

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

UNECE Work Session on Statistical Dissemination and Communication
(12-14 September 2006, Washington D.C., United States of America)

Topic (iv) Managing revisions and version control to maintain credibility

**INFORMING THE USER ABOUT CORRECTIONS:
THE POLICY OF STATISTICS NETHERLANDS**

Supporting Paper

Submitted by Statistics Netherlands¹

I. INTRODUCTION AND SUMMARY

1. Statistics is about information. However, informing the users on scheduled and unscheduled revisions in official statistics is less common than it should be. National statistical institutes have a long tradition of treating their results as the fixed outcome of the statistical process. But in fact revisions in official statistics are not uncommon in many countries. Some revisions are scheduled and others are not. Scheduled revisions are based on a pre-announced release policy consisting at least of provisional results and final results. Unscheduled revisions or corrections are the results of errors within the statistical process or, for economic statistics, business data which are unexpectedly received too late to be included in the first estimates.

2. In 2005, Statistics Netherlands started a project to improve the information about corrections in statistics on our website. The starting point is that statistical institutes should promote transparency on corrections. We started a new heading on our website www.cbs.nl which was specifically devoted to corrections. Our website contains about 1500 multi-dimensional tables, comprising all core statistics of Statistics Netherlands. Every month, there are about 200 regular updates for a new time period and about 25 corrections on previously published results. Every correction is shown separately. For internal purposes, a management summary of all corrections is compiled, showing the number of corrections per organisational unit. This paper gives an overview of our policy, including a few statistics on the number of corrections.

II. REVISIONS AND CORRECTIONS

3. Until 2004, Statistics Netherlands made a distinction in its statistical output between two types of changes in statistics:

- *New* statistics. A completely new statistical table.
- *Revised* statistics. All other changes in statistical output, i.e. including a new month or year, a scheduled revision or a correction after a mistake has been made in the statistical process.

¹ Prepared by Huib van de Stadt, Head of publications and communication, hsdt@cbs.nl.

Both internally and externally, this created some problems. For our users it was unclear why statistics were changed. Was this simply the result of a scheduled statistical process or was it the result of mistakes and errors made by Statistics Netherlands? In January 2004, we received some unfavourable press coverage after unemployment statistics for a certain period were changed the next month. It was therefore decided to improve the information about the nature of the changes on our website.

4. In April 2004, a new classification of changes in statistical output was made and approved by our Board of Directors. Four types of changes were distinguished:

- *New tables*. A completely new table, or a table with new variables or new classifications of existing variables (renewed table).
- *Updated tables*. A new time period is added (a month, a quarter or a year), but nothing else is changed.
- *Scheduled revisions* of tables. A revision of statistical data based on a release policy. A release policy is a pre-announced publication schedule of provisional and final data.
- *Corrected tables*. A revision of statistical data *not* based on a release policy.

5. Hence, in the new classification a clear distinction was made between revisions based on a release policy and revisions not based on a release policy. Corrections are revisions not based on a release policy. Usually they are the result of new information which becomes available unexpectedly: a new statement of a company, or a new version of a register. In other cases an error has been made in the statistical production process: errors in the software, wrong versions of the database, etcetera. As we all know from experience, these kinds of errors cannot be avoided entirely.

6. In the course of 2004, the new classification was introduced. All official output of Statistics Netherlands is placed on our website www.cbs.nl. For every new table to be placed on the website, the statistician has to classify the change into one or more of the four categories mentioned above. The publication department checks the table and the classification and places the table on the website.

7. As from 2005, every correction introduced includes a short explanation provided by the statistician involved to improve transparency further. This explanation is also placed on our website. So we have a list of explanations of corrections which is available for all users of our statistics. In order to keep this list within a reasonable length, explanations are deleted after one month. The following link shows the list of corrections: www.cbs.nl/nl-NL/menu/cijfers/statline/correcties/default.htm. Unfortunately, this part of our website is not translated into English. As an example, we have included the list of July 2006 below.

Centraal Bureau voor de Statistiek

Home | Thema's | **Cijfers** | Publicaties | Methoden | Informatie voor | Over het CBS | English | Sitemap

Home > Cijfers > StatLine databank > Correcties

StatLine correcties

Deze pagina bevat informatie over correcties op in StatLine gepubliceerde cijfers. N.B. reguliere bijstellingen van cijfers in het kader van een vast publicatieschema worden niet op deze pagina vermeld. Deze reguliere bijstellingen betreffen met name het vervangen van voorlopige cijfers door definitieve. Voorlopige cijfers zijn herkenbaar in StatLine doordat ze zijn gemarkeerd met een "**".

25-07-2006	Emissies naar lucht
25-07-2006	Emissies naar lucht; mobiele bronnen
25-07-2006	Emissies naar lucht; wegverkeer
25-07-2006	Gemeenterekeningen, balans 2004
25-07-2006	IPCC-emissies naar lucht
25-07-2006	NEC-emissies naar lucht
19-07-2006	Maandcijfers binnenlands spoorvervoer
19-07-2006	Maandcijfers internationaal spoorvervoer
18-07-2006	Detailgegevens
18-07-2006	VO; leerlingen en geslaagden
15-07-2006	Vestigingen
15-07-2006	Demografie van bedrijven
14-07-2006	Kerncijfers transport
14-07-2006	Kerncijfers Bedrijfsleven
14-07-2006	Milieudienstverlening
13-07-2006	EMU-saldo en EMU-schuld; Internationaal
07-07-2006	Energie en waterleiding met honderd of meer werknemers
07-07-2006	Energie en waterleiding, alle bedrijven
07-07-2006	Kerncijfers Bedrijfsleven
07-07-2006	Kerncijfers dienstverlening
07-07-2006	Kerncijfers groothandel en handelsbemiddeling
07-07-2006	Kerncijfers Sociale Zekerheid
07-07-2006	Omzetspecificaties dienstverlening
06-07-2006	Welzijnswerk en kindercentra
04-07-2006	Verpleeghuizen

© Centraal Bureau voor de Statistiek, Voorburg/Heerlen, 2006 | Copyright | Privacyverklaring | Disclaimer | Help | Contact

8. We think that our treatment of corrections is meaningful and in line with the views of international organisations. The concept of scheduled revisions is also used in the Statistical Data Description System (SDDS) of the IMF and in the RAMON-system of Eurostat. The recent OECD newsletter of July 2006 also stresses the importance of explicitly showing the magnitude of revisions. The eDamis-system of Eurostat is used for nearly all data flows from Statistics Netherlands to Eurostat. It has a provision for the replacement of data flows in the case of errors and revisions, but a systematic distinction between several types of replacements is not made.

III. STATISTICS ON STATISTICS

9. Statistics Netherlands has one central on-line database for all its statistical output. The database is called StatLine and is accessible to the public through our website www.cbs.nl. One central database has several advantages. A maybe somewhat unexpected advantage is that it makes it much easier to compile management information on the production of statistics, and also on the number of corrections and revisions. Statistics on statistics, so to speak.

10. Table 1 summarizes the content of and the changes in the database StatLine for the first half of 2006.

Table 1. StatLine, January-June, 2006

	Business statistics	Social and regional statistics	Macro-economic statistics	Total Statistics Netherlands
Number of StatLine tables, December 31, 2005	563	704	255	1522
Number of StatLine tables, June 30, 2006	608	735	256	1599
New tables (including renewed)	38	59	43	140
Updates	541	407	406	1354
Scheduled revisions	40	16	29	85
Corrections	83	56	18	157

11. Table 1 shows that at the end of 2005, the official output of Statistics Netherlands consisted of 1,522 multi-dimensional tables in StatLine. During the first half of 2006, this number increased to 1,599. We had 140 new tables in this period, including about 60 tables which were only renewed (new variables or classifications in existing tables). Table 1 further shows that 1,354 tables were updated adding a new month or year (a monthly table usually counts for six, because the table covers six months). We had 85 scheduled revisions and 157 corrections.

12. The number of corrections is not negligible. In 157 cases in six months we had to make a correction, because something had gone wrong. As mentioned before, part of these corrections was made because somewhere else things went wrong (new information from companies or holders of registers) and another part is because things went wrong within the production process of Statistics Netherlands. For the user, this distinction is not very important, because the fact that there is a correction is the only thing that matters. For us, the distinction seems more important. However, one might also argue that a sound production process takes into account the fact that information is becoming available all the time.....

13. To conclude this section, we also give some statistics on corrections in the data flows to Eurostat by the eDamis system. In the first half of this year, 489 data sets were sent to Eurostat. From these data sets, 314 were new and 175 were replacements of previous data flows. These replacements include both large corrections (in data) and small corrections (changes in texts). Still, the number of replacements is fairly large.

IV. CONCLUSIONS

14. Making errors is human. The statistical production process is no exception in this respect. In the Netherlands, we have extended the public information on errors and corrections in statistics. This policy has started in 2004 and has now been completed. All corrections are shown and explained on our website.