UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

Making Data Meaningful

Part 2

A guide to presenting statistics

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

Making Data Meaningful

Part 2:

A guide to presenting statistics

UNITED NATIONS
Geneva, 2009

NOTE

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontier or boundaries.

ECE/CES/STAT/NONE/2009/3

Table of Contents

Introduction v
1. Getting the message across 1
2. Visualization of statistics 7
3. Tables 12
4. Charts 17
5. Maps 30
6. Emerging visualization techniques 41
7. Accessibility issues 46
8. References and further reading 51

Introduction

The *Making Data Meaningful* guides have been prepared within the framework of the United Nations Economic Commission for Europe (UNECE) Work Sessions on the Communication and Dissemination of Statistics[1], under the programme of work of the Conference of European Statisticians[2].

These guides are intended as a practical tool to help managers, statisticians and media relations officers in statistical organizations, particularly those organizations that are in the process of developing their communication strategies. The guides provide advice on the use of text, tables, charts, maps and other devices to bring statistics to life for non-statisticians. They contain suggestions, guidelines and examples - but not strict rules or rigid templates.

An effective data release uses a combination of text, tables and graphics to maximize its strength in conveying various types of information. *Making Data Meaningful Part 1: A guide to writing stories about numbers* (issued in 2006) focused on the use of effective writing techniques. *Making Data Meaningful Part 2: A guide to presenting statistics* aims to help readers find the best way to get their message across to non-specialists, using the most suitable set of tools and skills now available from a dazzling array of communication methods.

This guide recognizes that there are many practical and cultural differences among statistical organizations and that approaches may vary from country to country.

A group of experts in the communication and dissemination of statistics prepared this guide. They are (in alphabetical order):

- Petteri Baer, UNECE
- Colleen Blessing, United States Energy Information Administration
- Eileen Capponi, Organisation for Economic Co-operation and Development
- Jerôme Cukier, Organisation for Economic Co-operation and Development
- Kerrie Duff, Australian Bureau of Statistics
- John Flanders, Statistics Canada
- Colleen Flannery, United States Census Bureau
- Jessica Gardner, UNECE
- Martine Grenier, Statistics Canada
- Armin Grossenbacher, Swiss Federal Statistical Office
- David Marder, United Kingdom Office for National Statistics
- Kenneth Meyer, United States Census Bureau
- Terri Mitton, Organisation for Economic Co-operation and Development
- Eric St. John, Statistics Canada
- Thomas Schulz, Swiss Federal Statistical Office
- Anne-Christine Wanders, UNECE

The contributions of Martin Lachance (Statistics Canada), Andrew Mair (Australian Bureau of Statistics), Alan Smith (United Kingdom Office for National Statistics) and Steven Vale (UNECE) are greatly appreciated.

[1] Information about the UNECE Work Sessions on the Communication and Dissemination of Statistics are available from the UNECE website at http://www.unece.org/stats/archive/04.05.e.htm.
[2] Information about the Conference of European Statisticians is available from the UNECE website at http://www.unece.org/stats/archive/act.00.e.htm.

1. Getting the message across

1.1 The written word

News releases are often the vehicle through which your statistical organization communicates key findings of its statistical and analytical programmes to the intended audience, which is most probably the general public. The text is the principal vehicle for explaining the findings, outlining trends and providing contextual information.

In this chapter, we will provide many suggestions for preparing an "effective" news release or other document, such as a report or an analytical article.

What makes a news release, report or analytical article effective? Perhaps the best explanation comes from the first *Making Data Meaningful* guide, *Part 1: A guide to writing stories about numbers*, which provides an initial set of recommendations for getting your message across. An effective news release is one that:

- tells a story about the data;
- has relevance for the public and answers the question "Why should my audience want to read about this?";
- catches the reader's attention quickly with a headline or image;
- is easily understood, interesting and often entertaining;
- encourages others, including the media, to use statistics appropriately to add impact to what they are communicating.

Here are some tips to help you get started on your text.

1.2 Target audience: your first decision

The first important decision you must make is to pinpoint an audience: who are you writing for? Quite simply, the audience is in the driver's seat. By and large, what the audience wants is what you should be giving them. You have to listen to your audiences to find and select the right narratives, language, and visual and graphic devices that will capture their attention.

The choice of an audience is more complex these days because of the Internet. Most statistical organizations have a mandate to communicate to the general public, who are non-specialized, fairly well-educated laypeople. In the days of printed news releases, the principal target audience was likely to be the media, on which organizations relied to transmit key findings to the public.

Nowadays, however, statistical organizations have developed a significant direct readership through their websites, e-mail and other forms of Internet-based distribution. This means that they are communicating with a host of audiences simultaneously: the public, data users, bankers, financial analysts, university professors, students and so on, each with their own data requirements.

The communications world is constantly evolving. Successful commercial media know this and constantly monitor - often in real time - which of their stories get the most attention. They then target their resources to create richer content by using devices such as video, additional photos or more analysis to encourage greater interaction with each audience.

In any case, the message here is that before throwing precious resources into any method of communication - new or established - it is important to decide first who your audiences or stakeholders are, what it is they want from you and how they want it.

Should you wish to address several audiences, you must select the most appropriate method to reach each of them, by transmitting your message through appropriate channels and using appropriate communication techniques.

Often though, lack of time and resources mean that it is not possible to reach all of your audiences all of the time. You have a choice: you can prioritize or, if you want to reach the widest audience, you can find the clearest common ground.

This is what many statistical organizations do. They target the general public, but make a concerted effort to reach this audience by using journalists as a kind of 'conduit'. The intended audience is the public, but journalists are the means of communicating with that audience. Experts and specialists also benefit from this approach. Often, the simple clear techniques used to reach a wide audience are warmly welcomed by even the most specialized audience.

1.3 Understand the context in which you are communicating

Statistical communication does not occur in isolation. Therefore, it is important that you understand the context in which you are communicating. The way in which audiences consume media is constantly changing. There are also distinct differences between generations, in their technical abilities and understanding of statistics.

When planning statistical communication, you should keep in mind four particular trends in online media consumption, which represent both opportunities and risks:

1. The World Wide Web is increasingly becoming a medium for entertainment. Any message that is not presented in an interesting way risks not engaging with younger audiences.
2. Society has developed a “snack culture” in relation to information consumption. Audiences increasingly want smaller snippets of information that can be consumed quickly.
3. Audiences using the Internet tend to “satisfice”: they find a vaguely relevant piece of information and stop there, rather than look further for the most relevant piece of information.
4. In addressing different audiences and presentation styles, try not to exclude important audiences in the process of making your statistical communication more entertaining or easier to consume.

So what can you do to make the best use of the Internet? You must use the most appropriate tools of language, structure and presentation to get your message across. The following sections will illustrate how.

1.4 Narrative: telling the story

First and foremost, find a story

For data to be meaningful to a general audience, it is important to find meaning in the numbers. The word "story" often alarms people in the statistical/scientific world, because it has overtones of fiction or embellishment that might lead to misinterpretation of the data. This view might be justified if analysts do not approach the data with care and respect.

However, the alternative, i.e. avoiding a story, may be far worse. People often distrust statistics and feel they are misleading, because they cannot understand the data. This occurs because we, the people who produce data, fail to make them relevant and explain them in terms that people can understand. Without a story line, a release becomes just a simple description of numbers.

A statistical story must be based on sufficient knowledge of the data and the phenomenon under study. Otherwise, it may be interesting, but in fact all wrong. When preparing a statistical story, you must also remember the Fundamental Principles of Official Statistics[3]:

1. Impartiality;
2. Professionalism;
3. Metadata;
4. Comment on erroneous interpretation;
5. Diverse sources;
6. Confidentiality;
7. Transparency;
8. National coordination;
9. International standards;
10. International cooperation.

It is vital that statistical organizations remain impartial and ensure the confidentiality of respondents and small sub-populations.

Your text should place the most important and significant findings in the context of short- and longer-term trends. It should explore relationships, causes and effects, to the extent that they can be supported by evidence. It should show readers the significance of the most current information.

Write in journalistic style

Use the writing style adopted by journalists: the "inverted pyramid". Present the most important facts first, followed by subsidiary points in decreasing order of importance. Readers lose interest quickly, so the most critical information must be at the beginning of the text.

[3] These principles were adopted by the United Nations Statistical Commission in 1994. They are described in detail on the UNECE website at http://www.unece.org/stats/archive/docs.fp.e.htm.

Avoid starting your text with methodology and ending it with a conclusion. You can put key points regarding methodology in a note to readers - the less complex the explanation of methodology, the better. The conclusion should become your lead or opening paragraph.

The lead is the most important element of your text. It should tell a story about the data. It summarizes the story line concisely, clearly and simply, and sets the story in context. It should concentrate on one message or theme and contain a minimum of data.

GOOD EXAMPLE of a lead paragraph:

> Net profits of non-financial companies in the Netherlands amounted to 19 billion euros in the second quarter of 2008. This is the lowest level for three years. Profits were 11 percent lower than in the second quarter of 2007. The drop in net profits is the result of two main factors: higher interest costs - the companies paid more net interest - and lower profits of foreign subsidiaries.
>
> Source: Statistics Netherlands

Do not burden the reader with too many numbers in the body of the text and use only key rounded figures. Less important numbers should be relegated to accompanying tables. Use the text to present analysis, trends and context, not to repeat values in the tables.

Pay attention to structure

Structure your text so that each component makes sense on its own and also contributes to the overall story you are telling. Subheadings are an effective tool for strengthening the organization of a release. They break it into manageable and meaningful sections.

A concise subheading summarizes the main finding in the subsection. It may be more engaging and understandable when it contains a verb.

GOOD EXAMPLES of subheadings:

> "Inventory levels ease slightly"
>
> "Growth in energy products leads the rise in imports"

For Internet-based communications, each subsection should make sense on its own, which means that terms should be spelled out and sources should be noted. Search engines tend to drive users to deep links within websites, rather than to the home page or other gateways you have created to channel visitors to their desired destinations.

Your messages should also be layered so they cater to the different information needs of your audiences. Start each subsection with a topic sentence that states clearly the main finding in the subsection. You can elaborate on this finding in subsequent paragraphs.

This structure will provide a clear pathway into more detailed data, analysis and technical information for audiences that are interested in such detail. Also, in electronic publications, you can always use links to take your readers to more complex analysis.

1.5 Language: keep it clear, concise and simple

Simple language is at the heart of any successful communication. The old adage "keep it short and simple" (KISS) is as true now as it has ever been. This does not apply just to your text. It is also relevant for tables, visuals and graphics where, with so many opportunities to be clever, the temptation is to be too elaborate.

Strive for plain language. Too often, we accept material written in an "official" style that is inefficient and often unfriendly or unclear.

Using plain language is not patronizing, trivializing or over-simplifying. Nor does it mean letting grammar slip. Plain language conveys a clear and concise message. It is used with the reader in mind and with the right tone of voice.

Plain language is faster to read and gets your message across more often, more easily and in a friendlier way.

Remember: clear and simple messages are **NOT** the same as "dumbing down".

Some tips for clear writing:

- Use short sentences;
- Aim for one idea per sentence;
- Break up long sentences;
- Start each paragraph with the most important message;
- Keep paragraphs short;
- Keep your writing crisp.

In the case of verbs, avoid the passive voice; use the active voice. Passive verbs can be confusing and make writing long-winded and less direct.

BAD EXAMPLE:

"The unemployment rate increase was caused by the economic crisis."

GOOD EXAMPLE:

"The economic crisis caused the unemployment rate to rise."

Avoid jargon wherever possible. Jargon is language that has specific meaning to a certain group of people. It is almost always unsuitable for a wider audience. Use the simplest words that fit your message and that your reader will understand. However, if your audience consists entirely of specialists in a particular field, jargon might be the most appropriate language to use.

1.6 Evaluating the impact: media analysis

Did your audience make correct and intelligent interpretations of your data? Could they use the analysis to make sound economic, environmental and social judgments and policies?

You won't know unless you analyze what your audience did with your data release. It is a good idea for statistical organizations to monitor how both the media and other audiences use their information. Some methods include:

- Establish a “Voice of the Customer Programme” or some other customer feedback-gathering process;
- Gauge public reaction through user-groups and focus-groups consisting of your key stakeholders;
- Conduct user surveys or consult influential clients on a one-to-one basis;
- Employ market research services if you need in-depth understanding of a sensitive issue or an audience that is difficult to access;
- Monitor Internet traffic to help determine which data and stories are most in demand from audiences;
- Analyse search keyword patterns and terms for a powerful, yet simple, technique for understanding what information your audience is seeking.

You can work directly with the media to ensure messages are reported accurately and fairly. The media are the simplest, cheapest and most effective way of getting your statistical messages to a wide audience.

Unfortunately, some media have their own agenda, which can result in misuse, misinterpretation and misunderstanding of your data. It is vital to monitor the reach of your messages through the media, as well as the tone and quality of the coverage. In the case of any clear misuse, you should respond rapidly, consistently and firmly.

Now that we have covered the key elements of the text, let us turn our attention to visual elements that help us explain the data better.

2. Visualization of statistics

2.1 Why a picture is still worth a thousand words

We have all heard the old adage: “a picture is worth a thousand words”. One of the best techniques for understanding data is to visualize the numbers as a picture. This can make it far easier to see a pattern or it can expose patterns that might otherwise have been concealed.

You can visualize data in many different ways, from simple bar charts to more complex scatterplots, thematic maps and animated population pyramids. There is also no shortage of technical help: books have been written on visualizing data; there are scores of websites devoted to the subject; and there is a wide range of software and downloadable programs available for every purpose.

This chapter provides a brief historical look at data visualization and guidelines for developing good data visualizations.

2.2 Visualization is an integral part of statistical business processes

You should make the effective presentation of data an integral part of your statistical production process. As data visualization is such an important part of communicating statistical trends and relationships, it must be an on-going activity, not an afterthought. Visualization is included in the “disseminate” phase of the Generic Statistical Business Process Model[4], being developed by the UNECE/ Eurostat/OECD Steering Group on Statistical Metadata.

Readers can easily understand visual presentations. Charts and maps have an impact on nearly everybody through newspapers, television, the Internet and books. It is much easier to understand statistics presented as a chart or a map, rather than long lists of numbers - assuming, of course, that the visual presentations are produced correctly.

The presentations should illustrate trends and relationships quickly and easily. They are an efficient way of getting information out of a database and into the reader's head.

But take care. Poor visualizations of statistical information can be misleading. There are many ways to provide misleading information, whether deliberately or, as is more often the case, unintentionally. There must be a balance between design and function. Complicated visualizations often fail to communicate. As interpreting charts can be demanding, don’t force your readers to have to “dig out” the message.

Misunderstandings and misinterpretations also can result from different cultural traditions. Colours, for example, may have different symbolic meanings in different parts of the world.

[4] http://www1.unece.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model

2.3 A Scot's historical influence on data visualization

The history of statistical charts, as we understand them today, is less than 300 years old. Even though ancient Romans and Arabs were good at using numbers, they did not use bar charts or curves to present figures visually.

René Descartes (1596-1650), the French encyclopedist and philosopher, was the first to use a rectangular coordinate system to visualize his observations. But it was an engineer and economist from Scotland, William Playfair (1759-1823), who was the first to produce graphic presentations of statistical data in ways that are familiar today.

Playfair published two books on the development of world trade at the turn of the 19th century (1786 and 1801), in which he used what he called the "graphic method". He was a keen marketer of this new approach, advocating the advantages of visual presentations with the following arguments:

1. Graphical presentations help to simplify complicated relations that may be difficult to observe:

   "It is hoped that, with the assistance of these charts, information will be got, without the fatigue and trouble of studying the particulars of which it is composed."

2. Business owners, politicians and decision-makers need a visual shortcut to the essence of statistical information, because they do not have the time to study the detail:

   "Men of great rank, or active business, can only pay attention to general outlines [...]."

3. Tables, especially voluminous tables, can be boring and difficult to perceive. Graphical presentations help the user:

   "[...] as much information may be obtained in five minutes as would require whole days to imprint on the memory, in a lasting manner, by a table of figures."

4. Graphics attract the eye. Playfair emphasized this in underlining the usefulness of his method for understanding relations and quantities.

5. The graphic method attracts and challenges the mind, assisting in not only perception but also in understanding structures and relations.

Although there has been a revolution in methods of visualization since Playfair's time, there is still room for refinements and new developments (see chapter 6). However, it is important to avoid presentations that are too flashy and that distort the observed content. **Remember that technology is merely a servant.** You should not add useless notes and obscure elements just because you can. Keep the message simple for the reader.

New masters at promoting a better understanding of graphic presentations and human perceptions today include Jacques Bertin, William Cleveland, Jan-Erik Kristiansen, Vesa Kuusela, Hans Rosling, Edward Tufte and Howard Wainer. Notwithstanding the work of these more recent specialists, the fundamentals of statistical graphics continue to be built on the innovations of William Playfair.

2.4 Basic facts about human perception

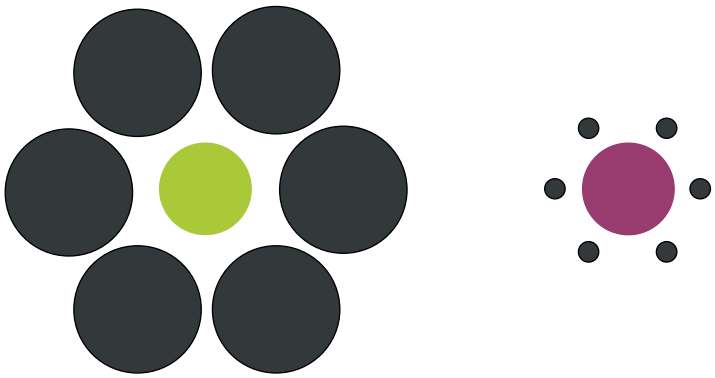
Our capacity to make visual observations rapidly and easily is based on the brain's ability to perceive regularities and irregularities. Much of this ability works unconsciously. The comparison is done almost before we start to think about it.

Look at the figure below. Which end of the horizontal bar in the centre is darker?

Source: Helmholtz, H. (1821-1894), "Simultaneous Contrast Illusion", In *Wikipedia, The Free Encyclopedia*, retrieved in July 2009 from http://en.wikipedia.org/wiki/Optical_illusion.

The bar is the same shade at both ends. The differences in the background produce the misperception. Cover the background with a piece of paper and you will easily notice this.

Now, consider these two images. Which circle is larger, the one in the centre of the diagram on the left or the one in the centre of the diagram on the right?



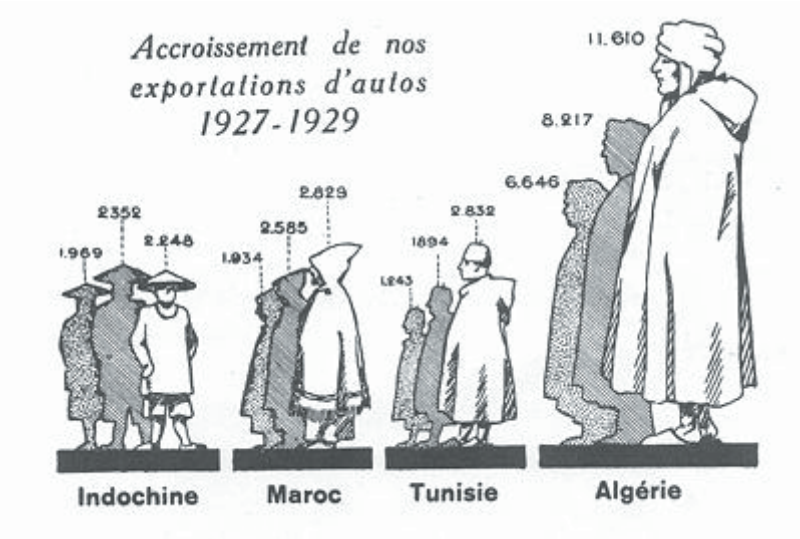
Source: Ebbinghaus, H. (1850-1909), "The Ebbinghaus Illusion", In *Wikipedia, The Free Encyclopedia*, retrieved in July 2009 from http://en.wikipedia.org/wiki/Hermann_Ebbinghaus.

The circles in the centre of each diagram are the same size. If you do not believe this, measure them!

These two examples illustrate how the human mind automatically sees objects in the context of their surroundings. It is the contrast with the surrounding matter that leads you to draw conclusions about the shade of the bar and the size of the two circles. Eliminate the surroundings and you are more likely to come to the correct conclusion.

The message for statisticians is this: you have to be careful when producing visual presentations of statistical observations. The context in which findings are presented may distort the user's perception.

Let us look at a third example. The graphic below is intended to convey the growth of car exports in Indochina, Morocco, Tunisia and Algeria from 1927 to 1929. How well do the sizes of the male figures reflect the real change in the data?



Source: Satet, R. (1932), *Les Graphiques*, Paris. Quoted in Tufte (2001).

In observing a picture, our mind compares the relative size of each object. In this graphic, the data on car exports between 1927 and 1929 are illustrated by the height and volume of the four sets of drawings. The relative values are misrepresented in two ways:

1. The scale is not accurate. The height of the figures for Algeria should be much greater. Something four times larger should be shown four times bigger or four times higher.
2. By overlaying the figure for the most recent year against the earlier years, the difference in values between the years is more difficult to calculate. The figure in the foreground appears much larger than those in the background. It is only their relative height that should be indicating the change between the years.

2.5 Perception is also based on experience

Experience also plays a role in how graphics are perceived. Know your audience and their abilities, experiences and possible differences. Do not assume that they know what you know, either about statistics or about the subject matter.

Statisticians must be aware of their own experience in identifying patterns in numbers, as opposed to the possible inexperience of their readers. As professionals in analysing statistical figures, statisticians are likely to be better than the average person at seeing the underlying message. Statisticians use visualizations to explain their findings, because wider audiences need more help to perceive the statistical information. A visual presentation of data should make the main findings easy to observe and understand.

2.6 Do not misuse the tools

Numerous technological tools are available to assist you in visualizing numeric information. Make sure your focus is on the substance and message of the chart, rather than on the methodology, design or technology of the graphic presentation.

Basic spreadsheet software has more than 70 standard chart types to choose from, plus the ability to create customized charts. This exciting range of options often inspires information providers to use all sorts of different ways to visualize data - just because they can. The result, however, can often be over-complication or distortion of statistical information.

This danger increases with the ever-growing flashiness and variety of new tools available. When using these tools, you should keep in mind that technology is simply a means to an end. The message contained within statistical information is far more important than showing how many flashy tools you may know how to use.

2.7 Checklist for developing good data visualizations

When producing visual presentations, you should think about:

- **The target group:** different forms of presentation may be needed for different audiences (e.g. business or academia, specialists or the general population).
- **The role of the graphic in the overall presentation:** analysing the big picture or focusing attention on key points may require different types of visual presentations.
- **How and where the message will be presented:** a long, detailed analysis or a quick slideshow.
- **Contextual issues that may distort understanding:** expert or novice data user.
- **Whether textual analysis or a data table would be a better solution.**
- **Accessibility considerations:**
  - Provide text alternatives for non-text elements such as charts and images.
  - Don't rely on colour alone. If you remove the colour, is the presentation still understandable? Do colour combinations have sufficient contrast? Do the colours work for the colour blind (red/green)?
  - Ensure that time-sensitive content can be controlled by the user (e.g. pausing of animated graphics).
- **Consistency across data visualizations:** ensure that elements within visualizations are designed consistently and use common conventions where possible (e.g. blue to represent water on a map).
- **Size, duration and complexity:** Is your presentation easy to understand? Is it too much for the audience to grasp at a given session?
- **Possibility of misinterpretation:** test your presentation out on colleagues, friends or some people from your target group to see if they get the intended messages.

3. Tables

3.1 Why tables are important

Good tables are an integral part of your package, whether this is a news release, an analytical article or a research paper. Using tables effectively helps minimize the number of data values in your text. It also eliminates the need to discuss less significant variables that are not essential to the story line.

In her book on writing about numbers, Miller (2004) gives the following guidelines on how to design good tables:

- Make it easy for your audience to find and understand numbers within your tables.
- Design both the layout and the labeling of your tables in a straightforward and unobtrusive fashion, so the attention is on the substantive points to be conveyed by your data, rather than on the structure of the table.

In this chapter, we cover what makes an effective table.

3.2 Two types of tables

You should be concerned with two types of tables. The first are smaller tables, called presentation (or demonstration) tables. They can be used to highlight key figures in a press release, web page or analytical publication.

The second type consists of larger tables, called reference tables. They are increasingly being replaced with interactive databases that allow users to generate their own tables online. As reference tables are more of an analytic tool, they will not be discussed here.

In presentation tables, data should be presented in a concise, well-organized way to support the accompanying analysis. A small, well-formatted table can provide a great deal of information that readers can quickly absorb.

Tables should be able to stand alone, whether published within a report, article, publication or web page. Each table should contain enough metadata, such as a descriptive title and indication of source, to allow it to be copied and pasted into another document and still make sense. If you ensure that your tables can stand alone, they are more likely to be understood correctly, within or outside their original context.

3.3 Checklist for designing a good table

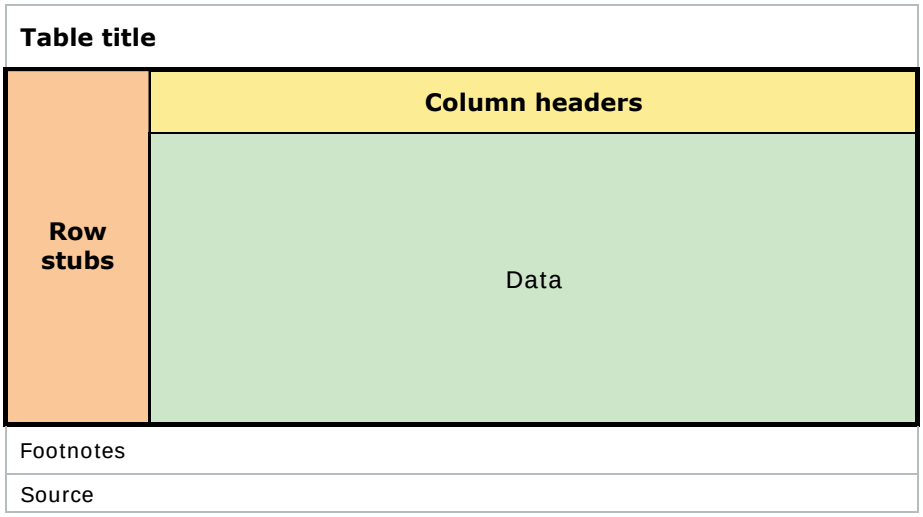
Five support components are needed to describe the data displayed in a table:

- The **table title** should give a clear and accurate description of the data. It should answer the three questions "what", "where" and "when". Be short and concise, and avoid using verbs.

GOOD EXAMPLE
"Planned age of retirement by occupation, Canada, 2007."
This text gives all the information needed to understand and use the data correctly.

- **Column headers**, at the top of the table, should identify the data presented in each column of the table and provide any relevant metadata (e.g. unit of measurement, time period or geographic area).
- **Row stubs**, in the first column of the table, should identify the data presented in each row of the table.
- **Footnotes**, at the bottom of the table, may provide any additional information needed to understand and use the data correctly (e.g. definitions).
- The **source line**, at the bottom of the table, should provide the source of the data, i.e. the organization that produced the data and the data collection method (e.g. population census or labour force survey).

The figure below shows how these table components should be displayed.



In a presentation table, you should display only a small subset of your data selected to best communicate your message, as illustrated in the example below.

GOOD EXAMPLE of a presentation table

Manufacturing sales in Canada, provinces and territories, June-July 2008
Seasonally adjusted

	June 2008[r]	July 2008[p]	June-July 2008
	$ millions		% change[1]
Canada	**52 685**	**54 105**	**2.7**
Newfoundland and Labrador	692	674	-2.5
Prince Edward Island	123	115	-6.1
New Brunswick	1 914	1 872	-2.2
Quebec	13 019	13 280	2.0
Ontario	23 902	25 015	4.7
Manitoba	1 360	1 445	6.2
Saskatchewan	1 079	1 108	2.8
Alberta	6 298	6 316	0.3
British Colombia	3 347	3 306	-1.2
Yukon	3	4	45.5
Northwest Territories and Nunavut	4	3	-27.4

[r] Revised
[p] Preliminary
[1] The percentage change is calculated from data in thousands of Canadian dollars

Source: Statistics Canada

Data values should be set out so key information can be extracted easily. Users may find it easier to scan down columns or across rows, depending on your message. You should consider this when deciding whether to present your table in portrait or landscape orientation. Lines or subtle shading can also be used to encourage users to read horizontally, as well as vertically. Spacing and shading can change the way a table is read.

Columns should be evenly spaced and not too far apart. The table should only be as wide as the data content requires.

To ensure that your tables are easy to understand, you should consider the following guidelines:

- Avoid unnecessary text.
- Display your data either by chronological order for time series or by using some standard classification. For longer time series, it may be more appropriate to use the reverse chronological order (i.e. starting with the most recent period and going backwards) in some cases, such as for monthly unemployment.
- Use a minimum of decimal places.
- Use thousand separators. Using a space instead of a symbol can avoid the problem of having to translate between languages.
- Align the numbers on the decimal point (or on the right in the absence of decimal places) so their relative value is clear. Do not centre the numbers in a column, unless they are all the same magnitude.
- Do not leave any data cell empty. Missing values should be identified as “not available” or “not applicable”. The abbreviation “NA” can apply to either, so it needs to be defined.

Some of these guidelines are discussed and illustrated in the next section.

When producing a series of tables for a publication or a website, you should use the same layout in all tables. Consider how much information needs to be provided in table titles (what is obvious and what is not) and be consistent in the use of abbreviations.

3.4 The use of rounding and decimals

Many non-statistical users find it difficult to see the difference between numbers when three or more digits vary. You can help them by rounding the values presented in your tables. Rounding can also be used when the data do not have a high degree of accuracy. In some cases, only rounded data are reliable and should therefore be displayed in tables. You should, however, take care not to lose too much information when rounding your data.

GOOD EXAMPLE	BAD EXAMPLE
1 320 000	1324567
1 670 000	1673985
1 830 000	1829456

In the example above, the rounded numbers on the left are easier to understand and memorize than the exact numbers on the right. The use of a space as a thousand separator is also illustrated in this example.

If you need to display values with varying numbers of decimal places, you should align them on the decimal point, not on the right. In the example below, the values on the left are easier to read than those on the right. This example also shows that it is much better to display the same number of decimal places in all values.

GOOD EXAMPLE	BAD EXAMPLE
93.2	93.2
1045.0	1045
385.6	385.63

Numeric values should be right justified. Using the same example, notice how difficult it is to read the values when the numbers are justified to the left margin as shown below.

GOOD EXAMPLE	BAD EXAMPLE
93.2	93.2
1045.0	1045.0
385.6	385.6

3.5 Example of how to improve a table

To illustrate the effectiveness of the guidelines presented in section 3.3, we show below an example of a bad table and how it can easily be improved.

BAD EXAMPLE

Final energy consumption by sector - Percentages

	1980	1985	1990	1995	2000	2002	2003
Transport	27.81	27.92	28.24	31.12	36.82	39.48	39.13
Residential	31.11	33.91	30.41	27.61	24.33	23.71	23.97
Industry	31.47	27.21	23.86	22.11	21.41	19.53	18.78
Agriculture	n/a	n/a	3.51	3.7	3.11	2.91	2.82
Services	9.61	10.96	13.98	15.46	14.33	14.37	15.3
Total	100	100	100	100	100	100	100

What is wrong in the table above?

- We do not know which geographic area the data refer to.
- The data source is not identified.
- The values are centered rather than right-aligned.
- The values should not be displayed with two decimal places (too much information).
- The total values should have the same number of decimal places as the other values.
- The abbreviation "n/a" is not explained.
- The grey shading and the lines of the same size between each row and each column do not help to understand the different data presented in the table.
- The table is unnecessarily spread across the width of the page.

GOOD EXAMPLE

Share of total energy consumption, by sector (in percent)
Ireland, 1980-2003

	1980	1985	1990	2000	1995	2002	2003
Transport	27.8	27.9	28.2	31.1	36.8	39.5	39.1
Residential	31.1	33.9	30.4	27.6	24.3	23.7	24.0
Industry	31.5	27.2	23.9	22.1	21.4	19.5	18.8
Agriculture	n/a[1]	n/a[1]	3.5	3.7	3.1	2.9	2.8
Services	9.6	11.0	14.0	15.5	14.4	14.4	15.3
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

[1] Data on energy consumption for the agricultural sector was not collected until 1990.
Source: Department of Public Enterprise, Ireland

How has the table been improved?

- All the information needed to understand the data is provided in the title and subtitle.
- The data source is identified.
- All values are right-aligned and displayed with one decimal place.
- The abbreviation "n/a" is explained in the footnote.
- Only the lines that separate the different components of the table (header, data, footnote and source) are displayed and the unhelpful shading has been removed.
- The table is not wider than needed to display all the headings and data.

4. Charts

4.1 Why use charts?

Statistics can often be better understood when they are presented in a chart than in a table. A chart is a visual representation of statistical data, in which the data are represented by symbols such as bars or lines. It is a very effective visual tool, as it displays data quickly and easily, facilitates comparison and can reveal trends and relationships within the data.

A chart generally takes the form of a one- or two-dimensional figure, such as a bar chart or a line chart. Although there are three-dimensional charts available, they are usually considered too complex to be easily understood.

Charts can be used to illustrate patterns in a large amount of data or to communicate a key finding or message. You should consider using charts if you want to show:

- **Comparison:** How much? Which item is bigger or smaller?
- **Changes over time:** How does a variable evolve?
- **Frequency distribution:** How are the items distributed? What are the differences?
- **Correlation:** Are two variables linked?
- **Relative share of a whole:** How does one item compare to the total?

In this chapter, we examine the most common types of charts and give guidelines to producing good charts.

4.2 Checklist for designing a good chart

If you decide that a chart is the most appropriate way to present your data, then no matter what type of chart you use, you need to keep the following three guidelines in mind:

1. **Define your target audience**: What do they know about the issue?
2. **Determine the message you want to communicate**: What do the data show? Is there more than one message?
3. **Determine the nature of your message**: Do you want to compare items, show time trends or analyze relationships in your data?

A good chart:

- grabs the reader's attention;
- presents the information simply, clearly and accurately;
- does not mislead;
- displays the data in a concentrated way (e.g. one line chart instead of many pie charts);
- facilitates data comparison and highlights trends and differences;
- illustrates messages, themes or storylines in the accompanying text.

4.3 When it may not be appropriate to use charts

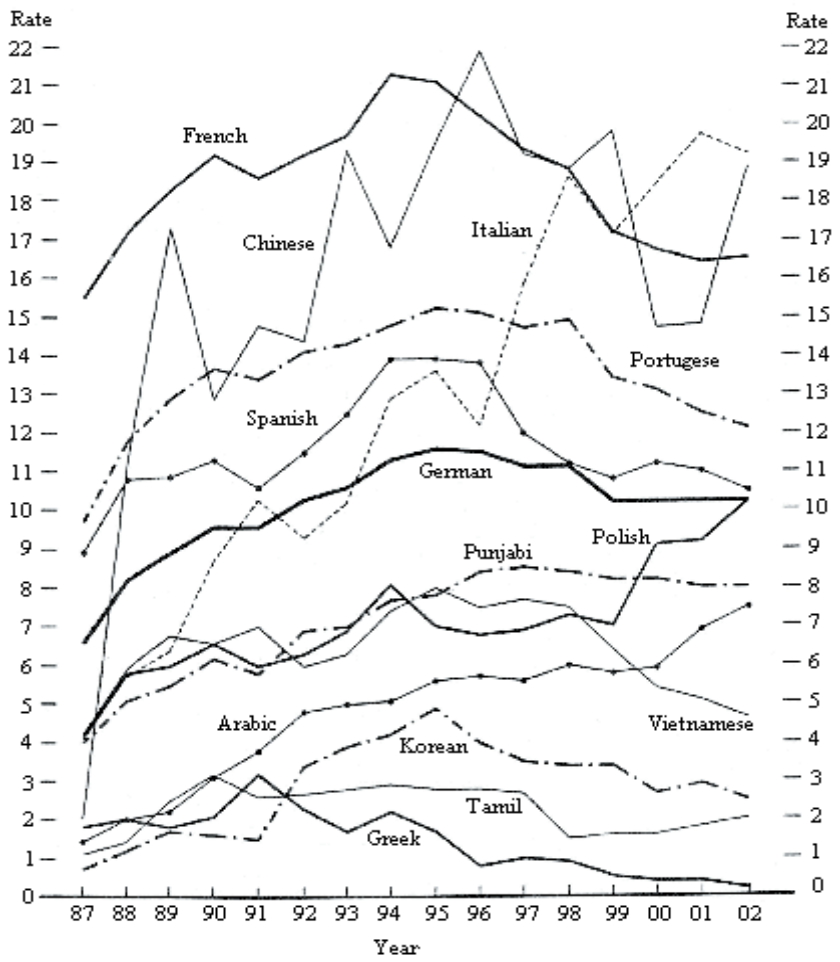
A chart is not always the most appropriate tool to present statistical information. Sometimes a text and/or data table may provide a better explanation to your audience and save you considerable time and effort.

You should reconsider using charts when your data:

- are very dispersed;
- have too few values;
- have too many values;
- show little or no variation.

BAD EXAMPLE of a line chart

Number of students taking English as a second language at West High School, by first language spoken, 1987 to 2002



Source: Statistics Canada, *Learning Resources: Using graphs*[5].

You should avoid anything resembling the line chart above. The data are far too numerous and whatever storylines the analyst hoped to illustrate are lost in the jungle of lines.

[5] http://www.statcan.gc.ca/edu/power-pouvoir/ch9/using-utilisation/5214829-eng.htm

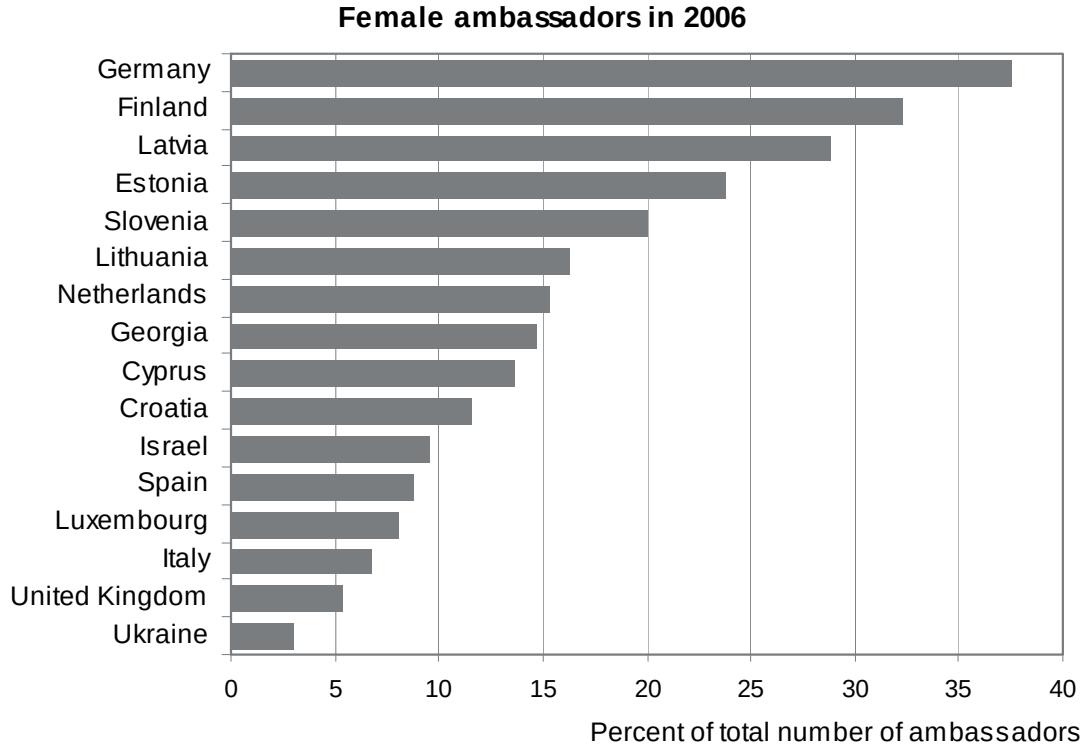
4.4 Selecting the appropriate type of chart

Knowing what type of charts to use with what type of information is crucial. Some charts are more appropriate than others, depending on the nature of the data. In this section, we provide guidelines for the most common types of charts: bar charts and population pyramids, line charts, pie charts and scatter plots.

Bar charts

A bar chart is the simplest type of chart to draw and read. It is used to compare frequencies or values for different categories or groups.

GOOD EXAMPLE of a bar chart



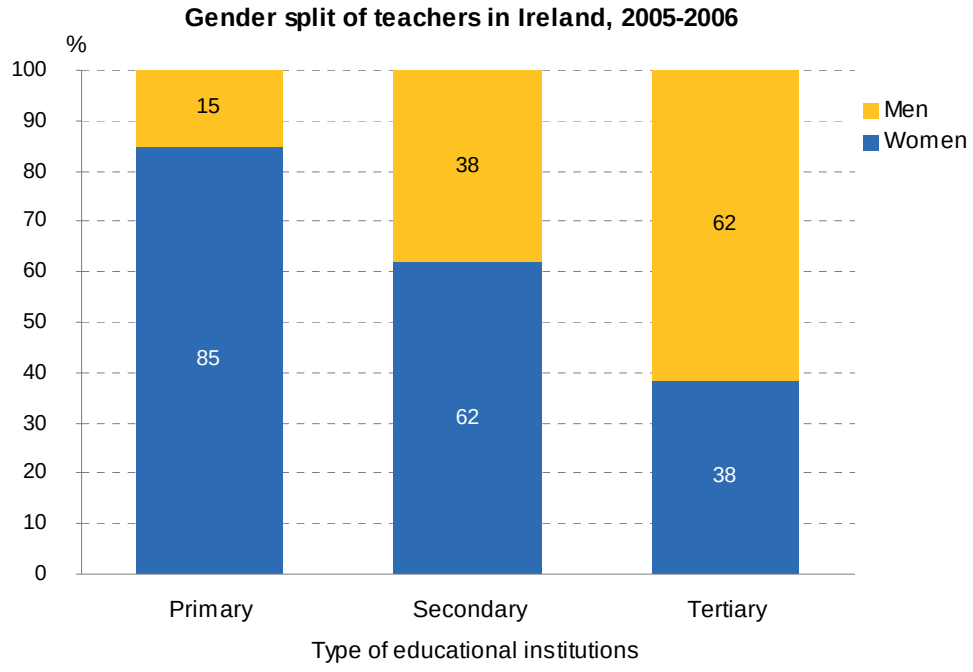
Source: UNECE Statistical Database

The bars can be either vertically or horizontally oriented. In the horizontal orientation, the text is easier to read, as in the example above. It is also easier to compare the different values when the bars are ordered by size from smallest to largest, rather than displayed arbitrarily.

The bars should be much wider than the gaps between them. The gaps should not exceed 40% of the bar width.

A stacked bar chart can be used to show and compare segments of totals. Caution should be exercised when using this type of chart. It can be difficult to analyze and compare, if there are too many items in each stack or if many items are fairly close in size.

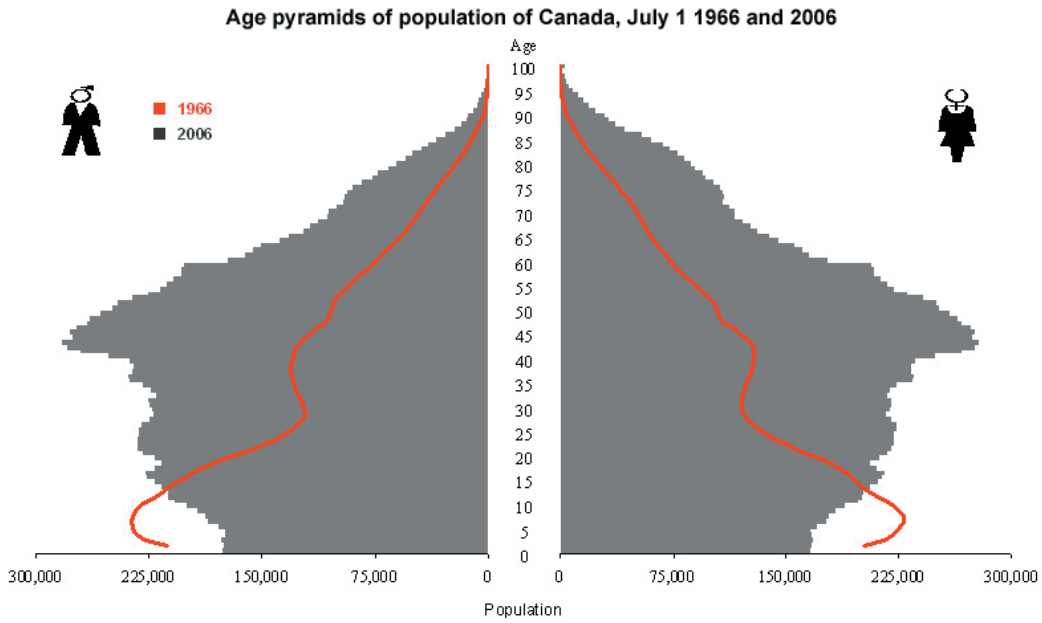
GOOD EXAMPLE of a stacked bar chart



Source: UNECE Statistical Database

A population pyramid is a combination of two horizontal bar charts, representing the age structure of the female and male population of a country or region. Men are conventionally shown on the left and women on the right. When you want to compare different population pyramids, it is usually better to represent the percentage of men and women in the total population, rather than their number.

GOOD EXAMPLE of a population pyramid



Source: Statistics Canada[6]

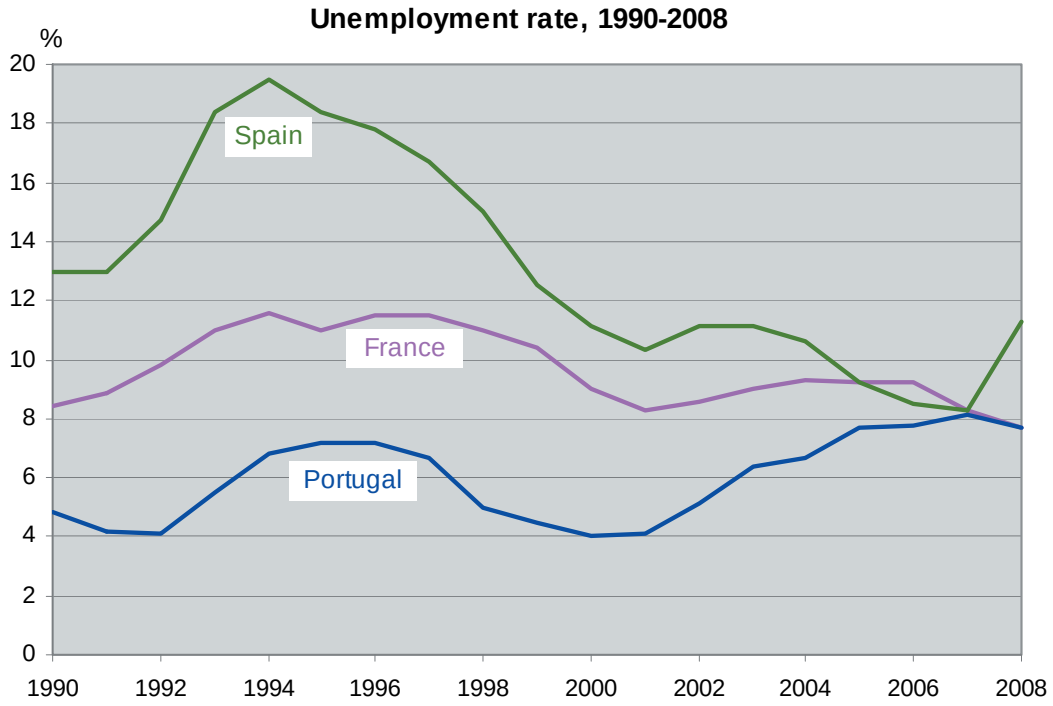
[6] http://www.statcan.gc.ca/daily-quotidien/061026/figure.htm

For most European countries, population pyramids do not take the shape of a pyramid any more, but they remain a very effective way of displaying a great deal of information on the age and sex structure of populations, even more so when they are "animated", i.e. moving through time.

Line charts

A line chart is an effective tool for visualizing trends in data over time and is therefore the most appropriate type of chart for time series. You can adjust the chart parameters to better communicate your message, but you should be careful not to distort the data. This issue is discussed and illustrated in section 4.6.

GOOD EXAMPLE of a line chart



Source: UNECE Statistical Database

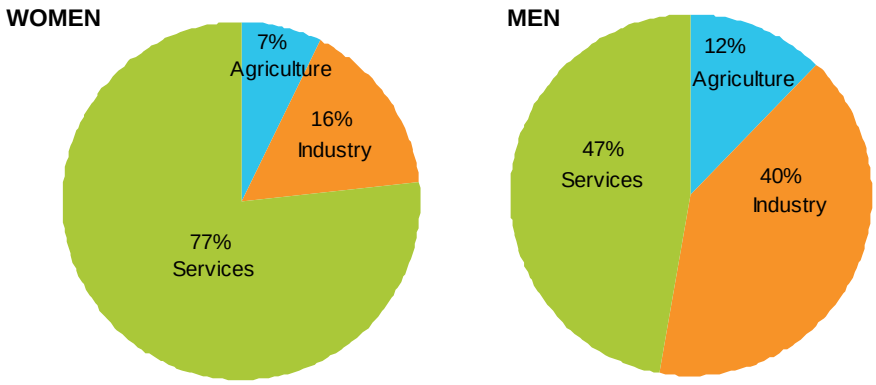
Pie charts

A pie chart can be used to show the percentage distribution of one variable, but only a small number of categories can be displayed, usually not more than six. The use of this type of chart is not recommended by many statisticians, as it can be difficult to compare the different segments of the pie and, even more, to compare data across different pie charts. To overcome this problem, the segments can be labeled with their actual values. In some cases, the category names can also be written as labels on the chart, so that the legend is not necessary. Segments are usually best presented in order from smallest to largest segments, rather than interspersing small and large segments.

In most cases, other types of charts (e.g. bar charts) are more appropriate, but pie charts should not be completely ruled out, as they are effective to visualize the relative importance of one category in the total. Pie charts can be well suited to provide an overview of a situation, such as in the example below.

GOOD EXAMPLE of a pie chart

Employment by major sectors in Latvia, 2007



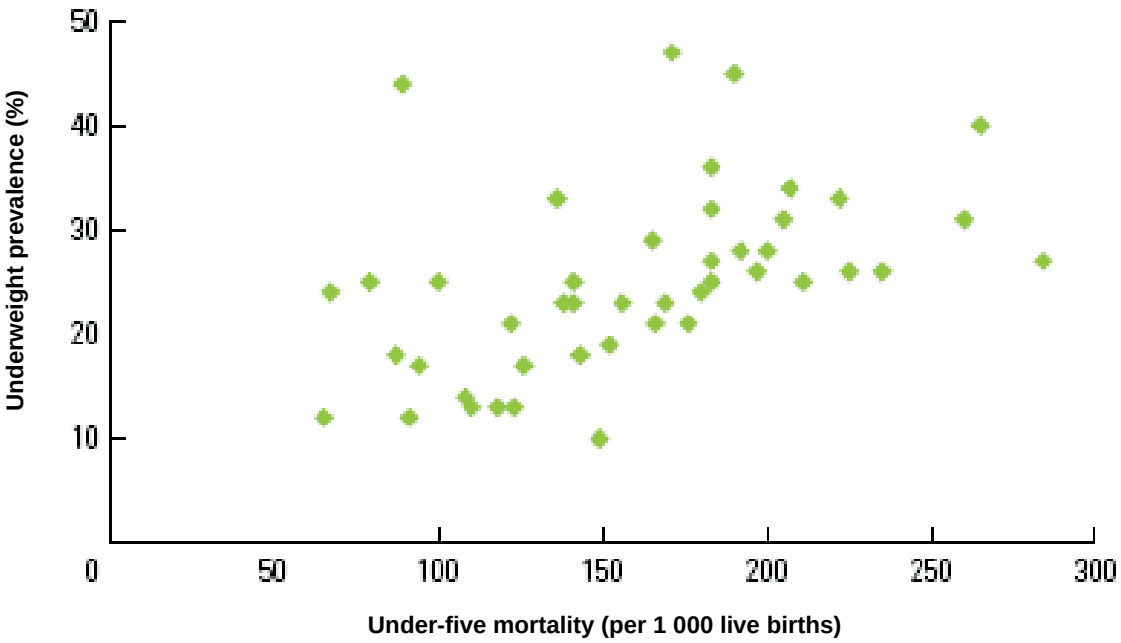
Source: UNECE Statistical Database

Scatter plots

A scatter plot is used to show the relationship between two variables. It is the most accurate way to display correlations, as illustrated in the example below. However, some analysts prefer to use bar charts, as scatter plots can be difficult to interpret.

GOOD EXAMPLE of a scatter plot

Under-five mortality and underweight prevalence in Sub-Saharan African countries, 2003



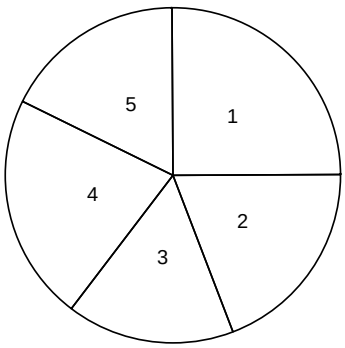
Source: Jamison et al. (2006) *Disease and Mortality in Sub-Saharan Africa, 2nd edition*, Washington D.C., The World Bank[7].

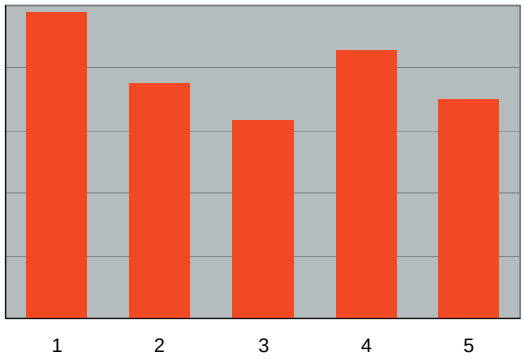
[7] http://www.dcp2.org/file/66/Disease and Mortality in SSA.pdf

Experimenting with different types of charts

Which type of chart should you use? You have to make that decision. A good practice is to experiment with different types of charts to select the most appropriate tool to communicate your message.

Below are two different ways to graph the same data. Which one is clearer?





Can you tell which segment on the pie chart is the biggest one? Some readers tend to find it more difficult to compare angles than bars or lines. On the pie chart, segments 1 and 4 look practically the same, while the difference in their relative size is immediately clear on the bar chart.

4.5 What makes an effective chart

Chart components

The different chart components compete with each other for the reader's attention. The more features you include, the harder it becomes to see your point.

Chart components fall into three categories:

1. **Data components** that represent the data: bars, lines, areas or points.
2. **Support components** that assist in understanding the data: title, legend, data labels, gridlines, footnotes and data source.
3. **Decorative features** that are not related to the data.

Data components alone are never self-sufficient. To ensure correct understanding of your charts, you need to include the following support components:

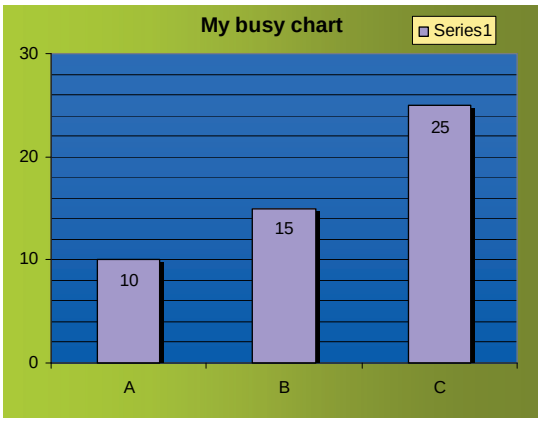
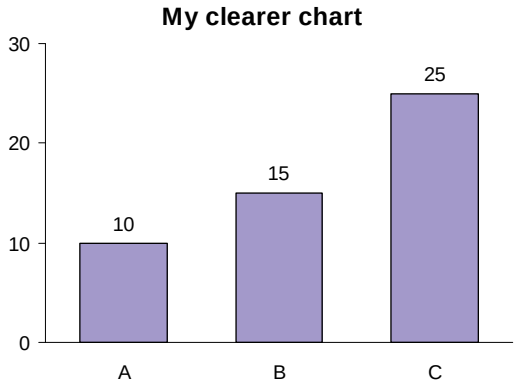
- The **chart title** should give a clear idea of what the chart is about. It has to be short and concise. You can have two types of titles:
  - An **informative title** provides all the information needed to understand the data. It should answer the three questions “what”, “where” and “when”.
  - A **descriptive title** is a caption that highlights the main pattern or trend displayed in the chart. It states in a few words the story that the chart illustrates.

- The **axis labels** should identify the values displayed in the chart. The labels are displayed horizontally on both axes.
- The **axis titles** should identify the unit of measure of the data (e.g. “in thousands”, “%”, “age (in years)” or “$”). You do not need to include an axis title when the unit of measure is obvious (e.g. “years” for time series).
- **Gridlines** can be added in bar and line charts to help users read and compare the values of the data.
- The **legend and data labels** should identify the symbols, patterns or colors used to represent the data in the chart. The legend should not be displayed when only one series of values is represented in the chart. Whenever possible, you should use data labels rather than a legend. Data labels are displayed on or next to the data components (bars, areas, lines) to facilitate their identification and understanding.
- A **footnote** may be used to provide definitions or methodological information.
- The **data source** should be identified at the bottom of the chart.

It’s all about the data

To maximize the efficiency of a chart, data should take centre stage. Support components should:

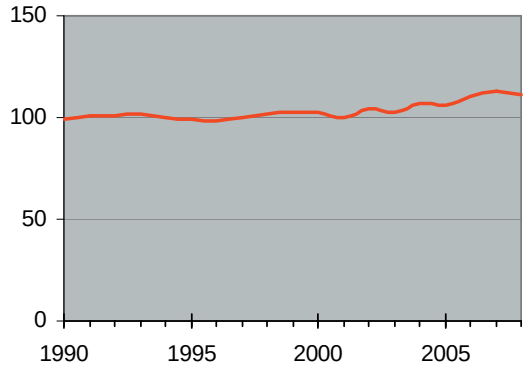
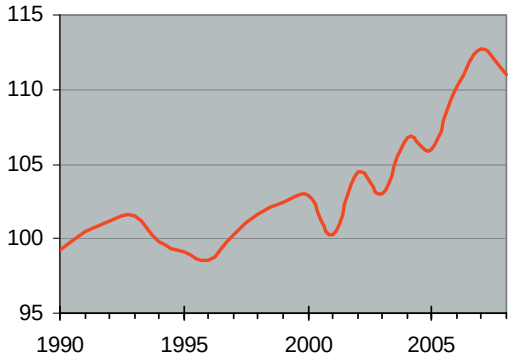
- **Only be present if relevant.** Title axes, legend and data labels may be essential for the correct understanding of your chart or may not be needed at all, depending on the nature of your data.
- **Be subtle.** Use lighter lines for axes and gridlines than for data components. Decorative feature should not distract the reader’s attention.

BAD EXAMPLE	GOOD EXAMPLE
	
All components have maximum impact. The result is a busy chart, difficult to read, even though it shows only three values.	This chart is much easier to read. Minimal use of support components ensures that data take centre stage.

Data components can also conflict with each other. The more variables and values you want to display, the more difficult it is to present the data clearly. An effective chart has a clear, visual message. If a chart tries to do too much, it becomes a puzzle that requires too much work to understand. In the worst case, it is just plain misleading.

4.6 Adjusting the chart parameters

When designing a chart, you can adjust the scales to best convey your message. The two line charts below display the same data, but they provide very different pictures:

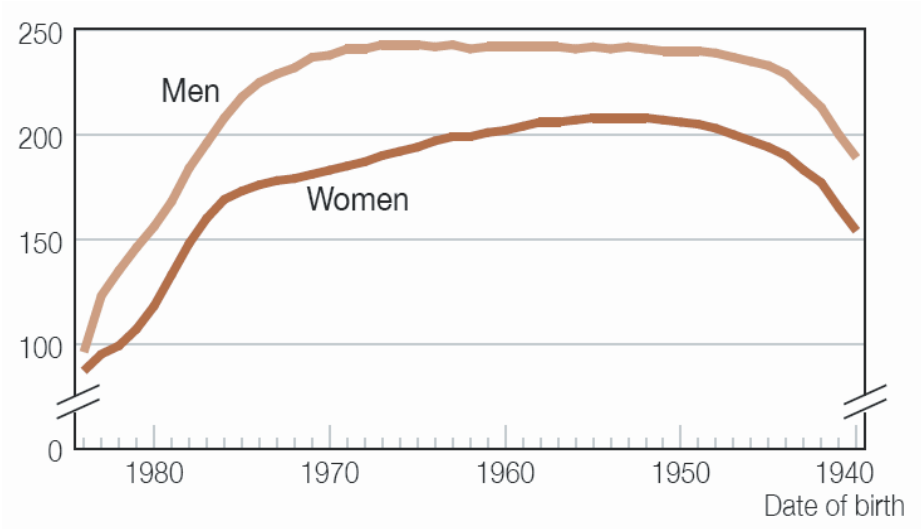
	
This chart shows an overall trend of stability for about 10 years, followed by moderate growth.	By reducing the scale on the y-axis, you can highlight the changes. Small increases and decreases appear in the first 10 years and the growth at the end of the period seems more dramatic.

It is good practice to use some symbol to indicate when the scale of values does not start at zero, such as on the y-axis in the right-hand side example above. The best option is to start from zero and put either a zigzag line or a break, as illustrated in the example below.

GOOD EXAMPLE of a chart with a y-axis not starting at zero

Pensionable income for persons aged 20-64 in Sweden, 2004

Average income in thousands Swedish Krona



Source: Statistics Sweden (2006), *Women and Men in Sweden: Facts and figures 2006*[8].

[8] http://www.scb.se/statistik/_publikationer/LE0202_2006A01_BR_X10ST0602.pdf

4.7 Controlling the cognitive load of your charts

Your data may contain several messages that you want to highlight using a chart. Charts, like every element in a publication, can be assigned a "cognitive load". Cognitive load basically means how hard the reader has to work to understand what you are trying to communicate. A chart with a high cognitive load will be hard to understand and to remember. Its message will be difficult to communicate. A chart with a low cognitive load will be understood at a glance. Its message will be obvious. Most guidelines about effective chart design are meant to keep cognitive load low.

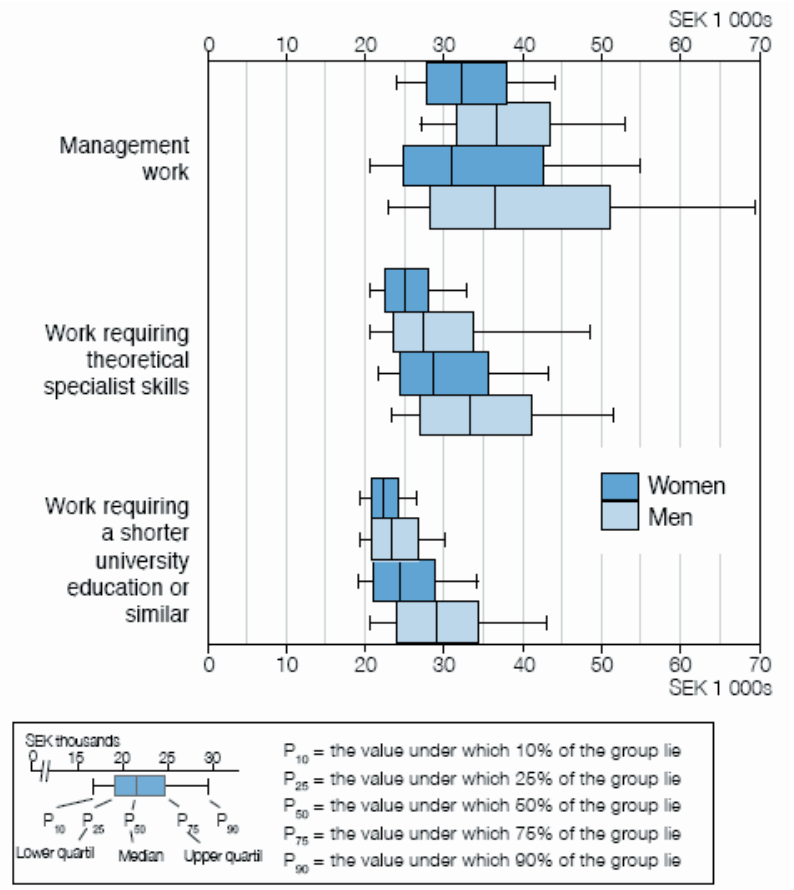
When you design a chart, you control its cognitive load. You can lower it and send a clear message, by using appropriate conventions and formats. You can also deliberately attempt to raise the cognitive load of your chart, if you want your audience to focus on a subtle aspect of your data. By raising the cognitive load, you force readers to consider the chart from a different point of view. Below is an example of a chart with a high cognitive load.

GOOD EXAMPLE of a chart with a high cognitive load

Wage dispersion in occupational groups that require higher education in Sweden, 2004

Monthly salary in Swedish Krona (SEK)

The two upper bars in each occupational group show women and men in the public sector, the two lower bars show the private sector.



Source: Statistics Sweden (2008), *Women and Men in Sweden: Facts and figures 2008*[9].

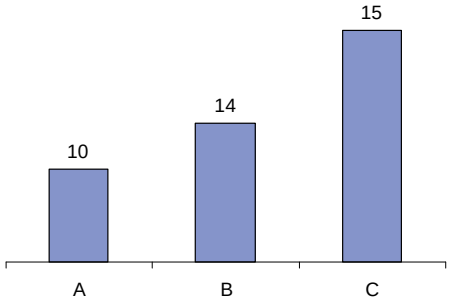
[9] http://www.scb.se/statistik/_publikationer/LE0202_2008A01_BR_X10BR0801ENG.pdf

4.8 Suggestions to improve your charts

Be accurate

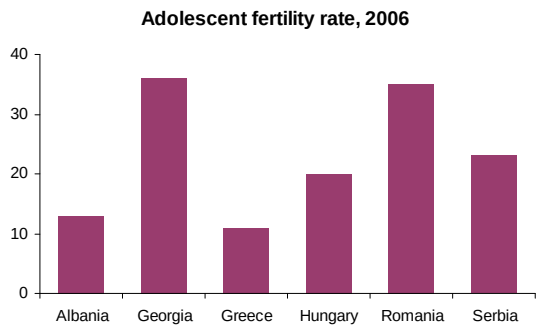
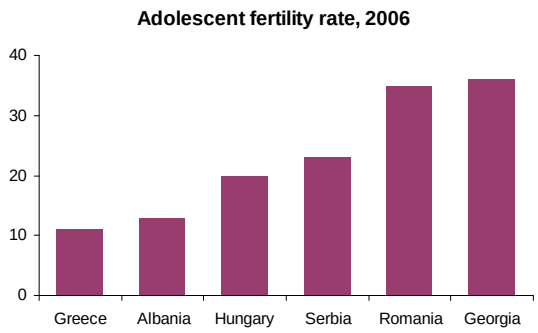
Graphical objects must be sized to present ratios accurately. A chart that displays data as objects with disproportionate sizes, such as in the example below, is misleading.

BAD EXAMPLE of relative size between chart objects



Sort your data

When using bar or pie charts, you should sort your data from smallest to largest values, so they are easier to compare.

BAD EXAMPLE	GOOD EXAMPLE
	
The data are presented by alphabetical order of countries. The values are very difficult to compare. Attention is on the first and last values, which have no specific relevance.	The data are presented in order from smallest to largest values. It is easy to compare them. Attention is focused on the minimum and maximum values of the dataset.

Avoid misleading correlations

Plotting variables with different scales on the same chart is likely to result in erroneous conclusions. The fact that two curves move together is not sufficient to establish a correlation.

The chart below attempts to link suicide and long-term unemployment rates in Japan. Both curves appear to move together, but the two variables are different. One is the number of suicides **per 100 000** population, while the other is the number of unemployed for 12 months or more as a **percentage** of the total number of unemployed. Both variables happen to have values between 10 and 35

over the whole period, but any slight change in definition or scale would have produced a very different chart. This chart is successful at conveying an impression that both variables are linked, but it cannot and does not prove it.

BAD EXAMPLE: misleading correlation between two variables

Suicide and long-term unemployment in Japan

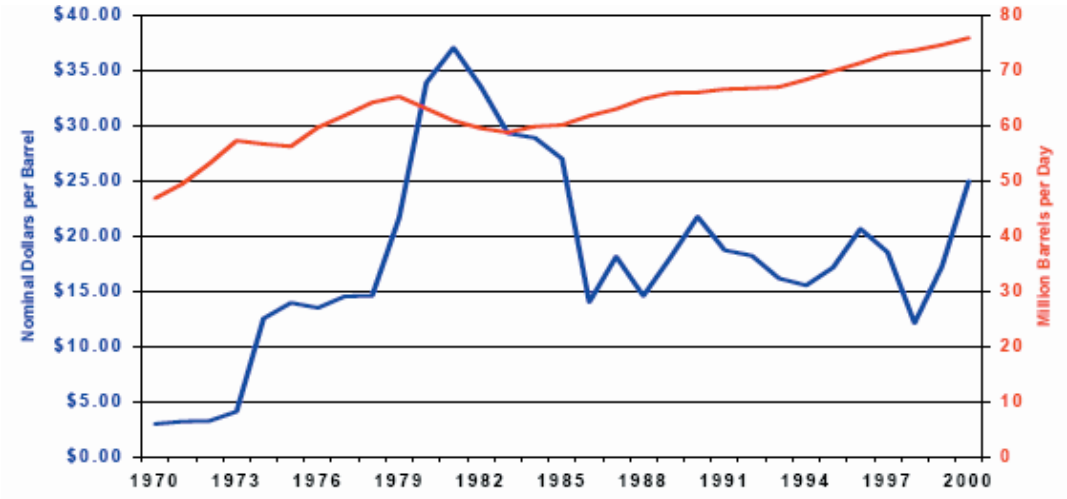
Source: Swivel[10]

Use dual y-axes with caution

Dual y-axes have the potential to cause confusion. You may be able to use this type of chart successfully if you have two different variables (e.g. price and quantity). But you should be very careful with your labels and show each data line in the same colour as the axis the user needs to consult, as shown in the example below.

GOOD EXAMPLE of a dual y-axes chart

World demand and world oil prices



Source: Blessing et al. (2003), *Cognitive Testing of Statistical Graphs: Methodology and Results*[11].

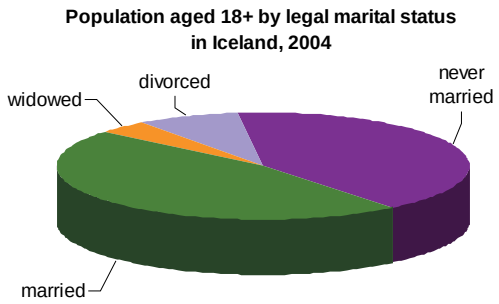
[10] http://www.swivel.com/graphs/show/28847825
[11] http://www.fcsm.gov/03papers/BlessingBradsher.pdf

When you have similar units on two different scales in the same chart, you can confuse even experts in the subject. Lines might “cross” in a dual y-axis chart that, if drawn on the same scale, would not be anywhere close to each other.

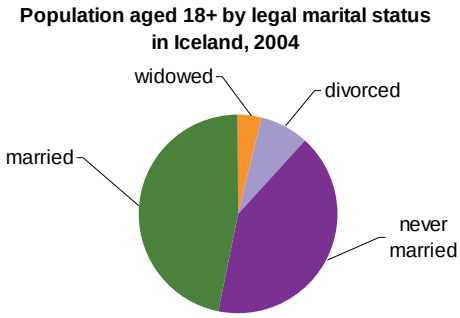
Avoid unnecessary graphic features

Any graphic features that do not show something about the dataset you are representing will make your chart less legible. This is especially true when using three dimensions for charts that represent simple datasets. Although software is available to easily produce three-dimensional charts and some people may consider them attractive, they often distort the data. The message is much easier to understand when presented in a simple two-dimensional format, as shown below.

BAD EXAMPLE



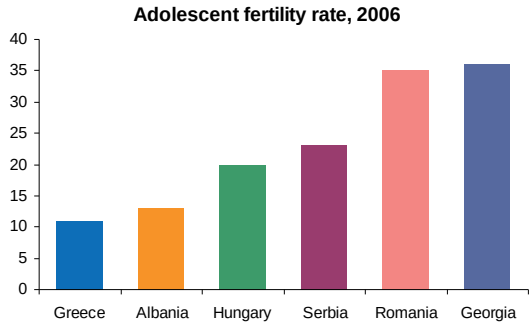
GOOD EXAMPLE



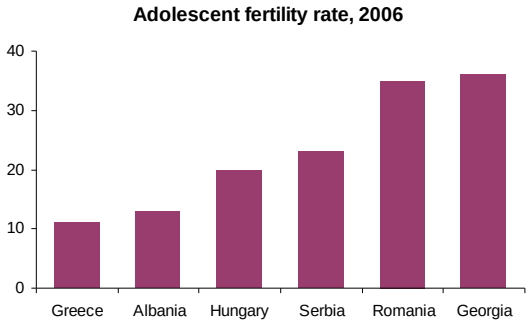
Three-dimensional charts rarely add value and often confuse readers. The images have depth, making some parts appear closer and others further away. The brain compensates for this by believing the objects that look like they are in the distance are larger. However, when representing data that have more than two attributes, using depth can in fact allow readers to understand your point more clearly.

You should generally avoid adding any feature that does not carry any data. For instance, using a different colour for each value in a bar chart makes it harder to read, as shown in the example below. You should stick with one colour.

BAD EXAMPLE



GOOD EXAMPLE



5. Maps

5.1 Why a map is worth a thousand numbers

Geographic information is an integral part of all statistical data. Geographic areas have boundaries, names and other information that make it possible to locate them on the ground and relate statistical information to them. This spatial relationship is particularly important for census data.

Maps are the most efficient tools to visualize spatial patterns. When carefully designed and presented, they are more than just decorative features in a statistical presentation. They can help people identify and highlight distributions and patterns that might not be apparent from tables and charts.

If "a picture is worth a thousand words", then "a map is worth a thousand numbers". In our visual era, maps are a powerful information medium. They serve as valuable decision-making tools for experts, politicians and the general public, and meet a growing demand for information in all parts of society.

The power of maps

Maps are well-designed if they are easy to grasp. They should help people understand a large amount of information at a glance. They can summarize voluminous data tables or long and complicated texts. When you want to present statistical information for all regions of your country, you can produce a whole set of charts, or you can show all the information in a single map.

Many cartographic technologies are now available, from geographic information systems (GIS), which offer a broad range of analytical functions and integrate map components, to high-end cartographic information systems (CIS) for professional map and atlas makers in statistical offices. During the age of paper publications, maps were often underused by statisticians, because they did not show exact numbers. This drawback disappeared with the emergence of interactive mapping tools that allow the user to retrieve the actual data "behind the map".

Using maps in statistics

Maps can be very useful both in the preparation of censuses and surveys and in the analysis and reporting of results. You should consider using maps if you want to:

- show the geographical location and spatial distribution of your data;
- compare different areas;
- summarize a large volume of data and reduce their complexity;
- communicate a clear message;
- validate your findings;
- attract people's attention;
- store spatial information in a geographical information system.

In this chapter, we examine the most common types of maps and give guidelines to producing good maps.

5.2 Checklist for designing a good map

Mapmaking is a mixture of art, science and technology. It is a complex task, as there are unlimited possibilities for organizing the layout.

With the growth of Web 2.0 technologies (see sections 6.2 and 6.4), many interactive mapping tools are available online that allow users to upload their data and retrieve thematic maps instantly. The production of thematic maps has become much cheaper and faster, but it does not automatically result in well-designed maps that communicate your message accurately.

To design a good map, you need to consider the following four guidelines:

1. **Define your target audience:** How and in which context will the map be used? Are there any accessibility constraints?
2. **Determine the message you want to communicate:** What do the data show? Is there more than one message?
3. **Determine the nature of your data:** How many variables should be mapped? Is there a time dimension?
4. **Determine the appropriate mapping technique, colours and symbols:** What is the nature of your data (quantitative or qualitative, absolute or relative values)? Is there any technical constraint (e.g. format or black and white reproduction)? What are the conventions for colours or classifications?

A good map:

- is simple and easily understood;
- has a clear and objective message;
- gives an accurate representation of the data and does not mislead;
- attracts the reader's attention to the most important information;
- is well presented and attractive;
- fits the output format and your audience;
- can stand by itself without further explanations;
- is accessible to colour-blind persons.

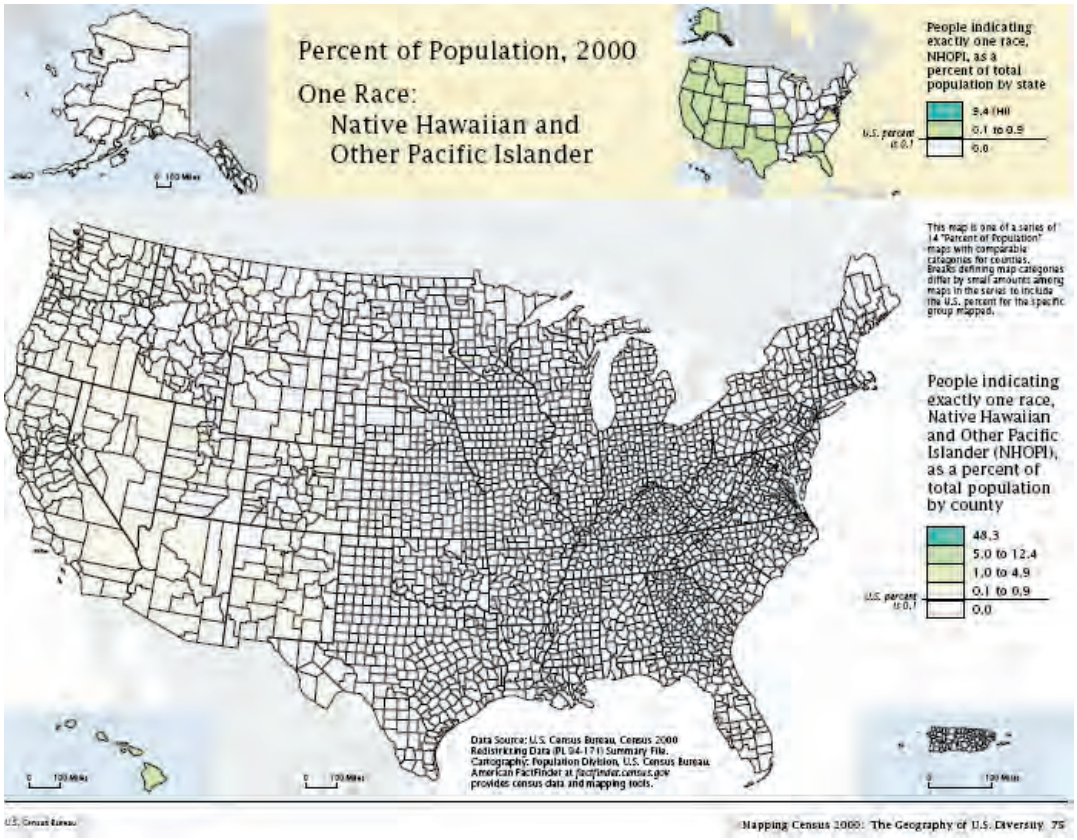
5.3 When it does not make sense to produce a map

Before starting to produce a map, you should consider whether it is the most appropriate tool to present your data. Do not waste your time and effort if a chart or a data table can provide a better way to communicate your message.

There is no point in mapping your data if:

- the data have no geographical breakdown;
- there is no significant variation in the data;
- your target audience may have difficulty understanding your map;
- there is not enough space available to present the map so it can be properly read and understood.

BAD EXAMPLE of a map



Source: Brewer, C.A. and Suchan, T.A., U.S. Census Bureau (2001), *Mapping Census 2000: The Geography of U.S. Diversity*, U.S. Government Printing Office, Washington DC[12].

The example above illustrates how a map can be wasted when there is not enough information to map. It aims to represent the spatial distribution of a small ethnic group in the United States of America: the Native Hawaiians and Other Pacific Islanders (NHOPI). The map is almost empty, because this group represents less than 1% of the total population in most counties. The few counties that have more than 1% of NHOPI cannot be seen easily.

5.4 Different types of maps

Maps can be classified according to scale, function, design, production technology or the way they are used in a publication.

In general, there are two types of maps:

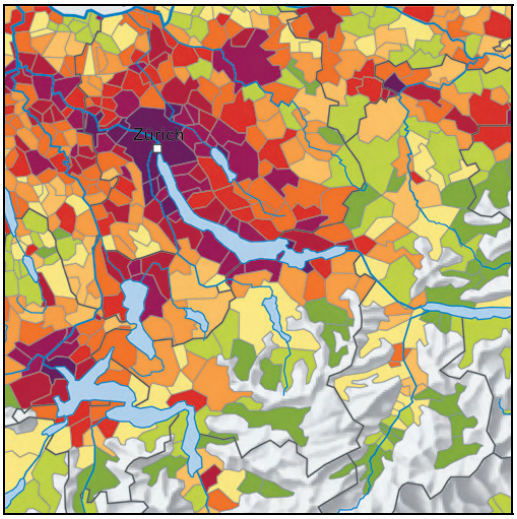
- **General reference (topographic) maps** are used to support orientation in space and show the location of a variety of different features, such as lakes, rivers, mountains, coastlines, roads, etc. They help users identify the boundaries of geographic areas.
- **Thematic (statistical) maps** are used to show the spatial distribution of one or more statistical attributes. A thematic map is always designed to serve a purpose and answer specific questions about political, social, cultural, economic, agricultural or natural phenomena.

[12] http://www.census.gov/population/www/cen2000/atlas/index.html

Example of a general reference map

Source: swisstopo, the Federal Geo-Information center[13]

Example of a thematic map



Source: Swiss Federal Statistical Office[14]

Static or interactive maps

Maps can be static or interactive. Static maps cannot be edited by the user. Interactive maps offer flexibility and give the user the ability to alter the design, select and retrieve data, animate the map, and change the topic or focus on aspects that are of key interest.

Maps can be image- or data-based. If a map is solely image-based, it is pre-produced and static in display. When you produce a data-based map, you store all the information (data and metadata) that is needed to create the map in a database. The map itself is only produced when the user requests it on the Internet. With this technique, the user can easily update and change the data and map parameters without having to re-create the map.

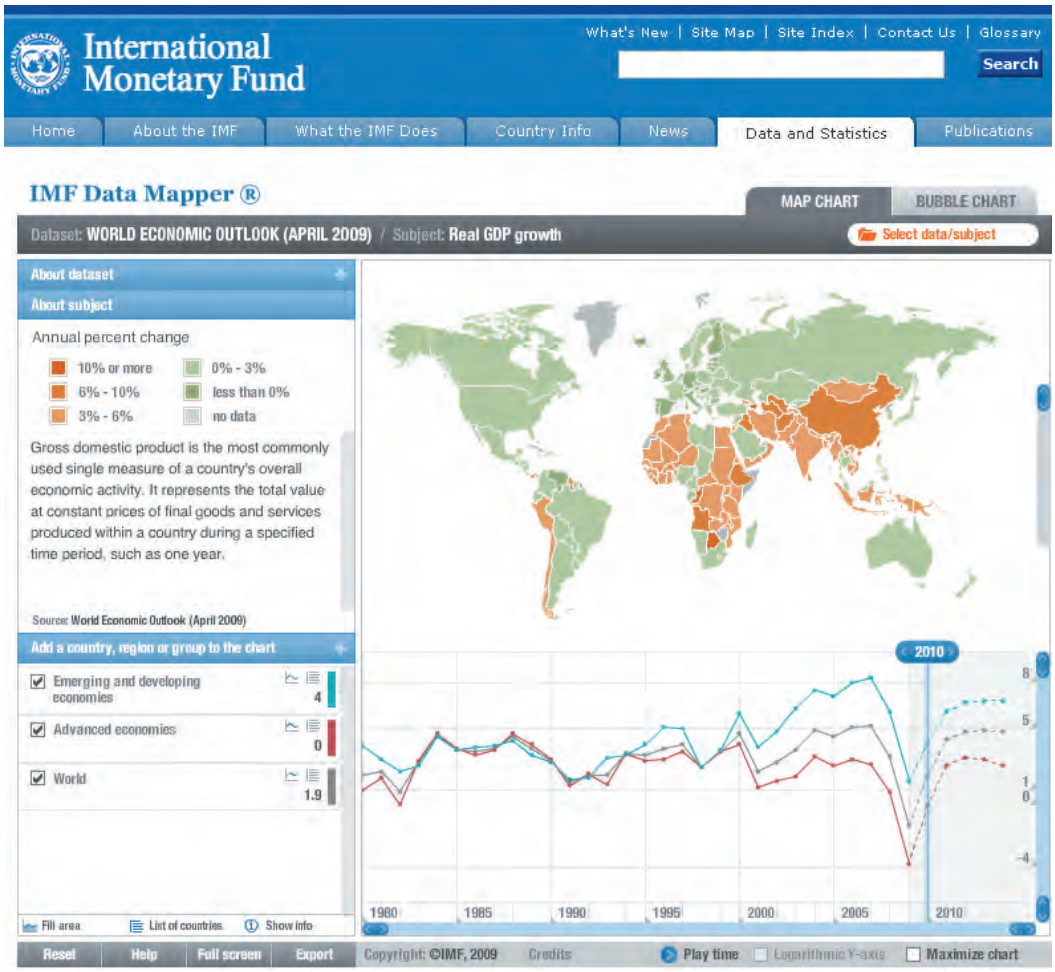
Thematic atlases

In its simplest form, an atlas is a bound collection of maps. Thematic atlases present statistics in a comprehensive way and are accompanied and enriched by valuable information in the form of text, charts and tables. Most statistical organizations recognize their potential for conveying data and are producing popular census atlases or thematic atlases (for example, on population, health or the economy).

Atlas technologies have significantly improved over the past decade. Modern online atlas information systems (AIS) allow the user to explore the data behind the maps, click on regions, “tailor” their own maps, integrate their own data, and communicate with the map author or office. Behind the scenes, new production processes have emerged that facilitate the integration of different types of expertise, such as map-making, graphic design, data analysis, writing and translating.

[13] http://www.swisstopo.admin.ch/internet/swisstopo/en/home.html
[14] http://www.bfs.admin.ch/bfs/portal/en/index/regionen/thematische_karten/maps.html

GOOD EXAMPLE of an online thematic atlas



Source: International Monetary Fund, *IMF Data Mapper*[15].

5.5 Selecting the appropriate type of map

The same advice given for charts also applies to maps: it is crucial to know what type of map to produce with what type of information. The selection of the appropriate mapping technique depends on the nature of the data. This section provides guidelines for the three most commonly used types of thematic maps: choropleth maps, dot maps and proportional symbol maps.

Choropleth maps

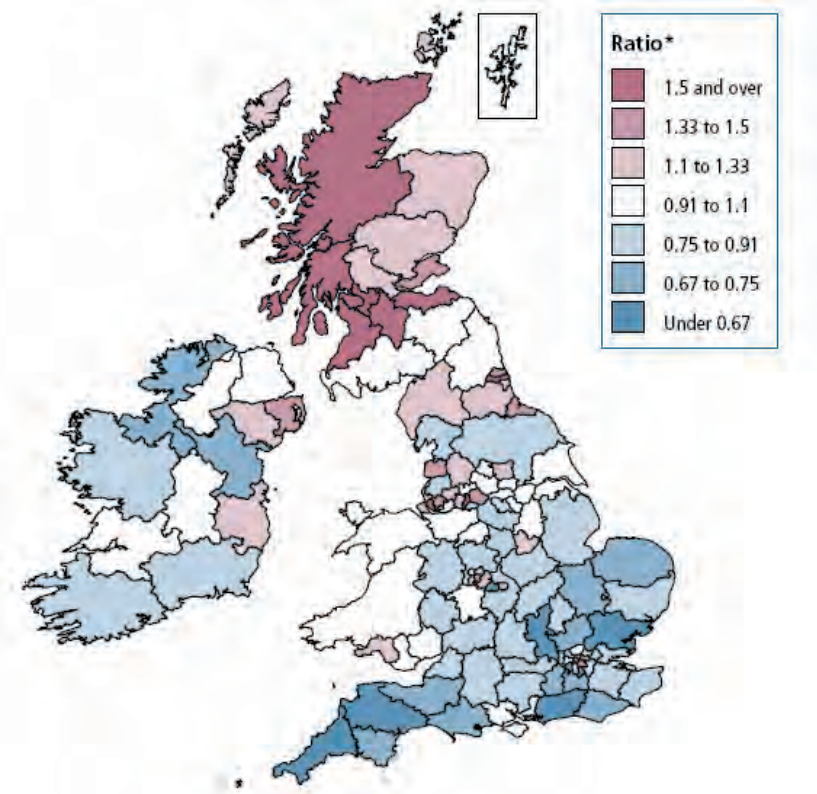
The most common type of map is the choropleth map, in which areas are shaded in proportion to the value of the variable being displayed. This kind of map provides an easy way to visualize patterns across space, as shown in the example below.

The abundance of choropleth data and the ease of design using GIS has unfortunately led to frequent misuse of the choropleth mapping technique. Only ratios (i.e. proportions, rates or densities) can be mapped with this technique. You should NOT use it to represent absolute values, such as population size.

[15] http://www.imf.org/external/datamapper/index.php

GOOD EXAMPLE of a choropleth map

Cancer of the larynx: incidence by health authority
Males, UK and Ireland, 1991-99



** Ratio of directly age-standardised rate in health authority to UK and Ireland average*

Source: Office for National Statistics (2005), *Cancer Atlas of the United Kingdom and Ireland 1991-2000*, London[16].

A choropleth map is based on data aggregated over pre-defined areas, such as the United Kingdom and Ireland health authorities in the example above. This map type is well suited for discussion about these health authority regions. However, real-world patterns often do not conform to pre-defined areas, which can lead to major misinterpretation. Choropleth maps should, therefore, preferably be used to display phenomena that are evenly distributed within each spatial unit.

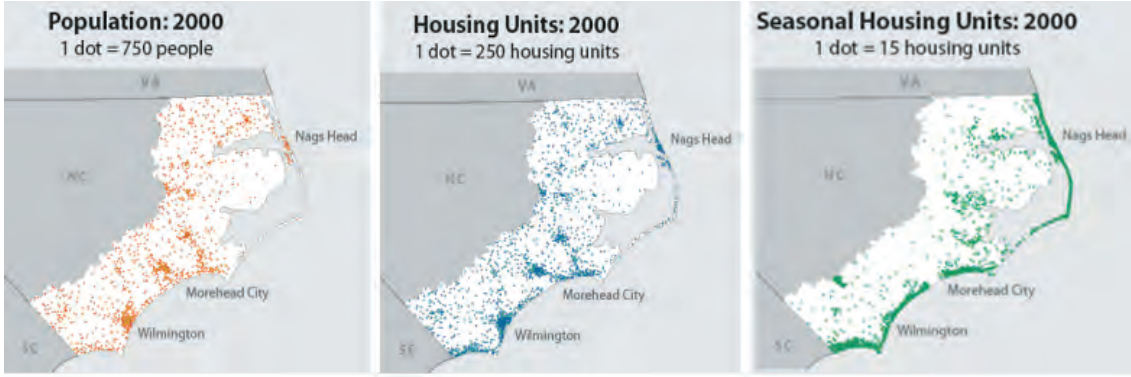
Data classification is a major issue with the choropleth technique. The spatial patterns displayed in the map are determined by the grouping of the data values. You can produce different maps with the same data by applying different classification methods, dividing the data into different numbers of classes or selecting different class limits. There are many techniques, but no rules, so you should determine the most appropriate method to classify your data.

Dot maps

A dot (or point) map displays the location and density of a population or phenomenon using symbols. It enables users to quickly grasp the general magnitude of the data, as well as their concentration or dispersion. Each dot represents a discrete value, usually a large number of entities as shown in the examples below.

[16] http://www.statistics.gov.uk/statbase/Product.asp?vlnk=14059

GOOD EXAMPLE of dot maps



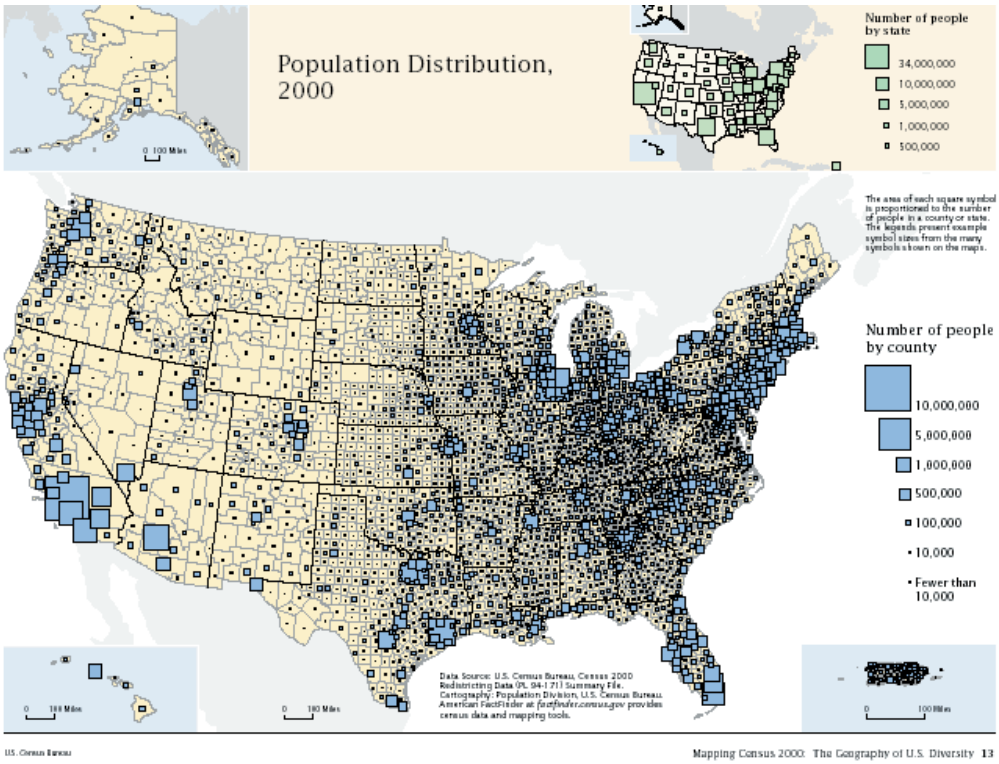
Source: U.S. Census Bureau, *Census Data and Emergency Preparedness*[17].

Proportional symbol maps

A proportional (or graduated) symbol map is used to display absolute values. The size of the symbol is proportional to the size of the population or phenomenon being represented. Each symbol is attached to a specific point within the spatial unit, usually either the centre or the capital.

Circles are most commonly used, because they are compact and easy to scale. But other geometric shapes, such as squares or triangles, can also be used, as shown in the example below.

GOOD EXAMPLE of a proportional symbol map



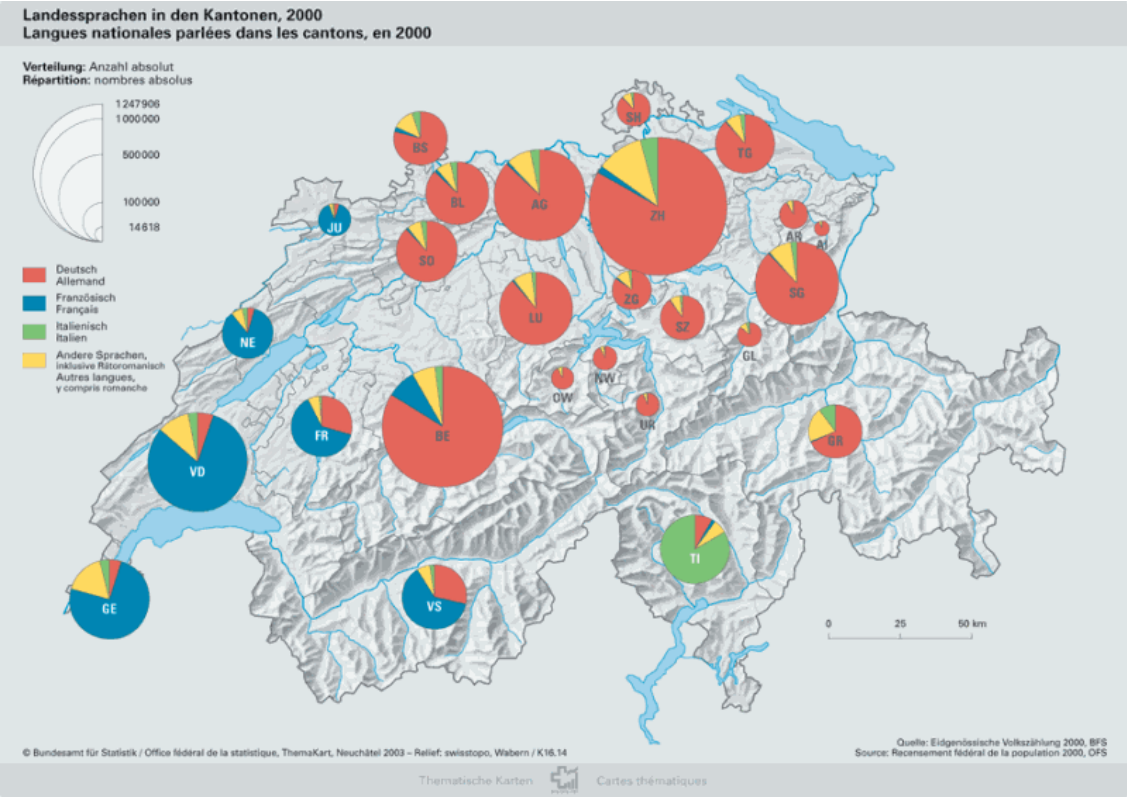
Source: Brewer, C.A. and Suchan, T.A., U.S. Census Bureau (2001), Mapping Census 2000: The Geography of U.S. Diversity*, U.S. Government Printing Office, Washington DC*[18].

[17] http://www.census.gov/Press-Release/www/emergencies/
[18] http://www.census.gov/population/www/cen2000/atlas/index.html

The example above shows that expert mapping skills may be required to create a well-designed proportional symbol map. When the size of the symbol is bigger than the size of the corresponding spatial unit, it may be difficult to identify the unit that the symbol refers to. This difficulty of interpretation becomes even harder when many symbols overlap, as illustrated in this map of the population distribution in the United States of America.

Chart and mapping techniques can be combined to display the distribution of the different categories of a population on the same map. In complex symbol maps, pie charts or bar charts are used as symbols. In the map of Switzerland below, the pie chart represents the share of the population speaking each of the main national languages (German, French or Italian) or any other language in the 26 cantons. The size of the symbol represents the total population of each canton.

GOOD EXAMPLE of a complex symbol map



Source: Swiss Federal Statistical Office, *Map Gallery Switzerland – Languages and religions*[19].

This complex technique should be used with great care, as the map and the legend can easily become overloaded. Such maps can only be produced at the national level (e.g. countries in Europe) or regional level (e.g. Swiss cantons in the map above). Moreover, only a small number of categories (not more than five) can be displayed in the pie or bar chart.

[19] http://www.bfs.admin.ch/bfs/portal/en/index/regionen/thematische_karten/maps.html

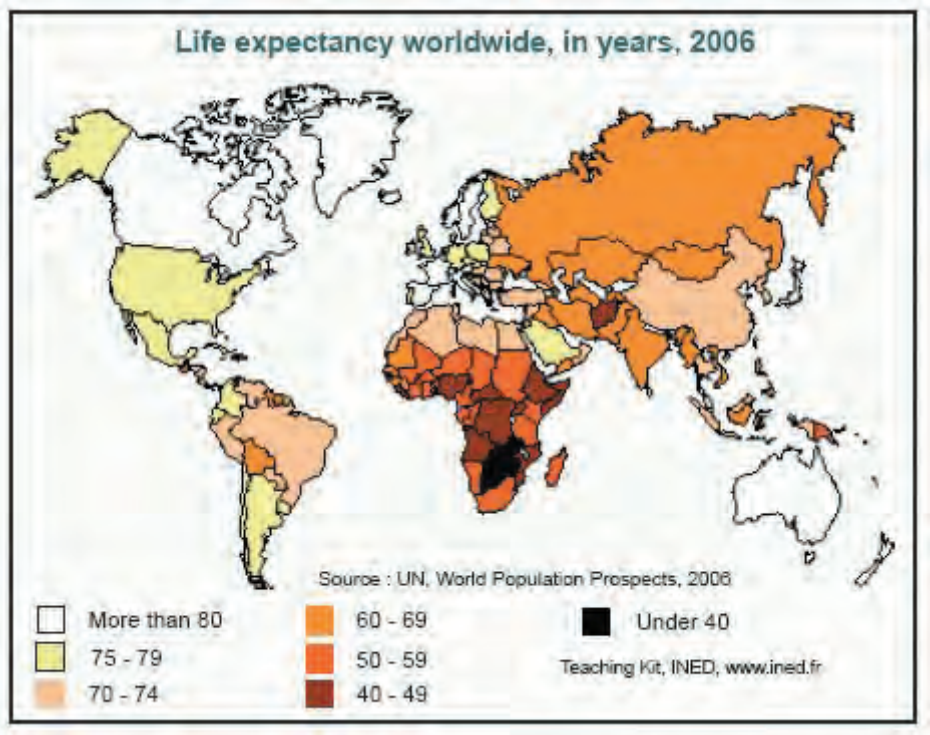
5.6 Design tips: keep it simple!

As for all data visualizations, the most important recommendation to ensure that your message gets across is: **keep your map simple!** Be careful not to distract or confuse your audience by displaying too much information or too many visually conflicting elements.

Knowing your audience is also of crucial importance. What is their background? Are they familiar with maps in their everyday life? Could your map offend them in any way? You should be aware of the sensitivities of your target audience. Maps have a powerful visual impact, and some colours or symbols may have a negative connotation for some people.

You should always design your maps so they are **independent of story texts or data tables**. Just like a chart, a map should be understood on its own, without further reference to surrounding text or notes. Once a map is published, it can be scanned or downloaded and used out of context. It is, therefore, essential to include the different map components described below.

GOOD EXAMPLE of a map that does not require any further reference



Source: Institut National d'Etudes Démographiques (INED)

Map components

Map components compete with each other for the reader's attention. To maximize the efficiency of your message, ensure that your data are the focus of your layout, especially when other information (water, altitude, etc.) are also presented. The map should cover 80-85% of the total presentation space.

The following components are needed to assist the user in understanding the map:

- The **map title** should give a clear idea of what the map is about. It needs to be short and concise. Subtitles may be added to provide more detailed information (e.g. unit of measure).
- The **legends** should identify all the symbols, patterns and colours used to represent the data in the map.
- The **geographic units** at which the data are represented in the map should be identified, either in the title (or subtitle) or in a legend.
- **Text labels** may be added on the map to identify important or relevant places or other information.
- The **map scale** may be provided to help the user measure distances and compare different maps.
- **A footnote** may be used to provide definitions or other methodological information.
- The **data source** should be identified at the bottom of the map.
- The **copyright information** should identify the author responsible for its content at the bottom of the map.

Other components may be included in some maps, but they are not essential:

- A **north arrow** is only needed when the map is not orientated to the north.
- **Latitudes and longitudes** are only needed on world or continental maps.
- A **location map** is a small replica of the base map that sets the mapped area in its wider context. It may be useful if your audience is not familiar with the geography of the area.
- **Charts** may be added if they enhance the understanding of the map.

Use only key words in title and legends

Although maps communicate a visual message, the associated text lines are also important. The wording of the title and any legend should be considered carefully, as it determines the user's understanding and interpretation of your map. Here are some guidelines for the usage of text in maps:

- Be accurate, but keep it simple.
- Use only key words and avoid repeating the same words in the title, legends or footnotes.
- Use neutral language.
- Avoid abbreviations and acronyms.
- Use smaller font for the legends than for the title and even smaller font for the footnotes, but ensure that all text lines are legible.

Design the legend carefully

The design of your legend must ensure correct understanding and interpretation of your map. Each map type requires a different type of legend. But there are a few basic rules for choropleth and proportional symbol maps:

- All class limits should be unambiguous: avoid ranges such as 100-200, 200-300, 300-400.
- There should be no gap between classes: avoid ranges such as 1.0-1.5, 2.0-2.5, 3.0-3.5.
- Areas for which data are not available should be identified.
- In a map representing more than one variable, the legend should be displayed in descending order of importance of the variable.

Colour: another important choice

Colour is one of the most powerful graphic features. Select with great care the colours you use in your map, as they can influence and mislead.

The choice of colour in a thematic map depends on the data and on the type of map. But three other aspects should also be considered. First, you should be aware of any existing convention associated with any given colour, as well as possible positive or negative connotations.

Then, you should make sure that everyone can understand your message with the colours used in your map. For example, colour-blind people have difficulty distinguishing between some colours. The most common case is red-green blindness. If you use red and green to show a difference between two types of areas, such as growth and decline, colour-blind people will not see it. There is an easy solution: these people will see the difference if you replace red with purple.

Finally, when there are relatively few classes of data for values on a continuous scale (e.g. population density), you should consider using different shades of the same colour rather than different colours. If you have data in discrete classes, or with positive and negative values, different colours are often more appropriate.

6. Emerging visualization techniques

6.1 Why visualization is more than just a picture

Emerging tools and techniques are providing new opportunities for visualizing data and making them more interesting to users. Dynamic table, chart and map generators allow users to manipulate data and create their own visualizations. Animation and video are engaging formats, somewhat like television. They do a good job of illustrating changes over time and include verbal or textual descriptions to explain the meaning behind the numbers. New types of visualizations are also emerging, such as sparklines and tag clouds (see section 6.4), providing alternative ways to illustrate information.

The latest web technologies, and the expectations they create amongst the user community, are changing the way statistical organizations communicate statistics. The Internet is now a two-way communication forum where users can share their own data visualizations and discuss their findings. Websites such as Many Eyes[20], Swivel[21] and Data Place[22] are examples of online communities that discuss and share data and graphics.

Whilst increasing user flexibility, the development of new visualization techniques and more interactive websites can also cause problems for statistical organizations. It is becoming increasingly easy for users, whether by accident or design, to distort or misrepresent statistics, and then make these distortions and misrepresentations widely available to others. It is therefore necessary for statistical organizations to have a clear policy on how they will apply and offer new visualization techniques.

This chapter provides an overview of these new visualization tools and techniques.

6.2 Dynamic visualizations

With the advent of the Internet and growth of Web 2.0 technologies[23], users can interact with data and create their own visualizations. Many statistical organizations now provide access to their databases through their websites, allowing users to query and download statistical information themselves. This functionality is increasingly being complemented by a suite of visualization tools that allow users to create tables, charts or maps online, without having to download the data and work in another application.

There are concerns about the consequences of giving this level of control to users. There is the possibility that they will create nonsensical graphics or derive inappropriate correlations. However, it is surely preferable that users are accessing and working with the data. Potential problems can be minimized by providing key metadata in a clear and obvious way, offering support to less experienced users, and by monitoring and correcting any misuse.

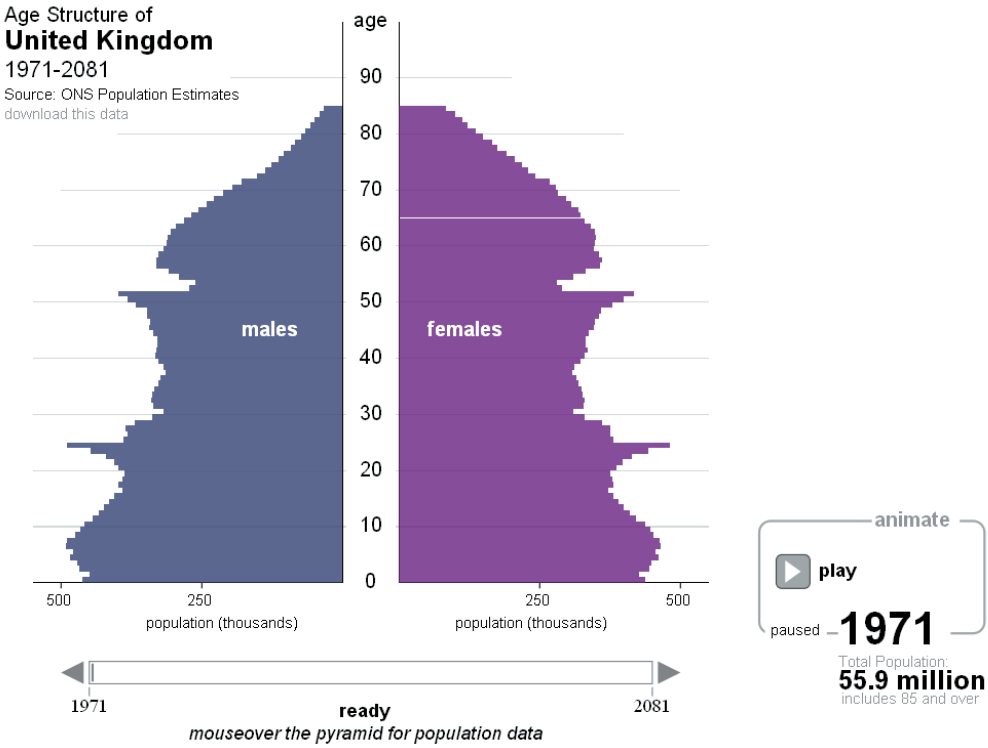
---

[20] Many Eyes is a website where users can upload data, create charts and other visualizations and discuss their findings (see http://manyeyes.alphaworks.ibm.com/manyeyes/).
[21] Swivel offers similar functionality to Many Eyes (see www.swivel.com).
[22] Data Place is a website that provides users with statistics on cities, towns and states in the United States (see www.dataplace.org/).
[23] "Web 2.0" is a term to describe a new wave of internet technologies that allow users to do more than just access information online: they can add, change or influence web content. Examples include wikis, such as Wikipedia, blogs and social networking sites, such as Facebook or LinkedIn.

6.3 Animation and video

Animation and video are two important emerging data visualization techniques. When you consider the popularity of television and film, it is not surprising that users like the idea of receiving messages via moving pictures. This format makes it easier to tell the story, by combining audio or textual descriptions with graphical illustrations to communicate the meaning behind the numbers.

GOOD EXAMPLE of using animation to communicate statistics



Source: Office for National Statistics, United Kingdom[24]

The dynamic population pyramids developed by several statistical organizations, including the UK Office for National Statistics and Statistics Canada, are good examples of combining animation with interactivity in a simple interface. Users can click on play to watch how the shape of the population pyramid changes over time. They can interact with the chart by selecting age groups and noting the detailed numbers and proportion of total population.

Hans Rosling, co-founder of Gapminder[25], has had great success with using animation to illustrate data, as a way of communicating statistics. Rosling has achieved a massive audience through online video, an increasingly popular feature of the Internet. His presentation at TED Conference[26] in 2006 has been viewed online and downloaded thousands of times. He has built upon the popularity of this medium for communicating statistics by developing "gapcasts", which are short video lectures on issues, such as maternal mortality, globalization, energy and human development trends.

[24] http://www.statistics.gov.uk/populationestimates/svg_pyramid/uk/index.html
[25] http://www.gapminder.org/
[26] TED stands for Technology, Entertainment, Design. It is an annual conference that attracts leading people to speak on current issues in these fields. The presentations are published on their website at www.ted.com.

GOOD EXAMPLE of combining animation and new web technologies to communicate statistics

Source: Gapminder[27]

6.4 Web 2.0 and building communities around data

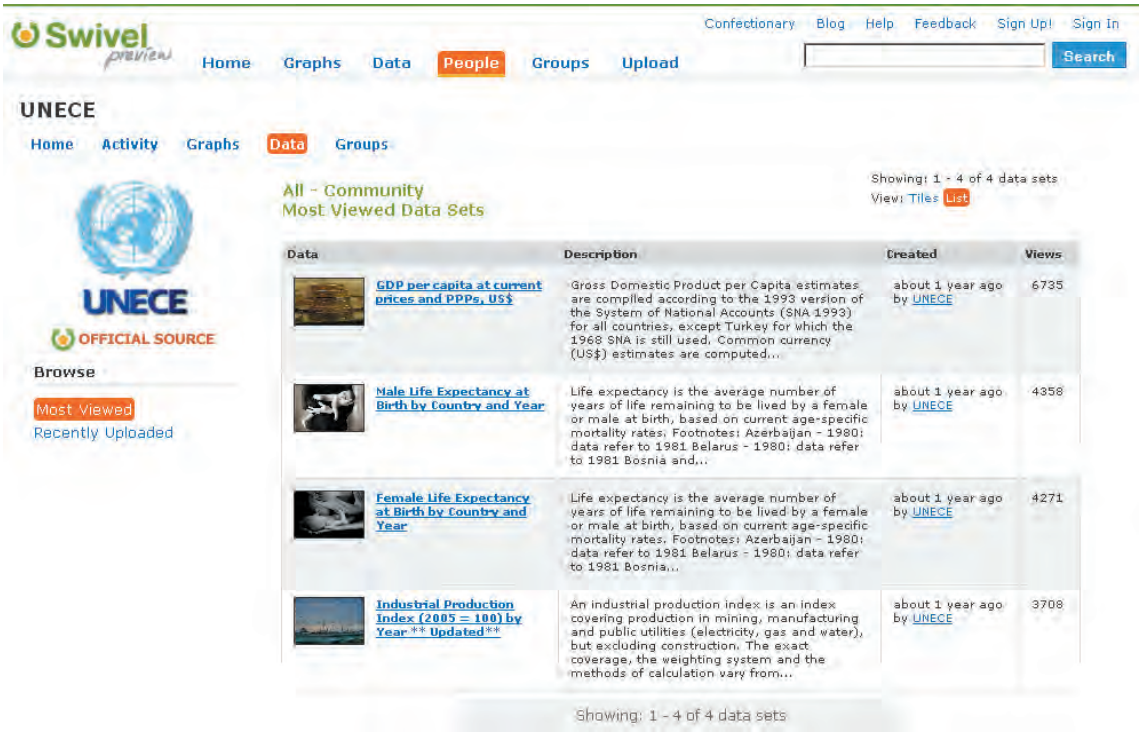
Websites such as Many Eyes, Swivel and Data Place are adding a new dimension to visual presentations by building online communities around data visualization and sharing. These websites allow users to upload datasets and create graphics for sharing and discussion with other users. Other applications, often referred to as "mashups", combine data or functionalities from two or more sources to create a new service. An example of a mashup is the combination of statistical and cartographic data to create a geographic view of data[28].

Some producers of official statistics are experimenting with the potential of these Web services for communicating with a wider audience. For example, the UNECE has uploaded four datasets to Swivel, with mixed results. On the positive side, these datasets have each been viewed over 5 000 times during the first 18 months, indicating that the data have reached many users. However, they have only generated two user comments and the links to the UNECE statistical database website have only been followed 10 times during this period, so virtually no information has been gathered about these new users.

[27] http://www.gapminder.org/video/gap-cast/

[28] For example, the latest versions of the PC-Axis suite of statistical dissemination software (http://www.pc-axis.scb.se/) provide the option to combine data with both Google Maps and Google Earth.

GOOD EXAMPLE of using the new Web services to reach wider audiences



Source: Swivel[29]

Although success has been mixed so far, this sort of online community clearly provides a relatively easy way to reach additional users and is therefore an emerging area of data visualization worth following.[30]

6.5 Other new visualization techniques

Sparklines

Sparklines are small, word-sized line charts that show trends over time. They have the benefit of showing a great deal of information at a glance and can be placed alongside words that explain their meaning.

Tufte (2006) first proposed sparklines. The following example shows sparklines used to illustrate fluctuations in the Euro exchange rate against other currencies. These "intense, simple, word-sized graphics" enhance tabular data with a visual representation without taking up much space.

GOOD EXAMPLE of a sparkline

Source: Tufte, E.R. (2006), *Beautiful Evidence*, Cheshire CT, Graphics Press.

---

[29] http://www.swivel.com/users/show/1005968

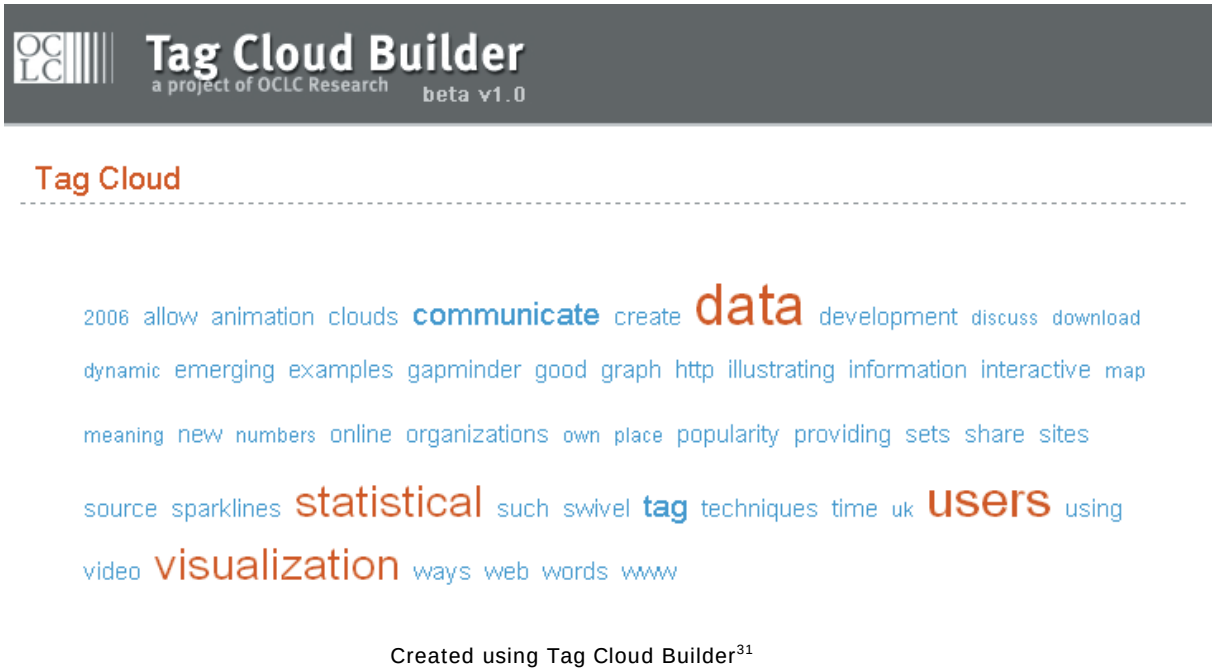
[30] For further reference, see the special issue of the *Statistical Journal of the IAOS*, vol. 25, nos 3-4, 2008: "Web 2.0 and Official Statistics", available at: http://iospress.metapress.com/content/v0376364l348/?p=fc2e171758ee4053a01be16bbbae10eb&pi=0

Tag clouds

A tag cloud (sometimes also called a word cloud) is a visual representation of the frequency of a word or tag in a particular text or dataset. Varying size and colours illustrate the popularity or importance of tags. They are often seen on websites as a list of categories, with each word being a clickable link that takes the user to more information relating to that word.

Tag clouds are a useful way to identify common terms in text and build keyword taxonomies. The example below has been created using the text from this chapter, clearly illustrating the key words.

GOOD EXAMPLE of a tag cloud



Created using Tag Cloud Builder[31]

[31] http://tagcloud.oclc.org/tagcloud/TagCloudDemo

7. Accessibility issues

For basic statistical information to be broadly and easily used, it must be universally accessible. This means everyone must be able to consult and understand the information, regardless of the technology they use or any disabilities they may have. Therefore, an important goal of a communications strategy is to generate information that is accessible.

To attract and reach large audiences, the information should be presented in a structured and well organized way that meets established and generally recognized standards. Three partly overlapping aspects of accessibility are considered in this chapter:

- Providing suitable metadata to help users understand the data;
- Providing data in a range of formats, including via new media such as mobile phones or other hand-held devices;
- Ensuring the maximum accessibility to information for people with disabilities, often in compliance with specific legal or policy requirements.

Making statistical information accessible to everyone requires extra awareness, extra work and extra resources. However, the principle of equal access to basic statistical information is an important one. A well thought-out accessibility strategy will benefit everyone.

In this chapter, we discuss the main components of good information that should be considered when communicating statistical information.

7.1 Text

Text must be concise, coherent and well-structured so users can easily find the information they are looking for. Presenting your text in distinct, logical sections with headings and subheadings will make searching it and converting it to other formats easier.

To serve all target groups effectively, text should be available in multiple formats, e.g. Braille, audio or oversized letters. When creating the original text, you should anticipate possible format conversions to minimize their impact. For example, adding a text description for each piece of graphic information will reduce the task of converting it to Braille, audio or even HTML[32] format for a website.

Accessibility rules developed by the international World Wide Web Consortium (W3C)[33] were designed to ensure access to information in electronic format using adaptive technologies. These include screen readers, character enlargement systems and portable data access devices (PDA), to name but a few.

[32] Hyper Text Mark-up Language.
[33] http://www.w3.org/

7.2 Tables

You should also pay attention to accessibility when using tables to present your data. If the data are displayed without all the information needed to understand and interpret them, they are unusable or misleading.

Consider, for example, a table spread over several pages. If the column headers or row stubs are no longer visible, the user will hardly be able to read the table. It is therefore important to repeat the column headers and row stubs on each new page.

GOOD EXAMPLE of a table that meets accessibility requirements

Livestock inventory in Canada, 2007 and 2008

Type of livestock	2007		2008	
	Thousands of head	%	Thousands of head	%
Cattle	15 885	50.2	15 195	52.0
Hogs	14 690	46.4	12 985	44.4
Sheep	1 096	3.5	1 062	3.6
TOTAL	**31 671**	**100**	**29 242**	**100**

Source: Canadian Cattle Statistics August 2008[34]

Each value is not only associated with a type of livestock and a year, but it also refers either to the number of head (in thousands) or the percentage. A screen reader must be able to give all this information for each value in the table. Users would therefore hear: “In 2007, the percentage of cattle is 50.2”.

In electronic format such as HTML, tags can be included for column headers and row stubs, along with a short description of their content when necessary. Using tags is good for everyone. When tables are downloaded, the column headers and row stubs are transferred directly, avoiding confusion. Moreover, when the structure of your table is complex, it is important to describe precisely how the data are organized so users can get the maximum information with a minimum of effort.

7.3 Charts

When producing charts, it is worth considering that not all users have access to pictures. A text description can be created to provide the same information as displayed in your chart. Below is an example from the National Braille Association Tape Recording Manual.

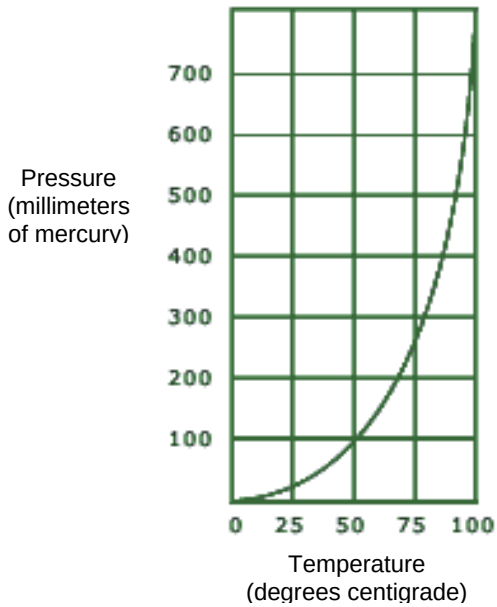
On a website, the use of an “alt text” tag (a brief text equivalent) or a “longdesc” attribute (a long description stored on a different page and accessible through an e-mail address) are techniques recognized and described by W3C. These descriptions may also be used in the production of documents in audio, Braille or other formats.

You may have so many charts that you have difficulty producing a text equivalent manually. Projects are underway to design tools that will be able to automatically extract information from a chart. They will also be able to generate a descriptive page at the same time as the chart is created.

[34] http://www.thedairysite.com/articles/1613/canadian-cattle-statistics-august-2008

GOOD EXAMPLE of a chart with a text equivalent

Relationship between the vapor pressure of water and its temperature



"The relationship between the vapor pressure of water and its temperature." This is a line graph whose x-axis is temperature in degrees centigrade, running from zero to one hundred degrees. The y-axis is pressure in millimeters of mercury and runs from zero to 800 millimeters. The curve starts at the origin and rises so that when x is 25 degrees, y is approximately 40 millimeters. When x is 50, y is 100. When x is 75, y is just under 300. When x is 100, y is about 760."

Source: Accessible Digital Media: Design Guidelines for Electronic Publications, Multimedia and the Web[35].

7.4 Maps

Maps constitute major technical challenges in terms of accessibility for people with certain disabilities. Information is transmitted by a combination of images and colours, two methods virtually incompatible with accessibility standards. We should therefore consider offering a text alternative to provide the same information as represented in the map and/or give access to the data tables.

Another option is to provide a search tool that would let users select parameters, such as an area, a town or a street. A database search could be something like this: “Search the number of male individuals aged 15 to 49 living in a specific area and working in the agro-industry sector”. The results should be generated in the form of a statistical table providing the same information as shown visually.

[35] http://ncam.wgbh.org/publications/adm/guideline_f.html

GOOD EXAMPLE of a map accessible to blind people

Braille map of a portion of Washington DC

Source: US Geological Survey Library, courtesy of Flickr[36]

7.5 Metadata

Statistical organizations should ensure that users are provided with the metadata they need to understand the data, including their strengths and limitations. These metadata must be kept up-to-date by incorporating the latest changes in definitions, classifications and methodology.

You should consider the following recommendations[37] to ensure that your metadata can be accessed and used by everyone:

- Provide access to the metadata via a range of different media, such as paper publications, CD-ROMs, etc. All metadata should be available on the Internet, as it is the most accessible medium for users to find the most up-to-date metadata.
- Present the metadata in such a way as to meet the needs of a range of users with different requirements and/or statistical expertise.
- Disseminate the metadata free of charge on the Internet, even if there is a charge for paper versions or for the statistics they describe.
- Ensure active linkage of the metadata to the tables and charts they describe - and vice versa.
- Make metadata available not only in the national language but also, if possible, in a common language such as English.
- Provide a local search engine based on free text search.
- Implement a mechanism to ensure the stability of URLs (Uniform Resource Locators) or providing links between the old and new URLs that will redirect users to the new address. This is a key issue given the importance of links between websites.

[36] http://www.flickr.com/photos/98169608@N00/3296197787/

[37] The guidelines for the reporting and dissemination of metadata are adapted from those included in the 2007 OECD publication "Data and Metadata Reporting and Presentation Handbook", available at http://www.oecd.org/bookshop?9789264030329.

- Provide the names of contact persons or email addresses where further information about concepts, definitions and statistical methodologies may be obtained. In some organizations, the “contact” would be a generic corporate contact point or referral service for all client enquiries.

Some statistical organizations present metadata as layers of a pyramid. With this approach, metadata become more detailed as the user moves down from the top of the pyramid:

- **At the top of the pyramid:** metadata that are essential for a basic understanding of the statistics, including information on the status of the data (preliminary or final). Casual users will rarely go beyond this type of metadata.
- **Below:** explanatory notes that provide a brief description of the statistics (definitions, key issues, limitations, etc.) that can impact the use of the data. Intermediate users may consult this type of metadata.
- **At the base of the pyramid:** the most detailed methodological information, typically in the form of methodological manuals or similar documents. This type of metadata is likely to be of interest only to the most experienced users and those using the data for detailed research.

8. References and further reading

Bertin, J. (1981), Graphics and Graphic Information-processing, New York, de Gruyter.

Bertin, J. (1983), *Semiology of Graphics: Diagrams Networks Maps*, Madison Wis., University of Wisconsin Press.

Blessing, C., Bradsher-Fredrick, H., Miller, H., Miller, R. and Rutchik, R. (2003), *Cognitive Testing of Statistical Graphs: Methodology and Results*, Washington D.C., U.S. Energy Information Administration.

Bosch ten, O. and Jonge de, E. (2008), "Visualising official statistics", in *Statistical Journal of the IAOS*, vol. 25, nos 3-4, pp. 103-116, available at: http://iospress.metapress.com/content/v0376364l348/?p=fc2e171758ee4053a01be16bbbae10eb&pi=0.

Cleveland, W.S. and McGill, R. (1984), "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods", in *Journal of the American Statistical Association*, 79, pp. 531-554, available at: https://secure.cs.uvic.ca/twiki/pub/Research/Chisel/ComputationalAestheticsProject/cleveland.pdf.

Cleveland, W.S. and McGill, R. (1987) "Graphical Perception: The Visual Decoding of Quantitative Information on Statistical Graphs (with Discussion)", in *Journal of the Royal Statistical Society Series A*, 150, pp. 192-229, available at: http://www.wjh.harvard.edu/~kwn/Kosslyn_pdfs/1987Cave_JRoyStatSocA_CommentaryClevelandArticle.pdf

Few, S. (2004), *Show Me the Numbers: Designing Tables and Graphs to Enlighten*, Oakland CA, Analytics Press.

Gardner, J. (2008), "Blogs, wikis and official statistics: New perspectives on the use of Web 2.0 by statistical offices", in *Statistical Journal of the IAOS*, vol. 25, nos 3-4, pp. 81-92, available at: http://iospress.metapress.com/content/v0376364l348/?p=fc2e171758ee4053a01be16bbbae10eb&pi=0

Harris, R.L. (2000), *Information Graphics*, New York and Oxford, Oxford University Press.

Kennedy, D. (2007), *Research Paper: Data Visualization*, Canberra, Australian Bureau of Statistics, available at: http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/7F8E375FC22D26A5CA25731C0022DF1B/$File/1211055001_jul%202007.pdf.

Miller, J.E. (2004), *The Chicago Guide to Writing About Numbers*, Chicago, University of Chicago Press.

Playfair, W. (1786), *The Commercial and Political Atlas: Representing, by Means of Stained Copper-Plate Charts, the Exports, Imports, and General Trade of England, at a Single View*, London.

Playfair, W. (1801), *Statistical Breviary: Shewing, on a Principle Entirely New, the Resources of Every State and Kingdom in Europe*, London, Wallis.

Robbins, N.B. (2005), *Creating More Effective Graphs*, Hoboken NJ, John Wiley & Sons.

Schulz, T. (2009), *Guidelines on the Presentation of Statistical Maps*, UNECE Work Session on the Communication and Dissemination of Statistics, Warsaw, available at: http://www.unece.org/stats/documents/ece/ces/ge.45/2009/crp.1.e.pdf.

Smith, A. and Rogers, S. (2008), "Web 2.0 and official statistics: The case for a multi-disciplinary approach", in *Statistical Journal of the IAOS*, vol. 25, nos 3-4, pp. 117-123, available at: http://iospress.metapress.com/content/v0376364l348/?p=fc2e171758ee4053a01be16bbbae10eb&pi=0.

Tufte, E.R. (2001), *The Visual Display of Quantitative Information (2nd ed.)*, Cheshire CT, Graphics Press.

Tufte, E.R. (2006), *Beautiful Evidence*, Cheshire CT, Graphics Press.

Wainer, H. (1984), "How to Display Data Badly?", in *The American Statistician*, vol. 38, no. 2, pp. 137-147, available at: http://www.soc.washington.edu/users/bpettit/soc504/wainer_display.pdf.

Ware, C. (2004), *Information Visualization: Perception for Design, San Francisco CA*, Morgan Kaufmann Publishers.

Further useful information and discussion on the presentation of statistics can be found at: http://blogstats.wordpress.com/.

Making Data Meaningful

A guide to presenting statistics

A picture is worth a thousand words. Patterns in data are often more clearly revealed when you see the numbers presented as a picture. There are many ways to present data, from simple bar charts to more complex scatterplots, thematic maps and animated graphics.

This guide is a practical tool to help producers of statistics present data in a clear and meaningful way. It provides advice on preparing effective tables, charts and maps, and using other forms of visualizations to bring statistics to life. It also suggests how to avoid bad or misleading visual presentations. Clear presentations of data, prepared with the target audience in mind, will increase the use of statistics and unlock the valuable information contained inside.