

Data journalism is the new punk

Simon Rogers

*Come on, anyone can play, the results might even be raucously
memorable in years to come, says a member of the band*

This is a chord ... this is another ... this is a third. NOW FORM A BAND. So went the first issue of British punk fanzine *Sideburns*, in 1977, in the “first and last part in a series”. It might be 35 years old, but this will do nicely as a theory of data journalism in 2012. Why? Arguably punk was most important in its influence, encouraging kids in the suburbs to take up instruments, with little or no musical training. It represented a DIY ethos and a shake-up of the old established order. It was a change. Crucial to it was the idea: anyone can do it.

Is the same true of data journalism? Do you need to be part of a major news operation, working for a big media company to be a data journalist? Now is the time to examine this. In May 2010, we published a piece at *The Guardian* on how journalists would be flooded with a “tsunami of data”. A few years on and data journalism is part of the fabric of what we, and many other news organisations, do.

What is it? I would say data journalism incorporates a wide range of styles – from visualisation to long-form articles. The key thing they have in common is that they are based on numbers and statistics – and that they should aim to get a “story” from that data. The ultimate display of that story, be it words or graphics, is irrelevant, I think: it’s more about the process. There are even different streams now – short-form, quick-and-dirty data visualisations of the kind we do every day on the Datablog, right through to complex investigations and visualisations such as our riots data analysis or the kind of projects from around the world, which made the

31



shortlist of the Data Journalism Awards.

So, can we still say that anyone can do data journalism; in the first and last part in a series, would this work?

- 1) This is a dataset.
- 2) Here's another.
- 3) Here are some free tools.
- 4) Now be a data journalist.

Why data journalism is the new punk

OK, it lacks a certain 1976 grittiness, but the theory is there. You don't have to be a developer or a coder to be a data journalist. We asked our Twitter followers what they thought. A couple stand out to me: "Maybe everyone can do it, but not everyone can do it well." "Like so many other things, done well is a mix of art and science." Mutual disregard for shared constructs of authority? Shared overarching aim of revealing reality away from the facade? But is that enough? The thing about data journalism is that there are so very many "chords" – just the free ones could fill several training manuals: Google fusion tables, Tableau, Gephi, OutWit Hub, Google Refine ... Can anyone really do it?

Dan Sinker knows about both data and punk: he heads up the Knight-Mozilla News Technology Partnership and is a former editor of *Punk Planet*. He says there are some parallels – with a crucial difference. "While I agree with the premise – it's never been easier to do this stuff than it is right now – I think there are a few steps beyond just learning three chords when doing data journalism. For one, Legs [McNeil, who coined the word "punk"] didn't really say a band needed to be 'good' but I'd like to think we'd require that for data journalism. The theory goes that the punk bands we remember best are the ones that were good – but there needed to be a whole lot of kids experimenting and sounding awful before they got there. For what it's worth, I like the fact that there are many just trying stuff out, even if it is forgettable, because some of it will be amazing."

In fact, data journalism is a great leveller. Many media groups are starting with as much prior knowledge and expertise as someone hacking away from their bedroom. Many have, until very recently, no idea where to start and great groups of journalists are still nervous of the spreadsheets they increasingly are confronted with. It's rare for the news site reader to find themselves as powerful as the news site editor, but that's where we are



right now – and that power is only increasing as journalists come to rely more and more on their communities for engagement and stories. Says Sinker: “Where I think there are more parallels is in the fact that this is a young community (in years if not always age), and one that’s actively teaching itself new tricks every day. That same vitality and excitement that motivated punk, it’s motivating news hackers right now.”

Meanwhile, more and more news teams are discovering that data equals stories and bulking up their teams. Some would say it’s just an extension of work they’ve always done, but that’s to ignore the huge shift in power the web has created. “Some people think that this stuff is instant,” says Sinker. “Even though there are incredible tools now, there is still a learning curve.”

Out there in the world, there are lots of people who have just formed a band and got on with it – despite the obstacles. Take the data team at Argentinian *La Nacion*, recently shortlisted for the Data Journalism awards for their work on transport subsidies. When the team started, it was sparse, to say the least, says Florencia Coelho: “We had no web programmer or CAR [computer assisted reporting] people in our newsroom. We gathered an interactive designer and we self taught Tableau with their free training videos in what we called our Tableau days, in a Starbucks at a shopping mall in Buenos Aires.”

The team is still not exactly huge, but it is easily the best data journalism site in South America and one of the most innovative around. It’s not all about investigative reporting. First, all reporting probably counts as investigative journalism, but if you want to play semantics, then I will see your “investigative” and raise you “analytical”. Not all data journalism has to bring down the government – it’s often enough for it to shine a light in corners that are less understood, to help us see the world a little clearer. And if that’s not investigative, what is?

Democratisation of data

There’s a great democratisation of data going on. Rather than the numbers belonging to the experts, they belong to all of us – and data journalism is part of that reclaiming of the facts. Even at the OECD, users’ voices are part of the process, making up the core analysis that lies at the heart of the Better Life Index on wellbeing. And, just to be clear, data journalism doesn’t have to mean data visualisation. It is not about producing charts or intricate graphics, the results of data journalism just



happen to lend themselves to that. Sometimes a story is best told in images and infographics, other times it works as words and stories. It's the ultimate in flexible formats.

But, when it comes to visualisations, what really comes across from an analysis of Visual.lys' most viral infographics is how the simplest things can flood the web sometimes. Single charts are likely to be successful because they are easy to consume: the viewer only needs to learn how to read one "chunk" of visualisation to get the whole story. Simplicity lends itself to quick understanding and sharing, whereas complexity can prevent a viewer from reaching those points. Curiously, mixed charts, which is what we commonly think of as the typical form of an infographic, is the least successful here, perhaps because they take more mental work to consume completely, again pointing to simplicity and brevity as strengths in visual communication.

However, sometimes things done messily can still be hits – it's the information that's vital. People are willing to forgive a lack of perfection; they are much less forgiving for those who get the facts wrong. Data visualisation experts will always say: allow the data to choose the visualisation, that it's crucial for the visualisation to fit the numbers – and not the other way around. That question equally applies itself to whether something needs a visualisation in the first place.

Of course, for some people, this will never be journalism. But then, who cares? While they are worrying about the definitions, the rest of us can just get on with it. Punk eventually turned into new wave, new wave into everyday pop and bands that just aren't as exciting. But what it did do is change the climate and the daily weather. Data journalism is doing that too. In the words of Joe Strummer: people can do anything.

Simon Rogers is Data Editor at Twitter in San Francisco. He was previously at The Guardian for 15 years and created its Datablog. He is also the author of Facts are Sacred (Faber and Faber, 2013). This piece is reprinted from Data Journalism, Mapping the future, edited by John Mair and Richard Lance Keeble with Paul Bradshaw and Teodora Beleaga and published by abramis academic publishing.

