



New  
Issue

**1211.0.55.001**

**Research Paper**

**Data Visualisation**

**Australia**

**July 2007**



**Research Paper**

**Data Visualisation**

**Australia**

**July 2007**

**Damien Kennedy  
Integrated Collection and Dissemination Services  
Division  
Australian Bureau of Statistics**

**This Research Paper is intended to make the results of current research available to other interested parties. The aim is to present accounts of developments and research work or analysis of an experimental nature so as to encourage discussion and comment.**

**Views expressed in this paper are those of the author and do not necessarily represent those of the Australian Bureau of Statistics. Where quoted or used, they should be attributed clearly to the author.**

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30AM (CANBERRA TIME) THURS 19 JUL 2007

ABS Catalogue No. 1211.0.55.001

ISBN 9780642483560

© Commonwealth of Australia 2007

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced by any process without prior written permission from the Commonwealth. Requests and inquiries concerning reproduction and rights in this publication should be addressed to The Manager, Intermediary Management, Australian Bureau of Statistics, Locked Bag 10, Belconnen ACT 2616, by telephone (02) 6252 6998, fax (02) 6252 7102, or email: <intermediary.management@abs.gov.au>.

In all cases the ABS must be acknowledged as the source when reproducing or quoting any part of an ABS publication or other product.

Produced by the Australian Bureau of Statistics

## INQUIRIES

- For further information about the contents of this publication contact Damien Kennedy on Canberra (02) 6252 7547.

# CONTENTS

---

*page*

## CONTENTS

Executive Summary . . . . .	1
What is Data Visualisation? . . . . .	2
Present Use of Data Visualisation by the ABS . . . . .	3
Data Visualisation and statistical stories . . . . .	4
Applications of Data Visualisation . . . . .	5
Conclusion . . . . .	42
Reference List . . . . .	43



## EXECUTIVE SUMMARY

---

### EXECUTIVE SUMMARY

The aim of this research paper is to analyse existing data visualisation tools and techniques. It will seek to highlight these examples in a practical context, incorporating issues of statistical storytelling and benefits for displaying statistical data. It will build upon the previously released paper, "*Data Communication - Emerging International Trends and Practices of the Australian Bureau of Statistics.*"

This paper should not be viewed from a technical perspective but rather from a design point of view. Although reference will be made to various technical applications this does not fall within the primary scope of this paper.

## WHAT IS DATA VISUALISATION?

---

### WHAT IS DATA VISUALISATION?

Data visualisation is defined as "the set of techniques used to turn a set of data into visual insight. It aims to give the data a meaningful representation by exploiting the powerful discerning capabilities of the human eye. The data is displayed as 2D or 3D images using techniques such as colorization, 3D imaging, animation and spatial annotation to create an instant understanding from multi-variable data." (Fisher, 1999) A more concise explanation may be that of Palace (1996) as "the visual interpretation of complex relationships in multidimensional data." Essentially it is presenting data in a way that is aesthetically pleasing to the eye and has the ability to inform and provide value to the user.

Data visualisation deals with the concept of perception. The main issue with perception is "how humans attach meaning to the sensory information they receive" (Australian Bureau of Statistics, 2006). It could be suggested that attaching meaning refers to the ability to create and develop stories that are associated with the data or that arise depending on how it is displayed. This perception is critical as the greater amount of meaning that a user can obtain, the more likely they are to use and store this information.



## PRESENT USE OF DATA VISUALISATION BY THE ABS

---

### PRESENT USE OF DATA VISUALISATION BY THE ABS

Currently the standard or most common form of graphical displays used by the ABS on its website and within publications are static two-dimensional black and white images. These generally take the form of column, line or dot graphs and are developed using internal software packages. All ABS published graphs are subject to publishing guidelines.

The ABS has recently introduced a number of new data visualisation techniques to its website. These include:

- *Animated Population Pyramids*
- *National Regional Profiles - By Location Map*
- *Mapstats*
- *Microdata Usage Graphs (Dashboard)*

## DATA VISUALISATION AND STATISTICAL STORIES

---

### DATA VISUALISATION AND STATISTICAL STORIES

A pivotal concept behind the introduction of data visualisation is the aim to produce statistical stories. "A statistical story is one that doesn't just recite data in words. It tells a story about the data" (UNECE, 2006). Statistical data is only useful or relevant if knowledge or stories are gained or interpreted from it. An audience is more likely to learn an idea within a story rather than remember the actual data. By visually displaying data there are many more opportunities for this to occur. Statistical stories should grab a user's attention, invoke thought, be informative and ideally be entertaining.

The primary objective of any statistical story should be to inform its audience and be newsworthy. It must use the statistics available to provide substance and stimulate interest. It should seek to delve through the large pool of data and only surface those details which will be useful and pertinent to the needs of the user. Once this data has been uncovered the next step must be to ensure that the presentation of the story is in a format that is understandable and easy to use. All statistical stories have a target audience and it is critical that their needs are considered.

Data visualisation invokes many new elements compared to common written statistical story techniques. With the media and graphical options available, data visualisation has a great advantage in being able to attract the attention of users and encourage statistical literacy.

# APPLICATIONS OF DATA VISUALISATION

## DASHBOARDS

A dashboard is "a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance." (Stephen Few, 2004)

Figure 1 illustrates various sales measures for a wine business in a dashboard format. Key information on quantitative and qualitative factors are comparable through the use of column graphs with colour and symbols. Instead of users having to trawl through this information in a standard report format, with a quick look at the dashboard they are presented with an overall impression of performance.

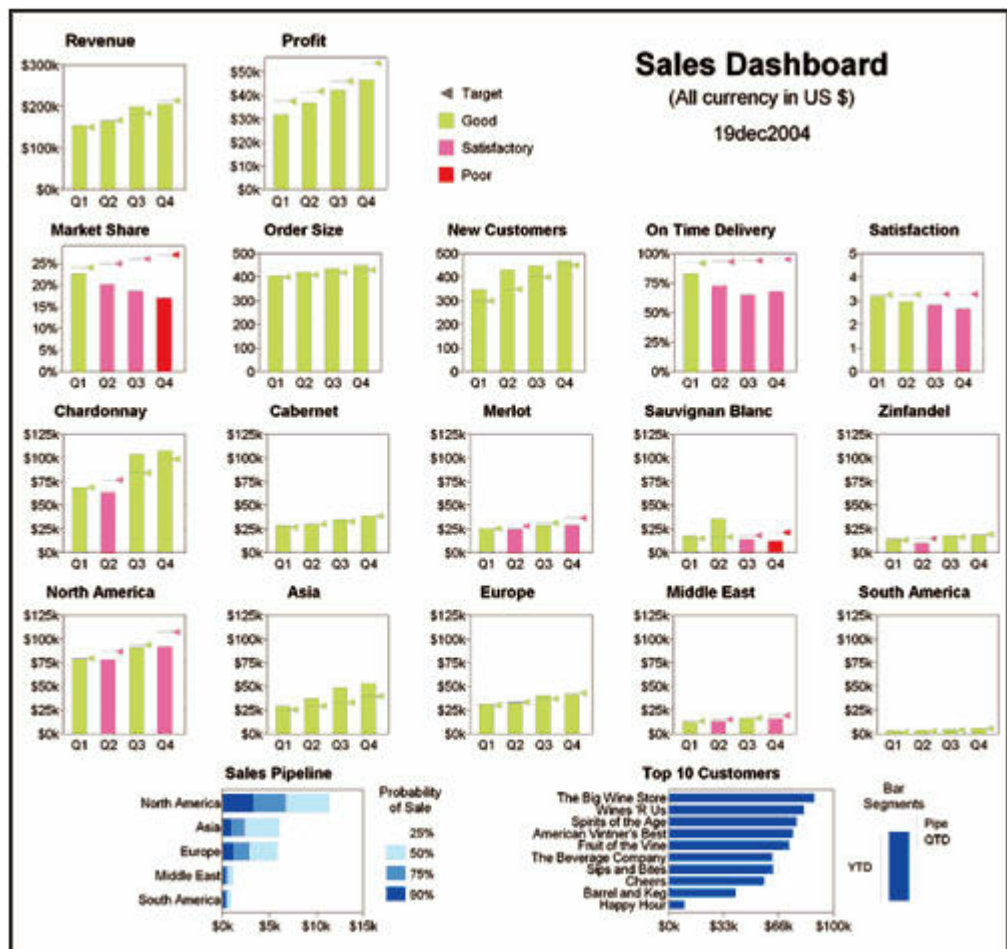


Figure 1 - Source: Downloaded from <http://www.math.yorku.ca/SCS/Gallery/allison/scen3b.htm>

There are many features that contribute to this dashboard being able to present a statistical story. Firstly each graph in the first four rows represents the last four quarters, or the annual performance. They are all striving towards telling the one story. This attribute of consistency across the graphs is very important. Secondly, the use of column graphs help to make the data and the changes stand out. Other statistical graphing techniques could have been used, (e.g. line or dot graphs) however for impact and instant recognition, column graphs provide more value. The use of colour is effective in helping to highlight levels of performance, however in this example it would be more beneficial to have greater contrast between the green, pink and red in a similar way to how the blue is contrasted in the Sales Pipeline bar graph. The target levels may be

## DASHBOARDS *continued*

useful for the management of this company, however may not be relevant for ABS statistics.

There could be a misconception about dashboards that a few graphs on a page equals a dashboard. This is not the case as there are many issues to be considered. For example Figure 1 illustrates how graphs have been meaningfully arranged on the page. The revenue and profit figures, usually a high priority for a company's management, are placed on their own row at the top of the dashboard. If the user requires a further breakdown of information by different categories (eg. type of wine, continent) they can look in the subsequent rows below. The two graphs at the base are also useful. For example, the stacked bar graph, which is not often utilised in ABS statistical publications, is used effectively here through analysing probabilities of sales. Another feature, available through a html overlay, is the ability to link to the supporting data by clicking on the various graphs. This is a great advantage for users who may seek more detailed information than that presented in the dashboard.

Dashboards rely upon simplicity. Johnson (2006, p.25) believes that dashboards are "a tool that simplifies multiple sources of information and allows us to focus on what really matters". Modern business practices rely upon quick and decisive actions and the information required to achieve these goals need to be presented in a format that allows this to occur. With information overload a very real concern Johnson (2006, p.25) believes that dashboards "sort through the chaos of overconnectedness and replace it with "meaningful" connectedness."

Simplicity in dashboards also refers to its design. How can information be simple to analyse if it is not presented in a format that is recognisable and understood by its users? There are many examples used and marketed by organisations with irrelevant dials and graphical techniques that serve little purpose in achieving dashboard objectives. Stephen Few (2005) believes this is a real concern in this field, as indicated by his statement that "most dashboards I've seen, especially vendor examples, suggest little concern for communication, but a great deal of concern for entertainment." Figure 2 could be described as an example of this concern.



Figure 2 - Source: Downloaded from <http://www.8e6.com/products/pdfs/8e6-Threat-Analysis-Reporter-Data-Sheet.pdf>

DASHBOARDS *continued*

Dashboard creation is a difficult task, as condensing a large amount of information into a readable and informative format can result in a variety of issues. The best dashboards communicate information easily. An example of this is Figure 3. There is nothing complicated about its design. Column graphs are well recognised as a graphical format and the colour scheme does not reflect all the colours of the rainbow. The most important information or annual overview is displayed in the large graph at the top with this being broken down even further for each department in the lower graphs. With the top graph being this size, the use of numerical text and the hover-over function, to indicate the exact figure, are useful additions. This may not work with smaller graphs as it may appear too cluttered or the columns may be so small that many users will have difficulty in allowing the mouse to rest over them.

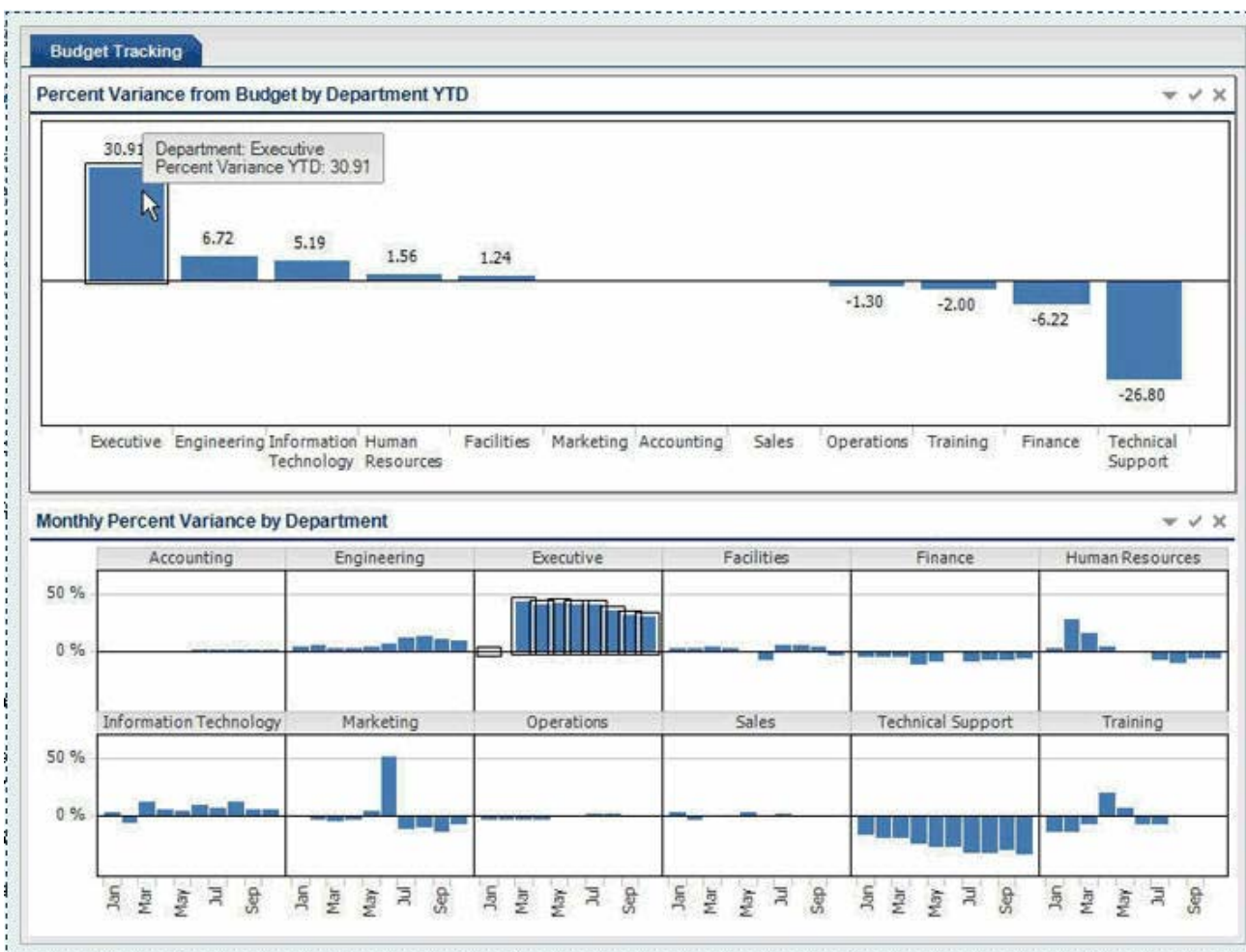


Figure 3 - Source: Downloaded from <http://www.b-eye-network.com/view/3224#thumb>

Another common issue with dashboards is the ability for users to place too much emphasis or reliance upon them. Few's (2004) definition used the phrase "monitored at a glance" and this is critical. Dashboards should only be used to give a quick understanding of the data and users should not become "over-reliant" (Bednarz 2006, p. 1) upon them. They must be supported by the report or underlying data if they are going to have any meaning. Therefore a link to the original data should always be supplied with a dashboard. Consequentially, users of dashboards should acknowledge this issue and

DASHBOARDS *continued*

only use them for a brief overview or as a guide for searching through background information.

SPARKLINES

Sparklines are "data-intense, design-simple, word-sized graphics" as defined by Edward Tufte (2006). As Gibbs (2006, p. 36) states they are intended to be "instantly understandable without adding unnecessary detail". Essentially they are small graphs that provide an instant story-line or trend that provides context to the surrounding data. Examples of sparklines can be seen in Figure 4.



Figure 4 - Source: Downloaded from [http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg\\_id=0001OR&topic\\_id=1&opic=](http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001OR&topic_id=1&opic=)

Sparklines are about communicating "trends rather than detailed data." (Gibbs, 2006, p. 36) If there is a requirement for a detailed graphic illustrating all aspects of the data being analysed, sparklines are not the solution. However this is not their purpose. They exist to assist an audience to gain a greater understanding without disrupting thought processes, whilst adding value to the data. Sparklines can be included within sentences, in tables and even around other graphics to provide background information and trends. Often large graphics distract the audience away from the messages being presented. Sparklines instead compliment and assist in the story-telling mechanism.

The need for maximising data density is critical. It refers to taking into account the size of the graphic in relation to the amount of data displayed (Tufte, 2001, p. 105). He believes that through shrinking a graphic to the minimum size whereby its meaning is not lost, greater value is provided. Tufte qualifies this issue through his analysis on the remarkable nature of our eyes to make numerous distinctions within a small area. This should be taken advantage of, and sparklines do this. Why should our eyes have to scan and evaluate more space than they have too? Simplicity is the key.

Sparklines are not restricted to time-series line graphs. Numerous examples of common types of graphs, expressed as sparklines, are discussed below.



Column graph - Standard graphical format that is common and understood by most users.

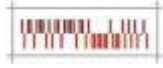
SPARKLINES *continued*



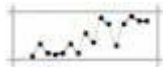
Column graph with negative values - Very useful graph for highlighting negative periods.



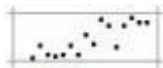
Column graph illustrating periods - Can be used to highlight quarters or cycles.



Win and loss graph - Often used in sporting analysis to highlight a teams performance. Could be used for any data where there are two distinct options.



Line graph with points - Adds further context to a line graph by highlighting periodic results.



Point graph - Useful for determining a line of best fit over the period.



Line graph with open, close and high values - Adds value to a line graph. See the discussion below.



Line graph with normal band - The normal band was suggested by Tufte (2006) as a way of highlighting extreme values or outliers.



Combination line and column graph - Can be used to provide time period context to a line graph.



Pie graph - Well-known format that should be used with caution. May be beneficial to show the percentage difference between two values but any more multiples and it can lose its meaning.

SPARKLINES *continued*



Bar graph - Similar to the Pie graph in that it is well known and cannot be used with too many values.

Figure 5 - Source: Downloaded from <http://www.bonavistasystems.com/index.html>

The examples above illustrate that there are numerous ways of adding even more value to a standard sparkline. Besides changing the graphical depiction of the sparkline additional text can be placed in and around it, within reason, to achieve this goal. This text is also used to compensate for the lack of an axis on a sparkline.

The additional text that supports a sparkline generally takes the form of showing the highest, lowest and current figures or the opening and closing figures. The option chosen will depend on the data being displayed. If a time period context is required, using opening and closing figures will be more beneficial. However for a vertical scale or size context, using highest, lowest and current figures will be the best option. Figures 6 and 7 illustrate stockmarket figures and use the highest, lowest and current option as it will provide more value to their users who are after trend value information.



Figure 6 - Source: Downloaded from <http://www.stockmorph.com/sparklines-remote-module-or-gadget-for-google-home-page/>



Figure 7 - Source: Downloaded from <http://www.bissantz.de/sparklines/index.asp>

The concern with providing additional contextual information to a sparkline is the issue of space. Firstly the size of the text must be readable for users. Secondly there must be appropriate space between the text to ensure a lack of clutter. Thirdly it must be ensured that it is placed and formatted in a way that can easily be understood. Consider Figure 8. The lowest point and the current point are so close together that it appears disorganised and difficult to understand in "a glance." The eye must take those extra few moments to distinguish between the two data points and evaluate what it is trying to show. This is not an example of an effective sparkline.



Figure 8 - Source: Downloaded from <http://www.stockmorph.com/sparklines-remote-module-or-gadget-for-google-home-page/>



From an accessibility perspective the size of the sparklines is an issue. However enlarging their size removes their usefulness. Therefore to overcome this problem, enlarged graphics should be provided or available from another location. Colours are also very important. There is not a set colour pattern for how the data points should be represented. It is also relative to the number of points being referred to on the sparkline. A very common colour scheme is the use of red for the lowest, green for the highest and blue for the current points. Various tests should be undertaken to ensure that not only are the colours acceptable from an accessibility perspective but that the reference points are of a size that these colours are distinguishable.

The simplicity of the sparkline design has led to many unique adaptations and uses. One of the most practical and easy to create designs is the ability to make in-cell bar graphs using a spreadsheet package. With the use of a column and a simple formula, in a glance users can gain an understanding without the need to scan up and down the numerical data. Consider the spreadsheet in Figure 9. It illustrates batting statistics for a number of baseball players. Although player data is available in a column, through a simple formula a bar graph can be created to highlight a particular statistic. In this example the bar graph shows the number of time the batter has received four balls and walked to first base, BB (Base on Balls). This technique is particularly useful when analysing a lengthy list of data.

	A	B	C	D	E	F	G
1	<b>Name</b>	<b>AB</b>	<b>AVG</b>	<b>HR</b>	<b>BB</b>	<b>K</b>	
2	Nomar Garciaparra	305	0.33	12	28	17	=REPT("T",E2)
3	Albert Pujols	291	0.33	32	57	26	
4	Scott Hatteberg	268	0.32	9	43	21	
5	Lance Berkman	335	0.32	26	50	64	
6	Travis Hafner	325	0.31	29	72	74	

Figure 9 - Source: Downloaded from <http://www.juiceanalytics.com/weblog/?p=236>

Sparklines may be criticised for their inability to appropriately display the time variable. As one of their primary uses is to display time-series data this is a particular concern. Figure 10 not only is able to alleviate this problem through an effective design, but combines a few other visualisation tools to present a very effective story. As mentioned above, column graph sparklines are particularly useful when analysing data that only has two options, in this case a victory or a loss. The time period along the base helps to illustrate from when Isiah Thomas became president of the New York Knicks until the end of the 2005-6 season. It is based on the number of basketball games within the seasons, not days and years. There is a clear separation and distinction between these periods and there is no system of numbers like in Figure 6, which relies upon either knowledge or common-sense from the user.



### SPARKLINES *continued*

Figure 10 - Source: Downloaded from

[http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg\\_id=0001lh&topic\\_id=1](http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001lh&topic_id=1)

There are a few other useful elements in this design. The use of colour is particularly commendable. Within the sparkline a very light blue colour indicates a victory whereas a dark navy blue is used to signify a loss. This contrast is beneficial for not only those who are colour-blind but also reflects the cultural associations that a dark colour often signifies an unfortunate event. This dark colour also stands out more, which reflects the aim of the graphic, to display the poor performance or losses by the New York Knicks over this period. Colour is also used in the coaches row. Notice how Herb Williams does not have a light blue background behind him. The reason for this is to signify that he was only an interim coach and the club was in a transition period. Another useful feature of this graphic is the text that helps to explain the meaning of the story being displayed. A bold text headline is used, to attract those users who are familiar with Isiah Thomas, and then a small description is used to help explain the story. As important and useful that the graphic is, users may not fully recognise its purpose without this brief description.

### SEARCH CLOUDS

Search or tag clouds are a growing trend amongst many web sites. An example can be seen in figure 11. They are a visual depiction of the most popular terms that people have tagged or visited and link to associated pages. The larger the term, the larger the popularity.



Figure 11 - Source: Downloaded from <http://www.flickr.com/explore/>

There are two types of search clouds. They look the same however the reasons for the size of the tags distinguishes them. The size of the words can either be based upon how many times a tag is allocated or applied by those operating the website or by the popularity associated with external users searching or selecting a tag.

Search clouds can do more than just highlight key terms or illustrate the most popular items. Nagy (2006) proposes that one of their most effective uses is in providing contextual relevance after a search has been conducted. The mock-up he created to illustrate this point is seen in Figure 12.

SEARCH CLOUDS *continued*

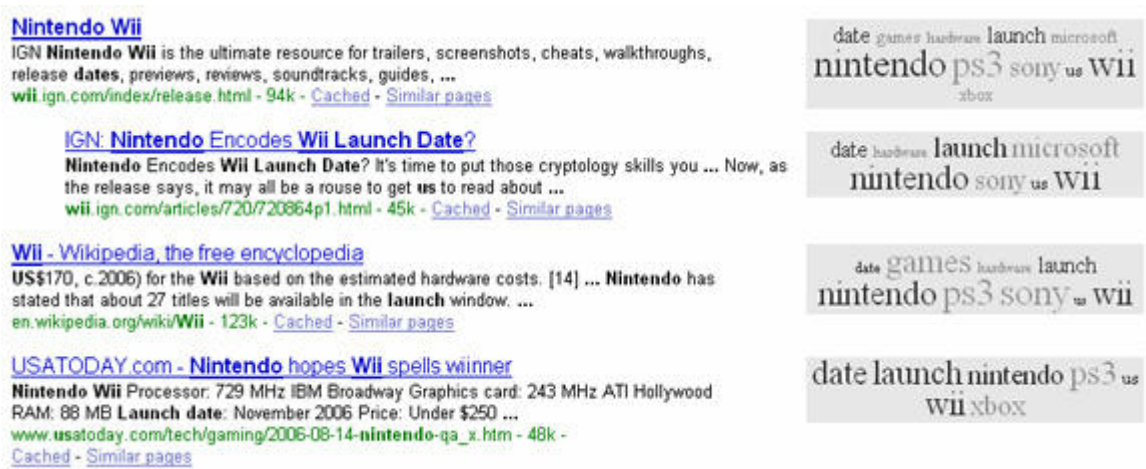


Figure 12 - Source: Downloaded from [http://lab.arc90.com/2006/10/search\\_clouds.php](http://lab.arc90.com/2006/10/search_clouds.php)

This example used a search for "Nintendo Wii Launch Date". This would provide him with a large amount of results to choose from. However through the use of search clouds, which show the frequency of related words on each page, with a quick glance he can narrow his search down even further. Although it is the fourth site on the list, this result highlights all the key terms being searched for and may be a better solution than the first result. The third item on the list highlights 'games' as a major term and although it may not have been the desired result for the user, it may attract their attention and invite them to visit this site. Search clouds used this way simply provide more context or background information to assist with a search.

As well as providing assistance in finding appropriate sites, search clouds also have the ability to eliminate irrelevant search results. For example, Hoekstra (2006) highlights how search clouds would help in searches using terms that have multiple meanings. For example she discusses how a search for RSS would generally produce results on Really Simple Syndication, the news or blog update service, however in one particular search it brought back a result on a problem with Macbook computers known as Random Shutdown Syndrome. Through further investigation she found out that there are at least 40 different abbreviations that are known as RSS. With a simple search cloud next to these search results relating to Macbooks can instantly be eliminated.

There are many other uses of search clouds. Clusty Cloud creator (2006) highlights many useful applications of the search cloud:

Make a Cloud about a person so that visitors to your site can find out more information about them



SEARCH CLOUDS *continued*

Figure 13 - Source: Downloaded from <http://cloud.clusty.com>

Make a Vanity Cloud about your website to see how it is in the news or what people think about it

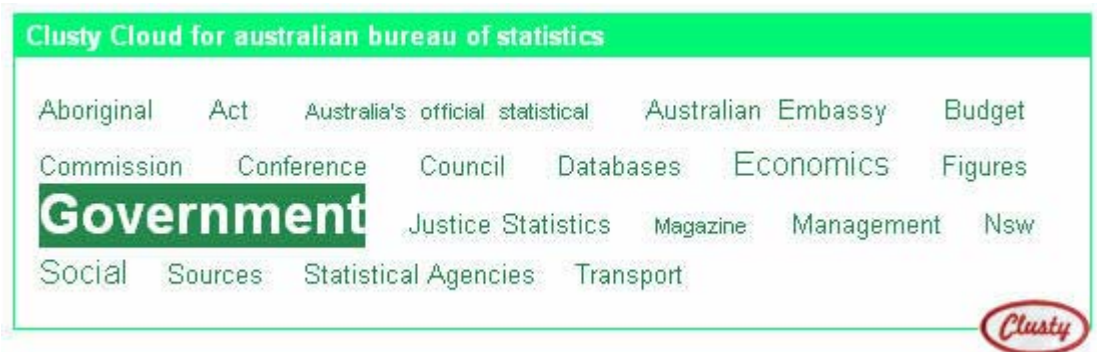


Figure 14 - Source: Downloaded from <http://cloud.clusty.com>

Include a cloud after a page or document to show all the highlights or most popular issues

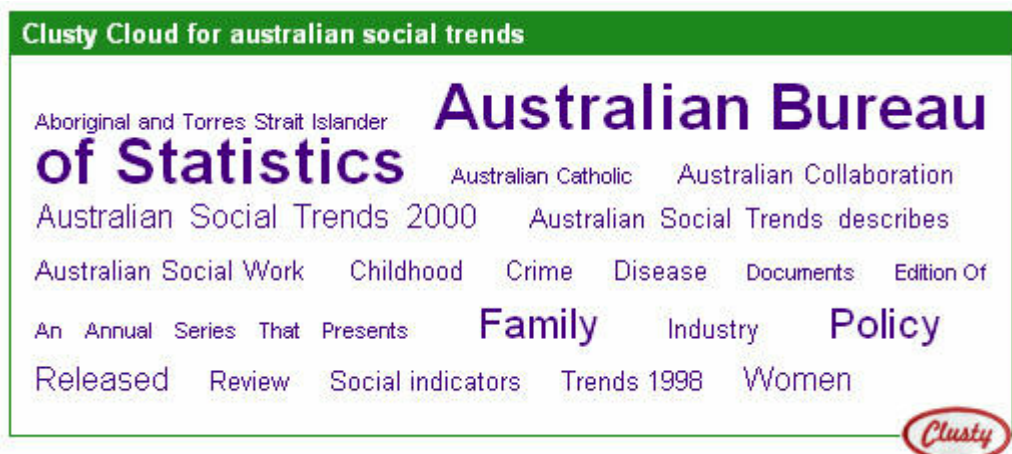


Figure 15 - Source: Downloaded from <http://cloud.clusty.com>

### SEARCH CLOUDS *continued*

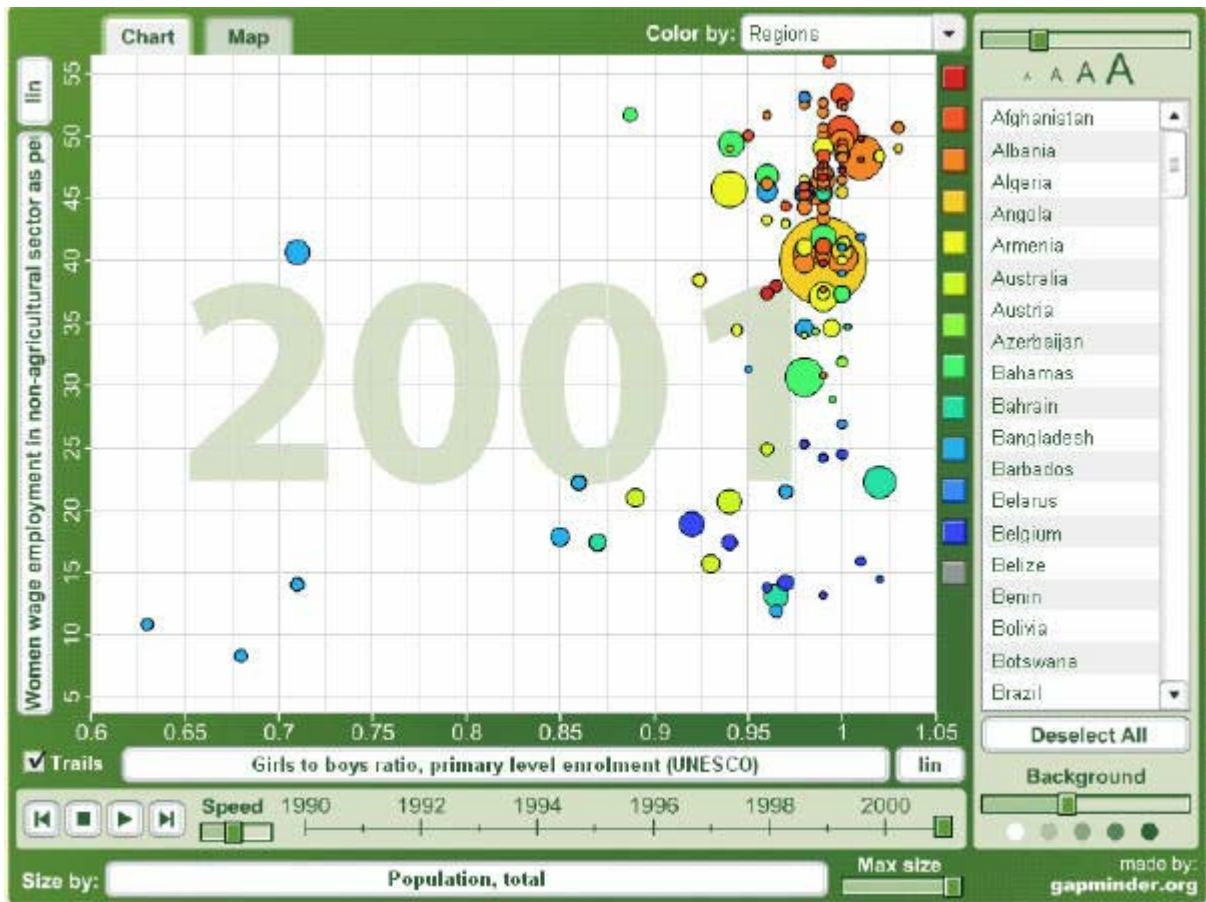
Search clouds do have some concerns. Although they are growing in popularity many users will still be unfamiliar with their look and purpose. To see a box with some large and small words in what appears to be a random order may alienate many users as it can appear untidy and unnecessary on a web page. If the users do select a term, their vision of where this takes them may be different to the page they arrive at. They may be taken to a list of search results or to a related page, e.g. (Consumer Price Index product page) Appropriate metadata should explain the linked page. Another consideration should be the font size. The smallest font should not be so small that sight-impaired users will not be able to see them. Relevance is lost if these smaller terms cannot be seen.

### GAPMINDER

Gapminder is a non-profit venture for development and provision of free software that visualise human development. It began in 1998 when Ola Rosling, Anna Rosling Ronnlund and Hans Rosling had an idea to enhance the understanding of world health. Since then it has grown into an organisation that develops one of the most talked-about data visualisation techniques available today. With their vision to make sense of the world by having fun with statistics, their work is widely recognised and applauded.

There are a variety of Gapminder examples, although most follow a similar layout. This discussion will focus around the example seen in Figure 16 which relates to the millennium development goal indicators. It has a variety of features that allow for the customisation of the graph. There are the two axes that can be changed to illustrate certain variables and the scale, certain countries or regions can be highlighted, the size and shading of the background can be modified, the graph or map option exists and the animated aspects can be adjusted for speed and year details.



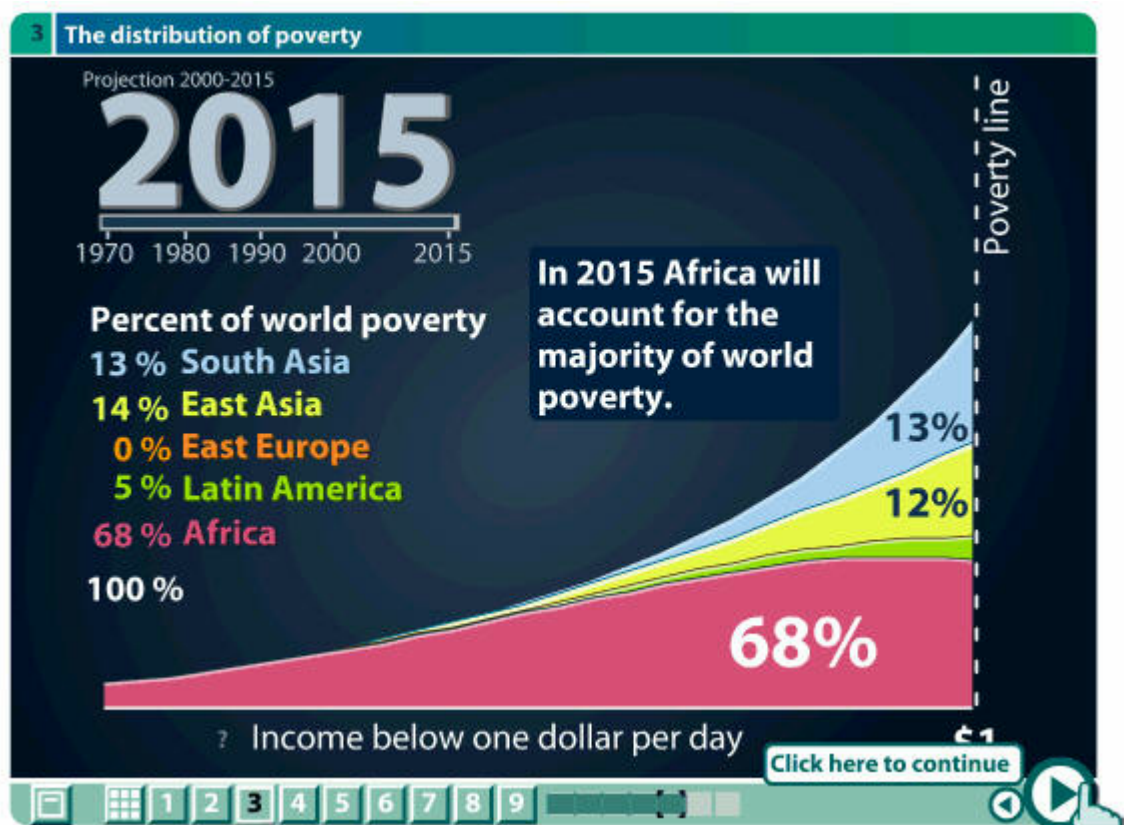


GAPMINDER *continued*

Figure 16 - Source: Downloaded from <http://mdgs.un.org/unsd/mdg/Default.aspx>

There are numerous advantages and benefits associated with the Gapminder product. Its colourful and attractive presentation can capture a user's imagination and gain interest where a standard graph may not. Its interactivity and animation with numerous variables and features provides a visualisation that should not just be seen with a glance. Instead it should be explored and played with to uncover stories and find interesting facts relevant to the user. The usefulness of this product lies in its ability to show relationships. These relationships lead to stories and these stories lead to knowledge and a greater understanding of statistics. For Gapminder, and any users, it is this ability to create awareness that links back to the original objectives of this product.

The second type of Gapminder presentation is displayed in Figure 17. This technique provides the user with the complete story. It works in a similar fashion to a visual slide presentation, however it always includes animation and certain slides are interactive. It 'walks the user through' the story, rather than giving them free reign to modify and create their own story like the example seen in Figure 16. In a way it provides a prepackaged product. This may benefit many users who may find it difficult to find their own stories, however it takes away the freedom and may introduce a level of bias. The best examples are when this type of story is presented but the last slide is an application like Figure 16, which will allow the user to also locate and discover their own story.



GAPMINDER *continued*

Figure 17 - Source: Downloaded from [www.gapminder.org](http://www.gapminder.org)

The technology or animation behind Gapminder is through the use of Flash. Gapminder developed the free software Trendalyzer which imports data and shows moving graphics on the screen as exported Flash files. However the majority of the projects are stand-alone software (.exe files) and can be downloaded and used on a PC without the need for any players or programs to be run. Gapminder are currently developing a product for external use.

Blum's (2006) article about Gapminder stated that having better tools to analyse data "encourages closer examination". Gapminder is certainly one of these tools. It helps to illustrate statistical stories in such a way that it may uncover unique links or relationships that should be explored further. It has enormous possibilities as a data visualisation technique, however it must be ensured that it is used appropriately. It must be user-friendly and accessible to the common user if it is to succeed as more than just a presentation tool. Blum states that the "intended audience is not lay people, but researchers, civil servants, journalists and activists who will then present their graphical analyses to a broader public". However if it truly is to succeed as a tool to enhance understanding and educate the public, it must work at an appropriate level for a general audience.

Although Gapminder is a very unique product, it is possible to reproduce its general attributes using alternative software. There are a few examples that reflect many of its attributes. One such example can be seen in Figure 18. This Business Cycle Tracer visualises the key national statistical trends using animation to illustrate the time variable. It is an unusual concept in that each quadrant of the graph indicates a particular part of

GAPMINDER *continued*

the trends performance as seen on a line graph. So it is actually visualising a visualisation in that the data is actually the statistics line graph. It has many similar features, such as the ability to toggle the desired boxes and the time bar across the base of the graph.

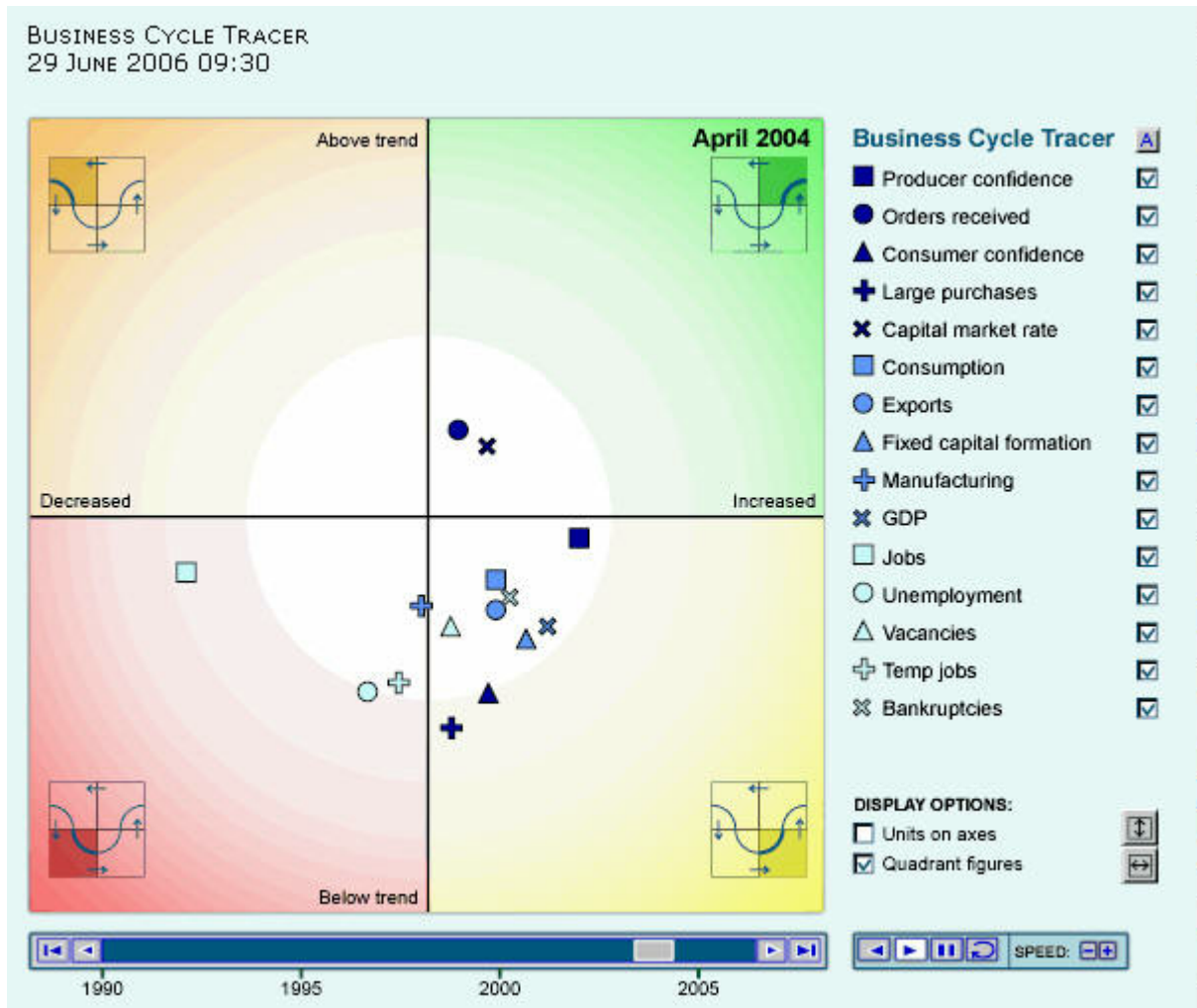


Figure 18 - Source: Downloaded from <http://www.cbs.nl/en-GB/menu/themas/macro-economie-financie-instellingen/conjunctuurgegevens/publicaties/conjunctuurbericht/klok/ck-homepage.htm>

MINDMAP SEARCHING

Searching the internet or a website has become one of the most natural web navigational techniques. It is one of the most important design elements on any website. Searching has evolved in recent years with the introduction of a number of different search tools and techniques. Advertising, tagging and advanced searching tools are just some examples of ways in which search engines are doing more than just finding a list of relevant web sites.

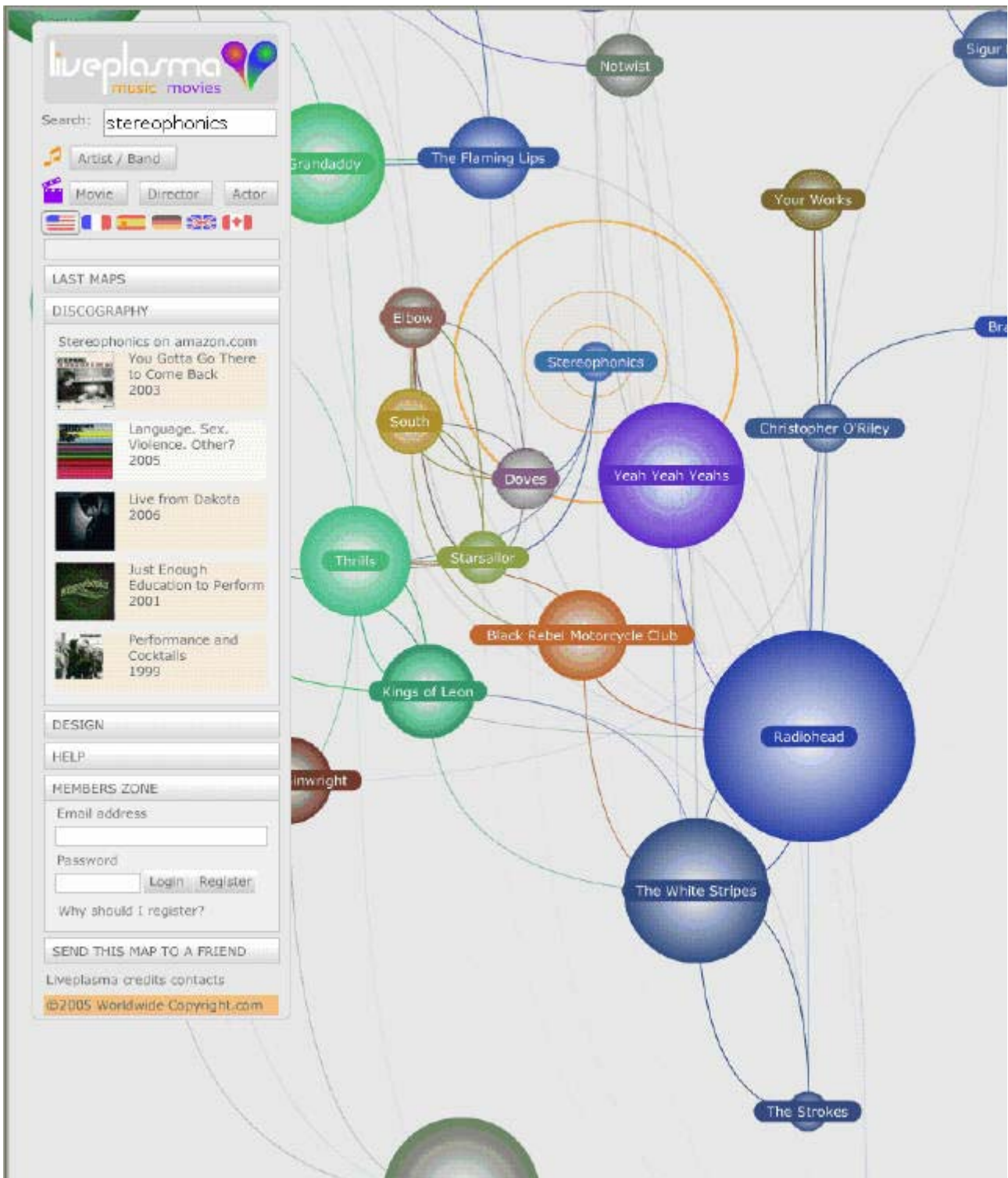
One current technique that is gaining popularity is a form of searching that creates mindmaps or links between various terms rather than just producing a list of results. The primary benefit of this type of searching is that users may find alternatives or options that they had not considered searching or looking for. In a way it is advertising. That is, it is trying to take the user somewhere or make them do something they had not planned on



### MINDMAP SEARCHING *continued*

doing. This makes for a very useful tool and can provide benefits to both the user and the website producer. Nielsen (2003) stated that the reason why search ads work so well is because "search engines are the one type of website that people visit with the explicit goal of finding someplace to go." This is the basis behind mindmap searching.

Figure 19 illustrates a unique mindmap searching tool for music or movies. This example illustrates various links between a number of artists based upon a criteria. The size of the circle or bubble represents the popularity of the artist. As this website does not appear to explain its methodology this may be based upon the amount of searches, album sales, recognisability or a number of factors. The colour of a bubble represents the style of music that each band has been classified into and similar colours will reflect similar types of artists. Lines are used to illustrate which bands are linked to each other.



MINDMAP SEARCHING  
*continued*

Figure 19 - Source: Downloaded from <http://www.liveplasma.com/>

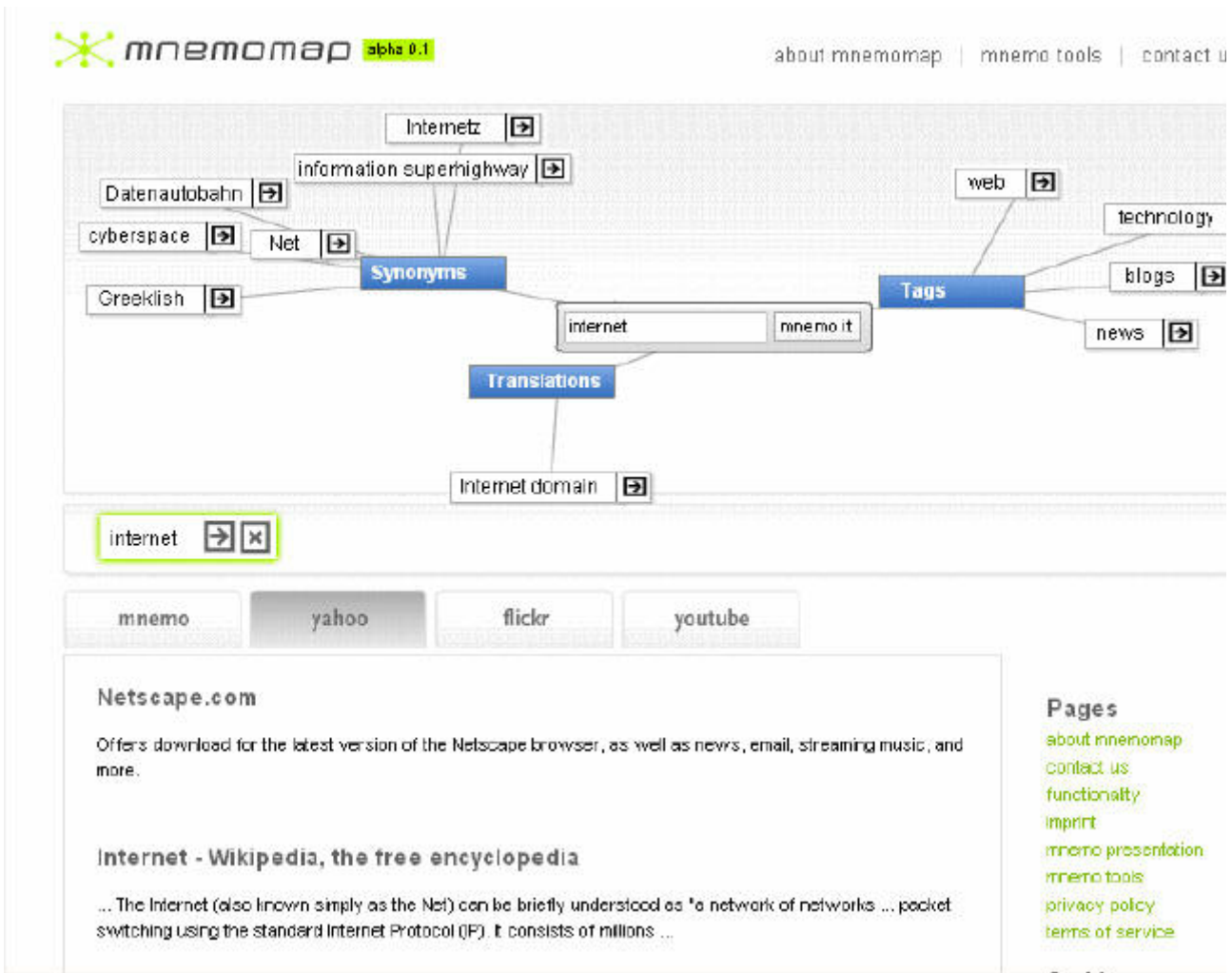
There is more to this example than simply showing links between various artists. By hovering over an artists bubble, a star will appear which will then allow the user to place them in a favourites list. This favourites list will allow them to come back and find these artists in the future or send them to their friends. Another great feature available to members is the ability to receive news about these favourite artists that have been bookmarked sent straight to an email address, a similar function to Really Simple

### MINDMAP SEARCHING *continued*

Syndication (RSS). Another interesting feature is the discography section on the left hand navigation which advertises products (CDs) from the artist that was searched for. This provides added context and possible future assistance in learning more about the artist.

These interesting facets of this search engine can all be used to benefit the search for statistics, enhance statistical literacy and even tell a few stories along the way. Instead of artists various statistical products or concepts could be used. The popularity based upon search results or downloads could be used to signify the size of the bubble, colours could represent a statistical category or topic and the lines would again link these bubbles. A members section could be created to allow favourites to be selected and possibly be incorporated into the current RSS system. Advertising to various statistical products could also be used in the left hand navigation to attract users to products that are associated with their search which they had not thought about.

Another example can be seen in Figure 20. It is more of a traditional search engine and has a few added functions. The first is the additional branches that allow for such items as synonyms, translations, definitions and tags. This could be especially beneficial for a metadata vision that would allow users to find out a definition of a statistical term or link to how others have used this term in news articles, blogs etc. Another useful function is the little arrow next to a branched search term. For example clicking on the arrow next to "Internetz" will redirect the search so that this term is the focal point and branches stem off from it. Furthermore the bar beneath the search area acts like a memory of recent searches or like a breadcrumb trail.



MINDMAP SEARCHING  
*continued*

Figure 20 - Source: Downloaded from <http://mnemo.org/>

Mindmap searching will not necessarily find a desired page, normal searching is probably best for this. Instead it should be used as a tool to interest, guide and promote users around a website. With the enormous array of existing statistical products and terminology, this type of visual tool may solve many problems. The concept of a mindmap is not new, that is one of its best qualities. Users are familiar with the concept and will hopefully quickly adapt and enjoy the benefits of this technique. There are numerous possibilities in its design. Ultimately it could become a key navigational technique around a website, as if designed appropriately any statistical search term entered could link to all its uses on the website from information to definitions to downloads etc.

Mindmaps do not necessarily have to arise from entering a search term. Instead they can be used for their more visual technique of showing linkages and relationships between various entities. Figure 21 is an example of this from a website looking at the influence of Exxonmobil on the issue of climate change. This reflects the common mindmap studying technique of taking ideas, concepts or entities and showing how they connect to each other. Many students use them as a visual aid to assist them in remembering different aspects of the subject they are studying. In this example users are able to import organisations or people into the map and the software will show how they are linked

MINDMAP SEARCHING  
*continued*

together. These entities can then be moved and shaped around so that the user has greater control over the design of the mindmap. These factors help to tell a story by allowing users to visually conceive who has influence, power and the contacts to explain why events unfolded as they did.

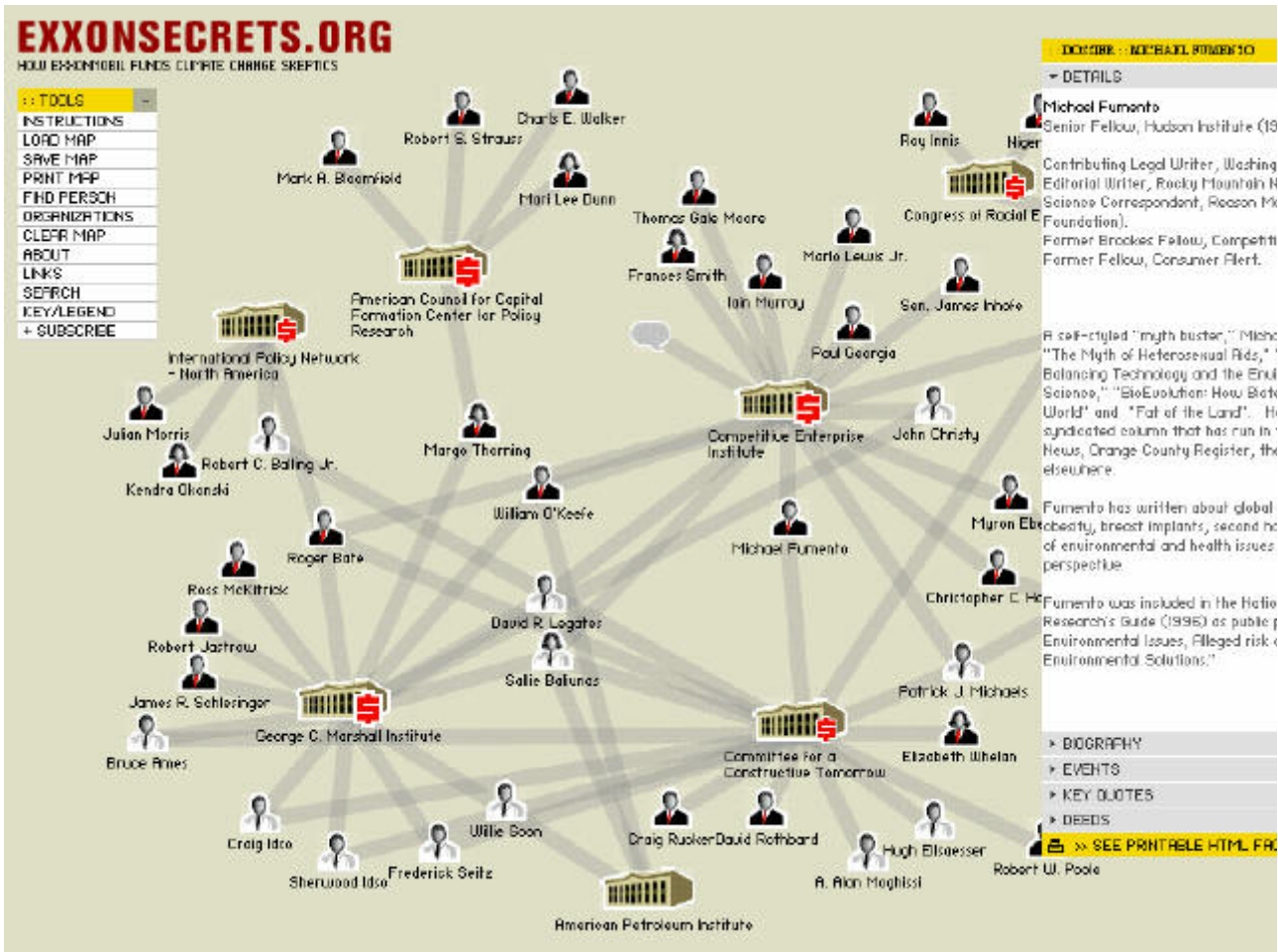


Figure 21 - Source: Downloaded from www.exxonsecrets.org

The added advantage of this type of mindmap over a student's mindmap is that this is interactive and further data and information can be accessed at the click of a button. Users are able to select people or organisations and bring up information boxes like the one seen on the right above. This provides context or background information to further enhance the stories being created. It is essentially a visual playground from which stories can be constructed and manipulated to the advantage of the user.

The disadvantages of these mindmapping tools lie in their accessibility. This sort of design would be very difficult to reproduce in HTML as well as in a format that would be readable by sight-impaired users using screen readers. In terms of usability there may also be concerns of a user's ability to use and understand the movement and links that they may not be used to in a search engine. Therefore it is essential that appropriate help and documentation is available to help users take the advantages available away from this technique.

### TREEMAPS

A treemap "is a method for displaying information about entities with a hierarchical relationship, in a "space-constrained" environment" (Wikipedia, 2006) The idea for their construction arose in the early 1990s when a university professor, Ben Shneiderman, had difficulty managing the small amount of hard disk space he had available on his server and required a way of showing his tree diagrams "in a space-constrained layout." (Shneiderman, 2006)

Treemaps display rows of data as groups of shapes that can be "arranged, sized and colored to graphically reveal underlying data patterns." (Wikipedia, 2006) It not only condenses information into a more compact and manageable form but it also allows for the recognition of relationships or patterns.

An example of one company that has taken a different approach to treemaps and expanded on the traditional layout can be seen below. Figure 22 takes information from the Nasdaq 100 to highlight a great deal of information in a small amount of space. At a wide level the size or importance of each stock is related to the size of the shape that it occupies. Similarly each categorised group (eg. Technology, Healthcare etc.) is grouped according to its relevance and size. Different stocks are colour-coded to indicate a certain type of performance and when selected, information on the selected company is displayed on the left of the treemap. Furthermore the drop-down boxes in the bottom left hand corner allow the treemap to be customised to a greater extent.





TREEMAPS *continued*

Figure 22 - Source: Downloaded from <http://www.labescape.com/>

However the most attractive statistical addition to the lab escape example is the use of area graphs within each shape. In terms of a data visualisation technique it works in a similar way to sparklines in that in a glance the user can gain an insight into the recent performance of a certain indicator. The ability to quickly look at the different stock and see a trend or story of performance is informative and easy on the eye. Treemaps are designed to condense information, however if it is too dense, meanings and derivations are lost.

A criticism of treemaps may be that they look colourful and interesting but they fail to simply explain the information being presented. Asahi, Turo and Shneiderman (1995) discussed their capabilities as a decision-making tool. If they fail to simply explain the information this capability is lost. Figure 23 is an illustration of one of the most attractive and useable examples of a treemap. It shows the most popular songs being downloaded and can be used as a decision-making tool to help users select songs they may wish to purchase. The musical genre tags help to categorise and break down the songs into a readable format. A good technique here is that only songs with a large enough shape or box state the name or a part of the name. This helps to avoid clutter and places more

TREEMAPS *continued*

emphasis on these most popular songs. The problem that this arises is if the user is only interested in country music as this genre is not as popular and their songs are difficult to identify. To correct this problem, the ability to click on the country tag and a similar treemap for country music only is provided. This ability to drill-down through treemaps adds further detail and more information for the user.

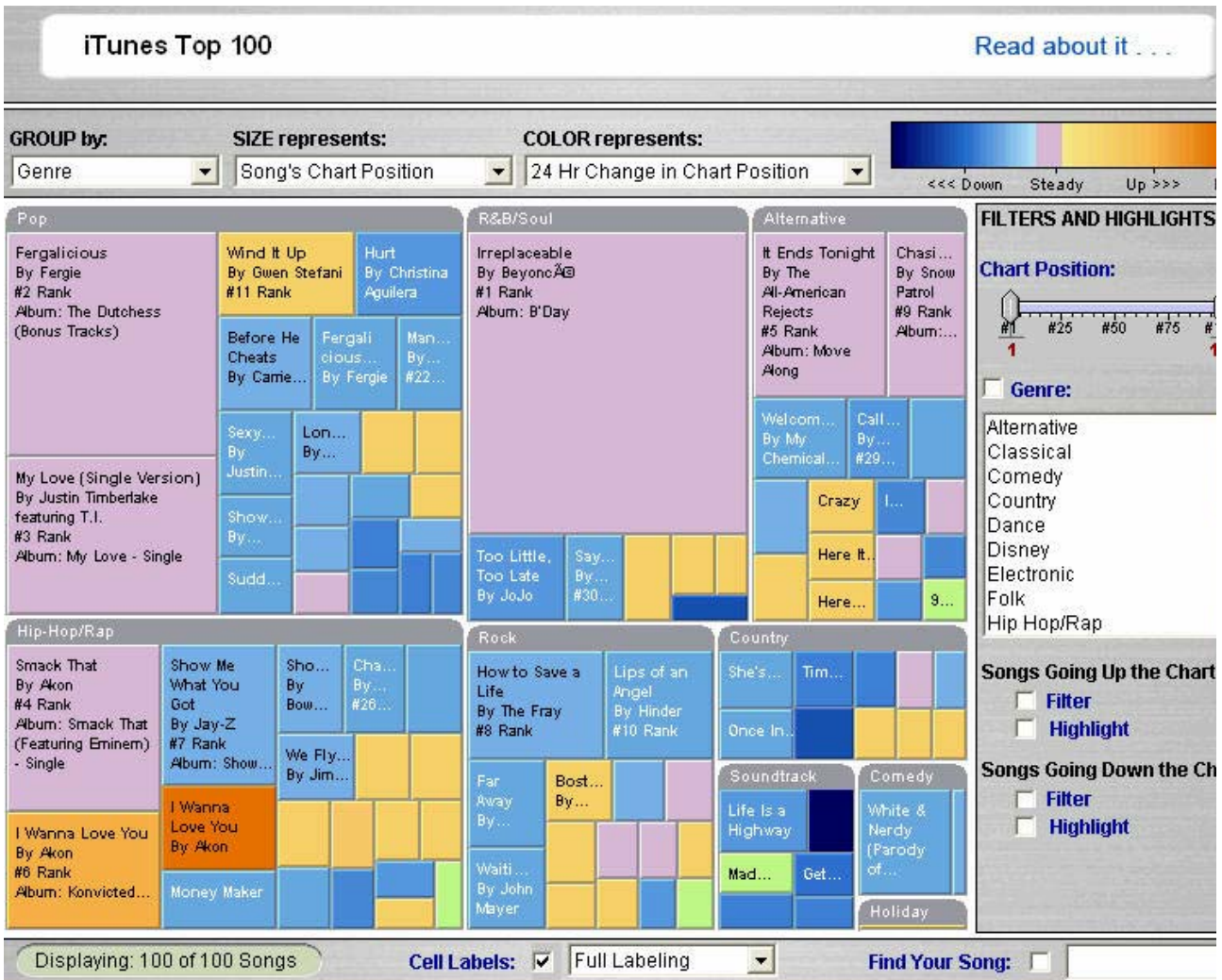


Figure 23 - Source: Downloaded from www.hivegroup.com

The other useful additions to this treemap are all the variables and additional information that can be gathered from the surrounding options. The ability to change the grouping, size and colour as well as the filtering and searching options are all beneficial tools. The colour guide in the top right corner is also very easy to understand and use in identifying various aspects of the treemap.

In a way treemaps are similar to search clouds in that they are trying to quickly highlight the most popular or important aspects of data. One example that reflects many of the properties of a search cloud can be seen in Figure 24. This example highlights the most popular pages or tags. Instead of just using a single box with different words or phrases in bolder text to indicate their level of significance, many different sized boxes are used and the subsequent text size is proportionate to this size. The reason why a treemap may



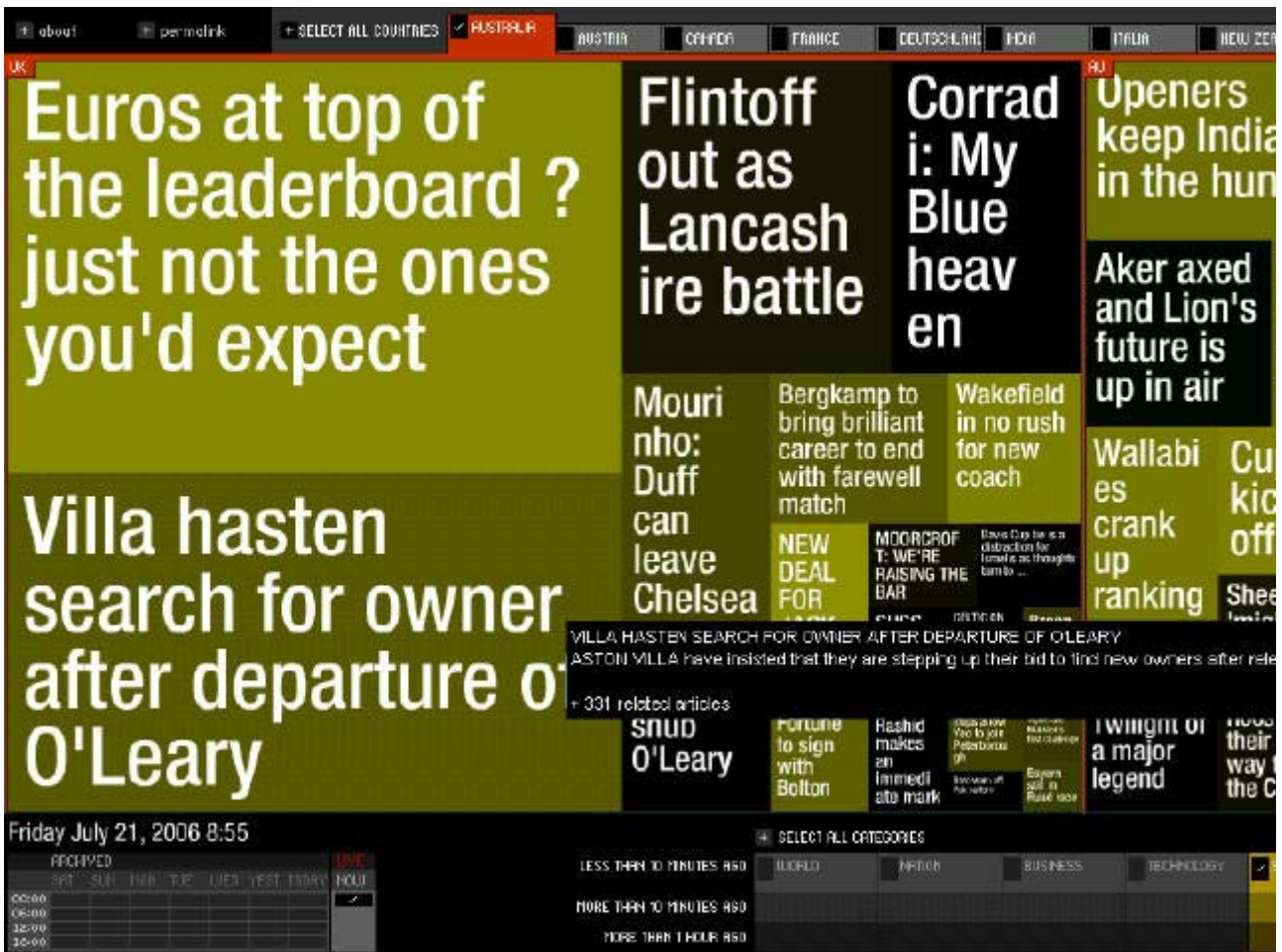
TREEMAPS *continued*

be used over a search cloud in this case is the fact that the pages being linked to are described in more than just a single word or two. Instead the headline or major aspects of the page take up slightly more space to entice readers.



Figure 24 - Source: Downloaded from <http://codecubed.com/map.html>

An example that further builds on these ideas can be seen in Figure 25. The same principles apply however many more variables and options can be taken advantage of. This site only deals with news stories and headlines are presented in the treemap boxes. An added feature is that when the user hovers their mouse over the headline they will be given the first line of the article to provide them with more context to enhance their decision of whether to read the entire article. There is also the option of selecting the related articles expansion which will allow the user to focus in more detail on the issue or news story they are trying to learn about.



TREEMAPS *continued*

Figure 25 - Source: Downloaded from <http://www.marumushi.com/apps/newsmap/newsmap.cfm>

The additional options available with this example are very useful in this news context. The ability to choose articles from specific countries as well as the news genre are very useful and allow for the filtering or subsetting of the data. However the addition of the time variable adds even more value to this treemap. The option of being able to find the most recent stories or from an archived date provides a great database of information.

The concerns over these headline treemaps reflect those of all treemaps in that they may appear too cluttered for many users. This is especially the case when dealing with large amounts of text, such as displaying headlines in this format. However if a balance can be achieved between font size, colour and the number of boxes or shapes their ability to convey a message will succeed. The idea and opportunities for treemaps are certainly there, it is just a case of gaining the most out of them through appropriate design.

STORIES CREATED THROUGH USER INPUT

As Walker and Antanies (2006) discuss, data visualisation tools have the ability to empower those that use them. Role plays are an example of a learning technique and their advantages lie in participant involvement. The same can be said about data visualisation tools. Users will gain a great deal more and discover their own stories if they have some input into the statistical presentation process.

# APPLICATIONS OF DATA VISUALISATION *continued*

STORIES CREATED  
THROUGH USER INPUT  
*continued*

Figure 26 illustrates an example where a user can answer a few simple questions and the subsequent data is visualised and explained. As a user enters how often they undertake the various activities, the bar beneath expands or contracts to illustrate how much energy is used and further information is provided. It is a simple, yet effective technique that can engage a user.

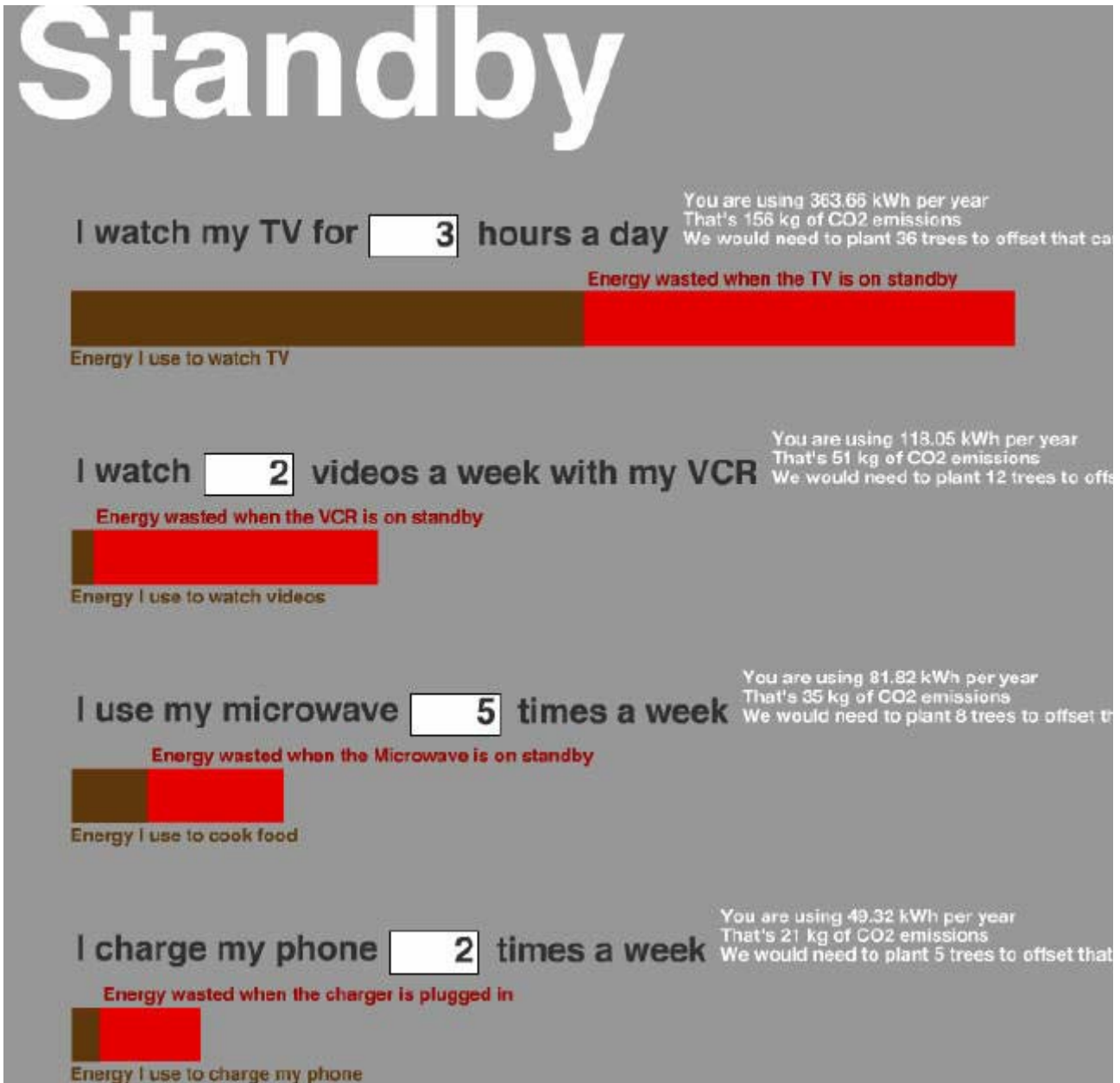


Figure 26 - Source: Downloaded from <http://energy.failedrobot.com/standby.html>

To enhance this technique further a scale on the expanding bar could be adopted. Depending on the data being visualised, the need for text may be removed if appropriate metadata surrounding the bar is supplied. It may also be more beneficial to use some other form of visual effect other than a bar. A line or pie graph may be more effective or even the use of sparklines.

STORIES CREATED  
THROUGH USER INPUT  
*continued*

A relevant statistical example, which requires user input, has recently been developed by the Federal Statistical Office Germany. It allows users to customise their own Consumer Price Index (CPI). The Index calculator, illustrated in Figure 27, provides users with the option to adjust the average consumption habits that make up the CPI, in accordance with their own spending activities. Sliders are used to adjust the percentages and this is reflected in an individualised line graph that overlays the overall CPI graph.



Figure 27 - Source: Downloaded from <http://www.destatis.de/basis/e/preis/start.htm>

Statistics Norway has also introduced a CPI calculator, see Figure 28. Although it does not have a visual aspect to its design, a graph could easily be developed to display the data requested by a user.



## How much is it worth...?

Would you like to know how much 1000 Norwegian kroner in 1930 equals in 2002 kroner, or what 5000 kroner in 2002 would have been worth in 1930? You can find this out by using the computing areas below. The monthly consumer price indices from Statistics Norway.

- Fill in the kroner amount that is the starting point:**
- Fill in the year and, if desired, month of the amount:**  
 If the month is not stated, a yearly average will be used.  
 Average
- Fill in the year and, if desired, month that you want the amount computed into:**  
 September

STORIES CREATED  
THROUGH USER INPUT  
*continued*

Figure 28 - Source: Downloaded from [http://www.ssb.no/kpi\\_en/kpicalc.html](http://www.ssb.no/kpi_en/kpicalc.html)

The Office of National Statistics (ONS) website has another variation on visualising the CPI. The advantage to this design, as illustrated in Figure 29, is that users have the ability to input monetary figures. This allows users to see a more direct relationship between the changes in CPI and the impact upon their spending habits. A line graph, a table and a bubble graph then provide an interactive element to displaying the data.

**national STATISTICS**      **your spending**      **Personal Inflation Calculator**      [about this calculator...](#)

- Estimated MONTHLY Expenditure** on Regularly Purchased Items:
- How much of this is on:**
  - Food:
  - Meals Out:
  - Alcohol:
  - Tobacco:
  - Phone Charges:
  - Clothing and Footwear:
  - Rail and Bus Fares etc.:
  - Education/Child Care:
  - Chemists Goods:
  - Fuel for Transport:
  - Heating and Lighting:
- Calculated Other Monthly Expenditure** (edit this value if necessary):
- Calculated Monthly Total (£):** **1,133**
- Accommodation Expenses**
  - a. if you own your property:**
    - Value of Outstanding Mortgage:  (Est. Annual Interest) **1270**
    - Value of Your Property:  (Est. Annual Depreciation) **1022**
    - Where You Live:  East Midlands
  - b. if you pay rent:** MONTHLY Rent:
  - c. Utilities and Insurance:** ANNUAL Council Tax:  ANNUAL Water Charges and House Insurance:
- ANNUAL Spending on**
  - Housing Repairs, Maintenance and DIY:
  - Vehicle Repair/Maintenance:
  - Vehicle Tax/Insurance:
  - UK and Foreign Holidays and Other Airfares:  (Est. Car Expenditure) **1759**
- Spending in LAST THREE YEARS on** Furnishings and Electrical Goods:
- Calculated Annual Total (£):** **24,002**

**10 Month Inflation Rates (%) - Jun 2015-May 2007**

MONTH	INFLATION
Jun 2007	2.2
Jul 2007	2.4
Aug 2007	2.1
Sep 2007	2.1
Oct 2007	2.4
Nov 2007	2.4
Dec 2007	2.5
Jan 2008	2.4
Feb 2008	2.1
Mar 2008	2.1
Apr 2008	2.0
May 2008	2.0
Jun 2008	2.0
Jul 2008	2.1
Aug 2008	2.1
Sep 2008	2.1
Oct 2008	2.1
Nov 2008	2.1
Dec 2008	2.1
Jan 2009	2.1
Feb 2009	2.1
Mar 2009	2.1
Apr 2009	2.1
May 2009	2.1
Jun 2009	2.1
Jul 2009	2.1
Aug 2009	2.1
Sep 2009	2.1
Oct 2009	2.1
Nov 2009	2.1
Dec 2009	2.1
Jan 2010	2.1
Feb 2010	2.1
Mar 2010	2.1
Apr 2010	2.1
May 2010	2.1
Jun 2010	2.1
Jul 2010	2.1
Aug 2010	2.1
Sep 2010	2.1
Oct 2010	2.1
Nov 2010	2.1
Dec 2010	2.1
Jan 2011	2.1
Feb 2011	2.1
Mar 2011	2.1
Apr 2011	2.1
May 2011	2.1
Jun 2011	2.1
Jul 2011	2.1
Aug 2011	2.1
Sep 2011	2.1
Oct 2011	2.1
Nov 2011	2.1
Dec 2011	2.1
Jan 2012	2.1
Feb 2012	2.1
Mar 2012	2.1
Apr 2012	2.1
May 2012	2.1
Jun 2012	2.1
Jul 2012	2.1
Aug 2012	2.1
Sep 2012	2.1
Oct 2012	2.1
Nov 2012	2.1
Dec 2012	2.1
Jan 2013	2.1
Feb 2013	2.1
Mar 2013	2.1
Apr 2013	2.1
May 2013	2.1
Jun 2013	2.1
Jul 2013	2.1
Aug 2013	2.1
Sep 2013	2.1
Oct 2013	2.1
Nov 2013	2.1
Dec 2013	2.1
Jan 2014	2.1
Feb 2014	2.1
Mar 2014	2.1
Apr 2014	2.1
May 2014	2.1
Jun 2014	2.1
Jul 2014	2.1
Aug 2014	2.1
Sep 2014	2.1
Oct 2014	2.1
Nov 2014	2.1
Dec 2014	2.1
Jan 2015	2.1
Feb 2015	2.1
Mar 2015	2.1
Apr 2015	2.1
May 2015	2.1
Jun 2015	2.1

**10 Month Inflation Rates (%) - Jun 2015-May 2007**

MONTH	INFLATION
Jun 2007	2.2
Jul 2007	2.4
Aug 2007	2.1
Sep 2007	2.1
Oct 2007	2.4
Nov 2007	2.4
Dec 2007	2.5
Jan 2008	2.4
Feb 2008	2.1
Mar 2008	2.1
Apr 2008	2.0
May 2008	2.0
Jun 2008	2.0
Jul 2008	2.1
Aug 2008	2.1
Sep 2008	2.1
Oct 2008	2.1
Nov 2008	2.1
Dec 2008	2.1
Jan 2009	2.1
Feb 2009	2.1
Mar 2009	2.1
Apr 2009	2.1
May 2009	2.1
Jun 2009	2.1
Jul 2009	2.1
Aug 2009	2.1
Sep 2009	2.1
Oct 2009	2.1
Nov 2009	2.1
Dec 2009	2.1
Jan 2010	2.1
Feb 2010	2.1
Mar 2010	2.1
Apr 2010	2.1
May 2010	2.1
Jun 2010	2.1
Jul 2010	2.1
Aug 2010	2.1
Sep 2010	2.1
Oct 2010	2.1
Nov 2010	2.1
Dec 2010	2.1
Jan 2011	2.1
Feb 2011	2.1
Mar 2011	2.1
Apr 2011	2.1
May 2011	2.1
Jun 2011	2.1
Jul 2011	2.1
Aug 2011	2.1
Sep 2011	2.1
Oct 2011	2.1
Nov 2011	2.1
Dec 2011	2.1
Jan 2012	2.1
Feb 2012	2.1
Mar 2012	2.1
Apr 2012	2.1
May 2012	2.1
Jun 2012	2.1
Jul 2012	2.1
Aug 2012	2.1
Sep 2012	2.1
Oct 2012	2.1
Nov 2012	2.1
Dec 2012	2.1
Jan 2013	2.1
Feb 2013	2.1
Mar 2013	2.1
Apr 2013	2.1
May 2013	2.1
Jun 2013	2.1
Jul 2013	2.1
Aug 2013	2.1
Sep 2013	2.1
Oct 2013	2.1
Nov 2013	2.1
Dec 2013	2.1
Jan 2014	2.1
Feb 2014	2.1
Mar 2014	2.1
Apr 2014	2.1
May 2014	2.1
Jun 2014	2.1
Jul 2014	2.1
Aug 2014	2.1
Sep 2014	2.1
Oct 2014	2.1
Nov 2014	2.1
Dec 2014	2.1
Jan 2015	2.1
Feb 2015	2.1
Mar 2015	2.1
Apr 2015	2.1
May 2015	2.1
Jun 2015	2.1

STORIES CREATED  
THROUGH USER INPUT  
*continued*


Figure 29 - Source: Downloaded from <http://www.statistics.gov.uk/PIC/index.html>

IMPROVING STATIC  
TWO-DIMENSIONAL  
GRAPHS

Graphing raw data "often leaves important aspects of data undiscovered" (Cleveland, 1993, p. 1). This is often the case with many graphs failing to appropriately portray their message, either through inappropriate or neglectful metadata or by simply not highlighting stories evident within the data. Static two-dimensional graphs have existed for centuries and they will always have a place and be needed in statistical presentations. However this does not mean that improvements and new techniques can not be implemented to help bring these stories to the surface.

A simple technique to improve a time-series graph is to provide context to various data points through the use of news stories and background information. The example from the BBC news website below is a very good example of this. An area or line graph is used to illustrate the US Presidents approval rating over the past ten years, which in itself is a useful graph. However three different pieces of metadata are used to add further value to this graph.





**7 August 1998: Embassy bombings**

The US embassies in Nairobi, Kenya and Dar-es-Salaam, Tanzania were hit simultaneously by car bombs. More than 230 people were killed and at least 4,000 were injured, most of them African citizens. An al-Qaeda cell was linked to the bombings and awareness of Osama Bin Laden began to rise around the world.

IMPROVING STATIC  
TWO-DIMENSIONAL  
GRAPHS *continued*

Figure 30 - Source: Downloaded from [http://news.bbc.co.uk/2/hi/in\\_depth/629/629/5305868.stm](http://news.bbc.co.uk/2/hi/in_depth/629/629/5305868.stm)

### IMPROVING STATIC TWO-DIMENSIONAL GRAPHS *continued*

The most notable piece of metadata is the addition of clickable buttons that open news stories relating to time periods in the graph. This simple feature adds a great deal of value in that they help to explain the graph or show what occurred in response to an event. For example, notice the story that unfolds in this example when George Bush's presidential approval rating climbed dramatically as a result of how he responded to the September 11 attacks.

The second piece of metadata is the brief story that is told explaining the key outcomes of the graph. This paragraph above the graph provides a quick description of the story being presented. It is critical however, that this story does not exceed a few sentences as otherwise its meaning will be lost. Its purpose should be complimentary to, and provide users with the main points of the graph.

The third piece of metadata is the tabular format used in this example. Note how the presidential rating is only a small part of the wider terrorism story and the various tabs provide further value. More often than not stories are part of a bigger picture and providing linkage or context to similar stories is very beneficial.

Another example of providing contextual information to a graph can be seen in Figure 3.9.2. However instead of using news stories and clickable buttons, transactions and reference numbers are used. This format would be simpler to implement and may work better for illustrating statistical changes, however it would not have the visual impact of the BBC example.

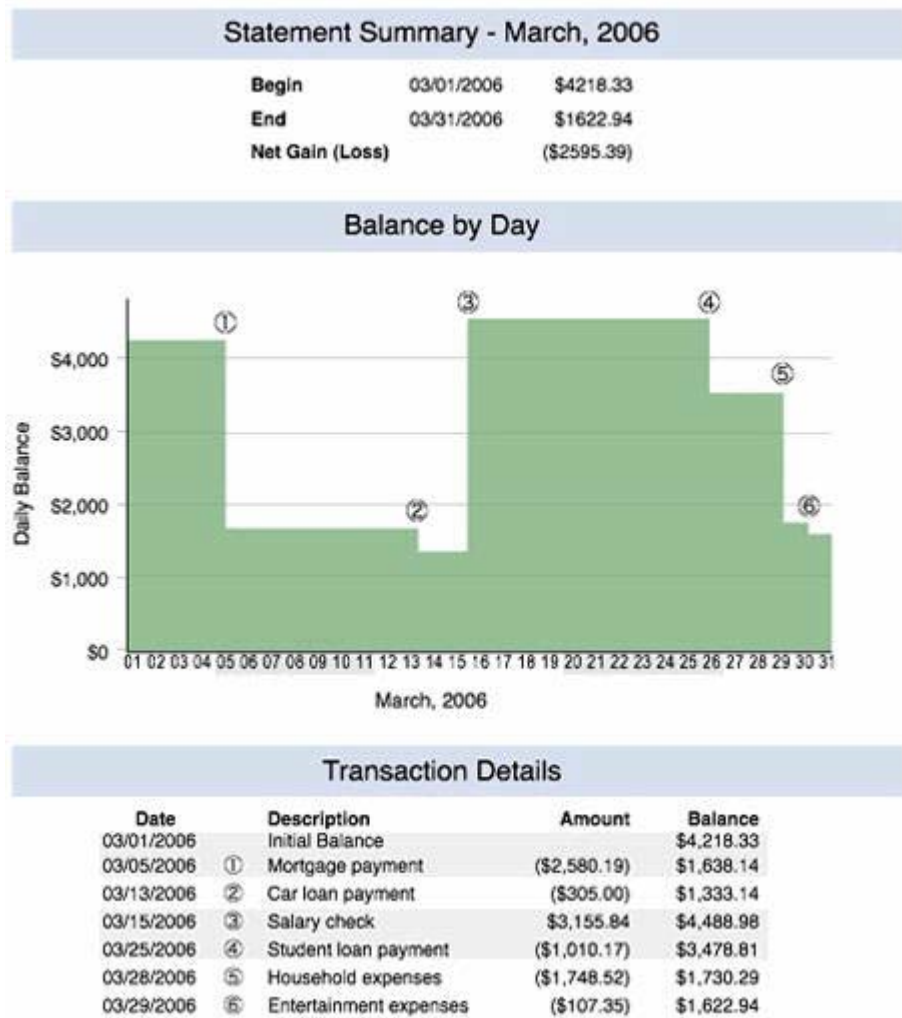


Figure 31 - Source: Downloaded from <http://www.b-eye-network.co.uk/view-articles/3354?PHPSESSID=f7405bbe901fd588946927f5a4ab3c93>

A problem with many static two-dimensional graphs is their inability to display multiple variables. Although interactive graphs may have the ability to change the variables on each axis, move the data around, change the time period and show animation, static 2D graphs still have their place and role in telling statistical stories.

Showing multiple variables on a graph is not necessarily a difficult task. However the difficulty lies in presenting the information in a format that is easy to view, understand and take stories away from. Figure 32 is an excellent example of this. In simple terms the graph is trying to display when various companies advertise on television over a certain weekend.



IMPROVING STATIC  
TWO-DIMENSIONAL  
GRAPHS *continued*

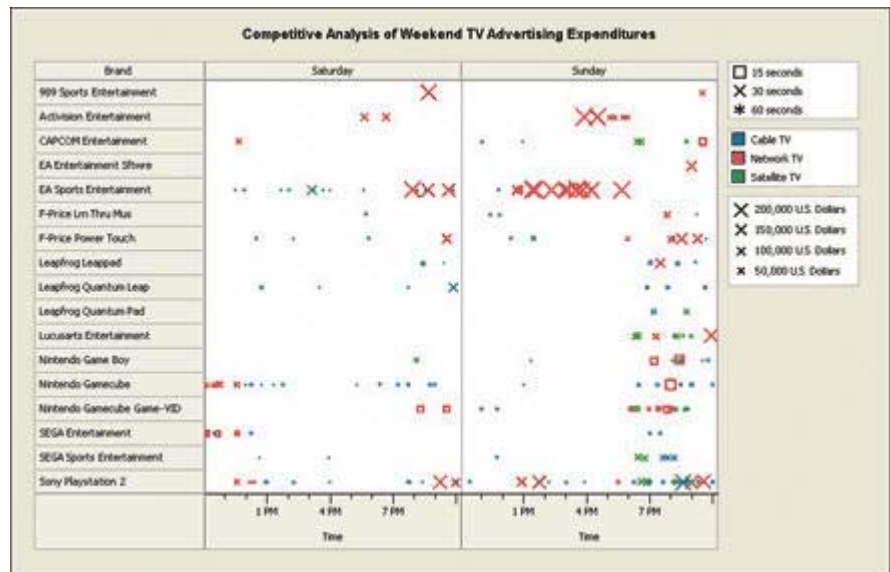


Figure 32 - Source: Downloaded from [http://www.dmreview.com/article\\_sub.cfm?articleId=1038100](http://www.dmreview.com/article_sub.cfm?articleId=1038100)

Note how time is incorporated into this graph. Instead of a standard x-axis time scale the days are illustrated at the top and the hours being analysed can be seen at the base of the graph. It also only shows the time from 12pm to 12am each day, which is the prime advertising period. This is very beneficial in that narrowing the time period down to the smallest amount possible without losing any critical data will assist with usability.

The second key feature to illustrating this data can be seen in the three legends on the right hand side of the graph. They help to show that the points within the graph each have an element of data associated with them based on their shape, colour and size. The shape illustrates how long the commercial was, the colour indicates on what type of television the commercial was aired on and the size illustrates how much money was paid for the advertising. Based on all these factors stories unfold within the graph. For example it can be seen that EA Sports Entertainment spend a lot of money and focus heavily on advertising on a Sunday afternoon on Network television. If thought about, a strong marketing ploy unfolds that focuses on selling sporting games to an audience enjoying watching the football or baseball on a Sunday afternoon.

The key to adding multiple variables to a graph lies in its balance. A balance must be achieved so that one variable does not overpower the rest of the graph and that all graphs are distinguishable in shape, colour, size etc. Reference points and scales are also very important. For example it is difficult to identify the data points associated with Leapfrog Quantum Pad due to the fact that they only advertise late on a Sunday. Many users may find it difficult to move their eyes straight across the page and reference the data points.

Another example of showing multiple variables in this format can be seen in Figure 33. It takes great advantage of the x-axis to show time-series data by dividing it by department and maintaining a constant time period for each. In a glance users can witness how each department of this business has performed and at an approximate expense level. The colour coding of total, exempt and non-exempt expenses is useful, although additional

IMPROVING STATIC  
TWO-DIMENSIONAL  
GRAPHS *continued*

value could be provided by illustrating targets or allocated expense budgets through either a dotted line graph or column graph in each department.

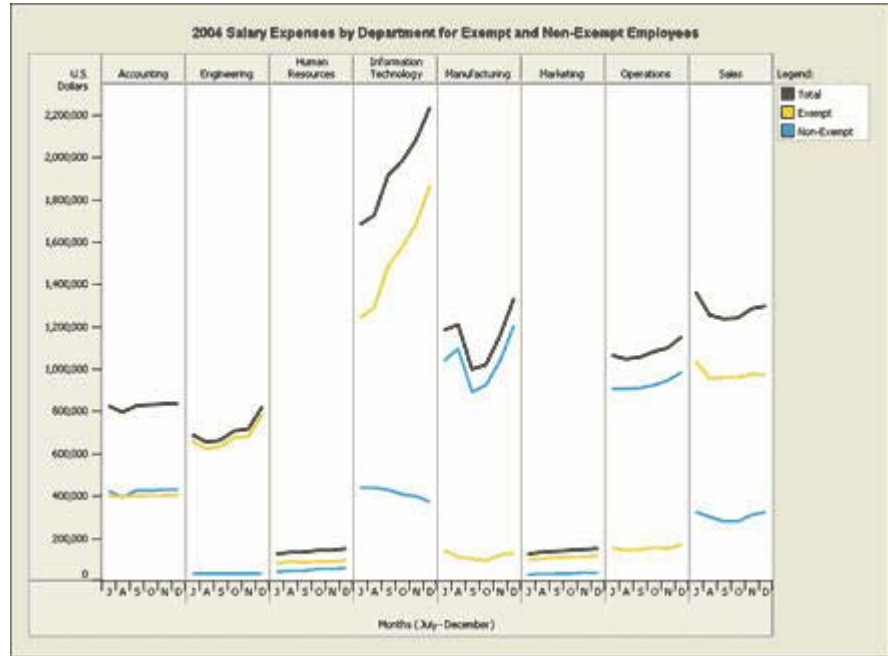


Figure 33 - Source: Downloaded from [http://www.dmreview.com/article\\_sub.cfm?articleId=1031173](http://www.dmreview.com/article_sub.cfm?articleId=1031173)

NATIONMASTER

NationMaster is a website that describes itself as "a handy way to graphically compare nations." (NationMaster, 2006) It collects data from a wide range of respected statistical sources, including the CIA World Factbook, the UN and OECD, to provide a "central data source."

Figure 34 illustrates one effective way to search for statistics on this website. Through drop-down boxes and a selection box, users are able to quickly find and possibly compare statistics. The additional information in brackets assists users in making sure they are presented with the statistics they have requested. For accessibility a text version of these lists is also presented, as well as a search function so that all users should be able to find their desired statistics.

**FACTS & STATISTICS** [Simple view](#)

- Select countries to view:** (hold down Control key and click to select several)
  - All countries
  - OECD
  - Central America and the Caribbean
  - Europe
  - Oceania
  - Asia
- Select fact or statistic:** (\* = graphable)
  - Food
  - Beer consumption \*
- (OPTIONAL) Compare to statistic:** (both need to be graphable)
  - Lifestyle
  - Happiness level > Very happy \*
- View result as:**
  - Scatterplot with ...
  - Flags

Compare: All

**Go »**

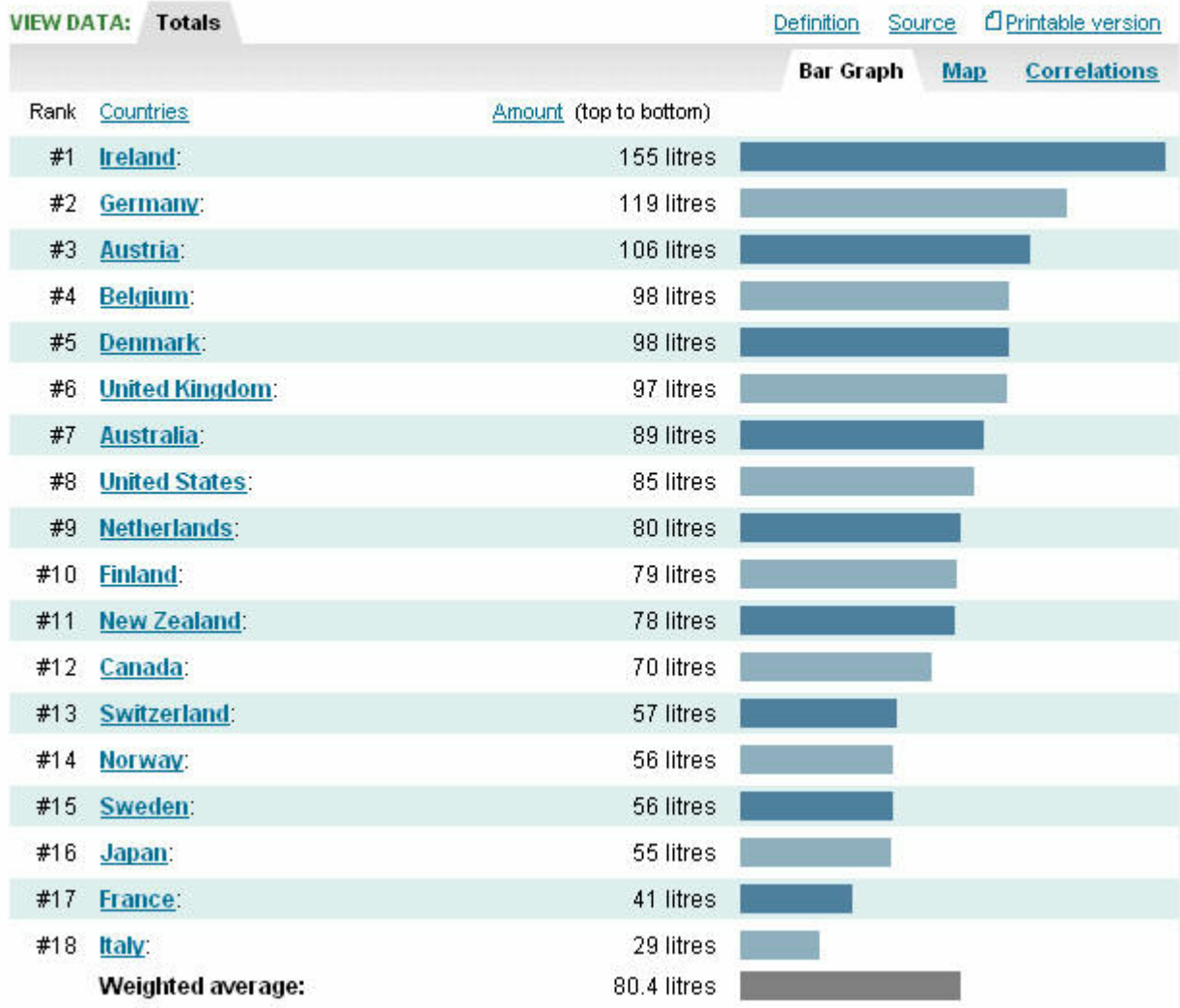
(OR) SEARCH ALL encyclopedia, stats & forums:  **Search »**

NATIONMASTER *continued*

Figure 34 - Source: Downloaded from www.nationmaster.com

There are four graphical means of displaying statistical data with NationMaster, available by selecting the 'View result as:' drop-down list. The entry point and most basic of these is the Bar Graph. It reflects the ideas of sparklines as presented in Figure 35 relating to in-cell bar graphs. The length of the bar refers to the percentage value of the statistic in relation to the first ranking country.

**Food Statistics > Beer consumption by country**



NATIONMASTER *continued*

Figure 35 - Source www.nationmaster.com

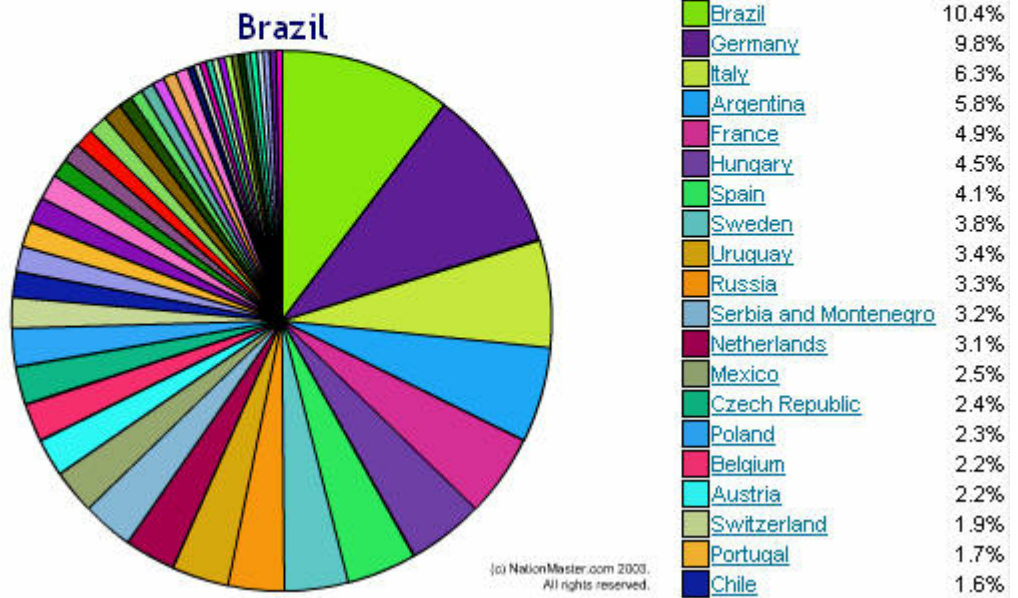
A Pie Graph is presented as another way of displaying the statistical data in NationMaster. It has a mouse-over function, which highlights the country being represented by a segment, plus the ability to click on a segment and drill-down to further information. Figure 36 is an example. A major problem with using a pie graph can be seen in this example. With so many countries being represented by a segment it appears crowded, cluttered and difficult to distinguish. However, NationMaster assists with this problem by the ability to view only the top 5 countries, thus reducing the clutter associated with the smaller figures.

**Sports Statistics > World Cup Totals > Goals for by country**

VIEW DATA: **Totals**

[Definition](#) [Source](#) [Printable version](#)

[Bar Graph](#) [Pie Chart](#) [Map](#)



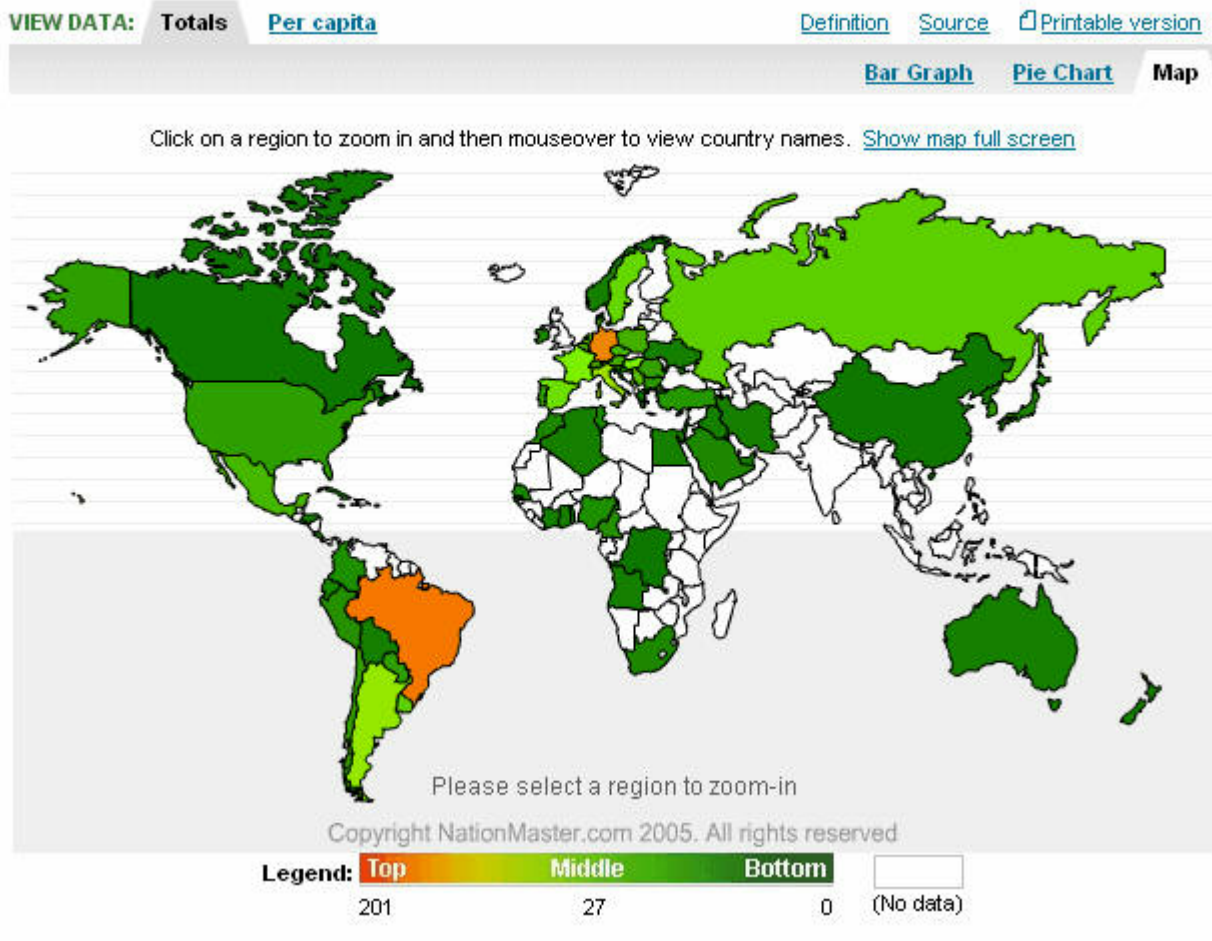
NATIONMASTER *continued*

Figure 36 - Source: Downloaded from www.nationmaster.com

The third way of displaying data is through a map. It essentially works the same way as the pie graph in that colours are used to represent the country's data, by hovering over an area the name of the country is displayed and by clicking on the country the user will be able to drill-down to further information. The other useful features include the ability to zoom in and the full screen options. The zoom in function works really well in that it operates quickly and allows the user to get a closer look at regions like Europe where there are many small nations. The full screen option is a simple feature that does exactly what its name suggests. From a usability perspective, these ideas are only small, but provide additional value for many users.



**Sports Statistics > World Cup Totals > Goals for by country**



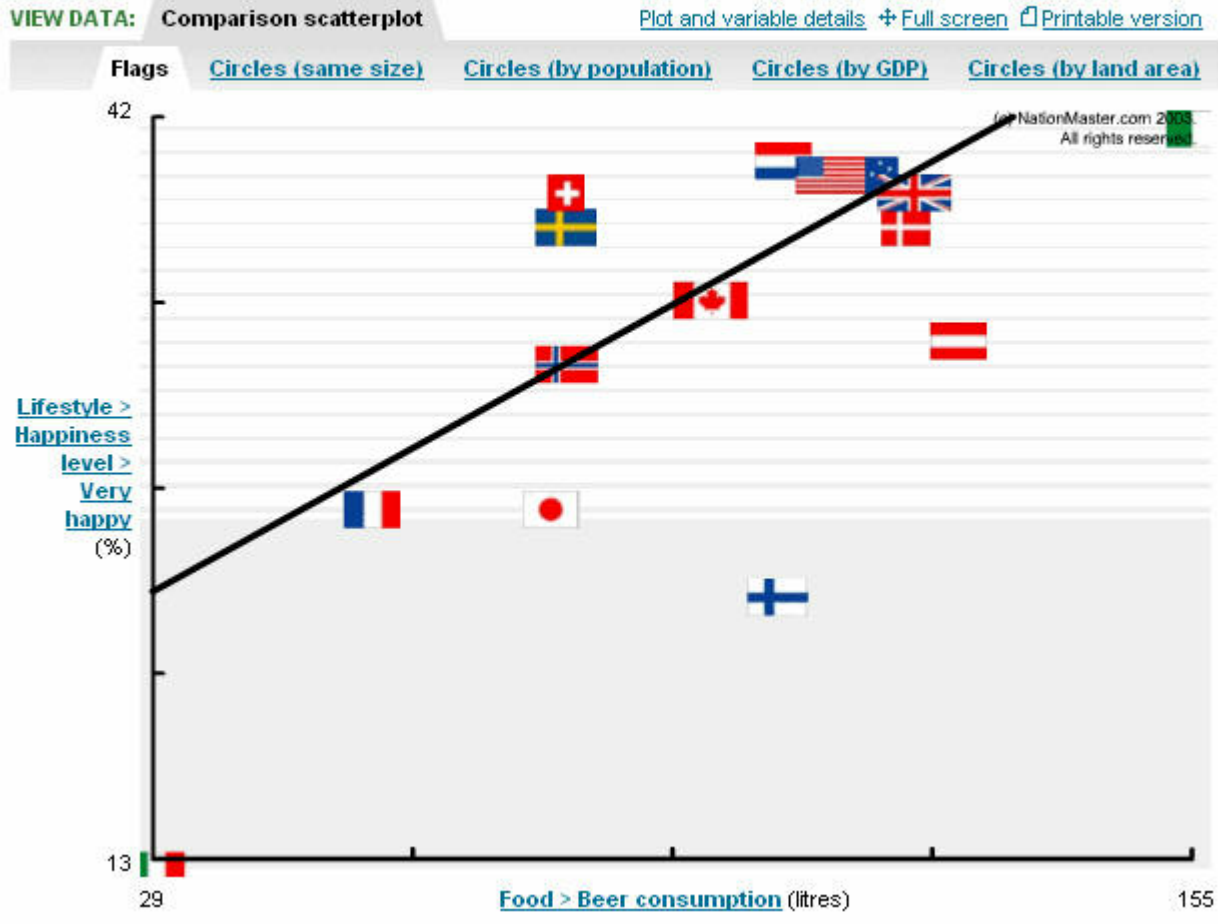
NATIONMASTER *continued*

Figure 37 - Source: Downloaded from [www.nationmaster.com](http://www.nationmaster.com)

A fourth graphical technique is the Comparison scatterplot, as seen in Figure 38. It reflects the Gapminder concept of comparing any two variables on a graph with the only difference being that there is no animation of the time variable. This may appear to be major disadvantage, however from a technical and usability perspective this tool has its advantages. Without a Gapminder product currently available this option may present a simpler technique that could be produced and disseminated quickly. It also takes away the movement and animation that may confuse or distract many users.

To compare two variables the search tool in Figure 34 can be used or alternatively when viewing certain statistical pages the option is available to view correlations, as seen in the tab structure of Figure 35. This correlations list provides the user with a list of the most accurate correlations that exist in relation to the chosen statistic. Any two statistics with graphical capabilities have the ability to be compared in this format.

**Food > Beer consumption (vs) Lifestyle > Happiness level > Very happy**



NATIONMASTER *continued*

Figure 38 - Source: Downloaded from www.nationmaster.com

Another great feature of this technique is the ability to change the icon relating to the data points. The use of flags for countries is instantly recognisable for many users, although it does rely upon some user knowledge. The other option to use circles also works well and allows for the introduction of a further variable in population, GDP or land area. Further drill-down possibilities, by clicking on the flags or circles, is also available.

## CONCLUSION

---

### CONCLUSION

There are currently a variety of Data Visualisation tools and techniques available or being used on the web. Each aims to tell a story or inform its audience in a quicker or more intuitive way. However audience reactions will always vary to any technique that is used. The key is to find those applications that grasp the attention, invoke thought, inform and possibly entertain the target audience.

Although many applications will either look attractive or function appropriately that does not mean they should automatically be adopted. The technique must fit smoothly with the raw statistical data and be applied in the right context.

Data Visualisation for statistical organisations such as the ABS should be about statistical storytelling. It should be about changing public perception and adding alternatives that will enable users to more fully understand and use statistics.



## REFERENCE LIST

### REFERENCE LIST

#### Text references

- Asahi, T, Turo, D & Shneiderman, B 1995, *Visual decision-making: Using treemaps for the Analytic Hierarchy Process*, University of Maryland, College Park, viewed 15 October 2006, <[http://www.sigchi.org/chi95/proceedings/videos/ta\\_bdy.htm](http://www.sigchi.org/chi95/proceedings/videos/ta_bdy.htm)>
- Australian Bureau of Statistics 2006, *Data Communication - Emerging International Trends and Practices of the Australian Bureau of Statistics*, Australian Bureau of Statistics, Canberra, viewed 20 June 2007, <<http://www.abs.gov.au/Ausstats/abs@.nsf/Latestproducts/1211.0Contents12006?opendocument&tabname=Summary&prodno=1211.0&issue=2006&num=&view=>>>
- Bednarz, A & Dubie, D 2006, 'Dashboards called a key to net mgmt', *Network World*, vol. 23, iss. 25, pp. 1-2, viewed 30 July 2006, ABI/INFORM Global, document ID: 1077209041.
- Blum, A 2006, *Graphing the Development Gap*, *Business Week*, Red Oak, viewed 20 September 2006, <[http://www.businessweek.com/innovate/content/feb2006/id20060222\\_010217.htm](http://www.businessweek.com/innovate/content/feb2006/id20060222_010217.htm)>
- Cleveland, W 1993, *Visualizing Data*, Hobart Press, Summit, New Jersey.
- Clusty 2006, *Clusty Cloud Creator*, Clusty, Pittsburgh, viewed 26 October 2006, <<http://cloud.clusty.com/>>
- Few, S 2004, *Dashboard Confusion*, Perceptual Edge, Berkeley, viewed 30 July 2006, <<http://www.perceptualedge.com/article1.htm>>
- Few, S 2005, *Data Visualization: Intelligent Dashboard Design*, *DM Review*, Brookfield, viewed 30 July 2006, <[http://www.dmreview.com/article\\_sub.cfm?articleId=1035522](http://www.dmreview.com/article_sub.cfm?articleId=1035522)>
- Fisher, R 2006, *Edinburgh Online Graphics Dictionary*, University of Edinburgh, Edinburgh, viewed 25 July 2006, <<http://homepages.inf.ed.ac.uk/rbf/GRDICT/grdict.htm>>.
- Gibbs, M 2006, 'Data-intense, design-simple graphics', *Network World*, vol. 23, iss. 36, p. 36, viewed 30 July 2006, ABI/INFORM Global, document ID: 1136353571.
- Hoekstra, M 2006, *Search Clouds Increase Relevance of Search Results*, *Clever Clogs*, Rijswijk, viewed 26 October 2006, <[http://www.cleverclogs.org/2006/10/search\\_clouds\\_i.html](http://www.cleverclogs.org/2006/10/search_clouds_i.html)>
- Johnson, J 2006, 'Technology's next wave: the dashboard', *Network World*, vol. 23, iss. 26, p. 25, viewed 30 July 2006, ABI/INFORM Global, document ID: 1081404141.
- Nagy, J 2006, *Search Clouds*, arc90, New York, viewed 30 September 2006, <[http://lab.arc90.com/2006/10/search\\_clouds.php](http://lab.arc90.com/2006/10/search_clouds.php)>
- NationMaster 2006, *Nationmaster*, Rapid Intelligence, Sydney, viewed 7 December 2006, <<http://www.nationmaster.com/index.php>>
- Nielsen, J 2003, *Will plain-text ads continue to rule?*, Jakob Nielsen's Alertbox, 23 April 2003, useit.com, Fremont, viewed 30 September 2006, <<http://www.useit.com/alertbox/20030428.html>>

## REFERENCE LIST *continued*

### *Text references continued*

Palace, B 1996, *What is Data Mining?*, Anderson Graduate School of Management at UCLA, Los Angeles, viewed 22 June 2007, <<http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining.htm>>

Shneiderman, B. 2006, *Treemaps for space-constrained visualization of hierarchies*, University of Maryland, College Park, viewed 15 October 2006, <<http://www.cs.umd.edu/hcil/treemap-history/>>

Tufte, E 2001, *The Visual Display of Quantitative Information*, Graphic Press, Cheshire, Conn.

Tufte, E 2006, *Beautiful Evidence*, The work of Edward Tufte and Graphics Press, Cheshire, viewed 2 September 2006, <[http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg\\_id=0001OR&topic\\_id=1](http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001OR&topic_id=1)>

United Nations Economic Commission for Europe (UNECE) 2006, *Making Data Meaningful*, United Nations Economic Commission for Europe, Geneva, viewed 29 October 2006, <<http://www.unece.org/stats/documents/writing/what.htm>>

Walker, C & Antanies, J 2006, 'Data Visualization Tool Empowers Operators through Easier Data Access', *Pulp & Paper*, vol. 80, iss. 3, pp. 40-43, viewed 15 October 2006, ABI/INFORM Global, document ID: 1009615141.

Wikipedia 2006, *Treemapping*, Wikipedia, San Diego, viewed 15 October 2006, <<http://en.wikipedia.org/wiki/Treemapping>>

### *Image references*

*Ask E.T.: Sparklines: theory and practice.* (2006) [Online image]. Available from: <[http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg\\_id=0001OR&topic\\_id=1](http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001OR&topic_id=1)&topic=> [accessed 1 September 2006].

*Ask E.T.: Sports graphics.* (2006) [Online image]. Available from: <[http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg\\_id=0001lh&topic\\_id=1](http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001lh&topic_id=1)> [accessed 21 September 2006].

*BBC News.* (2006) [Online image]. Available from: <[http://news.bbc.co.uk/2/hi/in\\_depth/629/629/5305868.stm](http://news.bbc.co.uk/2/hi/in_depth/629/629/5305868.stm)> [accessed 30 July 2006].

*Bissantz sparkline tools.* (2006) [Online image]. Available from: <<http://www.bissantz.de/sparklines/index.asp>> [accessed 20 September 2006].

*BonaVista Systems.* (2006) [Online image]. Available from: <<http://www.bonavistasystems.com/index.html>> [accessed 1 September 2006].

*Business Cycle Tracer.* (2006) [Online image]. Available from: <<http://www.cbs.nl/en-GB/menu/themas/macro-economie-financiele-instellingen/conjunctuurgegevens/publicaties/conjunctuurbericht/klok/ck-homepage.htm>> [accessed 29 July 2006].

## REFERENCE LIST *continued*

### *Image references continued*

*Clusty Cloud Creator*. (2006) [Online image]. Available from: <<http://cloud.clusty.com/search?v%3Aproject=clusty-cloud&cc=1&query=dennis+trewin&borderstyle=simple&border-color=123C92&linkstyle=color&color=123C92&width=auto&width-px=500>> [accessed 26 October 2006].

*Clusty Cloud Creator*. (2006) [Online image]. Available from: <<http://cloud.clusty.com/search?v%3Aproject=clusty-cloud&cc=1&query=australian+bureau+of+statistics&borderstyle=simple&border-color=00FF7F&linkstyle=color&color=2E8B57&width=auto&width-px=500>> [accessed 26 October 2006].

*Clusty Cloud Creator*. (2006) [Online image]. Available from: <<http://cloud.clusty.com/search?v%3Aproject=clusty-cloud&cc=1&query=australian+social+trends&borderstyle=simple&border-color=228B22&linkstyle=color&color=4B0082&width=fixed&width-px=500>> [accessed 26 October 2006].

*Data Visualization: Creative Visualization*. (2006) [Online image]. Available from: <[http://www.dmreview.com/article\\_sub.cfm?articleId=1031173](http://www.dmreview.com/article_sub.cfm?articleId=1031173)> [accessed 28 September 2006].

*del.icio.us most popular treemap*. (2006) [Online image]. Available from: <<http://codecubed.com/map.html>> [accessed 1 November 2006].

*Exxon Secrets*. (2006) [Online image]. Available from: <<http://www.exxonsecrets.org>> [accessed 2 September 2006].

*Flickr: Explore*. (2006) [Online image]. Available from: <<http://www.flickr.com/explore/>> [accessed 30 September 2006].

*Gapminder*. (2006) [Online image]. Available from: <<http://www.gapminder.org>> [accessed 12 July 2006].

*Graph Designs for Rapidly Assessing Budget Performance*. (2006) [Online image]. Available from: <<http://www.b-eye-network.com/view/3224#thumb>> [accessed 28 August 2006].

*Graph Designs for Reviewing Transactions and the Changing Balance*. (2006) [Online image]. Available from: <<http://www.b-eye-network.co.uk/view-articles/3354?PHPSESSID=f7405bbe901fd588946927f5a4ab3c93>> [accessed 28 September 2006].

*How much is it worth?* (2006) [Online image]. Available from: <[http://www.ssb.no/kpi\\_en/kpicalc.html](http://www.ssb.no/kpi_en/kpicalc.html)> [accessed 29 September 2006].

*Index calculator*. (2007) [Online image]. Available from: <<http://www.destatis.de/basis/e/preis/start.htm>> [accessed 10 January 2007].

*Lab Escape*. (2006) [Online image]. Available from: <<http://www.labescape.com/>> [accessed 1 November 2006].

*Lightweight data exploration in Excel*. (2006) [Online image]. Available from: <<http://www.juiceanalytics.com/weblog/?p=236>> [accessed 21 September 2006].

*liveplasma music, movies, search engine and discovery engine*. (2006) [Online image]. Available from: <<http://www.liveplasma.com/>> [accessed 20 September 2006].

## REFERENCE LIST *continued*

---

### *Image references continued*

*mnemomap*. (2006) [Online image]. Available from: <<http://mnemo.org/>> [accessed 20 September 2006].

*Nationmaster*. (2006) [Online image]. Available from: <<http://www.nationmaster.com>> [accessed 7 December 2006].

*newsmap*. (2006) [Online image]. Available from: <<http://www.marumushi.com/apps/newsmap/newsmap.cfm>> [accessed 21 July 2006].

*Personal Inflation Calculator*. (2007) [Online image]. Available from: <<http://www.statistics.gov.uk/PIC/index.html>> [accessed 20 June 2007].

*Sales Dashboard*. (2006) [Online image]. Available from: <<http://www.math.yorku.ca/SCS/Gallery/allison/scen3b.htm>> [accessed 30 July 2006].

*Search Clouds*. (2006) [Online image]. Available from: <[http://lab.arc90.com/2006/10/search\\_clouds.php](http://lab.arc90.com/2006/10/search_clouds.php)> [accessed 30 September 2006].

*Standby*. (2006) [Online image]. Available from: <<http://energy.failedrobot.com/standby.html>> [accessed 11 September 2006].

*StockMorph*. (2006) [Online image]. Available from: <<http://www.stockmorph.com/sparklines-remote-module-or-gadget-for-google-home-page/>> [accessed 20 September 2006].

*The Hive Group*. (2006) [Online image]. Available from: <<http://www.hivegroup.com>> [accessed 1 November 2006].

*Threat Analysis Reporter*. (2006) [Online image]. Available from: <<http://www.8e6.com/products/pdfs/8e6-Threat-Analysis-Reporter-Data-Sheet.pdf>> [accessed 30 July 2006].

*unstats* | *Millenium indicators*. (2006) [Online image]. Available from: <<http://mdgs.un.org/unsd/mdg/Default.aspx>> [accessed 11 July 2006].







## FOR MORE INFORMATION . . .

*INTERNET* **www.abs.gov.au** the ABS web site is the best place for data from our publications and information about the ABS.

*LIBRARY* A range of ABS publications are available from public and tertiary libraries Australia wide. Contact your nearest library to determine whether it has the ABS statistics you require, or visit our web site for a list of libraries.

## INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our web site, or purchase a hard copy publication. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

*PHONE* 1300 135 070

*EMAIL* [client.services@abs.gov.au](mailto:client.services@abs.gov.au)

*FAX* 1300 135 211

*POST* Client Services, ABS, GPO Box 796, Sydney NSW 2001

## FREE ACCESS TO STATISTICS

All statistics on the ABS web site can be downloaded free of charge.

*WEB ADDRESS* [www.abs.gov.au](http://www.abs.gov.au)



2000001566534

ISBN 9780642483560

RRP \$11.00