User notes for PMA2017/Burkina Faso 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: Most variables in PMA2020 have consistent values for option choices across countries (e.g. fees ==1 is equivalent to charging contraceptive fees in all countries). See the codebook for exceptions, such as geographic variables (e.g. region, county) and facility_type.

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.

Sampling
PMA2017/Burkina Faso Round 5 used a two-stage cluster design with urban-rural strata. A sample of 83 enumeration areas (EAs) was drawn from the l'Institut National de la Statistique et de la Démographie (INSD) master sampling frame. Each EA was listed and mapped. Public facilities were included if a selected EA fell within the catchment area. Private facilities were included if they fell within the boundaries of the EA. Data collection was conducted between November 2017 and January 2018. The final completed sample was 130 SDPs.

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