Encyclopedia of Aesthetics, Michael Kelly, Editor-in-Chief
(New York: Oxford University Press, 2014, second edition)

Data Visualization.

Aesthetics, broadly defined as the sensuous perception of material, plays a central role in the emerging practice of data visualization, which aims to render abstract information intelligible to human perception through the use of images and other forms. Contemporary data visualization typically concerns three interrelated areas: data collection, analytics, and visual representation. Data, which refer to the values corresponding to specific qualitative or quantitative measurements of physical or virtual phenomena, are the raw material from which information is generated. Analytics refer to processes or algorithms that operate on a set of raw data to synthesize higher-level information. Data analytics often rely on data models with which hypotheses are tested or decisions or predictions are made. Although data visualizations sometimes map data directly, they often use analytics as an interpretive layer to produce an image. While visualization typically uses images to render concepts or models comprehensible, other techniques, such as sonification and tactile representation, have been explored as well. This essay focuses on the visual as it is the most common form and has played a particularly central role in shaping a new web aesthetic, changing the look and feel of online content. The number of potential visualizations corresponding to a given set of data is limitless, and aesthetic judgments are inevitably involved in the process of creating intelligible and compelling visualizations.

Historical Roots of Data Visualization.

The history of data visualization can be divided broadly into analog and digital periods, the latter beginning with the appearance of programmable computers in the second half of the twentieth century. Analog methods applied to the visual display of information can be found in such disparate disciplines as astronomy, cartography, and statistical analysis. The charts and graphs that comprise the early visualizations leveraged the aesthetic qualities of visual images and the capacity of human vision to generate an intuitive comprehension of underlying data that might otherwise remain abstract.

Visualization generally seeks to reconcile the accurate and faithful representation of data with aesthetically compelling images. Even where objective representation of data is a goal, aesthetic decisions require subjective human influence. For instance, cartographers developed methodologies and symbols to organize data collected from a geographical territory into the spatially intelligible format of the map. Although mapping may seem to be an objective exercise in translating data onto the spatial coordinates, even the earliest cartographers found it necessary to use aesthetic methods to establish hierarchies of that same information. These methods enabled them to prioritize among geographical characteristics, roads, political boundaries, and the like to suit their particular purpose. Indeed, the early maps are aesthetically arresting due in part to the obvious subjectivity legible in their composition.

Innovations during the seventeenth century, including the establishment of the Cartesian coordinate system and analytic geometry, laid the groundwork for the visual representation of data in a range of disciplines outside of cartography. During the nineteenth century, visualization techniques were used in the application of diverse disciplines from statistics to commerce and planning, and by the twentieth century such techniques became widespread in areas such as academic research and military science, incorporating dynamic as well as static representations. Examples such as sonar, which plots acoustical information as “pings” on a field, enabling human vision to distinguish persistent patterns of an approaching enemy vessel from the acoustical disturbance it produced, hint at the potential that real-time data processing would assume in the digital age.

Analog to Digital.

A major shift in visualization occurred with the advent of programmable computers in the second half of the twentieth century, which enabled digital data collection and storage and greatly enhanced the capacity for statistical analysis. The shift from analog to digital data collection and display enabled designers and data scientists to explore more interactive and dynamic ways to display information. The 1970s and 1980s saw an explosion of experimentation with interactive and dynamic graphics through the development of novel computer algorithms. One early example is E. B. Fowlkes’s work at AT&T Bell Laboratories in 1971, which sought to develop an interactive system to explore probability plotting. Drawing in part on Fowlkes’s early work, in the 1980s and early 1990s graphic designer Muriel Cooper, director of the Visible Language Workshop at the Media Laboratory at the Massachusetts Institute of Technology, explored the aesthetic potential of dynamic and interactive visualizations based on stock market data. Cooper and her team conceived of datasets as forming points in an interactive three-dimensional world. Users could fly through the data using the mouse and keyboard, viewing different aspects of the data mapped to specific spatial dimensions in a continuous, yet finite, data world. Early computational explorations of data visualization like those elaborated above brought about new methods to experience data. Although these early computer-generated visualizations were in some cases dynamic and interactive, the datasets were small by contemporary standards. Their static nature led to approaches in which the entire data space could be theoretically viewed with sufficient time to navigate through the interface.

Web-Based Data Visualization.

The growth of massive data during the early 2000s, coupled with the popularity of interactive graphics, brought about a shift toward techniques capable of dealing with datasets that were functionally limitless. These developments fueled a shift from experiments in screen-based three-dimensional dynamic data worlds to an emphasis on enabling large numbers of distributed users to simultaneously manipulate and interact with multiple large datasets. Data visualization expanded from simple visual representations of abstract information to include interactive, web-based forms. These techniques afforded a greater measure of control than previously possible to a large number of users connected through the network. Nevertheless, the underlying code dictated the aesthetic aspects of the visualization. Even as users manipulated the content, system designers scripted the formal components that determined the overall aesthetic of the output.

One seminal example of the web-based approach to visualization of large datasets described above is Gapminder, developed by Swedish medical doctor and statistician Hans Rosling in 2006. His interactive system allows users to explore data related to the health and wealth of nations through the access and display of data from multiple sources in a unified and manipulable time-series graph interface. Gapminder appears simple, employing two-dimensional axes on which users select and plot data points from a set of given filters. This shifts the focus from a dynamic display intended to promote an aesthetic experience of the data, limited to the decisions made by the designer of the visualization, to an interactive software tool that enables users to filter or manipulate data directly. This represents a fundamental shift in the use of visualization in the context of large datasets. Although interactive tools like Gapminder provide greater freedom to the viewer to manipulate data, the designer still frames the data, determining the overall scope as well as individual options and even the style of display. Even here, where increased control is given to the user, the designer still exerts significant influence over the aesthetic qualities of the visualization.

Datasets to Data Streams.

Exponential growth in data production, collection, and distribution has required new techniques to engage with increasingly dynamic data streams. These methods transform the practice of data visualization from producing discrete pictures of static datasets to fluid interfaces capable of processing and displaying malleable streams of real-time data. The shift toward representing dynamic data streams enables visualization techniques to expand into new territories, merging with interface design and blurring distinctions between disciplines. Visualizations embedded in web browsers provide a visual history of past sites visited, subtly (or not so subtly) influencing future choices. Powerful analytics running behind popular social media sites attempt to predict which activities might be of most interest, providing new constellations expressed visually on news feeds and profile pages. Online retailers use metadata collected surreptitiously by sophisticated online trackers to tailor the display of merchandise to individual users visiting e-commerce sites. Eli Pariser's concept of the "filter bubble" suggests that the effect of such mass personalization might be increased isolation of users in specific cultural and ideological bubbles, sometimes without their consent or control. Together, these techniques represent a fundamentally different approach to data visualization in which the relationship of the front end (visualization) to the back end (data and the analytics that process them) reflects the dynamic nature of contemporary data. The aesthetic implications of this shift are far-reaching: data visualizations, in addition to their powerful role in creating meaningful information in editorial or scientific contexts, begin to inform the online experience in ways that are not always immediately apparent, or desired.

Data is collected, analyzed, and visualized at a breakneck pace for a wide range of purposes in the contemporary context. Early examples of data visualization were often intended to create meaning for limited audiences within particular disciplines. Data visualization is now so embedded in the current online experience that it approaches a mass medium, communicating across cultures and contexts, with a profound effect on the aesthetic character of the web. The ubiquity of data collection and visualization as a facet of everyday life is a self-perpetuating cycle. As the prevalence of data visualization grows, the public increasingly expects and relies on data visualization to inform basic choices. A structural shift in aims accompanies the widespread deployment of visualizations across the web. Rather than attempting to make meaning in large datasets, visualizations embedded in web browsers, e-commerce sites, and social media seek to directly influence human decision making in real time across contexts.

As the motivations underlying data visualization expand to encompass different objectives, it is possible to reduce them to two principal approaches, each of which produces a unique aesthetic outcome. The *rational* approach favors a positivist version of objective accuracy and faithfulness to the data, often in the pursuit of testing hypotheses or otherwise "finding" meaning in complex datasets. Such visualizations are often impersonal and descriptive, targeting discovery, and often assume a narrow sophisticated audience capable of interpreting images. One example of the rational approach is the meteorological visualizations used in weather prediction. These visualizations are constructed by analyzing data from hundreds of weather stations, ships, and aircraft that report readings of temperature, pressure, wind, moisture, and precipitation. Meteorologists use these models to produce complex visualizations of cloud cover, infrared radiation, and precipitation, which are typically rendered using standard methods of representation to avoid confusion among viewers. Such visualizations aspire to economy and clarity and make strategic use of graphic conventions in service of these aims.

In contrast to a rational approach to visualization, an *emotive* approach is projective, targeting communication of a particular perspective to a broad audience. Emotive approaches tend to be highly personal and subjective and are intended to convince, giving viewers reason to care about a particular problem or set of issues. Terminal Air, a visualization by the Institute for Applied Autonomy and experimental geographer Trevor Paglen, is an example of the emotive approach. Terminal Air enables users to track flights employed by the CIA for "extraordinary rendition," a secret program to transport terrorism suspects to overseas military bases around the world for interrogation and torture. Independent aviation enthusiasts exposed the program to professional journalists. Although based on raw data collected about the frequency of flight arrivals and destinations, Terminal Air nevertheless editorializes in a way that is as evocative as it is informative. The artists adopt the visual language of air traffic control and military logistics interfaces, employing aesthetic qualities to appear as accurate as those systems, and the institutional character of the visualization is juxtaposed uneasily with the reverberations of jet engines and a dark color scheme. These attributes create an ominous atmosphere that leaves little doubt about the artists' perspective on the subject matter.

While it is clear that an emotive approach incorporates the perspective of the visualizer, the rational approach too is often subject to bias based on the motivations and perspective of the visualizer. The question of motivation can be especially problematic when visualizations are used to advocate for a particular perspective, using the visual language of a more objective approach and cultivating a value-neutral appearance, when they are instead attempting to influence behavior.

The Neutrality of Data.

As data visualization is employed to reach wide audiences across contexts, it can appear as a value-neutral representation of "truth," borrowing from the verity of the scientific method to reveal what is implicit in a particular dataset. However, data visualization occurs within a particular context and is defined by cultural constraints particular to that context. Before its widespread prevalence, data visualization operated largely within disciplinary boundaries with shared conventions and mutually agreed-upon methodologies. However, as visualization enters the mainstream, often as a technique to influence behavior, recognizing the role of cultural context points to the need for transparency of visualization method and new technologies of accountability. This can be difficult or impossible when corporate policy or government regulation restricts access to the raw data.

The importance of acknowledging context is by no means unique to data visualization and analytics. The discipline of anthropology has grappled with the problematic aspects of judging one culture by the standards of another, generally described as enthnocentrism, since at least the middle part of the twentieth century. The institutional critique movement in contemporary art arose during the 1970s in response to the degree of influence exercised by galleries and museums in the determination of aesthetic taste. Like these other disciplines, data visualization is context-specific and rarely (if ever) a neutral act. As data visualization increasingly operates across disciplinary boundaries as a mass medium, transparency regarding methods of collection and visualization, acknowledging the possibility of multiple perspectives, and being forthright about the position of author in the visualization acquires a new importance. Together, these activities require the visualizer to account for context as well as their own role and bias in creating the visualization. Even positivist perspectives regarding the role of science in uncovering truth must be tempered by consciously acknowledging the difficulty of achieving value-neutrality in research. However, these values can sometimes be at odds with proponents of visualization when the objective shifts to influencing behaviors, as is often the case when visualization is used as a mass medium.

Ownership and Access.

A second contentious area of data visualization pertains to the access to the data itself. Although analyzing real-time datasets requires significant hardware to process complex analytics, contemporary data visualization is constrained more by access to data than to processing power. This shift is likely to continue in the future owing to increased processing capability of individual computers, accompanied by the rise of distributed computation, which harnesses the power of many machines across a network to accomplish a single objective. As the processing capability of individual users increases, government regulation and corporate policies that restrict access to data are a primary concern.

Data ownership can be especially problematic when it is collected without the knowledge or consent of users. Basic activities such as browsing the Internet, visiting social media sites, and using cell phones provide a stream of data to a second layer of less visible technologies. For instance, web trackers persistently log and store data about users' activities online, forming unique virtual profiles through the aggregation and analysis of these activities. These virtual profiles are combined with others to form networks of interactions that can be analyzed for a multitude of purposes. This enormous mass of data is constantly updated but is rarely provided in raw form to the individuals who have knowingly or unwittingly supplied it. Rather, ownership of data generally rests with its aggregators, who collect, mine, process, analyze, and distill these massive, dynamically changing real-time data streams for use in the form of small interface widgets, layout adjustments, and visual filters that we encounter through our experiences online. As aggregators collect more and more information, ownership of these data, and access to them, assumes greater importance. The open data movement has arisen in part due to concerns about the importance of making data accessible to the dissemination of knowledge and government accountability.

Data visualization is likely to play an increasingly important role in the future. A pervasive array of sensors gathering data at an unprecedented rate will accompany the rise of ubiquitous computing, increasing the relevance of effective analytics and techniques capable of finding meaning in large datasets. Accountability, methodological transparency, and acknowledgment of cultural bias and competing agendas will be essential to ensuring that visualization as a mass medium clarifies issues rather than obfuscates them.

Data collection, analytics, and its display through visualization and interactive interfaces increasingly define our aesthetic experiences online and in the world, changing the visual content of the web and influencing behaviors in subtle ways. The roots of data visualization can be traced back close to five hundred years within disparate disciplines. However, in the past few decades, fundamental changes have occurred due to the rise of processing power coupled with data collection on a scale greater than any point in history. As data analysis, visualization, and interface design merge, they provide customized, personalized interfaces that predict desire in ways that enable users to navigate efficiently through a vast array of online options. However, a more critical interpretation might hold that the prescriptive aspects of back-end analytics using data not knowingly supplied influence the aesthetic character of experiences in undesirable ways. Whether positive or negative, contemporary data visualization is increasingly functioning as a mass medium, in which visualization techniques are merging with data analytics and interaction design to exert a great influence on lives online and off. This new constellation has a significant effect on the practical as well as aesthetic aspects of contemporary life.

[See also Computing, Aesthetic; and Information Theory.]

Bibliography

Anderson, Chris. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *Wired*, 23 June 2008. http://www.wired.com/science/discoveries/magazine/16-07/pb_theory.

Bostock, Michael, Vadim Ogievetsky, and Jeff Heer. "D3: Data-Driven Documents." *IEEE Transactions on Visualization and Computer Graphics* 17, no. 12 (2011): 2301–2309.

Diamond, Sara. "Lenticular Galaxies: The Polyvalent Aesthetics of Data Visualization." http://www.ctheory.net/articles.aspx?id=651.

Frid-Jimenez, Amber, and Ben Dalton. "Data Is Political: Investigation, Emotion and the Accountability of Institutional Critique." In *Accountability Technologies: Tools for Asking Hard Questions*, edited by Dietmar Offenhuber and Katja Schechtner, pp. 14–22. Heidelberg, Germany: Springer-Verlag, 2013.

Friendly, Michael. "A Brief History of Data Visualization." In *Handbook of Computational Statistics: Data Visualization*. Vol. 3, edited by Chun-Houh Chen, Wolfgang Hardle, and Antony Unwin, pp. 1–34. Heidelberg, Germany: Springer-Verlag, 2006.

"Gapminder World." http://www.gapminder.org/world.

Harley, J. B., and David Woodward. *The History of Cartography*, 3 vols. Chicago: University of Chicago Press, 1987–2007.

Institute of Applied Autonomy and Trevor Paglen. "Terminal Air." http://www.appliedautonomy.com/terminalair/index.html.

Manovich, Lev. "The Anti-Sublime Ideal in Data Art." 2002. http://www.manovich.net. Originally published as "The Anti-Sublime Ideal in New Media," *Chair et Métal 7* (2002).

Manovich, Lev. "Introduction to Info-Aesthetics." http://manovich.net/Manovich.InfoAesthetics.2008.pdf.

Pariser, Eli. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin, 2011.

Shannon, Claude E. “A Mathematical Theory of Communication.” *Bell System Technical Journal* 27 (October 1948): 379–423, 623–656.

Tufte, Edward. *The Visual Display of Quantitative Information*. Cheshire, Conn.: Graphics Press, 1983.

Valéry, Paul. “The Conquest of Ubiquity.” In *Aesthetics*, translated by Ralph Manheim. London: Routledge and Kegan Paul, 1964.

van der Velden, Daniel, and Vinca Kruk. “Captives of the Cloud: Part I.” e-flux (September 2012). http://www.e-flux.com/journal/captives-of-the-cloud-part-i/.

Amber Frid-Jimenez and Joseph Dahmen

Oxford University Press