

STORYTELLING WITH DATA

PATRICK BOILY | STEPHEN DAVIES | JEN SCHELLINCK



COURSE DESCRIPTION

This course is for participants who wish to become savvy consumers of data presentations and learn how to work with teams to communicate useful, evidence-based messages and ideas. The focus is on understanding how dashboards, reports and data visualizations convey descriptive, diagnostic, predictive, and prescriptive data insights to decision-makers.

COURSE DESCRIPTION

Topics covered include:

- identifying and gathering data presentation requirements
- storyboarding
- critiquing dashboards and data presentations
- definition of a story
- relationship between data and story
- identifying and holding the audience's attention
- lessons in storytelling with data
- anatomy of dashboards and data visualizations
- effective visuals
- decluttering
- thinking like a designer with the Gestalt principles
- making data presentations accessible

LEARNING OBJECTIVES

Understand why stories are important in the communication of data and information

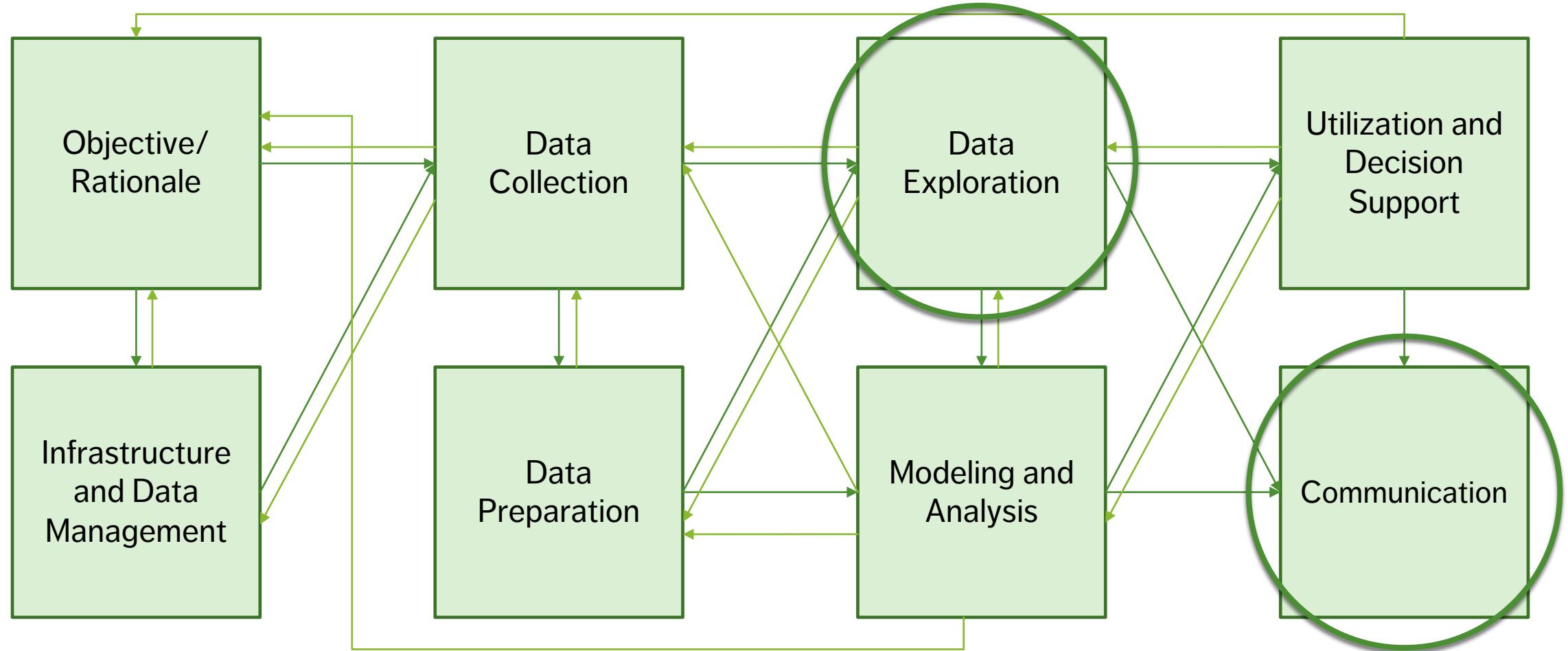
Understand how data stories differ from other types of stories

Identify best practices when building visualizations for storytelling

Understand what tools are useful when building data stories

Learn techniques for maximizing the effectiveness of storytelling visualizations

THE (MESSY) ANALYSIS PROCESS



OUTLINE

Part I – Stories and Storytelling

1. Stories & Storytelling
2. Storytelling Risks
3. Elements of Storytelling
4. Story Forms and Structures
5. Data Stories
6. Visual Storytelling

Part II – Effective Storytelling Visuals

7. Data Visualization ABC
8. Data and Stories
9. Evolving a Storytelling Chart
10. Anatomy of Storytelling Dashboards
11. Chart Aesthetics & Accessibility
12. Data Stories in the Wild

RECOMMENDED...

Something to take notes on (about e.g., definitions, story examples)

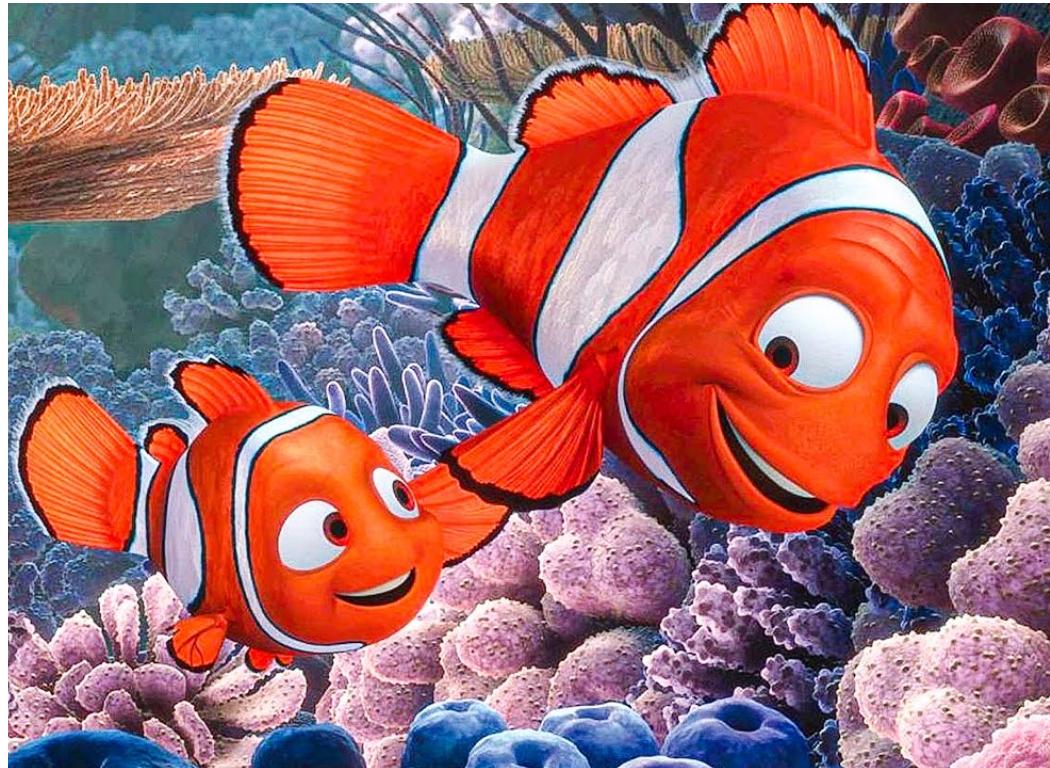
Access to Zoom emojis for polling and interaction purposes

Story listening props (popcorn?)

PART I – STORIES AND STORYTELLING

STORYTELLING WITH DATA

STORIES AND STORYTELLING



There once was a fish named Marlin, who loved his son Nemo more than anything. Every day he tried to protect Nemo from the ocean, which Marlin feared.

One day Nemo decided his dad was wrong and he swam away. But he was captured by a diver.

Because of that, Marlin had to leave the safety of his reef to find his son.

And because of that, he learned to let go of his fears and trust that Nemo had what it takes to take care of himself.

Ever since that day, Marlin gave Nemo the space to learn on his own.

STORIES AND STORYTELLING

Is this a **story**?

If so, what is its **purpose**?

Its **moral or message**?

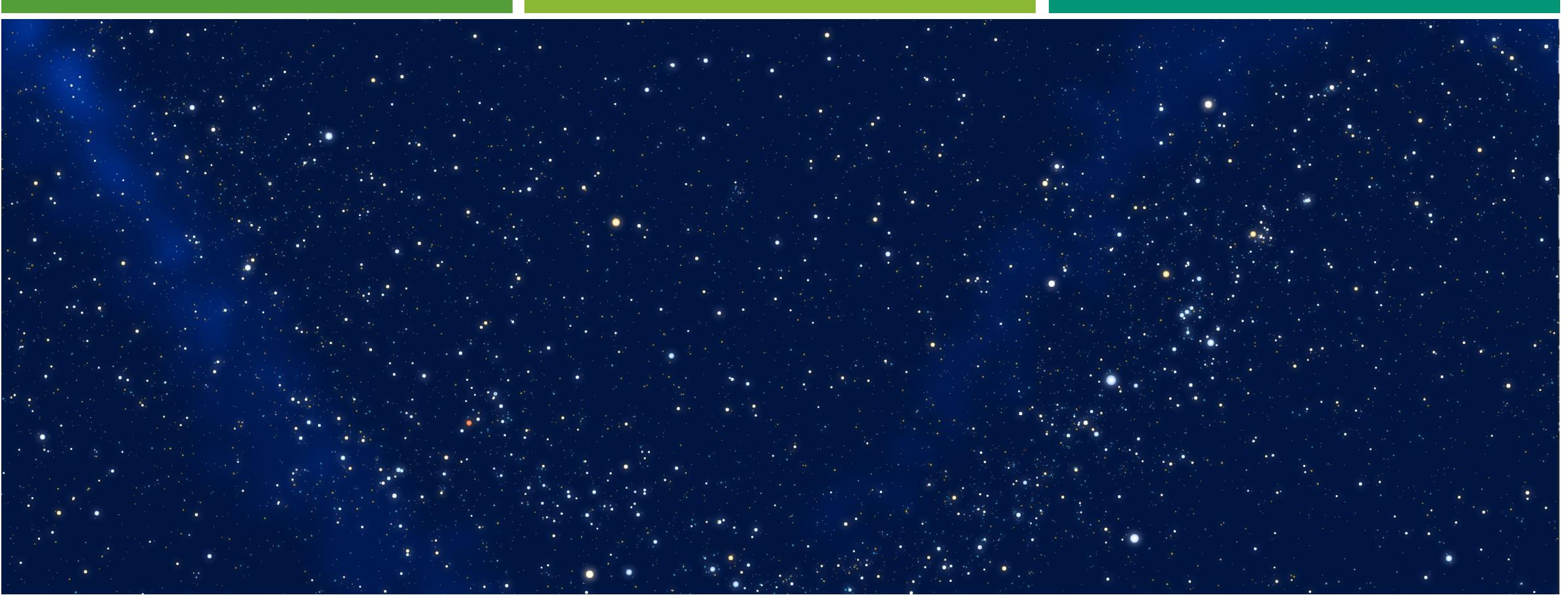
WHY STORIES?

Human beings are **social animals**, and they **communicate with each other**.

Communication is an evolutionary trick, which played a crucial role in the brain's development and its ability to house a mind: transfer of ideas is **much quicker** than the transfer of genes.

And how do we communicate? **We tell stories**.

Are there other ways to communicate? Other modes? Other tools?



A SKY FULL OF STARS

We have always had a drive to paint stories onto the Universe. When humans first looked at stars, which are great flaming suns an unimaginable distance away, they saw amongst them giant bulls, dragons, and local heroes. [...] Humans think in stories. [Cohen, Pratchett, Stewart]

STORY TIME

Queen Cassiopeia was the wife of King Cepheus of Ethiopia. She boasted that she was more beautiful than the Nereids, the 50 sea nymphs. They were enraged by her comments and appealed to Poseidon to punish Cassiopeia for her boastfulness.

The sea god obliged and sent Cetus, a sea monster, to ravage the coast of Cepheus' kingdom. Cepheus turned to an oracle for help: in order to appease Poseidon, he and Cassiopeia had to sacrifice their daughter Andromeda to the sea monster.

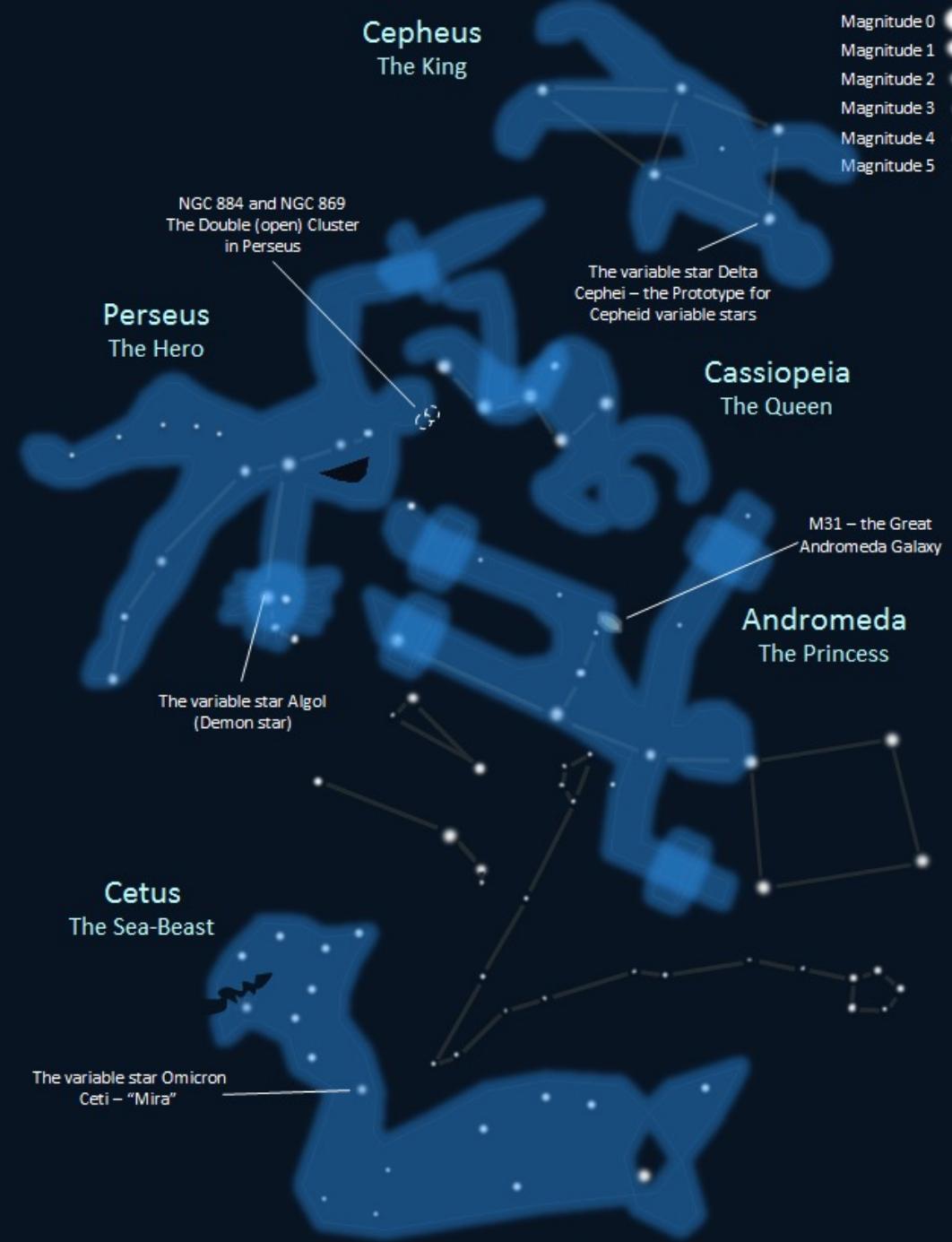
They did so reluctantly, leaving Andromeda chained to a rock for Cetus to find. However, she was saved in the last minute by Perseus, a Greek hero.

STORY TIME

Perseus and Andromeda were later married. At the wedding, one of her former suitors claimed that he was the only one who had the right to marry her.

There was a fight and Perseus, outnumbered, used the head of Medusa to defeat his opponents. One look at Medusa's head turned them all into stone. The king and queen also met their end.

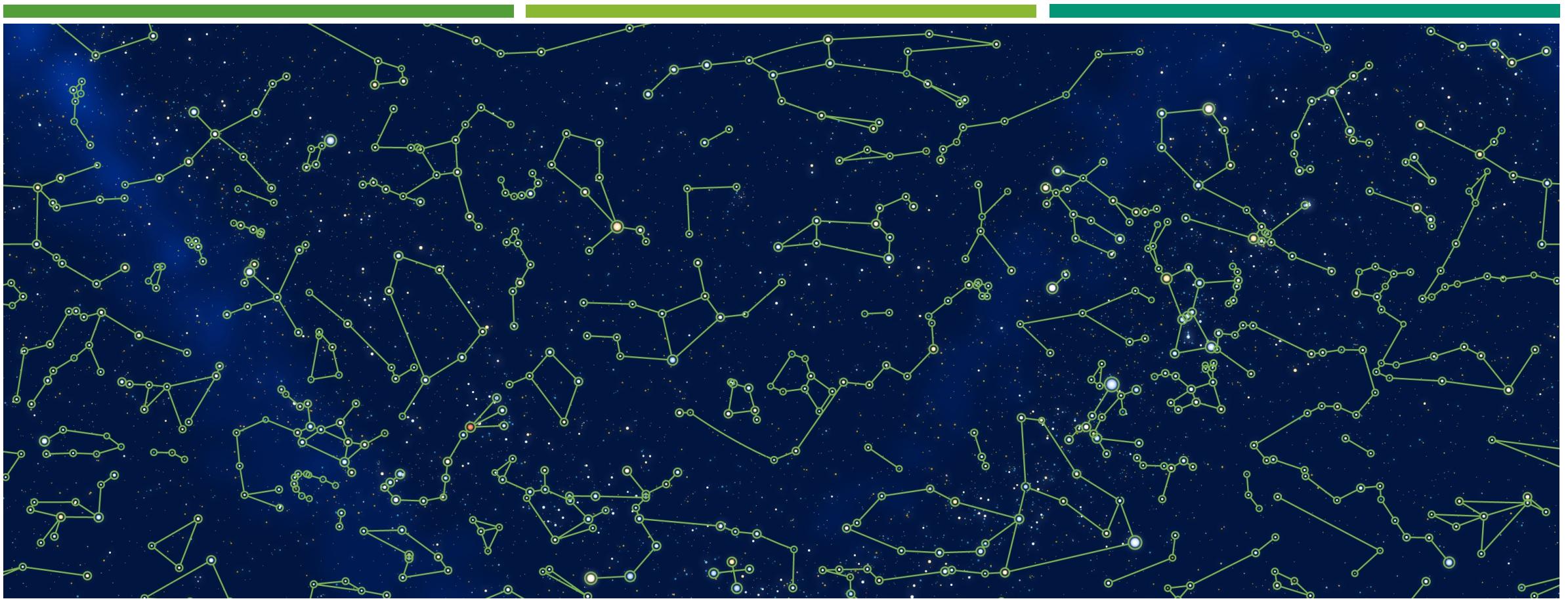
Poseidon then placed Cassiopeia and Cepheus in the sky: Cassiopeia was condemned to circle the celestial pole forever, and spends half the year upside down in the sky as punishment for her vanity.





A SKY FULL OF STARS – ANCIENT GREEK STORIES

We have always had a drive to paint stories onto the Universe. When humans first looked at stars, which are great flaming suns an unimaginable distance away, they saw amongst them giant bulls, dragons, and local heroes. Humans think in stories. [Cohen, Pratchett, Stewart]



A SKY FULL OF STARS – WESTERN [MODERN] CONSTELLATIONS

We have always had a drive to paint stories onto the Universe. When humans first looked at stars, which are great flaming suns an unimaginable distance away, they saw amongst them giant bulls, dragons, and local heroes. Humans think in stories. [Cohen, Pratchett, Stewart]

WHAT IS A STORY?

To paraphrase U.S. judge Potter Stewart: “I may not be able to define what a story is, but I know one when I see one”.

A **story** is the telling of a temporal sequence of “events”, either true or fictional. It is “told” so that the audience experiences or learns something from it. It is a means of transferring information, experiences, attitudes, or points of view. [M.W. Travis, *The Wrap*]

Stories are used to **explain, describe, argue, persuade, teach, entertain**, etc.

STORIES AS MEMES

Stories are **memes** (in the Dawkins sense): ideas, behaviours, styles

- spreads by means of imitation from person to person within a culture
- often carries symbolic meaning representing a particular phenomenon or theme.

Memes act as **unit** for carrying:

- cultural ideas, symbols, or practices,
- transmitted from one mind to another through writing, speech, gestures, rituals, etc.

Memes are cultural analogues to **genes**:

- they self-replicate, mutate, and respond to selective pressures

PRACTICAL DEFINITION OF STORIES

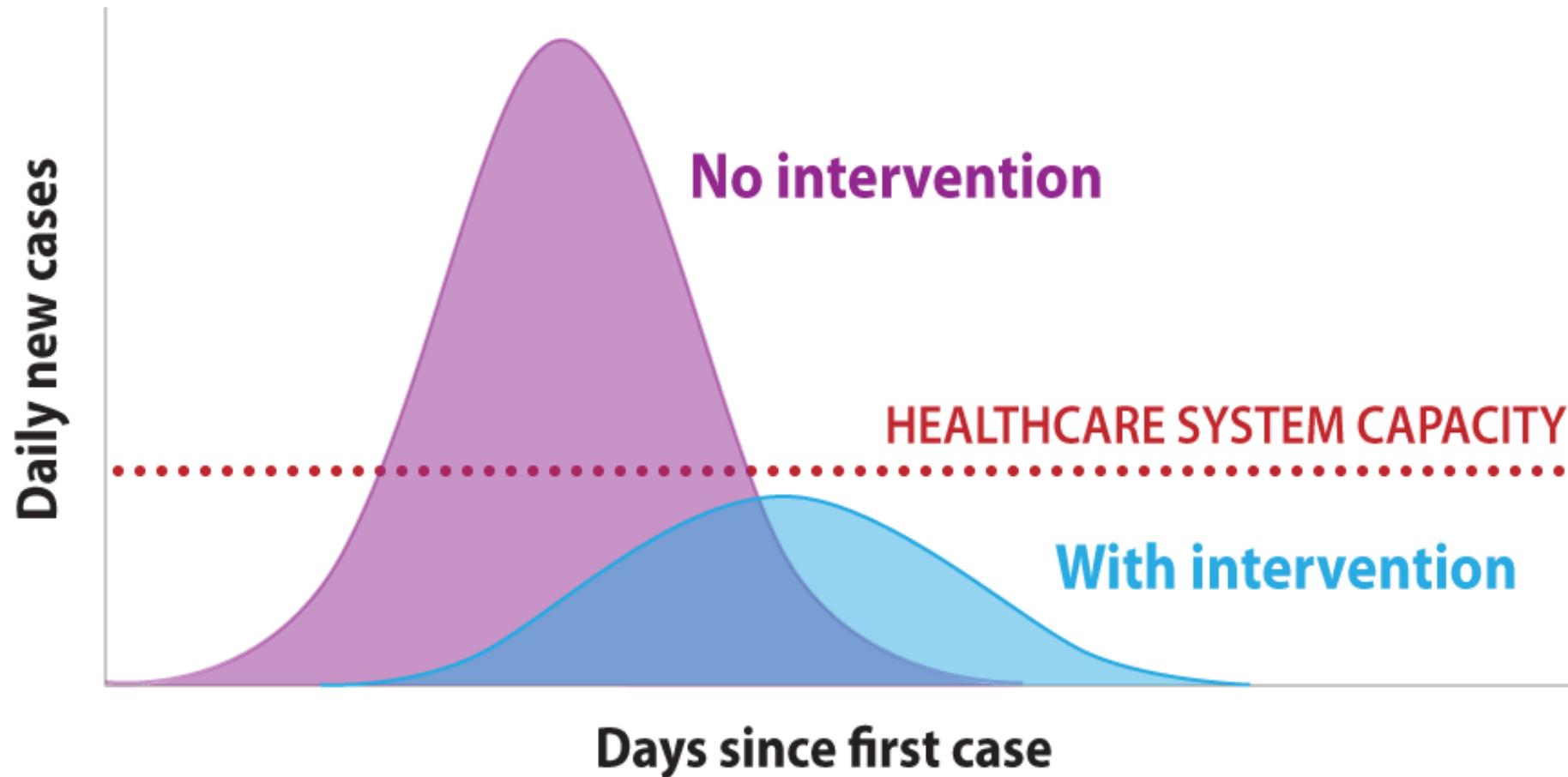
A story consists of:

- context,
- actors (typically),
- location(s) (sometimes),
- series of events, and
- outcome, result, consequence, or resolution.

Is the next slide showing a story according to this definition?

FLATTENING THE CURVE

A look at the importance of slowing the spread of a virus, so that the rate of infection doesn't outpace the resources to fight against it.



WHERE ARE STORIES FOUND?

- news
- books, magazines
- art and music industry
- television, movie studios, Netflix, HBO, Disney+, etc.
- social media: Facebook, Instagram, Snapchat, etc.
- sports and video games
- evidence: data, science experiments, etc.
- religions, ideologies, belief systems, etc.
- enduring coherent groups: cultures, countries, cities, etc.
- commerce: adverts

WHERE DO STORIES ORIGINATE?

In popular culture stories can be fictional, real work or a blend of both.

In the workplace we can think about stories in a focused way (**active** and **potential**).

Active stories – these are happening all of the time around us, we need to understand context to understand if they are important (are they worth telling).

- Some already have outcomes – “We got our funding”
- Some outcomes are yet to happen – “Will we get our funding? What happens if we don’t?”

Potential stories – these are stories we want to happen. We craft these stories to attempt to convince the actors to a specific conclusion – “please give us some funding”

EXERCISE – INTERPRETING A STORY

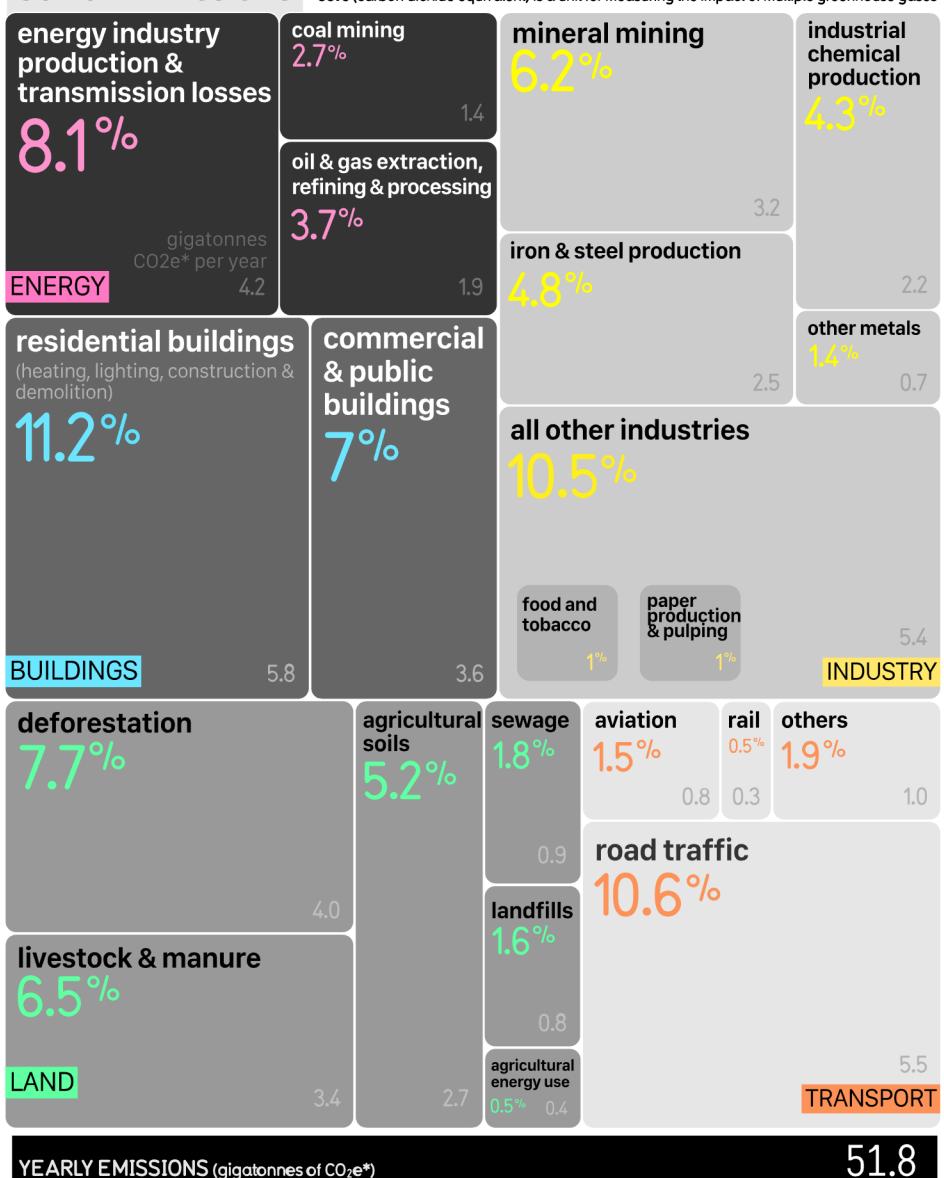
Group discussion:

- Is the dashboard on the next slide telling a story?
- If so what is the story and how easy is it to work out what the story actually is?

How Do We Get to Zero Greenhouse Gas Emissions?

ENERGY INDUSTRY BUILDINGS TRANSPORT LAND OTHER

Current Emissions



YEARLY EMISSIONS (gigatonnes of CO₂e*)

Emissions from where they are used, not where they are produced.

version 1.0 / May 2019

created by David McCandless, Duncan Geere, Hazel Healy, New Internationalist Magazine

Halving by 2030

gigatonnes reduced per year

increased solar energy	-1.1
widespread wind energy	-1.4
other renewables	-0.7
better recycling of raw materials	-3.2
less materials to make the same things (product materials efficiency)	-1.0
disassembling old products to make new ones (circular business models)	-1.1
waste reduction in production of steel, plastics and other industrial materials	-1.9
reduced use of refrigeration gases (CFCs and HFCs)	-1.6
low-carbon heating and cooling	-2.1
low-carbon construction	-0.9
retrofitting buildings with better insulation, energy efficient lighting, etc.	-1.5
automation of temperature and lighting	-0.4
electric vehicles	-1.3
increased use of public transport	-0.7
bikes, car-sharing, scooters	-0.5
more efficient shipping of goods with reduced air transport	-1.0
low-emission trucks	-0.5
halting deforestation, planting trees	-2.0
sustainable agriculture techniques	-0.9
plant-based diets	-1.7
reduced food waste	-0.7
other measures	-0.9

EMISSIONS REMAINING

24.3

Policies to Zero by 2050

Global Carbon Tax / Carbon Pricing	<ul style="list-style-type: none"> A progressive tax on fossil fuel producers and users dramatically reduces global emissions.
Energy	<ul style="list-style-type: none"> % renewable electricity by law and subsidies. Early retirement of fossil-fuel power plants. Grid-scale electrical storage. Reduced individual consumption in richer countries.
Industry	<ul style="list-style-type: none"> New efficiency standards. Switch from coal to biogas, biomass and other sustainable alternatives. Facilities designed to reduce waste.
Buildings	<ul style="list-style-type: none"> Solar incentives. Electrification of heating, furnaces, stoves. Higher energy-efficiency standards.
Transport	<ul style="list-style-type: none"> Fossil fuel vehicles phased out. Taxes on inefficient vehicles and fuel. Higher fuel economy standards.
Land	<ul style="list-style-type: none"> Methane capture and destruction. Improved forest and livestock management. More reforestation.
Justice and Equity	<ul style="list-style-type: none"> Climate finance flows to Global South. New jobs and training for affected workers. International co-operation secures strong climate deal.

EMISSIONS REMAINING

0.0

EXERCISE – STORIES FROM YOUR WORKPLACE

Write down 2 or 3 examples of stories from your workplace

- Were there entities (actors) involved, e.g., people, groups or organizations.
- What key events happened, e.g., decisions, crisis, delays.
- Was there a location (geographic or maybe part of an organization)
- Was there a sequence (narrative) of these events, e.g., a wrong decision caused a delay
- Was there an outcome, e.g., success, failure, business as normal

Example: our operations group delayed an investment decision which in turn delayed hiring which finally caused a key project to fail. It's not the most exciting story but it has most of the elements. **Actors** (operations group), **events** (investment, delay, crisis) and **narrative** leading to an **outcome** (failure)

STORYTELLING CONCLUSIONS?

In a workplace context we are:

- Uncovering active stories that are unfolding as we work (e.g., the project I am on just got cancelled)
- Crafting potential stories to achieve/obtain an outcome (e.g., creating and presenting a TB Submission to obtain funding)
- Working out what possible outcomes an in-process, active story might have (will I have a higher likelihood of success if I do A instead of B?)
- In all cases attempting to identify the context, actors, events, outcomes will help us to identify the data we need to create the data story which, in turn helps us to tell the actual story (we hope!)

STORYTELLING RISKS

A good story can help shed insights on a situation, but storytelling requires **choices**, and the outcome is affected by what is **included** and what is **omitted** in the telling.

It is easy to mislead by **accident**; it is also easy to mislead by **design**.

With data stories, there is an additional complication: we usually only have access to the **available data**. The data that was not collected is, by definition, not available. Some of the data that was collected may also be unavailable for a variety of reasons.

This implicit bias can lead to compelling yet **fundamentally flawed** data stories.

During WWII, mathematician **A. Wald** undertook a study to help protect British bombers flying over enemy territory.



Data included: the **number** and **location** of **bullet holes** on returning aircraft, and the goal was to use this information to determine where to add armor to best protect the plane's structure.

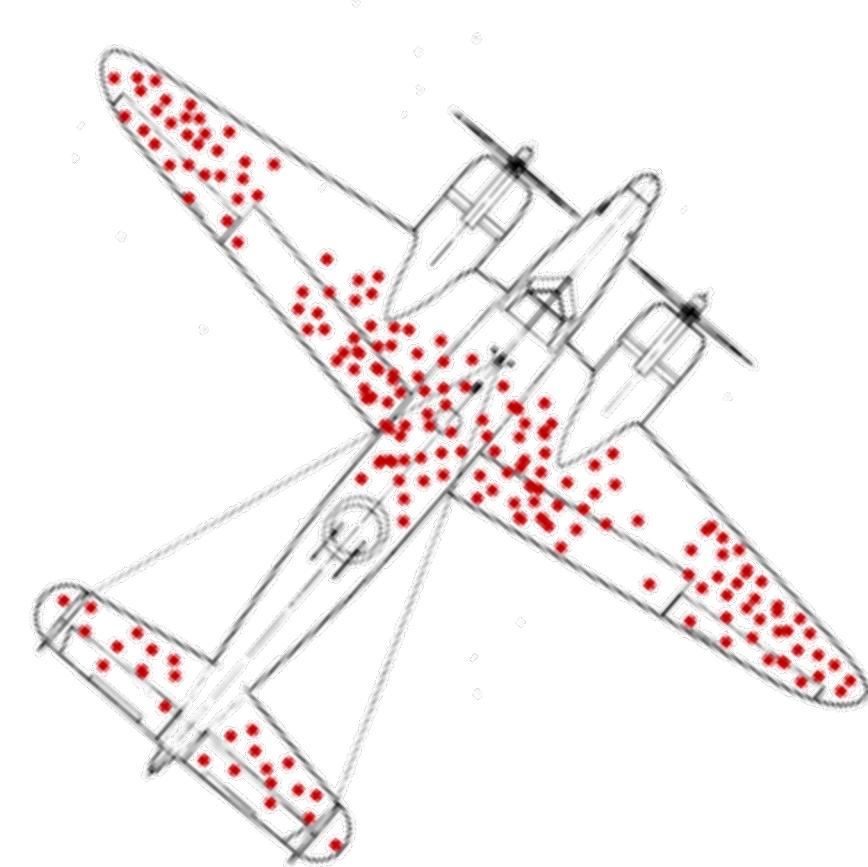
A chart was created to show where the maximum number of bullet holes were located on **returning aircraft**. This chart showed greatest damage on the **aircraft extremities**, not on the main wing and tail spars, engines, and core fuselage areas.

STORYTELLING RISKS

As such, the Air Ministry wanted to add armor to the **extremities**. Wald suggested they were **dead wrong**.

To avoid “**survivorship bias**”, armor should be added to the areas with the **fewest holes**: if no returning planes had holes in their wing spars and engines, then even a few holes in those locations were **deadly**.

Take-Away: the data that is missing may be as important to story than the data that is there. Storytelling is not always an obvious endeavour.



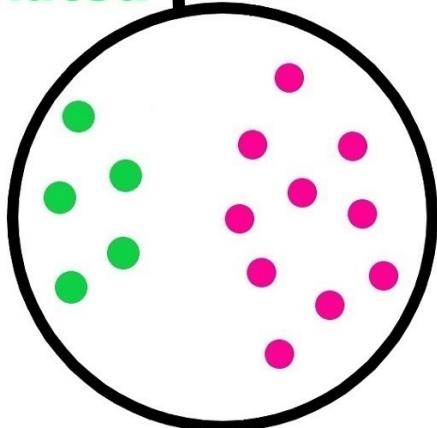
STORYTELLING RISKS

“Open any newspaper, watch any TV news show, and you find experts who forecast what's coming. Some are cautious. Most are bold and confident. A handful claim to be Olympian visionaries able to see decades into the future. With few exceptions, they are not in front the camera because they possess any skill at forecasting.

Accuracy is seldom even mentioned. [...] The one undeniable talent that talking heads have is their skill at **telling a compelling story with conviction**, and that is enough. Many have become wealthy peddling forecasting of untested value to corporate executives, government officials and ordinary people who would never think of swallowing medicine of unknown efficacy and safety but who routinely pay for forecasts that are as dubious as elixirs sold from the back of a wagon.” [Tetlock]

Hospitalized with Covid

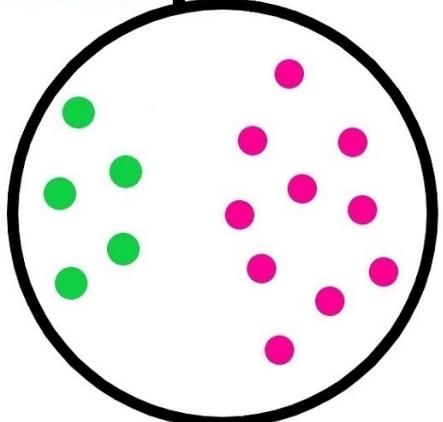
Un-
vaccinated Vaccinated



More vaccinated than
unvaccinated people
in the hospital

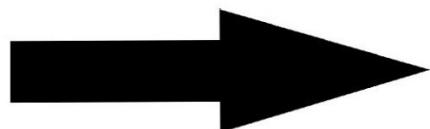
Hospitalized with Covid

Un-vaccinated Vaccinated



More vaccinated than unvaccinated people in the hospital

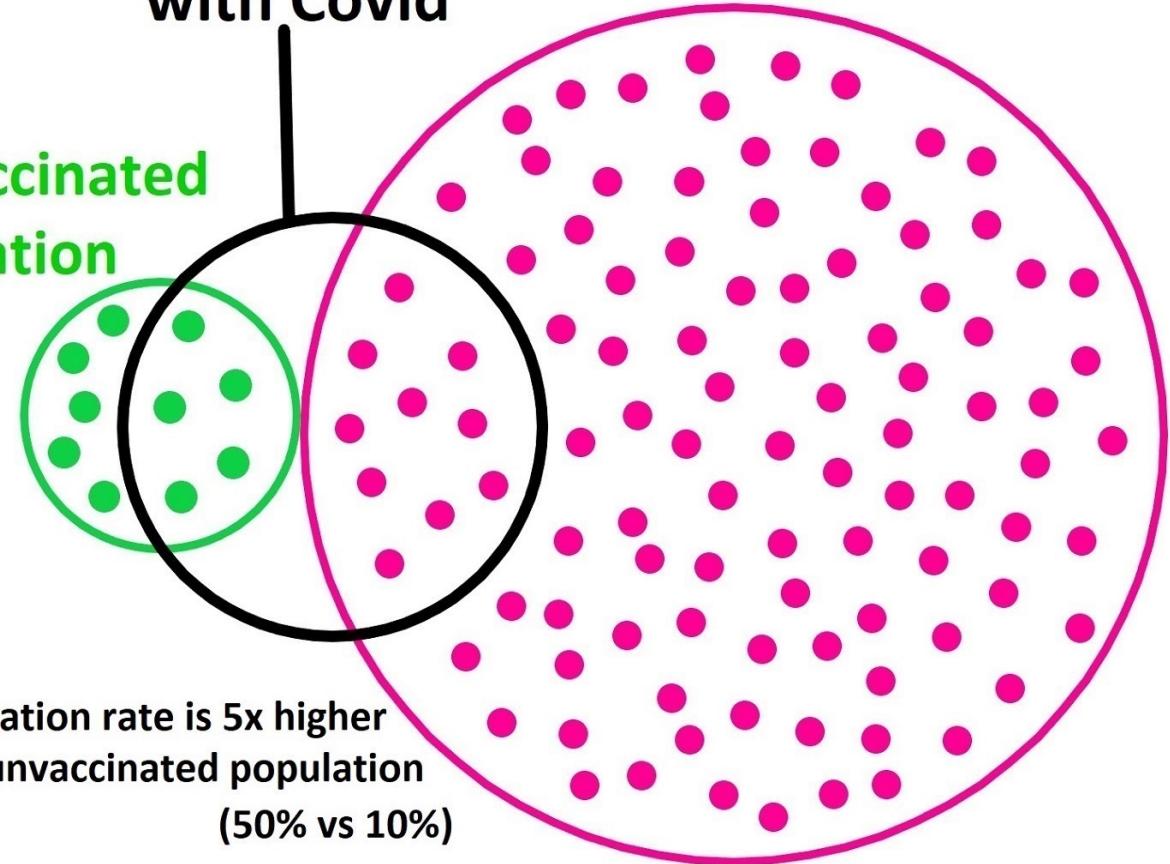
Source: Twitter.com/MarcRummey



But look at the rate of each group in the total population

Hospitalized with Covid

Un-vaccinated population



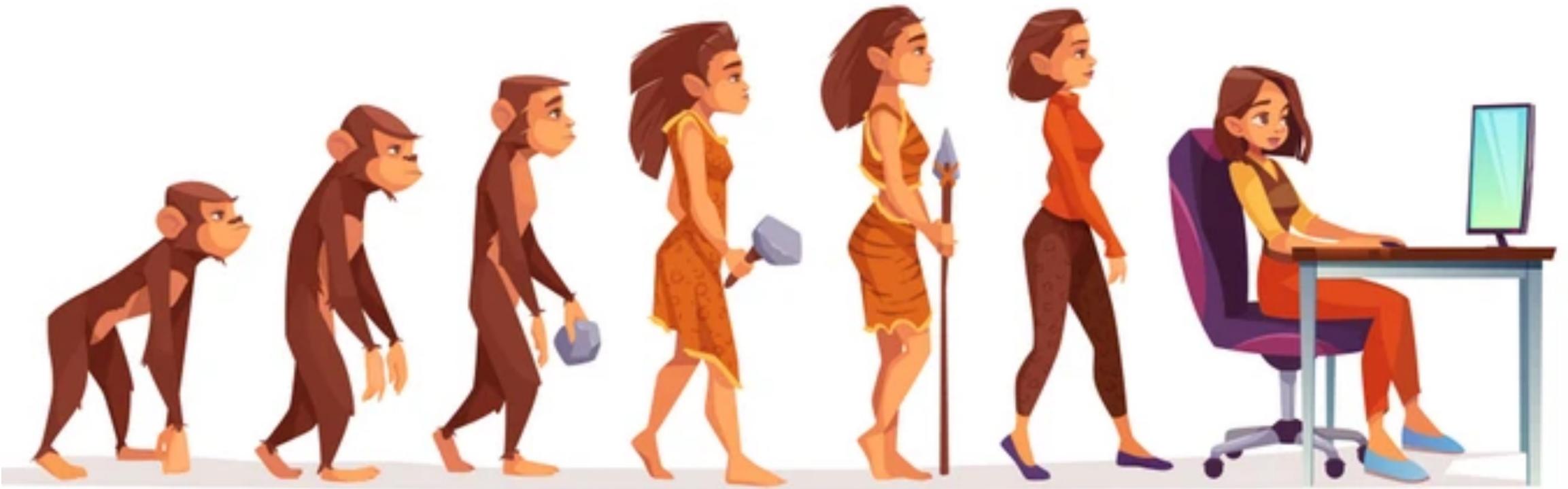
Hospitalization rate is 5x higher in unvaccinated population (50% vs 10%)

Note: The ratios presented are made to illustrate the concept of the base rate fallacy when the vaccination rate is high

ELEMENTS OF STORYTELLING

PART I – STORIES AND STORYTELLING





HUMAN STORIES

Humans *love* humans! They can't get enough of themselves. They crave the company of humans, they value the opinions of humans, and they love hearing stories about humans! [McCloud]

STORYTELLING GOALS

Cultural Stories

- entertain, inform, teach, explore, shock

Data (Scientific) Stories

- describe, diagnose, predict, prescribe, persuade

Any overlap?

Anything missing?

STORYTELLING AUDIENCE

Storytelling requires a **teller** and a **story**, but also an **audience**.

The **teller**'s job is to convince the audience to accept:

1. the premise (“I’m about to tell you a really interesting story, so listen up!”)
2. the contents (“All these things happened, honest!”)
3. the conclusion (“And that’s why you should never put peanut butter in your laundry.”)

The **story**'s must first and foremost not come in the way of the teller's job.

STORYTELLING AUDIENCES

The **audience** is a more nebulous entity.

In many cases, the teller never interacts directly with the audience. For all they know, the audience could be a single child, or the entire nation of Finland.

This **ambiguity** typically leads to storytellers imagining the largest possible audience. A story for the ages, which will be all things to all people.

This is a common mistake: **less is more**. It always pays to know the audience if we are able to.

STORYTELLING AUDIENCES

What is required of a storytelling audience?

What is expected of a storytelling audience?

What kind of audiences exist for stories?

For storytelling with data? (we will discuss this again at a later stage)

AUDIENCES AND GOALS

Who is the audience?

Try to avoid general audiences, we can do this e.g., by addressing **Lines of Business** (finance, engineering, HR, different Programs etc.)

Identify **decision-makers** and the various audience **roles**.

Ask the following questions:

- what relationship do you have with them?
- how do they perceive you?
- how do you establish trust and credibility?

AUDIENCES AND GOALS

What does the audience need the data storytelling to do?

To answer this question, we need to know how the results will be used (**actions**):

- what decisions are people going to make from the analysis?
- how often are they going to be looking at the data?
- how often do they expect the data to be refreshed?

What does the audience **need to know**?

AUDIENCES AND GOALS

What does the audience need to know about data **availability**?

- is the data clean?
- can it be accessed?
- is it being “massaged” (e.g., changed to paint a rosy picture?)

How much will the audience need/want to **interact** with the charts?

- are they passive?
- can they run limited filtering?
- what data can they download (if any)

IDENTIFYING AND GATHERING PRESENTATION REQUIREMENTS

The requirements for a dashboard, report, or presentation are driven by the **primary consumers** (the stakeholders that will primarily be getting “value” from using it).

A very common mistake is to **cast the net too wide** and to build something for too many consumer types (all things to all people). Care needs to be taken to identify the primary consumers.

Once that group has been identified, a **formal process** should be followed to gather requirements as accurately as possible.

IDENTIFYING AND GATHERING REQUIREMENTS

Typical requirement questions include (but should not be limited to):

- what is the name of the story you want to tell?
- who are the target audience?
- what is the high-level objective of telling the story?
- when does it need to be published?
- what is the data update frequency?
- what kind of business decisions will be made by the target audience?
- what is the source of the data that could tell the story?
- is the data/information duplicated anywhere else (e.g., by a 3rd party)?
- what is the sensitivity level of the source data?
- what is the sensitivity level of the final product?
- how is the source data gathered?
- what quality assurance is performed on it?

EXERCISE – STORYTELLING AUDIENCE

1. Make a list of a few different audience groups from your organization.
2. Identify a story from your organization that is unique to each group. Take one of the stories/audience groups and try and answer the requirements questions from the previous slide

STORYTELLING CONTEXT

Telling the same story to different groups means the context changes (or can change)

A given action may be seen as positive or as negative by audiences with different pre-existing feelings/knowledge concerning the agent/situation.

- Would you be able to recognize nobility in a political enemy's actions?
- Could a fan of the Maple Leafs/Habs ever have something worthy to say about hockey?

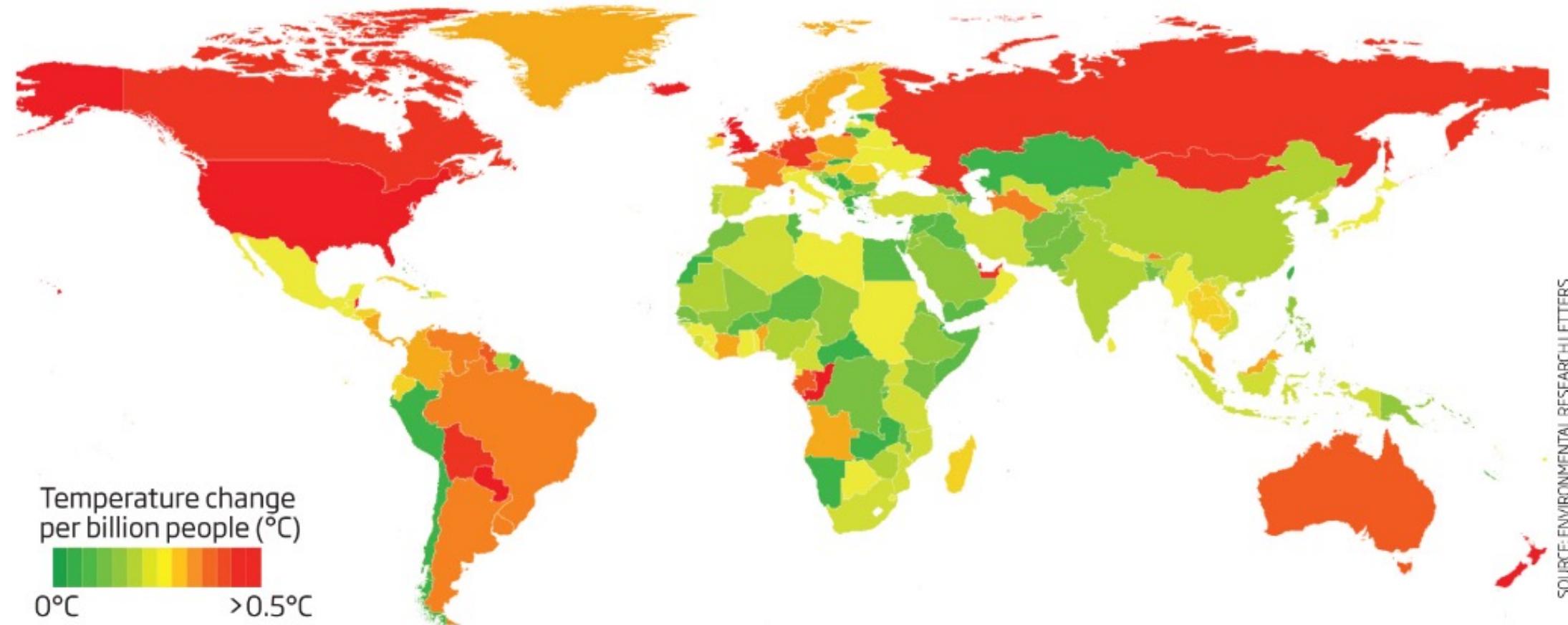
Similarly, a story may have different **outcomes/impacts** in different contexts.

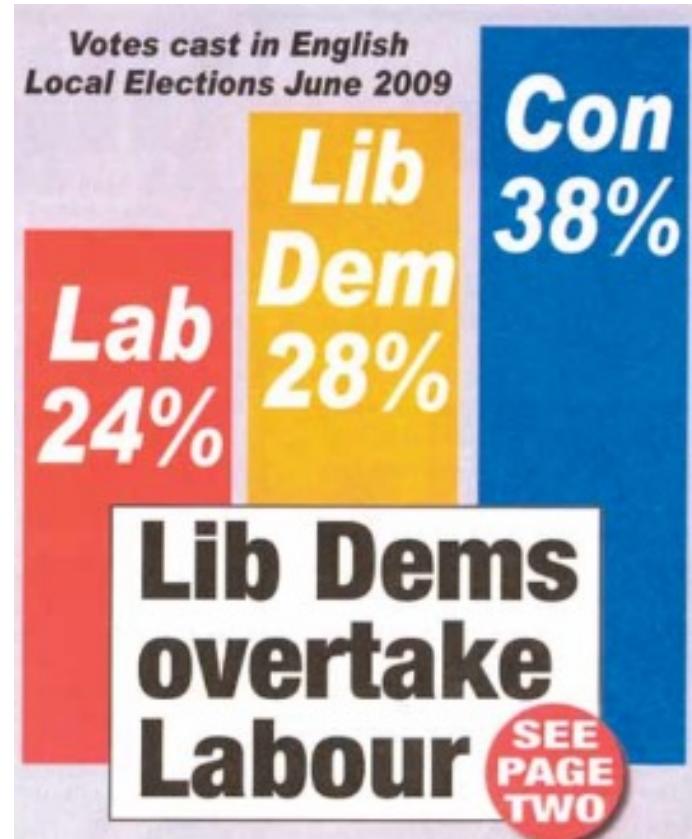
EXERCISE – GUESS THE STORIES

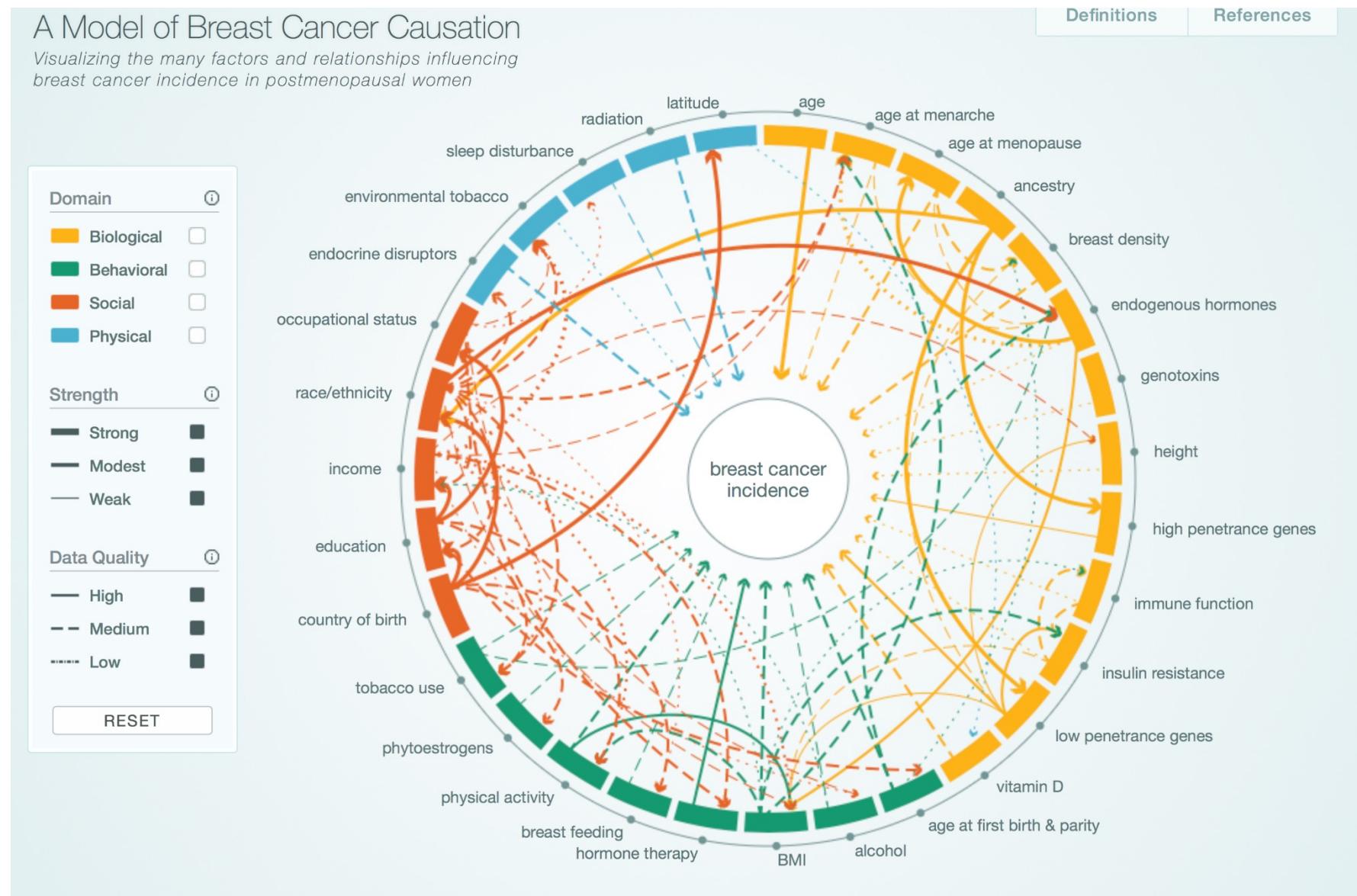
1. Guess the story:
 - a) A group of friends spends 9 hours returning jewellery.
 - b) A talking frog convinces a son to kill his father.
 - c) A young woman with mental illness talks to furniture and marries her kidnapper.
 - d) A depressed, widowed father teams up with an injured woman to find his disabled son.
2. In the following charts, who is the intended audience? What are the goals? Are the outcomes universal?

Global warming culprits, judged by population

Countries that have caused more global warming per billion people are coloured red and low-emitters are dark green

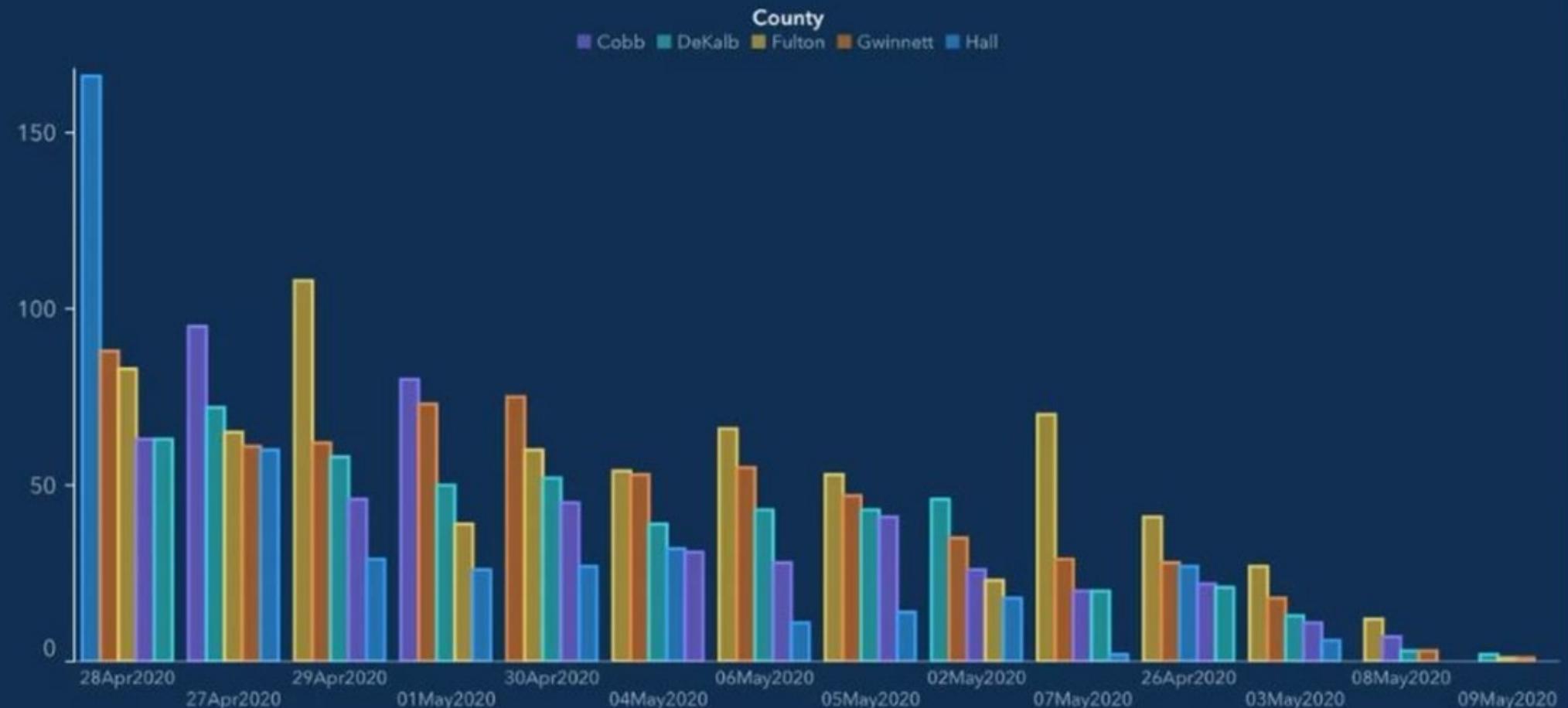






Top 5 Counties with the Greatest Number of Confirmed COVID-19 Cases

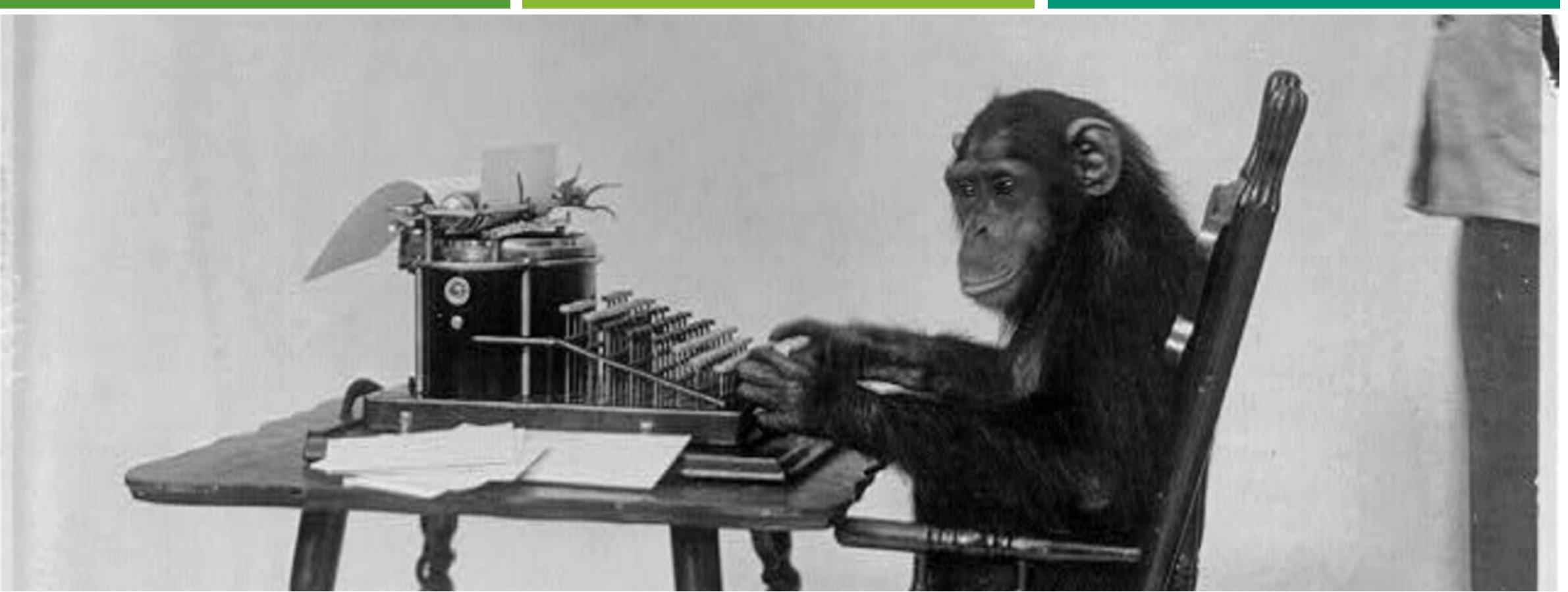
The chart below represents the most impacted counties over the past 15 days and the number of cases over time. The table below also represents the number of deaths and hospitalizations in each of those impacted counties.



EXERCISE – SAME STORY, DIFFERENT AUDIENCE

Identify a story from your organization that all would interest all audience groups. How would you have to change this story for each of the different groups?





STORYTELLING IS IN OUR BLOOD

The anthropologists got it wrong when they named our species *Homo sapiens* ('wise man'). [...] In reality, we are *Pan narrans*, the storytelling chimpanzee. [Cohen, Pratchett, Stewart]

THE ROLE OF TIME IN STORIES

Stories are (necessarily?) **dynamic**.

There must be events happening for a story to be a story, even if the sequence of these events is presented out of order.

Story illustrations may depict one or several moments of a story (graphic novels and comics take this to the next level).

TROPES

In storytelling, a **trope** is a conceptual figure of speech, a storytelling shorthand for a concept that the audience will recognize and understand instantly (e.g., convention):

- plot trick;
- setup;
- narrative structure;
- character type;
- linguistic idiom;
- etc.

Commonly-used tropes become **clichés**: elements that are expected to be part of any story in a given genre.

TROPES

Tropes are **patterns** in storytelling, not only within the works themselves, but also:

- behind-the-scenes aspects of creation;
- technical features of a medium, and
- audience experience and expectations

The idea being that storytelling is not just writing, it is the **whole process of creating and telling/showing a story.**

We have been identifying and discussing patterns in media **for centuries**. Aristotle wrote the *Poetics*, studying tragic plays and epics, making him the first troper of whom we have knowledge. He first diagnosed many of the tropes still in use.

TROPES DISCUSSED IN ARISTOTLE'S *POETICS*

- Acceptable Breaks from Reality
- Anti-Hero
- Bittersweet Ending
- Contrived Coincidence
- Deus ex Machina
- Downer Ending
- Emotional Torque
- Happy Ending
- Random Events Plot
- Reality Is Unrealistic
- The Reveal
- Rule of Cool
- Special Effects Failure
- Spectacle
- Three-Act Structure
- Twist Ending
- Unsympathetic Comedy Protagonist
- Willing Suspension of Disbelief

STORY SPINE: PIXAR

Once upon a time there was __.
Every day, __.
One day __.
Because of that, __.
Because of that, __.
Until finally __.

The Story Spine

We can have various spines in a given work, one for each story.

Where does Dory come in?

The sharks?

Nemo's friends in the dentist's aquarium?

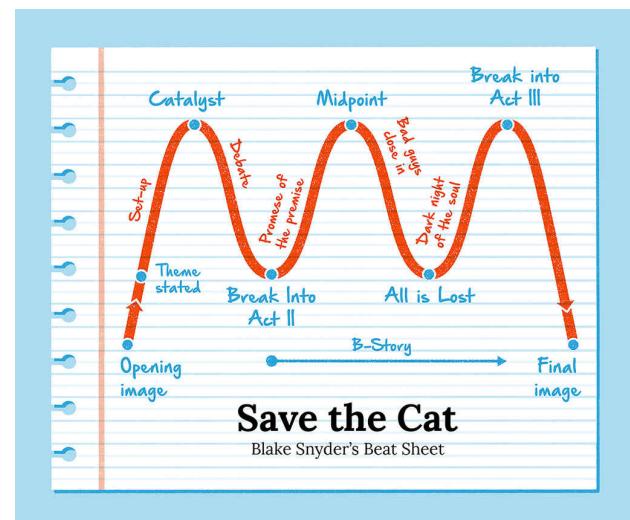
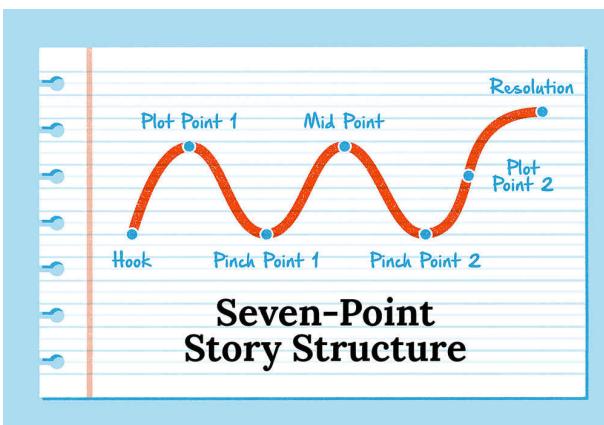
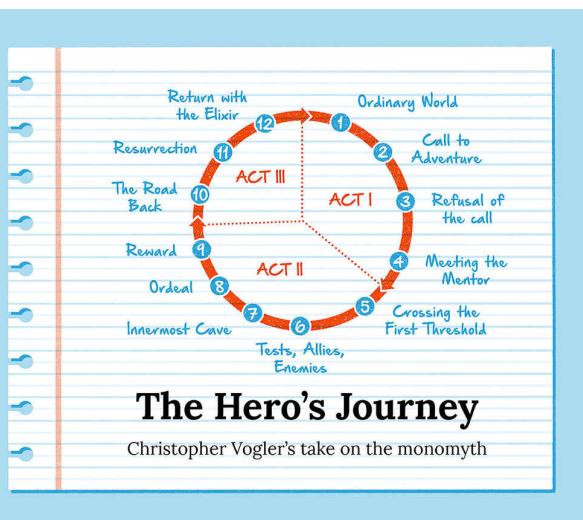
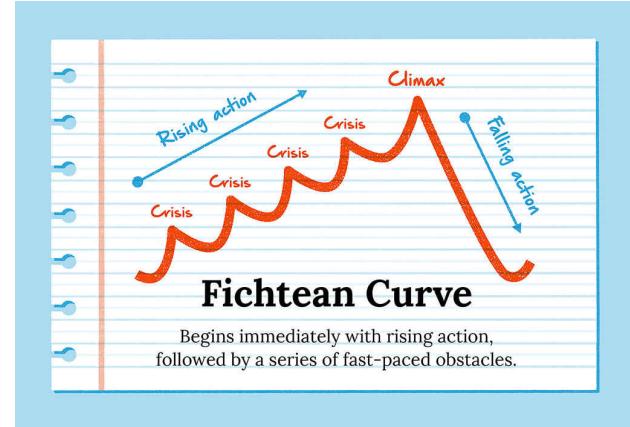
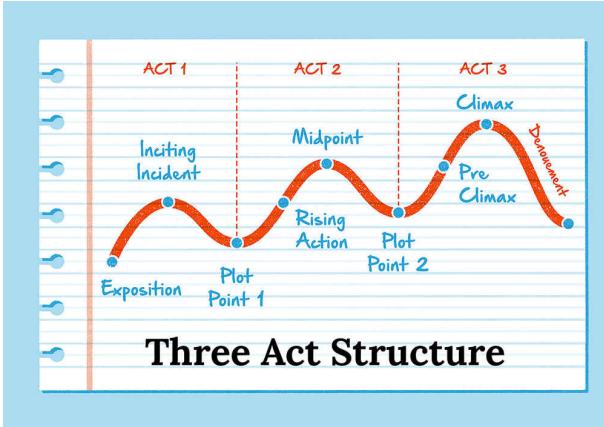
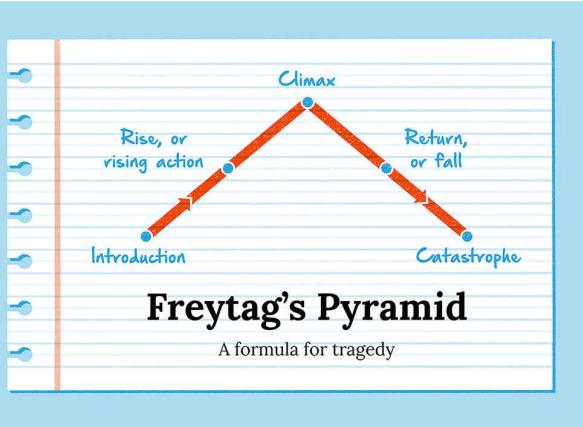
NARRATIVE STRUCTURES

“Narrative structure is the **order** in which events are organized into a beginning, middle, and ending. A story’s structure directly affects the way the plot **unfolds** and how its **driving forces** (characters, obstacles, setting, etc.) are introduced.

Tightly controlled narrative structure results in all questions being answered, provides a climax followed by resolution and information at the end of the story, furthers the characters’ development, and unravels any central conflicts (**humans prefer those**).

Structure helps creators draw connections between ‘things that happen’ and ‘things that matter.’ For instance, a tale about two vastly different people falling in love **can also** be about the value of compromise.”

NARRATIVE STRUCTURES



STORYBOARDING

Stories from our workplace are no different, we can tell the same story with multiple **different** narrative structures.

We can use the narrative structures described earlier to help along.

Storyboarding is another approach that can help.

STORYBOARDING

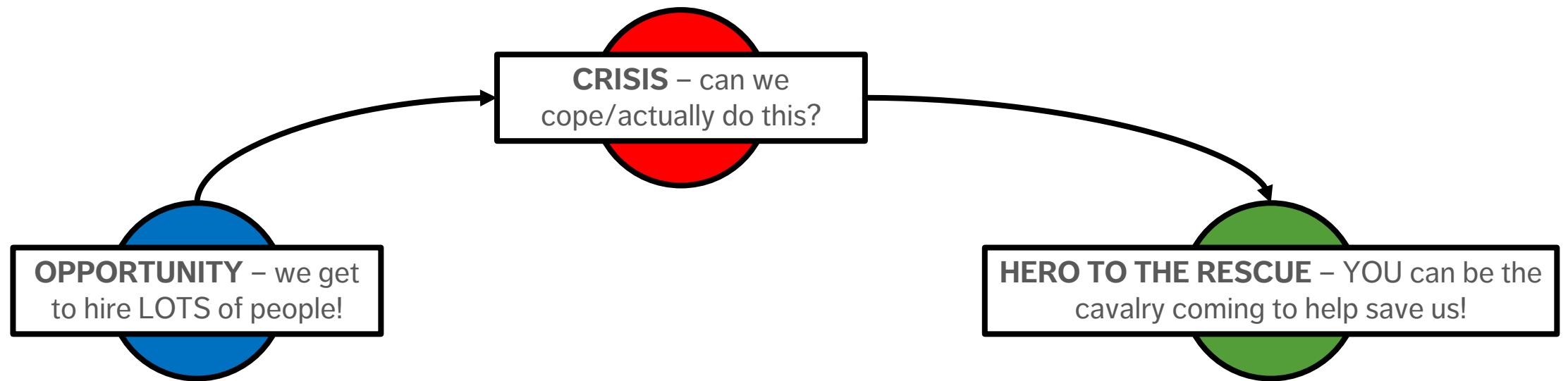
Example:

A department has an aggressive hiring goal for the year, imposed by new legislative requirements. Funding for the hiring has been approved but it will severely impact not just hiring managers but also various corporate service groups (IMIT, HR, Real Property etc.).

STORYBOARDING

Example:

A simple narrative structure might look like the following:



STORYBOARDING

Example:

Obviously we need a little more than this to help us build our reports or dashboards so enter the storyboard. This simply takes the narrative structure and breaks it down into more understandable steps.

We put each step in a box like in a cartoon strip. If we were making a movie we would also use images – feel free to use images if you want!!

STORYBOARDING

Example:

1. State intended hiring goal for the year

2. Describe what is driving the hiring (Fed Gov't Init)

3. Show how close/far the goal is as of today

4. Show which branches have the highest requirements

5. Demonstrate which corporate service groups are impacted the most

6. Ask/tell the audience how they can help

EXERCISES

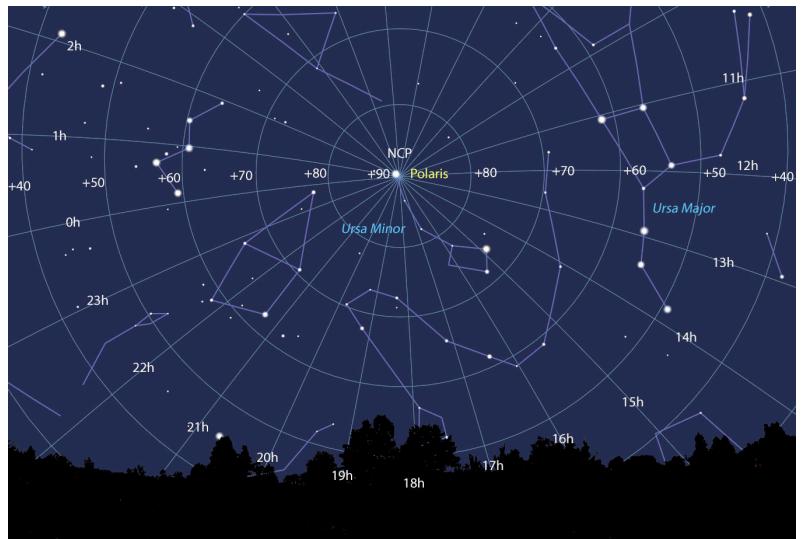
1. Think of a simple story from your workplace.
2. Think of a narrative structure.
3. Build a storyboard around that structure, using as many (or as few) boxes as necessary.

DATA STORIES

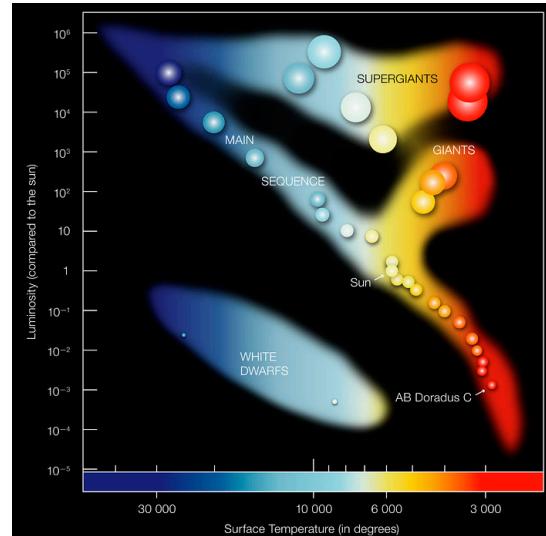
PART I – STORIES AND STORYTELLING

WHAT IS A DATA STORY?

If we look back at our map of the sky we don't just have patterns (constellations) that we can build stories around, we also have underlying data that we can use, e.g.:



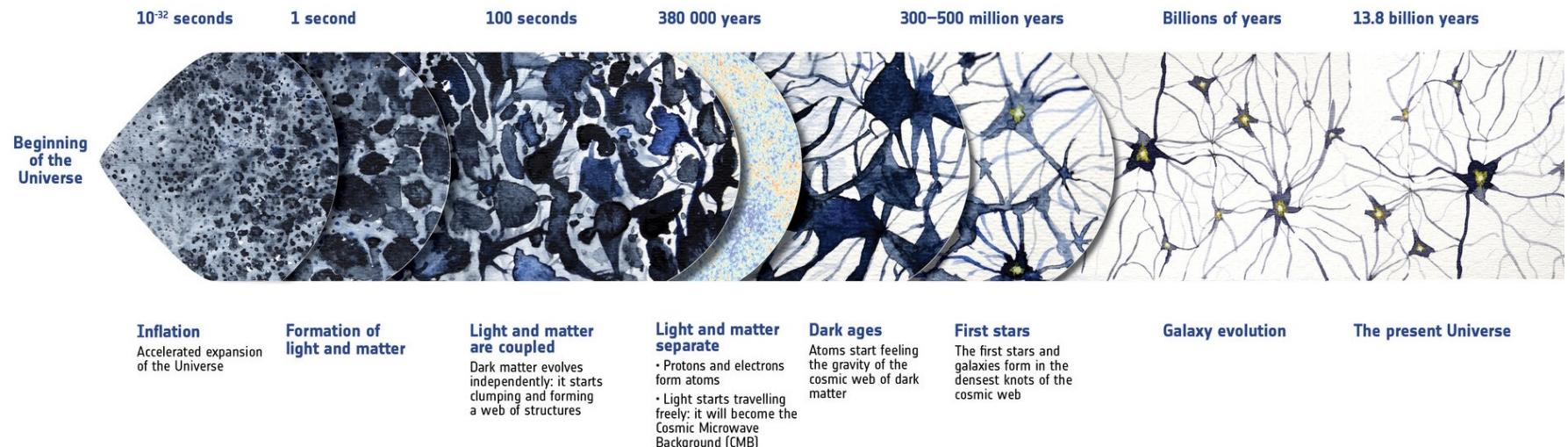
Constellation coordinates



Hertzsprung-Russell (star temperature – luminosity) Diagram

WHAT IS A DATA STORY?

When we do this it becomes obvious that our fantastical interpretation of the stars is **meaningless** (shame) but other **meaningful** stories begin to emerge based on our analysis and interpretation of the data!



WHAT IS A DATA STORY?

Data storytelling is the ability to effectively communicate insights from a dataset using narratives and visualizations. It can be used to put data insights into context for and inspire action from your audience (Catherine Cote).

There are 3 key components:

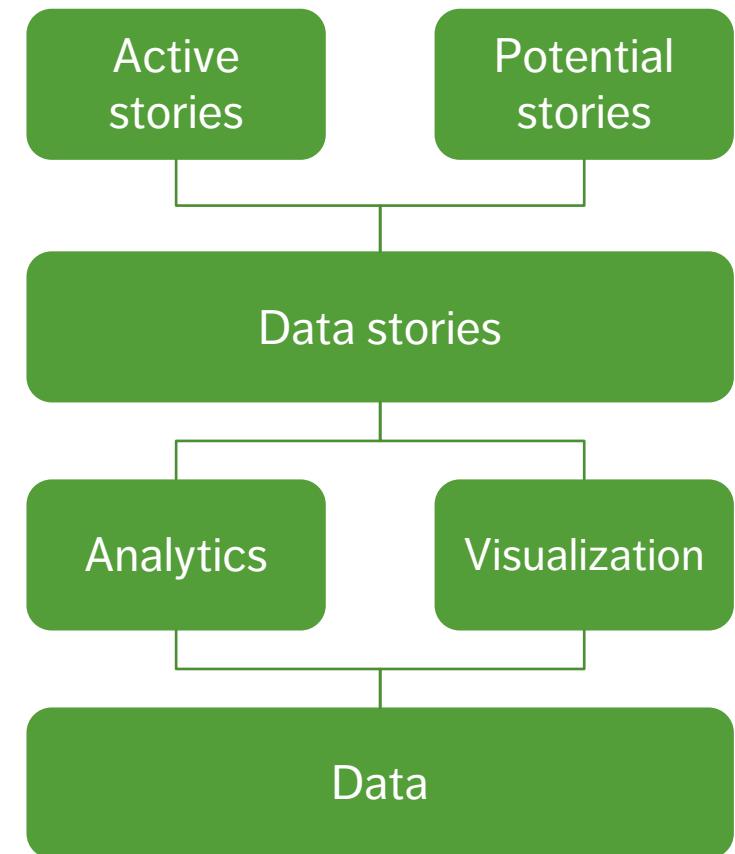
1. **data:** foundation of data story (descriptive, diagnostic, predictive, prescriptive analysis)
2. **narrative:** storyline used to communicate the insights gleaned from data and context, and recommended actions
3. **visuals:** representations of data, analysis results, and narratives, which are used to communicate stories clearly and memorably (charts, graphs, diagrams, pictures, or videos)

WHAT IS A DATA STORY?

Data stories help us explain the actual stories that exist, or to articulate stories we want to tell.

But we might not have all the data required to do this.

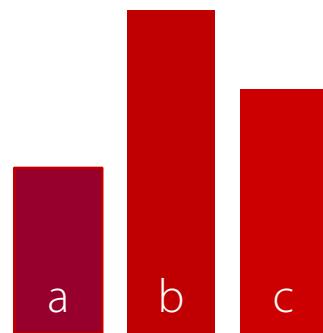
Thus, we might not fully understand existing stories or be able to clearly articulate stories we want to tell **effectively**.



ANALYTICS MODES

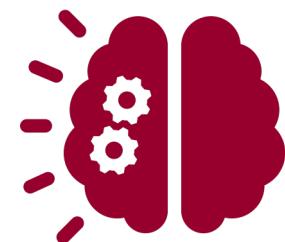
Analytics can be broken down into four core **key modes**:

Descriptive



Show **what** happened

Diagnostic



Explain **why** something happened

Predictive



Guess **what will** happen

Prescriptive



Suggest **what should** happen

Low Value
Low Difficulty



High Value
High Difficulty

Same Data, Different Conclusions

Twenty-nine research teams were given the same set of soccer data and asked to determine if referees are more likely to give red cards to dark-skinned players. Each team used a different statistical method, and each found a different relationship between skin color and red cards.

Referees are
three times as
likely to give red
cards to
dark-skinned
players

Statistically
significant results
showing referees are
more likely to give red
cards to dark-skinned
players

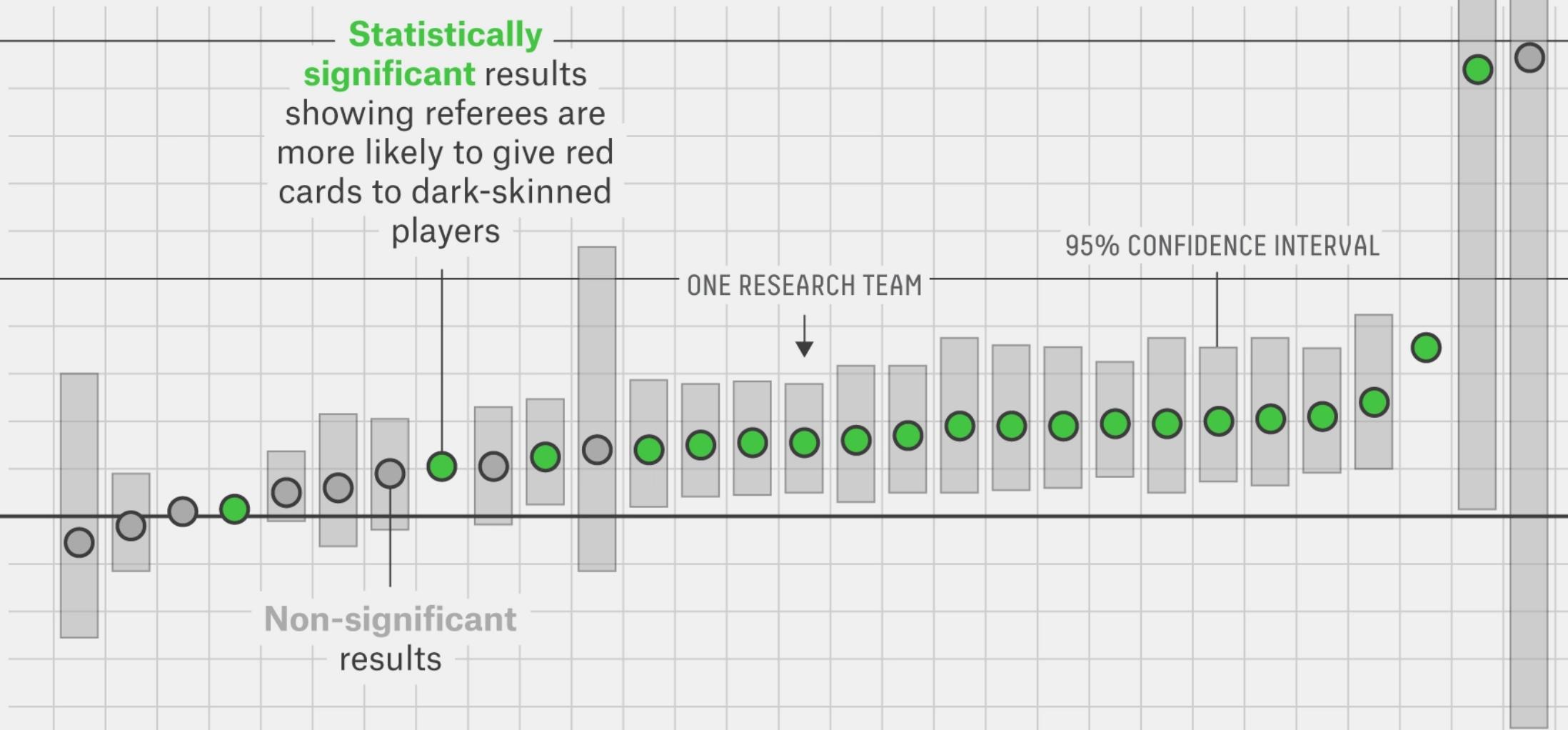
Twice as likely

ONE RESEARCH TEAM

95% CONFIDENCE INTERVAL

Equally likely

Non-significant
results



no. of
constellations

30 –

20 –

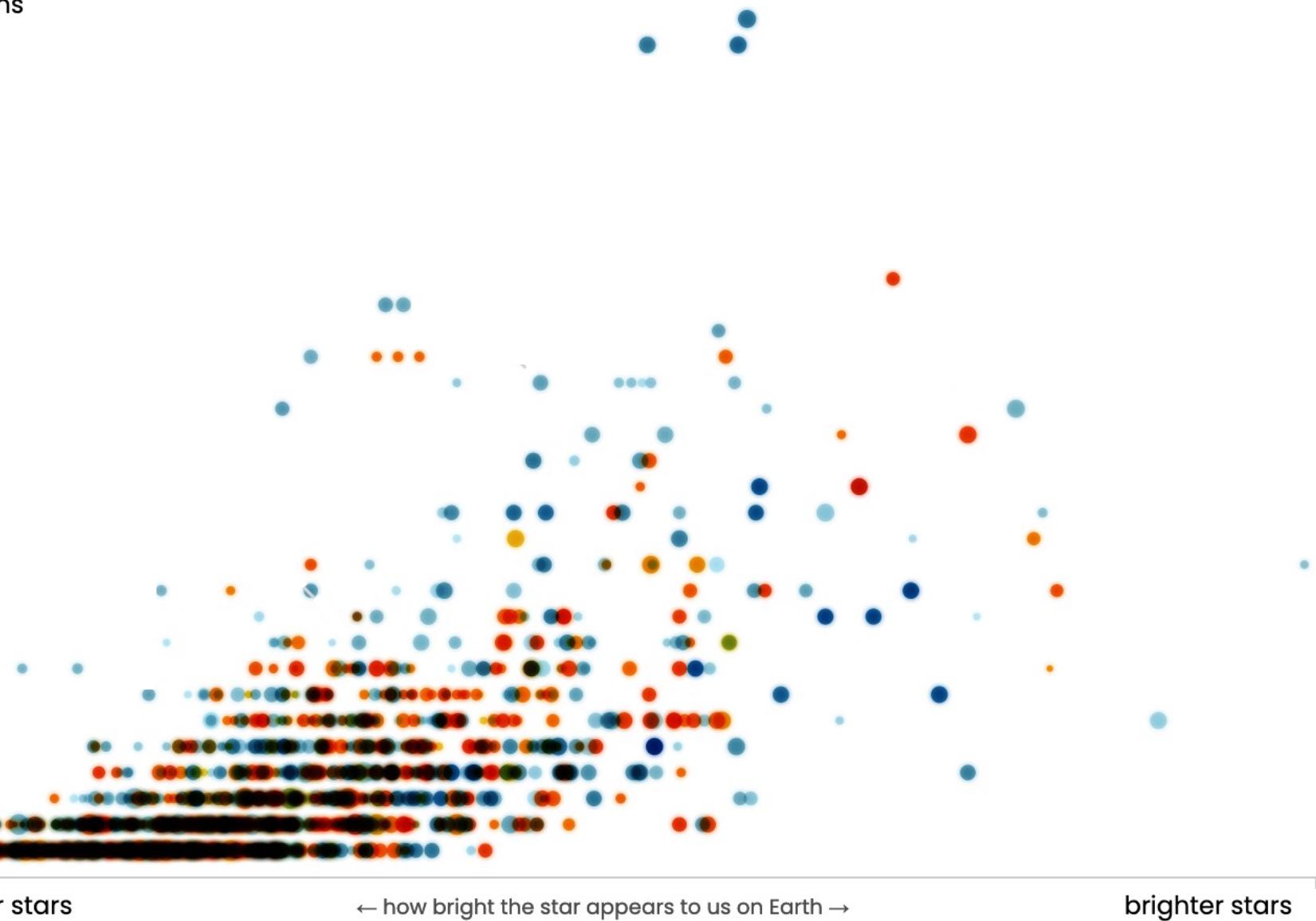
10 –

0 –

fainter stars

← how bright the star appears to us on Earth →

brighter stars



no. of constellations

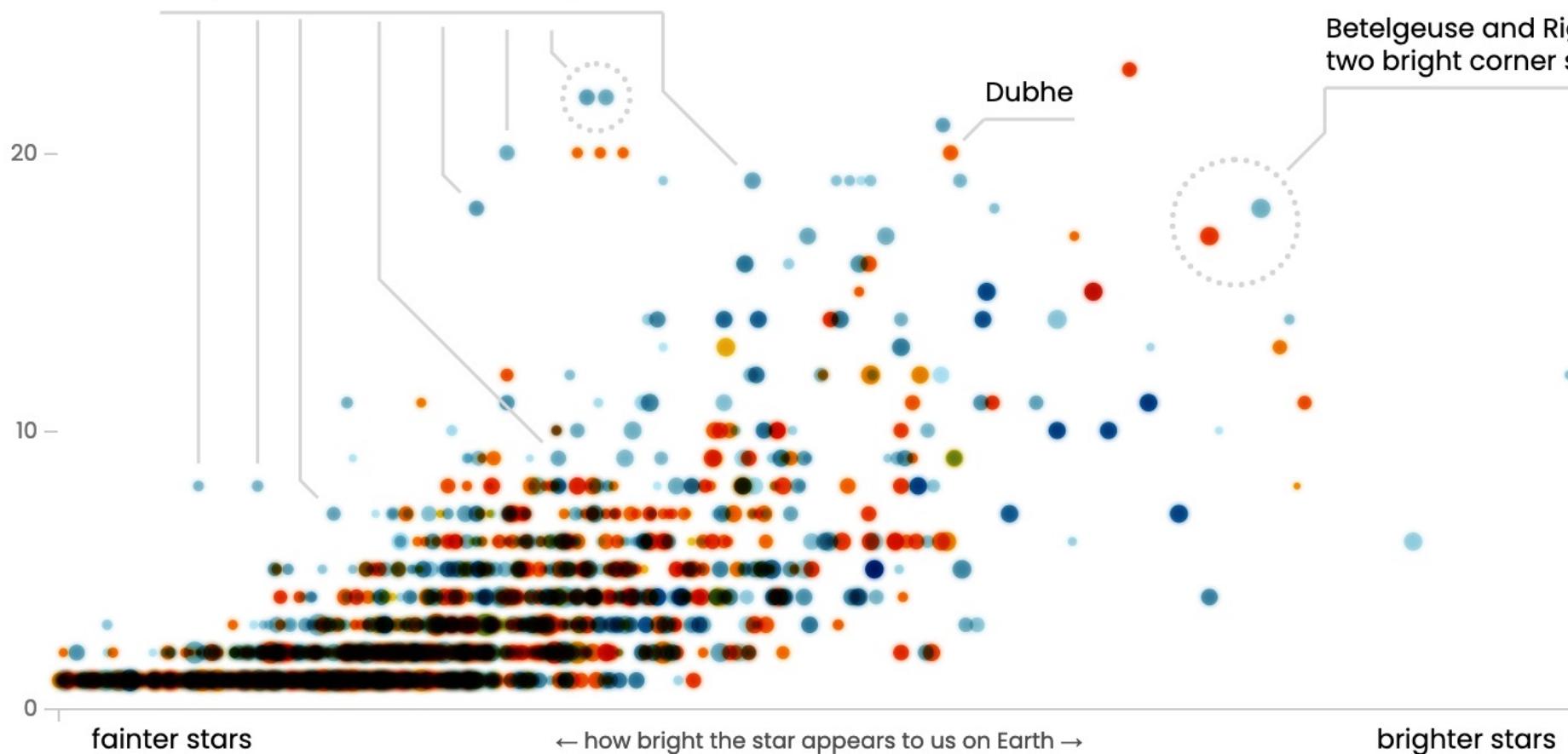
Pleiades

These 9 tightly packed stars are used in constellations more often than expected for their brightness. Most likely due to their ease of recognition



Orion's belt

The 3 stars that make up 'Orion's belt' are used in a constellation across most cultures. Some even more than once per culture



EXERCISE – DATA STORIES FROM YOUR WORKPLACE

Identify data that could tell part or all of one of the stories from the previous exercise.

Examples: from the previous exercise the underlying data for our “Data Story” might look something like:

- Project budgets regarding required O&M investments were consistently submitted late (we need project budget \$\$, submission date all for O&M category).
- Salary forecasts did not align with hiring plans (we need salary forecasts dates and hiring plan dates).
- Hiring lead time information from HR was ignored (we need date HR submitted information to stakeholders and date budgets were submitted).



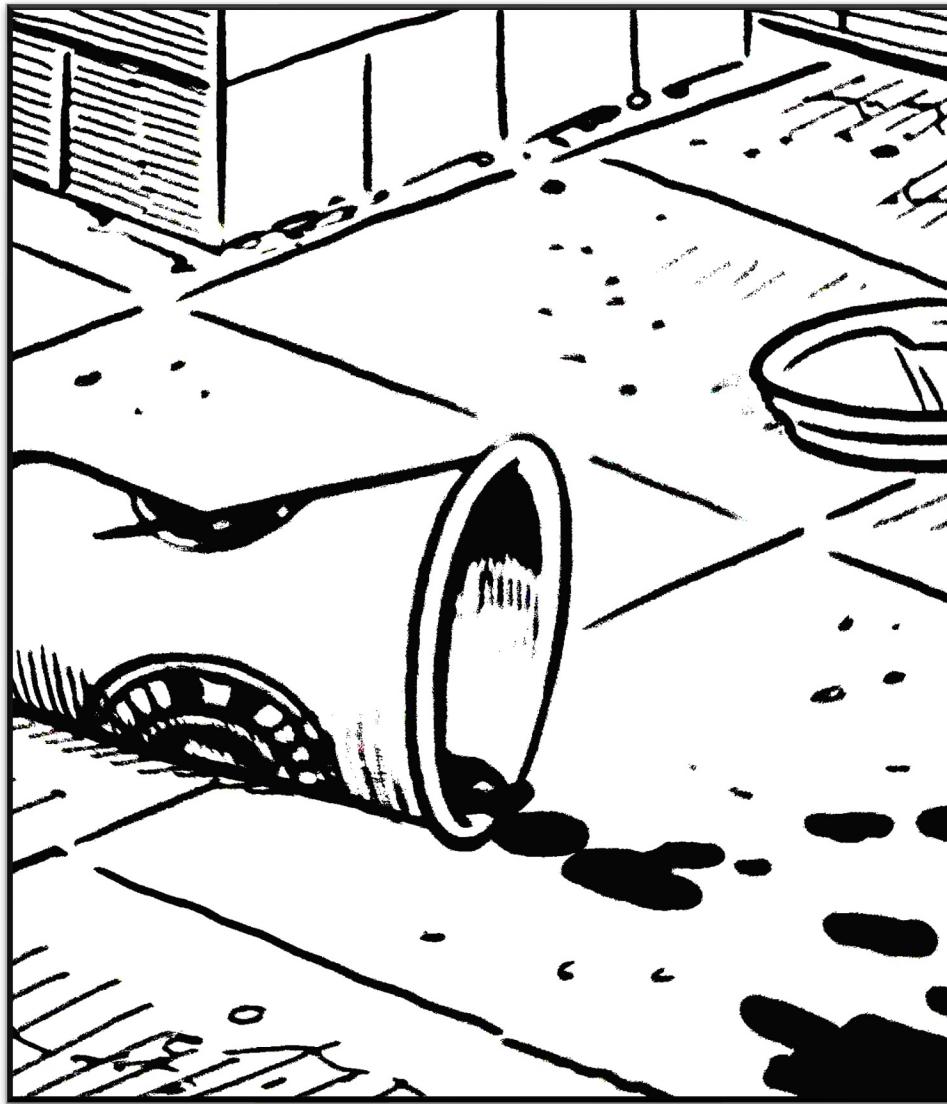
VISUAL STORYTELLING

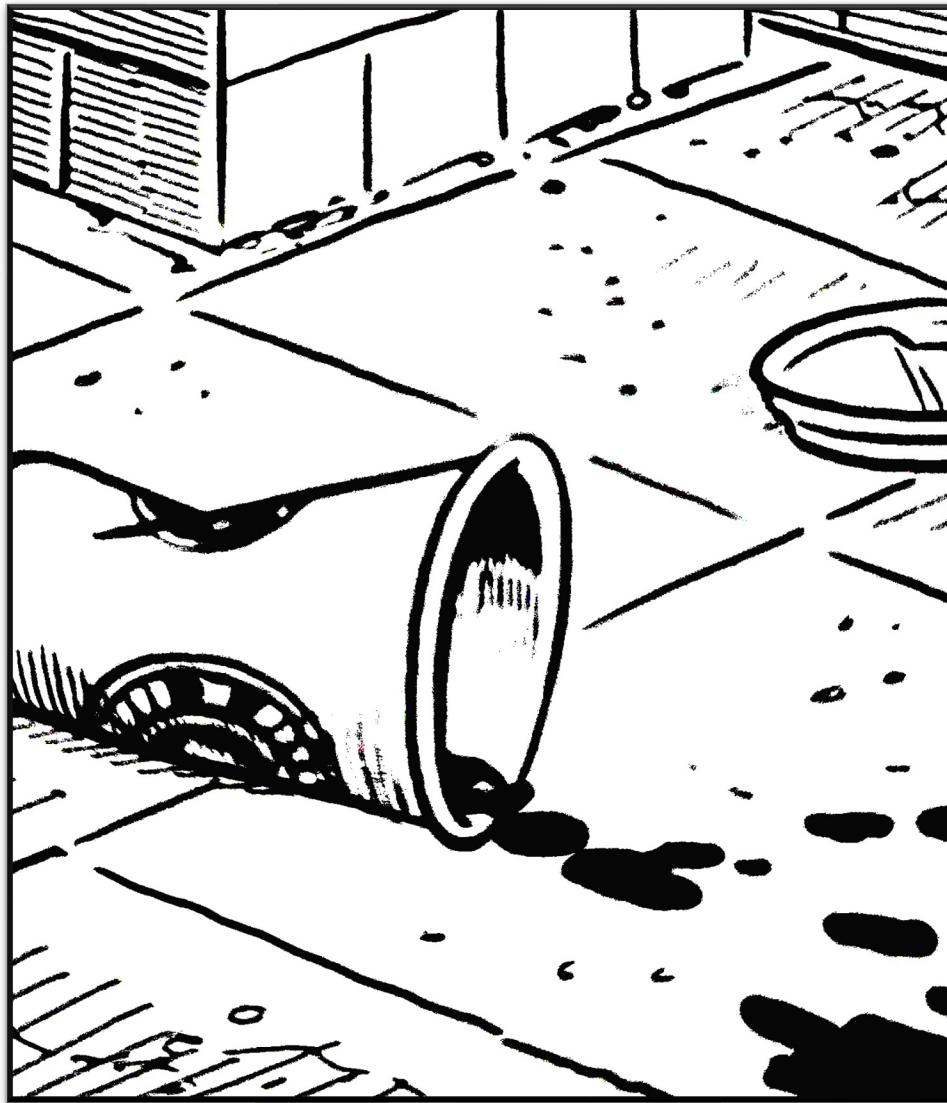
PART I – STORIES AND STORYTELLING



VISUAL STORYTELLING

Visual storytelling requires us to make a constant stream of choices regarding imagery, pacing, dialogue, composition, gesture, and a ton of other options. These choices break down into 5 basic types: choice of **moment**, choice of **frame**, choice of **image**, choice of **word**, choice of **flow**. These are the 5 areas where your choices can make the difference between clear, convincing storytelling and a confusing mess. [McCloud]





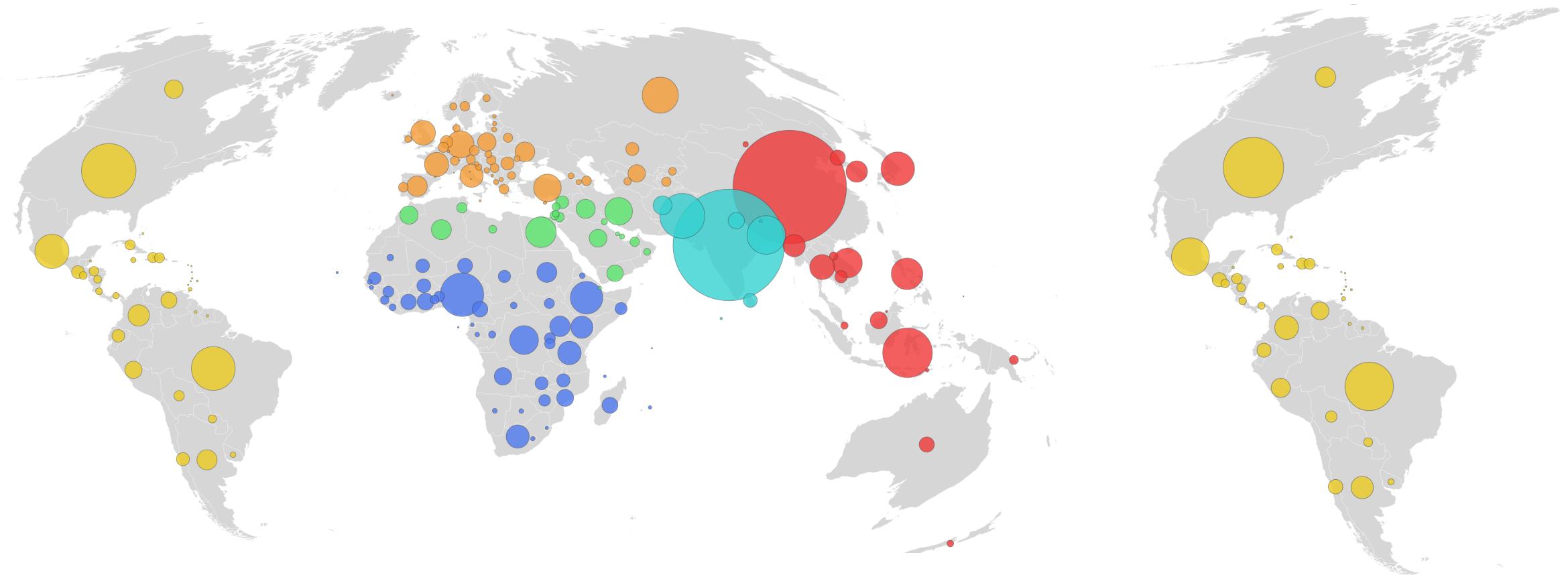
“On the bright side, I got my caffeine. On the not-so-bright side, we got mugged on the way home.”

VISUAL STORYTELLING CHOICES (PRE-GESTALT PRINCIPLES)

Communicating with **clarity** means that audience comprehension the **ultimate goal**:

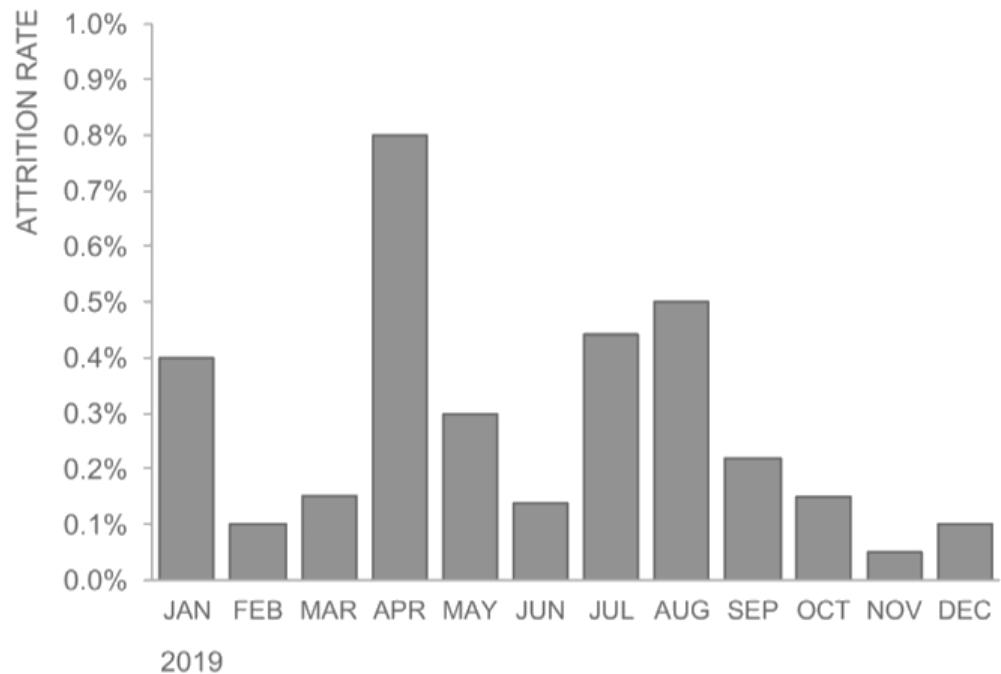
- choice of **moment** is ‘connecting the dots’, showing only what matters to the story;
- choice of **frame** is creating and directing the audience’s focus;
- choice of **image** is selecting the right charts for the story, with emphasis on simplicity and ability to convey the message;
- choice of **word** is clearly and persuasively communicating ideas in seamless combination with the charts;
- choice of **flow** is guiding the audience from one chart to the next, from one page to the next, and creating a transparent and intuitive ‘reading’ experience, by arranging pages in a dashboard, charts on a page, and elements within charts intelligently.

CHOICE OF MOMENT

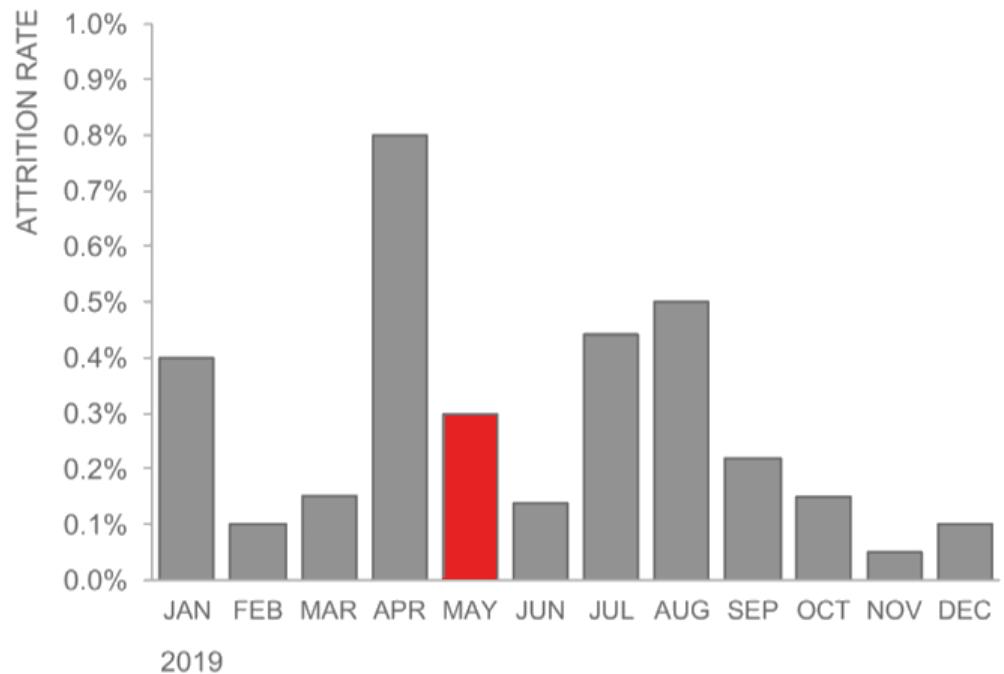


CHOICE OF FRAME

2019 monthly voluntary attrition rate

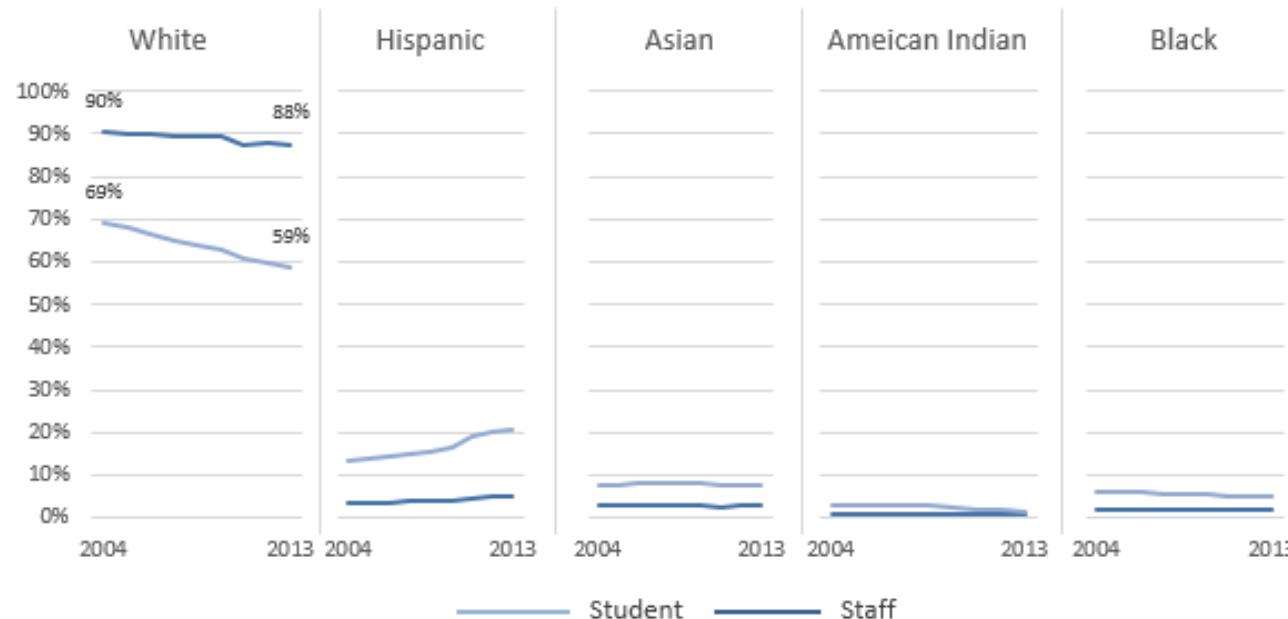


2019 monthly voluntary attrition rate

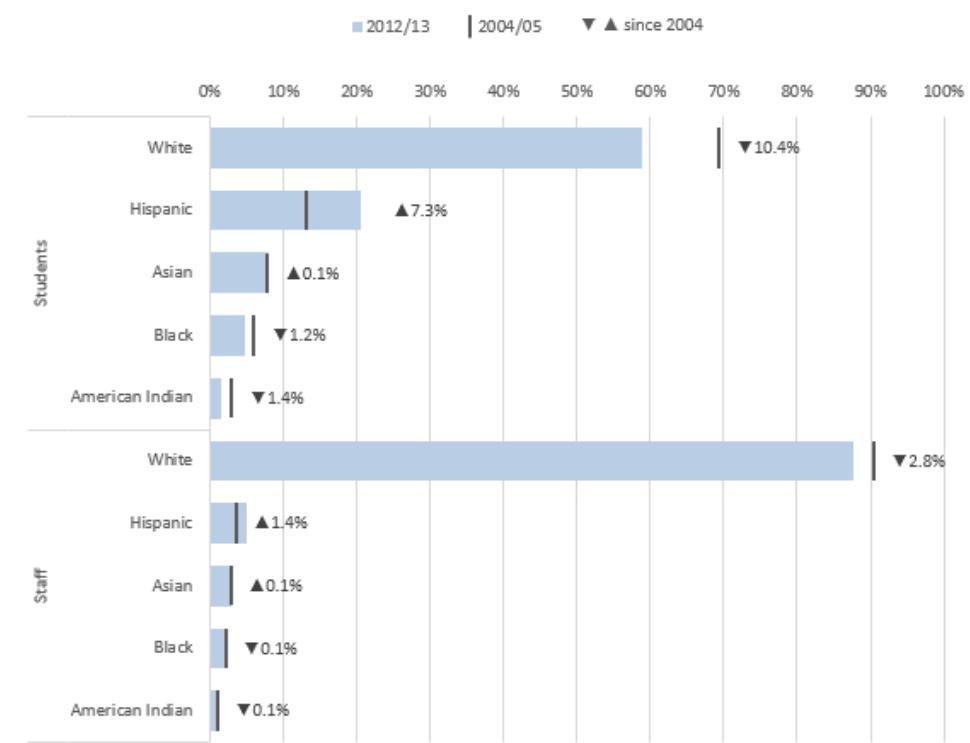


CHOICE OF IMAGE

Washington State Percentage Staff and Student by Ethnicity 2004 to 2013

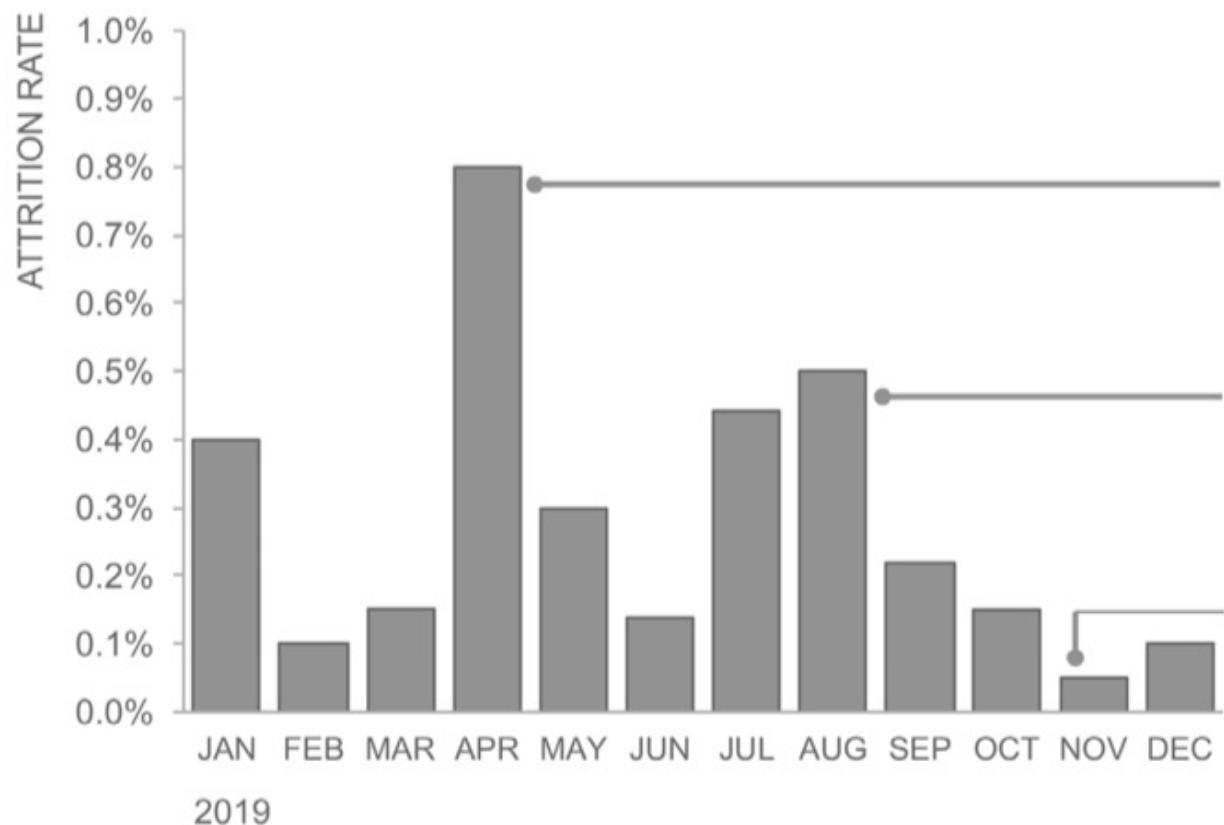


Washington State % of Staff and Student by Ethnicity 2004 to 2013



CHOICE OF WORD

2019 monthly voluntary attrition rate



Highlights:

In April there was a reorganization. No jobs were eliminated, but many people chose to leave.

Attrition rates tend to be higher in the Summer months when it is common for associates to leave to go back to school.

Attrition is typically low in November & December due to the holidays.

VISUAL STORYTELLING CHOICES

Decisions having to do with *moment*, *frame*, and *flow* are likely to be made in the dashboard **planning stages**, while *image* and *word* decisions are usually being made right **up to the finish line**.

We can:

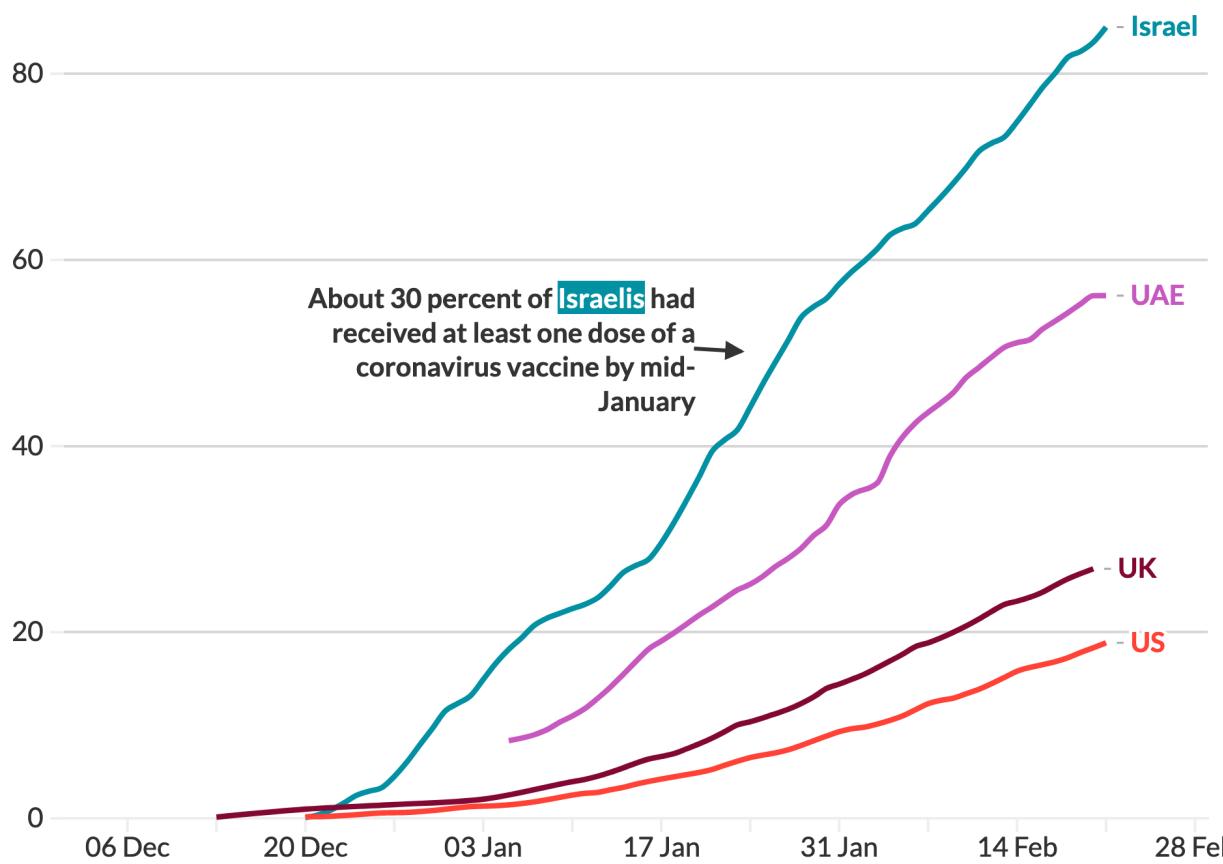
- start with a rough sketch of the dashboard (moment, frame, flow), then come up with the narrative (word), then populate the dashboard with charts (image);
- start with a full ‘script’/storyboard (moment, word), then use that to do a rough layout of the dashboard (frame, flow), then populate the dashboard with charts (image);
- create a finished chart (moment, frame, image, word) with no idea as to what else will show up on the dashboard until you create another chart (flow), and so on (not recommended)!

VISUAL STORYTELLING COMBINATIONS

- **text-specific**, where text provides all that is needed to know and the charts illustrate some aspects of the story that is described
- **chart-specific**, where the charts provide all that is needed to know and the text accentuates some aspects of the story that is shown
- **duo-specific**, where text and charts are both telling roughly the same story
- **intersecting**, where text and charts work together in some respects but also contribute to the story independently
- **interdependent**, where text and charts combine to convey an aspect of the story that neither could convey alone
- **parallel**, where words and charts follow seemingly different storylines, without intersecting

Cumulative vaccination doses administered in Israel, UAE, UK and US

Cumulative doses administered per 100 residents • Data last updated 24 Feb



Source: [ECDC/OWID](#) • Graphic: [Flourish](#) • [Embed this](#)



I have a story I'd like to tell you. It's about a train, and a group of people who live on that train and know of nothing else.

This train has been moving since anyone can remember. The people on the train can't imagine a time when the train wasn't moving, and when they were not on the train. Everyone works to keep the train moving. The train never stops.



It never stops. It cannot stop.

People on the train live in constant churn. The work to keep the train moving is hard, and inhumane. On the train, people are treated with cruelty and oppression. Some are treated worse than others. But nobody is truly living.



Sometimes they get breaks, but it is hard.

One day, a fire breaks out in one of the carriages of the train.



There is panic. The fire spreads throughout the whole train... Without getting off the train everyone is going to die.

Then the impossible happens.



The brakes no-one believed existed start to work. In the emergency, no-one notices how extraordinary it is that the train is stopping. They're too focused on the fire. The old rules go out the window.

For years on the train, the "worker class" of people have been dying from the awful conditions of the work they have to do on the train. They sleep in the aisles and sometimes have nowhere to sleep at all.

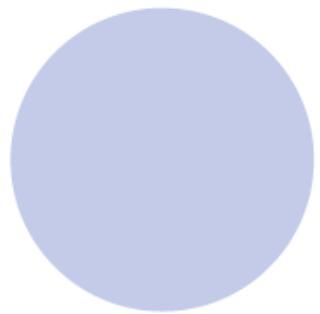
Suddenly, there are orders to house them and treat their ailments.

The train stops, and people begin to get off. Apart from the sound of the fire, suddenly there is a great silence.

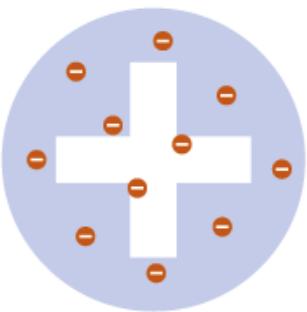
A HISTORY OF THE ATOM: THEORIES AND MODELS

How have our ideas about atoms changed over the years? This graphic looks at atomic models and how they developed.

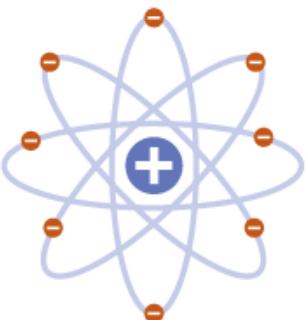
SOLID SPHERE MODEL



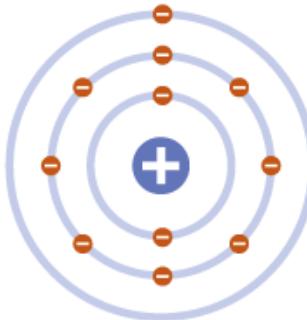
PLUM PUDDING MODEL



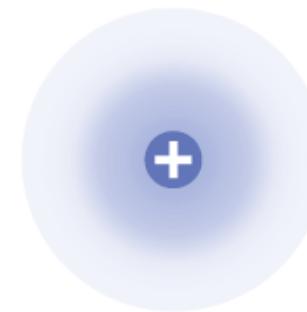
NUCLEAR MODEL



PLANETARY MODEL



QUANTUM MODEL



JOHN DALTON



1803

Dalton drew upon the Ancient Greek idea of atoms (the word 'atom' comes from the Greek 'atomos' meaning indivisible). His theory stated that atoms are indivisible, those of a given element are identical, and compounds are combinations of different types of atoms.

- + RECOGNISED ATOMS OF A PARTICULAR ELEMENT DIFFER FROM OTHER ELEMENTS

- ATOMS AREN'T INDIVISIBLE - THEY'RE COMPOSED FROM SUBATOMIC PARTICLES

J.J. THOMSON



1904

Thomson discovered electrons (which he called 'corpuscles') in atoms in 1897, for which he won a Nobel Prize. He subsequently produced the 'plum pudding' model of the atom. It shows the atom as composed of electrons scattered throughout a spherical cloud of positive charge.

- + RECOGNISED ELECTRONS AS COMPONENTS OF ATOMS

- NO NUCLEUS; DIDN'T EXPLAIN LATER EXPERIMENTAL OBSERVATIONS

ERNEST RUTHERFORD



1911

Rutherford fired positively charged alpha particles at a thin sheet of gold foil. Most passed through with little deflection, but some deflected at large angles. This was only possible if the atom was mostly empty space, with the positive charge concentrated in the centre: the nucleus.

- + REALISED POSITIVE CHARGE WAS LOCALISED IN THE NUCLEUS OF AN ATOM

- DID NOT EXPLAIN WHY ELECTRONS REMAIN IN ORBIT AROUND THE NUCLEUS

NIELS BOHR



1913

Bohr modified Rutherford's model of the atom by stating that electrons moved around the nucleus in orbits of fixed sizes and energies. Electron energy in this model was quantised; electrons could not occupy values of energy between the fixed energy levels.

- + PROPOSED STABLE ELECTRON ORBITS; EXPLAINED THE EMISSION SPECTRA OF SOME ELEMENTS

- MOVING ELECTRONS SHOULD EMIT ENERGY AND COLLAPSE INTO THE NUCLEUS; MODEL DID NOT WORK WELL FOR HEAVIER ATOMS

ERWIN SCHRÖDINGER



1926

Schrödinger stated that electrons do not move in set paths around the nucleus, but in waves. It is impossible to know the exact location of the electrons; instead, we have 'clouds of probability' called orbitals, in which we are more likely to find an electron.

- + SHOWS ELECTRONS DON'T MOVE AROUND THE NUCLEUS IN ORBITS, BUT IN CLOUDS WHERE THEIR POSITION IS UNCERTAIN

- + STILL WIDELY ACCEPTED AS THE MOST ACCURATE MODEL OF THE ATOM



EXERCISES

1. Look at your storyboard
2. Think about which visual storytelling combination(s) might work best to visualize the story and narrative

PART II – EFFECTIVE STORYTELLING VISUALS

STORYTELLING WITH DATA



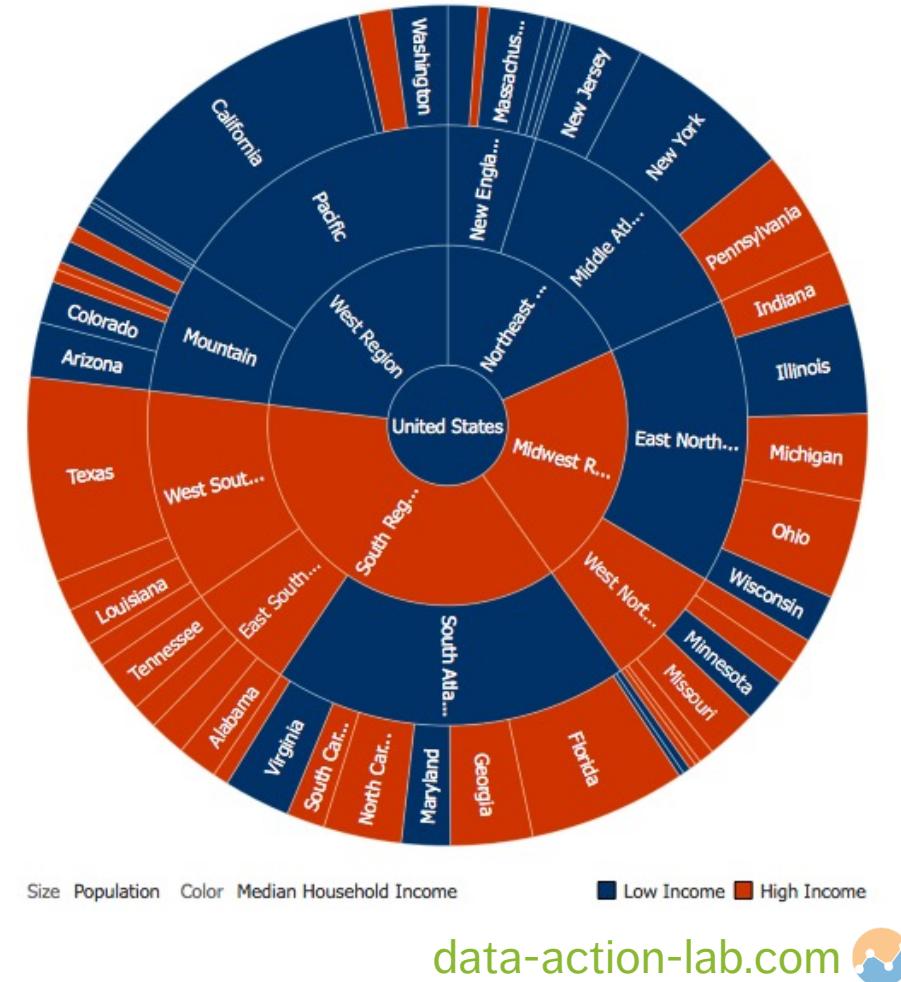
THE ABC OF DATA VISUALIZATION

PART II – EFFECTIVE STORYTELLING VISUALS

DATA VISUALIZATION AND INFOGRAPHICS

Data Visualization

- A **method**, as well as an item (**objective**)
- Typically focuses on the **quantifiable**
- Used to make sense of the data or to make it **accessible** (datasets can be massive and unwieldy)
- May be generated **automatically**
- The look and feel are less important than the **insights conveyed** by the data



DATA VISUALIZATION AND INFOGRAPHICS

Infographics

- Created for **story-telling** purposes (**subjective**)
- Intended for a **specific** audience
- **Self-contained** and discrete
- **Graphic design** aspect is key
- **Cannot** usually be re-used with other data
- Can incorporate **unquantifiable** information



LEARNING FROM HISTORY

Data visualization is not confined to the recent past: charts have been used for many years to help **communicate information and tell stories**.

Due to the absence of technical tools, a lot of thought had to go into the design and creation of these visualizations.

Consequently, there is a lot we can (and **should**) learn to bring into the development of charts from a **design and storytelling perspective**.

London's Cholera Outbreak of 1854

Physician John Snow links the outbreak to a contaminated well by plotting number of cases on a map, jump-starting the science of epidemiology.

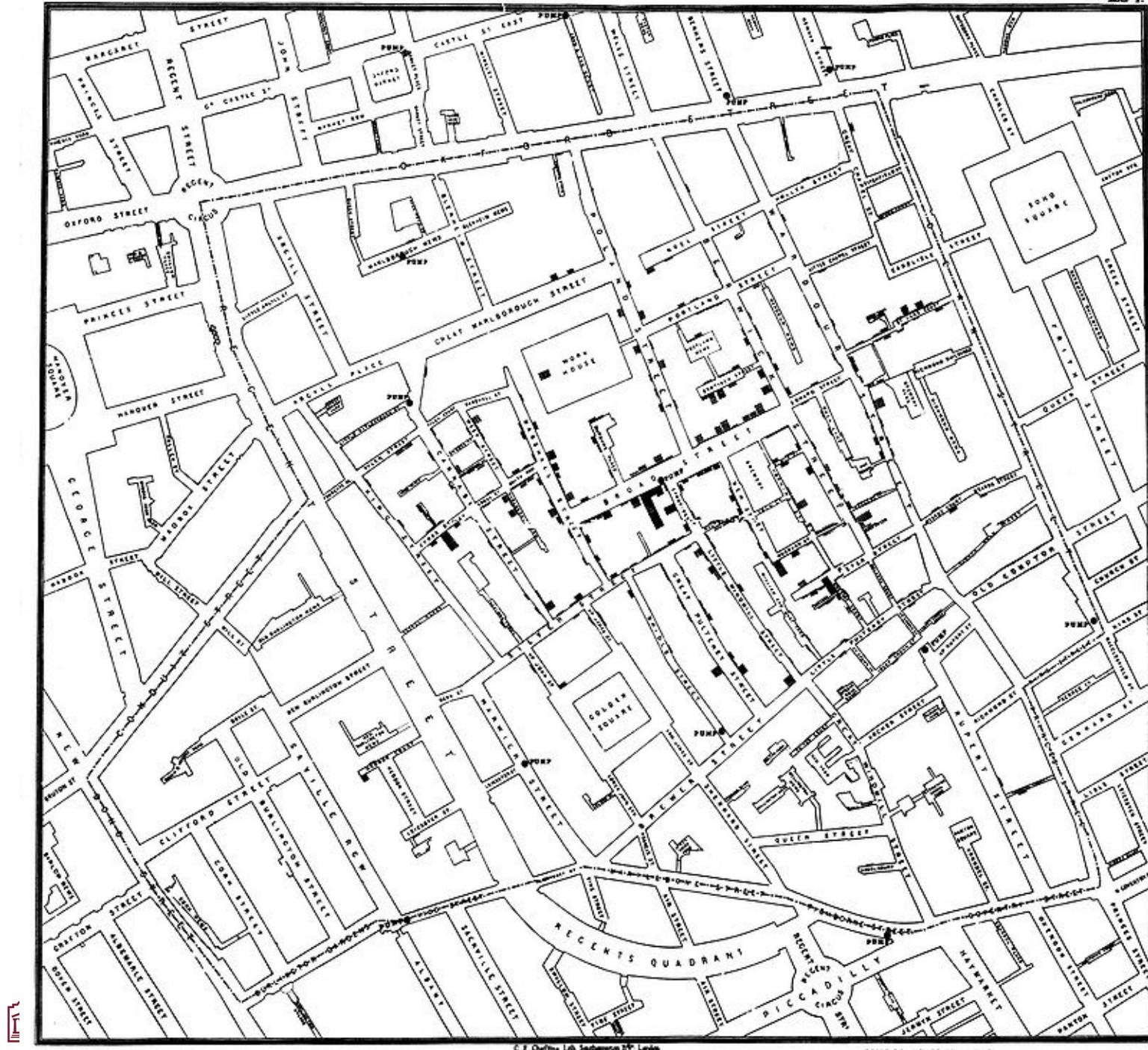
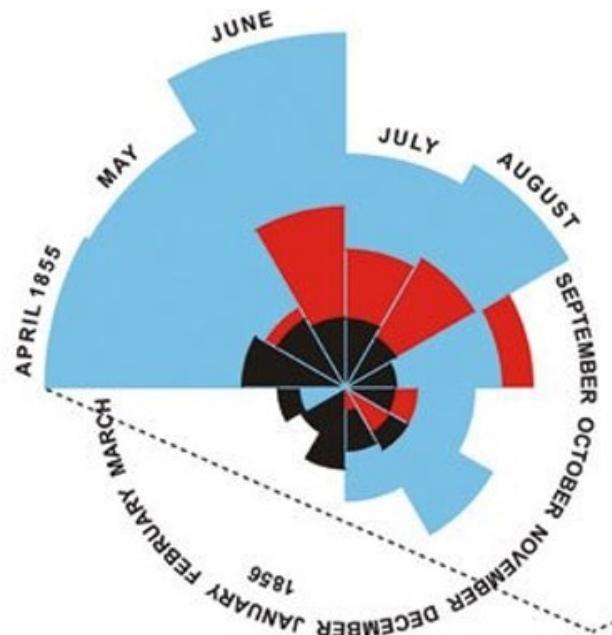
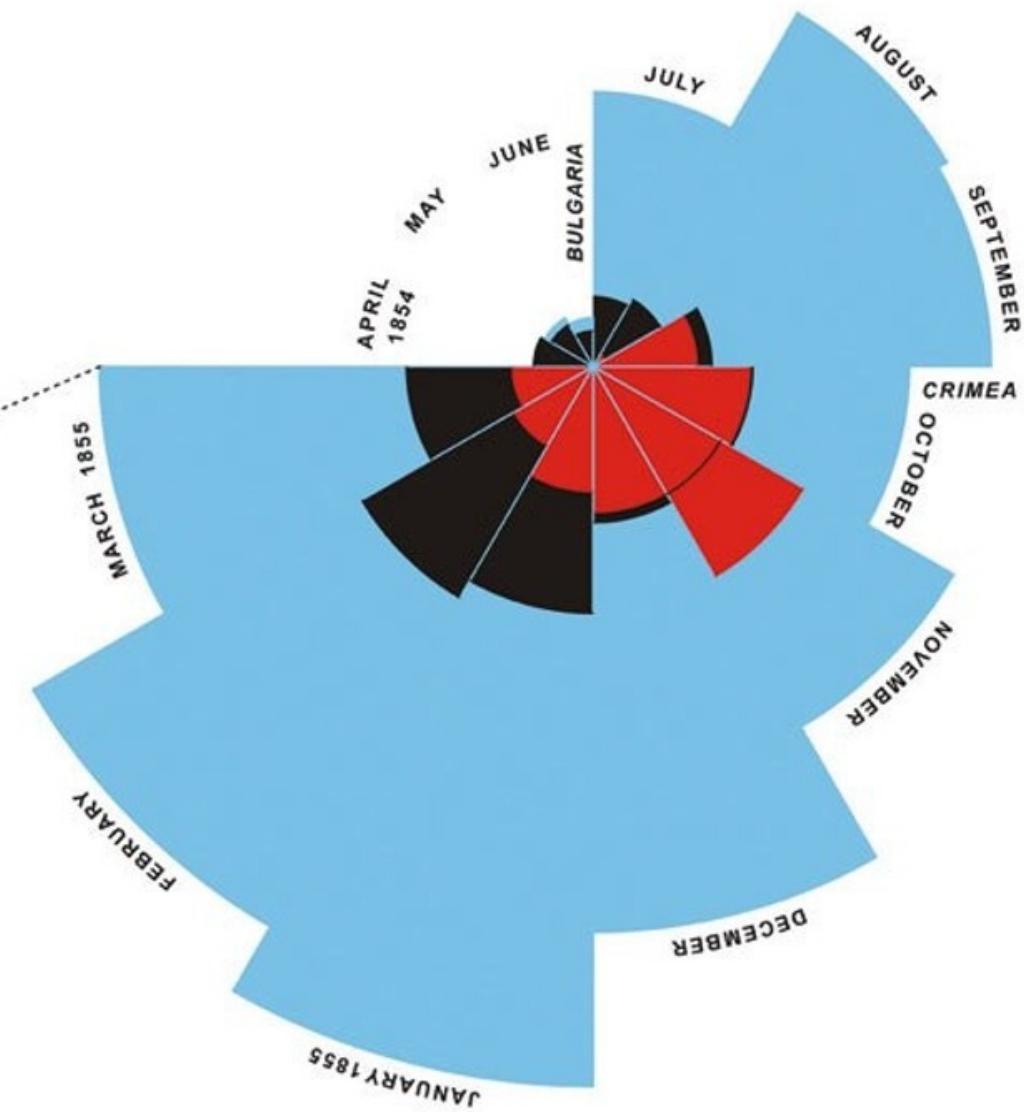


DIAGRAM OF THE CAUSES OF MORTALITY
IN THE ARMY IN THE EAST.

2.
APRIL 1855 TO MARCH 1856.



1.
APRIL 1854 TO MARCH 1855.



*The Areas of the blue, red, & black wedges are each measured from
the centre as the common vertex*

*The blue wedges measured from the centre of the circle represent area
for area the deaths from Preventible or Mitigable Zymotic Diseases, the
red wedges measured from the centre the deaths from wounds, & the
black wedges measured from the centre the deaths from all other causes*

*The black line across the red triangle in Nov' 1854 marks the boundary
of the deaths from all other causes during the month*

*In October 1854, & April 1855, the black area coincides with the red,
in January & February 1856, the blue coincides with the black*

*The entire areas may be compared by following the blue, the red & the
black lines enclosing them. ©hugh-small.co.uk*

EXERCISE – WHAT STORY IS MINDARD TELLING

- Take a look at the two upcoming slides **only** - no peeking at the one after (the first is in French and the second is in English)
- With no prompting take 5 mins in a group to work out what story Minard is telling

Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite
Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie; le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk en Mohilow et se rejoignaient vers Orscha et Wilebsk, avaient toujours marché avec l'armée.

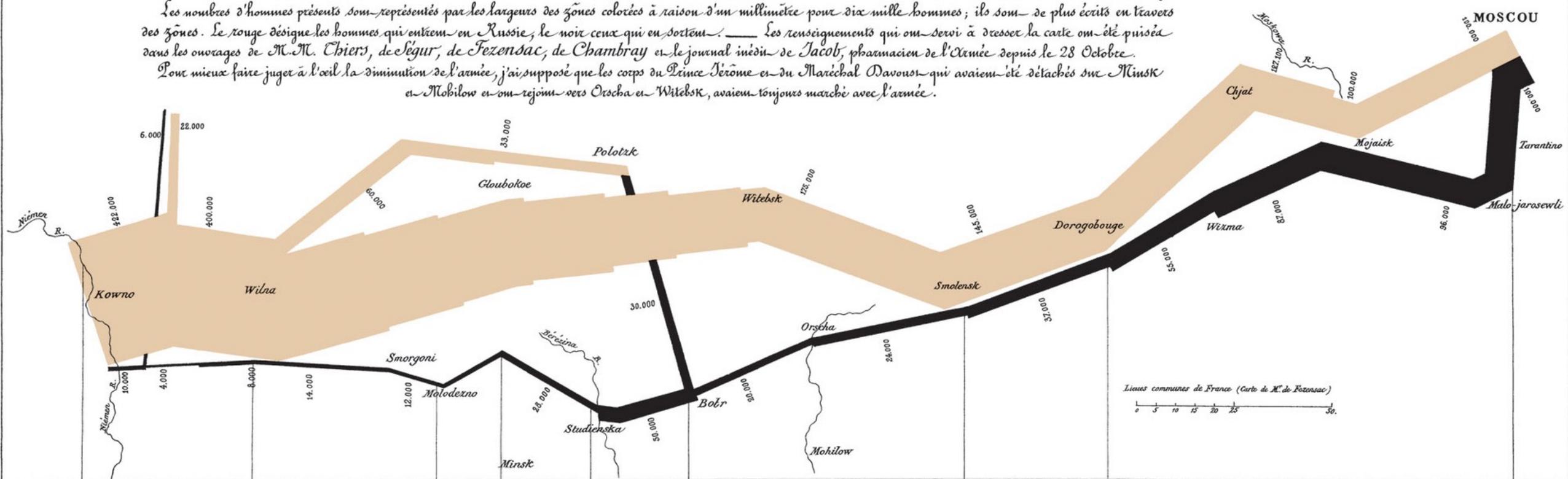
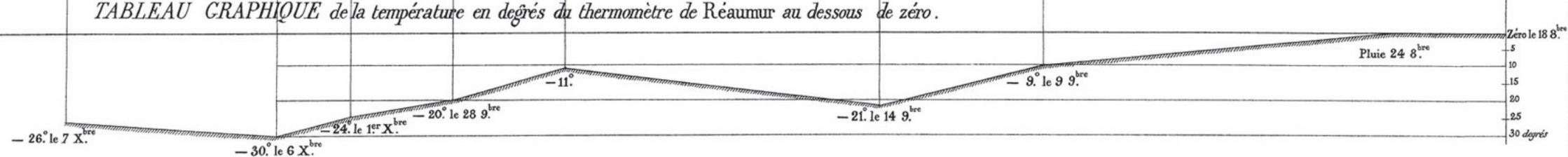


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

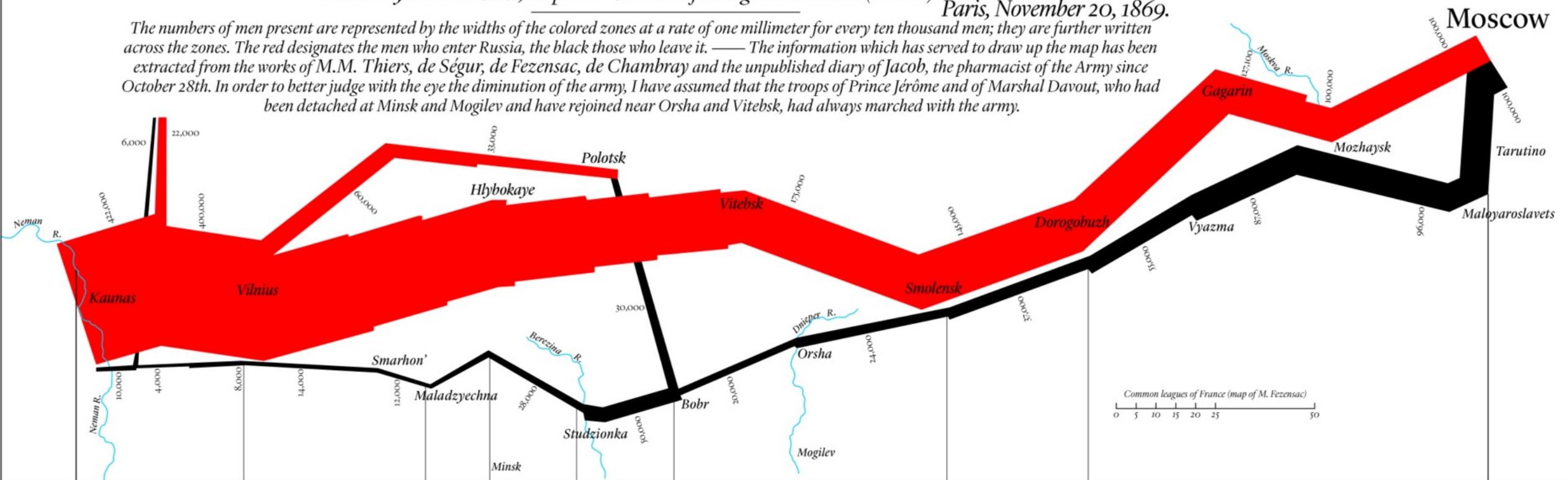
Les cosaques passent au galop le Niemen gelé.



*Figurative Map of the successive losses in men of the French Army in the Russian campaign 1812 ~ 1813
Drawn by M. Minard, Inspector General of Bridges and Roads (retired).*

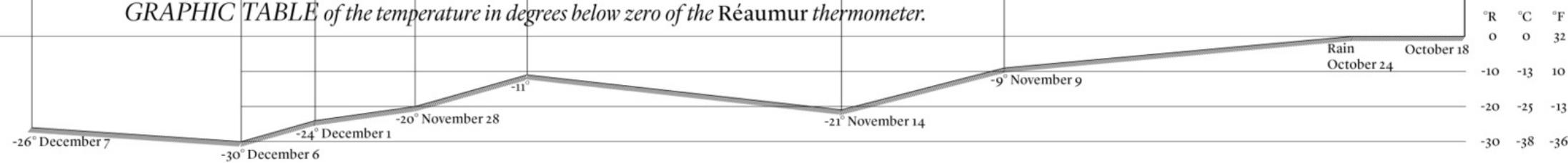
Paris, November 20, 1869.

The numbers of men present are represented by the widths of the colored zones at a rate of one millimeter for every ten thousand men; they are further written across the zones. The red designates the men who enter Russia, the black those who leave it. — The information which has served to draw up the map has been extracted from the works of M.M. Thiers, de Ségur, de Fezensac, de Chambray and the unpublished diary of Jacob, the pharmacist of the Army since October 28th. In order to better judge with the eye the diminution of the army, I have assumed that the troops of Prince Jérôme and of Marshal Davout, who had been detached at Minsk and Mogilev and have rejoined near Orsha and Vitebsk, had always marched with the army.



GRAPHIC TABLE of the temperature in degrees below zero of the Réaumur thermometer.

The Cossacks pass the frozen Neman at a gallop.



Minard's March to Moscow

1. Number of troops
2. Distance
3. Temperature

4. Location
5. Direction of travel
6. Location relative to dates

Russian campaign 1812 ~ 1813

Roads (retired).

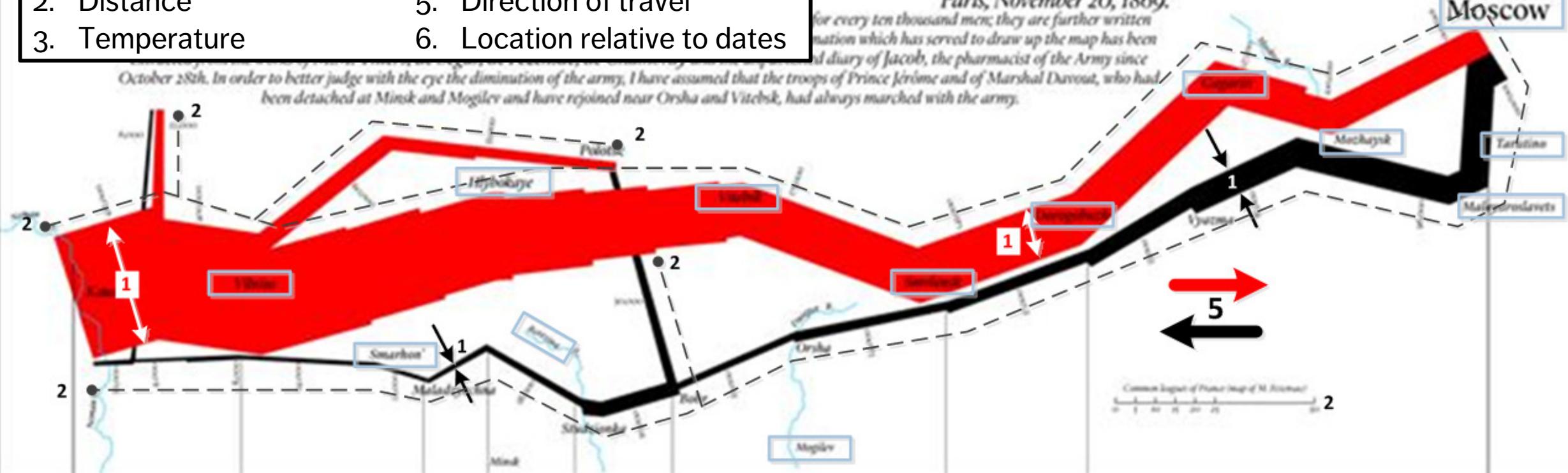
Paris, November 20, 1869.

for every ten thousand men; they are further written
information which has served to draw up the map has been
and diary of Jacob, the pharmacist of the Army since

4

Moscow

October 28th. In order to better judge with the eye the diminution of the army, I have assumed that the troops of Prince Jérôme and of Marshal Davout, who had been detached at Minsk and Mogilev and have rejoined near Orsha and Vitebsk, had always marched with the army.



GRAPHIC TABLE of the temperature in degrees below zero of the Réaumur thermometer.

The Cossacks pass the frozen
Neman at a gallop.



3

	°R	°C	°F
0	0	0	32
-10	-10	-13	10
-20	-20	-25	-12
-30	-30	-38	-36

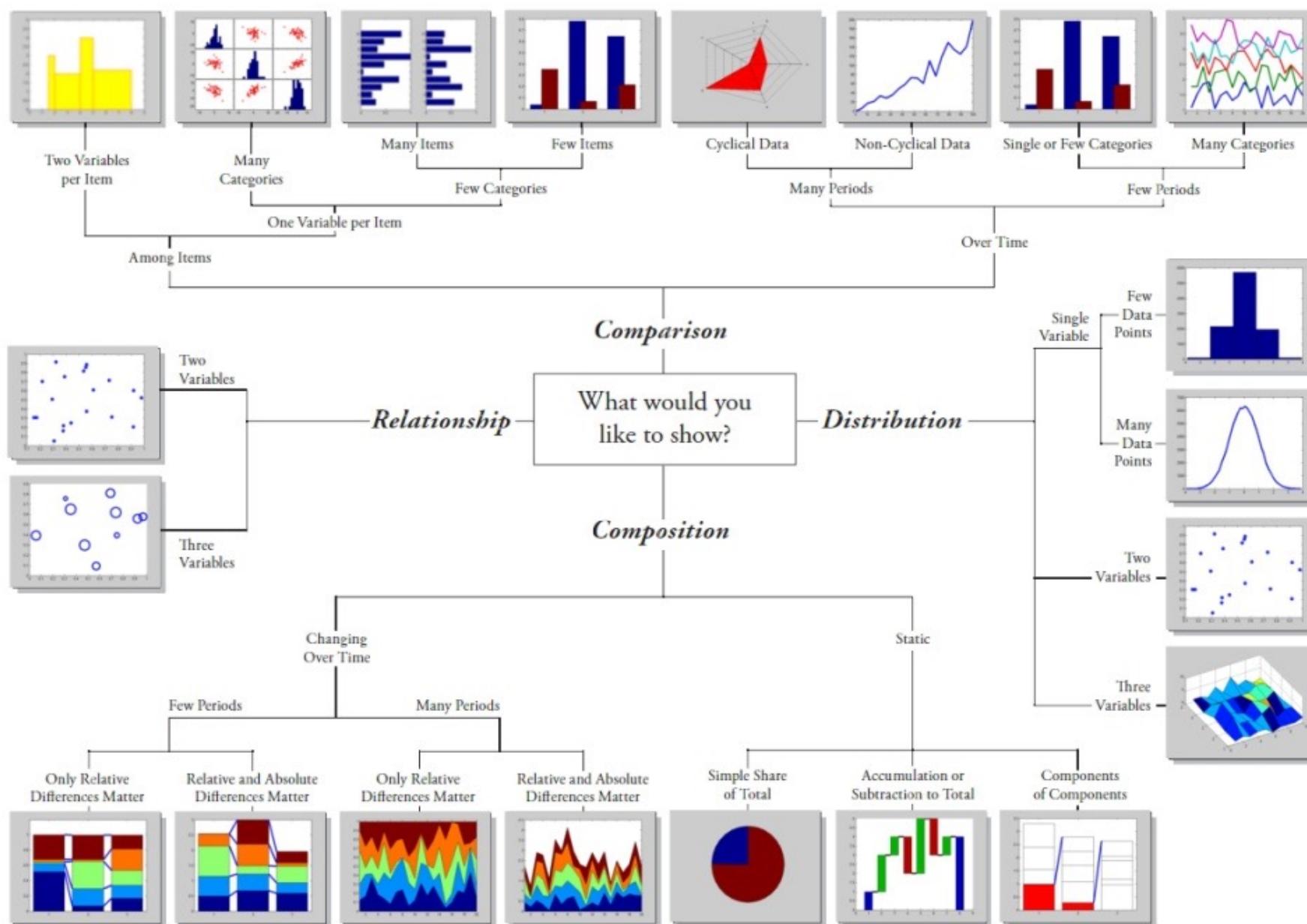
Minard's March to Moscow

TYPES OF CHARTS

With data visualizations, we want to highlight:

1. a **relationship** – show a connection or correlation between two or more variables, such as the impact of an aging population on health care;
2. a **comparison** – set some variables apart from others, and display how those two variables interact, or merely differ, such as the number of fans attending hockey games for different teams in a season;
3. a **composition** – collect different types of information that make up a whole and display them together, such as the various search terms that visitors used to land on your site, or how many visitors came from various sources (links, search engines, or direct traffic), and
4. a **distribution** – lay out a collection of related or unrelated information to see how it correlates (if at all), and to understand if there's any interaction between the variables, such as the number of bugs reported during each month after a new software release.

Chart Suggestions—A Thought-Starter



Modified with permission -Doug Hull
blogs.mathworks.com/videos
hull@mathworks.com 2009

www.ExtremePresentation.com
 © 2009 A. Abela — a.v.abela@gmail.com

TYPES OF CHARTS

Workhorse Data Visualizations

- Line Chart/Rug Chart/Number Line (data exploration)
- Histogram (data exploration)
- Boxplots (data exploration)
- Line Graph (data presentation + data exploration)
- Bar Chart (data presentation + data exploration)
- Scatterplot (data presentation + data exploration)

EXERCISE – HIGHLIGHTS AND CHARTS

Taking your previous example where we were looking at visual storytelling combinations:

1. Identify what items you might want to highlight (relationship, comparison, composition, distribution)
2. Think of some of the charts that might best represent both your items and the data that you use



LIMITATIONS OF DATA STORIES

What **constraints** exist on data stories?

Some constraints may be tied to the **function** (education vs. entertainment, say).

In this case, we are constrained to only tell stories that are **supported by the data**.

We can't tell just any old story we want to, even if we think it is the “right” story.

IMPACT OF CHOICES WHEN STORYTELLING WITH DATA

Data analysts have **agency**. They select:

- the question to answer;
- what data to collect;
- how to clean that data;
- which analytical method(s) to use;
- on what part(s) of the data to focus, etc.

This impacts the stories that **can be told** with data, relative to the stories that **could be told** about the situations and events represented by the data.

SCOPING VS. EXPLORATION VS. EXPLANATION VS. PERSUASION

When working with data, we create visualizations at **multiple stages** in the process.

This is reminiscent of the process behind **investigative journalism**:

1. initially, we **scope out** the area of investigation (data collection, story);
2. then we **explore** the situation and then **explore** the data we have collected about it
3. we may use the outcome of this exploration to **explain** the situation to our satisfaction;
4. and/or to **persuade** others about some course of action that should be taken with respect to the situation.

FALSIFICATION

Karl Popper differentiated science and pseudo-science by saying that scientific theories had to be **falsifiable** – this didn't mean they *had* to be false, but that it had to be *possible* for them to be false.

Similarly, with data storytelling, it should be possible for us to imagine some type of data that could *in principle* falsify the story we are telling.

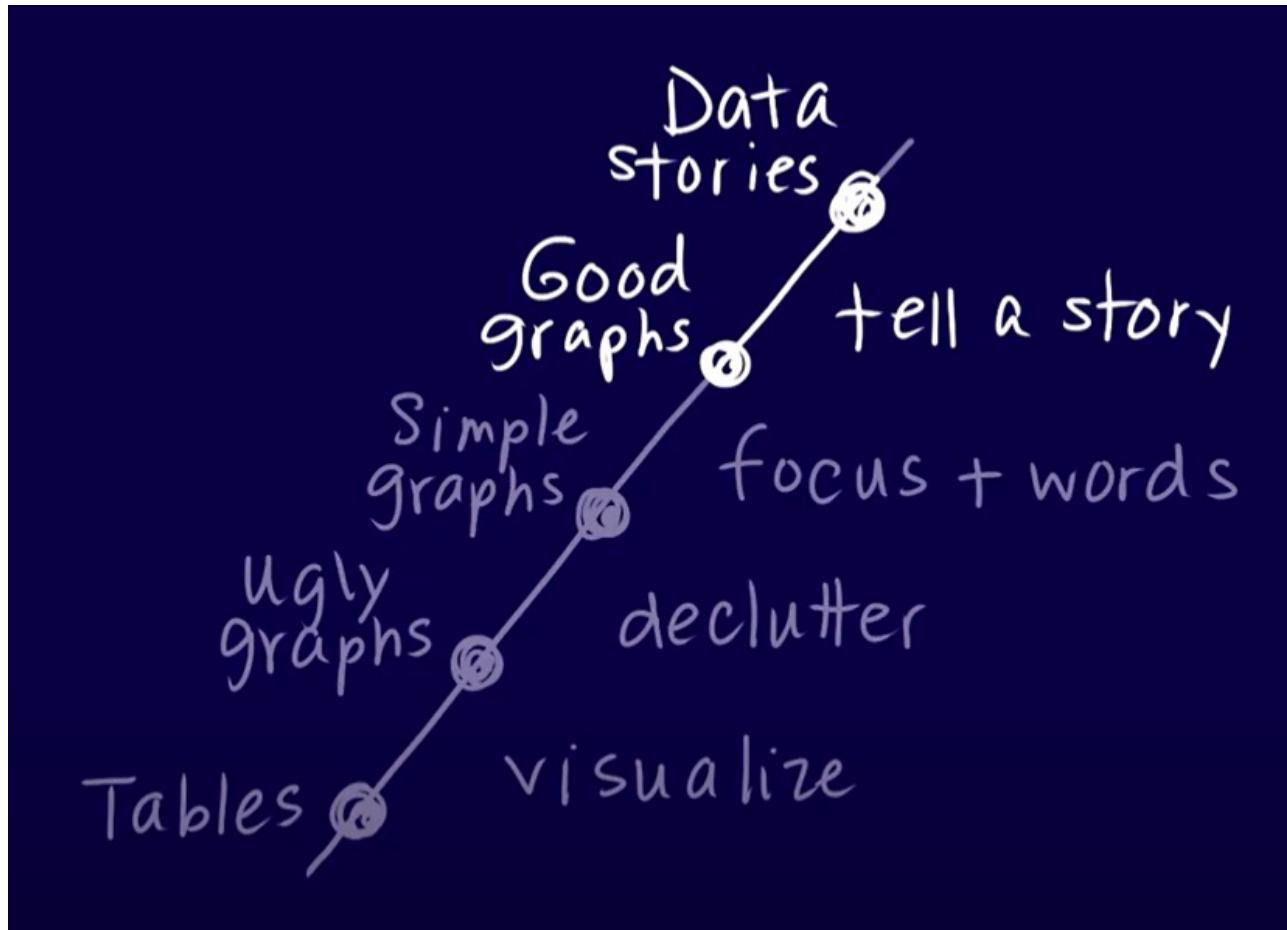
If we cannot do that, then the story and the data are not really connected.

EXERCISES

1. Identify instances of scoping, exploration, explanation, persuasion among the dashboards and charts from the two previous sections (main and exercises).
2. What do you think the underlying dataset structure and limitations are?
3. What analytical and data focus choices are at play?
4. Are the charts falsifiable?



EVOLVING A VISUALIZATION

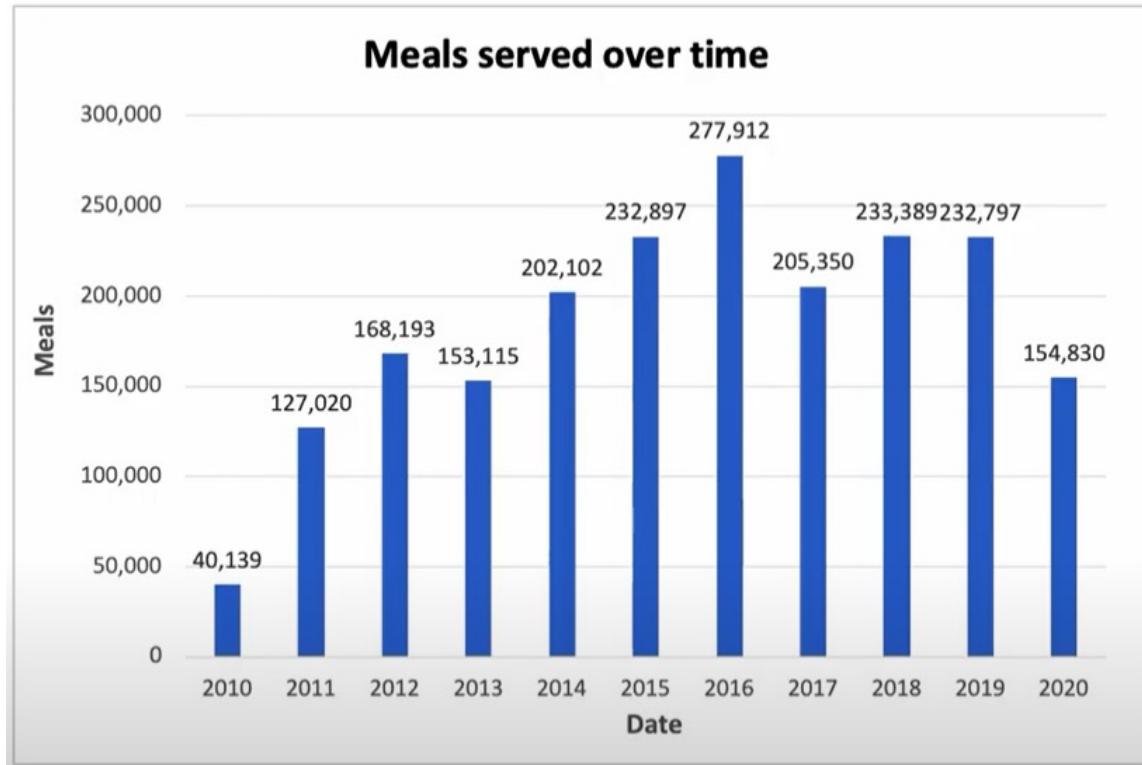


EVOLVING A VISUALIZATION – TABLE

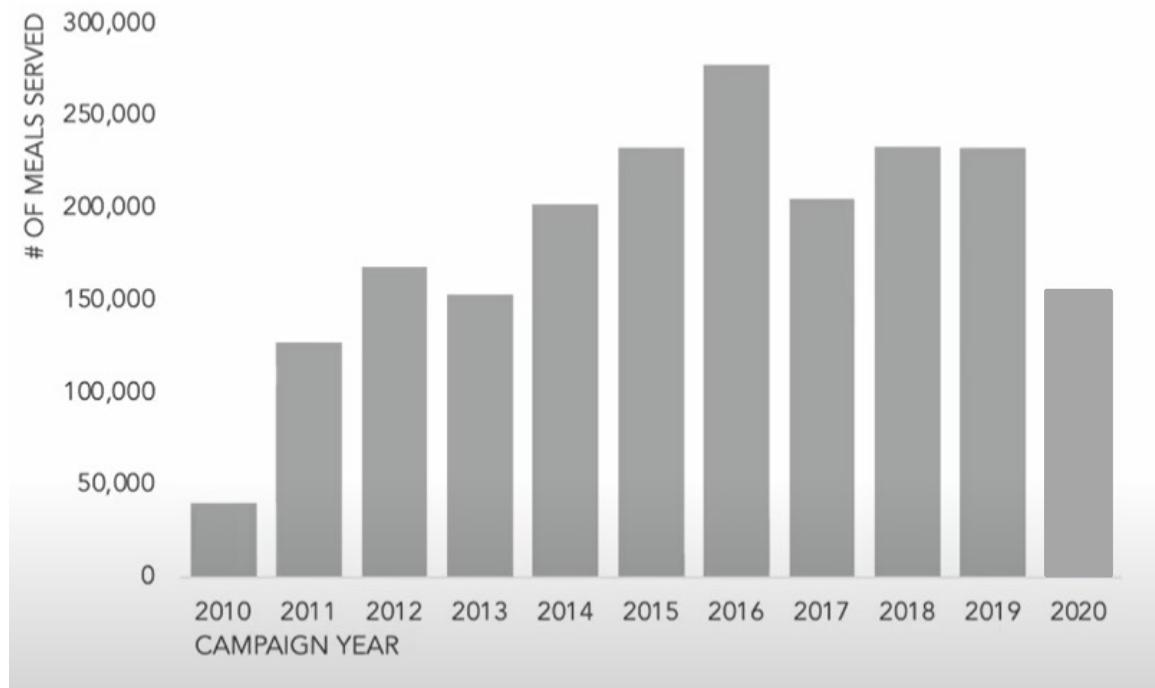
Meals served over time

Campaign Year	Meals Served
2010	40,139
2011	127,020
2012	168,193
2013	153,115
2014	202,102
2015	232,897
2016	277,912
2017	205,350
2018	233,389
2019	232,797
2020	154,830

EVOLVING A VISUALIZATION – UGLY GRAPH

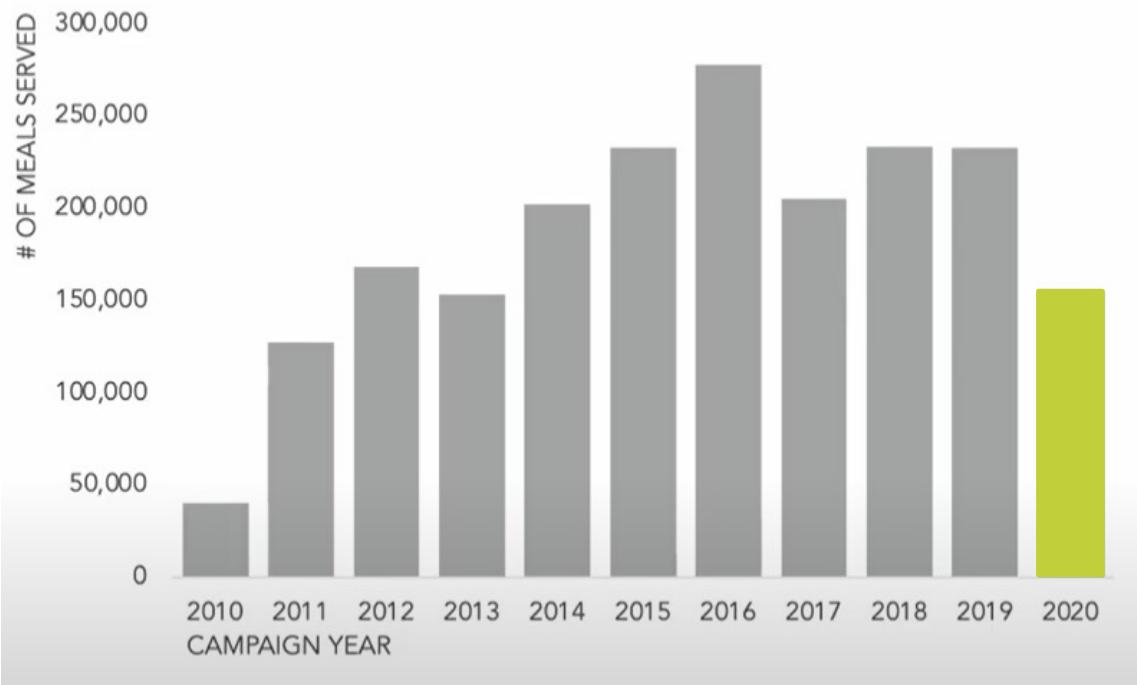


EVOLVING A VISUALIZATION – SIMPLE GRAPH



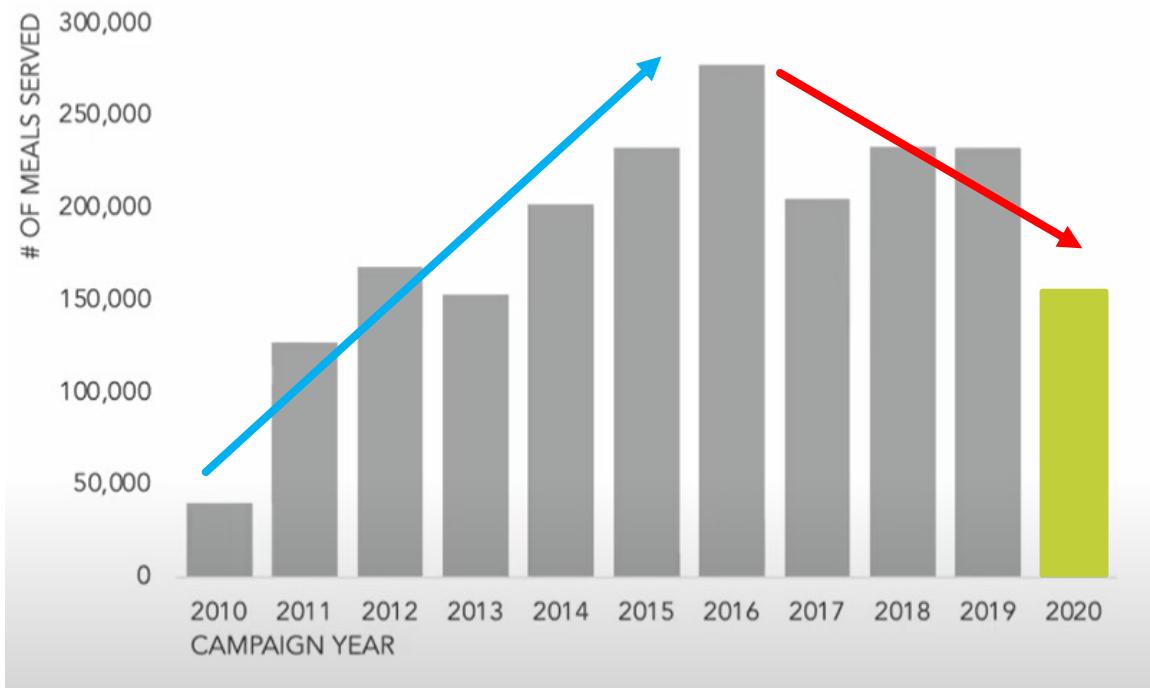
EVOLVING A VISUALIZATION – GOOD GRAPH

Meals served over time: **big drop in 2020**



EVOLVING A VISUALIZATION – DATA STORY

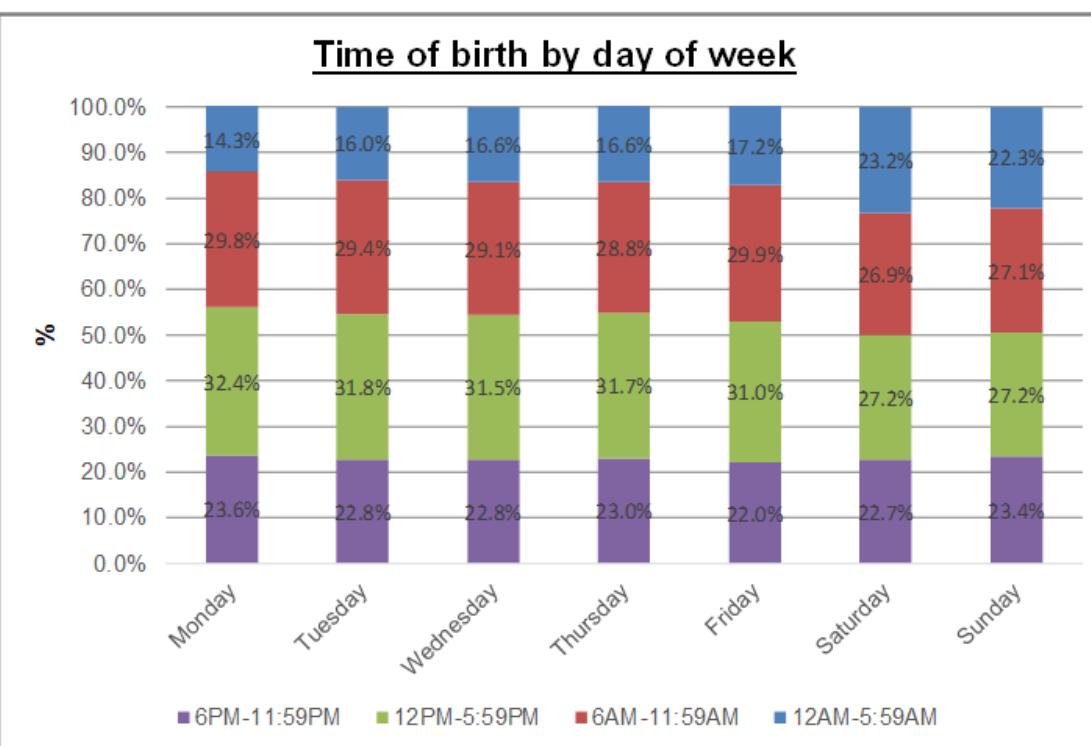
Meals served over time: **big drop in 2020**



Even though it might seem obvious that there would be a **big drop in meals served in 2020** due to the pandemic, note that the **2017-2019 numbers** were already going against the **2010-2016 trend** – we should not be planning for a return to 2016 levels without first understanding what happened in 2017-2019.

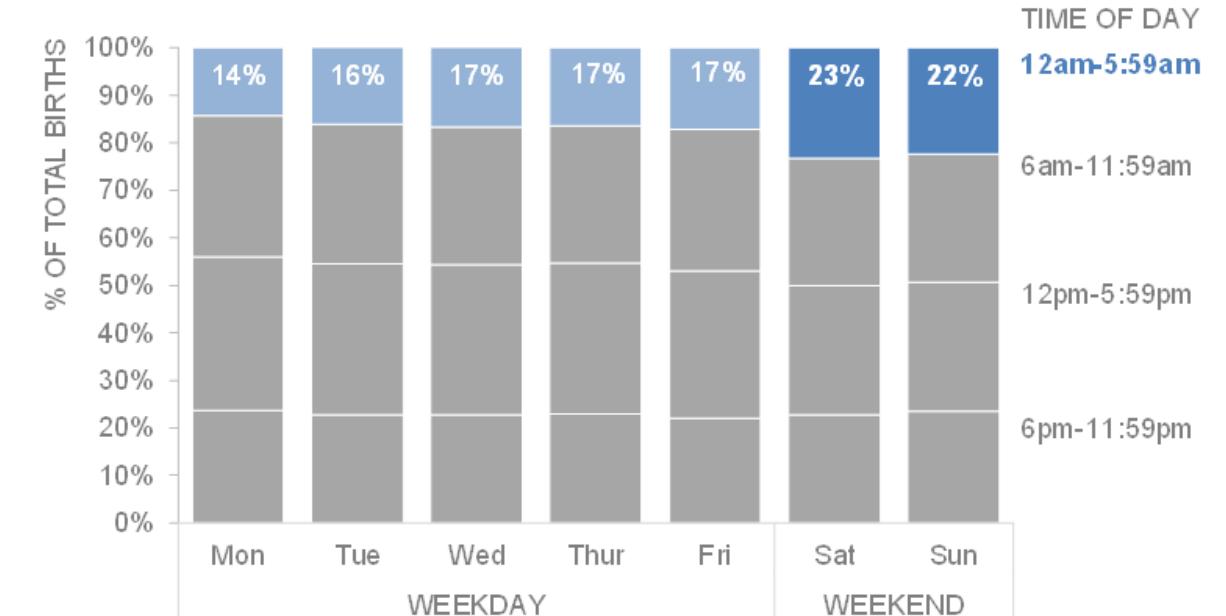
EXAMPLES

BEFORE



AFTER

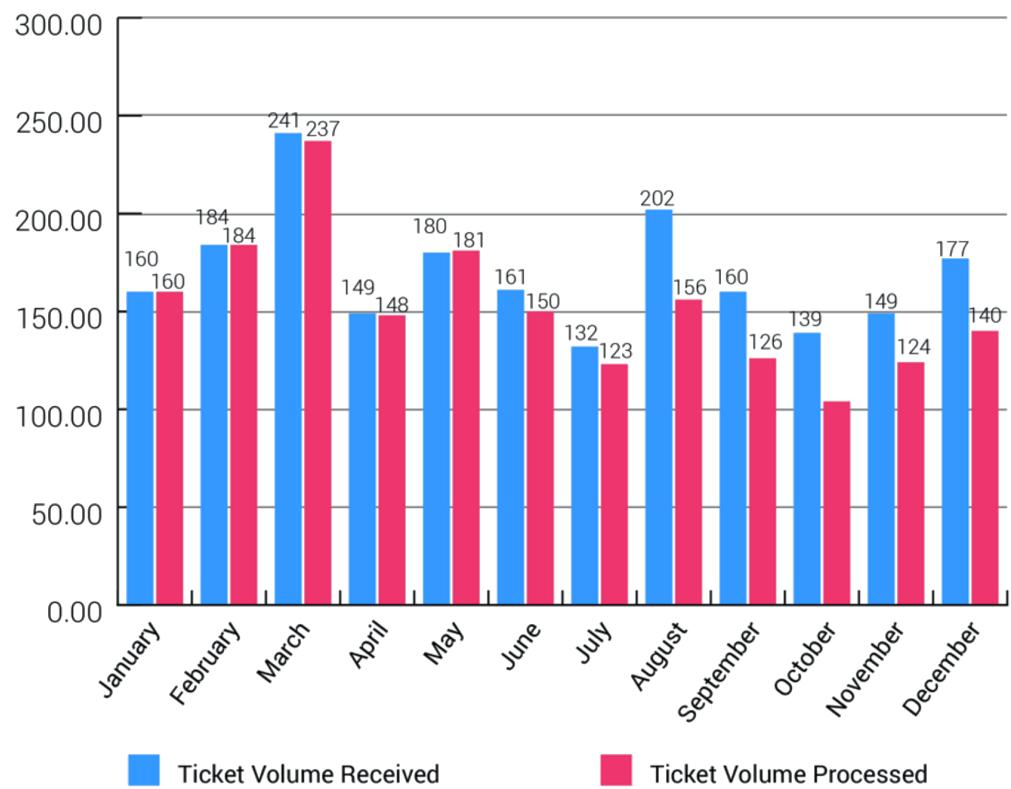
When babies are born
Weekend deliveries are more likely to be in early morning, compared to weekdays



EXAMPLES

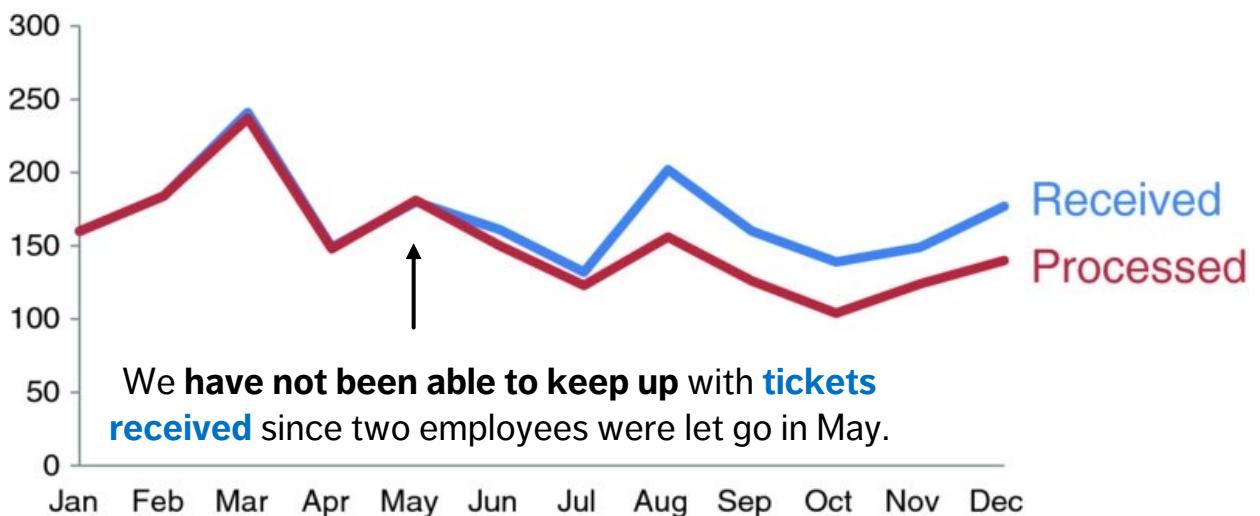
BEFORE

TICKET TREND



AFTER

Lag in Tickets Processed Since May Layoffs



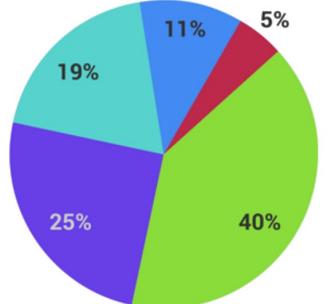
EXAMPLES

BEFORE

Survey Results

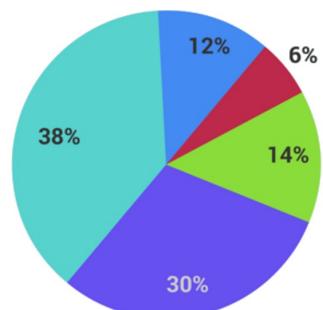
PRE: How do you feel about doing science?

■ Bored ■ Not great ■ OK ■ Kind of interested ■ Excited



POST: How do you feel about doing science?

■ Bored ■ Not great ■ OK ■ Kind of interested ■ Excited

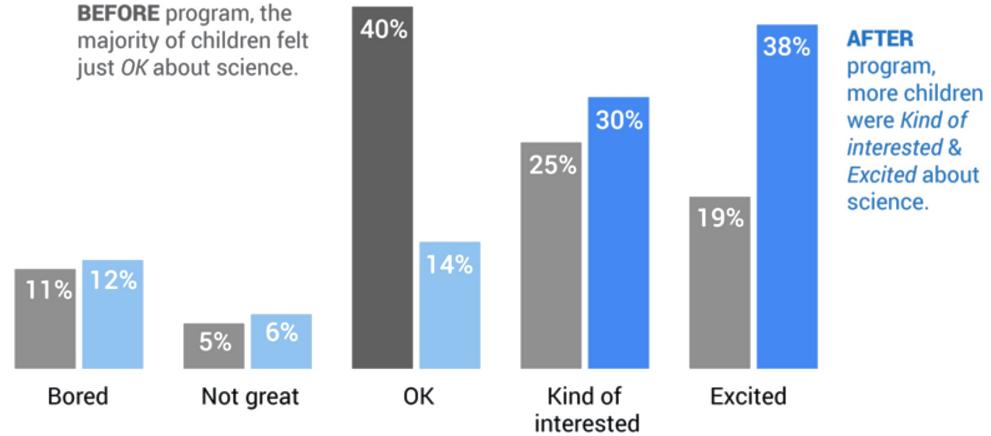


AFTER

Pilot program was a success

How do you feel about science?

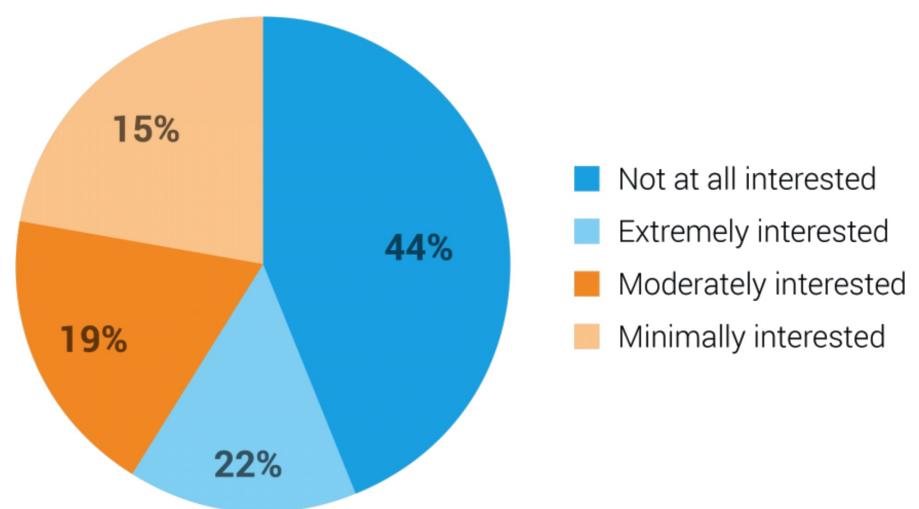
BEFORE program, the majority of children felt just *OK* about science.



EXAMPLES

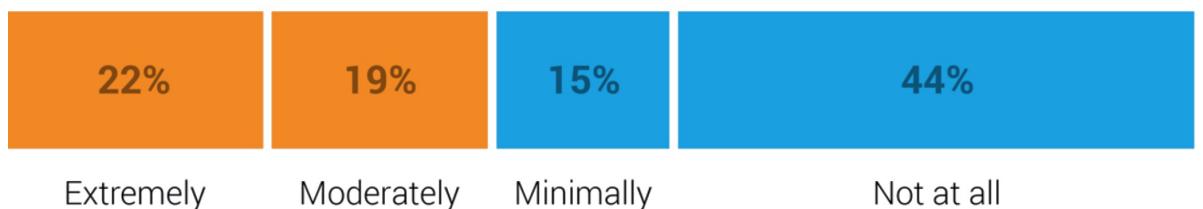
BEFORE

**HOW INTERESTED ARE YOU
IN THIS PRODUCT?**



AFTER

HOW INTERESTED ARE YOU IN THIS PRODUCT?



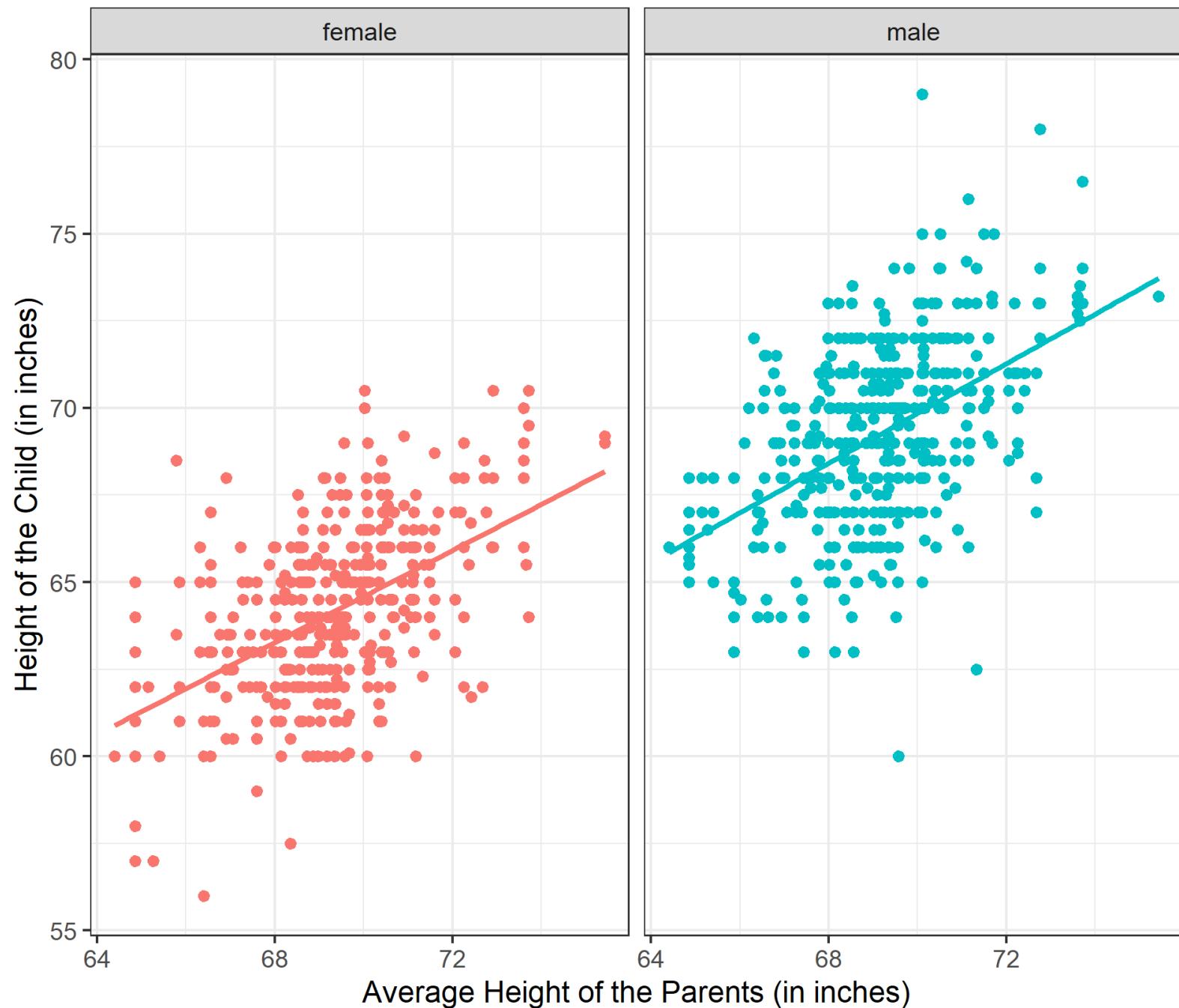
DATA STORYTELLING TROPES

Some data visualizations patterns are so familiar they have become **tropes** (icons):

- a scatterplot with a trend line going straight up or straight down
- a cluster bar chart with two categories where one is always lower than the other
- a line chart with the two lines crossing in one place
- pie charts being used all over the place
- red for republican, blue for democrat (US); red for left-leaning, blue for right-leaning (ROW)
- using broken axes to exaggerate effects
- etc.

Scatterplot matrix of Galton Family Data by Gender of the Child

<https://www.chsglobe.com/13376/cover-stories/sexualharassment>



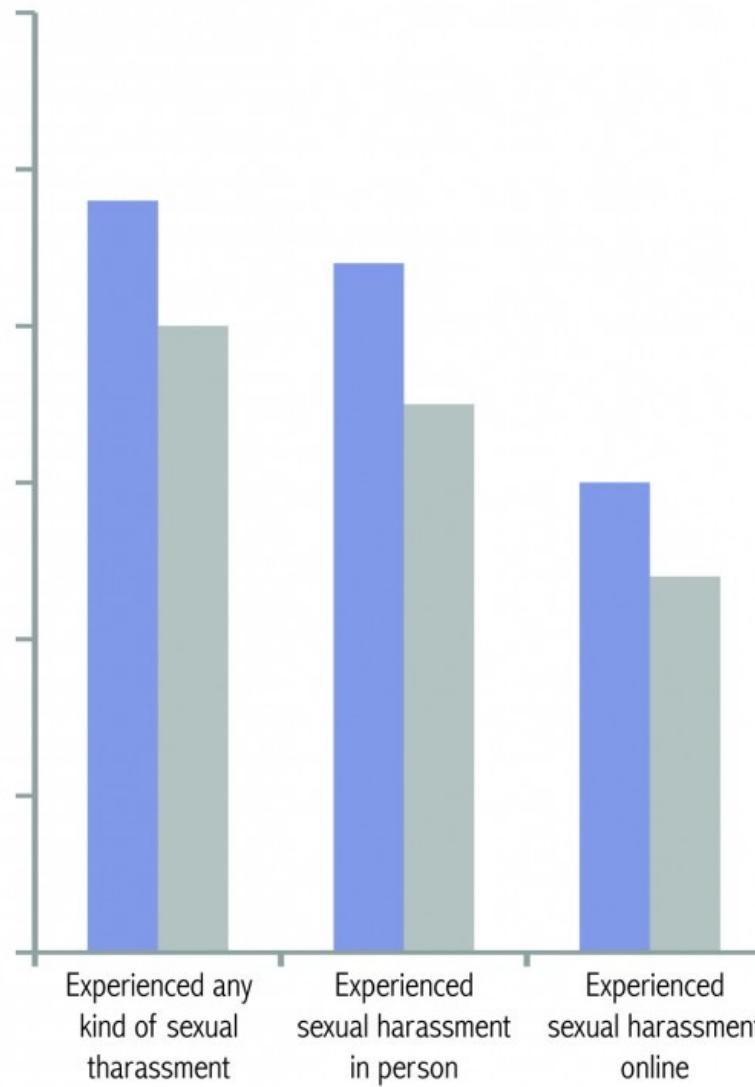
NATIONAL CRISIS STUDENT SEXUAL HARASSMENT

7-12 graders, %

SOURCE: AAUW report

Boys

Girls

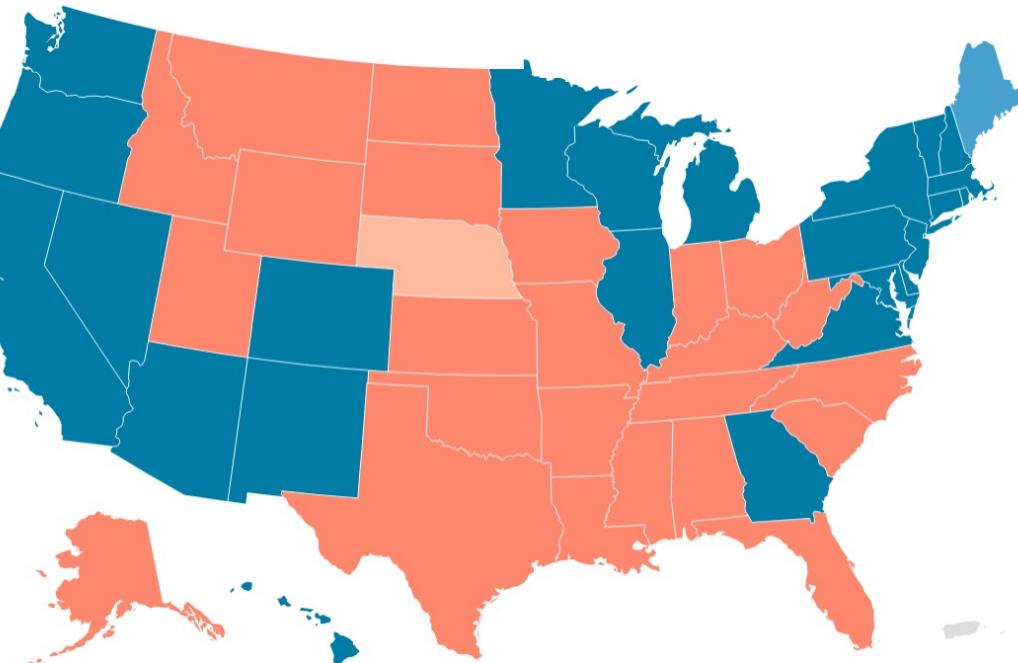


DATA STORYTELLING TROPES – EXAMPLES

Conventional Map of 2020 US Presidential Election Results

Maine and Nebraska allow some electoral votes to be split by district

Biden Biden + Trump Trump Trump + Biden

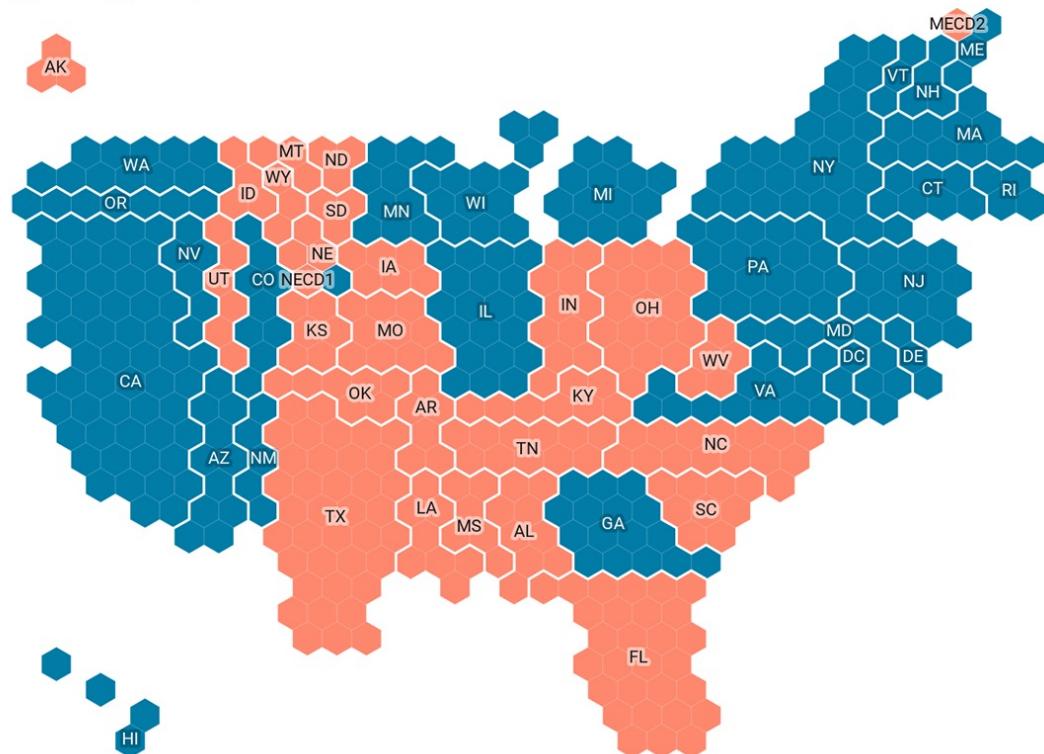


Created with Datawrapper

Cartogram of 2020 US Presidential Election Results

Each hexagon represents one electoral college vote

Biden Trump

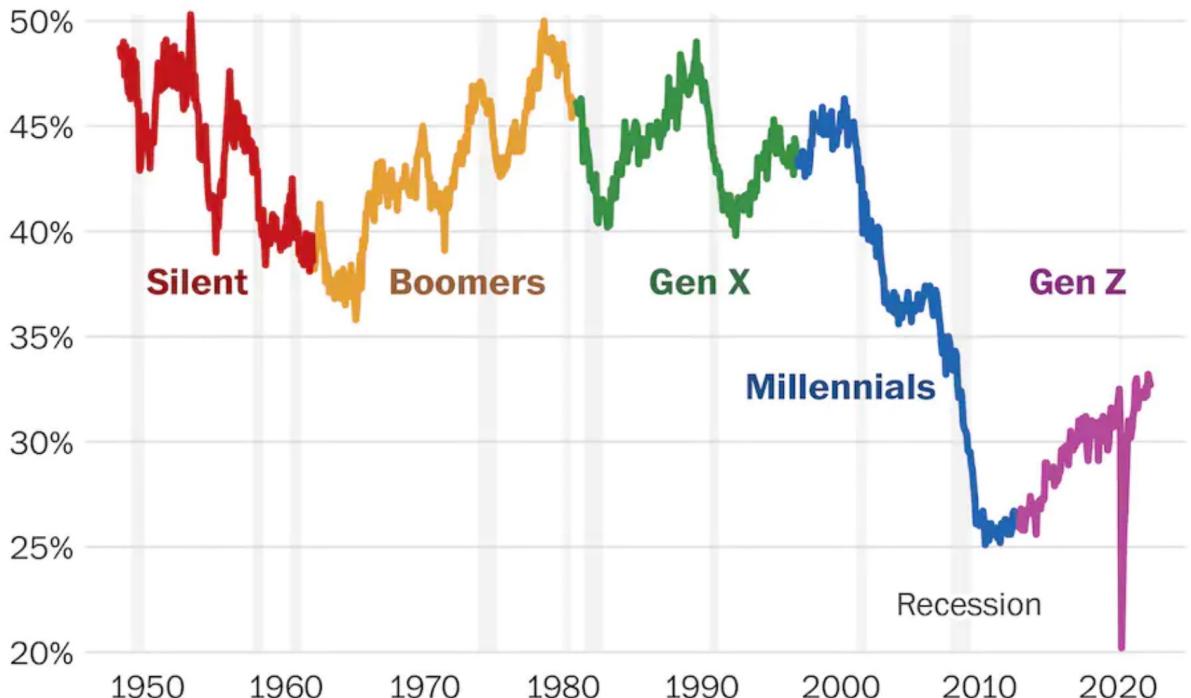


EXERCISE – CHARTS TO STORIES

Evolve the following charts into data stories. Focus on the message and how to avoid misleading the audience.

Teen work makes the dream work

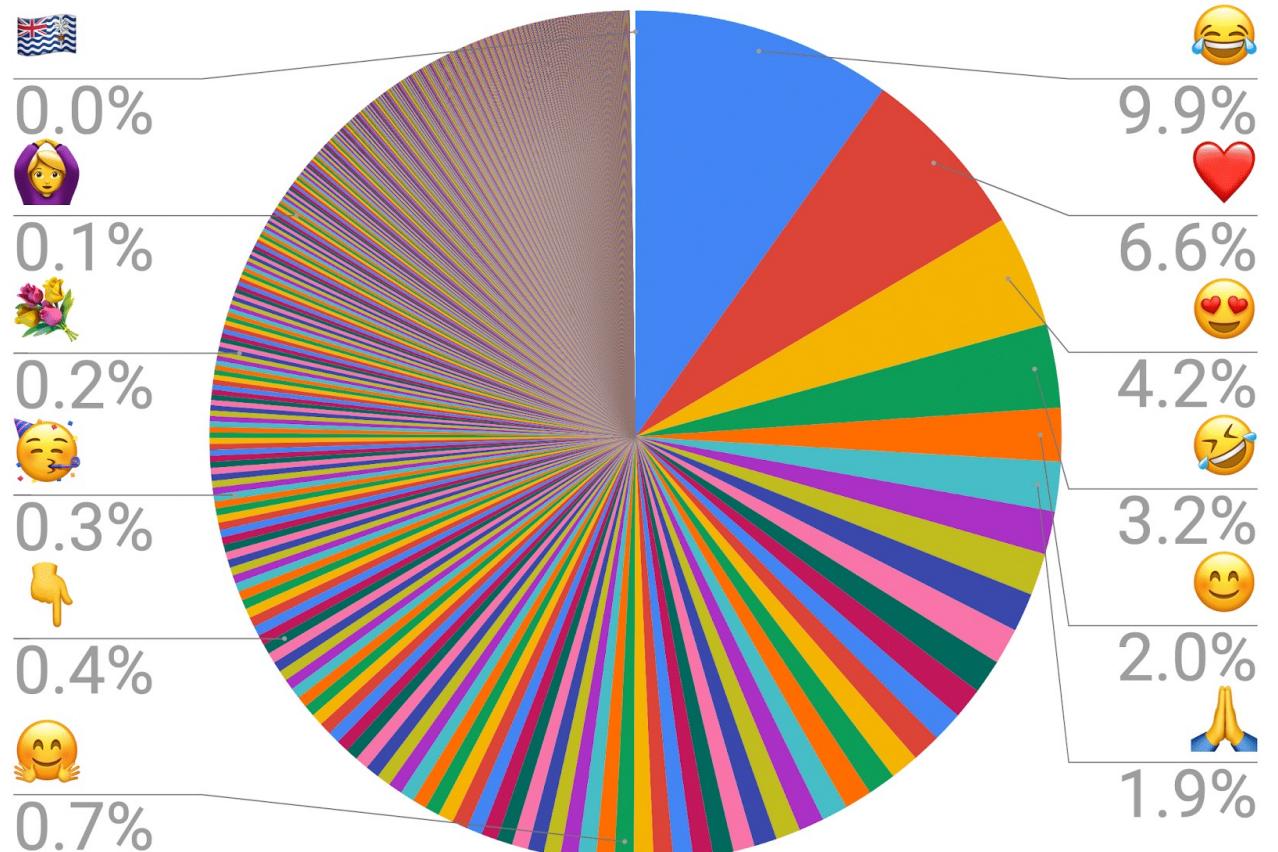
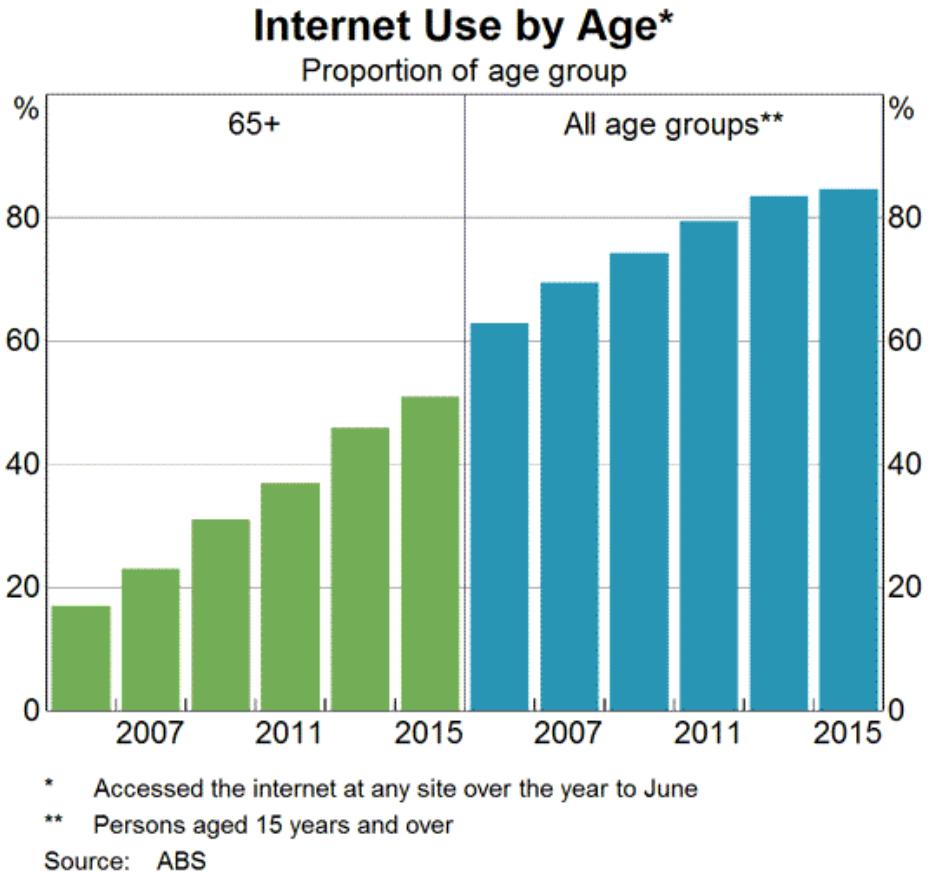
Employment-to-population ratio for those ages 16 to 19



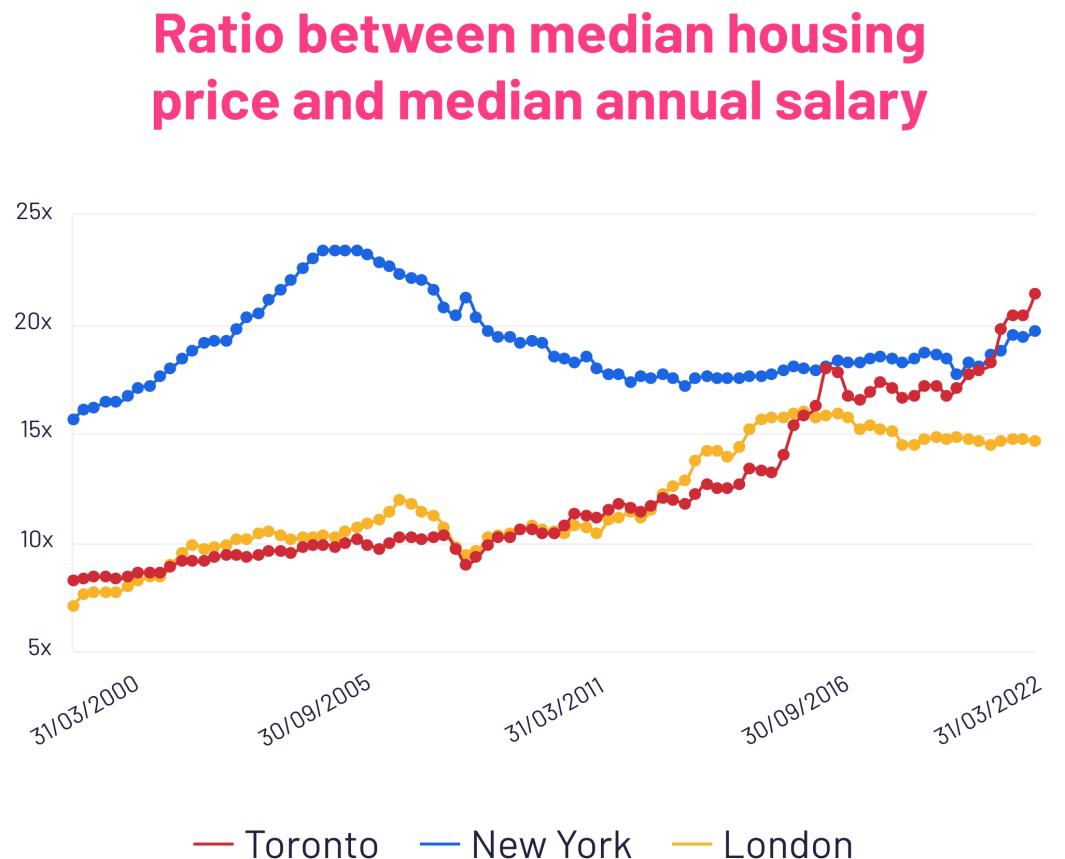
Note: Seasonally adjusted

Source: Bureau of Labor Statistics

EXERCISE – CHARTS TO STORIES



EXERCISE – CHARTS TO STORIES

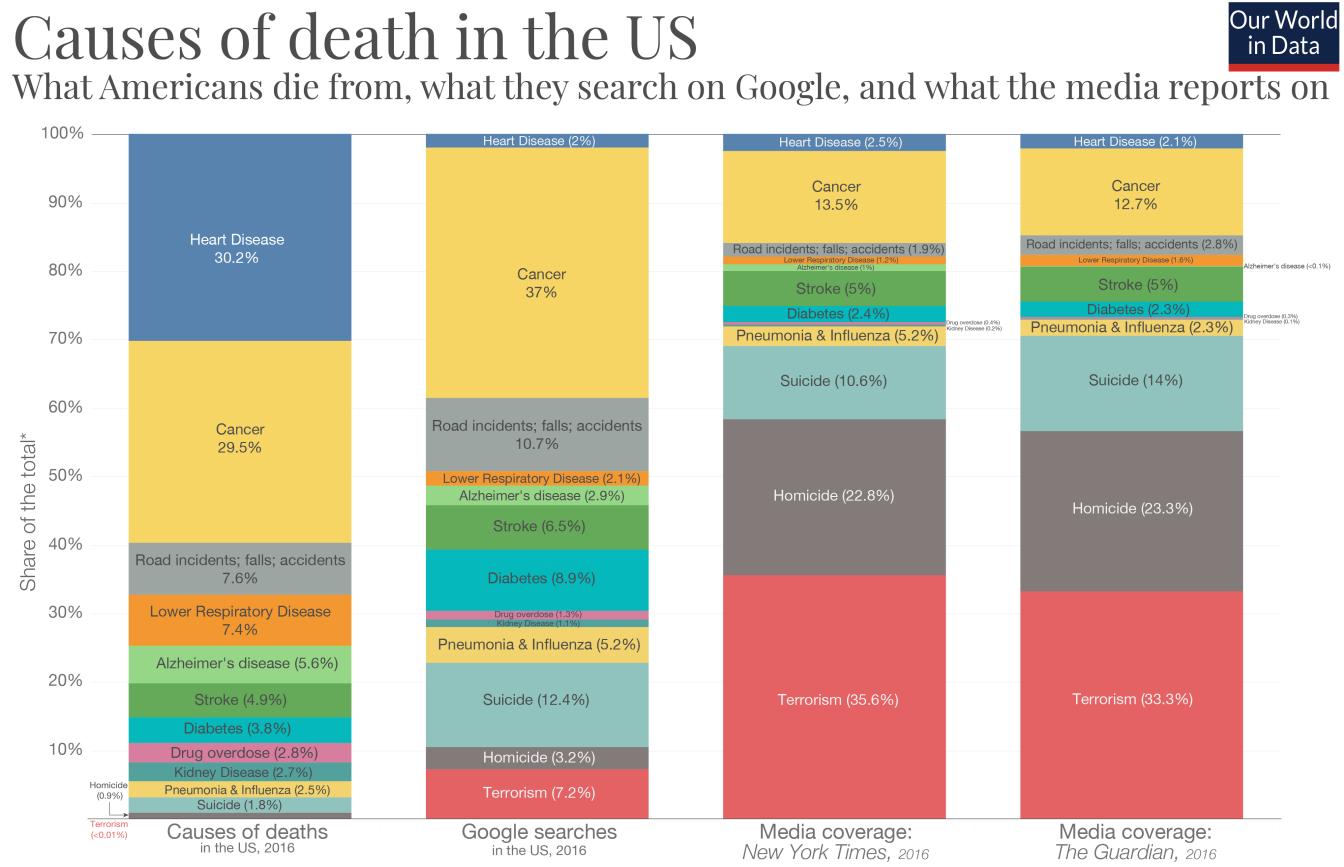


Methodology: Median housing price in each city divided by median pre-tax annual salary

Source: Bloomberg

More charts at (link in bio): genuineimpact.substack.com

Created by genuine impact



*This represents each cause's share of the top ten causes of death in the US plus homicides, drug overdoses and terrorism. Collectively these 13 causes accounted for approximately 88% of deaths in the US in 2016. Full breakdown of causes of death can be found at the CDC's WONDER public health database: <https://wonder.cdc.gov/>

Based on data from Shen et al (2018) – Death: reality vs. reported. All data available at: <https://owenshen24.github.io/charting-death>

All data refers to 2016.

Not all causes of death are shown: Shown is the data on the ten leading causes of death in the United States plus drug overdoses, homicides and terrorism.

All values are normalized to 100% so they represent their relative share of the top causes, rather than absolute counts (e.g. 'deaths' represents each cause's share of deaths within the 13 categories shown rather than total deaths). The causes of death shown here account for approximately 88% of total deaths in the United States in 2016.

This is a visualization from OurWorldInData.org, where you find data and research on how the world is changing.

Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

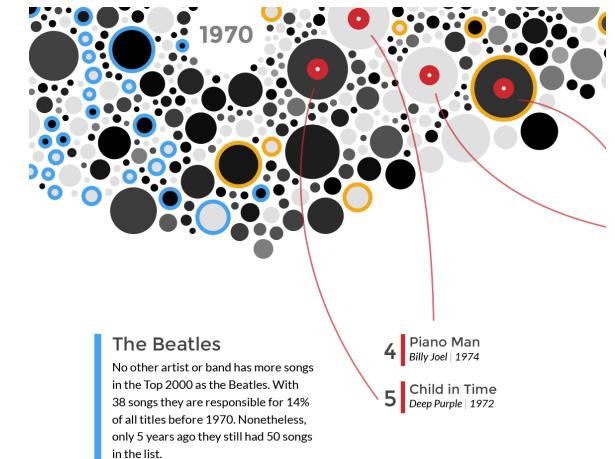
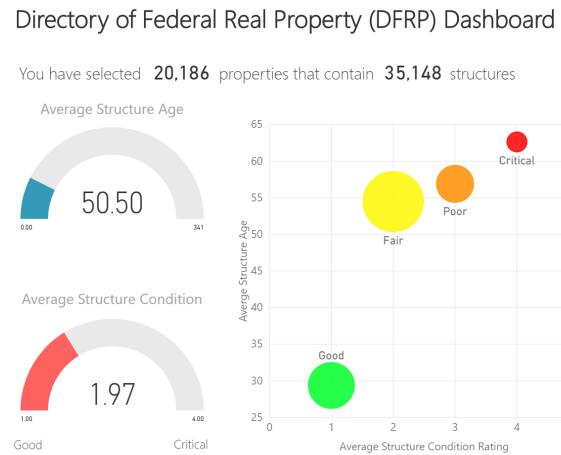
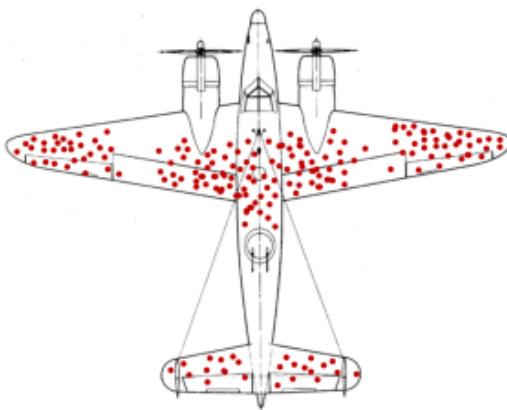


ANATOMICAL CONSIDERATIONS

The composition of a (storytelling) dashboard must consider various components:

- the audience
- the goals
- the dashboard's narrative
- the narrative's logic
- iconic memory
- short-term memory
- long-term memory

DEFINING CONTEXT



Seconds

Minutes

Fraction of Hour

Hours

← Infographics/Data Viz →

← Dashboards →

← Reports and Exploration →

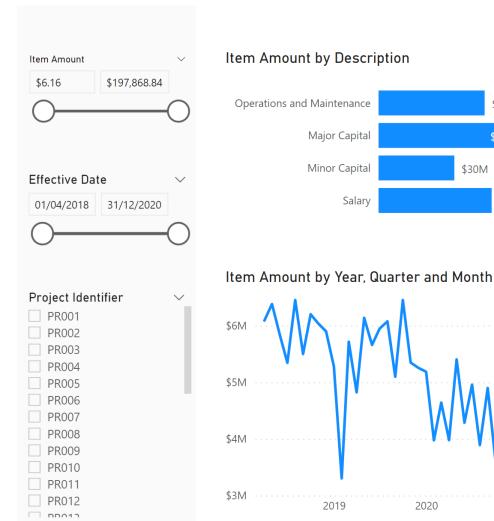
← Data Art →

EXPLORATION VS STORYBOOK VS SITUATIONAL AWARENESS

Exploration: using visualizations as a tool to explore data

- high level of interactivity
- high level of detail
- all aspects of data should be represented (tables, columns, calculations etc.)
- no annotations or explanations required

Financial Data Exploration



\$59.78K

Average of Item Amount

\$173.78M

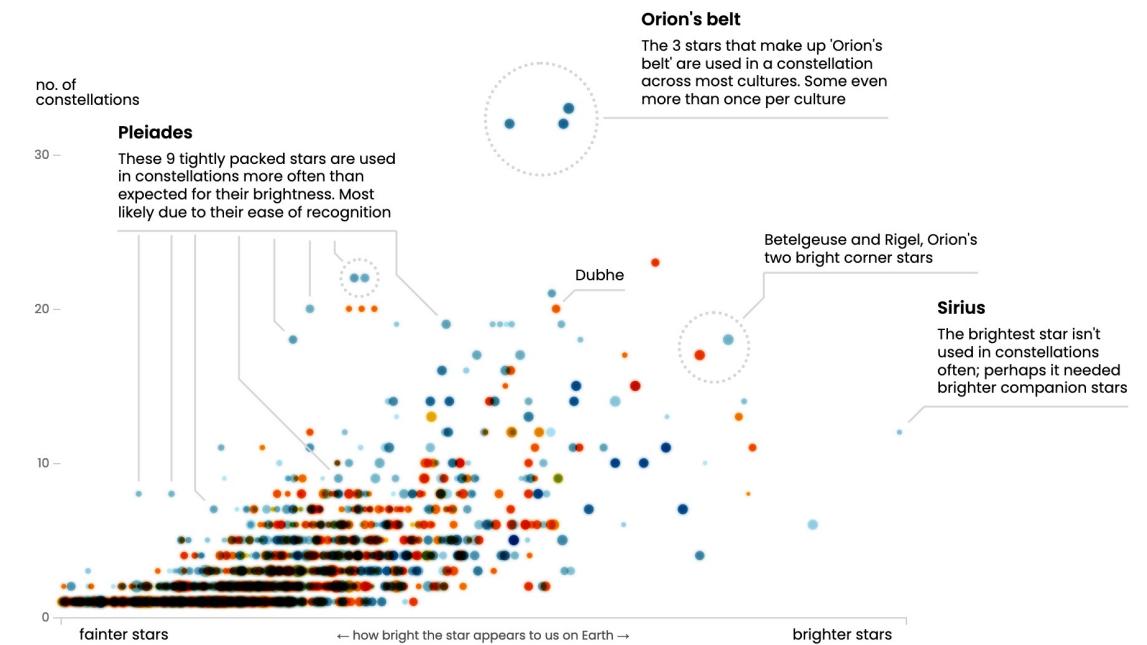
Sum Item Amount

Journal Voucher Type Code	2018	2019	2020	Total
MC	\$18,110,444.61	\$21,810,187.90	\$17,448,129.75	\$57,368,762.27
PR001	\$687,677.21	\$1,792,024.46	\$907,127.55	\$3,386,829.23
PR002	\$788,825.39	\$565,031.07	\$813,175.69	\$2,167,032.15
PR003	\$1,517,664.95	\$612,091.00	\$1,093,131.35	\$3,222,887.30
PR004	\$800,174.27	\$719,551.46	\$1,155,498.57	\$2,675,224.30
PR005	\$611,844.01	\$1,559,623.99	\$505,962.54	\$2,677,430.55
PR006	\$869,847.19	\$1,142,078.50	\$567,309.21	\$2,579,234.90
PR007	\$1,254,247.56	\$1,202,463.46	\$1,121,613.47	\$3,578,324.48
PR009	\$536,301.11	\$1,466,714.57	\$654,848.18	\$2,657,863.87
PR010	\$1,025,185.44	\$1,124,411.66	\$810,384.12	\$2,959,981.22
PR011	\$1,323,665.62	\$947,916.20	\$951,129.63	\$3,222,711.45
PR012	\$894,949.35	\$1,321,602.78	\$1,142,398.09	\$3,358,950.22
PR013	\$810,720.00	\$1,397,946.44	\$943,871.63	\$3,152,538.13
PR015	\$1,115,244.24	\$1,238,919.57	\$1,211,122.76	\$3,565,286.57
PR017	\$1,163,245.06	\$1,346,151.02	\$595,533.30	\$3,104,929.39
PR018	\$868,426.84	\$1,297,779.23	\$1,177,356.88	\$3,362,962.95
PR019	\$942,777.50	\$1,028,710.89	\$748,386.14	\$2,719,874.53
PR022	\$842,076.88	\$697,992.57	\$1,105,900.34	\$2,645,969.79
PR023	\$1,219,843.67	\$1,143,895.90	\$1,115,052.77	\$3,478,792.34
PR027	\$817,728.27	\$1,205,883.13	\$828,327.52	\$2,851,938.92
MIC	\$8,733,325.92	\$11,316,310.76	\$9,855,321.54	\$29,904,958.22
PR001	\$488,147.03	\$447,373.91	\$493,012.00	\$1,428,532.94
PR002	\$288,526.70	\$794,250.21	\$275,485.45	\$1,358,262.36
PR003	\$249,707.20	\$301,928.04	\$339,914.44	\$891,549.68
Total	\$53,750,707.93	\$65,112,880.21	\$54,913,391.39	\$173,776,979.54

EXPLORATION VS STORYBOOK VS SITUATIONAL AWARENESS

Storybook: using visualizations as a tool to explain data

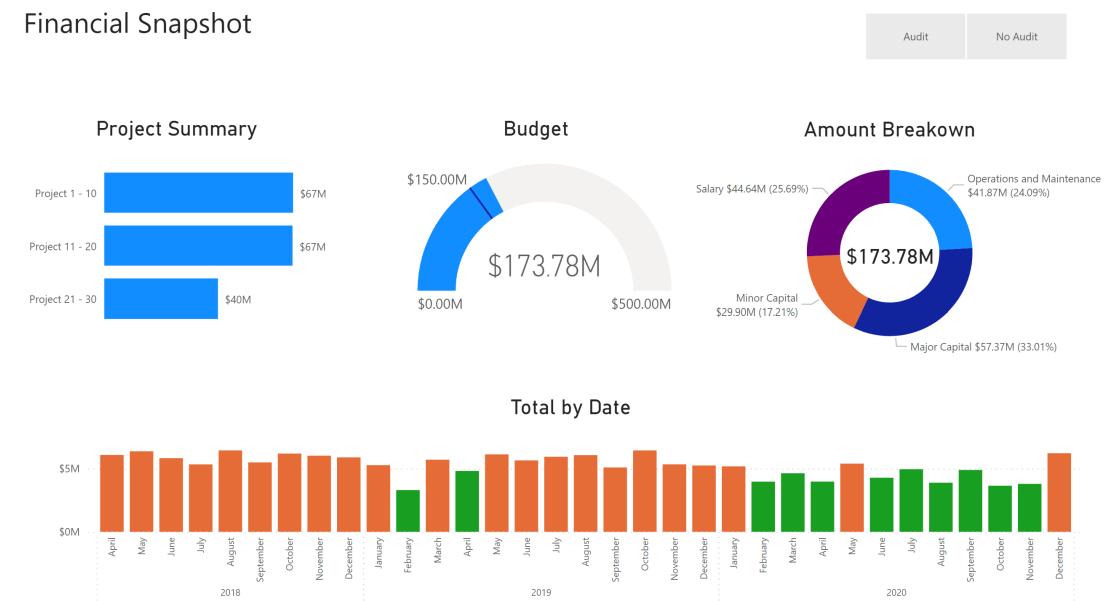
- low level of interactivity
- low level of detail
- key aspects of data should be represented
- annotations and explanations drive the “story”



EXPLORATION VS STORYBOOK VS SITUATIONAL AWARENESS

Situational Awareness: using visualizations as a tool to provide a snapshot of the data

- medium level of interactivity
- not “scripted” but well organized (e.g., categorized)
- summary data should be represented
- anomalies are highlighted



CREATING A NARRATIVE

There are a number of ways of constructing a **narrative**, including:

- chronological
- most important first, or least important first
- begin with the end
- success first, bad news last, or bad new first, success last

Advice: tell the story of the data in a number of different ways

Some dashboards are temporary but some will be a constant reference: this has an impact on how the data should be presented.

MAINTAINING A CLEAR NARRATIVE

Horizontal logic:

- if your visualizations span many pages then the title of each page should tell you the story
- reinforce with an executive summary dashboard or report at the beginning

Vertical logic:

- one page or many, the content should reinforce the title and *vice versa* (self-reinforcement)
- there should be a logical link between all the elements, tags and visual aids on the page

VISUAL PROCESSING

Perception is fragmented – eyes are continuously scanning.

Visual thinking seeks patterns

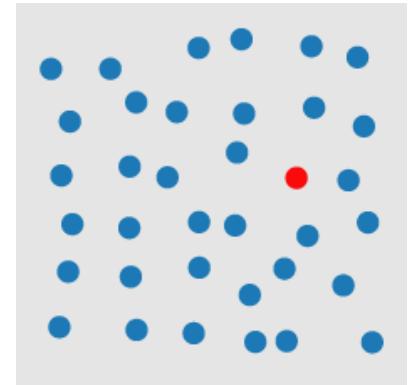
- **pre-attentive processes:** fast, instinctive, efficient, multitasking
gather information and build patterns:

features → patterns → objects

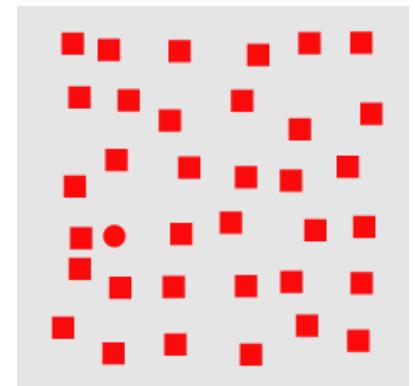
- **attentive process:** slow, deliberate, focused
discover features in the patterns:

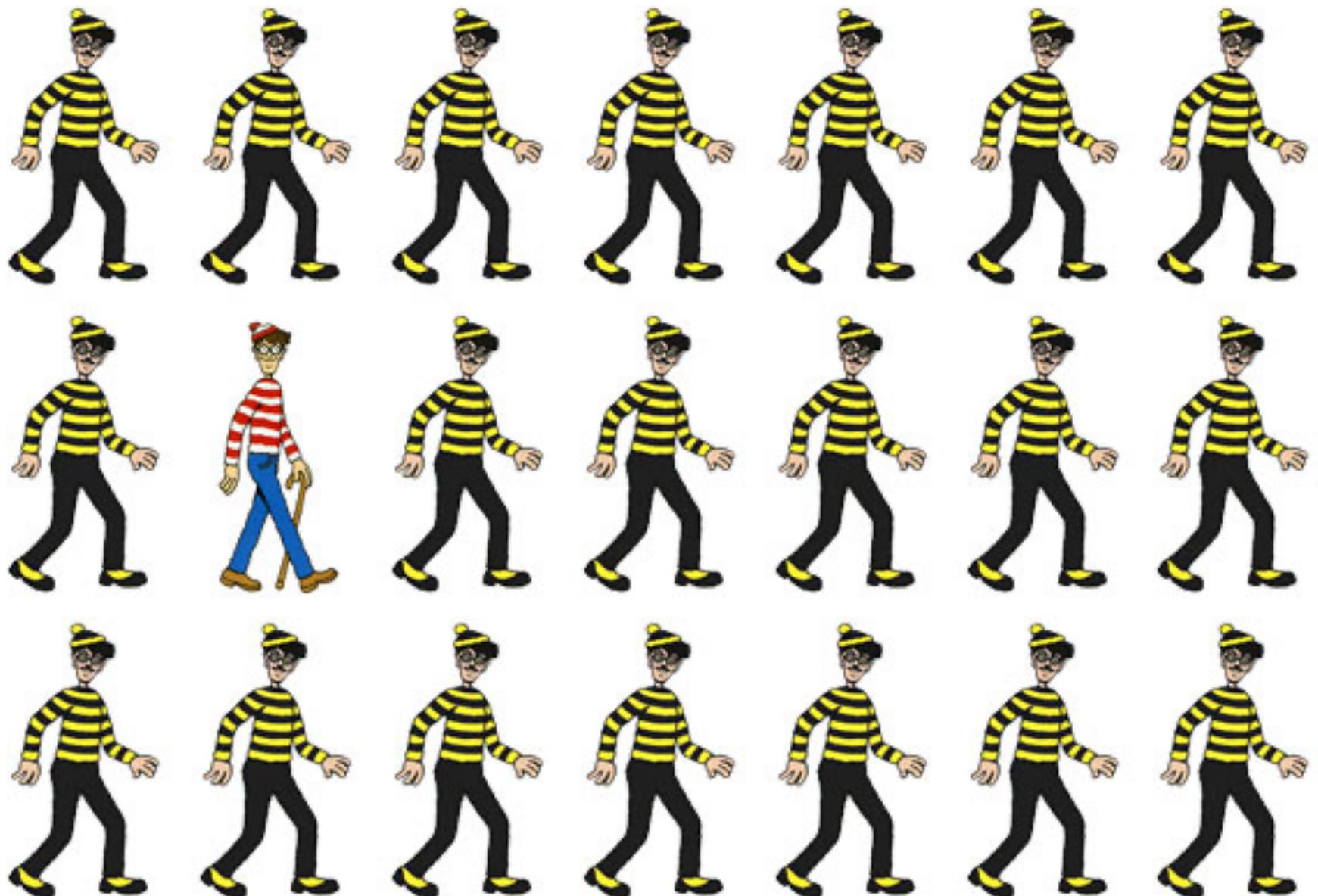
objects → patterns → features

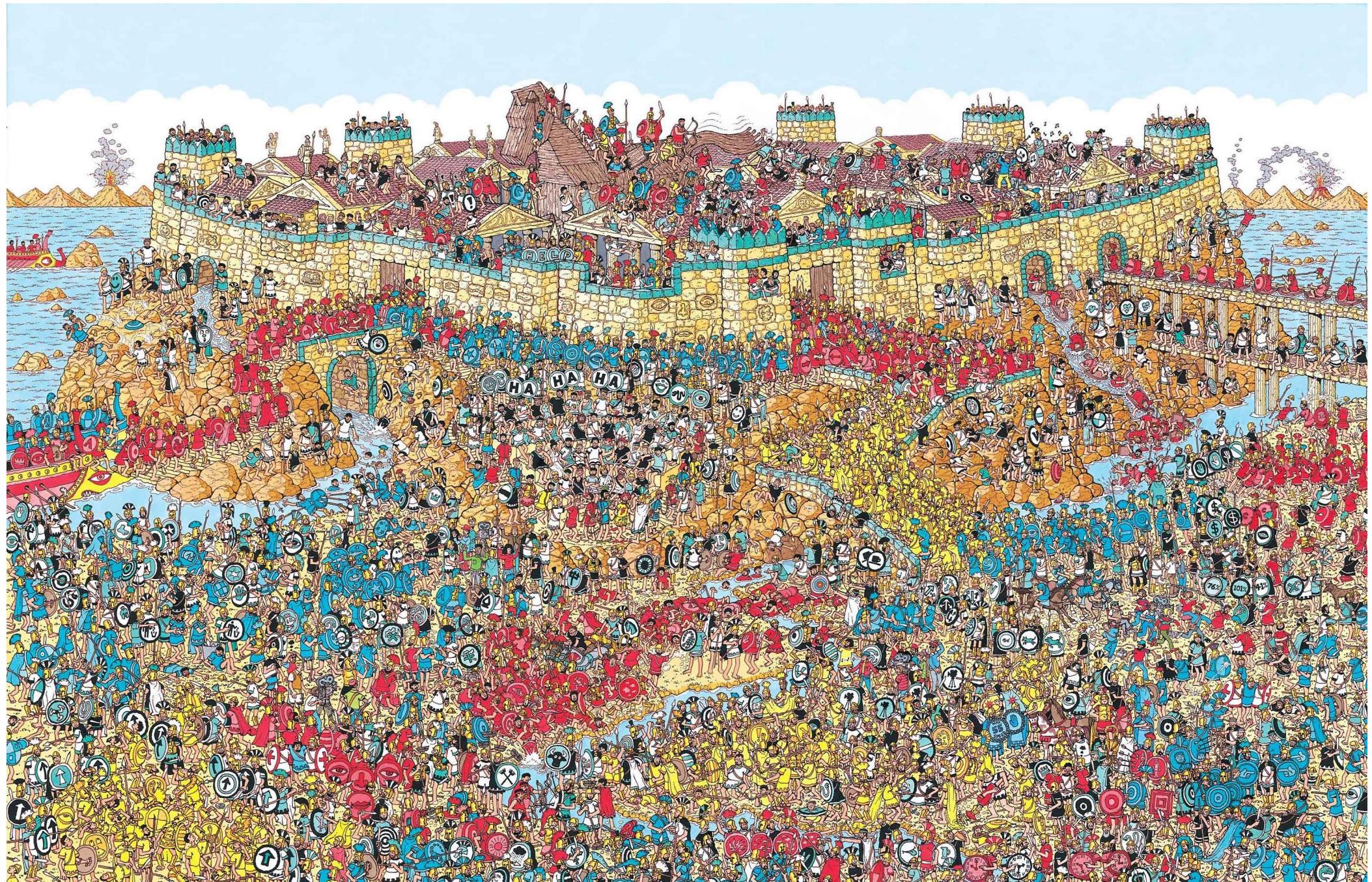
pre-attentive



attentive







ENGAGING MEMORY

Different types of memories are engaged when we tell stories:

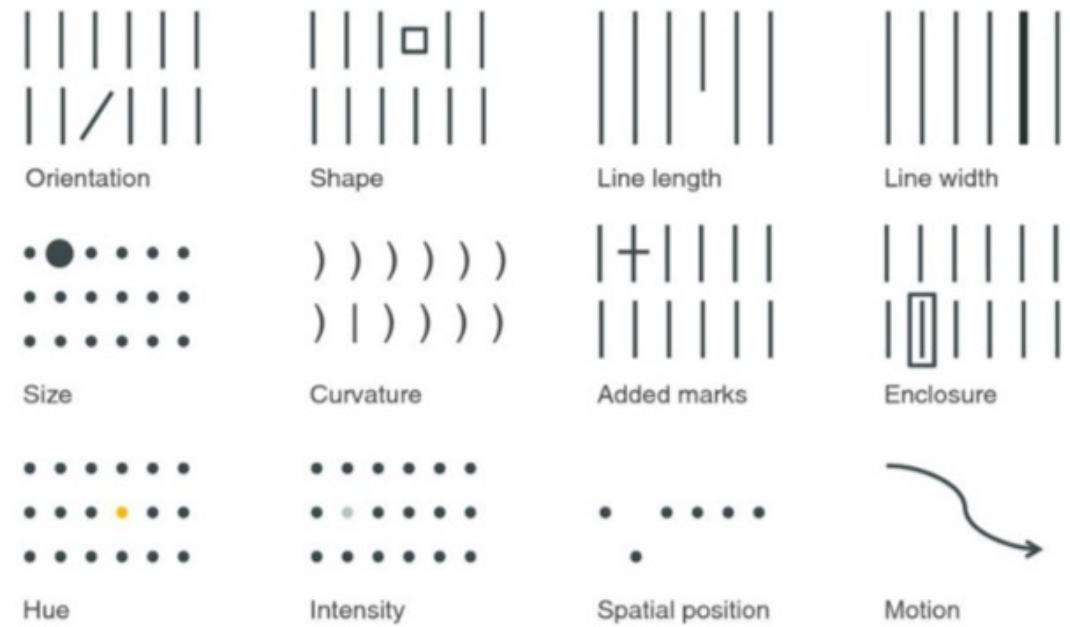
1. **iconic memory** – directs the eye
2. **short-term memory** – limits how many charts are found in dashboards
3. **long-term memory** – helps the audience remember what they saw

ENGAGING MEMORY

Iconic memory is the **visual sensory memory** (SM) register relating to the visual domain and a fast-decaying, high-capacity store of visual information.

Iconic memory is **very brief** (< 1000 ms) and provides a **coherent representation of our entire visual perception**.

Tuned to **pre-attentive attributes** (subconscious accumulation of information from the environment).



ENGAGING MEMORY

Short Term Memory can hold ~4 chunks of visual information in **short-term memory** at a given time.

When presented with more chunks (such as data points on a graph), chunks need to be **processed in and out of memory**.

Generally, we try to form **bigger, focused** hierarchies of chunks (Gestalt principles).

ENGAGING MEMORY

Long-term memory is built up over a lifetime and is the basis for pattern recognition and general cognitive processing.

It is an aggregate of **visual** memory and **verbal** memory.

Images help us recall long-term memory, making the story “**stick**”.

Context-providing text also makes a difference:

You have currently selected 28,711
ATIP requests totaling 6,597,612
pages of information

VS

ENGAGING MEMORY

Long-term memory is built up over a lifetime and is the basis for pattern recognition and general cognitive processing.

It is an aggregate of **visual** memory and **verbal** memory.

Images help us recall long-term memory, making the story “**stick**”.

Context-providing text also makes a difference:

vs

ATIP Requests		
30K requests	6.6M pages	230 pages/request

WEEKLY number of boats sold (20X6) – Store #16

2869408609876
9348586748676
2967303986739
3967496749674

Yearly goal: 290
20X6 total: 307

Do these numbers look
reasonable?

2869408609876
934**8586748676**
29673039**86739**
3967496749674

Another frequent
weekly number of
boats sold: 8

Occurred: **5 times**
immediately before a
6 (out of 7)
(surprising)

286940860987**6**
934858674867**6**
296730398**6**739
396749674967**4**

Most frequent weekly
number of boats sold:
6
(11 times)

Occurred: **randomly**
(as expected)

2869408609876
9348586748676
2967303986739
3967496749674

Another frequent
weekly number of
boats sold: 7

Occurred: **7 times**
immediately after a
6 (out of 8)
(surprising)

VERDICT: The two last charts suggest that the weekly sale numbers **are not random**, and that they may have been falsified. We recommend **performing an audit** of sales for store #16.

EXERCISE – NARRATIVE LOGIC

Taking the storyboard, highlights and chart choices from our previous exercises

1. Identify what type of narrative and logic do you think would best serve your needs?
 - Is your dashboard a single page or does it need to span many pages, if so, what is the horizontal logic behind the page layout?
 - Start to brainstorm what the vertical logic might be on one of your pages.

CHART AESTHETICS

PART II – EFFECTIVE STORYTELLING VISUALS

GESTALT PRINCIPLES

The **Gestalt principles** are the “laws” of human perception.

They describe how humans group similar elements, recognize patterns and simplify complex images when they perceive objects.

Designers use them to organize content on charts, dashboards, websites, and other interfaces so that they be **aesthetically pleasing** and **easy to understand**.

GESTALT PRINCIPLES

“Gestalt” is German for “unified whole”.

The first principles were devised in the 1920s by German psychologists Wertheimer, Koffka (“the whole is greater than the sum of the parts”) and Kohler.

Aim: understand how humans gain meaning from the chaotic stimuli around them.

The Gestalt principles are a set of “laws” which address the natural compulsion to find order in disorder. According to this, the mind “informs” what the eye sees by **perceiving a series of individual elements as a whole.**

GESTALT PRINCIPLES

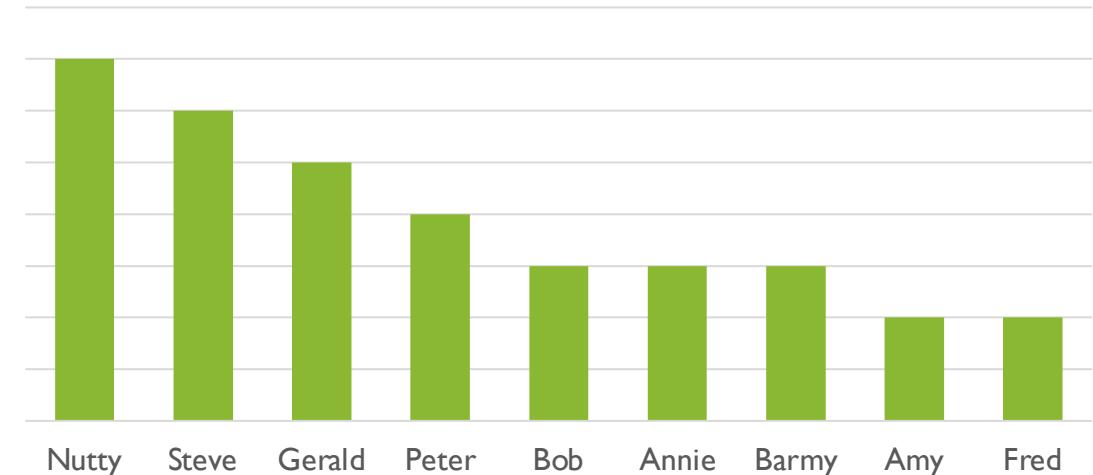
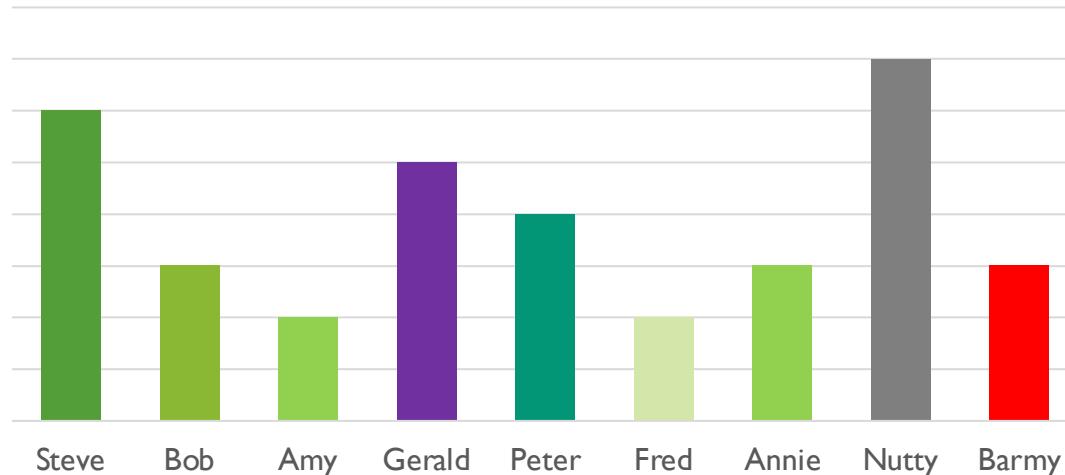
- simplicity
- continuation
- proximity
- similarity (invariance)
- focal point
- isomorphic correspondence
- figure / ground duality
- common fate*
- closure*
- uniform connectedness*



GESTALT PRINCIPLES – SIMPLICITY

The brain has a preference for **simplicity** – it tends to process simple patterns faster than patterns that are more complex.

Lesson: arrange data simply and logically wherever possible.

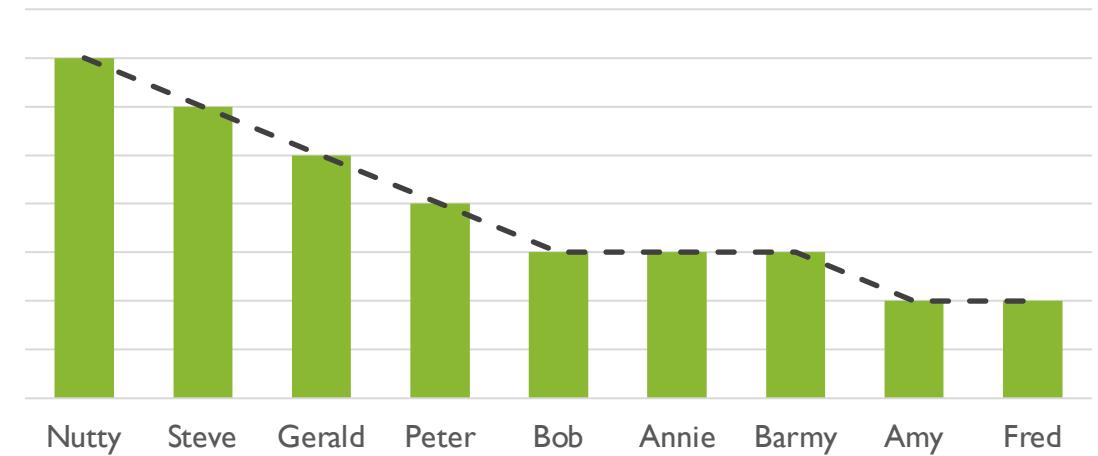
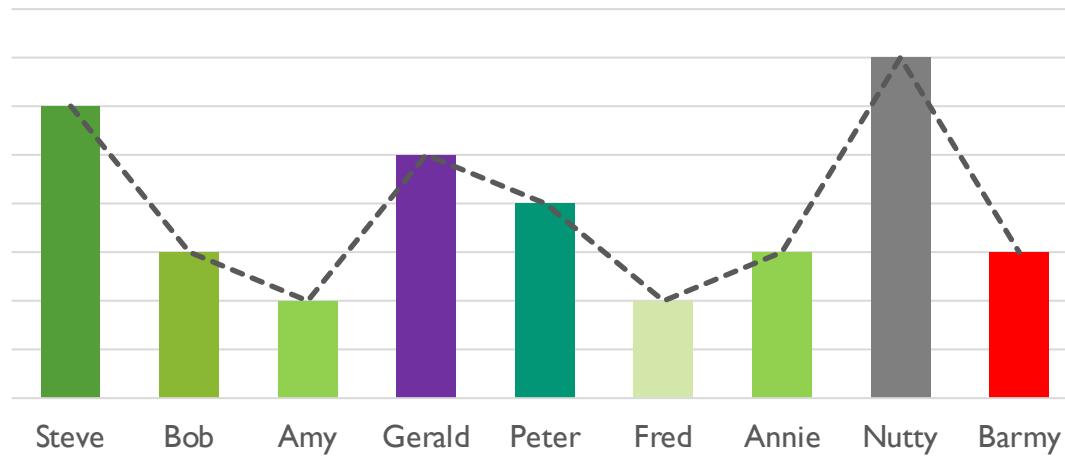


GESTALT PRINCIPLES – CONTINUATION

Our eyes group things that are **aligned** (e.g. sorted from high to low) with each other.

In the chart on the right the eyes follow a **continuous path**; it makes the whole chart more readable because of the continuous downward direction

Lesson: arrange objects in a line to facilitate grouping and comparison.



GESTALT PRINCIPLES – PROXIMITY

Objects/shapes that are **in proximity** (close) to one another appear to form **groups**.

The effect generated by the collected group is more “powerful” than that generated by separate elements.

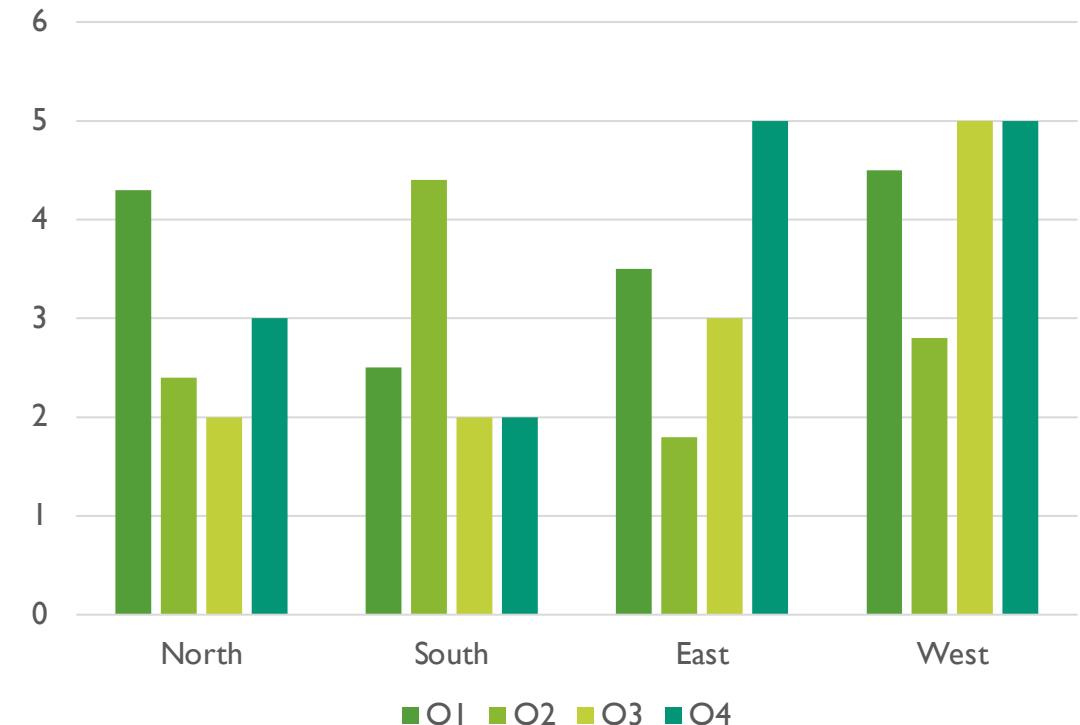
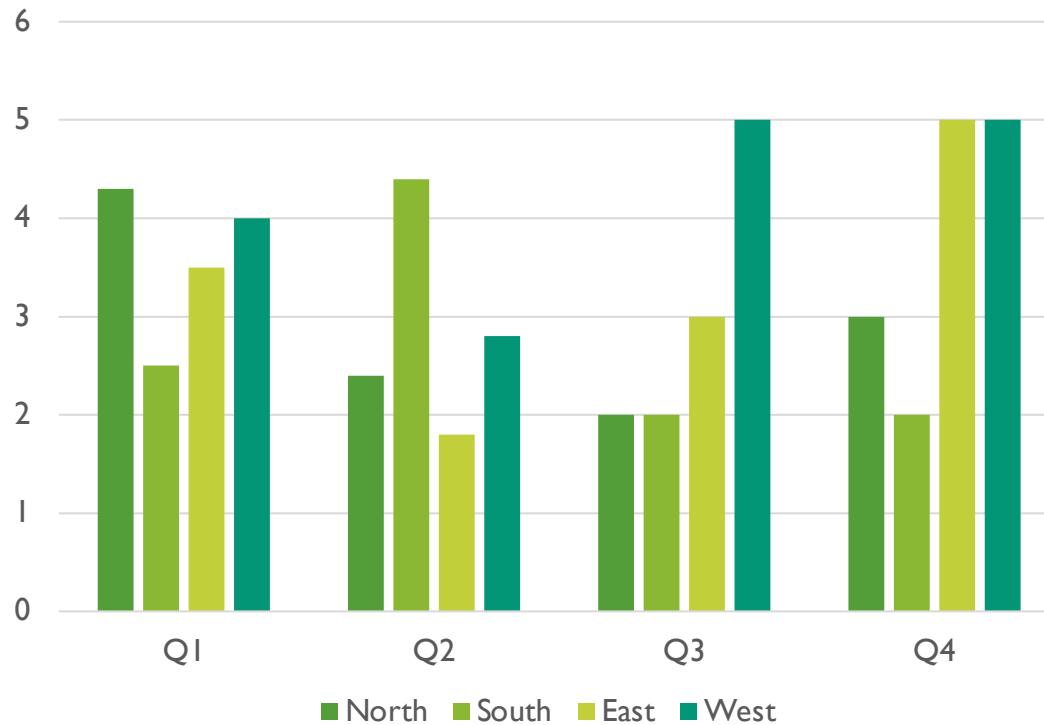
Elements which are grouped together create the **illusion** of shapes/planes in space, even if the elements are not touching.

Lesson: understand the chart’s priorities and create groupings through proximity that support those priorities.

GESTALT PRINCIPLES – PROXIMITY



GESTALT PRINCIPLES – PROXIMITY



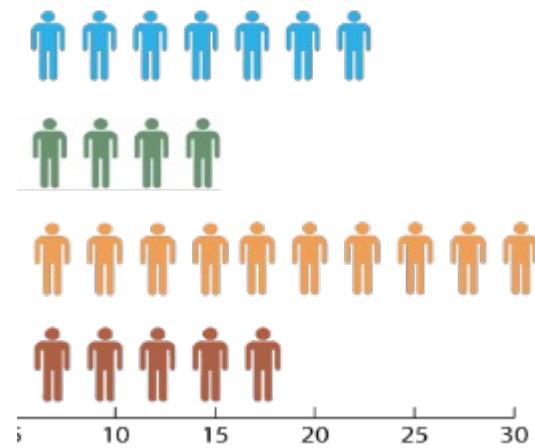
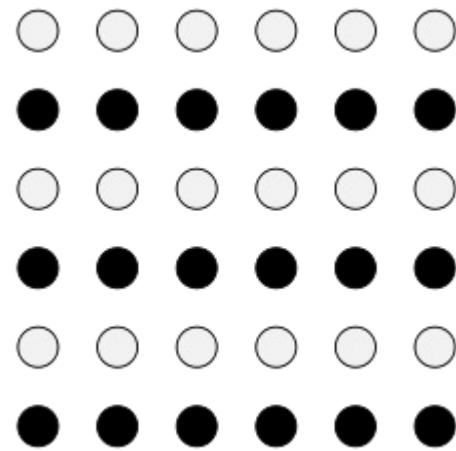
GESTALT PRINCIPLES – SIMILARITY (INVARIANCE)

Similarity: stimuli that physically resemble each other are viewed as part of the same object; stimuli that don't are viewed as part of a different object.

Similarity and proximity often come together to form a **Visual Hierarchy**. Either principle can dominate the other, depending on their application and combination.

Lesson: use similar characteristics to establish relationships and to encourage groupings of objects.

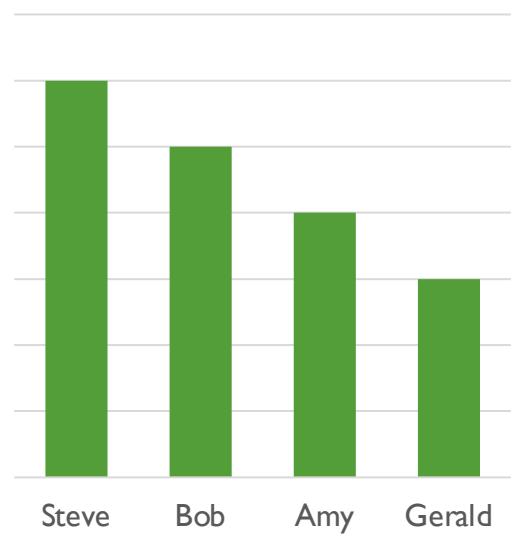
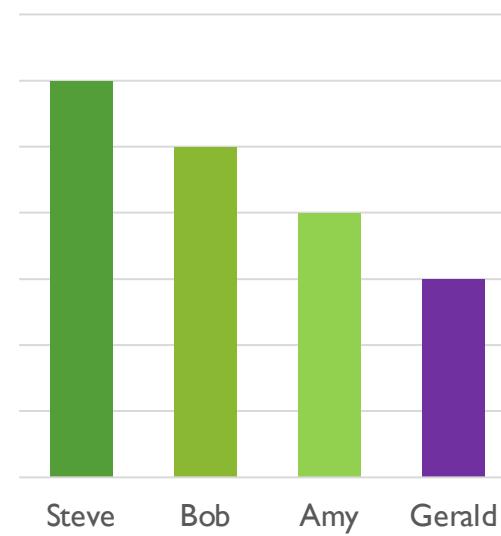
GESTALT PRINCIPLES – SIMILARITY (INVARIANCE)



In these examples, similarity dominates over proximity: we see rows before we see columns.

GESTALT PRINCIPLES – SIMILARITY (INVARIANCE)

Making things similar can reduce cognitive load (cf. last graph colour).

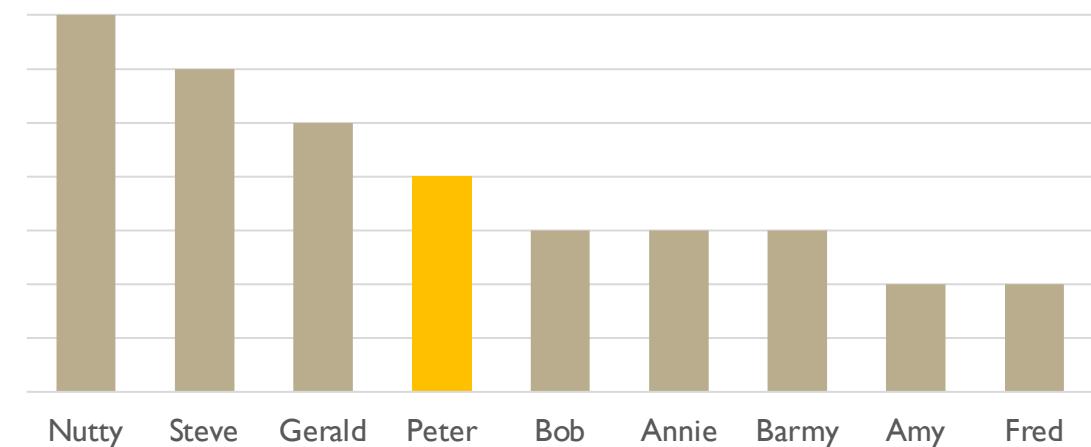
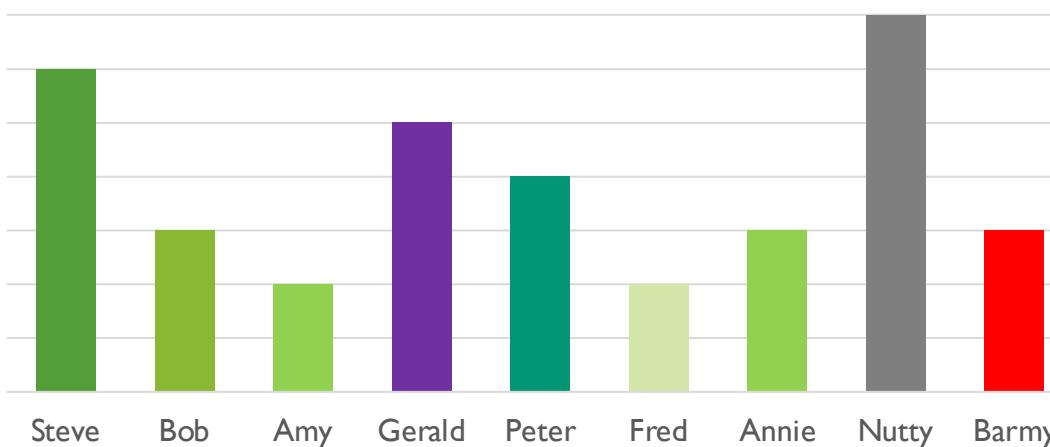


GESTALT PRINCIPLES – FOCAL POINT

In opposition to similarity, the **focal point** principle states that distinctive-looking objects can create a focal point.

To highlight one salesperson's performance, make their bar graph color different.

Lesson: use different characteristics to highlight and create focal points.

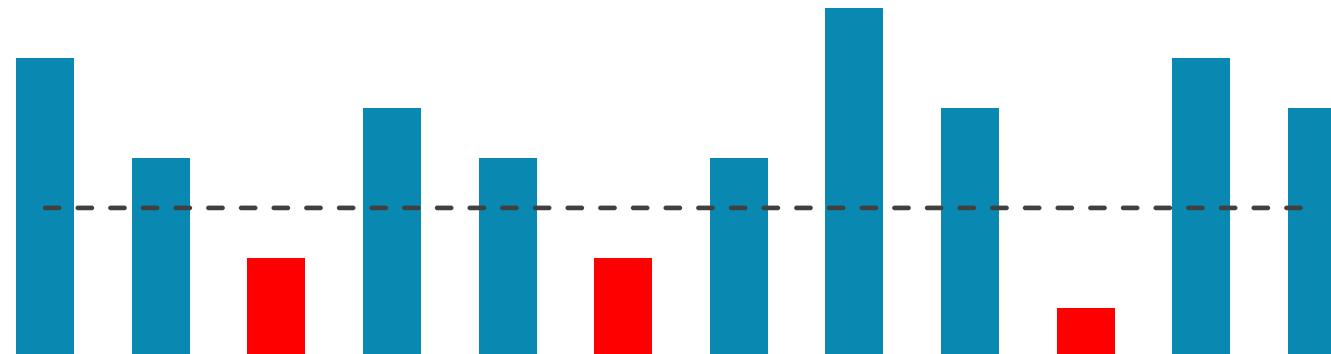


GESTALT PRINCIPLES – ISOMORPHIC CORRESPONDENCE

People interpret and respond to images based on past/shared experiences (in particular, for the selection of chart colours).

Red is often associated with **bad** and **green** with **good** (colour-blindness?). We can colour-code charts accordingly.

Lesson: stick to well-established conventions and best practices (even if boring!)



GESTALT PRINCIPLES – FIGURE/GROUND DUALITY

Chart elements are either perceived as **figures** (focus) or as (back)**ground**.

Foreground objects are **promoted** by the brain, background objects are **demoted**.

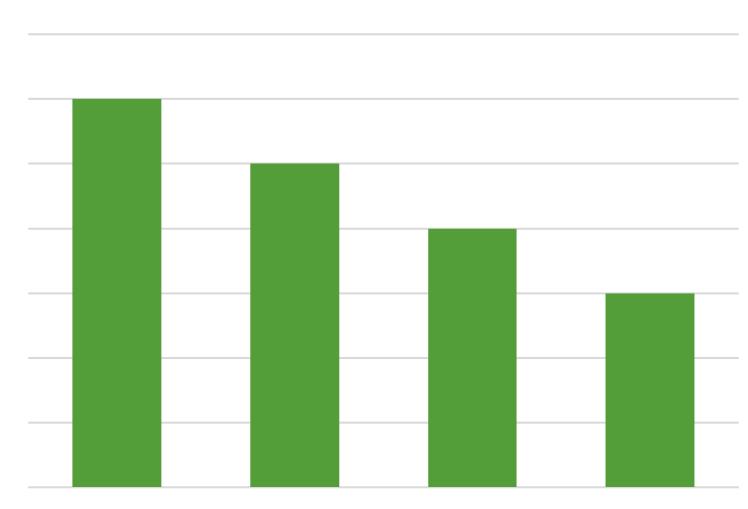
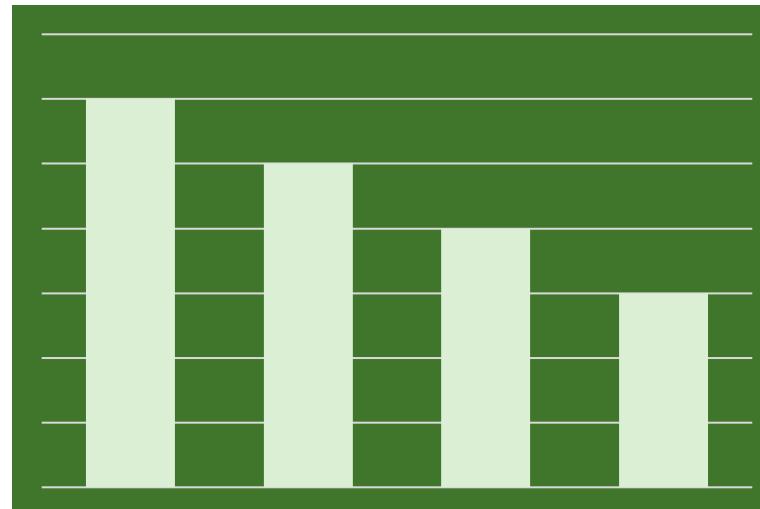
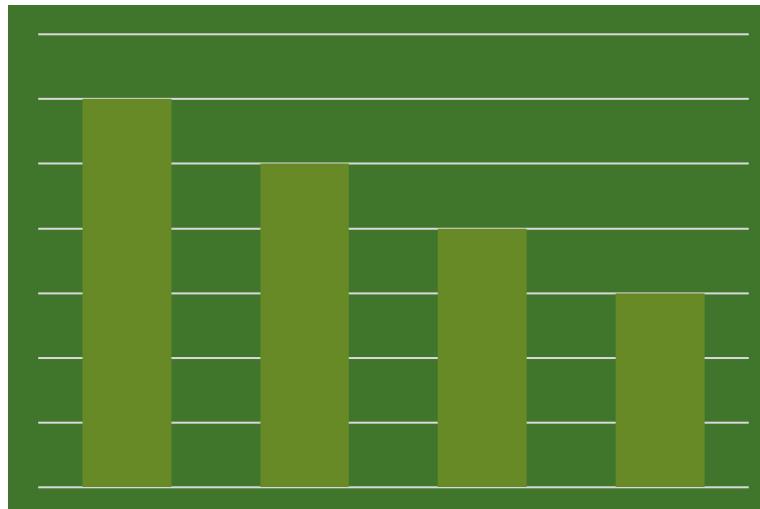
Strong contrast makes it easier to distinguish between the two types of objects.

Lesson: ensure there is enough contrast between the chart foreground (figures) and their background.

GESTALT PRINCIPLES – FIGURE/GROUND DUALITY

Because of the low contrast between the figure and background in the chart on the left, there is an **additional cognitive load**.

Increasing the contrast on the right improves readability.



EXERCISE – HIGHLIGHTING THE CRITICAL ITEMS

Taking the previous exercise:

1. Identify pre-attentive attributes that might be helpful in telling the story.
2. Identify Gestalt principles that might be helpful in telling the story.

DECLUTTERING

CLUTTER IS THE ENEMY!

Every element on a page adds **cognitive load**

- identify and **remove** anything that isn't adding value
- think of cognitive load as mental effort required to process information (lower is better)

Tufte refers to the **data to ink ratio** – “the larger the share of a graphic’s ink devoted to data, the better”

In *Resonate*, Duarte refers to this as “**maximizing the signal-to-noise ratio**” where the signal is the information or the story we want to communicate.

DECLUTTERING

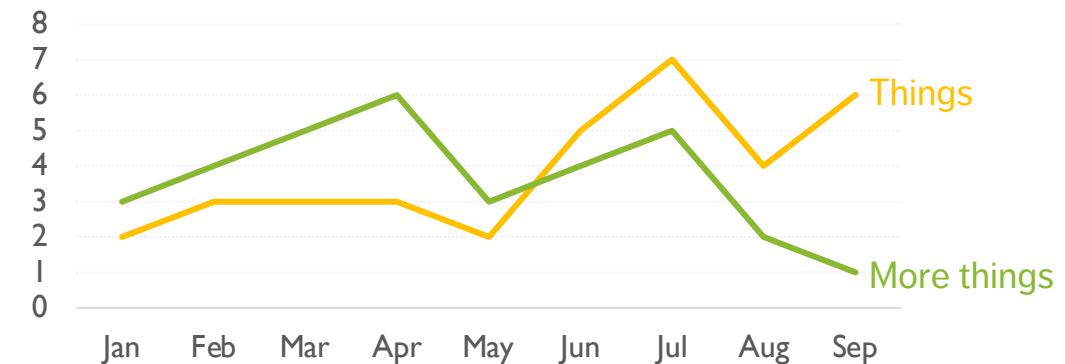
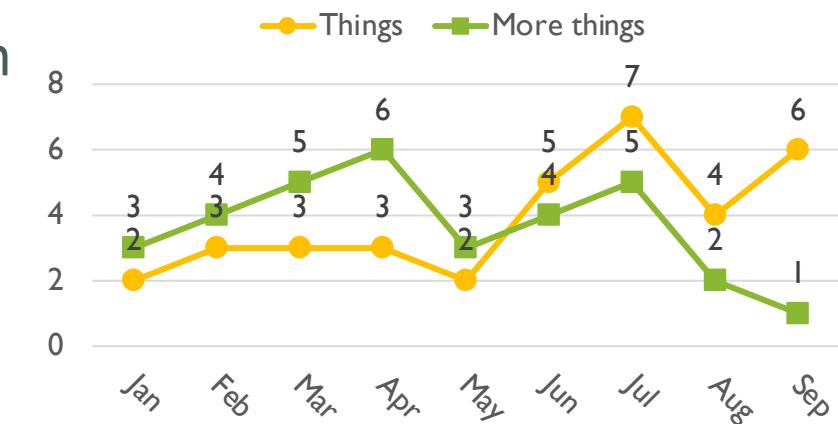
Use the **Gestalt Principles** to organize/highlight data in the chart.

Align all the elements (graphs, text, lines, titles, etc).

- DON'T rely on eye, use position boxes and values

Charts:

- remove border, gridlines, data markers
- clean up axis labels
- label data directly



DECLUTTERING

Use **consistent** font, font size, colour and alignment.

Don't rotate text to anything other than 0 or 90 degrees.

Use **white space**:

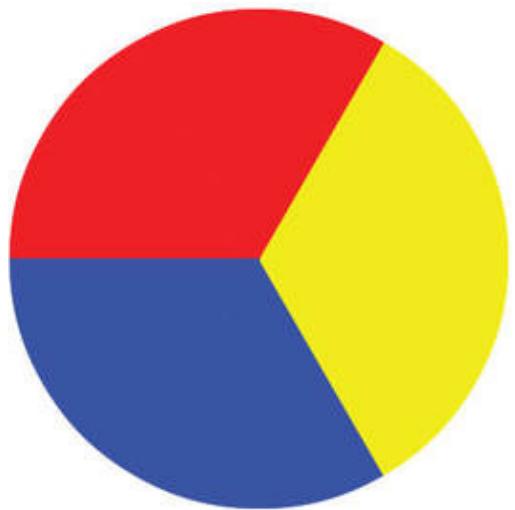
- margins should remain free of text and visuals
- don't stretch visuals to edge of page or too close to other visuals
- think of white space as a border

COLOUR THEORY

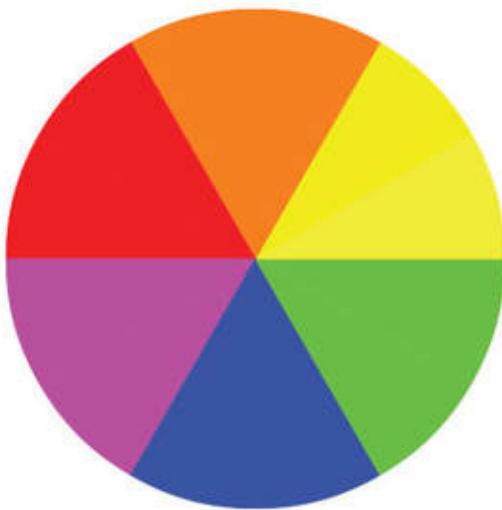
Colour theory (complicated topic – here is a start):

- <http://www.deanenettles.com/webexamples/colorexamples/>
- <https://www.sessions.edu/color-calculator/>

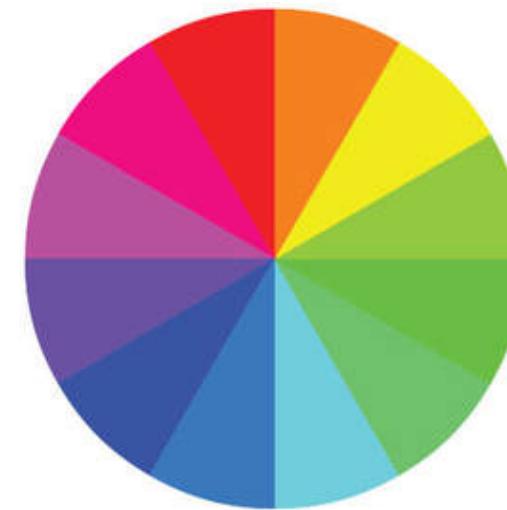
Colour wheels:



Primary Colours



Secondary Colours



Tertiary Colours

COLOUR SCHEMES

Achromatic (colourless, using only blacks, whites and grays)



Monochromatic (1-colour schemes)



Complementary (colours directly across from each other on the colour wheel)



Split complementary (2 of the 3 colors are adjacent; 1 of the colours is opposite)



COLOUR TIPS

When it comes to colour, **less is more**: use it sparingly (graphic designers are taught to “get it right, in black and white”).

Based on the Gestalt Principles, **monochrome** schemes can be particularly effective.

When appropriate, pick scheme based on corporate identity (this maximizes buy in).

Create a template (and stick to it).

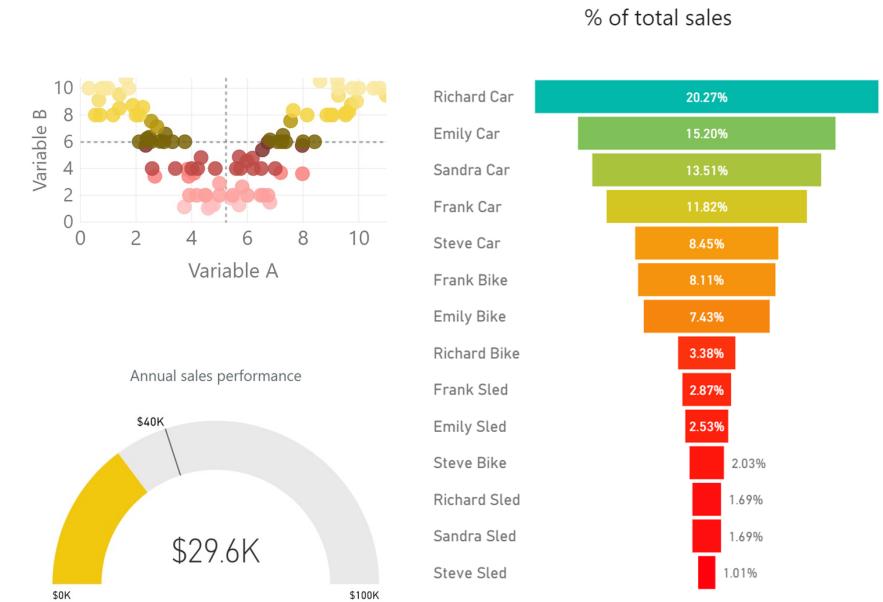
Upload images to see what charts look like in various flavours of colour-blindness:

- <https://www.color-blindness.com/coblis-color-blindness-simulator> (there are other tools)

SIZE CONSIDERATIONS

Size: assuming that the chart has been decluttered

- things of equal importance size similarly
- other things scale to importance



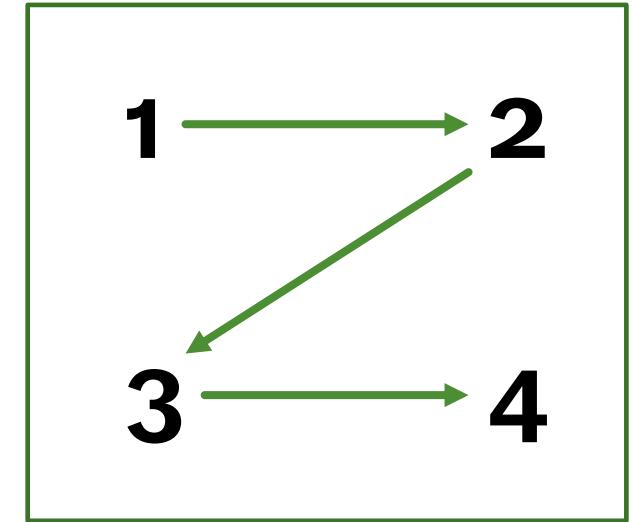
POSITION CONSIDERATIONS

How should the elements be placed in a chart/dashboard?

In the West, most people start at the top left and zig-zag all the way to the bottom right.

Simple rule: don't make people work too hard

- main message: top left/top right
- info in order of preference
- people concentrate less as they scan so get less complex as you move to bottom corner



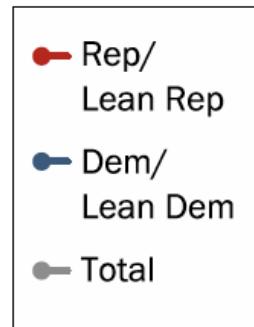
EXERCISES

Comment on the aesthetics of the following charts, according to:

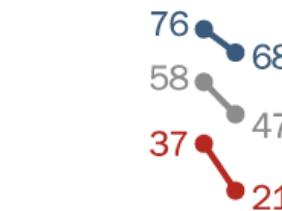
- Gestalt principles
- use of colours
- lack of clutter
- size and position
- etc.

Republican concern about the budget deficit increases sharply; Democratic concern declines

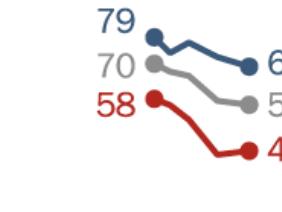
% who say ____ is a *very big problem* in the country today



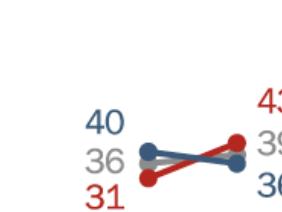
The coronavirus
outbreak



The affordability
of health care



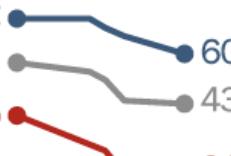
Quality of
public schools



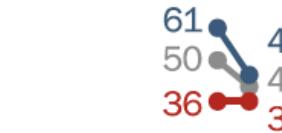
The federal
budget deficit



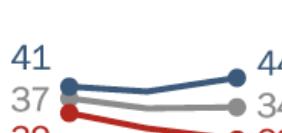
Economic
inequality



Unemployment



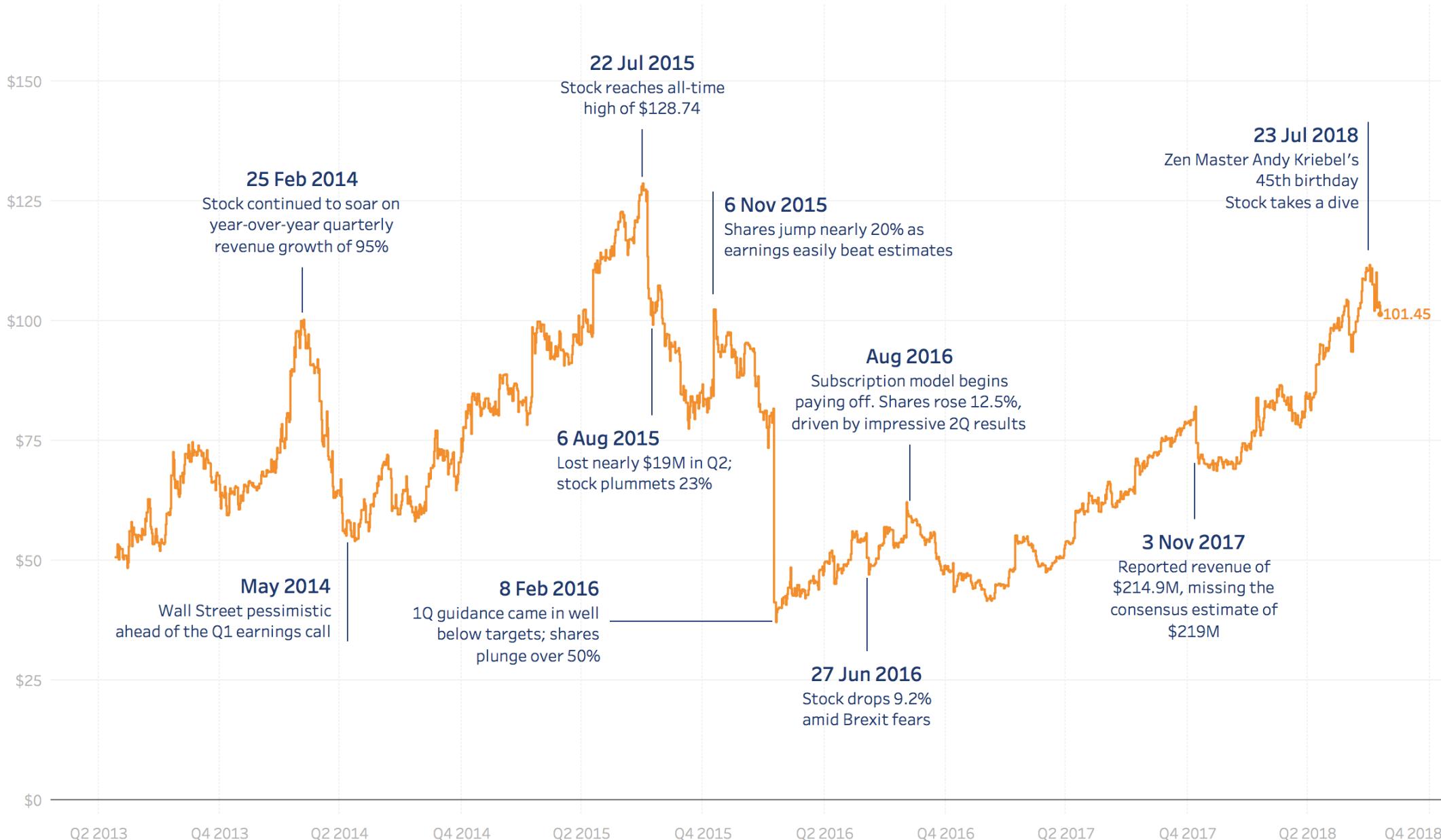
Condition of
infrastructure



Note: March 2019 and earlier wording for economic inequality was "The gap between the rich and poor." See topline for details.

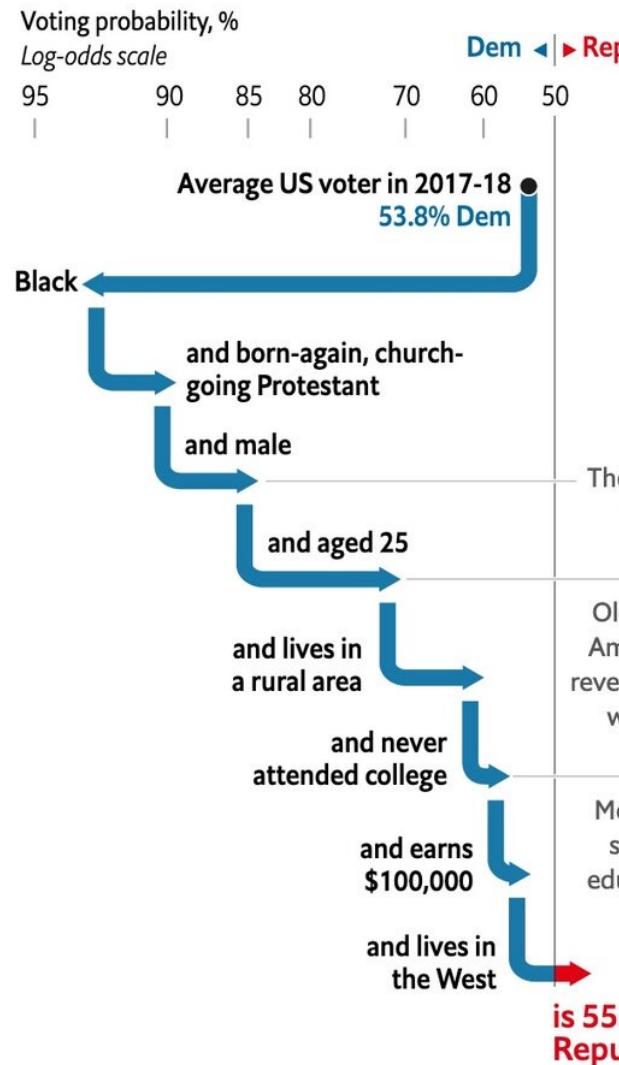
Source: Survey of U.S. adults conducted April 5-11, 2021.

The Roller Coaster Ride of Tableau's Stock

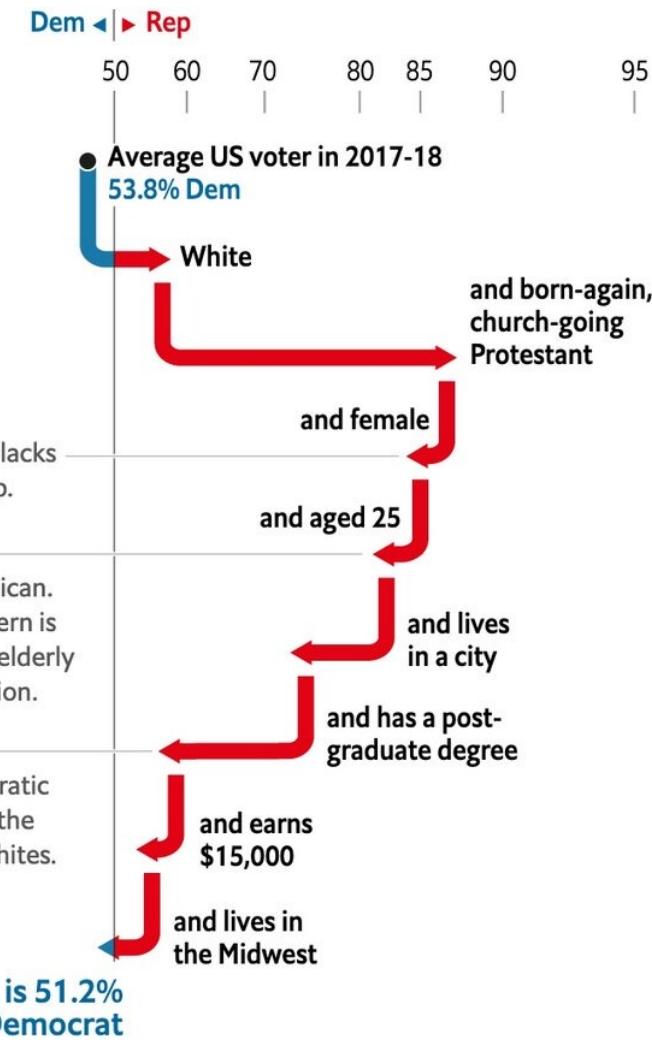


What would it take...

... to make a black voter
a likely **Republican**?

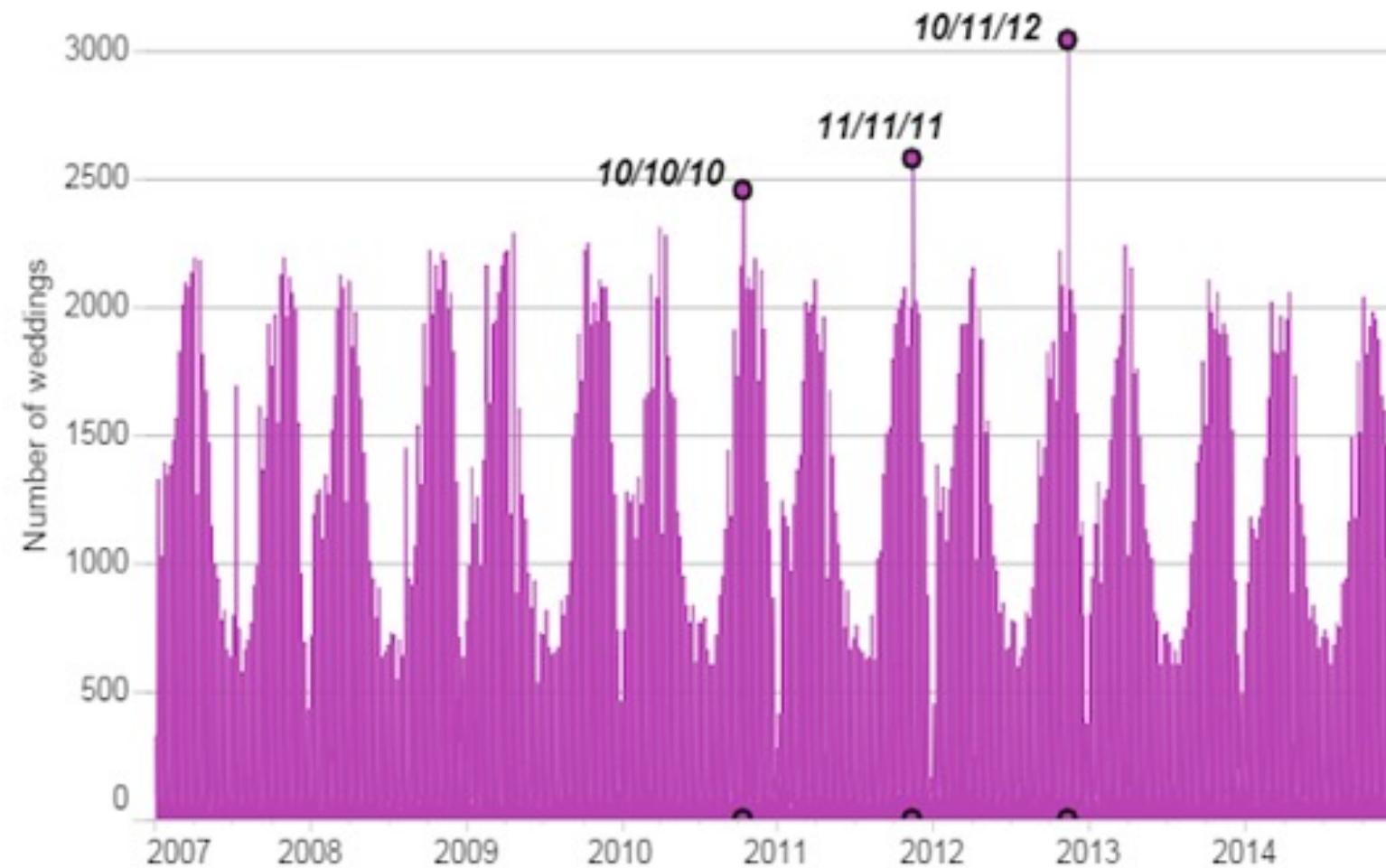


...to make a white born-again Protestant a likely **Democrat**?



Weddings in Australia

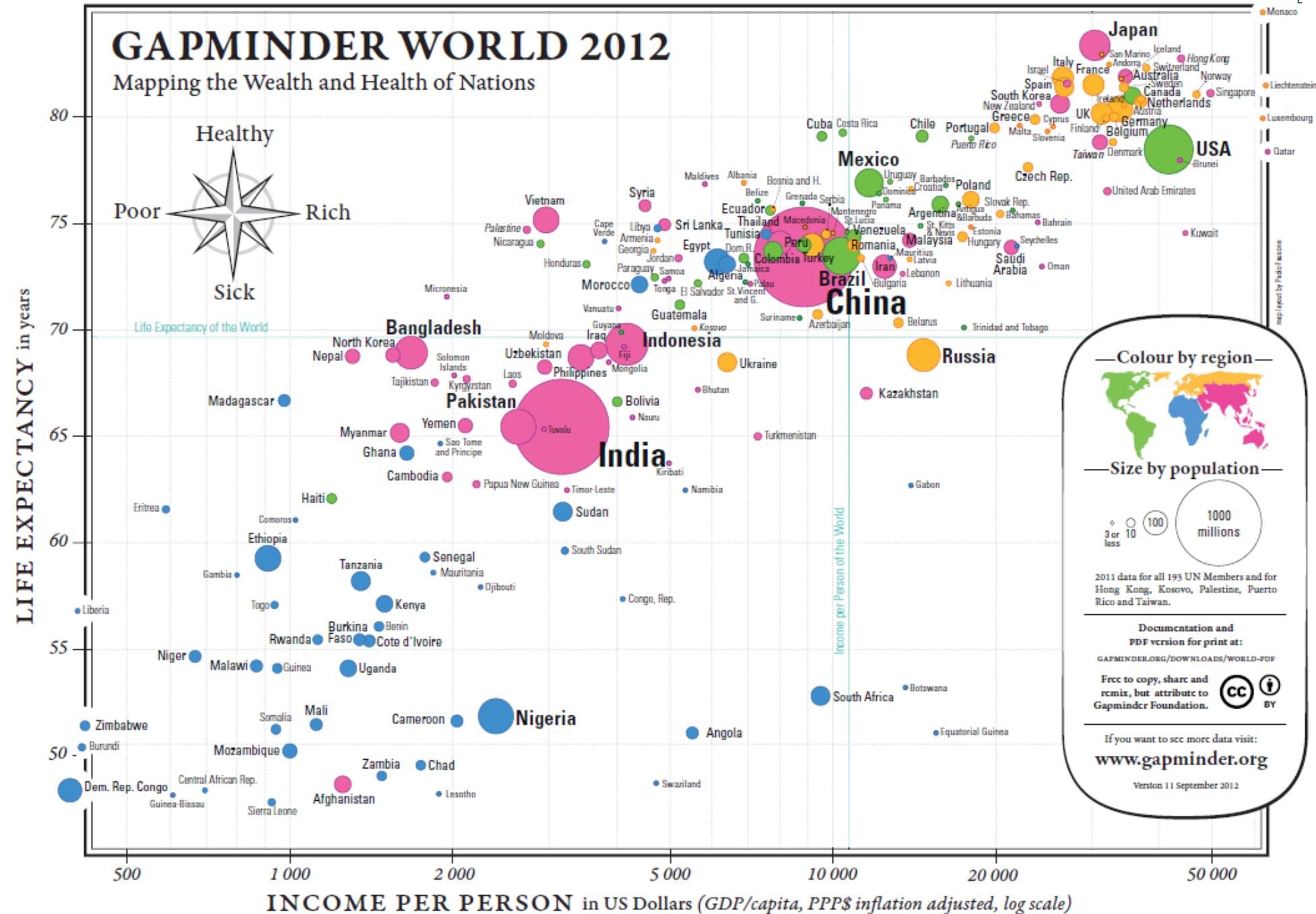
The most popular wedding dates form repeating or sequential number patterns.



Graphic: Inga Ting | Source: ABS 2015

GAPMINDER WORLD 2012

Mapping the Wealth and Health of Nations



ACCESSIBILITY

PART II – EFFECTIVE STORYTELLING VISUALS

A WORD ABOUT ACCESSIBILITY

A table can be translated to Braille, but that's not always possible for charts.

Describing the features and emerging structures in a visualization is a possible solution... **if they can be spotted.**

Analysts must produce clear and meaningful visualizations, but they must also describe them and their features in a fashion that allows all to "see" the insights.

This requires them to have "seen" all the insights, which is not always necessarily the case (if at all possible).

A WORD ABOUT ACCESSIBILITY

Data Perception:

- texture-based representations
- text-to-speech
- sound/music
- odor-based or taste-based representations (?!?)

Sonifications:

- [TRAPPIST Sounds : TRAPPIST-1 Planetary System Translated Directly Into Music](#)
- [Listening to data from the Large Hadron Collider, L. Asquith](#)

A WORD ABOUT ACCESSIBILITY



Frank ✎ @FrankElavsky · 1h

•••

Resources we could use more of (1/?):

Low vision (~30% of all people):

- High contrast text
- High contrast elements
- Using texture, shape, units
- Designing with zoom/magnification
- Using Hierarchy and Focus
- Using annotations or guides

1

7

42



A WORD ABOUT ACCESSIBILITY



Frank ✎ @FrankElavsky · 1h

...

Resources we could use more of (2/?):

Functional/motor impairment (~13% of all people in US):

- Keyboard interactivity/navigation
- UI alternatives to in-chart controls (brushing, subselecting, etc)
- Alternative data navigation schemes
- Scrollytelling alternates

1

5

30



A WORD ABOUT ACCESSIBILITY



Frank ↗ @FrankElavsky · 1h

•••

Resources we could use more of (3/?):

Cognitive disability (~11% of all people in US):

- Captions, summaries, clear titles, and plain text alternatives
- Reducing visual complexity
- Forgivable user interactions
- Use of hierarchy
- Assistive design (how-to-read guides, help)

1

6

31



A WORD ABOUT ACCESSIBILITY



Frank ↗ @FrankElavsky · 1h

...

Resources we could use more of (4/?):

Attention deficit/hyperactive disorder (~9% of all people in US):

- Clear, short text summaries
- Object constancy
- Motion design and animation
- Use of breadcrumbs
- Interaction history (with undo/redo functions)

2

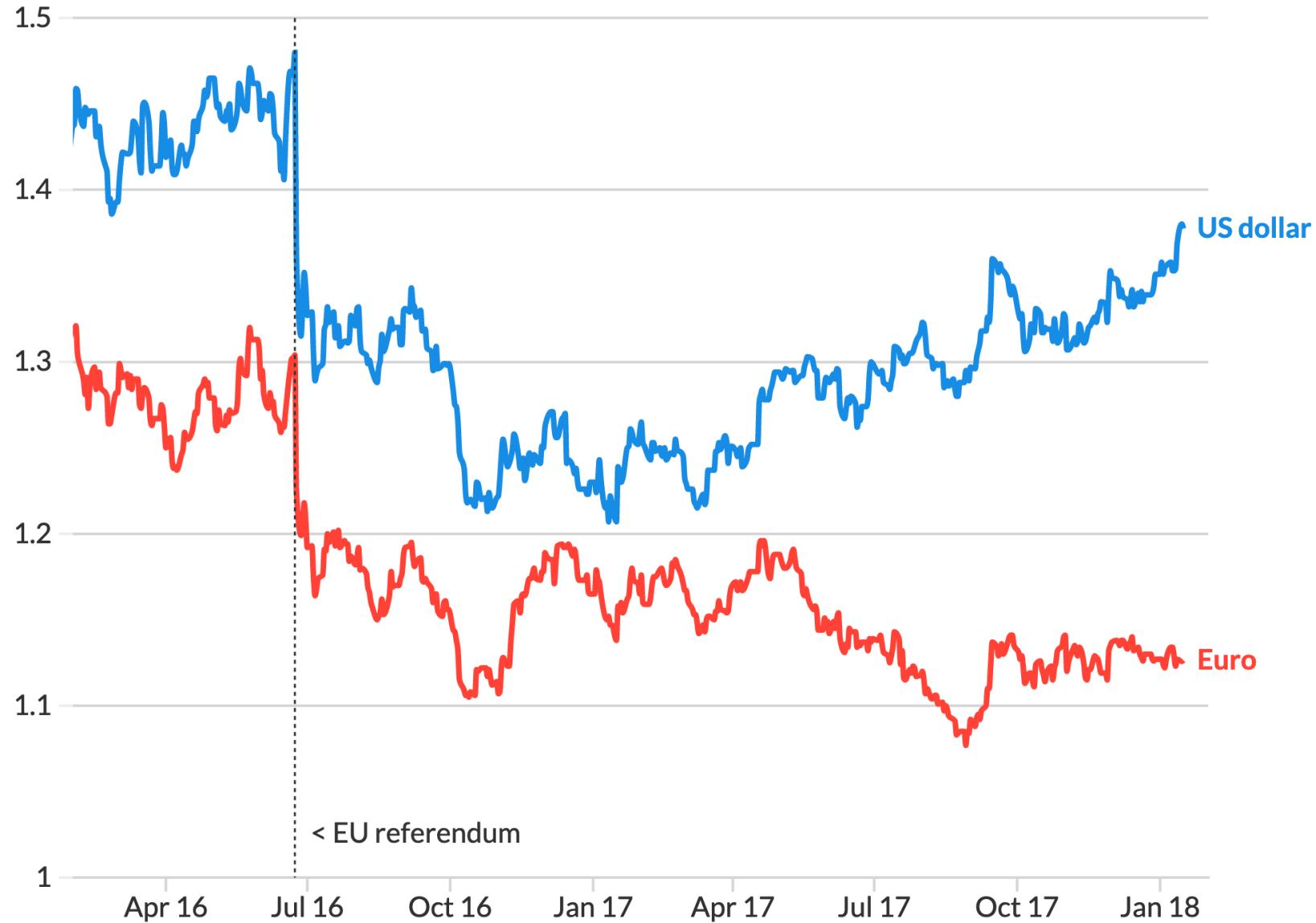
4

21



The value of the pound has fallen, particularly since the EU referendum

Euros and US \$ per £



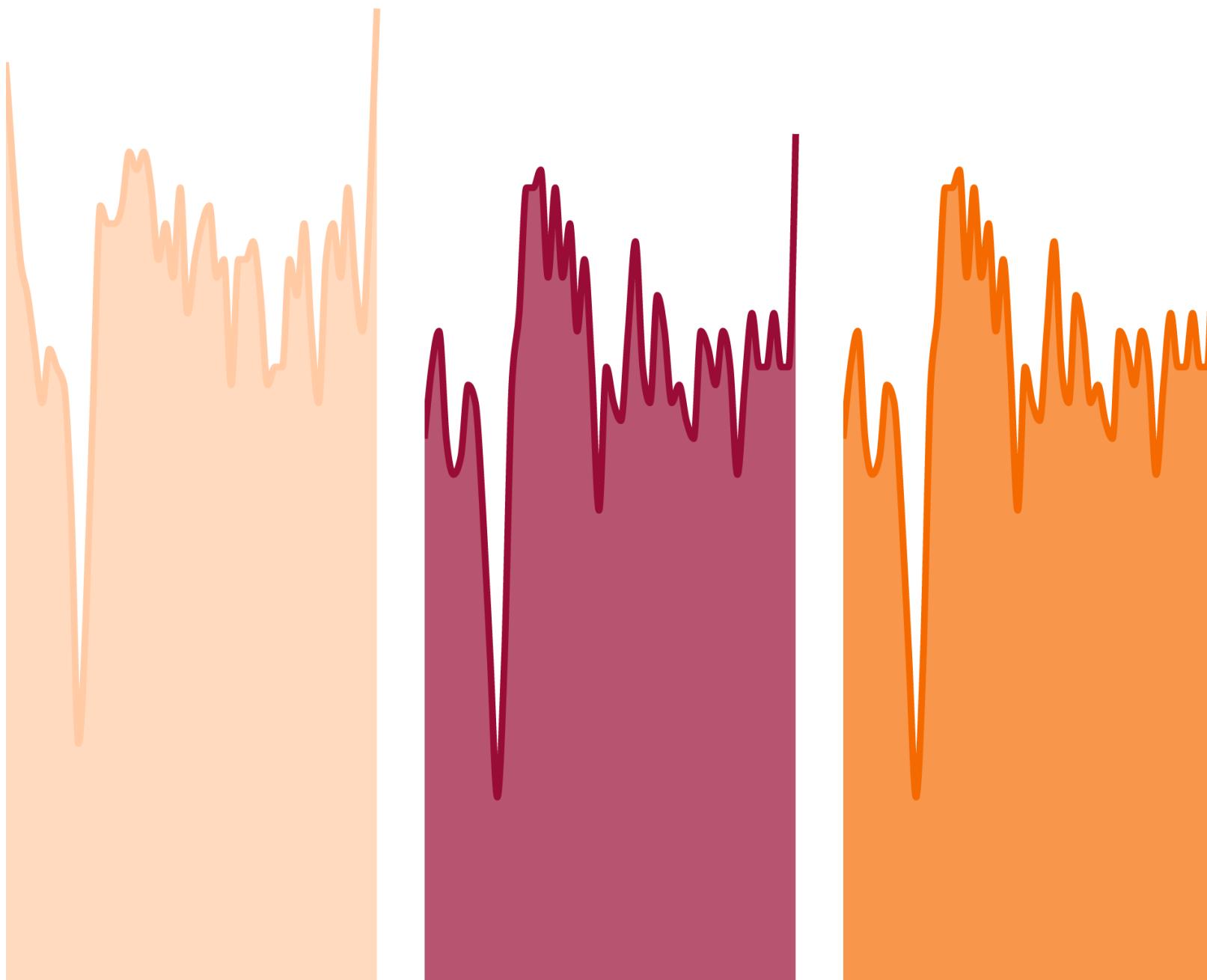
A line chart showing the value of the pound in relation to euros and US dollars. A large drop is visible after the EU referendum in June 2016. Just before the referendum you could get 1.48 US dollars and 1.3 euros for each pound. After the referendum it fell to 1.29 US dollars and 1.16 euros – a fall of around 12%.

Fail (1.46:1)

Pass (8.52:1)

Partial pass (3.02:1)

[<https://flourish.studio/blog/accessible-chart-design>]



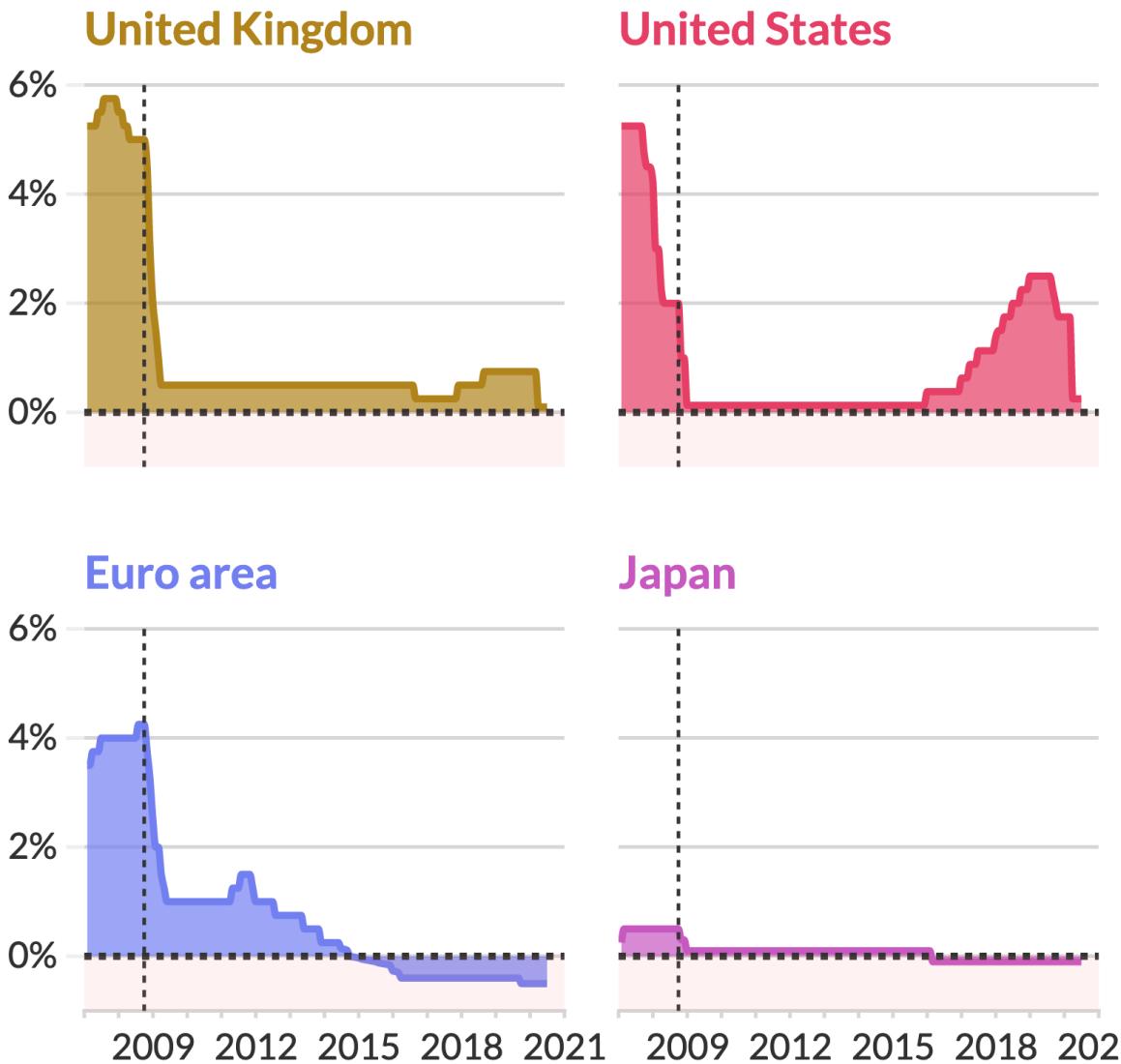
Use colors that are bold and clear enough for people to see both text and graphical elements.

[Web Content Accessibility Guidelines \(WCAG\)](#) suggest meeting the WCAG AA requirements.

To check if your color (and font size) choices are AA accessible you can use a [contrast checker website](#).

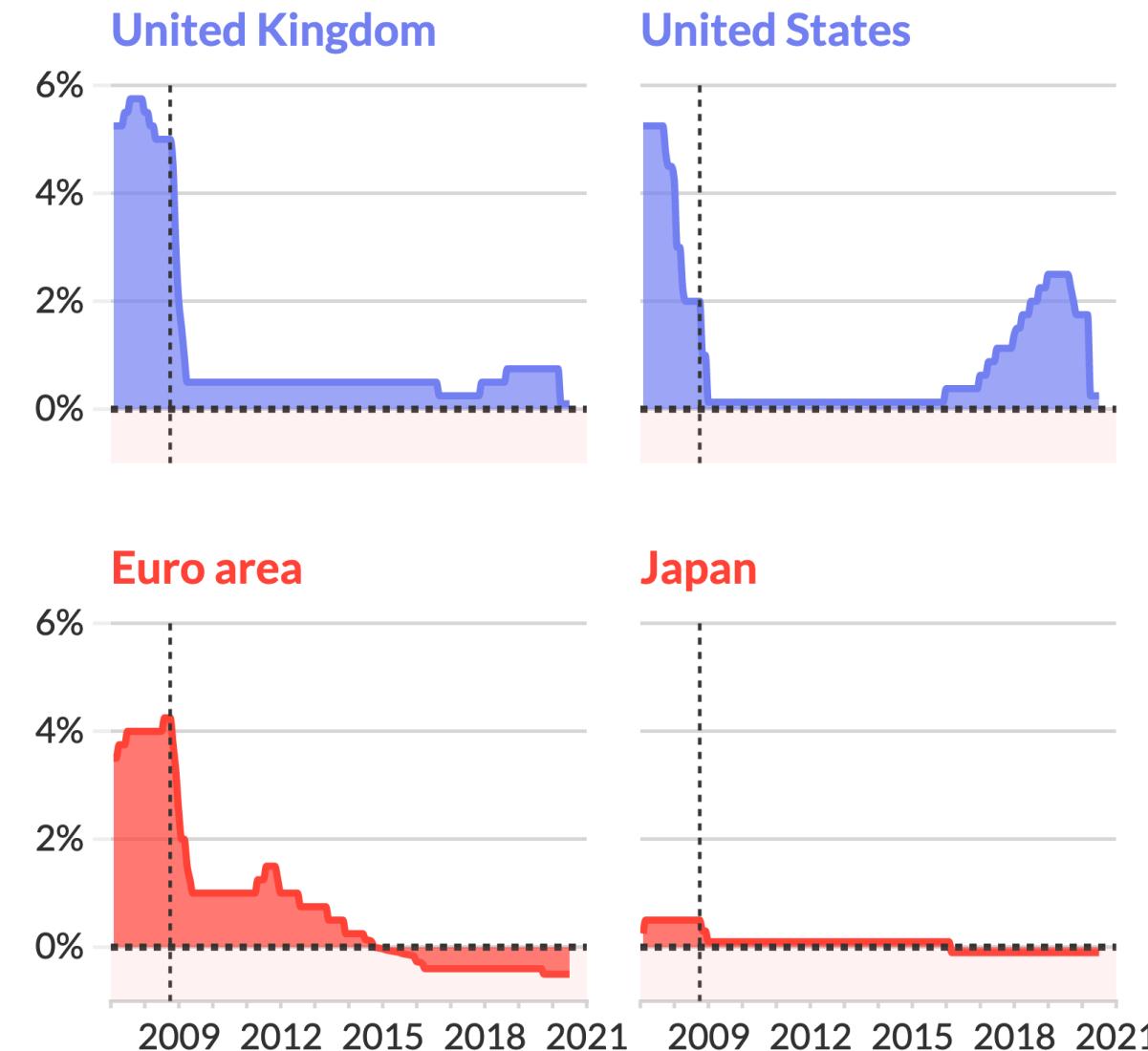
For colours to be AA accessible they need to have a contrast ratio of at least **3:1** for **graphical elements**, and **4.5:1** for **normal text**.

Interest rates have been falling since the financial crisis, and have even gone negative in some countries



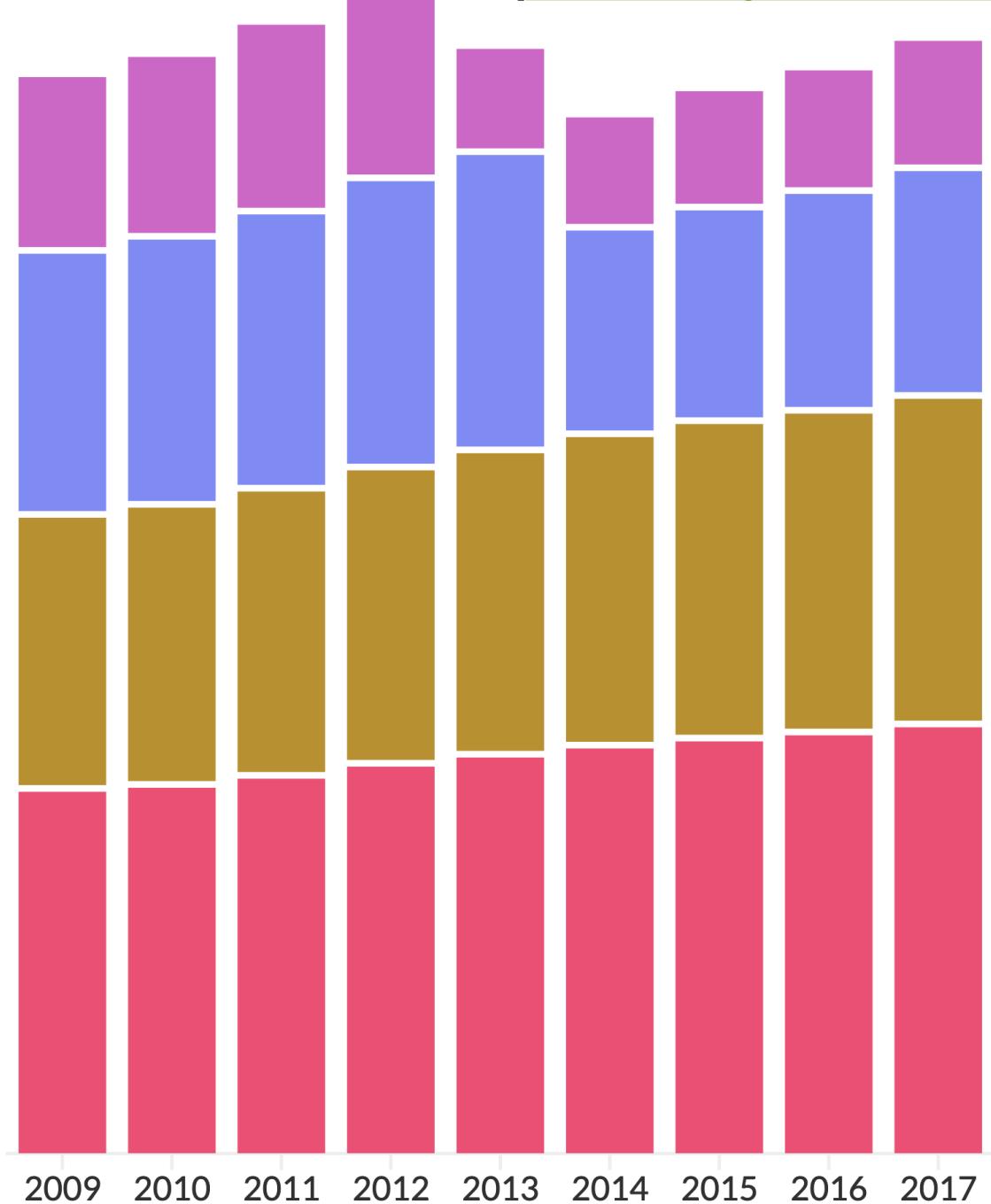
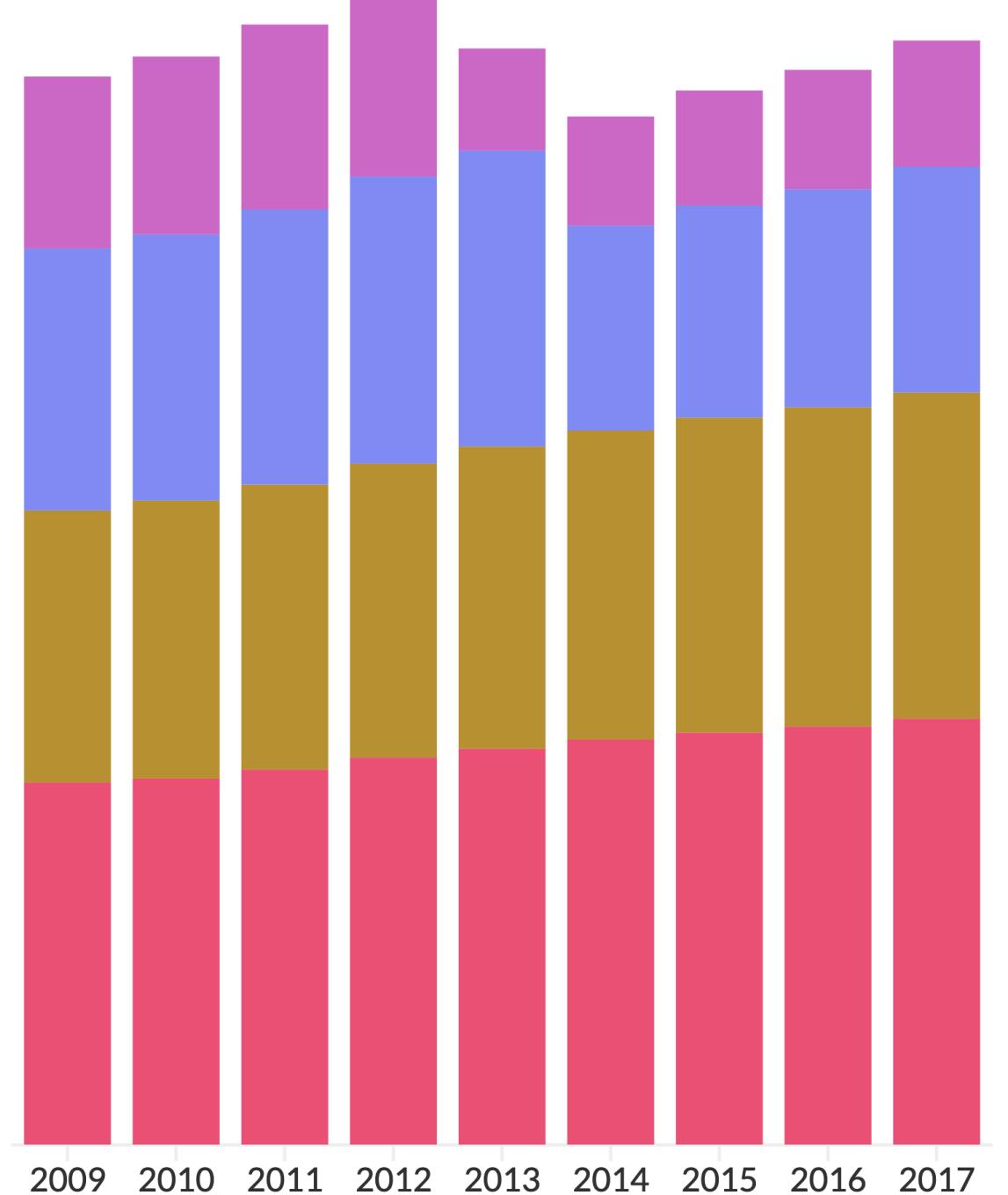
Source: [ONS](#)

Interest rates have been falling since the financial crisis, and have even gone **negative** in some countries



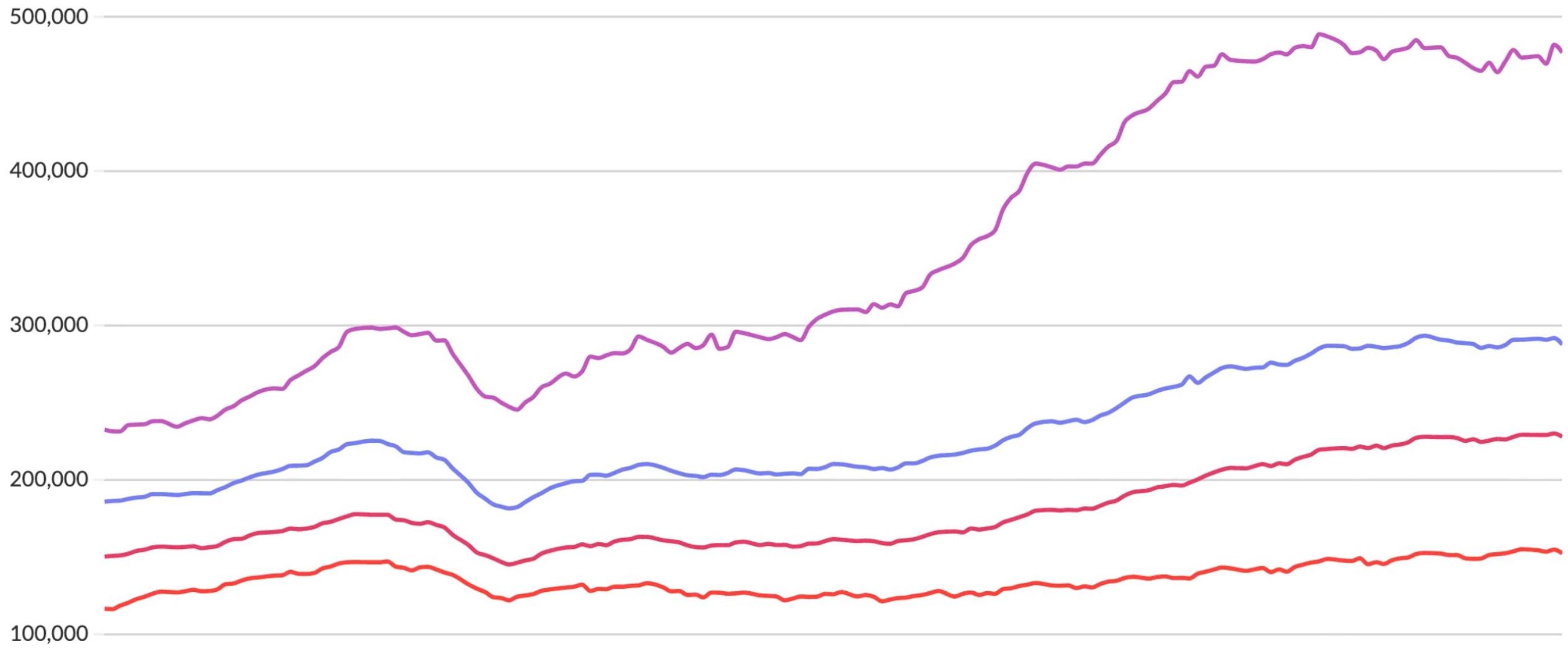
Source: [ONS](#)

[<https://flourish.studio/blog/accessible-chart-design>]



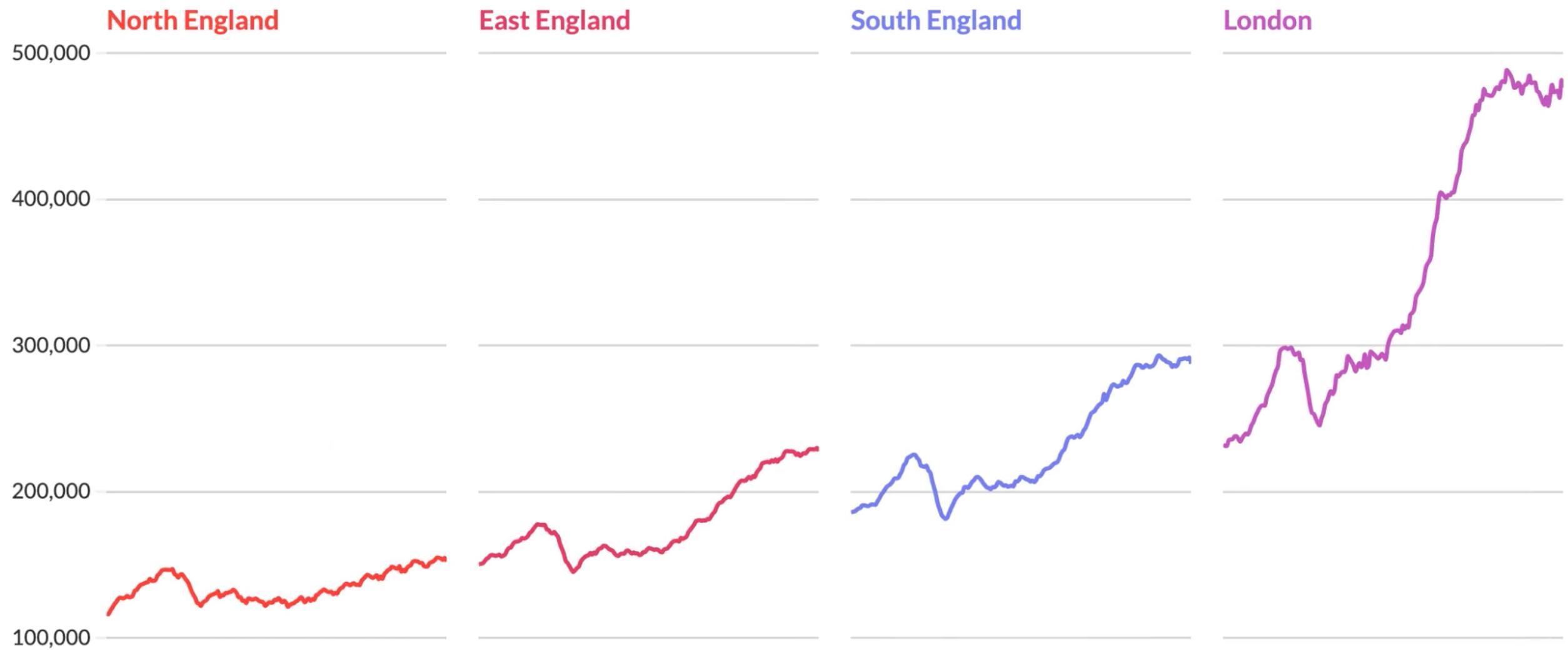
House prices have been increasing in England since 2005, but vary across regions

House prices in England



House prices have been increasing in England since 2005, but vary across regions

House prices in England





DATA STORIES IN THE WILD

PART II – EFFECTIVE STORYTELLING VISUALS

EXERCISES

Consider the following examples of charts found in the wild.

Are they examples of exploration, storytelling, situational awareness with data?

Are they data stories? If not, how would you turn them into stories?

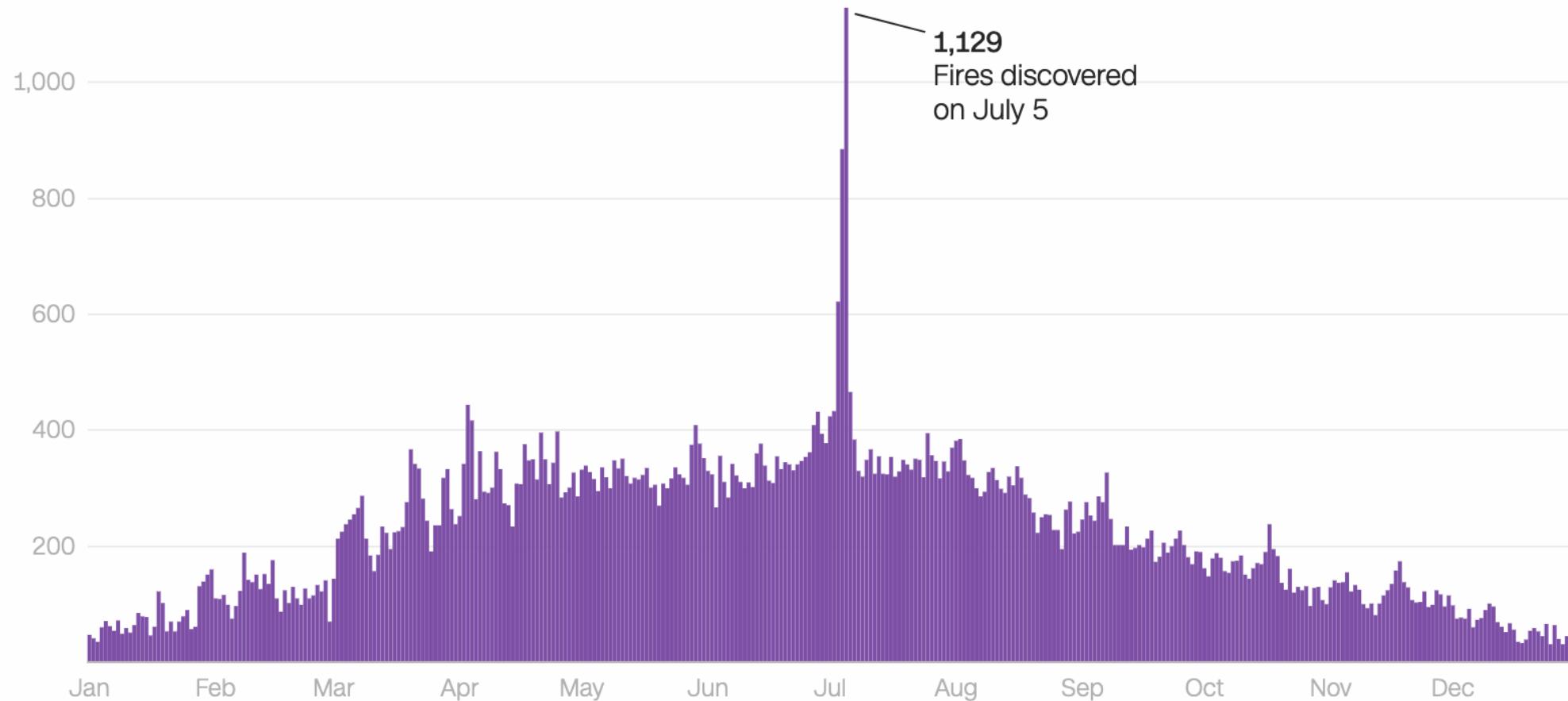
If so, are they good stories? Bad ones? Ugly ones?

If they are not good stories, how would you improve them?

Wildfires spike around July 4 holiday

Human-caused wildfires in the United States jump around Independence Day.

Total wildfires discovered each day of the year since 2014

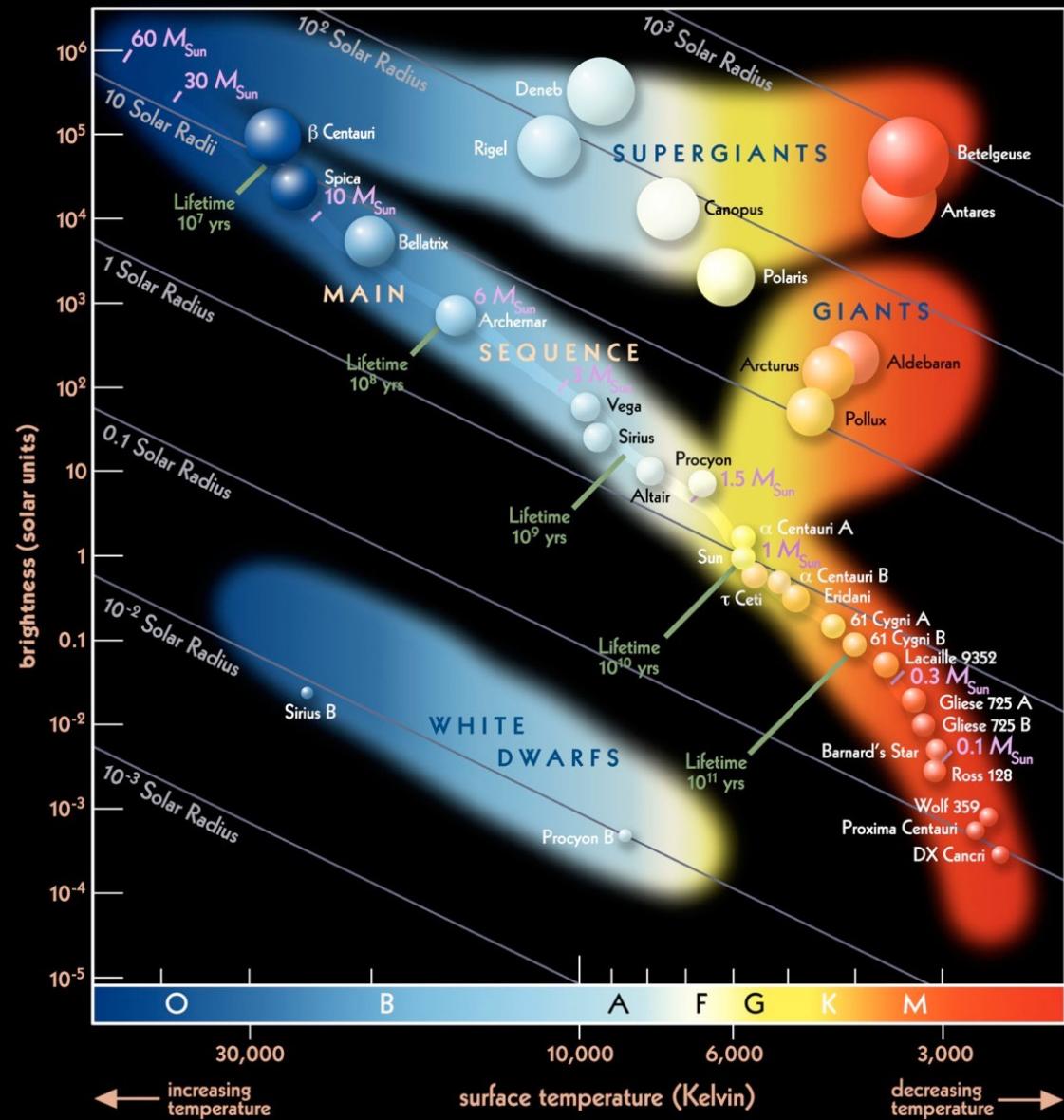


Sources: CNN analysis of data from the National Interagency Fire Center

Graphic: John Keefe, CNN

Human-caused fires, excluding prescribed fires. 2022 fires included through June 30. All incident times Eastern.

Hertzsprung-Russell Diagram



Data Elements

- star radius (x 2)
- surface temperature (x 2)
- spectral class
- brightness
- mass
- lifetime
- name

Underlying Structure

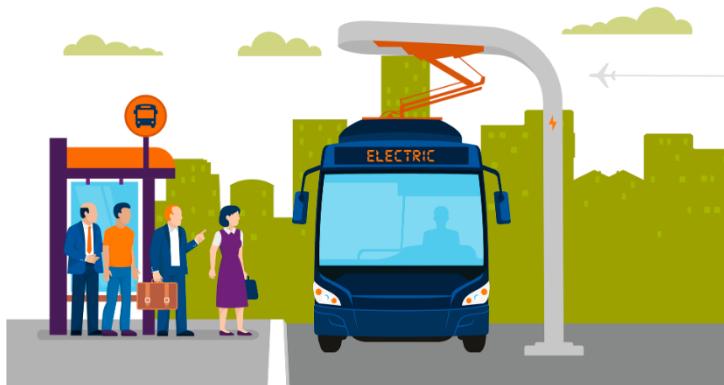
- 4 clusters/group
- lifetime, mass and radius are related to brightness and surface temperature on the Main Sequence

Only a subset of all the stars is shown in the HR diagram.

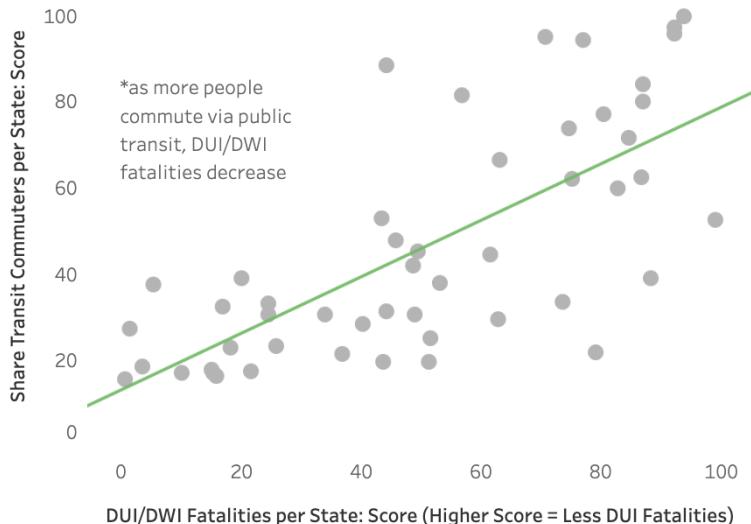
Public Transit and Complete Streets — How Do They Relate to Safety?

All data taken from The Bureau of Transportation Statistics (BTS), part of the Department of Transportation (DOT)

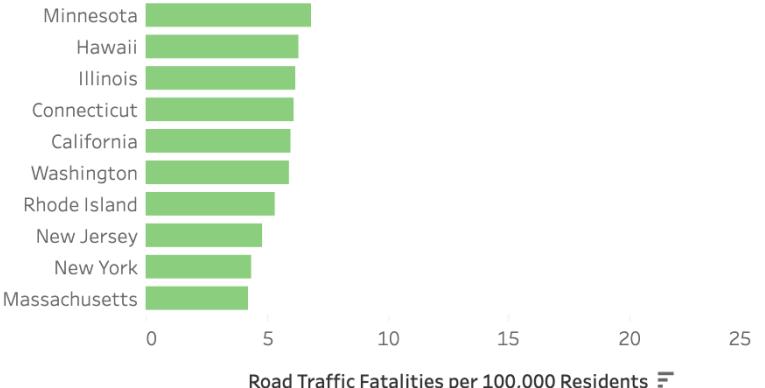
Complete Streets Policy?
 Policy in Place
 No Active Policy



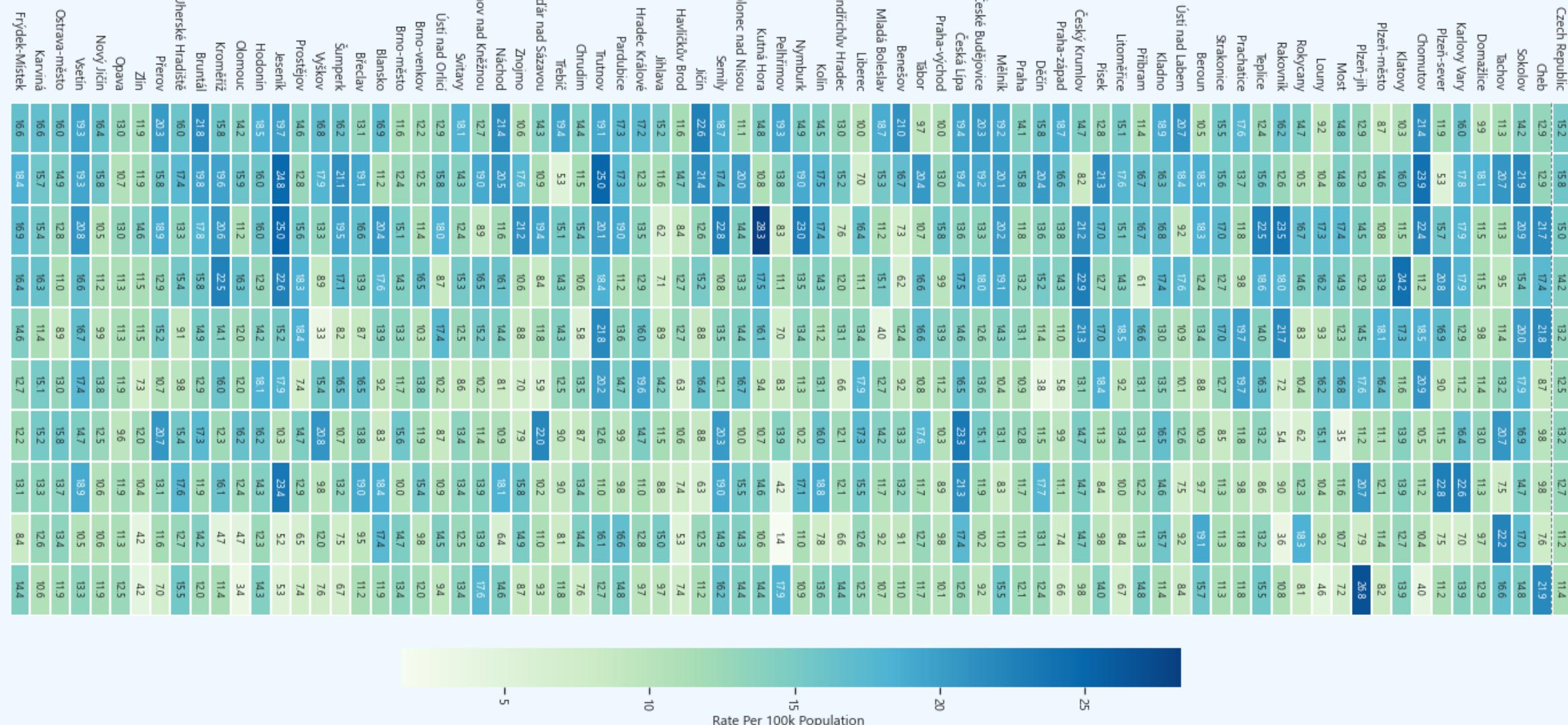
Top 10 Highest Auto Fatalities per 100k

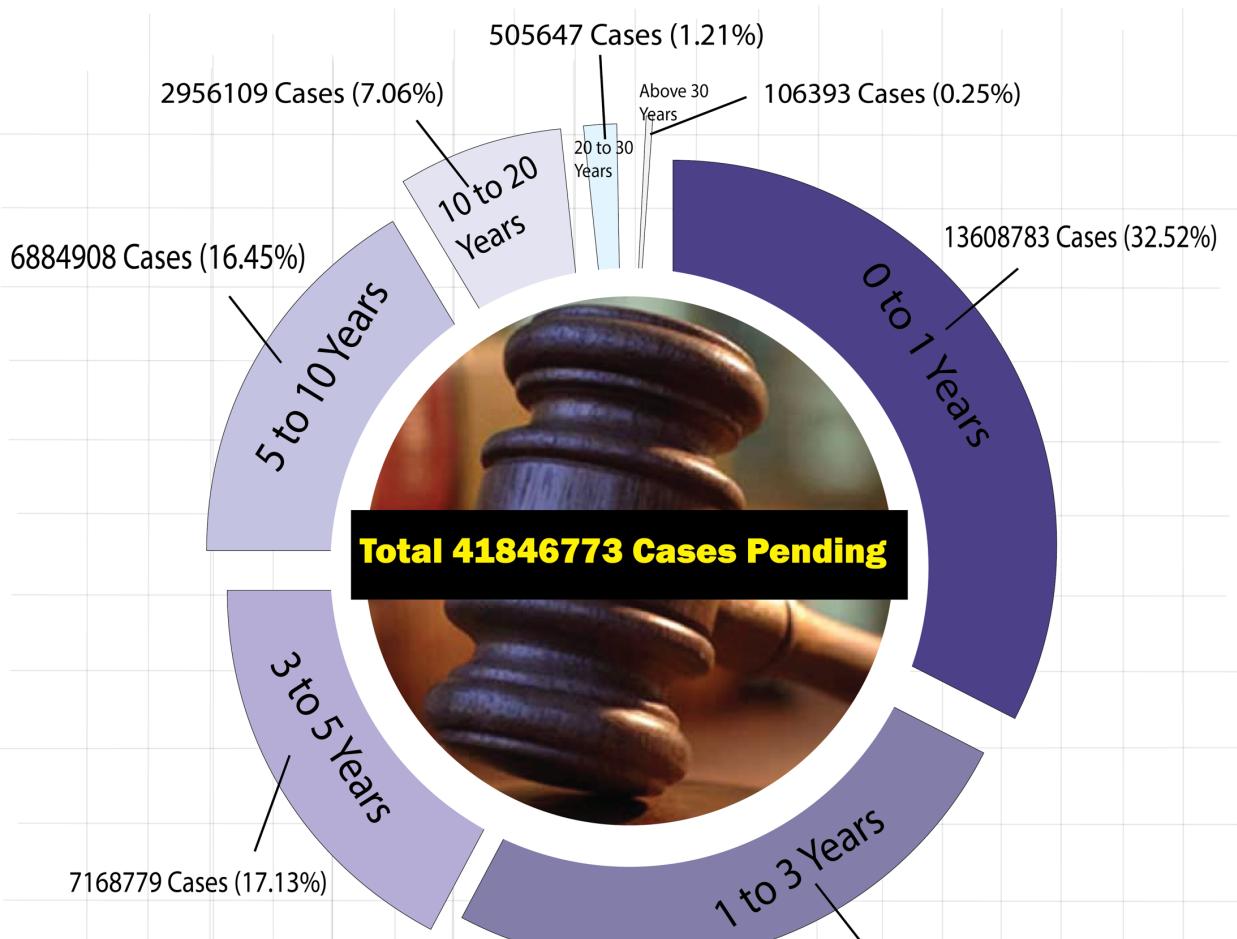


Top 10 Lowest Auto Fatalities per 100k



Heatmap for Suicide Rate 2011-2020

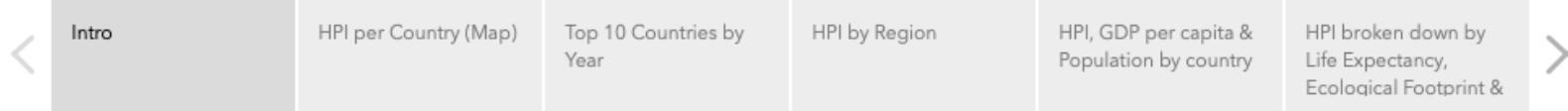




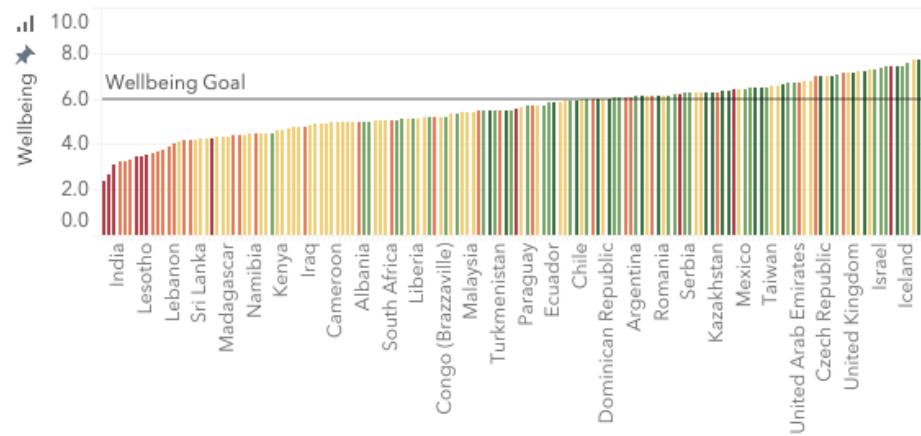
**How Many Cases
are pending in
'Indian Lower Courts' &
for how long ?**



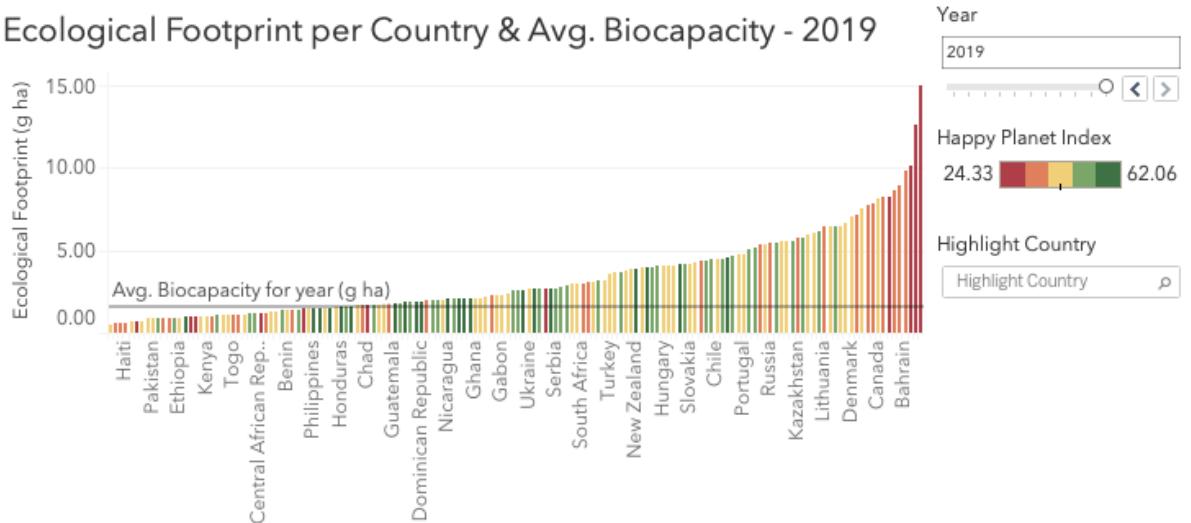
Happy Planet Index Vizualisation Story



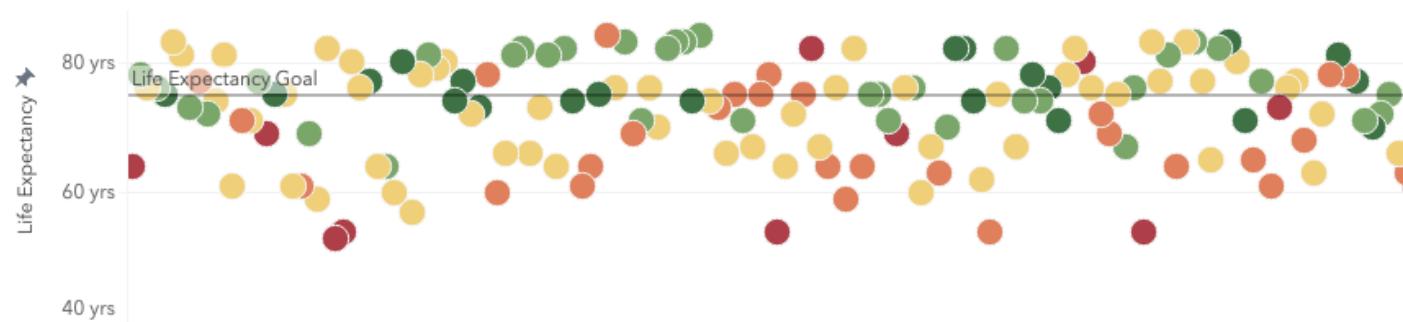
Wellbeing & HPI per Country - 2019



Ecological Footprint per Country & Avg. Biocapacity - 2019



Life Expectancy & HPI per Country



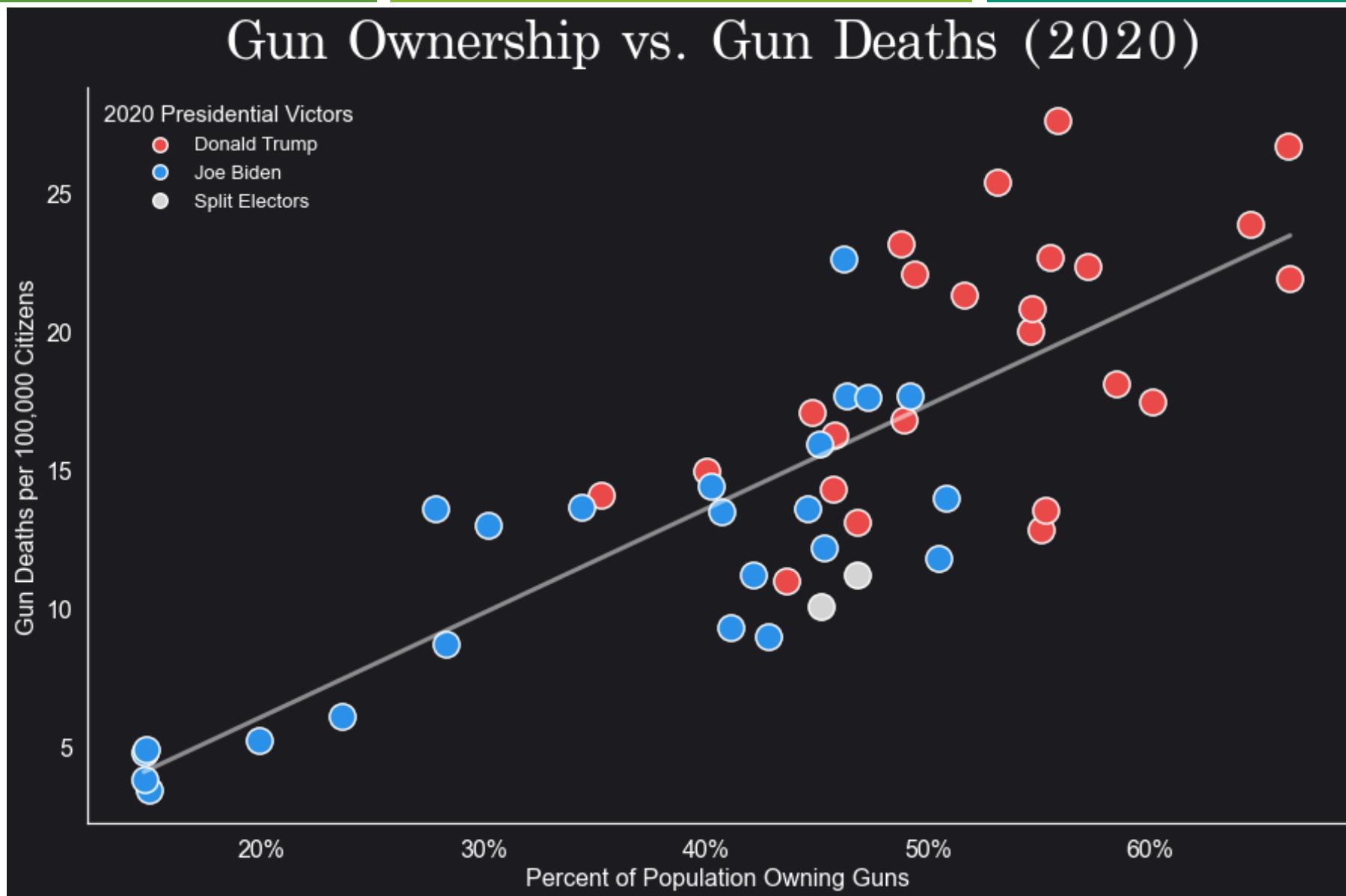
What is the Happy Planet Index?

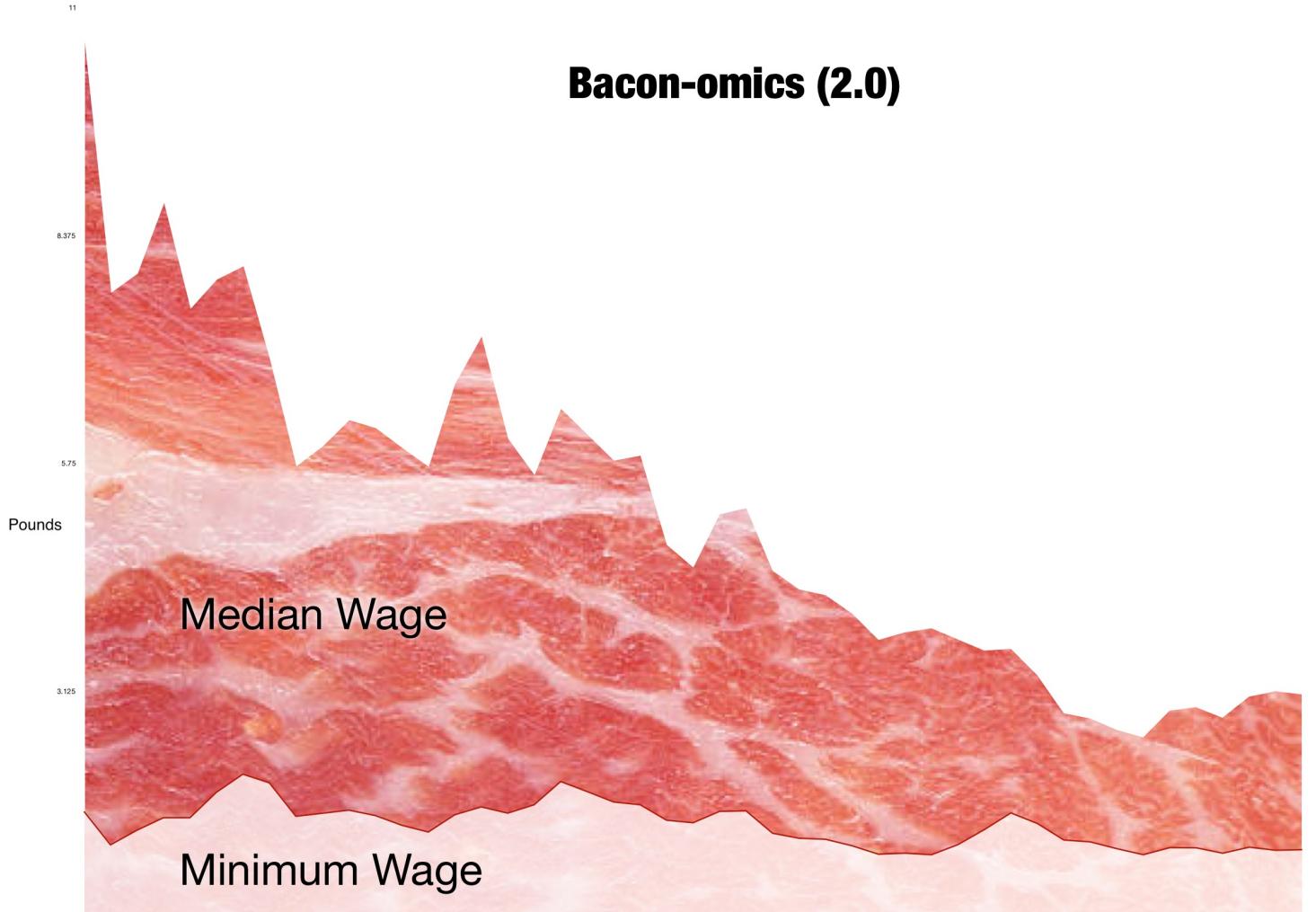
The Happy Planet Index is a measure of sustainable wellbeing, ranking countries by how efficiently they deliver long, happy lives using our limited environmental resources.

"Is it possible to live good lives without costing the Earth?"

*Learn more by clicking the HPI logo >







How Many Pounds of Bacon (12mo average price) you could buy working 1 hour @ Median Wage and Minimum Wage

Source: CPI Data for Price of Bacon Per Pound 1970 to Present (BLS Beta Labs)

Minimum Wage for [dol.gov](https://www.dol.gov)

Data for Median Wage: fred.stlouisfed.org - Real Median Personal Income Table Median Wage divided by 2080 hours (40 hours per week, 52 Weeks Per Year) To Get Hourly Rate

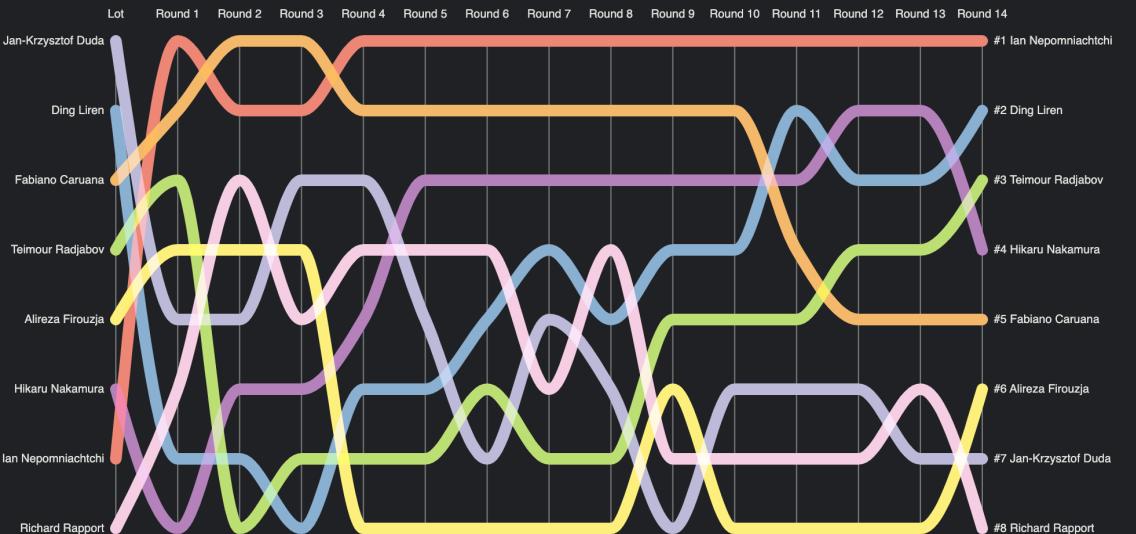
2022 Chess Candidates Tournament

u/boxer-collar  ebemunk  ThinkingThroughTheParty
data from [lichess.org](#)

Nepomniachtchi wins for a second time, **Radjabov** surprises in the last minute, **Caruana** implodes in the latter half. It will be **Nepomniachtchi** vs **Ding** if **Carlsen** doesn't defend his title.

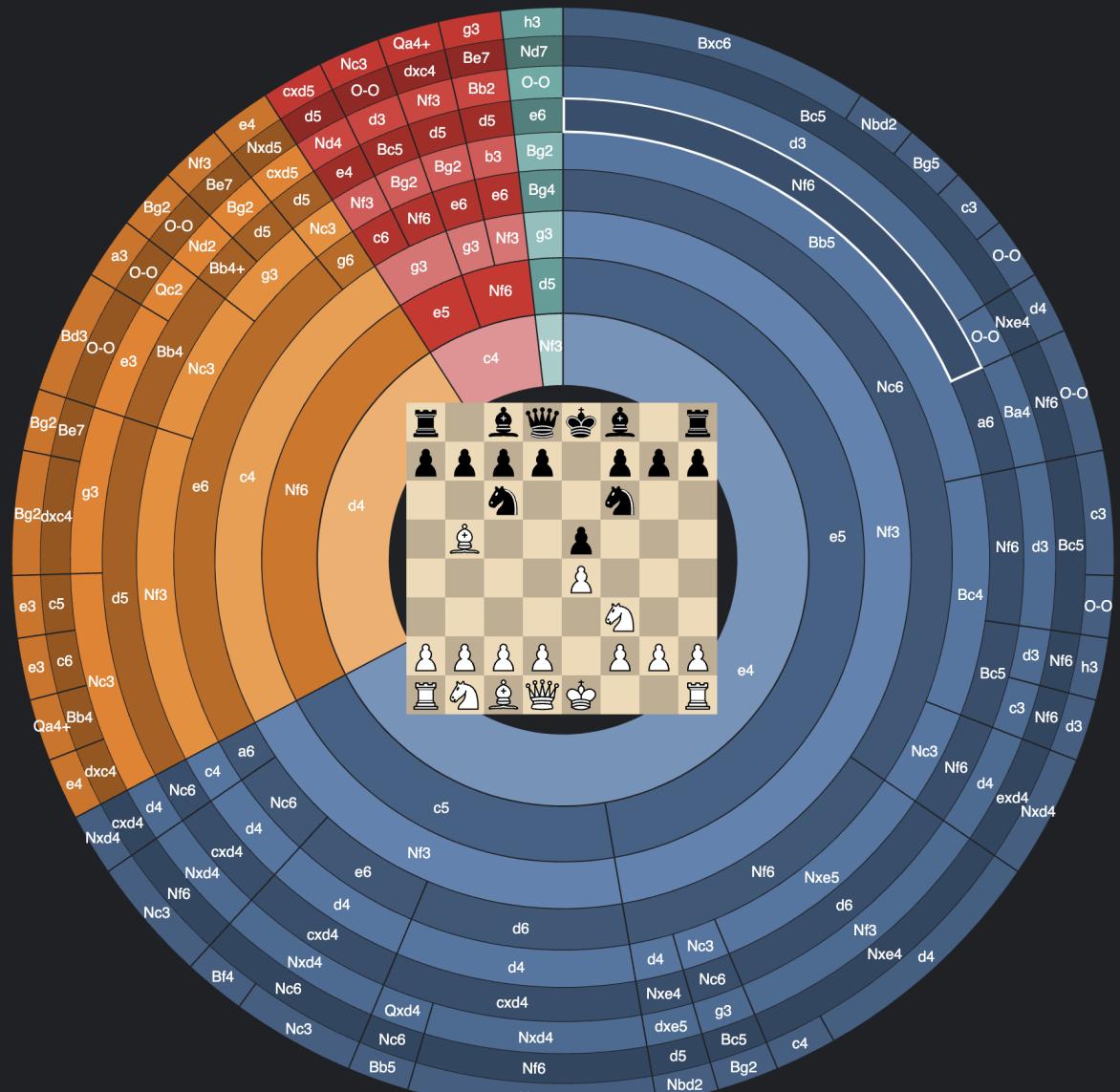
Only Rounds 3 and 5 saw all-draws. **Nepomniachtchi** was the only player with no back-to-back black games while **Ding** had 2 pairs and **Firouzja** had 2 pairs of back-to-back whites. **Caruana** was the only one with no back-to-back whites.

Nepomniachtchi kept his lead throughout the whole tournament without a single loss. **Caruana** had good chances but nosedived after Round 10. **Ding** had a slow burn but finished 2nd. **Radjabov** started winning after round 9 to end up 3rd. **Nakamura** lost out on €31,000 with his last round loss.



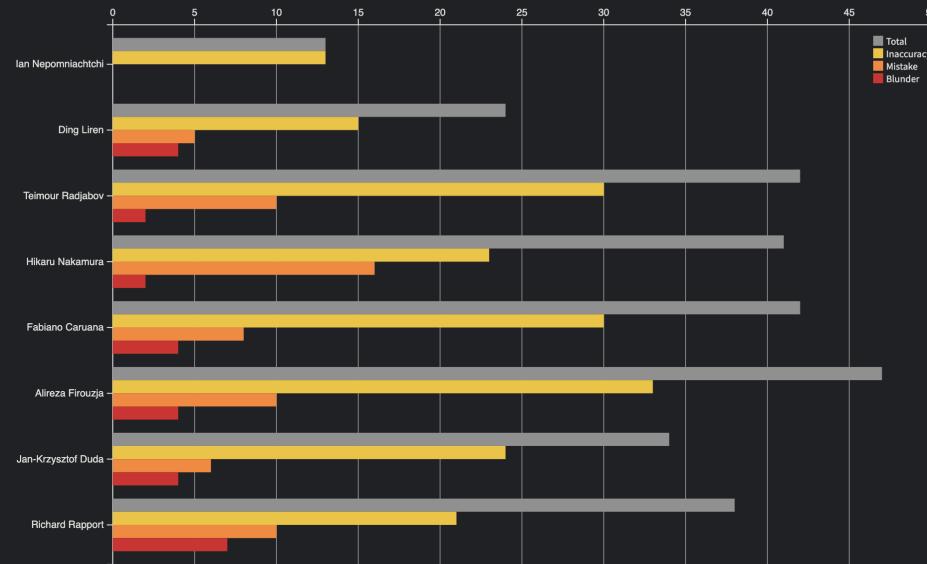
Openings

Berlin Defense (C65) is still a favorite at this level, followed by **Petrov Defense (C42)** and **Sicilian Najdorf (B90)**. Below is the openings chart of the first 9-ply from every game, with the most popular line highlighted.



Inaccuracies, Mistakes and Blunders

Count of mistakes from engine evaluation. Nepomniachtchi outclassed the field with <1 inaccuracy per game throughout the tournament. Even though Firouzja had more total mistakes, his second win gave him the edge over Duda and Rapport.



Time Management

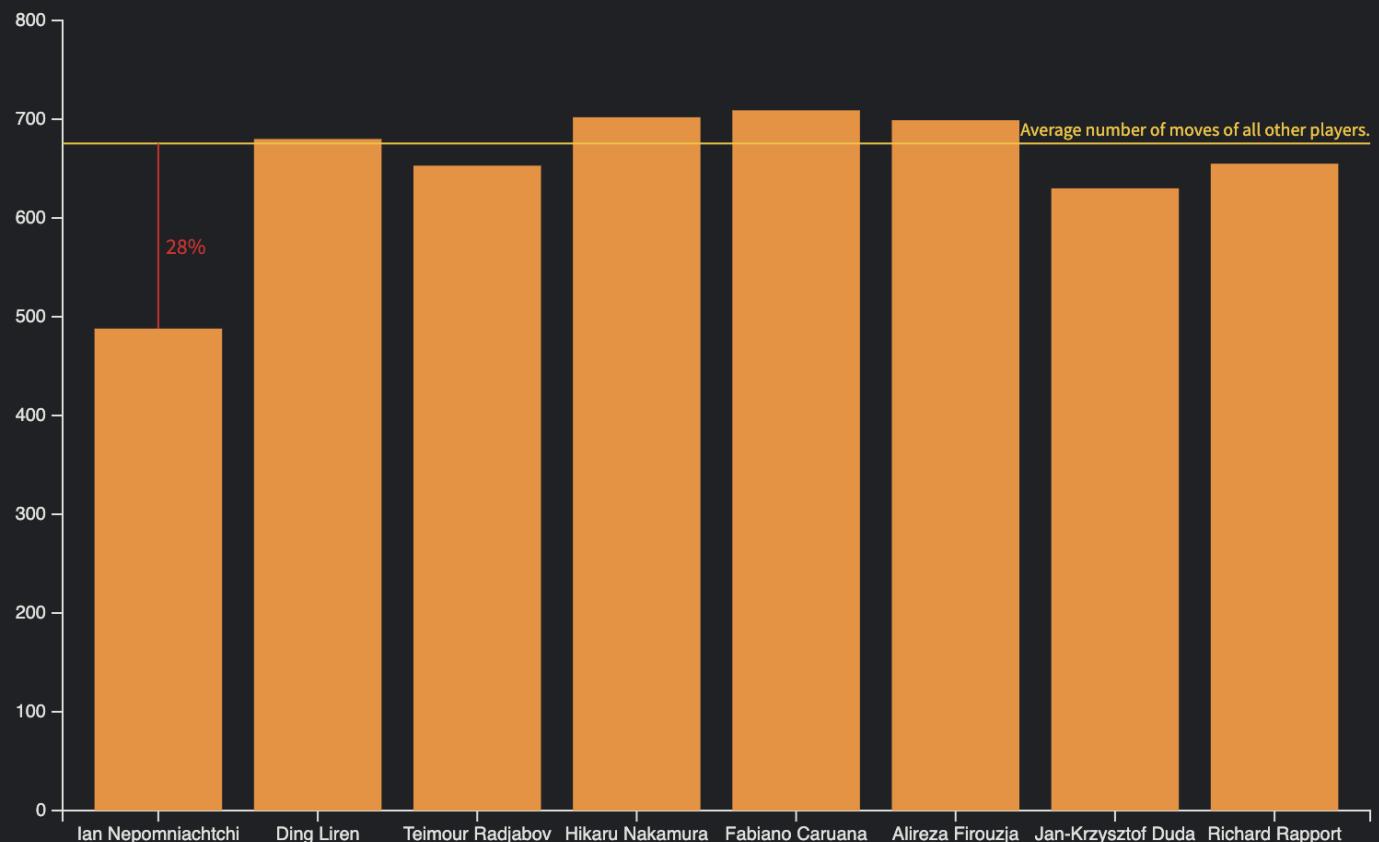
Remaining time after every move in each round. Players generally got better after every round.

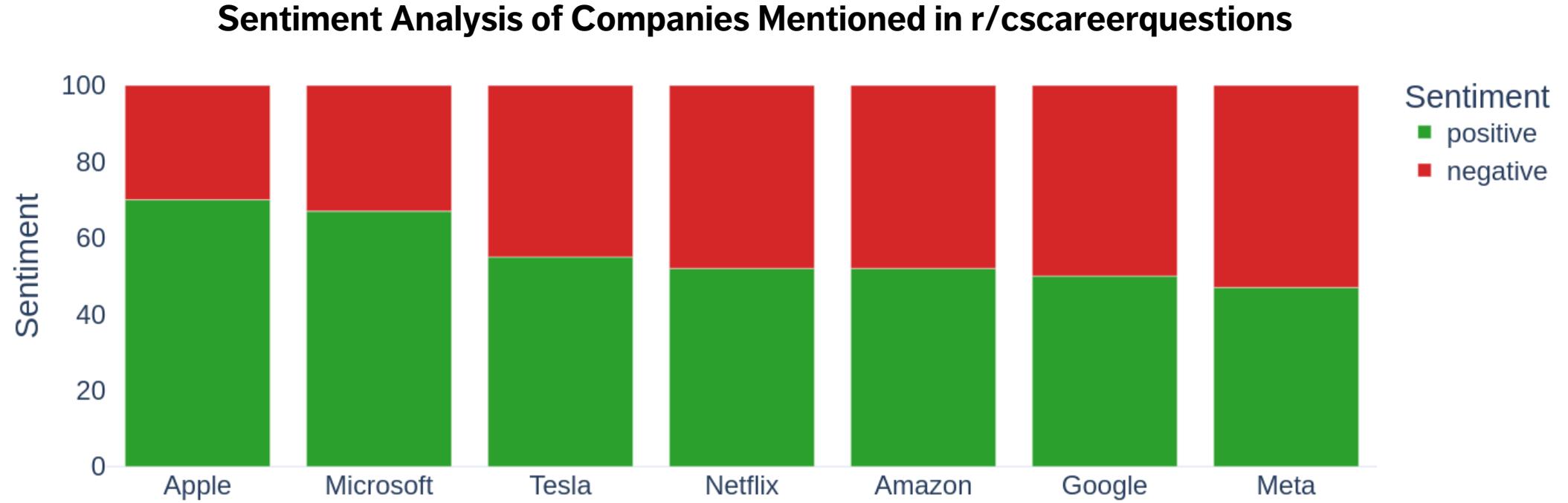


https://old.reddit.com/r/dataisbeautiful/comments/vsuy99/oc_i_visualized_the_games_from_the_2022/

Number of Moves

Nepomniachtchi's trick to winning? Just play less! He made **28%** fewer moves than the other players.





@RobLawrencium

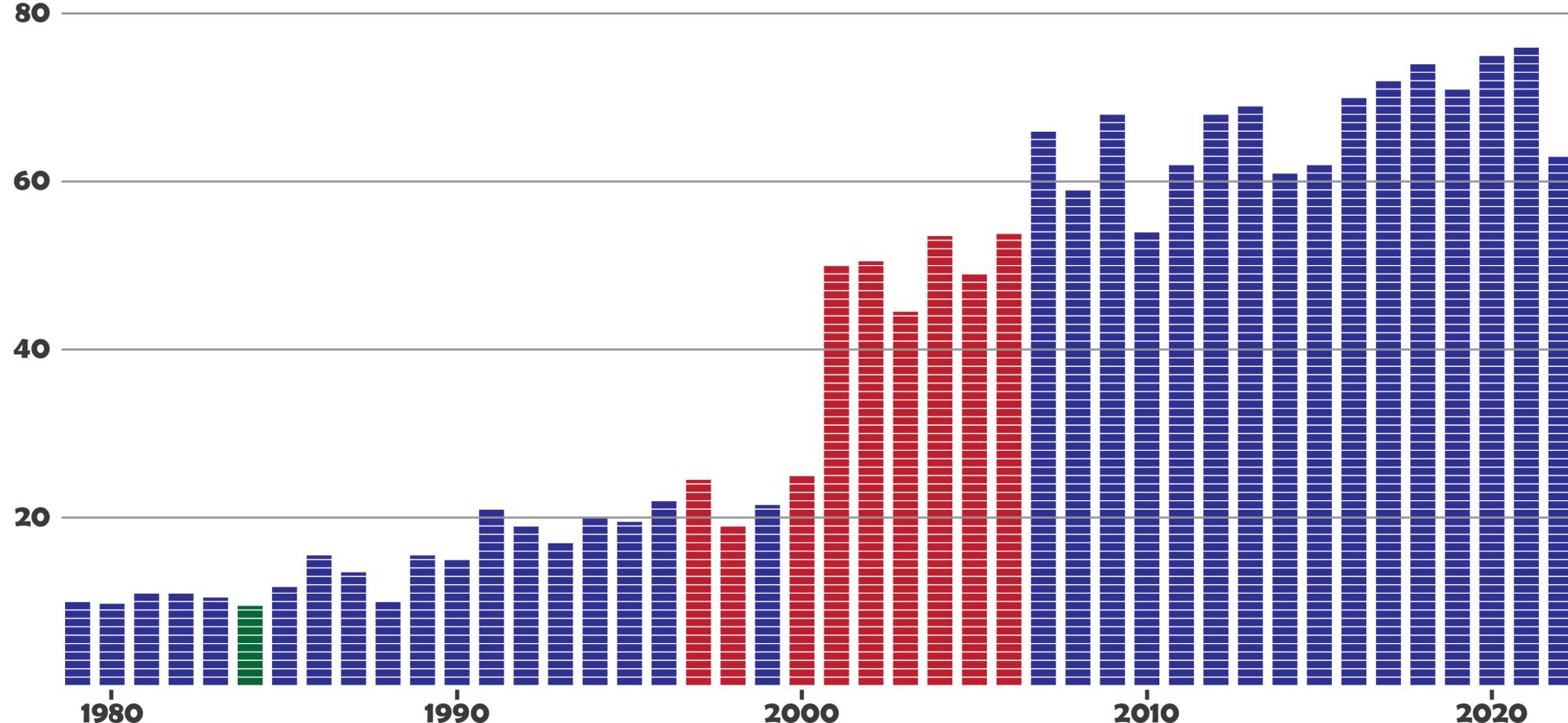
Nathan's Hot Dog Eating Contest

Hot Dog + Bun Count for Independence Day Winners

U.S. Winner

Germany Winner

Japan Winner

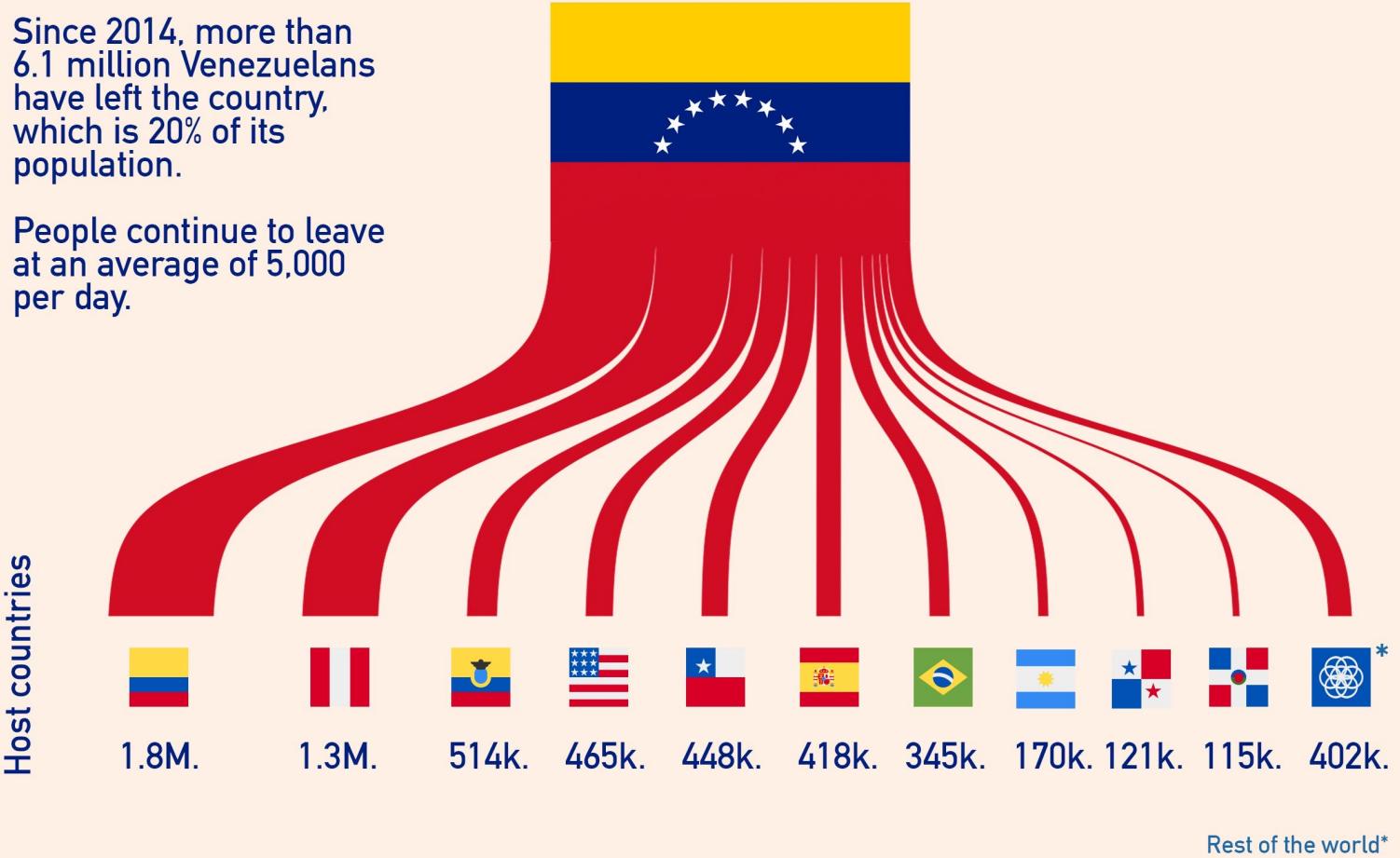


Venezuela is bleeding.

Source :UNHCR
Twitter PietroViolo

Since 2014, more than 6.1 million Venezuelans have left the country, which is 20% of its population.

People continue to leave at an average of 5,000 per day.



REFERENCES

STORYTELLING WITH DATA

REFERENCES

W. Battle-Baptiste and B. Rusert, *W.E.B. DuBois's Data Portraits: Visualizing Black America.* Princeton Architectural Press, 2018.

P. Boily, S. Davies, J. Schellinck, *The Practice of Data Visualization.* Data Action Lab, 2022.

M. Bowles, C. Burns, J. Hixson, S. Austin Jennes, and K. Tellers, *How to Tell a Story.* Crown, 2022.

A. Cairo, *The Functional Art.* New Riders, 2013.

A. Cairo, *The Truthful Art.* New Riders, 2016.

REFERENCES

- S. Evergreen, *Effective Data Visualization: the Right Chart for the Right Data*, Second edition. Thousand Oaks, California: SAGE Publications, Inc.
- M. Friendly and H. Wainer, *A History of Data Visualization and Graphic Communication*. Harvard University Press, 2021.
- Z. Gemignani and C. Gemignani, *Data Fluency: Empowering Your Organization with Effective Data Communication*. Wiley, 2014.
- Z. Gemignani and C. Gemignani, *A Guide to Creating Dashboards People Love to Use*.
- K. Healey, *Data Visualization: A Practical Introduction*, 2018.

REFERENCES

- S. McCloud, *Understanding Comics: The Invisible Art*. Harper, 1994.
- S. McCloud, *Making Comics: Storytelling Secrets of Comics, Manga and Graphic Novels*. Harper, 2006.
- I. Meireilles, *Design for Information*. Rockport, 2013.
- C. Nussbaumer Knafllic, *Storytelling with Data*. Wiley, 2015.
- S. Rendgen, *The Minard System : the complete statistical graphics of Charles-Joseph Minard, from the collection of the École Nationale des Ponts et Chaussées*. Princeton Architectural Press, 2018.