Introduction

It's not what you look at that matters, it's what you see.

—Henry David Thoreau

Professional writers and speakers like me live interesting lives. I'd hazard to guess that most of us work from home, although some maintain proper offices. And when you work from home, strange things can happen. For one, it can become difficult to separate work from leisure. There's no boss looking over your shoulder to see if you've completed that TPS report. *Did you get that memo?* If you want to take a nap in the early afternoon as I routinely do, no one's stopping you. In a way, people like me are always *at work*, even though we're not always working. It's fair to say that the notion of work-life balance can be challenging. Lines usually blur. Maybe they're even obliterated.

In many ways, working from home could not be more different from working for "the man." Even today, many rigid corporate environments block employees from visiting certain websites via services like Websense. And forget the obvious sites (read: porn). At many companies, there's no guarantee that employees can access websites that serve legitimate business purposes, at least without a call to the IT help desk to unblock them. Examples include Twitter, Facebook, Tumblr, and Pinterest. Of course, many employees in industrialized countries sport smartphones these days, minimizing the effectiveness of the Websenses of the world. As a result, many companies have reluctantly embraced the Bring Your Own Device movement. That genie is out of the bottle.

We home-based employees, though, don't have to worry about these types of restrictions. No one stops us from wasting as much time as we want on the Web, the golf course, or anywhere else for that matter. In an increasingly

blurry world, though, what does it *really* mean to waste time? That's a bit existential. Let me rephrase: Are my tweets generally work related? Have they changed over time? If so, how?

ADVENTURES IN TWITTER DATA DISCOVERY

Twitter tells me that, since 2010, I have tweeted more than 17,000 times as of this writing, or about ten times per day. I'd wager that more than 70 percent of my tweets were work related. (Yes, I have been paid to tweet. Lamentably, I don't command Kim Kardashian-type rates for my 140 characters.* *Maybe some day.*) Twitter has let me connect with interesting people and organizations, many of whom you'll meet in this book. In the course of researching this book, I searched Twitter for a random sample of thoughts, typically with the hashtag `#dataviz`. At least to me, Twitter is an exceptionally valuable business service that I would gladly pay to use. While we're at it, let's put Twitter client HootSuite in that same boat.

At the same time, though, I unabashedly use Twitter for reasons that have absolutely no connection to work. If you go to `@philsimon` and follow me (please do), there's a good chance that you'll see a few tweets with `#Rush` and `#BreakingBad`, my favorite band and TV show, respectively. What's more, I've tweeted many of these things during times and days when I probably should have been working. I could delude myself, but I won't. A few of my favorite celebrities and athletes have engaged with me on Twitter, bringing a smile to my face. I'll say it: Twitter is fun.

But let's stick with work here. Based on what I'm doing, I suspect that my tweets have evolved over time, but how? It's presumptuous to assume that the content of my tweets is static. (I like to think that I have a dynamic personality.)

To answer this question, I could have accessed my archived tweets via Twitter.com. The company made user data available for download in December 2012. I could have thrown that data into Microsoft Excel or Access and started manually looking for patterns. Knowing me, I would have created a pivot table in Excel along with a pie chart or a basic bar graph. (Yes, I am a geek and I always have been.) The entire process would have been pretty time consuming even though I've been working with these productivity staples for a long time. Let's say that Twitter existed in 1998. If I wanted to visualize and understand my tweets back then, I would have had to go the Microsoft route.

Of course, it's not 1998 anymore. Answering these simple questions now requires less thought and data analysis than you might expect. Technology today is far more powerful, open, user-friendly, ubiquitous, and inexpensive compared to the mid-1990s.

---

* Reportedly, a mind-blowing $10,000 per tweet.

Like many companies today, Twitter relies upon a *relatively* open application programming interface (API).* At a high level, APIs allow devices, apps, and Web services to easily interact with one another. They also facilitate the near-instant flow of data. Lately, APIs have become all the rage. Myriad people use them every day, whether they know it or not. Facebook, LinkedIn, FourSquare, Google, and scores of other companies effectively use APIs for all sorts of reasons. And forget massive tech companies with billion-dollar valuations. Many start-ups are based on "the Twitter fire hose," including the aforementioned HootSuite. Open APIs encourage development of third-party products and services, a topic I discussed in great detail in *The Age of the Platform*.

One such service is Vizify, a start-up founded in 2011 and based in Portland, Oregon. The company is a proud graduate of both Seattle TechStars and the Portland Seed Fund. I fittingly "met" company cofounder and CEO Todd Silverstein over Twitter in June 2013 while researching this book. Vizify quickly and easily lets users connect to different social networks like Facebook, Twitter, FourSquare, and LinkedIn.

It took about three minutes for Vizify to pull my photos, education history, current occupation, work history, home page, tweets, and other key profile data that I've chosen to make publicly available. Of course, users aren't obligated to connect to any individual network. (I passed on FourSquare.) After the initial load, users can easily remove pictures or other information they would prefer not to share. By accessing open APIs, Vizify allows users to create free and interactive visual profiles. Mine is shown in Figure I.1.

If you want to see my full multipage profile, go to https://www.vizify.com/phil-simon. In case you're wondering, users can change the colors on their profiles. I went with that particular shade of green as a homage to *Breaking Bad*.

A snazzy visual profile is all fine and dandy, but it still didn't answer my Twitter question. Fortunately, Vizify also allowed me to effortlessly see the evolution of my tweets over time. A screenshot from that part of my profile is shown in Figure I.2.

Figure I.2 proved what I had suspected. First, I use Twitter for both business and personal reasons. Second, my tweets for `#BigData` began to increase in October 2012. At that time, I was knee-deep into the research for my previous book, *Too Big to Ignore: The Business Case for Big Data*. Before then, I didn't tweet about `#BigData` very often, much less the title of the book (`#TooBigToIgnore`).

But not everything changes—at least with me. For instance, my tweets about `#BreakingBad` and `#Rush` have remained fairly constant over time, with a few notable exceptions. (*Did I really go a whole month in early 2013 without mentioning Canada's finest export on Twitter?)*

---

* It used to be more open and has recently earned the ire of many developers for allegedly heavy-handed tactics. For more on the Twitter API, see https://twitter.com/twitterapi.

Figure I.1 Vizify Phil Simon Profile
Image courtesy of Vizify

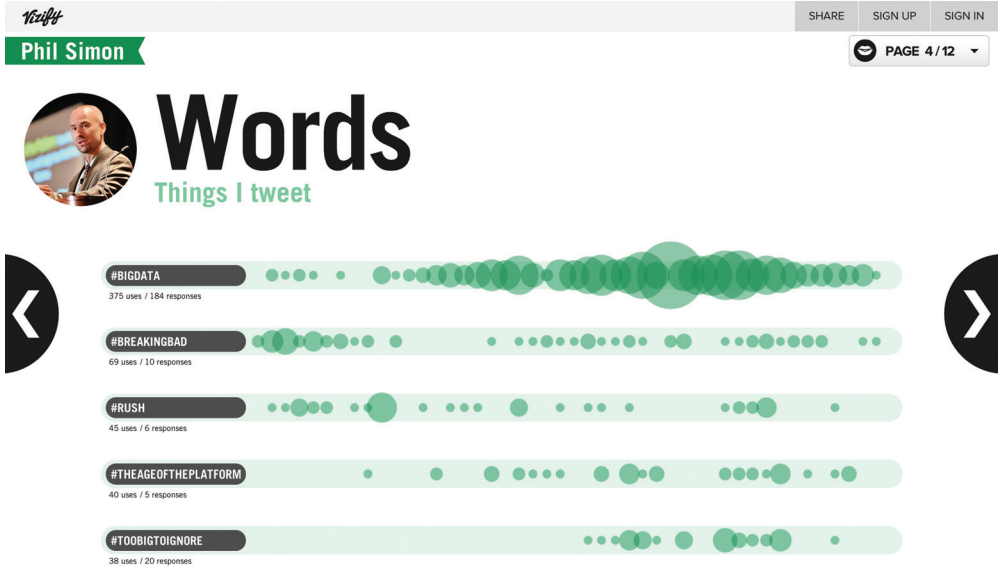

Figure I.2 Vizify Representation of `@philsimon` Tweets
Image courtesy of Vizify

NOTE

Vizify allows users to customize their public profiles, as well as see their other frequently used Twitter hashtags. Figure I.2 shows a snapshot of my top-five hashtags as of June 2013.*

Even though this was a one-time experiment, I could see using Vizify on a regular basis. My tweets will continue to evolve, probably mirroring my professional endeavors and newfound personal interests. Case in point: my publisher has scheduled this book to be released in early 2014. If I run Vizify again around that time, I would assume that many of my tweets will contain the hashtag `#dataviz`. Fortunately, that won't be difficult to discern.

So, my tweets have changed over time, but (as you can probably tell) this process just whet my appetite. I was still curious about my Twitter habits, and other questions remained, like this one: what were my peak tweeting hours? It took only a few clicks to answer that question. Vizify created a personal 30-second video analyzing my tweets.† A screenshot from that video is presented in Figure I.3.

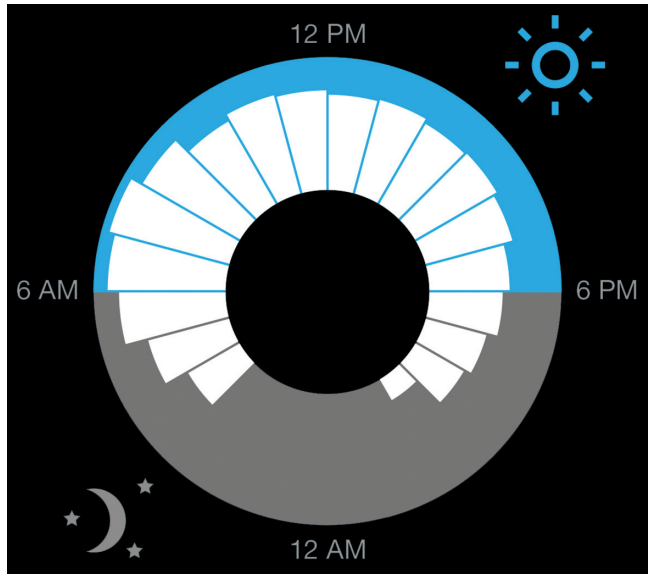

Figure I.3 `@philsimon` Tweets by Hour of Day
Image courtesy of Vizify

* Vizify even lets users create 30-second Twitter videos based on pictures tweeted. To see mine, go to https://www.vizify.com/phil-simon/twitter-video.

† To see my video, go to https://www.vizify.com/phil-simon/twitter-video.

Since I have always been a morning person, it's no shocker that my first tweets start as early as 6 a.m. On most days, I wake up by 5 a.m. and promptly make myself a cup of coffee. I check my e-mail and tweet new posts or articles I've written for my clients or my own sites. I intentionally break up my normal day to give my weary eyes a rest by hitting the gym around 10 a.m. By 6 p.m., I've already put in more than a full day. I'm rarely in front of my computer after that time, although, like many people, I have recently embraced the two-screen experience of tweeting when I watch television. (It's a sign of the times. For many TV viewers today, "It is a common practice to tweet while watching. Nielsen has new research that confirms for the first time that tweets can increase a TV program's ratings."[1])

Vizify confirmed what I expected: I am not much of a late-night tweeter, although I occasionally schedule tweets and let HootSuite auto-tweet for me. (I generally try to space out my tweets, and I don't follow anyone who tweets 34 times per hour. It's fair to say that I have developed my own Twitter philosophy. I'd even call myself a bit of a Twitter snob.)

Aside from my most frequently used hashtags, Vizify also identified my most frequent targets—that is, the people about whom I tweeted most often. I have particular affection for author and professor Terri Griffith (`@terrigriffith`) and blogger Jim Harris (`@ocdqblog`). In Jim's case, the feeling is mutual.*

Vizify let me indulge in what was mostly an intellectual exercise. (I can't say that my boss forced me to geek out.) I was curious about my tweeting history and decided to play around with a new toy, hardly an uncommon occurrence for me. And there is a slew of other toys. For instance, Ionz lets users easily "self-visualize" their Twitter data, and Visually lets users do something similar with their Facebook data. But data visualization is anything but the sole purview of geeks like me with admittedly too much time on their hands. Social networks are becoming more interactive, data driven, and visual.

Twitter senior management pays close attention to what its ecosystem and competition are doing, as it should. Not that Twitter is alone here. For instance, in June 2013, Facebook added verified accounts, support for hashtags, and Vine-like video capabilities to its Instagram app. Seem familiar? Facebook clearly "borrowed" these features from Twitter. Such is life in the Age of the Platform; frenemies and coopetition are the norm. During that same month, Twitter added enhanced native analytics of its own.† I have presented my own in Figure I.4.

I'll spare you any more analysis of my tweets. You get it. This little yarn only serves to illustrate one of the key points in this book: it's never been easier or more essential to visualize data.

---

* You can watch Jim's video here: https://www.vizify.com/jim-harris-1/twitter-video.

† To see yours, just go to http://tinyurl.com/analytics-twitter and log in.

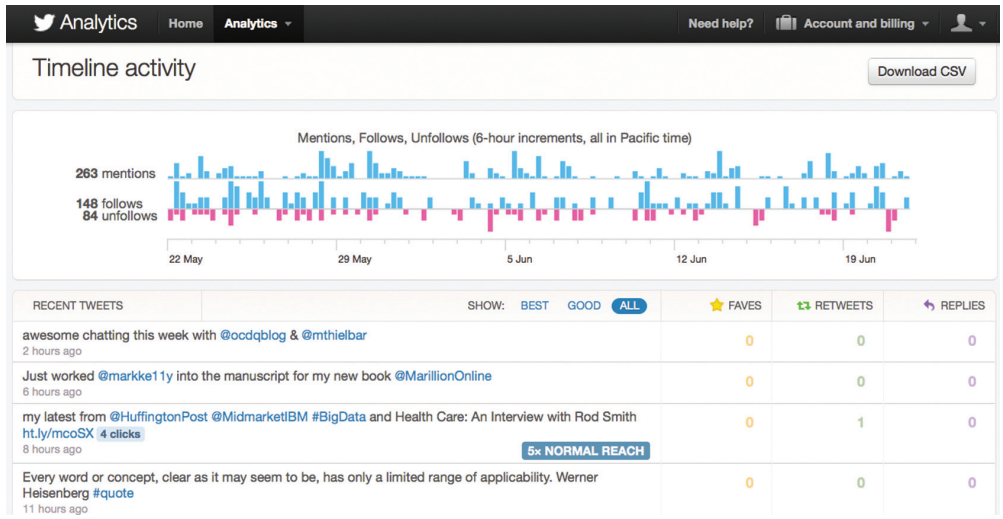

Figure I.4 `@philsimon` Twitter Analytics
Source: Twitter

CONTEMPORARY DATAVIZ 101

Incessant social media, memes, and nonstop content permeate our lives. With seemingly every new hot topic or trend, there's no shortage of definitions, many of which come from people and organizations with vested interests in *their* definition winning (read: consulting firms, software vendors, and thought leaders).

In both *The Age of the Platform* and *Too Big to Ignore*, I devote a fair amount of space to defining in plain English my key terms *platforms* and *Big Data*, respectively. There's so much noise and confusion out there on each topic. I feel the need to do the same here with *data visualization*.

NOTE

In this book, contemporary *data visualization*, or *dataviz*, signifies the practice of representing data through visual and often interactive means. An individual dataviz represents information after it been abstracted in some schematic form. Finally, contemporary data visualization technologies are capable of incorporating what we now call Big Data.

Primary Objective

There's a surfeit of data-oriented terms in the business world right now because data is just plain hot. Let me be absolutely clear here: modern-day dataviz is

not just a synonym or a fancy term for data mining, business intelligence, the many forms of analytics,* or enterprise reporting.

Delineating among all these terms isn't terribly important here. Chapter 2 returns to this subject. For now, suffice it to say that these concepts aren't completely unrelated to one another. In fact, there's a great deal of overlap among them. The most obvious: each is predicated on data in one form or another.

Views on the "proper" goal of dataviz vary considerably. For instance, consider the words of Vitaly Friedman, the editor-in-chief of *Smashing Magazine*, an online periodical for professional Web designers and developers:

> The main goal of data visualization is its ability to visualize data, communicating information clearly and effectively. It doesn't mean that data visualization needs to look boring to be functional or extremely sophisticated to look beautiful. To convey ideas effectively, both aesthetic form and functionality need to go hand in hand, providing insights into a rather sparse and complex dataset by communicating its key aspects in a more intuitive way. Yet designers often tend to discard the balance between design and function, creating gorgeous data visualizations which fail to serve its main purpose—communicate information.[2]

Communicating information is unquestionably important, but not everyone believes in its primacy (read: that it should be *the* goal of dataviz). Dataviz pioneers Fernanda Viégas and Martin M. Wattenberg have suggested that the ideal dataviz goes beyond promoting understanding and communication. Those are short-term goals that should buttress the long-term aims of making better business decisions and even prediction. We'll see in Part II how Visual Organizations use dataviz to do much more than understand what's currently happening.

Dataviz applications are certainly important, but it's best to think of data visualization as more than the output of some software program. So argues Nathan Yau in his 2013 book *Data Points: Visualization That Means Something*. Yau stresses the importance of thinking of dataviz more as a medium than a specific tool. "Visualization is a way to represent data, an abstraction of the real world, in the same way that the written word can be used to tell different kinds of stories," he writes. "Newspaper articles aren't judged on the same criteria as novels, and data art should be critiqued differently than a business dashboard."[3]

I could quote other dataviz experts *ad infinitum,* but I won't belabor the point: opinions on the topic are far from unanimous. For the purposes of this book, dataviz shares the same *ultimate* goal with data mining, business intelligence (BI), analytics, and enterprise reporting: to make more informed

* These include standard analytics, Big Data analytics, visual analytics, not to mention industry-specific analytics like retail, health care, and manufacturing.

business decisions. Contemporary dataviz is primarily a means of exploring data and discovering valuable insights. It is not about reporting *per se* nor is it about creating pretty graphs or charts for the sake of doing so. In other words, the most valuable data visualizations today are often based on the premise that employees don't knowing exactly what they're looking for, much less what they'll find. By exploring the data, employees are *more likely* to discover interesting tidbits or revelations that should drive better decision-making and outcomes. (You won't find too many absolutes and guarantees in this book. I'm a big fan of probabilistic thinking.)

Benefits

To be sure, data doesn't always need to be visualized, and many data visualizations just plain suck. Look around you. It's not hard to find truly awful representations of information. Some work in concept but fail because they are too busy; they confuse people more than they convey information, to paraphrase the late George Carlin.* Visualization for the sake of visualization is unlikely to produce desired results—and this goes double in an era of Big Data. Bad is still bad, even and especially at a larger scale.

John Sviokla serves as the vice chairman of Diamond Management & Technology Consultants. As he writes on the *Harvard Business Review* blog,[4] dataviz confers three general benefits:

1. Great visualizations are efficient. They let people look at vast quantities of data quickly.
2. Visualizations can help analysts or groups achieve more insight into the nature of a problem and discover new understanding.
3. A great visualization can help create a shared view of a situation and align folks on needed actions.

At a high level, Sviokla is spot-on. Consider the following example, as it demonstrates how quickly even a simple dataviz can communicate information. Figure I.5 shows a comical visual of six prominent companies' 2011 org charts.

Would it be hard to write a few sentences on each organization's structure? Of course not. In early 2011, Apple revolved around one iconic man. Even after Steve Jobs's death, his presence is deeply felt throughout the company. For its part, Oracle is still a litigious company. Microsoft is composed of warring factions. Looking at the six images in Figure I.5 represents a quicker and certainly more humorous way of summarizing each company than even my witty text probably could.

---

* You can watch his rant on language here: http://tinyurl.com/carlin-language.

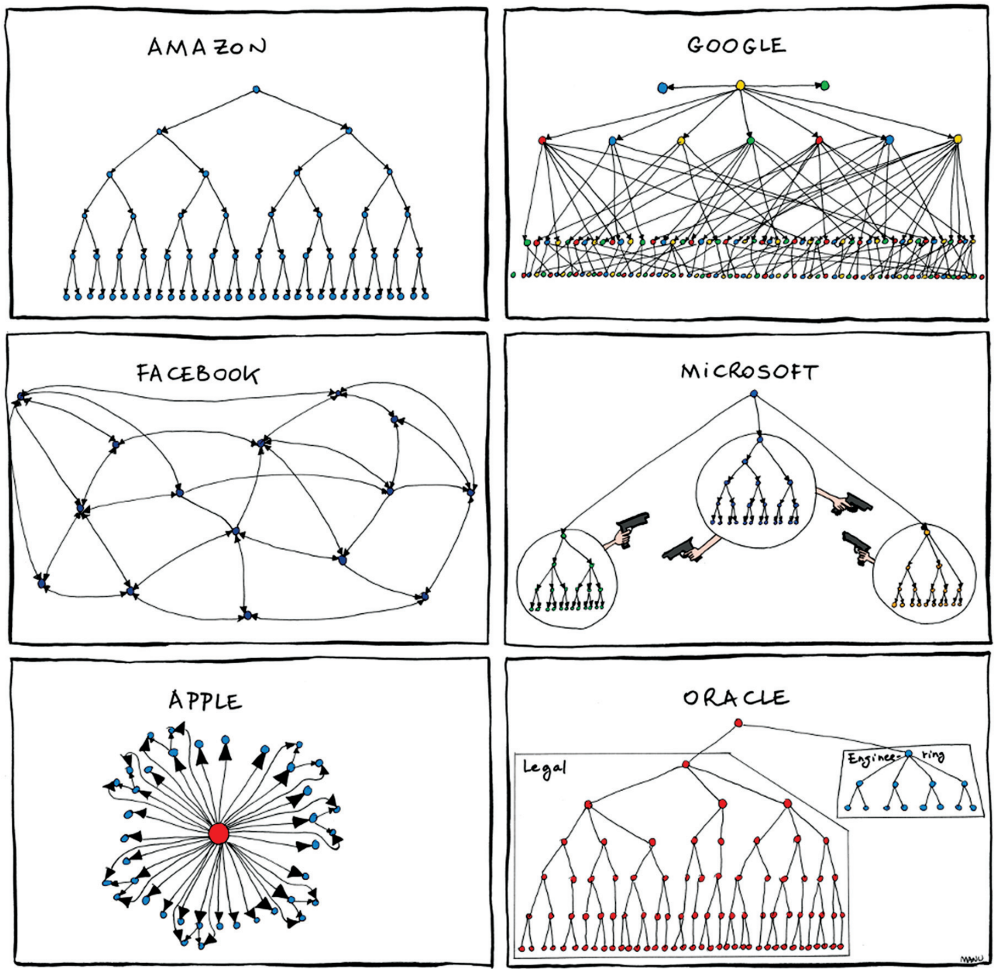

Figure I.5 Organizational Charts (2011)
Source: Manu Cornet

I would add that, for the purposes of understanding large, unpredictable datasets, *interactive* data visualizations are generally superior to static infographics, dashboards, and standard reports. (I should know. I've designed thousands of the latter in my consulting career for my clients.) By definition, presenting even Small Data in predetermined, static, noninteractive formats limits what users can do with—and ultimately get from—data. This has always been the case. In other words, these types of formats generally preclude people from interacting with the data.* They can't drill down and around. They can't explore, nor can they ask iterative and better questions, and ultimately find answers.

* Good report writers know that it's not terribly difficult to add some level of interactivity to static reports. For one example of how to do this, see http://tinyurl.com/phil-crystal.

More Important Than Ever

There are myriad questions that we can ask from data today. As such, it's impossible to write enough reports or design a functioning dashboard that takes into account every conceivable contingency and answers every possible question.

Like all sentient beings, we humans have *always* processed information in different ways, or at least attempted to do so. Many researchers have proven that there is something unique about how we *see* information, as opposed to how we hear it. Many excellent studies and books have informed our current understanding of the workings of the human brain, and I certainly won't attempt to summarize them all here. *The Visual Organization* is in no way a book about the behavioral sciences, neurology, or cognitive psychology. I will, however, concisely mention a few of the more important works in those fields.

"The human visual system is a pattern seeker of enormous power and subtlety," writes Colin Ware in his classic text *Information Visualization: Perception for Design*. "The eye and the visual cortex of the brain form a massive parallel processor that provides the highest bandwidth channel into human cognitive centers. At higher levels of processing, perception and cognition are closely interrelated, which is why the words *understanding* and *seeing* are synonymous."[5] Our brains are wired to process information better in a visual manner.

That humans tend to comprehend visual information quicker than raw data doesn't mean that all visualizations are created equal. On the contrary, we understand certain types of graphical representations better than others. Researchers William S. Cleveland and Robert McGill proved as much in September 1984. Cleveland and McGill published a paper in the *Journal of the American Statistical Association* titled "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods." Cleveland and McGill studied the visual clues that people are able to decode most accurately. The two ranked these clues in the following list:

1. Position along a common scale, e.g., scatter plot
2. Position on identical but nonaligned scales, e.g., multiple scatter plots
3. Length, e.g., bar chart
4. Angle and slope (tie), e.g., pie chart
5. Area, e.g., bubbles
6. Volume, density, and color saturation (tie), e.g., heat map
7. Color hue, e.g., news map

A slightly modified visual of this list is presented in Figure I.6. In English, it means that not all people comprehend and decode all visual cues equally. For instance, we tend to understand data positioned along a

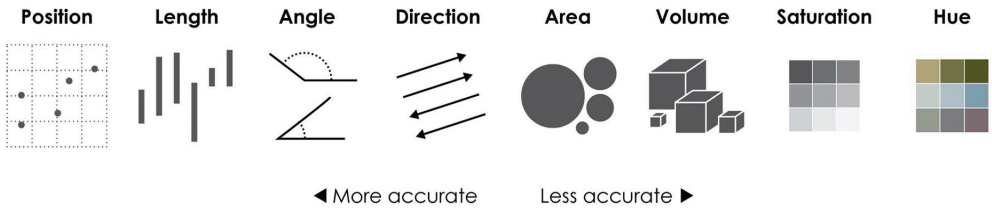

Figure I.6 Visual Cues Ranking
Source: Reprinted with permission from *The Journal of the American Statistical Association*. Copyright 1984 by the American Statistical Association. All rights reserved.

common scale better than data shown on heat maps. Note, however, that the findings of Cleveland and McGill should not be seen in absolute terms. The study suggests that absolutes are a myth and that the ability to understand visual clues is situational. For example, some people will understand a bar chart better than a bubble chart. The Cleveland and McGill recommendations are just general guidelines.

In their 2012 book *Infographics: The Power of Visual Storytelling*, Jason Lankow, Josh Ritchie, and Ross Crooks demonstrate how even very simple formatting can make certain data stand out at the expense of other data. Consider Figure I.7, a series of random numbers. Go ahead and find each instance of the number 7.

Now, with simple formatting, repeat the same exercise with Figure I.8.

In professional settings, data has always mattered, although some departments and industries have been more likely to embrace it than others. In this

2	1	4	3	9	5	6	7	8	2	3	6	5	9	4	0	1
6	7	9	3	4	9	0	5	6	2	5	8	4	0	5	2	6
9	8	2	6	3	5	9	3	2	9	3	7	2	6	3	4	8
8	1	6	2	3	8	7	9	5	0	2	3	9	2	8	4	3
0	9	1	8	5	4	2	9	4	7	4	6	8	4	0	2	9
3	9	2	7	3	6	6	5	2	9	4	0	4	9	4	8	6
5	2	4	3	6	4	8	1	0	3	9	4	8	4	7	3	2
8	6	2	3	0	8	7	3	6	2	5	4	4	8	3	5	0

Figure I.7 Preattentive Processing Test 1
Source: Lankow, Ritchie, and Crooksa

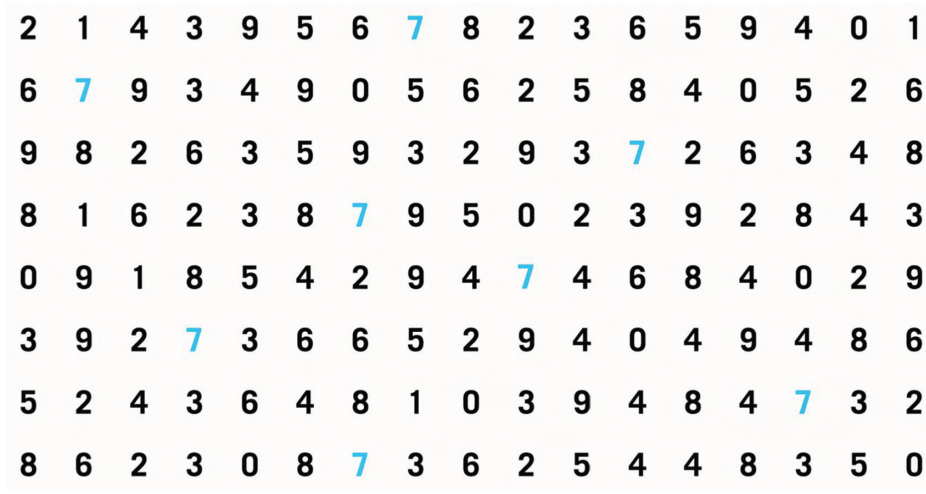

Figure I.8 Preattentive Processing Test 2
Source: Lankow, Ritchie, and Crooks

book, I contend that data visualization has never been more important. Chapter 1 will have a great deal more to say about the rise of the Visual Organization. For now, suffice it to say that representing information in schematic forms has always been essential to human understanding.

We acquire more information through our visual system than we do through all our other senses *combined.* We understand things better and quicker when we see them.

Revenge of the Laggards: The Current State of Dataviz

Fifteen years ago, the presentation of data wasn't terribly democratic, sophisticated, and interactive, especially compared to today. Tech-savvy analysts and IT professionals generated static diagrams, graphs, and charts for quarterly or annual meetings or "special events." Back then, cutting-edge dataviz wasn't part and parcel to many jobs. There just wasn't that much data, especially compared to today.

In a way, this was entirely understandable. Yes, the late-1990s saw the advent of modern enterprise reporting and BI applications adroit at representing mostly structured data. In most organizations, however, relatively few people regularly visualized data, at least not *on a regular basis*.

My, how times have changed. Now data is everywhere. As I wrote in *Too Big to Ignore*, we are living in the era of Big Data, and many things are changing. In the workplace, let's focus on two major shifts. First, today it is becoming incumbent upon just about *every* member of a team, group, department, and organization to be, at a minimum, comfortable with data. Fewer and fewer knowledge workers can hide from quantitative analysis. Second, pie charts, bar charts,

and other simple data visualizations of 15 years ago now seem quaint. They don't remotely resemble anything that qualifies as contemporary dataviz. More important, today they often fail to tell the stories that need to be told.

Next, data no longer needs be presented on an occasional or periodic basis. We are *constantly* looking at data of all types—a trend that will only intensify in the coming years. Before our eyes, we are seeing the ability to effectively present quantitative information in a compelling manner become a professional *sine qua non*. Hidden in the petabytes of structured and unstructured data are key consumer, employee, and organizational insight. *If found and unleashed*, those insights would invariably move the needle.

The PwC survey confirmed what I have long assumed. Although notable exceptions exist, only a minority of organizations and professionals currently do very much with dataviz. Most enterprises fail to present data in visually compelling ways. Far too many rely upon old standbys: bar charts, simple graphs, and the ubiquitous Excel spreadsheet. And their business decisions suffer as a result.

Why the widespread lack of adoption? I'd posit that several factors are at play here. First, while dataviz is hardly new, the landscape is. Many of the applications and services detailed in Chapter 2 are recent advents. Second, I have little doubt that the explosions of dataviz and Big Data left many CXOs overwhelmed. In this way, dataviz is much like cloud computing. With myriad options, it's natural for those in control of the purse strings to ask, "Where do we even start?"

Next, many organizations suffer from downright ignorance. Many lack the knowledge that better tools exist, not to mention that enterprises are successfully using them. (Hopefully, this book will change that, at least to some extent.) Then there are organizations whose cultures systematically ignore data and analysis. I have seen my share of those As such, their employees generally lack the willingness to try, buy, deploy, and use contemporary dataviz tools. When corporate fiat, culture, and politics dominate decision-making, what's the point of even looking at data?

For these reasons, it should be no surprise that Big Data is still in its infancy. Brian McKenna tackles this subject in an April 2013 ComputerWeekly article. About the state of Big Data, he writes that "Analytics firm SAS and SourceMedia surveyed 339 data-management professionals about their organizations' use of

NOTE

The hype around Big Data and, to a lesser extent, dataviz still far exceeds their business realities. To quote former Notre Dame coach Lou Holtz, "When all is said and done, more is said than done." Rather than hem and haw, organizations should recognize the vast opportunity that the status quo represents. Those that act now can realize significant benefits that won't be available to them once their competition wakes up.

THE CURRENT DATA ON DATAVIZ

Sadly, most employees—and, by extension, departments and organizations—don't capitalize on the massive opportunities presented by Big Data and data visualization. So says consulting firm PricewaterhouseCoopers (PwC) in its fifth annual Digital IQ Survey (titled "Digital Conversations and the C-suite").

In 2013, PwC surveyed 1,108 respondents from 12 countries across a variety of industries. Respondents were equally split between IT and business leaders. More than 75 percent worked in organizations with revenues of more than $1 billion.*

FINDINGS

A majority of respondents (62 percent) think that Big Data can provide a competitive advantage. That's not exactly surprising, but believing in the power of Big Data is hardly the same as turning it into actual business insights—and then acting upon them. Nearly the same number of respondents (58 percent) agreed that moving from data to insight is much easier said than done.

Only 26 percent of global survey respondents are currently using dataviz. (I suspect that many of these "forward-thinking" organizations aren't exactly Google-like in their execution.) Interestingly, though, adoption—or lack thereof—is not evenly distributed among all respondents. Specifically, those that reported revenue growth in excess of 5 percent led the pack—and weren't letting up. They planned to invest more in data visualization in 2013. The same can be said of organizations in the top quartile for revenue, profitability, and innovation. The gap between the dataviz haves and have-nots seems to be growing.

OBSTACLES

Organizations face four major obstacles with respect to Big Data:

1. They are blind to the importance of visualization.
2. They are investing more in gathering data than analyzing it.
3. They are facing a talent gap.
4. They are struggling with insufficient systems to rapidly process information.

"The amount of information and data that we're collecting now is truly enormous, [especially] the volume that is outside the four walls of the organization," says Anand Rao, principal at PwC. "Organizations don't have the right people, they don't have the right structure in place, and they're still struggling with some of the tools and techniques."[6]

Rao points out that many organizations do a passable job at looking backward—that is, "hindsight analysis." Far fewer, though, predict very well. As we'll see throughout this book, dataviz can be useful in this regard.

* Access the entire report at http://tinyurl.com/pwc-dv-2.

data-management technology in December 2012, discovering reality still lags behind its hype. Only 12 percent of information management professionals are doing Big Data, according to a recent survey."[7] Remarkably, only 14 percent of respondents categorized their organizations as "very likely" to begin working with Big Data in 2014. Nearly one in five responded "not likely at all."

BOOK OVERVIEW

Big Data is here, leaving many organizations and their employees overwhelmed. Fortunately, new data-visualization applications are helping enterprises isolate the signal in the noise.

For instance, through interactive dataviz tools, Netflix discovers trends, diagnoses technical issues, and unearths obscure yet extraordinarily valuable customer insights. Employees at Autodesk use a remarkable and interactive tool that visualizes current and historical employee movement. From this, they identify potential management issues and see what a corporate reorg *really* looks like. Through cutting-edge dataviz, start-up Wedgies instantly serves up real-time poll results while monitoring poll traction and site issues. The University of Texas is bringing a visual type of transparency to academia. It makes unprecedented amounts and sources of institutional data available on its website. Anyone with the desire and an Internet connection can slice and dice a mountain of its data in myriad ways. And then there's eBay. Powerful data-discovery tools allow its employees to effectively "see" what ebay.com would look like as a brick-and-mortar store.

These progressive organizations are the exceptions that prove the rule. Most enterprises are woefully unprepared for Big Data. Far too many erroneously believe and act like nothing has really changed. As such, they continue to depend exclusively on reporting stalwarts like Microsoft Excel, static dashboards, basic query applications, and even traditional business intelligence tools. In so doing, they are missing out on the tremendous opportunities that new data sources and dataviz tools can provide.

Amidst all the hype and confusion surrounding Big Data, though, a new type of enterprise is emerging: the Visual Organization. An increasing number of organizations have realized that today's ever-increasing data streams, volumes, and velocity require new applications. In turn, these new tools promote a different mind-set—one based upon data discovery and exploration, not on conventional enterprise reporting. Interactive heat maps, tree maps, and choropleths promote true data discovery more than static graphs and pie charts.

Today, a growing number of enterprises have turned traditional dataviz on its head. In their stead, they are embracing new, interactive, and more robust tools that locate the signals in the noise that is Big Data. As a result, these enterprises are asking better questions of their data—and making better business decisions.

The Visual Organization is a largely positive and forward-thinking book. I focus more on profiling the organizations and employees who get it, not excoriating the ones that don't. (Trust me. There is no shortage of the latter.) Where warranted, I do attempt to explain the reasons behind certain types of stasis, dysfunction, and failure. These observations are based upon both my research for this book and the decade I spent as an enterprise IT consultant. Let me be clear: my goal here is *not* to harp on the negative. Rather, I merely want readers to understand the ways in which Visual Organizations differ from less progressive enterprises. As Bill Gates once said, "It's fine to celebrate success, but it is more important to heed the lessons of failure."

In the following pages, you'll meet some amazing companies and people who recognize the power of Big Data and dataviz. They are pushing the envelope and looking at problems very differently than their data-challenged counterparts. And they are seeing their efforts bear fruit.

Defining the Visual Organization

While useful and informative, many of the texts on data visualization emphasize theory more than practice. *The Visual Organization* does not. The forthcoming chapters introduce some fascinating practitioners who regularly visualize data to understand it, interpret it, and ultimately take action on it. You'll discover, as I did in researching this book, that Visual Organizations have moved well beyond simple charts, graphs, and dashboards that play nice with structured, transactional data—aka, *Small Data*.* They are using new tools to make sense of unstructured data, metadata (data about data), and other emerging data types and sources. And, as you'll see, the results are impressive.

NOTE

Since this is a book about Visual Organizations, a short, formal definition is in order:

A Visual Organization is composed of intelligent people who recognize the power of data. As such, it routinely uses contemporary, powerful, and interactive dataviz tools to ask better questions and ultimately make better business decisions. As we'll see in Chapter 6, the notion of a Visual Organization is not binary; there are four levels. More advanced enterprises use interactive data-visualization applications to analyze Big Data. They recognize the inherent limitations of Small Data and static dataviz.

Central Thesis of This Book

The Visual Organization is based on a simple premise. The Data Deluge has arrived, and it isn't going anywhere. More than ever, employees and organizations

*Examples include a list of sales or employees. Think orderly and Excel-friendly data.

have to process and understand unprecedented amounts of information—*or at least try*. Complicating matters, new types and sources of data are flying at us faster than ever. Consider this amazing fact from *The Human Face of Big Data*, a fascinating book by Rick Smolan and Jennifer Erwitt. Today the average man is exposed to more data in a single day than his fifteenth-century counterpart was in his entire lifetime! According to an oft-cited March 2013 U.N. study, today more people can access cell phones than toilets.* Out of an estimated 7 billion people on the planet, roughly 6 billion can use mobile phones. Only 4.5 billion can say the same about working commodes.

Alternatively stated, data is streaming at us with increasing variety, velocity, and volume, with no discernible end in sight. These are the well-documented three *v's* of Big Data. Against this backdrop, intelligent organizations have realized several things. First, data visualization is becoming essential, and not just to manage discrete events. Visual Organizations benefit from *routinely* visualizing many different types and sources of data. Doing so allows them to garner a better understanding of what's happening and why. Equipped with this knowledge, employees are able to ask better questions and make better business decisions. As companies like Amazon, Apple, Facebook, Google, Twitter, Netflix, and others have shown, discoveries from Big Data can represent a huge competitive advantage. To do this, they have had to buy and build new tools. Yes, old standbys like Microsoft Excel spreadsheets, charts, dashboards, key performance indicators, and even mature business intelligence tools still matter. *By themselves, however, they are no longer sufficient to cope with the Data Deluge*†.

This is not a book about how to visualize data *per se*. Rather, it is a book about Visual Organizations.

Bottom line: we live in a world rife with Big Data. Organizations and their employees need different applications to find the needles buried in the haystacks, comprehend immense and dynamic datasets, and ultimately make better business decisions.

Cui Bono?

In any given month, I typically talk to a wide variety of people: CXOs, consultants, freelancers, mid-level managers, entry-level employees, unemployed professionals, journalists, fellow authors and speakers, professors, and college and graduate students. Some live in the United States, others abroad. They work at organizations that run the gamut: tiny start-ups, small businesses, and large corporations. And they work for nonprofits, government agencies, and the private sector. Although the conversations vary, I have noticed a recurring

*To read more, go to http://tinyurl.com/un-toilets.

†Doug Laney of Gartner coined the three v's in February of 2001. For more on this, see http://tinyurl.com/gartnervs.

theme over the past few years: most people are simply overwhelmed by data. They are struggling to cope with this deluge.

I wrote *The Visual Organization* for all of these people.

At its core, this book demonstrates how intelligent people and organizations are making better business decisions via contemporary dataviz new data visualization applications. Contemporary dataviz is no longer just nice to have or fodder for quarterly presentations. Organizations are increasingly embracing new dataviz tools, Big Data, and, most important, a new, data-driven mind-set. Visual Organizations and their employees are handling the Data Deluge better than their "visually challenged" counterparts. Finally, they distinguish between traditional reporting and data discovery.

In the forthcoming chapters, I'll demonstrate that dataviz is becoming indispensable, but make no mistake: it is no elixir. It does not solve every conceivable business problem. No matter how insightful, no matter how much data they present, data visualizations do not always provide the right answer, much less guarantee flawless execution. Often a dataviz only serves to clarify an existing issue, and there's no guarantee that it will shed light on every possible problem.

Limitations aside, the need for—and power of—dataviz has never been more pronounced, a fact that the Visual Organizations profiled in this book and their employees completely understand.

Methodology: Story Matters Here

Of all the companies started around the time of the dot-com boom, Amazon remains one of its few survivors. Calling it a *survivor*, however, is the acme of understatement. The company is nothing short of a titan—the Walmart of the Internet. And Amazon is causing unexpected ripple effects for a slew of companies and industries.

As I write these words, Oracle and its CEO Larry Ellison are forging partnerships with longtime rivals Microsoft and Salesforce.com.* The companies are putting aside their acrimonious histories with one other. They have struck an important alliance that attempts to preserve their footholds in the enterprise. At the core of their newfound and unexpected cooperation: a common fear of Jeff Bezos's firm. Amazon is a threat to them all. Ellison, Salesforce.com head Eric Benioff, and Microsoft big kahuna (at least, as of this writing) Steve Ballmer clearly understand the old Arabic proverb, "The enemy of my enemy is my friend."

Despite Amazon's longstanding prominence, the purportedly definitive text on was only recently written. Magazines like *Wired* have covered different

* The Oracle alliance put Salesforce.com's cloud-based CRM software atop Oracle apps and infrastructure. Cats and dogs living together…

aspects of the company very well and in some depth. Nearly 20 years after its founding, Amazon lacks the equivalent of an authorized tell-all, a comprehensive window into its vast business. Up until recently, the books written about the company have been at best incomplete and at worst disappointing. That finally changed in October 2013. *Bloomberg Businessweek* reporter Brad Stone published his much-anticipated book *The Everything Store: Jeff Bezos and the Age of Amazon*. The book is the closest thing available to a comprehensive company biography. Stone interviewed hundreds of former executives and operations, and I have eagerly followed the status of his book since it was announced.

As an author and occasional journalist, I am familiar with these types of press-related obstacles. (Maybe privacy isn't *completely* dead after all.) In researching previous books, I have contacted folks at high-profile companies, some of whom I would even call *friends*. My requests to speak on the record to employees in the know were politely denied, whether it was about privacy at Google or about Big Data at Facebook. In each case, these folks kindly told me that, as much as they may want to help me, their employers took controlling the message very seriously. I was disappointed but not offended. I understood. Access to senior management about proprietary or sensitive subjects isn't easy to come by, especially if the result is a book or an article.

None of this should be surprising. Steve Jobs only agreed to an authorized biography with Walter Isaacson when the former faced his imminent mortality. For years Jobs denied requests by authors and publishers to do the same thing. Like Jobs, Bezos is by all accounts a very private person, and Amazon follows the lead of its iconic CEO. Letting journalists and authors into their walled gardens ultimately serves no real business purpose. The risks far outweigh the rewards. Companies on that level aren't exactly hurting for PR, and flying under the radar suits them just fine. Sanctioned books like *In the Plex: How Google Thinks, Works, and Shapes Our Lives* by Steven Levy are the exceptions that prove the rule. (Levy's access to Google was unprecedented.)

In 2009, the AMC Network launched a new slogan: Story Matters Here. I couldn't agree more. For a book like *The Visual Organization* to work, I would have to do a good bit of research. That meant identifying organizations visualizing their data in interesting ways, making better business decisions as a result. In the Internet age, I knew that that wouldn't be terribly hard to do. Aside from my personal connections, I could use Google, Facebook, Twitter, LinkedIn, and other indispensable sites for research purposes.

But that wasn't all. To do this book right, I needed to do two other things. First, I would have to find dataviz practitioners doing cutting-edge things—and

then talk to them. The dozens of conversations I had with dataviz professionals inform the pages that follow, whether or not I ultimately profiled their companies and clients. I learned a great deal, as I hope you will.

I have always aspired to write more books that are more "show me don't tell me." To that end, I knew that the book's case studies would be key—and I set that bar relatively high. I was clear with interviewees from the get-go. For their organizations to be featured in the text, they would have to get specific. I would need them to provide actual examples of the dataviz tools they used to do their jobs. Platitudes just wouldn't cut it.

Now, this wasn't my first rodeo. I knew that my self-imposed second requirement would pose more challenges than my first. For instance, United Parcel Service uses technology and data in truly amazing ways. UPS routes its trucks to millions of homes and businesses each day in an efficient manner. This process requires incredibly sophisticated algorithms. I am quite certain that the company's use of data visualization is book-worthy. In June of 2013, I reached out to a friend of mine, a UPS employee for more than two decades. My friend told me exactly what I expected: UPS keeps a low profile and does not like to be featured in magazine articles and books. Including fresh UPS material and examples of its proprietary tools in this book would require approval at the highest level of the company. (Can someone say *lawyers*?) Unfortunately, my efforts to include UPS went nowhere, as did similar attempts to pioneer new dataviz research on universally recognizable organizations like the National Basketball Association, Facebook, Twitter, ESPN, Pandora, and a few others. Just because these companies are not profiled in *The Visual Organization* doesn't mean that they're not doing fascinating things with data and dataviz.

With the exception of Netflix, the case studies in Part II meet *both* of my two criteria. (I'll explain the reasons for the slightly different methodology for Netflix in Chapter 3.) Profiling only relatively forthcoming organizations with remarkable dataviz stories has resulted in a better book. Such examples will, I hope, teach the reader important lessons about the subject, including what to do, what not to do, how to do it, and more. To me, a "story-centric" approach just made sense. It is superior to one that emphasized company notoriety at the expense of specifics and transparency. In the end, I believe that how, why, and what are more important than who. *The Visual Organization* benefits from profiling organizations with compelling and specific examples of contemporary data visualization, even if a few of those organizations aren't necessarily household names. And, as I'll argue in the following pages, these lesser-known enterprises may well become more recognized and successful precisely because they understand the tremendous value that data and dataviz offer.

The Quest for Knowledge and Case Studies

One day in August 2013, a graduate student by the name of James Eichinger tweeted at me. @Ikejames101 is studying predictive analytics at Northwestern University. In a subsequent e-mail, Eichinger informed me that wanted to do his final project on data visualization, but he was encountering a major problem. In his words, "Most of the case studies [he] found are either weak or tangential to the subject."[8] As for the blog posts on sites like *Harvard Business Review*, "Not a single one [shows] impacts on business decisions, management culture, or information efficiency." I didn't entirely concur with Eichinger's assessment, but our exchange piqued my curiosity about the prevalence of proper data-visualization case studies on the Web. I slept on it.

The next morning, I performed three specific Google searches. I queried existing case studies related to three different types of major enterprise technologies. The results are presented in Table I.1, and is displayed graphically in Figure I.9.

Even though Table I.1 and Figure I.9 confirmed my suspicions, it should *not* be taken as gospel or proof that the dataviz case study landscape is entirely barren. For three reasons, I wouldn't go that far.

First, think of Google's search data here as a proxy of sorts.* Without carping over the proper definition of the term *case study*, I have no doubt that there are more than 23 dataviz stories on the Web. (How many of them are actually good, useful, and vendor neutral is another matter altogether.)

Second, quantity should never be mistaken for quality. Many of the ERP and CRM case studies on the Web aren't terribly instructive. Third, by default, Google provides increasingly personalized results based upon factors like user geography, known demographic information, individual browsing history, and others.† Sometimes identical Google searches from ostensibly similar users yield wildly different results.

Table I.1 Google Search Results on Three Different Types of Case Studies

Google Search Term (with Quotes)	Approximate Results	Notes
"Data Visualization case studies"	23	Interestingly, two of the 23 results came from www.philsimon.com
"ERP case studies"	6,670	Enterprise resource planning
"CRM case studies"	16,500	Customer relationship management

Source: Google, as of August 31, 2013

* As any experienced Googler knows, small changes in search terms can yield vastly different results.

† Users can easily turn this "feature" off if they like.

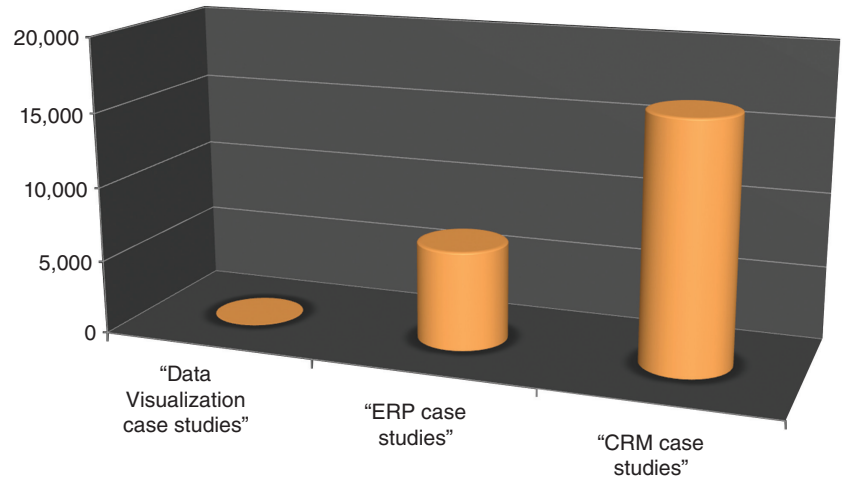

Figure I.9 Graph of Google Search Results on Three Different Types of Case Studies
Data Source: Google, as of August 31, 2013

Despite these qualifications, the differences in my little case study experiment are irrefutable. The number of CRM and ERP profiles are larger than their dataviz counterparts by orders of magnitude (717 and 290 times, respectively). Put differently, Table I.1. and Figure I.9 only illustrate what I, Eichinger, and countless others have found: profiles of organizations using contemporary dataviz and new forms of data not written by marketing departments are lacking. Data visualization is becoming critical, but how have organizations are *really* doing it? And what are the lessons? What are the pitfalls?

And that's where *The Visual Organization* steps in. It is my sincere hope that the observations, framework, case studies, interpretation, and original research in this book will help organizations use dataviz to move their needles. Academic studies and journal articles are certainly beneficial, but as of this writing there is a paucity of vendor-neutral case studies on the subject. Along these lines, perhaps this book will also make a meaningful contribution to the field of data visualization.

Differentiation: A Note on Other Dataviz Texts

The Visual Organization is hardly the first book about dataviz. On the contrary, many other researchers, authors, and practitioners have contributed a great deal to the field. It's no understatement to say that a vast body of work has been done on the topic.

In their books, Stephen Few, Edward Tufte, Alberto Cairo, Colin Ware, and Nathan Yau explain how to effectively visualize data very well. They cover the mechanics of creating graphs, charts, and, more recently, infographics,

heat maps, tree maps, and choropleths. Many of these authors' books illustrate best design practices and serve as how-to guides, and I recommend checking them out. For their parts, dataviz researchers like Marek Walczak, Martin M. Wattenberg, and Fernanda Viégas have gone way beyond extending our current understanding of dataviz. They have created exciting new ways to visualize data. *The Visual Organization* does not attempt to replicate their work here.

Nor is this a text primarily about how the human brain processes data. I don't cover the science behind the mind's ability to understand information represented in a visual form. I'm the furthest thing from a neurologist. Again, a panoply of excellent books has already been written on the subject. The bottom line, as data journalist John Burn-Murdoch writes in *The Guardian*, is that "Humans are visual creatures. Peer-reviewed studies have shown that we can consume information more quickly when it is expressed in diagrams than when it is presented as text."[9]

The Visual Organization demonstrates how and why a growing number of organizations are visualizing their data to diagnose issues, discover new customer insights, and make better decisions.

Plan of Attack

The Visual Organization consists of four parts. Part I, "Book Overview and Background," examines the reasons behind the ascent of the Visual Organization. It also covers the five general categories of contemporary dataviz applications and services.

Part II, "Introducing the Visual Organization," introduces a number of diverse Visual Organizations. You'll discover how Netflix, Wedgies, Autodesk, and other enterprises have embraced Big Data and dataviz, and not just as discrete one-time "projects." We'll see how Visual Organizations have garnered profound customer insights and solved thorny business problems through new dataviz techniques and applications.

Part III, "Getting Started: Becoming a Visual Organization," takes a step back. It begins by providing a framework for readers to understand the four different levels of Visual Organizations. It then asks a key question before extrapolating a series of lessons, best practices, myths, and mistakes from the case studies in Part II. No, it's not a checklist to follow for becoming a Visual Organization, but it does present sage advice for readers interested in both reaping the benefits of dataviz and avoiding their common pitfalls.

Part IV, "Conclusion and the Future of Dataviz," concludes the book. It offers a number of careful predictions about current trends, Visual Organizations, Big Data, and the future of data visualization.

NEXT

Chapter 1 examines the ascent of the Visual Organization. It explains what's happening and why. We'll soon see that important business, cultural, technological, and human shifts are collectively causing enterprises of all kinds to change the way they think about data and traditional reporting.

NOTES

1. Snider, Mike, "Twitter Can Boost TV Ratings," *USA Today*, August 6, 2013, http://www.usatoday.com/story/tech/personal/2013/08/06/nielsen-twitter-affects-tv-ratings/2613267, Retrieved August 27, 2013.
2. Friedman, Vitaly, "Data Visualization and Infographics," *Smashing Magazine*, January 14, 2008, http://www.smashingmagazine.com/2008/01/14/monday-inspiration-data-visualization-and-infographics, Retrieved June 12, 2013.
3. Yau, Nathan, *Data Points: Visualization That Means Something*, Hoboken, NJ: Wiley, 2013.
4. Sviokla, John, "Swimming in Data? Three Benefits of Visualization," *Harvard Business Review* Blog Network, December 4, 2009. http://blogs.hbr.org/sviokla/2009/12/swimming_in_data_three_benefit.html, June 11, 2013.
5. Ware, Colin, *Information Visualization: Perception for Design,* Morgan Kaufmann, 2000.
6. Olavsrud, Thor, "4 Barriers Stand Between You and Big Data Insight," CIO.com, April 9, 2013, http://www.cio.com/article/731503/4_Barriers_Stand_Between_You_and_Big_Data_Insight, Retrieved August 27, 2013.
7. McKenna, Brian, "SAS: Data Quality, Data Governance Concerns Impede Big Data Programmes," ComputerWeekly.com April 3, 2013, http://www.computerweekly.com/news/2240180600/SAS-data-quality-data-governance-concerns-impede-big-data-programmes, Retrieved August 30, 2013.
8. E-mail from Eichinger, August 31, 2013.
9. Burn-Murdoch, John, "Why You Should Never Trust a Data Visualisation," theguardian.com, July 24, 2013, http://www.guardian.co.uk/news/datablog/2013/jul/24/why-you-should-never-trust-a-data-visualisation, Retrieved July 24, 2013.