

Effective Data Visualization for Scientific Researchers

VLAG PhD week
Baarlo, October 17th, 2024

Irene van den Broek, PhD

 @IrenevdBroek

 @JeBentWatJeMeet

 @IreneVDB

Hi there, I'm Irene!



Irene van den Broek

irenevandenbroek@gmail.com

I have a PhD in biomarker research

Liquid chromatography coupled to tandem mass spectrometry for the quantitative bioanalysis of bioactive and potential biomarker peptides



Irene van den Broek 2010



2005-2010
Utrecht University

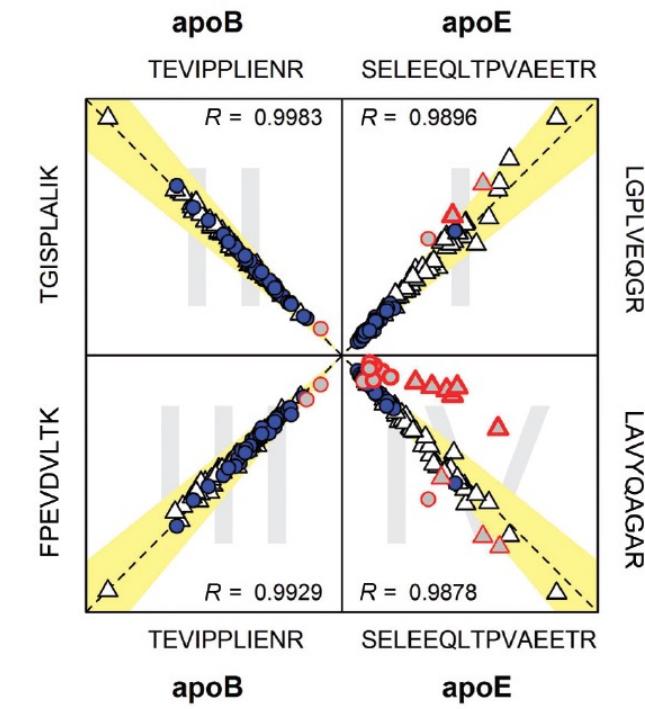
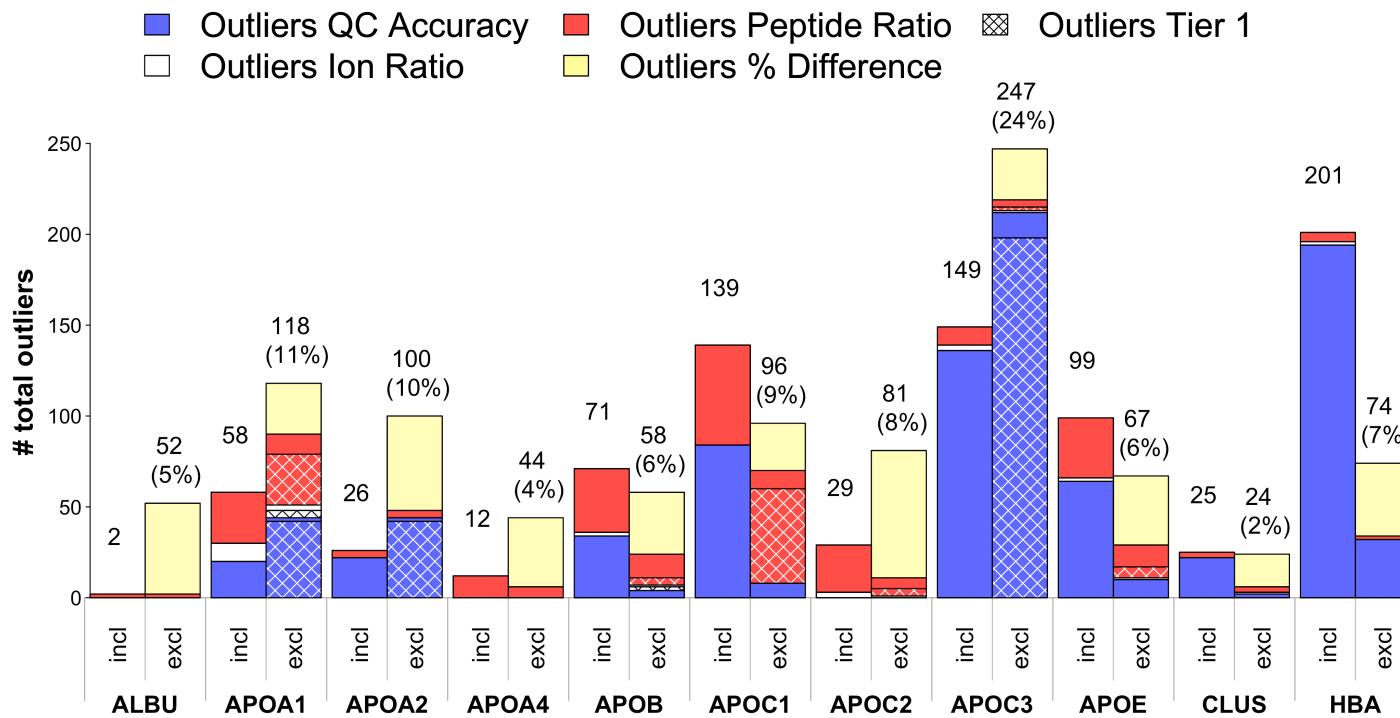


2013-2015
Leiden University
Medical Center

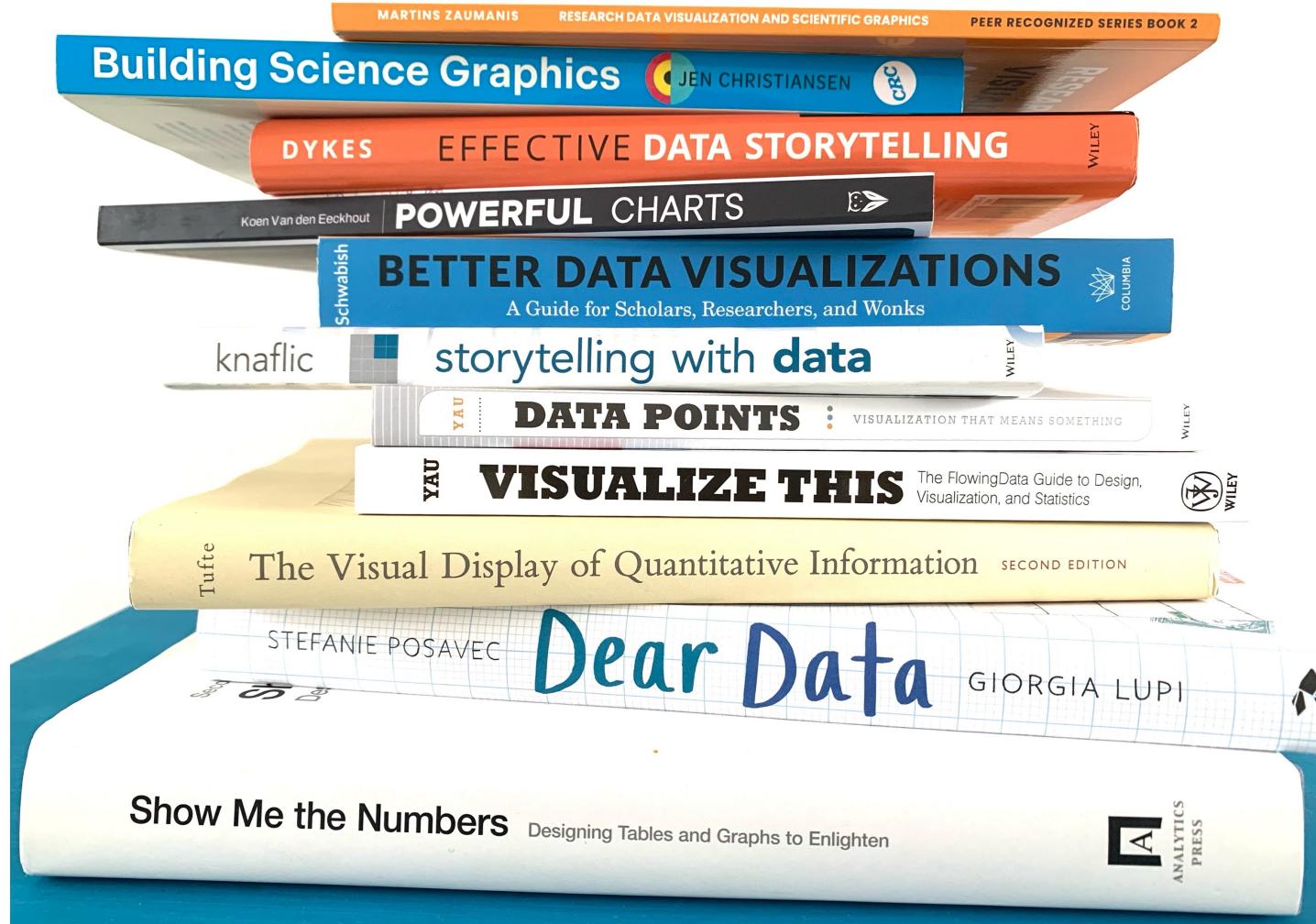


2015-2019
Cedars-Sinai
Medical Center

During my PhD and post-docs, I loved creating charts, preferably as complicated as possible...

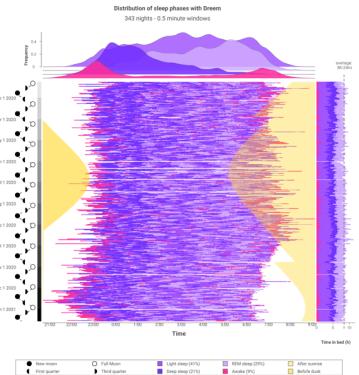


When I learned **Data Visualization** is a field on its own, I was hooked!

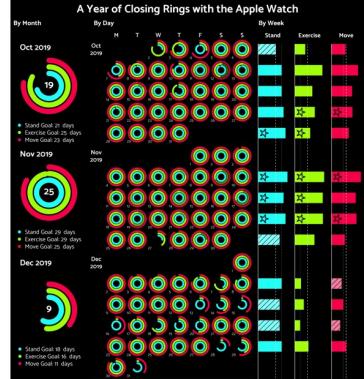


I started visualizing my Health and Habits.

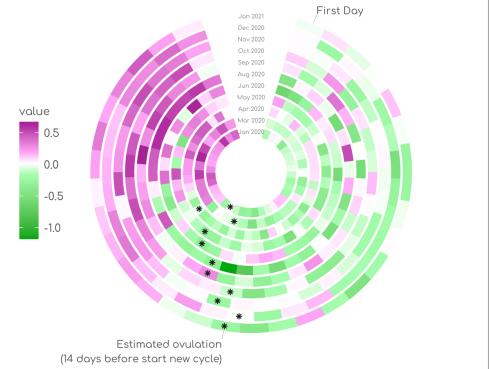
Sleep



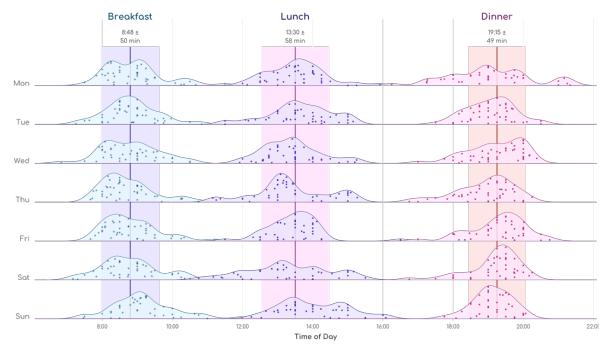
Activity



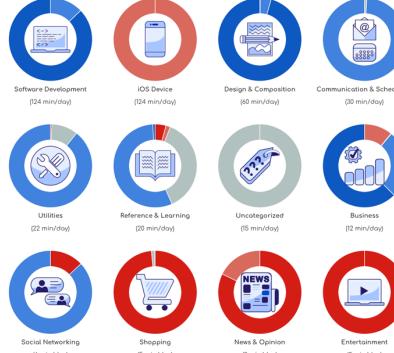
Body



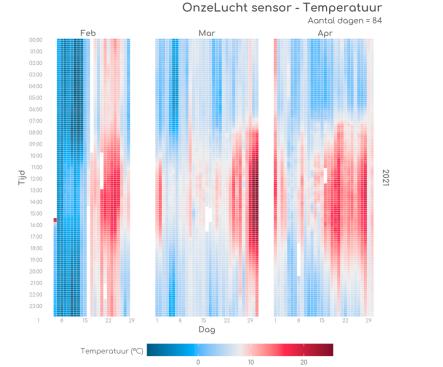
Food



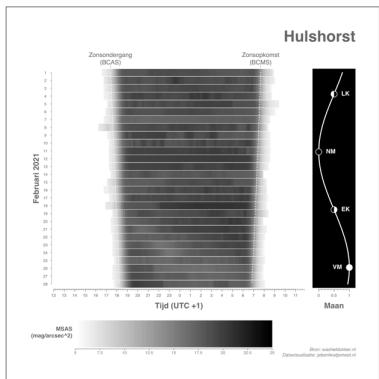
Habits



Environment



I now combine my passions for teaching and data visualization



www.washetdonker.nl



www.onzelucht.nl

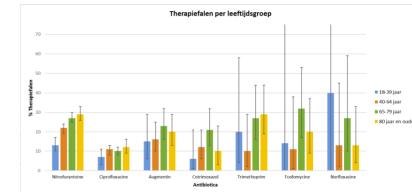


Tabel 1. AUC analyses van het basis- en final model, in vrouwen en mannen in totaal, en over etnische groepen.

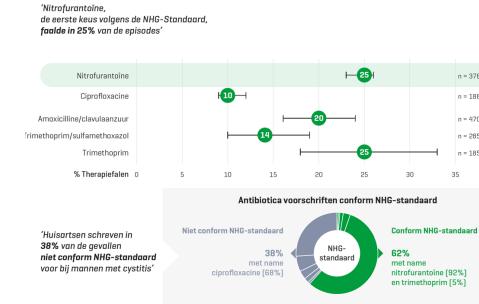
	Vrouwen (n=5852)		Mannen (n=4392)	
	AUC [95%CI]	AUC [95%CI]	AUC [95%CI]	AUC [95%CI]
Totalle populatie	.563 [.539, .584]	.682 [.664, .699]	.610 [.595, .627]	.664 [.644, .681]
Nederlands	.604 [.558, .640]	.807 [.772, .838]	.609 [.582, .638]	.721 [.691, .757]
Hindoestaans-Surinaams	.581 [.541, .623]	.756 [.711, .793]	.613 [.568, .646]	.667 [.627, .708]
Afro-Surinaams	.591 [.554, .633]	.701 [.669, .731]	.642 [.620, .665]	.666 [.633, .693]
Ghanees	.598 [.528, .642]	.678 [.631, .733]	.574 [.539, .613]	.617 [.580, .659]
Turks	.678 [.631, .733]	.707 [.650, .744]	.617 [.580, .659]	.680 [.644, .741]
Marokkaans	.609 [.568, .661]	.809 [.773, .836]	.631 [.584, .672]	.727 [.702, .754]

*Basis model: obesitas, roken, en een familiegeschiedenis van HVZ. **Final model: obesitas, roken, en een familiegeschiedenis van HVZ, opleidingsniveau en arbeidsstatus.

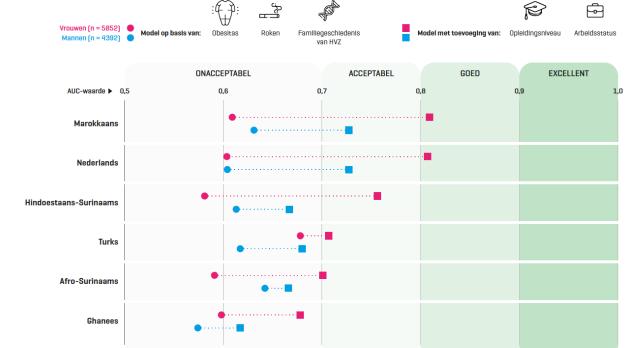
Figuur 1 Therapiefalen voor verschillende antibiotische regimes per leeftijdsgroep



Voorschriften en faalpercentages van verschillende antibiotica bij mannen met cystitis



Bij zowel vrouwen als mannen droegen opleidingsniveau en arbeidsstatus bij aan de predictie van een verhoogd HVZ-risico
Toegevoegde voorspelende waarde [AUC] lijkt groter bij vrouwen dan bij mannen en varieerde tussen etnische groepen



Huisarts en Wetenschap

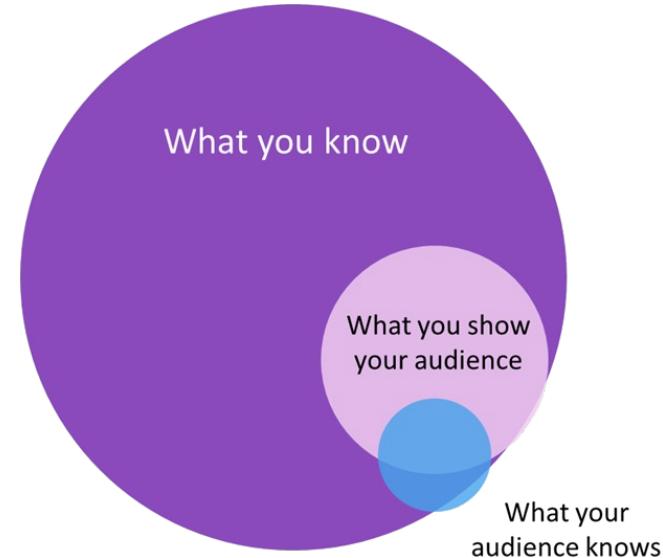
I'm excited to share with you what I wish I'd learned during my PhD!

Resist the urge to show everything!



Don't underestimate how smart you are!

The curse of knowledge

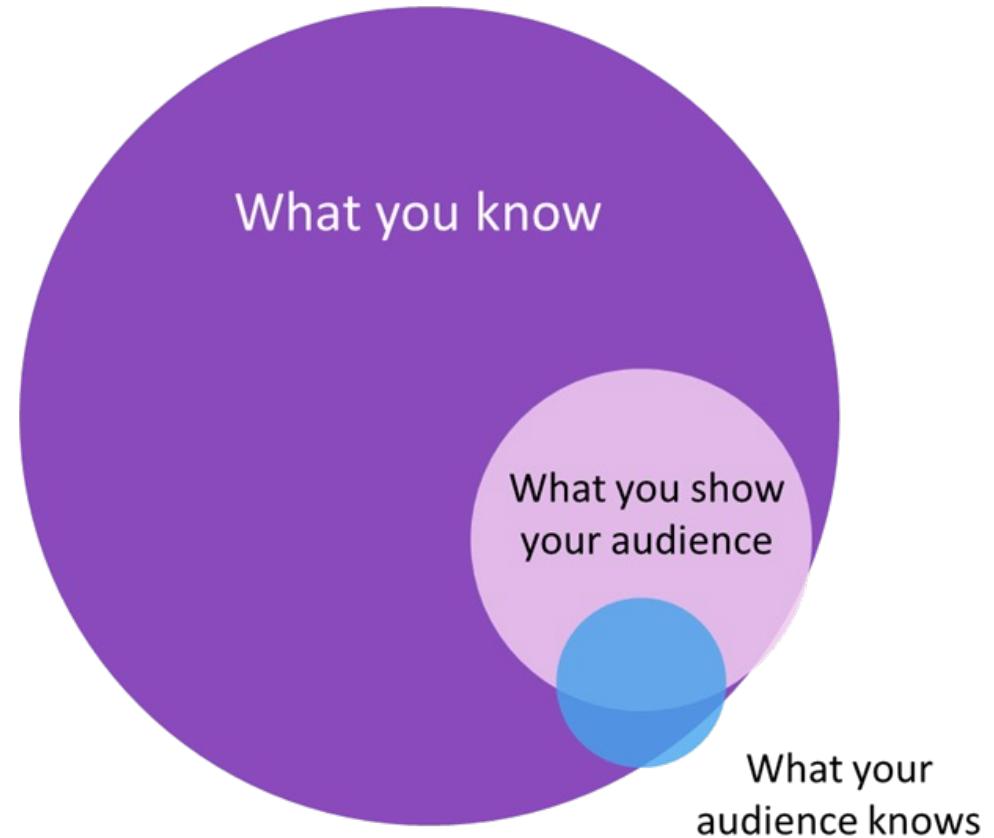


Resist the urge to show everything!

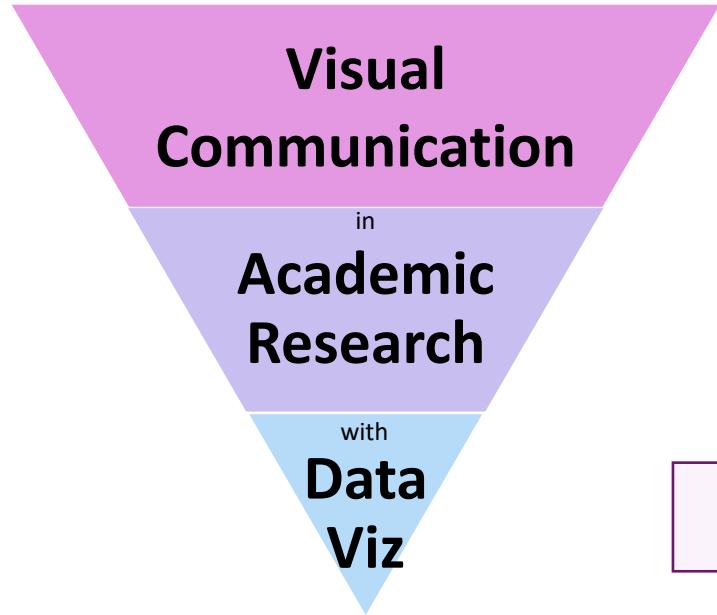


Don't underestimate how smart you are!

The curse of knowledge



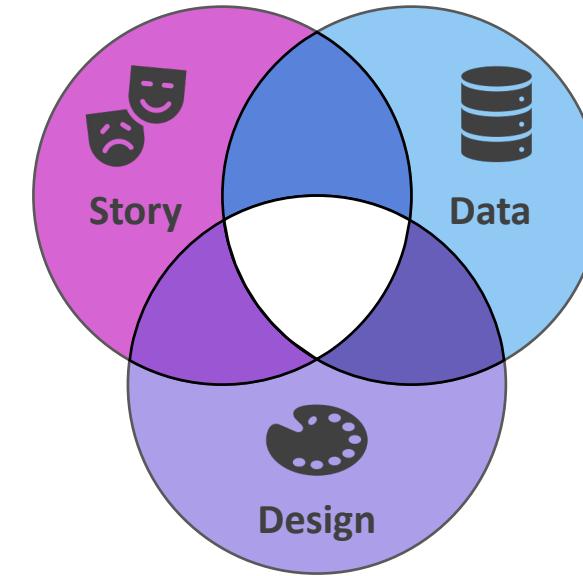
An overview for today:



WHY?

WHAT?

HOW?





Your turn!

A look at your work.

Nice to meet **YOU**!

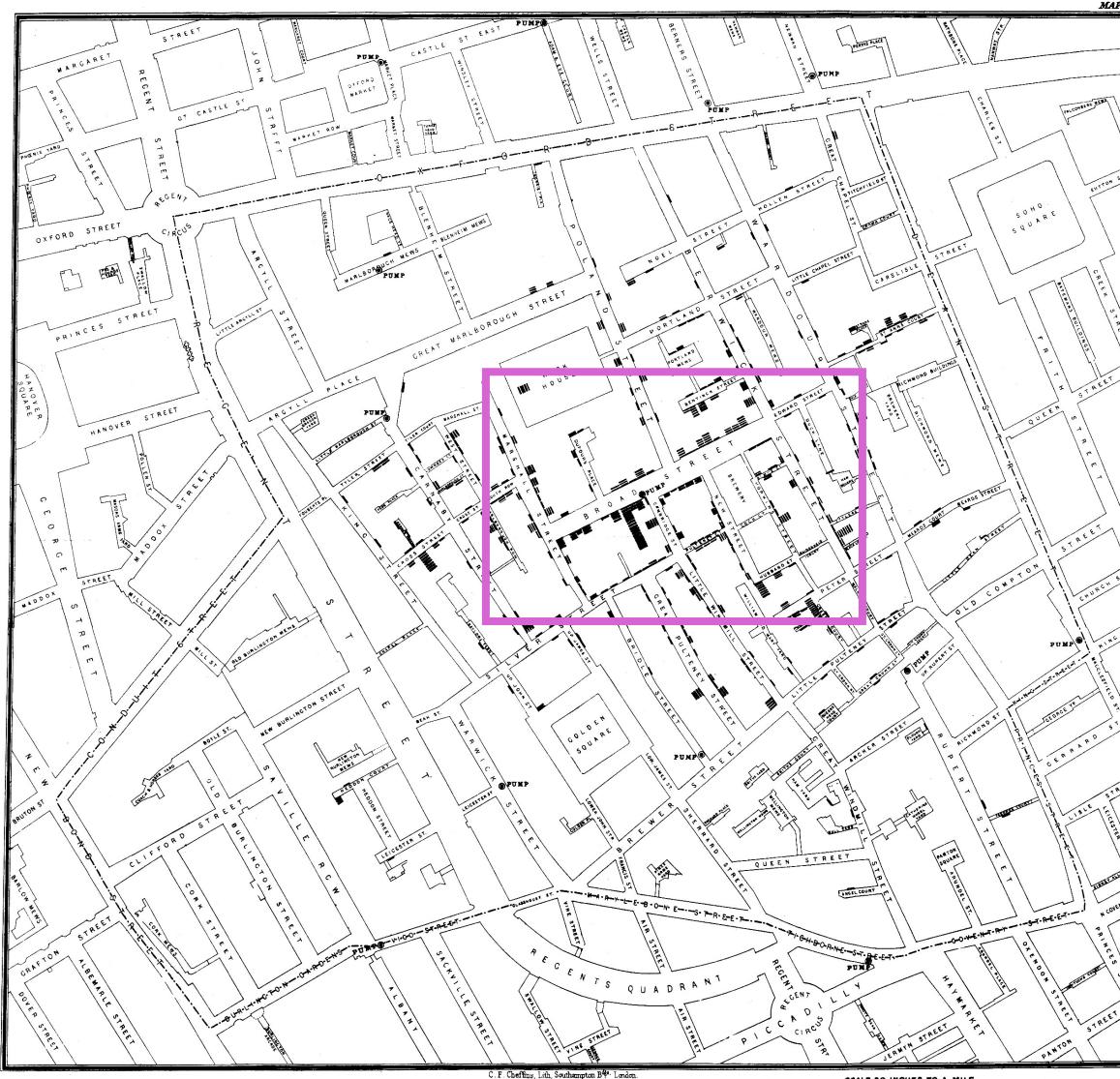


Explain in 1 sentence:

- The **topic** of your PhD as explained to your **neighbor, friend, or family member**
- Your selected **visual**: (if applicable)
 - One thing you **like** or **dislike** about it

Why visual communication?

1854: The cholera map from John Snow



1854: The cholera map from John Snow



What can you see in these numbers?

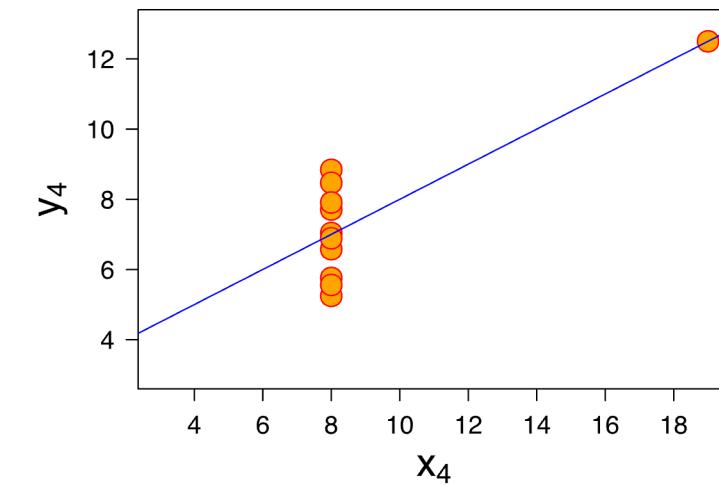
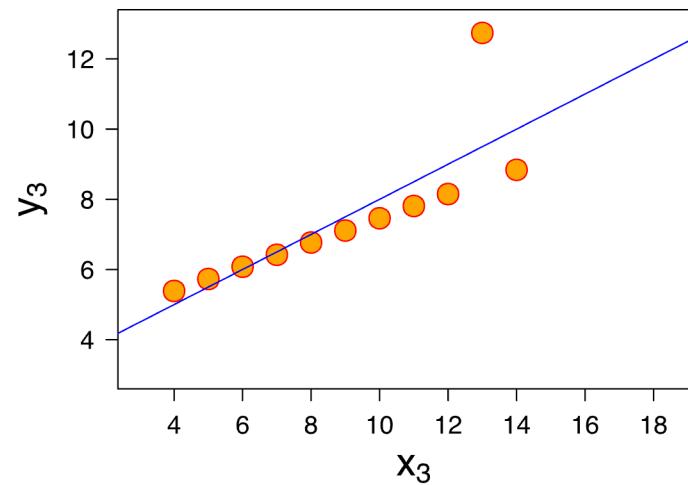
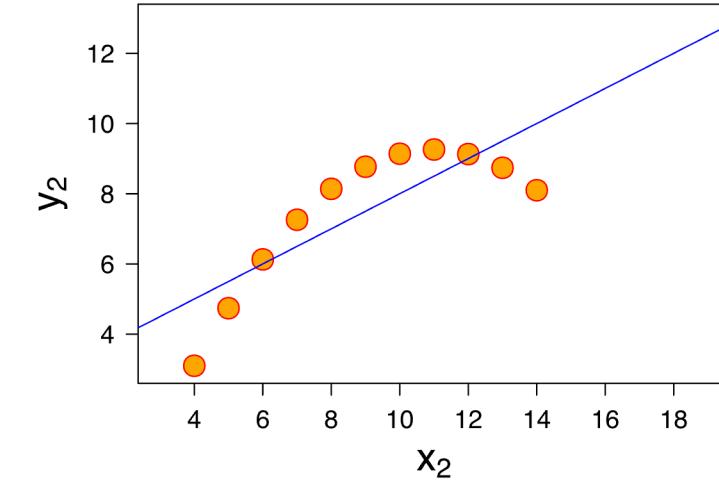
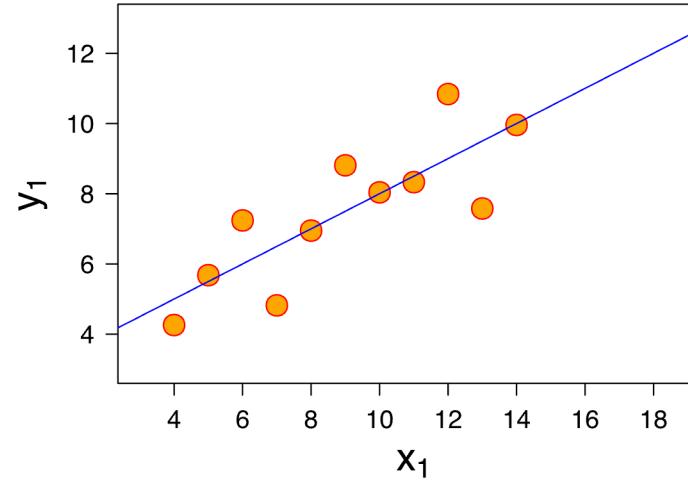
Anscombe's quartet:
4 datasets
11 x-values
11 y-values
Same total
Same mean
Same standard deviation

	I	II	III	IV				
	x	y	x	y	x	y	x	y
	10	8,04	10	9,14	10	7,46	8	6,58
	8	6,95	8	8,14	8	6,77	8	5,76
	13	7,58	13	8,74	13	12,74	8	7,71
	9	8,81	9	8,77	9	7,11	8	8,84
	11	8,33	11	9,26	11	7,81	8	8,47
	14	9,96	14	8,1	14	8,84	8	7,04
	6	7,24	6	6,13	6	6,08	8	5,25
	4	4,26	4	3,1	4	5,39	19	12,5
	12	10,84	12	9,13	12	8,15	8	5,56
	7	4,82	7	7,26	7	6,42	8	7,91
	5	5,68	5	4,74	5	5,73	8	6,89
SUM	99,00	82,51	99,00	82,51	99,00	82,50	99,00	82,51
AVG	9,00	7,50	9,00	7,50	9,00	7,50	9,00	7,50
STDEV	3,32	2,03	3,32	2,03	3,32	2,03	3,32	2,03

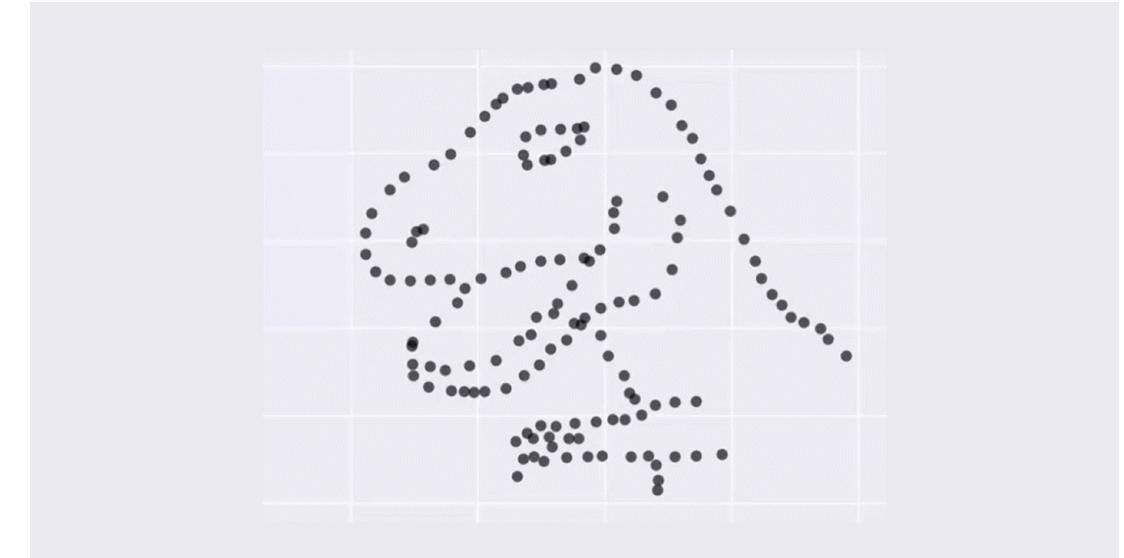
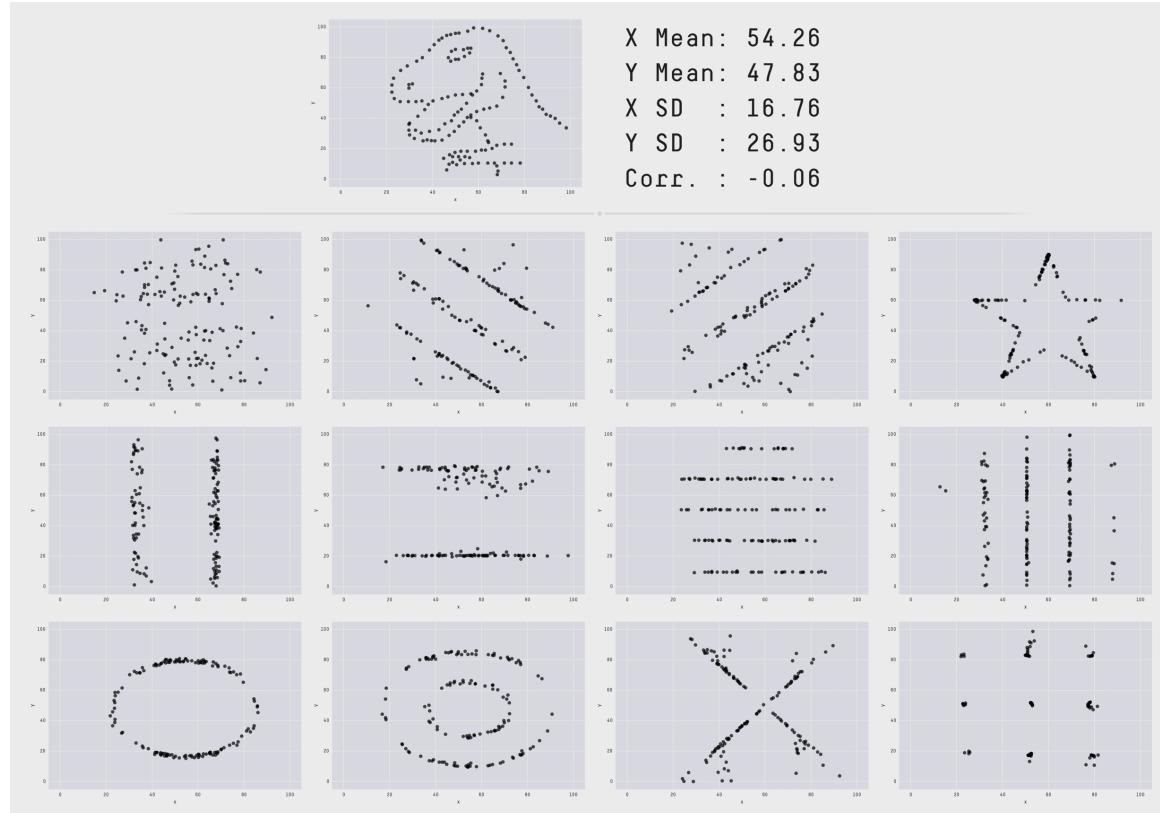
A picture is worth a thousand words!

Anscombe's quartet:

4 datasets
11 x-values
11 y-values
Same total
Same mean
Same standard deviation

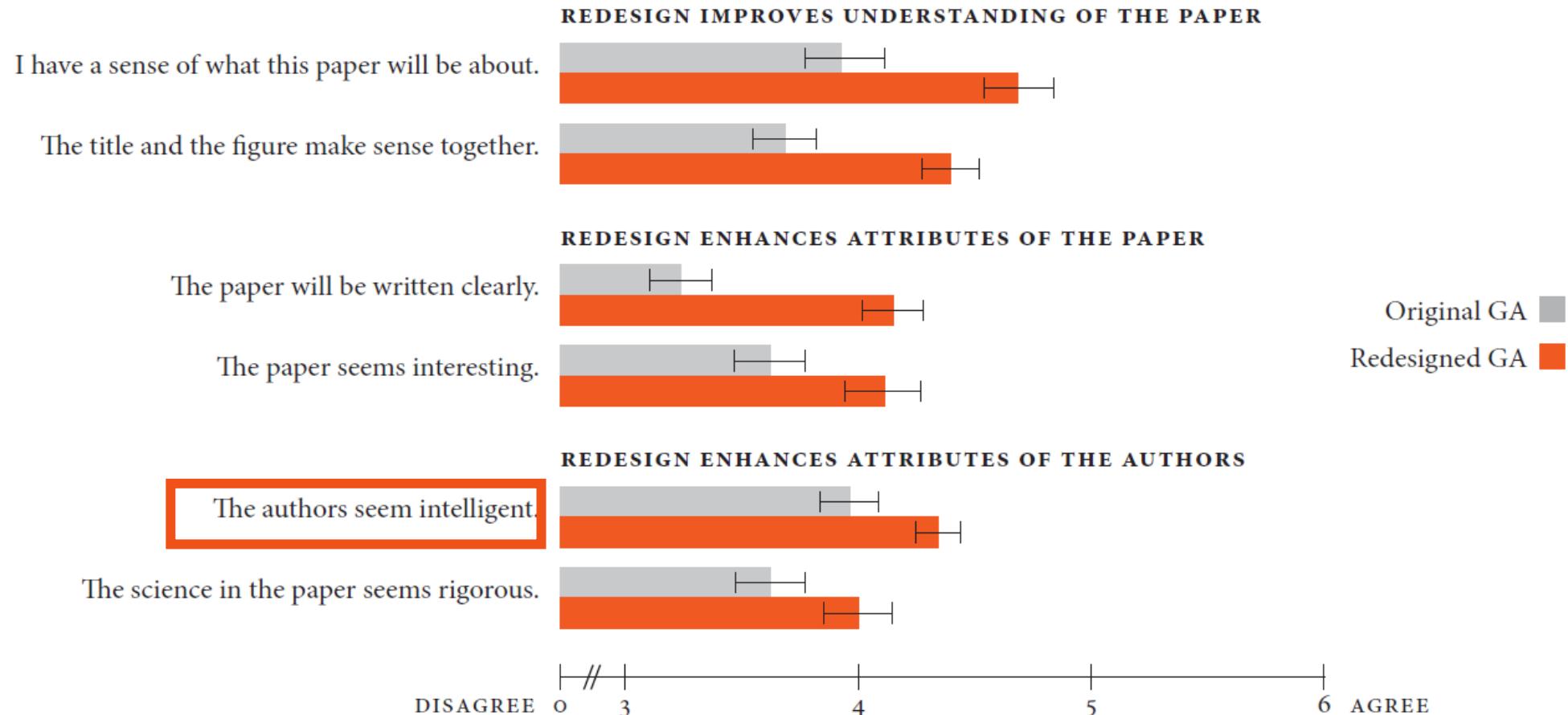


The Datasaurus Dozen 😊: Never trust summary statistics alone!



Matejka J & Fitzmaurice G (2017). *Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing*.
<https://www.autodeskresearch.com/publications/samestats>.

A good visual makes you (look) smarter!

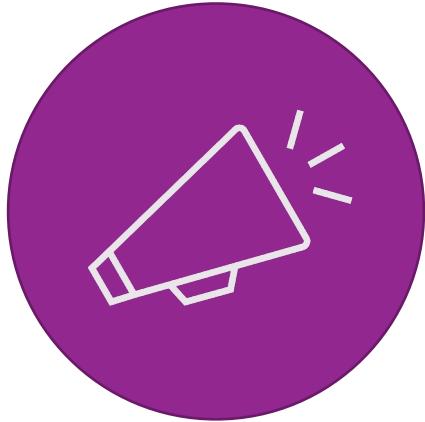


Why use visuals in academic research?



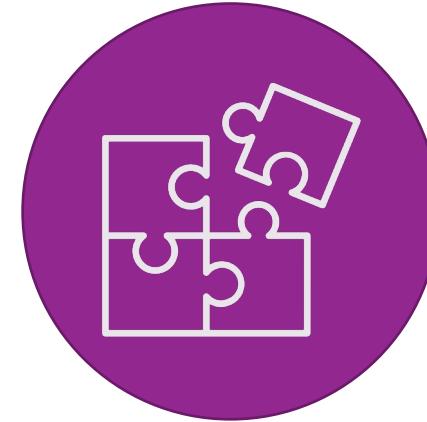
Provide **insight**

-
- Publications
 - Oral presentations
 - Poster presentations



Grab **attention**

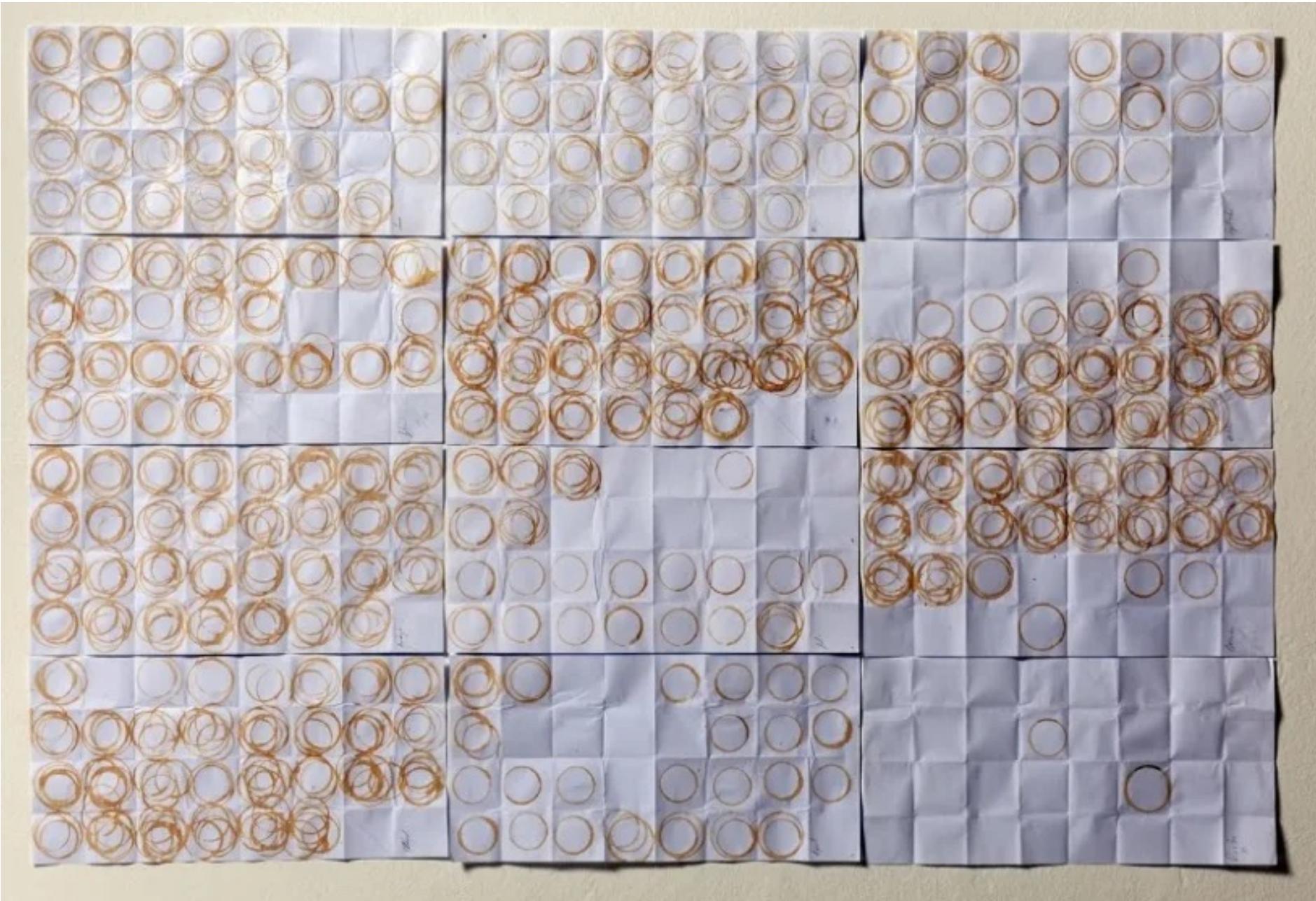
-
- Graphical abstracts
 - Social media
 - General media



Show **competence**

-
- Grant applications
 - Portfolio / CV
 - Pitch

What is visual communication ?



Jaime Serra Palou Café Diario

Talking about reaching a broad public!

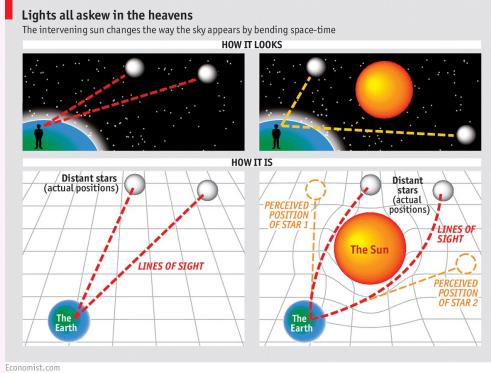
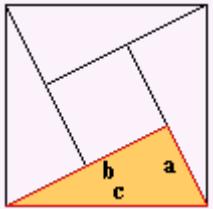


Prof. Ed Hawkins
#showyourstripes



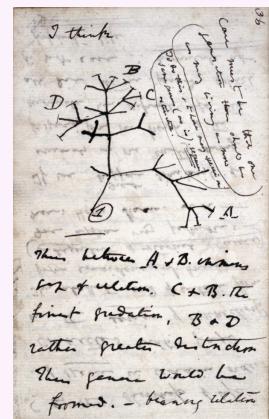
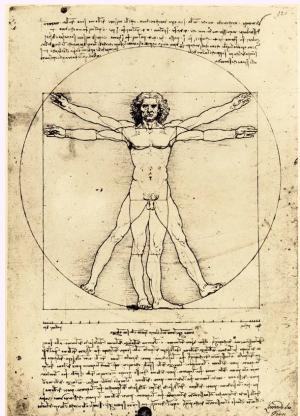
Four types of visualizations

Idea illustration



DECLARATIVE

Idea generation

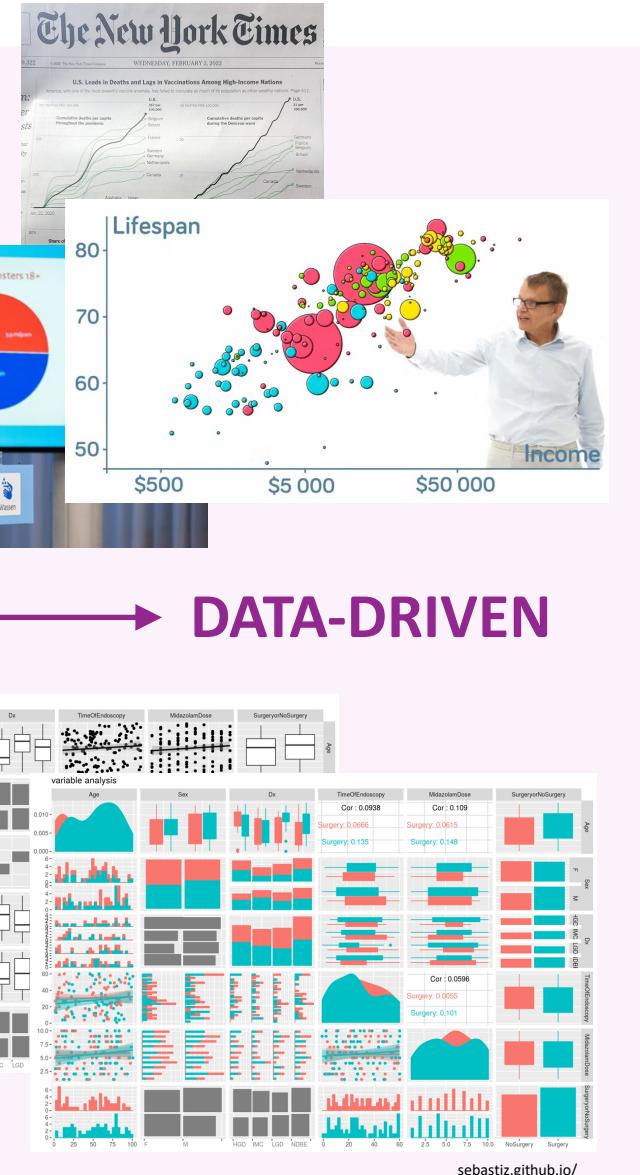


CONCEPTUAL



DATA-DRIVEN

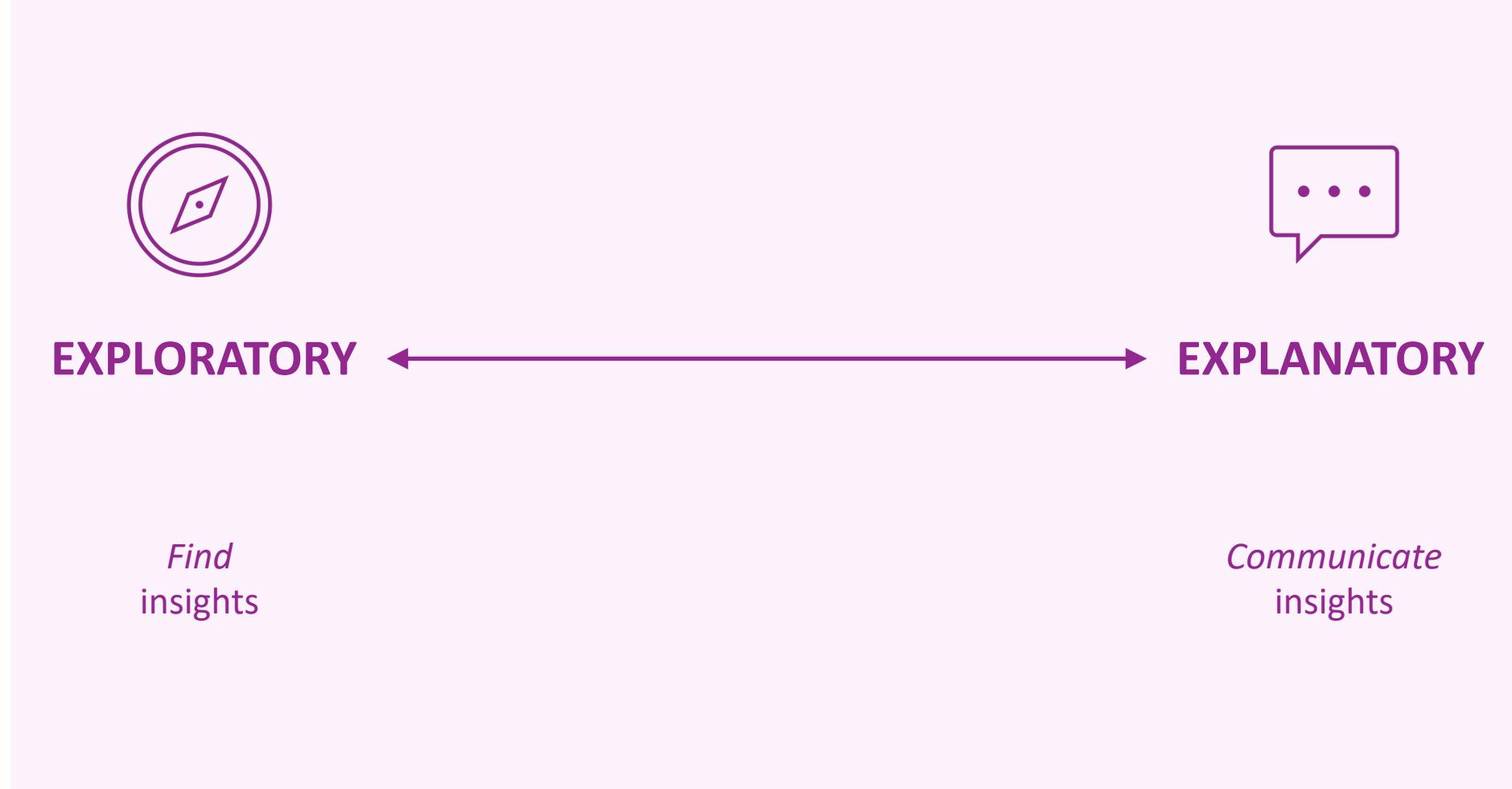
EXPLORATORY



'Everyday'
DataViz

Visual
discovery

Exploratory vs. explanatory data visualization



Exploratory vs. explanatory data visualization



EXPLORATORY

Understand

You



EXPLANATORY

Communicate

Other people

Less familiar (others)

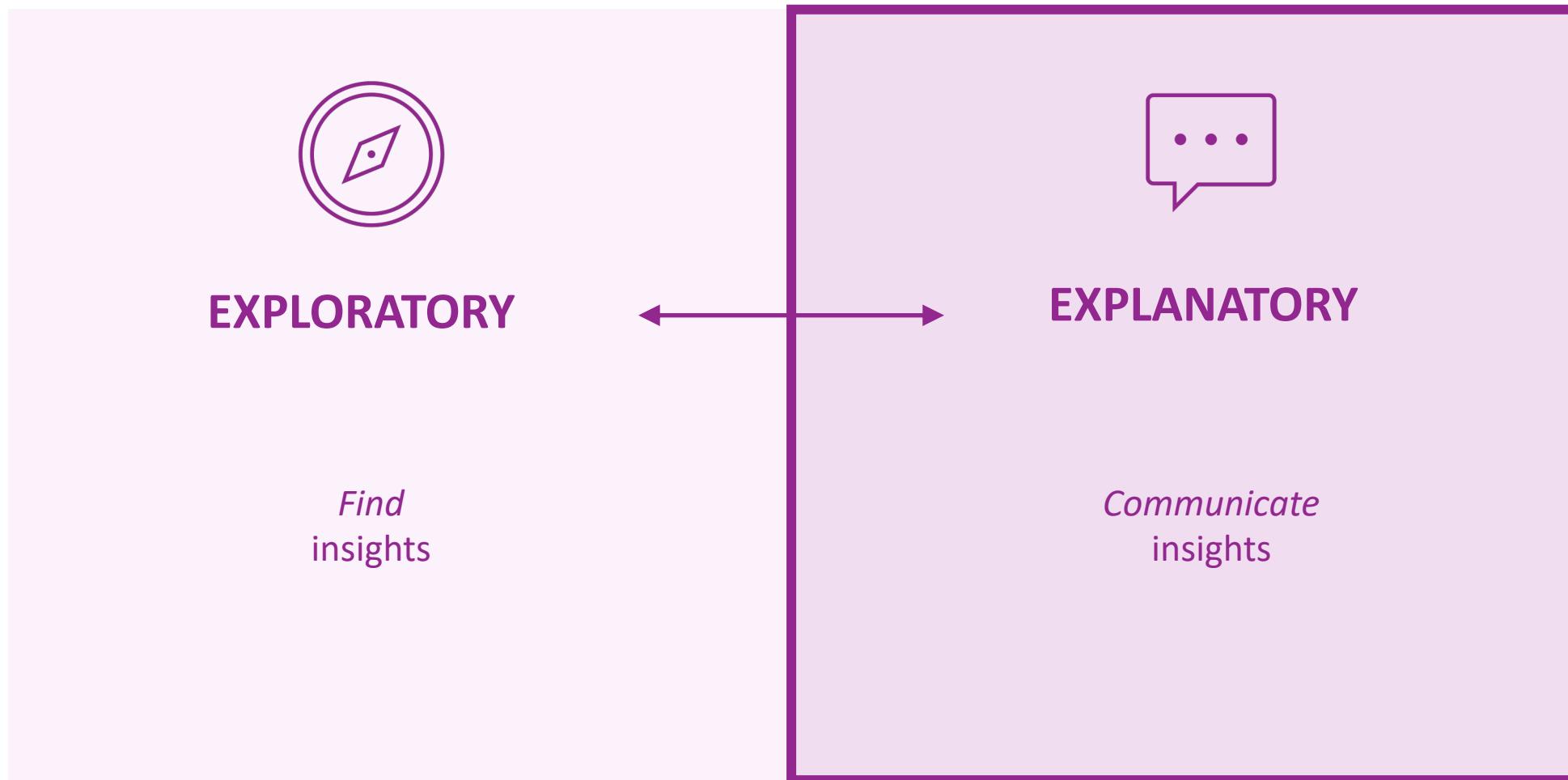
Simplicity, clarity, and cohesion

Known

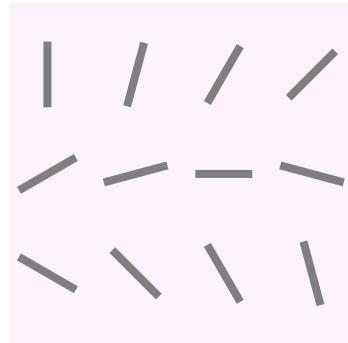
Action

Goal:	Understand
Audience:	You
Data familiarity:	Very familiar (you)
Visualization focus:	Flexibility and speed
Narrative:	Unknown
Outcome:	Insight

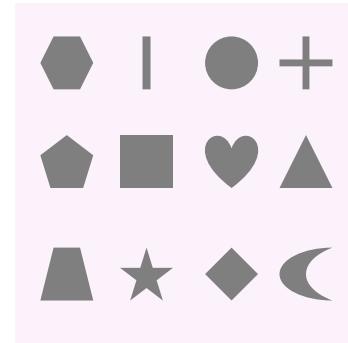
We will focus on **explanatory** data visualization



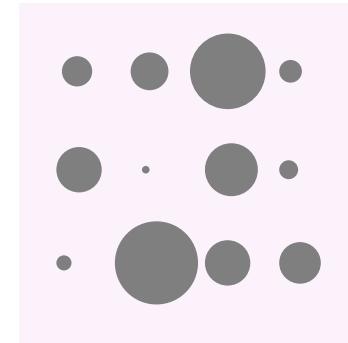
Visual attributes: from data to image



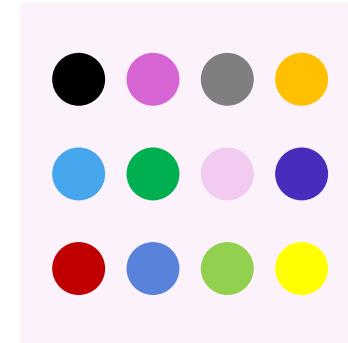
Orientation



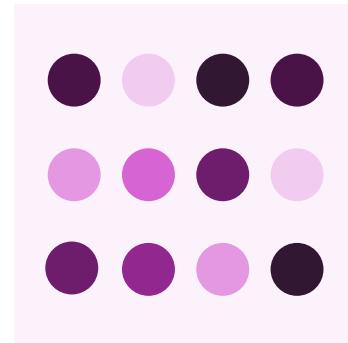
Shape



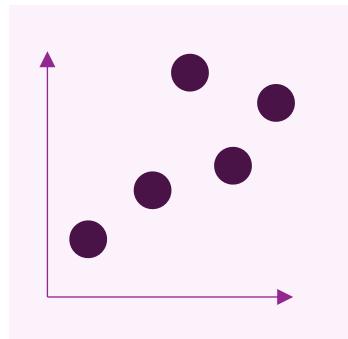
Size



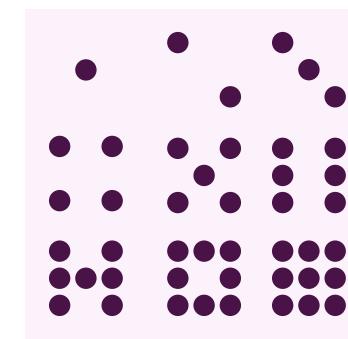
Color (hue)



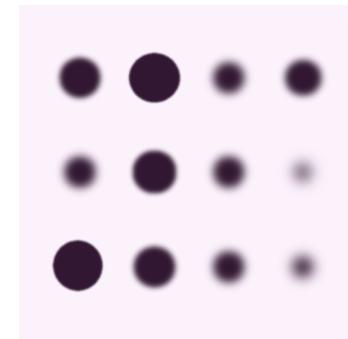
Color (saturation)



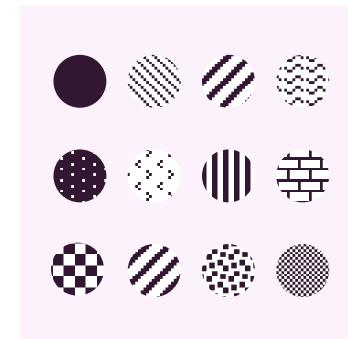
Position



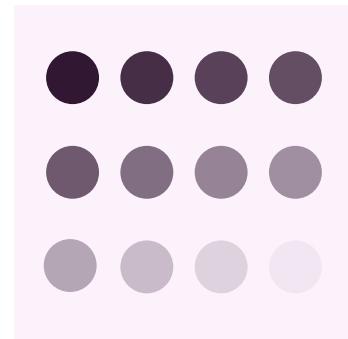
Number



Fuzziness

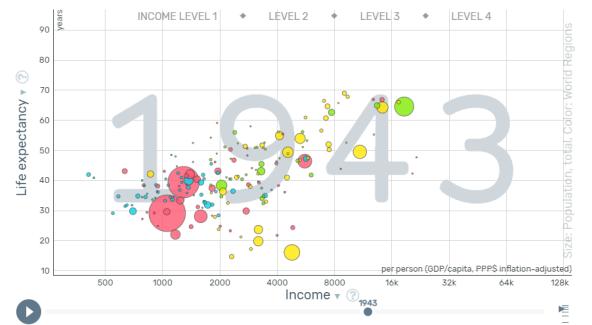
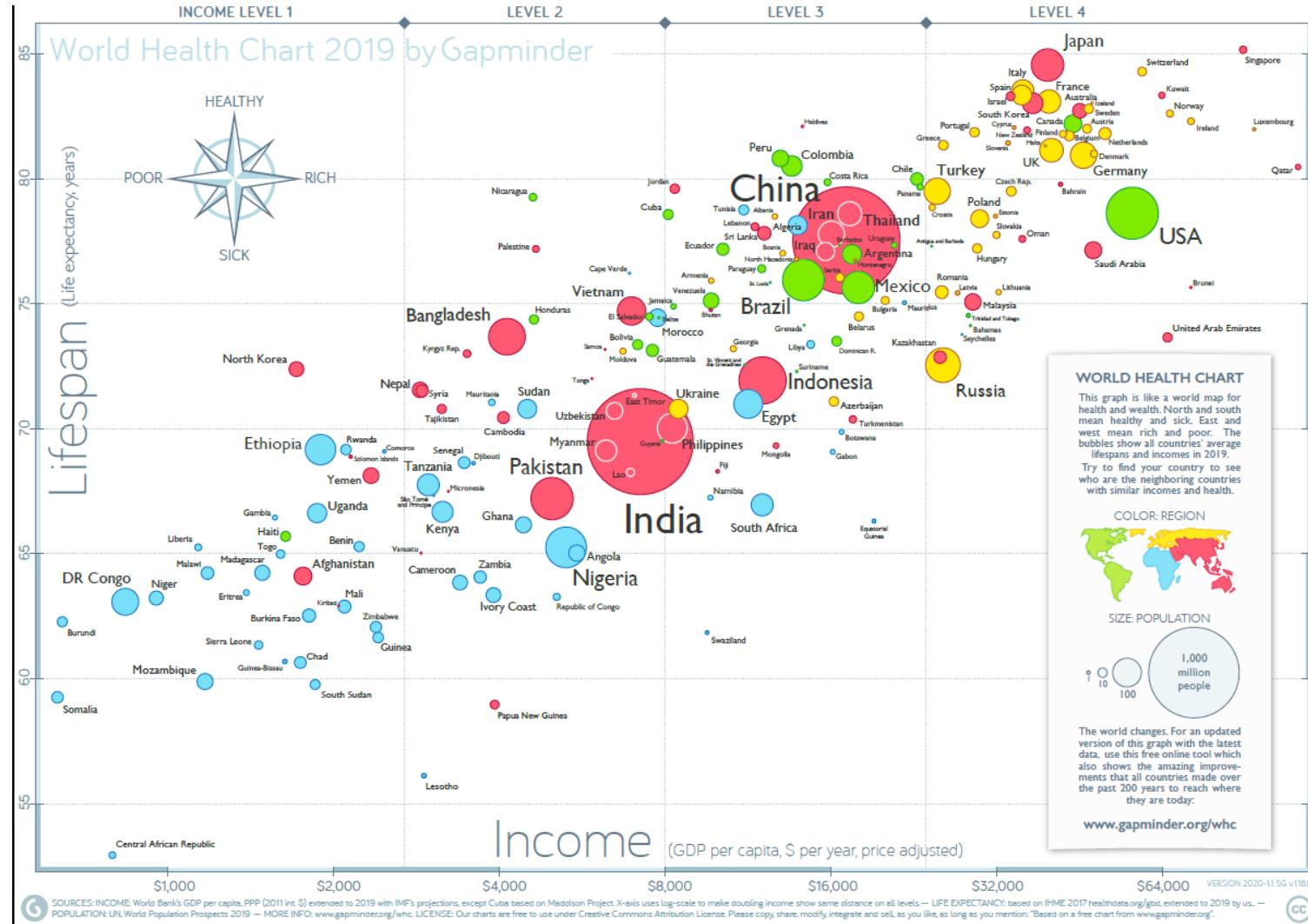


Texture



Transparency

Hans Rosling's Bubble chart

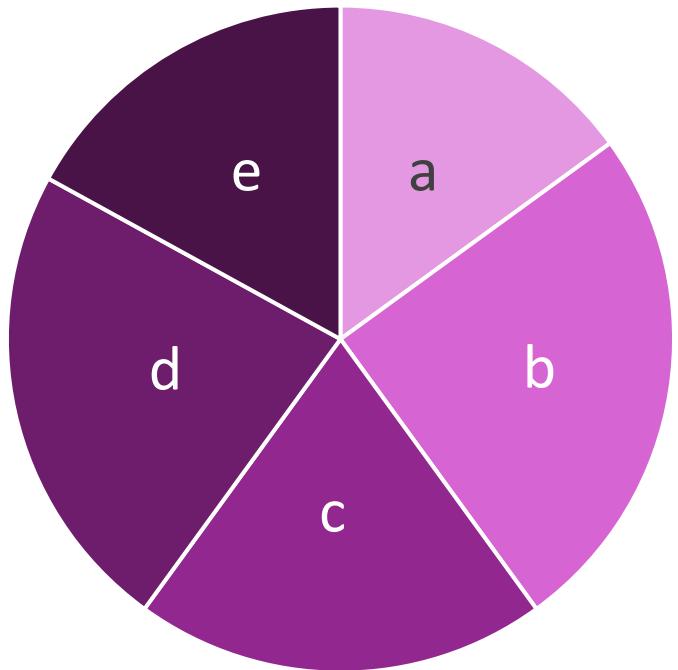


<https://www.gapminder.org/tools>

<https://www.youtube.com/watch?v=jbkSRLYSojo>

Some visual attributes are **more effective** than others

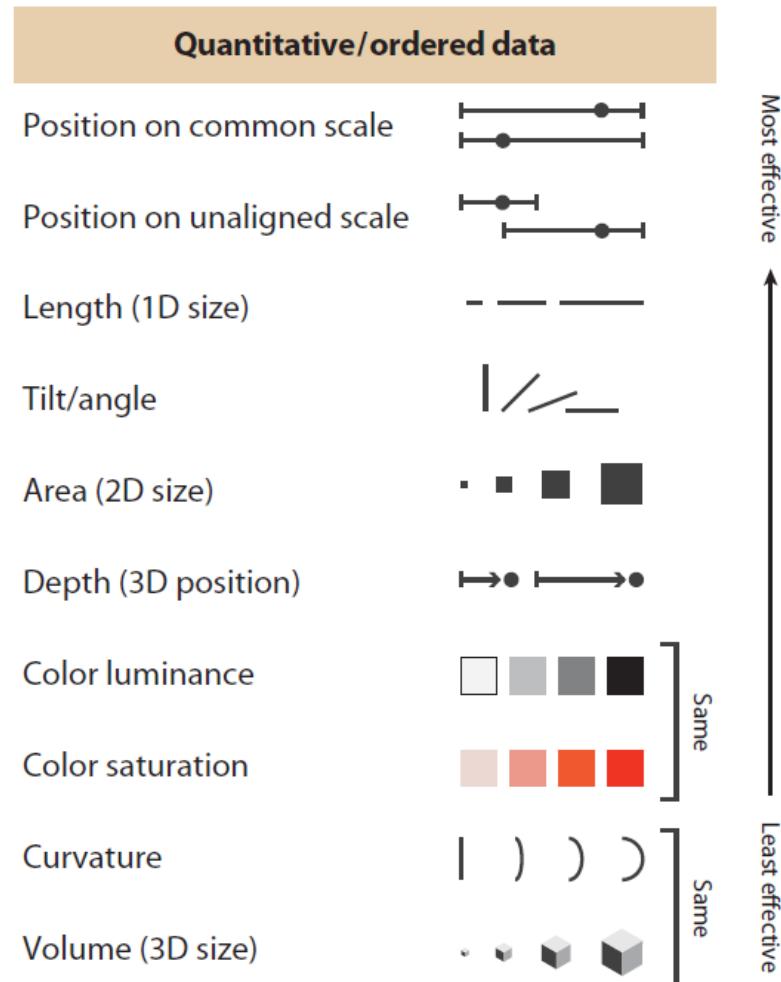
Which slice is biggest?



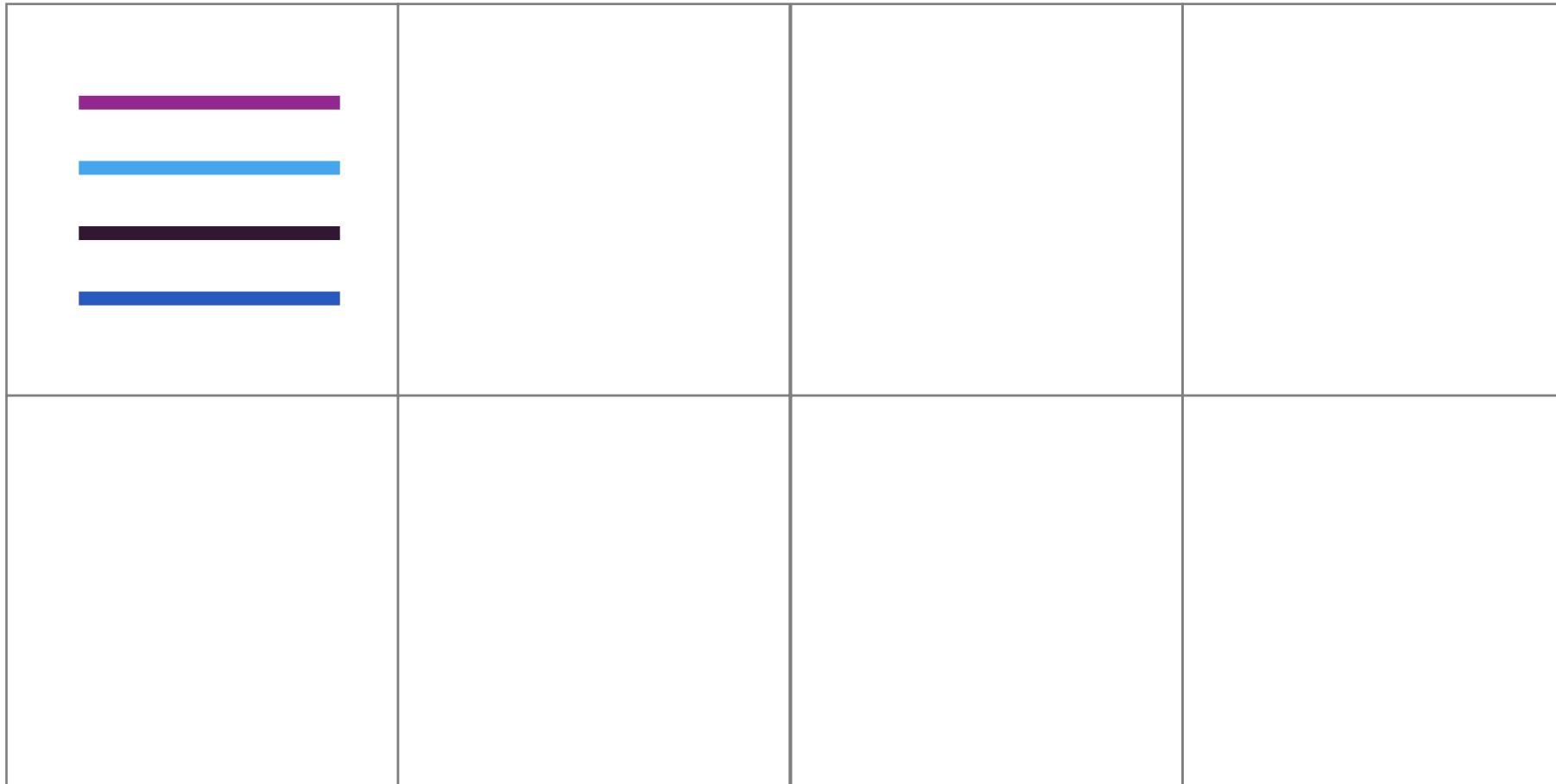
Which bar is highest?



Some visual attributes are **more effective** than others



Warm-up: draw 4 lines using 1 visual attribute at a time to differentiate them*

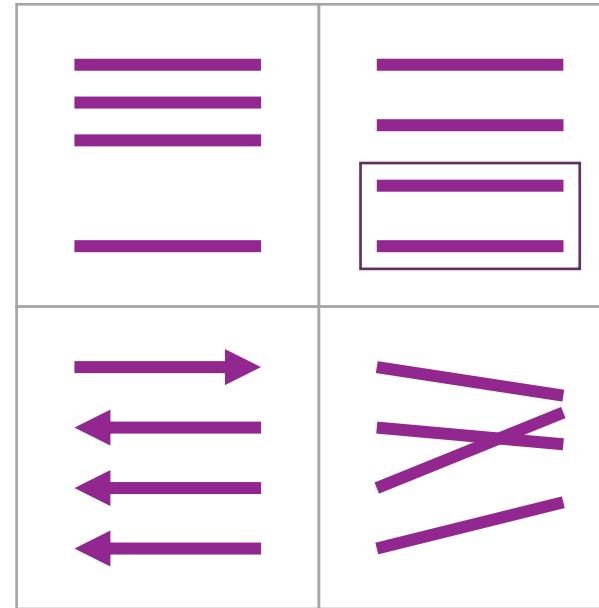
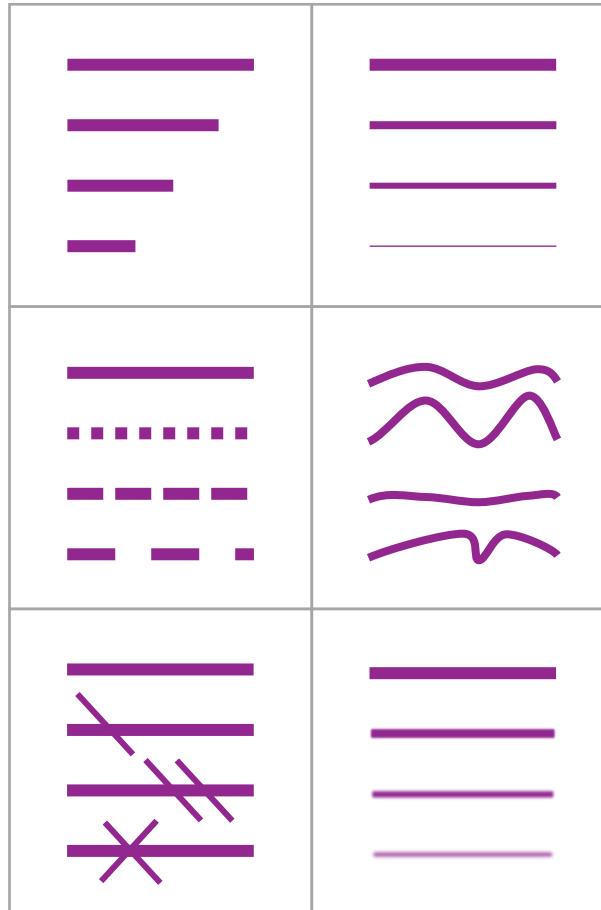


* From each other or at least one should be seen as different from the others



3 min

Exercise: draw 4 lines using 1 visual attribute at a time to differentiate them*



HOW visual communication ?

Show the *right* data and show the data *right*.

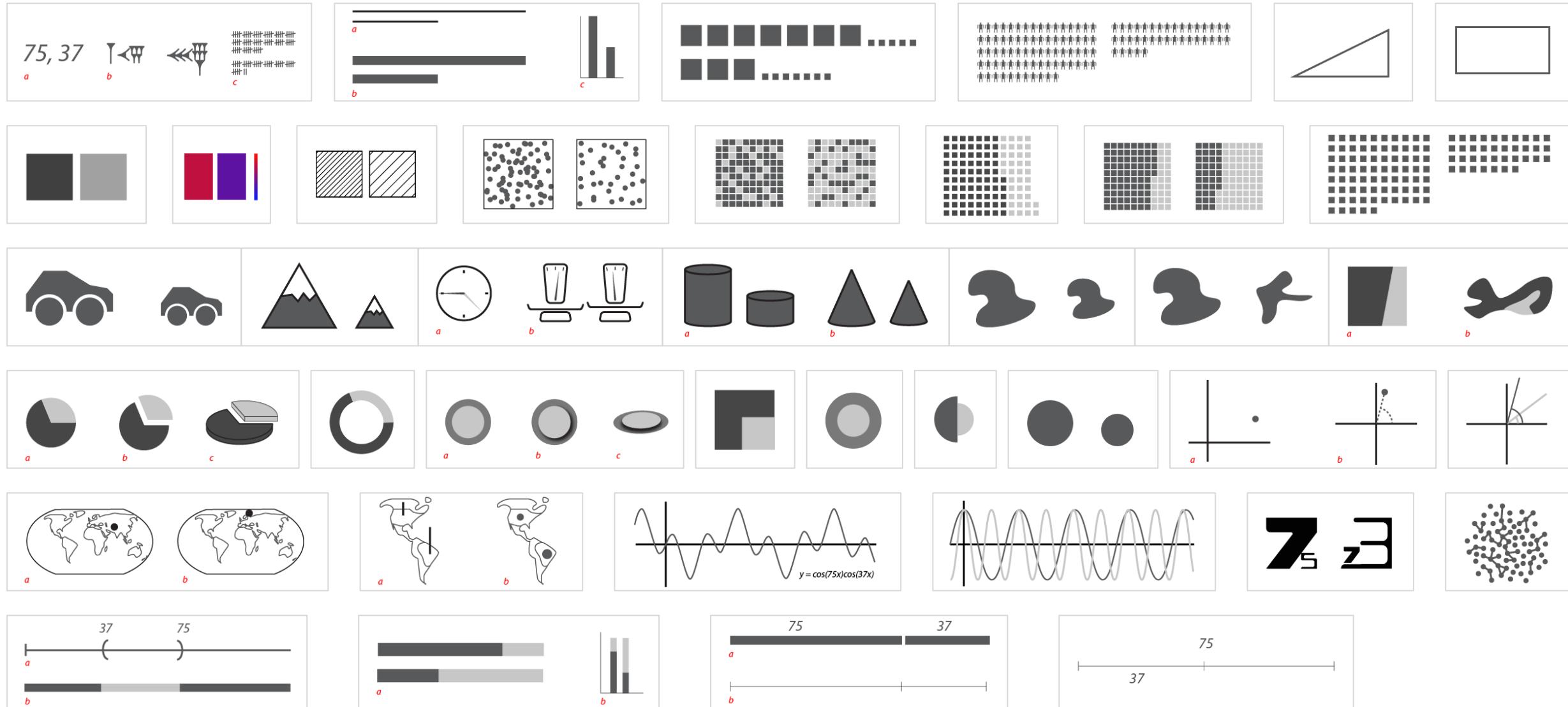
Exercise: In how many ways can you visualize two numbers?

75 and 37



5 min

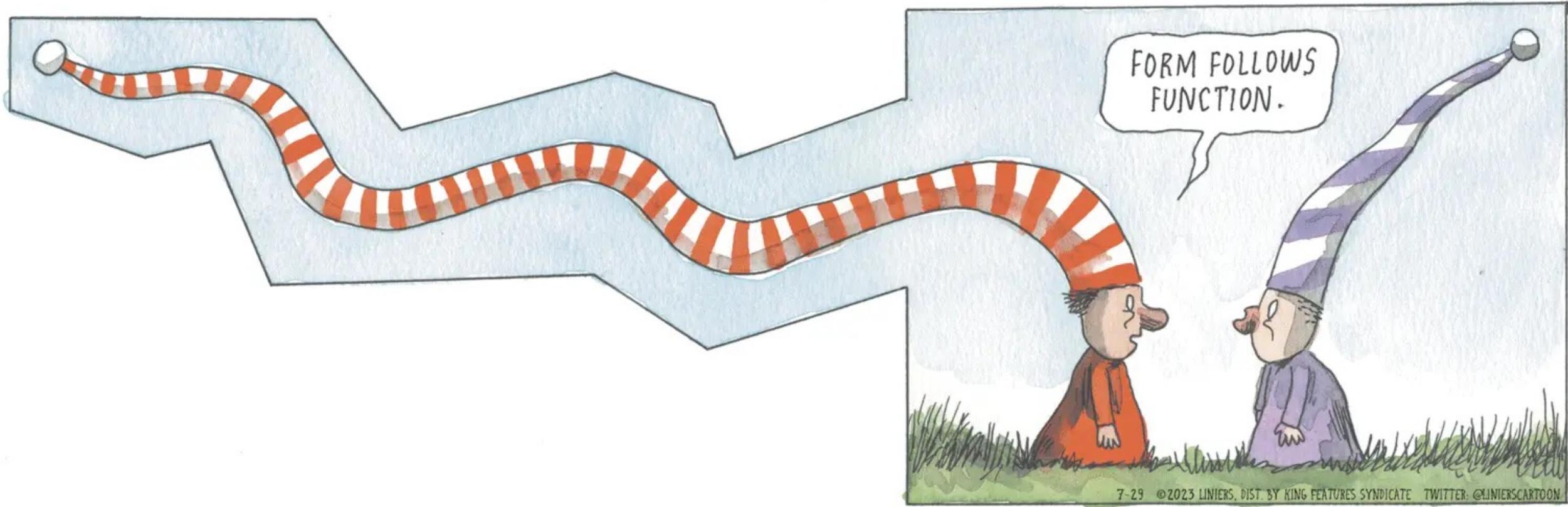
45 ways to communicate two quantities...



So many possibilities. How to choose?



Form follows function!

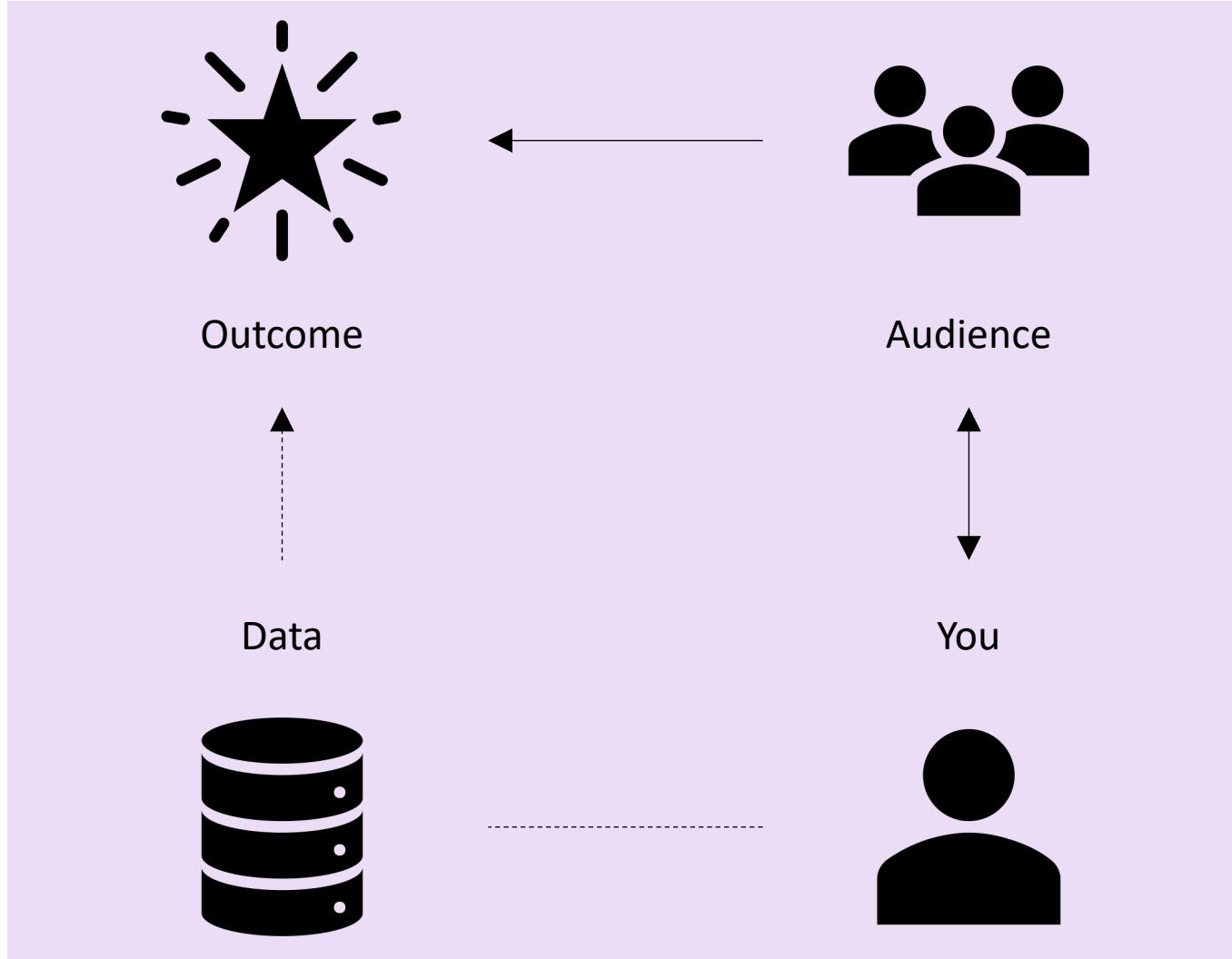


7-29 © 2023 LINIERS, DIST. BY KING FEATURES SYNDICATE TWITTER: @LINIERSCARTOON

<https://comicskingdom.com/macanudo/2023-07-29>

Start with the end in mind

Why?

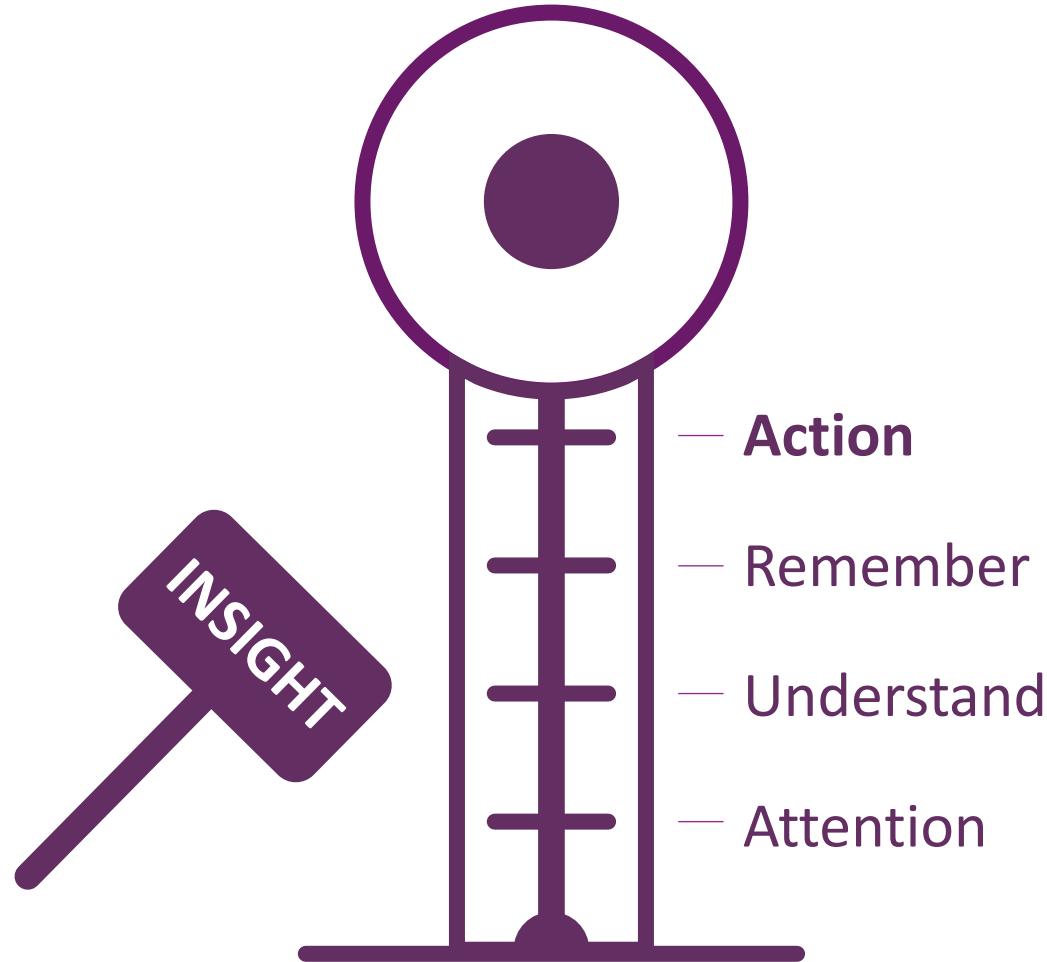


What?

Who?

Where?

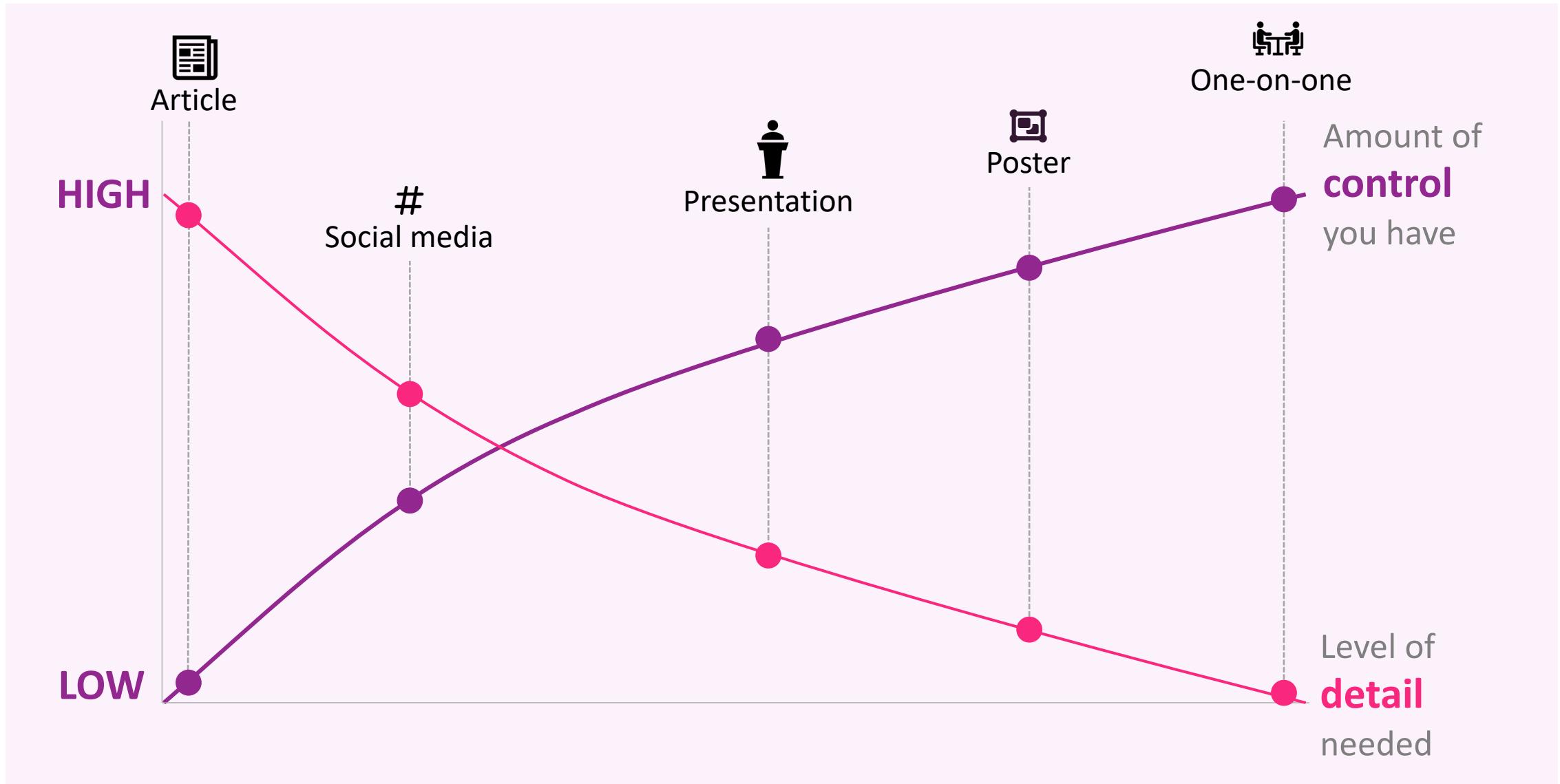
Why should your audience care?



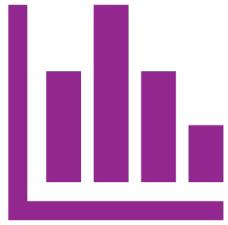
What is your **purpose**?
What **action** do you want your audience to take?

← **Why?**
← **Why?**
← **Why?**

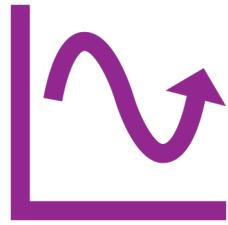
Where do you present?



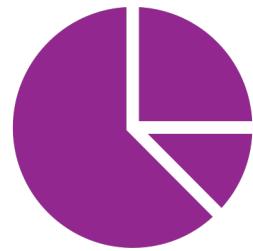
What do you want to show?



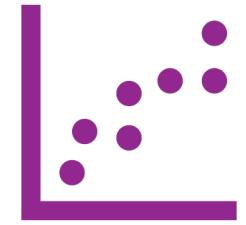
Comparison



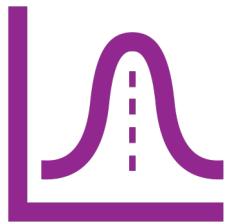
Trend



Part of a whole



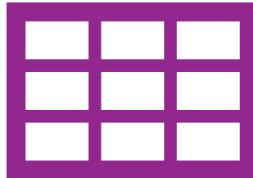
Relationship



Distribution



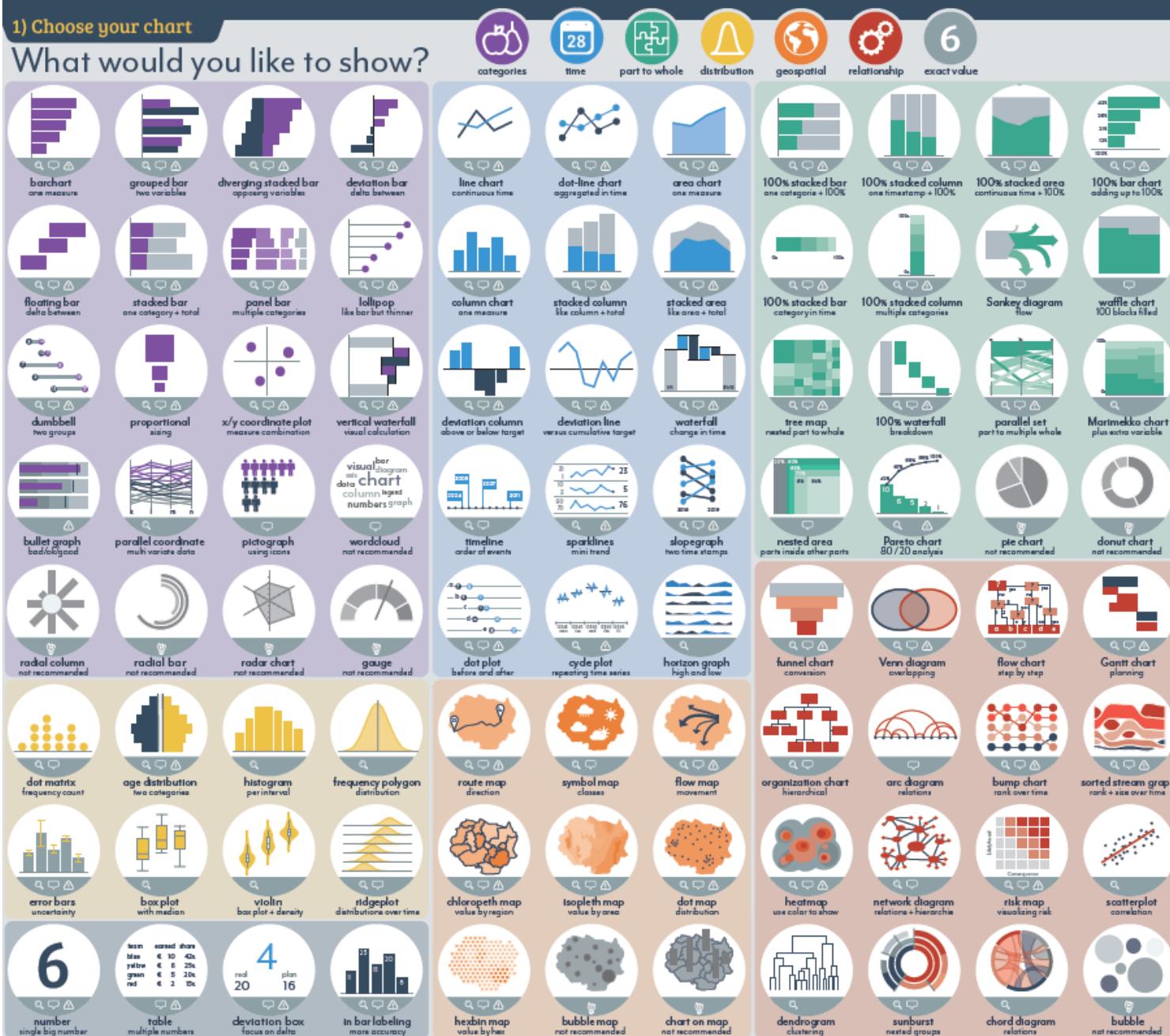
Location



Exact values

1) Choose your chart

What would you like to show?



Use this chart for: analysing, communicating, monitoring, confusing

Find more tips to choose and design your Perfect Chart at:

2) Design your chart

Let your data speak

do's

Show in context

Rank your data in a relevant way so patterns and outliers become visible.

relevant ranking

Enrich your chart with target or benchmarks, to give the values more meaning.

Support easy comparison

When you have more than 4 series, change your chart to small multiples.

small multiple

Gridlines make it easier to compare length of columns and bars and help to see the sharpness of lines.

gridlines

Visual hierarchy

De-emphasize all non-data elements like axis and legend. The data is more important.

less is more

Highlight the most important element, to make it stand out.

emphasize

Show and tell

Create a title to name the insight in the chart.

descriptive title

Show as little decimals as possible. Always show the same number of decimals.

too much details

Labels and annotations help your data to tell a story.

annotation

Please print/poster as PDF available at www.chart-guide.poster

Copyright, poster and chart design: www.Chart.Guide - Version 4.0.2

don'ts

Misleading design

A truncated axis in a column or bar chart distorts the relative size of the columns.

cutting of Y axes

Multiple Y-axes give a false meaning to lines crossing or to the space between lines.

Beauty over accuracy

Do not use 3D to make a chart prettier, it will make it harder and more confusing to read.

3D effect

Although fluid lines might look nicer, they are not representing the data properly.

fluid lines

Too much

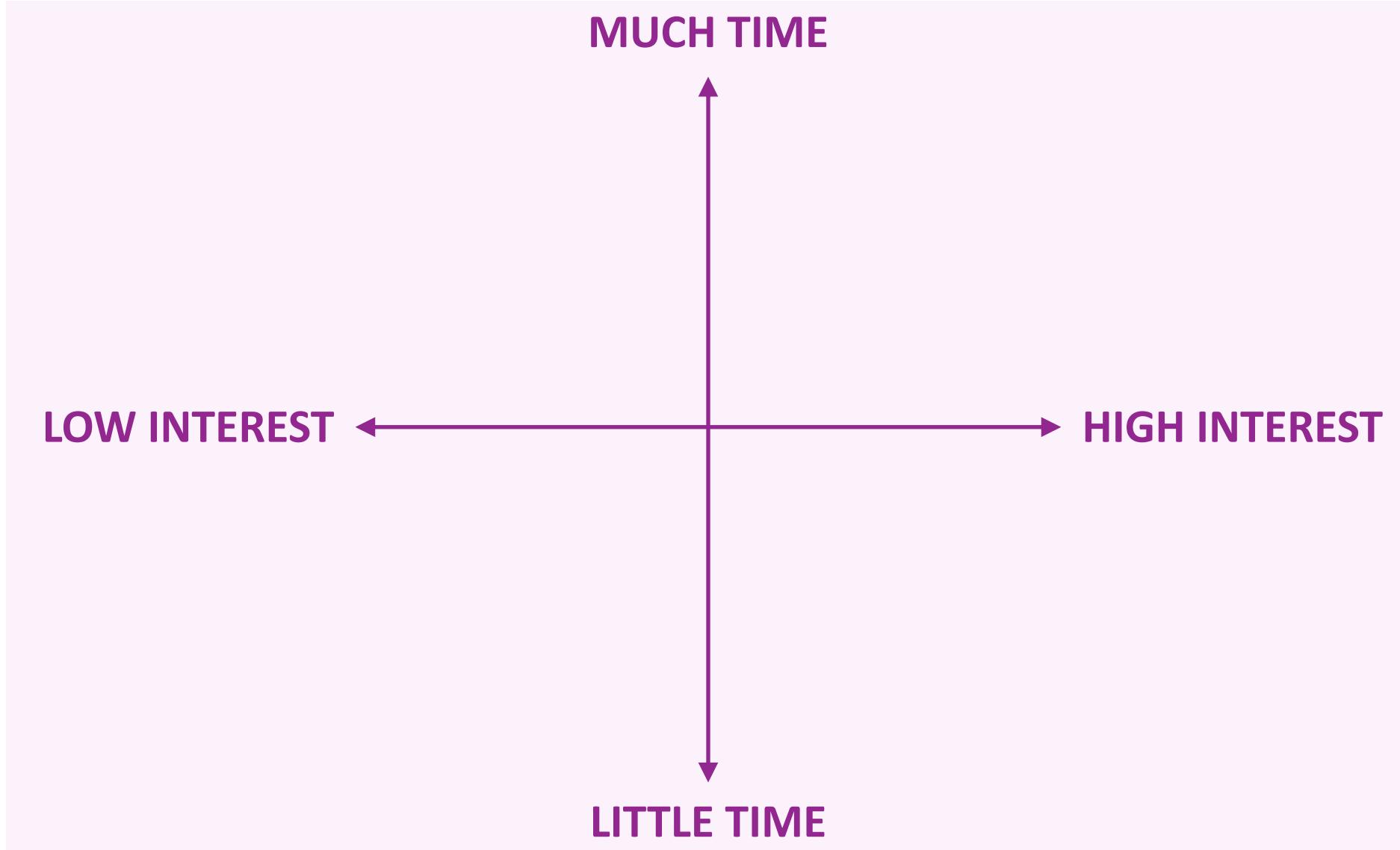
The human brain can only process a maximum of 4 series in a chart.

more than 4 series

Things that are the same, should have the same color

too much or meaningless color

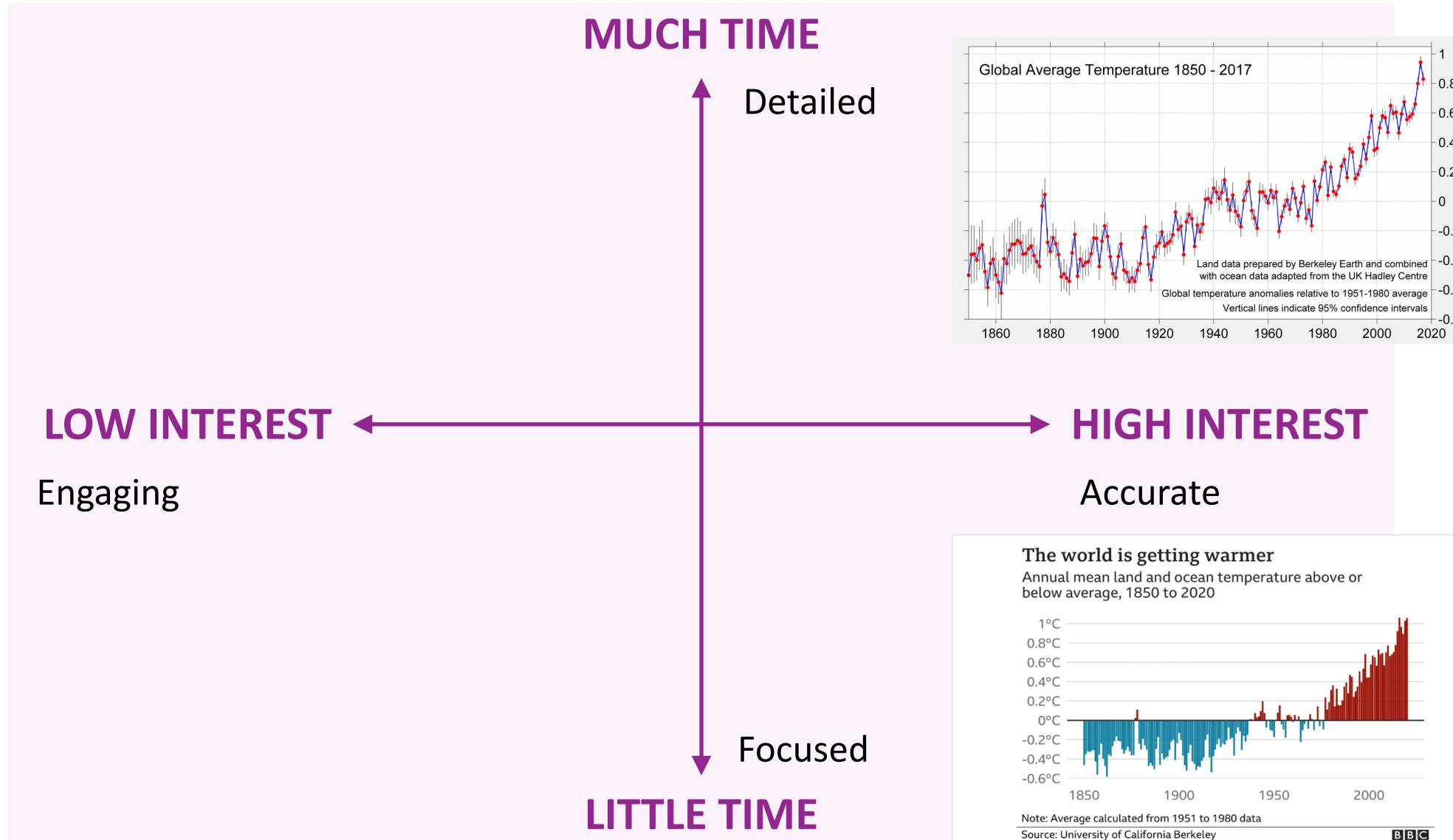
Who is your audience?



Who defines the level of detail and accuracy

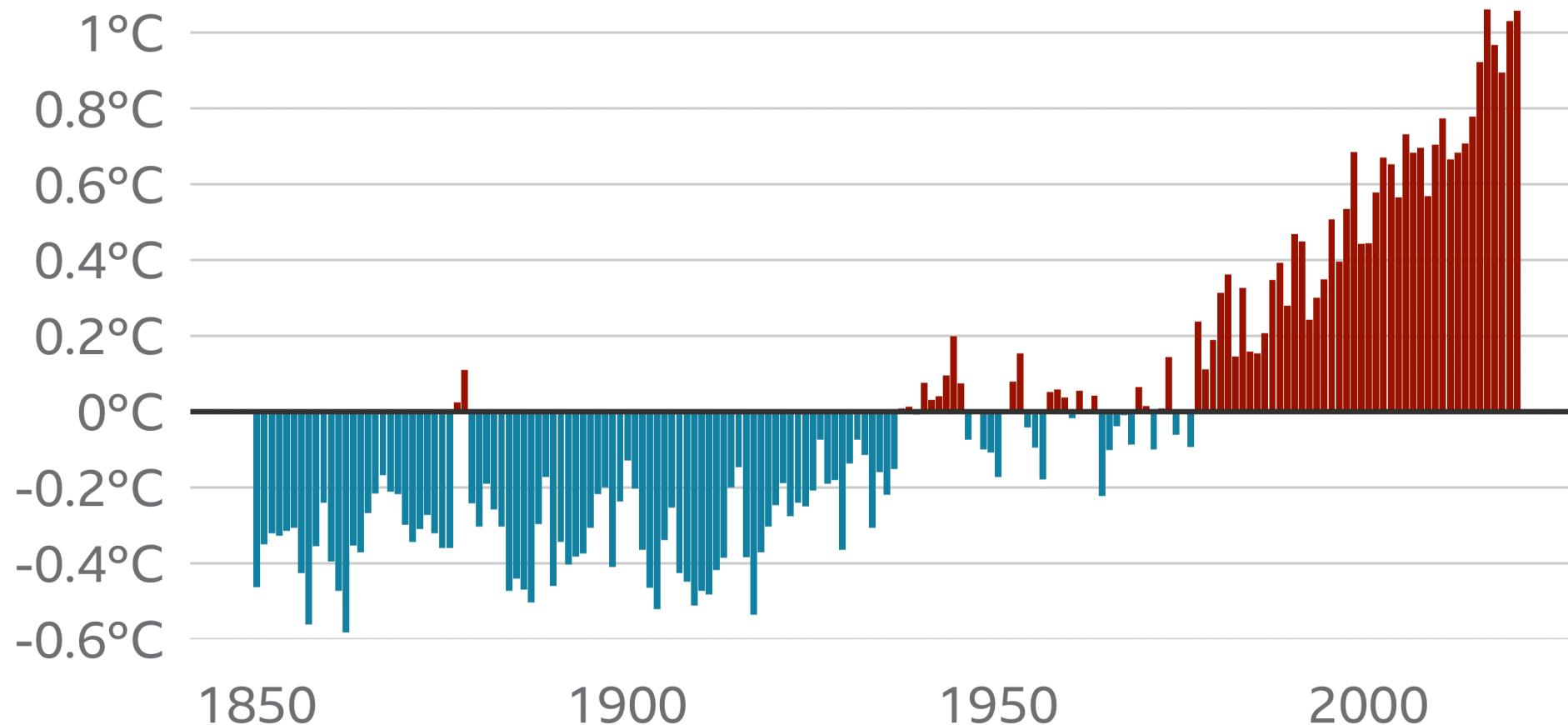


Who defines the level of detail and accuracy



The world is getting warmer

Annual mean land and ocean temperature above or below average, 1850 to 2020

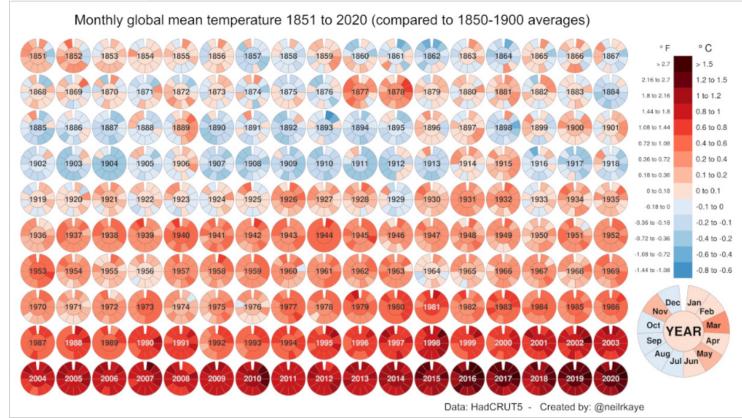


Note: Average calculated from 1951 to 1980 data

Source: University of California Berkeley

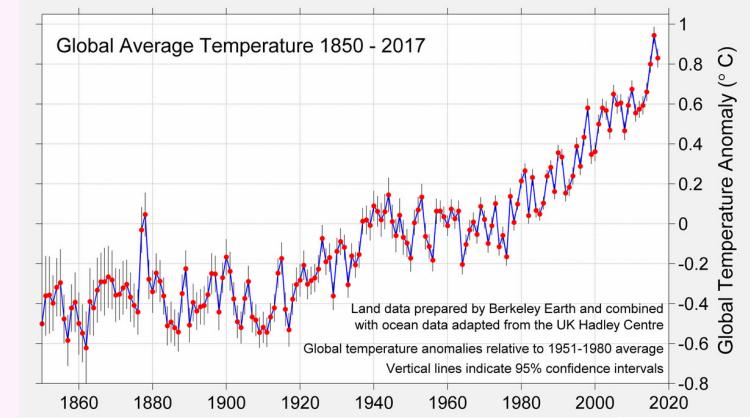
BBC

Who defines the level of detail and accuracy



MUCH TIME

Detailed



LOW INTEREST

Engaging

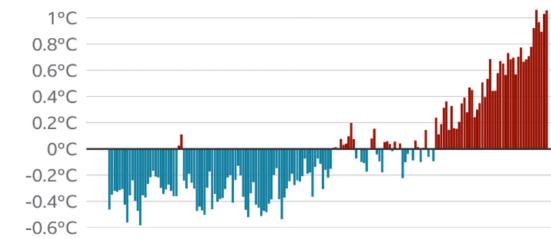
HIGH INTEREST

Accurate

LITTLE TIME

Focused

The world is getting warmer
Annual mean land and ocean temperature above or below average, 1850 to 2020

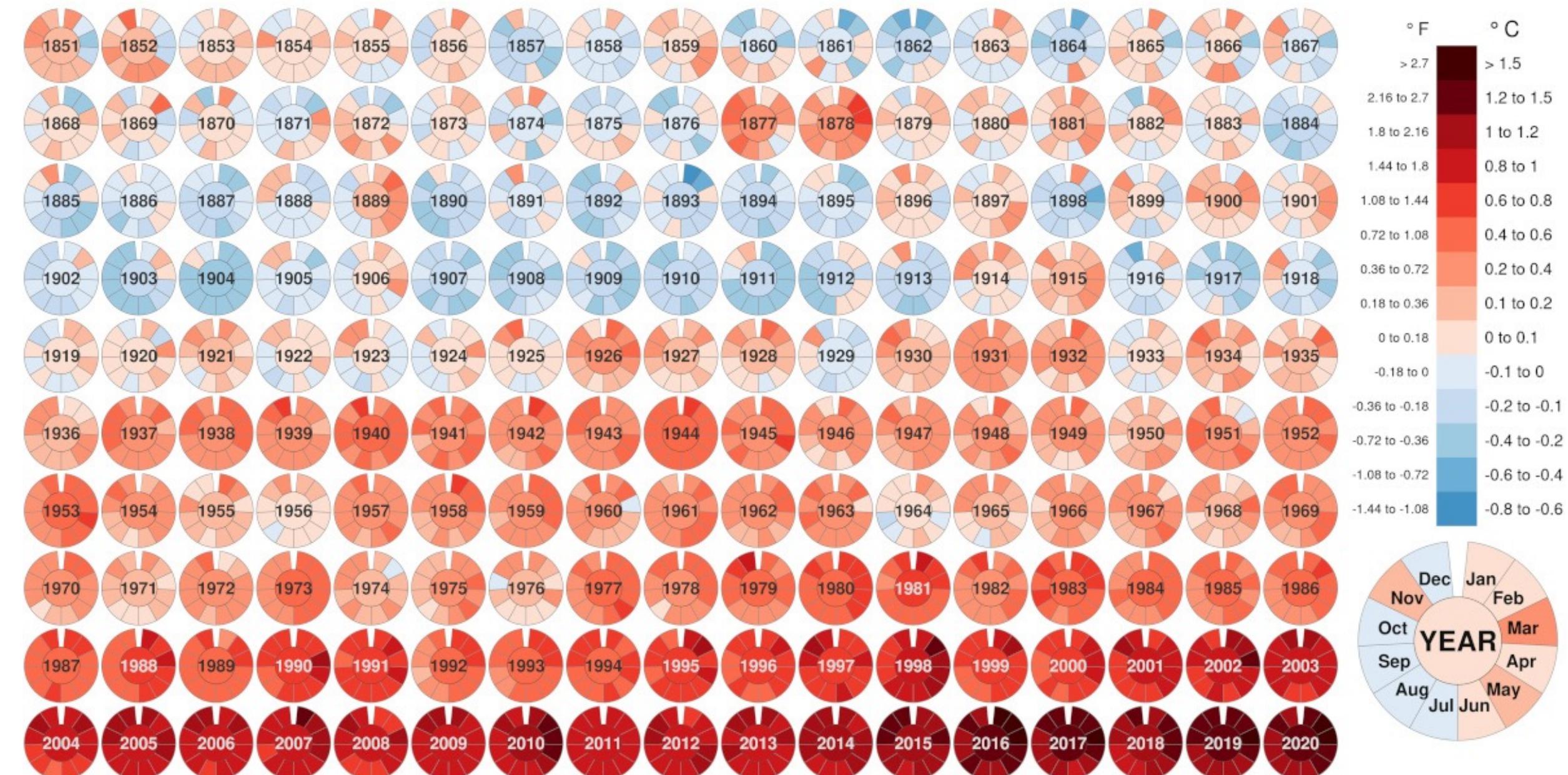


Note: Average calculated from 1951 to 1980 data

Source: University of California Berkeley

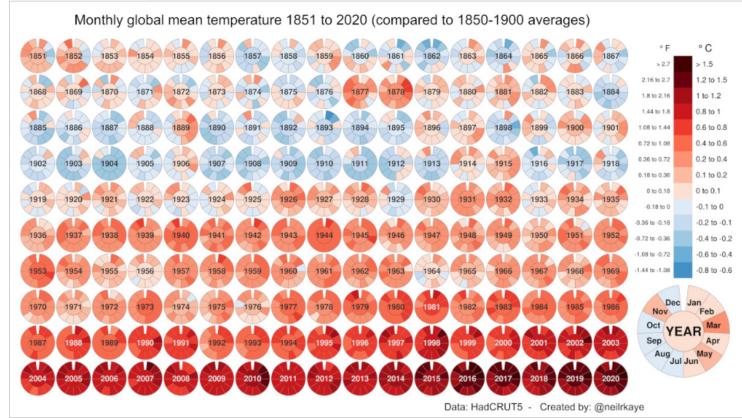
BBC

Monthly global mean temperature 1851 to 2020 (compared to 1850-1900 averages)



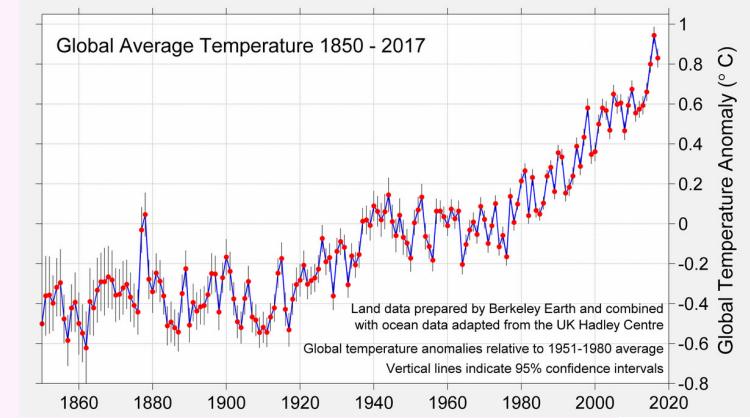
Data: HadCRUT5 - Created by: @neilrkaye

Who defines the level of detail and accuracy



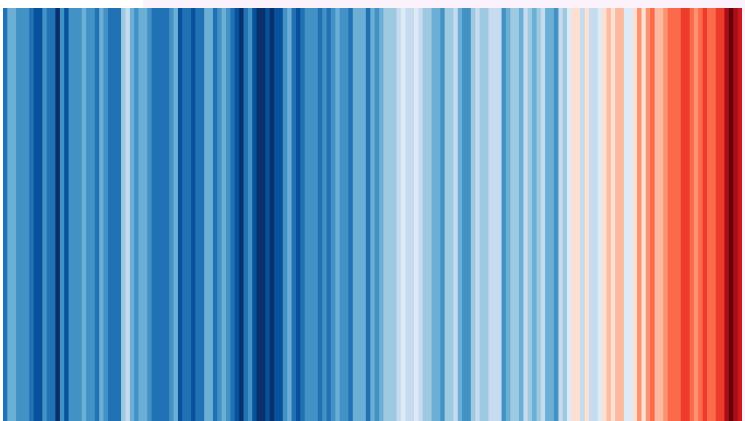
MUCH TIME

Detailed



LOW INTEREST

Engaging



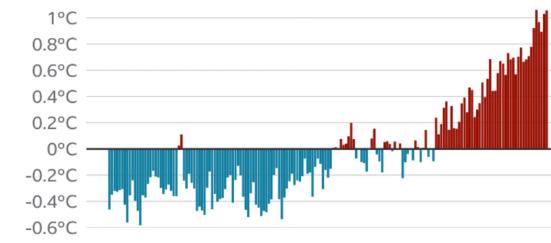
HIGH INTEREST

Accurate

LITTLE TIME

Focused

The world is getting warmer
Annual mean land and ocean temperature above or below average, 1850 to 2020



Note: Average calculated from 1951 to 1980 data

Source: University of California Berkeley

BBC



Exercise



Exercise: One dataset, multiple stories

No. infections	City
250	Amsterdam
110	Rotterdam
125	Den Haag
23	Utrecht
125	Eindhoven
55	Groningen
155	Leeuwarden
85	Enschede
120	Maastricht
15	Wageningen

5 min

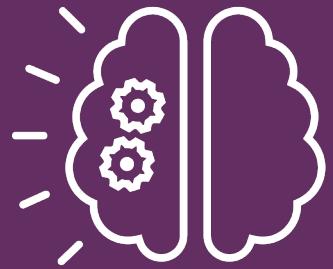


5 Scenarios:

The reader wants to know:

1. How many infections per city?
2. How does Maastricht rank?
3. How does the number of infections in Amsterdam relate to the other cities?
4. Which cities have more than 100 infections?
5. Is there a relation between location and the number of infections?

- Create a visual tailored to the reader's question
- Discuss: How can you adjust this visual for another (type of) audience?



How does the brain perceive
visual information?

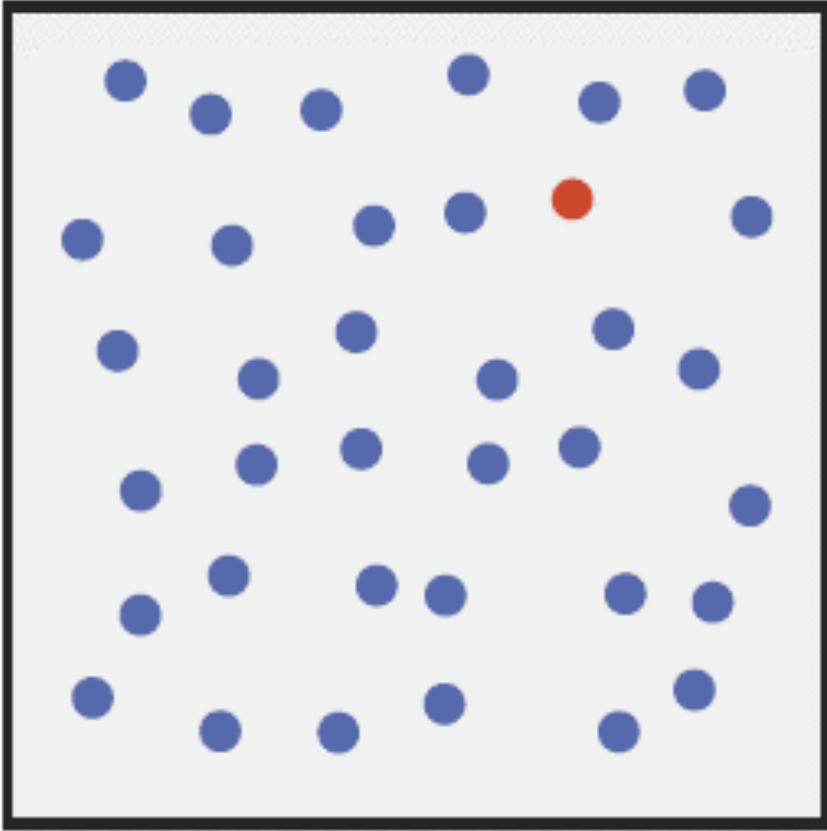
LEFT?

YES?

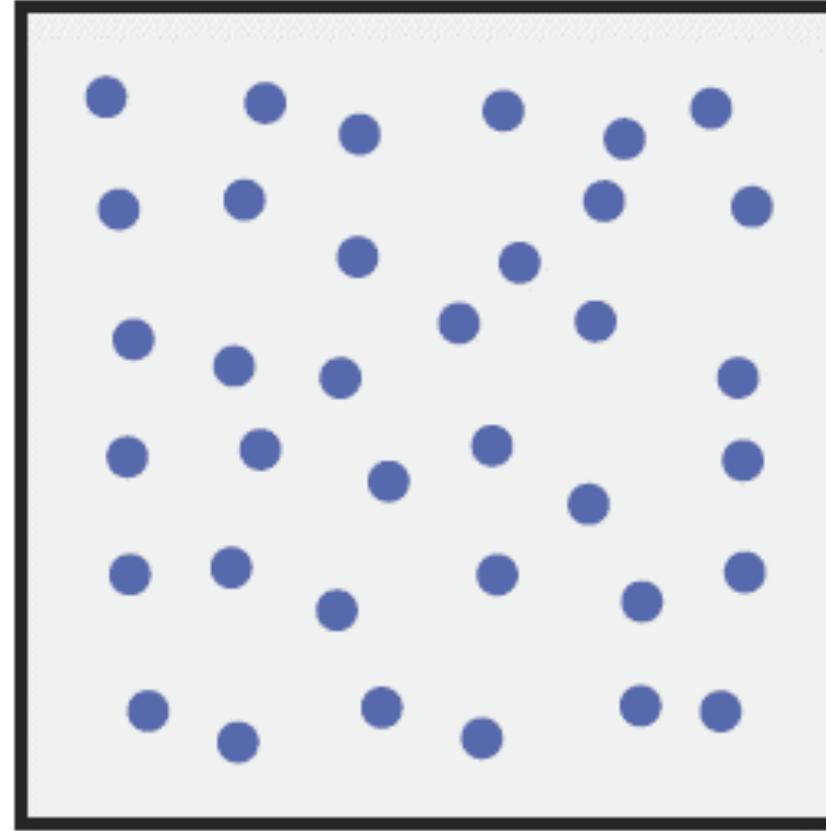


Which panel has a red dot?

A



B



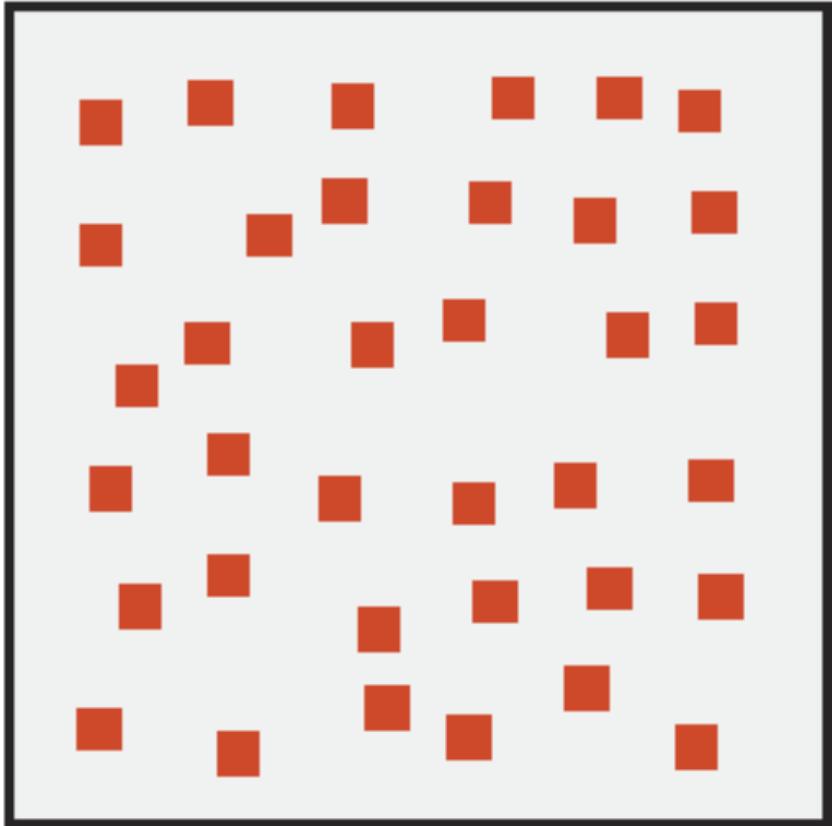
LEFT?

=

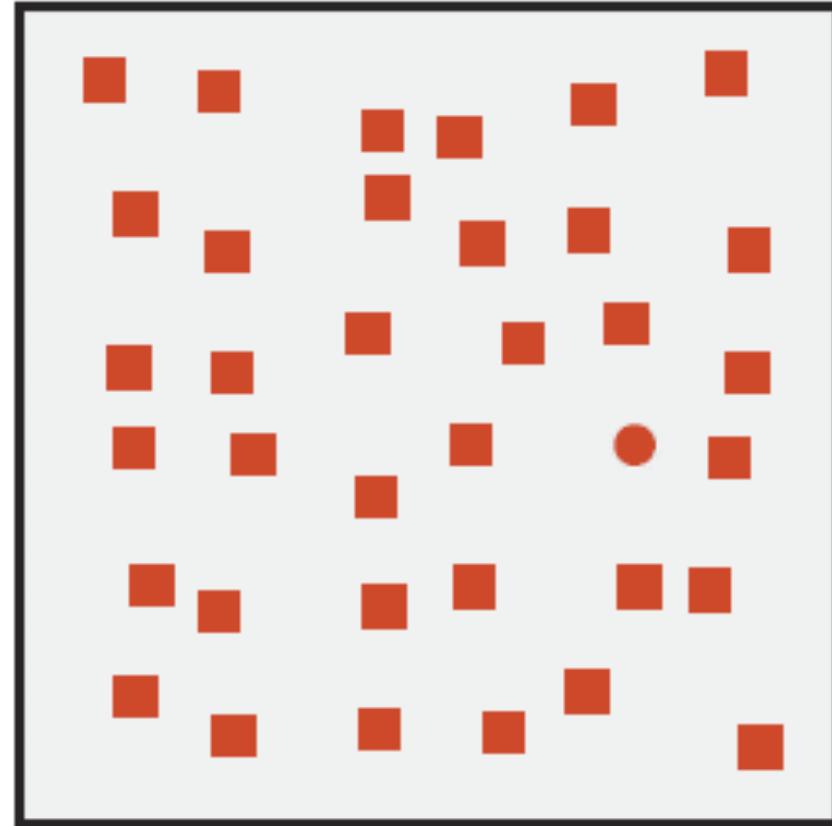
UP!

Which panel has a red dot?

A



B



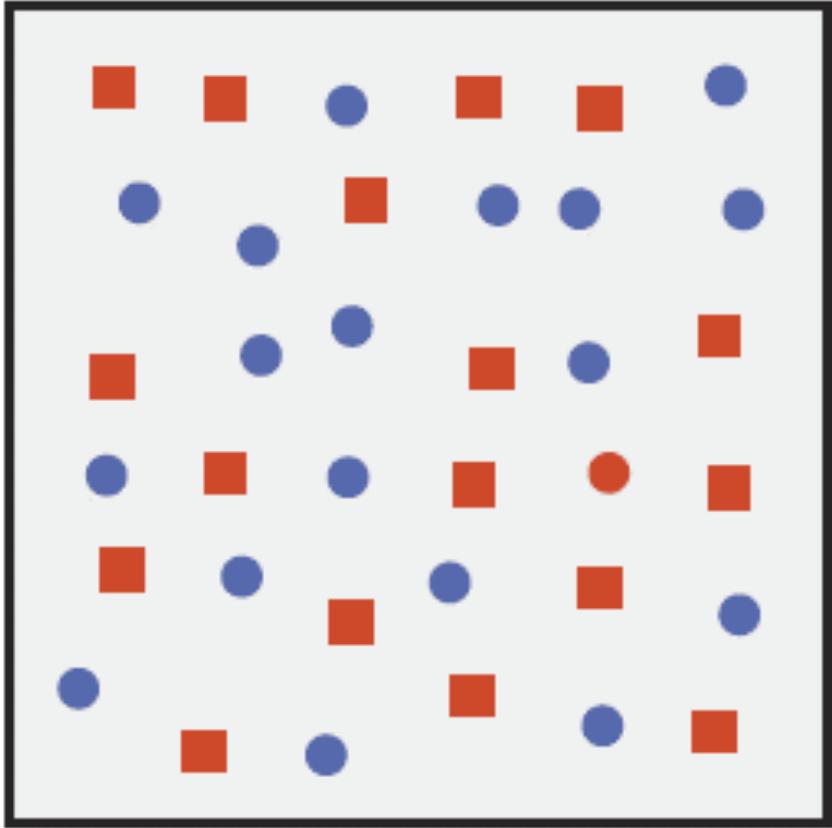
LEFT?

=

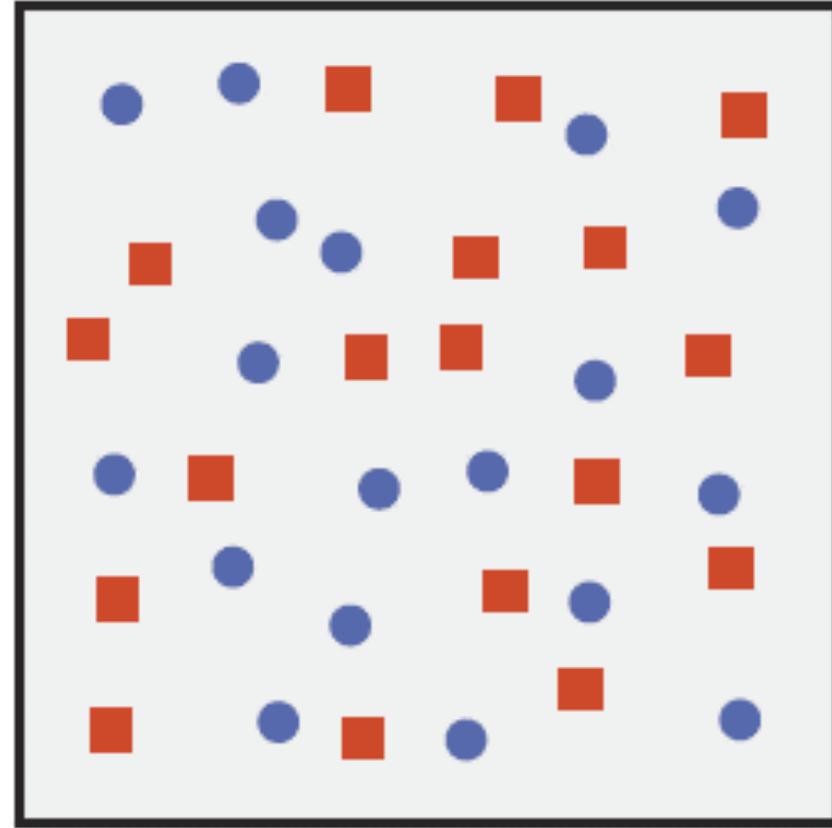
UP!

Which panel has a red dot?

A



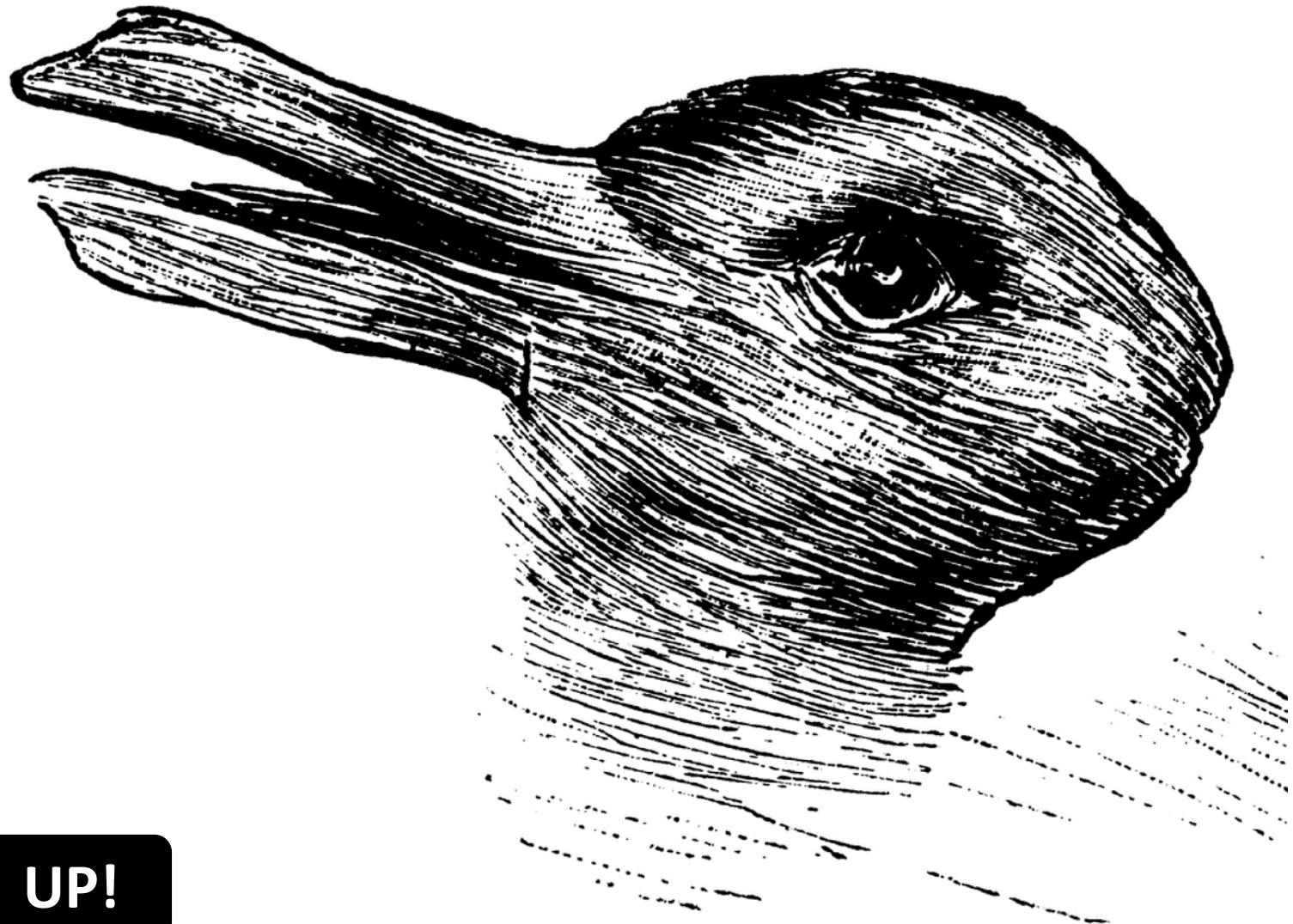
B



LEFT? =

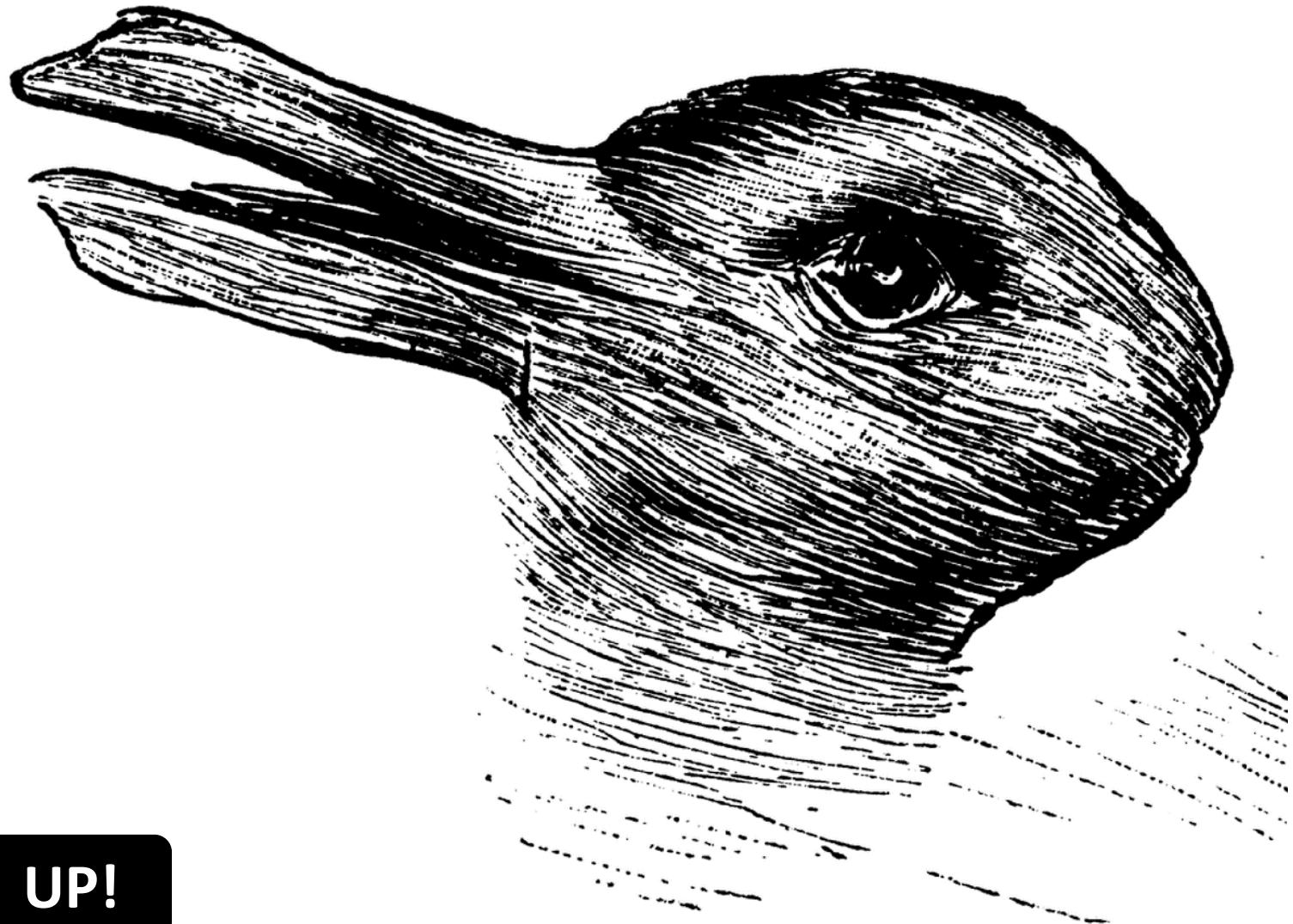
UP!

Do you see a Rabbit?



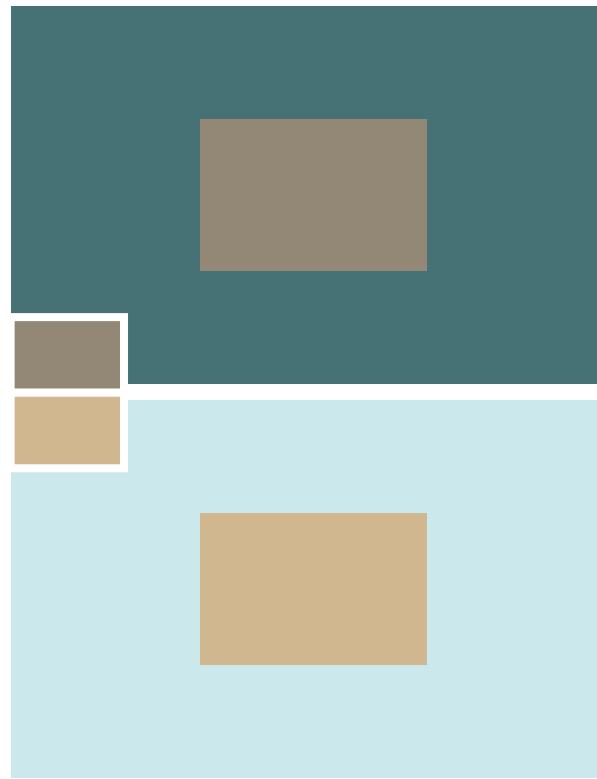
YES? = UP!

Do you see a Duck?



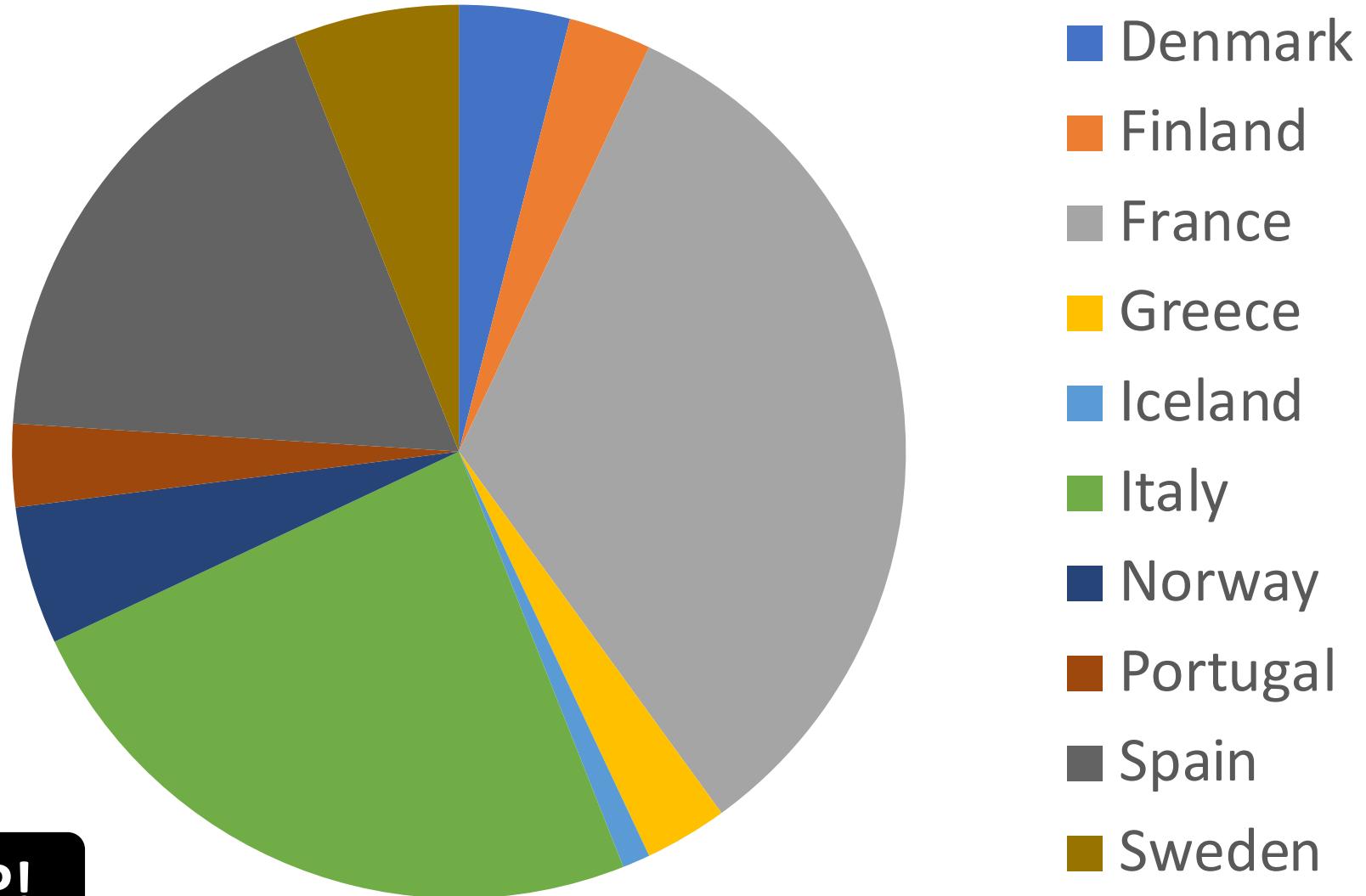
YES? = UP!

The inner squares have the same color



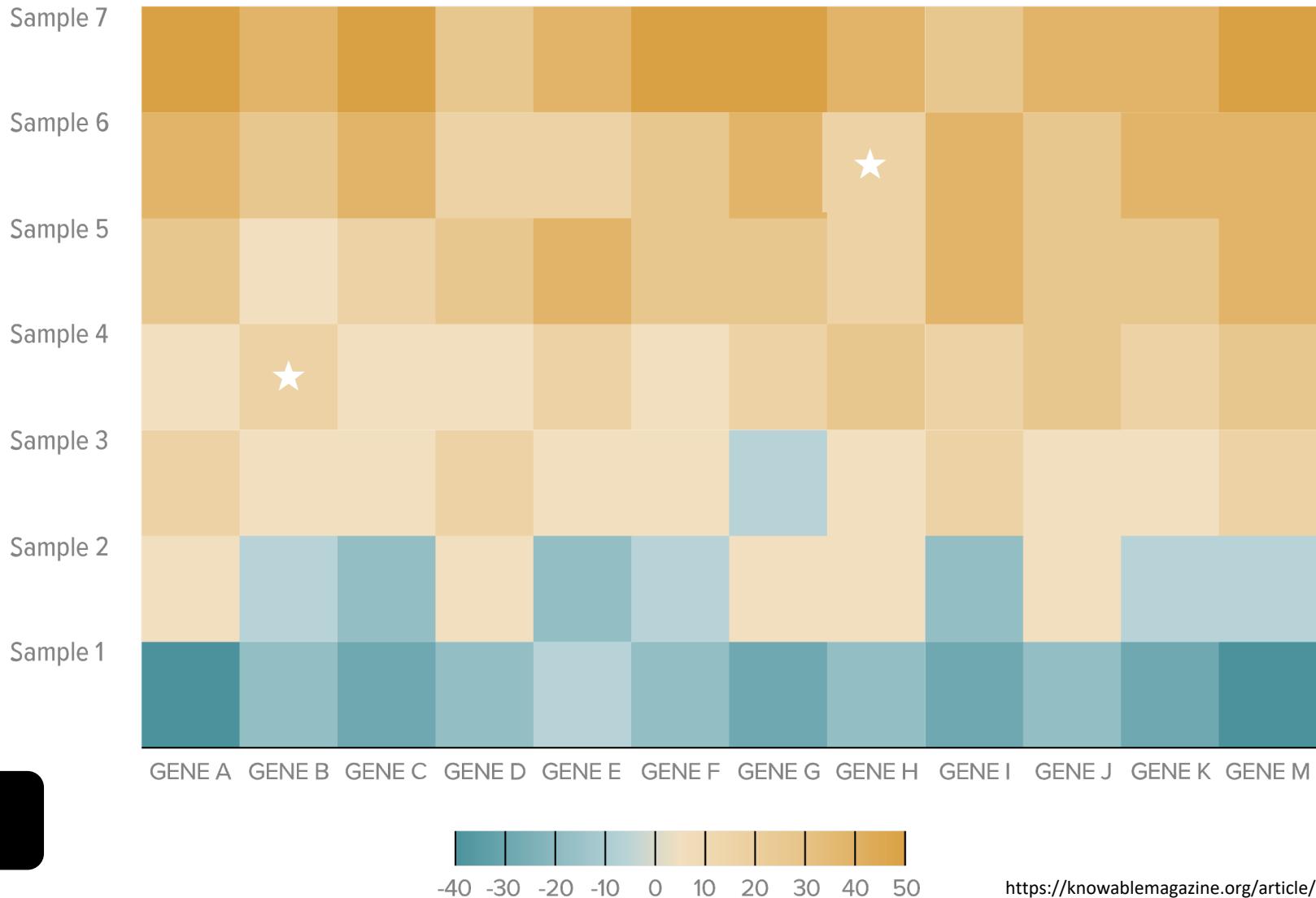
YES? = **UP!**

Norway has more widgets than Finland

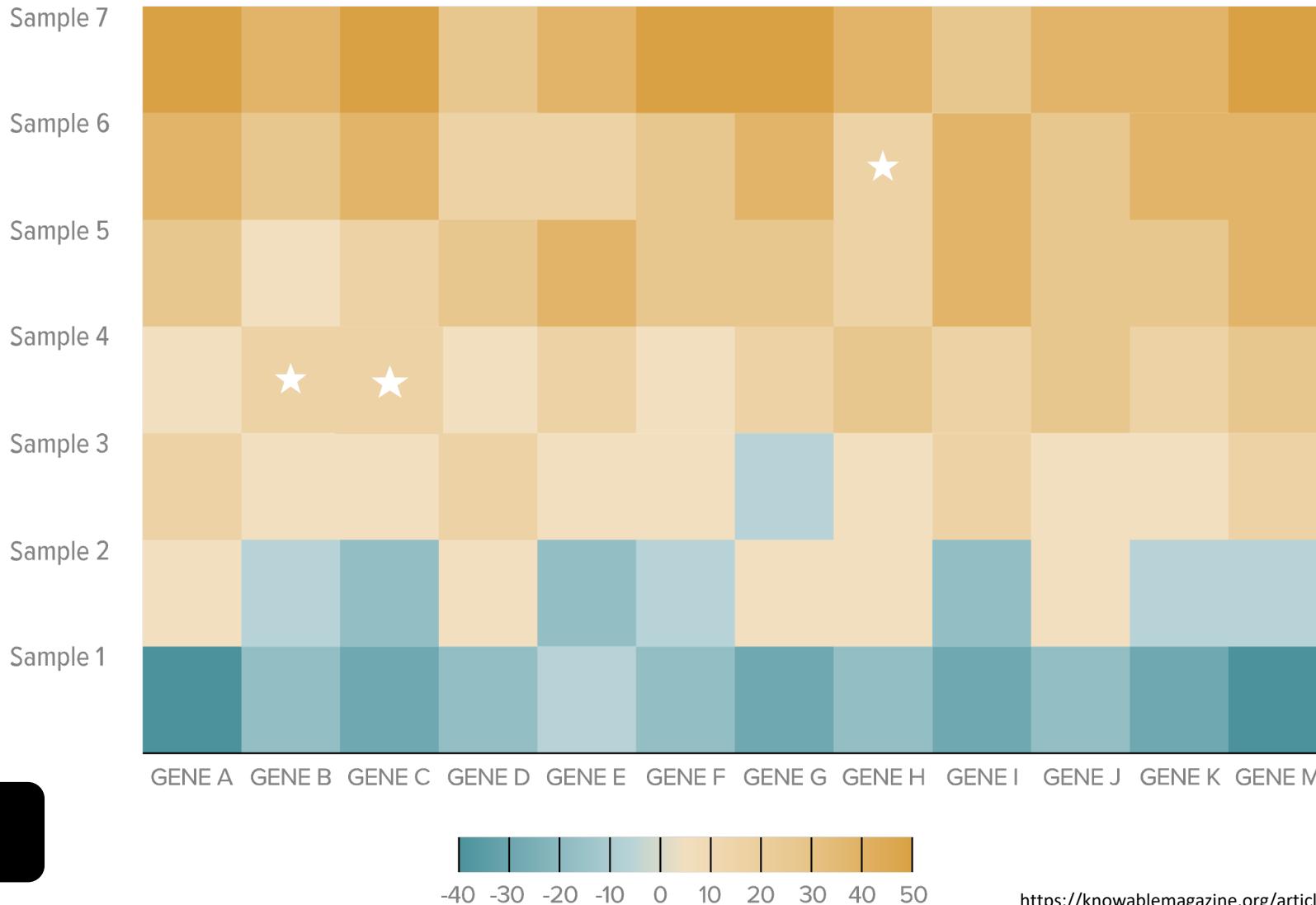


YES? = UP!

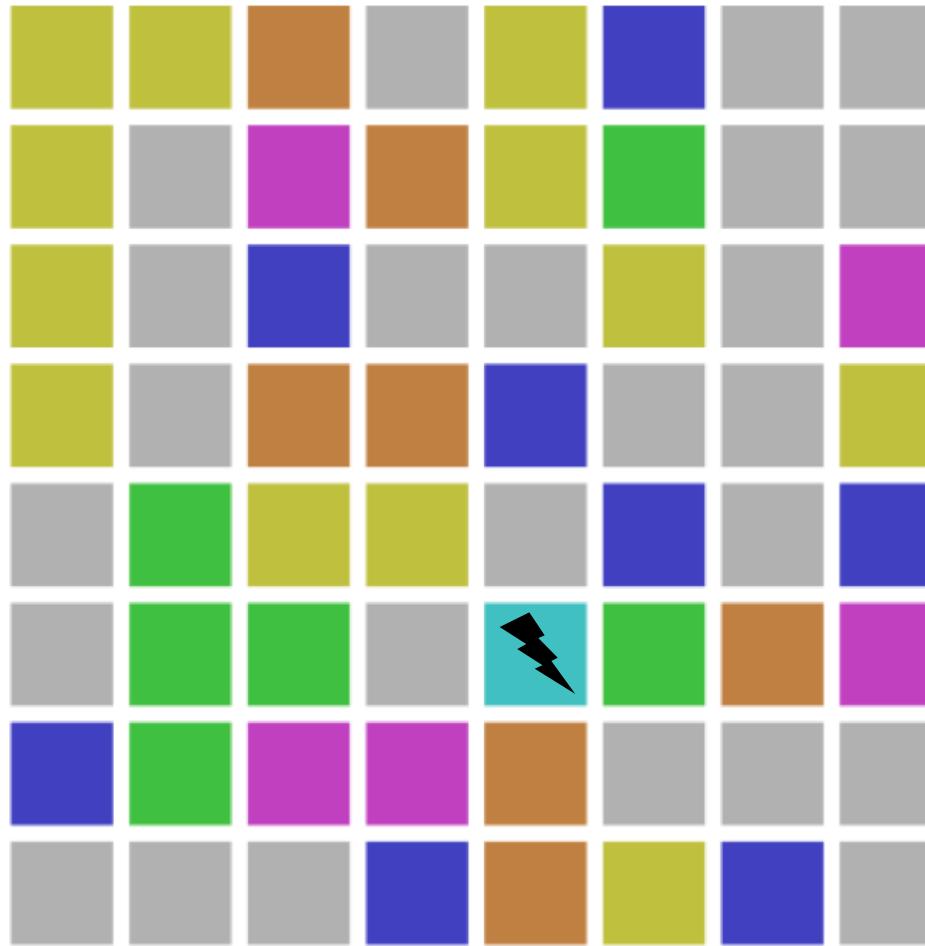
The areas with a star have the same color



The areas with a star have the same color

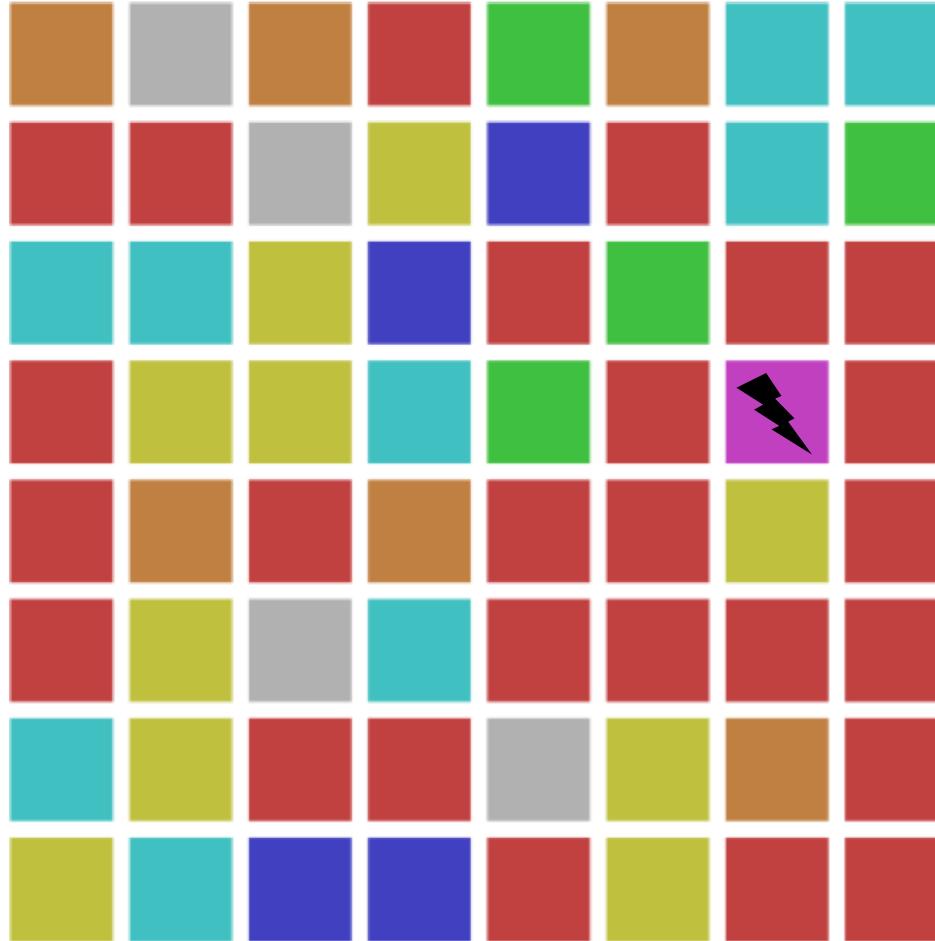


One square has a unique color



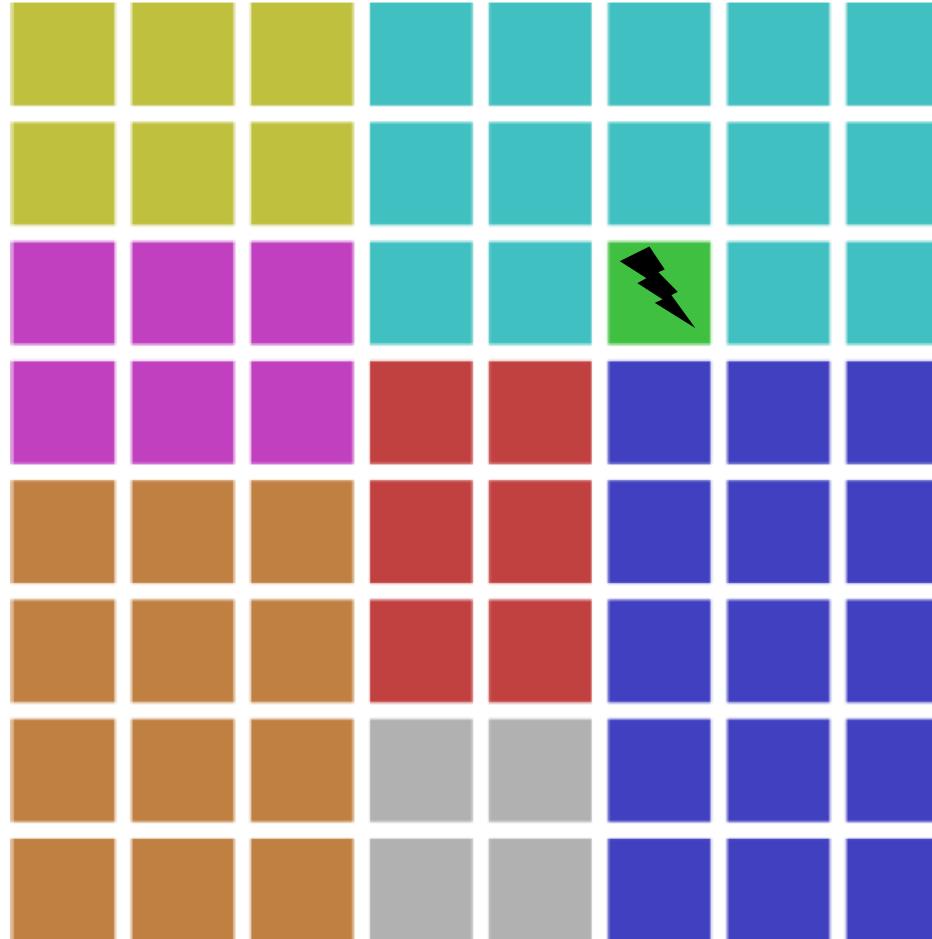
YES? = UP!

One square has a unique color



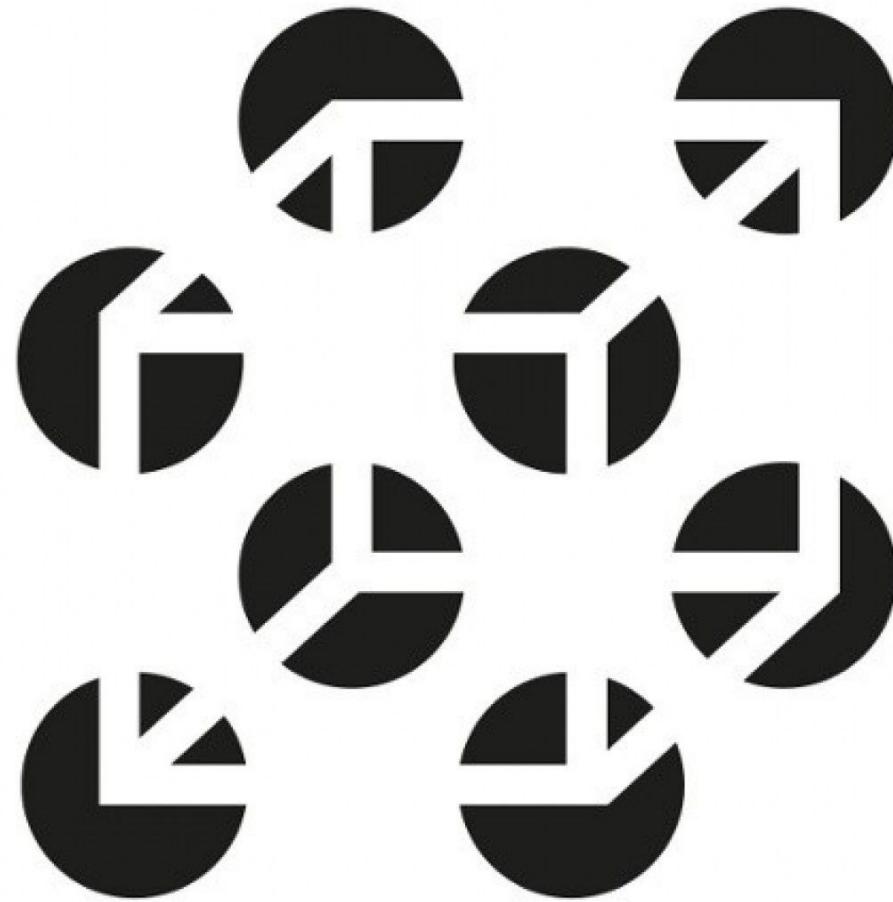
YES? = UP!

One square has a unique color



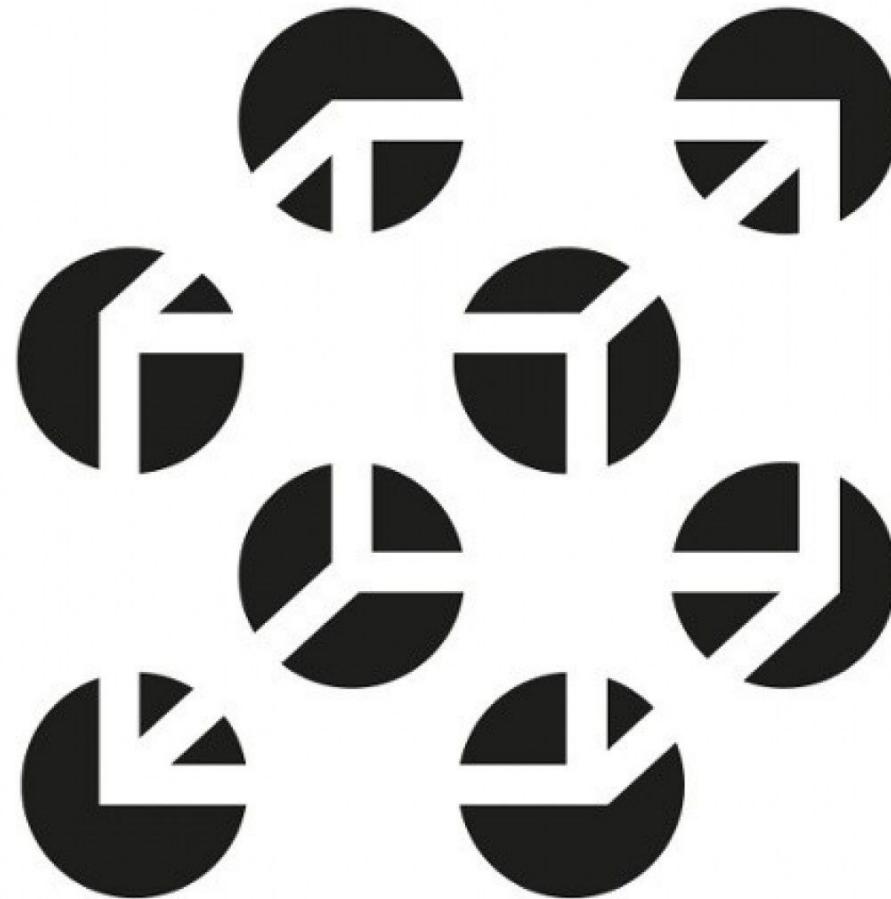
YES? = UP!

Do you *see* a cube?



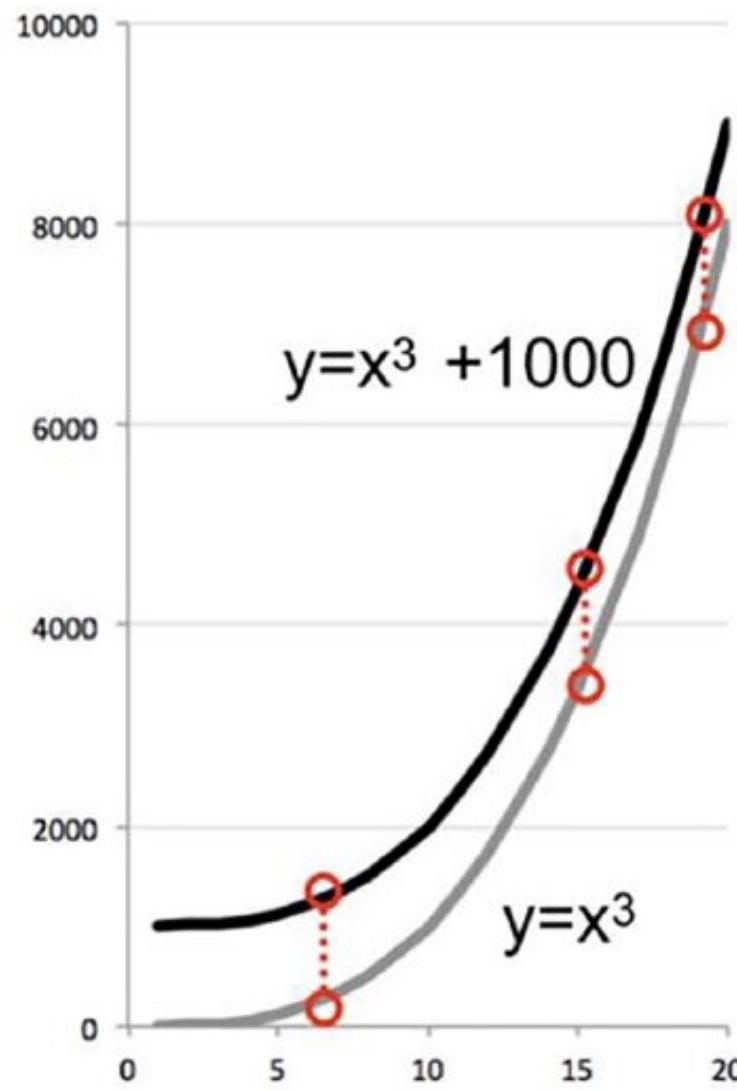
YES? = UP!

Is there a cube?



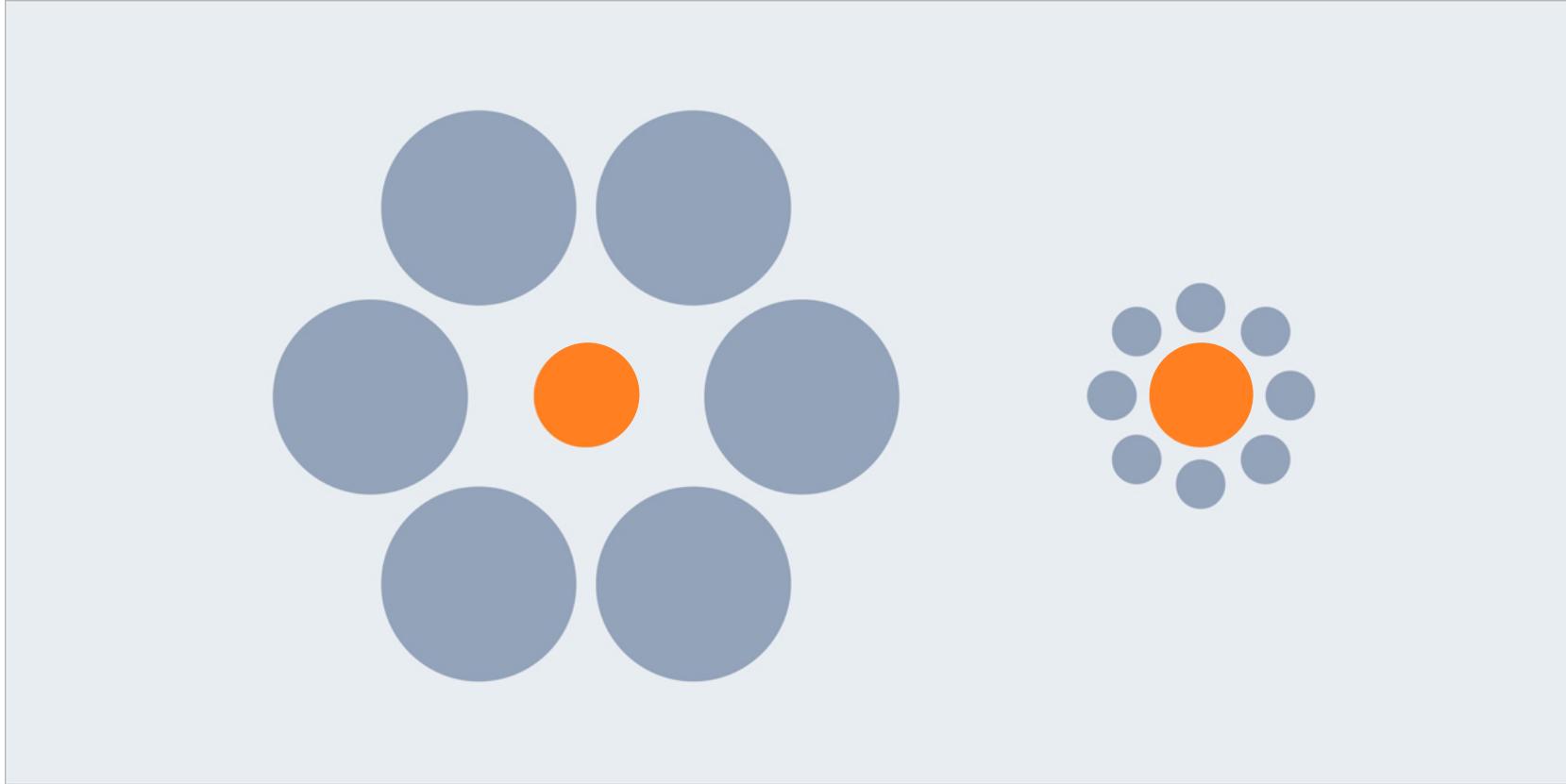
YES? = UP!

Where is the difference between the lines bigger? Left or right?



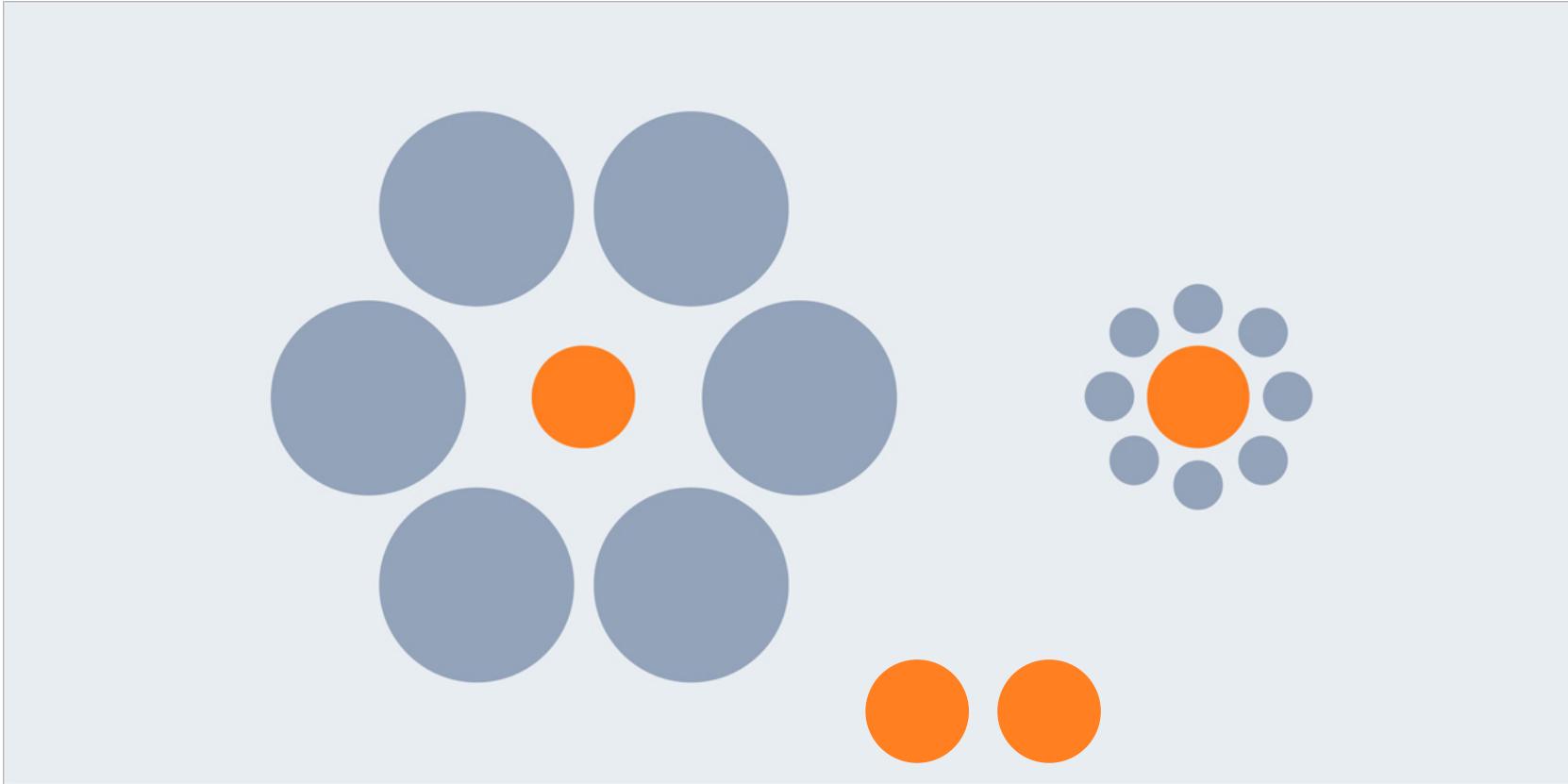
LEFT? = UP!

Which orange circle is smaller?



LEFT? = UP!

Which orange circle is smaller?



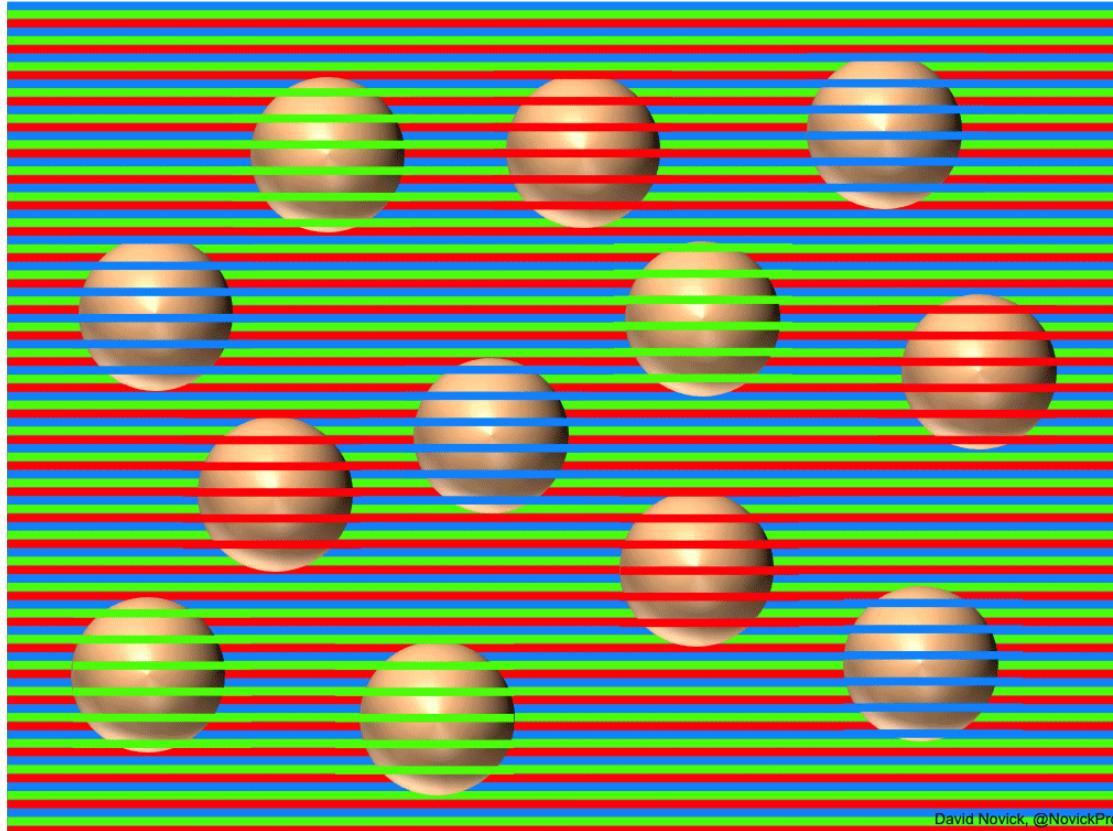
LEFT? = UP!

These spheres have the same color



YES? = UP!

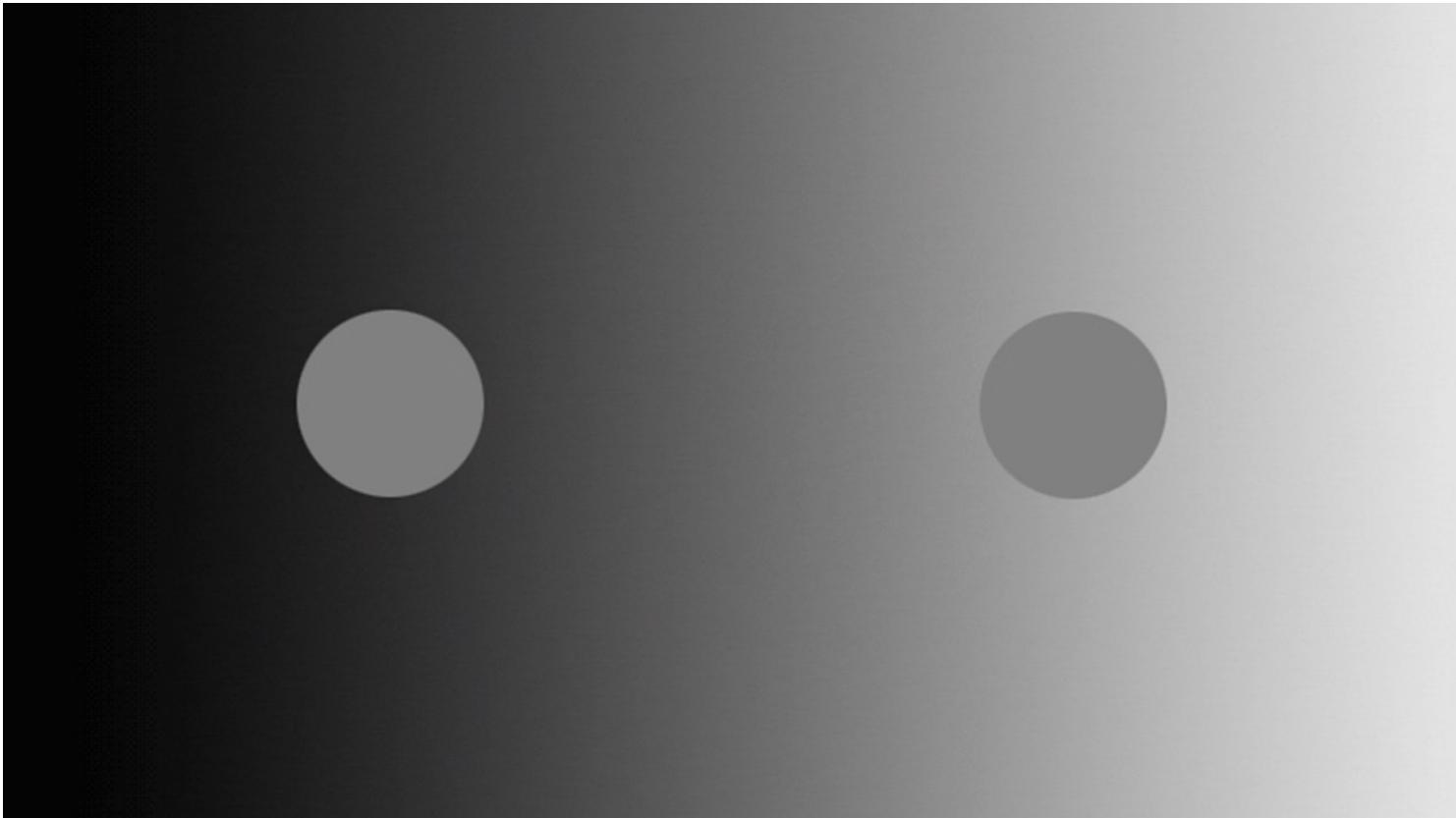
These spheres have the same color



David Novick, @NovickProf

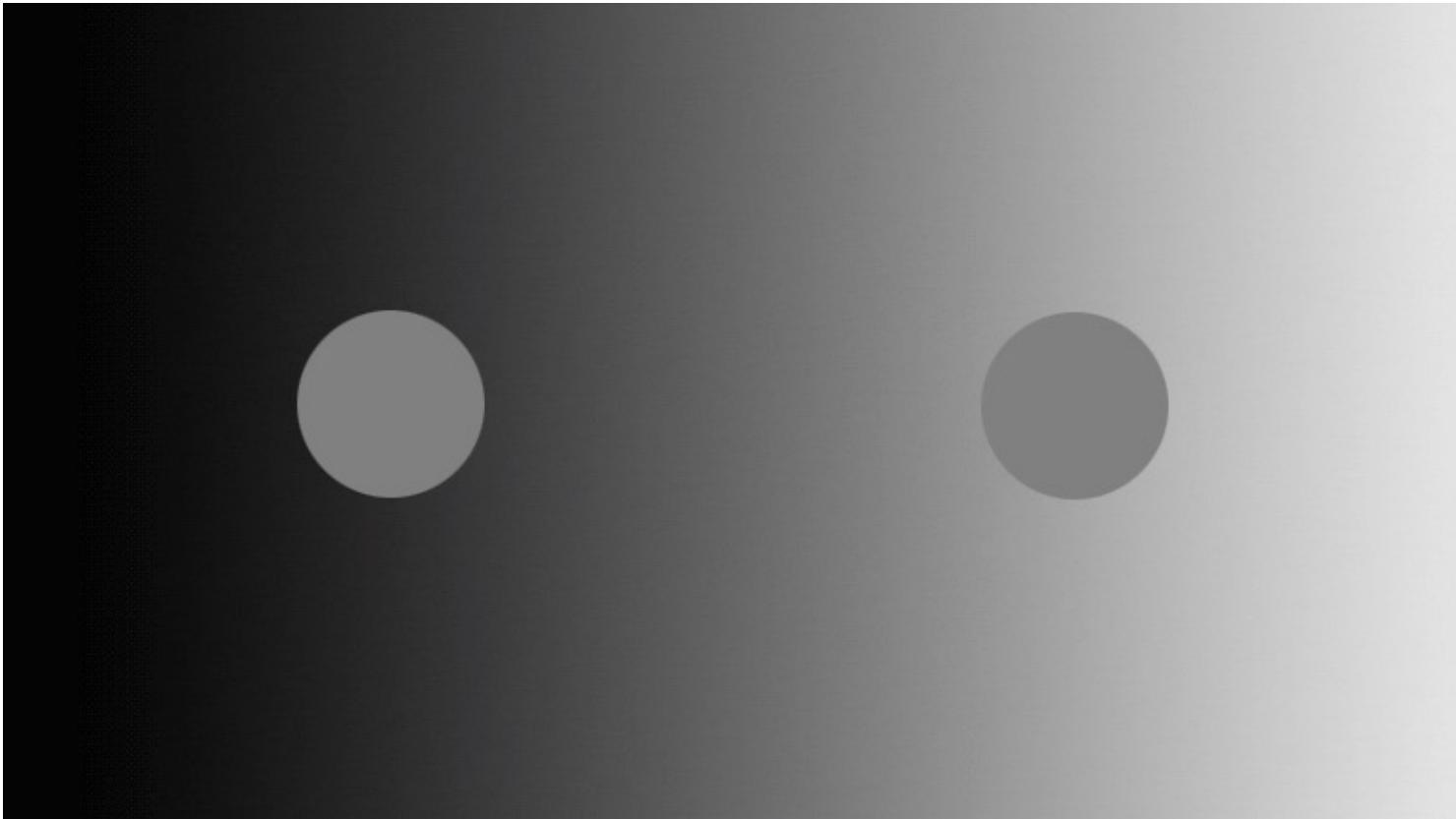
YES? = UP!

Which dot is darker?

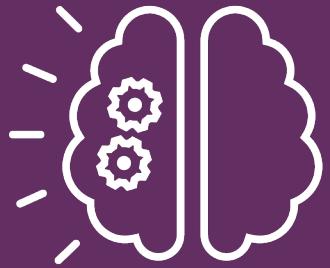


LEFT? = UP!

Are you sure...?



LEFT? = UP!



*“It’s not what you look at that
matters, it’s what you see”*

Henry David Thoreau



The strengths of our brain

“The eyes only see what the mind is prepared to comprehend.” – Henri Bergson.

How many fives do you see?

321654643216465132168461321
646432136748651349687463123
668432616266984892536496874
651687964649843616984616957
465167412323161689312439873
421432875943869234987596387

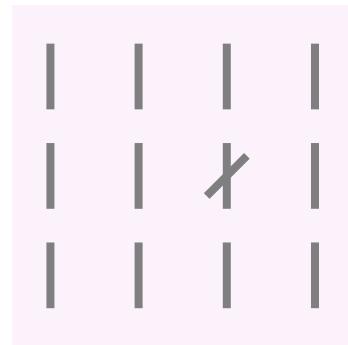
How many fives do you see?

321654643216465132168461321
646432136748651349687463123
668432616266984892536496874
651687964649843616984616957
465167412323161689312439873
421432875943869234987596387

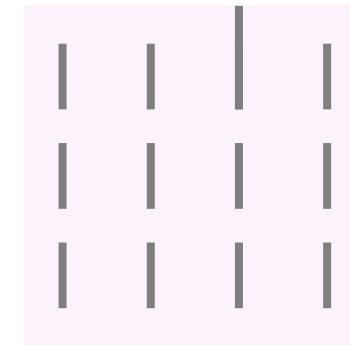
Pre-attentive attributes make your audience see what you want them to see before even seeing it!

321654643216465132168461321
646432136748651349687463123
668432616266984892536496874
651687964649843616984616957
465167412323161689312439873
421432875943869234987596387

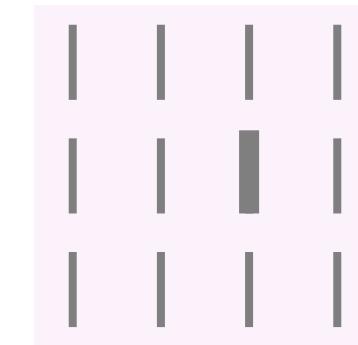
Pre-attentive attributes make your audience see what you want them to see before even seeing it!



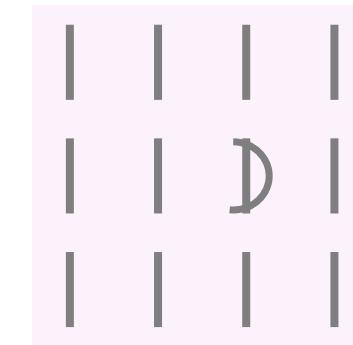
Orientation



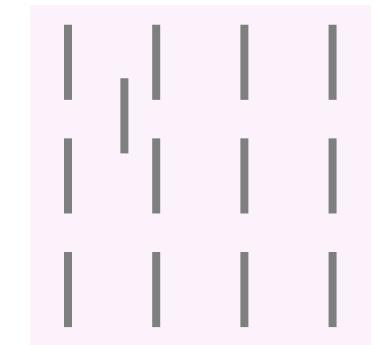
Length



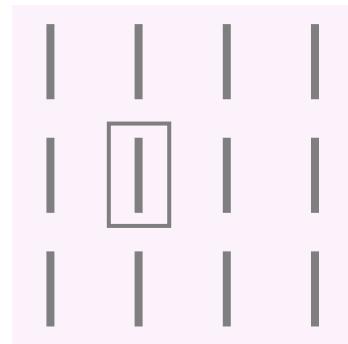
Width



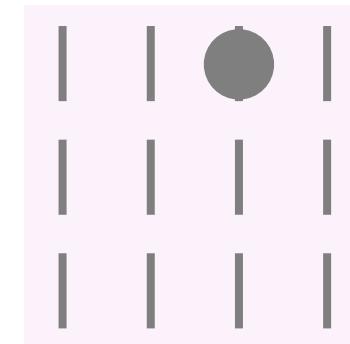
Curvature



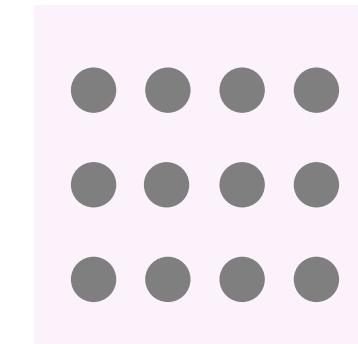
Position



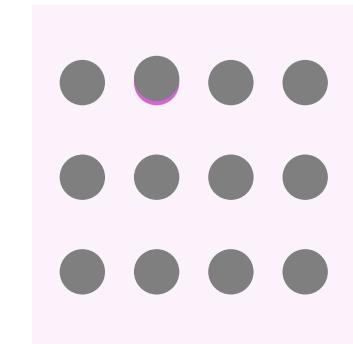
Enclosure



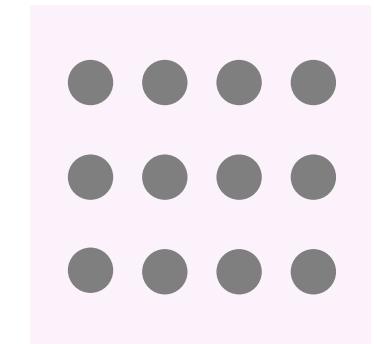
Shape



Size



Color



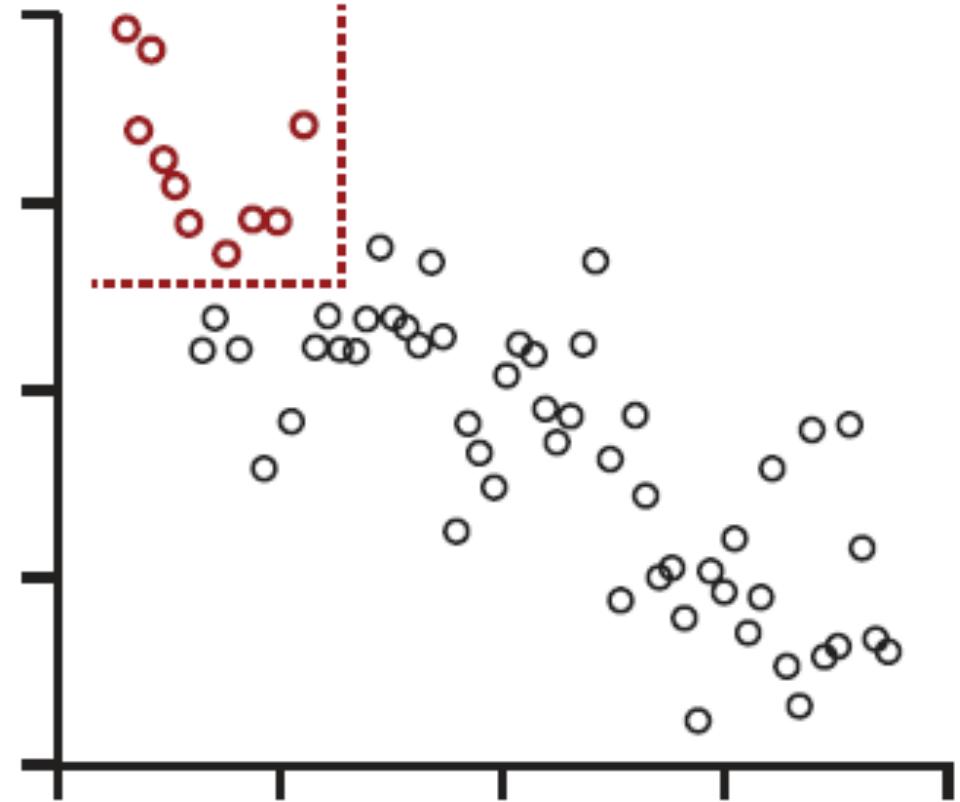
Motion

You can use pre-attentive attributes in **tables**

Color name	RGB (1-255)
Black	0, 0, 0
Orange	230, 159, 0
Sky blue	86, 180, 233
Bluish green	0, 158, 115
Blue	0, 114, 178
Vermillion	213, 94, 0

You can use pre-attentive attributes in tables and **charts**

Color name	RGB (1-255)
Black	0, 0, 0
Orange	230, 159, 0
Sky blue	86, 180, 233
Bluish green	0, 158, 115
Blue	0, 114, 178
Vermillion	213, 94, 0



You can also use pre-attentive attributes in text.

How can you leverage the use of pre-attentive attributes?

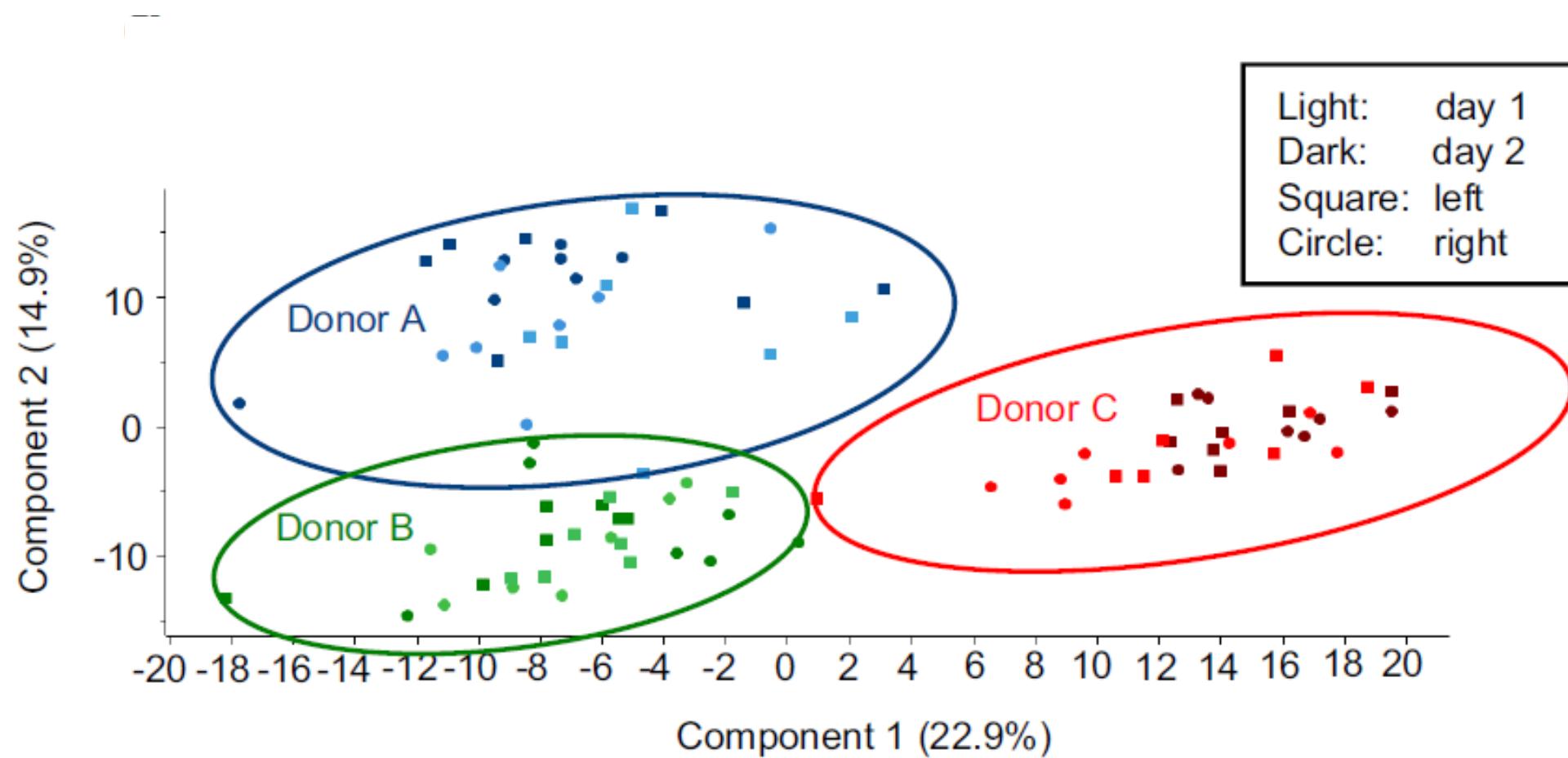
Without pre-attentive attributes you just have plain text. Or a plain figure. There is no guidance for the reader or viewer.

With pre-attentive attributes, you can:

- (1) *Draw your audience attention to where you want them to look*, and
- (2) *Create a visual hierarchy of information*

This counts for **graphs** as well as **text**!

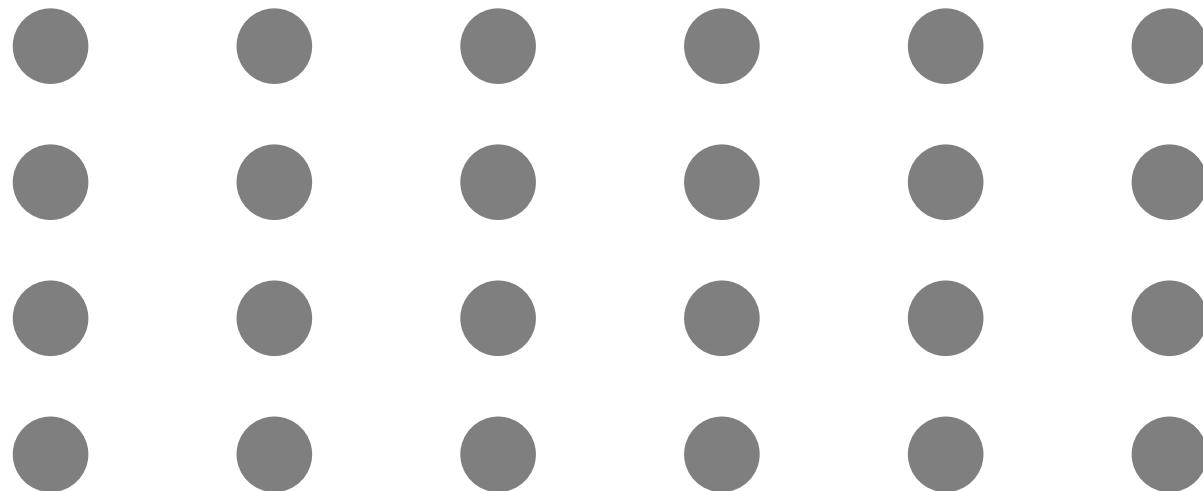
Pre-attentive attributes can create a visual hierarchy



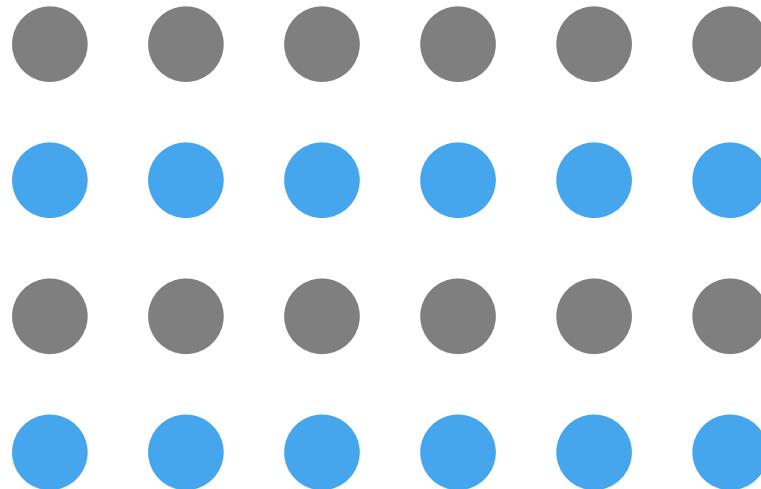
Do you see 4 rows or 6 columns?



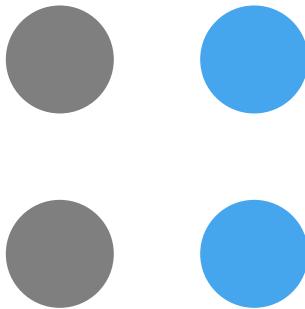
And now?



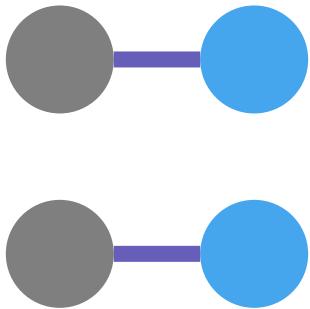
Do you see 4 rows or 6 columns?



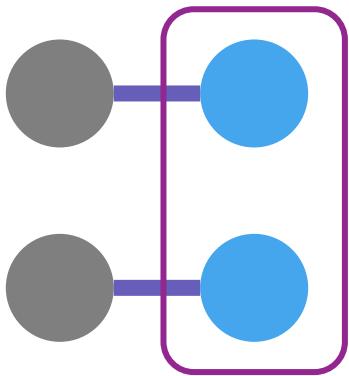
Which dots belong together?



And now...?



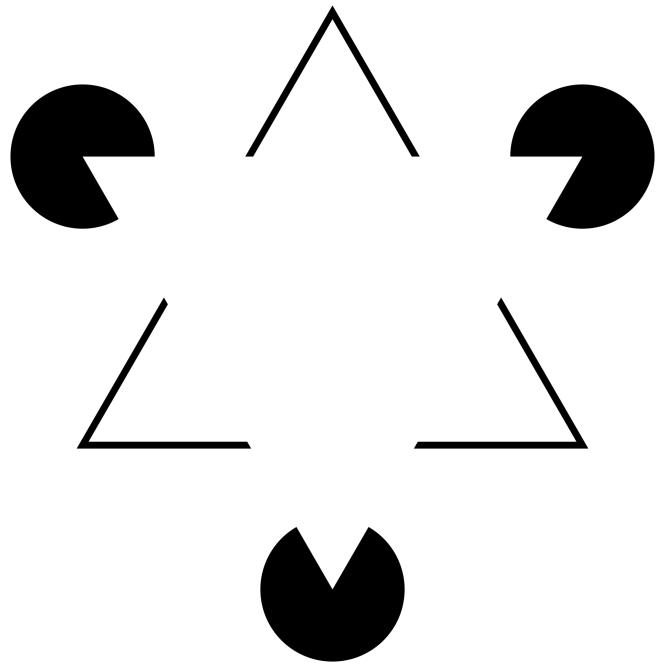
And now...?



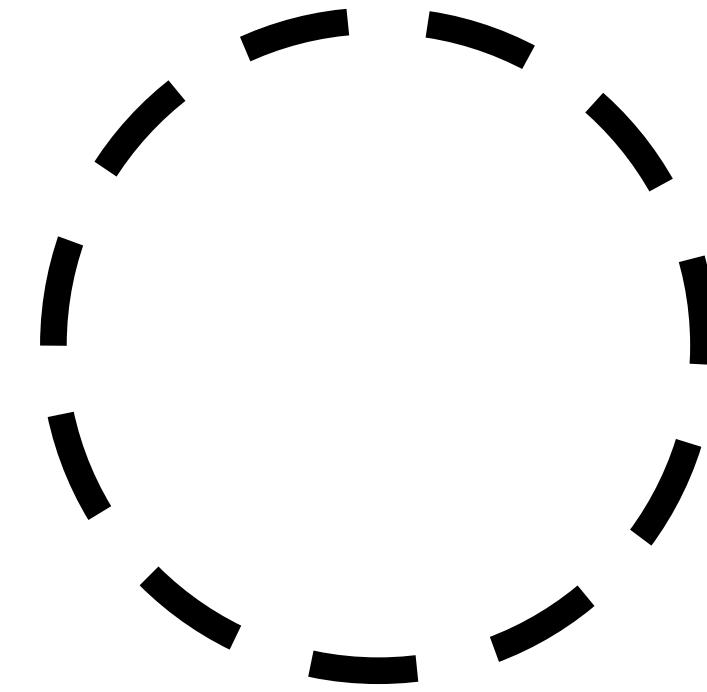
Can you read this?

Aoccdrnig to a rsceearh at Cmabrgide Uinervtisy, it deosn't
mttaer in what order the ltteers in a wrod are, the olny
iprmoatnt tnihg is taht the frsit and lsat ltteer be at the
rghit pclae. The rset can be a toatl mses and you can siltl
raed it wouthit porbelm. Tihs is bcuseae the huamn mnid
deos not raed ervey lteter by istlef, but the wrod as a
wlohe.

We're good at seeing things that may not explicitly exist!



Continuity



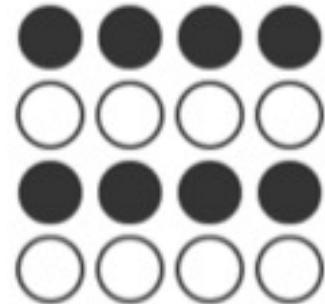
Closure

Gestalt principles: The whole is other than the sum of the parts.

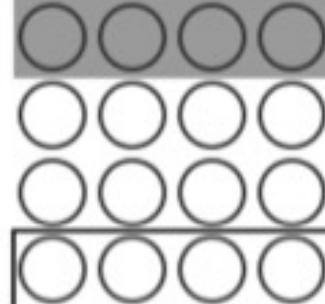
Proximity



Similarity



Enclosure



Symmetry

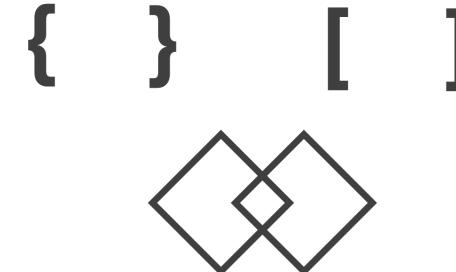
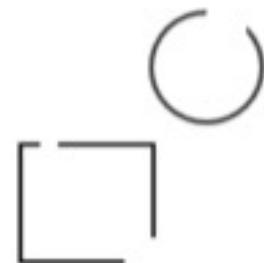


Figure-ground



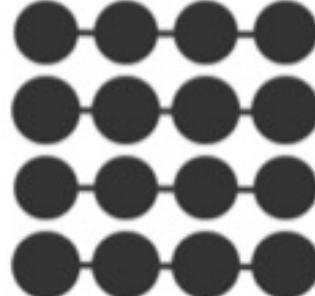
Closure



Continuity



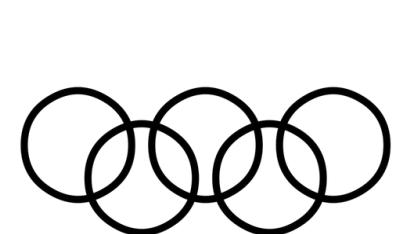
Connection



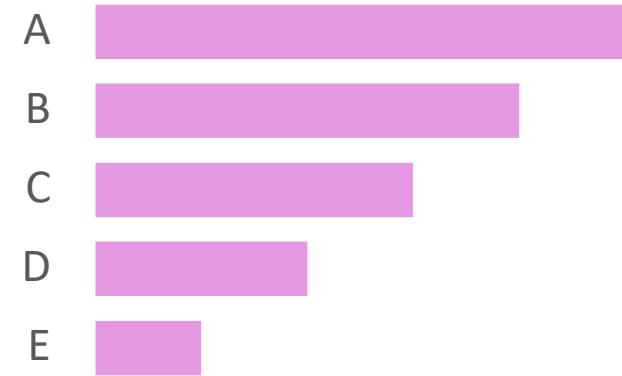
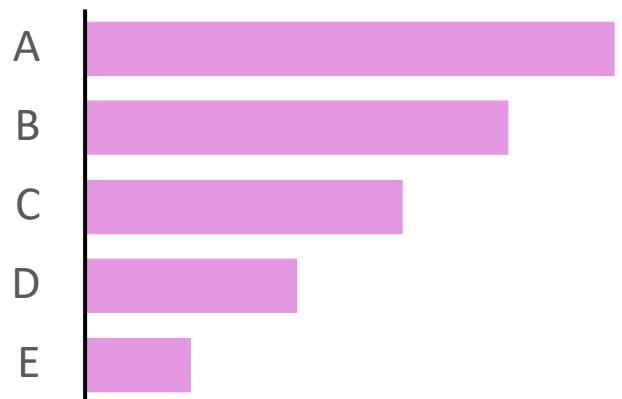
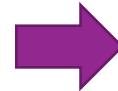
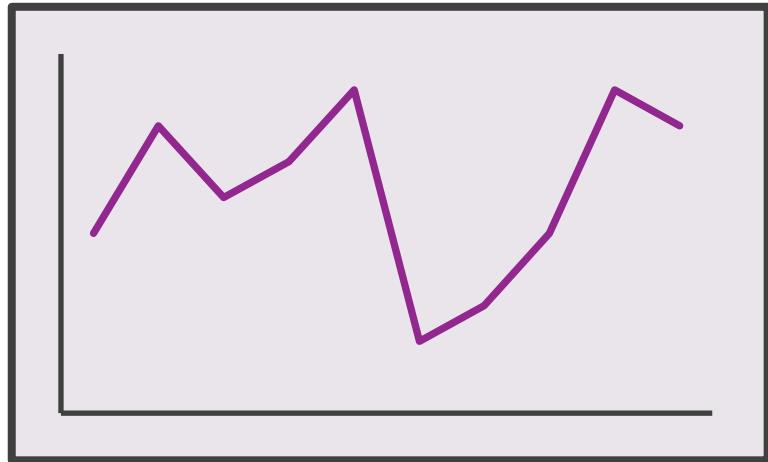
Common Fate



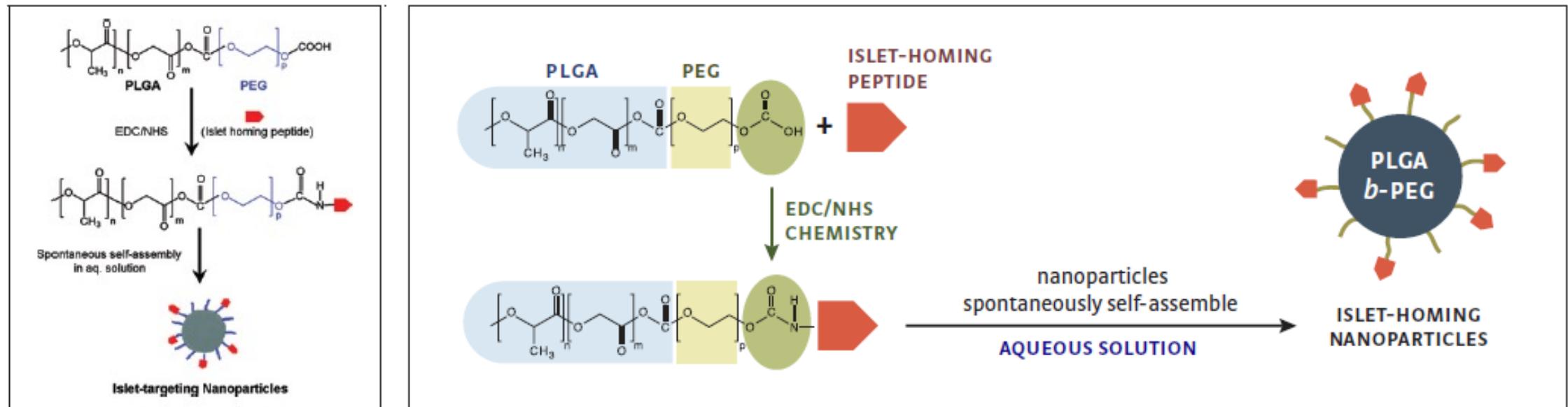
Simplicity



You can use the **Continuity** and **Closure** principle
to remove unnecessary elements in your chart



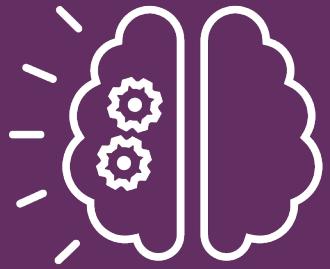
The Gestalt principles also apply to graphical abstracts



Shapes and colors used to identify similar structures.

But...

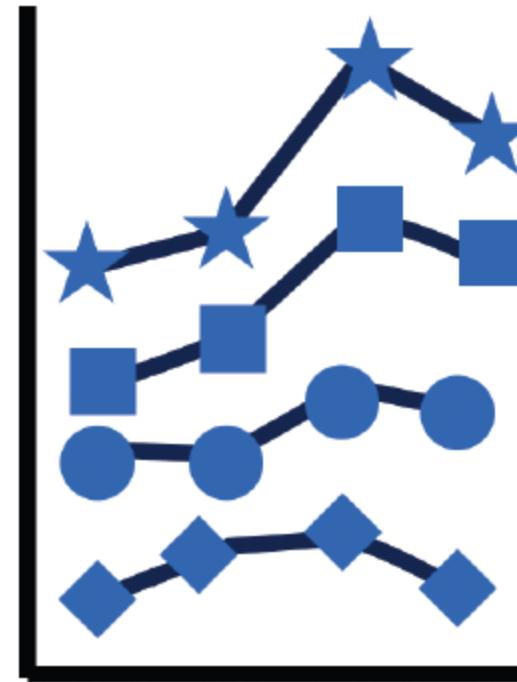
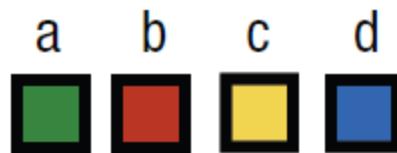
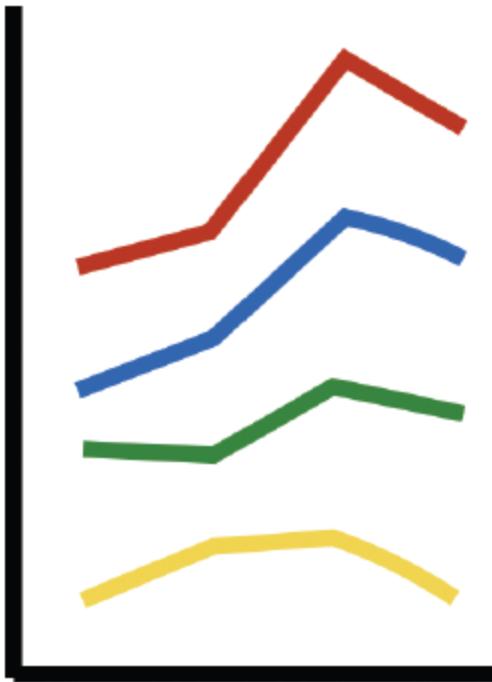
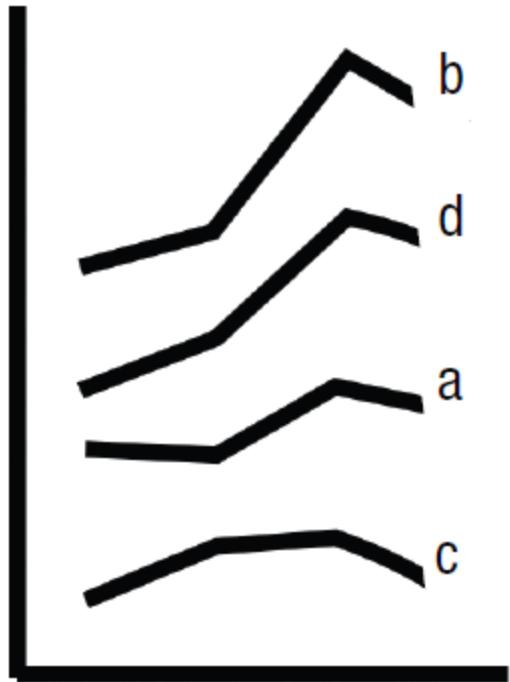




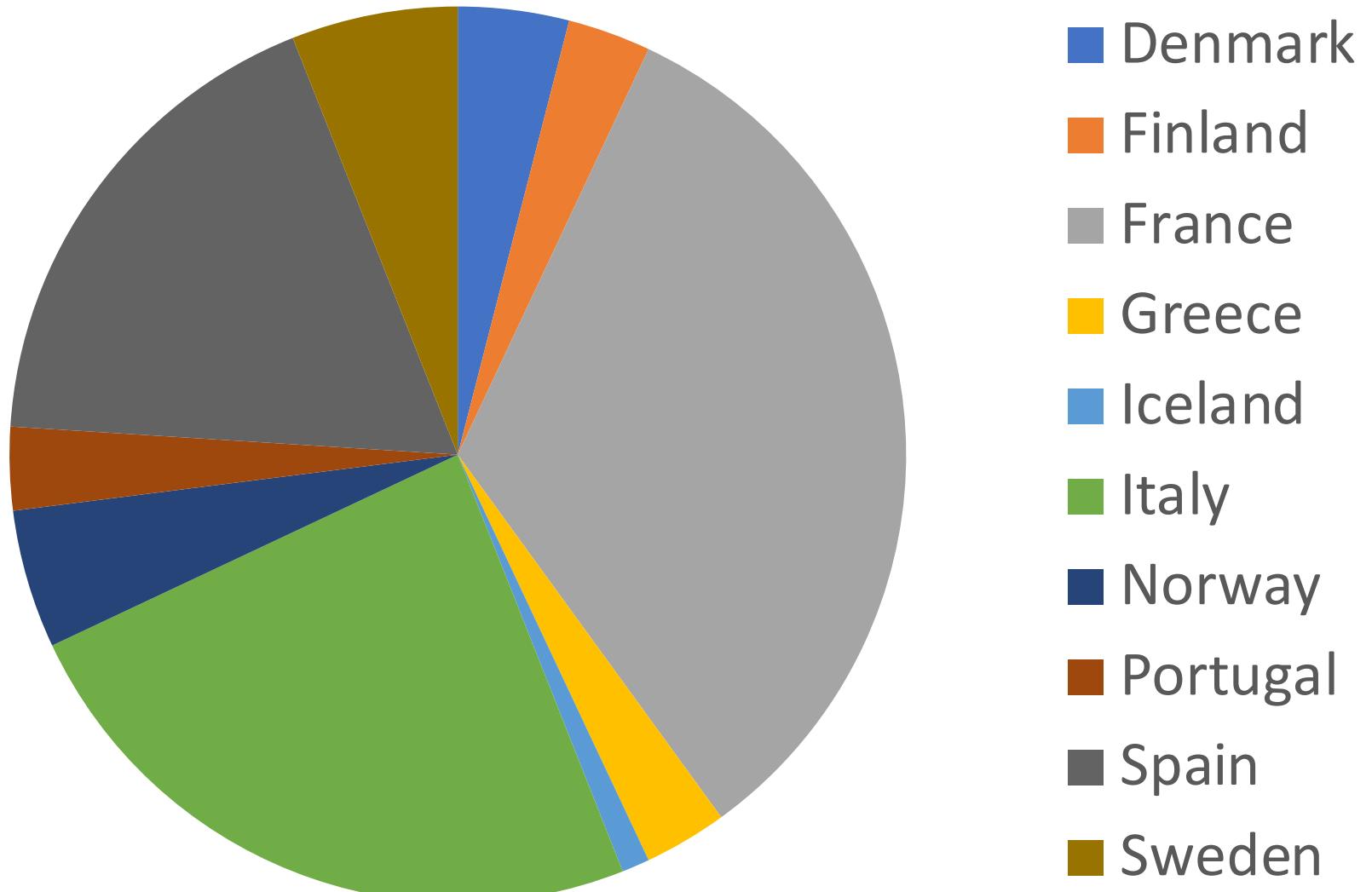
The weaknesses of our brain

“The eyes only see what the mind is prepared to comprehend.” – Henri Bergson.

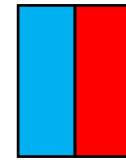
What is the order from top-to-bottom?



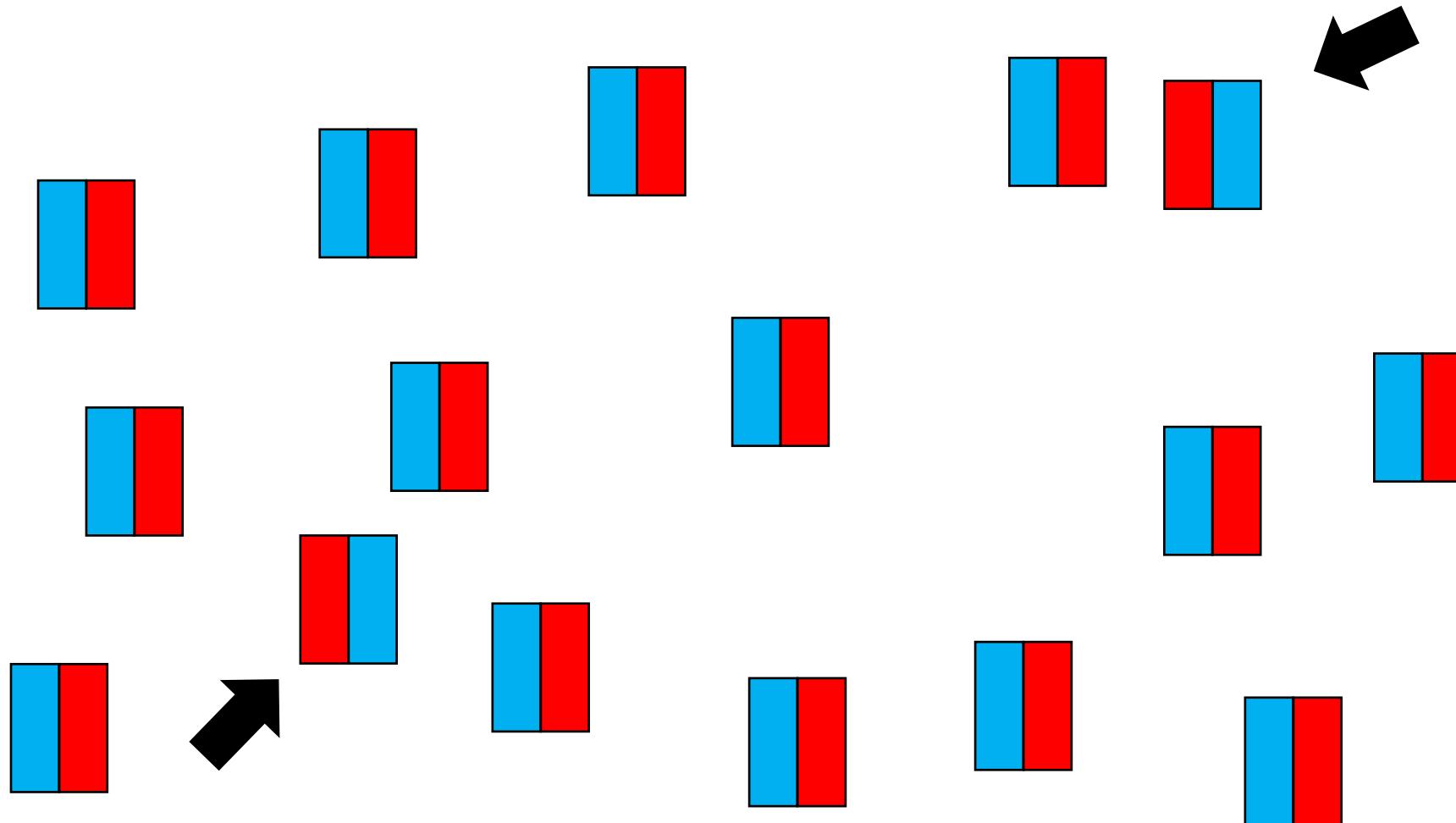
Norway has more widgets than Finland?



Comparing one set of colors is easy...



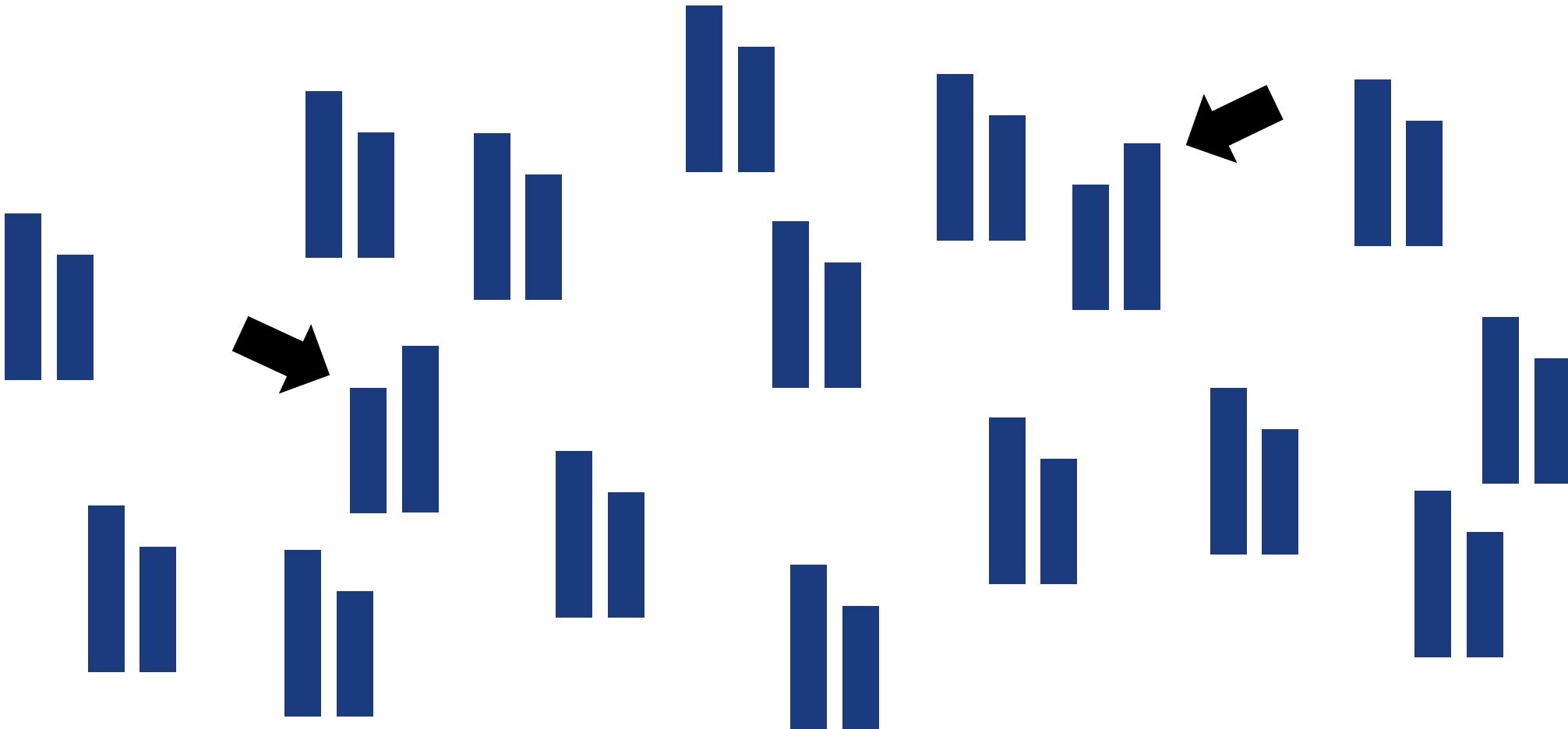
But: which one is not like the others?



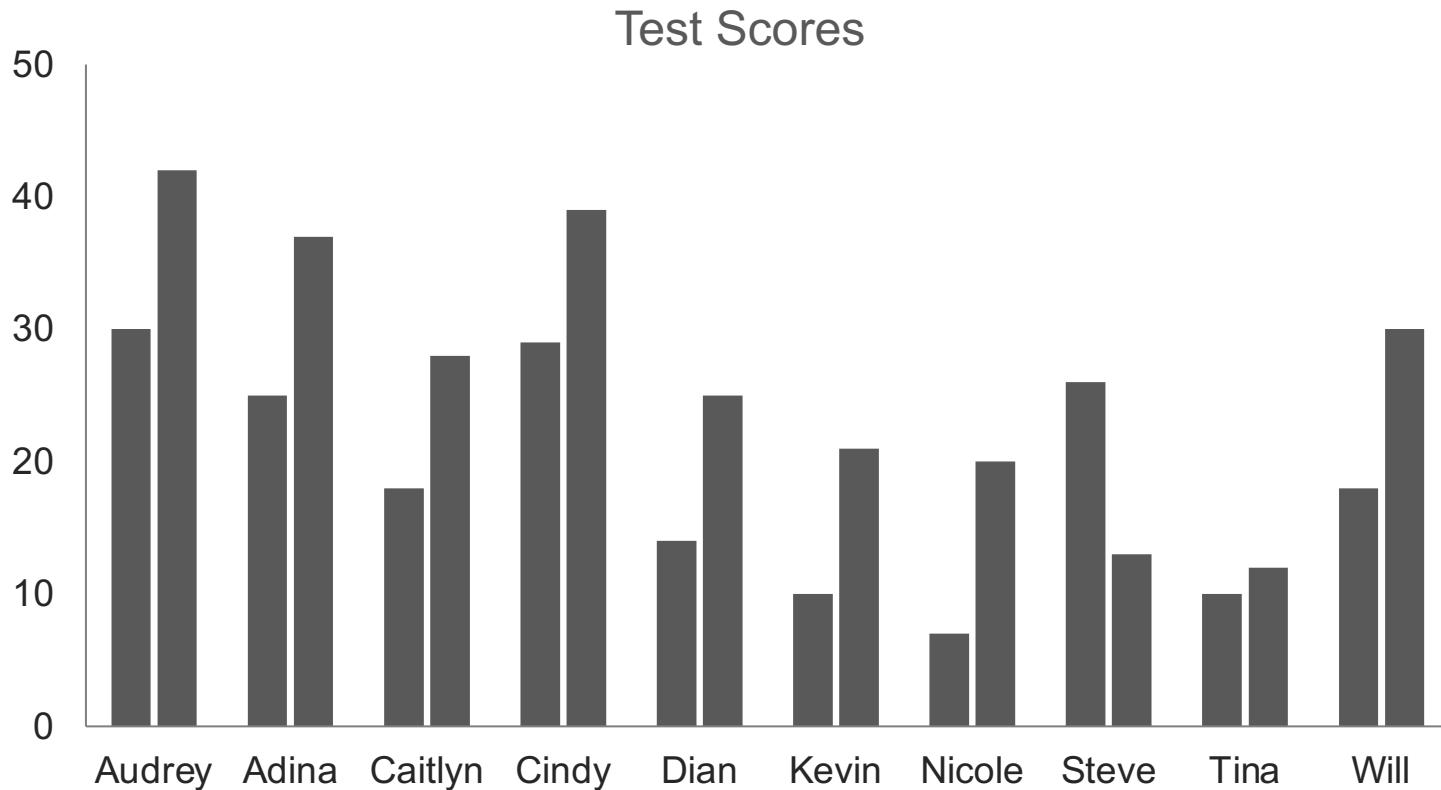
Do you see small-big or big-small?



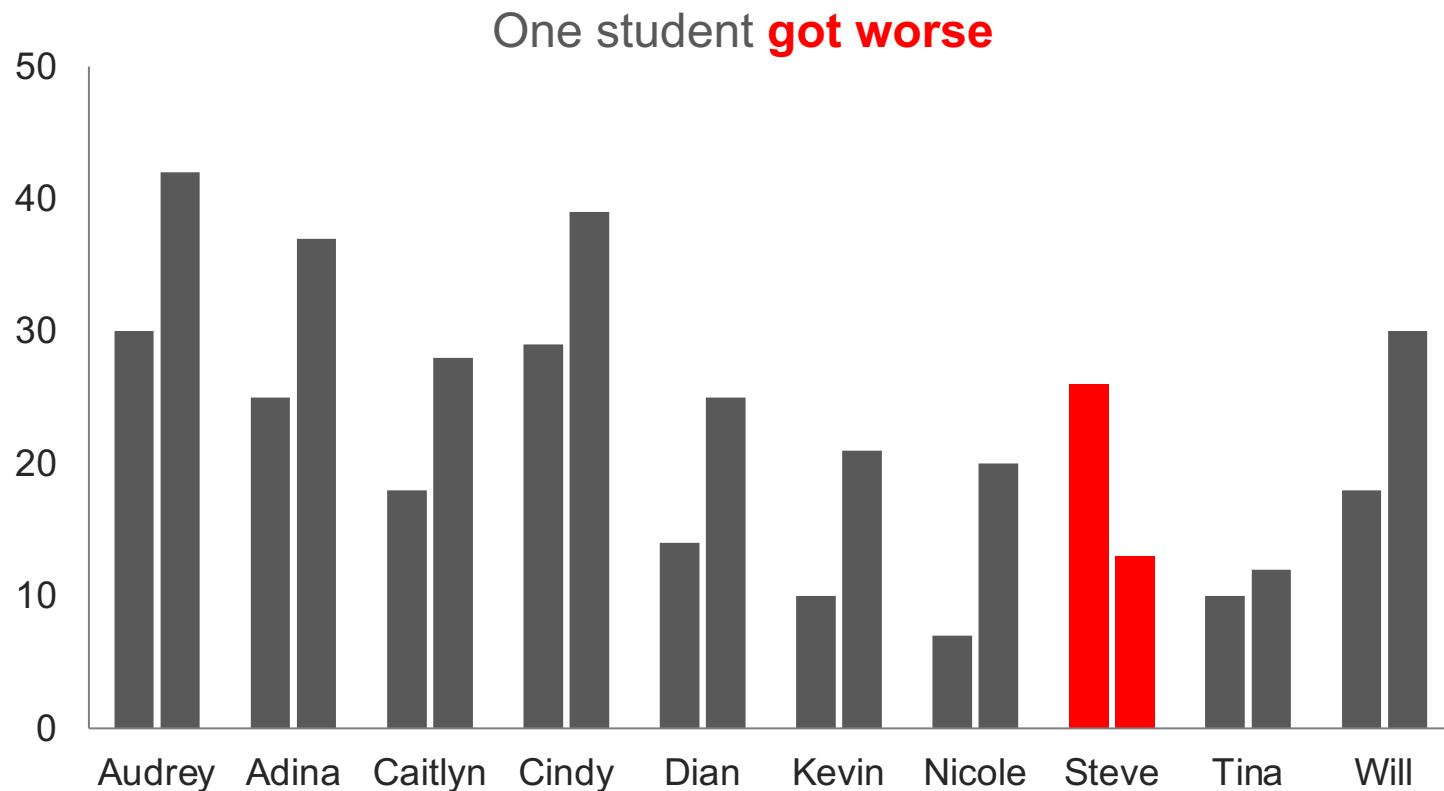
Which one is not like the others?



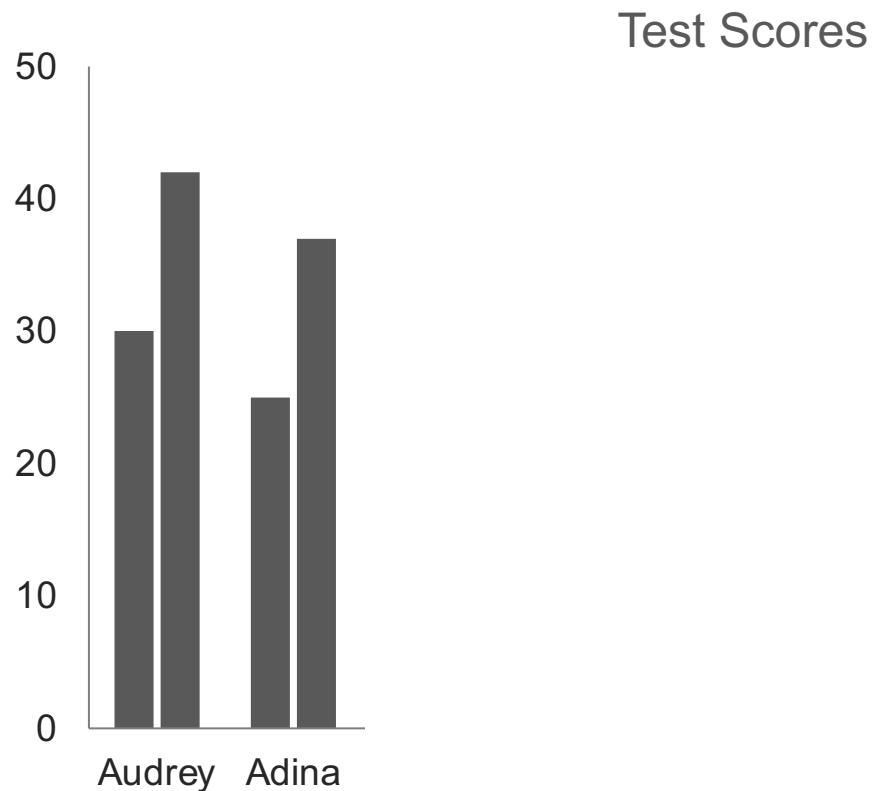
Which student got worse?



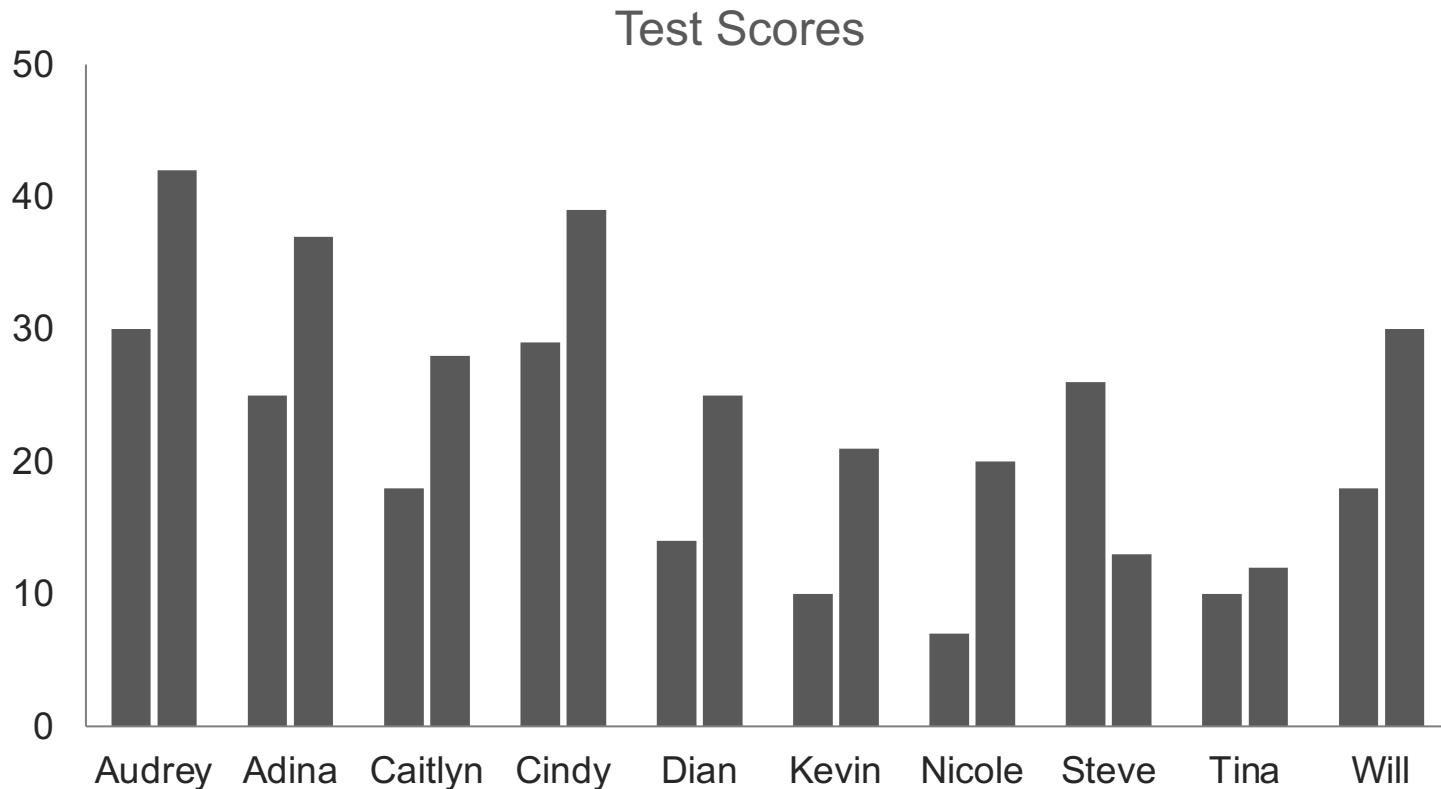
Which student got worse?



What is the average improvement?



What is the average improvement?



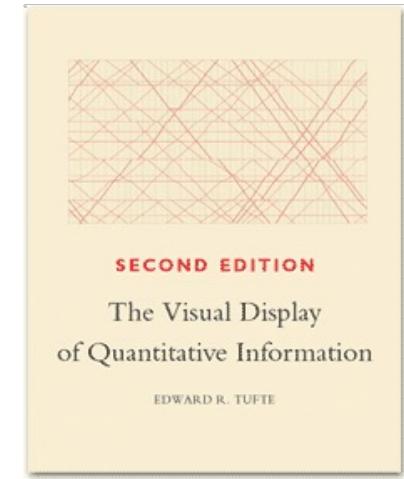
Where to look?



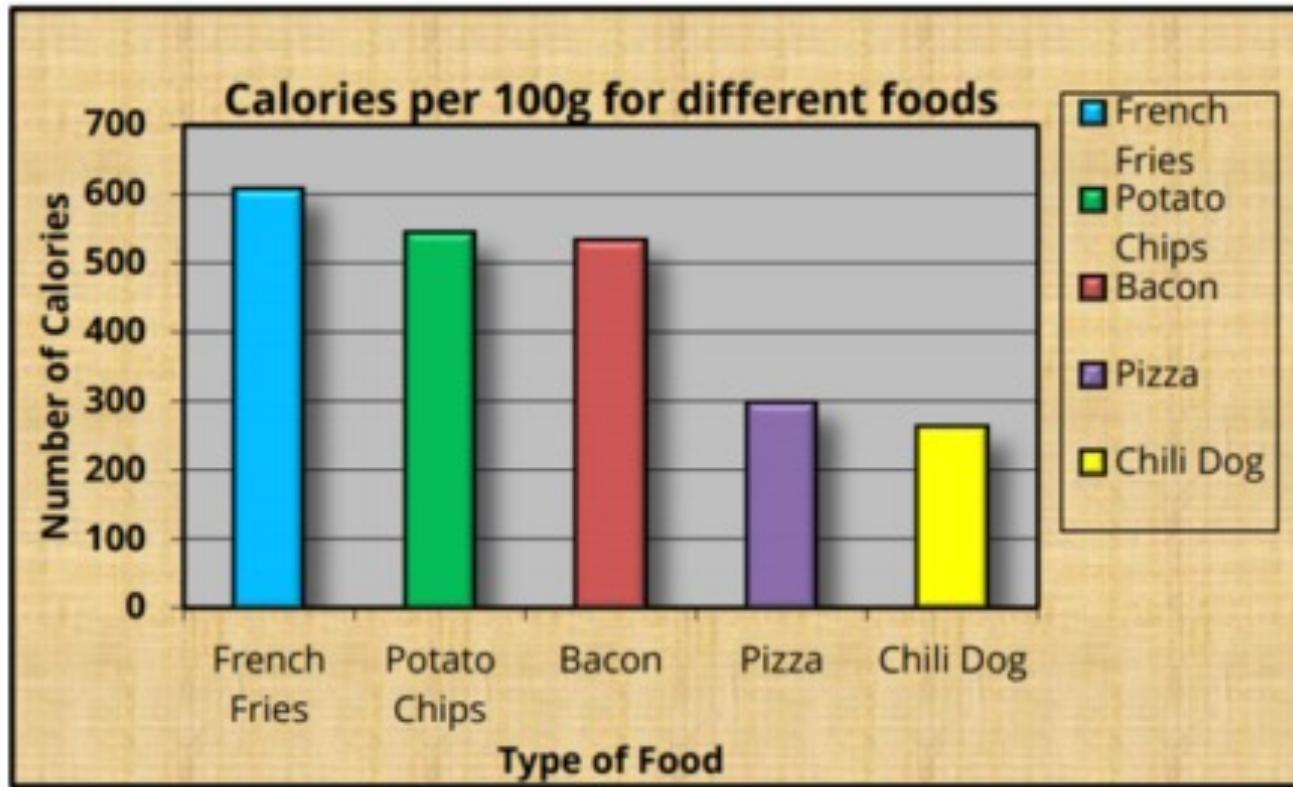
The data-ink ratio

Data-ink ratio =
$$\frac{\text{Data-ink}}{\text{Total ink used to print the graphic}}$$

- = proportion of a graphic's ink devoted to the non-redundant display of data-information
- = $1.0 - \text{proportion of a graphic that can be erased}$



Minimize distraction: maximize data-ink ratio



The data-ink ratio: remove to improve

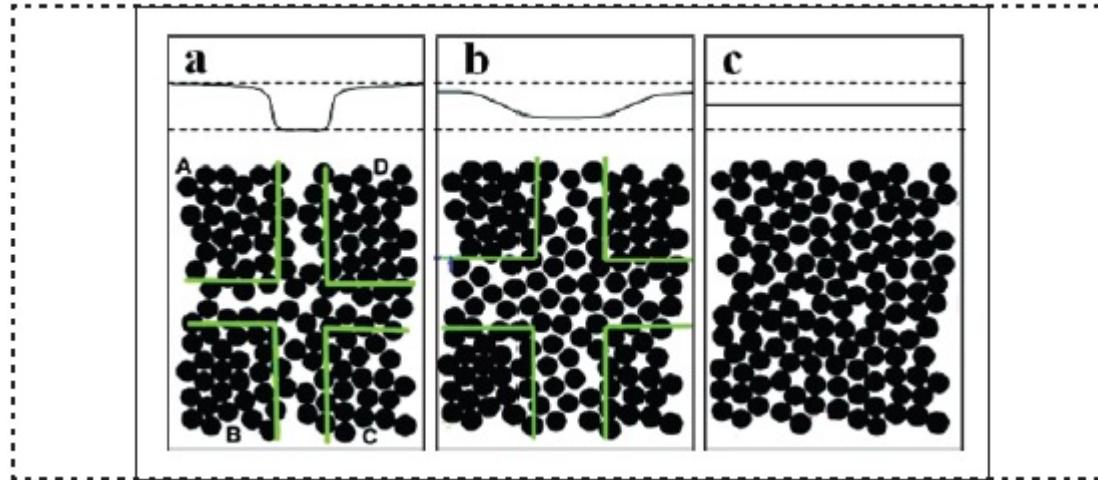
Remove
to improve
(the **data-ink** ratio)

Data-ink ratio also applies to tables

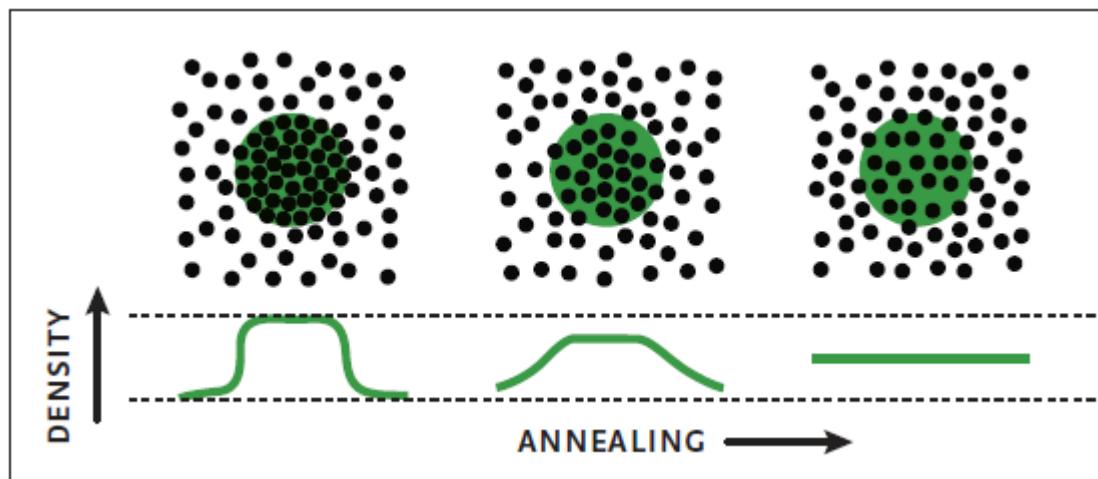
Student	Test1	Test2	Difference
Audrey	30.00	42.00	12.00 pts
Adina	25.00	37.00	12.00 pts
Caitlyn	18.00	28.00	10.00 pts
Cindy	29.00	39.00	10.00 pts
Dian	14.00	25.00	11.00 pts
Kevin	10.00	21.00	11.00 pts
Nicole	7.00	20.00	13.00 pts
Steve	26.00	13.00	-13.00 pts
Tina	10.00	12.00	2.00 pts
Will	18.00	30.00	12.00 pts

Student	Test1	Test2	Difference (pts)
Audrey	30	42	12
Adina	25	37	12
Caitlyn	18	28	10
Cindy	29	39	10
Dian	14	25	11
Kevin	10	21	11
Nicole	7	20	13
Steve	26	13	-13
Tina	10	12	2
Will	18	30	12

And to graphical abstracts



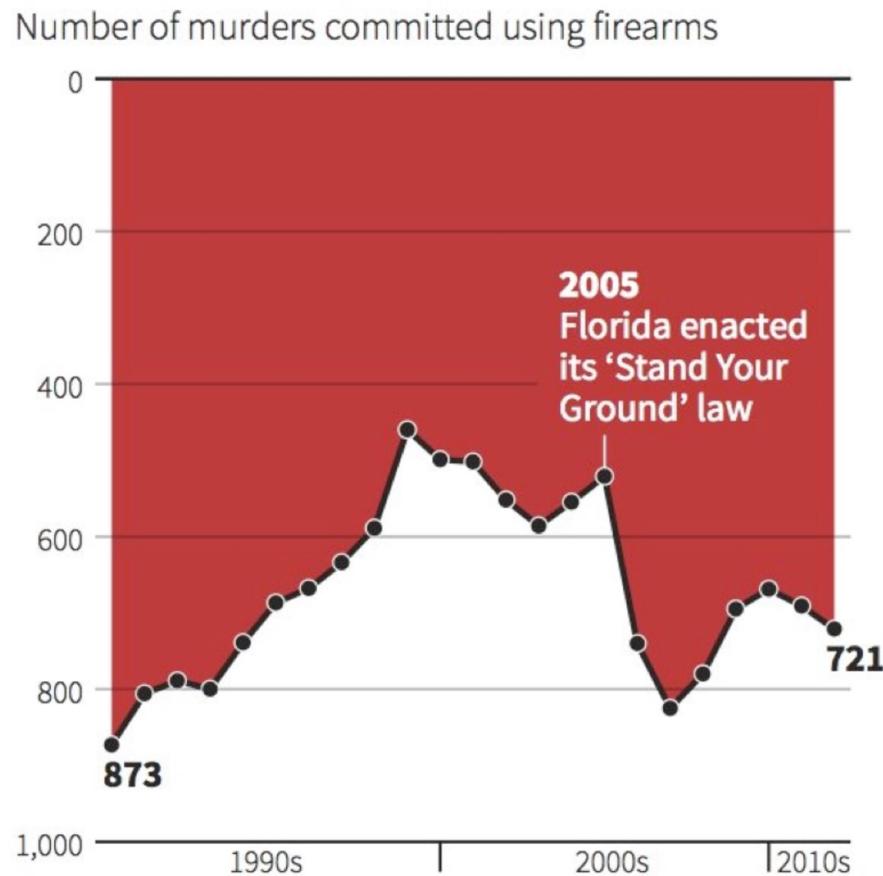
Too many interior boxes clutter the layout.



Boxes removed and arrows simplified.

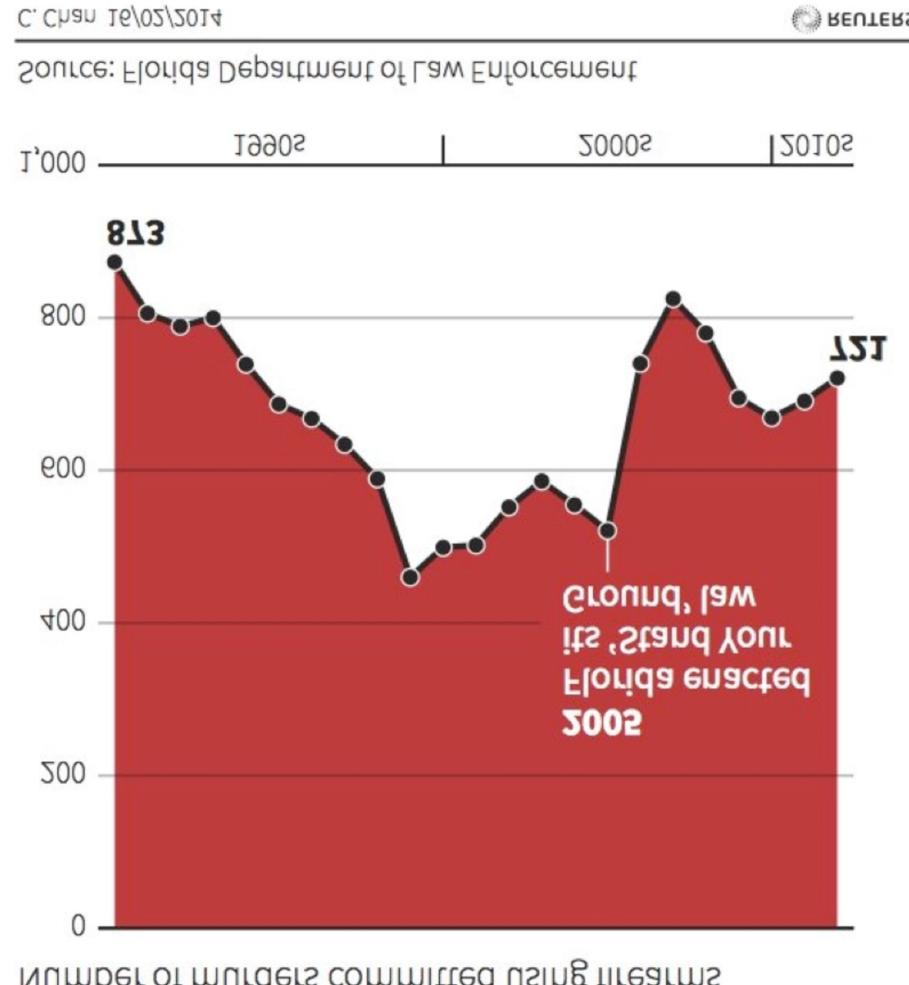
Downwards means...?

Gun deaths in Florida



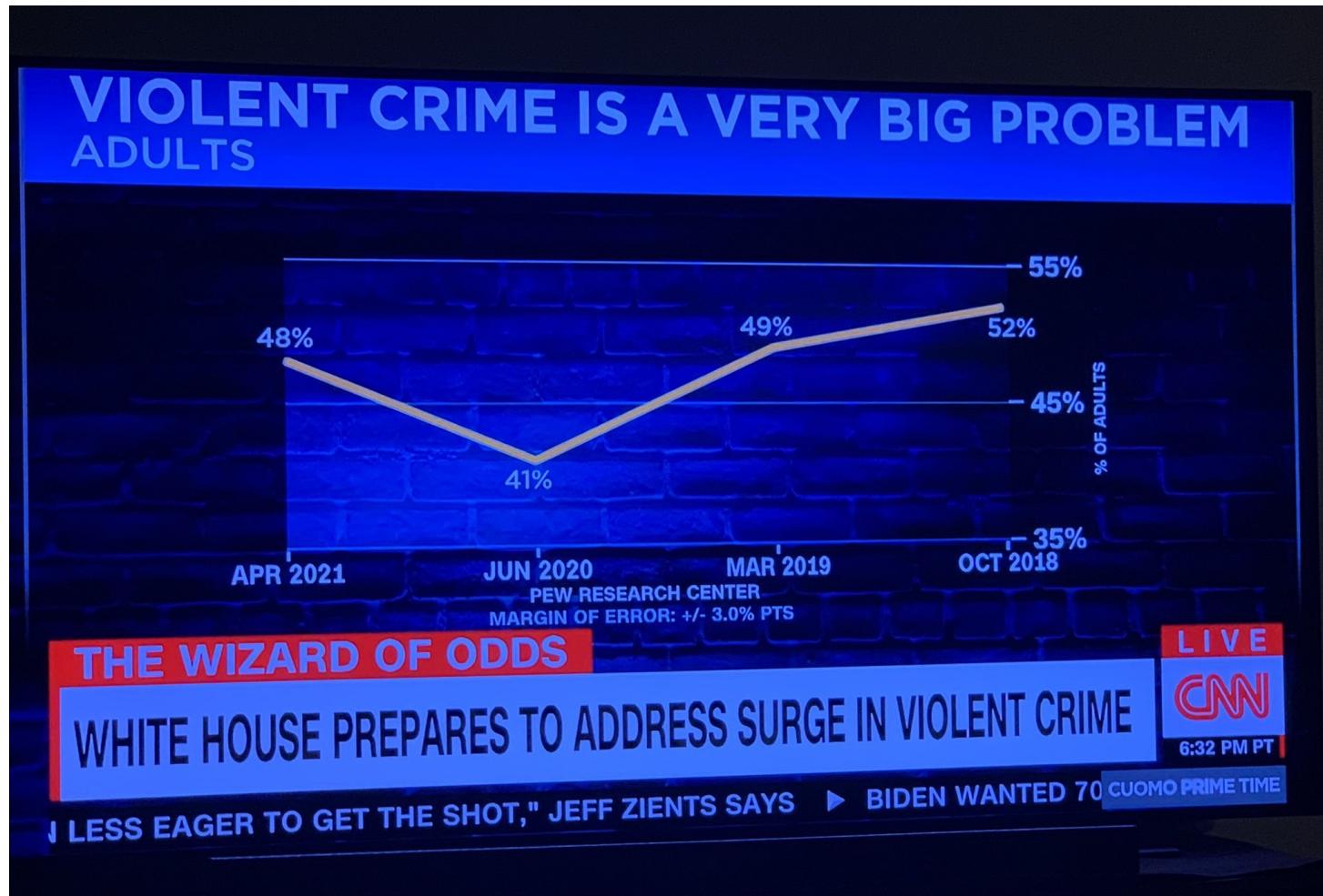
Source: Florida Department of Law Enforcement

Downwards means...?

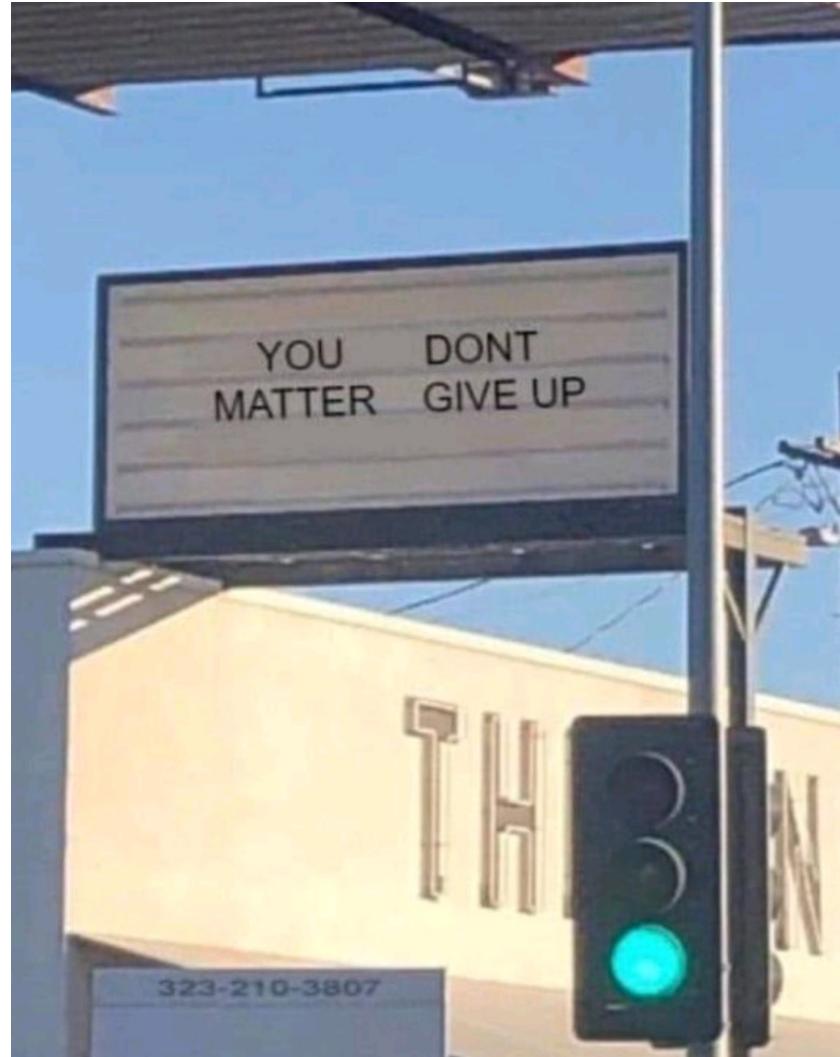


Eltona in safety

Don't go against our natural flow!

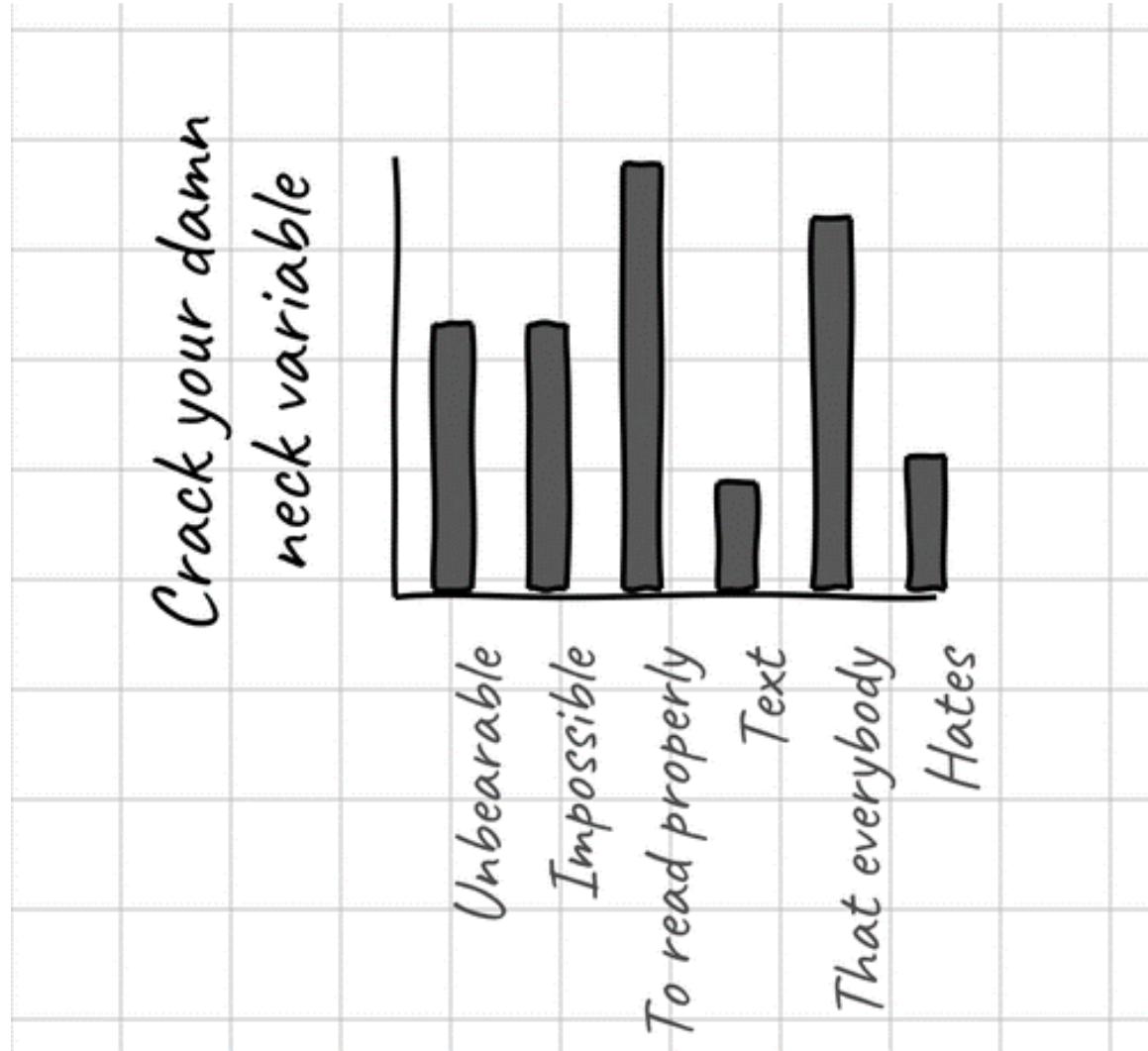


We read from left-to-right...

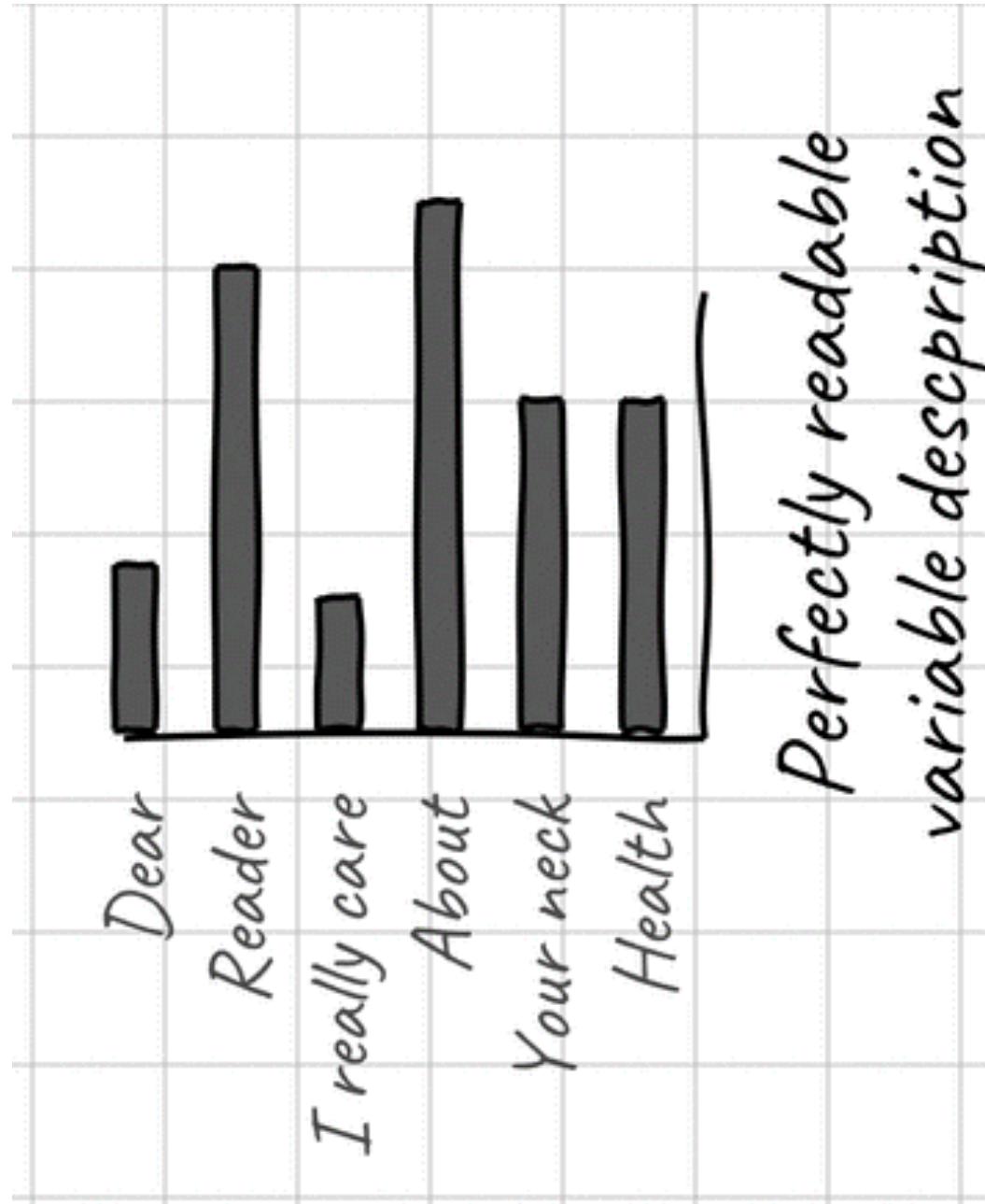


..right?

How's reading these labels?



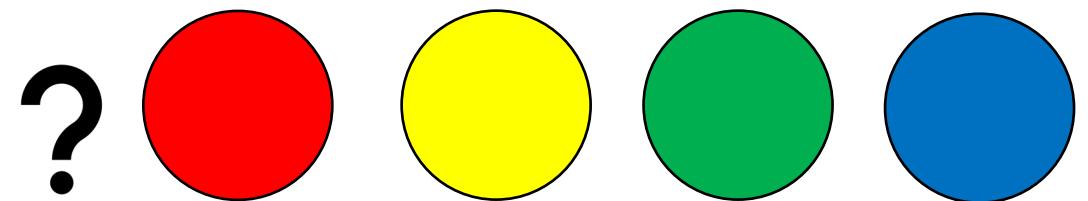
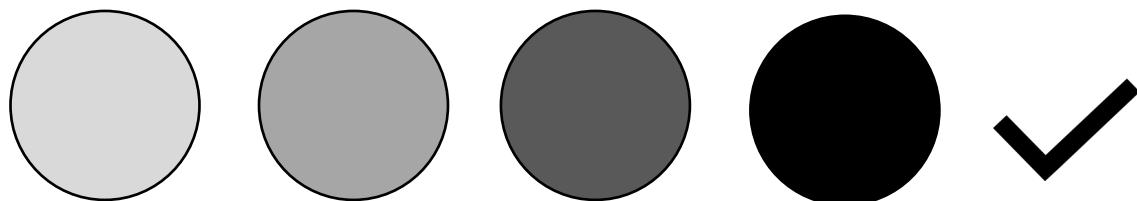
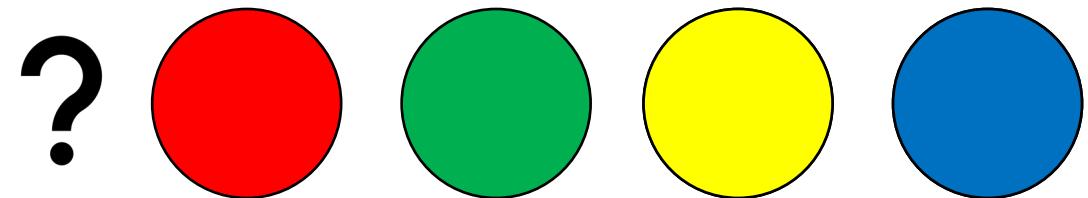
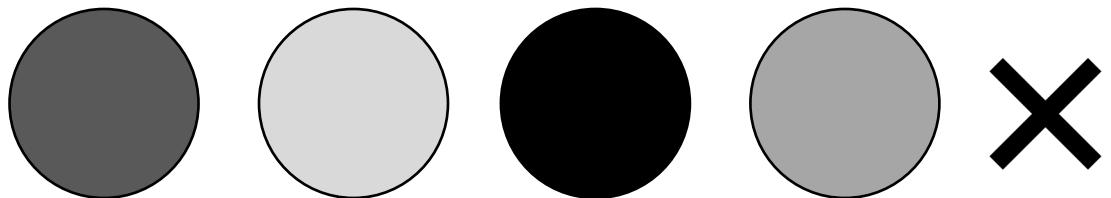
Better?



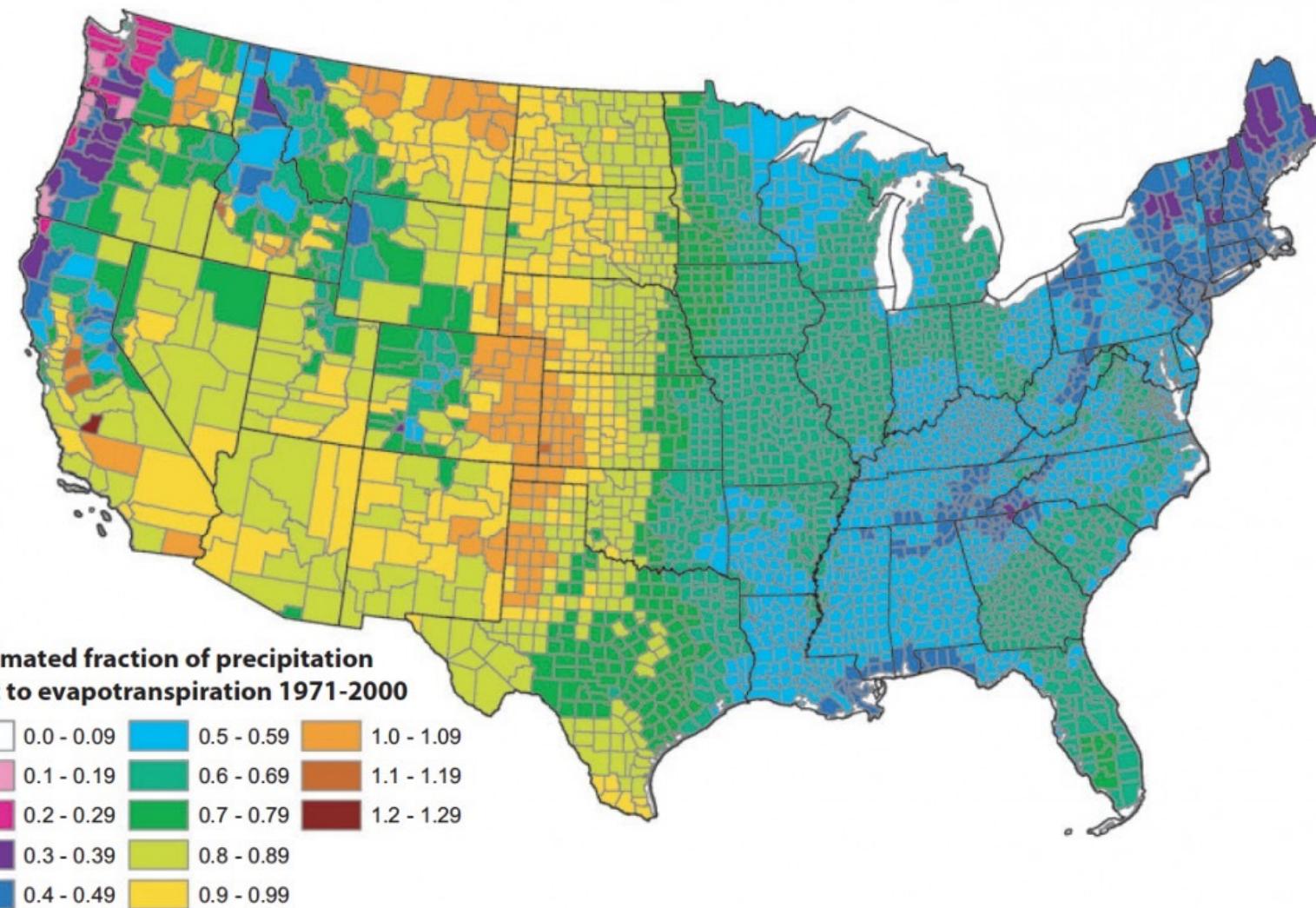
How to rank those circles from low to high?



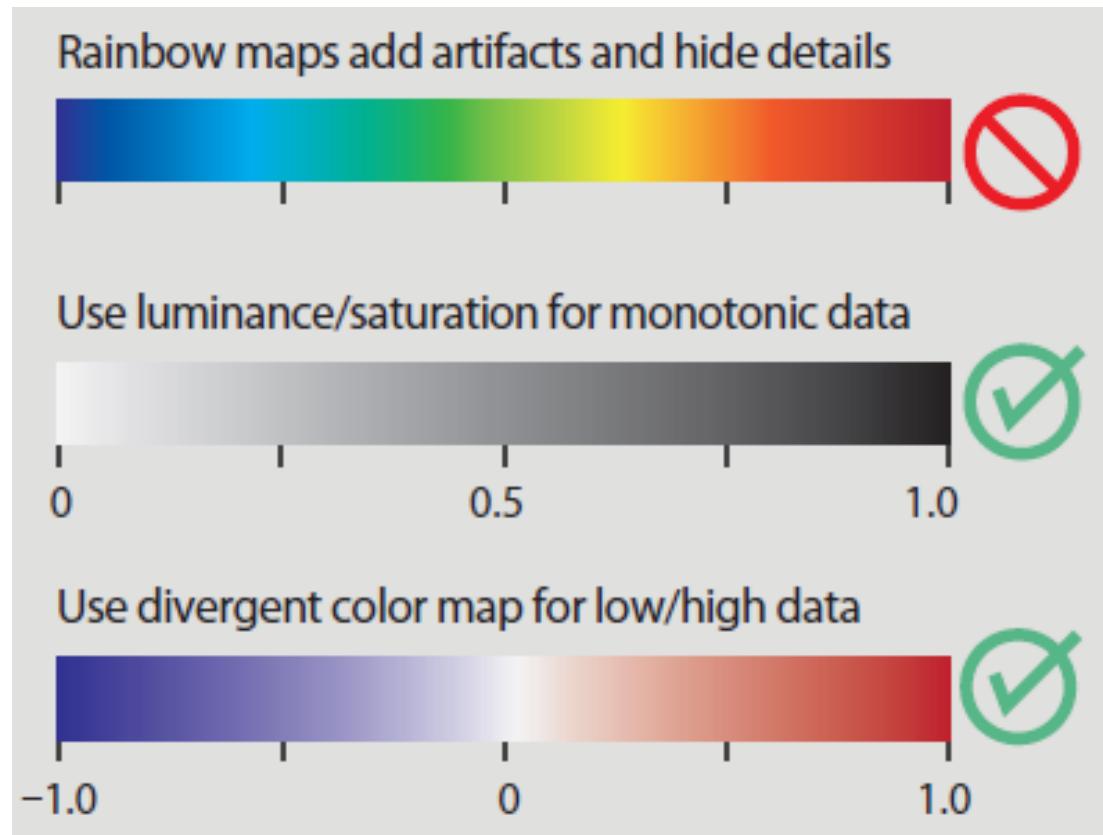
How to rank those circles from low to high?



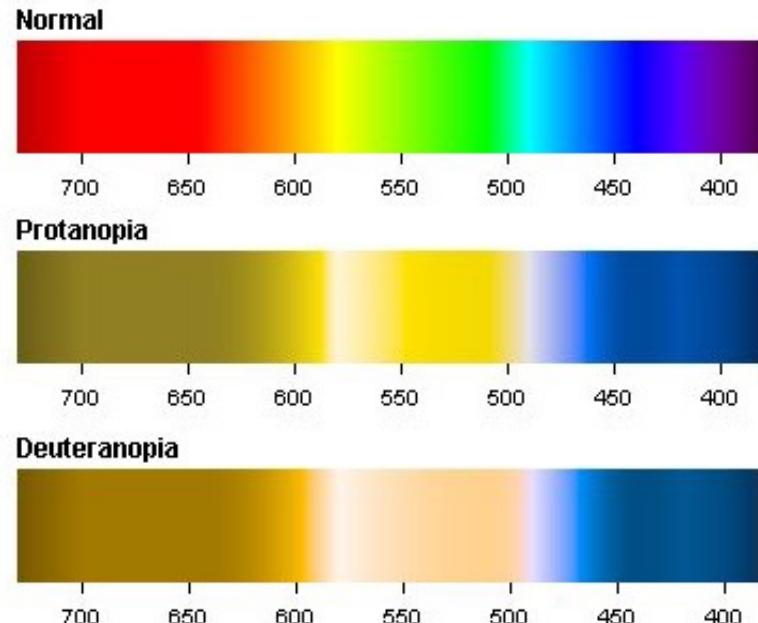
Avoid rainbow color maps for quantities!



Use a monochromatic or a diverging color scale instead

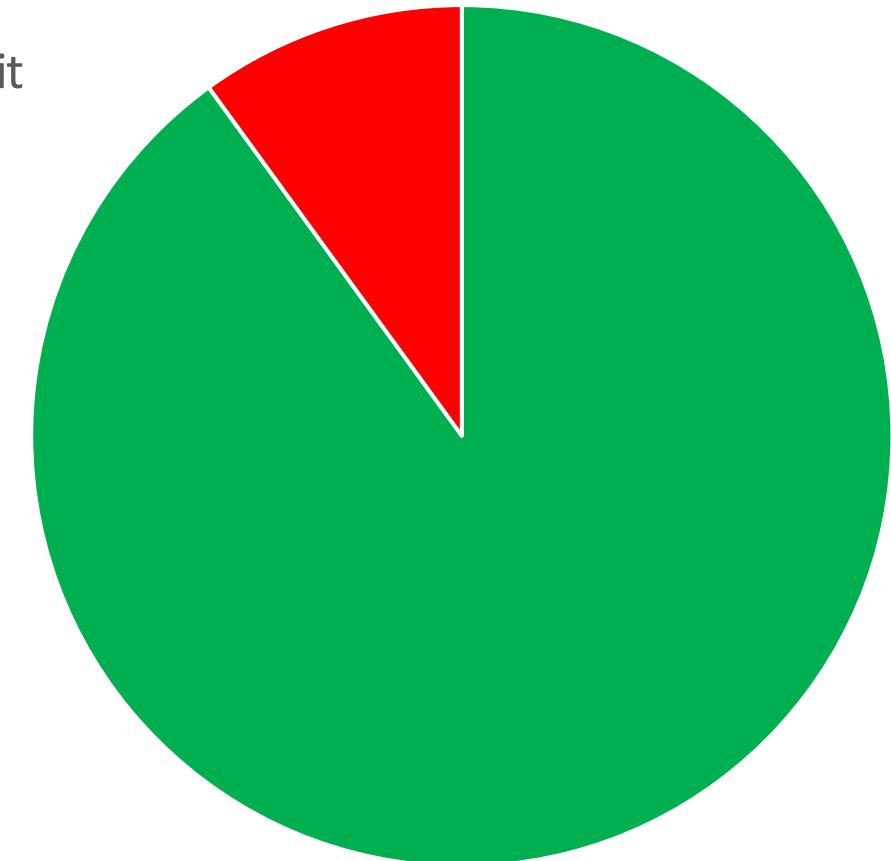


Another issue: color vision deficiency



Red-green color blindness in men

- Don't have it
- Have it



Color issues are common in science communications!

Hydrol. Earth Syst. Sci.

797 papers from 2005-2020

Papers with color issues
(total % per journal)



Nature Scientific Reports

Top 100 papers from 2019



Nature Communications

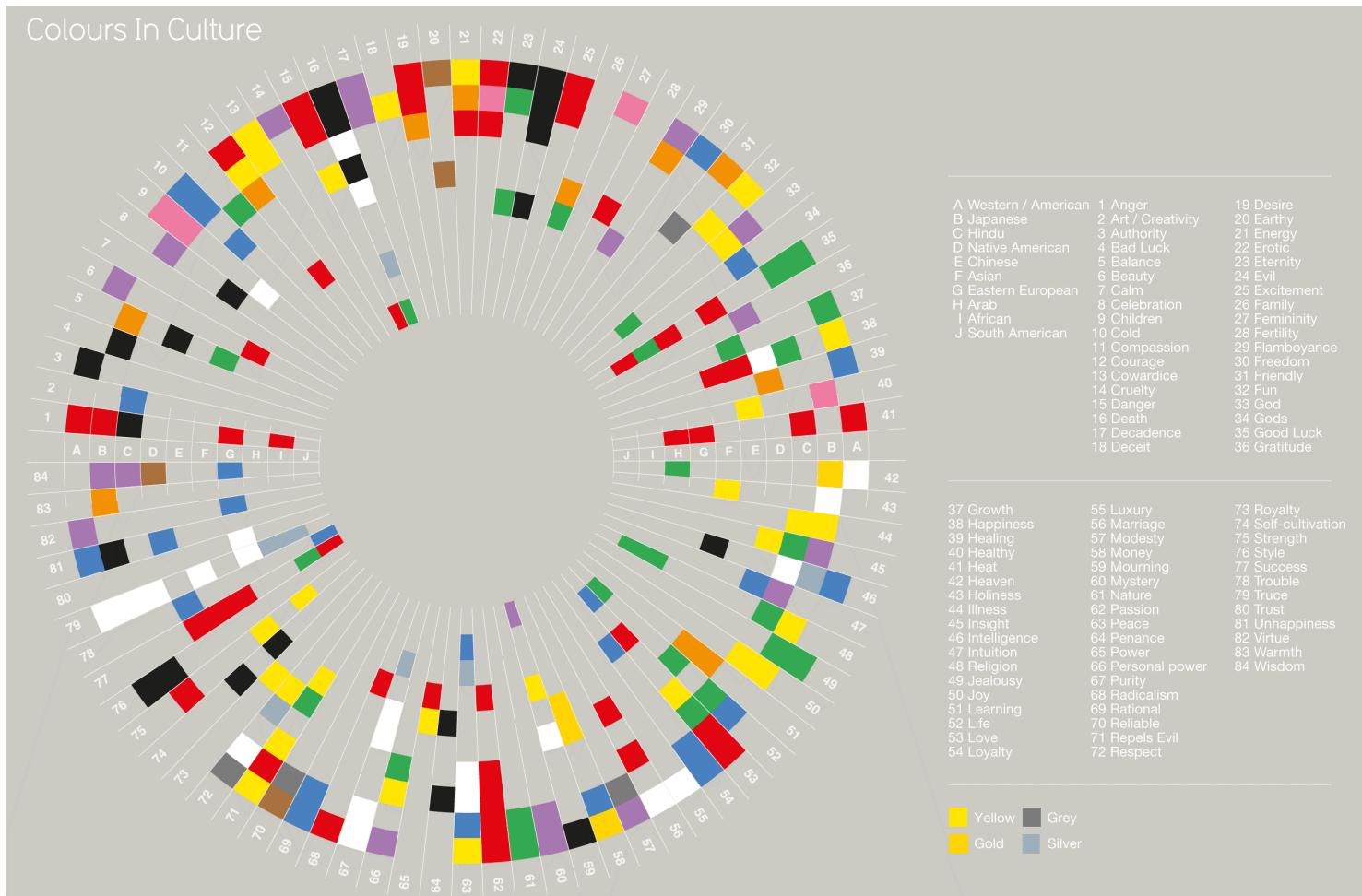
100 papers from 4-6 Nov 2020



Lastly, color is associated to feelings...

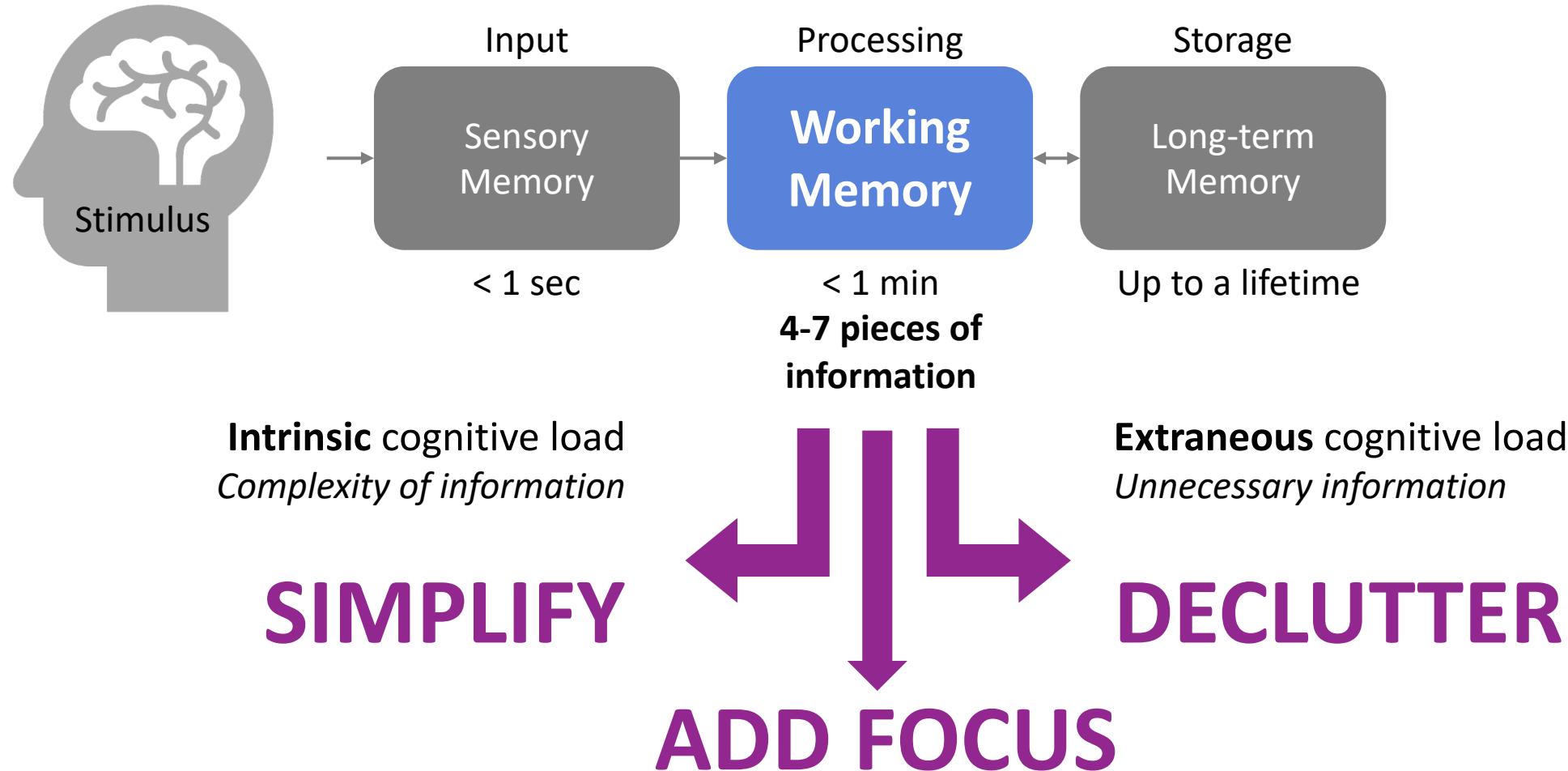


...and differs between cultures



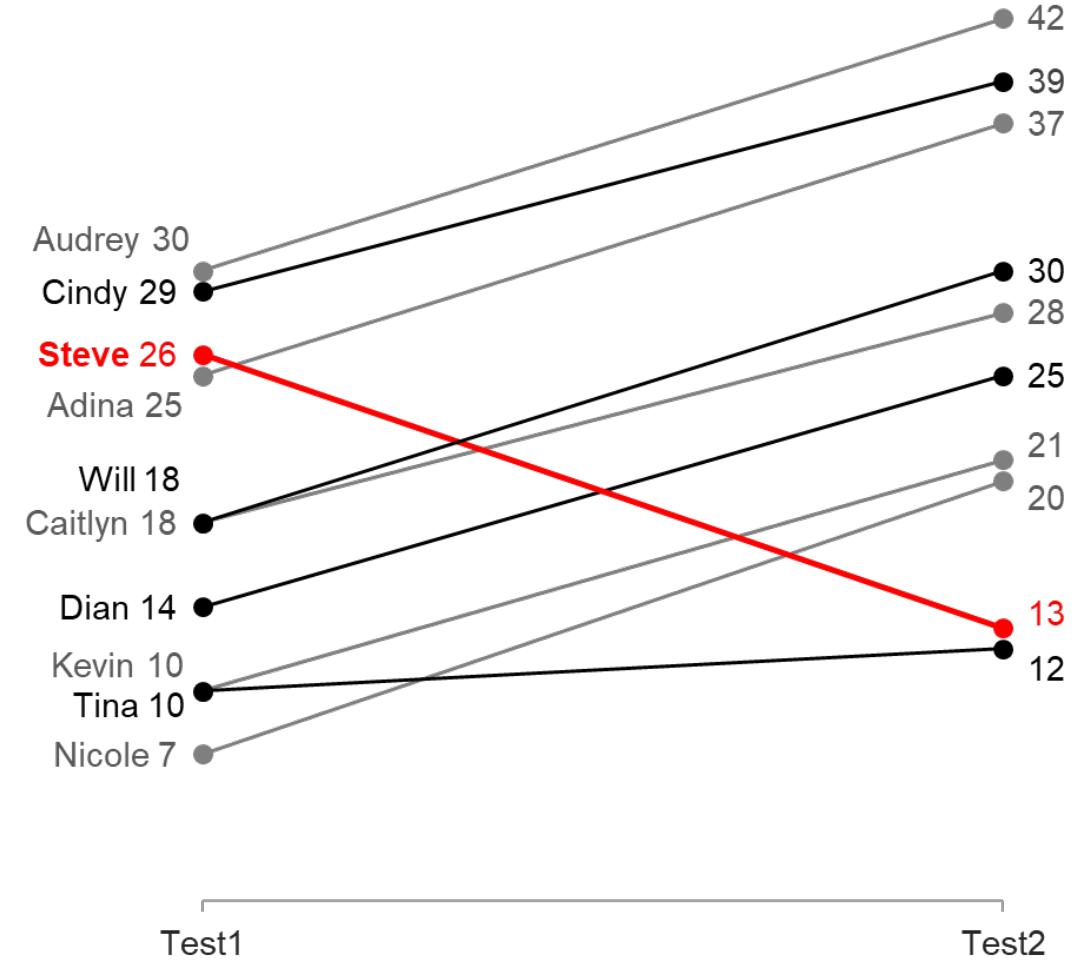
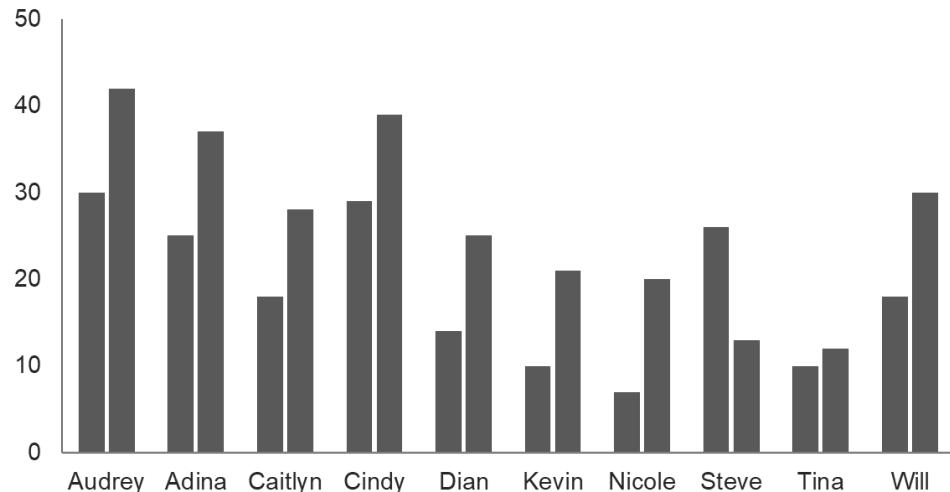
Putting it all together

Optimize visual perception: minimize *intrinsic* and *extraneous cognitive load*

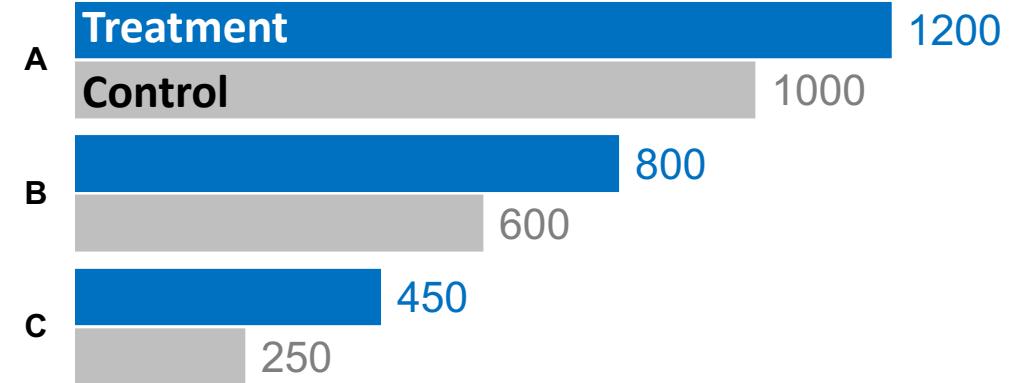
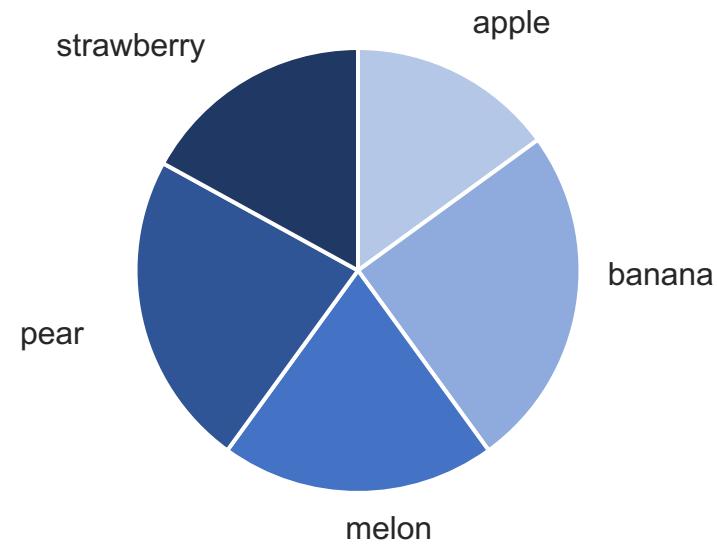


Simplify: Enable direct comparisons

Which student got worse?

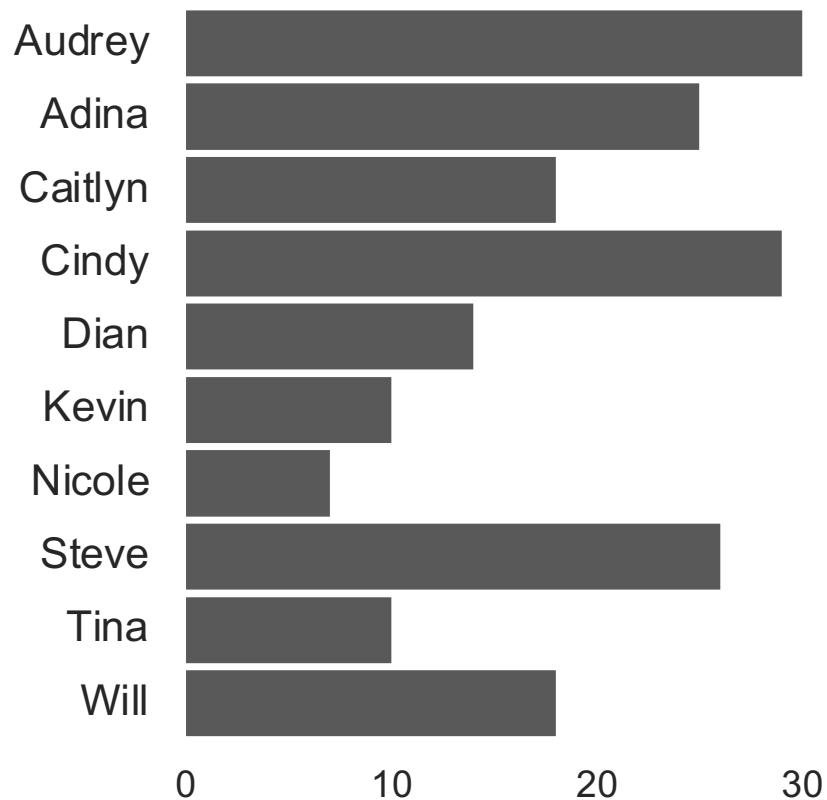


Simplify: Label your data directly

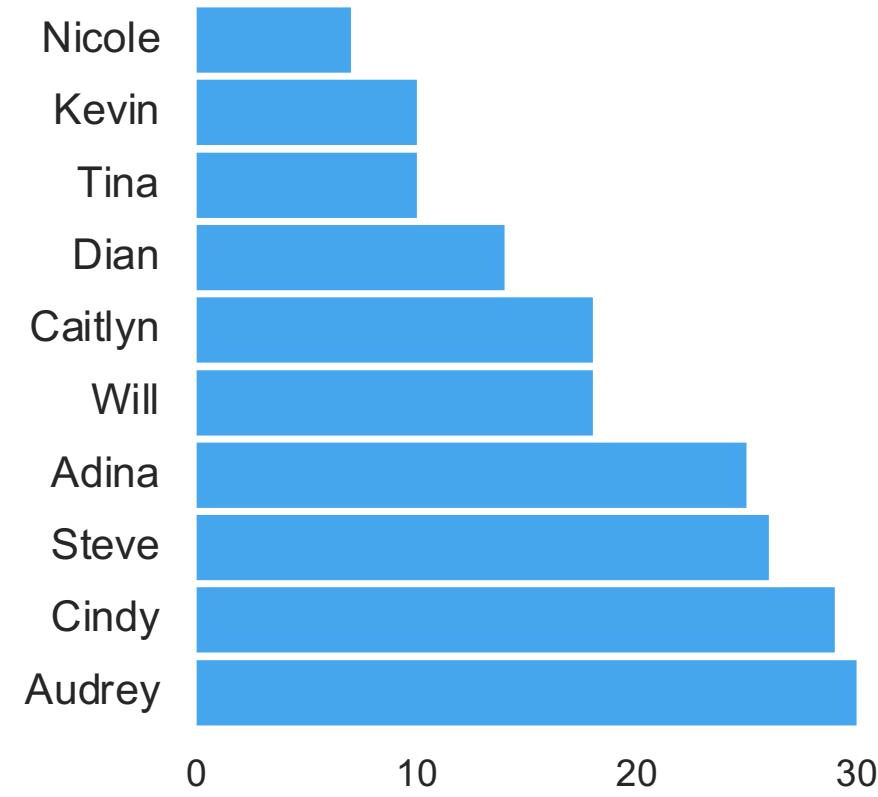


Simplify: Order your data logically, e.g., alphabetically or by number

Student's scores for test 1



Student's scores for test 1



Declutter: Minimize distractions



Effective. Not optimal.

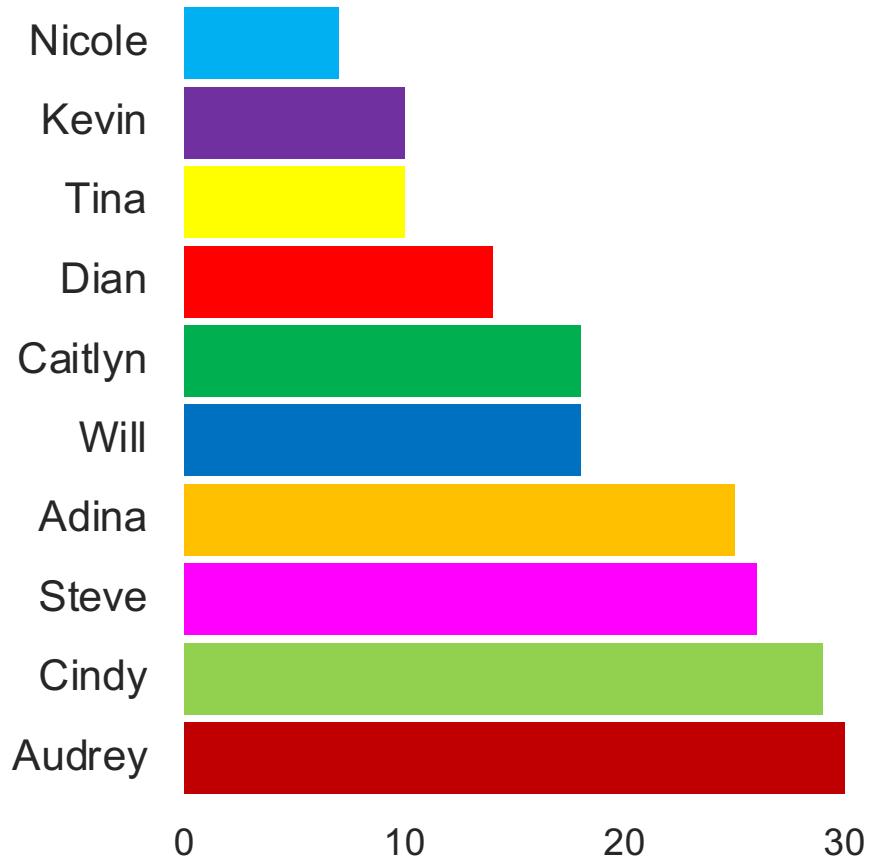


Simple, effective, optimal.



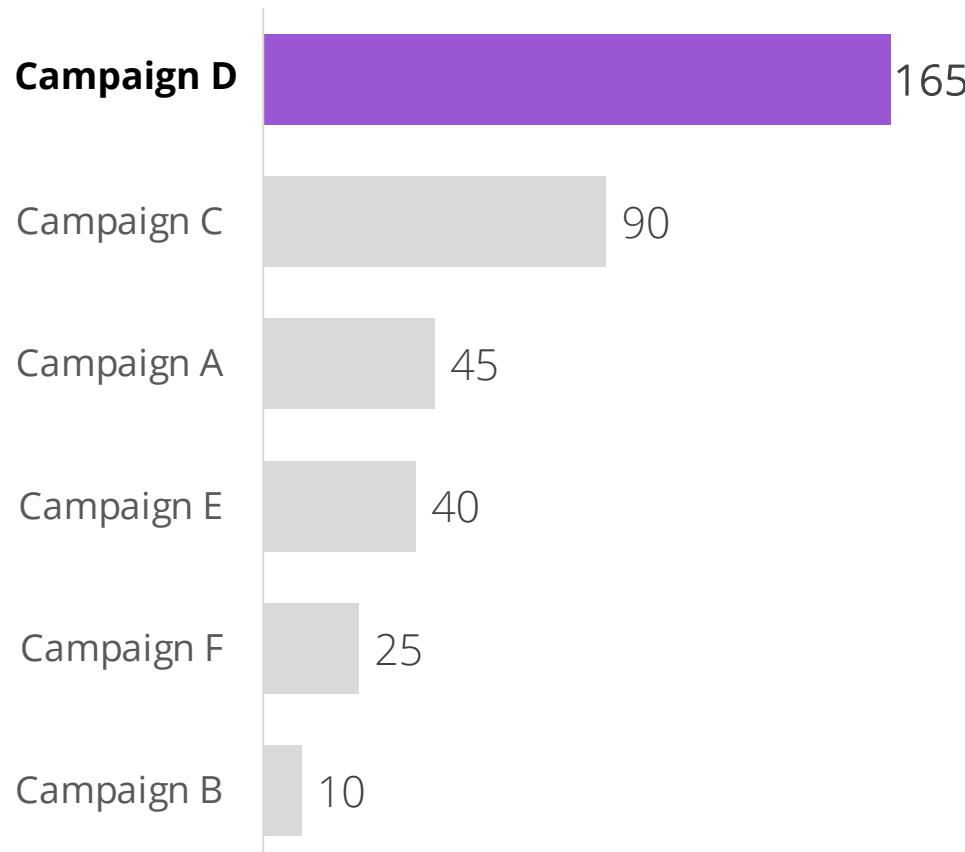
Bad.

Declutter: Avoid too many (useless) colors

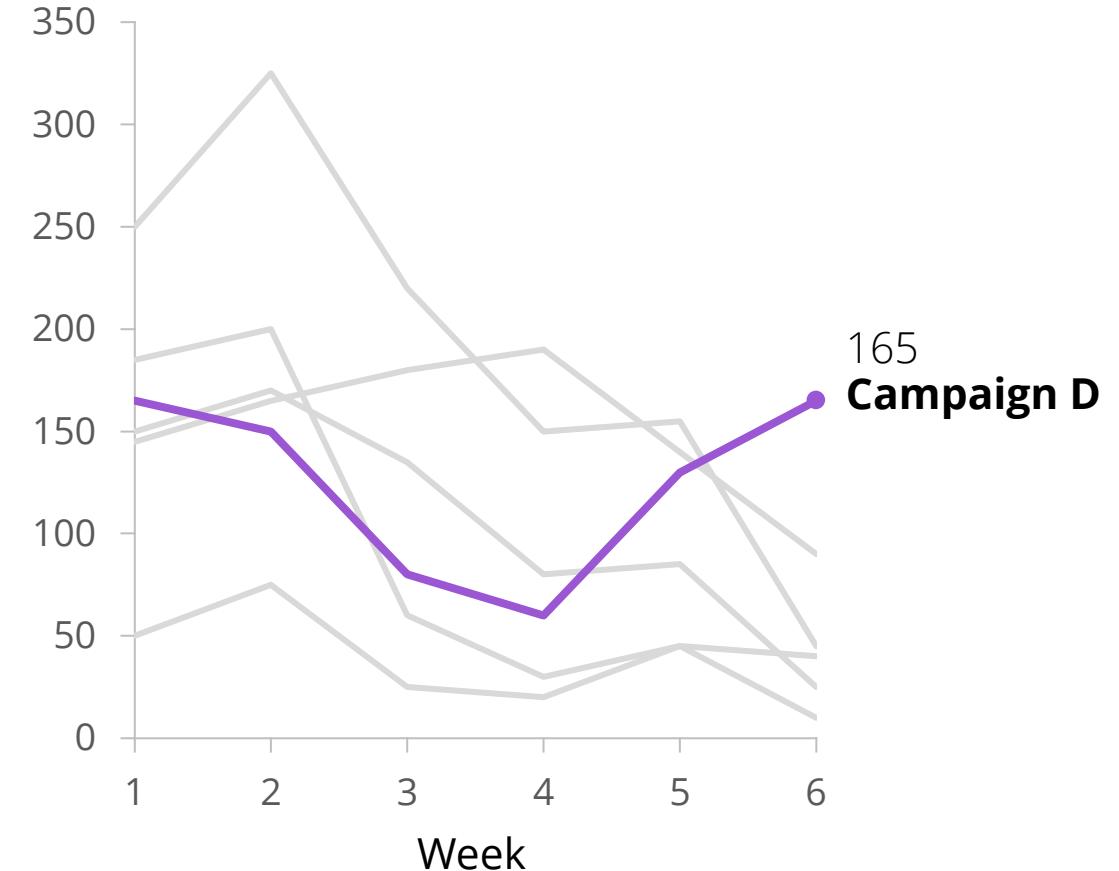


Add focus: Highlight what's important

Leads Generated by Campaign in Sixth Week

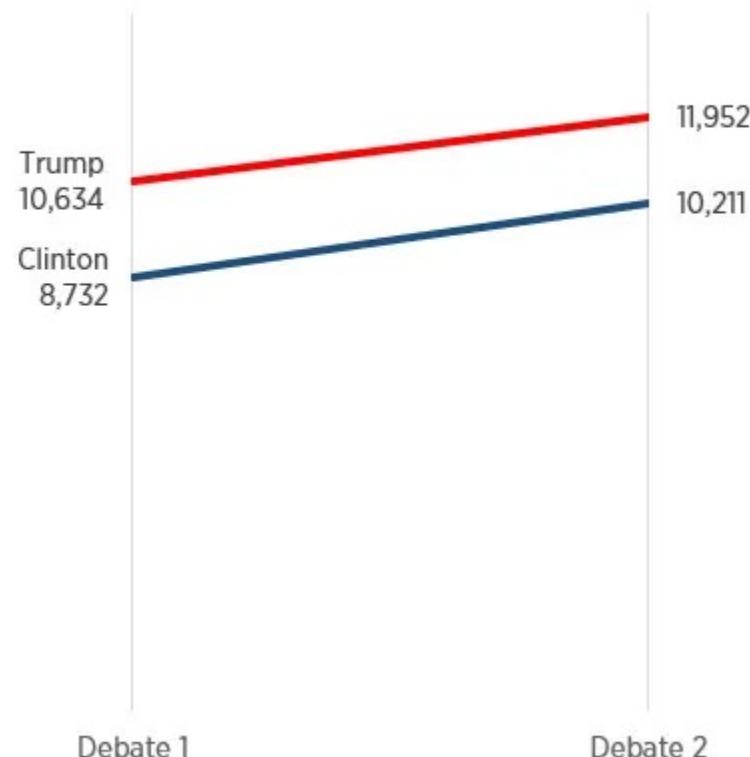


Leads Generated by Campaign

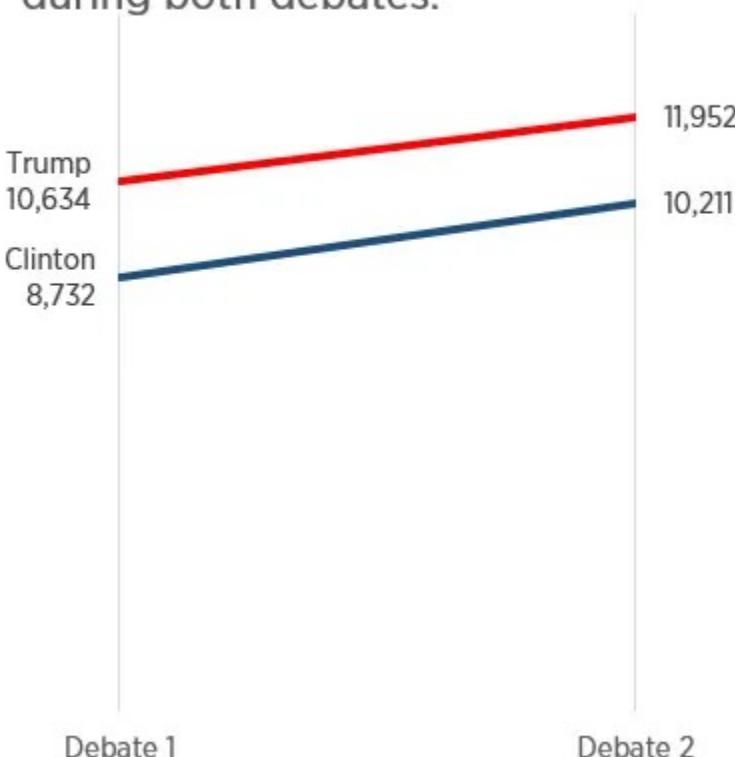


Add focus: Describe ≠ Judge

Words Spoken at Debate

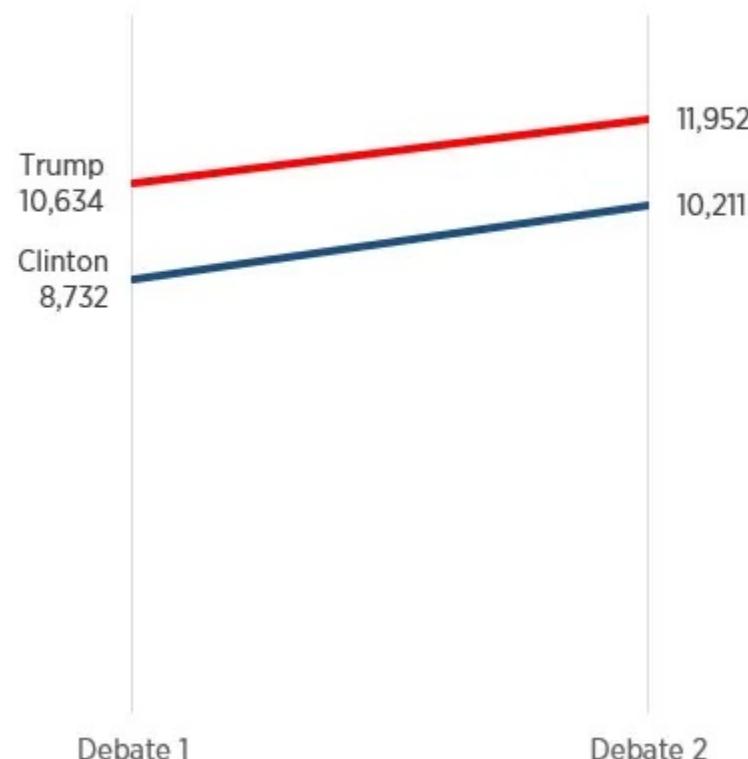


Trump spoke more words during both debates.

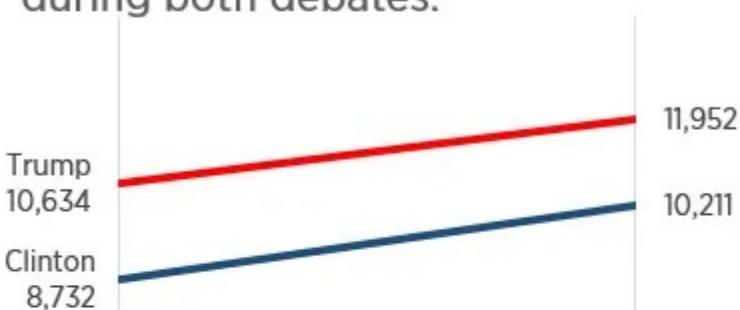


Add focus: Describe ≠ Judge

Words Spoken at Debate

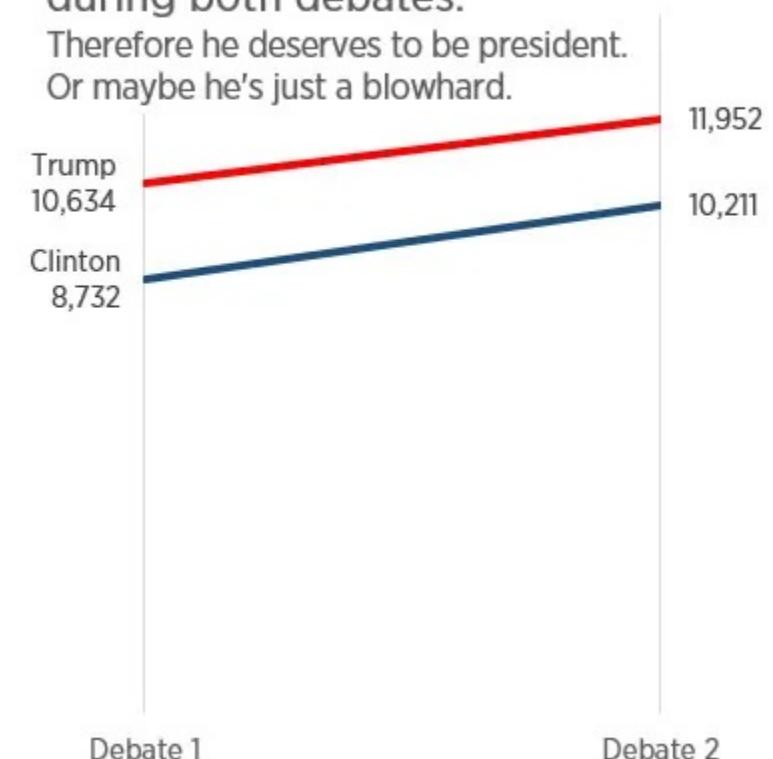


Trump spoke more words during both debates.



Trump spoke more words during both debates.

Therefore he deserves to be president.
Or maybe he's just a blowhard.



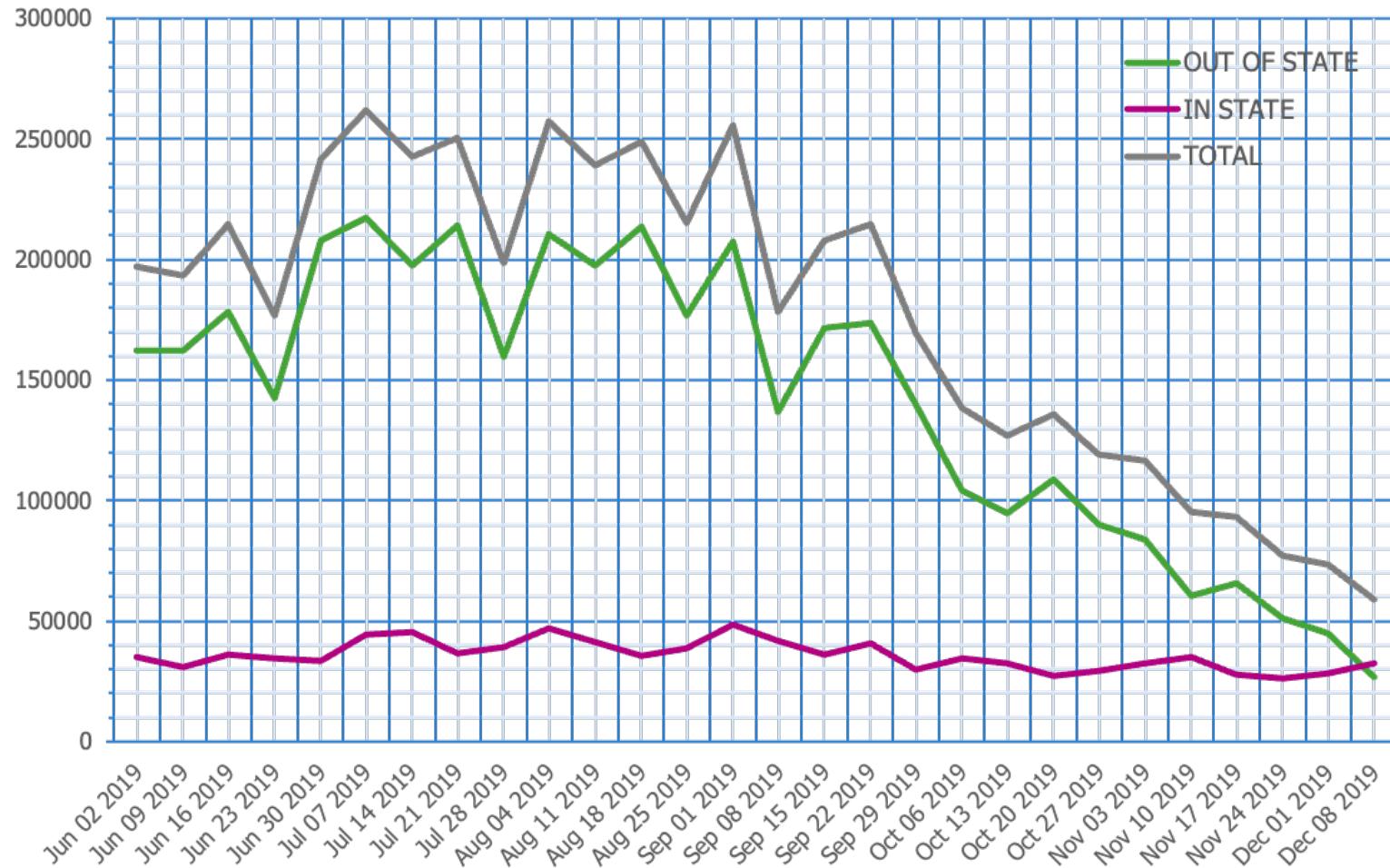
Exercise

Simplify and declutter.



Exercise: Simplify, Declutter, Add focus

Daily Average Park Visitors By Week



Peer-feedback on your own work.

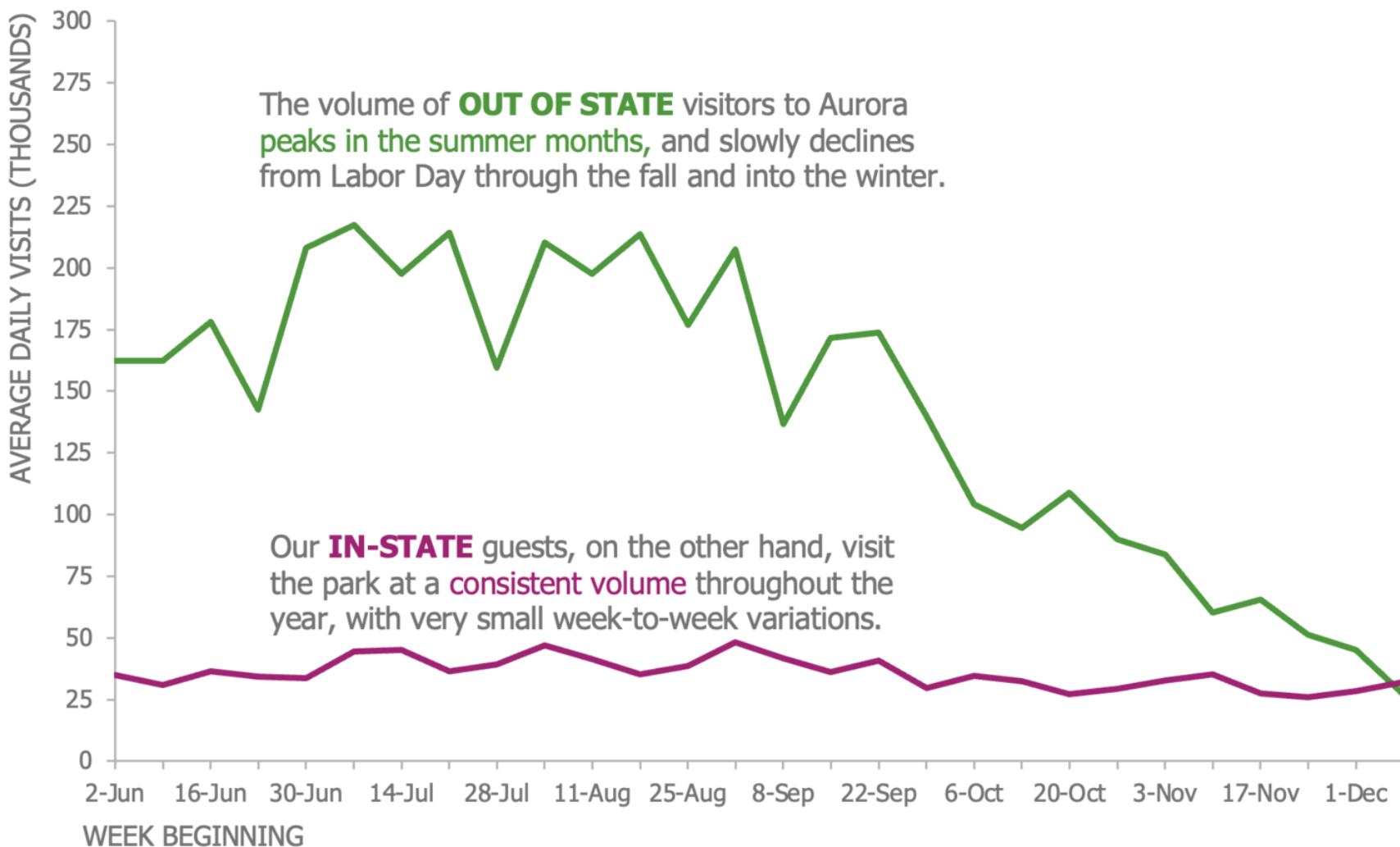


1. **Simplify:** Enable direct (fast) comparisons
2. **Declutter:** Eliminate irrelevant distractions
3. **Add focus:** Draw attention to where you want it

- Is the most appropriate chart type used to present the data/message?
- What takes most time to process? Could this process be simplified?
- Are there any unnecessary, redundant or distracting elements that can be removed/reduced?
- How could the message be more emphasized?
- Are all text and labels legible?

Daily visitors to Aurora Park in summer/fall 2019

VALUES ARE CALCULATED WEEKLY AS A 7-DAY AVERAGE

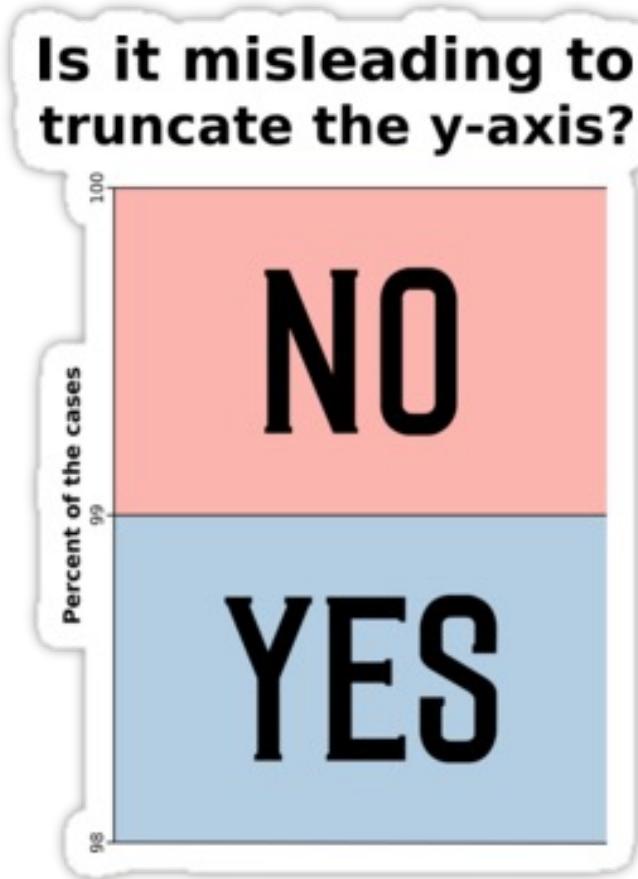


Do's and Don'ts

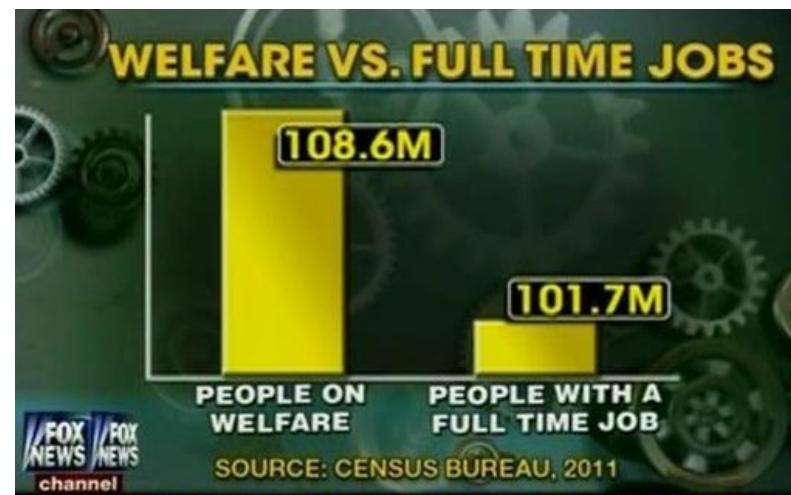
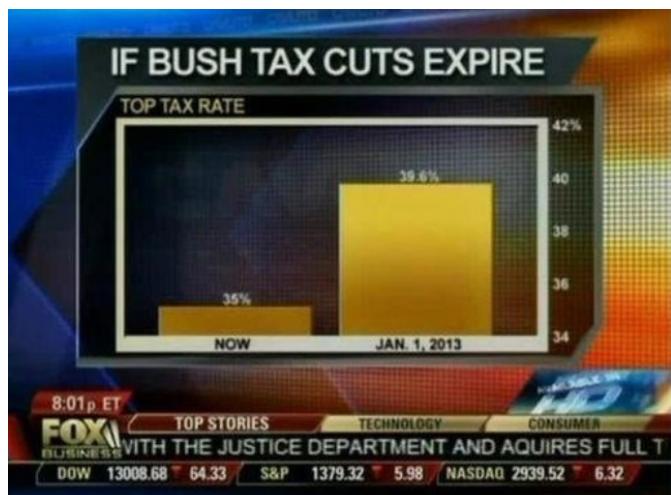
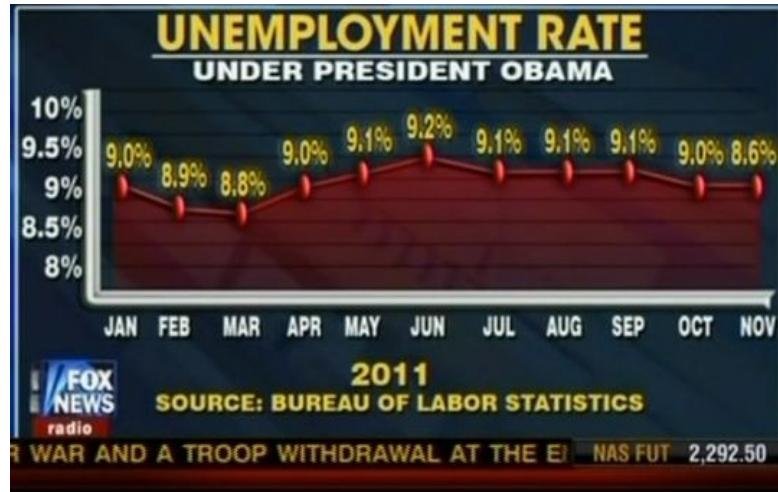
Show the *right* data, show the *data* right!



Don't: Truncate the y-axis



Who does that, you say?

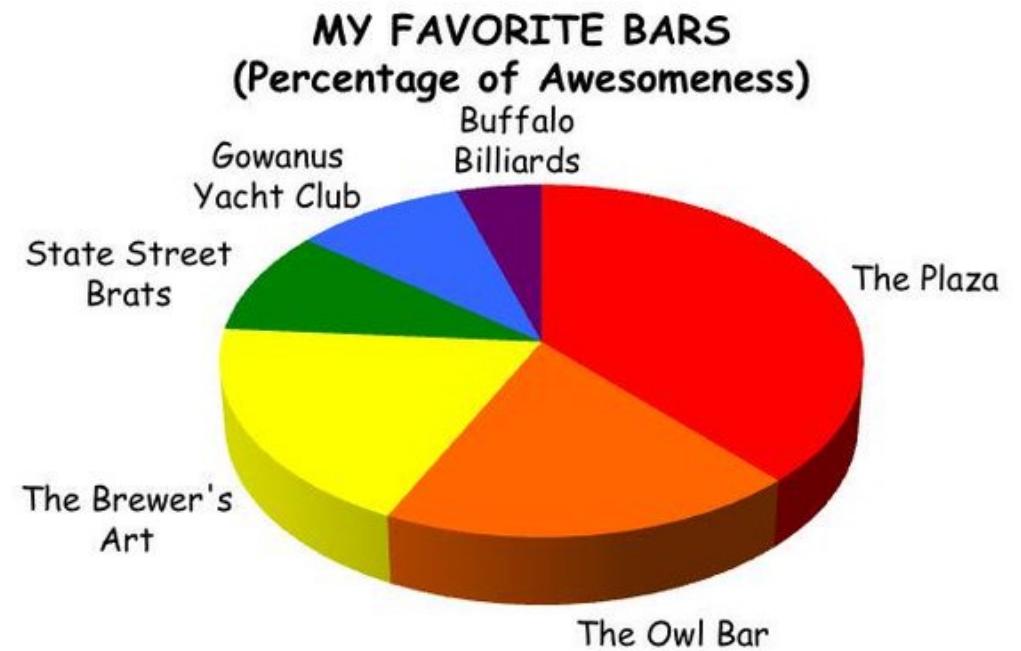
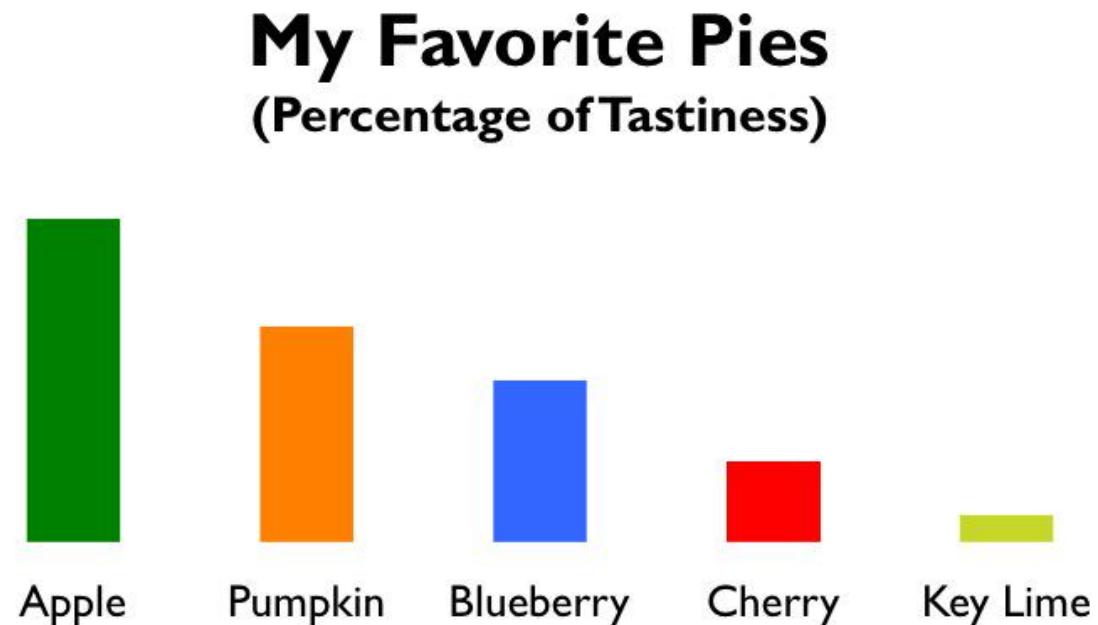


Don't: use double axes

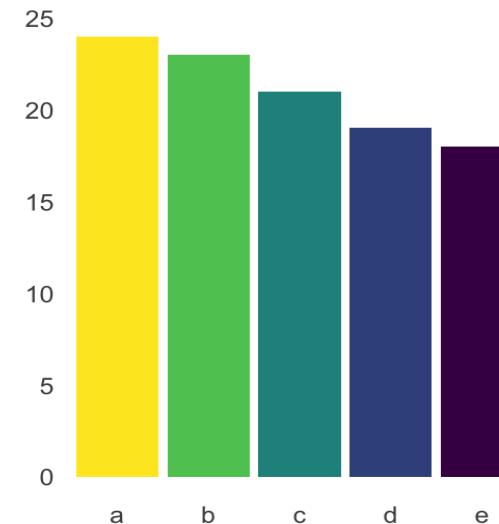
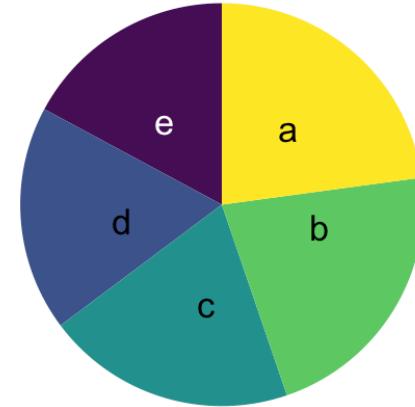
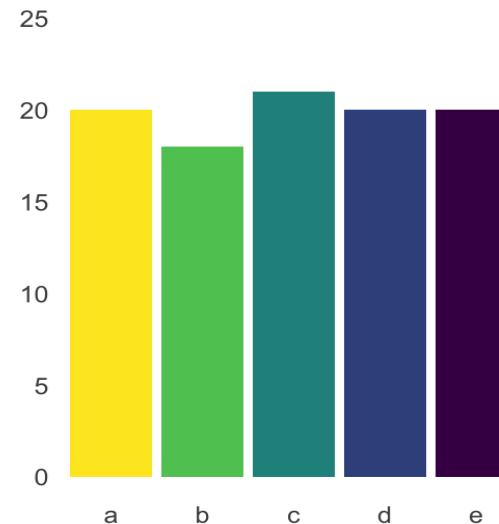
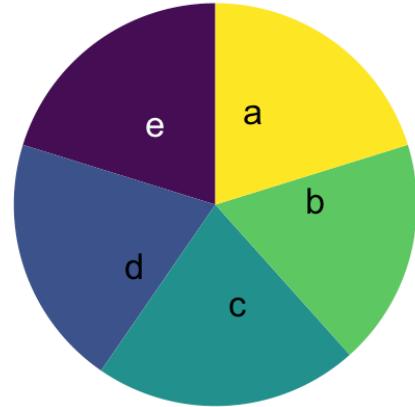
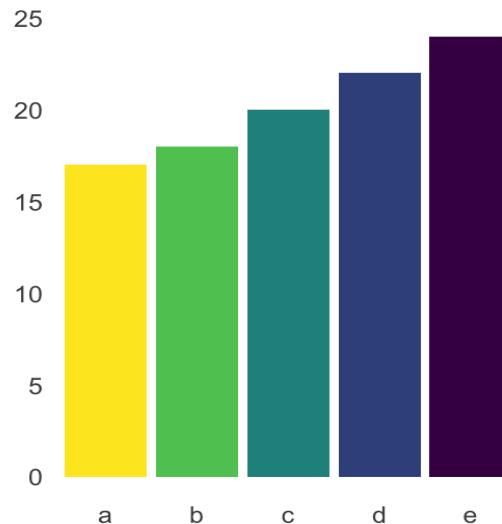
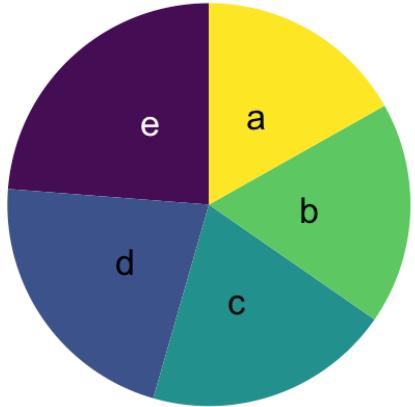
Number of people who drowned by falling into a pool
correlates with
Films Nicolas Cage appeared in



Do: Know when to choose for *bars* or *pies*

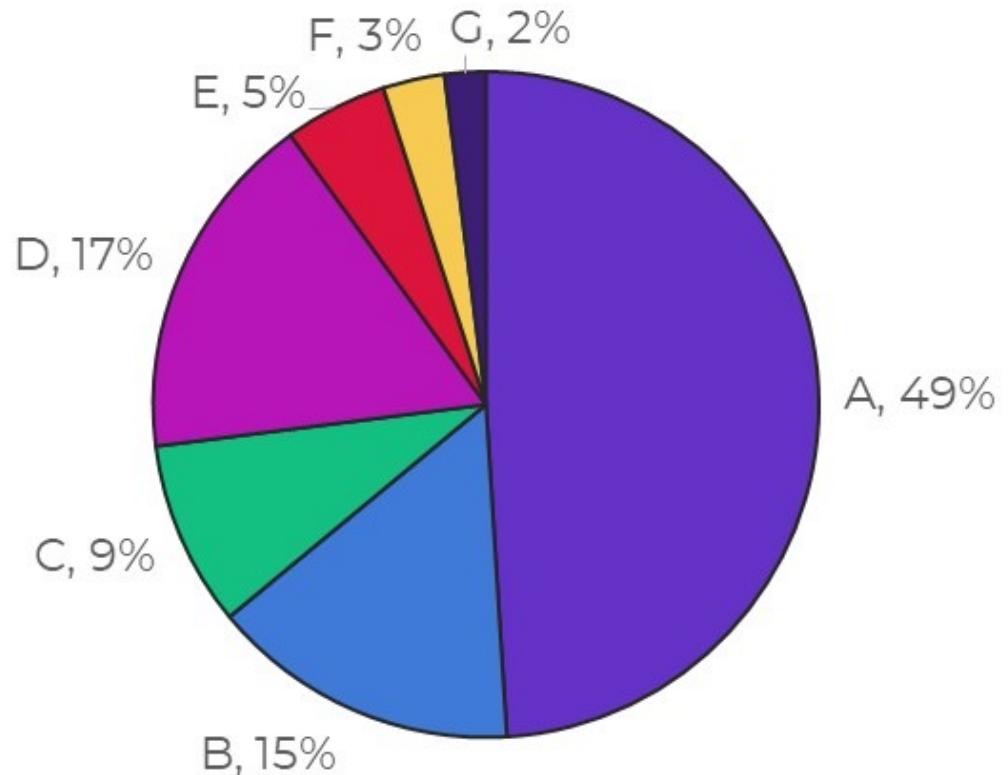


The problem with pie charts...

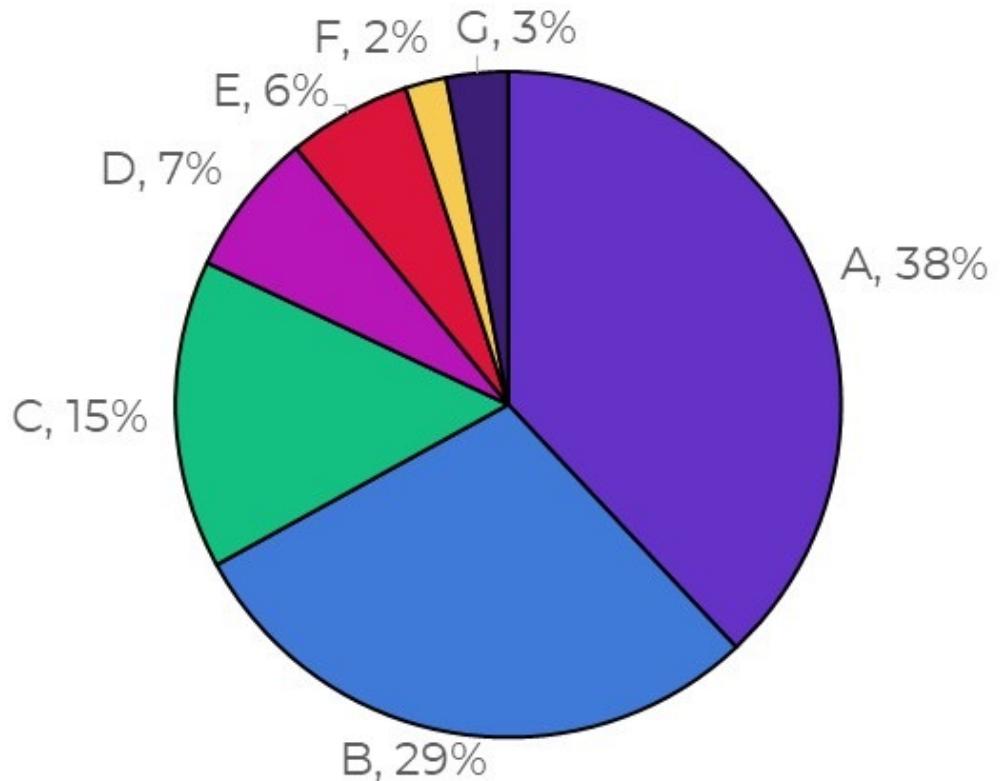


The problem with pie charts...

ABC-Funded Research
Projects by Topic, 2012-2014

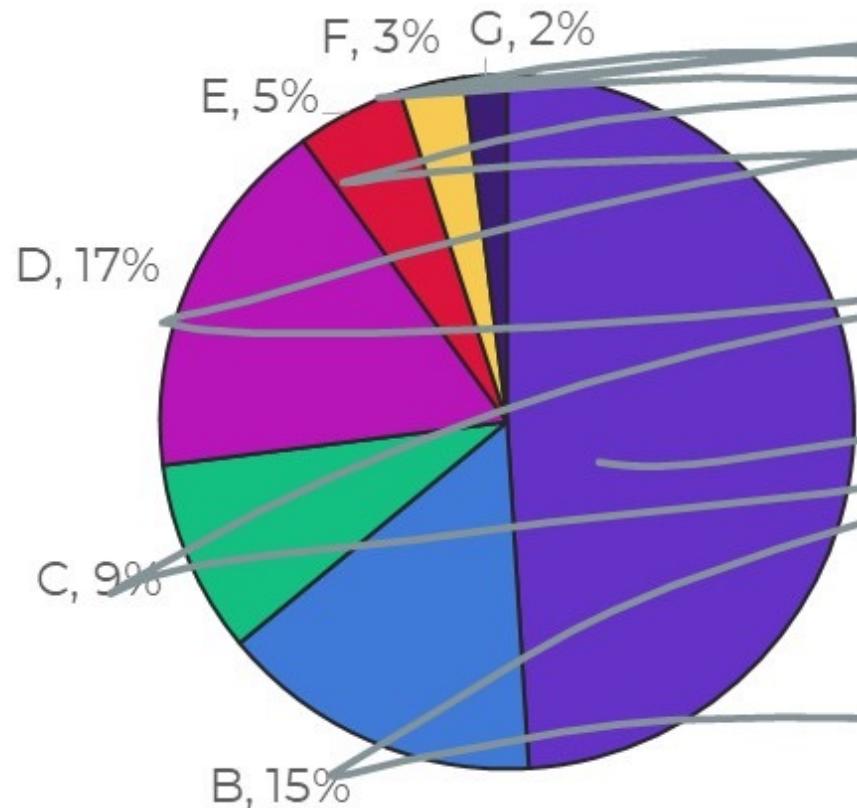


ABC-Funded Research
Projects by Topic, 2015-2017

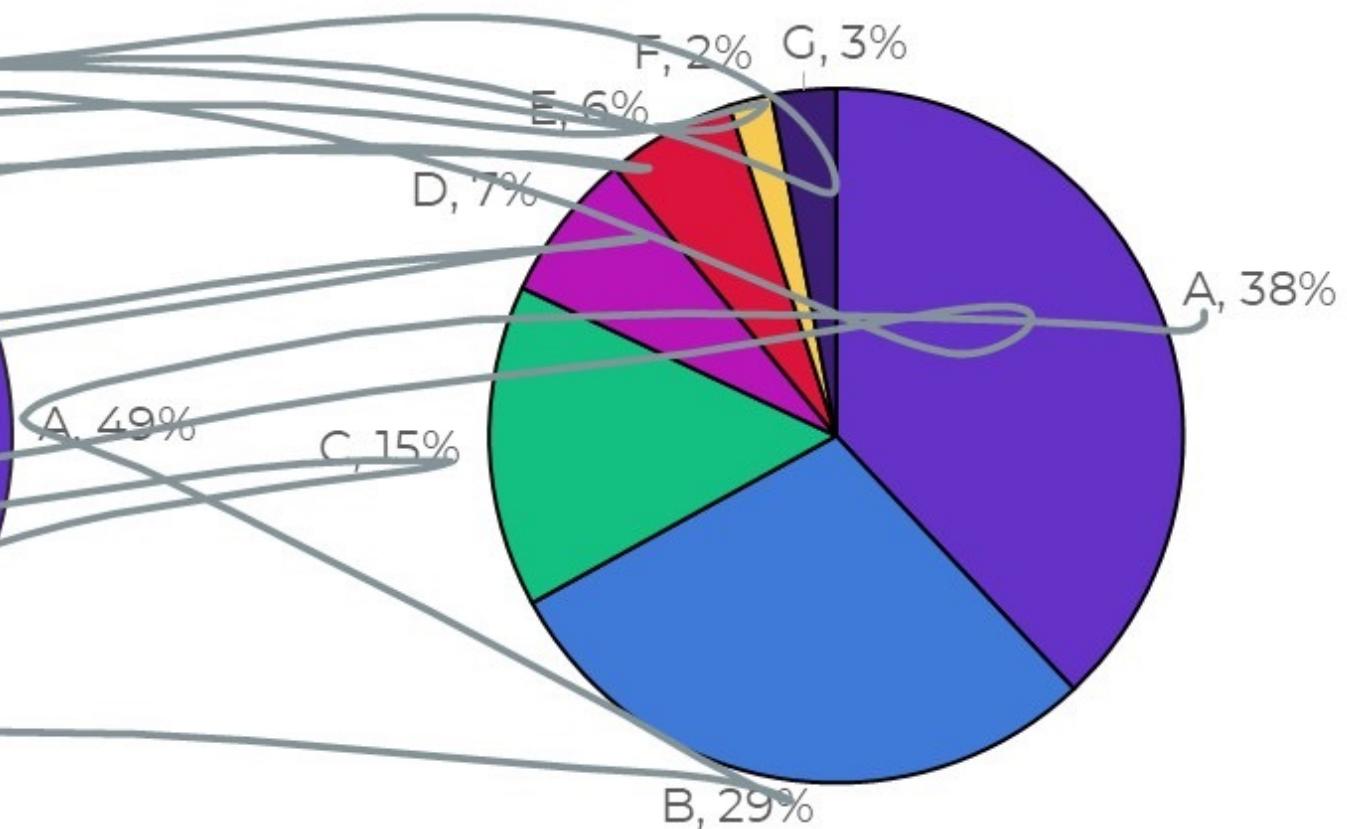


The problem with pie charts...

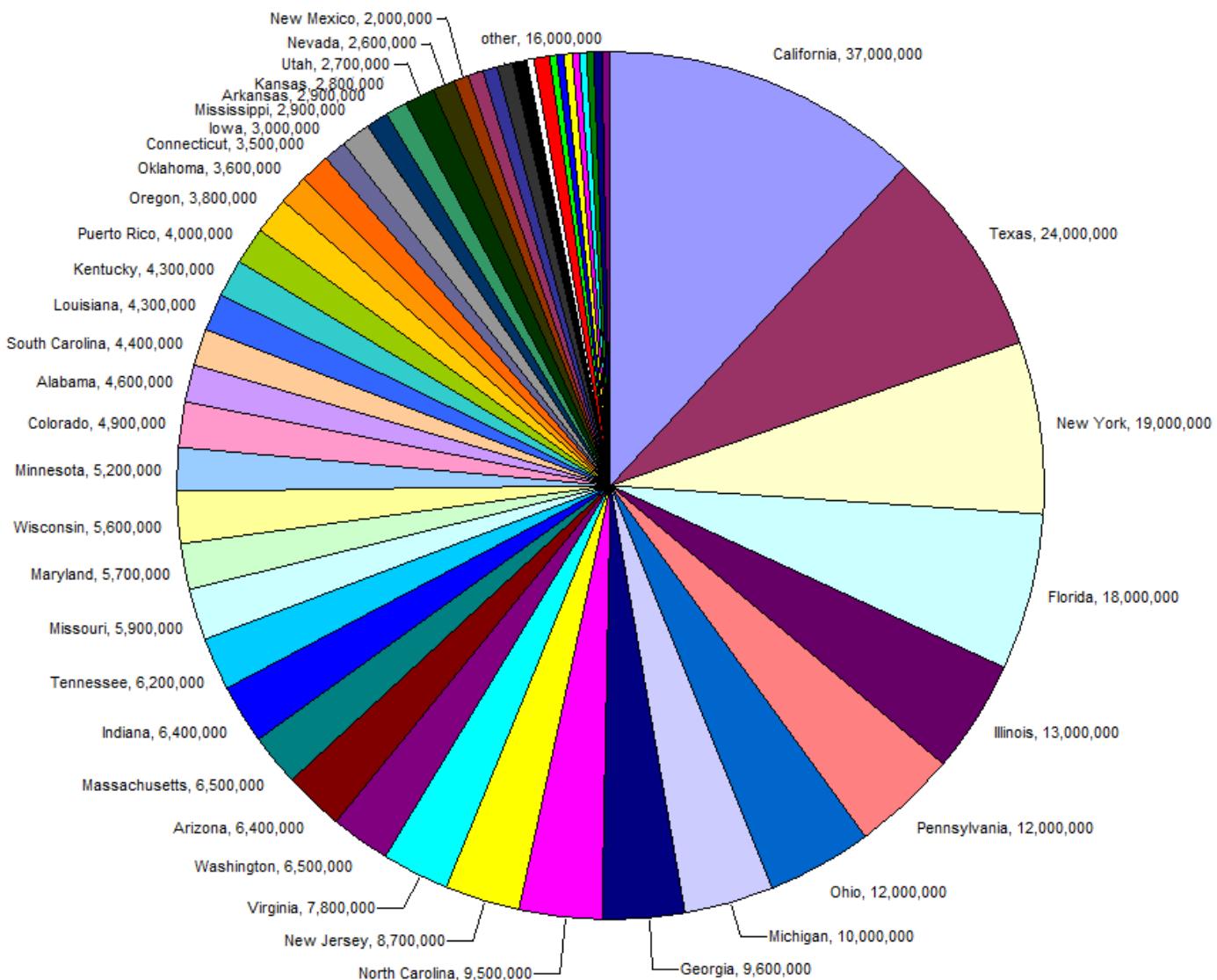
ABC-Funded Research
Projects by Topic, 2012-2014



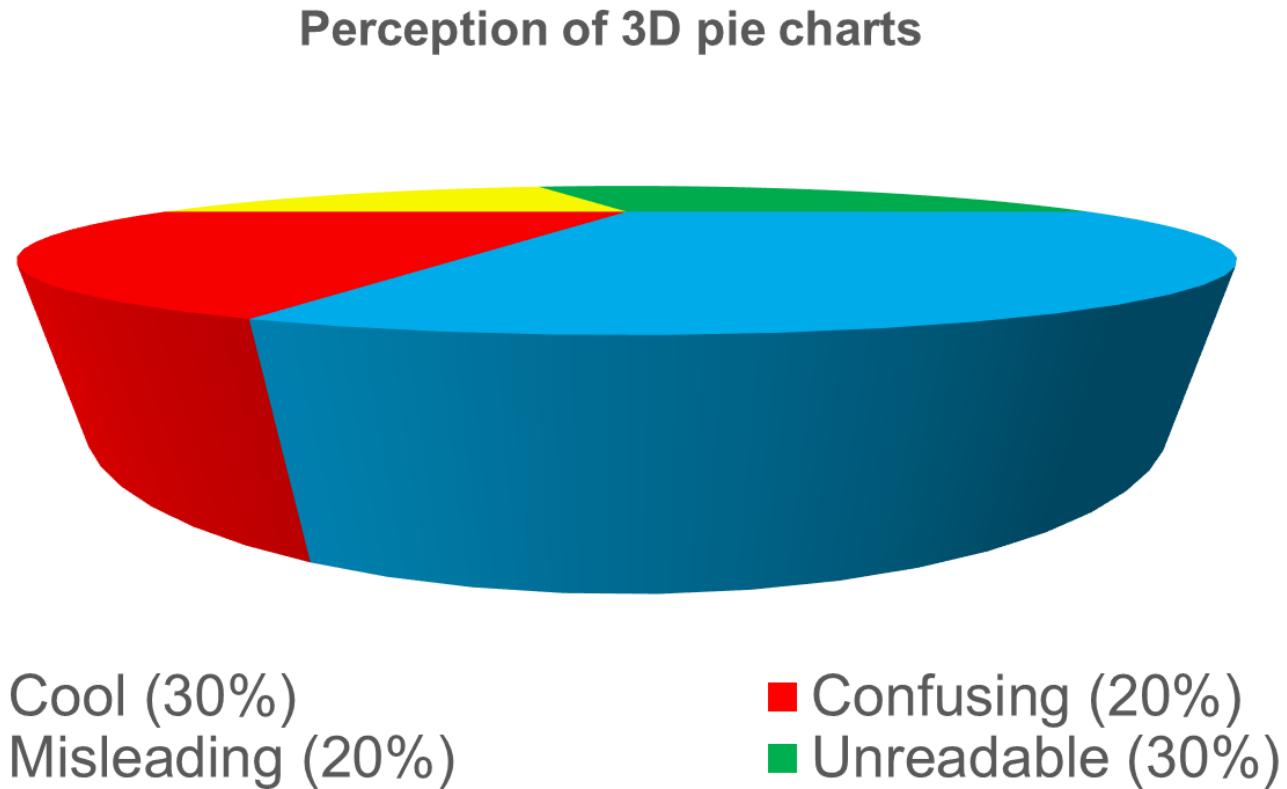
ABC-Funded Research
Projects by Topic, 2015-2017



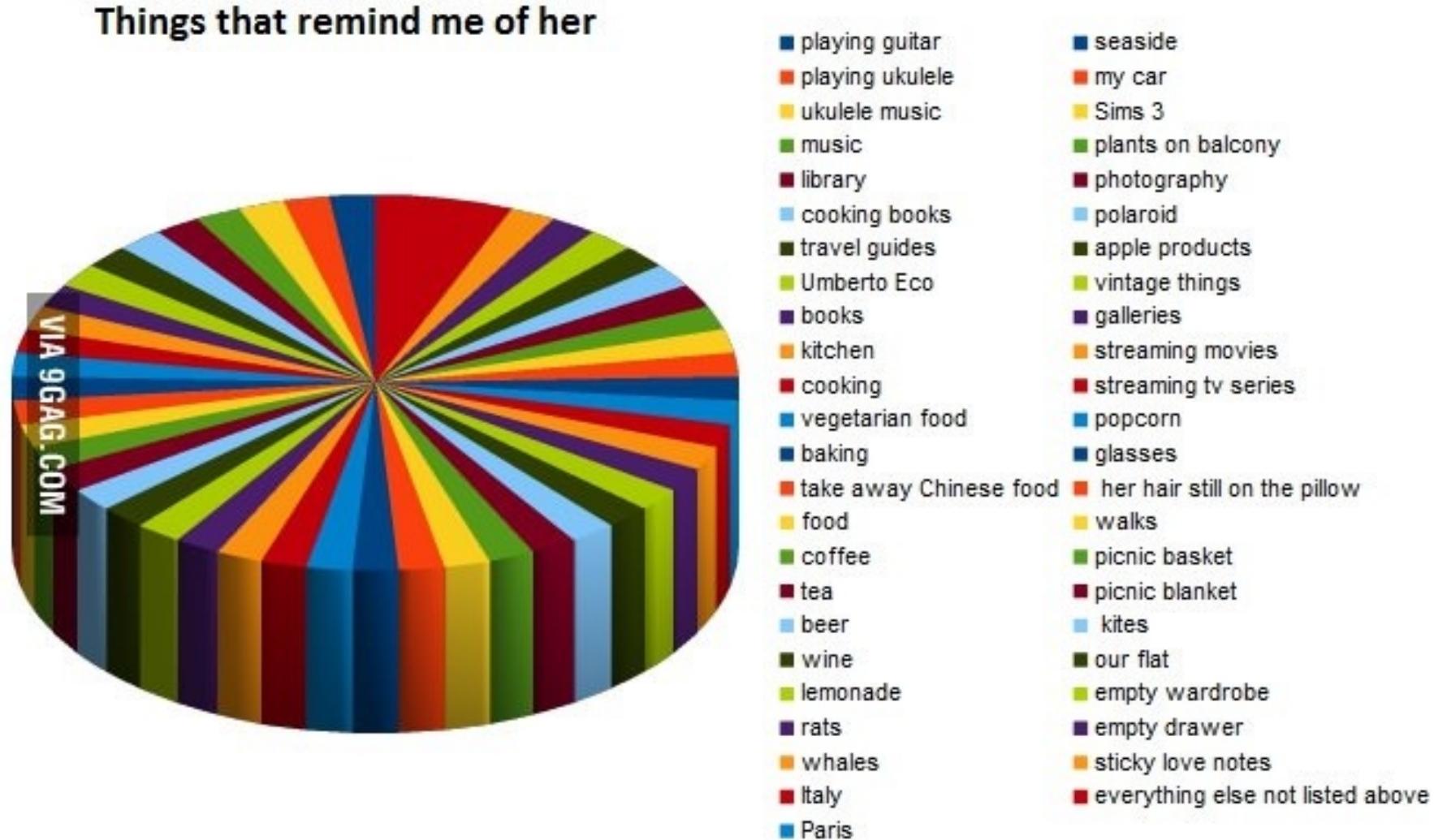
Don't: Use pie charts with >5 categories



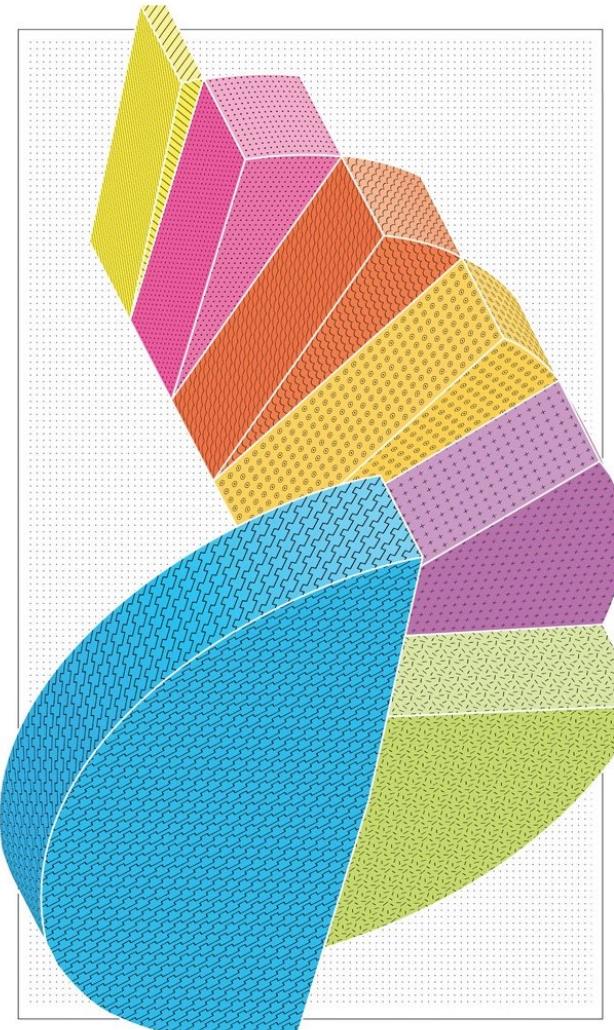
Don't: Use 3D pie charts



Don't: Use 3D pie charts with >5 categories



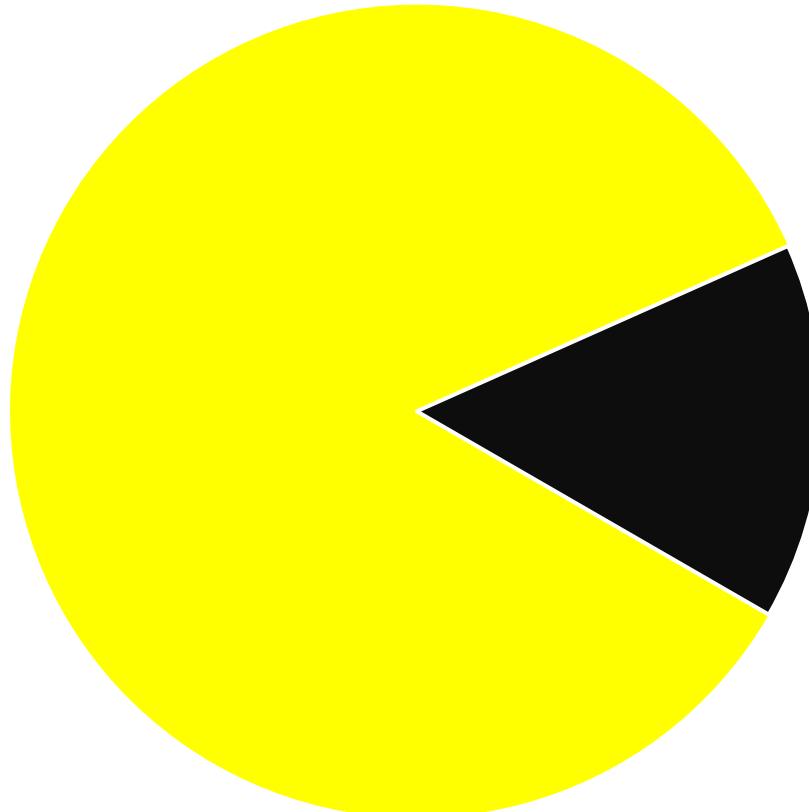
Don't: Use patterned 3D winding staircase lying on its side pie diagrams with >5 categories!



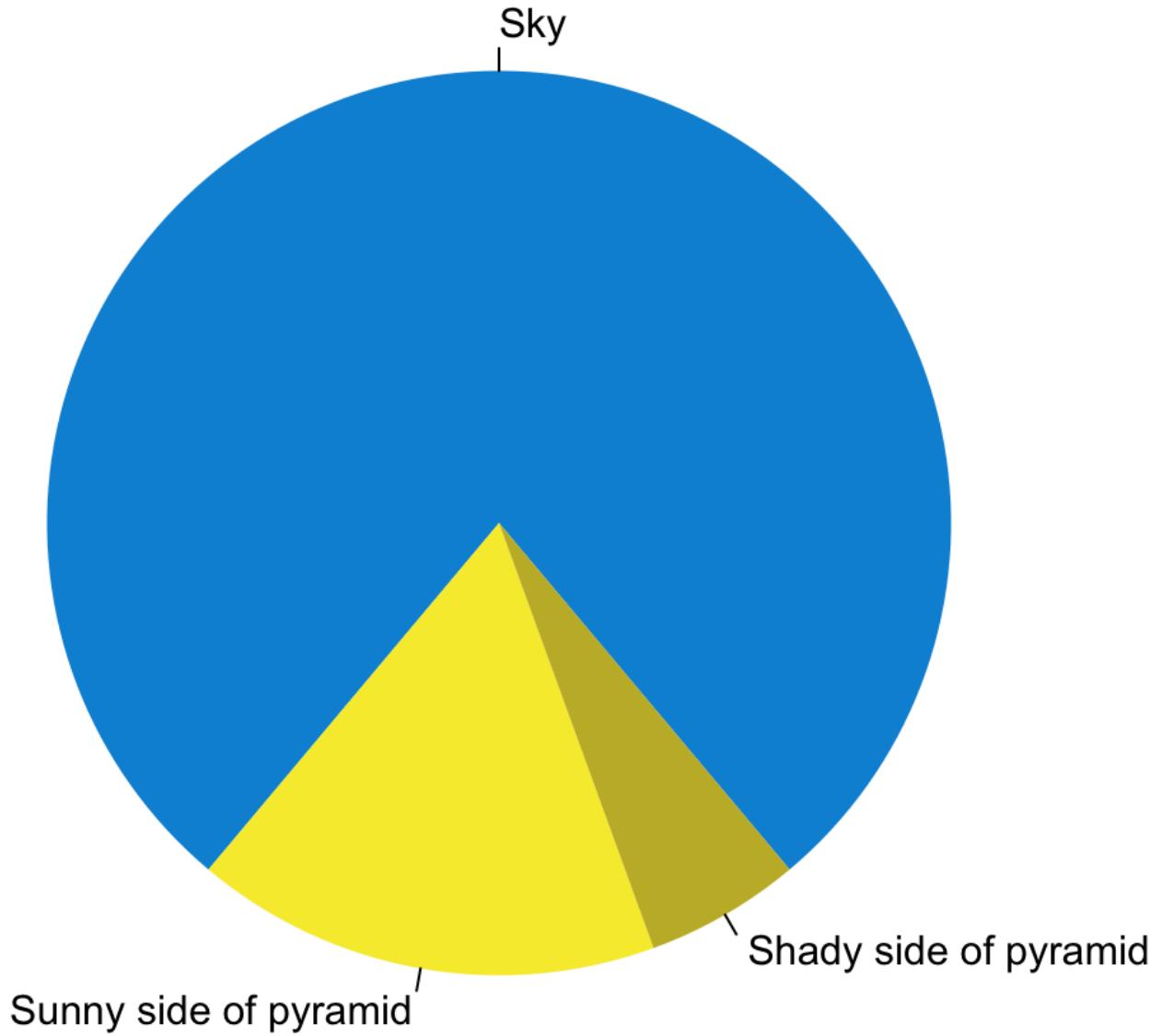
Pie charts are great to show the part-to-whole relation of 2 categories

Percentage of chart that looks like Pacman

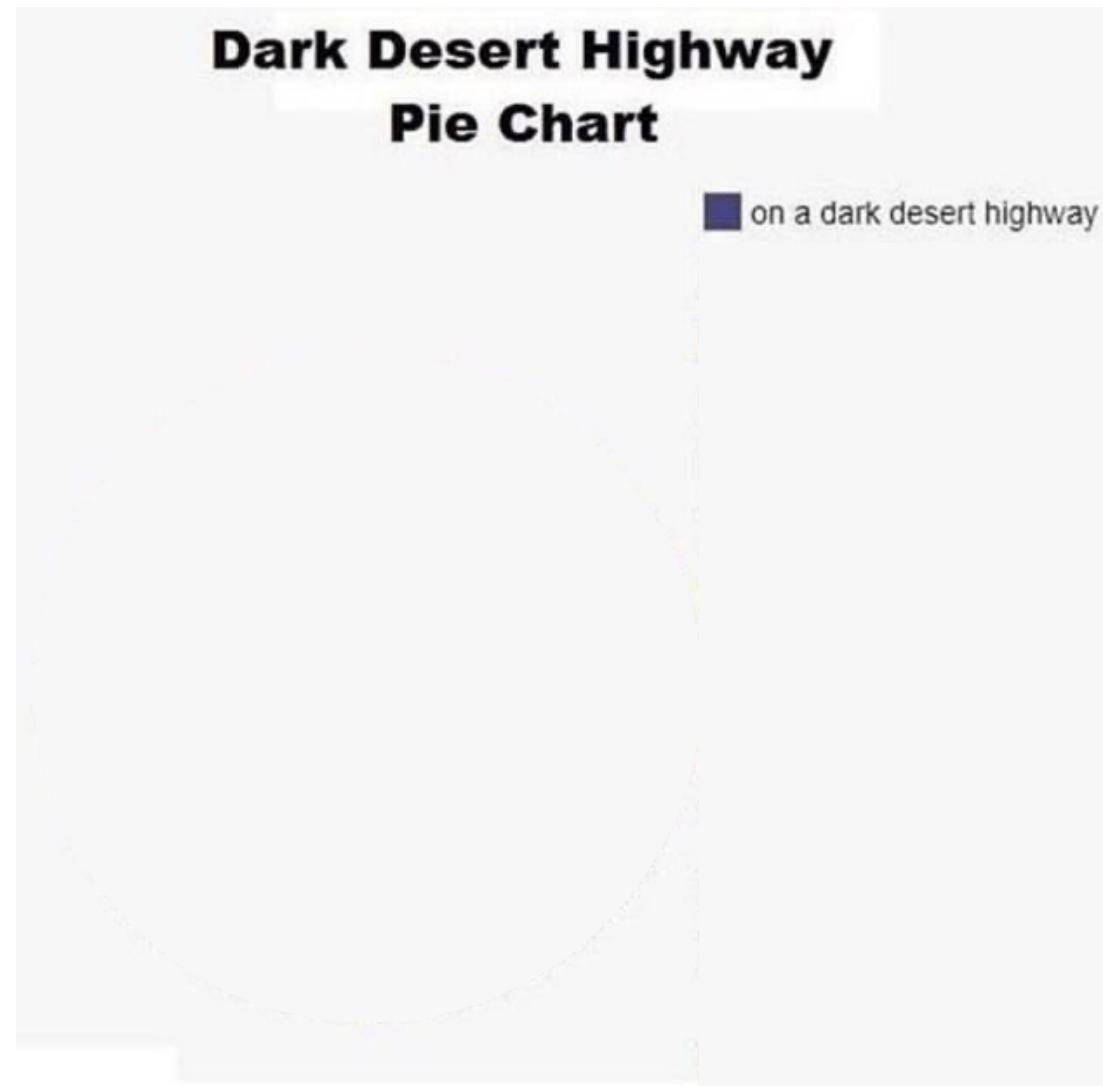
- Pacman
- Not Pacman



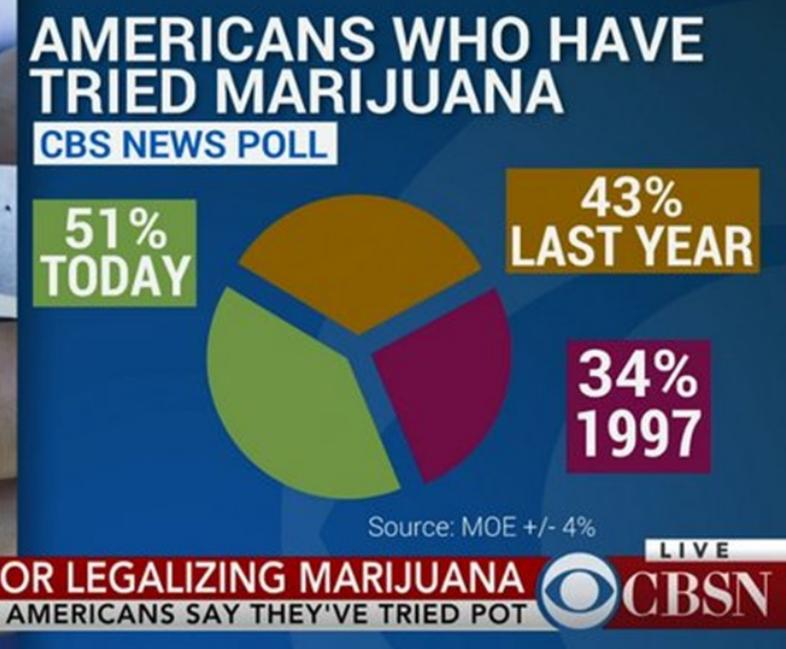
Pie charts for 3 categories are okay too



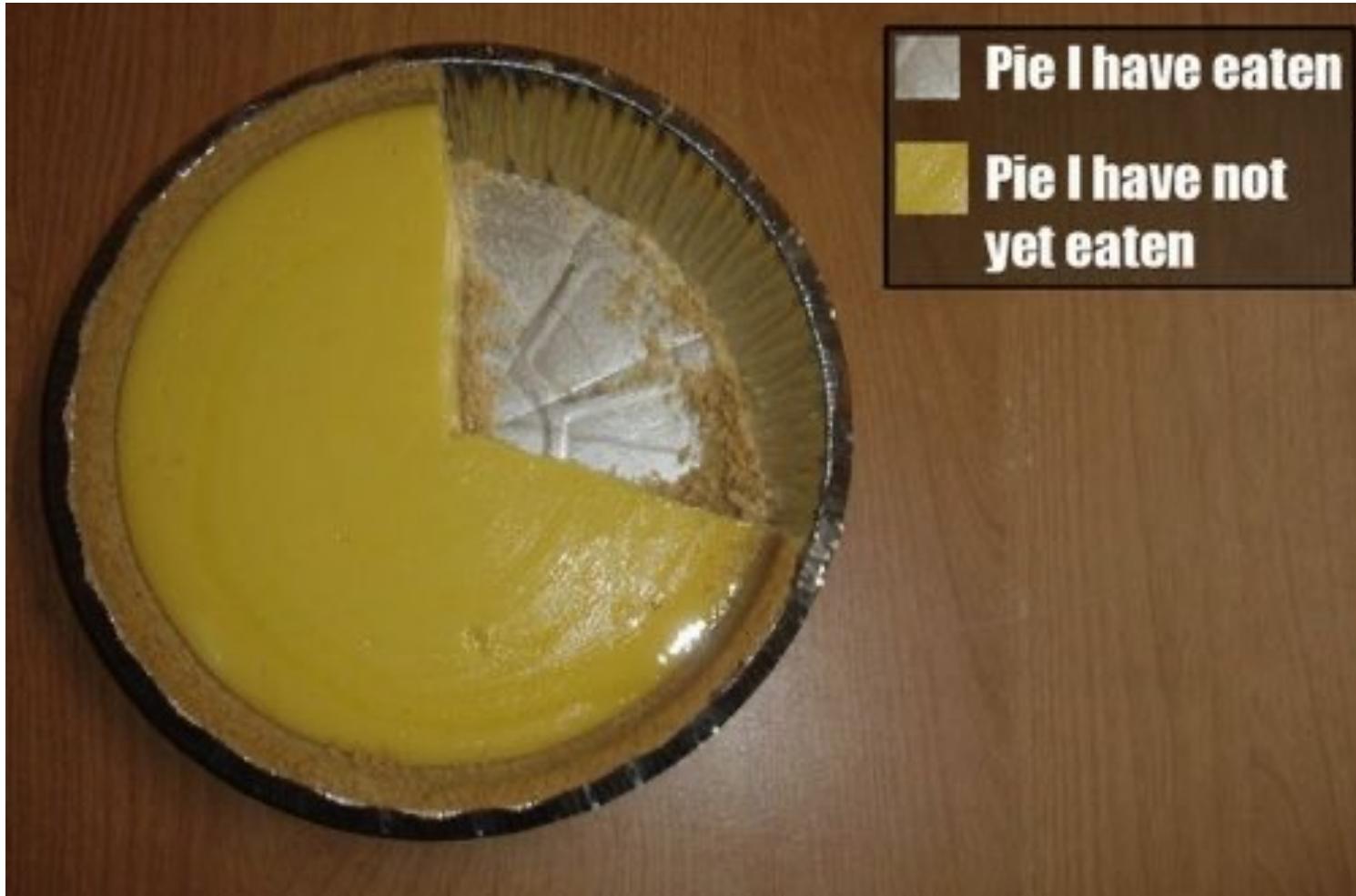
Sometimes pie charts with >5 groups work 😊



The *worst* pie chart ever...?



The *best* pie chart ever?



Thank you!



Irene van den Broek, PhD

✉ Irenevandenbroek@gmail.com

LinkedIn icon @IrenevdBroek

Twitter icon @JeBentWatJeMeet

Happy Viz-ing!

Links to slides and resources:

<https://github.com/IreneVDB/IreneVDB.github.io/blob/main/VLAG/>

- Presentation.pdf
- ReferenceList.pdf



<https://irenevdb.rbind.io>