

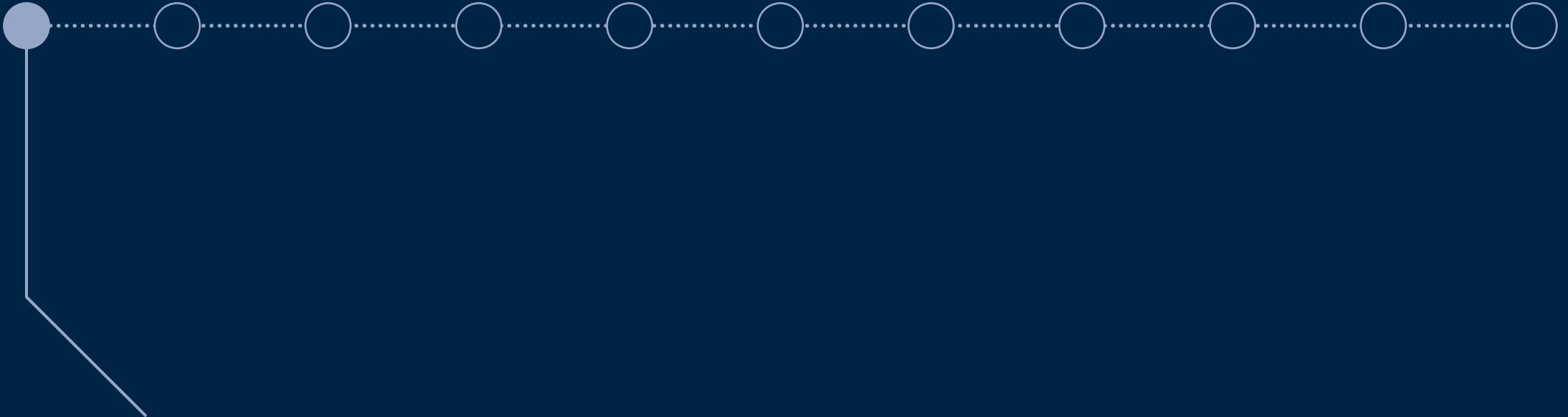


EFFECTIVE VISUALS

Helping readers see what you see in your data

AN INTENSIVE WORKSHOP ON THE HUMBLE,
YET POWERFUL, SCIENTIFIC CHART

BY KATIE PEEK



0 THE HUMBLE SCIENTIFIC CHART, DEFINED

0 DEFINITIONS

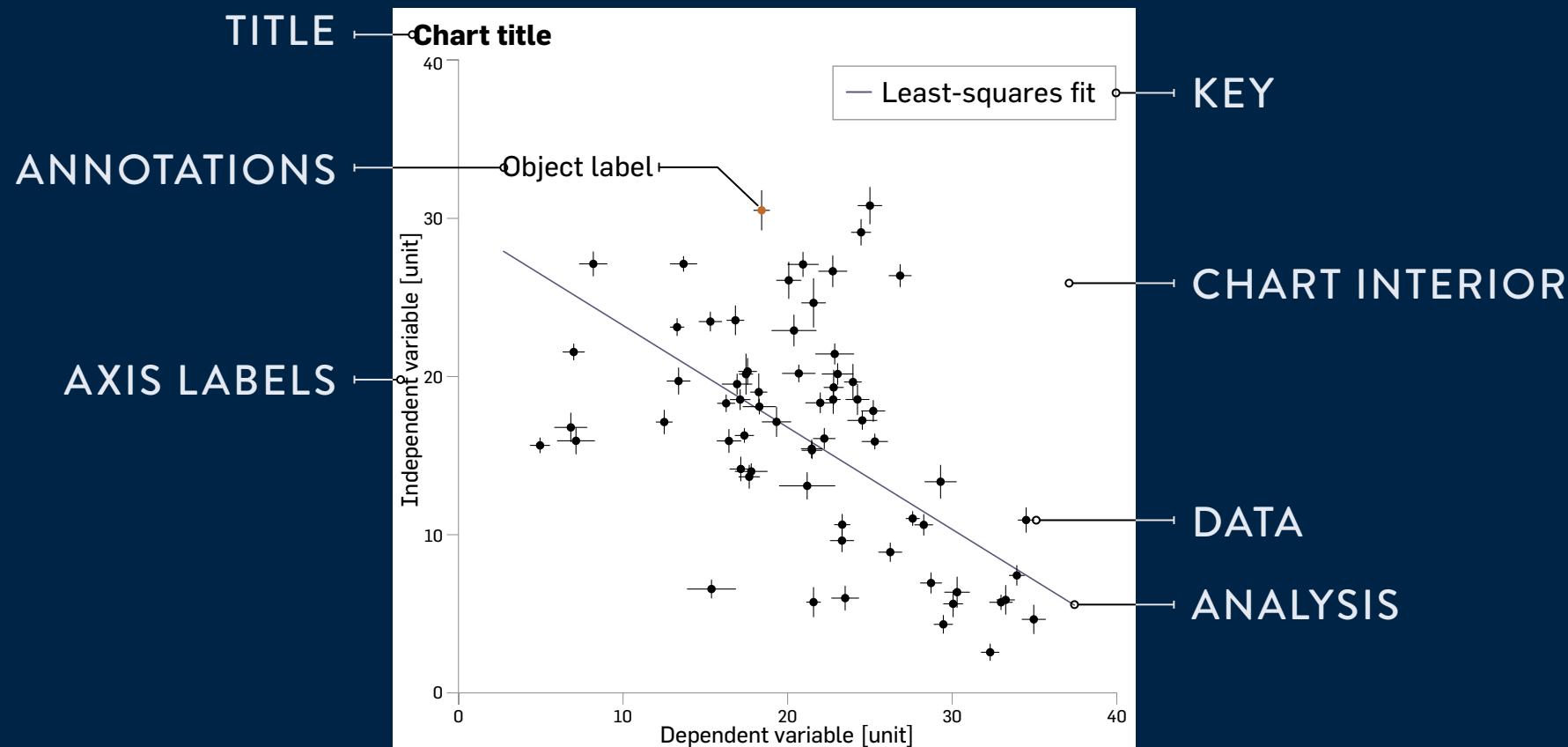
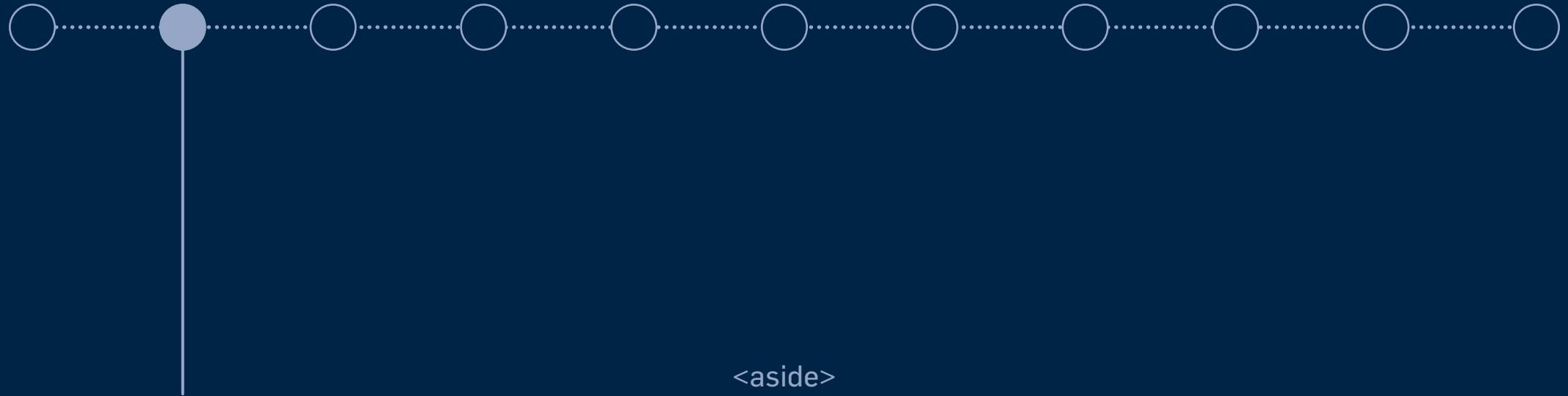


Fig. 1 This talk will focus on the elements of that humble, yet powerful, scientific chart. Terms I'll use are displayed in white.



<aside>

1 WHO I AM—AND MY PERSPECTIVE ON CHARTS

GRAPHIC SCIENCE

Text and Graphic by Katie Peek

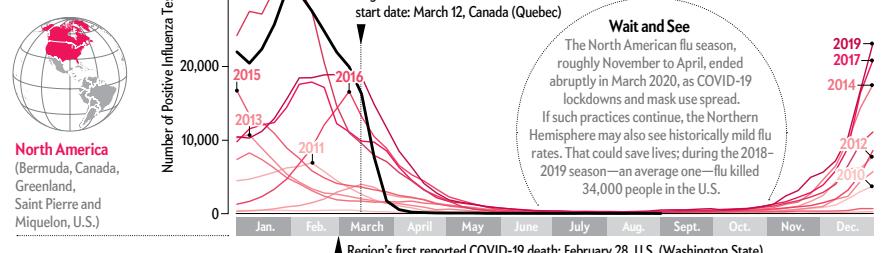
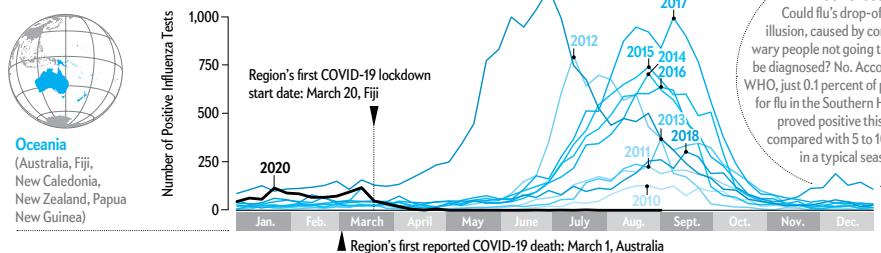
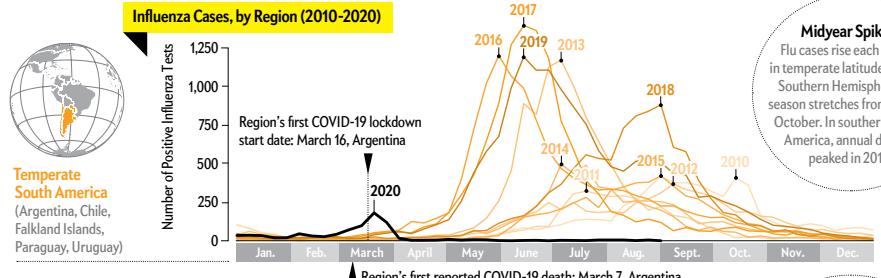
The Missing Flu Season

COVID-19 response has thwarted influenza in the Southern Hemisphere

Region by Region

The World Health Organization tracks influenza activity in 18 transmission zones. Three of those regions appear here. Only people who get tested for influenza-like illnesses—typically about 5 percent of those who fall ill—are tallied.

In March, as coronavirus widened its global sweep, one health statistic quickly flattened: influenza cases. In the Southern Hemisphere, flu season would have been just taking off, but cases were virtually nonexistent. “Never in my 40-year career have we ever seen rates … so low,” says Greg Poland, an influenza expert at the Mayo Clinic. Although researchers need to study the reasons further, several told *Scientific American* that coronavirus prevention measures—hand washing, mask wearing and social distancing—are working against flu transmission. If those measures continue, Poland says, countries could see the most dramatic drop in influenza cases in modern human history. U.S. health experts still recommend flu shots, however, because not everyone in the country is observing measures to contain the virus and because COVID-19 could perhaps be more threatening in people who contract flu.



SOURCE: ILLINOIS GLOBAL INFLUENZA SURVEILLANCE AND RESPONSE SYSTEM, WORLD HEALTH ORGANIZATION (INFLUENZA CASES)

GRAPHIC SCIENCE

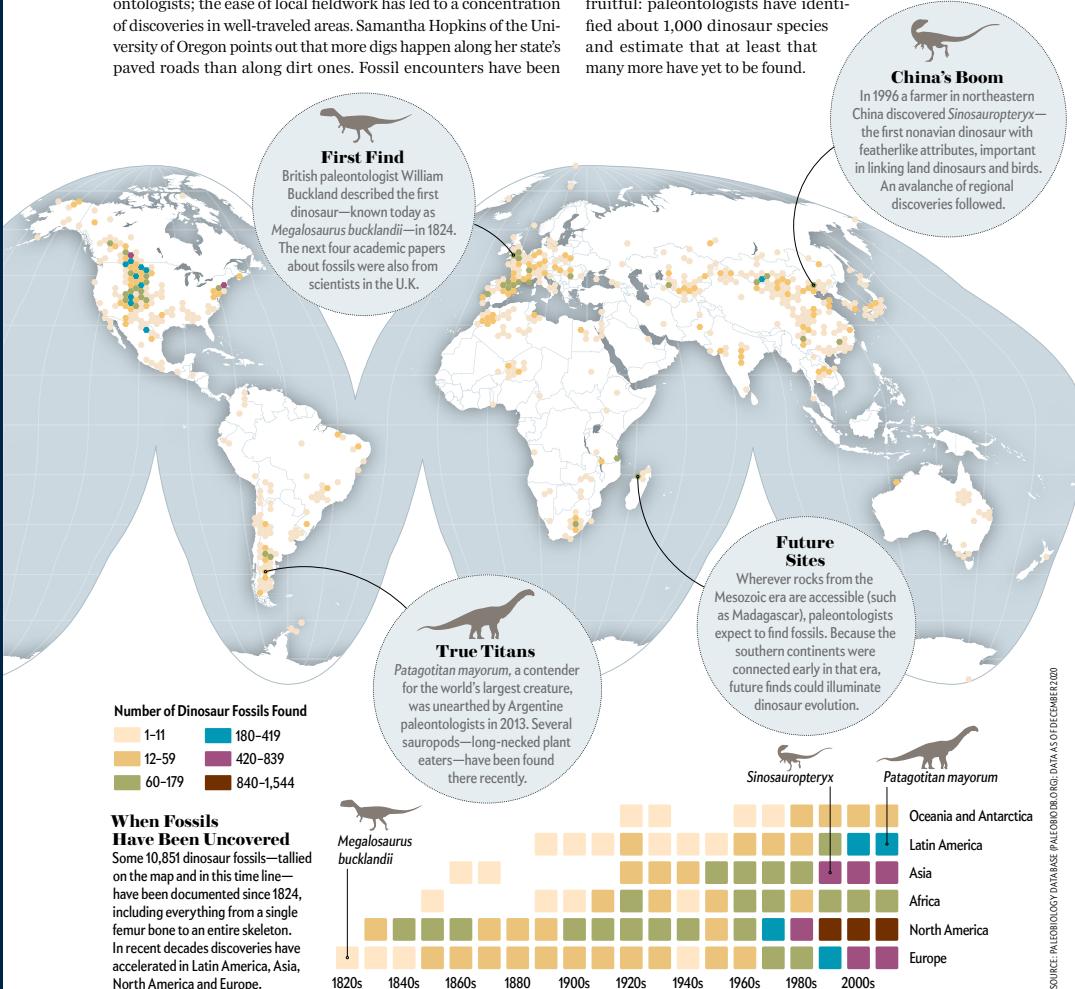
Text and Graphic by Katie Peek

Where the Dinosaurs Are

Discovery is booming, and there are plenty more bones to find

In the two centuries since the first dinosaur bones were identified in England, nearly 11,000 dinosaur fossils have been unearthed worldwide, two thirds of them in North America and Europe. Most of the finds have been made in the home countries of paleontologists; the ease of local fieldwork has led to a concentration of discoveries in well-traveled areas. Samantha Hopkins of the University of Oregon points out that more digs happen along her state's paved roads than along dirt ones. Fossil encounters have been

expanding geographically, however, notably in East Asia and southern South America. Widening the scope depends on building local expertise—a tricky task for a fairly niche (and not particularly lucrative) field. Investment could be fruitful: paleontologists have identified about 1,000 dinosaur species and estimate that at least that many more have yet to be found.



1 AN EMOJI TIMELINE



.....



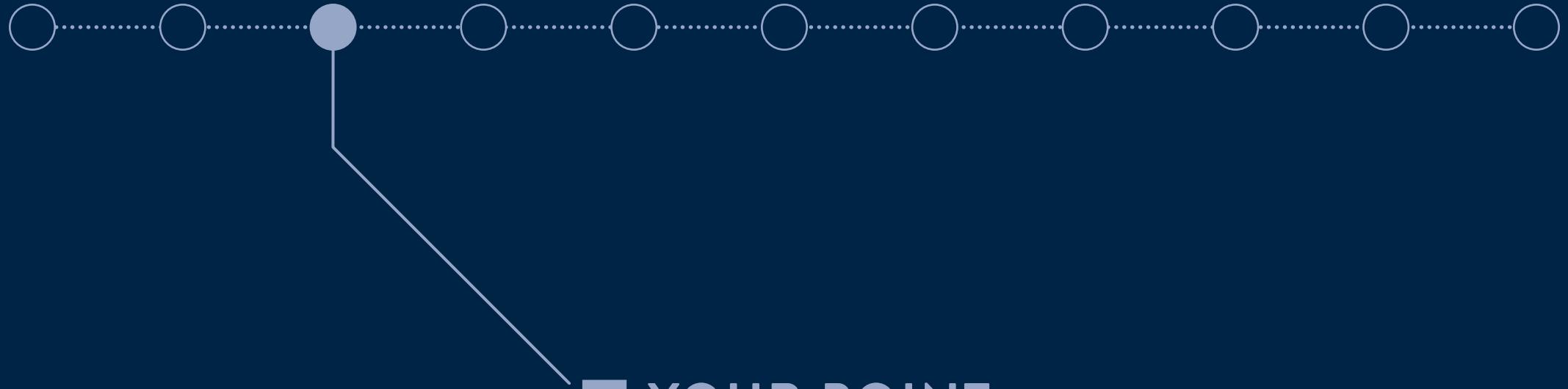
.....



1 A CONFESSION



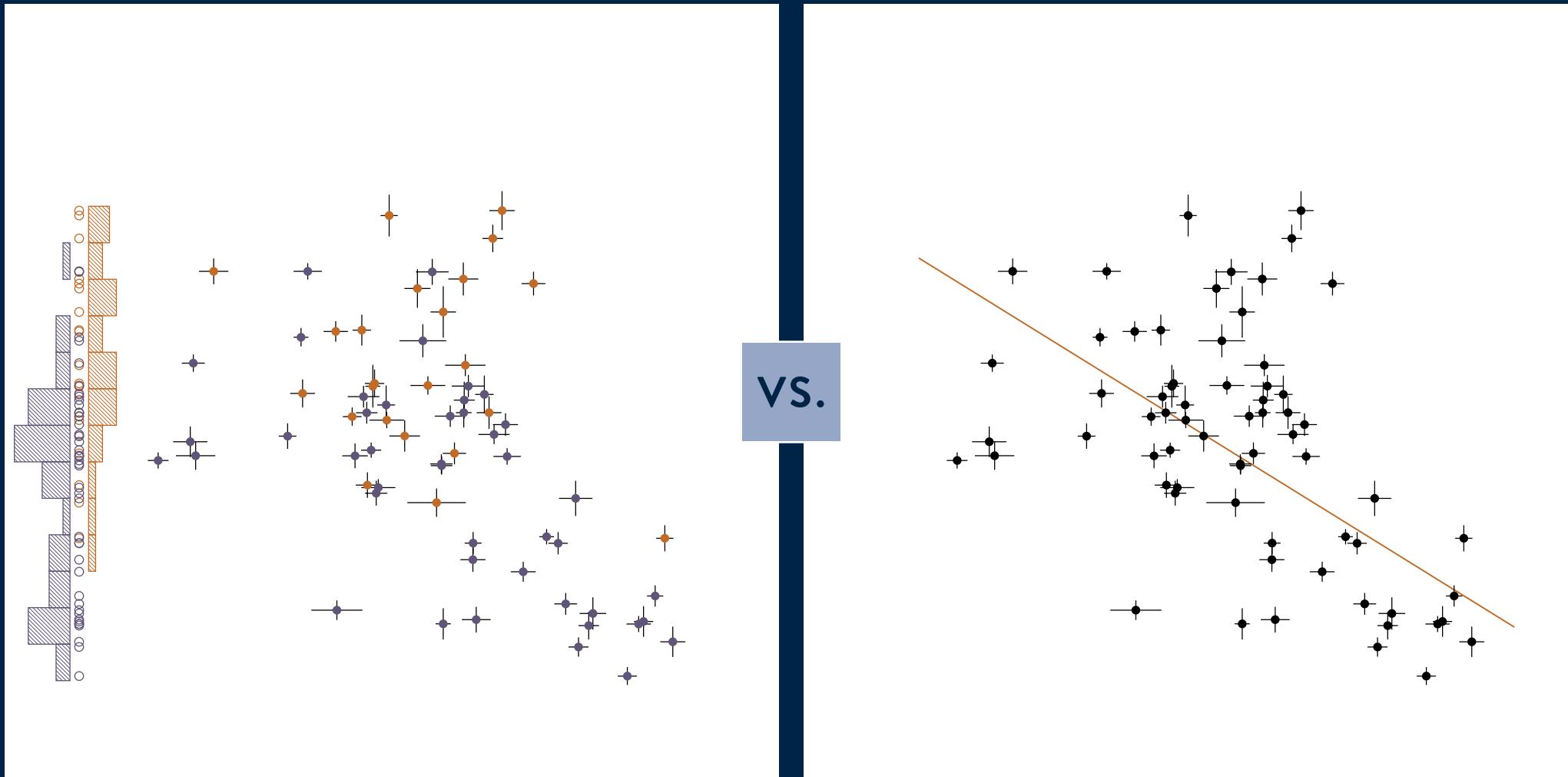
</aside>



2 YOUR POINT

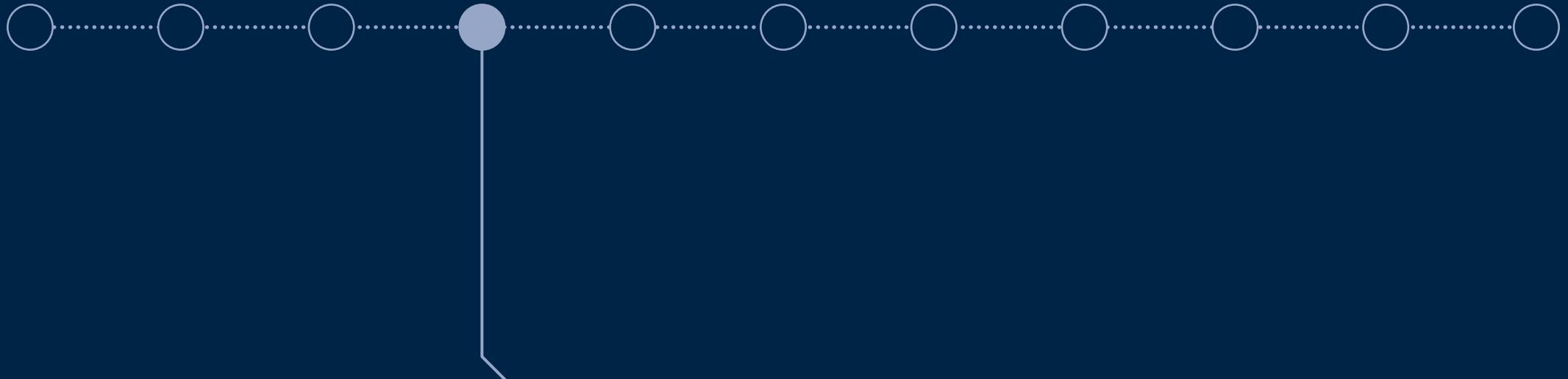
Ask thyself: What are you trying to say? What is this chart's point?

2 DESIGN CHOICES SHIFT VISUAL EMPHASIS



2 WHAT POINT DO YOU WANT YOUR CHART TO MAKE?

- YOU MIGHT FRAME THE POINT IN ONE OR TWO SENTENCES
- TRY THINKING OF THE POINT AS THE ARGUMENT THE CHART WILL MAKE—AN IDEA YOU'RE TRYING TO CONVINCE YOUR READER OF
- BY DEFINING YOUR POINT AT THE OUTSET, IT CAN GUIDE YOUR CHOICES



3 YOUR AUDIENCE

Ask thyself: Whom are you speaking to?

3 SOME POSSIBLE AUDIENCES

FELLOW SUB-FIELD
SPECIALIST

Perhaps for a workshop talk

FELLOW
ASTRONOMER

A chart for an *ApJ* paper, maybe

SCIENTIST IN
ANOTHER FIELD

Perhaps for a paper in *Science*

EDUCATED
LAYPERSON

A figure in a press release

3 AUDIENCES AS ACTUAL PEOPLE



FELLOW SUB-FIELD
SPECIALIST

Perhaps for a workshop talk



FELLOW
ASTRONOMER



A chart for an *ApJ* paper, maybe



SCIENTIST IN
ANOTHER FIELD



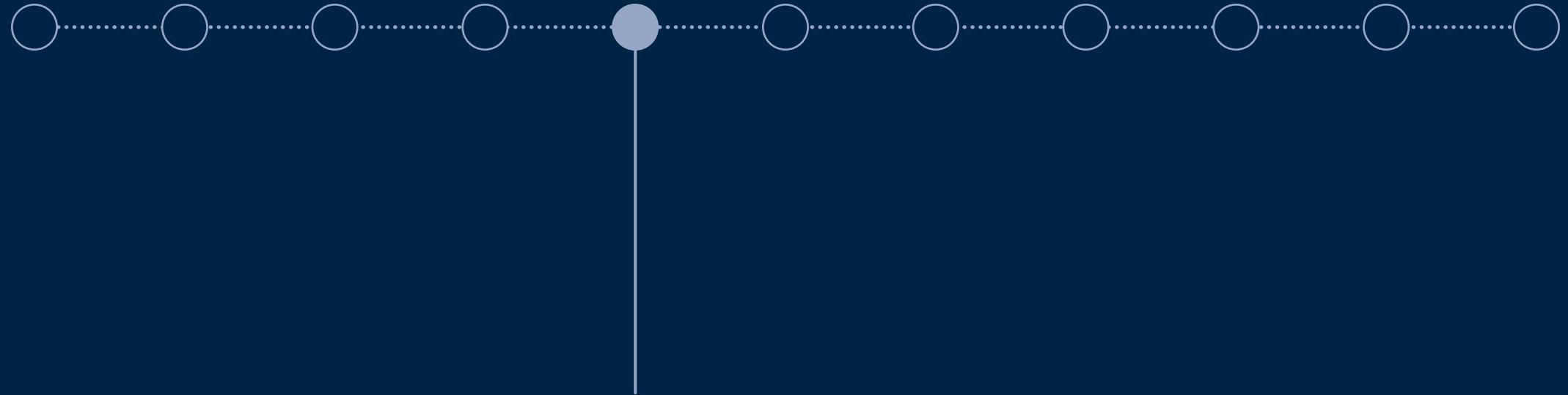
Perhaps for a paper in *Science*



EDUCATED
LAYERPERSON



A figure in a press release



4 CHART TYPES

4 ALL THE CHARTS!

Visual vocabulary

Designing with data

There are so many ways to visualise data. How do we know which one to pick? Use this guide to assess the best way to decide which data representation is most appropriate for your story, then look at the different types of chart, whether the categories form some kind of hierarchy or whether they work better. This is by no means an exhaustive list, but it's a useful starting point for making informed and meaningful data visualisations.

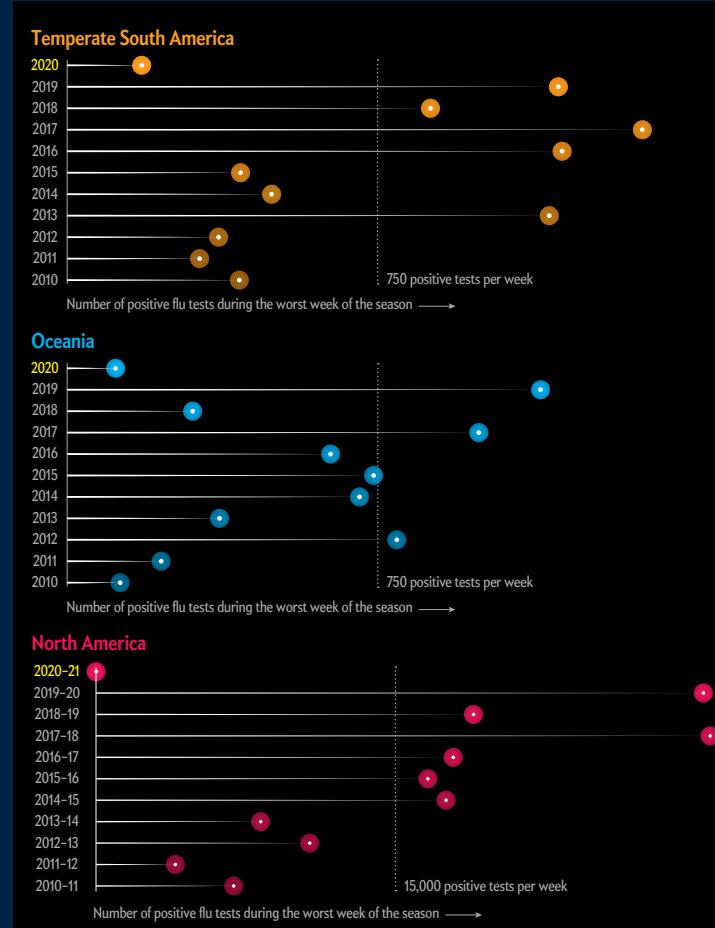
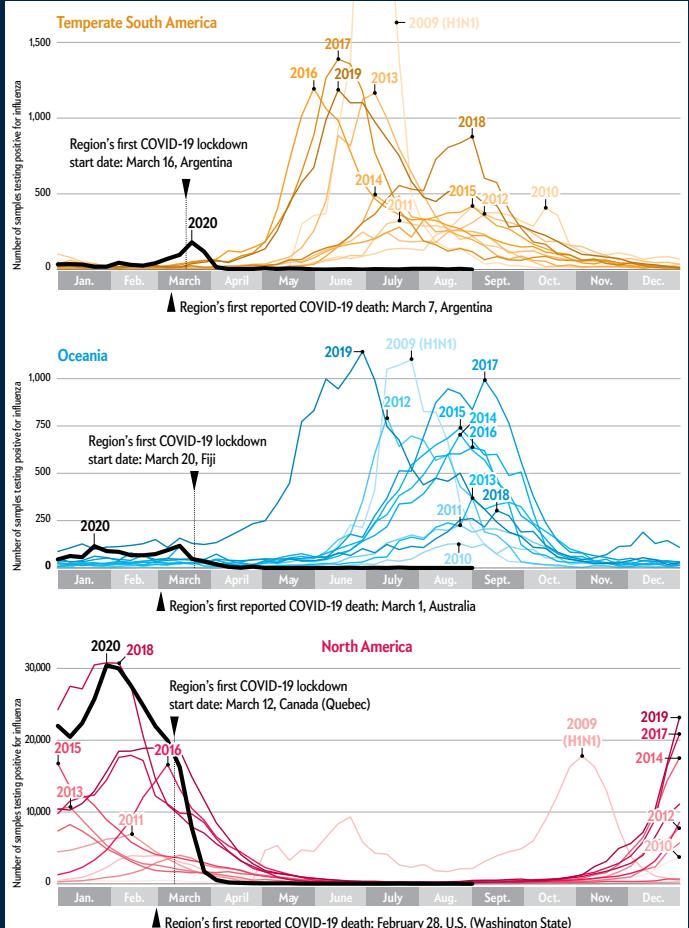


ft.com/vocabulary



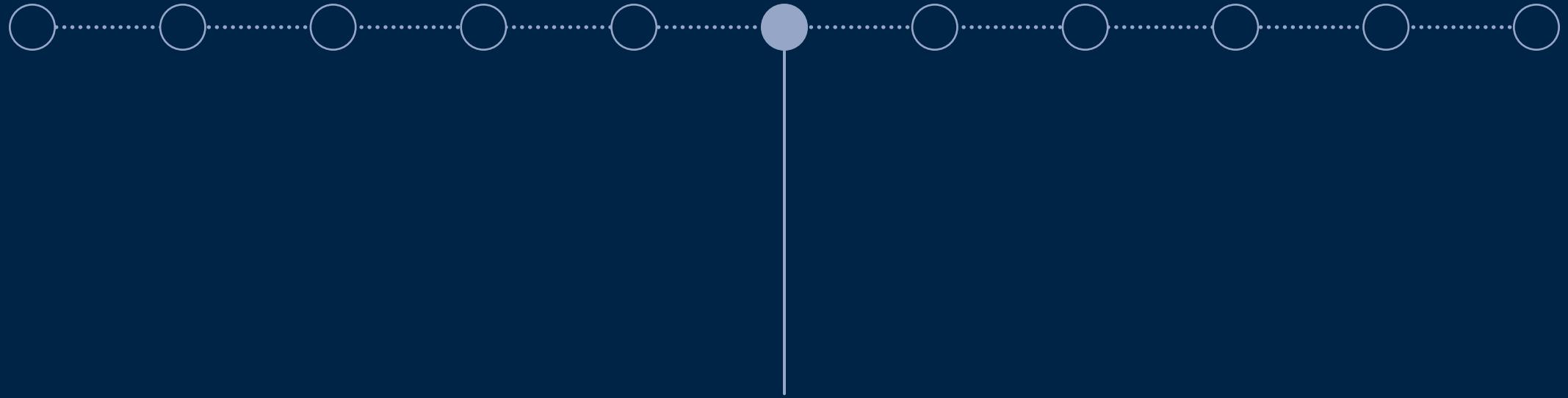
FT

4 PRIORITY: DOES THE CHART MAKE YOUR POINT?



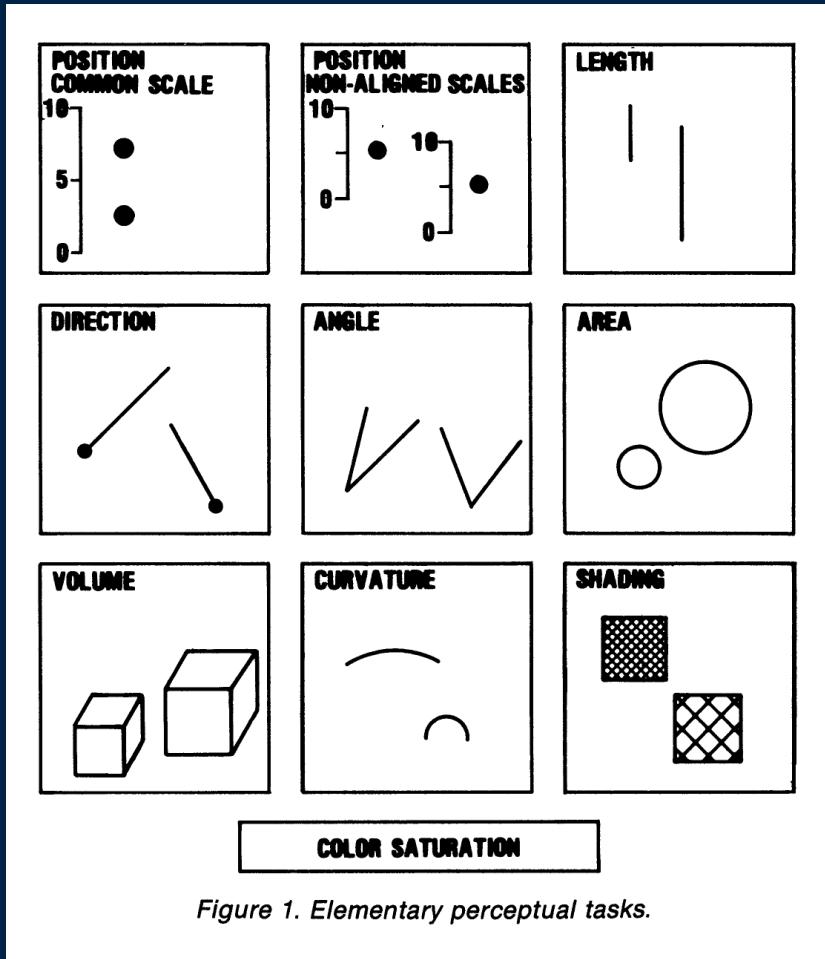
4 CHOOSE THE TYPE THAT MAKES YOUR POINT

- THE BEST CHART DESIGN PLACES YOUR POINT FRONT AND CENTER, VISUALLY...
- YET ALLOWS READERS TO MAKE THEIR OWN CLEAR AND FAIR ASSESSMENT TO VERIFY
- MAKE YOUR ARGUMENT WITH THE STRONGEST VISUAL ELEMENTS IN THE CHART...
- BUT INCLUDE ENOUGH EVIDENCE THAT YOU CAN SHOW, DON'T TELL



5 VISUAL PERCEPTION

5 THE HIERARCHY OF VISUAL PERCEPTION



EASIEST

Comparing lengths or positions along a single scale

MIDDLE

Comparing angles, directions, and areas

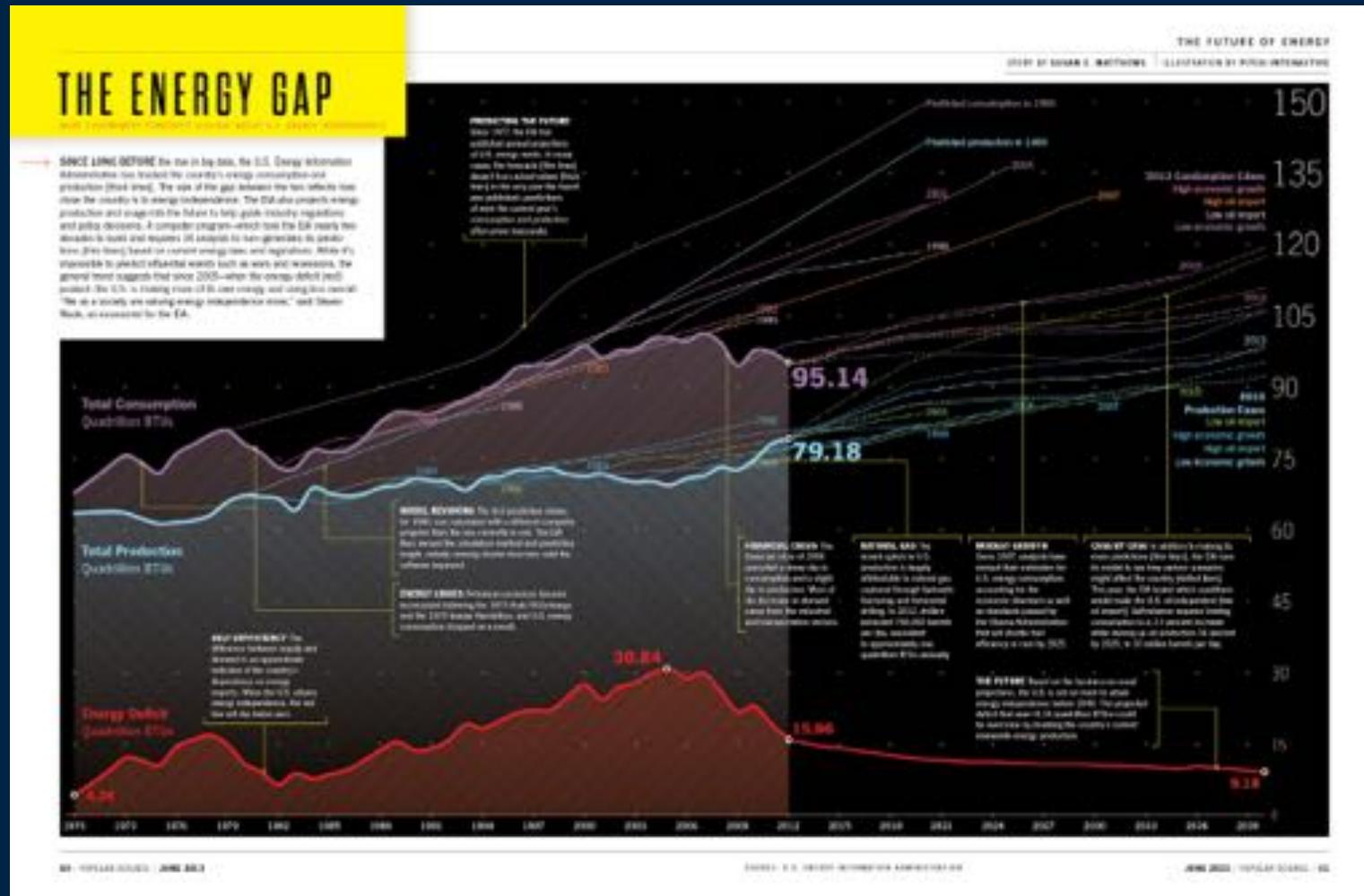
HARDEST

Comparing volumes, curves, and colors



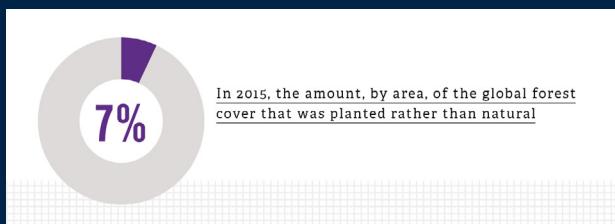
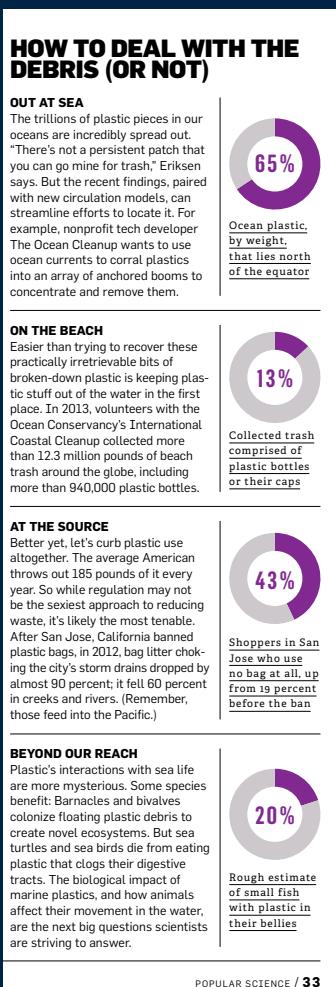
"Graphical Perception:
Theory, Experimentation, and
Application to the Development
of Graphical Methods,"
William S. Cleveland
and Robert McGill,
*Journal of the American
Statistical Association*,
September 1984

5 HELP THE READER'S EYE

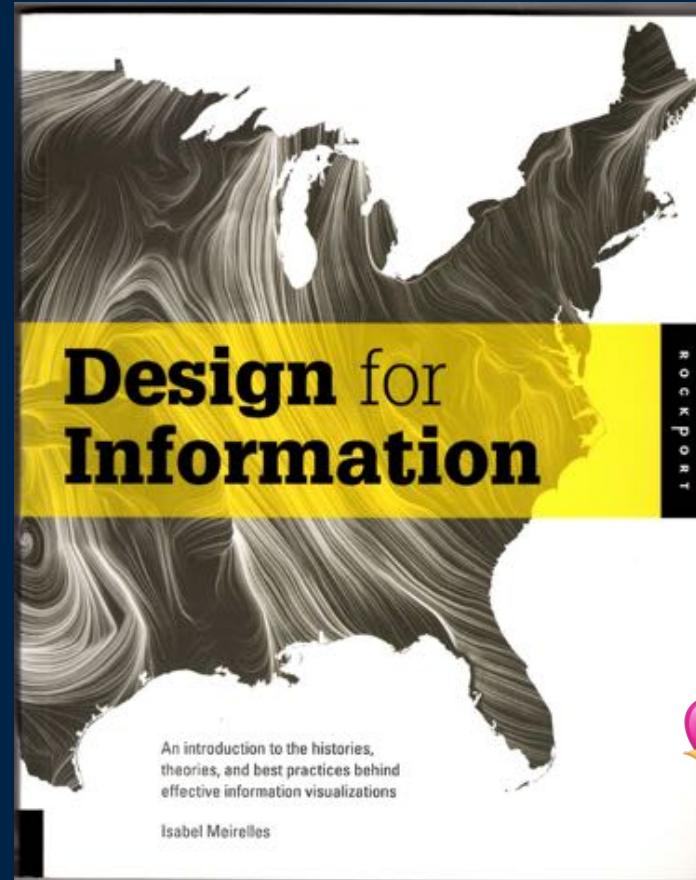
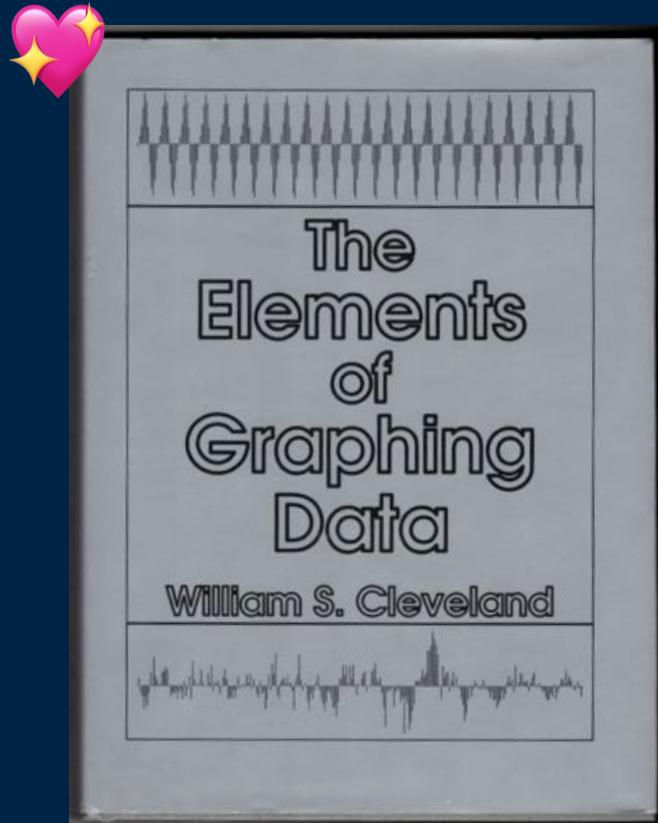


Bigger → Smaller
Thicker → Thinner
Brighter → Darker

5 FAMILIARITY COUNTS



5 FOR INSPIRATION & IDEAS

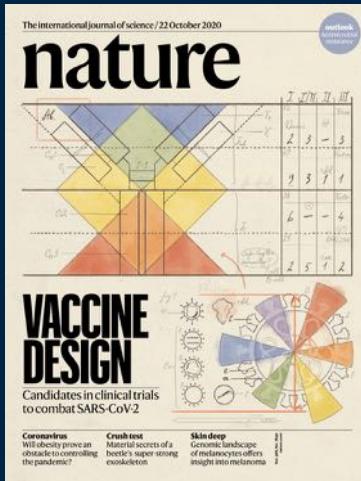




6 COLORS

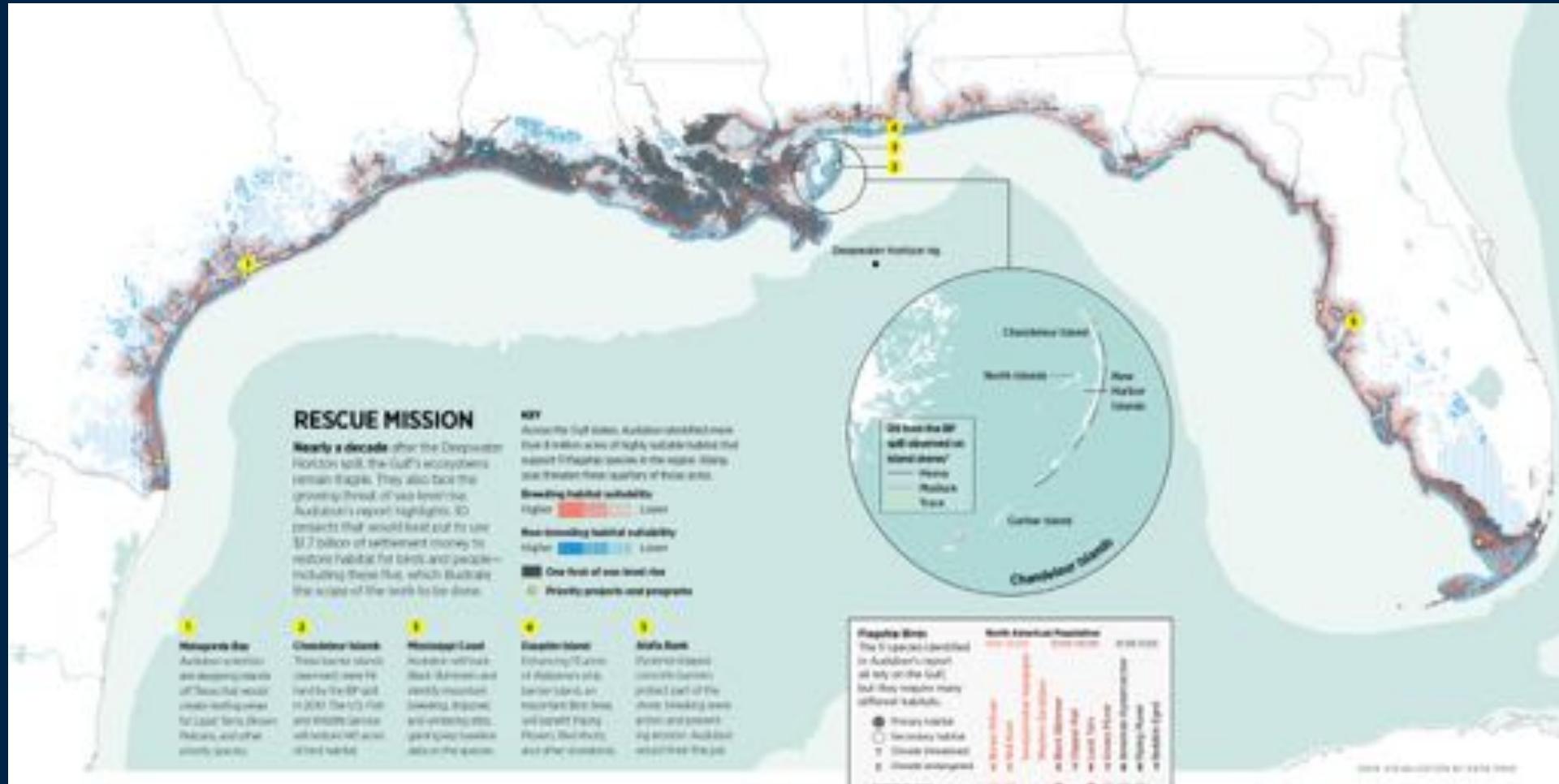
6 WHAT COLOR DOES

- DRAWS ATTENTION
- CONVEYS IMPORTANCE
- ENCODES CATEGORY
- CREATES A FEELING
- STRENGTHENS A VISUAL HIERARCHY

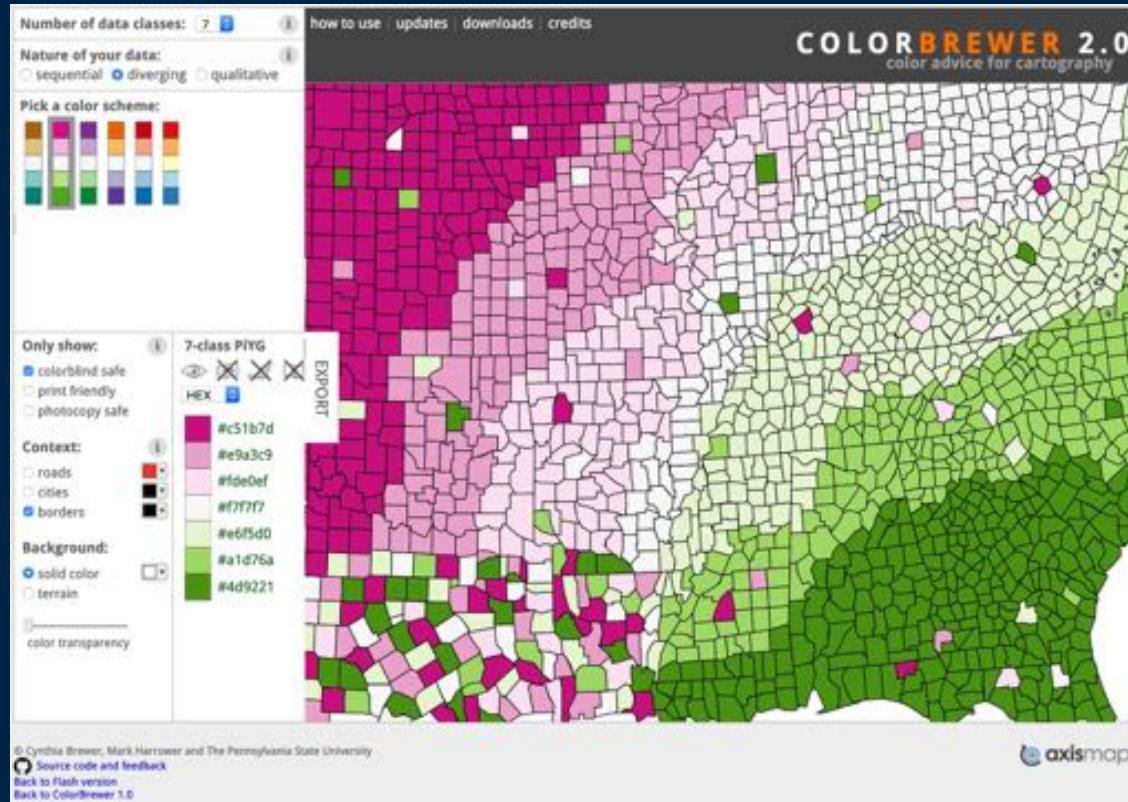


6 WHAT COLOR DOES

AUDUBON Spring 2019



6 COLOR BREWER



ColorBrewer2.org ❤️

6 COLOR CAUTIONS

- COLORBLIND PEOPLE ARE EVERYWHERE (~1 IN 20)
- PEOPLE PRINT THINGS SOMETIMES (IN B&W, NO LESS!)
- NON-ADJACENT COLORS CAN BE DIFFICULT TO COMPARE
- DIFFERENT PEOPLE JUDGE COLOR DIFFERENTLY



ColorOracle.org





7 FONTS

7 CHOOSING FONTS

SERIF

Garamond 1234567890AaBbCcDdEeFfGgHhIiJjKkLlMmNnOoPpQqRrSsTtUuVvWwXxYyZz

Georgia 1234567890AaBbCcDdEeFfGgHhIiJjKkLlMmNnOoPpQqRrSsTtUuVvWwXxYyZz

Times 1234567890AaBbCcDdEeFfGgHhIiJjKkLlMmNnOoPpQqRrSsTtUuVvWwXxYyZz

SANS SERIF

Arial 1234567890AaBbCcDdEeFfGgHhIiJjKkLlMmNnOoPpQqRrSsTtUuVvWwXxYyZz

Trebuchet 1234567890AaBbCcDdEeFfGgHhIiJjKkLlMmNnOoPpQqRrSsTtUuVvWwXxYyZz

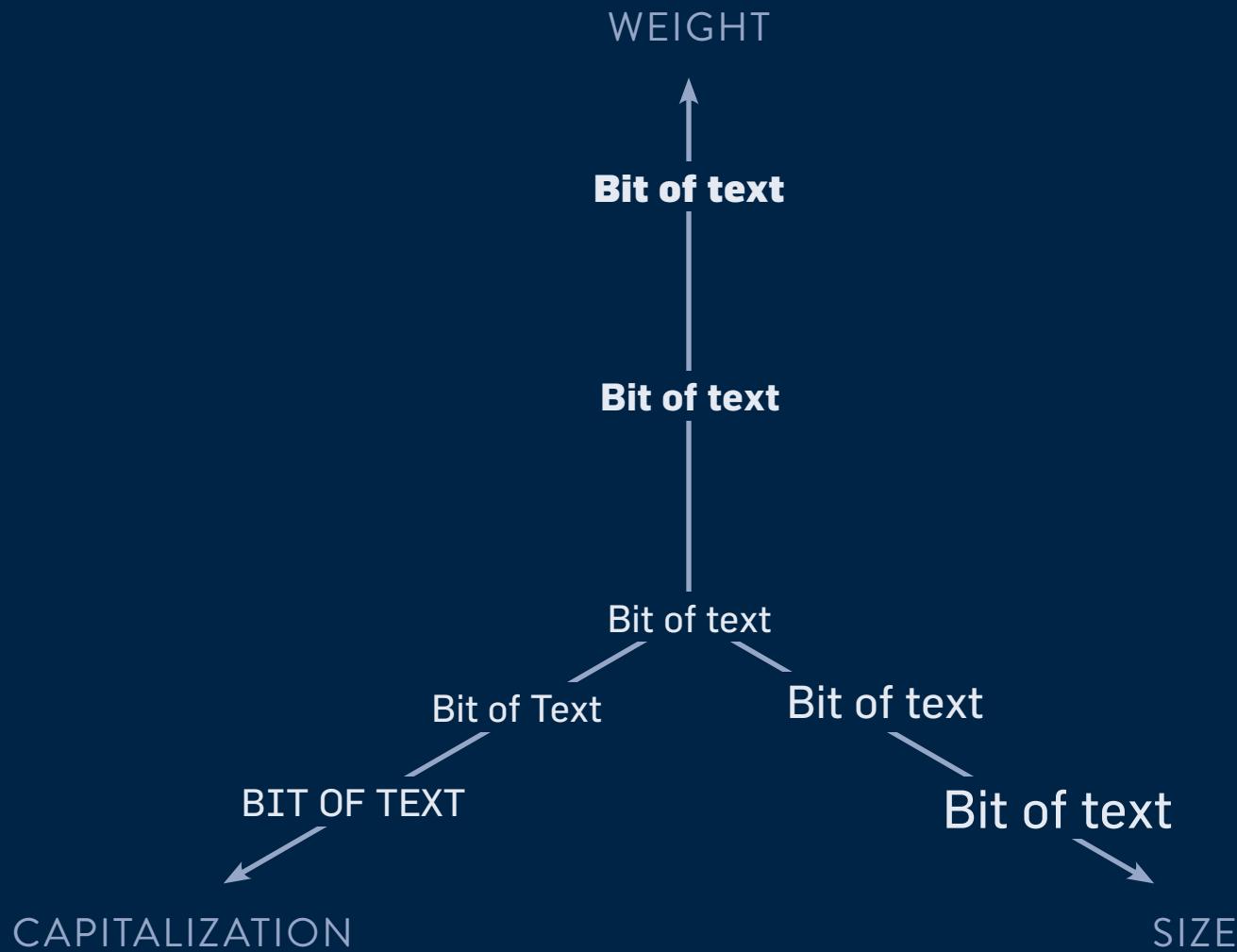
Verdana 1234567890AaBbCcDdEeFfGgHhIiJjKkLlMmNnOoPpQqRrSsTtUuVvWwXxYyZz

MONOSPACE

💖Courier 1234567890AaBbCcDdEeFfGgHhIiJjKkLlMmNnOoPpQqRrSsTtUuVvWwXxYyZz

Monaco 1234567890AaBbCcDdEeFfGgHhIiJjKkLlMmNnOoPpQqRrSsTtUuVvWwXxYyZz

7 TEXT HIERARCHY





8 ERROR BARS

8 ERROR BAR USAGE

- BE SUPER CLEAR ABOUT WHAT THEY SHOW
- FOR AN AUDIENCE OF GENERALISTS, CONSIDER USING A 95% CONFIDENCE INTERVAL

POINTS OF SIGNIFICANCE

Error bars

The meaning of error bars is often misinterpreted, as is the statistical significance of their overlap.

Last month in Points of Significance, we showed how samples are used to estimate population statistics. We emphasized that, because of chance, our estimates had an uncertainty. This month we focus on how uncertainty is represented in scientific publications and reveal several ways in which it is frequently misinterpreted.

The uncertainty in estimates is customarily represented using error bars. Although most researchers have seen and used error bars, misconceptions persist about how error bars relate to statistical significance. When asked to estimate the required separation between two points with error bars for a difference at significance $P = 0.05$, only 22% of respondents were within a factor of 2 (ref. 1).

In light of the fact that error bars are meant to help us assess the significance of the difference between two values, this observation is disheartening and worrisome.

Here we illustrate error bar differences with examples based on a simplified situation in which the values are means of independent (unrelated) samples of the same size and drawn from normal populations with the same spread. We calculate the significance of the difference in the sample means using the two-sample *t*-test and report it as the familiar *P* value. Although reporting the exact *P* value is preferred,

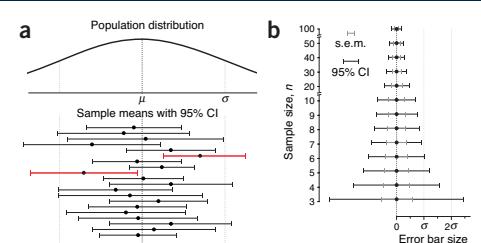
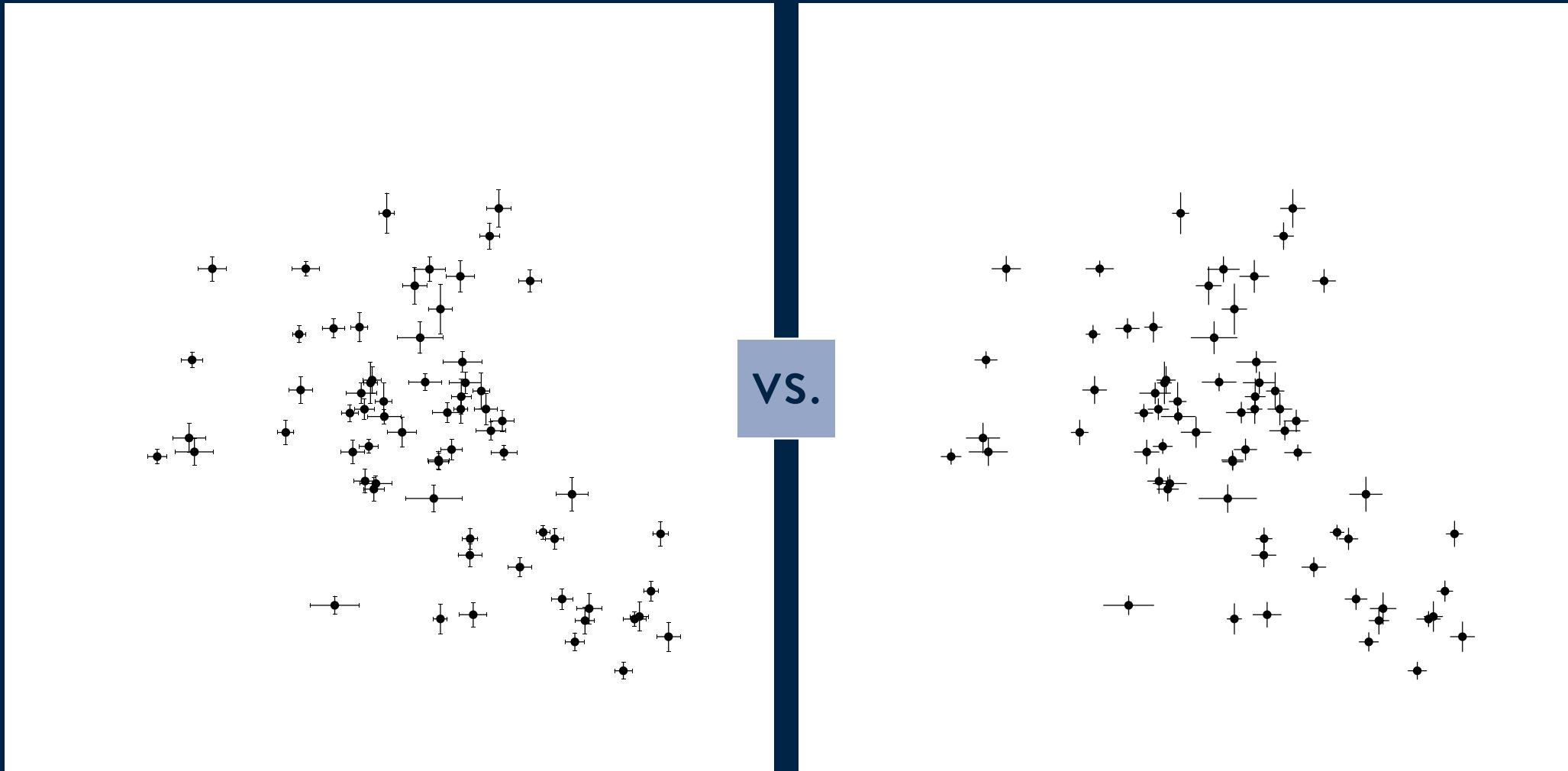


Figure 2 | The size and position of confidence intervals depend on the sample. On average, 95% of intervals are expected to span the mean—about 19 in 20 times for 95% CI. (a) Means and 95% CIs of 20 samples ($n = 10$) drawn from a normal population with mean μ and s.d. σ . By chance, two of the intervals (red) do not capture the mean. (b) Relationship between s.e.m. and 95% CI error bars with increasing n .

does not ensure significance, nor does overlap rule it out—it depends on the type of bar. Chances are you were surprised to learn this unintuitive result.

The first step in avoiding misinterpretation is to be clear about which measure of uncertainty is being represented by the error bar. In 2012, error bars appeared in *Nature Methods* in about two-thirds of the figure panels in which they could be expected (scatter and bar plots). The type of error bars was nearly evenly split between s.d. and s.e.m. bars (45% versus 49%, respectively). In 5% of cases the error bar type was not specified in the legend. Only one figure² used bars based on the 95% CI. CIs are a more intuitive measure of uncertainty and are popular in the medical literature.

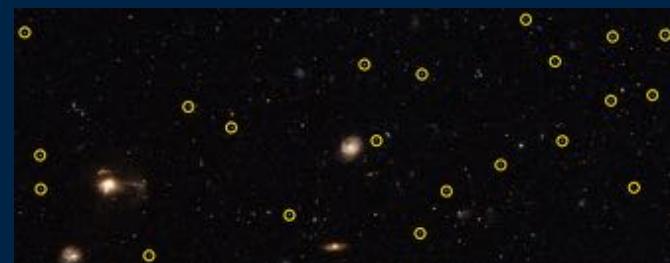
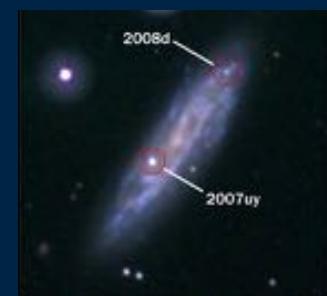
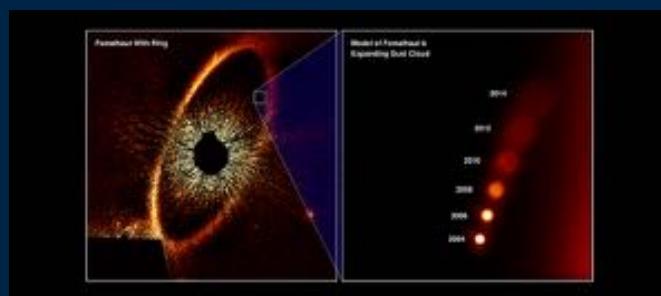
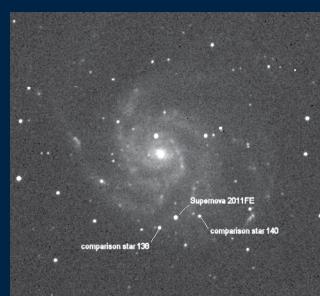
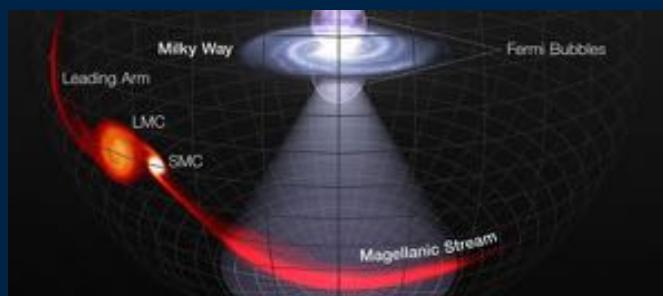
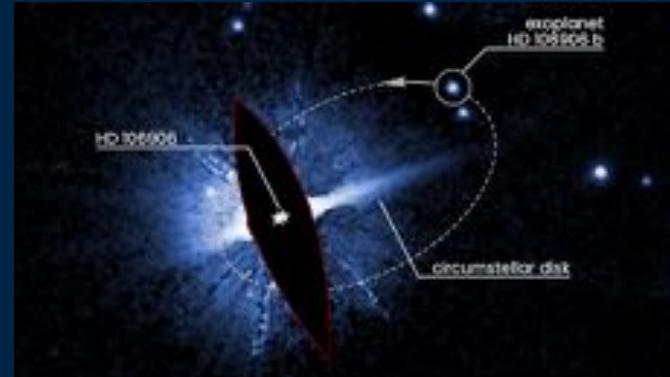
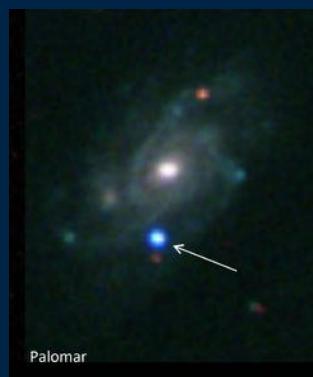
8 ERROR BAR STYLES





9 ANNOTATIONS

9 ANNOTATIONS 😍



9 SIMPLE LABELS, COMPLEX EXPLANATION



9 ANNOTATION IS POWERFUL

- THE ANNOTATION LAYER IS YOUR CHANCE TO CALL SOMETHING OUT EXPLICITLY
- DIRECT LABELS ARE QUICKER TO PARSE THAN A KEY OR SENTENCES IN THE CAPTION
- USE ANNOTATION TO MAKE YOUR POINT ABUNDANTLY CLEAR—MAYBE EXCESSIVELY SO!
- HELP YOUR READER SEE WHAT YOU SEE



10 THE CAPTION

10 WHEN TO DRAFT A CAPTION?

CAPTION AS
AFTERTHOUGHT

VS.

CAPTION AS AN
INTEGRATED PART OF
THE COMMUNQUÉ.

10 CAPTION CRAFT

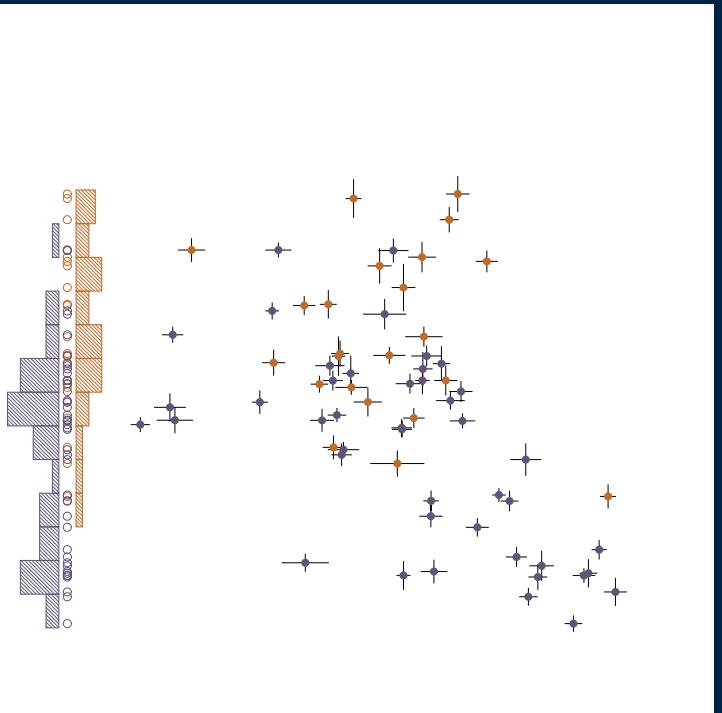


Fig. 3 Stars of the main sample, plotted as brightness versus distance. Error bars, in black, represent are one-sigma errors. Stars with oranger properties are in orange, stars in purpler properties are in purple.

VS.

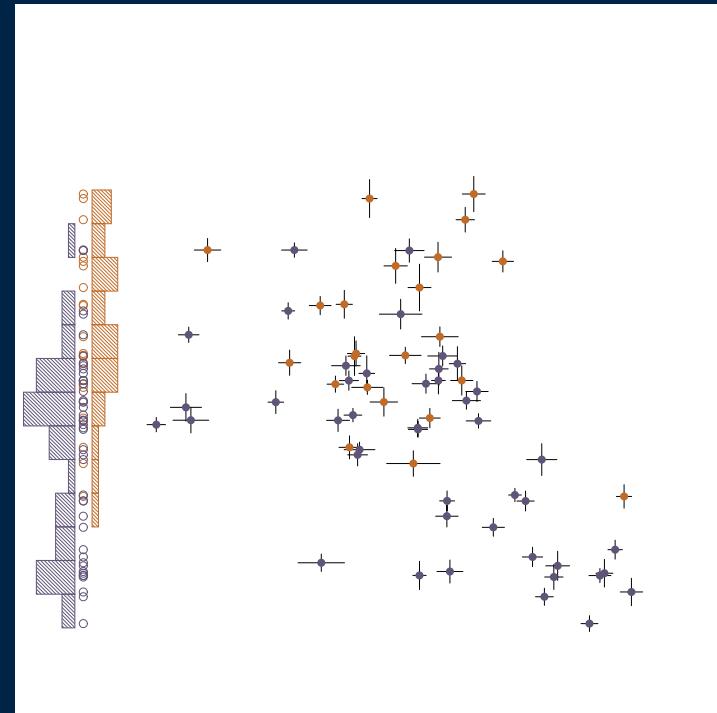


Fig. 3 Stars fall into two main populations: one (orange dots) is oranger, and the other (purple dots) is purpler. Dots here represent the mean value of a single star, with measurement errors displayed as black bars.

10 THE FUNCTION OF A CAPTION

- USE THE CAPTION SPACE TO YOUR ADVANTAGE! MAKE THE POINT YOU WANT TO BE SURE YOUR READER GETS.
- CAPTION ALSO NEEDS TO CARRY WEIGHT OF EXPLAINING TECHNICAL DETAILS. INCLUDE THEM, OF COURSE!
- BUT MAYBE DON'T LEAD WITH THOSE DETAILS. HELP THE READER FIND THE KEY TAKEAWAY EASILY.



FIN