



# Manuel de procédure pour les enquêtes et Recensements de l'INSD

Ministère de l'économie,  
des finances et de la prospective

Secrétariat général  
Institut national de la statistique  
et de la démographie

BURKINA FASO  
Unité - Progrès - Justice



# Manuel de procédures pour les enquêtes et recensements de l'INSD

Réalisé avec l'appui du Programme d'appui au renforcement de la gestion des finances publiques et des statistiques (Par-Gs) financé par l'Union européenne.



Mars 2024

# AVANT PROPOS

L'Institut National de la Statistique et de la Démographie (INSD) a le plaisir de mettre à la disposition des producteurs et utilisateurs de données statistiques son manuel de procédures pour les opérations d'enquêtes et de recensements.

Ce manuel de procédures est élaboré à la suite de l'adoption du cadre d'évaluation de la qualité des statistiques produites par l'INSD. Il précise les orientations à suivre pour garantir la qualité de la production statistique lors des enquêtes et recensements. Le manuel est conçu de telle sorte que toutes les dimensions de la qualité des statistiques soient couvertes.

Cette première édition n'est pas une description détaillée des différentes phases du processus de production statistique. Il précise pour chaque phase les principales activités à réaliser et les documents factuels qui attestent de leur effectivité. Elle pourrait être améliorée lors des prochaines éditions.

L'INSD exprime sa profonde reconnaissance à tous les services qui ont contribué à son élaboration. Il formule également ses remerciements à l'endroit de tous ceux qui, d'une manière ou d'une autre, ont contribué à l'élaboration du présent manuel, en particulier l'Union Européenne qui a apporté un appui technique et financier à travers le Programme d'Appui à la Gestion Publique et aux Statistiques (PAGPS).

**Le Directeur Général**

**Boureima OUEDRAOGO**

*Chevalier de l'ordre du mérite de l'Economie et des Finances*

# SOMMAIRE

AVANT PROPOS .....	I
SOMMAIRE .....	II
LISTE DES TABLEAUX .....	III
ABREVIATIONS.....	IV
DEFINITION DES CONCEPTS .....	VI
INTRODUCTION .....	1
1 CADRE JURIDIQUE ET INSTITUTIONNEL DU PROCESSUS DE PRODUCTION DES STATISTIQUES.....	2
2 DESCRIPTION DU MODELE DE GESTION DE L'INSD.....	3
3 DESCRIPTION DU MODELE GENERIQUE DU PROCESSUS DE PRODUCTION STATISTIQUE .....	4
4 DEFINITION DES ROLES ET RESPONSABILITES.....	7
5 PHASE DE DEFINITION DES BESOINS.....	10
6 PHASE DE TRAVAUX PRELIMINAIRES .....	12
7 PHASE DE CONCEPTION.....	15
8 PHASE DE CONSTRUCTION.....	19
9 PHASE DE DEMANDE DE VISA STATISTIQUE.....	22
10 PHASE DE COLLECTE .....	24
11 PHASE DE TRAITEMENT.....	28
12 PHASE D'ANALYSE.....	33
13 PHASE DE DIFFUSION .....	35
14 PHASE D'ARCHIVAGE .....	37
15 PHASE D'EVALUATION.....	41
CONCLUSION.....	42
BIBLIOGRAPHIE .....	43
ANNEXES .....	44
TABLE DES MATIERES .....	47

## LISTE DES TABLEAUX

Tableau 1: le modèle générique de gestion de l'INSD.....	3
Tableau 2: les phases du processus de production statistique.....	4
Tableau 3 : Responsabilités des acteurs selon les phases .....	7
Tableau 4: Schéma du plan de diffusion.....	17
Tableau 5: la fréquence des archivages .....	38

## ABREVIATIONS

AC	Agence comptable
CCI-SE	Cellule de contrôle interne et de suivi-évaluation
CNS	Conseil national de la statistique
DAAF	Direction des affaires administratives et financières
DCMEF	Direction du contrôle des marchés et engagements financiers
DCSFR	Direction de la coordination statistique, de la formation et de la recherche
DD	Direction de la démographie
DGCOOP	Direction générale de la coopération
DGESS	Direction générale des études et statistiques sectorielles
DIMIS	Direction de l'informatique et du management des informations statistiques
DRH	Direction des ressources humaines
DSCVM	Direction des statistiques sur les conditions de vie des ménages
DSSE	Direction des statistiques et des synthèses économiques
DUE	Délégation de l'Union européenne
INSD	Institut national de la statistique et de la démographie
ISO	International Organization for Standardization
ISSP	Institut supérieur des sciences de la population
MAAH	Ministère de l'agriculture et des aménagements hydro-agricoles
MENAPLN	Ministère de l'éducation nationale, de l'alphabétisation et de la promotion des langues nationales
MESRSI	Ministère de l'Enseignement Supérieur, de la Recherche Scientifique et de l'Innovation
PAGS	Programme d'appui à la gestion publique et aux statistiques
PNDES	Plan national de développement économique et social
PNUD	Programme des Nations unies pour le développement
PTF	Partenaires techniques et financiers
SCEAM	Service des comptes économiques et des analyses macroéconomiques
SCFV	Service de la cartographie et du fichier villages
SCS	Service de la coordination statistique

SDS	Schéma directeur de la statistique
SERDM	Service des études, de la recherche et du développement méthodologique
SFPR	Service de la formation, du perfectionnement et du recyclage
SP	Secrétariat Permanent
SPAC	Service de la prévision et de l'analyse de la conjoncture
SPED	Service de la population et des études démographiques
SPS	Sous-programme Statistique (PAGPS)
SSN	Système statistique national

# DEFINITION DES CONCEPTS

<b>Concepts</b>	<b>Définitions</b>
<b>Enquête</b>	: une opération qui consiste à collecter des informations sur une partie des unités statistiques d'une population donnée appelée échantillon
<b>Recensement</b>	: une enquête au cours de laquelle des informations sont collectées sur toutes les unités statistiques d'une population donnée
<b>Processus de production statistique</b>	: toute activité visant à acquérir, à recueillir ou à manipuler des renseignements en vue de produire des données statistiques
<b>Besoin</b>	: lacune en produits statistiques nécessitant d'être comblée.
<b>Métadonnées</b>	: information sur les concepts, définitions et classifications sous-jacents utilisés, méthodologie de collecte et de traitement des données, et indicateurs ou mesures de l'exactitude des informations statistiques.
<b>Système statistique national</b>	le cadre regroupant les fournisseurs, les producteurs et les utilisateurs de statistiques publiques ainsi que les organes de coordination des activités statistiques et les institutions publiques de formation de statisticiens et de démographes
<b>Unité statistique</b>	unité d'observation de base, à savoir une personne physique, un ménage, une entreprise, etc., à laquelle se rapportent les données

# INTRODUCTION

L'une des missions de l'Institut national de statistique et de la démographie (INSD) est d'élaborer les outils et instruments d'analyse et d'aide à la décision, de diffuser l'information à caractère statistique, économique et démographique, et d'assurer la coordination technique et institutionnelle du Système statistique national (SSN). Conformément à sa mission, l'INSD produit et met à la disposition des utilisateurs des statistiques de qualité dans le but d'apporter une information objective pour les projets de développement dans les domaines politique, économique, social et culturel. Les statistiques permettent de définir les orientations des programmes d'intervention, de définir les objectifs, d'identifier les indicateurs permettant d'évaluer les progrès et avancements des différents programmes et projets, et d'évaluer leurs effets et résultats. Il est donc clair que la disponibilité de données de qualité contribuera fortement à l'amélioration des prises de décisions et garantira aux utilisateurs la pertinence, l'exactitude, la complétude des statistiques exploitées.

Pour garantir la qualité de ses produits statistiques, l'INSD s'est doté d'un cadre d'évaluation. Ce cadre d'assurance qualité définit un ensemble d'objectifs, de normes et principes nécessaires pour garantir la couverture de toutes les dimensions de la qualité des statistiques.

La mise en œuvre de ce cadre implique de nombreuses mutations dans le processus de production statistique. Parmi ces changements, il faut noter essentiellement :

- La capitalisation et la vulgarisation des bonnes pratiques au sein de l'INSD ;
- L'harmonisation ou l'uniformisation des concepts, des méthodes, des nomenclatures utilisées par les directions techniques de l'INSD ;
- La convergence vers l'automatisation de toutes les étapes du processus de production statistique ;
- La garantie de la transparence tout au long du processus de production statistique ;
- etc.

L'atteinte de tous ces objectifs nécessite l'élaboration d'un manuel de procédures pour les enquêtes et les recensements pilotés par l'INSD.

Ce présent manuel se veut être un guide pratique général pour la conception et la conduite d'enquêtes et recensements sur n'importe quel type d'unités statistiques étudiées (individu, ménage, entreprise, ...). L'objectif principal est de réunir dans un seul et même document les principales indications relatives à la conception et la conduite d'enquêtes et de recensements afin de collecter, traiter et publier des statistiques utiles, d'une qualité et d'une exactitude suffisantes. C'est pourquoi ce manuel utilise des méthodes rationnelles fondées sur la théorie statistique impliquant l'utilisation d'un échantillonnage probabiliste à chacune des phases du processus de sélection.

De façon spécifique, le manuel définit, pour toutes les étapes du processus de production :

- Les tâches à exécuter ;
- Le niveau de responsabilité ;
- Les différentes étapes de traitement ;
- Les lieux de réalisation ;
- Le mode d'exécution.

Le manuel est structuré conformément aux phases nécessaires pour conduire une enquête par échantillonnage ou un recensement en partant de la définition des besoins de statistiques jusqu'à la phase de l'évaluation de la qualité des données. Chaque phase du processus de production est décomposée en des sous-phases et les sous-phases sont accompagnées d'orientations techniques qui contribuent à leur réalisation. Mais, ces phases du processus de production statistique sont précédées de la présentation du cadre juridique et institutionnel, de la description du modèle de gestion de l'INSD et de la description du modèle générique du processus de production statistique. Le manuel n'est pas une description détaillée des différentes phases du processus de production des statistiques. Dans chaque phase, il donne les orientations précises sur les principales activités à réaliser et les documents factuels qui attestent de leur effectivité.

# **1 CADRE JURIDIQUE ET INSTITUTIONNEL DU PROCESSUS DE PRODUCTION DES STATISTIQUES**

## **1.1. Au plan institutionnel**

Sur le plan institutionnel, Le Système statistique national (SSN) est le cadre administratif réunissant les producteurs et les utilisateurs des statistiques officielles, y compris les organes de coordination et les institutions nationales de formation de statisticiens et de démographes.

Au niveau de l'INSD, les enquêtes et les recensements sont conduits par les directions techniques centrales et les directions régionales qui travaillent, si nécessaire, en collaboration avec les autres membres du SSN.

## **1.2. Au plan juridique**

La production de statistiques à travers les enquêtes et recensements est encadrée au niveau national par les textes suivants :

La Loi n°036-2021/an du 6 décembre 2021 portant organisation et réglementation des activités statistiques. Elle définit les principes fondamentaux qui régissent les activités statistiques et définit le Système statistique national.

Le décret n° 2000-508/PRES/PM/MEF du 27 octobre 2000 portant érection de l'Institut national de la statistique et de la démographie en établissement public à caractère administratif ;

Le décret n°2015-1518/PRES-TRANS/PM/MEF du 18 décembre 2015 portant approbation des statuts particuliers de l'Institut national de la statistique et de la démographie.

Le décret 2016-27/PRES/PM/SGG-CM du 23 février 2016 portant organisation type des départements ministériels ;

Le décret n°2023-0198/PRES-TRANS/PM/MEFP du 13 mars 2023 portant organisation du Ministère de l'Economie, des Finances et de la Prospective ;

L'arrêté n°2023-00307/MEFP/SG/INSD portant attributions, organisation et fonctionnement de l'Institut national de la statistique et de la démographie (INSD).

L'arrêté n°2020-00553/MINEF/SG/INSD portant attributions, organisation et fonctionnement de l'Institut national de la statistique et de la démographie (INSD) ;

Au niveau international, le Burkina Faso a ratifié la charte africaine de la statistique qui guide aussi les productions statistiques.

## 2 DESCRIPTION DU MODELE DE GESTION DE L'INSD

D'après l'organigramme<sup>1</sup> de l'INSD, on peut reconstituer le modèle générique de gestion de la façon suivante :

**Tableau 1: le modèle générique de gestion de l'INSD**

Rôle dans la gestion	Acteur de l'INSD
<b>Stratégie et direction</b>	
Définir la vision	DG, Comité stratégique
Administrer et diriger	CA, Conseil de direction
Gérer la coopération et la collaboration stratégiques	SCS
<b>Gestion des capacités</b>	
Planifier le renforcement des capacités	SFPR, DRH
Élaborer le renforcement des capacités	SFPR
Suivre le renforcement des capacités	SFPR, DRH
Appuyer la mise en œuvre du renforcement des capacités	DRH
<b>Services d'appui</b>	
Gérer l'organisme et les performances	CCI-SE
Gérer les finances	DAAF, AC, PRM
Gérer les ressources humaines	DRH
Gérer les technologies de l'information	SITGBD
Gérer la méthodologie statistique	SERDM
Gérer l'information et les connaissances	SERDM, SDA
Gérer les consommateurs	Service courrier, SMIS, SDA, DTS <sup>2</sup>
Gérer les fournisseurs de données	DTS
Gérer les bâtiments et les locaux	SEI
Gérer la qualité	SERDM

<sup>1</sup> L'arrêté n°2023-00307/MEFP/SG/INSD portant attributions, organisation et fonctionnement de l'Institut national de la statistique et de la démographie (INSD).

<sup>2</sup> Par DTS, on entend les 3 directions centrales sectorielles, à savoir la DD, la DSCVM et la DSSE.

### 3 DESCRIPTION DU MODELE GENERIQUE DU PROCESSUS DE PRODUCTION STATISTIQUE

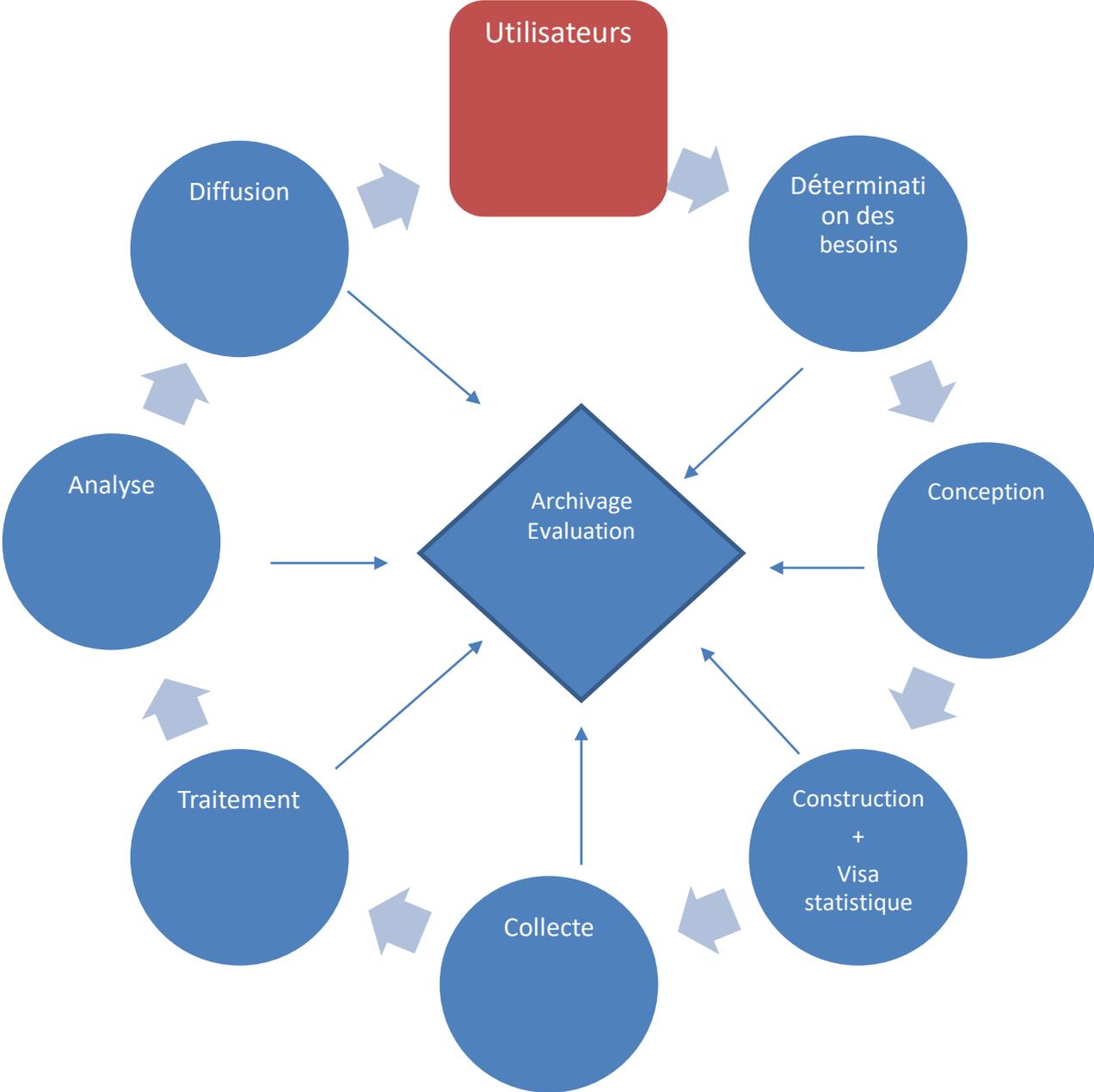
Le processus de production statistique lors des opérations d'enquêtes et recensements suit en effet un cycle, toujours le même, qui va de la détermination des besoins des utilisateurs à leur satisfaction par la diffusion des statistiques de qualité qu'ils souhaitaient. Au total, 11 étapes fondamentales sont nécessaires pour satisfaire les utilisateurs des statistiques. Chaque phase comprend un ensemble de sous-étapes ou tâches qui concourent à sa réalisation. Le tableau Tableau 2 retrace les différentes phases et les tâches associées.

**Tableau 2: les phases du processus de production statistique**

Phases	Sous-phases
I. Définition des besoins	Détermination des besoins d'information
	Elaboration du répertoire des utilisateurs potentiels
	Consultation des utilisateurs sur les besoins
	Validation des besoins des utilisateurs
	Etablissement des objectifs en termes de produits (indicateurs)
	Identification des concepts et des variables
	Vérification de la disponibilité des données
II. Travaux préparatoires	Elaboration de l'étude de faisabilité
	Elaboration du protocole de l'opération /document de projet/ Convention
	Elaboration des plans de tabulation et des plans d'analyse
	Mobilisation des ressources humaines, financières et matérielles
	Planification des activités
III. Phase de conception	Elaboration du calendrier de l'opération
	Sensibilisation
	Conception des produits/ indicateurs
	Élaboration de la description des variables
	Définition de la méthode de collecte des données
	Définition de la méthodologie d'échantillonnage
	Conception de la méthodologie de traitement statistique (codification, apurement et tabulation)
	Conception de l'analyse des données
Conception du plan de diffusion	
IV. Phase de construction	Conception du plan d'archivage
	Elaboration ou actualisation des outils de collecte (manuel, questionnaire, masque de saisie ou application de collecte)
	Elaboration des programmes de traitements (programme de codification, programmes d'apurement de données, programme de construction des indicateurs, programme de tabulation, plan d'analyse)
	Test du système de production (test des outils de collecte)
	Test du processus de production (test des programmes de traitements)
V. Phase de demande	Finalisation du système de production (finalisation des outils de collecte et programmes de traitements)
	Constitution ou actualisation de la base de sondage
	Préparation du dossier de visas statistique

<b>Phases</b>	<b>Sous-phases</b>
de visa statistique	Envoi du dossier
	L'obtention du visa statistique
VI. Phase de collecte	Tirage de l'échantillon
	Mise à jour de l'échantillon
	Mise en place du processus de collecte (recrutement, formation, déploiement du personnel de collecte)
	Collecte des données
	Finalisation de la collecte des données
VII. Phase de traitement	Intégration des données
	Classification et codification des données
	Contrôle, qualification, redressement des données
	Imputation des données
	Dérivation des nouvelles variables et unités statistiques
	Calcul des pondérations
	Calcul des agrégats
VIII. Phase d'analyse	Préparation de la version préliminaire des résultats
	Validation des résultats
	Analyse descriptive des données
	Analyse explicative des données (en cas de besoin)
	Application des contrôles de confidentialité
	Finalisation des résultats
IX. Phase de diffusion	Mettre à jour les systèmes de diffusion
	Elaboration des produits de diffusion
	Gestion de la sortie des produits de diffusion
	Promotion des produits de diffusion
	Gestion l'assistance aux clients
X. Phase d'archivage	Définition des règles d'archivage
	Gestion de l'entrepôt des archives
	Conservation des données et les métadonnées associées
XI. Phase d'évaluation	Recueil des contributions d'évaluation
	Réalisation de l'évaluation
	Recherche de consensus sur un plan d'actions

Processus statistique d'une Enquête



## 4 DEFINITION DES ROLES ET RESPONSABILITES

Dans toute opération statistique, enquête et recensement en particulier, la définition du rôle des acteurs dans chaque sous-phase est importante pour sa réussite dans les délais impartis. Le tableau donne les rôles des acteurs suivant toutes les étapes de l'opération.

**Tableau 3 : Responsabilités des acteurs selon les phases**

Phases	Sous-phase	Responsabilité
<b>I. Définition des besoins</b>	·Détermination des besoins d'information	Tous les services porteurs du dossier et/ou le commanditaire
	Elaboration du répertoire des utilisateurs potentiels	Tous les services porteurs du dossier
	Consultation des utilisateurs sur les besoins	SERDM
	Validation des besoins des utilisateurs	Tous les services porteurs du dossier, le commanditaire de l'enquête
	Etablissement des objectifs en termes de produits (indicateurs)	Tous les services porteurs du dossier
	Identification des concepts et des variables	Tous les services porteurs du dossier
	Vérification de la disponibilité des données	Tous les services porteurs du dossier
	Elaboration de l'étude de faisabilité	Tous les services porteurs du dossier
<b>II. Travaux préparatoires</b>	Elaboration du protocole de l'opération/document de projet/ Convention	Service concerné + DAAF
	Elaboration des plans de tabulation et des plans d'analyse	
	Mobilisation des ressources humaines, financières et matérielles	Service concerné + DAAF + DG
	Planification des activités	Service concerné + DAAF
	Elaboration du calendrier de l'opération	Service concerné + DAAF
	Sensibilisation	
<b>III. Phase de conception</b>	Conception des produits/ indicateurs	Service concerné + SERDM
	Élaboration de la description des variables	Service concerné
	Définition de la méthode de collecte des données	Service concerné + SERDM
	Définition de la méthodologie d'échantillonnage	Service concerné + SERDM
	Conception de la méthodologie de traitement statistique (codification, apurement et tabulation)	Service concerné + DIMIS
	Conception de l'analyse des données	Service concerné
	Conception du plan de diffusion	DIMIS + service concerné
	Conception du plan d'archivage	DIMIS(SDA) + service concerné
<b>IV. Phase de construction</b>	Elaboration ou actualisation des outils de collecte (manuel, questionnaire, masque de saisie ou application de collecte)	Service concerné + DIMIS (SITGBD)

Phases	Sous-phase	Responsabilité
	Élaboration des programmes de traitements (programme de codification, programmes d'apurement de données, programme de construction des indicateurs, programme de tabulation, plan d'analyse)	Service concerné + DIMIS (SITGBD)
	Test du système de production (test des outils de collecte)	Service concerné + DIMIS
	Test du processus de production (test des programmes de traitements)	Service concerné + DIMIS
	Finalisation du système de production (finalisation des outils de collecte et programmes de traitements)	Service concerné + DIMIS
	Constitution ou actualisation de la base de sondage	Service concerné +DD(SCFV)
<b>V. Phase de demande du visa statistique</b>	Préparation du dossier de visas statistique	Service concerné
	Envoie du dossier	Service concerné + DG
	L'obtention du visa statistique	Service concerné + DG
<b>VI. Phase de collecte</b>	Tirage de l'échantillon	Service concerné + SERDM
	Mise à jour de l'échantillon	Service concerné
	Mise en place du processus de collecte (recrutement, formation, déploiement du personnel de collecte)	Service concerné +DRH +DG
	Collecte des données	Service concerné
	Finalisation de la collecte des données	
<b>VII. Phase de traitement</b>	Intégration des données	Service concerné + DIMIS(SITGBD)
	Classification et codification des données	Service concerné + DMIS
	Contrôle, qualification, redressement des données	Service concerné + DMIS
	Imputation des données	Service concerné + DMIS
	Dérivation des nouvelles variables et unités statistiques	Service concerné + DMIS
	Calcul des pondérations	Service concerné + DMIS +SERDM
	Calcul des agrégats	Service concerné
<b>VIII. Phase d'analyse</b>	Préparation de la version préliminaire des résultats	Service concerné
	Validation des résultats	Service concerné
	Analyse descriptive des données	Service concerné
	Analyse explicative des données (en cas de besoin)	Service concerné++SERDM
	Application des contrôles de confidentialité	Service concerné
	Finalisation des résultats	Service concerné
<b>IX. Phase de diffusion</b>	Mettre à jour les systèmes de diffusion	Service concerné + DIMIS
	Elaboration des produits de diffusion	Service concerné + DIMIS
	Gestion de la sortie des produits de diffusion	Service concerné + DIMIS
	Promotion des produits de diffusion	Service concerné + SDA
	Gestion l'assistance aux clients	Service concerné + DIMIS
<b>X. Phase d'archivage</b>	Définition des règles d'archivage	Service concerné + DIMIS
	Gestion de l'entrepôt des archives	DIMIS

<b>Phases</b>	<b>Sous-phase</b>	<b>Responsabilité</b>
	Conservation des données et les métadonnées associées	DIMIS
<b>XI. Phase d'évaluation</b>	Recueil des contributions d'évaluation	Service concerné
	Réalisation de l'évaluation	SERDM
	Recherche de consensus sur un plan d'action	Directeur technique + Direction G +SERDM+ Service concerné

## **5 PHASE DE DEFINITION DES BESOINS**

La première étape de tout processus de production statistique, les enquêtes et recensements en particulier, consiste à en déterminer les principaux objectifs qui émanent impérativement de l'expression des besoins divers de statistiques. La détermination des besoins en statistiques s'impose à chaque opération afin de prendre en compte les besoins de nouvelles statistiques ou les préoccupations particulières relatives aux statistiques actuelles. Cette phase de définition des besoins comporte huit sous phases.

### **5.1. Détermination des besoins d'information**

Dans cette sous-étape, plusieurs tâches sont à exécuter : Il s'agit de faire un état des lieux des besoins de statistiques auxquels l'opération statistique doit répondre. Au niveau national, les besoins de statistiques peuvent provenir des plans, projets et programmes nationaux de développement (ENP 2025, le SNADDT, le RND, Les PRD et PCD), les politiques sectorielles issues des différents départements ministériels. Au niveau international, l'activité statistique peut répondre aux besoins de formulation et d'évaluation des ODD des Nations unies, l'agenda 2063 de l'UA ...

A cette sous-étape, il faut également déterminer et examiner l'existence des données nécessaires pour répondre aux besoins et capitaliser les bonnes pratiques d'autres instituts de statistiques. Il faut identifier et déterminer les besoins particuliers de divers groupes spécifiques d'utilisateurs comme les personnes vivant avec un handicap, les femmes, les enfants, les personnes âgées, les autres personnes vulnérables ...

### **5.2. Elaboration de la liste des utilisateurs potentiels**

Les besoins de statistiques varient fondamentalement selon les différents types d'utilisateurs. Ainsi, il apparaît nécessaire d'élaborer une liste des utilisateurs potentiels (utilisateurs institutionnels) des principales statistiques à produire dans le cadre de chaque enquête ou recensement. La liste des structures membres du SSN, le répertoire des PTF et l'enquête satisfaction peuvent contribuer à l'élaboration de la liste des utilisateurs potentiels.

### **5.3. Consultation des utilisateurs sur les besoins**

Sur la base de la liste des utilisateurs potentiels, il faut organiser des consultations nationales élargies aux producteurs de statistiques semblables, aux PFT, aux centres de recherche... Ces consultations permettent non seulement de clarifier et de confirmer leurs besoins dans les moindres détails, mais également de savoir où, quand, comment et pourquoi il faut produire les données.

### **5.4. Validation des besoins des utilisateurs**

La revue de la littérature et les larges consultations conduisent à une panoplie de besoins de statistiques. Dans cette sous-étape, il faut les examiner et les valider. Les besoins non retenus doivent être motivés et portés à la connaissance des demandeurs. Dans tous les cas, il faut tenir compte de l'impératif équilibre entre les mesures de qualité recherchée et les ressources disponibles pour l'activité. Il faut prendre en compte les contraintes opérationnelles telles que la période de référence, les coûts et les méthodes d'acquisition des données.

### **5.5. Etablissement des objectifs et des produits**

À cette sous-étape, il faut transformer les besoins validés de statistiques en objectifs clairs, à assigner à l'enquête ou au recensement. Les objectifs de l'opération permettent d'aboutir à une liste d'indicateurs à calculer à l'issue de l'opération. Dans cette sous-étape, il faut établir une liste provisoire d'indicateurs accompagnés de leurs métadonnées.

## 5.6. Identification des concepts et des variables

A cette sous-étape, les concepts, les nomenclatures et les variables susceptibles d'être utilisés doivent être identifiés.

## 5.7. Vérification de la disponibilité des données

Cette sous-étape consiste à vérifier la disponibilité des données susceptibles de répondre aux besoins validés des utilisateurs. Les données disponibles peuvent aussi servir à des données intermédiaires. Dans tous les cas, les éventuelles données disponibles doivent faire l'objet d'une évaluation pour savoir si elles sont d'une qualité minimale et si elles répondent aux besoins et aux objectifs exprimés. Les sources de données nationales sont : le site de l'INSD ([www.insd.bf](http://www.insd.bf)), le site du CNS ([www.cns.bf](http://www.cns.bf)), les sites web des différents départements ministériels...

## 5.8. Elaboration de l'étude de faisabilité

Cette sous-étape consiste à documenter et à analyser tous les types possibles de processus de production pour obtenir les données statistiques souhaitées et à déterminer, en fonction des ressources disponibles, celui qui est le plus apte à répondre aux besoins des utilisateurs ou des clients. Généralement présenté sous forme de document d'analyse de rentabilisation, le processus de production statistique choisi devra satisfaire à plusieurs normes préétablies et être approuvé par la haute gestion avant d'être mis en œuvre.

### Conclusion partielle

Cette étape importante « définition des besoins » du processus de production statistique contribue à améliorer la pertinence des données statistiques, une des dimensions de la qualité. Pour cette raison, la bonne conduite de cette étape doit être vérifiée par les éléments factuels. Ainsi, à la fin de cette phase, les documents suivants doivent être élaborés et disponibles :

- la liste des utilisateurs potentiels des principales statistiques ;
- la liste des besoins de statistiques à satisfaire ;
- le rapport détaillé des consultations ;
- la liste des indicateurs clés et leurs métadonnées ;
- le rapport d'analyse de rentabilisation du processus établissant l'écart existant entre les besoins des utilisateurs et les résultats visés.

## **6 PHASE DE TRAVAUX PRELIMINAIRES**

### **6.1. Elaboration du protocole de l'opération**

Le protocole décrit les différentes phases d'une étude, les règles qui seront observées dans les relations entre l'équipe, les pouvoirs publics et la population concernée par l'étude. Ainsi, le protocole de recherche est un engagement, un cahier de charges, une convention, un guide qui dicte les règles de conduite pendant tout le processus de l'étude. Cette phase comporte quatre sous étapes.

#### **6.1.1 Contexte et justification**

Cette partie a pour rôle de convaincre l'utilisateur de l'utilité ou de l'intérêt de réaliser l'étude pour résoudre un problème. Le contexte doit décrire le cadre géographique, démographique, sociologique, politique ou économique de l'étude. La justification doit permettre d'expliquer pourquoi avoir choisi de s'intéresser à une thématique et une population particulière en précisant l'importance du sujet. Il faut aussi préciser les buts et résultats attendus de cette opération.

#### **6.1.2 Objectifs de l'étude**

La formulation des objectifs se fait en deux phases :

L'objectif général qui fixe un résultat global à atteindre.

Les objectifs spécifiques qui sont formulés avec une précision maximale permettant d'identifier les moyens exacts pour les atteindre. Ils permettent de savoir quelles informations on va chercher à collecter pour atteindre les résultats attendus de l'étude.

#### **6.1.3 Méthodes**

##### **6.1.3.1 Identification de la population/groupe cible**

Un groupe cible est une population précisément définie qui peut ainsi faire l'objet d'une intervention ou d'une étude spécifique. Il peut s'agir d'un groupe spécifique de la population qui partage des caractères communs, ou d'une population plus générale. Il faut ainsi décrire les critères d'inclusion des unités statistiques dans la population cible.

##### **6.1.3.2 Échantillonnage**

L'échantillonnage consiste à constituer un sous-ensemble de la population cible qui soit le plus représentatif possible.

L'équipe technique devra :

- définir la base de sondage ;
- définir le plan de sondage avec précision ;
- préciser la taille de l'échantillon.

Avant de commencer l'enquête, il faut définir la taille souhaitée de l'échantillon, c'est-à-dire le nombre de répondants minimum à interroger afin de pouvoir généraliser les résultats obtenus.

##### **6.1.3.3 Outils de collecte**

Il s'agit du nombre et types d'outils de collecte (questionnaire, manuel, document de méthodologie...)

#### **6.1.4 Ressources humaines**

Il s'agit de définir le nombre, la qualité du personnel nécessaire pour la conduite de l'opération, à toutes les phases.

#### **6.1.5 Ressources matérielles**

Il s'agit de définir le nombre, la qualité de toutes les ressources nécessaires pour la conduite de l'opération, à toutes les phases.

#### **6.1.6 Budget et plan de financement**

Il est nécessaire d'élaborer, de défendre et de mobiliser un budget réel précisant la totalité des coûts (en personnel, équipement, frais de fonctionnement, locaux...) engendrés par l'étude avant de s'engager sur le terrain. Les ressources existantes, leur valeur financière et leur origine doivent être identifiées.

Il est aussi nécessaire de préciser les dépenses prévues pour chaque ligne du programme, ce qui permet aux financiers d'échelonner leur "effort" de soutien.

#### **6.1.7 Chronogramme**

Une fois que toutes les activités à réaliser sont identifiées, il faut évaluer leur durée et les séquencer convenablement. Ce travail va aboutir au chronogramme de l'opération qu'il faut suivre scrupuleusement.

#### **6.1.8 Validation du protocole par un comité d'éthique**

Pour les enquêtes et recensements qui intègrent un volet santé, le protocole doit être soumis à la validation d'un comité d'éthique de la santé.

#### **6.1.9 Validation du protocole par la commission d'informatique et de liberté**

Pour les enquêtes et recensements qui intègrent des aspects sensibles à l'éthique nationale sur les libertés, le protocole doit être soumis à la validation de la commission d'informatique et de liberté.

### **6.2. Elaboration du document de projet/ convention**

Le service porteur du dossier en collaboration avec les autres services, élaborent un document de projet. Le document de projet répond à deux objectifs différents :

- guider la mise en œuvre du projet avec un minimum de précision. Il servira en outre de référence pour le suivi du projet et son évaluation ;
- convaincre un bailleur de financer le projet.

Un document de projet doit contenir :

- ▶ Une description du projet, résumant le diagnostic réalisé et la solution technique et organisationnelle retenue
- ▶ Un résumé synthétique du projet (par ex. un cadre logique simplifié)
- ▶ Le dispositif de pilotage et de mise en œuvre du projet
- ▶ Le chronogramme du projet
- ▶ Le budget

Ce document de projet servira de support pour la signature de convention pour les études commanditées.

### **6.3. Elaboration du document d'une convention**

Pour les études commanditées par les partenaires (UEMOA, Banque mondiale, etc.), l'élaboration d'une convention est requise. Ce document décrit sans ambiguïté les accords entre les parties prenantes, précisent les conditions de collaborations, les obligations de chaque acteur. L'élaboration d'un tel document exige les services d'un expert juriste qui est le garant de la conformité des différentes activités de l'entreprise vis-à-vis de la loi, la rédaction des articles du protocole pour assurer une sécurité à l'INSD en cas de litige avec un partenaire.

### **6.4. Mobilisation des ressources financières**

Cette phase vise à mobiliser les ressources nécessaires à la réalisation de l'opération statistique. Elle est essentielle car elle conditionne la suite du processus.

### **6.5. Planification des activités**

Il s'agit de recenser toutes les activités à réaliser avec leur durée, leur agencement.

### **6.6. Élaboration du calendrier de l'opération**

Le calendrier de travail établit un chronogramme des activités à mener en précisant l'ordre des priorités, le type d'activité, la période de réalisation, le responsable de l'activité et les moyens de vérification ou d'évaluation. Le calendrier de l'étude est un tableau de bord qui prévoit la réussite dans les délais souhaités. Le calendrier de l'étude conditionne le budget de l'étude.

Le temps nécessaire pour réaliser une opération statistique dépend des objectifs, de la population cible, de la répartition géographique et des contraintes topographiques et logistiques d'accès aux unités tirées. Toutes les étapes de l'opération doivent être planifiées. Le diagramme de Gant peut être utile pour la représentation du calendrier.

Le calendrier doit être publié et toute modification ultérieure doit être notifiée.

## **Conclusion partielle**

Les travaux préliminaires conditionnent la suite effective des activités de l'opération. Il est donc nécessaire de les mener convenablement. A la fin de cette phase, les documents suivants doivent être élaborés et disponibles :

- le protocole de l'opération ;
- le document de projet de l'opération ;
- la convention ;
- le budget détaillé de l'opération accompagné de son plan de financement ;
- le calendrier détaillé de l'opération.

## **7 PHASE DE CONCEPTION**

La conception constitue une phase importante dans le processus de production statistique. Elle est la phase créative qui consiste à décrire les activités, les tâches et les recherches existantes pour parvenir aux produits statistiques clés, aux concepts, aux variables, aux méthodologies, aux canaux d'acquisition des données et aux processus opérationnels. Elle comprend huit sous-phases.

### **7.1. Conception des produits/ indicateurs**

Cette sous-phase commence par une validation définitive des listes des besoins d'informations issues des concertations sectorielles en prenant en compte les exigences de l'étude de faisabilité. Ensuite, il faut établir la liste exhaustive des produits/indicateurs clés qui permettent de répondre aux besoins validés. Cette liste d'indicateurs est accompagnée de leurs métadonnées détaillées (confère canevas en annexe).

### **7.2. Élaboration de la description des variables**

Les métadonnées des indicateurs/produits clés précisent les principales variables nécessaires pour les calculer. Dans cette sous-phase, il faut les définir clairement sans ambiguïté. A cet effet, il faut impérativement se référer, si possible, aux concepts, normes et classifications nationales, régionales et internationales.

Dans tous les cas, la description des variables comprend :

- la définition opérationnelle des variables comprenant la période de référence de l'enquête, les unités d'intérêt ;
- le type de variables (qualitative, quantitative) ;
- le type de mesure des variables (nombre, quantité, taux, proportion, ratio, ...)
- les modalités harmonisées ou standardisées des variables ;
- les normes et classifications utilisées ;
- l'utilité de la variable.

### **7.3. Définition de la méthode de collecte des données**

La collecte des données est le processus qui permet d'obtenir l'information nécessaire sur chaque unité statistique sélectionnée de l'enquête. Les principales méthodes de collecte de données sont : les enquêtes par auto administration, les interviews directes, les interviews téléphoniques et les collectes en ligne.

Le choix de la méthode de collecte doit être objectivement motivé et comprendre également la conception de toute entente formelle se rapportant à l'acquisition de données, comme les protocoles d'entente et la confirmation du fondement juridique pour l'acquisition des données.

La méthode de collecte des données choisie devrait donner un taux de participation élevé et les données obtenues devraient être les plus complètes et précises possibles. De même, la méthode devrait aussi minimiser le fardeau pour les répondants et tenir compte du budget et des limites opérationnelles du client.

Le choix de la méthode de collecte varie en fonction de plusieurs paramètres :

- la disponibilité dans la base de sondage d'informations sur la variable d'intérêt ;
- les caractéristiques de la population cible ;
- le type de questions à poser ;
- les ressources disponibles ;

- la facilité à remplir le questionnaire ;
- les considérations sur la vie privée ;
- le temps imparti ;
- les exigences de qualité, pertinence des données.

## **7.4. Définition de la méthodologie d'échantillonnage**

La planification de l'opération et les études de rentabilisation doivent aboutir au choix d'un recensement ou d'une enquête. Dans le cas d'une enquête, il faut prévoir donc comment sélectionner l'échantillon c'est-à-dire le sous-ensemble d'unités dans une population statistique aux fins de la collecte de l'information sur ces unités pour réaliser des inférences sur l'ensemble de la population. Plusieurs méthodes statistiques permettent de choisir une technique d'enquête sur un phénomène auprès d'un échantillon d'individus sélectionné selon des règles scientifiques pour représenter toute la population. De façon générale, il existe deux grands groupes : les méthodes probabilistes et les méthodes non probabilistes. La méthodologie choisie doit être expliquée et clairement motivée.

## **7.5. Conception de la méthodologie de traitement statistique (codification, apurement et tabulation)**

Cette sous-phase consiste à identifier les méthodes statistiques solides pour détecter les incohérences internes et externes afin d'apurer les données. Il s'agit essentiellement des méthodes techniques de :

- codification des données ;
- vérification de la conformité des données ;
- détection et correction d'erreurs relevées dans les données ;
- imputation des données manquantes totales ou partielles ;
- désaisonnalisation des données, si nécessaire ;
- déflation ou l'étalonnage, si nécessaire ;
- modélisation de la variable dépendante ;
- validation et de la mise au point des données finales ;
- tabulation des données ;
- ...

## **7.6. Conception de l'analyse des données**

Il s'agit de choisir et de motiver la méthode d'analyse appropriée à utiliser sur les données définitives. Ensuite, il faut décrire la méthodologie de l'analyse données (analyses descriptives, exploratoire, explicative ...)

## **7.7. Conception du plan de diffusion**

Cette sous-phase consiste à planifier la distribution des données de l'enquête aux utilisateurs par l'intermédiaire de divers canaux un communiqué, une interview radio ou télédiffusée, une réponse téléphonique ou télécopiée à une demande spéciale, la publication d'un document, une microfiche, un média électronique, y compris Internet, ou un fichier de microdonnées sur CD, etc.

Les données collectées et compilées par les services de statistiques sont rendues publiques. Dans le but d'assurer une bonne diffusion des données statistiques, l'INSD a mise en place une stratégie de gestion de ses données en s'appuyant sur les points suivants :

Le plan de diffusion définit l'organisation des différentes publications et leur support en fonction des publics cibles visés.

### Schéma du plan de diffusion

Ce schéma présente l'adéquation des différents types de publication aux préférences des groupes d'utilisateurs. Le nombre de "+" traduit la préférence.

**Tableau 4: Schéma du plan de diffusion**

	Publication synthétique	Analyses détaillées	Tableaux de résultats	Micro-données	Exploitation ad hoc des données
Décideurs publics et privés	+++	+	+++		+
Responsables des suivi-évaluation des politiques publics (1)		+++	+++		++
Universitaires et chercheurs		+++	+++	++	+
Grand public et média	+++	+			
Organismes internationales		++	+		+++

Le plan de diffusion doit également définir le rôle joué par internet dans la diffusion des données. En effet ce media est de plus en plus utilisé pour la diffusion des données statistiques. Mais son utilisation ne pose pas de problèmes aux utilisateurs et aux producteurs.

### ***Les publics***

Les utilisateurs de données statistiques ont des besoins assez différenciés selon leur type, aussi bien en ce qui concerne le contenu de l'information que le support de diffusion à utiliser.

L'expérience permet de regrouper les utilisateurs en 5 publics différents :

- Les décideurs publics et privés : ils ont besoin d'une information synthétique et d'une présentation des grandes tendances et des évolutions récentes ;
- Les responsables du suivi et de l'évaluation des politiques publiques : ils ont besoin des indicateurs statistiques qui ont été définis pour le suivi des différents programmes (CSLP, appui budgétaire, programmes sectoriels) ;
- Les universitaires et chercheurs : ils ont un besoin de données assez détaillées et ce sont les principaux utilisateurs des micro-données.
- Le grand public et les médias : ils ont besoin d'une information synthétique et d'explications très pédagogiques sur les phénomènes observés.
- Les organismes internationaux : outre leur rôle de bailleurs, ce sont aussi des collecteurs de données statistiques assez détaillées qu'ils rediffusent dans le cadre de comparaisons internationales.

### **Conclusion partielle**

A la fin de la phase de conception, les documents suivants doivent être disponibles :

- la liste exhaustive des produits/indicateurs de l'enquête/recensement assortie de leurs métadonnées ;
- la liste exhaustive des variables à collecter accompagnée d'une description des caractéristiques et de leur utilisation ;
- le document de méthodologie de l'enquête/recensement comprenant les justifications des choix des méthodes de collecte et méthodes d'échantillonnage ;
- le plan de diffusion des données ;
- le plan d'archivage des données

- le plan d'analyse.

## **8 PHASE DE CONSTRUCTION**

L'étape de la construction consiste à élaborer les outils, le processus et le système de collecte et à les mettre à l'essai de façon pratique sur le terrain. Elle se base sur les produits annoncés à l'étape de la conception. Si les outils existent, ceux-ci doivent être préconisés quitte à y apporter des modifications nécessaires. Cela permet d'améliorer la comparabilité des données issues de différentes opérations de collecte. Dans le cas contraire, ces outils devront être élaborés en se conformant aux directives contenues dans les six sous phases qui suivent.

### **8.1. Elaboration ou actualisation des outils de collecte**

Les outils de collecte à concevoir dans le cadre d'une enquête ou d'un recensement sont : les questionnaires, le masque de saisie/application de collecte et les manuels.

#### **8.1.1 Elaboration du questionnaire**

Le questionnaire doit être axé sur le sujet de l'étude, être le plus bref possible et comporter des questions qui suivent un ordre logique. Il ne doit contenir que des questions pertinentes à l'atteinte des objectifs mentionnés. Il doit ainsi être conçu sur la base de la liste exhaustive des variables à collecter établie lors de la phase de conception. Dans la mesure du possible, il faudra veiller à utiliser les concepts et la terminologie déjà en usage ou le contenu harmonisé. Au besoin, il faudra utiliser les questions déjà élaborées pour assurer la cohérence et la comparabilité par rapport aux résultats d'enquêtes similaires qui ont déjà mesurées adéquatement les mêmes concepts.

Le questionnaire doit être professionnel, attrayant, convivial et facile à administrer ou à remplir pour le répondant. Il doit comporter une page de garde où figurent le logo et le timbre de l'INSD, le titre de l'opération, la référence de la loi statistique (obligation de réponse et confidentialité), les références du visa statistique, la référence de l'attestation du comité d'éthique de la santé (si disponible). La page suivante doit être réservée à une présentation synthétique de l'opération statistique (le contexte, les objectifs, les périodes de collecte) et le consentement. Les pages suivantes sont dédiées aux différentes sections spécifiques à l'étude. Elles doivent contenir des entêtes à chaque section ainsi que des instructions courtes et précises à l'attention du répondant ou de l'enquêteur. Les sauts et filtres doivent y être clairement indiqués. Une traduction du questionnaire dans les principales langues du pays peut s'avérer utile. Enfin, il doit être prévu à la dernière page, une page pour les éventuelles observations et une formule de salutation et remerciement.

#### **8.1.2 Elaboration du masque de saisie/application de collecte**

Le masque de saisie est une application permettant d'introduire les données à l'ordinateur. Conçu avec un logiciel de saisie de données adéquat parmi les plus connus (CSPRO, ODK, Kobo collect, Surveys solutions, Excel, Access, etc.). Il doit être à l'image du questionnaire physique : les sections, les questions, les modalités et les sauts doivent être identiques à ceux du questionnaire physique. Le logiciel choisi doit offrir la possibilité d'implémenter des procédures de contrôle permettant de déceler les incohérences et faire des sauts automatiques. Le déploiement des masques dans des smartphones ou tablettes pour une saisie directe sur le terrain (collecte mobile) doit être préconisé.

#### **8.1.3 Elaboration des manuels**

Le manuel constitue un guide de référence dans lequel sont consignées les directives concernant le travail de chaque catégorie de personnel de collecte. Il doit comporter une page de garde où figurent le logo, le timbre de l'INSD et le titre de l'opération. Les pages suivantes sont réservées aux éléments suivants : la définition des termes clés, les objectifs et le fondement légal de l'opération, les rôles, tâches, interdits et sanctions du personnel de collecte, les consignes de remplissage pour chaque question (pour le manuel des agents enquêteurs/recenseurs), les consignes de contrôle (pour le manuel du contrôleur) et de supervision (pour le manuel du superviseur) puis en annexe le questionnaire, les nomenclatures, les catalogues d'images pour certaines modalités du questionnaire et les canevas des rapports.

## **8.2. Elaboration des programmes de traitements**

Les principaux programmes de traitements à réaliser sont : les programmes de codification, programmes d'apurement de données, les programmes de construction des indicateurs, les programmes de tabulation. L'ensemble de ces programmes doivent être conçus sous un logiciel adapté (SPSS, Stata, R, etc.).

- les programmes de codification permettent d'affecter des codes aux réponses sur les questions ouvertes et semi ouvertes. Ils peuvent permettre aussi de recodifier les questions fermées en regroupant ou en éclatant certaines modalités ;
- les programmes d'apurement de données permettent de détecter et corriger les incohérences ;
- les programmes de construction des indicateurs permettent de calculer les indicateurs ;
- les programmes de tabulation permettent de sortir les tableaux et graphiques nécessaires aux analyses ;

## **8.3. Elaboration du plan d'analyse**

Cette sous-phase consiste à élaborer le plan d'analyse détaillé de chaque thématique abordée par l'opération statistique. Les plans d'analyse doivent utiliser toutes les variables contenues dans le questionnaire. Au cas échéant, les variables non retenues doivent être justifiées. Chaque sous-point du plan d'analyse contient les variables à utiliser. Le plan d'analyse doit être accompagné de programmes conçus sous les logiciels statistiques (SPSS, Stata, R, etc.).

Il doit comporter les jeux de données à utiliser, les critères d'inclusion / d'exclusion, les variables à utiliser pour l'analyse principale, les méthodes statistiques et logiciels à utiliser, la liste des tableaux et graphiques à réaliser, les maquettes des tableaux et l'estimation du temps et des ressources nécessaires pour les analyses.

Dans cette sous-phase, il faut également élaborer le projet de recherche (confère canevas en annexe) de l'analyse approfondie de la variable dépendante d'intérêt de l'enquête.

## **8.4. Test du système de production**

Ce test est réalisé au travers d'une opération pilote et vise à expérimenter les dispositifs et outils et tirer les enseignements en vue de rectifier le tir le cas échéant. Il porte sur un échantillon réduit choisi de façon aléatoire ou raisonnée de manière à mettre en évidence les difficultés à surmonter lors de la collecte et des autres phases de l'opération et adopter des stratégies appropriées. Il doit être exploité intégralement afin d'en tirer bénéfice.

## **8.5. Test du processus de production**

Ce test vise à évaluer l'efficacité des programmes de traitement de données. En utilisant les données collectées lors du test du système de production, tous les programmes conçus (codification, apurement, construction des indicateurs, tabulation) doivent être mis en œuvre.

## **8.6. Finalisation du système de production**

Cette étape consiste à finaliser les outils, les procédures de collecte au regard des tests. Les lacunes constatées lors de ces tests doivent être comblées.

## 8.7. Constitution ou actualisation de la base de sondage

Dans le cadre d'un recensement, chaque unité est enquêtée alors que dans une enquête par sondage seul un échantillon d'unités l'est et les informations fournies par l'échantillon sont utilisées pour faire des estimations concernant l'ensemble des unités. Avant de procéder au tirage de l'échantillon, il est indispensable de disposer d'une base de sondage qui est une liste exhaustive sans omission et sans répétition des unités (individus, ménages, entreprises par exemple). Elle doit être complète et à jour. Les données de recensement constituent une base de sondage très utilisée. Cependant, elles sont vite désuètes (de nouveaux ménages se créent, se déplacent ou cessent d'exister). Il faut ainsi obligatoirement actualiser la base de sondage au préalable avant de procéder à un échantillonnage.

### Conclusion partielle

A l'issue de cette phase de construction qui comprend principalement l'élaboration des outils de collecte et la réalisation des tests, les documents et outils suivants doivent être élaborés et disponibles :

- le(s) questionnaire(s) ;
- les manuels ;
- le masque de saisie ;
- nomenclature ;
- autres outils de collecte ;
- les programmes de codification ;
- programmes d'apurement de données ;
- programmes de construction des indicateurs ;
- programmes de tabulation ;
- Le plan détaillé et documenté d'analyse ;
- la base de sondage actualisée.

## 9 PHASE DE DEMANDE DE VISA STATISTIQUE

Le décret<sup>3</sup> n°2007-720/PRES/PM/MEF fixant les conditions et les procédures de réalisation des recensements et des enquêtes statistiques par les services et organismes statistiques publics auprès de personnes ne faisant pas partie de ces structures stipule, qu'avant leurs réalisations, ils doivent obtenir l'autorisation préalable du Ministre dont ils relèvent et du Ministre chargé de la statistique tel défini à l'article 2 de la loi n°012-2007/AN du 31 mai 2007 portant organisation et réglementation des activités statistiques.

En rappel, cinq (5) Commissions spécialisées ont été créée par l'arrêté n°2007-785/MEF/CAB du 31 décembre 2007. Ce sont :

- la Commission spécialisée « Organisation du système statistique, législation, ressources humaines et financement » ;
- la Commission spécialisée « Statistiques démographiques et sociales » ;
- la Commission spécialisée « Statistiques économiques et financières » ;
- la Commission spécialisée « Statistiques du secteur rural et de l'environnement » ;
- la Commission spécialisée « Traitement, archivage, diffusion des données et technologies de l'information et de la communication ».

Conformément au décret ci-dessus, le Président du Conseil national de la statistique (CNS) saisit le Président de la Commission spécialisée (CS) appropriée (c'est-à-dire en fonction du domaine de l'enquête) afin qu'il convoque ses membres en vue d'examiner les documents techniques de l'opération qui lui a été soumis pour demander le visa statistique.

### 9.1. Préparation du dossier de visa statistique

La structure demandeuse de visa statistique doit préparer un dossier qui se compose de :

- une correspondance adressée au Ministre de l'économie, des finances et du développement, Président du CNS ;
- une fiche<sup>4</sup> de renseignement pour l'octroi de l'autorisation préalable (à retirer au Secrétariat du CNS à l'INSD) ;
- la méthodologie de l'enquête ;
- le budget de l'enquête ;
- le personnel de l'enquête ;
- le(s) manuel(s) de l'enquête ;
- le(s) questionnaire(s)<sup>5</sup>.

### 9.2. Envoi du dossier

Le dossier de demande de visa statistique est envoyé au Ministère chargé de la statistique (Président du CNS).

### 9.3. Programmation de réunion(s) de Commission spécialisée et l'octroi du visa statistique

Après réception du dossier de demande, et conformément aux dispositions des articles 9 et 10 de l'arrêté n°2007-785/MEF/CAB, le Président du CNS adresse une correspondance au Président de la

<sup>3</sup> Un avant-projet de loi actualisant ces décrets et lois a été introduit pour adoption en Conseil des ministres

<sup>4</sup> Voir annexe....

<sup>5</sup> Version(s) papier(s) de(s) questionnaire(s) dans le cas du CAPI ou PAPI

Commission spécialisée (CS) appropriée afin de requérir leur avis dans les plus brefs délais possibles. Ladite CS tient une ou des réunions sur convocation de son président pour examiner les documents de l'opération et décide enfin de l'autorisation ou non de la réalisation de l'opération. En effet, la loi prévoit un délai maximum de trente (30) jours pour statuer sur l'octroi du visa .

De l'organisation pratique de la réunion par secrétariat du CNS (INSD)

- Préparation de l'avis technique par l'INSD (Direction technique concernée par la thématique)
- Organisation administrative de la réunion (Service de la coordination Statistique). Le SCS est chargé de préparer les correspondances administratives (Lettre du ministre en charge de la statistique au président du CS concerné, lettre du président du CS concerné aux membres du CS pour convoquer la réunion, réservation de salle, la pause).

A issu de la réunion, le secrétariat de la commission spécialisée produit le rapport de la réunion et une lettre motivée d'octroi ou de refus du visa, qui sera soumis à la signature du président du CNS. Cette lettre comporte le numéro du visa statistique qui doit être porté sur les questionnaires.

## **9.4. Validation du protocole par un comité d'éthique**

Les comités d'éthiques étant des instances indépendantes et pluridisciplinaires qui émettent des avis concernant les protocoles d'étude.

Un avis favorable de la part d'un comité d'éthique a pour objet :

- de garantir l'obtention de données de qualité en évaluant la pertinence et la cohérence scientifique du protocole ;
- de vérifier les aspects éthiques de l'étude en évaluant notamment le respect des droits des participants à l'étude. Cette validation est obligatoire lorsque l'on envisage de publier les résultats de l'étude.

Pour les enquêtes qui intègrent un volet santé, le protocole doit être soumis au comité d'éthique en santé rattaché au Ministère en charge de la santé. Les éléments constitutifs du dossier sont le protocole, les outils de collecte, le formulaire de consentement, la note d'intention, les cv du personnel impliqué, le reçu de paiement.

## 10 PHASE DE COLLECTE

### 10.1. Tirage de l'échantillon

Les enquêtes sont en général conçues pour collecter des informations sur un grand nombre de thèmes auprès d'entités opérant dans un large éventail de secteurs. Pour des raisons de coût et de fardeau de réponse, les pays souhaiteront plutôt constituer un échantillon représentatif de la population des entités (base de sondage). Cette base de sondage n'est le plus souvent parachevée que si le nombre d'entités est minime et limité.

Les entités du champ de l'enquête (les populations cibles) varient d'un pays à l'autre et au sein même d'un pays. Comme pour toute enquête le champ et la population cible des enquêtes statistiques sont en général définis selon des variables clefs (branche d'activité, taille d'entité, implantation géographique...). Le recours à des classifications internationales (CITI, CITP, ...) améliore la comparabilité des résultats statistiques d'un pays à l'autre. Dans un même pays, ce champ devrait idéalement couvrir les milieux urbain et rural. C'est un point particulièrement important pour marquer les fractures entre zones urbaines et rurales (Confère : Principes et recommandations sur le classement des localités en « urbain » ou « rural » concernant les enquêtes et recensements de la population et de l'habitat, 2e édition ; UNSD, 2008).

#### 10.1.1 Bases de sondage et couverture

La base de sondage d'une enquête correspond à la forme opérationnelle de la population cible et consiste en une liste exhaustive de toutes les unités statistiques pertinentes. Elle sert habituellement à extraire des échantillons d'unités (échantillons aléatoires pour des sous-populations spécifiques comme par exemple le répertoire des entreprises, des registres administratifs...)

L'exhaustivité suppose une base sans doublons d'unités et d'une couverture totale et même géographique intégrant toutes les unités dans la population cible. D'où nécessité de mise à jour des répertoires recourant aux informations provenant de recensements périodiques des populations, entreprises et établissements, d'enquêtes ponctuelles ou régulières auprès des unités, pour éviter tout phénomène de sur-/sous-couverture, d'une chasse aux doublons et d'une vérification croisée avec d'autres registres (fiscaux ou administratifs, immatriculations, autorisations, sécurité sociale, etc.).

Pour garantir la représentativité de l'échantillon sélectionné, il faut faire appel à des techniques probabilistes pour le constituer. Seul un échantillonnage probabiliste (aléatoire) permet d'estimer l'erreur de sondage (ou « erreur d'échantillonnage ») – définie comme la différence avec la « vraie » valeur, et qui est imputable à l'utilisation d'un échantillon au lieu d'un recensement complet de la population.

Aussi, un plan de sondage aléatoire doit reposer sur des considérations de structure de la population d'entités (stratification), de coût de collecte des données et d'erreur statistique maximale acceptable dans les estimations.

La stratification est la technique qui permet de diviser les unités en groupes relativement homogènes (appelés « strates ») à des fins de plans d'échantillonnage et d'estimation. Correctement effectuée, la stratification minimisera la variance d'échantillonnage des estimations pour une taille d'échantillon donnée. Les stratégies de stratification seront optimales si elles reposent sur des variables étroitement liées aux variables mesurées comme leurs caractéristiques socioéconomiques, taille, position géographique, ... Le recours à des variables de stratification permet en général de réduire la variance globale des estimations ainsi que les désagrégations par secteurs ou entités administratives.

#### 10.1.2 Taille de l'échantillon

La taille de l'échantillon – à savoir le nombre d'unités statistiques à partir desquelles (ou sur lesquelles) des informations seront collectées – est calculée en fonction de la stratification de la population d'entreprises pour laquelle les estimations doivent être effectuées. Si le plan de tabulation comporte la diffusion des données par secteur, taille et/ou implantation, alors l'échantillon devra être suffisamment important pour que ces estimations de données désagrégées aient un niveau acceptable d'erreur d'échantillonnage. Il faudra trouver un juste équilibre entre d'une part le degré de précision (reflétant la taille de l'erreur d'échantillonnage) des estimations effectuées et d'autre part les

ressources disponibles. Plus l'échantillon sera grand, plus les coûts de collecte et de traitement des données seront élevés.

Pour une strate donnée, le fait de doubler la précision (c'est-à-dire de diminuer de moitié l'erreur d'échantillonnage) pour l'estimation d'une proportion implique de multiplier par quatre la taille d'échantillonnage dans cette strate. La taille et le plan d'échantillonnage sont obtenus à l'aide de procédures ascendantes qui calculent la taille minimale d'un échantillon pour que les estimations de sortie concernant les variables les plus importantes aient une erreur d'échantillonnage maximale fixée. Les estimations de l'erreur d'échantillonnage pour ces variables peuvent servir à constituer l'échantillon en s'appuyant éventuellement sur des enquêtes précédentes ou des tests pilotes.

NB : Il n'existe pas de recommandations internationales quant à la précision à obtenir pour des indicateurs quelconques. La précision est habituellement exprimée en coefficient de variation (l'écart type divisé par la valeur de l'estimation, le plus souvent exprimé en pourcentage). La taille finale de l'échantillon correspondra à la somme des échantillons par strate. Si l'échantillon final est trop grand, il peut être judicieux de reconsidérer les erreurs de certains groupes de population puis de recalculer la taille totale. L'approche habituelle pour déterminer la précision voulue des estimations consiste à fixer une erreur d'échantillonnage maximale pour les ventilations unidimensionnelles d'indicateurs et des erreurs d'échantillonnages supérieures pour les ventilations croisées. Les recommandations d'Afristat à l'attention de ses États membres préconisent un coefficient maximal de variation de 2 % pour les proportions globales et de 5 % pour les proportions liées aux différents sous-groupes de la population lorsque ces sous-groupes constituent au moins 5 % de la population totale du champ de l'enquête. Une procédure descendante sera parfois utile pour calculer la taille maximale de l'échantillon en fonction du budget affecté à l'enquête et du coût unitaire de la collecte de données auprès d'une unité statistique puis pour allouer (distribuer) l'échantillon par strate en fonction d'une règle opérationnelle précise. La répartition optimum de Neyman (fondée sur le coût et la variance des variables dans chaque strate) garantit par exemple que l'erreur globale est minimisée. Ces méthodes peuvent être adaptées aux besoins de l'enquête. Certaines modifications sont assez fréquentes : 1) fixation d'une taille minimale de l'échantillon dans certaines strates, afin de calculer des estimations avec une précision minimale ; 2) recensement exhaustif pour certaines strates importantes (comme celles recouvrant les grandes entités ; 3) sélection d'un échantillon plus grand que la taille optimale afin d'anticiper la réduction de l'échantillon réel de fait des non-réponses.

Rappelons que, quelle que soit la taille de l'échantillon, un taux important de non réponse signifie que les estimations calculées sont probablement biaisées, et que ce biais va augmenter avec le taux de non-réponse si les unités statistiques sans réponses ne sont pas très différentes de celles qui répondent. L'un des objectifs importants de l'enquête consistera donc à minimiser le taux de non-réponse (sachant que ce taux ne donne pas une information complète sur le biais des estimations).

### **10.1.3 Méthodes d'échantillonnage**

À partir d'une population organisée selon les strates décrites ci-dessus, comment procéder à l'échantillonnage au sein de ces strates ?

L'échantillonnage systématique et l'échantillonnage aléatoire sont les deux méthodes les plus simples et les plus courantes (voir documentation). Rappelons que la méthode d'échantillonnage doit correspondre à la méthode d'estimation. Ainsi, si l'on considère adéquat de sélectionner les unités d'une strate donnée avec des probabilités inégales (proportionnelles à leur taille par exemple), les estimations devront pondérer les unités par un coefficient qui correspond à ces probabilités dans leurs formules.

En synthèse, le tirage de l'échantillon se fera en disposant de la base de sondage exhaustive et adéquate, en respectant la formule mathématique approuvée, en définissant proprement les strates convenables, en respectant la méthode d'échantillonnage prédéfinie dans la méthodologie conceptuelle (SAS, R, Stata, etc.) et en justifiant toute autre méthode d'affinement de l'échantillon.

## **10.2. Mise à jour de l'échantillon**

Elle consistera, une fois sur le terrain, d'énumérer à nouveau les entités tirées afin d'avoir des effectifs actualisés avant de procéder aux tirages à différents degrés.

Mais avant d'aller sur le terrain, des préalables sont nécessaires.

## **10.3. Mise en place du processus de collecte**

### **10.3.1 Recrutement du personnel de collecte**

Les recrutements doivent respecter un protocole général de sélection dans toute enquête ou recensement. Le processus de recrutement doit respecter – la définition des termes de références pour le recrutement – la mise en place du comité de recrutement, la formulation et diffusion du communiqué de recrutement selon les normes audiovisuelles, la réception des candidatures, le dépouillement des candidatures par le comité de recrutement, la sélection selon les critères prédéfinis, la validation des listes par l'autorité compétente, la publication des listes validées des candidats retenus et la transmission des listes aux équipes techniques de formation.

### **10.3.2 Identification des locaux pour la formation des agents recenseurs ;**

Il est important de prévoir les lieux de formation d'une enquête et s'assurer des bonnes conditions de formation (nombre suffisant de formateurs, salles bien éclairées avec des tables/bans pour déposer les outils utilitaires et des prises de courant et rallonges pour faciliter l'accès à l'électricité, ...)

### **10.3.3 Evaluation et sélection définitive des participants formés ;**

Il est également nécessaire de prévoir des évaluations pour s'assurer de la bonne portée de la formation et sélectionner les meilleurs agents parmi les formés, établir la liste des admis, soumettre à validation et afficher/communiquer après validation. S'en suivra la signature des contrats avant le départ sur le terrain.

## **10.4. Sensibilisation**

La sensibilisation se fait en deux étapes :

La première étape consiste à lancer des communications ou communiqués pour prévenir, sensibiliser et susciter l'adhésion de la population cible de l'opération de collecte (diffusion des spots audio, vidéo, des pages et des encarts publicitaires dans des journaux de la place, banderoles...).

La seconde étape, qui se déroule à la phase terrain, nécessite un contact avec les autorités administratives et locales pour avoir leurs accompagnements et, pendant l'enquête proprement dite, les agents expliquent le bien-fondé de l'enquête pour avoir le bon assentiment des populations pour favoriser une bonne interview.

## **10.5. Collecte des données**

La collecte des données nécessite certains préalables :

### **10.5.1 Matériel de la collecte (quantité, distribution, etc.)**

Le matériel de collecte préparé en amont par la DIMIS en coordination avec le service responsable de l'enquête, sera récupéré par les superviseurs de l'enquête et distribué (avec bordereaux pour traçabilité).

### **10.5.2 Mise en place du personnel du terrain ;**

La mise en place des agents sur le terrain doit se faire progressivement dans le respect du protocole hiérarchique (Superviseur, contrôleur, agent). Une fois sur place, les contrôleurs avec leurs agents, munis des croquis adéquats, feront la reconnaissance des limites des ZD à eux assignées.

### **10.5.3 Numérotation des concessions et des ménages ;**

Une fois les ZD connues dans leurs limites, à partir de la fiche d'énumération version électronique ou papier, l'agent enquêteur procède à l'énumération des unités un par un.

#### **10.5.4 Remplissage des questionnaires**

Les préliminaires de salutations et d'explication de l'enquête ayant permis l'acceptation de la population cible à donner les informations, l'agent enquêteur, en déroulant les questions qui ressortent dans son smartphone ou son questionnaire papier, note les réponses données par l'enquêté et fait de même pour toutes les unités statistiques de sa zone de couverture.

#### **10.5.5 Contrôle des questionnaires**

Le marquage serré des contrôleurs et superviseurs permet toujours d'assurer une collecte franche des données ; sinon un remplacement des agents malveillants est vite nécessaire pour assurer la bonne collecte.

#### **10.5.6 Supervision du dénombrement**

La supervision des enquêtes se fait en grande partie, par les rencontres et échanges sur le terrain ou appels quotidiens des contrôleurs et agents. Elle est ponctuée par la convocation et le management des contrôleurs et agents ayant des manquements dans leur travail, la révocation des agents et contrôleurs défaillants, les interventions dans les ménages pour insister sur le bien-fondé de l'enquête et l'obtention de leur acceptation à se laisser enquêter...

#### **10.5.7 Transfert des données**

Les données électroniques ou papiers doivent être transférées au jour le jour, par Bluetooth ou dans une chemise cartonnée aux contrôleurs et récupérées après par les superviseurs. Si le dispositif est possible, les données seront envoyées directement au bureau central de l'opération. De nos jours, avec les supports électroniques, l'assistance d'informaticiens pour pallier aux difficultés éventuelles, est toujours à prévoir.

#### **10.5.8 Finalisation de la collecte des données**

Elle consistera à assurer la complétude des données à collecter soit par retour terrain soit par appels téléphoniques, ....

#### **10.5.9 Rapport de collecte**

Toute opération d'enquête nécessite à la fin un rapportage des activités menées, rapports qui serviront à mieux réfléchir les opérations futures et à peaufiner les analyses de données. Les rapports pourront ressortir : la capacité de stockage des smartphones et leur autonomie ; la fiabilité des questionnaires papiers ou la performance des smartphones ; la vitesse d'action de l'application de collecte et sa stabilité ; l'opérationnalité du menu de gestion de l'application ; les difficultés dans le recrutement et la formation et l'assiduité des agents formés ; les difficultés dans la sensibilisation ; les difficultés dans l'obtention et la distribution du matériel d'enquête ; les difficultés cartographiques et de reconnaissance des ZD ; les difficultés ou cours de la collecte des données auprès des enquêtés ; les contrôles et supervisions internes et externes et leurs impacts sur la collecte ; et toutes autres situations ayant influencé d'une façon ou d'une autre le déroulement de l'opération d'enquête.

# 11 PHASE DE TRAITEMENT

## 11.1. Profilage et préparation des données

L'étape du profilage et de la préparation des données consiste à modifier les renseignements sources originaux au moyen d'un processus par lequel des données incomplètes, incorrectes, inexacts ou non pertinentes sont détectées par ce qu'on appelle le profilage et sont ensuite modifiées, corrigées ou supprimées. Elle peut également impliquer la suppression d'erreurs typographiques et la correction de valeurs selon une liste d'entités connues. Une autre pratique courante de préparation des données consiste en leur amélioration par l'ajout de renseignements connexes ou complémentaires. L'objectif de cette étape est de s'assurer que les renseignements sources acquis sont non seulement cohérents, mais aussi prêts à passer aux prochaines étapes. L'étape du profilage et de la préparation des données comporte sept sous-étapes, qui peuvent se dérouler en ordre séquentiel ou de manière parallèle et qui peuvent également être itératives. Les sept sous-étapes sont les suivantes :

- Profilage des données ;
- Standardisation, classification et codage ;
- Contrôle et imputation ;
- Ajustement et repondération ;
- Dérivation de nouvelles variables et d'unités statistiques ;
- Évaluation et documentation de l'incidence des changements ;
- Mise au point finale des fichiers de données

## 11.2. Profilage des données

Cette sous-étape du processus de la préparation consiste à examiner et à récolter le plus de renseignements possibles sur les données qui ont été acquises pour découvrir leurs différentes caractéristiques et en dresser un profil. Les principaux objectifs des renseignements recueillis sont les suivants :

- vérifier si les renseignements contenus dans les données peuvent être facilement utilisés à d'autres fins ;
- déterminer les sous-ensembles d'enregistrement et de variables provenant de sources administratives et confirmer lesquels sont dans le champ du programme statistique ;
- avoir des mesures sur la qualité des données et sur leur conformité par rapport aux normes de l'INSD;
- évaluer les risques engendrés par l'intégration de ces données à d'autres sources, y compris les défis liés au couplage d'enregistrements ;
- vérifier si les métadonnées décrivent correctement les renseignements contenus dans la source de données ;
- évaluer les modèles de distribution des valeurs et les dépendances fonctionnelles des renseignements contenus dans la source de données ;
- avoir une bonne compréhension de l'enjeu des renseignements sur les projets à venir pour éviter les mauvaises surprises ;
- avoir une vue globale des renseignements quant aux utilisations telles que la gestion des bases de données ou l'intendance des données pour améliorer leur qualité.

## 11.3. Standardisation, classification et codage

Cette sous-étape consiste en la normalisation de la structure, du format et des ensembles de codes et d'abréviations des variables relatives aux données sources pour garantir leur uniformisation, leur

comparabilité et leur stabilité dans tous les fichiers. Cette sous-étape comprend aussi les activités permettant d'attribuer une valeur numérique ou alphanumérique à une réponse texte selon un système de classification préétablie. Bien qu'elles soient fréquemment automatisées, les activités exécutées à cette sous-étape requièrent parfois une intervention manuelle parce qu'elles peuvent exiger une prise de décision complexe. La classification et le codage des données permettent de produire des données formatées qui seront utilisées aux étapes subséquentes.

## **11.4. Contrôle et imputation**

Cette sous-étape est une suite logique de la sous-étape du profilage des données. Elle a lieu lorsque, durant le profilage, les données sont déclarées incorrectes, manquantes ou peu fiables. L'imputation est donc le processus qui permet d'attribuer des valeurs de remplacement à des données manquantes, invalides ou incohérentes qui ont été rejetées à l'étape de la validation des données. Ce processus doit suivre certaines règles et il comprend les phases suivantes :

- la détermination de la pertinence d'ajouter ou de modifier des données ;
- le choix de la méthode à utiliser ;
- l'ajout ou la modification des valeurs des données à imputer ;
- la production de métadonnées sur le processus d'imputation.

## **11.5. Ajustement et repondération**

Cette sous-étape sert en particulier au rajustement des poids de sondage et a lieu essentiellement dans le cas des processus statistiques d'enquêtes par échantillonnage lorsque l'échantillon observé est plus petit que celui initialement sélectionné. Ce phénomène est observé en cas de non-réponse totale ou à cause des unités hors champ sélectionnées. Les facteurs d'ajustement doivent être fondés sur la probabilité de réponse de chaque unité échantillonnée dans le cas où la non-réponse est reliée aux variables mesurées.

## **11.6. Dérivation de nouvelles variables et d'unités statistiques**

Cette sous-étape consiste à créer de nouvelles variables et unités statistiques à partir des valeurs de celles obtenues de l'ensemble des données sources. Ces variables et unités sont le plus souvent nécessaires à la création du produit statistique souhaité. La dérivation est réalisée au moyen de l'application de formules mathématiques à l'une ou à plusieurs des variables présentes dans l'ensemble de données ou de l'application de différentes hypothèses de modélisation. Il est parfois nécessaire de procéder à la dérivation de manière itérative, puisque certaines variables qui servent à dériver d'autres variables peuvent elles-mêmes avoir été créées par dérivation. Il importe donc de veiller à ce que les variables soient dérivées dans le bon ordre.

## **11.7. Évaluation et documentation de l'incidence des changements**

Cette sous-étape consiste à évaluer et à documenter l'incidence des différentes transformations opérées tout au long des sous-étapes précédentes. Il s'agit d'un exercice important parce qu'il permet non seulement de dresser un portrait global des améliorations apportées aux données sources, mais également de préciser l'incidence qu'ont eue les transformations sur la qualité des données.

## **11.8. Mise au point finale des fichiers de données**

C'est à cette sous-étape que l'ensemble des résultats des autres sous-étapes du traitement sont mis ensemble pour produire le fichier de données devant servir à l'étape suivante. Cette sous-étape consiste à créer le fichier de traitement final, qui constitue la base de diffusion de tous les fichiers de données, à savoir le fichier principal, le fichier de microdonnées à grande diffusion (FMGD) et le fichier partagé. Les fichiers de diffusion sont les fichiers qui seront utilisés à l'intérieur et à l'extérieur de l'INSD pour analyser les données de l'enquête. Le FMGD est une version abrégée du fichier principal

qui a été minutieusement évalué par les statisticiens pour éliminer tous les risques raisonnables d'identification des répondants pouvant présenter des caractéristiques passablement uniques. Le fichier partagé, quant à lui, ne contient que les enregistrements des répondants qui ont consenti à ce que leurs enregistrements soient partagés.

## **11.9. Intégration, estimation et compilation**

L'intégration des données regroupe un ensemble d'activités qui consistent à extraire les données provenant de différentes sources (bases de données, fichiers, applications, services Web, courriels, etc.), à leur faire subir certaines transformations (jointures, déduplication, concaténation, calculs, etc.) pour finalement les consolider dans une seule et même base de données. La compilation consiste à créer des données agrégées et des chiffres de population à partir des microdonnées ou d'agrégats de niveau inférieur. Dans le cadre des processus de production statistique impliquant plusieurs sources de données d'entrée, cette étape consiste à mettre en correspondance plusieurs types de statistiques, concepts, classifications et conventions au moyen d'un processus itératif conduisant à des données agrégées. Dans ce cas, l'intégration et la compilation se réalisent de manière simultanée. Dans le cas des processus d'enquêtes-échantillons, si des données auxiliaires sont disponibles, on peut améliorer la précision des estimations. L'intégration de données auxiliaires dans ces processus d'estimation est appelée calage. L'étape de l'intégration et de la compilation comporte cinq sous-étapes, qui peuvent se dérouler en ordre séquentiel ou de manière parallèle et qui peuvent être itératives. Ces sous-étapes sont les suivantes :

- Détermination des éléments, des règles et de la stratégie d'intégration
- Évaluation et ajustement de la stratégie d'intégration
- Chargement, application des mises en correspondance et intégration des données sources
- Estimation, compilation et application des méthodes statistiques
- Vérification et ajustement d'amélioration de la qualité

## **11.10. Détermination des éléments, des règles et de la stratégie d'intégration**

Cette sous-étape consiste à déterminer les éléments d'intégration. Il s'agit essentiellement d'une variable ou d'une combinaison de variables qui rendent faisable l'intégration des données de sources multiples. Ce sont les identifiants de base de données ou clés. Certains fichiers de données comportent des clés uniques pour chaque enregistrement, ce qui peut faciliter leur intégration lorsque ces clés sont les mêmes pour tous les fichiers de l'intégration. Dans le cas d'un processus statistique par couplage d'enregistrements, cette sous-étape comprend la détermination d'un ensemble de paires d'enregistrements correspondant parfaitement à une clé donnée ou à un critère particulier et représentant des paires potentiellement jumelées, que l'on appelle également paires potentielles. Cette sous-étape consiste aussi à élaborer et à définir les règles et la stratégie qui seront utilisées pour l'intégration des données. Dans le cadre d'un processus de couplage d'enregistrements, cette sous-étape commence par la comparaison des champs et des enregistrements où, pour chaque paire, les attributs des enregistrements couplés sont comparés. La comparaison d'attributs est en outre fondée sur des fonctions de comparaison qui diffèrent selon la manière dont elles traitent les valeurs manquantes. Les résultats de la comparaison servent ensuite à la prise de décision en ce qui a trait à la stratégie de couplage. Cette sous-étape se termine généralement par la production des clés de couplage préliminaires. Dans les cas des processus impliquant plusieurs types de données d'entrée tels que les comptes macroéconomiques et les estimations de population, il arrive souvent que cette sous-étape se limite à confirmer qu'il n'y a rien à modifier. Cependant, dans certains cas, de nouveaux types de transactions peuvent émerger pour lesquels aucune structure ni règle n'a encore été créée ou élaborée. Par conséquent, des règles et un traitement devront être déterminés avant que les composantes ou les comptes soient compilés.

## **11.11. Évaluation et ajustement de la stratégie d'intégration**

Cette sous-étape consiste à évaluer la qualité des éléments d'intégration. Dans le cas d'un processus par couplage d'enregistrements, elle porte sur l'évaluation de la qualité du couplage (validation interne) et de l'adéquation à l'utilisation de l'ensemble de données couplées (validation externe). Le but de cette étape est de veiller à ce que l'ensemble de données couplées présente un niveau de qualité adéquat pour l'utilisation prévue. Toute limite des données couplées relevées au cours de cette sous-étape devrait contribuer à la décision relative à leur utilisation. Cette sous-étape permet également de perfectionner, au besoin, la stratégie de couplage d'enregistrements en fonction des résultats de la validation interne et externe. Cela peut consister à perfectionner les règles de couplage ou les seuils utilisés pour vérifier si les paires correspondent ou non. Cette sous-étape se termine généralement par la production des clés de couplage finales.

En ce qui a trait aux processus macroéconomiques, cette sous-étape suppose l'analyse des comptes pour en vérifier la cohérence et la nature raisonnable. Différentes perspectives ou dimensions sont examinées, comme la dimension relative à la cohérence temporelle et la dimension relative à la structure. Les restrictions comptables et la mise en équilibre sont également examinées dans le cadre de ce sous-processus pour assurer une cohérence globale. À la suite de cette étape, il est possible qu'on revienne à la phase de conception, puisque les connaissances acquises au cours de cette étape peuvent entraîner une modification de la conception du compte. Une fois les ajustements apportés à la conception, il faut revenir à la sous-étape VII.1 pour mettre à jour les structures comptables, les règles et la stratégie d'intégration. Pour ce qui est des estimations et des projections de population, cette sous-étape suppose l'analyse des tendances ainsi que de l'accroissement démographique pour en vérifier la cohérence et la nature raisonnable. La structure selon l'âge et le sexe ainsi que la distribution géographique feront également partie de l'analyse.

## **11.12. Chargement, application des mises en correspondance et intégration des données sources**

Cette sous-étape consiste en l'utilisation d'identifiants de base pour créer de nouvelles données à partir de deux ensembles de données sources ou plus. Dans les cas de processus statistique comme les comptes macroéconomiques, les estimations ou les projections de population, cette sous-étape consiste à charger les données sources acquises et les transformer de manière à ce qu'elles soient conformes aux concepts et aux mises en correspondance de la comptabilité macroéconomique ou de la démographie. Tout renseignement acquis au cours des premières étapes est également pris en considération, puisqu'il s'agit de renseignements importants dans le processus général de création de comptes macroéconomiques ou des composantes démographiques. En ce qui a trait au processus statistique par couplage d'enregistrements, l'intégration comprend généralement le couplage systématique, l'utilisation de clés de couplage pour créer un ensemble de données couplées et le rapprochement des variables lorsque deux ensembles de données sources ou plus contiennent les mêmes variables. Pour examiner le processus de fusion, il convient d'effectuer des vérifications de la fréquence ou des enregistrements sur chaque ensemble de données sources avant le processus de fusion, puis à nouveau après la fusion. On compare ensuite le nombre d'enregistrements afin de déceler tout écart. Pour ce qui est des processus d'enquêtes-échantillons, l'intégration a lieu avec des données auxiliaires et est appelée calage (voir l'ajustement et la repondération). Elle consiste à rajuster les poids initialement calculés pour que les estimations de la ou les variables auxiliaires correspondent à des totaux connus.

## **VIII.4 Estimation, compilation et application des méthodes statistiques**

Cette sous-étape consiste à créer des données agrégées et des chiffres de population à partir de microdonnées ou d'agrégats de faible niveau. Elle comprend la sommation des données pour les enregistrements qui ont certaines caractéristiques en commun, la détermination des mesures de la moyenne et de la dispersion, et l'application des poids de la sous-étape (ajustement et repondération) pour calculer les totaux appropriés. Dans le cas des enquêtes-échantillons, les erreurs d'échantillonnage peuvent également être calculées dans cette sous-étape, et associées aux agrégats pertinents. La compilation est un processus itératif qui peut être une bonne méthode pour évaluer la qualité des données parce qu'elle permet de faire ressortir certaines incohérences. C'est également à cette sous-étape que les méthodes statistiques, telles que la dessaisonnalisation, la déflation ou l'étalonnage, sont appliquées.

### **11.13. Vérification et ajustement d'amélioration de la qualité**

Cette sous-étape consiste à vérifier que les données intégrées ou compilées répondent aux objectifs de production et peuvent être utilisées de manière fiable. Les activités de vérification consisteront notamment à confronter le produit statistique à d'autres données pertinentes (internes et externes), à confronter les statistiques aux attentes et aux connaissances du domaine ou à étudier les incohérences présentes dans les statistiques. Dans le cadre des processus de couplage d'enregistrements, cette sous-étape est importante parce qu'elle fournit des renseignements supplémentaires sur l'adéquation à l'utilisation de l'ensemble de données couplées. Les erreurs ou les limites décelées à cette sous-étape peuvent exiger de plus amples ajustements de la stratégie de couplage d'enregistrements, de l'évaluation de la qualité ou des ajustements d'amélioration de la qualité exigeant la répétition de sous-étapes antérieures.

## **12 PHASE D'ANALYSE**

L'étape de l'analyse consiste, d'une part, à examiner, à interpréter et à valider les données aux fins de diffusion et, d'autre part, à élaborer des réponses à certaines questions.

### **12.1. Préparation de la version préliminaire des résultats**

Cette sous-phase consiste à calculer les principaux produits/indicateurs de l'opération conformément aux métadonnées.

### **12.2. Validation des résultats**

Cette sous-étape consiste à valider la qualité des données produites en les soumettant aux normes et standards du cadre d'évaluation de l'assurance de l'INSD. Il faut aussi vérifier si les données sont conformes aux objectifs espérés assignés au processus statistique. Elle consiste également à collecter les données antérieures et à les comparer ensuite aux données acquises pour repérer toute divergence par rapport aux attentes et pour permettre des analyses basées sur des données probantes. La base de données définitive et apurée est enregistrée des formats exploitables (SPSS ou STATA).

### **12.3. Analyse descriptive des données**

Les données définitives issues de l'opération statistique peuvent servir à des fins d'analyses descriptive, exploratoire ou explicative. Les utilisations descriptives ciblent l'estimation des paramètres récapitulatifs de la population (les moyennes, les fréquences, variance, dispersion...)

L'analyse descriptive des données consiste à :

- Etablir et examiner des tableaux, des graphiques, des cartes et diverses mesures sommaires ;
- Etablir une comparaison entre les estimations de l'enquête et les résultats d'autres enquêtes ou sources de données ;
- Relever et commenter les points saillants ;
- Faire des tests statistiques de comparaison (moyenne, proportion, variance, ...), d'indépendance (khi-deux, coefficients, rapports de corrélation, ...) et pour commenter les écarts ;
- Donner des pistes d'explications sur certains résultats.

### **12.4. Analyse explicative des données (en cas de besoin)**

L'inférence statistique sert à utiliser des techniques statistiques solides (régressions, modèles de classification, ...) pour vérifier les hypothèses ou étudier les liens entre des caractéristiques ou identifier les facteurs explicatifs la variable d'intérêt à l'aide d'autres variables.

### **12.5. Elaboration rapport de l'opération**

Il s'agit de produire un rapport de l'opération conformément à un plan d'analyse pour répondre aux besoins d'informations exprimés par les utilisateurs. Il est recommandé d'utiliser un style d'écriture facile à comprendre, d'interpréter correctement les chiffres et d'éviter le jargon statistique.

### **12.6. Application des contrôles de confidentialité**

Le contrôle de confidentialité vise à protéger des fournisseurs de données pour éviter la publication directe ou indirecte de données à caractère individuel.

## 12.7. Finalisation des résultats

Il s'agit de soumettre à amendement le rapport provisoire à l'ensemble des utilisateurs potentiels identifiés lors d'un atelier. Les amendements sont pris en compte si possible pour élaborer le rapport définitif de l'enquête qui fera l'objet de diffusion.

A la fin de la phase, il faut disposer de :

- la base de données définitive apurée ;
- le rapport provisoire de l'enquête ;
- le rapport définitif de l'enquête.

## **13 PHASE DE DIFFUSION**

L'étape de la diffusion consiste à mettre à la disposition des utilisateurs les données recueillies et traitées dans le cadre d'une activité statistique par divers moyens. La diffusion doit être accompagnée d'une communication efficace visant à offrir aux utilisateurs un large accès à l'information en vue d'optimiser l'utilisation des données. Toutes les sous étapes doivent être réalisées par le service concerné en collaboration avec la DIMIS.

### **13.1. Mettre à jour les systèmes de diffusion**

Cette sous-étape consiste à faire les dernières mises au point des systèmes de diffusion avec les données et métadonnées qui vont être publiées. Elle comprend les activités de formatage des données et des métadonnées de manière à ce qu'elles soient prêtes à être chargées dans le module de sortie, et de la vérification de l'existence d'un lien entre les données et les métadonnées pertinentes.

### **13.2. Elaboration des produits de diffusion**

Cette sous-étape consiste à finaliser la production des données selon la manière qu'elles ont été conçues. Elle comporte les activités de préparation des composantes du produit, telles que le texte explicatif, les tableaux et graphiques ou les énoncés sur la qualité. Elle inclut aussi l'assemblage des composantes préparées et une vérification finale pour s'assurer que le produit répond aux objectifs de départ et qu'il est conforme aux normes de diffusion.

### **13.3. Gestion de la sortie des produits de diffusion**

La politique de diffusion de l'INSD indique les principaux canaux à utiliser :

Le courrier ordinaire qui est un canal de diffusion permettant de transmettre les informations statistiques (les publications au format papier et/ou CD-Rom) par courrier officiel signé du DG aux utilisateurs institutionnels. Elle nécessite la tenue d'une liste de destinataires prioritaires. L'INSD retient comme destinataires institutionnels prioritaires : le cabinet du MEF, les structures membres du SSN, les bibliothèques centrales des universités, les partenaires techniques et financiers, les représentations diplomatiques, les structures abonnées aux publications de l'INSD.

Les sites web de l'INSD ([www.insd.bf](http://www.insd.bf)) et du CNS ([www.cns.bf](http://www.cns.bf)) : Toute publication doit être systématiquement mise en ligne à la date indiquée dans le calendrier de parution. Les nouvelles publications doivent être mises en évidence (à la une) pour être facilement repérables. En même temps, elles doivent être répertoriées et bien archivées sur le site. Les bases de données doivent également être diffusées sur le site web. Toutes les ressources disponibles sur le site web de l'INSD sont gratuites et libres d'accès. Le catalogue des publications doit être actualisé et le calendrier de diffusion doit être mis en ligne. Le site doit offrir la possibilité de s'inscrire sur les listes de diffusion des publications et tous les communiqués de presse sur les statistiques doivent y être publiés.

Le mailing ou le E-mailing est la diffusion d'un message électronique personnalisé à un public ciblé afin de lui présenter un produit. Avec l'utilisation de l'Internet qui ne cesse de croître, le mailing est devenu un puissant canal de diffusion statistique. A cet effet, il est utilisé, soit pour envoyer directement le fichier d'une publication aux utilisateurs, soit pour annoncer la mise à disposition d'une publication en indiquant son lien Internet ou les conditions de son accessibilité s'il n'est pas disponible en ligne. On parle dans ce deuxième cas d'une alerte. Le mailing permet de toucher individuellement un public très large parmi les personnes disposant d'une adresse électronique. A la parution d'une publication, de façon systématique, le fichier électronique au format PDF non modifiable doit être envoyé à toutes les adresses de la liste correspondante. Toute mise à jour du site web donne lieu à l'envoi d'une alerte indiquant les mises à jour opérées.

La salle de lecture : La consultation des publications de l'INSD doit être possible dans une salle de lecture aménagée à son siège et ouverte en permanence pendant les heures de service. Des possibilités de réaliser des photocopies et impressions doivent être offertes, de même que celles de réaliser une recherche documentaire sur le site web de l'INSD.

## **13.4. Promotion des produits de diffusion**

Elle peut être réalisée au travers des masses médias et des ateliers de dissémination.

Les Mass médias (Tv, Radio, presse écrite) : Lorsque les résultats de grandes opérations sont disponibles ou lorsqu'une publication est disponible, l'INSD doit envoyer un communiqué de presse aux médias et/ou organiser un point de presse. Le communiqué de presse annonce que de nouvelles données statistiques sont disponibles. Il contient quelques résultats parmi les plus importants et leur signification. Le point de presse permet d'avoir un échange avec les journalistes sur les nouvelles statistiques. Il permet aux médias de livrer la bonne information aux populations en minimisant les risques d'interprétations erronées.

Ateliers de dissémination des résultats : A l'occasion de la parution des résultats des grandes opérations (enquêtes et recensements), l'INSD doit organiser des ateliers de dissémination pour diffuser les informations statistiques. Des responsables institutionnels et les médias y sont invités. Ce mode de diffusion touche un public limité et doit forcément être associé à un mode de diffusion qui touche un large public.

## **13.5. Assistance aux utilisateurs**

L'assistance aux utilisateurs peut concerner l'explication des statistiques produites (données disponibles) ou la production des statistiques complémentaires pour étayer celles qui sont publiées ou simplement pour des besoins spécifiques (données sur commande). Une équipe compétente doit être mise en place pour centraliser et coordonner l'assistance aux utilisateurs en collaboration avec les services techniques qui élaborent les publications.

## **14 PHASE D'ARCHIVAGE**

Cette phase consiste à utiliser l'ensemble des techniques et moyens pour recueillir, classer, conserver et exploiter des documents relatifs au processus statistique jusqu'à leur destruction éventuelle.

### **14.1. Définition du système d'archivage des données**

Le système d'archivage est un ensemble de techniques et de procédés mises en place par une structure pour une meilleure organisation de ses archives. Le système d'archivage se repose sur des impératifs de protection, d'accessibilité et d'intégrité des données.

Ainsi, le système doit garantir la protection et l'intégrité des données contre la perte ou l'altération et d'offrir aux utilisateurs un accès facile aux données

### **14.2. Archivage des questionnaires remplis**

Après exploitation, les questionnaires remplis des enquêtes de grandes envergures tels-que le RGPH doivent faire l'objet d'un archivage spécifique puisqu'ils contiennent généralement les noms et les prénoms des personnes enquêtées. Afin de résoudre le problème d'espace de stockage, ils doivent être transférés au Archives Nationales puis systématiquement numérisés et conservés sous forme de fichiers à la fin de l'enquête. La confection de CD ROM comportant tous les documents techniques de l'enquête est obligatoire à la fin de l'opération. Ces supports sont élaborés par la Direction en charge de l'informatique en collaboration avec les directions techniques concernées.

### **14.3. Collecte des données à archiver**

Très souvent, l'archivage des données est considéré comme l'étape ultime de l'opération statistique. Il est réalisé dans l'urgence, lorsqu'il est effectivement réalisé. Réunir les données à archiver constitue alors un exercice difficile car certaines données ont pu être déjà perdues ou dispersées. Or l'archivage est plus efficace et plus facile à mettre en œuvre si c'est une préoccupation permanente tout au long de l'opération : les données qui vont faire partie de l'archive vont être collectées au fur et à mesure de leur disponibilité

### **14.4. Information à archiver**

Le champ de données à archiver couvre les données traitées et produites par les différents types d'opérations statistiques réalisées par l'INSD, lorsqu'il s'agit d'opérations ponctuelles (enquêtes, recensements), de l'exploitation régulière des données fournies par des sources administratives (déclarations douanières, déclarations statistiques et fiscales, ...) ou des opérations statistiques de synthèse (comptes nationaux, études conjoncturelles).

#### **14.4.1 Documents réglementaires de l'opération**

Il s'agit des documents juridiques spécifiques qui encadrent tout le processus.

#### **14.4.2 Outils de collecte de données**

Il s'agit des questionnaires, manuels, méthodologie, catalogues...

#### **14.4.3 Données de base**

Il s'agit des bases de données collectées et apurées, celles utilisées pour produire les résultats et indicateurs statistiques. Pour faciliter leur utilisation ultérieure, elles peuvent être stockées sous différents formats facilement exploitables : Excel, SPSS, Stata et R.

#### 14.4.4 Métadonnées des indicateurs

Elles peuvent décrire les résultats et indicateurs statistiques aussi bien du point de vue conceptuel que des traitements. Les métadonnées sont fournies par différents documents.

#### 14.4.5 Notes méthodologiques

Elles décrivent les concepts retenus et les méthodes utilisées pour collecter l'information et mesurer les phénomènes observés. Dans le cas des enquêtes, elles seront complétées par le plan de sondage et les questionnaires.

#### 14.4.6 Nomenclatures et classifications utilisées

Les documents relatifs à l'organisation de l'opération : manuel de l'enquêteur, du contrôleur, du superviseur, plan d'informations.

#### 14.4.7 Rapports, plans de traitement

Il s'agit des documents issus des traitements effectués sur les données : contrôles de validité, contrôle de vraisemblance, redressements, calculs de nouvelles variables.

#### 14.4.8 Produits statistiques de l'opération

Il s'agit de tous les produits issus du processus statistique.

La fréquence de constitution de l'archivage dépend du type d'opération statistique : l'archive sera disponible en fin d'opération pour les enquêtes ponctuelles et recensements ; il sera réalisé généralement annuellement pour les autres opérations. Dans ce dernier cas, il comportera également toutes les informations faisant références aux productions des données infra annuelles (mensuelles, trimestrielles).

Le tableau Tableau 5 récapitule les principales opérations prévues au SDS devant faire l'objet d'un archivage et la fréquence de ces archivages. A celles-ci il faudra ajouter les opérations ponctuelles, généralement des enquêtes.

**Tableau 5: la fréquence des archivages**

Opération statistique	Fréquence de l'archivage
Comptes nationaux	Annuelle
Statistiques du commerce extérieur	Annuelle
Enquête sur les prix de détails(IHPC)	Annuelle
Enquête sur les productions industrielles	Annuelle
Enquête sur l'opinion des chefs d'entreprise	Annuelle
Exploitation des DSF	Annuelle
Répertoire des entreprises	Annuelle
Tableau de bord de l'économie	Annuelle
Recensement industriel et commercial	2 ans (fin d'enquête)
Enquête QUIBB	Annuelle (fin d'enquête)
Enquête prioritaire	4 ans (fin d'enquête)
Enquête démographique et de santé (EDS)	5 ans (fin d'enquête)
MICS	5 ans (fin d'enquête)
RGPH	10 ans (fin d'enquête)

## **14.5. Archivage des programmes informatiques**

Afin d'assurer la pérennité dans la gestion des systèmes au sein de l'INSD une documentation est faite lors de la conception et de l'acquisition programmes et progiciels. Cette documentation concerne également les travaux menés sur les réseaux informatiques.

## **14.6. Archivage des documents méthodologiques et des rapports d'analyse**

Les documents méthodologiques, c'est l'ensemble des étapes permettant de chercher, d'identifier et trouver des documents relatifs à un sujet par l'élaboration d'une stratégie de recherche tandis que le rapport d'analyse est un document qui fait une description des points forts et les limites des différentes mesures.

L'archivage des documents méthodologiques est systématiquement effectué après l'opération, sous la responsabilité de la Direction en charge de l'archivage et de la direction technique concernée. Les documents sur supports papiers sont déposés au centre de documentation de l'Institut pour archivage et ceux sous format électronique sont archivés sur CDROM. Une synthèse de ces documents sera publiée sur le site de l'INSD en format PDF. Ces fichiers ne peuvent faire l'objet d'aucune rétrocession sans l'avis du comité scientifique.

## **14.7. L'archivage et la sécurisation des données**

L'archivage est essentiel pour permettre à la fois de sécuriser les données produites et les mettre à la disposition des utilisateurs. Les bases de données finales des enquêtes et recensement des différentes directions sont systématiquement centralisés pour les besoins d'archivage et de sécurisation au niveau de la direction en charge de l'archivage des données. Un serveur des bases de données est dédié à cet effet. Des copies de ces bases de données sont également faites sur des serveurs de secours logés dans des structures externes moyennant la signature de protocole d'accord entre les différentes parties. Les outils utilisés sont :

- TOOLKIT ;
- IMIS/REDATAM.

La direction en charge de l'archivage des données est responsable de la centralisation et de la gestion des bases de données. Des sessions de formations à l'usage de ces outils ont lieu chaque année à l'intention des agents pour assurer l'expertise en matière de gestion des bases de données et leur mise à jour régulière.

## **14.8. Système d'archivage électronique**

Le Système d'Archivage électronique (SAE) des données de l'INSD est conçu donc comme un système consistant à recevoir, conserver, traiter, restituer des lots de données, des paquets d'informations, et qui s'appuie sur une plateforme informatique.

Le système d'archivage électronique des données de l'INSD répond aux objectifs suivants :

Il permet une conservation durable des micro-données et des métadonnées créées dans le cadre de chaque opération statistique, qu'il s'agisse des enquêtes, des recensements ou de l'exploitation de données administratives ; la diffusion de ces données est assurée selon les règles définies par l'INSD et le cadre réglementaire.

Il permet la conservation dans une base de données unique des principales données statistiques produites (macro-données économiques et sociales) et facilite leur diffusion sous forme de publications ou via internet.

Il garantit la continuité de la sécurité des données traitées par l'INSD sur les postes de travail et dans les espaces partagés sur le réseau local par des mécanismes de contrôle des accès et de sauvegarde.

Le stockage à long terme des documents et données numériques constitue le principe fondamental de l'archivage électronique des données, cependant les problématiques liées sont le coût et la durée de vie des supports, mais aussi l'accès au contenu malgré les avancées technologiques rendant les anciens supports obsolètes. Ainsi la gestion électronique des documents (GED) contribue à une bonne diffusion des données.

## **15 PHASE D'ÉVALUATION**

Au cours de cette étape, il s'agit d'évaluer le Système Statistique National (SSN), l'environnement institutionnel, la chaîne de production des données ou des produits statistiques selon les grilles d'évaluation consacrées par le cadre d'assurance qualité et de produire un rapport qualité des données statistiques. Elle permet à l'équipe de cerner les forces et les faiblesses du processus et de faire des recommandations sur les améliorations potentielles dans l'avenir. Cette étape, la dernière de toute activité de production statistique, est généralement effectuée par l'ensemble de l'équipe y ayant participé.

Structures responsables : équipe chargée de l'évaluation.

Autres structures concernées : SERDM, Les Directions Techniques plus les Directions Générales.

L'étape de l'évaluation comporte trois sous étapes, qui se déroulent généralement en ordre séquentiel. Ces sous étapes sont les suivantes :

### **15.1. Recueil des contributions d'évaluation**

Cette sous étape consiste à rassembler toute la documentation produite durant les différentes étapes du processus de production statistique et la mettre ensuite à la disposition de l'équipe chargée de l'évaluation. Cette documentation se trouve sous plusieurs formes, y compris les commentaires des utilisateurs, les métadonnées et para données sur les processus, les mesures des systèmes, et les suggestions du personnel. Les rapports de suivi du plan d'action du projet doivent également servir de données d'entrée à l'évaluation.

### **15.2. Réalisation de l'évaluation**

Cette sous étape consiste à analyser les données recueillies à la sous étape précédente et d'en faire la synthèse dans un rapport d'évaluation. Ce dernier devra faire état de tous les problèmes rencontrés au cours du processus, comment ils ont été gérés et les résultats engendrés, et recommander des changements au besoin. Ces recommandations peuvent être sur les changements à apporter à une étape ou une sous étape quelconque du processus ou peuvent consister à proposer que le processus statistique en entier ne soit pas reconduit. Il est également important de noter les processus qui ont bien fonctionné dans l'évaluation.

### **15.3. Recherche de consensus sur un plan d'action**

Cette sous étape implique de chercher le consensus auprès de la haute gestion sur la nécessité d'élaborer un plan d'action basé sur les recommandations du rapport et sur les moyens de sa mise en œuvre. Ce plan doit inclure certains mécanismes de surveillance des répercussions des mesures prévues, lesquels peuvent, à leur tour, servir de base aux évaluations futures du processus.

## CONCLUSION

Ce présent manuel est un guide pratique général pour la conception et la conduite d'enquêtes et recensements sur n'importe quel type d'unités statistiques (individu, ménage, entreprises, ...). Le manuel n'est pas une description détaillée des différentes phases de processus de productions statistiques. Dans chaque phase, il donne les orientations précises sur les principales activités à réaliser et les documents factuels qui attestent de leur effectivité. Le manuel est structuré conformément aux phases nécessaires pour conduire une enquête ou un recensement en partant de la définition des besoins de statistiques jusqu'à la phase de l'évaluation de la qualité des données. Chaque phase du processus de production est décomposée en des sous-phases et les sous-phases sont accompagnées d'orientations techniques qui contribuent à leur réalisation.

L'application de manuel répond aux exigences d'encadrement de la qualité des produits statistiques au sein de l'INSD et leur amélioration progressive.

## BIBLIOGRAPHIE

- [1.] CEDEAO (2011), cadre d'évaluation des normes et standards des systèmes d'information de gestion de l'éducation dans la région CEDEAO ;
- [2.] CNUCED (2010) Manuel pour la production de statistiques sur l'économie de l'information
- [3.] DGESS/MENA (2017), référentiel sur l'harmonisation et la cohérence de la collecte des données statistiques du ministère de l'éducation nationale et de l'alphabétisation ;
- [4.] INSD (2020), Cadre d'assurance qualité
- [5.] INSD (2014), charte graphique
- [6.] INSD (2014), politique de diffusion
- [7.] STATISTIQUE CANADA (2019) Méthodes et pratiques d'enquête ;
- [8.] STATISTIQUE CANADA (2019), Lignes directrices concernant la qualité
- [9.] STATISTIQUE Congo, Manuel de Procédures de Production des Statistiques de l'Emploi
- [10.] UNION AFRICAINE (2015), Cadre d'Assurance de la Qualité des Statistiques pour le Système Statistique Africain ;

# ANNEXES

## Fiche type de métadonnées

Les informations fournies pour chaque indicateur, dans la mesure de leur disponibilité, sont les suivantes :

Indicateur	Dénomination de l'indicateur
Code	Code de l'indicateur
Définition	Définition de l'indicateur
Objectif et cible visés	Si disponibles, objectifs et cibles
Raison d'être	Bien fondé de l'indicateur
Méthode de calcul	Formule de calcul de l'indicateur
Sources de données	Indications des données rentrant dans la formule de calcul de l'indicateur, des opérations par lesquelles elles sont obtenues et des structures qui les produisent
Périodicité des mesures	Fréquence à laquelle l'indicateur est produit
Considérations sexo-spécifiques	Situation de l'indicateur par rapport aux questions liées au sexe
Désagrégation	Niveaux de désagrégation possibles de l'indicateur (désagrégation géographique et autres)
Observations et limites	Clarifications, spécifications, réserves, limites ou toute information supplémentaire permettant de mieux comprendre l'indicateur
Propositions d'amélioration	Suggestions pouvant permettre une amélioration de l'indicateur
Structure responsable	Structure qui produit ou diffuse l'indicateur

**BURKINA FASO**  
**CONSEIL NATIONAL DE LA STATISTIQUE**  
**COMMISSIONS SPECIALISEES DU CONSEIL NATIONAL DE LA**  
**STATISTIQUE**

**FICHE DE RENSEIGNEMENT POUR L'OCTROI DE L'AUTORISATION**  
**PRELABLE (VISA STATISTIQUE)**

<b>IDENTITE DE LA STRUCTURE DEMANDEUSE</b>	
Ministère	
Nom de la structure	
Boîte Postale de la structure	
Adresse mail de la structure	
Nom et Prénom du premier responsable	
Numéro de téléphone du premier responsable	
Adresse mail du premier responsable	
Nom et Prénom du responsable de l'enquête	
Numéro de téléphone du responsable de l'enquête	
Adresse mail du responsable de l'enquête	

<b>INFORMATIONS SUR L'ENQUETE</b>
-----------------------------------

Titre de l'enquête	
Existence dans le Schéma directeur de la statistique (SDS)	Encercler le OUI ou le NON <i>OUI</i> <i>NON</i> <b>Si NON pourquoi ?</b>
Existence dans le Programme statistique national (PSN)	Encercler le OUI ou le NON <i>OUI</i> <i>NON</i> <b>Si NON pourquoi ?</b>
Objectif principal de l'enquête	
Principaux résultats attendus	Résultat 1 :  Résultat 2 :
Univers (Population concernée) de l'enquête	
Taille de l'échantillon	
Taux de sondage	
Dates prévisionnelles de début-fin de l'enquête	
Observations	

Liste Indicateurs du PNDES

Liste Indicateurs clés sectoriels : santé, démographie, économie, emploi, éducation, environnement, agriculture, eau et assainissement,

Agenda 2030

Agenda 2063

# TABLE DES MATIERES

AVANT PROPOS .....	I
SOMMAIRE.....	II
LISTE DES TABLEAUX .....	III
ABREVIATIONS.....	IV
DEFINITION DES CONCEPTS .....	VI
INTRODUCTION .....	1
<b>1 CADRE JURIDIQUE ET INSTITUTIONNEL DU PROCESSUS DE PRODUCTION DES STATISTIQUES.....</b>	<b>2</b>
1.1. Au plan institutionnel .....	2
1.2. Au plan juridique .....	2
<b>2 DESCRIPTION DU MODELE DE GESTION DE L'INSD.....</b>	<b>3</b>
<b>3 DESCRIPTION DU MODELE GENERIQUE DU PROCESSUS DE PRODUCTION STATISTIQUE .....</b>	<b>4</b>
<b>4 DEFINITION DES ROLES ET RESPONSABILITES.....</b>	<b>7</b>
<b>5 PHASE DE DEFINITION DES BESOINS.....</b>	<b>10</b>
5.1. Détermination des besoins d'information .....	10
5.2. Elaboration de la liste des utilisateurs potentiels .....	10
5.3. Consultation des utilisateurs sur les besoins.....	10
5.4. Validation des besoins des utilisateurs .....	10
5.5. Etablissement des objectifs et des produits .....	10
5.6. Identification des concepts et des variables .....	11
5.7. Vérification de la disponibilité des données.....	11
5.8. Elaboration de l'étude de faisabilité .....	11
Conclusion partielle.....	11
<b>6 PHASE DE TRAVAUX PRELIMINAIRES .....</b>	<b>12</b>
6.1. Elaboration du protocole de l'opération .....	12
6.1.1 <i>Contexte et justification</i> .....	12
6.1.2 <i>Objectifs de l'étude</i> .....	12
6.1.3 <i>Méthodes</i> .....	12
6.1.4 <i>Ressources humaines</i> .....	13
6.1.5 <i>Ressources matérielles</i> .....	13
6.1.6 <i>Budget et plan de financement</i> .....	13
6.1.7 <i>Chronogramme</i> .....	13
6.1.8 <i>Validation du protocole par un comité d'éthique</i> .....	13
6.1.9 <i>Validation du protocole par la commission d'informatique et de liberté</i> .....	13
6.2. Elaboration du document de projet/ convention .....	13
6.3. Elaboration du document d'une convention .....	14
6.4. Mobilisation des ressources financières.....	14
6.5. Planification des activités.....	14
6.6. Élaboration du calendrier de l'opération .....	14
Conclusion partielle.....	14
<b>7 PHASE DE CONCEPTION.....</b>	<b>15</b>

7.1.	Conception des produits/ indicateurs .....	15
7.2.	Élaboration de la description des variables .....	15
7.3.	Définition de la méthode de collecte des données .....	15
7.4.	Définition de la méthodologie d'échantillonnage.....	16
7.5.	Conception de la méthodologie de traitement statistique (codification, apurement et tabulation).....	16
7.6.	Conception de l'analyse des données.....	16
7.7.	Conception du plan de diffusion.....	16
	Conclusion partielle.....	17
<b>8</b>	<b>PHASE DE CONSTRUCTION.....</b>	<b>19</b>
8.1.	Elaboration ou actualisation des outils de collecte .....	19
8.1.1	<i>Elaboration du questionnaire .....</i>	<i>19</i>
8.1.2	<i>Elaboration du masque de saisie/application de collecte .....</i>	<i>19</i>
8.1.3	<i>Elaboration des manuels .....</i>	<i>19</i>
8.2.	Elaboration des programmes de traitements.....	20
8.3.	Elaboration du plan d'analyse.....	20
8.4.	Test du système de production.....	20
8.5.	Test du processus de production.....	20
8.6.	Finalisation du système de production .....	20
8.7.	Constitution ou actualisation de la base de sondage.....	21
	Conclusion partielle.....	21
<b>9</b>	<b>PHASE DE DEMANDE DE VISA STATISTIQUE.....</b>	<b>22</b>
9.1.	Préparation du dossier de visa statistique .....	22
9.2.	Envoi du dossier.....	22
9.3.	Programmation de réunion(s) de Commission spécialisée et l'octroi du visa statistique.....	22
9.4.	Validation du protocole par un comité d'éthique .....	23
<b>10</b>	<b>PHASE DE COLLECTE .....</b>	<b>24</b>
10.1.	Tirage de l'échantillon.....	24
10.1.1	<i>Bases de sondage et couverture .....</i>	<i>24</i>
10.1.2	<i>Taille de l'échantillon.....</i>	<i>24</i>
10.1.3	<i>Méthodes d'échantillonnage.....</i>	<i>25</i>
10.2.	Mise à jour de l'échantillon .....	25
10.3.	Mise en place du processus de collecte.....	26
10.3.1	<i>Recrutement du personnel de collecte .....</i>	<i>26</i>
10.3.2	<i>Identification des locaux pour la formation des agents recenseurs ; .....</i>	<i>26</i>
10.3.3	<i>Evaluation et sélection définitive des participants formés ; .....</i>	<i>26</i>
10.4.	Sensibilisation .....	26
10.5.	Collecte des données .....	26
10.5.1	<i>Matériel de la collecte (quantité, distribution, etc).....</i>	<i>26</i>
10.5.2	<i>Mise en place du personnel du terrain ; .....</i>	<i>26</i>
10.5.3	<i>Numérotation des concessions et des ménages ; .....</i>	<i>26</i>
10.5.4	<i>Remplissage des questionnaires.....</i>	<i>27</i>
10.5.5	<i>Contrôle des questionnaires .....</i>	<i>27</i>
10.5.6	<i>Supervision du dénombrement .....</i>	<i>27</i>
10.5.7	<i>Transfert des données.....</i>	<i>27</i>
10.5.8	<i>Finalisation de la collecte des données.....</i>	<i>27</i>
10.5.9	<i>Rapport de collecte.....</i>	<i>27</i>
<b>11</b>	<b>PHASE DE TRAITEMENT.....</b>	<b>28</b>

11.1.	Profilage et préparation des données .....	28
11.2.	Profilage des données .....	28
11.3.	Standardisation, classification et codage .....	28
11.4.	Contrôle et imputation .....	29
11.5.	Ajustement et repondération.....	29
11.6.	Dérivation de nouvelles variables et d'unités statistiques .....	29
11.7.	Évaluation et documentation de l'incidence des changements .....	29
11.8.	Mise au point finale des fichiers de données.....	29
11.9.	Intégration, estimation et compilation.....	30
11.10.	Détermination des éléments, des règles et de la stratégie d'intégration .....	30
11.11.	Évaluation et ajustement de la stratégie d'intégration .....	31
11.12.	Chargement, application des mises en correspondance et intégration des données sources	31
11.13.	Vérification et ajustement d'amélioration de la qualité .....	32
<b>12</b>	<b>PHASE D'ANALYSE.....</b>	<b>33</b>
12.1.	Préparation de la version préliminaire des résultats .....	33
12.2.	Validation des résultats .....	33
12.3.	Analyse descriptive des données.....	33
12.4.	Analyse explicative des données (en cas de besoin) .....	33
12.5.	Elaboration rapport de l'opération .....	33
12.6.	Application des contrôles de confidentialité.....	33
12.7.	Finalisation des résultats .....	34
<b>13</b>	<b>PHASE DE DIFFUSION .....</b>	<b>35</b>
13.1.	Mettre à jour les systèmes de diffusion .....	35
13.2.	Elaboration des produits de diffusion .....	35
13.3.	Gestion de la sortie des produits de diffusion.....	35
13.4.	Promotion des produits de diffusion.....	36
13.5.	Assistance aux utilisateurs .....	36
<b>14</b>	<b>PHASE D'ARCHIVAGE .....</b>	<b>37</b>
14.1.	Définition du système d'archivage des données.....	37
14.2.	Archivage des questionnaires remplis.....	37
14.3.	Collecte des données à archiver.....	37
14.4.	Information à archiver.....	37
14.4.1	<i>documents réglementaires de l'opération</i> .....	37
14.4.2	<i>Outils de collecte de données</i> .....	37
14.4.3	<i>Données de base</i> .....	37
14.4.4	<i>métadonnées des indicateurs</i> .....	38
14.4.5	<i>notes méthodologiques</i> .....	38
14.4.6	<i>nomenclatures et classifications utilisées</i> .....	38
14.4.7	<i>Rapports, plans de traitement</i> .....	38
14.4.8	<i>produits statistiques de l'opération</i> .....	38
14.5.	Archivage des programmes informatiques .....	39
14.6.	Archivage des documents méthodologiques et des rapports d'analyse .....	39
14.7.	L'archivage et la sécurisation des données.....	39
14.8.	Système d'archivage électronique .....	39
<b>15</b>	<b>PHASE D'ÉVALUATION.....</b>	<b>41</b>
15.1.	Recueil des contributions d'évaluation .....	41
15.2.	Réalisation de l'évaluation .....	41
15.3.	Recherche de consensus sur un plan d'action .....	41

<b>CONCLUSION.....</b>	<b>42</b>
<b>BIBLIOGRAPHIE .....</b>	<b>43</b>
<b>ANNEXES .....</b>	<b>44</b>
Fiche type de métadonnées .....	44
Agenda 2030 .....	46
Agenda 2063 .....	46
<b>TABLE DES MATIERES .....</b>	<b>47</b>

DCSFR/2024/03  
mars 2024

Institut national de la statistique et de la démographie (INSD)  
Avenue Pascal ZAGRE, Ouaga 2000  
01 B.P : 374 Ouagadougou 01 – Burkina Faso  
Tél : (00226) 50 37 62 04 - Fax : (00226) 50 37 62 26  
Site internet : [www.insd.bf](http://www.insd.bf) - Email : [insd@insd.bf](mailto:insd@insd.bf) ou [insbf@yahoo.fr](mailto:insbf@yahoo.fr)