



FICHE MÉTHODOLOGIQUE

Champ de l'enquête

Conduite conjointement par le Bureau d'études, de recherches, et de consulting international (Berci), Ebuteli et le Groupe d'Etudes sur le Congo (GEC) de l'Université de New York, ce sondage a été réalisé du 24 au 29 avril 2024 auprès de 2000 personnes âgées de 18 ans et plus. L'échantillon tiré était proportionnel au poids démographique de chacune des 26 provinces du pays. Au moins 200 sites d'enquêtes (communes urbaines, collectivités locales, villages ou territoires) ont constitué notre unité primaire de sondage (UPS), sélectionnés de manière aléatoire.

L'objectif de cette enquête nationale était de collecter des informations auprès des Congolais de tous les coins du pays afin de mieux comprendre le contexte sociopolitique et économique de la RDC dans une période post-électorale. Des enquêteurs formés en interne ont collecté les données à l'aide des interviews directes par téléphone (CATI) après avoir obtenu le consentement éclairé et informé des participants. Le logiciel de collecte des données électroniques mobiles ODK (Open Data Kit) a été utilisé et les données collectées ont été directement transférées dans un serveur en ligne hébergé par la plateforme Ona.io

Le questionnaire d'enquête comprenait 38 questions subdivisées en thèmes et sous-thèmes. Pour faciliter la compréhension et la bonne administration de l'enquête, le questionnaire a été traduit en langues nationales (kikongo, lingala, swahili et tshiluba), en plus du français. Les interviewés avaient le choix de répondre dans l'une de ces langues.

Méthode d'échantillonnage

Base de sondage

Ce sondage d'opinion a été réalisé en utilisant une base de données téléphoniques provenant des grandes enquêtes de l'Institut nationale de la statistique (INS) et de Berci/GEC, réalisées au domicile des personnes enquêtées. Cette base de données contient près de 400 000 numéros de téléphone, obtenus grâce à un échantillonnage stratifié, divisant la population en sous-groupes ou strates en fonction de certaines caractéristiques d'intérêts en utilisant

Technique d'échantillonnage

La méthode utilisée était celle de l'échantillonnage avec probabilité à taille (PPT) - un échantillon de probabilité stratifié en plusieurs étapes, représentatif au niveau national et provincial. Chaque strate est ensuite échantillonnée indépendamment pour créer un échantillon représentatif de la population en termes de caractéristiques d'intérêt telles que la province, le milieu, la ville ou le territoire.

L'objectif était de donner à chaque enquêté les mêmes chances d'être choisi ou non. Les chances d'atteindre cet objectif ont été améliorées en utilisant des méthodes de sélection aléatoire à chaque étape du processus d'échantillonnage, de manière indépendante à chaque strate.

Taille de l'échantillon

La taille initiale de l'échantillon était de 2000 personnes. Pour atteindre cette taille, 2420 entretiens ont été lancés, et 2041 personnes ont accepté de répondre, soit un taux de participation de 84 %. Au final, 1788 questionnaires ont été validés après le nettoyage des données. Les raisons des refus étaient principalement : la méfiance due au caractère politique de certaines questions, la peur d'être appréhendés par la suite, le manque d'habitude d'être interviewé, le manque de temps et d'intérêt.

La taille de l'échantillon avait été obtenu en utilisant la formule suivante : $n/N = nh/Nh$

Avec :

n : la taille totale de l'échantillon ;

N : la taille de la population (liste qu'on aura) ;

nh : la taille de l'échantillon dans la province ;

Nh : la taille de la province.

TABLE 1 : PRINCIPAL RÉSULTAT EN TERMES DE SITES ET DE RÉPONDANTS

UNITÉ D'ÉCHANTILLONNAGE PRIMAIRE ATTENDUE ET LANCÉE	NOMBRE (%)
<i>Nombre total de sites échantillonnés</i>	2420
<i>Nombre total de répondants observés avant le nettoyage du fichier</i>	2041
<i>Nombre total de répondants observés après le nettoyage du fichier</i>	1788
<i>Nombre (et %) des interviews éliminées après le nettoyage du fichier</i>	253 (10,5%)
<i>Nombre de refus (et %)</i>	379 (16%)

Sélection des répondants

Une seule personne par ménage ou numéro de téléphone contact sélectionné a été interrogée. Si la personne sélectionnée refusait de répondre, les intervieweurs pouvaient remplacer le ménage par suivant sur leur liste téléphonique. Avant de remplacer un ménage, les intervieweurs avaient été invités à faire plusieurs tentatives - au minimum trois tentatives - le même jour. Pour assurer une représentation égale entre les sexes, un quota de genre a été utilisé.

Pondération

Les pondérations ont été appliquées pour ajuster les probabilités différentielles de sélection, rééquilibrant ainsi l'échantillon afin qu'il corresponde à la distribution de la population sur les variables clés tels que l'emplacement géographique, le sexe et le milieu de résidence. Cette méthode a aidé à réduire l'erreur d'échantillonnage et à augmenter la précision des estimations.

Profil démographique comparatif de nos répondants

L'analyse comparative des résultats du sondage avec l'annuaire de l'INS 2020 et la Commission électorale nationale indépendante (Ceni) 2018 semble indiquer que les caractéristiques de nos

répondants concernant leur localisation géographique sont plus en phase avec la répartition géographique de la Ceni 2018, étant donné la temporalité de notre base de données et celle de l'enrôlement (cf. Tableau ci-dessous). Pour les variables sexe, milieu, âge et niveau d'études, seules les données de l'annuaire INS 2020 nous permettent d'avoir des données comparatives.

TABLE 2 : RÉPARTITION DE L'ÉCHANTILLON SELON CERTAINES CARACTÉRISTIQUES
SOCIO-DÉMOGRAPHIQUES

VARIABLES DÉMOGRAPHIQUES	INS 2020 (*) MICS 2019	ÉCHANTILLON NON PONDÉRÉ (N= 1788)
Milieu		
§ Urbain	40%	60%
§ Rural	60%	40%
Sexe		
§ Homme	49%	59%
§ Femme	51%	41%
Âge(*)		
§ 18-24 ans	27%	5%
§ 25-34 ans	29%	29%
§ 35-44 ans	20%	36%
§ 45-54 ans	12%	22%
§ 55-64 ans	7%	7%
§ Plus de 65 ans	10%	2%
Niveau d'Études (*)		
§ Sans/pré-primaire	11%	6%
§ Primaire	27%	4%
§ Secondaire	54%	48%
§ Universitaire	7%	40%

Traitement et analyse des données

Le nettoyage des données a été effectué après vérification des incohérences, une variable à la fois, pour corriger les écarts non détectés et effacer les doubles entrées. Au cours de ce processus, 253 (soit 10,5%) des données ont été éliminées, selon les critères suivants :

- Absence d'indication sur l'heure de début ou l'heure de fin de l'interview ;
- Absence d'indication sur la localisation géographique ou encore impossibilité de déterminer la localisation géographique exacte des répondants ;
- Durée de l'interview irréaliste et questionnaire incomplet.

L'analyse des données s'est faite avec le logiciel STATA 18. Les chercheurs du GEC et d'Ebuteli ont ensuite rédigé la note d'analyse.

Contrôle qualité

L'assurance de la qualité a été effectuée en temps réel au fur et à mesure que les sondages étaient téléchargés sur le serveur ONA. Des feedbacks ont été fournis aux équipes régulièrement. Les intervieweurs ont été suivis pour détecter toute précipitation et autres indicateurs de qualité, tels que les omissions et les incohérences dans les données. Près de 10 % des effectifs ont fait l'objet d'un contrôle a posteriori par nos superviseurs.

Travaux de terrain

L'équipe de terrain était composée de 75 intervieweurs expérimentés et de 10 superviseurs à Kinshasa. Avant de lancer le travail sur le terrain, une formation de deux jours avait été organisée entre l'équipe de recherche, les superviseurs, les intervieweurs et les contrôleurs. Cette formation incluait l'examen du questionnaire et de la méthodologie, les directives du code de déontologie professionnelle des enquêtes sur le terrain, l'utilisation des tablettes, et les instructions de remplissage du questionnaire. Une attention particulière a été accordée aux moyens de minimiser les non-réponses et de respecter la confidentialité des répondants, ainsi qu'à l'importance d'obtenir leur consentement informé. Cette dernière partie a été réalisée conjointement avec le GEC et en conformité avec les recommandations du comité d'éthique de l'Université de New York.

Défis rencontrés

- Écart important entre les projections du recensement de l'INS de 2019 et celles de 2020, indiquant des changements substantiels dans le pourcentage de la population dans les provinces de Kinshasa, Kwango, Kwilu, Kongo Central et, dans une moindre mesure, Équateur, Nord-Ubangi et Sud-Ubangi après le lancement de l'enquête ;
- Faible bande passante des connexions téléphoniques dans les sites ruraux, causant une surreprésentation du milieu urbain.

Précision d'intervalle ou marge d'erreur

La marge d'erreur varie au niveau des provinces en fonction de la taille de l'échantillon. Pour les analyses provinciales, il est recommandé d'utiliser la grille de 11 provinces plutôt que celle de 26 pour réduire la marge d'erreur.

Ex : Intervalle de confiance d'une proportion par la méthode de Wilson avec Correction de Continuité

On cherche à calculer l'intervalle de confiance d'une proportion p avec une probabilité de $(1-\alpha)$. On note p la proportion expérimentale observée, et n la taille de l'échantillon. La population-mère est supposée infinie.

TABLE 6: MARGES D'ERREUR ET INTERVALLE DE CONFIANCE À 95 %

P	Province	n	1-a	z	Borne inférieure	Borne supérieure
15,0 %	Bas-Uélé	11	95,0 %	1,960	2,1 %	49,1 %
15,0 %	Équateur	22	95,0 %	1,960	4,3 %	37,5 %
15,0 %	Haut Katanga	102	95,0 %	1,960	9,0 %	23,6 %
15,0 %	Haut Lomami	63	95,0 %	1,960	8,1 %	25,6 %
15,0 %	Haut-Uélé	91	95,0 %	1,960	8,7 %	24,2 %
15,0 %	Ituri	87	95,0 %	1,960	8,7 %	24,2 %
15,0 %	Kasaï	80	95,0 %	1,960	8,4 %	25,1 %
15,0 %	Kasaï Central	82	95,0 %	1,960	8,5 %	24,7 %
15,0 %	Kasaï Oriental	88	95,0 %	1,960	8,6 %	24,6 %
15,0 %	Kinshasa	256	95,0 %	1,960	11,0 %	20,0 %
15,0 %	Kongo Central	76	95,0 %	1,960	8,3 %	25,2 %
15,0 %	Kwango	47	95,0 %	1,960	6,8 %	28,9 %
15,0 %	Kwilu	109	95,0 %	1,960	9,2 %	23,3 %
15,0 %	Lomami	50	95,0 %	1,960	7,1 %	28,0 %
15,0 %	Lualaba	16	95,0 %	1,960	3,4 %	41,4 %
15,0 %	Maindombe	31	95,0 %	1,960	5,4 %	33,2 %
15,0 %	Maniema	37	95,0 %	1,960	6,4 %	30,1 %
15,0 %	Mongala	56	95,0 %	1,960	7,4 %	27,3 %
15,0 %	Nord Ubangi	29	95,0 %	1,960	5,1 %	34,4 %
15,0 %	Nord-Kivu	121	95,0 %	1,960	9,5 %	22,7 %
15,0 %	Sankuru	54	95,0 %	1,960	7,4 %	27,3 %
15,0 %	Sud Ubangi	50	95,0 %	1,960	9,0 %	23,6 %
15,0 %	Sud-Kivu	98	95,0 %	1,960	7,0 %	28,3 %
15,0 %	Tanganika	39	95,0 %	1,960	6,2 %	30,8 %
15,0 %	Tshopo	42	95,0 %	1,960	6,4 %	30,1 %
15,0 %	Tshuapa	51	95,0 %	1,960	7,3 %	27,5 %