

PRINCIPLES OF INFORMATION VISUALIZATION FOR BUSINESS RESEARCH

Ioan I. ANDONE^{*}

Abstract: *In the era of data-centric-science, a large number of visualization tools have been created to help researchers understand increasingly rich business databases. Information visualization is a process of constructing a visual presentation of business quantitative data, especially prepared for managerial use. Interactive information visualization provide researchers with remarkable tools for discovery and innovation. By combining powerful data mining methods with user-controlled interfaces, users are beginning to benefit from these potent telescopes for high-dimensional spaces. They can begin with an overview, zoom in on areas of interest, filter out unwanted items, and then click for details-on-demand. With careful design and efficient algorithms, the dynamic queries approach to data exploration can provide 100 msec updates even for million-record databases. Visualizations of business information are therefore widely used in actually business decision support systems, and by business researchers also. Visual user interfaces called dashboards are tools for reporting the status of a company and its business environment to facilitate business intelligence and performance management activities. In this study, we examine the research on concepts, and the principles of business information visualization, because we hope to be using correctly by business Ph.D. students in their researches. Visual representations are likely to improve business managers, and business researchers efficiency, offer new insights, and encouraging comparisons.*

Keywords: *business quantitative data, visual presentation, visual interfaces, performance management, business intelligence*

1. INTRODUCTION

Two previous revolutions have relied on visual presentation of business information – the spread of PCs, which began reaching a much wider audience with the invention of the GUI (Graphical User Interface), and the spread of the Web (World Wide Web). When the business manager, business analyst or business researcher shift from desktop to the Web, it goes from a small and highly controlled environment to a vast, chaotic one, and yet the visual tools for handling the complexities of the Web are much less well-developed than the desktop. Even the desktop is a glorified filing system, all but useless for extracting patterns from large amounts of business information. Likewise, the ability to handle the large and

^{*} Al.I.Cuza University of Iași, Faculty of Economics and Business Administration, iandone@uaic.ro, +40-232-201600

growing quantities of business data stored in databases is very limited, as most of the tools we have for searching or extracting business information are visually thin.

Recent advances in information technology have enabled the automatic collection of massive amounts of business, accounting and non-accounting data from a wide variety of sources, including the Web. The challenge is for business information users, especially managers, analysts and researchers to think hard and elegantly. The rapid increase in the size of modern data stores also imposes new challenges for the efficient presentation of information. The data contained within a large data store can reveal significant, valuable and even previously unknown facts to people, if they are summarized and presented in an appropriate and illustrative manner. A visual format is often the best choice for this purpose.

Information visualization is an old concept, because the business statistical data has been visualized for over two centuries already (Tuft,2001). The goal of visualization is to create a graphical representation of abstract quantitative data that is concise and easy to interpret even when the amount of presented data is very large. Its foundation lies in the study of human visual perception. Understanding the properties of the visual system may explain why one image is considered clear and simple by most people and another one very difficult to perceive, even if those two images are merely different presentations of the same information. The presentation format is especially important in complex situations, when a lot of business information needs to be displayed in a small space, such as a single computer screen. An efficient visualization at its best is an extremely powerful cognitive tool that integrates the ingenious pattern-finding mechanisms of the human visual system with the computational power and information resources of modern computer systems (Ware, 2004).

Visualizing information involves not only collecting and processing the business data that are to be displayed, but also defining the graphical elements that will display them on the screen as well as the relationships between these elements. Displaying business data outside its logical context, without any context or even in the wrong context might lead to incorrect interpretation of the information it is supposed to convey.

However, when the appropriate data are presented together in the right context, they can help the reader to understand the situation they are depicting or even discover new relationships that have been previously hidden in the data. In addition to the context, the type of graphical element chosen to display certain data is of great importance.

There are numerous different types of graphical elements available for displaying business data, such as line graphs, bar graphs, sparklines and bullet graphs (Tuft,2001; Few,2006). Not all elements are suitable for different kinds of data, however. This only emphasizes the importance of choosing the display

element. On the other hand, numerical values do not always require a graphical representation - sometimes the information is best conveyed by showing the actual numbers instead of an image.

A decision support system is a general term for an information system used for acquiring business information for decision-making purposes. Decision support systems have been used for several decades in many different fields, such as industry, agriculture, medicine, or environmental crisis management, but they have had an important role especially in assisting business decision-making. For historical reasons, business decision support systems are known by many names, such as management information systems or executive information systems. *Recently, decision support systems in business have been focused to support business intelligence (BI) activities and performance management.* These concepts are related to the strategic and operational management of a company and involve collecting data both inside the company and from its business environment. The large quantities of data create a need for effective presentation of summarized information in order to avoid an “information overload”.

Information visualization techniques provide a solution to this problem, and therefore most tools for reporting (called dashboards) and analysis are based on visualization. Table 1 provides examples of commonly available visualization tools with applications for marketers and consumers. For each tool, Table 1 indicates whether it affects the visual perspective and/or the information context. We propose that the visual perspective and information context influence decision processes and outcomes by changing the decision-making frame—that is, what information a decision maker uses and how he or she uses it to gain insights and make decisions.

Table 1
Examples of Visualization Tools

Visualization Tool	Characteristics Affected	
	Visual Perspective	Information Context
TableLens (http://www.inxight.com/products/sdks/tl/) creates a visual representation of large amounts of tabular (e.g., spreadsheet) data, including an interactive interface that enables the user to sort columns, expand and contract rows, and drill down for more details.	✓	
SmartMoney.com’s MarketMap (http://www.smartmoney.com/marketmap/). A Treemap (i.e., a two-dimensional representation of hierarchical data in which each element is represented by a cell whose arrangement, size, and color represent attributes of that data element) application used for the reporting of stock portfolio information.	✓	✓
Newsmap (http://www.marumushi.com/apps/newsmap/newsmap.cfm). A Treemap application that visually reflects patterns in news reporting.	✓	✓
ArcGIS (http://www.esri.com/products.html). Geographic information software used for business-mapping applications, such as displaying results by sales territory or other regions.		✓
Lands’ End’s My Virtual Model (http://www.landsend.com/). An interactive virtual reality application that enables customers to build a virtual image of themselves and then “try” on clothing.	✓	

Visualization Tool	Characteristics Affected	
	Visual Perspective	Information Context
Fish-Eye Visualizations. Nonlinear magnification enables the user to see details of immediate interest (i.e., focus) and the overall picture (i.e., context). Examples include maps, charts, and text-based applications.	✓	
Spotfire DecisionSite v. 8.2.1 (http://support.spotfire.com/release/DecisionSite821/). Platform for integration with many different 3 rd party statistical tools, enabling experts to distribute statistically rich workflows throughout the organization	✓	✓

Note: Table 1 identifies which characteristics of visual representation a given tool is likely to affect: (1) the visual perspective (i.e., interactivity or depth of field) or (2) the information context (i.e., vividness, evaluability, or framing).

This study has two main objectives: 1) to review literature on the concepts and principles of information visualization and the underlying principles of visual perception in order to understand which factors influence the effectiveness of visual presentations of business data, and 2) to learn how these principles can be utilized in designing visual reporting and analysis interfaces in business intelligence and performance management systems.

The sources referred to in this study include Edward Tufte (Tufte, 2001; Tufte, 2006), Colin Ware (Ware, 2004), Ben Shneiderman (Shnei&Plais, 2004) and Stephen Few (Few, 2006). Tufte and Shneiderman are perhaps the best known specialists in the field of visualization; references to them ideas are found in numerous books and research articles throughout the field, as well as in many articles of cognitive science and psychology. Ware has a less general point of view, related more to computer science than Tufte's, and he also considers the technical aspects of visualization. His work, too, is very often cited. Few is not very well known in the academic community, since his playground is the business world. His ideas rely on Tufte and Ware to a great extent, but he has made some significant contributions of his own as well, especially related to dashboards. In addition, he tries to bridge the gap between academic visualization research and the business world, which, he argues, unfortunately are too far apart¹.

2. CONCEPTS AND PRINCIPLES OF INFORMATION VISUALIZATION

Information visualization is defined as *the use of interactive visual representations of abstract, nonphysically based data to amplify cognition* (Card et al,1999). This definition captures several key aspects of the field and explains its importance. First of all, constructing visual representations of business data takes full advantage of the capabilities of human visual perception and enable rapid finding of interesting and unknown patterns and relationships. Secondly, the nature of the data is abstract - such as revenue from sales information or the gain or loss on

¹ The IEEE InfoVis 2007 conference, <http://vis.computer.org/vis2007/session/tutorials.html#t1>

disposition. In fact, the term scientific visualization is used for visualization of data based on physical measurements, for example the depreciation of equipment of a company or costs of materials. Moreover, the overall purpose of visualization is to assist the user in understanding the meaning of the business data; to provide insight and increase the user's knowledge. It aims to help the user to complete cognitive tasks with little effort compared to e.g. textual representations.

Although the result is meant to be as simple to the user as possible, designing visualizations is anything but simple. Figure 1 illustrates different phases of the visualization process. It shows that visualization includes not only encoding the original data in a visual format by using different visual attributes, such as shape, size, position, orientation and color.

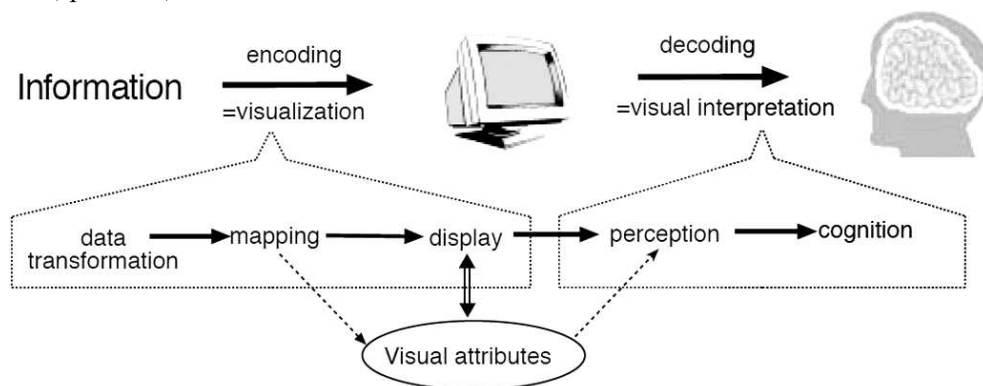


Figure 1 Information visualization process (Wünsche, 2004)

It also includes a decoding step: transforming the visual attributes into a mental representation in the brain of the viewer, and the patterns perceived in this representation are combined with knowledge stored in the long-term memory to finally construct the interpretation of the image.

This fact emphasizes the inter-disciplinary nature of the information visualization field: mastering it requires expertise in several fields, such as cognitive psychology, computer science, mathematics and statistics, and even art and architecture (Erbacher,2007). Also knowledge of the specific domain for which the visualization is created is crucial in order to make it effective.

2.1 Maximizing the business information content

Examples from existing statistical graphics indicate that many graphical displays of data are filled with irrelevant or redundant visual information, which only complicates understanding the actual content of the display and creates undesired "visual clutter". Tufte (Tufte,2001) has therefore presented the *concept of data-ink ratio* to measure the proportion of the graphic's ink (the data-ink²) that is

² Naturally the concept can be applied in the computer world just by replacing "ink" with "pixels"

used to present the actual information in the graphic: *Data-ink ratio = data-ink/ total ink used to print the graphic.*

One of the main principles in business statistical graphics should be to maximize the proportion of data-ink, which can be done e.g. using a method called erasing (Tuft,2001): editing the graphical content by removing all unnecessary components that represent the non-data-ink. Figure 2a shows a box plot, in which each data point is presented by a box and dashed lines.

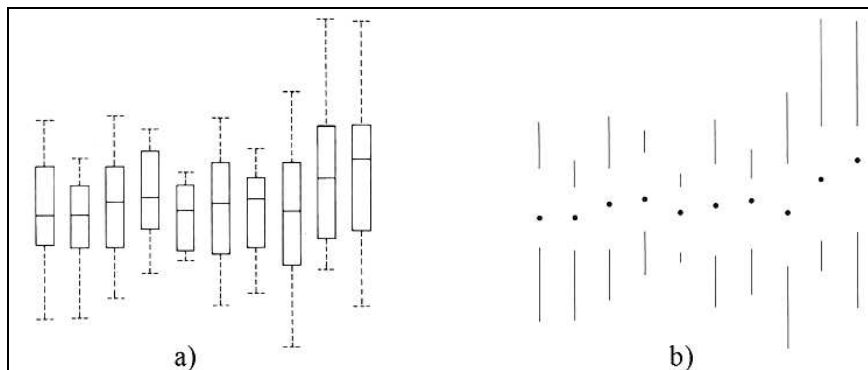


Figure 2 Maximizing the data-ink ratio (Tuft, 2001)

A single data point thus describes five numbers: median, high and low quartiles and minimum and maximum values. These elements are quite commonly used to present e.g. data from physical measurements or stock markets. Nevertheless, after erasing all non-data-ink the graphic in Figure 2b is what results. Both graphs now display the same five values for each data point, but the erased version looks much “lighter” and clearer. The median dots are now perceived as a continued line due to the Gestalt principle of continuity (Palmer,1999), allowing one to assess the trend of change in the data. Such a trend line is not easily seen in the original graph, because the large boxes distract the display.

It has been claimed (Tuft,2001) that the numerous graphical tools and effects available in commercial spreadsheet software have dramatically increased the amount of irrelevant information in statistical graphs. These tools and effects are provided to assist in creating different kinds of decorations, but while looking impressive and attractive, they severely obfuscate the information content. This is why Tuft (Tuft,2001) has introduced the term *chartjunk*, which is divided in three categories:

- *Unintentional optical art.* Using texture fill effects, such as thin parallel lines of different orientations, to fill up certain areas of the graph (for example the bars in a bar graph) often cause moiré effects³ that bring a sense of “vibration” or movement to the display. This is an effective way to cause irritation and to

³ A moiré effect is an optical illusion that occurs when thin lines are close to each other (Spillman,1993).

draw the viewer's attention away from the main target of interest of the data. Research (Tufte,2001) indicates that this is a surprisingly common phenomenon found in scientific journals, user manuals for computer graphics programs and even handbooks of statistical graphics.

- *The Grid.* Very often dark gridlines are present in statistical graphics, causing distraction and competing with the data (Tufte,2001). In many cases, however, the grid is very useful and helps to read and interpolate the data. Nevertheless, if a grid is present, it should not be too dense and preferably light in color.
- *Self-promoting graphics: The Duck.* An analogy to architecture describes this phenomenon: in United States, there is a store called the “Big Duck”, and the building itself has the form of a duck (Tufte,2001). This means that the graphic is taken over by decoration and the data measures and structures become merely design elements. Such a graph is only meaningful as an exhibit of graphical style. For example, the graph in Figure 3 vividly resembles a mountain landscape (there are even “snowcaps” on the “blue mountain” in the center), and the overall impression efficiently draws attention away from the main issue: the actual revenues.



Figure 3 Exhibiting graphical style rather than data

Tufte's principles and opinions about effective statistical graphics (Tufte,2001; Tufte,2006) may seem strange and perhaps provocative at first, but his observations are very sharp and his arguments well-founded.

2.2 Color coding

Humans can generally distinguish only about 200 different hues (Ware,2004; Goldstein,2007). The millions of colors in modern computer screens are therefore meaningful only for situations which do not require exact discrimination between each hue, e.g. scientific visualizations or viewing digital photographs of the real world. For purposes of color coding in graphical displays of statistical data where the accurate discrimination of hues is important, one must be careful when selecting the color palette. Colors that are too similar in terms of hue, lightness or contrast

might be easily confused. One way to avoid confusion is to choose all colors used for coding should from different categories of hues, meaning that for example multiple shades of green should not be used as color codes (Ware,2004). Figure 4 displays a set of 12 colors that are distinct enough so that they can “safely” be used for color coding without causing judgmental errors (Ware,2004).



Figure 4 The “safe” categories for color coding (Ware,2004)

The primary colors red, green, blue and yellow are naturally the most distinctive colors, so they should be the first four categories. After them, black and white are also easy to distinguish - although the use of black and white is somewhat questionable, since white is very often the color of the background and black the color of text.

One benefit of using color to label information is that it may ease the classification of data into separate categories that are in no particular order (Ware,2004). However, it easy to misuse color; the most common mistake is to use too much of it. A useful guideline is that if color is to be used for highlighting, the display should be quite homogeneous with respect to other colors, and the highlighting color should have a great contrast to other colors (Ware,2004; Tufte,2001). The extensive highlighting adds noise to the display, thus decreasing the signal-to-noise ratio (Tufte,2006): if every item is highlighted, in fact none are. It has even been proposed that in statistical graphics the data should always be presented in varying shades of gray (Tufte,2001), thus reserving other colors only for highlighting purposes and perhaps categorizing data values.

A serious problem with color is that some people have color deficiencies. The issue of using color in displays of business quantitative data is quite complicated and its usefulness seems to be somewhat questionable. So and Smith (So&Smith, 2002) point out that the effect of color in visual representations has not been studied very much, and most of the existing studies focus on educational or search and identification tasks. Thus very little is actually known about how color affects the perception of statistical graphics in decision-making tasks; it is often considered self-evident that color naturally eases the comprehension of graphical displays. However, this was not confirmed in So and Smith's experiment (So&Smith, 2002). They concluded that color coding results in performance benefit only when the task at hand is complex and that even in complex tasks the benefit applies only to females and has only a small effect.

We therefore conclude that in general, the use of color is not recommended in business statistical graphics, with the exception of special highlighting and categorization purposes as mentioned above. However, it has been pointed out that gray shades are not the only possible option (Few,2006): any palette containing a single hue with varying saturation and lightness levels is perceptually equivalent with the grayscale palette even for a color-deficient viewer, but may add to the visual aesthetics of the display.

2.3 Scaling

Scale is a very important factor in presenting business data. A good example of the importance of scale is a geographical map. If we look at a map of Europe, for instance, it is impossible to see any details of a certain city. On the other hand, if we look at a map of a single city, we miss the information about the surrounding regions outside the city. Just like the scale of a map, the scale of a graphical representation of business quantitative data may reveal important details (or hide them, if chosen poorly). The correct scale naturally depends on the characteristics of the data and the information that should be conveyed.

An example of a positive effect achieved by rescaling a graph is shown in Figure 5. It contains two graphs that display historical changes in the number of sunspots during 175 years. Both graphs clearly show that solar activity has peaks once in every eleven years. However, the graph in Figure 5b reveals another interesting fact: the higher peaks rise very fast and decline slowly, while the lower peaks are less dramatic.

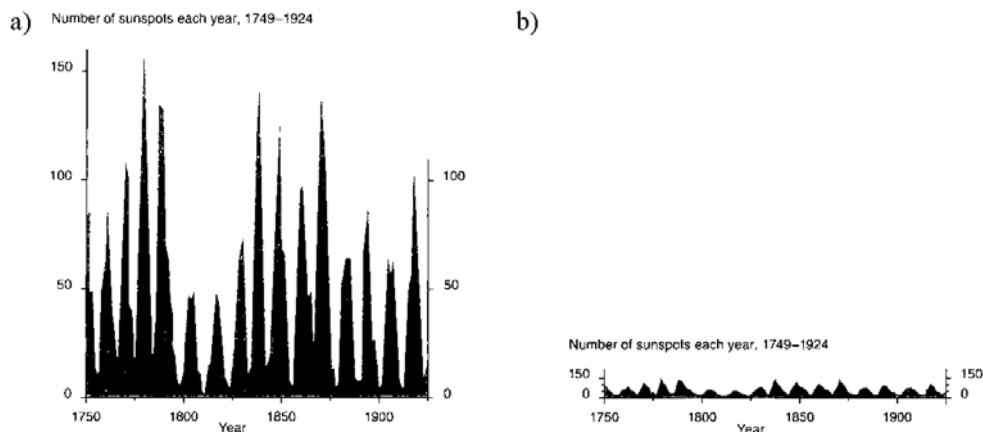


Figure 5 Two scales for historical solar activity data (Tuft,2001)

This cannot be easily seen in the graph of Figure 5a, which has a larger scale on the value axis. The relative heights of the peaks are still visible despite the change of scale, so the small graph presents the same information as the large one, and a little more.

The method used to change the scale is called banking to 45. It maximizes the discriminability of the orientations of line segments in the graph by finding an aspect ratio (width/height) such that the slopes are as close to 45 as possible. This method works especially well with cyclic data sets when the inclinations and declinations of the slopes are relatively similar, such as the sunspot data in Figure 5.

If the data contains many different kinds of slopes, some of which are steep and some gentle, a choice must be made about whether to emphasize the gentle slopes (and make the high peaks very sharp) or clarify the steep slopes. Naturally this choice depends on the data set and also the information that is considered most valuable for that data.

2.4 Ordering the business data set

Sometimes simply changing the ordering of business data points may reveal previously undiscovered patterns in the data. If the business data includes a time dimension “which is very common” it is an easy solution to use time as the ordering criteria. But the data usually includes many other dimensions as well, and ordering the data by some of those dimensions might be more useful than just ordering by time. Friendly and Kwan (Friendly&Kwan,2003) have studied the methods for ordering different kinds of multivariate data in visual displays. Their idea is that the data could be sorted by the effects to be observed. Several statistical methods for effect-ordering both numerical and categorical data. They may be generalized as optimization problems whose solutions may be expressed in terms of singular vectors, and the angles between these vectors provide the ordering for the data (Friendly&Kwan,2003). The ordering of categorical data for visualization has also been studied by Beygelzimer et al. (Beygelzimer et al, 2002), who have developed an algorithm for efficiently finding the optimal ordering of the values of two categorical variables in large data sets. Friendly and Kwan (Friendly&Kwan,2003) note that the ordering of data has great significance especially when the user's task is to perform comparison or to detect patterns, trends or anomalies in the graph. Information may be available in the display, but it might not be accessible if the ordering of the data is ineffective. Similar arguments can be found in other sources as well (Spence, 2001; Beygelzimer et al, 2002). Friendly and Kwan (Friendly&Kwan,2003). have studied the methods for ordering different kinds of multivariate data in visual displays. Their idea is that the data could be sorted by the effects to be observed. The ordering of categorical data for visualization has also been studied by Beygelzimer who have developed an algorithm for efficiently finding the optimal ordering of the values of two categorical variables in large data sets. Friendly and Kwan note that the ordering of data has great significance especially when the user's task is to perform comparison or to detect patterns, trends or anomalies in the graph. Information may be available in the display, but it might

not be accessible if the ordering of the data is ineffective. Similar arguments can be found in other sources as well (Spence, 2001; Beygelzimer et al, 2002).

Tufte (Tuft,2001) has introduced the term small multiples for displays of multivariate data. The idea is equivalent to that of Cleveland's: a series of small graphics that show multiple combinations of several variables is very effective for comparisons and clearly display the relationships between variables (Tuft,2001). The psychological properties of the human visual system explain the usefulness of small multiples. Because each of the graphs in the display are quite small, it is easy to perceive the overview of a single graph without moving the point of focus; eye movement is only required when shifting the attention from one graph to another, which reduces the cognitive effort required in interpretation. Fortunately the interaction capabilities provided by computers enable the dynamic rearrangement of the graphic with little effort.

2.5 Interactive visual displays

One of the most important features of computer-based visual displays is that they allow the reader to interact with the business information and perform dynamic queries on the data. Especially in business intelligence systems there are vast amounts of summarized information available. This requires that users must be able to focus on the details on some part of the data that seems important and to dynamically explore the properties of individual pieces of that information. This operation is usually referred to as drill-down (Ware,2004). The famous visual information seeking mantra condenses the idea in a few simple words (Shneiderman,1996): *Overview first, zoom and filter, then details-on-demand*.

Another similar suggestion is that a statistical graphic should contain at least three viewing depths (Tuft,2001):

1. What is seen from a distance - overview of the data
2. What is seen up close and in detail - the fine structure of the data
3. What is seen implicitly, underlying the graphic - the "story" being told by the data.

The visual information seeking mantra is a result of business empirical research, whereas the proposition of the three viewing depths is purely intuitive, but they have a lot in common. In most visual problem-solving tasks the users indeed want to take these steps when they are exploring and analyzing the data set; this is called exploratory data analysis (Ware,2004). The large-scale overview facilitates comparisons between data points and the detailed views provide the lookup functionality that is often necessary for analyzing the data set (Ware,2004). These two together help to reveal the "story" of the data to the viewer.

There are a lot of methods to choose from when designing and implementing the interaction in visual displays. The traditional approach for moving from the

overview to the detailed level is changing the magnification of the display, i.e. zooming, which can be done by either selecting an area to focus on or sometimes simply clicking the mouse on the desired focus point in the graph (Spence, 2001).

In addition to these, Cleveland (Cleveland,1993) has introduced a technique called *brushing*, which means that the details of a data point are displayed when the mouse pointer is moved on top of that point. The details displayed may be just a single value of the data point, but sometimes the popup may include several attributes of the data point or even information about other data points related to the current point. Such extended popups are referred to as hover queries, since they actually present additional information about the data set (Ware,2004). Brushing and hover queries enable very fast exploration of the data set as the user can retrieve information by simply moving the mouse pointer over the display.

Another popular method of interaction in visualizations is the use of different distortion techniques. Distorting the graphic means that the scale of the data is intentionally varied in different parts of the display (Spence, 2001). A subset of the data is in focus and it has a smaller scale to enable displaying more details of the focused data, while other parts outside the focused subset have a larger scale thus showing the context of the focused data.

The *Table Lens* introduced by Card (Card et al,1999) is one of the many applications using distortion in visualization, and it is well suited for displaying multivariate relational information. It combines the traditional spreadsheet with graphical representations by showing a small graph in each cell. Figure 6 displays statistics for two players in a Table Lens. It has one row for each player and 17 variables for each player in the columns. In columns containing numerical values, each cell contains a horizontal bar, whereas in categorical columns (there are six of them in this example) the cell's graphical element is a colored and positioned small rectangle.

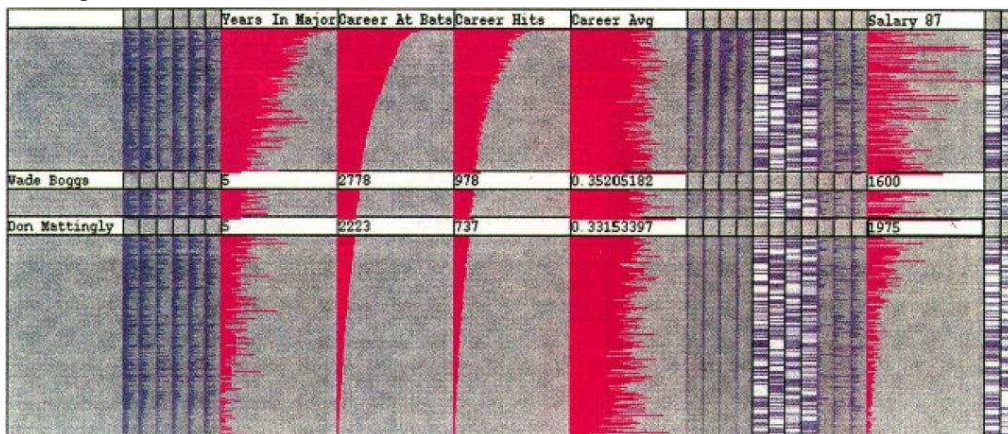


Figure 6 The Table Lens (Rao&Card,1999)

As we see, multiple rows and/or columns may be focused at the same time independently, and focused cells also display the value of the cell. Columns can be sorted by clicking the column title, which reveals correlations between variables, clearly visible. The Table Lens thus provides the overview of the entire data set and details about selected individual objects in a single display.

3. DASHBOARD - VISUAL INFORMATION INTERFACE FOR BUSINESS INTELLIGENCE AND PERFORMANCE MANAGEMENT TOOLS

The advantages of displaying numerical data in a graphical format have been recognized centuries ago, and even the first management information systems produced results in a graphical format. Today, most vendors of Business Intelligence and Performance Management tools now offer dashboards as the graphical user interface for business reporting, but none of them really explain what a dashboard is (Few,2006).

The origin of the term is in the French performance measurement framework *tableau de bord*, which literally means “dashboard”. The thinking behind the dashboard metaphor is probably that performance management is often regarded similar to driving a car, flying an aeroplane or steering a ship. Nonetheless, the term has been adopted by the end users to the extent that this tool is not likely to be known by any other name, even if its current name is quite misleading and non-descriptive. S. Few is the first who decided to create a definition applicable to all 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.*

This definition is very general in nature and carefully avoids any specific association with performance measurement or business intelligence; in fact, considering that after all a dashboard is a graphical user interface, it could be used to display any information that is necessary, meaningful and possible to represent in a concise visual format. However, the applications of dashboards are at the moment limited to business intelligence and performance management, since dashboards were originally introduced in this context.

The dashboard is only a “tip of the iceberg” behind it there is a massive infrastructure in which the business data are collected in their operational sources (databases, spreadsheets, text files or the web pages), moved into a data warehouse or mart through an extract-transform-load (ETL) process, and then processed in OLAP (On-line Analytical Processing) cubes that provide an integrated, multi-dimensional view on the data. Although the underlying structures play a significant role, we may say that the dashboard is the most important part of the information processing chain in the sense that it is the interface through which the human interacts with the computer system and the business data. The user only sees the

consolidated results on the dashboard, not the intermediate components of the business data processing chain. The dashboard is thus a tool for reporting, not a full-scale analysis tool (unless the user is allowed to build the dashboard from scratch, which is quite laborious), although some means of simple interaction with the data are usually featured. In order to achieve its goal, to provide information about the data that is turned into knowledge in the human mind, the dashboard must communicate the information as efficiently as possible and preferably in a manner that requires little cognitive effort. A visual presentation format makes it possible to fulfil these requirements, but not any given visual format. As we have discussed earlier, there are certain principles that the human brain follows when the visual image is interpreted; therefore some formats are more suitable than others, depending on the context. We will see here, how these principles can be taken into account in the design of a dashboard.

A business intelligence system is usually based on data warehouse technology (Watson&Wixon,2007). Collecting all data in a single repository offers an integrated view on all information regardless of its original source, and it also facilitates improving and controlling the quality of data, especially when some data are transferred from legacy systems that may have different encodings for the same information (Howson, 2007) .In large companies, the data may be further replicated to subunits as data marts, which are similar to the main warehouse, but each data mart contains only a subset of the warehouse data (based on e.g. geographic location or business function) that is relevant to that subunit (Turban et al, 2005); smaller companies might resort to independent data marts that are essentially “small warehouses”. Data in the warehouse or data marts can be utilized for business intelligence activities in many ways (see Figure 7). These include data mining, ad hoc querying, reporting and predictive analysis (Watson&Wixon,2007).

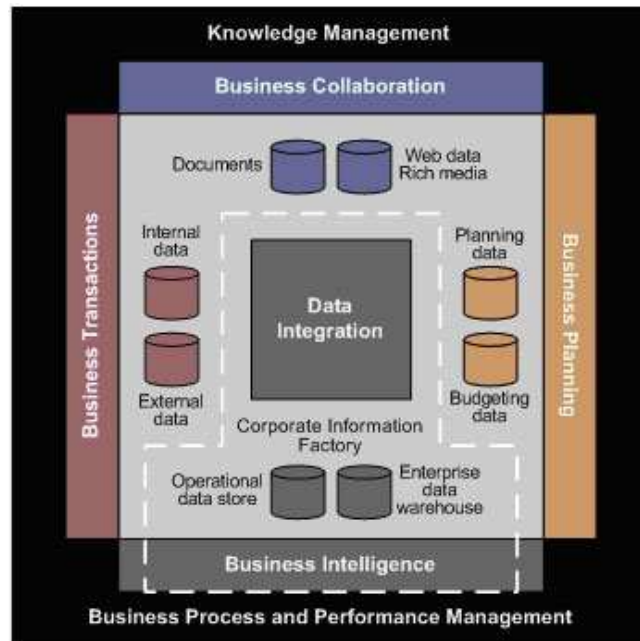


Figure 7 The Smart BI Framework (White, 2005)

As we see in Figure 7, key to collaboration and the sharing of information is knowledge management, which brings together portals, content management and collaboration tools. The growing importance of business intelligence also means that it too must be integrated into the knowledge management environment. Especially in reporting, visual displays of summarized data, i.e. dashboards, seem to be the most common format for user interfaces; some software products also include information visualization techniques for analysis functions. This may indicate either that the advantages of visual displays over text and tables have been implicitly recognized by software vendors and business users, or that they have become popular just because they make the interfaces more appealing. Whatever the reason, the trend is that - business intelligence methods and tools are highly visual in nature. At present, business intelligence is data-centric, but as it becomes more integrated with business operations it will need to become more process-centric so that business intelligence results can be more easily related to business processes and their associated business activities. This involves integrating performance management and process management technologies. In Europe, business intelligence is generally understood as gathering and processing both internal and external information (Pirttimäki,2007). Performance management and business intelligence are related to each other, but not synonyms. For example, the Balanced Scorecard includes measures for both external information (customer satisfaction, market share etc.) and internal information, and collecting data for these measures may be regarded as an intelligence activity. On the other hand, the business intelligence process itself can be subject to performance measurement, as suggested by Pirttimäki (Pirttimäki,2007). One differentiating

character between these two is that performance measures serve as a reporting function: the values of the measures are based on static predefined rules and formulas, whereas business intelligence is an analytical process that dynamically refines the raw information (by automatic data mining and modeling processes or through user interaction, such as ad hoc querying) to produce a deeper insight for decision-making (Pirttimäki,2007). Business intelligence can thus also be seen as the operation that provides the data for performance management. However, both functions are important and useful for managing a business. We therefore consider both performance management and business intelligence systems as the latest step in the continuum of decision support systems' evolution. In Figure 8 we present a concret visual interface to support an exploratory relationship between the users and the data with *Spotfire's DecisionSite* software, an analytic application environment based on user behavior. *DecisionSite* is designed to be able to import data from a variety of sources, then offer users several different possible representations of that data. Users can pose and alter queries and receive instant responses, enabling them to shape their questions accordingly and to manipulate the interface to better support their needs as they learn the data set.

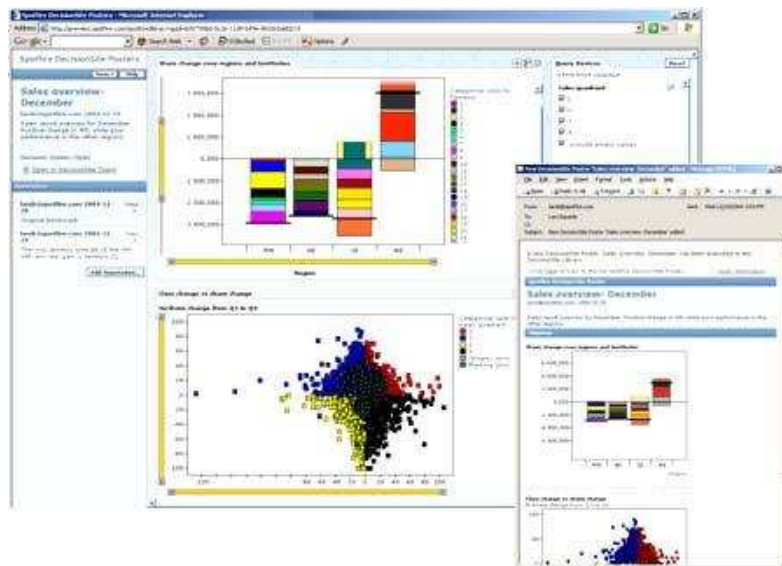


Figure 8 Business managers optimize investment projects based on performance (*DecisionSite*)

4. CONCLUSION

The business world needs information visualization. My paper in this field is important – too important to remain isolated in small enclaves. It's hard to imagine that any of us don't want our work to count for as much as possible in the management and research of business world. In my role as an Ph.D. students

supervisor, teacher, and writer, I work with a broad range of businesses, research groups on business intelligent systems, students, and universities. This puts me in touch with diverse communities that have one thing in common: they must make sense of and present business data to do their jobs, and most of them must do so in ways that require little statistical or technical sophistication. These aspects don't receive enough attention from the information visualization research community. Business researchers tend either to work on dissertations and articles or other research projects that are technically interesting to themselves as computer scientists or to focus on the needs of professions that are statistically and technically sophisticated. All of them must be interested by the using of information visualization tools and techniques.

REFERENCES

1. Beygelzimer, A., Perng, C.-S., Ma, S., (2002), Fast ordering of large categorical datasets for visualization, in *Intelligent Data Analysis*, 6,4 (2002), pp. 379-395.
2. Card, S.K., Mackinlay, J. D., Shneiderman, B., (1999), Information visualization, in *Readings in information visualization: using vision to think*, Card, S. K., Mackinlay, J. D. and Shneiderman, B., (Editors), Morgan Kaufmann Publishers, San Francisco, 1999, pp. 1-34.
3. Erbacher, R.F., *Exemplifying the inter-disciplinary nature of visualization research*, IV '07: Proceedings of the 11th International Conference on Information Visualization, IEEE Computer Society, 2007, pp. 623-630.
4. Few, S., *Information Dashboard Design: The Effective Visual Communication of Data*, 2006, O'Reilly Pub.
5. Friendly, M., Kwan, E., Effect ordering for data displays, in *Computational Statistics & Data Analysis*, 43,4 (2003), pp. 509-539.
6. Goldstein, E.B., *Sensation and perception*, Thomson Wadsworth, 7th edition, 2007.
7. Graham, D., *Information visualisation theory and practice*, in Proceedings of Ninth International Conference on Volume, Issue , 6-8 July, 2005, pp. 599 – 603.
8. Howson , C., *Successful Business Intelligence: Secrets to Making BI a Killer App*, 2007, McGraw-Hill.
9. Lurie, N.H., Mason, Ch.H., *Visual Representation: Implications for Decision Making*, in Journal of Marketing, American Marketing Association, Vol. 71, January, 2007, pp.160–177.
10. Palmer, S.E., *Vision science: Photons to phenomenology*, MIT Press, 1999.
11. Pirttimäki, V., *Business Intelligence as a Managerial Tool in Large Finnish Companies*, Ph.D. Thesis, Tampere University of Technology, Publication 646, 2007.
12. Shneiderman, B., *The eyes have it: a task by data type taxonomy for information visualizations*, Proceedings of the IEEE Symposium on Visual Languages, IEEE Computer Society, 1996, pp. 336-343.
13. Spence, R., *Information Visualization*, Addison-Wesley, 2001
14. Spillmann, L., The perception of movement and depth in moiré patterns. *Perception*, 22,3(1993), 2003, pp. 287-308.
15. So, S., Smith, M., Colour graphics and task complexity in multivariate decision-making, *Accounting, Auditing & Accountability Journal*, 15,4(2002), 2002, pp. 565-593.
16. Shneiderman, B., Plaisant, C., *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 4th Edition, Addison Wesley, 2004

17. Turban, E., Aronson, J. E., Liang, T.-P., *Decision Support Systems and Intelligent Systems*, Pearson Education, 7th edition, 2005
18. Tufte, E., *The Visual Display of Quantitative Information*, Graphics Press, 2nd edition, 2001
19. Tufte, E., *Beautiful Evidence*, Graphics Press, 2006
20. Ware, C., *Information visualization: Perception for design*, Morgan Kaufmann, 2nd edition, 2004
21. White, C., *The Smart Business Intelligence Framework*, in On-Line Business Intelligence Network, July 18, 2005, <http://www.b-eye-network.com/view/1182>
22. Watson, H. J., Wixon, B. H., The current state of business intelligence, *Computer*, 40,9(2007), pp. 96-99.