

# Improving visualisation tools in EUSTAT: Explaining the data

**Jorge Aramendi Rique (Eustat), Anjeles Iztueta Azkue (Eustat),  
Cristina Prado Valle (Eustat), Yosu Yurramendi Mendizabal (EHU-UPV)**

## ABSTRACT

Statistical offices are one of the largest producers of statistical information and they face the challenge of disseminating data clearly, providing relevant data to users who are not statistical experts. An alternative to the classic data-table are graphics that try to synthesise the released data.

Aware of this need, Eustat set up a working group to improve visualisation and in particular, the graphics on the Website: our display window. Visualisation can serve as a tool to provide better understanding of the data and to boost the use of our official statistics.

The group established two premises: Ease and interactivity for our users and integration and automation in the production process.

Short-term business and population statistics were prioritised in this revitalisation process due to the impact of their dissemination. The result has been the inclusion and the standardisation of interactive graphics in our press releases.

In a second phase we have incorporated interactive population pyramids where the user can interact with the main elements of the visual display.

Current work is aimed at incorporating interactive maps that allow geographical multivariate analysis with applications to structural statistics.

## 1 OVERVIEW

Official statistics are amongst the main sources of information production and are usually disseminated in the form of attached data tables. The tables are themselves a summary; however, when the volume of information increases, there is a risk that comprehension will decrease. An alternative method of disseminating the information is via the visualisation of the data in the form of graphics. Graphics make data more accessible to end users with a wide range of profiles – irrespective of their statistical background.

Eustat's objectives include the dissemination of statistical information in a clear and comprehensible manner and data visualisation was a weak point within our organisation. Aware of this shortcoming, a working group was set up, which remains in operation, in order to improve visualisation, particularly the graphical elements of our website, our display window. Visualisation can serve as a tool that improves understanding and encourages the use of data afforded by official statistics. In order to accomplish this, it is concentrating research efforts on the visualisation of statistical data in order to make it easier for users to analyse and manage them in a dynamic and interactive manner.

The group established two premises: Ease and interactivity for our users and integration and automation in the production process.

As a starting point, short-term business and population statistics were prioritised within this renovation due to their greater impact and dissemination. Prior to the renovation of graphical elements, each departmental head determined the nature and format of each graphic on an individual basis, in the absence of any standard. A global vision on graphics within the organisation had not been established. Currently, graphics are provided with approximately 15 short-term statistics, normally around two graphics for each statistic, with a similar methodology and appearance.

## **2 PROCESS OF REVITALISING GRAPHIC ELEMENTS**

### *2.1 FIRST PHASE*

The group charged with development is made up of technical officers from different areas (IT, production, methodology and dissemination) within Eustat and researchers from the University of the Basque Country. To date, one of the factors contributing to its success is the diverse composition of the group.

As a first step towards standardisation and revitalisation, the group initially reviewed the current practices within the European statistical system in this area. It was decided to take an approach affording clarity and interaction with the user, with graphics that enable different levels of interactivity, adapted to serve particular needs.

We paid particular attention to large European statistical organisations that have made advances in this area and in open-source experiences. Although initially *Google visualization API* was used, our organisation requires greater control over visual displays, whereby we opted for a custom development of the visualisation tool.

The result of the process has been the standardisation and inclusion of interactive graphics in press releases for all short-term statistics. The graphics included in press releases can be customised by users, affording an interpretation of the data in accordance with their interests and context, enabling selection and adaptation in keeping with their preferences.

## 2.2 VISUALISATION REQUIREMENTS

Graphical formats that summarise a data series are not, a priori, easy to use, sometimes are complex and usually require a learning period. Aware of this, we have attempted to minimise the length of the learning process by standardising graphical formats as much as possible. Once, the functionality of one graphic is fully learned/understood, the remainder prove similar. At present, the main graphics in current press releases follow the same basic structure and design principles:

- A clean and uncluttered appearance, avoiding unnecessary information and adornment whilst remaining appealing. In terms of data, clear scales are employed that can be adapted to each graphic, avoiding the use of too many classes.
- Refrain from using too many labels and headings, and only display them where they are required by the user (when moused over). Flexibility with regards to the editing and positioning of labels and graphical elements as well as combining of different graphic elements (bars, lines, areas, etc). Similar colours are employed in all graphics.
- Given that we are a bilingual Spanish/Basque administration and, furthermore, we present some of our information in English, there is a multilingual approach
- The majority of the graphics present very long data series, whereby we have established a simple method of selecting the time scale. Moreover, as each graphic presents several series of data, the option of activating/deactivating the visualisation of a series is provided.
- Graphics are scalable and can be adapted to suit different browsers, screen resolutions and devices.
- Independence of the data and graphic code. General configuration files for all graphics.
- Respect for the rules of accessibility and computer-based requirements of our organisation.

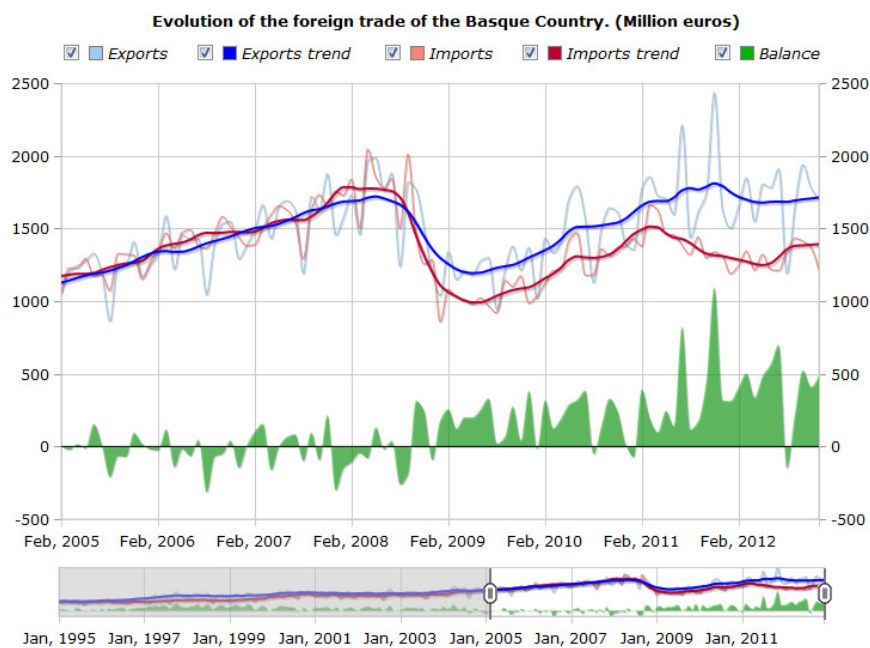
### 2.3 STATISTIC PRODUCER REQUIREMENTS

Short-term statistics entail a very tight dissemination schedule of results and any change is extremely costly in terms of time. To obtain graphics, we have introduced a system that facilitates the current process, wherein all values are predetermined and only the value of the new period requires updating.

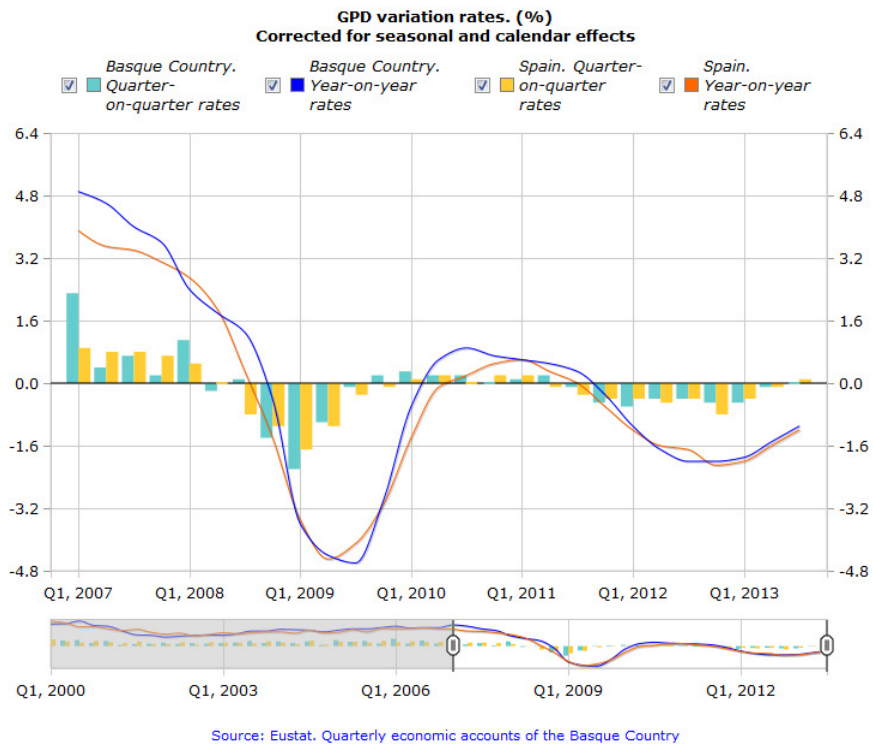
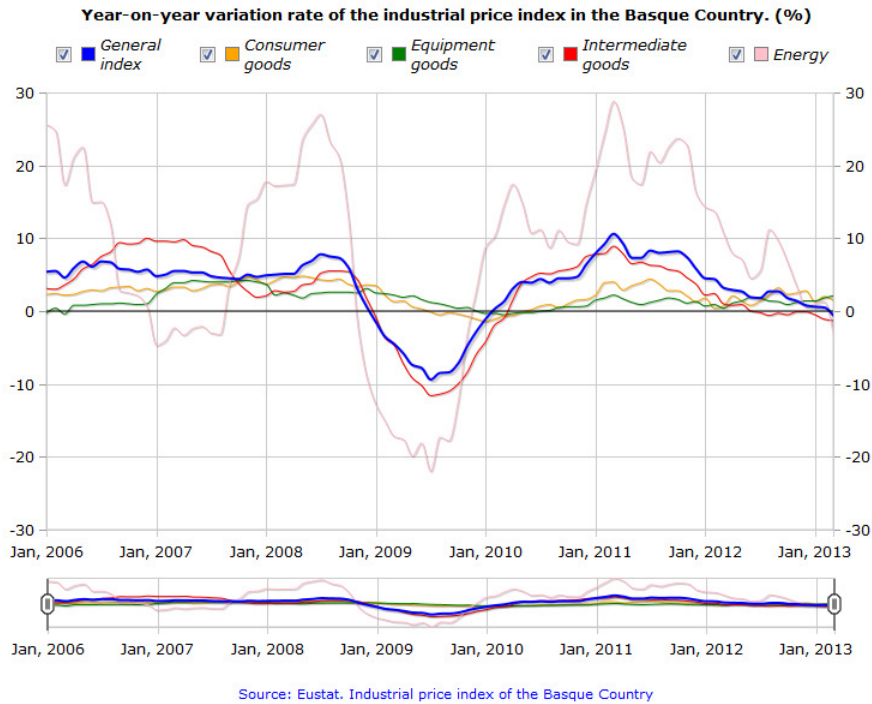
A program was developed within our SAS environment that, for each statistic, generates the necessary files for web publication, reads the data from different sources and produces a shared output with a predefined structure.

The data and the code that produces the graphic remain separate: The files that the producer generates only contain the data and the text element of the graphic are stored in a JSON file, the remainder of the elements are available on the web. The server employs JavaScript public graphic libraries, such as JQuery and others of our own creation. Such architecture for the graphics greatly facilitates the maintenance and the creation of new graphics.

Here we show 3 of the main models/templates for interactive graphics currently employed in press releases for short-term statistics: The first mixes lines and areas, the second include only lines and the third employs lines and bars.



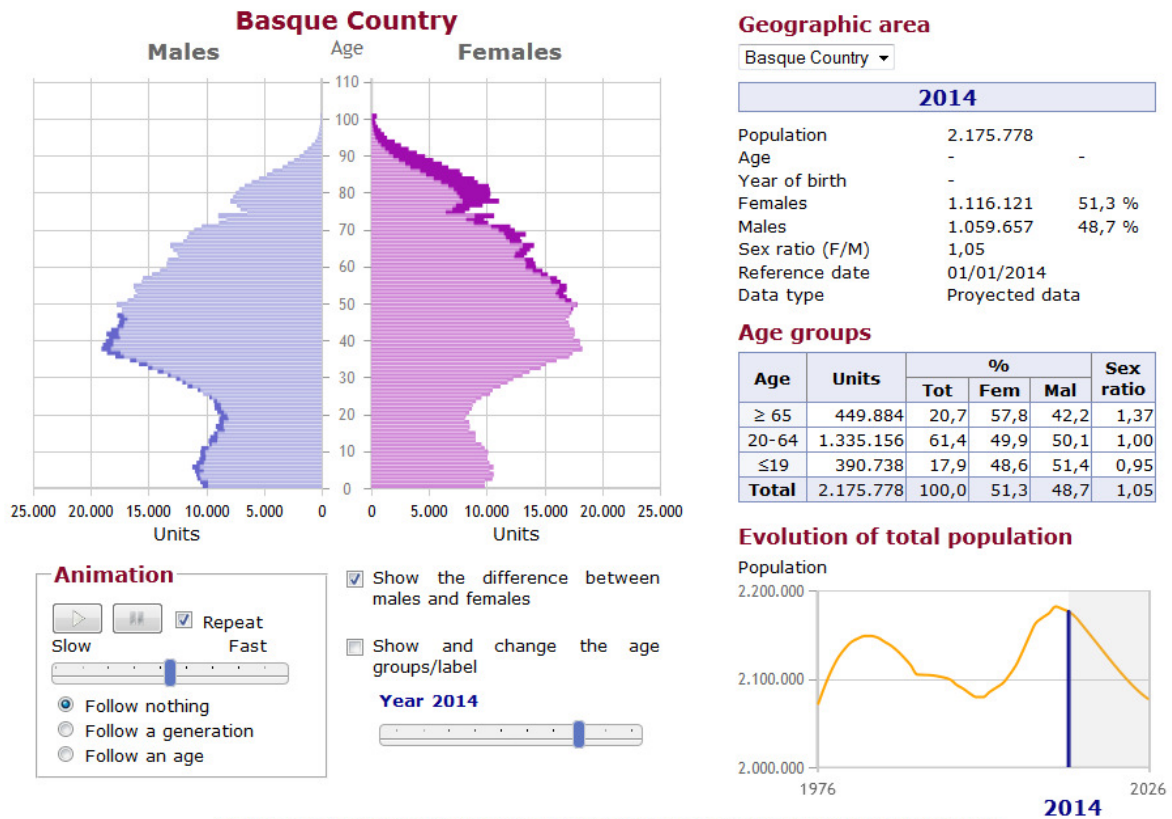
Source: Eustat. Foreign trade statistics of the Basque Country



## 2.4 SECOND PHASE

In the second phase, we incorporated a section for the visualisation of dynamic demographic pyramids wherein the user can interact with the main elements of the visual display. For these

graphics we have gathered the experiences of other institutions, such as INSEE and DESTATIS.



Source: Eustat. Demographic indicators. Municipal inhabitants statistics. Population projections

The display area is divided into 4 elements: The pyramid, the data, the visualisation controls and a support graphic with population data. The data presented on the right-hand side are linked to the graphical interactions and selections. The geographical area of the graphic is an additional parameter within the selection and values are referred to the Basque country, provinces (3) and municipalities.

### 3 RESULTS AND FUTURE STEPS

Initial results suggest that the number of hits recording website access increases when interactive graphics are employed. Both the area of dissemination and the statistical producers are satisfied with the results of this project. In the next *User satisfaction evaluation survey* of Eustat on an annual basis, a question on the graphic media will be added in order to receive feedback. Initial measures of visits and contact with frequent users encourages us to continue

with our efforts to improve graphics. We hope to add new functionalities to the current graphics and proceed with the standardisation of new interactive graphics.

The current work that is being undertaken is aimed at the incorporation of interactive maps, which enable multivariate analysis at geographical level with applications to structural statistics.