Becoming a Data Detective

How to Recognize Forms of Bias in Data Storytelling

David Cuervo, John Montgomery, Zach Gonzales

TEXAS
Department of Family and Protective Services
Analytics for Change & Evaluation

Objectives

- Introduction of our work
- How design impacts our perception
- Cognitive Bias in data
- Transparency in data
- Data Detective tips

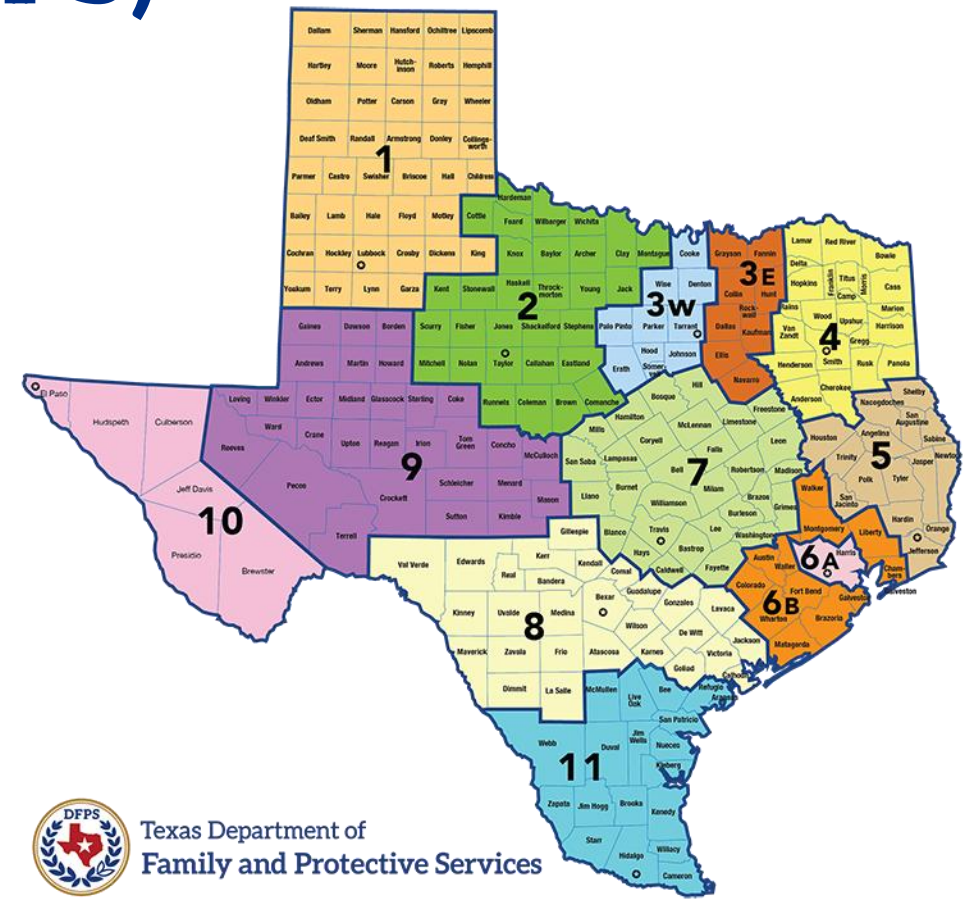
Who We Are

TEXAS
Department of Family
and Protective Services

Texas Department of Family and Protective Services (DFPS)

- Mission: We promote safe and healthy families and protect children and vulnerable adults from abuse, neglect, and exploitation.

- ~28,000 children in DFPS custody during FY2024
- ~23,000 families received Family Preservation Services in FY2024



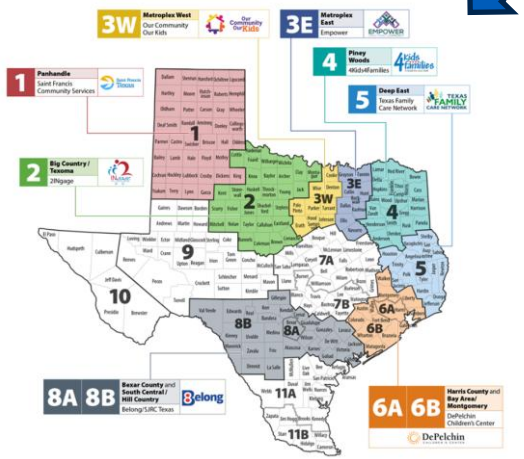
DFPS Programs

Program	Description
Adult Protective Services	Protects the elderly and people with disabilities from abuse, neglect, and exploitation through investigations and services.
Child Protective Services	Protects children from abuse and neglect through services, foster care, and adoption.
Child Protective Investigations	Investigates reports of child abuse or neglect to determine if any child in the family has been abused or neglected. If the children aren't safe, the investigator starts protective services.
Statewide Intake	Takes reports of abuse, neglect, and exploitation from across the state through its Texas Abuse Hotline (1-800-252-5400) and through the website TxAbuseHotline.org 24-hours a day, every day of the year. Professional, passionate, and innovative.

Our Work

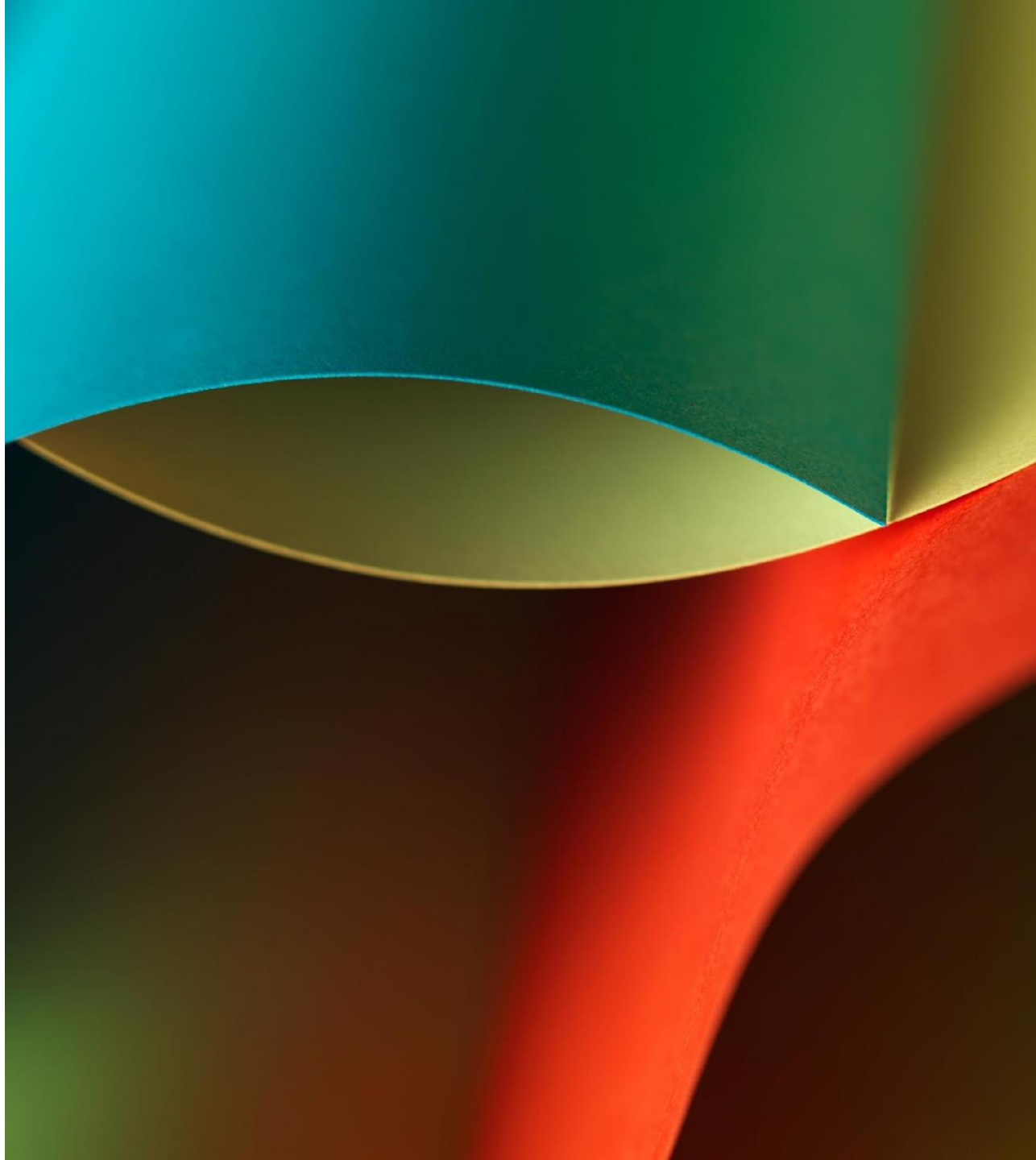
Visualization and Interpretation for Change and Enhancement

TEXAS Department of Family and Protective Services



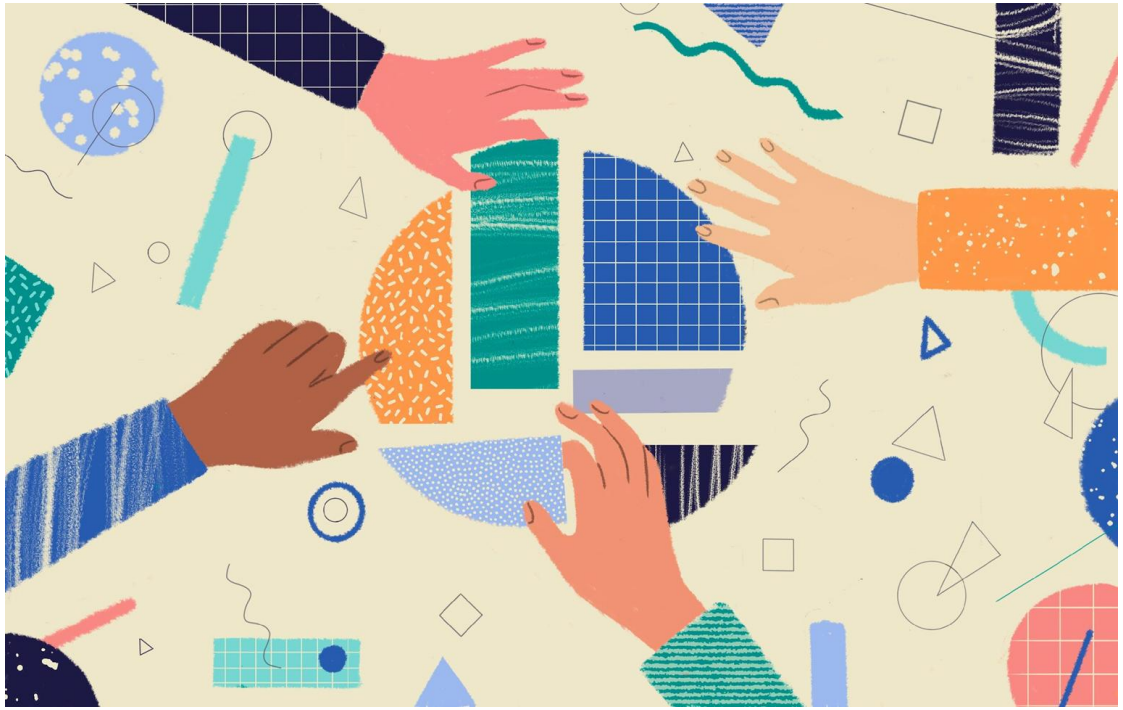
ADMINISTRATION FOR CHILDREN & FAMILIES

Data Visualization Design



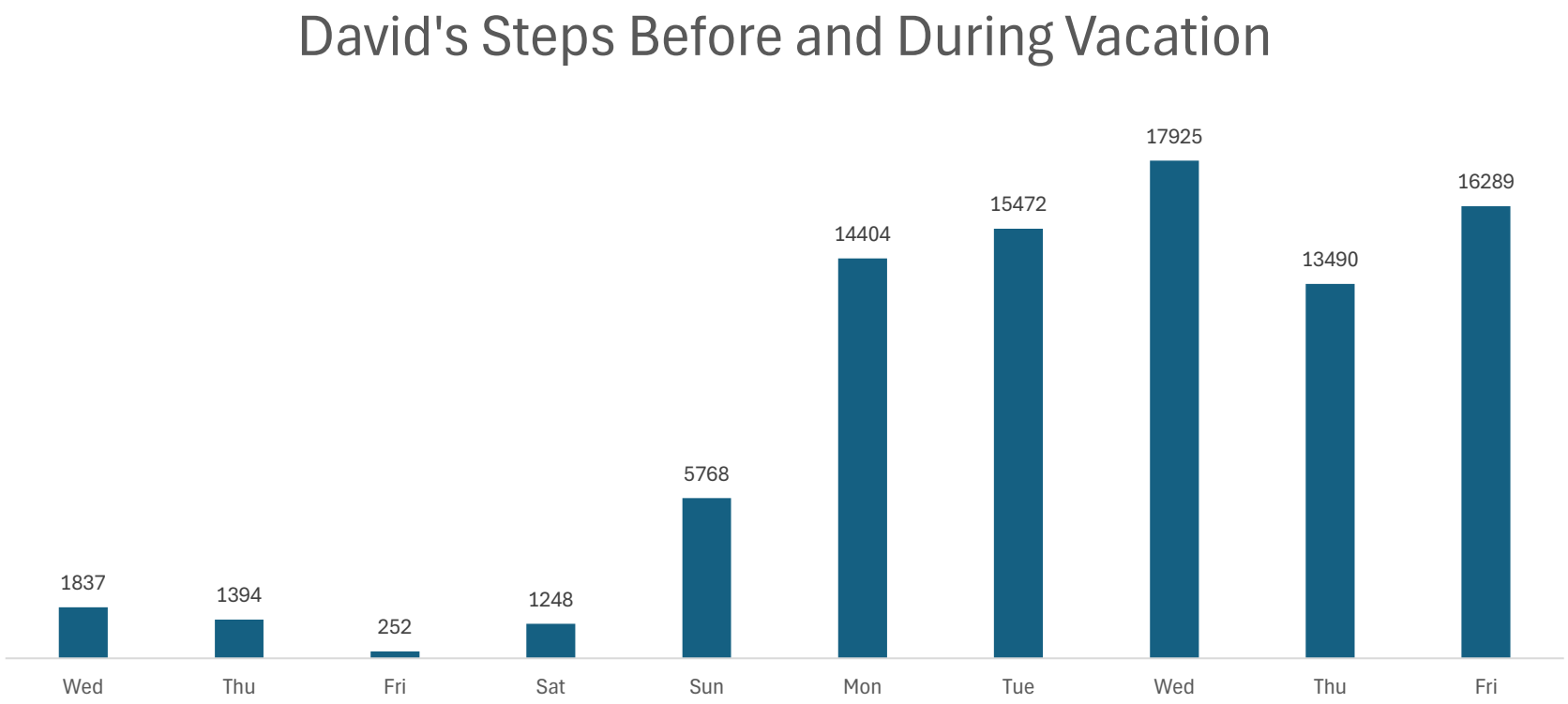
Storytelling through Data

- Data is always being spoken for
- Many decisions shape the creation of a chart, graph, infographic, etc.
    - What is included, what is left out, scale, colors
- Many opportunities for bias to creep in



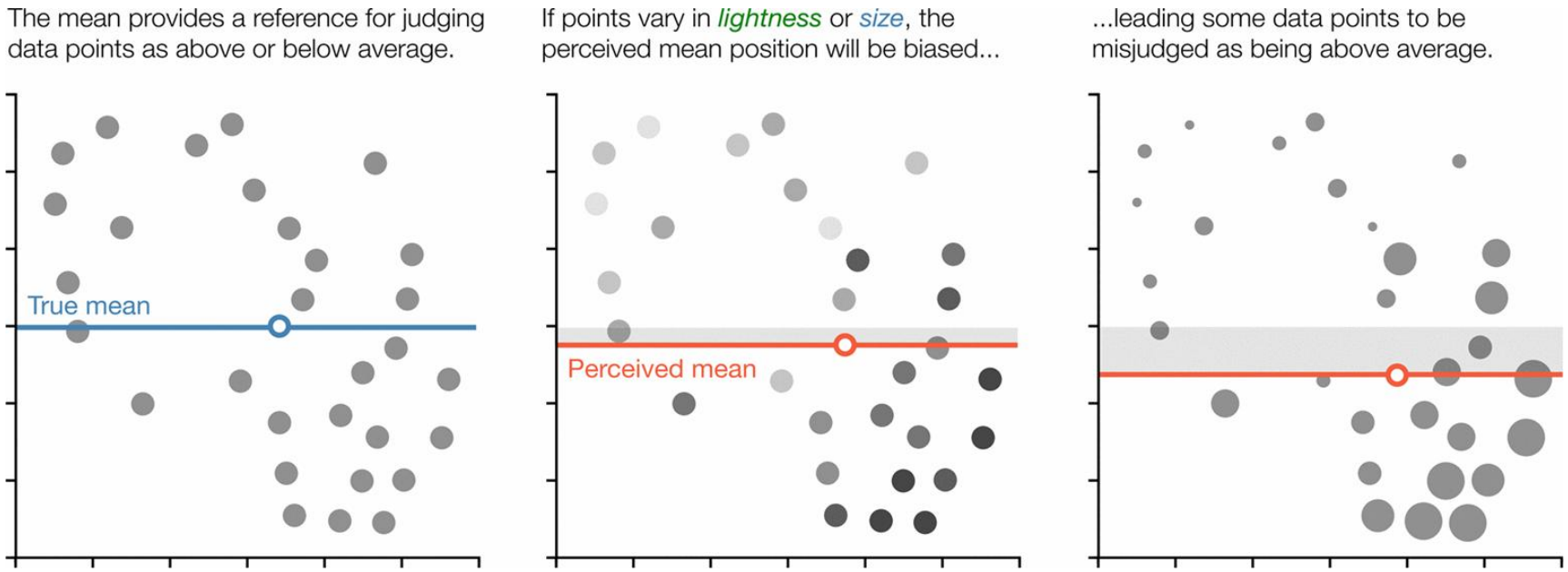
What Bias Can be Introduced?

Type of Bias	Description
Sampling bias	What gets included/excluded
Selection bias	Choosing data that support a narrative
Axis manipulation	Truncating axes or using non-zero baselines to exaggerate differences
Visual encoding bias	Making use of color, shape, or size to draw attention



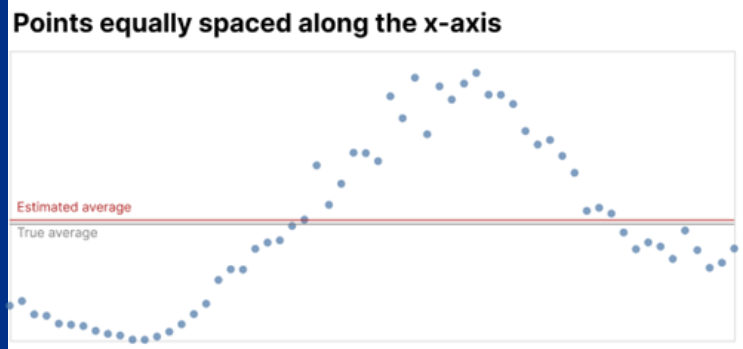
Psychological Impacts of Visualization Techniques

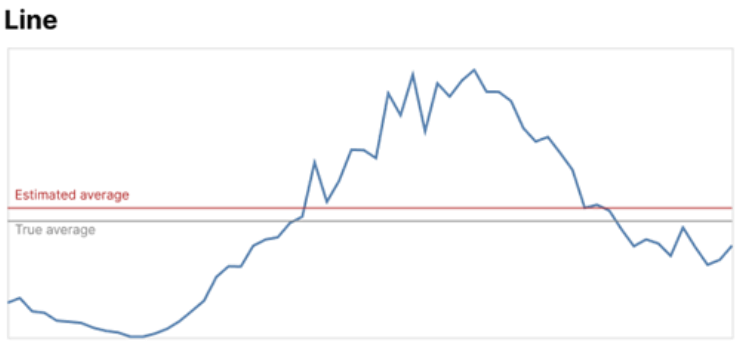
The Weighted Average Illusion: Biases in Perceived Mean Position in Scatterplots (Hong et al, 2022)

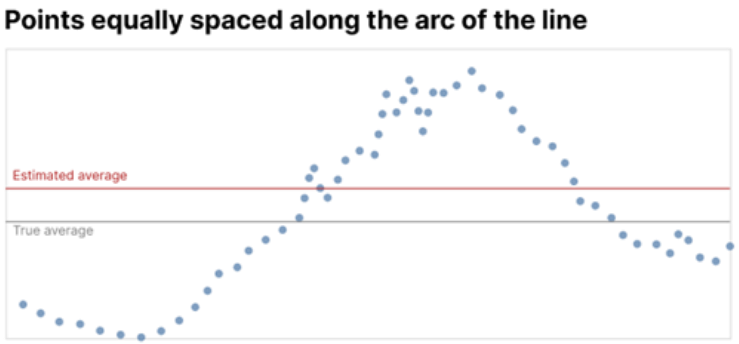


M. Hong, J. K. Witt and D. A. Szafir, "The Weighted Average Illusion: Biases in Perceived Mean Position in Scatterplots" in IEEE Transactions on Visualization & Computer Graphics, vol. 28, no. 01, pp. 987-997, Jan. 2022, doi: 10.1109/TVCG.2021.3114783.

Change in Perception



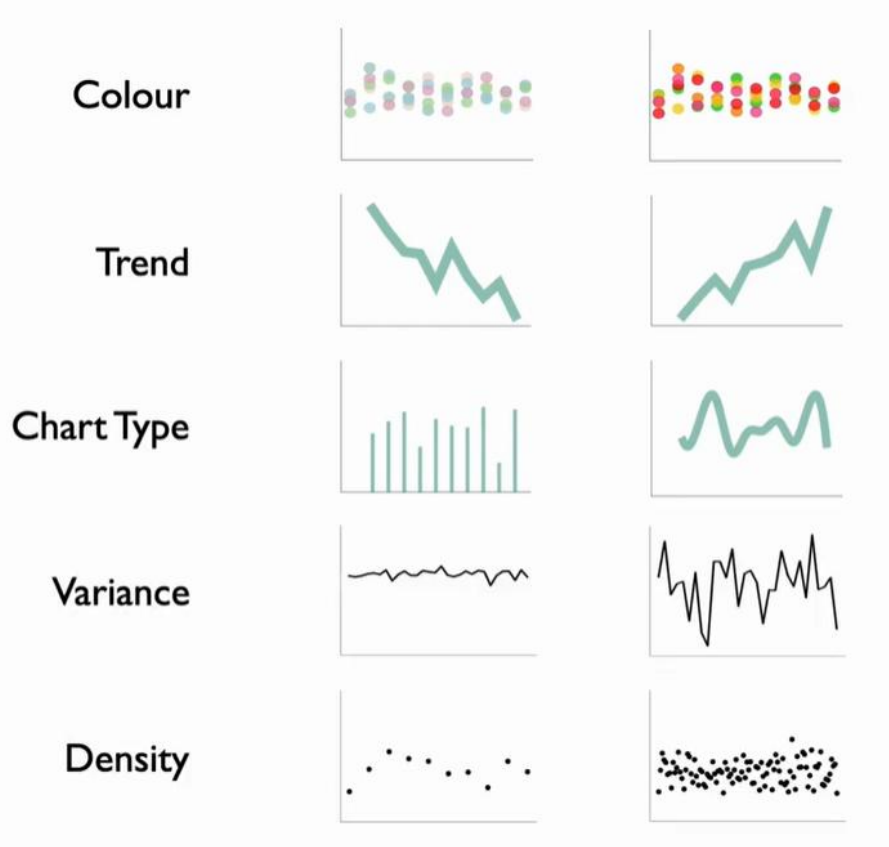


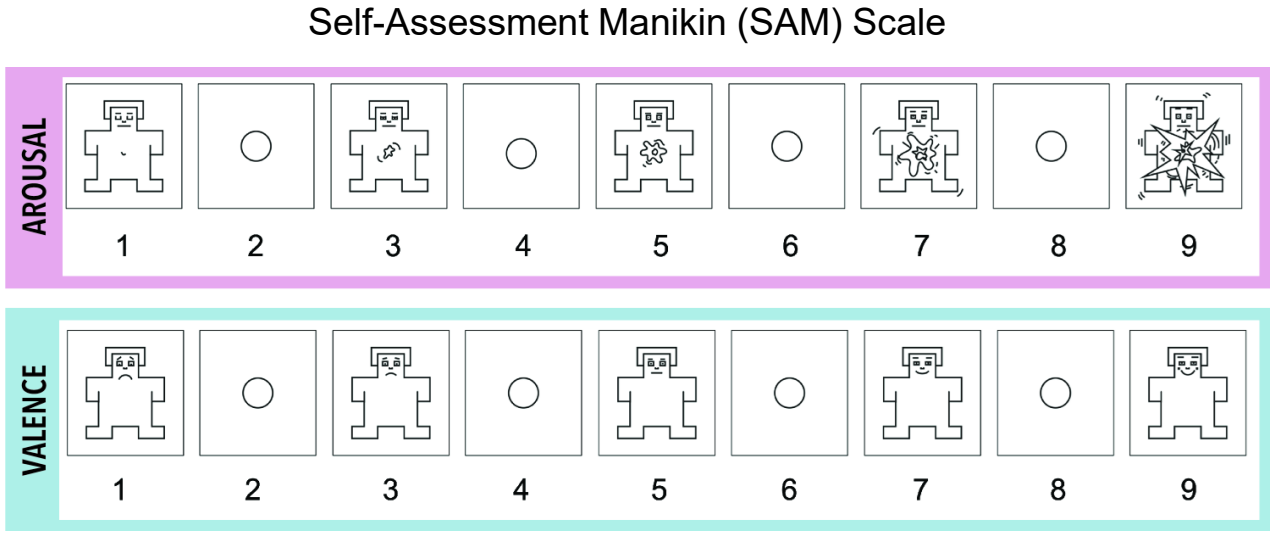


- People's estimates of the average value tends to be pulled toward sections of the graph where the line is more "wiggly"
- More visually salient

D. Moritz, L. M. Padilla, F. Nguyen and S. L. Franconeri, "Average Estimates in Line Graphs Are Biased Toward Areas of Higher Variability," in IEEE Transactions on Visualization and Computer Graphics, vol. 30, no. 1, pp. 306-315, Jan. 2024, doi: 10.1109/TVCG.2023.3326589.

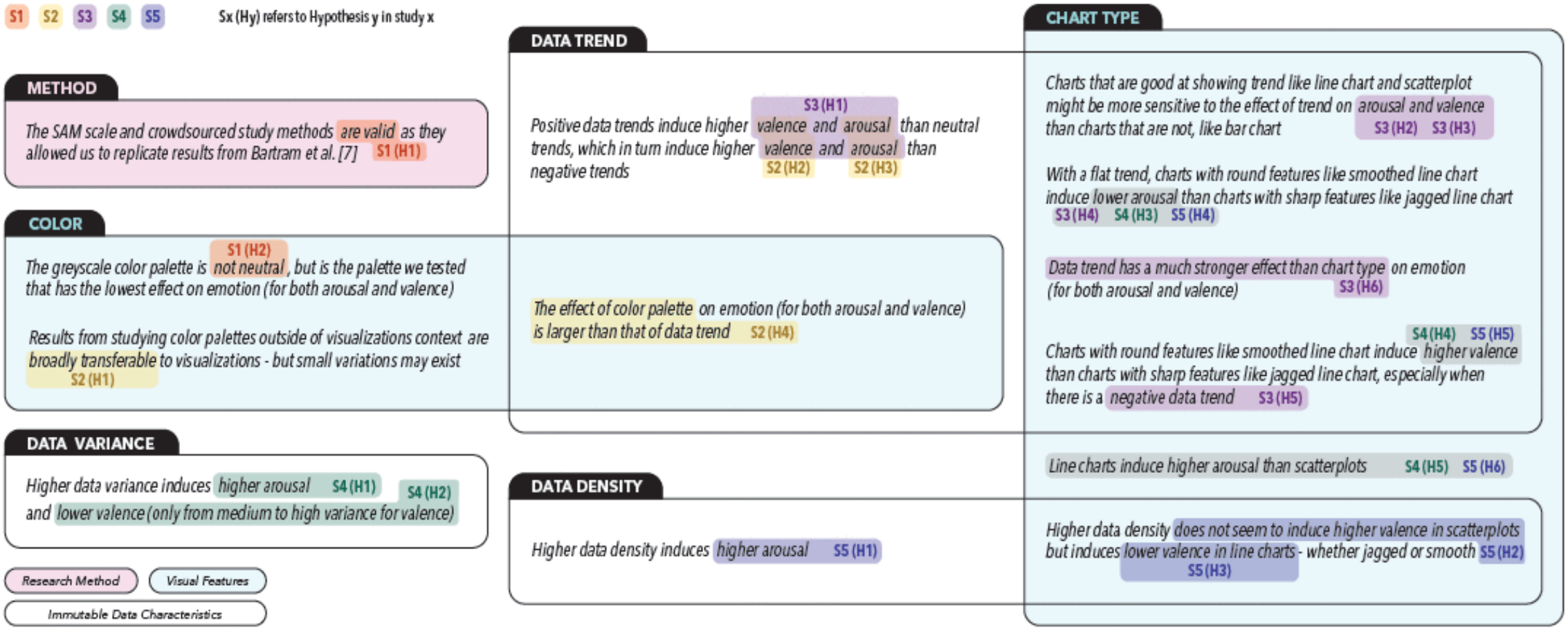
Emotional Responses to Data Characteristics





Blair C, Wang X, Perin C. Quantifying Emotional Responses to Immutable Data Characteristics and Designer Choices in Data Visualizations. IEEE Trans Vis Comput Graph. 2025 Jan;31(1):1006-1016. doi: 10.1109/TVCG.2024.3456361. Epub 2024 Nov 25. PMID: 39255151.

All 5 Factors Have an Impact on Our Emotions



Blair C, Wang X, Perin C. Quantifying Emotional Responses to Immutable Data Characteristics and Designer Choices in Data Visualizations. IEEE Trans Vis Comput Graph. 2025 Jan;31(1):1006-1016. doi: 10.1109/TVCG.2024.3456361. Epub 2024 Nov 25. PMID: 39255151.

What is Cognitive Bias?

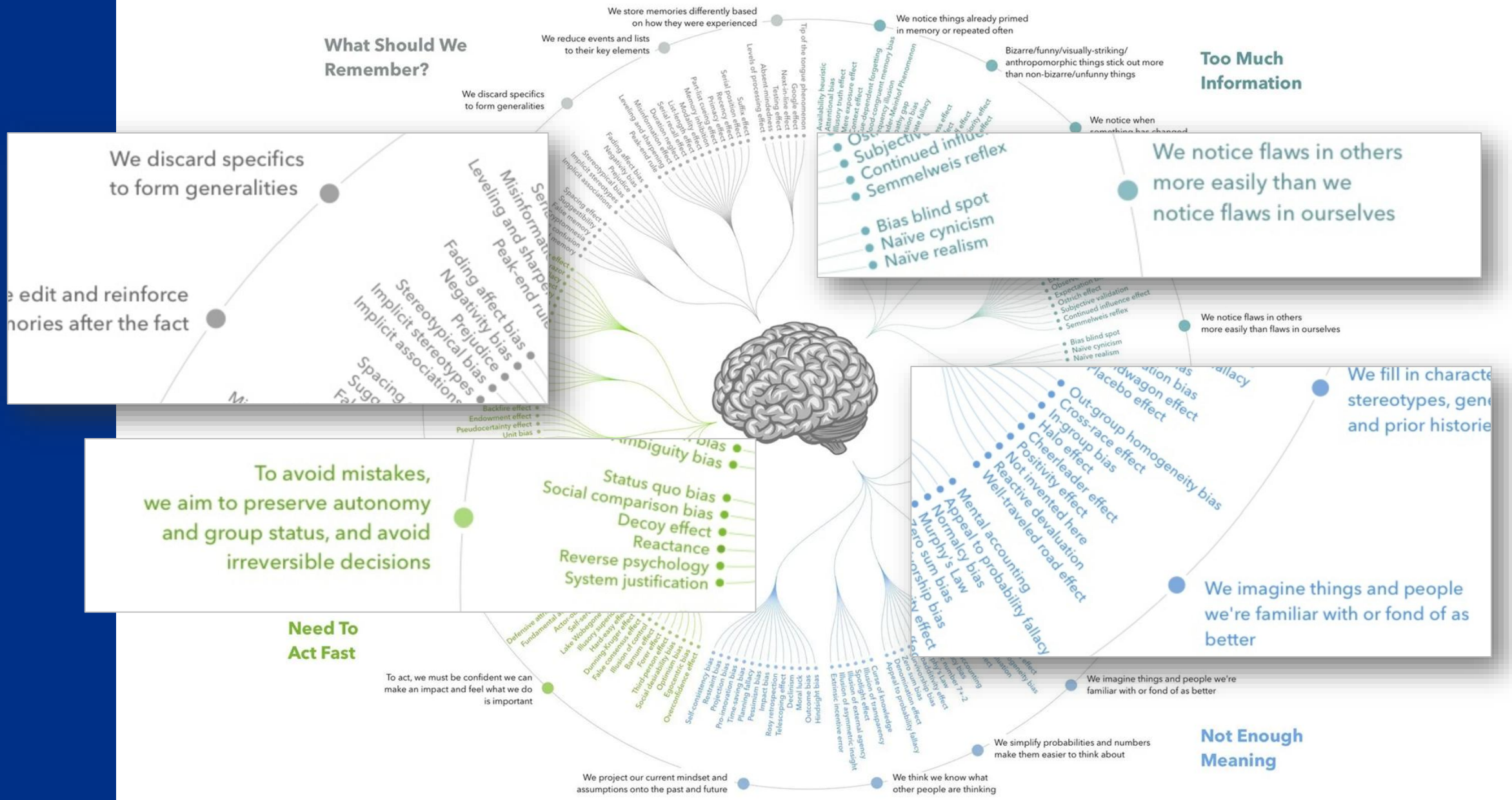
Cognitive bias is a mental shortcut that influences our thinking and decision-making, leading us to process information in a selective and subjective manner, often resulting in inaccurate or irrational judgements.

Hello, I’m Mango!

Think about it…

What formative events in your life are you aware of that influenced building your character and values?

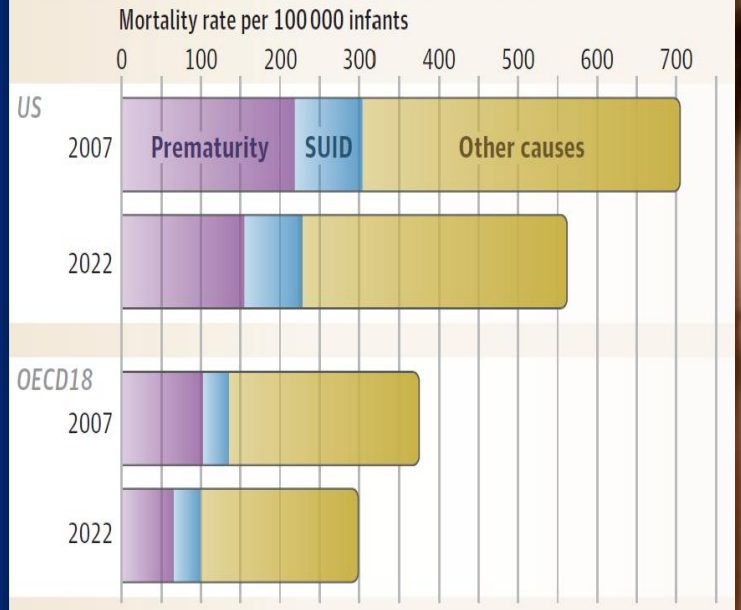
COGNITIVE BIAS CODEX, 2016



ALGORITHMIC LAYOUT + DESIGN BY JM3 - JOHN MANOOGIAN III // CONCEPT + METICULOUS CATEGORIZATION BY BUSTER BENSON // DEEP RESEARCH BY WIKIPEDIANS FAR + WIDE

Mortality rate of infants in the US vs the OECD18[a]

The all-cause mortality rate decreased for all infants <1 y of age between the years 2007 and 2022 but remains higher in the US.

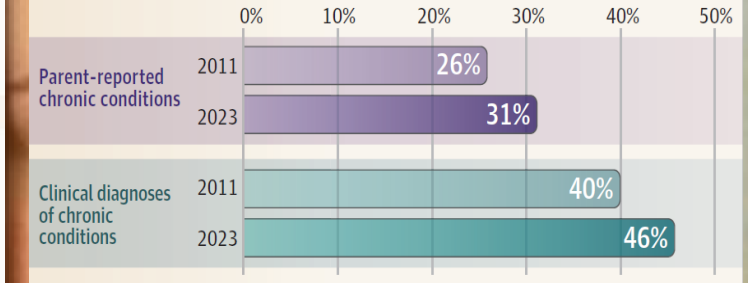


[a]The Organisation for Economic Co-operation and Development (OECD18) comprises 18 comparator nations including Australia, Austria, Belgium, Canada, Demark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, and the United Kingdom.

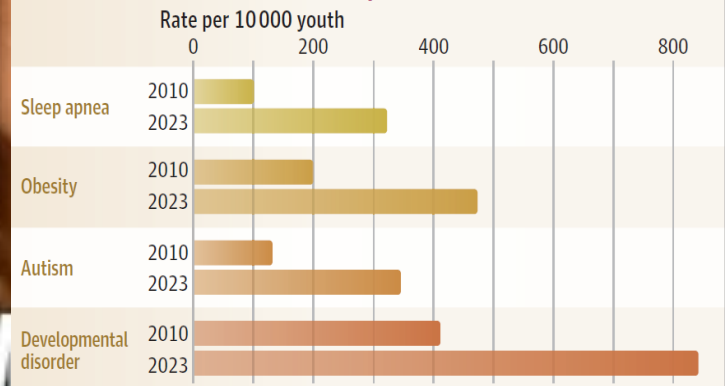
SUID indicates sudden unexpected infant death.

Prevalence of chronic conditions in US youth

Percentage with chronic conditions increased in those aged 3 to 17 y between 2011 and 2023, according to both parents and clinicians.

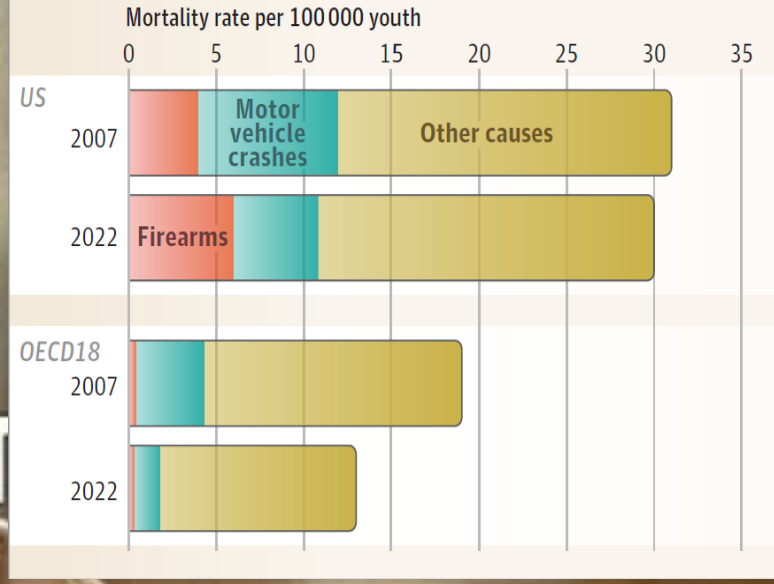


Some chronic conditions increased in prevalence between 2010 and 2023.



Mortality rate of youth in the US vs the OECD18

Individuals aged 1 to 19 y were more than twice as likely to die in the US as those in the OECD18 in 2022.



gest discrepancy was for firearm deaths, which were more than 15 times higher in the US. Motor vehicle crashes, the second most common cause of death in the US, were 2.5 times higher than in the OECD18.

What Christopher B. Forrest, MD, PhD, lead author of the study and a professor of pediatrics at Children's Hospital of Philadelphia, found particularly "disturbing" was the growing mortality gap be-

Cognitive Biases that Enable Storytelling Bias

Despite ongoing medical advances, the US may be failing to adequately support children's health. A new study in *JAMA* examined pediatric mortality and chronic disease to assess the state of children's health in the US over the past 2 decades compared with other high-income countries.

JAMA Data Brief

More US Children Die Than in Other High-Income Nations

Samantha Anderer; Karen Bucher, MA

surveys, a
atric healt
Altho
1 year in b
mained comparatively high
infants in the US were 2.2 tin
2.4 times more likely to di
(SUID), which includes sudd

sess the
compare

ronic disease to as-

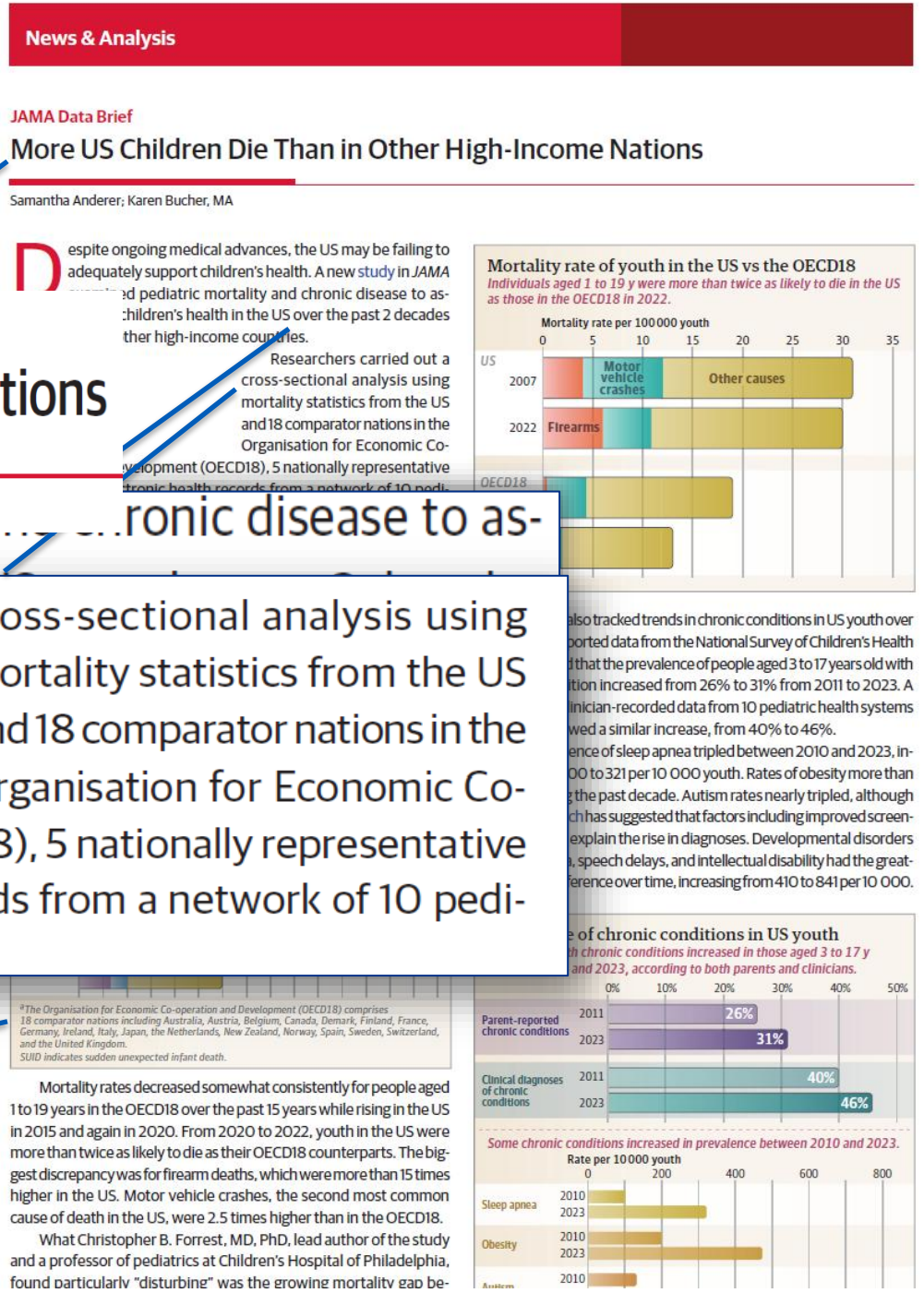
Editorial

Related article

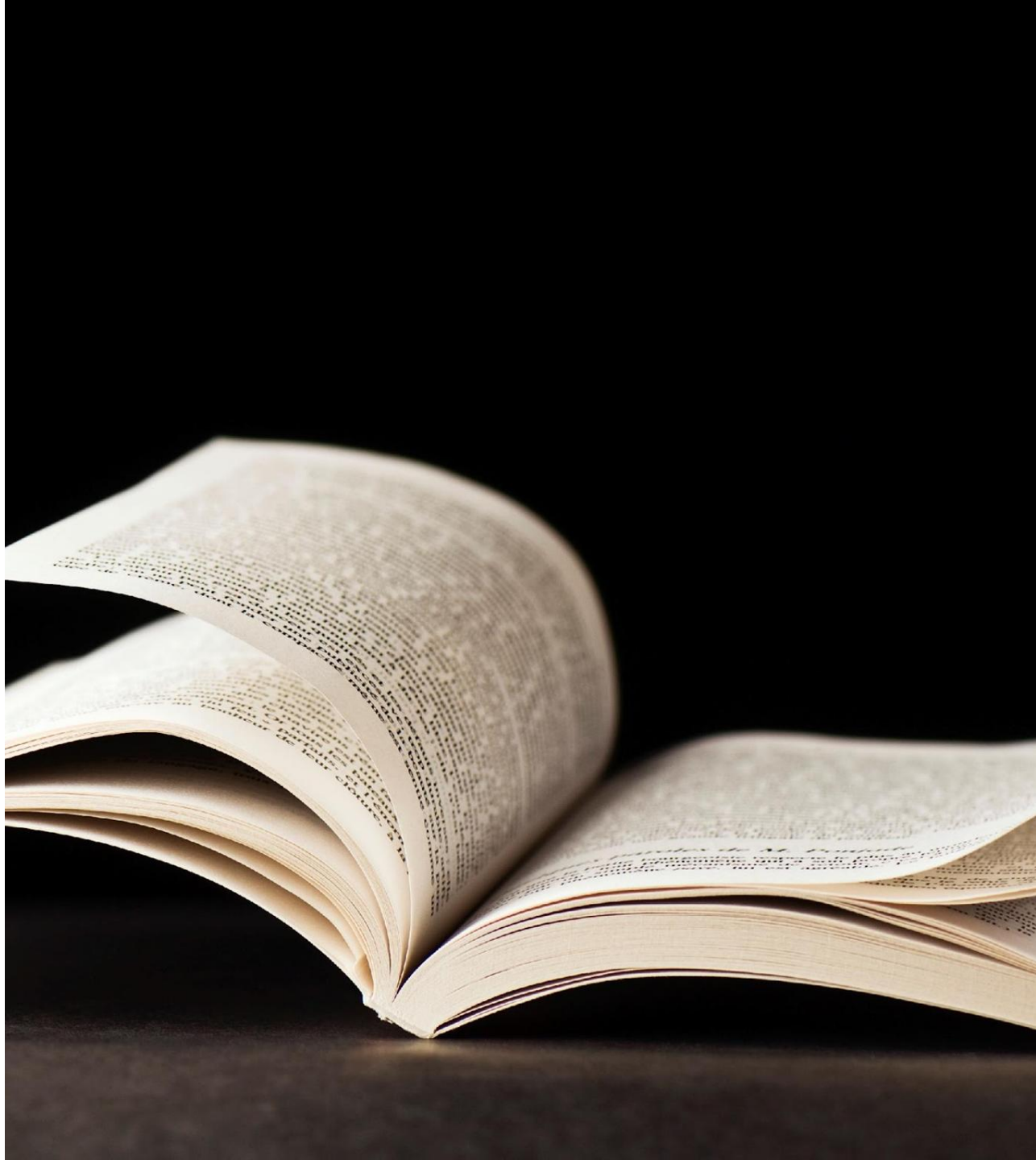
cross-sectional analysis using mortality statistics from the US and 18 comparator nations in the Organisation for Economic Co-operation and Development (OECD18), 5 nationally representative surveys, and electronic health records from a network of 10 pediatric health systems.

[a]The Organisation for Economic Co-operation an
18 comparator nations including Australia, Aust
Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Spain, Sweden, Switzerland, and the United Kingdom.
SUID indicates sudden unexpected infant death.

The Organization for Economic Co-Operation and Development (OECD)



Qualities of Good Storytelling in Data



Jargon: BAD Storytelling

Definition:

Jargon refers to specialized language used by a particular group, profession, or field that may be difficult for outsiders to understand.

Medical Jargon: "The patient was diagnosed with CVA and requires an MRI to assess the extent of the ischemia."

What is the statistician trying to say?

Term	Standard English	Statistician English
“Significant”	“Important”	“An effect has been observed”
“Uniform”	“Without Variation. Sameness”	“Evenly spread over a range (e.g. uniform distribution or design”
“Random”	“Haphazard. Wild. Unpredictable”	“an experiment where the assignment of treatments is deliberately determined by chance”
“Normal”	“Conforming to general standards”	“related to the bell-shaped curve (e.g., the Normal distribution”

Full Disclosure of Methodology

- A good analyst/analysis will provide their/its methodology in a clear, understandable way
    - Examples include showing and explaining Outliers, Nulls, Variance
- An analyst/analysis should be clear in what their/its limitations were
    - Example – providing information on the sample used in the original data collection
    - Example – informing if any data was omitted when performing the analysis

Takeaways

What to be mindful of moving forward

Detective Tips

- Don’t Accept Data at face value
- Be inquisitive
    - “Is this graph employing tricks to be more visually exciting?”
    - “Is this graph showing me the whole picture? Is this the right graph to display this information?”
- Be mindful of Bias
    - “Who created this data I’m seeing? Do they have blind spots? Do I have blind spots?”
- Look for transparency
    - “Is there language used that I don’t understand?” “Can I easily access the steps and methods taken to get to the outcomes of this analysis”

Key Takeaways

- Data can be manipulated visually – look out for common tricks
- Storytellers and listeners come with bias – be aware
- Clear communication is a sign that data can be trusted

Q&A

Thank you!