

## TECH WE'RE USING

# In Data Journalism, Tech Matters Less Than the People

Ben Casselman, an economics reporter, uses a programming language called R and works with vast data sets. But he says interviews still make for the best stories.



Featuring Ben Casselman

Nov. 13, 2019

**Sign up for the On Tech newsletter.** Get our best tech reporting from the week. [Get it sent to your inbox.](#)

*How do New York Times journalists use technology in their jobs and in their personal lives? Ben Casselman, an economics and business reporter, discussed the tech he's using.*

**You compile complicated data sets and distill them into stories for our readers. What are the tech tools and methods behind this madness?**

“Madness” is tough but fair. If you walk by my desk on any given day you’ll find my computer monitor littered with charts, spreadsheets and way more Chrome tabs than any sane person would consider reasonable. And my physical desktop is just as cluttered with half-used notebooks, printed-out economics papers and my trusty TI-86 calculator (a holdover from college calculus that I still find inexplicably useful — apparently a theme among economics reporters at The Times).

Really, though, the most important piece of technology on my desk is my landline telephone. I think some people have the idea that “data journalism” means staring at spreadsheets until a story magically appears, but in the real world that almost never happens. The best stories almost always emerge from talking to people, whether they are experts or just ordinary people affected by the issues we write about. They’re the ones who pose the questions that data can help answer, or who help explain the trends that the data reveals, or who can provide the wrinkles and nuances that the data glosses over.

For example, one of my favorite stories from the past year is one that I wrote with my colleague Conor Dougherty about investors buying up single-family homes. We had access to a huge data set — millions and millions of real estate transactions — that showed how investors had come to dominate the market for starter homes in many cities.

But what really made the story come to life was when we zoomed in on one house that had changed hands several times and got to talk to all the people who had touched it — the investor who flipped it, the family that bought it, the would-be buyer who kept losing out to investors.

At the end of the day, data isn't the story; people are the story.

**Fair enough. But I assume you aren't doing all that data analysis on your phone. Or on your TI calculator, for that matter.**

Yes, that's true. I do most of my data analysis in a statistical programming language called R. It lets me work with data sets that have hundreds of thousands or even millions of rows, much too large for a spreadsheet program like Excel. It also makes it easy to automate tasks that I perform regularly — so when the Labor Department's monthly jobs report comes out, for example, I can download the new data and update my analysis with just a few keystrokes.

R also has a great suite of tools for making charts, which is a critical part of my work. The Times has the best graphics team in the business, and I can't come close to doing what it does in terms of making beautiful, readable charts and interactive graphics. But I'm not trying to make charts like that. I just need something that lets me spot a trend or relationship, or that helps identify examples that are worth exploring further. For that, quick-and-dirty is just fine.

In terms of hardware, I wish I could tell you I was running some hotshot rig with multicore processors and a boatload of RAM. But these days I do pretty much all my work on a standard-issue laptop. I store some data in the cloud, but that's mostly because I don't want to lose it all if my computer melts down.

That's mostly a reflection of how quickly technology has improved. A few years back, when I was at The Wall Street Journal, I had to go begging for extra processing power, and even then my co-workers used to complain that my computer sounded like a jet airplane struggling to gain altitude. These days, my computer can manage much larger data sets without breaking a sweat.



Mr. Casselman coding in R, the statistical programming language, with an office phone and his TI-86 calculator close at hand. Karsten Moran for The New York Times

**There has been a running debate in journalism circles about whether reporters should learn to code. I take it you're firmly in the "yes" camp?**

Not at all! Coding has been a valuable tool for me, and I'm glad that more journalists are learning to code (and that more coders are getting interested in journalism). But I was a reporter long before I learned how to code, and most of the reporters I admire most have never written a line of code in their life.

I do think that pretty much all reporters need to have a basic comfort with numbers and statistics. Not that everyone needs to be able to run regressions or calculate p-values, or even define what a p-value is (which isn't easy, by the way). But they should be fluent enough in the language of statistics to understand when an argument makes sense, and when it's suspect enough that they should start probing more deeply. And, yes, I think you should be able to do a percent change calculation without turning to Google for help.

In the last couple of years, The Times has developed a course to teach basic data skills to reporters and editors — things like how to vet data to make sure it's legitimate, how to evaluate statistical claims and how to use a spreadsheet program to explore a data set. One of the core messages is that you can do a lot of this stuff *without* learning to code. I do a session where I walk through how to replicate a story of mine using nothing but Google Sheets and a few basic functions.

## Outside of work, what tech product are you personally obsessed with? What do you do with it?

Despite my love of data, I'm not much of a techie in my nonwork life.

I don't have a smart watch or a smart thermostat or anything like that, and I still had an old cathode-ray TV until my in-laws finally broke down and sent us a flat screen. (I did give my wife a HomePod, Apple's smart speaker, for her birthday a couple of years ago, but it mostly just terrifies me by having Siri pipe up seemingly at random.)

Probably my most prized digital possession is my kitchen scale. I've gotten really into baking sourdough bread, without any commercial yeast, which means it's about the lowest-tech form of baking imaginable.

**Ben Casselman** writes about economics, with a particular focus on stories involving data. He previously reported for FiveThirtyEight and The Wall Street Journal.

---

A version of this article appears in print on , Section B, Page 6 of the New York edition with the headline: Fleshing Out the Data With the Help of People