

The book version

Ben Klemens

16 November 2012

Hi. I'm back.

In case you hadn't worked it out, I was out working on turning that tip-every-other-day series into a book, made from paper. That's now done, and the next several entries are going to be about statistical technique. This column here will be a few notes on the book and its production, in case you're interested in such things.



Figure 1: My mock-up of the cover, which never got used.

The book came out a week or two ago—just over a year and a month after I Entry #049. Here it is on Amazon¹, or buy it directly from O'Reilly², or buy the paper book from Amazon, then register the print book at O'Reilly and (at the Your Products → Print Books page) upgrade to an electronic version.

You've got options, because print is increasingly just another medium for viewing PDFs. According to the acquiring editor at O'Reilly, *most* of their book sales are now non-paper editions. A tech press like O'Reilly is clearly ahead of the curve on this, but it's increasingly clear that that's just how it's gonna be, and we are in the transition period from paper-norm + electronic-niche to electronic-norm + paper-niche.

The O'Reilly back-end is of course all smoothly automated. I check out from their Subversion repository, add text, check back in, and next thing you know I have a publication-quality PDF on hand. They make some changes over there, and then what I wrote has an index. When we're all happy with the product, they change a few settings on the back-end, and the edition gets sent out to the world in mobi/PDF/&c. formats.

[Yes, they still use Subversion, which was odd to me. Every time the editors and I had logistic troubles, it was over Subversion's trouble with merging simultaneous work.]

So writing a book is now largely like printing any other PDF, plus the additional organization that goes into any major project.

¹<http://t.co/xIQKSzbs>

²<http://shop.oreilly.com/product/0636920025108.do>

Book format vs article/blog format All that said, the book format is still unique and essential, for a few reasons.

The first is in the concept of a largely self-contained work that goes into great detail about a subject. A series of articles or blog posts makes it so easy to leave some topics unexplored, because either the author will get to them later, or can just add a link or a reference to some other place that covers that topic in detail. Our expectations about a book are that it will cover the ground it chose to cover, in one place and with one consistent worldview and notation. If there are big holes, people will complain and leave two-star reviews on Amazon.

[Have you gotten a login to goodreads³? I've found that reviews there tend to be a little more sincere and positive than reviews on larger sites. The reviews feel more like they are notes written to friends, as opposed to Amazon reviews, which read like an address being read from on stage.]

The second is that correcting an error that has been printed in a thousand copies is rather difficult. That forces another sort of discipline on the author—especially with a technical book where every page will have a few dozen factual statements that could be wrong.

So on both the table-of-contents level and the micro-detail level, there's a stronger incentive to get things right and produce a better quality of output.

Reconciling the two books This is a blog about *Modeling with Data*, and here I am plugging another book.

They're both books involving C, so if you left it at that, they'd be overlapping. I really hope you aren't leaving it there.

But if you were using both as a C textbook, you might want to start in *Modeling with Data*, Chapter 2, which is a pretty complete basic C tutorial. Then, work through *21st Century C*, which assumes basic C knowledge, and covers the environment and still more C tricks. [I like the coverage of makefiles in *21st Cent. C* better than Appendix A of *MwD*, so I hope you didn't buy *MwD* for the appendix on makefiles.] At that point, the rest of *MwD* will look really easy and feel really comfortable because—maybe I've mentioned this before—better computing technique makes you a better statistician.

At the end of all that, you'll know everything I do, on top of your own prior knowledge. Books are great that way.

³<http://www.goodreads.com/book/show/14514281-21st-century-c>