UNECE Experiences in Statistical Dissemination

UNECE Statistical Division [1]

I.   Introduction

1.   The United Nations Economic Commission for Europe (UNECE) started disseminating statistics to the public online via an interactive database in February 2006. Like the National Bureau of Statistics of Moldova, the UNECE decided to adopt the PC-Axis suite of software tools as the basis for statistical data dissemination. PC-Axis was chosen for two main reasons. Firstly, the software is developed in a collabora tive, non-commercial environment, and secondly it is used successfully by a number of other statistical organisations.

2.   This paper describes the UNECE experiences since that date, as well as ideas for further developing data dissemination systems and pr inciples over the next few years.

II.   A General Look at the UNECE Statistical Database

3.   The UNECE Statistical Database covers a wide range of data for UNECE member countries and can be accessed by the public via both English and Russian interfaces at http://w3.unece.org/pxweb/Dialog/.

4.   The data are clustered according to five main categories:

- Country Overviews – This section brings together all of the key indicators from the rest of the database. It presents them in a way that facilitates cross -country comparisons. It was introduced in response to user demand for a simple overview of key data. Although it is composed entirely of existing content, it has quickly become the most popular dataset. This shows that simply re-organising existing content can greatly improve usability.

- Economic Statistics – This was the first statistical domain p ublished online. It provides a structured set of economic indicators for countries of the UNECE region, focusing on National Accounts, employment, price indices, ind ustrial production and external economic relations.

- Forestry Statistics – This is the newest part of the UNECE database, and the result of cooperation between the UNECE and United Nations Food and Agriculture Organization (FAO). It includes indicators on f orestry resources, forestry health and vitality, biological diversity, productive and socioeconomic functions etc.

- Gender Statistics – A key priority globally for the United Nations is the advancement of gender equality. To support this key policy aim, the UNECE has developed gender breakdowns for social and demographic data. This part of the database has seen a lot of development during recent years, with the addition of data on computer use, work - life balance and the gender pay gap.

- Transport Statistics – This part of the database contains data collected by statisticians in the UNECE Transport Division. It includes information on road safety, transport by

---

[1] Paper prepared by: Lidia Bratanova, Marlen Jigitekov and Steven Vale for the seminar on dissemination of statistics and launching a new web solution in the National Bureau of Statistics of Moldova, Chisinau, 27 May 2010.

road, rail, inland waterways and oil pipelines, transport infrastructures, and railway employment.

5.　　In accordance with the Fundamental Principles of Official Statistics[2], access to the UNECE statistical database is free of charge. Emphasis has been placed on making data ready for international comparison and immediate analytical use, as well as on maki ng them easy to find, manipulate and download.

6.　　Users may be particularly interested in comparing data for countries of the Commonwealth of Independent States with those for other European and North American countries. It is often difficult, or even im possible to find other data sources with this sort of information held in a single integrated database.

7.　　The UNECE compiles data from a wide range of different official national and international sources. This involves active cooperation with other international organizations to share data, and minimise the reporting burden on individual countries. Where we have to collect data directly, we either take them from national web sites, or use questionnaires pre - filled as far as possible with existing data.

8.　　The data update frequency varies from monthly for some macroeconomic statistics to annual for gender and transport statistics and once every five years for forestry statistics. Improving data timeliness has been a major focus of UNECE work over the last two years.

9.　　As well as being available directly to the public via the Internet, the UNECE Statistical Database is used as the source for various other statistical outputs, notably the biennial publication "UNECE Countries in Figures[3]". Data of topical interest are also regularly extracted for short analytical articles published in the UNECE Weekly[4] newsletter, which is widely disseminated to customers across the whole range of UNECE activities .

III.　　Modernizing Statistical Production

10.　　A general theme that cuts across UNECE database work, and the work of several Conference of European Statisticians (CES) expert groups, is the modernization of the statistical production process. Typically this involves moving from domain -based statistical production processes to a more harmonised, but also flexible and modular, process -oriented production system.

11.　　The CES expert group on Statistical Metadata (METIS) has produced a "Generic Statistical Business Process Model"[5] (GSBPM), which has been adopted by many national statistical organisations around the world. It is used as a basis for developing and describing a more harmonised approach to statistical production. It is also used by expert groups on managing statistical information systems (MSIS) and softw are sharing, as a basis for describing business architectures, and facilitating the development of common software tools.

12.　　The basic principle behind the GSBPM is that all statistical production can be mapped to a single process model. This means that a range of sub-processes can be identified that are common to production processes both within and between statistical organisations (for example, "impute" or "calculate aggregates"). These sub-processes are described using common terminology, facilitating comparisons between processes and organisations. This makes it easier to develop common methods and software.

---

[2] http://www.unece.org/stats/archive/docs.fp.e.htm
[3] http://www.unece.org/stats/profiles2009/Welcome.html
[4] http://www.unece.org/highlights/unece_weekly/unece_weekly_index.htm
[5] http://www.unece.org/stats/gsbpm

13.   The GSBPM has been used as the basis for the work of the current European Union "ESSnet" project on a Common Reference Architecture (CORA) . This project is gathering useful information on statistical software and development projects, classified according to the GSBPM, which may be suitable for sharing between organisations. The resulting inventory has been published on the MSIS Wiki[6].

Applying process-oriented principles to re-engineer the UNECE Statistical Database

14.   After the development of the statistical data dissemination system based on PC-Axis was completed, the UNECE started to re-engineer its own statistical data production system following the principles of the GSBPM. The new system will follow a flexible and modular system architecture (see figure 1 below).

15.   The multi-dimensional cube concept, a central feature of PC-Axis, is being extended to all stages of data processing. This should make data management easier, improve data quality and considerably shorten the time required for data to pass from collecting through processing to disseminating in PC-Axis.

16.   This approach gives UNECE statisticians more direct control over statistical production than before. They can now develop data validation and computation routines themselves. As a result, scarce IT resources are used more efficiently and are focused more on the development of new functionalities. An example of the sort of new functionality being developed is the automatic upload of data from key sources, reducing manual intervention and improving timeliness.

Figure 1 – High-level database system architecture in the UNECE

---
[6] http://www1.unece.org/stat/platform/display/msis/Software+Inventory
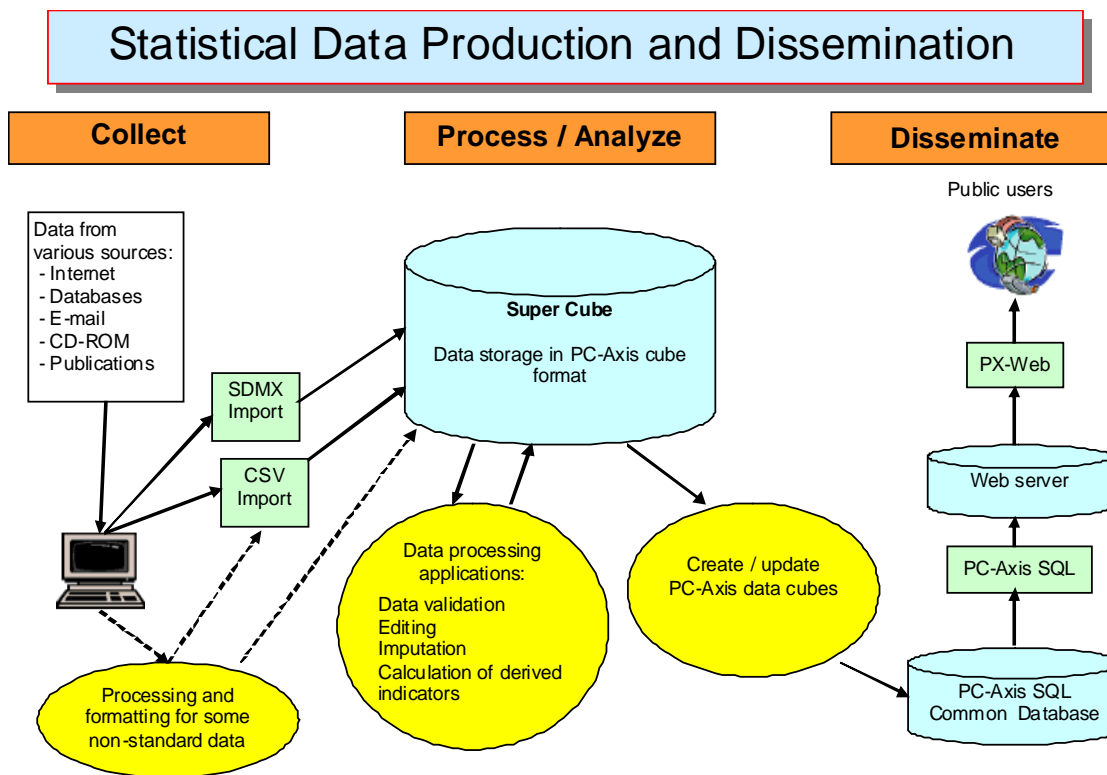
Applying SDMX standards

17.   The Statistical Data and Metadata eXchange (SDMX) standards have been endorsed by the United Nations Statistical Commission as preferred standards for exchanging data and metadata. Although originally conceived as a set of standards for exchange between organisations, SDMX is now being applied to data and metadata exchange between components of the statistical production process, within organisations.

18.   The use of common exchange standards between components is seen as a way of further increasing the flexibility and modularity of statistical production systems, and facilitating the development of common software tools. The UNECE is currently investigating the use of SDMX in this way, making maximum use of existing open-source tools, such as those developed under various Eurostat projects. this is in line with current thinking in several leading national statistical organisations. Initial results look very promising, particularly if the complexity of the SDMX format can be reduced.

IV.   Monitoring Database Use and Collecting Customer Feedback

19.   The number and range of users for the UNECE Statistical Database has grown significantly since it was released to the public four years ago. Database use is monitored both in terms of user downloads of datasets using monitoring facilities built into the PX-Web software, and in terms of "hits" to the database website using the free "Google Analytics" application.

20.   Data downloads per year are shown in Figure 2. They have more than tripled between 2006 and 2009, and this growth seems to be continuing during 2010 with a new monthly record of over 22,000 downloads in March.

Figure 2 – Data Downloads per year

Downloads Number
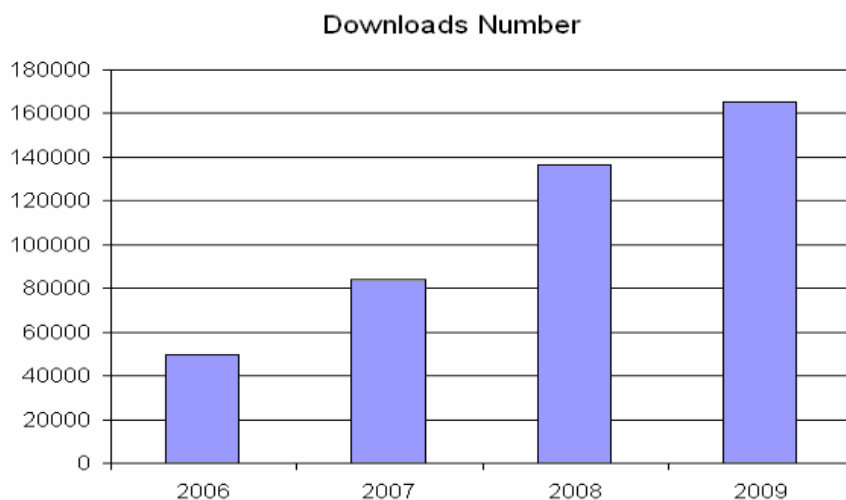
21.   Whilst the results are not strictly comparable, Google Analytics gives a range of complementary information about users, including geographic location, frequency of use, type of Internet browser and how they arrived to the database.

Figure 3 – Location of database users, January to April 2010
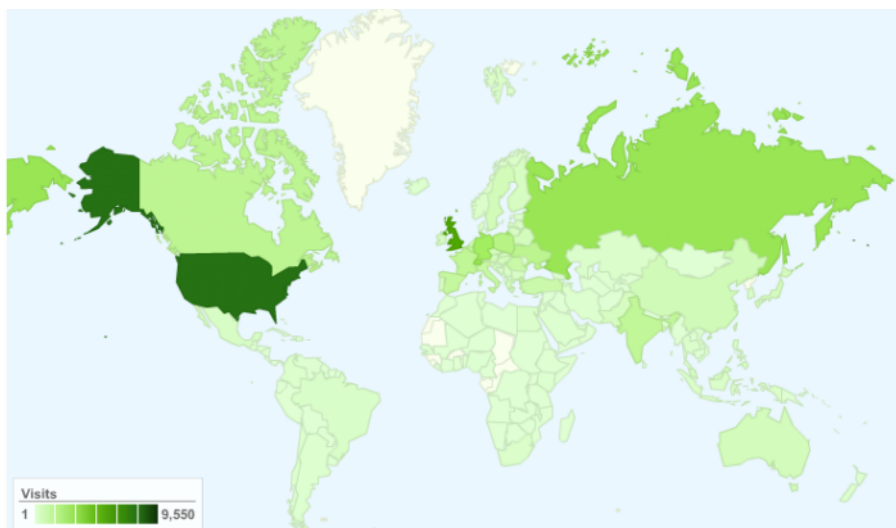
64,918 visits came from 189 countries/territories

22.   Each year since 2007, the UNECE has conducted database user survey via a web-based questionnaire. A link to this questionnaire is added in a prominent place on the database home page. Users are asked ten questions about themselves and what they think of the database. The results of the 2009 survey are summarized below. These results feed into the annual data quality improvement plan, and have a strong influence on database development priorities, ensuring that user needs are met as far as possible.

Key findings of the 2009 UNECE Database User Survey

- Over 84% of respondents rated the overall quality of UNECE data as good or excellent;
- Over 78% thought the usability of the database web sit e was good or excellent;
- 40% of respondents found the information in the database to be highly relevant to their needs, whilst 55% found it to be partly relevant;
- 77% of respondents were satisfied with the timeliness of UNECE data, and 91% were satisfied with the metadata provided;
- 53% of respondents wanted data on population, 48% on other economic indicators, 48% on other social indicators, 43% on employment and 42% on National Accounts

---

V.   International Cooperation and Technical Assistance

23.   The UNECE takes an active part in the PC-Axis consortium, which organizes annual meetings where participants discuss issues and priorities for the future development of PC-Axis tools. UNECE experts are happy to share experiences and any of the applications they have developed.

24.   Over the past few years, the UNECE has organized workshops on data dissemination using PC-Axis in Geneva and Kazakhstan. We have also provided training, demonstrations and technical assistance relating to PC-Axis to the national statistical organisations of Ethiopia, Kazakhstan, Kyrgyzstan, Mongolia, Switzerland and Tajikistan, as well as to colleagues in the United Nations Food and Agriculture Organization (FAO) and the United Nations Office on Drugs and Crime. It is important to note that the UNECE Statistical Division can provide technical assistance activities in both English and Russian languages.

VI.   Future Plans

25.   Over the next few years the UNECE plans to complete the database modernisation programme, including the full implementation of SDMX standards and upgrading to the latest versions of PC-Axis software tools. The contents of the database will be expanded by the addition of data on progress towards the Millennium Development Goals. We are currently exploring the possibility of adding key indicators relating to energy use and the environment.

26.   There are plans to further develop the work of CES groups of experts relating to modernising the statistical production process. This will include providing mechanisms for the exchange of ideas and experiences, and tools to help implement new approaches. The existing inventory of statistical software suitable for sharing between organisations will be maintained and expanded, and the GSBPM will be enhanced by additional information describing its implementation in practice.

VII.   Conclusion

27.   The UNECE is happy with the choice of PC-Axis software for statistical dissemination. PC-Axis tools have been successfully integrated with the existing technical infrastructure. They also fit well with future plans for a more modular, process-oriented systems architecture.

28.   The UNECE strongly believes and promotes the principle of disseminating data free of change. Data should be disseminated in a range of formats, but the focus should always be on providing easy access to key data through a simple, user-friendly interface. This approach is driven by user needs. It is facilitated by the growth of open standards such as SDMX, and the availability of high quality dissemination tools such as PC-Axis.