



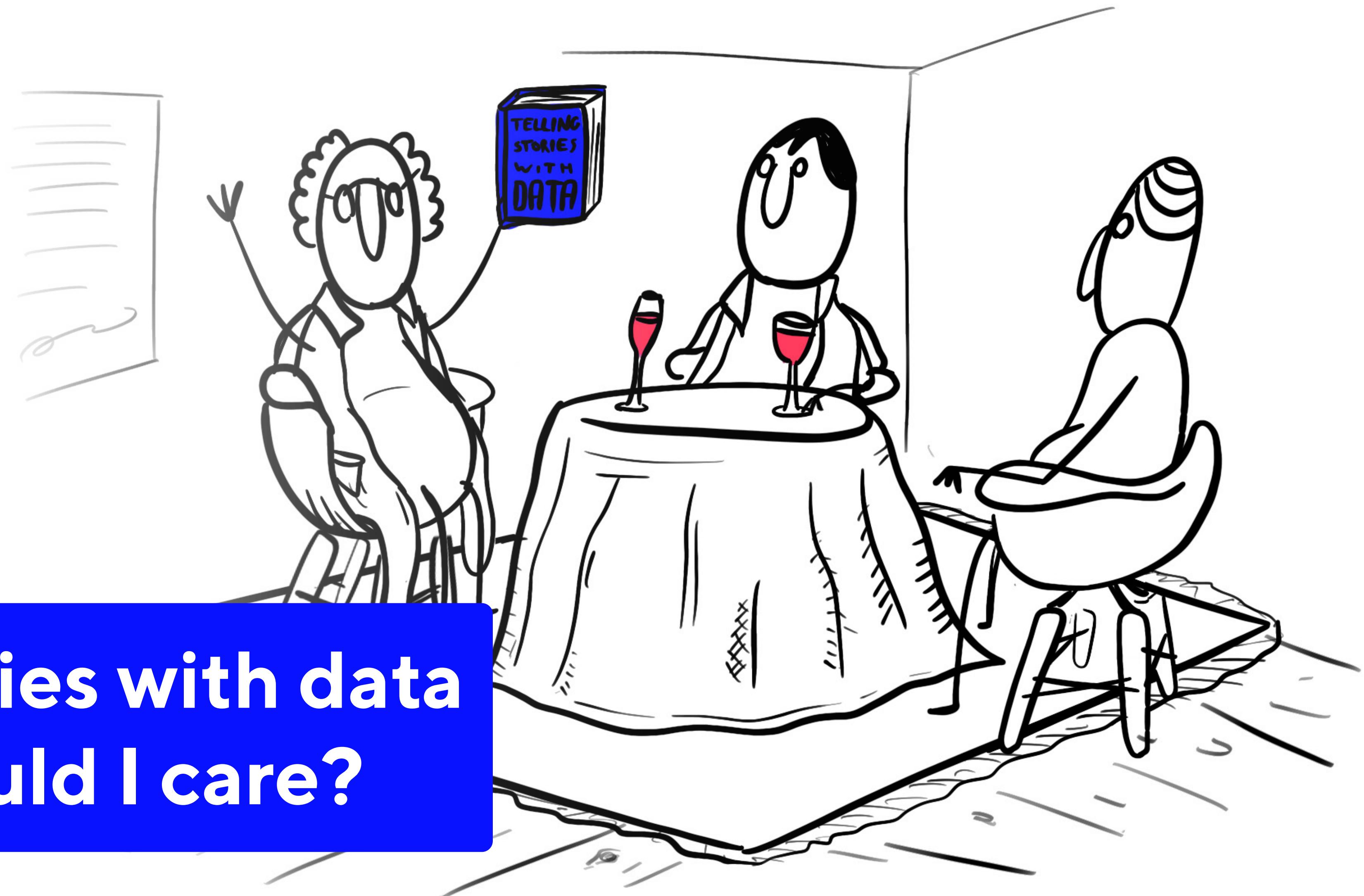
How to tell your story with data.

The Storyteller by ParCos

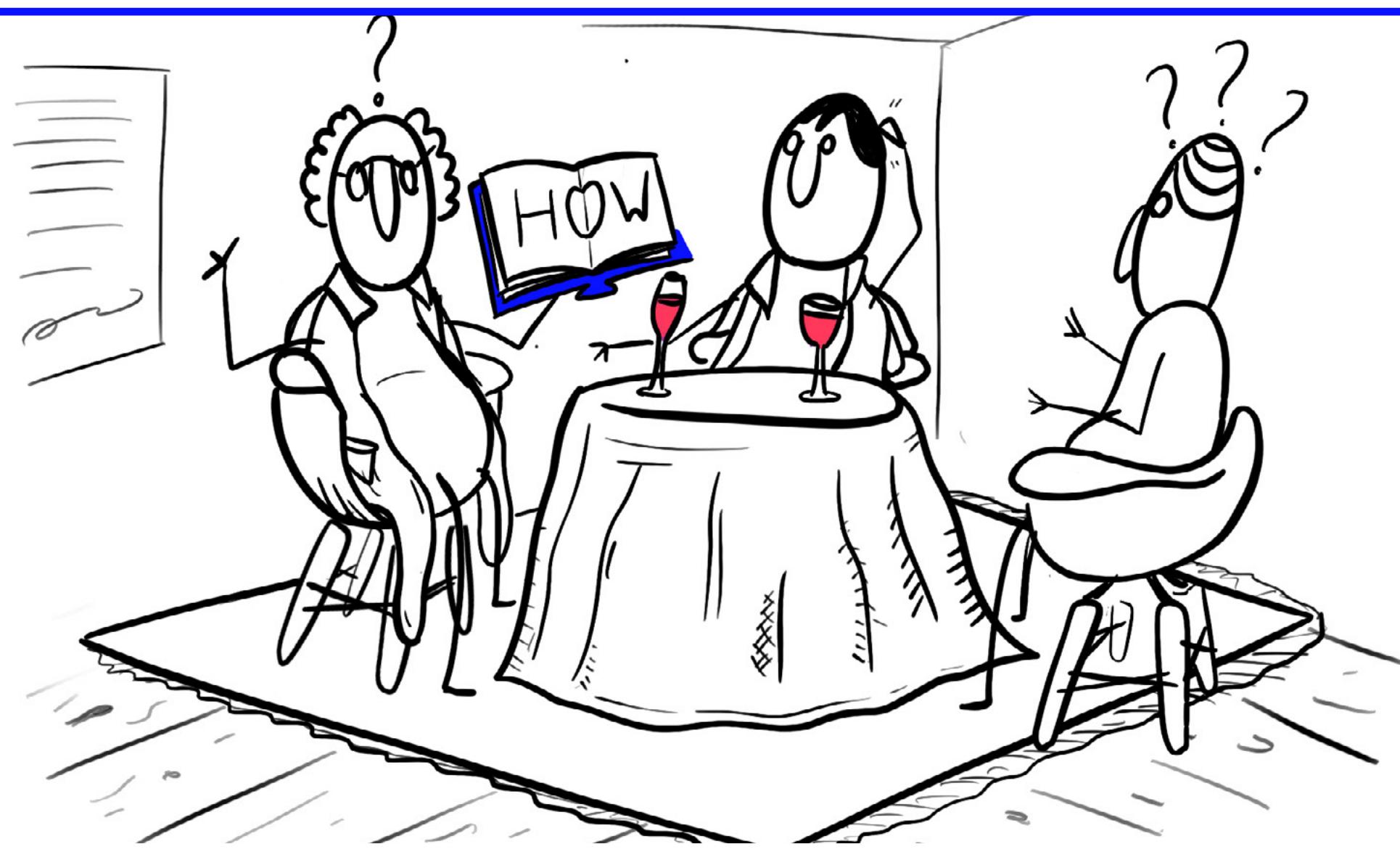
Table of contents:

Why use data in storytelling	1 – 3
Creating a data story in 5 steps	4 – 22
1. Finding the data	4 – 5
2. Finding the story	6 – 8
3. Developing the narrative	9 – 12
4. Presenting the data story	13 – 17
5. Involving the reader	18 – 22

Telling Stories with data — Why should I care?

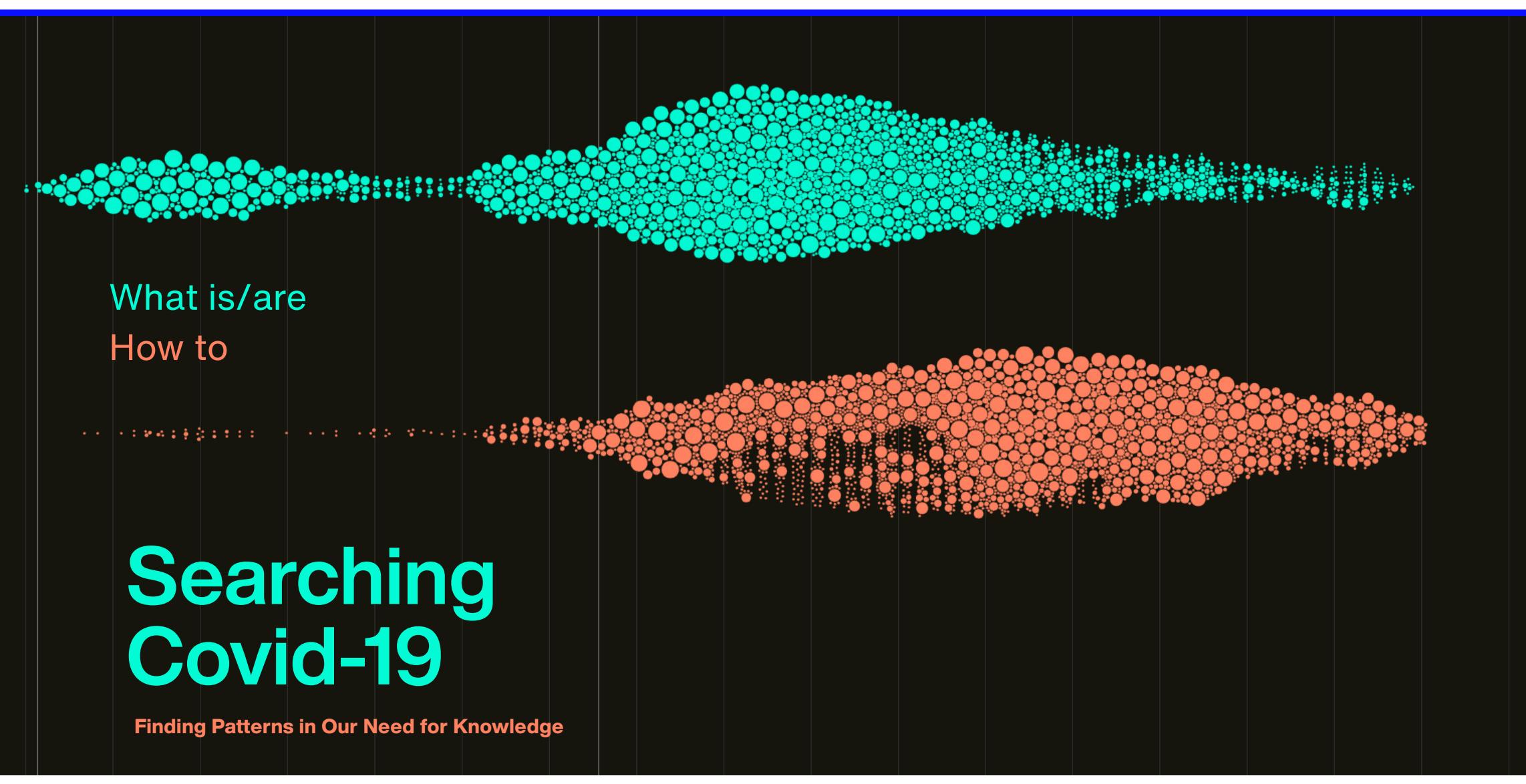


Source:
Meijer, I.C. (2013) WHEN NEWS HURTS, Journalism Studies, 14:1, 13-8, DOI:10.1080/1461670X.2012.662398



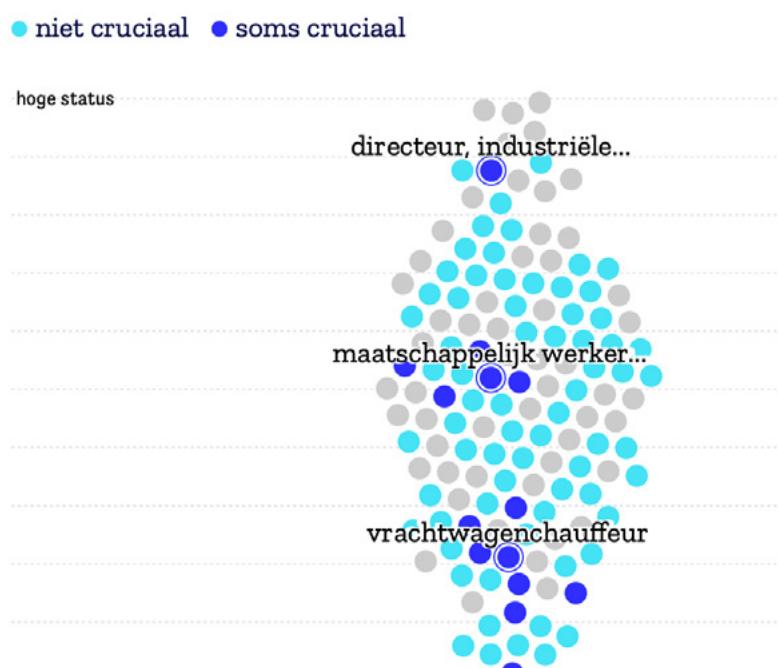
**It's a new way to tackle misinformation.
+ it supports data literacy skills**

Source:
Wolff, A., Gooch, D., Montaner, J. J. C., Rashid, U., & Kortuem, G. (2016). Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics, 12(3)*.



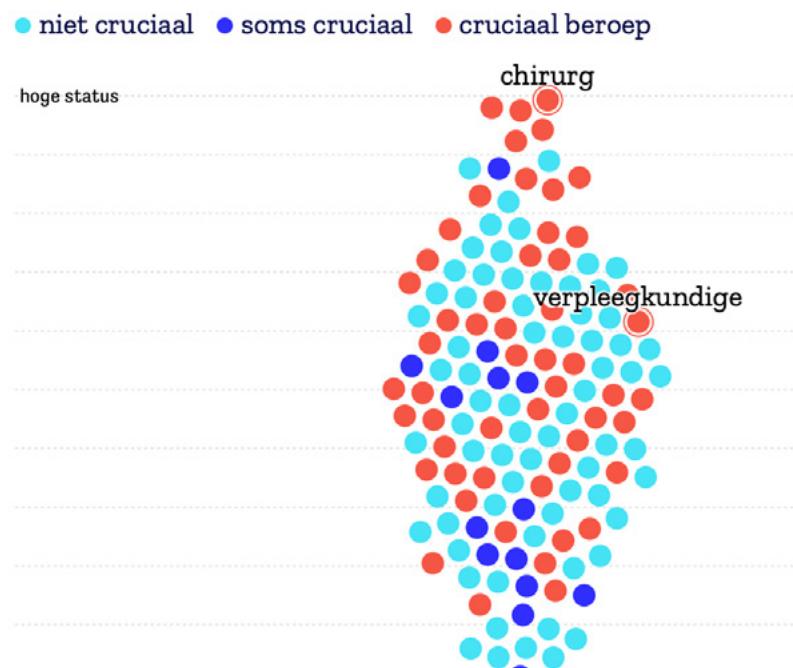
You can build trust in the presented story by revealing the underlying data.

Source:
searchingcovid19.com - Finding Patterns in Our Need for Knowledge (<https://searchingcovid19.com>)



2. Soms cruciale beroepen

Bij sommige beroepen hangt het ervan af in welke sector zij werken.



3. Cruciale beroepen

Hier zie je de cruciale beroepen van vuilnisophaler, waarvan de status als laagste ingeschat wordt naar verpleegkundige en tot aan chirurg, met de hoogste status.

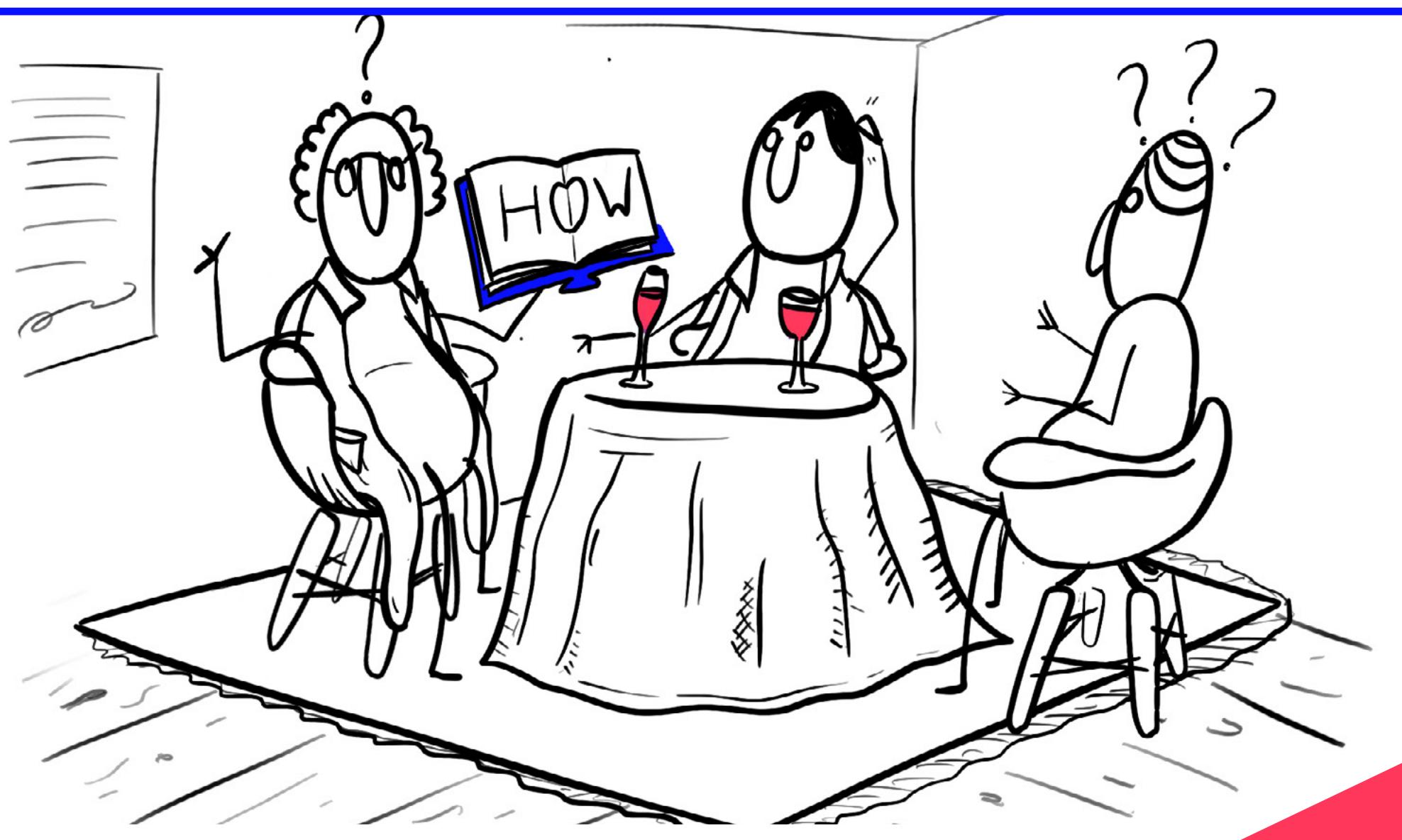
Example:
<https://pointer.kro-nrcv.nl/artikelen/cruciale-beroepen-zijn-onmisbaar-in-corona-tijd-maar-de-onderlinge-salaris-verschillen>

By presenting the main story in text, you can allow readers to further explore data for themselves.

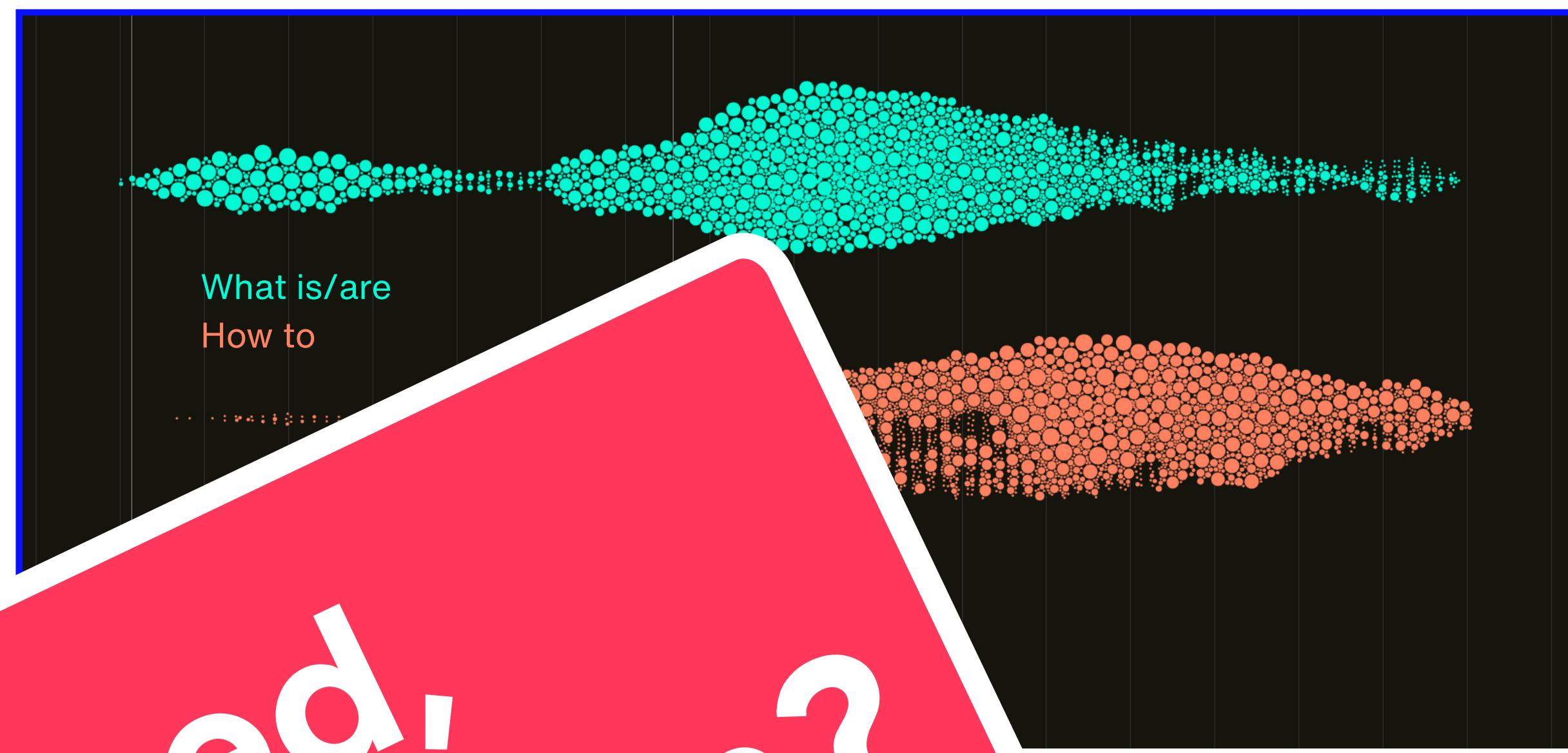


Adding data representations to stories contributes to data literacy skills by making data feel familiar.

Source:
<https://www.newscaststudio.com/2019/04/09/weather-channel-climate-change-imr/>



**It's a new way to tackle misinformation.
+ it supports data literacy skills**



g
Source:
searchingcovid19.com - Finding Patterns in Our Need for Knowledge (<https://searchingcovid19.com>)

I'm convinced, Where do I begin?

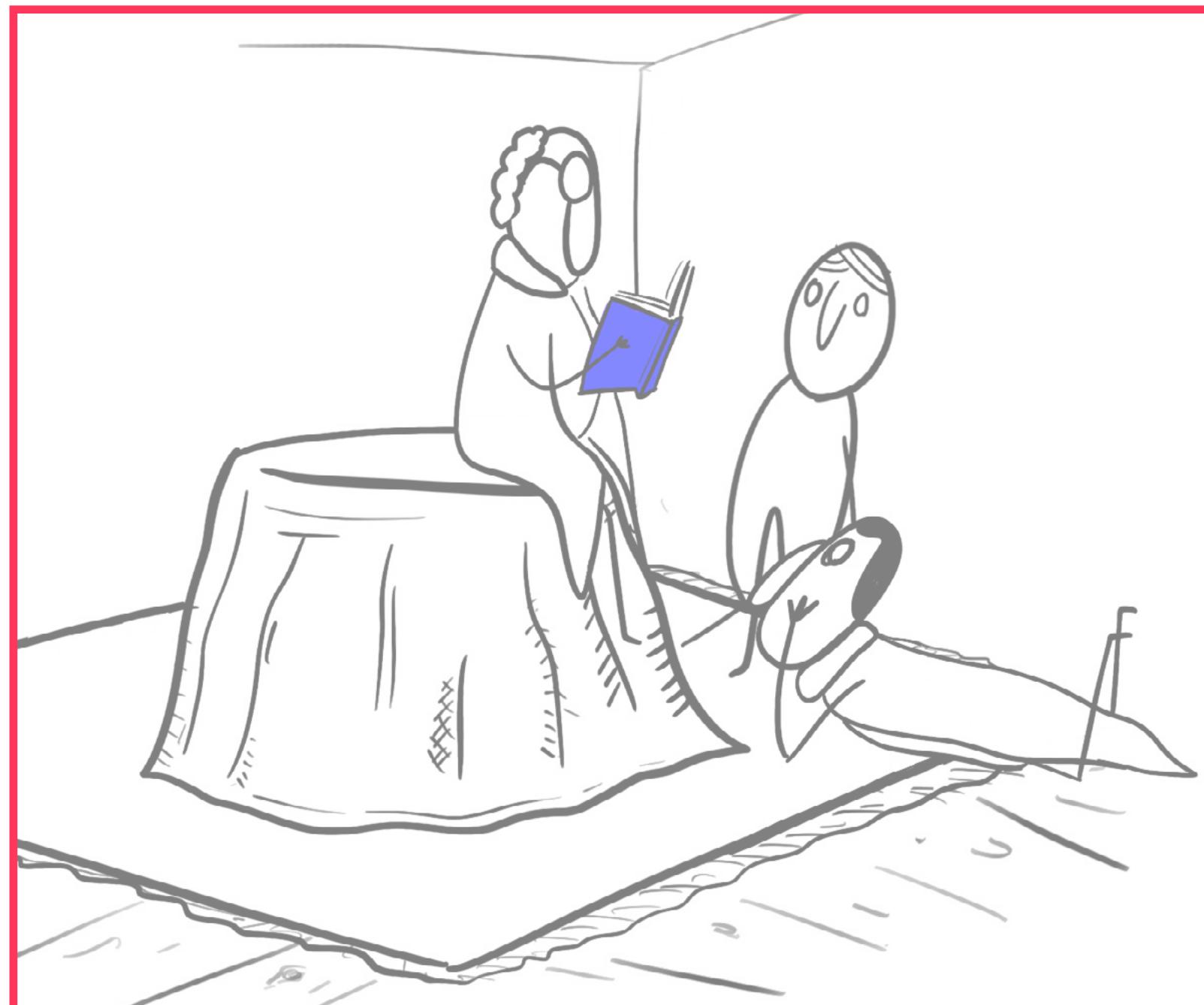
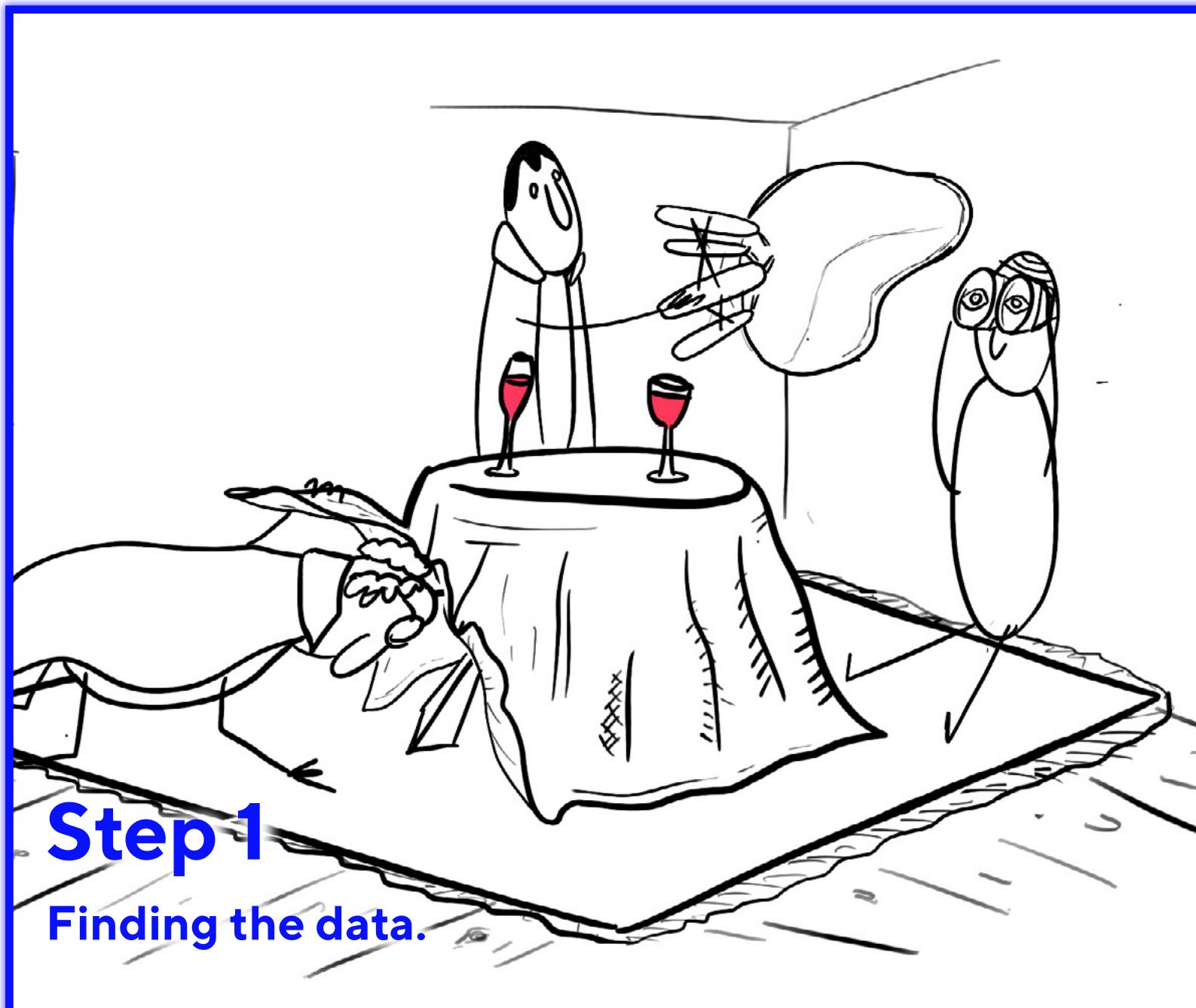


By presenting the main story in text, you can allow readers to further explore data for themselves.

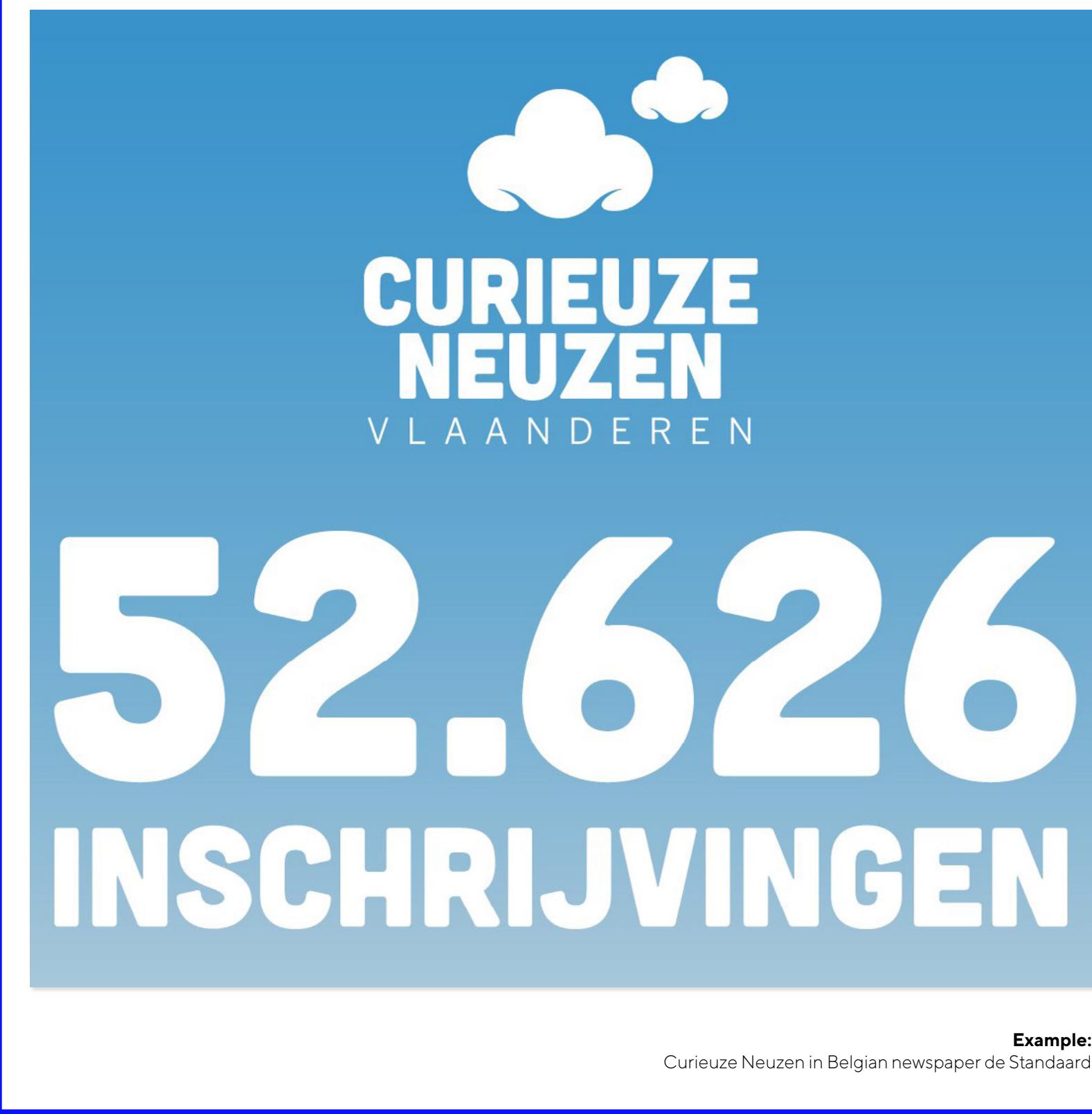


Adding data representations to stories contributes to data literacy skills by making data feel familiar.

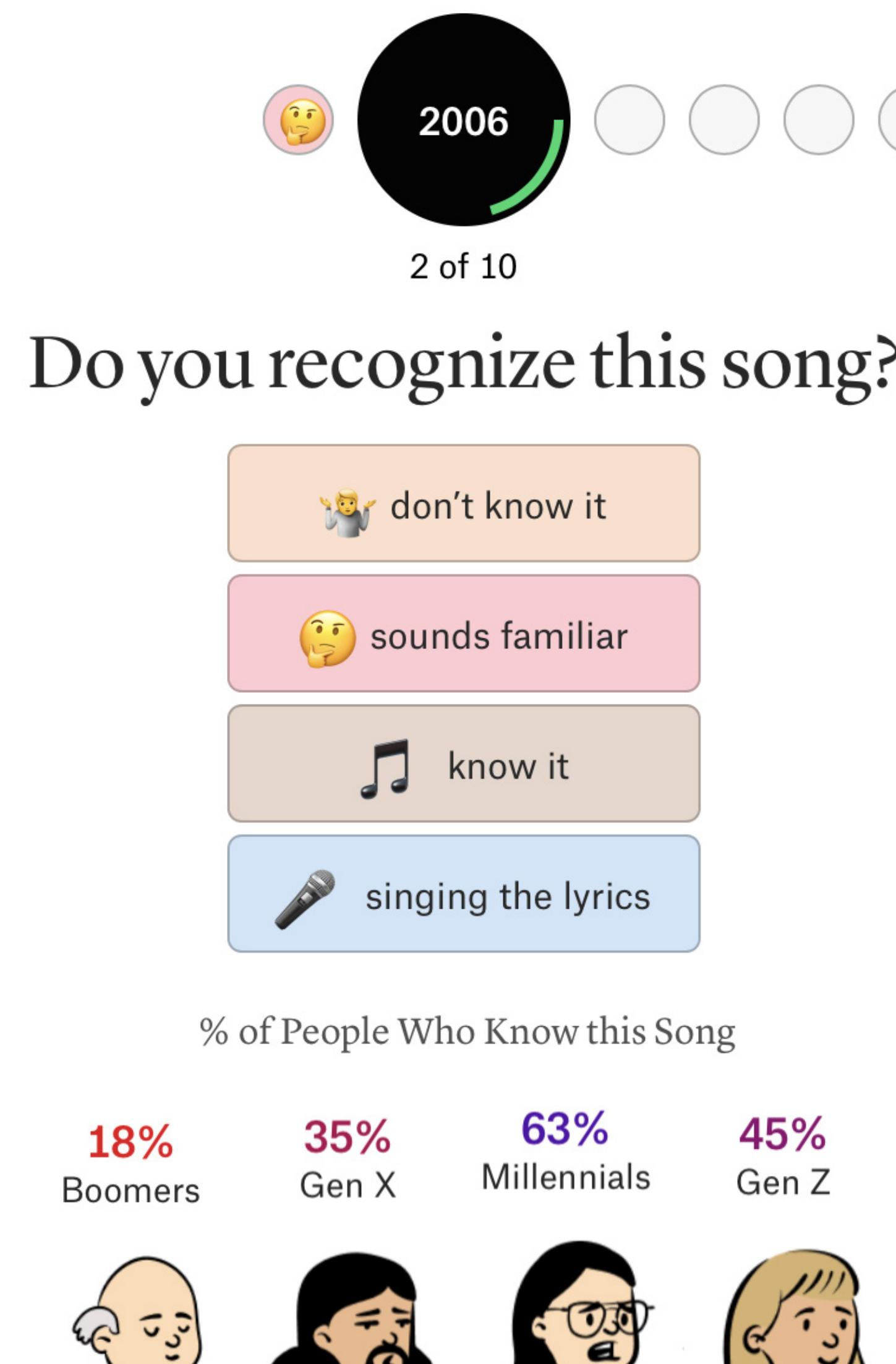
Source:
<https://www.newscaststudio.com/2019/04/09/weather-channel-climate-change-imr/>



Besides exploring existing data sources, you may also ask the audience to contribute data!
Collecting citizen science data to inform the story.



The collected data does not always need to be about serious topics!



Don't forget to share those data sources!
Plus how the data is collected (methods).

Data and Methods

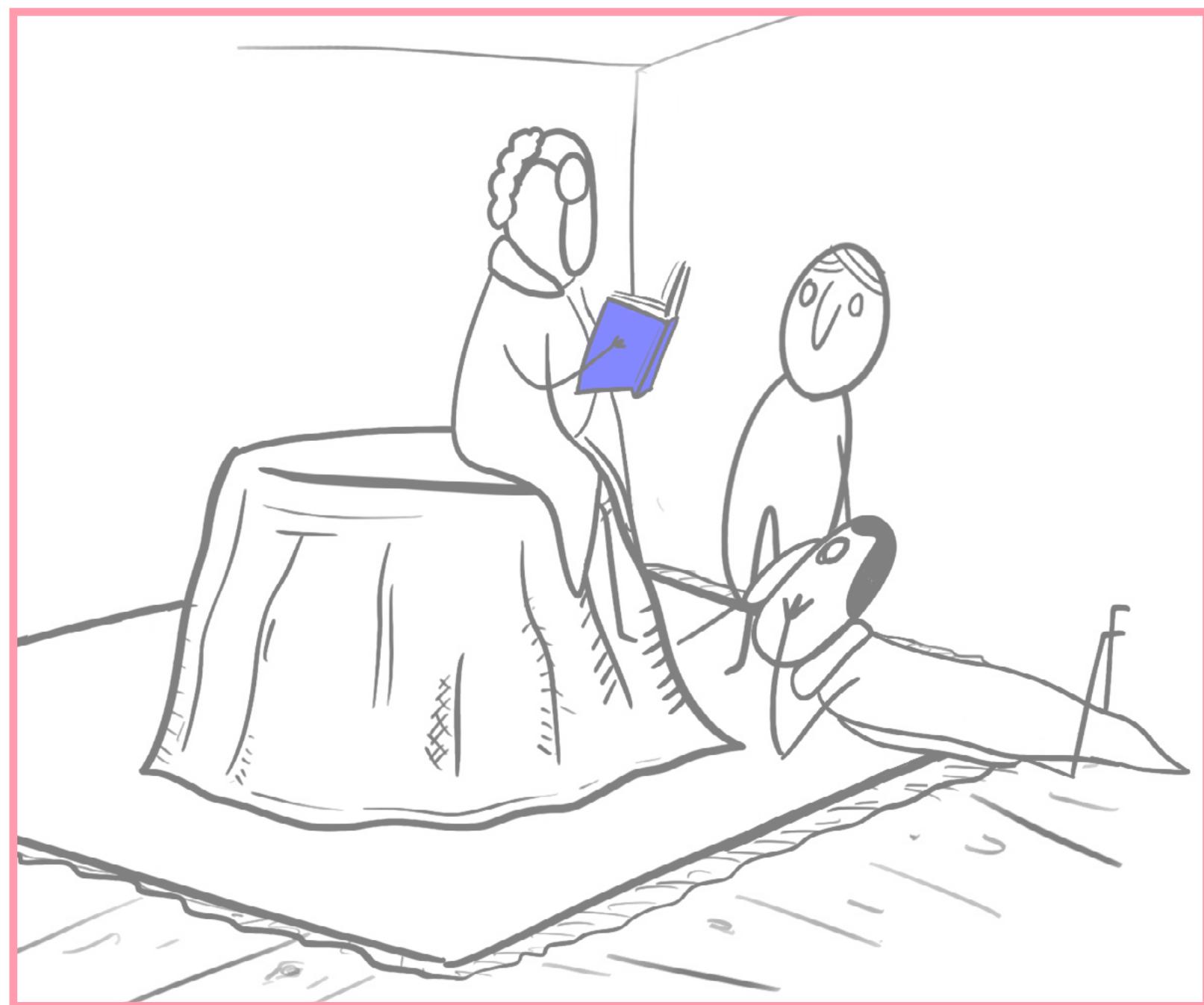
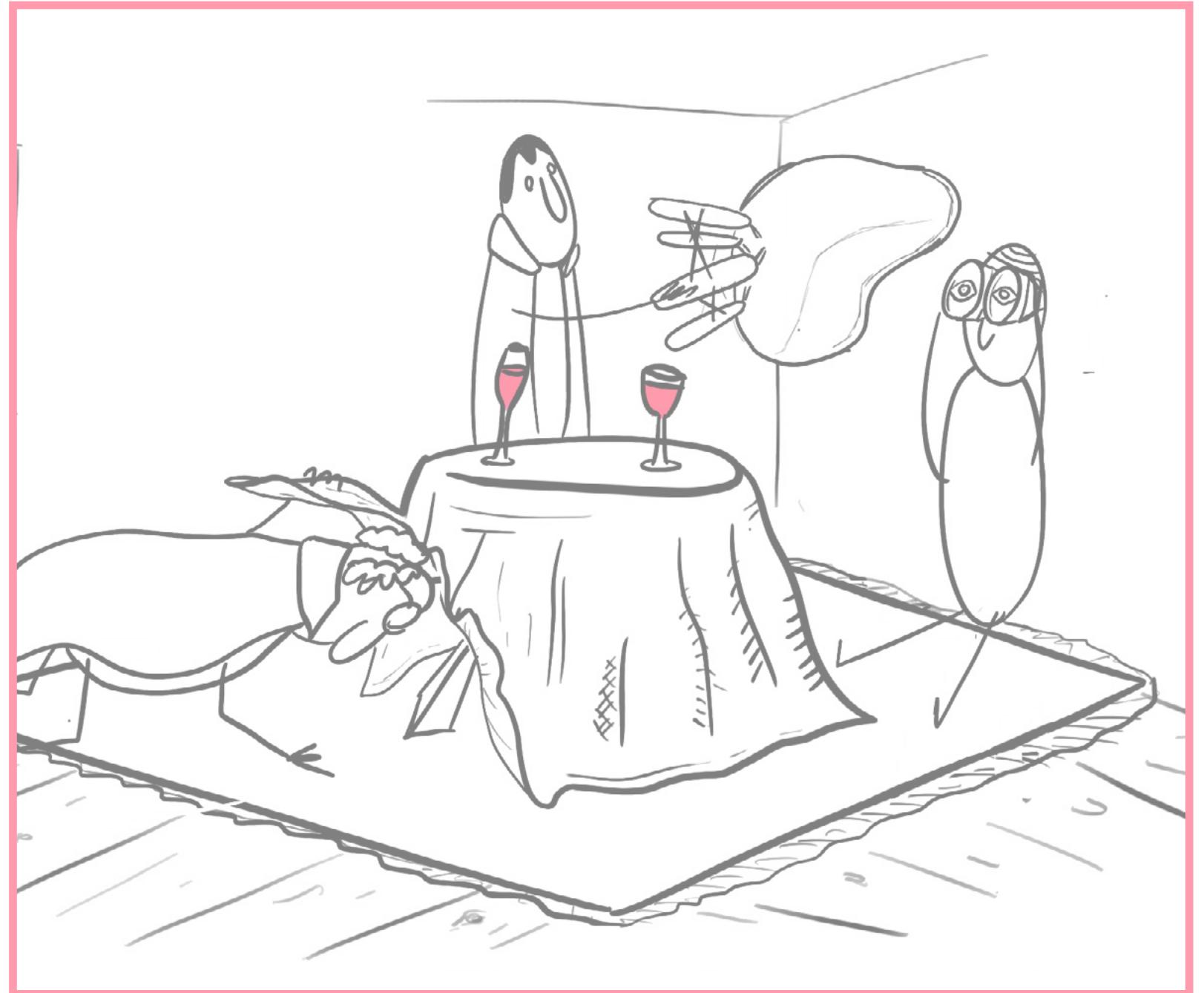
Data for this story were collected and processed using the Wikipedia API. The period of collection was from July 1, 2015–September 13, 2018, from English Wikipedia. Any person who appeared in the top 1,000 pages for at least one day in that range was considered. The full source is on [Github](#).

Wikipedia's aggregation of notable births was the starting point to decide who is thought to be a celebrity or not, a foundation of over 40,000 people. Additionally, each top 1,000 page with "born" in the text, which is a consistent characteristic of people pages on Wikipedia, was also added to this database of people, to ensure no one not yet notable enough to be added to the births pages was still included.

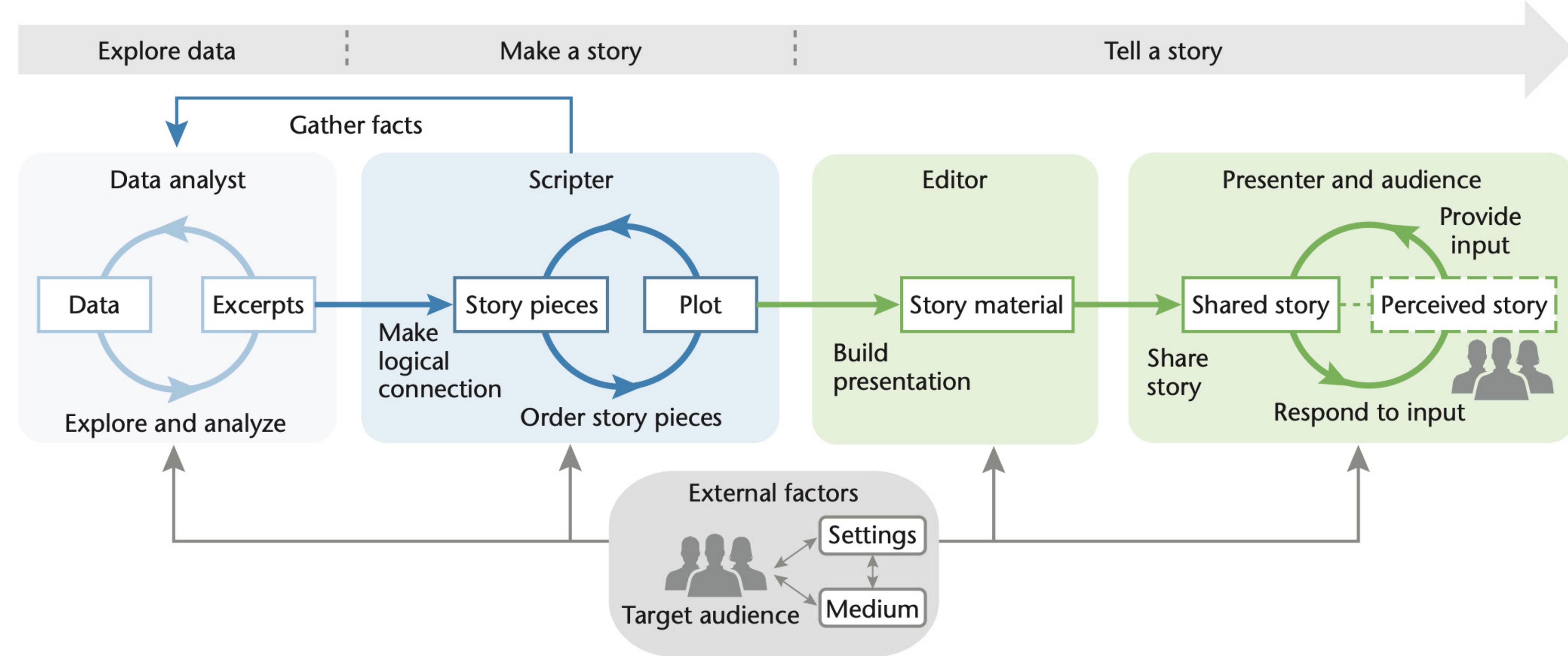
We started with those who had little to no pageviews in the second half of 2015, eliminating already known celebrities. The methodology to define rising celebrities was centered on a series of levels of sustained pageviews. Levels were assigned to based on monthly averages. There were eight levels (like Karate belts): (1) 50, (2) 100, (3) 200, (4) 500, (5) 1,000, (6) 2,000, (7) 5,000, and (8) 10,000, pageviews. If someone hit a new level of pageviews and never dipped below that level's threshold again, they were assigned the level, hence the term "sustained pageviews." If a person hit level 5 (1,000 pageviews), for example, but then dropped below 1,000 pageviews the following month, they would still be a level 4.

After assigning levels, anyone with 1. a beginning level lower than 4, 2. a level change of more than 4 levels, and 3. less than level 2 in 2015, was included in the final list. People above level 6 were considered those who have risen to fame. Anyone who satisfied those parameters but still remained below level 6 was considered rising.

By [Russell Goldenberg](#) and [Caitlyn Ralph](#). For questions, comments, etc., sup@pudding.cool.



It is a non-linear process!

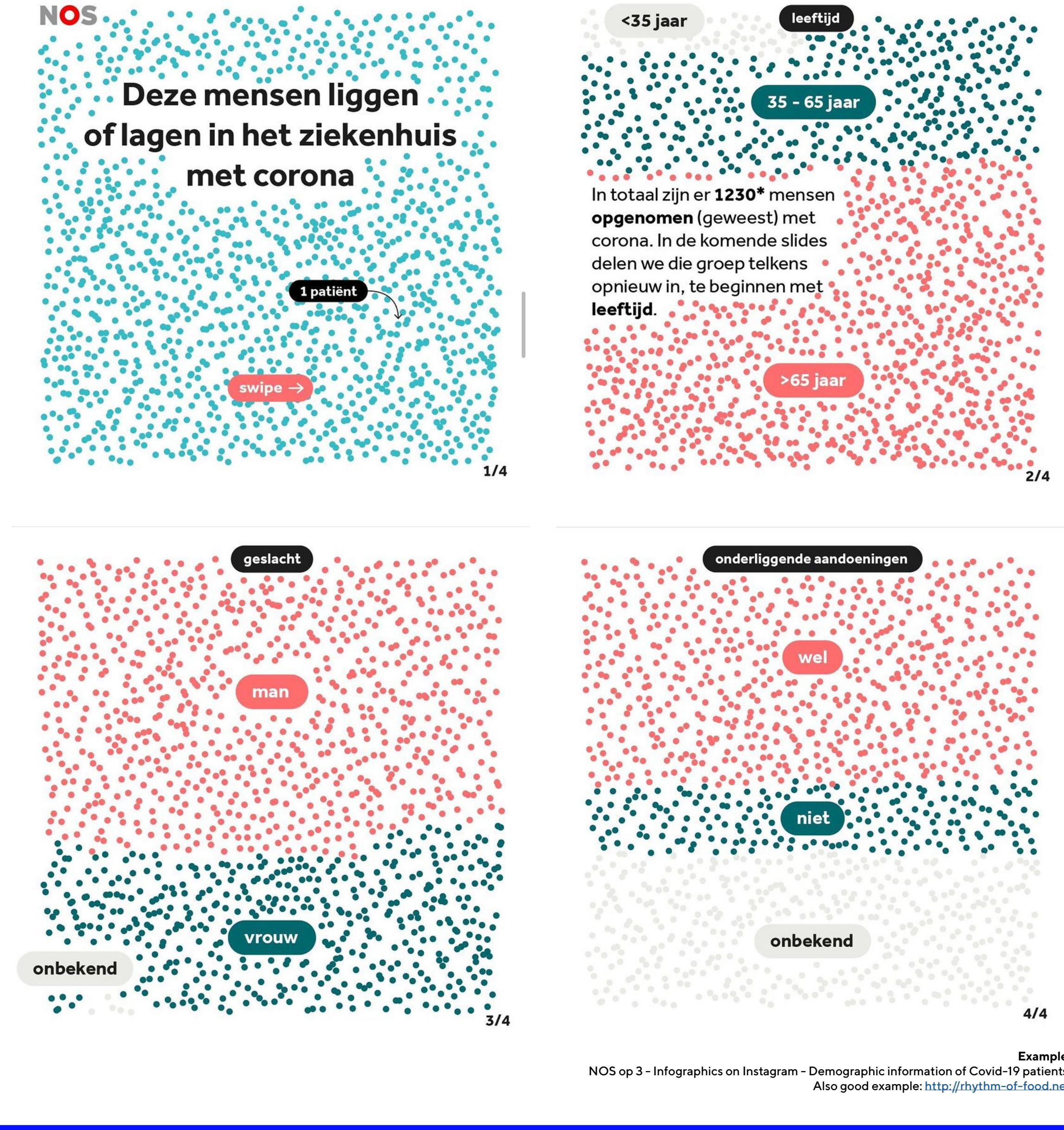


The data may be the starting point:
> Analyze to find insights.

The story may be the starting point:
> Find data to support the narrative insights.

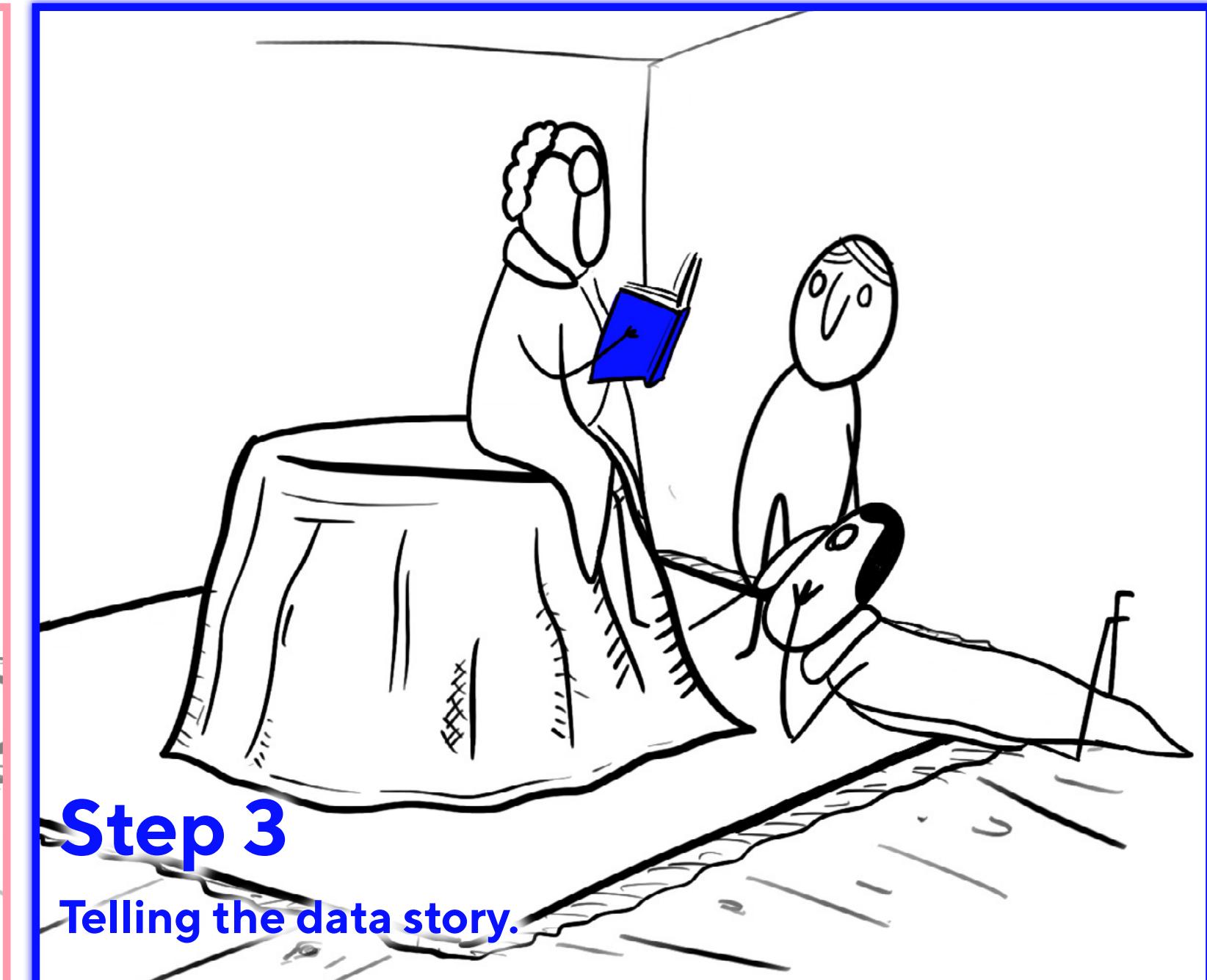
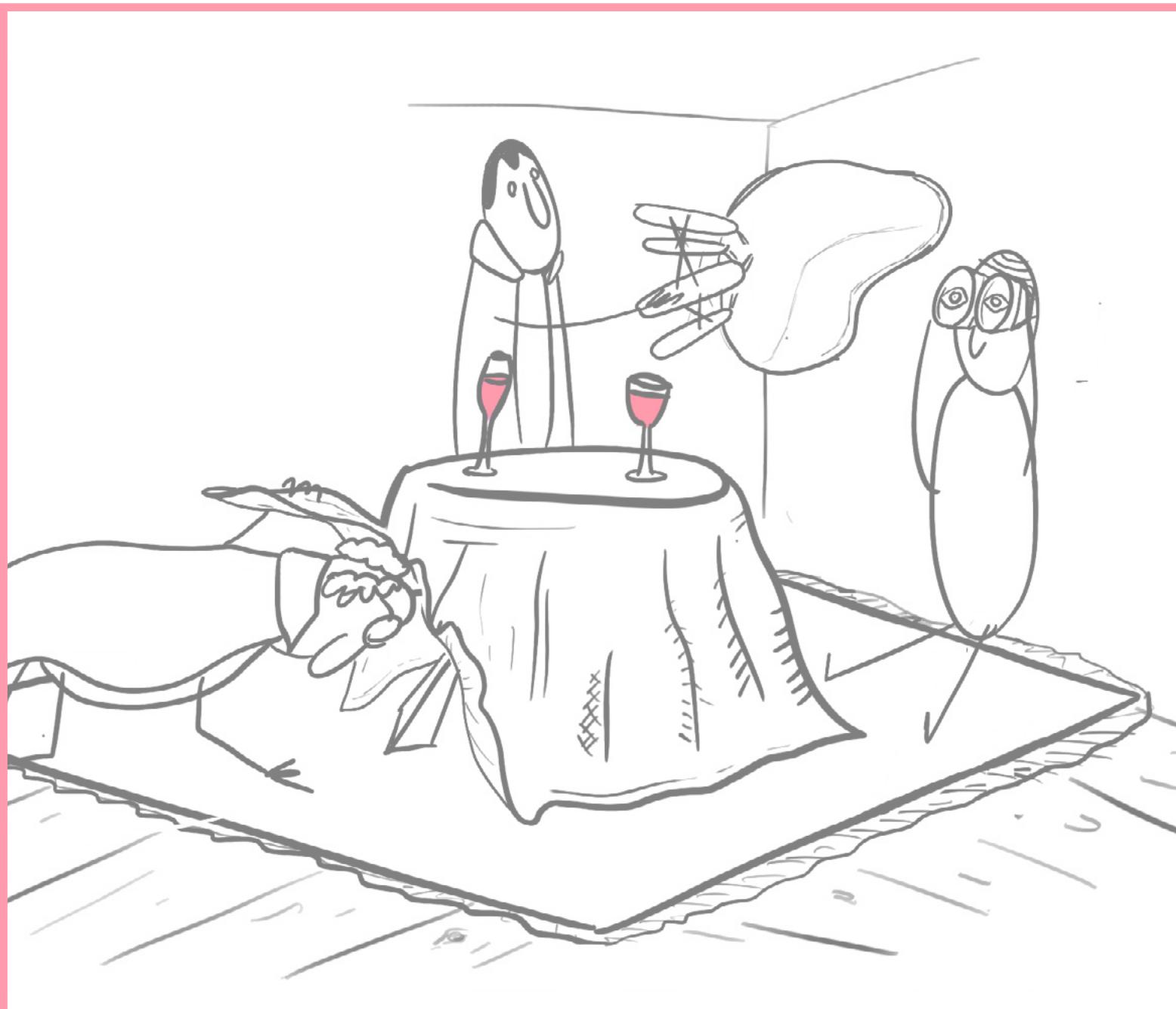
Source:
More than Telling a Story: A Closer Look at the Process of Transforming
Data into Visually Shared Stories - Bongshin Lee

The main story may already be set out, and data may be used to support it, which is often the case in stories that use citizen science. As citizen science data offers both quantitative data and qualitative data in the form of anecdotes, quotes, etc.

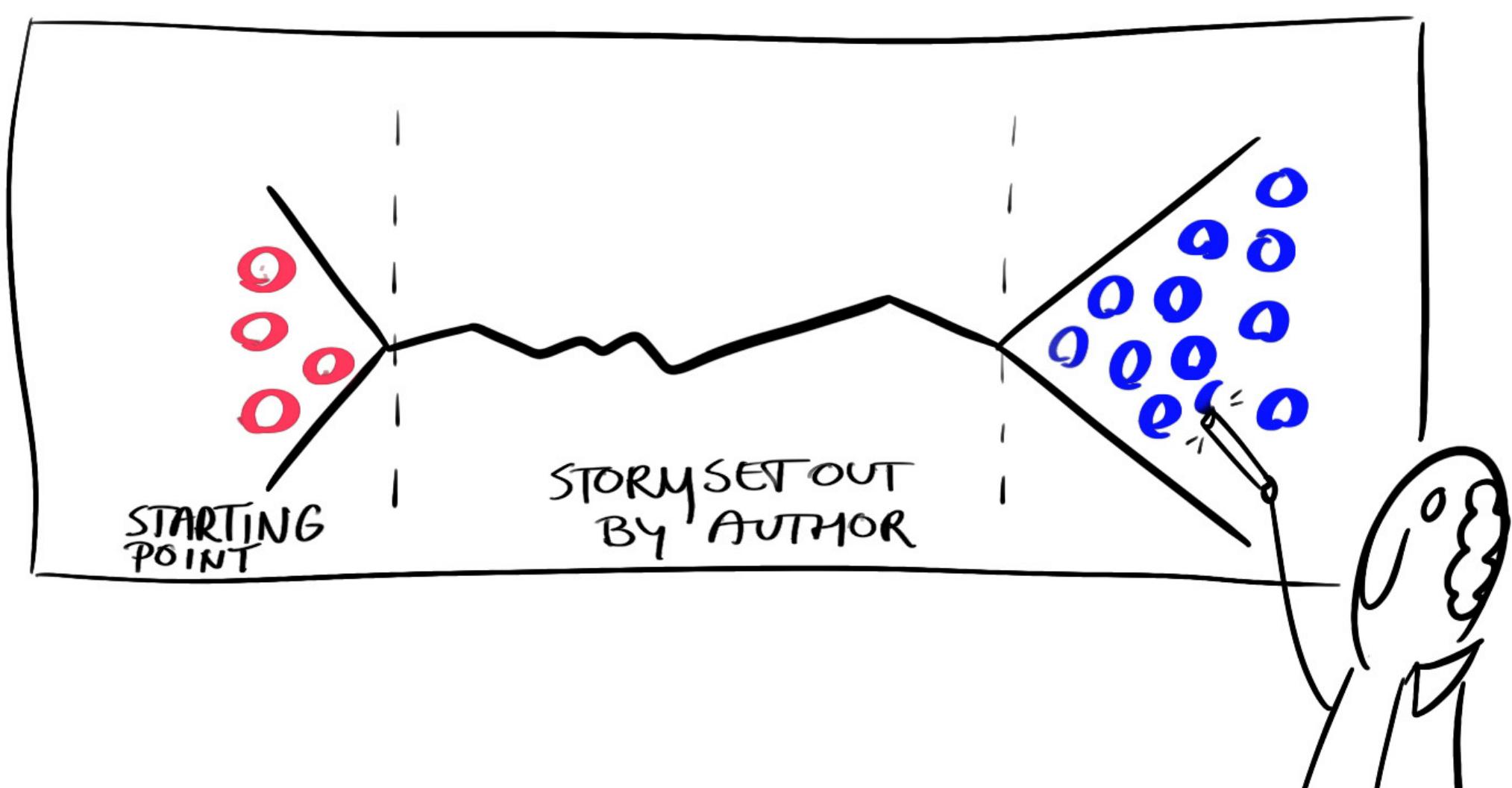


Data may be the starting point: the story may reveal itself when analyzing and visualizing the data. This example is seemingly simple yet it incorporates narrative elements as conflict (*in the readers' expectation*) and empathy (*as you can identify with it*).





Martini Glass (on its back): Scrollytelling examples are typical for the Martini Glass structure:



This Martini glass structure is most often used for data stories.
Starting with a story that is fixed, scrolling down to the point that the reader may interact freely with the data.

Immo Royal

In België kan je meer dan 7.500 hectare aan koninklijk vastgoed terugvinden. Journalisten voerden voor het eerst in 90 jaar diepgaand onderzoek naar het patrimonium van de Koninklijke Schenking.

DE TIJD apache Knack

BEGIN TE LEZEN

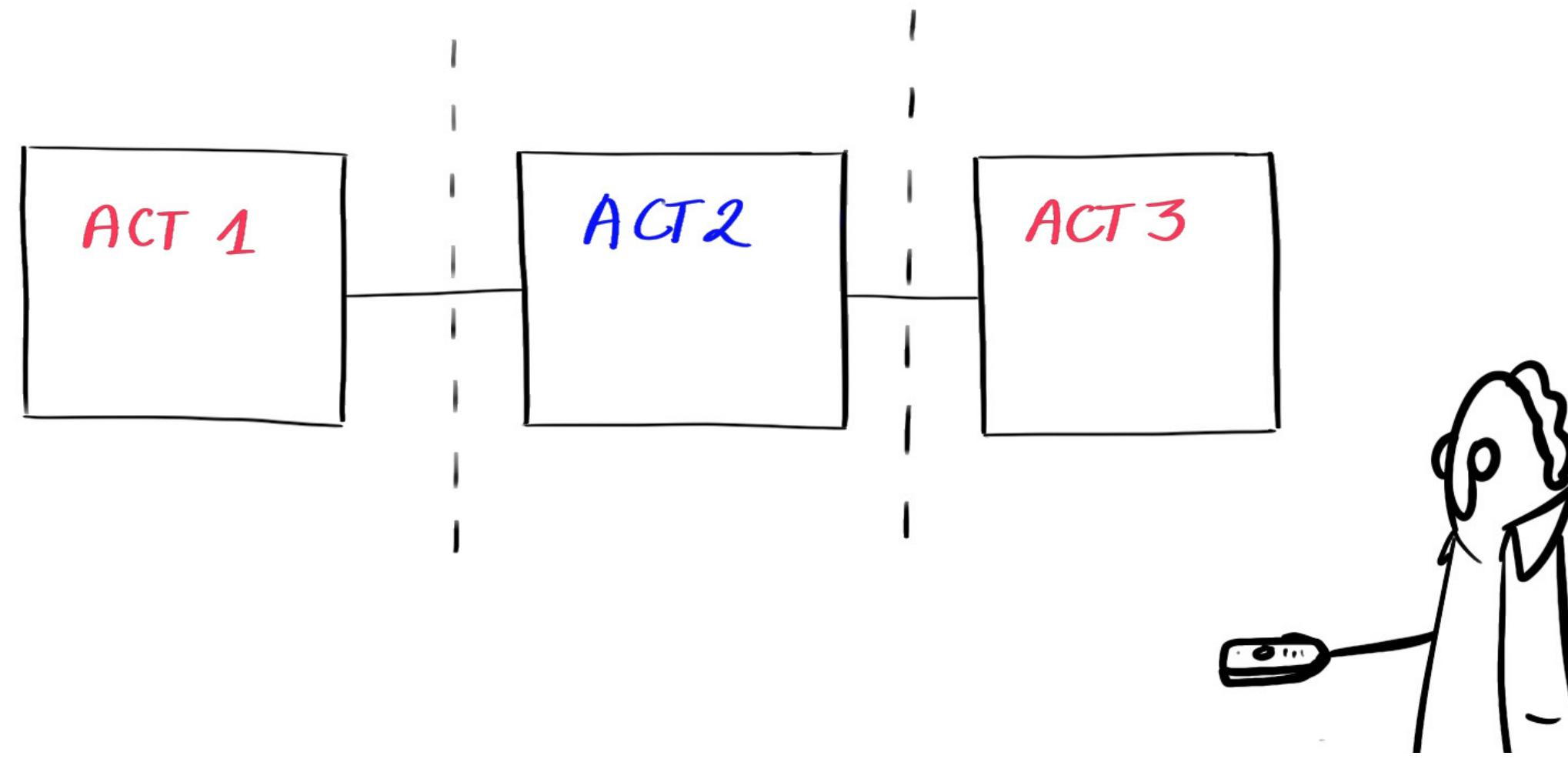
Example:
koningshuizen.be (Immo Royal) about the properties of Belgian monarchies

Similar examples:
<https://www.juiceanalytics.com/writing/20-best-data-storytelling-examples>



The Slideshow:

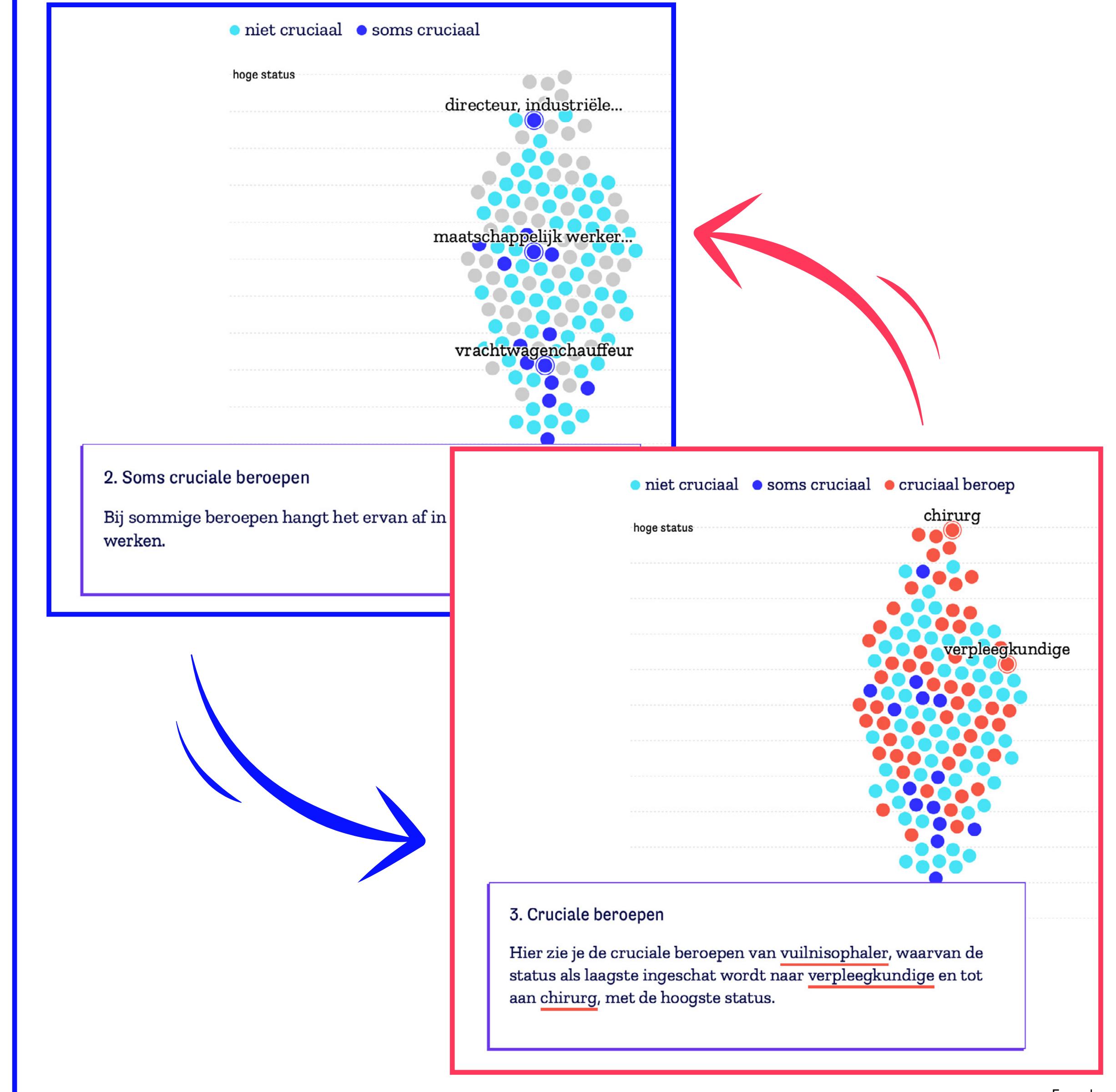
A second type of narrative structure for data stories is meant for more complex stories



It allows data to be exposed step-by-step.

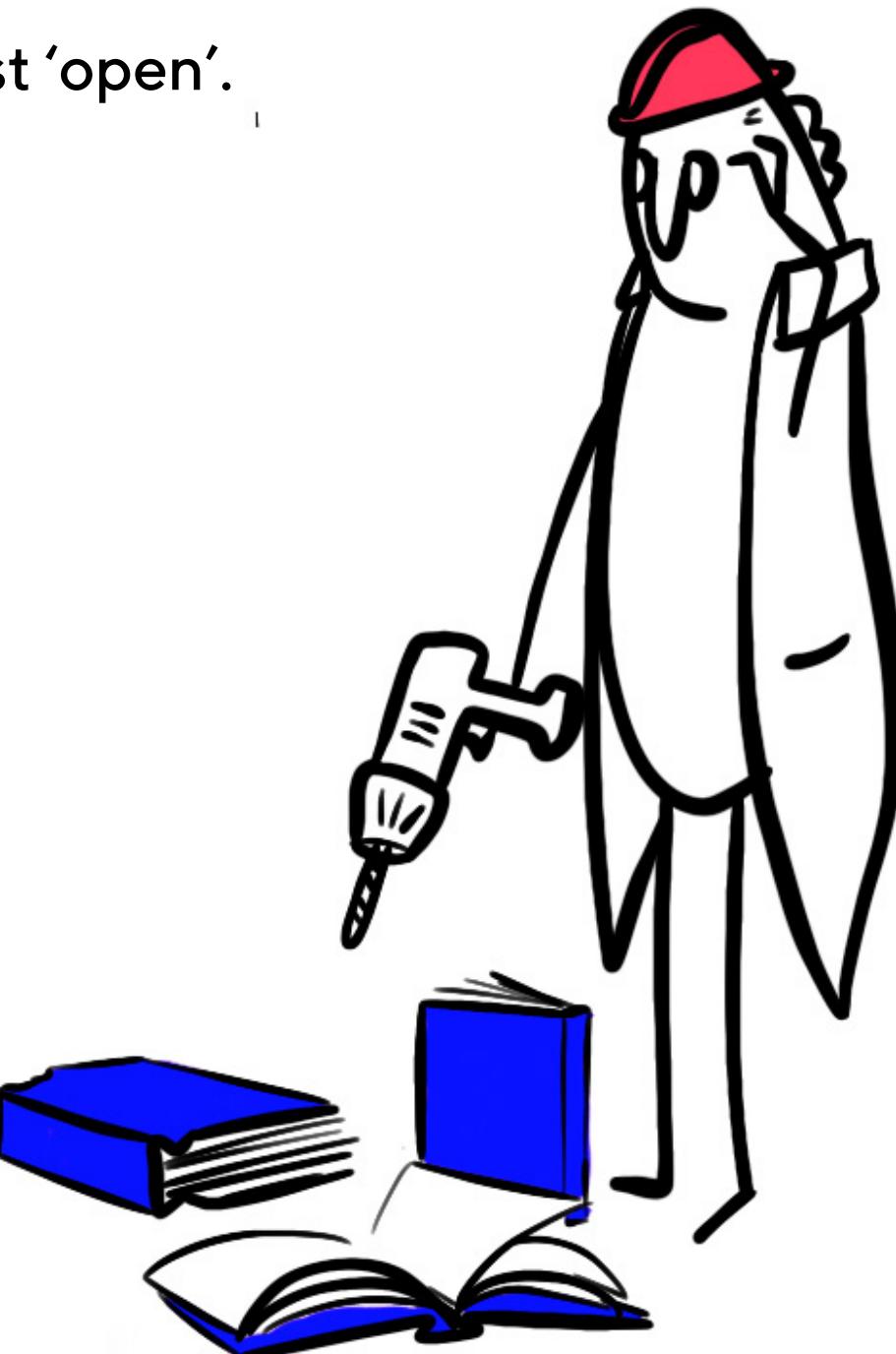
The reader can choose to continue, go back or select something else.

This example repeatedly switches between textual story elements with interactive plot charts that allow the reader to further explore



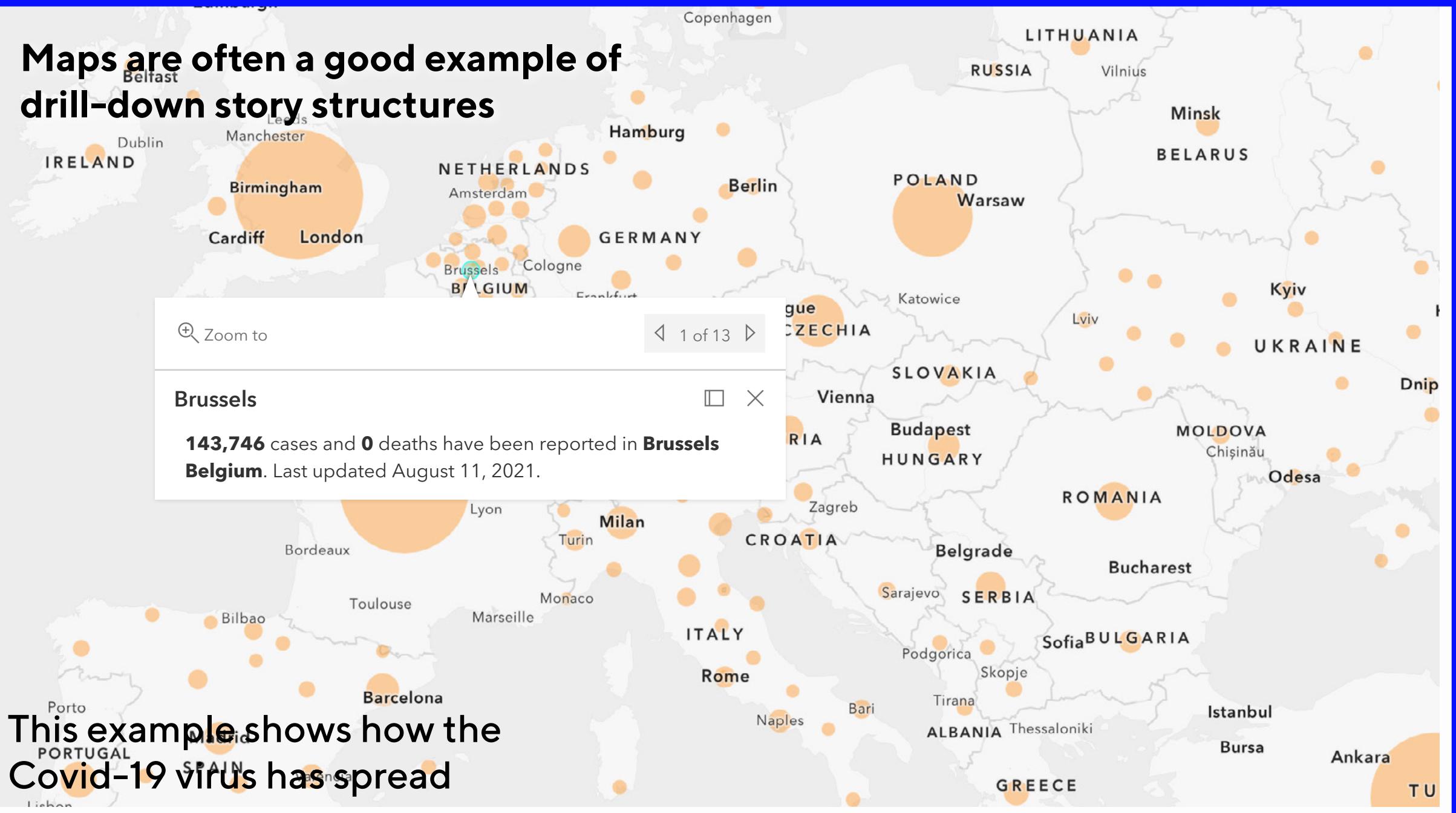
Example:
<https://pointer.kro-nrcv.nl/artikelen/cruciale-beroepen-zijn-onmisbaar-in-corona-tijd-maar-de-onderlinge-salarisverschillen>

Drill-down:
This third structure is most 'open'.



*Readers may interact at any point in the story.
Participants can zoom in, ask details, explore, etc.*

Maps are often a good example of drill-down story structures

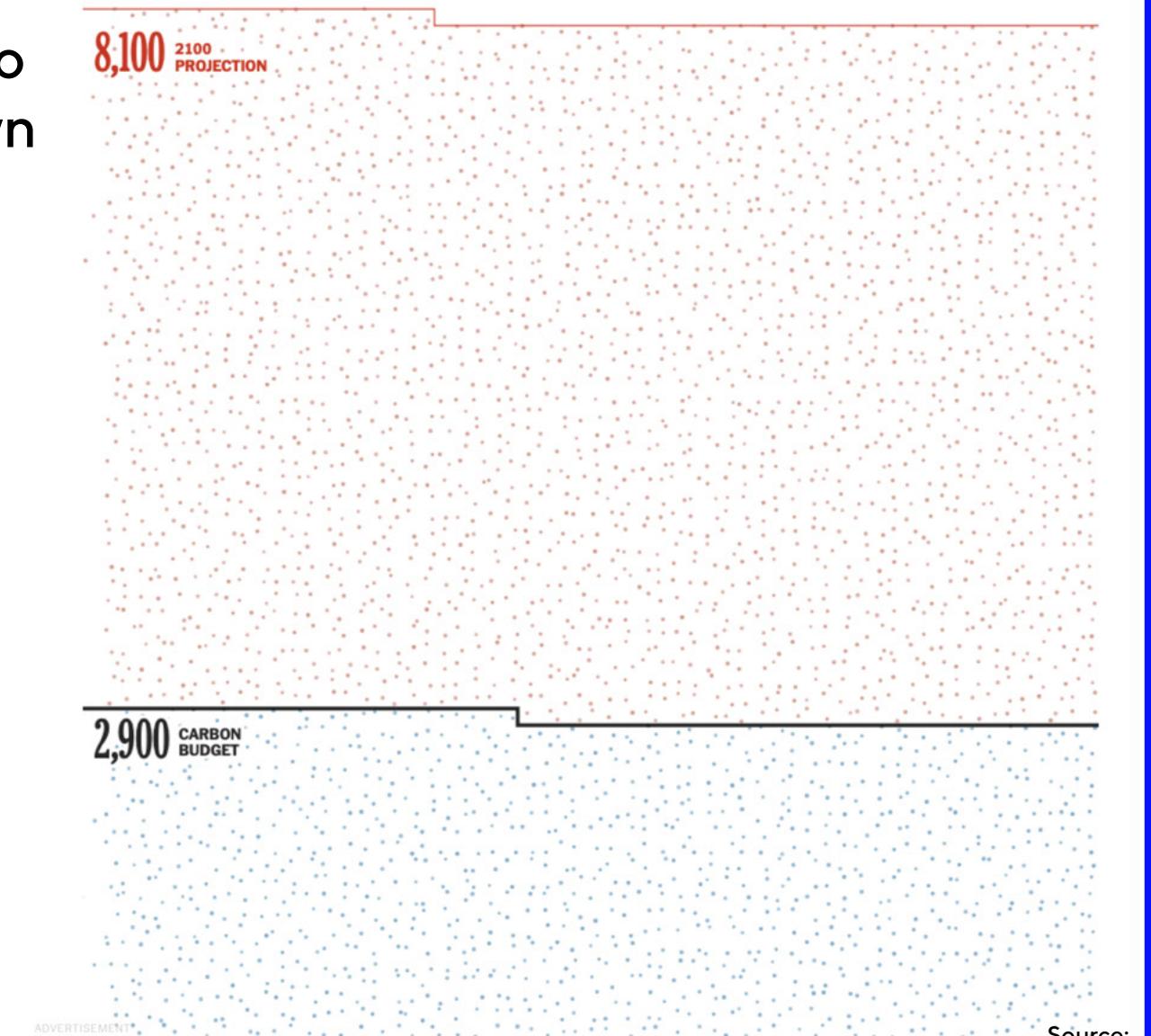


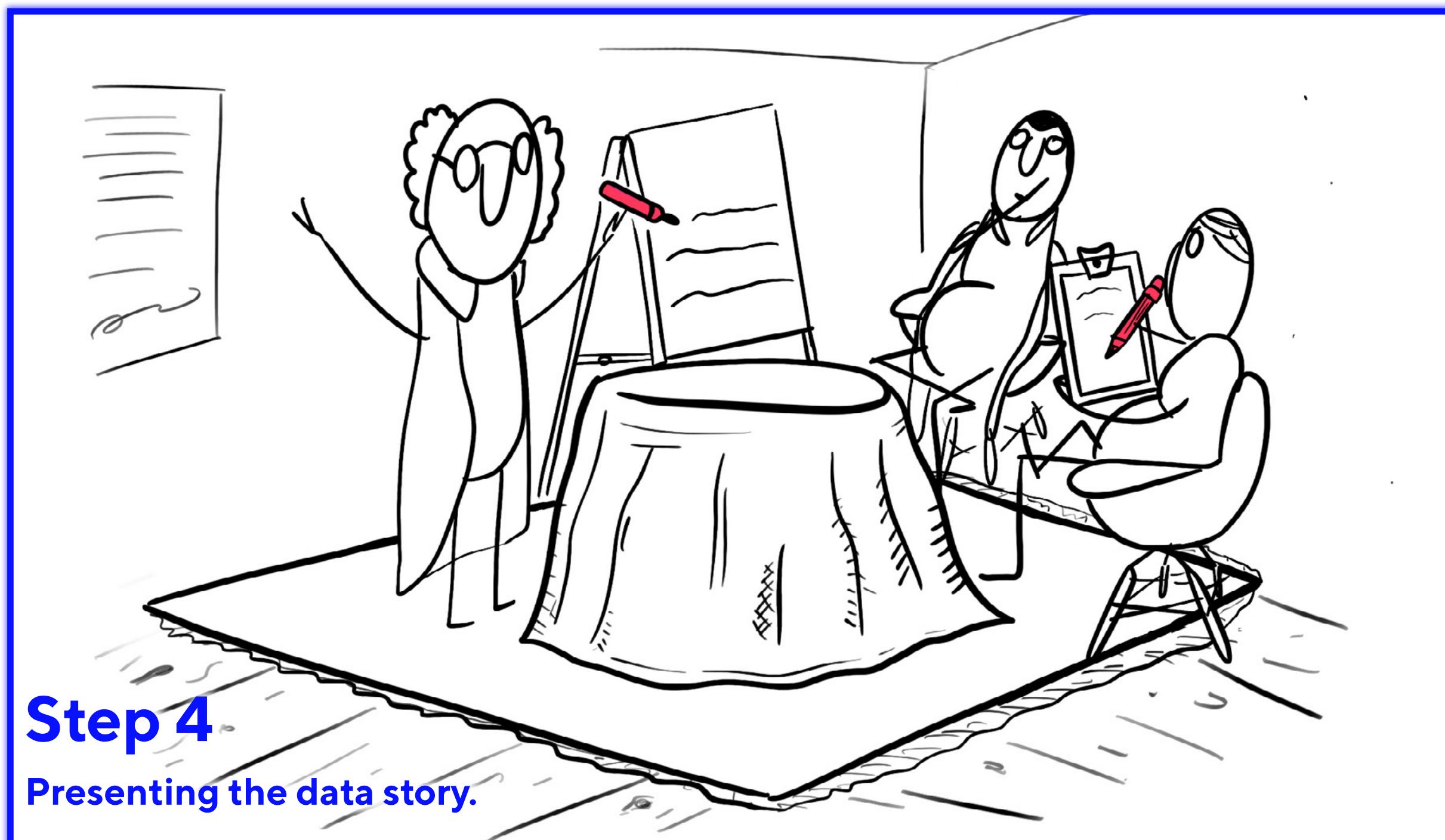
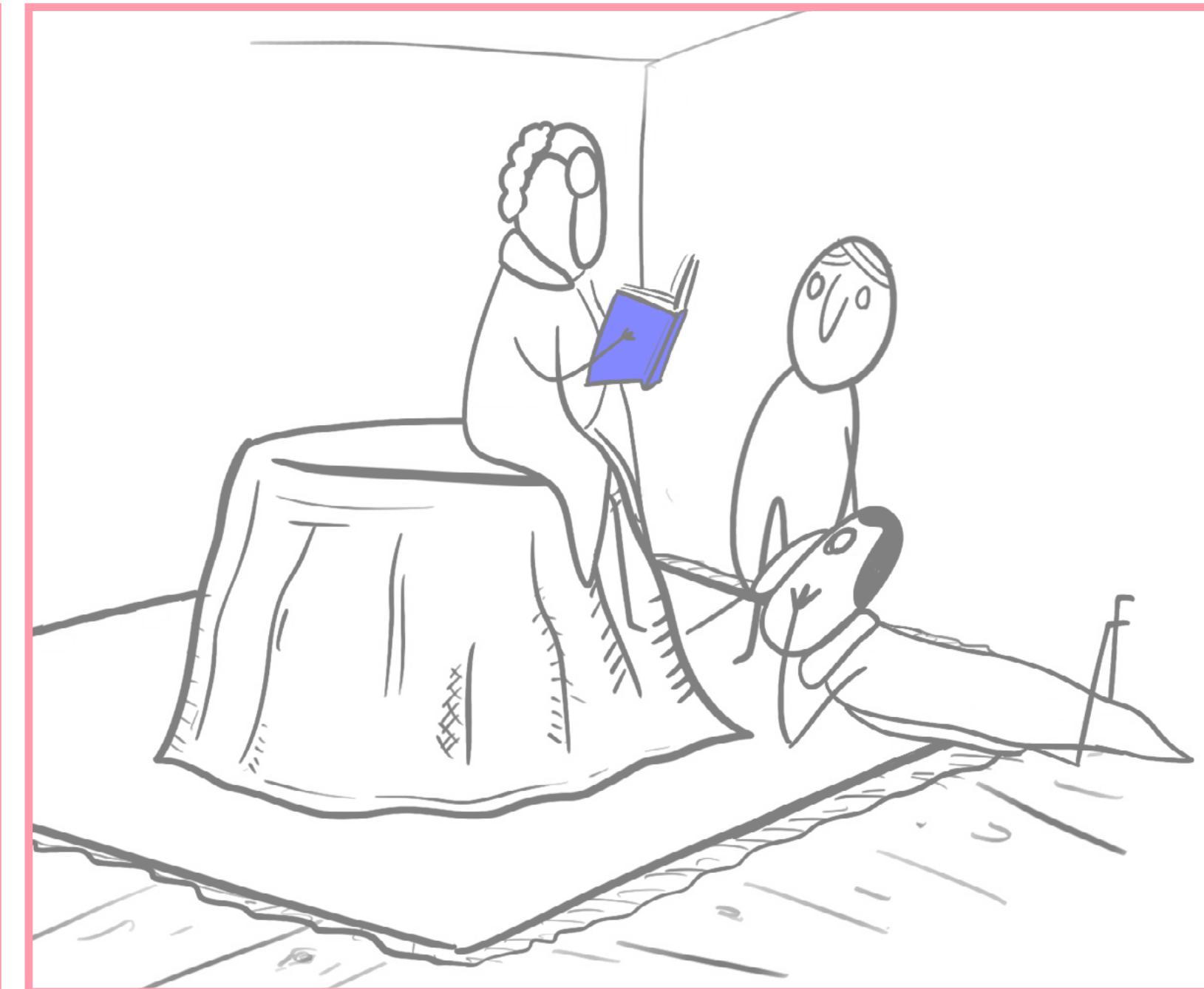
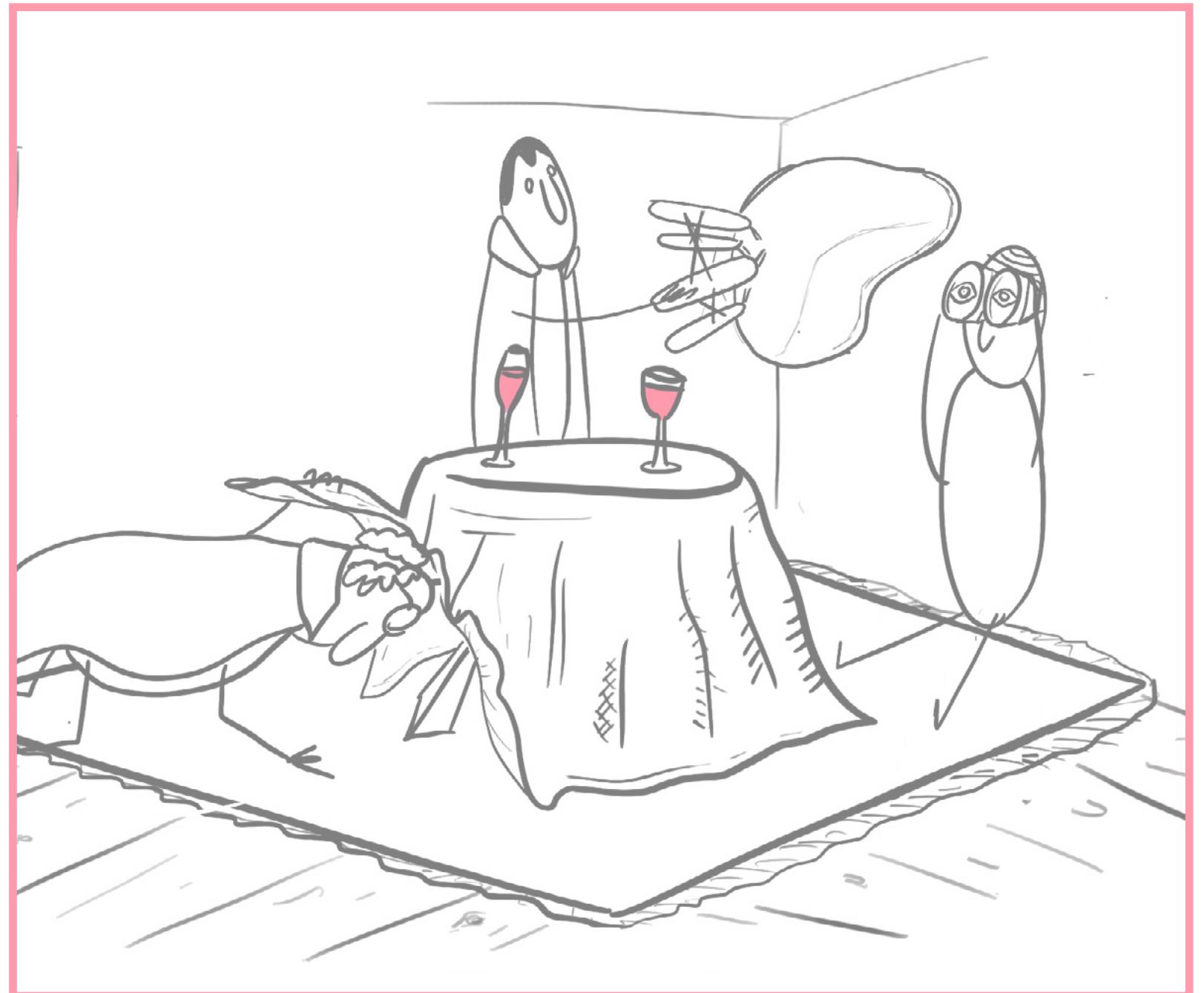
But drill-down story structures can also take other visualization forms, as shown in this example:

We've already used about 73% of our budget.

The world has emitted 2,100 gigatons of CO₂ since 1870, mostly from developed countries that prospered and polluted from the Industrial Revolution to today.

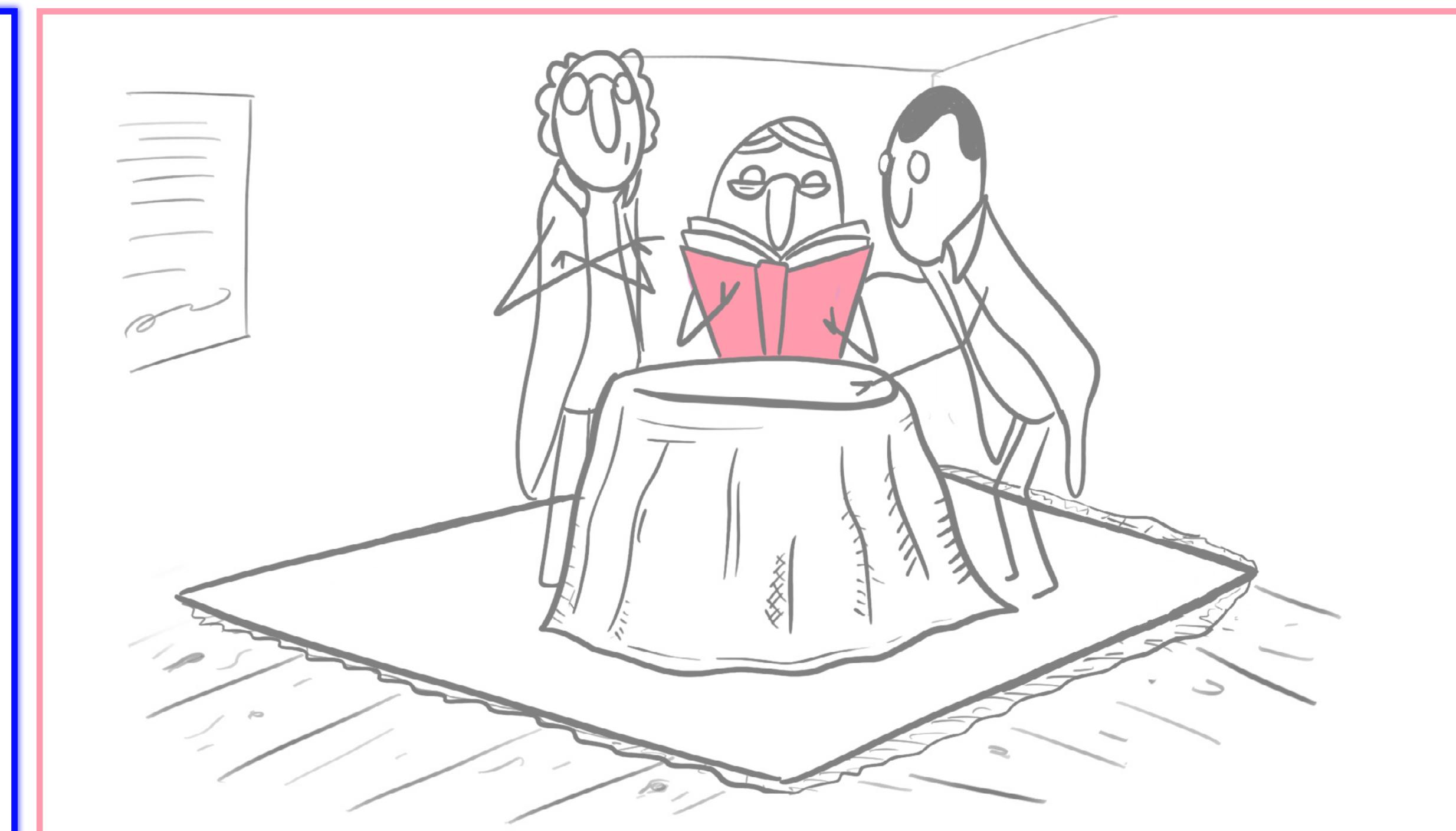
The United States played an outsized role here, responsible for about 20% of emissions despite having just 4.4% of the world's population. Overall, countries in the developed world account for 19% of the world's population but are responsible for more than half of all emissions to date.



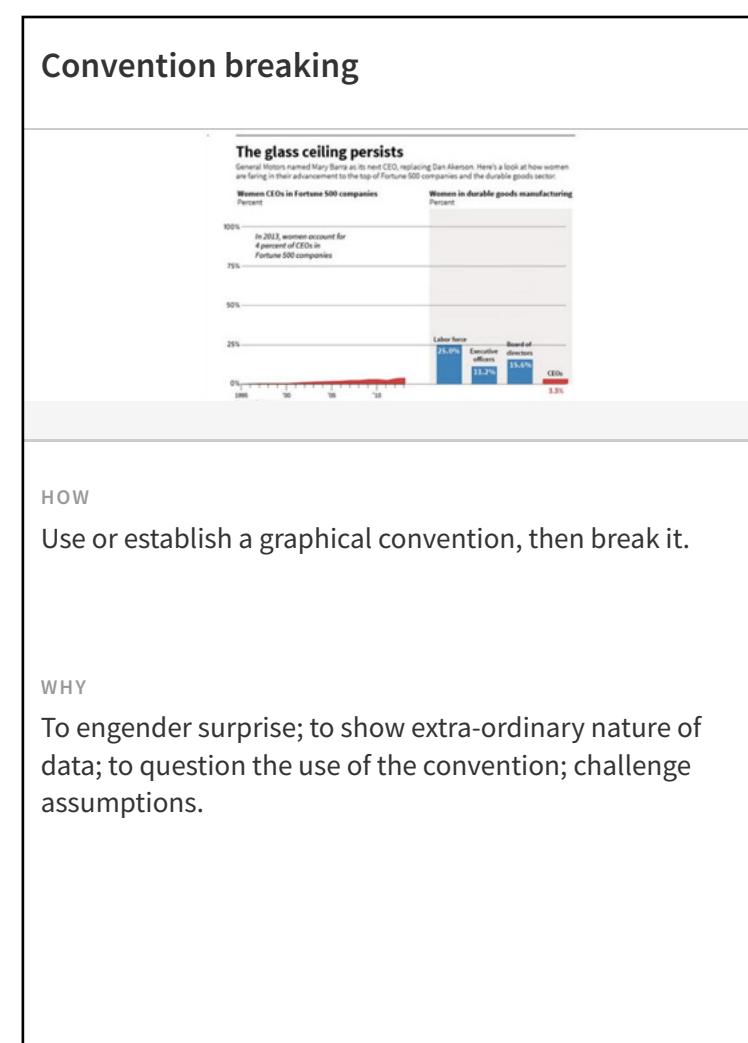
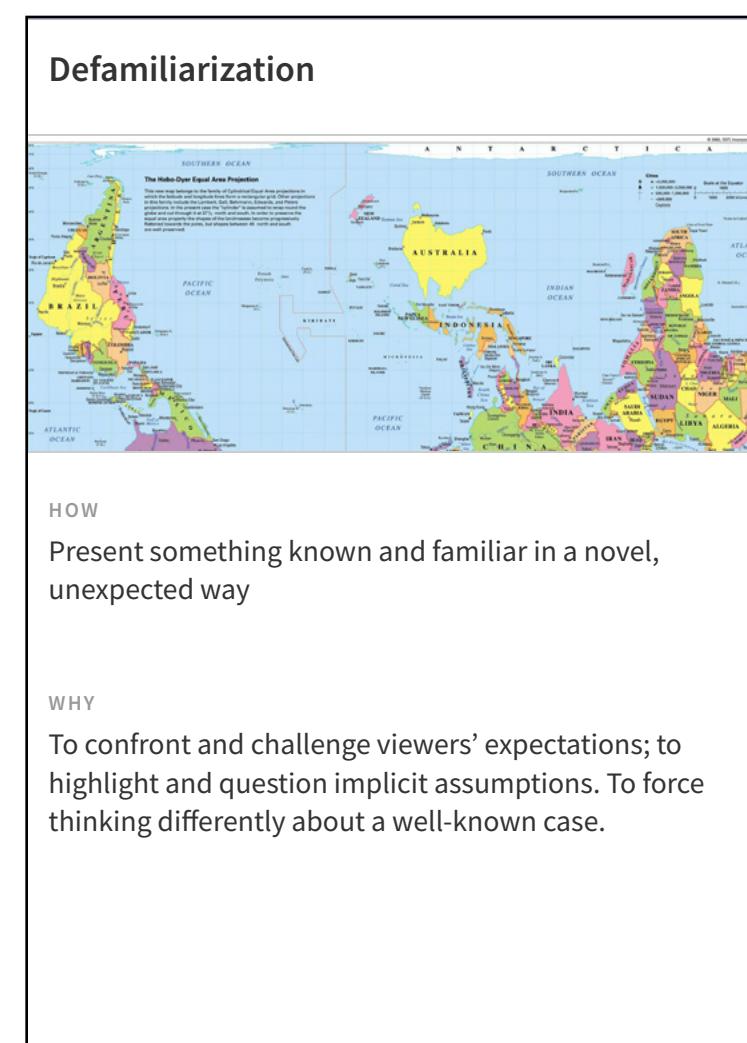
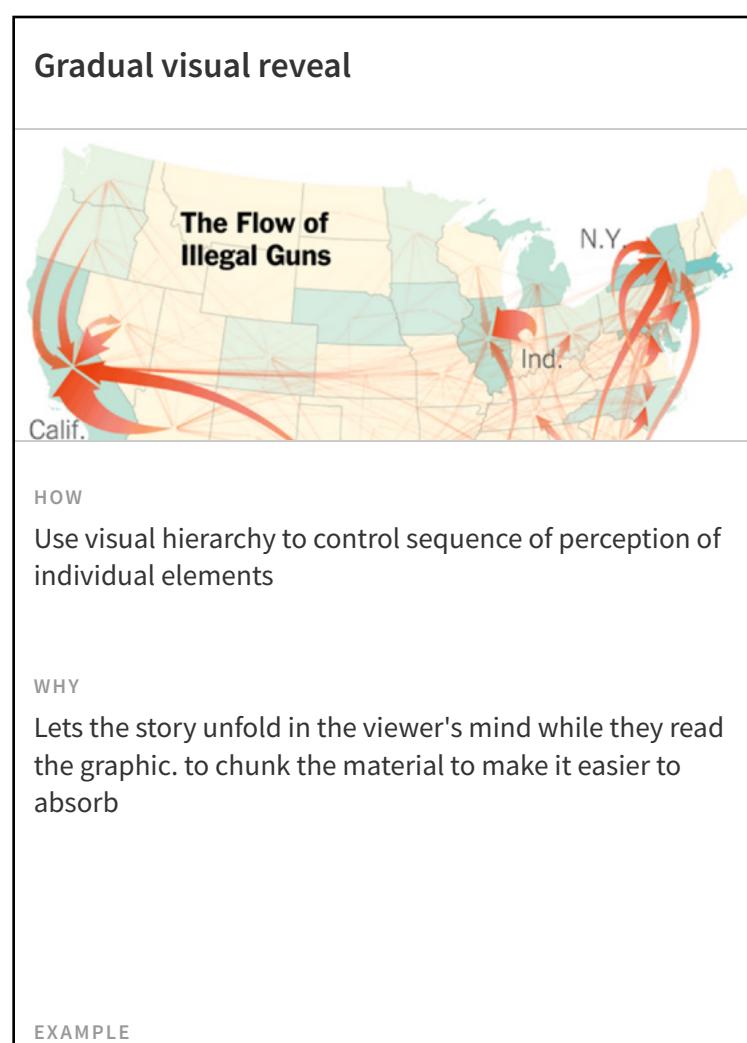


Step 4

Presenting the data story.



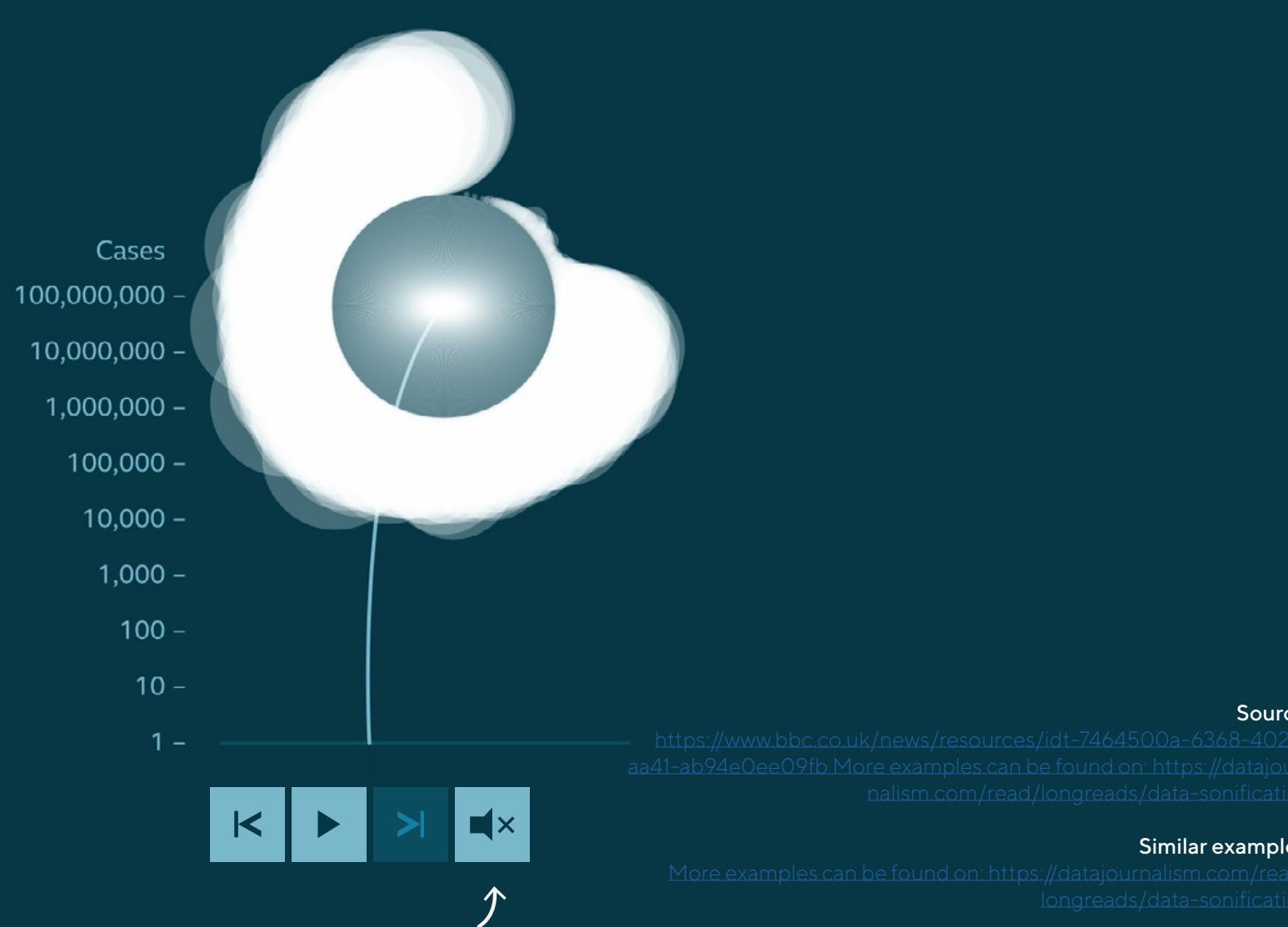
Think of how the visual language will support the story



NAPA Cards can help you decide! See: <http://napa-cards.net/>



Also consider other data representation forms; such as sound, to include everyone. BBC NEWS allows both visuals and sound to explore data.



Or use physical data representations.

These offer a friendly and familiar way to engage people that disregard digital visualizations as bar charts and line graphs.

Hans Rosling became famous for telling data stories with boxes.



Wait.

Some people might disregard nice data visualizations?



Wait.

Some people might disregard nice data visualizations?

Today, storytellers, artists and academic scholars are combining data representations with more familiar media (such as boxes in a video talk!) to engage a varied audience.



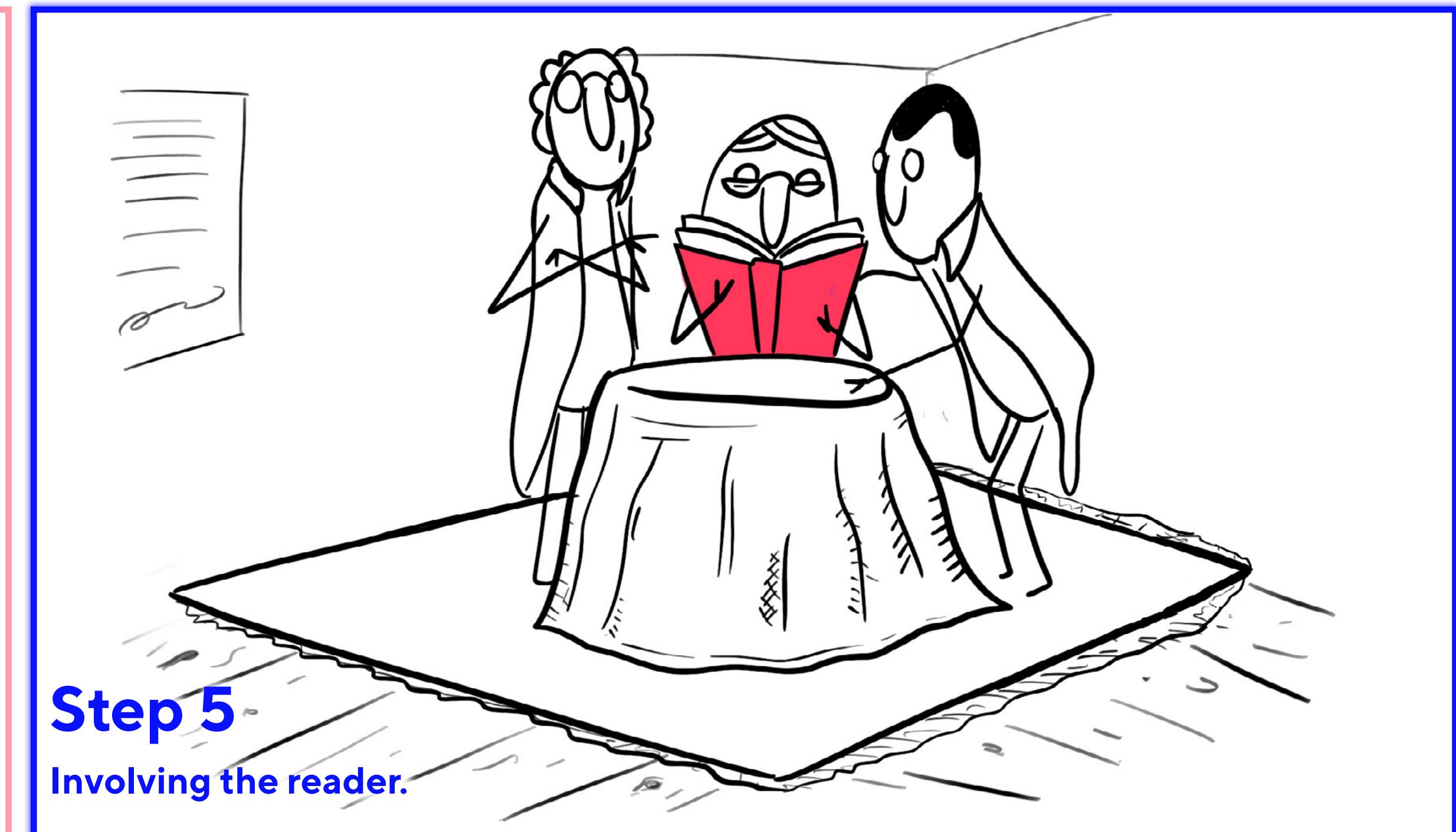
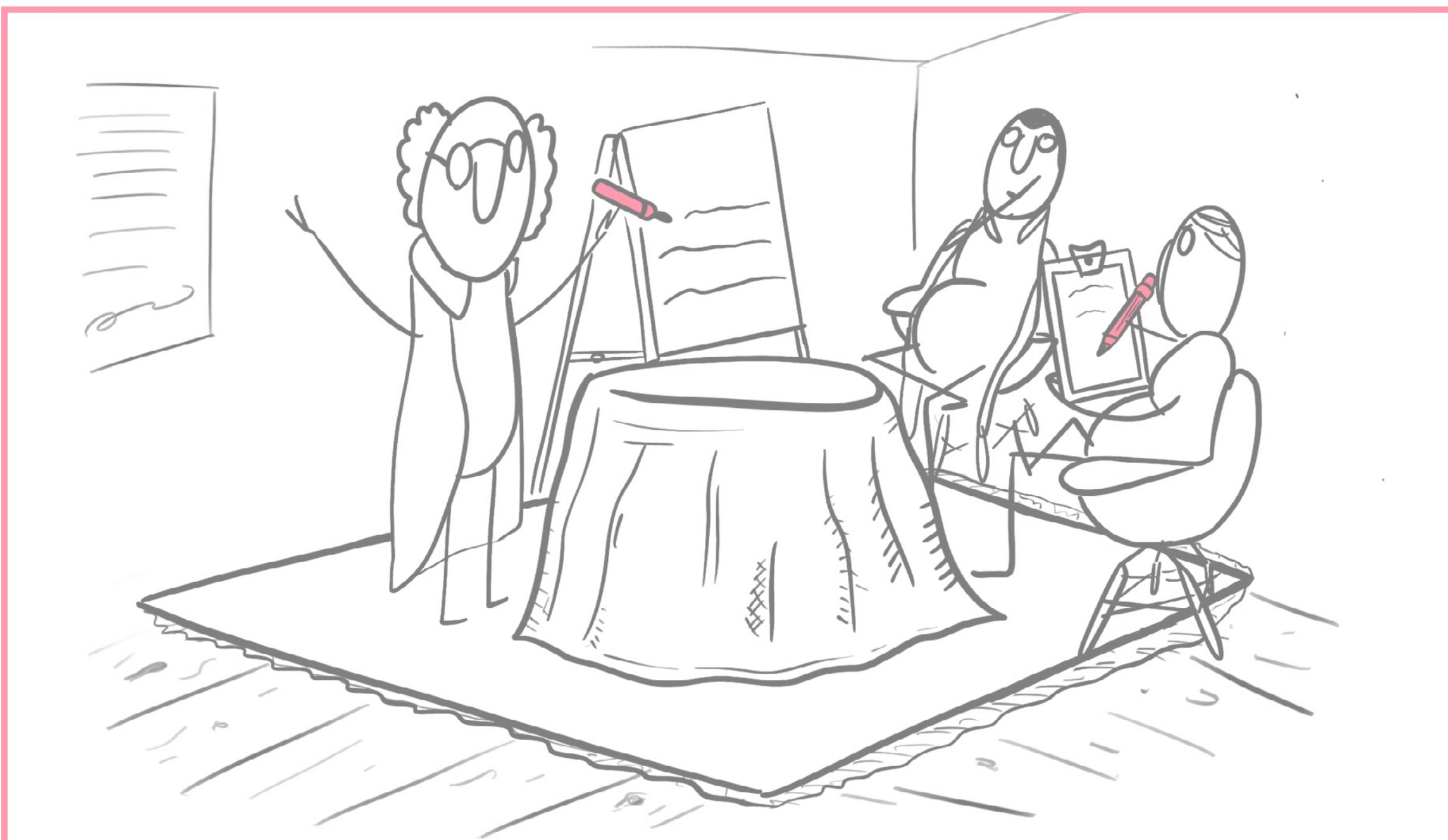
Wait.

Some people might disregard nice data visualizations?

Today, storytellers, artists and academic scholars are combining data representations with more familiar media (such as boxes in a video talk!).

Thus, data story structures and visualisation formats still apply, it is only combined with more qualitative information.





Step 5
Involving the reader.

When presenting data, the reader is already involved in some way of sensemaking.
Key is to trigger active participation to truly facilitate data literacy skills and combat misinformation

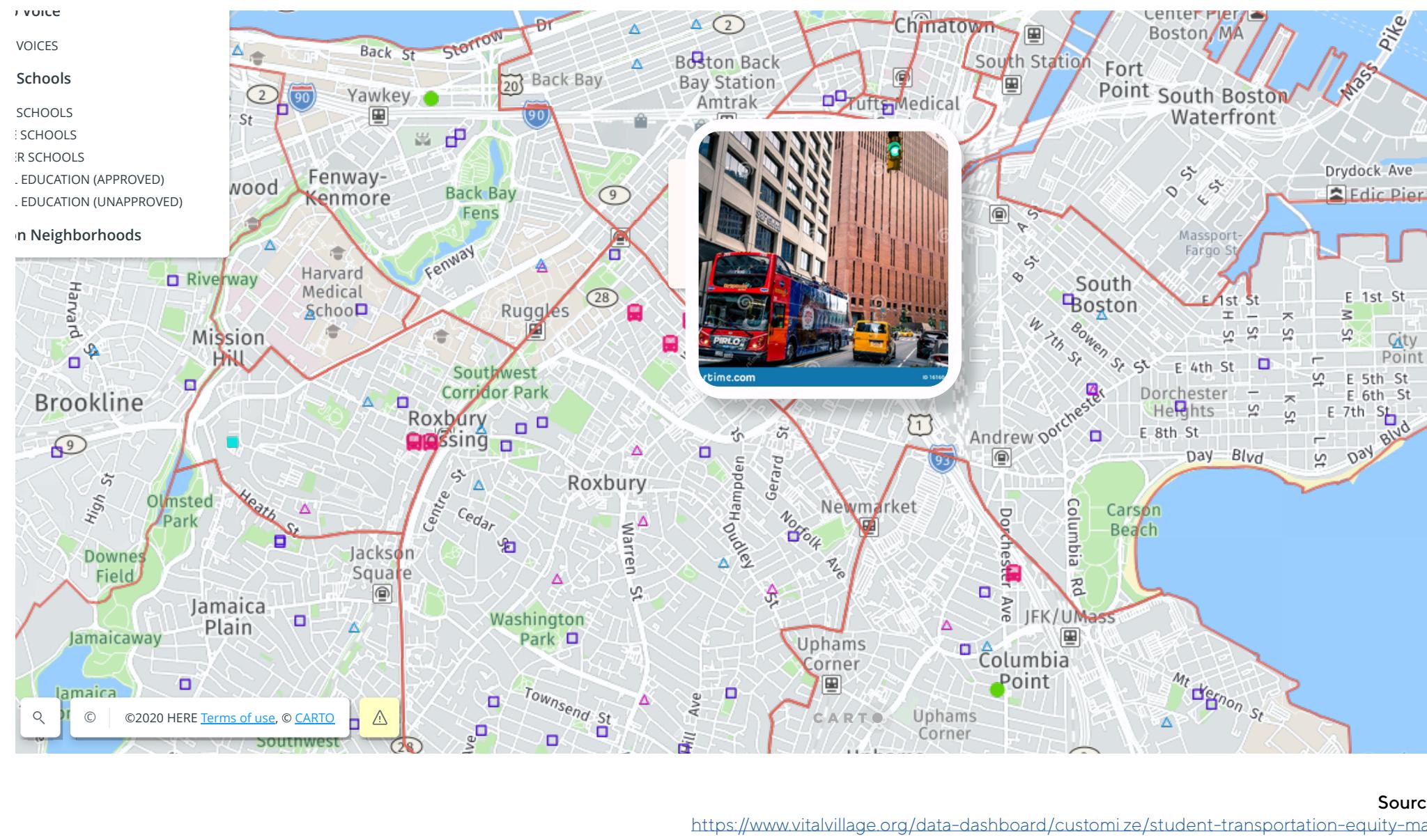
participation:



Add data storyline.

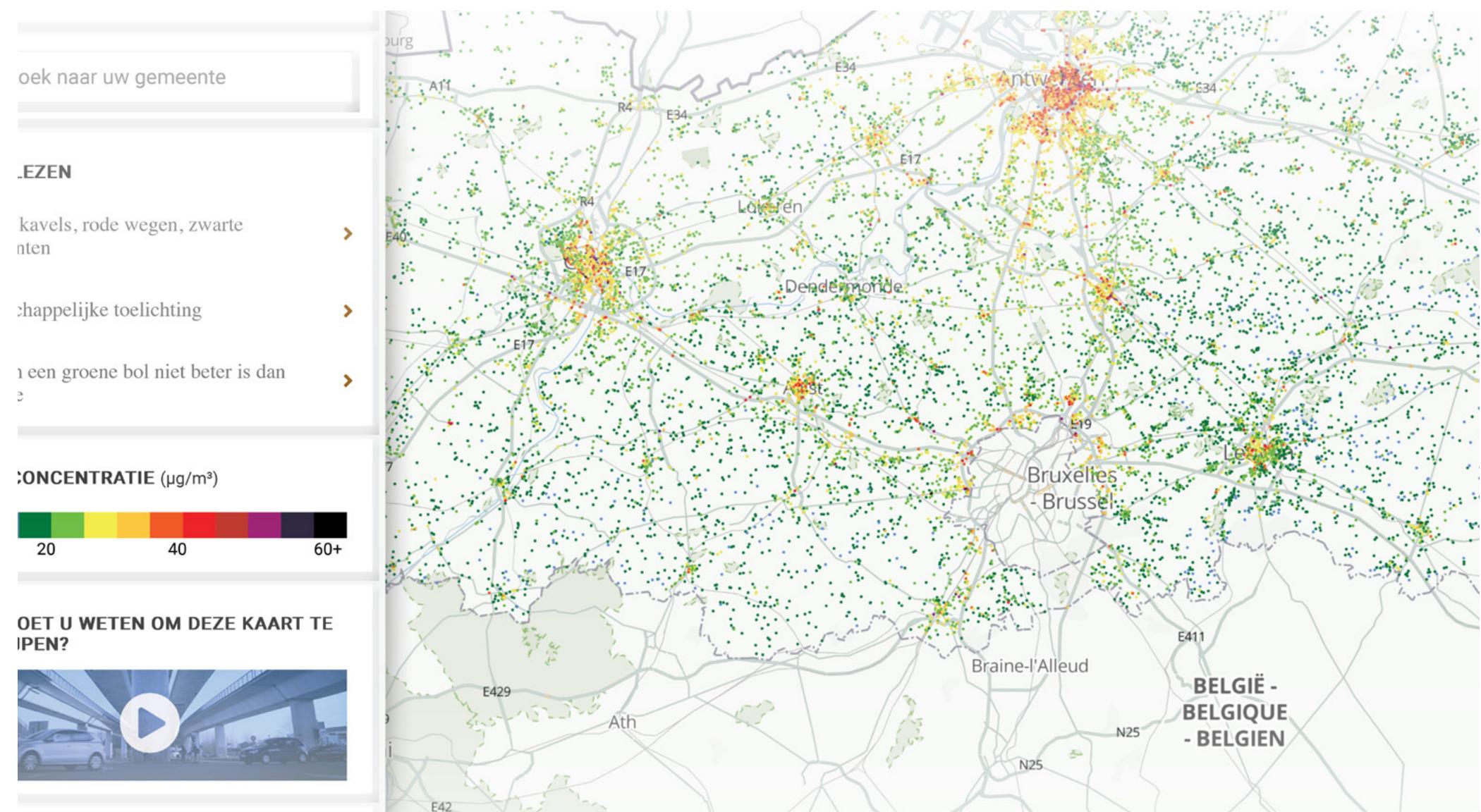
The ultimate form of participation! Readers may contribute their own data interpretation through anecdotes, photos...

...or both, as in this example that shows the issues with student transportation from students' perspectives.



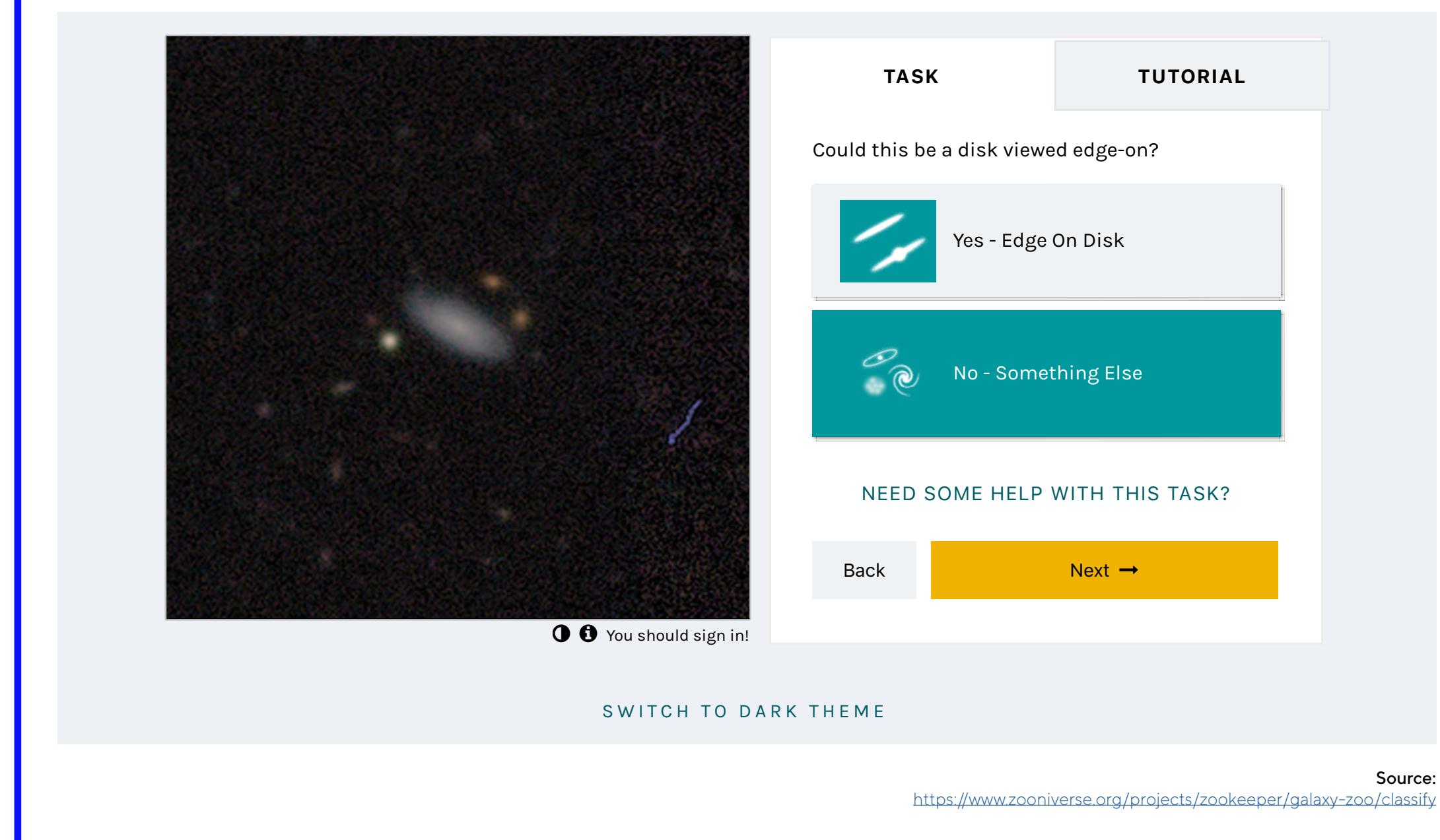
Add data to storyline.

It is another way to establish a personal connection with the story and the presented information.



Interpret data to progress storyline.

The reader has part in the story, which connects on a personal level.
For example; by identifying galaxies, the user is involved in a small part of the larger story that entails the complexity of space



Could this be a disk viewed edge-on?

Yes - Edge On Disk

No - Something Else

NEED SOME HELP WITH THIS TASK?

Back Next →

SWITCH TO DARK THEME

Source: <https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/classify>

Interpret data to discover passive participation.

Make data relevant by connecting it to the reader's personal situation.

in this example, they made a documentary film about 'Brooke' who lives near the user, depending on the users' geolocation data. As Brooke is a 'real' person, it is also easier to have empathy for her.



Concannon, S., Rajan, N., Shah, P., Smith, D., Ursu, M., & Hook, J. (2020, April). Brooke leave home: Designing a personalized film to support public engagement with open data. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-14).

Interpret data to discover passive participation.
Make data relevant by connecting it to the reader's personal situation.

Enter personal situation.

Location (County)
Poweshiek

How long will you be interacting with each person (on average)?
15

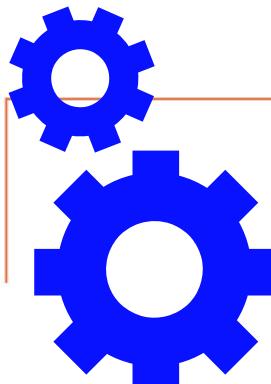
How many people will you interact with?
100

Will everyone be practicing physical distancing?
Yes

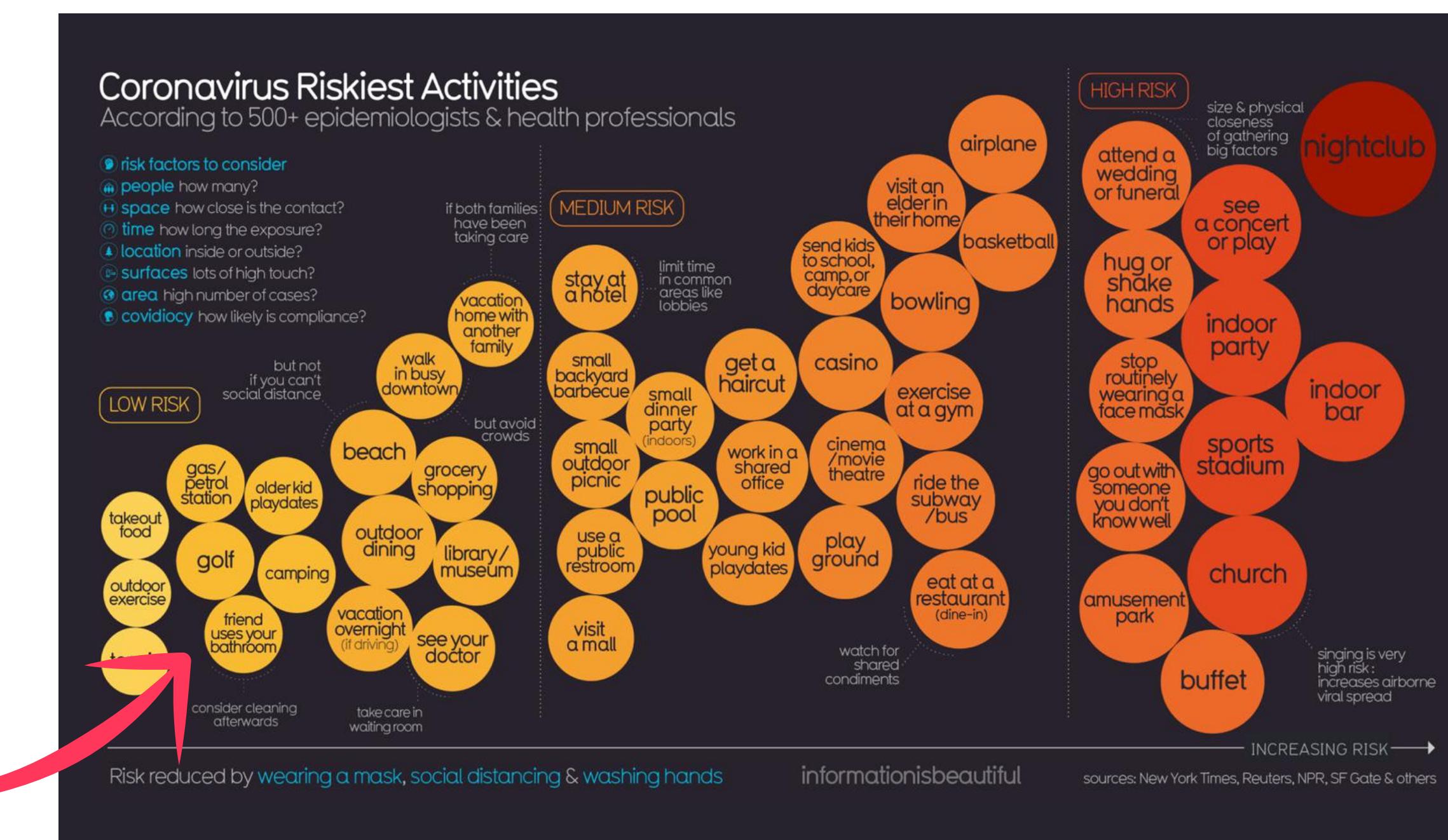
Will everyone be wearing a mask or shield?
Yes

Will this activity be indoors or outdoors?
Outdoors

Does this activity include rapid breathing (for example singing or exercise)?
Yes

 **High Risk**

Get results that are based on your situation.



PARCOS

Participatory Communication of Science



Learn more: parcos-project.eu
Presentation & Illustrations by: tristan@heemels.be



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 872500.