

REPUBLIQUE DU CAMEROUN

Paix-Travail-Patrie

Ministère de l'Economie, de la
Planification et de l'Aménagement
du Territoire

Conseil National de la Statistique



REPUBLIC OF CAMEROON

Peace-Work-Fatherland

Ministry of Economy, Planning and
Regional Development

National Statistics Council

MISE EN ŒUVRE DU CADRE NATIONAL D'ASSURANCE QUALITE

Manuel de lignes directrices pour les enquêtes et les recensements

Première édition

Juin 2019



Secrétariat du Conseil National de la Statistique

BP: 134 Yaoundé Tél: (237) 22 22 04 45 Fax: (237) 22 23 24 37
Site Web: www.statistics-cameroon.org



PARIS21

Sommaire

Avant-propos	4
Contexte	6
La démarche qualité du SNIS.....	6
Le Cadre National d'Assurance Qualité	6
But et portée des lignes directrices	7
Démarche d'élaboration des lignes directrices.....	7
Utilisation du manuel.....	8
Section 1 : la démarche qualité et le processus de production statistique	9
I. La démarche qualité en statistique.....	9
II. Le Modèle Générique du Processus de Production Statistique	10
Définition et objectif	10
Origine et historique	10
Présentation	11
Préalables.....	11
Section 2 : Les lignes directrices pour la qualité des processus	13
I. Détermination des besoins.....	14
I.1. Identification et validation des besoins.....	14
I.2. Identification des produits statistiques.....	15
I.3. Détermination des concepts	16
I.4. Vérification de la disponibilité des données.....	16
I.5. Etude de faisabilité	17
II. Conception.....	19
II.1. Conception des produits	19
II.2. Conception et description des variables	20
II.3. Conception de la collecte.....	20
II.4. Conception de la base de sondage et de l'échantillon	22
II.5. Conception du traitement et de l'analyse	23
II.6. Conception du système de production	24
II.7. Conception de la diffusion et de la dissémination	25
II.8. Conception de l'archivage	26
II.9. Conception de l'évaluation	26
III. Elaboration.....	27
III.1 Elaboration de l'instrument de collecte.....	27
III.2 Elaboration du processus de production.....	28
III.3 Elaboration du processus de diffusion.....	29
III.4 Configuration du déroulement des activités	30
III.5 Mise à l'essai du système de production	31
III.6 Mise à l'essai du processus de production statistique	31
III.7 Mise au point finale du système et du processus de production	32
IV. Collecte.....	32
IV.1. Création/mise à jour de la base de sondage et sélection de l'échantillon	33
IV.2. Préparation de la collecte	34
IV.3. Exécution de la collecte	35
IV.4. Mise au point finale de la collecte et centralisation des données	37
V. Traitement	38
V.1 Intégration des données	38

V.2 Classification et codage	39
V.3 Examen et validation de données	40
V.4 Contrôle et imputation	41
V.5 Dérivation de nouvelles variables et unités	41
V.6 Calcul des poids de sondage	42
V.7 Calcul des indicateurs et estimation des paramètres	43
V.8 Mise au point finale des fichiers de données	44
VI. Analyse	44
VI.1. Elaboration des projets de produits	45
VI.2. Validation des produits statistiques	45
VI.3. Interprétation et explication des données	46
VI.4. Contrôle de la non divulgation des données individuelles	47
VI.5. Mise au point finale des produits	48
VII. Diffusion	48
VII.1. Mise à jour des systèmes de diffusion (préparation de la diffusion)	49
VII.2. Production des produits de diffusion	49
VII.3. Mise au point du processus de diffusion	50
VII.4. Promotion des produits de diffusion	50
VII.5. Gestion du soutien aux utilisateurs	51
VIII. Evaluation	52
VIII.1 Recueil des données d'évaluation	52
VIII.2 Conduite de l'évaluation	53
VIII.3 Adoption du plan d'actions	53
Bibliographie	55
Annexes	56
Annexe 1 : Nomenclature du processus de production statistique	56

Avant-propos

Les politiques et stratégies qui sont conçues et mises en œuvre en vue d'atteindre les objectifs de développement du Cameroun doivent être fondées sur des statistiques officielles de qualité à savoir des statistiques pertinentes, exactes, actuelles, accessibles, intelligibles et cohérentes. La question de qualité est depuis quelques années au centre des préoccupations du Système National d'Information Statistique (SNIS) qui connaît des améliorations certaines de sa production avec la mise en œuvre des Stratégies Nationales de Développement de la Statistique (SNDS). A ce jour, l'Institut National de la Statistique qui assure le Secrétariat technique du Conseil National de la Statistique (CNS), a mis à la disposition du SNIS, un ensemble d'outils de normalisation pour encadrer la production. Il s'agit entre autres des nomenclatures, du manuel des concepts et des définitions utilisées dans les statistiques officielles et du dictionnaire des enquêtes statistiques. Plus récemment encore, l'INS a produit avec la participation des administrations sectorielles et l'appui multiforme d'AFRISTAT et de Statistique Canada, le Cadre National d'Assurance Qualité (CNAQ) qui constitue le cadre d'orientation de la politique qualité du SNIS.

Le Ministre de l'Economie, de la Planification et de l'Aménagement du Territoire, Président du Conseil National de la Statistique a le plaisir de vous présenter la toute première édition du manuel des lignes directrices pour les enquêtes et recensements au Cameroun, un des outils opérationnels du CNAQ.

Ce document présente les bonnes pratiques identifiées pour la mise en œuvre efficace d'un processus de production statistique. Cette première édition principalement adossée sur les enquêtes et les recensements devrait dans un avenir proche intégrer la production des statistiques de sources administratives, qui sont appelées à devenir les leviers de la production statistique du SNIS.

Le présent document de lignes directrices n'a pas la prétention de remplacer l'expertise du personnel chargé de la conception des enquêtes et recensements. Il a d'ailleurs été élaboré de manière participative par les acteurs majeurs du SNIS, et est inspiré des bonnes pratiques observées aussi bien au niveau national qu'international. Ce manuel vise surtout à développer la culture qualité du personnel tout au long du processus de production statistique, en les amenant à se conformer aux normes en la matière. Il est aussi destiné à aider les nombreux professionnels non statisticiens qui participent à la production statistique, et en sont même parfois responsables dans certaines structures.

Je voudrais saisir cette occasion pour remercier tous ceux qui ont contribué à la conception et à l'élaboration de ce document, notamment les administrations du SNIS. J'adresse mes félicitations à l'INS qui a su conduire le projet et encadrer toutes les contributions. Je voudrais enfin, inviter et encourager tous les praticiens (statisticiens, démographes et tous les autres) à se référer à ces lignes directrices dans leurs travaux de production statistique. C'est aussi à ce prix que l'objectif de l'amélioration continue de la qualité des données statistiques pourra être atteint.

Le Ministre de l'Economie, de la
Planification et de l'Aménagement du
Territoire,
Président du Conseil National de la
Statistique

ALAMINE OUSMANE MEY

Contexte

La démarche qualité du SNIS

La démarche qualité du SNIS consiste à doter ses composantes des instruments institutionnels, des cadres ainsi que des outils techniques pour gérer la qualité dans l'activité statistique nationale. Ces instruments et outils peuvent se regrouper en quatre (4) grandes catégories:

- une Structure administrative (Division, Cellule, etc.) dédiée entièrement à la qualité;
- un cadre national d'assurance de la qualité;
- des lignes directrices pour la production des données ;
- des rapports de suivi de la qualité des produits statistiques ;
- d'autres cadres de gestion pouvant contribuer au développement et à l'amélioration de la qualité, qui se réfèrent à des aspects n'ayant pas trait directement à la statistique, mais desquels dépendent l'efficacité du système statistique dans la fourniture des produits et services. Ces cadres ont trait par exemple au cadre législatif et réglementaire qui encadre l'activité statistique nationale, à la gestion du personnel dans la production statistique, au financement de l'activité statistique, aux questions de planification stratégique et de gestion du rendement et de la promotion du résultat, etc.

La mise en œuvre de tous ces éléments et leur suivi permanent permettront de développer au sein du système statistique une culture de qualité. La réussite de cette initiative nécessite l'engagement inconditionnel du Conseil National de la Statistique (CNS) sur la question de la qualité, ainsi qu'une forte gouvernance en la matière. Cela nécessitera également une forte adhésion au projet de la part de tous les acteurs du système, les producteurs en particuliers. Il conviendra donc de mettre en place une stratégie efficace de communication et de sensibilisation autour de la démarche qualité, du CNAQ et de tous les outils opérationnels pour sa mise en œuvre.

Le Cadre National d'Assurance Qualité

Le Cadre National d'Assurance Qualité (CNAQ) décrit un ensemble d'outils, de documents et de mesures mises (ou à mettre) en place par le Système Statistique National pour gérer la qualité des données produites. Il constitue le cadre d'orientation de la politique qualité du SNIS. C'est un document qui présente la vision de notre Système en termes de qualité, et qui donne les grandes orientations pour la mise en œuvre de la démarche qualité.

Le CNAQ a été élaboré en 2017 sur la base d'un diagnostic du système qui a permis de relever les points forts et les points faibles du système sur tous les aspects de la production statistique. Il a été développé dans la logique de la vision du SNIS consignée dans la Stratégie Nationale de Développement de la Statistique (SNDS 2015 - 2020) qui est d'asseoir à l'horizon 2020, un SNIS stable, efficient, produisant et diffusant à temps des données statistiques fiables, actuelles, répondant aux besoins de tous les utilisateurs.

Le CNAQ est développé autour des quatre (4) dimensions retenues par les Nations Unies pour bâtir un CNAQ, à savoir la gestion du système statistique, la gestion de l'environnement statistique, la gestion des processus et la gestion des produits

But et portée des lignes directrices

Le CNAQ constitue le cadre d'orientation de la politique qualité du SNIS. Il définit les idéaux de qualité auxquels le système aspire dans l'accomplissement de ses missions. La mise en œuvre de ce cadre nécessite la production d'un ensemble d'outils opérationnels pour accompagner les praticiens dans tous les aspects de la production statistique. Les lignes directrices pour les enquêtes et recensements constituent ainsi l'un de ces outils techniques.

Le principal objectif de ces lignes directrices est de fournir aux producteurs du SNIS une liste plus ou moins exhaustive de principes directeurs et de pratiques exemplaires à appliquer lors de la conception et la mise en œuvre d'enquêtes. Ces principes et pratiques ont été répertoriés auprès des systèmes statistiques de référence, des cadres et standards internationaux régissant l'activité statistique, ainsi que des bonnes pratiques relevées par les praticiens au niveau national. Les lignes directrices sont recommandées mais pas obligatoires, l'idée étant d'aider le producteur de statistique publique à se conformer aux normes en matière de qualité édictées dans des politiques et directives (instructions officielles obligeant à instaurer (ou à éviter) une mesure particulière), ou encore dans les bonnes pratiques recensées.

Démarche d'élaboration des lignes directrices

Sur le plan technique, la méthodologie a tenu en plusieurs étapes :

- L'adaptation au contexte national du modèle générique du processus de production statistique (MGPPS) adopté au niveau international, modèle qui définit et décrit l'ensemble des activités nécessaires à l'établissement de statistiques officielles : cela a consisté principalement à comprendre chacune des huit étapes proposées, à les modifier si nécessaire et à les valider pour notre contexte, à décliner chaque étape en sous processus ;

- L'identification et la définition des livrables pour chaque sous processus, notamment l'output technique attendu de la mise en œuvre de chaque sous processus et devant permettre de poursuivre le processus de production ;
- L'identification et la définition des activités élémentaires à mener pour l'atteinte des objectifs à chaque niveau du processus ;

Cette première édition se limite à identifier et à définir les activités élémentaires nécessaires à chaque sous processus. Pour la suite du processus, il sera question de répertorier les normes, d'identifier les principes directeurs associés à chaque sous processus, ainsi que les bonnes pratiques pour la mise en œuvre des activités identifiées.

Le document est organisé autour de deux sections. La première est consacrée à la présentation du modèle générique du processus de production statistique, tandis que les développements sur la portée et l'objet de chaque sous processus et le contenu des activités proposées sont faits dans la deuxième section.

Utilisation du manuel

Le manuel des lignes directrices peut être utilisé à plusieurs fins

- Outil de gestion de la qualité des processus et des produits ;
- Outil de planification opérationnelle des activités ;
- Outil de communication (gouvernance, redevabilité, information des utilisateurs) ;
- Outil de formation (des intervenants dans la chaîne de production des statistiques : responsables et cadres de l'INS et des sectoriels, gestionnaires des registres administratifs,)
- Outil de budgétisation.

Section 1 : la démarche qualité et le processus de production statistique

I. La démarche qualité en statistique

La démarche–qualité en statistique est un processus mis en œuvre par l'office statistique à l'effet d'améliorer la performance du système, de maîtriser et garantir la qualité dans la mise en œuvre des activités statistiques. Ce processus doit être continu dans le temps, progressif et demande l'adhésion de tous, en particulier du top management de l'office.

La démarche qualité fait référence à plusieurs notions qu'il convient de définir.

Qualité

La qualité est un concept subjectif qui tente de décrire dans quelle mesure un produit ou un service est bon ou mauvais. En statistique, la qualité désigne l'adaptation des données à leur utilisation, i.e l'adéquation aux besoins exprimés et aux objectifs visés.

La qualité d'un produit statistique est appréciée à travers six dimensions, à savoir la pertinence, l'exactitude, l'accessibilité, l'actualité, l'intelligibilité et la cohérence.

Bien que le produit statistique soit l'extrait d'un processus, très souvent on ne considère pas la qualité comme une activité unique, mais plutôt comme un ensemble complet d'activités grâce auquel on arrive à un produit acceptable.

Assurance qualité

L'assurance qualité s'entend comme l'ensemble des activités planifiées nécessaires pour s'assurer qu'un produit ou un service statistique convienne à l'utilisation prévue. Elle est caractérisée par l'ensemble des activités de planification, d'orientation et de contrôle destinées à établir, maintenir ou à améliorer la qualité de la production statistique. L'assurance de la qualité s'efforce d'anticiper les problèmes avant qu'ils ne surviennent et vise à assurer la qualité grâce à la prévention. Dans cette perspective, l'assurance qualité déplace essentiellement la qualité en amont.

Dans le contexte des programmes statistiques, l'assurance qualité peut avoir lieu à n'importe quelle étape de la production.

Contrôle qualité

Le Contrôle qualité désigne la vérification de la conformité d'un produit à sa définition ou à ses spécifications. C'est une procédure de réglementation qui consiste à mesurer la performance réelle en matière de qualité à l'aide d'un

échantillon représentatif, comparer la mesure à une norme préétablie et à prendre une décision objective en fonction des mesures.

Dans le contrôle qualité, il est question de réagir aux problèmes observés et en ce sens, la qualité est traitée en aval. Il convient de souligner que les procédures de contrôle de la qualité, ainsi que les stratégies de rétroaction, peuvent empêcher que les problèmes se reproduisent. Par conséquent, bien que les méthodes d'assurance de la qualité et de contrôle de la qualité soient essentiellement de nature corrective, elles peuvent aussi être préventives lorsque la rétroaction est donnée.

Le contrôle qualité est généralement utilisé au stade de la production et s'applique seulement à certaines phases de l'enquête.

Tableau 1: Comparaison entre l'Assurance Qualité et le Contrôle Qualité

Assurance de la qualité	Contrôle de la qualité
Menée au stade de la planification	Mené au stade de la production
Anticipe les problèmes avant qu'ils ne surviennent	Réagit aux problèmes observés
Action préventive	Action corrective et/ou préventive
S'applique à toutes les phases d'un programme statistique	S'applique à certaines phases d'un programme statistique

II. Le Modèle Générique du Processus de Production Statistique

Définition et objectif

Le Modèle Générique du Processus de Production Statistique (MGPPS, en anglais Generic Statistical Business Process Model (GSBPM)) définit et décrit l'ensemble des processus d'activités nécessaires à l'établissement de statistiques officielles. Il définit une architecture commune pour le processus de production statistique, et offre un cadre type et une terminologie harmonisée pour aider les organismes de statistique à évaluer et à moderniser leurs processus de production statistique.

Origine et historique

La première version du GSBPM a été établie par le Groupe de haut niveau sur la modernisation de la production et des services statistiques de la Commission Economique Européenne. Il a été établi initialement sur la base du modèle générique de processus d'activité de Statistics New Zealand, que les experts ont reconnu au cours d'un atelier en 2007 comme bon point de départ pour l'élaboration d'un modèle générique. Ce modèle de base a été complété par les apports d'autres organismes de statistique possédant une expérience reconnue.

La version 4.0 du GSBPM a été publiée en avril 2009, et a depuis été largement adoptée par les organismes de statistique. Ce modèle organisait le processus de production statistique en neuf étapes, à savoir la détermination des besoins, la conception, l'élaboration, la collecte, le traitement, l'analyse, la diffusion, l'archivage et l'évaluation.

La révision dès 2013 de la version 4.0 en vue de l'améliorer a conduit à la publication en 2014 de la version 5.0 (version actuelle). Les principaux changements ont été le passage de neuf à huit étapes (la phase d'archivage a été supprimée, au motif qu'elle se fait à tous les stades du processus), quelques sous processus ont été ajoutés, certains ont été rebaptisés pour plus de clarté, et les descriptifs ont été actualisés et étoffés pour prendre en compte le recours croissant aux autres sources de données (sources administratives, etc.).

Présentation

Les travaux de compréhension et d'adaptation du modèle générique aux réalités de la production statistique au Cameroun, tout en tenant compte des idéaux de qualité définis dans le Cadre National d'Assurance Qualité, ont conduit à s'accorder sur le découpage du processus de production en huit étapes. En effet, bien que l'archivage soit conçu au même titre que les huit étapes, il se fait tout au long du processus. Cependant, des lacunes encore importantes sont observées à ce niveau.

Plus généralement, si on peut noter que la présentation des étapes et des sous processus suit un ordre logique, cela ne correspond pas forcément à un ordre chronologique, bien que certains sous processus soient préalables à d'autres. Il faut surtout garder à l'esprit que le plus grand intérêt de ce découpage réside dans l'ordonnement des activités dans un souci de suivi de la mise en œuvre et d'évaluation de la qualité, et que des activités relatives à des étapes distinctes puissent se réaliser en même temps ou même dans certains cas dans un ordre chronologique inverse.

Préalables

Le modèle générique est un cadre ou un outil technique qui guide les praticiens dans la production de statistiques, et leur permet de s'assurer que toutes les activités techniques essentielles à la production des statistiques de qualité sont efficacement mises en œuvre. A l'issue de chaque activité, sous processus ou étape, un livrable est obtenu, validant la mise en œuvre et servant généralement d'intrant pour l'étape suivante. C'est le tableau de bord du technicien.

On suppose ainsi que tous les préalables, c à d tous les éléments de l'environnement, fondamentaux pour la bonne mise en œuvre du processus mais

qui ne relèvent pas directement des compétences des techniciens, sont satisfaits tout au long du processus. Ces fondamentaux ont trait au cadre législatif et réglementaire, à la budgétisation, la planification stratégique, la disponibilité et la gestion des ressources, etc. Le technicien devrait donc pouvoir bénéficier de toutes les autres compétences nécessaires pour satisfaire aux besoins exprimés en vue de la mise en œuvre efficace du processus de production statistique.

Cependant, même si ces éléments d'accompagnement et de facilitation du processus ne relèvent pas directement de la compétence du technicien, il est de sa responsabilité de s'assurer que les activités destinées à les rendre disponibles évoluent normalement, et que la mise en œuvre du processus ne risque pas d'être entravée. C'est pour cela que toutes les compétences utiles au projet doivent être correctement identifiées et prises en compte dans la définition et la constitution des équipes.

Section 2 : Les lignes directrices pour la qualité des processus

Cette section est organisée en huit sous sections, correspondant aux huit étapes retenues pour le processus de production statistique. Chaque étape est décrite dans une introduction qui présente son objet, son rôle et son importance dans le processus de production, le livrable attendu à l'issue de sa mise en œuvre. Il est aussi évoqué les différents sous processus qui la composent, qui seront eux même présentés dans leurs paragraphes respectifs.

La présentation de chaque sous processus se décline en quatre points :

Portée et objet : qui décrit le sous processus, son rôle et présente les grandes activités qui y sont attendues.

Produit attendu : entendu comme le livrable ou l'output technique (s'il y a lieu) permettant de poursuivre le processus. Bien entendu, en plus de ce livrable technique, chaque sous processus doit être documenté et donner lieu à la production d'un rapport décrivant sa mise en œuvre.

Lignes directrices : Comme annoncé à l'introduction, cette première édition des lignes directrices propose et décrit sommairement les activités élémentaires.

Dimensions de la qualité des données concernées : qui fait le lien entre la qualité de chaque sous processus et la qualité des données, en énonçant lesquelles des six dimensions de la qualité des données le sous processus contribue à atteindre. Ces dimensions sont :

- **La pertinence** : qui renvoie à la mesure dans laquelle les besoins réels des utilisateurs sont satisfaits. Les statistiques seront donc pertinentes si elles éclairent les utilisateurs sur les questions les plus importantes à leurs yeux ;
- **L'exactitude (fiabilité)** : qui fait référence à la mesure dans laquelle l'information décrit correctement le phénomène qu'elle doit mesurer, le degré de fidélité avec la réalité qu'on veut présenter. Le degré d'exactitude découle de la méthodologie mise en œuvre, et est habituellement exprimé en termes d'erreur dans les estimations ;
- **L'actualité (ponctualité)** : qui se rapporte 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. L'actualité constitue aussi élément d'appréciation de la pertinence, et souvent un compromis doit être trouvé entre l'actualité et l'exactitude de l'information ;

- **L'accessibilité** : définie comme la facilité avec laquelle on peut se procurer les données produites. L'accessibilité intègre aussi la facilité avec laquelle on peut constater que l'information existe, le caractère approprié de la présentation de l'information ou du média au moyen duquel on peut avoir accès aux données, et même le coût pour certains utilisateurs ;
- **L'intelligibilité** : renvoie à la disponibilité de renseignements supplémentaires et de métadonnées nécessaires à l'interprétation et à l'utilisation appropriée de ces données. Il s'agit en général de renseignements sur les variables, les classifications et les concepts sous-jacents utilisés, sur les méthodes de collecte et de traitement des données et sur les indicateurs de l'exactitude des données statistiques.
- **La cohérence (comparabilité)** : traduit la mesure dans laquelle les données produites peuvent être jumelées à d'autres renseignements statistiques dans un vaste cadre analytique au fil du temps. L'utilisation de concepts, de classifications et de populations cibles types favorise la cohérence, tout comme l'utilisation de méthodes d'enquêtes communes. La cohérence ne renvoie pas nécessairement à la concordance numérique parfaite.

I. Détermination des besoins

Cette phase constitue le point de départ de tout processus statistique, et s'amorce lorsqu'un besoin de nouvelles statistiques est déterminé ou lorsqu'un examen est entrepris à la suite de rétroaction au sujet de statistiques actuelles. Elle comprend toutes les activités liées à la consultation des utilisateurs en vue d'identifier leurs besoins détaillés en matière de statistique, et de proposer des solutions optimales pour répondre à ceux-ci. Cette phase donne lieu à l'élaboration du document de projet qui précise pourquoi on met en œuvre le projet, en quoi il va consister et comment il va être mené.

I.1. Identification et validation des besoins

Portée et objet : Ce sous-processus consiste à cerner les besoins des utilisateurs en vue de déterminer les statistiques nécessaires pour y répondre. Il peut être déclenché par une nouvelle demande d'informations de la part des principaux utilisateurs, par l'émergence des besoins particuliers de divers groupes cibles, par un changement dans l'environnement social, économique, etc.

La consultation des utilisateurs à ce niveau est primordiale, pour bien comprendre leurs besoins et confirmer en détail les besoins statistiques. Il s'agit non seulement de savoir ce qu'il faut produire mais également quand, comment et, surtout

pourquoi. Pour des opérations statistiques itératives, cette phase consiste principalement à déterminer si les besoins cernés précédemment ont changé

Produit attendu : à l'issue de ce sous processus, l'organisme statistique doit disposer d'un rapport d'analyse des besoins, qui décrit les besoins identifiés et présente les utilisateurs réels ou potentiels des statistiques à produire.

Lignes directrices : l'identification et la validation des besoins passent par 4 activités essentielles :

- Identifier les utilisateurs : par rapport à la thématique et à l'objectif de l'activité, il faut identifier les utilisateurs réels (habituels) ou potentiels ;
- Identifier les besoins d'information internes : il s'agit pour les utilisateurs identifiés sur le plan national de dégager les besoins potentiels. Il peut s'agir par exemple des besoins en termes de suivi des politiques publiques, des besoins du secteur privé, ou de la société civile ;
- Identifier les besoins d'information externes : ceux-ci peuvent découler des orientations des organisations internationales, pour le suivi des programmes de développement auxquels le pays a souscrit ;
- Organiser des concertations avec les utilisateurs : en vue de disposer d'un ensemble cohérent de besoins en statistiques, des concertations (bilatérales ou multilatérales) sont menées avec le maximum d'utilisateurs pour analyser et valider les besoins.

Dimension de la qualité des données: Pertinence

I.2. Identification des produits statistiques

Portée et objet : ce sous processus permet de déterminer les produits statistiques nécessaires pour satisfaire aux besoins validés, et s'entendre avec les utilisateurs sur l'utilité des produits proposés et les mesures de la qualité, en vue de mieux orienter les étapes du projet. Certains éléments tels que la disponibilité des ressources peuvent orienter la prise de décision au cours de ce sous processus.

Produit attendu : il est attendu ici la liste des indicateurs prioritaires et des produits de diffusion des statistiques à produire, qui va constituer un élément du rapport d'analyse des besoins. Ces produits peuvent être un rapport d'analyse, une base de données (statique ou dynamique), un système d'informations, un atlas, une carte, etc.

Lignes directrices : ce sous processus consiste principalement en une concertation avec les utilisateurs pour arrêter la liste des indicateurs prioritaires et des produits statistiques. Elle peut se faire dans le cadre des concertations pour l'identification et

la validation des besoins, en fonction de l'envergure de l'opération ou de la densité des besoins identifiés.

Dimension de la qualité des données concernée: Pertinence

I.3. Détermination des concepts

Portée et objet : ce sous processus permet de préciser les concepts devant être mesurés par le processus de production. Ces précisions sur les concepts sont données par les utilisateurs, et ces concepts sont au besoin alignés aux normes statistiques existantes. Ces normes peuvent être soit les cadres statistiques internationaux, soit ceux reconnus ou déjà adaptés au niveau national. Les précisions ici concernent aussi le choix et la définition des variables à mesurer.

Produit attendu : à l'issue de ce processus, un document définissant l'ensemble des concepts à mesurer ou utiles pour le processus de production doit être disponible.

Lignes directrices : la définition des concepts doit s'inspirer des documents de normalisation disponibles au niveau national (manuel des concepts en l'occurrence), et des cadres de normalisation au niveau international, qu'il peut être utile de contextualiser. Une concertation avec les experts sectoriels est indispensable à cette étape.

Dimensions de la qualité des données concernées: Pertinence, cohérence (comparabilité)

I.4. Vérification de la disponibilité des données

Portée et objet : Lorsque les besoins en information sont clairement identifiés, il convient de vérifier si les sources de données existantes peuvent permettre d'y répondre, même partiellement. Les sources à examiner peuvent être administratives, provenir d'autres opérations statistiques déjà menées, ou même d'opérations en cours ou déjà envisagées. Cet exercice inclut l'évaluation des conditions dans lesquelles ces sources peuvent être capitalisées pour répondre aux besoins actuels, plus précisément si elles peuvent être traitées et utilisées à des fins statistiques. Il convient aussi d'identifier les restrictions éventuelles, et d'évaluer minutieusement les méthodologies en relation avec les objectifs de production déjà établis. Dans un contexte de contraintes budgétaires et le souci de limiter le fardeau du répondant, cette évaluation permet d'identifier le gap à combler dans les données et d'élaborer une stratégie pour combler ces lacunes, dont la mise en place d'une nouvelle opération. Elle peut même conduire à une proposition de changements dans les systèmes de collecte existants (surtout les sources

administratives) en vue de mieux capitaliser ces sources de données pour la satisfaction de besoins identifiés.

Produit attendu : à l'issue de ce sous processus, un rapport d'analyse documentaire est nécessaire pour présenter l'état des lieux des sources de données potentielles par rapport aux besoins (en tenant compte de l'actualité).

Lignes directrices :

- Identifier les sources administratives, par des consultations avec les producteurs du SNIS et fournisseurs du secteur privé, etc. : les données administratives sont souvent le fait du fonctionnement d'organismes et de la mise en œuvre de programmes gouvernementaux. On peut identifier toutes les structures dont les résultats des programmes ou du fonctionnement peuvent rentrer dans le champ des besoins détectés ou l'inverse ;
- Examiner les sources (données) administratives, en vue de comprendre les données (qui peuvent être documentées ou pas), d'en évaluer la pertinence et l'opportunité (potentiel par rapport aux besoins à satisfaire) et les conditions d'utilisation, d'examiner les procédures de collecte et de traitement, de déceler tous les éléments qui peuvent entraver la fiabilité, etc. une collaboration avec les responsables des systèmes administratifs est indispensable ;
- Vérifier la disponibilité d'autres sources (autres opérations statistiques) et les évaluer en ce qui concerne les objectifs, la méthodologie, la couverture, les concepts, les classifications, les dates de référence, les normes et Dimensions de la qualité, etc. ;
- Vérifier la disponibilité des bases de données, pouvant être exploitées (ou combinées) pour produire les statistiques requises ;
- Faire la Synthèse des sources identifiées, en soulignant les éventuelles insuffisances et les questions à régler en vue de les améliorer et les rendre plus utiles.

Dimensions de la qualité des données concernées: Pertinence, actualité (ponctualité)

I.5. Etude de faisabilité

Portée et objet : les précédents sous processus de cette phase ont permis de cerner les besoins et d'explorer les différentes sources de données compatibles avec ces besoins. Il s'agit maintenant de documenter les résultats sous forme d'une analyse coût opportunité pour soutenir et de justifier la solution qui se dégage en vue de combler les attentes des utilisateurs. Celle-ci peut consister à mettre en œuvre un nouveau processus de production statistique, à modifier un processus existant, à

exploiter des sources de données existantes, ou une combinaison de toutes ces options. La documentation des résultats précédents doit donc permettre de comprendre la solution adoptée en vue d'une mise en œuvre optimale. Il convient de préciser que quelle que soit l'option retenue, cela ne remet pas en question la validité des besoins identifiés.

Produit attendu : A l'issue de ce sous processus et de cette première étape du processus de production, le document de projet doit être élaboré, intégrant une analyse coût (en ressources humaines, matérielles, financières) opportunité (pertinence, actualité, cohérence, exactitude) de l'option choisie pour répondre aux besoins identifiés.

Lignes directrices :

- Faire la synthèse des sources par rapport aux besoins, par l'élaboration d'une table de correspondance présentant pour chaque besoin dans quelle mesure il peut être satisfait par les sources disponibles ;
- Evaluer les actions à entreprendre pour répondre aux besoins : il s'agit, au regard du potentiel des sources existantes, d'identifier toutes les actions nécessaires (sur le plan scientifique et technique) pour combler le gap (capitaliser les sources existantes, améliorer un processus de production, mettre en œuvre un nouveau processus, etc.) ;
- Procéder à une analyse coût avantage, notamment l'évaluation d'une part des ressources à mobiliser et d'autre part du gain en qualité de chaque option qui se dégage ;
- Evaluer la possibilité d'utiliser les proxys, pour prendre en compte les besoins identifiés, pertinents, mais pour lesquels aucune source de données n'est disponible, et il n'existe pas de méthodologie connue. Les proxys peuvent aussi être justifiés par les coûts élevés de production d'une information ;
- Evaluer les contraintes externes de mise en œuvre des options envisagées ;
- Présenter le choix stratégique (solution proposée), notamment l'option à choisir, qui doit être la plus pertinente au regard de l'analyse coût avantage. Cette solution doit être assortie d'une présentation détaillée de la manière dont le processus de production statistique sera mis au point pour produire les statistiques nouvelles ou révisées requises.

Dimensions de la qualité des données concernées: pertinence, actualité (ponctualité), cohérence (comparabilité), exactitude (fiabilité)

II. Conception

La phase de conception est celle où tout le processus de production est pensé. Il s'agit de concevoir, de définir et de développer tous les éléments méthodologiques, les produits et les services nécessaires à la production des statistiques recherchées. Tous les outils nécessaires à chacune des étapes du processus de production sont conçus à ce niveau, de la collecte à l'évaluation, de manière à les mettre en exécution le moment venu. Cette phase débute par la conception des produits puis des variables, et se poursuit par la conception de tous les éléments constitutifs des différentes étapes : collecte, traitement, analyse, archivage, diffusion et évaluation. Ces sous processus sont généralement séquentiels, mais peuvent aussi se dérouler en parallèle et peuvent être itératifs.

Les activités de conception s'appuient majoritairement sur les normes internationales et nationales afin de réduire le coût (temps, ressources) du processus de conception et de renforcer la comparabilité et la possibilité d'utilisation des produits.

Les différents produits de cette phase constituent ce qui est généralement appelé « document de méthodologie », qui peut être décliné en plusieurs sections correspondant chacune à la mise en œuvre d'un ou plusieurs sous processus.

II.1. Conception des produits

Portée et objet : à ce niveau, il est question de concevoir de manière détaillée les produits statistiques en vue de préparer le processus de production et de diffusion. Les produits sont conçus et définis de manière à répondre aux besoins identifiés, et à satisfaire les diverses catégories d'utilisateurs identifiés.

Produit attendu : Document décrivant chacun des produits en liaison avec ses utilisateurs, à fournir à l'issue du processus de production

Lignes directrices : la conception des produits consiste à :

- Définir les formats de diffusion : en fonction des utilisateurs et des utilisations envisagées, il peut s'agir de rapports, Bases de Données, atlas, cartes, tableaux de bord, indicateurs clés : penser par exemple à la manière de présenter les indicateurs dans un site web, etc.;
- Concevoir la tabulation des indicateurs, qui consiste à définir la manière de présenter les indicateurs pour alimenter les différents produits, identifier les variables clés, les variables de classification, le niveau de désagrégation, etc. ;
- Concevoir les méthodes de contrôle de divulgation des données individuelles ;
- Concevoir le calendrier d'exécution.

Dimension de la qualité des données concernée: Pertinence

II.2. Conception et description des variables

Portée et objet : les résultats statistiques sont généralement des indicateurs ou des paramètres calculés à partir de variables recueillies lors de la collecte. Il est question ici d'identifier, de définir et de décrire ces variables. Celles-ci peuvent être soit collectées directement auprès des unités cibles, soit dérivées à partir des variables collectées. La description porte sur les éléments tels que la nature, l'unité, la population cible, le niveau de désagrégation, ainsi que les classifications statistiques qui seront utilisées.

Produit attendu : Document décrivant toutes les variables (Dictionnaire des variables) utilisées dans le processus de production, y compris les variables de la base de sondage et les variables de qualité

Lignes directrices : permettent de :

- Identifier et décrire les unités statistiques, sur lesquelles va porter la collecte des données ;
- Identifier ou élaborer (au besoin) les classifications à utiliser : On s'attache en principe à respecter les normes nationales et internationales existantes tant que c'est possible, mais pour les besoins de l'étude on peut avoir besoin d'adapter, ou d'élaborer de nouvelles classifications, tout en essayant au maximum de se conformer aux standards existant ou aux autres opérations pour des besoins de comparaison ;
- Décrire les variables, en précisant pour chacune d'elles les éléments caractéristiques tels que sa nature, l'échelle de mesure, les modalités, l'unité de mesure, etc ;
- Décrire les agrégats de la population : la description des groupes cible, l'élaboration des formules d'agrégation, la définition des conditions d'agrégation, etc.) ;
- Définir les variables dérivées, issues de la combinaison d'autres variables : (regroupement des variables/modalités, méthode de création de variables, conditions et ordre de dérivation des variables, etc.).

Dimensions de la qualité des données concernées: Pertinence, exactitude (fiabilité)

II.3. Conception de la collecte

Portée et objet : au cours de ce sous-processus, il est question de déterminer les méthodes et instruments les plus appropriés. Il s'agit en l'occurrence du choix du

mode (papier ou assisté par ordinateur) et de la méthode (auto administré, entrevue directe, à distance (téléphone, internet, etc.) etc.), de l'identification des questions pertinentes pour renseigner les variables d'intérêt, de la définition de la structure des questionnaires et autres outils de collecte. Ce sous processus comprend aussi la conception de tous les éléments de mise en œuvre, d'accompagnement et de facilitation de la collecte tels que les ressources, les aspects linguistiques et éthiques, les accords formels en vue de la fourniture des données, la sensibilisation, etc.

Produit attendu : les outils, instruments et infrastructures pour la collecte sont conçus (structure détaillée du questionnaire, identification et organisation des ressources, etc.).

Lignes directrices : la conception de la collecte consiste généralement à :

- Concevoir le plan/stratégie de collecte dont les éléments sont entre autres l'organisation de la collecte, la définition des éléments de sensibilisation et de promotion de la collecte, la conception du déploiement, etc ;
- Choisir le mode et la méthode de collecte : papier ou assisté par ordinateur, auto administré, entrevue directe, à distance, en tenant compte de toutes les conditions et contraintes liées à la mise en œuvre de chaque méthode ;
- Concevoir les outils de collecte, en définissant la structure et le contenu des questionnaires, fiches, manuels, documents de classification, supports de collecte, fiches d'instruction, masque de saisie, etc. ;
- Concevoir le soutien matériel à la collecte, notamment la logistique, et la gestion de la logistique ;
- Concevoir les applications de gestion de la collecte pour le suivi du bon déroulement de la collecte par exemple, la définition des contrôles primaires pendant la collecte ;
- Concevoir les ressources humaines pour la collecte, qui intègre la définition des critères d'affectation des ressources humaines (qui en plus de la charge de travail peuvent être les distances, l'enclavement, la langue parlée, le genre, la sécurité). Il s'agit aussi de définir le plan de formation, d'évaluation et sélection des ressources humaines (critères de sélection, profils des agents, experts et formateurs, etc.), sans oublier le recrutement pour la collecte dans ses aspects administratifs, en pensant à établir pour les services compétents la fiche technique pour le recrutement ;
- Déterminer les infrastructures pour la collecte : serveurs, tablettes, locaux, etc. Il convient de penser aux infrastructures adéquates aux nouvelles sources de

données émergentes telles que les big data, les données en ligne, les données satellitaires, etc. ;

- Choisir le mode de saisie de données (double saisie, saisie réseau, outil de saisie, etc.) ;
- Concevoir les autres éléments de la collecte tels que les aspects linguistiques, les aspects éthiques le cas échéant, les éléments de collaboration qui peuvent être des conventions avec les médias et des fournisseurs pour faciliter la collecte ;
- Concevoir l'utilisation des sources de données émergentes (big data) : données en ligne, données satellitaires, etc.

Dimensions de la qualité des données concernée: Exactitude (fiabilité)

II.4. Conception de la base de sondage et de l'échantillon

Portée et objet : la conception de la base de sondage et de l'échantillon consiste à définir et à décrire la base de sondage ainsi que la population d'intérêt, i.e l'ensemble des unités sur lesquelles on désire obtenir des renseignements et pour lesquelles il faut produire des estimations. Il s'agit de définir sa constitution (variables d'intérêt), et de préciser comment différentes sources peuvent être combinées au besoin pour constituer la base de sondage. Ces sources peuvent être des registres administratifs et statistiques, les données de recensements et les informations tirées d'autres enquêtes par sondage. Il s'agit aussi de déterminer et cerner la population d'intérêt, d'établir un plan de sondage pour la détermination de l'échantillon, et de déterminer la méthode et les critères d'échantillonnage les plus appropriés.

NB : Cette étape ne concerne pas les recensements

Produit attendu : méthodologie de construction de la base de sondage, ainsi que sa structure

Lignes directrices :

- Concevoir la base de sondage, par l'identification et la description des variables d'intérêt qui la composent, la précision des éléments de localisation au besoin (cartes, coordonnées géo, etc.). Il peut aussi s'agir d'une mise à jour de la base de sondage, si celle-ci est déjà disponible ;
- Identifier les sources de données pour les variables de la base de sondage, y compris la définition des traitements et combinaisons nécessaires, s'étant au préalable rassuré qu'aucune base de sondage ne peut assurer, à elle seule, une couverture suffisante ;

- Concevoir la méthode d'échantillonnage : en fonction des contraintes opérationnelles, de la base de sondage, des objectifs visés et d'autres facteurs tels que degré de précision et de détail visé pour les informations à livrer, la disponibilité de variables auxiliaires permettant la stratification et la sélection de l'échantillon, les méthodes d'estimation qui seront appliquées, le budget alloué, etc. ;
- Déterminer la taille et la répartition de l'échantillon : en tenant compte de la précision visée, des taux de réponse attendus, etc.
- Concevoir l'extraction (tirage) de l'échantillon : il s'agit généralement de définir l'algorithme de mise en œuvre de la méthode d'échantillonnage.

Dimension de la qualité des données concernée: Exactitude (fiabilité)

II.5. Conception du traitement et de l'analyse

Portée et objet : l'objet de ce sous processus est de définir les méthodes qui seront utilisées pour le traitement statistique des données collectées, puis leur analyse. Il s'agit précisément de définir tous les méthodes et outils qui vont servir à vérifier, nettoyer, transformer les données d'entrée. On y retrouve principalement la définition des routines de codage, d'édition, d'imputation, d'estimation, d'intégration, de validation et de finalisation des ensembles de données.

Produit attendu : Méthodes et outils utilisés dans les processus de traitement et d'analyse définis.

Lignes directrices :

- Concevoir la codification, qui consiste à identifier les variables à codifier, à apprêter les nomenclatures. Il peut s'agir d'un pré codage pour les questions fermées, ou d'un codage après la collecte. Cette activité inclut la conception de l'évaluation du codage, pour s'assurer de l'exactitude des données codées ;
- Concevoir la détection des erreurs : il s'agit de définir tous les outils nécessaires pour mettre en évidence les enregistrements des données qui sont susceptibles de contenir des erreurs. La vérification englobe une vaste gamme d'activités et se situe à différents niveaux du processus de collecte et de traitement. Elle intègre les vérifications des intervieweurs sur le terrain et les avertissements générés par ordinateur au moment de la collecte ou de la saisie des données, etc. ;
- Concevoir la gestion des erreurs : principes et les méthodes de correction, méthodes et principes d'imputations, types de correction à différents niveaux, etc.) ;

- Concevoir la validation de l'échantillon corrigé (critères de validation de l'échantillon corrigé) ;
- Concevoir l'intégration des fichiers : Il peut s'agir des fichiers de même source pour lesquels il convient de définir les clés d'appariement, ou des fichiers de sources différentes qui nécessite l'identification des variables pour les clés d'appariement des différentes bases de données ;
- Concevoir le calcul des poids : probabilités d'inclusion ou poids de sondage, poids corrigés de la non réponse, choix des variables auxiliaires pour le calage qui doivent être théoriquement corrélées à la variable d'intérêt, poids calés, etc. ;
- Concevoir les estimateurs, en intégrant le cas échéant la mise à jour des séries chronologiques, les ajustements saisonniers et les effets calendaires ;
- Concevoir la correction des estimateurs, en décrivant comment prendre en compte les éléments tels que les non réponses ; les changements constatés de la base de sondage, les écarts de taille constatés entre de l'échantillon théorique et l'échantillon empirique, etc ;
- Concevoir les estimateurs de la variance ;
- Concevoir le contrôle de la confidentialité statistique (Sécurisation des données lors des traitements, des analyses et de la diffusion) ;
- Concevoir l'analyse, qui peut consister à décrire exactement les éléments de contexte à capitaliser, l'association des données des cycles précédents, etc.

Dimensions de la qualité des données concernées: Exactitude (fiabilité), cohérence (comparabilité)

II.6. Conception du système de production

Portée et objet : à ce niveau, il est question de définir le déroulement des opérations, de la collecte à la diffusion des données, en veillant à ce qu'elles s'enchaînent de manière efficace, sans lacunes ni redondances. C'est à cette étape qu'il convient de définir et répartir les rôles et responsabilités des intervenants, ainsi que les interactions entre eux et avec le système.

Produit attendu : Architecture de l'équipe, des activités, et calendrier de travail

Lignes directrices :

- Concevoir l'architecture de l'équipe, qui consiste à identifier les différents niveaux d'intervention et les responsabilités, les compétences requises, les interactions, etc.
- Concevoir le cahier de charges des intervenants, par la définition des activités à mener à chaque niveau, les outputs attendus,

- Concevoir l'enchaînement des activités : interdépendances, ordre de mise en œuvre ;
- Concevoir le dispositif de suivi du bon déroulement des opérations ;
- Elaborer le calendrier de travail, en tenant compte des spécificités de l'opération, une collecte de données dans un secteur ou pour un thème précis pouvant être mieux adaptée à une période qu'à une autre, par exemple (éducation, dépenses de consommation, prix, etc.). A ce niveau le calendrier est prévisionnel, et sera ajusté au fur et à mesure, avec l'évolution des activités. Il est aussi utile d'évaluer les risques encourus si le calendrier n'est pas respecté, et de concevoir les stratégies de gestion des risques.

Dimensions de la qualité des données concernées : exactitude (fiabilité), actualité (ponctualité)

II.7. Conception de la diffusion et de la dissémination

Portée et objet : l'objet de ce sous processus est de penser la diffusion des données une fois produites. Il s'agit d'identifier, de définir et de décrire tous les outils nécessaires à la diffusion, mais aussi à la communication sur les données, de façon à optimiser leur utilisation.

Produit attendu : Plan de diffusion des données produites

Lignes directrices :

- Concevoir les métadonnées : identifier, décrire et présenter les informations utiles (sur les données) devant accompagner la diffusion des données ;
- Concevoir le fichier de microdonnées de diffusion : si les microdonnées sont destinées à être diffusées, il s'agit de définir la structure et le contenu du fichier, en veillant à la prise en compte du secret statistique, à la gestion des données sensibles (sécurité et défense, santé publique, etc.), à la définition des supports de diffusion, etc. ;
- Concevoir le calendrier de diffusion ;
- Concevoir le catalogue des produits (maquette du catalogue en relation avec les produits à diffuser, liste des produits et description du contenu, langues de diffusion, dates de publication, dépliants)
- Concevoir la dissémination (cibles, plan de dissémination), par la définition des objectifs de diffusion et de dissémination des différents produits (données, microdonnées, métadonnées, etc.) ;
- Définir les supports de diffusion des produits ;
- Concevoir la communication sur les produits : Identifier et décrire tous les canaux et actions de promotion des produits.

Dimensions de la qualité des données concernées: accessibilité, intelligibilité

II.8. Conception de l'archivage

Portée et objet : l'objet de ce sous processus est d'organiser l'archivage de tout le processus de production, de la définition des besoins à l'évaluation. Il s'agit de définir les règles d'archivage, tant pour les produits physiques que numériques obtenus à chaque niveau de la mise en œuvre du processus, que ce soit les données, les rapports, les métadonnées, etc. L'intérêt de l'archivage est de constituer une mémoire pour le processus de production, afin d'assurer la sécurité des données, la pérennité et la traçabilité du processus.

Produit attendu : Architecture et plan d'archivage.

Lignes directrices :

- Définir les règles de gestion des dépôts d'archives : ces dépôts peuvent être des bases de données ou des emplacements physiques où des copies de données ou de métadonnées sont stockées. Les règles de gestion de ces dépôts d'archives doivent être définies de manière à garantir la récupération des données individuelles ou des ensembles de métadonnées, garantir l'intégrité des données archivées et des métadonnées, etc ;
- Définir les règles de préservation des données et métadonnées associées, qui consiste par exemple à définir les données et métadonnées à archiver (liste des produits d'archivage), définir les procédures de vérification de la qualité de l'archivage, etc. ;
- Définir les règles d'élimination des données et métadonnées associées ;
- Définir les responsabilités dans le processus d'archivage.

Dimensions de la qualité des données concernées: accessibilité, intelligibilité

II.9. Conception de l'évaluation

Portée et objet : l'objet de ce sous processus est de définir les outils devant permettre d'évaluer la qualité de la mise en œuvre du processus de production statistique. Son intérêt réside surtout dans l'amélioration de la qualité des opérations futures. Il s'agit d'identifier les normes et de définir les critères d'évaluation de la qualité de chaque étape du processus de production.

Produit attendu : document d'assurance qualité (normes et critères de validation).

Lignes directrices : la conception de l'évaluation d'un processus de production statistique consiste à :

- Définir les critères d'évaluation et de validation de la qualité de chaque niveau du processus ;
- Identifier (ou élaborer) les normes associées ;
- Définir les sources de vérification.

Dimensions de la qualité des données concernées: toutes les dimensions.

III. Elaboration

L'étape d'élaboration est celle pendant laquelle tous les outils qui ont été identifiés et décrits à la phase de conception sont réellement produits. A la phase de conception, les produits ainsi que leur contenu et leur forme sont définis de manière schématique et détaillée et consignés dans le document de méthodologie. La phase d'élaboration permet d' « élaborer » et de mettre à l'essai le produit retenu jusqu'à ce qu'il soit prêt à être utilisé dans des conditions réelles. La différence entre la phase de conception et la phase d'élaboration est similaire à celle entre l'algorithme et le programme, le premier étant le concept et le second l'implémentation du concept.

Il est parfois difficile dans la pratique de marquer la limite réelle entre les activités de conception et d'élaboration, dans la mesure où les activités de conception n'aboutissent pas souvent à des produits palpables. Il est cependant primordial de bien les dissocier et de s'assurer de bien documenter chaque phase et d'archiver leurs produits, dans la mesure où l'équipe de conception peut ne pas être la même que celle chargée de l'élaboration, ou même que la détection et la correction des erreurs dans la mise en œuvre du processus se trouvera facilitée si chaque étape est bien menée et bien documentée.

Cette phase comporte sept sous-processus généralement séquentiels mais qui peuvent aussi se dérouler en parallèle ou de manière itérative.

III.1 Elaboration de l'instrument de collecte

Portée et objet : ce sous-processus permet de fournir les instruments et éléments qui seront utilisés durant la phase de la collecte. Il peut s'agir des questionnaires sur support papier, sur support électronique ou sur Internet, des plates-formes SDMX, des routines d'extraction des données, des métadonnées pour les questionnaires et supports de collecte, des éléments d'organisation, de suivi et contrôle de la collecte.

Produits attendus : tous les outils de collecte, à savoir les questionnaires, supports, outils de recrutement du personnel, etc.

Lignes directrices : les activités consistent à implémenter ce qui a été conçu précédemment à savoir:

- Elaborer les questionnaires, des métadonnées pour les questionnaires et des supports de collecte;
- Elaborer les paramètres de suivi et contrôle de la collecte de données (paradonnées) ;
- Elaborer le plan/stratégie de collecte (organisation de la collecte, sensibilisation, déploiement, etc.) ;
- Elaborer la méthodologie de collecte ;
- Elaborer la fiche de soutien matériel à la collecte (logistique, gestion de la logistique) ;
- Elaborer la fiche technique pour le recrutement du personnel externe (y c les quantités) ;
- Elaborer le plan d'affectation des ressources humaines(en fonction de la charge de travail, des distances, de l'enclavement, la langue parlée, le genre, la sécurité, etc.) ;
- Elaborer le plan de formation, d'évaluation et sélection des ressources humaines (outils de formation, guide, critères de sélection, profils des agents, experts et formateurs, etc.) ;
- Développer / mettre en place une infrastructure pour la collecte (fiches techniques architecture et fonctionnement de l'infrastructure de transfert des données, caractéristiques des locaux pour le serveur) ;
- Elaborer les accords de collaboration (conventions avec les médias, et autres accords pour faciliter la collecte) ;
- Développer les actions pour la promotion de la collecte (sensibilisation, communication,);
- Développer le cas échéant des aspects éthiques ;
- Elaborer les contrôles primaires (pendant la collecte)

Dimensions de la qualité des données concernées: exactitude (fiabilité), pertinence

III.2 Elaboration du processus de production

Portée et objet : le processus de production des données concerne toutes les activités qui utilisent des données collectées pour produire un produit. Il comprend les activités de constitution ou de mise à jour de la base de sondage (puisque'elle intervient à la suite d'une collecte effectuée à cet effet), les activités de traitement et les activités d'analyse. Les aspects organisationnels, de suivi, d'évaluation ou de contrôle de ces activités, de même que les activités de collecte, de diffusion, d'évaluation et d'archivage ne font pas partie. Elaborer le processus de production

consiste à préparer et rendre prêt un ensemble d'éléments, conformément aux orientations du document de conception, qui seront tout simplement exécutés lors de ce processus. Ces éléments peuvent être sous forme d'instructions dans des documents papier ou sous forme d'applications informatiques.

Produit attendu : Manuels et programmes de traitement et d'analyse

Lignes directrices :

- Elaborer le manuel ou l'application informatique de construction ou de mise à jour de la base de sondage ;
- Elaborer la méthode d'échantillonnage et le programme de tirage de l'échantillon ;
- Elaborer le guide ou le programme de codification (il s'agit de décrire comment codifier les questions ouvertes. La codification des questions fermées est déjà prise en compte dans les manuels de collecte) ;
- Elaborer le manuel et les programmes de spécification des erreurs (permet de savoir qu'est ce qui peut être qualifié d'erreur) ;
- Elaborer le manuel ou le programme de gestion des erreurs (comment s'organiser pour repérer et corriger les erreurs) ;
- Elaborer la méthodologie de validation de l'échantillon corrigé (peut être accompagné par un programme informatique) ;
- Elaborer les méthodologies d'intégration des fichiers de la même source ;
- Elaborer les méthodologies d'intégration des fichiers de différentes sources (conditions préalables nécessaires, construction des clés d'appariement) ;
- Elaborer la méthodologie et le programme de calcul des poids ;
- Elaborer la méthodologie et le programme de calcul des estimateurs ;
- Elaborer la méthodologie (formule) et le programme la correction des estimateurs ;
- Elaborer la méthodologie et le programme du calcul des estimateurs de la variance ;
- Elaborer la méthodologie et le programme de contrôle de la confidentialité statistique (identification et contrôle de tous les risques (variables ou croisement de variables, appariements) ;
- Elaborer le guide d'analyse.

Dimensions de la qualité des données concernées: exactitude (fiabilité)

III.3 Elaboration du processus de diffusion

Portée et objet : l'objectif de la diffusion est de mettre les produits à la disposition des utilisateurs. A chaque groupe d'utilisateurs, doit correspondre une manière de

présenter l'information, de communiquer et le moment de le faire. Le processus de diffusion concerne donc les maquettes des produits de diffusion ou de communication (éléments de présentation des résultats où les informations ne sont pas encore chargées). L'élaboration du processus de diffusion concerne toutes les activités de description détaillée des produits de diffusion ou de communication et du calendrier de diffusion.

Produit attendu : Maquettes et plan de diffusion

Lignes directrices :

- Elaboration des produits de diffusion ;
- Maquette du rapport (structure, tabulation (tel que pensé à la conception), et tous les éléments constitutifs du rapport, y c les annexes) : le rapport doit être formaté et n'attendre que les données et les commentaires.
- Développement de la BD, tableau de bord, tous les outils de présentation des indicateurs sur site web ;
- Elaboration (description, présentation) des métadonnées des produits de diffusion ;
- Construction du fichier de microdonnées de diffusion (prise en compte du secret stat, supports de diffusion) ;
- Elaboration du calendrier de diffusion ;
- Elaboration du catalogue des produits (maquette du catalogue en relation avec les produits à diffuser, liste des produits et description du contenu, langues de diffusion, dates de publication, dépliants (à finaliser avec les images des couvertures définitives des produits) ;
- Développement des outils de dissémination (cibles, plan de dissémination) ;
- Développement des outils de communication sur les produits (points de presse, communiqués de presse : fournir tous les éléments techniques aux services compétents).

Dimensions de la qualité des données concernées: accessibilité (et clarté)

III.4 Configuration du déroulement des activités

Portée et objet : la configuration du déroulement des travaux ne concerne pas les processus de production ou de diffusion mais concerne la description des activités d'organisation, de suivi et de déroulement des travaux de la collecte à la diffusion des données.

Produit attendu : Plan de configuration du déroulement des activités

Lignes directrices :

- Elaborer le plan de formation des ressources humaines ;
- Elaborer le plan de désignation des intervenants ;
- Elaborer le cahier de charges des intervenants (les responsabilités, les interactions entre les intervenants, ...) ;
- Décrire les activités et organiser leur enchaînement (comment les activités vont s'enchaîner et s'intégrer : collecte, traitement, diffusion) ;
- Elaborer le dispositif de suivi du déroulement des activités ;
- Elaborer le calendrier de travail.

Dimension de la qualité des données concernées: exactitude (et fiabilité)

III.5 Mise à l'essai du système de production

Portée et objet : le système de production concerne tous les outils qui accompagnent les processus de production et de diffusion. Ce sous-processus permet de s'assurer que chaque élément du système de production fonctionne dans la pratique. Dans la pratique, un certain nombre d'activités prévues à ce niveau peuvent logiquement être reliées au sous processus 3.2 (élaboration du processus de production).

Produit attendu : outils de production testés et approuvés

Lignes directrices :

- Tester les outils de sélection de l'échantillon et de construction de la base de sondage ;
- Tester les outils de collecte (sur un certain nombre d'unités, avant l'enquête pilote) ;
- Tester les outils de traitement ;
- Tester les outils de diffusion.

Dimensions de la qualité des données concernées: exactitude (fiabilité), pertinence, cohérence (comparabilité).

III.6 Mise à l'essai du processus de production statistique

Portée et objet : ce sous-processus permet de s'assurer que le processus de production et les composantes des programmes informatiques associés fonctionnent dans la pratique. Il s'agit en pratique de réaliser un essai sur le terrain ou un essai pilote du processus de production statistique. Il prévoit en général une collecte de données à petite échelle, puis le traitement et l'analyse des données recueillies, et ce, afin de s'assurer que le cycle statistique fonctionne comme prévu. Après la phase expérimentale, il peut être nécessaire d'ajuster les instruments, systèmes ou composantes. En fonction de l'envergure de l'opération, plusieurs

itérations peuvent être nécessaires avant que le processus ne fonctionne de manière satisfaisante.

Produit attendu : processus de production testé, ajusté et validé

Lignes directrices :

- Réaliser l'enquête pilote : L'enquête pilote doit permettre de tester les instruments de collecte et toutes les étapes du processus de production jusqu'à la production des indicateurs, afin de s'assurer que tout fonctionne comme prévu ;
- Tester conjointement les composantes des programmes informatiques associés au processus de production.

Dimensions de la qualité des données concernées: exactitude (fiabilité), pertinence, cohérence (comparabilité).

III.7 Mise au point finale du système et du processus de production

Portée et objet : ce sous-processus permet de s'assurer que tous les éléments nécessaires à la collecte sont prêts et documentés.

Produit attendu : processus de production finalisé et documenté.

Lignes directrices :

- Effectuer la formation des ressources humaines ;
- Effectuer les ajustements du système de production ;
- Effectuer les ajustements du processus de production ;
- Documenter et archiver tous les résultats de ce sous-processus.

Dimensions de la qualité des données concernées: exactitude (fiabilité), pertinence, cohérence (comparabilité).

IV. Collecte

Cette phase comprend la collecte ou le rassemblement de toute l'information nécessaire (données et métadonnées), ainsi que 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. Dans le cas des produits statistiques qui sont produits sur une base régulière, cette phase a lieu à chaque itération.

La phase de la collecte comprend quatre sous-processus qui se déroulent généralement en ordre séquentiel, mais pouvant aussi se dérouler parallèlement et être interactifs.

IV.1. Création/mise à jour de la base de sondage et sélection de l'échantillon

Portée et objet : ce sous-processus permet d'établir la base de sondage et de sélectionner l'échantillon sur lequel portera la collecte, tel que précisé au sous-processus 2.4 Conception de la base de sondage et de l'échantillon. Il comprend également la coordination des échantillons entre occurrences du même processus de production statistique (par exemple pour gérer le chevauchement ou la rotation), ainsi qu'entre différents processus utilisant un même plan de sondage ou registre (pour gérer le chevauchement ou pour répartir le fardeau de réponse par exemple). Ce sous-processus intègre l'assurance de la qualité ainsi que l'approbation de la base de sondage et de l'échantillon sélectionné.

Produit attendu : il est attendu de ce processus un échantillon sélectionné.

Lignes directrices : ce sous-processus met en action le plan de sondage retenu à la phase de conception. Les différentes tâches sont les suivantes :

- Collecter toutes les bases de sondages identifiées comme nécessaires pour l'échantillonnage ;
- Evaluer la qualité de chacune de ces bases de sondage, et ne retenir que celles fournissant de bonnes garanties de qualité et de pertinence (présence de toutes les variables retenues pour la base de sondage, couverture satisfaisante) ;
- Procéder aux traitements (suppression de doublons) et à la mise à jour si nécessaire (par exemple avec des opérations de cartographie) des bases retenues ;
- Apparier les bases de sondage (en s'assurant de la qualité des variables d'appariement notamment), et traiter la base de sondage obtenue ;
- Evaluer la base de sondage obtenue (couverture, représentativité de la population) et répéter si nécessaire les étapes jusqu'à l'obtention d'une base de sondage satisfaisante ;
- Tirer l'échantillon selon le plan de sondage retenu à la phase de conception ;
- Valider l'échantillon, en tenant compte des unités sélectionnées lors des éditions précédentes ou lors d'autres enquêtes dans l'idée d'une meilleure coordination des enquêtes et d'une meilleure gestion des relations avec les répondants (assurer une rotation dans l'échantillon pour diminuer le fardeau des répondants) ;
- Ajuster l'échantillon jusqu'à l'obtention d'un échantillon satisfaisant ;

- Documenter toutes les actions du sous-processus.

Dimension de la qualité des données concernée: Exactitude (et fiabilité).

IV.2. Préparation de la collecte

Portée et objet : ce sous processus permet de veiller à ce que les ressources humaines, les processus et les technologies soient prêts à procéder à la collecte des données et des métadonnées, selon tous les modes, tels que conçus. Les activités de ce sous-processus interviennent à différents moments du processus de production, puisqu'il comprend les activités d'élaboration de la stratégie de collecte, de planification et de formation.

Produit attendu : il est attendu de ce sous-processus un ensemble de résultats, dont les plus importants sont :

- Le personnel de collecte est recruté;
- Le matériel de collecte est effectivement acquis et disponible ;
- Les ressources matérielles (y compris logistiques et logicielles) et financières pour la collecte sont effectivement mobilisées et disponibles ;
- L'infrastructure de collecte est installée et fonctionnelle ;
- Les contrats sont signés avec toutes les institutions intervenant dans la collecte ;
- Le plan de déploiement disponible ;
- Une bonne communication a été faite autour de la collecte.

Lignes directrices : les activités à mettre en œuvre ici doivent s'inscrire en droite ligne des prescriptions de la stratégie de collecte élaborée lors de la phase de conception. Il s'agit :

- De former et de recruter personnel de collecte (de codification et de saisie éventuellement) : tous les intervenants au niveau de la collecte sont recrutés selon les procédures réglementaires en vigueur. Il convient de s'assurer de déclencher la procédure suffisamment à temps de manière à ce que les processus de recrutement et de formation de ce personnel n'induisent pas de retard sur la mise en œuvre du projet. Il faut s'assurer de garantir les éléments tels que l'assurance, et toutes les dispositions contractuelles et réglementaires à prendre lors du recrutement du personnel. Pour certaines opérations, il peut être nécessaire de recruter un personnel spécifique pour appuyer la formation (personnel de santé par exemple) ;
- D'acquérir le matériel : tout le matériel (véhicule, ordinateurs, tablettes, papier, sacs, parapluies, stylos, markers...) de l'enquête est acquis selon les procédures réglementaires en vigueur. Ici également, les délais pour

l'acquisition du matériel doivent être pris en compte pour ne pas induire des retards sur la mise en œuvre du projet ;

- D'assurer l'effectivité de la mise à disposition des ressources matérielles et financières : une fois acquis, toutes les procédures doivent être mises en œuvre pour effectivement mettre le matériel à la disposition de l'équipe de collecte. Dans le cas de collecte assistée par ordinateur, cette activité permet d'obtenir l'application informatique qui servira à la collecte et à la gestion des données. Il est important de s'assurer de la disponibilité de la totalité du financement de la collecte à cette étape ;
- D'installer l'infrastructure de collecte : il s'agit éventuellement des bâtiments nécessaires pour le stockage des questionnaires, des pools de saisie, des serveurs, ordinateurs, call center... ;
- De s'assurer de tous les arrangements nécessaires pour la collecte : la collecte de données peut nécessiter l'intervention des organismes intermédiaires, par exemple les sous-traitants pour les interviews téléphoniques, les administrations pour la fourniture de certaines informations nécessaires, les fournisseurs d'accès à internet pour la circulation des données... Ces questions sont réglées ici ;
- De mettre sur pied de l'équipe de collecte, conformément à la stratégie définie à la phase de conception, et d'effectuer le déploiement ;
- De distribuer le matériel de collecte selon la répartition retenue ;
- D'exécuter des actions de promotion pour la collecte : les activités de sensibilisation et de communication devant faciliter l'accueil des équipes de collectes sont gérées ici (lettres d'information des unités sélectionnées, dépliants sur des statistiques clés pouvant inciter la participation de répondants, stratégies facilitant la communication des informations destinées au public...) ;
- De prendre toutes les dispositions pratiques nécessaires pour le bon déroulement de la collecte : sécurité (au besoin) des biens et des personnes, interprètes, etc ;
- De documenter la mise en œuvre du sous-processus.

Dimension de la qualité des données concernée: Conditions préalables, exactitude (fiabilité)

IV.3. Exécution de la collecte

Portée et objet : Ce sous-processus est celui de mise en œuvre de la collecte au moyen de différents instruments retenus pour recueillir l'information, qui peut comprendre des microdonnées brutes ou des agrégats produits à la source, ainsi

que toutes métadonnées connexes. Il comprend le premier contact avec les fournisseurs de données et toute mesure de suivi ou de rappel. Il peut comprendre l'entrée manuelle des données au point de contact, ou la gestion des travaux sur le terrain, selon la source et le mode de collecte. Il permet d'enregistrer quand et comment on a communiqué avec les fournisseurs de données et s'ils ont répondu. Ce sous-processus comprend également la gestion des fournisseurs de données qui participent à la collecte en cours et sert à veiller à ce que les relations entre l'organisme statistique et les fournisseurs de données demeurent positives, ainsi qu'à enregistrer les commentaires, demandes de renseignements et plaintes et à y répondre.

Produit attendu : il est attendu de ce sous-processus des supports de collecte remplis pour tout l'échantillon et centralisés, de même que le rapport décrivant de déroulement de l'opération de collecte.

Lignes directrices : Dans ce sous-processus, l'équipe de collecte doit s'atteler à :

- Établir et entretenir de bonnes relations avec les répondants, afin d'obtenir des taux de réponse satisfaisants. Les activités de promotion relevées au sous-processus 4.2 seront d'une utilité certaine, et il reviendra aux interviewers de poursuivre les efforts afin de mettre en confiance les unités à enquêter.
- Choisir le moment opportun, pour appeler ou visiter les unités d'enquête, en se référant aux parodonnées produites lors des cycles précédents de l'enquête ou à celles d'une enquête similaire, et veiller à ce que le nombre de visites ou d'appels n'excède pas les limites acceptables.
- Garantir la confidentialité des données en respectant les règles relatives à la diffusion de l'information contenues dans la loi statistique de 1991.
- Bien manager les interviewers en mettant en place des conditions (physiques, matérielles, etc.) optimales pour un meilleur résultat
- Assurer le suivi et le contrôle de la collecte, qui consiste principalement à vérifier que les unités enquêtées sont celles sélectionnées, appliquer les procédures de contrôle de la qualité afin d'assurer la fiabilité et la cohérence des données recueillies, considérer la possibilité de lancer un programme visant la répétition des interviews afin d'évaluer l'exactitude des opérations de manière globale ;
- Suivre la progression de la couverture et des coûts (par exemples les dépenses liées au déplacement de l'interviewer, au transfert des données...) et les comparer aux estimations prévues. Ces indicateurs permettent de prendre de meilleures décisions pendant la période de collecte, mais aussi d'évaluer les

coûts et les efforts additionnels à consentir pour l'accroissement des taux de réponse plus particulièrement vers la fin de la collecte ;

- Mettre en place des procédures pour minimiser la non-réponse. Ceci est en fait un des enjeux principaux des opérations de collecte de données ;
- Recueillir (à la fin de la collecte) pour un échantillon des unités sélectionnées n'ayant pas répondu des informations nécessaires pour l'ajustement de la non-réponse (comme la taille du ménage par exemple) ;
- Documenter la mise en œuvre du processus de collecte, notamment avec l'élaboration d'un rapport de collecte.

Dimension de la qualité des données concernée des données concernée : Exactitude (fiabilité)

IV.4. Mise au point finale de la collecte et centralisation des données

Portée et objet : ce sous-processus concourt à transférer les données recueillies et les métadonnées dans un environnement électronique approprié aux fins de traitement. Il peut également comprendre la saisie manuelle ou automatique de données, par exemple en utilisant des employés administratifs ou des outils de reconnaissance optique de caractères pour extraire des données de questionnaires papier, ou la conversion des formats des fichiers de données reçus d'autres organismes. Il peut également comprendre l'analyse des métadonnées du traitement associées à la collecte (paradonnées), de manière à évaluer si les activités de collecte ont répondu aux exigences.

Produit attendu : il est essentiellement attendu de cette étape une ou plusieurs bases de données prêtes pour les traitements (fichiers de données brutes)

Lignes directrices : les modes de collecte ont une influence sur les activités à mener dans ce sous processus, la différence étant l'intégration de la saisie des données lorsque la collecte n'est pas assistée par ordinateur. Ainsi, ces activités peuvent consister à :

- Saisir des données : le sous-processus 4.2 a prévu le recrutement et la formation du personnel qui sera chargé de la saisie le cas échéant. Le choix peut également être porté sur une saisie automatique basée sur des procédures de numérisation et de reconnaissance intelligente de caractères. Il est important de prendre des mesures pour minimiser et corriger le cas échéant des erreurs de frappe, comme la double saisie par exemple. Il faut mettre à l'essai les systèmes automatisés de saisie des données qui sont basés sur la reconnaissance intelligente des caractères à partir d'images numérisées avant de les utiliser.

- Documenter le sous-processus.

Dimension de la qualité des données concernée: exactitude (fiabilité)

V. Traitement

Cette phase décrit l'apurement de la base de données issue de la collecte et leur préparation aux fins d'analyse. Elle se compose de sous-processus qui permettent de vérifier, de nettoyer et de transformer les données d'entrée, de manière à pouvoir les analyser et les diffuser comme des produits statistiques. Elle peut être répétée plusieurs fois au besoin. Dans le cas de produits statistiques qui sont élaborés sur une base régulière, cette phase a lieu à chaque itération.

Les phases « Traitement » et « Analyse » peuvent être itératives et parallèles. L'analyse peut fournir d'autres précisions sur les données, qui peuvent révéler la nécessité d'autres traitements. Les activités des phases « Traitement » et « Analyse » peuvent commencer avant que la phase « Collecte » ne soit terminée. Cela permet de compiler les résultats provisoires lorsque l'actualité est une préoccupation importante pour les utilisateurs.

V.1 Intégration des données

Portée et objet : ce sous-processus permet d'intégrer les données provenant d'une ou de plusieurs sources. C'est ici que les résultats des sous-processus de la phase de la collecte sont combinés. Les données d'entrée peuvent provenir d'un mélange de sources de données externes ou internes, et de divers modes de collecte. Il en résulte un ensemble de données harmonisées.

L'intégration des données peut avoir lieu à n'importe quel point durant cette phase de traitement, avant ou après l'un quelconque des autres sous-processus. En outre, un processus de production statistique peut comprendre plusieurs occurrences d'intégration de données.

Produit attendu : ce sous-processus permet d'obtenir une base de données unifié, intégrant toutes les sources.

Lignes directrices : la combinaison des données provenant de sources multiples (enquêtes à plusieurs niveaux comme ménage – femme –enfant), de différents modes de collecte (collecte avec un questionnaire papier plus collecte assistée par ordinateur) ou encore de différentes couches d'informations (coordonnées géographiques, photos, questionnaire, etc.) comprend habituellement les activités suivantes :

- Etablir les priorités, lorsque deux ou plusieurs sources contiennent des données pour la même variable (dont les valeurs peuvent être différentes) ;
- Traiter les identifiants des unités des différentes sources: contrôler les changements d'identifiants, assurer pour chaque individu l'unicité de l'identifiant dans l'ensemble des fichiers de données...
- Créer des clés d'appariement : appariement exact ou appariement statistique
- Exécuter les routines d'appariement ou de couplage d'enregistrements
- Valider les résultats de l'appariement, en évaluant les éléments de contrôle (effectif attendu, doublons, etc.). Si ces éléments ne sont pas satisfaisants, revoir les routines d'appariement et répéter le processus.
- Documenter le déroulement du sous-processus.

Dimension de la qualité des données concernée: Exactitude (et fiabilité)

V.2 Classification et codage

Portée et objet : ce sous-processus permet de classer et de coder les données d'entrée, dans l'idée d'attribuer des codes numériques à des réponses textes selon un système de classification préétabli. Lors de la conception du questionnaire, il est préférable de codifier (élaborer les classifications) autant que possible les questions fermées. Les données relatives aux questions ouvertes doivent être codifiées après la collecte, par l'équipe de conception après le dépouillement. Le codage quant à lui, effectué par le personnel recruté et formé, peut s'effectuer manuellement ou alors de façon automatique (prévoir dans ce cas l'acquisition du matériel nécessaire). Cet aspect doit être pris en compte dans la gestion du personnel et dans l'acquisition du matériel. Elle doit être conforme aux procédures mises sur pied lors de la phase de conception (sous-processus 2.5).

Produit attendu : à l'issue de ce processus, il est attendu une base de données intégrée et codifiée.

Lignes directrices : lors des opérations de codage, effectuées par le personnel recruté et formé, il convient de suivre les procédures de manière constante pour toutes les unités d'enquête, afin d'éviter, le plus possible, de commettre des erreurs. Les cas difficiles doivent être soumis à un petit groupe d'experts en la matière. Le codage post collecte doit être centralisé afin de réduire les coûts et de pouvoir bénéficier plus facilement des connaissances des experts. Il est important de mettre sur pied des procédures de contrôle, par exemple la vérification à posteriori d'un échantillon de données codées ;

Ce sous-processus comprend usuellement les activités suivantes :

- Etablir les nomenclatures et les systèmes de classification à partir du dépouillement d'un échantillon des réponses aux questions ouvertes, ainsi que des modalités non prévues pour les questions fermées correspondant généralement à une modalité dénommée "Autre", dont le contenu est précisé comme une réponse ouverte. Les nomenclatures sont établies sur la base de la récurrence des réponses observées principalement ;
- Exécuter les routines de codage à partir des nomenclatures élaborées ;
- Soumettre les cas ambigus au groupe d'experts commis à cet effet ;
- Vérifier la qualité de la codification, par exemple à partir de l'examen d'un échantillon de données codifiées ;
- Documenter de la mise en œuvre du sous-processus ;

Dimensions de la qualité des données concernées: exactitude (fiabilité), cohérence (Comparabilité)

V.3 Examen et validation de données

Portée et objet : ce sous-processus examine les données à l'effet de déterminer d'éventuels problèmes, erreurs et divergences comme les valeurs aberrantes, la non-réponse partielle et le codage erroné. On l'appelle également validation des données d'entrée. Il peut être exécuté d'une manière itérative, en validant les données par rapport à des règles de vérification prédéterminées lors de la phase de conception. Il peut indiquer les données devant faire l'objet d'une inspection ou d'une révision automatique ou manuelle. Cette activité doit démarrer pendant la collecte, dès les premières remontées des données, de manière à repérer les problèmes et à répercuter les constats aux équipes de terrain, qui peuvent y apporter des corrections.

Alors que ce sous-processus concerne la détection des erreurs réelles ou potentielles, toute activité de correction qui modifie réellement les données s'effectue au sous-processus 5.4 Contrôle et imputation.

Produit attendu : ce sous-processus n'apporte pas de modification à la base de données. Il est donc attendu de ce dernier un rapport de détection des erreurs (potentielles).

Lignes directrices : la principale activité ici c'est l'exécution des routines de contrôle pour la détection des erreurs tel qu'élaborées lors du sous-processus 3.2. Ces routines peuvent avoir trait à :

- La détection des données manquantes ;
- La détection des valeurs aberrantes ;

- La détection des incohérences ;
- La détection des codes erronés.

Dimensions de la qualité des données concernées: exactitude et fiabilité, cohérence et comparabilité.

V.4 Contrôle et imputation

Portée et objet : lorsque l'étape de vérification révèle des données invalides, manquantes ou incohérentes, ce sous processus permet d'attribuer, dans la mesure du possible, des valeurs de remplacement à celles-ci, afin de produire des estimations de qualité, conformes aux caractéristiques de la population.

Produit attendu : le produit ce sous-processus est la base de données intégrée, codifiée et apurée (c'est-à-dire celle intégrant les données imputées).

Lignes directrices : le contrôle et l'imputation des données visent toute une gamme de méthodes basées sur des règles bien précises. Les diverses étapes sont habituellement les suivantes :

- Analyser les erreurs et évaluer la pertinence de modifier des données, selon les principes de la gestion des erreurs retenus à la phase de conception, sous-processus 2.5 ;
- Sélectionner la méthode à utiliser, qui comprend le choix des variables auxiliaires à utiliser, la modélisation à l'aide d'autres sources de données ;
- Imputer les valeurs de données ;
- Valider les imputations, qui consiste à évaluer la cohérence des données après imputation, comparer les estimateurs (avec et sans valeurs imputées), etc.
- Documenter le sous-processus, pour rendre compte de toutes les corrections qui ont été apportées aux données, ainsi que des méthodes utilisées, etc.

Dimensions de la qualité des données concernées : Exactitude (et fiabilité), cohérence (et comparabilité).

V.5 Dérivation de nouvelles variables et unités

Portée et objet : ce sous-processus permet de dériver les valeurs de variables et d'unités statistiques qui ne sont pas explicitement fournies dans la collecte, mais qui sont nécessaires pour livrer les produits requis. Les nouvelles variables peuvent être obtenues en appliquant des formules arithmétiques à l'une ou à plusieurs des variables qui sont déjà présentes dans l'ensemble de données, ou en appliquant différentes hypothèses de modèle.

Produit attendu : le produit ce sous-processus est la base de données intégrée, codifiée et apurée, et complétée de nouvelles variables et unités statistiques.

Lignes directrices : il pourra être nécessaire de procéder de manière itérative, puisque certaines variables dérivées peuvent elles-mêmes être fondées sur d'autres variables dérivées. Il importe donc de veiller à ce que les variables soient dérivées dans l'ordre voulu. Dériver des ménages lorsque les unités de collecte sont des personnes, ou des entreprises lorsque les unités de collecte sont des unités légales, en sont des exemples. Les activités de ce sous-processus peuvent être :

- Vérifier et valider les conditions de dérivation des nouvelles unités et variables ;
- Appliquer les formules arithmétiques ou d'agrégation pour la dérivation de nouvelles unités statistiques et de nouvelles variables ;
- Documenter le sous-processus, pour présenter toutes les nouvelles variables et unités, ainsi que les éléments méthodologiques pour leur calcul.

Dimensions de la qualité des données concernées : Exactitude (et fiabilité), cohérence (et comparabilité).

V.6 Calcul des poids de sondage

Portée et objet : les poids sont utilisés pour permettre l'inférence adéquate des résultats observés sur l'échantillon à la population cible. Le calcul des poids de sondage n'a de signification que dans le cas des enquêtes probabilistes, c'est-à-dire celles où tous les individus de la population possèdent une probabilité connue d'être sélectionnés dans l'échantillon.

Ce sous-processus permet de créer et d'appliquer des poids aux données individuelles selon la méthode retenue lors de la phase de conception, au sous-processus 2.4 Conception de la base de sondage et de l'échantillon. Ce sous processus consiste aussi à réajuster les poids de sondage afin de tenir compte de la non-réponse dans les dénombrements totaux.

Ce sous-processus peut être déclenché aussitôt après le sous-processus 2.4 pour le calcul des poids de sondage, la méthodologie utilisée permettant d'obtenir les probabilités d'inclusion. La suite du processus a lieu une fois que la collecte est terminée.

Produit attendu : selon le cas, ce sous-processus aboutit aux poids de sondage si on n'utilise pas de calage, ou alors au poids de calage. Dans les deux cas, ces poids sont corrigés pour tenir compte de la non-réponse.

Lignes directrices : les activités de ce sous-processus consistent à :

- Calculer les poids de sondage selon la méthodologie retenue lors de la phase de conception au sous-processus 2.4 (Conception de la base de sondage et de l'échantillon) ;
- Réajuster les poids de sondage pour tenir compte des non réponses ;
- Vérifier la disponibilité des variables auxiliaires de calage : Si des données auxiliaires sont disponibles, il est possible d'améliorer la précision des estimations. L'intégration de données auxiliaires dans les processus d'estimation est appelée calage. Le calage consiste à rajuster les poids, afin que les estimations de la ou des variables auxiliaires correspondent à des totaux connus, aussi appelés totaux de contrôle.
- Vérifier de manière empirique la corrélation entre les variables auxiliaires de calage et la variable d'intérêt ;
- Calculer les poids de calage, en contrôlant la fourchette des poids pour éviter qu'ils ne soient très gros ou négatifs ;
- Documenter le sous-processus, pour rendre compte de tous les travaux effectués pour aboutir aux poids définitifs appliqués aux individus de l'échantillon.

Dimensions de la qualité des données concernées: exactitude (et fiabilité), cohérence (et comparabilité).

V.7 Calcul des indicateurs et estimation des paramètres

Portée et objet : les principaux paramètres d'intérêt lors d'une enquête sont habituellement les suivants : totaux, moyennes, proportions, ratios et percentiles. Ce sous-processus permet d'obtenir ces paramètres sur la population cible à partir de microdonnées collectées et traitées. Les estimations des indicateurs retenus lors du sous-processus 2.1. Conception des produits devraient prendre en compte le poids de sondage, le recours à la repondération, et le calage aux marges.

Les éléments de précision des estimateurs sont également calculés ici. La fiabilité d'une estimation est évaluée en calculant l'erreur-type ou le coefficient de variation de l'estimateur utilisé. L'intervalle de confiance pour un paramètre est calculé en utilisant soit l'erreur-type de l'estimation, soit la distribution statistique de l'estimateur choisi.

Produit attendu : il est attendu de ce sous-processus les estimations des paramètres de la population, ainsi que les éléments de précision associés.

Lignes directrices : les activités de ce sous-processus intègrent les suivantes :

- Exécuter les routines de calcul des estimations selon la tabulation retenue au processus 2.1. Conception des produits, en tenant compte des poids de sondage, de la repondération, et du calage éventuellement ;
- Calculer les éléments de précision des estimations (erreur type, coefficient de variation, intervalle de confiance...) ;
- Valider les estimations : comparaison avec des informations connues, en tenant compte des poids de sondage et des poids de calage. Lorsque les résultats ne concordent pas avec l'information existante, vérifier les procédures de calcul des poids de sondage, de repondération et de calage et faire à nouveau les estimations.
- Documenter le processus.

Dimensions de la qualité des données concernées: Exactitude (et fiabilité), cohérence (et comparabilité).

V.8 Mise au point finale des fichiers de données

Portée et objet : ce sous-processus permet de rassembler les résultats des autres sous-processus de cette phase et de produire un fichier de données et de microdonnées qui sert d'intrant à la phase d'analyse. A ce niveau, on s'assure que tous les traitements effectués ont été correctement intégrés à la base des données brutes.

Produit attendu : à l'issue de ce sous-processus, on obtient un fichier de microdonnées (intégrant tous les traitements et prêt pour les analyses plus poussées), ainsi que les principaux indicateurs recherchés par l'enquête. Il est également attendu le détail de tous les traitements effectués sur la base de données, de préférence dans un rapport spécifique.

Lignes directrices :

- Vérifier que tous les traitements nécessaires ont été effectués sur la base de données ;
- Elaborer le rapport sur les différents traitements effectués sur la base de données.

Dimension de la qualité des données concernée des données concernée: intelligibilité.

VI. Analyse

Cette phase est celle où les produits statistiques sont obtenus, examinés dans le détail et préparés en vue de leur diffusion. Il s'agit ici d'élaborer et de valider le

contenu statistique des produits en rapport avec les besoins et les objectifs fixés, avant leur diffusion. Ce contenu comprend non seulement les indicateurs et statistiques produites, mais aussi les notes techniques et toutes les observations nécessaires pour comprendre les statistiques. Cette phase consiste globalement à élaborer les produits statistiques, les interpréter et procéder au contrôle de la non divulgation des données individuelles.

NB : les résultats à interpréter concernent aussi bien les indicateurs que les modèles statistiques et les analyses exploratoires des données

VI.1. Elaboration des projets de produits

Portée et objet : ce sous-processus consiste à transformer les données (issues du traitement) en statistiques, et à produire toutes les informations supplémentaires et mesures additionnelles pour l'analyse des statistiques, telles que les tendances, les indices, les séries corrigées, etc. Le calcul des caractéristiques de qualité constitue aussi un élément important de cette étape.

Produit attendu : à l'issue de ce sous processus, le draft de produit (contenu statistique) doit être disponible, pour examen et validation

Lignes directrices : ce sous processus consiste à :

- Mettre au point des tableaux des estimations ;
- Préparer les ébauches des produits, qui consiste principalement à alimenter les tableaux de la maquette, charger les bases de données, les fichiers de microdonnées pour les utilisateurs, les métadonnées de diffusion ;
- Mettre à jour les séries chronologiques au besoin, pour les opérations périodiques ;
- Effectuer les ajustements saisonniers, effet calendaire (au besoin) ;
- Documenter.

Dimension de la qualité des données concernée des données concernée: Exactitude (et fiabilité)

VI.2. Validation des produits statistiques

Portée et objet : à ce niveau, les statisticiens doivent valider et adopter le contenu des produits. Il s'agit particulièrement de détecter et examiner toute divergence par rapport aux attentes et avec d'autres données pertinentes en rapport avec le domaine statistique concerné, afin de faire des analyses et interprétations éclairées.

Produit attendu : à l'issue de ce sous processus, les statistiques doivent être validées non seulement sur le plan « purement statistique » (Dimensions de la qualité),

mais aussi du point de vue de la cohérence avec d'autres statistiques pertinentes du domaine ou des domaines connexes.

Lignes directrices : la validation des statistiques passe par les activités suivantes :

- Vérifier la couverture et l'examen des taux de réponse, pour s'assurer qu'ils répondent aux prescriptions, de manière à permettre le calcul des indicateurs, surtout au niveau désagrégé ;
- Comparer les statistiques avec celles des itérations précédentes : pour les opérations périodiques, qui portent généralement sur les phénomènes structurels, il convient à chaque itération de s'assurer que les statistiques suivent la tendance générale, ou que les écarts peuvent être justifiés par des événements bien identifiés ;
- Comparer les statistiques à d'autres données pertinentes sur les plans internes et externes, en vue de s'assurer qu'elles constituent un ensemble de données cohérentes (le cas par exemple des statistiques sur la morbidité et l'accès aux soins de santé, la prévalence des maladies hydriques et l'accès à l'eau potable, etc.) :
- Rechercher d'autres incohérences dans les statistiques :
- Vérifier les macrodonnées, à savoir les statistiques sur des groupes ou des agrégats tels que les proportions, les moyennes, les totaux, les fréquences.
- Valider les statistiques par rapport aux attentes et aux renseignements sur le domaine ;
- Documenter.

Dimension de la qualité des données concernée des données concernée: cohérence (et comparabilité).

VI.3. Interprétation et explication des données

Portée et objet : il s'agit à ce niveau de présenter les données produites de manière pertinente, de façon à permettre aux utilisateurs de comprendre les résultats de l'opération statistique. C'est à ce niveau que les statisticiens s'appuient sur des connaissances approfondies pour expliquer les statistiques en évaluant la mesure dans laquelle elles répondent aux attentes initiales, et effectuent des analyses plus ou moins approfondies. Les données doivent être présentées clairement de manière à éviter les erreurs d'interprétation.

Produit attendu : au terme de ce sous processus, les statistiques validées sont résumées sous forme de texte, commentées et analysées.

Lignes directrices : une bonne présentation et interprétation des statistiques obtenues consiste à :

- Elaborer et adopter une démarche pour l'analyse, qui consiste par exemple à orienter l'interprétation des tendances à court, moyen et long terme, à choisir des points de référence pertinents et significatifs dans le temps ou dans l'espace, etc. ;
- Adopter la présentation des données (graphiques, tableaux, cartes, etc.) : les choix s'opèrent en fonction des données, du type de message à passer, et du public cible ;
- Identifier les sujets pertinents d'analyse : en rapport avec les thématiques abordées, il convient de détecter dans les statistiques les informations saillantes autour desquelles va se structurer l'analyse et l'interprétation des données ;
- Identifier les éléments de contexte liés aux résultats, en l'occurrence ceux qui peuvent permettre de comprendre les statistiques obtenues (une politique mise en œuvre, un évènement majeur, une position géographique, etc.), sans traduire nécessairement un lien direct de causalité ;
- Utiliser les données des itérations précédentes pour les séries ;
- Documenter.

Dimensions de la qualité des données concernées: Pertinence, Accessibilité (et clarté), intelligibilité

VI.4. Contrôle de la non divulgation des données individuelles

Portée et objet : au cours de ce sous processus, il est question de veiller à ce que les produits à diffuser (données, microdonnées et métadonnées) respectent les règles en vigueur en matière de confidentialité, tout en préservant l'utilité des données. Il s'agit d'effectuer des contrôles systématiques de non divulgation à tous les niveaux (primaire et secondaire), et d'appliquer les techniques nécessaires en fonction du type de données à publier et de leur utilisation. Une bonne mise en œuvre de ce sous processus contribue à accroître la confiance des répondants envers le système statistique et à assurer des taux de réponse élevés.

Produit attendu : les données, microdonnées et métadonnées à diffuser sont débarrassées de tous les risques d'identification et de divulgation des données confidentielles.

Lignes directrices : le contrôle de la non divulgation des données individuelles consiste à :

- Identifier et évaluer les risques d'identification, c à d repérer toutes les données (catégories de données, croisement de données, désagrégations de données) susceptibles de contribuer à l'identification des répondants (cas par

exemple des secteurs à peu d'entreprises). Penser aussi au risque dû au recoupement de données (à publier) avec d'autres renseignements accessibles (publications antérieures, données connexes, etc.) ;

- Identifier le degré de risque et de la méthode de contrôle : en fonction du niveau des risques identifiés et du type de produits, il peut s'agir d'anonymisation des variables d'identification, de suppression ou de perturbation des données, etc. ;
- Appliquer les techniques appropriées.

Dimensions de la qualité des données concernées: Exactitude, accessibilité.

VI.5. Mise au point finale des produits

Portée et objet : la finalisation des produits consiste globalement à s'assurer que les statistiques et l'information associée sont de qualité, répondent aux objectifs, sont utilisables et donc prêtes à être diffusées. Il s'agit de s'assurer que tous les sous processus de cette étape ont été efficacement mis en œuvre.

Produit attendu : le produit est validé (sur le plan de la qualité des données statistiques) et prêt à être diffusé.

Lignes directrices :

- Appliquer les derniers contrôles de cohérence ;
- Appliquer les mises en garde voulues pour la diffusion ;
- Assembler toutes les métadonnées nécessaires (notes techniques, informations d'appui, commentaires, etc.) en vue de produire les documents de métadonnées ;
- Organiser les discussions préalables à la diffusion, avec les spécialistes : (en vue d'approuver le contenu pour diffusion) ;
- Rédiger les autres parties connexes : avant-propos, introduction, acronymes, conclusion, résumé, encadrés, etc. au besoin ;
- Valider :
 - *Relire (vérifier la cohérence globale des analyses, forme, etc.) ;*
 - *valider et prendre en compte les amendements ;*
 - *Produire et valider les annexes ;*
- Documenter.

VII. Diffusion

La diffusion consiste à mettre les données obtenues dans le cadre d'une activité statistique à la disposition des utilisateurs par divers moyens. Elle rassemble toutes les activités liées non seulement à la remise des produits obtenus aux utilisateurs,

mais aussi à la communication sur ces données et aux interactions avec les utilisateurs, afin de les aider à accéder aux produits et à bien les utiliser.

VII.1. Mise à jour des systèmes de diffusion (préparation de la diffusion)

Portée et objet : organiser les données et les métadonnées aux fins de diffusion, et s'assurer que toutes les métadonnées nécessaires sont en place et prêtes à être diffusées.

Produit attendu : il est attendu ici que les métadonnées soient disponibles, corrélées aux données à diffuser, et prêtes pour être mises à la disposition des utilisateurs

Lignes directrices : les activités de ce sous processus devraient s'opérer plus en amont de préférence, mais c'est à ce stade que l'on procède au contrôle final tendant à vérifier que toutes les métadonnées nécessaires sont en place prêtes à être diffusées. En général, il s'agit de :

- Formater les données et les métadonnées prêtes à être saisies dans les bases de données produites;
- Télécharger les données et métadonnées dans les bases de données produites;
- Faire en sorte que les données soient corrélées aux métadonnées pertinentes.

Dimensions de la qualité des données concernées: accessibilité, intelligibilité.

VII.2. Production des produits de diffusion

Portée et objet : au cours de ce sous-processus, il est question de produire ou de rendre disponibles les produits de diffusion sous divers formats et sur divers supports, tels que conçus au préalable pour répondre aux besoins des utilisateurs (imprimés, CD, bases de données, fichiers de microdonnées, fichiers téléchargeables, atlas, etc.).

Produit attendu : Différents produits de diffusion accessibles aux utilisateurs

Lignes directrices : le processus comprend généralement les activités suivantes :

- Formater les produits de diffusion selon les outils conçus (texte explicatif, tableaux, graphiques, énoncés sur la qualité, etc., (le produit de diffusion obtenu en 8.1 n'épouse pas forcément le format de l'outil de diffusion tel que conçu)
- assembler les composantes en produits ;
- Vérifier et valider les produits, dans le fond et la forme, en s'assurant le cas échéant de la fonctionnalité (liens pour les produits électroniques) ;
- Traduire, en veillant à ce que les deux versions concordent en ce qui a trait aux données et au texte ;

- S'assurer du respect des normes de publication
- Editer.

Dimensions de la qualité des données concernées: accessibilité

VII.3. Mise au point du processus de diffusion

Portée et objet : après avoir validé le produit sur le plan technique, il convient de s'assurer que tous les éléments sont en place aux fins de la diffusion grand public. Cette étape est celle de la gestion du calendrier de publication, de l'information préalable aux groupes spécifiques ayant l'accès privilégié (et réglementé) à l'information (les responsables ministériels par exemple), de la préparation des éléments de communication sur les produits, la gestion de l'accès aux données confidentielles par les groupes d'utilisateurs agréés, etc.

Produit attendu : Calendrier de diffusion validé

Lignes directrices : la mise au point du processus de diffusion consiste à :

- Valider le choix du moment de la diffusion : ce choix peut dépendre par exemple de certains facteurs liés au contexte socioéconomique, de nature à influencer l'interprétation des résultats ;
- Informer les groupes particuliers d'utilisateurs comme les membres de la presse ou les ministres ainsi que tout embargo avant la diffusion, à condition que l'accès privilégié à l'information soit prévu et bien réglementé dans la politique nationale de diffusion ;
- organiser la gestion de l'accès aux données sensibles, par les groupes d'utilisateurs autorisés, comme les chercheurs. Il s'agit de définir le niveau et les conditions d'accès en fonction des données et des utilisateurs ;
- S'assurer de la conformité aux procédures et protocoles standards de diffusion et de publication pour les données, métadonnées, la documentation et les indicateurs de la qualité

Dimensions de la qualité des données concernées: accessibilité (clarté), actualité (ponctualité), cohérence (comparabilité)

VII.4. Promotion des produits de diffusion

Portée et objet : au cours de ce sous processus, il est question de mettre les produits statistiques à la disposition du public par divers moyens et d'en faire activement la promotion, afin d'atteindre le plus vaste auditoire possible et optimiser l'utilisation des produits. Cette diffusion concerne aussi les métadonnées.

Produit attendu : Produits diffusés.

Lignes directrices :

- Charger les supports de diffusion : impression, gravure, site web, etc.
- Cibler les utilisateurs éventuels des produits : bien que les principaux utilisateurs (gouvernement, partenaires) aient été identifiés et associés à la détermination des besoins, la diffusion doit être élargie et cibler tous les groupes d'utilisateurs potentiels (organismes, entreprises, chercheurs, etc.) ;
- Présenter les produits aux utilisateurs, y compris toutes les informations utiles telles que les formats disponibles, les supports, etc. ;
- Communiquer et mettre en œuvre les actions de promotion via les médias, les sites Web, ateliers, etc.

Dimensions de la qualité des données concernées: pertinence, accessibilité (clarté).

VII.5. Gestion du soutien aux utilisateurs

Portée et objet : après avoir mis les produits statistiques à la disposition des différents groupes d'utilisateurs, il est indispensable d'interagir de manière étroite et efficace avec ces derniers. Il s'agit de garantir que leurs requêtes par rapport aux données publiées sont enregistrées, suivies et traitées, de leur fournir toutes les informations complémentaires pour une bonne compréhension des données, et en même temps de recenser et documenter les erreurs qui ont échappé aux diverses vérifications. C'est aussi le cadre pour entrevoir de nouveaux besoins des utilisateurs ou l'évolution des besoins existants.

Produit attendu : le rapport de suivi des utilisateurs, qui présente les informations telles que la fréquence des requêtes, le type de requêtes, les erreurs signalées, les avis sur la pertinence et la qualité des produits, etc.

Lignes directrices :

- Mettre à disposition des coordonnées utiles aux utilisateurs : les utilisateurs doivent savoir qu'ils ont la possibilité d'exprimer des requêtes par rapport aux produits, et avoir connaissance de tous les canaux disponibles (centre de documentation, boîte aux lettres dédiée, site internet, etc.) ;
- Enregistrer, analyser et traiter les requêtes des utilisateurs
- Recueillir, évaluer et faire la synthèse des commentaires des utilisateurs
- Documenter

Dimension de la qualité des données concernée: pertinence, accessibilité (clarté)

VIII. Evaluation

La phase ultime du processus statistique est celle de l'évaluation. En effet, le processus statistique ne s'arrête pas avec la diffusion des données produites, il convient de déterminer les points forts et les points faibles à chacune des étapes, afin d'apprécier globalement la qualité de la mise en œuvre du processus. La qualité du processus est révélateur de la qualité des données produites, bien qu'une évaluation de la qualité des données elles même reste nécessaire selon les outils reconnus. Bien qu'elle ait lieu à la fin du processus, l'évaluation repose sur un ensemble de données recueillis tout au long des diverses phases. L'intérêt de cette évaluation réside dans l'identification des améliorations possibles à apporter au processus lors des prochaines itérations ou pour des opérations similaires. Bien qu'elle soit particulièrement utile pour les nouvelles opérations, l'évaluation est nécessaire même pour les cycles réguliers et bien établis, même si dans certains cas cette évaluation n'est pas officielle et peut ne pas se faire de manière constante. Elle peut par exemple dans ces cas servir à déterminer si l'itération suivante doit démarrer dès la phase de l'identification des besoins, ou à partir d'une phase ultérieure.

VIII.1 Recueil des données d'évaluation

Portée et objet : à ce niveau, il est question de rassembler toutes les données d'entrée et les Dimensions de la qualité établis tout au long du processus de production pour préparer l'évaluation. Ces données peuvent prendre diverses formes, réactions des utilisateurs, métadonnées de processus (paradonnées), paramètres de mesure des systèmes ou même les suggestions du personnel. Une liste de contrôle est dans tous les cas conçue à cet effet, pour chaque phase du processus. Toutes ces données sont rassemblées et mises à la disposition de l'équipe chargée de l'évaluation.

Produit attendu : Données d'évaluation disponibles auprès de l'équipe en charge de l'évaluation.

Lignes directrices : le recueil des données d'évaluation consiste à :

- Exploiter et compiler les rapports, en vue de renseigner les indicateurs, quantitatifs ou qualitatifs, définis pour chaque sous processus, chaque étape;
- Exploiter les métadonnées et les paradonnées, en vue de déterminer les sources d'erreurs dans la mise en œuvre du processus ;
- Recueillir les données de diverses sources : opinions et avis des experts, réactions des utilisateurs, opinions des techniciens, etc.

Dimensions de la qualité des données concernées: intelligibilité.

VIII.2 Conduite de l'évaluation

Portée et objet : l'évaluation consiste à analyser les données d'entrée (indicateurs, opinions, métadonnées, etc.) pour évaluer la mise en œuvre du processus de production, afin de recommander les changements qui s'imposent pour l'amélioration de la mise en œuvre des opérations futures. L'évaluation donne lieu à la production d'un rapport qui, en plus de présenter les résultats de l'évaluation proprement dite, consigne toutes les sources d'erreurs.

Produit attendu : Rapport d'évaluation de la qualité du processus

Lignes directrices : l'évaluation consiste à :

- Synthétiser les données d'évaluation : il s'agit, à défaut d'établir un indice de qualité unidimensionnel et unique, de résumer les différents indicateurs de la qualité et les comparer, sur le plan de leur importance relative et de leurs conséquences ;
- Identifier les principales sources d'erreurs (points faibles, manquements) ;
- Formuler les recommandations, en vue d'améliorer les itérations futures ou les opérations similaires ;
- Rédiger le rapport d'évaluation, qui fait ressortir le niveau de qualité de la mise en œuvre du processus, les difficultés et les recommandations.

Dimensions de la qualité : toutes les dimensions de la qualité : Pertinence, actualité, intelligibilité, exactitude, cohérence, accessibilité.

VIII.3 Adoption du plan d'actions

Portée et objet : l'évaluation d'un processus statistique se solde par l'élaboration d'un plan d'actions fondé sur les résultats de l'évaluation. Ce plan d'actions intègre l'adoption d'un mécanisme de surveillance de la mise en œuvre des mesures adoptées, ainsi que de l'impact de ces mesures, lequel pourra, à son tour, contribuer aux évaluations des itérations futures du processus.

Produit attendu : Plan d'actions d'amélioration de la qualité

Lignes directrices : les activités à mener pour élaborer peuvent être :

- Analyser les problèmes, qui consiste à déterminer les causes exactes des manquements observés tout au long du processus ;
- Identifier et élaborer les solutions adéquates, liées aux causes sous-jacentes ;
- Identifier les conditions de mise en application des solutions, en vue de distinguer celles qui peuvent être facilement appliquées de celles dont la mise

en application sont soumises à de fortes contraintes (qui vont généralement au-delà des seules compétences de l'appareil statistique) ;

- Elaborer le dispositif de mise en œuvre et de suivi des mesures adoptées :

Dimensions de la qualité : toutes les dimensions de la qualité : Pertinence, actualité, intelligibilité, exactitude, cohérence, accessibilité.

Bibliographie

NATIONS UNIES. 2014. Conférence des statisticiens européens. Modèle générique du processus de production statistique, Paris, France.

STATISTIQUE CANADA. 2009. «lignes directrices concernant la qualité», cinquième édition.

ESPAGNE, Institut National de la Statistique. 2015. «lignes directrices sur la qualité qualité», version actualisée.

Annexes

Annexe 1 : Nomenclature du processus de production statistique

1. DETERMINATION DES BESOINS

1.1. Identification et validation des besoins

- 1.1.1. Identification des utilisateurs
- 1.1.2. Identification des besoins internes
- 1.1.3. Identification des besoins externes
- 1.1.4. Concertations avec les utilisateurs

1.2. Identification des produits statistiques

1.3. Détermination des concepts

- 1.3.1. Identification des documents de normalisation au niveau international
- 1.3.2. Identification des cadres de normalisation au niveau national
- 1.3.3. Concertations avec les experts

1.4. Vérification de la disponibilité des données

- 1.4.1. Identification et examen des sources administratives
- 1.4.2. Identification et examen des autres sources
- 1.4.3. Vérification de la disponibilité des bases de données
- 1.4.4. Synthèse des sources identifiées

1.5. Etude de faisabilité

- 1.5.1. Synthèse des sources identifiées
- 1.5.2. Evaluation des actions à entreprendre
- 1.5.3. Réalisation de l'analyse coût – avantage
- 1.5.4. Evaluation des contraintes externes
- 1.5.5. Présentation du choix stratégique

2. CONCEPTION

2.1. Conception des produits

- 2.1.1. Conception des formats de diffusion
- 2.1.2. Conception de la tabulation des indicateurs
- 2.1.3. Conception des méthodes de contrôle de divulgation des données individuelles
- 2.1.4. Conception du calendrier d'exécution

2.2. Conception et description des variables

2.2.1. Identification et description des unités statistiques

2.2.2. Identification/élaboration des classifications

2.2.3. Description des variables

2.2.4. Description des agrégats de population

2.2.5. Définition des variables dérivées

2.3. Conception de la collecte

2.3.1. Conception du plan/stratégie de collecte

2.3.2. Choix du mode et de la méthode de collecte

2.3.3. Conception des outils de collecte

2.3.4. Conception du soutien matériel à la collecte

2.3.5. Conception des applications de gestion de la collecte

2.3.6. Conception des ressources humaines pour la collecte

2.3.7. Détermination des infrastructures pour la collecte

2.3.8. Choix du mode de saisie

2.3.9. Conception des autres éléments pour la collecte (langue, aspects éthiques, conventions, etc.)

2.4. Conception de la base de sondage et de l'échantillon

2.4.1. Conception de la base de sondage

2.4.2. Identification des sources de données pour la base de sondage

2.4.3. Conception de la méthode d'échantillonnage

2.4.4. Détermination de la taille et de la répartition de l'échantillon

2.4.5. Conception de l'extraction (tirage) de l'échantillon

2.5. Conception du traitement et de l'analyse

2.5.1. Conception de la codification

2.5.2. Conception de la détection et de la gestion des erreurs

2.5.3. Conception de la validation de l'échantillon corrigé

2.5.4. Conception de l'intégration des fichiers

2.5.5. Conception du calcul des poids

2.5.6. Conception des estimateurs et de la correction des estimateurs

2.5.7. Conception des estimateurs de la variance

2.5.8. Conception du contrôle de la confidentialité statistique

2.5.9. Conception de l'analyse

2.6. Conception du système de production

- 2.6.1. Conception de l'architecture de l'équipe
- 2.6.2. Conception du cahier de charges des intervenants
- 2.6.3. Conception de l'enchaînement des activités
- 2.6.4. Conception du dispositif de suivi
- 2.6.5. Elaboration du calendrier de travail

2.7. Conception de la diffusion et de la dissémination

- 2.7.1. Conception des métadonnées
- 2.7.2. Conception du fichier de microdonnées de diffusion
- 2.7.3. Conception du calendrier de diffusion
- 2.7.4. Conception du catalogue des produits
- 2.7.5. Conception de la diffusion
- 2.7.6. Conception de la communication sur les produits

2.8. Conception de l'archivage

- 2.8.1. Conception des règles d'archivage
- 2.8.2. Définition des produits d'archivage
- 2.8.3. Définition des procédures d'archivage

2.9. Conception de l'évaluation

- 2.9.1. Définition des critères d'évaluation
- 2.9.2. Identification des normes associées aux critères d'évaluation
- 2.9.3. Définition des sources d'évaluation

3. ELABORATION

3.1. Elaboration de l'instrument de collecte

- 3.1.1. Elaboration des questionnaires, métadonnées et supports de collecte
- 3.1.2. Elaboration des paramètres de suivi et contrôle de la collecte
- 3.1.3. Elaboration du plan/stratégie de collecte
- 3.1.4. Elaboration de la méthodologie de collecte
- 3.1.5. Elaboration du plan d'affectation des ressources humaines
- 3.1.6. Elaboration du plan de formation, d'évaluation et de sélection des ressources humaines
- 3.1.7. Développement/mise en place de l'infrastructure de collecte
- 3.1.8. Elaboration des contrôles primaires
- 3.1.9. Elaboration des accords de collaboration pour la collecte
- 3.1.10. Développement des actions de promotion de la collecte

3.2. Elaboration du processus de production

- 3.2.1. Elaboration du manuel/application de construction de la base de sondage
- 3.2.2. Elaboration de la méthode d'échantillonnage et du programme de tirage de l'échantillon
- 3.2.3. Elaboration du guide/programme de codification
- 3.2.4. Elaboration du manuel et des programmes de spécification et de gestion des erreurs
- 3.2.5. Elaboration de la méthodologie de validation de l'échantillon corrigé
- 3.2.6. Elaboration des méthodologies d'intégration des fichiers
- 3.2.7. Elaboration de la méthodologie et du programme de calcul des poids
- 3.2.8. Elaboration de la méthodologie et du programme de calcul des estimateurs
- 3.2.9. Elaboration de la méthodologie et du programme de correction des estimateurs
- 3.2.10. Elaboration de la méthodologie et du programme de calcul des estimateurs de la variance
- 3.2.11. Elaboration de la méthodologie et du programme de contrôle de la confidentialité statistique
- 3.2.12. Elaboration du guide d'analyse

3.3. Elaboration du processus de diffusion

- 3.3.1. Elaboration des produits de diffusion et leurs métadonnées
- 3.3.2. Construction du fichier de microdonnées de diffusion
- 3.3.3. Elaboration du calendrier de diffusion
- 3.3.4. Elaboration du catalogue des produits
- 3.3.5. Développement des outils de dissémination et de communication sur les produits

3.4 Configuration du déroulement des activités

- 3.4.1. Elaboration du plan de formation des ressources humaines
- 3.4.2. Elaboration du plan de désignation des intervenants
- 3.4.3. Elaboration du cahier de charges des intervenants
- 3.4.4. Description des activités et de leur enchaînement
- 3.4.4. Elaboration du dispositif de suivi du déroulement des activités
- 3.4.4. Elaboration du calendrier de travail

3.5. Mise à l'essai du système de production

- 3.5.1. Test des outils de sélection de l'échantillon et de construction de la base de sondage
- 3.5.2. Test des outils de collecte
- 3.5.3. Test des outils de traitement
- 3.5.4. Test des outils de diffusion

3.6. Mise à l'essai du processus de production statistique

- 3.6.1. Enquête pilote
- 3.5.2. Test des composantes et des programmes informatiques associés au processus de production

3.7. Mise au point finale du système et du processus de production

- 3.7.1. Formation des ressources humaines
- 3.7.2. Ajustements du système de production
- 3.7.3. Ajustements du processus de production
- 3.7.4. Documentation et archivage de tous les résultats de ce sous-processus

4. COLLECTE

4.1. Création/mise à jour de la base de sondage et de l'échantillon

- 4.1.1. Collecte et évaluation de la qualité des bases de sondage
- 4.1.2. Traitement des bases de sondage
- 4.1.3. Appariement des bases de sondage
- 4.1.4. Traitement et évaluation de la qualité de la base de sondage
- 4.1.5. Tirage de l'échantillon
- 4.1.6. Evaluation/Validation de l'échantillon
- 4.1.7. Ajustement de l'échantillon

4.2. Préparation de la collecte

- 4.2.1. Formation et recrutement du personnel de collecte
- 4.2.2. Acquisition et distribution du matériel de collecte
- 4.2.3. Vérification de la disponibilité des ressources matérielles et financières
- 4.2.4. Installation de l'infrastructure de collecte
- 4.2.5. Mise sur pieds de l'équipe de collecte
- 4.2.6. Autres actions : promotion de la collecte, arrangements et conventions pour la collecte, etc.

4.3. Exécution de la collecte

- 4.3.1. Formation et recrutement du personnel de collecte

4.3.2. Acquisition et distribution du matériel de collecte

4.3.3. Vérification de la disponibilité des ressources matérielles et financières

4.3.4. Installation de l'infrastructure de collecte

4.4. Mise au point finale de la collecte et centralisation des données

4.4.1. Saisie des données

5. TRAITEMENT

5.1. Intégration des données

5.1.1. Traitement des identifiants

5.1.2. Création des clés d'appariement

5.1.3. Appariement, couplage d'enregistrements

5.1.4. Validation des résultats de l'appariement

5.2. Classification et codage

5.2.1. Création des nomenclatures et des systèmes de classification

5.2.2. Exécution des routines de codage

5.2.3. Vérification de la qualité du codage

5.3. Examen et validation des données

5.3.1. Détection des données manquantes

5.3.2. Détection des valeurs aberrantes

5.3.3. Détection des incohérences

5.3.4. Détection des codes erronés

5.4. Contrôle et imputation

5.4.1. Analyse des erreurs

5.4.2. Sélection de la méthode d'imputation

5.4.3. Imputation des valeurs

5.4.4. Validation des imputations

5.5. Dérivation des nouvelles variables et unités

5.5.1. Validation des conditions de dérivation des nouvelles variables et unités

5.5.2. Dérivation des nouvelles variables et unités

5.6. Calcul des poids de sondage

5.6.1. Calcul des poids de

5.6.2. Réajustement des poids de sondage

5.6.3. Vérification de la disponibilité des variables auxiliaires de calage

5.6.4. Validation des variables auxiliaires de calage

5.6.5. Calcul des poids de calage

5.7. Calcul des indicateurs et estimation des paramètres

5.7.1. Calcul estimations

5.7.2. Calcul des éléments de précision des estimations

5.7.3. Validation des estimations

5.8. Mise au point finale des fichiers de données

5.7.1. Vérification de l'effectivité et de l'exhaustivité des traitements

5.7.2. Production du rapport sur les traitements effectués

6. ANALYSE

6.1. Elaboration des projets de produits

6.1.1. Mise au point des tableaux des estimations

6.1.2. Préparation de l'ébauche des produits

6.1.3. Mise à jour des séries chronologiques

6.1.4. Ajustements saisonniers

6.2. Validation des produits statistiques

6.2.1. Vérification de la couverture et examen des taux de non réponse

6.2.2. Comparaison des statistiques avec d'autres données

6.2.3. Recherche des incohérences

6.2.4. Vérification des macrodonnées

6.2.5. Vérification de cohérence par rapport aux renseignements sur le domaine

6.3. Interprétation et explication des données

6.3.1. Elaboration et adoption de la démarche d'analyse

6.3.2. Adoption de la présentation des données

6.3.3. Identification des sujets pertinents d'analyse

6.3.4. Identification des éléments de contexte liés aux résultats

6.3.5. Vérification de cohérence par rapport aux renseignements sur le domaine

6.4. Contrôle de la non divulgation des données individuelles

6.4.1. Identification et évaluation des risques

6.4.2. Identification de la méthode de contrôle

6.4.3. traitement/élimination des risques

6.5. Mise au point finale des produits

6.5.1. Application des derniers contrôles de cohérence

6.5.2. Assemblage de toutes les métadonnées nécessaires

6.5.3. Discussions préalables à la diffusion

6.5.4. Validation des produits

7. DIFFUSION

7.1. Mise à jour des systèmes de diffusion

7.1.1. Vérification de la disponibilité des métadonnées

7.1.2. Vérification de la conformité des métadonnées aux données à publier

7.1.3. Vérification de l'accessibilité des métadonnées

7.2. Production des produits de diffusion

7.2.1. Formatage des produits de diffusion

7.2.2. Assemblage des composantes en produits

7.2.3. Vérification et validation des produits

7.2.4. Vérification du respect des normes de publication

7.2.5. Edition et reproduction

7.3. Mise au point du processus de diffusion

7.3.1. Validation du choix du moment de la diffusion

7.3.2. Information aux groupes particuliers

7.3.3. Gestion de l'accès aux données sensibles

7.3.4. Vérification des procédures et protocoles standards de diffusion et de publication

7.4. Promotion des produits de diffusion

7.4.1. Chargement des supports de diffusion

7.4.2. Ciblage des utilisateurs

7.4.3. Présentation des produits aux utilisateurs

7.4.4. Communication et actions de promotion

7.5. Gestion du soutien aux utilisateurs

7.5.1. Communication des coordonnées utiles aux utilisateurs

7.5.2. Gestion des requêtes des utilisateurs

7.5.3. Gestion des commentaires des utilisateurs

7.5.4. Documentation

8. EVALUATION

8.1. Recueil des données d'évaluation

8.1.1. Exploitation et compilation des rapports

8.1.2. Exploitation des métadonnées et paradonnées

8.1.3. Exploitation de données d'autres sources

8.2. Conduite de l'évaluation

8.2.1. Synthèse des données d'évaluation

8.2.2. Identification des sources d'erreurs

8.2.3. Formulation des recommandations

8.2.4. Rédaction du rapport d'évaluation

8.3. Adoption du plan d'actions

8.3.1. Analyse des problèmes

8.3.2. Identification et élaboration des solutions

8.3.3. Identification des conditions de mise en application des solutions

8.3.4. Elaboration du dispositif de mise en œuvre et de suivi des mesures

Annexe 2 : Equipe de rédaction

Supervision

1. TEDOU Joseph, Directeur Général, INS
2. SHE ETOUNDI Joseph, Directeur Général Adjoint de l'INS

Coordination technique

1. OKOUDA Barnabé, Chef de Département de la Coordination Statistique, de la Coopération et de la Recherche, INS
2. KANA KENFACK Christophe, Chef de Division de la Coopération, de la Révision et de la Recherche Appliquée, INS

Equipe technique de rédaction

1. ABANDA Ambroise, INS
2. DZOSSA Anaclet Désiré, INS
3. NDEFFO Guy Ferdinand, INS
4. TCHOMTHE Séverin, INS
5. ESSAMBE BOME Vincent Ledoux, INS
6. TCHAKOUTE NGOHO Romain, INS
7. TCHOUDJA Victorien, INS
8. TIOBO'O PAPOU Sédric Edmond, INS
9. KAMGAING YOUGBISSI Léonie, INS
10. KAMGUE Max, INS
11. NTOUDA BETSOGO, MINADER
12. BOGMIS Marcel, BUCREP
13. MODOU SANDA, INS
14. TALLA Jacques, INS

Contact:

Site web: www.stat.cm

Tél. (237) 222 22 04 45

Fax. (237) 222 23 24 37

B.P. 134 Yaoundé-Cameroun