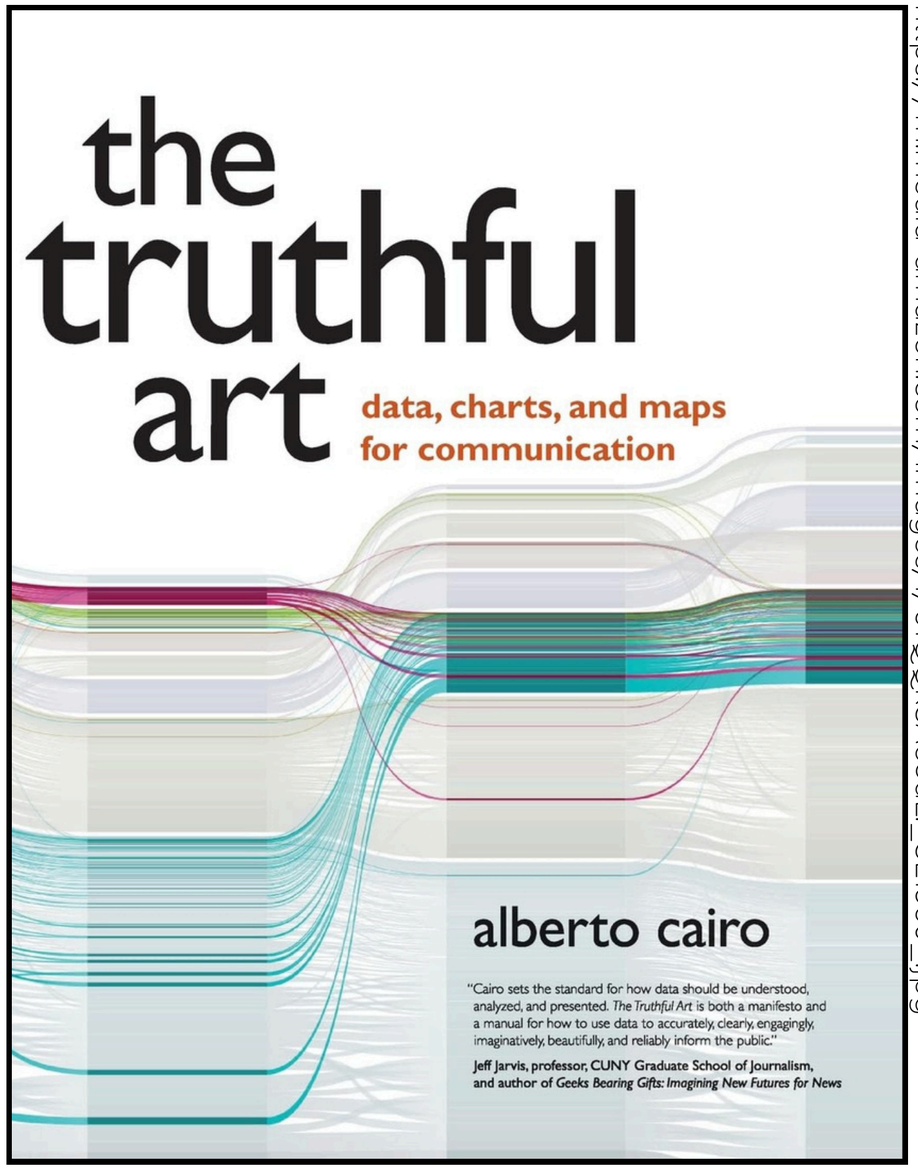
Data Visualization

How to design information in a trustworthy manner

(and why even bother)

Krzysztof Tomasz Stawarz
Social Informatics
Faculty of Humanities AGH
08.10.2024, Cracow



https://m.media-amazon.com/images/I/61QQkSAo5dL._SL1360_.jpg

https://com.miami.edu/profile/alberto-cairo/

Cairo, A. (2016).
The Truthful Art: Data, Charts, and Maps for Communication.
New Riders.

What is Data Visualization?

"the use of computer-supported, interactive,
visual representations of **data** to **amplify cognition**"

Card, S. K., Mackinlay, J., & Shneiderman, B. (1999). Readings in information visualization: Using Vision to Think. (p. 6) Morgan Kaufmann.

"a transformation of quantified **data** which is not
visual into a visual **representation**"

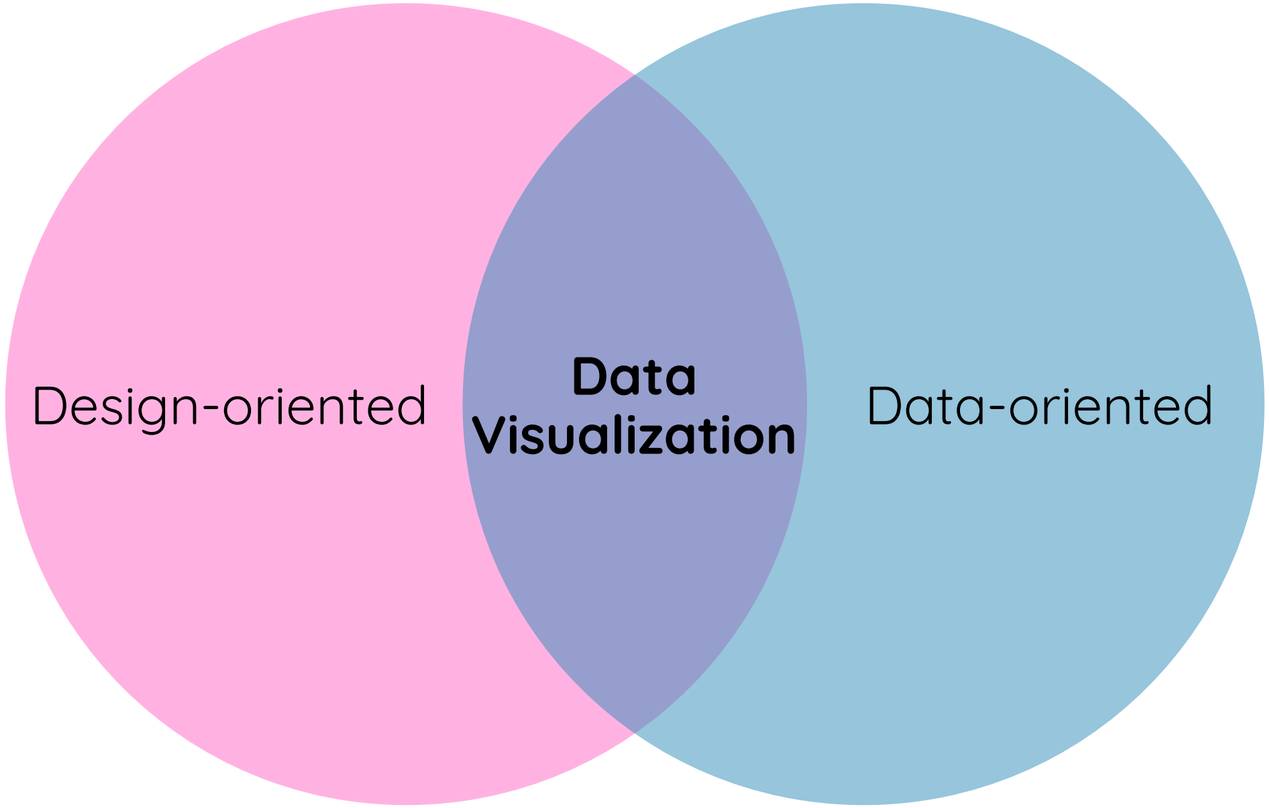
Manovich, L. (2010). What Is Visualization? (p. 20) https://manovich.net/index.php/projects/what-is-visualization

"**information** which has been **abstracted** in some schematic form,
including **attributes** or **variables** for the units of information"

Friendly, M., Denis, D., & Truman, H. (2001). Milestones in the history of thematic cartography, statistica, graphics, and data visualization. (p. 2).

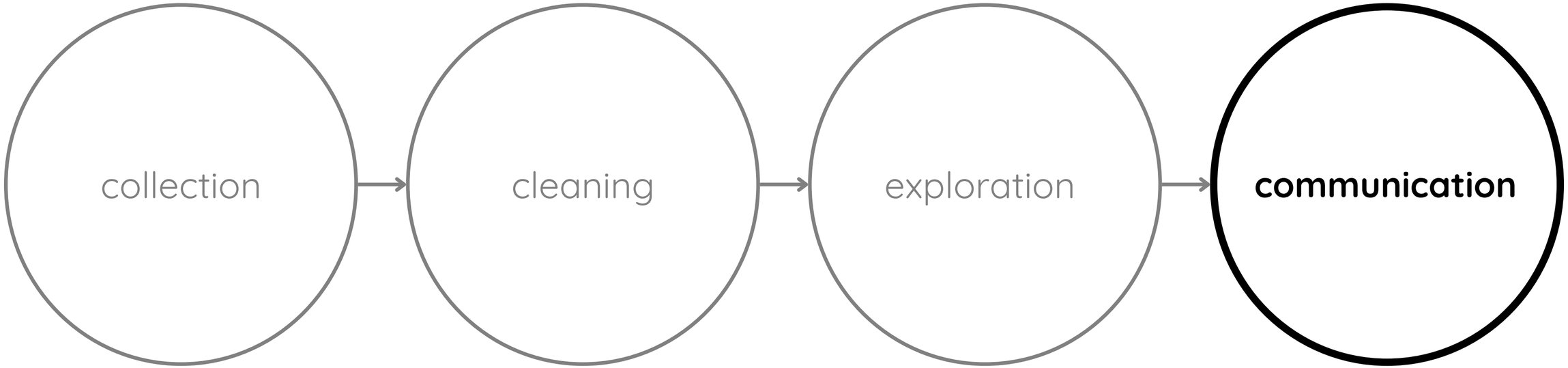
Li, Q. (2020).
Overview of Data Visualization. In Q. Li, Embodying Data (pp. 17–47).
Springer Singapore. https://doi.org/10.1007/978-981-15-5069-0_2

What is Data Visualization?



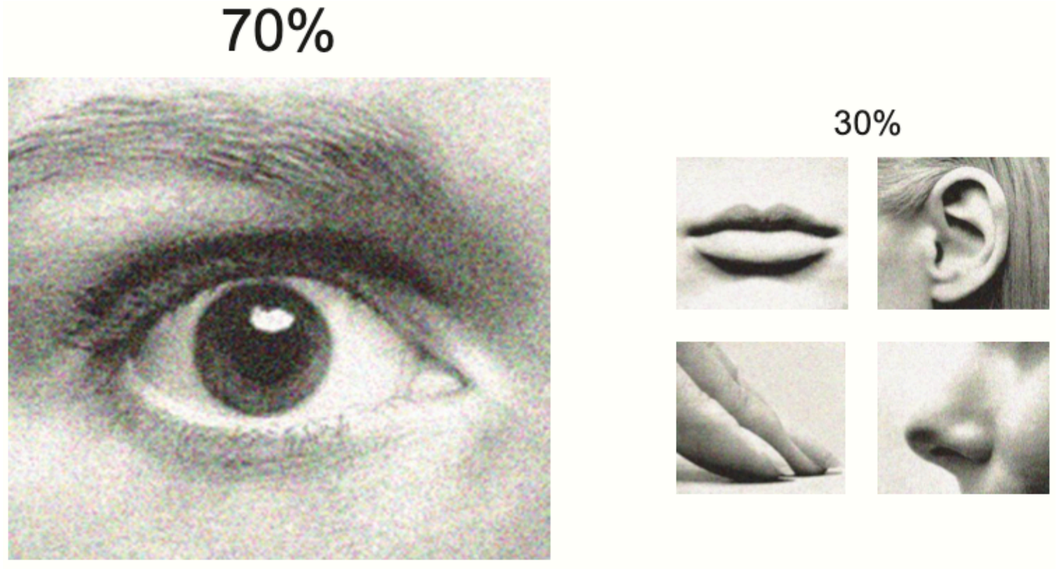
If we think about Social Informatics in terms of two main ideas of approaching the problem, Data Visualization would sit perfectly in the **middle** - it requires both **design literacy and data literacy**.

What is Data Visualization?



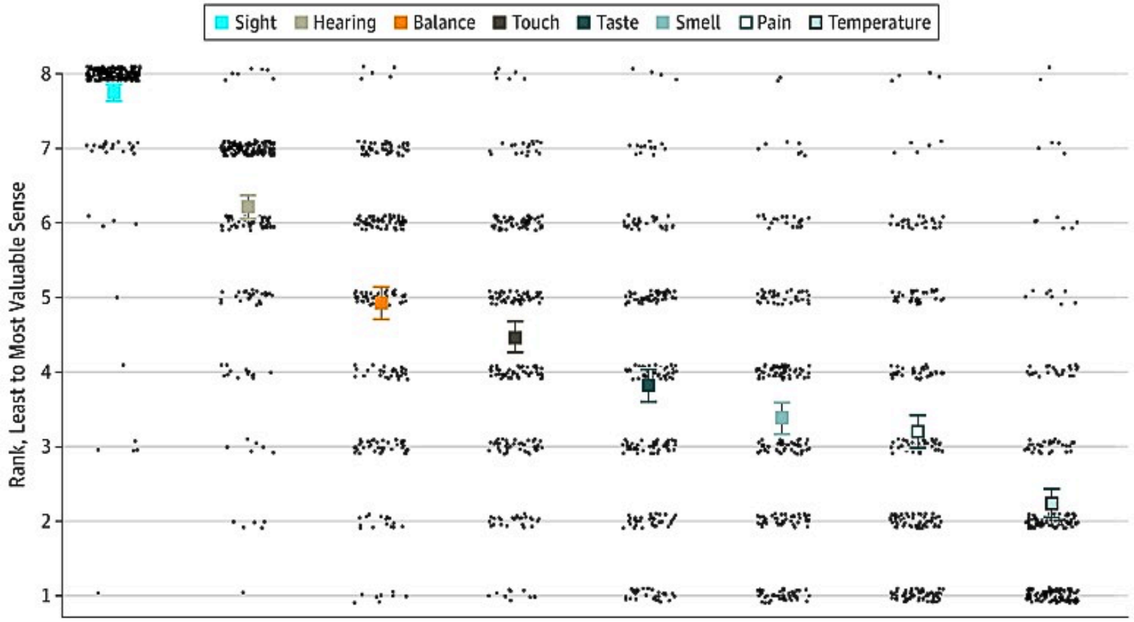
For me, it is the final step in the process of gaining insights from the data:
a clean, informative and trustworthy visual presentation of the information.

Why visually?



Few, S. (2012).
Show me the numbers: Designing Tables and Graphs to Enlighten. (p. 61)

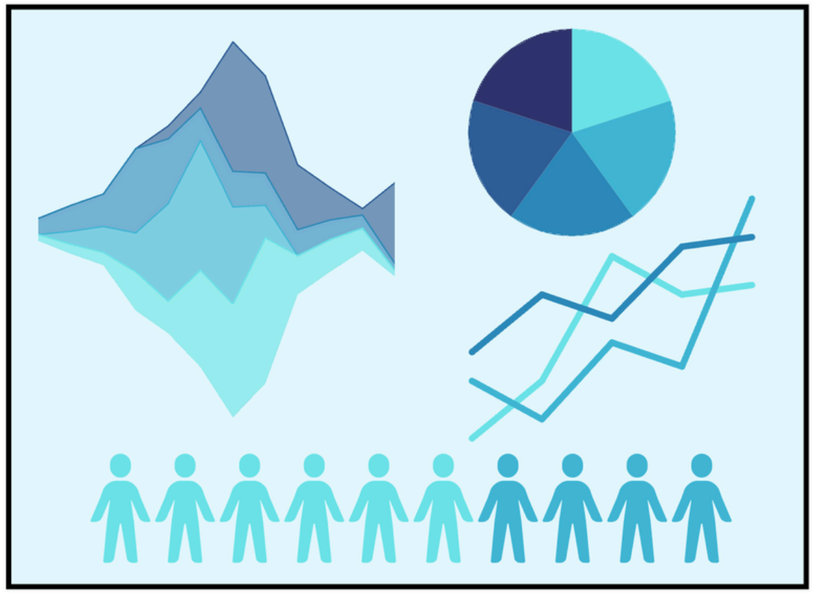
Rehman, I., Hazhirkarzar, B., & Patel, B. C. (2023).
Anatomy, Head and Neck, Eye. In StatPearls. StatPearls Publishing.



Enoch, J., McDonald, L., Jones, L., Jones, P. R., & Crabb, D. P. (2019).
Evaluating Whether Sight Is the Most Valued Sense. JAMA Ophthalmology, 137(11), 1317.

Biologically speaking, our brain processes approximately **50-70%** of all the **information** it receives **through our eyes**. Moreover, when asked about general usability, the average person tends to rank **sight as the most valuable of all senses**.

Why visually?



Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

In general, an information is more likely to grab our attention if it is presented in a simple, graphical way rather than in written or spoken form. **We subconsciously look for graphical simplicity** when, for example, scrolling, watching a keynote, or reading academic articles.

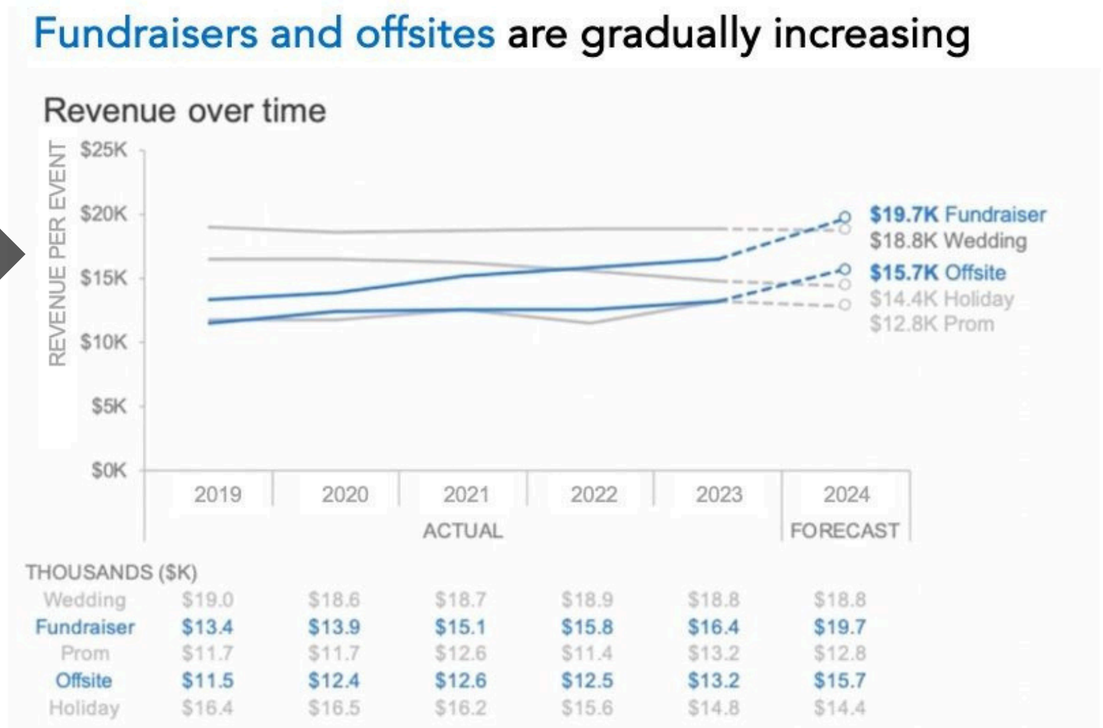
Why visually?

ORIGINAL VIEW

Revenue in thousands by event type					
Date	Wedding	Fundraiser	Prom	Offsite	Holiday
12/31/19	$ 19.0	$ 13.4	$ 11.7	$ 11.5	$ 16.4
12/31/20	$ 18.6	$ 13.9	$ 11.7	$ 12.4	$ 16.5
12/31/21	$ 18.7	$ 15.1	$ 12.6	$ 12.6	$ 16.2
12/31/22	$ 18.9	$ 15.8	$ 11.4	$ 12.5	$ 15.6
12/31/23	$ 18.8	$ 16.4	$ 13.2	$ 13.2	$ 14.8
12/31/24	$ 18.8	$ 19.7	$ 12.8	$ 15.7	$ 14.4

https://www.linkedin.com/posts/storytelling-with-data-llc_dataviz-presentation-activity-7244357862498430979-MH4h?utm_source=share&utm_medium=member_desktop

AUGMENTED VIEW



Why bother about the design?

How many **‘n’**s are there?

mmnmmmmmmm

mmmmmmmmmm

mmmmmmmmnm

mmnnmmmmmm

mmmmmmmmmm

mmmnmmmmmm

Why bother about the design?

How many **‘n’**s are there?

mmnmmmmmmm

mmmmmmmmmm

mmmmmmmmmnm

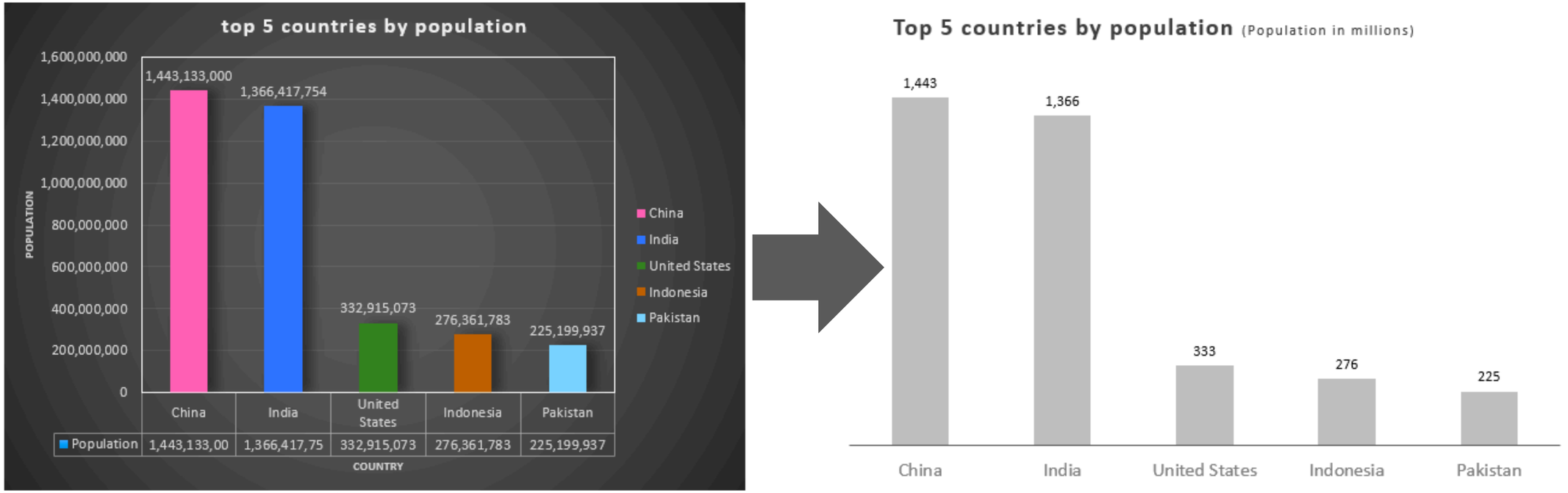
mmnnmmmmmm

mmmmmmmmmm

mmmnmmmmmm

And how about now?

Why bother about the design?



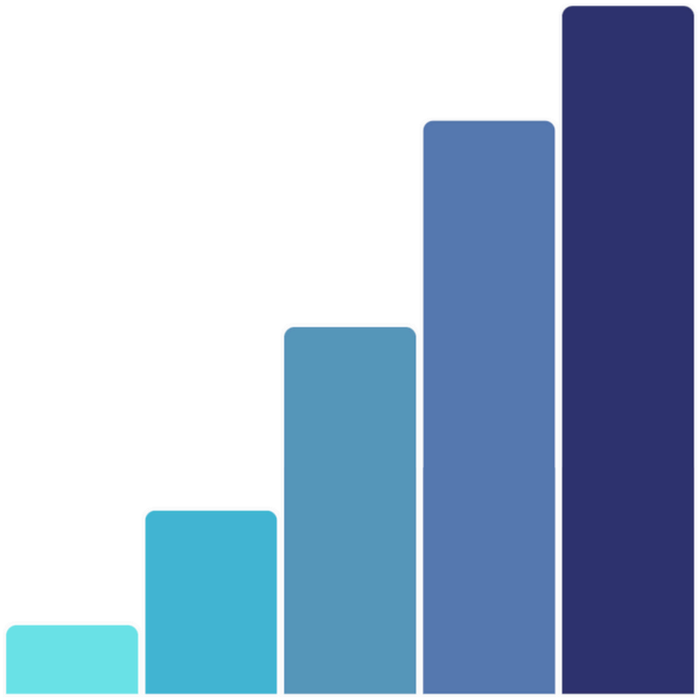
https://www.linkedin.com/pulse/improving-data-visualization-using-principle-data-ink-wijesinghe/

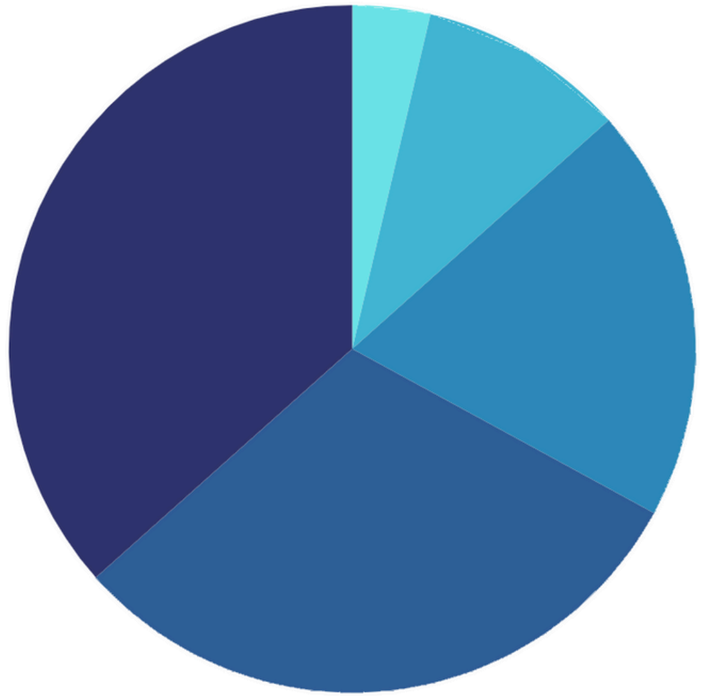
By adopting **design schemas** such as **Gestalt** principles or the **Graphical Integrity and Graphical Excellence** principles proposed by Edward R. Tufte (2001), we can enhance our ability to communicate data in a **clear, informative, and trustworthy** way.

Top 3 tips on how to make our visualizations genuine

(and how to avoid being misled when reading them)

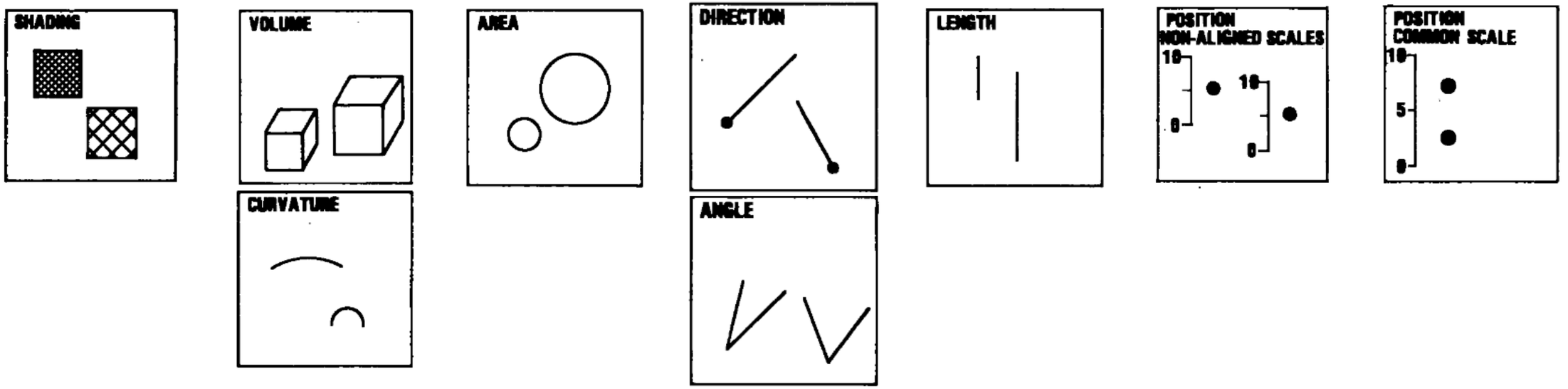
Tip 1: Bar chart is almost always better than a pie chart





Least accurate

Most accurate

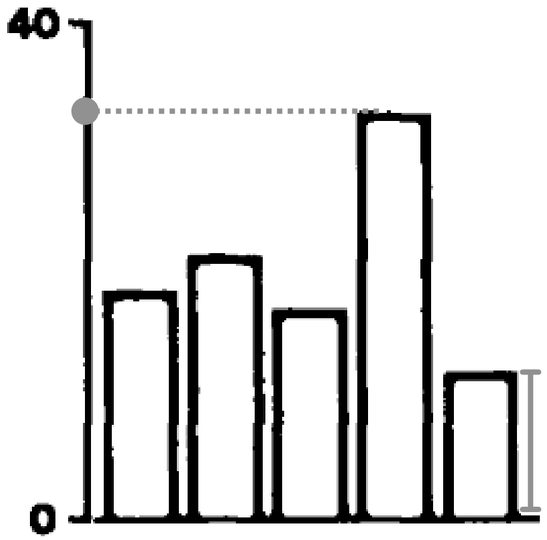


In 1984 William S. Cleveland and Robert McGill proposed the ranking of **visual variables**
(term created by Jacques Bertin in 1967 “Semiology of Graphics”),
based on the mean error people expierience while **decoding** values encoded with them.

all the theory and graphics come from the original paper:

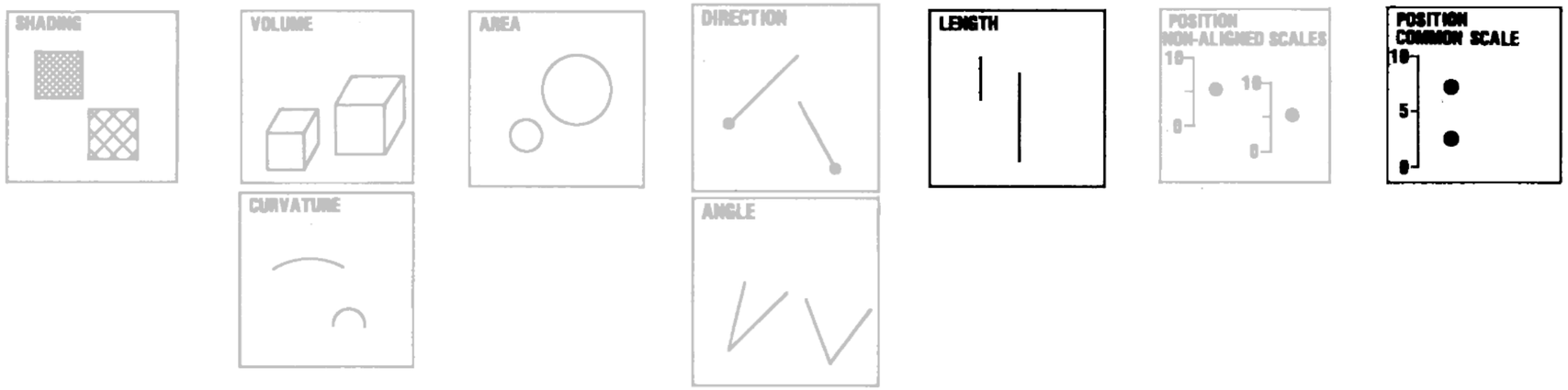
Cleveland, W. S., & McGill, R. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. Journal of the American Statistical Association, 79(387), 531–554. https://doi.org/10.1080/01621459.1984.10478080

Bertin, J., & Berg, W. J. (1967). Semiology of Graphics: Diagrams, Networks, Maps (1st ed). ESRI Press : Distributed by Ingram Publisher Services.



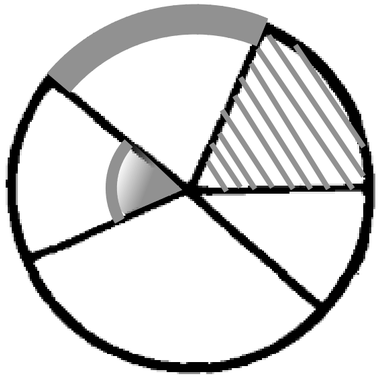
Least accurate

Most accurate

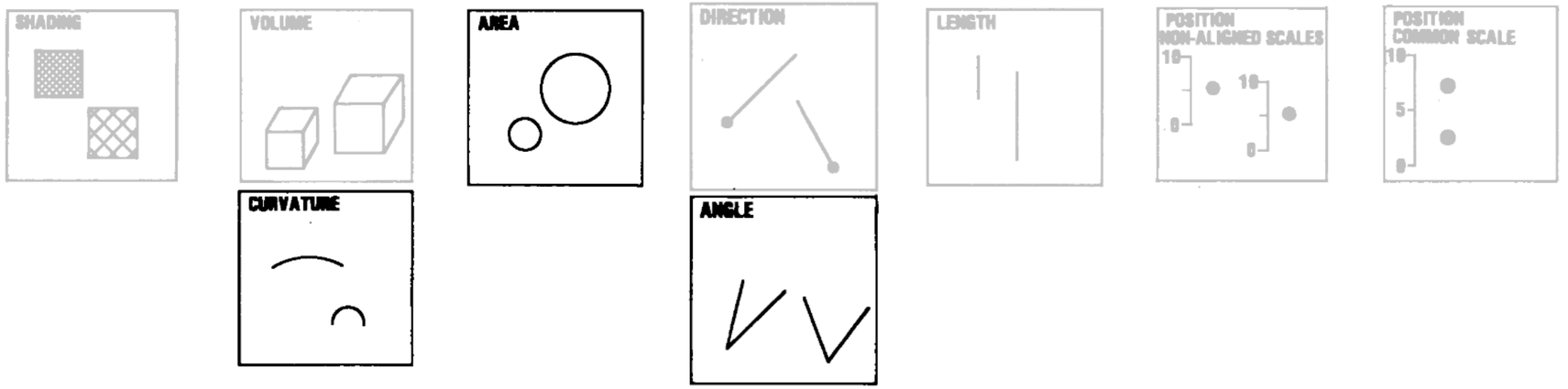


When reading a **bar chart**, people subconsciously decode variables such as:
position on a common scale and a bar **length**,
which are relatively high on the Cleveland's ranking.

On the other hand, **pie charts** require us to decode their **curvature/arc length, angle, and pie slice area**. This forces us to use significantly more cognitive resources to interpret them correctly.

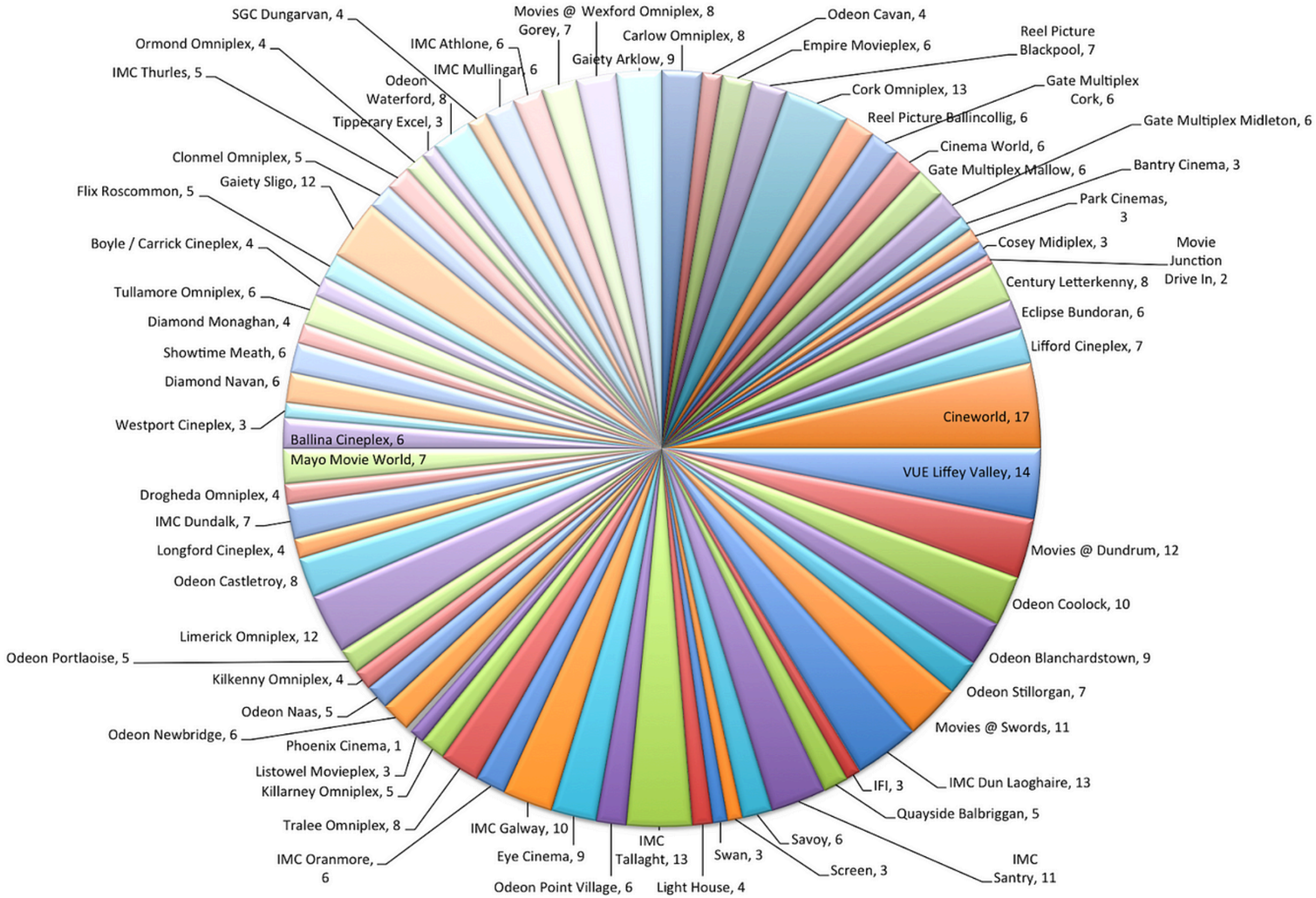


Least accurate → Most accurate



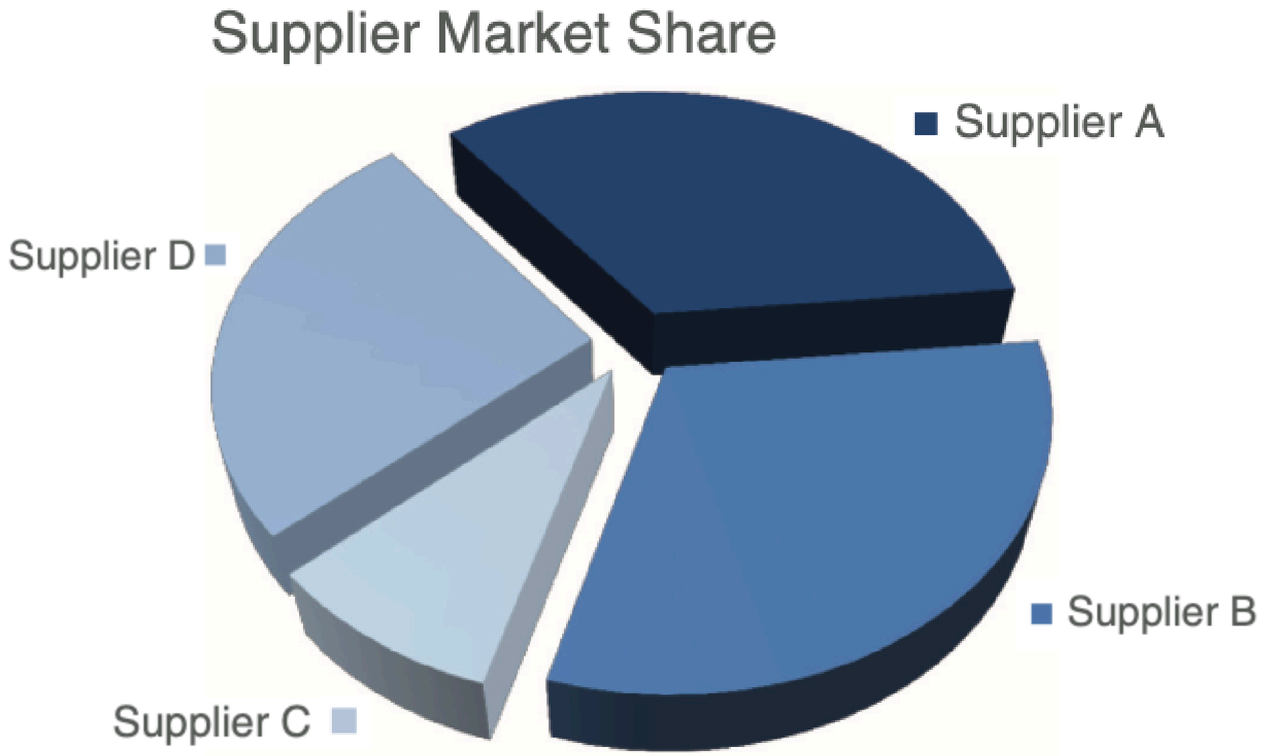
Although they may look “cooler”, **use pie charts instead of bar charts only when comparing values that add up to a specific, arbitrary whole that you want to emphasize**.
Otherwise, in most real life scenarios, bar charts are objectively the better choice.

https://medium.com/@stevoscript/why-not-pie-4752f086c32



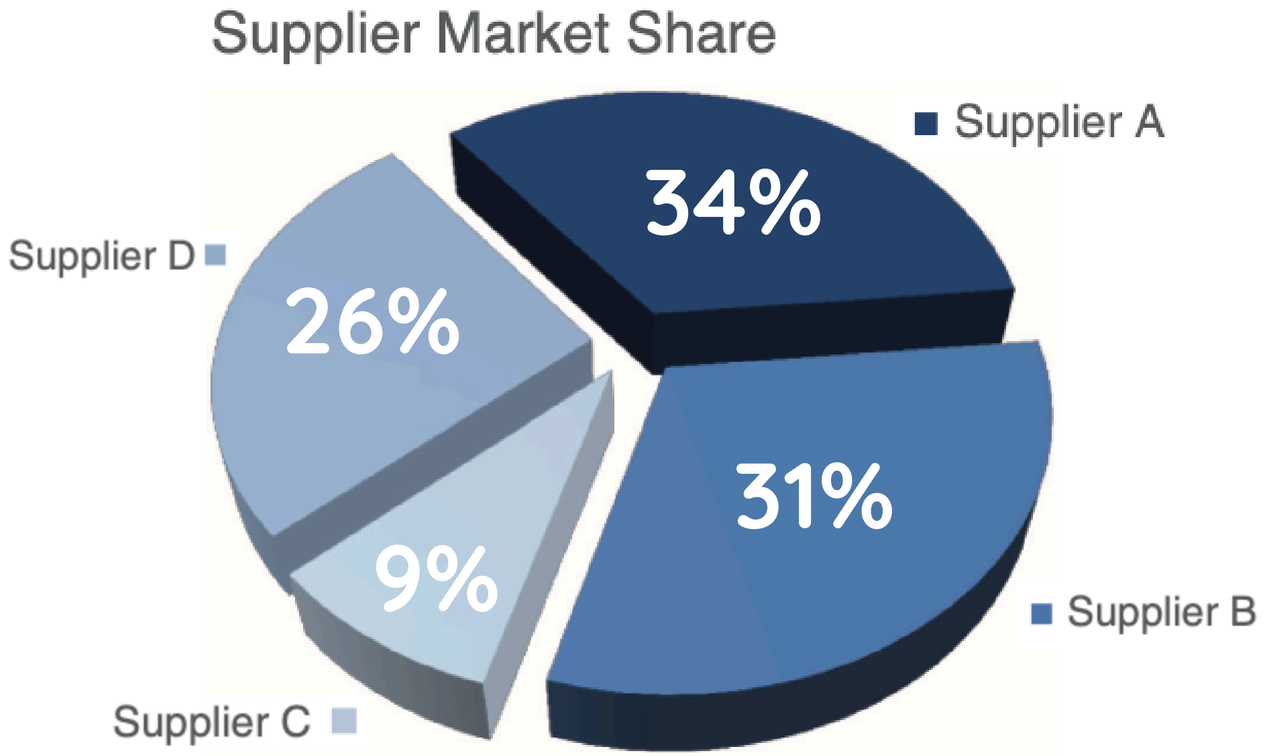
Oh, almost forgot: **Use up to 4-5 five slices per pie!**
If your dataset has more observations, combine them into “Others” category or even better - reconsider your chart choice ;)

Which supplier has the biggest market share?



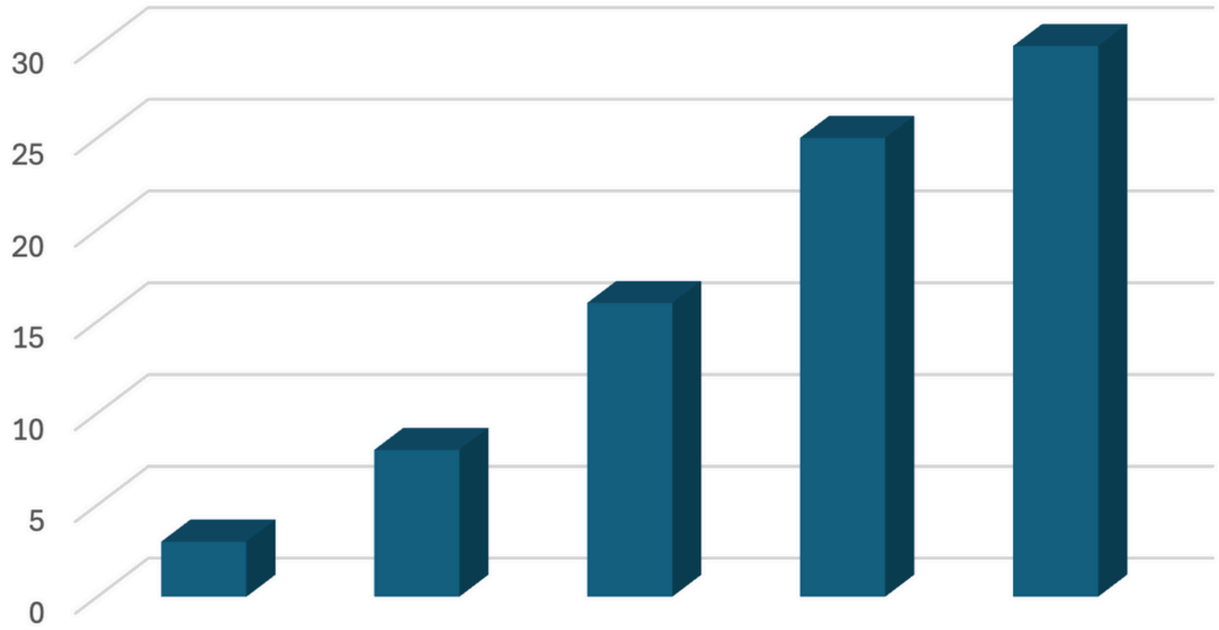
Nussbaumer Knaflic, C. (2015). Storytelling with data: A data visualization guide for business professionals. (p. 62-65) Wiley.

Which supplier has the biggest market share?



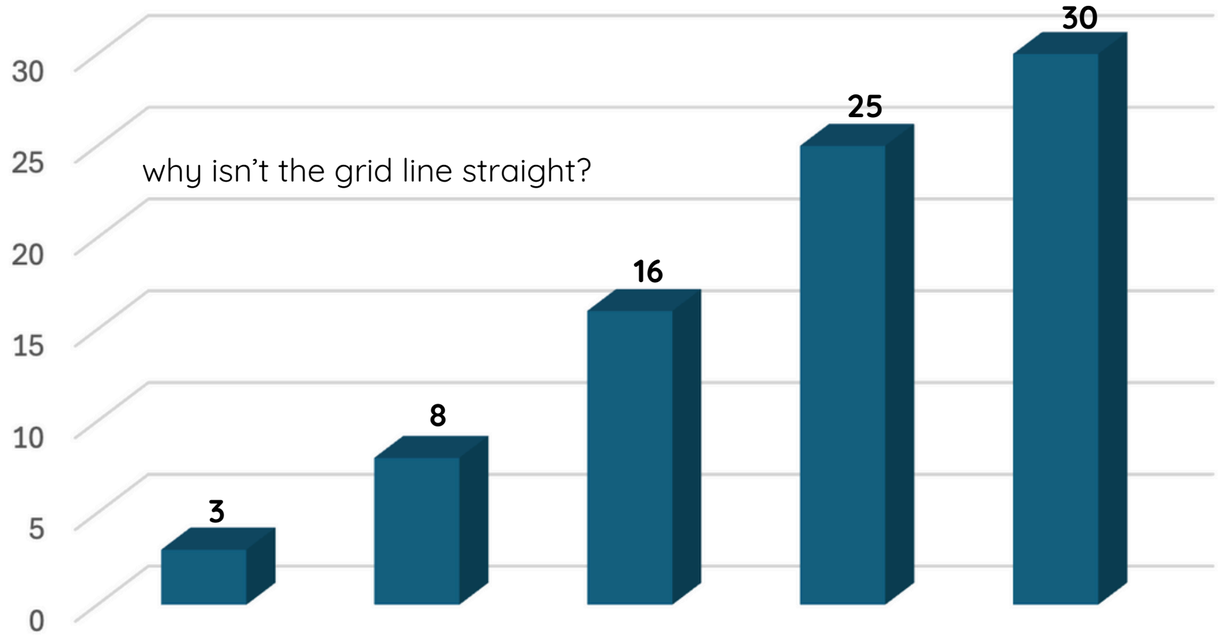
Nussbaumer Knaflic, C. (2015). Storytelling with data: A data visualization guide for business professionals. (p. 62-65) Wiley.

What is the value of the middle bar?



What is the value of the middle bar?

Label	Value
l1	3
l2	8
l3	16
l4	25
l5	30



It’s **16**. But Microsoft Excel generated the grid that apparently is moved one unit up due to 3D perspective.

Tip 2: Avoid 3D perspective

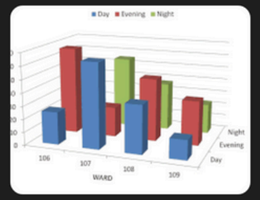
https://stackoverflow.com/questions/22369224/3d-pie-chart-in-highcharts-javascript

Google

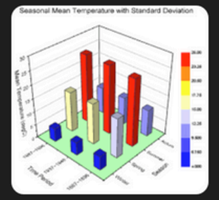
3d chart

Wszystko | Grafika | Produkty | Wideo | Wiadomości | Książki | Sieć | Więcej | Narzędzia

Bar | Pie | Scatter | Tableau | Powerpoint | Highcharts | Excel | Wpf | Telerik | Three js | 3d infographic

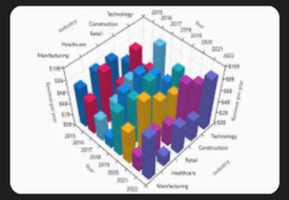


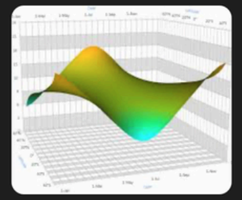
PowerPoint and Presenting Stuff - …
Problems with 3D Charts

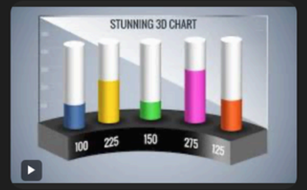


OriginLab
Help Online - Origin Help …

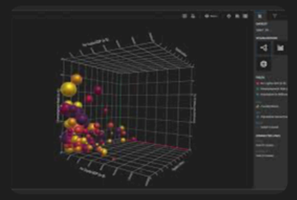

Data Viz Project
Multiple Series 3D Bar…

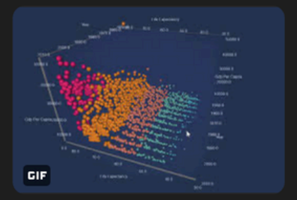

Telerik.com
WPF 3D Chart Control | Telerik UI fo…


Telerik.com
New 3D Chart in UI for WPF

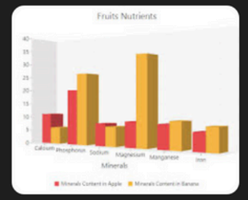


YouTube
Stunning 3D Chart tutorial in PowerPoi…

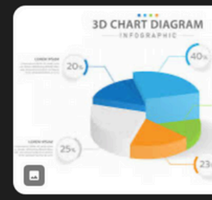

Microsoft
Charts 3D | Microsoft Garage



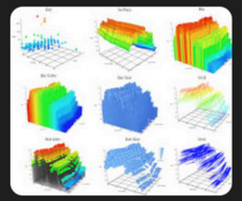
SciChart
SciChart.JS - Fast Javascript 3D Ch…

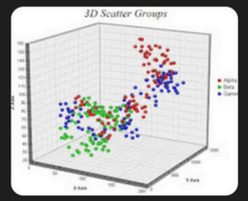


Help.Syncfusion.com
3D Chart types available in Sy…



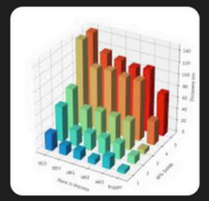
Adobe Stock
3D Graph Chart Images – Bro…


Kx Systems
3D Chart - KX Dashboards

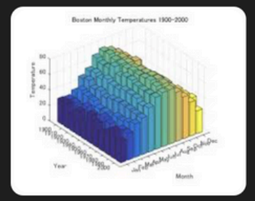


ChartDirector
3D Scatter Charts - ChartDire…

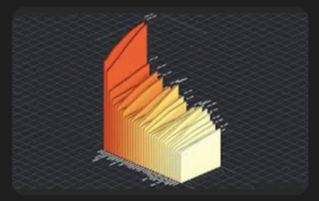

4D Blog
How to use Google Charts with 4D –…


Syncfusion
3D chart Support in Blaz…

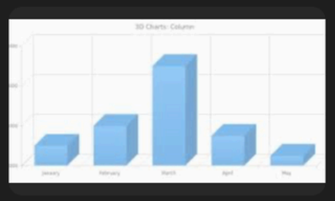

Stack Overflow
3D Pie chart in Highcharts/Javasc…



Data Viz Project
Multiple Series 3D Bar Chart | …

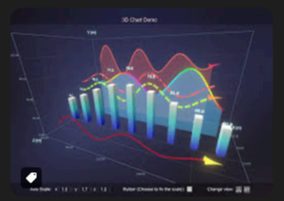

three.js forum
Creating a 3D Area Chart with D3.js & T…

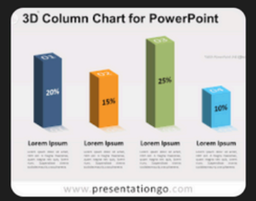

Adobe Stock
3D Graph Images – Browse 647…



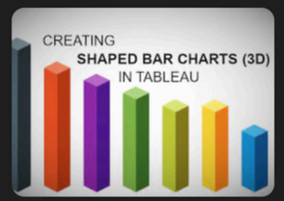
AnyChart Documentation
Overview | 3D | Basic Charts | AnyChart Do…

iStock
Statistics Busines

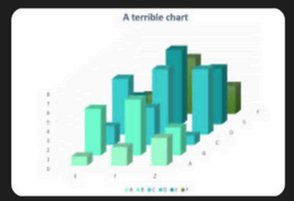

Unity Asset Store · W magazynie
3D Chart Simulation | Modeling | U…



PresentationGO
3D Column Chart for PowerPoi…



Toan Hoang
Shaped Bar Charts (3D) in Tableau…



Towards Data Science
When to use 3D in data visualisation …

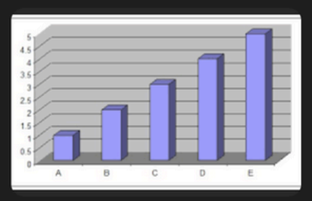
Podobne wyszukiwania

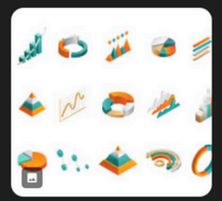
graph 3d

3d chart png

3d graphs

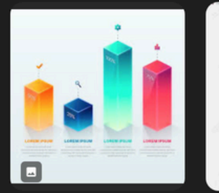
YouTube
Beautiful 3D BAR CHART Design Tutori…


Consultant Journal
Use 3D charts and graphs at your own …


Vecteezy
3d Chart Vector Art, Icons…


Free Vector Art & Graphics
3 D Charts Graphics Vector Art …


Pngtree
3d Chart PNG T

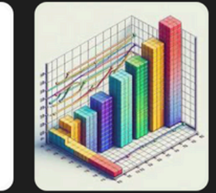

Freepik
3d Chart: wektory i ilu…


123RF
3d Wykres Wykres…

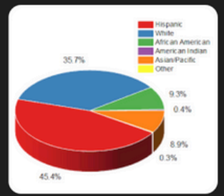
Prezibase · W magazynie
3D Bar Chart Maker Prezi Template | P…

Medium
Data Visualization: Why 3D charts are a terrible idea | b…

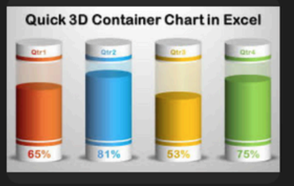

Vecteezy
3d minimal bar chart. data ana…


Easy-Peasy.AI
Color-Coded 3D Bar …


Adobe Stock
3D Graph Chart Images – Bro…

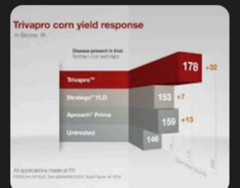


OriginLab
Help Online - Origin Help - …



PK: An Excel Expert
Quick 3D Container chart in Excel - …


Medium
Data Visualization: Why 3D charts a



Podobne wyszukiwania

infographic 3d pie chart

graph 3d icon

Tip 2: Avoid 3D perspective

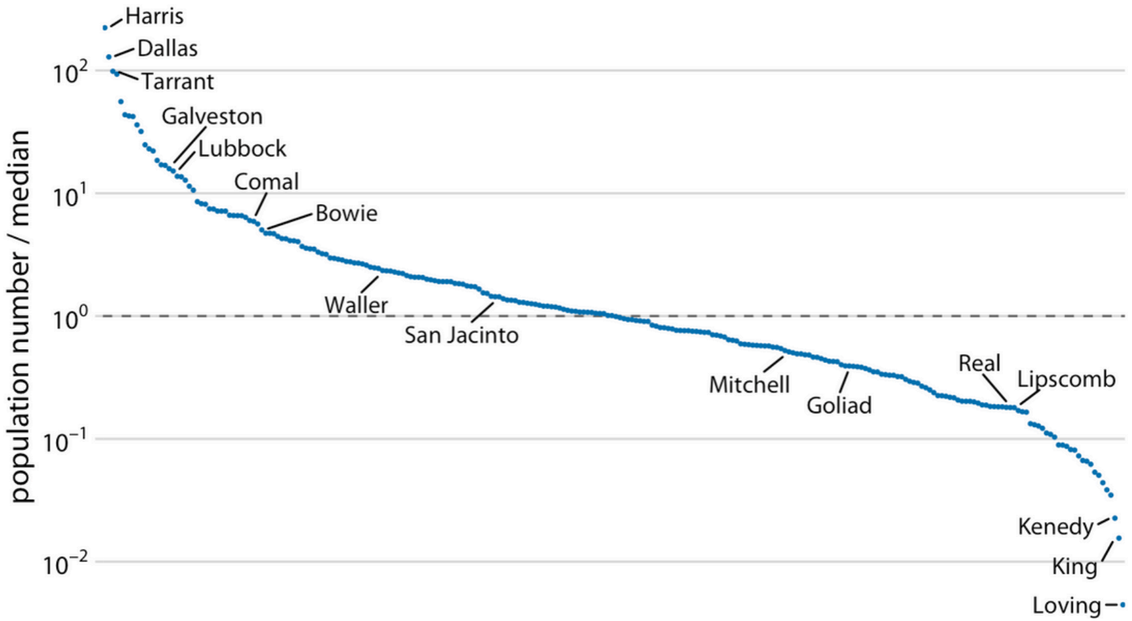
"[…] Notably, **in none of these cases does the third dimension convey any actual data**. **3D is used simply to decorate and adorn the plot**. I consider this use of 3D as gratuitous. It is unequivocally bad and should be erased from the visual vocabulary of data scientists."

Wilke, C. O. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures (First edition; chap. 26, par. 1). O'Reilly.

Read more about **Liefactor** and **Chartjunk** in:

Tufte, E. R. (2013). The Visual Display of Quantitative Information (2nd ed., 8th print). Graphics Press.

How would you name the general trend of the data?

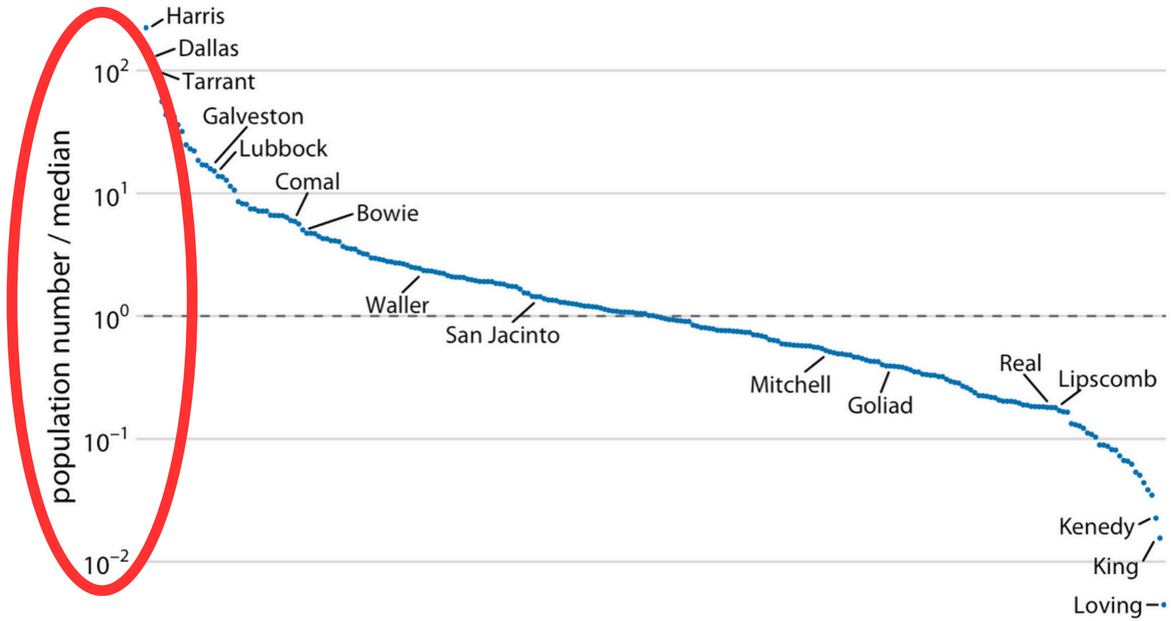


Wilke, C. O. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures (First edition; chap. 3.2). O'Reilly.

It is linear, right?

How would you name the general trend of the data?

Look at the y-axis ticks!

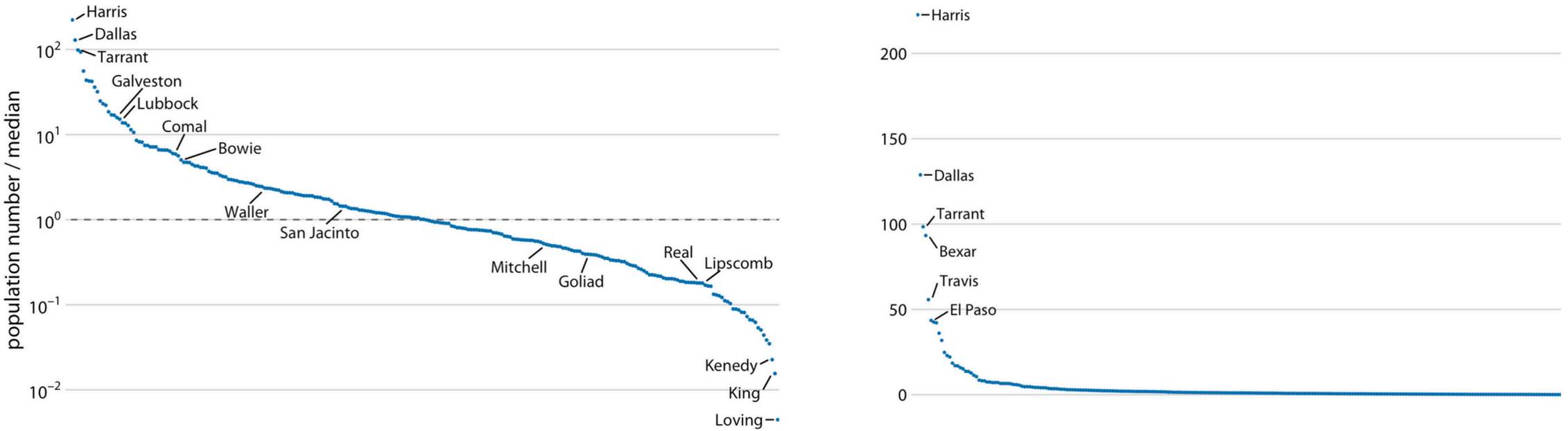


Wilke, C. O. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures (First edition; chap. 3.2). O'Reilly.

It is linear, right?

WRONG!

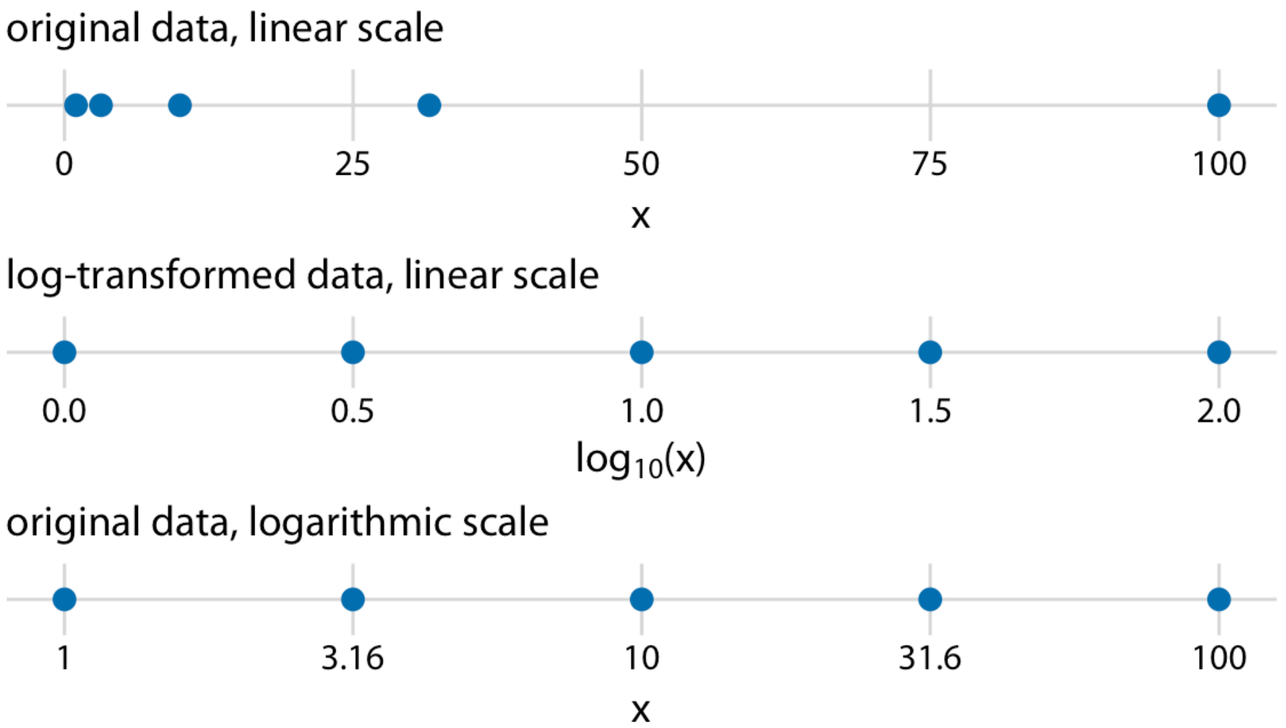
How would you name the general trend of the data?



Wilke, C. O. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures (First edition; chap. 3.2). O'Reilly.

It LOOKS linear because of the **logarithmic transformation** on the data.

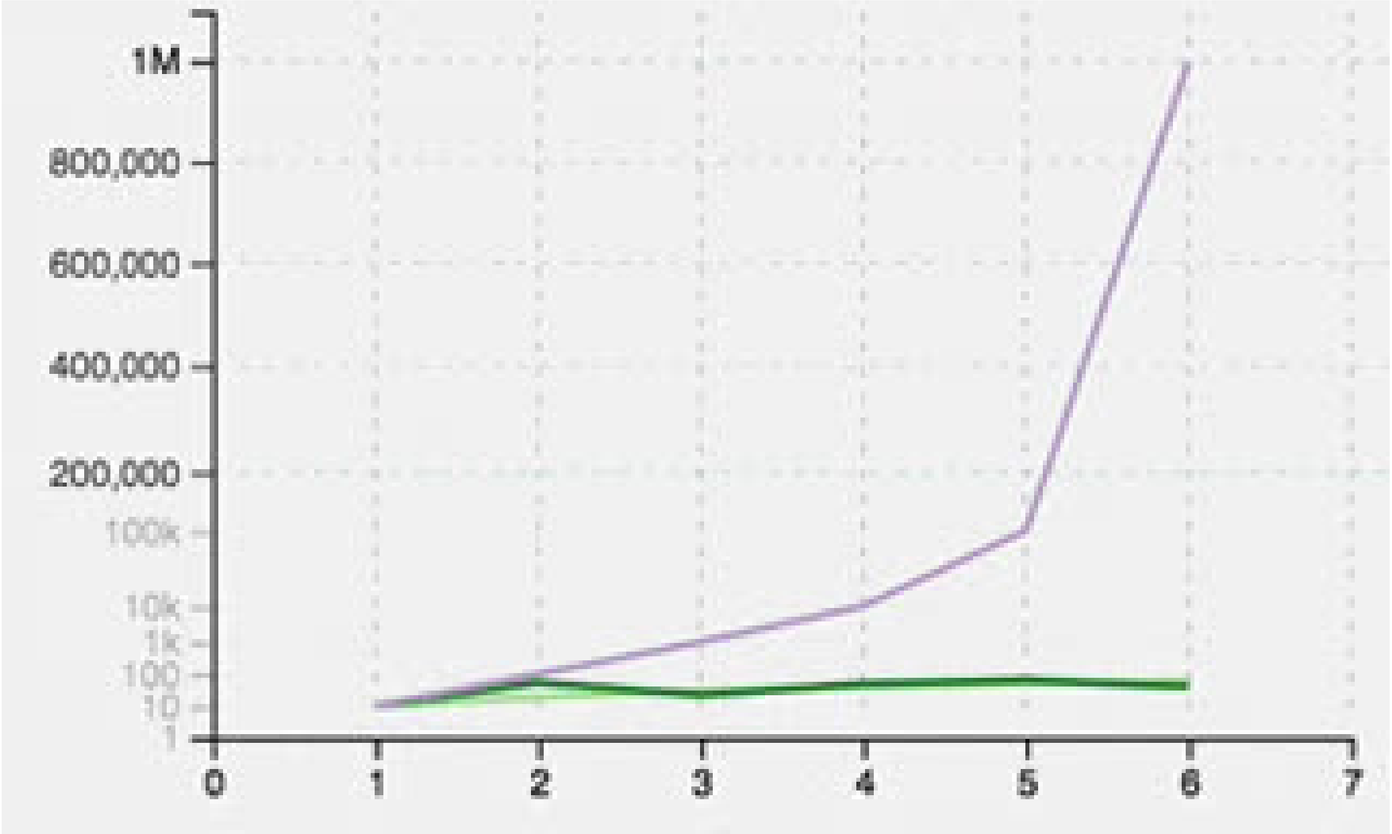
Tip 3: Check your axes twice



Wilke, C. O. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures (First edition; chap. 3.2). O'Reilly.

There are two ways the data can be non-linearny presented: either the **data** itself is **log-transformed** or the **scale** is. Either way, it is important to **know your audience** and **avoid using log-transformations when communicating with non-technical people**.

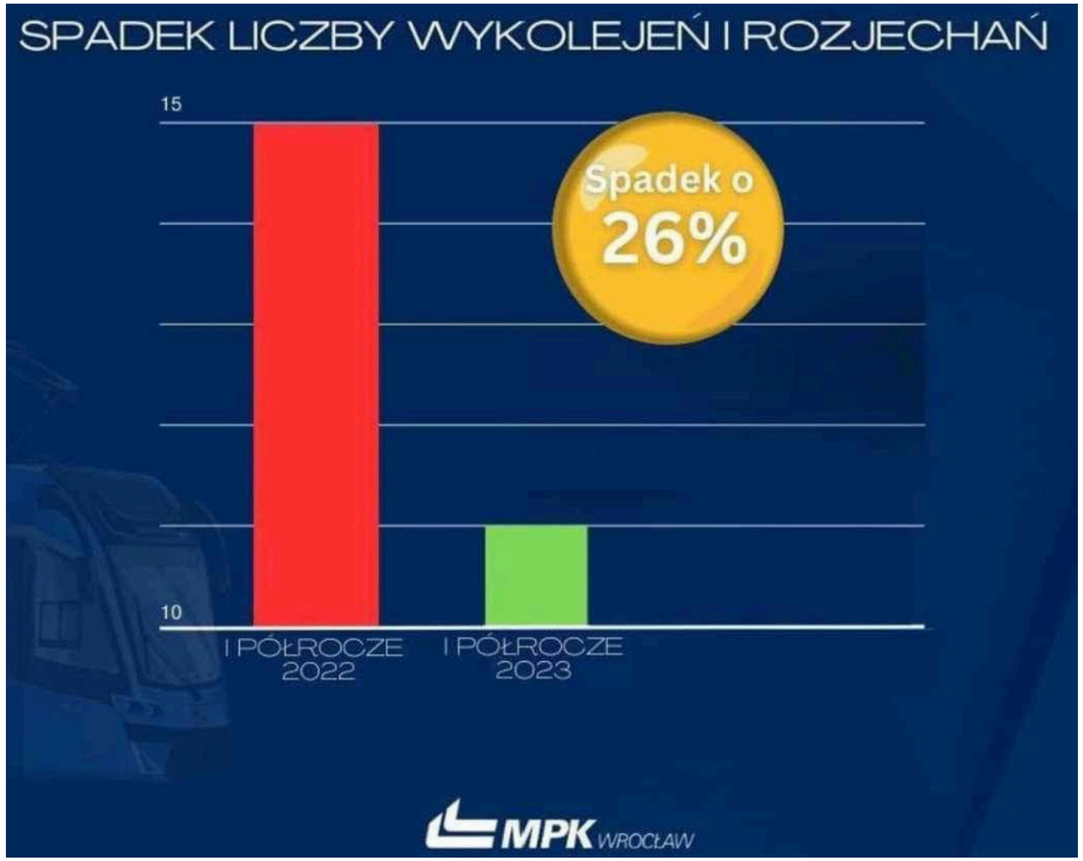
Tip 3: Check your axes twice



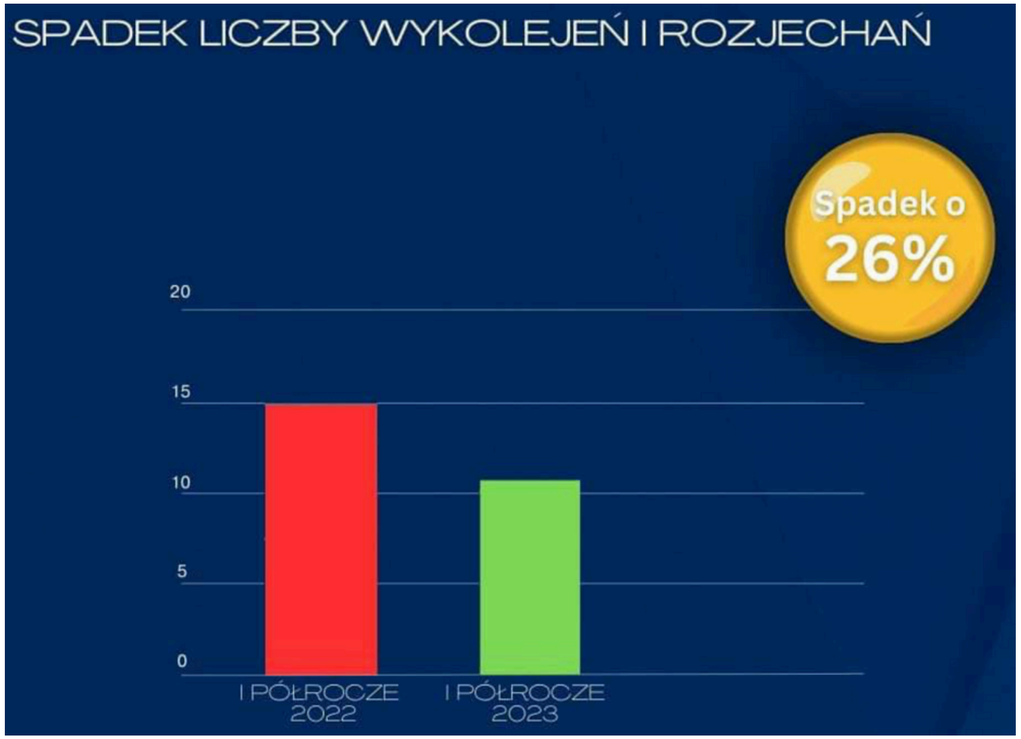
https://scienceprimer.com/log-scale-overview

As for viewers - it is crucial to **check the axes and scales** before judging the shape of the plot.
Draw your conclusions only after meticulous inspection.

Tip 3: Check your axes twice



original screenshot from July 2023

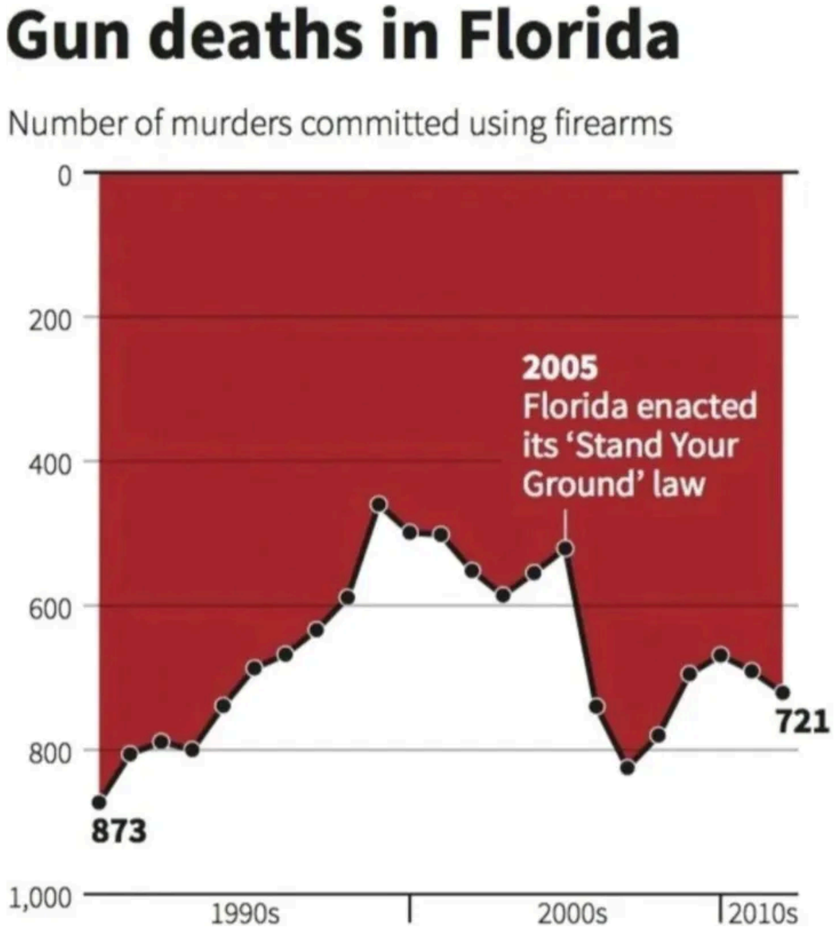


they fortunately corrected it couple days later

https://www.facebook.com/photo.php?fbid=604336548507785&set=pb.100067943363854.-2207520000&type=3

26% drop? I'm seeing at least 400% ;)

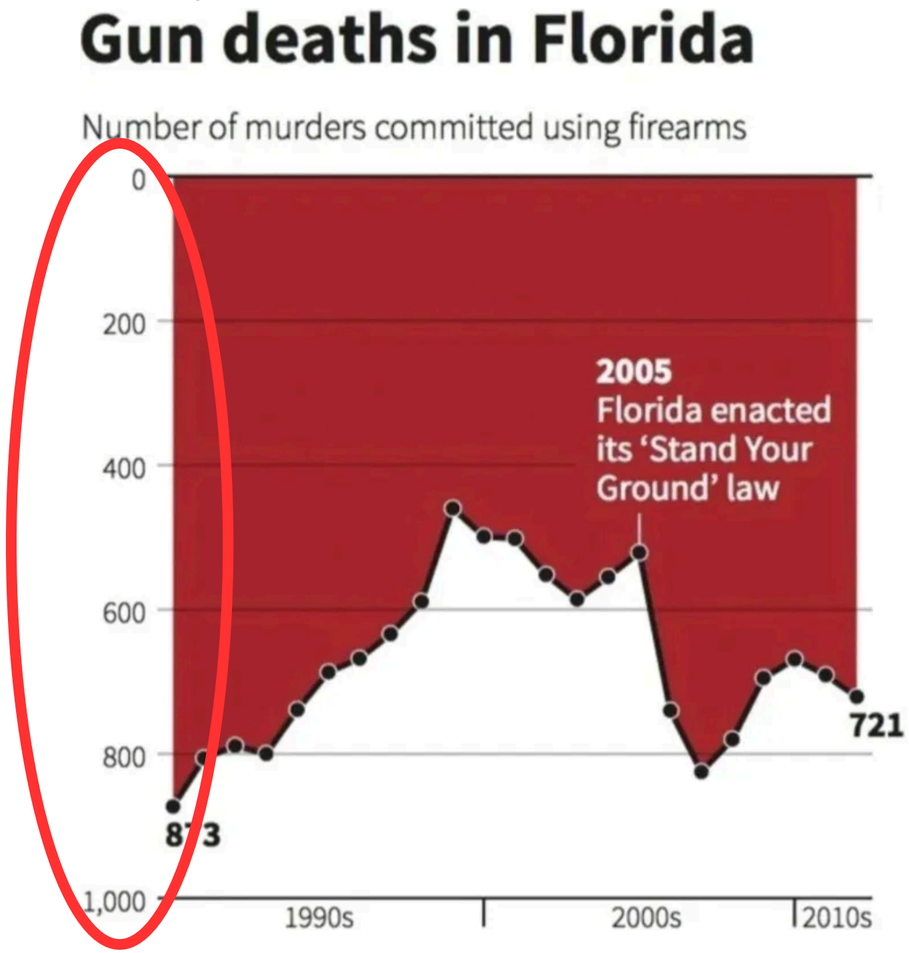
Tip 3: Check your axes twice

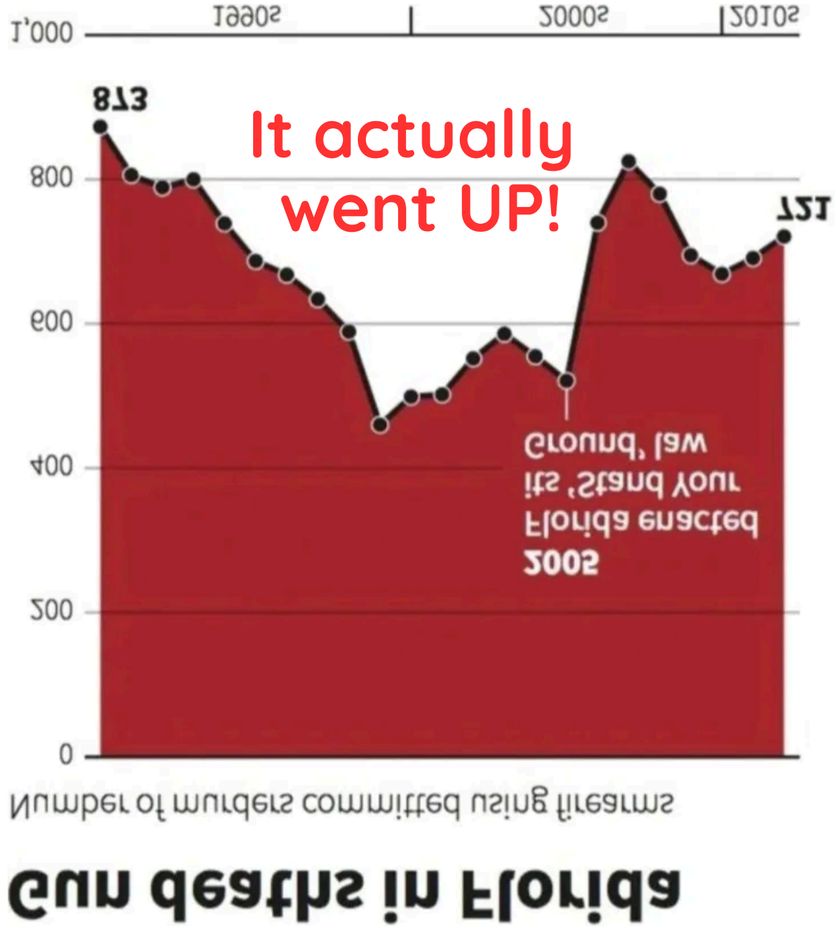


Murders went down after the law was enacted?

https://www.researchgate.net/figure/Reversed-Y-axis-figure-showing-the-relationship-between-the-population-of-species-X-and_fig3_330945404

Tip 3: Check your axes twice





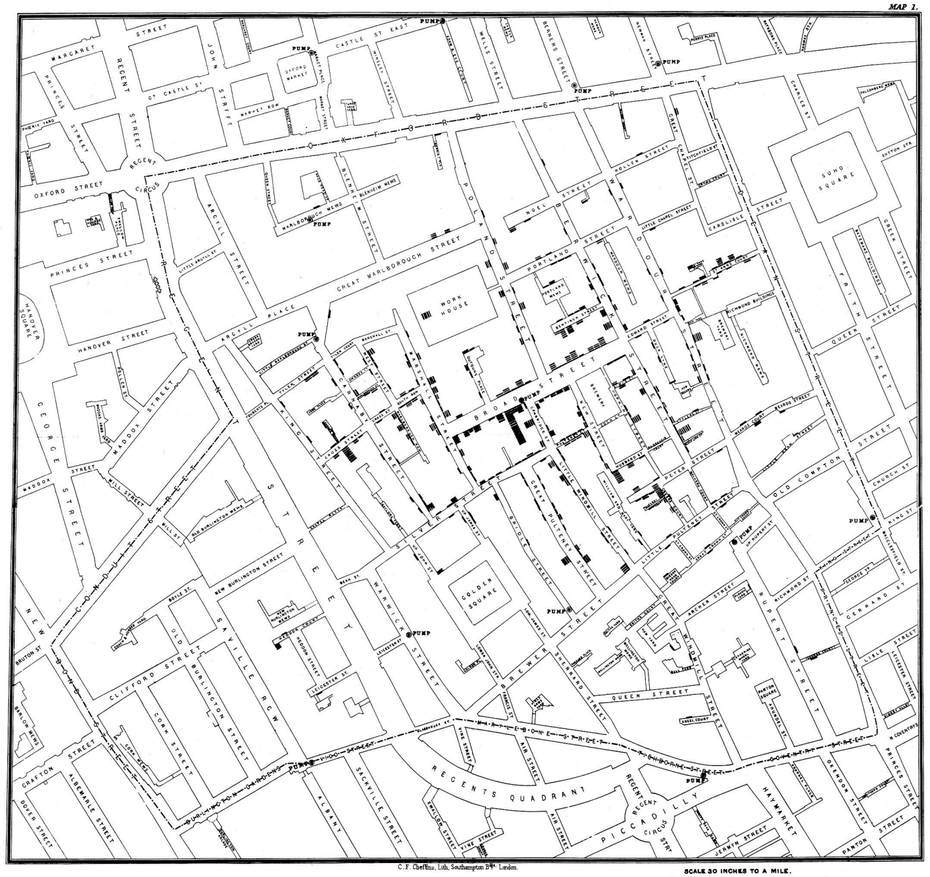
Murders went down after the law was enacted?

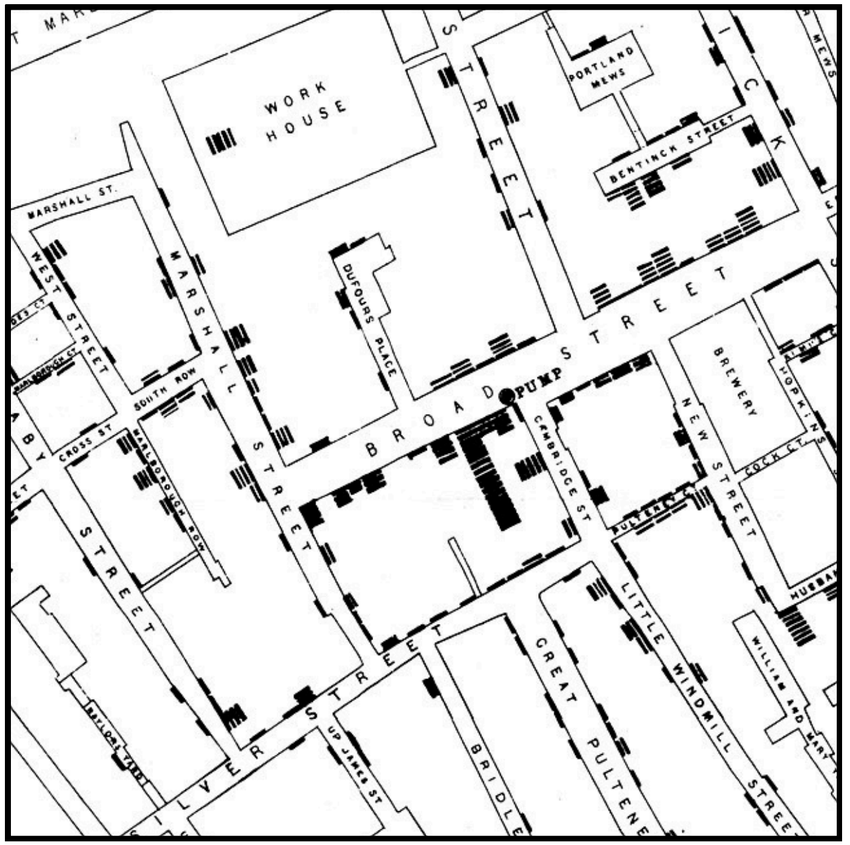
https://www.researchgate.net/figure/Reversed-Y-axis-figure-showing-the-relationship-between-the-population-of-species-X-and_fig3_330945404

Why bother about the truthfulness?

https://www.youtube.com/watch?v=dQw4w9WgXcQ

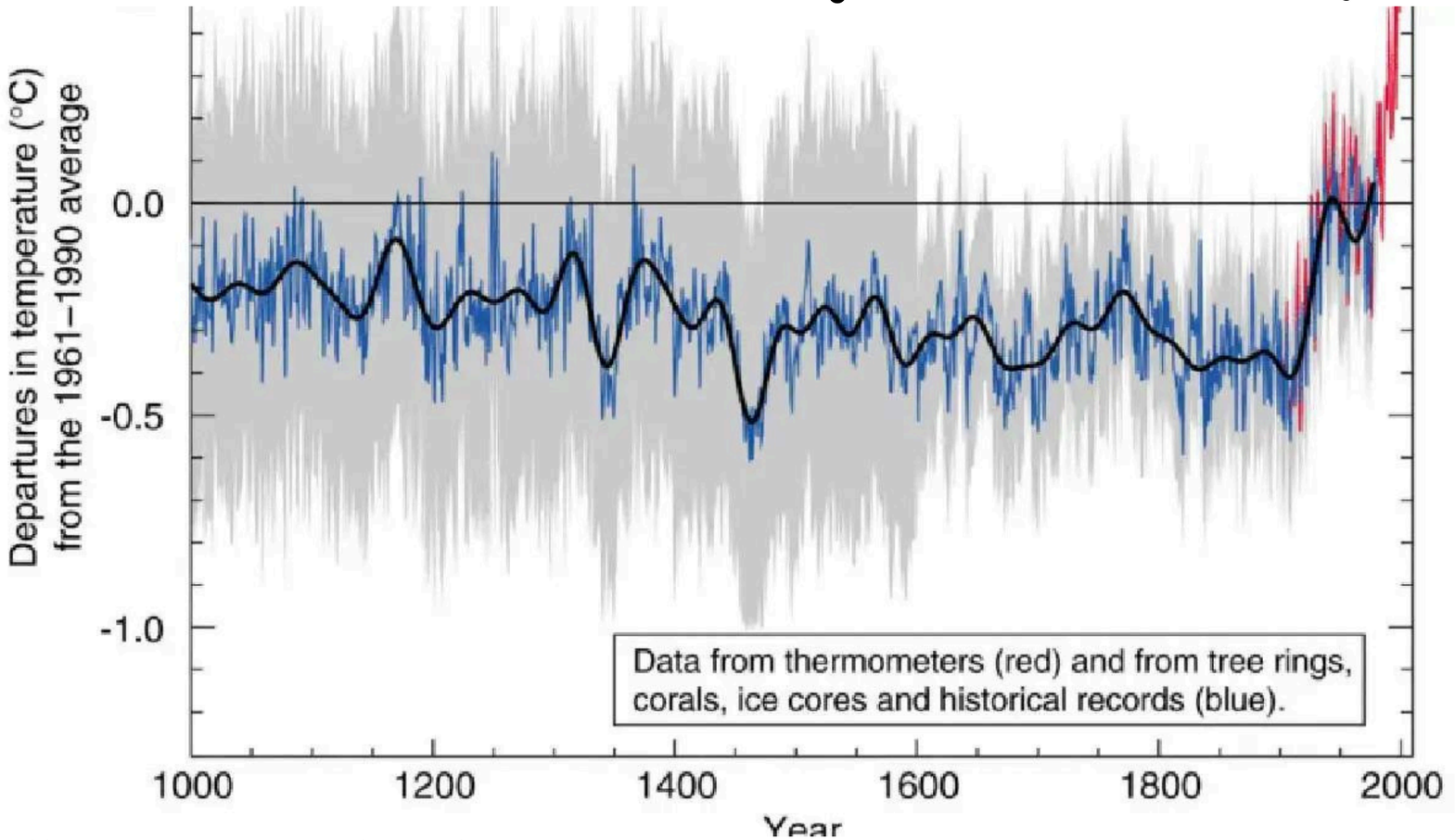
John Snow map that saved countless lives (1854)





Snow, J. (1855). On the Mode of Communication of Cholera. John Churchill.

“Hokey stick chart” that may have saved our planet (2001)



https://www.newscientist.com/article/2167127-why-the-hockey-stick-graph-will-always-be-climate-sciences-icon/

What will happen tomorrow?

https://www.cyberjournalist.net/2024/03/31/data-journalism-tells-the-story-through-numbers-with-unprecedented-clarity/

https://www.axios.com/2023/07/10/ai-misinformation-response-measures

Data journalists and critical-thinking, data-literate citizens

VS

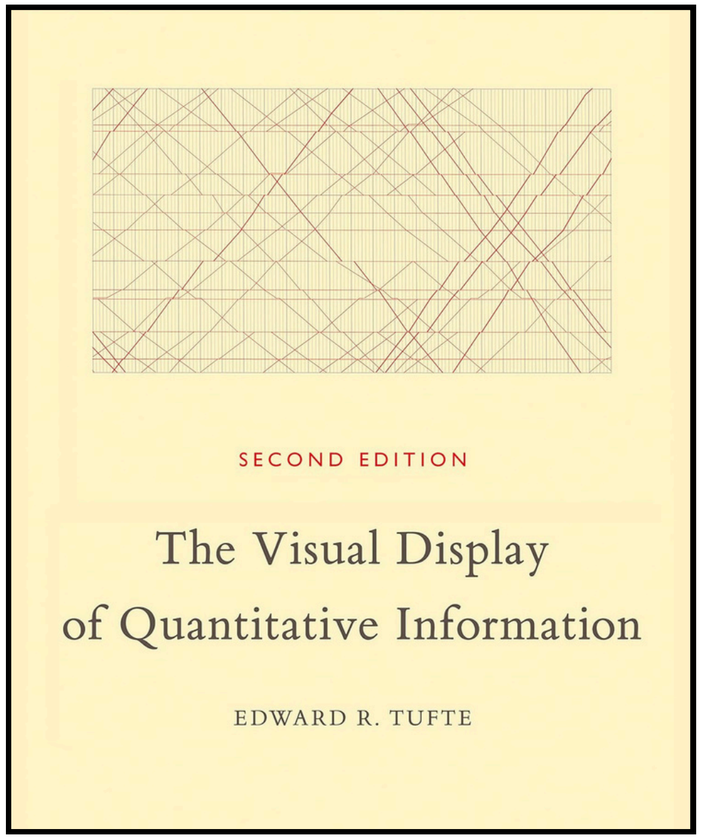
Trolls, liars and AI misinformations

“Look after truth and goodness,
and beauty will look after herself.”

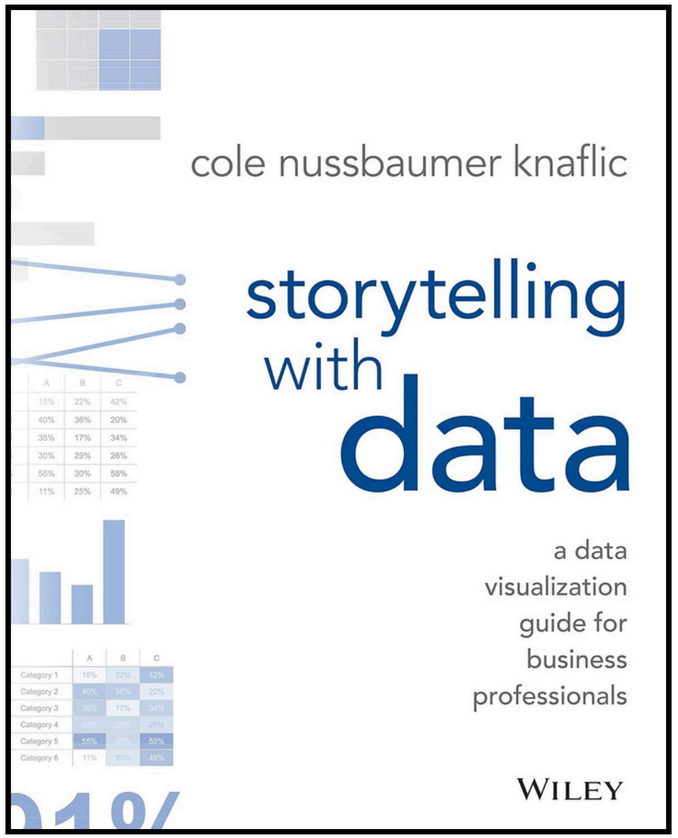
“You want to see to learn something, not to confirm something. We usually see to confirm things. It’s very economizing for the brain. How can we see not to confirm, but to see to learn?”

Tufte, E. R. (2013). The Art of Data Visualization | Off Book | PBS Digital Studios.
URL: https://www.youtube.com/watch?v=AdSZJzb-aX8 (7:02-7:20)

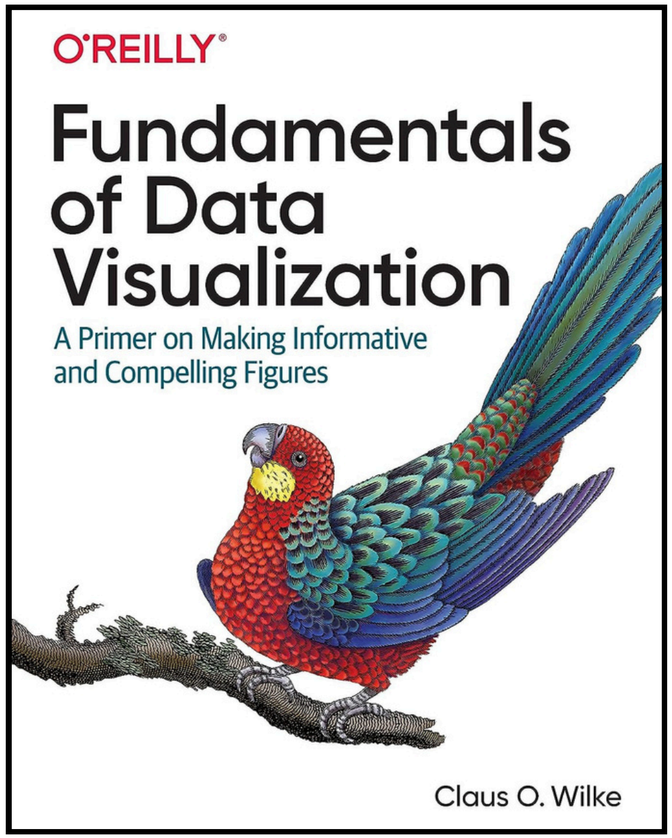
Further reading



Tufte, E. R. (2013). The Visual Display of Quantitative Information (2nd ed., 8th print). Graphics Press.



Nussbaumer Knaflic, C. (2015). storytelling with data: a data visualization guide for business professionals. Wiley.



Wilke, C. O. (2019). Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures (First edition). O'Reilly.

Thank you for your attention!

Full presentation with bibliography available on github:

https://github.com/stawarzkrzysztof/spec_ang