

Module 4.1 - What data can do for your stories

Hi, this is Denise Malan. I'm the senior training director with investigative reporters and editors.

Welcome to Module 4. We're going to talk about one of my favorite topics, which is using data in your investigative reporting. This first video this week is going to talk about the concepts of using data and what data can do for your stories. In the second video, we'll talk about the technical tools and some resources to learn those tools.

I want to start with a definition of data and the type of data that we typically use in investigative reporting is structured data that looks like this. It might come in a spreadsheet, perhaps in a text file, but it has columns and rows and those are well-defined. And the source of your data tells you everything that is in those columns and rows and what they mean.

Let's talk about the process of incorporating data into your reporting. First, you need to spend some time identifying what data might be useful to your reporting and what is available out there. You need to gather that data and we'll talk about the sources here in a minute.

A very crucial step is evaluating the data. What source did you get that from? Are they a trustworthy source? What methodology did they use? And is that a sound methodology? Do you have everything that you need? Are there other data sets that you might need to give this better context? How dirty is the data? Sometimes names are misspelled or places are misspelled and they don't allow you to analyze it correctly until you actually spend some time cleaning the data.

After you've done all of those steps, then you can start analyzing it and do your reporting from the questions that you find in your data. Then you can optionally, you can visualize your data set. You might have something that is good for mapping or you can make some charts or perhaps a searchable database that people might need to search and find information in before you finally publish your project. Now, this process depends on the data set that you're using in your project. Sometimes it might take months or years, but sometimes you might find the data online and you can do something in a breaking new situation the same day. So, the process can really vary depending on your project.

Let's also talk about the why journalists use data and incorporate it in their reporting. One of the most important reasons is that data can give context to your reporting. It helps you show the big picture. You might show how often an issue is happening or the total amount spent on a project in your area. Now, at the same time it shows the big picture, data often helps you zoom in and find an anecdote as well, or different examples that can help you tell good stories, tell a good narrative and find characters that help your audience connect with the story.

Data is a crucial part of watchdog reporting as well. It helps you test government procedures or test something that an industry is claiming. And it helps you uncover information that maybe no one else has looked at before.

It also helps you make your stories more interactive. Sometimes you might be able to make a map or some kind of interactive that helps people find their own information in your story. People might be able to find their own crime rate in their own neighborhood or the

rating for their child's school, for example, that helps them really personalize and make the news more useful for them.

I want to look at a couple of examples as well and show how journalists are using data in their stories. The first example is from Education Week. This project is about how Hurricane Maria disrupted the life of school children and their parents. After numerous schools were closed across the island. The bar chart here shows the big picture of how many students had to travel certain distances to their new school. And then in the map, you can zoom in to find the smaller picture of how far each distance was for each school.

The second project is from Reuters. This is an example of testing a government procedure. They come across court, court data and police data to find 1000 incidents in which people died after police use tasers on them. This project really countered the police claim that tasers are a good non-lethal option.

Our third example is from Rutas del Conflicto in Colombia, which was a project, a massive project that pulled together original reporting for every massacre during armed conflicts going back to 1982 in the country. They pulled together maps, timelines, information about victims to really show the personal human toll of conflicts across their country. It was a massive project, received a lot of awards internationally and very well deserved.

Our federal example is from Canada. CBC and the Toronto Star. And they took a look at an issue that is a big problem for anyone who's used Ticketmaster or other sites to buy tickets to sporting events or concerts. They scraped data from the site to show the big disparity in ticket prices and how scalpers can inflate the prices. They did a huge series using this data. There's a link on these slides as well to show how they scraped the data.

So how do you find data to try to incorporate it into your stories? A good first step is to Google it and to try to find data that is that deals with your issue. There are many open data portals as well across the world and for different countries. So online sleuthing is a good first step. If you're a government agency doesn't have data available, you might try to find some from different journalism organizations or even other NGOs that collect their own data. Just remember to verify your source and their methodology.

One good example of a journalism organization that provides data is OCCRP, the Organized Crime and Corruption Reporting Project. And they put together the site Aleph, which has data from many different countries to try to help you find connections between business people and politicians and try to find some possibly corrupt situations.

You also might have to use your sources and try to find data that's not available online and use your open records, laws that are available to you in your country to request those. If no one has collected this data before, then you might have to actually collect your own. You can do this through crowdsourcing or if there are paper documents available, you can build your own database from those.

One of the resources we've provided to you in this course is an international data tip sheet from our IRE Tip sheet database. One of our international members provided this and it has many different links to data sets across the world. So hopefully you can find something there that interests you as well.

All of these slides from this lecture are available on the Journalism Courses website so that you can explore these links on your own. I'll see you in the next video.