

# User Notes for PMA Kenya Phase 1 Service Delivery Point Baseline 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.*

## PMA

Performance Monitoring for Action (PMA), formerly PMA2020, builds on the previous success of PMA2020 surveys in Kenya and focused on collecting routine data on key global indicators in family planning and reproductive health, while expanding content area to address questions of contraceptive decision-making and autonomy, in order to better understand the determinants and consequences of unique contraceptive use and patterns of use in Kenya. These are measured through three-related data collection activities: household and female surveys (HQFQ) producing both cross-sectional and longitudinal data, Service Delivery Point panel surveys (SQ), and an SDP client exit surveys (CQ).

In Kenya, a cross-sectional and panel household and female surveys (HQFQ) are conducted annually, with follow-up for the panel occurring at Year 2 and 3. The Service Delivery Point Survey (SQ) panel baseline data is collected at Year 1 and follow-up data will be collected annually. The Service Delivery Point Client Exit Survey (CQ) is conducted biannually with a baseline and a follow-up occurring 6 months after the baseline enrollment each year.

## Sampling

Kenya Phase 1 (KEP1) Service Delivery Point Baseline Survey includes 308 enumeration areas (EAs) of selected counties. The national sample is composed of the existing 115 PMA2020 EAs which were refreshed and supplemented with the new EAs. The new EAs were drawn using the same stratified cluster design with urban-rural strata by the Kenya Bureau of Statistics following the pre-existing sampling frame used for PMA2020. The results are national and county-level representative. The final sample included 945 SDPs (94.6%) which completed the interview. Data collection was conducted between November and December 2019.

## 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. **year\_open** and **year\_openSIF**).

**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 service delivery point respondent said that the facility offers female sterilization counseling, provision, and charges for the provision, the response for the variable `offered_female_ster` would read "counseled provided charge"). 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 all PMA countries (e.g. fees ==1 is equivalent to charging contraceptive fees in all PMA countries). Exceptions include the variables which have country-specific options and numbering. For examples:

- Geographic variable (e.g. **region, county**): geographic variable names and response options vary across countries
- **facility\_type**: facility types vary across countries
- **postpartum**: options of items discussed during postpartum visits vary slightly across countries

**Select multiple:** Similarly, most select multiple variables have the same response options across all PMA countries, with exceptions for variables with choices that are country specific.

See the PMAET SQ Master Codebook for complete details on variables and answer choices for each survey.

## Specific Variables

**EA\_ID:** The primary sampling unit masked with a random number for anonymity. The same random number is applied to the same EA across multiple survey years.

**EAserved#:** Some SDPs serve more than one EA. The EAserved# variables indicate the additional EAs that a given facility serves, if any. Information regarding which EAs an SDP serves comes from the country/local government. Only public facilities are assigned to serve more than one EA.

**RE\_ID:** Identification number of the resident enumerator (RE), or interviewer. RE names are masked with the PMA-Ethiopia random numbers in the SDP dataset. The same random number is applied to the same REs across different surveys of the PMA-Ethiopia grant.

**PMA2020\_RE\_ID:** The resident enumerators (REs), or interviewers, who were involved in the previous PMA2020 surveys (2014-2018), also had PMA2020 IDs. This ID is consistent for all survey rounds of the PMA2020 grant.

**facility\_ID:** Facility names are masked with randomly generated values. The same random number is applied to the same facility across different surveys of the PMA-Ethiopia grant.

**PMA2020\_facilityID:** The facilities that were also included in the previous PMA2020 surveys (2014-2018), also had PMA2020 IDs. This ID is consistent for all survey rounds of the PMA2020 grant.

**metainstanceID:** A unique ID generated by ODK for each form submitted to the central server. This variable is unique for each SDP.

## GPS Variables

GPS coordinates are not released in this dataset.

## 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.

### **Dataset Citations**

Suggested citation: International Centre for Reproductive Health Kenya (ICRHK) and The Bill & Melinda Gates Institute for Population and Reproductive Health at The Johns Hopkins Bloomberg School of Public Health. *Performance Monitoring for Action (PMA) Kenya Phase 1: Service Delivery Point Baseline Survey PMA/Kenya-P1-SQ-Baseline*. 2019. Kenya and Baltimore, Maryland, USA.

### **To report errors or inconsistencies:**

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