

Error correction policy

1. Définitions

1.1. Definition and contours of error

An error is an incorrect², unexpected result, discovered after publication, resulting from a defect in the source data collected, a deviation from the planned statistical production process or an anomaly in dissemination.

It may relate to the data themselves (including those reported in a commentary, table, graph, diagram or infographic), but also to the text (presentation, explanation, analysis) or the accompanying metadata (any methodological indication: definition, scope, source, hypothesis, method).

All distribution media can be affected by an error: paper or electronic publication (including database and data visualization), webservice, website page or blog post, social network publication, press release, as well as detailed files (made available internally or intended for researchers), data delivered to Eurostat and *open source* code.

This scope does not include :

- Revision. Like an error, this involves modifying data/information that has already been published, but differs in the following ways:
 - planning (a revision is planned, announced in advance³ ; an error is unexpected);
 - origin (a revision is due to the inclusion of an improved version of source data, or to a change in methodology⁴ ; an error results from an “accident” in the production or distribution process).
- Statistical uncertainty (random error - linked to the fact of surveying only part of the population studied) inherent in any statistical production process, as well as measurement bias (coverage error, non-response error, etc.) are not covered by the present policy.

In short, **an error** is a phenomenon that **can be corrected by applying the planned production process or making minor adaptations**.

² In other words, they do not meet the reliability criteria required for data distribution.

³ Except in the special case of unplanned revisions due to unexpected changes in source data.

⁴ A “routine” revision is due to the availability of more complete or better quality data, while a “methodological” revision is due to changes in legislation, concepts, standards, methods, sources, classification or base year.

1.2. Minor and major errors

An error is considered :

- **major** if it modifies in a non-trivial⁵ and potentially specious way the interpretation of the data or information disseminated and presents (estimated) significant consequences for users;
- **minor** in other cases.

The consequences of the error cannot be perfectly evaluated ex-ante. They are approximated by combining :

- the type of indicator affected and its sensitivity in economic or social analysis;
- the known uses concerned data and the anticipated consequences of the error;
- the visibility of concerned data (as a proxy for their degree of dissemination);
- the age of the error at the time of discovery.

The type of error (minor/major) determines how the error is handled, and how the error and its correction are communicated (see below).

2. Rapid error correction is the rule

In accordance with the European code of best practice, **errors are corrected as quickly as possible**, if possible simultaneously on all the media concerned (cf. 1.1).

From the user's point of view, corrections represent a gain, but they can also be costly (the cost of integrating data, and of reprocessing and analyzing it). The advantages and disadvantages must be weighed up, so as to limit the number and frequency of redistributions of data and information already published.

Two principles have been adopted:

- a **major** error involves a **new release as soon as possible**;
- a minor error can be corrected without delay if the balance of advantages and disadvantages is beneficial to the user. If this is not the case, a certain amount of time may be allowed for the correction to be incorporated, particularly if the correction can be made on the occasion of a new data release (in the case of data released frequently), or if the current production workload requires the correction to be deferred, so as to minimize the risk of delay in making subsequent estimates available. Provided that the difference between the published value of an indicator and its correctly estimated value is insignificant, a minor error may exceptionally not be corrected.

3. User information

3.1. Prior to correction

When their correction requires a delay, major errors are reported as soon as possible, before they are corrected :

- in a conspicuous place on the media concerned and, in some cases, in any other location likely to attract the public's attention;
- an *ad hoc* message aimed at known users of the data (notably subscribers in the case of an error on a publication, and in some cases the media), transmitted via the usual distribution channels.

⁵ This is not the case if the correction of this error, even a significant one, is obvious to the reader, whether it concerns an element of form (typography, grammatical error, etc.) or an element of content (data, comments, etc.).

The error report includes :

- a precise description of the error ;
- the date by which the error will be corrected (or at least a maximum deadline);
- the origin of the error (when identified);
- its potential impact (where this can be assessed): rough estimate, effect on the message transmitted, etc. ;
- any media providing more precise information on the error.

If the **risk of confusion** for the user remains significant despite these warnings, **incorrect information may also be removed** from the website until it has been corrected.

Minor errors are not reported before they are corrected.

3.2. At the time of correction

When an error is corrected, whether major or minor :

- the erroneous information is replaced by the corrected version and is permanently deleted, except in the case of special follow-up (consumer price index, etc.);
- the corrected version is accompanied by an **acknowledgement of the correction**, except in the case of formal errors;
- in the case of correction of major errors only, **known users (as described in 3.1) are informed**, as described in 3.1.

Mention of the correction is based on the following three principles:

① Are indicated

- the date on which the correction was made;
- the extent of the correction;
- for major errors, an explanation of the error (origin) and the correction made.

② It is **placed at the location of the error or close to it**, depending on possibilities (e.g. databases). In the case of major errors, it is also placed in a clearly visible position on the medium (e.g. warning on the presentation page for an electronic publication), and possibly in any other location likely to attract attention.

③ As far as possible, it is permanent. This is particularly true for publications.

Special case of paper publications :

Minor errors of form are corrected in new copies only.

Other errors are :

- in new copies: a note of the correction made at the location of the error;
- in copies already printed but not yet distributed: an erratum. Paper publications are not reprinted.

3.3. After correction

All major errors and their resolution are documented (nature, origin, correction made) and archived centrally.

The aim of identifying errors is to :

- continuously improve source data, processes and the quality of results obtained ;
- inform users about corrections made, and more generally about the quality of statistics.

Every year, all errors recorded in the output of the official statistical service (INSEE and ministerial statistical services) are analyzed and presented to the Strategic Committee for the Quality of the Official Statistical Service, which may recommend any measure aimed at improving the quality of output.

Minor errors are not documented centrally.