ECONOMIC COMMISSION FOR EUROPE   **STATISTICAL COMMISSION**

CONFERENCE OF EUROPEAN STATISTICIANS

Fifty-fourth plenary session
Paris, 13-15 June 2006
Item 8 of the provisional agenda

SEMINAR ON HUMAN RESOURCES AND TRAINING
SESSION II

Educating the public – the role of e-learning and visual communication of official data[1]

Submitted by Eurostat

I.   INCREASING PUBLIC INTEREST IN DATA FROM OFFICIAL STATISTICS

1.   Reliable statistical data are indispensable for evidence-based policy making. They play a crucial role in planning political decisions, in quantifying goals to be benchmarked against the real-world observations and in assessing the performance of political strategies. Public institutions at international or national level (European Commission, European Parliament, ECB, OECD, UN agencies, ministries and other government bodies) as well as further players in society closely monitoring political developments (trade unions, employers' associations, political parties) have been for a long time almost the sole customers of information disseminated by statistical offices. In the past, producers of statistics at international level traditionally geared their services to the needs of this important group of data users and left it more or less to the media to communicate relevant statistical information to the public at large. Hence, official data have been for a long time a public good that was not fully exploited. The available statistical information has been either not directly accessible or even unknown to the potentially interested lay public, such as the school teacher, the student or the self-employed.

---

[1] This paper has been prepared at the invitation of the secretariat.

2.   In recent years, this situation has changed dramatically. National and supranational statistical offices have experienced a fast transition of their dissemination strategies towards a policy giving open web access to statistical information. This development has been driven by advances in ICT and by the requirements and expectations of today's knowledge- and information-based society. Making data available via the Internet free of charge creates new demands and attracts new groups of customers. Eurostat, for example, switched to a free online dissemination of publications and non-confidential data in October 2004. The monthly number of user sessions involving Eurostat's web site amounted to over 930,000 in February 2006 – an increase of more than 530 % since the beginning of the free dissemination. The immediacy of online access to data and publications obviously attracts additional users and places the data producers more and more under the spotlight. This opens a window of opportunity to attain greater visibility in society, to become a trusted agent for the citizen in general and to be a key contributor to statistical literacy.

3.   The flip side of the issue is that the homogeneity of the user community is fading away whilst sound empirical information about new groups of customers and their needs is rarely available. The existing information gap involves a risk which again refers to public trust. A non-professional data user will not consider a data producer as really trustworthy and credible if the disseminated data are not presented attractively in a user-friendly way, easy to understand and accompanied by a minimum of meta information and guidance. Clear and understandable statistics are trust-enhancing factors for official statistics or even a precondition for public trust ([12]). The European Statistics Code of Practice ([4]), an embodiment of the UN Fundamental Principles of Official Statistics for the European Statistical System, addresses this item under the quality dimension "accessibility and clarity". Nowadays, with a multi-user group framework, it becomes more difficult to attain the goal of providing statistical information which is comprehensible for everyone.

## II.   ADDRESSING THE NEEDS OF THE NON-PROFESSIONAL USER

4.   In order to meet the expectations of such a broad audience, statistical offices have started to identify evolving new target groups and to learn about their needs by implementing suitable strategies for measuring user satisfaction with data and services. Some offices, for example that of Finland, evaluated the level of trust and confidence in official statistics by means of questionnaires. Others, such as those of New Zealand and Latvia, tried to identify the composition of its audiences and to learn about their expectations from the dissemination of official data.

5.   It is indisputable that successful communication with the increasingly important group of non-professional customers requires that statistical offices go far beyond the simple provision of tables and other purely static information. The visual presentation of data through comprehensible and flexible graphical tools, possibly embedded in a story-telling environment and connected with maps for the presentation of spatial data, crucially contributes to meeting the needs of the non-expert. This is not meant as a plea for turning official statistics into "infotainment" but as an appeal for exploiting hitherto unexploited potential by making use of complementary dissemination channels. Communicating statistical information in a way easily understood not only by the traditional customers will lead to a broader interest in work done in statistical offices. Official data will be more often applied in educational settings if they are

presented by means of interactive visualization tools that offer space for ad-hoc exploration and communicate messages behind statistical information at first glance.

6.   Figure 1 illustrates this by means of a prototypical interactive visualization tool, applied to a time series on unemployment rates for European countries, Japan and the USA, showing how user-controlled environments for the display and spontaneous exploration of data could appear. The Java-based tool is a result of adapting developments from a non-commercial German multimedia initiative aiming at improving the quality of statistics education ("New Statistics" project [1]). The tool is ready for online and offline use, for example as an integral part of a PowerPoint presentation in a press conference. The example presented in Figure 1 and further applications are publicly accessible [2].

Figure 1: Self-contained interactive visualization of unemployment rates

a. Time series for user-selected countries   b. Bar chart for a user-defined year
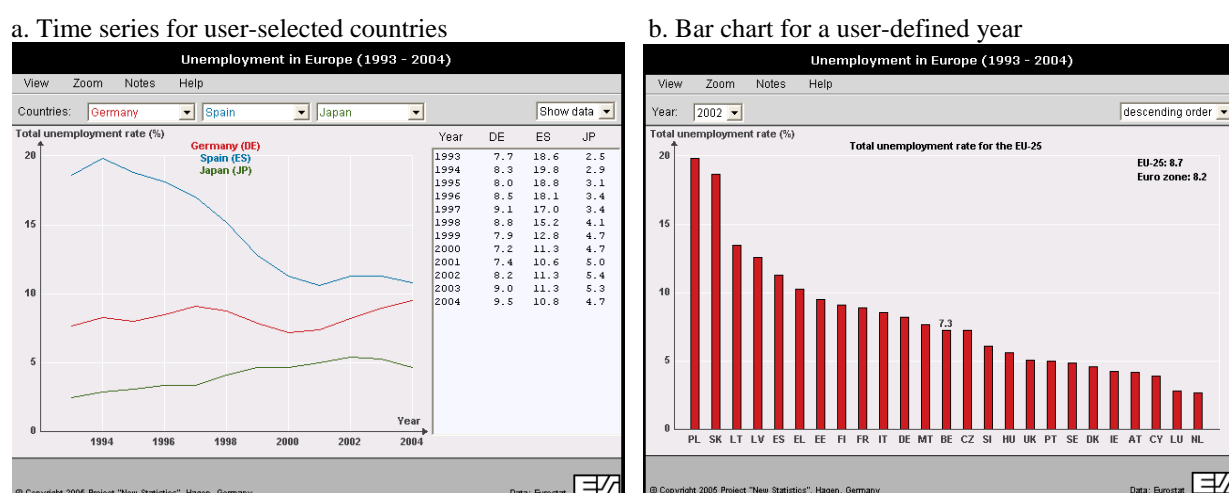
7.   The actual data can be either displayed, as shown in Figure 1a, or suppressed, depending on the user's choice. The visualization tool is self-contained with built-in meta information accessible via a second layer. Furthermore, it offers the option of viewing the data from different perspectives by applying different graphical instruments, for example time series graphics, bar charts or boxplots. Figure 1b refers to a bar chart presentation for a reference year which has to be selected via a menu window. The user may re-order the countries or slot in the numerical value defining a specific bar as illustrated for Belgium. Screenshots of any of the multitude of potentially available graphics can be saved, for example by using the "Paint" functionality of Windows or a specialized screenshot capturing program, and exported to other working environments. In principle, the tool presented in Figure 1 is applicable to any data set and can be directly connected via an appropriate interface to a database, thus enabling automated regular updates.

8.   Figure 2 provides a second example illustrating the flexibility of the same visualization tool, but now applied in different settings. Figure 2a uses the tool as a self-contained element for communicating European earnings data (for details see [10]). The screenshot in Figure 2a presents a bar chart referring to gross hourly earnings, expressed in euro, for a user-selected economic activity. Again, the user can re-order the countries, display the numerical value

belonging to a bar or slot in the country code as shown for Sweden. Furthermore, the user is free to apply different styles to the way the data is presented.

9.   Figure 2b illustrates that the visualization tool can be likewise embedded into an online publication, thus opening up a completely new dimension for web-based publication. The incorporation of dynamic graphs into hypertext environments leads to interactive publications with obvious added value compared to traditional online publications. The reader of such a publication is invited to choose a personal path for data exploration and discovery within a predefined framework. The screenshot in Figure 2b shows an excerpt from a publicly accessible experimental Statistics in Focus edition recently disseminated by Eurostat (see [3], first entry) concerning the main results of the recent European structure of earnings survey. This interactive publication offers graphics defaults which can be changed, providing access to a multitude of further user-selected graphics. Embedding interactive graphics displaying official data into hypertext environments will certainly be applied to a large extent by online journalists if statistical offices are ready to provide the graphics.

Figure 2: Visualization of data on gross earnings in different environments

a. Bar chart for a user-selected economic activity (self-contained element for online and offline use)
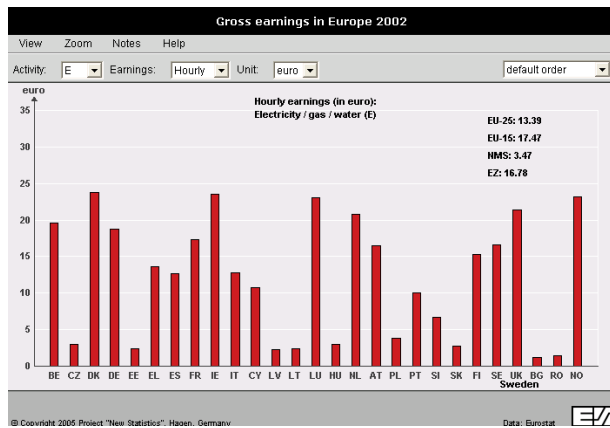
b. Box plots displaying earnings dispersion within countries (interactive graph embedded in a hypertext environment)
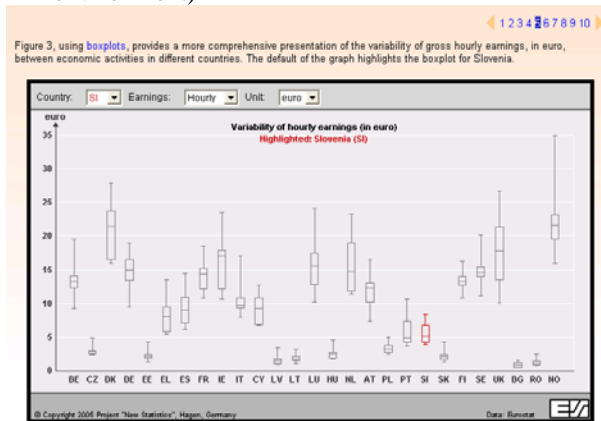

10.   The data driven visualization tool presented so far represents a starting point for discussion, further reflection and refinement. There exist differently designed developments stressing visual data communication. An impressive example is the dynamic population pyramid, first developed for the population of England and Wales, later for that of Germany, Norway, Australia and a few more countries ([5]). The pyramid presented in Figure 3a refers to the Norwegian population and is based on scalable vector graphics (SVG), a promising and rather new web standard appreciated by the open source community. The Office for National Statistics of the UK has been a forerunner in successfully experimenting with this format. Unfortunately, a native support for the SVG format in web browsers is still not always ensured (for details, see ([13]). SVG graphs can be animated and enriched with highly interactive features. The Norwegian pyramid displays past and current population figures as well as projections into the future. The changes over time can be viewed like a movie but the user can also interact with the graphs, for example by displaying figures for a user-defined age group. This is illustrated in

Figure 3a by means of the age class "35 years" for the reference year 2006.

11. Figure 3b shows an interactive map for the German State Baden-Württemberg. The map represents the key element of a geographical information system applied by the statistical office of Baden-Württemberg for displaying information on spatial data related to numerous variables (see [6]). The final results of the German elections in September 2005 and that of the preceding elections have been officially presented by using an interactive map which has been linked to dynamic statistical graphics ([7]). The online editions of a well-known German weekly magazine made direct use of this attractive communication tool on its web site.

Figure 3: Visualization of population data by means of animated graphics or interactive maps

a. Animated population pyramid
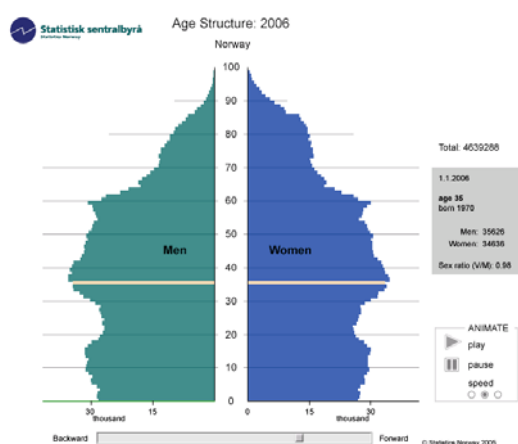
b. Display of spatial data via an interactive map
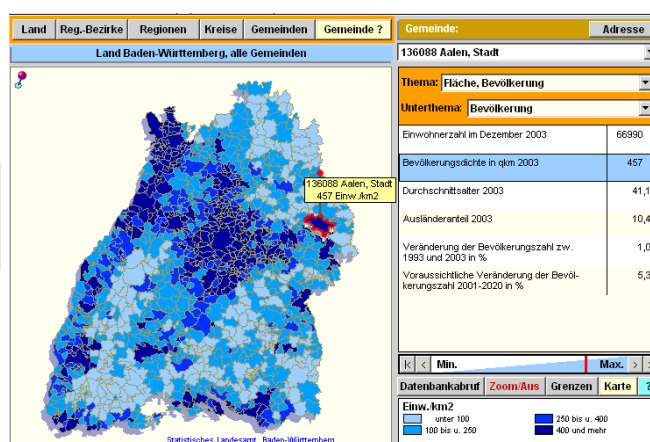
12. Advanced visualization tools for communicating official data by means of graphics and maps, freely accessible via search engines and web sites of statistical offices, attract the lay public and contribute to promoting statistical literacy. It will be part of any efficient customer relations management in statistical offices to complement such eye-catchers by space for user feedback. The latter serves modern quality management strategies not only to meet but to exceed the expectations of customers via permanent improvement of products and services.

13. The added value of dynamic graphics and interactive maps goes beyond providing more user-friendliness and immediate availability of an almost unlimited amount of statistical information. Interactivity contributes to the discovery and unlocking of information which is more difficult or even impossible to grasp from tables or static graphics. Environments for user-controlled data exploration, as shown in Figures 1-2, enable the user to look at a data set from different perspectives and to better exploit the potential of information behind the data. An animated population pyramid as presented in Figure 3a communicates structural changes, such as the overageing of a society, in a way easier to remember by focusing on visual perception. The same statement holds for dynamic maps which are unbeatable in unlocking striking features behind regional data, such as concentration phenomena. Furthermore, modern visualization tools

can be applied internally by statistical agencies for checking data consistency and for identifying processing errors.

III. ESTABLISHING VIRTUAL LEARNING ENVIRONMENTS FOR EDUCATION AND TRAINING

14.   Statistical literacy is part of the knowledge and information which represents the key driving force in today's fast-moving globalized economies. National and supranational statistical offices ought to be active players in the process of producing and distributing this part of knowledge. Modern environments for e-learning, complementing traditional modes of education and training, offer innovative and efficient ways of promoting the use of official data and fostering critical awareness as regards the appropriate presentation and interpretation of data. The media will continue to be important partners for attaining this task but the ubiquitous availability of Internet access enables statistical agencies to also address the public directly.

15.   Therefore, statistical offices should offer, within the limitations defined by constraints in human resources, web-based statistical information tailor-made for important user groups, such as educational institutions and the media. The goal is to build up closer partnerships, to increase visibility in public life and to promote statistical literacy. With a view from Eurostat, an appropriate measure could be to establish an "educational corner" on Eurostat's web site presenting, for example, by means of user-friendly visualization tools, selected data sets of general interest, complemented by didactical comments and meta information on methods or sampling frameworks. Several national offices have already gained experience with a web-based free dissemination of learning resources ([8]). The Canadian statistical agency offers a variety of learning resources for teachers, students, postgraduates and pupils ("kid's zone"). The repository contains traditional as well as new media, for example flash animations and slide shows, although most of them with rather moderate interactivity. Statistics Finland also successfully networked with the educational sector and offers a few e-learning modules presenting important areas of work in official statistics, such as population statistics or national accounts. For the time being, these modules only include a few dynamic elements, but they provide links to educational web sites of other statistical offices, such as those of Ireland or Portugal ([8]).

16.   Close cooperation between educational institutions and official statistics, often neglected due to lack of interest or time, is of mutual benefit to all parties involved (see [11]). On the one hand, attractive and ready-for-use educational material from official statistics is highly appreciated by teaching staff at universities and schools. It helps to illustrate that statistical methods are of paramount importance for backing decision-making in all areas of life. On the other hand, statistical offices can make use of existing know-how as regards the development of e-learning content and virtual learning environments for the public.

17.   The public multimedia project "New Statistics" ([1]) mentioned earlier in this paper gives an example of how cooperation can be organized. The project output contains a virtual library with three sub-libraries giving access to user-controlled experiments visualizing statistical methods and another sub-repository with interactive elements visualizing official data. The fourth sub-library, covering amongst others the data displayed in Figures 1a-b and Figure 2a, is shared by Eurostat and the project consortium of ten universities. It suggests that statistical offices can re-use and adapt high-quality e-learning content originating from the educational

world not only for addressing the lay public but also for vocational training of their own staff.

IV.   A PLEA FOR INTERNATIONAL COOPERATION

18.   Any innovative framework to be built for promoting statistical literacy of the citizen and efficient statistical training inside statistical agencies will make use of current ICT technologies, advanced visualization tools and modern concepts for e-learning. An intensified cooperation between national and supranational offices as well as with mapping agencies and educational institutions is needed, aiming at a systematic exchange and re-use of self-contained e-learning components (Flash animations, Java- and SVG-based applications, complete learning modules). A far-reaching approach is the implementation of a document management system for non-proprietary e-learning components suitable for educating the public and for statistical training on the job. Such a system shall be based on international standards for learning object metadata (for a prototype of such a catalogue see [9]). The experiences of Korea in building up a nationwide information system for e-learning content including metadata can be studied in this context.

19.   International cooperation between statistical agencies and close ties with educational institutions helps to avoid re-inventing the wheel and to increase cost-efficiency. Systemic cooperation might start with developing and sharing eye-catching elements related to issues of core interest for the public, the media and for the statistical offices. A potential eye-catcher is a population pyramid for the EU or another set of countries. Such a pyramid will go beyond that shown in Figure 2a by including the option to change the country or to simultaneously display the pyramids for two user-selected countries. The approach can be adapted to displaying other multinational data sets, for example data on labour forces, unemployment and job vacancies. Furthermore, one might visualize, similarly to Figure 1, statistical indicators for a set of countries but in connection with interactive maps and with the option of language change. Press releases can be complemented by web links providing access to such visualization tools.

20.   Apart from visual communication of data via self-contained and regularly updated eye-catchers, statistical offices could cooperate in exploiting the potential connected with online publications that leave the user the ability to genuinely interact with the data. Interactive publications could incorporate not only dynamic graphics, as presented in Figure 2b, but animated elements or dynamic maps as well.

V.   WEB LINKS AND OTHER REFERENCES

[1]   Web site of an interdisciplinary German multimedia project for statistics education:
       http://www.fernuni-hagen.de/newstatistics
[2]   Virtual library illustrating interactive visualization of Eurostat data:
       http://forum.europa.eu.int/irc/dsis/wages/info/data/index.htm
[3]   Web site with interactive publications dealing with Eurostat data:
       http://forum.europa.eu.int/irc/dsis/wages/info/data/interactive.htm
[4]
       http://epp.eurostat.cec.eu.int/portal/page?_pageid=2273,1,2273_47140765&_dad=portal&_
       schema=PORTAL
[5]   Population pyramids for different countries or regions (United Kingdom, Germany,

Norway and Australia):
http://www.statistics.gov.uk/populationestimates/svg_pyramid/PP6101_4.svgz
http://www.destatis.de/basis/d/bevoe/src/svg-01/start.php
http://www.ssb.no/english/subjects/02/03/folkfram_en/pyramid/
http://www.abs.gov.au/websitedbs/d3310114.nsf/home/population%20pyramid

[6]   Interactive map displaying various statistics for the German State Baden-Wuertemberg:
http://www.statistik.baden-wuerttemberg.de/intermaptiv/archiv/home.asp

[7]   Interactive map displaying German election results:
http://www.bundeswahlleiter.de/wahlen/bundestagswahl2005/onlineatlas/btwClientKarte.htm

[8]   Learning resources offered by the statistical offices of Canada, Finland, Ireland and Portugal:
http://www.statcan.ca/english/edu/index.htm
http://www.stat.fi/tup/verkkokoulu/index_en.html
http://www.cso.ie/studentscorner/
http://alea-estp.ine.pt/ingles/index.html

[9]   Repository for open source multimedia assets:
http://mmc.mmk-hagen.de/

[10]   Grünewald, W., and Mittag, H.-J. (2006). The use of advanced visualization tools for communicating European data on earnings to the citizen. Invited paper to be presented at the 7th International Conference on Teaching Statistics (ICOTS 7), San Salvador, Brazil, July 2006.

[11]   Helenius, R. (2005). Co-operation with educational institutions: a strategic challenge for statistical offices. Paper presented at the 55th ISI session, Sydney, Australia, April 2005.

[12]   Plovsing, J. (2005). Closing speech of the 91st Conference of the Directors-General of theNational Statistical Offices (DGINS Conference), Copenhagen, May 2005. In: Conference Proceedings, p. 175-176, Eurostat / Office for Official Publications of the European Communities, Luxembourg.

[13]   Smith, A., and Calder, A. (2003). Unlocking information through data visualization – some case studies in using SVG. In: A Beginner's Guide to Scalable Vector Graphics. CD-ROM, Office for National Statistics, United Kingdom.

* * * * *