Book Reviews

ELIZABETH ZWICKY

R for Everyone

R for Everyone is an introduction to using R that does not assume that you know, or want to know, a lot about programming, and concentrates on covering the range of common uses of R, from simple calculations to high-end statistics and data analysis.

In three straightforward lines of R, you can load 200,000 lines of data, calculate the minimum, maximum, mean, median, standard deviation, and quartiles for every numeric field and the most popular values with counts for the strings, and then graph them all in pairs to see how they correlate. Then you can spend 30 minutes trying to remember the arcane syntax to do the next thing you want to do, which will be easy as pie once you figure out where all the parts go. As a result, using R when manipulating numbers will not only simplify your life, it will also give you a reputation as an intellectual badass.

Of course, to get there, a guide will help. Many guides, however, are written by people who want to teach you statistics or a particular programming style, or who seem to find R intuitive, which is lucky for them but does not help the rest of us. The good and bad news is that *R for Everyone* does not want to teach you statistics. This is good news because it frees the book up to teach you R, and there are plenty of other places to learn statistics. It is bad news if you are going to feel sad when it casually mentions how to get R to produce a Poisson distribution, and you have no idea what that is. You shouldn't feel sad. Just move on, and it will be there when you need it.

Because I am more of a programmer than a statistician, I can't vouch for whether *R for Everyone* is actually sufficient for non-programmers. It is plausible that it would give a statistician enough programming background to cope, although I certainly wouldn't recommend it as a programming introduction for anyone who didn't find the prospect of easy ways to regress to the mean enticing. It is certainly gentler as a programming introduction than other R books with that aim.

It is also unusually comprehensible for an R book. I would recommend it as a first R book. You still also probably want a copy of O'Reilly's *R Cookbook*; the two books are mostly complementary, although their graphics recommendations are moderately incompatible. *R for Everyone* is more up-to-date, and the more traditional format is easier to learn from, while the *R Cookbook* is more aimed at specific problems, which makes it easier to skip through in panic looking for that missing clue.

My one complaint about *R for Everyone* is that some of the early chapters are insufficiently edited. There's some odd word usage and at least one example that is puzzlingly wrong (fortunately, it's the example of doing assignment right to left instead of left to right, which you should pretend not to know about anyway. Nobody does that; it's just confusing).

Threat Modeling: Designing for Security

This is a great book for learning to think about security in a development environment, as well as for learning to do threat modeling itself. It's a practical book, written from the point of view of an experienced practitioner, and it presents multiple approaches. (If you believe there is exactly one right way to do things, you will be annoyed. In my experience, however, people who believe in exactly one right way are people who enjoy righteous indignation, so perhaps you should read it anyway.) Also, it's written in a gently humorous style that makes it pleasant to read.

The basic problem with writing a book on threat modeling is that you have two choices. You can talk about threats, but not what you might want to do with them, which is kind of like writing an entire cookbook without ingredients lists—you get lots of techniques, sure, but you have no idea how to actually make anything. On the other hand, you can talk about threats and what to do about them, which leaves you trying to cover all of computer security in one book and still discuss threat modeling somewhere. Shostack goes for the latter approach, which is probably the better option, but it does make for a large and thinly spread book. And, because the focus is threat modeling, some great advice is buried in obscure corners.

I'd recommend this book to people new to the practice of security, or new to threat modeling, or unsatisfied with their current threat modeling practice, including people who are in other positions but are not being well served by their security people.

There are, of course, some points I disagree with or feel conflicted about. I agree that "Think like an attacker" in practice leads people down very bad roads, because they think like an attacker who also is a highly competent engineer who understands the product internals, rather than like any real attacker ever. On the other hand, there's an important kernel of truth there that needs to get through to otherwise intelligent programmers who say things like, "Oh, we don't call that interface any

more, so it doesn't matter." (What matters is not what the software does when you operate it as designed, but what is possible for it to do.)

Data Protection for Photographers
Patrick H. Corrigan
Rocky Nook, 2014. 359 pages
ISBN 978-1-937538-22-4

If you are not a photographer you may be wondering why photographers would need to think specially about data protection. If you are a digital photographer, you're used to watching your disks fill and wondering how you're going to keep your pictures safe when most people's backup solutions are sized for a hundredth the amount of data you normally think about. You spend a lot of time realizing that data storage intended for homes is not going to meet the needs of your particular home. In fact, even small business systems may not suffice.

This is not a problem if your first love is system administration and photography is just your day job, but for the rest of us, it is at best annoying and at worst incomprehensible. This book will fix the incomprehensible part, explaining enough about disk and backup systems to allow photographers to make good decisions. Sadly, there's still no easy answer. You will have to choose for yourself which annoying tradeoffs to make, but at least you will make them knowingly.

I would have liked more emphasis on testing restores, and some coverage of travel options. However, this is the book you need if your data at home is outstripping the bounds of your current backup solution in either size or importance.