



**Economic and Social
Council**

Distr.
GENERAL

ECE/CES/GE.30/2008/9
11 July 2008

Original: ENGLISH

ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Group of Experts on Gender Statistics

Fifth Session

Geneva, 6-8 October 2008

Item 3 of the provisional agenda

USAGE OF STATISTICAL DATA FOR
GENDER ANALYSIS AT NATIONAL AND INTERNATIONAL LEVELS

Visual Presentation of Data:

Making data meaningful through effective tables, graphs and maps

Note by the United Nations Economic Commission for Europe (UNECE)

Summary

Gender statistics provide a wealth of information about the differences between women and men, but not everyone can understand these statistics without a little help. Simple visual presentations and good statistical story writing can guide readers into understanding data and grasping the relevance of the information being presented. In this paper, some general principles and practical tips on effective visual presentation of data and graphics are presented, along with suggested further readings.

I. INTRODUCTION

1. Tables, graphs and maps have been used to visualize data for hundreds of years. An invaluable skill for anyone working with statistics is to create a visual representation that accurately communicates the data's underlying message. Gender statistics provide a wealth of information about the differences between women and men, but not everyone can understand these statistics without a little help. Simple visual presentations and good statistical story writing can guide readers into understanding data and grasping the relevance of the information being presented.

II. GOOD VISUAL PRESENTATION - WHAT WORKS AND WHY

2. The most important principle of good visual presentation is simplicity. Advancements in computing enable almost anyone to create a colourful graph or table with a few clicks of the mouse, but there are some general rules of good design that should be followed. To put this in context, some understanding of the human ability to perceive visualizations is useful.

3. People are good at recognizing patterns, shapes and colours. Research into human thinking suggests we can absorb between five and nine pieces of information at a time (Kennedy, 2007), so visual presentations should be limited in what they attempt to illustrate. They should be clear and simple, even when they are targeted at audiences with statistical expertise. Design and visualization guru, Edward Tufte (1983:13) explains that excellent graphics should "show the data" and encourage the reader to think about the substance of the message rather than being distracted by the design.

4. Colour is a powerful way to communicate messages, but less can be best. People associate meaning with colours, such as red for heat or blue for cold, and using these appropriately can maximise the impact of a visual presentation. The human eye can effectively distinguish between five shades of grey, or five shades of any colour (Bessler, 2004); useful information if choosing a monochromatic colour scheme. The most common form of colour blindness results in the inability to distinguish red from green; a good reason to avoid the often used traffic light colours, red, yellow and green (ibid). Often the most effective colour combinations are black, grey and white, offering high-contrast and being suitable for both screen and print.

III. A MODEL FOR MAKING DATA MEANINGFUL

5. Visual presentations are a way to complement textual analysis to communicate the story or message being told by the data. A statistical story is one that does more than just describe the data. They should describe and illustrate a message and make it clear why the information is relevant and important. "Statistical stories should grab a user's attention, invoke thought, be informative and ideally be entertaining (Kennedy, 2007:4).

6. A group of dissemination and communication experts from national statistical offices recommend how to do this in *Making Data Meaningful: a guide to writing stories for numbers*, published by the UNECE in 2006. The guide provides practical tips to "help managers,

statisticians and media relations officers to use text, tables, graphics and other information to bring statistics to life using effective writing techniques” (UNECE, 2006:iv).

7. There are a range of skills needed to make statistical data meaningful to a reader. They include being able to analyse statistical data, understand the audience, develop the message or story, draft the text, create a corresponding graphic and publish the final product. Visual presentations are usually just a part of an overall package. As this paper concentrates on the visual presentations only, readers are encouraged to review the tips in *Making Data Meaningful* (www.unece.org/stats/documents/writing) for more information on effective writing techniques.

A. Identify the audience

8. When people look at statistical information, one of the first things they will want to know is how it is relevant to them. Producers can add value to the statistics they publish by clearly explaining the relevance to a reader’s area of work or interest. This may be particularly important for gender statistics, being an emerging area that readers might be unfamiliar with or needs more explanation to be properly understood. To be able to explain relevance, the target audience must be identified.

9. The level of statistical knowledge or understanding, often called statistical literacy, will depend on the audience. For example, information targeted at the general public is often expressed differently to that aimed at experts in economics or gender studies. Wherever possible, jargon and acronyms should be avoided in headings and labels. Even terminology that is widely used and accepted, such as ‘gender pay gap’ or ‘gross domestic product’, may need to be explained or expressed differently so particular audiences can understand it.

10. Before preparing any visual presentation, the first step is to decide on the intended audience and understand their information needs. Only then is the presentation of data and graphics is likely to be relevant and appropriate for the reader.

B. Formulate the message

11. To engage with the reader, data and graphics must convey a clear message. Gender statistics provides a wealth of information about the differences between men and women, but not everyone can understand these statistics by themselves. Simple visual presentations and good statistical story writing can help readers realise the relevance of this information.

12. Visual presentations should be clear and simple, conveying one finding or a single concept (Godin, 2008). Graphs that try to show too much can become complicated and difficult to understand. The presentation should be simple, using few or monochromatic (different shades of the same colour) colour schemes. Journalistic techniques can be employed to grab the reader’s attention with a catchy heading. Any accompanying text should be written in plain language, related to a theme or topic rather than simply describing the data.

C. Choose the visualization technique

13. The type of graph or presentation to choose will depend on the data and the message. Tables present data values in a structured format, a line graph is used to show trends over time,

pie graphs show relative shares of a whole, bar graphs show differences in categorical data and maps are for illustrating differences or similarities across geographical areas. Gender statistics show the relationship between two sets of data, that for women and that for men. Suitable chart formats include horizontally grouped bar charts (different coloured bars for women and men), back-to-back bar charts (eg population pyramids) and scatter diagrams (ONS, 2006:102).

14. Visualizations help people to digest statistical information and to remember the message that is being communicated (ABS, 2007:4). Poorly prepared graphics can mislead or confuse the reader and so it is important to follow some basic guidelines:

- (a) for graphs, avoid using three-dimensional formats as they often distort the data, the Y axis should start at zero in order not to exaggerate trends and there should usually be only one unit of measurement per graph.
- (b) tables should be sorted in an appropriate order (e.g. highest to lowest), one decimal place is usually enough and numbers should be right justified to clarify their value.
- (c) maps should include a title and a legend that explains the statistical units, the date the data was collected or produced and the geographic area being presented.

D. Monitor the impact

15. Agencies can improve their visual presentations of data by collecting information on what has worked in the past, or what may have proven misleading. The impact of different presentation formats might be understood by monitoring the breadth of media coverage about particular releases and checking whether the associated graphic is reproduced directly, or changed. Analysing website usage, such as downloads of publications or links from external sites is also a useful way to learn about the popularity of certain information. To learn about which visual presentations work best, agencies could ask customers and media representatives through surveys and informal interviews or conduct tests to find out how people interpret particular presentations and whether it matches the intention.

IV. EXAMPLES

16. There are an abundance of examples that statistical agencies can draw upon to inspire good visual presentations. They include looking at products of other national statistical offices, particularly those that have dedicated visualization / graphic design teams or are well-resourced and trained in the area of visual presentation. Other good examples include graphics presented in reputable books, newspapers, magazines and professional journals. To remain viable, commercial mass media has to communicate simple and engaging messages in a limited word-space, making them a potential source of good examples, but beware; they may also be a source of bad examples!

17. Some examples of good graphics are provided in Annex 1, but examples to illustrate the value in following these guidelines will be given in the presentation at the October meeting.

V. EMERGING VISUALIZATION TECHNIQUES

18. With the advent of the Internet and growth of new Web 2.0 technologies, users can interact directly with data and create their own visualizations. Many statistical offices now provide access to statistical databases through their website, allowing users to query and download statistical information. This is often complemented by a suite of visualization tools enabling users to create graphs, tables or maps on-line, without having to download the data and work in another application. There may be concerns about the consequences of giving this level of control over to the user. There is the possibility that they will create nonsense graphs or inappropriate correlations, however, it is surely preferable that people are accessing and using the data. Offering support, monitoring and correcting any misuse are ways to minimize potential problems.

20. Animation and video are two important emerging data visualization techniques. Hans Rosling, co-founder of Gapminder (www.gapminder.org), a tool that uses animation to illustrate data, has demonstrated great success with this approach to communicating statistics. Rosling has achieved a massive audience through on-line video, an increasingly popular feature of the Internet. His presentation at TED Conferences in 2006 has been viewed on-line and downloaded thousands of times and he has built upon the popularity of this medium for communicating statistics by developing gapcasts, short video lectures on issues such as maternal mortality, globalization, energy and human development trends (<http://www.gapminder.org/video/gap-cast/>).

21. Websites such as Many Eyes (<http://services.alphaworks.ibm.com/manyeyes/home>), Swivel (www.swivel.com) and Data Place (<http://www.dataplace.org/charttable/>) are adding a new dimension to visual presentations by building on-line communities around data visualization and sharing. The sites allow their users to upload data sets and create graphs for sharing and discussion with other users. Some producers of statistics, including UNECE (<http://www.swivel.com/users/show/1005968>), are experimenting with the potential of these web services for communicating with a wider audience. Although success has been limited to date, these on-line communities are certainly an emerging area of data visualization worth following.

VI. RESOURCES AND FURTHER READING

22. There is a vast range of reading material available on this topic. For those wishing to explore in more detail, a list of recommended reading and resources is provided below.

Books

- (a) *The Visual Display of Quantitative Information*, Edward R. Tufte, 1983
- (b) *Show Me the Numbers*, Stephen Few, 2004
- (c) *The Tiger that Isn't: Seeing through a world of numbers*, Blastland, M. and Dilnot, A., 2007

Websites and on-line resources

- (d) *Making Data Meaningful: A guide to writing stories about numbers*, UNECE, 2006
www.unece.org/stats/documents/writing

- (e) *Research Paper: Data Visualization*, Damien Kennedy, Australian Bureau of Statistics, 2007
<http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1211.0.55.001Jul%202007>
- (f) *Sparklines: Intense, Simple, Word-Sized Graphics*, Edward R. Tufte, 2006
http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001OR&topic_id=1
- (g) *More or Less*, British Broadcasting Corporation (BBC)
http://news.bbc.co.uk/2/hi/programmes/more_or_less/
- (h) Weblogs about data visualization:
 - Flowing Data <http://flowingdata.com/>
 - Junk Charts <http://junkcharts.typepad.com/>
 - Information Aesthetics <http://infosthetics.com/>
 - Many Eyes Blog <http://manyeyes.alphaworks.ibm.com/blog/>
 - Tasty Data Goodies <http://blog.swivel.com/>
 - Well-formed Data <http://well-formed-data.net/>
 - Pictures of Numbers <http://www.numberpix.com/>

23. The UNECE is currently working with its group of experts on statistical dissemination and communication to develop a style guide on the visualization of statistical data. This short practical guide will provide statistical offices with advice on creating good visual presentations. It will form part of the Making Data Meaningful series and is planned for print and on-line publication in early 2009.

References

- Bessler, L., (2004), *Communication Effective Use of Color for Web Pages, Graphs, Tables, Maps, Text and Print*, SUGI Poster Section, SAS Institute Inc. Available on-line at <http://www2.sas.com/proceedings/sugi29/176-29.pdf> accessed 15 July 2008.
- Godin, S., (2008), *The three laws of great graphs*, available on-line at http://sethgodin.typepad.com/seths_blog/2008/07/the-three-laws.html, accessed 14 July 2008.
- Kennedy, D., (2007), *Research Paper: Data Visualization, Australia*, Australian Bureau of Statistics.
- Office for National Statistics, United Kingdom, (2006), *ONS House Style and Author's Guide*, 2nd Edition, Office of Public Sector Information, Norwich.
- TED Conferences (2006), *Hans Rosling: Debunking third-world myths with the best stats you've ever seen*, available on-line at: http://www.ted.com/index.php/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html, accessed 15 July 2008.
- TED Conferences (2008), *Announcing the Top 10 TEDTalks*, available on-line at <http://www.ted.com/index.php/talks/top10>, accessed 15 July 2008.
- Tufte, E.R., (1983), *The Visual Display of Quantitative Information*, Graphics Press.
- United Nations Economic Commission for Europe (UNECE) (2006), *Making Data Meaningful: A guide to writing stories about numbers*, United Nations.

ANNEX 1: EXAMPLES OF GOOD PRACTICE

Table 1

Tertiary teachers: by sex and academic year, France

Data format: percent (%)

	Academic year						
	1999 -2000	2000 -2001	2001 -2002	2002 -2003	2003 -2004	2004 -2005	2005 -2006
Women	33.0	45.7	33.9	37.8	39.3	38.0	37.1
Men	67.0	54.3	66.1	62.2	60.7	62.0	62.9
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Definition: Teachers are classified according to levels of ISCED 1997. Primary corresponds to level 1, secondary to levels 2, 3 and 4; tertiary to levels 5 and 6.

Source: UNECE Statistical Database, compiled from national and international (EUROSTAT and UNESCO Institute for Statistics) official sources.

What makes this a good table?

- Clear and descriptive heading and labels, including units of measurement
- Values are right aligned
- Definitions and source are indicated
- Although it may seem redundant, including the total clarifies that all relevant dimensions are being presented

Table 2 (below) provides a more complex example. Note the following good points:

- Sub-headings and totals are in bold
- Shading is used to assist with readability
- Data is in hours and minutes, rather than decimals, to aid understanding
- Leading zeros have been removed to reduce clutter
- Footnotes provide definitions of ambiguous terms

Table 2**Use of free time¹: by activity, country and sex, 2003²**

Data format: hours : minutes

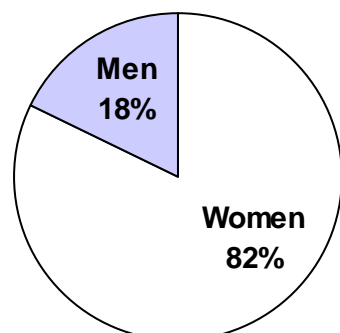
Activity	Latvia		Lithuania		Spain		United States	
	Women	Men	Women	Men	Women	Men	Women	Men
TV and video	1 : 55	2 : 18	1 : 58	2 : 36	1 : 46	2 : 00	2 : 15	2 : 42
Socializing ³	34	34	31	31	49	57	54	45
Reading ⁴	30	25	22	22	12	16	21	16
Sport ⁵	19	31	13	21	37	52	11	19
Hobbies and games ⁶	4	10	4	12	7	21	18	21
Volunteer work and help ⁷	10	10	12	15	13	7	22	17
Other free time activities	31	34	22	28	40	40	36	42
Total free time activities	4 : 03	5 : 02	3 : 45	4 : 46	4 : 25	5 : 16	5 : 00	5 : 24

- 1 Free time includes all other kinds of activities, e.g, volunteer work and meetings, helping other households, socializing and entertainment, sports and outdoor activities, hobbies and games, reading, watching TV, resting or doing nothing.
- 2 Time use represents the average time spent on an activity per day (hours and minutes per day). All persons are included, whether they have performed this activity or not, and all days of the week, as well as working and holiday periods are included.
- 3 Socializing includes visiting and receiving visitors, feasts, phone calls and conversation with family or friends, etc.
- 4 Reading includes reading periodicals, brochures, books, etc.
- 5 Sport includes walking and hiking, all kinds of sports (jogging, ball games, water sports, etc.) productive exercises (hunting, fishing, picking berries, etc.) and sport related activities.
- 6 Hobbies and games – collecting, correspondence, solo games and play, parlour games and play, gambling, etc. Includes also visual, performing and literary arts.
- 7 Volunteer work and help – includes work for/through an organization and informal help to other households.

Source: UNECE Statistical Database (www.unece.org/stats/data)

Graph 1

Part-time workers in France are mostly women, 2006



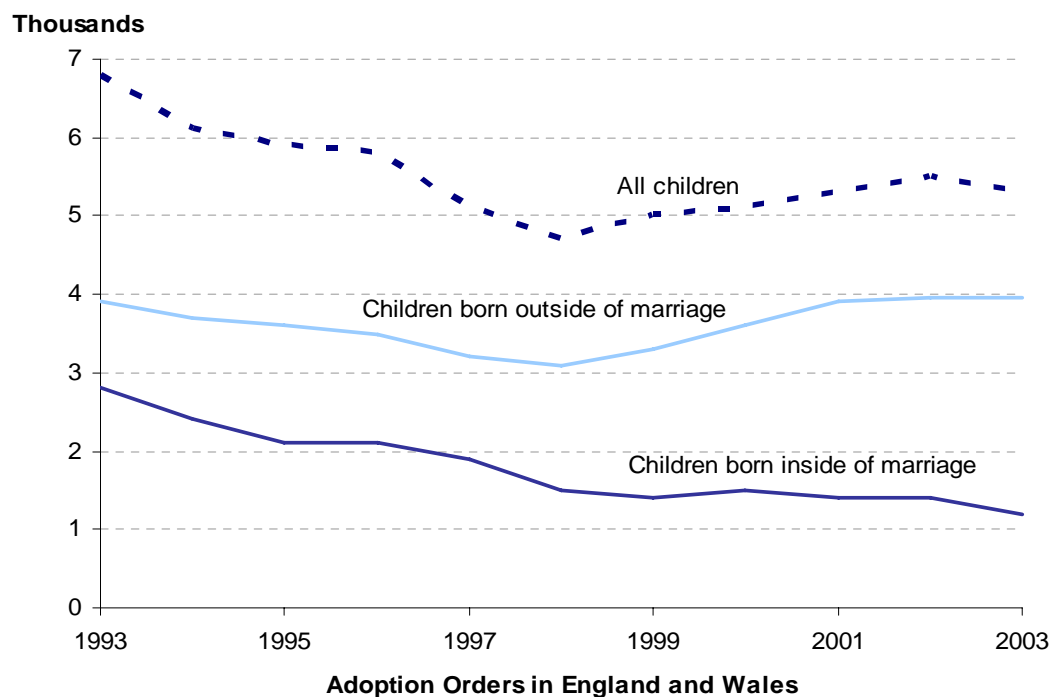
Source: UNECE Statistical Database, compiled from national and international (EUROSTAT) official sources.

What makes these good graphs?

- Simple presentation and colour schemes
- Single message
- Eye-catching, news-style headings
- Using labels instead of a legend make them easier to read and interpret
- Sources are clearly indicated

Graph 2

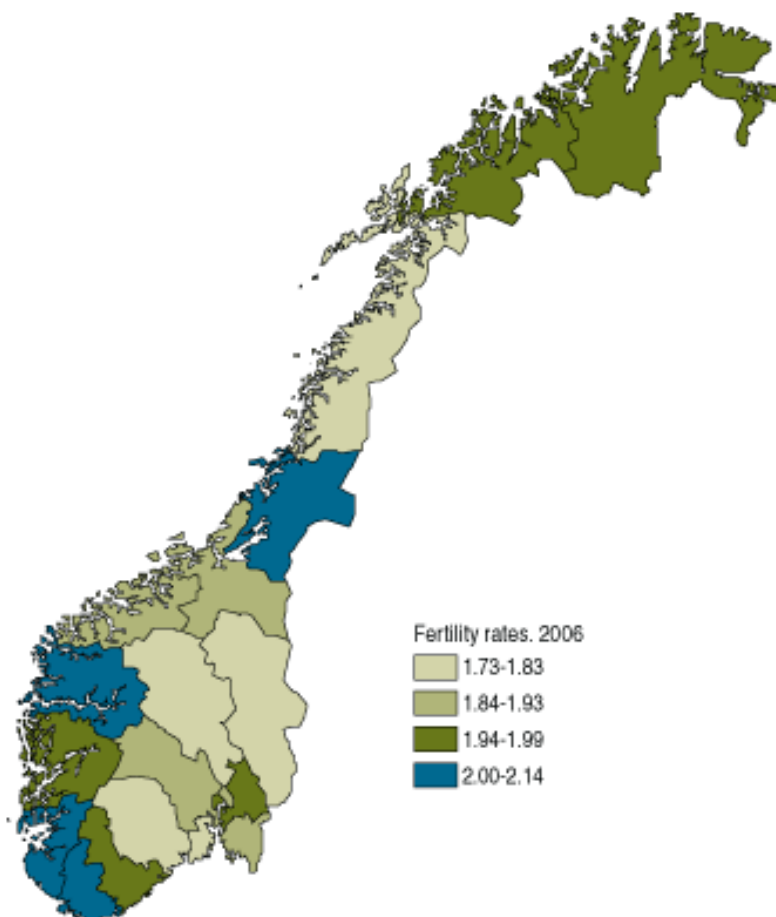
Adoptions fall by 2.4% in 2003



Source: United Kingdom Office of National Statistics. Available online at <http://www.statistics.gov.uk/cgi/nugget.asp?ID=592> [accessed 28 September 2005]. Reproduced from Making Data Meaningful, UNECE, 2006.

Map 1

Norway: Fertility rates by county, 2006



Source: Statistics Norway. Map data: Norwegian Mapping Authority.

More information: http://www.ssb.no/fodte_en/

Available online at http://www.ssb.no/english/subjects/00/minifakta_en/en/main_03.html [accessed 22 July 2008].

What makes this a good map?

- Uncomplicated colour scheme
- There is a link to more information on the topic, which takes the user to a summary of statistics on fertility in Norway
- Single message
- Clear and descriptive heading
- Source is indicated