec\_bl | Baseline other specify | %1s |
| a19\_bl | Baseline 19. at what time do you leave for school in the morning? | %9.0g |
| a20\_bl | Baseline 20. at what time do you get to school? | %9.0g |
| a32e\_bl | Baseline 32. does your family have e) fridge | %8.0g |
| a32j\_bl | Baseline 32. does your family have j) car/van/taxi/bus | %8.0g |
| a32m\_bl | Baseline 32. does your family have m) sheep | %8.0g |
| a37a\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (a. English) | %8.0g |
| a37b\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (b. Math) | %8.0g |
| a37c\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (c. Chichewa) | %8.0g |
| a37d\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (d. Science) | %8.0g |
| a37e\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (e. Social St | %8.0g |
| a37f\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (f. Religious | %8.0g |
| a37g\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (g. Life Skil | %8.0g |
| a37h\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (h. Expressiv | %8.0g |
| a37i\_bl | Baseline 37. Which subjects do you study at school? Do you study<U+FFFD> (i. Agricultu | %8.0g |
| a37j\_bl | Baseline 37. which subjects do you study at school. do you study…. j. other | %8.0g |
| a37\_spec\_bl | Baseline specify | %20s |
| a47\_bl | Baseline 47. do you have learning materials (pens, pencils, notebooks) | %8.0g |
| q62a\_bl | Baseline 62. do you like your english teacher? | %8.0g |
| q62b\_bl | Baseline 62. do you like your math teacher? | %8.0g |
| q62c\_bl | Baseline 62. do you like your chichewa teacher? | %8.0g |
| q63a\_bl | Baseline 63. does english teacher provide you with examples if you do not unders | %8.0g |
| q63b\_bl | Baseline 63. does math teacher provide you with examples if you do not understan | %8.0g |
| q63c\_bl | Baseline 63. does chichewa teacher provide you with examples if you do not under | %8.0g |
| q64a\_bl | Baseline 64. does the english teacher help you if you have any difficulty in the | %8.0g |
| q64b\_bl | Baseline 64. does the math teacher help you if you have any difficulty in the le | %8.0g |
| q64c\_bl | Baseline 64. does the chichewa teacher help you if you have any difficulty in th | %8.0g |
| source\_bl | Baseline | %8.0g |
| A2\_STRING\_bl | Baseline school\_id\_a34 | %9s |
| base\_bl | Baseline | %9.0g |
| stu\_age\_square\_bl | Baseline square of student age | %9.0g |
| travel\_time\_bl | Baseline On a typical day, how long does it take for you to get to school from h | %9.0g |
| live\_with\_no\_parent\_bl | Baseline live with neither mom nor dad | %9.0g |
| mom\_literate\_bl | Baseline mom literate (y or n) | %9.0g |
| dad\_literate\_bl | Baseline dad literate (y or n) | %9.0g |
| perm\_room\_bl | Baseline Permanent structure (mostly the house is made of concrete, brick, mud e | %9.0g |
| SES\_10th\_bl | Baseline 10th percentile of asset index | %9.0g |
| SES\_20th\_bl | Baseline 20th percentile of asset index | %9.0g |
| SES\_40th\_bl | Baseline 40th percentile of asset index | %9.0g |
| SES\_40th\_sch\_bl | Baseline 40th percentile of asset index, by school | %9.0g |
| SES\_60th\_bl | Baseline 60th percentile of asset index | %9.0g |
| SES\_75th\_bl | Baseline 75th percentile of asset index | %9.0g |
| a\_72\_bl | Baseline Correct answer for Which figure follows in FLASHCARD 1? Y or N | %9.0g |
| a\_73\_bl | Baseline Correct answer for Which figure follows in FLASHCARD 2? Y or N | %9.0g |
| a\_74\_bl | Baseline Correct answer for Which figure follows in FLASHCARD 3? Y or N | %9.0g |
| a\_75\_bl | Baseline Correct answer for Which figure follows in FLASHCARD 4? Y or N | %9.0g |
| a\_76\_bl | Baseline Correct answer for Which figure follows in FLASHCARD 5? Y or N | %9.0g |
| a\_77\_bl | Baseline Correct answer for Which figure follows in FLASHCARD 6? Y or N | %9.0g |
| get\_hw\_yes\_bl | Baseline student got English, Math or Chichewa homework last week (y or n) | %9.0g |
| stu\_verb\_abuse\_bl | Baseline Student was recently called names | %9.0g |
| stu\_oth\_abuse\_bl | Baseline Student recently had things taken away or destroyed | %9.0g |
| qq1\_bl | Baseline QQ number | %8.0g |
| class\_id\_bl | Baseline Class ID: | %8.0g |
| enumerator\_name\_bl | Baseline Enumerator name & Code | %50s |
| start\_time\_hours\_bl | Baseline Start Time(hours) | %8.0g |
| start\_time\_min\_bl | Baseline Start Time(Min) | %8.0g |
| end\_time\_hours\_bl | Baseline End Time(hours) | %8.0g |
| end\_time\_min\_bl | Baseline End Time(Min) | %8.0g |
| studnet\_full\_name\_bl | Baseline Studnet Full Name | %9s |
| gender\_bl | Baseline Gender: | %1s |
| standard\_bl | Baseline Standard | %8.0g |
| stream\_bl | Baseline Stream: | %3s |
| e21a\_bl | Baseline 7a) Write sentences using the following words\_\_\_ ; WATER ; | %41s |
| e21b\_bl | Baseline 7b) Write sentences using the following words\_\_\_ ; EAT ; | %40s |
| studentid1\_bl | Baseline Student ID1 | %8.0g |
| entry1\_id\_bl | Baseline 1st Data entry (ID) | %18.0g |
| dd1\_bl | Baseline Day [1st Data Entry] | %8.0g |
| yr1\_bl | Baseline Year [1st Data Entry] | %8.0g |
| entry2\_id\_bl | Baseline 2nd Data entry (ID) | %18.0g |
| dd2\_bl | Baseline Day [2st Data Entry] | %8.0g |
| yr2\_bl | Baseline Year [2st Data Entry] | %8.0g |
| missing\_page\_bl | Baseline | %8.0g |
| version1\_bl | Baseline | %9.0g |
| version2\_bl | Baseline | %8.0g |
| version3\_bl | Baseline | %8.0g |
| version4\_bl | Baseline | %8.0g |
| version5\_bl | Baseline | %8.0g |
| version6\_bl | Baseline | %8.0g |
| date\_bl | Baseline Date of Interview | %10s |
| AEnumerator\_name\_bl | Baseline Enumerator Name | %19s |
| ra\_id\_bl | Baseline Enumerator ID | %8.0g |
| gender\_2016\_bl | Baseline Gender | %8.0g |
| pq1\_bl | Baseline 1). Circle the missing letter of the alphabet | %1s |
| pq2\_bl | Baseline 2) Circle the correct name for the objects below | %1s |
| pq3\_bl | Baseline 3) Pezani nambala yomwe ikusowa mundondomeko ya manambala awa | %1s |
| pq4\_bl | Baseline 4) Zungulizani nambala yopezeka mukaphatikiza nambala mwapatsidwazi | %1s |
| pq5\_bl | Baseline 5) Chithunzi ichi ndi cha\_\_\_\_\_\_\_\_\_\_\_ | %9s |
| pq6\_bl | Baseline 6) Zungulizani mawu omwe akuyimira dzina mchiganizo ichi | %1s |
| start\_time\_min\_2016\_bl | Baseline | %9s |
| AEnd\_time\_hours\_bl | Baseline | %10.0g |
| AEnd\_time\_min\_bl | Baseline | %10.0g |
| EnglishTotal\_Percent\_bl | Baseline percentage score for English | %9.0g |
| MathsTotal\_Percent\_bl | Baseline percentage score for Maths | %9.0g |
| ChichewaTotal\_Percent\_bl | Baseline percentage score for Chichewa | %9.0g |
| Total\_percent\_bl | Baseline overall percentage score for English, Math and Chichewa | %9.0g |
| division\_num\_bl | Baseline Division\_num | %10.0g |
| division\_nam\_bl | Baseline Division\_nam | %16s |
| district\_num\_bl | Baseline District\_num | %8.0g |
| district\_nam\_bl | Baseline District\_nam | %19s |
| strata\_bl | Baseline strata | %16.0g |
| wtg\_sch\_bl | Baseline school weights | %9.0g |
| wtg\_stu\_bl | Baseline student weights | %9.0g |
| wtg\_tch\_bl | Baseline teacher weights | %9.0g |
| total\_klg\_score\_bl | Baseline mean average score across 3 subjects) | %9.0g |

Table 4. Variable Names, Labels, and Format of the Derived Data at the Teacher Level.

| var\_name | label | format |
| --- | --- | --- |
| school\_id | school id | %10.0g |
| teacher\_id | teacher id | %8.0g |
| ODL | Highest level of teacher training = ODL | %9.0g |
| IPTE | Highest level of teacher training = IPTE | %9.0g |
| ODL2 | 1=ODL, 0=IPTE | %9.0g |
| leader\_yes | Has Leadership role in school | %9.0g |
| tch\_present | Teacher Present | %9.0g |
| tch\_present\_tch | Teacher present and teaching | %9.0g |
| a63a | 63 a) preparing lesson plans. | %9.0g |
| a63b | 63 b) teaching students of my own class | %9.0g |
| a63c | 63 c) marking and evaluating homework/ classroom work | %9.0g |
| a63d | 63 d) school administrative tasks | %9.0g |
| a63e | 63 e) helping other teachers | %9.0g |
| a63f | 63 f) other | %9.0g |
| time\_prepare\_lesson | time spent on preparing lesson plan(in hours) | %9.0g |
| time\_tch\_stu | Time spent teaching students | %9.0g |
| time\_mark\_hw | Time spent marking hw | %9.0g |
| time\_sch\_adm | Time spent on school admin | %9.0g |
| time\_help\_oth | Time spent helping others | %9.0g |
| time\_other | Time spent on other activities | %9.0g |
| time\_all |  | %9.0g |
| lesson\_based\_on\_txt | lesson plan is based on textbooks | %9.0g |
| ht\_observe |  | %37.0g |
| tch\_reward |  | %15.0g |
| perm\_tch | Permanent teacher | %9.0g |
| PT4 | Grade 4 | %9.0g |
| PT3 | Grade 3 | %9.0g |
| PT2 | Grade 2 | %9.0g |
| PT1 | Grade 1 | %9.0g |
| grade\_missing | Grade missing | %9.0g |
| tenure | Tenure as teacher | %9.0g |
| tenure\_sch | Tenure at this school | %9.0g |
| male | Teacher is male | %9.0g |
| tch\_married | Teacher Married | %9.0g |
| tch\_educ\_years | Years in education | %9.0g |
| tch\_educ\_years\_squared | Years in education (squared) | %9.0g |
| tch\_age | Teacher age | %9.0g |
| tch\_age\_square | Teacher age (squared) | %9.0g |
| tch\_same\_vill\_sch | Teacher same village | %9.0g |
| tch\_same\_district\_sch | Teacher same district | %9.0g |
| tch\_commu\_time | Average commute time | %9.0g |
| num\_pri\_age\_child | Number of primary aged children | %9.0g |
| pri\_age\_child\_yes | Primary aged children | %9.0g |
| child\_attend\_same\_sch | Children attend same school | %9.0g |
| basic\_qualification | Teacher has basic qualifications | %9.0g |
| above\_high\_qualification | Teacher has above high-school qualifications | %9.0g |
| transfer\_better\_amenity | Transferred from school because of better amenities | %9.0g |
| tch\_trained | Teacher trained | %9.0g |
| years\_since\_trained | Years since trainng | %9.0g |
| district\_num | District number - name is in a separate table | %8.0g |
| district\_nam | District Name | %19s |
| same\_district\_salary | Salary received in same district | %9.0g |
| tch\_salary | tch\_salary | %9.0g |
| hsa\_yes | HAS received | %9.0g |
| amount\_hsa | Amount of HAS | %9.0g |
| unpaid\_claim\_yes | teacher has unpaid claim including salary delays | %9.0g |
| salary\_delay\_yes | Salary has been delayed | %9.0g |
| num\_salary\_delay | number of salary delays in the past 12 months | %9.0g |
| other\_non\_paid\_claim\_yes | teacher has any other non-paid claims other than salary delay | %9.0g |
| eff\_score\_self | Teachers self perception | %9.0g |
| eff\_score\_ht | Perception of head teacher | %9.0g |
| eff\_score\_lower\_pri | Perception of lower primary teachers | %9.0g |
| eff\_score\_upper\_pri | Perception of upper primary teachers | %9.0g |
| output\_effi\_score | % of students teacher expects will continue to secondary | %9.0g |
| in\_service\_train | Received in-service training | %9.0g |
| tch\_age\_pt4 | age of teacher if in grade 4 | %9.0g |
| source | source | %8.0g |
| division\_num | Division\_num | %10.0g |
| division\_nam | Division\_nam | %16s |
| EnglishTotal\_Percent | English percent score | %9.0g |
| MathsTotal\_Percent | Math percent score | %9.0g |
| ChichewaTotal\_Percent | Chichewa percent score | %9.0g |
| Total\_score\_percent\_tch | Total percent score | %9.0g |
| ENG\_hard\_item\_percent | English hard item percent score | %9.0g |
| MATH\_hard\_item\_percent | Math hard item percent score | %9.0g |
| tch\_female | Teacher female | %9.0g |
| strata | strata | %16.0g |
| wtg\_sch | school weights | %9.0g |
| wtg\_stu | student weights | %9.0g |
| wtg\_tch | teacher weights | %9.0g |
| km\_to\_nid |  | %10.0g |
| status\_distance | Distance status | %27.0g |

1. At the time of the report writing in November 2021. [↑](#footnote-ref-1)
2. Not clear what the process is from harmonized to clean [↑](#footnote-ref-2)
3. This dofile imports a file that seems to created outside the system 01\_MLSS\_SAMPLE/02\_Data/962\_IE\_IDENTIFICATION\_COMPONENT\_101118\_complete\_infor [↑](#footnote-ref-3)
4. The support of the World Bank to access these secondary sources may be required. [↑](#footnote-ref-4)