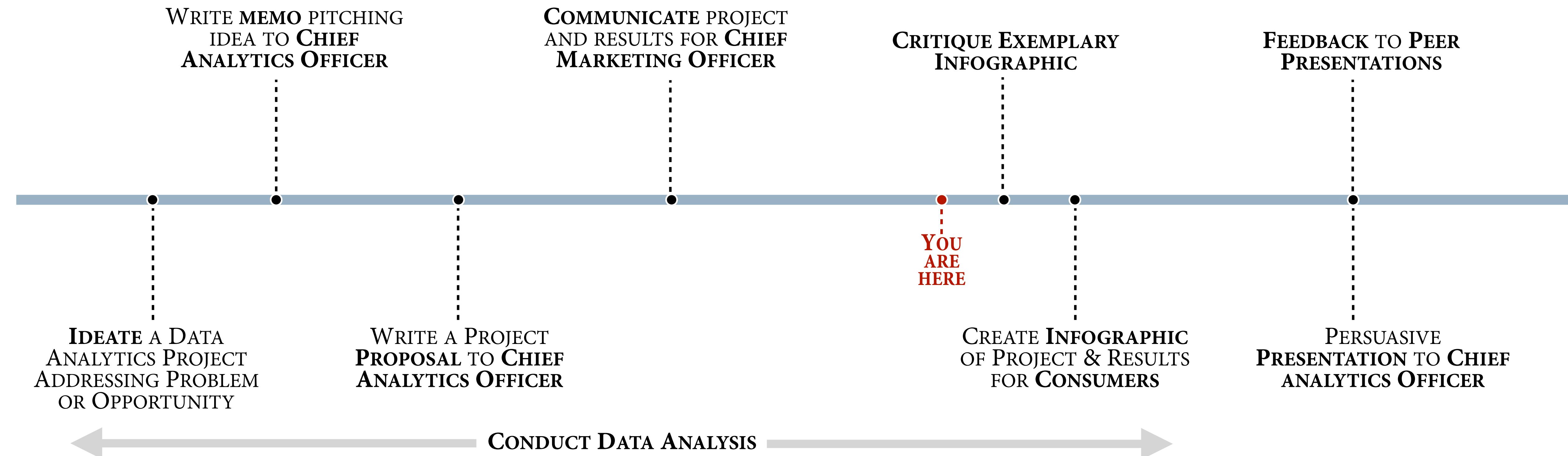


Storytelling With Data
information graphics; uncertainty

Conceptual project timeline



What are information graphics?

What are information graphics? | *dictionary definition is too broad, generic*

INFOGRAPHIC n. a visual image such as a chart or diagram used to represent information or data in an easily understandable form.

What are information graphics? | *viewpoints at Malofiej — the infographics world summit*



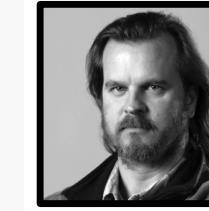
Gregor Aisch

Infographics is an abbreviated form of “information graphics”. It seems to mean a lot of different things to different people. I rarely use the term.



Federica Fragapane

A **visual translation of data** and information: a language to communicate topics, contents and **stories** to people.



Laris Karklis

Infographics is . . . using **visuals** to tell a **story**.

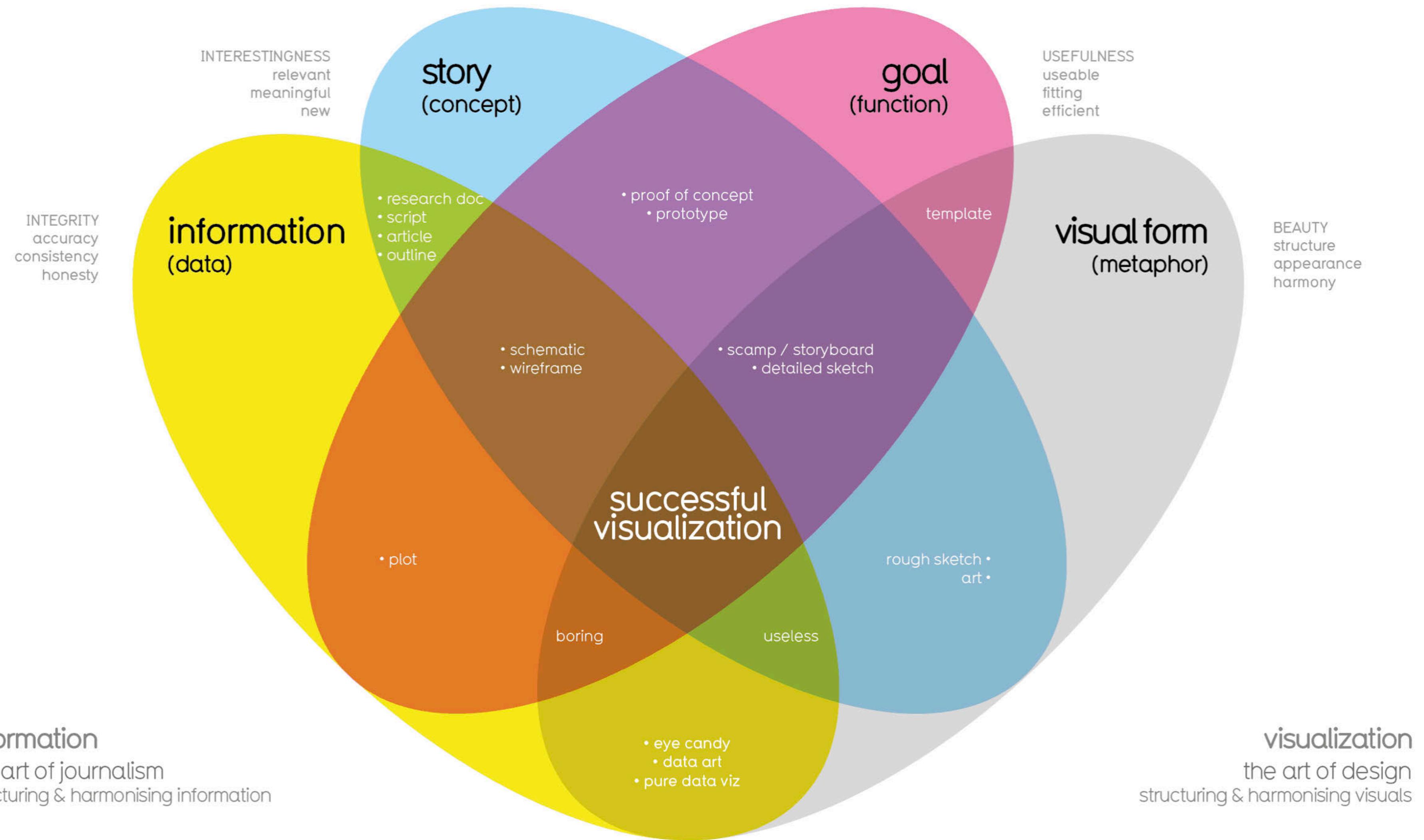


Nadieh Bremer

Infographics . . . combine graphical elements, such as a drawn portion of an animal, human, map, etc. with small mini **data visualizations** (a small bar chart for example) and annotations to tell a **story**.

What Makes a Good Visualization?

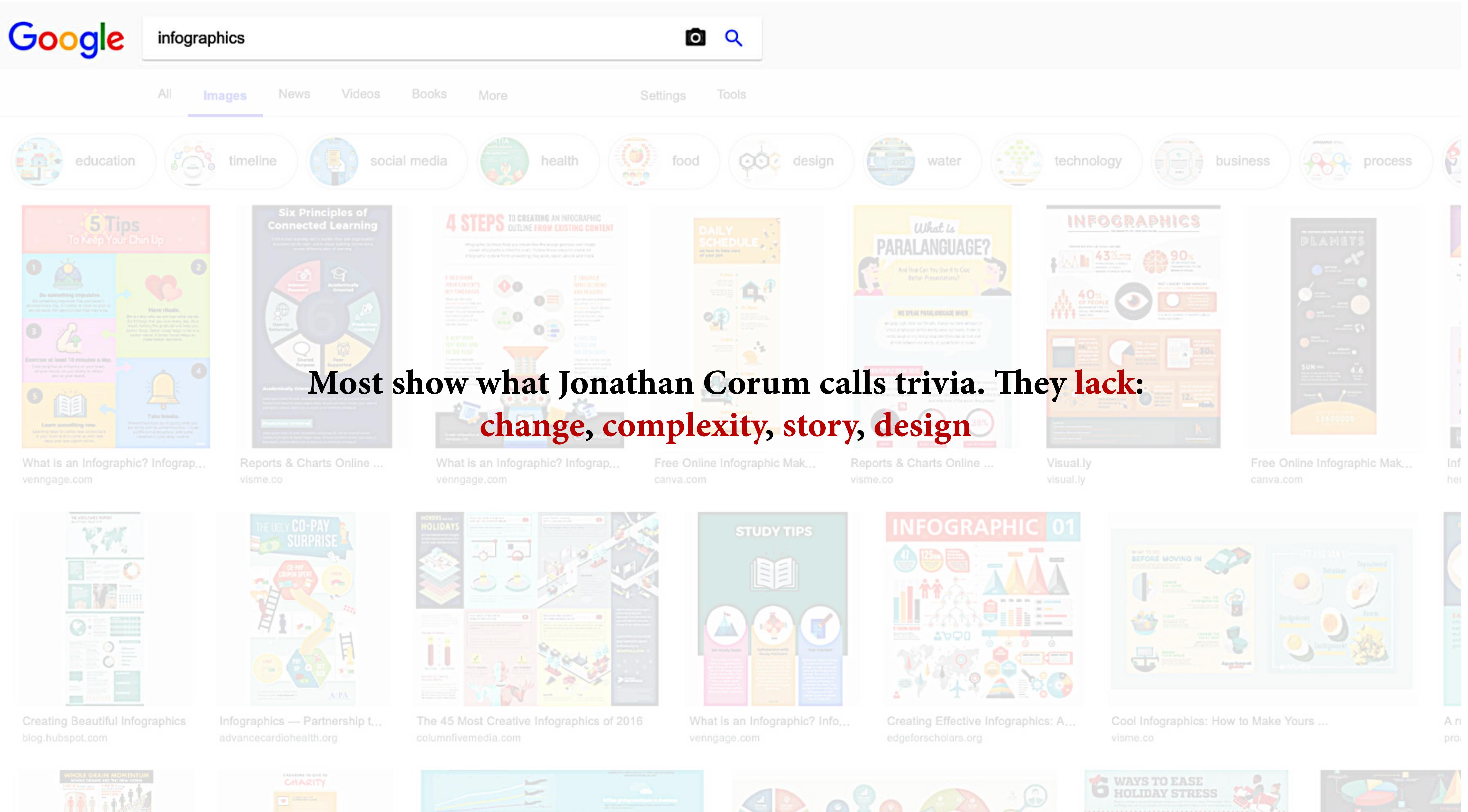
explicit (implicit)



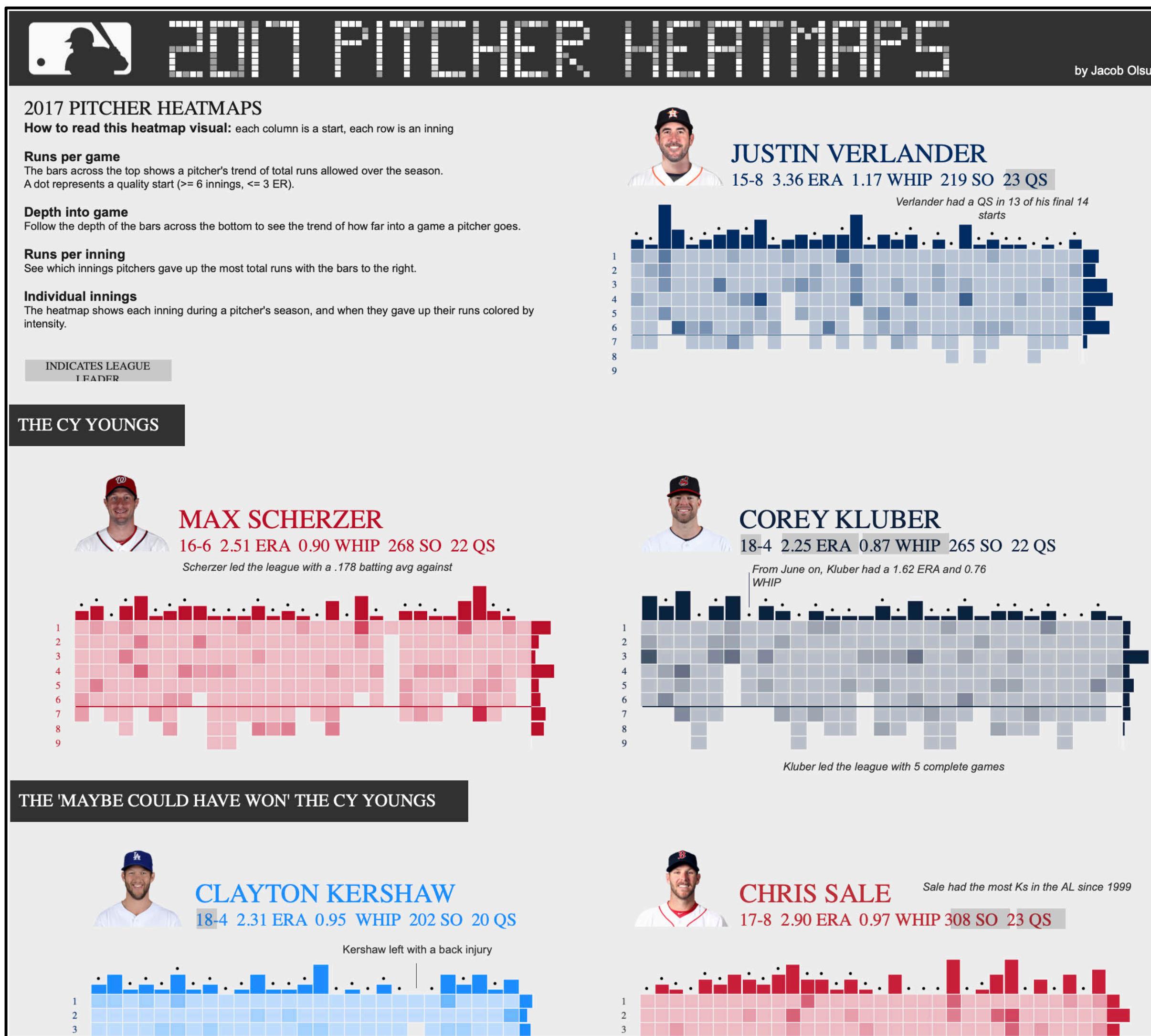
information
the art of journalism
structuring & harmonising information

visualization
the art of design
structuring & harmonising visuals

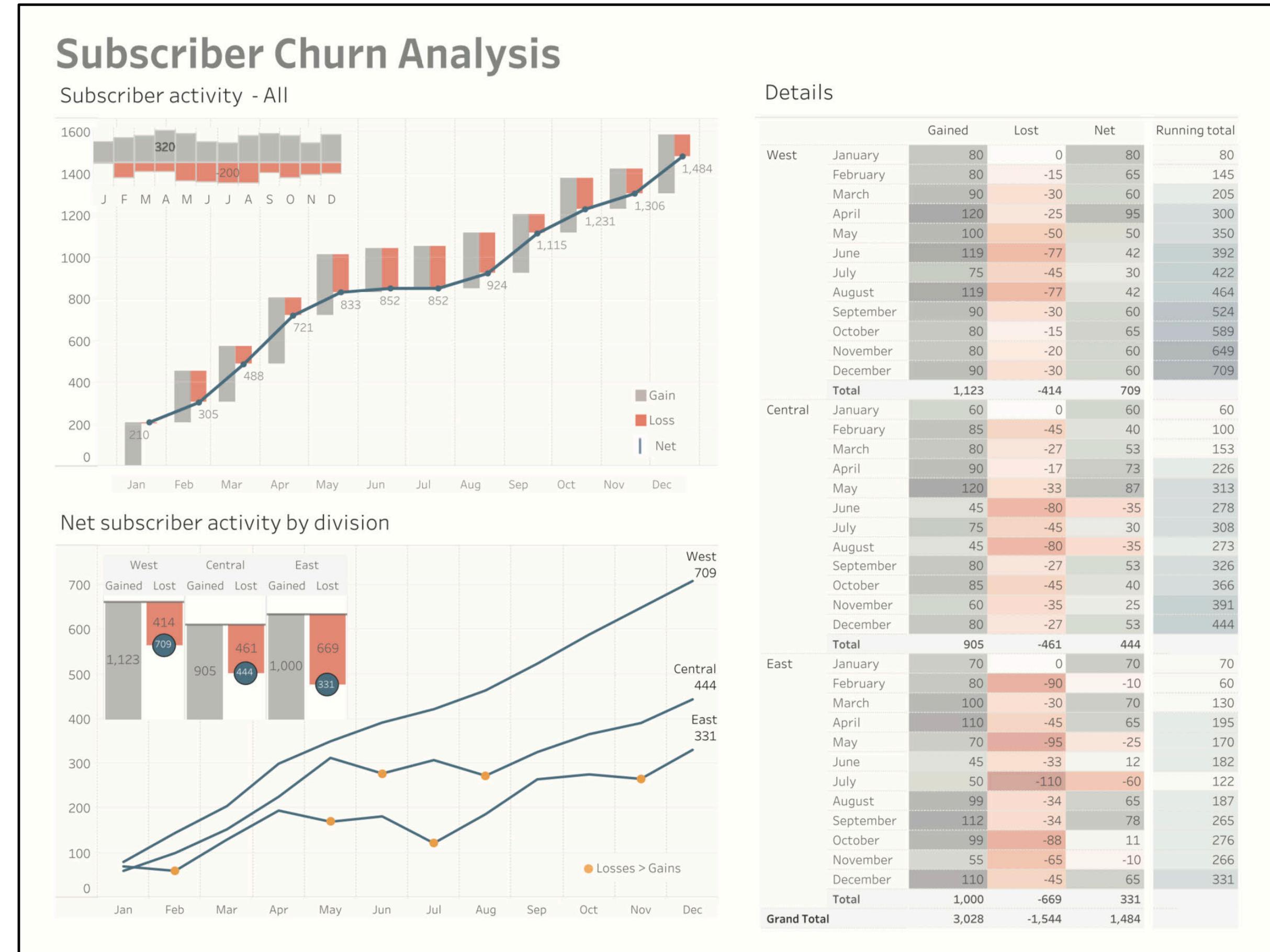
What are information graphics? | *to play the roulette wheel of **bad** examples, just google “infographics”*



What are information graphics? | *without a narrative, it's just trivia, list of facts*



What are information graphics? | *dashboards, unlike infographics of interest, lack narrative or story*



We differentiated ... “list of fact” infographics from the infographics medium as a whole, and chose to *exclude* them because this specific submedium *lacks authorial narrative*.



Riche, co-editors

TELL A COMPLETE STORY where the purpose is to inform, entertain or persuade the audience (to act). It should have:

- clear, focused messages revealed through narrative
- new, surprising information
- credible data sources

- visually coherent and integrated graphics and narrative
- comparisons and contrasts for context and meaning
- principles of information design, be organized

Are infographics storytelling appropriate?

Are infographic stories appropriate? | yes, here's how and why

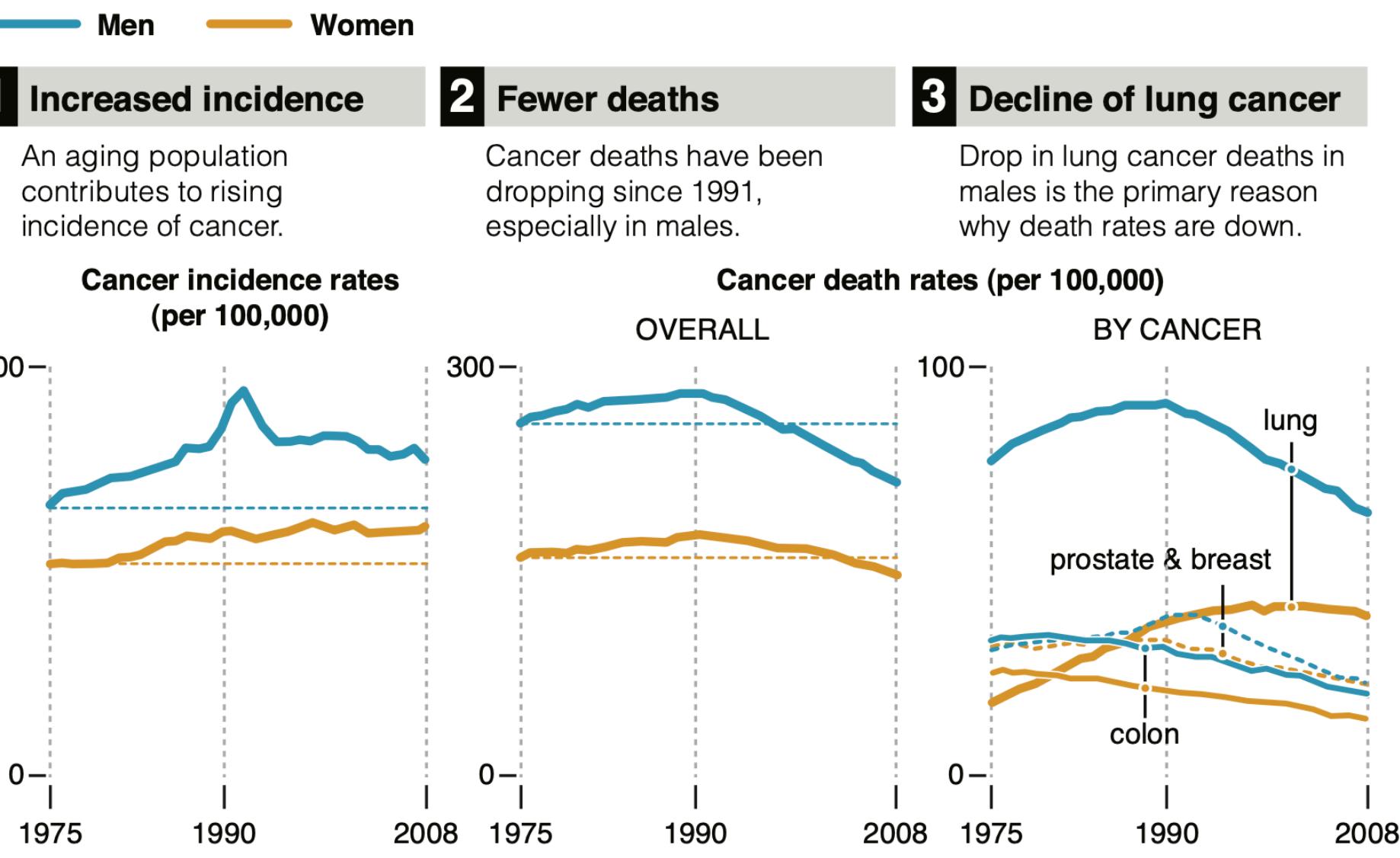
Maintain focus ... by leaving out detail that does not advance the plot.
Distinguish necessary detail from minutiae; do not give in to the
desire to show all your hard-won data. Provide sufficient support for
your story, but stick to the plot.



A story adds meaning and clarity to complex statistics. Use multiple panels to establish flow, and use colloquial language when addressing a general audience. . . . Always be accurate, but balance qualitative and quantitative expositions. An occasional tangent . . . adds texture to the presentation without diluting the message.

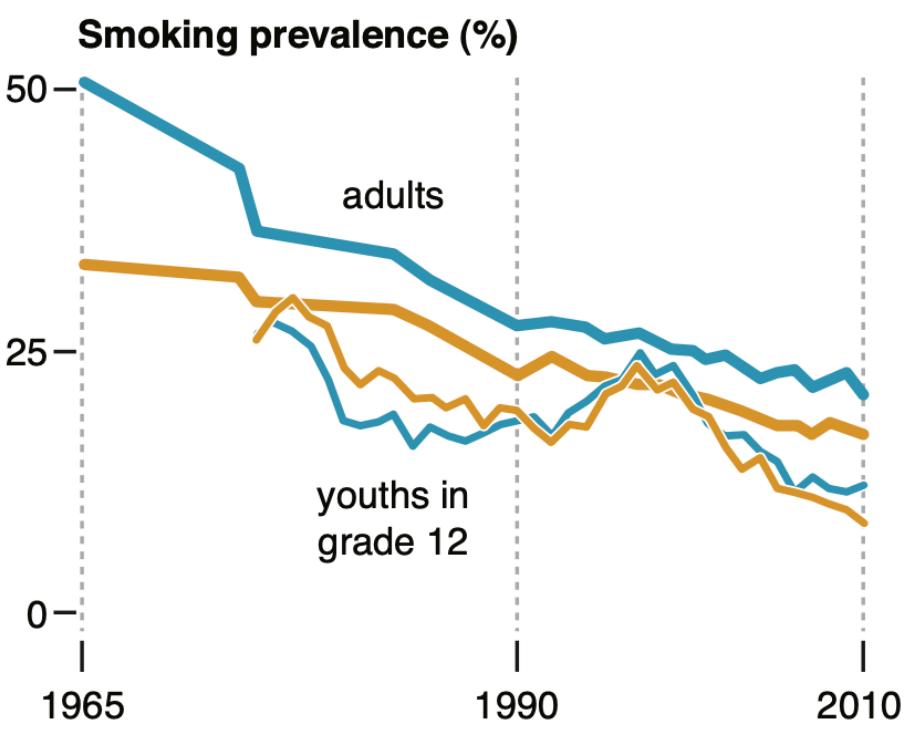
WHERE THERE'S SMOKE—THERE'S CANCER

Cancer rates are up, but mortality is down. New diagnostics and treatments are responsible for part of this trend. But the greatest single contributing factor is the decline in smoking—rates are at their lowest level in 50 years.



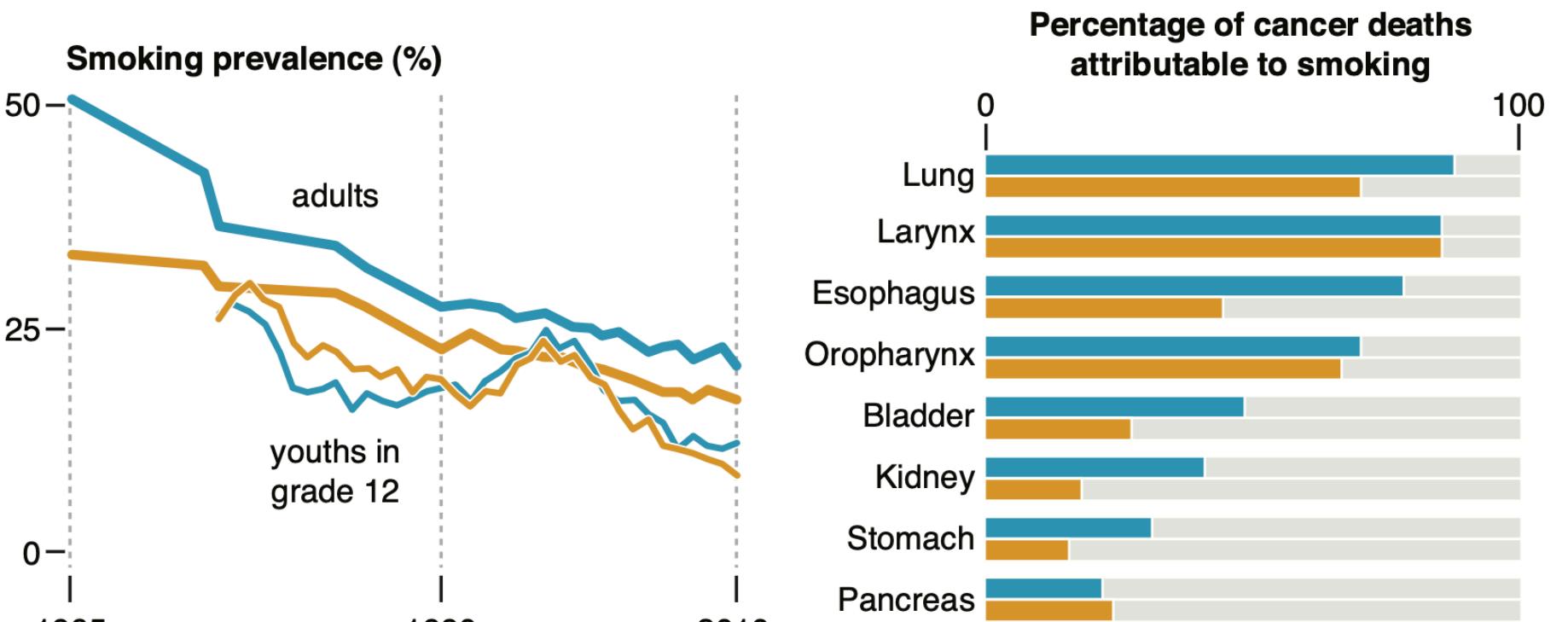
4 Decline in smoking

Since the 1964 first Surgeon General's report, smoking rates have been dropping. By 2010, the rate among males was down to 20%, from 50% at its peak. Among youths, rates have been on an even steeper decline since 1997.



5 Impact of smoking on cancer deaths

Smoking is a major risk factor for many types of cancer and significant contributor to cancer-related deaths. It remains the single largest preventable cause of disease and premature death in the US.



source: American Cancer Society Cancer Statistics 2012; Monitoring the Future (University of Michigan).

Are infographic stories appropriate? | *critics say no*

“Against storytelling of scientific results”

Use of graphics storytelling often result in a *distorted* and *unrepresentative* display of data.

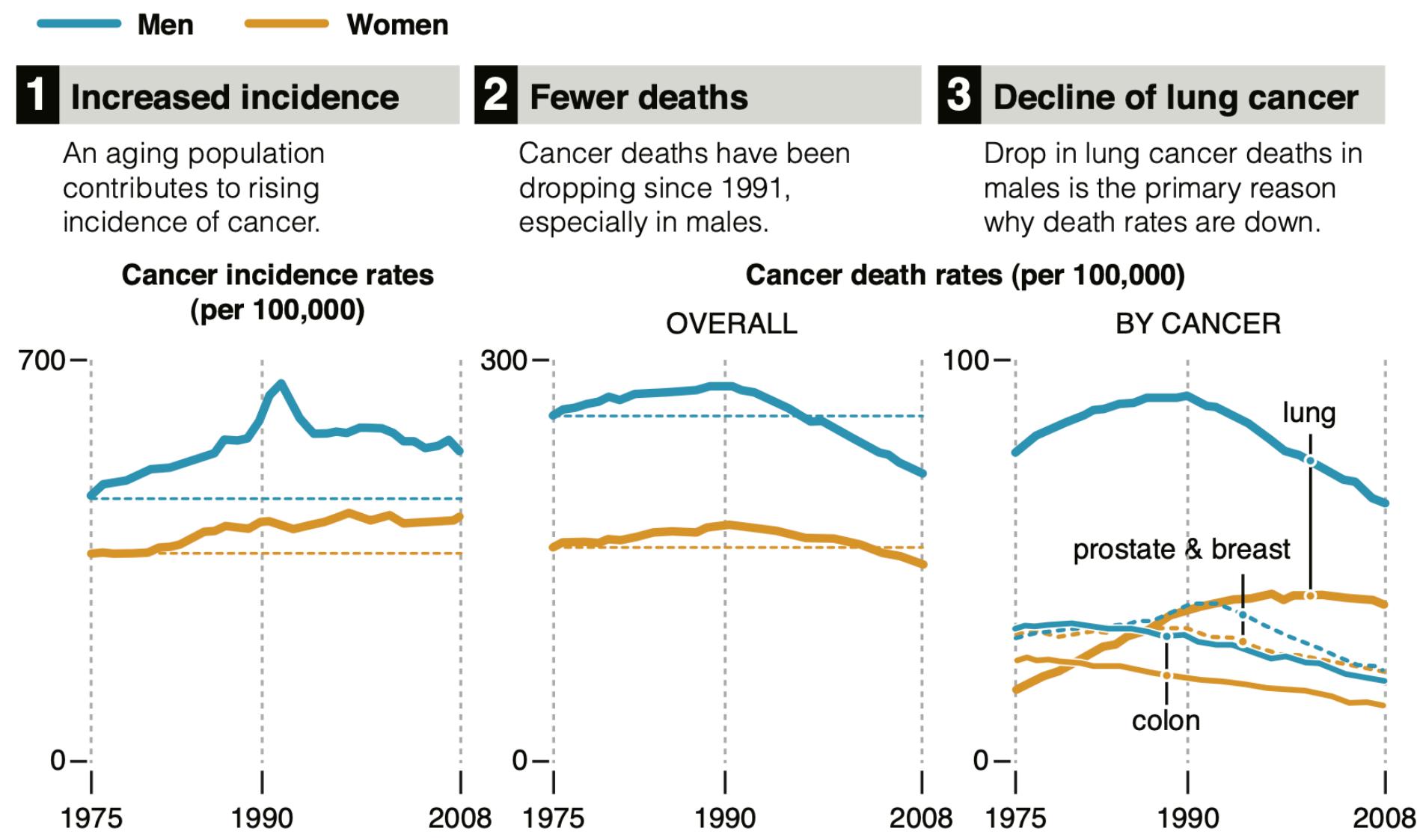


Great storytellers *embellish* and *conceal* information to evoke a response in their audience. Inconvenient truths are swept away, and marginalities are spun to make a point more spectacular. A storyteller would plot the data in the way most persuasive rather than most informative or representative.

How do Cairo and Krzywinski reconcile the conflicting views?

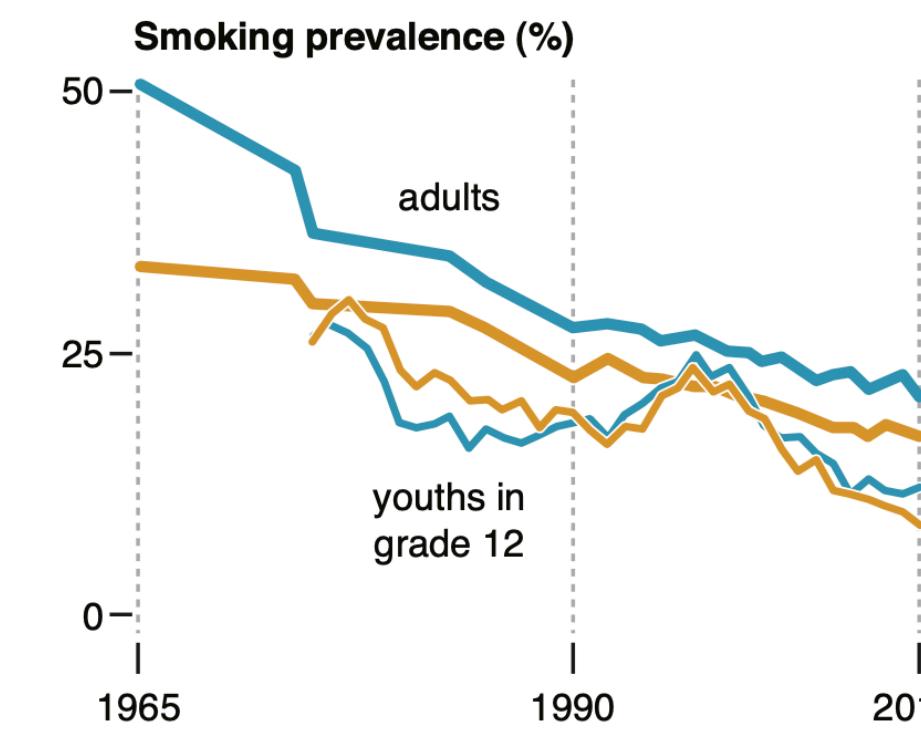
WHERE THERE'S SMOKE—THERE'S CANCER

Cancer rates are up, but mortality is down. New diagnostics and treatments are responsible for part of this trend. But the greatest single contributing factor is the decline in smoking—rates are at their lowest level in 50 years.



4 Decline in smoking

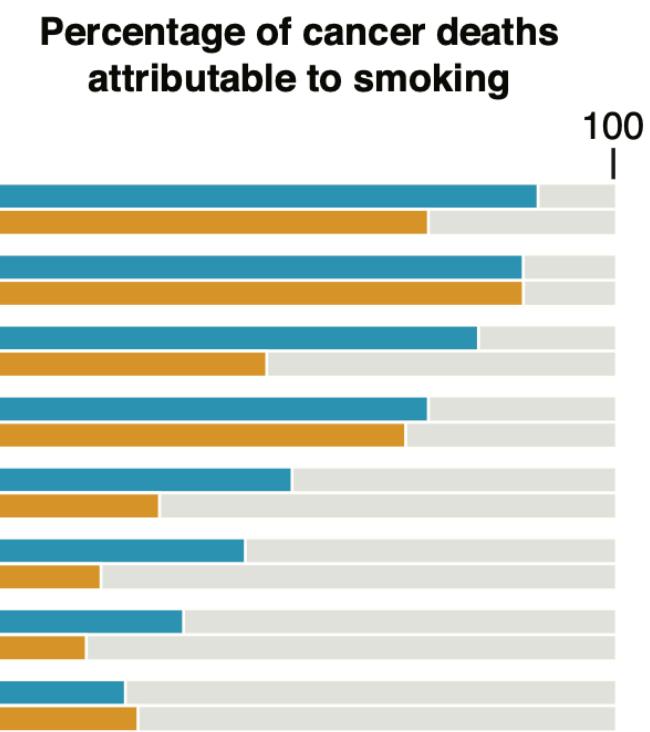
Since the 1964 first Surgeon General's report, smoking rates have been dropping. By 2010, the rate among males was down to 20%, from 50% at its peak. Among youths, rates have been on an even steeper decline since 1997.



source: American Cancer Society Cancer Statistics 2012; Monitoring the Future (University of Michigan).

5 Impact of smoking on cancer deaths

Smoking is a major risk factor for many types of cancer and significant contributor to cancer-related deaths. It remains the single largest preventable cause of disease and premature death in the US.



critiquing infographics examples

infographics examples for discussion

Audience?

Does this infographic seem designed to communicate with an identified audience? If so, who?

Purpose?

Do you see a purpose? If so, what is it trying to inform, entertain, or persuade the audience to act? Or something else?

Narrative?

Does it use narrative? If so, what structure? Examples? Metaphors? Test with tools from past lectures.

Encoding, decoding?

What data is encoded? How? Any issues of perception in decoding?

Comparison or change?

Does the infographic describe comparisons or change? If so, what?

Color, coherency?

Is color used? If so, for what purpose(s) are its hue, chroma, or luminance used?

Hierarchy, annotation?

Does it have a hierarchy of information? If so, how is that hierarchy made? Are data encodings explained? If so, how?

Layering, layout?

Is the information organized? If so, how?

Credibility, transparency?

Are data sources identified, explained? Limitations or issues discussed?

Let's critique

Get specific

(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

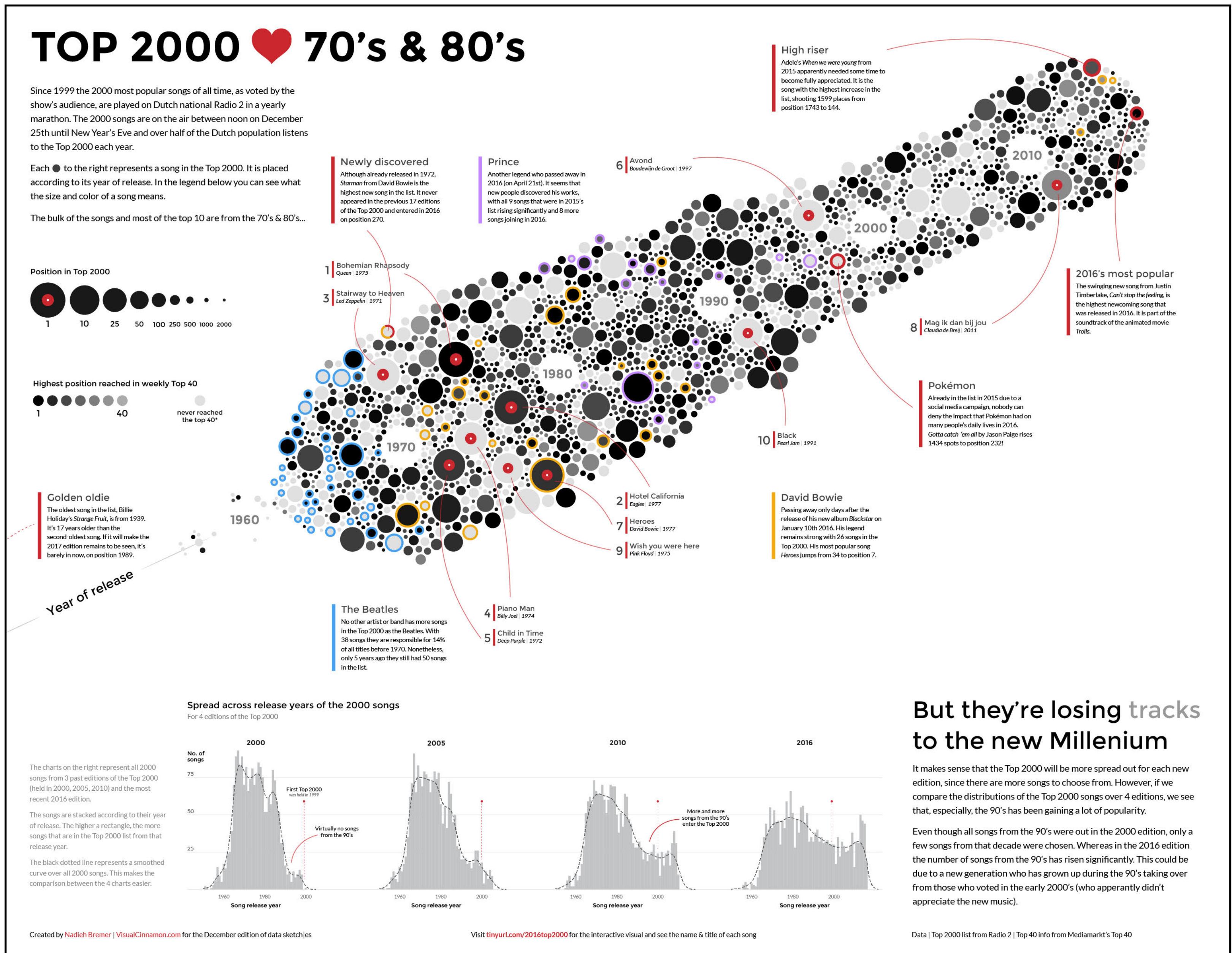
Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?



But they're losing tracks
to the new Millennium

It makes sense that the Top 2000 will be more spread out for each new edition, since there are more songs to choose from. However, if we compare the distributions of the Top 2000 songs over 4 editions, we see that, especially, the 90's has been gaining a lot of popularity.

Even though all songs from the 90's were out in the 2000 edition, only a few songs from that decade were chosen. Whereas in the 2016 edition the number of songs from the 90's has risen significantly. This could be due to a new generation who has grown up during the 90's taking over from those who voted in the early 2000's (who apparently didn't appreciate the new music).

(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

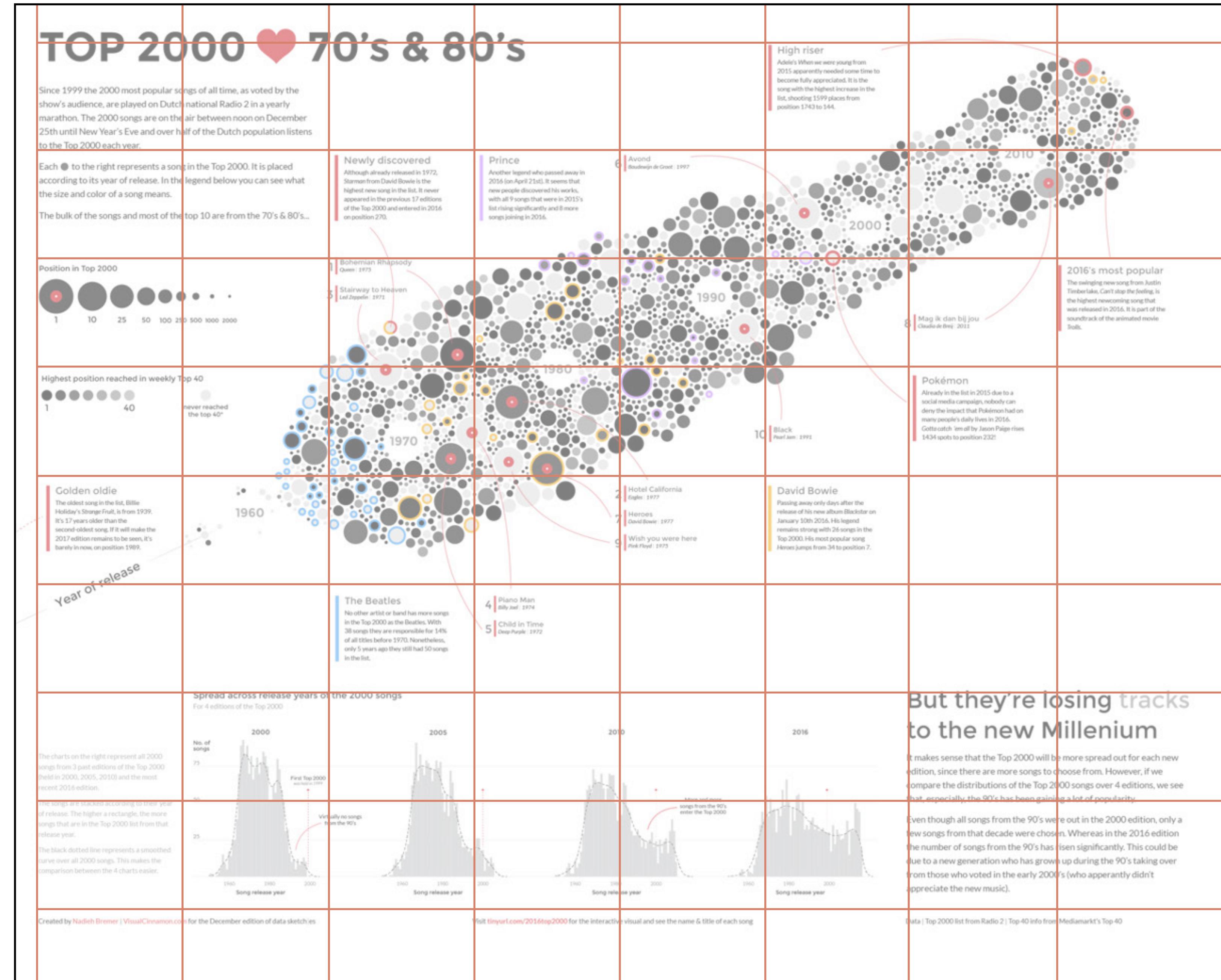
Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?



(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

Data encodings, decodings?

Comparison or change?

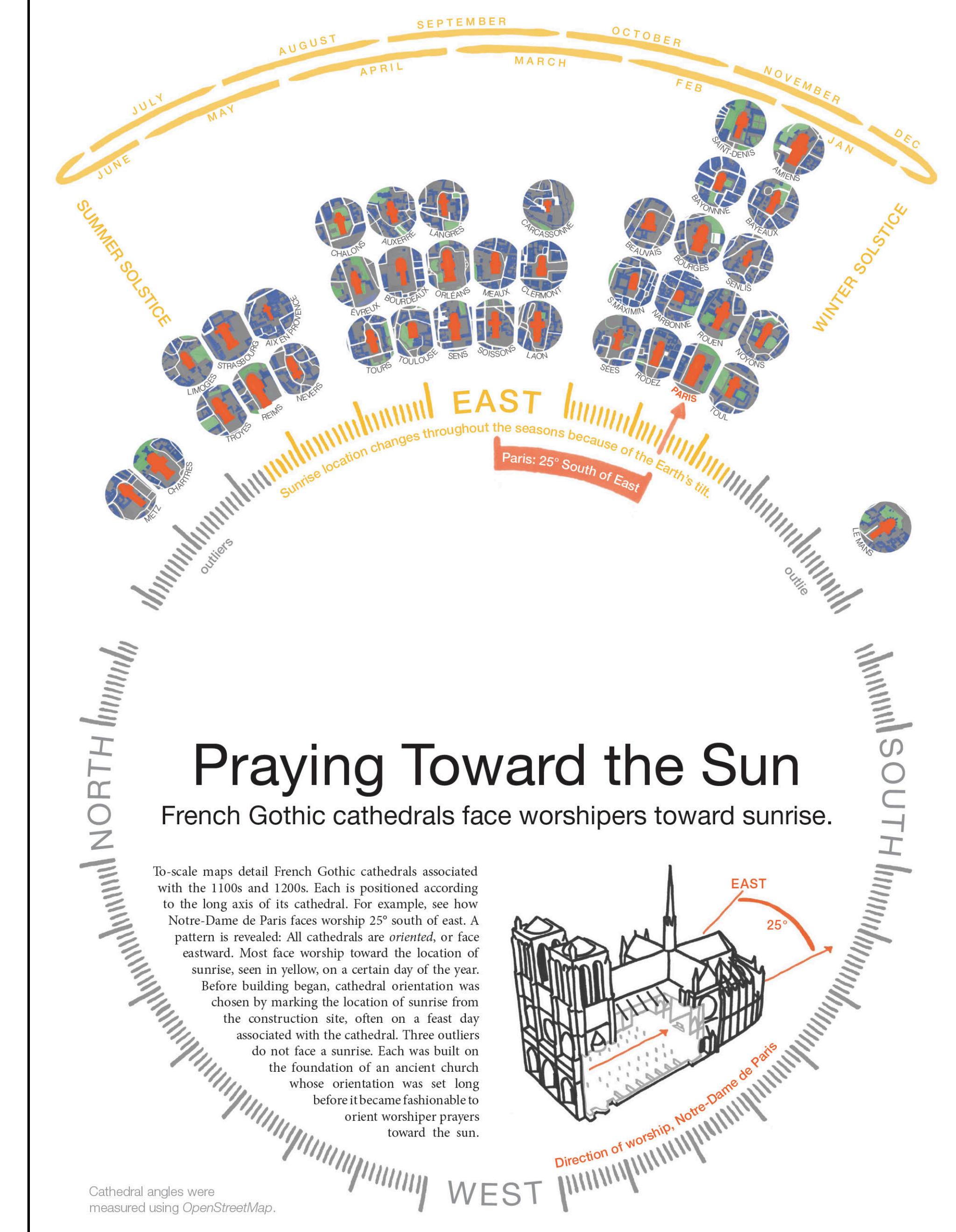
Color, coherency?

Layering, layout?

Credibility, transparency?



Andrews, R.J.



(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

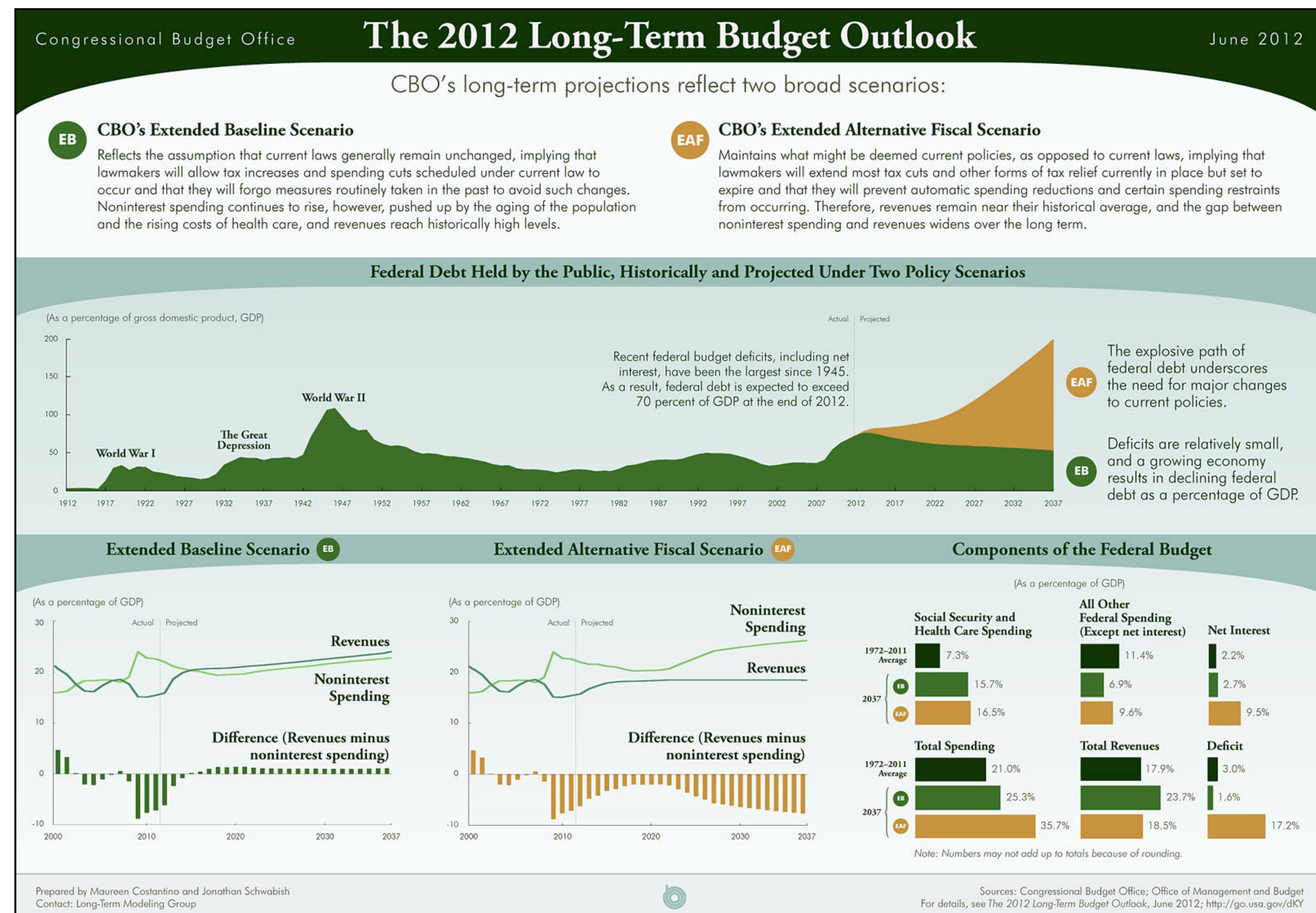
Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?



(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

Data encodings, decodings?

Comparison or change?

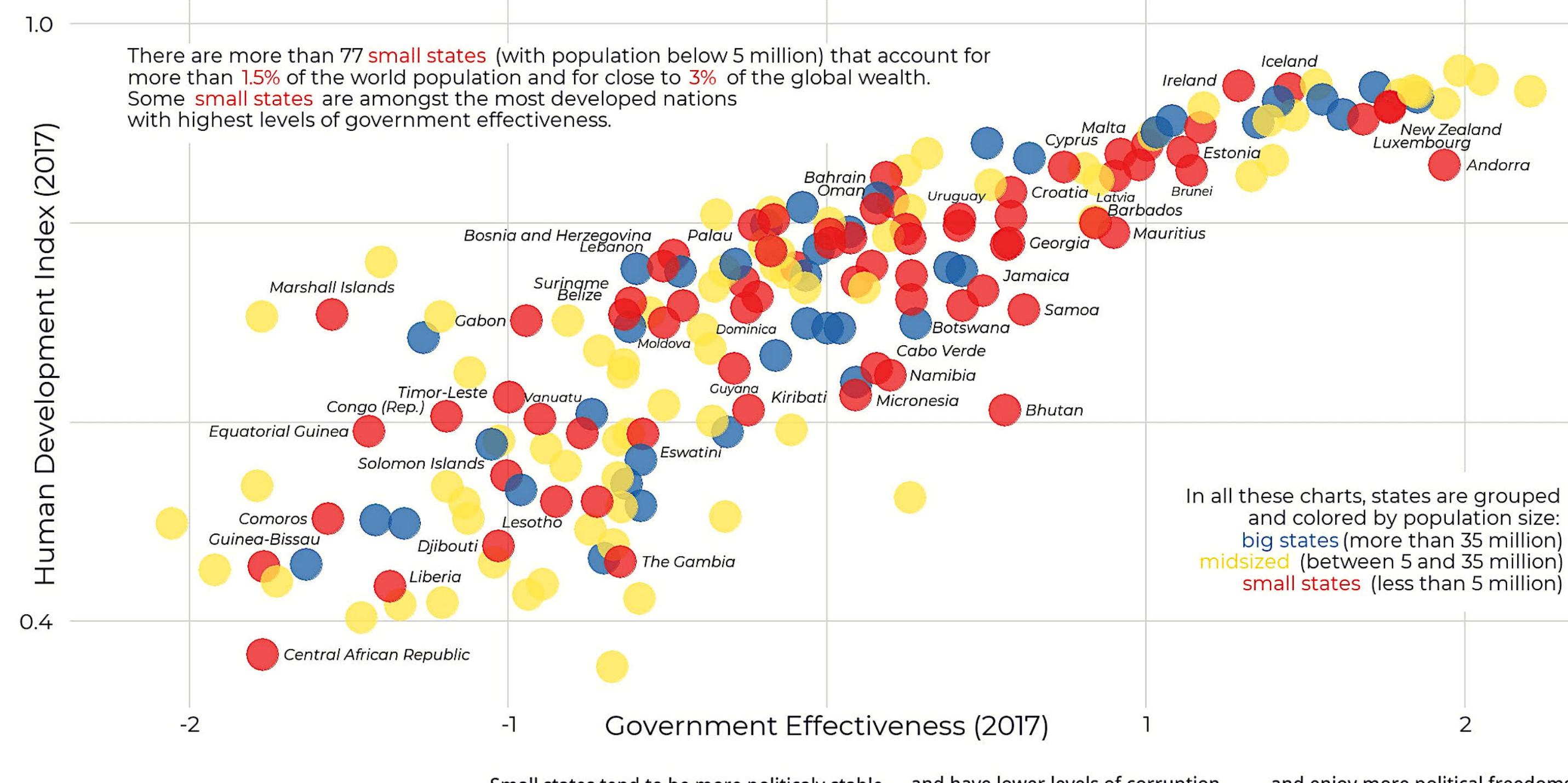
Color, coherency?

Layering, layout?

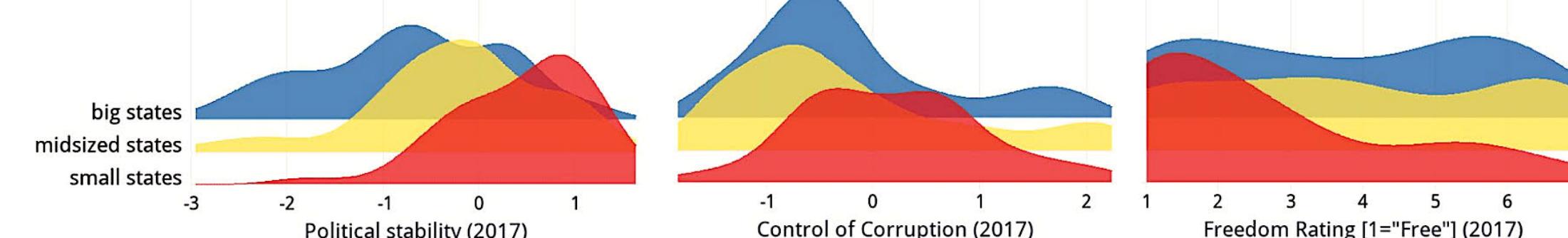
Credibility, transparency?



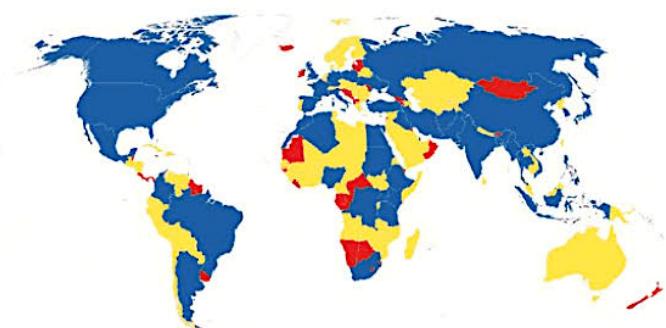
Small States Can Be Big Players in Development and Good Governance



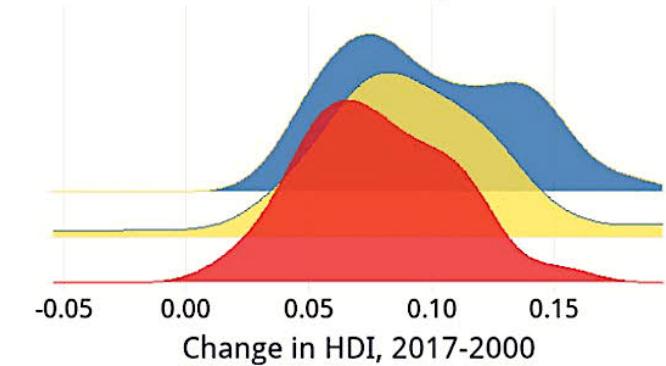
These charts show the distribution densities of different variables by three groups based on state size:



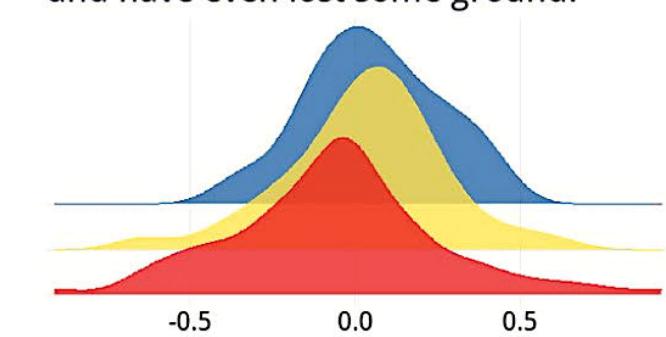
Data: Human Development Index from Human Development Reports. Government Effectiveness, Control of Corruption, and Political Stability from Worldwide Governance Indicators. Freedom Rating from Freedom House.



But small states have improved less



and have even lost some ground.



(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

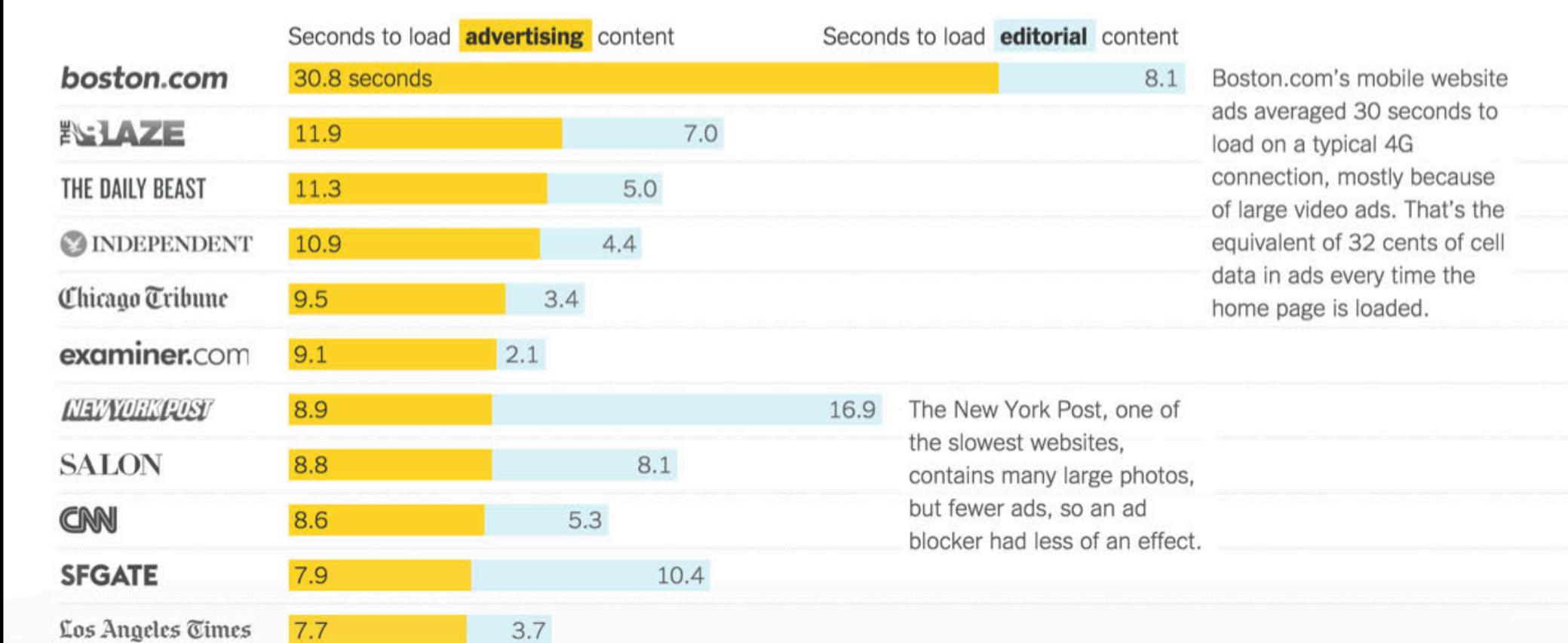
Credibility, transparency?



The Cost of Mobile Ads on 50 News Websites

By GREGOR AISCH, WILSON ANDREWS and JOSH KELLER OCT. 1, 2015

Ad blockers, which Apple first allowed on the iPhone in September, promise to conserve data and make websites load faster. But how much of your mobile data comes from advertising? We measured the mix of advertising and editorial on the mobile home pages of the top 50 news websites – including ours – and found that **more than half of all data came from ads** and other content filtered by ad blockers. Not all of the news websites were equal. [RELATED ARTICLE](#)

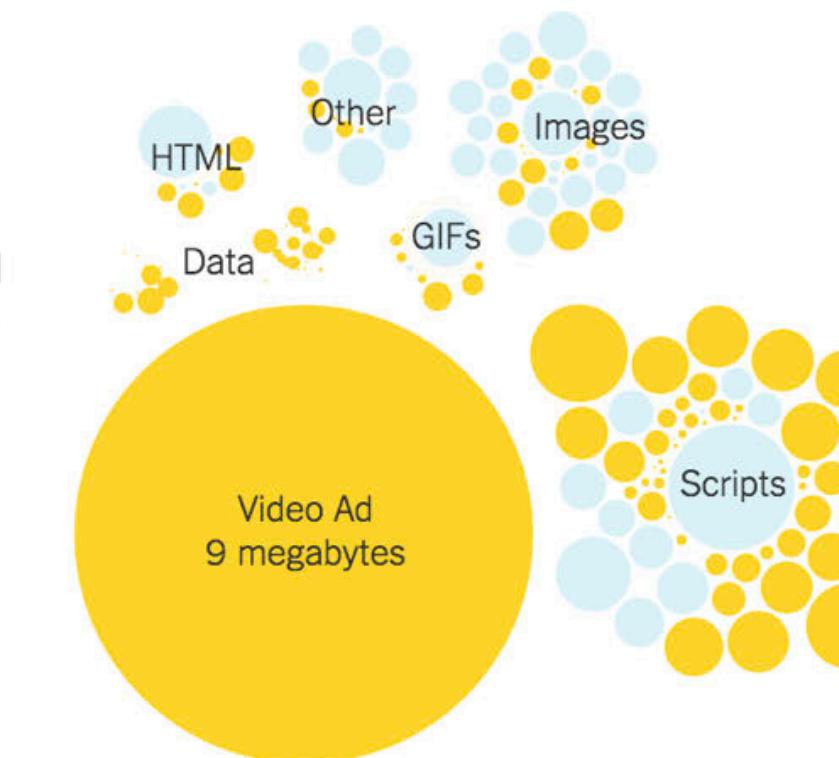


boston.com

Here are all the files that made up the Boston.com data during one visit, including one large video ad and many script files used by ad networks. With an ad blocker, those files were gone.

Without ad blocker

389 files, 16.3 megabytes, 33 seconds



With ad blocker

52 files, 3.5 megabytes, 7 seconds



(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?



For the past six springs, New Yorkers pedaled past colorful blossoms on their way to work, home, or just cruising.

Yet, some cruisers wearied whilst scouting a docking station []. And some on foot languished curbside without saddle to straddle.

Empty and full docking stations sprout like dandelions under the sun and moon, shown in 10 minute increments. Availability waxes and wanes by time and place.

A ride against the flow is a joy ride for us all.



Ride against the flow

Data from 2019 January 12 - 31, sourced from NYC Open Data, and The Open Bus project, <https://www.theopenbus.com/raw-data.html>
Analysis and design by Scott Spencer. April 2019. <https://ssp3nc3r.github.io>

(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

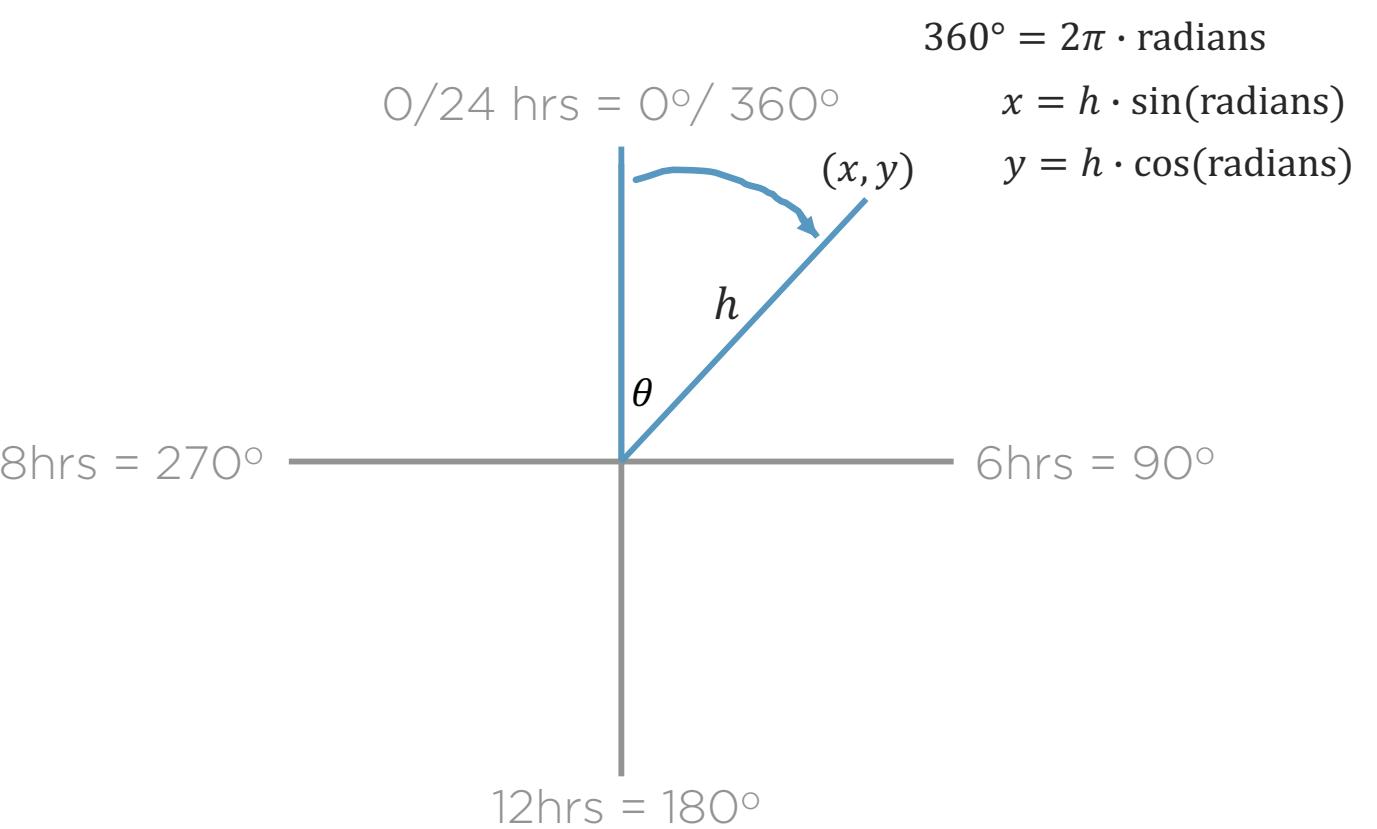
Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?

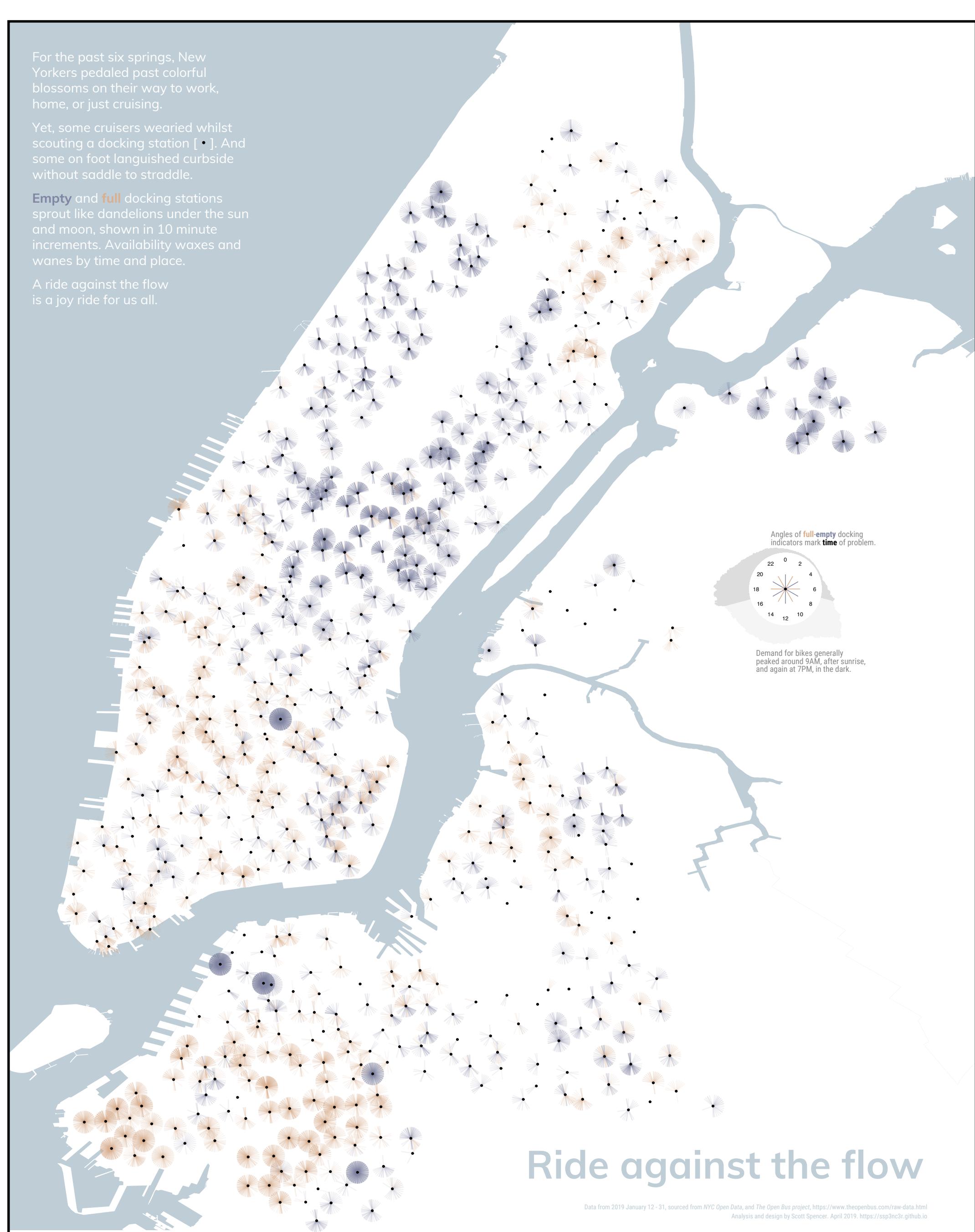


For the past six springs, New Yorkers pedaled past colorful blossoms on their way to work, home, or just cruising.

Yet, some cruisers wearied whilst scouting a docking station []. And some on foot languished curbside without saddle to straddle.

Empty and full docking stations sprout like dandelions under the sun and moon, shown in 10 minute increments. Availability waxes and wanes by time and place.

A ride against the flow is a joy ride for us all.



Ride against the flow

Data from 2019 January 12 - 31, sourced from NYC Open Data, and The Open Bus project, <https://www.theopenbus.com/raw-data.html>
Analysis and design by Scott Spencer. April 2019. <https://ssp3nc3r.github.io>

(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

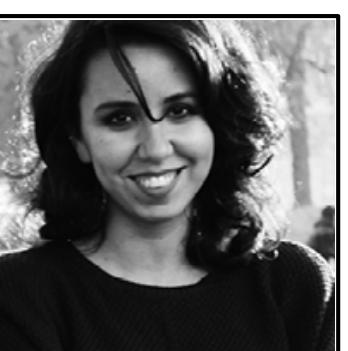
Data encodings, decodings?

Comparison or change?

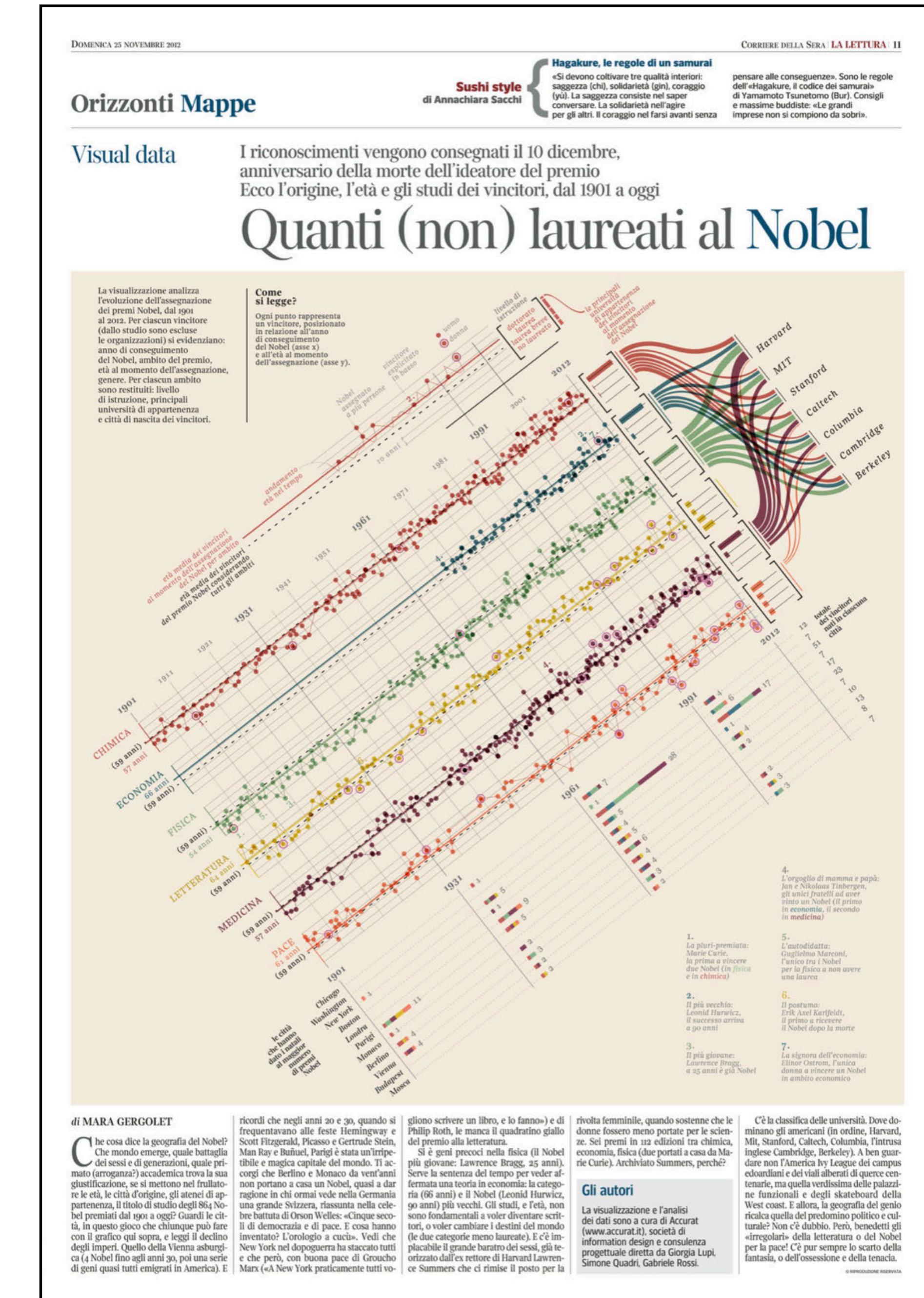
Color, coherency?

Layering, layout?

Credibility, transparency?



Lipi, Giorgio. Fragapane, Federica.



(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

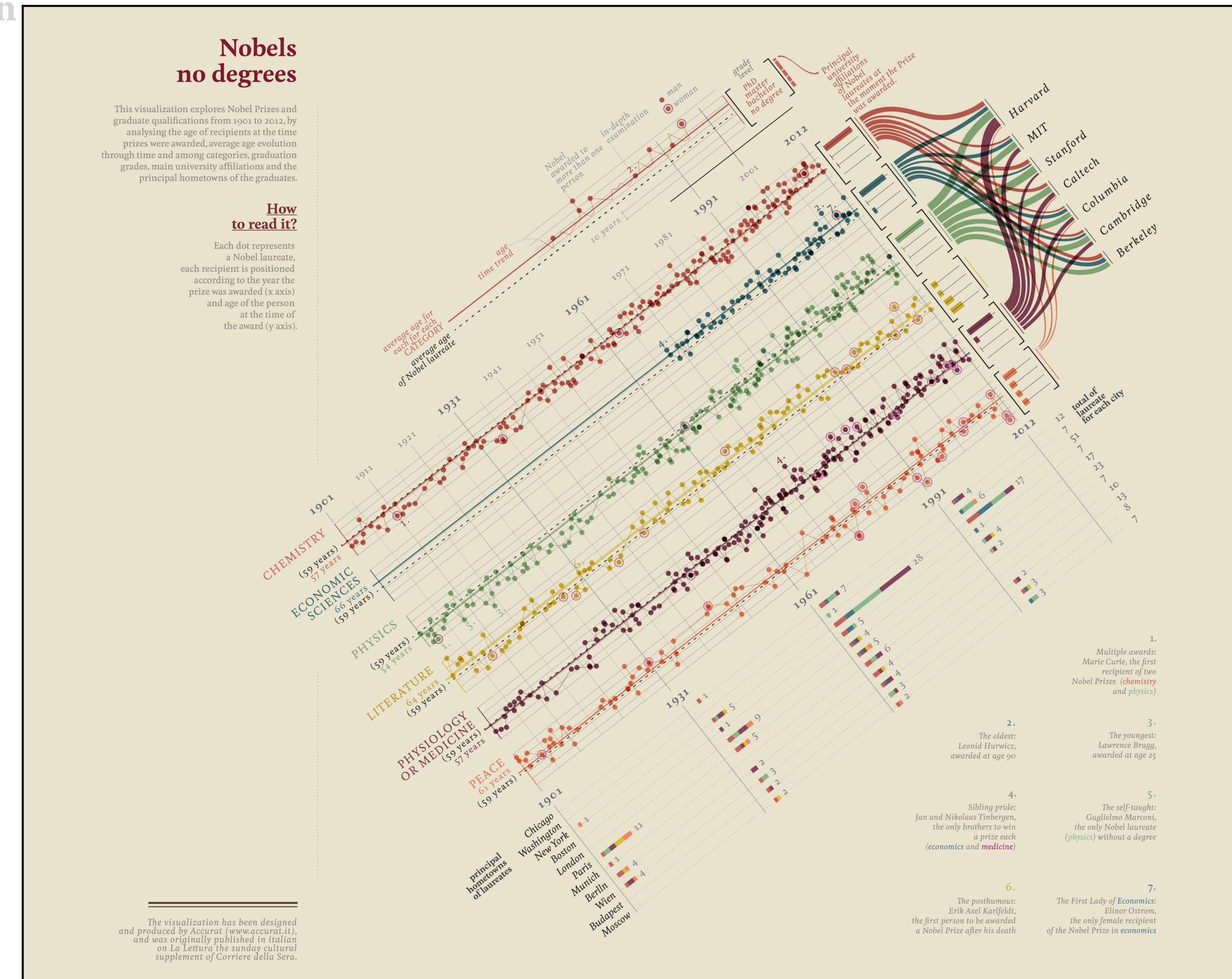
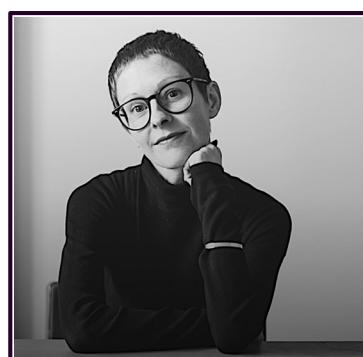
Data encodings, decodings?

Comparison or change?

Color, coherency?

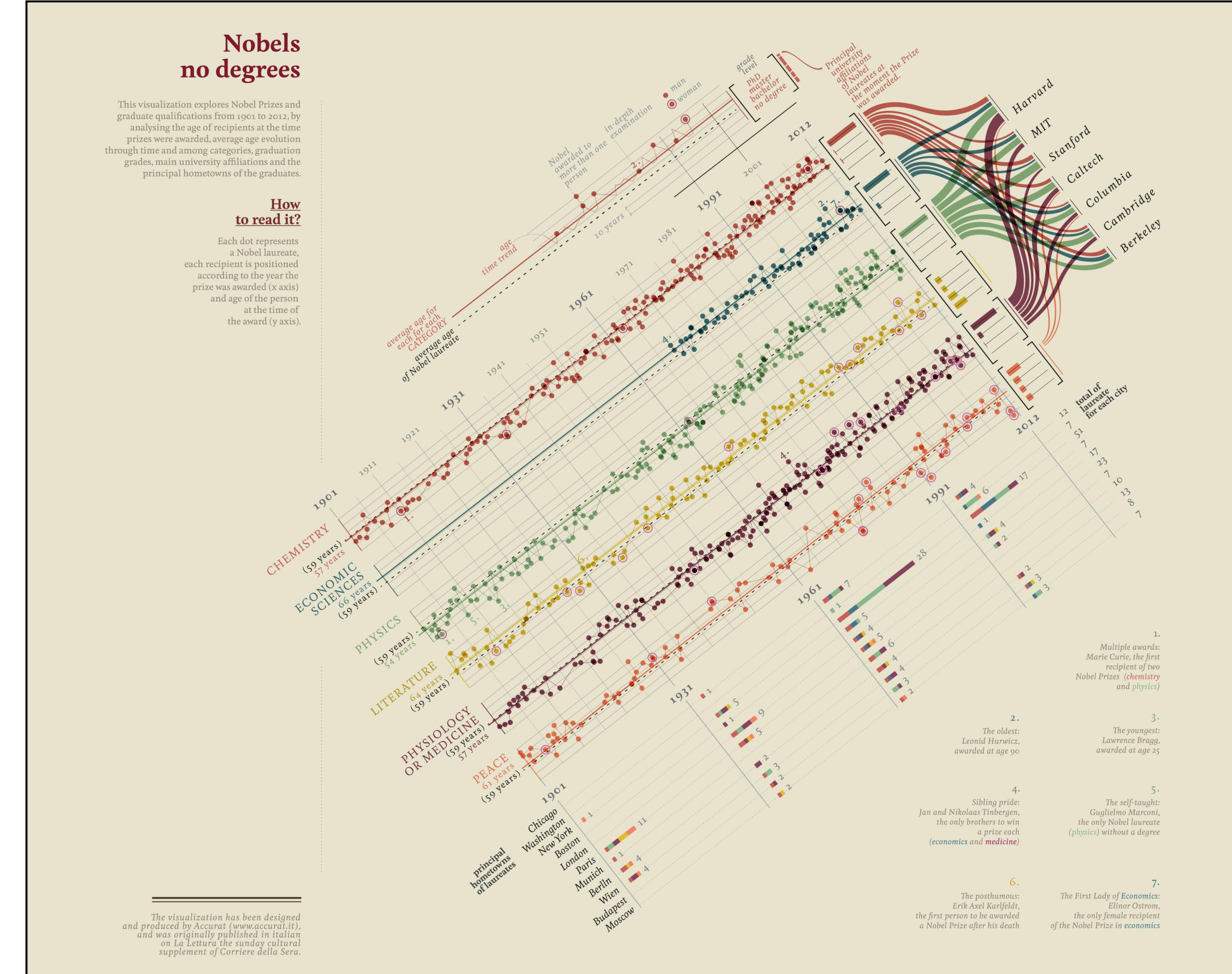
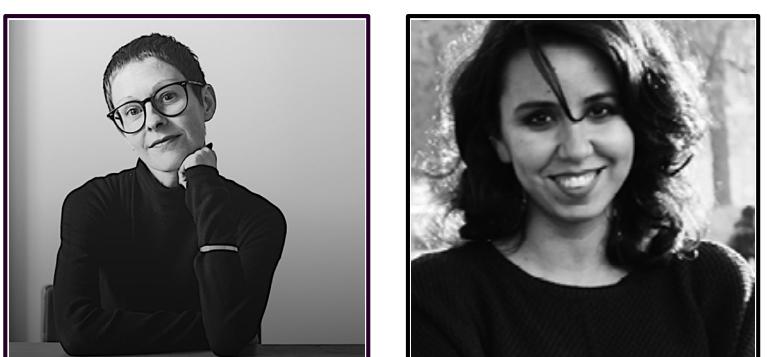
Layering, layout?

Credibility, transparency?



practice in the studio

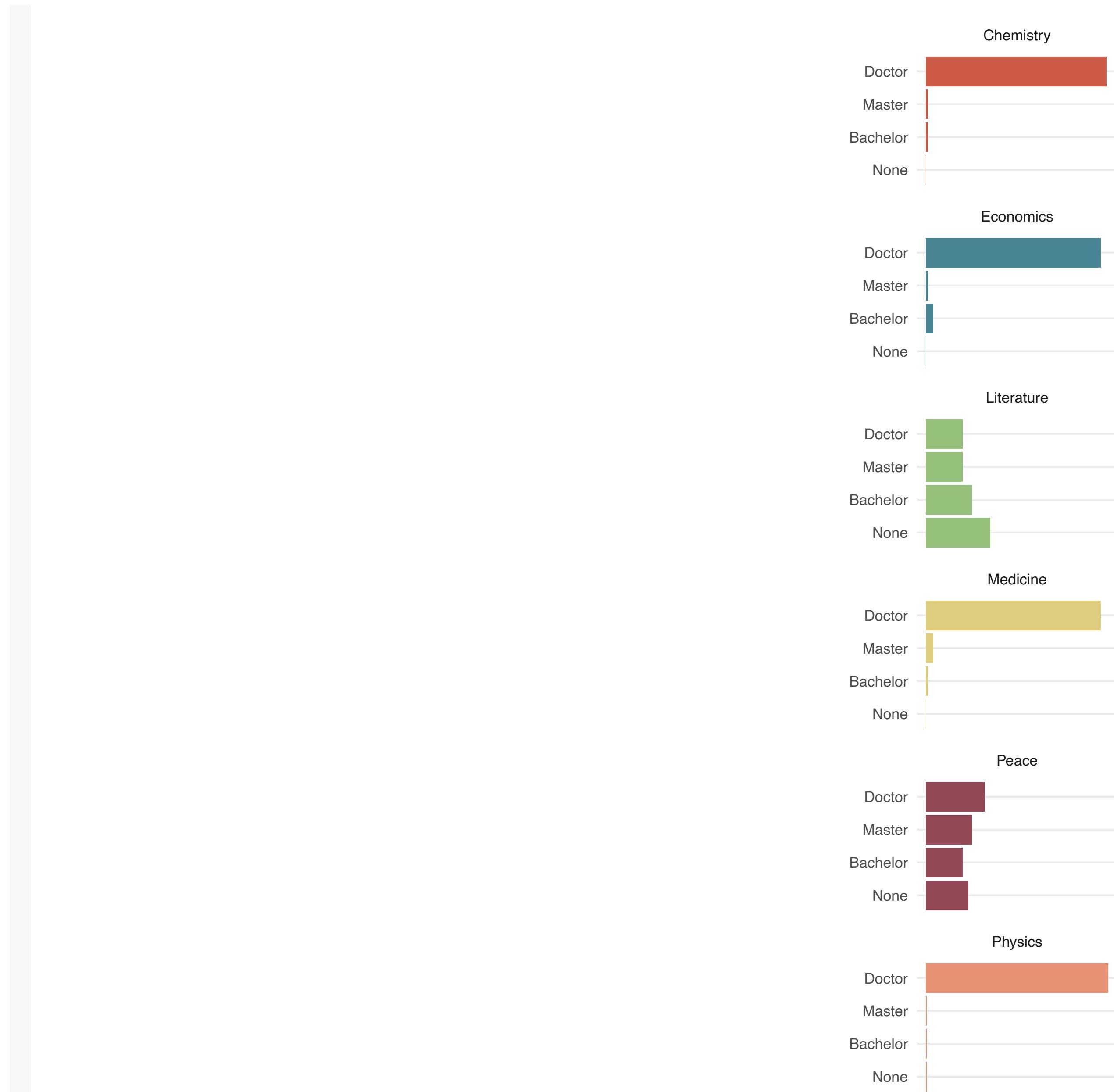
reconstructing an infographic



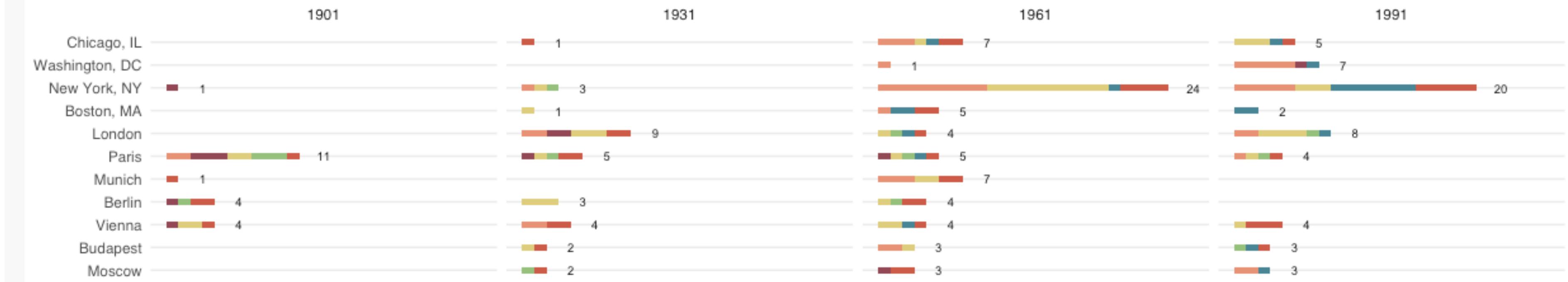
approximating the components | *encoding prizes by color, type, and points and lines by time and age*



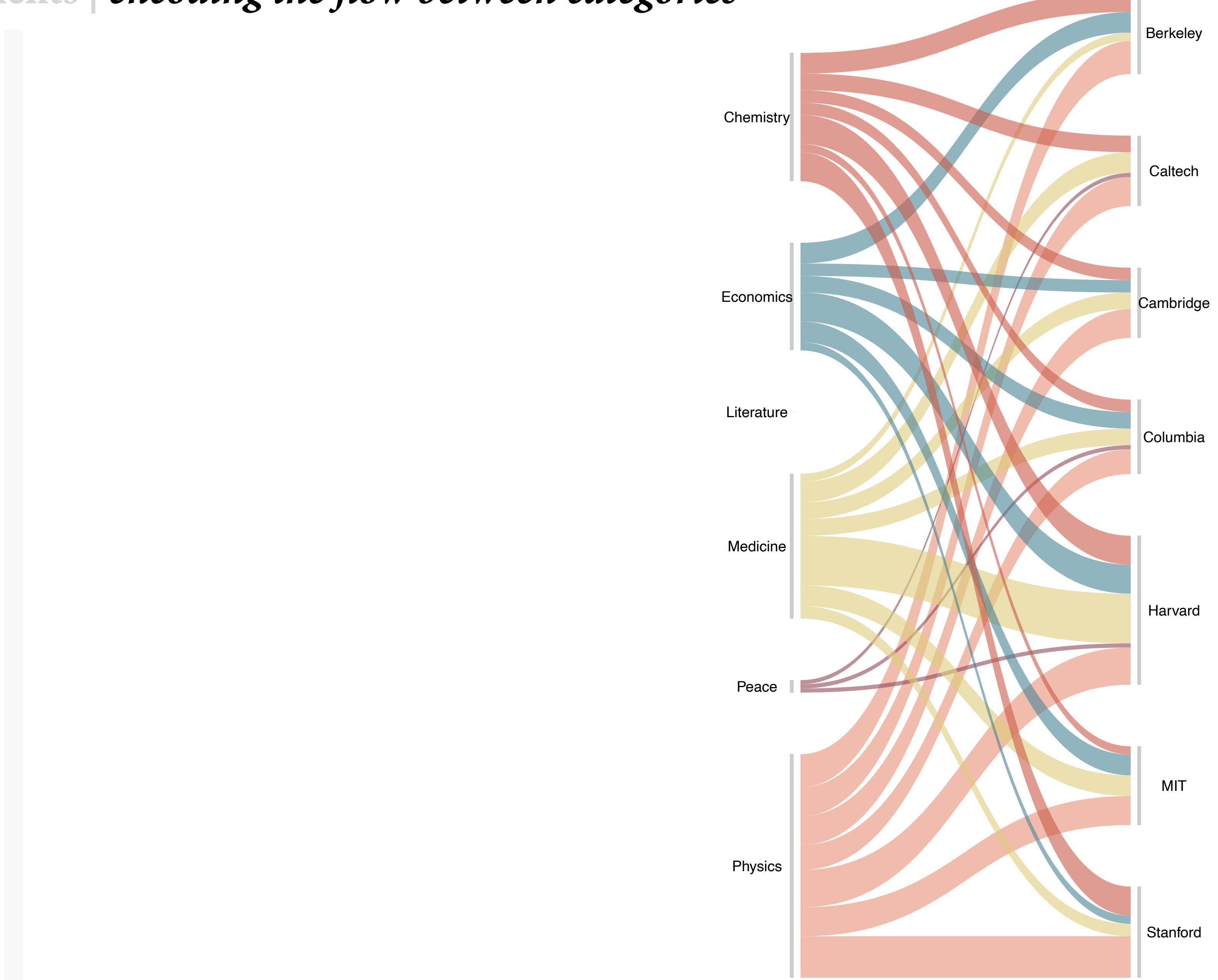
approximating the components | *encoding prizes by color, type, and histogram length by education type*



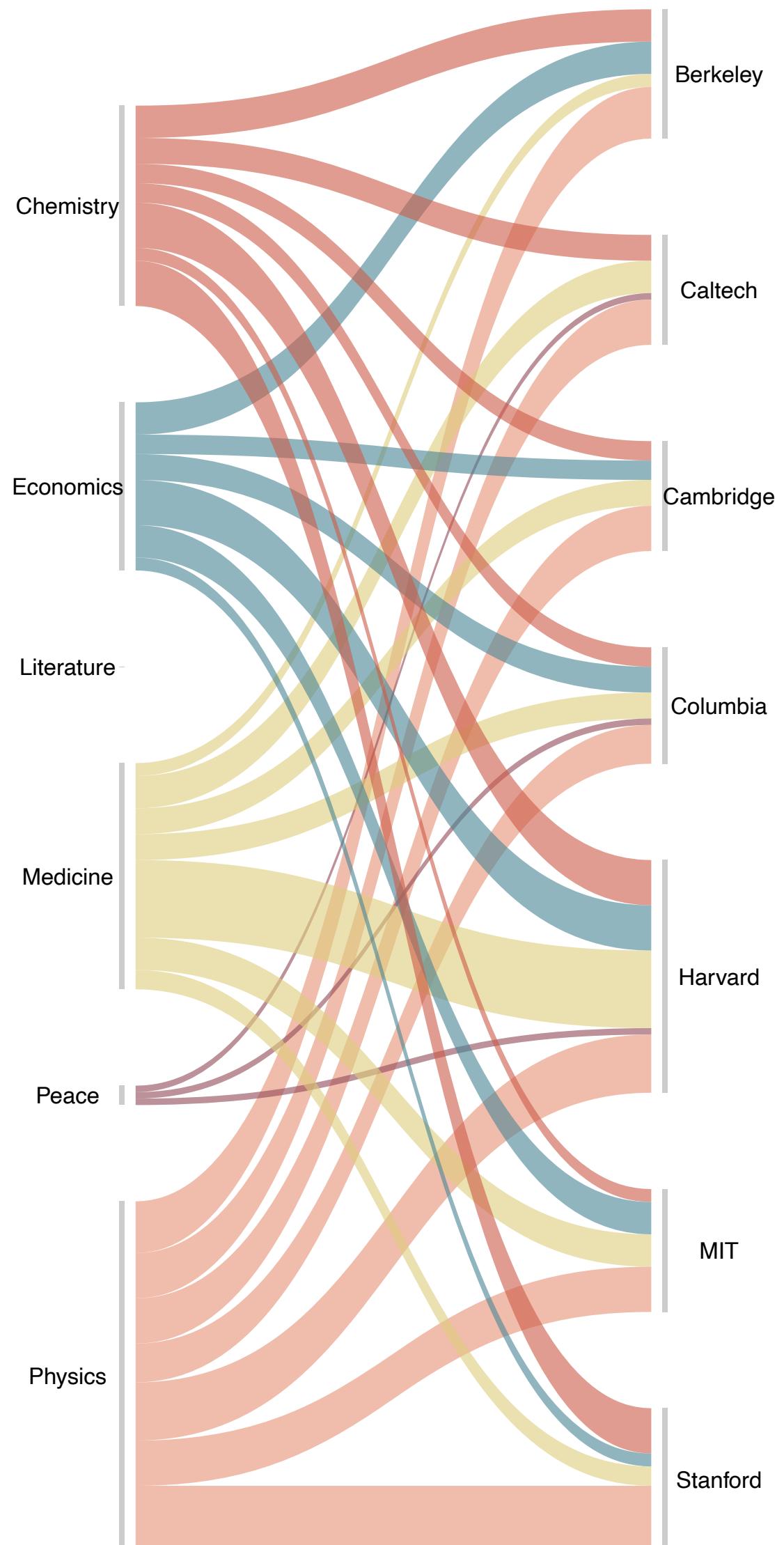
approximating the components | *encoding prizes by color, decade, and stacked bar length by city*



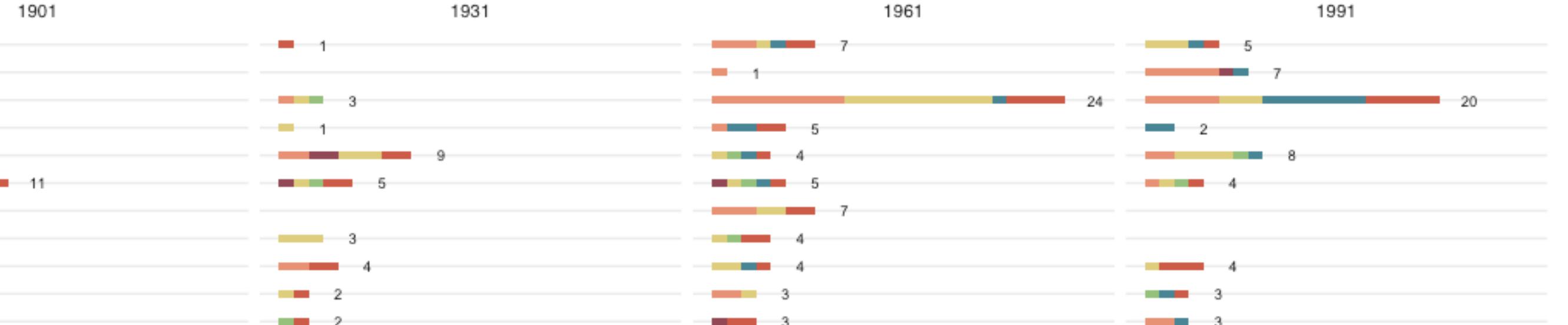
approximating the components | *encoding the flow between categories*



approximating the components | *arranging components and aligning axes*



Chicago, IL
Washington, DC
New York, NY
Boston, MA
London
Paris
Munich
Berlin
Vienna
Budapest
Moscow



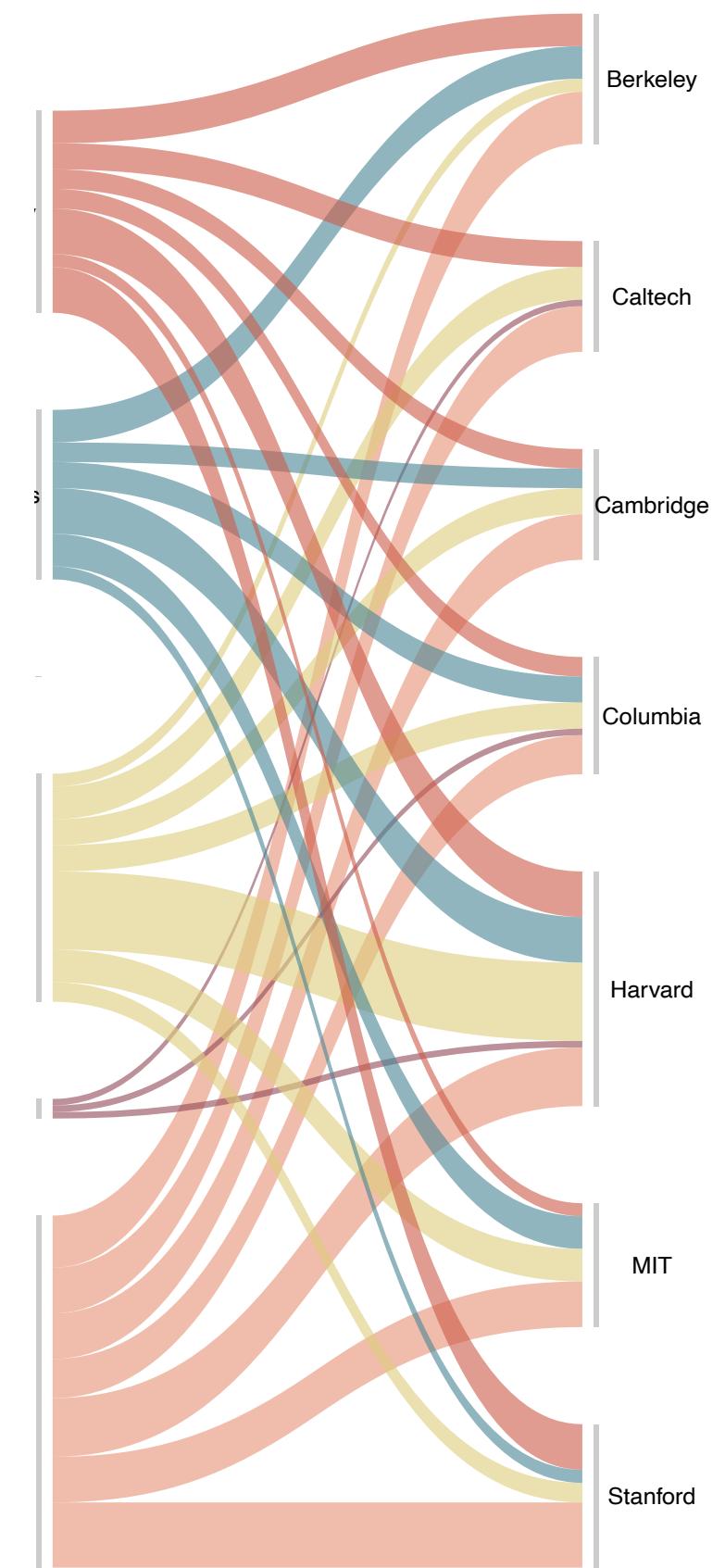
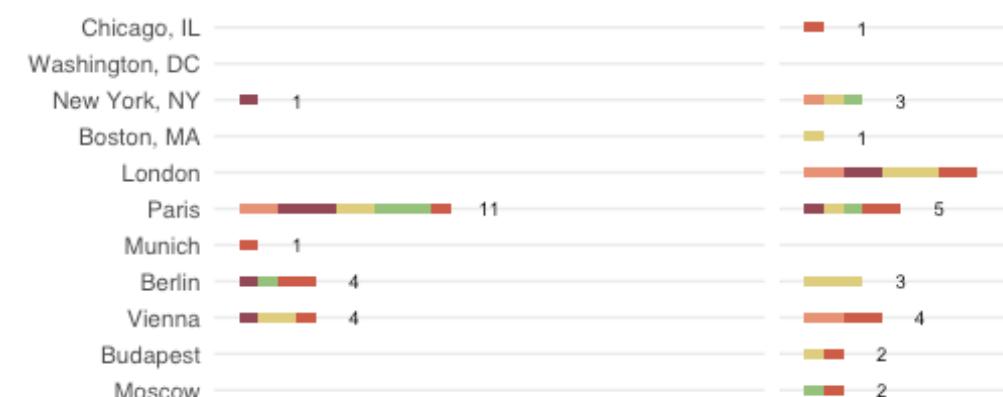
approximating the components | *layering in the remaining annotations, linking to graphics*

Nobels, no degrees

This visualization explores Nobel Prizes and graduate qualifications from 1901 to 2012, by analyzing the age of recipients at the time prizes were awarded, average age evolution through time and among categories, graduation grades, main university affiliations and the principal hometowns of the graduates.

How to read it?

Each dot represents a Nobel laureate, each recipient is positioned according to the year the prize was awarded (x axis) and age of the person at the time of the award (y axis).



1. Multiple awards:
Marie Curie, the first
recipient of two Nobel
Prizes (**chemistry** and
physics)

2.
The oldest:
Leonid Hurwicz,
awarded at age 90

3.
The youngest:
Lawrence Bragg,
awarded at age 25

4.
Sibling pride:
Jan and Nikolaas Tinbergen,
the only brothers to win a prize
each (**economics** and **medicine**)

5.
The self-taught:
Guglielmo Marconi,
the only Nobel laureate
(**physics**) without a degree

6.
The First Lady of **Economics**:
Elinor Ostrom,
the only female recipient
of the Nobel Prize in **economics**

help from colleagues

uncertainty

model specifications and selections

(Do the models (parameters, data, functions) represent the underlying process intended for inference and account for data collection?)

estimations in model parameters

(parameters represent variation in observations, measurement error, etc)

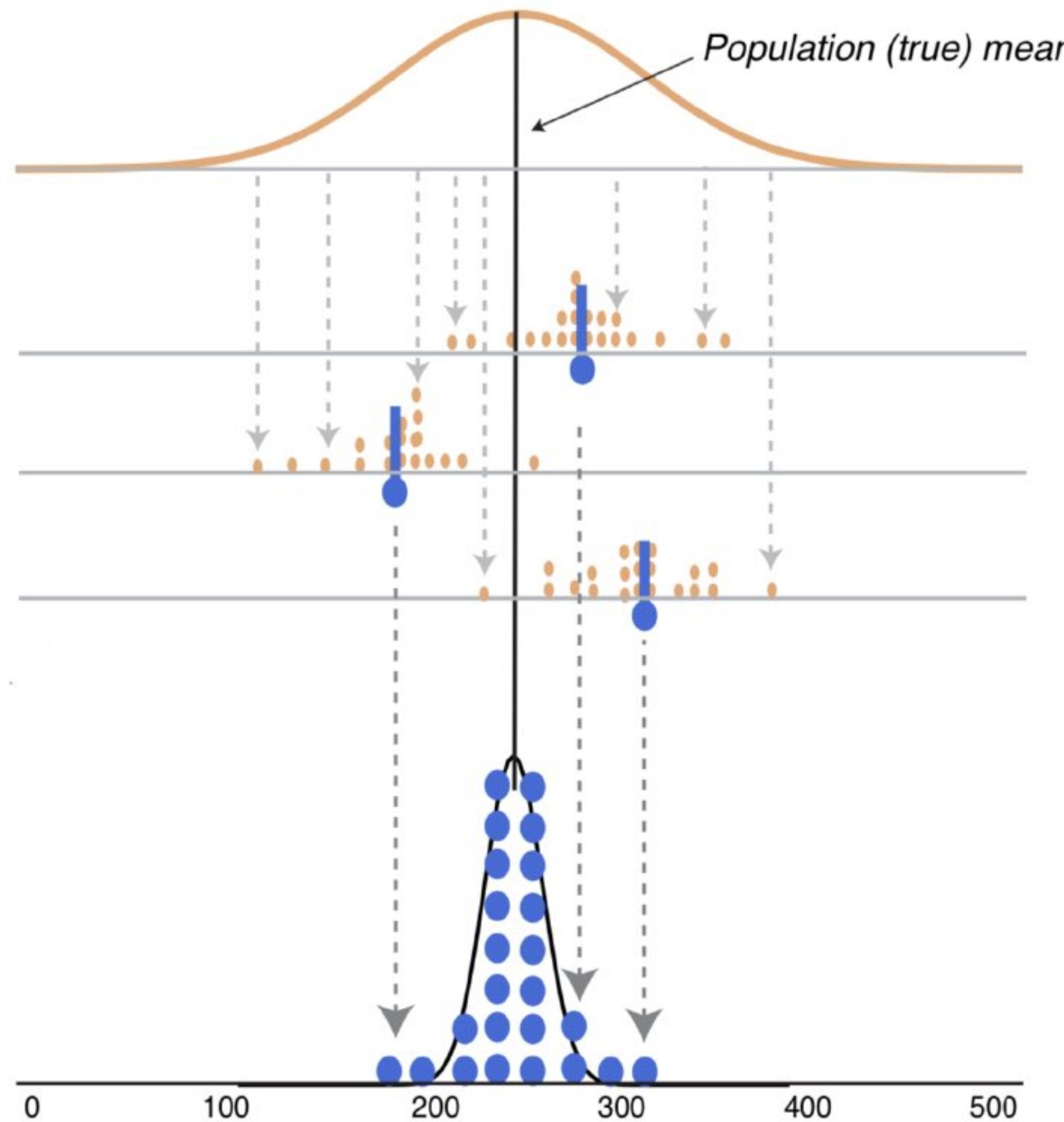
whether computations work as intended

(e.g., calculation overflows, underflows, coding mistakes)

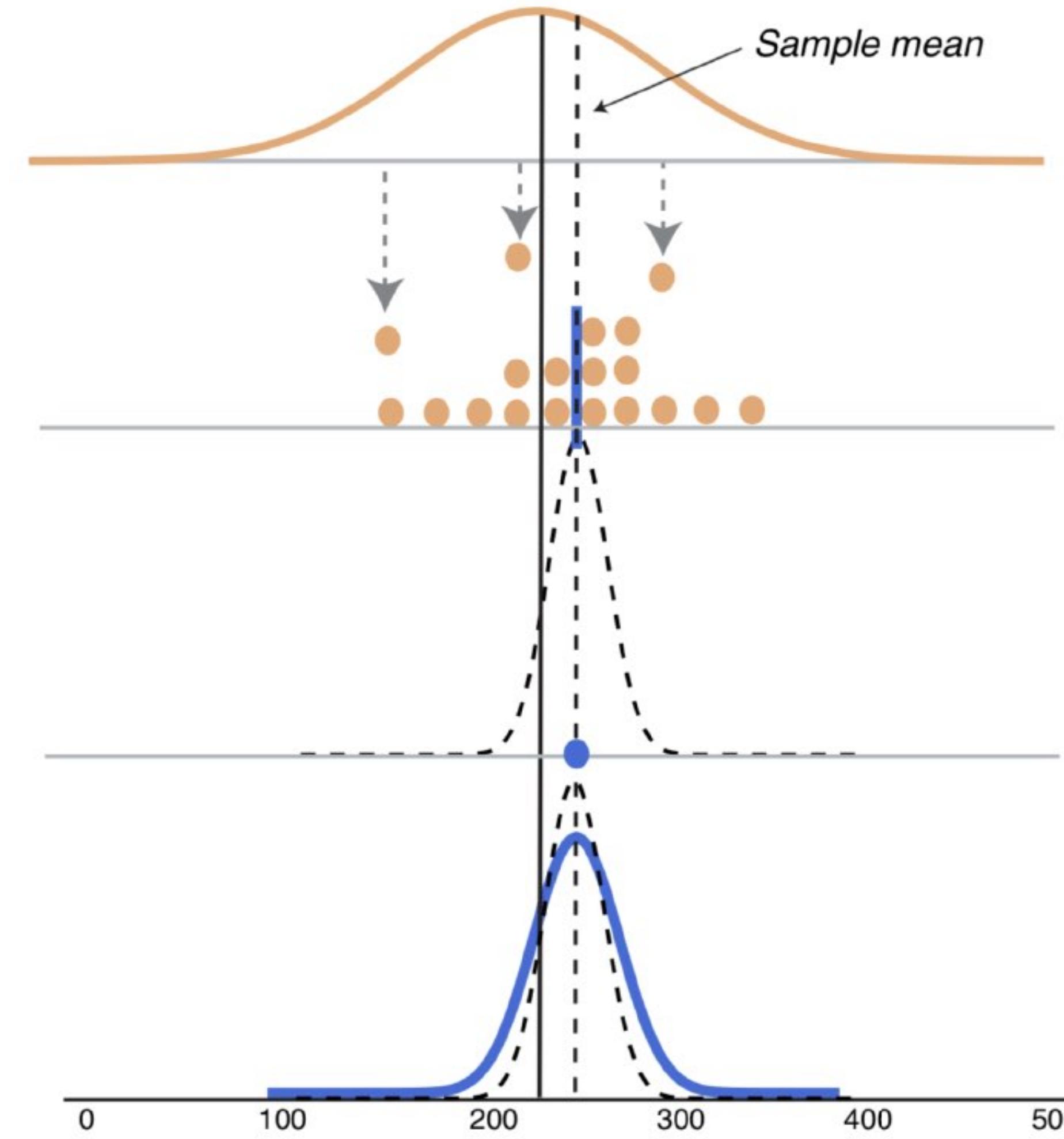
decisions from model outputs

(look to decision theory, utility functions)

Perfect Knowledge of World



Imperfect Knowledge of World (Experimental Paradigm)



uncertainty | *overcoming concerns with communicating uncertainty*

Concern | people will misinterpret quantities of uncertainty, inferring more precision than intended.

Response | Most people like getting quantitative information on uncertainty, from them can get the main message, and without them are more likely to misinterpret verbal expressions of uncertainty. Posing clear questions guide understanding.



Fischhoff, Baruch

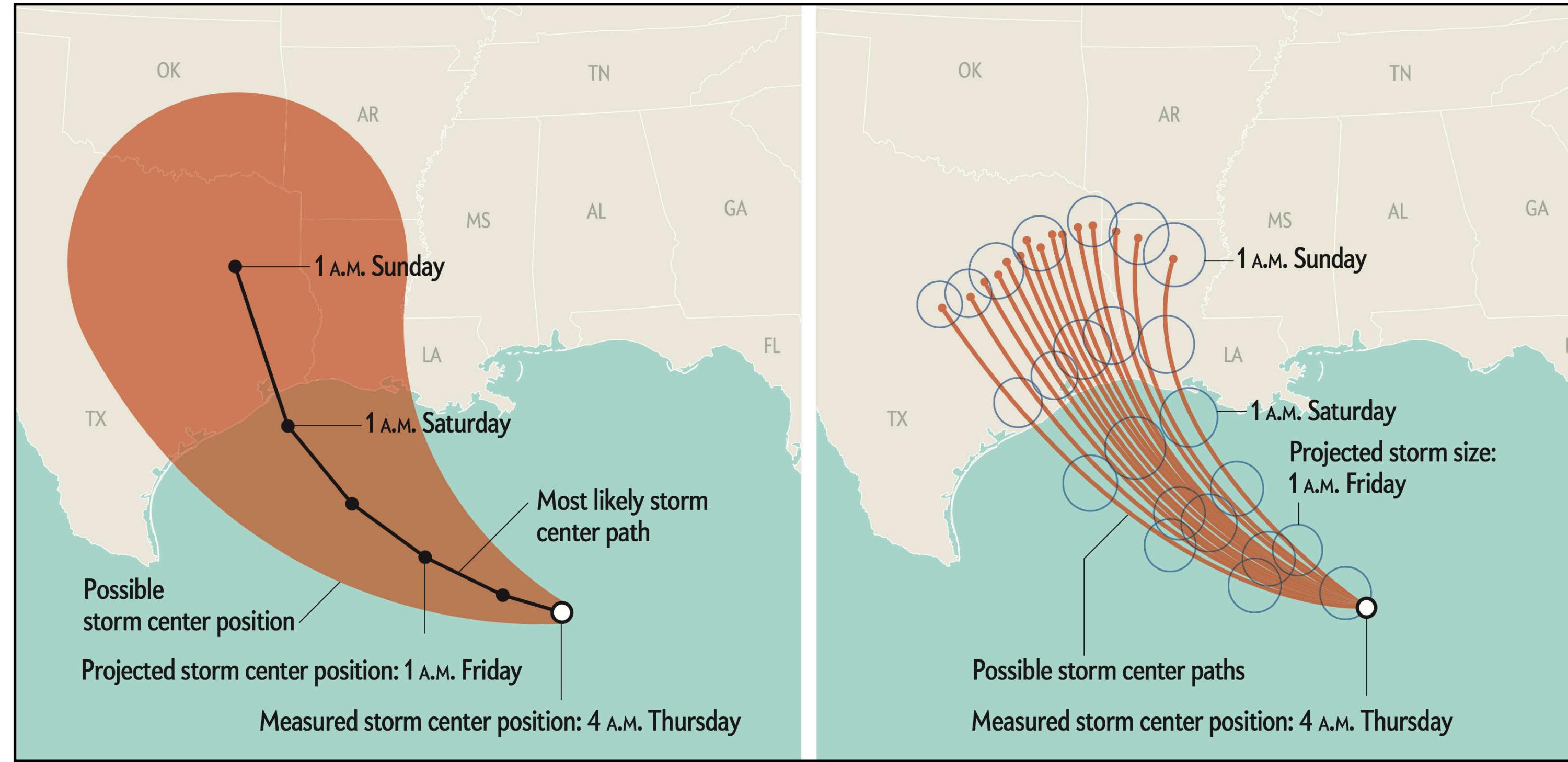
Concern | people cannot use probabilities.

Response | laypeople can provide high-quality probability judgments, if they are asked clear questions and given the chance to reflect on them. Communicating uncertainty protects credibility.

Concern | credible intervals may be used unfairly in performance evaluations.

Response | probability judgments give us more accuracy about the information; *i.e.*, won't be too confident or lack enough confidence.

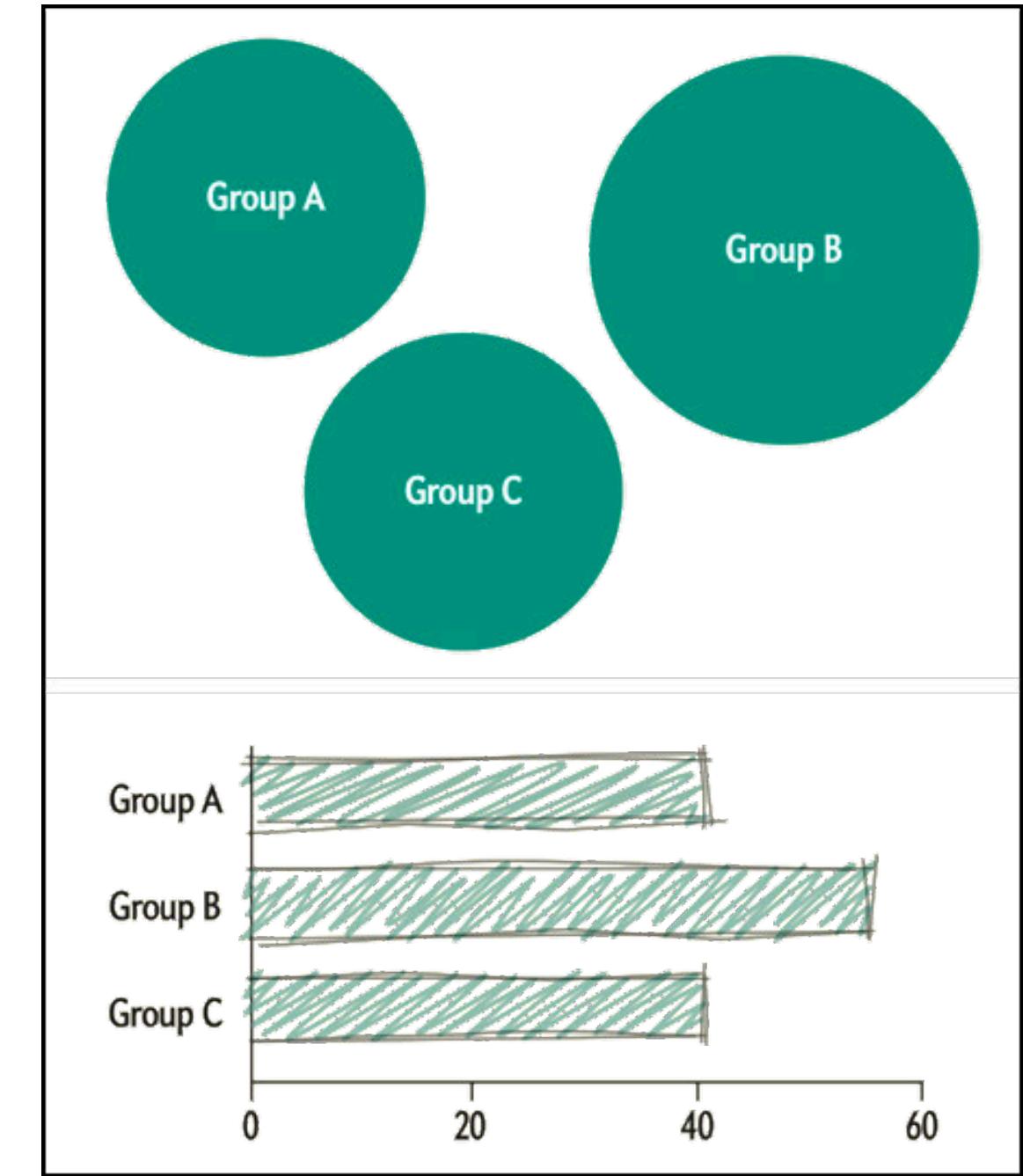
uncertainty | graphically encoding uncertainty



Hullman, Jessica

uncertainty | *graphically encoding uncertainty*

no quantification



Hullman, Jessica

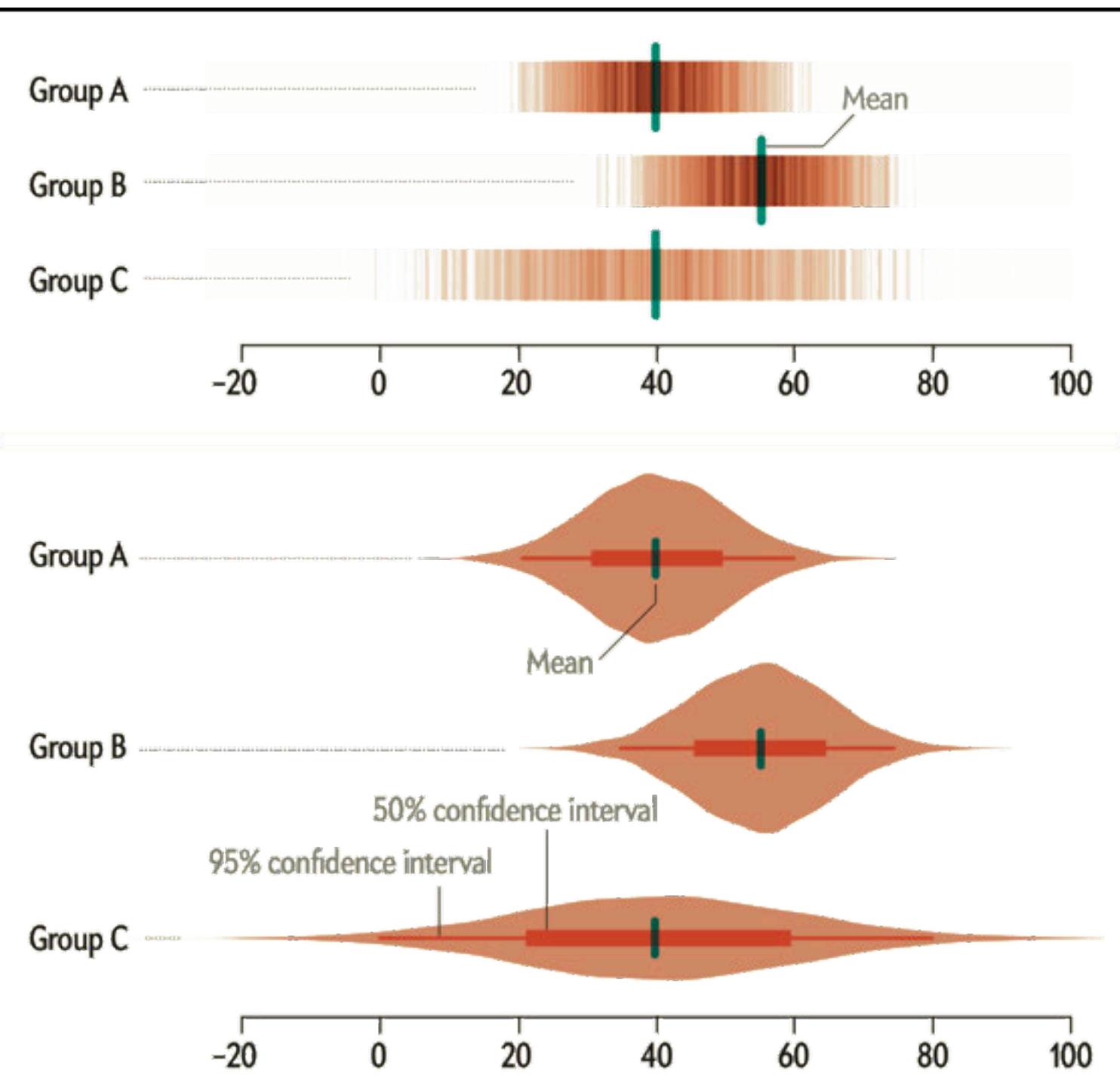
uncertainty | graphically encoding uncertainty

intervals



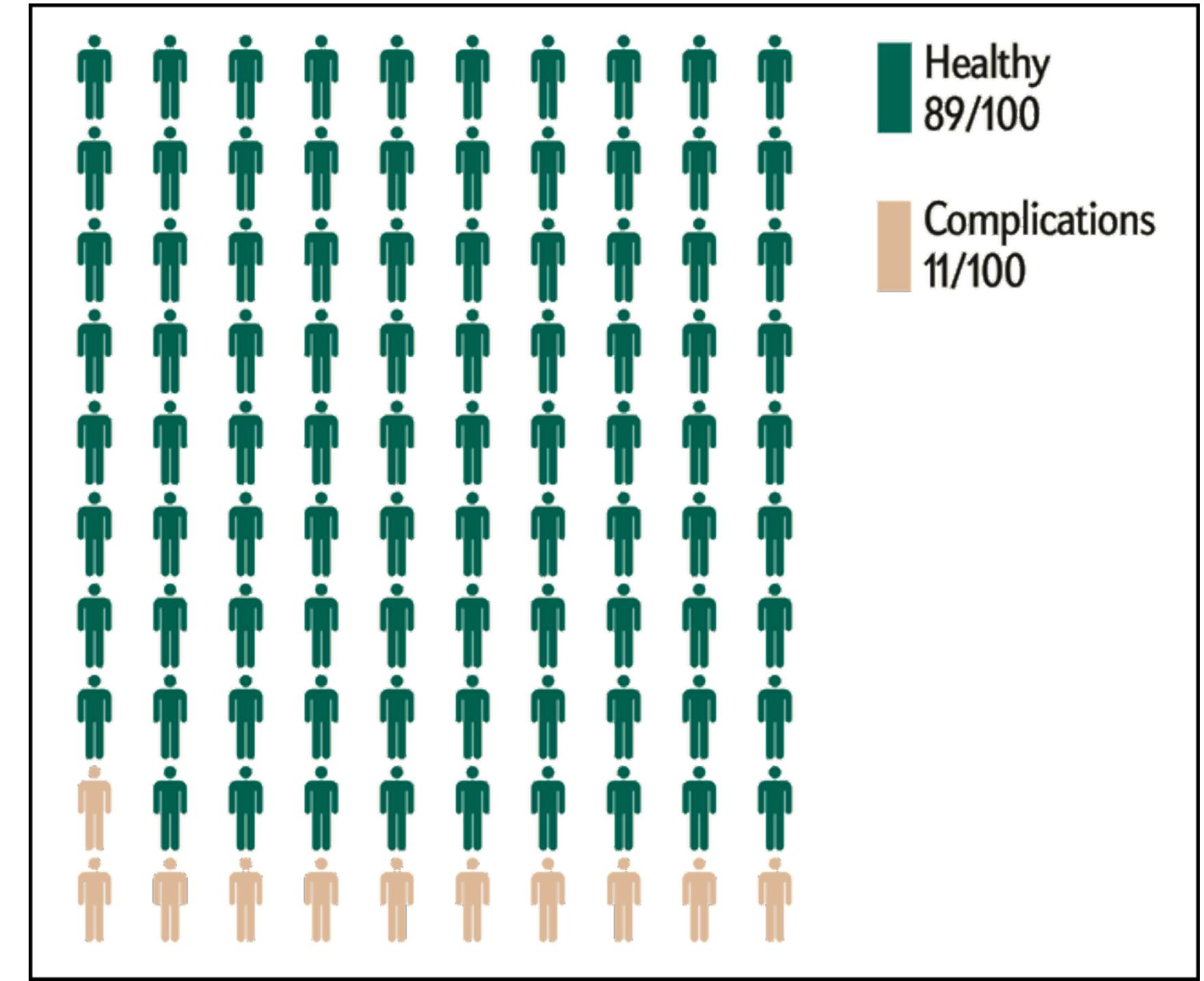
uncertainty | graphically encoding uncertainty

probability density

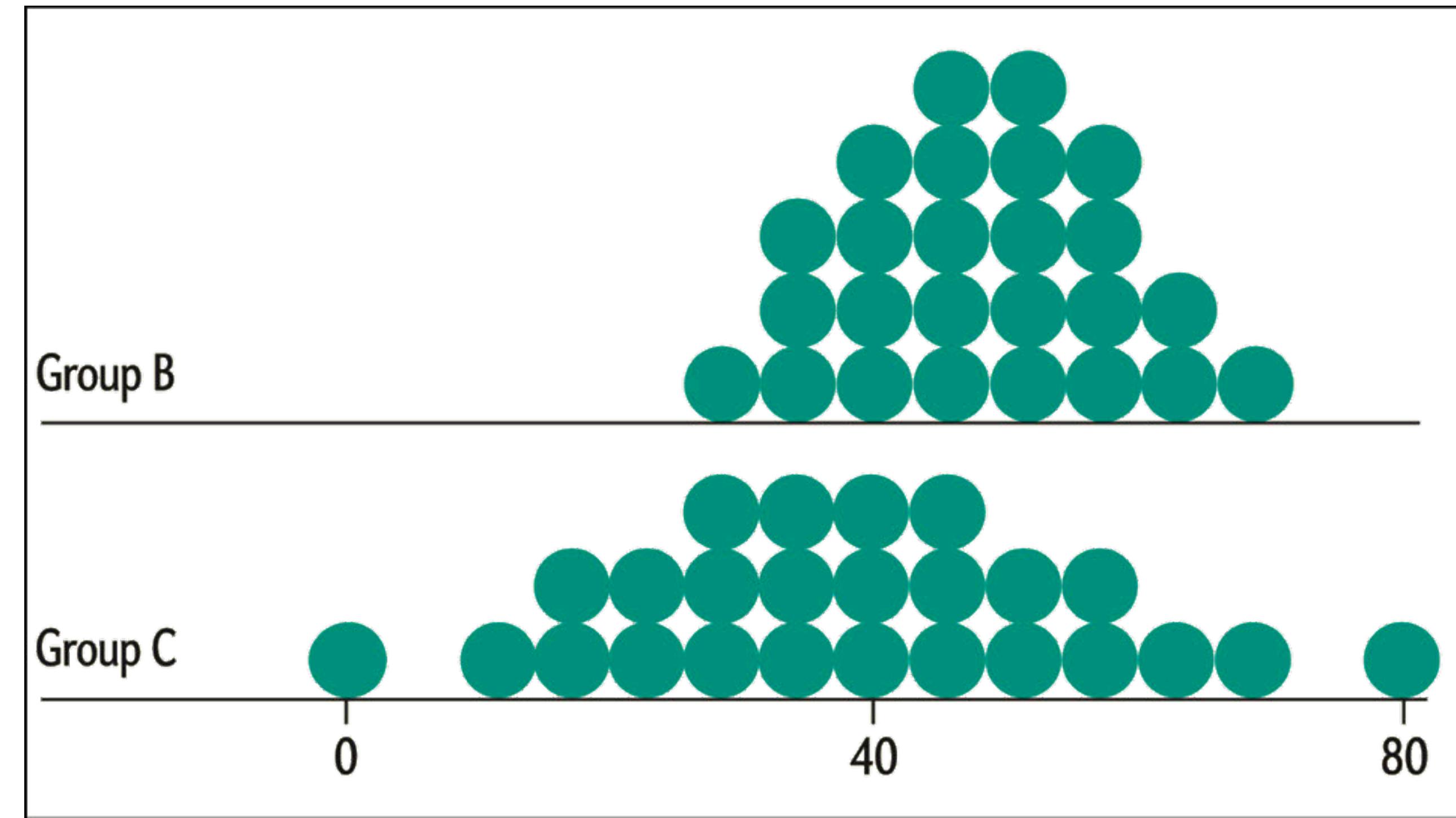


uncertainty | graphically encoding uncertainty

arrays of icons (isotypes)



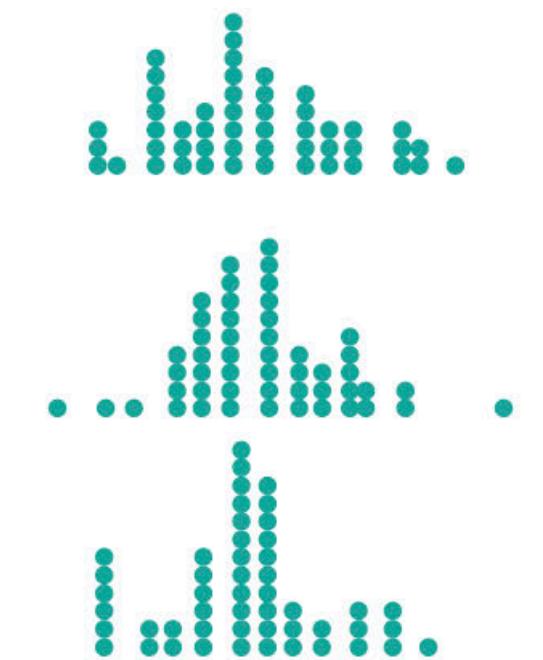
multiple samples in space (dot plots)



Hullman, Jessica

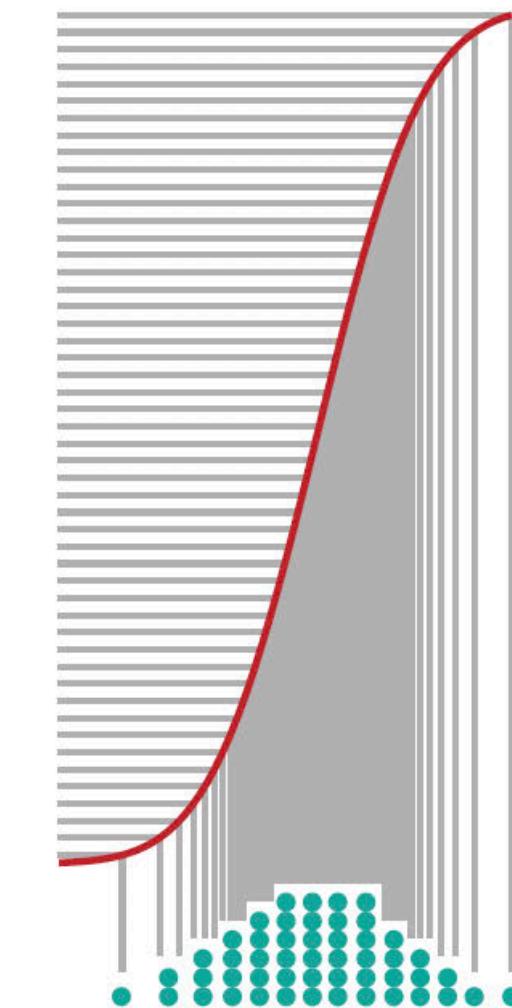


To generate a discrete plot of this distribution, we could try taking **random draws** from it. However, **this approach is noisy**: it may be very different from one instance to the next.



quantile dot plots

Instead, we use the **quantile function (inverse CDF)** of the distribution to generate "draws" from evenly-spaced quantiles.

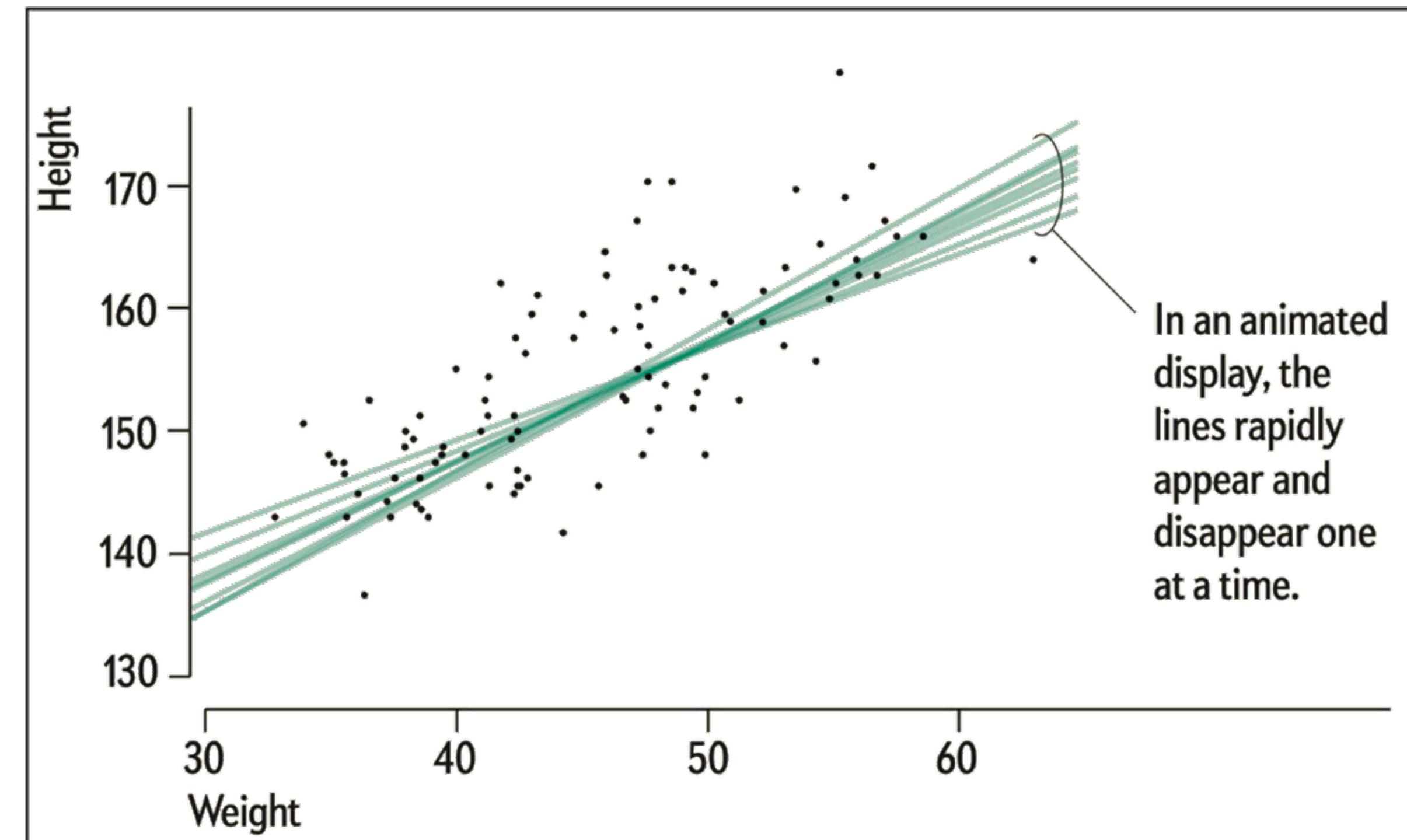


We plot the quantile "draws" using a Wilkinsonian dotplot, yielding what we call a **quantile dotplot**: a consistent discrete representation of a probability distribution.

By using quantiles we facilitate interval estimation from frequencies: e.g., knowing there are 50 dots here, if we are willing to miss our bus **3/50** times, we can count **3 dots** from the left to get a one-sided **94% ($1 - 3/50$) prediction interval** corresponding to that risk tolerance.



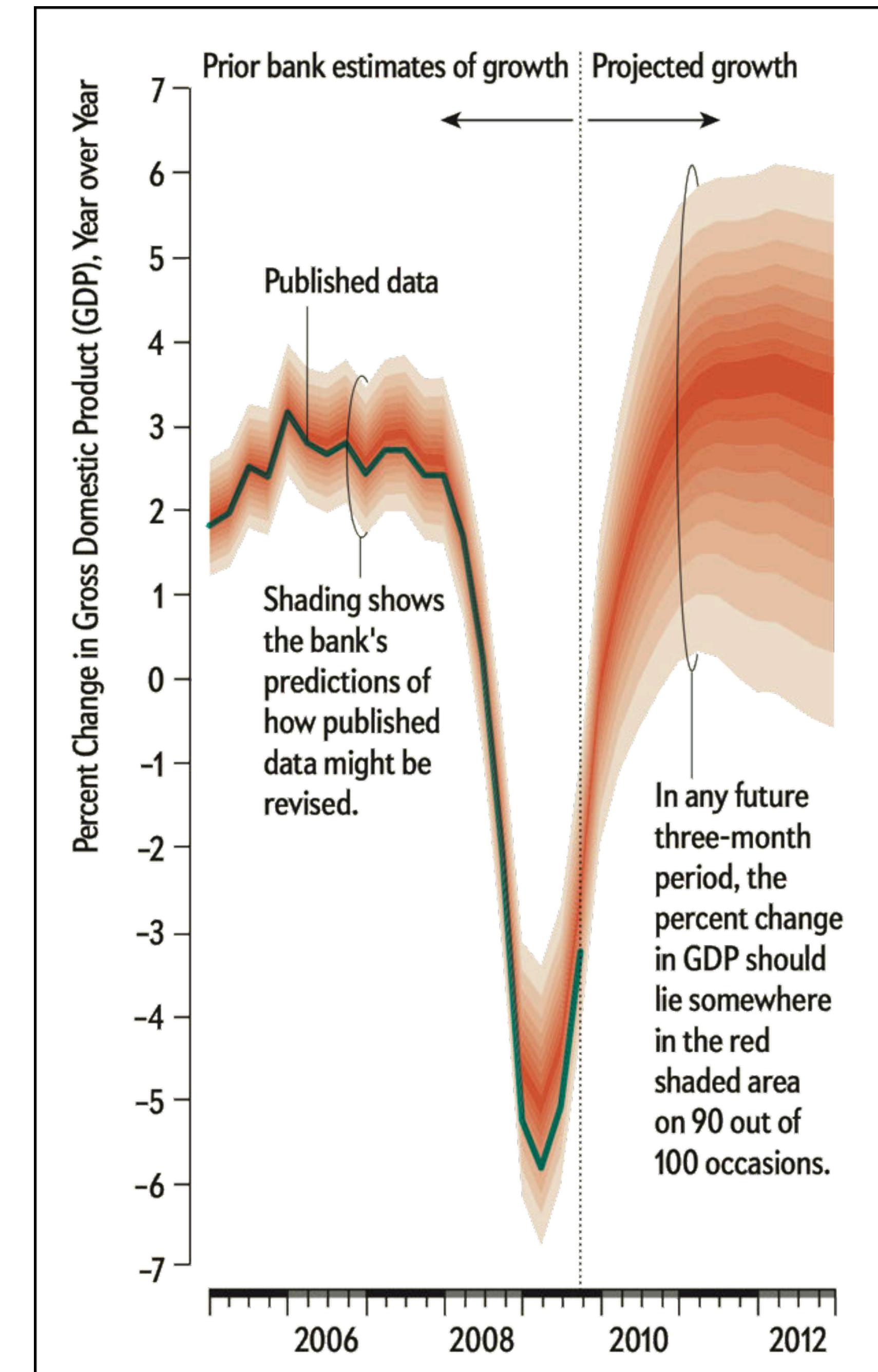
displaying samples in time



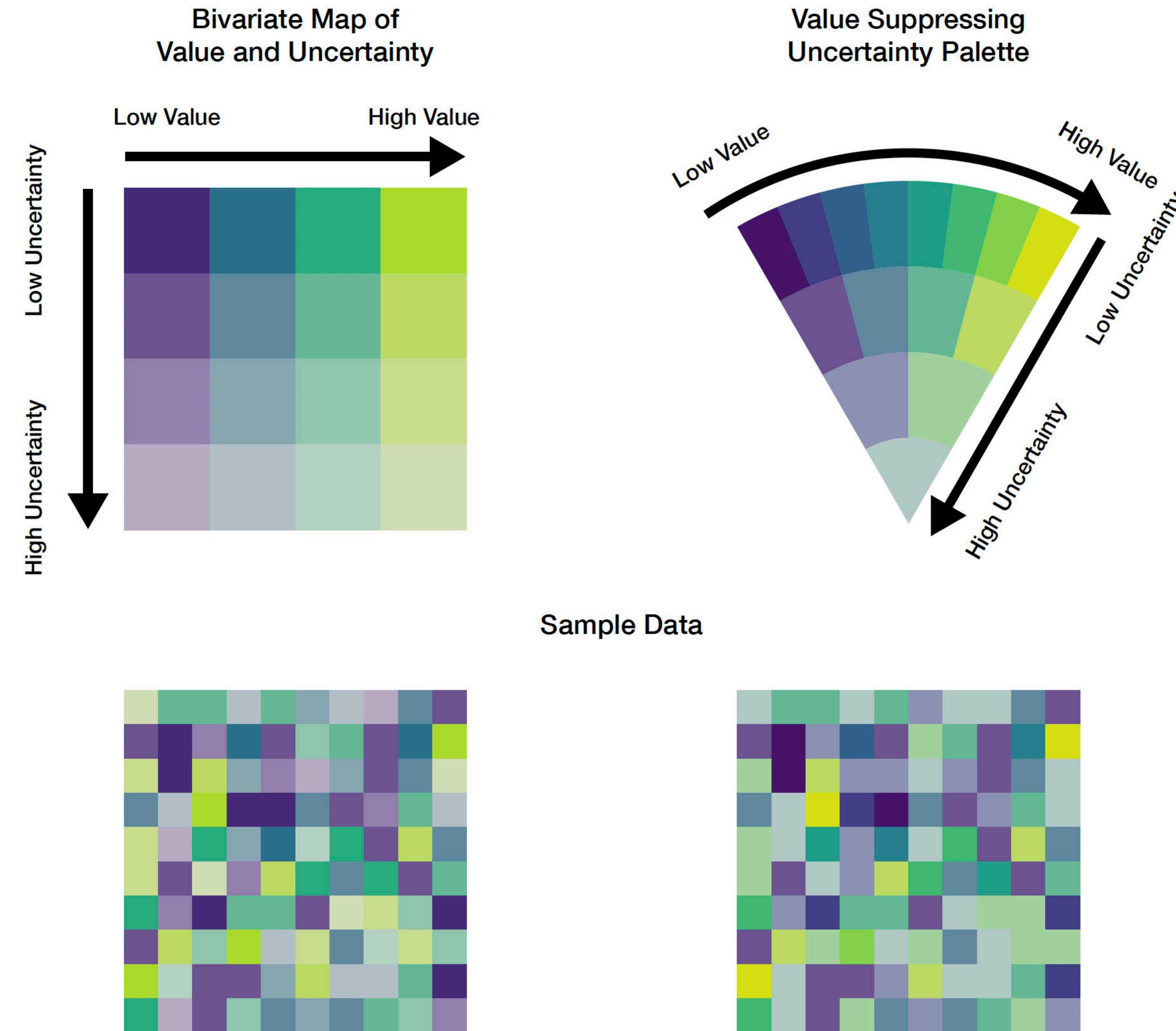
hybrid approaches



Hullman, Jessica



uncertainty | *encoding uncertainty with color*



Correll, Michael & co-authors

practice in the studio

References

- Aisch**, Gregor, Wilson Andrews, and Josh Keller. “*The Cost of Mobile Ads on 50 News Websites*.” New York Times, October 1, 2015, Online edition, sec. Business. <https://www.nytimes.com/interactive/2015/10/01/business/cost-of-mobile-ads.html>.
- Andrews**, R J. *Info We Trust: How to Inspire the World with Data*. Wiley, 2019.
- Bremer**, Nadieh. “*The Top 2000 Loves the 70s & 80s*.” Personal. Visual Cinnamon, December 2016. <https://www.visualcinnamon.com/portfolio/top2000>.
- Congressional Budget Office**. “*The 2012 Long-Term Budget Outlook: Infographic*.” Congressional Budget Office, June 5, 2012. <https://www.cbo.gov/publication/43289>.
- Correll**, Michael, Dominik Moritz, and Jeffrey Heer. “*Value-Suppressing Uncertainty Palettes*.” In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI ’18, 1–11. Montreal QC, Canada: ACM Press, 2018.
- Corum**, Jonathan. See, *Think, Design, Produce 3*. 13pt Information Design, March 28, 2016. <http://style.org/stdp3/>.
- Dimiter**, Toshkov. *World Data Visualization Prize: Small States Can Be Big Players in Development and Good Governance*. Personal. Dimiter Toshkov, 2019. http://www.dimiter.eu/Visualizations_files/WDP.html.
- Fischhoff**, Baruch. *Communicating Uncertainty: Fulfilling the Duty to Inform*. Issues in Science and Technology 28, no. 4 (August 2012): 63–70.
- Hullman**, Jessica. *Confronting Unknowns: How to Interpret Uncertainty in Common Forms of Visualization*. Scientific American, September 2019.
- Information Is Beautiful Awards**. *Voting: What Our Judges (and You) Look out for in Great Visualization*. Accessed April 2, 2020. <https://www.informationisbeautifulawards.com/news/236-voting-what-our-judges-and-you-look-out-for-in-great-visualization>.
- Katz**, Yarden. *Against Storytelling of Scientific Results*. Nature Publishing Group 10, no. 11 (November 2013): 1045–1045.
- Kay**, Matthew, Tara Kola, Jessica R Hullman, and Sean A Munson. “*When (ish) Is My Bus? User-Centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems*.” In The 2016 CHI Conference, 5092–5103. New York, New York, USA: ACM Press, 2016.
- Krzywinski**, Martin, and Alberto Cairo. *Storytelling*. Nature Publishing Group 10, no. 8 (August 2013): 687–687.
- . *Reply to: Against Storytelling of Scientific Results*. Nature Publishing Group 10, no. 11 (November 2013): 1046–1046.
- Malofiej**. *Quick Questionnaire*. Accessed April 2, 2020. <https://www.malofiejgraphics.com/category/quick-questionnaire/>
- McCandless**, David. *What Makes a Good Visualization?* Information is Beautiful. Accessed April 2, 2020. <https://informationisbeautiful.net/visualizations/what-makes-a-good-data-visualization/>.
- Ojo**, Adegboyega, and Bahareh Heravi. “*Patterns in Award Winning Data Storytelling: Story Types, Enabling Tools and Competences*.” Digital Journalism 6, no. 6 (July 3, 2018): 693–718.
- Riche**, Nathalie Henry, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale. *Data-Driven Storytelling*. CRC Press, 2018.
- Spencer**, Scott. *Approximating the Components of Lupi’s Nobels, No Degrees*. p(ssp3nc3r | Columbian) (blog), March 15, 2019. <https://ssp3nc3r.github.io/post/approximating-the-components-of-lupi-s-nobel-no-degrees/>.
- . *Ride Against the Flow*. 2019. Computer graphics. <https://www.informationisbeautifulawards.com/showcase/4367-ride-against-the-flow>.

additional material for discussion

(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

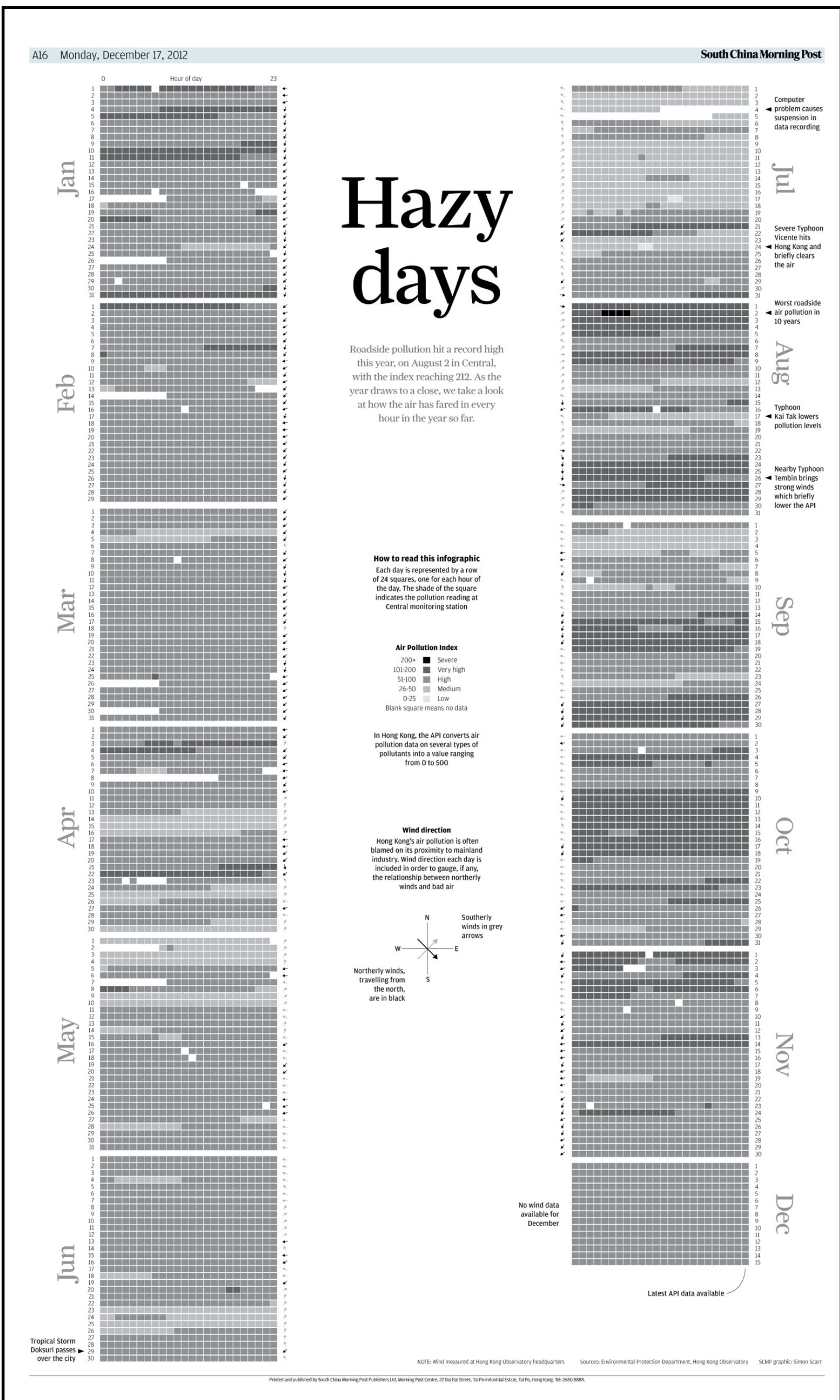
Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?



(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

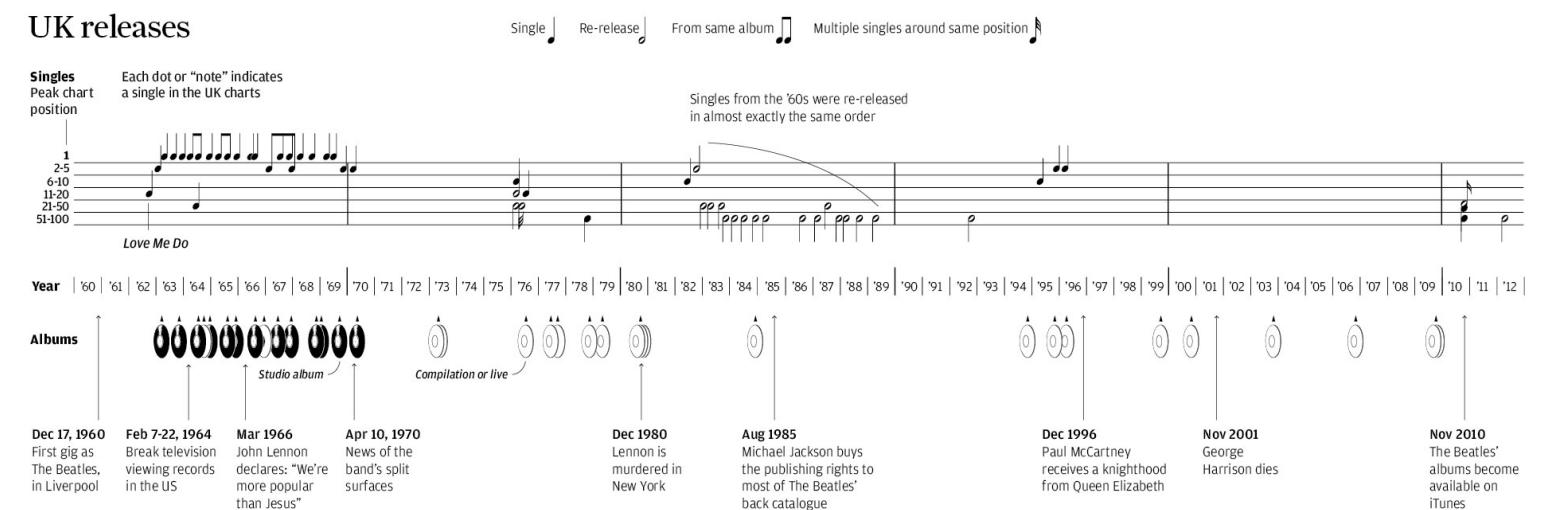
Credibility, transparency?

A18 Saturday, October 13, 2012

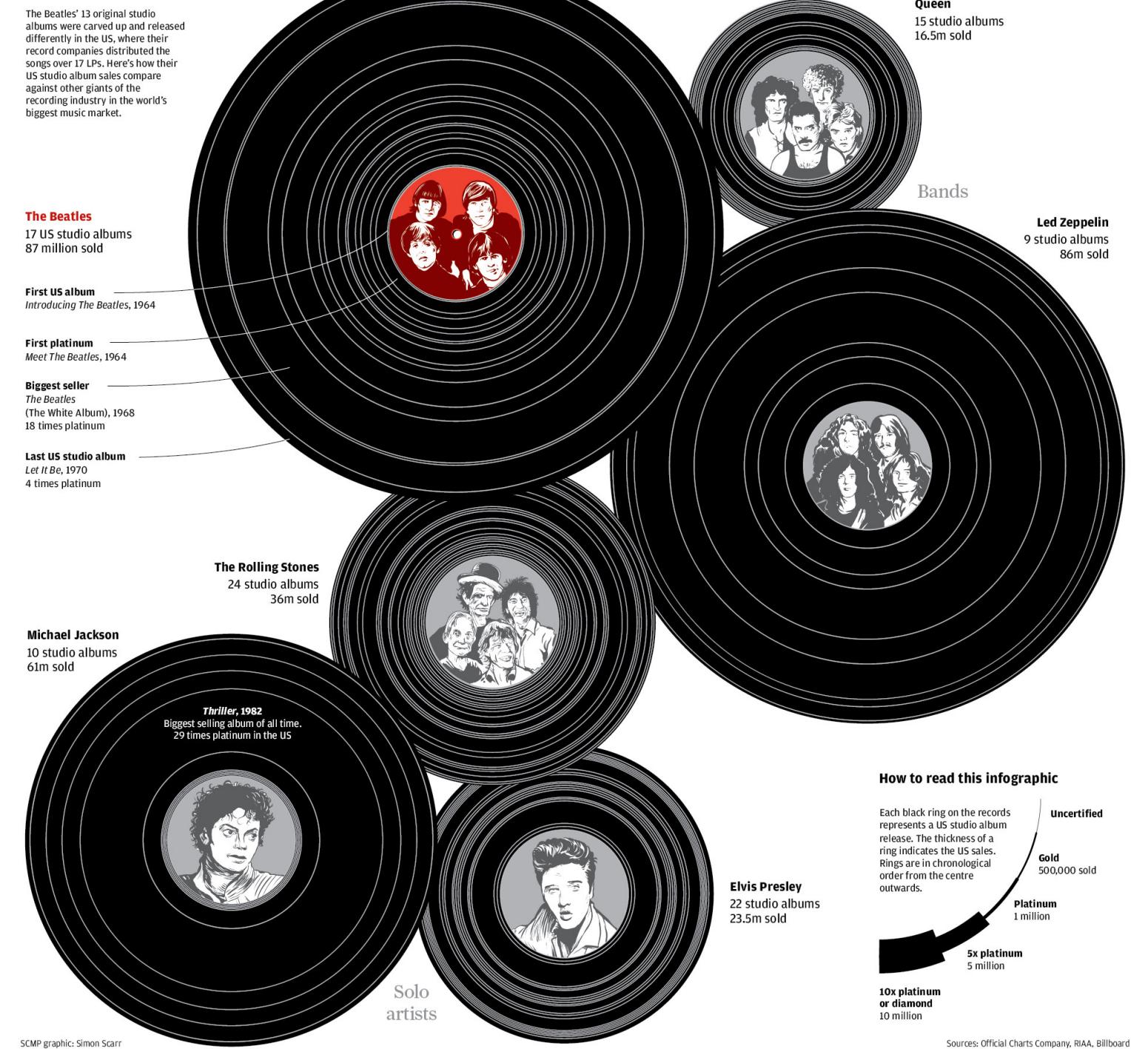
South China Morning Post

Love Me Do

Fifty years ago today The Beatles made their chart debut with their first single, *Love Me Do*. They would go on to have 17 number one singles in their native Britain and sell over 80 million albums in the United States, all the while changing the face of popular culture.



US album sales



(info)graphics examples for discussion

Audience?

Annual report 2018

page 11

Purpose?

Expansion into the next adjacent segment: Mid-market

Historically, we have been focused on the needs of large multinational merchants. However, we believe that we are well-positioned to also grow the business with mid-market merchants, which we view as the next adjacent segment to enterprise, as a result of the following:

- Industry-leading functionality: All merchants that use our platform gain access to the same performance and functionality as the world's largest multinational companies. This makes the platform attractive to both large domestic merchants and local merchants with international ambitions, allowing them to effectively future-proof their payments.
- Serving local heroes: Over the last 11 years, we've been able to establish a truly global footprint, with 18 offices worldwide, which provide local presence, payment methods and expertise in key markets. This allows us to focus on further serving local mid-market merchants.

Narrative?

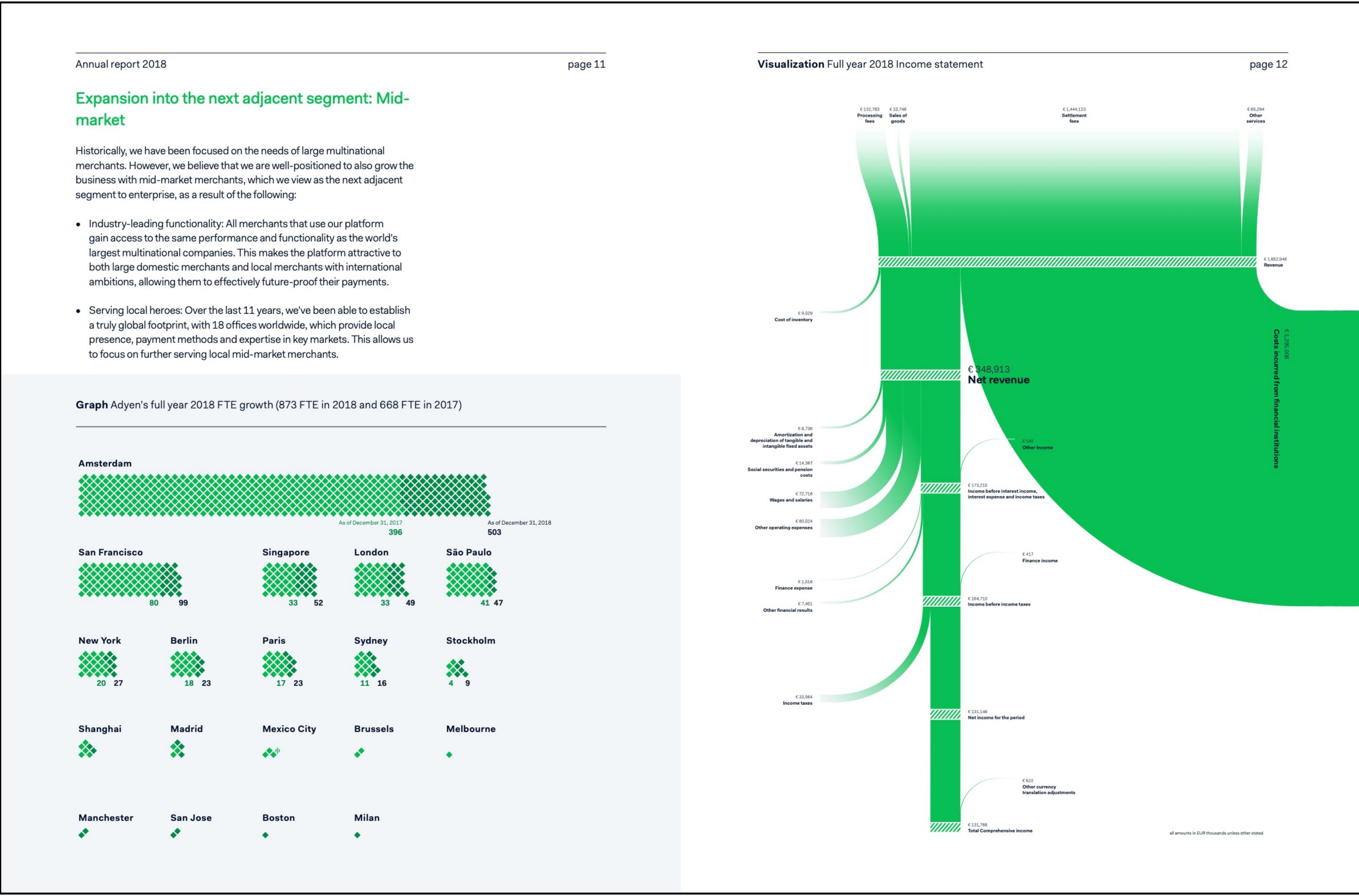
Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?



(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?

López, Alberto Lucas

A14 Thursday, June 26, 2014

South China Morning Post

How to read this infographic

Each square represents HK\$10 million market value. The 16 teams with the highest market value playing at the World Cup. Market value of the highest paid player for the top 16 teams.

The 16 teams with the lowest market value playing at the World Cup.

Player (club where he plays) Personal market value

Percentage approx. on the total market value of his national team

National football team (low to high market value)

HKS millions

Brazil Neymar (FC Barcelona) 710

Spain Juan Mata (Chelsea) 496

Argentina Lionel Messi (FC Barcelona) 1,455

Germany Mario Götze (Bayern Munich) 635

France Karim Benzema (Real Madrid) 341

Belgium Eden Hazard (Chelsea) 480

Italy Claudio Marchisio (Juventus) 308

England Wayne Rooney (Manchester United) 443

Portugal Cristiano Ronaldo (Real Madrid) 1,131

Colombia Radamel Falcao (Monaco) 593

Uruguay Edinson Cavani (PSG) 607

Russia Alan Dzagoev (CSKA Moscow) 254

Croatia Luka Modrić (Real Madrid) 338

Netherlands Robin van Persie (Man United) 398

Chile Arturo Vidal (Juventus) 443

Ivory Coast Yaya Touré (Man City) 349

Costa Rica Costa Rica 219

Iran Iran 231

Australia Australia 249

Honduras Honduras 256

South Korea South Korea 487

Mexico Mexico 497

United States United States 541

Algeria Algeria 665

Greece Greece 683

Ecuador Ecuador 731

Ghana Ghana 890

Nigeria Nigeria 945

Japan Japan 1,117

Bosnia Herzegovina Bosnia Herzegovina 1,278

Cameroon Cameroon 1,296

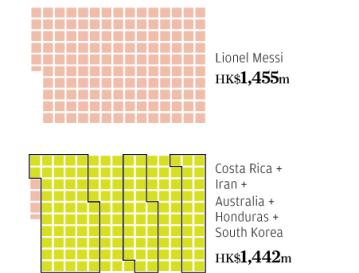
Switzerland Switzerland 1,310

The World Cup's great divide

While you can't put a price on a fan's passion for the beautiful game, we do know how much clubs are willing to pay for players. By combining the value of individual footballers, we have estimated the market value of each of the 32 teams at the World Cup in Brazil.



Individual vs country
Lionel Messi is the world's most valuable player. Financially, the market value of Argentina's worth alone could finance the five teams with the lowest market value. None of the 16 lowest-ranked teams have players among the top 100 highest paid.



SCMP graphic: Alberto Lucas López

Sources: PLURB Consultancy report "Os 100 jogadores mais valiosos do mundo em 2013" (The 100 most valuable players in the world in 2013), PLURB Consultancy report "Valor do Mercado das 32 Seleções que disputarão a Copa do Mundo Brasil 2014" (Market value of the 32 teams at the World Cup in Brazil), Football Finance

The data provided in the reports are in euro.

Conversion value at time of infographics 1 EUR = HK\$10.5412

Note: the report about the market value of the players was published on January 14, 2014. Data and their clubs refers to that time.

Printed and published by South China Morning Post Publishers Ltd, Morning Post Centre, 22 Des Voeux Street, Tai Po Industrial Estate, Tai Po, Hong Kong, Tel: 2660 8888.

(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

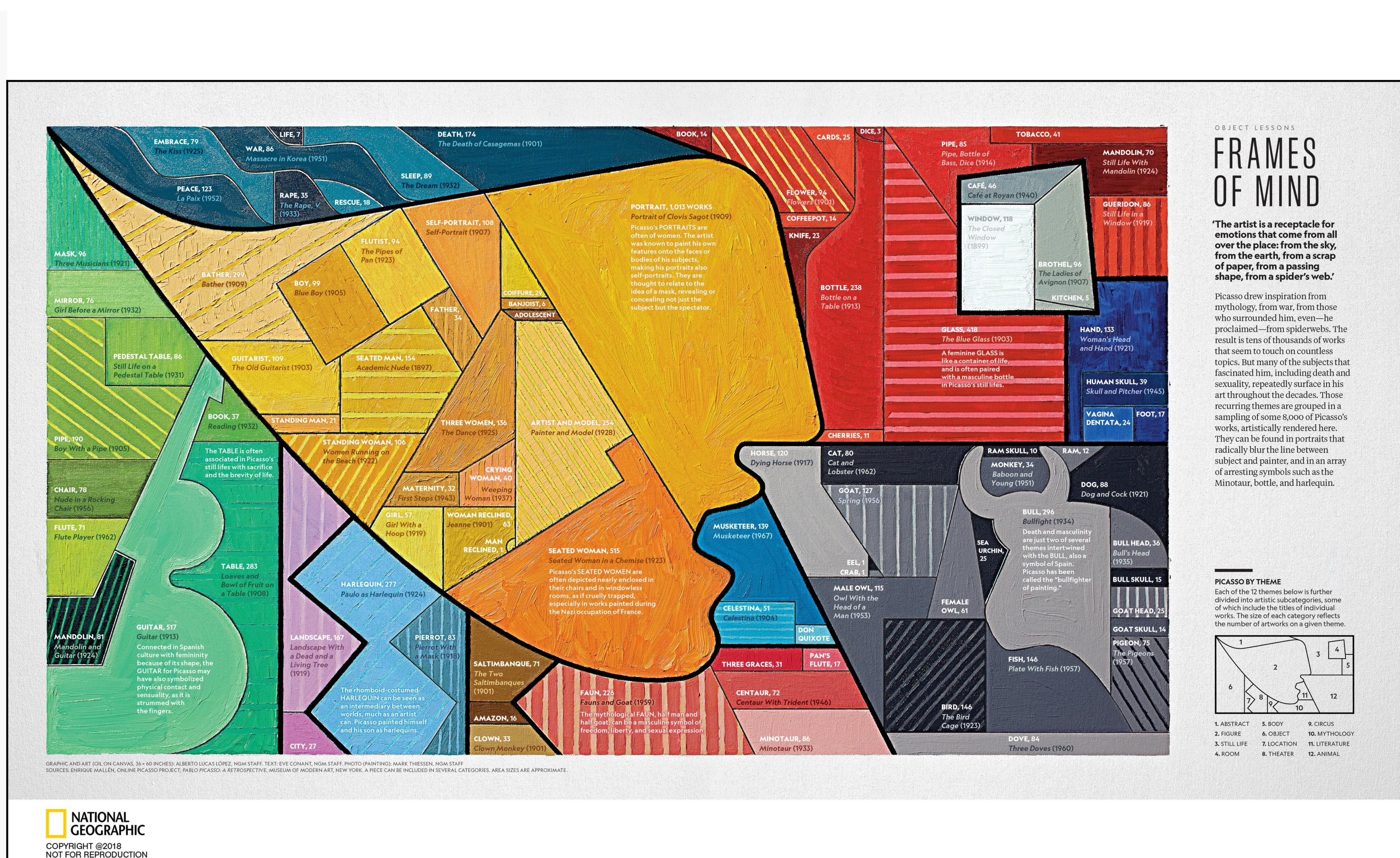
Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?



López, Alberto Lucas

(info)graphics examples for discussion

Audience?

Purpose?

Narrative?

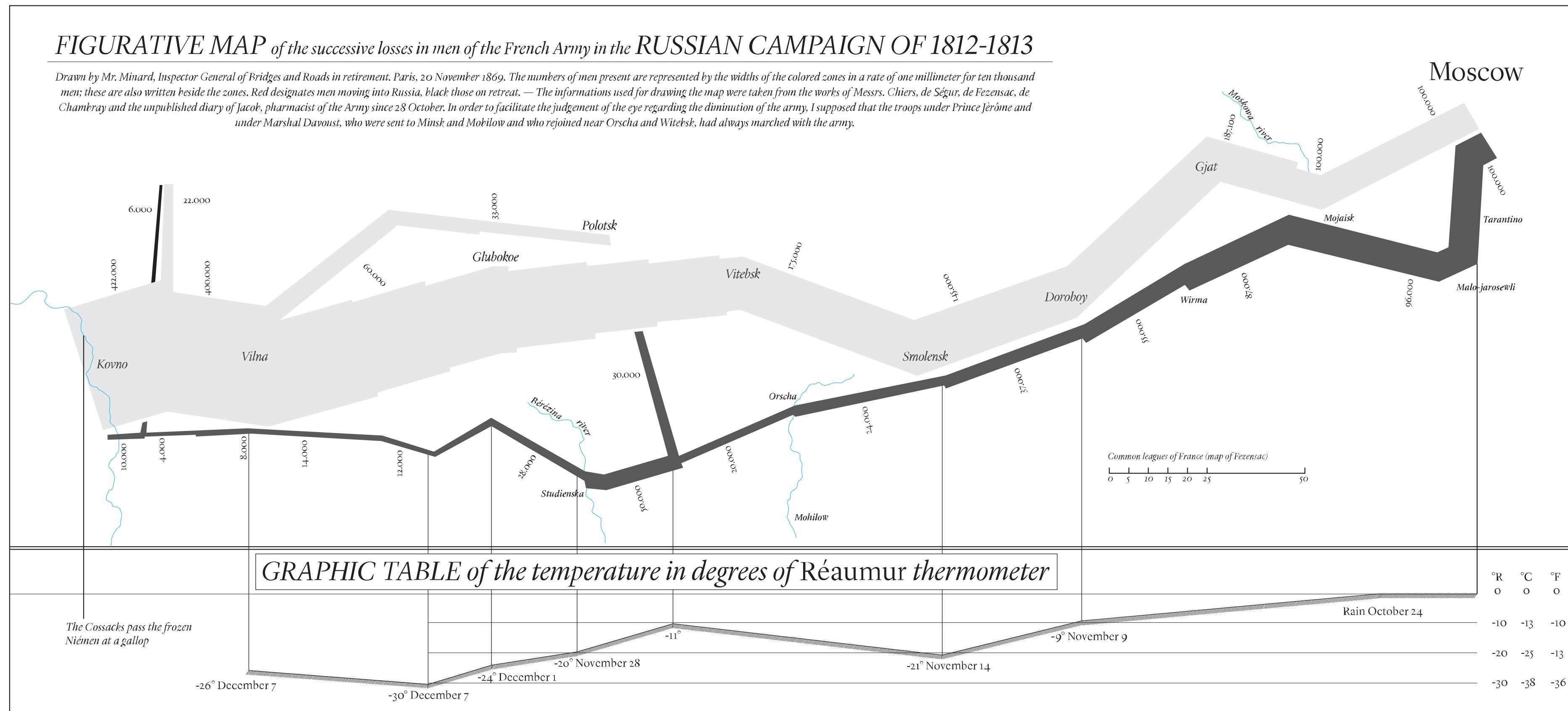
Data encodings, decodings?

Comparison or change?

Color, coherency?

Layering, layout?

Credibility, transparency?



Minard