WP. 5
ENGLISH ONLY

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

Workshop on the dissemination of census results
(Geneva, 16 May 2008)

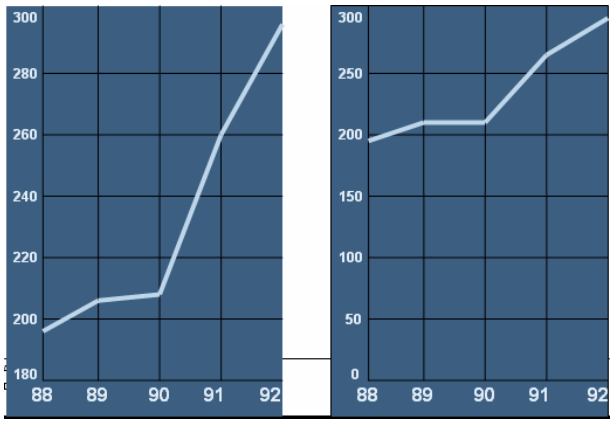
"MAKING DATA BE UNDERSTOOD – AND EASILY PRODUCED" INTERACTIVE VISUALIZATION FROM A USER AND PRODUCER PERSPECTIVE

Submitted by Federal Statistical Office of Switzerland[1]

> "With every human being a new world is born which did not exist before he saw it, which will never exist again when death closes his eyes. … the world , which is nothing but life as seen by the individual .."
>
> *Sadakichi Hartmann, cited by Penelope Niven in Steichen. A Biography, 1997.*

1. Our understanding of the world around us largely depends on images. Images determine how we see and explain facts, images help us reduce complexity and gain orientation. As individuals we live in a realm of images and build our (more or less individual) "image of the world" for ourselves. Modern media make us depend more and more on images when navigating and interpreting the world and thus images are a very powerful instrument and on the other hand a very dangerous one as well. To visually present statistical information implies that such images will have an influence on how the public understands reality (whatever that is).

2. A visualization of religious affiliations in a country, or of economic development or criminality helps in shaping opinions. How we choose, for instance, to visually depict the development of criminality can have a huge effect on the message and, consequently on opinion building.



Offences against the Criminal Justice Act: 2 representations with the same data but contradictory messages

3. Nowadays, most statistical organizations and institutes enjoy the benefit of outstanding tools and modern techniques for the visualization of statistical data in the form of graphs or maps. The communication of statistics has consequently become much easier and statistics themselves have gained in popularity (at least we like to think so).

4. The quality of a visual representation is measured according to how these prevailing tools on the one hand, and user-friendliness as well as political correctness on the other hand, will ensure didactic value and

[1] By Armin Grossenbacher, Federal Statistical Office of Switzerland. With thanks to Daniel von Burg and Alain Nadeau.

quality. Techniques and trends in visualizing have to be checked against this. In fact we carry a significant responsibility when visualizing data.

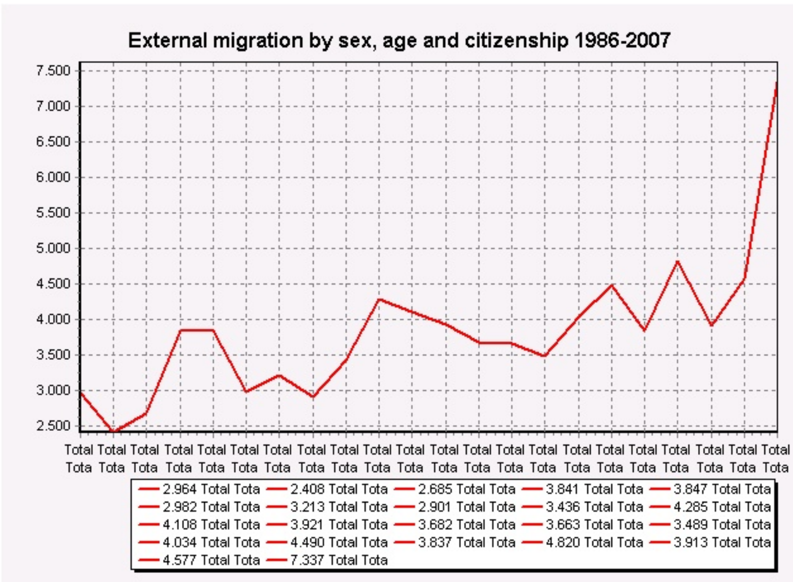
5. So what about the trend towards interactive visualizations of statistical information?

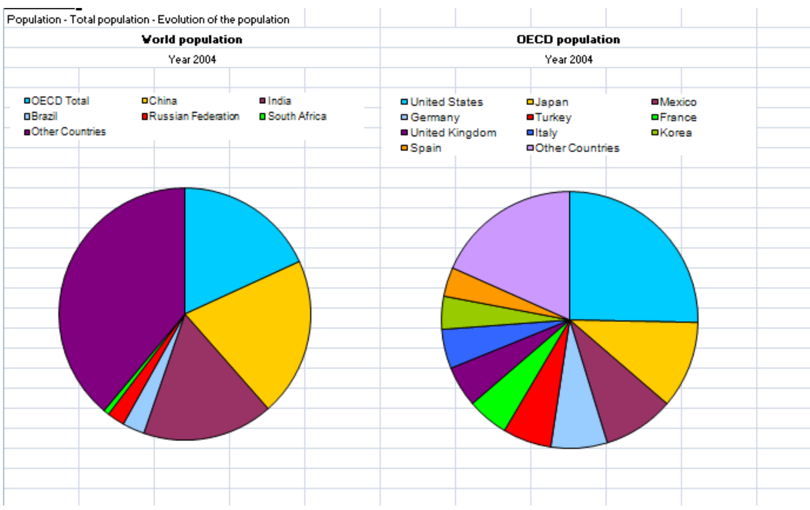
User-side interactivity

6. Newer trends tend to avoid the display of static, predefined graphs and maps, preferring instead to provide users with the possibility of defining their individual needs and modifying predefined presentations. What are the best kinds of user interactivity in this field? Where are the chances, where the risks?

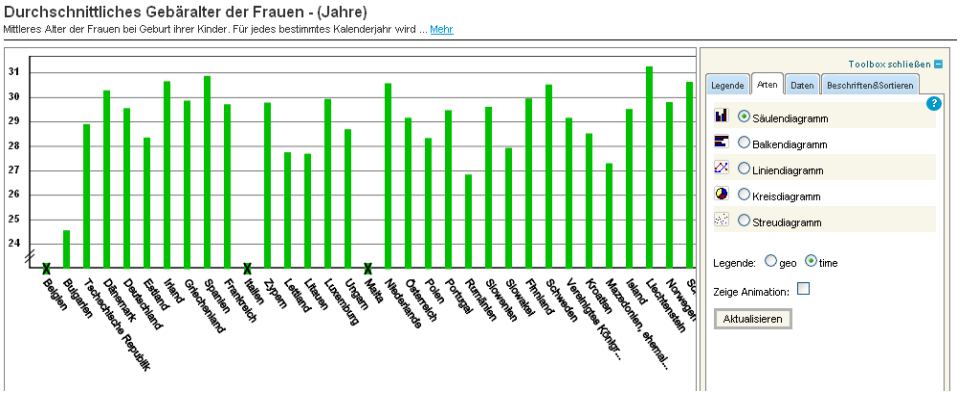
7. Graphs created by skilled specialists ensure a certain degree of quality. But their production is slow and time-consuming. Providing users with hands-on interactive graphical applications opens new possibilities, but it also carries the danger of constructing images without controlling quality aspects.

- Populating graphs with data in an automated way is **one widespread type of interactivity**. Interactivity is restricted here to building a graph on the basis of predefined tables; EXCEL or online datbases such as PC-Axis and others allow this functionality.

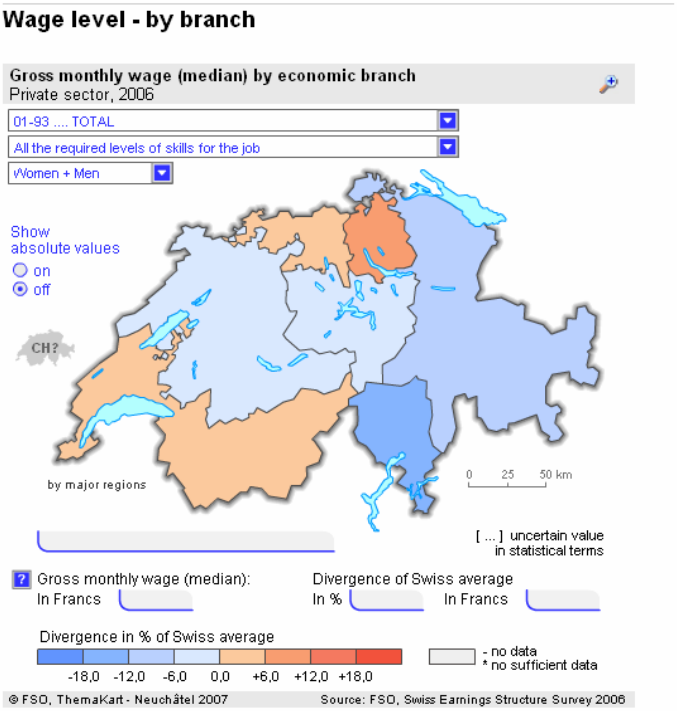




- **Another type of interactivity** gives users the possibility of changing the appearance of graphs. Choosing bar or pie chart presentation, colours and size are part of this interactivity. This is not a full interactivity: it is more a playground for adapting presentation aspects to user preferences.



- A **third type of interactivity** does offer a choice in information. In predefined graphs variables can be changed and new information elements displayed. Some examples:



http://www.bfs.admin.ch/bfs/portal/en/index/themen/03/04/blank/key/lohnstruktur/nach_branche.html
see also:
Gapminder: http://www.gapminder.org/
IMF datamapper: http://www.imf.org/external/datamapper/index.php

- All of these graphic types can also be **dynamic**. Users can decide to let variables or developments in time be presented in a film. But such films are usually difficult to follow. The effect is very attractive, but our intellect normally cannot memorize the differences and grasp the information.
For didactical purposes it would be better here to offer the differences side by side or to build graphs and maps that show the changes as such.

- A **new and special kind of interactivity** is a combination of interactivity, dynamics and explanation. It is interactivity conceived, controlled and presented with commentary to the users by intermediate persons. This television-like presentation of statistics was elaborated by Hans Rosling. By playing with his trendalyzer tool (now Google) he delivers an outstanding example and makes fascinating statistics courses for a wide public.

http://www.gapminder.org/video/gap-cast

Producer-side interactivity

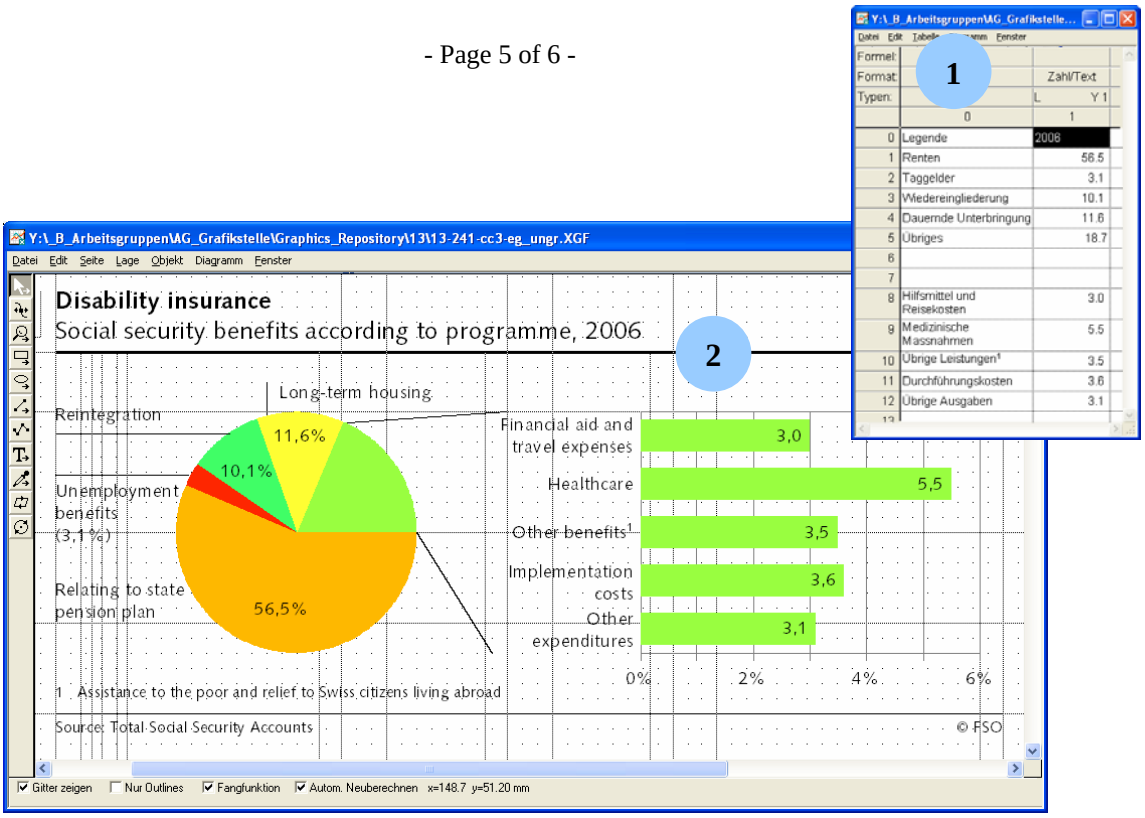
8. Automated production of graphs with software like Excel or some Online-Databases (see user-side interactivity above) leads to graphs offering a first insight in the meaning of the visualized data. These graphs normally are “working graphs” but not necessarily “publishing graphs” as they come with some flaws.

9. High-quality graphs need authoring. But traditional manual production of high-quality visualizations by specialized services becomes increasingly difficult given the ever-present budget cuts in statistical offices on the one hand and the increasing amount of visualization needs on the other hand. So we observe an increasingly important trend in making the publishing of high quality graphs and maps both cheaper and....faster.

10. Following this approach authors have to be provided with all the necessary interactive tools to assemble and publish graphs. A step further would be to use specific procedures so as to ensure automatic or semi-automatic production at given quality levels.

Template-driven graphics production

11. Specialists can use semi-automatic tools to build templates for graphics and to introduce data via i.e. Excel files or xml. This technique allows for the fast and easy updating of an application: it is kid's play to add an additional year to a graphic. This is one of the techniques favored by Statistics Switzerland. Using the program Xact we have been producing high-quality templates with professional results, all managed by a small central unit. The same procedure is used or made available for other graphs, for instance Flash based graphs or graphs on external platforms like Goggle spreadsheet with graph types like Trendalyzer (now: Motion Chart, see: http://blogstats.wordpress.com/2008/04/13/trendalyzer-becomes-motion-chart/).

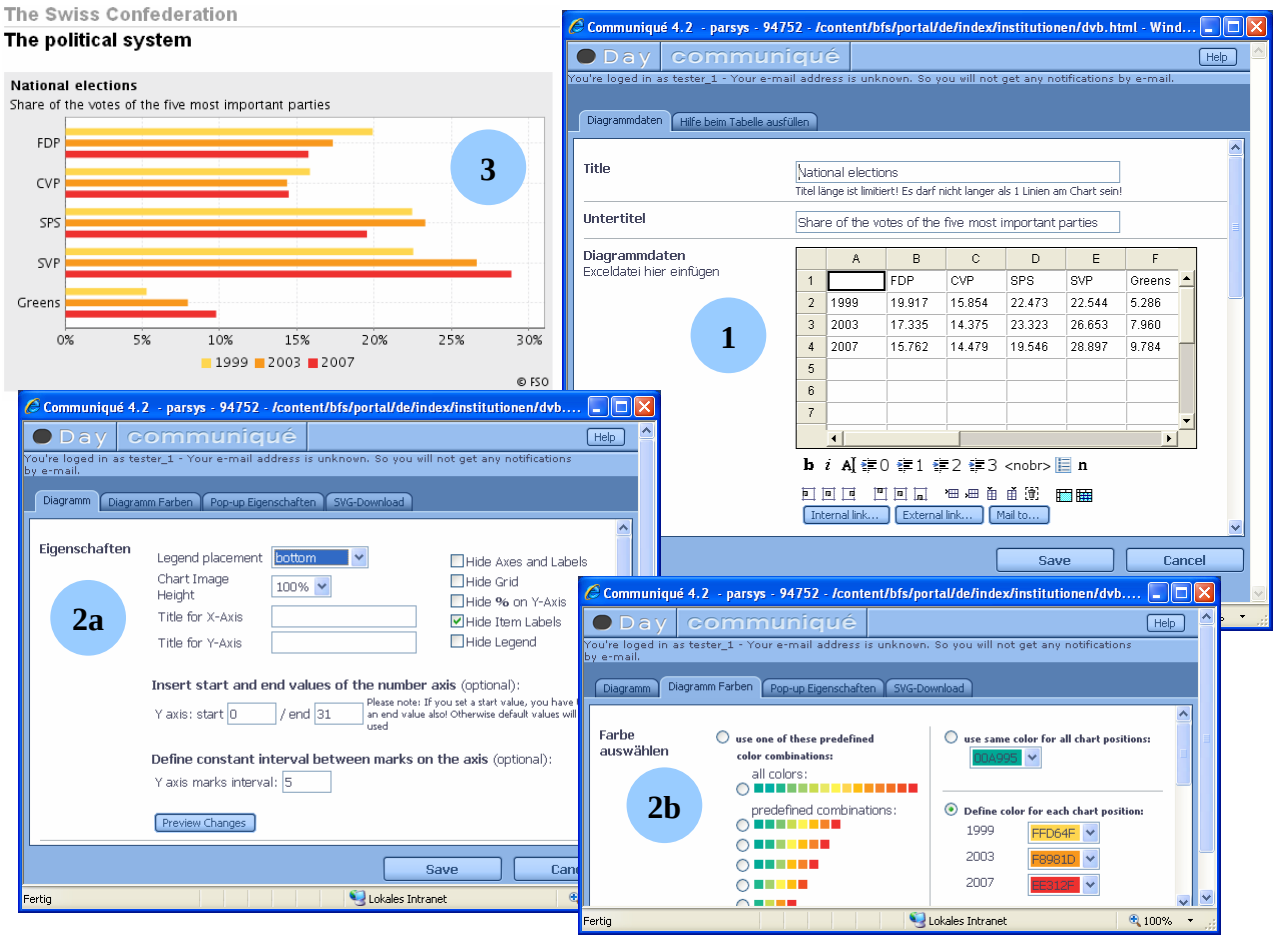


Xact is very useful for rapid edition of charts by specialists. The explorer view helps finding the file wanted. It is easy to offer the template in different languages. To visualize another year just paste the new data into the table file: lines 1–5 are used for the pie, line 8–12 for the bar chart. The dotted lines in the background show the grid of a standard FSO publication, to produce a GIF for the web we just drag and drop the EPS for a print publication into Photoshop and press a key to run a macro.

Template and graphics power to the authors

12. This producer-side semi-automatic approach still requires the presence of specialists in the publishing workflow. Granting authors the direct publishing power in the field of visualizations is a further approach. Here, Web Content Management Systems (CMS) open new possibilities. Today, at Statistics Switzerland, more than 100 authors are editing their thematic webpages using such a powerful CMS. They can edit texts and introduce downloads of files of different formats. A new template will now give them the possibility to build graphs interactively and to publish them directly on their pages. Predefined templates for author usage and the central activation service ensure quality before publishing.

13. In a future step these graphs will be exported via xml to other publishing tools and so opening cross media opportunities.



The Web CMS is very useful for authors lacking training in typography. All data is edited in **a CMS template**. Authors may fine-tune some of the chart properties to make the data easier to understand, e.g. axis and colour properties.

Automatically generated high-quality graphs

14. The last step i.e. the automatic production of high quality graphs is not yet made. Hereby quality controlled templates would be filled and updated with data via a direct database access. (i.e. Flash Dynamic). So graphs presented in a web page or pdf-reports could be produced and presented easily and rapidly.

15. Projects aiming to achieve this are underway and an international discussion about this is of fundamental interest for us.

*****