perceptual edge

Software Support for Immersive Business Intelligence

Stephen Few
June 13, 2006

Incorporating best practices of data analysis and communication into the software would result in better business intelligence.

The most frequent complaint that I receive about my books and classes is that they don't apply what I teach about data analysis and communication to a particular software product. Most often, the product that people want featured in my work is Microsoft Excel, but other products, such as those from Business Objects, Cognos, Hyperion, MicroStrategy, or SAS are sometimes the means of choice. This omission is intentional. I don't teach people how to use software. Instead, I teach people how to make sense of data and how to clearly communicate what they discover to others. I teach practical concepts—principles and practices—that can be applied to any software. It's not that I don't value software training; it simply doesn't interest me. I find classes that involve procedural instructions such as "Now, select the item labeled *Chart* from the *Insert* menu, and then pick *Bar Chart* from the list" rather boring, at least from a teacher's perspective. I'll leave it to the vendors to teach these classes.

Learning the mechanics of a particular software product is quite different from learning the concepts (principles and practices) and developing the skills needed by knowledge workers. Knowing how to use Excel does not make you a knowledge worker. Knowing how to analyze and communicate data does. Knowing how to use Excel simply might give you the means to perform the work more accurately and efficiently. Software tools cannot enable you to do what you don't already know how to do conceptually.

Somewhere along the timeline that extends from the advent of the PC in the early 1980s to the present, we started believing that knowing how to use software was equal to knowing how to do what the software was developed to help us do. This misunderstanding has resulted in a decrease in several of the basic skills that are needed in the business place, especially skills required for knowledge work. Knowing how to make a chart appear in Excel is not the same as knowing how to communicate graphically, and certainly is not the same as knowing how to analyze data.

What we face today is a bit like what would result if someone with no training in construction were sold a complete set of house-building tools, instructed where to hold each and turn it on, and then set loose to build his dream home. If the fellow survived, would the house be livable? Business intelligence products offer complex sets of tools that are no less complicated to those who haven't been trained than a cabinet full of drill bits and augers in every imaginable shape and size to someone like me with no construction experience. I find it easy to empathize with people who are given responsibility for making sense of data and producing reports to present data without ever having been trained in the principles and practices of data analysis and communication.

Recently, a thoughtful fellow responded to something that I wrote in my blog with the following question: "Are you suggesting that the BI vendors should teach us how to be data analysts?" To

say that I believe BI vendors *should* teach us the concepts and skills of data analysis would overstate my position. I don't believe that vendors bear this *responsibility*, but I do think that they waste their time when they teach people how to use BI software if these people haven't already learned the concepts and skills of data analysis and communication. By addressing this need, the vendors would better equip their customers to make good use of their software, resulting in happier, more successful customers, increased sales, and fewer technical support calls. Everyone would win. Greater focus on this need would place vendors in a better position to legitimately claim that they are "solution providers."

Good business intelligence software supports and augments our thinking, helping us to better understand what data means. It doesn't do the thinking for us, however. Software should provide guidance that helps people focus, not on the mechanics of using the tools, but on the meaning of the data and the questions they should ask to investigate the data. Software that supports the analysis and communication of data could support us in the following uncommon ways:

- By helping us to select the appropriate medium of display, and doing so in the context of the data and our objectives rather than the mechanics of the process.
- By identifying and allowing us to highlight aspects of the data that are meaningful in light of our analytical intentions, thereby making it easy to spot them as specific visual patterns and to understand what they mean.

I believe that these steps in the data analysis process should be supported in a manner that keeps the user focused on the meaning of the data rather than the mechanics of the task.

Become immersed in the meaning of the data,..	...not in the mechanics of the software
How are sales changing through time?	How do I switch this to a line graph?
What are the relationships of each department's expenses to the whole and to one another?	How do I get the largest slice to begin at the top of the pie chart?
Is there a correlation between newspaper ads and sales revenues?	How can I get the newspaper ads to appear on the vertical axis of the scatter plot?
Let's consider only Web traffic during this week.	How do I construct a query to select this week?
What product-level factors are contributing to the number of returns for this product family?	Do I right click or double-left-click to drill down to the product level?
Have salaries increased equally overall or is the increase limited to a small group of people?	What do I need to do to see the high, median, and low salaries for each year at once?
Which is increasing at a faster rate, international or domestic sales?	What can I do to compare the rate of increase between high-volume and low-volume sales?

Figure 1: Good BI software should help us stay focused on the meaning of the data, not the mechanics of the software.

Selecting the Best Medium of Display

One of the primary challenges of data analysis and communication involves selecting the best medium of display. If your task involves looking up and comparing individual values, a table would do the job nicely. If your task involves searching for and interpreting patterns, trends, or exceptions in the data, a properly designed graph would support discoveries that might never be found in a table of the same data.

Choosing a graph rather than a table, or vice versa, is just the first step. If a graph is required, what kind of graph? Do the patterns you're searching for reveal themselves better when visualized as lines, bars, individual data points, or perhaps some other display object? If a bar graph would do the job best, should you use vertical or horizontal bars? If you are investigating sales revenues compared to the sales plan by region, should the distinction between actual and planned revenues be labeled along the axis or should the regions be labeled along the axis? If regions are labeled along the axis, how should you visually distinguish actual from planned sales? These are typical questions that must be answered to select the best medium of display.

Business intelligence software should make these choices easy and efficient. Best practices for visual display, which have emerged from extensive research, could be applied to every one of these choices. Why shouldn't the software provide expert guidance when it's needed, based on these rules? Why shouldn't the software discourage bad choices, such as by eliminating those types of graphs that are not appropriate, given the data and your objectives? For instance, when you wish to examine sales revenue per region, which is a categorical variable that consists of discrete items (for example, north, east, south, and west), it wouldn't make sense to connect the values across those regions with a line, so why not eliminate this choice?

And why not provide this kind of guidance in a way that supports your thinking process in conceptual terms, rather than in terms that reduce the process to a series of mechanical choices? For instance, rather than presenting a list of chart types or pictures of charts from which to choose, why not ask you to identify your objectives—what you want to communicate or the conditions you're interested in detecting in the data? When you present data in a graph, the essence of communication isn't that a line is going up or down or that one bar is longer or shorter than another; it is that sales are rising or falling or that expenses exceeded or came in under budget for the quarter. By assisting you in these terms, you can remain immersed in the meaning of the data rather than the software-specific mechanics of the process.

Think of what I'm suggesting as an attempt to achieve a similar effect to what has been achieved by modeling data dimensionally for reporting and analysis, rather than in complex relational database structures. What Ralph Kimball and others strove to achieve, among other things, through the paradigm of measures and dimensions, was the expression of data in terms that were closely aligned with the way we understand and use the data in the real world. Interacting with dimensionally structured data, or with data that is presented through a semantic layer that is dimensionally organized, allows us to remain engaged with the data in a meaningful way, with less distraction from the technical implementation or an artificial mechanical process.

Highlighting Meaningful Patterns

When you analyze data in an attempt to understand it, you search for, examine, and interpret meaningful conditions in the data. For instance, when analyzing sales behavior through time, you might want to see if any behaviors exist that are cyclical in nature; behaviors that appear periodically and predictably at particular intervals. When data is visualized in particular ways (for

example, as a line graph), these patterns of behavior exhibit themselves as particular visual patterns that can be searched for, recognized, and understood. Wouldn't it be helpful if your data analysis software was aware of those visual patterns that are useful, given the nature of the data and the objectives of your analysis, so that it could suggest meaningful conditions to search for and make them easy to find and examine if they exist? You bet it would.

Furthermore, wouldn't it be nice if the software allowed you to indicate your objectives in conceptual terms, to ask your questions more directly, interacting with you in ways that keep you immersed in the meaning of your analysis, rather than the mechanics of the software? Imagine using software that allowed you to ask the question "Do customers of a particular type dominate sales in the west region?" without having to translate that question into a series of mechanical steps, because the software has the intelligence to do that for you. This kind of support would allow you to forget about the software almost entirely as you become lost in thoughts about the data, fluidly moving between thought, interaction, and observation, without interruption. Software of this type would support the emergence of intelligence about the business (that is, true business intelligence), requiring you to be an expert in the business, rather than in the mechanics of a particular software product.

What I'm proposing is not science fiction. If people like me, who specialize in the use of data visualization for analysis and communication, can codify the rules for translating meaningful conditions in data into appropriate visual media and associated patterns, why can't these heuristics be built into the software? Why force users to learn and then spend most of their time engaged in the mechanics of the process, something a computer is capable of handling, rather than fully engaged in thinking and communicating, which is something that only human beings can do?

Stop Making Excuses

Don't mistake my opinion that software products should do more to support data analysis and communication for permission to let the vendors do all the work or to blame the vendors when you don't know what you're doing. Ultimately, if it's your job to make sense of data and communicate your findings to others, the necessary skills are your responsibility. Just as it was your responsibility to learn the material in the classes you attended in school, no matter how untalented and undedicated the teacher, it is your responsibility to make the best possible use of software, no matter how poorly it supports the process.

If the tools that you use aren't up to the task and you have a choice in the matter, find better tools—they're out there (perhaps not the ideal that I've described, but probably much better than what you have). If the choice of tools is out of your hands, then raise your voice in protest to put pressure on your vendor to improve. Don't waste your vendor's time by demanding silly features that add no real value (such as most 3-D graphs). As long as customers continue doing this, vendors can rely on the excuse that if the software doesn't work well, it's not their fault, because they were only being faithful to customer requirements, no matter how absurd. Take the time to learn what works and then demand that your software supports it.

As a business intelligence professional, you must be an expert in your business and possess skills in the analysis and communication of information. Expertise in a particular software product is secondary, and without the pertinent conceptual skills, it is useless. It could be even worse than useless, because it might be mistaken for genuine expertise that doesn't exist, leading to bad decisions. If you provide the intelligence and your software does its part to support and augment your intelligence, worthwhile business intelligence will emerge. Make it so.

About the Author

Stephen Few has worked for over 20 years as an IT innovator, consultant, and teacher. Today, as Principal of the consultancy Perceptual Edge, Stephen focuses on data visualization for analyzing and communicating quantitative business information. He provides training and consulting services, writes the monthly *Visual Business Intelligence Newsletter*, speaks frequently at conferences, and teaches in the MBA program at the University of California, Berkeley. He is the author of two books: *Show Me the Numbers: Designing Tables and Graphs to Enlighten* and *Information Dashboard Design: The Effective Visual Communication of Data*. You can learn more about Stephen's work and access an entire library of articles at www.perceptualedge.com. Between articles, you can read Stephen's thoughts on the industry in his blog.