

# Data Storytelling

**SMI105: Week 3**

**Dr. Calum Webb**

Sheffield Methods Institute, the University of Sheffield  
[c.j.webb@sheffield.ac.uk](mailto:c.j.webb@sheffield.ac.uk)

**Sign in**



# Learning outcomes

---

What will I learn?

---

By the end of this week you will know:

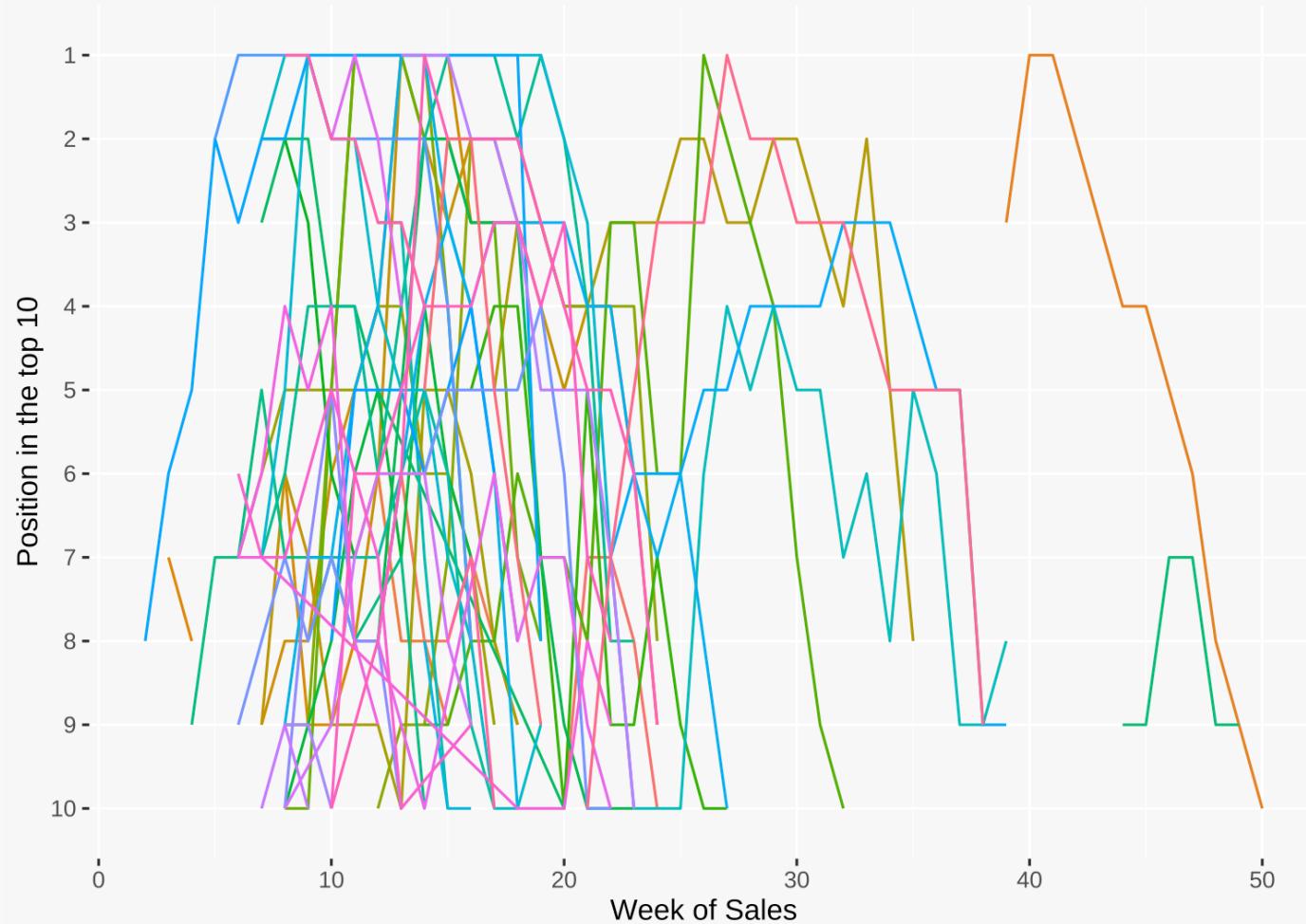
- Lecture: How to apply principles of storytelling to data visualisation and presenting data and critically assessing others' presentation of data.
- Workshop: How to extend our use of scatterplots in order to incorporate more information or manage complexity.

**Part 1**

# What is a story?



# The data underlying any particular data visualisation contains many possible stories



There were 53 different songs in the top ten during the year 2000, how do we choose what our story is here?

Consider the data backstory:

- Are there any particularly interesting patterns or anomalies?
- Is there something people will care about?

A **story** helps the audience to engage with the data and your data visualisations.

# What is a story?

We can think of a story of having a few essential components:

- An audience
- A plot
- A character/characters

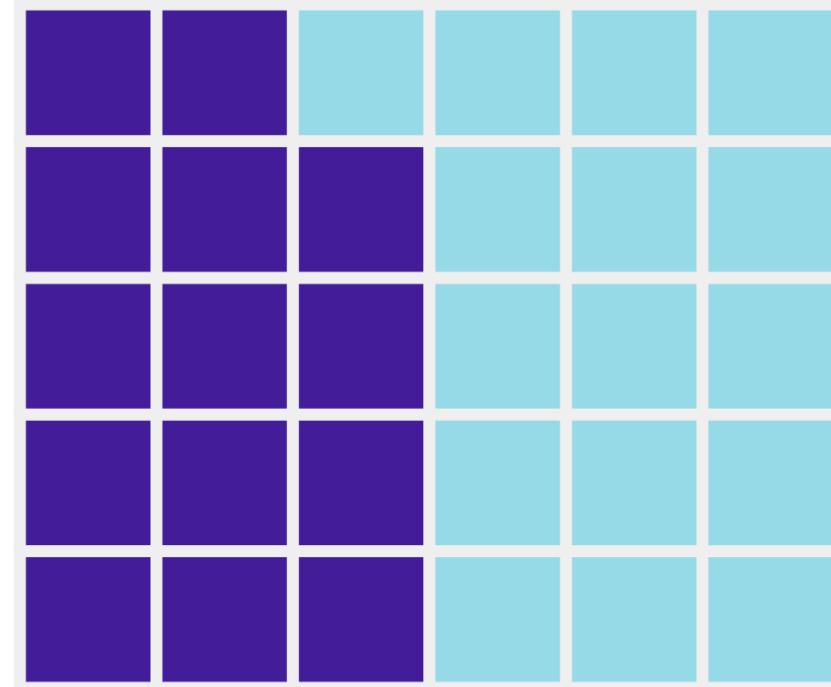
Feigenbaum and Alamalhodaei (2020), *the Data Storytelling Workbook*

# What is a story?

We can think of a story of having a few essential components:

- An audience
  - Different audiences will be drawn to different presentations of data.
  - Example 1: Establishes a connection at the personal, employee, level — people can *feel* how long 14 days is
- A plot
- A character/characters

**By the end of the year, the average Brit will spend 14 days — almost half a month — commuting to and from work.**



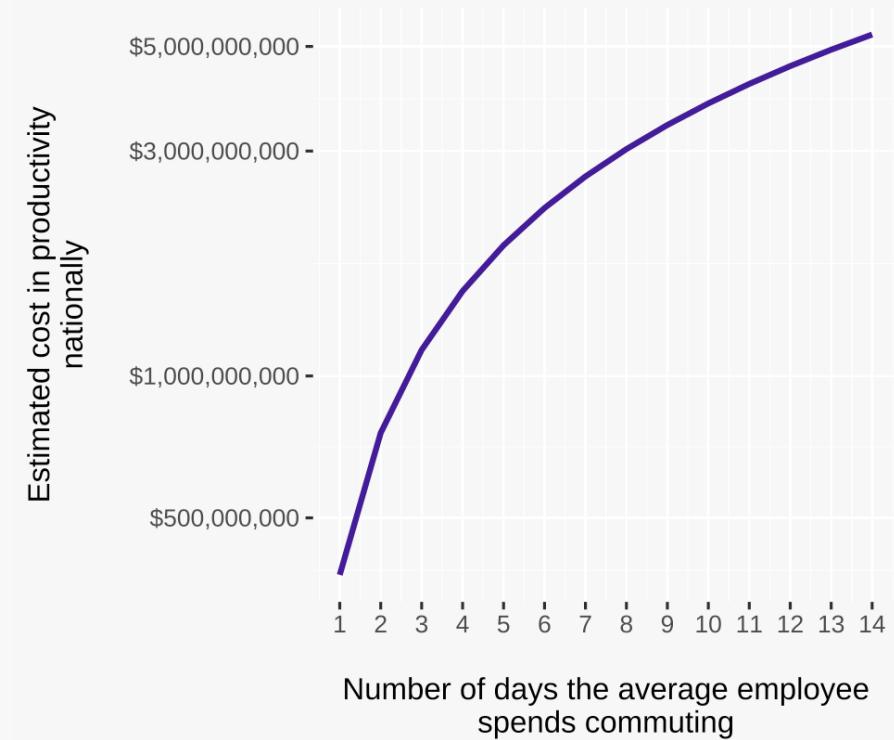
Feigenbaum and Alamalhodaei (2020), *the Data Storytelling Workbook*

# What is a story?

We can think of a story of having a few essential components:

- **An audience**
  - Different audiences will be drawn to different presentations of data.
  - Example 1: Establishes a connection at the personal, employee, level — people can *feel* how long 14 days is.
  - Example 2: £5.2billion is not really a scale of money most people can imagine — framing more targeted towards a politically informed audience, policymakers, or large businesses.
  - Data may be "framed" for different audiences, this can inform the kind of data visualisation you choose.
- A plot
- A character/characters

**Time spent commuting by UK workers is worth £5.3billion in lost productivity annually**



# What is a story?

We can think of a story of having a few essential components:

- An audience
  - Different audiences will be drawn to different presentations of data.
  - Example 1: Establishes a connection at the personal, employee, level — people can *feel* how long 14 days is.
  - Example 2: £5.2billion is not really a scale of money most people can imagine — framing more targeted towards a politically informed audience, policymakers, or large businesses.
  - Data may be "framed" for different audiences, this can inform the kind of data visualisation you choose.
- A plot
- A character/characters

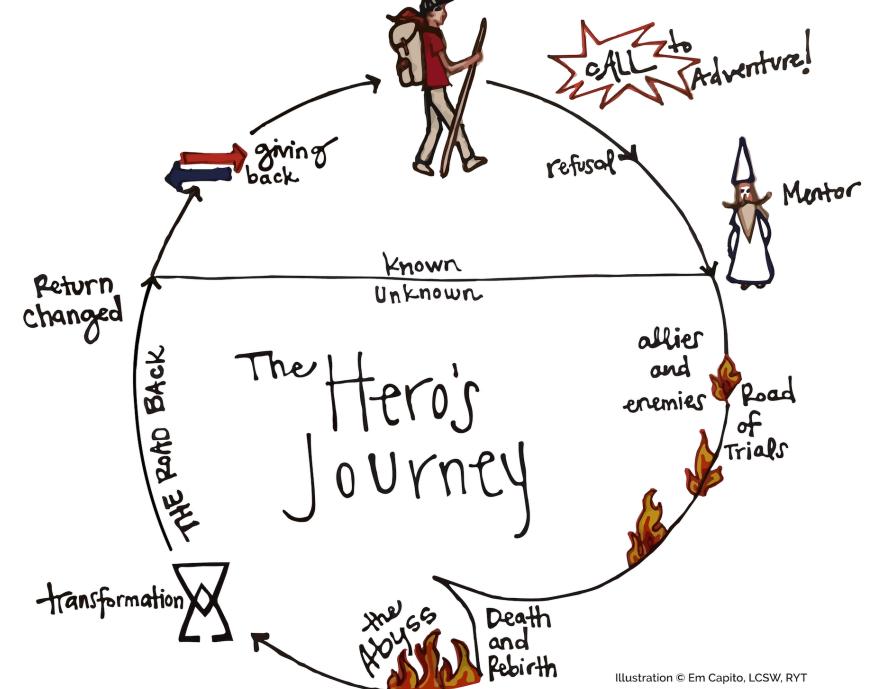
"Think, feel, do" Feigenbaum and Alamalhodaei, (2020).

- Who are your audience?
- What do they **know** already? What do you want them to **think** after viewing your data visualisation?
- How do you want them to **feel** after reviewing your data visualisation?
- What do you want them to **do**? What is your **call to action**?
- How much **time do they have to hear your story**?

# What is a story?

We can think of a story of having a few essential components:

- An audience
- **A plot**
  - Not a 'data visualisation' type of plot, but an actual story plot.
  - Not just **what** the story is, but **how you tell it**
- A character/characters



Feigenbaum and Alamalhodaei (2020), *the Data Storytelling Workbook*



# What is a story?

We can think of a story of having a few essential components:

- An audience
- **A plot**
  - Not a 'data visualisation' type of plot, but an actual story plot.
  - Not just **what** the story is, but **how you tell it**
  - Examples of using 'conventional' narrative story archetypes:
    - **Hero's Journey:** *Something disturbs our comfortable knowledge of a topic, we struggle to make sense of this new information, we transform our existing knowledge and return changed.*
- A character/characters

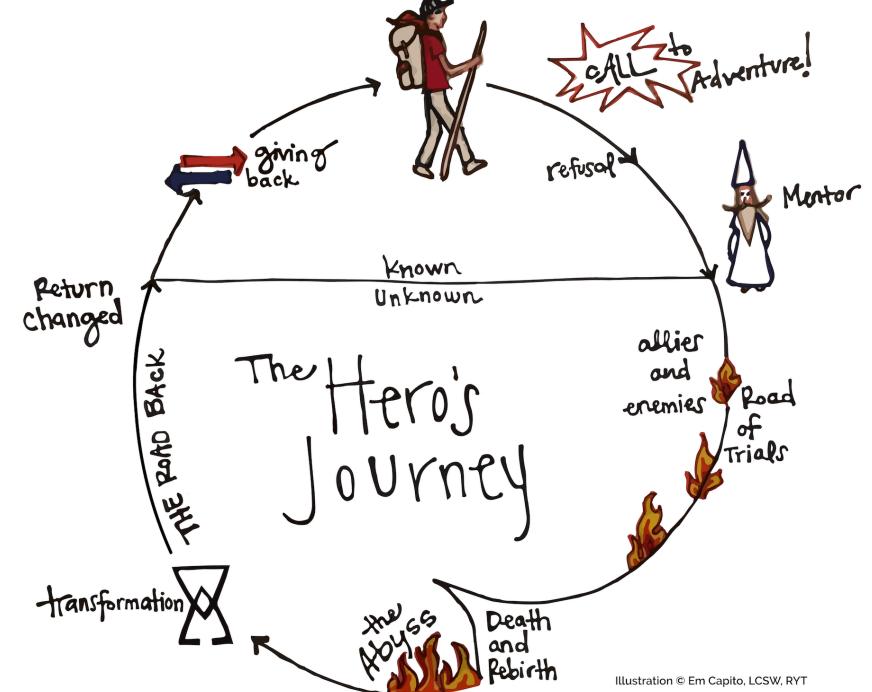


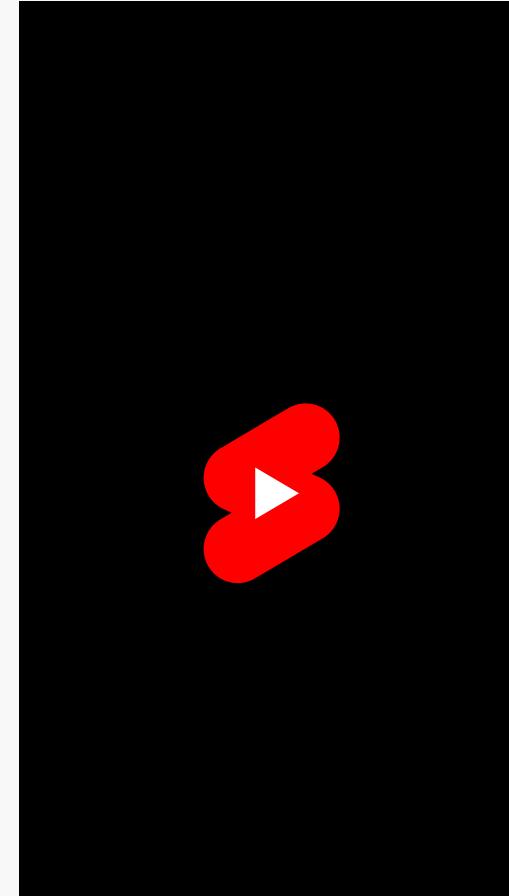
Illustration © Em Capito, LCSW, RYT

Feigenbaum and Alamalhodaei (2020), *the Data Storytelling Workbook*

# What is a story?

We can think of a story of having a few essential components:

- An audience
- **A plot**
  - Not a 'data visualisation' type of plot, but an actual story plot.
  - Not just **what** the story is, but **how you tell it**
  - Examples of using 'conventional' narrative story archetypes:
    - **Voyage and Return:** *We set out to a strange land not knowing what we might find, before returning with new wisdom and life experience. Typical for a lot of very "new" areas of research, e.g. new online movements (Think documentary)*
- A character/characters



Feigenbaum and Alamalhodaei (2020), *the Data Storytelling Workbook*

# What is a story?

We can think of a story of having a few essential components:

- An audience
- **A plot**
  - Not a 'data visualisation' type of plot, but an actual story plot.
  - Not just **what** the story is, but **how you tell it**
  - Examples of using 'conventional' narrative story archetypes:
    - **Mystery:** *There is a puzzling event or observation. You follow along with the investigator who uncovers clues and makes discoveries. At the end of the story there is a big revelation where all of the clues fit together.*
- A character/characters



# What is a story?

We can think of a story of having a few essential components:

- An audience
- A plot
- **A character/characters**
  - A person, an event (including a year), or a place can all be characters.
  - When there are many possible characters in a story, one or two can be used to act as representatives for a wider group in order to make it more relatable to an audience (e.g. Barbie and Oppenheimer)

1. **Event as Character** – 2020 Olympics: Climbing faces challenge for acceptance (BBC 22 September 2017 <https://www.bbc.co.uk/sport/olympics/41348585>)
2. **Person as Character** – Dina Asher-Smith: Why British sprinter could win Olympic gold at Tokyo 2020 (BBC 12 Aug 2018 <https://www.bbc.co.uk/sport/athletics/45163546>)
3. **Place as Character** – Tokyo 2020 Olympics: Japan debates daylight saving to avoid heat (BBC 6 Aug 2018 <https://www.bbc.co.uk/news/world-asia-45080980>)



# What is a story?

We can think of a story of having a few essential components:

- An audience
- A plot
- A character/characters
- **Dialogue**
  - It is also common to have some "dialogue" between data visualisation and either people affected by the issue or experts on the matter. This can be very effective for engaging audiences and making the subject more personally relatable.

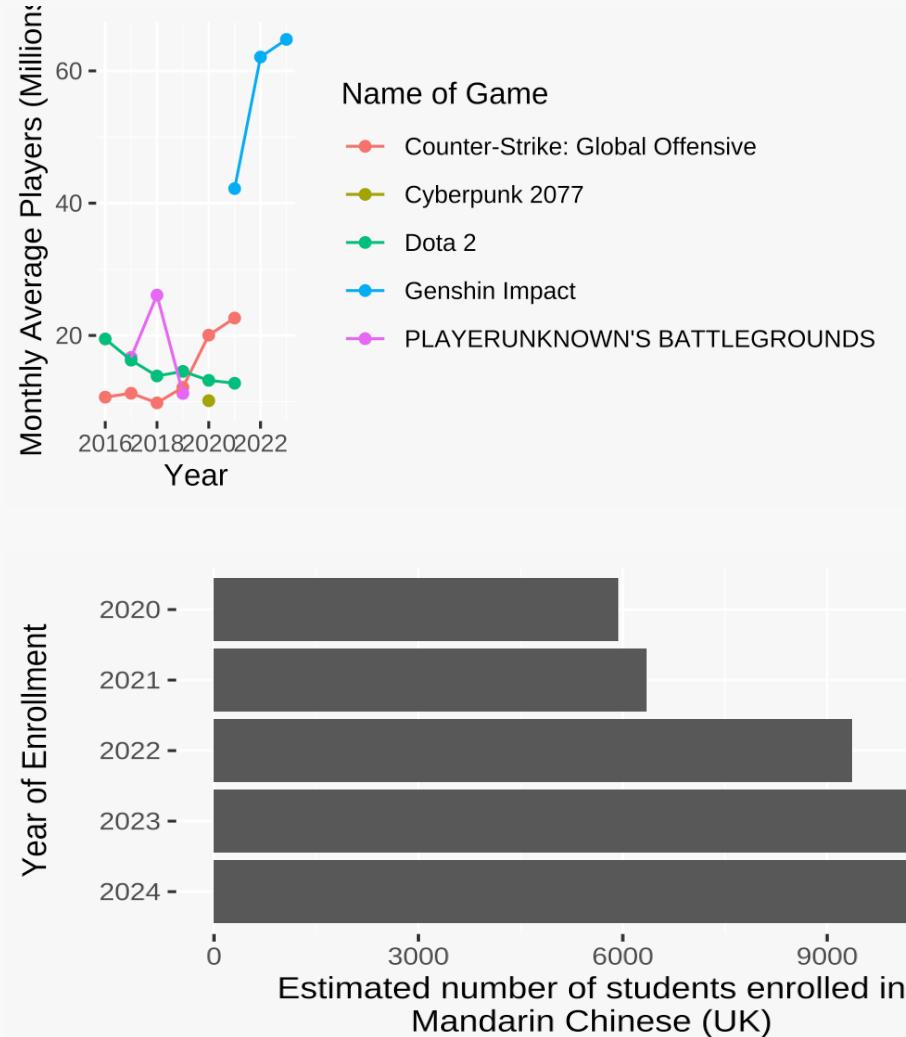


Feigenbaum and Alamalhodaei (2020), *the Data Storytelling Workbook*

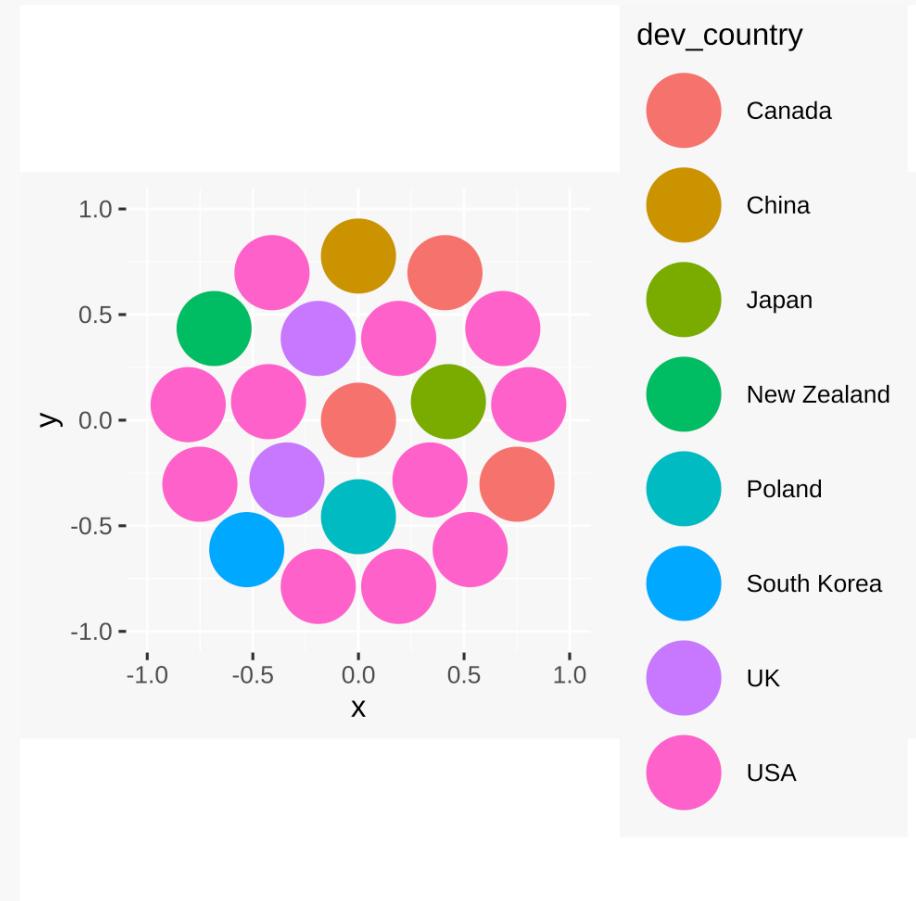
Part 2

Let's hear a story.



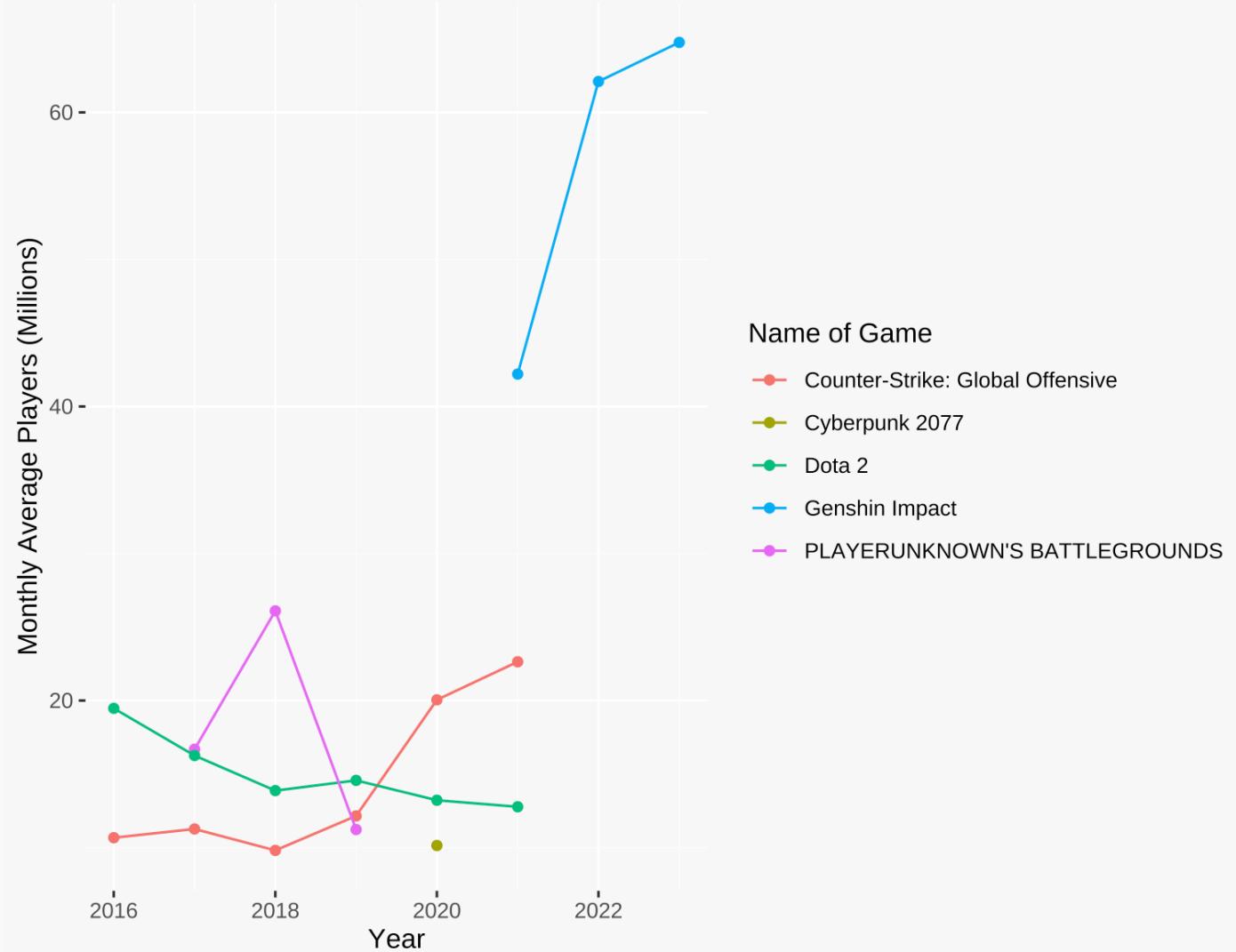


I'm going to tell a story using these three data visualisations.



# Let's start with this chart

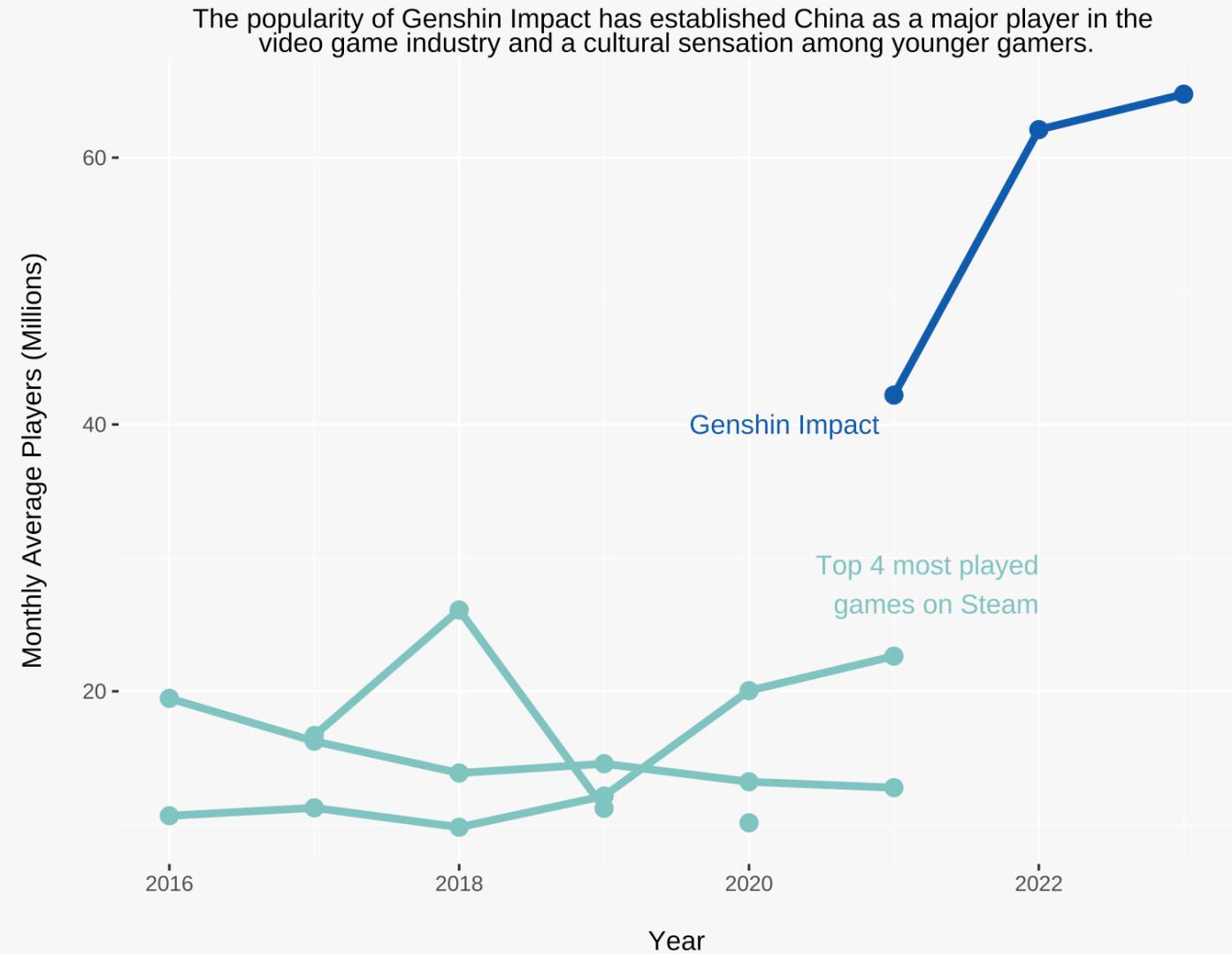
- What is interesting about it?
- How can I draw attention to what is interesting about it?
- How can I focus the audience on one specific "character"?



# Turning this visualisation into a story

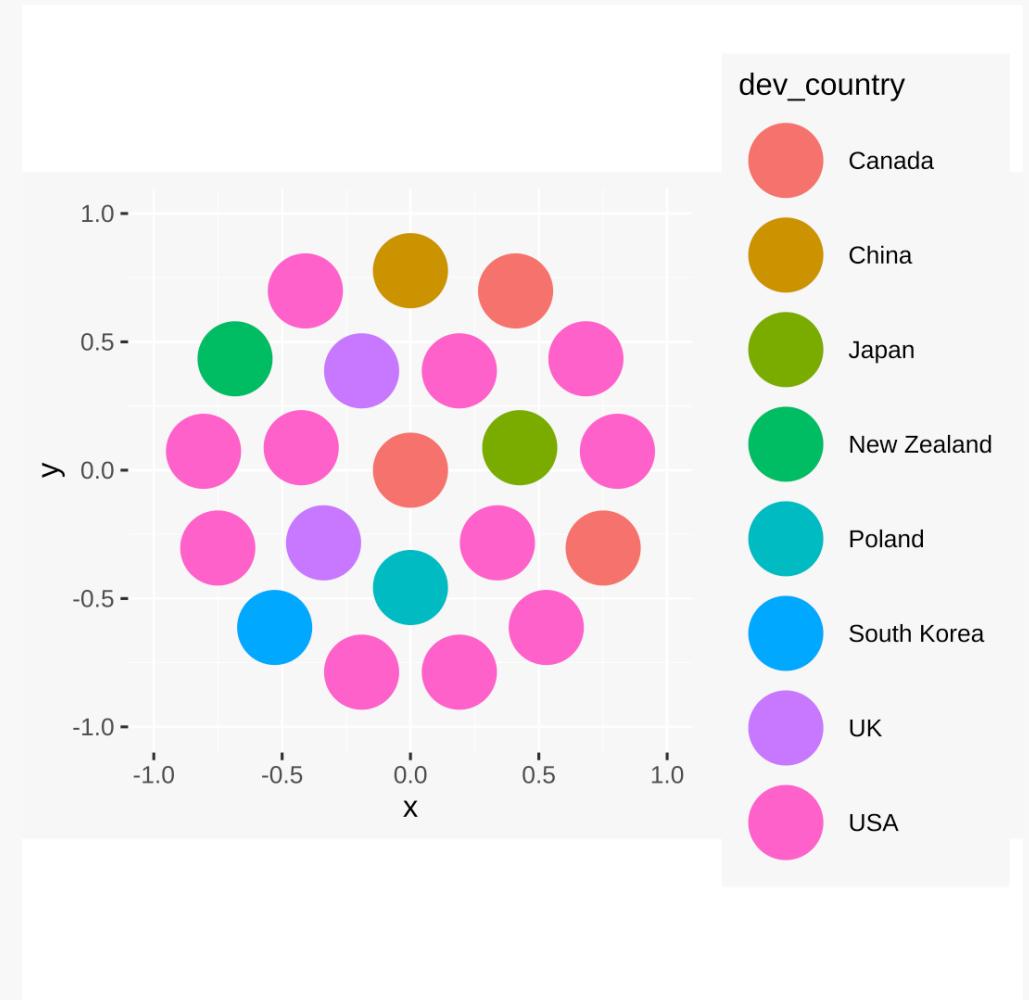
- Add an active title
- Highlight important characters

"Since its launch in 2020, the video game Genshin Impact, from Chinese developer miHoYo, has become a staple app on many young peoples' smartphones across the world. The scale of the game's success outstrips even the most successful PC games on the Steam platform in the years before."



# Developing our story

- Placing the data within a broader context
- Highlighting variation between two things
- Why is this new and interesting?



## Developing our story

- Improved communication by using instantly recognisable flags, avoiding the need for a key
- Placed the dot representing Genshin Impact (the Chinese studio developed game) in the centre of the visualisation, rather than at the top, to make it the focal point (the main character).

Taking a wider view, few of the most popular games on the Steam platform have been developed in China, with the United States and Canada dominating the 'top-played' list.

*"Other games have Chinese translations and can be loads of fun, but very few (probably none) have such high production values with such a vast amount of high quality relatable dialogue/audio for \$0 up front." - Genshin Impact Player*

The most popular 20 video games have been dominated by United States and Canada-based developers.



Few games, however, have also generated so much interest in Chinese culture, as games from Japan did for Japanese culture in the 1980s and 1990s:

*"It also means there are cultural norms hidden in it because it was created by Chinese citizens in China. ... there is a bunch of simple scenarios mixed in with the Dragons, Infinite teapots and dimension travelling."* - Genshin Impact Player





# Genshin Impact × McDonald's

Exclusive Rewards  
Available on the McDonald's App

09/17 ▶ 09/29

ONLY IN THE U.S.

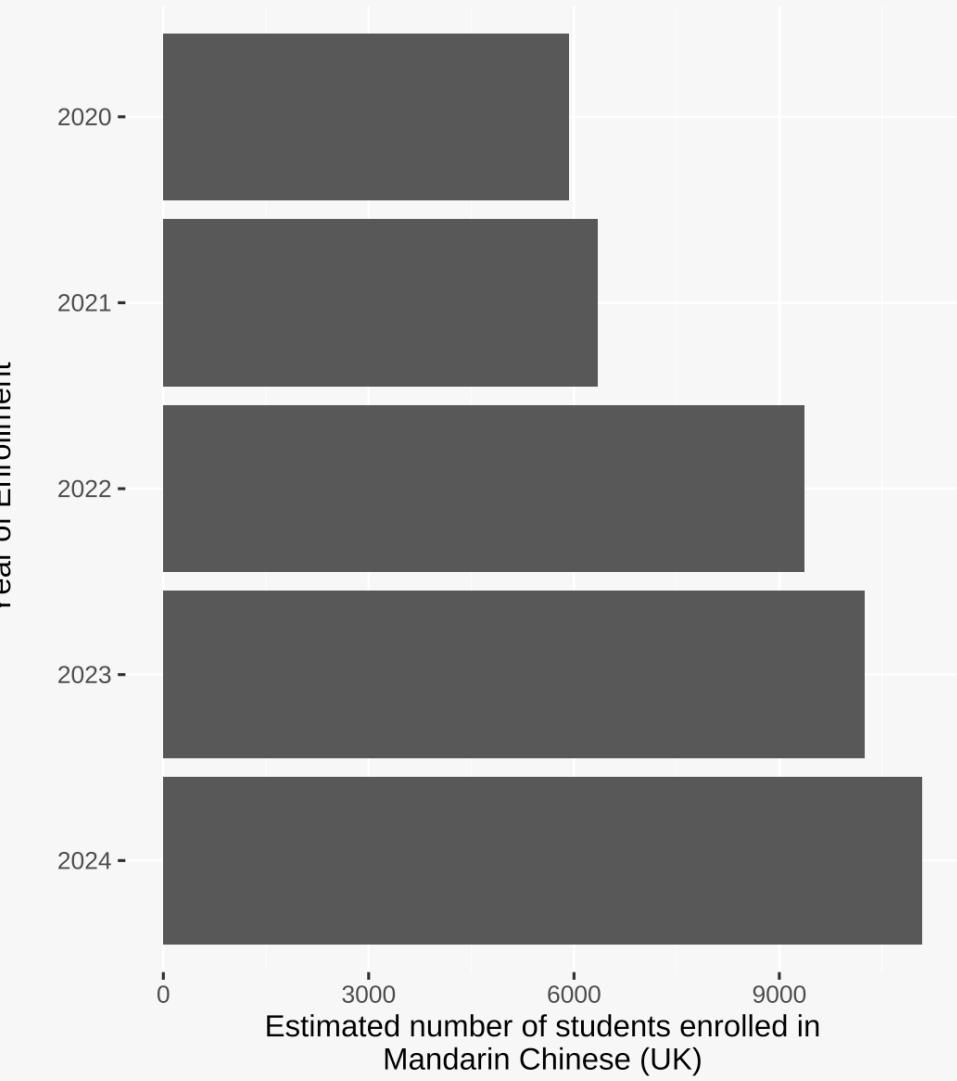
Valid 1x per apple pie or Genshin Meal purchase at participating McDonald's for a limited time.  
Must opt in to email and Rewards before purchase.

(C) McDonald's 2024.  
(C) COGNOSPHERE.



# Call to action: changing the way people think about video games

- How to draw a link between two pieces of data in a way that tells an interesting story?
- What message do I want people to take away that will change their thinking or behaviour?



# Call to action: changing the way people think about video games

- Cross-reference the 'character' from the first two visualisations at a specific point in the data on a different visualisation
- Add frequency counts to the bars to make the change more of the focus

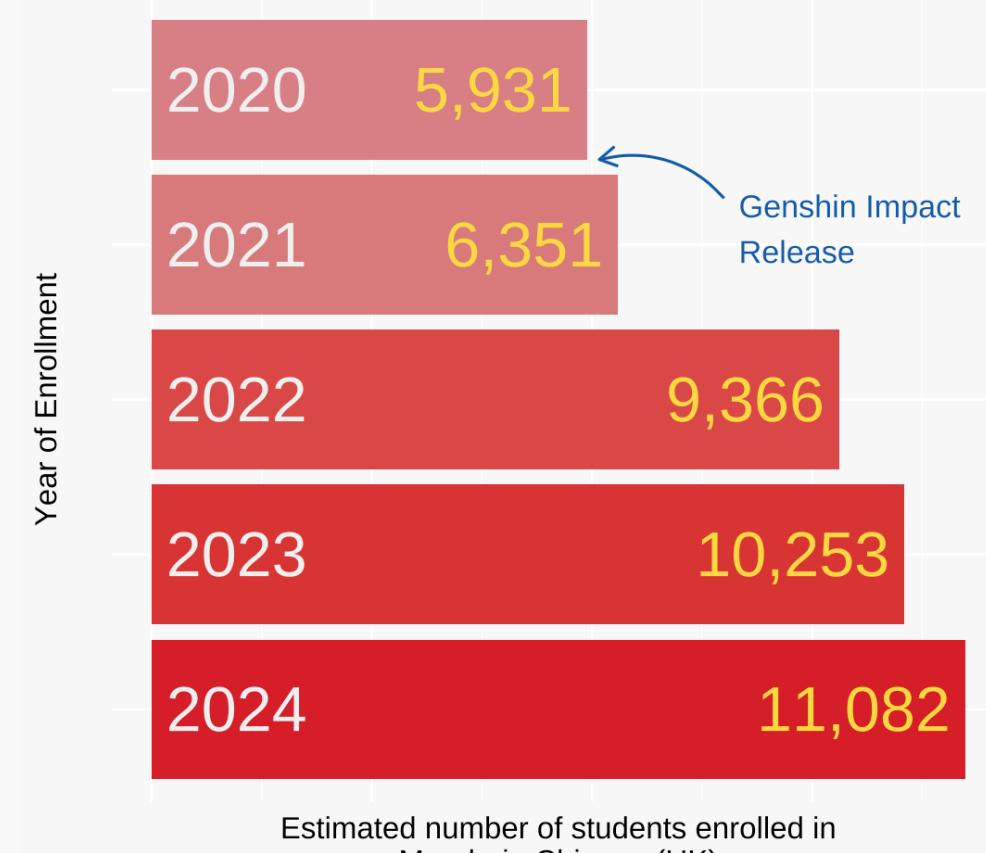
The success of Genshin Impact has coincided with a large increase in the numbers of young people studying for GCSE and A-level qualifications in Mandarin Chinese in the UK in a way that other strategies to increase uptake in the language have failed to accomplish.

*"[A]fter I started learning Chinese, I switch from JP to CN audio. It's helpful. And at the same time you're learning and enjoying the game." - Genshin Impact Player*

*"I've always been a bit interested in Chinese culture. I'm not gonna lie when I say that Genshin Impact introduced me to its culture in a new light." - Genshin Impact Player*

*"Genshin got way more people interested in Chinese culture." - Genshin Impact Player*

There has been significant growth in the number of young people in the UK studying Mandarin Chinese in secondary education since the release of the game.



# Activity: Breaking down the storytelling of "The Genshin Impact on young peoples' interest in Chinese language in the UK"

- Go to [app.wooclap.com/TKOOEY](https://app.wooclap.com/TKOOEY) and join the first event.

Think about the following:

- Who was the audience? How did I want them to change their views after the story?
- What kind of storytelling plot archetype might this data story have fallen into?
- Who/what could be considered a character in this story?
- How could the audience and plot be changed if I presented the last plot in my story first instead?



# The rest of this week:

## This week's workshop:

Let's extend graphs where we show the relationship between a continuous and a categorical variable: box plots, density curves, histograms, and so on. What's the easiest way to understand this information? Why would we use one of these graphs over another?

# The rest of this week:

## This week's workshop:

Let's extend graphs where we show the relationship between a continuous and a categorical variable: box plots, density curves, histograms, and so on. What's the easiest way to understand this information? Why would we use one of these graphs over another?

## Before week 4:

### Core tasks:

- **Before lecture:** Bookmark or download a copy of the FT's Visual Vocabulary: <https://github.com/Financial-Times/chart-doctor/blob/main/visual-vocabulary/poster.png>
- **Before lecture:** Find a chart in the news that relates to the topic given to you at the end of week 3 lecture. Then, share on Piazza:  
(i) the chart you find, (ii) where it was published (if that's not obvious) and (iii) which category from the 9 listed across the top of FT Visual Vocabulary you think best describes the data relationship in it.
- **Before workshop:** Work through the task at the end of the workshop handout from week 3 Supplementary tasks:
- **Before workshop:** Read chapter 3/Make a plot of the Healy book