The Big Book of Dashboards

Visualizing Your Data Using Real-World Business Scenarios
THE
BIG BOOK
OF
DASHBOARD
Cover image: Course Metrics Dashboard by Jeffrey Shaffer
Cover design: Wiley

Copyright © 2017 by Steve Wexler, Jeffrey Shaffer, and Andy Cotgreave. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the Web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at http://booksupport.wiley.com. For more information about Wiley products, visit www.wiley.com.

A catalog record for this book is available from the Library of Congress.

ISBN 978-1-119-28271-6 (Paperback)
ISBN 978-1-119-28278-5 (ePDF)
ISBN 978-1-119-28273-0 (ePub)

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1
## Contents

Acknowledgments vii
About the Authors ix
Introduction xi

**PART I  A STRONG FOUNDATION**

Chapter 1  Data Visualization: A Primer 2

**PART II  THE SCENARIOS**

Chapter 2  Course Metrics Dashboard 38
Chapter 3  Comparing Individual Performance with Peers 48
Chapter 4  What-If Analysis: Wage Increase Ramifications 62
Chapter 5  Executive Sales Dashboard 70
Chapter 6  Ranking by Now, Comparing with Then 80
Chapter 7  Are We on Pace to Reach Our Goals? 92
Chapter 8  Multiple Key Performance Metrics 98
Chapter 9  Power Plant Operations Monitoring 106
Chapter 10  Showing Year-to-Date and Year-over-Year at the Same Time 118
Chapter 11  Premier League Player Performance Metrics 130
Chapter 12  RBS 6 Nations Championship Match Performance Analysis 138
Chapter 13  Web Analytics 146
Chapter 14  Patient History Analysis of Recent Hospital Admissions 156
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Hospitality Dashboard for Hotel Management</td>
<td>164</td>
</tr>
<tr>
<td>16</td>
<td>Sentiment Analysis: Showing Overall Distribution</td>
<td>174</td>
</tr>
<tr>
<td>17</td>
<td>Showing Sentiment with Net Promoter Score</td>
<td>186</td>
</tr>
<tr>
<td>18</td>
<td>Server Process Monitoring</td>
<td>202</td>
</tr>
<tr>
<td>19</td>
<td>Big Mac Index</td>
<td>210</td>
</tr>
<tr>
<td>20</td>
<td>Complaints Dashboard</td>
<td>224</td>
</tr>
<tr>
<td>21</td>
<td>Hospital Operating Room Utilization</td>
<td>236</td>
</tr>
<tr>
<td>22</td>
<td>Showing Rank and Magnitude</td>
<td>246</td>
</tr>
<tr>
<td>23</td>
<td>Measuring Claims across Multiple Measures and Dimensions</td>
<td>258</td>
</tr>
<tr>
<td>24</td>
<td>Showing Churn or Turnover</td>
<td>268</td>
</tr>
<tr>
<td>25</td>
<td>Showing Actual versus Potential Utilization</td>
<td>282</td>
</tr>
<tr>
<td>26</td>
<td>Health Care Provider Productivity Monitoring</td>
<td>294</td>
</tr>
<tr>
<td>27</td>
<td>Telecom Operator Executive Dashboard</td>
<td>306</td>
</tr>
<tr>
<td>28</td>
<td>Economy at a Glance</td>
<td>316</td>
</tr>
<tr>
<td>29</td>
<td>Call Center</td>
<td>328</td>
</tr>
<tr>
<td></td>
<td><strong>PART III</strong> SUCCEEDING IN THE REAL WORLD</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Want to Engage People? Make Your Dashboards Personal</td>
<td>338</td>
</tr>
<tr>
<td>31</td>
<td>Visualizing Time</td>
<td>352</td>
</tr>
<tr>
<td>32</td>
<td>Beware the Dead-End Dashboard</td>
<td>382</td>
</tr>
<tr>
<td>33</td>
<td>The Allure of Red and Green</td>
<td>390</td>
</tr>
<tr>
<td>34</td>
<td>The Allure of Pies and Donuts</td>
<td>396</td>
</tr>
<tr>
<td>35</td>
<td>Clouds and Bubbles</td>
<td>404</td>
</tr>
<tr>
<td>36</td>
<td>A Journey into the Unknown</td>
<td>410</td>
</tr>
<tr>
<td></td>
<td><strong>Glossary</strong></td>
<td>419</td>
</tr>
<tr>
<td></td>
<td><strong>Bibliography</strong></td>
<td>423</td>
</tr>
<tr>
<td></td>
<td><strong>Index</strong></td>
<td>425</td>
</tr>
</tbody>
</table>
Acknowledgments

From the three of us

Stephen Few, whose books have made a profound and lasting impression on us.

Alberto Cairo for his invaluable feedback and for his leadership in the data visualization community.

Our technical reviewers greatly improved our first drafts. Thanks to Troy Magennis, Andy Kirk, Jon Schwabish, Ariel Pohoryles, Trudy Weiss Craig, Michael Fry, Andy Kriebel, and a special thanks to Cole Nussbaumer Knaflic for introducing us to the Wiley team and who went far beyond our expectations with her detailed edits and comments.

All the contributors to this book gave significant time to tweak their dashboards according to our requests. We thank you for allowing us to include your work in the book.

Thanks, also, to Mark Boone, KK Molugu, Eric Duell, Chris DeMartini, and Bob Filbin for their efforts.

Our stellar team at Wiley: acquisitions editor Bill Falloon for fighting so hard on our behalf; editor Christina Verigan for her deft reworking and invaluable help optimizing flow; senior production editor Samantha Hartley for overseeing the daunting process of making this book a beautiful, tangible thing; copy editor Debra Manette for such detailed editing and insights; proofreader Hope Breeman for her meticulous proof check; the team at WordCo for a comprehensive index and marketing manager Heather Dunphy for her exceptional expertise in connecting author with audience.

From Steve

My wife, Laura, and my daughters, Janine and Diana, for the never-ending support and love.

Ira Handler and Brad Epstein, whose friendship, encouragement, and example have been a godsend for the past dozen years.

Joe Mako, who has always been willing to help me with “the difficult stuff” and provided much needed encouragement when I was starting out.

The Princeton University Triangle Club, where I learned how to bring talented people together to make wonderful things. Without my experiences there I don’t know if I would have had the insight and ability to recruit my fellow authors.

Jeff and Andy, who not only made the book way better than it would have been had I tackled it on
my own, but for providing me with one of the most rewarding and enriching experiences of my career. Your abilities, candor, humor, grit, patience, impatience, thoughtfulness, and leadership made for a remarkable ride.

**From Andy**

I would like to thank Steve and Jeff for approaching me to join this project. I’d been procrastinating on writing a book for many years, and the opportunity to work with two passionate, skilled leaders was the trigger I needed to get going. I would like to thank them both for many hours of constructive debate (argument?) over the rights and wrongs of all aspects of dashboards and data visualization. It has been an enriching experience.

Finally, to Liz, my wife, and my daughters, Beatrice and Lucy. Thank you for your support and the freedom to abandon you all on weekends, mornings, and evenings in order to compete this project. I could not have done it without you.

**From Jeff**

Thank you, Steve and Andy. It was a pleasure working with you guys. I will miss the collaboration, especially our many hours of discussion about data visualization and dashboard design.

A special thank you to Mary, my wife, and to Nina and Elle, my twin daughters, for sacrificing lots of family time over many long nights and weekends. I would not have been able to complete this project without your support.
About the Authors

Steve Wexler has worked with ADP, Gallup, Deloitte, Convergys, Consumer Reports, The Economist, ConEd, D&B, Marist, Tradeweb, Tiffany, McKinsey & Company, and many other organizations to help them understand and visualize their data. Steve is a Tableau Zen Master, Iron Viz Champion, and Tableau Training Partner.

His presentations and training classes combine an extraordinary level of product mastery with the real-world experience gained through developing thousands of visualizations for dozens of clients. In addition to his recognized expertise in data visualization and Tableau, Steve has decades of experience as a successful instructor in all areas of computer-based technology. Steve has taught thousands of people in both large and small organizations and is known for conducting his seminars with clarity, patience, and humor.

Website: DataRevelations.com

Jeffrey A. Shaffer is Vice President of Information Technology and Analytics at Recovery Decision Science and Unifund. He is also Adjunct Professor at the University of Cincinnati, where he teaches Data Visualization and was named the 2016 Outstanding Adjunct Professor of the Year.

He is a regular speaker on the topic of data visualization, data mining, and Tableau training at conferences, symposiums, workshops, universities, and corporate training programs. He is a Tableau Zen Master, and was the winner of the 2014 Tableau Quantified Self Visualization Contest, which led him to compete in the 2014 Tableau Iron Viz Contest. His data visualization blog was on the shortlist for the 2016 Kantar Information is Beautiful Awards for Data Visualization Websites.

Website: DataPlusScience.com

Andy Cotgreave is Technical Evangelist at Tableau Software. He has over 10 years’ experience in data visualization and business intelligence, first honing his skills as an analyst at the University of Oxford. Since joining Tableau in 2011, he has helped and inspired thousands of people with technical advice and ideas on how to build a data-driven culture in a business.

In 2016 he ran the MakeoverMonday (http://www.makeovermonday.co.uk/) project with Andy Kriebel, a social data project which saw over 500 people make 3,000 visualizations in one year. The project received an honourable mention in the Dataviz
Project category of the 2016 Kantar Information is Beautiful Awards.

Andy has spoken at conferences around the world, including SXSW, Visualized, and Tableau’s customer conferences. He writes a column for Computerworld, Living with Data (http://www.computerworld.com/blog/living-data/), as well as maintaining his own blog, GravyAnecdote.com.

Website: GravyAnecdote.com
Introduction

We wrote *The Big Book of Dashboards* for anyone tasked with building or overseeing the development of business dashboards. Over the past decade, countless people have approached us after training sessions, seminars, or consultations, shown us their data, and asked: “What would be a really good way to show this?”

These people faced a specific business predicament (what we call a “scenario”) and wanted guidance on how to best address it with a dashboard. In reviewing dozens of books about data visualization, we were surprised that, while they contained wonderful examples showing why a line chart often works best for time-series data and why a bar chart is almost always better than a pie chart, none of them matched great dashboards with real-world business cases. After pooling our experience and enormous collection of dashboards, we decided to write our own book.

**How This Book Is Different**

This book is not about the fundamentals of data visualization. That has been done in depth by many amazing authors. We want to focus on proven, real-world examples and why they succeed.

However, if this is your first book about the topic of data visualization, we do provide a primer in Part I with everything you need to know to understand how the charts in the scenarios work. We also dearly hope it whets your appetite for more, which is why this section finishes with our recommended further reading.

**How This Book Is Organized**

The book is organized into three parts.

**Part I: A Strong Foundation.** This part covers the fundamentals of data visualization and provides our crash course on the foundational elements that give you the vocabulary you need to explore and understand the scenarios.

**Part II: The Scenarios.** This is the heart of the book, where we describe dozens of different business scenarios and then present a dashboard that “solves” the challenges presented in those scenarios.

**Part III: Succeeding in the Real World.** The chapters in this part address problems we’ve encountered and anticipate you may encounter as well. With these chapters—distilled from decades of real-world experience—we hope to make your journey quite a bit easier and a lot more enjoyable.
How to Use This Book

We encourage you to look through the book to find a scenario that most closely matches what you are tasked with visualizing. Although there might not be an exact match, our goal is to present enough scenarios that you can find something that will address your needs. The internal conversation in your head might go like this:

“Although my data isn’t exactly the same as what’s in this scenario, it’s close enough, and this dashboard really does a great job of helping me and others see and understand that data. I think we should use this approach for our project as well.”

For each scenario we present the entire dashboard at the beginning of the chapter, then explore how individual components work and contribute to the whole.

By organizing the book based on these scenarios and offering practical and effective visualization examples, we hope to make *The Big Book of Dashboards* a trusted resource that you open when you need to build an effective business dashboard. To ensure you get the most out of these examples, we have included a visual glossary at the back of this book. If you come across an unfamiliar term, such as “sparkline,” you can look it up and see an illustration.

We also encourage you to spend time with all the scenarios and the proposed solutions as there may be some elements of a seemingly irrelevant scenario that may apply to your own needs.

For example, Chapter 11 shows a dashboard used by a team in the English Premier League to help players understand their performance. Your data might have nothing to do with sports, but the dashboard is a great example of showing current and historical performance. (See Figure I.1.) That might be something you have to do with your data. Plus, if you skip one scenario, you might miss a great example of the exact chart you need for your own solution.

We also encourage you to browse the book for motivation. Although a scenario may not be a perfect match, the thought process and chart choices may inspire you.

Succeeding in the Real World

In addition to the scenarios, an entire section of the book is devoted to addressing many practical and psychological factors you will encounter in your work. It’s great to have theory- and evidence-based research at your disposal, but what will you do when somebody asks you to make your dashboard “cooler” by adding packed bubbles and donut charts?

The three of us have a combined 30-plus years of hands-on experience helping people in hundreds of organizations build effective visualizations. We have fought (and sometimes lost) many “best practices” battles. But by having endured these struggles, we bring an uncommon empathy to the readers of this book.

We recognize that at times readers will be asked to create dashboards and charts that exemplify bad practice. For example, a client or a department head may stipulate using a particular combination of colors or demand a chart type that is against evidence-based data visualization best practices.

We hear you. We’ve been there.
Although the dashboard in Figure I.1 pertains to sports, the techniques are universal. Here the latest event is in yellow, the five most recent events are in red, and older events are in a muted gray. Brilliant.

**FIGURE I.1** A player summary from an English Premier League Club

(Note: Fake data is used.)
We’ve faced many of the hurdles you will encounter and the concepts you will grapple with in your attempt to build dashboards that are informative, enlightening, and engaging. The essays in this section will help smooth the way for you by offering suggestions and alternatives for these issues.

What to Do and What Not to Do

Although the book is an attempt to celebrate good examples, we’ll also show plenty of bad examples. We guarantee you will see this kind of work out in the wild, and you may even be asked to emulate it. We mark these “bad” examples with the cat icon shown in Figure 1.2 so that you don’t have to read the surrounding text to determine if the chart is something you should emulate or something you should avoid.

![Cat Icon](image)

**FIGURE 1.2** If you see this icon, it means don’t make a chart like this one.

Illustration by Eric Kim

What Is a Dashboard?

Ask 10 people who build business dashboards to define a dashboard and you will probably get 10 different definitions. For the purpose of this book, our definition is as follows:

A dashboard is a visual display of data used to monitor conditions and/or facilitate understanding.

This is a broad definition, and it means that we would consider all of the examples listed below to be dashboards:

- An interactive display that allows people to explore worker compensation claims by region, industry, and body part
- A PDF showing key measures that gets e-mailed to an executive every Monday morning
- A large wall-mounted screen that shows support center statistics in real time
- A mobile application that allows sales managers to review performance across different regions and compare year-to-date sales for the current year with the previous year

Even if you don’t consider every example in this book a true dashboard, we think you will find the discussion and analysis around each of the scenarios helpful in building your solutions. Indeed, we can debate the definition until we are blue in the face, but that would be a horrible waste of effort as it simply isn’t that important. What is important—make that essential—is understanding how to combine different elements (e.g., charts, text, legends, filters, etc.) into a cohesive and coordinated whole that allows people to see and understand their data.
Final Thought: There Are No Perfect Dashboards

You will not find any perfect dashboards in this book.

In our opinion, there is no such thing as a perfect dashboard. You will never find one perfect collection of charts that ideally suits every person who may encounter it. But, although they may not be perfect, the dashboards we showcase in the book successfully help people see and understand data in the real world.

The dashboards we chose all have this in common: Each one demonstrates some great ideas in a way that is relevant to the people who need to understand them. In short, they all serve the end users. Would we change some of the dashboards? Of course we would, and we weigh in on what we would change in the author commentary at the end of each scenario. Sometimes we think a chart choice isn’t ideal; other times, the layout isn’t quite right; and in some cases, the interactivity is clunky or difficult. What we recognize is that every set of eyes on a dashboard will judge the work differently, which is something you also should keep in mind. Where you see perfection, others might see room for improvement. The challenge all the dashboard designers in this book have faced is balancing a dashboard’s presentation and objectives with time and efficiency. It’s not an easy spot to hit, but with this book we hope to make it easier for you.

Steve Wexler
Jeffrey Shaffer
Andy Cotgreave
PART I

A STRONG FOUNDATION
Data Visualization: A Primer
This book is about real-world dashboards and why they succeed. In many of the scenarios, we explain how the designers use visualization techniques to contribute to that success. For those new to the field, this chapter is a primer on data visualization. It provides enough information for you to understand why we picked many of the dashboards. If you are more experienced, this chapter recaps data visualization fundamentals.

Why Do We Visualize Data?

Let’s see why it’s vital to visualize numbers by beginning with Table 1.1. There are four groups of numbers, each with 11 pairs. In a moment, we will create a chart from them, but before we do, take a look at the numbers. What can you see? Are there any discernible differences in the patterns or trends among them?

Let me guess: You don’t really see anything clearly. It’s too hard.

Before we put the numbers in a chart, we might consider their statistical properties. Were we to do that, we’d find that the statistical properties of each group of numbers are very similar. If the table doesn’t show anything and statistics don’t reveal much, what happens when we plot the numbers? Take a look at Figure 1.1.

Now do you see the differences? Seeing the numbers in a chart shows you something that tables and some statistical measures cannot. We visualize data to harness the incredible power of our visual system to spot relationships and trends.

This brilliant example is the creation of Frank Anscombe, a British statistician. He created this set

<table>
<thead>
<tr>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
<th>Group D</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>y</td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>10.00</td>
<td>8.04</td>
<td>10.00</td>
<td>9.14</td>
</tr>
<tr>
<td>8.00</td>
<td>6.95</td>
<td>8.00</td>
<td>8.14</td>
</tr>
<tr>
<td>13.00</td>
<td>7.58</td>
<td>13.00</td>
<td>8.74</td>
</tr>
<tr>
<td>9.00</td>
<td>8.81</td>
<td>9.00</td>
<td>8.77</td>
</tr>
<tr>
<td>11.00</td>
<td>8.33</td>
<td>11.00</td>
<td>9.26</td>
</tr>
<tr>
<td>14.00</td>
<td>9.96</td>
<td>14.00</td>
<td>8.10</td>
</tr>
<tr>
<td>6.00</td>
<td>7.24</td>
<td>6.00</td>
<td>6.13</td>
</tr>
<tr>
<td>4.00</td>
<td>4.26</td>
<td>4.00</td>
<td>3.10</td>
</tr>
<tr>
<td>12.00</td>
<td>10.84</td>
<td>12.00</td>
<td>9.13</td>
</tr>
<tr>
<td>7.00</td>
<td>4.82</td>
<td>7.00</td>
<td>7.26</td>
</tr>
<tr>
<td>5.00</td>
<td>5.68</td>
<td>5.00</td>
<td>4.74</td>
</tr>
</tbody>
</table>
Now can you see a difference in the four groups?

Identifying trends is as hard as it was with Anscombe’s Quartet. To read the table, we need to look up every value, one at a time. Unfortunately, our short-term memories aren’t designed to store many pieces of information. By the time we’ve reached the fourth or fifth number, we will have forgotten the first one we looked at.

Let’s try a trend line, as shown in Figure 1.2.
How Do We Visualize Data?

Now we have much better insight into the trends. Office supplies has been the lowest-selling product category in all but two quarters. Furniture trends have been dropping slowly over the time period, except for a bump in sales in 2015 Q4 and a rise in the last two quarters. Technology sales have mostly been the highest but were particularly volatile at the start of the time period.

The table and the line chart each visualized the same 48 data points, but only the line chart lets us see the trends. The line chart turned 48 data points into three chunks of data, each containing 16 data points. Visualizing the data hacks our short-term memory; it allows us to interpret large volumes of data instantly.

How Do We Visualize Data?

We’ve just looked at some examples of the power of visualizing data. Now we need to move on to how we build the visualizations. To do that, we first need to look at two things: preattentive attributes and types of data.
Preattentive Attributes

Visualizing data requires us to turn data into marks on a canvas. What kind of marks make the most sense? One answer lies in what are called “preattentive attributes.” These are things that our brain processes in milliseconds, before we pay attention to everything else. There are many different types. Let’s look at an example.

Look at the numbers in Figure 1.3. How many 9s are there?

How did you do? It’s easy to answer the question—you just look at all the values and count the 9s—but it takes a long time. We can make one change to the grid and make it very easy for you. Have a look at Figure 1.4.

Now the task is easy. Why? Because we changed the color: 9s are red, and all the other numbers are light gray.

Color differences pop out. It’s as easy to find one red 9 on a table of hundreds of digits as it is on a 10-by-10 grid. Think about that for a moment: Your brain registers the red 9s before you consciously addressed the grid to count them. Check out the grid of 2,500 numbers in Figure 1.5. Can you see the 9?

It’s easy to spot the 9. Our eyes are amazing at spotting things like this.
FIGURE 1.5  There is a single 9 in this grid of 2,500 numbers. We wager you saw it before you started reading any other numbers on this page.
Figure 1.6 Differences in size are easy to see too.

Figure 1.7 Coloring every digit is nearly as bad as having no color.

Color (in this case, hue) is one of several preattentive attributes. When we look at a scene in front of us, or a chart, we process these attributes in under 250 milliseconds. Let’s try out a couple more preattentive features with our table of 9s. In Figure 1.6, we’ve made the 9s a different size from the rest of the figures.

Size and hue: Aren’t they amazing? That’s all very well when counting the 9s. What if our task is to count the frequency of each digit? That’s a slightly more realistic task, but we can’t just use a different color or size for each digit. That would defeat the preattentive nature of the single color. Look at the mess that is Figure 1.7.

It’s not a complete disaster: If you’re looking for the 6s, you just need to work out that they are red and then scan quickly for those. Using one color on a visualization is highly effective to make one category stand out. Using a few colors, as we did in Figure 1.2 to distinguish a small number of categories, is fine too. Once you’re up to around eight to ten categories, however, there are too many colors to easily distinguish one from another.

To count each digit, we need to aggregate. Visualization is, at its core, about encoding aggregations, such as frequency, in order to gain insight. We need to move away from the table entirely and encode the frequency of each digit. The most effective way is to use length, which we can do in a bar chart. Figure 1.8 shows the frequency of each digit. We’ve also colored the bar showing the number 9.

Since the task is to count the 9s in the data source, the bar chart is one of the best ways to see the results. This is because length and position are best for quantitative comparisons. If we extend the example one final time and consider which numbers are most common, we could sort the bars, as shown in Figure 1.9.
How Do We Visualize Data?

We also know that 9 was the third most common digit in the table. We can also see the frequency of every other digit.

The series of examples we just presented used color, size, and length to highlight the 9s. These are three of many preattentive attributes. Figure 1.10 shows 12 that are commonly used in data visualization.

Some of them will be familiar to you from charts you have already seen. Anscombe’s Quartet (see Figure 1.1) used position and spatial grouping. The x- and y-coordinates are for position, while spatial grouping allows us to see the outliers and the patterns.

Preattentive attributes provide us with ways to encode our data in charts. We’ll look into that in more detail in a moment, but not before we’ve talked about data.

To recap, we’ve seen how powerful the visual system is and looked at some visual features we can use to display data effectively. Now we need to look at the different types of data, in order to choose the best visual encoding for each type.

FIGURE 1.8 There are 13 9s.

FIGURE 1.9 Sorted bar chart using color and length to show how many 9s are in our table.

This series of examples with the 9s reemphasizes the importance of visualizing data. As with Anscombe’s Quartet, we went from a difficult-to-read table of numbers to an easy-to-read bar chart. In the sorted bar chart, not only can we count the 9s (the original task), but we also know that 9 was the third most common digit in the table. We can also see the frequency of every other digit.

FIGURE 1.10 Preattentive features.
Types of Data

There are three types of data: categorical, ordinal, and quantitative. Let’s use a photo to help us define each type.

**Categorical Data**

Categorical (or nominal) data represents things. These things are mutually exclusive labels without any numerical value. What nominal data can we use to describe the gentleman with me in the Figure 1.11?

- His name is Brent Spiner.
- By profession he is an actor.
- He played the character Data in the TV show *Star Trek: The Next Generation*.
- Brent Spiner’s date of birth is Wednesday, February 2, 1949.
- He appeared in all seven seasons of *Star Trek: The Next Generation*.
- Data’s rank was lieutenant commander.
- Data was the fifth of six androids made by Dr. Noonien Soong.

Other types of ordinal data include education experience, satisfaction level, and salary bands in an organization. Although ordinal values often have numbers associated with them, the interval between those values is arbitrary. For example, the difference in an organization between pay scales 1 and 2 might be very different from that between pay scales 4 and 6.

**Quantitative Data**

Quantitative data is the numbers. Quantitative (or numerical) data is data that can be measured and aggregated.

- Brent Spiner’s date of birth is Wednesday, February 2, 1949.
- His height is 5 ft 9 in (180 cm) tall.
- He made 177 appearances in episodes of *Star Trek*.
- Data’s positronic brain is capable of 60 trillion operations per second.

You’ll have noticed that date of birth appears in both ordinal and quantitative data types. Time is unusual in that it can be both. In Chapter 31, we look in detail about how you treat time influences your choice of visualization types.

Other types of quantitative measures include sales, profit, exam scores, pageviews, and number of patients in a hospital.
Quantitative data can be expressed in two ways: as discrete or continuous data. Discrete data is presented at predefined, exact points—there’s no “in between.” For example, Brent Spiner appeared in 177 episodes of Star Trek; he couldn’t have appeared in 177.5 episodes. Continuous data allows for the “in between,” as there is an infinite number of possible intermediate values. For example, Brent Spiner grew to a height of 5 ft 9 in but at one point in his life he was 4 ft 7.5 in tall.

Encoding Data in Charts

We’ve now looked at preattentive attributes and the three types of data. It’s time to see how to combine that knowledge into building charts. Let’s look at some charts and see how they encode the different types of data. Sticking with Star Trek, Figure 1.12 shows the IMDB.com ratings of every episode of Star Trek: The Next Generation. Table 1.3 shows the different types of data, what type it is, and how it’s been encoded.

Table 1.3 Data used in Figure 1.12.

<table>
<thead>
<tr>
<th>Data</th>
<th>Data Type</th>
<th>Encoding</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Episode</td>
<td>Categorical</td>
<td>Position</td>
<td>Each episode is represented by a dot. Each dot has its own position on the canvas.</td>
</tr>
<tr>
<td>Episode Number</td>
<td>Ordinal</td>
<td>Position</td>
<td>The x-axis shows the number of each episode in each season.</td>
</tr>
<tr>
<td>Season</td>
<td>Ordinal</td>
<td>Color</td>
<td>Each season is represented by a different color (hue).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Position</td>
<td>Each season also has its own section on the chart.</td>
</tr>
<tr>
<td>IMDB rating</td>
<td>Ordinal</td>
<td>Position</td>
<td>The better the episode, the higher it is on the y-axis.</td>
</tr>
<tr>
<td>Average season</td>
<td>Ordinal</td>
<td>Position</td>
<td>The horizontal bar in each pane shows the average rating of the episodes in each season. There is some controversy over whether you should average ordinal ratings. We believe that the practice is so common with ratings it is acceptable.</td>
</tr>
<tr>
<td>rating</td>
<td>Quantitative</td>
<td>Position</td>
<td></td>
</tr>
</tbody>
</table>
Let’s look at a few more charts to see how preattentive features have been used. Figure 1.13 is from *The Economist*. Look at each chart and see if you can work out which types of data are being graphed and how they are being encoded.

Table 1.4 shows how each data type is encoded.

**Figure 1.13** “A terrible record” from *The Economist*, July 2016.  
Source: START, University of Maryland. *The Economist*, http://tabsoft.co/2agK3if

**Table 1.4** Data used in the bar chart in Figure 1.13.

<table>
<thead>
<tr>
<th>Data</th>
<th>Data Type</th>
<th>Encoding</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>Categorical</td>
<td>Position</td>
<td>Each country is on its own row (sorted by total deaths).</td>
</tr>
<tr>
<td>Deaths</td>
<td>Quantitative</td>
<td>Length</td>
<td>The length of the bar shows the number of deaths.</td>
</tr>
<tr>
<td>Death type</td>
<td>Categorical</td>
<td>Color</td>
<td>Dark blue shows deaths of victims, light blue shows deaths of the perpetrators.</td>
</tr>
<tr>
<td>Attacks</td>
<td>Quantitative</td>
<td>Size</td>
<td>Circles on the right are sized according to the number of attacks.</td>
</tr>
</tbody>
</table>