HCI
Lecture 10

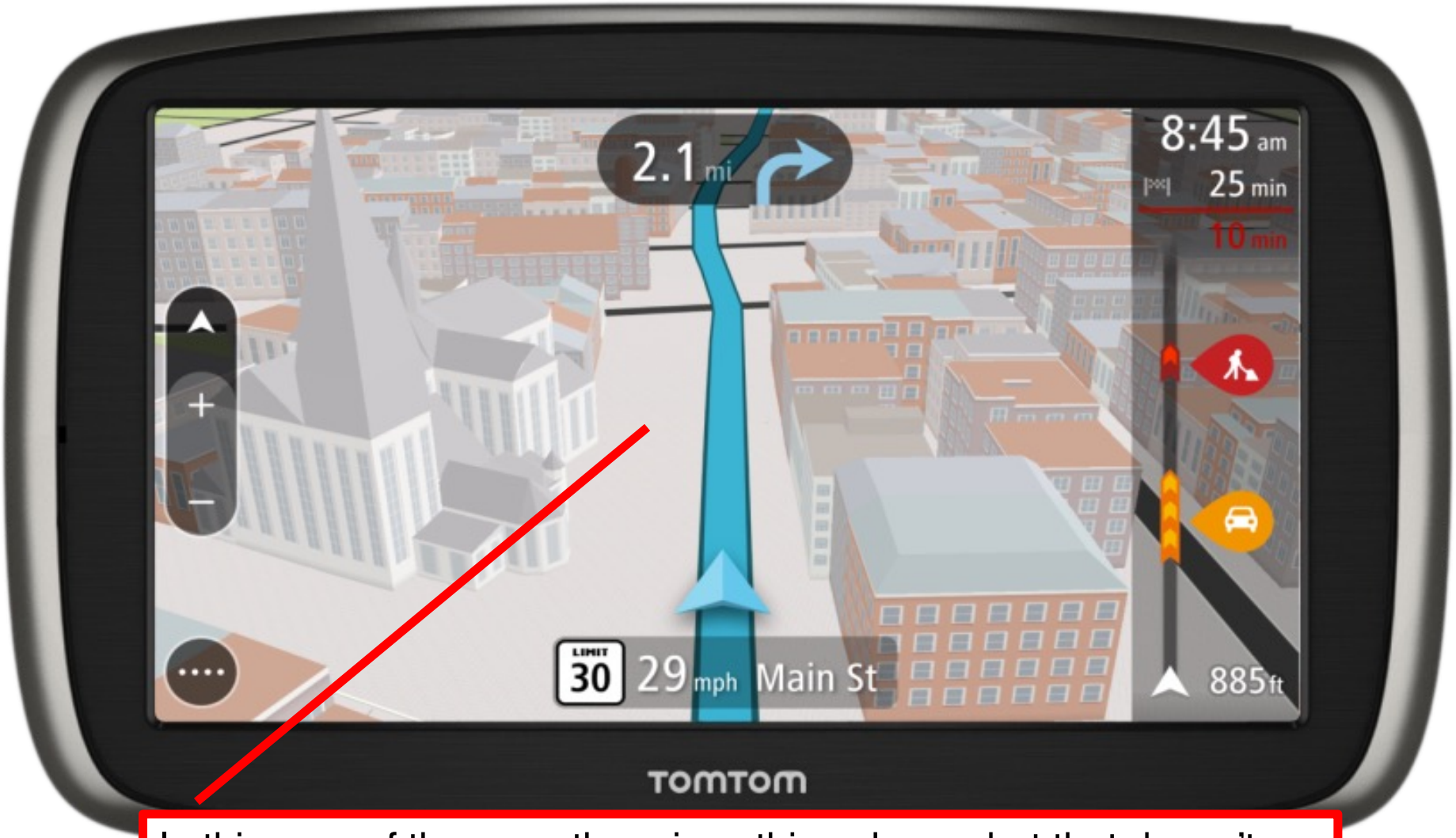
Mario Verdicchio
Università degli Studi di Bergamo
Anno Accademico 2024-2025

Models

- A **model** is a set of signs and relationships among them that describes, with varying degrees of accuracy, an aspect of reality.
- Based on this description, observers can construct explanations or predictions about the aspect of reality being described.
- Typically, the description of a model is a simplification of reality that omits elements unnecessary for subsequent explanations and predictions, while highlighting the elements considered important for these processes.

Visualizations are models

- All visual representations are models.
- Maps, for example, are models.
- As such, they leave out details that aren't needed for the purpose they were created for.
- A GPS map, for instance, doesn't tell you anything about the condition of the bushes along the roadside.



In this area of the map, there is nothing shown, but that doesn't mean there is an empty space in reality. It simply means that, according to the creators of this visualization, there are no elements in that part of the city that are useful for car navigation.

Omission as an advantage

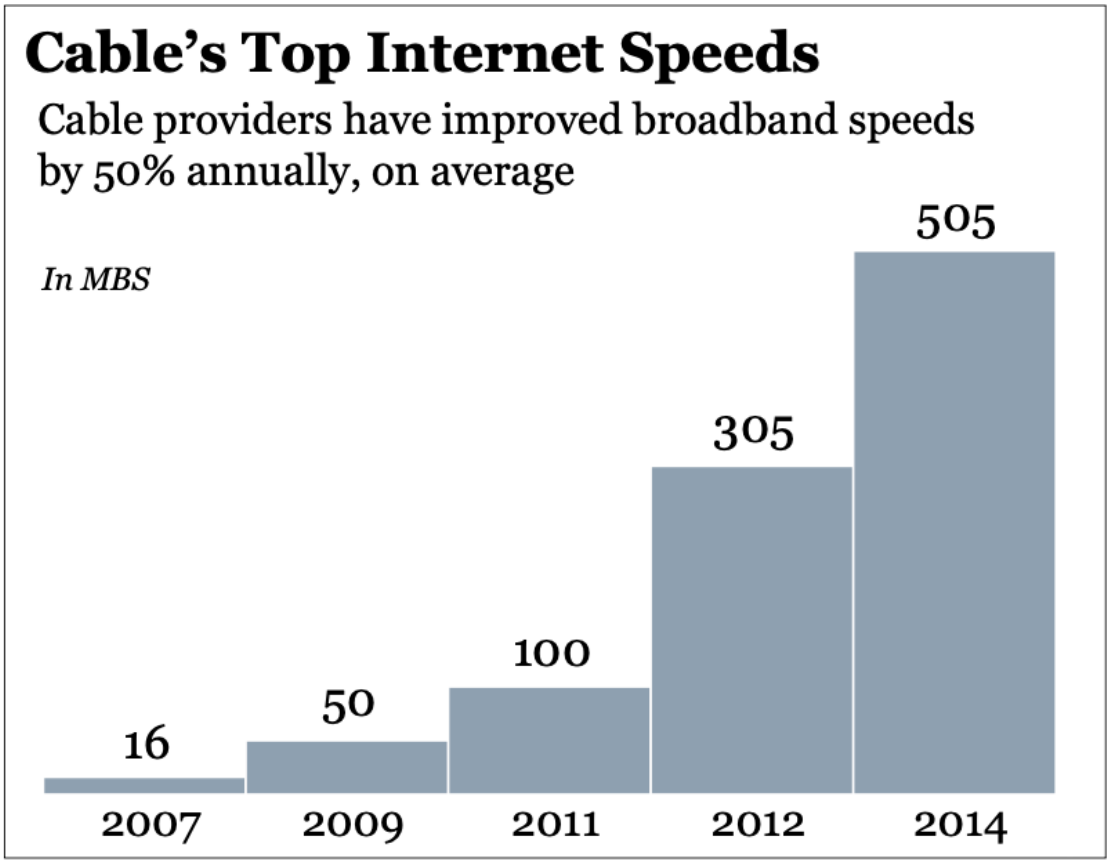
- Omitting elements allows viewers of the visualization to focus on what is important.
- A model is better when it closely represents what it is meant to show without being unnecessarily complex, and when it is easy for its viewers to interpret correctly.

Omission as a disadvantage

- Warning: it is the creators of the visualization who decide which elements to include and which to omit.
- Due to incompetence or worse, malicious intent, creators can produce misleading visualizations that give viewers a representation of reality that is not merely imprecise, but false.

NCTA, again

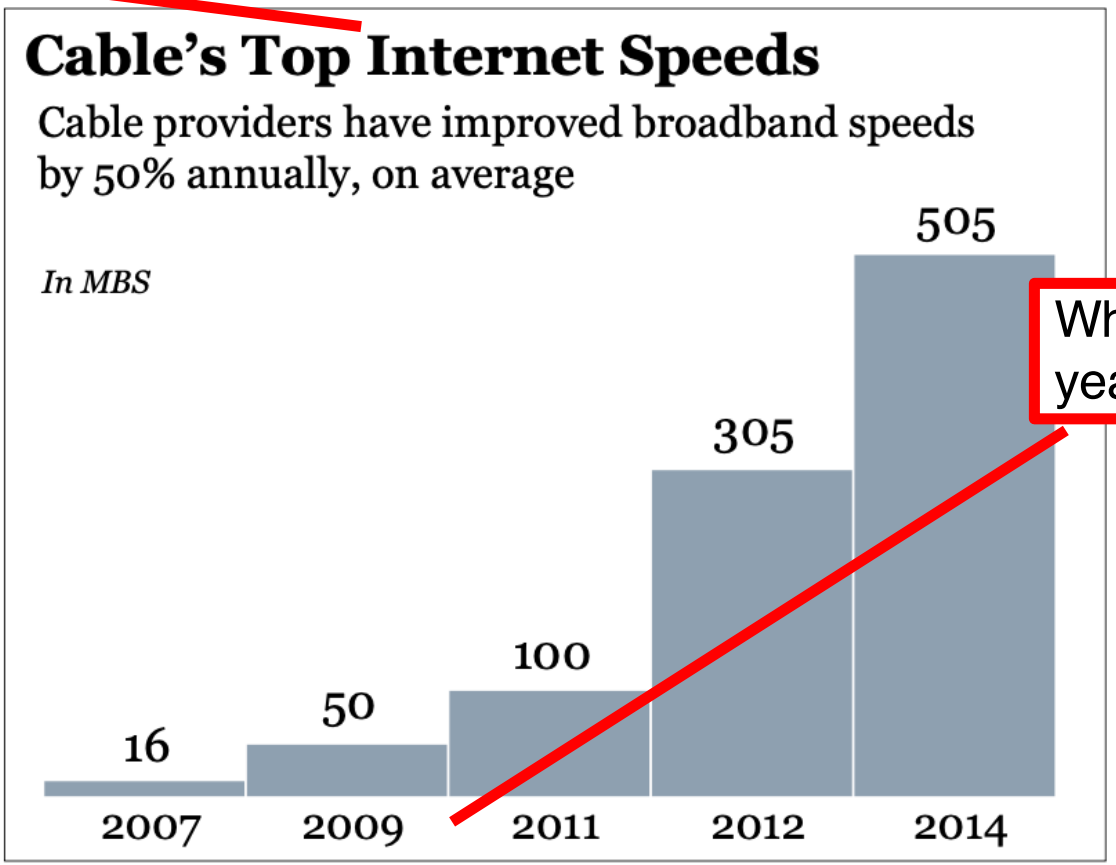
- The National Cable & Telecommunications Association provides yet another negative example



From "The Truthful Art" by A. Cairo

These are the maximum speeds. Very well. But:

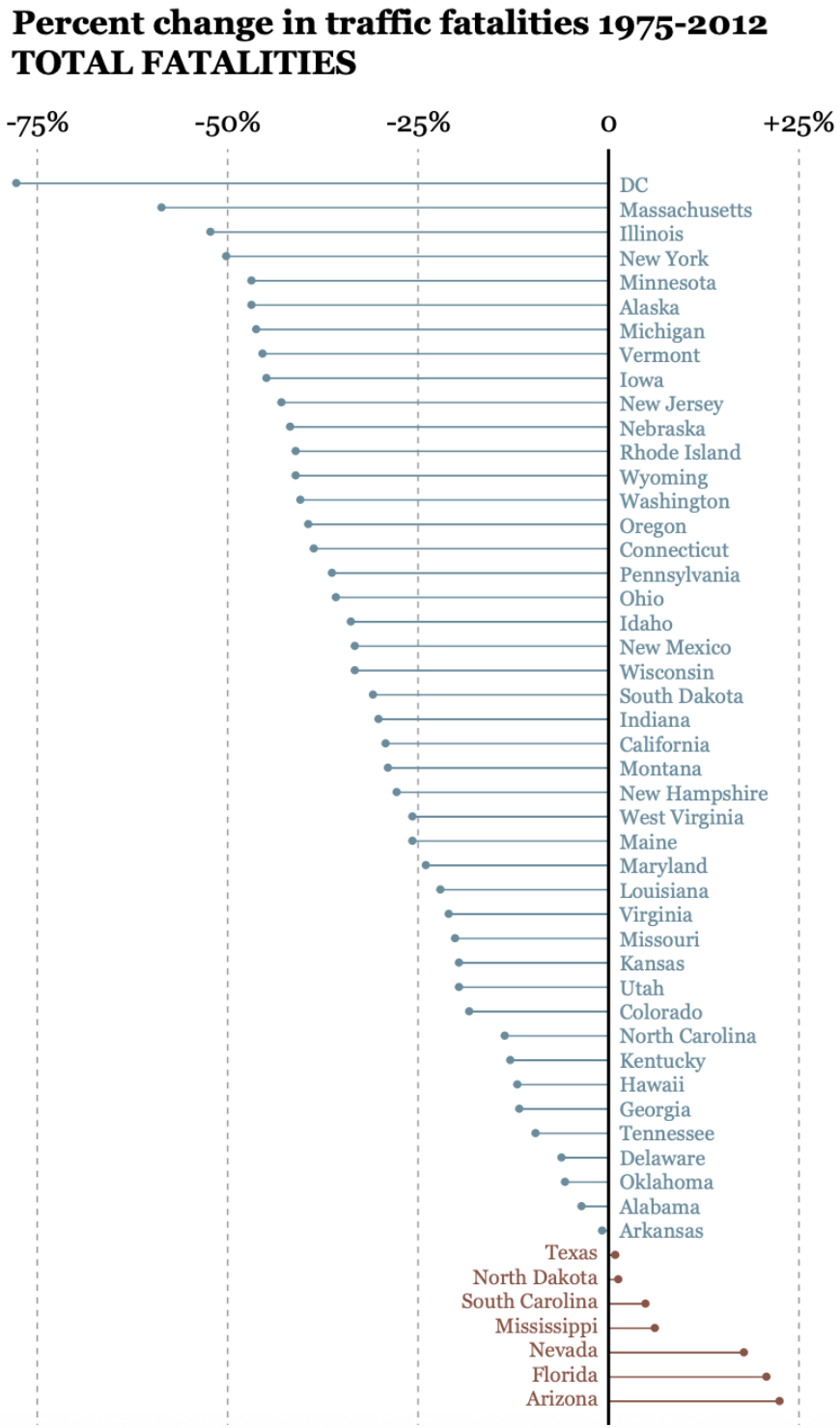
- What are the minimum speeds?
- What are the average speeds?
- How many people have access to the maximum speed?
- How is the speed distributed geographically?



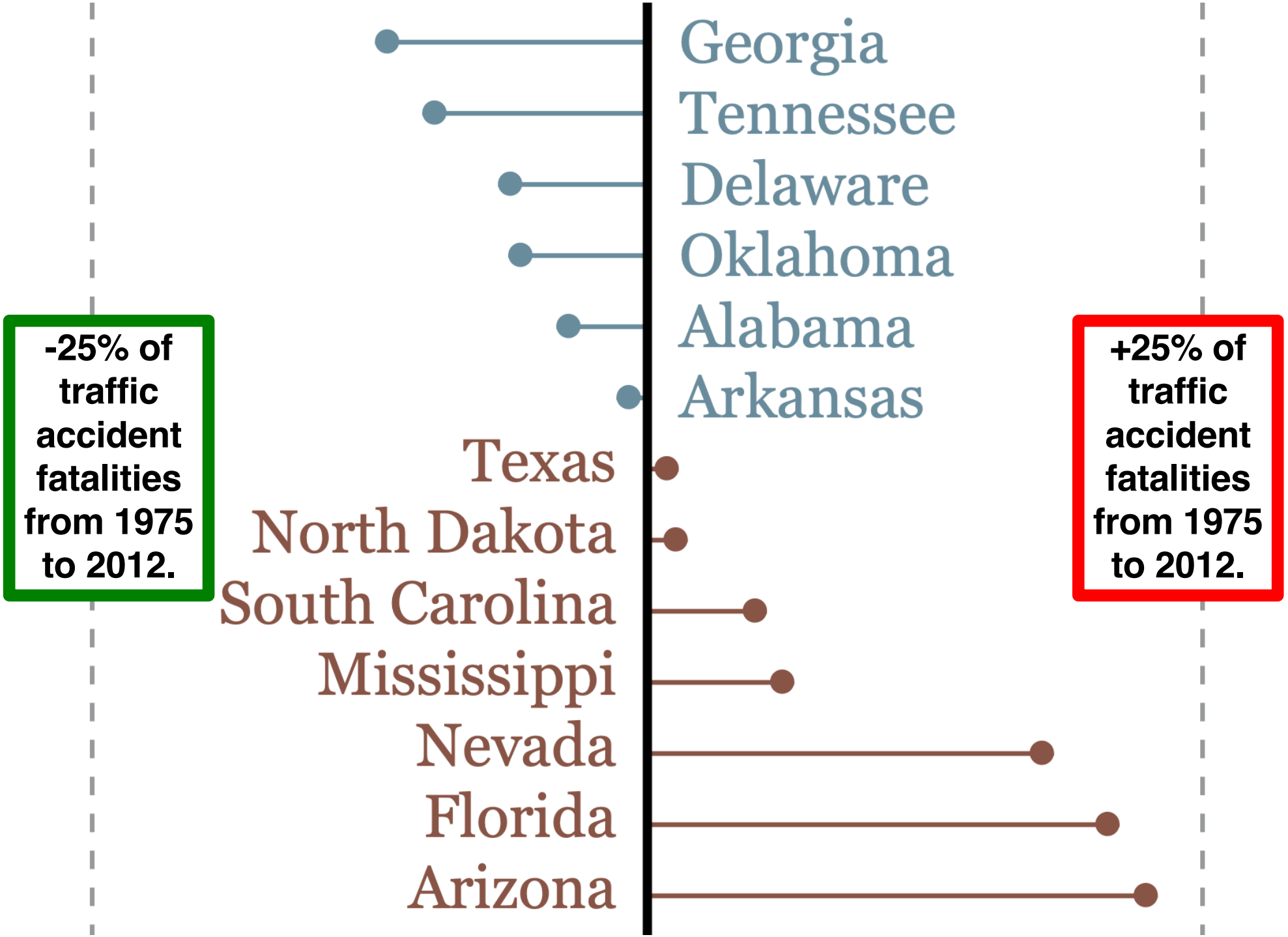
Why are some years missing?

Even those with good intentions…

- …can provide misleading visualizations.
- The risk is that useful policies may be developed, but not in the areas where they are most needed.
- Here is an example from the NHTSA (National Highway Traffic Safety Administration) on traffic accident fatalities between 1975 and 2012 across different U.S. states.



From “The Truthful Art”
by A. Cairo


-25% of traffic accident fatalities from 1975 to 2012.
+25% of traffic accident fatalities from 1975 to 2012.
Georgia
Tennessee
Delaware
Oklahoma
Alabama
Arkansas
Texas
North Dakota
South Carolina
Mississippi
Nevada
Florida
Arizona

Is this visualization truthful?

- It doesn't present false data (although the data sources should still be verified).
- However, it leads the viewer to draw incorrect conclusions—for example, thinking that the road fatality situation in Florida has significantly worsened.
- Why is that reasoning incorrect? Because essential data is missing for a proper interpretation of road fatality statistics:
  - Has the state's population increased in the meantime?
  - Has traffic in the state increased over the same period?
- Absolute values do have meaning, but when comparing two different time periods, it's necessary to account for other variables that make those values relative to the context they refer to.
- Otherwise, it would be like comparing the purchasing power of one dollar in 2012 with that of one dollar in 1975.

What Food a Dollar Could Buy the Year

1975: 1 pound of Chock full o'Nuts coffee

CHOCK FULL O'NUTS® COFFEE HEAVENLY ORIGINAL, MEDIUM ROAST, GROUND

$5.00

(54) **Reviews** | Write a Review

Size and Packaging

11.3 oz Can

Qty

– 1 +

11,3 oz are 0,7 pounds

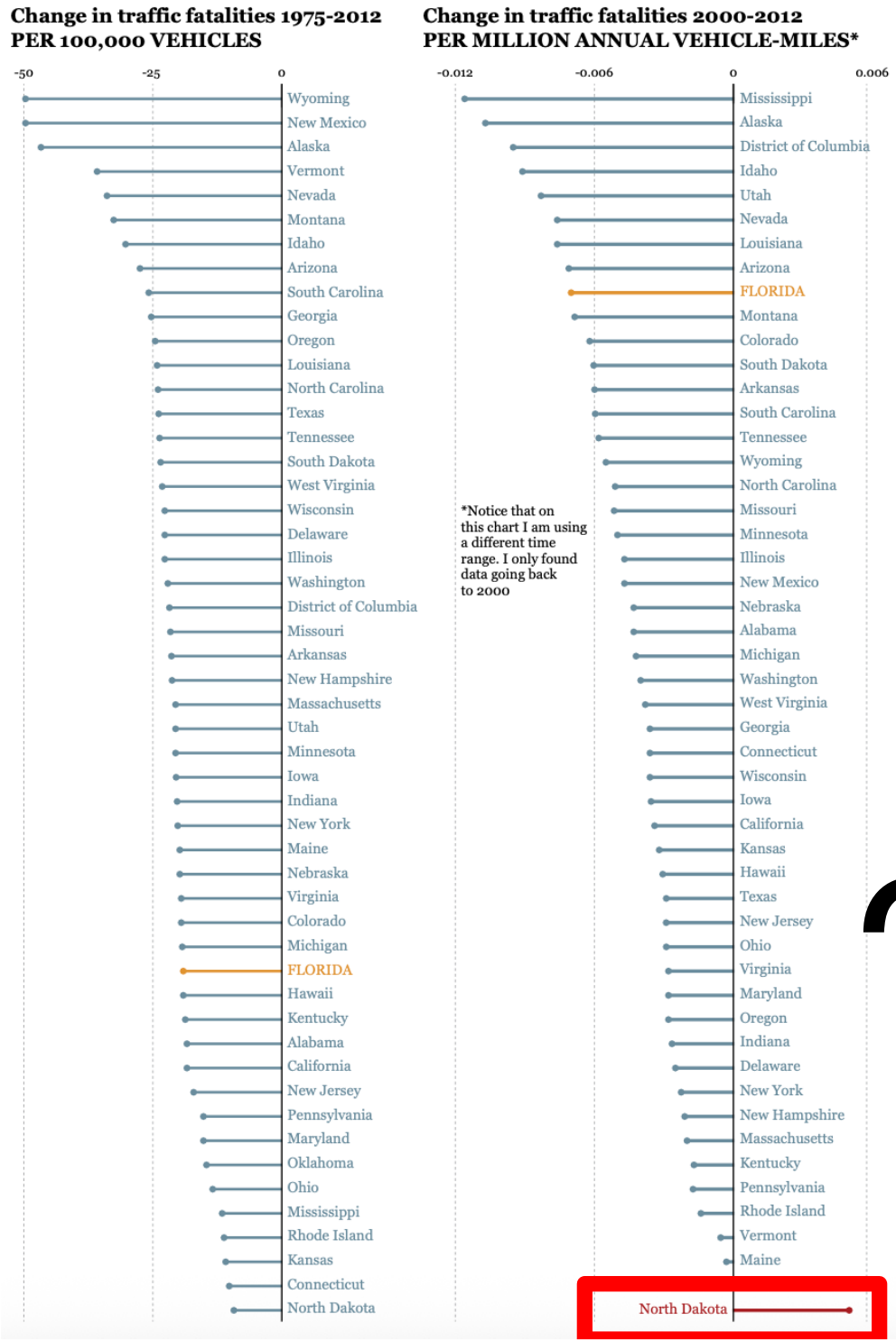
If 0,7 pounds cost 5 dollars, it means that 1 pound costs more than 7 dollars.

Surprise! When we take population data into account—specifically, the number of vehicles—not only do all states show improvement (with dozens fewer fatalities per 100,000 vehicles), but Florida also moves up a few positions in the ranking.

Watch out: the number of cars is a meaningful variable, but it doesn't tell us how much those cars are actually used by the population (they could be cars driven only 100 km a month by people who don't commute).

Even more significant is the data on fatalities relative to the distance travelled by car.

When looking at the change in the number of fatalities per one million miles travelled, Florida performs even better and ranks in the top 10 for improved road safety.


Change in traffic fatalities 1975-2012
PER 100,000 VEHICLES
-50
-25
0
Wyoming
New Mexico
Alaska
Vermont
Nevada
Montana
Idaho
Arizona
South Carolina
Georgia
Oregon
Louisiana
North Carolina
Texas
Tennessee
South Dakota
West Virginia
Wisconsin
Delaware
Illinois
Washington
District of Columbia
Missouri
Arkansas
New Hampshire
Massachusetts
Utah
Minnesota
Iowa
Indiana
New York
Maine
Nebraska
Virginia
Colorado
Michigan
FLORIDA
Hawaii
Kentucky
Alabama
California
New Jersey
Pennsylvania
Maryland
Oklahoma
Ohio
Mississippi
Rhode Island
Kansas
Connecticut
North Dakota
Change in traffic fatalities 2000-2012
PER MILLION ANNUAL VEHICLE-MILES*
-0.012
-0.006
0
0.006
*Notice that on this chart I am using a different time range. I only found data going back to 2000
Mississippi
Alaska
District of Columbia
Idaho
Utah
Nevada
Louisiana
Arizona
FLORIDA
Montana
Colorado
South Dakota
Arkansas
South Carolina
Tennessee
Wyoming
North Carolina
Missouri
Minnesota
Illinois
New Mexico
Nebraska
Alabama
Michigan
Washington
West Virginia
Georgia
Connecticut
Wisconsin
Iowa
California
Kansas
Hawaii
Texas
New Jersey
Ohio
Virginia
Maryland
Oregon
Indiana
Delaware
New York
New Hampshire
Massachusetts
Kentucky
Pennsylvania
Rhode Island
Vermont
Maine
North Dakota

??

The mysterious case of Nord Dakota

North Dakota, which was already low in the rankings in the first visualization, unlike Florida, not only does not rise in the rankings, but turns out to be the only state that saw an increase in deaths per million miles travelled from 2000 to 2012.

Image by TUBS, from Wikipedia.org

Why?

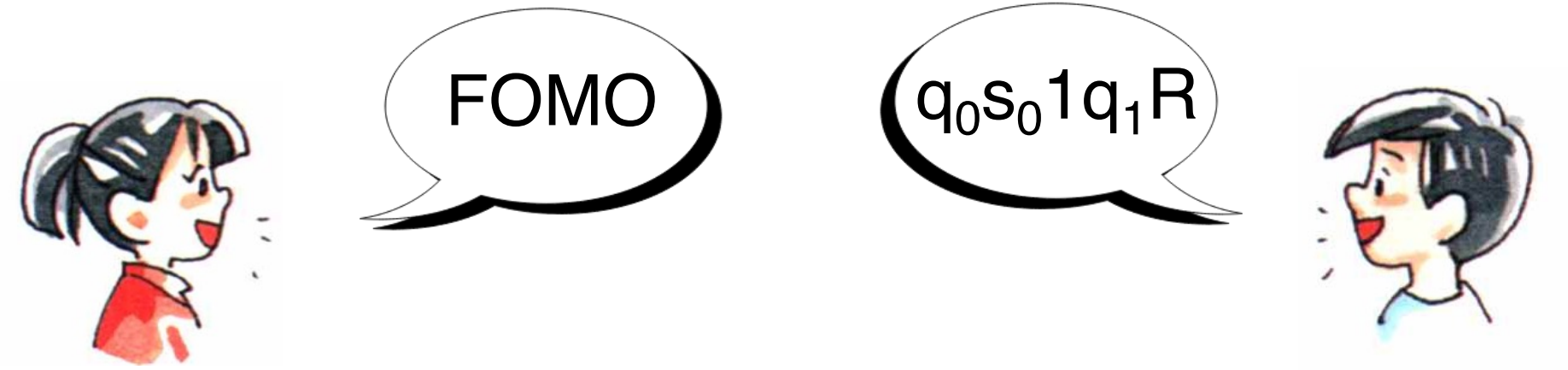
- We can only make hypotheses.
- One hypothesis is based on the oil industry boom in the state toward the end of the 2000s, as it turned out to have ideal soil for fracking. This led to:
  - an increase of 100,000 workers (from 2009 to 2014)
  - with long commuting times (because no one lives near the oil fields where they work)
- Is this hypothesis correct? Is it wrong? Only one thing is certain…

A road in North Dakota. Source: www.usnews.com

It won’t be the data on fatal accidents that give us an explanation.
It’s true that data can help the observer formulate explanations, but the causes of a phenomenon are often external to the phenomenon itself.
So, to verify the hypotheses about the increase in road deaths in North Dakota, we’ll need other data: a broader investigation is necessary.

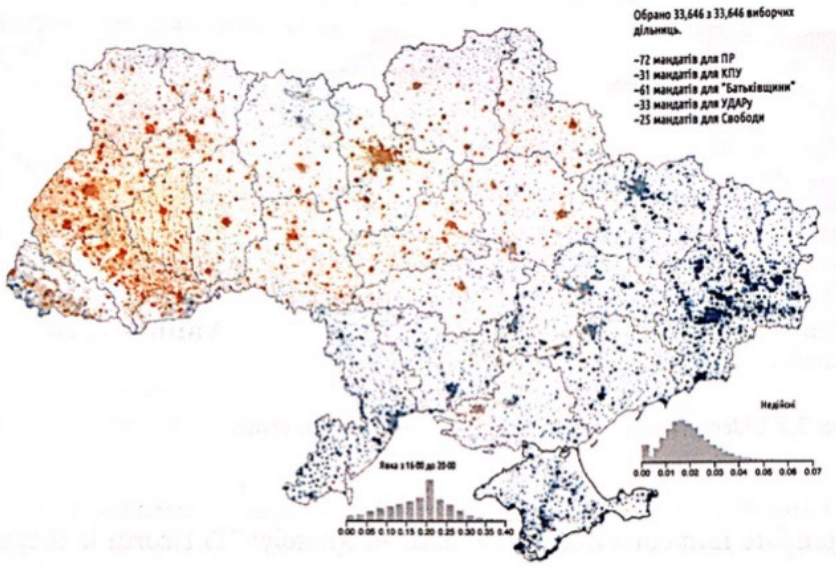
Visualizations and the human mind

- In communication, **context** is defined as the set of information that the sender and the receiver must share for the transmission of a message to be successful.



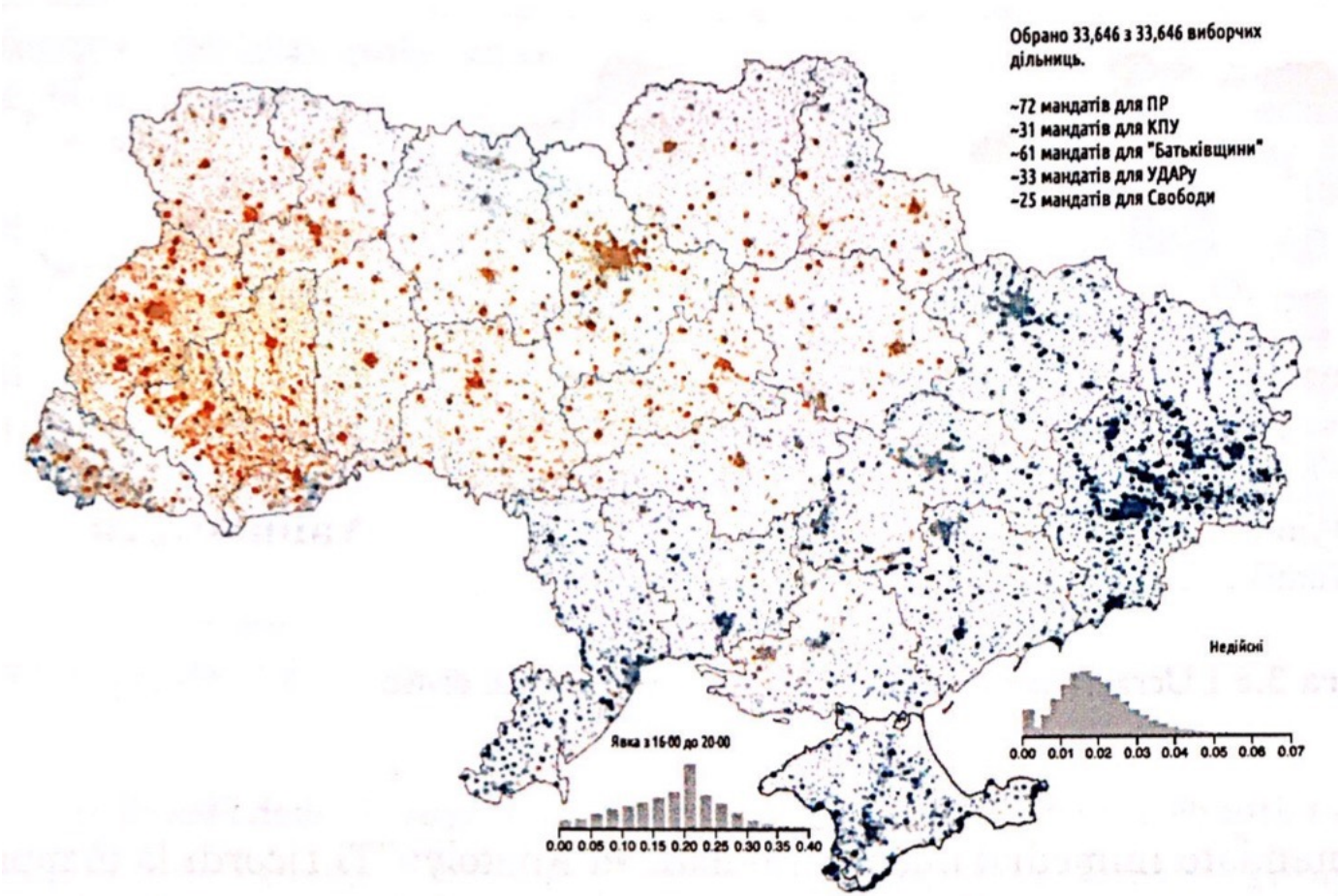
Visualizations and the human mind

- A very similar concept applies to visualizations: they are a model that acts as a bridge between the mind of the creator and the mind of the viewer.



From “The Truthful Art” by A. Cairo

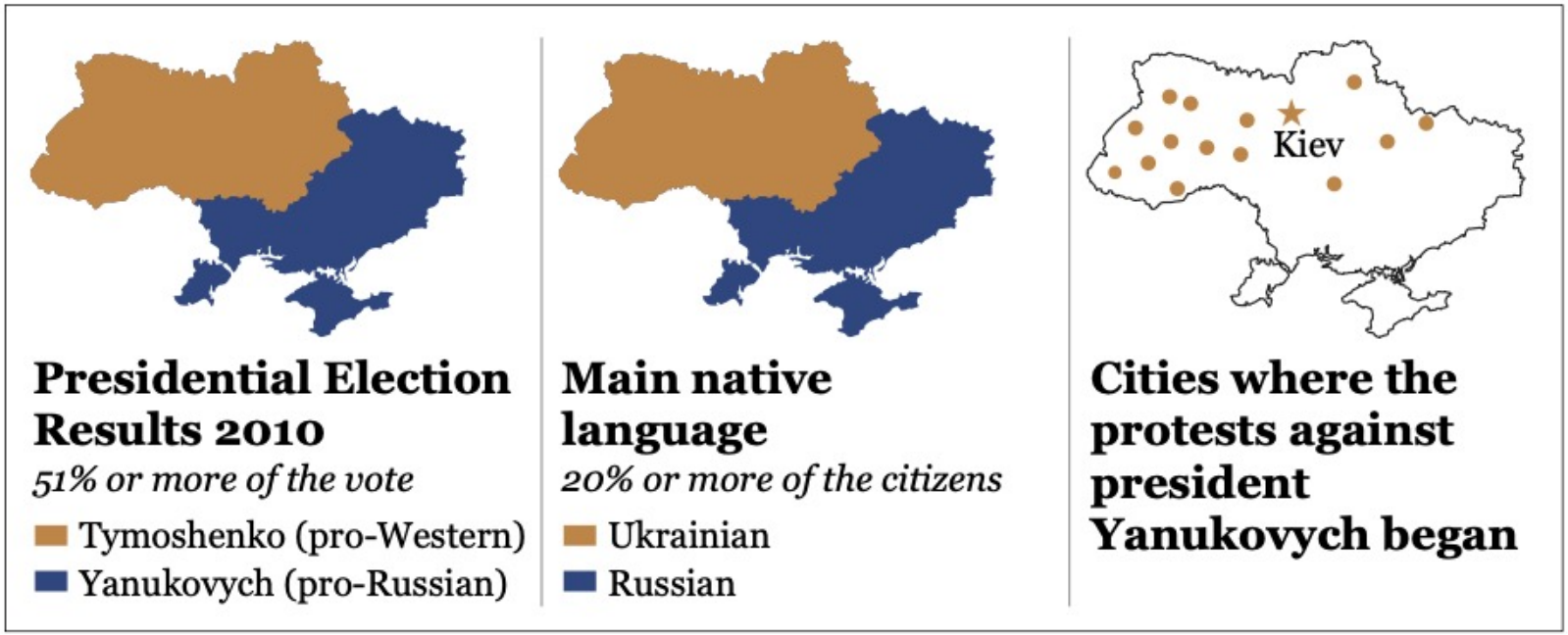
The Ukraine case: 2012



This visualization shows the results of the 2012 Ukrainian elections, which confirmed the parliamentary majority supporting the pro-Russian president Viktor Yanukovych. The circles on the maps represent electoral districts, and their size is proportional to the number of voters.

Their color is determined by the winning party: orange for the pro-Western opposition party, blue for the pro-Russian majority party.

The Ukraine case: 2013

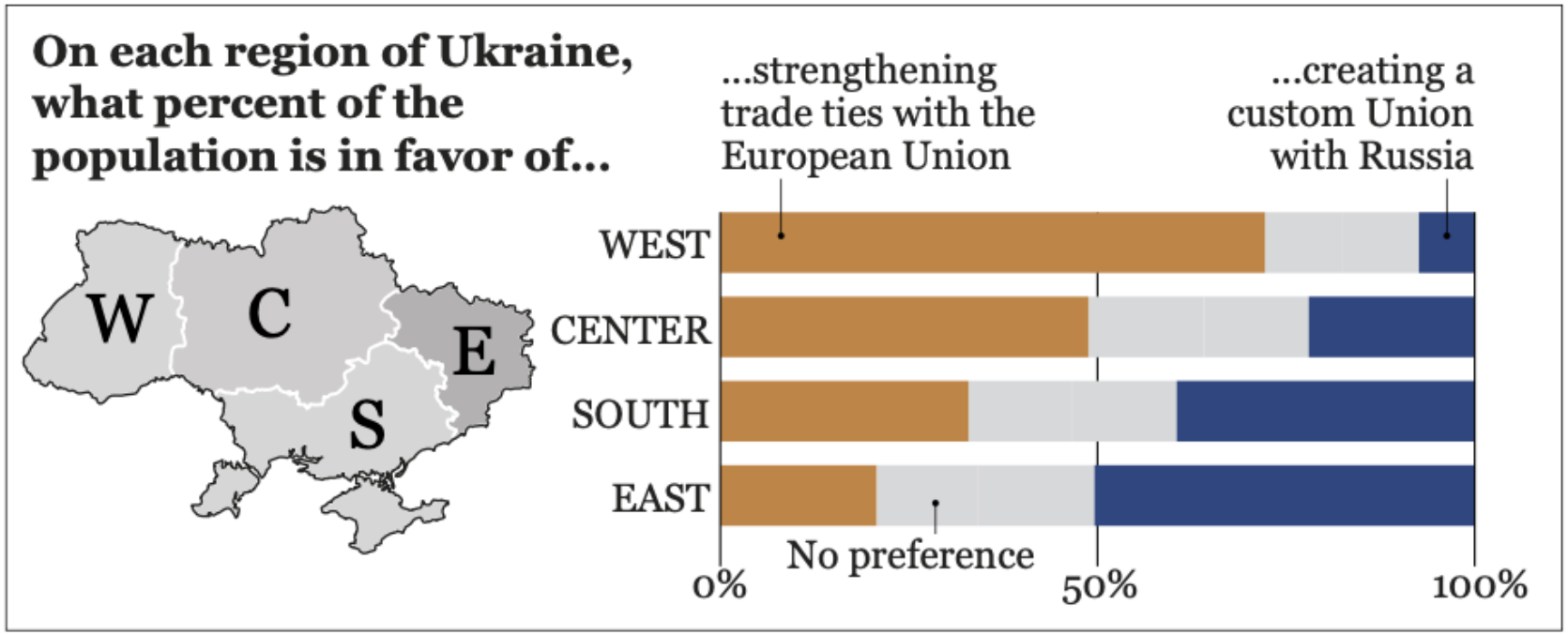


From “The Truthful Art” by A. Cairo

A year later, following the president’s refusal to make agreements with the European Union and his proposals for closer ties with Russia, protests broke out, culminating in Yanukovych’s flight to Russia and his replacement by an interim president pending new elections.

In reporting on the uprisings in Ukraine, many newspapers published visualizations like the ones above, showing a country sharply divided in two by political orientation and language.

The Ukraine case: it's complicated



From "The Truthful Art" by A. Cairo

In reality, by gathering more detailed data, the picture becomes more complicated. There is no clear division between the West and East, between pro-Europe and pro-Russia.

These orientations are clear only in specific, well-defined regions of the country. In the central and southern regions, there is a significant portion of the population that is indifferent to both options, but this group is also decisive in establishing a majority within the region.

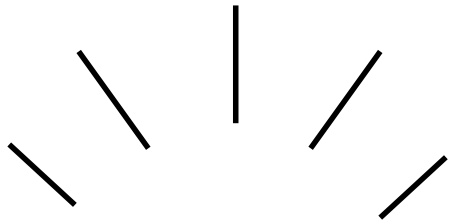
Visualizations and the human mind

- Visualizations are very helpful in selecting the significant elements of a reality to represent, and in eliminating other unimportant elements.
- However, when this simplification is taken to an extreme, it can oversimplify and provide an inaccurate representation or lead the viewer to make incorrect conclusions (like the division of Ukraine into two).
- In many cases, to clarify a topic, it is necessary to increase the information, not reduce it.

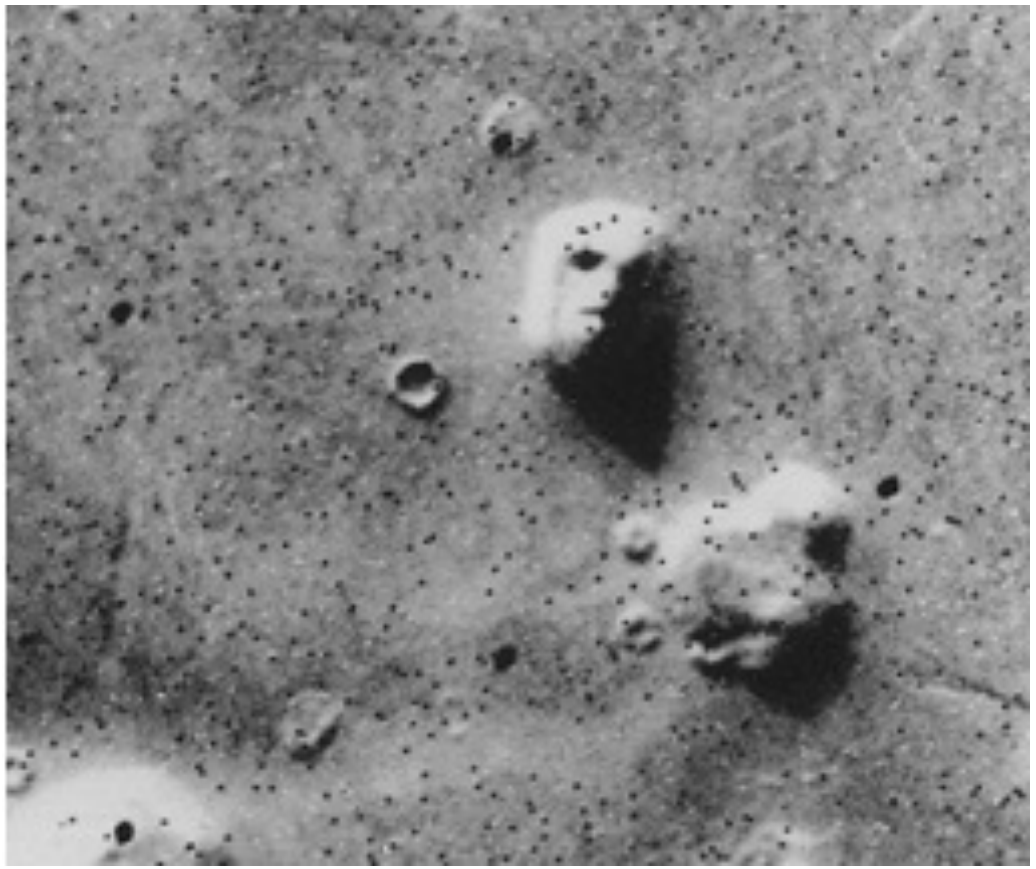
Visualizations and the human mind

- There will always be a discrepancy, albeit small, between what the designer has conceived and what the observer has understood.
- It is essential to minimize the possibility of misunderstandings, especially because the human mind has a natural tendency to make mistakes.

Apophenia, patternicity

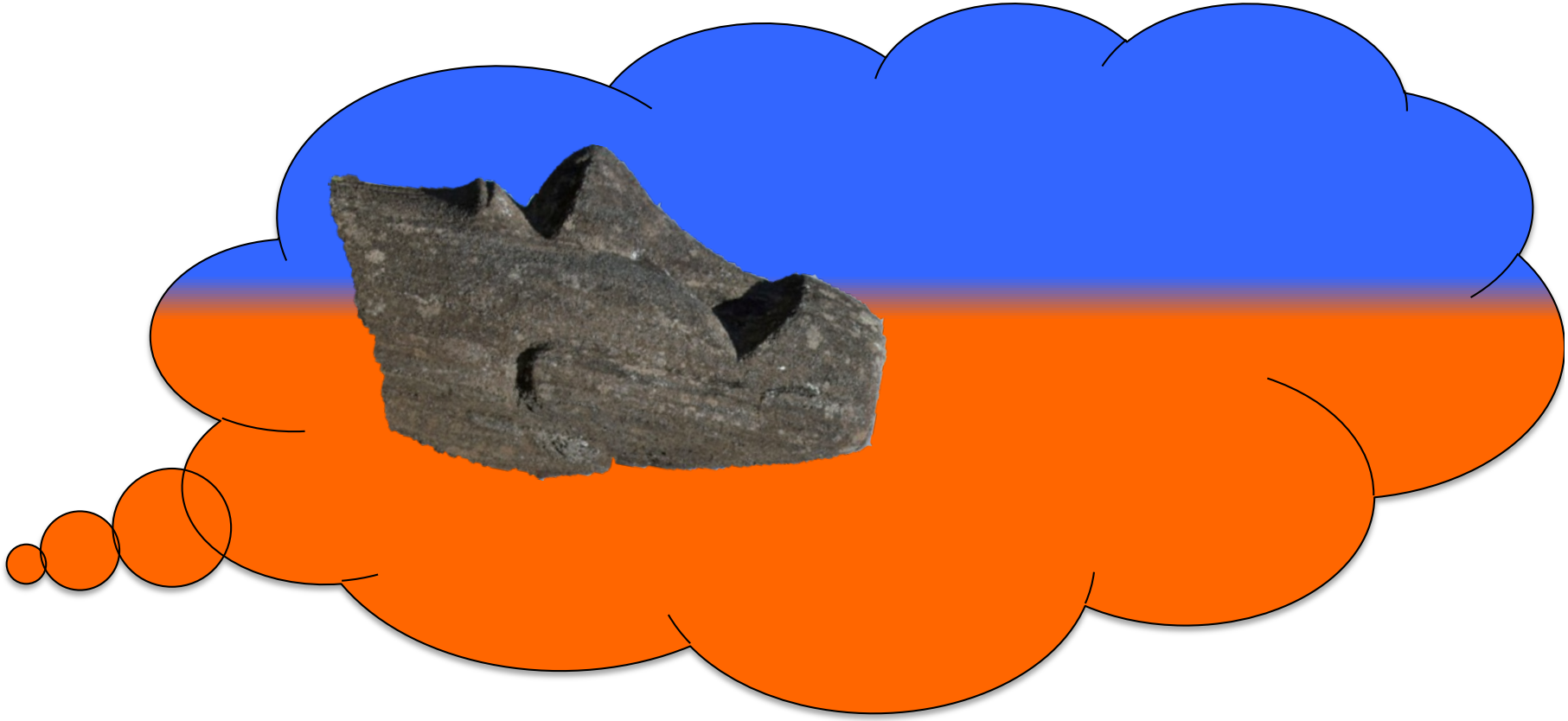


Mars surface (Viking-1, 1976). Source: NASA.



Apophenia is defined as the tendency to perceive patterns, regardless of whether they are real or not. It is commonly known also as patternicity.

Narration



There is also a tendency toward narration, meaning the inclination to immediately create a narrative to find a coherent explanation for the patterns discovered. (Caution: coherent does not necessarily mean true or realistic.)

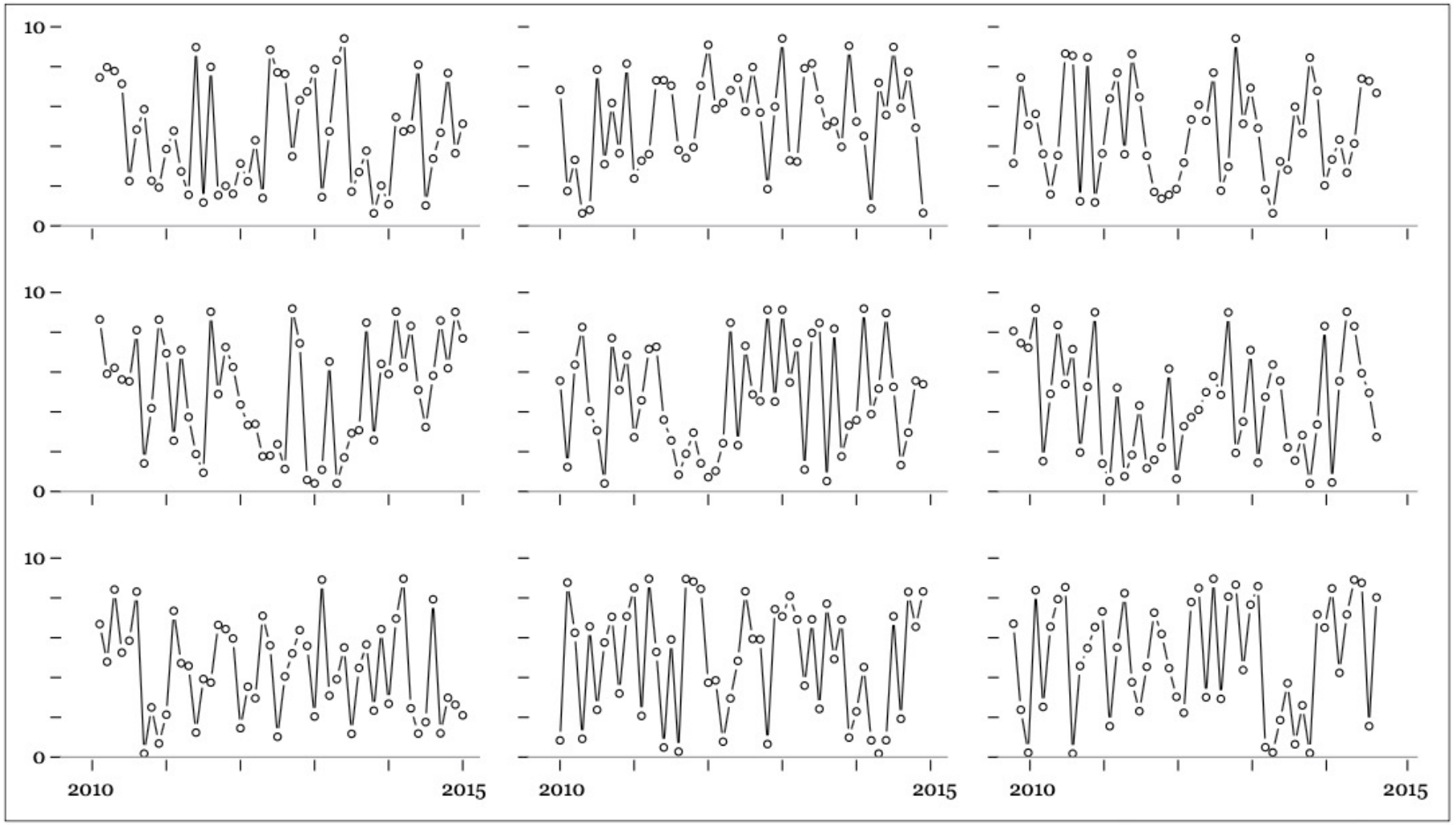
Confirmation

Confirmation bias is the tendency, once a narrative has been formed, to interpret every new piece of data as confirmation of the narrative, even those that conflict with it.

Why are we like this?

- There are no solid theories about the causes of these tendencies.
- One hypothesis is evolutionary in nature:
  - The living beings that survived were those who “saw” a dangerous predator and immediately fled.
  - Those who waited for confirmation did not survive.
- Jumping to conclusions that overestimate the risk is an advantage in the wild.
- However, jumping to conclusions often means arriving at incorrect conclusions.

Do you see patterns?

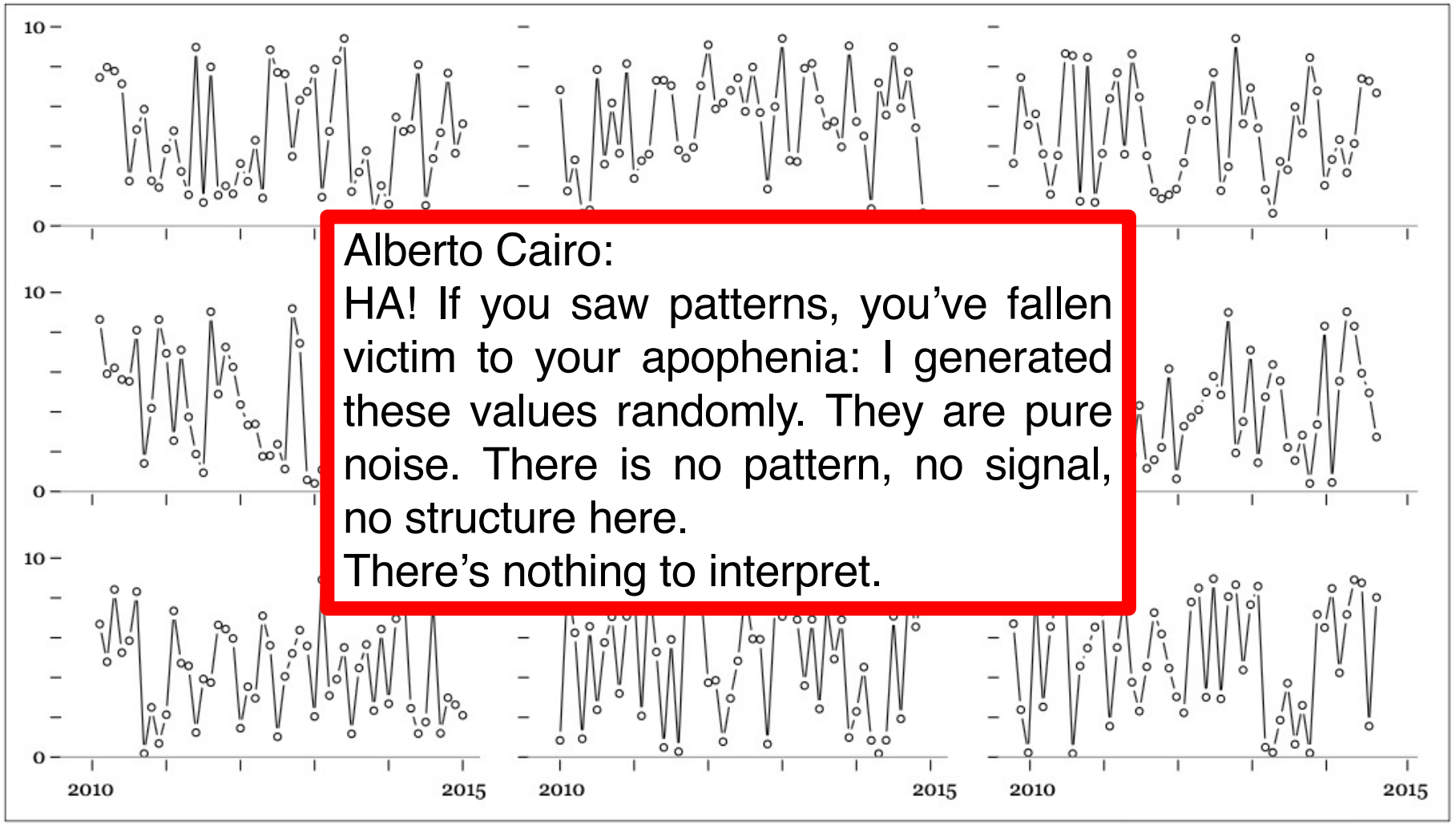


From “The Truthful Art” by A. Cairo

Pattern vs noise

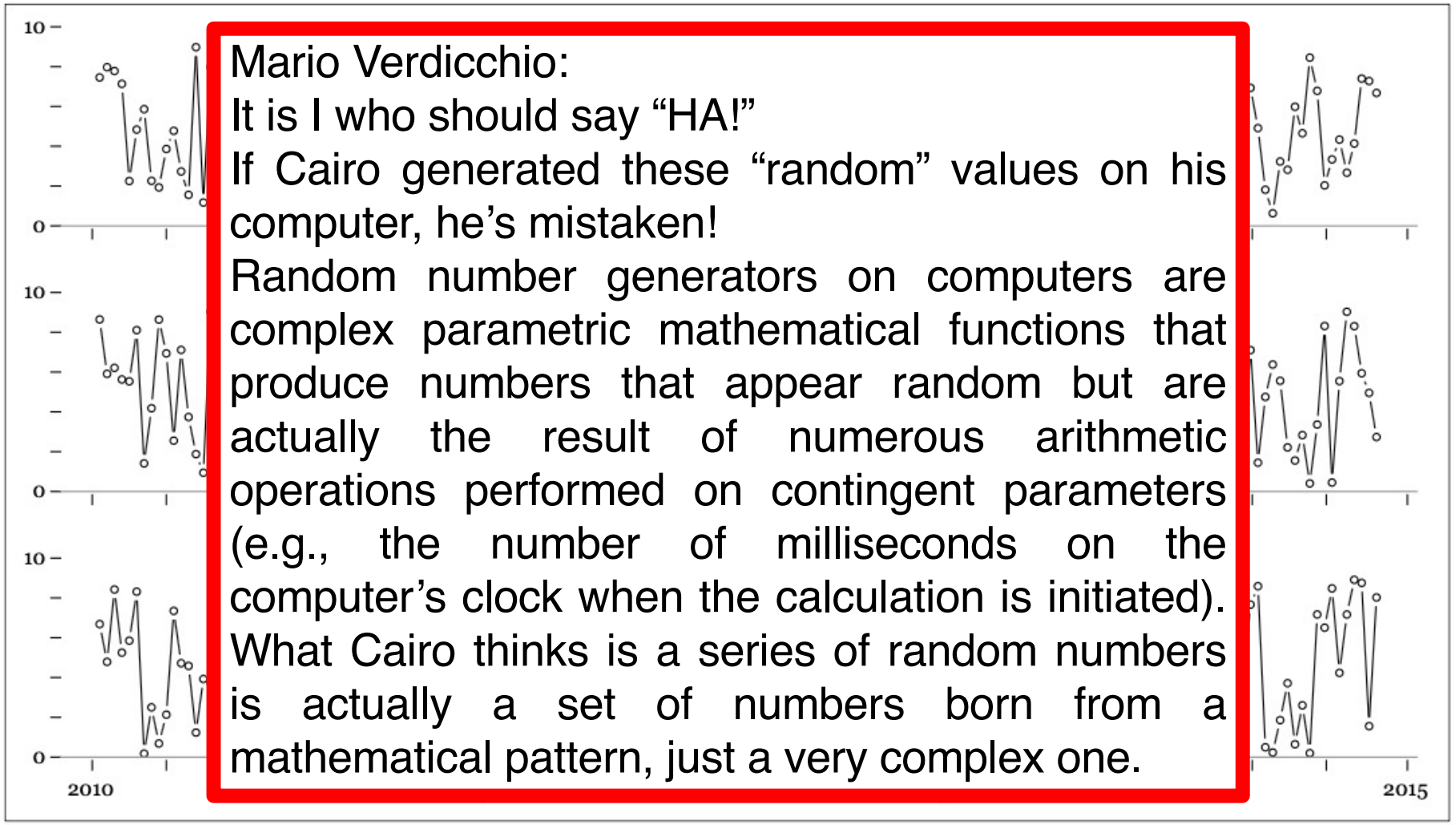
- We define noise as random fluctuations in a signal.
- More generally, we can identify noise with random values.
- Patterns, on the other hand, are determined by clear signals, by well-defined schemes.
- Apophenia is the tendency to mistake noise for patterns.

Back to the example



Alberto Cairo:
HA! If you saw patterns, you've fallen victim to your apophenia: I generated these values randomly. They are pure noise. There is no pattern, no signal, no structure here.
There's nothing to interpret.

Back to the example



Mario Verdicchio:
It is I who should say “HA!”
If Cairo generated these “random” values on his computer, he’s mistaken!
Random number generators on computers are complex parametric mathematical functions that produce numbers that appear random but are actually the result of numerous arithmetic operations performed on contingent parameters (e.g., the number of milliseconds on the computer’s clock when the calculation is initiated).
What Cairo thinks is a series of random numbers is actually a set of numbers born from a mathematical pattern, just a very complex one.

Plot twist!

From
“The sixth sense” by
M. Night Shyamalan

An observer of Cairo’s visualizations, due to their apophenia, sees patterns where there are none.
Cairo generates the visualization using a computer and does not see any patterns in the data because he generates them with a “random” function. In reality, since it’s a function running on a computer, its results can’t truly be random because the computer is a deterministic system that only works with precise mathematical rules.
So, Cairo, who accuses others of seeing patterns, could be accused of not seeing them himself!

Let’s clarify 1/3

- Cairo, in reality, is not entirely wrong, and he probably knows very well (since he is a data scientist) that the “randomly” generated data from a computer only appear to be random.
- They are more correctly called “pseudo-random” data.

Let’s clarify 2/3

- The “random” functions provided by programmers are specifically designed to make the mathematical patterns of the underlying functions completely invisible to the human eye: we would need to be able to calculate the exact time (to the millisecond) when the function is launched, perfectly know its code, and perform billions of calculations per second to recognize these patterns once the results are displayed on the screen.

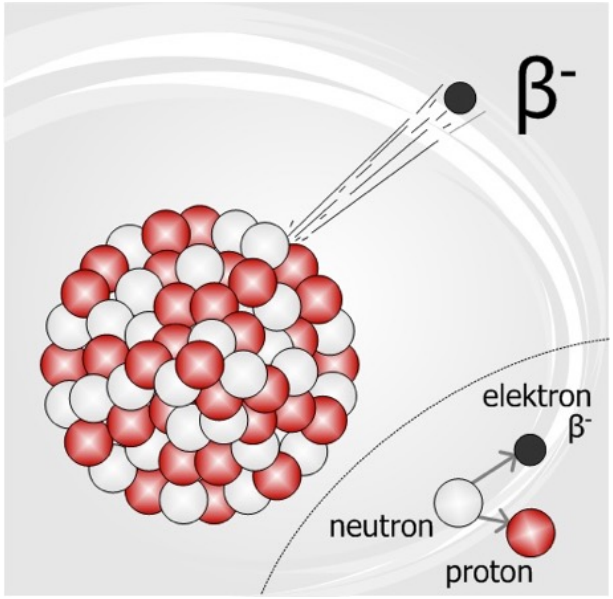
Let’s clarify 3/3

- With computer “random” functions, the patterns are indeed impossible to see, so we can easily accept the hypothesis that they do not exist.
- Moreover, even if we somehow managed to calculate them, they would only provide indications of processes happening within our computer and would not reflect any interesting aspect of the reality around us.

Patterns: to search or to ignore?

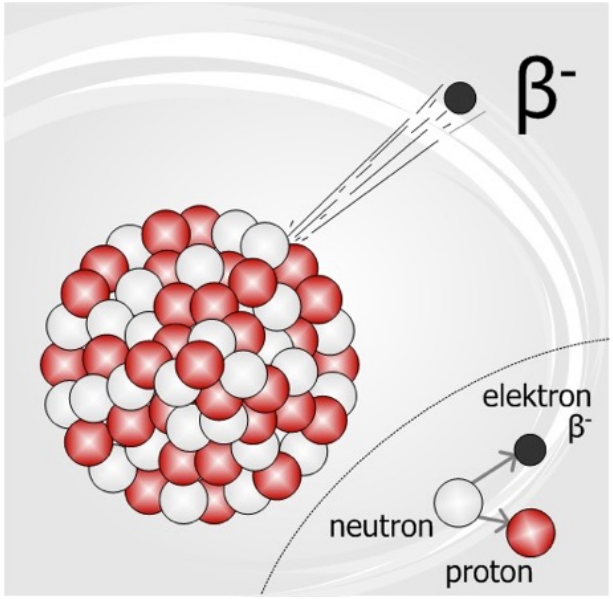
- Apophenia is problematic because it makes us see patterns where they don't exist.
- Visualizations must aim to be truthful, so apophenia should be avoided.
- It's fine to ignore patterns that don't interest us (like mathematical functions inside a computer).
- However, ignoring patterns that could be useful for our purposes can undermine the usefulness of a visualization.

True randomness?



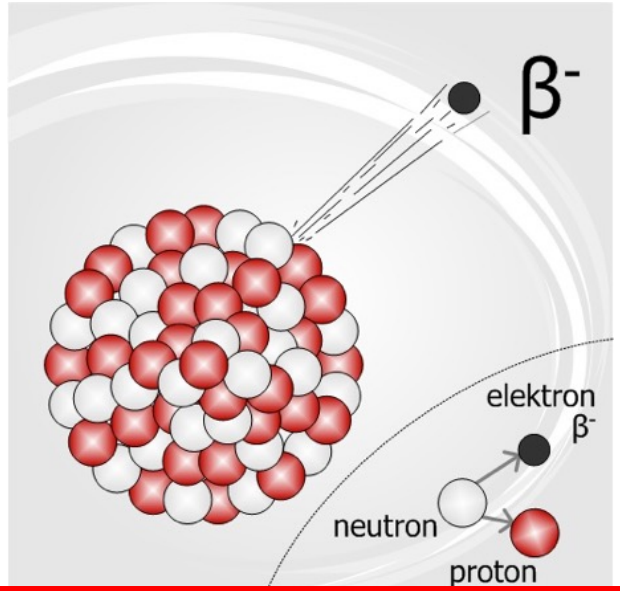
There are devices to generate true random numbers, and they are based on the decay of radioactive atoms.
In the nucleus of these atoms, a neutron occasionally decays into a proton and releases an electron and an electron antineutrino. The electron hits a sensor, and the decay is detected.
Consider 4 atoms (1, 2, 3, 4), let T_{12} be the time interval between the decay of 1 and the decay of 2; let T_{34} be the time interval between the decay of 3 and the decay of 4.
If $T_{12} < T_{34}$, we write a 0.
If $T_{12} \geq T_{34}$, we write a 1.
This digit is truly random because it is based on physical phenomena that have no pattern.

Who says that?



There are devices to generate true random numbers, and they are based on the decay of radioactive atoms.
In the nucleus of these atoms, a neutron occasionally decays into a proton and releases an electron and an electron antineutrino. The electron hits a sensor, and the decay is detected.
Consider 4 atoms (1, 2, 3, 4), let T_{12} be the time interval between the decay of 1 and the decay of 2; let T_{34} be the time interval between the decay of 3 and the decay of 4.
If $T_{12} < T_{34}$, we write a 0.
If $T_{12} \geq T_{34}$, we write a 1.
This digit is truly random because it is based on physical phenomena that have no pattern.

Who knows?



Isn't it possible that there is a pattern that exists but is invisible to us? Not a computational pattern inside the computer, but a pattern of a universal physical law that we haven't yet discovered or may never be able to discover?

There are devices to generate true random numbers, and they are based on the decay of radioactive atoms.

In the nucleus of these atoms, a neutron occasionally decays into a proton and releases an electron and an electron antineutrino. The electron hits a sensor, and the decay is detected.

Consider 4 atoms (1, 2, 3, 4), let T_{12} be the time interval between the decay of 1 and the decay of 2; let T_{34} be the time interval between the decay of 3 and the decay of 4.

If $T_{12} < T_{34}$, we write a 0.

If $T_{12} \geq T_{34}$, we write a 1.

This digit is truly random because it is based on physical phenomena that have no pattern.

Nature

The anomalous anomaly

Understanding the muon's magnetic moment holds the key for unlocking potential new physics, as Thomas Teubner shows.

1 Nov 2018

The Guardian

The big idea: are we about to discover a new force of nature?

Modern physics deals with some truly mind-boggling extremes of scale. Cosmology reveals the Earth as a tiny dot amid an observable universe...

16 Apr 2024

Physics World

The muon’s magnetic moment exposes a huge hole in the Standard Model – unless it doesn’t

The Muon g-2 Experiment at Fermilab found a significant discrepancy between the theoretical prediction and the experimental measurement of the anomalous...

18 Feb 2025

The creation of stories

- Here comes the role of narration, literally the creation of stories.
- The word “story” immediately suggests fiction, so it would be more appropriate to call them hypotheses.
- Returning to the example of radioactive atoms, we have several hypotheses:
  - Their decay is truly random.
  - Their decay is governed by laws not yet discovered.
  - Their decay is governed by laws that are impossible to discover.

We prefer stories

"*At its most basic level, a story is a description of something happening that contains some form of sensation, or drama. It is, in other words, an explanation of cause and effect that is soaked in emotion. Human thinking must take this form because we are biologically incapable of removing the feeling from it. That is how our thoughts are delivered. (…) We are natural-born storytellers who have a propensity to believe our own tales.*"

Will Storr
"The Unpersuadables: Adventures with the Enemies of Science"
(2014)

“Don’t catch a cold!”

- This is one of the hardest stories to die, because many people still believe it.
- Science has repeatedly confirmed through experiments that cold weather does not cause a “cold” (whose name is itself a consequence of the story):
  - Some viruses stay active longer when temperatures are low.
  - When it’s cold, people tend to stay inside longer, where they are more likely to come into contact with other people who have a virus.
- It’s the viruses that cause a cold (and worse), not cold weather.

The power of narration

- Narration can be a powerful tool for communicating effectively.
- However, it is also dangerous because it can blind us to potential evidence that should prompt us to reconsider or discard the models we have adopted so far.
- Radical activism, the uncritical defense of causes, and “opinion” journalism are all activities based on narratives.
- Confirmation bias strengthens the power of a narrative.

Confirmation bias

Confirmation bias

- Warning: the comfort of lies in the cartoon does not come from the fact that those lies say pleasant things (e.g., “I’m still in great shape,” “gray hair makes me more attractive,” “these slides are really interesting”).
- That comfort comes from the alignment between the lies and the model in the mind of the listener: they seem to confirm what we already believe, so we tend to want them to be true.
- Similarly, uncomfortable truths are not unpleasant because they are negative in themselves, but because they cause cognitive dissonance.

Cognitive dissonance

In *A Theory of Cognitive Dissonance* (1957), Leon Festinger was the first to propose the hypothesis that human beings have a desire to maintain psychological consistency in order to live well. A person who perceives an inconsistency between their beliefs, or between their beliefs and their behavior, experiences psychological discomfort and is driven to reduce this dissonance.


Doctors exaggerate the harms of smoking.
A statistical increase in the probability of cancer doesn't mean I'll definitely get it.
If I quit, I'd gain weight, and that's just as dangerous.
Smoking is truly enjoyable and worth the risk.

Each of these thoughts is formulated to reduce dissonance—that is, to soften the model that says smoking is harmful (the only scientifically supported one) with alternative models (based on emotions and rhetoric).
The desire to reduce dissonance makes a person more receptive to stories that confirm the alternative models. This is cognitive bias.

Back to climate change

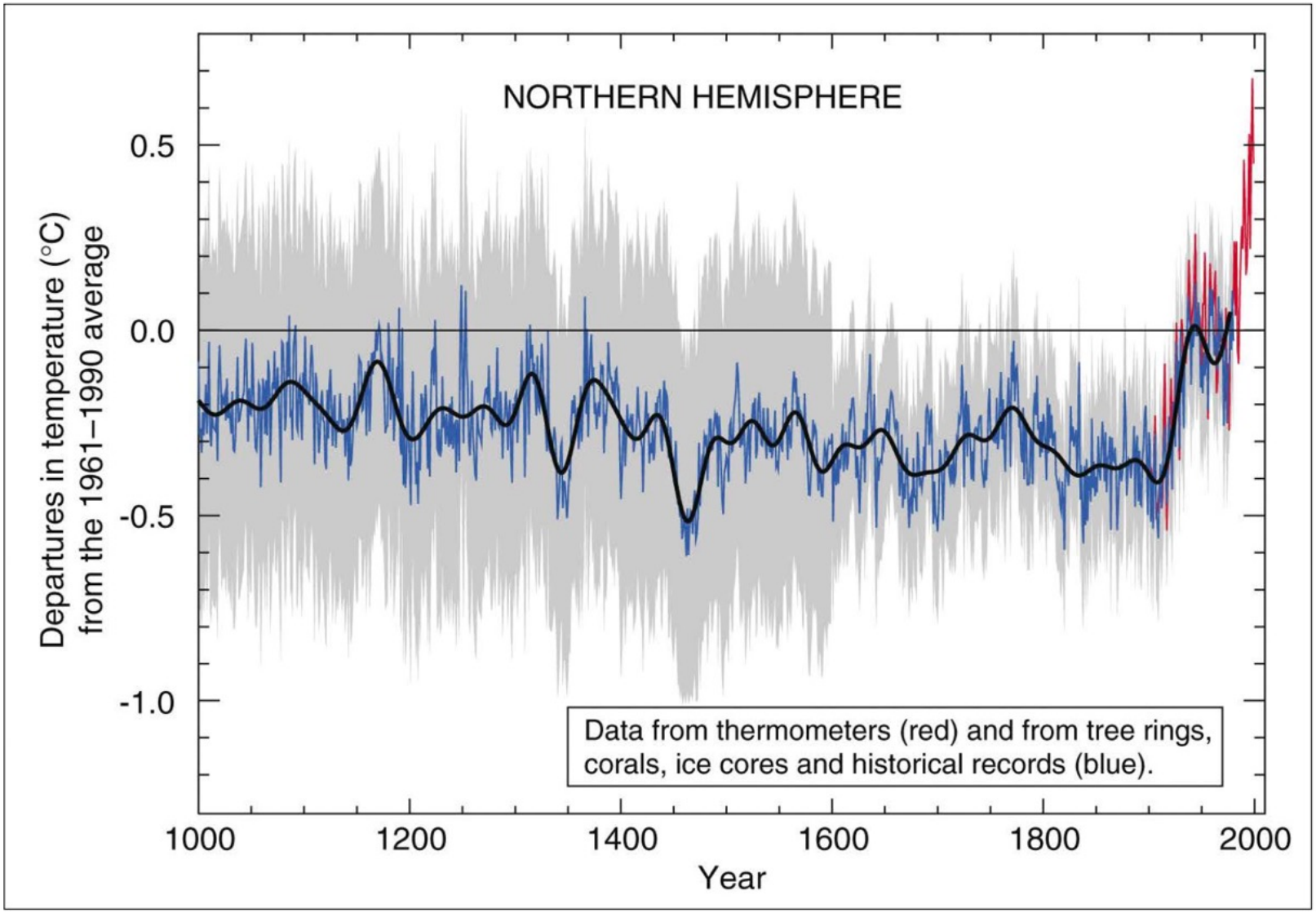


Chart by Mann, Bradley e Hughes. From “The Truthful Art” by A. Cairo

“The principal reason people disagree about climate change science is not that it has been communicated to them in forms they cannot understand.
Rather, it is that positions on climate change convey values—communal concern versus individual self-reliance; prudent self-abnegation versus the heroic pursuit of reward; humility versus ingenuity; harmony with nature versus mastery over it—that divide them along cultural lines.

“Merely amplifying or improving the clarity of information on climate change science won’t generate public consensus if risk communicators fail to [pay attention to] the cues that determine what climate-change risk perceptions express about the cultural commitments of those who form them.
In fact, such inattention can deepen polarization.

“Citizens who hold hierarchical and individualistic values discount scientific information about climate change in part because they associate the issue with antagonism to commerce and industry.

“ Individuals are prone to interpret challenges to beliefs that predominate with their cultural community as assaults on the competence of those whom they trust and look to for guidance.

“That implication—which naturally provokes resistance—is likely to be strengthened when communicators with a recognizable cultural identity stridently accuse those who disagree with them of lacking intelligence or integrity.”

From:
Dan Kahan et al. “The tragedy of the risk-perception commons: culture conflict, rationality conflict, and climate change.”
Temple University legal studies research paper (2011).

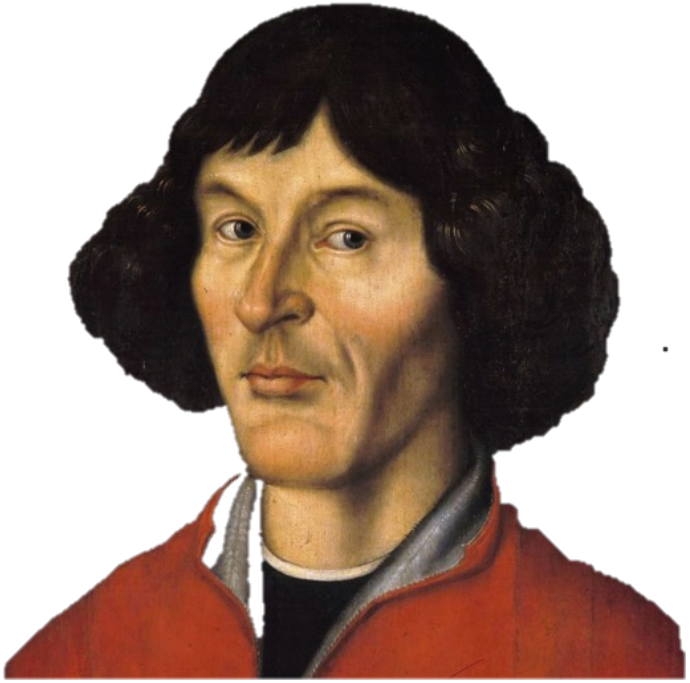
- The way we present information is just as important as the reliability of the information itself.

Models

- Returning to the concept of models, we've seen that it can refer to visualizations, but also to personal beliefs or scientific theories.
- The goal of honest designers, researchers, and scientists is to find models that are as close to reality as possible.
- Given the similarity to science, one idea could be to borrow characteristics of the scientific method.

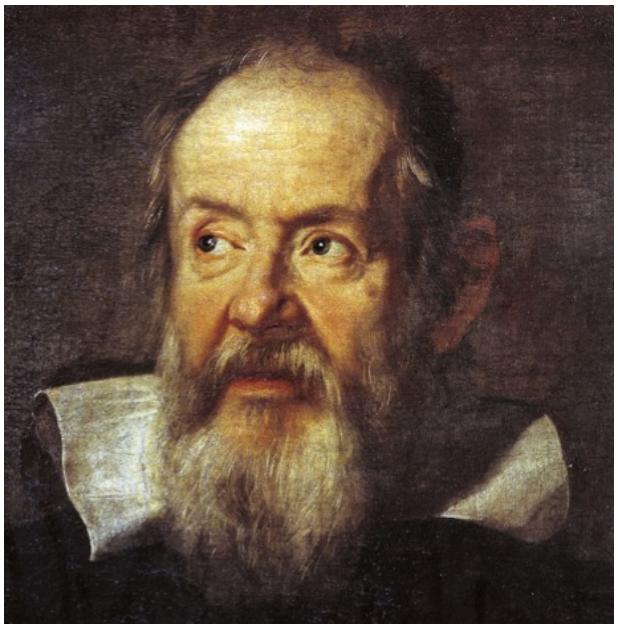
The scientific method

- Francis Bacon (1561–1626) is generally credited with being the first to formalize the scientific method.
- A successful lawyer and prominent London philosopher, he proposed a new approach to scientific inquiry, which he published in 1621 as the *Novum Organum Scientiarum*.
- He advocated for inductive reasoning (from the particular to the general) as the foundation of scientific thinking.
- He believed that only a clear and precise system of scientific inquiry would ensure humankind's mastery over the world.



Nicolaus Copernicus
(1473–1543)

The works of Copernicus and Galileo had a significant influence on Bacon.



Galileo Galilei
(1564 – 1642)

From the scientific method to visualizations

- To make visualizations as truthful as possible, we apply rigorous thinking tools:
  - logic
  - statistics
  - experimentation
- More abundant and better information leads to better models.
- 400 years of empirical and experimental science have shown that these tools work.
- Even if perfect understanding of reality is unattainable (e.g., atomic decay), just as better scientific theories replace previous ones and provide a better description of reality, we can strive to have visualizations that increasingly approximate reality.

Example: journalists in Georgia

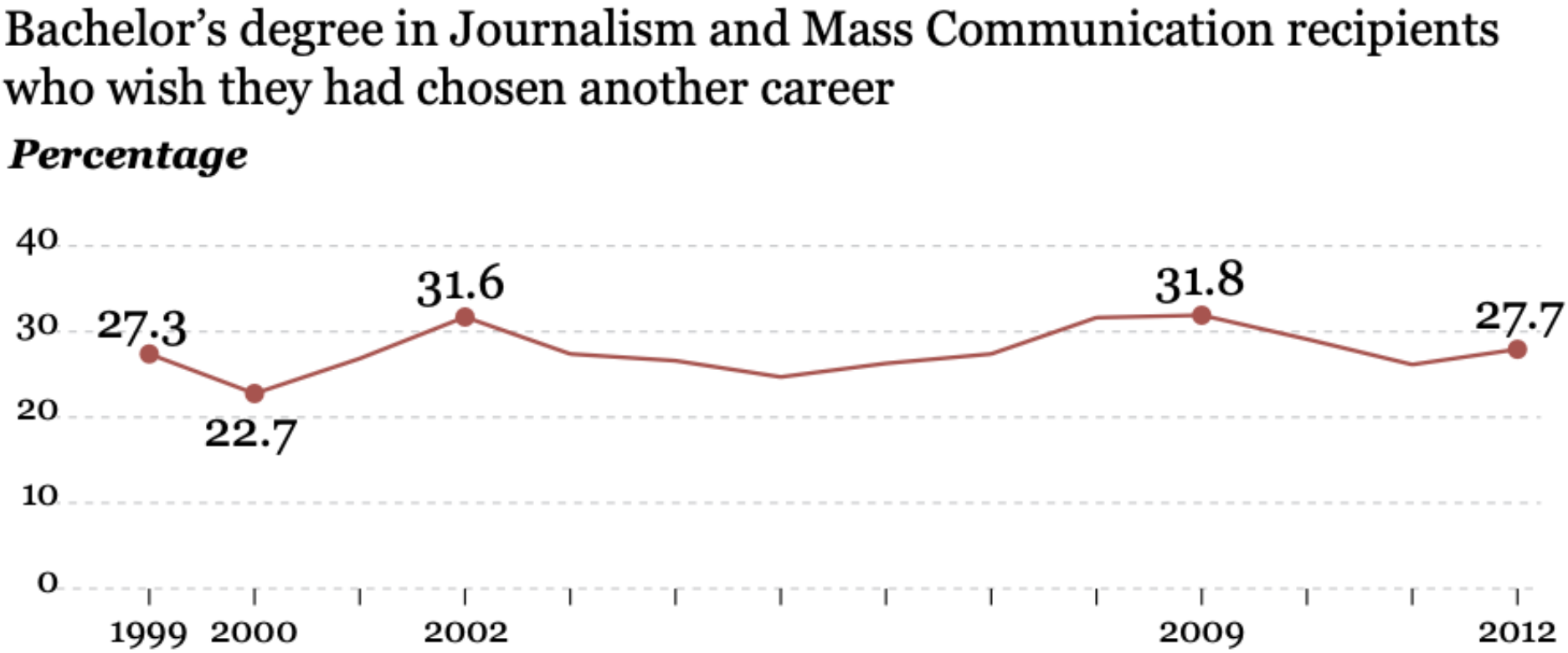
Figure by TUBS, from Wikipedia.org

Example: journalists in Georgia

- Headline from a newspaper on 15/08/2013: "*A study reveals that more than a quarter of journalism graduates wish they had chosen a different career.*"
- The article begins with a reference to the data source:
  "About 28% of journalism graduates wish they had chosen a different field, according to the annual graduate survey from Grady College at the University of Georgia."
- Is it true? Is it false? Or rather, are the data from this survey a good approximation of reality?

Possible improvement 1

- The data cited by the article referred to the recently concluded academic year.
- A potential improvement comes from looking at similar data from previous years.
- Cairo did this and discovered what is shown in this graph.

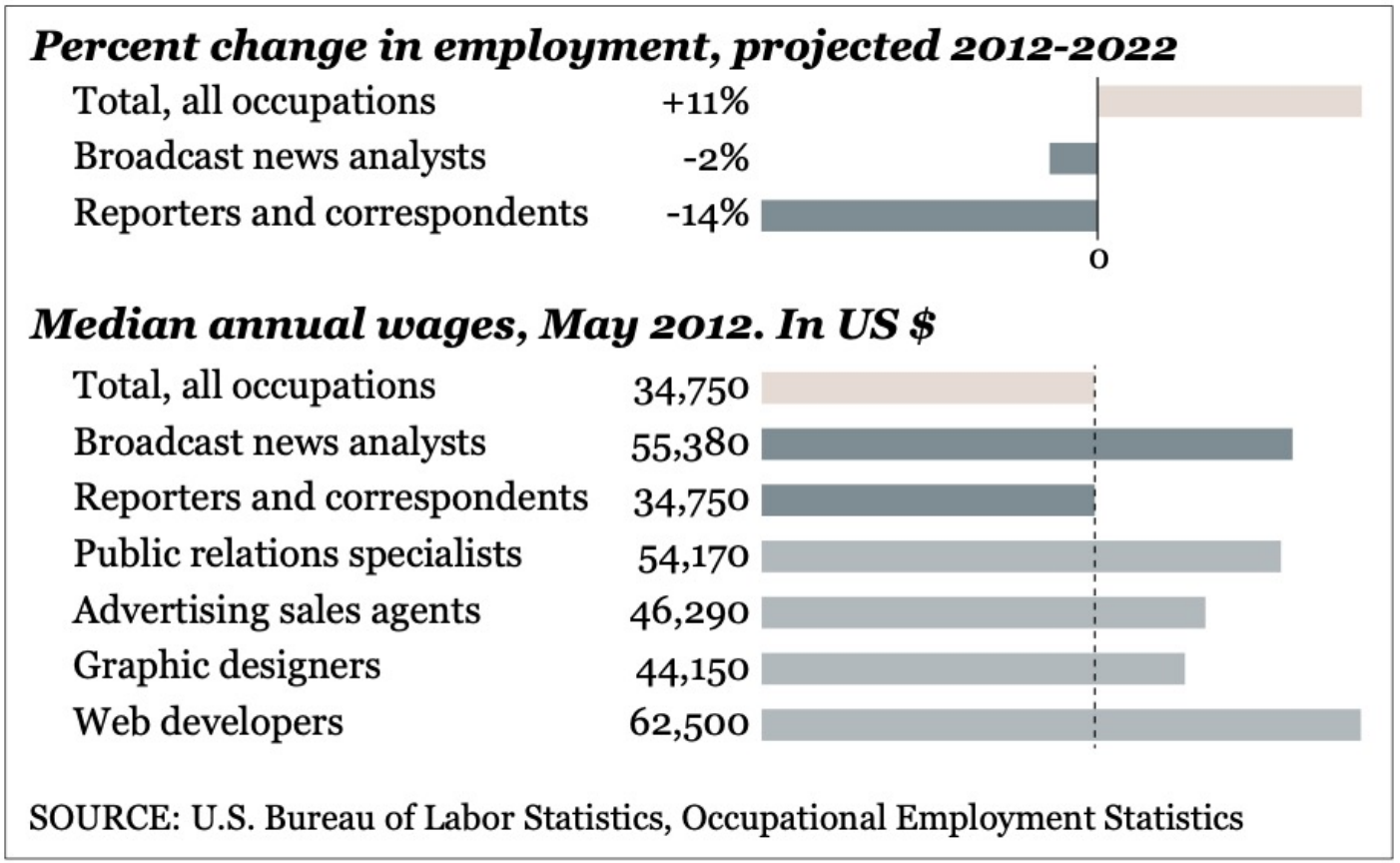


From "The Truthful Art" by A. Cairo

- We find that this percentage has remained almost unchanged (except for a dip in 2000 that warrants further investigation) since 1999.
- So, even a headline that seems almost sensationalist, suggesting a new phenomenon or a collapse, turns out inadequate and unrepresentative of the reality.

Possible improvemente 2

- The percentage has not changed significantly since 1999, but many things have changed since 1999:
  - The Internet boom
  - 9/11
  - The great crisis in publishing
  - The 2008 economic crisis
- A possible improvement to the model is to expand the context by considering other data that are connected to the initial ones.
- Example: How much do journalists earn?



From "The Truthful Art" by A. Cairo

- With additional data from the Bureau of Labor Statistics (the one behind the "hockey stick" graph of COVID-19-related unemployment claims), we confirm that the job market for journalists is in crisis. Not only that…
- We find that the most common profession for journalism graduates, i.e., reporter, is not very well paid: exactly the average salary in the USA, with the average calculated across all jobs, including the lowest-paid ones.

Possible improvement 3

- There is no data on this, but how do other faculties fare?
- Who tells us that the 28% of dissatisfied journalism graduates isn't a much smaller percentage compared to the 40% of law graduates or the 50% of humanities graduates? (All hypotheses to be verified)
- Expanding the context with additional data certainly helps to better interpret the data we have and create more truthful visualizations.

A better title

- With all the additional information we've uncovered from other data connected to the initial dataset, the one mentioned in the article, we are able to propose a better headline.
- No longer:
  "*A study reveals that over a quarter of journalism graduates wish they had chosen a different career.*"
- But rather:
  "*A study reveals that the shocks to the job market over the past ten years have not shaken the resolve of journalism graduates.*"
- A whole different story.

Lesson learned

- Don’t rush to write a headline or a story, nor to design a visualization immediately after finding a pattern, a data point, or an interesting fact.
- Pause and think.
- Look for other sources and people who can help you remove blinders and eliminate confirmation bias, seeking external factors that might help you explain what you’ve discovered.
- Only then can you decide:
  - what to say
  - how to say it
  - how many details to provide