

Performance Monitoring for Action (PMA) Kenya Covid-19 Female Follow-up Survey (2020) Dataset, Version 1.1

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 and Analytic Sample

PMA Kenya collects nationally and county-level representative data on knowledge, practice, and coverage of family planning services in 308 enumeration areas selected using a multi-stage stratified cluster design with urban-rural strata. The COVID-19 phone survey was conducted in 11 counties in June 2020 among females aged 15-49 at the time of the COVID-19 Survey who were interviewed at the PMA Kenya Phase 1 survey, consented to follow-up and owned or had access to a phone (68% of the Phase 1 population). Of the 6,377 eligible respondents, 5.1% were not reached. Of those reached, 5,986 completed the survey for a response rate of 93.9% among contacted women.

PMA Kenya is led by the Ministry of Health in collaboration with International Centre for Reproductive Health Kenya (ICRH-K), National Council for Population and Development, and Kenya National Bureau of Statistics. Overall direction and support are provided by the Bill & Melinda Gates Institute for Population and Reproductive Health at the Johns Hopkins University and Jhpiego. Funding is provided by the Bill & Melinda Gates Foundation.

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**).

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 respondent said that they learned about Covid-19 from two sources of media, such as social media and radio, the value of the observation would read “social_media radio”). 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. marital_status==1 is equivalent to currently married in all PMA countries). An exception include the variable **school**, which has country-specific options and numbering.

Select multiple: Similarly, most select multiple variables have the same response options across all PMA countries. Some select multiple variables, however, have answer options that vary across countries.

See the PMA Kenya Covid-19 Codebook for complete details on variables and answer choices for the 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 surveys.

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

PMA2020_RE_ID: PMA2020 identification number of resident enumerator (RE), or interviewers. Only available for resident enumerator (RE), or interviewers, who were involved in the PMA2020 surveys. This value is missing for those who were not involved in PMA2020 surveys.

metainstanceID: metainstanceID is the unique ID generated by ODK for each form submitted to the central server.

Female_ID: A unique ID of each female respondent. The same ID is given to the same female respondent across different PMA surveys.

current_methodnum: The numbering scheme for contraceptive methods is consistent across all PMA countries. For example, female sterilization is equal to 1 in every PMA 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 PMA 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 caveat:

1. Women who report not being a current user of contraception (**current_user=0**), but who report using EC since Coronavirus (Covid-19) restrictions began (**covid_ec_used =1**) are coded as **cp=1** and **mcp=1**. During analysis, current method is classified as EC in the method mix. The variable **current_methodnum_rc** reflect this.
2. No recoding of LAM and female sterilization was done (unlike in other PMA and PMA2020 Household and Female Datasets), since questions on the data needed to recode was not asked in the Covid-19 survey.

FQweights: COVID-19 survey weights were generated for women aged 15-49 at the time of the COVID-19 survey, who completed the Kenya Phase 1 survey, consented to follow-up, provided a valid phone number, and completed the

COVID-19 survey. These weights were calculated using the female weight from the Phase 1 survey, adjusting for loss-to-follow-up weight, that is, the inverse of predicted probability of having a completed COVID-19 survey. The log odds of having completed the COVID-19 survey was modeled as a linear combination of age, education, marital status, wealth, and residence at the Phase 1 survey. The COVID-19 survey weight was further adjusted for selectivity due to phone number ownership using a similar inverse probability weighting approach.

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 Covid-19 form is not marked as completed (COV_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. 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. 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. 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.

On 28 February 2021, a new version of this dataset has been released with the following updates:

- Covid-19 female follow-up survey weight variable (**cFQFUweight**), which was missing the previous version, was added.
- Female survey weight from the phase 1 baseline survey (**FQweight**) was removed in this version since this is not the weight of this survey and this variable could be obtained from the phase 1 baseline dataset.

Dataset Citations

International Centre for Reproductive Health Kenya (ICRHK), the Bill & Melinda Gates Institute for Population and Reproductive Health at The Johns Hopkins Bloomberg School of Public Health, and Jhpiego. *Performance Monitoring for Action (PMA) Kenya Covid-19 Survey (Version 1)*. 2019. Kenya and Baltimore, Maryland, USA. doi: 10.34976/b8cb-j456.

To report errors or inconsistencies:

Please email datamanagement@pmadata.org