Title
Contemporary Data Visualization: A Cultural History and Close Readings

Author
Zepel, Tara

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Contemporary Data Visualization: A Cultural History and Close Readings

A Dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Art History, Theory and Criticism

by

Tara Zepel

Committee in charge:

Professor Lev Manovich, Co-Chair
Professor Peter Lunenfeld, Co-Chair
Professor Benjamin Bratton
Professor William N. Bryson
Professor James Hollan
Professor Elizabeth Losh

2018

Signature Page

The Dissertation of Tara Zepel is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Co-Chair

University of California, San Diego

2018

TABLE OF CONTENTS

TABLE OF FIGURES

VITA

2006 Bachelor of Arts, Duke University

2008-2016 Teaching Assistant, Department of Visual Arts, University of California, San Diego

2007-2011 Teaching Assistant, Culture, Art and Technology Writing Program, University of California, San Diego

2016-2017 Lecturer, Media Arts, Society & Technology minor, Cal Poly, San Luis Obispo

2018 Doctor of Philosophy, University of California, San Diego.

PUBLICATIONS

“How to Compare One Million Images?” with Lev Manovich and Jeremy Douglass, 2011. Online at https://www.academia.edu/9813363/How_To_Compare_One_Million_Images.

“Revisiting the Twitter Experiment (In Context)." *The New Everyday* (March 2011): Online at http://mediacommons.futureofthebook.org/tne/pieces/revisiting-twitter-experiment-context-0.

“Cultural Analytics at Work: The 2008 U.S. Presidential Online Video Ads.” *Video Vortex Reader II*, edited by Geert Lovink and Somers Miles, 234-249. Amsterdam: Institute of Network Cultures, 2011.

FIELDS OF STUDY

Major Field: Art History

Studies in Media Studies, Digital Humanities, and Cognitive Science

ABSTRACT OF THE DISSERTATION

Contemporary Data Visualization: A Cultural History and Close Readings

by

Tara Zepel

Doctor of Philosophy in Art History, Theory and Criticism

University of California, San Diego, 2018

Professor Lev Manovich, Co-Chair

Professor Peter Lunenfeld, Co-Chair

This dissertation is the first in-depth study of a new important area of contemporary visual and digital culture - data visualization. First developed at the end of the 18th and early 19th century, data visualization until recently has been understood as an analytic tool for expert use. However, a growing number of projects have challenged these assumptions. The expansion of a data visualization into art (including many exhibitions in leading art museums), social activities, and nearly every dimension of life that begins around 2004 indicates a far more complex set of interactions between representation, viewer, and data than it was assumed earlier. While a small handful of scholars have begun to investigate data visualization's untraditional or alternative uses, there is still no in-depth study of how and why data visualization functions in contemporary society and culture.

My work lays the groundwork for seeing data visualization as a socially and culturally situated medium and practice. I examine my subjects by combing methods and concepts from a number of disciplines: media studies, art history, cognitive science, and design. These disciplines have not been brought together so far in investigating contemporary data visualization culture, so this is a methodological innovation of the dissertation.

The presentation of the material is organized into two parts. The first part presents a cultural history of data visualization as it has developed alongside digital culture and technology since 1970s until the present. In the second part, I analyze how data visualization functions today in different contexts via close reading of select projects. Such close analysis is common in art history, film studies or literary studies, but has not yet been applied to data visualization projects. My readings test theoretical ideas of the dissertation, while also showing how we can how we can think of data visualizations as complex cultural objects not unlike paintings, films or novels.

INTRODUCTION

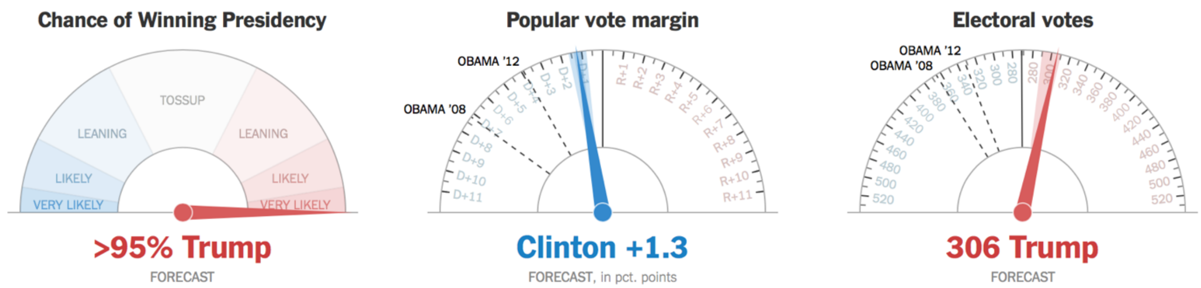

Figure 0.1 Screenshot of *The New York Times* interactive election graphic (2016).

November 8th, 2016. It was the evening of the U.S. Presidential election. The three gauges shown above featured prominently on the New York Times' website. As the various precincts began to report, the graphic took on life. Needles jittered back and forth. Numbers for each candidate changed. By eleven o'clock at night, it was clear who would become the next president of the United States.[1] I remember the experience of watching this graphic. Notably, it was on a computer screen. And notably, it made a large amount of information easier to understand. What I remember the most vividly, however, is the feeling with which I watched. Every time the dials moved, I felt. I felt captivated and alert. I felt my hopes for the country grow and fade. I felt the polarized nation. And I felt my stomach move. One viewer summarized the

[1] Several people, including the graphic's designers, posted about its design and effect after the fact. For a blogpost with screenshots at different intervals throughout the night, please see: J.K. Trotter, "The New York Times Live Presidential Election Meter Is Fucking With Me," Gizmodo.Com, November 8, 2016, http://gizmodo.com/the-new-york-times-live-presidential-meter-is-fucking-w-1788732314.

feeling well by calling the visual “an online widget [that] brilliantly correlates to my blood pressure.”[2] It seemed as though this graphic was determining the country’s political fate. Except it was not. It was simply representing data in a way that evoked tremendous social and political impact.

Data visualization is the visual representation of data of any type. *The New York Times* election meter is just one example of how data visualization and contemporary society interact. Once contextualized as a tool for expert use, data visualization now pervades. We have data visualizations about the human genome, data visualizations about our buying habits, data visualizations about our social networks, and data visualizations about our art. We use data visualizations analytically, to help us better understand large amounts of information; socially, to communicate and explore relationships with others; and culturally, to express identity and as art. When a NASA scientist plots temperature data to model the atmosphere on Mars, that is data visualization.[3] When the feature on your smartphone app lets you map your run and share it with friends, that is data visualization.[4] When a digital humanities scholar represents Shakespeare’s tragedies as network graphs, that is

---

[2] Trotter.

[3] See Greg Shirah, Mars’ Lost Atmosphere, Data Visualization, 2015, https://svs.gsfc.nasa.gov//12046.

[4] Several apps give you the ability to track exercise, set goals and share your progress with a larger community. Some, such as Map My Run and the Nike+ Running App, are for social network oriented. Others, such as Strava or vivoactive®HR, add the ability to visualize more than GPS data and run analytics.

also data visualization.[5] Data visualization has essentially become another language through which we communicate and experience life.

This dissertation is about situating data visualization in its larger social, cultural, and historical contexts. What do I mean by this? At the most basic level, I mean recognizing that data visualization, just like any other representational medium and practice, is conditioned by a particular set of circumstances, a particular set of goals, and particular ways of knowing. The pages that follow offer the first book-length investigation of contemporary data visualization—that is data visualization as it has developed alongside computer culture and technology from the late 1970s to the present—in order to make the point that data visualizations must be understood as situated in the circumstances in which they are designed and seen. Combining ideas and methods from media studies, art history, cognitive science, and design, I look at how data visualization has become a pervasive socio-cultural form, what assumptions and influences are at play, and what this means for the way we conceptualize and approach data visualization research and design. I approach the study of data visualization much like the study of photography or film, through close reading and visual analysis. Why did contemporary data visualization emerge when it did? What factors were at play? What occurs in the relationship between viewer, representation, and environments? And how do formal as well as contextual elements influence the

[5] See Martin Grandjean, "Network Visualization: Mapping Shakespeare's Tragedies," Personal Website, Martin Grandjean, (December 23, 2015), http://www.martingrandjean.ch/network-visualization-shakespeare/.

meaning, reception, and design of visualizations today? These are just some of the questions this dissertation asks.

In doing so, I intend to make a methodological argument about how to best move forward with visualization design and research. The past fifteen to twenty years have been characterized by data visualization moving outside of traditional research environments and into nearly every dimension of life. While a wide variety of scholars have noted this shift, only a handful have posed alternative or updated approaches to visualization research and design. These tend to be framework-driven or remain abstract. In beginning with the concrete, in-depth analysis of existing projects, I hope to perform that background work necessary to build such an approach. Data visualizations are not simply analytic representations of data, and it is time we stop seeing them as such.

Data Visualization as a Socio-Cultural Form

Recognizing that data visualization has become a pervasive socio-cultural form requires little work. The extent to which data visualization is implicated in how we understand and experience the world is vast. The examples given at the beginning of this introduction are just a few out of many. From the latest scientific research to how we interact with friends, data visualization plays a role. The question is what role does it play? Recent developments in visualization design and research have challenged the assumption that data visualization serves as an objective, analytic tool for data's display. While we do use data visualizations analytically to help us understand data, we also use them to express, to communicate, to feel, and to persuade.

Take, for example, Stefanie Prosavec and Giorgia Lupi's *Dear Data* project. Created shortly following the designers met in 2014, it quite literally answers the question of can two people get to know each other through data alone. Every week, for an entire year, Prosavec and Lupi collected and measured a particular type of data about their personal lives (e.g. the times they laughed, the time they complained, the times they engaged in certain activities, etc…). They then represented this data on a postcard-sized sheet of paper and dropped the postcard in a mailbox to be delivered to the other's address. Prosavec lives in London; Lupi in New York. What resulted is a collection of rather intimate hand-drawn visualizations (Figure 0.2) that both record and give insight into a person's day-to-day life. Originally published as a website (dear-data.com/theproject), the project has since been turned into a book.[6] Taken as a whole, the experience of viewing Dear Data is a bit like the experience of looking through a box of photographs or similar artifacts. Each is personal, each is data-based, and each takes on additional meaning because of something that is not physically represented on the screen (or page).

[6] Giorgia Lupi and Stefanie Prosavec, *Dear Data* (Princeton Architectural Press, 2016), https://www.amazon.com/Dear-Data-Giorgia-Lupi/dp/1616895322/ref=sr_1_1?ie=UTF8&qid=1486748686&sr=8-1&keywords=dear+data+book.

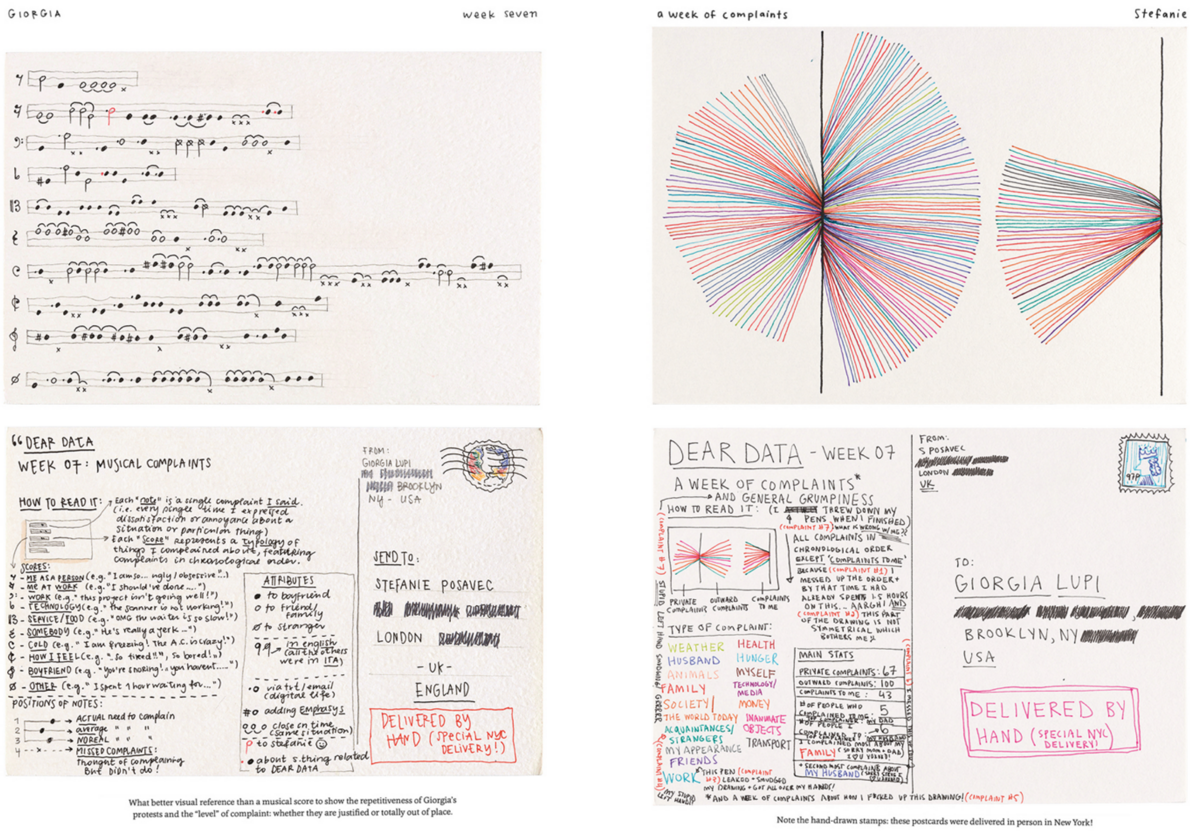

Figure 0.2 Stefanie Prosavec and Giorgia Lupi. Screenshot of *Dear Data* (2014).

This dissertation proceeds from the basic premise that data visualizations, like any other visual media, both reflect and affect how we *see*. In 1972, art historian John Berger wrote and starred in a television miniseries called *Ways of Seeing*.[7] In it, he presents his insights on how our very sense of sight has been transformed. To see, he argues, is a political act. To look at an image is to engage in a historically and socially constructed process such that when and where we see something affects what we

[7] John Berger, *Ways of Seeing* (London: BBC and Penguin Books, 1972).

understand from looking at it. In other words, seeing requires context. The same photograph of a woman driving means something different depending on where it is displayed, who is looking at it, and the circumstances in which it is seen. The idea that images are constructions and carry social and cultural meaning is an important tenant of visual culture, art history, and film studies, but one that is often not considered when looking at data visualizations. Why?

Data visualization has reached a critical point where it needs the same. As data visualization has become more and more a part of everyday life, questions about its values and ways of constructing knowledge arise. One prevalent critique coming out of the digital humanities has been that data visualization is overly positivistic.[8] I would argue that the first step to addressing this is not to assume that positivism is inherent to the practice but rather the result of values and assumptions that have been put into place at specific points in time. In her forthcoming book *Data by Design*, Lauren Klein traces the rise of modern data visualization techniques to show how historical epistemologies, as much as form, continue to influence the design, reception, and rhetoric of visualization today. In many ways, this dissertation does the same except that the period and examples I focus on are drawn from the recent past and present. Instead of focusing on the development of techniques, I focus on the development of a contemporary visualization culture. It is only by understanding this culture and the various events and values that have influenced it that we can begin to understand our

---

[8] Anne Burdick et al, *Digital_humanities* (Cambridge, MA: MIT Press, 2012), 44.

present relationship with data visualization and what we want out of it in the future. As data visualization becomes increasingly popular and pervasive, this is an increasingly important agenda. It is important to understand how we “see” and interact with data representations. This feedback is equally important for future visualization research and design. And it important for developing a basic visualization literacy.

Motivations and Related Work

The motivations behind my research are threefold. First and foremost, a cohesive history of contemporary data visualization has not been written. Most histories of data visualization begin in the late 1700s with the invention and popularization of many of the most common graphical forms still used today (e.g. bar chart, line graph, etc…). They then continue to lay out the work of key individuals and projects, which in retrospect explain the linear development of visualization as a cognitive tool for data analysis and comprehension. The official birth of the field in 1986 and its consequent explosion throughout the 1990s is then taken as a near ending point. There are two problems with this. First, the development of data visualization as a field and practice is and never has never been linear. Second, it stops short. While such overviews will often acknowledge data visualization’s continued diversification and growth, they stop short of examining and theorizing the present. The handful of articles and discussions that do exist remain scattered on websites, in conferences, and on blogs. No in-depth, comprehensive account of contemporary data visualization exists.

Secondly, this dissertation responds to a call in the existing visualization literature for more critical reflection and theoretical work. The idea that we might need to change the way we conceptualize and approach data visualization is not my own. The rise of several untraditional forms of data visualization in the late 1990s and early 2000s raised questions about differing audiences, ideals, and goals. If data visualization could serve as art, or as a social facilitator, and reach non-expert audiences (among many other things), then traditional approaches to research and design were likely out of date. Most advancements to this regard occurred in the design of the projects themselves with only a handful of scholars stopping to reflect or theorize the changes that were taking place. Meanwhile, data visualization continued to change. Over the past fifteen years, data visualization has become more and more important to how we experience life. It makes sense that we might want to think about how data visualizations affect our understanding of the world, what values and assumptions they communicate, and what direction we want their future development to take.

There is a small body of literature coming out of the digital humanities and similarly interdisciplinary research that has begun to perform this work. Critical and feminist approaches to data visualization aim to uncover taken-for-granted assumptions about visualization and reflect on what values are portrayed. The same data can be represented in infinite ways. By focusing on how a particular visualization is constructed and what expectations dominate the field, such approaches emphasize the need to think about the situational factors that influence a given display. Critical and feminist approaches to data visualization also aim to create more inclusive and nuanced experiences of interacting with data. This intention has been echoed by recent calls to humanize data and its representation.[9] There seems to be a growing recognition that merely technical approaches to visualization need to change if we want to use data visualization to its fullest capacity. This dissertation is in conversation with this line of thought, which is explained in more detail in Chapter 2.

There is also a way that this dissertation is in conversation with a paradigm in cognitive science known as situated cognition. The idea that what we know and what we think are dependent on context spans a broad range of scholarly work in psychology, anthropology and cognitive science. In 1987, anthropologist Lucy Suchman wrote a book called *Plans and Situated Actions* in which she argues that every course of human action is constructed from dynamic interactions with the material and social world.[10] Suchman grounds her investigation in the observation of human-machine communication, specifically the communication between novice users and the double-sided function of a copier. What she found was that users figured out

---

[9] Giorgia Lupi, "Data Humanism: The Revolutionary Future of Data Visualization," *Print Magazine*, January 30, 2017, http://www.printmag.com/information-design/data-humanism-future-of-data-visualization/; Giorgia Lupi, "Data Humanism – The Revolution Will Be Visualized" (IEEE VIS 2017, Pheonix, AZ, October 6, 2017), https://www.youtube.com/watch?v=S0YkTtLFIDs.

[10] Lucy A. Suchman, *Plans and Situated Actions: The Problem of Human-Machine Communication* (Cambridge University Press, 1987).

how to use copy feature not as a result of any pre-conceived, rational plan but as an emergent property of the moment to moment interactions between actors and between actors and their environment. In other words, people use their circumstances to achieve mutual understanding and intelligent action. This approach to the study of context is known as the situated action model. It has been expanded upon by others and emphasizes the emergent and contingent nature of human activity.[11]

Two related approaches in cognitive science that developed around the same time are activity theory and distributed cognition. Activity theory maintains that the constituents of an activity are not fixed but can dynamically change. It proposes a very specific notion of context – that of the activity system itself. Context is constituted through the enactment of relations involving people, operations and artifacts. There is no context out there inside which this interaction takes place.[12] Distributed cognition is an approach to cognitive science that studies the representation of knowledge both inside and outside individuals' heads. It sees cognition as distributed across a collection of individuals, artifacts, and/or time in a particular work setting.[13] All three

---

[11] Jean Lave, *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life* (Cambridge University Press, 1988); Lucy A. Suchman and Randy Trigg, "Understanding Practice: Video as a Medium for Reflection and Design," in *Design at Work: Cooperative Design of Computer Systems*, ed. Joan Greenbaum and Morten Kyng (Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1991).

[12] For summaries of activity theory, see: A. Leont'ev, "The Problem of Activity in Psychology`," *Soviet Psychology* 13, no. 2 (1974): 4–33; Susanne Bødker, "A Human Activity Approach to User Interfaces," *Hum.-Computer Interact.* 4, no. 3 (September 1989): 171–195, https://doi.org/10.1207/s15327051hci0403_1; K. Kuutti, "Activity Theory and Its Applications to Information Systems Research and Development," in *Information Systems Research*, ed. H. E. Nissen (Amsterdam: Elsevier Science Publishers, 1991), 529–49.

[13] Edwin Hutchins and Nick Flor, "Analyzing Distributed Cognition in Software Teams: A Case Study of Team Programming during Perfective Software Maintenance," in *Proceedings of the*

approaches to the study of context have contributed to the development of data visualization as a field and as a practice. However, context is something that is rarely accounted for in the visualization literature outside of material or technical specifications.

Developing an Updated Approach

If data visualization is constructed, if it must be read as situated in the circumstances in which it is designed and seen, then an updated, more reflective and human-focused approach to data visualization design and research is needed. The pages that follow largely perform the background work necessary for beginning this task. However, unlike existing research, I do not begin with a particular perspective or framework. Rather, I begin with close readings of the projects themselves. While frequently applied to the study of literature, film, and art, such close reading has not yet been applied to data visualization as a systematic method of research. Data visualization is an inherently interdisciplinary field. By beginning with close, concrete analysis of example projects, I hope to avoid the problem of starting with a disciplinary theory vocabulary that might not be applicable to all data visualizations or understood by all participants. In many ways, this dissertation makes the argument that all data

---

Fourth Annual Workshop on Empirical Studies of Programmers, ed. Jürgen Koenemann-Belliveau, Thomas G. Moher, and Scott P. Robertson (Norwood, N.J.: Ablex Publishing, 1991), 36–59; Edwin Hutchins, *Cognition in the Wild* (MIT Press, 1996).

visualizations must be read individually. The initial approach of close reading is an attempt to emphasize this specificity.

The dissertation proceeds in two parts. Part 1 (Chapters 1 and 2) charts the trajectory of what I am calling contemporary data visualization. In it, I analyze how key events, projects, and trends both inform and reflect the course of data visualization's development from the late 1970s to the present. Importantly, no single storyline or perspective is given preference. Rather, visualization's development is told through a complex array of social, cultural, technological, and political factors that were dominant at certain points in time. My aim is not to provide a comprehensive history. Rather my aim is to highlight some of the main influences in contemporary data visualization's development, and show how these, at least in part, are reflected in the way visualization is used, designed, and conceptualized today.

Chapter 1 examines the period leading up to and immediately following contemporary data visualization's birth. It spans from roughly the late 1970s to 2000. The chapter begins by introducing the term contemporary data visualization as a way to differentiate between earlier forms of data representation and the pervasive socio-cultural form that exists today. I then move on to examine three broad trends that set the stage from which contemporary data visualization emerged – the development of supercomputing, a timely interest in visual representation, and the development of computer graphics. No single storyline or perspective is given higher status. Social, cultural and technological factors are considered alongside together. The chapter then proceeds to tell the story of contemporary data visualization's birth, pointing to the

1986 *Visualization in Scientific Computing* (ViSC) Report as its official decree, and outline its earliest development.[14] Scientific visualization and information visualization are examined as the two dominant genres that emerged during this period.

Chapter 2 picks up where the previous chapter left off. It examines the various ways in which contemporary data visualization has expanded outside of traditional research environments and into nearly every dimension of life. Several technological and social developments over the past fifteen years such the popularization of the Internet and lowered barriers to access for data and software have open visualization up to new audiences, settings, and goals. The chapter is structured as a survey of three broad areas in which this has arguably been the case – visualization as art, social data visualization, and visualization and the everyday. While these are not the only three areas in which data visualization has expanded, they are the most prominent. By conversing with example projects and relevant literature, I show in very concrete terms what is meant by the claim that data visualization has become a pervasive socio-cultural form. More importantly, I show a recognized need to rethink existing approaches to visualization design and research. The chapter concludes by introducing alternative approaches that have recently emerged in an effort to make data visualization more reflective and human-focused.

---

[14] Thomas A. Defanti, M axine D. Brown, and Bruce H. McCormick, "Visualization in Scientific Computing" (New York: ACM SIGGRAPH, 1987), http://www.sciencedirect.com/science/article/pii/S0065245808601680.

Part II (chapters 3, 4, and 5) shifts the discussion to the single project scale. It offers close readings of select visualization projects in order to expand upon ideas and concepts introduced in Part I. Part II additionally serves to model where any current approach to visualization must begin. Beginning with a preconceived framework risks not accounting for the true complexity and interdisciplinary that many contemporary data visualizations entail. Similarly, beginning with abstract concepts does not account for the specificity and context of individual projects. It is only by beginning with careful and close visual analysis of the projects themselves that we can begin to construct and theorize what an alternative or updated approach to data visualization research and design might be.

Chapter 3 analyzes Dustin Cable's *Racial Dot Map* through the lens of current socio-cultural events and personalized experience.[15] It argues that data visualizations, much like documentaries, are representations that give evidence of the culture that produced them and the things represented in them. Importantly, what is represented – that is what is seen – takes on meaning based on context, audience and an array of external factors that may or may not be directly represented on the page or screen. In depicting racial diversity and distribution across America based on 2010 census data, the *Racial Dot Map* involves a whole lot more. It involves the country's consciousness at the time, it involves a shared racial history, and it involves personal geographic

[15] https://demographics.virginia.edu/DotMap/

experience. Such external factors play a key role in the visualization's effectiveness and meaning.

One important aspect of seeing visualization as a socially and culturally situated practice is recognizing that visualizations may have an agenda. This does not necessarily mean that they are not accurate representations. Chapter 4 investigates how rhetorical purpose and emotion are incorporated into the design of *U.S. Gun Deaths*.[16] Through first-person accounts and close visual and textual analysis, I show how the careful incorporation of emotive details can be used strengthen the message of a visualization without necessarily obscuring or misrepresenting the data. To this regard, the chapter addresses questions of ethics, data selection and social activism through data visualization.

[16] http://guns.periscopic.com/?year=2013

PART I

CHAPTER 1

Contemporary Data Visualization (1970s – 1990s)

In October 1986, the National Science Foundation (NSF) sponsored a meeting of the newly organized Panel on Graphics, Image Processing and Workstations. Their goal? To provide advice on the acquisition and establishment of graphics and image processing hardware and software at research institutions doing advanced scientific computing. Of particular interest were five recently established national supercomputing centers. Researchers at the centers had repeatedly been requesting funds towards better graphics resources and research.[1] The 1980s were a period of extreme technological growth, particularly in relation to computing power and speed. Newer, faster, and more powerful forms of computing technology were streaming in data at quicker rates and in greater volume than ever before. The situation was one of overwhelming influx. Nowhere was this more acutely felt than in the sciences. Scientists and other researchers at the forefront of using computational technology faced an ever-growing problem—too much data. The sheer amount was difficult to read, let alone understand. Graphically representing all of this data offered a way to help.

A few short months following their initial meeting, the Panel organized a workshop to explore the situation surrounding the scientists' repeated requests. On

[1] Thomas A. Defanti, Maxine D. Brown, and Bruce H. McCormick, "Visualization in Scientific Computing" (New York: ACM SIGGRAPH, 1987).

February 9th and 10th, 1987, researchers from across industry, academia, and government met in Washington D.C. to discuss what they saw as an emergent field called visualizing in scientific computing. The ensuing report, known simply as the *Visualization in Scientific Computing* (ViSC) Report, summarized the Panel's findings:

> Visualization in Scientific Computing (ViSC) is emerging as a major computer-based field, with a body of problems, a commonality of tools and terminology, boundaries, and a cohort of trained personnel. As a tool for applying computers to science, it offers a way to see the unseen. As a technology, Visualization in Scientific Computing promises radical improvements in the human/computer interface and may make human-in-the-loop problems approachable.[2]

As such, better resources were direly needed.

This report and the events leading up to it are now largely considered to be the birth of what I am calling *contemporary data visualization*—that is, data visualization as it has developed in an inextricable relationship with computing culture and technology. While practices of graphically representing data certainly existed prior to this point, it was not until the mid-1980s that the dedicated field and study of how and why we represent data emerged. That emergence is the story of this first chapter. What follows charts the trajectory of the contemporary data visualization field, from the period leading up to its birth in 1987 through the first decade of its development. The chapter begins by offering a brief overview of three important social and technological trends that set the stage for the field's birth. It then moves on to provide a detailed account of the publication and energy surrounding the ViSC Report. The chapter concludes by tracing the initial growth of scientific visualization and information

[2] Defanti, Brown, and McCormick, "Visualization in Scientific Computing," ix.

visualization, the two primary genres that emerged within the field. Emphasis is placed on multiple stories and multiple disciplinary perspectives. The story of contemporary data visualization's initial development is far messier and more complex than the often condensed and partial accounts show. Science, data inundation, and computing did play a role, but so did art, Hollywood, and an increasing interest in visual representation.

The Set-Up: Pre 1987

There is no one place to begin. The build-up to contemporary data visualization follows a series of interlocking social, cultural, and technological events and trends that characterized the second half of the twentieth century. The difficulty of writing a complete history is made evident by the lack of detailed accounts that exist. Most histories of visualization focus on the work of a few key players or take a single disciplinary perspective and gloss over the rest. Because of the field's interdisciplinary and multiple areas of influence, no truly complete history can exist. The first part of this chapter tries to offer a more representative account by exploring three broad areas of influence: supercomputing and the problem of too much data; an increased interest in visual representation; and the development of computer graphics. Specific attention is paid to where they overlap and converge.

Supercomputing and the Problem of Too Much Data

The years immediately leading up to and following the birth of contemporary data visualization were characterized by extreme technological growth, particularly in relation to computing power and speed. By the 1980s, computers and other telecommunication technologies were streaming in data at such unprecedented rates that all researchers could do was gather and store the numbers they generated. The authors of the ViSC Report identify what they call several "*fire hoses* of information."[3] While no single data source was responsible for the feeling of inundation, there was one class of machine that loomed high above the rest: supercomputers. Supercomputers, simply put, are the highest performing computers at any given point in time. They are often large, exceed existing limits of power, speed, and capacity, and stand at the forefront of computing technology. The story of contemporary data visualization is intimately tied to their development.

The history of supercomputing dates back to the late 1950s with the Atlas at Manchester University.[4] Its development marks the first of many firsts in what would become an ongoing arms race to build the world's fastest and most powerful computer. The introduction of FORTRAN as the first high-level programming greatly reduced the amount of time needed to write and assemble code.[5] This, combined with

---

[3] Defanti, Brown, and McCormick, "Visualization in Scientific Computing," 4.
[4] A short film produced by Google to commemorate the Atlas' 50th anniversary provides a nice compliment to the academic sources that follow. Available at https://www.youtube.com/watch?v=6TRfy70DqD8.
[5] "The FORTRAN Programming Language," accessed July 21, 2016, http://groups.engin.umd.umich.edu/CIS/course.des/cis400/fortran/fortran.html.

the high-performance needs of the scientific and engineering applications it helped run, introduced a demand for more powerful computers. In 1959, the British engineering firm Ferranti teamed up with the university's high-speed computing project (MUSE) to build an ambitious machine that would exponentially improve upon existing models. The project, renamed Atlas, marks the first of many firsts in what would become a decades-long arms race to build the world's fastest and most powerful computer.[6] Part of the motivation behind building such a machine was the physical need. The other was the desire to beat out a handful of other companies, particularly those in the United States that were attempting similar projects. In December 1962, the Atlas won the race. Performance in supercomputing is measured in floating-point operations per second, or flops. The more flops a computer has, the more calculations it can process. When first installed and successfully run, the Atlas ran at an unprecedented one million flops, making it one hundred times faster than existing Ferranti models and nearly four times more powerful than the most powerful computer in the world at the time.[7] The era of supercomputing had begun.

Supercomputers, however, were not called supercomputers yet.[8] The first machine to receive this designation was the Cray-1. It ranked as the world's fastest

---

[6] Simon Lavington, "The Atlas Story," (Atlas Symposium, Computer Science, Kilburn Building, Manchester University, 2012).

[7] Simon Lavington, "The Manchester Mark I and Atlas: A Historical Perspective," *Communications of the ACM* 21, no. 1 (January 1978): 9.

[8] This is a pattern that emerges again with the introduction of the term visualization. New eras, trends, and technologies are frequently given a title only after they have been in existence for some time.

from 1976 to 1982. Designed and engineer by Seymour Cray, the so-called "father of supercomputers," the Cray-1 looked and performed like no other computer. A unique "C" shaped design enabling wiring to be kept closer together combined with a novel cooling system using Freon gave the machine a very high-performance level. The first CRAY-1® system, installed at Los Alamos National Laboratory in 1976, boasted a top speed of one-hundred and thirty-three million flops (133 megaflops) and an eight megabyte (1 million words) main memory.[9] High performance, however, came at a high cost. The system installed at Los Alamos cost approximately 8.8 million dollars.[10] For the first few decades, from the 1950s through the 1970s, supercomputers remained so rare and expensive that only large government agencies could afford them. Still, demand grew. The new breed of computers proved incredibly useful, first to aid in the design of nuclear weapons and breaking code and then in setting up experiments to create thermonuclear fusion, predict weather, and model climate. Governments continued to encourage the development of more and more powerful machines.[11]

As is often the case with developing technology, costs eventually went down. Capital-intensive industries in the United States, such as oil and automotive, began to

---

[9] "Cray 1," The History of Computing Project, accessed July 22, 2016, http://thocp.net/hardware/cray_1.htm; "The Cray-1 Computer System" (Cray Research, Inc., 1997), http://archive.computerhistory.org/resources/text/Cray/Cray.Cray1.1977.102638650.pdf.
[10] "Cray 1."
[11] William J. Kaufmann and Larry L. Smarr, *Supercomputing and the Transformation of Science* (New York, NY: Scientific American Library 1993), 19.

employ supercomputers in the late 1970s. Throughout the 1980s, supercomputing rapidly spread across industry and into academic research. Engineers and researchers in fields as diverse as aerospace, pharmaceuticals, and energy, among others, used the powerful machines to run simulations of performance and model natural phenomena.[12] By the end of the decade, supercomputing was considered fundamental to cutting-edge scientific and engineering research. Other countries experienced similar growth. While supercomputing was being adopted by industry in the United States, a number of open scientific research centers in Europe acquired the machines. Supercomputing, however, remained expensive. Academic researchers in the United States had to travel to Europe or federal research facilities to gain access. In 1985, the NSF initiated funding for five national supercomputing centers: The John Von Neumann Center at Princeton University; Cornell Theory Center; Pittsburgh Supercomputing Center; National Center for Supercomputing Applications at the University of Illinois in Urbana-Champaign; and San Diego Supercomputer Center at the University of California, San Diego. These centers would serve as the heart of the nation's high-performance computing and development for the next five to ten years.[13] The official birth of contemporary data visualization was not far behind.

---

[12] Ibid.

[13] Supercomputing centers were established at the University of California, San Diego, Cornell, Carnegie Mellon/University of Pittsburgh, Princeton, and the University of Illinois at Urbana-Champaign. All still exist today with the exception of the center at Princeton, which closed in 1990 because the company that made the two computers housed at the center (ETA systems) went out of business.

In their book *Supercomputing and the Transformation of Science*, William J. Kauffman and Larry Smarr argue that supercomputers transformed what it meant to do engineering and science. What supercomputing offered that previous technologies had not was the ability to run an enormous number of calculations. More calculations meant more complex models that could better approximate conditions in the real world. The ability, for example, to model a colliding arrangement of stars as millions of numbers, as opposed to hundreds, created a digital reality that allowed scientists to simulate natural phenomena in infinitely more detail. Variables that had previously been unaccounted for could now be calculated and plugged back in. More complex models, however, also meant more output. This is where the need for something like visualization began to form. The amount of data generated by running supercomputer simulations very quickly began to add up. By the end of the 1980s, contemporary supercomputers were running at a rate of up to one billion flops. The amount of data output by such a machine is immense. Kauffman and Smarr offer a perceptible idea of just how much data this is:

> In one second, a modern supercomputer can generate one billion numbers, which if printed out in 10 columns of 50 lines a piece of each page, would require a pile of paper over 50 stories tall![14]

[14] Kaufmann and Smarr, *Supercomputing and the Transformation of Science,* 3.

Supercomputing simulations typically run for hours. Multiplying the fifty stories by the appropriate number of seconds clarifies the problem the scientists faced.

Scientists became very good at flipping through pages of numbers, but at some point, the data become too much.[15] It is not difficult to imagine the problem scientists and other researchers at the supercomputing centers faced. They had mounds upon mounds of data, but a very inefficient way of looking at it. So, they started doing what they do best: They began to experiment. While there was no official visualization field yet, scientists and researchers had an intuitive sense that seeing their data in a different light would help them to better understand it. Early experiments included building three-dimensional models and rough graphic displays. One particularly innovative group of researchers working with storm data turned to wires and glue. Scientists had already identified the external features of a storm system, but they knew relatively little about its internal structure. Using supercomputers and data collected from a group of sixteen tornadoes that struck Oklahoma on May 20, 1977, Peter Ray, Robert Wilhelmson, and Kenneth Johnson were able to construct a digital map of the rain intensity and velocity inside a storm at six different elevation levels.[16] To better understand what this might look like in three-dimensions, they then glued different colored arrows on a grid of wires to represent measurements at each elevation and

---

[15] "Visualization: A Way to See the Unseen," in *America's Investment in the Future* (NSF, 2000), 92, https://www.nsf.gov/about/history/nsf0050/visualization/visualization.htm.
[16] P. S. Ray et al., "The Morphology of Several Tornadic Storms on 20 May 1977," *Journal of the Atmospheric Sciences* 38, no. 8 (August 1, 1981): 1643–63.

stacked the grids.[17] The resulting model revealed a coherent internal structure—something that nobody had been able to see before.

A Timely Interest in Visual Representation

Much of the turn towards visual representation came from an increased interest in and recognition of the role vision plays in cognition. There has long been a known link between seeing and understanding. When something is clear, we say that we understand it. What is opaque or hidden is synonymous with what is yet unknown. Philosophers, artists, and scientists from the ancient Egyptians and Greeks all the way up to the present have been interested in exploring the relationship between perception and cognition. In his book, *Information Visualization: Perception for Design* (2004), Colin Ware offers a biological explanation. The human visual system has a remarkable ability to take in large amounts of information and quickly recognize any patterns or trends. This is, on one hand, part of our brain's internal structure. Visual displays provide the highest bandwidth channel through which information can pass between environment and human. In fact, we take in more information through our sight than all of our other senses combined. Some twenty billion neurons—nearly half of the neurons in the human brain—are devoted to processing and analyzing visual information.[18] Still, taking in large amounts of information is cognitively taxing, and our

---

[17] Kaufmann and Smarr, *Supercomputing and the Transformation of Science*, 11.
[18] Colin Ware, *Information Visualization: Perception for Design*, 2nd ed., The Morgan Kaufmann Series in Interactive Technologies (San Francisco, CA: Elsevier, 2004), 2.

visual system has developed a short-cut to help. Many of the same neurons devoted to visual processing also structure a pattern-finding mechanism that allows us to discern likes from a group. At higher levels of visual processing, this pattern-finding mechanism and cognitive activity overlap. The result is that at the neuronal level, there are instances in which to see is to understand.

Importantly, the link between perception and cognition extends beyond the brain. Visual representations are a perfect example of external cognition, which posits that internal and external representations work together to formulate thought. By using graphical aids, humans can increase the ease and complexity of their thoughts. Take, for, example, the task of multiplying two two-digit numbers. The task is made substantially easier by performing the operation on pencil and paper. Using the paper to hold and carry the partial results until they are ready to be used frees the brain to focus on higher-levels of cognition. The implication is that people can think better with visual support. Full theories of external cognition were not made popular until the 1990s, after contemporary data visualization was born. Still, the understanding that external, particularly visual, representations could increase cognitive capacity had been gaining traction.

In 1969, German psychologist and art critic Rudolph Arnheim published his influential book *Visual Thinking*, in which he argued that visual perception and intellect are "indivisibly intertwined."[19] Without the stuff of the senses, Arnheim reasoned, the

---

[19] Arnheim's notion challenged existing models that distinguished between two different modes of cognition – thinking and perceiving, or as Arnheim refers to them *intellect* and *intuition*.

mind has nothing to think with. In other words, it is only because perception allows us to gather material for thought that we can even begin to think. He bases his argument on observations and examples, particularly those drawn from the art world. The ability to infer and draw meaning from an arrangement of shapes shows that thought is at work in vision itself. For example, recognizing a human figure in an abstract painting requires perception and symbolic thought. The lines and patches of color on the canvas are meaningless until someone perceives them. Without the material of perception, the mind has nothing with which to think. In other words, it is only because perception allows us to gather material for thought that we can even begin to cognize. Interestingly, this is a skill that is learned and becomes easier over time because we have accumulated more visual material through experience.[20] While Arnheim's argument is not specific to visualization, it does reflect a growing recognition of the relationship between vision and cognition in the years leading up to contemporary data visualization's birth. This recognition had also been gaining momentum in practice.

Visualization Precedents

While the problem of too much data might have been acutely felt during the years leading up to contemporary data visualization, it was certainly not new. Neither was the solution of visual representation. People have been visually representing data

---

[20] Many of Arnheim's examples rely on similar experiences of recognizing something that is or is not there. The tie between abstraction and thought is made primarily in the second half of the book. Chapters 8 and 9 define what abstraction is and is not. Chapter 14 explores the tie between art and thought. Chapter 5 discusses the relationship between the past and present.

for centuries. The first maps originated approximately eight thousand years ago. Ancient Egyptians used tables as early as the second century. Greek philosophers and scientists routinely drew diagrams to help them better understand celestial bodies. Arguably, the oldest cave paintings are the first examples of this practice.[21] However, it was not until the late eighteenth century that the graphical representation of quantitative information took hold. This milestone is largely attributed to the work of William Playfair who, over the course of a varied thirty-six-year career, developed and popularized many of the most common graphical forms still used today. His *Commercial and Political Atlas* (1786) introduced the first bar chart, as well as the first time-series line graph.

William Playfair

Playfair is often credited with the invention of visualization, or at least the graphical representation of data. His work marks the first in a long list of historical precedents that have retrospectively been associated with the contemporary data visualization field. In reality, Playfair was not the first to experiment with such graphical forms. Nor did the concept of visualization, as defined in this dissertation, exist yet. What has made Playfair's work and the work surrounding other visualization

[21] For a detailed account spanning the entire history of visual representations of data, see Michael Friendly and Daniel Denis' *Milestone Projects*: http://www.datavis.ca/milestones/index.php?page=milestones+project

precedents so noteworthy is the tendency to think of graphical representation as an entire mode of communication and analysis.

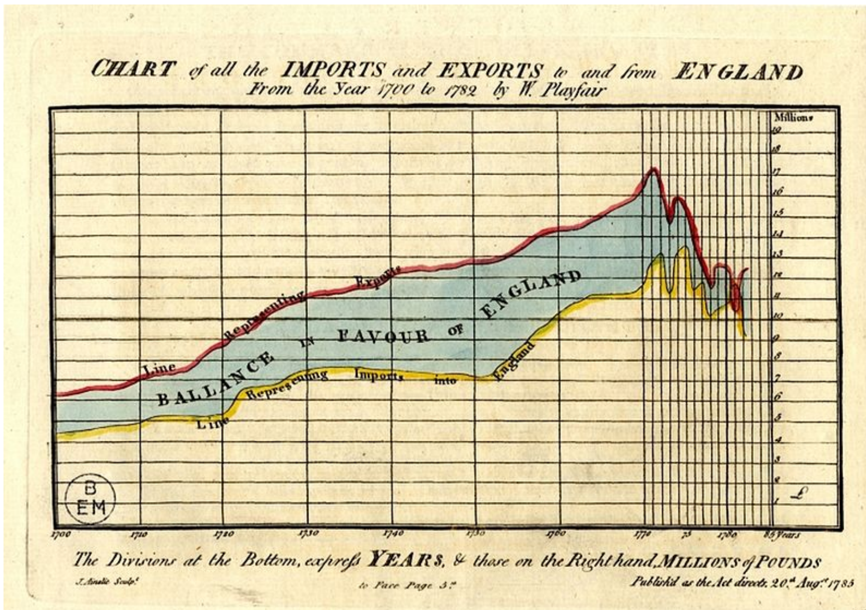

Figure 1.1. Opening chart from William Playfair's *Commercial and Political Atlas* (1786).

Figure 1.1 shows the line graph that opens *The Commercial and Political Atlas*. The chart provides an overview of the amount of all imports and exports to and from English between the years 1700 and 1782. Playfair's term for this type of representation is "linear arithmetic." That he sees his linear arithmetic as a novel and appealing form of communication is made apparent in the introductory text:

> That I have succeeded in proposing and putting in practice a new and useful mode of stating accounts, has been so generally acknowledged, that it remains only for me to request that those who do not, understand the manner of inspection the Charts, will read with attention the few lines of directions facing the first Chart, after which they will find all the difficulty entirely vanish, and as

> much information may be *obtained in five minutes as would require whole days to imprint on the memory, in a lasting manner, by a table of figures.*[22]

The description reads as insightful and modern even today, indicating two things: first, that Playfair had a good grasp of the relationship between perception and cognition long before any psychological or scientific studies were done; second, that the basic rationale behind the graphic representation of data—obtaining further insight—has not changed. What has changed is the type of data represented and the degree to which analysis or communication is prioritized.

Minard, Nightingale, and Snow

Other well-known precedents following Playfair include the work of Charles Joseph Minard, Florence Nightingale, and Dr. John Snow. Minard's famous map of Napoleon's 1812 campaign (1869) tells the story of the army's dwindling numbers as they marched into Russia during the winter months.[23] Florence Nightingale's rose diagrams (1858/59) similarly show the mortality rate of soldiers in army hospitals following the Crimean War, and Dr. John Snow's map of London's cholera epidemic (1854) is repeatedly held up in visualization literature as the epitome of gaining insight through graphic representation. The map purportedly helped solve the city-wide

---

[22] *Playfair's Commercial and Political Atlas and Statistical Breviary* (Cambridge University Press, 2005), xii.
[23] Arthur H. Robinson, "The Thematic Maps of Charles Joseph Minard," *Imago Mundi* 21 (January 1, 1967): 95–108.

epidemic by correctly identifying the Broad Street water pump as the contaminated water source.

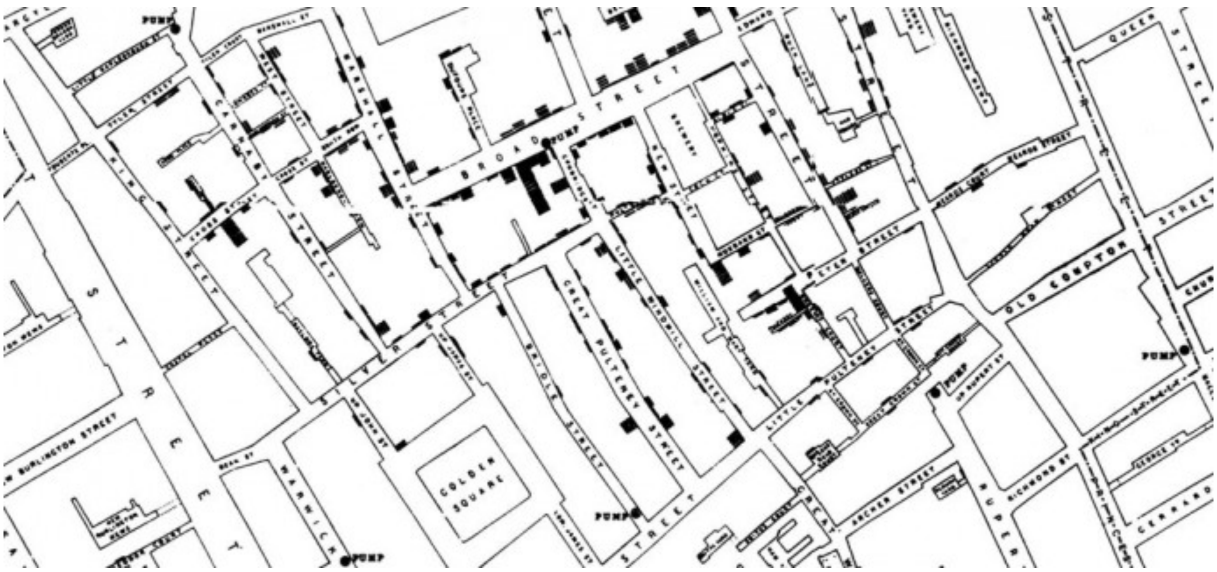

Figure 1.2. Detail from map of all Cholera cases as recorded by Dr. John Snow (1854).

Most histories of visualization will begin with Playfair and reference at least one of the nineteenth-century examples described above. The contemporary data visualization field clearly has origins in this early graphical work. However, it is important to remember that most if not all of what we know about visualization's history has been written in retrospect.[24] With this comes an inevitable degree of painting the past to fit the present. Visualization researcher Robert Kosara has a telling blog post in which he argues that the goal of visualization has shifted between analysis and

[24] This is true of any historical account, particularly one covers a rapidly evolving medium. One of the primary goals of this dissertation is to document and analyze the contemporary data visualization field as it continues to unfold.

presentation over time.[25] While many of the earliest visualization precedents have been painted as examples of visualization's analytic capacity, they were more geared towards presentation and communication at the time of their creation.

Dr. Snow's London Cholera Map (Figure 1.2) provides an illustrative case. The map, which plots the origin of cholera cases atop a street map of London, appears to solve the question of the epidemic's source. Many visualization histories tell the story of how Snow plotted the cases and thereby correctly identified the Broad Street water pump as the contaminated source. Looking at the map, this is a convincing story indeed. The problem is that Snow already knew the location of the pump. Details from his own reports show that he had already developed and tested his hypothesis prior to drawing the map. The map simply confirmed his findings. More importantly, it communicated the data and helped convince the appropriate officials to take action.[26]

Neurath and Bertin

The path towards contemporary data visualization continued to develop with an emphasis on presentation. The end of World War I brought significant social and political changes to Europe. Governments and other organizations suddenly found themselves in need of a quick and effective way to communicate relevant social and

---

[25] Robert Kosara, "The Changing Goals of Data Visualization," *Eagereyes* (blog), November 25, 2012, http://eagereyes.org/criticism/changing-goals-data-visualization.
[26] H Brody et al., "Map-Making and Myth-Making in Broad Street: The London Cholera Epidemic, 1854," *Lancet* 356, no. 9223 (July 1, 2000): 64–68, https://doi.org/10.1016/S0140-6736(00)02442-9.

economic information to their newly formed publics. It was with this intent that Austrian sociologist and life-long educator Otto Neurath designed a technique for visualizing statistics through pictorial means commonly known as Isotype (International System of TYpographic Picture Education).[27] Originally developed in Vienna in the 1920s, the technique made its first appearance at the Gesellschafts und Wirtschaftsmuseum (Museum of Society and Economy) where it was used to explain and illustrate social and economic matters to ordinary people.

Figure 1.3 shows one of the one hundred charts that was part of this collection. Titled ‘Development of Rubber Production Since 1985,’ it shows both the amount of wild and cultivated rubber (green) as well recycled rubber (orange) along with information as to where the different types of production take place. Even if one cannot read German, the subject reveals itself through the iconic units of tires. Each represents 25,0000 tons. Neurath and his team very much strove for this type of intuitive link. They saw what they were doing as creating a kind of visual Esperanto—a universally understood and agreed upon language for graphic communication.[28] Their work on Isotype continued, spreading to the Netherlands, Great Britain, the Soviet Union, and the United States among other locations. It seems to have only stopped in 1971, when Marie Neurath, Otto’s widowed wife, gave all archival and working materials to the University of Reading.

---

[27] Otto Neurath, *International Picture Language* (Reading: University of Reading, Reading & Language Information, 1981).
[28] Marie Neurath, “ISOTYPE,” *Instructional Science*, no. 3 (1974): 128–29.

By the 1960s, presentation and analysis began to take equal weight. Neurath's Isotype introduced the idea that some ways of graphically representing information might be better understood than others. French cartographer Jacques Bertin greatly furthered this line of thought in his 1967 publication of *Semiologie Graphique*. Bertin proposed that graphics is best approached as a system of signs and that, just like any other sign-system, it is subject to its own structural rules and regulations. He divides a given representation into two components: "the content (the INFORMATION to be transmitted)" and "the container (the PROPERTIES of the graphic system."[29]

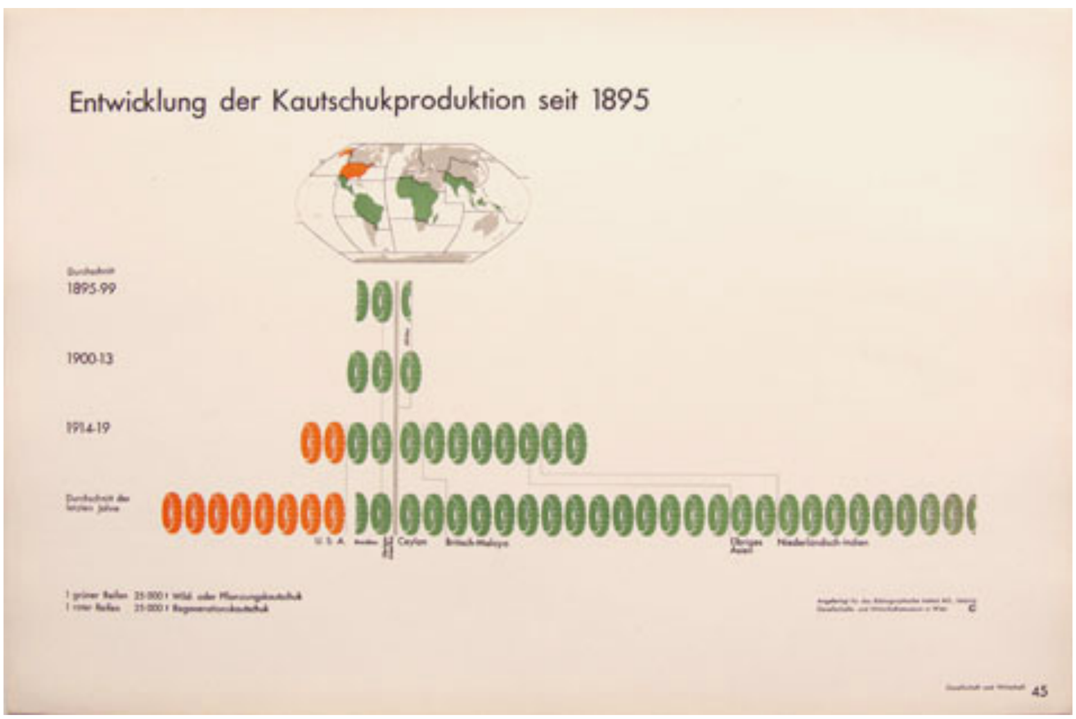

Figure 1.3. Otto Neurath, chart no. 45 from the *Gesellschaft und Wirtschaft* exhibition.

[29] Jacques Bertin, *Semiology of Graphics: Diagrams, Networks, Maps*, (ESRI Press, 2010), 5.

While the information to be transmitted remains constant, the properties of the graphic system can change. Bertin identifies these properties early on in the text as “visual variables” of which there are eight: the two dimensions of the plane (*x* and *y*); size; value; texture; color; orientation; and shape.[30] The book is filled with examples of the wide variety of ways in which the eight variables can be combined. In taking a semiological approach, he argues that some arrangements are better than others. What makes a particular design better, or "more compelling" in Bertin's words, is its relative efficiency. He defines efficiency as follows:

> Efficiency is defined by the following proposition: If, in order to obtain a correct and complete answer to a given question, all other things being equal, one construction requires a shorter period of perception than another construction, we can say that it is more *efficient* for this question.[31]

Efficiency, then, is a measure of being more attuned to human perception. While Playfair and Neurath, among many others, noted this advantage of graphical representation, Bertin was the first to systematically study it.[32] His *Semiologie Graphique* provided the first official theory linking graphic representation and visual

---

[30] Bertin, *Semiology of Graphics,* 7.
[31] Bertin, *Semiology of Graphics*, 9.
[32] In his introduction to the 1983 edition, Howard Wainer calls Bertin’s *Semiologie Graphique* “the most important work on graphics since the publication of Playfair’s *Atlas*” (ix). Interestingly, the book was not available in English until more than ten years following its publication. Its translation ultimately began in 1977 as part of the NSF-supported “Graphic Social Reporting Project.”

perception. Interestingly, as contemporary data visualization has grown into a field, no proposed theory has supplanted it.

Tukey and Tufte

The build-up to contemporary data visualization and the graphic revolution of which it is only part continued to waver between analysis and presentation. Efficiency came to be held up as the golden standard of functionality, and the idea that graphical representation could be used to analyze and validate large quantities of data took hold. The 1940s and '50s pushed computing and the analysis of large amounts of information to new levels of import. Much of the impetus for building large, automated computers like Colossus and Enigma came from the need to decipher high-level military intelligence during World War II.[33] Their success demonstrated the power of analytics. Large datasets similarly invigorated the development of statistics.

In 1977, John Tukey introduced an approach to analyzing large collections of data that changed the field. Exploratory data analysis (EDA) is an approach to statistics that employs a variety of techniques to describe and summarize a dataset's main characteristics, often by visual means.[34] It is more of a philosophy than anything else. Tukey argued that statisticians spent too much time testing hypotheses. Rather than making assumptions about what kind of model would best fit the data, they

---

[33] For more information on Colussus and Enigma as well as the entire history of computing, see Nicholas Metropolis, ed., History of Computing in the Twentieth Century (New York: Academic Press, 1980).
[34] John W. Tukey, *Exploratory Data Analysis* (Pearson, 1977).

should allow the data to speak for itself. In other words, by exploring the data, they could more accurately infer its internal structure. The majority of EDA's techniques for doing this are graphical in nature. Common techniques include:

Box Plot
Scatter Plot
Histogram
Multi-vari Chart
Stem-and-leaf Plot
Parallel Coordinates
Principle Component Analysis

The reason for the heavy reliance on graphics is clear to see in the context of this chapter. In harnessing our natural ability to find and recognize patterns, graphical representations of data get to the heart of EDA's proposition—they allow us to see what may not have otherwise be seen. Indeed, Tukey is often quoted as saying, "The greatest value of a graph is when it forces us to see what we never expected."[35] The championing of EDA led to an increased recognition of the power of graphical analysis, particularly in relation to large datasets. EDA also encouraged the development of statistical computing packages like S and R.[36] The belief that computation could be combined with graphics to "let the data speak for itself" became ingrained in information practices and culture. What resulted was a pragmatic, efficiency-based mindset in terms of visualizing patterns in data.

---

[35] Tukey, *Exploratory Data Analysis*, vi.
[36] R. A. Becker and J. M. Chambers, *S: An Interactive Environment for Data Analysis and Graphics* (Belmont, CA: Chapman and Hall/CRC, 1984); "R: The R Project for Statistical Computing," accessed June 3, 2016, https://www.r-project.org/.

The tie to visualization, and what perhaps most clearly set contemporary data visualization on the path towards this overarching mindset, came with Edward Tufte. If Tukey pushed the goal of visualization towards analysis, then Tufte ensured that presentation was purely analytic in its intent. The ideas outlined in his book *The Visual Display of Quantitative Information* (1983) introduced concepts such as graphical excellence and graphic integrity.[37] The book also presented the first in-depth evaluation of graphical representations over time. Tufte insists that there is a distinction between "good" and "bad" graphics. Good graphics tell the truth about the data and most efficiently communicate complex quantitative ideas.[38] Bad graphics, which he characterizes as full of *chartjunk* and inessential markings, do not—communicate efficiently, that is. His measurement of what he calls the "*data-ink ratio*" is a condensed example of the type of design that he promotes. Data-ink refers to "the non-erasable core of a graphic, the non-redundant ink arranged in response to variation in the numbers."[39] The data-ink ratio is then the fraction of data-ink to the total ink used to print graphic. In Tufte's own words:

> Above all else show the data.
> Maximize the data-ink ratio.
> Erase non-data-ink.
> Erase redundant-ink.
> Revise and edit.[40]

---

[37] Edward R. Tufte, *The Visual Display of Quantitative Information* (Graphics Press, 1983).
[38] Tufte, *The Visual Display of Quantitative Information*, chapters 1–2.
[39] Tufte, *The Visual Display of Quantitative Information*, 93.
[40] Tufte, *The Visual Display of Quantitative Information*, 105.

Anything else is superfluous and thus merely decorative. Tufte's ideas gained traction throughout the 1980s and 1990s. He later expanded upon his design recommendations and principles in several other books.[41] It is beyond the scope of this chapter to speculate why this might be the case, but what is clear, as far as visualization precedents are concerned, is that Tufte largely defined the expectations of graphical representations of data in the years immediately preceding the birth of contemporary data visualization. This was an expectation based on analysis, efficiency, and a sought-after objective truth at the expense of all else.

The Development of Computer Graphics

There is one other influence to consider when laying out the context of the years leading up to contemporary data visualization's birth, and that is computer graphics. Contemporary data visualization would not have developed at the time that it did and with the energy that it did if it were not for the related development in the field of computer graphics. Computer graphics involves the drawing of pictures with computers. Like visualization, it is an interdisciplinary field with origins in both art and science. And like visualization, there is no single event to point to in sketching its origins.

---

[41] Edward R. Tufte, *Envisioning Information* (Graphics Press 1990); Edward R. Tufte, *Visual Explanations: Images and Quantities, Evidence and Narrative* (Graphics Press, 1997); Edward R. Tufte, *Beautiful Evidence* (Graphics Press, 2006).

Infancy

The field's beginnings date back to the 1950s when several unrelated research projects opened up the possibility of drawing with computers.[42] Naval research projects Whirlpool and SAGE (Semi-Automatic Ground Environment) had demonstrated that cathode-ray tubes could be used as viable graphic displays and pioneered digital computers with interactive interface tools. The TX-2 computer (1959), for example, had a nine-inch display, light-pen, and panel of switches on which the first computer graphics system was based.[43] Around the same time, artists began experimenting with using computers to make art. Iowa-based artist and mathematician Ben Laposky famously used analog computers to create oscilloscope artwork as early as 1953.[44] His so-called "electronic abstractions" fall under the category of early Computer Art and represent some of the first instances combining programming with an open-ended variety of form.[45] Michael Noll and others at Bell Labs did similarly pioneering work

[42] See Wayne Carlson, "A Critical History of Computer Graphics and Animation," The Ohio State University, accessed October 9, 2015, http://design.osu.edu/carlson/history/lessons.html; Terrance Masson, "History of Computer Graphics (CG)," The Computer Graphics Book Of Knowledge, accessed October 9, 2015, http://www.cs.cmu.edu/~ph/nyit/masson/history.htm. Not surprisingly, many of the most comprehensive resources on the history of computer graphics exist as websites, not academic journals.
[43] Wayne Carlson, "Section 2: The Emergence of Computer Graphics," A Critical History of Computer Graphics and Animation, accessed November 3, 2015.
[44] Ben F. Laposky, "Oscillons: Electronic Abstractions," *Leonardo* 2, no. 4 (October 1, 1969): 345–54.
[45] Laposky, 347–48.

using computers to generate graphic designs.[46] It was not long before artists and scientists, sometimes in collaboration, were routinely drawing using code.

In 1960, William Fetter coined the term "computer graphics" to describe the type of work he was doing at Boeing.[47] Fetter was charged with exploring new graphic techniques that could lead to the development of computer drawings to support airplane design. He, along with a team of other designers and engineers, developed a way to visualize the human figure in the cockpit and moving through the aisles, which led to improved exactitude and flexibility in predicting the success of a given design.[48] Computer graphics quickly became a catch-all term used to describe any type of computer-aided or generated design. The field grew as an experimental area of interest with obvious artistic, scientific, and commercial applications.

Many of the early advances in computer graphics came out of university laboratories. In 1961, MIT graduate student Steve Russell created the first popular computer game (*Spacewar*).[49] In 1963, fellow graduate student Ivan Southerland

---

[46] A. M. Noll, "The Digital Computer as a Creative Medium," *IEEE Spectrum* 4, no. 10 (October 1967): 89–95; Frank Dietrich, "Visual Intelligence: The First Decade of Computer Art (1965-1975)," *Leonardo* 19, no. 2 (1986): 159–69.

[47] Fetter is conventionally credited with coining the term within the Computer Graphics literature; however, he admits that it might have been his supervisor Verne L. Hudson.

[48] W. A. Fetter, "A Progression of Human Figures Simulated by Computer Graphics," *IEEE Computer Graphics and Applications* 2, no. 9 (November 1982): 10.

[49] J. Martin Graetz, "The Origin of Spacewar," Creative Computing 7, no. 8 (August 1981): 56–67. Spacewar ran on a PDP-1 computer, which was donated to MIT by Digital Equipment Corporation (DEC) in hopes that the university could do something remarkable with their product. The ultimate goal was to increase its commercial viability and appeal. For more information on DEC and the PDP-1's functionality, see Carlson, "Section 3: The Industry Evolves."

introduced the first computer drawing program (*Sketchpad*).[50] *Sketchpad* ran on the same type of TX-2 computer previously used for SAGE. It was the first-ever program to utilize a graphical user interface (GUI), allowing people to interact with the computer through images rather than text or code. Using a light pen and a few simple controls, users could dynamically draw and manipulate shapes on screen. This not only demonstrated the artistic and technical power of the computer graphics, it opened up a new form of human-computer interaction. [51] *Sketchpad* and the ideas it propagated soon became an inspiration to anybody working in the computer graphics field. Researchers at university laboratories and computational research facilities including Harvard, the University of Utah, Ohio State University, and Bell Labs began to investigate techniques that could harness the power of the computer and its display, in combination with simple interaction devices to achieve artistic and technical feats.[52]

Maturation

By the mid-1970s, computer graphics had matured as a discipline, and a corresponding industry had begun to emerge. The federal government and companies

---

[50] Ivan E. Sutherland, "Sketch Pad a Man-Machine Graphical Communication System," in *Proceedings of the SHARE Design Automation Workshop*, DAC '64 (New York, NY, USA: ACM, 1964), 6.329–6.346.

[51] An important aspect of Sutherland's program that I have not emphasized here is that it opened up the possibility of using computers to those who did not know how to code. GUI interfaces drastically reduce the amount of technical knowledge required by substituting visual metaphors for programming commands. This reduction has been instrumental in defining the development of personal computing as well as HCI.

[52] For a summary of the work performed at different labs see Wayne Carlson, "Section 4: Research Moves the Industry," A Critical History of Computer Graphics and Animation.

interested in further developing computer graphics technology backed research centers solely focused on computer graphics technology.[53] The commercial potential of computer graphics was obvious to researchers and entrepreneurs alike. A number of first-generation companies that sought to bring computer graphics to entertainment (e.g., gaming, television, advertising, and film) also emerged. Perhaps the most famous of these is Pixar, which grew out Lucasfilm's Computer Division. A decade later, four of these companies, Digital Effects, Abel, MAGI, and Information International Inc., would team up to produce the 235 scenes of computer-generated imagery for *TRON* (1982).[54]

The 1980s brought the establishment of even more production companies, and the use of computer-generated imagery in television and film rapidly increased.[55] While the Hollywood world of CGI focused on entertainment, the technical development of computer graphics hardware and software continued to change the way both artists and researchers went about their work. Many of the same expectations that drove supercomputing and the adoption of a general computing culture—namely speed and capacity—pushed the computer graphics field. By the early 1980s, graphical interaction with computers was standard across scientific, artistic, and even personal computing use.

---

[53] Carlson, "Section 5: University Research Activities," A Critical History of Computer Graphics and Animation.
[54] Carlson, "Section 14: CGI in the Movies," A Critical History of Computer Graphics and Animation.
[55] Carlson, "Section 11: Production Companies," A Critical History of Computer Graphics and Animation.

Still, researchers and designers often had to wait long periods of time to see their work. Complex visuals, like the digital rendering of clouds used in both the creative and scientific application of computer graphics, rely on comprehensive mathematical models that require a large amount of computing power. It took even the most advanced computers a significant amount of time to process the number of calculations necessary to display images onscreen. Demand for decreased rendering time and the introduction of graphic co-processors eventually led to stand-alone workstations that could process high-quality displays in far less time. Silicon Graphics, Inc. and Sun Microsystems are just two of the major companies that formed around this trend.[56] While workstations provided designers and engineers with a near real-time front-end window, supercomputers were still often needed to process the data and generate image frames. Production studios, which did not have the budget to buy their own supercomputers, began buying time on some of the supercomputers at research centers and universities.[57]The result was cross-pollinated environment in which artists and scientists were working with the same machines in the same places. The official birth of contemporary data visualization was not far behind.

---

[56] One of the most important contributions in the area of graphics display hardware was the Geometry Engine, envisioned by Jim Clark at Stanford University. His idea was to create a collection of components in a very-large-scale-integration (VSLI) processor that perform all of the basics in the first stage of image generation based on geometric models. Having tried to shop his design around at various computer companies without any takers, he and some of his colleagues left to form their own company, Silicon Graphics, Inc. (1982). For specifics on the Geometry Engine, see Clark, "The Geometry Engine: A VLSI Geometry System for Graphics." Computer Graphics (SIGGRAPH 82 Proceedings) 16(3) July 1982.

[57] Thomas A. DeFanti, Larry Smarr, and Thomas DeFanti, Informal Interview with Thomas DeFanti and Larry Smarr at UCSD, September 18, 2013.

SIGGRAPH

Perhaps the most direct influence on contemporary data visualization's birth stemming from the field of computer graphics is SIGGRAPH. SIGGRAPH, or the Association for Computing Machinery's (ACM) Special Interest Group on Computer Graphics and Interactive Techniques, is an interdisciplinary organization dedicated to sharing and promoting the latest computer graphics research.[58] It emerged in the late 1960s when a group of members of the ACM, one of the two global societies concerned with scientific and educational computer research, took an active interest in the burgeoning computer graphics field. The group received an official designation in 1969 and has since grown. What was a small interest group that circulated a quarterly newspaper has become the largest arena for developing and displaying cutting-edge graphics research. The organization now runs two yearly conferences, puts out multiple publications, and serves as an umbrella association bringing together industry and academic leaders interested in furthering the possibilities within computer graphics. Herein lies the tie to contemporary data visualization.

One of the defining characteristics of SIGGRAPH is and always has been its wide scope. Conferences and journals tend to bring together researchers from across a vast range of interests and disciplinary work. Hollywood CGI companies show off their latest work in combination with media artists, engineers, and scientists doing innovative research. In the 1980s, this meant that SIGGRAPH offered the means and

---

[58] http://www.siggraph.org

encouragement for scientists who were beginning to experiment with visually representing data and computer graphics designers to not only use the same technology but also to see each other's work and comunicate. In other words, SIGGRAPH facilitated the same type of cross-pollinated environment that was being created by production companies buying time at supercomputing centers, with the added benefit of display. The cross-pollination resulted in hybrid work. In 1984, Larry Yaeger and Craig Upson at Digital Productions combined computational fluid dynamics with CGI to create a realistic simulation planet Jupiter for the movie *2010* (1984). Their presentation on the simulation at SIGGRAPH 1986 generated a great deal of excitement. The Jupiter video marked the first time the latest techniques in computer graphics were combined with scientific data and research. Scientists struggling with how to best understand and communicate their data suddenly saw the full potential of visually representing their research. The chain of events that followed gave name to and introduced the contemporary data visualization field.

Birth and Growth: 1987 – 1990s

Contemporary data visualization was born at precisely the time when the stories of supercomputing, increased interest in visual representation, and the development of computer graphics found significant overlap. SIGGRAPH, the need to visualize large amounts of scientific data, and the general fascination with graphic representation all added up to an environment in which something like visualization was ripe to emerge.

When the NSF called the panel described at this chapter's beginning in October 1986, there was arguably an already emergent field. Researchers from across industry, academia, medicine, and the government had already begun to experiment with using visual means to comprehend the large amounts of data generated by supercomputers and similarly data-intensive resources. Projects like the Jupiter simulation and the early experiments visualizing data at supercomputing centers are evidence of this. The *Visualization in Scientific Computing* (ViSC) Report summarizing the panel's findings merely provided the officiating stamp.

The ViSC Report

When the ViSC Report first came out in July 1987, it marked the legitimate birth of the contemporary data visualization field. The fourteen-page document, accompanied by a lengthy appendix, was actually published twice—once in July, following the NSF-sponsored workshop on Visualization in Scientific Computing, and again in November, as a special issue of SIGGRAPH's *Computer Graphics* quarterly.[59] This second edition was accompanied by two hours of videotape published as issues 28 and 29 of the *SIGGRAPH Video Review*.[60] While specific to visualization in scientific computing, the text and the visuals propelled the development of an entire visualization field. Visualization's potential for art, science, and the future of human-

[59] Defanti, Brown, and McCormick, "Visualization in Scientific Computing."
[60] Thomas A. Defanti, Maxine D. Brown, and Marshall C. Yovits, *Visualization in Scientific Computing: Domain and Visualization in Scientific Computing: Systems*, vol. 28/29, SIGGRAPH Video Review (Elsevier, 1987).

computer interaction are all highlighted. Part of what made the ViSC Report so influential is that it gave visualization a name. Prior to this point, researchers and technicians across various disciplinary boundaries had different names for what they called the practice of graphically representing data. In defining these practices under the umbrella of "visualization," the ViSC Report established a coherent identity and base.

The character of visualization as presented in the report largely determined the trajectory that contemporary data visualization would take. This character can best be summarized as methodic, pioneering, and effectiveness-based. The majority of the report's content is specific to science and engineering. As such, visualization adopts a character appropriate to its context. Broadly defined in the first few paragraphs as "a method of computing," visualization is more specifically painted as paving a path towards scientific insight.[61] One oft-quoted excerpt from the text reflects this line of thought:

> Richard Hamming observed many years ago that "The purpose of [scientific] computing is insight, not numbers." The goal of visualization is to leverage existing scientific methods by providing new scientific insight through visual methods.[62]

The tie made between visualization and scientific insight is resolute. Visualization is hailed as having potential applications in medical imaging, space exploration, and

---

[61] Defanti, Brown, and McCormick, "Visualization in Scientific Computing," 3.
[62] *Ibid.*, 3.

computational fluid dynamics among a slew of other areas that form an impressive list. Over twenty different opportunities for scientific and visualization research are laid out in detail in an appendix to the report.[63] It is almost as if visualization is synonymous with scientific progress. The report even goes as far to claim that visualization in scientific computing could have an effect on scientific productivity and the potential for breakthroughs "at a level of influence comparable to supercomputers themselves."[64]

Science, however, was not the only area of application mentioned in the ViSC Report. Recall that at this time Hollywood production companies were renting time on supercomputers to process all of the calculations required by the latest CGI. Tufte, among others, had shown the analytic and communicative power of graphical display. Two separate sub-sections, "Visualization and Society" and "Visualization and Interdisciplinary Teams," discuss visualization's artistic, communicative, and commercial appeal.[65] The potential to affect development across disciplines combined with the interdisciplinary collaboration developing the visualization field would take, certainly added to the report's significance and pioneering feel. Indeed, the high level of energy and forward-thinking mentality surrounding the report cannot be overstated. Tom DeFanti, one of the three editors of the report, recalls the excitement leading up

[63] Appendix A, titled "Scientific and Engineering Opportunities," clearly identifies the range of potential applications visualization might take. Each application is treated individually. The authors' make a specific argument for how visualization could help the said of research. By including this level of detail in the appendices, they strengthen their case for increased funding without bogging down the main body of the report with field-specific language that may detract from the wide-spread necessity of their cause.

[64] Defanti, Brown, and McCormick, "Visualization in Scientific Computing," vii.

[65] *Ibid.*, 6, 11.

to its distribution.[66] Contributors were up-all-night photocopying. Eager to reach a wider audience, they saw the ViSC Report not only as their chance to convince NSF administrators of the need to invest in visualization research, but also to communicate the field's potential to other artists and scientists around the world. The United States was in a unique position because of the extreme amount of money and personnel available in the movie industry.[67] It was not until people saw what could be done with the highest quality graphic machines and people trained in visual representation that visualization gained the attention of a wider, even global crowd.[68]

The argument worked. Researchers at supercomputing centers received increased federal funding "to put visualization tools into the hands and minds of scientists."[69] Better graphics resources led to increased collaboration and the development of visualization research. Graphics studios began working more closely with supercomputing centers, and soon, dedicated visualization groups within the centers formed to further develop visualization techniques. Some administrators even convinced Hollywood designers, like Craig Upson of the Jupiter simulation fame, to come work full-time with scientists at supercomputing centers. Interest in the field quickly grew. Visualization became an innovative topic in peer-reviewed journals and other publications. A special issue of the IEEE's flagship publication *Computer* was

---

[66] Thomas DeFanti, Informal Interview with Thomas DeFanti and Larry Smarr at UCSD, September 18, 2013.
[67] Larry Smarr, Ibid.
[68] Thomas DeFanti, Ibid.
[69] Defanti, Brown, and Yovits, *Visualization in Scientific Computing: Domain and Visualization in Scientific Computing: Systems*, 28/29.

dedicated to the emergent field in 1989.[70] The annual visualization conference, IEEE Vis, which is still in existence today, was established in 1990. SIGGRAPH and other conferences also helped to develop the field worldwide.

What should be obvious at this point is that contemporary data visualization emerged as a field with multiple disciplinary origins and applications. Nevertheless, certain areas did characterize its early development. Scientific visualization and information visualization arose as two subfields or "genres" of visualization that dominated the discussion and research in the field throughout the 1980s and 1990s.

Scientific Visualization

Scientific visualization is the interactive visualization of scientific data. What distinguishes it from other types of visualization is a matter of debate.[71] The genre's emergence coincided with contemporary data visualization's birth and primarily grew out of the supercomputing center infrastructure set up by the ViSC Report. The genre is still one of the main types of visualization that is prevalent today. In a 2000 Vision special issue of IEEE *Computer Graphics and Applications*, visualization researcher Jock Mackinlay stated that "scientific visualization…focuses on physical data such as the human body, the earth, molecules, and so on."[72] While the distinction was not

---

[70] IEEE Society, Computer 22, no. 8 (August 1989).
[71] Theresa-Marie Rhyne, "Panel: 'Information and Scientific Visualization: Separate but Equal or Happy Together at Last'" (Panel at IEEE Visualization 2003, Seattle, Washington, 2003), 611–14.
[72] J. Mackinlay, "Opportunities for Information Visualization," *IEEE Computer Graphics and Applications* 20, no. 1 (2000): 22.

necessarily defined in the late 1980s when the genre first appeared, it has retrospectively been used to characterize scientific visualization from the late 1980s up until to the present. The real significance of the physical tie is in its reference to physical phenomena that exist in world. In visualizing physical data, scientific visualizations almost always give shape to phenomenon that might otherwise be unobservable or difficult to see (e.g., the surface of Jupiter or the interaction between two molecules).[73] The other characteristic used to distinguish scientific visualizations is interactivity. Interactive, here, refers to the process of inquiry. Researchers will often use visualizations to explore simulated environments or guide calculations. The ability to interact with the represented data is part of the scientific process through which models and the data can be improved or better comprehended. As with any type of categorization, the genre also projects a distinct style. The typical scientific visualization is often designed to be viewed by experts, exceedingly detailed, and focused on the direct presentation of data, sometimes at the expense of legibility and aesthetics.

Robert Wilhelmson's simulation of a numerically modeled severe storm is a classic example. In fact, the video was shown so widely and with such frequency when

[73] If we follow this characteristic through to its conclusion, we could consider medical imaging, microscopy, and even remote photography (e.g. the Mars Rover) as part of scientific visualization. There is a very fine line between what some might consider digital imaging and scientific visualization. Pixels are, after all, data. The difference seems to lie in the intent of the visualization.

first released that it came to be known as the “teapot of scientific visualization.”[74] In 1990, Wilhelmson and a group of other researchers at the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Champagne-Urbana built upon previous models of storm systems to create a short animation that showed the progressive development of a storm.[75] What was so exciting about the video was that it allowed scientists to see the storm’s development in four dimensions (three-dimensions plus time) at an exceptional level of detail. The video, which lasts about three minutes, traces the modeled depiction of a storm from beginning to end. What begins as a single cumulous cloud quickly escalates into a severe storm system with high-speed winds and rain. Narration, combined with a transparent external structure, reveal details about velocity, air pressure, and temperature, among other variables, along with the structure of what is going on inside.[76]

Fig. 1.4 shows a still from a section of the video in which tracers are released. The time in the lower right is given in hours, minutes, and seconds and corresponds to the time elapsed during the storm. Dots, color, and opacity are used to represent different values of the variables shown. In this case, dot color is an indication of draft. Those dots colored orange-red were moving up at the particular time shown. Those

---

[74] M. Pauline Baker and Colleen Bushell, “After the Storm: Considerations for Information Visualization,” *IEEE Computer Graphics and Applications* 15, no. 3, 1995.
[75] Robert B. Wilhelmson et al., “A Study of the Evolution of a Numerically Modeled Severe Storm,” *International Journal of High Performance Computing Applications* 4, no. 2 (June 1, 1990): 20–36.
[76] A copy of the video as well as sample images can be found at the webpage “Study of a Numerically Modeled Severe Storm Sample Images,” accessed March 10, 2016, http://redrock.ncsa.illinois.edu/image_89video.html

colored blue were moving down.[77] Other sections of the video show individual slices from inside of the cloud, providing and even further detailed depiction of the storm's internal dynamics. The simulation was run on a CRAY supercomputer and initialized using storm data gathered from a particularly severe storm that struck Texas and Oklahoma on April 3, 1964.

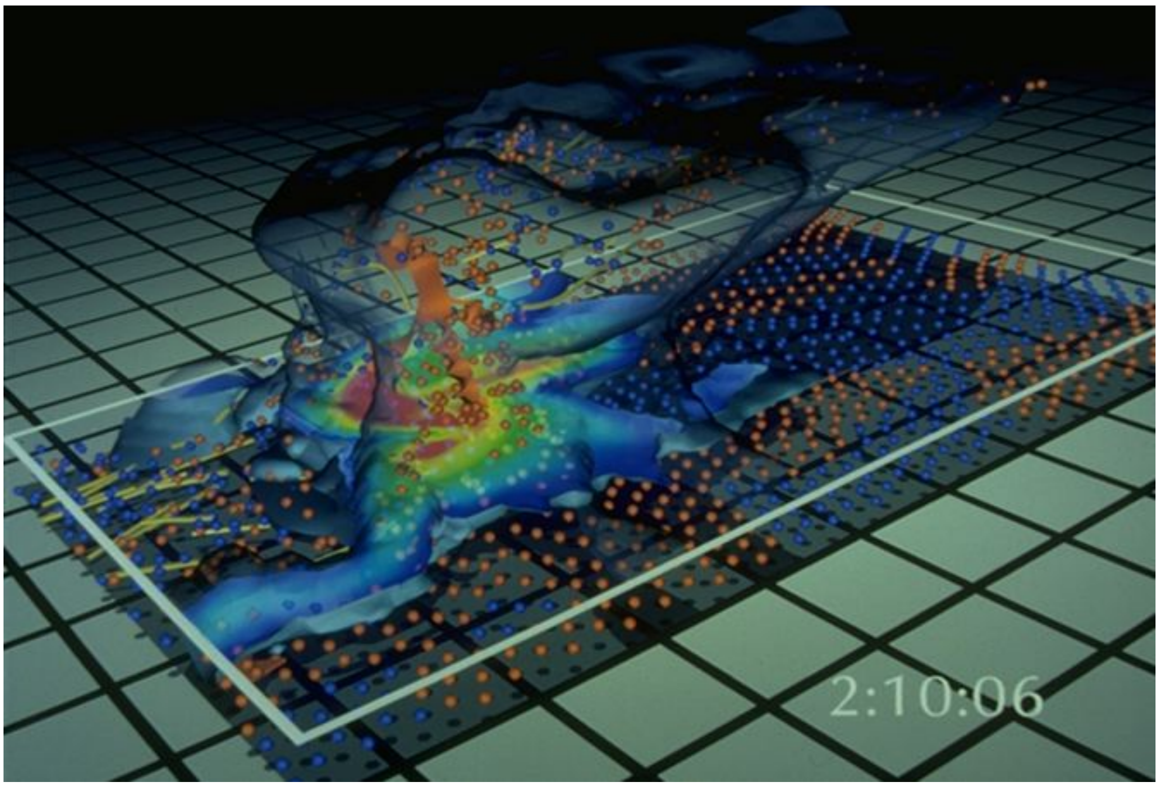

Figure 1.4. A video snapshot of storm flow from Wilhelmson et al's animation (1990).

Not surprisingly, it bears a remarkable similarity to the experimental storm-mapping detailed earlier in this chapter. Both projects have some of the same

[77] I am deducing the analysis from the caption provided for Figure 6 in the Wilhelmson et at. 1990 paper.

researchers on the team. And both represent the cumulative work of many different researchers and many different calculations over the years. This reveals an additional characteristic of scientific visualizations that might not be evident from just looking at the final representation—they are often collaborations over the long-term. The 1990 model offers an additionally interesting example in that it was later redesigned by Edward Tufte and a group of NCSA visualization professionals. The team studied the well-known example of scientific visualization, identifying scenes where principles of design, perception, and cognition suggested improvements.[78]

Information Visualization

The other genre of visualization to emerge following the ViSC Report was information visualization, or, as it is more commonly known, infovis. Infovis developed throughout the 1990s as a broadly applicable and increasingly popular subfield focused on the visual comprehension of all different types of data. In the introduction to the canonical volume *Readings in Information Visualization: Using Vision to Think*, editors Card, Mackinlay and Schneiderman define information visualization as: "The use of computer-supported, interactive, visual representations of **abstract** data to amplify cognition."[79]

[78] Baker and Bushell, "After the Storm."
[79] *Readings in Information Visualization: Using Vision to Think* (Morgan Kaufmann, 1999), 7.

The bolded term here carries weight. Information visualization is often distinguished from other types of data visualization based on the type of data represented. Whereas scientific visualization primarily deals with data that has a spatial or physical existence in the world, information visualization does not. Instead, it deals with data that is not concrete—in other words, abstract. This distinction is now somewhat obsolete. Pervasive use of embedded and geo-location technologies has significantly blurred the physical-digital divide. It is becoming increasingly difficult to discern between the two.[80] RFID tags and QR codes, for example, contain both physical and non-physical data. There is an ongoing debate within the visualization community as to what extent scientific visualization and information visualization should remain separate.[81] However, in the early 1990s, this is the distinction that was made.

Information visualization emerged as a genre at a very particular time in computing history when a range of technical and cultural developments opened the world of data up to non-scientific and non-technical realms. The introduction of the World Wide Web, widespread adoption of personal computing, and a general fascination with computing technology produced an environment in which a variety of

[80] Malcolm McCullough, *Digital Ground: Architecture, Pervasive Computing, and Environmental Knowing*, New Ed edition (Cambridge, Mass.: The MIT Press, 2005); Mark Shepard, ed., *Sentient City: Ubiquitous Computing, Architecture, and the Future of Urban Space* (New York City: Cambridge, MA: The MIT Press, 2011).
[81] Theresa-Marie Rhyne, "Panel: 'Information and Scientific Visualization: Separate but Equal or Happy Together at Last'"; Theresa-Marie. Rhyne, ed., "Does the Difference between Information and Scientific Visualization Really Matter?," *IEEE Computer Graphics and Applications* 23, no. 3 (2003): 6–8.

different types data started to amass. Financial data, demographic data, data associated with collections of documents—all of these different types of data also needed to be understood. Researchers in a broad array of information-heavy fields such as economics, Human Computer Interaction (HCI), and information science began to experiment with visualization techniques as a way to help. By 1995, the IEEE Visualization Conference had split into separate SciVis and InfoVis events. The first IEEE InfoVis Symposia opened with the definitively broadened scope of “All aspects of visualizing information (on and off the Internet and the Web) and how users interact with digital information.”[82]

The inclusion of different types of data and emphasis on how *users* interact with information gave infovis a slightly different character. The genre developed with significant influence from cognitive science and perceptual psychology, and as a result, is far more concerned with the human process of visually comprehending information. Information visualizations tend to be more design-focused, user-centered, and appeal to a more general audience in comparison to scientific visualizations. The first known use of the term already set this standard. Information Visualization was first introduced as a research domain which “uses 2D and 3D environments to represent the structure of information” in a 1989 article by researchers at Xerox PARC.[83] The

---

[82] “The First Information Visualization Symposium (InfoVis ’95),” accessed October 30, 2016, http://www.infovis.org/infovis/1995/.

[83] G. Robertson, S. K. Card, and J. D. Mackinlay, “The Cognitive Coprocessor Architecture for Interactive User Interfaces,” in *Proceedings of the 2nd Annual ACM SIGGRAPH Symposium on User Interface Software and Technology*, UIST ’89 (New York, NY, USA: ACM, 1989), 10–18.

term was conceptually expanded upon as the *Information Visualizer*, a prototype workspace that allows computer users to easily navigate and retrieve information by interacting with objects in an extended desktop metaphor.[84] Emphasis was placed on giving users an overview of the information structure and decreasing the amount of time necessary to retrieve information. The result is a user interface that builds knowledge through interactivity and, importantly, does so efficiently.

Interactivity and efficiency are often held up as early ideals of the genre. This is true regardless of the type of information being visualized or the technique used to visualize it. Stuart Card identifies four levels at which information visualization can function: (1) the infosphere; (2) the workspace; (3) sense-making tools; and (4) the document or visually enhanced object.[85] While the categories are not necessarily separate, they do offer an indication of just how varied information can be. The *SeeSoft* ® system offers an example of what is known as a sense-making tool in that it helps users to "make sense" of the data in a given collection, usually by encoding and representing information.[86] *SeeSoft* ® allows users to analyze up to 50,000 lines of

---

84 Stuart K. Card, George G. Robertson, and Jock D. Mackinlay, "The Information Visualizer, an Information Workspace," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91 (New York, NY, USA: ACM, 1991), 181–186.

85 Stuart K. Card, "Visualizing Retrieved Information: A Survey," *IEEE Comput. Graph. Appl.* 16, no. 2 (March 1996): 63–67.

86 Daniel M. Russell et al., "The Cost Structure of Sensemaking," in *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems* (ACM, 1993), 269–276, E.H. Chi and S.K. Card, "Sensemaking of Evolving Web Sites Using Visualization Spreadsheets" (IEEE Computer Society, 1999), 18–25, https://doi.org/10.1109/INFVIS.1999.801853.

computer code at once by mapping each line onto a document column.[87] In the figure below (Figure 1.5), colors represent how recently a line was modified. Each of the fourteen columns corresponds to a different file or document. Brushing over the recency scale highlights corresponding lines in the display. A separate inset window can also be called up to display the code and statistics associated with a selected segment of the display. The effect is to provide an overview of the patterns in the data while still maintaining a fine level of detail. New patterns can potentially be discovered as the user continues to reconfigure and interact with the representation.

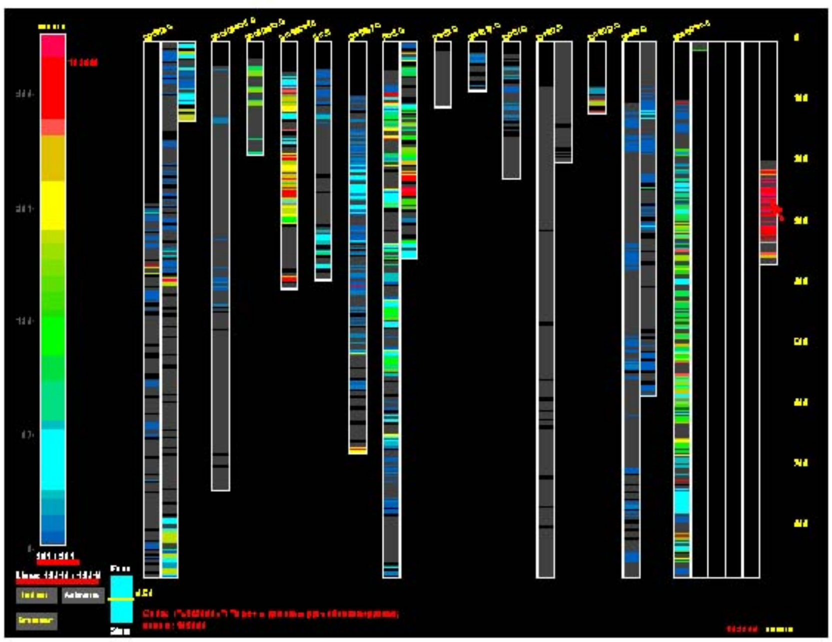

Figure 1.5. Screenshot of *SeeSoft* ® interface (1992). Those rows colored dark blue represent the oldest lines of code whereas those colored red represent lines that are new.

[87] Stephen G. Eick, Joseph L. Steffen, and Eric E. Sumner Jr., "Seesoft - A Tool for Visualizing Line Oriented Software Statistics," *IEEE Transactions on Software Engineering* 18, no. 11 (November 1992): 957–68.

SeeSoft ® also serves to illustrate what is commonly known as the Visual Information-Seeking Mantra—"overview first, zoom and filter, then details on demand." [88] The mantra, first identified by infovis researcher Ben Schneiderman in a 1996 article, reflects the overall design principle structuring many information visualization projects. It holds true for both the design process and process of reception.

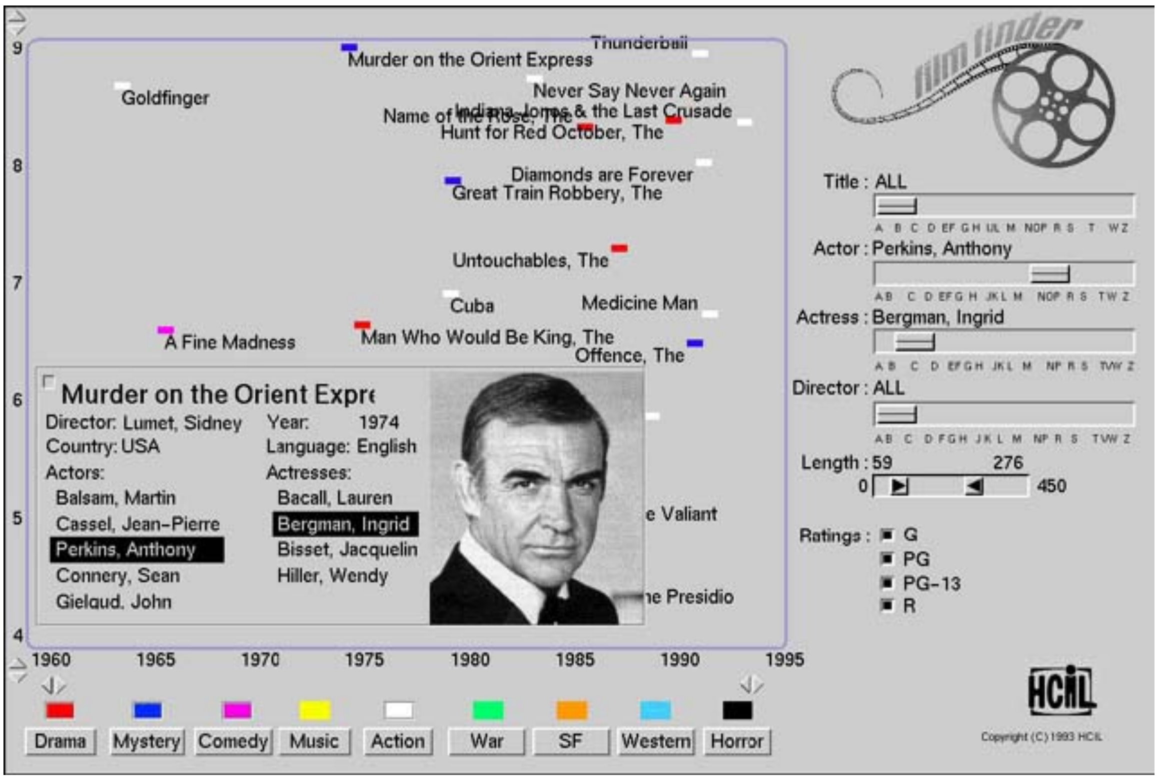

Figure 1.6. Screenshot of *FilmFinder* interface (1994).

[88] B. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," in *Visual Languages, 1996. Proceedings., IEEE Symposium On*, 1996, 336–37, https://doi.org/10.1109/VL.1996.545307.

One final example helps to cohere the discussion. *FilmFinder* (Figure 1.6) is a graphical search engine that allows users to search for certain films in a movie database without any prior or expert knowledge.[89] It combines the goals of efficiency and interactivity with a focus on the user in a way that earned it a similar status to Wilhemson's storm visualization.[90] What is so impressive about *FilmFinder* is that it allows the user to dynamically search through an unknown dataset and arrive at insight vis-à-vis visualization. The screenshot below (Figure 1.6) shows a search for films with Anthony Perkins and Ingrid Bergman between 59 and 276 minutes in length. Users can set search parameters by manipulating simple query devices (e.g., toggles, sliders) to the right.

The results are then displayed in an interactive scatterplot known as a starfield to the left.[91] Genre (color), title, production year (x-axis), and popularity (y-axis) are also shown. Query devices can then be further manipulated, thereby dynamically affecting the starfield display. Note the level of attention to what is intuitive and efficient for the user to use. Query parameters are easy to manipulate without having to enter

---

[89] Christopher Williamson and Ben Shneiderman, "The Dynamic HomeFinder: Evaluating Dynamic Queries in a Real-Estate Information Exploration System," in *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '92 (New York, NY, USA: ACM, 1992), 338–346; Christopher Ahlberg and Ben Shneiderman, "Visual Information Seeking Using the FilmFinder," in *Conference Companion on Human Factors in Computing Systems*, CHI '94 (New York, NY, USA: ACM, 1994), 433–434.

[90] This assessment is derived from person observation while browsing the information visualization literature. While *FilmFinder* was not explicitly dubbed "the teapot of information visualization," it and the projects leading up to it are frequently given as examples.

[91] Christopher Ahlberg and Ben Shneiderman, "Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '94 (New York, NY, USA: ACM, 1994), 313–317.

code. Black represents horror and drama is signified by red—two color associations that work well given the intended (primarily American) audience.[92] The search process, not coincidently, parallels that of wandering through a library or video store. People browse different categories, search for an interesting relationship or further information, then decide on a movie. This is a parallel Schneiderman and Ahlberg were careful to create.

Contemporary data visualization continues to grow. New technology, new understanding, and new techniques are constantly pushing the boundaries of what visualization is or can be. While scientific visualization and information visualization have remained the dominant genres within the field, new genres have begun to emerge. The remainder of this dissertation focuses on data visualization's more recent development. Nevertheless, the foundational period, spanning from the late 1970s through the 1990s, still plays a role. It is important to recognize the historical precedents and motivations as the field moves forward. The next chapter continues to chart contemporary data visualization's growth. Many of the more recent trends and developments are best understood in relation to its past.

---

[92] Different cultures will often attribute different meanings to the same color. For example, red symbolizes good fortune and joy in the Chinese culture as opposed to the American associations of warning, anger, and passion.

CHAPTER 2

Contemporary Visualization Expands (2000 - present)

When contemporary data visualization first emerged in the mid-1980s, it did so with a specific intent. Visualization was a tool that could be used by experts to help analyze and comprehend large amounts of data. While multiple disciplines played a role, visualization took on a scientific analytic mindset. This resulted in the prioritization of a certain type of viewer and set of ideals. Viewers were predominantly experts and goals pragmatically based. The typical setting for a data visualization in the 1980s and 1990s was a scientific research facility with access to powerful computing technology and an audience with a high level of level expertise—think NASA. The typical setting for a data visualization in the 2000s includes this and a great deal more.

Data visualization is now pervasive. In 2002, nearly half of the Net Art projects at the Whitney Biennale involved some form of visualization.[1] In 2017, the number of visualization projects that can also be categorized as art is too high to count. There are books and websites and entire companies dedicated to helping people visualize information. Newspapers like *The New York Times* and *The Guardian* frequently incorporate visualizations as part of their reporting and investigation. There are even award ceremonies for data visualization.

[1] Lev Manovich, "Data Visualization as New Abstraction and Anti-Sublime," 2002.

What changed? This chapter picks up where the previous chapter left off. It charts the trajectory of contemporary data visualization's development from the early 2000s to the present. This period can best be characterized as data visualization having expanded outside of traditional research environments and into nearly every dimension of life. While the existing genres of scientific visualization and information visualization have continued to grow, new genres, audiences, settings and goals have entered contemporary data visualization's purview. The chapter is structured by looking at three broad areas of practice or research that have emerged during this period. First, I examine data visualization as art and tease apart the relationship between art and aesthetics. Second, I look at what has been called social data visualization, which is a broad categorization of projects that either represent or facilitate human interaction and relationships. Third, I discuss the popularization and everyday use of data visualization. The chapter concludes by introducing digital humanities and humanistic approaches to data visualization that have gained popularity in recent years.

Visualization as Art

Perhaps the first area to embrace contemporary visualization outside of the realm of scientific and computational research was new media art. New media became an established label in the 1990s for a range of artistic practices incorporating digital and electronic media technology.[2] The mainstreaming of computers and computer graphics, as well as the introduction of the World Wide Web, meant that people could access, analyze, and display data on their computers at home.[3] No longer was it necessary to have specialized equipment or technical expertise to interface with data. This, combined with the cultural importance of digital technologies, led many artists to use data to create art. It was not long before artists started experimenting with some of the same visualization techniques used in scientific and social scientific research.

One early example is Lisa Jevbratt's *1:1* (1999/2002).[4] The project, which centers on a database that contains the IP addresses of every Web site in the world within a given year, offers five dynamic interfaces through which viewers can use the database and navigate the Web.[5] The "every IP" interface shown below (Figure 2.1) looks like it could be a traditional information visualization, but it has a slightly different intent. The goal is still to explore and comprehend the represented data but the intent of doing so is to experience the data. In the artist's own words:

[2] For a comprehensive resource on media art, see Christiane Paul's *Digital Art* (2008) or *New Media Art* (2009) edited by Reena Jana and Mark Tribe.
[3] Martin Wattenberg and Fernanda Viegas, "Design and Redesign in Data Visualization," Medium, March 27, 2015, https://medium.com/@hint_fm/design-and-redesign-4ab77206cf9#.a9nusg2ir.
[4] Lisa Jevbratt, 1:1 (2), 2002 1999, Web, 2002 1999, http://128.111.69.4/~jevbratt/1_to_1/index_ng.html.
[5] The original project, 1:1, consists of only one database generate in 1999. The second version, 1:1 (2), adds an additional database generated in 2001.

When navigating the Web through the databases, via the five interfaces, one experiences a very different Web than when navigating it with the "road maps" provided by search engines and portals. Instead of advertisements, pornography, and pictures of people's pets, this Web is an abundance of inaccessible information, undeveloped sites and cryptic messages intended for someone else. Search-engines and portals deliver only a thin slice of the Web to us, not the high-resolution image we sometimes think they do. The interfaces/visualizations are not maps of the Web but are, in some sense, the Web. They are super-realistic and yet function in ways images could not function in any other environment or time. They are a new kind of image of the Web, and they are a new kind of image.[6]

Figure 2.1. Lisa Jevbratt, “every IP” interfaces from *1:1* (1999/2002).

Jevbratt’s project was not isolated. It, along with many other similar projects, was included in prominent art and design exhibitions during the first part of the

[6] Lisa Jevbratt, “Description [1:1],” 1:1 (2), 1999/2002, http://128.111.69.4/~jevbratt/1_to_1/description.html.

decade. SFMOMA's *010101: Art in Technological Times* (2001), the Whitney Biennale (2002), and Ars Electronica (2001-2004) all include projects that could be categorized as data visualizations in that they graphically represent data. The appearance of such projects led Lev Manovich to conclude that data visualization, along with Graphical User Interface, the database, navigable space, and simulation, is one of the "genuinely new cultural forms enabled by computing."[7] As Manovich points out, data visualization does not require the use of a computer. There are many examples of graphical representation that do not intersect with computing either in terms of data-processing or display.[8] However, computing culture and technology, as Chapter 1 points out, is a primary player in the in the development of the contemporary visualization field. It turns such representations of data from the exception into the norm and engenders the possibility of new visualization uses and techniques.[9]

Artistic Data Visualization

The trend of artists using data visualization for artistic purposes continued. In a 2007 article, Martin Wattenberg and Fernanda Viegas gave a name to this body of work calling it *artistic data visualization*.[10] As is the case with scientific and information

---

[7] Manovich, "Data Visualization as New Abstraction and Anti-Sublime."
[8] The previous chapter points to a few in the section of visualization precedents.
[9] Manovich, "Data Visualization as New Abstraction and Anti-Sublime."
[10] F. Viégas and M. Wattenberg, "Artistic Data Visualization: Beyond Visual Analytics," in OCSC'07 *Proceedings of the 2nd International Conference on Online Communities and Social Computing* (Springer, 2007), 182–191.

visualization, the label comes with certain parameters and norms. Wattenberg and Viegas define artistic data visualization as “visualizations of data done by artists with the intent of making art.”[11] They are very careful to differentiate it from other types of data visualization (and other types of art) based on a few characteristics. First, artistic data visualizations must be data-based. That is, they must represent actual data, rather than metaphors or surface appearances. Second, they do not have to be pretty. The definition offered by Wattenberg and Viegas purposefully avoids the issue of beauty. Beautiful scientific visualizations, such as the picture of a cellular structure taken under a microscope, do not automatically qualify as artistic, nor do artistic visualizations have to be eye-pleasing. What distinguishes artistic data visualizations from other types of visualization is that they create a guided aesthetic experience. Visualization is often assumed to be a neutral, analytic tool. While artistic visualizations may employ some of the same techniques used in the scientific visualization community, they do not feign objectivity. Rather, artistic data visualizations, in Wattenberg and Viegas’ words, work to “actively guide analytical reasoning and encourage a contextualized reading of their subject matter.”[12] It is this guidance and work to contextualize the data that most strongly characterizes the genre.

Golan Levin’s *The Secret Lives of Numbers* (2002) offers a concrete example. The project, which visualizes the relative popularity of every integer between one and

[11] Viégas and Wattenberg, "Artistic Data Visualization," 182.
[12] *Ibid.*, abstract.

one million, was commissioned by New Radio and Performing Arts, Inc. for its Turbulence.org website. In 1997, Levin and a team of others created a script to search the Web at regular intervals for the number of web pages containing a given number.[13] The cumulative results (1997-2002) were then presented as an interactive visualization available to any viewer online. What becomes immediately apparent once visualized is that some numbers are more popular than others. Some numbers, like 1040, 1776, and 90210, enjoy a greater popularity than their neighbors because they are associated with tax forms, famous dates, or well-known zip codes. Others, like 12345 or 8888, appear to be more popular because they are easier to remember. We also seem to have a preference for intervals of ten.

Figure 2.2 shows a screenshot of the visualization with the number 1812 selected. Its ranking, details about its popularity, and common associations are made readily visible. Yet, the intent of the visualization is not to communicate any of this information or information about any one integer, for that matter. The intent of *The Secret Lives of Numbers* is to do exactly what the description in the upper left corner invites, "explore how the usage patterns of numbers reflect our culture, history, and biology."[14] The interactivity that holds parallel to information visualization is not characteristic of all projects in the genre. In fact, many artistic visualizations represent the data in static form. What is characteristic of the genre is that the work to contextualize the data

---

[13] Martin Wattenberg, Jonathan Feinberg, David Becker, David Elashoff, and Shelly Wynecoop also contributed to the project.
[14] http://turbulence.org/Works/nums/applet.html

uncovers its social and cultural meaning. *The Secret Lives of Numbers*, like many other artistic visualization projects, directs our attention to the cultural and subjective experience of living with data. It asks us not so much to understand as to reflect.

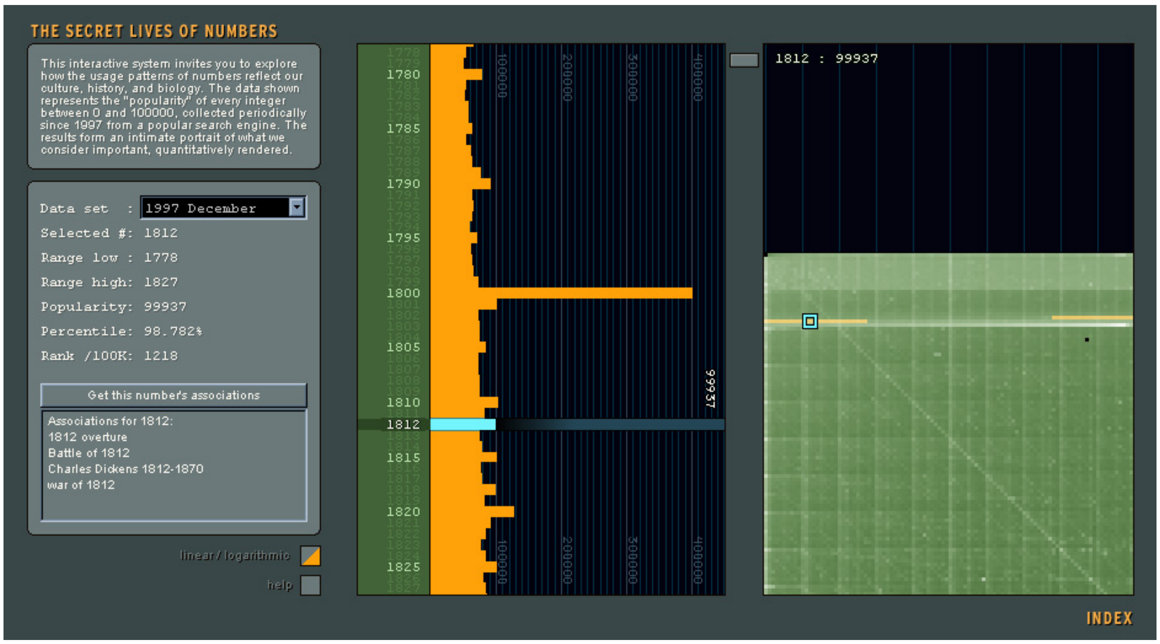

Figure 2.2. Screenshot of Golan Levin's *The Secret Life of Numbers* (2002).

A Broader Interest in Aesthetics

The projects discussed thus far in this chapter easily fit into the genre of artistic data visualization. They were created by artists with artistic intent. However, not all visualizations that guide aesthetic experience fall so neatly into the category. Ben Fry's *isometricblocks* (2002-2004/5) (Figure 2.3) serves as an interesting case in point.[15]

15 http://benfry.com/isometricblocks/.

The project, which visualizes single letter changes in the genome of three different populations, was included in the Museum of Modern Art's (MoMA) *Design and the Elastic Mind* exhibition in 2008. Yet, it was not originally intended to be art. Fry created the project as part of his graduate work in the Aesthetics + Computation Group at the MIT Media Laboratory. While acknowledging the importance of aesthetics, his self-stated goal was "to develop a single software piece that combines aspects of many of the preceding analysis and visualization tools for haplotype and LD data."[16] In other words, Fry's project was at least in part designed to serve a functional purpose. Genetic researchers have used it to this end. One visualization was even used to illustrate the cover of the HapMap issue of *Nature* in October 2005, three years prior to its display in a museum.

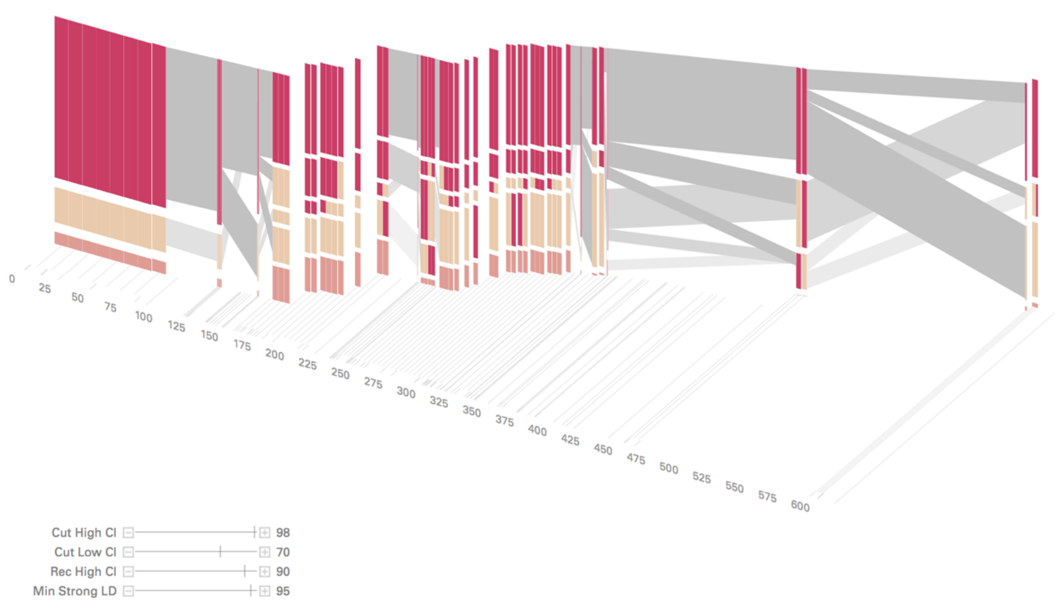

Figure 2.3. Screenshot of Ben Fry's *isometricblocks* (2002-2004/2005).

[16] The Art of Data Visualization, You Tube video, Off Book (PBS Digital Studio), accessed November 15, 2017, https://www.youtube.com/watch?v=AdSZJzb-aX8.

Fry's *isometricblocks* (2002-2004/5) hovers in a gray area between information visualization and visualization art. It is representative of a growing number of visualizations that, beginning in the first decade of the 2000s, worked to combine visualization goals with principles of art and design. The combination, while now not uncommon, was unconventional at the time. Art and aesthetics have been particularly slow to enter the visualization literature. The standard argument against including them in the discussion has been that that which is pleasing to the eye or introduces subjective experience is superfluous. Part of the reason for this line of thought stems from disciplinary boundaries. While data visualization is very much a field that borrows from both art and science, the heavy financial and infrastructural support provided by the science and computing industry following contemporary visualization's birth undoubtedly affected its character. Anything deemed outside the realm of functionality, such as aesthetics, was black-boxed out.

In the late 1990s and early 2000s, perspectives started to change. Researchers in fields such as human computer interaction (HCI) and cognitive science began to recognize that aesthetics and other social and cultural experiences have an effect on cognition.[17] Still, aesthetics remained outside the visualization literature. As late as 2005, Chaomei Chen identified aesthetics as one of the top ten unsolved problems in

[17] Edwin Hutchins, Cognition in the Wild (MIT Press, 1996); Donald Norman, Emotional Design: Why We Love (or Hate) Everyday Things. (Basic Books, 2004); Patrick W. Jordan, Designing Pleasurable Products, (CRC Press, 2002).

information visualization.[18] Figuring out how to account for subjective experience in a field that prided itself on objectivity proved challenging. Much of the reason for the prolonged absence of aesthetics in contemporary visualization culture has to do with the work of Edward Tufte. His book *The Visual Display of Quantitative Information* (1983) introduced some of the first guidelines as to how to best design visualizations. Using original concepts such as *chartjunk* and the *data-ink ratio*, Tufte effectively advocates for a pared-down aesthetic based on principles of truth and clarity. In his opinion, any design element that does not most efficiently show the data is confusing. Redundant or decorative markings are therefore deemed unnecessary, or worse misleading.

Tufte's influence on contemporary visualization's stylistic development cannot be overstated. His aesthetic, or lack thereof, set the standard for what counted as good visualization design well into the 2000s. One only needs to look at the astonishing number of times his work has been cited to see the evidence.[19] Part of the reason for Tufte's success has to do with timing. While Jacques Bertin's *Sémiologie Graphique* introduced some of the same recommendations based on similar principles when first published in 1967, the work was not translated into English until 1983 and therefore never widely disseminated. The other reason that Tufte's work persists has

---

[18] Chaomei Chen, "Top 10 Unsolved Information Visualization Problems," *IEEE Computer Graphics and Applications* 25, no. 4 (2005): 15.

[19] According to a Google Scholar search in November 2017, *The Visual Display of Quantitative Information* has been cited 9757 times. Tufte's later books receive similarly high rates of citation in relation to other works in the visualization design literature. By contrast, Nathan Yau's 2012 book *Visualize This* only has 184 citations.

to do with principle. In *The Visual Display of Quantitative Information* and later books, Tufte ties his standards for good graphical design to the belief that visualizations can and should tell the truth. The assumption largely stems from the publication of Darrell Huff's 1954 bestseller *How to Lie with Statistics*. The popular statistics textbook describes some of the most common visual sins that people can make when representing data, such as truncating the y-axis.[20] The result was a decades-long vigilance against creating misleading graphics.

So, does this mean that all artistic visualizations lie? That any attention to aesthetic experience obscures the truth? Certainly not. Projects that attend to aesthetics have different purposes, goals and audiences than other data visualizations. Each visualization must be looked at in its specific context. The underlying intent, whether this also involves creating art or not, is to represent data in a way that facilitates a deeper understanding of it. There are multiple ways to represent and thereby understand data. The assumption that there is a single, objective truth to be uncovered ignores the fact that data visualizations are representations. And as representations, they are *designed*. The choices inevitably made when deciding how to best represent the data are directly influenced by the multiple factors informing each visualization. Moreover, empirical research has shown that aesthetic factors such as what one likes or is pleasing to the eye cannot be

[20] *How to Lie with Statistics* (W. W. Norton & Company, 2010), chapter 5.

separated from functionality.[21] A 2007 study surveying two hundred and eighty-five participants found that perceived aesthetic quality has a measurable impact on a visualization's usability. Those designs that participants ranked as having a higher aesthetic quality performed better on metrics of efficiency and effectiveness as well as measures of task abandonment and latency in erroneous response.[22]

Visualization Criticism and Information Aesthetics

There have been a handful of papers that have tried to carve out a space for incorporating art and aesthetics into visualization design and research. In a 2007 paper, Robert Kosara, proposes visualization criticism as a way to bridge the gap between two different cultures of visualization—pragmatic visualization and artistic visualization.[23] Pragmatic visualization involves the technical application of visualization techniques for data analysis, whereas artistic visualization aims to evoke a deep emotional and intellectual response. Defining each culture, Kosara shows how they exist on a continuum between utility and sublimity. He further argues that because the two are on opposite ends of a spectrum and thus theoretically impossible

21 Norman, *Emotional Design*; Noam Tractinsky, "Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '97 (New York, NY, USA: ACM, 1997), 115–122.

22 N. Cawthon and A. V. Moere, "The Effect of Aesthetic on the Usability of Data Visualization," in Information Visualization, 2007. IV '07. 11th International Conference, 2007, 637–48.

23 R. Kosara, "Visualization Criticism - The Missing Link Between Information Visualization and Art," in Information Visualization, 2007. IV '07. 11th International Conference (Information Visualization, 2007. IV '07. 11th International Conference, IEEE, 2007), 631–36.

to reconcile, it does not mean we should not try.[24] There is a productive space in between. Kosara's suggestion for navigating this space is what he calls "visualization criticism," which is a type of art criticism applied to visualization. Researchers and designers would critique projects based on fact and alternative solutions. The reflection on design decisions is intended to open up a new space as well as provide the building blocks for a language and theory of visualization. How do we attach meaning to graphical objects? How should we represent different kinds of data to make them easier to understand? What role do metaphors play, and how can they be used?[25] These are the types of questions that visualization criticism aims to build a vocabulary for answering and asking. Precisely how a theoretical framework would be developed is not explained. However, the motivation and need for such a framework and critically examining data visualization from multiple perspectives is laid out.

Andrew Lau and Andrew Vande Moere propose a similar bridge with their model of information aesthetics.[26] The proposed model is based on analyzing existing projects for their interpretive intent and data-mapping technique. It reveals information aesthetics as a continuum between visualization art and information visualization that incorporates several different types of visualization, including social visualization, information art, and ambient visualization. The diagrams offered in the paper help to

---

[24] Kosara, "Visualization Criticism," section 4.4.
[25] Kosara, "Visualization Criticism," section 5.
[26] A. Lau and A. Vande Moere, "Towards a Model of Information Aesthetics in Information Visualization," in Information Visualization, 2007. IV '07. 11th International Conference, 2007, 87–92.

illustrate this (Figure 2.4). The model is more conceptual than practical. Its purpose is not so much to try to come up with a new way of categorizing visualization projects as it is to provide a framework, or rather the idea of a framework, for understanding the roles of aesthetics in visualization. Neither visualization criticism nor the information aesthetic model outline a clear path forward. What is important is that they are both deliberate attempts to expand the domain of visualization design and research outside traditional boundaries.

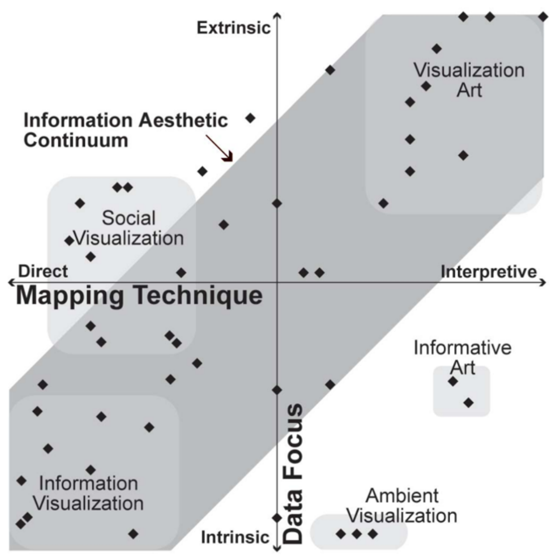

Figure 2.4. Model of Information Aesthetics from A. Lau and A. Vande Moere, “Towards a Model of Information Aesthetics in Information Visualization,” *IV ’07*.

The Present

The relationship between art, aesthetics and visualization has continued to grow. A simple internet search for the pairing art and visualization in November 2017 yields approximately ten million results.[27] Not all of the results pertain to art and visualization as discussed in this chapter—there are also entries about the art of data visualization[28] and visualizations that have been deemed beautiful or artistic[29]—but the sheer number provides an indication of just how much art and visualization are now intertwined. Contemporary art and design institutions have continued to include data visualization projects in their exhibitions and permanent collections. One museum, or more specifically, one curator, has been particularly influential in promoting the conceptualization of visualization as an artistic medium. Senior design curator at the Museum of Modern Art (MoMA) Paola Antonelli has repeatedly pointed to visualization's social and cultural significance across multiple exhibitions over the past ten years. Her 2008 exhibition *Design and the Elastic Mind* was the first to include visualization as one of eleven possible thematic categories. The exhibition, which explores the reciprocal relationship between science and design in the contemporary world, showcased twelve different data visualization projects. Visualization featured

---

[27] A Google search for art and visualization November 12, 2017 yielded approximately 97,900,000 hits.
[28] The Art of Data Visualization, You Tube video, Off Book (PBS Digital Studio), accessed November 15, 2017, https://www.youtube.com/watch?v=AdSZJzb-aX8.
[29] For example, a search on Pinterest for art and visualization yields images of hundreds of projects that users have categorized as data art.

similarly in MoMA's 2011 *Talk to Me* exhibition. The online version of the exhibition allows users to filter projects by selecting "visualization" as a tag.[30] More recently, Antonelli organized three capsule exhibitions as part of an ongoing collaboration with the Hyundai Card Design Library in Seoul. The second exhibition, Designing with Data, opened in July 2015.[31]

Many of the projects included in these exhibitions are intended to be art. Others are not. In 2016, MoMA acquired Stefanie Prosavec and Geogia Lupi's "slow data" visualization project *Dear Data* as part of its permanent collection.[32] The project, originally published as a website, documents correspondence between Prosavec and Lupi in the form of hand-drawn visualizations on postcards sent through the mail.[33] The two designers met at a conference in 2014 and wondered if it was possible to get to know each other through data alone. Intrigued by the idea, they embarked on a year-long experiment to find out. Every week, for an entire year, Prosavec and Lupi collected and measured a particular type of data about their personal lives (e.g., the times they laughed, the time they complained, the times they engaged in certain activities). They then represented this data on a postcard-sized sheet of paper and dropped the postcard in a mailbox to be delivered to the other's address (Figure 2.5).

---

[30]http://www.moma.org/interactives/exhibitions/2011/talktome/objects/#category=objects
[31] For more information see Past Exhibitions on the library's website http://library.hyundaicard.com/design/.
[32] Giorgia Lupi, "Dear Data Has Been Acquired by MoMA, but This Isn't What We Are Most Excited about.," Medium (blog), November 22, 2016, https://medium.com/@giorgialupi/dear-data-has-been-acquired-by-moma-but-this-isnt-what-we-are-most-excited-about-bdaa3376d9db#.m8r4jqw3d.
[33] http://www.dear-data.com/theproject

Prosavec lives in London; Lupi in New York. What resulted is a rather intimate collection of images that both records and gives insight into a person's day-to-day life.

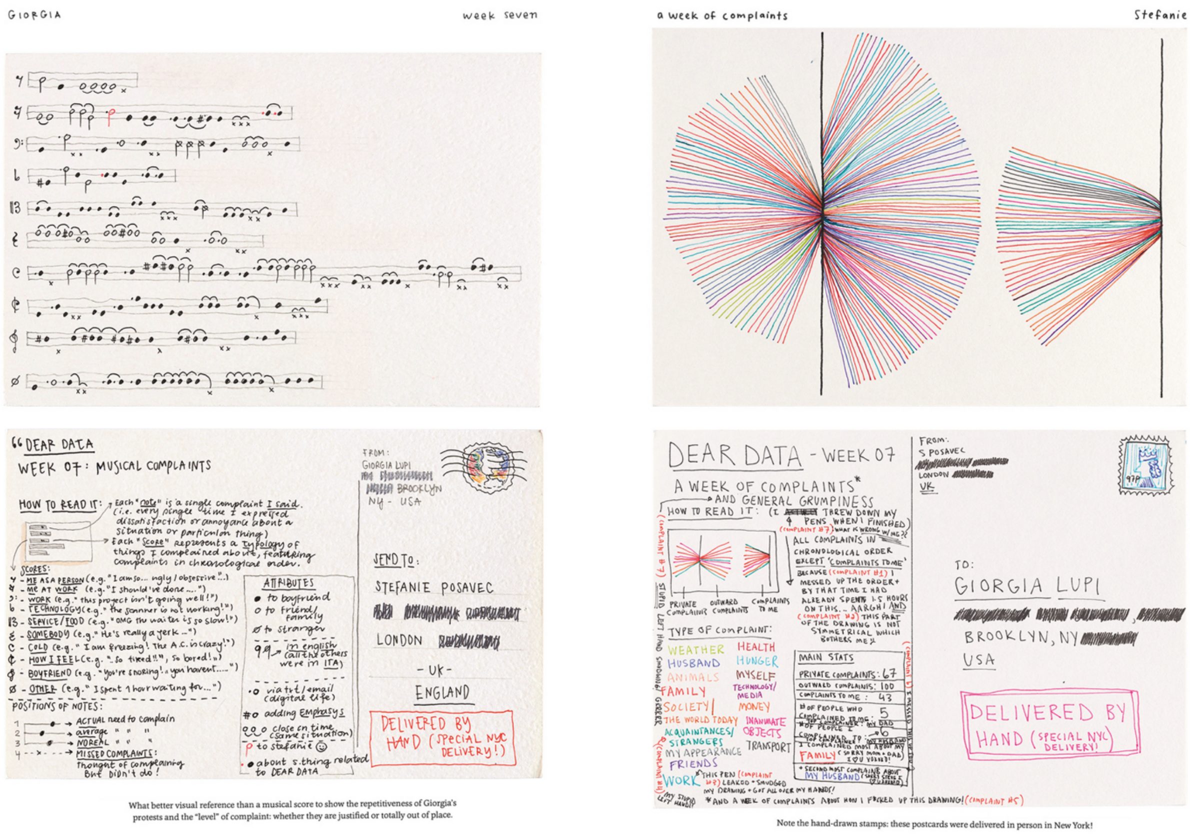

Figure 2.5. Geogia Lupi and Stefanie Prosavec, *Dear Data* (2014-15). The postcard on the left shows Lupi's record and visualization of complaints. The postcard on the right shows the same week's record and visualization by Prosavec.

The experience of viewing *Dear Data* is a bit like the experience of looking through a box of photographs or similar artifacts. Each is personal, each is data-based, and each takes on additional meaning because of something that cannot be seen in the representation alone. Interestingly, the project has since been turned into a book and inspired a bit of a following.[34] Readers of *The New York Times* statistics blog *FiveThirtyEight* were asked to send in visualizations of their podcast listening habits,[35] and other designers have created similar projects that visualize their own lives.[36] There is a long history of artists methodically documenting and representing their day-to-day behavior.[37]

Social Data Visualization

Art is known for pushing boundaries. While the expansion of data visualization into the realms of art and aesthetics made no claim of being avant-garde, it did push the boundaries of what was expected and considered acceptable in terms of a visualization audience, setting, and ideal. A visualization on display at a museum functions very differently than the same visualization on the cover of a scientific journal. Another area of practice and research that grew in the early 2000s and continues to challenge traditional visualization norms is what researchers broadly refer to as social data visualization. “Social” is a vague term that can be interpreted in many

---

[34] Giorgia Lupi and Stefanie Prosavec, Dear Data (Princeton Architectural Press, 2016).
[35] See FiveThirtyEight’s call for readers to visualize their podcast habits at http://fivethirtyeight.com/features/dear-data-and-fivethirtyeight-want-you-to-visualize-your-podcast-habits/.
[36] “Dear Data Two,” accessed January 20, 2017, http://www.dear-data-two.com/.
[37] For One Year Performance (1980–1981), Tehching Hsieh recorded himself punching a time clock in his New York studio every hour for an entire year. Other artists have similarly taken a photo every day for years on end and then stitched the images together into a time-lapse video.

different ways. The way it is used in the following section is to refer to the relationships and interactions between people and objects in a given environment.

Visualizing Social Relationships

The application of visualization to the study of social relationships is not new. Researchers, particularly those in the social sciences, have been using visual means to represent human relationships since the 1930s. In the 1930s, social psychologist Jacob L. Moreno pioneered the recording and analysis of social interactions within small groups. He further introduced some of the first node-link diagrams in his 1934 book *Who Shall Survive?*[38] The node-link diagrams, or "sociograms," as Moreno called them, proved immensely helpful in terms of tracking and analyzing individual preferences and behavior. Visually representing the relationships between group members allowed him to better see the entirety of the network and this greatly aided in understanding its patterns and structure.[39]

Moreno's work is some of the earliest social network analysis (SNA), which refers to the process of investigating social structures through the use of networks and graph theory. While social network analysis did not become a standard means for investigating human relationships until the late 1970s, visual techniques continued to

---

[38] Jacob L. Moreno, *Who Shall Survive?: Foundations of Sociometry, Group Psychotherapy, and Sociodrama* (Beacon House, 1977).
[39] Moreno; Linton C. Freeman, "Visualizing Social Groups," in *American Statistical Association*, 1999, 47–54.

evolve alongside research and technology.[40] For example, the introduction of the computer in the 1950s allowed for computationally demanding procedures such as multidimensional scaling and factor analysis. Computers also allowed for more complex visual components and interactions.[41] In the past two decades, visualization techniques have been widely applied to social networks to facilitate better accessibility and increased volume. Programs such as Pajek, NetVis, and UCINet allow users to create dynamic visualizations and perform social network analysis on a very large scale.[42] To date, there have been large-scale social network analyses of knowledge communities,[43] co-authorship in scientific publications,[44] and correspondence[45] among many other manifestations.

Such programs and projects are typically grounded in sociological research, but there are other ways in which visualization meets with understanding social networks and relationships. Perhaps no single factor has had more of an impact in driving social

---

[40] For a brief history of Social Network Analysis, see Mingxin Zhang, "Social Network Analysis: History, Concepts, and Research," in *Handbook of Social Network Technologies and Applications*, ed. Borko Furht (Springer Science & Business Media, 2010), 3–22.

[41] Ing-Xiang Chen and Cheng-Zen Yang, "Visualization of Social Networks," in *Handbook of Social Network Technologies and Applications*, ed. Borko Furht (Springer US, 2010), 585–86.

[42] Freeman, "Visualizing Social Networks." For a comprehensive list of available programs, see the Wikipedia entry for "Social network analysis software."

[43] For example, see http://www.visualcomplexity.com/vc/index.cfm?domain=Knowledge%20Networks. The Visual Complexity website also provide countless examples of other types of network visualizations.

[44] M. E. J. Newman, "Coauthorship Networks and Patterns of Scientific Collaboration," *Proceedings of the National Academy of Sciences* 101, no. 1 (April 6, 2004): 5200–5205.

[45] Danyel Fisher and Paul Dourish, "Social and Temporal Structures in Everyday Collaboration," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04 (New York, NY, USA: ACM, 2004), 551–558.

data visualization further than the introduction of the Web.[46] New forms of online sociality such as e-mail, blogging, and remote collaboration generated an enormous amount data.[47] Visualization applications have increasingly been used to help people make the most of their social relationships and access abundant resources online. Visualizing social networks is thus not simply an approach to research but also a method of accessing and connecting people and things in effective ways.[48]

Online Collaboration

The relationship between social data visualization and the Web is not black and white. It is not as if social data visualization suddenly became needed or trendy following the introduction of online forms of sociality (e.g., e-mail, blogging, social networking sites), but that the popularization of the Web at the beginning of the 21st century served as a tipping point. The amount data that pertained to interactions between people and resources magnified greatly. Some of the earliest projects to be classified as social data visualization mapped collaborative online activity.

---

[46] Tim Berners-Lee invented the fundamental technologies that remain the backbone of today's web in 1990. The number of browsers and web pages grew exponentially throughout the 90s. By 2001, the web had 51.3 million users worldwide. Today's estimate is well over three times as many. For more details, see "The History of UVA on the Web," The University of Virginia, May 17, 2017, http://www.virginia.edu/content/404; "History of the Web," World Wide Web Foundation, accessed September 5, 2017, https://webfoundation.org/about/vision/history-of-the-web/.

[47] The trend of ever increasing data is one of the reasons visualization has remained at the forefront of scientific, social, and technological research over the last thirty years. New bodies of data and new ways to understand them exist in a co-dependent relationship in which the continued development of visualization is the result.

[48] Chen and Yang, "Visualization of Social Networks," 587.

One such project, *Soylent*, is an infrastructure for examining and visualizing the social and temporal structures of collaboration left behind by user activity. It was created by University of California, Irvine researchers as part of a larger study to examine everyday ways to support collaboration in the workplace.[49] Collaborative online activity has an antecedent in Computer-Supported Cooperative Work (CSCW). The design-oriented field emerged in the late 1980s as a way to address how people working together in groups could be supported by computer systems and their associated technologies.[50] It is out of this tradition, combined with visualization, that *Soylent* emerged. Using network graphs and other visualization techniques, researchers were able to show that social and temporal patterns exist and operate across different applications. More importantly, users could identify and give relevance to the patterns they saw.[51]

Because *Soylent* was created in 2004, much of the collaborative activity that was studied took place online. Even prior to the popularization of the Web, networked computing and shared resources related to work set the stage for visualization pertaining to online activity and the social relationships between people. It is the visual representation of these relationships, not the fact that the relationship at least partially took place online, that qualifies the project as social data visualization. Many

---

[49] Fisher and Dourish, "Social and Temporal Structures in Everyday Collaboration."
[50] Peter H. Carstensen and Kjeld Schmidt, "Computer Supported Cooperative Work: New Challenges to Systems Design," in In K. Itoh (Ed.), *Handbook of Human Factors*, 1999, 619–636.
[51] Fisher and Dourish, "Social and Temporal Structures in Everyday Collaboration," 556.

visualization tools and projects created around the same time more directly explored online collaboration and social activity. *ContactMap* (Figure 2.6), for example, is an application designed in 2002 that combines a visual representation of important contacts with built-in tools and messages that help users identify important contacts and initiate actions associated with them.[52] The interface is much like a social network graph from which you can retrieve and create communications.

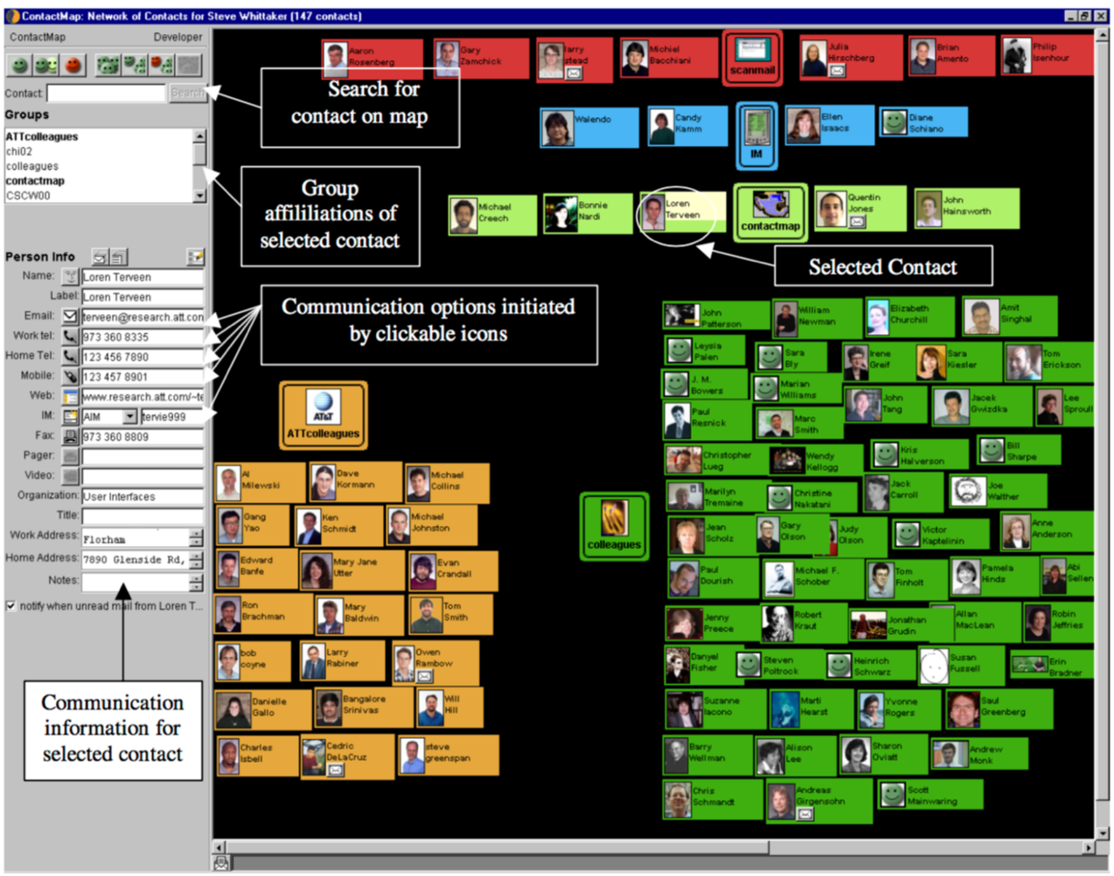

Figure 2.6. Screenshot of *ContactMap* (2002). Contact Information panel showing communication and initiation functions.

[52] Bonnie A. Nardi et al., "Integrating Communication and Information Through ContactMap," *Communications of the ACM* 45, no. 4 (April 2002): 89–95.

The ability to connect people, content, and resources through visualized connections is a common feature in many co-working tools that are popular today. Project management applications like Trello[53] and Asana[54] greatly benefit from visual representations of connections and tasks. For example, Asana's dashboard (Fig. 2.7) includes multiple charts that give users an overview of their team's collective progress and activity. Platforms like Google Drive and other shared file services also use visualization to aid remote collaboration. Because files are updated in real-time and have a user's signature tied to any changes, it is possible to see the amount of teamwork that went into creating a given document or presentation.

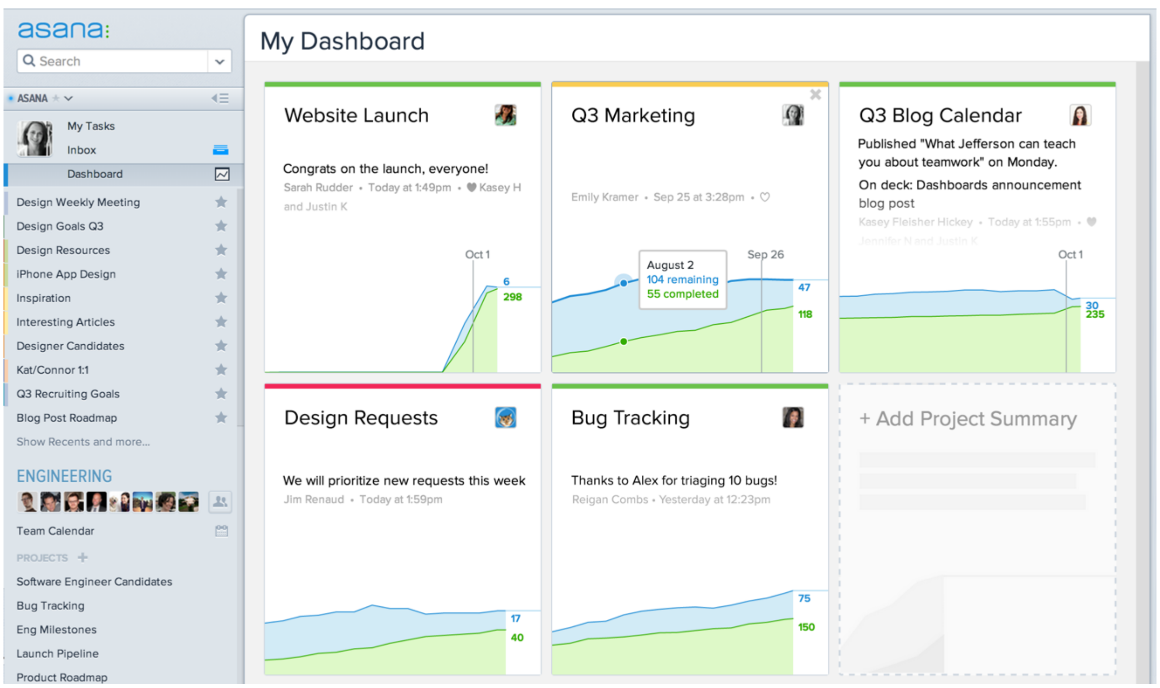

Figure 2.7. Asana Dashboard (2017).

[53] http://trello.com
[54] https://app.asana.com

Asana and Google Drive fall into the category of what are considered Web 2.0 technologies. Web 2.0 is a term that was coined in 2004 to describe a new generation of websites and services that emphasize user-generated content, collective intelligence and interoperability.[55] Much of social data visualization's growth over the past 10 years has been a result of Web 2.0 technologies. When an individual posts a video on YouTube, he or she becomes a visible part of the service's online community. Even the presentation of a username linked to a personal profile serves to visualize social activity. The number of likes, the invitation to subscribe and the comments section all visually reinforce the idea of connectivity. In a way, you can even think of an individual page as a visual representation of connections in the larger YouTube and Web communities. Each visualizes data—in these cases HTML and javascript code. Each contains links to other users, media and places (a.k.a., nodes). Each is itself a node in the massive network that is the Web. From this perspective, the Web can almost be seen as a large-scale and multiple-component social data visualization.

Social Media and Social Networking Sites

Another huge area of growth within social data visualization has been tied to social media and, in particular, social networking sites. In the early 2000s, these provided rich sources of largely social data that researchers turned into visualization to

---

[55] Tim O'Reilly, "What Is Web 2.0," O'Reilly, September 30, 2005, http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html.

help try to navigate and understand. Vizster (2003) is a visualization system designed to be used by the end-users of Friendster, an early, now-defunct social networking site.[56] While not the first of its kind, Friendster was the first social networking site to attain over 1 million users. Typical interfaces to social networking sites at the time gave users an idea of the scale of their network but remained largely linear. In 2003, researchers dana boyd and Jeffrey Heer created Vizster to allow users to see and build their social connections. Its goal, in boyd and Heer's own words, was simply "to build a visualization system that end-users of social networking services could use to facilitate discovery and increased awareness of their online community."[57] In other words, Vizster's motivation was not research- or work-related. This was a first in terms of the intention to bring social data visualization outside of a research setting and to a larger population of computer users.

Vizster matched the power of social network visualization to the practitioners of online social networking through several different views or interfaces. Figure 2.8 shows just one of the several presentations. The panel on the left displays an egocentric network view with the option for a keyword search. The panel on the right details the selected member's profile information. Jeff, for example, is from Stockton, CA. He lives in Berkeley, identifies as male, likes guitar and enjoys watching the

---

[56] J. Heer and d. boyd, "Vizster: Visualizing Online Social Networks," in *IEEE Symposium on Information Visualization*, 2005. INFOVIS 2005, 2005, 32–39, https://doi.org/10.1109/INFVIS.2005.1532126. A video demonstration of the system can be found at: https://vimeo.com/19278513.
[57] Heer and Boyd, "Vizster," 34.

History Channel. Words in the profile panel that appear in more than one profile will highlight on mouse-over; clicking on a highlighted word will initiate a search for that term. All members whose profile matches the search term will also become highlighted. Other views allow users to visualize gender, perform community analysis and explore connectivity.

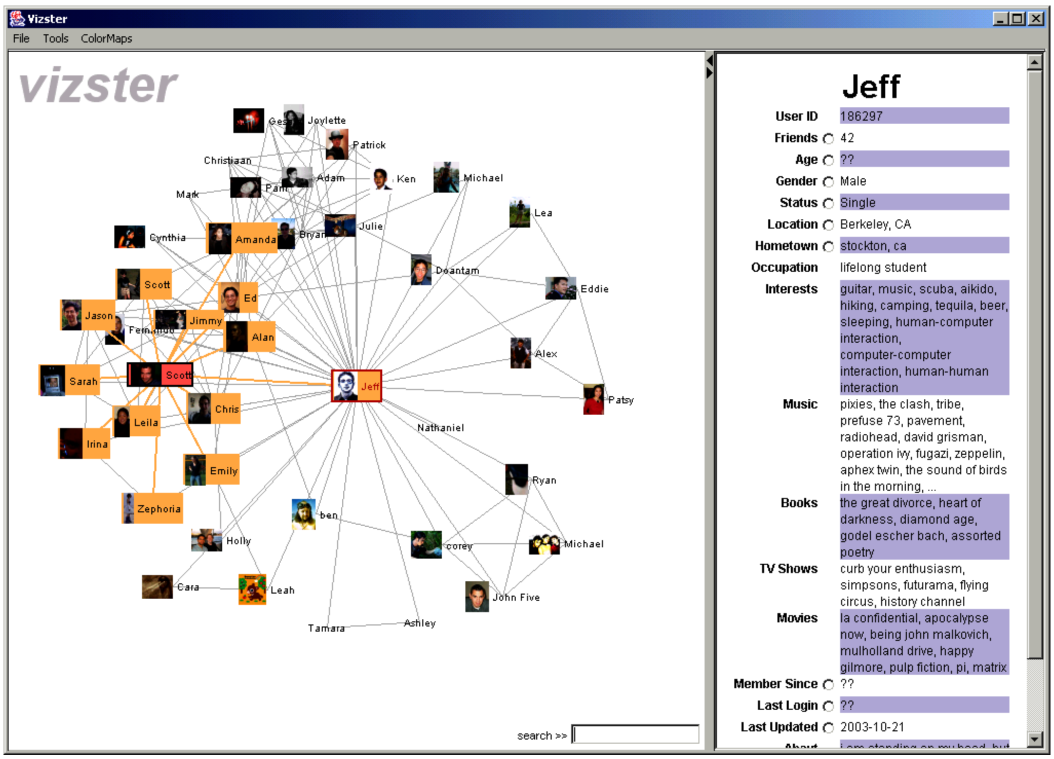

Figure 2.8. Vizster screenshot. Highlighted nodes show the friends that are shared between the two selected members.

Interestingly, in observing users, researchers found that interacting with the Vizster interface would often prompt exploration and further expansion of an individual's network through the discover of intermediary friends that were not yet visible.[58] While

[58] Heer and Boyd, "Vizster."

designed for Friendster end-users, the Vizster system never saw widespread use. Friendster famously declined in popularity following the introduction of competitive sites and the decision to turn down a $30 million buyout from Google in 2003.[59]

The concept, however, took hold. Similar visualization systems have been created for other social networking services. Nexus[60] and TouchGraph[61] are two Facebook applications that allow users to visualize and explore their network of friends. Both parse user profiles through the Facebook API and calculate layout based on different data dimensions. TouchGraph can actually be applied to any dataset that contains relationships. Interestingly, both applications as applied to Facebook are now defunct. Nexus found the application too niche and TouchGraph ran into trouble when Facebook changed its API and data retrieval policy. At the same time, Facebook outwardly supported development based on users' social graphs. Founder and CEO Mark Zuckerberg popularized the term at the Facebook F8 conference in May 2007 when he used it to explain how the newly introduced Facebook Platform would allow developers to take advantage of the networked connections between people using the service, thereby creating a richer online experience.[62] This dance between privacy, making the most of information, and the ephemerality of platform-specific services is

---

[59] Gavin Rivlin, "Wallflower at the Web Party," *New York Times*, October 15, 2006, http://www.nytimes.com/2006/10/15/business/yourmoney/15friend.html.
[60] http://nexus.ludios.net
[61] https://apps.facebook.com/touchgraph/
[62] "Social Graph," Wikipedia, November 24, 2016, https://en.wikipedia.org/w/index.php?title=Social_graph&oldid=751289213.

an area of social data visualization and the collection and distribution of data about people in general that is only starting to be investigated.

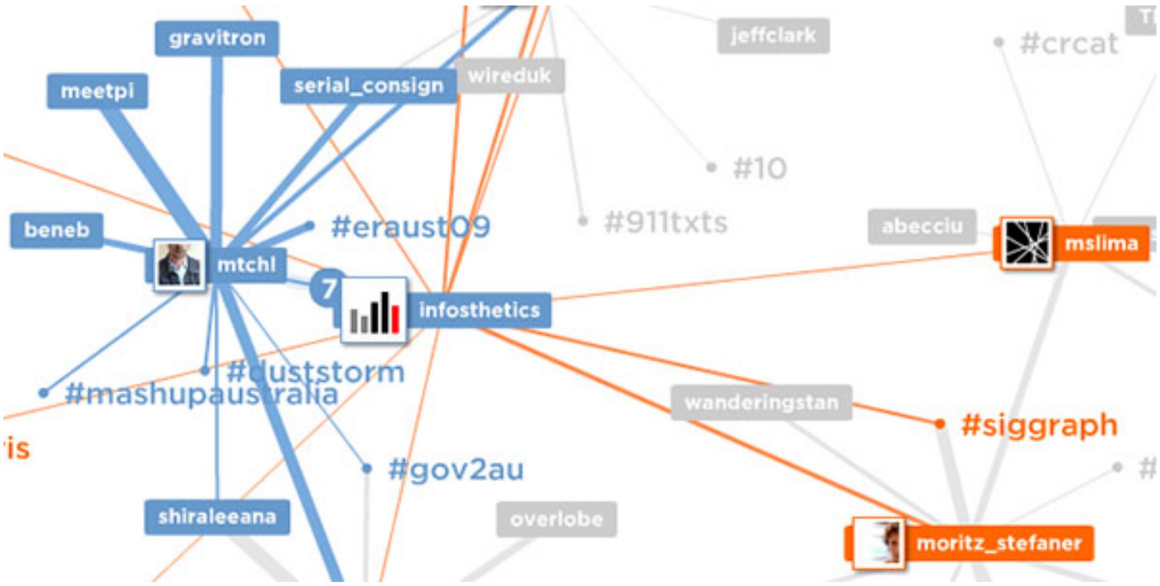

Figure 2.9. Detail of Mentionmap screenshot from infosthetics.com. Highlighted elements show the friends that are shared between the two selected members

Countless other tools and projects that visualize relationships on social media exist. Mentionmap is a web application that allows users to explore their Twitter network.[63] Users are presented with a network visualization of who they are talking to and what they are talking about. Figure 2.9 shows a screenshot of the Mentionmap for the data visualization's blog handle, @infosthetics. Mentions are represented as linear connections, and discussions between multiple users form clusters. The thicker the

[63] A. Vande Moere, "Mentionmap: Visualizing People and Conversations on Twitter," Information Aesthetics (blog), December 2009, http://infosthetics.com/archives/2009/12/mentionmap_visualizing_twitter_networks_and_interactions.html.

line, the more users have mentioned each other. The app additionally draws information from user profiles. Each node or line can be hovered over for further detail. Again, the data represented is drawn from Twitter's API. And again, the application no longer exists.

One example of a social data visualization tool that has persisted is Tweetsmap.[64] Tweetsmap provides similar network visualization to Mentionmap, but adds the functionality of publishing analytics. By monitoring and analyzing their Twitter campaigns, users can customize and strive to improve their social media presence. In a strange way, all of these different social data visualization tools act as interfaces to social interaction. People interact with and manage the connections to other people through these visual representations.

Facilitating Social Interaction

Many of the visualizations discussed thus far in this chapter do more than represent social relationships. They facilitate interaction around it. Recall that the primary goal of Soylent was to examine the social and temporal structures of collaboration left behind by user activity. Soylent was not designed as an end-user product but rather to aid in the design of a future co-working environment. While visualizing structure helped researchers to identify roles and relationships in social groups, it had an additional effect. Researchers found that seeing the visualizations of

[64] https://tweepsmap.com/.

activity helped the people within these groupings contextualize and find meaning in their work. Users would attribute connections between people and shifts in participation over time to particular stories and events. For example, one pattern might be explained by "the arrival of summer interns"; another by "the fall patent negotiation."[65] Heer and boyd found similar trends of off-screen engagement when observing Vizster's usage in both large-group and small-scale settings. Usage, particularly within the large-group setting, was often coupled with some form of social play. For example, people would challenge each other to hunt down somebody they knew. Observers tended to participate as well. Most users and observers would begin telling stories about specific people, connections and common experiences.[66] In other words, seeing the visualizations had very real social effects.

The ability to facilitate conversation and action around any type of data is another way to think about what is social about social data visualization. The project that largely spearheaded this line of thought was not a direct representation of social groups and relationships. In February 2005, Martin Wattenberg released a web-based visualization applet called the NameVoyager.[67] The project, which allows users to interactively explore the popularity of different baby names over time, was intended to help draw attention to his wife's new book.[68] It did. Within two weeks, the applet had

---

[65] Fisher and Dourish, "Social and Temporal Structures in Everyday Collaboration," 554.
[66] Heer and Boyd, "Vizster."
[67] http://www.babynamewizard.com/voyager
[68] The Baby Name Wizard (2005) by Lara Wattenberg offers a data-based approach to helping parents understand the trends and styles of American baby names.

drawn over 500,000 site visits. Two months later, it maintained an average of 10,000 visits per day. What surprised Wattenberg was not the widespread popularity of the site. What surprised Wattenberg was the way that the applet was used. In addition to serving as a tool, the visualization found itself at the center of an online social environment. [69] While certainly not the first project to do so, the NameVoyager set a precedent for seeing visualization as a social facilitator and activity.

To better understand what is meant by this, it is helpful to have an idea of the project's overall design. The NameVoyager (Fig. 2.10) visualizes a data set of 6,000 time series derived from the Social Security Administration's tracking of baby name trends in the United States. For each decade since 1900, the top one thousand boys and girls names are shown along with their rank. Boy names are colored blue and girl names are colored pink, in accordance with standard American convention. Users can interactively search the dataset by typing in a specific name or letter. Much like Ben Schneiderman's FilmFinder, the visualization dynamically responds. The project is compelling in and of itself and incorporates several techniques that are known to aid in viewer comprehension (e.g., interactivity, ease-of-use, an existing familiarity with the data).

[69] M. Wattenberg, "Baby Names, Visualization, and Social Data Analysis," in *IEEE Symposium on Information Visualization*, 2005. INFOVIS 2005, 2005, 1–7.

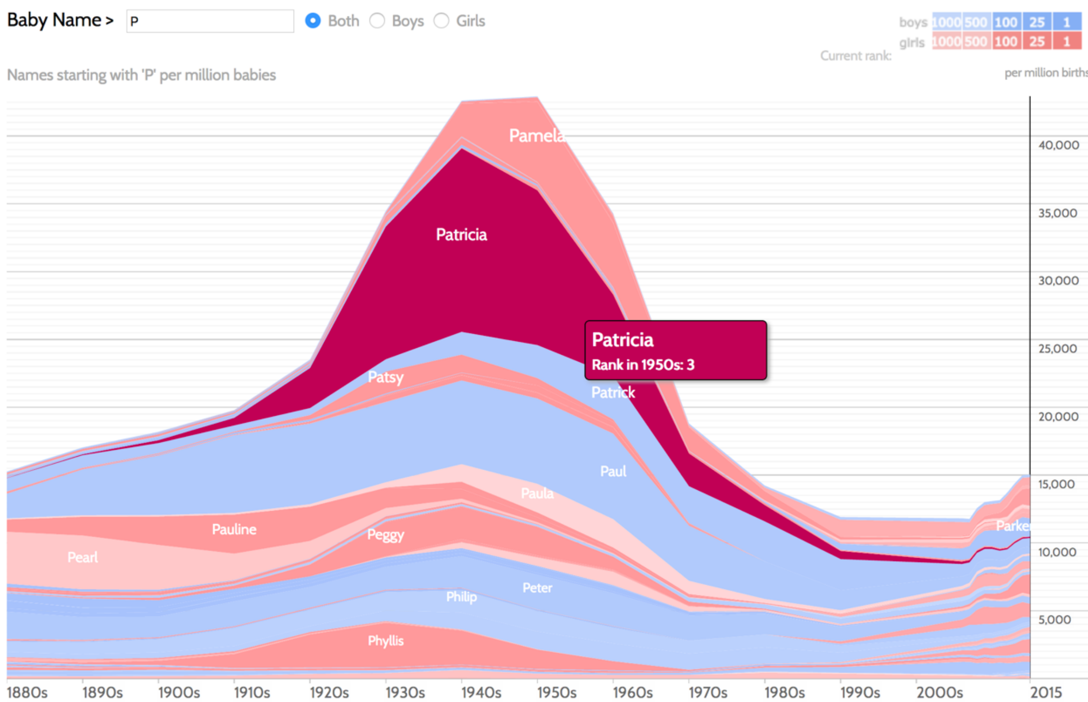

Figure 2.10. Martin Wattenberg, *Name Voyager*, 2005. Screenshot of dynamic response when searching for the letter "P" and the name "Patricia" selected.

Interestingly, much of this comprehension took place socially. Wattenberg noticed that soon after its release, people started using the NameVoyager to facilitate a type of "social" data analysis whereby using the applet or conversing about its dataset became a social activity.[70] At the most basic level, this has involved sharing or discussing the applet on the web. The NameVoyager has appeared in blogs, discussion forums, and on similar websites, even garnering attention outside its target audience.[71] User comments taken from public websites help to illustrate:

[70] Wattenberg, "Baby Names, Visualization, and Social Data Analysis," section 4.
[71] For example, see this post to an individual's personal blog: "Baby Name Voyager," Archshrk, August 15, 2007, http://archshrk.com/baby-name-voyager.

"This is perfect, as baby names weigh heavily on my mind these days."

"Surprisingly addictive"

"This rules, even if it's about baby names."[72]

User comments also reveal that groups of people have used the applet socially, as a sort of interface to collaborative data mining. The string between a group of users shown below illustrates how this can even take on a game-like quality:

> "For a challenge, try finding a name that was popular at the beginning of the sample (around 1900), went out of style, then came back into vogue recently."

> *"Take a look at Grace, #18 in the 1900s, #13 in 2003, and down in the 2000s and 300s during mid-century".*

> *"1900's comeback: Porter. Another one, with a mini-peak in through: Caroline".*[73]

The NameVoyager benefits from being embedded in an already existing cultural practice, that of choosing a baby name. The applet's success, both as a popular online visualization and as a social facilitator, undoubtedly benefits from this embeddedness as much as it does from an effective design. The shared cultural references provide a common ground from which conversations can emerge. In his paper on the project, Wattenberg hypothesizes that a combination of shared common ground and unique perspectives will help to encourage a social style of data analysis in any

[72] Wattenberg, "Baby Names, Visualization, and Social Data Analysis," section 3.2.
[73] Wattenberg, "Baby Names, Visualization, and Social Data Analysis," section 4.

visualization.[74] There are two additional reasons for the NameVoyager's social use. First, the social nature of the Web allows the applet to take on properties of the larger patterns of interaction already in use.[75] A blog begets comments. Second, the applet is able to take on the quality of a multiplayer online game. Wattenberg notes that users seem to fall into known categories of player types (i.e., explorer, achiever, socializer, or killer).[76]

As the Web has become increasingly social, the tendency to use data visualization to facilitate social interaction has only increased. People post visualizations to Twitter and Facebook, and conversation ensues. There are even entire services and groups dedicated to created communities around the visual representation of data. Another of Wattenberg's projects, Many Eyes, is a platform designed to support collaboration and conversation around visualization on a large scale.[77] Users can upload data, create interactive visualizations, and carry on discussions in a centralized place. We see similar communities form around competitions and tools. Tableau, one of the most popular visualization software packages, posts a "Viz of the Day" drawn from the many user projects submitted to its

---

[74] Wattenberg, "Baby Names, Visualization, and Social Data Analysis," section 5.1.
[75] While Wattenberg does recognize the importance of the public-nature of the web-based application, he seems to paint it as more of a user study benefit than a property of the visualization's context.
[76] Wattenberg, "Baby Names, Visualization, and Social Data Analysis," section 4.
[77] Fernanda B. Viégas et al., "Many Eyes: A Site for Visualization at Internet Scale," IEEE Transactions on Visualization and Computer Graphics 13, no. 6 (2007): 1121–1128.

gallery online.[78] Similarly, competitions like the VAST Challenge,[79] the Kantar Information Is Beautiful Awards,[80] and NASA's Visualization Storytelling Competition[81] unite designers and fans around outstanding projects.

Visualization and the Everyday

Artistic or aesthetic and social application of data visualization are just two broad areas within the types of practices and research that have developed over the past fifteen years. They are not always distinct. Many of the projects that could be classified as artistic data visualization also serve social functions. The reverse—social data visualizations that function as art—is more unusual. Such projects certainly do exist. Josh On's *They Rule* (2004), a project that interactively visualizes the invisible networks of power between individuals who are on the boards of multiple corporations, has been exhibited at Ars Electronica 2004 as well as in MoMA's *Talk to Me*, among many other exhibitions.[82] The connections between people that define our lived experience and culture are common subjects of interest for art. However, I would argue that what is occurring in this crossover is something more than inspiration. What is going on across all aspects of contemporary data visualization's expansion since the

---

78 https://public.tableau.com/en-us/s/gallery
79 The VAST Challenge is day-long competition that takes place during IEEE VizWeek each year. For more information about the challenge and past winners, see: http://www.vacommunity.org/About+the+VAST+Challenge
80 https://www.informationisbeautifulawards.com/about
81 https://www.nasa.gov/feature/agu-data-visualization-storytelling-competition
82 http://www.theyrule.net.

2000s is that data visualization is becoming another medium through which we communicate and experience life. The popularization and pervasion of data visualization is yet another area of expansion within contemporary data visualization's growth. There are a handful of projects and papers that have tried to engage with this.

Casual and Vernacular Visualization

Casual and *vernacular* visualization are two terms that highlight the untraditional use and user-base of information visualization. Both are drawn from specific papers and both have been used within the visualization community to refer to recent developments or trends. Casual is defined by visualization tools or systems that are designed to be used in casual rather than work-related settings. Pousman et al. include ambient, social, and artistic visualizations as well as other edge-case scenarios in this grouping.[83] The untraditional setting often correlates to a wide range of user populations. Casual visualizations are frequently geared toward the non-expert user and the everyday. A good example is the Ambient Orb™, which is a commercially available glass ball that changes color based on the movement of the stock market.[84] It is part of a larger class of peripheral displays that allow users to monitor remote activity while focusing on a separate, primary task.

---

[83] Zachary Pousman, John Stasko, and Michael Mateas, "Casual Information Visualization: Depictions of Data in Everyday Life," in *IEEE Transactions on Visualization and Computer Graphics* 13, no. 6 (2007).
[84] "Ambient Devices - Stock Orb," accessed September 26, 2017, https://ambientdevices.myshopify.com/products/stock-orb.

Vernacular refers more directly to the general user, as well as the visualization technique. Martin Wattenberg and Fernanda Viégas use the term to describe tag clouds, which they see as a type of "street-wise" visualization technique.[85] Tag clouds, or word clouds as they are sometimes called, visualize the frequency of words in a text by aggregating words and varying font style.[86] The technique, which became immensely popular in the first decade of the 2000s, has origins outside the visualization community, hence, the term "vernacular." The label of vernacular visualization also has associations with mass media and participatory culture.[87]

Participatory Visualization and Popular Culture

It is not just that data visualization has pervaded everyday life by seeping outside of work settings. General users have gained the ability to create visualizations on their own. Projects like IBM's Many Eyes or the tag cloud applets that Wattenberg and Viégas called "vernacular" afford users the ability to import data and make select design decisions about how it should be visualized. That is, they are given the opportunity to participate. Relatively few and minimalistic in the early 2000s, such software and platforms are now widely used. Excel graphs, Google Charts, Tableau

---

[85] Fernanda B. Viégas and Martin Wattenberg, "TIMELINES: Tag Clouds and the Case for Vernacular Visualization," Interactions 15, no. 4 (July 2008): 49–52.
[86] One popular tag-cloud applet is Wordle; which can be found at http://www.wordle.net.
[87] F. B. Viégas, M. Wattenberg, and J. Feinberg, "Participatory Visualization with Wordle," in *IEEE Transactions on Visualization and Computer Graphics* 15, no. 6 (November 2009): 1137–44.

and d3, and the sophisticated JavaScript library for visualization data on the Web, are all examples of such participatory visualization.[88] There are even entire companies dedicated to helping people visualize their data. Visual.ly is an online service that doubles as a showcase for infographics and a marketplace for the people who can help others create them.[89] The word "participatory" is drawn from Henry Jenkins's concept of participatory culture, which refers to a cultural system in which media consumers are also media producers.[90] He defines a participatory culture as one:

- with relatively low barriers to artistic expression and civic engagement;
- with strong support for creating and sharing one's creations;
- with some type of informal mentorship whereby what is known by the most experienced is passed along to novices;
- in which members believe their contributions matter;
- in which members feel some degree of social connection with one another (at the least they care what other people think about what they have created).[91]

This is worlds removed from the type of visualization projects and culture that defined the 1980s and '90s.

What changed? It is not so much that data visualization moved away from its analytic basis, but that data and analysis, and thereby data visualization, have become integral parts of contemporary culture. In the introduction to their article written for the

[88] Viégas, Wattenberg, and Feinberg.
[89] https://visual.ly
[90] Henry Jenkins, *Convergence Culture: Where Old and New Media Collide* (NYU Press, 2006).
[91] Henry Jenkins et al., "Confronting the Challenges of Participatory Culture: Media Education for the 21st Century," white paper, Building the Field of Digital Media and Learning (The MacArthur Foundation, 2006), https://www.issuelab.org/resources/830/830.pdf.

22nd Malofiej Infographic Awards, Wattenberg and Viégas go as far as to declare visualization a mass medium:

> Visualization is now a mass medium. It's not quite Hollywood, but information graphics have millions of viewers, awards ceremonies, and even their own celebrities with tens of thousands of Twitter followers. More important, from the perspective of journalism, is that data visualization is an essential part of the communication process. Today, a data-driven story without a chart is like a fashion story without a photo.[92]

The expansion of data visualization into nearly every dimension of life also includes the growth of several other untraditional types of data visualization that I have not yet had the opportunity to mention. Visual journalism, which involves the use of images and graphics to tell stories, has become a standard feature of journalistic investigation. *The New York Times*, *Washington Post*, and *The Guardian*, among many other internationally recognized news and media companies have entire teams dedicated to creating interactive graphics that engage and inform their audiences through visual explanations.[93] Visual journalism also has ties to narrative visualization and infographics. I will only briefly mention it here as its function and expanded use has been written about extensively by other scholars.[94]

There is also a growing class of visualization projects and platforms that aim to engage the public, either through data collection or social awareness or both. Two

---

92 Wattenberg and Viégas, "Design and Redesign."

93 The BBC website offers an in-depth definition of visual journalism that while specific to their own content, applies to other publications and companies. See, Amanda Farnsworth, "What Is Visual Journalism?" BBC News, May 10, 2013, http://www.bbc.com/news/blogs-the-editors-22483705.

94 Alberto Cairo, *The Functional Art: An Introduction to Information Graphics and Visualization*, (Berkeley, California: New Riders, 2012).

very different examples of this visualization type are The Valley of the Khans and Ushahidi.[95] The Valley of the Khans is an international collaboration between several institutions, including the National Geographic Society, to perform high-tech, non-invasive remote sensing to locate the tomb of Genghis Khan. In 2010, the project turned to crowdsourcing as one possible location technique. The public was invited to interact by downloading super-high-resolution satellite images of Mongolia and tagging parts of the terrain that look like they could contain a man-made burial site.[96] Ushahidi, which translates to "testimony" in Swahili, offers a similar form of civic engagement. This platform first emerged in 2008, following reports of violence in the aftermath of Kenya's disputed presidential election. Local observers were invited to submit eyewitness reports using their mobile phones or the Internet. The crowdsourced data was then mapped to a website for the public to see. The Ushahidi platform has since become a widespread tool for mapping natural disaster responses, human rights reporting, and election monitoring, among many other types of crisis situations.[97]

Figure 2.11 shows a screenshot of the Ushahidi platform as it was used following the 2010 earthquake in Haiti. Viewers can filter by type of report as well as hover over individual reports to get further details. The visualization is well designed. However, there is nothing particularly innovative or compelling about it. What is

---

95 https://www.ushahidi.com/

96 "Project - Valley of the Khans," IGERT-TEECH, accessed September 21, 2017, http://igert.ucsd.edu/igert-teech/projects/valley_of_the_khans.php.

97 Ushahidi Staff, "Where in the World," Ushahidi (blog), August 5, 2013, https://www.ushahidi.com/blog/2013/08/05/where-in-the-world.

compelling is that users are invited to participate as citizen scientists in large community projects. The projected expectation—that is, what viewers are presumably led to believe—is that they are helping out in a humanitarian response.

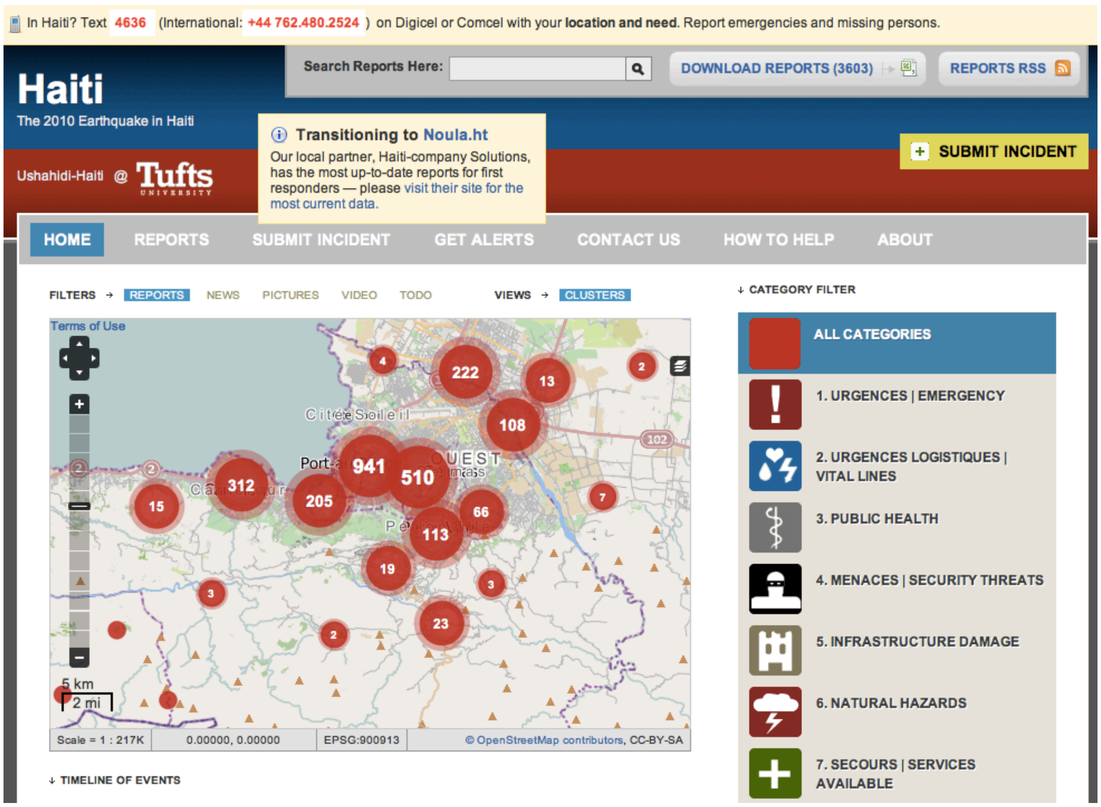

Figure 2.11. Screenshot of Ushahidi platform (2010).

There is likewise an aspect of social awareness. Even those viewers who are not on-the-ground observers are made aware of the scale and seriousness of the crisis situation. Whether such projects are effective in helping people on the ground is the subject of much debate, along with ethical questions about accuracy, privacy, and access. Similar mapping platforms have been used to document cultural sites or tell the stories of events and communities. HyperCities, which is a Google Earth-based digital humanities mapping platform, has been applied to projects as diverse as reconstructing ancient Rome, documenting social media during the Arab Spring, and empowering youth through digital media in underprivileged communities.[98]

Data Humanism and the Humanities

The expansion of data visualization into nearly every dimension of life, combined with repeated challenges to traditional visualization expectations and norms, has led to the realization that existing approaches to data visualization research and design are out of date. In 2017, Giorgia Lupi gave the capstone address at IEEE VIS, one of the largest and most important visualization conferences of the year. Her talk, titled “Data Humanism – The Revolution Will Be Visualized,” largely encapsulates this momentum to move forward.[99] Lupi has written about her ideas on data humanism, differentiating between a first and second wave of data visualization.[100] The first wave, which spans an unspecified number of years, was characterized by the pervasion of data visualization, broader audiences, and what Lupi argues is a superficial and merely technical approach to representing and understanding big data. The second

---

98 http://www.hypercities.com/; Todd Presner, David Shepard, and Yoh Kawano, HyperCities: Thick Mapping in the Digital Humanities, 2014 edition (Cambridge, MA: Harvard University Press, 2014).
99 Giorgia Lupi, “Data Humanism – The Revolution Will Be Visualized” (IEEE VIS 2017, Phoenix, AZ, October 6, 2017), https://www.youtube.com/watch?v=S0YkTtLFIDs.
100 Giorgia Lupi, “Data Humanism: The Revolutionary Future of Data Visualization,” Print Magazine, January 30, 2017.

wave, which she implies needs to begin right now, will “question the impersonality of a merely technical approach to data, and to begin designing ways to connect numbers to what they really stand for: knowledge, behaviors, people.”[101] In other words, the second wave will recognize that data represent real life and therefore need to be made more personal, more human. Lupi offers some suggestions on how we might do this, but the amount of research and scholarship on how and why remains minimal.

An area of potential interest that has emerged over the past few years is the intersection between data visualization and the humanities. Just like artistic data visualization and social data visualization, digital humanities projects often push what is considered acceptable content and use for data visualization.[102] There have also been a handful of workshops and papers that have discussed what the humanities can bring to visualization, and vice versa. There are two groupings that fall under the category of humanities-based approaches to visualization.

The first, which also has ties to the field of design, advocates for a more critical approach to visualization.[103] Critical, here, is a term borrowed from both critical theory and critical approaches to knowledge. Critical theory is a social theory first defined by Max Horkheimer in 1937 that attempts to critique and transform society as a whole,

---

[101] Lupi, "Data Humanism: The Revolutionary Future of Data Visualization."
[102] Anne Burdick and Johanna Drucker, Digital_humanities (Cambridge, MA: MIT Press, 2012). Specific papers and events are cited in the following paragraphs.
[103] Marian Dörk et al., “Critical InfoVis: Exploring the Politics of Visualization,” in CHI’13 Extended Abstracts on Human Factors in Computing Systems (ACM, 2013), 2189–2198; “Critical Visualization Network Workshop,” accessed November 19, 2017, http://torch.ox.ac.uk/critical-visualization-network; “Critical Visualization Event at CUNY” accessed November 19, 2017, http://patriksv.net/2016/04/critical-visualization-event-at-gc-cuny-on-june-9-2016/.

often by identifying assumptions and norms in relation to power. Critical approaches to knowledge are more varied. They are united by the fact that they are often reflexive about the tools, methodologies and frameworks they employ. Critical notions of pedagogy, for example, challenge the supposed neutrality of education. They reject ideas that "The Teacher knows everything and the students know nothing" and instead attempt to empower students to think and evaluate for themselves.[104] Critical cartography and geography challenge the scope of the discipline by linking maps and other representations of geographic knowledge to issues of social, environmental and political change. For critical cartographers, maps are only perceived neutralities.[105] They see ideologies and biases as hidden in between the lines. It is therefore possible for maps to serve social and political interests.[106] Critical visualization aims to uncover and question taken-for-granted assumptions about data visualization and reflect on how and what values are portrayed.

The proposal of a feminist data visualization takes on many of the same characteristics. Those who argue for a feminist approach to visualization research and design question how feminist principles and theory might be applied to the design process as well as the experience of viewing a visualization. Feminism, in this case, is not (just) about women, but rather about deviating from dominant ways of producing

---

[104] Henry A. Giroux, "Lessons From Paulo Freire," *Chronicle of Higher Education*, October 17, 2010.
[105] John Brian Harley, "Deconstructing the Map," *Cartographica: The International Journal for Geographic Information and Geovisualization* 26, no. 2 (1989): 1–20.
[106] Wood, *The Power of Maps*, (New York: The Guilford Press, 1992).

and communicating information.[107] Feminist approaches in other fields such as Science and Technology Studies and Human Computer Interaction deploy feminism to draw attention to and empower people and perspectives that are marginalized.

What data humanism or critical and feminist approaches to data visualization might look like is a question that does not have a definitive answer. However, some suggestions have been made (See Table 2.1). In outlining her vision for a data humanist future, Lupi puts forth several design principles or paths forward. Each embodies the desire to make data more intimate, unique, and relatable. Careful and thoughtful data visualization is seen as a way of doing this. The suggestions offered by critical and feminist approaches to visualization reflect similar strategies and values. Scholars who have advocated for a more critical approach to data visualization point to disclosure, plurality, contingency, and empowerment as principles for designing and viewing visualization through a critical lens. Feminist data visualization offers like principles, focusing on incorporating pluralism, as well as reflections on empowerment, affect, and visible labor as part of the design process and output. While the specific words used by each approach may vary, there is significant overlap. The italicized words in Table 2.1 are intended to emphasize this. The intent of creating more

[107] Catherine D'Ignazio, "What Would Feminist Data Visualization Look Like?," MIT Center for Civic Media (blog), December 20, 2015, https://civic.mit.edu/feminist-data-visualization; Catherine D'Ignazio and Lauren F. Klein, "Feminist Data Visualization," workshop on Visualization for the Digital Humanities (VIS4DH, Baltimore, MD, 2016); Lauren F. Klein, "Feminist Data Visualization; Or, the Shape of History" (MLA 2017, Philadelphia, PA, January 2017).

thoughtful and complex visualizations that reflect the reality of data and the experience of viewing is largely the same.

Table 2.12. Principles for Alternative Approaches to Data Visualization

Data Humanism[a]	**Critical Infovis**[b]	**Feminist Data Visualization**[c]
* Embrace *complexity* * Move *beyond standards* * Sneak *context* in (always) * Remember that *data is flawed* (as we are)	* *Disclosure* * *Plurality* * *Contingency* * *Empowerment*	* Rethink *binaries* * Embrace *pluralism* * Examine *power* and aspire to *empowerment* * Consider *context* * Legitimize embodiment and affect * Make *labor* visible

[a] Lupi, "Data Humanism: The Revolutionary Future of Data Visualization."
[b] Dörk et al., "Critical InfoVis: Exploring the Politics of Visualization."
[c] D'Ignazio and F. Klein, "Feminist Data Visualization."

The second part of this dissertation follows the same trajectory but takes an alternative approach. Building on the history of data visualization's expansion and the recognition that the way we represent and understand data needs to change, I offer close readings of how select contemporary data visualization projects function. Understanding how and why we use data visualizations and their relationship to the environments in which they are created and displayed is the first step towards conceptualizing how we view and interact with them. Data visualizations are not simply objective, analytic representations of abstract numbers, and it is time we stop seeing them as such.

PART II

CHAPTER 3

The *Racial Dot Map*: Context, Interaction, and Building Meaning

July 2013. Dustin Cable publishes the *Racial Dot Map* (Fig. 3.1).[1] Almost immediately, the mainstream media and visualization community begin to respond. People talk about the map on Facebook and Twitter. *Wired* calls it "The Best Map Ever Made of America's Racial Segregation."[2] *Slate* challenges, "Can You Find Yourself on this Map?"[3] *The Atlantic*, along with several visualization blogs, notes its use of cartographic pointillism, describing it as "strangely beautiful."[4] Indeed, the visualization is reminiscent of a painting by George Seurat.

Composed of over three million individual colored dots, the *Racial Dot Map* depicts racial identification and distribution across every neighborhood in the United States. Each dot represents a single person living in the country at the time of the 2010 U.S. Census. Dots are color-coded according to self-identified race. Blue corresponds to white while green represents black. Asian is colored red, Hispanic is orange, and brown is used to denote any other race. Viewers can zoom in or out to

---

[1] https://demographics.virginia.edu/DotMap/index.html

[2] Kyle Van Hemert, "The Best Map Ever Made of America's Racial Segregation | Wired Design | Wired.Com," *Wired Design*, accessed February 5, 2014, http://www.wired.com/design/2013/08/how-segregated-is-your-city-this-eye-opening-map-shows-you/.

[3] Jeremy Stahl, "This Amazing Map Shows Every Person in America," *Slate*, August 15, 2013, http://www.slate.com/articles/news_and_politics/map_of_the_week/2013/08/segregation_in_america_every_neighborhood_in_the_u_s_mapped_along_racial.html.

[4] Emily Badger, "A Strangely Beautiful Map of Race in America," *The Atlantic Cities*, August 14, 2013, http://www.theatlanticcities.com/neighborhoods/2013/08/strangely-beautiful-map-race-america/6534/.

change the area and scale of what is represented. As they do, trends and disparities in the data become clear. Dots, which are densely packed on the eastern seaboard, become noticeable sparser moving west. New gradations of color and population patterns are introduced. What looks like a smudge of blue and purple at the state-level turns into sections of cities divided by race, up close. Zoom in further and race seems to change with every city block.

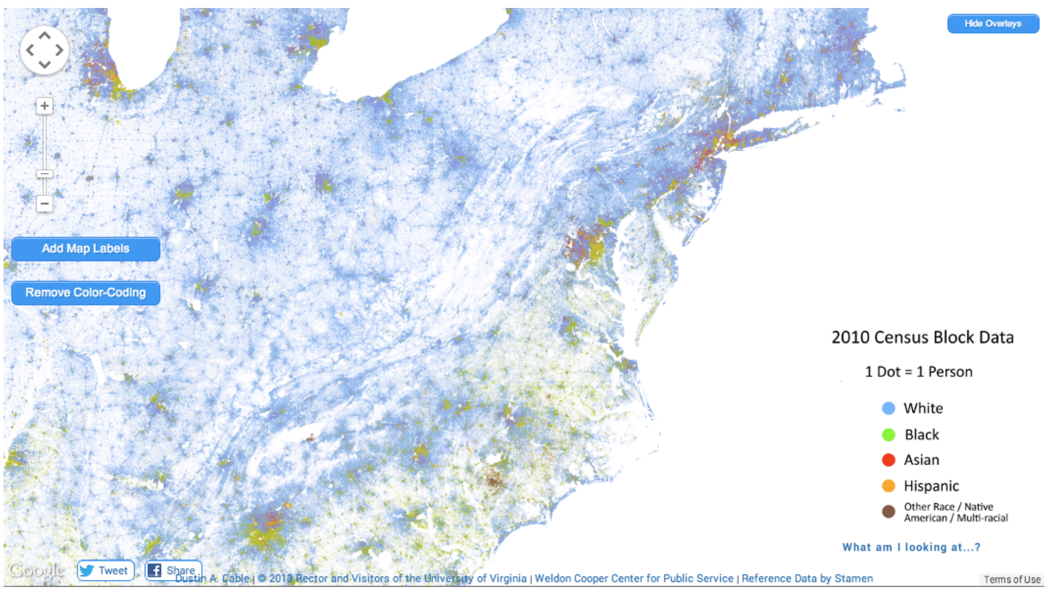

Fig. 3.1. Screenshot from the *Racial Dot Map*, Dustin Cable (2013). This screenshot shows the entry view presented to the viewer when first navigating the project's website.

The experience of looking at the *Racial Dot Map* is a bit like looking at a documentary or work of abstract art. There are elements that are recognizable, but much of the visual information presented is new or not yet understood. In many ways, this is true of the experience of looking at any data visualization. What makes the

Racial Dot Map unique is the personal identification it evokes and the context in which it was created and seen. As a U.S. citizen or citizen of any nation who has experienced racial tension, it is difficult to look at the representation and not want to explore it. It is difficult to look at the representation and not map one's personal experience and perception onto it. And it is difficult to look at the representation and not draw social meaning and significance from it. While certainly true of the way we look at other types of visual media, each of these characteristics is not so often true of the way we look at data visualization. This chapter closely reads the *Racial Dot Map* in the way one might read a documentary photograph. It looks at context, authorship, intertextuality, and the relationship between viewer, representation and space, to tease out the variety of different ways it is possible to interact with contemporary data visualization.

Context

One aspect of interaction that is often discussed in relation to photographs and film, but not in relation to data visualization, is context. Context shapes the environment in which any visual medium is seen. Photography, documentary, and even art get meaning from what has been written about them, the events going on around them, other text or objects, viewers' awareness, and the discussions going on about them and their subject matter.[5] For example, Dorothea Lange's photograph "Migrant Mother" (1936), perhaps the most iconic documentary image, is impossible to view without taking into account the motivation of social reform and the life that the photograph has had. The image has been retitled, reproduced, and widely embraced as a universal symbol of poverty and human suffering.[6] The experience of viewing it necessarily entails knowledge of this background. Data visualization has been criticized for imposing a lack of context on the data.[7] There is no particular reason why this has to be the case, and there has been recent interest and discussion on how to best incorporate context, particularly in relation to storytelling and narrative visualization.[8]

While the *Racial Dot Map* does not explicitly incorporate context into its representation, it relies on context nonetheless. July 2013 marked a particularly heightened time for Americans in relation to race. Eight months earlier, Barack Obama had been elected to his second term as President of the United States. As the

---

[5] Howard S. Becker, "Visual Sociology, Documentary Photography, and Photojournalism: It's (Almost) All a Matter of Context," *Visual Sociology* 10, no. 1–2 (January 1, 1995): 8, https://doi.org/10.1080/14725869508583745.
[6] James C. Curtis, "Dorothea Lange, Migrant Mother, and the Culture of the Great Depression," *Winterthur Portfolio* 21, no. 1 (1986): 1.
[7] Nick Diakopoulos, "Context for Data Visualization," *Nick Diakopoulos: Musings on Media* (blog), accessed November 28, 2017, http://www.nickdiakopoulos.com/projects/context-for-data-visualization/.
[8] Eui-Chul Jung and Keiichi Sato, "A Framework of Context-Sensitive Visualization for User-Centered Interactive Systems," in *International Conference on User Modeling* (Springer, 2005), 423–427; Enrico Bertini and Moritz Stefaner, *Visual Storytelling w/ Alberto Cairo and Robert Kosara*, Data Stories, accessed November 28, 2017, http://datastori.es/data-stories-35-visual-storytelling-w-alberto-cairo-and-robert-kosara/; Robert Kosara, "The Importance of Context," *Eagereyes* (blog), July 18, 2017, https://eagereyes.org/blog/2017/the-importance-of-context; Lupi, "Data Humanism: The Revolutionary Future of Data Visualization."

country's first and only non-white president, his election seemed to signal a step away from corrosive racial politics and towards equal opportunity for people of any race.[9] His reelection in November 2012 seemed to signal a continuation along the same path. Obama had remained relatively silent on race relations during his first term until the killing of Trayvon Martin in February 2012 made talking about the subject unavoidable. Martin, an unarmed African-American teenager, had been fatally shot by George Zimmerman, a 28-year-old man of mixed Hispanic descent. Martin had been visiting relatives in the gated community where Zimmerman served as neighborhood watch coordinator. On the evening of February 26, 2012, Zimmerman called the police to report a suspicious person wearing a "dark hoodie." He was not on active duty at the time. Zimmerman then followed the individual and was asked to remain in his car until the police arrived. At some point thereafter, a violent altercation occurred, resulting in Martin's death.[10] Zimmerman argued that Martin attacked him, pleading self-defense.

The incident sparked a national debate. Shortly following the shooting, President Obama issued a statement, reminding the American public that "if [he] had a son, he'd look like Trayvon."[11] Issues of racial profiling and minority bias came to the

---

[9] Michael Tesler and David O. Sears, *Obama's Race: The 2008 Election and the Dream of a Post-Racial America*, Chicago Studies in American Politics edition (Chicago: University of Chicago Press, 2010).

[10] The media coverage of the Martin-Zimmerman case is substantial. CBS News has put together a centralized website that links to many of the different articles and perspectives published as a result of the shooting and trial. For more, see http://www.cbsnews.com/feature/george-zimmerman-trial-trayvon-martin-case/.

[11] Byron Tau, "Obama: 'If I Had a Son, He'd Look like Trayvon,'" *POLITICO*, accessed June 12, 2017, http://www.politico.com/politico44/2012/03/obama-i-had-a-son-hed-look-like-trayvon-118439.html.

forefront. Some sided with Zimmerman, using the opportunity to highlight the need to protect one's community and gun rights. Others saw the incident as a clear case of racial profiling and used the opportunity to address what they saw as a problematic relationship between the African American community and the police.[12] Others stayed silent, not falling into either camp. Regardless of what people believed, the shooting and trial had an enormous effect on the American public. Discussions of race and bias pervaded the news. The United States entered a period that is still underway when discussions of race and bias have ignited following a related or like event.

Two such events occurred in the month leading up to the publication of the *Racial Dot Map*. In late June 2013, the U.S. Supreme Court ruled 5-4 in favor of striking down a section of the Voting Rights Act (VRA) that protected against racial discrimination in the voting booth.[13] The decision was contentious. It hinged on the question of whether race relations had changed dramatically enough since the act's enactment in 1965 to do away with federal oversight of select states. The court ruled yes. Many of the majority-opinion justices made a point of clearly condemning any racial discrimination in voting. Still, a large portion of the population saw the ruling as a step backward for civil rights.[14] Then, on July 13, 2013, George Zimmerman was found

---

[12] Krissah Thompson and Jon Cohen, "Trayvon Martin Case: Poll Finds Stark Racial Divide," *Washington Post*, April 10, 2012, https://www.washingtonpost.com/politics/trayvon-martin-case-poll-finds-stark-racial-divide/2012/04/10/gIQAEETX8S_story.html?_r=3&ref=us&utm_term=.7159d177ea09.

[13] U.S. Supreme Court, Shelby County, Alabama v. Holder, Attorney General, et al., No. 12-96 (U.S. Supreme Court June 25, 2013).

[14] Stephanie Drahan, "LWV Reacts to Supreme Court Decision on the Voting Rights Act | League of Women Voters," June 5, 2013, http://lwv.org/press-releases/lwv-reacts-supreme-court-decision-voting-rights-act.

not guilty of second-degree murder. His acquittal in the case stemming from the shooting of Trayvon Martin reignited vigorous debates. At the same time, it sparked calls for peaceful action and calm.[15]

None of these events are explicitly represented by the data or visual elements shown in the *Racial Dot Map*. However, the context of a nation embroiled in race relations remains. In July 2013, there was a communal need to take a hard look at conceptions of race, regardless of one's individual social or political beliefs. This underlying motivation for reflection and proof undoubtedly influences both the project's design and reception. Whether Cable or the viewer consciously make this connection is beside the point. The point is that published at a different time, the project would evoke a different reaction.

Another way of saying this is that the *Racial Dot Map* has a certain sense of timeliness.[16] It reflects its surrounding rhetorical situation, perhaps not consciously, but it is tied to it nevertheless. This awareness extends beyond the content and concept of race. Dustin Cable was working as a researcher at the University of Virginia's Weldon Cooper Center for Public Service at the time he created and published the *Racial Dot Map*. The center's stated mission is "to strengthen and preserve effective government in Virginia by leveraging the research, expertise, and resources of the University of

[15] Greg Botelho and Holly Yan, "George Zimmerman Found Not Guilty of Murder in Trayvon Martin's Death," *CNN*, July 14, 2013, http://www.cnn.com/2013/07/13/justice/zimmerman-trial/index.html.
[16] In classic and modern rhetoric, the idea that there is an opportune time for an act to occur based on surrounding circumstances and rhetorical situation is known as *kairos*.

Virginia to inform public policy, develop public leaders, and deliver strategic and technical assistance."[17] This is also part of the project's context. In an informal interview, Cable explains that a significant motivation in created the *Racial Dot Map* was to show people that collecting and interpreting census data matters:

> I've been disappointed by the conversation that is currently taking place about the value of census data and whether or not the decennial census should even continue to measure things like race, age, or gender. We've also recently had government officials and lawmakers threatening to get rid of these data collection efforts out of fears of invading privacy, or worse, stating that race does not matter like it used to. I think that one of the reasons this message resonates with people is that many don't know the value of data and what it says about the country and ourselves…. One of my goals for *the Racial Dot Map* project was to lift this veil and show, in a powerful way, how our communities are shaped by race and ethnicity.[18]

Cable's answer references a related question that concerned him at the time and that is, what is the value of census data? I have included a substantial part of the interview here because it clearly states how the issue of race and questions of the value of data intersect. The *Racial Dot Map* was published and viewed at a time when the country felt a need to look at local and national trends in relation to race. This connection—that is, the relationship between the desire to look at race in America and the conditions under which race is perceived, comprehended and received—is precisely what makes the visualization so effective. In other words, how and what we see in a data visualization is influenced by the environment in which it is created and received.

[17] "About," *Weldon Cooper Center for Public Service*, accessed November 29, 2017, https://coopercenter.org/about.
[18] Dustin Cable, Interview with Dustin Cable, e-mail correspondence, November 22, 2013.

If the Racial Dot Map had been published at a different time, if it represented a less comprehensive dataset, and if it had not been associated with a university public service center, the experience of seeing it would be different.

The Viewing Experience

Context is also drawn from viewer experience. To attribute the *Racial Dot Map*'s response and effectiveness solely to its contemporaneous social situation would be inaccurate and simplistic. Several aspects of the project's design and opportunities for viewer interaction add context to the data through other associations. The following section looks at three different ways that the map gains meaning—through exploration, personal identification, and zooming. All work in unison, often relying on each other to build significance and meaning.

Exploration

Like many contemporary data visualization projects, the *Racial Dot Map* is designed to be viewed online. When the website first loads (Fig. 3.1), the viewer sees a scattering of fine dots. Most are blue, but others appear to be purple or green. The dots in certain areas are densely packed. Other areas are contrastingly sparse. A stippled outline hints that what you are looking at is a map of the Northeastern part of the United States. The legend in the bottom, right-hand corner provides further detail. What you are looking at is a visual representation of 2010 census block data. Each dot represents a single person living in the United States. Dots are color-coded according to self-identified race. Clicking on the unassuming link “What am I looking at?...” opens a new page. Information about how the map is made and what patterns the data shows are laid out in close-ups and text.

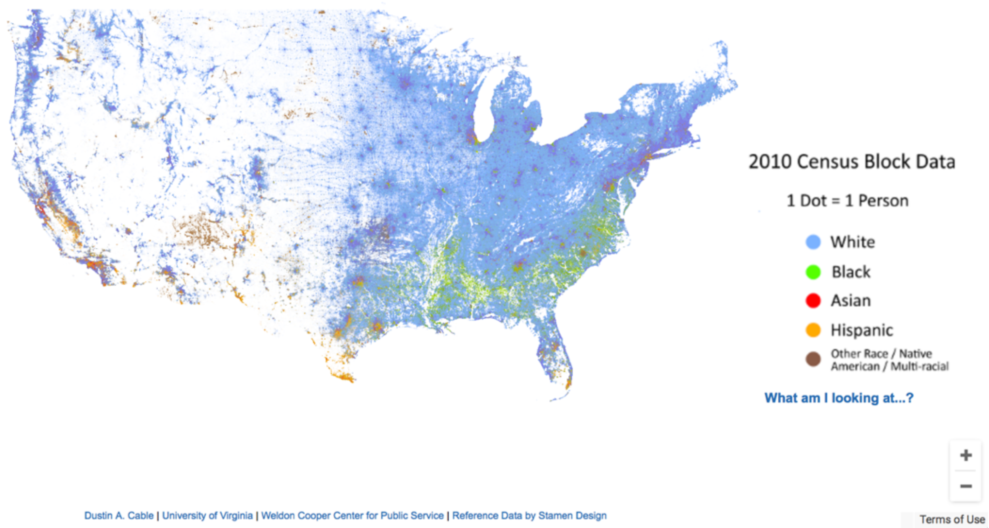

Fig. 3.2. Screenshot from the *Racial Dot Map* zoomed out to show the entire country.

Returning to the map invites use. The interface is familiar, borrowing from the Google Maps API to represent the underlying geographical shape.[19] Users are invited to explore the map by zooming in and out through the direct manipulating of simple graphical controls. Zooming out reveals thousands of more dots. In total, there are 308,745,538. With each click, the map of the country becomes more complete. The ability to discern between different-colored dots becomes difficult. Smudges of

[19] For more information about the Google Maps API and how it displays geographical information, see the Google Maps API website at https://developers.google.com/maps/.

composite color or lack thereof begin to appear. Countrywide trends begin to take shape. The Eastern half of the United States is distinctly more populated than the West. Dots become sparser moving west and only pack together again tightly at the coast. The southernmost tip of Texas and many areas in the Southwest introduce colors that were not present in large numbers on the East coast. Smudges of green and purple give way to varying shades of orange and red (Fig. 3.2). Zooming in offers a remarkably different representation. Color patches are once again discernable as individual dots. Patterns of roads and city blocks appear.

Each city tells a different story. San Francisco appears to be fairly well-integrated, as well as racially diverse. Two or more colors often intermingle within close proximity, resulting in a heterogeneous patch. The outlined region in the figure below (Fig. 3.3) shows an area of the city where this happens to be the case. On the other hand, there are clear areas of racial distinction. For those who know the city well, certain neighborhoods and racial majorities stand out. The Richmond Districts, to the far left, are home to a fairly even mixture of Asian and white residents. The Mission District, shown at the center-right, is primarily Hispanic, but also contains other colors of dots. San Francisco’s Chinatown is recognizable as the dense patch of red in the upper right. Noticeably absent is a large number of inhabitants who identify as Black.

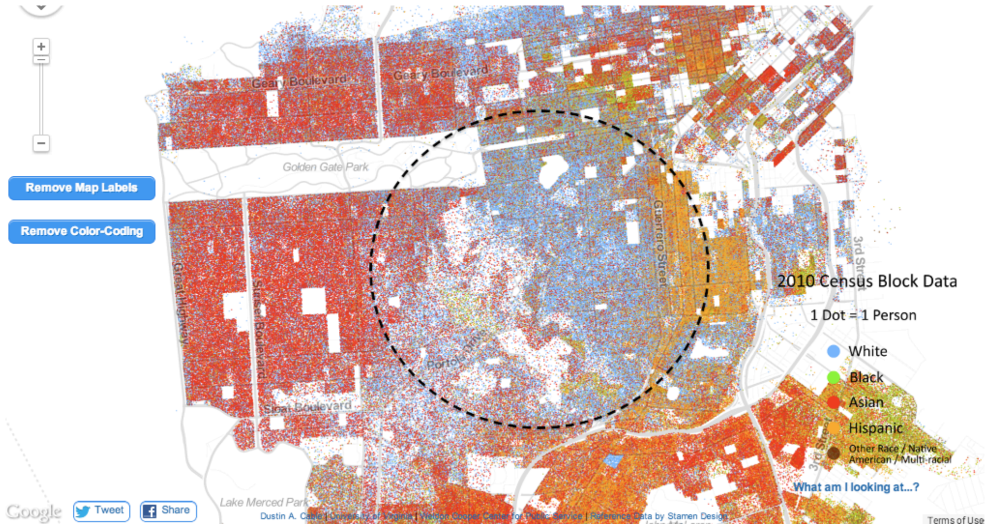

Fig. 3.3. Screenshot from the *Racial Dot Map*, San Francisco. The outlined region shows an area of the city that, at the zoom level, appears to be racially integrated, as well as diverse.

Other cities, like New York, are racially diverse, but not well-integrated. Fig. 3.4 shows a screenshot of the *Racial Dot Map* interface zoomed out to the regional level. The text identifying Harlem and the four visible Boroughs has been added. Unlike in San Francisco, vast patches of single colors dominate. While Manhattan appears to have inhabitants from all different racial identities, there are clear racial divides. Harlem and The Bronx are almost exclusively Black and Hispanic. Different communities within Brooklyn and Queens exhibit stark contrasts in color. Zooming into the closest level on lower Manhattan (Fig. 3.5) shows some integration. At the same time, there are neighborhoods and even city blocks that are largely characterized by a single race.

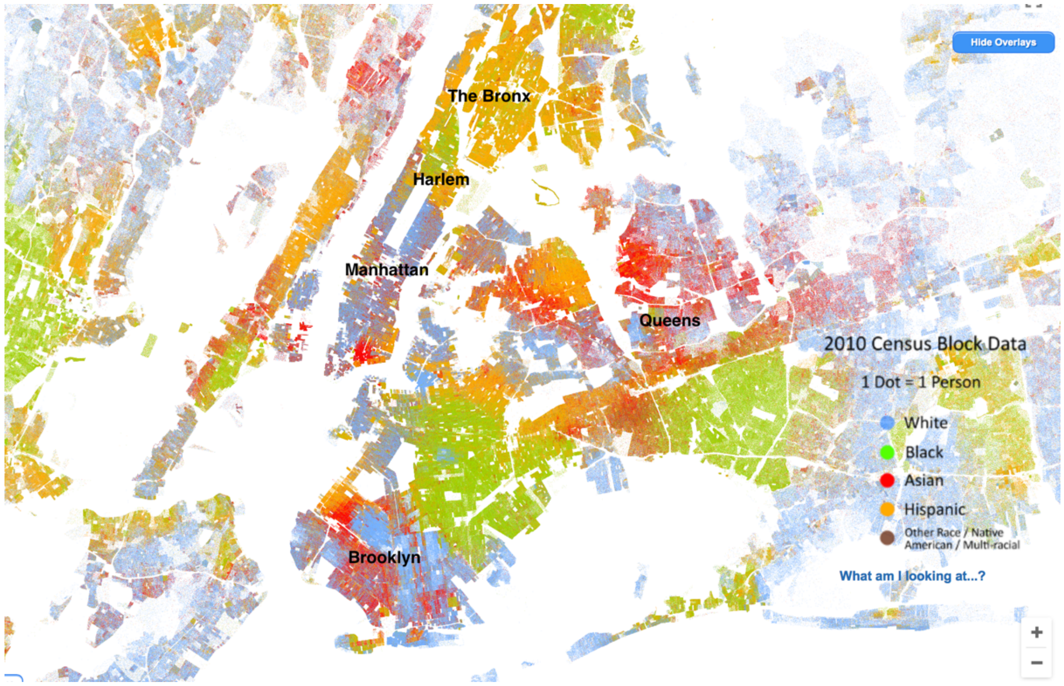

Fig. 3.4. Screenshot from the *Racial Dot Map* zoomed in to show the five boroughs of New York. Labels have been added to help explain the racial composition of neighborhoods in different areas.

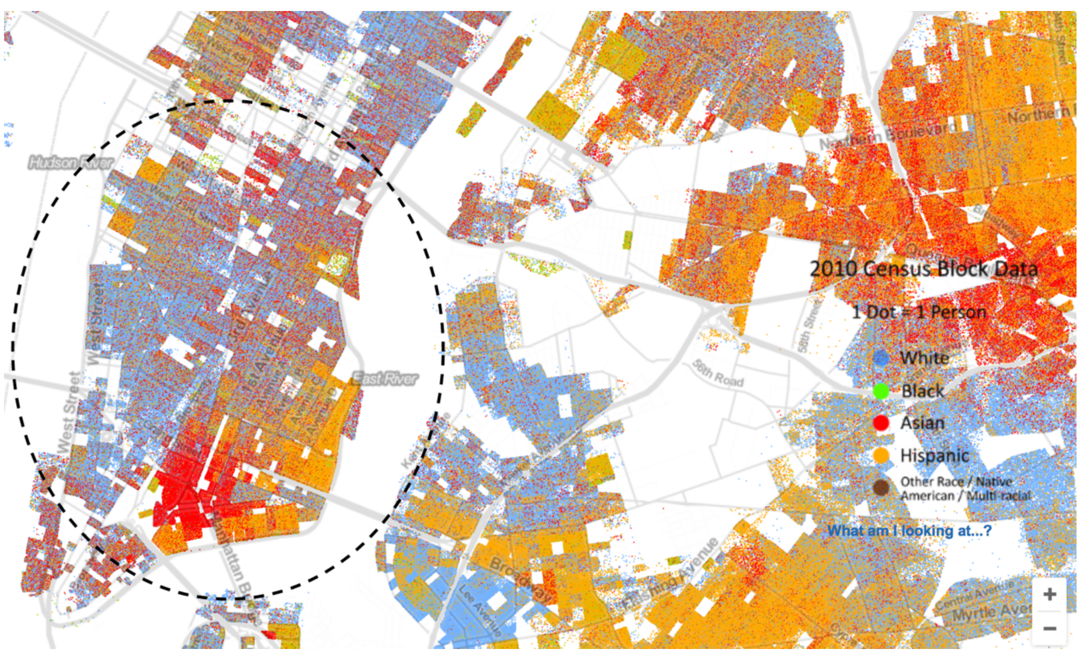

Fig. 3.5. Screenshot from the *Racial Dot Map* showing a close-up of lower Manhattan.

Detroit (Fig. 3.6) includes a particularly stark example where the two dominant races of White and Black appear to divide across a single street. For the most part, those inhabitants who identify as Black live to the south of 8 Mile Road, whereas those who identify as White live to the north. The exception is the small community of mixed Asian and White populations located in the lower-right quadrant of the map. This community, known as Hamtramck, has a history of being home to Polish immigrants and has recently seen an increased Asian population.[20]

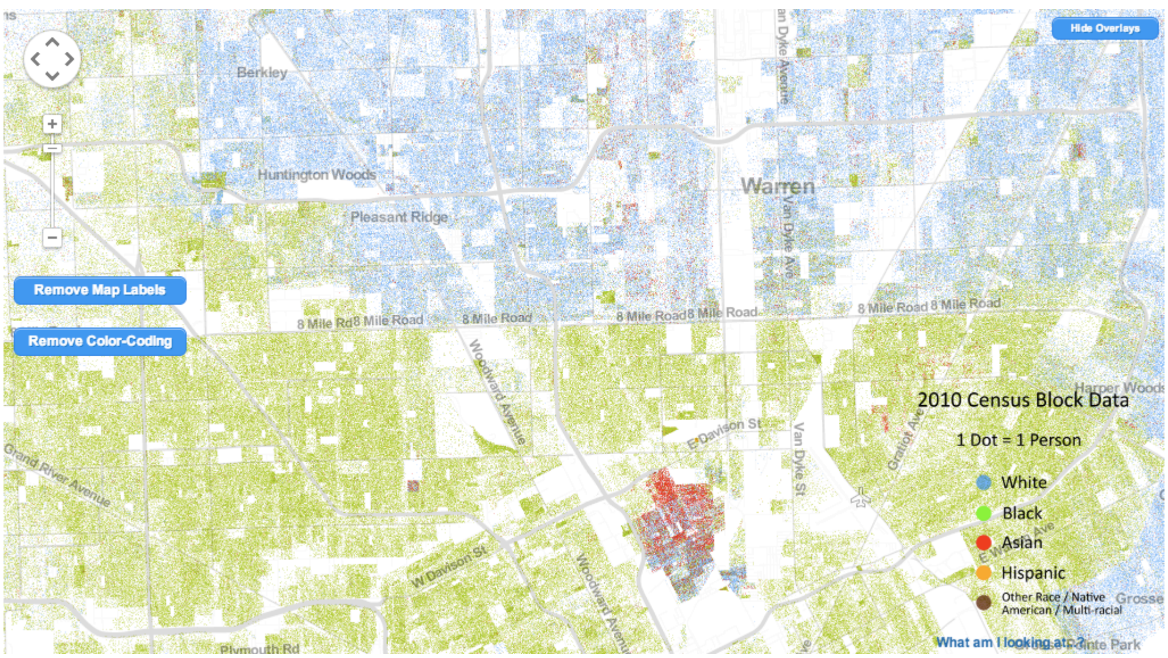

Fig. 3.6. Screenshot from the *Racial Dot Map*, Detroit.

[20] "Hamtramck, Michigan," *Wikipedia*, November 13, 2017, https://en.wikipedia.org/w/index.php?title=Hamtramck,_Michigan&oldid=810172144.

Personal Identification

But how does this build meaning? The insight gained by looking *the Racial Dot Map* is not necessarily new. Analysis of the 2010 US Census Data confirms that Detroit is one of the most segregated cities in the country and that, nationally, while residential segregation between different ethnicities in metropolitan areas is slowly declining, it still persists.[21] There is, however, something to be said about the value added to data by visualizing it at the level of the individual. In an interview with *National Geographic*, Cable is careful to recognize the higher level of engagement offered by interacting with the map, rather than scrolling through hundreds of census data tables. The map, he says, "puts complex data into context—you are a point on that map somewhere. You can look yourself up and look at yourself in the context of that neighborhood."[22] The importance of this personal identification with the data and representation cannot be overstated. Recent discussion and debate surrounding data visualization have pointed to the need to make data and how it is represented more human, more personal, and more identifiable.[23]

---

[21] John R. Logan, "The Persistence of Segregation in the 21st Century Metropolis," *City & Community* 12, no. 2 (2013): 160–168, https://doi.org/10.1111/cico.12021.

[22] Jaclyn Skurie, "Interactive Map Color-Codes Race of Every Single American," *National Geographic*, August 21, 2013, http://newswatch.nationalgeographic.com/2013/08/21/interactive-map-color-codes-race-of-every-single-american/.

[23] Lupi, "Data Humanism: The Revolutionary Future of Data Visualization"; Marian Dörk et. al., "Critical InfoVis: Exploring the Politics of Visualization," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (ACM, 2013), 2189–2198; Catherine D'Ignazio and Lauren F. Klein, "Feminist Data Visualization," in *Workshop on Visualization for the Digital Humanities (VIS4DH), Baltimore. IEEE*, 2016.

One way that the *Racial Dot Map* does this is through mapping techniques. The one-dot-per-person technique adds a sense of inclusion to the viewing experience. Viewers who participated in the 2010 U.S. Census are individually represented. They are not subsumed and thereby hidden by a larger trend or categorization. It is worth pointing out that the *Racial Dot Map* is not the first map to use this technique. The technique, known as "dot density mapping" or "dot distribution mapping," is the same technique used by Dr. John Snow in the famous street map illustrating the origin of the 1854 cholera outbreak in London. It even has its own Wikipedia page.[24] Bill Rankin's 2009/2010 dot maps of Chicago's racial and ethnic distribution, among many others, bear a remarkable similarity to Cable's representation.[25] In an explanatory article accompanying the visualization, Cable attributes his inspiration to Brandon Martin-Anderson's *Census Dotmap* (2011), which used the same one-dot-per person technique to map population density and distribution across the entirety of North America in a seamless zoomable interface.[26] Cable wondered what it would be like to add race to the representation and used Martin-Anderson's open-source code, made freely available on GitHub, to build the *Racial Dot Map*.

---

[24] "Dot Distribution Map," *Wikipedia*, November 11, 2017, https://en.wikipedia.org/w/index.php?title=Dot_distribution_map&oldid=809832601.
[25] Alberto Cairo, "Celebrating Pointillistic Cartography," *The Functional Art* (blog), September 4, 2013, http://www.thefunctionalart.com/2013/09/celebrating-pointillistic-cartography.html; Bill Rankin, "Chicago Boundaries," *Radical Cartography*, 2009, http://www.radicalcartography.net/index.html?chicagodots.
[26] Martin-Anderson's map draws data from the 2010 U.S., 2011 Canadian and 2010 Mexico census.

Another way that the *Racial Dot Map* evokes personal identification is through learned experience. For many viewers, the data and physical space represented on the screen overlap with past experience and knowledge related to, but not included in the representation. One of the first things I did when I came across the *Racial Dot Map* was zoom into the area where I grew up. I found the block where I lived, the block where I went to school, and the areas of town that I perceived as racially segregated. In some cases, the representation and my lived experience were the same. In others, they did not overlap. In either case, the effect was the same. Not only was I represented on the map, but I also knew the map. The result was a type of highly contextualized interactivity that added personal meaning to my relationship with the visualization.

Take, for example, the map of lower Manhattan zoomed into the city block level (Fig. 3.5). When I interact with the map, I know that the bright red patch on the lower east side is Chinatown. I know this not just because of the legend but also because of personal experience. I know that the Asian population in this area has grown dramatically over the last twenty years and that there used to be a deli on Broadway and Franklin St. where I would get a bagel almost every morning. It is as if I can add another layer, complete with street-view memories and contextualized information, to the representation. This ability of digital maps to facilitate highly contextualized interaction is an ability that has been explored by scholars in the field of geographic visualization.[27] It is particularly powerful in its ability to generate meaning associated with spatial and demographic data.

Zoom In / Zoom Out

Much of the power of the *Racial Dot Map* does not exist unless the viewer takes the initiative to zoom in and out. The map has a certain content, context and technique independent of a viewer's interactions. However, the ability to build personal context and meaning is largely dependent on the viewer's decision to engage in this simple act. What is gained, however, is not merely a different perspective on the data, but the ability to see into new worlds. In their short film *Powers of Ten* (1977), Charles and Ray Eames take the viewer on a journey of magnitudes. The film begins with a man and woman picnicking by the lakeside in Chicago. The screen then cuts to a view of one square meter above the ground, and the camera starts to gradually zoom out by powers of ten. We see the park, the entire city, the earth from space, and so forth until we arrive at a view of 1024 meters across—or just big enough to see the size of the observable universe. Then, the pattern reverses. The camera speeds back to the opening scenes. We arrive back at the picnic, then zoom to a close up of the man's hand. The camera continues through his skin and eventually stops at a view of

[27] For a summary article on different types of interactivity and exploration in geovisualization, see Jeremy W. Crampton, "Interactivity Types in Geographic Visualization," *Cartography and Geographic Information Science* 29, no. 2 (January 1, 2002): 85–98, https://doi.org/10.1559/152304002782053314.

10–16 meters across—small enough to see the quarks in a proton of a carbon atom. The journey ends.

The experience of viewing the *Racial Dot Map* is similar. What is visible at a city or state level is not always visible when zoomed out to account for the entire country. The example of zooming in on lower Manhattan, or any area to the individual dot level, for that matter, shows the deeper understanding gained by changing magnitudes. It is also a poignant reminder that the same representation of the same data can appear different to different people and under different circumstances. There is an interesting way in which the code of the *Racial Dot Map* contributes to this.

Because of the incredibly large number of dots, each is smaller than one pixel at maximum zoom levels. The way the code behind the *Racial Dot Map* resolves this is to display a color based on the ratio of different dots each pixel contains. For example, a pixel containing a relatively even mixture of red and blue dots appears purple. A pixel containing mostly blue and green dots appears teal. This color-by-combination technique is ultimately what allows the viewer to see dots at such a granular level when zoomed into the smallest scale.

Yet, it also means that some areas might appear to have a different racial composition than the data entails. In a short article detailing the map and his methodology, Cable offers the Minneapolis-St. Paul metro area as an example.[28]

[28] Dustin Cable, "The Racial Dot Map | Demographics," *Weldon Cooper Center for Public Service*, accessed December 9, 2017, http://demographics.coopercenter.org/racial-dot-map/.

When viewed from a state-wide zoom level, the area is purple, slowly fading to blue as the population moves further out from the city center. The immediate assumption is that Minneapolis-St. Paul area is racially integrated as well as diverse. However, zooming in tells a different story. Fig. 3.7 shows the Minneapolis-St. Paul area at progressively closer scales. The purple smudge that that seemed to characterize the city quickly gives way to neighborhoods and even city blocks that are dominated by a single race.

Fig. 3.5. Screenshots from the *Racial Dot Map* offering a progressively zoomed in view of the Minneapolis-St. Paul area.

The change in perception in this case directly corresponds to a change in cognition. By exploring and zooming in further, the viewer is forced to recognize existing perceptions and beliefs with new visualization representations. In many ways, this type of reconciliation is what is occurring at all levels of viewer interaction with the representation. This raises interesting questions about how much of a visualization's effectiveness is the responsibility of the designer versus the responsibility of the viewer. The designer can build the functionality into the design or code, but it is ultimately up to the viewer to add meaning through action.

So how does this all tie back to the timeliness and effectiveness of the *Racial Dot Map*? In a very unique way, the *Racial Dot Map* impels viewers to think and visually reconcile their existing beliefs with the visual representation of the data. That racial relations in the United States were part of the national consciousness at the time of the project's publication does not ensure that a viewer will look at the map and see racial tension. However, the experience of viewing the map incorporates several different opportunities for the viewer to interact with the data either through exploration, personal identification, or zooming that at least prompt the viewer to reflect and reconcile perceived assumptions about the racial composition of the United States with actual data, if he or she chooses.

CHAPTER 4

4. U.S. Gun Deaths: Emotion, Ethics and Visualization

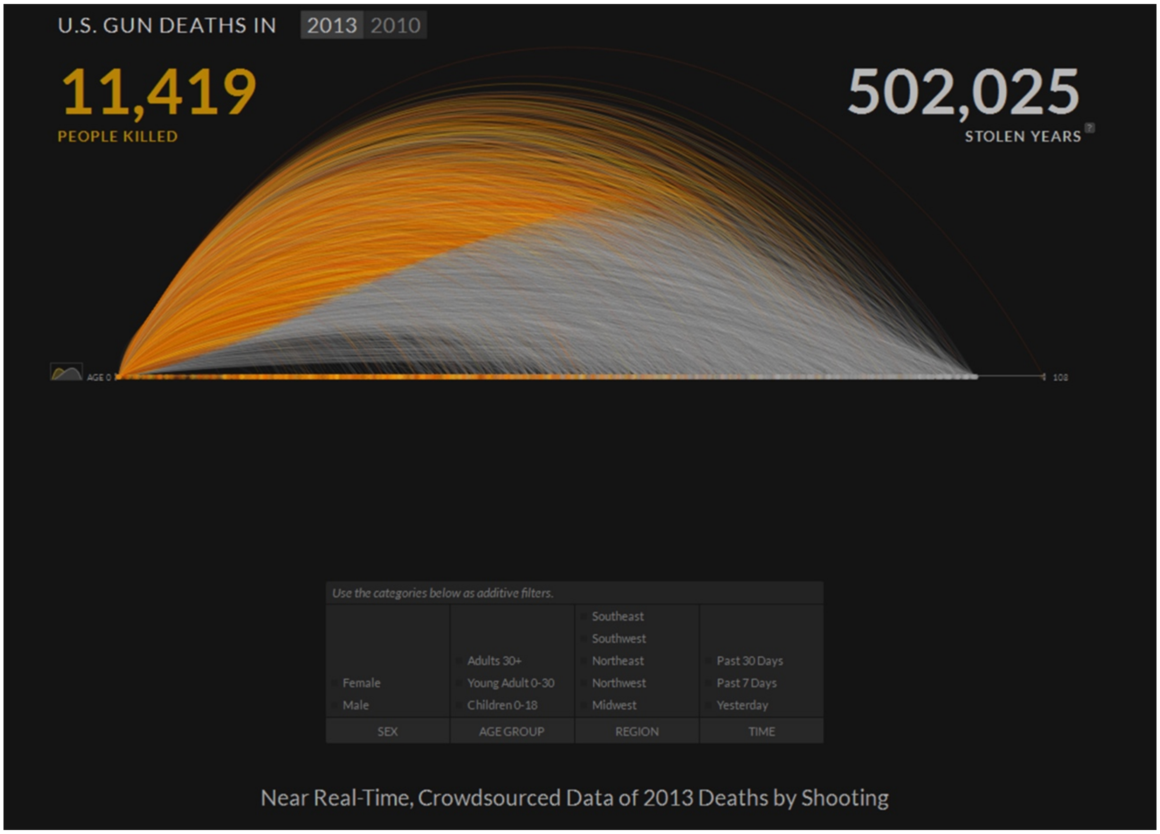

Figure 4.1. Screenshot of Periscopic's *U.S. Gun Deaths* (2013) once the animation is complete.

U.S. Gun Deaths is a data visualization project created by the design firm Periscopic shortly after the Sandy Hook Elementary shooting massacre in December 2012. Like many contemporary data visualization projects, it is located on the Web. When the page first loads, an empty graph with an x-axis labeled Age 0 to 108 appears. A thin orange line appears. Moving slowly, it begins to arc across an otherwise blank background. The background is contrastingly black. A label indicates that the line being drawn represents a person's life. His name is Alexander Lipkins. He was killed by gunfire at the age of 29. With this information, the line halts. Orange fades to gray and a small dot is left to fall back down to the baseline. The line, now gray, continues on. Again, a label appears explaining the second part of the arc: "could have lived to be 93." Five more lines are drawn in this same way (Figure 4.2a). Seconds later, the animation speeds up. Hundreds of lines begin to erupt from the x-axis (Figure 4.2b). Each follows a similar path—arcing, halting, and moving on. A grueling three minutes later, the animation is done. 11,419 lines have been drawn. Inevitably, all fade to gray.

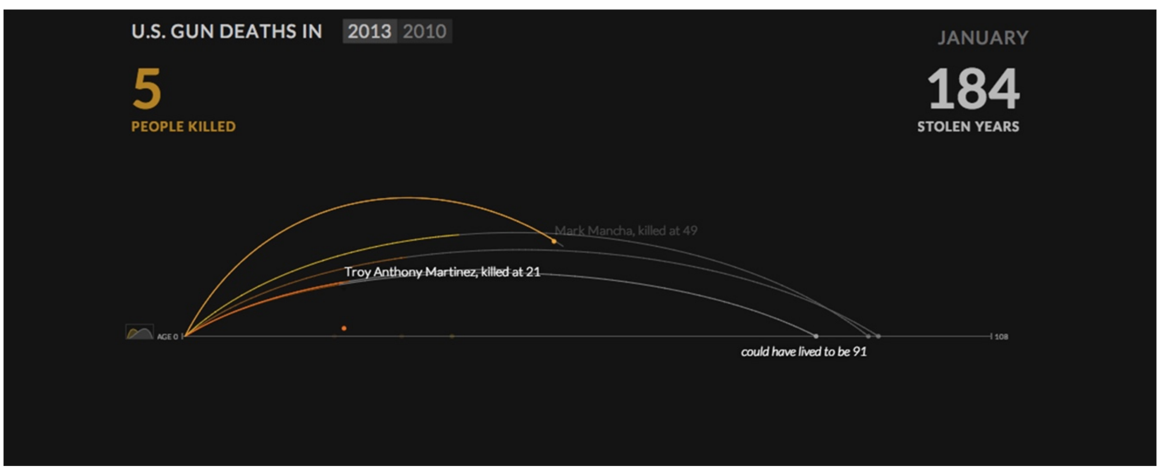

Figure 4.2a. Screenshot of *U.S. Gun Deaths* 2013 visualization at 32 seconds. Labels detailing the victim's name and age at the time of death and projected life expectancy are built into the animation for the first five lines. Labels for the following lines are only accessible by rolling over the line once the animation is complete.

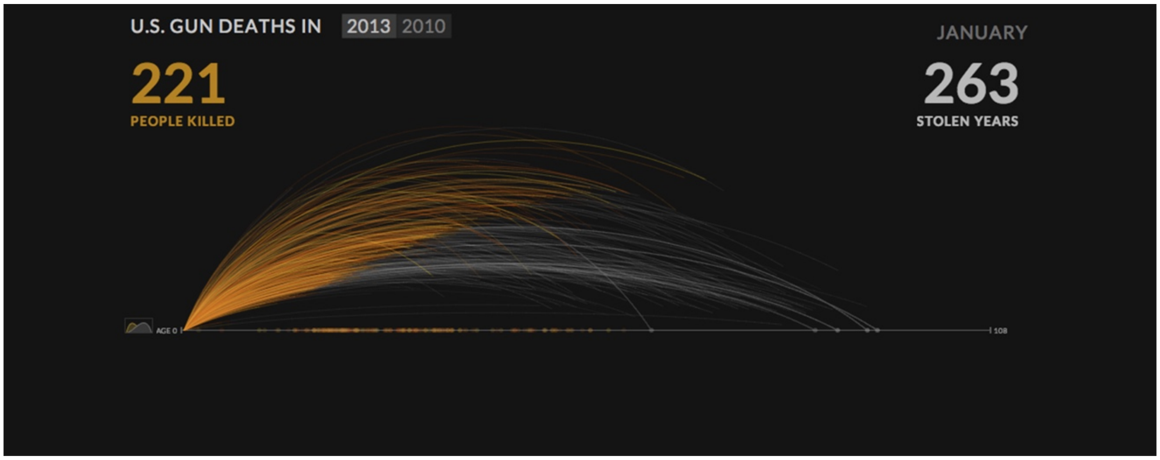

Figure 4.2b. Screenshot of U.S. Gun Deaths 2013 visualization at 38 seconds. Hundreds of lines in varying shades of orange and yellow begin to erupt for the x-axis. The entire animation takes approximately three minutes.

Emotion and data visualization. The two terms do not easily intersect. Until quite recently, there has been a perceived tension between the two. Data visualization, traditionally construed as the objective representation of qualitative information, is everything that emotion is not. Data visualization is rational, detached, and exact. Emotion is subjective, open to change, and felt. However, the growing role and popularity of data and data visualization have led to projects and research that challenge this assumption. Projects like Jonathan Harris and Sep Kamvar's *We Feel Fine* and Paul and Eve Ekman's *Atlas of Emotions* explicitly represent emotion-based data in visual form.[1] There has been a growing recognition that emotions and feelings play key roles in how we perceive and understand—even at the most rational level.[2] While other fields, such as user experience design and interaction design, have been

[1] http://www.wefeelfine.org/; http://atlasofemotions.org/
[2] Antonio R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain*, 1st ed. (Harper Perennial, 1995); Norman, *Emotional Design*.

quick to incorporate emotion into concepts of effectiveness, data visualization has not. This chapter asks why. It reads Periscopic's *U.S. Gun Deaths* (2013) as a way to examine the relationship between emotion and contemporary data visualization design. That the project is intended to evoke an emotional response on the part of the viewer is not hard to see. Emotion is incorporated at almost every level. It is present in the data. It is present in the design. And it is present in the publication context. The chapter begins by briefly introducing the recent perspective on emotion, cognition, and design. I then move on to examine the context and content behind the data visualization project, its specific design choices, and questions of ethics.

Emotion and Data Visualization

Understanding the relationship between emotion and data visualization begins with understanding recent strides in the relationship between design, cognition and emotion. In 2004, usability expert Donald Norman published an influential book titled *Emotional Design* in which he argues that "the emotional side of design may be more critical to a product's success than its practical elements."[3] Norman's background is in cognitive science with an emphasis on what is known as human-centered design. His inclusion of emotion marks a significant departure from previous work and traditions of how we learn and understand information. While artists and philosophers have long debated the relationship between emotion and cognition, modern approaches to

---

[3] Norman, *Emotional Design*, 5.

effective design have traditionally favored principles of functionality and practicality. If a product is not practical, then it cannot be used. The modernist design principle "form follows function" is both a manifestation and perpetrator of this mentality.[4]

Norman positions his claim in a growing cross-disciplinary research trend that sees technology as experience, of which emotion is a significant part.[5] This trend, which has origins in what is known as "the affective turn" in the humanities, and an increased interest in emotion in the sciences and social sciences, has been immensely influential in other fields, such as user experience and interaction design.[6] Both see emotion as central to how, why, and in what ways we interact with objects in the world around us.[7] Strangely, data visualization has been slow to do the same. The exception has been when visualizations are considered art. There are many possible reasons for this, including the relative infancy of data visualization when compared to product design as well as the proselytizing emphasis on rationality and objectivity espoused by visualization scholars such as Edward Tufte in the 1990s. Whatever the reason for past opposition, the relationship between data visualization and emotion is starting to

---

[4] The phrase "form follow function" was originally coined by architect Louis Sullivan in an article titled "The Tall Office Building Artistically Considered" in *Lippincott's Magazine* (March 1896): 403–409.

[5] John McCarthy and Peter Wright, *Technology as Experience* (Cambridge, MA: The MIT Press, 2004); Rosalind W. Picard, *Affective Computing* (Cambridge, MA: The MIT Press, 1997); Thomas Markussen and Jonas Fritsch, eds., "Exploring Affective Interations," *The Fibreculture Journal*, no. 21 (2012).

[6] The term applies specifically to the phenomenon of an increased cross-disciplinary interest in pre-cognitive bodies coming, which originated in the fields of cultural studies and critical theory in the 1990s. For more on this origin and theory, see *Affective Turn: Theorizing the Social* (Durham, NC: Duke University Press, 2007).

[7] Jenny Preece, Helen Sharp, and Yvonne Rogers, "5. Emotional Interaction," in *Interaction Design: Beyond Human-Computer Interaction*, 4th ed. (Wiley, 2015).

change. Designers and scholars advocating for a more human approach to visualization have cited the viewer's emotional response as a key factor for consideration.[8] Discussions of emotion and empathy have been particularly prominent in the fields of visual journalism and narrative visualization.[9] This chapter takes an initial step towards making this change by asking what it means to design for emotional content and emotional viewer response.

Content and Context

The first thing to point out about *U.S. Gun Deaths* is that it has specific rhetorical intent. The project was designed in January 2013 by Periscopic, a data visualization firm with a bold motto—"do good with data." When business partners Kim Rees and Dino Citraro decided to convert their interactive data firm into a visualization-based design firm in 2004, they did so with the explicit intent of using their skills as data analysts and designers to "convert raw data into visual and interactive experiences that allow people to empathize and understand."[10] At the time, this marked a unique move. While a select number of artists and designers had begun to

[8] Dörk et al., "Critical InfoVis"; D'Ignazio and Klein, "Feminist Data Visualization"; Lisa Charlotte Rost, "A Data Point Walks Into a Bar: Designing Data for Empathy" (Open Vis Conference, Boston, MA, April 24, 2017), https://openvisconf.com/#videos.
[9] For example, see Lisa Charlotte Rost's presentation "A Data Point Walks Into A Bar" at the 2017 OpenVis Conference in Boston, available at: https://openvisconf.com/#videos. There have also been a handful of workshops on emotion and visualization including Emovis 2016, which coincided with the ACM International Conference on Intelligent User Interfaces.
[10] "Periscopic: Do Good With Data," http://www.periscopic.com. For more information about the firm and its founding, see Kim Rees and Dino Citraro, "Creative Voices: Periscopic," December 2013, http://www.adobe.com/inspire/2013/12/interview-periscopic.html.

experiment with using visualization techniques in their own work, data visualization largely remained inside the lab and inside strict boundaries of expertise.[11] The idea that visualization could serve as a mass-communication medium and could be designed to evoke certain types of responses on the part of the viewer, such as empathy or understanding, was new.

U.S. Gun Deaths marks Periscopic's response to a very specific and emotionally laden event—the mass-shooting massacre at Sandy Hook Elementary. On December 14, 2012, a lone gunman carrying a semi-automatic assault rifle entered an elementary school in Newtown, Connecticut, killing six adults and twenty children. All children killed were between the ages of six and seven. The incident, which marked the deadliest mass shooting in a K-12 school in U.S. history and the second-deadliest shooting by a single person, was seen as particularly tragic.[12] Phrases like "no words" and "our hearts are broken" echoed in the country's media response.[13] Not surprisingly, so did discussions of statistics and data.[14] The events at Sandy Hook sparked a national debate. While nobody questioned the tragedy of the specific event,

---

[11] Michael Danziger, "Information Visualization for the People" (Masters Thesis, Massachusetts Institute of Technology, 2008).

[12] "Sandy Hook Elementary Shooting: What Happened?" *CNN*, accessed February 26, 2015, http://www.cnn.com/interactive/2012/12/us/sandy-hook-timeline/index.html.

[13] "'No Words' for Senseless Newtown Tragedy," The Huffington Post, accessed February 26, 2015, http://www.huffingtonpost.com/brian-normoyle/newtown-school-shooting_b_2303919.html; "Video: President Obama on School Massacre: 'Our Hearts Are Broken,'" *ABC News,* accessed February 26, 2015, http://abcnews.go.com/Nightline/video/president-obama-school-massacre-hearts-broken-17981785.

[14] For example, see Ezra Klein, "Twelve Facts about Guns and Mass Shootings in the United States," *Washington Post*, December 14, 2012.

making larger claims about gun violence, regulations, and rights required more than anecdotal support.

The visualization of this data was a logical next step. Seeing the need for a project that communicated and helped people to better understand the existing data on gun violence Periscopic decided to take this step.[15] Periscopic is a commercial visualization firm. They are usually hired by clients to visualize specific content, but *U.S. Gun Deaths* was designed as an internal project. A few weeks following the events at Sandy Hook, Periscopic posted a tweet asking that anybody who was interested in doing a visualization about gun violence contact the firm. Surprisingly, they got no response. This was despite their offer to work for free. Why this was the case is an interesting question. The second chapter of this dissertation talks about how visualizations are often in conversation with one another. Much like artworks, they often borrow, respond to, and build off of each other, both in terms of content and design. The eventual catalyst for Periscopic to pursue the project internally as well as many of the design decisions they chose to make can be traced back to one visualization in particular.[16]

---

[15] In the months following Sandy Hook, other news organizations also visualized data on gun violence and regulations. Two notable examples are "Mapping the Dead," *Huffington Post,* accessed February 26, 2015, http://data.huffingtonpost.com/2013/03/gun-deaths; and Karen Yourish et al., "State Gun Laws Enacted in the Year Since Newtown," accessed February 26, 2015, http://www.nytimes.com/interactive/2013/12/10/us/state-gun-laws-enacted-in-the-year-since-newtown.html.

[16] In a presentation at Stanford, Rees refers to William M. Briggs' visualization as the "straw that broke the camel's back." See Kim Rees, "Living, Breathing Data" (Seminar Presentation presented at the Stanford Human-Computer Interaction Seminar, Stanford University, February 15, 2013), https://www.youtube.com/watch?v=xTQdR3Fnrd0.

In early January 2013, self-proclaimed “statistician to the stars,” William M. Briggs (@mattstat) posted a link to his blogpost titled “Adjusted Mass Public Shootings per Decade Uncorrelated With Gun Ownership.” The post contained the following figure (Figure 4.3):

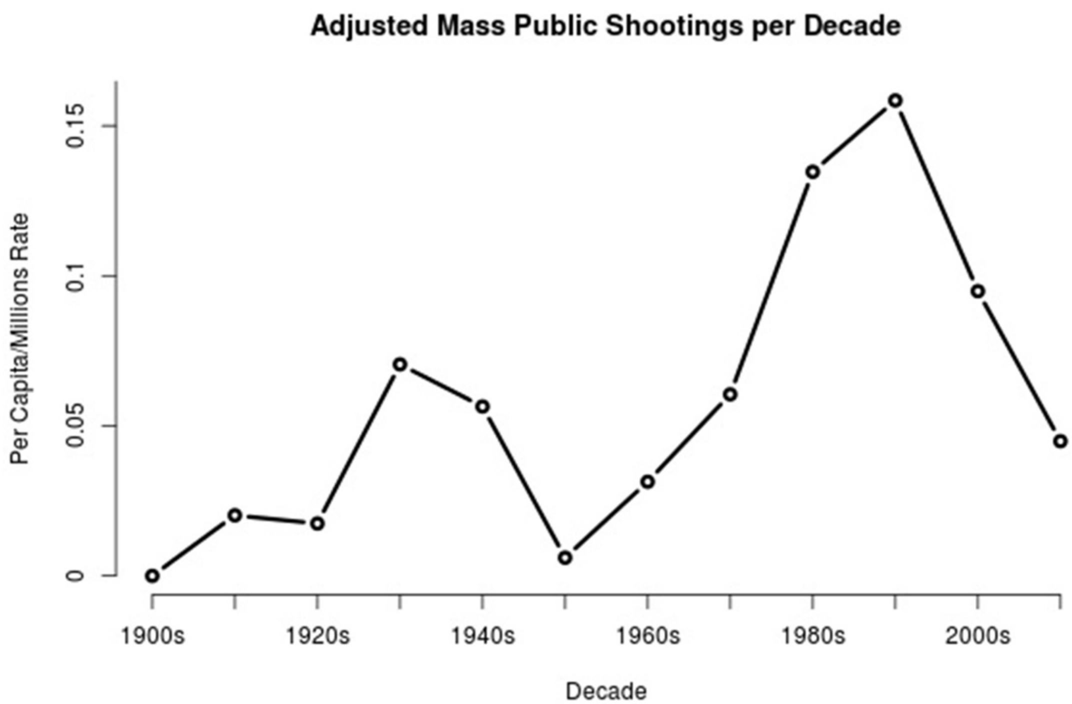

Figure 4.3. Screenshot of graph from William M. Briggs’ blogpost, January 13, 2013.

The graph shows the rate of mass-shootings per million per decade (1900 through 2012) as adjusted for the average population over each ten-year period. Its design is simple—a single black line formed by twelve data points. The data behind it is similarly sparse. Briggs draws data from a *Washington Post* article citing research conducted by Grant Duwe, then-director of evaluation and research at the Minnesota Department of Corrections.[17] Duwe collected data on mass murder in the United States going back 100 years for his 2007 book *Mass Murder in the United States: A History*. While the article presents the data in tabular form, Briggs presents it visually adjusted for population growth.[18] The mass shooting rate rose steadily from the 1950s through 1980s then fell drastically in the 1990s. If the number of mass shooting incidents was correlated with gun ownership, we would expect the rate to drop during the '70s, '80s, and '90s. Gun ownership fell steadily during these years.[19] Clearly, this is not the case.

The graph serves the purpose of creating rhetorical and visual impact. Briggs' post makes the point that mass shootings are a complex social phenomenon and that no one factor is to blame. While certainly true, nothing about this complex array of factors is evident from looking at the graph. The text offers little to no help in terms of further analysis or detail. We know twelve numbers and only what information we can glean based on our existing knowledge base and inclinations. Periscopic took issue with this—not necessarily Briggs' graph, in particular, but many of the data analysis and design principles bound into its representation.

---

[17] Glenn Kessler, "Bill Clinton's over-the-Top 'Fact' on Mass Shootings," *The Washington Post - Blogs*, January 11, 2013.
[18] Neither Brigg's blogpost nor the *Washington Post* article explain or look into the validity of Duwe's data.
[19] Sabrina Tavernise and Robert Gebeloff, "Rate of Gun Ownership Is Down, Survey Shows," *The New York Times*, March 9, 2013.

(Against) Aggregate Data

Aggregate data is data about collectivities. It is perhaps described best by what it is not—individual. The term, which is used in a variety of quantitatively based disciplines that deal with large amounts of data, refers to data that combines multiple variables (e.g., level of democracy or heart health score) or is summarized across individual entities (e.g., measurements from a single patient). Exactly what constitutes aggregate data depends on the specific context in which it is used. For example, aggregate data in medical studies often refers to data that is not limited to a single patient but tracked across time, organization, or some other variable. Aggregate data in political science often refers to measurements that represent bodies of interest as a whole (e.g., nations, regions, or levels of democracy) as opposed to their individual components (e.g., voters, workers, or inhabitants of a geographical area).[20] It is almost always secondary, meaning that these numbers are not collected by researchers directly, but rather are calculated based on recorded data using a variety of statistical averaging and distribution techniques. Let me be clear: there is nothing wrong with aggregate data. They give us a model and an overview; and when dealing with large and unwieldy data sets, aggregating data can help to examine the complex relationships between variables across measurements and across time. What Periscopic took issue with in Briggs' graph was that the data was presented in such an

---

[20] The term aggregate data is sometimes used to refer to data compiled from different sources. However, this is not the way it is used in this chapter.

aggregated, simplified form that none of the original complexity and detail informing the data was visible. Without this, the data becomes almost unrelatable. Being able to relate to data is precisely what drives people to action. Intellectually, the viewer might recognize a trend but none of the emotional aspects related to cognition are addressed. While Periscopic did not want to be a torch for any particular cause in creating the *U.S. Gun Deaths* visualization, they did want to present the data on an issue of national concern in a way that was more relatable and could potentially lead to informed action—whatever that action might be.[21]

Represented Data

To do this, they needed more comprehensive data. As is often the case with large-scale visualization projects, the entirety of data informing the project comes from multiple sources that were assembled and interrelated in a rather ad-hoc manner. The complete *U.S. Gun Deaths* project represents a combination of three different datasets: the Federal Bureau of Investigation's (FBI) Unified Crime Report, World Health Organization (WHO) demographic statistics, and *Slate*'s interactive, crowdsourced tally of gun deaths in the United States.

[21] It should be noted that many of the designers at Periscopic likely take an anti-gun or pro-gun regulation stance. The ethical question of if they allow this to influence their design decisions is addressed in the "Ethics" part of this chapter.

2010 Dataset

Every year, the FBI collects voluntarily reported data from over 18,000 law enforcement agencies across the country. Some of this data is summarized and published in their annual *U.S. Crime Report*, made available online.[22] The information included tends to take on an extremely aggregated form. Summary numbers for each state, region, crime type, or other broad categorical variable are included without any additional information on the data from which they are comprised. It is possible, however, to obtain more detailed data. Around the same time that Periscopic was first was thinking about doing a visualization related to gun violence, Rees received a data file from a friend. The email, plainly titled "all 2010 gun homicides," contained incident levels from the FBI Database including detailed information about each person killed, their age, their race, their gender, their relationship to the killer, and more. This is the dataset from which the 2010 visualization extends.

What the FBI dataset did not contain was any information on the potential lives these victims might have lived had their years not been cut short by gunfire. To calculate this, Periscopic turned to predictive analytics. Using WHO demographic statistics collected by the United Nations, they first performed an age prediction, weighted according to the age distribution of deaths in the United States. Then, using this age, they predicted a likely cause of death.

---

[22] http://www.fbi.gov/about-us/cjis/ucr/ucr-publications

2013 Dataset

The numbers behind the 2013 visualization come from a different source. At the time of the project's publication, the 2010 FBI data was the latest information released by the federal government. To get a more updated picture of gun deaths across the nation, Periscopic had to turn to a less "refined" source—*Slate* magazine's crowd-sourced tally of US gun deaths. Anonymous Twitter user @USGunDeaths began counting gun deaths via Twitter feed shortly following the movie theatre shooting in Aurora, Colorado. Anything verifiable by report was included. In January 2013, *Slate* partnered with @USGunDeaths to create an interactive, online project documenting how many people had been killed by gun since Sandy Hook. Readers were encouraged to contribute known incidents via email to help ensure the most accurate count. Reports were published daily, and the 2013 visualization was updated to correspond with the current count.[23] The use of WHO data to predict the future life expectancy and cause of death of 2013 victims functioned the same as with the 2010 dataset. It is important to note the limitations of this data. Crowd-sourced, real-time data is necessarily incomplete and may suffer from certain inaccuracies that raise legitimate cause for concern. These, as well as other concerns with the dataset, are addressed later in the "Ethics and *U.S. Gun Deaths*" section of this chapter.

---

[23] *Slate*'s tally project was retired on December 31, 2013, a year after its start. The count has been picked up by Michael Klein's Gun Violence Archive, which continues to collect information on every gun death across the nation. See http://www.gunviolencearchive.org/ for more information on the Klein's endeavor.

Design

Periscopic's goal in visualizing this data is, of course, to visualize more than just numbers, and they make very specific design choices that are meant to highlight the non-aggregate nature of the data as well as draw out an emotional response from the viewer. The following section looks at three aspects of *U.S. Gun Death*'s design (Graphical Features, Use of Language, and Interactivity), and analyzes them in relation to the visualization's larger rhetorical purpose.

Graphical Features

The first and perhaps most obvious of these is Periscopic's decision to show the one among the many. Each individual—each victim—gets his or her own line. Each line is accompanied by a label detailing the circumstances of that person's life and death. For the first few lines drawn, this information is given. Text labels accompanying each line appear as the animation moves forward. We know, for example, that the first line drawn belongs to Alexander Lipkins. We know that he was killed at age 29 but that he could have lived to be age 93. By giving this otherwise abstract line a named identity, the visualization's design encourages viewers to see each line not as a number or aggregated statistic, but as a once-living, individual person.

By the time the animation is finished, there are 11,419 lines—11,419 individuals. It takes a surprisingly long time for all of these to be drawn out. For a full three minutes, the viewer must watch and wait as new lines are continuously drawn. Meanwhile, the ticker showing the number of people killed in the top left corner and the ticker showing the number of years stolen in the upper right corner keeps going up. Powerfully, we see that not only have 11,419 been people been killed by gunfire in 2013, but 502,025 years have been stolen from them. These numbers are 9,595 and 414,046 respectively for the year 2010. The lines, en masse, have a sobering effect. The individual would be lost if it were not for the slight variations in color that make each line separately visible. As they continue to build up, they visually reiterate the sheer volume of the numbers.

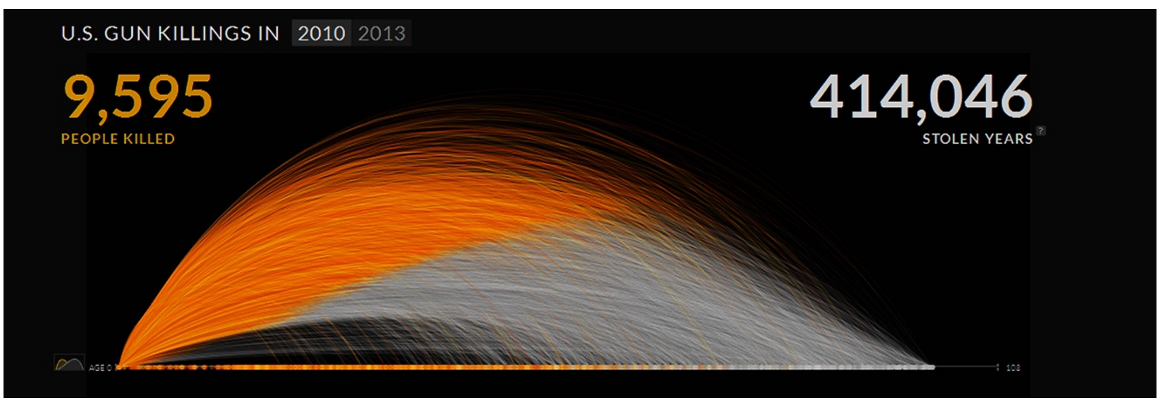

Figure 4.4. Screenshot of Periscopic's *U.S. Gun Deaths* 2010 visualization once the animation is complete. Note the tallies of people killed and total years lost in the upper left and right-hand corners. Bright orange and yellow are used to represent life. Dull shades of gray are used to represent life not lived.

This is made even more palpable by the visualization's strategic use of color and contrast. Figure 4.4 shows a screenshot of the completed 2010 visualization, in which the contrast is exaggerated. In the context of the data's subject matter, the black canvas sets an underlying ominous tone that contrasts sharply with varying vibrant shades of orange. It, however, contrasts less sharply with the single shade of gray used to represent the years lost. The association between color and life and absence of color and death is not hard to tease out. Once all the lines have been drawn, the divide between the zone of orange on the left and the zone of gray on the right is equally stark. It is almost as if a line of death by gunfire has been drawn. The orange and gray areas are approximately the same size, but this is not what stands out. What stands out is the way the gray almost seems to impinge upon the zone of orange, as opposed to the other way around.

Death creeps closer to birth in this visualization. Some victims live out the majority of their predicted lives before they are killed, but what you cannot help but notice is how many lines fade before they reach their halfway point. The visual understanding is that not only have as many years been lost as lived, but many victims were killed at a relatively young age. Whether this attentiveness to the relationship between gun deaths and youth was conscious on the part of Periscopic's design or simply true to the data, or both, the visualization's publication shortly following the Newtown incident places its reception in this context. There is something very powerful in this for the argument of situated visualizations that this dissertation makes. Not only are how and why data is visualized dependent upon the social, cultural and historical contexts in which a visualization is made. How and why we interpret the visualization is dependent upon the social, cultural and historical contexts in which it is seen.

Language and Text

Text is not something that we typically think about when we think about data visualization, but language, or at least textual explanation, often plays an important role in framing and contextualizing the data that is visualized. Usually, this takes the form of textual labels (e.g., chart labels, axis labels, series labels, key text, rollover labels) that offer some indication of how the data is represented or should be interpreted. Text is also often used to provide further insight or detail. Interactive graphics with rollover labels often use text in this manner. By rolling over certain data points or other graphical attributes, viewers can reveal additional information about the data represented.

The way text is primarily used in *U.S. Gun Deaths* functions in this informative way. Users can learn more about each victim by rolling over individual lifelines. This consists of two labels revealing information that is not apparent (even implicitly) by looking at the graphical elements alone. Exactly what information is contained in the labels varies between the datasets, but for both the 2010 and 2013 visualizations, one label reveals the victim’s identity and circumstances of death, the other details their potential future life. The text is short and the language direct. One of the few pairs of labels belonging to a female victim reads (Figure 4.5):

Tina M. Clark, a **48** year-old **woman**, was killed on **Thursday, January 3rd** in **Hayden, Idaho**.

*She might have lived to be **91** and died of **nervous system disease**.*"

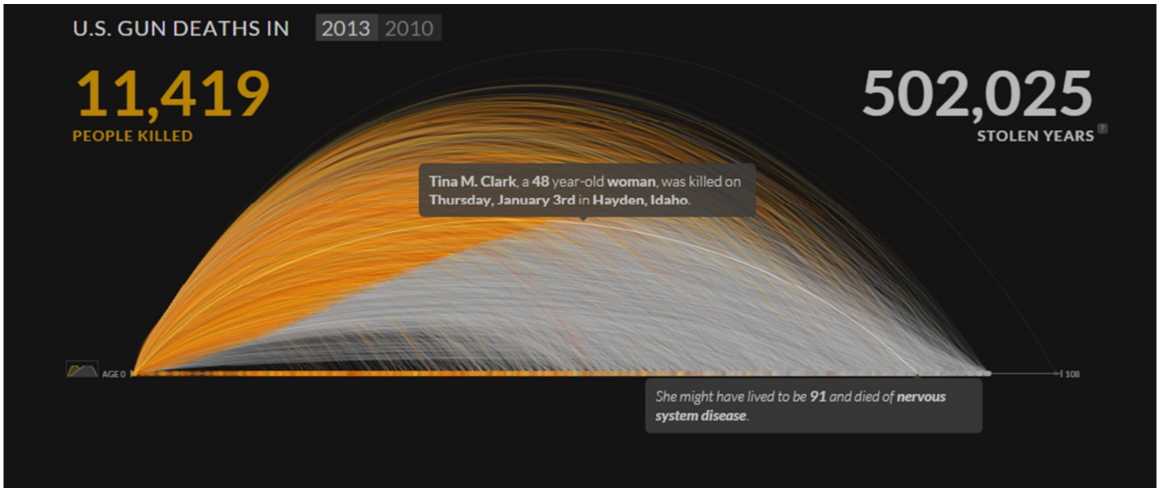

Figure 4.5. Screenshot of roll-over labels in *U.S. Gun Deaths* (2013). Labels can be uncovered by rolling over an individual line once the animation is complete. Direct language and font formatting are used to provide the viewer with information while highlighting certain aspects of the dataset.

Together, they read as short, pithy obituaries—one recorded and one only a possibility—intending to identify and honor each victim on a remarkably personal level. As viewers, not only do we now understand that each line is tied to an individual life, but we understand something about that person. It becomes more difficult to see them as a statistic.

In designing *U.S. Gun Deaths*, the designers and data scientists at Periscopic spent a great deal of time discussing what language they should be using.[24] They

[24] Rees, "Living, Breathing Data."

wanted to communicate in the most realistic and human way possible. This meant using the language people use every day. Their guide for what constituted this was *The Economist* magazine. Known by for its tight and engaging style of prose, *The Economist* puts out an online style guide that makes basic recommendations for how text should be written—namely that it should be clear, to the point, and the language of everyday speech.[25] For the most part, these are the stylistic guidelines Periscopic tried to follow.

Some of the language used in *U.S. Gun Deaths* has strong emotional connotations. Years are "stolen." People are "killed." "Stolen" as opposed to "lost" or "not lived," brings with it a stronger implication of injustice. "Killed" emphasizes the violent circumstances of their death. Given the proximity of the visualization to Sandy Hook, there is an impulse to see these words as biased and contradictory to the objective presentation of data. They certainly paint a negative reaction towards gun violence and the implications surrounding its regulation. Given the emotional potency of the language, it may be easy to conclude that Periscopic goes over the top in making an overtly political appeal. This assumption, however, requires caution.

The language Periscopic uses is direct, but it is also the same language people use to talk about gun-related deaths in everyday life. It is difficult to parse out what viewers' reactions may be, due to the timeliness of the project, the specific use of language, or viewers' preexisting views. There is no such thing as an entirely objective

[25] *The Economist* Style Guide is available at http://www.economist.com/styleguide.

visualization. The presentation, and even collection, of data is always already informed by the conditions surrounding viewers' expectations and designers' intent. To what extent certain design choices influence or manipulate viewers is dependent upon a lot more than the detached presentation of data. This is exactly Periscopic's point. The presentation of data, especially data that represents people, should not be detached because attachment is what effects change. Periscopic had a very specific goal in creating this project—to increase people's awareness of the existing data on gun violence while acknowledging the emotional and human aspects of these numbers—and the language used, regardless of an individual's beliefs on gun rights or regulation, is undeniably effective at doing this.

Interactivity

One other design component that works to de-aggregate the data and reinsert its emotional capacity is interactivity. Interactivity has become central to the way we think about and engage with data visualizations, especially on the web. As a hybrid online, animated-interactive project, *U.S. Gun Deaths* uses interactivity strategically to further its rhetorical goal. The rollover labels described in the section above serve as an interesting bridge between the discussion of interactivity and text. The first five labels to appear are part of the project's initial animation. When the website loads, the animation begins and continues until all lines for that year are drawn. Once the animation is complete, however, *U.S. Gun Deaths* becomes a very different visualization. The viewer becomes the user. As users, they must actively roll over an individual line to reveal the labels that otherwise remain hidden. The uncovering of a victim's identity, age, location and cause of death—all the details in the dataset that give information about a particular person—is thus tied to the viewer's actions. While rolling over a line to reveal more information arguably constitutes a very simple form of interactivity, there is rhetorical power in it. In a small way, the viewer-turned-user feels some sense of getting to know an individual person through his or her own action, which stands in marked contrast to passively observing an aggregated data point along a trend line. Arguably, this action is both physical and mental. Research in cognitive science has shown a strong connection between our physical movements and gestures and our cognitive processes.

Interactive search filters play a more nuanced role in getting the viewer to see and feel the data in a non-aggregated way. Once the animation has finished, users can interactively explore the dataset and see the results visualized. Categorical lists displayed below the x-axis serve as search filters (Figure 4.6). Each pertains to different data dimension (e.g., age group, race) and breaks down this dimension into smaller sub-categories. "Age Group," for example, is broken down into "Adults 30+," "Young Adults 0-30," and "Children 0-18." By selecting different combinations of sub-categories, users can set specific constraints on the data. The results of their search are then displayed in the altered visualization above.

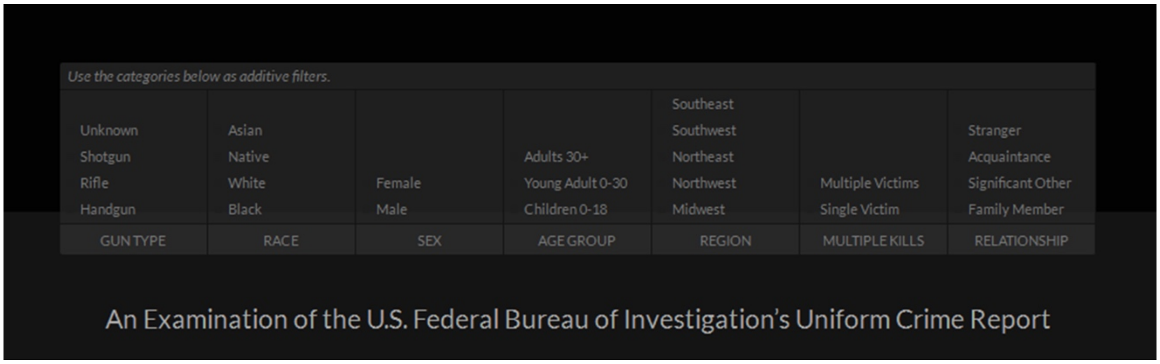

Figure 4.6. Screenshot of filters available for the *U.S Gun Deaths* 2010 visualization. Fewer filters exist for the 2013 visualization (see Figure 4.7). The discrepancy reflects different variables included in the different datasets.

This particular type of interaction, known as dynamic querying, is especially effective at communicating the complexity and nuance that characterize the dataset. Dynamic queries were first developed by Ben Schneiderman and Christopher Ahlberg in the early 1990s to give users a way to visually "fly through" databases.[26] Most database systems at the time required users to create complex queries in order to search the data, which presumed that they had some degree of familiarity with high-level query languages (such as SQL and QUEL) as well as the structure of the data.[27] In other words, they already knew the different components and complexity of the data. What dynamic queries allowed users to do was turn these language-based queries into graphical ones. By making simple adjustments to graphical widgets (such as sliders or buttons), users could search and see the database visually. The next

[26] Christopher Ahlberg and Ben Schneiderman, "Dynamic Query Filters with Starfield Displays," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '94 (New York, NY, USA: ACM, 1994), 313.
[27] James A. Larson, "A Visual Approach to Browsing in a Database Environment," *Computer* 19, no. 6 (June 1986): 62–71.

chapter looks at an expansion of this technique combining methods and modes of visualization from cultural analytics.

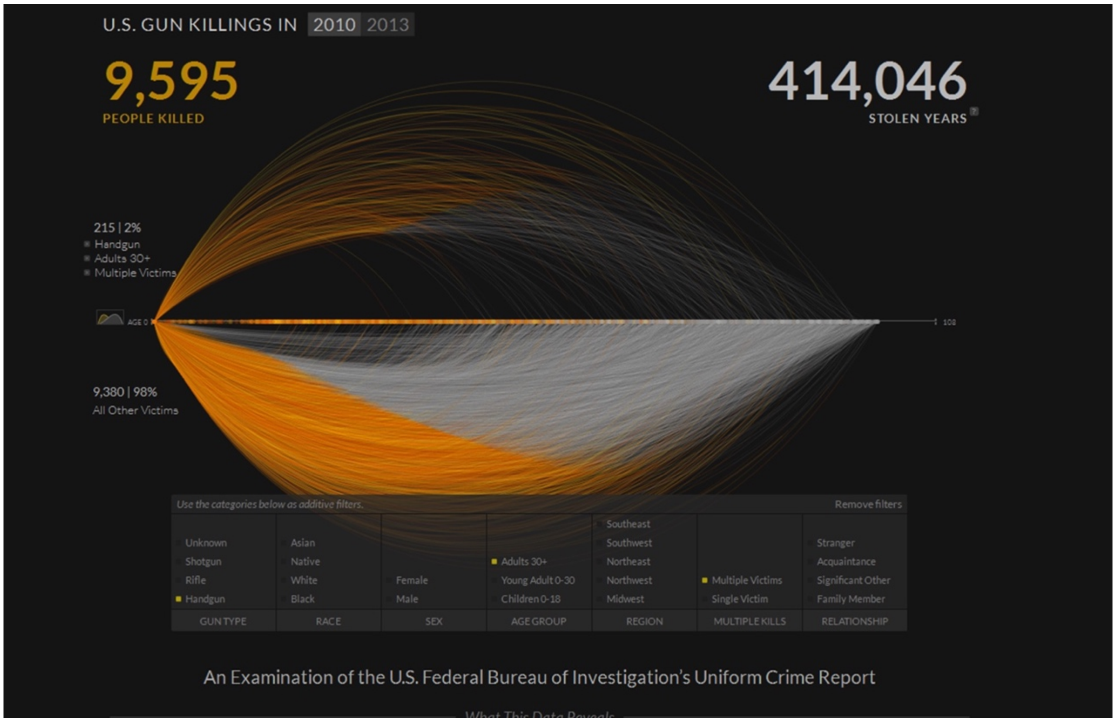

Figure 4.7. Screenshot of filters selected, *U.S. Gun Deaths* (2013). Only those lines that meet the selected criteria are shown above the x-axis. A reflection of all lines in that year's dataset are shown below the x-axis.

In *U.S. Gun Deaths*, the act of selecting filters dramatically transforms the visualization. What was a single graph with many lines becomes two. The graph depicting all lines drawn throughout a given year becomes reflected across the x-axis, and a like graph, showing only those lines that meet the requirements set by the filters takes its place (Figure 4.7). So, what you see are two graphs, one showing all victims, one showing only those victims that meet the selected criteria. Selected criteria can be combined, altered and reset at will by clicking on different sub-categories within the group or pressing the “Remove Filters” button at the top of the filter window.

What the labels and search filters add to *U.S. Gun Deaths* is more than providing the viewers with a way to functionally search the dataset. What the labels and search filters allow is a greater level of insight through interaction. The insight that evolves is particular to understanding the structure and the details of the dataset, which in this case facilitates seeing the data as a complex, interrelated network of individual measurements about individual shootings rather than numbers of some nondescript aggregate form. This, along with the content and other design choices that individualize the dataset, works to open the viewer up to an active emotional response to the project.

Ethics and Visualization

Existing Literature

There is, of course, an ethics to all of this. Ethics have become a recent topic of debate in the visualization field, particularly as they relate to representing and communicating information to a non-expert audience. We tend to believe that what we see equals the truth, and viewers can easily be persuaded to believe one way or another about the same data given different representations. The idea that there is a “right” and “wrong” way to represent data can be traced back to a general distrust of graphical representations of statistics during most of the nineteenth and twentieth centuries.[28] As Darrell Huff's popular 1954 book *How to Lie With Statistics* showed, statistical techniques can be used to obfuscate, sensationalize and confuse. Statistical graphs can even further distort our perception of importance and relationality in the data.[29]

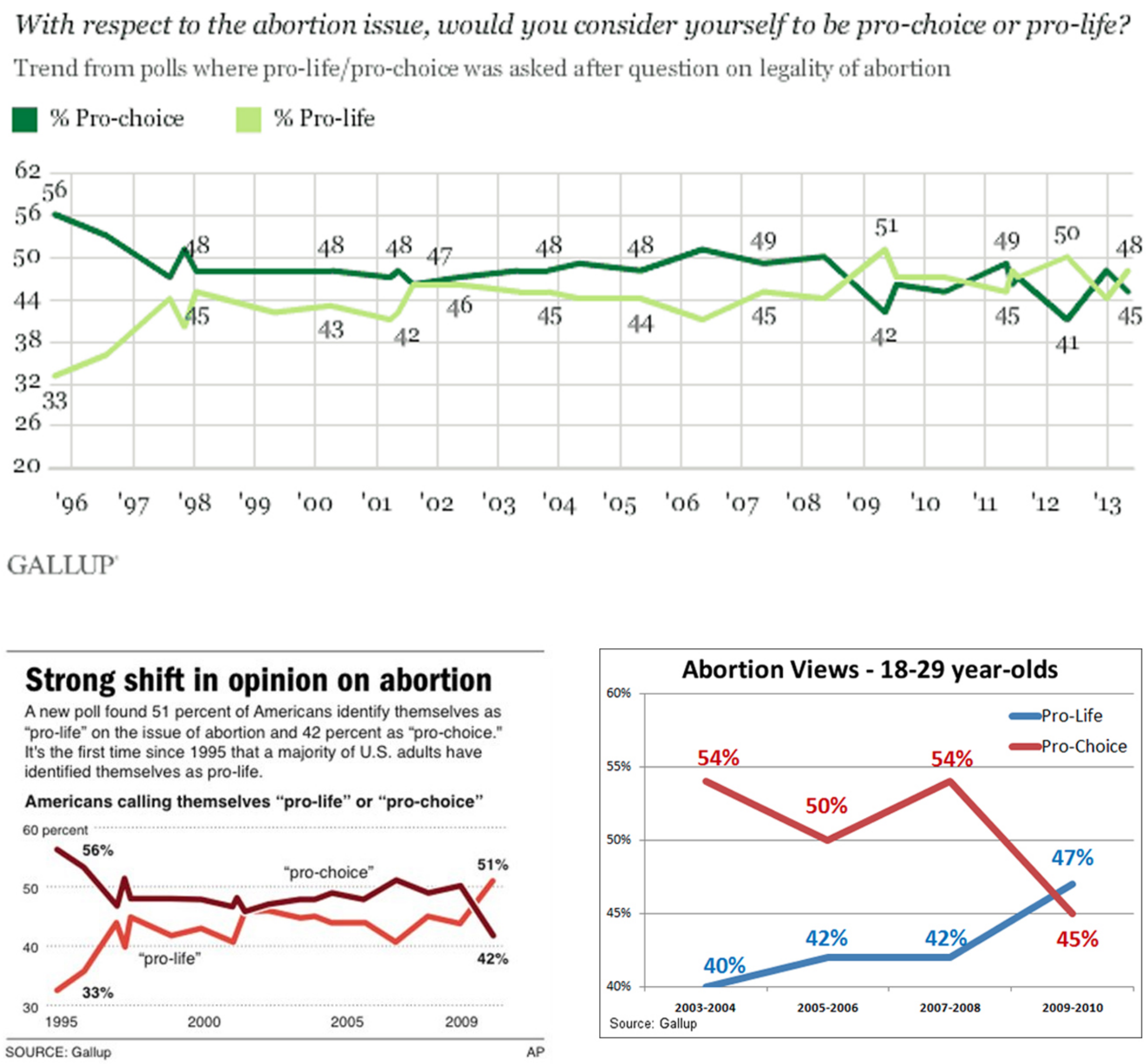

Figure 4.8a (top), 4.8b (bottom left), and 4.8c (bottom right). These charts, drawn from Mushon Zer-Aviv's article on "disinformation visualization," clearly show the power to influence viewers' beliefs through representation. Although all three charts represent the same Gallup Poll survey data, they encourage drastically different conclusions.

[28] Tufte, *The Visual Display of Quantitative Information*, 53.

[29] Darrell Huff, *How to Lie with Statistics*, (New York: W.W. Norton & Company, 1993).

Designer and media activist Mushon Zer-Aviv has an excellent article in which he talks about this. The article, cleverly titled "Disinformation Visualization," traces the representation of Gallup poll data surveying public opinion on abortion through three very different representations with very different intended interpretations.[30] The first (Figure 4.8a), created by Gallup to showcase the two opposing trends in their data, shows fluctuating lines with data points for all years measured. Numbers placed along the lines at evenly selected intervals put emphasis on the yearly totals and the data is represented at such a scale that it does not vary, at least not dramatically. The second (Figure 4.8b) narrows the time span, thickens the trend lines and introduces "pro-life" and "pro-choice" labels. Designed by the Associated Press, the graph presents a clear ideological shift, which makes for a much more interesting news story. The third graph (Figure 4.8c) shows an even smaller subset of the data. Only those responses given by 18-29 year-olds during the years 2003 through 2010 are shown. This representation was created by LiveAction.org, a website dedicated to building a pro-life culture. It gives the clear impression that pro-choice sentiment is on the decline. None of these charts necessarily lie about the data, but they do make certain design choices in line with their intended audience and presentation context.

The ethical concern in question here is not only that designers of visualizations might interfere with or influence viewers' perceptions. In his 1983 book *The Visual*

[30] Mushron Zer-Aviv, "Disinformation Visualization: How to Lie with Datavis," *Visualising Information for Advocacy* (blog), January 31, 2014, https://visualisingadvocacy.org/blog/disinformation-visualization-how-lie-datavis.

Display of Quantitative Information, Edward Tufte introduced the concept of graphical integrity. "Graphical excellence," he writes "begins with telling the truth about data."[31] Not telling the truth about data is not only morally wrong; it can lead to disastrous results. His famous example for illustrating this is the series of charts used by Morton Thiokol engineers to communicate launch conditions to NASA prior to the Challenger Space Shuttle launch in 1986. His claim is extreme. Because the engineers did not properly communicate their concerns, they are partially responsible for the astronauts' deaths. The style advocated by Tufte as a remedy for this is famously clean, simple and unadulterated. While Tufte's assessment of the Challenger charts has recently been challenged, many of his stylistic recommendations and the ideal of graphical integrity remain.[32] In fact, Tufte's recommendations and accompanying theory are cited as general rules to follow by many in the visualization community.

Much of the more recent discussion of the relationship between ethics and visualization has been taken up in the field of visual journalism. Visual journalism, which entails the practice of combining words and graphics to convey information to a largely non-expert public audience, is a particularly apt field for the discussion of ethics.[33] Journalism professor and information design Alberto Cairo has recently

[31] Tufte, 53. It should be noted that graphical integrity is only a component of graphical excellence in Tufte's work. Graphical excellence involves the larger claim that complex ideas in statistical graphics should be communicated with clarity, precision, and efficiency. Truth is thus tied to following these three principles.

[32] In their paper "Representation and misrepresentation: Tufte and the Morton Thiokol engineers on the Challenger," Robinson et al. make the claim that Tufte both misunderstood the argument and the true cause of the equipment failure. Moreover, he makes the case that it is the engineer's fault when readily available information shows what really happened.

[33] http://www.bbc.com/news/blogs-the-editors-22483705

become almost an advocate for the cause. For Cairo, visualization is an act of journalism (communicating information) plus an act of engineering (designing and structuring how information is assessed by the viewer). In recent articles and presentations, he has argued that the basic ethical value on which all of journalism is founded—presenting the public with information while minimizing the potential harm that this activity might cause—must also be extended to visualization. As Cairo puts it:

> Journalists who design visualizations need to address questions related to what they should display and why, but also pay increased attention to how they should display it. In other words, visualization designers must think about the structures, styles, and graphic forms that let audiences access information successfully in every situation.[34]

As people who decide and structure the communication of information, this is not only their moral responsibility. It is their civic duty.[35]

Cairo's argument to date has been more focused on advocating for a general awareness of the importance of ethics in visualization design than providing any overarching framework, but his utilitarian strategy of "do more good than harm" in representing information has translated to a few recommendations specific to individual projects. Among these are warning to base design decisions on more than

---

[34] Alberto Cairo, "Ethical Infographics: In Data Visualization, Journalism Meets Engineering," *The IRE Journal*, no. Spring (2014).
[35] Alberto Cairo, "The Journalist, the Artist and the Engineer: The Ethics of Data Visualization" (ONA14, Chicago, September 26, 2014).

visual appeal, be careful of confounding variables, and clarify exactly what data is represented.[36] Visual.ly, a platform for connecting designers, journalists and developers with clients to provide high-quality visual content at unprecedented scale and speed, has published a code of ethics that makes similar recommendations at a larger scale.[37] It is replicated in full in Figure 4.9. There has even been the suggestion of a Hippocratic oath for visualization designers.[38]

---

[36] Cairo also illustrates his concerns by providing an example of what he considers to be ethical slippages in visually representing data content. His example comes from visualizations in two blogposts about Nigerian Kidnappings on *FiveThirtyEight* blog that misrepresent data by applying it out of context. *FiveThirtyEight* has since updated the posts to correct the problems pointed out by Cairo's critique.

[37] Drew Skau, "A Code of Ethics for Data Visualization Professionals," *Visual.ly Blog*, 2012, http://blog.visual.ly/a-code-of-ethics-for-data-visualization-professionals/.

[38] Robert Kosara, "Visualization Is Growing Up," *Eagereyes: Visualization and Visual Communication* blog, November 6, 2011, https://eagereyes.org/blog/2011/visualization-is-growing-up.

1. Data collection
Data is pretty easy: data sources must be reliable and verifiable, attribution should be given whenever possible, dates should be included, etc. For more on finding reliable and verifiable sources, read our blog post on researching and sourcing infographics.

2. Data Analysis
This is where you find the "story" that goes into your visualization, and depending on what you are creating, the steps you take in your analysis can vary greatly. Sometimes, the data source is very simple and there isn't much analysis necessary. Other times, the data has multiple complex stories in it, and the analysis must be done carefully to only find truths.
It is important to leave out assumptions and only look at what the source data actually shows. If you have to make some basic assumptions, and if these assumptions aren't obviously visible in the finished product, you need to make them known with annotations.
Because the data analysis happens behind closed doors, so to speak — a viewer can't see what exactly it is that you did — this is the stage where the viewer needs to trust the presenter to have done their job well.

3. Design
The final stage is actually creating the visuals. Since the cognitive processes that make visualization work are still being researched, creating a comprehensive guide to ethical visualization is difficult. Still, we have plenty to work with to create a solid base of ethics requirements.
When designing, try to accurately portray the data and analysis, using the visuals you choose.
Be aware of things like the hierarchy of importance of visual properties and best labeling practices. Colors alone have a huge range of issues, from cultural meaning to isoluminance to colorblindness.
To really do visualization responsibly, immerse yourself in the world of visualization. Do lots of reading on the subject, examine any visualization you see with a critical eye, and be open to criticism yourself.

Figure 4.9. Visual.ly "Code of Ethics."

The Ethics of U.S. Gun Deaths

So, does *U.S. Gun Deaths* function ethically? The underlying point running throughout this entire discussion of an ethics of visualization is that visualizations are constructed. Visualization designers make choices in every step of the design process, from choosing which data to represent to how to best represent it. While there is no overarching framework for determining what constitutes an ethical visualization, there is a general consensus that ethical design should do its best to do what is commonly vocalized as “stay true to the data.” Staying true to the data, however, involves the acknowledgment that visualizations are designed for particular audiences and particular contexts. Working within these bounds, while maintaining a high degree of transparency and accuracy in representation is what is important. *U.S. Gun Deaths* was designed following the Sandy Hook Elementary School Shooting. It was meant to inform the public about the existing data on gun violence and it was meant to evoke an emotional response. The existing literature on ethics in visualization says nothing directly about designing for emotion as part of viewer response, but the implied concern is that doing so might somehow bias or influence the understanding of the data. While it is true that designers’ and viewers’ emotions towards a particular subject or design might influence their understanding of the data, this is not how emotion is used in *U.S. Gun Deaths*.

Emotion is used in a very specific way in *U.S. Gun Deaths*—and that is to de-aggregate the data and accentuate the individual human lives interwoven into its content. The design decisions made regarding graphical elements, text and language, and interaction, might influence an individual’s course of action after observing the visualization, but they in no way change the representation of the relationships and patterns in the data.

These two elements—what a viewer comprehends about the data and what a viewer decides to do or think upon seeing the data—are hard to tease out. What

Periscopic did do in designing *U.S. Gun Deaths* was incorporate some data that raised some ethical concern. There are two aspects of the visualization that people conversationally take issue with.[39] The first is its use of predictive data; the second is the addition of the crowd-sourced 2013 data. Both the 2010 and the 2013 visualizations rely on predictive data to represent the years victims might have lived had they not been killed by gunfire. None of the data represented by the gray portion of each lifeline was ever recorded. It was, instead, calculated based on existing numbers and trends.[40] To some extent, this may seem prescient, but it is not. What Periscopic is doing is using the same predictive analytic techniques used by Amazon.com, the Center for Disease Control (CDC), and any number of institutions or research groups that work with big data that is subject to variation over time. Based on these variables, this is what will probabilistically happen.

The second issue—the addition of the 2013 data—presents a less easily resolved concern. When the project was created, the 2010 data drawn from the FBI's database represented the latest nationwide data on gun deaths. Because it was tied to the federal government, it came with an assumed degree of certainty and validity. The release of such government data, however, has a long lag time. In deciding to introduce the crowdsoured 2013 data gathered by *Slate*, the designers at Periscopic made a decision to use contrasting datasets internal to the same project and

---

[39] I am basing this on reactions I have observed when discussing the project with others and the details put forth by Rees in her presentations discussing the project.

[40] The calculations performed to predict this data are fully explained in the "Methods and Sources" section of the project's webpage.

incorporate less reputable data. *Slate*'s dataset necessarily introduces new issues of ethical consideration. For one, it is incomplete. *Slate*'s project counted only the number of gun deaths reported in the media or otherwise verified, not the actual number of actual deaths. Suicides, which account for sixty percent of gun-related deaths in the United States, but rarely get reported, and similarly unreported incidents, were thus left out. As with any crowdsourced tally, there is also the possibility that some of the data reported was inaccurate.

Incorporating these two datasets were also design decisions. Much of the emotional power of the project is tied to their inclusion. The contrast in line color and textual labels juxtaposing how a victim died and what they might have died from would not be possible without the use of the predictive data. Likewise, tracking the number of gun fatalities that occurred in the year of the project's creation (2013) in near real-time and comparing them the numbers from a previous year (2010) would not be possible without the addition of the *Slate* crowd-sourced data. There is something to be said for the immediacy of near real-time data. The interactive filters for time in the 2013 visualization allowing viewers to explore which deaths had occurred in the past 30 days, the past 7 days, and the past day are proof of this. There is an affective difference between seeing data related to a gun death that occurred three years ago versus yesterday. Several years later, these filters no longer have direct effect. Instead, they serve as vestigial reminders of the power of near real-time data and the specific time and context in which the project was created.

Including these two datasets and their corresponding elements might be considered unethical, but this would only be true if they were represented deceptively. There is no such thing as a purely objective visualization, and Periscopic is remarkably open and transparent about their choice of data and what they have done to modify it. The "Sources and Methods" link at the bottom of the visualization's webpage opens a full explanation of what data is included and the potential pitfalls it may have. This is not unusual for large-scale visualization projects published online. What is unusual about *U.S. Gun Deaths* is that opening the description does not even lead the viewer away from the visualization. Clicking on the "Sources and Methods" link opens a pop-up window containing a textual explanation. The now dimmed visualization is still faintly visible in the background. This is a subtle move, but one that underscores Periscopic's intent of de-aggregating the data to evoke an emotional response on the part of the viewer. Part of their ethical responsibility in doing this is to provide the viewer with the information they need to be informed.

Despite what our assumptions about data visualizations might lead us to believe, emotion and visualization are not separate. They are not incompatible experiences and we should not strive to artificially keep them so. Looking at and exploring representations of data, especially representations of data that is tied to intimately human contexts and content, necessarily evokes some kind of emotional response on the part of the viewer. This is not to say that data must be about humans to evoke an emotional response. For a biologist, genetic data on fruit fly mutations might evoke the same intensity of emotional response, but this is largely dependent on the viewer's intimate understanding of the data and knowledge of its context. What is important is that this intimacy and data is built into the data's perception.

In the case of *U.S. Gun Deaths*, the data analysts and designers at Periscopic make conscious decisions about how to do this. They de-aggregate the data, tie it to the shared experience of human individuality, make choices about what data should be represented, and play with graphical and textual elements to create the type of experience they want. Importantly, they do this in a way that does not threaten the data or the project's integrity. This is not to say that there are not ethical precautions to be taken when designing for emotion in visualization. As the affective turn in design research shows, emotion significantly influences our cognition and our perception. Visualization can be and has been used to represent what we want to see in the data. But this does not mean that emotion should be excluded. It is an integral part of how we draw meaning from data and pretending otherwise is almost absurd. Much of our uneasiness with the pairing of emotion and visualization has to do with our expectations about data representation. We expect data representation to be absolute and true. When questions of uncertainty or emotion come into play, they threaten this.

CONCLUSION

This dissertation has largely been about situating contemporary data visualization in its surrounding social, cultural, and historical contexts. I have shown, through close reading and visual analysis, how data visualization, just like any other representational medium, is always conditioned by a particular set of circumstances, a particular set of goals, and particular ways of knowing and interpreting what we see. Part I charts the trajectory of contemporary data visualization in order to show how this is true at the level of practice. Data visualization is painted not so much as something that emerged out of scientific computing in the mid-1980s but as a visual form of representation that could not help but emerge given the confluence of social and technological pressures present at the time. It's goals, direction, and expansion to new audiences and genres in the late 1990s and early 2000s, are revealed to be more a product of this cultural history than anything inherent to the medium. Part II maps the idea that data visualizations are inseparable from the conditions and contexts in which they are created and seen to individual visualization projects.

This realization, I argue, has become increasingly important in recent years. As data visualization continues to play a role in how we understand and experience life, the assumption that data visualization is an objective, analytic tool for data's display becomes more difficult and dangerous to hold. In the introduction, I stated that the primary goal of this dissertation is to perform the background work necessary to open the door for more reflective and human-focused approaches to data visualization design and research. The last section of Chapter 2 outlines two existing approaches – critical visualization and feminist visualization. Both have ties to humanities concepts, and both challenge the assumption that data visualization serves as an objective, analytic tool for data's display. By uncovering what is at stake and what is implicit in choices related to data, representation, interaction, and reception, they reveal the hidden politics and expectations of data visualization. Their goal? To promote awareness, awareness about how we represent data, what we want representations to do, and what values are communicated in the process. Giorgia Lupi's call for a second wave of data visualization characterized by 'data humanism' similarly emphasizes these levels of awareness, advocating for more thoughtful and meaningful interactions.

When I first started working on this dissertation, I thought that I could build such an approach. Inspired by Dunne and Raby's critical design and academic background in the humanities, I thought it was important that we think about the role data visualization currently plays in everyday life and if we want it to change. I thought it was important that we think about questions of power, value, implications, and access in continuing to develop new techniques and interactions with representations of data. It is. I also thought that introducing a set of critical design principles would go a long way towards engendering this type of awareness. It would. However, what I found out in the process of researching and writing was that this is a nearly impossible task. It is nearly impossible, not because there is no merit or need, but because some of the basic foundations do not yet exist. These are summarized below in brief:

No common theory or vocabulary exists.

Data visualization is an inherently interdisciplinary field. This is both its greatest weakness and its greatest strength. Any attempt to build more reflection into the way we conceptualize and approach data visualization first needs a common base. Concepts and theories from a certain disciplinary perspective can serve as guides, but they must be relatable to all aspects of visualization reception and design.

No Recent History.

Much of the impetus for creating an update, more human-focused, approach to data visualization research and design stems from the path that contemporary data visualization's development has most recently taken. Yet, if there is no documentation or history of the recent past and present, understanding what changes need to be made is a stab in the dark.

No Close Reading.

Most peculiar is the fact that very few in-depth studies of individual visualization projects exist. Most reflection is done by designers on their own projects, with little external perspective or room for letting the projects speak for themselves.

I thus resolved to perform some of this background work. Having done so, there are a few themes or common ideas that have emerged. These serve as potential directions for future research.

Visualizations in Conversation

The first theme concerns the observation that data visualizations are in conversation. Beginning with Wilhelmson's storm visualizations in Chapter 1, it became quite apparent that each new representation was in conversation with previous work. Over the course of nearly twenty years, Wilhelmson and a team of researchers built off of their previous models and understanding in order to represent storm data in new ways. That is, decisions about data, representation, interaction, and reception environment were not new.

The *Racial Dot Map* (2015) quite literally illustrates this concept. Cable readily acknowledges that he borrowed from Brandon Martin-Anderson's *Census Dotmap* (2011) both in terms of inspiration and implementation. Martin-Anderson made his code available on GitHub for others to use. In this way, data visualizations can be seen to function like software, building off of component pieces created within a larger community. There is also the fact that the *Racial Dot Map*, and any visualization for that matter, visually references other data visualizations that have used the same technique or represented similar content. *The New York Times* also published a dot density map of racial segregation in July 2015.[1] Given the proximity in timing, one has to wonder if the two designers were in conversation.

[1] "Mapping Segregation," *New York Times*, interactive graphic, https://www.nytimes.com/interactive/2015/07/08/us/census-race-map.html

One possible direction for future research is to use Bhaktin's notion of the dialogic or similar theories of intertextuality to show how all visualizations carry on a dialogue with other works and designers. The theme of conversation could also be extended to show how all data visualizations are in conversation at multiple levels. They are in conversation with the data, conversation with viewers, conversation with certain design goals, and conversation with the time and contexts in which they are created and seen. Latour's Actor Network Theory might be useful in such a conceptualization.

Visualizations as Documentaries

Another idea is that we should see data visualizations more like documentaries. Chapter 3 briefly raises this possibility in the discussion of how the experience of viewing the *Racial Dot Map* necessarily relies on shared knowledge and human experience. There are several parallels to be drawn between the two forms of representation. Both are constructed but assumed to entail some degree of factual representation. That is, they are non-fiction. Both represent something that already exists in the world but for varying reasons remains unexplored or unseen. Both represent recorded data in a way that is intended to facilitate comprehension.

Many of the early experiments in photography and film were intended to document or better understand something. In 1872, the former governor of California Leland Stanford hired Edward Muybridge to help study a horse's gait. Stanford had reputedly taken a bet on whether all four of a horse's hooves left the ground simultaneously during a race. Muybridge set up twelve cameras side by side along a track. When the horse ran past, each camera would be tripped. The resulting image sequence definitively settled the debate. Data visualizations function similarly in that they facilitate comprehension through visual means. The difference is that the data represented is not always inherently visual.

There are differences. Like all analogies, the similarity between data visualization and documentary eventually breaks down. Documentaries are often more ideologically directed than data visualization. Visualizations are usually viewed individually rather than in a group. Rather than simply mapping documentary onto data visualization or vice versa, it might be more productive to investigate the similarities and differences between the two. Any such investigation would have to recognize visualizations as constructed. It would also open the door for a reception theory of data visualization, where viewers are needed to complete the process of communication and can have different interpretations of the same representations of data depending on their personal and past experiences.

Visualizations as Situated / Situated Visualization

The third theme and direction for future research concerns the idea that data visualizations are situated. I have hinted at this theme throughout the text using words like “embedded,” “context,” and “situated” to describe what occurs in the relationship between representation, viewer, and environment. At several points, it has been explicitly stated. The picture I have painted is one of data visualizations as cultural objects that are inseparable from the circumstances in which they are created and seen. This is more a media studies or humanities-based perspective. It would be interesting to explore the same idea from a cognitive science perspective, focusing on the specific mechanism or pathways that what is cognized and understood in designing or looking at a given visualization happen. This could take the form of both literacy and interaction studies, where users and designers are observed in situ, or constructing a distributed and embodied knowledge framework that spans multiple, minds, tools, and bodies.

SELETED BIBLIOGRPAHY

"About." *Weldon Cooper Center for Public Service*. Accessed November 29, 2017. https://coopercenter.org/about.

Affective Turn: Theorizing the Social. Durham, NC, USA: Duke University Press, 2007.

Ahlberg, Christopher, and Ben Shneiderman. "Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 313–317. CHI '94. New York, NY, USA: ACM, 1994. https://doi.org/10.1145/191666.191775.

———. "Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 313–317. CHI '94. New York, NY, USA: ACM, 1994. https://doi.org/10.1145/191666.191775.

———. "Visual Information Seeking Using the FilmFinder." In *Conference Companion on Human Factors in Computing Systems*, 433–434. CHI '94. New York, NY, USA: ACM, 1994. https://doi.org/10.1145/259963.260431.

Badger, Emily. "A Strangely Beautiful Map of Race in America." *The Atlantic Cities*, August 14, 2013. http://www.theatlanticcities.com/neighborhoods/2013/08/strangely-beautiful-map-race-america/6534/.

Baker, M. Pauline, and Colleen Bushell. "After the Storm: Considerations for Information Visualization." *IEEE Computer Graphics and Applications* 15, no. 3 (1995):12-15. https://doi.org/10.1109/38.376601.

Becker, Howard S. "Visual Sociology, Documentary Photography, and Photojournalism: It's (Almost) All a Matter of Context." *Visual Sociology* 10, no. 1–2 (January 1, 1995): 5–14. https://doi.org/10.1080/14725869508583745.

Becker, R. A., and J. M. Chambers. *S: An Interactive Environment for Data Analysis and Graphics*. Belmont, CA: Chapman and Hall/CRC, 1984.

Bertin, Jacques. *Semiology of Graphics: Diagrams, Networks, Maps*. ESRI Press, 2010.

Bertini, Enrico, and Moritz Stefaner. "Visual Storytelling w/ Alberto Cairo and Robert Kosara." *Data Stories*. April 16, 2014. Podcast accessed November 28, 2017. http://datastori.es/data-stories-35-visual-storytelling-w-alberto-cairo-and-robert-kosara/.

Brody, H, M R Rip, P Vinten-Johansen, N Paneth, and S Rachman. “Map-Making and Myth-Making in Broad Street: The London Cholera Epidemic, 1854.” *Lancet* 356, no. 9223 (July 1, 2000): 64–68. https://doi.org/10.1016/S0140-6736(00)02442-9.

Burdick, Anne, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jerry Schnapp. *Digital_humanities*. Cambridge, MA: MIT Press, 2012.

Cairo, Alberto. “Ethical Infographics: In Data Visualization, Journalism Meets Engineering.” *The IRE Journal*, Spring (2014).

———. *The Functional Art: An Introduction to Information Graphics and Visualization*. Berkeley, CA: New Riders, 2012.

———. “The Journalist, the Artist and the Engineer: The Ethics of Data Visualization.” Presentation at Online News Association '14, Chicago, IL, September 26, 2014.

Card, Stuart K. “Visualizing Retrieved Information: A Survey.” *IEEE Computer Graphics and Applications* 16, no. 2 (March 1996): 63–67. https://doi.org/10.1109/38.486683.

Card, Stuart K., Jock D. Mackinlay, and Ben Shneiderman. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, 1999.

Card, Stuart K., George G. Robertson, and Jock D. Mackinlay. “The Information Visualizer, an Information Workspace.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 181–186. CHI ’91. New York, NY, USA: 1991. https://doi.org/10.1145/108844.108874.

Carlson, Wayne. “A Critical History of Computer Graphics and Animation.” The Ohio State University. Accessed October 9, 2015. http://design.osu.edu/carlson/history/lessons.html.

Carstensen, Peter H., and Kjeld Schmidt. “Computer Supported Cooperative Work: New Challenges to Systems Design.” In *Handbook of Human Factors*, edited by K. Ito, 619–636. 1999.

Cawthon, N., and A. V. Moere. “The Effect of Aesthetic on the Usability of Data Visualization.” In *Information Visualization ’07*, 637–48. 11th International Conference 2007. https://doi.org/10.1109/IV.2007.147.

Chen, Chaomei. "Top 10 Unsolved Information Visualization Problems." *IEEE Computer Graphics and Applications* 25, no. 4, (2005): 12-16. https://doi.org/10.1109/MCG.2005.91.

Chen, Ing-Xiang, and Cheng-Zen Yang. "Visualization of Social Networks." In *Handbook of Social Network Technologies and Applications*, edited by Borko Furht, 585–610. USA: Springer, 2010. https://doi.org/10.1007/978-1-4419-7142-5_27.

Chi, E.H., and S.K. Card. "Sensemaking of Evolving Web Sites Using Visualization Spreadsheets." In *Proceedings of Information Visualization*, 18–25. IV '9. San Francisco, CA, USA: IEEE Computer Society, 1999. https://doi.org/10.1109/INFVIS.1999.801853.

Crampton, Jeremy W. "Interactivity Types in Geographic Visualization." *Cartography and Geographic Information Science* 29, no. 2 (January 1, 2002): 85–98. https://doi.org/10.1559/152304002782053314.

Curtis, James C. "Dorothea Lange, Migrant Mother, and the Culture of the Great Depression." *Winterthur Portfolio* 21, no. 1 (1986): 1–20.

Damasio, Antonio R. *Descartes' Error: Emotion, Reason, and the Human Brain*. Harper Perennial, 1995.

Danziger, Michael. "Information Visualization for the People." Masters thesis, Massachusetts Institute of Technology, 2008.

Defanti, Thomas A., Maxine D. Brown, and Bruce H. McCormick. "Visualization in Scientific Computing." New York, NY: ACM SIGGRAPH, 1987.

Defanti, Thomas A., Maxine D. Brown, and Marshall C. Yovits. *Visualization in Scientific Computing: Domain* and *Visualization in Scientific Computing: Systems*. Vol. 28/29. SIGGRAPH Video Review. Elsevier, 1987.

Dietrich, Frank. "Visual Intelligence: The First Decade of Computer Art (1965-1975)." *Leonardo* 19, no. 2 (1986): 159–69. https://doi.org/10.2307/1578284.

D'Ignazio, Catherine, and Lauren F. Klein. "Feminist Data Visualization." In *Workshop on Visualization for the Digital Humanities (VIS4DH),* Baltimore, MD: IEEE, 2016.

Dörk, Marian, Patrick Feng, Christopher Collins, and Sheelagh Carpendale. "Critical InfoVis: Exploring the Politics of Visualization." In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 2189–2198. ACM, 2013.

Eick, Stephen G., Joseph L. Steffen, and Eric E. Sumner Jr. “Seesoft - A Tool for Visualizing Line Oriented Software Statistics.” *IEEE Transactions on Software Engineering* 18, no. 11 (November 1992): 957–68.

Fetter, W. A. “A Progression of Human Figures Simulated by Computer Graphics.” *IEEE Computer Graphics and Applications* 2, no. 9 (November 1982): 9–13. https://doi.org/10.1109/MCG.1982.1674468.

Fisher, Danyel, and Paul Dourish. “Social and Temporal Structures in Everyday Collaboration.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 551–558. CHI ’04. New York, NY, USA: ACM, 2004. https://doi.org/10.1145/985692.985762.

Freeman, Linton C. “Visualizing Social Networks.” *JoSS: Journal of Social Structure* 1. Accessed August 21, 2016. http://www.cmu.edu/joss/content/articles/volume1/Freeman.html.

Fry, Benjamin Jotham. “Computational Information Design.” PhD dissertation, Massachusetts Institute of Technology, 2004.

Giroux, Henry A. “Lessons From Paulo Freire.” *Chronicle of Higher Education*, October 17, 2010.

Graetz, J. Martin. “The Origin of Spacewar.” *Creative Computing* 7, no. 8 (August 1981): 56–67.

Harley, John Brian. “Deconstructing the Map.” *Cartographica: The International Journal for Geographic Information and Geovisualization* 26, no. 2 (1989): 1–20.

Heer, J., and D. Boyd. “Vizster: Visualizing Online Social Networks.” In *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005*, 32–39, 2005. https://doi.org/10.1109/INFVIS.2005.1532126.

Huff, Darrell. *How to Lie with Statistics*. New York, NY: W. W. Norton & Company, 2010.

Hutchins, Edwin. *Cognition in the Wild*. Camridge, MA: MIT Press, 1996.

Jenkins, Henry. *Convergence Culture: Where Old and New Media Collide*. New York, NY: NYU Press, 2006.

Jordan, Patrick W. *Designing Pleasurable Products*. CRC Press, 2002.

Jung, Eui-Chul, and Keiichi Sato. “A Framework of Context-Sensitive Visualization for User-Centered Interactive Systems.” In *International Conference on User Modeling*, 423–427. Edinburgh, Scotland, UK: Springer, 2005.

Kaufmann, William J., and Larry L. Smarr. *Supercomputing and the Transformation of Science*. New York, NY: Scientific American Library, 1993.

Klein, Ezra. “Twelve Facts about Guns and Mass Shootings in the United States.” *The Washington Post*, December 14, 2012.

Klein, Lauren F. “Feminist Data Visualization; Or, the Shape of History.” Presentation at The Annual Convention of the Modern Language Association, Philadelphia, PA, January 5-8, 2017.

Kosara, R. “Visualization Criticism - The Missing Link Between Information Visualization and Art.” In *Information Visualization, 2007. IV ’07. 11th International Conference*, 631–36. IEEE, 2007. https://doi.org/10.1109/IV.2007.130.

Laposky, Ben F. “Oscillons: Electronic Abstractions.” *Leonardo* 2, no. 4 (October 1, 1969): 345–54. https://doi.org/10.2307/1572117.

Larson, James A. “A Visual Approach to Browsing in a Database Environment.” *Computer* 19, no. 6 (June 1986): 62–71. https://doi.org/10.1109/MC.1986.1663255.

Lau, A., and A. Vande Moere. “Towards a Model of Information Aesthetics in Information Visualization.” In *Information Visualization, 2007. IV ’07. 11th International Conference*, 87–92, 2007. https://doi.org/10.1109/IV.2007.114.

Lavington, S. H. “The Manchester Mark I and Atlas: A Historical Perspective.” *Communications of the. ACM* 21, no. 1 (January 1978): 4–12. https://doi.org/10.1145/359327.359331.

Logan, John R. “The Persistence of Segregation in the 21st Century Metropolis.” *City & Community* 12, no. 2 (2013): 160–168. https://doi.org/10.1111/cico.12021.

Lupi, Giorgia. “Data Humanism – The Revolution Will Be Visualized.” Capstone Address at the IEEE Visualization Symposium, Pheonix, AZ, October 6, 2017.

———. “Data Humanism: The Revolutionary Future of Data Visualization.” *Print Magazine*, January 30, 2017. http://www.printmag.com/information-design/data-humanism-future-of-data-visualization/.

Lupi, Giorgia, and Stefanie Prosavec. *Dear Data*. Princeton Architectural Press, 2016.

Mackinlay, J. "Opportunities for Information Visualization." *IEEE Computer Graphics and Applications* 20, no. 1 (2000): 22–23. https://doi.org/10.1109/38.814540.

Manovich, Lev. "Data Visualization as New Abstraction and Anti-Sublime," 2002. http://manovich.net/content/04-projects/040-data-visualisation-as-new-abstraction-and-anti-sublime/37_article_2002.pdf.

Markussen, Thomas, and Jonas Fritsch, eds. "Exploring Affective Interactions." *The Fibreculture Journal*, no. 21 (2012).

McCarthy, John, and Peter Wright. *Technology as Experience*. Cambridge, MA: The MIT Press, 2004.

McCullough, Malcolm. *Digital Ground: Architecture, Pervasive Computing, and Environmental Knowing*. Cambridge, MA: The MIT Press, 2005.

Metropolis, Nicholas, ed. *History of Computing in the Twentieth Century*. New York: Academic Press, 1980.

Moreno, Jacob L. *Who Shall Survive?: Foundations of Sociometry, Group Psychotherapy, and Sociodrama*. Beacon House, 1977.

Nardi, Bonnie A., Steve Whittaker, Ellen Isaacs, Mike Creech, Jeff Johnson, and John Hainsworth. "Integrating Communication and Information Through ContactMap." *Communications of the ACM* 45, no. 4 (April 2002): 89–95. https://doi.org/10.1145/505248.505251.

Neurath, Marie. "ISOTYPE." *Instructional Science*, no. 3 (1974): 127–50.

Neurath, Otto. *International Picture Language*. Reading: University of Reading, UK: Reading & Language Informat, 1981.

Newman, M. E. J. "Coauthorship Networks and Patterns of Scientific Collaboration." *Proceedings of the National Academy of Sciences* 101, no. 1 (April 6, 2004): 5200–5205. https://doi.org/10.1073/pnas.0307545100.

Noll, A. M. "The Digital Computer as a Creative Medium." *IEEE Spectrum* 4, no. 10 (October 1967): 89–95. https://doi.org/10.1109/MSPEC.1967.5217127.

Norman, Donald. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books, 2004.

Paul, Christiane. *Digital Art*. Thames & Hudson, 2008.

Phillips, R. L. “Distributed Visualization at Los Alamos National Laboratory.” *Computer* 22, no. 8 (August 1989): 70–77. https://doi.org/10.1109/2.35203.

Picard, Rosalind W. *Affective Computing*. Cambridge, MA: MIT Press, 1997.

Playfair, William, Howard Wainer, and Ian Spence. *Playfair’s Commercial and Political Atlas and Statistical Breviary*. Cambridge University Press, 2005.

Pousman, Zachary, John Stasko, and Michael Mateas. “Casual Information Visualization: Depictions of Data in Everyday Life.” *IEEE Transactions on Visualization and Computer Graphics* 13, no. 6 (2007).

Preece, Jenny, Helen Sharp, and Yvonne Rogers. *Interaction Design: Beyond Human-Computer Interaction*, 4th ed. Wiley, 2015.

Presner, Todd, David Shepard, and Yoh Kawano. *HyperCities: Thick Mapping in the Digital Humanities*. Cambridge, MA: Harvard University Press, 2014.

Ray, P.S, B.C. Johnson, K.W. Johnson, J.S. Bradberry, J.J. Stephens, K.K. Wagner, R.B. Wilhelmson, and J.B. Klemp. “The Morphology of Several Tornadic Storms on 20 May 1977.” *Journal of the Atmospheric Sciences* 38, no. 8 (August 1, 1981): 1643–63. https://doi.org/10.1175/1520-0469(1981)038<1643:TMOSTS>2.0.CO;2.

Rhyne, Theresa-Maries, ed. “Does the Difference between Information and Scientific Visualization Really Matter?” *IEEE Computer Graphics and Applications* 23, no. 3 (2003): 6–8. https://doi.org/10.1109/MCG.2003.1198256.

Rhyne, Theresa-Marie (Chair), Tory, M., Munzner, T., Ward, M., Johnson, C., Laidlaw, D.H. (Panelists), "Panel: Information and Scientific Visualization: Separate but Equal or Happy Together at Last", IEEE Visualization, 611-614. Seattle, Washington, 2003.

Robertson, G., S. K. Card, and J. D. Mackinlay. “The Cognitive Coprocessor Architecture for Interactive User Interfaces.” In *Proceedings of the 2nd Annual ACM SIGGRAPH Symposium on User Interface Software and Technology*, 10–18. UIST ’89. New York, NY, USA: ACM, 1989. https://doi.org/10.1145/73660.73662.

Robinson, Arthur H. “The Thematic Maps of Charles Joseph Minard.” *Imago Mundi* 21 (January 1, 1967): 95–108.

Rost, Lisa Charlotte. “A Data Point Walks Into a Bar: Designing Data for Empathy.” Conference Presentation presented at the Open Vis Conference, Boston, MA, April 24, 2017. https://openvisconf.com/#videos.

Russell, Daniel M., Mark J. Stefik, Peter Pirolli, and Stuart K. Card. "The Cost Structure of Sensemaking." In *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems*, 269–276. ACM, 1993. http://dl.acm.org/citation.cfm?id=169209.

Shepard, Mark, ed. *Sentient City: Ubiquitous Computing, Architecture, and the Future of Urban Space*. Cambridge, MA: The MIT Press, 2011.

Shneiderman, B. "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations." In *Proceedings of the IEEE Symposium On Visual Languages*, 336–43. Boulder, CO: 1996. https://doi.org/10.1109/VL.1996.545307.

Sutherland, Ivan E. "Sketch Pad a Man-Machine Graphical Communication System." In *Proceedings of the SHARE Design Automation Workshop*, 6.329–6.346. DAC '64. New York, NY, USA: ACM, 1964. https://doi.org/10.1145/800265.810742.

Tesler, Michael, and David O. Sears. *Obama's Race: The 2008 Election and the Dream of a Post-Racial America*. Chicago Studies in American Politics edition. Chicago, IL: University of Chicago Press, 2010.

Tractinsky, Noam. "Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 115–122. CHI '97. New York, NY, USA: ACM, 1997. https://doi.org/10.1145/258549.258626.

Tribe, Mark, and Reena Jana. *New Media Art*. Taschen, 2009.

Tufte, Edward R. *Beautiful Evidence*. Graphics Press, 2006.

———. *Envisioning Information*. Graphics Press, 1990.

———. *The Visual Display of Quantitative Information*. 2nd ed. Graphics Press, 1983.

———. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press, 1997.

Tukey, John W. *Exploratory Data Analysis*. Pearson, 1977.

Viegas, F. B., M. Wattenberg, and J. Feinberg. "Participatory Visualization with Wordle." *IEEE Transactions on Visualization and Computer Graphics* 15, no. 6 (November 2009): 1137–44. https://doi.org/10.1109/TVCG.2009.171.

Viégas, F., and Martin Wattenberg. "Artistic Data Visualization: Beyond Visual Analytics." In *Proceedings of the 2nd International Conference on Online Communities and Social Computing*, 182–191. OCSC'07. Springer, 2007. https://doi.org/10.1007/978-3-540-73257-0_21.

Viégas, Fernanda B., and Martin Wattenberg. "TIMELINES: Tag Clouds and the Case for Vernacular Visualization." *Interactions* 15, no. 4 (July 2008): 49–52. https://doi.org/10.1145/1374489.1374501.

Viegas, Fernanda B., Martin Wattenberg, Frank Van Ham, Jesse Kriss, and Matt McKeon. "Many Eyes: A Site for Visualization at Internet Scale." *IEEE Transactions on Visualization and Computer Graphics* 13, no. 6 (2007): 1121–1128.

Ware, Colin. *Information Visualization: Perception for Design*. 2nd ed. The Morgan Kaufmann Series in Interactive Technologies. San Francisco, CA: Elsevier, 2004.

Wattenberg, M. "Baby Names, Visualization, and Social Data Analysis." In *IEEE Symposium on Information Visualization, 2005,* 1–7, INFOVIS '05. https://doi.org/10.1109/INFVIS.2005.1532122.

Wattenberg, Martin, and Fernanda Viégas. "Design and Redesign in Data Visualization." Medium, March 27, 2015. https://medium.com/@hint_fm/design-and-redesign-4ab77206cf9#.a9nusg2ir.

Wilhelmson, Robert B., Brian F. Jewett, Crystal Shaw, Louis J. Wicker, Matthew Arrott, Colleen B. Bushell, Mark Bajuk, Jeffrey Thingvold, and Jeffery B. Yost. "A Study of the Evolution of a Numerically Modeled Severe Storm." *International Journal of High Performance Computing Applications* 4, no. 2 (June 1, 1990): 20–36. https://doi.org/10.1177/109434209000400203.

Williamson, Christopher, and Ben Shneiderman. "The Dynamic HomeFinder: Evaluating Dynamic Queries in a Real-Estate Information Exploration System." In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 338–346. SIGIR '92. New York, NY, USA: ACM, 1992. https://doi.org/10.1145/133160.133216.

Wood. *The Power of Maps*. New York, NY: The Guilford Press, 1992.

Zhang, Mingxin. "Social Network Analysis: History, Concepts, and Research." In *Handbook of Social Network Technologies and Applications*, edited by Borko Furht, 3–22. Springer Science & Business Media, 2010.