APPENDIX 1

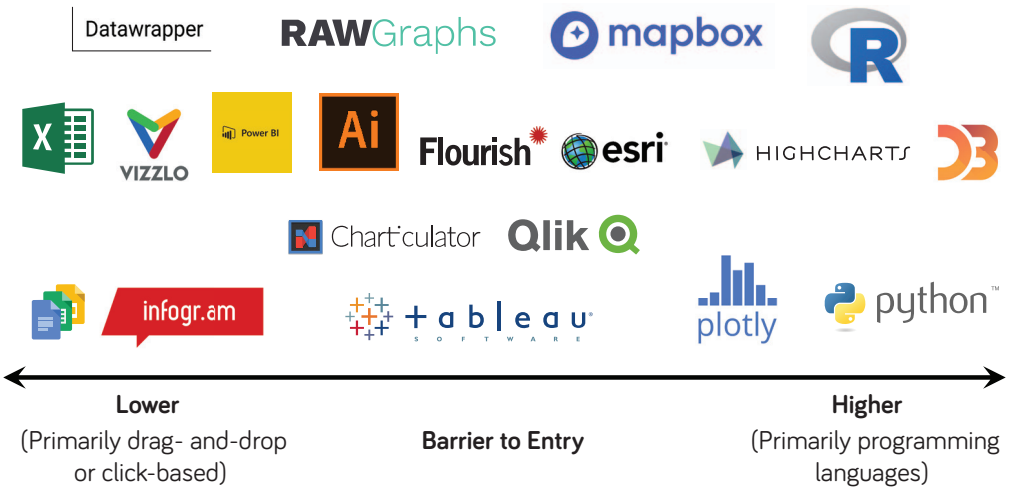
DATA VISUALIZATION TOOLS

This book is tool agnostic. My goal is not to show you how to create each of the eighty-plus graphs we explored. There are far too many tools you can use and far too many approaches within each tool. What matters is not which tool you use, but that it helps you create the graphs you need to best serve the needs of your audience.

There are many, many data visualization tools available, for use on different platforms and with different purchasing and subscription options. The number, types, and capabilities of these tools is constantly changing to reflect updated underlying technologies and coding languages. Which tool you use will depend on your personal preferences and skills, as well as those of your colleagues and support within your organization.

Data visualization tools live along a spectrum. On one end are click or drag-and-drop tools like Excel, which allow the user to click and insert a chart. On the other end are programming languages like R and JavaScript that require written code to create data visuals. The barrier to entry is much lower in Excel—virtually anyone can create a line chart or bar chart in seconds. It's much different on the other end, where programming languages require an understanding of how to write code and the different syntax across the different languages. Yet programming languages give you substantially more flexibility, while tools like Excel can box you in within a small subset of graphs.

How "difficult" these tools are depends on the person and sometimes the organization. You may have an affinity for computer programming languages, in which case JavaScript, Python, or R may be appropriate, but for a variety of reasons, your organization may not allow you to use these open source tools. Many social science researchers use



statistical languages like SAS, SPSS, and Stata, but, in my opinion, the graphing capabilities of these tools lag behind those of other languages. Some of the drag-and-drop tools like Excel and Tableau are easier to start with but designing more bespoke visualizations is either impossible or may require some coding (for example, Calculated Fields in Tableau).

I used a variety of tools to create the graphs in this book. Many, if not all, could probably be created in programming language tools like R or JavaScript; a sizable proportion could be created in drag-and-drop tools like Excel and Tableau. By utilizing a variety of tools, I found that some were easier to create the graph, but harder to style the way I liked and some were more difficult to learn while others were more intuitive. This list of tools is not exhaustive (by far) and I based my decision on which tools to include on my experience in the field, not on any formal survey or data set. Before you invest substantial time or money to learn any of these, you should explore the landscape of available tools and products.

PRIMARILY DRAG-AND-DROP OR CLICK-BASED

Adobe Illustrator. This is primarily a design tool. Adobe Illustrator, and the rest of the Adobe Creative Suite like Photoshop and InDesign, are the workhorse tools for designers. The graphing library is actually pretty poor in Illustrator, but you can insert graphs made in

other tools to add more styling, labels, and annotation. The Adobe Creative Cloud is now primarily subscription-based, but is not a particularly cheap tool to purchase.

Charticulator. Launched in 2018, Microsoft's Charticulator is an online tool that enables you to create a custom chart layout. It differs from some other tools in that you don't select from predefined charts but instead the creators have transformed chart specifications into mathematical layout parameters, such as marks (e.g., rectangle, line, or text) and axes (e.g., property of one direction in a plot). Charticulator charts are primarily static, but can be integrated with Microsoft's PowerBI tool to create interactive visualizations. At the time of this writing, Charticulator is free to use.

Datawrapper. An online tool from a team based in Germany where you can upload your data, select a graph template, refine and style, and publish or download. Created charts can be embedded in websites and can be made interactive. Datawrapper is free for most purposes, and paid versions allow for custom themes and additional exporting options. There are many tools in this same vein (Flourish and RAW below are two other examples)—some are better than others with different default options and user experience.

Excel. Likely the primary data and data visualization tool for many people around the world. As part of the Microsoft Office suite, Excel is not free, but it is your basic click-based tool. At the time of this writing, you can create more than 16 basic chart types in Excel, with variations within most of them. There is a basic library of charts that you can expand with clever "hacking" or additional coding using the Visual Basic for Applications (VBA) programming language that sits behind all Microsoft Office tools.

Flourish. Launched in 2016, Flourish is an online tool primarily aimed at newsrooms to help journalists create both static and interactive data visualizations in a drag-and-drop framework. There are options to customize and further develop Flourish graphs using the underlying JavaScript framework. Pricing options range from the free Public version to the paid Personal version (which gives you additional features) and the paid Business version (targeted towards large teams and organizations). Through a partnership with the Google News Lab, Flourish provides newsrooms free premium accounts.

Google Sheets. Akin to Excel, Google Sheets is part of the Google suite of tools. It works very similarly to Excel, though without some of the more sophisticated options. Because it is based online, the sharing capabilities are somewhat better than Excel (with the side effect being you need to have internet access to use it).

PowerBI. Microsoft's business intelligence tool that allows you to create interactive dashboards and visualizations. It directly links with the rest of the Microsoft Office suite

(especially Excel) and can be modified and customized in ways similar to Tableau. There is the free Power BI Desktop version, the paid Power BI Pro, and the Power BI Premium package for organizations.

RAW. Created by DensityDesign Research Lab in Italy in 2013, RAW was an early project designed to help creators link spreadsheet tools like Excel to graphic editing tools like Adobe Illustrator. It is an open source tool, which means you can download the code to further customize the visualization options. There is also an online platform in which, like other tools, you upload your data, select the graph, and customize. RAW has a variety of options for certain non-standard graphs (like streamgraphs and bump charts) that are not typically available in other tools. RAW is free to use.

Tableau. Perhaps the most popular business intelligence dashboarding tool, Tableau's drag-and-drop interface enables you to create interactive dashboards and visualizations. Like Excel, users have customized their Tableau work to create an array of visualizations outside the basic graph menu. There are a number of versions of Tableau, from the free Tableau Public (but which means you save your work to the Tableau website) to paid versions like Tableau Desktop and Tableau Server (for large organizations).

ONLINE TOOLS (CLICK-BASED)

Infogram, Venngage, and Vizzlo. These are just three of the many click-based online tools that are aimed more for people who want to quickly create infographics and reports. In my experience, these tools sometimes have more graph options than other online tools, but they are not always based on best practices. Pricing varies from free packages that usually mean your data and visualizations can be viewed by anyone, to enterprise packages for large teams and organizations.

PROGRAMMING LANGUAGES

D3. We first need to define JavaScript. JavaScript is programming language that allows you to implement information onto a webpage. Every time a web page does something, like display updates, animate graphics, or play videos, JavaScript is probably involved. D3 is a JavaScript library for manipulating objects based on data and was developed by Mike Bostock along with Jeff Heer and Vadim Ogievetsky at Stanford University in the early 2010s. Most of the interactive data visualizations we currently see on the web are run on D3—virtually

every interactive graph you play with on the *New York Times*, *Washington Post*, and *Guardian* websites is built with D3. Like other programming languages, there is a steep learning curve to using D3, but you are basically unlimited in the kinds of visualizations you can create. D3 is an open source language, which means it is free to use.

Highcharts. Launched in 2009 by a team in Norway, Highcharts—and its cousin tools Highstock, Highmaps, Highcharts Cloud, and Highslide—is a suite of interactive data visualization tools rooted in JavaScript. You do need to know a bit of coding to use Highcharts, but templates and libraries help you create the basics of a graph and then add additional styling and formats. Highcharts is free for personal use and nonprofit organizations; pricing then varies by number of licenses and package.

Python. Python is used in a wide number of fields and industries from basic and complex data analysis to web application to artificial intelligence. Like D3 and R, the language is open source, which means there are a lot of open, free libraries to help you create data visualizations including, for example, Matplotlib, Seaborn, Bokeh, and ggplot.

R. Conceived in 1992 and initially released in 1995, R is a free, open source programming language for statistical computing and graphics. R is becoming more and more widely used for data visualization, especially since the launch of the "ggplot2" package by Hadley Wickham in 2005 (based on the "Grammar of Graphics" by Leland Wilkinson). R allows you to conduct statistical analyses and create customizable data visualizations; additional tools and packages can enable you to create interactive visualizations.