

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

UNECE Work Session on the Communication of Statistics
(30 June – 2 July 2010, Paris, France)

- (ii) Improving outreach using Web 2.0, social media and multimedia

Statistics Explained: reaching new audiences

Submitted by Eurostat, European Union¹

I. INTRODUCTION

1. Eurostat, the statistical office of the European Union, has publicly launched Statistics Explained, a new way of publishing statistics, on 28 September 2009. Statistics Explained was designed to solve two major problems of traditional methods using paper or pdf output:

- workflows are long and cumbersome;
- non-specialists cannot find or use the publications easily.

2. To solve the second problem, publications must explain statistics in a way understandable to all and, even more importantly, they should be present in a readily usable form where people want them nowadays: on the first page of their search result, immediately quotable in discussion forums and blogs, easily transferable via Facebook, Twitter or similar social media, as a trustworthy source or reference in Wikipedia.

3. Although Statistics Explained has been open to the public for a relatively short time only, its combination of encyclopaedia, glossary and portal seems to attract an ever wider audience previously unserved by traditional publications. This paper, after briefly introducing Statistics Explained, presents an overview of the many ways in which it is used, by people who may not have known they had a use for statistics too.

II. STATISTICS EXPLAINED: A BRIEF HISTORY

4. Traditional publications on paper have many disadvantages compared to information on the internet: search and navigation possibilities are limited, connecting to other information sources is impossible and, due to long and cumbersome production processes, data are not the latest ones available and usually outdated when finally published. Pdf files cannot really be considered on-line publishing; they remain essentially paper publications rendered electronically with a limited number of added functionalities.

5. Mediawiki software, as exemplified by its use for Wikipedia, offers a possible solution: statistical (or any) information can be published in a stable and controlled way, yet providing all linking and layering possibilities while combining efficiency and flexibility for producers with ease of access, understanding and navigation for users. For a statistical organization, ultimately responsible for the data it releases, one crucial difference with Wikipedia is **control**: data should not be changed by just anyone who considers the inflation rate a bit exaggerated! Mediawiki offers the necessary tools to combine freedom for contributors with sufficient safeguards.

¹ Prepared by Marc Debusschere, marc.debusschere@ec.europa.eu.

6. Statistics Explained (http://epp.eurostat.ec.europa.eu/statistics_explained) was set up as a prototype in the summer of 2008, opened up to contributors (all Eurostat staff, 870 persons) in April 2009 and went on line for the general public, as a beta version, at the end of September 2009. No incident or accident has occurred yet, neither content-wise nor with its operation. The original set of governance rules is still functioning without a change.

7. Statistics Explained is a combination of three elements, each of them essential:
- an online **encyclopaedia** of European statistics, consisting of articles explaining all possible statistical topics in a way understandable to all;
 - a **glossary** briefly defining all terms used in the articles which might not be readily understood;
 - a **portal** to ever more detailed data and metadata, via context-specific and deep links.

Although originally not intended as a channel for releasing new data, a repository of static 'magazine' articles or a 'home page' to a statistical domain, Statistics Explained now performs these functions as well (for unemployment data, articles from Sigma (general-public magazine on European statistics) and the Environmental Data Centres set up jointly with the European Environment Agency, respectively).

8. Some figures on content and use of Statistics Explained (May 2009):
- **content**: some 160 statistical articles, 17 background articles (on methodology and context) and 21 Sigma magazine articles, 900 glossary items, as well as numerous pages for navigation, site management and contributors;
 - **contributors**: about 130 Eurostat staff members (almost 1 in 6 of staff) are logged-in users, plus an additional 6 external ones (contractors) with temporary access at this moment; updating and expanding content continuously: since April 2009 some 27,000 page edits were done (135 per working day, on the average);
 - **viewers**: 1.4 million page views in the 7 months since opening, visitor sessions more than doubled since the start of monitoring in November 2009, to well over 2000 a day; most accessed topics are tourism, national accounts and GDP, external trade, unemployment and environment.

III. THE USES AND USERS OF STATISTICS EXPLAINED

9. The main objective of Statistics Explained is reaching the un-served audience of non-specialists not used to or inclined to visit statistical websites or boldly access on-line databases. They are the general public in need of specific figures, for whatever purpose at hand: study, work, discussing, sharing information with others, or just interest and curiosity. The question, after 7 months: has Statistics Explained succeeded in providing bloggers and chatters, journalists, school children and students, policy makers, and so on ... with the statistical information they can understand and use?

Below are some first results and impressions, mainly based on user statistics produced in real time by a Mediawiki plug-in called 'Piwik'.

10. Some overall conclusions from the 2010 (January-April) user statistics:
- visitors are a mixed bunch indeed, entering from many different countries (mainly European, however) and languages (judging from Google translate results), from general-public sites (including search and 'social') as well as official, media, academic and business ones, using thousands of different search terms and accessing just about every page and file in significant numbers;
 - about half of visitor sessions 'bounce' (leave the site having viewed only one page), but the other half look at 6 Statistics Explained pages on average; this is perhaps explained by visitors reading an article and clicking on glossary links, 'see also' pages or category pages.

A. Search engines

11. People looking for particular information will only find your particular web page if it ends up in the top-10, on the first result page of a relevant Google search. As a rule, Statistics Explained's rankings are very high, even for non-Europe-specific search terms like '*biodiversity statistics*' (result 1 out of 3,060,000), '*ISCED levels*' (3rd of 38,900) or '*mobile subscription*' (10th of 58,000,000) and the trend is still upward, maybe because the number of links to Statistics Explained from high-quality sites is increasing continually.

12. The importance of ranking is demonstrated by the high percentage, 58% of visitors, entering via a search engine - against 33.5% directly, via a bookmark, and 8.5% from a referring site link. The 58% entering via search is actually an underestimation, because the 8.5% referring site links also include many Google links, mainly from Google translate. In the 'search engine' category Google is overwhelmingly dominant, accounting for 96.4%.

13. Apart from the typical keyword searches, two other specific types are surprisingly frequent too:

- **Google Images** accounts for almost 10% of searches, making the Statistics Explained maps, graphs and tables an important additional entry point;
- **Google translate** (not included in 'search engines' but in 'site links') in its many language versions is used very frequently to display Statistics Explained pages in other languages than English; quality is acceptable and gives a fairly accurate idea of the original content; most translations were, in order of importance, into Polish, Spanish, Russian, Lithuanian, French, Slovak and Italian; but also in Arabic, Japanese, Korean and Chinese. Statistics Explained, unlike Wikipedia, is as yet only available in English, but the technical implementation for multiple-language is foreseen for mid-2010.

B. Blogs and discussion forums

14. The overview of referring site links shows that Statistics Explained material is being used regularly in blogs or discussion forums, to make a point supported by statistical data. The clearest example as yet is a discussion on broadband availability in Italy which in 3 days sent a staggering 2000 visitors to the relevant Statistics Explained article.

15. Statistics Explained material is also used as a reference and source by media articles, blogs and forums. Some examples are the Guardian (UK - crime rates), Der Spiegel (GermanyE - 'Greek' problem), Gazeta (Poland – patent applications), BBC on line (UK – external trade), Aftenposten (Norway - social benefits).

C. Social media

16. Although Statistics Explained is represented in the social media, absolute figures are not impressive. It has been 'twittered' and 'friendfeeded' in a limited way (especially after having been 'site of the week' on some sites, see below). Facebook and e-mail sites send through a small but steady trickle of visitors.

17. It is hard to tell, at present, whether the use of Statistics Explained is still below a critical level and will take off in the near future if virtual mouth to mouth marketing gets a chance to make its impact. Or if the social media, because of their nature (leisure, informal contacts) are inherently unsuitable as a medium for transmitting 'serious' statistical information in a significant way.

D. Wikipedia and academia

18. The highest-scoring external referring site is Wikipedia, with some 50 pages of the English-language version linking and sourcing to Statistics Explained; most of those pages are in the domain of national accounts and public finance. The French-language Wikipedia has three pages linking and in the Hungarian and Estonian versions there is one. In addition, a number of academic, library and documentation centre sites, either academic or business-oriented, show links to Statistics Explained. The number of visitors from Wikipedia is relatively high and regular and both the number of referring pages and visitors is increasing at a steady pace. Wikipedia is a promising source of visitors that seems to be only taking off; the potential is high, because numerous Wikipedia articles treat topics for which trustworthy, official and more detailed as well as recent data can be found in Statistics Explained, presented furthermore in a way totally compatible as to look and navigation.

E. Some comments by information professionals

19. **Stempra newsletter** (http://www.stempra.org.uk/newsletter/09_winter/05.htm), "set up to bring together people working in science communication", in its New Year 2010 issue:

"At last... statistics explained

Last but not least, more useful information. The European Union's Statistics (Eurostat) website contains all the stats you will ever need on all things European, but unless you are statistically inclined, or have the patience of a saint, getting to the stats you want can be somewhat challenging.

But now EuroStat has produced a 'wiki' sub-site: Statistics Explained, which is considerably more user-friendly. It consists of a series of themed articles put together by statistical experts and then edited by more communication-orientated people. It might just be the Euro-resource you've been waiting for!"

20. **ResourceShelf**, "A daily newsletter with resources of interest to information professionals, educators and journalists", chose Statistics Explained as its 'Resource of the week' in February 2010 (<http://www.resourceshelf.com/2010/02/15/resource-of-the-week-eurostat-statistics-explained/>) and again in May 2010 (<http://www.resourceshelf.com/2010/05/17/resources-of-the-week-five-niche-databases-and-directories/>):

"(...) As with many sites like this — regardless of geographic location — the reader/searcher/user can be easily overwhelmed by the sheer volume of available information. Which is why we were quite taken with the "Statistics Explained" site, and felt it was worth sharing. (...) It's a straightforward, easier-to-use, gateway to Eurostat. (...) What we particularly like about this page is that it highlights new and especially "meaty" statistical reports. (...) If you click some of the links in the vertical navigation bar on the lefthand side, you'll soon realize that what you're dealing with here is a wiki — which can be edited/updated rapidly. What an interesting way of presenting new information!"

21. **James J. Hill Reference Library Blog** (<http://blog.hillsearch.org/index.cfm/2010/3/16/Business-Website-of-the-Week--Statistics-Explained-your-guide-to-European-Statistics>), "Your source for useful tips and advice on how to find business information", elected Statistics Explained 'Business Website of the Week' in March 2010:

"So often government statistical sites are rich in data but hard to navigate and decipher. No longer can this be said for Eurostat, the official compiler and publisher of just about all things statistical for the European Union, with the release of Statistics Explained.

Using a Wiki-type platform, Statistics Explained makes European data more easily accessible and understandable, with charts and graphs along with any background needed for understanding the numbers, links to related information, and a glossary of terms. (...) It's rare that a website can successfully fill the needs of both novices and expert users, but Statistics Explained may well be one that does. If you do business or have an interest in the European Union, be sure to bookmark it today."

F. Questions for discussion and comment

- Will the use of Statistics Explained or any type of statistical information in the social media like Facebook and Twitter remain marginal?
- Will search results remain high, get even higher or decline as more publishers turn to this kind of integral web publishing?
- Will referencing and sourcing by Wikipedia take off as anticipated? Or will Wikipedia editors just rip the information and integrate it in their article?
- Will bloggers, on-line journalists and forum discussions develop a reflex to come and fetch the instantly usable and transmittable data to support their arguments?
- Will more publishers, of statistics or any other content, start using Mediawiki to publish on line? And if they do, will this create competition for the attention of users?