

## Note sur échantillonnage et les coefficients d'extrapolation de l'EDAM4

La principale base de données et de cartes pour la compilation de la base de sondage de l'EDAM4 est celle du Recensement général de la population et de l'habitat de Djibouti 2009 sur la population et le logement. La base de sondage des Zones de dénombrement (ZD) ou Unités primaires d'échantillonnage (UPE) du recensement est stratifiée par région, zone urbaine et zone rurale pour la plupart des enquêtes nationales. Bien que les zones urbaines et rurales de chaque région soient des strates d'échantillonnage individuelles pour améliorer l'efficacité de la conception de l'échantillon, le total des résultats des autres régions urbaines et rurales sera limité au niveau national.

Un modèle d'échantillonnage stratifié en deux étapes a été utilisé pour l'EDAM4. Au cours de la première étape, les échantillons de ZD ont été sélectionnés systématiquement avec des probabilités proportionnelles à la taille au sein de chaque strate (Tableau 1). Dix strates de sondages ont été définies : les cinq arrondissements de Djibouti ville et les cinq régions de l'intérieur du pays.

Avant de procéder au tirage des ménages échantillon à la deuxième phase, une mise à jour des listings des ménages des ZD échantillons a été effectuée en avril 2017. Au niveau de la deuxième phase, un échantillon aléatoire systématique de 12 ménages a été sélectionné dans chaque ZD ou échantillon ou grappe pour Djibouti ville et 15 ménages par grappe pour les autres régions. Dans le cas des échantillons du milieu rural, l'échantillon de 15 ménages devait comprendre les ménages sédentaires et nomades. Étant donné le grand pourcentage de ménages nomades dans les zones rurales selon les résultats publiés du recensement de 2009 (plus de 50%), cette procédure d'échantillonnage devait fournir un échantillon suffisant de ménages nomades pour l'analyse au niveau national. Au total, l'échantillon de l'EDAM4 comprend 376 ZD (251 en milieu urbain ; 125 en milieu rural). Le Tableau 1 présente la distribution de ZD de l'EDAM4.

**Tableau 1. Zones de dénombrement (ZD) sélectionnés dans l'échantillonnage**

Région	Total	Urbain		Rurale	
	Total ZD	Total ZD	ZD sélectionnés	Total ZD	ZD sélectionnées
Djibouti Ville					
Arrondissement 1	74	74	42	-	-
Arrondissement 2	120	120	42	-	-
Arrondissement 3	38	38	38	-	-
Arrondissement 4	156	156	42	-	-
Arrondissement 5	142	142	42	-	-
Région Ali Sabieh	46	27	16	19	18
Région Dikhil	64	20	9	44	25
Région Tadjourah	57	12	4	45	30
Région Obock	45	8	8	37	26
Région Arta	51	10	8	41	26
Total	793	607	251	186	125

Pour que les estimations d'échantillons à partir des données de l'EDAM4 de 2017 soient représentatives de la population, il faudra multiplier les données par un poids d'échantillonnage ou un facteur d'expansion. Le poids de base pour chaque ménage échantillon est égal à l'inverse de sa probabilité de sélection (calculée en multipliant les probabilités à chaque étape d'échantillonnage).

Sur la base de la conception de l'échantillon stratifié en deux étapes, la probabilité globale de sélection pour les ménages échantillon EDAM 2017 peut être exprimée comme suit :

$$p_{hi} = \frac{n_h M_{hi}}{M_h} \times \frac{m_{hi}}{M'_{hi}},$$

où:

$p_{hi}$ =	Probabilité d'échantillonnage globale pour les ménages sélectionnés pour EDAM4 dans le i-ème groupe d'échantillons dans la strate h
$n_h$ =	nombre de clusters d'échantillons sélectionnés dans la strate h pour l'EDAM4
$M_{hi}$ =	nombre total de ménages dans le i-ème groupe d'échantillons dans la strate h du cadre d'échantillonnage basé sur le recensement 2009 de Djibouti (mis à jour en partie)
$M_h$ =	nombre total de ménages dans la strate h de la base de sondage
$m_{hi}$ =	12 or 15 = nombre de ménages échantillons sélectionnés pour EDAM4 dans le i-ème groupe d'échantillons dans la strate h
$M'_{hi}$ =	nombre total de ménages dans la nouvelle liste pour le i-ème groupe d'échantillons dans la strate h

Le poids de base pour les ménages échantillon de l'EDAM4 est l'inverse de cette probabilité de sélection, exprimée comme suit :

$$W_{hi} = \frac{1}{p_{hi}} = \frac{M_h M'_{hi}}{n_h M_{hi} m_{hi}},$$

où:

$W_{hi}$ =	poids de base pour les ménages-échantillons EDAM4 dans le i-ème groupe d'échantillons dans la strate h
------------	--

Après la collecte des données EDAM4, il sera nécessaire d'ajuster les poids de base pour tenir compte des non-entrevues, comme suit :

$$W'_{hi} = W_{hi} \frac{m_{hi}}{m'_{hi}},$$

où:

$W'_{hi}$ =	poids ajusté pour les ménages-échantillons EDAM4 dans le i-ème groupe d'échantillons dans la strate h
$m'_{hi}$ =	nombre de ménages échantillonnés ayant terminé des entrevues dans le i-ème groupe d'échantillons dans la strate h.