ChangeThis THE PRESENTATION OF DATA GRAPHICS IN EVERYDAY LIFE Kristen Sosulski, Ed.D

In 1964, Marshall McLuhan published Understanding Media: Extensions of Man and coined the phrase, "The medium is the message."

He argued that the medium itself, not the content that it carries, should be the focus of study. That is because characteristics of the medium, not just the content it carries, influence our behaviors, responses, and actions.

Consider how our smartphones, social media platforms, or even search engines alter the pace, pattern, and scale of our daily activities. Think about your own process for selecting restaurants, documenting moments, and sharing your thoughts with others and the role that your media plays in these activities.

Next, consider pictorial representations of data, or data graphics as a medium. How have data graphics proliferated their presence in our everyday life? Navigation systems guide us from one location to another through narrated directions accompanied by digital maps, annotated with points of interests like the nearest coffee shops, shoe stores, or art museums as alternative pathways to peak our interest and reroute us from our original destination. Aside from

navigation aids, data graphics accompany news reports, research findings, and advertisements as both supporting evidence and alternative narratives. We have come to expect statistics and numerical data to be organized for us visually. Rightly or wrongly, the juxtaposition of text and data graphics implies that the content is data-driven and supported by legitimate evidence.

Not surprisingly, data graphics have become an incredibly fashionable medium. So much so that it is the de facto interface to navigate the endless deluge of streaming data we deal with every day. From smart cities to smart phones, data graphics facilitate both the exploration of data and the presentation of it. When used with integrity and purpose, data graphics have the potential to help us, as humans, make sense of this intensely data focused world. Yet, the current state of data graphics is that they are often overused and underwhelming.

Data and information visualization have provided humans a lens upon which to view information in aggregate. William Playfair (1821), John Snow (1854), and Charles Joseph Minard (1861) are credited with introducing this medium to our modern vernacular. Later, John Tukey, Edward Tufte, and Stephen Few made significant contributions to the research and practice. Their contemporaries have evolved it with massive and largely unstructured data. Among others, Fernanda Bertini Viégas and Martin Wattenberg led the way for popularizing the practice of information visualization in the new millennium.

Characteristics of the medium, not just the content it carries, influence our behaviors, responses, and actions.

In research for my book, *Data Visualization Made Simple*, I discuss how data graphics accelerate the pace by which data can be understood and used. It enables the aggregation and summation of massive data into digestible graphic presentations. These presentations take the form of static, animated, and interactive charts and graphs. At the foundational level, data graphics reveal trends, patterns, results, and even answers using encodings such as lines, bars, shapes, points, and color—revelations of insights or findings in the form of data graphics. Those insights inform our understanding of phenomena and the actions we may take with new information.

For example, temporal trends such as stock price are easily observed as lines that are increasing, decreasing, stable, or cyclic. Categories, such as the number of apartment rentals by neighborhood, are easily compared through bars or areas. The best and worst performers in a tennis tournament are shown by the length of the bar, the size of the area over time, or for a single point in time. Ratios and densities of geographic locations show population, for instance, and reveal their proportions through shaded maps. The darker areas indicate higher populations, while lighter areas show lower densities. Distances between locations are encoded by the length of line connecting two or more entities. Outliers are easily spotted in scatter plots, and linear relationships are apparent through trend lines. The shape of the data is understood through density plots and histograms. The range and median of asking prices for single family homes in a given city are easily seen through a box plot. Moreover, even non-numeric data can be understood at a glance as word clouds or text overlays that show sentiment, themes, popularity, and frequencies. Text data is encoded by size or color to show similarities and differences of customers' critiques of their dining experiences for a particular restaurant. Relationships amongst entities are highlighted through directed and undirected networks of nodes and edges. Line thickness portrays the strength of the relationship and lines show the connectivity, such as people in a social network. These are some of the nascent types of data graphics that, while ordinary, serve a role in rapid information presentation and interpretation. As an illustration, think about your own fitness and nutritional health. There are key daily metrics that you surely want to know such as your weight, number of calories consumed and burned, and exercise performance. If you are a runner, for instance, you may want to know the distance you ran today, your average pace, and fastest mile. These individual metrics represent a snapshot of your health and nutritional performance at a single point in time. What if you wanted to see how it has changed since last Monday or last year? This is where you would no doubt expect to see a line chart or spark line showing a trend of your performance to date. The lulls and sprints are made apparent by the dips and spikes of the line.

The shape, direction, size, and orientation of graphic encodings form the visual language used to construct data graphics. As humans, we are especially attuned to perceiving these encodings as the core building blocks of our visual system. This makes it easy for us to perceive many data points at once.

Data graphics accelerate the pace by which data can be understood and used. It enables the aggregation and summation of massive data into digestible graphic presentations.

The Process of Visualizing Data With Purpose

The designing of a purposeful and useful data graphic is contingent on both the data and the audience. The data source must be of high quality, and the insights derived reliable and valid. With the audience in mind, the design of the data graphic in then focused on optimizing data insights and maximizing interpretation of the information.

Broadly speaking, the steps are:

Step 1. The data graphics creative process begins with a question. For example: What is the median household income for the zip code 11201?

Step 2. Identify the appropriate data sources and query it for an answer to your question.

The data needed to answer this question is trivial to identify; it is available from the U.S. Census. The answer is a single number: \$109,472. This is the middle household income value for the sample of households in 11201.

Yet, while this answers the question, it does not make for an interesting or particularly insightful visualization. This is because the question itself is too narrow. Usually, if the initial question begins with "what," the results will be a single number. Asking "how" or "why" questions tend to produce more descriptive and detailed findings. Refine the question as needed. A more interesting line of questioning may involve the distribution of household income at either a more birds-eye level, or at a more granular level. For example, "How does the median household income of 11201 compare to all the other neighborhoods in New York City?" This contextualizes

the answer to the first question and presents the single value, for 11201 as "\$109,472" in comparison to the other neighborhoods in New York City.

Ensure the question provides a sufficient answer by analyzing the necessary data. This structured exploratory process is iterative and commonly begins with a question to guide data querying and analysis.

Step 3. Once a satisfactory answer is reached, the answer requires interpretation. This involves an evaluative statement that helps the audience understand and see the results presented in the graphic. Such as: The median income of 11201 is one of the highest in all of Brooklyn at \$109,472, but well behind many of the neighborhoods in Manhattan, which exceed \$250,000.

This is the interpretation of the data values based on findings.

Step 4. Present the answer using a data graphic as evidence. Design the data graphic for your audience, select the appropriate chart type, and highlight the most important information.

In this example, a simple choropleth map of median income by zip code puts the single neighborhood of 11201 in context. A choropleth is a thematic map where the filled regions are shaded in proportion to the measurement of the statistical variable, which is the median household income. A static choropleth would present a snapshot, while an animated graphic can show the change in the median household income over decades. Furthermore, adding interactively involves the audience exploring the graphic through filtering by zip code, year, or borough, allowing them to experience a human reality (the history of how wealth and poverty has concentrated in various New York neighborhoods over the years) through a graphic representation. These decisions require understanding how your audience will engage with the new information. Will it be shown to them in person to view alongside the presenter, online to interact with, or printed together with an article for them to review at their own pace? **Step 5. Refine the graphic.** The goal is to maximize the audience's retention and minimize overload. First, the visual information should be easily perceived. Obviously, it is naive to assume your audience will just see the key insight. Don't make your audience work too hard. Are there additional data points you'd like to highlight or call out? A historical context where were on this day last year compared to today, this year. Is there a way to lead the viewer through your visualization through animation, call outs, etc.? Colin Ware , in *Information Visualization: Perception for Design* (2013) defines four pre-attentive visual properties. These include color, form, movement, and spatial position. Second, reinforce the key insight as the takeaway. Generally, there are three types of explanations that can accompany a chart based on the presentation type:

Oral presentation. Use spoken words to describe what the audience is seeing.

Oral static narration. For video-based or recorded presentations, oral narration guides the viewer in a way similar to a live presentation, but without direct engagement with the audience.

Written explanations. Written communication for chart titles and descriptions. Essentially, anything you want the audience to read.

Words help reinforce a visual message. For live presentations, the words spoken should differ from those written to avoid information overload and redundancy. In addition, the data graphics selected should be easy for the audience to interpret. Ask yourself if your audience has seen this chart type before? Is the display type you are showing in line with other charts or graphs they have seen in the past? For example, a map clearly shows geography. The encodings, such as fills or shading, are decoded by the user through a legend, along with verbal or written explanations. Researchers continue to evaluate how the design of data graphics can interfere and hinder the desired result from your audience. This involves the chart format, color usage, appropriate text and labels, readability, scales or units of analysis, integrity of the data, non-data elements such as chartjunk, how much data is presented (data density), the data richness (or context), and source identification and attribution.

Step 6. Seek feedback from others. Collect empirical evidence to know the degree to which your data graphic conveys your intent.

Conclusion

There is a huge opportunity ahead of us. The knowledge economy of the future is a world in which data is easily accessible, integrated in systems that make for easy querying. These structures create the possibilities for advancement of data graphics to aid in human decision-making and analysis. The advancement of virtual and augmented reality systems also promises transformation of how we view the world and ourselves.

The gestalt of the visualization process is the encoding of data into symbols, the data we choose to show and omit, how we show it (the graphic), when we show it (presentation, timing), where we show it (internet, newspaper, slide presentation), and ultimately how those data graphics impact our audience and their decisions, behaviors, and understandings. **Ultimately, the way information is consumed by the audience, and their impact on our understanding of the world, is on the shoulders of the creator and the medium they choose to convey it.**

Info



Ready to dig deeper into this idea? Buy a copy of <u>Data Visualization Made Simple</u>.

Want copies for your organization or for an event? We can help: customerservice@800ceoread.com 800-236-7323

About the author

As a leading expert on data visualization, Kristen Sosulski regularly consults, delivers seminars, and leads workshops on data visualization techniques and best practices. You can find her speaking on the subject at events like Social Media Week NYC, plot.ly's PlotCon conference, and Tableau's events and to organizations like the National Association of Public Opinion and the National Economic Research Association, Kristen Sosulski is an Associate Professor of Information, Operations and Management Sciences at New York University's Stern School of Business and the Director for the NYU Stern Learning Science Lab where she leads team to design immersive learning environments for professional business school education. She teaches MBA and executive courses in data visualization. computer programming, and operations management. Sosulski's research and practice focus is on technological progress in education and visualization. Her specific area of expertise is learning technologies and data visualization.

Share this

Pass along a copy of this manifesto to others.

Subscribe

Sign up for e-news to learn when our latest manifestos are available.



Powered by the love and tender care of 800-CEO-READ, ChangeThis is a

vehicle for big ideas to spread. Keep up with the latest developments in business books and ideas at 800ceoread.com

This document was created on September 19, 2018 and is based on the best information available at that time.

The copyright of this work belongs to the author, who is solely responsible for the content. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License. To view a copy of this license, visit Creative Commons. Cover image from Adobe Stock.