Introduction

In early 2020, countries across the world struggled to contain the spread of COVID. One country, though, succeeded where others did not: New Zealand. There are many reasons why New Zealand was so successful in tackling COVID. One of these was R (yes, R).

How did a humble tool for data analysis help New Zealand fight COVID? R helped a team at the Ministry of Health to generate daily reports on cases throughout New Zealand. These reports (there were three each day, each for a slightly different audience) were essential in helping officials develop policies that kept New Zealand largely COVID-free. It was a big lift for a small team. Producing these reports every day with a tool like Excel would not have been feasible. As team leader Chris Knox told me, “Trying to do what we did in a point-and-click environment is not possible.” But with R, a few staff members wrote R code that they could re-run every day to produce updated reports.

The reports that the New Zealand Ministry of Health produced did not involve any complicated statistics – they were literally counts of COVID cases. The value that the team got was from everything else R can do: data analysis and visualization, report creation, and automating workflows. Many people think of R as simply a tool for statistics. But, over a quarter century since its creation, R can do much more than statistical analysis, and New Zealand used R to keep its residents safe from COVID.

I used to feel ashamed about the way I use R. As someone with an extremely non-quantitative background (I did a PhD in anthropology) who never used R in graduate school, I use R, a tool for statistical analysis, but I don’t use it for complex statistical analysis. For a long time, I felt like I wasn’t a “real” R user. Real R users, in my mind, used R for hardcore stats; I “only” used R for descriptive stats.

But eventually, I realized that, no matter what else you do in R, you have to **illuminate** your findings and **communicate** your results. And, the more you use R, the more you’ll find yourself wanting to **automate** things you used to do manually. I realize now that the things that I use R for are the things that everyone uses R for. R was created for statistics. But today people are just as likely to use R without statistics.

I’m excited to be your guide on this journey through the ways you can use R without statistics. If I, a qualitatively-trained anthropologist whose most complex statistical use for R is calculating averages, can find value in R, so can you. No matter your background or what you think about R right now, using R without statistics can transform your work.

Who This Book is For

This book is for you if you are either a current R user keen to explore new ways of using R or a non-R user wondering if R is right for you. I’ve written R Without Statistics so that it should make sense even if you’ve never written a line of R code. But if you have written many lines of R code, the book should help you learn plenty of new techniques to up your R game.

About This Book

This book shows the many ways that people use R without statistics. Each chapter focuses on one novel use of R. You’ll begin by learning about R users who have transformed their work using R. You’ll learn about a problem they had and how R helped them to solve it. We’ll dive into their code, breaking it down to help you understand how they used R. Each chapter will conclude with a short summary, offering lessons you can take from this novel way of using R. The book has three parts:

Part 1: Illuminate

In the first part, you’ll learn about ways to use R to illuminate your findings.

Chapter 2: Principles of Data Visualization This chapter breaks down a visualization by Cédric Scherer and Georgios Karamanis on drought in the United States. In doing so, it shows important principles that can help you to make high-quality data visualization.

Chapter 3: Making Your Own Theme This chapter shows how journalists at the BBC made a custom theme for the data visualization package known as ggplot2. We’ll break down the bbplot package and in the process you’ll learn how to make your own theme.

Chapter 4: Creating Maps This chapter walks through the code that Abdoul Madjid used to make a map showing COVID rates in the United States in 2021. You’ll learn how to use the ggplot2 package to make high-quality maps.

Chapter 5: Creating High-Quality Tables This chapter will show you how to use the gt package to make high-quality tables in R. Based on a conversation with Tom Mock, you’ll learn to apply design principles to ensure your tables communicate effectively.

Part 2: Communicate

The second part of the book focuses on using R Markdown to communicate efficiently.

Chapter 6: Writing Reports in R Markdown This chapter introduces R Markdown through a conversation with Alison Hill. A tool that allows you go from data import to final report, all in R, R Markdown can transform how you communicate. This chapter will introduce the basics to help you get started with R Markdown.

Chapter 7: Parameterized Reporting One of the advantages of using R Markdown is that you can produce multiple reports at the same time using a technique called parameterized reporting. In this chapter, I speak with staff members at the Urban Institute about how they used R to produce fiscal briefs for all 50 U.S. states. In the process, you’ll learn how parameterized reporting works and how you can use it.

Chapter 8: Making Slideshow Presentations with xaringan In addition to traditional reports, R Markdown can be used to make slides. You’ll come away from this chapter, which is based on my conversation with Silvia Canelón, ready to make your own presentations with the xaringan package.

Chapter 9: Building Websites with distill R Markdown can also make websites. In this chapter, I speak with Matt Herman about how he used the distill package to make a website about COVID-19 rates in Westchester County, New York. The chapter will show you how to create your own website with R Markdown and distill.

Part 3: Automate

The last part of the book focuses on ways you can use R to automate your work.

Chapter 10: Accessing Online Data In addition to working with data you already have, R can help you to automatically access data. This chapter shows two packages that can bring in data: googlesheets4 for working with Google Sheets and tidycensus for working with United States Census Bureau data. Through conversations with Meghan Harris and Kyle Walker, you’ll learn how the packages work, and how you can use them to automate the process of accessing data.

Chapter 11: Code Once, Run Twice: Creating Your Own Functions One of the major benefits of R is that you can create your own functions to automate common tasks. In this chapter, I show a few example functions that I and others have made. You’ll come away ready to make your own R functions.

Chapter 12: Bundle Your Functions Together in Your Own R Package Once you have a set of functions that you use regularly, you’ll want to bundle them into a package. Doing so makes it easy for you and others to use the code you’ve written. I speak with Travis Gerke and Garrick Aden-Buie about how they created packages to improve the work of researchers at the Moffitt Cancer Center. This chapter will set you up to make your own R package.

Before we dive into the book, I have a favor to ask. This book is called R Without Statistics. But it’s not meant to be taken literally. Of course it’s true that if you’re making a graph you’re using statistics. Before you start typing an angry email, please know that R Without Statistics is a mindset, not a statement meant to be taken literally. We’re all using R with statistics already. Let’s learn to use R without statistics.