

## User notes for PMA2017/Kenya Nutrition Round 1 Household and Female survey, version 2

**Disclaimer:** PMA2020 cannot provide in-depth support for data analysis or data related questions, however, to assist the end-user, some explanation of the data is provided below.

### Variables

**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”). For every string select multiple variable, binary numeric yes/no variables for each response option were generated (e.g. the aforementioned respondent would have “yes” values for the generated variables protected\_dug\_well and rainwater and “no” values for generated variables for other water sources).

**Dropped variables:** Some variables were dropped from the dataset due to incorrect skip patterns or choice options in the questionnaire. In Kenya Round the following variables were dropped: FQpreviousPMA, first\_birth\_twins, first\_birth\_alive, first\_birth\_death, first\_birth\_deathSIF, nb\_death, nb\_deathSIF, ner\_extra\_food, and ne\_extra\_food were dropped.

### Sampling

The PMA2020/Kenya Nutrition Round 1 survey used a multi-stage stratified cluster design with urban-rural and selected 11 counties as strata. A sample of 151 enumeration areas (EAs) was drawn by the Kenya National Bureau of Statistics from its master sampling frame for the Kenya Round 5 PMA2020 family planning survey and these EAs were used for the nutrition survey. In each EA, 140 households were randomly selected. Occupants in selected households were enumerated using the household roster section of the household questionnaire. In eligible households, defined as a household with both a child under two and a female age 10-49, the rest of the household questionnaire was administered. In these eligible households, all women age 10-49 were contacted and consented for the female questionnaire interview.

Data collection was conducted between May and August 2017.

### Dataset structure

As explained above, two questionnaires were administered during this survey. The household questionnaire collected information about household occupants, and for eligible households (i.e. a household with both a child under two and a female age 10-49), about household characteristics including assets, water, sanitation, and hygiene, and food security. The female questionnaire included questions about the female respondent and all

children under 5 years she cares for. The female questionnaire included questions pertaining to:

- The woman age 10-49 in the household who responded to this questionnaire, including:
  - The respondent's current pregnancy. These variable names begin with "ac\_".
  - The respondent's most recent pregnancy completed in the last 24 months. These variable names begin with "ar\_".
  - The respondent's nutritional status and diet. These variable names begin with "nh\_".
- The youngest child under 2 years old the respondent cares for (biological or non-biological). These variable names begin with "ny\_".
- The youngest child between 2 and 5 years old the respondent cares for (biological or non-biological). These variable names begin with "ner\_".
- All other children between 2 and 5 years old the respondent cares for who were not previously discussed in the survey (biological or non-biological). These variable names begin with "ne\_".

For example, if a woman cares for a newborn, a 2 year old, a 3 year old, and a 4 year old, the survey would ask her about her birth history, her most recent pregnancy (variables beginning in ar\_), the newborn (variables beginning in ny\_), the 2 year old (variables beginning in ner\_), and the 3 year old and the 4 year old (variables beginning in ne\_). Each child is a separate row in the dataset (i.e. children from the same household are not found in the same row of data).

The folder of materials for this study includes three datasets:

- The Household dataset: each observation is a member of the household. The dataset includes variables from the household questionnaire about household characteristics and members of the household from the household roster section of the questionnaire.
- The Female dataset: each observation is a member of a household who is female and age 10-49 years. The dataset includes variables from the female questionnaire that are about the woman: her birth history, her current pregnancy if she is pregnant, her most recent pregnancy if the pregnancy was within the last two years, and her nutritional status and diet.
- The Child dataset: each observation is one child and includes variables on the nutritional status of that child. Each child is a separate row in the dataset (i.e. children from the same household are not found in the same row of data).

Datasets from this study can be merged using FQmetainstanceID. Merge the Female dataset to the Household dataset using FQmetainstanceID. Merge the Child dataset to the Female dataset using FQmetainstanceID. Note that if the Female dataset and Child dataset are merged together, the female respondent will be repeated in the merged dataset.

### Sample size

The final sample included 19,930 households that completed a household roster (98.2% response rate), of which 4,624 were eligible to participate in the study and completed the rest of the household survey (23.2% eligible rate). The final sample also included 7,260 complete female surveys (99.2% response rate) including:

- 179 currently pregnant women
- 4,434 recently pregnant women (pregnant within the past 24 months; note: some women may have been currently pregnant and pregnant in the past 24 months)
- 1,242 females 10-14 years old
- 982 females 15-19 years old
- 5,036 females 20-49 years old

Female respondents who completed the female survey provided information on the following number of children:

- 4,602 children 0-23 months old
- 2,048 children 24-59 months old.

### Dataset version updates

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

In December 2018, version 2 of the dataset was released. Version updates are as follows:

- The variable **FQmetainstanceID** was added to the Household dataset.
- Duplicate female observations were removed from the Household dataset.

### To report errors or inconsistencies:

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