

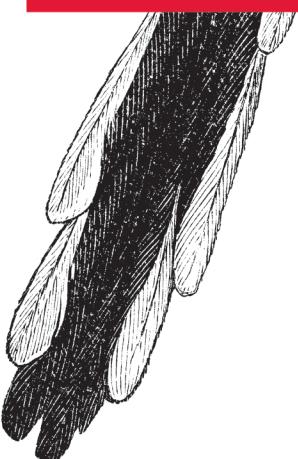
An Introduction to Designing with D3



Free Sampler

Interactive Data Visualization

for the Web



O'REILLY®

Scott Murray

O'Reilly Ebooks—Your bookshelf on your devices!



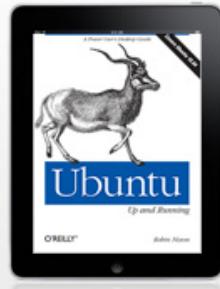
PDF



ePub



Mobi



APK



DAISY

When you buy an ebook through oreilly.com you get lifetime access to the book, and whenever possible we provide it to you in five, DRM-free file formats—PDF, .epub, Kindle-compatible .mobi, Android .apk, and DAISY—that you can use on the devices of your choice. Our ebook files are fully searchable, and you can cut-and-paste and print them. We also alert you when we've updated the files with corrections and additions.

Learn more at ebooks.oreilly.com

You can also purchase O'Reilly ebooks through the iBookstore, the [Android Marketplace](http://Android.Marketplace), and Amazon.com.

O'REILLY®

Spreading the knowledge of innovators

oreilly.com

Interactive Data Visualization for the Web

by Scott Murray

Copyright © 2013 Scott Murray. All rights reserved.

Printed in Canada.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://my.safaribooksonline.com>). For more information, contact our corporate/institutional sales department: 800-998-9938 or corporate@oreilly.com.

Editor: Meghan Blanchette

Indexer: Judith McConville

Production Editor: Melanie Yarbrough

Cover Designer: Karen Montgomery

Copyeditor: Teresa Horton

Interior Designer: David Futato

Proofreader: Linley Dolby

Illustrator: Rebecca Demarest

March 2013: First Edition

Revision History for the First Edition:

2013-03-04: First release

See <http://oreilly.com/catalog/errata.csp?isbn=9781449339739> for release details.

Nutshell Handbook, the Nutshell Handbook logo, the cover image, and the O'Reilly logo are registered trademarks of O'Reilly Media, Inc. *Interactive Data Visualization for the Web*, the cover image of a long-tail bushtit, and related trade dress are trademarks of O'Reilly Media, Inc.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and O'Reilly Media, Inc., was aware of a trademark claim, the designations have been printed in caps or initial caps.

While every precaution has been taken in the preparation of this book, the publisher and author assume no responsibility for errors or omissions, or for damages resulting from the use of the information contained herein.

ISBN: 978-1-449-33973-9

TI

Table of Contents

Preface.....	ix
1. Introduction.....	1
Why Data Visualization?	1
Why Write Code?	2
Why Interactive?	2
Why on the Web?	3
What This Book Is	3
Who You Are	4
What This Book Is Not	5
Using Sample Code	5
Thank You	6
2. Introducing D3.....	7
What It Does	7
What It Doesn't Do	8
Origins and Context	9
Alternatives	10
Easy Charts	10
Graph Visualizations	12
Geomapping	12
Almost from Scratch	13
Three-Dimensional	13
Tools Built with D3	14
3. Technology Fundamentals.....	15
The Web	15
HTML	17
Content Plus Structure	18

Adding Structure with Elements	19
Common Elements	20
Attributes	22
Classes and IDs	22
Comments	23
DOM	24
Developer Tools	24
Rendering and the Box Model	27
CSS	29
Selectors	29
Properties and Values	31
Comments	31
Referencing Styles	31
Inheritance, Cascading, and Specificity	33
JavaScript	35
Hello, Console	35
Variables	35
Other Variable Types	36
Arrays	36
Objects	37
Objects and Arrays	38
Mathematical Operators	40
Comparison Operators	41
Control Structures	41
Functions	43
Comments	44
Referencing Scripts	44
JavaScript Gotchas	45
SVG	49
The SVG Element	50
Simple Shapes	50
Styling SVG Elements	53
Layering and Drawing Order	54
Transparency	55
A Note on Compatibility	57
4. Setup.....	59
Downloading D3	59
Referencing D3	60
Setting Up a Web Server	61
Terminal with Python	61
MAMP, WAMP, and LAMP	62

Diving In	62
5. Data.....	63
Generating Page Elements	63
Chaining Methods	65
One Link at a Time	66
The Hand-off	67
Going Chainless	67
Binding Data	67
In a Bind	67
Data	68
Please Make Your Selection	72
Bound and Determined	73
Using Your Data	76
High-Functioning	77
Data Wants to Be Held	78
Beyond Text	79
6. Drawing with Data.....	81
Drawing divs	81
Setting Attributes	82
A Note on Classes	83
Back to the Bars	83
Setting Styles	84
The Power of data()	85
Random Data	87
Drawing SVGs	89
Create the SVG	89
Data-Driven Shapes	90
Pretty Colors, Oooh!	92
Making a Bar Chart	92
The Old Chart	93
The New Chart	93
Color	98
Labels	101
Making a Scatterplot	103
The Data	103
The Scatterplot	104
Size	105
Labels	106

Next Steps	107
7. Scales.....	109
Apples and Pixels	109
Domains and Ranges	110
Normalization	111
Creating a Scale	111
Scaling the Scatterplot	112
d3.min() and d3.max()	112
Setting Up Dynamic Scales	114
Incorporating Scaled Values	114
Refining the Plot	115
Other Methods	119
Other Scales	119
8. Axes.....	121
Introducing Axes	121
Setting Up an Axis	122
Cleaning It Up	123
Check for Ticks	126
Y Not?	127
Final Touches	128
Formatting Tick Labels	130
9. Updates, Transitions, and Motion.....	133
Modernizing the Bar Chart	133
Ordinal Scales, Explained	134
Round Bands Are All the Range These Days	136
Referencing the Ordinal Scale	136
Other Updates	137
Updating Data	137
Interaction via Event Listeners	138
Changing the Data	139
Updating the Visuals	139
Transitions	142
duration(), or How Long Is This Going to Take?	143
ease()-y Does It	144
Please Do Not delay()	145
Randomizing the Data	147
Updating Scales	150
Updating Axes	152
each() Transition Starts and Ends	153

Other Kinds of Data Updates	161
Adding Values (and Elements)	161
Removing Values (and Elements)	166
Data Joins with Keys	169
Add and Remove: Combo Platter	174
Recap	176
10. Interactivity.....	177
Binding Event Listeners	177
Introducing Behaviors	178
Hover to Highlight	179
Grouping SVG Elements	184
Click to Sort	185
Tooltips	188
Default Browser Tooltips	189
SVG Element Tooltips	190
HTML div Tooltips	192
Consideration for Touch Devices	195
Moving Forward	195
11. Layouts.....	197
Pie Layout	198
Stack Layout	202
Force Layout	206
12. Geomapping.....	213
JSON, Meet GeoJSON	213
Paths	215
Projections	216
Choropleth	219
Adding Points	222
Acquiring and Parsing Geodata	226
Find Shapefiles	226
Choose a Resolution	226
Simplify the Shapes	228
Convert to GeoJSON	228
13. Exporting.....	233
Bitmaps	233
PDF	234

SVG	235
A. Appendix: Further Study.....	241
Index.....	245

CHAPTER 1

Introduction

Why Data Visualization?

Our information age more often feels like an era of information overload. Excess amounts of information are overwhelming; raw data becomes useful only when we apply methods of deriving insight from it.

Fortunately, we humans are intensely visual creatures. Few of us can detect patterns among rows of numbers, but even young children can interpret bar charts, extracting meaning from those numbers' visual representations. For that reason, data visualization is a powerful exercise. Visualizing data is the fastest way to communicate it to others.

Of course, visualizations, like words, can be used to lie, mislead, or distort the truth. But when practiced honestly and with care, the process of visualization can help us see the world in a new way, revealing unexpected patterns and trends in the otherwise hidden information around us. At its best, data visualization is expert storytelling.

More literally, visualization is a process of *mapping* information to visuals. We craft rules that interpret data and express its values as visual properties. For example, the humble bar chart in [Figure 1-1](#) is generated from a very simple rule: larger values are *mapped* as taller bars.

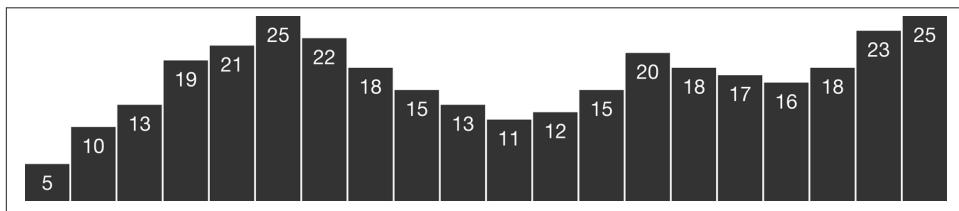


Figure 1-1. Data values mapped to visuals

More complex visualizations are generated from datasets more complex than the sequence of numbers shown in [Figure 1-1](#) and more complex sets of mapping rules.

Why Write Code?

Mapping data by hand can be satisfying, yet is slow and tedious. So we usually employ the power of computation to speed things up. The increased speed enables us to work with much larger datasets of thousands or millions of values; what would have taken years of effort by hand can be mapped in a moment. Just as important, we can rapidly experiment with *alternate mappings*, tweaking our rules and seeing their output re-rendered immediately. This loop of write/render/evaluate is critical to the iterative process of refining a design.

Sets of mapping rules function as *design systems*. The human hand no longer executes the visual output; the computer does. Our human role is to conceptualize, craft, and write out the rules of the system, which is then finally executed by software.

Unfortunately, software (and computation generally) is extremely bad at understanding what, exactly, people want. (To be fair, many humans are also not good at this challenging task.) Because computers are binary systems, everything is either on or off, yes or no, this or that, there or not there. Humans are mushier, softer creatures, and the computers are not willing to meet us halfway—we must go to them. Hence the inevitable struggle of learning to write software, in which we train ourselves to communicate in the very limited and precise syntax that the computer can understand.

Yet we continue to write code because seeing our visual creations come to life is so rewarding. We practice data visualization because it is exciting to see what has never before been seen. It is like summoning a magical, visual genie out of an inscrutable data bottle.

Why Interactive?

Static visualizations can offer only precomposed “views” of data, so multiple static views are often needed to present a variety of perspectives on the same information. The number of dimensions of data are limited, too, when all visual elements must be present on the same surface at the same time. Representing multidimensional datasets fairly in static images is notoriously difficult. A fixed image is ideal when alternate views are neither needed nor desired, and required when publishing to a static medium, such as print.

Dynamic, interactive visualizations can empower people to explore the data for themselves. The basic functions of most interactive visualization tools have changed little since 1996, when Ben Shneiderman of the University of Maryland first proposed a

“Visual Information-Seeking Mantra”: *overview first, zoom and filter, then details-on-demand.*

This design pattern is found in most interactive visualizations today. The combination of functions is successful, because it makes the data accessible to different audiences, from those who are merely browsing or exploring the dataset to those who approach the visualization with a specific question in search of an answer. An interactive visualization that offers an overview of the data alongside tools for “drilling down” into the details may successfully fulfill many roles at once, addressing the different concerns of different audiences, from those new to the subject matter to those already deeply familiar with the data.

Of course, interactivity can also encourage engagement with the data in ways that static images cannot. With animated transitions and well-crafted interfaces, some visualizations can make exploring data feel more like playing a game. Interactive visualization can be a great medium for engaging an audience who might not otherwise care about the topic or data at hand.

Why on the Web?

Visualizations aren’t truly visual unless they are *seen*. Getting your work out there for others to see is critical, and publishing on the Web is the quickest way to reach a global audience. Working with web-standard technologies means that your work can be seen and experienced by anyone using a recent web browser, regardless of the operating system (Windows, Mac, Linux) and device type (laptop, desktop, smartphone, tablet).

Best of all, everything covered in this book can be done with freely accessible tools, so the only investment required is your time. And everything we’ll talk about uses open source, web-standard technologies.

By avoiding proprietary software and plug-ins, you can ensure that your projects are accessible on the widest possible range of devices, from typical desktop computers to tablets and even phones. The more accessible your visualization, the greater your audience and your impact.

What This Book Is

This book is a practical introduction to merging three practices—data visualization, interactive design, and web development—using D3, a powerful tool for custom, web-based visualization.

These chapters grew out of my own process of learning how to use D3. Many people, including myself, come to D3 with backgrounds in design, mapping, and data visualization, but not programming and computer science.

D3 has a bit of an unfair reputation for being hard to learn. D3 itself is not so complicated, but it operates in the domain of the Web, and the Web *is* complicated. Using D3 comfortably requires some prior knowledge of the web technologies with which it interacts, such as HTML, CSS, JavaScript, and SVG. Many people (myself included) are self-taught when it comes to web skills. This is great, because the barrier to entry is so low, but problematic because it means we probably didn't learn each of these technologies from the ground up—more often, we just hack something together until it seems to work, and call it a day. Yet successful use of D3 requires understanding some of these technologies in a fundamental way.

Because D3 is written in JavaScript, learning to use D3 often means learning a lot about JavaScript. For many datavis folks, D3 *is* their introduction to JavaScript (or even web development generally). It's hard enough to learn a new programming language, let alone a new tool built on that language. D3 will enable you to do great things with JavaScript that you never would have even attempted. The time you spend learning both the language and the tool will provide an incredible payoff.

My goal is to reduce that learning time, so you can start creating awesome stuff sooner. We'll take a ground-up approach, starting with the fundamental concepts and gradually adding complexity. I don't intend to show you how to make specific kinds of visualizations so much as to help you understand the workings of D3 well enough to take those building blocks and generate designs of your own creation.

Who You Are

You may be an absolute beginner, someone new to datavis, web development, or both. (Welcome!) Perhaps you are a journalist interested in new ways to communicate the data you collect during reporting. Or maybe you're a designer, comfortable drawing static infographics but ready to make the leap to interactive projects on the Web. You could be an artist, interested in generative, data-based art. Or a programmer, already familiar with JavaScript and the Web, but excited to learn a new tool and pick up some visual design experience along the way.

Whoever you are, I hope that you:

- Have heard of this new thing called the “World Wide Web”
- Are a bit familiar with HTML, the DOM, and CSS
- Might even have a little programming experience already
- Have heard of jQuery or written some JavaScript before
- Aren't scared by unknown initialisms like CSV, SVG, or JSON
- Want to make useful, interactive visualizations

If any of those things are unknown or unclear, don't fear. You might just want to spend more time with [Chapter 3](#), which covers what you really need to know before diving into D3.

What This Book Is Not

That said, this is definitely not a computer science textbook, and it is not intended to teach the intricacies of any one web technology (HTML, CSS, JavaScript, SVG) in depth.

In that spirit, I might gloss over some technical points, grossly oversimplifying important concepts fundamental to computer science in ways that will make true software engineers recoil. That's fine, because I'm writing for artists and designers here, not engineers. We'll cover the basics, and then you can dive into the more complex pieces once you're comfortable.

I will deliberately *not* address every possible approach to a given problem, but will typically present what I feel is the simplest solution, or, if not the simplest, then the most understandable.

My goal is to teach you the fundamental concepts and methods of D3. As such, this book is decidedly *not* organized around specific example projects. Everyone's data and design needs will be different. It's up to you to integrate these concepts in the way best suited to your particular project.

Using Sample Code

If you are a mad genius, then you can probably learn to use D3 without ever looking at any sample code files, in which case you can skip the rest of this section.

If you're still with me, you are probably still very bright but not mad, in which case you should undertake this book with the full set of accompanying code samples in hand. Before you go any further, please download the sample files from [GitHub](#).

Normal people will want to click the ZIP link to download a compressed ZIP archive with all the files. Hardcore geeksters will want to clone the repository using Git. If that last sentence sounds like total gibberish, please use the first option.

Within the download, you'll notice there is a folder for each chapter that has code to go with it:

```
chapter_04  
chapter_05  
chapter_06  
chapter_07  
chapter_08  
...
```

Files are organized by chapter, so in [Chapter 9](#) when I reference *01_bar_chart.html*, know that you can find that file in the corresponding location: *d3-book/chapter_9/01_bar_chart.html*.

You are welcome to copy, adapt, modify, and reuse the example code in these tutorials for any noncommercial purpose.

Thank You

Finally, this book has been handcrafted, carefully written, and pedagogically fine-tuned for maximum effect. Thank you for reading it. I hope you learn a great deal, and even have some fun along the way.

Want to read more?

You can [buy this book](#) at [oreilly.com](#)
in print and ebook format.

Buy 2 books, get the 3rd FREE!

Use discount code: OPC10

All orders over \$29.95 qualify for **free shipping** within the US.

It's also available at your favorite book retailer,
including the iBookstore, the [Android Marketplace](#),
and [Amazon.com](#).



O'REILLY®

Spreading the knowledge of innovators

[oreilly.com](#)