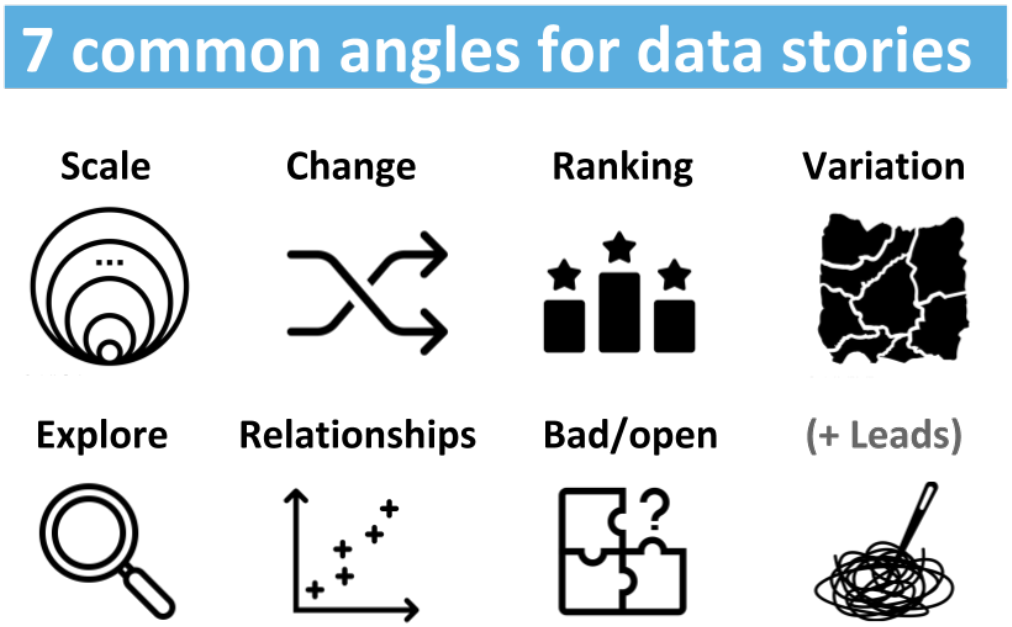
Here are the angles journalists use most often to tell the stories in data

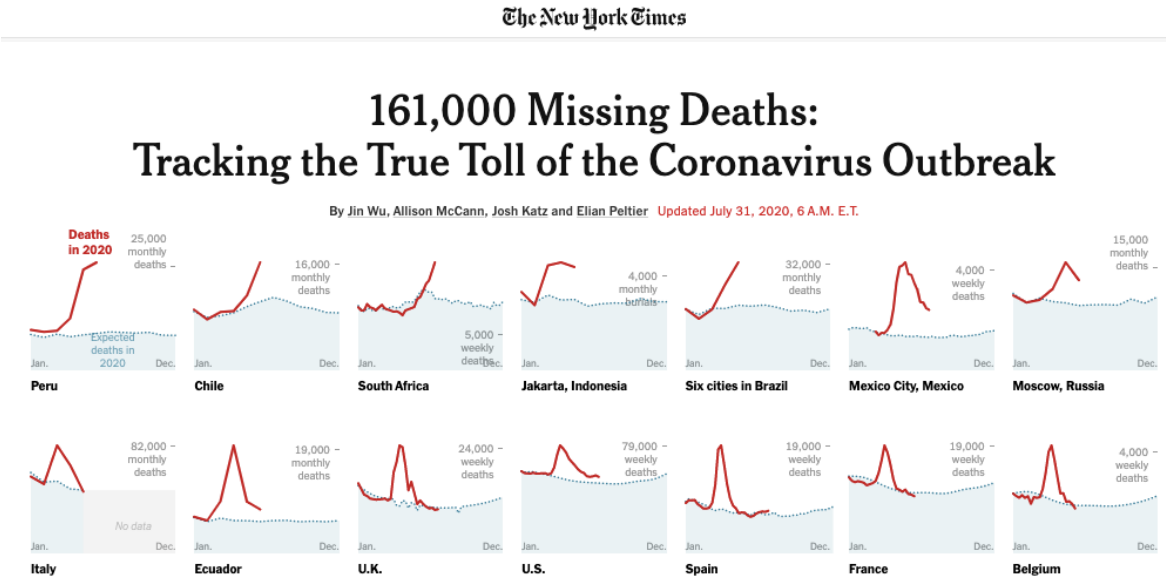


ONLINE JOURNALISM BLOG

Icons: the Noun Project: Becris (scale), Adrian Coquet (change and ranking), Kirby Wu (variation), Aradila Studio (explore) Trevor Dsouza (relationships), Iconpai (bad data), Kirill Ulitin (leads)

In my data journ ism teaching and training I often talk about common types of stories that can be found in datasets — so I thought I would take 100 pieces of data journalism and analyse them to see if it was possible to identify how often each of thos story angles is used. I found that there are actually broadly **seven core data story angles**. Many incorporate other ar gles as secondary dimensions in the storytelling (a change story might go on to talk about the scale of something, for ample), but all the data journalism stories I looked at took one of these as its lead. In the first of a two-part series I wal through how the four most common angles can help you identify story ideas, the variety of their execution, and the c siderations to bear in mind.

Data angle 1: Scale — "This is how big a problem is"



The pandemic has been full of stories about scale: this New York Times piece outlines the scale of the death toll

Perhaps the most common type of story found in data is the **scale story**: these are stories that identify a big problem the size of an issue which has become topical. At their most simple scale stories provide an **update** on new numbers ing released: it could be the latest unemployment figures, the amount of crime, air pollution, money spent on some a births, deaths or marriages. During the first months of the pandemic, for example, we had daily scale stories on the numbers of cases, deaths, and tests, among other things. Examples of scale stories include '*Death toll in UK care home from coronavirus may be 6,000, study estimates*', but also stories like '*Unduly lenient sentences review scheme 'inadequate*' where the lead is based on **reaction** to the scale of an issue you have identified. Sometimes scale is provided as ***back groud*** to a single-event story, as in '*Drone causes Gatwick Airport disruption*' (how many near misses are there?) or to a icy proposal, such as in '*New drivers could be banned from driving at night, ministers say*' (how many new drivers are unc 19). Scale stories are one of the easier genres to write: in many cases no calculation is needed. Indeed, the main work volved is likely to be in setting **context** to that scale — at its worst a scale story simply becomes a 'big number' story (" lot of money was spent on stuff" or "Something happens to a lot of people"), and the reader is left unclear whether thi actually newsworthy or just normal. For that reason it's important to put scale into context by using **percentages or p portions** (e.g. "one in five") or comparisons and analogies ("The money spent on the scheme is the equivalent of the wages of 500 teachers"). You might also bring in change and/or variation as a **secondary angle:** establishing historica context to the scale you've just outlined, or how that scale varies. In the New York Times piece above, for example, the "true toll" (scale) of the coronavirus outbreak is immediately contextualised by charts which show how that has chang since the start of the year, in different parts of the country.

Data angle 2: Change and stasis — things are going up, things are going down, things aren't happening

Image: Belfast Telegraph

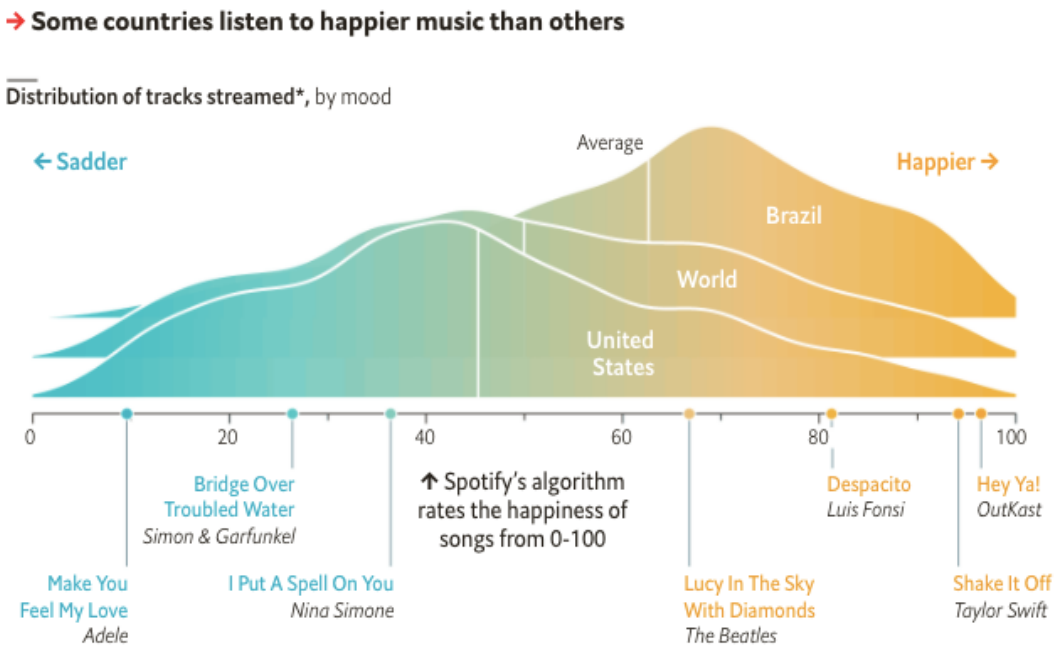
Change stories are almost as common as scale stories — and probably more straightforward to pitch. After all, change inherently newsworthy and gives you the verb ("rises", "plummets", "[goes] up") that you need in a headline. Once you identified some sort of change in your data it's likely you will need further reporting to answer the "**why?**" question. W are those numbers going up or down? You might also add a secondary angle to your story which explores **ranking** in that trend – the areas or categories where those numbers have gone up or dropped the most and least. This can help you direct your reporting on 'Why?' because chances are that the areas affected most will be those most aware of the sue, and able to comment on it. When reporting on change it's important to be aware of two considerations: **seasona** and **margins of error.** Seasonality is the role that (typically predictable and normal, and therefore non-newsworthy) s sonal factors can play in numbers, such as the end of a financial year or school term, the release of new cars or simply changing temperatures. Year-on-year comparisons (this August compared to last August, for example) or seasonal ad justment is often used to prevent this effect. The **margin of error**, meanwhile, is the range within which the *real* num bers actually lie. Because many datasets are based on *samples* which are then generalised to the rest of the populatio being looked at, a margin of error (or confidence intervals) is used to indicate how accurate that generalisation actuall is. If any change is within that margin of error then we can't really report that anything has changed. If your analysis fi a **lack of change**, that can also be newsworthy. In this story on company insolvencies, for example, we fully expected find an increase during the pandemic but the absence of any change in companies going bust was surprising — and le us to interview experts about why this might be the case. In other cases, a lack of change can also be significant becau it means things aren't improving — especially where resources have been invested, or promises made.

Data angle 3: Ranking and outliers — who's best and who's worst? Who's unusual and why?

Sad songs say so much

Data from Spotify suggest that listeners are gloomiest in February

Around the world, the most popular tunes this month will be depressing ones



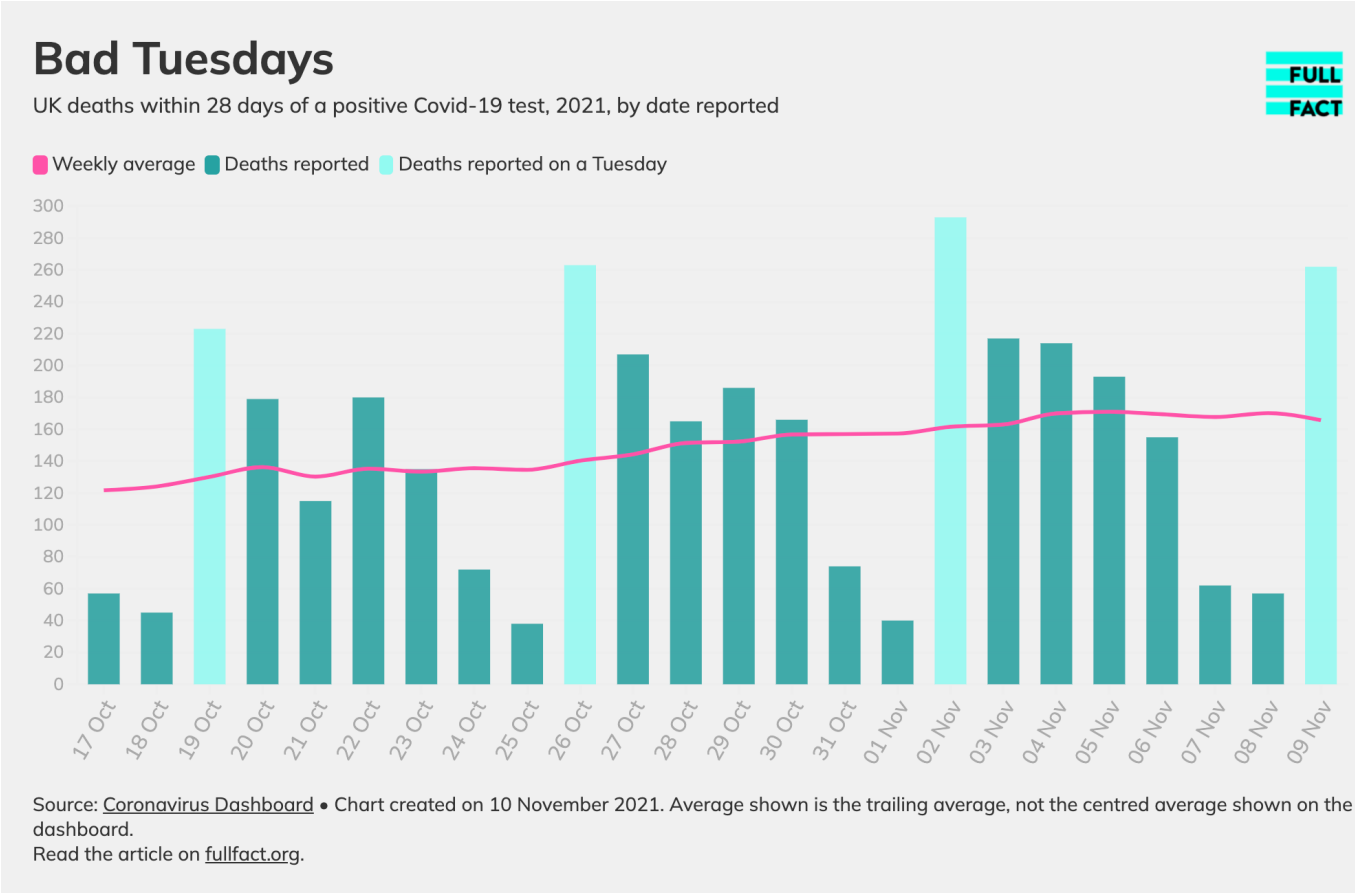
This Economist article is a 'ranking' story because it identifies the "gloomiest" month

Ranking stories are all about who or what comes out **worst or best** in a dataset, or where a particular entity of intere (the local police force, schools or teams, or an industry if it's the specialist press) **sits in comparison** to others. Typica stories in this category might include "Local area one of worst areas for crime" or "Local schoolchildren get third best r sults in the country". You might focus on the places 'worst-hit', as in "*The parts of Birmingham in top 10 UK areas worst-h by Universal Credit advances*", or you might look at where your sector compares to others, as in "*Construction is third m dangerous UK industry*." But ranking stories can also be about the best or worst **times, places or categories** that a dataset 'reveals'. The Economist article above, for example, is about the top-ranked month for listening to gloomy son A Birmingham Live story, on the other hand, leads on *The most common crimes in Sandwell – and where you're most like to be a victim*. The Economist, by the way, dedicated part of one data journalism newsletter to 'How to compile an inde

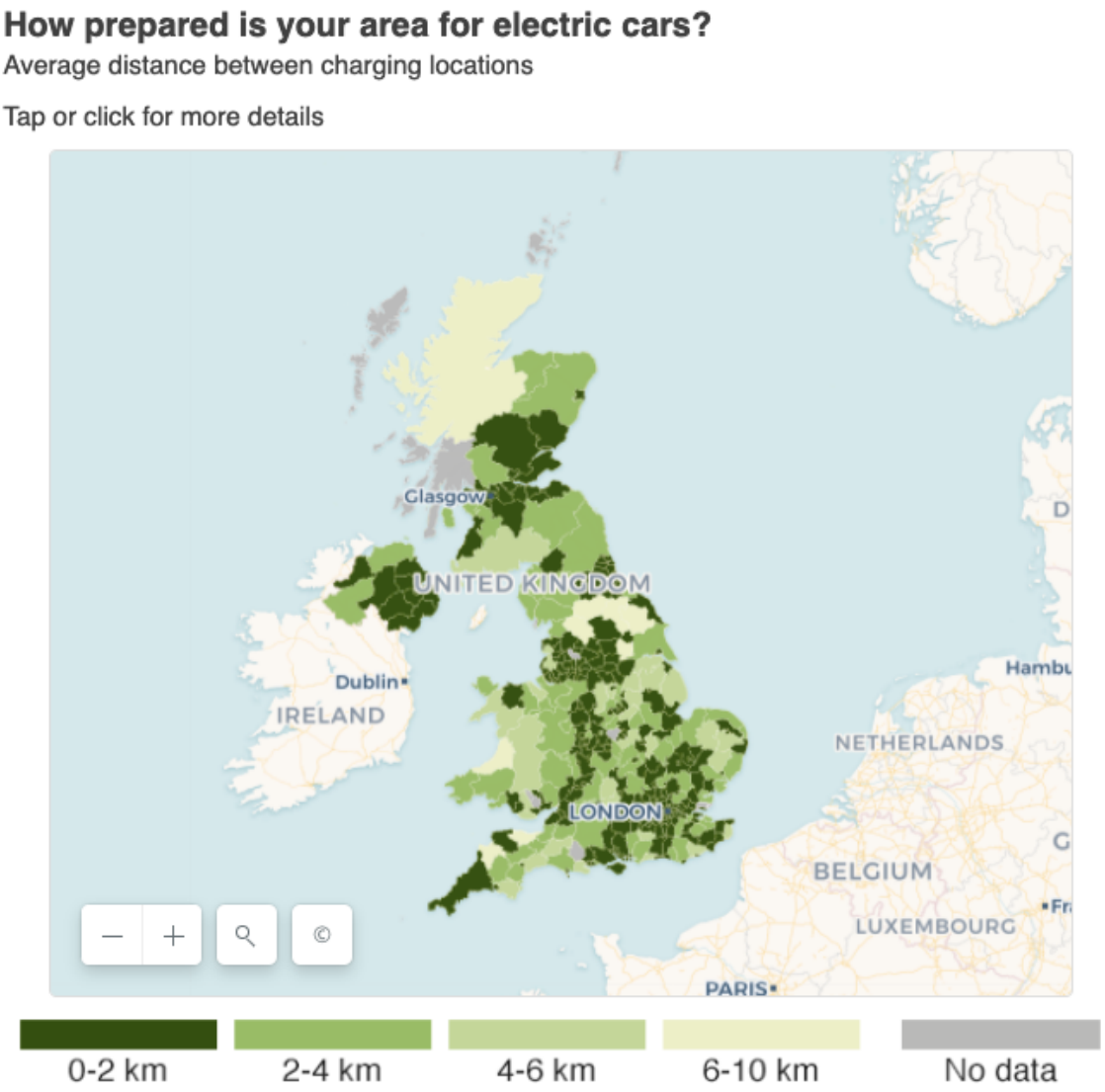
> *"How useful are such indices? Any ranking that isn't built on objective criteria is open to criticism. Qualitative rankings are built on subjective measures. Perhaps "tolerable" means almost the same to someone as "uncomfortable"—whereas "intolerable" might feel twice as bad as "undesirable"? On ordinal scales the distance between these measures is subjective—and yet they have to be assigned a numerical score for the ranking to work. "The Economist has been publishing its Big Mac index, a measure of currency valuations, since 1986. In 2011 we published the Shoe-Thrower's index, which assessed the potential for unrest across the Arab world. And this year, we've created a global normalcy index, which is tracking countries' recovery from covid-19. An imperfect measure is better than having no means of com-*

parison at all."

Ranking stories need to be careful about **context**: an area may have the most crime, disease, or pollution simply because it also has the most *people*. Reporting dates can skew data, too: COVID case rates tended to peak on Tuesdays because the figures "include many deaths not reported over the weekend", as FullFact pointed out.



Data angle 4: Variation — "postcode lotteries", maps and distributions



This BBC Shared Data Unit story by Aimee Stanton focused on the variation in access to electric car charging point

Variation stories are similar to ranking stories — but with a key difference: the story here isn't who or what is top or bo tom. The story leads instead on the very fact that there are wide differences at all. For that reason, this story angle wo best when we expect there to be equal treatment or consistency. The classic example uses a **choropleth map or heatmap** to show how some parts of a country have less access to something, or more demand for something, than other parts. The phrase "postcode lottery", for example, reflects the sense that a person's access to something which i supposed to be equally distributed is actually a game of chance. The BBC data unit story "*IVF: NHS couples 'face social r tioning*"', for example, maps out how where you live in England can mean the difference between being able to access fertility treatment or not. A variation story may be revealing that the unfairness exists — or, if people are aware of it, p cisely how and where it plays out (particularly in their area). **Algorithmic accountability** stories such as ProPublica's Machine Bias series are often about variation and the unfairness that is revealed when an algorithm is unpicked: it ma be people being sentenced differently, or given different insurance quotes, despite no meaningful difference betweer them on the dimensions that matter. A variation story can equally be used to highlight areas of underserved demand, lack of supply: one story that I worked on for the BBC Shared Data Unit about electric car charging points involved ide fying how much infrastructure existed in the country, and where. The picture that the data painted provided a founda tion for case studies and reaction. *In the second part of this series I look at the other three angles: exploratory stories, thos that focus on data quality, existence or absence, and angles about relationships. A version of the diagram is also available ir Finnish.*

This entry was posted in data journalism, online journalism and tagged change, data journalism, margin of error, no data, Ranking, scale on August 11, 2020 [https://onlinejournalismblog.com/2020/08/11/here-are-the-7-types-of-stories most-often-found-in-data/] by Paul Bradshaw.

---

8 thoughts on “Here are the angles journalists use most often to tell the stories in data”

Pingback: Here are the angles journalists use most often to tell the stories in data – Il blog del corso di giornalismo

Pingback: 3 more angles most often used to tell data stories: explorers, relationships and bad data stories | Online Journalism Blog

Pingback: 3 more angles most often used to tell data stories: explorers, relationships and bad data stories – Totally ne

Pingback: 3 more angles most often used to tell data stories: explorers, relationships and bad data stories – BUZZINCLICK

Pingback: Veri Gazeteciliği Projelerinde 7 Ortak Açı - Açık Veri ve Veri Gazeteciliği Türkiye

Pingback: Online Journalism Blog: Here are the angles journalists use most often to tell the stories in data | ResearchBuzz: Firehose

Pingback: Crowded Planet, Instagram, Language Learning, More: Monday Afternoon ResearchBuzz, August 24, 2020 – ResearchBuzz

Pingback: New Here are the angles journalists use most often to tell the stories in data – Stephen's Lighthouse