

Politique de correction d'erreur

1. Définitions

1.1. Définition et contours de l'erreur

Une erreur est un résultat incorrect¹, inattendu, découvert après publication, résultant d'un défaut des données sources collectées, d'un écart au processus de production statistique prévu ou d'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'avance² ; une erreur est inattendue) ;
 - son origine (une révision est due à la prise en compte d'une version améliorée de données sources ou à une modification méthodologique³ ; 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 C'est-à-dire ne répondant pas aux critères de fiabilité prévalant à la diffusion de la donnée.

² Sauf dans le cas particulier d'une révision non planifiée due à des changements inattendus dans les données sources.

³ Une révision « de routine » est due à la disponibilité de données plus complètes ou de meilleure qualité et une révision « méthodologique » est due à des changements de législation, de concepts, de normes, de méthodes, de sources, de classification, d'année de base.

1.2. Erreurs mineures et erreurs majeures

Une erreur est considérée comme :

- **majeure** si elle modifie de façon non triviale⁴ et potentiellement spécieuse 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.

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 des erreurs 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. À condition que l'écart entre la valeur diffusée d'un indicateur et sa valeur correctement estimée soit non significatif, une erreur mineure peut exceptionnellement ne pas être corrigée.

3. Information des utilisateurs

3.1. En amont de la correction

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

⁴ Ce n'est pas le cas si la correction de cette erreur, même importante, est évidente pour le lecteur, qu'elle porte sur un élément de forme (typographie, erreur grammaticale, etc.) ou sur un élément de contenu (données, commentaires, etc.).

- à un endroit bien visible sur les supports concernés et, dans certains cas, à tout autre emplacement susceptible d'attirer l'attention du public ;
- par un message *ad hoc* destiné aux utilisateurs connus des données (notamment aux abonnés dans le cas d'une erreur sur une publication et dans certains cas aux médias), 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 (ou *a minima* un délai maximal) ;
- l'origine de l'erreur (lorsqu'elle a pu être identifiée) ;
- son impact potentiel (lorsqu'il a pu être évalué) : ordre de grandeur, conséquence sur le message transmis, etc. ;
- les supports éventuels donnant des informations plus précises sur l'erreur.

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

Les erreurs mineures ne sont pas signalées en amont de leur correction.

3.2. *Au moment de la correction*

Lors de la correction d'une erreur, qu'elle soit majeure ou mineure :

- les informations erronées sont remplacées par leur version corrigée et sont définitivement supprimées sauf suivi particulier (indice des prix à la consommation, etc.) ;
- la version corrigée est accompagnée d'une **mention de la correction**, sauf dans le cas d'erreurs de forme ;
- par ailleurs, dans le cas de la correction des erreurs majeures uniquement, **les utilisateurs connus (tels que décrits au 3.1) en sont informés**, selon les modalités décrites au 3.1.

La mention de la correction suit les trois principes suivants :

Elle indique :

- la date à laquelle la correction a été effectuée ;
- l'ampleur de la correction ;
- pour les erreurs majeures, l'explication de l'erreur (origine) et de la correction effectuée.

Elle est **placée à l'endroit de l'erreur ou à proximité immédiate** selon les possibilités (bases de données notamment).

Pour les erreurs majeures, elle est en outre 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.

3.3. 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 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 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 centralisée.