

FONDS EUROPEEN DE DEVELOPPEMENT (FED) NIGER
**RENFORCEMENT DU SYSTEME D'INFORMATION STATISTIQUE
DE L'EDUCATION-FORMATION (SISEF)**

**OPERATIONNALISATION D'UN SYSTEME DE CONTROLE
QUALITE RENFORCE ET D'UN SYSTEME DE VALIDATION DU
SISEF : MANUEL DE CONTROLE QUALITE ET DE VALIDATION DU
SISEF**

DOCUMENT FINAL

JUIN 2020

1. INTRODUCTION	3
1.1 Contexte	3
1.2 Objectif.....	3
2. APPUI DE L'INS AU SUIVI DE LA QUALITE DES DONNEES DU SISEF.....	5
3. CADRE D'EVALUATION DE LA QUALITE	6
3.1 Introduction de la démarche méthodologique	6
3.2 Cadre National d'Assurance Qualité des Statistiques Publiques.....	6
3.3 Choix des dimensions de la qualité	7
3.3.1 Rigueur méthodologique	7
3.3.2 Exactitude et fiabilité	8
3.3.3 Actualité et ponctualité	8
3.3.4 Cohérence et comparabilité.....	8
3.4 Choix des phases du processus statistique.....	9
3.4.1 Elaboration.....	9
3.4.2 Collecte	9
3.4.3 Traitement des données	9
3.4.4 Analyse des résultats	10
3.4.5 Diffusion.....	10
4. INDICATEURS D'ÉVALUATION	10
5. PROCESSUS DE VALIDATION.....	13
5.1 Indicateurs de validation	13
5.2 Processus.....	13
6. ANNEXES.....	15
6.1 Annexe 1 : Liste des questions à double saisir.....	15
6.2 Annexe 2 : Contenu du rapport d'exécution.....	16
6.3 Annexe 3 : Contenu du rapport de traitement	17
6.4 Annexe 4 : Contenu du rapport méthodologique	18
6.5 Annexe 5 : Format du calendrier du recensement.....	19
6.6 Annexe 6 : Calendrier indicatif.....	20
6.7 Annexe 7 : Fiche de contrôle des questionnaires	21
6.8 Annexe 8 : Fiche d'apurement des questionnaires.....	23
6.9 Annexe 9 : Niveaux 1 et 2 du modèle statistique général du processus opérationnel.....	24
6.10 Annexe 10 : Liste des documents consultés	25

1. INTRODUCTION

1.1 Contexte

L'INS en tant qu'organe central du système statistique national (SSN) est chargé de valider les productions statistiques des services sectoriels. Dans la pratique autrefois, cette validation par l'INS intervenait en fin de processus, sans informations complémentaires sur ses modalités de production et sur tout le processus de collecte, d'exploitation des données et l'élaboration des indicateurs. Les directions des Statistiques (DS/DSI) envoyaient souvent leurs publications en produit fini à la Direction de la Coordination et du Management des Informations Statistiques de l'INS (INS/DCMIS) pour validation. Aucun délai légal n'était fixé à l'INS pour répondre. Le processus incluait toujours un atelier de validation organisé par les directions des statistiques sectorielles auquel l'INS participait.

En novembre 2019, un Comité technique a été créé au sein de l'INS. Ce comité technique est un organe de mission, rattaché à la Direction Générale de l'INS, et est composé de cadres du système statistique national (SSN). Ce comité est chargé de la validation des principaux indicateurs socioéconomiques, des agrégats macroéconomiques et des différentes publications du SSN. Le comité a pour principale mission d'examiner, en amont, les documents de projet des opérations statistiques d'envergure, les dispositifs et les outils techniques mis en œuvre, de réaliser le suivi des opérations de collecte de données statistiques et d'en valider les résultats. Si ces attributions sont définies, ses modalités de fonctionnement ne sont pas encore précisées.

Dans le cadre du renforcement du système d'informations du secteur de l'éducation et de la formation (SISEF), il est prévu l'opérationnalisation d'un système de contrôle qualité renforcé, ainsi que celle d'un système de validation. Le SISEF couvre les ministères de l'Enseignement Primaire, de l'Alphabétisation, de la Promotion des Langues Nationales et de l'Education Civique (MEP/APLN/EC), des Enseignements Secondaires (MES), de l'Enseignement Supérieur de la Recherche et de l'Innovation (MESR/I), des Enseignements Professionnels et Techniques (MEP/T), de la Jeunesse et des Sports (MJ/S) et de la Renaissance Culturelle, des Arts et de la Modernisation Sociale (MRCA/MS).

1.2 Objectif

Le dispositif d'appui au secteur prévoit l'accompagnement de l'INS dans les phases de préparation, de collecte et de traitement des recensements. Cet appui est important pour sensibiliser et former les personnels des directions sectorielles, mais il doit être accompagné de la production d'un rapport qualité en fin de processus sur lequel s'appuiera le comité pour la validation d'un annuaire. Ce manuel décrit le contenu du rapport qualité attendu.

L'objectif, de ce document est d'identifier un ensemble d'indicateurs permettant d'évaluer la qualité des statistiques produites par le SISEF et de définir les modalités de leur validation, tant en termes de critères, que de processus.

Ce document est divisé en quatre (4) parties :

- La première partie présente l'appui de l'INS aux structures sectorielles pour améliorer la qualité des données.
- La seconde partie propose un cadre d'analyse s'appuyant sur le Cadre d'assurance qualité des statistiques publiques développé par l'INS et sur les dimensions d'un processus telles que définies au niveau international.
- La troisième partie propose une liste d'indicateurs permettant d'évaluer la qualité des indicateurs statistiques.
- La dernière partie traite des modalités de validation, à savoir qui validera et selon quels critères ?

Avertissement : ce manuel a été développé en s'appuyant sur les enseignements tirés des différents travaux de vérification menés par une équipe de la DS/MEP, de la CAMOS, de la CAPEG et de l'INS, sur les données du recensement des établissements d'enseignement primaire. Il est donc très orienté par les résultats de ces travaux. Les modèles des fiches de contrôle et d'apurement proposés en annexe sont spécifiquement adaptés aux recensements des établissements d'enseignement primaire et des enseignements secondaires.

Il s'agit d'un document expérimental qui sera testé, révisé et complété pour prendre en compte les spécificités des différentes sous-composantes du secteur de l'éducation et de la formation.

Ce manuel sera finalisé au cours de la période couverte par la subvention de l'Union européenne pour l'appui au renforcement du système d'information statistique du secteur de l'éducation et de la formation.

2. APPUI DE L'INS AU SUIVI DE LA QUALITE DES DONNEES DU SISEF

Dans le cadre de la subvention UE au SISEF, il est prévu que, l'INS fournisse un appui au SISEF dans les différentes phases du processus de production au travers notamment de ses directions régionales (DR). Il s'agira d'un accompagnement de l'INS qui travaillera en étroite collaboration avec les structures sectorielles. L'INS ne se substituera en aucun cas aux structures sectorielles qui ont la responsabilité de la production statistique sectorielle.

Cet appui de l'INS interviendra :

- Au niveau des formations des différents personnels impliqués dans le processus de collecte et d'exploitation des données statistiques du SISEF auxquelles les DR de l'INS participeront.
- Au niveau des contrôles des questionnaires tout le long de la chaîne de collecte des données (par sondage aléatoire), dans les secteurs pédagogiques, dans les inspections, les DREP et les DRES.

Ces contrôles porteront sur la complétude du questionnaire, et sur la cohérence des données sur les variables clés relatives aux salles de classe, aux enseignants et aux effectifs.

Ces contrôles seront réalisés au cours des missions de supervision, menées sur l'ensemble des inspections. Ces missions seront menées conjointement par les DR de l'INS et les DREP. Selon le type d'anomalies constatées les corrections pourront être réalisées directement, ou nécessiteront de contacter les directeurs.

Une fiche de contrôle sera annexée à ce document, en annexe 6. Elle sera testée lors des prochains recensements et sera révisée après la refonte des questionnaires.

- Au niveau de la saisie des données en réalisant une double saisie d'un lot de questionnaires. Cette opération sera destinée à vérifier la qualité de la saisie des données, à la fois en termes d'erreurs et de complétude (saisie complète du questionnaire).

La double saisie sera effectuée sur 10% des questionnaires sur les variables clés : la localisation et l'identification de l'établissement, les salles de classe, le personnel enseignant et les effectifs. Elle sera effectuée quotidiennement, en s'assurant que tous les opérateurs sont effectivement contrôlés, et ce pendant toute la durée de la saisie.

La liste des variables concernées par la double saisie pour le recensement des établissements primaires est produite en annexe 1.

Le taux d'erreur quotidien acceptable sur les variables clés sera fixé à 5%. Cela signifie que si, sur 100 questionnaires saisis par un agent, des erreurs ou omissions (informations non saisies) sont observées sur un maximum de 5 questionnaires la saisie est acceptée. Dans le cas contraire la saisie du jour effectuée par l'agent sera totalement reprise.

La fiche d'apurement sera proche de celle du contrôle, elle est présentée en annexe 8. Elle sera testée lors des prochains recensements et sera révisée après la refonte des questionnaires.

- Au niveau de l'apurement des données, pour limiter les délais de production, il est conseillé d'apurer les données en continu, et non en fin de saisie. La mise en œuvre de toutes les procédures décrites ci-dessus réduira substantiellement la durée de l'apurement, qui est aujourd'hui d'environ 3 mois pour le recensement du MEP.

Actuellement, l'apurement au niveau du MEP repose sur des requêtes automatiques, développées sous StatEduc 2. Il s'agit principalement de contrôles de cohérence des données, entre deux ou plusieurs questions. Par exemple :

- Un enseignant fonctionnaire ne peut pas faire la classe dans une école privée.
- Il ne peut pas y avoir 2 directeurs dans une école.

L'application n'intègre pas de corrections automatiques, ceci ralentit considérablement la phase de correction dans sa durée. Elle n'inclut pas de contrôle de vraisemblance (cas d'une école qui avait 150 élèves l'année précédente et 16 élèves seulement cette année, il y a de grandes chances que 16 soit une erreur de saisie). Cette anomalie n'est pas détectée pendant la saisie dans l'application actuelle.

Il est prévu, dans le cadre de la subvention de l'UE, une révision de StatEduc. Ce sera l'occasion d'intégrer ces différents contrôles et des corrections automatiques.

- Au niveau des outils de collecte, avec la révision des questionnaires, et la production de manuels d'enquêtes (manuels de remplissage, de contrôle, de traitement, ...).

Ce dispositif sera complété par l'élaboration d'un rapport qualité, qui accompagnera la production de l'annuaire statistique, et qui permettra d'évaluer sa qualité et en facilitera le processus de validation.

L'objectif de ce rapport sera également de servir de guide, lors de la réalisation du recensement. Ainsi, le taux de couverture du recensement, qui constitue un indicateur clé de validation, devra être suivi pendant toute la durée de la collecte.

Outre sa fonction d'outil de pilotage pour les ministères sectoriels, ce rapport permettra de constituer une mémoire des opérations de production. Cette mémoire est nécessaire, en raison notamment de la mobilité du personnel.

Finalement, ce dispositif, d'auto-évaluation par les producteurs, favorisera la crédibilité des données et donc leur utilisation.

3. CADRE D'EVALUATION DE LA QUALITE

3.1 Introduction de la démarche méthodologique

La démarche a consisté à produire un cadre d'analyse à partir :

- D'une part, de l'identification des dimensions de la qualité qui se concentreront sur les processus et les produits statistiques ;
- D'autre part, de la sélection des étapes du processus statistique pertinentes pour l'exercice, au sens où ces étapes concerneront des activités mises en œuvre à chaque vague de recensement. Par exemple, l'élaboration du questionnaire est réalisée ponctuellement, alors que la mise à jour du fichier des établissements scolaires devra être faite chaque année.

Une fois le cadre d'analyse déterminé, il s'agira d'identifier des indicateurs simples à renseigner et non ambigus. L'accent sera mis sur :

- Des informations déjà disponibles et qui ne nécessiteront pas un travail conséquent pour les structures productrices ;
- Des indicateurs factuels n'impliquant aucune perception mais uniquement des réalisations.

3.2 Cadre National d'Assurance Qualité des Statistiques Publiques

Au sein d'un système statistique national, la gestion de la qualité s'appuie sur un cadre national d'assurance de la qualité (CNAQ). Il existe plusieurs cadres d'évaluation de la qualité des données. Le cadre d'évaluation de la qualité des données (CEQD) développé par le FMI en 2003 basé sur les principes fondamentaux de la statistique officielle. Ce CEQD est une démarche globale qui tient compte des contextes institutionnels, des processus statistiques et des caractéristiques des produits statistiques. Il a été décliné pour certains secteurs tels que l'éducation. L'OCDE, Eurostat (Code de bonnes pratiques de la statistique européenne), Statistique Canada (Cadre d'Assurance Qualité), ont développé leur propre cadre d'évaluation de la qualité.

L'Institut National de la statistique du Niger a récemment produit son Cadre National d'Assurance Qualité des Statistiques Publiques (CNAQ/SP). Le CNAQ/SP, adopté en août 2017, présente les principaux domaines, les principales dimensions ainsi que les indicateurs relatifs à l'évaluation de la qualité des données statistiques officielles. La qualité statistique s'appuie sur 4 domaines interdépendants qui doivent être pris en compte dans le processus de production. Ces domaines sont relatifs à : (i) la gestion du Système Statistique National (SSN) ; (ii) la gestion de l'environnement institutionnel ; (iii) la gestion des processus de production des données statistiques publiques officielles ; et (iv) la gestion des produits statistiques du SSN.

Le CNAQ/SP du Niger compte huit (8) dimensions principales dimensions :

- Conditions préalables de la qualité ;

- Exactitude et la fiabilité ;
- Rigueur méthodologique ;
- Intégrité ;
- Utilité ;
- Accessibilité ;
- Actualité, ponctualité ;
- Cohérence et comparabilité.

Les **conditions préalables de la qualité** font référence à l'appréciation du cadre institutionnel, aux ressources humaines, matérielles et financières.

L'**exactitude et fiabilité** se rapportent au principe selon lequel les données produites donnent une image adéquate de la réalité socio-économique.

La **rigueur méthodologique** relève du principe selon lequel la production de statistiques doit reposer sur une base méthodologique rigoureuse, ce qui peut être assuré par l'application de normes, principes ou bonnes pratiques acceptés au niveau international.

L'**intégrité** consiste au respect du principe d'objectivité dans le processus de production statistique.

L'**utilité** se réfère à ce que l'information statistique doit répondre aux besoins des utilisateurs et être produite à temps.

L'**accessibilité** se réfère à ce que les données et métadonnées doivent être présentées de manière claire et intelligible.

L'**actualité et ponctualité** font référence à la diffusion des statistiques en temps utile et aux moments prévus.

La **cohérence et comparabilité** se réfère à ce que les statistiques doivent présenter une cohérence interne et dans le temps et permettre la possibilité de comparer entre les régions et les pays.

Pour chacune de ces dimensions, le cadre identifie des éléments et des indicateurs normatifs qui doivent être évalués.

3.3 Choix des dimensions de la qualité

Le choix se concentrera sur des dimensions qui permettent d'évaluer la qualité des procédures et des résultats.

Concernant les **procédures**, la dimension rigueur méthodologique garantit que la production statistique respecte les bonnes pratiques.

La qualité des **résultats** pourra être appréhendée par les dimensions : exactitude et fiabilité, actualité et ponctualité, cohérence et comparabilité.

3.3.1 Rigueur méthodologique

La rigueur méthodologique est illustrée par l'application de normes, principes ou bonnes pratiques acceptés au niveau international. Elle se compose de la conformité de quatre éléments aux normes internationales :

- Les concepts et définitions ;
- Le champ d'application ;
- Les classifications ;
- La base d'enregistrement.

Cette dimension s'intéressera aux éléments qui sont susceptibles de changements d'une vague d'un recensement à l'autre, il s'agira du champ d'application et de la base d'enregistrement.

- ✓ Les concepts et définitions seront évalués au travers du manuel méthodologique et de sa mise à jour. L'existence/mise à jour des autres outils de collecte (manuels de remplissage, de contrôle, ...) sera également considéré.

- ✓ Le champ d'application pose la question de l'existence d'un fichier des établissements. Sa mise à jour (selon la création ou la fermeture d'une école) relèvera de la dimension exactitude et fiabilité.
- ✓ La base d'enregistrement se référera à la période de référence qui peut changer d'une année à l'autre suite à la mobilisation des financements.

3.3.2 Exactitude et fiabilité

Cette dimension reflète la spécificité des sources et de leurs traitements. Les éléments de cette dimension sont :

- Les données de base sont appropriées pour la production des statistiques ;
- Les techniques statistiques sont saines pour traiter les données de base ;
- L'évaluation et validation des résultats intermédiaires et des produits statistiques ;
- Les études de révision.

Les indicateurs suivants seront considérés :

- ✓ La mise à jour du fichier des établissements (données de base) ;
- ✓ Le respect de la période de collecte : le % d'établissements ayant renseigné le questionnaire dans les délais (données de base) ;
- ✓ La documentation de l'enquête et l'existence des différents manuels d'enquête (données de base) ;
- ✓ Le taux de non-réponse et de données erronées pour les variables de base (données de base) ;
- ✓ Les méthodes d'imputation des données manquantes (techniques statistiques) ;
- ✓ L'existence d'un rapport d'erreurs et d'incohérences à partir des contrôles manuels et informatiques (évaluation et validation).

3.3.3 Actualité et ponctualité

L'actualité fait référence au délai entre le moment où l'information est collectée et celui où elle est diffusée. La ponctualité se réfère au respect du calendrier de diffusion.

- ✓ L'indicateur choisi sera le respect de la date de diffusion. En cas de non-respect, il sera demandé d'indiquer les raisons.

3.3.4 Cohérence et comparabilité

Cette dimension s'intéresse à :

- La cohérence entre les niveaux de données et à la cohérence des séries chronologiques pour une période suffisante ;
- La reconstruction des séries lorsque les données de base, les méthodologies ou les techniques statistiques sont modifiées ;
- L'existence de notes méthodologiques qui expliquent les ruptures et discontinuités dans les séries chronologiques ;
- La confrontation des données issues de sources différentes (données de routine et données issues d'enquêtes).

L'indicateur se limitera à :

- ✓ Concordance des données clés avec celles des 4 années précédentes au niveau régional : salles de classe, enseignants et effectifs. Les raisons qui expliqueront la non-concordance des données seront à préciser.

3.4 Choix des phases du processus statistique

Les groupes de travail communs UNECE / Eurostat / OCDE sur les métadonnées statistiques (METIS) ont élaboré le modèle générique du processus de production statistique (GSBPM). Le GSBPM définit l'ensemble des processus d'activité nécessaires à l'établissement de statistiques officielles.

Le GSBPM comprend trois niveaux :

- Le niveau 0, qui est celui du processus de production statistique ;
- Le niveau 1, qui correspond aux huit phases du processus de production statistique ;
- Le niveau 2, qui correspond aux sous-processus de chaque phase.

Les neuf étapes du processus de production statistique sont les suivantes :

1. Définition des besoins ;
2. Conception ;
3. Elaboration ;
4. Collecte ;
5. Traitement ;
6. Analyse ;
7. Diffusion ;
8. Archivage ;
9. Evaluation.

Un diagramme indiquant les phases (niveau 1) et les sous-processus (niveau 2) est reproduit à l'annexe 9.

Les trois premières phases ont généralement lieu lors de la première mise en œuvre ou en cas de révision des outils. L'étude de la conformité des processus se concentrera sur les phases de l'élaboration (3), la collecte (4), du traitement (5), de l'analyse (6) et de la diffusion (7). La phase d'élaboration a été conservée pour analyser les modalités de préparation du recensement.

3.4.1 Elaboration

Ce sous-processus inclut :

- L'élaboration de l'instrument de collecte ;
- L'établissement de la documentation sur les composantes du processus, y compris la documentation technique et les manuels d'utilisation ;
- Le renforcement des capacités des acteurs aux modalités de fonctionnement du processus.

3.4.2 Collecte

La collecte (4) englobe le tirage de l'échantillon, la mise en place du processus de collecte et la réalisation de la collecte.

- Le tirage de l'échantillon : s'il s'agit d'un recensement donc exhaustif, la question pourra concerner le fichier des établissements et notamment les modalités de sa mise à jour ;
- La mise en place du processus de collecte concerne la formation des personnels et la préparation des outils (impression) ;
- La réalisation de la collecte ;
- La finalisation de la collecte qui correspond à la saisie des données.

3.4.3 Traitement des données

Il s'agit des phases d'apurement et de préparation des données pour l'analyse. Cette étape comprend :

- L'intégration des données provenant d'une ou plusieurs sources, par exemple des données démographiques nécessaires au calcul des principaux indicateurs ou de données produites par d'autres directions techniques des ministères (DRH, Direction de l'évaluation, ...).
- La classification et le codage des données, par exemple des routines de codage automatique (ou manuel) peuvent attribuer des codes numériques selon une classification pré-déterminée (code établissement).
- L'examen et la validation en vue d'identifier les problèmes, erreurs et disparités potentiels comme les valeurs aberrantes, les cas de non-réponse et les erreurs de codage.
- L'édition et imputation des données jugées incorrectes, manquantes ou non fiables, qui sont remplacées par de nouvelles valeurs.
- Le calcul de nouvelles variables et unités statistiques qui ne sont pas explicitement prévues dans la collecte, mais qui sont nécessaires pour fournir les produits attendus (par exemple les indicateurs de scolarisation).
- Le calcul des pondérations (spécifique aux enquêtes par sondage).
- Le calcul des agrégats (création des données agrégées à partir des micro-données).
- La finalisation des fichiers de données (création du fichier de données utilisé dans la phase d'analyse).

3.4.4 Analyse des résultats

L'analyse des résultats regroupe la création des produits statistiques, leur validation et leur préparation en vue de leur diffusion.

Les activités de validation peuvent comprendre :

- La vérification de la couverture du recensement ;
- La comparaison avec les résultats des recensements précédents ;
- La confrontation des résultats avec d'autres données pertinentes (par exemple issues d'enquêtes) ;
- La recherche d'incohérences dans les statistiques ;
- La validation des statistiques en procédant par exemple à des échanges avec des experts internes.

3.4.5 Diffusion

Elle correspond à la mise à disposition des résultats aux utilisateurs, elle comprend entre autres l'élaboration des produits de diffusion.

4. INDICATEURS D'ÉVALUATION

La combinaison des dimensions de la qualité et des processus statistiques a permis d'identifier une liste d'indicateurs pour évaluer la qualité des productions statistiques. Ces indicateurs sont repris dans le tableau 1 page suivante. Une rubrique additionnelle relative aux moyens mis en œuvre a été ajoutée à ces indicateurs.

Il est à noter que la constitution d'une documentation technique figure parmi les activités qui ont été programmées pour le renforcement du SISEF. La question qui se posera sera donc la mise à jour de cette documentation suite à des changements méthodologiques, plutôt que son existence.

En soi le remplissage de ce tableau ne nécessitera pas un travail conséquent. Toutes les informations qui permettront de renseigner les indicateurs figurent dans la documentation technique du recensement. Il s'agira principalement des rapports d'exécution et de traitement du recensement. Le plan de ces rapports est présenté en annexes 2 et 3.

Tableau 1 : Liste des indicateurs, types et sources des informations

	Thématiques	Indicateurs	Sources des informations	Type d'informations	Niveau administratif	Dimensions de la qualité	Commentaires
Elaboration	Documentation du recensement	Modifications du questionnaire par rapport au recensement précédent	Rapport d'exécution	Narratif	Central	Rigueur méthodologique	Préciser les modifications opérées et leurs raisons
		Existence/mise à jour de la documentation technique : Manuel de formation Manuel de remplissage Manuel de contrôle Manuel de méthodologie		Narratif	Central		Les manuels ont-ils été mis à jour et lesquels
		Existence d'un calendrier de réalisation du recensement		Calendrier complet de l'opération	Central, régional		De la phase de préparation jusqu'à la diffusion
	Existence d'un fichier national des établissements	Narratif		Central	Description du fichier central (liste des variables)		
	Fichier des établissements scolaires	Fichier des établissements à jour (Par région et type d'établissement)		Narratif Tableau récapitulatif des créations et des établissements qui ne sont pas ou plus en activité par rapport à l'année précédente	Central, régional, inspection	Exactitude et fiabilité	Description du processus de mise à jour
Collecte	Couverture du recensement	Taux de retour des questionnaires (Par région et type d'établissement)	Rapport d'exécution	Tableaux récapitulatifs	Central, régional	Exactitude et fiabilité	Rapport entre le nombre de questionnaires distribués et le nombre de questionnaires restitués et complets
	Période de référence	% d'établissements ayant transmis leurs questionnaires dans les délais (Par région et type d'établissement)		Tableaux récapitulatifs	Central, régional, inspection		Rapport entre le nombre d'établissements qui ont transmis leurs questionnaires à la DREP dans les délais et le nombre d'établissements enquêtés.

	Thématiques	Indicateurs	Sources des informations	Type d'informations	Niveau administratif	Dimensions de la qualité	Commentaires
Traitement	Examen et validation des données	Nombre de questionnaires ayant nécessité des corrections par variable clé (effectifs, enseignants, salles de classe)	Rapport de traitement	Narratif Tableaux récapitulatifs des corrections pour chaque variable	Central, régional	Exactitude et fiabilité	Description du type de corrections effectuées sur chacune des variables. Tableaux des corrections par région et type d'établissement
	Edition et imputation des données incorrectes	Nombre de questionnaires ayant présenté des données manquantes par variable clé (effectifs, enseignants, salles de classe)		Narratif Tableaux récapitulatifs du nombre de questionnaires pour chaque variable	Central, régional		Description du type de données manquantes. Tableaux par région et type d'établissement
		Méthode d'imputation des données manquantes (variables clés)		Narratif	Central, régional		Décrire la méthode pour chaque type
		Complétude du rapport de traitement		Narratif	Central, régional		
Analyse	Validation des produits	Complétude du rapport d'exécution	Rapport d'exécution	Narratif	Central, régional, inspection		
		Concordance des données clés avec celles des années précédentes et raisons de non concordance	Rapport de traitement	Narratif Tableaux récapitulatifs des effectifs, enseignants, salles de classe sur 5 ans par région	Central, régional	Cohérence et comparabilité	Décrire les raisons de non concordance.
Diffusion	Diffusion de l'annuaire	Respect de la date de diffusion	Rapport d'exécution	Oui / Non	Central	Actualité et ponctualité	Si non respect expliciter les raisons
Moyens		Nombre de personnels engagés dans la collecte et fonction	Rapport d'exécution	Narratif	Central, régional		
		Nombre de personnels formés (et modalités de la formation)					Préciser les modalités de formation

Le format du calendrier de réalisation du recensement est indiqué en annexe 5 et un calendrier indicatif est produit en annexe 6.

5. PROCESSUS DE VALIDATION

5.1 Indicateurs de validation

Il est proposé que la remise d'un rapport de résultats soit assortie des documents permettant de décrire l'ensemble du processus et la fiabilité des données, il s'agira :

- Description processus :
 - Rapport de qualité ;
 - Rapport d'exécution ;
 - Rapport de traitement.
- Fiabilité des données

Les erreurs non dues à l'échantillonnage peuvent se produire à toutes les étapes du processus de production, dans tous les types d'enquêtes et de recensement, et elles sont difficiles à mesurer. Il s'agit des :

- Erreurs de couverture : une partie de la population à enquêter n'est pas couverte par l'enquête ;
- Erreurs de réponse : le directeur ne répond pas correctement à la question formulée ;
- Erreurs de non-réponse : le directeur ne répond pas à une ou plusieurs questions ;
- Erreurs de traitement : dans la saisie ou l'imputation inappropriée des données manquantes.

Le produit statistique sera validé si les conditions suivantes sont réunies :

- ✓ Rapport qualité complet avec tous les indicateurs renseignés ;
- ✓ Rapport d'exécution du recensement complet conformément au plan en annexe 2 ;
- ✓ Rapport de traitement complet conformément au plan en annexe 3 ;
- ✓ Taux de couverture minimum de 95% au niveau régional ;
- ✓ Taux de non-réponse inférieur à 10% au niveau régional pour chacune des variables clés : salles de classe, enseignants et effectifs avant imputations.
- ✓ Taux d'incohérences inférieur à 10% au niveau régional pour chacune des variables clés : salles de classe, enseignants et effectifs avant corrections.

L'INS appuiera les ministères du secteur éducation-formation pour l'élaboration de ces documents techniques.

5.2 Processus

La production du SISEF concernera à la fois les niveaux nationaux et régionaux.

- Niveau régional : le processus intégrera une vérification initiale des annuaires régionaux du MEP/APLN/EC et du MES avant la consolidation nationale. Cette vérification des données sera menée par un comité régional de contrôle qualité constitué de :
 - DREP : chargé des statistiques, directeur, équivalent du DEP ;
 - DR de l'INS, le statisticien/démographe en charge des statistiques de l'éducation ;
 - Un représentant du Gouvernement de la région ;
 - Un représentant du Conseil Régional ;
 - Directeur Régional de Développement Communautaire ;
 - Un représentant de la société civile en lien avec l'éducation.

Les données contrôlées seront validées au niveau national, ce qui implique que l'annuaire régional ne pourra être produit qu'après la validation de l'annuaire national. Le non-respect de cette condition pourrait déboucher sur l'existence de divergences entre l'annuaire régional et l'annuaire national.

- Niveau national : il n'est pas conseillé de créer un dispositif de validation parallèle à celui mis en place à l'INS, mais plutôt d'adapter le dispositif récemment adopté aux caractéristiques sectorielles. Il est proposé la création d'un sous-groupe SISEF au sein du comité technique qui sera en charge de valider la production du secteur.

Ce sous-groupe SISEF sera constitué de plusieurs membres parmi lesquels :

- Le sous-comité des statistiques sociales du comité de validation créé par l'INS ;
 - 3 cadres des ministères concernés, à savoir 2 cadres de la DS/DSI et 1 cadre de la DEP ;
 - Les directeurs des statistiques des autres ministères en charge de l'éducation et de la formation ;
 - 4 personnes ressources, dont 1 représentant des partenaires techniques et financiers, 1 représentant de la société civile en lien avec l'éducation, 1 représentant de la CAMOS, 1 représentant de la CAPEG.
- Délai : la validation des données interviendra dans un délai maximum de 2 semaines après la transmission de l'annuaire à l'INS.
 - A l'issue de la réunion un rapport de validation sera élaboré indiquant le résultat de la séance.

En conclusion, la validation reposera sur l'évaluation de la qualité mais comme le souligne Statistique Canada : "*Bien que le respect de bonnes pratiques en matière d'assurance de la qualité des données ne garantisse pas leur validité, il réduit la probabilité qu'il y ait des erreurs*".

6. ANNEXES

6.1 Annexe 1 : Liste des questions à double saisir

Recensements des établissements primaires

- Section 1 : Localisation et identification de l'établissement.
- Section IV : Caractéristiques et état des locaux.
- Section VIII : Nouveaux inscrits en 1^{ère} année d'études (CI) (tableau répartition des nouveaux inscrits).
- Section IX : Répartition des élèves par groupe pédagogique.
- Section X : Aire de recrutement de l'école.
- Section XI : Effectifs provenant des classes passerelles.

6.2 Annexe 2 : Contenu du rapport d'exécution

Objectif du rapport

D'une manière générale les modalités de réalisation du recensement seront exposées dans le manuel de méthodologie. Le rapport d'exécution précisera les modalités de réalisation du recensement actuel :

- Les éventuels changements apportés dans les outils et méthodes ;
- Les moyens mobilisés ;
- Les problèmes rencontrés et les solutions apportées ;
- Les propositions d'amélioration.

Ce rapport sera établi par chaque DREP et consolidé par la DS pour produire un rapport national.

Plan du rapport

I. Rappel de l'objectif du recensement

II. Méthodologie

- Champ de l'enquête
 - Mise à jour du fichier des établissements ;
 - Taux de couverture du recensement (ventilation par région et type d'établissement) ;
 - Conformité du taux de couverture avec les recensements précédents (raisons des potentiels écarts) ;
 - Tableau récapitulatif : nombre de questionnaires distribués, renseignés, saisis.
- Questionnaire
 - Eventuelles modifications par rapport au précédent recensement et raisons.
- Collecte
 - Formation des personnels (modalités) ;
 - Contrôle des questionnaires (liste des erreurs et difficultés rencontrées) ;
 - Durée de la collecte.
- Saisie
 - Organisation de la saisie (nombre d'agents, organisation de la supervision) ;
 - Application de saisie et contrôles ;
 - Principales erreurs et difficultés rencontrées ;
 - Durée de la saisie.
- Apurement des données
 - Résultats des requêtes d'apurement sur les variables principales (incohérences, données manquantes) ;
 - Imputation des données manquantes ;
 - Durée de l'apurement.

III. Problèmes rencontrés et propositions d'amélioration

- Problèmes (en général et éventuelles spécificités régionales) ;
- Propositions d'amélioration.

6.3 Annexe 3 : Contenu du rapport de traitement

Objectif du rapport

Le rapport de traitement fournira une description détaillée des phases d'apurement et d'analyse des données.

Plan du rapport

I. Organisation de l'apurement des données

- Moyens ;
- Durée.

II. Description des requêtes et les résultats

- Requêtes de contrôle et les résultats de ces requêtes ;
- Règles adoptées pour la correction des erreurs ;
- Méthodes d'imputation des données manquantes.

III. Indicateurs

- Définition des principaux indicateurs ;
- Méthode de calcul des principaux indicateurs.

IV. Propositions d'amélioration

- Collecte ;
- Saisie.

6.4 Annexe 4 : Contenu du rapport méthodologique

Objectif du rapport

Le rapport présentera une description complète des différentes étapes du recensement : de la préparation à la diffusion des résultats du recensement. Il devra être mis à jour après chaque changement méthodologique.

Plan du rapport

I. Introduction

- Présentation du document ;
- Objectifs du recensement.

II. Méthodologie

- Champ du recensement ;
- Mise à jour du fichier des établissements ;
- Taux de couverture habituel (couverture des établissements privés) ;
- Calendrier du recensement.

III. Outils de collecte

- Questionnaire (présentation, éventuels changements) ;
- Manuels (manuels de remplissage, manuel de contrôle des questionnaires).

IV. Collecte

- Dispositif de collecte ;
- Formation des personnels ;
- Contrôle des questionnaires ;
- Saisie des données (organisation, application de saisie, contrôle des données) ;
- Difficultés rencontrées ;
- Durée de la collecte.

IV. Traitement

- Apurement (organisation, requêtes) ;
- Imputation des données manquantes ;
- Indicateurs (définition, méthode de calcul).

V. Diffusion

- Accès aux données (calendrier, supports de diffusion).

VI. Annexes

- Questionnaire.

6.5 Annexe 5 : Format du calendrier du recensement

Elaboration

- Cette phase correspond à la préparation des outils de collecte, à leur reproduction, ainsi qu'à la mise à jour des établissements.

Collecte

- Formation des acteurs (directeurs et chargé des statistiques) ;
- Recueil des informations (dates de démarrage et de fin de collecte) ;
- Contrôle des questionnaires (dates de démarrage et de fin de contrôle) ;
- Saisie des questionnaires (dates de démarrage et de fin de saisie).

Traitement

- Apurement des données ;
- Consolidation des bases de données régionales ;
- Production de l'annuaire ;
- Validation des résultats.

Diffusion

- Diffusion de l'annuaire ;
- Diffusion des autres produits.

6.6 Annexe 6 : Calendrier indicatif

Elaboration

- Mise à jour du fichier des établissements, révision des outils de collecte : Octobre – Novembre ;
- Reproduction des questionnaires : Novembre – Décembre.

Collecte

- Formation des acteurs (directeurs et chargé des statistiques) : Décembre ;
- Collecte des données dans les écoles : Décembre ;
- Récupération des questionnaires et contrôle des questionnaires par les chargés de secteur : Décembre – Janvier ;
- Récupération des questionnaires et contrôle des questionnaires par les inspections : Janvier ;
- Saisie des questionnaires : Février – Mars.

Traitement

- Apurement des données : Mars – Avril ;
- Consolidation des bases de données régionales : Avril.
- Production de l'annuaire : Avril – Mai ;
- Validation des résultats : Juin.

Diffusion

- Diffusion de l'annuaire : 31 Juillet.

6.7 Annexe 7 : Fiche de contrôle des questionnaires

Cette fiche sera destinée au suivi des résultats du contrôle des questionnaires avant leur saisie. Elle fournira un indicateur de la qualité de remplissage du questionnaire en identifiant les incohérences ainsi que les données manquantes. Cette fiche de contrôle devra être renseignée par les chefs de secteur ou par les chargés des statistiques de l'inspection. Elle sera établie au niveau du secteur et sera agrégée par inspection.

Pour chaque questionnaire présentant des incohérences ou des données manquantes, il sera reporté leur nature selon la codification ci-dessous.

Cette codification a été établie sur la base des incohérences identifiées lors du contrôle des recensements des établissements d'enseignement primaire. Elle sera testée et adaptée aux spécificités des recensements des autres sous-composantes du secteur de l'éducation et de la formation.

- Salles de classe :
 - Type de local non ou partiellement renseigné (question 4.1) ;
 - Nature des murs non ou partiellement renseigné (question 4.7) ;
 - Etat des murs non ou partiellement renseigné (question 4.8) ;
 - Financement non ou partiellement renseigné (question 4.14) ;
 - Financement non ou partiellement renseigné (question 4.14) ;
 - Tableau récapitulatif non ou partiellement renseigné ;
 - Incohérences entre le détail des classes et le tableau récapitulatif ;
 - Incohérences dans le tableau récapitulatif (Taux de couverture du recensement (nombre de salles dans la partie répartition est différent du nombre de salles utilisées, ...)).
- Enseignants :
 - Personnel non ou partiellement renseigné (tableau VIII) ;
 - Etablissement sans enseignant (type d'emploi = administration ou appui) ;
 - Incohérences entre statut de l'établissement et statut de l'enseignant (fonctionnaire dans établissement privé ou communautaire, ...) ;
 - Incohérences entre type d'emploi et fonction (un directeur déchargé a une fonction d'enseignement administrative, un gardien une fonction d'appui ; ...)
 - Incohérences entre le dernier diplôme professionnel et la qualification ;
 - Incohérences entre la liste des enseignants et le tableau récapitulatif du personnel enseignant par statut ;
 - Incohérences entre la liste des enseignants et le tableau récapitulatif du personnel enseignant par qualification ;
 - Tableau récapitulatif du personnel enseignant par qualification et/ou par statut non ou pas renseigné...).
- Effectifs des élèves :
 - Nouveaux entrants non renseignés (tableau 8.1) ;
 - Groupe pédagogique non renseigné (tableau IX) ;
 - Aire de recrutement non renseignée (tableau X) ;
 - Erreur d'indexation dans le groupe pédagogique (tableau IX) ;
 - Nouveaux entrants non ou partiellement reportés dans le tableau IX groupes pédagogiques ;

- Incohérences entre les effectifs de nouveaux entrants t (tableau 8.1) et ceux reportés dans le tableau IX groupes pédagogiques ;
 - Incohérences entre le tableau IX répartition par groupe pédagogiques et le tableau XII structures pédagogiques et groupes d'élèves ;
 - Incohérences entre le tableau IX répartition par groupe pédagogiques et le tableau X aire de recrutement ;
 - Incohérences entre le tableau IX répartition par groupe pédagogiques et le tableau XI élèves venant des classes passerelles.
- Autres
- Autre relatif aux salles ;
 - Autre relatif aux enseignants ;
 - Autre relatif aux effectifs.

6.8 Annexe 8 : Fiche d'apurement des questionnaires

Cette fiche sera destinée au suivi des résultats de la saisie et de l'apurement. Elle fournira un indicateur de la qualité de la saisie en identifiant les questions renseignées qui n'auront pas été saisies ou qui auront été mal saisies. A ce stade certaines incohérences qui auraient échappées au contrôleur pourront être identifiées.

Elle sera établie au niveau de la DREP par le chargé des statistiques et le superviseur de la saisie.

Pour chaque questionnaire présentant des incohérences ou des données manquantes, il sera reporté leur nature selon une codification similaire à celle de la fiche de contrôle, à laquelle seront ajoutées la liste des questions clés non ou partiellement saisies.

Cette fiche a été établie sur la base des incohérences identifiées lors du contrôle des recensements des établissements d'enseignement primaire. Elle sera testée et adaptée aux spécificités des recensements des autres sous-composantes du secteur de l'éducation et de la formation.

6.9 Annexe 9 : Niveaux 1 et 2 du modèle statistique général du processus opérationnel

Gestion de la qualité / Gestion des métadonnées								
1 Précision des besoins	2 Conception	3 Mise en oeuvre	4 Collecte	5 Traitement	6 Analyse	7 Diffusion	8 Archivage	9 Évaluation
1.1 Détermination du besoin d'information	2.1 Résultats	3.1 Instrument de collecte des données	4.1 Sélection de l'échantillon	5.1 Normalisation et préservation de l'anonymat	6.1 Acquisition de renseignements au sujet du domaine	7.1 Mise à jour des résultats	8.1 Définition des règles d'archivage	9.1 Collecte des avis liés à l'évaluation
1.2 Consultation et confirmation	2.2 Base de sondage et méthodologie d'échantillonnage	3.2 Éléments du processus	4.2 Mise en place du processus de collecte	5.2 Intégration des données	6.2 Préparation des résultats provisoires	7.2 Production des résultats	8.2 Gestion du dépôt des archives	9.2 Préparation de l'évaluation
1.3 Établissement des objectifs du produit	2.3 Variables	3.3 Configuration du flux	4.3 Lancement du processus de collecte	5.3 Classification et codification	6.3 Vérification des résultats	7.3 Gestion de la diffusion des produits	8.3 Préservation des données et des métadonnées connexes	9.3 Établissement d'un plan d'action par consensus
1.4 Vérification de la disponibilité des données	2.4 Collecte de données	3.4 Mise à l'essai	4.4 Chargement des données dans l'environnement de traitement	5.4 Vérification et imputation des données	6.4 Interprétation et explication	7.4 Commercialisation et promotion des produits	8.4 Élimination des données et des métadonnées connexes	
1.5 Préparation de l'analyse de rentabilisation	2.5 Méthode de traitement statistique	3.5 Finalisation des systèmes de production		5.5 Dérivation de nouvelles variables	6.5 Contrôle de la divulgation	7.5 Gestion des demandes de la clientèle		
	2.6 Système de traitement et flux			5.6 Calcul de la pondération	6.6 Finalisation des résultats pour la diffusion			
				5.7 Calcul des agrégats				

6.10 Annexe 10 : Liste des documents consultés

- INS : "Guide d'application du cadre national d'assurance qualité des statistiques publiques", novembre 2019.
- INS : "Cadre national d'assurance qualité des statistiques publiques".
- Eurostat : "Code de bonnes pratiques de la statistique européenne", novembre 2017.
- Statistique Canada : Lignes directrices concernant la qualité, 5ème édition. Octobre 2009.
- OCDE : "Recommandations du Conseil de l'OCDE concernant les bonnes pratiques statistiques", 2015.
- UNESCO / UIS : Cadre d'évaluation de la qualité des données de l'éducation (CEQD-Ed) pour évaluer les systèmes de données administratives courantes : Manuel de conduite d'une évaluation par une équipe technique nationale.
- Nations Unies, Commission économique pour l'Europe : Modèle générique du processus de production statistique. Document établi par le Groupe de haut niveau sur la modernisation de la production et des services statistiques. Janvier 2014.