

Notes aux utilisateurs des données ménages et femmes de PMA2018/Burkina Faso Vague 6, Version 1

Attention : PMA2020 n'est pas en mesure d'apporter un soutien approfondi à l'analyse des données ou aux questions liées aux données, cependant, pour aider les utilisateurs, des explications de certaines variables sont fournies ci-dessous.

Variables générales

Variables SIF : Les variables de la date et de l'heure sont fournies à la fois au format chaîne de caractères, et comme valeurs SIF (Stata Internal Format). Le signe SIF est ajouté à la fin du nom de toute variable ayant été changée en SIF (ex : **system_date** et **system_dateSIF**). Pour toutes les questions nécessitant d'enregistrer une date, si l'enquêté(e) a répondu « Ne sait pas » ou refusé de répondre à la question, la date a été enregistrée au 1^{er} Janvier 2020.

Variables des réponses à choix multiple : Certaines questions permettent la sélection de plusieurs réponses. Ces variables sont au format chaîne et les valeurs sont la concaténation des modalités de réponse (ex : si un répondant du ménage déclare que son ménage utilise deux sources d'eau, comme par exemple un puits protégé et l'eau de pluie, la valeur de cette observation sera "protected_well rainwater"). Les réponses à choix multiples sont généralement, bien que pas toujours, transformées en variables binaires pour l'analyse.

Variables des modalités de réponses

Une seule réponse : La plupart des variables numériques qui ne permettent qu'une seule réponse ont les mêmes valeurs de modalités de réponses entre tous les pays (ex : **marital_status==1** correspond à « actuellement marié(e) » dans tous les pays). Les variables suivantes font exception : **school, fp_provider, floor, roof, et walls**, qui comprennent des modalités et numérotations de réponses propres à chaque pays.

Plusieurs réponses possibles : de la même manière, la plupart des variables à choix multiple ont les mêmes modalités de réponses dans tous les pays. Certaines variables à choix multiple, cependant, telle que **assets**, ont des modalités de réponses qui peuvent varier d'un pays à l'autre.

Consultez le Livre de code des questionnaires ménages et femmes pour les variables et modalités de réponses détaillées de chaque pays et vague d'enquête de PMA2020.

Variables spécifiques

EA_ID: Unité d'échantillonnage primaire masquée par un numéro aléatoire pour préserver l'anonymat. Le même numéro aléatoire est assigné à la même zone de dénombrement sur toutes les vagues d'enquête.

wealth: Certaines bases de données des pays comprennent des quintiles de bien-être économique (**wealthquintile**) tandis que d'autres utilisent des tertiles de bien-être économique (**wealthtertile**). Le **score** continu de la variable est inclus pour permettre la construction de différentes catégories de bien-être économique.

metainstanceID: metainstanceID est l'identifiant unique généré par ODK pour chaque formulaire envoyé sur le serveur central. Pour PMA2020, la variable metainstanceID est unique pour chaque ménage, mais se répète au sein du ménage. **memberID** assigne un identifiant unique à chaque personne au sein du ménage.

FQmetainstanceID: FQmetainstanceID est l'identifiant unique généré par ODK pour chaque formulaire femme envoyé sur le serveur central. Pour PMA2020, la variable FQmetainstanceID est unique pour chaque femme enquêtée.

current_recent_methodnum, current_methodnum, recent_methodnum: La numérotation des méthodes contraceptives est la même dans tous les pays de PMA2020. Par exemple, la stérilisation féminine est égale à 1 dans tous les pays de PMA2020, qu'il y ait ou non des cas déclarés d'utilisation de la stérilisation féminine dans l'ensemble des données. Dans certains pays, la numérotation ne sera donc pas consécutive si certaines options de méthode ne sont pas sélectionnées.

cp, mcp, tcp: Les variables identifiant les utilisatrices actuelles de méthodes contraceptives (**cp**), de méthodes contraceptives modernes (**mcp**), et de méthodes contraceptives traditionnelles (**tcp**) sont incluses dans les bases de données accessibles au public pour que les estimations de PMA2020 impliquant l'utilisation de l'utilisation contraceptive actuelle et de la distribution des méthodes puissent être répliquées. Les valeurs pour ces variables sont 0 (non) ou 1 (oui). PMA2020 codifie **cp**, **mcp**, et **tcp** en fonction de la variable **current_methodnum**, avec les précautions suivantes :

1. Les femmes déclarant ne pas utiliser actuellement la contraception (**current_user=0**), mais déclarant avoir utilisé la contraception d'urgence dans les 12 derniers mois (**recent_methodnum=8. emergency**) sont codifiées comme **cp=1** et **mcp=1**. Pendant l'analyse, la méthode actuelle est classifiée comme étant la contraception d'urgence dans la distribution des méthodes utilisées. Les variables **current_methodnum_rc** et **recent_methodnum_rc** reflètent cela.
2. Les femmes déclarant utiliser la méthode MAMA comme méthode de contraception actuelle (**current_methodnum=14. LAM**) doivent remplir les trois conditions ci-dessous pour être codifiées comme **mcp=1**. Si au moins une de ces conditions n'est pas remplie, ces femmes sont codifiées comme **tcp=1**. Pendant l'analyse, la méthode actuelle est classifiée comme MAMA ou méthode traditionnelle. La variable **current_methodnum_rc** reflète cela.
 - a. Moins de six mois post-partum
 - b. Aménorrhée
 - c. Déclare utiliser la méthode MAMA avec l'intention d'éviter une grossesse.
3. Les femmes déclarant utiliser la stérilisation féminine comme méthode contraceptive principale (**first_methodnum=1. female sterilization**), mais ne déclarant pas utiliser la stérilisation féminine actuellement sont codifiées comme **cp=1** et **mcp=1**. Pendant l'analyse, la méthode actuelle est classée comme stérilisation féminine dans la distribution des méthodes utilisées. La variable **current_methodnum_rc** reflète cela.

Variables GPS

Les coordonnées GPS ne sont pas publiées dans la base de données. Pour certains pays PMA2020, une base de données séparée des coordonnées GPS déplacées à l'échelle des zones de dénombrement est publiée.

Notes pour les données manquantes

Sous Stata, les données manquantes sont exprimées par un "." dans la cellule. Généralement, les commandes Stata effectuent des calculs de tout type pour traiter les données manquantes en omettant la rangée avec la valeur manquante. Cependant, ceci peut varier selon les commandes. PMA n'impute pas de valeur manquante. Les données manquantes dans les ensembles de données devraient être étudiées et/ou traitées avant de procéder à l'analyse.

Raisons des données manquantes :

Situations normales :

1. **Formulaires incomplets :** Si un formulaire ménage, femme ou SPS n'est pas enregistré comme étant terminé (HHQ_result, FRS_result, SDP_result non égal à 1), l'observation ne comportera probablement pas la plupart des informations. Les formulaires incomplets ne devraient pas être inclus dans l'analyse.
2. **Observations inéligibles aux formulaires subséquents :** Seul(e)s les enquêté(e)s éligibles se verront attribuer un formulaire subséquent. Par exemple, les hommes et les femmes non éligibles n'auront aucun formulaire femme dans les enquêtes sur la planification familiale, c'est pourquoi leurs observations afficheront des valeurs manquantes dans les formulaires femmes.
3. **Question non posée en raison de la logique de saut des questions :** les enquêtes PMA utilisent la fonction de saut de questions logique d'ODK. Les questions subséquentes sont posées de manière sélective en fonction des réponses précédentes du/de l'enquêté(e). Les questions non applicables ou non pertinentes sont sautées. Par exemple, une femme qui n'est pas utilisatrice de la contraception ne se verra pas poser de question subséquentes sur l'usage de la contraception.

Situations non courantes :

1. **Formulaires perdus :** En raison de contraintes techniques dans certaines zones difficiles où les données sont collectées, des formulaires peuvent être perdus dans le processus d'envoi des données. Bien que la plupart des formulaires aient été par le passé récupérés, quelques-uns n'ont pu être trouvés. Par exemple, une observation d'une femme éligible ayant un formulaire femme complet mais des informations manquantes dans son formulaire ménage, ou vice versa. Ces observations peuvent être abandonnées selon les besoins de l'analyse.
2. **Données manquantes en raison d'une logique de saut de question incorrecte :** les enquêtes PMA ont été conduites sous un contrôle qualité rigoureux. Cependant, dans de rares cas, il peut y avoir des logiques de saut inexacts qui ont omis une question qui aurait dû normalement être posée, donnant lieu à une donnée manquante. Ces erreurs sont documentées dans le livre de code de PMA, disponible en téléchargement sur le site

internet de PMA. Il n'est pas nécessaire d'abandonner l'observation entière, dans la mesure où ceci ne devrait affecter que quelques questions.

Distinguer les données manquantes des valeurs négatives :

1. -99 : Pas de réponse. On a posé à l'enquêté(e) la question mais ce(tte) dernier/ère n'y a pas répondu. L'enquête PMA requiert le consentement de la personne enquêtée, qui a le droit de refuser de répondre à toute question, à tout moment de l'enquête. -99 est enregistré pour refléter le fait que l'enquêté(e) n'a pas donné de réponse à une certaine question.
2. -88 : Ne sait pas. L'enquêté(e) a consenti à répondre à une question spécifique mais sans connaître la réponse.
3. -77 : Non applicable. La question a été posée à l'enquêté(e) mais elle n'était pas applicable à la situation du/de l'enquêté(e).

Échantillonnage

PMA2018/Burkina Faso Vague 6 a utilisé une méthode d'échantillonnage stratifiée à 2 niveaux (urbain et rural). L'échantillon de 83 zones de dénombrement (ZD) a été tiré à partir de la base de sondage de l'Institut National de la Statistique et de la Démographie (INSD). Dans chaque ZD, 35 ménages ont été sélectionnés. Les ménages ont été sélectionnés systématiquement de façon aléatoire à l'aide du logiciel « Random Number Generator ». Les ménages sélectionnés et qui ont des femmes éligibles, c'est-à-dire des femmes d'âge reproductif (15-49 ans), ont été contactés afin d'obtenir leur consentement ainsi que le consentement des femmes pour l'interview. La collecte de données s'est déroulée de décembre 2018 à janvier 2019 et a concerné au total 2763 ménages and 3388 femmes.

Échantillon de l'analyse

Les analyses de PMA2020 incluent uniquement les observations des entretiens terminés avec les ménages. L'échantillon de femmes inclut uniquement les entretiens terminés avec les femmes dans les ménages terminés. La majorité des indicateurs comprennent uniquement les femmes de facto (femmes ayant passé la nuit précédente dans le ménage). Toutes les observations, néanmoins, sont incluses dans la base de données pour permettre aux utilisateurs de calculer les taux de réponse.

Mises à jour de la base de données

Toute mise à jour apportée à une base de données après sa publication initiale sera documentée ici.

En janvier 2017, toutes les bases de données publiées précédemment ont été modifiées comme suit :

1. La valeur de **age_at_first_use_children** est 0 pour les femmes qui ont déjà utilisé la planification familiale au moins une fois et qui n'ont jamais accouché d'un enfant. Auparavant, ces femmes avaient une valeur manquante pour **age_at_first_use_children**.
2. Les valeurs de **water_sources_main_drinking** et **water_sources_main_other** sont égales à la valeur de **water_sources_all** si un ménage n'a qu'une seule source d'eau. Auparavant, ces ménages pouvaient avoir une valeur manquante pour ces variables.



3. La valeur de **sanitation_main** est égale à la valeur de sanitation_all si un ménage n'a qu'une installation sanitaire. Auparavant, ces ménages pouvaient avoir une valeur manquante pour cette variable.

Toutes les bases de données publiées après janvier 2017 comprendront ces changements.

Pour signaler des erreurs ou incohérences :

Veuillez écrire à datamanagement@pma2020.org