User notes for PMA2018/Niger (Niamey) Round 5 Service Delivery Point 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. year_open and year_openSIF). 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 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 (e.g. counseled_female_ster, provided_female_ster, etc.).

Country specific variables: All variables in PMA2020 have consistent values for option choices across countries (e.g. fees ==1 is equivalent to charging contraceptive fees in all countries) with the exception of the following

1. Geographic variable (e.g. region, county): geographic variable names and response options vary across countries
2. facility_type: facility types vary across countries
3. postpartum: options of items discussed during postpartum visits vary slightly across countries

Specific variables

EA: The primary sampling unit, i.e. enumeration area (EA). EAs are masked with random numbers in the household dataset and the SDP dataset. The random numbers are consistent in the two datasets, in future rounds of datasets, and can be used to match facilities with households in the same enumeration area.

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: The resident enumerator (RE), or interviewer. RE names are masked with random numbers in the household dataset and the SDP dataset and the random numbers are consistent in the two datasets and in future rounds of datasets.

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 SDP.

facility_ID: Randomly generated values mask facility names. Values will be consistent across rounds for the same SDP. New facility IDs indicate additional facilities selected between rounds.

GPS Variables

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

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.

Sampling

For the PMA2018/Niamey survey, a total of 33 enumeration areas (EAs) were selected using probability promotional to size, from a sampling frame provide by the Fourth General Census of Population and Housing, conducted by Niger's National Statistics Institute in 2012. Structures, households, and service delivery points (SDPs) were enumerated in each selected EA. The SDP sample included up to three public SDPs that serve an EA and up to three private SDPs within the EA. Data collection occurred from June to August 2018. In Niamey, a total of 30 service delivery points were interviewed, of which 21 are public.

Analytic sample

PMA2020 analyses include only observations from completed SDP interviews. However, all observations 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.

To report errors or inconsistencies:

Please email datamanagement@pma2020.org