**By MaryJo Webster**

**@MaryJoWebster**

[**Mjwebster71@gmail.com**](mailto:Mjwebster71@gmail.com)

**What is a “data state of mind”?**

Journalists just learning how to analyze and visualize data always ask me how to incorporate their new skills into their everyday work.

I’ve learned over the years that, in addition to learning how to punch the buttons in their software of choice and understanding some basic analysis concepts, they also need to train their brain to have a “data state of mind.”

To me, that’s the biggest key to making data a staple in your reporting.

“Data state of mind” is a phrase that I’ve sort of stolen, and slightly adapted, from the “documents state of mind” that great investigative journalists Donald Bartlett and James Steele regularly promoted at Investigative Reporters and Editors conferences in years past.

The modern twist is that instead of constantly being on the lookout for documents (i.e. pieces of paper) that will support or prove your key point, you should be looking for data stored in computers. Usually that’s going to be data that you analyze in order to quantify or measure something or otherwise prove your point.

The common thread between the two concepts is that you’re not relying solely on people to tell you what’s going on. Documents and data tend to have a better memory and be more honest.

You’ll also usually get a much better story.

Let me give you a few story examples of a “data state of mind” in action, and then I’ll give you some tips on how to train your brain.

* A suburb of St. Paul, Minn., was considering a new law that would set prohibitions on where high-level sex offenders could live. The reporter had done a simple, traditional story saying the council was going to discuss this at an upcoming meeting. I suggested that we try to measure what might happen if they passed this proposed law. The big question: Would there be anywhere for the sex offenders to live in this small suburb? Using mapping software, we plotted locations of the places the sex offenders would have to stay a certain distance away from under this new law – schools, churches and parks being the big ones – and added buffers that represented the prohibited areas. The resulting map showed the prohibited areas covered almost the entire city. A few days after our story ran, the council dropped the proposal.
* A reporter colleague of mine got a press release a few years ago about an increase in children being treated in U.S. hospitals for medication overdoses, so he decided to see if that held true in Minnesota. Instead of the usual approach of calling experts, he first turned to the state’s death certificate database to find poisoning-related deaths. He didn’t see the trend he was looking for, but found an even better one. It wasn’t primarily children dying from overdoses of over-the-counter medications, it was young adults. When he started calling state and federal experts, they told him this was a topic that had just started hitting their radar. It was too new for a press release.
* Another reporter wanted to know how often people were criminally charged (and convicted) for killing a pedestrian in a traffic accident. It didn’t take long for him to learn that nobody in Minnesota tracked this. There weren’t reports out there. There wasn’t even a database. The so-called experts had no answer to his question. But the pieces of data existed. Our state compiles a crash database of traffic accidents, including all that result in injury or death and many others that did not. But it doesn’t have any names or other identifying information about the crash, other than date and county (sometimes the city). But we also have a death certificate database that includes date of injury/death and very detailed information about where and how the person died. Since pedestrian deaths are so rare, it was easy to match the two datasets to find the victims. The crash data also included the official case number with the police agency that handled the crash. That led him to reports that included names and made it possible to search court records to look for charges and/or convictions. In the end, he essentially built his own database, but it gave him a one-of-a-kind story. (And the answer is that they are rarely charged)

Hopefully you noticed that all three of those examples don’t rise to the level of a “project” and all came out of beat reporting. All were done within a matter of days, maybe weeks (especially the third). All of them also benefited from the fact that my newsroom had most of the required data in-house and up-to-date, or the data we needed was readily available on the Internet.

And most importantly, all of them had something they were trying to measure: How much of the city would be left for sex offenders to live in? Are more young people dying from medication overdoses? How often are people convicted for killing pedestrians?

Some tips:

Think of data as people too.

The database you get from a public agency is just like a human source. It can answer questions, raise questions, point you in the right direction, or even mislead you if you’re not careful in not only how you ask the question, but also in how you interpret the answer. It can be a tipster or it could be the key source in your story or it could just offer some background or context.

Define your story ideas as questions, not statements

Instead of saying, I want to do a story about unsafe bridges…say I want to find out what percentage of bridges in the state are unsafe? This frames your story into something that is quantifiable and helps you figure out how to get started.

Tune your radar to pick up opportunities to quantify

Some example to watch out for:

Trend stories that vaguely say something has changed over time or is bigger or different here versus there

When interviewing sources, listen for times when they refer to something that was measured – when they talk about something increasing or decreasing; when they clearly make it sound like there was some analysis to get that answer. Ask them, where did you get that? And then, perhaps, you might also want to ask to get the raw data yourself.

When you find yourself asking a source to give you summary numbers for a story, stop and think whether it would be better to get the data that those numbers are based on. For example, if you are writing a story about how overtime has caused the police department’s budget to go through the roof… instead of asking the city to give you total figures on overtime this year versus last year (or some other breakdown), ask yourself if you’d get a better story if you had data showing how much overtime each employee got?

A local government (city, county, state, etc) created a program to do something last year. Can we find a way to measure whether it has succeeded?

A breaking news story. How often has this occurred in the past?

IRE’s Extra Extra to find ideas that others have done

Look at your own past stories and look for missed opportunities – could you go back and do a deeper or broader look at the topic using data?

That rumor/myth that is always circulating. Can you figure out whether it’s true?

Learn what datasets are available on your beat

What is collected and why? How does the government agency use the data? What reports are generated from it? This will prove useful in a couple ways…first, it might give you more insight into your beat and second, it might set the stage in case you need some of this data down the road (especially if the story is time-sensitive).