

La politique de correction des erreurs au DSED

1. Définitions

1.1 Définitions et contours de l'erreur

Une erreur est un résultat incorrect, inattendu, découvert après publication, lié à un défaut des données sources collectées, à un écart au processus de production statistique prévu ou à une anomalie dans la diffusion.

Elle peut porter sur les données elles-mêmes (y compris celles reportées dans un commentaire, un tableau, un graphique, un schéma ou une infographie) mais aussi sur le texte (présentation, explication, analyse) ou les métadonnées les accompagnant (toute indication méthodologique : définition, champ, source, hypothèse, méthode).

Tous les supports de diffusion peuvent être concernés par une erreur : publication papier ou électronique (y compris base de données et visualisation des données), page du site Internet, publication sur les réseaux sociaux, communiqué de presse, de même que les codes en open source.

Le champ d'application retenu ici n'inclut pas :

- **La révision.** Celle-ci induit, comme une erreur, la modification de données/informations déjà publiées mais s'en distingue par :
 - sa planification (une révision est prévue, annoncée à l'avance¹ ; une erreur est inattendue) ;
 - son origine (une révision est due à la prise en compte d'une version améliorée de données sources² ou à une modification méthodologique³ alors qu'une erreur résulte d'un « accident » dans le processus de production ou de diffusion).
- **L'incertitude statistique** (erreur aléatoire - liée au fait de sonder une partie de la population étudiée seulement) inhérente à tout processus de production statistique, ainsi que les biais de mesure (erreur de couverture, erreur due à la non-réponse, etc.) ne relèvent pas de la présente politique.

En résumé, **une erreur est un phénomène qui se corrige en appliquant le processus de production prévu ou en l'adaptant à la marge.**

1.2 Erreurs mineures et erreurs majeures

Une erreur est considérée comme :

- **majeure** si elle modifie de façon non triviale et potentiellement spéculative l'interprétation des données ou informations diffusées et présente des conséquences (estimées) significatives pour les utilisateurs ;
- **mineure** dans les autres cas.

Les conséquences de l'erreur ne peuvent pas être parfaitement évaluées ex-ante. Elles sont approchées par la combinaison :

- du type d'indicateur touché et de sa sensibilité dans l'analyse économique ou sociale ;
- des usages connus des données concernées et des conséquences anticipées de l'erreur qui en découlent ;
- de la visibilité des données concernées (comme proxy de leur degré de diffusion) ;
- de l'ancienneté de l'erreur au moment de sa découverte.

¹ Sauf dans le cas particulier d'une révision non planifiée due à des changements inattendus dans les données sources

² Cela se produit dans les cas de disponibilité de sources plus complètes ou de meilleure qualité.

³ Une révision méthodologique est liée à des changements de réglementation, de concepts, de normes, de méthodes de calcul ou d'estimation, de sources, de classification...

Le type de l'erreur (mineure/majeure) détermine le traitement de l'erreur ainsi que la communication sur l'erreur et sa correction (cf. infra).

2. La correction rapide est la règle

Conformément au code européen de bonnes pratiques, les erreurs sont corrigées dans les meilleurs délais, si possible simultanément sur l'ensemble des supports concernés (cf. 1.1).

Du point de vue des utilisateurs, les corrections présentent un gain, mais peuvent également présenter un coût (coût d'intégration des données, de reprise de leurs traitements et analyses). Les avantages et inconvénients doivent être mis en balance, de façon à limiter le nombre et la fréquence de rediffusion de données et informations déjà publiées.

Ainsi, deux principes sont retenus :

- **une erreur majeure fait l'objet d'une nouvelle diffusion dans les meilleurs délais ;**
- une erreur mineure peut être corrigée sans délai si la balance avantages/inconvénients est bénéfique pour l'utilisateur. Dans le cas contraire, un certain délai peut être envisagé pour intégrer la correction, notamment si celle-ci peut intervenir à l'occasion d'une nouvelle diffusion de données (dans le cas de données diffusées fréquemment) ou si la charge de production courante nécessite de différer la prise en charge de la correction, de façon à minimiser le risque de retard dans la mise à disposition des estimations suivantes.

3. Information aux utilisateurs

1.3 En amont de la correction

Lorsque la correction nécessite un délai, les erreurs majeures sont signalées, dès que possible, en amont de leur correction :

- A un endroit bien visible sur les supports concernés et à tout autre emplacement susceptible d'attirer l'attention du public ;
- Par un message ad hoc destiné aux utilisateurs connus des données (notamment des abonnés dans le cas d'une erreur sur publication), transmis par les canaux habituels de diffusion.

Le signalement de l'erreur comporte :

- La description précise de l'erreur ;
- La date à laquelle la correction sera effectuée (a minima un délai maximal) ;
- L'origine de l'erreur (si elle a pu être identifiée) ;
- Sin impact potentiel (s'il a pu être évalué) : ordre de grandeur, conséquence sur le message transmis, etc...
- Les supports éventuels donnant une description plus précise de l'erreur.

Dans le cas où le risque de confusion pour l'utilisateur est non négligeable malgré ces signalements, les informations incorrectes peuvent être retirées du site Internet en attendant leur correction.

Les erreurs mineures sont corrigées sans signalement en amont de la correction.

1.4 Au moment de la correction

Lors de la correction d'une erreur mineure ou majeure :

- Les informations erronées sont remplacées par leur version corrigée et sont définitivement supprimées ;
- La version corrigée est accompagnée d'une mention de correction ;
- Dans le cas d'une erreur majeure, les utilisateurs connus en sont informés.

La mention de la correction suit les trois principes suivants :

- Indication de la date à laquelle la correction a été effectuée
- Ampleur de la correction effectuée
- Pour les erreurs majeures, explication de l'erreur (origine) et de la correction effectuée.

Pour les erreurs majeures, la mention de la correction est portée à un endroit bien visible sur le support (par exemple, avertissement sur la page de présentation pour une publication électronique) et éventuellement à tout autre emplacement susceptible d'attirer l'attention.

Elle est, dans la mesure du possible, pérenne. C'est le cas, en particulier, pour les publications.

Cas particulier des publications papier :

Les erreurs mineures de forme sont corrigées dans les nouveaux exemplaires tirés uniquement.

Les autres erreurs font l'objet :

- dans les nouveaux exemplaires tirés : d'une mention de la correction effectuée à l'emplacement de l'erreur ;
- dans les exemplaires déjà tirés non encore distribués : d'un erratum. Les publications papier ne font pas l'objet de retirages

1.5 Après la correction

Toutes les erreurs majeures, ainsi que leur résolution, sont documentées (nature, origine, correction effectuée) et archivées de façon centralisée.

Le recensement des erreurs majeures vise à :

- améliorer en continu les données sources, les processus et la qualité des résultats obtenus ;
- renseigner les utilisateurs sur les corrections effectuées, et plus généralement sur la qualité des statistiques.

L'ensemble des erreurs majeures répertoriées dans la production du service statistique public (Insee et services statistiques ministériels) fait chaque année l'objet d'une analyse présentée au Comité stratégique de la qualité du service statistique public, qui peut recommander toute disposition visant à améliorer la qualité des productions.

Les erreurs mineures ne font pas l'objet d'une documentation.