Data Visualization, Data Interpreters, and Storytelling

Hugh J. Watson

Hugh J. Watson is a professor of MIS and holds a C. Herman and Mary Virginia Terry Chair of Business Administration in the Terry College of Business at the University of Georgia. He is a TDWI Fellow and senior editor of the *Business Intelligence Journal*.

hwatson@uga.edu

Data visualization (or visual data discovery) is one of the hottest topics in BI and is being used to understand customer purchasing decisions, determine behavioral patterns that suggest fraud, and identify supply chain inefficiencies. Increasingly, managers and professionals are expected to make their arguments and decisions based on data and visualizations.

The science behind data visualization shows that humans can readily distinguish differences in line length, shape orientation, size, placement on page, and color with minimum cognitive effort. Software products such as Tableau and Qlik can be used to take advantage of this through a multitude of different visualizations and an amazing palette of colors. Especially in this age of big data and analytics, data visualization has the potential for making sense out of, and finding relationships in, huge data sets.

However, I have mixed feelings about some of the newer visualizations—such as treemaps and starbursts—and when and how they are used. Many of them result in screens that "pop" but are difficult to interpret—especially by senior executives. Are we creating visualizations that are too difficult for users to understand? If so, what can we do about it?

In her keynote address at the 2015 TDWI Conference in Orlando, Claudia Imhoff talked about the need for *data interpreters* to explain the key takeaways from analytics-based visualizations. Her insight resonated with me because of my experiences with visualizations for senior executives. As we talked later, Claudia and

I discovered that we had many similar thoughts and experiences, and we decided to collaborate and share them with *Journal* readers. Hopefully, you will find them helpful in your company's data visualization efforts.

VISUALIZATION: INFORMATION VS. DATA

A useful distinction can be made between *information visualization* and *data visualization*. Information visualization focuses on presenting information about relationships that are already understood, such as the graphical displays in dashboards. By contrast, data visualization displays data about *newly discovered* relationships, such as a previously unknown customer segment or a sequence of events that can predict customer churn. Therefore, data visualization is more often associated with the discovery or exploratory analytics of big data.

TWO STORIES

In our conversations, Claudia and I shared experiences that illustrated the need for data professionals to be careful with their data and information visualizations.

The Retinal Scan (a data visualization example). Claudia tells the story of meeting with a senior executive. After rapport was established, the executive pulled a graphic from his desk and asked if she could help him interpret it. He said it "looked like a retinal scan" to him and that he didn't have a clue as to what it was trying to convey.

The MTV Screen Designers (an information visualization example). Several years ago, I was helping a company develop a system for senior executives. An important consideration was ease of use—the executives had to be able to easily find, view, and understand the information.

Two relatively young contractors were hired to build the screens using software that had been acquired for the project. I found myself having to constantly remind them to use consistent colors and formats across the screens rather than the more colorful, exotic screens they initially developed.

I could tell that the "MTV screen designers" (as I jokingly called them) did not always appreciate my insistence on certain colors (i.e., blues and grays) and graphical formats (i.e., bar and line charts). Nonetheless, I was vindicated at the initial rollout of the system when one of the executives said, "I really like the executive feel of the system."

THE TAKEAWAYS

The lesson from these stories is that familiar visualizations may be easier to interpret—especially by senior executives and line-of-business (LOB) managers—and that there is value in adhering to carefully established screen design standards. Keep your audience in mind and don't become so enamored with the visualization that you forget the targeted users' capabilities or the business problem.

Stephen Few (2006), a highly regarded data visualization guru, makes specific recommendations for dashboard design, including:

- Don't use meaningless variety
- Don't misuse or overuse color
- Don't use useless decoration

Though newer kinds of visualizations should be used with caution, they may command rapid adoption and be easily understood in the near future. For example, sparklines and treemaps

were uncommon several years ago but are now included in Excel.

Even so, some of the fanciest visualizations are just vendor "eye candy." They are attractive but lack any real value beyond their looks—and may even lessen a viewer's understanding of the information displayed (3-D pie charts fall in this category). On a more positive note, many vendors now offer products that automatically present data using appropriate visualizations.

> Some of the fanciest visualizations are just vendor "eye candy." They are attractive but lack any real value beyond their looks.

WHAT THE RESEARCH TELLS US

Everyone is familiar with the saying "a picture is worth a thousand words." A more contemporary, business-oriented version is "a picture is worth a thousand spreadsheets." You have probably participated in an exercise that illustrates this. Data is presented in a table and you are asked to identify the relationships between the variables. After struggling for a while, you are then shown graphs that make the relationships readily apparent.

Data visualization is a heavily researched topic (Tufte, 1983; Few, 2006) and the consensus is

that graphical representations of quantitative data boost understanding. In a recent study, Wakeling et al. (2015) considered different visualizations in terms of accuracy, speed, and confidence, which are the commonly used measures to assess graph literacy. Some of the most salient of their findings are:

- The accuracy of interpretation of tabular data (remember, tables are a type of visualization) is surprisingly low.

- More familiar, simple charts (e.g., bar, line, pie) are more accurately interpreted than less familiar advanced ones (e.g., Sankey diagram, stacked bar, bubble).

- Simple, familiar charts are interpreted most quickly.

- User confidence is highly correlated with how frequently charts are encountered; the more people use a particular kind of visualization, the more confident they are interpreting it.

- User confidence does not always correlate with the accuracy of interpretation. Just because someone is confident with a type of visualization does not guarantee that the data will be interpreted correctly.

Although these findings generally support the use of common simple visualizations, they do not mean that newer, more sophisticated visualizations aren't useful. For example, simple visualizations display only two or three dimensions and their associated values, whereas

advanced visualizations, such as heat maps, can show many more dimensions on the same chart. You trade simplicity for the ability to convey more information. Users, however, are slower, less confident, and more likely to make errors with visualizations that are more advanced and less familiar.

THE DATA INTERPRETER

Addressing the user learning curve while breaking away from basic visualizations requires the introduction of a *data interpreter*. This person explains visualizations to executives, LOB managers, and other users who need help interpreting them. The role may be performed by someone whose job it is or possibly as an additional task for administrative assistants, BI or business analysts, or data scientists.

> **People are using storytelling concepts to clearly communicate a message using carefully created visualizations.**

Regardless of who performs this role, the data interpreter must understand the audience, the business domain, the problem the visualization is meant to address, and the visualization itself, in addition to having storytelling ability and excellent communication skills. Data interpreters can become trusted advisers to the C-suite.

Onetime visualizations need to be handled differently than recurring ones, such as those in dashboards. With the former, the data interpreter explains the visualization and if problems in understanding persist, he or she can create a new more easily understood visualization. With recurring visualizations, though, the focus is more on training users, who will understand more complex visualizations when they are used repeatedly in their jobs.

The development methodology used for creating applications is also a factor. When an agile method is employed, the user sees the visualization early on and can influence the design and learn what information the visualization is meant to display.

For example, in a recent issue of the *Journal*, we looked at how Mark Jackson of Piedmont Healthcare uses Tableau to produce dashboards. He will often discuss the information needs, build a dashboard, and make sure the user understands the information displayed and how to manipulate it—all in a single meeting (Watson and Jackson, 2016).

STORYTELLING

We are all familiar with "death by Power-Point"—presentations that are boring and ineffectual. These presentations often include too much data, fail to communicate a clear message, and don't lead to action.

In response to the failings of traditional presentations, people are using storytelling concepts to clearly communicate a message using carefully created visualizations. The value of this approach is based on the premise that stories:

- Help people get the gist of an idea quickly

- Help them remember facts much better

- Are memorable and enjoyable

Storytelling can be used in many different settings with a variety of audiences, including a self-running presentation; a live presentation to a large audience; or a live presentation to an individual or small group, which allows interaction with the listener or listeners (Kosara and Mackinlay, 2013). Stories can be used to *explain* or to help the audience *explore* a topic. Each of these scenarios affects the techniques used, the way the presentation is structured, and the amount of interaction anticipated.

The following is a set of storytelling best practices.

Focus on the story. The focus of storytelling is to capture listeners' attention with a narrative that is supported by visualizations. Stories have a beginning, middle, and end. Although the details will vary, a common story flow is to describe the current situation, provide predictions should it continue, lay out the options, make a recommendation, and describe the likely outcome from implementing the recommendation.

Know your audience. Your listeners will ask themselves what the story means to them. Put yourself in their place when creating the story. Include specific, personalized examples that people will remember. For example, politicians often name and describe an individual to illustrate a point, such as someone who is benefiting from Obamacare.

Establish the setting. This makes the story relevant. Describe the situation in a way that is meaningful to the listener. Identify the business issues and challenges. Support this description with data and let listeners know its source, accuracy, and relevance. You want people to be confident that you understand the business and issues. If possible, take a unique and meaningful perspective.

Define the characters. Let the audience know who and what is being affected by the current situation. Is it the company's bottom line? Perhaps it is the organization's management, customers, or employees. This is where a personal experience can be effectively used to put a face on the people being impacted.

Define the problem or conflict. Visualizations are critical here because they can identify a problem or issue or prove a point. Be strategic with your visualizations and keep them simple, uncluttered, and tasteful. Highlight the critical parts. The more interesting the visualization, the more time and attention your audience will give it.

> Presentations should do more than identify a problem. They should suggest a resolution or at least lead to a discussion.

Show the resolution and future. Presentations should do more than identify a problem. They should suggest a resolution or at least lead to a discussion. Use data to make predictions, lay out alternatives, identify a solution, create a call to action, and show likely outcomes. However, do all of this in a visual, artistic, and thoughtful way.

An excellent storytelling example is the *Boston Globe*'s presentation of the competition between Starbucks and Dunkin' Donuts (Chang and Carroll, 2013). The presentation visually describes the distribution of each company's franchises across the country and provides insights about why the two powerhouse chains are popular in specific geographical areas (e.g., Starbucks started in Seattle and has an especially strong presence in the Northwest and affluent areas such as Manhattan).

CONCLUSION

Data visualization is growing in popularity, but be cautious when using newer, less familiar visualizations—especially with senior executives—as these can be very difficult to interpret and understand. Introduce the data interpreter role to ensure that visualizations are understood. Introduce storytelling best practices to increase interest and comprehension of presentations. ●

REFERENCES

Chang, A., and M. Carroll [2013]. "Split Country: Dunkin' vs. Starbucks," Boston.com. http://www.boston.com/ yourtown/specials/starbucks_vs_dunkin_ donuts/

Few, S. [2006]. *Information Dashboard Design: The Effective Visual Communication of Data*, O'Reilly Media.

Kosara, R., and J. Mackinlay [2013]. "Storytelling: The Next Step for Visualization," available at http://kosara.net/ papers/2013/Kosara_Computer_2013.pdf

Tufte, E. R. [1983]. *The Visual Display of Quantitative Information*, Graphics Press.

Wakeling, S., P. Clough, J. Wyper, and A. Balmain [2015]. "Graph Literacy and Business Intelligence: Investigating User Understanding of Dashboard Data Visualizations," *Business Intelligence Journal*, Vol. 20, No. 4, pp. 8–19.

Watson, H. J., and M. Jackson [2016]. "Piedmont Healthcare: Using Dashboards to Deliver Information," *Business Intelligence Journal*, Vol. 21, No. 3, pp. 5–9.