

2016

REPUBLIQUE DU BURUNDI



MINISTRE A LA PRESIDENCE CHARGE DE
LA BONNE GOUVERNANCE ET DU PLAN

LIGNES DIRECTRICES SUR LA QUALITÉ DE PRODUCTION DES STATISTIQUES AU BURUNDI

Première édition



Devise : "En marche avec les Statistiques Fiables au Service du Développement"

INSTITUT DE STATISTIQUES ET
D'ÉTUDES ÉCONOMIQUES DU
BURUNDI

AVANT PROPOS

La qualité des données est au cœur des statistiques produites dans tout le Système Statistique National (SSN) au Burundi. En tant qu'organe central de coordination technique des activités du SSN du Burundi, l'Institut de Statistiques et d'Études Économiques du Burundi a l'obligation de veiller au respect de la qualité dans la production de l'information statistique au sein du SSN.

Afin d'effectuer correctement les activités de production des données statistiques pertinentes, fiables et crédibles, différents documents ont déjà été élaborés. Il s'agit notamment de la Stratégie Nationale de Développement de la Statistique, du Recueil des concepts, nomenclatures, classifications, définitions, indicateurs et méthodologies utilisés dans le SSN ainsi que certains textes d'application de la loi N°1/17 du 25 septembre 2007 portant organisation du Système Statistique au Burundi.

Le principal objectif du présent document est de fournir des lignes directrices permettant l'amélioration de la qualité des données produites dans le système statistique national.

La gestion de la qualité des données doit donc jouer un rôle central dans la gouvernance globale du SSN. Ce document traite donc les principales dimensions qualitatives des statistiques et est complété par un code de bonnes pratiques en ce qui concerne :

- L'environnement institutionnel du SSN ;
- Les procédures statistiques et
- Les résultats statistiques.

En outre, il rappelle le contexte de la démarche qualité que le SSN, en général, et l'ISTEEBU, en particulier, doivent suivre. Enfin, il présente le modèle retenu pour la gestion de la qualité, définit chacune de ses dimensions, et énonce les règles de son contrôle.

Cette première édition a été élaborée en concertation avec les producteurs et utilisateurs des données. Ce document devra être amélioré chaque fois que de besoin. Toutes les observations, critiques et suggestions destinées à son amélioration sont donc les bienvenues.

Le Gouvernement du Burundi, par le biais du Ministère à la Présidence Chargé de la Bonne Gouvernance et du Plan, tient à remercier tous ceux qui, de près ou de loin, ont contribué à l'élaboration de ce document, spécialement Statistique Canada et l'Observatoire Économique et Statistique de l'Afrique Subsaharienne (AFRISTAT).

**LE MINISTRE A LA PRESIDENCE CHARGE DE LA
BONNE GOUVERNANCE ET DU PLAN,**

Ir Serges NDAYIRAGLIE.-



TABLE DES MATIERES

AVANT PROPOS.....	2
TABLE DES MATIERES.....	3
LISTE DES ABREVIATIONS ET SIGLES UTILISES.....	4
PREMIERE PARTIE : APERCU GENERAL.....	5
1.1. Définition de la qualité de l'information statistique.....	5
1.2. Dimensions de la qualité des données statistiques.....	5
1.2.1. Revue des dimensions de la qualité adoptées ailleurs.....	5
1.2.2. Définition des dimensions de la qualité des données statistiques.....	5
1.3. Dimensions de la qualité adoptées par le Burundi.....	8
DEUXIEME PARTIE : GESTION DES PRODUITS STATISTIQUES.....	8
2.1. Gestion de la pertinence de l'information statistique.....	8
2.2. Gestion de l'exactitude et de la fiabilité.....	9
2.3. Gestion de la ponctualité et de l'actualité.....	10
2.4. Gestion de l'accessibilité et de la clarté.....	11
2.5. Gestion de l'intelligibilité (ou des métadonnées).....	12
2.6. Gestion de la comparabilité et de la cohérence.....	12
TROISIEME PARTIE : PROCESSUS DE PRODUCTION STATISTIQUE : LES ETAPES DES ENQUETES STATISTIQUES ET DES RECENSEMENTS.....	14
3.1. Détermination des besoins et formulation des objectifs.....	15
3.1.1. Objet et finalité.....	15
3.1.2. Principes d'approche et d'orientation.....	15
3.1.3. Lignes directrices.....	16
3.1.4. Indicateurs de qualité/Vérification de qualité.....	16
3.2. Conception.....	17
3.2.1. Objet et finalité.....	17
3.2.2. Principes d'approche et d'orientation.....	17
3.2.3. Lignes directrices.....	17
3.2.4. Indicateurs de qualité/Vérification de qualité.....	18
3.3. Construction, mis à l'essai et échantillonnage.....	18
3.3.1. Objet et finalité.....	18
3.3.2. Principes d'approche et d'orientation.....	19
3.3.3. Lignes directrices.....	20
3.3.4. Indicateurs de qualité/Vérification de qualité.....	23
3.4. Collecte des données.....	24
3.4.1. Objet et finalité.....	24
3.4.2. Principes d'approche et d'orientation.....	24
3.4.3. Lignes directrices.....	24
3.4.4. Indicateurs de qualité/Vérification de qualité.....	25
3.5. Traitement des données.....	25
3.5.1. Objet et finalité.....	25
3.5.2. Principes d'approche et d'orientation.....	26
3.5.3. Lignes directrices.....	27
3.5.4. Indicateurs de qualité/Vérification de qualité.....	29
3.6. Analyse.....	29
3.6.1. Objet et finalité.....	29
3.6.2. Principes d'approche et d'orientation.....	30
3.6.3. Lignes directrices.....	30
3.6.4. Indicateurs de qualité/Vérification de qualité.....	31

3.7. Diffusion	31
3.7.1. Objet et finalité.....	31
3.7.2. Principes d'approche et d'orientation.....	32
3.7.3. Lignes directrices	32
3.7.4. Indicateurs de qualité/Vérification de qualité	32
3.8. Evaluation	33
3.8.1. Objet et finalité.....	33
3.8.2. Principes d'approche et d'orientation.....	33
3.8.3. Lignes directrices	34
3.8.4. Indicateurs de qualité/Vérification de qualité	34
REFERENCES BIBLIOGRAPHIQUES	35

LISTE DES ABBREVIATIONS ET SIGLES UTILISES

AFRISTAT	: Observatoire Economique et Statistique d'Afrique Subsaharienne
CDD	: Classification Décimale de Dewey
CD-ROM	: Compact Disc Read Only Memory
CEE	: Communauté Economique Européenne
CNAQD	: Cadre National d'Assurance Qualité des Données
CNIS	: Conseil National de l'Information Statistique
CTIS	: Comité Technique de l'Information Statistique
EUROSTAT	: Office statistique de l'Union européenne
GPS	: Global Positioning System
ICSP	: Interagency Council on Statistical Policy
IMIS	: Integrated Management Information Systems
ISTEEBU	: Institut de Statistiques et d'Etudes Economiques du Burundi
MGPPS/GSBPM	: Modèle générique du processus de production statistique
NADA	: National Data Archive
OCDE	: Organisation de coopération et de Développement Economique
ONG	: Organisation Non Gouvernementale
ONU	: Organisation des Nations Unies
OSC	: Organisation de la Société Civile
PTF	: Partenaire Technique et Financier
SAS	: Statistical Analysis System
SPSS	: Statistical Package for Social Sciences
SSN	: Système Statistique National

PREMIERE PARTIE : APERCU GENERAL

Cette partie est consacrée aux définitions et aux dimensions liées à la qualité des données statistiques.

1.1. Définition de la qualité de l'information statistique

La qualité de l'information statistique est l'ensemble des propriétés et caractéristiques d'un produit statistique qui lui confèrent l'aptitude à satisfaire des besoins exprimés et implicites. En d'autres termes, il s'agit d'une adaptation des données à leur utilisation.

1.2. Dimensions de la qualité des données statistiques

1.2.1. Revue des dimensions de la qualité adoptées ailleurs

a) Cadre générique des Nations Unies sur la qualité

Les éléments du modèle générique des Cadres Nationaux d'Assurance de la Qualité des Nations unies pour la gestion des produits statistiques sont la pertinence, l'exactitude et la fiabilité, l'actualité et la ponctualité, l'accessibilité et clarté, la cohérence et la comparabilité ainsi que la gestion des métadonnées.

b) Dimensions de la qualité adoptées par certains institutions/pays

Les différentes dimensions de gestion de la qualité des produits statistiques varient selon les spécificités de chaque institution/pays et, dans certains cas, certaines dimensions sont couplées.

A titre indicatif, l'EUROSTAT distingue huit dimensions, à savoir : (i) la pertinence, (ii) la précision, (iii) la comparabilité, (iv) l'actualité et la ponctualité, (v) l'accessibilité et la clarté et (vi) la cohérence.

L'interagency Council on Statistical Policy (ICSP) des Etats Unis d'Amérique distingue, quant à elle, sept dimensions, à savoir : (i) la pertinence, (ii) l'actualité, (iii) l'accessibilité et la comparabilité, (iv) l'objectivité et (v) la transparence et la reproductivité.

Le Canada distingue six dimensions qui sont : (i) la pertinence, (ii) l'exactitude, (iii) l'accessibilité, (iv) l'actualité, (v) l'intelligibilité et (vi) la cohérence.

La charte africaine de la statistique recommande douze dimensions, au titre de la qualité : la pertinence, la pérennité, les sources de données, l'exactitude et la fiabilité, la continuité, la cohérence et la comparabilité, la ponctualité, l'actualité, les spécificités et enfin la sensibilisation¹.

1.2.2. Définition des dimensions de la qualité des données statistiques

a) Pertinence

La pertinence fait référence à l'adéquation de la production statistique aux besoins réels des utilisateurs. Pour être qualifiées de pertinentes, les données statistiques doivent éclairer les utilisateurs sur les questions les plus importantes à leurs yeux. Elles apportent une réponse à un problème particulier en comblant une demande institutionnelle ou sociale.

¹ The African Statistical Journal, volume 8, May 2009, Principe 2

L'évaluation de cet élément est subjective. Dans la pratique, elle se résume à la façon la plus optimale de concilier les deux difficultés ci-après :

- satisfaire les besoins actuels et potentiels des utilisateurs,
- élaborer un programme qui répond aux besoins dans un contexte de ressources limitées.

b) Exactitude et fiabilité

L'exactitude de l'information statistique exprime « la mesure dans laquelle l'information décrit correctement le phénomène qu'elle devrait évaluer ». Elle est évaluée par l'erreur dans les estimations statistiques (erreur systématique et erreur aléatoire). Elle peut aussi être définie par rapport aux sources principales d'erreurs susceptibles de mener à des données imprécises (par exemple : couverture, échantillonnage, non-réponse, réponse, les méthodes et les instruments de collecte et de mesure, le traitement, ...).

L'évaluation de l'exactitude fait donc référence au degré de rapprochement se situant entre la valeur finale retenue et la valeur réelle, souvent ignorée de la population à l'étude.

Quant à la fiabilité, elle sous-entend que les données de base sont raisonnablement conformes aux critères de définition, de champ d'application, de classification, d'évaluation et de date d'enregistrement prescrits pour l'établissement des statistiques.

c) Actualité et ponctualité

L'actualité et la ponctualité des données statistiques correspondent au délai entre le point de référence (ou la fin de la période de référence) auquel se rapporte l'information et la date à laquelle les données sont disponibles. La notion d'actualité est liée avec celles d'exactitude et de pertinence. L'évaluation de l'actualité est évidente et tient compte des normes de diffusion : écart entre la période de référence et la période de publication

Les statistiques nationales doivent être diffusées en temps utile et, dans toute la mesure du possible, selon un calendrier annoncé à l'avance, et doivent prendre en compte les événements courants et être d'actualité.

d) Accessibilité et clarté

La notion d'accessibilité renvoie à la facilité avec laquelle tout utilisateur peut obtenir l'information dont il a besoin auprès du producteur.

En plus, l'information doit être présentée sous une forme claire et compréhensible, diffusée d'une manière pratique et adaptée, disponible et accessible pour tous et en même temps. Les éléments suivants permettent de l'évaluer :

- l'existence de l'information,
- la forme ou le support d'accès à l'information,
- le coût de l'information pour les utilisateurs.

e) Cohérence et comparabilité

Par **cohérence** des données statistiques, on entend la mesure dans laquelle celles-ci peuvent être jumelées à d'autres renseignements statistiques dans un cadre analytique au fil du temps.

La **comparabilité** renvoie au caractère par lequel les données peuvent se prêter à la comparaison avec celles produites par d'autres sources.

Les statistiques nationales doivent présenter une cohérence interne dans le temps et permettre la comparaison entre les régions et avec d'autres pays. À cette fin, il doit être possible de combiner et d'utiliser conjointement des données connexes provenant de sources différentes. Les concepts, classifications, terminologies et méthodes établis et reconnus au niveau international, doivent être utilisés.

L'évaluation de ces deux éléments se réfère à l'uniformisation des normes et concepts et à l'utilisation de méthodologies communes au sein du SSN. Des efforts doivent être consentis pour l'utilisation de concepts normalisés et de méthodologies communes d'une enquête à l'autre.

f) Intelligibilité et interprétabilité

L'**intelligibilité et l'interprétabilité** des données statistiques renvoient à la disponibilité de renseignements supplémentaires et/ou métadonnées nécessaires à l'interprétation et à l'utilisation appropriée de ces données.

Ceci peut se vérifier par l'existence des métadonnées sur « les variables, les classifications et les concepts sous-jacents utilisés, la méthode de collecte des données et le traitement, etc. »

g) Pérennité

Par pérennité, on entend la conservation de la mémoire de la documentation des processus de production statistique.

Les statistiques officielles doivent être conservées sous une forme aussi détaillée que possible afin d'en garantir l'utilisation par les générations futures, tout en préservant les principes de confidentialité et de protection des répondants.

h) Continuité

Les autorités statistiques doivent garantir la continuité et la comparabilité dans le temps des informations statistiques dans le temps et dans l'espace.

i) Sources de données

Les sources de données statistiques utilisées à des fins statistiques sont : les recensements, les enquêtes statistiques et/ou les fichiers administratifs.

Les organismes responsables de la statistique doivent choisir leur source en tenant compte des besoins des utilisateurs, de la qualité des données qu'elle peut fournir, de leur actualité, et particulièrement, la charge qui pèse sur les répondants et les coûts y relatifs. Les données individuelles collectées pour l'établissement des statistiques par les services qui en ont la responsabilité, qu'elles concernent les personnes physiques ou morales doivent être strictement confidentielles et ne doivent être utilisées qu'à des fins statistiques.

j) Spécificité

Les méthodes de production et d'analyse de l'information statistique doivent tenir compte des spécificités de chaque institution/pays.

k) Sensibilisation

Les institutions/pays doivent sensibiliser le public, et en particulier, les fournisseurs des données statistiques sur l'importance de la statistique.

1.3. Dimensions de la qualité adoptées par le Burundi

Les dimensions de la qualité se chevauchent et se recoupent à telle enseigne que leur gestion oblige à les considérer toutes.

Le cadre National d'Assurance Qualité des Données (CNAQD) du Burundi a retenu dix dimensions à savoir : (i) la pertinence ; (ii) l'exactitude, (iii) la fiabilité ; (iv) la ponctualité ; (v) l'actualité ; (vi) l'accessibilité ; (vii) la clarté ; (viii) l'intelligibilité (métadonnées) ; (ix) la comparabilité ; (x) la cohérence.

DEUXIEME PARTIE : GESTION DES PRODUITS STATISTIQUES

Au Burundi, le Cadre d'assurance de la qualité des produits statistiques est l'ensemble des pratiques, des procédures et des mécanismes de gestion, d'exploitation et de consultation ainsi que d'autres bonnes pratiques utilisées afin de gérer la qualité des produits statistiques.

Il établit un lien entre les besoins des utilisateurs et les produits, et représente un outil de rétroaction, d'examen du rendement, ainsi que de planification et d'élaboration permanentes.

La gestion et le suivi de ce cadre incombent à l'ISTEEBU en tant qu'organe central du SSN.

Le cadre offre également des orientations aux gestionnaires des produits statistiques et à leurs équipes leur permettant de remplir leurs missions en tenant compte des situations conflictuelles et des contraintes liées à la gestion des différentes dimensions.

Les sections suivantes déclinent les lignes directrices pour assurer la gestion des dites dimensions ainsi que les indicateurs associés.

2.1. Gestion de la pertinence de l'information statistique

La pertinence de l'information statistique se confond à l'opportunité de l'information cherchée ou donnée. Une information statistique est pertinente quand elle répond aux besoins réels des utilisateurs. Quant à sa gestion, elle englobe tous les mécanismes qui sont mis en œuvre par l'organisme pour identifier et satisfaire les besoins des utilisateurs. Cette dimension doit aussi tenir compte des ressources humaines, matérielles et financières qui doivent être utilisées pour chaque programme.

L'opportunité et la gestion de l'information statistique permettent donc de traduire les programmes à mettre en œuvre en tenant compte des besoins des utilisateurs et des ressources disponibles. Les mécanismes ou processus qui servent à assurer la pertinence permettent aussi d'exercer un contrôle de base sur les autres aspects de la qualité et ainsi, d'évaluer les besoins des utilisateurs par rapport à ces autres dimensions.

Pour répondre efficacement aux multiples besoins des utilisateurs qui changent ou qui naissent avec le temps, il faut mettre en place un système régulier de communication pour déterminer leurs besoins et évaluer leur niveau de satisfaction. Ceci permettra de procéder à l'identification des besoins non satisfaits, d'effectuer un examen périodique de la pertinence des programmes réguliers et d'apporter les réajustements nécessaires tout en tenant compte des ressources disponibles.

a. Lignes directrices pour assurer la pertinence

- Enoncer et indiquer clairement les buts et les objectifs des programmes statistiques;
- S'assurer, au moment de la conception du questionnaire d'une opération statistique, que toutes les questions soient pertinentes et permettent d'atteindre les objectifs du programme ;
- Ne pas alourdir le fardeau des répondants et éviter les questions sensibles touchant la population ;
- Identifier les utilisateurs potentiels des produits des activités statistiques ;
- Maintenir une liaison bilatérale ou multilatérale avec les principaux utilisateurs ;

- Avoir de façon régulière des consultations avec le Groupe Sectoriel Statistique rassemblant les professionnels gouvernementaux de la Statistique, le Secteur Privé, la Société Civile, les Organisations Non Gouvernementales(ONG) et les Partenaires Techniques et Financiers(PTF);
- Avoir des consultations chaque fois que de besoin avec le comité technique de l'information statistique² dont les membres proviennent des divers secteurs ;
- Procéder à des évaluations annuelles de tous les programmes ;
- Confirmer et réajuster les priorités des programmes dans le cadre d'un processus de revue à mis parcours en examinant les besoins en information émergente et la pertinence des programmes courants ;
- Mesurer les utilisations des données et le niveau de satisfaction des utilisateurs.

b. Indicateurs/Liste de vérification pour mesurer la pertinence du produit

- % d'utilisateurs qui utilisent le produit statistique;
- % d'utilisateurs qui sont satisfaits du produit statistique ;
- Quelles sont les catégories socioprofessionnelles d'utilisateurs du produit statistique ?

2.2. Gestion de l'exactitude et de la fiabilité

L'exactitude et la fiabilité des données statistiques font référence à leur conformité à la réalité qu'elles doivent décrire. Cette dimension de la qualité garantit aux utilisateurs que les données statistiques fournies sont les plus exactes possibles. C'est-à-dire que ces données ont été produites, d'une part, de manière à minimiser les erreurs typiques (erreur d'échantillonnage et biais) qui affectent les données et d'autre part, de façon à maximiser la rigueur scientifique dans les interprétations, analyses et synthèses qui en sont faites. L'objectif est que ces travaux rendent compte de la réalité, avec le plus de justesse possible.

Après l'identification des programmes à mettre en œuvre, leurs objectifs généraux et les paramètres de ressources dans lesquels ils doivent être réalisés (gestion de la pertinence) vient la gestion de l'exactitude et de la fiabilité qui nécessite une attention particulière pendant les phases de conception, de mise en œuvre et de validation.

a. Lignes directrices pour assurer l'exactitude et la fiabilité

- Planifier et mettre en place des méthodes explicites et les processus d'assurance de la qualité intégrés pour détecter et contrôler les erreurs éventuelles décrites dans les diverses étapes du processus de production des statistiques;
- Minimiser et évaluer les erreurs de couverture et de classement des unités statistiques dans la base de sondage utilisée dans la sélection de l'échantillon ;
- Effectuer des pré-tests de toutes les versions des instruments des enquêtes (questionnaire, plateforme de saisie, programmes/logiciel d'imputation et d'estimation, etc.) jusqu'à la validation;
- Utiliser des échantillons probabilistes et sélectionner un plan de sondage qui permet d'optimiser l'exactitude des données produites pour les besoins de l'enquête;
- Vérifier l'échantillon finalement enquêté par rapport au plan de sondage initial ;
- Vérifier la représentativité et surtout les remplacements et les non répondants ;
- Minimiser et évaluer des erreurs de traitement (erreurs de saisie, erreurs de vérification et d'imputation des données, erreurs de manipulation des fichiers...);
- Vérifier le calcul des pondérations (spatiales notamment) ;
- Calculer les taux de non réponse;

² Décret N° 100/58 du 18 mars 2008 portant création, attributions, composition et fonctionnement du Conseil National de l'Information Statistique (CNIS)

- Identifier et décrire les raisons principales des non réponses (par exemple non contact, refus, inéligibles, d'autres raisons) et évaluer l'ampleur des erreurs dues à la non réponse;
- Pour chaque estimation produite, fournir une estimation de la variance qui tient compte du plan d'échantillonnage et des diverses manipulations effectuées sur les données d'enquête (ajustement de la non réponse, imputation, calibration...);
- Utiliser d'autres sources afin de comparer, évaluer et valider les estimations calculées;
- Utiliser les nouvelles technologies afin de minimiser les erreurs introduites par les manipulations manuelles et non cohérentes des données.

b. Indicateurs/Liste de vérification pour mesurer l'exactitude et la fiabilité

- Taux de couverture de la population cible;
- Taux de réponse;
- Estimation de la variance (coefficient de variation).

2.3. Gestion de la ponctualité et de l'actualité

La ponctualité et l'actualité de l'information statistique se rapportent à l'intervalle de temps entre le point de référence (ou la fin de la période de référence) et la date à laquelle l'information est publiée. En réalité, elles se trouvent en relation avec l'exactitude et influenceront sur la pertinence de cette information. Elles sont supposées être observées à toutes les étapes de la production et de diffusion.

Etant donné que la fixation des délais tient compte de l'exactitude et des coûts, il est probable de s'attendre à des améliorations des délais de diffusion au fur et à mesure que les nouvelles technologies sont utilisées.

a. Lignes directrices pour assurer la ponctualité et l'actualité

- Bien planifier les étapes de production statistique et avoir des mécanismes de suivi réguliers de leur mise en œuvre ;
- Annoncer au préalable le calendrier des publications tout comme les changements des dates de diffusion prévues à plus long terme ;
- Adopter la stratégie de diffusion des données préliminaires, suivies de chiffres révisés et définitifs pour les rendre plus actuels ;
- Utiliser les nouvelles technologies de collecte et de traitement pour raccourcir les délais de diffusion ;
- Informer les utilisateurs au sujet des contraintes associées à l'actualité ;
- Expliquer et fixer à l'avance une nouvelle date de parution pour tout décalage par rapport au calendrier prévu ;
- Une publication mensuelle doit être publiée au plus tard le 10 du mois suivant³ ;
- Une publication trimestrielle doit être publiée au plus tard le 15^{ème} jour du deuxième mois du trimestre suivant ;
- Une publication semestrielle doit être publiée au plus tard le 30^{ème} jour du troisième mois du semestre suivant ;
- Une publication annuelle doit être publiée au plus tard à la fin de l'année suivante ;
- Les comptes économiques provisoires d'une année N doivent être disponibles à la fin de l'année N+1 ;
- Les comptes économiques définitifs d'une année N doivent être disponibles à la fin de l'année N+2.

³ Décret N° 100/227 du 8 octobre 2014 portant Cadre Nationale d'Assurance Qualité des Données (CNAQD) au Burundi.

b. Indicateurs/Liste de vérification pour mesurer la ponctualité et l'actualité

- Existence d'un calendrier de publication ;
- Degré de ponctualité selon le calendrier de publication arrêté, nombre de jours, de mois ou d'années de retard selon la publication ;
- Période entre la publication des résultats préliminaires et la publication de la version définitive.

2.4. Gestion de l'accessibilité et de la clarté

L'accessibilité de l'information fait référence à la facilité avec laquelle les utilisateurs peuvent connaître l'existence de cette information, sa localisation et son importation dans leur propre environnement de travail. En ce qui concerne la clarté, il faut noter que l'information doit être présentée sous une forme claire et compréhensible, accompagnée de métadonnées et d'explications.

Par exemple, les principaux instruments de diffusion de données utilisés actuellement par l'ISTEEBU sont les documents sur papier, CD-ROM, site web de l'ISTEEBU (www.isteebu.bi), les bases de données BurundInfo, IMIS et NADA, portail de données (africadata).

a. Lignes directrices pour assurer l'accessibilité et la clarté

- Mettre en place un système de documentation et d'archivage des données statistiques dans tous les services constituant le SSN ;
- Examiner les publications archivées périodiquement pour vérifier l'accessibilité aux utilisateurs. Les producteurs des données doivent veiller à ce que les documents statistiques produits soient bien intégrés aux systèmes de catalogage du SSN⁴ ;
- Établir des procédures dès le début du cycle de vie des enquêtes à caractère répétitif afin de s'assurer que tous les fichiers importants sont conservés ;
- S'assurer que lors de la publication et de la transmission à des tiers des résultats statistiques agrégés des opérations de collecte et de traitement qu'aucune identification directe ou indirecte des personnes physiques et morales concernées n'est possible ;
- Produire et mettre en œuvre des procédures permettant l'accès aux microdonnées anonymisées pour des fins statistiques ;
- Mettre en place des mécanismes de centralisation et d'accès aux macrodonnées anonymisées pour des fins statistiques ;
- Informer les utilisateurs sur les publications disponibles, leur fréquence de publication ainsi que les principaux modes de diffusion utilisés ;
- Organiser des séminaires de diffusion des résultats des enquêtes ou d'autres rapports de résultats statistiques disponibles pour les utilisateurs potentiels ;
- Rédiger et diffuser des communiqués de presse clairs et concis qui renvoient au produit final et mentionner un correspondant susceptible de répondre à des demandes de clarification ;
- Comme pour la pertinence, effectuer des recherches sur le niveau de satisfaction des utilisateurs concernant la facilité d'accès et les modes de diffusion des produits du SSN ;

b. Indicateurs/Liste de vérification pour mesurer l'accessibilité et la clarté

- Est-ce que le produit statistique est archivé ?
- Est-ce que le produit statistique est disponible sous plusieurs formats ?
- Est-ce que le produit statistique est accessible sous plusieurs formats ?
- Est-ce qu'il existe un référent du produit statistique au sein du service producteur ? ;
- Est-ce qu'un séminaire/atelier de diffusion du produit statistique a été organisé ? ;
- Est-ce qu'un communiqué de presse sur le produit a été diffusé ?
- Est-ce qu'il existe une métadonnée pour chaque produit ?

⁴ Classification Décimale de Dewey (CDD)

2.5. Gestion de l'intelligibilité (ou des métadonnées)

L'intelligibilité de l'information statistique reflète la disponibilité de l'information et des métadonnées supplémentaires nécessaires à l'interprétation et à l'utilisation appropriées des renseignements. Normalement, cette caractéristique prend en compte les variables, les classifications et les concepts sous-jacents utilisés, le plan d'échantillonnage, la méthode de collecte et le traitement des données, ainsi que les indications ou les mesures de l'exactitude de l'information statistique.

Une des responsabilités des organismes du SSN consiste à offrir suffisamment des informations pour permettre aux utilisateurs d'interpréter efficacement les renseignements statistiques. La gestion de l'intelligibilité porte principalement sur la communication de métadonnées ou de l'information sur l'information.

L'information nécessaire à la compréhension des données statistiques se divise en trois catégories générales :

- les concepts, les variables et les classifications qui sous-tendent les données;
- la méthodologie utilisée pour recueillir et compiler les données;
- les indicateurs de précision des données.

a. Lignes directrices pour assurer l'intelligibilité (ou des métadonnées)

- Actualiser régulièrement le recueil des concepts, définitions, nomenclatures et méthodologies utilisées dans le SSN du Burundi en fonction des innovations statistiques en se référant aux documents internationaux ;
- Rendre disponible les informations sur les usages des données et les objectifs qu'elles sont censées satisfaire ;
- Pour chaque produit statistique :
 - Inclure les définitions, les concepts de même que le questionnaire de collecte utilisé;
 - Expliquer la méthodologie utilisée pour recueillir l'information nécessaire à la diffusion du produit statistique ;
 - Mettre l'accent, pour les enquêtes et recensements sur la population cible, la base de sondage, la couverture, la période de référence, le plan d'échantillonnage, la taille de l'échantillon de même que sa méthode de sélection, la méthode de collecte et les étapes de traitement des données collectées;
 - Fournir des renseignements quant à la qualité des données : erreurs dues ou non à l'échantillonnage, taux de réponse, variance, imputation, estimation et vérification.

b. Indicateurs/Liste de vérification pour mesurer l'intelligibilité (métadonnées)

- Est-ce que la fiche méthodologique décrivant l'opération statistique est mise à la disposition des utilisateurs?
- Est-ce qu'un rapport d'évaluation du processus statistique est disponible à l'interne ?

2.6. Gestion de la comparabilité et de la cohérence

Les données observées peuvent être comparées à d'autres ensembles de données décrivant le même phénomène dans le temps et dans l'espace. Les organismes du SSN doivent ainsi s'assurer autant que possible, que les mêmes données statistiques de différentes sources et périodes ou de différents endroits géographiques soient comparables.

La cohérence reflète la mesure dans laquelle on peut réussir à regrouper une information avec d'autres renseignements statistiques dans un cadre analytique général et au fil du temps.

Par cohérence des données statistiques, on entend la concordance entre les différentes données élémentaires rattachées à un même point dans le temps et dans l'espace. L'utilisation de classifications, de concepts normalisés favorise la cohérence, tout comme l'utilisation d'une méthode commune d'une enquête à l'autre.

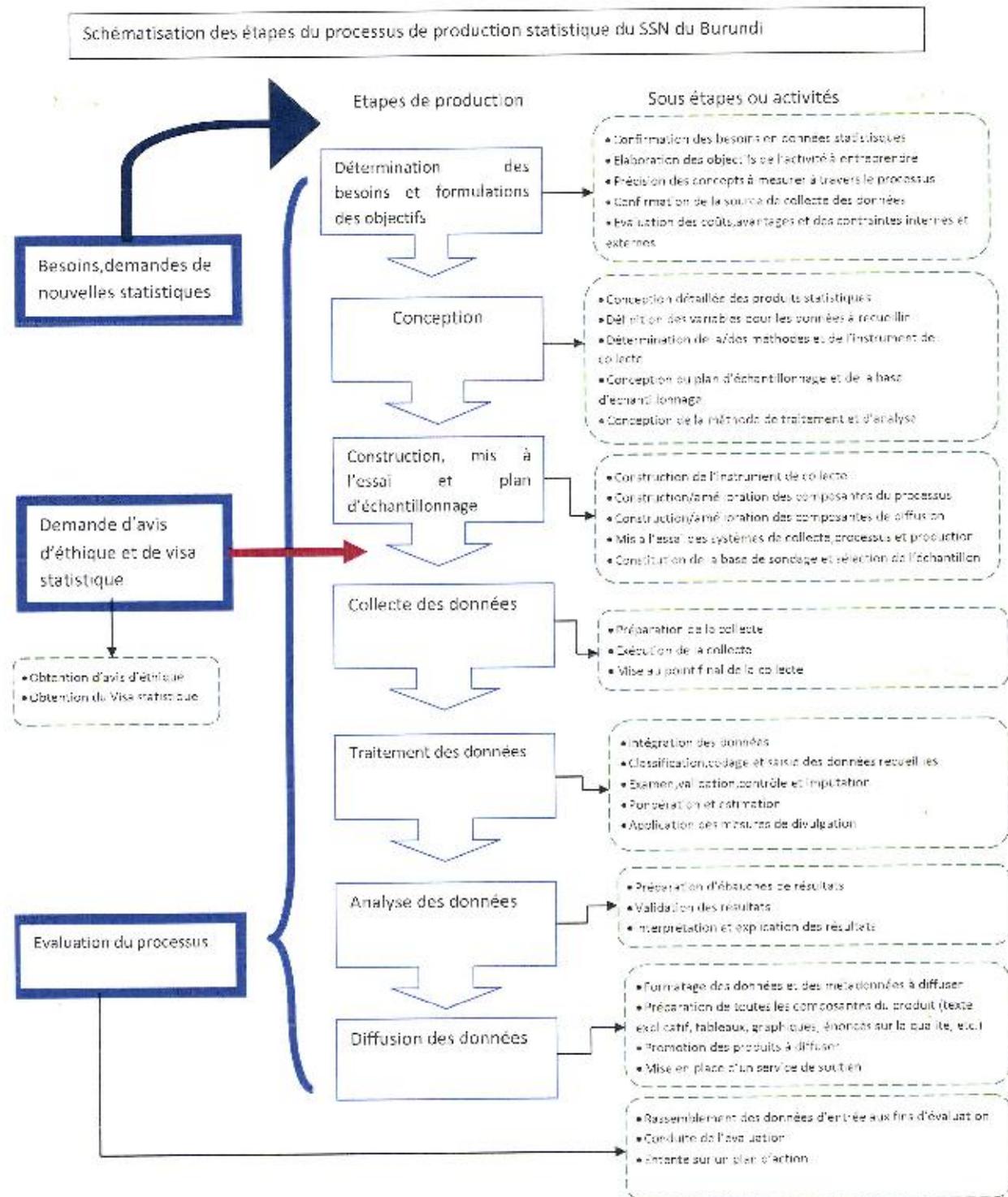
a. Lignes directrices pour assurer la comparabilité et la cohérence

- Maintenir à jour des documents des méthodologies et canevas de collecte des données administratives et celles provenant des opérations de collecte;
- Documenter l'impact de l'introduction de nouvelles classifications et de nouveaux concepts sur les séries statistiques ;
- Tous les producteurs de statistiques officielles du Burundi doivent se conformer aux normes établies. Sinon, ils doivent documenter clairement les causes de leur non utilisation et l'impact sur la comparabilité et la cohérence du produit ;
- Clarifier les différences, en cas de besoin, entre les concepts et les méthodes utilisés à différentes périodes pour les mêmes variables;
- Evaluer et documenter les effets dus aux différences entre les méthodes et les outils utilisés à des périodes différentes ;
- Mesurer les différences entre les pratiques locales et les pratiques internationales si elles existent.

b. Indicateurs/Liste de vérification pour mesurer la comparabilité et la cohérence

- Est-ce que les normes, concepts, définitions et classifications utilisés sont conformes au Recueil des concepts, définitions, nomenclatures et méthodologies adoptés par le SSN ?;
- Est-ce que vous appliquez le code de bonnes pratiques nationales en matière de méthodes statistiques ?
- Est-ce que vous avez souscrit aux codes de bonnes pratiques régionales et internationales en matière de méthodes statistiques ?
- Est ce que vous avez souscrit aux codes de bonnes pratiques internationales en matière de méthodes statistiques ?

TROISIEME PARTIE : PROCESSUS DE PRODUCTION STATISTIQUE : LES ETAPES DES ENQUETES STATISTIQUES ET DES RECENSEMENTS



3.1. Détermination des besoins et formulation des objectifs

3.1.1. Objet et finalité

«La détermination des besoins et formulation des objectifs engagent le commanditaire, les utilisateurs et l'organisme statistique chargé de l'opération. Les besoins sont déterminés à deux niveaux. D'une part, un commanditaire considéré comme utilisateur d'une enquête peut ressentir un besoin de données et d'autre part, l'organisme statistique, compte tenu de sa mission, mène des investigations statistiques pour mettre à disposition un ensemble de données indispensables pour les décideurs»³.

L'approbation de produire des nouvelles statistiques ou de réorienter un programme de production existante, doit répondre à un besoin d'information. Ainsi, la formulation d'objectifs, première étape de planification d'une activité statistique, est toujours dictée en fonction de la nature des informations requises. Leurs énoncés doivent être clairement précisés afin de permettre de mieux orienter les autres étapes subséquentes de l'activité statistique. Une révision des objectifs peut être faite à tout moment de l'élaboration de l'activité et autant de fois que cela sera nécessaire. Les besoins en information tels qu'énoncés dans les objectifs doivent aussi justifier le fardeau de réponse qu'ils engendreront.

Enfin, il est très important de s'assurer que la pertinence des résultats attendus par l'activité statistique pour la collectivité d'utilisateurs ciblée soit clairement établie. Les utilisateurs étant le gouvernement burundais incluant les services statistiques centraux et sectoriels, les partenaires techniques et financiers, le secteur privé, la société civile, les universités et instituts de recherches, les autres organismes et/ou groupes et les personnes qui utilisent l'information produite.

À la fin de cette étape du processus, l'ISTEEBU ou tout autre producteur de statistiques du SSN, devrait être en mesure de :

- décrire, dans le cas de statistiques actuellement produites, le processus de production et tous les renseignements sur la façon dont elles sont produites en soulignant toute inefficacité ainsi que les questions à régler;
- proposer, dans le cas de nouvelles statistiques, la solution «à venir» en détaillant la façon dont le processus de production sera conduit afin de produire ces statistiques;
- évaluer les coûts et les avantages ainsi que toutes contraintes internes et externes.

Les résultats de cette étape sont : liste des besoins en données statistiques, liste des objectifs arrêtés,, les concepts à utiliser, la source des données identifiée et enfin coûts, avantages et contraintes internes et externes évalués.

3.1.2. Principes d'approche et d'orientation

L'établissement des objectifs et la précision des utilisations de toute nouvelle activité statistique ou de la réorientation d'un programme existant, doivent être réalisés en étroite consultation avec les utilisateurs de l'information à recueillir et tous les partenaires participant activement au projet. Leur implication peut permettre d'une part, d'aider à bien décrire tout le contenu de l'activité statistique et d'autre part à déterminer et à anticiper certains problèmes conceptuels, méthodologiques et opérationnels et en proposer des solutions adéquates.

³ Tiré du Modèle générique du processus de production statistique (MGPPS/GSEPPM), publié par le groupe de travail conjoint de la CEE-OMU, d'Eurostat et de l'OCDE sur les métadonnées statistiques (METIS) en avril 2009

Par ailleurs, l'énoncé des objectifs doit se faire après une fine analyse et une bonne compréhension des besoins en données exprimés par les partenaires/utilisateurs. Cet énoncé doit être clair et précis tout en comportant les hypothèses à mettre à l'épreuve, la qualité prévue, les contraintes en ressources budgétaires et humaines ainsi que les dates prévues pour la publication des données.

Enfin, il est important pour tous les producteurs de données du SSN de s'assurer, avant de se lancer dans la production de toute nouvelle activité statistique, que les sources de données actuelles ne peuvent pas répondre aux besoins des utilisateurs et les conditions dans lesquelles elles sont disponibles, y compris toutes restrictions concernant leur utilisation.

3.1.3. Lignes directrices

Les lignes directrices déclinées ci-après doivent être suivies par toutes les composantes du SSN du Burundi en vue de déterminer les besoins en information statistique et la formulation des objectifs :

- Analyser les besoins des utilisateurs sur la recherche des solutions les plus rationnelles à court et à long terme ;
- Vérifier si des sources de données actuelles peuvent répondre aux besoins des utilisateurs ;
- Établir et entretenir des relations avec les utilisateurs des données dans tous les secteurs (public, privé, PTF, ONG, OSC, Chercheurs et tout autre utilisateur en général) afin de s'accorder sur le contenu pour rehausser la pertinence de l'information produite et d'améliorer la diffusion des produits et des services ;
- Susciter un appui des partenaires à l'égard du projet ;
- Évaluer le coût d'opportunité d'une nouvelle activité de production statistique par rapport aux statistiques disponibles pouvant répondre aux besoins des utilisateurs ;
- Établir les objectifs de l'enquête conjointement avec les utilisateurs et les principaux intervenants ;
- Déterminer dans quelle mesure une nouvelle activité de production statistique répond aux besoins des utilisateurs ;
- Trouver un équilibre raisonnable entre les besoins et les questions liées au fardeau de réponse et à la protection des renseignements personnels des répondants ;
- Reformuler les objectifs des activités statistiques périodiques en tenant compte des nouvelles exigences ;
- Fournir des précisions sur les concepts devant être mesurées à travers l'activité statistique ;
- Intégrer à l'énoncé des objectifs de l'activité statistique les aspects mesurables de la qualité qui doivent être satisfaits ;
- En déterminant les objectifs et les résultats d'une activité statistique précise, tenir compte des objectifs et des besoins des activités statistiques subséquentes ou parallèles.

3.1.4. Indicateurs de qualité/Vérification de qualité

Dimension de la qualité : Pertinence

Indicateurs sur le processus /Liste de vérification:

- Les utilisateurs courants et potentiels des produits ont-ils été identifiés? (*Oui, Partiellement, Non; Sinon ou partiellement pourquoi*);
- Des consultations formelles ont-elles eu lieu afin de définir les besoins en statistiques? (*Oui, Non; Sinon pourquoi*);
- Les coûts, les avantages et les contraintes internes et externes ont-ils été évalués et mesurés? (*Oui, Partiellement, Non; Sinon ou partiellement pourquoi*).

3.2. Conception

3.2.1. Objet et finalité

La conception est la détermination des produits statistiques, des concepts, des méthodologies, des instruments de collecte et des étapes du processus opérationnel de production statistique. Les produits de la conception sont inspirés des normes nationales et internationales. Ils peuvent être réutilisés ou réadaptés lors des opérations ultérieures. Lors de la conception, il est essentiel aussi de préciser les métadonnées pertinentes, les procédures d'assurance qualité ainsi que toute entente formelle se rapportant à la fourniture de données, comme les fiches de consentement, la règle de confidentialité et la confirmation du fondement juridique de la collecte des données.

A la fin de cette étape, les services et organismes statistiques doivent s'assurer que toutes les étapes du processus de production sont exhaustivement déterminées de manière à garantir leur enchaînement efficace et l'absence de lacunes ou de redondances dans les données à recueillir, la détermination de la/des méthodes et l'instrument.

Les résultats de cette étape sont : liste détaillée des produits statistiques, liste des variables des données à recueillir, les méthodes et l'instrument de collecte arrêtés, le plan d'échantillonnage et base de sondage, la méthode de traitement et d'analyse ; bref tous les systèmes de production statistique sont fixés. En outre, il est important de se pencher sur la disponibilité du personnel, la répartition des responsabilités et le calendrier.

3.2.2. Principes d'approche et d'orientation

Les principaux objectifs des organismes du SSN en tant que producteurs de données statistiques sont de s'assurer que l'information produite dresse un portrait logique et cohérent de l'économie, de la société et de l'environnement du Burundi. Cet ensemble des données doit pouvoir, par ailleurs, être mis en cohérence avec d'autres sources.

Afin de tirer des conclusions à partir d'un ensemble de données, il est extrêmement important que les utilisateurs aient une connaissance des définitions, concepts, méthodologies, classifications, nomenclatures et normes qui sous-tendent les données.

Seule l'utilisation des définitions harmonisées, de cadres conceptuels adaptés, de variables et de classifications au niveau international permet d'atteindre cesdits objectifs. Cependant, il est tout aussi nécessaire de modifier certaines définitions dans le but de mieux les adapter au contexte national pour pouvoir respecter les utilisations prévues des données.

3.2.3. Lignes directrices

Les lignes directrices déclinées ci-après doivent être suivies par toutes les composantes du SSN du Burundi en vue de procéder à la conception d'une opération statistique:

- Définir clairement les concepts et les variables et indiquer leur utilisation prévue ;
- Utiliser les concepts, définitions, classifications, unités statistiques et populations en se référant aux normes établies dans le Recueil des Concepts, Définitions et Nomenclatures du SSN du Burundi pour permettre la comparaison des données recueillies à partir de différentes sources et pour intégrer les données d'une source à l'autre ;
- Définir les nouveaux concepts différents en fonction des nomenclatures et classifications internationales normalisées produites par AFRISTAT, le Bureau de statistique des Nations Unies, le Bureau international du Travail, et d'autres organismes régionaux et internationaux ;

- Inscrire les nouveaux concepts dans le Recueil des Concepts, Définitions et Nomenclatures du SSN du Burundi ;
- Coder les microdonnées et conserver les fichiers au plus faible niveau de classification appropriée possible pour optimiser la souplesse d'utilisation ;
- Agréger les données à un niveau plus élevé pour répondre aux besoins particuliers de l'analyse ou pour respecter les contraintes de confidentialité ou de fiabilité des données ;
- Utiliser des unités d'observation normalisées pour faciliter la comparaison des données ;
- Les classifications sont généralement conçues à l'aide d'unités d'observation particulières comme les unités locales de mesure. Ainsi, l'ISTEEBU doit mettre à la disposition du SSN les équivalences documentées avec les unités normalisées ;
- Définir et déterminer les instruments et la ou les méthodes de collecte les plus appropriés à l'activité statistique ;
- Définir la population cible et la population d'enquête ;
- Établir le plan d'échantillonnage et déterminer la méthode et les critères de sélection d'échantillonnage les plus appropriés à l'activité statistique ;
- Déterminer et concevoir la méthode de traitement statistique et d'analyse des données à recueillir. Il s'agit de spécifier les routines de codages, d'imputation, d'estimation, de validation et de finalisation de l'ensemble des résultats ;
- Concevoir et établir le déroulement des opérations de l'activité statistique, de la collecte à la diffusion des données.

3.2.4. Indicateurs de qualité/Vérification de qualité

Dimension de la qualité : exactitude, intelligibilité, ponctualité.

Indicateurs sur le processus /Liste de vérification:

- Est-ce que la méthodologie permet d'atteindre les objectifs de l'activité statistique? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les concepts, les classifications et les unités statistiques en usage correspondent aux concepts normalisés utilisés au niveau national, régional et international? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les ressources financières, humaines ainsi que les outils de collecte utilisées permettent de respecter les délais? (*Oui, Non; Sinon pourquoi*);
- Est-ce que la base de sondage est exhaustive et à jour? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le tirage a été effectué selon la méthode la plus appropriée? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le calcul de la taille de l'échantillon tient compte des éléments pertinents (la précision, le taux de réponse, le coût et le plan d'échantillonnage,...)? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les plans de tabulation et d'analyse présentent une cohérence avec les variables du questionnaire qui sera utilisé? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le calendrier des activités a été établi? (*Oui, Non; Sinon pourquoi*);

3.3. Construction, mis à l'essai et échantillonnage

3.3.1. Objet et finalité

Cette phase permet de construire et de mettre à l'essai les outils de production (documents techniques, les programmes informatiques, les systèmes de diffusion) jusqu'à ce qu'ils soient prêts à être utilisés en milieu «réel» ainsi que la constitution de la base de sondage et la sélection de l'échantillon.

Le choix de l'échantillon doit se faire à partir d'une base de sondage. Celle-ci est une liste, un document ou un dispositif qui délimite et identifie les éléments de la population observée et qui en permet l'accès.

Les produits de la phase de conception orientent la sélection de processus, d'instruments, d'information et de services réutilisables, qui sont assemblés et configurés à cette phase pour créer l'environnement opérationnel. Les nouveaux services sont créés pour combler des lacunes dans le catalogue de services existants provenant de l'organisme producteur. Ces nouveaux services sont construits de manière à pouvoir être réutilisés au sein de l'architecture commune de la production de statistiques.

Les résultats de cette étape sont: les questionnaires, les différents manuels, les programmes de traitement et d'analyse ; les outils de dissémination ou diffusion et l'échantillon.

A la fin de ce processus, les outils de collecte et les programmes informatiques sont révisés selon les résultats des tests. Le personnel de l'opération est ensuite formé pour une mise à niveau afin de rentabiliser les leçons tirées de l'exercice précédent.

3.3.2. Principes d'approche et d'orientation

La construction d'un questionnaire doit tenir compte des éléments suivants :(i) les objectifs de l'enquête, (ii) la longueur et la formulation des questions, (iii) les caractéristiques des répondants, (iv) le fardeau du répondant, (v) les instructions sur les questions complexes ; (vi) les notes sur la confidentialité et le consentement.

L'échantillonnage probabiliste sous-tend une sélection aléatoire représentative d'unités de la base de sondage. Il est utilisé pour choisir un échantillon aléatoire de la population observée. L'objectif consiste à recueillir de l'information utile à partir des unités échantillonnées afin de permettre les inférences liées à la population observée.

Au niveau des programmes informatiques, le schéma logique du système devra également faire apparaître les principales opérations à effectuer sur les fichiers, comme le codage des données, l'entrée, la vérification et la validation des données, la sélection, la projection, le tri et le rapprochement des fichiers, la dérivation de nouvelles variables, l'agrégation, la tabulation et la présentation graphique.

Enfin, il faut configurer le déroulement des opérations, les systèmes et les transformations utilisés au sein des processus de production statistique, allant de la collecte des données à la diffusion.

Il importe de constituer une documentation claire et complète (c'est-à-dire une série de métadonnées) aussi bien pour la mise au point que pour la maintenance des systèmes de traitement des données. Il faut par conséquent documenter les fichiers et les différentes opérations réalisées de sorte que les personnes qui n'ont pas pris part à la mise en œuvre du système originel puissent également les utiliser.

La diffusion est l'objectif final de toute opération, elle présente de manière attractive et compréhensible un rapport final comprenant les éléments clés de l'étude. Elle prend les formes les plus variées en mettant à profit les nouvelles Technologies de l'information et de communication.

Encadré : Demande d'avis d'éthique et de visa statistique

En application du Décret N°100/261 du 31 octobre 2013 portant institution du Visa Statistique et de l'Avis d'éthique pour les enquêtes statistiques et recherches biomédicales et comportementales au Burundi, tout promoteur d'une opération d'enquête statistique à visée officielle requiert l'obtention du visa statistique auprès du Ministre ayant les statistiques dans ses attributions. En plus, toute enquête ou recherche biomédicale ou comportementale doit obligatoirement obtenir un avis d'éthique délivré par le Comité National d'Éthique pour la protection des êtres humains participant à la recherche biomédicale ou comportementale.

Ces visa et avis doivent être obtenus avant le début des travaux de terrain ou de recherche.

3.3.3. Lignes directrices

Les lignes directrices déclinées ci-après doivent être suivies par toutes les composantes du SSN du Burundi en vue de:

➤ Procéder à la construction du questionnaire

- Utiliser des mots et des concepts qui ont la même signification pour les répondants et les concepteurs du questionnaire et, dans le cas des entreprises, choisir des questions, des périodes de référence et des catégories de réponse qui sont compatibles avec les pratiques de tenue des livres de l'établissement ;
- Dans la mesure du possible, harmoniser les concepts et la terminologie avec ceux qui sont déjà utilisés ;
- Si certains concepts ont déjà été mesurés adéquatement par d'autres enquêtes, utiliser les questions déjà élaborées afin d'assurer la cohérence et la comparabilité entre les résultats d'enquête ;
- Utiliser une formulation et une terminologie qui encouragent les répondants à répondre aux questions le plus exactement possible. Dans cet esprit, le questionnaire doit se concentrer sur le sujet de l'enquête, être le plus bref possible, comporter des questions qui suivent un ordre logique afin de faciliter le rappel des répondants et diriger ces derniers vers la source d'information appropriée ;
- Tenir compte du fardeau du répondant dans la conception du questionnaire (que ce soit au niveau du contenu du questionnaire, de la durée et de l'inclusion des questions délicates) ;
- Établir un lien direct entre le questionnaire et le système de métadonnées statistiques, de manière à faciliter la saisie des métadonnées durant la phase de collecte ;
- Sur la page couverture du questionnaire :
 - indiquer le titre ou le sujet de l'enquête ;
 - identifier le commanditaire ;
 - expliquer l'objectif de l'enquête ;
 - demander la collaboration des répondants ;
 - Indiquer également l'autorité sous laquelle l'enquête est administrée et les mesures de protection de la confidentialité ;
- Les questions préliminaires d'ouverture doivent s'appliquer à tous les répondants, être faciles à comprendre et intéressantes à répondre, et établir le fait que le répondant fait partie de la population observée ;
- S'assurer que la communication de l'information aux répondants est clairement indiquée sur le questionnaire et expliquer la raison pour laquelle il est important de remplir le questionnaire ainsi que la façon dont les données de l'enquête seront utilisées ;

- A l'intérieur du questionnaire :
 - donner des titres ou des entêtes à chaque section du questionnaire ;
 - intégrer des directives et des espaces de réponse qui facilitent la réponse exacte aux questions ;
 - utiliser le gras, des ombrages, des illustrations et des symboles pour attirer l'attention des répondants ou des enquêteurs et pour les guider dans les parties du questionnaire qui doivent être lues et pour indiquer où doivent figurer les réponses ;
 - laisser un espace à la fin du questionnaire pour les commentaires additionnels des répondants et intégrer une formule d'appréciation destinée aux répondants.
- Concevoir des questionnaires professionnels, attrayants et faciles à administrer ou à remplir soi-même par l'enquêté. Si le questionnaire doit être administré par un enquêteur, il doit être convivial pour ce dernier.
- Pour réduire au minimum la possibilité d'erreurs de déclaration, s'assurer que les instructions à l'intention des répondants ou des enquêteurs sont courtes, précises et faciles à trouver.
- Donner des définitions au début du questionnaire ou en face de questions précises, s'il y a lieu ;
- S'assurer que les périodes de référence en temps et les unités de réponse sont claires pour le répondant ;
- S'assurer que les catégories de réponse sont exhaustives et mutuellement exclusives ;
- Éviter les questions à deux volets, les questions suggestives et les négations doubles ;
- Porter une attention particulière aux questions à caractère sensible ou délicat ;
- Regrouper les catégories de réponse ou précéder la question d'une introduction qui atténuera les craintes des répondants ;
- Effectuer des essais pilotes du questionnaire afin d'observer la façon dont se déroulent, en pratique, les opérations des enquêtes, ce qui comprend l'administration du questionnaire ;

➤ **Procéder à la construction ou amélioration des composantes du processus et de diffusion**

- Construire la plateforme de saisie des données ou adapter celle existante aux besoins de l'activité statistique ;
- inclure dans la plateforme des mécanismes empêchant la digitalisation des données non valides ;
- Établir un lien direct avec les métadonnées et les données à saisir ;
- Construire ou/et adapter les composantes du système nécessaire au traitement et à l'analyse des données ;
 - Inclure les algorithmes permettant d'imputer les données ;
- Effectuer des essais pilotes des composantes du processus et de diffusion ;
- Adapter les supports de diffusion selon le public cible.

➤ **Procéder à la constitution de la base de sondage**

- Évaluer la fiabilité et la pertinence de différentes bases de sondage possibles à l'étape de la planification d'une enquête avant d'arrêter un choix approprié ;
- Utiliser une même base de sondage pour les enquêtes ayant la même population cible afin d'éviter les incohérences, faciliter la comparaison des estimations tirées des enquêtes et réduire les coûts de la mise à jour et de l'évaluation de la base de sondage ;
- Conserver et stocker l'information sur l'échantillonnage afin de procéder à une coordination des enquêtes et mieux gérer les relations avec les répondants et le fardeau de réponse. La rotation de l'échantillon pour la collecte des données est préconisée ;
- Réduire les erreurs de base de sondage en offrant une formation efficace au personnel chargé de l'échantillonnage ;

- Procéder, pour les bases aréolaires, à des vérifications de carte pour assurer la délimitation précise et sans chevauchement des aires géographiques utilisées dans la conception de l'échantillonnage (p. ex., par des vérifications sur le terrain ou l'utilisation d'autres sources de carte). S'il y a lieu, utiliser le Système automatisé de GPS pour créer des unités de base de sondage géographiques ;
- Pour les activités statistiques rattachées à des sources administratives ou pour les activités statistiques dérivées où les changements de couverture peuvent échapper au contrôle du gestionnaire immédiat, déterminer et surveiller la couverture, et se mettre d'accord sur les changements requis avec le gestionnaire de la source ;
- Apporter des réajustements aux données ou utiliser des données supplémentaires tirées d'autres sources pour compenser l'erreur de couverture de la base de sondage.
- Intégrer, dans la documentation sur l'enquête, la description des populations cible et observée, les différences entre la population cible et la population observée, ainsi que la description de la base de sondage et ses erreurs de couverture.

➤ Procéder à l'échantillonnage

- Favoriser la stratification afin d'avoir des strates qui contiennent des unités les plus homogènes possibles à l'égard de l'information souhaitée ;
- Pour les enquêtes longitudinales, choisir des variables de stratification correspondant à des caractéristiques qui demeurent stables au fil du temps ;
- Pour les populations dont la distribution est très asymétrique, créer une strate de grandes unités qui seront incluses dans l'enquête ;
- Pour les plans d'échantillonnage nécessitant plus de deux degrés, déterminer le nombre de degrés d'échantillonnage nécessaires et l'unité statistique appropriée à chaque degré.
- Pour chaque type d'unité possible, examiner la disponibilité d'une base adéquate ou la possibilité d'en créer une, la facilité de la prise de contact et de la collecte/mesure des données, la qualité des données fournies par l'unité et le coût de la collecte ;
- Pour déterminer la taille de l'échantillon, tenir compte :
 - des niveaux de précision nécessaires à la production des estimations de l'enquête ;
 - du type de plan et d'estimateur à utiliser ;
 - de l'accessibilité à l'information auxiliaire ;
 - des contraintes budgétaires ; et
 - des facteurs d'échantillonnage (p. ex., mise en grappe, stratification) et des facteurs hors échantillonnage (p. ex., non-réponse, présence d'unités hors du champ de l'enquête, attrition/déperdition dans les enquêtes longitudinales).
- Pour déterminer la répartition et la taille des échantillons stratifiés, tenir compte des taux d'erreur de classification prévus dans les unités et des autres anomalies de la base de sondage ;
- À l'étape de la mise en œuvre :
 - comparer la taille et les caractéristiques de l'échantillon réel par rapport aux attentes ;
 - comparer la précision des estimations par rapport aux objectifs prévus.
- Pour les enquêtes périodiques qui sont fondées sur un plan d'échantillonnage dans lequel la taille de l'échantillon augmente parallèlement à la taille de la population, élaborer une méthode visant à maintenir stable la taille de l'échantillon et, par conséquent, les coûts de collecte.
- Pour le remaniement sélectif des strates qui ont subi une grave détérioration, mettre en place une stratégie d'actualisation. Si des estimations sont requises pour des domaines d'intérêt précis, former les strates en combinant de petites unités stables liées aux domaines identifiés (p. ex., petites régions géographiques), si possible ;

- En ce qui concerne les enquêtes à panel longitudinal :
 - déterminer la longueur du panel (sa durée dans l'échantillon) en équilibrant la nécessité des données de durée et les effets de structure ;
 - utiliser un plan fondé sur des panels chevauchants en vue de produire des estimations transversales.

3.3.4. Indicateurs de qualité/Vérification de qualité

Dimension de la qualité : exactitude, clarté, pertinence, cohérence, ponctualité, fiabilité, comparabilité et intelligibilité.

Indicateurs sur le processus/Liste de vérification:

- Est-ce que les parties prenantes ont participé à l'élaboration du questionnaire? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le questionnaire répond aux normes de qualité? (*Oui, Non; Sinon pourquoi*);
- Est-ce que tous les guides (enquêteur, contrôleur, superviseur, codification et de saisi) ont été élaborés? (*Oui, Non; Sinon pourquoi*);
- Est-ce que tous les programmes informatiques ont été élaborés? (*Oui, Non; Sinon pourquoi*);
- Est-ce que tous les fichiers des différentes opérations ont été archivés? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le calendrier de diffusion a été élaboré? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les données et les métadonnées à mettre à la disposition du public ont été préparées? (*Oui, Non; Sinon pourquoi*);
- Est-ce que la succession des différentes étapes de l'opération minimise la durée de mise à disposition des résultats? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les différents programmes informatiques sont mis à l'essai? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les différents outils de collecte ont été testés? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les questions posées dans l'enquête pilote sont bien agencées? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les capacités des agents de terrain sont adéquates pour l'enquête proprement dite? (*Oui, Non; Sinon pourquoi*);
- Est-ce que tous les outils ont été validés? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le rapport de l'enquête pilote a été produit? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les recommandations de l'enquête pilote ont été prises en compte lors de la finalisation des différents outils de collecte? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les programmes informatiques ont été reconfigurés en tenant comptes des recommandations de l'enquête pilote? (*Oui, Non; Sinon pourquoi*);
- Est-ce que l'approche d'imputation a été validée lors d'une mise à l'essai? (*Oui, Non; Sinon pourquoi*);
 - Est-ce que la population cible est bien identifiée? (*Oui, Non; Sinon pourquoi*);
 - Est-ce que l'échantillon choisi est représentatif? (*Oui, Non; Sinon pourquoi*);
 - Est-ce que la base de sondage est exhaustive? (*Oui, Non; Sinon pourquoi*);
 - Est-ce que la base de sondage est à jour? (*Oui, Non; Sinon pourquoi*);

3.4. Collecte des données

3.4.1. Objet et finalité

La collecte des données est le rassemblement de toute l'information nécessaire (données et métadonnées) au moyen de différents modes de collecte (y compris les extractions à partir de registres et de bases de données statistiques et administratives et d'autres registres et bases de données non statistiques), et leur chargement dans l'environnement approprié à des fins de traitement approfondi. Bien qu'elle puisse comprendre la validation des formats de l'ensemble de données, elle ne comprend pas les transformations des données recueillies, qui sont toutes faites à la phase du traitement.

Les résultats de cette étape sont : plan de déploiement du personnel de terrain, sensibilisation, enquêteurs formés et recrutés, données collectées et rapport de l'étape de collecte produit.

Ce dernier résultat provient du rassemblement de toutes les informations en rapport avec le déroulement de la collecte. Le rapport est archivé dans un fichier de données (habituellement des microdonnées) qui sert d'intrants à la phase analyse.

3.4.2. Principes d'approche et d'orientation

Les répondants, ou fournisseurs de données sont les ressources essentielles de tout service ou organisme menant des enquêtes. Pour assurer une collaboration constante, il est important de minimiser le fardeau des répondants et de les rassurer quant au caractère confidentiel des informations qu'ils communiquent. Les organismes statistiques ont l'obligation de mettre tout en place afin de satisfaire les exigences relatives à la protection des renseignements personnels à toutes les étapes de collecte et du traitement des données. Leur crédibilité et la coopération des répondants en dépendent.

Tout au long de la collecte, il faut se rassurer de la qualité et du rendement pour pouvoir communiquer des mesures correctives aux superviseurs, aux contrôleurs et aux agents de collecte.

3.4.3. Lignes directrices

Les lignes directrices déclinées ci-après doivent être suivies par toutes les composantes du SSN du Burundi en vue de procéder à la :

➤ Collecte des données

- S'assurer que les enquêteurs reçoivent la formation appropriée et qu'ils disposent des outils de travail adéquats notamment les manuels d'instructions et tout le matériel individuel nécessaire ;
- Exploiter la technologie accessible pour rehausser l'efficacité et la qualité des processus de collecte (Par exemple, l'utilisation des androïdes et des tablettes lors des enquêtes, la collecte électronique des données par Internet) ;
- Contrôler attentivement les opérations de distribution et de récupération des questionnaires papier au cours des enquêtes en veillant à ce que chaque unité qui a été sélectionnée dans l'enquête reçoive le questionnaire qui convient ;
- En cas d'interview direct, contrôler attentivement les opérations de distribution, d'administration et de récupération des questionnaires papier au cours des enquêtes ;
- Lorsque le questionnaire est rempli, vérifier l'exactitude de l'information de couverture et la qualité des données fournies. Dans certains cas, il peut être nécessaire d'effectuer des contre-enquêtes pour se rassurer de la qualité du remplissage des fiches ;

- Lorsque le questionnaire n'est pas reçu, il faut chercher à savoir pourquoi (refus, disparition de l'enquêté, fermeture, etc.) ;
- Mettre en place un système de contrôle pour avoir des informations concernant chaque situation et pallier à l'information manquante ;
- Établir des procédures adéquates de contrôle des échantillons à l'égard de toutes les opérations de collecte des données ;
- Prendre des rendez vous au besoin avec le répondant pour augmenter le taux de réponse;
- Effectuer une évaluation rétrospective de toutes les opérations de collecte et documenter les résultats en prévision d'une utilisation ultérieure.

3.4.4. Indicateurs de qualité/Vérification de qualité

Dimension de la qualité : fiabilité, intelligibilité, exactitude et ponctualité.

Indicateurs sur le processus/Liste de vérification:

- Est-ce que la planification du travail de terrain permet de terminer l'opération dans des délais optimaux ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le personnel de collecte a reçu une formation ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le personnel de collecte a reçu les outils de travail ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le travail sur le terrain a fait l'objet d'une supervision ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que l'évaluation rétrospective de toutes les opérations de collecte a été documentée ? (*Oui, Non; Sinon pourquoi*);

3.5. Traitement des données

3.5.1. Objet et finalité

Cette étape décrit le nettoyage des données et leur préparation aux fins d'analyse. Elle se compose de sous-processus qui permettent de vérifier, de nettoyer et de transformer les données, de manière à pouvoir les analyser et les diffuser comme des produits statistiques. Concrètement, les différentes actions à mener à ce niveau sont : l'intégration des données, la classification, le codage et la saisie des données recueillies, l'examen, la validation, le contrôle et l'imputation, l'application de mesures de contrôle de la divulgation et enfin la pondération et l'estimation.

L'intégration consiste à combiner des données provenant de sources multiples en vue de la création d'une base de données harmonisées obtenues en exécutant des routines d'appariement/de couplage d'enregistrements des microdonnées ou des macrodonnées de différentes sources (données d'enquêtes, de recensements ou de sources administratives). L'intégration peut se faire à n'importe quelle étape du traitement et doit veiller aux exigences en matière de protection des données et de confidentialité des renseignements.

L'incidence des opérations de collecte et de saisie des données sur la qualité des données est directe et essentielle, car ces données sont les principaux intrants lors du traitement. La classification et le codage des données d'entrée réfèrent à tout le processus qui attribue une valeur numérique à une réponse. Bien que fréquemment automatisée, cette étape requiert parfois l'intervention humaine car elle peut demander de prendre des décisions complexes.

La saisie des données renvoie à tout processus qui permet de mettre et de sauvegarder sur un support électronique l'information communiquée par un répondant.

Il arrive parfois que la saisie se fasse dans le cadre d'un processus automatisé et intégré avec la collecte de données. Mais généralement, elle est une opération distincte de la collecte qui consiste à introduire manuellement sur clavier les informations collectées.

L'examen et la validation des données consistent à déterminer d'éventuels problèmes, erreurs et divergences comme les valeurs aberrantes, la non-réponse partielle et le codage erroné. Ils s'appliquent aux données de n'importe quel type de source avant et après l'intégration.

Au cours du traitement, de mesures doivent être prises afin de contrôler que les données (et les métadonnées) devant être diffusées n'enfreignent pas les règles appropriées en matière de confidentialité. Il peut comprendre des vérifications pour déceler la divulgation primaire et secondaire ainsi que l'application de méthodes de suppression de données ou de techniques de perturbation. Le degré et la méthode de contrôle de la divulgation peuvent varier pour différents types de produits, par exemple l'approche utilisée pour des ensembles de microdonnées à des fins de recherche sera différente de celle pour les tableaux ou les cartes publiés.

Dans le cas des statistiques produites sur une fréquence régulière, cette étape a lieu à chaque itération.

Les résultats de cette étape sont : les questionnaires codifiés, la base de données apurée et mesures de contrôle de divulgation appliquées.

L'estimation permet de créer des données agrégées et des chiffres de population à partir de microdonnées ou d'agrégats calculés à partir de l'échantillon. Il comprend la sommation des données pour les enregistrements qui ont certaines caractéristiques en commun, la détermination des mesures de la moyenne et de la dispersion.

3.5.2. Principes d'approche et d'orientation

Pour chaque enquête, il faut veiller à intégrer les résultats d'autres opérations précédentes et le cas échéant les données administratives afin de pouvoir comparer certains indicateurs.

Les catégories de réponse aux questions fermées peuvent être codées d'avance dans le questionnaire. Des cases pour les codes numériques peuvent être disposées à côté de la réponse à coder ou dans la marge pour les questionnaires sur support papier.

Compte tenu de sa forte incidence sur l'exactitude des données, l'opération de saisie doit se faire à l'aide d'outils qui permettent de mesurer la qualité et le rendement. Il faut également favoriser la double saisie pour pouvoir comparer les résultats

Après la saisie des données, il faut passer à l'apurement de ces dernières qui vise trois objectifs :

- l'amélioration future de l'instrument d'enquête,
- la mise à disposition de l'information sur la qualité des données de l'enquête,
- nettoyer les données.

En cas de données manquantes ou aberrantes, les personnes qui ont pleinement accès aux micro-données et qui possèdent une bonne information auxiliaire procèdent à l'imputation.

Les principes qui sous-tendent les activités de contrôle de la divulgation sont régis par la loi n°1/17 du 25 septembre 2007 portant organisation du SSN du Burundi qui stipule en son article 16 que « Les données individuelles recueillies par les services et organismes relevant du système statistique national ne peuvent faire l'objet de divulgation sauf autorisation explicite accordée par les personnes physiques ou morales concernées. »

Il existe deux principaux types de divulgation qui suscitent des préoccupations : la divulgation de l'identité et celle des attributs. Le premier intervient lorsqu'un répondant en particulier (personne, ménage, entreprise, etc.) peut être identifié à partir des données diffusées tandis que le second a lieu lorsqu'il est possible de révéler l'information confidentielle au sujet d'un répondant, à partir des données diffusées.

3.5.3. Lignes directrices

Les lignes directrices déclinées ci-après doivent être suivies par toutes les composantes du SSN du Burundi en vue:

➤ De procéder à la saisie des données

- S'assurer que les agents de saisie des données reçoivent la formation appropriée et qu'ils disposent des outils de travail adéquats notamment les manuels d'instructions et tout le matériel individuel nécessaire ;
- Exploiter la technologie accessible pour rehausser l'efficacité et la qualité du processus de saisie des données. Par exemple, l'utilisation de la saisie optique des données et du scanner sont autant de possibilités fournies par les nouvelles technologies ;
- Mettre en place des systèmes de contrôle efficaces pour assurer la sécurité de la saisie, de la transmission et de la manipulation des données ;
- Intégrer des vérifications automatiques pour les erreurs que l'opérateur de saisie des données peut corriger (c.-à-d). vérifications qui mettront en évidence les erreurs d'introduction sur clavier). Enregistrer ces cas pour analyse et examen ultérieurs ;
- Effectuer une évaluation rétrospective de toutes les opérations de saisie des données, et documenter les résultats en prévision d'une utilisation ultérieure.

➤ De procéder à la vérification

- Fournir un guide des règles de vérification aux agents de recodification ;
- Offrir une formation adéquate à tout le personnel concerné par la vérification et effectuer un suivi des travaux en soi ;
- S'assurer que l'interprétation des règles de vérification est identique et cohérente pour l'ensemble des vérificateurs ;
- Appliquer de nouveau les vérifications aux unités qui ont subi des corrections pour s'assurer qu'aucune autre erreur n'a été introduite directement ou indirectement par le processus de correction ;
- Utiliser un logiciel généralisé et réutilisable (type SAS, SPSS ou autres) bien qu'une certaine intervention manuelle soit nécessaire ;
- Déterminer les valeurs des données extrêmes d'une période d'enquête ou entre les périodes d'enquête (cet exercice est appelé le processus de détection des valeurs aberrantes).

➤ De procéder à l'imputation

- Évaluer le type de non-réponse, c'est-à-dire tenter de déterminer les variables auxiliaires qui peuvent expliquer le ou les mécanismes de non-réponse afin de les utiliser pour enrichir la méthode d'imputation. Intégrer ces variables auxiliaires dans la méthode d'imputation ;
- Élaborer et mettre à l'essai avec soin l'approche d'imputation ;
- Étudier la qualité et la pertinence des variables disponibles afin de déterminer celles qu'on peut utiliser comme variables auxiliaires ou variables de couplage ou pour établir des classes d'imputation ;
- Se rassurer que les enregistrements imputés ressemblent à l'enregistrement rejeté à l'étape de contrôle et de vérification ;
- Marquer les valeurs imputées et désigner clairement les méthodes et les sources d'imputation ;
- Conserver les valeurs imputées et non imputées des champs de l'enregistrement pour les besoins de l'évaluation ;
- Évaluer le degré et les effets de l'imputation ;
- Mesurer efficacement la variance d'échantillonnage dans le cadre de l'imputation et mesurer la variance/ou valeur ajoutée introduite par cette dernière ;
- Tenir compte du degré et de l'incidence de l'imputation au cours de l'analyse des données.

➤ De procéder à l'estimation

- Prendre en compte tous les aspects du plan d'échantillonnage en intégrant les poids d'échantillonnage dans le processus d'estimation (Stratification, phases et/ou degrés multiples de l'échantillon, etc.);
- Utiliser le redressement et les méthodes de calage si des données auxiliaires fiables sont disponibles pour les unités échantillonnées ;

➤ De procéder au contrôle de la divulgation⁶

- Déterminer le degré de confidentialité des informations par rapport aux règles établies à l'ISTEEBU et dans le SSN;
- Ne pas diffuser de tableaux de données quantitatives contenant des valeurs avec des informations jugées confidentielles ;
- Vérifier si les informations à basse fréquence posent un problème de confidentialité ;
- Anonymiser les informations confidentielles;
- Utiliser, s'il y a lieu, l'une des techniques d'arrondissement des valeurs pour protéger la confidentialité (arrondissement déterministe ou classique et aléatoire) ;
- Veiller à ce que des responsable des bases soient désignés officiellement ;
- En ce qui a trait à la diffusion de micro-données :
 - évaluer le risque de divulgation des fichiers de micro-données ;
 - appliquer l'une des deux *méthodes des données à diffusion restreinte* généralement utilisées pour contrôler le risque de divulgation (*réduction des données et modification des données*);
 - s'assurer que la population est assez grande dans le cas de certains groupes identifiables;

⁶ Mark Elliot, Anco Hiendepool, Eric Schulte Nordholt, Jean Louis Tambay, Tomas Wende : Glossaire du contrôle de la divulgation statistique, 4 novembre 2003

- élargir les catégories de variables;
- regrouper les valeurs extrêmes supérieures et inférieures ;
- supprimer certaines variables pour préserver la confidentialité du répondant;
- supprimer les renseignements fournis par certains répondants et les remplacer par les valeurs imputées, le cas échéant.

3.5.4. Indicateurs de qualité/Vérification de qualité

Dimension de la qualité : Exactitude, Fiabilité, Cohérence et Accessibilité.

Indicateurs sur le processus/Liste de vérification :

- Est-ce que le personnel de saisie a reçu une formation ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le personnel de saisie a reçu le guide de codage? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le travail de saisie a fait l'objet d'une supervision ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que l'évaluation rétrospective de toutes les opérations de saisie a été documentée ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que un guide d'apurement a été élaboré ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le personnel d'apurement a reçu une formation ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le personnel d'apurement a fait l'objet d'une supervision ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que l'évaluation rétrospective de toutes les opérations d'apurement a été documentée ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que la méthode d'imputation a été documentée? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le taux d'imputation par variable a été documenté? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le calcul des poids tient compte du plan d'échantillonnage? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le calcul de la variance tient compte du plan d'échantillonnage et de la méthode d'estimation? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les estimations produites s'accompagnent d'une mesure de la variance? (*Oui, Non; Sinon pourquoi*);
- Est que le processus d'estimation au niveau de l'échantillon tient compte du plan d'échantillonnage ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les règles de confidentialité ont été appliquées ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que des mesures ont été prises pour protéger des données sensibles? (*Oui, Non; Sinon pourquoi*);

3.6. Analyse

3.6.1. Objet et finalité

L'analyse consiste à transformer les données recueillies en produits statistiques et à les examiner en détail, préparer les commentaires, les notes techniques, etc., de sorte que ces produits soient adaptés à l'usage prévu avant la diffusion. Elle comprend également la validation des produits qui consiste à la prise en compte de la somme des connaissances accumulées sur le domaine et qui sont ensuite appliquées au cours de l'analyse pour améliorer la qualité des produits par rapport aux attentes.

Pour étayer l'analyse, sont également produits des outils supplémentaires comme des indices, des ratios, des tendances ou des séries désaisonnalisées. L'interprétation des résultats permet, au moyen d'analyse statistique approfondie, d'explicitier les statistiques produites.

Les résultats de cette étape sont : les tableaux, les résultats préliminaires, les résultats définitifs et les rapports des analyses thématiques.

3.6.2. Principes d'approche et d'orientation

Avant de produire toutes les statistiques descriptives permettant de faire une analyse sommaire des phénomènes étudiés, il importe préalablement d'évaluer la qualité des données produites afin d'aviser les utilisateurs sur cette dernière.

Les évaluations de la qualité des données sont impératives pour les services statistiques. Il s'agit ici d'examiner, par exemple, l'efficacité des plans d'enquête, l'importance de la non-réponse, ou les pratiques d'imputation douteuses.

Pour autant que ce soit possible, les résultats de l'évaluation doivent être suffisamment actuels pour aider les utilisateurs à analyser les données et le personnel à concevoir les prochaines enquêtes.

L'interprétation des produits statistiques consiste souvent à examiner des tableaux et graphiques de diverses mesures de récapitulation, notamment, les distributions de fréquences, les moyennes et les étendues. D'autres genres d'analyses de données plus perfectionnées sont aussi possibles, et l'inférence statistique peut être appliquée pour vérifier des hypothèses ou étudier des liens entre des caractéristiques.

3.6.3. Lignes directrices

Les lignes directrices déclinées ci-après doivent être suivies par toutes les composantes du SSN du Burundi en vue de procéder à la :

➤ Validation des résultats

- Vérifier la cohérence par rapport aux sources de données externes, par exemple d'autres enquêtes antérieures, ou des données administratives;
- Vérifier la cohérence interne en calculant des ratios qui sont censés se situer à l'intérieur de certaines limites (proportions hommes-femmes, valeur moyenne des biens, etc.);
- Analyser individuellement les unités les plus importantes pour ce qui est de leurs contributions aux estimations globales (généralement appliquées aux enquêtes-entreprises);
- Calculer les indicateurs de la qualité des données, par exemple les taux de non-réponse, les taux d'imputation et les coefficients de variation;
- Vérifier les tableaux croisés pour assurer l'uniformité et l'exactitude des variables clés et des domaines importants;
- Analyser la cohérence des données sur la base des événements récents dans l'actualité.

➤ Analyse des résultats

- Choisir une méthode analytique qui convient à la question examinée ;
- Si plus d'une source de données servent à l'analyse, vérifier si elles sont cohérentes et quelle est la façon de les combiner efficacement;
- Déterminer si les valeurs imputées doivent être incluses dans l'analyse et, le cas échéant, la façon dont elles doivent être traitées ;
- Déterminer la façon de traiter la non-réponse totale et/ou partielle dans l'analyse ;
- Éviter de faire des liens de causes à effets avec des phénomènes ou des facteurs externes aux données;

- Prendre garde à étudier les tendances à court terme en n'oubliant pas de tenir également compte des tendances à moyen et à long terme ; Les tendances à court terme ne représentent souvent que de légères fluctuations d'une tendance plus importante à moyen ou à long terme.
- Éviter les points de référence arbitraires dans le temps, comme le changement entre l'année précédente et l'année en cours ;
- Privilégier l'utilisation de points de référence comportant une plus grande signification tels que le dernier point tournant pour les données économiques, les différences intergénérationnelles pour les statistiques démographiques et les changements législatifs pour les statistiques sociales ;
- Consulter des experts au sujet des domaines spécialisés et des méthodes statistiques ;
- Expliquer les pratiques ou les méthodes d'arrondissement. Dans la présentation des données arrondies, le nombre de chiffres significatifs ne doit pas être trop grand de manière à être cohérent avec la précision des données ;
- Prendre soin d'établir la distinction entre le taux de variation et la variation en points de pourcentage lorsqu'on présente des détails sur les taux ;
- Définir la base de calcul des taux ;

3.6.4. Indicateurs de qualité/Vérification de qualité

Dimension de la qualité : exactitude, pertinence, cohérence et comparabilité

Indicateurs sur le processus/Liste de vérification :

- Est-ce que les données produites ont été évaluées? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les statistiques descriptives nécessaires ont été produites? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les données des autres sources sont cohérentes à celles produites? (*Oui, Non; Sinon pourquoi*);
- -Est ce que les autres sources de données ont été intégrées? (*Oui, Non; Sinon pourquoi*);
- Est-ce que le taux de couverture a été évalué ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les données ont été confrontées à d'autres sources à l'interne ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les différences observées ont été analysées et expliquées ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les réunions à l'interne de validation des différents produits ont eu lieu ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les réunions de validation avec les parties prenantes ont eu lieu ? (*Oui, Non; Sinon pourquoi*);

3.7. Diffusion

3.7.1. Objet et finalité

La diffusion consiste à mettre l'information obtenue, dans le cadre d'une activité statistique, à la disposition des utilisateurs. Divers modes de diffusion peuvent être envisagés : supports électroniques accessible sur site web; CD-ROM, publications sur papier; microfiches; fichiers de micro-données de réponse d'enquête anonymisées; données transmises par téléphone ou par télécopieur en réponse à une demande spéciale; allocutions prononcées en public; exposés ou entrevues aux médias.

Les résultats de cette étape sont : les données et des métadonnées à diffuser formatées, données mises en ligne, rapport imprimé, communiqués de presse préparés, rapports des ateliers de validation disponible, etc.

3.7.2. Principes d'approche et d'orientation

Il convient de s'assurer que l'information communiquée aux utilisateurs est exacte, complète, accessible, compréhensible, utilisable et actuelle; qu'elle satisfait aux exigences en matière de protection de la confidentialité. En matière de diffusion, il faudra tenir compte des besoins des utilisateurs et tirer parti des progrès technologiques afin de leur permettre un accès aisé et un usage facile.. Il convient également de prendre en considération les attentes des utilisateurs, par exemple en se fondant sur leurs commentaires et observations ou les activités de plaidoyer .

3.7.3. Lignes directrices

Les lignes directrices déclinées ci-après doivent-être suivies par toutes les composantes du SSN du Burundi en vue de procéder à la diffusion des données :

- S'assurer que les données diffusées à la suite de toutes les étapes de traitement correspondent aux données de source obtenues ;
- Tester les produits électroniques avant leur diffusion pour s'assurer qu'ils fonctionnent comme prévu ;
- S'assurer que les mesures de la qualité des données ou, si possible, les outils nécessaires à leur calcul (p. ex. les tables de validation des coefficients de variation, les programmes de variance d'échantillonnage) accompagnent le produit diffusé ;
- Documenter les produits à diffuser et fournir les documents méthodologiques utilisés;
- Etablir un calendrier de mise à disposition des données pour les publications régulières ;
- Mettre en place des mécanismes d'accord institutionnels explicites sur les modalités d'accès et d'utilisation des microdonnées ;
- Etablir des partenariats avec la presse en vue de promouvoir les produits statistiques.
- Désigner une personne-ressource avec ses coordonnées de contact(téléphone, adresse mail, etc).

3.7.4. Indicateurs de qualité/Vérification de qualité

Dimension de la qualité : fiabilité, accessibilité, ponctualité et intelligibilité

Indicateurs sur le processus /Liste de vérification:

- Est-ce que les données diffusées à la suite de toutes les étapes de traitement correspondent-elles aux données de source obtenues ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les produits statistiques sont accompagnés des métadonnées ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les utilisateurs ont été avisés sur la façon d'obtenir plus d'information sur les produits ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les bases des données sont anonymisées pour garantir la confidentialité? (*Oui, Non; Sinon pourquoi*);
- Est-ce que des accords institutionnels explicitant les modalités d'accès aux microdonnées sont mis en place ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que un calendrier des publications statistiques établi a été respecté ? (*Oui, Non; Sinon pourquoi*);

3.8. Evaluation

3.8.1. Objet et finalité

L'objectif de l'évaluation est de pouvoir apprécier comment l'opération a été réalisée, la façon dont elle a permis de satisfaire aux besoins des utilisateurs et de pouvoir tirer des leçons en vue d'améliorer les futures opérations. Elle peut concerner une étape du processus organisationnel statistique, en s'appuyant sur un large éventail d'intrants quantitatifs et qualitatifs, ainsi qu'en cernant les améliorations potentielles et en établissant les priorités.

La gestion de l'évaluation du processus de production statistique repose sur les intrants recueillis tout au long des différentes phases. Elle a lieu logiquement tout au long du processus mais elle peut intervenir à la fin.

Dans le cas des produits statistiques qui sont produits sur une base périodique, l'évaluation devrait avoir lieu pour chaque itération, de manière à déterminer si d'autres itérations sont nécessaires dans l'avenir. Dans le cas particulier des opérations de production statistique de routine, l'évaluation n'est pas effectuée formellement pour chaque itération. En pareil cas, chaque itération commence à partir de la phase de collecte.

L'évaluation s'effectue à chaque étape de l'opération. Elle nécessite des intrants de plusieurs formes dont les commentaires des utilisateurs, les métadonnées sur les processus (paradonnées), l'appréciation de la qualité de tous les outils de collecte, et les suggestions du personnel.

Les rapports des progrès réalisés par rapport à un plan d'action convenu lors d'une itération précédente peuvent également servir d'intrants aux nouvelles évaluations.

Un consensus sur un plan d'action est dégagé par les pouvoirs décisionnaires (généralement le comité de pilotage) à partir du rapport d'évaluation. Ce plan d'action doit prendre en considération les mécanismes de suivi évaluation, qui peuvent, à leur tour, servir d'intrants aux évaluations d'itérations futures.

Les résultats de cette étape sont : les rapports d'évaluation par étape, rapport d'évaluation final et un plan d'action pour les opérations futures.

3.8.2. Principes d'approche et d'orientation

Il est impératif de rédiger les rapports de chaque étape de l'opération. Ces rapports doivent donner aussi les détails du déroulement de chaque étape de l'opération et les problèmes rencontrés ainsi que les stratégies adoptées pour les surmonter. Ceux-ci doivent comprendre les éléments des commentaires des utilisateurs sur les métadonnées, les processus et l'appréciation de la qualité.

Les utilisateurs sont rarement capables d'exécuter l'évaluation. Celle-ci incombe pour la plus part du temps au service statistique responsable de l'opération, qui en communique les résultats aux utilisateurs, le plus tôt possible. Elle porte essentiellement sur la qualité des données et précise dans quelle mesure elles sont pertinentes et représentatives.

Outre la pertinence des données, l'évaluation de la qualité permet de vérifier s'il existe un lien entre certains types d'erreurs et certaines étapes du processus d'enquête. Cette démarche permet d'améliorer la qualité d'une prochaine édition de l'enquête ou d'enquêtes similaires.

Après la production du rapport d'évaluation, les partenaires intéressés par l'opération élaborent un plan d'action pour les opérations futures. Des mécanismes de suivi et d'évaluation de la mise en œuvre de ce plan doivent être également mis en place.

3.8.3. Lignes directrices

Les lignes directrices déclinées ci-après doivent être suivies par toutes les composantes du SSN du Burundi en vue de procéder à l'évaluation du processus de production statistique :

- Mettre en place une équipe qui sera chargée de mener l'évaluation du programme (processus de production récurrente) ou du projet (processus de production spéciale);
- S'assurer que le gestionnaire responsable de chacune des étapes/processus rassemble toutes les données d'entrées et les mette à la disposition de la personne ou de l'équipe d'évaluation. Il s'agit essentiellement:
 - des commentaires des utilisateurs ;
 - des métadonnées sur les processus ;
 - des taux d'erreurs calculés dans le système de saisie ;
 - des suggestions du personnel ;
 - des rapports de progrès réalisés par rapport à un plan d'action ;
 - etc. ...
- Impliquer les utilisateurs des résultats d'évaluation, qu'ils viennent d'un organisme statistique ou d'ailleurs, dans la définition des objectifs du programme d'évaluation de la qualité des données. Faire de même pour le processus d'évaluation, lorsque les circonstances le permettent.
- Évaluer la qualité en se fondant sur l'opinion des spécialistes externes qui sont familiarisés avec l'enquête, pour qu'ils vérifient si ses résultats sont plausibles ;
- Synthétiser le rapport d'évaluation de l'analyse des données d'entrées, indiquer tout problème de qualité particulier relevé au cours du processus de production statistique et recommander des changements au besoin ;
- Rassembler les pouvoirs décisionnaires nécessaires pour établir un plan d'action fondé sur les recommandations du rapport d'évaluation et s'entendre sur ce plan ;
- Mettre en place des mécanismes de suivi évaluation pour la mise en œuvre de ce plan.

3.8.4. Indicateurs de qualité/Vérification de qualité

Dimension de la qualité : pertinence, intelligibilité et fiabilité.

Indicateurs sur le processus/Liste de vérification :

- Est-ce qu'une équipe d'évaluation a été mise en place ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les données d'entrée ont été collectées(les rapports des différentes étapes) ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les utilisateurs des résultats d'évaluation ont été impliqués dans la définition des objectifs du programme d'évaluation de la qualité des données ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que un plan d'action fondé sur les recommandations du rapport d'évaluation a été adopté par les pouvoirs décisionnaires ? (*Oui, Non; Sinon pourquoi*);
- Est-ce que les mécanismes de suivi évaluation pour la mise en œuvre du plan ont été établis ? (*Oui, Non; Sinon pourquoi*);

REFERENCES BIBLIOGRAPHIQUES

1. Guide de la Statistique de la Coopération au développement de la commission européenne Edition 2013;
2. Cadre d'assurance de la qualité de statistique canada 2002 ;
3. Charte Africaine de la Statistique Mai 2009;
4. Code de bonnes pratiques de la statistique européenne, septembre 2011 ;
5. Guidelines for the template for a generic national quality assurance framework ;
8 February 2012
6. Guide pratique pour la conception d'enquêtes sur les ménages; Nations Unies New York, 2010
7. Handbook of quality guidelines for statistical production in Tanzania, November 2012;
8. Méthodes et pratiques d'enquête du Canada Edition 2010;
9. Statistique Canada : Lignes directrices concernant la qualité, Cinquième édition – Octobre 2009.