Best Practice for Whom?

SMI105: Week 10

Dr. Calum Webb
Sheffield Methods Institute, the University of Sheffield
c.j.webb@sheffield.ac.uk

Sign in

Learning outcomes

What will I learn?

By the end of this week you will know:

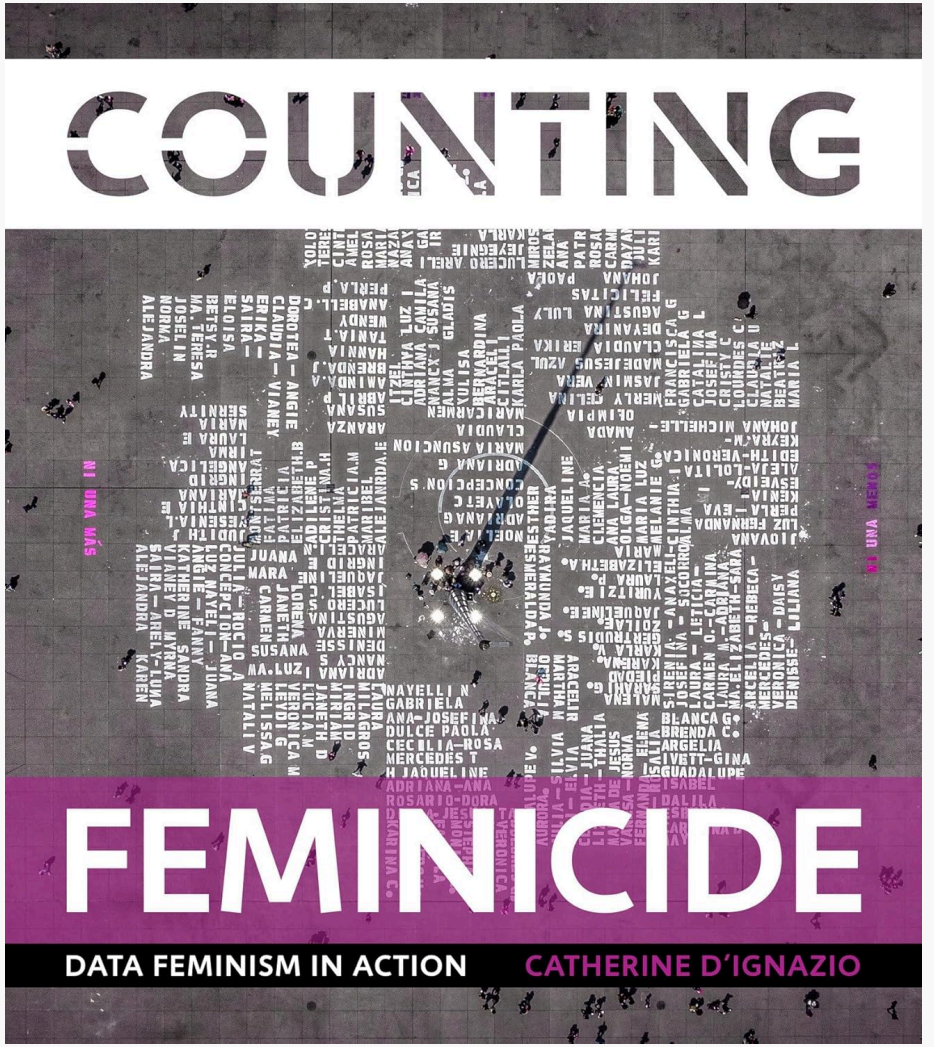
- So far we've been talking about best practice in terms of graphic design and coding. This week, we're going to focus specifically on good practice for critically reflecting on **the role that data visualisations play in society**.
- You'll learn some core principles from data feminism and other critical scholarly research so that we can **do better** when we make data visualisations, and **ensure that our data visualisations play a more positive role in society than a negative one**.

Both the data used in our data visualisations, and our data visualisations themselves, often use processes that often reinforce and reproduce societal biases.

More critical awareness of those biases and their impact can help us improve our data visualisations to benefit society more equitably.

(Some) Principles of Feminist Data (D'Ignazio & Klein, 2016; D'Ignazio, 2024) and further challenges (Feigenbaum & Alamalhodaei, 2020)

- Rethink binaries & classification
- Embrace pluralism
- Examine power & encourage empowerment
- Consider context
- Legitimise embodiment & affect
- Inclusion, exclusion & accessibility
- Sanitisation & impersonalisation



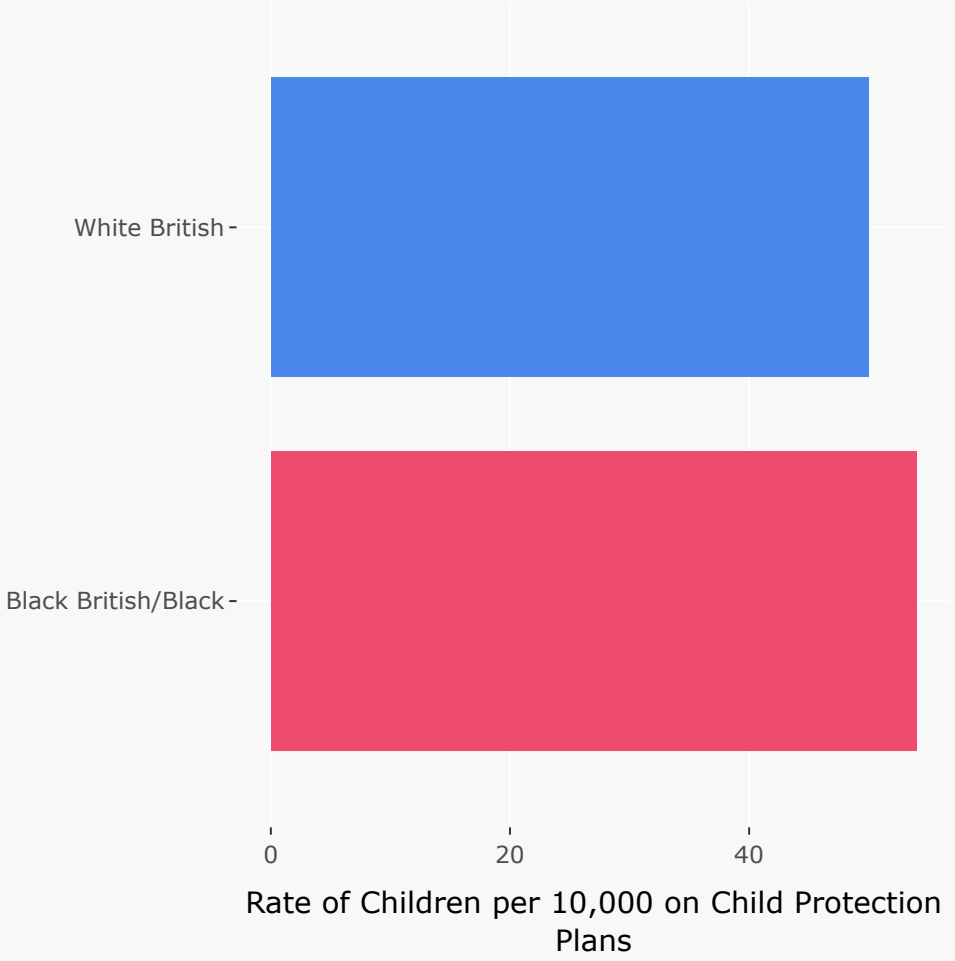
Rethinking binaries and classification

Often for the sake of convenience or simplicity, data collection and data visualisations will use limited numbers of categories and, often **binaries**.

These often compress and do not fully represent peoples' opinions, identities, or experiences

Some examples of binaries/limited classification:

- Thumbs up / Thumbs down
- Ethnicity and nationality



- Disability

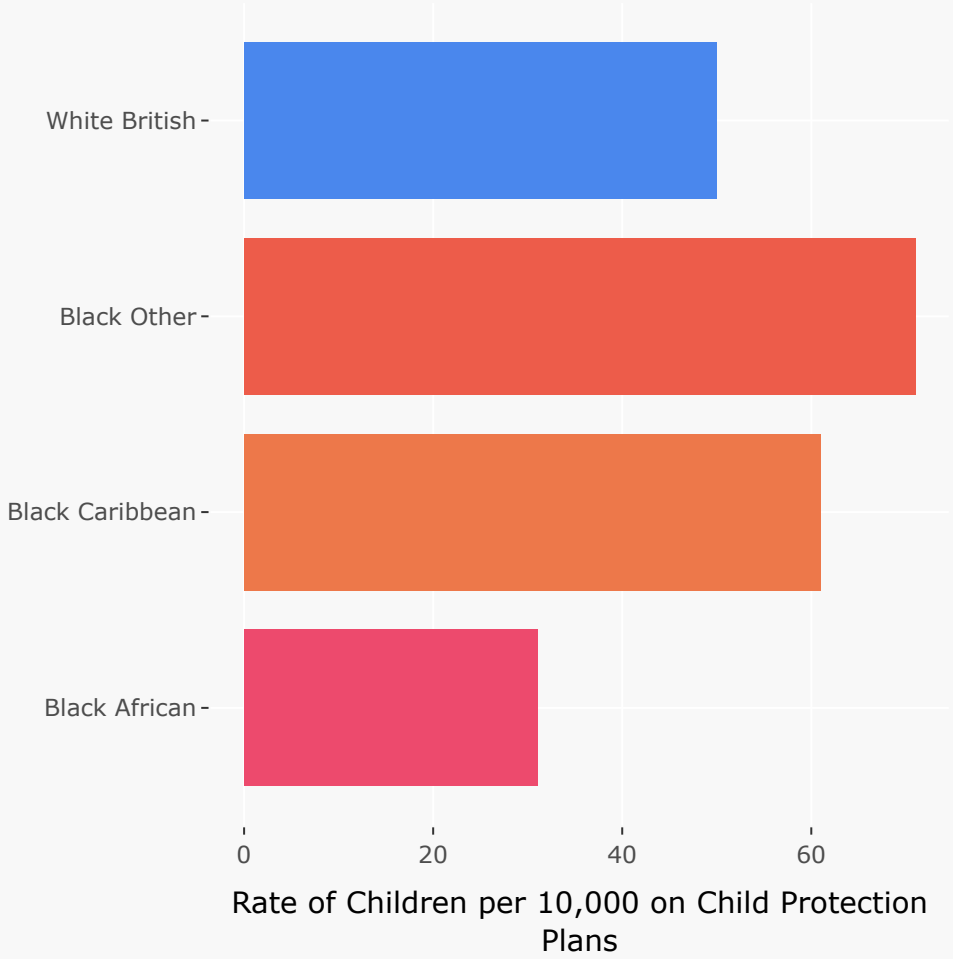
Rethinking binaries and classification

Often for the sake of convenience or simplicity, data collection and data visualisations will use limited numbers of categories and, often **binaries**.

These often compress and do not fully represent peoples' opinions, identities, or experiences

Some examples of binaries/limited classification:

- Thumbs up / Thumbs down
- Ethnicity and nationality



Based on Bywaters, et al. 2019.

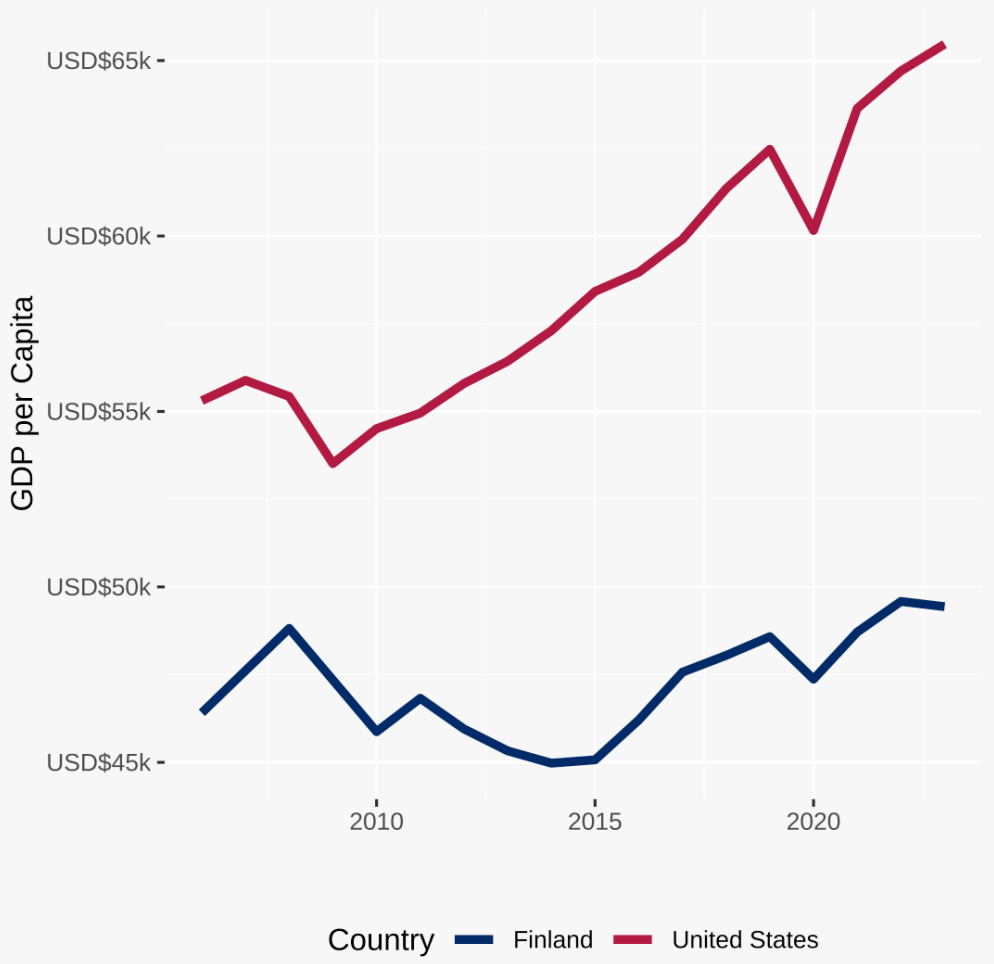
- Disability

Embrace pluralism

When we visualise certain data, **what is it we are actually trying to talk about**, and what other ways might it be considered?

For example, we often use GDP growth when **what we are talking about** is prosperity and quality of life; how else might these concepts be represented and **who benefits most from our emphasis on GDP**?

How can we create ranges of data visualisations that recognise multiple points of view?



Helliwell, et al. (2024). World Happiness Report 2024.

Embrace pluralism

When we visualise certain data, **what is it we are actually trying to talk about**, and what other ways might it be considered?

For example, we often use GDP growth when **what we are talking about** is prosperity and quality of life; how else might these concepts be represented and **who benefits most from our emphasis on GDP**?

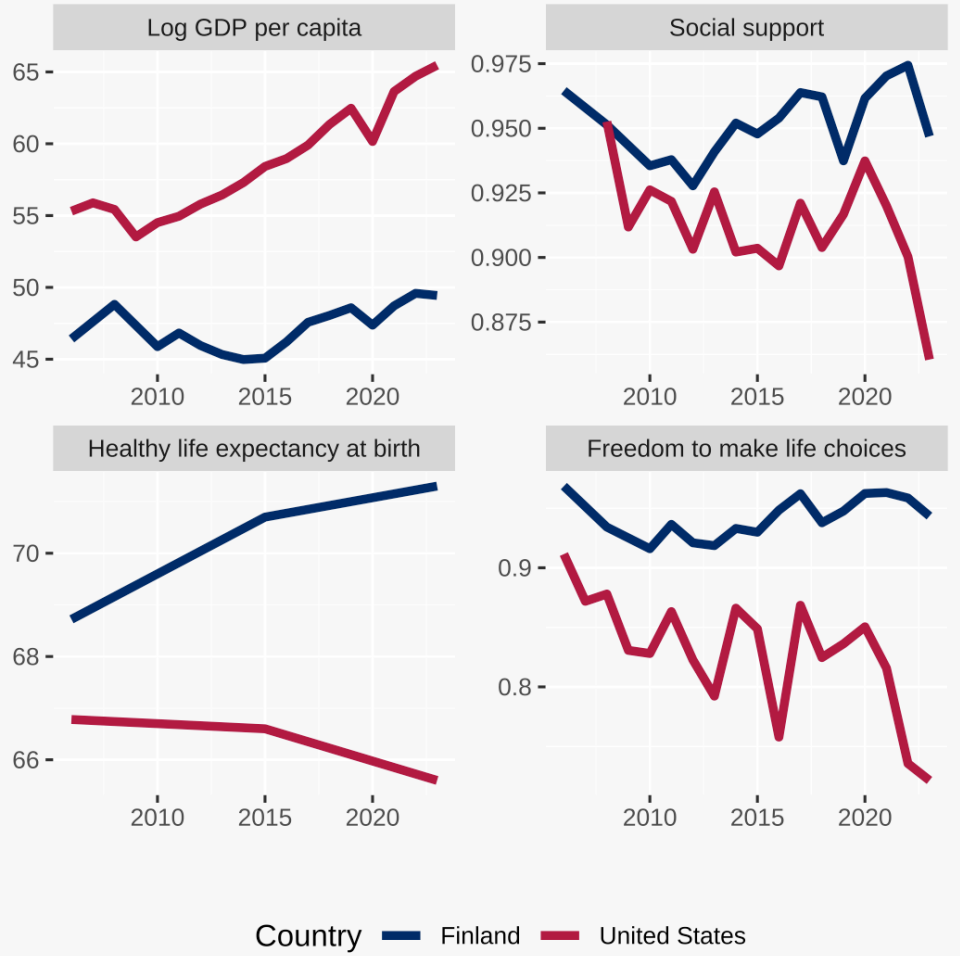
How can we create ranges of data visualisations that recognise multiple points of view?

Embrace pluralism

When we visualise certain data, **what is it we are actually trying to talk about**, and what other ways might it be considered?

For example, we often use GDP growth when **what we are talking about** is prosperity and quality of life; how else might these concepts be represented and **who benefits most from our emphasis on GDP**?

How can we create ranges of data visualisations that recognise multiple points of view?



Helliwell, et al. (2024). World Happiness Report 2024.

Embrace pluralism

When we visualise certain data, **what is it we are actually trying to talk about**, and what other ways might it be considered?

For example, we often use GDP growth when **what we are talking about** is prosperity and quality of life; how else might these concepts be represented and **who benefits most from our emphasis on GDP**?

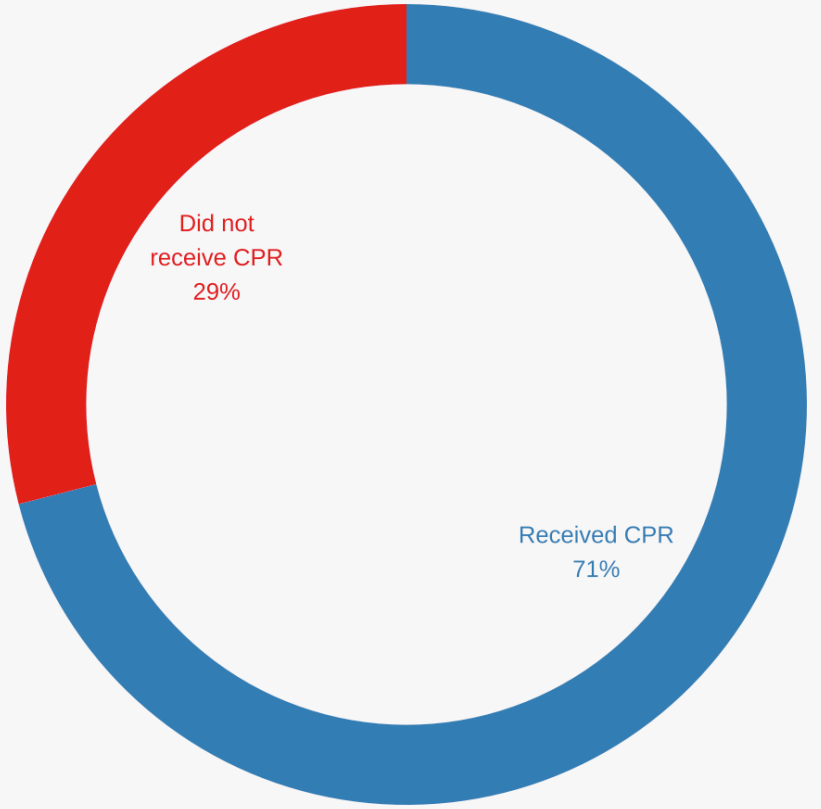
How can we create ranges of data visualisations that recognise multiple points of view?

Examine power & encourage empowerment

Do your data visualisations, or the data visualisations of others, consider **which groups in society might be most effected by the topic being visualised**?

In what ways are you trying to make your data visualisations a tool that *they* can use to challenge unjust outcomes and experiences?

- **Are your data visualisations accessible**: do they clearly tell a story? Is technical language explained? Are they likely to attract the engagement of those groups most affected? Are they accessible to the groups most affected (especially in terms of publication method)?
- Is there a **call to action**? What can people do to create a change when something there is a problem in society that you have visualised?



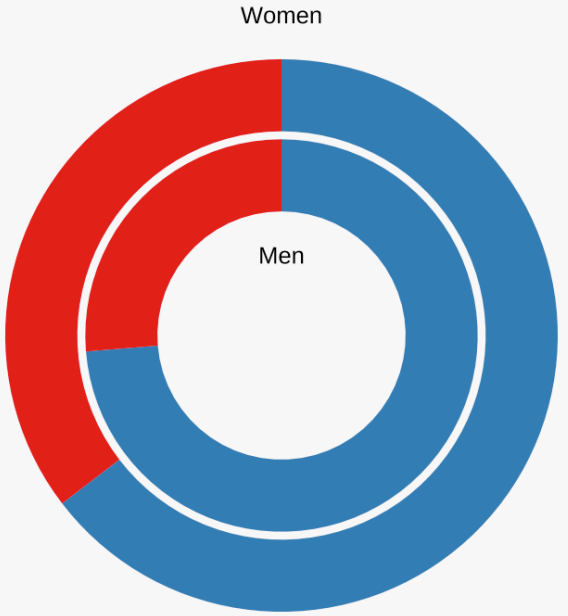
Munot, et al. (2024). Bystander cardiopulmonary resuscitation differences by sex.

Examine power & encourage empowerment

Do your data visualisations, or the data visualisations of others, consider **which groups in society might be most effected by the topic being visualised**?

In what ways are you trying to make your data visualisations a tool that *they* can use to challenge unjust outcomes and experiences?

- **Are your data visualisations accessible**: do they clearly tell a story? Is technical language explained? Are they likely to attract the engagement of those groups most affected? Are they accessible to the groups most affected (especially in terms of publication method)?
- Is there a **call to action**? What can people do to create a change when something there is a problem in society that you have visualised?



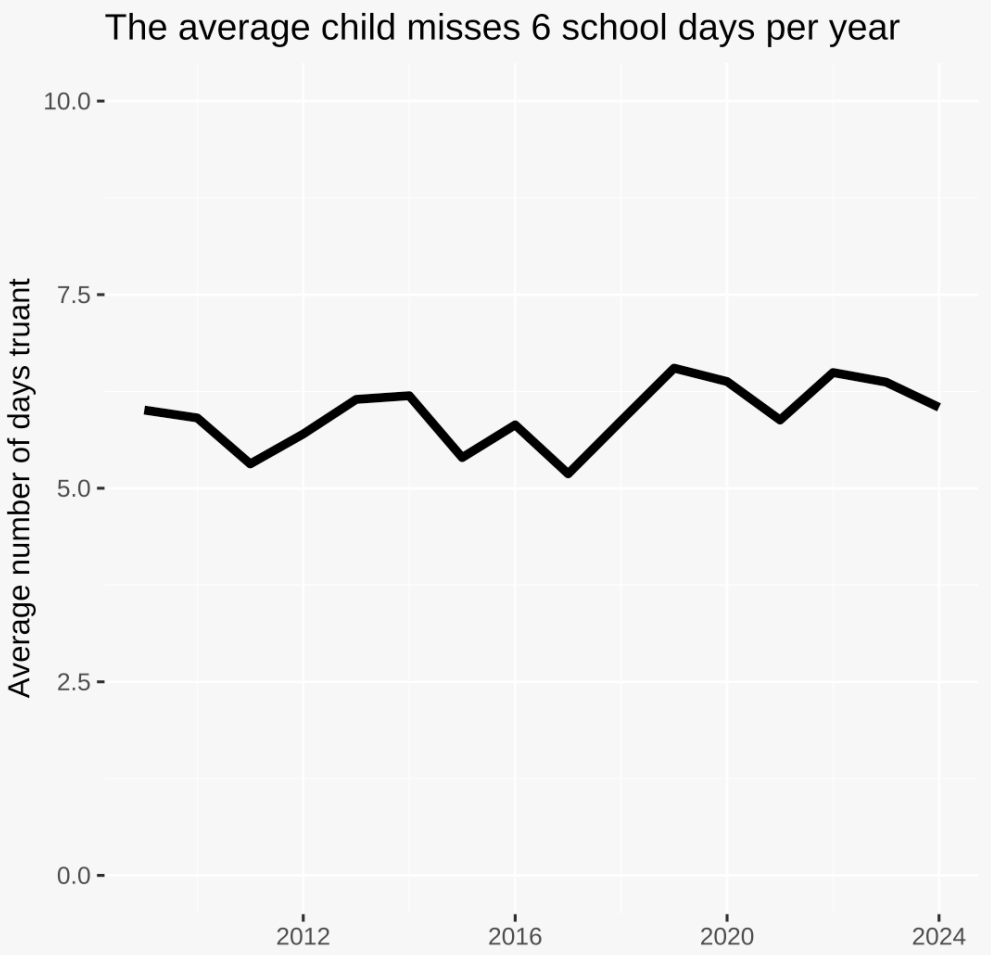
Munot, et al. (2024). Bystander cardiopulmonary resuscitation differences by sex.

Consider context

What is the context in which data is generated, and **how would better engagement with the people who the data represents change our understanding of the data being generated** and what matters to them?

For what reason is the data collected, why are some types of data not collected, and who does that benefit or disadvantage?

Could we do the following:



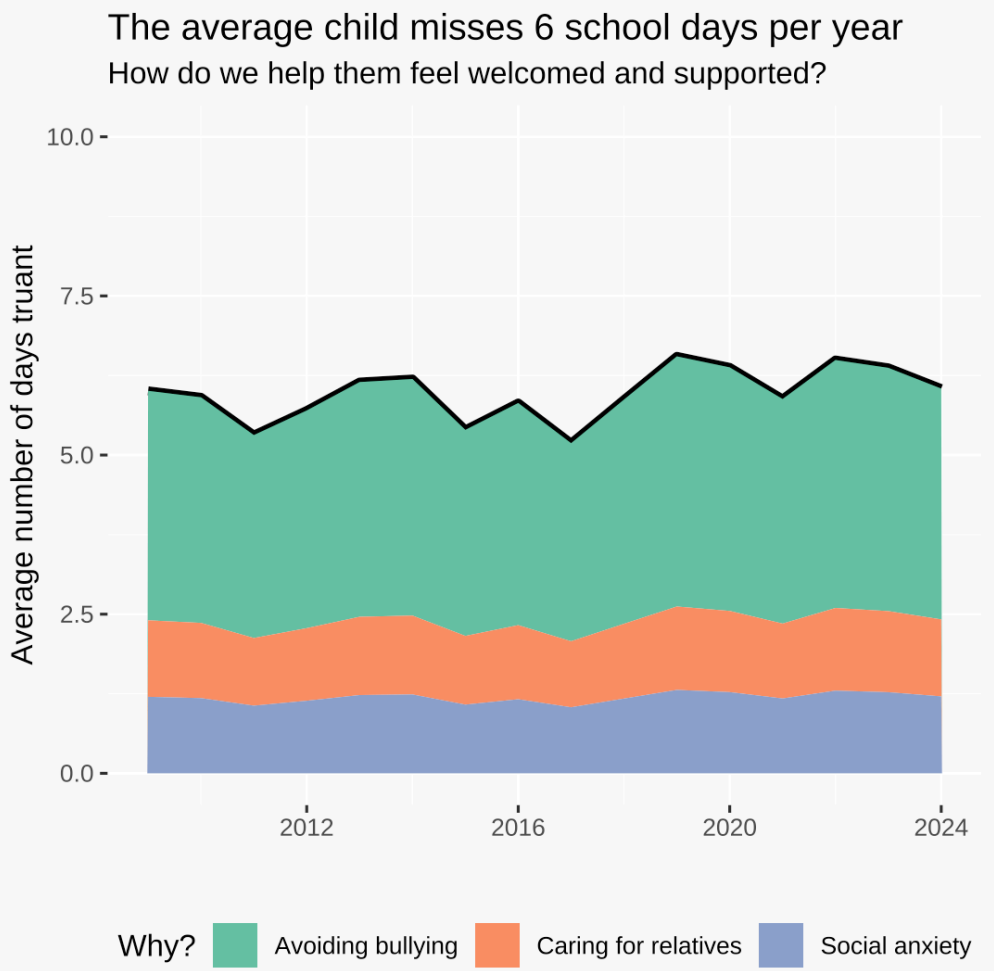
- Work together with the people who the data is about to find out what things matter to them and how they should be measured and visualised?
- Consider the complexity of the data that *is* collected: e.g. when a

Consider context

What is the context in which data is generated, and **how would better engagement with the people who the data represents change our understanding of the data being generated** and what matters to them?

For what reason is the data collected, why are some types of data not collected, and who does that benefit or disadvantage?

Could we do the following:



- Work together with the people who the data is about to find out what things matter to them and how they should be measured and visualised?
- Consider the complexity of the data that *is* collected: e.g. when a

Legitimise embodiment and affect

Don't be a data visualisation snob.

The principles and practice of data visualisation can be applied to many different materialities that draw upon individual peoples' skills and strengths.

Data visualisation is often a rapid, fast process of creating and publishing data visualisations — that may be good for some things, but as part of the way that data and data visualisation fits within our daily lives, "slow" forms of data science and analysis can also help is

"**reflect on everyday events that happen all around us**" (Schoof, 2022)

These types of "**data visceralisations**" can also cause people who

> This graph shows the length of daily delays on one woman's 40-minute commute between Munich and a town in the German countryside during 2018...
>
> The commuter knitted two rows each day. Gray for delays under five minutes, pink for up to 30 minutes, and red for a delay of more than a half-hour or delays in both directions.

New York Times (2019). What's Going On in This Graph? | Feb. 27, 2019

Legitimise embodiment and affect

Don't be a data visualisation snob.

The principles and practice of data visualisation can be applied to many different materialities that draw upon individual peoples' skills and strengths.

Data visualisation is often a rapid, fast process of creating and publishing data visualisations — that may be good for some things, but as part of the way that data and data visualisation fits within our daily lives, "slow" forms of data science and analysis can also help is

> Just 158 families have provided nearly half of the early money for efforts to capture the White House.

New York Times (Oct. 10, 2015). Buying Power

"**reflect on everyday events that happen all around us**" (Schoof, 2022)

These types of "**data viscoralisations**" can also cause people who

Legitimise embodiment and affect

Don't be a data visualisation snob.

The principles and practice of data visualisation can be applied to many different materialities that draw upon individual peoples' skills and strengths.

Data visualisation is often a rapid, fast process of creating and publishing data visualisations — that may be good for some things, but as part of the way that data and data visualisation fits within our daily lives, "slow" forms of data science and analysis can also help is

"**reflect on everyday events that happen all around us**" (Schoof, 2022)

These types of "**data visceralisations**" can also cause people who

> This group of students in my class decided to experiment with flavor as a way to represent air quality data. ... To surprise the audience they invite participants to taste different brownies but didn't telling them that the amount of salt had been increased based on how much pollution in the air there is in different cities.

Rahul Bhargava, MIT Media Lab (2019). Thoughts On Designing Data Sculptures

Inclusion, exclusion & accessibility

What don't you need to make data visualisations and tell stories using data?

You don't need to be a statistician, computer scientist, a graphic designer, or to have a lot of money to make data visualisations.

You are all proof of this.

- R is free
- RStudio is free
- All of the data we have been visualising is free

- Many of the books we have been using (e.g. Healy, 2018,

Inclusion, exclusion & accessibility

What don't you need to make data visualisations and tell stories using data?

You don't need to be a statistician, computer scientist, a graphic designer, or to have a lot of money to make data visualisations.

You are all proof of this.

- R is free
- RStudio is free
- All of the data we have been visualising is free

What **do** you need to make data visualisations and tell stories using data?

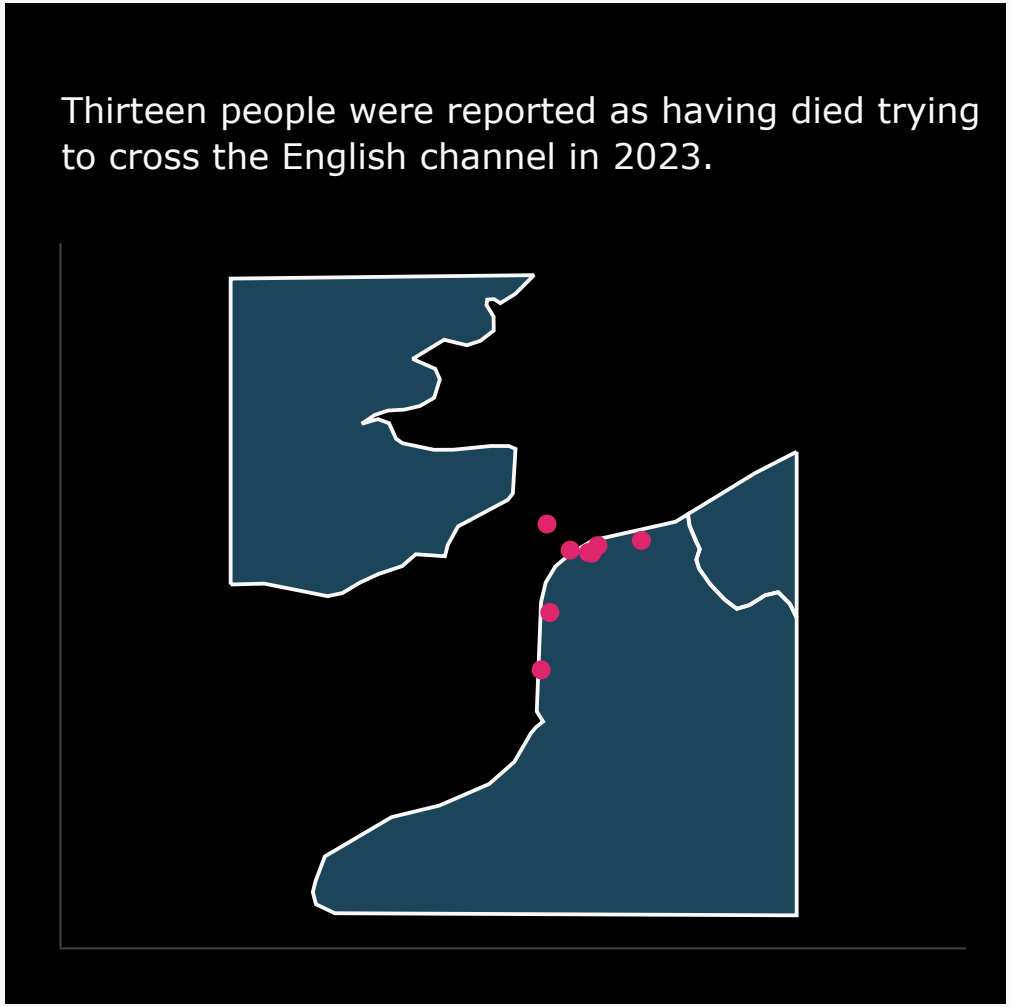
- A story that you want to tell.
- A willingness to learn how to create.
- A commitment to representing data in an accessible, fair, way.
- The empathy to put yourselves into the shoes of your audience.

- Many of the books we have been using (e.g. Healy, 2018,

Sanitisation & impersonalisation

The process of data collection, analysis and visualisation often **divorces the data subjects from their data**. Data become sterilised: we often lose any sense of identity, personality, or emotion from the "data subjects" — a.k.a. the people who the data is about.

This means that typical *best practice* approaches to data visualisations often make it **difficult or impossible to develop empathy for the people whom it is about** — removing us from the positionalities of the subjects (Kennedy, et al. 2016). Data depersonalises individuals.



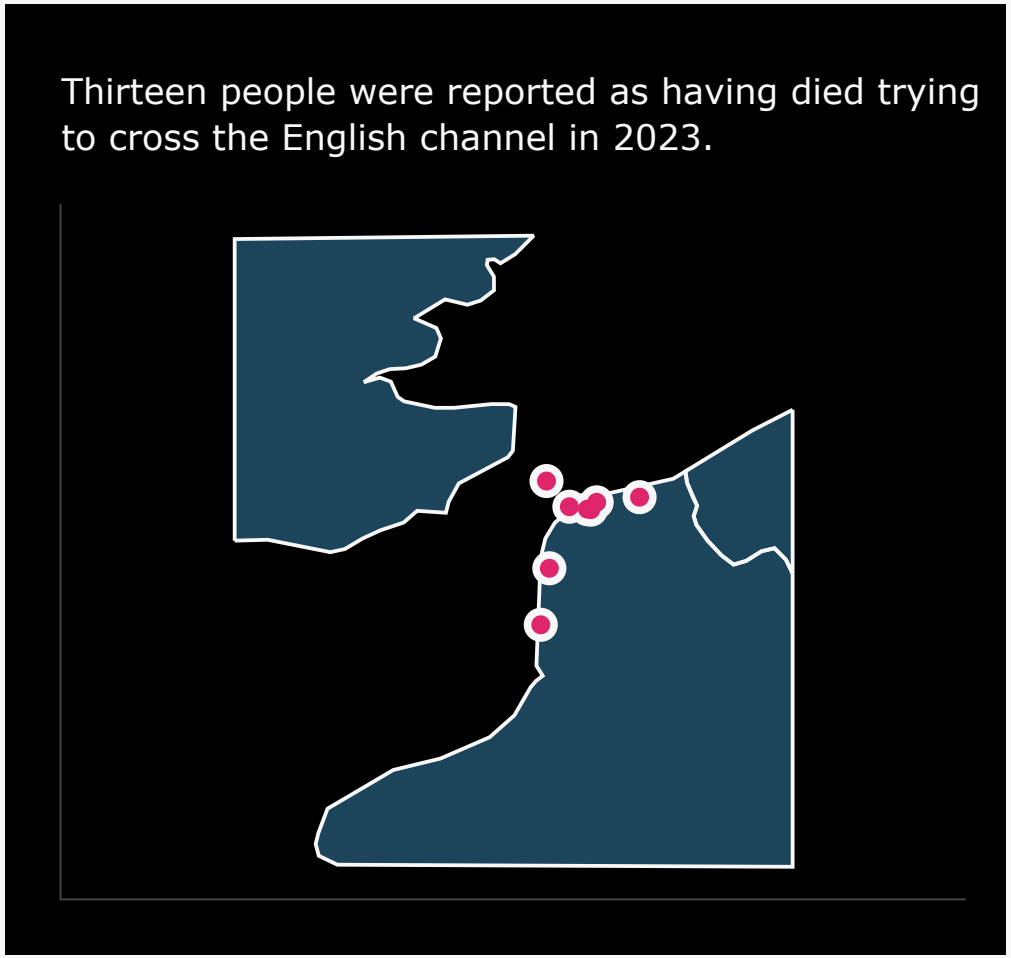
Data on incidents from The Missing Migrants Project

Lack of incorporation of voice and identity invites us to project our own, or dominant, narratives about the people concerned.

Sanitisation & impersonalisation

The process of data collection, analysis and visualisation often **divorces the data subjects from their data**. Data become sterilised: we often lose any sense of identity, personality, or emotion from the "data subjects" — a.k.a. the people who the data is about.

This means that typical *best practice* approaches to data visualisations often make it **difficult or impossible to develop empathy for the people whom it is about** — removing us from the positionalities of the subjects (Kennedy, et al. 2016). Data depersonalises individuals.



Data on incidents from The Missing Migrants Project, overlayed with quotes from migrants and refugees sourced from ITV.com, Big Issue, Migrant Voice, The Voices Network, Ordinary Things

Lack of incorporation of voice and identity invites us to project our own, or dominant, narratives about the people concerned.

It's impossible to make a perfect data visualisation. But thinking about these things can help us *do better*.

Best practice is not just about good design principles, but being a good and conscientious member of society. That means paying due respect to all of the people who are affected by data visualisation.

Workshop this week

This week — our last workshop where we'll be learning something new — I promised you a novel approach to data visualisation. We'll be learning how to scrape data from the internet — **without getting into trouble** — and how to create interactive data visualisations using the `plotly` package.

TASKS TO LOOK AT IN ADVANCE OF WEEK 11:

Core tasks:

- **Before lecture**: Read the assessment brief for assessment 2, including any articles you might be asked to critique. This will mean that you can come prepared for week 11. (I'm not expecting as much from you this week, as I know you've got imminent deadlines.)