

User Notes for Uganda Round 6 Follow-up Household and Female Survey Dataset, Version 1

Disclaimer: PMA cannot provide in-depth support for data analysis or data related questions, however, to assist the end-user, explanation of some variables is provided below.

Sampling

PMA2019/ Uganda Round 6 Follow-up (UGR6-FU) survey was designed to collect a longitudinal data by administering a follow-up survey to the female participants of PMA2018/Uganda Round 6 survey.

PMA2018/Uganda Round 6 (UGR6) survey used a two-stage cluster design with urban-rural and region as strata. It used the same set of 110 enumeration areas (EAs) as those that were selected in the previous round and drawn by the Uganda Bureau of Statistics from its master sampling frame. In each EA and households were listed and mapped, with 44 households randomly selected. Households were surveyed and occupants enumerated. Women participants were asked if they would be willing to participate in future surveys. UGR6 data collection was conducted between April and May 2018.

For UGR6-FU survey, the resident enumerators went back to the same households (44 per each EA) randomly selected in UGR6 survey and sought consent for participation to the UGR6-FU. In addition, the women who consented to the future surveys in UGR6, who were still at the same households and were still willing to be interviewed, were interviewed. Women who had moved to different households, which were not part of the randomly selected household sample but still located within the EAs, were followed and interviewed. However, no interviews were done for women who had moved outside the EAs.

The final sample for UGR6-FU survey included 3,597 households and 2,823 de facto females. UGR6-FU data collection was conducted in May and June of 2019.

Analytic Sample

PMA2020 analyses include only observations from completed household interviews. The female sample includes only completed female interviews from completed households. The majority of indicators include only de facto women (women who slept in the household the night before). All observations, however, are included in the dataset to allow end users to calculate response rates.

Type of Datasets released

The UGR6-FU HHQFQ Survey data is released as 3 datasets:

1. PMA2019_UGR6FU_HHQFQ_Cross-sectional_v#_Date
2. PMA2019_UGR6FU_HHQFQ_Panel_v#_Date
3. PMA2019_UGR6FU_HHQFQ_Panel_MergeKey_v#_Date

The Cross-sectional dataset contains data of respondents which could be used for cross-sectional analyses. The Panel dataset contains the data of respondents which could be used to merge with UGR6 HHQFQ survey dataset for panel analyses. The Panel Merge Key dataset contains the keys which could be used use to merge UGR6 and UGR6-FU datasets

General Variables

SIF variables: Date and time variables are provided in both string format and as Stata Internal Format (SIF) values. The variable name of any variable that has been changed into SIF is appended with SIF (e.g. **system_date** and **system_dateSIF**). For all questions requiring a date entry, if the respondent answered either “Do Not Know” or refused to answer the question, the date was recorded as January 1, 2020.

Select multiple variables: Some questions allow for the selection of multiple answers. These variables are in string format and the values are the concatenation of answer choices (e.g. if a household respondent said that they use two sources of water, such as a protected well and rainwater, the value of the observation would read “protected_well rainwater”). Multi-select options are generally, though not always, transformed into binary variables for analysis.

Variable Response Options

Select one: Most select one numeric variables have consistent values for option choices across countries and rounds (e.g. **marital_status==1** is equivalent to currently married in all countries). Exceptions include the variables **school**, **fp_provider**, **floor**, **roof**, and **walls**, which have country-specific options and numbering.

Select multiple: Similarly, most select multiple variables have the same response options across countries. Some select multiple variables, however, such as **assets**, have answer options that vary across countries.

See the *UGR6 Follow-Up HHQFQ Codebook* for complete details on variables and answer choices.

Specific Variables

metainstanceID: The unique ID generated by ODK for each household form submitted to the central server. Thus, it is unique for each household but will be repeated for members within the same household.

memberID: The a unique ID for each member within the household.

FQmetainstanceID: The unique ID generated by ODK for each female form submitted to the central server. It is unique for each female interviewed in *the Uganda Round 6 Follow-up Survey*.

HH_type: This variable indicates whether the household surveyed contributes to the *Uganda Round 6 follow-up* panel only dataset or both cross-sectional and panel datasets.

FQ_type: This variable indicates whether the female surveyed contributes to the *Uganda Round 6 follow-up* panel only dataset or both cross-sectional and panel datasets.

EA_ID: The primary sampling unit, enumeration area (EA), masked with a random number for anonymity. The same random number is applied to the same EA across multiple rounds.

wealthquintile: Categorization of wealth into categories. The continuous variable **score** is included to allow for construction of various wealth categories.

calendar_full: This 25-character string variable was generated using the data from contraceptive calendar.

current_recent_methodnum, current_methodnum, recent_methodnum: The numbering scheme for contraceptive methods is consistent across all PMA2020 countries. For example, female sterilization is equal to 1 in every PMA2020 country, whether or not there are any reported uses of female sterilization in the dataset. In some countries, therefore, the numbering will be non-consecutive if some method choices are not selected.

cp, mcp, tcp: Variables that identify current users of any contraceptive method (**cp**), a modern contraceptive method (**mcp**), and a traditional contraceptive method (**tcp**) are included in publicly available datasets so that PMA2020 estimates involving current contraceptive use and method mix can be replicated. Values for these variables are 0 (no) or 1 (yes). PMA2020 codes **cp**, **mcp**, and **tcp** based on the variable **current_methodnum** with the following caveats:

- Women who report not being a current user of contraception (**current_user=0**), but who report using EC in the past 12 months (**recent_methodnum=8**. emergency) are coded as **cp=1** and **mcp=1**. During analysis, current method is classified as EC in the method mix. The variables **current_methodnum_rc** and **recent_methodnum_rc** reflect this.
- Women who report using LAM as a current method (**current_methodnum=14**. LAM) must satisfy the three conditions listed below to be coded as **mcp=1**. If any of these conditions are not met, these women are coded as **tcp=1**. During analysis, current method is classified as LAM or traditional method. The variable **current_methodnum_rc** reflects this.
 1. Less than six months post-partum
 2. Amenorrheic
 3. Indicating that they are using LAM with the intention of preventing pregnancy
- Women who report female sterilization as their first contraceptive method (**first_methodnum=1**. female sterilization), but who do not report currently using female sterilization are coded as **cp=1** and **mcp=1**. During analysis, current method is classified as female sterilization in the method mix. The variable **current_methodnum_rc** reflects this.

Contraceptive Calendar Description

Panel Merge Key Variables

R6FU_FQmetainstanceID: This variable is the same as **FQmetainstanceID** variable of the *Round 6 Follow-up Panel (UGR6-FU) dataset*. This variable is used to merge in the Panel Merge Key data to the UGR6-FU dataset to prepare the UGR6-FU dataset to merge with the *Uganda Round 6 (UGR6) dataset* for panel analysis.

R6FQmetainstanceID: This variable is the same as the **FQmetainstanceID** variable of the *Uganda Round 6 (UGR6) dataset*. This variable is used to merge the *Uganda Round 6 Follow-up Panel (UGR6-FU) dataset* with the *Uganda Round 6 (UGR6) dataset*. See *Dataset Merging Guide* below for merge instruction.

fup_code: This variable indicates the status of female who consented to be followed up in the *Uganda Round 6 survey*.

Dataset Merging Guide for Panel Analysis

In order to do panel analyses, perform the following steps to merge UGR6 and UGR6-FU HHFQ datasets:

1. Prepare the UGR6-FU dataset by merging in the Keys:
 - 1.1. In the UGR6-FU dataset, drop observations with no female data, rename the variable **FQmetainstanceID** as **R6FU_FQmetainstanceID** and save the dataset.

Example Stata code:

```
drop if FQmetainstanceID==.
rename FQmetainstanceID R6FU_FQmetainstanceID
merge 1:m R6FU_FQmetainstanceID using "PMA2019_UGR6FU_HHQQFQ_Panel_MergeKey_v#_Date.dta",
    gen (keymerge)
save "filepath/UGR6FU_PanelPrep.dta", replace
```

2. Prepare the UGR6 dataset and merge in UGR6-FU dataset.
 - 2.1. In UGR6 dataset, drop observations with no female data, and rename the variable **FQmetainstanceID** as **R6FQmetainstanceID**.

Example Stata code:

```
drop if FQmetainstanceID==.
rename FQmetainstanceID R6FQmetainstanceID
```

- 2.2. Rename other variables of interest which are found in both UGR6 and UGR6-FU datasets by adding prefixes or suffixes to them to prevent the loss of data after the merge. Reference the UGR6-FU codebook to identify those common variables. Keep only the variable of interest for panel analyses.

Example Stata code:

```
rename current_method R6current_method
```

```
rename FQweight R6FQweight
keep R6FQmetainstanceID R6current_method R6FQweight
```

2.3. Merge UGR6 dataset with the UGR6-FU dataset prepared with key in step 1.1

Example Stata code:

```
merge m:1 R6FQmetainstanceID using "(UGR6FU_PanelPrep.dta)", gen (R6merge)
```

GPS Variables

GPS coordinates are not released in this dataset. For some PMA2020 countries a separate dataset of displaced GPS coordinates at the EA level is released.

Notes for Missing Data

In Stata, missing data is expressed as "." in the cell. Generally, Stata commands perform computations of any type handle missing data by omitting the row with the missing values. However, this may vary across commands. PMA does not impute missing values. Missing data in datasets should be studied and/or treated before proceeding to analysis.

Reasons for missing data:

Normal situations:

1. Incomplete forms: If a household, female, or SDP form is not marked as completed (HHQ_result, FRS_result, SDP_result not equal to 1), the observation is likely to miss most of the information. Incomplete forms should not be included in the analysis.
2. Observations that are ineligible for subsequent forms: Only eligible respondents will receive subsequent forms. For example, males and ineligible females will not receive female questionnaires in family planning surveys, hence their observations will have all missing values in female forms.
3. Question not administered due to skip logic: PMA surveys use ODK's skip logic function. The subsequent questions are administered selectively based on the respondent's previous answers. Irrelevant or inapplicable questions are skipped. For example, a woman who is not a contraceptive user will not be asked questions about contraceptive usage subsequently.

Uncommon situations:

1. Lost forms: Due to technical constraints in some challenging data collection areas, forms can be lost in the process of data submission. Although most forms were recoverable, there are occasionally a few that cannot be found. For example, an observation from an eligible woman with completed female form information but missing household form information, or vice versa. These observations may be dropped based on analysis needs.
2. Missing due to incorrect skip logic: PMA surveys were conducted under rigorous quality control. However, in rare cases, there can be incorrect skip logic, which skipped a question that was supposed to be administered, resulting in missing values. These errors are documented in the PMA codebook, which can be

downloaded from PMA website. It's not necessary to drop the entire observation since this will likely affect only a few questions.

Distinguish missing data from negative values:

1. -99: No response. The respondent was administered with the question but did not provide an answer. PMA survey requires consent from the respondent and the respondent has the right to refuse to answer any questions at any point. -99 is recorded to reflect that the respondent did not provide an answer to a certain question.
2. -88: Did not know. The respondent consented to answer a specific question but without knowing the answer.
3. -77: Not applicable. The question is administered to the respondent but not applicable to the respondent's situation.

Dataset Version Updates

Any updates made to datasets after their initial release will be documented here.

In January 2017, all previously released datasets were modified as below:

- The value of **age_at_first_use_children** is 0 for women who have ever used family planning and who have never given birth. Previously, such women had a missing value for `age_at_first_use_children`.
- The values for **water_sources_main_drinking** and **water_sources_main_other** equal the value of `water_sources_all` if a household has one water source. Previously, such households may have had a missing value for these variables.
- The value for **sanitation_main** equals the value of `sanitation_all` if a household has one sanitation facility. Previously, such households may have had a missing value for this variable.

All datasets released after January 2017 will have these changes included.

Suggested Citation

Suggested citation: Makerere University, School of Public Health at the College of Health Sciences and The Bill & Melinda Gates Institute for Population and Reproductive Health at The Johns Hopkins Bloomberg School of Public Health. *Performance Monitoring and Accountability 2020 (PMA2020) Household and Female Survey Round 6 Follow-up, PMA2019/Uganda-R6FU*. 2019. Uganda and Baltimore, Maryland, USA. doi:10.34976/9x2b-nd72

To report errors or inconsistencies:

Please email datamanagement@pma2020.org