



Revue du Schéma directeur de la Statistique du Mali

Problématique de la qualité des données statistiques

La qualité : « l'ensemble des propriétés et caractéristiques d'un produit ou d'un service qui lui confèrent l'aptitude à satisfaire les besoins exprimés et implicites »

Norme ISO 8402 de 1986

Bamako (Mali), les 26 et 27 mai 2009



Plan de la présentation

- Introduction
- La notion de qualité en Statistique
- Références internationales
- Mise en œuvre de la recherche de qualité dans les activités statistiques
- Conclusion



Introduction

- La Statistique, outil de pilotage des politiques de développement, doit respecter la notion de qualité
- La réputation d'un Institut national de la Statistique repose sur sa capacité à produire des données de qualité
- Pour assurer cette qualité (Produits et services), les activités d'élaboration doivent respecter des normes internationales de qualité
- Le système doit permettre le contrôle de la qualité (contrôle interne ou contrôle externe)



I. La notion de qualité en Statistique

- Les organismes internationaux, retiennent les 6 dimensions suivantes pour définir la qualité des statistiques officielles :
 1. *pertinence*
 2. *Exactitude - précision*
 3. *Actualité et ponctualité*
 4. *Accessibilité et clarté*
 5. *Interprétabilité*
 6. *Cohérence*



I.1. La pertinence

- La **pertinence** de l'information statistique exprime comment l'information répond aux besoins réels des utilisateurs.
 - Est-ce que l'information disponible permet à l'utilisateur de mieux comprendre le phénomène étudié ?
- L'évaluation de la pertinence est subjective et est basée sur l'examen des besoins des utilisateurs.
- Les difficultés :
 - satisfaire les besoins actuels et potentiels des utilisateurs
 - élaborer un programme qui répond aux besoins dans un contexte de ressources limitées.



I.2. L'exactitude / la précision

- **L'exactitude** de l'information statistique exprime « la mesure dans laquelle l'information décrit correctement le phénomène qu'elle devrait évaluer. »
 - Avec quelle précision la mesure du phénomène est faite ?
- Elle est évaluée par l'erreur dans les estimations statistiques (erreur systématique et erreur aléatoire).
- Les erreurs majeures conduisant à l'inexactitude ont pour causes :
 - l'échantillonnage,
 - la mesure (méthodes et instruments de collecte)
 - Les non-réponses
 - le traitement.



I.3. L'actualité et la ponctualité

- L'**actualité** de l'information statistique tient compte de sa date de publication par rapport à sa période de référence
- La notion d'actualité est liée avec celles d'exactitude et de pertinence.
 - Est-ce que l'information fournie est actuelle ?
- L'évaluation de l'actualité est évidente et tient compte des normes de diffusion : écart entre la période de référence et la période de publication.



I.4. L'accessibilité et la clarté

- **L'accessibilité** de l'information statistique renvoie à la facilité avec laquelle on peut l'obtenir auprès du producteur.
 - Est-ce que l'information est mise à la disposition de tous les utilisateurs ?
- Plusieurs éléments pour évaluer l'accessibilité :
 - l'existence de l'information,
 - la forme ou le média d'accès à l'information,
 - le coût de l'information pour les utilisateurs.



I.5. L'interprétabilité

- **L'interprétabilité** de l'information statistique se caractérise par la disponibilité d'informations supplémentaires nécessaires à son interprétation.
 - Est-ce que l'utilisateur dispose d'informations supplémentaires pour mieux comprendre le phénomène étudié ?
- Ceci peut se vérifier par l'existence des métadonnées sur « les variables, les classifications et les concepts sous-jacents utilisés, la méthode de collecte des données et le traitement, etc. »



I.6. La cohérence

- La **cohérence** de l'information statistique est assurée lorsqu'« on peut réussir à recouper cette information avec d'autres renseignements statistiques dans un cadre analytique général »
 - Existe-t-il d'autres indicateurs qui permettent de confirmer par exemple la tendance observée ?
- La vérification se fait par l'existence d'informations annexes au sujet étudié (dans l'espace et dans le temps)
- Difficultés : utilisation de concepts normalisés, de méthodologie commune d'une enquête à l'autre.



I.7 Conditions préalables de la qualité

- Cadre juridique et institutionnel favorable
- Ressources adaptées aux besoins des programmes (humaines & financières)
- Il faut aussi admettre que :
 - les statistiques contiennent de l'information pertinente pour le domaine
 - La qualité est une condition qui gouverne tout travail de production de statistiques



II. Références internationales

- Plusieurs organismes font référence en la matière :
 1. Le FMI :
 - Le Système général de diffusion des données (SGDD)
 - La Norme spéciale de diffusion des données (NSDD)
 - Le Cadre d'évaluation de la qualité des données (CEQD)
 2. Paris21:
 - Le questionnaire sur les Indicateurs de renforcement des capacités statistiques (IRCS)
 3. Des instituts de statistique :
 - Eurostat, Statistique Canada, Insee, etc.



III. Mise en œuvre de la recherche de qualité dans les activités statistiques

- La qualité sera recherchée à plusieurs niveaux :
 - Système de production :
 - Aspects institutionnels
 - Ressources humaines
 - Produits :
 - Outils
 - Méthodologie
 - Opérations



III.1. Système de production : aspects institutionnels (suite)

- L'élaboration du SDS doit intégrer la démarche « qualité » afin de faire du système statistique malien, notamment l'INSTAT un centre d'excellence de production des statistiques.
- Pour cela, il faut :
 - Fournir une gamme élargie et complète d'informations statistiques pour le développement du pays ;
 - Élargir l'utilisation des technologies de l'information dans les travaux du SSN ;
 - Améliorer le cadre et les conditions de vie du personnel du SSN.



III.1. Système de production : aspects institutionnels (fin)

- Pour atteindre ces objectifs, il devra mis en place :
 - un cadre institutionnel favorable :
 - Conseil national de la statistique
 - Comité technique de suivi
 - Comités sectoriels, Groupe des PTF
 - une législation fonctionnelle
 - Les lois réglementaires (collecte, exploitation et diffusion des données)
 - Les protocoles d'accord pour l'échange des données entre les administrations (sources administratives et données sectorielles)



III.2. Evaluation du cadre institutionnel

- Analyse de la réglementation mise en place :
- Il s'agira de s'assurer que les lois existantes définissent clairement :
 - La responsabilité en matière d'élaboration et de diffusion des statistiques
 - L'accès aux données des administrations
 - La collecte et la confidentialité des données individuelles



III.2. Evaluation du cadre institutionnel (suite)

- Analyse de la coordination du système :
 - Responsabilité en matière de coordination
 - Existence d'une programmation concertée d'activités statistiques
 - Fréquence des réunions de coordination du SSN
 - Promotion des cadres de référence (concepts, nomenclatures, méthodologies, etc.)



III.3. Evaluation du cadre institutionnel (suite)

- Analyse des ressources humaines :
 - Adéquation du nombre de ressources humaines aux activités statistiques programmées
 - Niveau de rémunération du personnel
 - Locaux, équipements et installations



III.3. Evaluation du cadre institutionnel (fin)

- Analyse de la programmation et suivi et évaluation :
 - Plan stratégique à moyen terme ;
 - Programmes de travail annuels ;
 - Rapports d'activités annuels ;
 - Suivi des dépenses et des recettes ;
 - Evaluation des financements extérieurs.



III.4. Analyse de la qualité des données statistiques

- Utilisation des nomenclatures dérivées des normes internationales :
 - NCOA/ NCAC → COICOP (10 ou 12 fonctions)
 - NAEMA → CITI Rev3
 - NOPEMA → CPC
 - Comptabilité nationale → SCN68, 93, 2008
 - Etc.

Concepts et indicateurs normés garantissant la comparabilité des données dans le temps et dans l'espace



III.4. Analyse de la qualité des données statistiques (suite)

- L'univers étudié : bien délimiter le champ de la statistique :
 - Échantillonnage explicite
 - Plan de sondage compréhensible
 - Pondérations spatiales disponibles
 - Limites du champ géographique
 - Taux de couverture



III.4. Analyse de la qualité des données statistiques (suite)

- Collecte des données : bien identifier la source des données :
 - Unités statistiques étudiées
 - Unité de collecte
 - Niveau d'observations
 - Thèmes / sections d'observations
 - Distinguer niveaux d'observations et niveaux d'analyse



III.4. Analyse de la qualité des données statistiques (suite)

- Collecte des données d'enquête :
 - Manuel de l'enquêteur
 - Manuel du contrôleur
 - Calendrier de collecte
 - Rapports d'enquête (remplacements, d'enquêtés, taux de non réponse, résolution des problèmes rencontrés, etc.)



III.4. Analyse de la qualité des données statistiques (suite)

- Saisie et traitement des données :
 - Programme de saisie contrôle
 - Logiciels informatiques utilisés accessibles
 - Manuel d'instructions aux agents de saisie
 - Organisation de l'atelier de saisie
 - Audit de la saisie (exhaustivités aux niveaux individus statistiques et des variables, etc.)
 - Personnel formés à la saisie et à la correction



III.4. Analyse de la qualité des données statistiques (suite)

- Collecte et utilisation des données administratives :
 - Protocole d'accord pour l'accès aux données
 - Formats standards des fichiers
 - Logiciels utilisés disponibles
 - Procédures d'évaluation des données
 - Procédures de confrontation des résultats



III.4. Analyse de la qualité des données statistiques (suite)

- Traitement des données :
 - Apurement des données d'enquêtes (programmes informatiques détaillés)
 - Traitement des non réponses (imputations et ajustements, etc.)
 - Automatisation des traitements des données d'enquêtes
 - Manuels de traitement disponibles



III.4. Analyse de la qualité des données statistiques (suite)

- Diffusion des résultats :
 - Respect du calendrier de diffusion des données
 - Résultats disponibles sous diverses formes
 - Résultats selon les niveaux de la nomenclature
 - Niveaux de signification des résultats
 - Précision des indicateurs calculés
 - Niveaux d'analyse précisés
 - Possibilités d'analyses complémentaires



III.4. Analyse de la qualité des données statistiques (fin)

- Diffusion des résultats (suite) :
 - Possibilité de production de résultats particuliers
 - Base de données consultables
 - Résultats consultables à travers le site Internet du producteur
 - Archivage des données sources 'anonymisées'



Conclusion

- La démarche qualité doit devenir une partie intégrante des activités statistiques
- Des contrôles internes et externes doivent se faire à toutes les étapes du processus de programmation et de production statistique
- L'évaluation de la qualité reste subjective
- Des tentatives existent pour l'évaluation de l'efficacité des SSN



Je vous remercie