Interactive web visualisation of Eurostatistics via R: enhancing the quality of data presentation through storyboarding

Rosa Ruggeri Cannata[1], Piotr Ronkowski[1], Anette Sundstroem[1]

[1]Eurostat, European Commission

Abstract

Eurostat has transformed Eurostatistics from a static PDF format to a Statistics Explained article, complemented by an interactive web visualisation developed in R. The Eurostatistics visualisation uses Flexdashboard, which is built for R Markdown and simplifies the process of publishing groups of data charts and their analysis as a dashboard. The storyboard layout is a powerful feature of Flexdashboard organising information into a sequence of slides. This results in a more comprehensive step-by-step narrative for readers.

The visualisation tool for Eurostatistics will be significantly enhanced through the development of the R package eurostatRtool. This innovative tool is specifically designed to improve the presentation and analysis of statistical data, offering several new features and enhancements that streamline processes and enhance data interactivity. Its ease of use, combined with powerful customisation and efficient update capabilities, makes the eurostatRtool an excellent choice for anyone looking to improve their data visualisation practices.

Keywords: visualisation, dashboard, R

1. Introduction

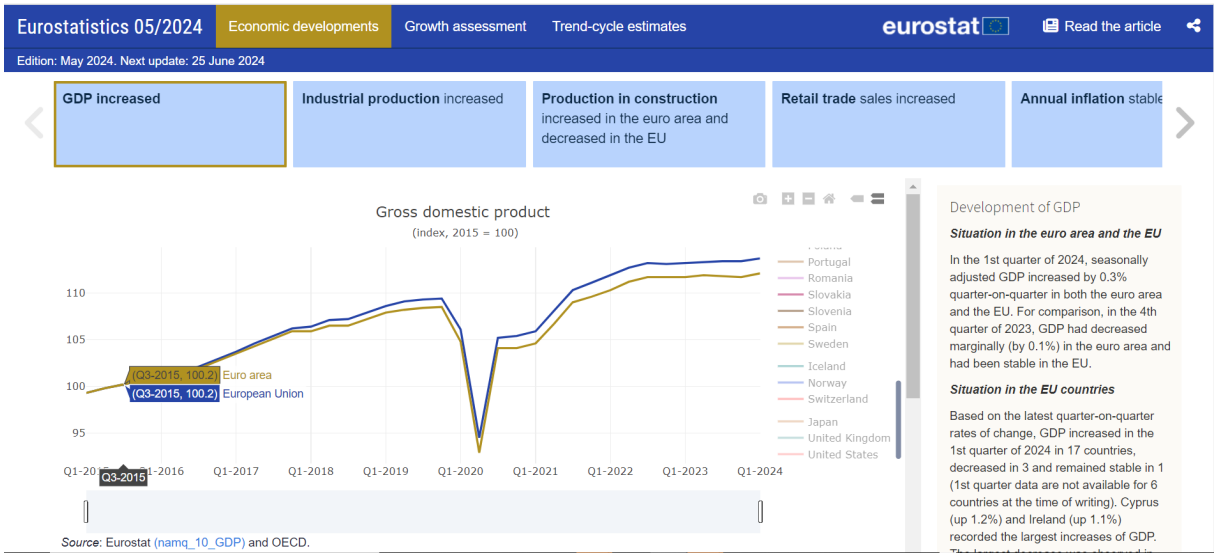
In today's digital landscape, there is an increasing demand to make statistics both more accessible and more meaningful to the public. The increasing demand for dynamic and interactive data presentations calls for innovative solutions that not only inform but also captivate the reader and contribute to improving statistical literacy.

Responding to this need, Eurostat Unit C1 has transformed Eurostatistics from a static PDF format to a Statistics Explained article (Eurostat, 2024), complemented by an interactive web visualisation developed in R, a widely recognised tool for statistical analysis and data visualisation (R Core Team, 2024). This shift from a static to an interactive presentation marks a quality enhancement in how statistical data are communicated. By using the Flexdashboard package, which is built for R Markdown, data started to be presented in a more narrative and engaging manner through its storyboard feature. Additionally, the integration of the Plotly package enriches this experience by adding interactive charts and graphs. These enhancements significantly improve the clarity and impact of the presented data, making them more accessible to a broader audience.

2. Eurostatistics and the interactive visualisation tool

Eurostatistics is a monthly publication released by Eurostat since the 80s. It offers a monthly snapshot of the macroeconomic situation in the European Union, the euro area and individual EU countries, as well as in several other key countries. It includes a wide array of short-term indicators that span multiple domains such as national accounts, business statistics, consumer prices and labour market. These indicators are complemented by a few cyclical coincident indicators, trend cycle decompositions of GDP, industrial production and employment, and a summary of the forecasts from the European Commission (DG ECFIN), the ECB, the IMF and the OECD. The current version of Eurostatistics includes a key feature: an interactive visualisation which is shown in Figure 1, allowing end users to engage with the data in an intuitive way. This visualisation offers a number of possibilities, which stem from the R packages used. Firstly, any user can explore the data in a non sequential way, directly picking what is considered to be more interesting: they could, for example, look at the evolution of GDP and then to inflation, or industrial production, without being concerned with production in construction. Secondly, the user can select the country/ European aggregate of interest, and modify the graphs following personal choice. Thirdly, the user can focus on a specific time interval using the time selector present under the graphs, and read the proposed comments on economic evolution. Finally, the user can download the graphs in a png format, or share Eurostatistics on the social channels. There are further visualisation possibilities not mentioned here, and we invite the reader to have a look at the tool on the Eurostat website.

Figure 1: Interactive visualisation for Eurostatistics



3. R Markdown integrating narrative text with embedded code

We start now to illustrate which packages form the basis of the Eurostatistics visualisation tool. The tool uses R Markdown, a powerful tool that integrates narrative text, written in plain-text markup language, with embedded R code and its outputs, such as charts. This seamlessly generates reports with integrated data, ensuring straightforward updates by simply re-running the R Markdown script. In this way, the production process is particularly rapid and suitable for content which must be frequently delivered.

R Markdown supports multiple output formats, enabling the production of reports in HTML, PDF or MS Word format from the same source document. This versatility makes it an invaluable tool for communicating data with their analyses across different platforms and audiences.

Furthermore, R Markdown simplifies the process of web page creation, requiring no prior knowledge of HTML. Users can focus on the content and data analysis, while R Markdown handles the underlying code necessary to format and present data.

4. Flexdashboard: creating dashboards with minimal coding and maximum impact

Flexdashboard, which is built for R Markdown, simplifies the process of publishing groups of data charts and their analysis as a dashboard or interactive publications.

It generates outputs that are optimised for viewing on both desktop and mobile devices. This ensures that dashboards are device independent, thereby broadening user reach.

Flexdashboard operates entirely on the client-side, meaning it does not require server-side scripting. It gives the possibility of creating and sharing dashboards (or interactive publications) without the need for complex server configurations, making Flexdashboard a practical solution.

One of the key features of Flexdashboard is its choice of predesigned layouts for HTML pages. These layouts allow to quickly assemble and customise dashboards without extensive web development skills. Flexdashboard is therefore an ideal solution for anyone looking to create dashboards with minimal coding and maximum impact.

A standout feature of Flexdashboard is its storyboard layout which strengthens the message presented in the data.

5. Storyboard enhancing data narratives

The storyboard layout is a powerful feature of Flexdashboard organising information into a sequence of slides, although a sequential reading is not an obligation. This results in a more comprehensive narrative for readers.

Its user-friendly navigation allows viewers to move back-and-forth through the data story at their own pace, thereby engaging them in a more interactive exploration of the data story.

Additionally, the storyboard layout supports a variety of interactive elements, such as charts with commentary text integrated into each slide. This complements the general narrative of the storyboard with more in-depth analysis, explaining underlying patterns and insights within the data.

Storyboard and commentary go hand to hand with “frozen” data. It is not possible to comment live data, which could not correspond to data used when writing the related comment. Although this could look like a limitation, monthly data rarely change more than once a month and commenting on data better fit the targets mentioned in the introduction.

6. Plotly for creating professional-quality graphics

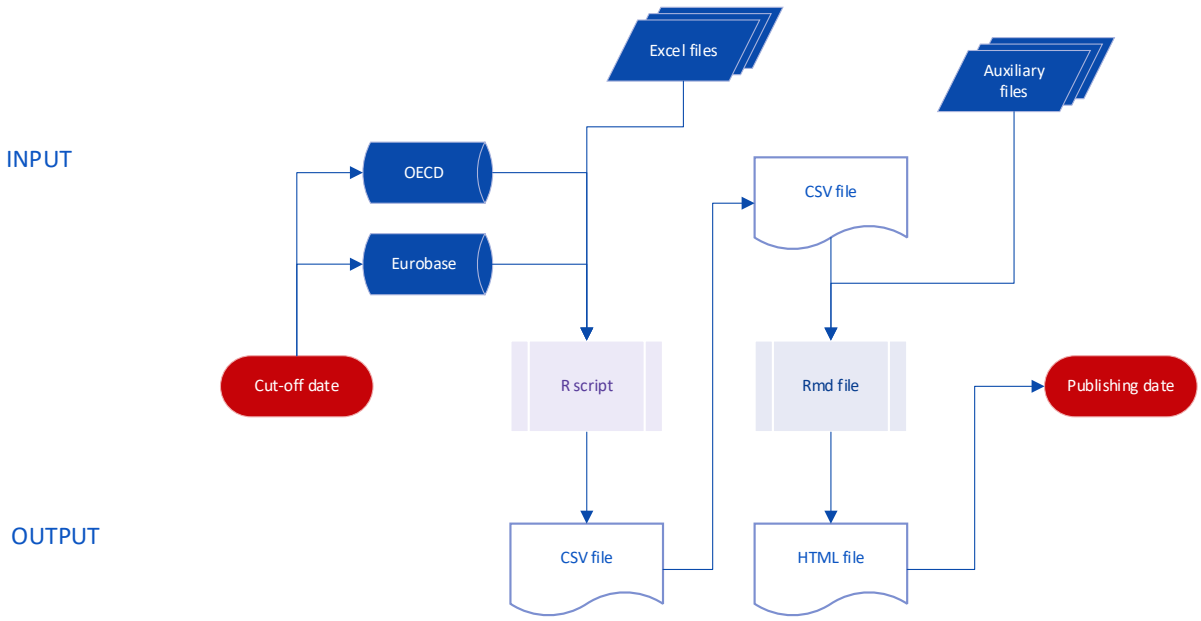
Plotly is a versatile graphing library that the production of interactive and aesthetically appealing charts. The interactivity is particularly useful for web-based analytics, as it allows end-users to engage with the data.

In addition to its core plotting capabilities, Plotly offers a high degree of customisation, allowing tailoring of the visualisations to meet specific aesthetic or branding requirements. From adjusting the layout and style of the charts to modifying the interactivity options, Plotly provides all tools necessary to create professional-quality graphs.

7. Workflow for creating Eurostatistics’ visualisation

The workflow for creating the visualisation outlines a structured process for managing data from initial input to final output. Figure 2 shows a detailed description of each step in the workflow. The production part of the workflow is particularly fast. Most of the production time is spent in controls ensuring the publication is in line with Eurostat’s standards.

Figure 2: Eurostatistics production workflow



We can distinguish two main stages:

- **Data downloading and integration**: Data are downloaded from the online databases of Eurostat and OECD, along with Excel files with data not available via a web service. These diverse data sources are then integrated into a single CSV file, ensuring a unified dataset for further processing.
- **Visualisation creation with R Markdown**: Web pages are generated using R Markdown. This process uses as input data contained in a CSV file plus auxiliary files with the narrative text and a few additional HTML elements. The generated web pages are ready for deployment without the need for further adjustments.

8. Enhancing the tool and preparing the code for sharing

The current version of Eurostatistics has been online for around two years by now, and we decided both to improve it and to share the developed code with other potential users. We contacted R professional developers, asking them first to assess if the chosen packages were fit for our purposes; when they confirmed this was the case, we asked them to improve the code in order to minimise the R knowledge required to deploy the tool. The “tool administrator” will need to configure at the start what input is to be used, the layout of different tabs on the visualisations, the logo, etc. The original script has been simplified to facilitate those steps as much as possible, and the code will be accompanied by documentation helping this configuration process. Once this initial configuration has been performed, updating the visualisation, for example at monthly frequency, will be rapid and simple and could be achieved with basic R knowledge.

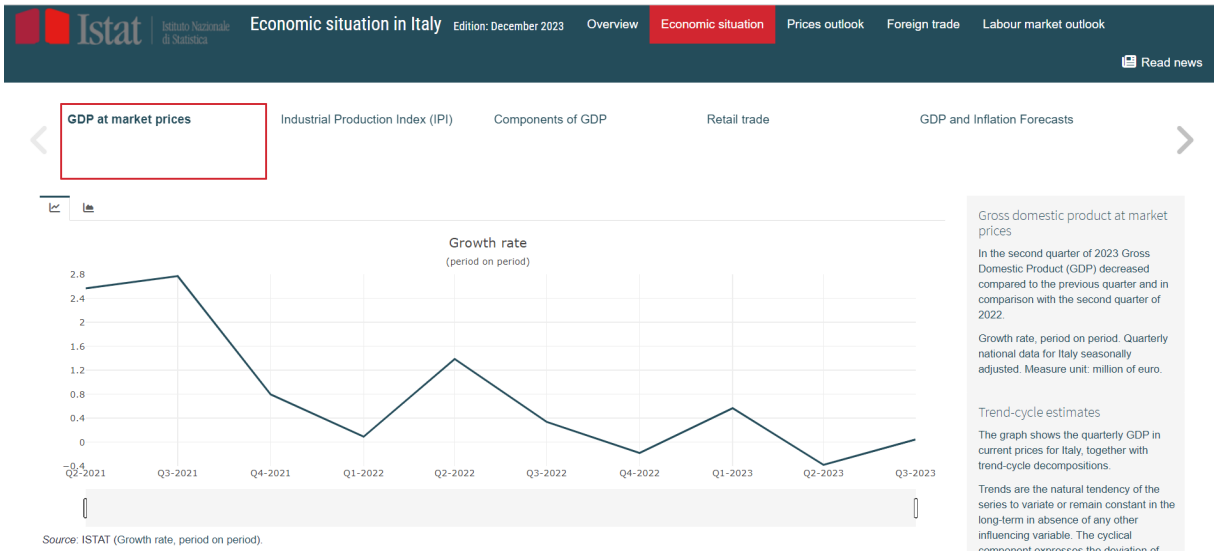
The newly developed package is called eurostatRtool and will significantly enhance the visualisation tool of Eurostatistics. This innovative tool is specifically designed to improve the presentation and analysis of statistical data, offering several new features and enhancements that streamline processes and enhance data interactivity.

The eurostatRtool includes new functionalities for displaying data through interactive maps and comprehensive tables. These features allow for a more nuanced exploration of data, providing clearer insights and improved geographic contextualisation.

The process of constructing content in an input file has been streamlined. This improvement simplifies data manipulation and reduces the time required to prepare data for analysis and presentation, enhancing productivity and user experience.

NSIs and other organisations can customise the layout of their statistical outputs, including options for colours and logos. This capability allows any organisation to tailor the outputs to be aligned with its branding and communication strategies, ensuring a consistent visual style. In Figure 3 you can see an interactive tool generated by eurostatRtool looking at ISTAT website. This was done with the only aim of exploring feasibility aspects and it should not be considered as an official ISTAT output.

Figure 3: A non-official example of output generated via the eurostatRtool



The development of the eurostatRtool is a collaborative effort involving consultations with NSIs in Estonia, Italy, and Romania, who have helped us in improving the available functionalities and in testing. The first version of this package is expected to be available in Summer 2024. It

will be fully available for anybody requiring it; however, please note that this is not an official Eurostat tool for which future maintenance is guaranteed.

9. Conclusions

We wished to share our experience with the use of R for visualisation tools which need a frequent and rapid updating. In this context, we recommend using Flexdashboard built for R Markdown together with Plotly for interactive graphs, based on our positive experience with using them in our visualisation tool for Eurostatistics.

The visualisation tool for Eurostatistics will be significantly enhanced through the development of the R package eurostatRtool. This innovative tool is specifically designed to improve the presentation and analysis of statistical data, offering several new features and improvements that streamline processes and enhance data interactivity. The key features and benefits of the eurostatRtool include:

- **Interactive visualisations**: eurostatRtool is crafted to streamline the development of dynamic and interactive visualisations, empowering the ability to convey data in an intuitive manner. This feature is particularly beneficial for engaging a wide audience and facilitating a deeper understanding of the data presented.
- **Suitable for R beginners**: The package requires only basic knowledge of R, making it accessible to a broader range of potential users.
- **Efficiency in updates**: Once the visualisations are set up and customised, updating them with new data or adjustments is straightforward and rapid. This feature is especially valuable for projects requiring frequent data refreshes, ensuring that all visualisations are current and accurate without extensive manual intervention.
- **Effective storytelling**: The storyboard layout facilitates an effective storytelling approach to data presentation. This layout structures data narratives in a coherent flow, making it easier for readers to follow complex analyses and derive insights.
- **Diverse data presentation capabilities**: The package supports an extensive array of visualisation types, including various charts, maps, and tables. This versatility allows users to select the most appropriate graphical representations for their data, catering to specific analysis needs and audience preferences.
- **Customisation and branding**: The eurostatRtool provides options for customisation that align with organisational branding and visual identity. Users can tailor the appearance of their visualisations to reflect their organization’s style guide.

In summary, its ease of use, combined with powerful customisation and efficient update capabilities, makes the eurostatRtool an excellent choice for anyone looking to improve their data visualisation practices.

References

Eurostat (2024). Eurostatistics – data for short term economic analysis (Statistics Explained). European Commission. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Eurostatistics_-_data_for_short-term_economic_analysis

R Core Team (2024). The R Project for Statistical Computing. The R Foundation. https://www.r-project.org/