

## User notes for PMA2014/Nigeria Round 1 (Kaduna & Lagos) Household and Female data, version 1

**Disclaimer:** PMA2020 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.

### Generic

**SIF variables:** Data 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. The values for these variables are the concatenation of answer choices (e.g. if a household respondent said that they use two sources of water, choices a and c, the value of the observation would read “a c”). Multi-select options are generally, though not always, transformed into binary variables for analysis.

**Country specific variables:** All variables in PMA2020 have consistent values for option choices across countries (e.g. marital\_status==1 is equivalent to currently married in all countries) with the exception of the following

1. Livestock questions: The specific livestock options (cow, rabbit etc) vary across countries
2. school: education categories for female schooling vary across countries
3. fp\_provider: provider of current or most recent method of family planning vary across countries
4. roof/wall/floor: Household materials vary across country
5. assets: The household assets used to construct wealth scores vary across countries as do the binary variables that are created from the multi-select asset question
6. wealthquintile/wealthtertile: In some countries, wealth quintiles are provided, in other wealth tertiles. The continuous variable score is included to allow for reconstruction of various wealth categories.

### Specific variables

**EA\_ID/ Cluster\_ID:** The primary sampling unit. In most countries, EA\_ID identifies the primary sampling unit. In Nigeria, enumeration areas were be too small to serve

as the primary sampling unit. In this case, EAs were clustered and the variable Cluster\_ID serves as the primary sampling unit.

**metainstanceID:** metainstanceID is the unique ID generated by ODK for each form submitted to the central server. For PMA2020, the variable metainstanceID is unique for each household but will be repeated within the household. **memberID** will provide a unique ID for each person within the household.

**FQmetainstanceID:** FQmetainstanceID is the unique ID generated by ODK for each female form submitted to the central server. For PMA2020, the variable FQmetainstanceID is unique for each female surveyed.

**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:

1. 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; however, current method is not changed in the data that is publicly available.
2. 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; however, current method is not changed in the data that is publicly available.
  - a. Less than six months post-partum
  - b. Amenorrheic
  - c. Indicating that they are using LAM with the intention of preventing pregnancy
3. 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; however, current method is not changed in the data that is publicly available.

### **GPS Variables**

No GPS coordinates for either household or service delivery points will be released for any reason.

### **Sampling**

The PMA2014/Nigeria Round 2 survey in Lagos used a two-stage cluster design. A sample of 37 index enumeration areas (EAs) in the urban stratum was drawn from the National Population Commission's master sampling frame along with a list of EAs contiguous to the index EA. The EAs in Nigeria are usually small, with approximately 48 households on average. To create clusters with a minimum of 200 households, each index EA was listed and mapped. If there were fewer than 200 households identified, the next EA in the list of contiguous EAs was listed and mapped in its entirety. If the total number of households was still less than 200, the entirety of the third EA was listed and so on. Each cluster of EAs serves as the primary sampling unit from which 35 households and up to 3 private health facilities were randomly selected. Households were surveyed and occupants enumerated. All eligible females age 15 to 49 were contacted and consented for interviews. The final completed sample included 974 households, 764 females and 93 health facilities. Data collection was conducted between September and October 2014.

The PMA2014/Nigeria Round 2 survey in Kaduna used a two-stage cluster design with urban-rural as strata. A sample of 66 index enumeration areas (EAs) was drawn from the National Population Commission's master sampling frame along with a list of EAs contiguous to the index EA. The EAs in Nigeria are usually small, with approximately 48 households on average. To create clusters with a minimum of 200 households, each index EA was listed and mapped. If there were fewer than 200 households identified, the next EA in the list of contiguous EAs was listed and mapped in its entirety. If the total number of households was still less than 200, the entirety of the third EA was listed and so on. Each cluster of EAs serves as the primary sampling unit from which 35 households and up to 3 private health facilities were randomly selected. Households were surveyed and occupants enumerated. All eligible females age 15 to 49 were contacted and consented for interviews. The final sample included 2,194 households, 2,569 females and 136 health facilities. Data collection was conducted between September and October 2014.

### **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.

## 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:

1. 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.
2. 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.
3. 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.

### To report errors or inconsistencies:

Please email [datamanagement@pma2020.org](mailto:datamanagement@pma2020.org)