

Effective Data Visualization for Scientific Researchers

VLAG PhD week
Baarlo, April 6th 2023

Irene van den Broek, PhD

 @IrenevdBroek

 @JeBentWatJeMeet

 @IreneVDB

Hi there, I'm Irene!

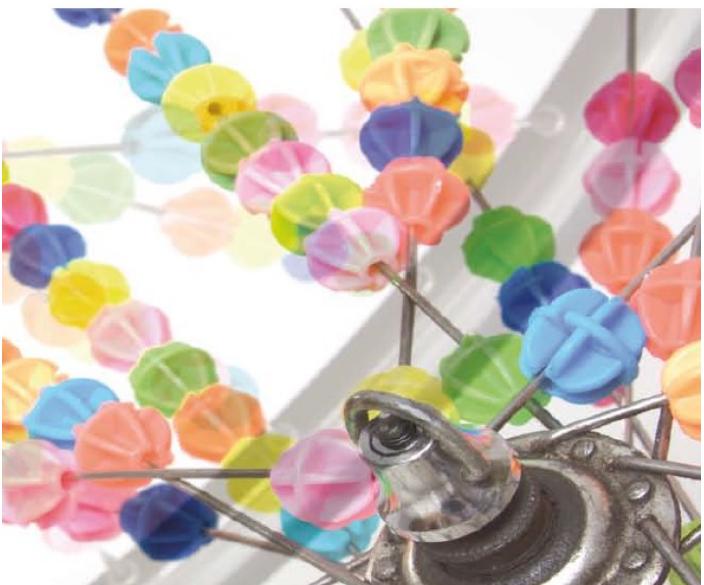


Irene van den Broek

irenevandenbroek@gmail.com

I have a background 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

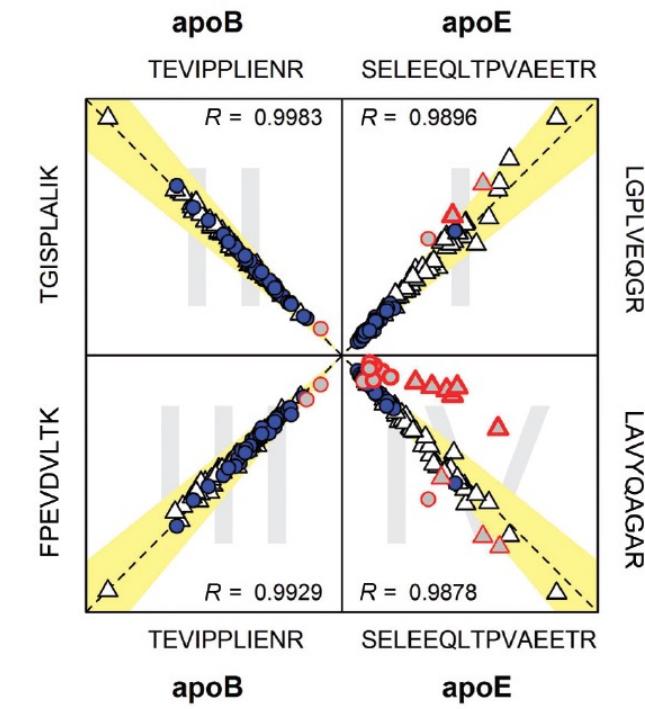
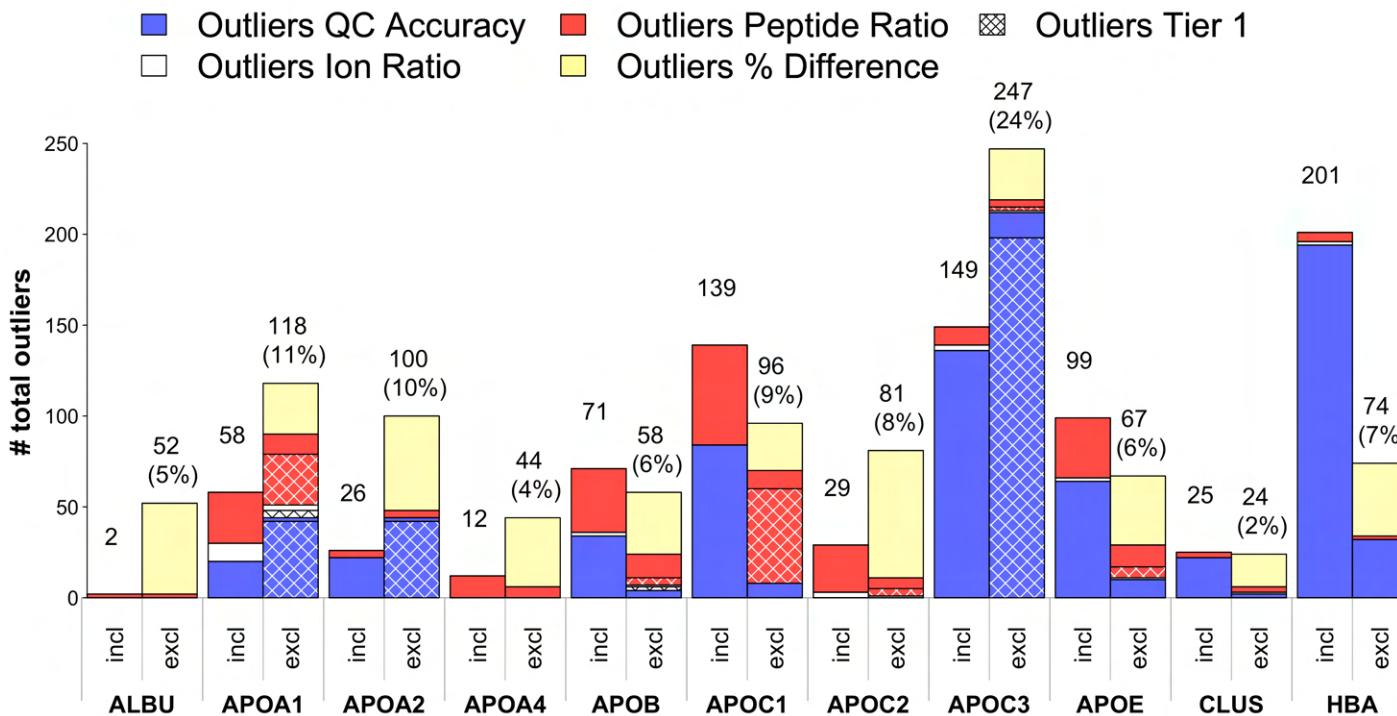


2013-2015
Leiden University
Medical Center

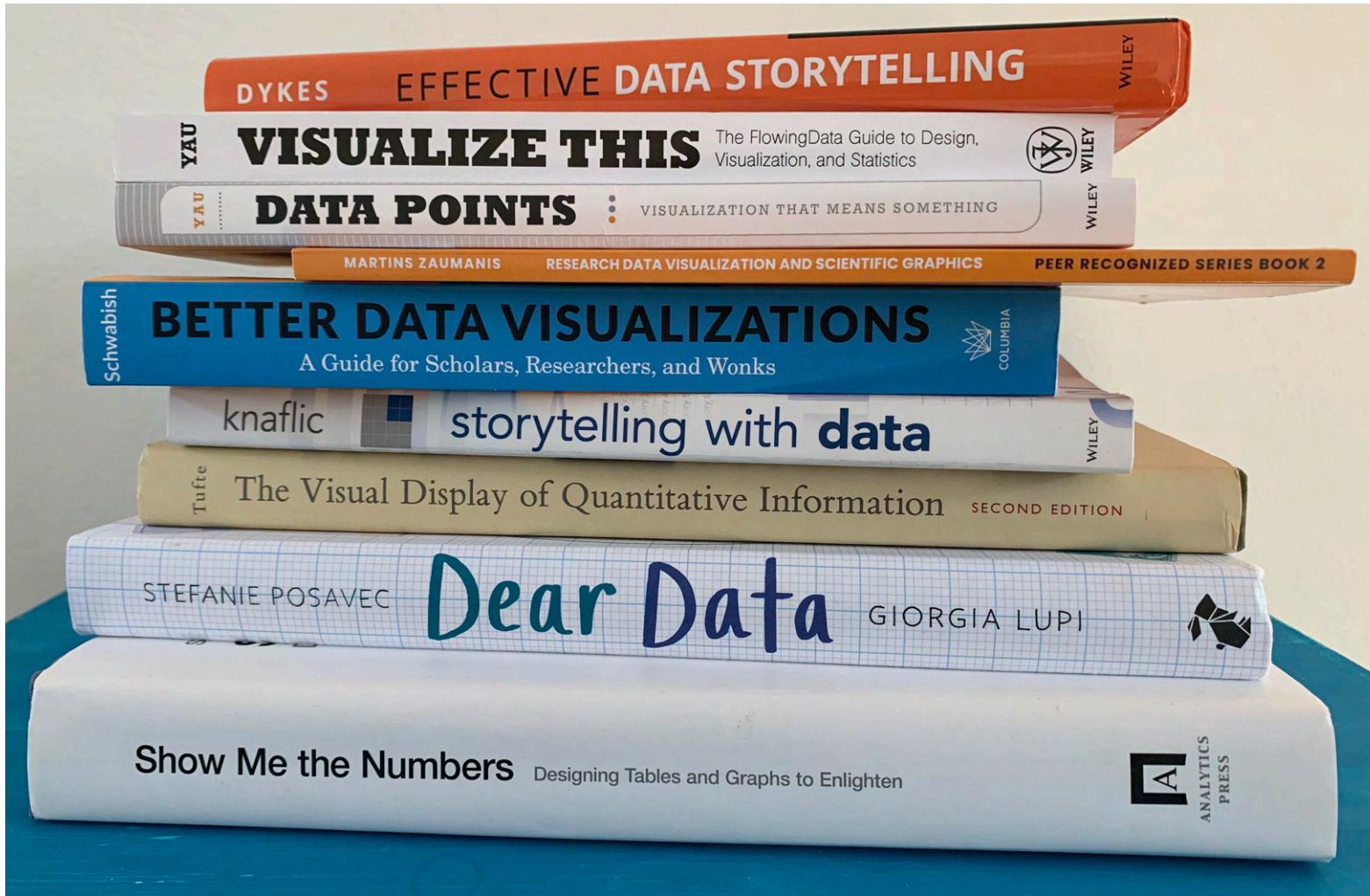


2015-2019
Cedars-Sinai
Medical Center

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

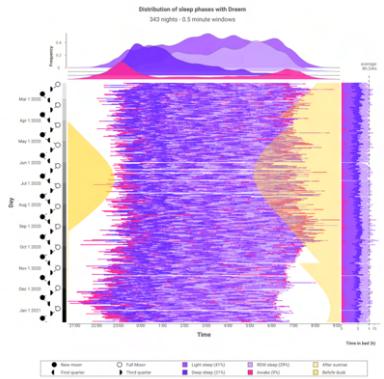


When I learned **Data Visualization** is a field on its own, I was hooked!

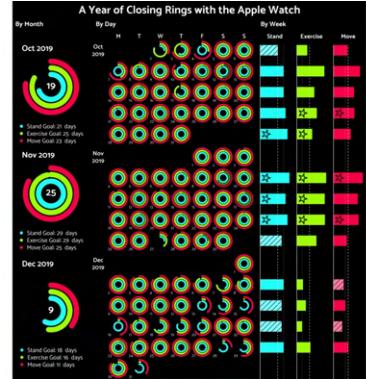


I started visualizing my Health and Habits.

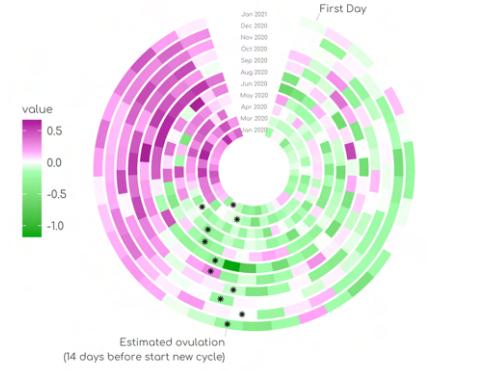
Sleep



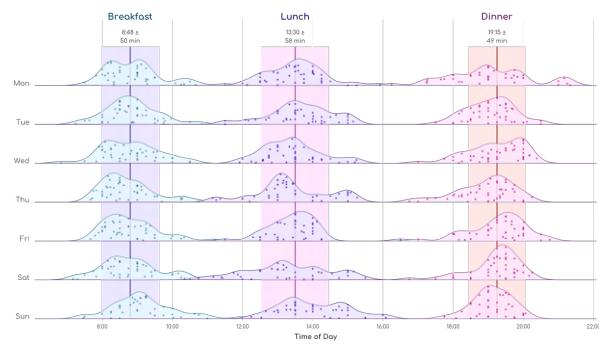
Activity



Body



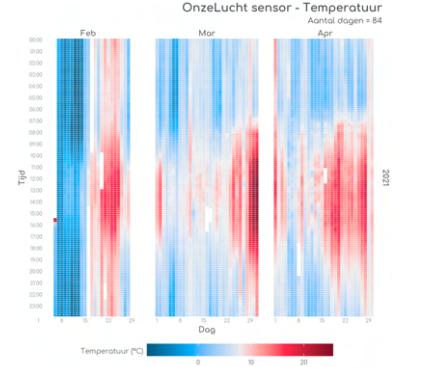
Food



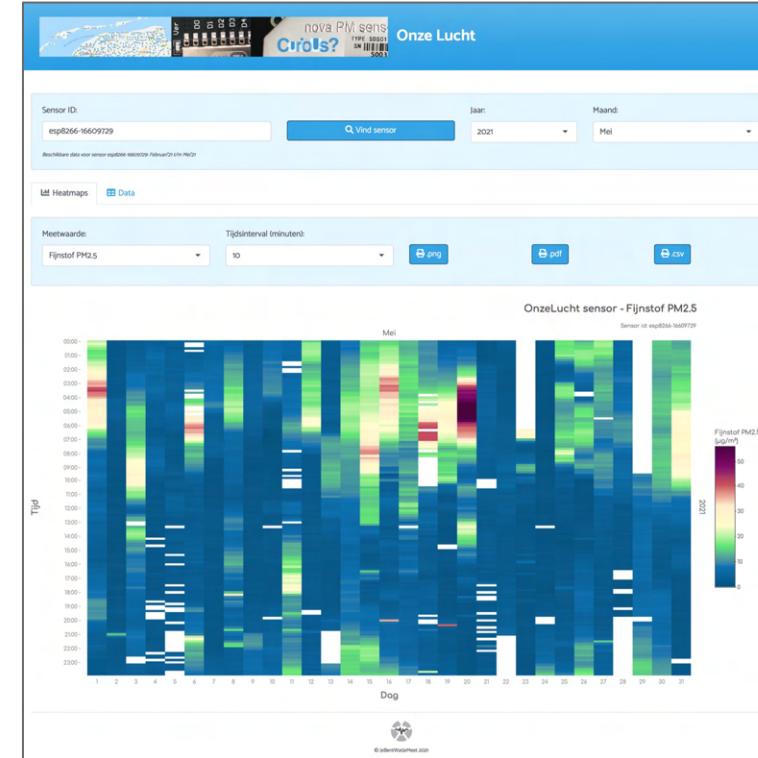
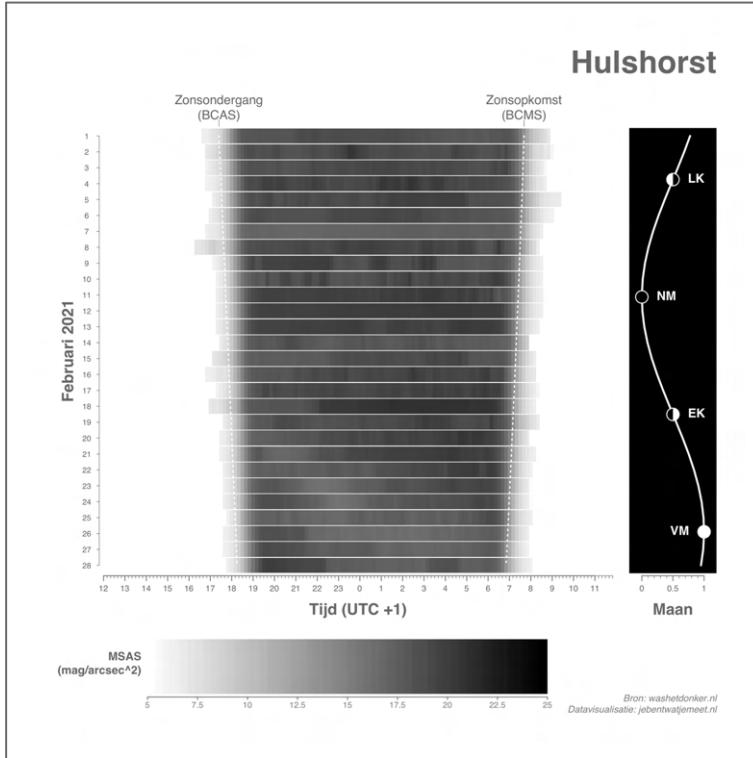
Habits



Environment



And now combine my passions for teaching and data visualization



www.washetdonker.nl



www.onzelucht.nl

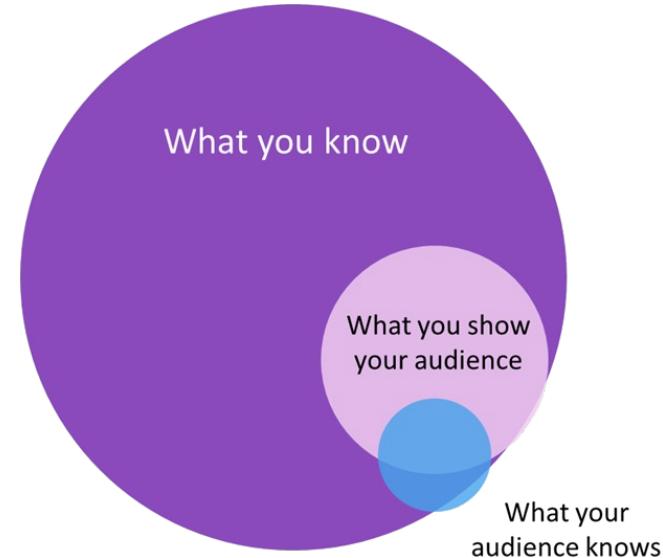
I'm excited to share with you what I wish I'd learned during my PhD!

Resist the urge to show everything!



Don't underestimate how smart you are!

The curse of knowledge

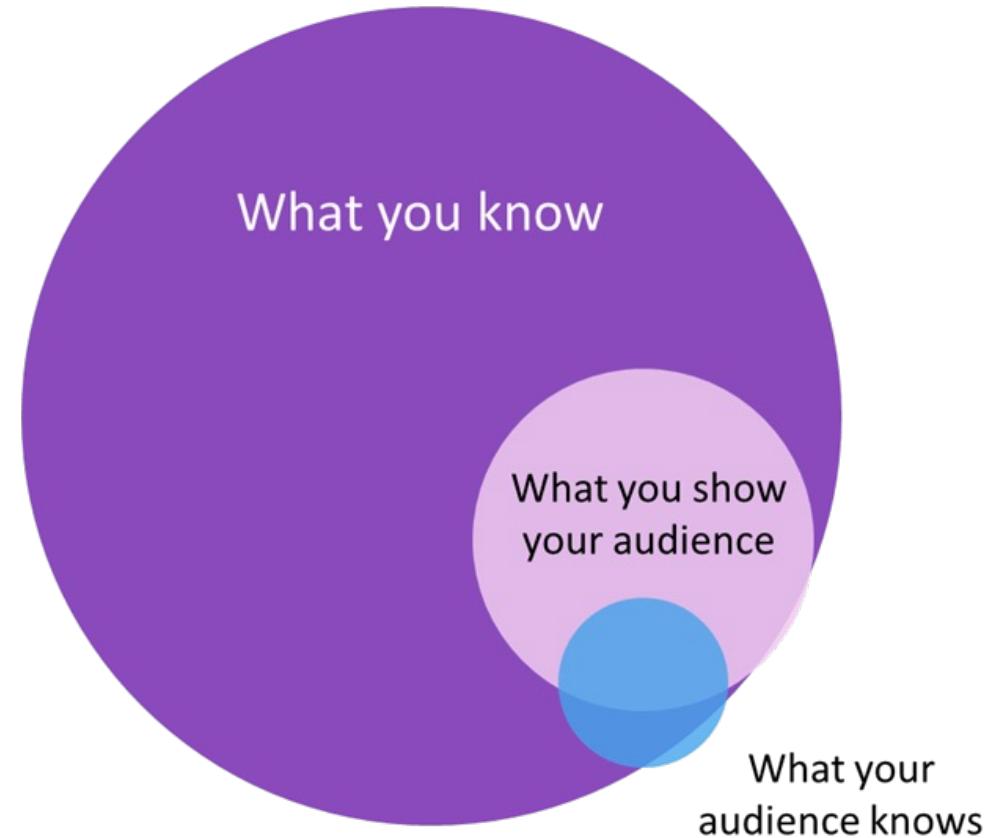


Resist the urge to show everything!

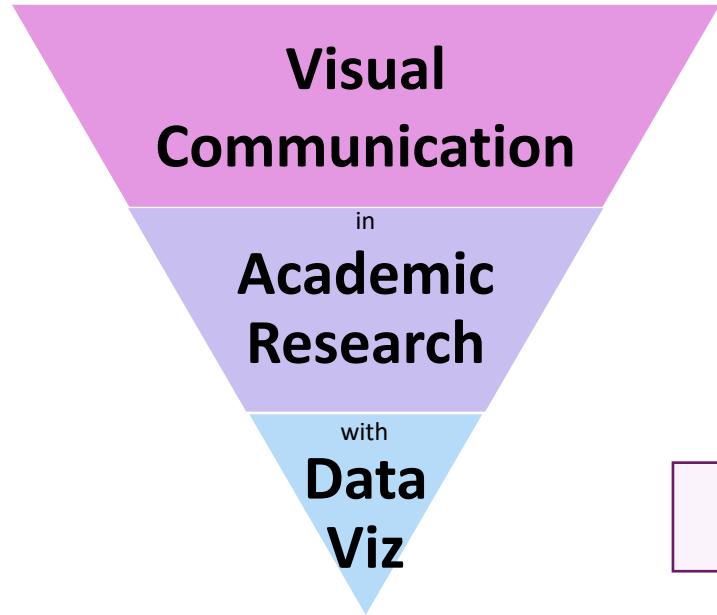


Don't underestimate how smart you are!

The curse of knowledge



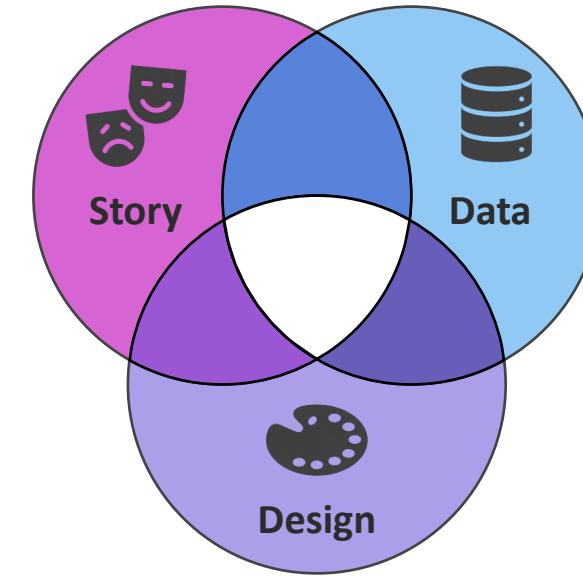
An overview for today:



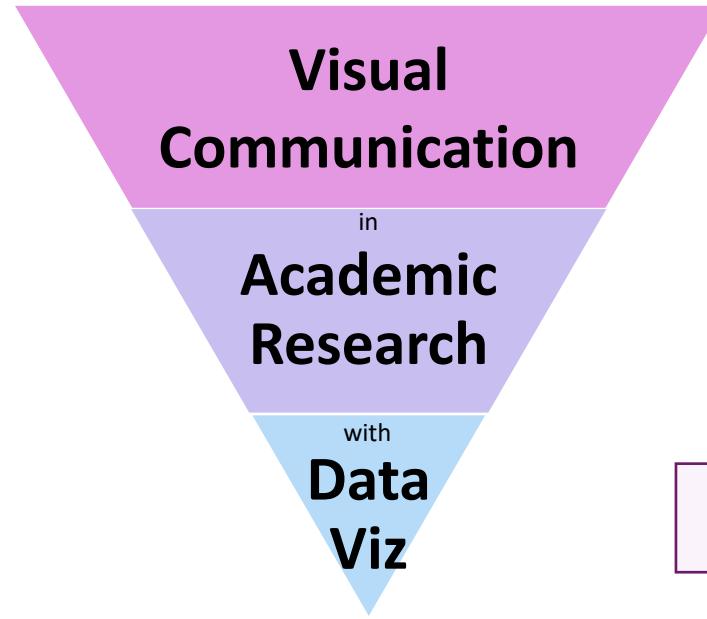
WHY?

WHAT?

HOW?



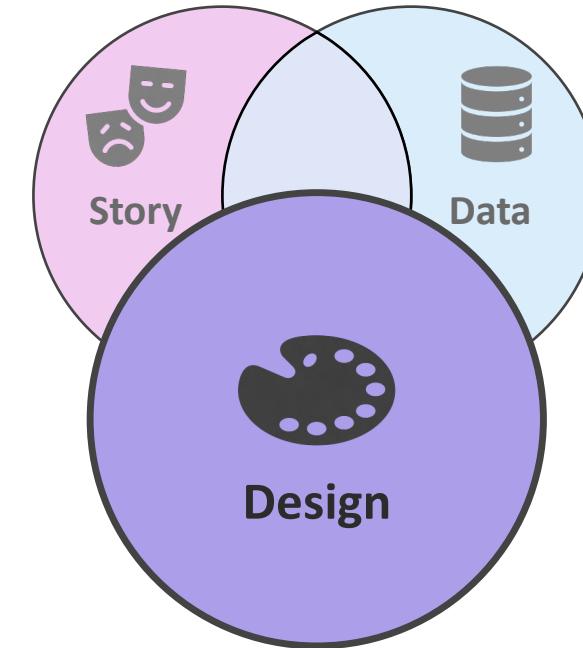
An overview for today:



WHY?

WHAT?

HOW?



🧠 Understand the brain

👍 Do's and 👎 Don'ts

✍ Practice 🕳️

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

Warm-up!

And stand up

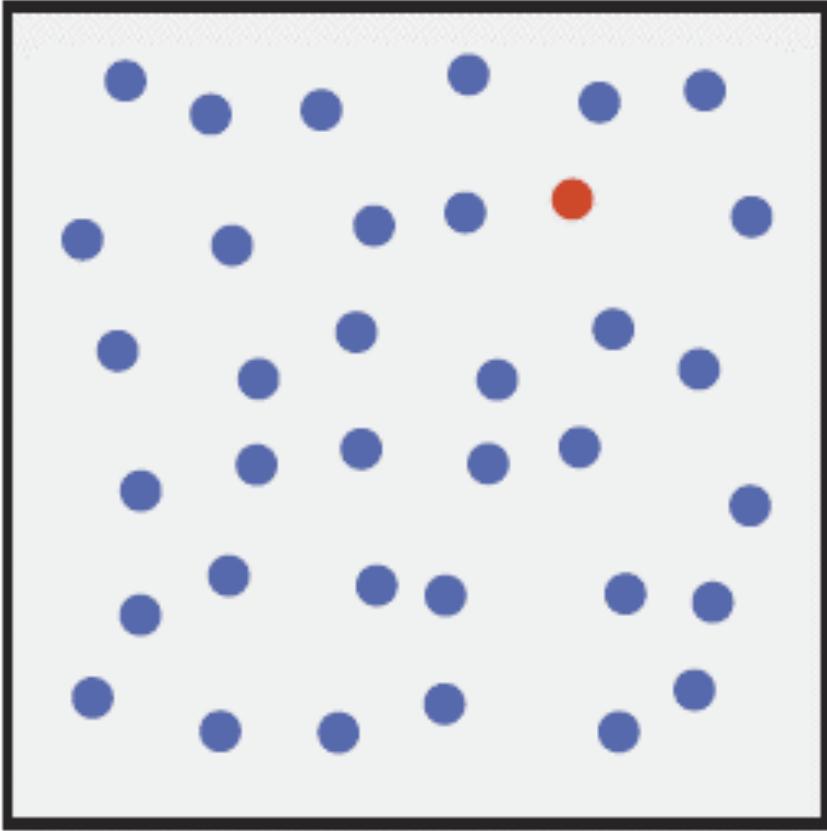
LEFT?

YES?

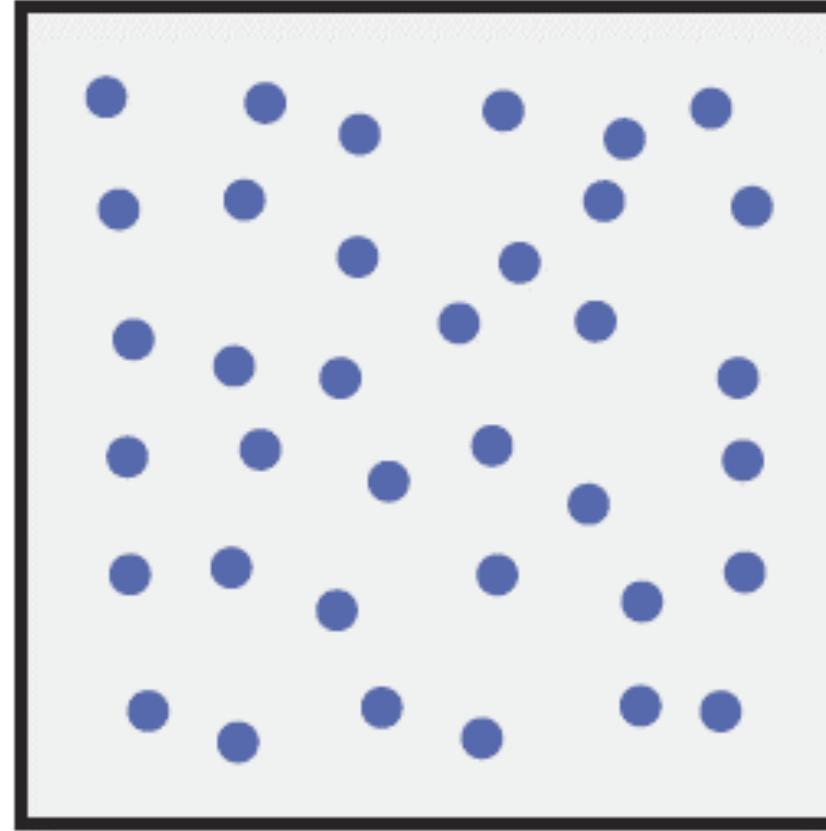


Which panel has a red dot?

A



B



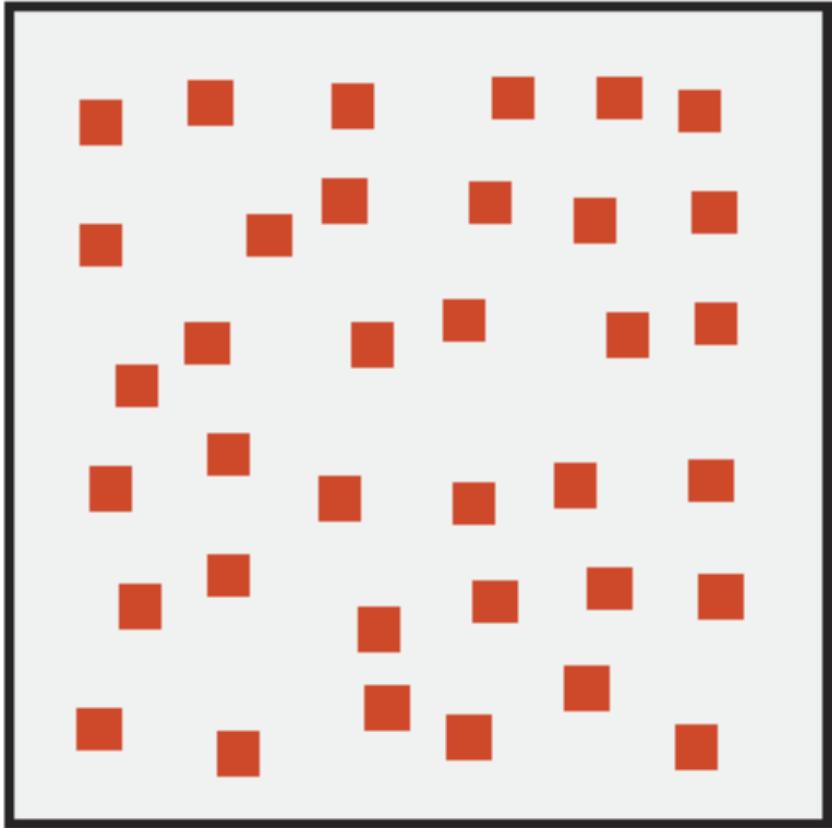
LEFT?

=

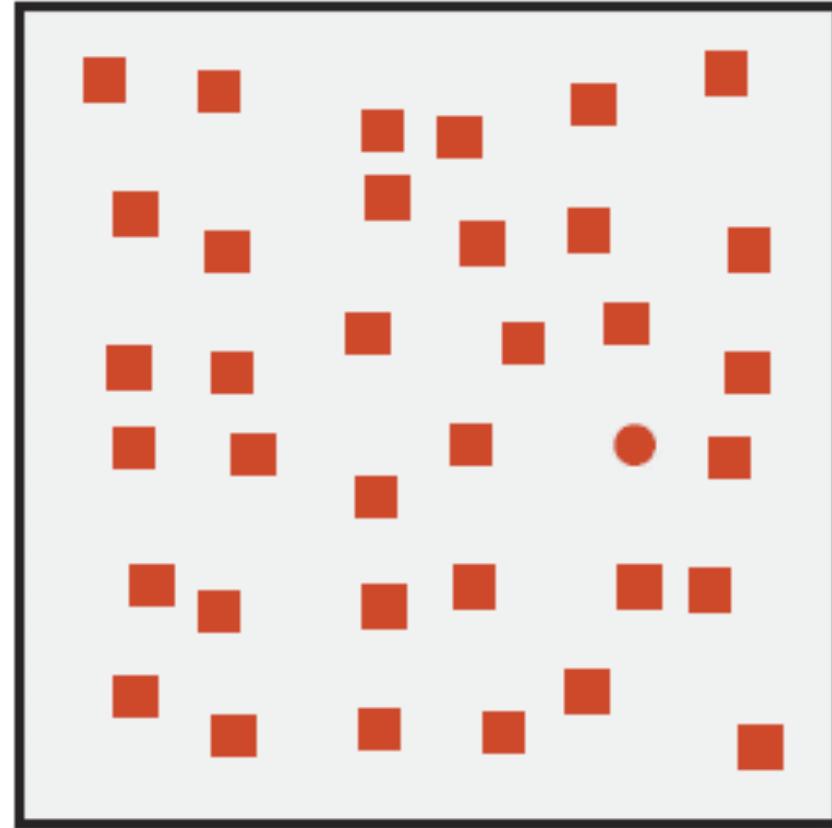
UP!

Which panel has a red dot?

A



B



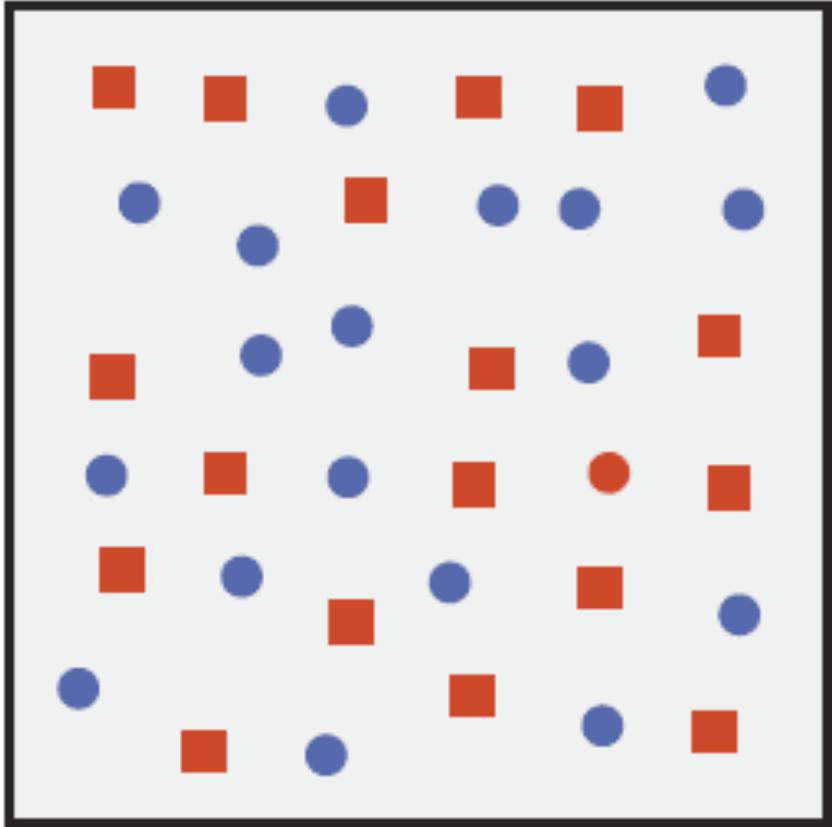
LEFT?

=

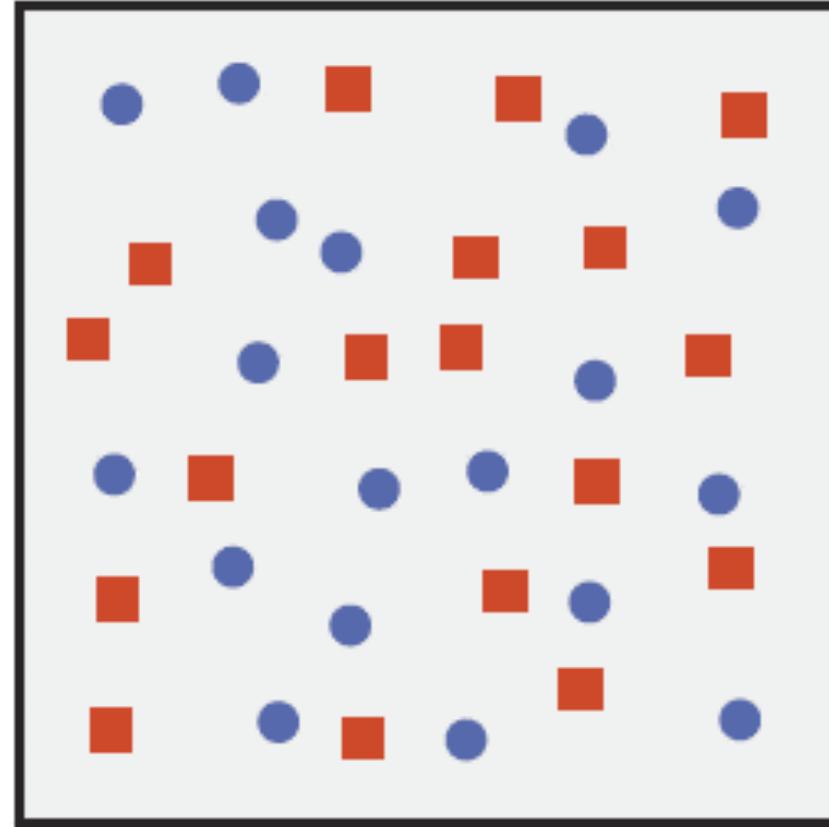
UP!

Which panel has a red dot?

A



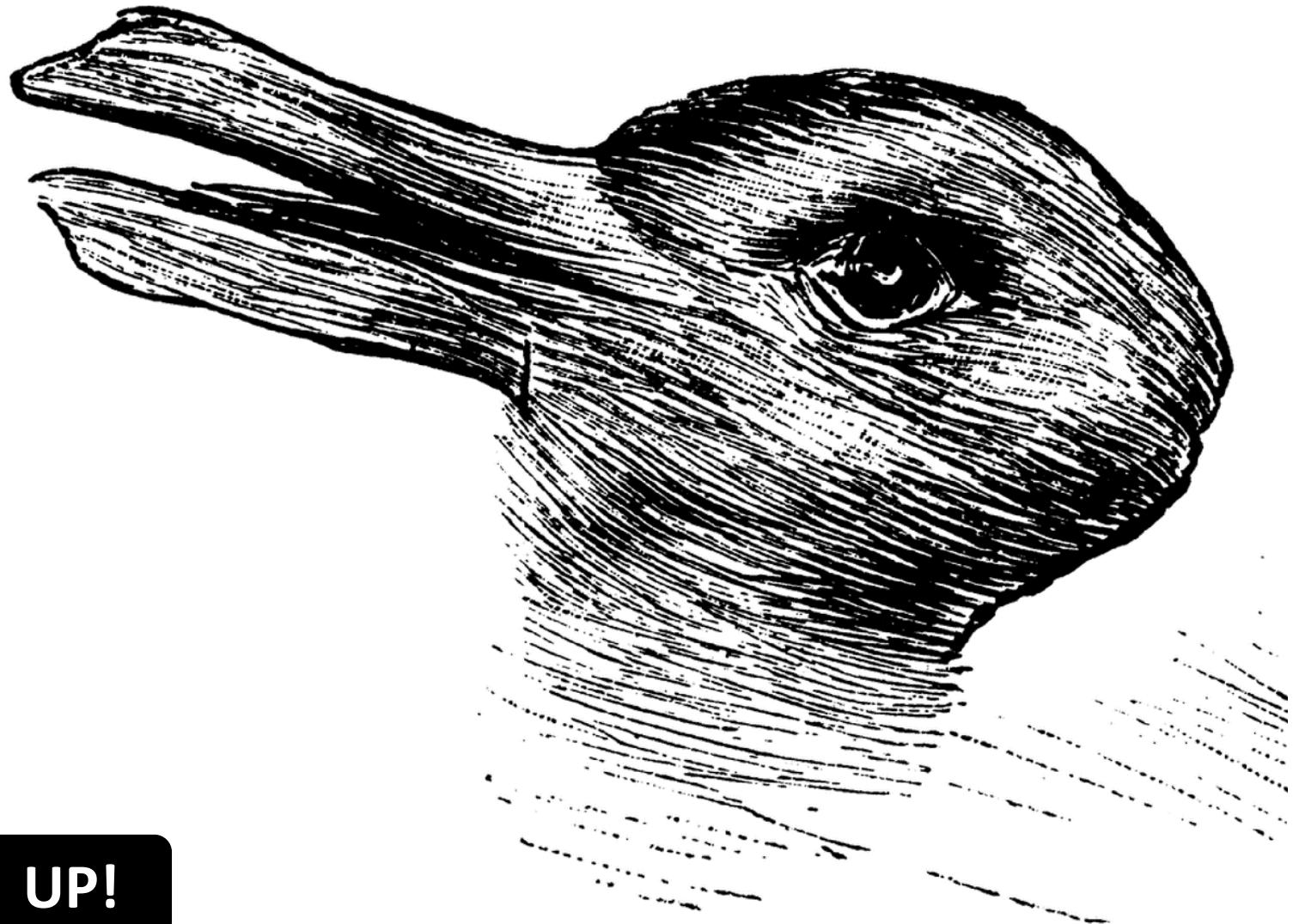
B



LEFT? =

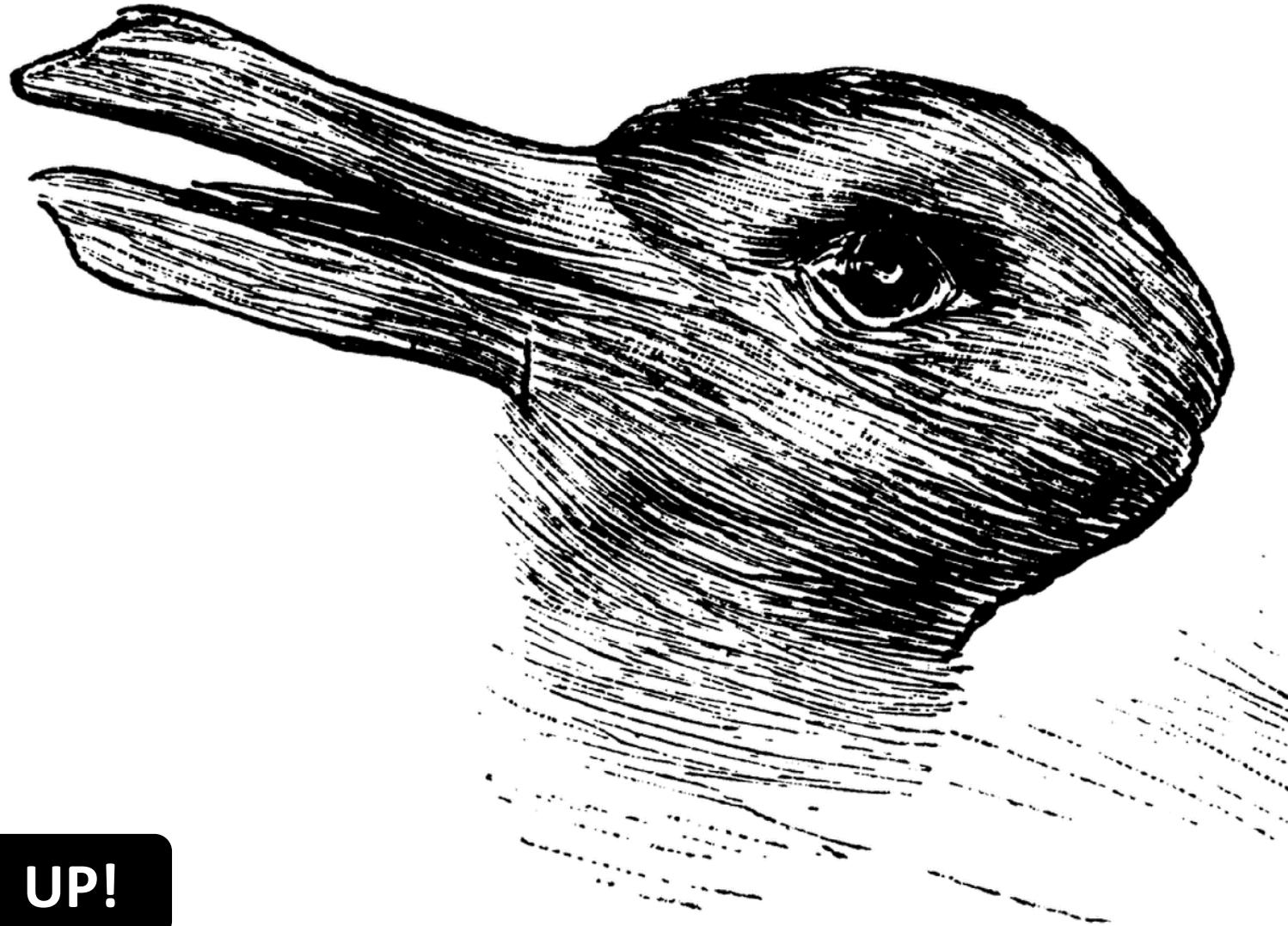
UP!

Do you see a Rabbit?



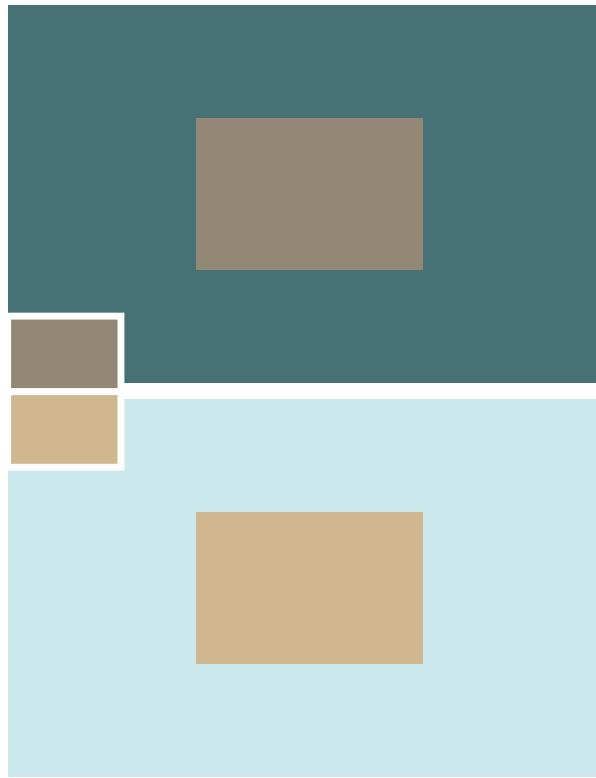
YES? = UP!

Do you see a Duck?



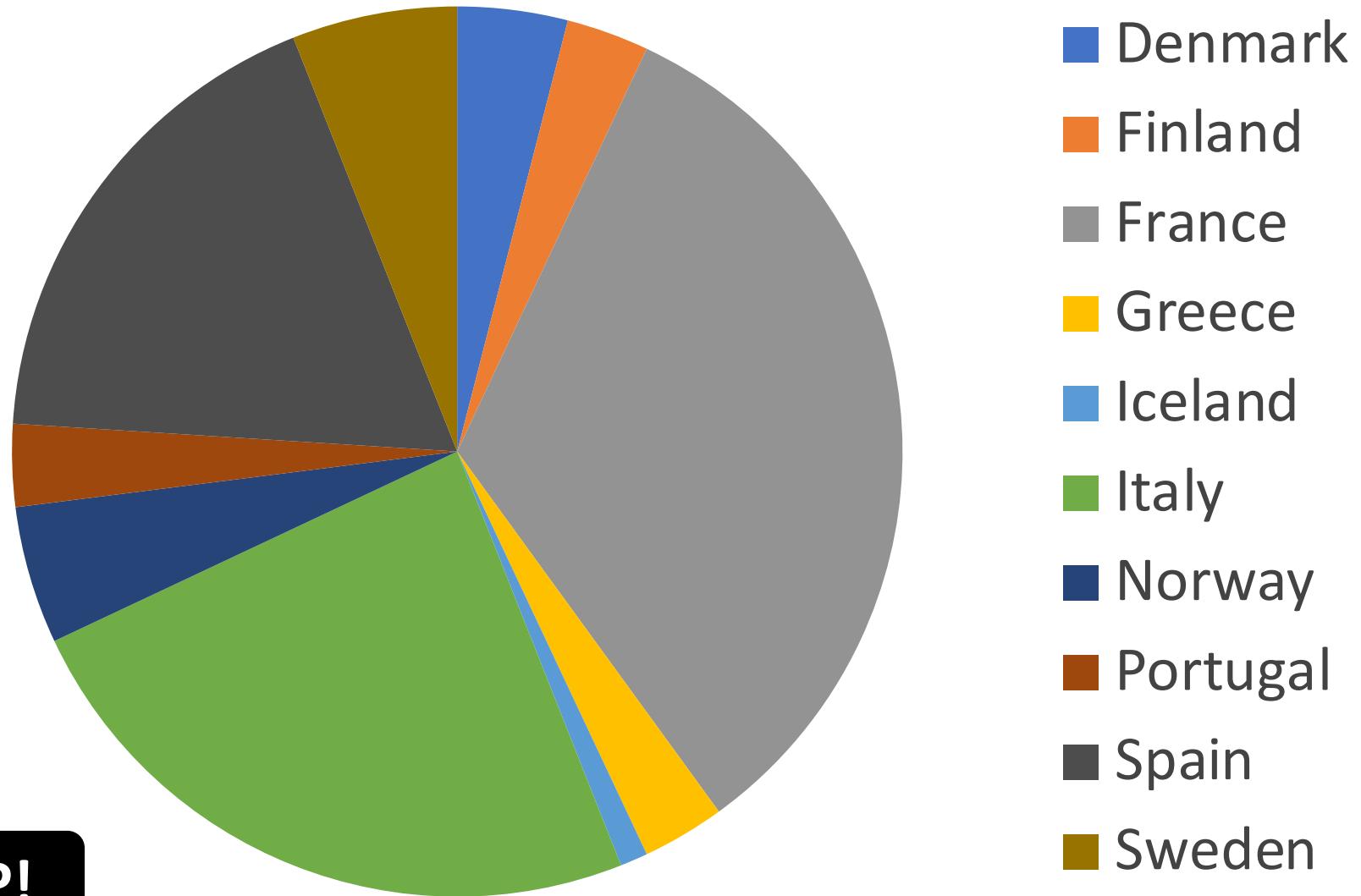
YES? = UP!

The inner squares have the same color



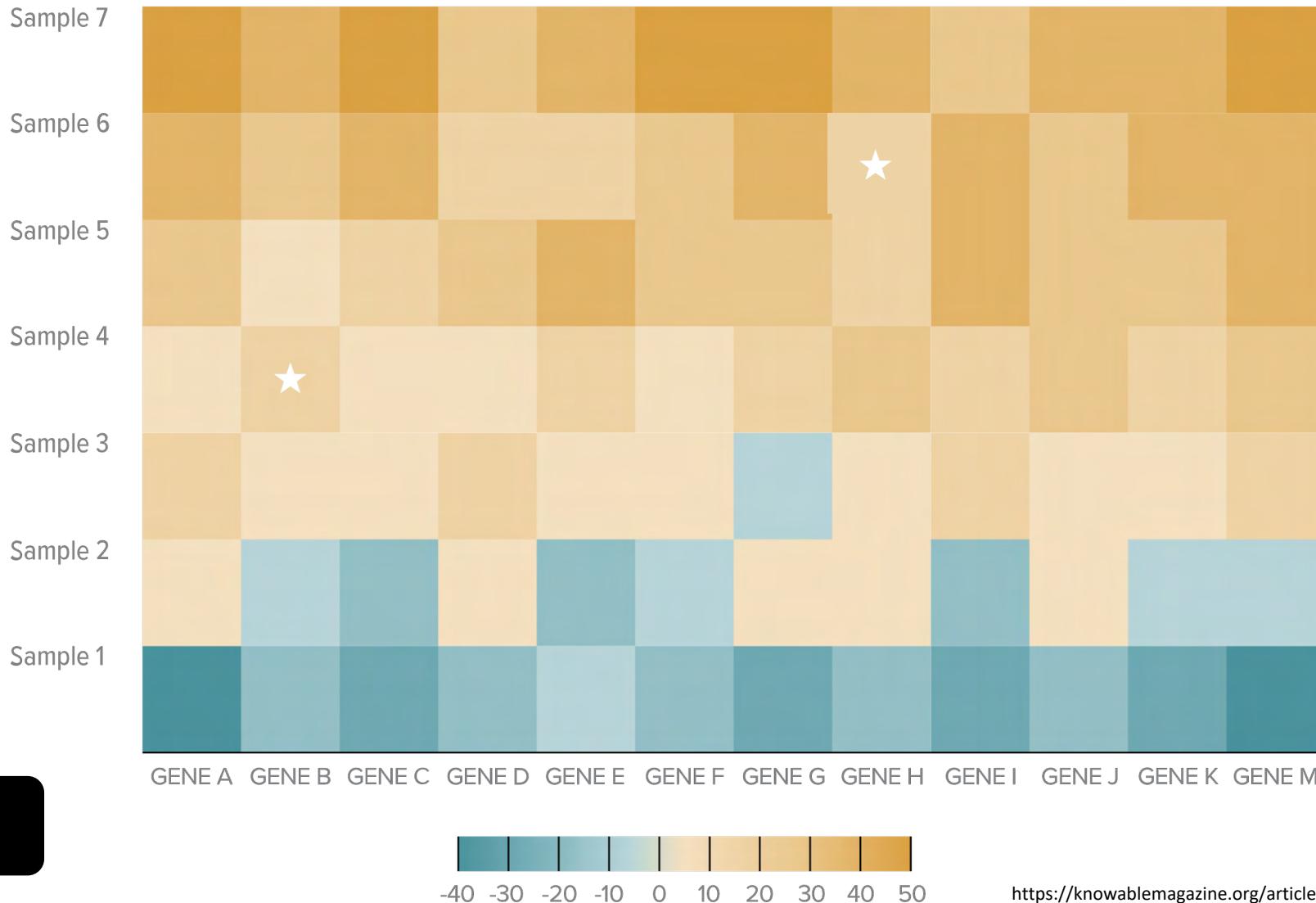
YES? = **UP!**

Norway has more widgets than Finland

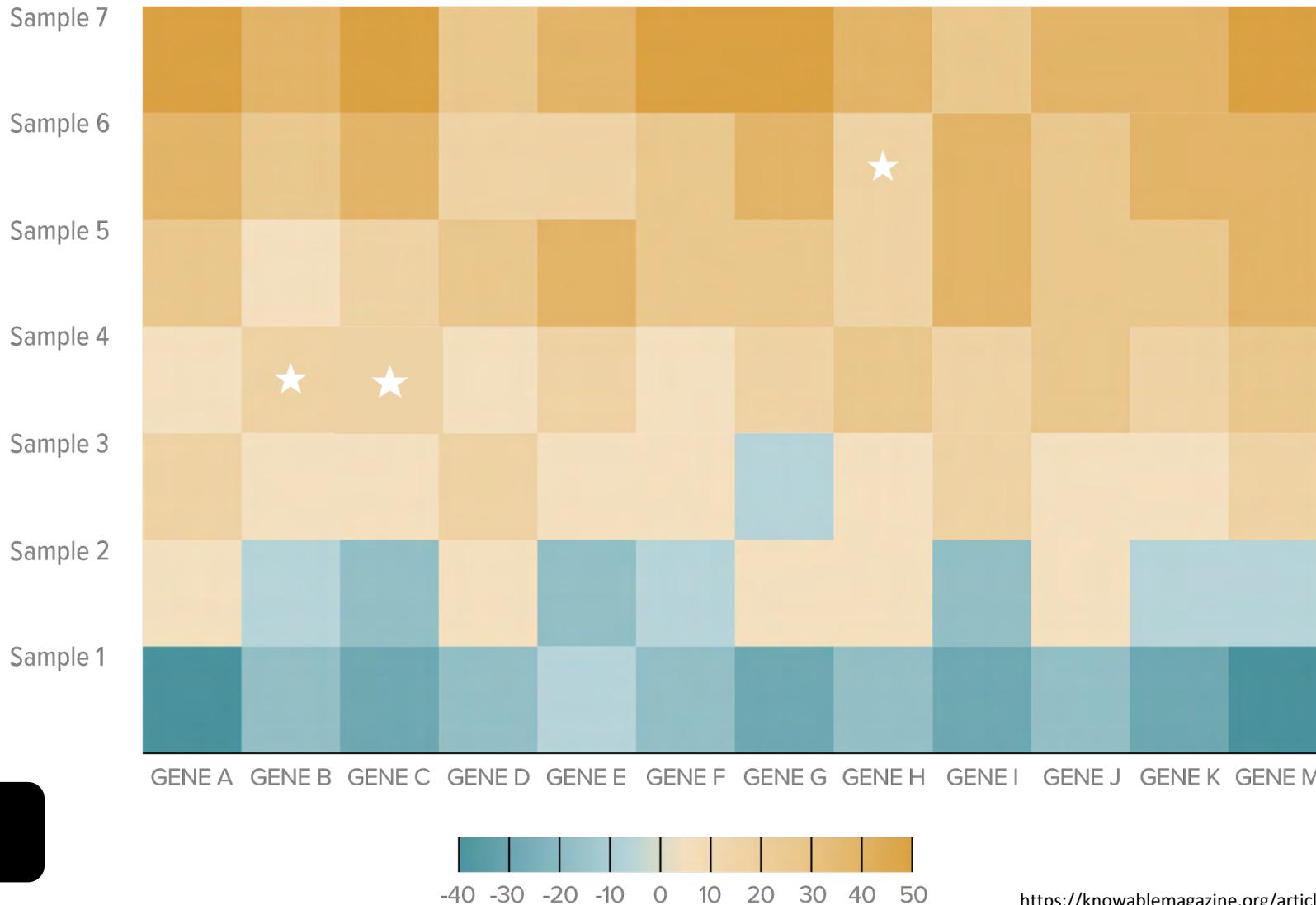


YES? = UP!

The areas with a star have the same color

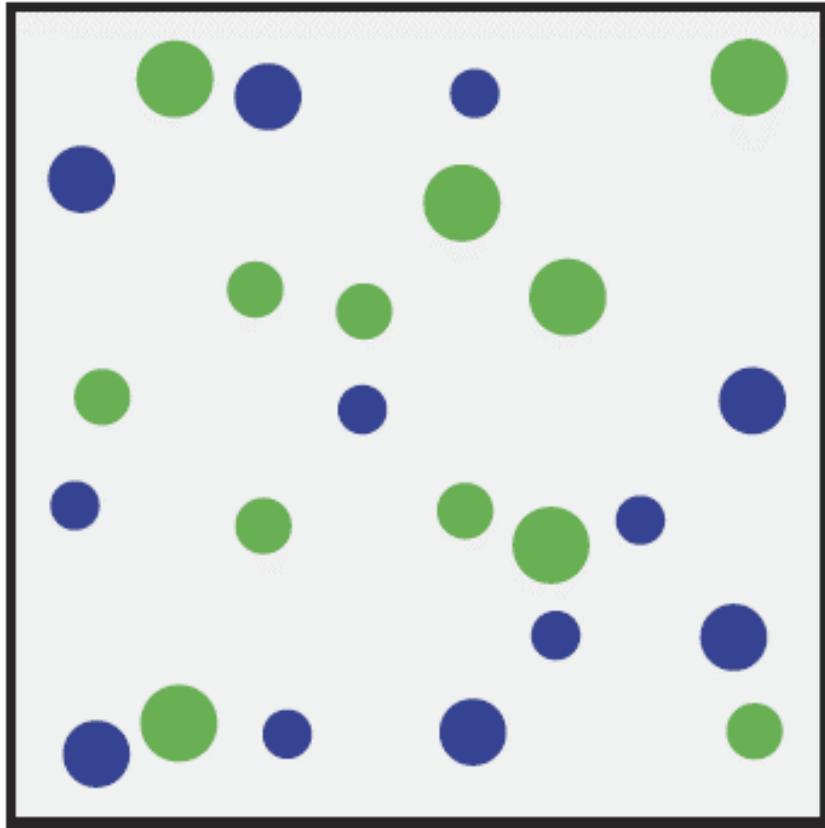


The areas with a star have the same color

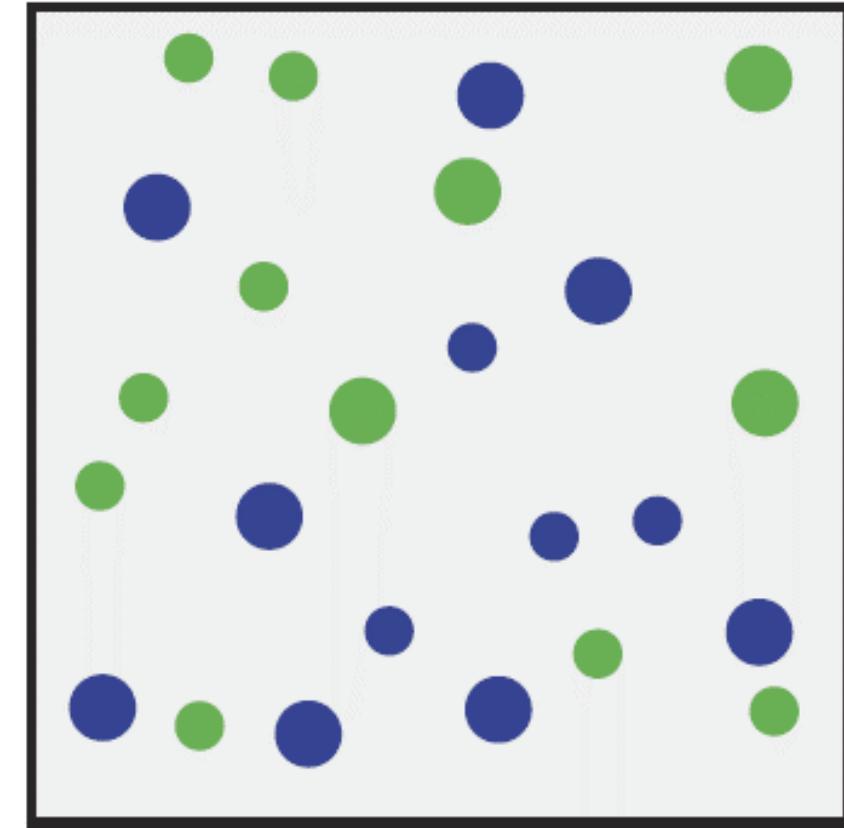


Which panel has the **biggest green** dots?

A



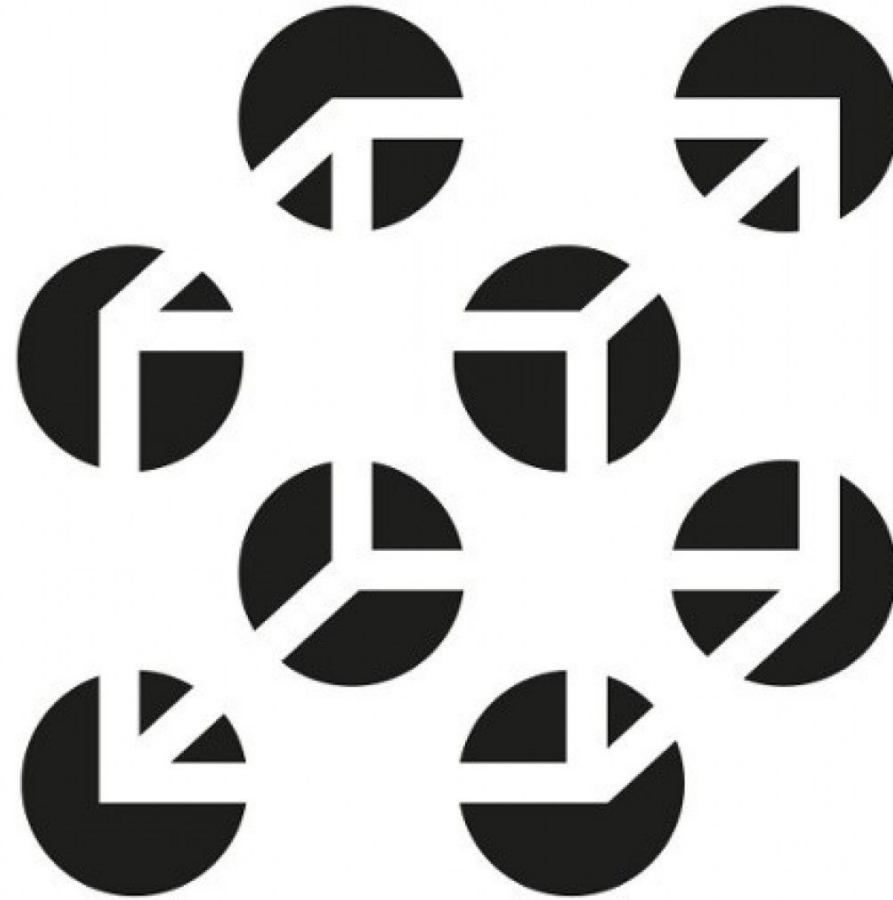
B



LEFT? =

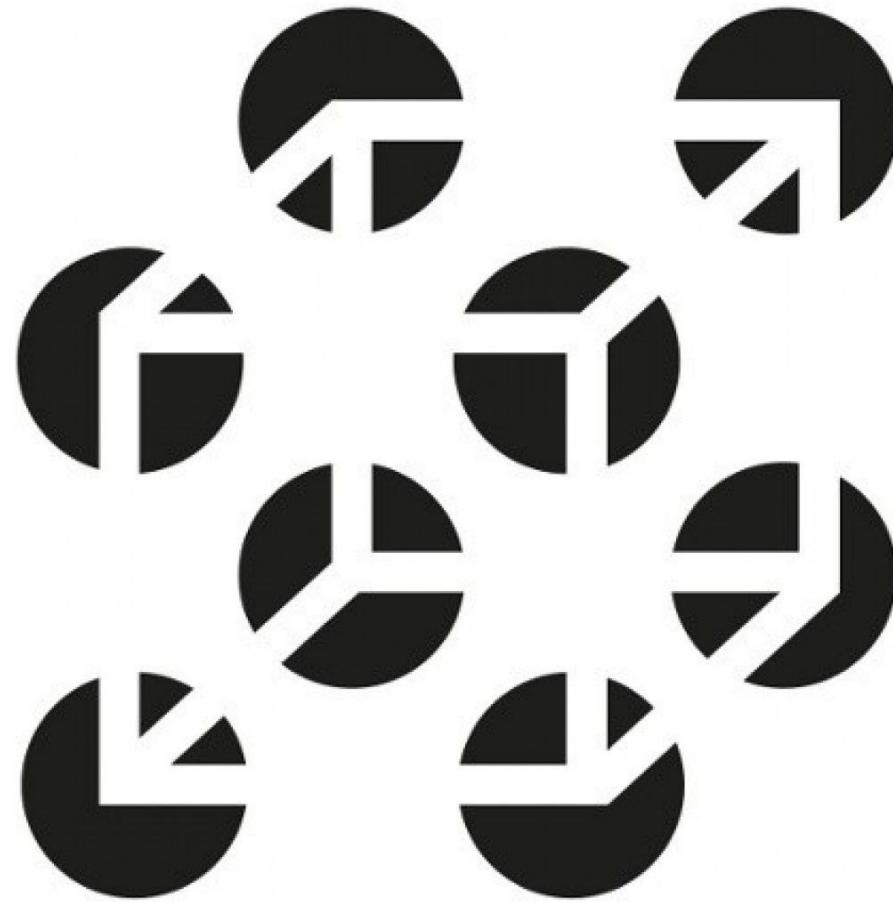
UP!

Do you *see* a cube?



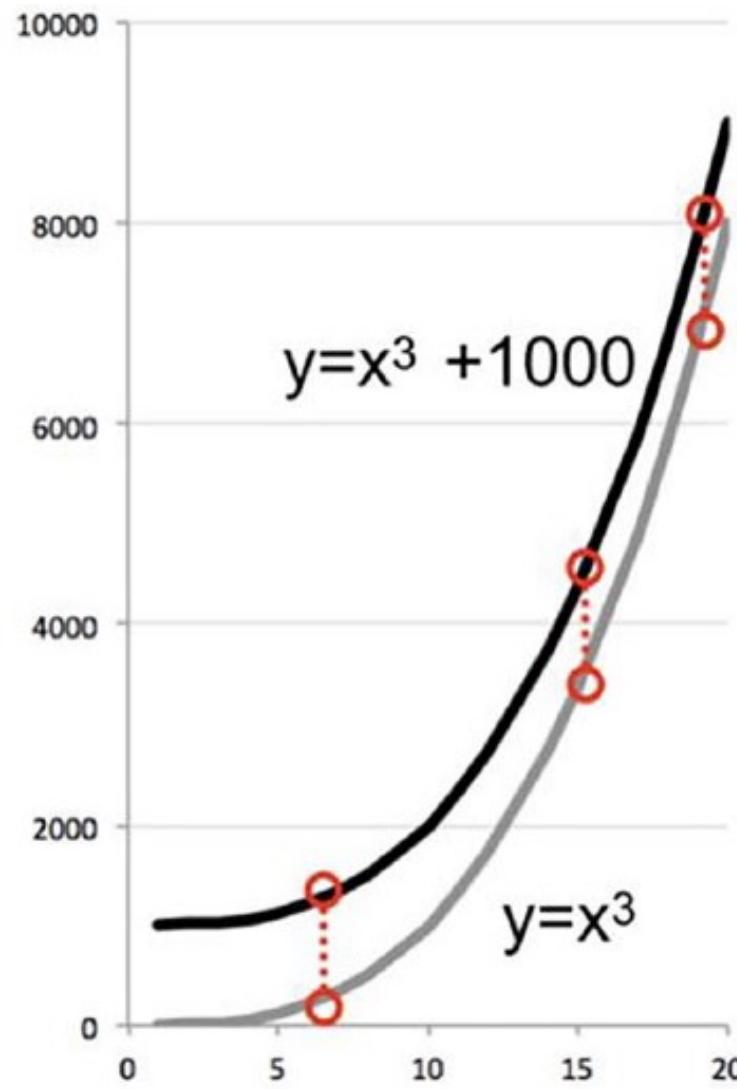
YES? = UP!

Is there a cube?



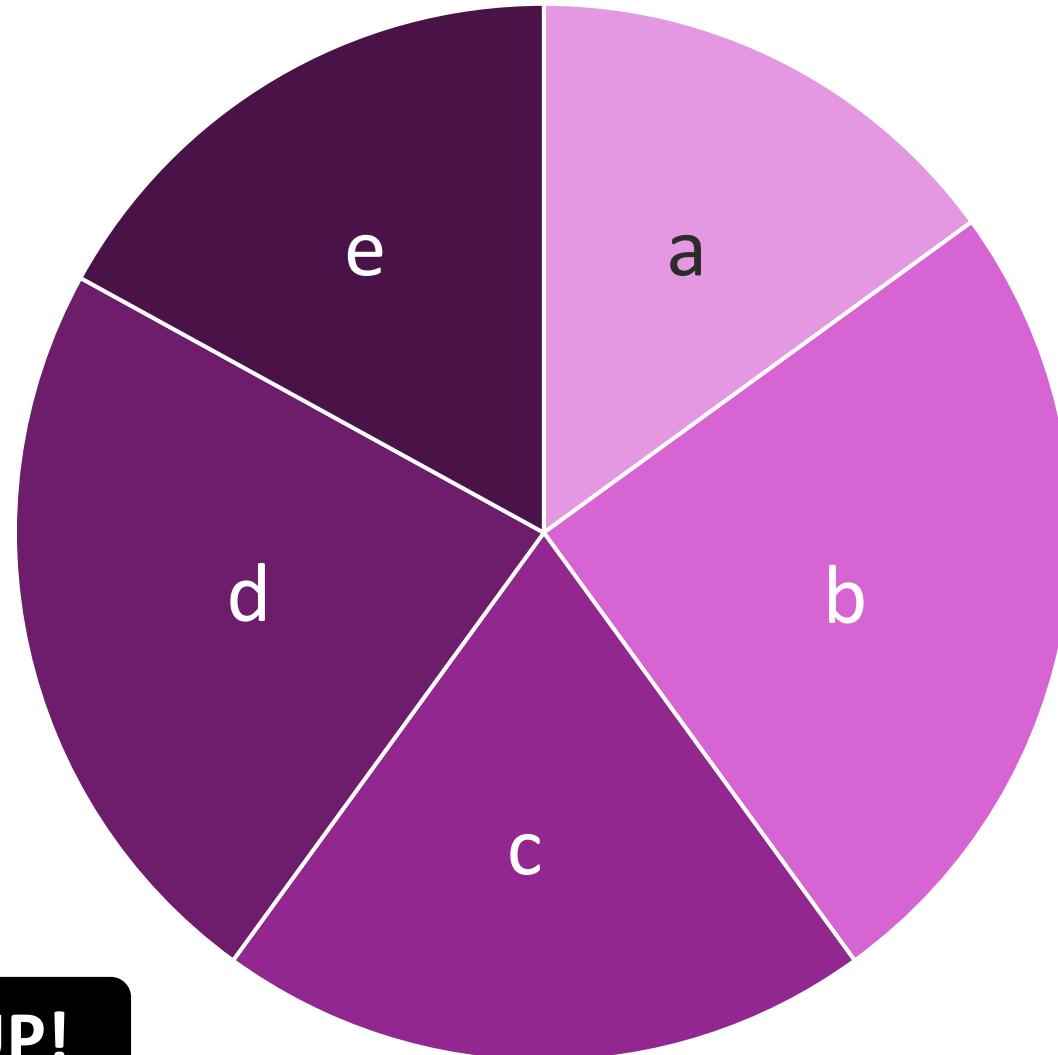
YES? = UP!

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



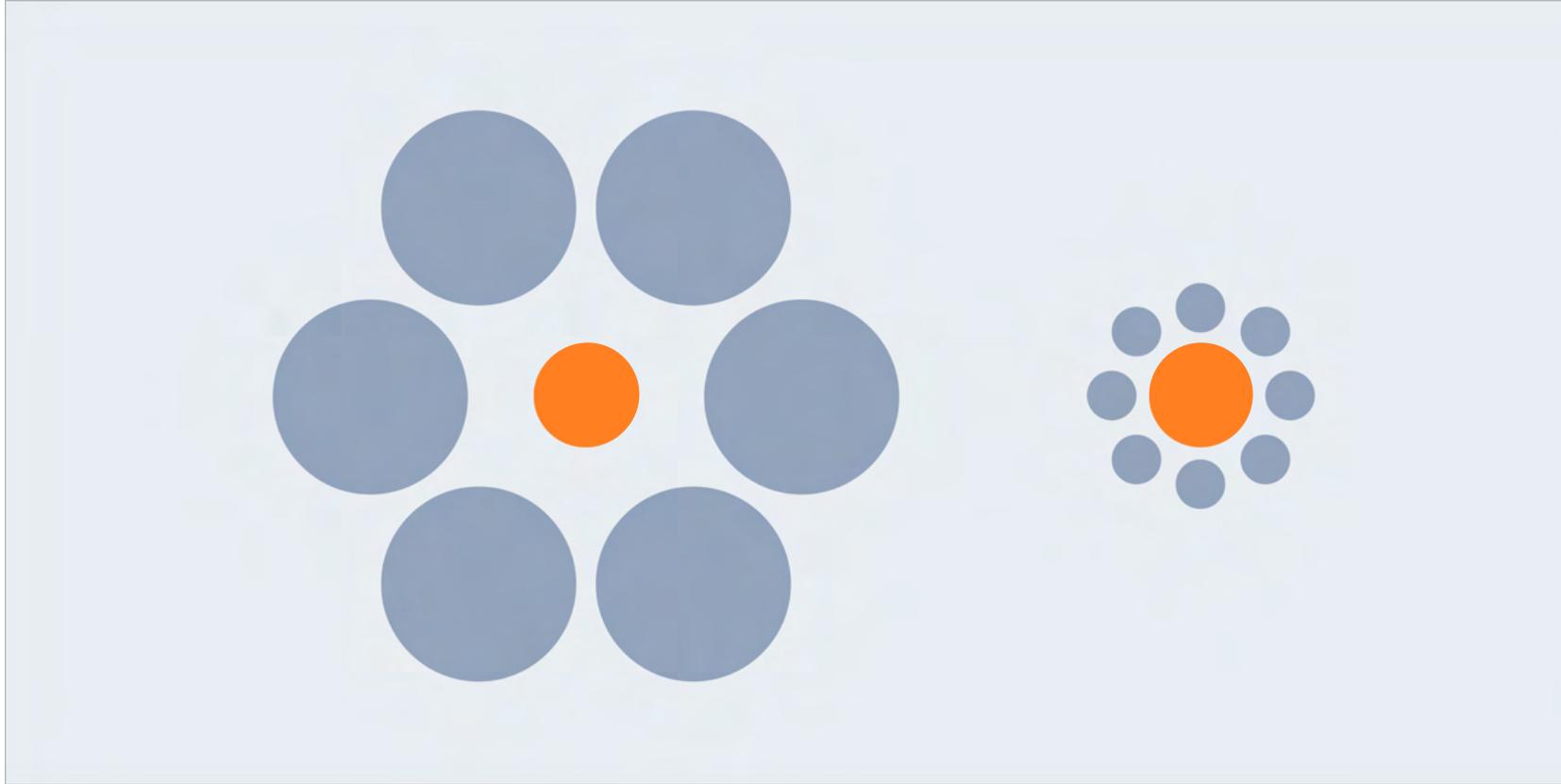
LEFT? = UP!

The order from high to low is b > d > c > e > a



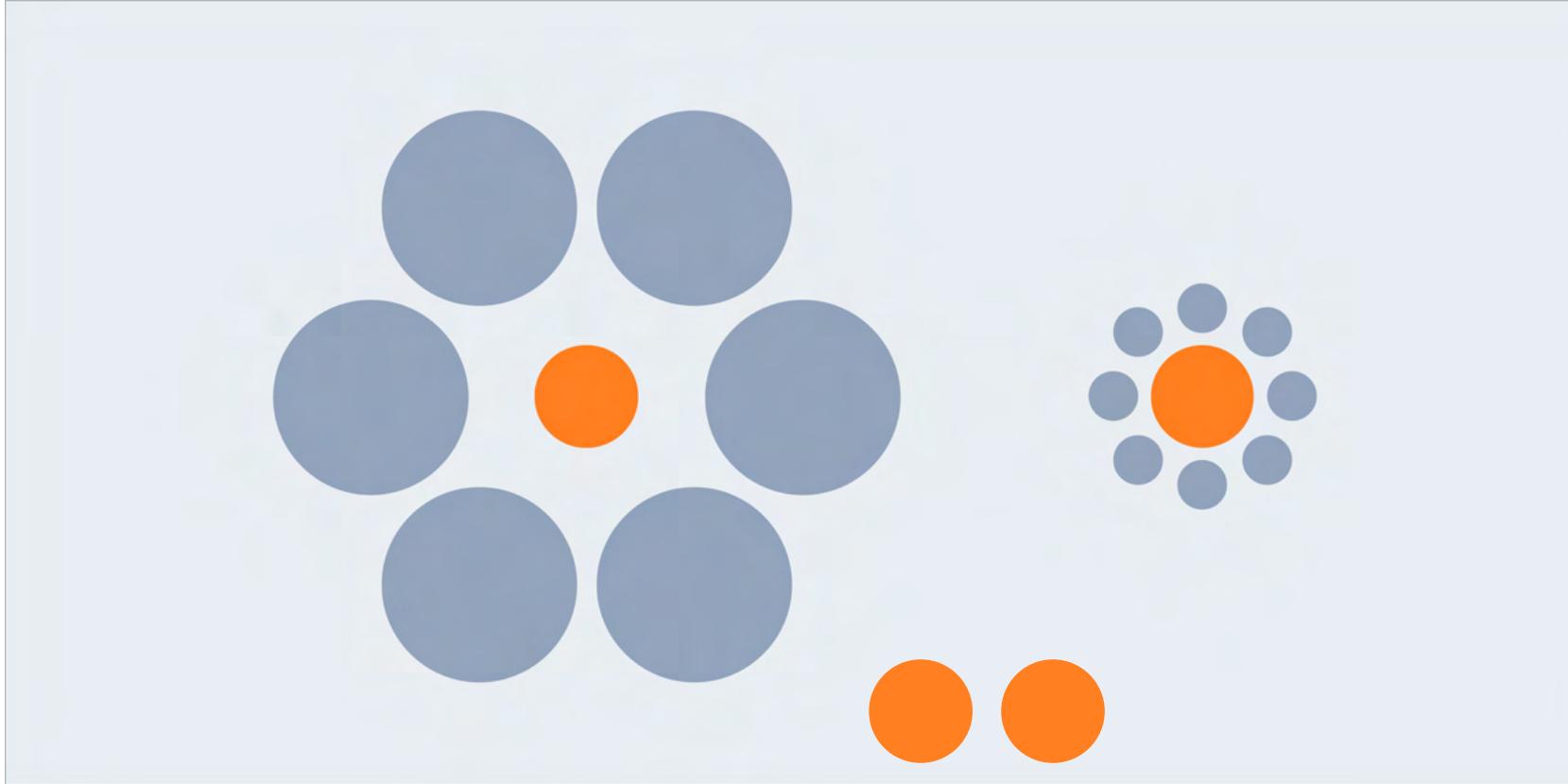
YES? = UP!

Which orange circle is smaller?



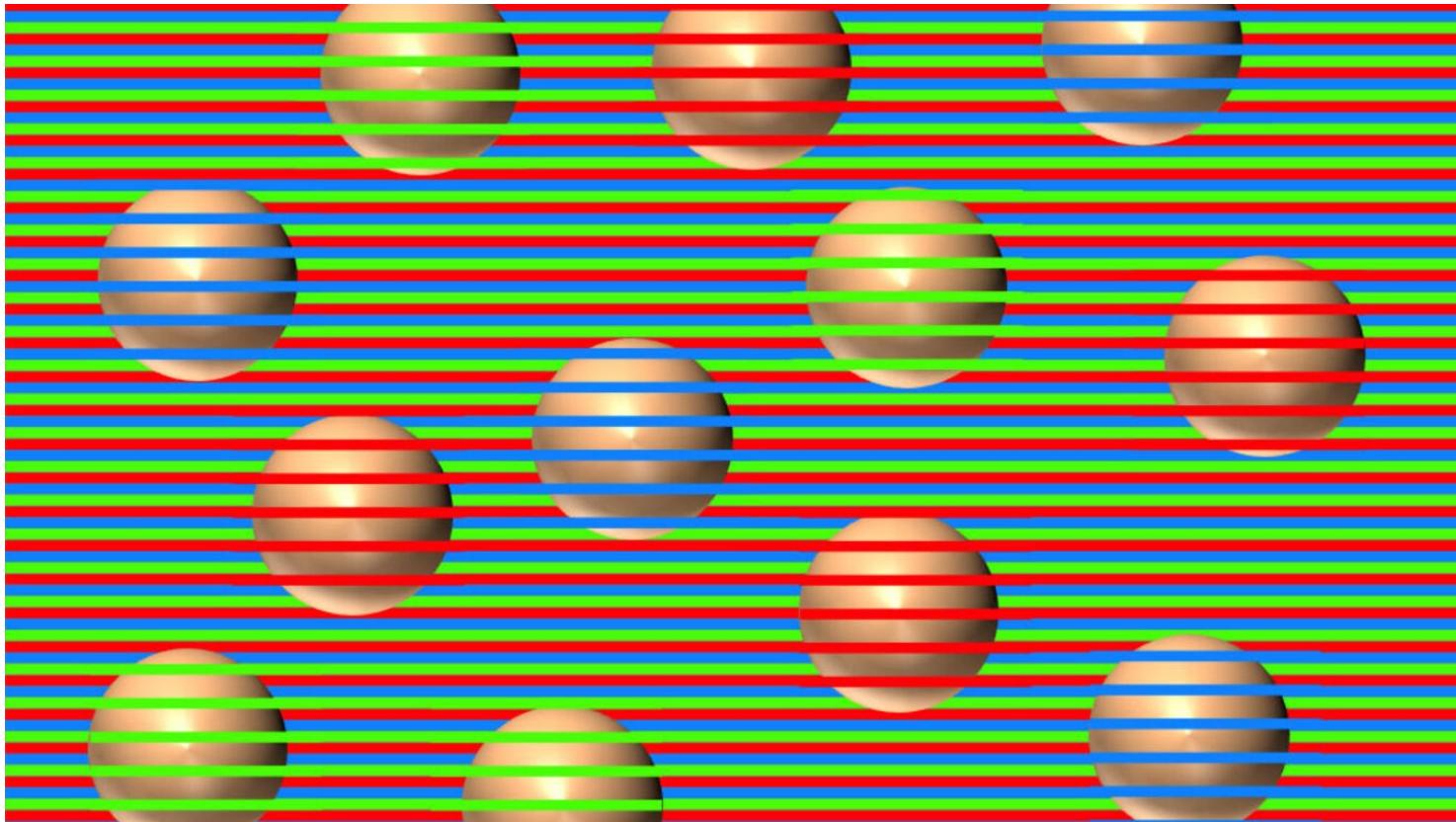
LEFT? = UP!

Which orange circle is smaller?



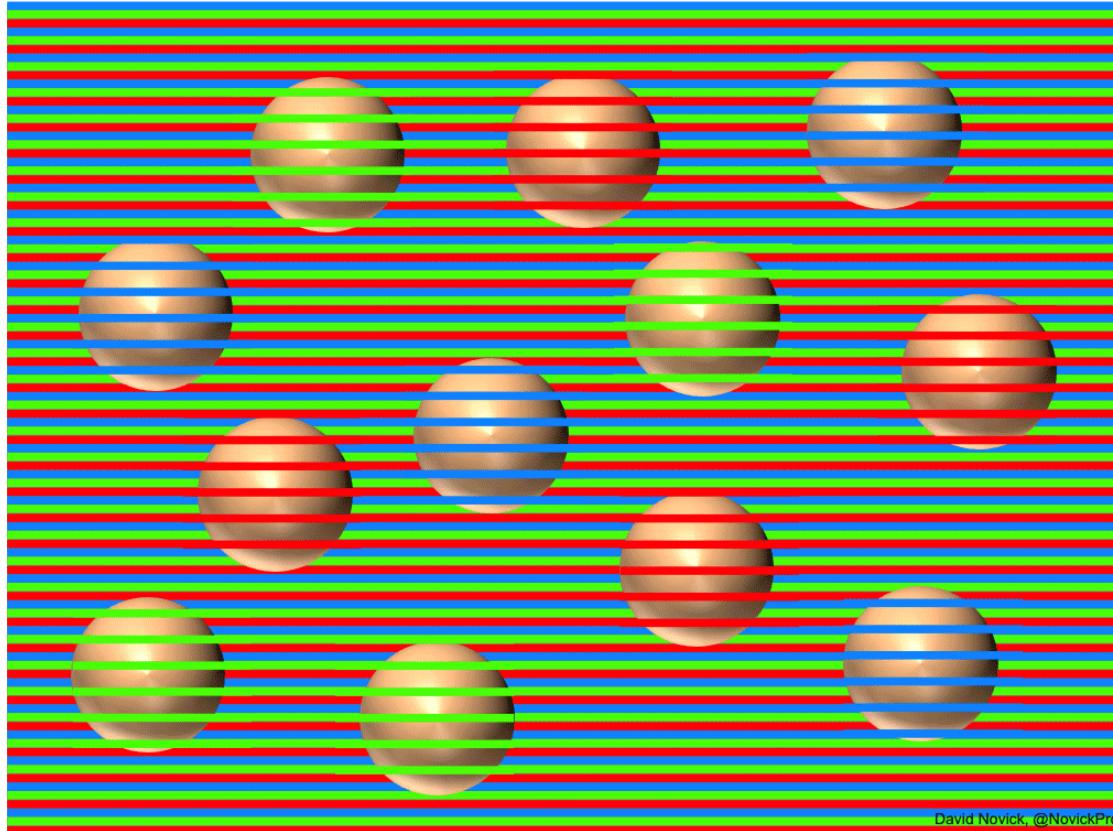
LEFT? = UP!

These spheres have the same color



YES? = **UP!**

These spheres have the same color



David Novick, @NovickProf

YES? = UP!

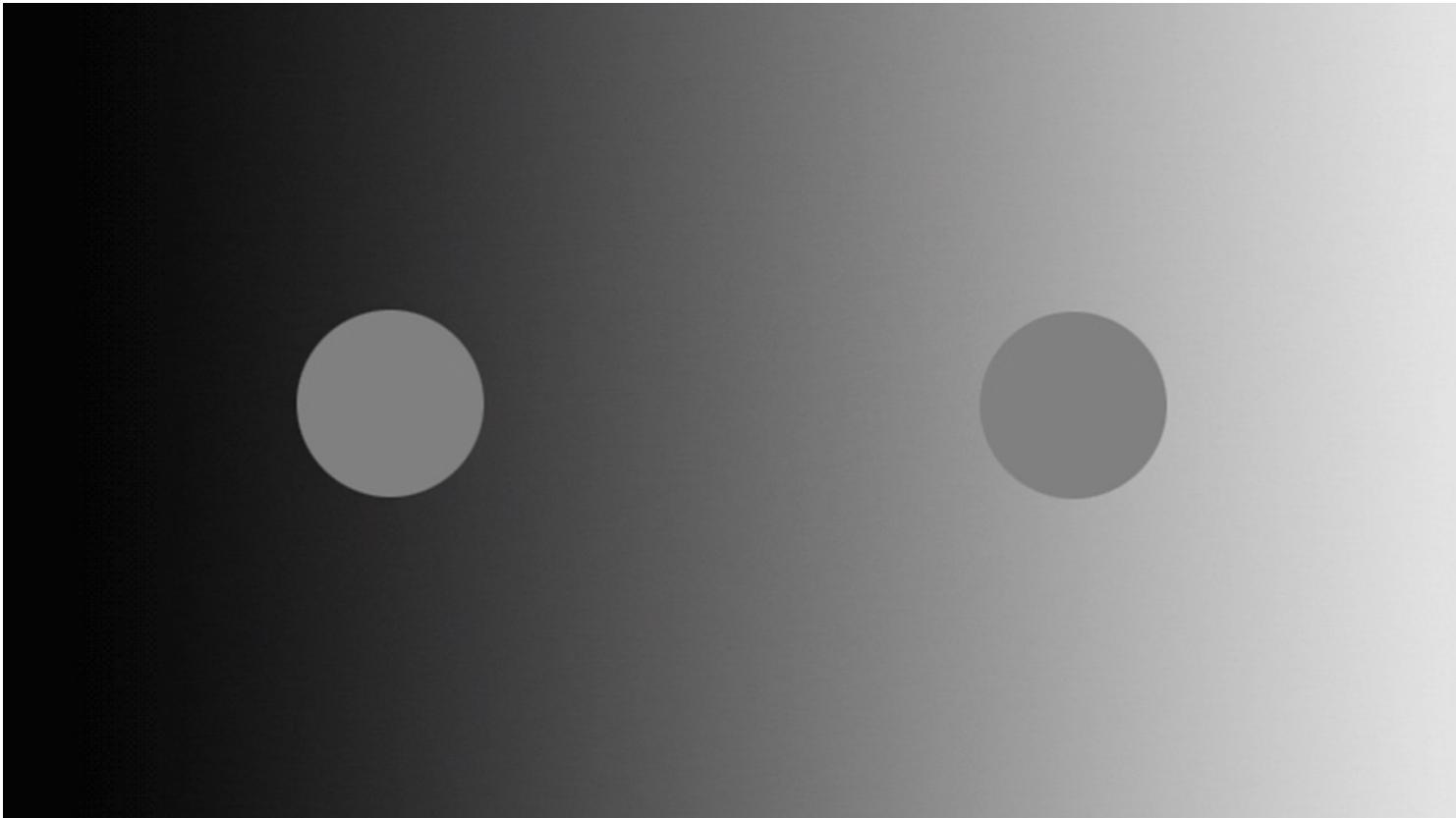
Can you read this?

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.

YES?

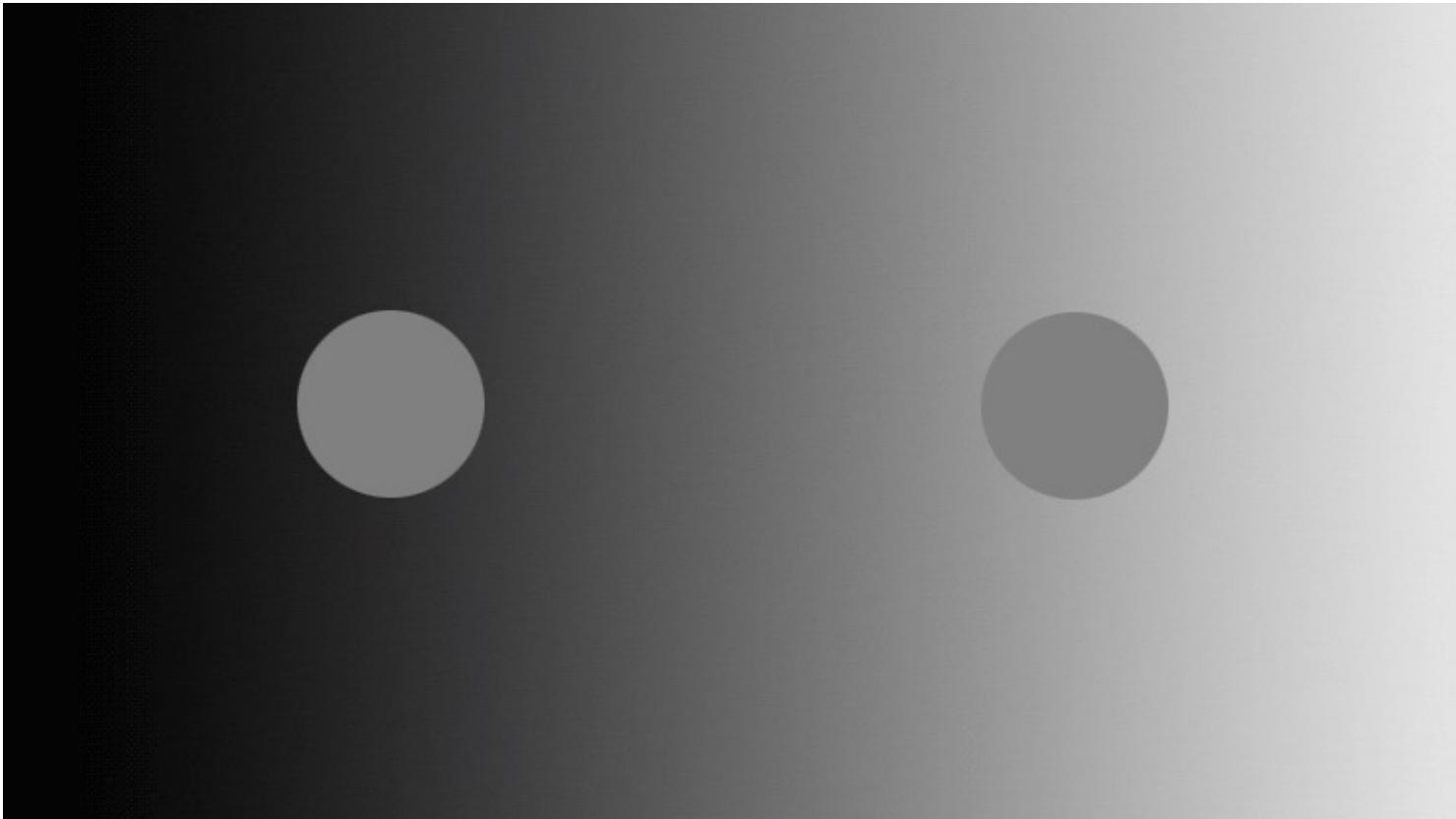
= UP!

Which dot is darker?

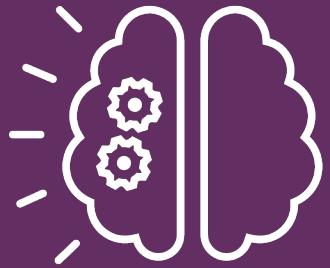


LEFT? = UP!

Are you sure...?



LEFT? = UP!

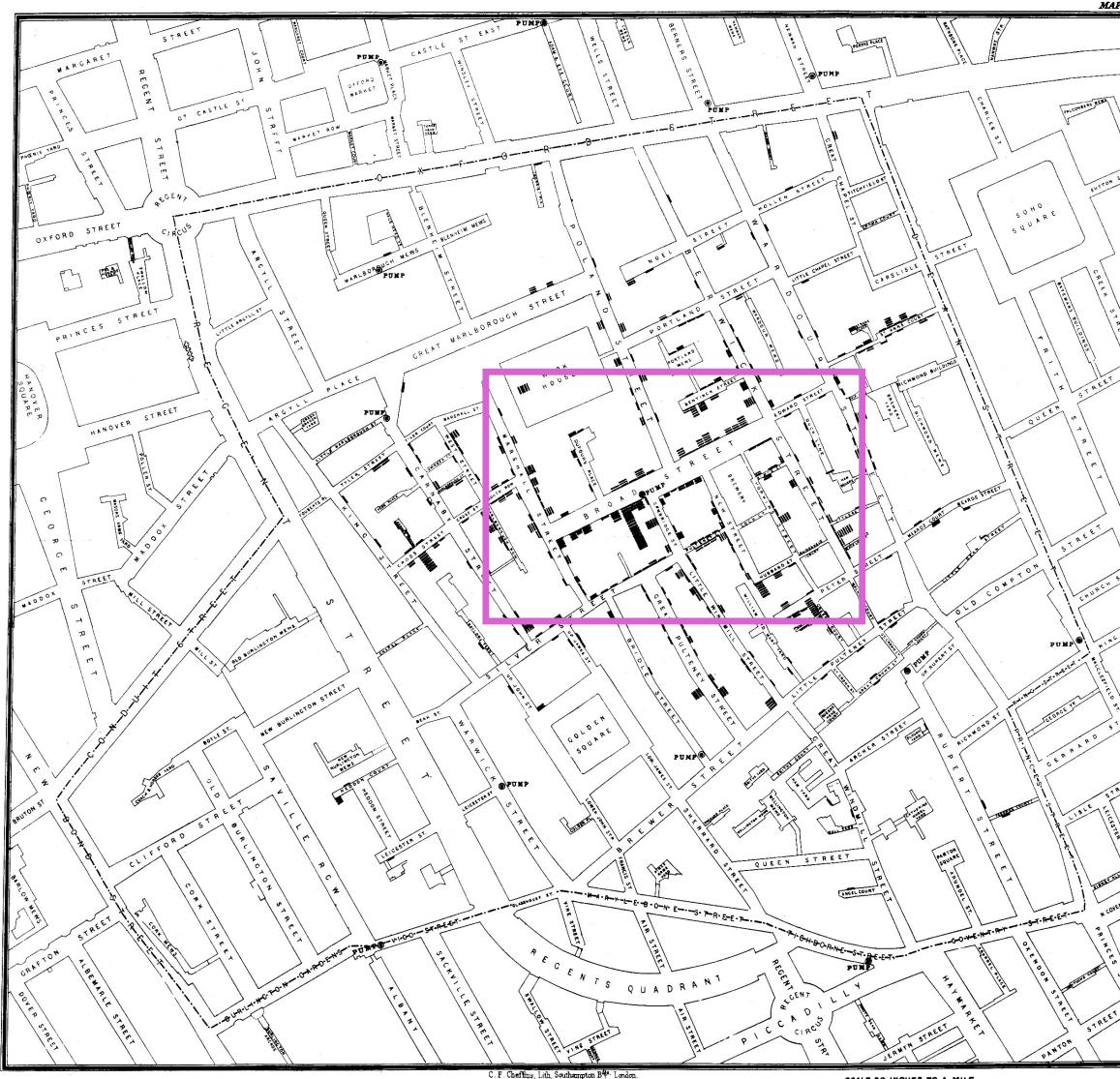


*“It’s not what you look at that
matters, it’s what you see”*

Henry David Thoreau

Why visual communication?

1854: The cholera map from John Snow



1854: The cholera map from John Snow



What can you see in these numbers?

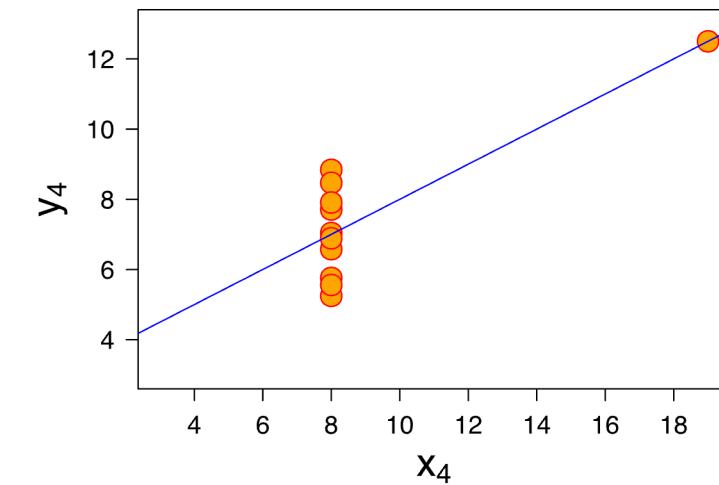
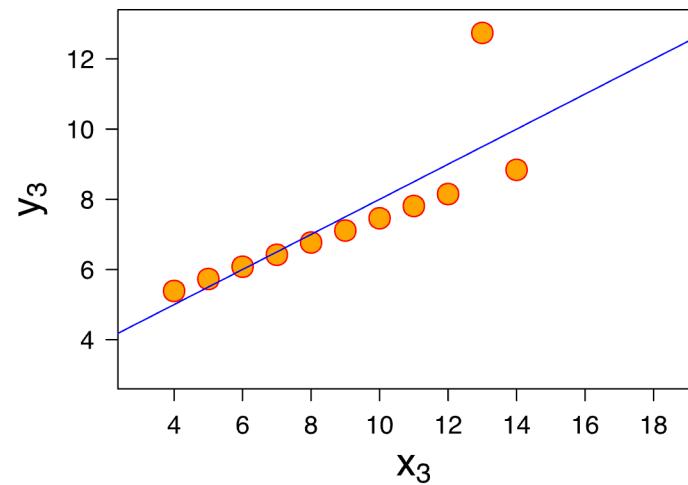
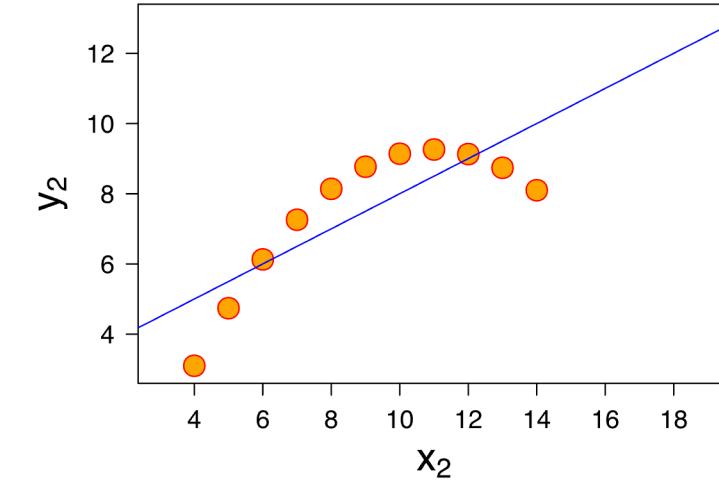
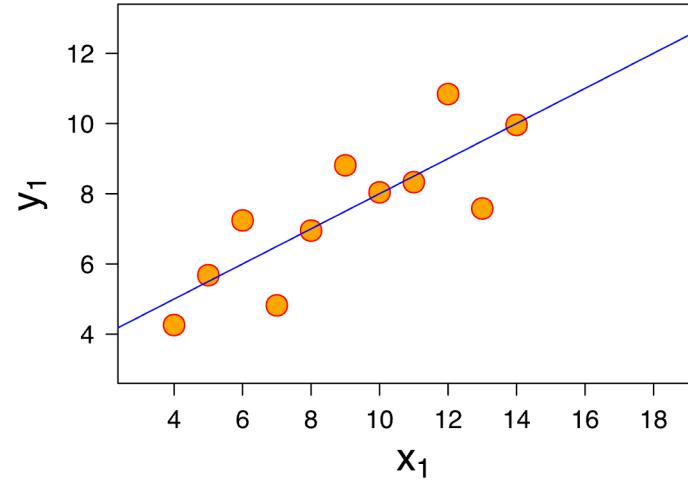
Anscombe's quartet:
4 datasets
11 x-values
11 y-values
Same total
Same mean
Same standard deviation

	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

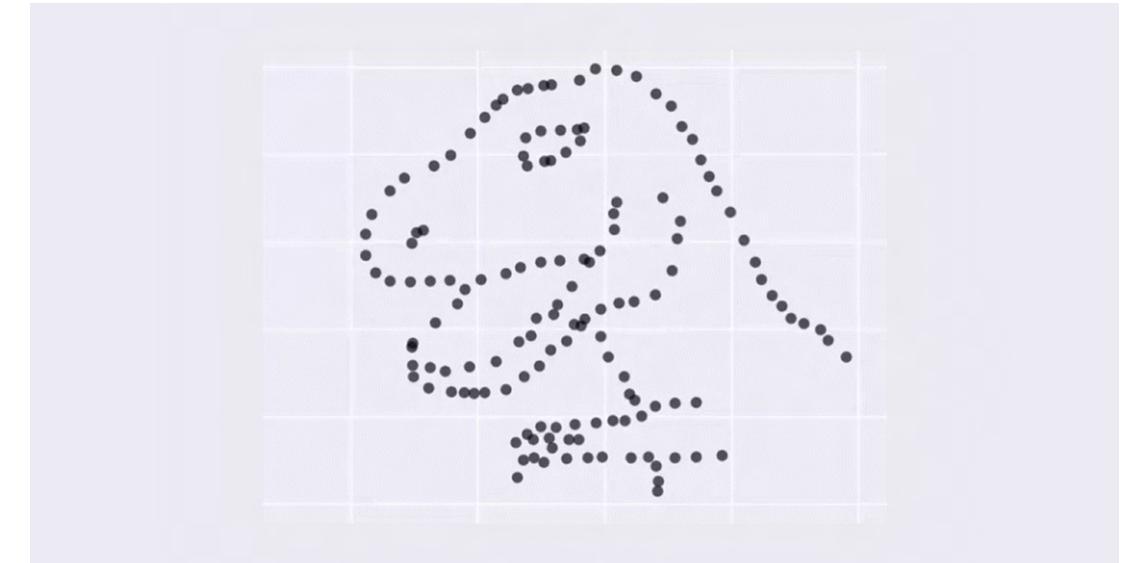
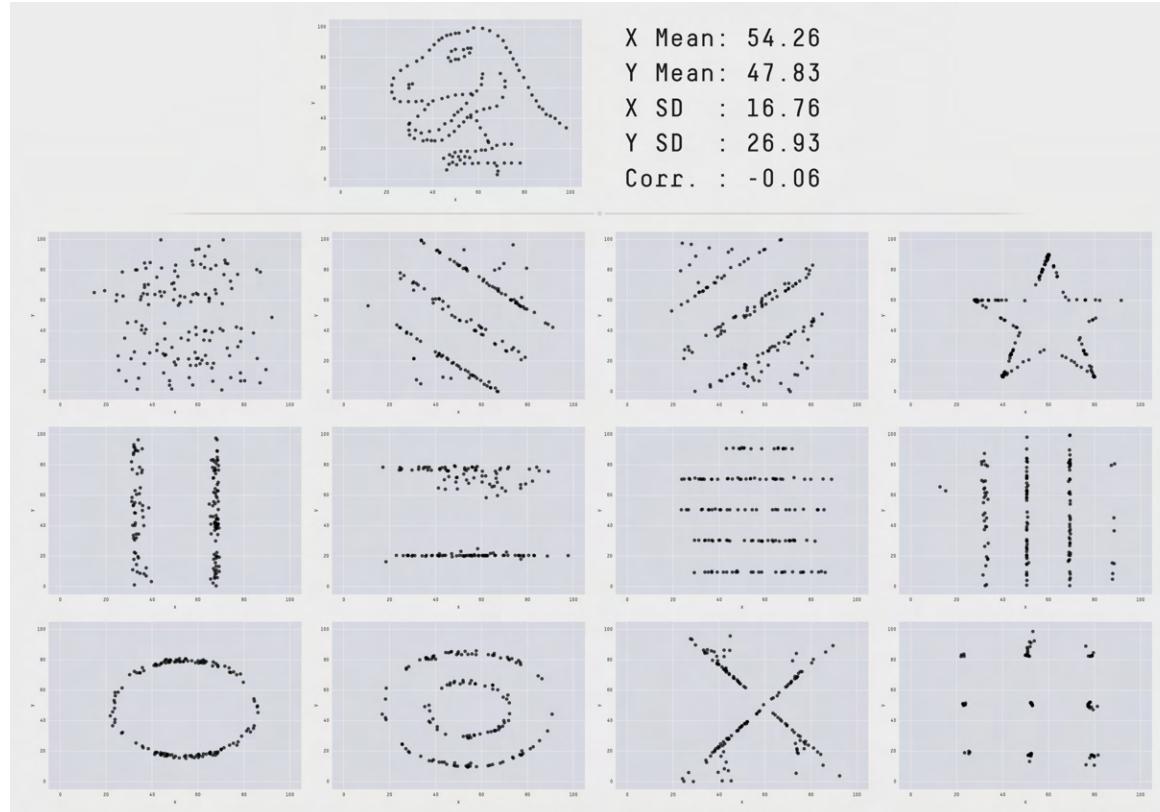
A picture is worth a thousand words!

Anscombe's quartet:

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

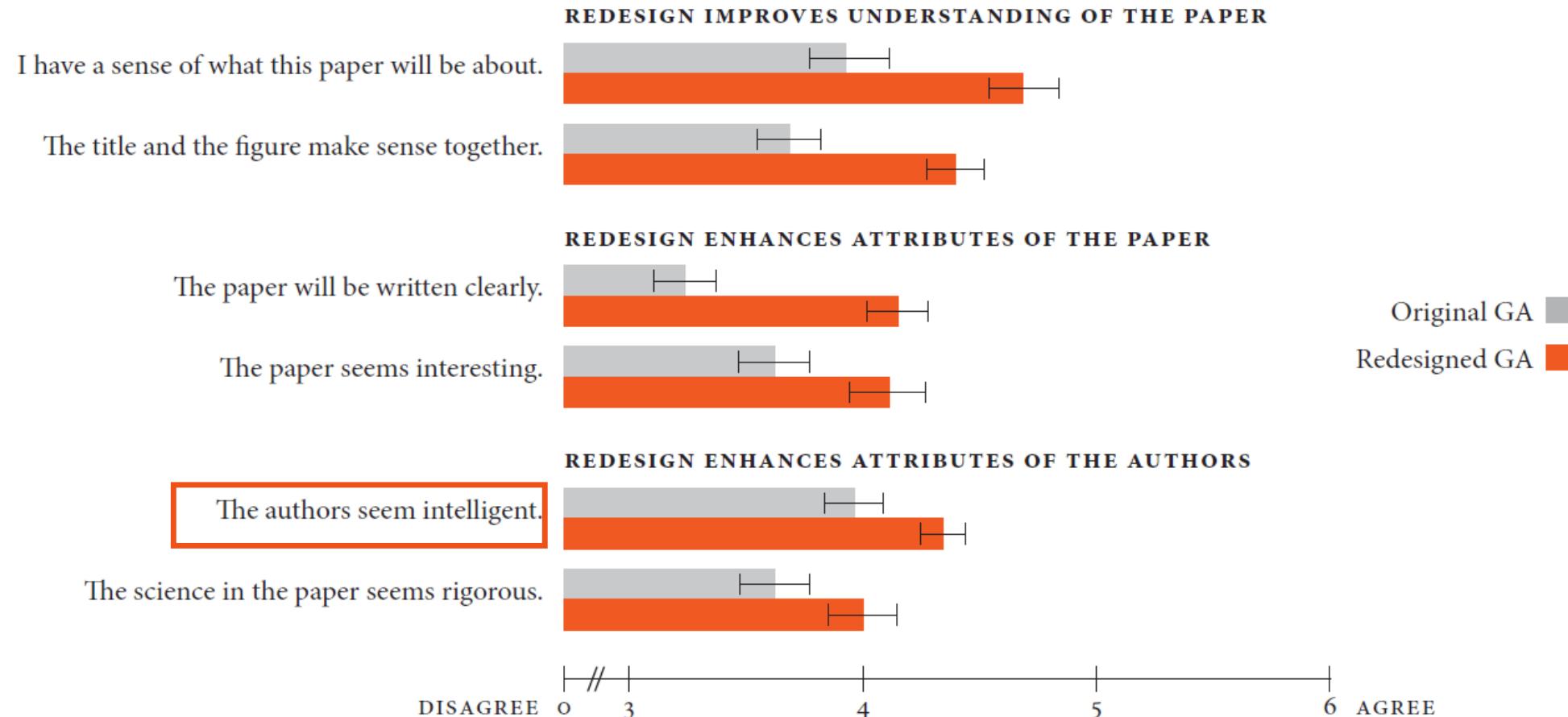


The Datasaurus Dozen 😊: Never trust summary statistics alone!



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?



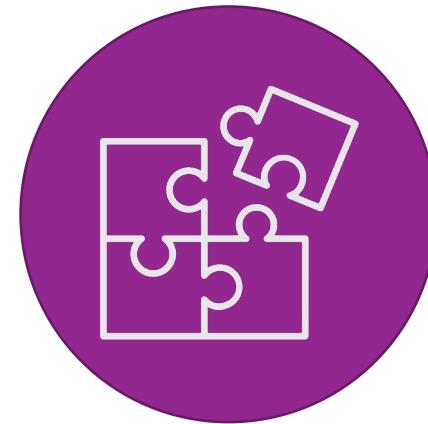
Provide **insight**

-
- Publications
 - Oral presentations
 - Poster presentations



Grab **attention**

-
- Graphical abstracts
 - # Social media
 - General media

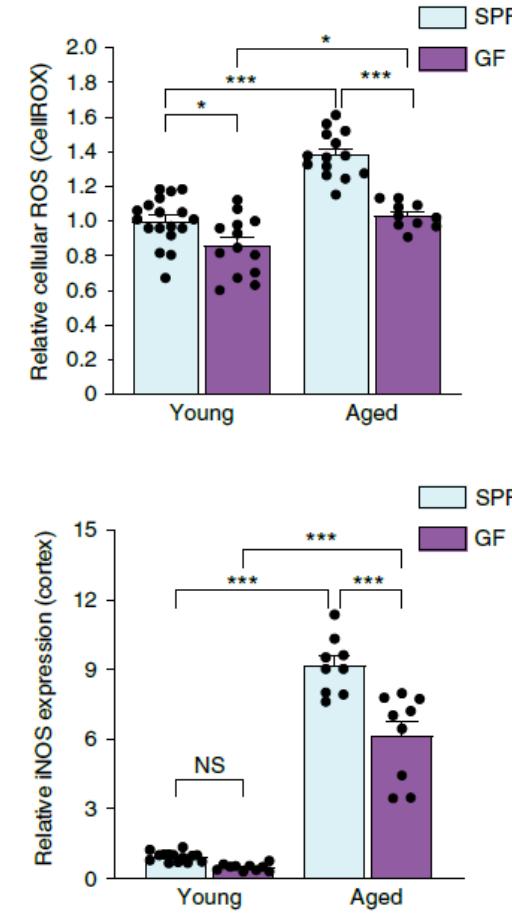
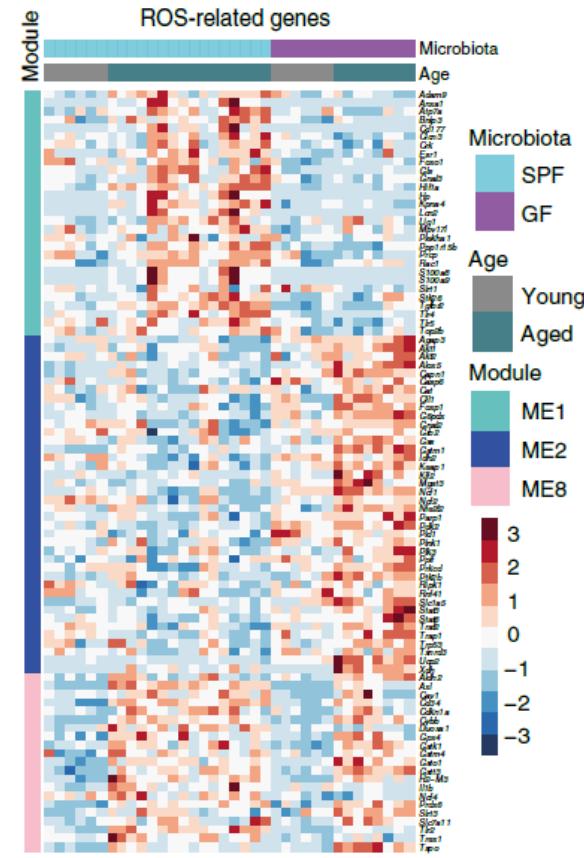
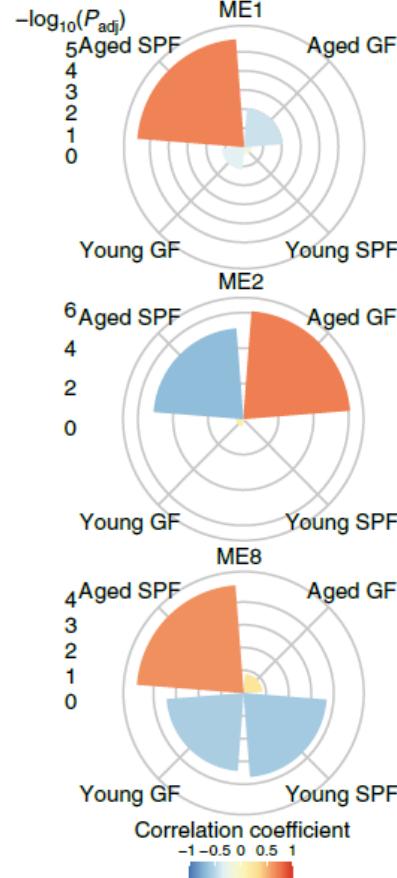


Show **competence**

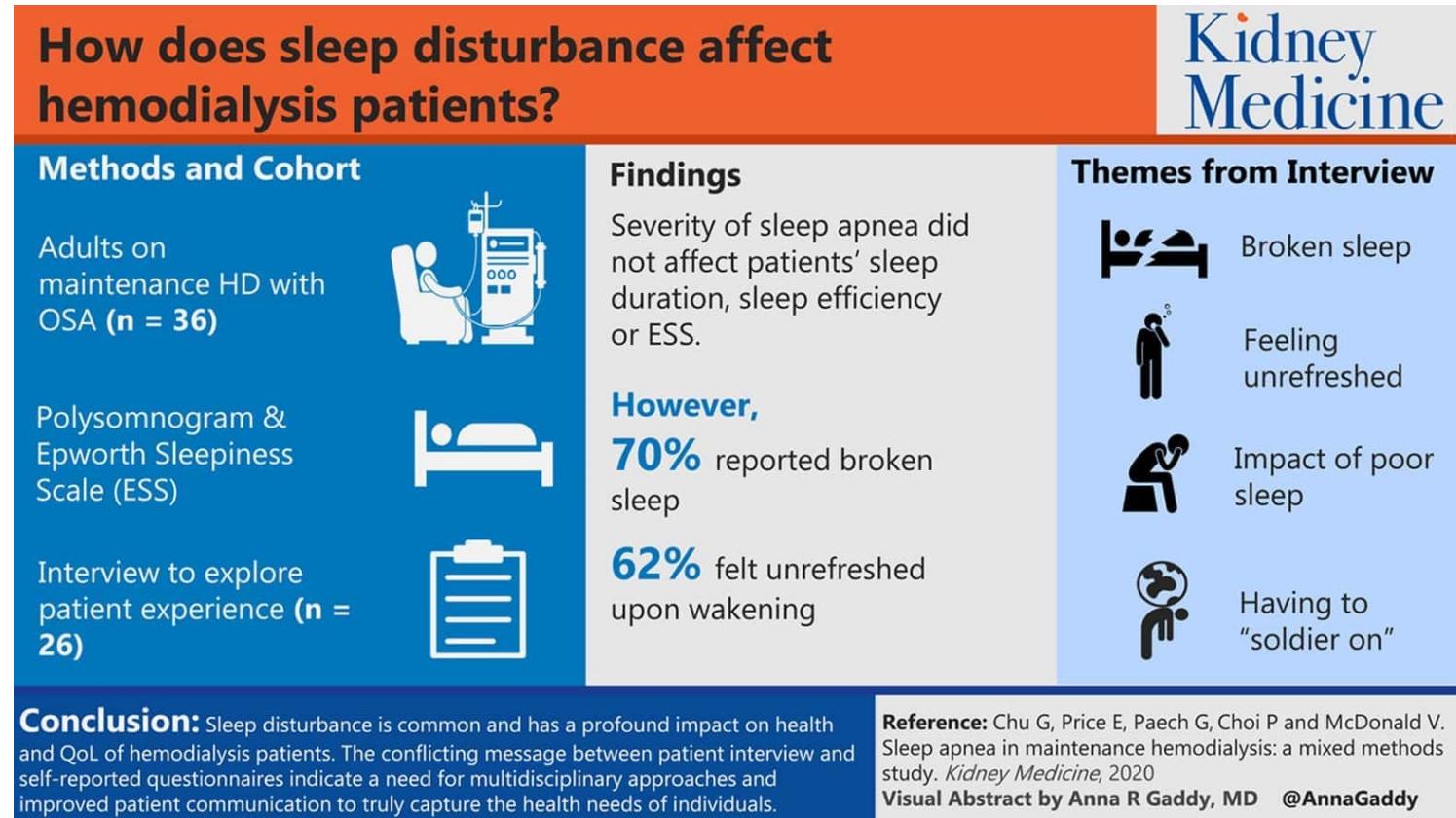
-
- Grant applications
 - Portfolio / CV
 - Pitch

What is visual communication ?

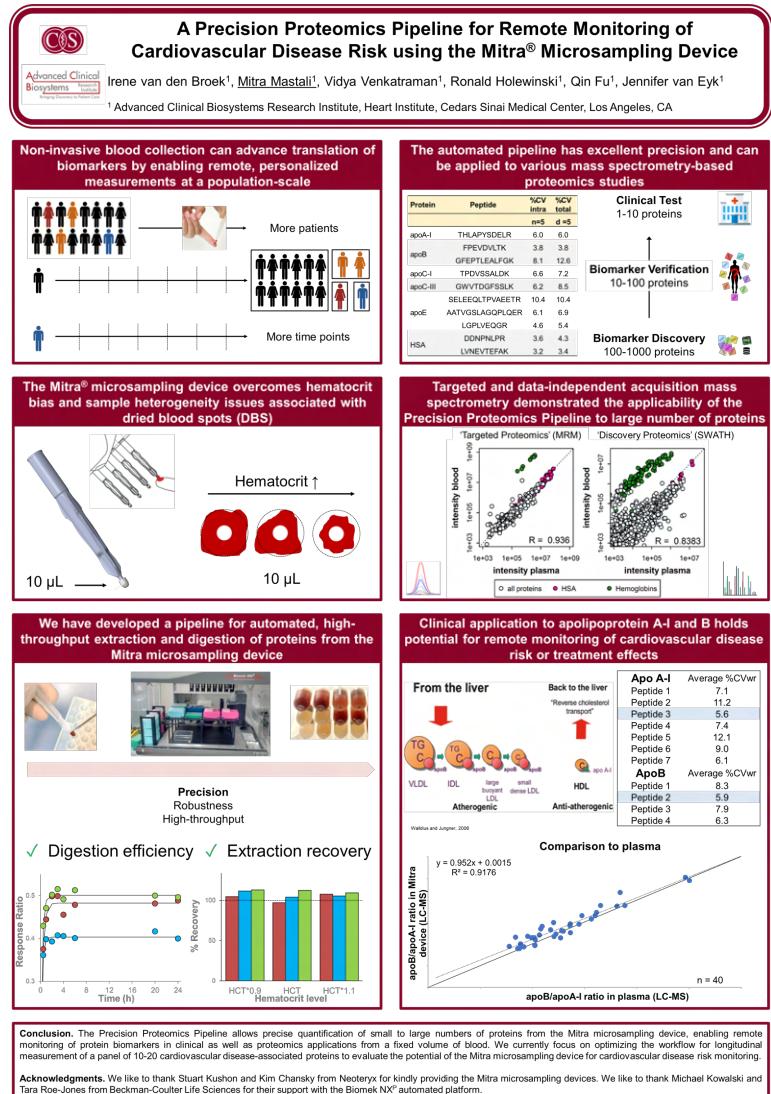
Visual communication in academic research is not *just* **data** charts...



You can also think of graphical abstracts...

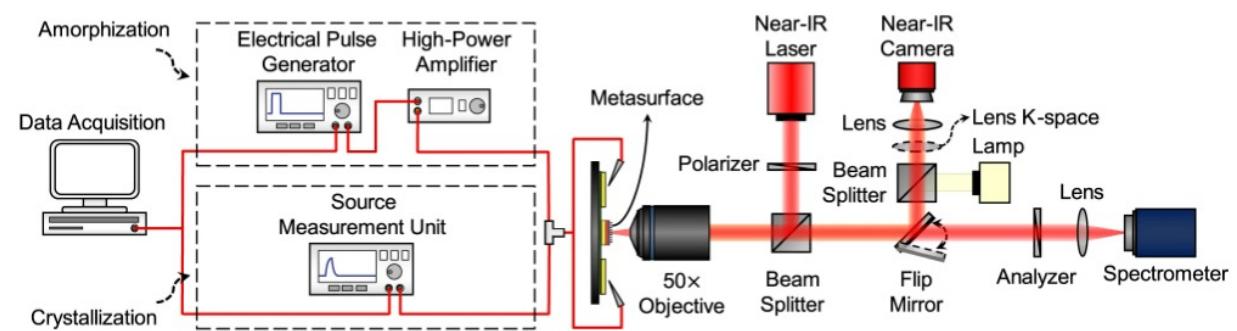
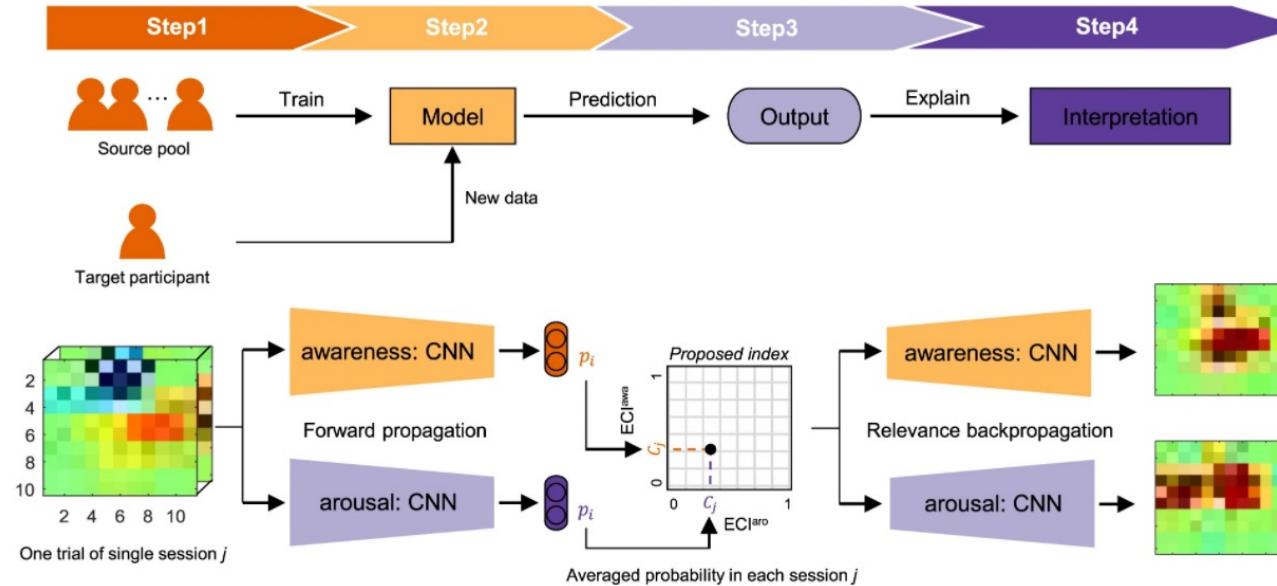


...or posters...



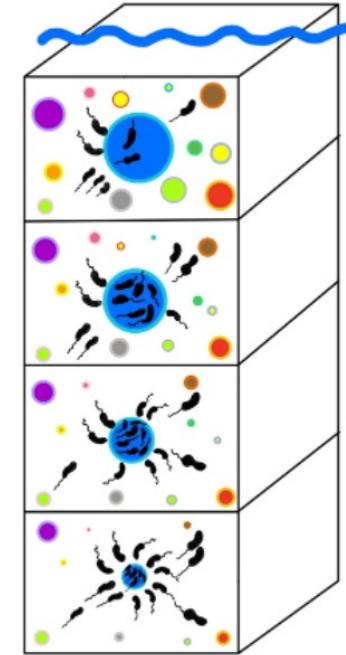
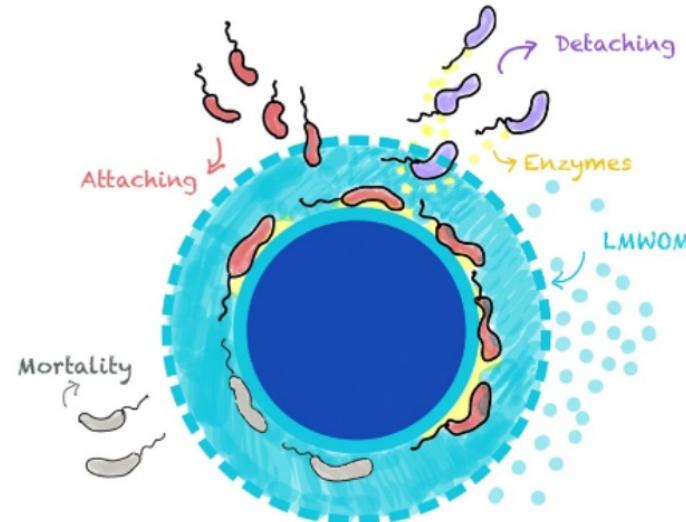
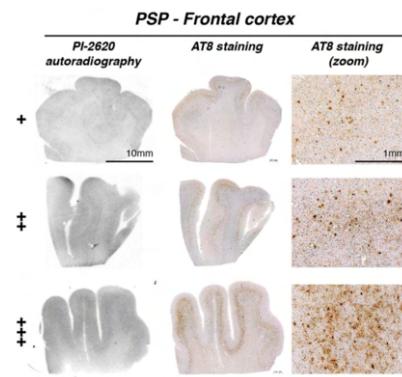
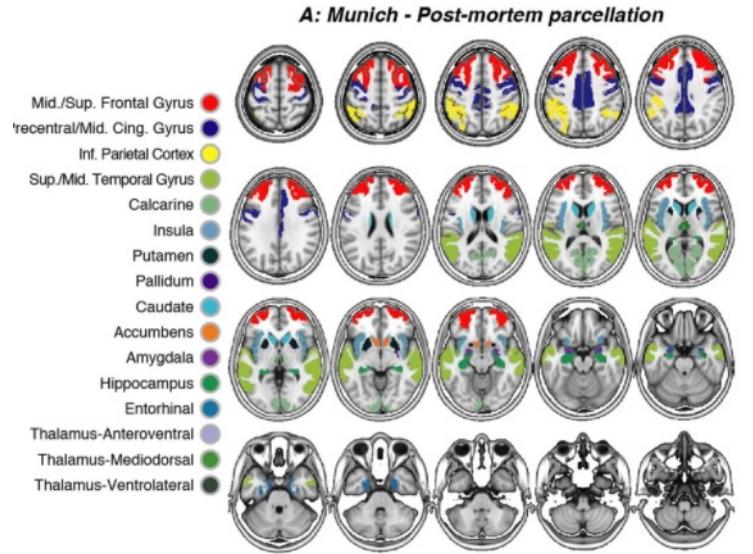
<https://phdcomics.com/>

Flow charts or schematics...



Abdollahramezani, S et al. Nat Commun **13**, 1696 (2022).

Illustrations, drawings, images...



Talking about reaching a broad public!

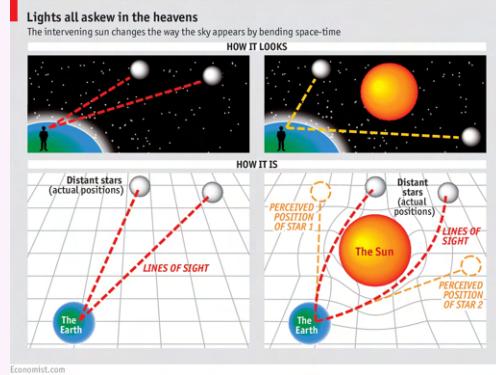
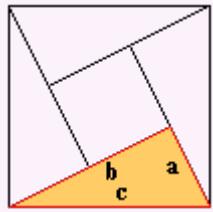


Prof. Ed Hawkins
#showyourstripes



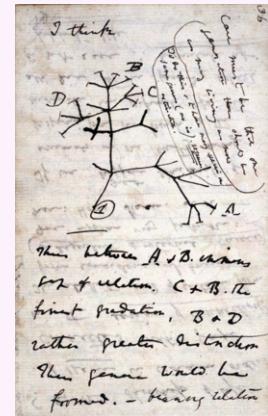
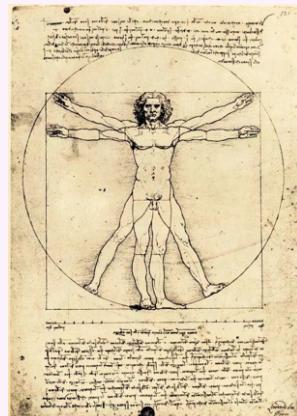
Four types of visualizations

Idea illustration



DECLARATIVE

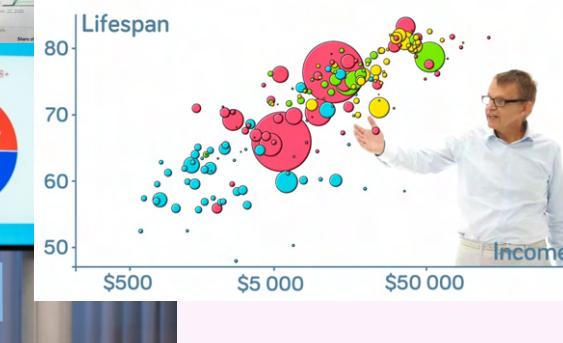
Idea generation



CONCEPTUAL



DATA-DRIVEN



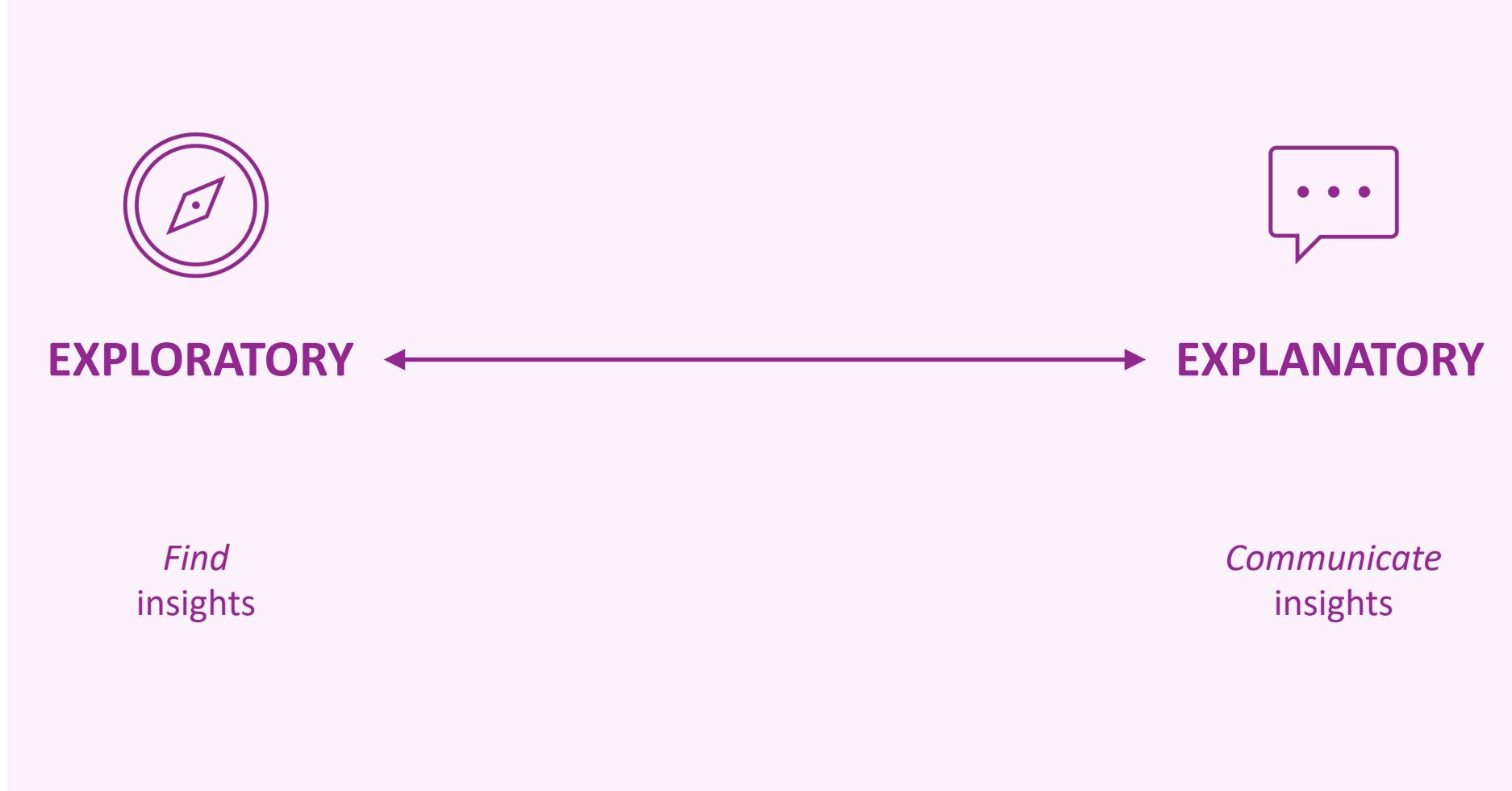
EXPLORATORY

'Everyday' DataViz

Visual discovery



Exploratory vs. explanatory data visualization



Exploratory vs. explanatory data visualization



EXPLORATORY

Understand

You



EXPLANATORY

Communicate

Other people

Less familiar (others)

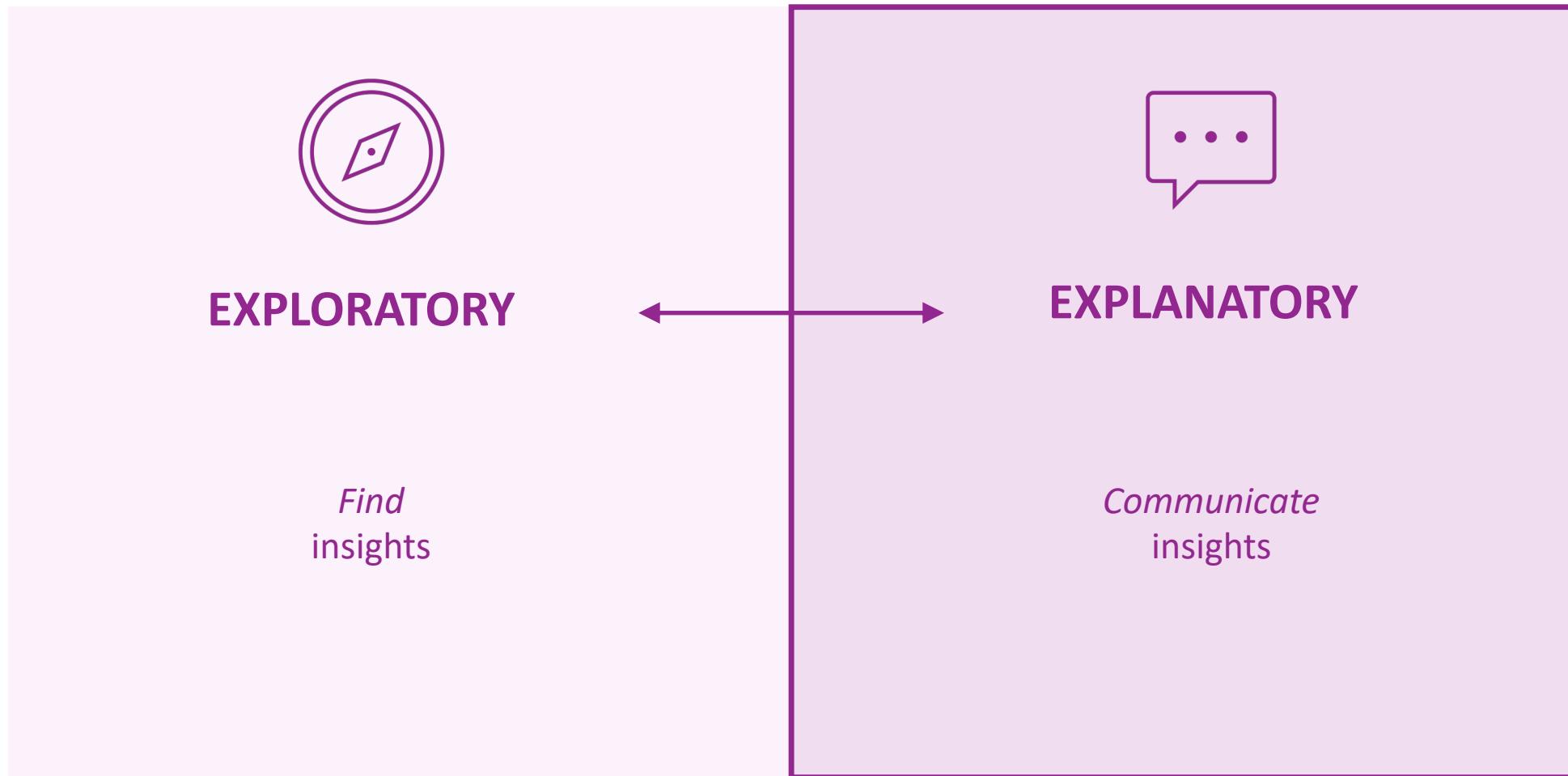
Simplicity, clarity, and cohesion

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



HOW visual communication ?

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

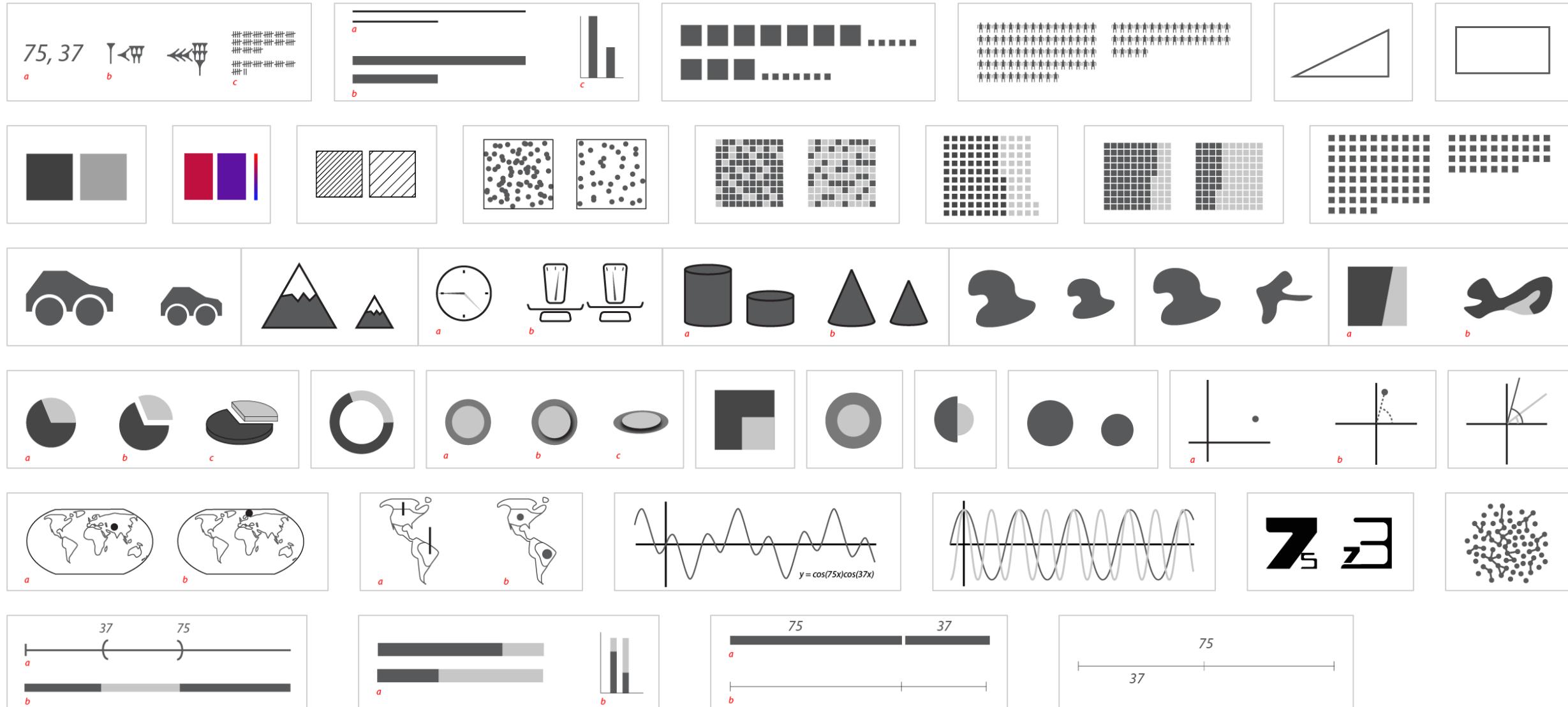
Exercise: In how many ways can you visualize two numbers?

75 and 37



5 min

45 ways to communicate two quantities...

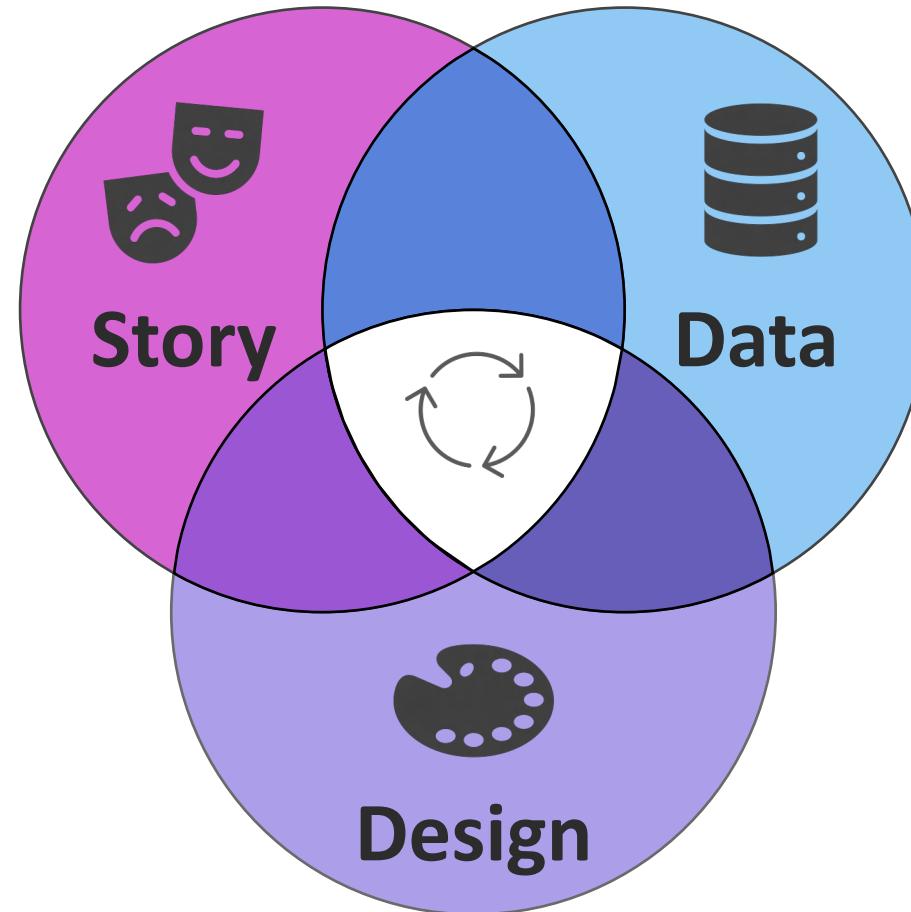


So many possibilities. How to choose?



The components of an effective data story

What's the context?
What's the message?

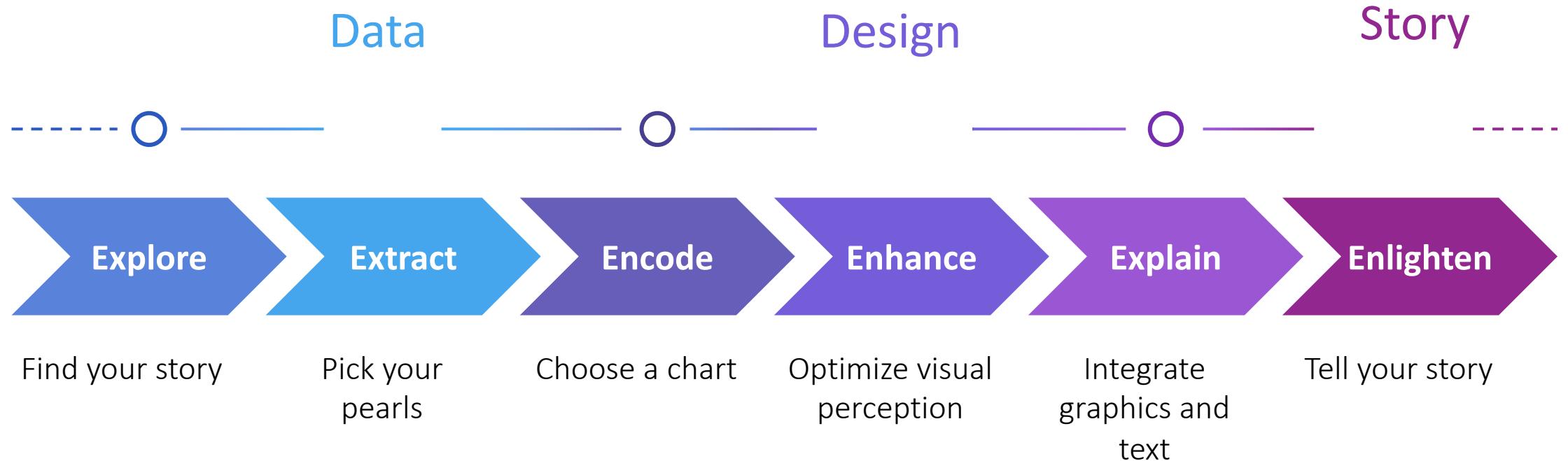


What do you want to show?
How do you want to show it?

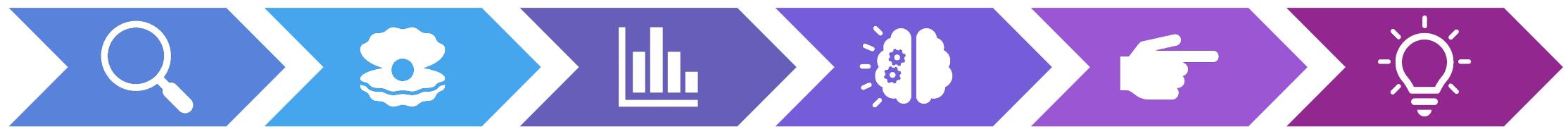
How to optimize visual perception?

Adapted from Effective Datastorytelling by Brent Dykes

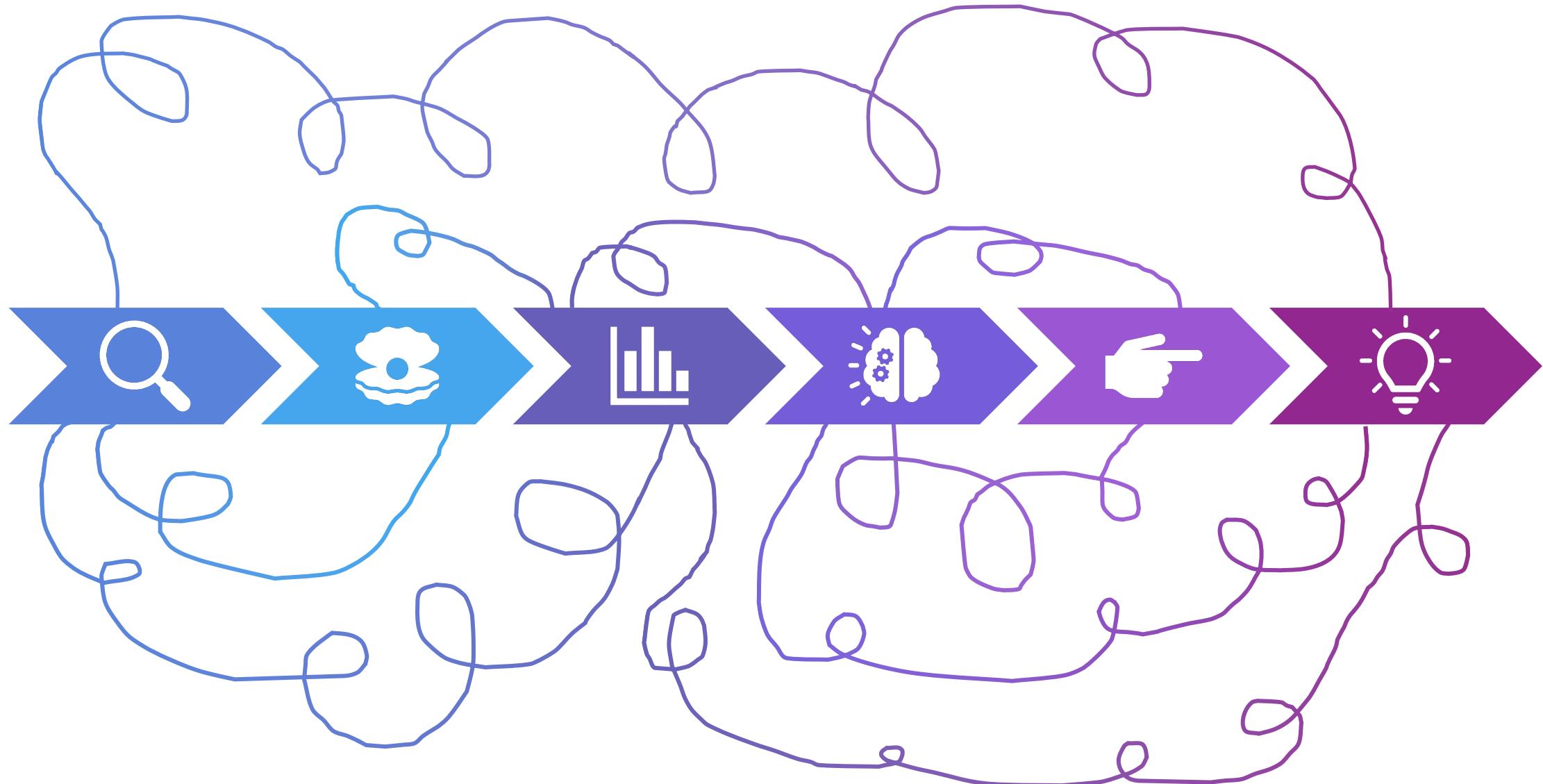
Six steps to integrate Data, Design and Story

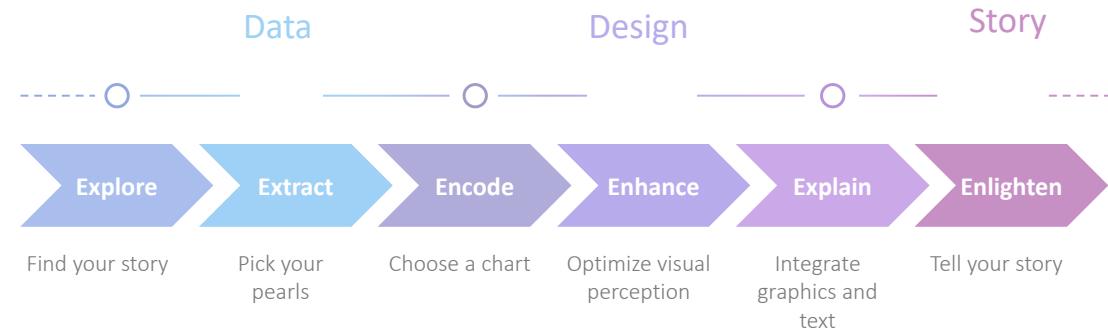


The ideal situation...



...reality!



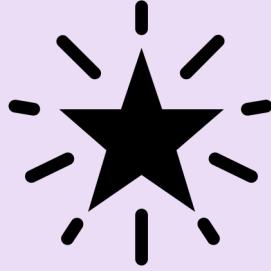


Explore: find your story

Begin with the end in mind

Know your context!

Why?



Outcome



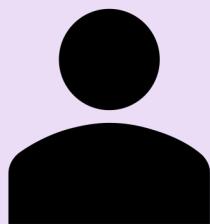
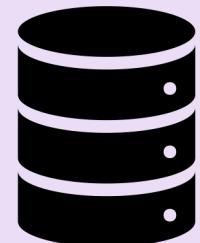
Audience

Who?



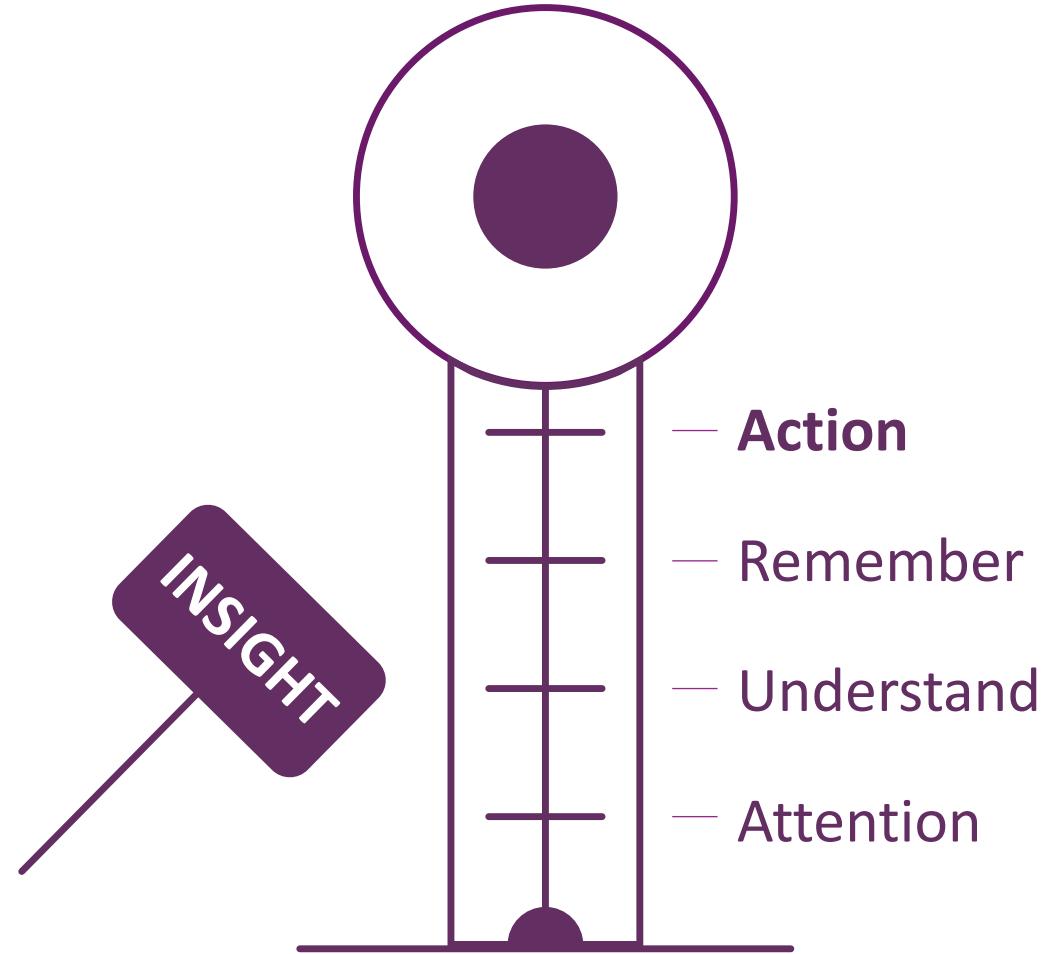
You

What?



Where?

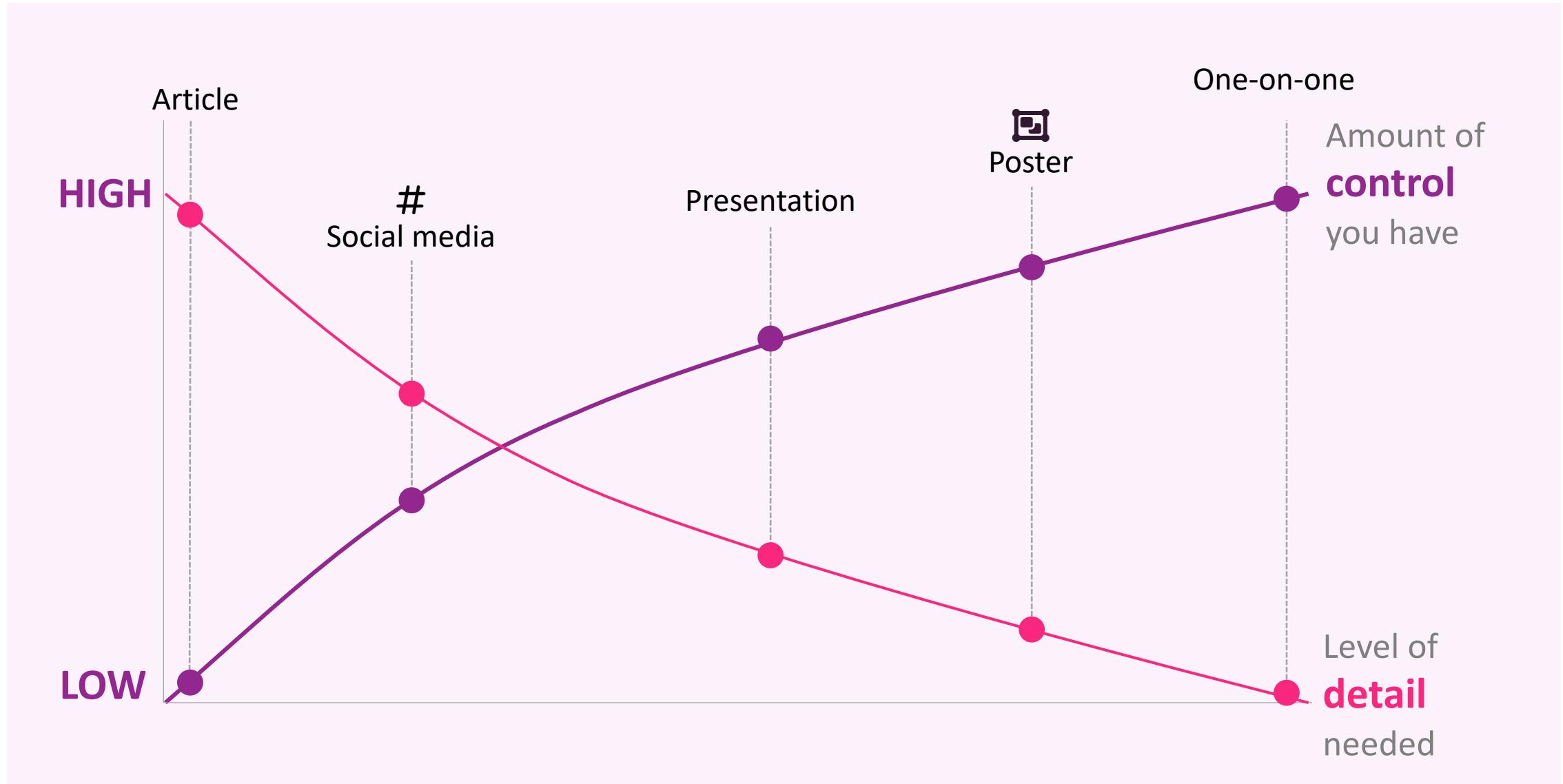
Why should your audience care?



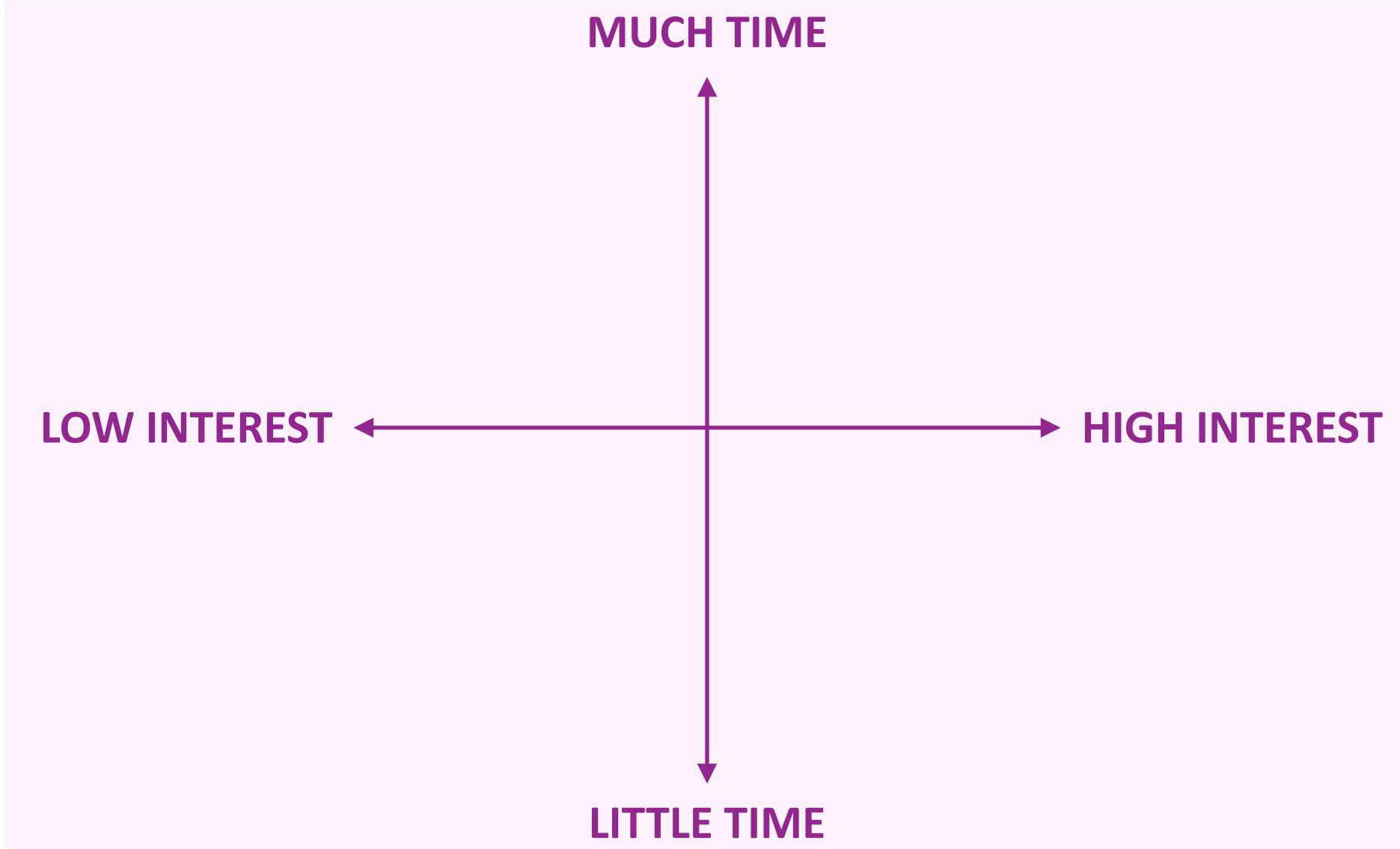
What is your **purpose**?
What **action** do you want your audience to take?

← **Why?**
← **Why?**
← **Why?**

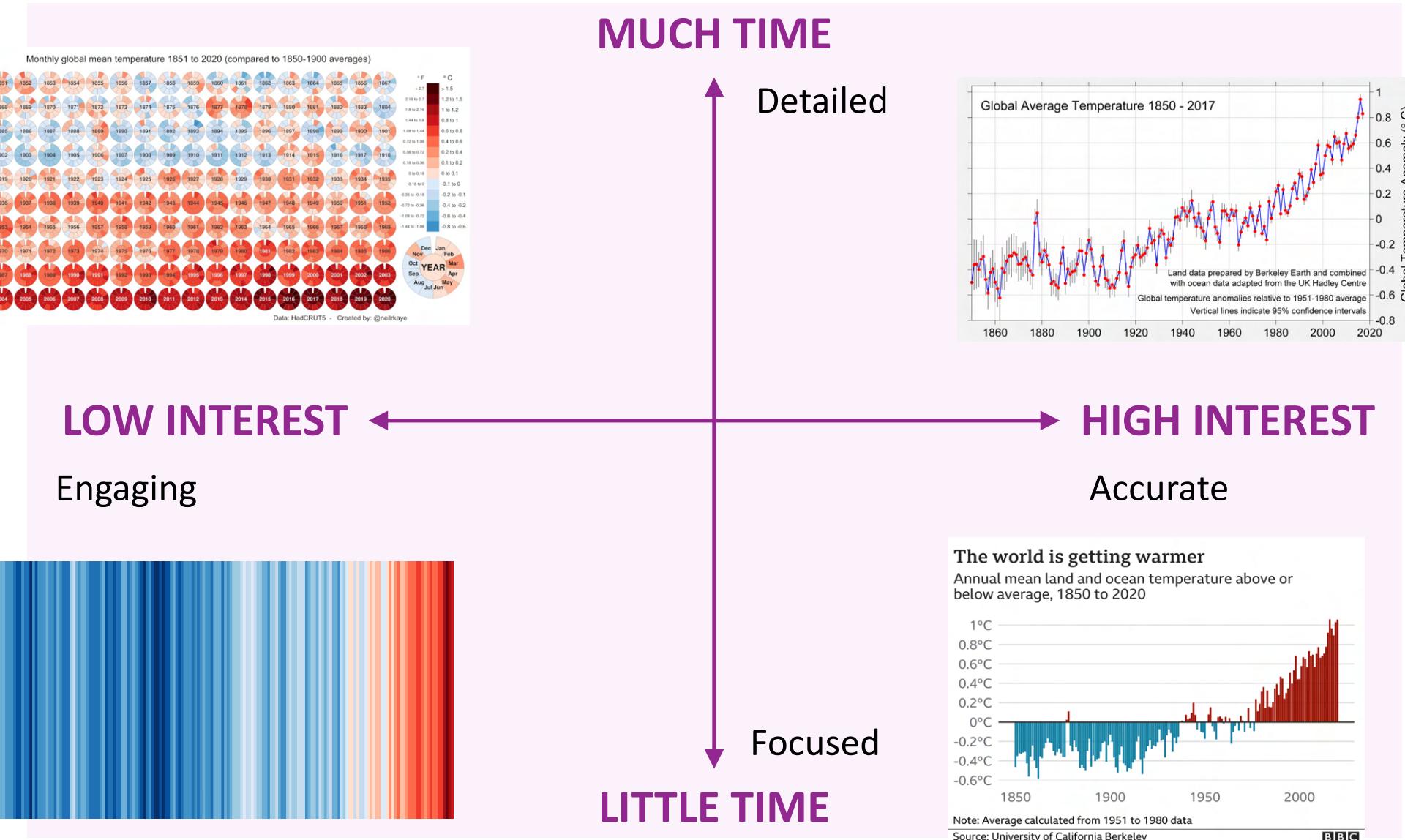
Where you present defines *how* you present

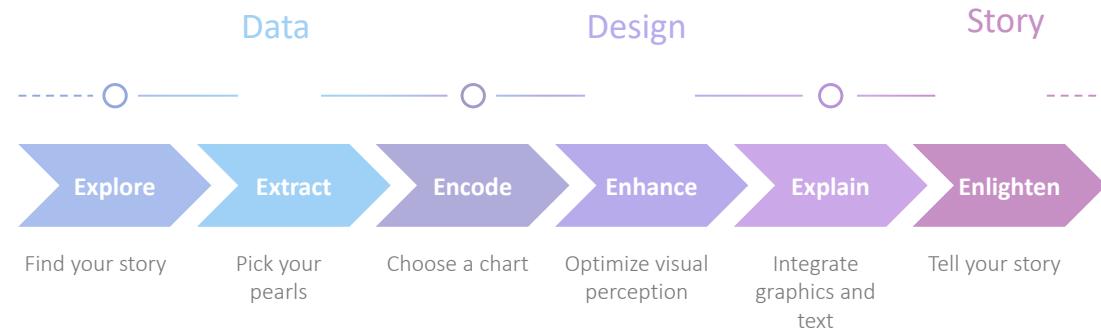


Who is your audience?



Who defines the level of detail and accuracy





Extract: pick your pearls

A wealth of information creates a poverty of attention

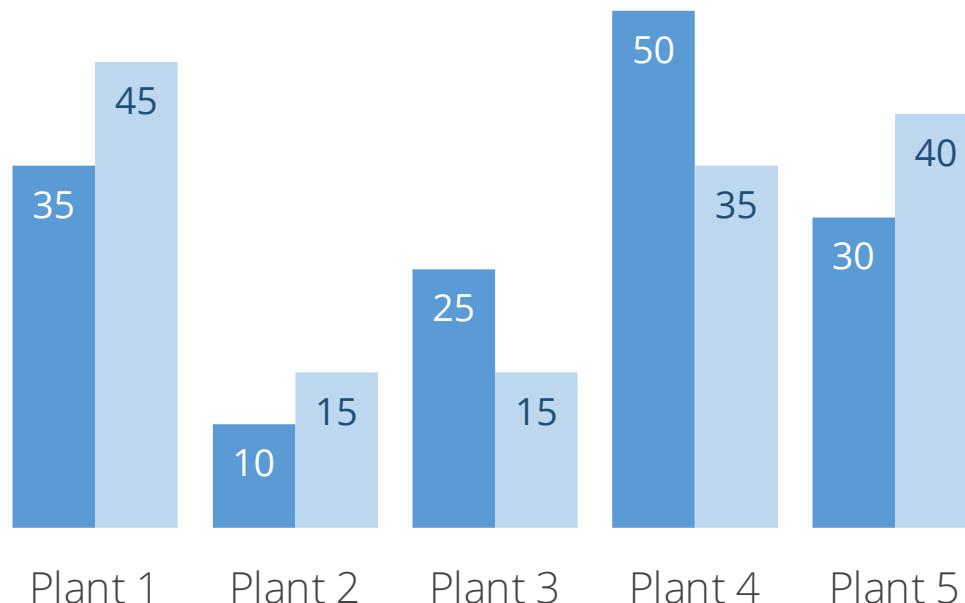
Resist the urge to show everything...!



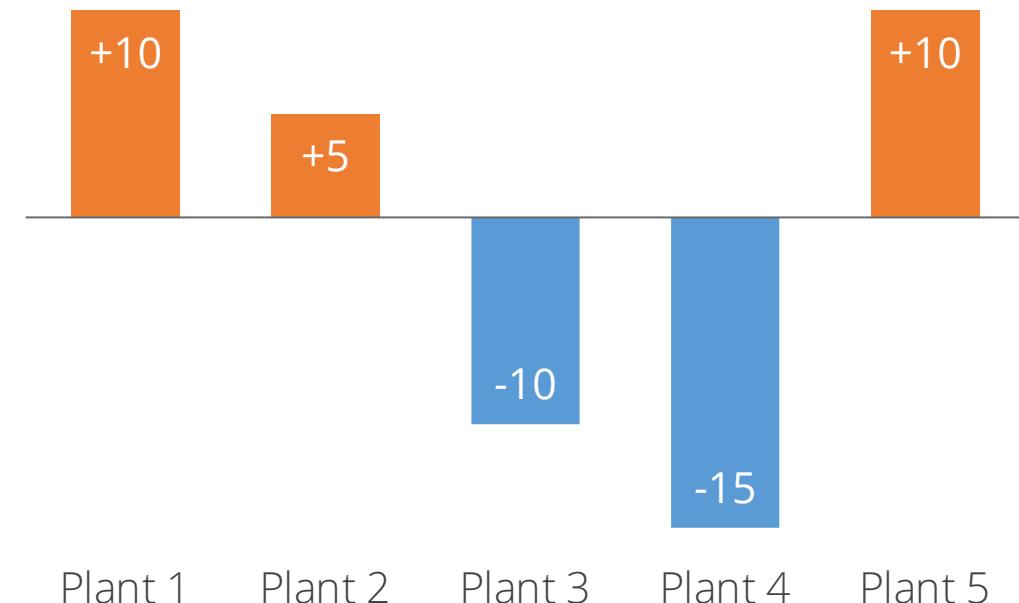
Avoid mental math: consider data variations

Safety Incidents by Plant

■ 2017 ■ 2018

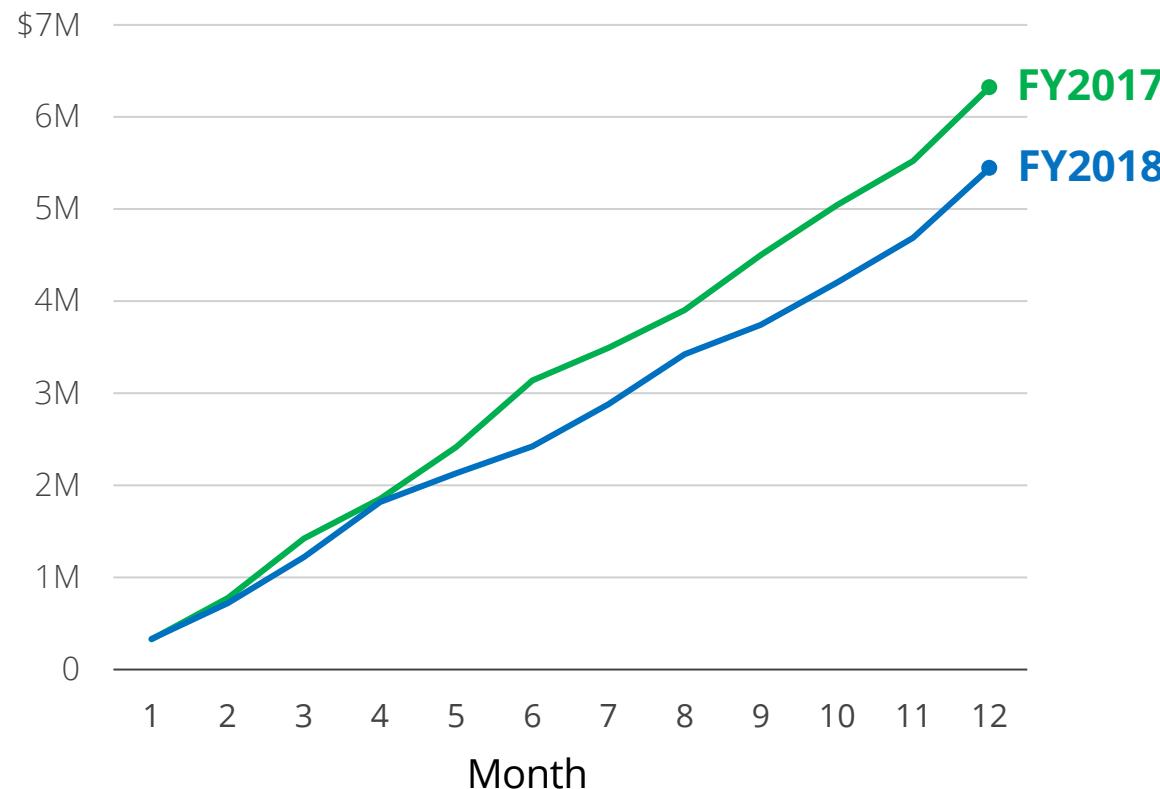


Year-to-Year Variance in Safety Incidents by Plant (2017-2018)

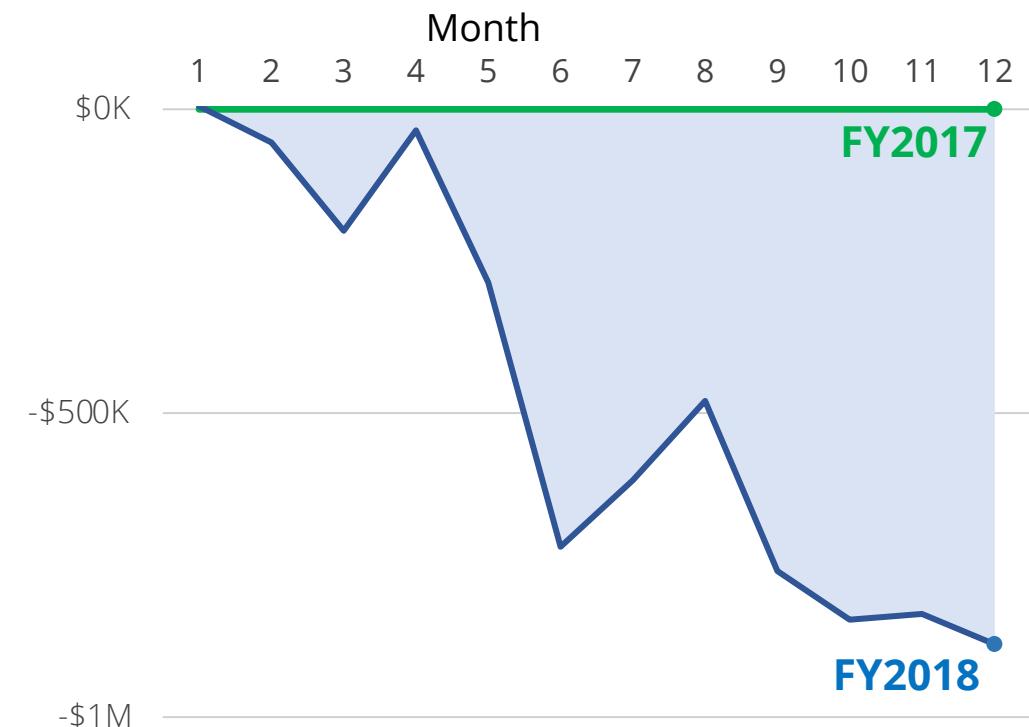


Avoid mental math: consider data variations

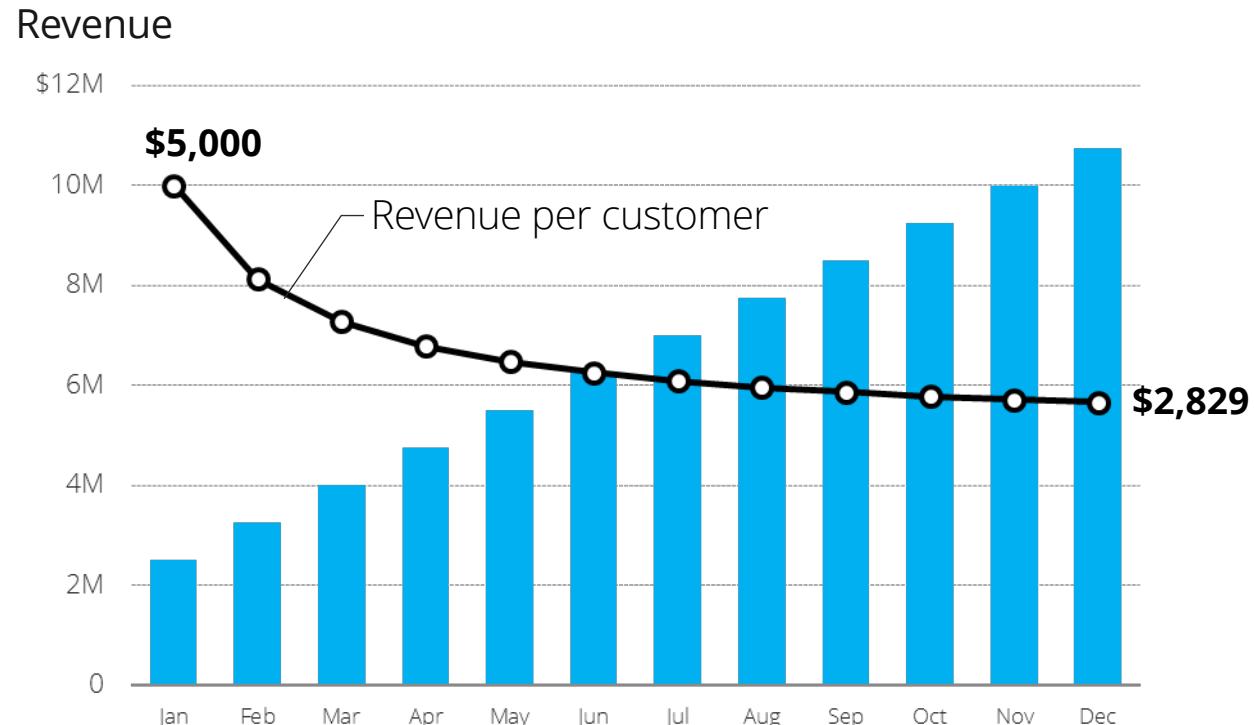
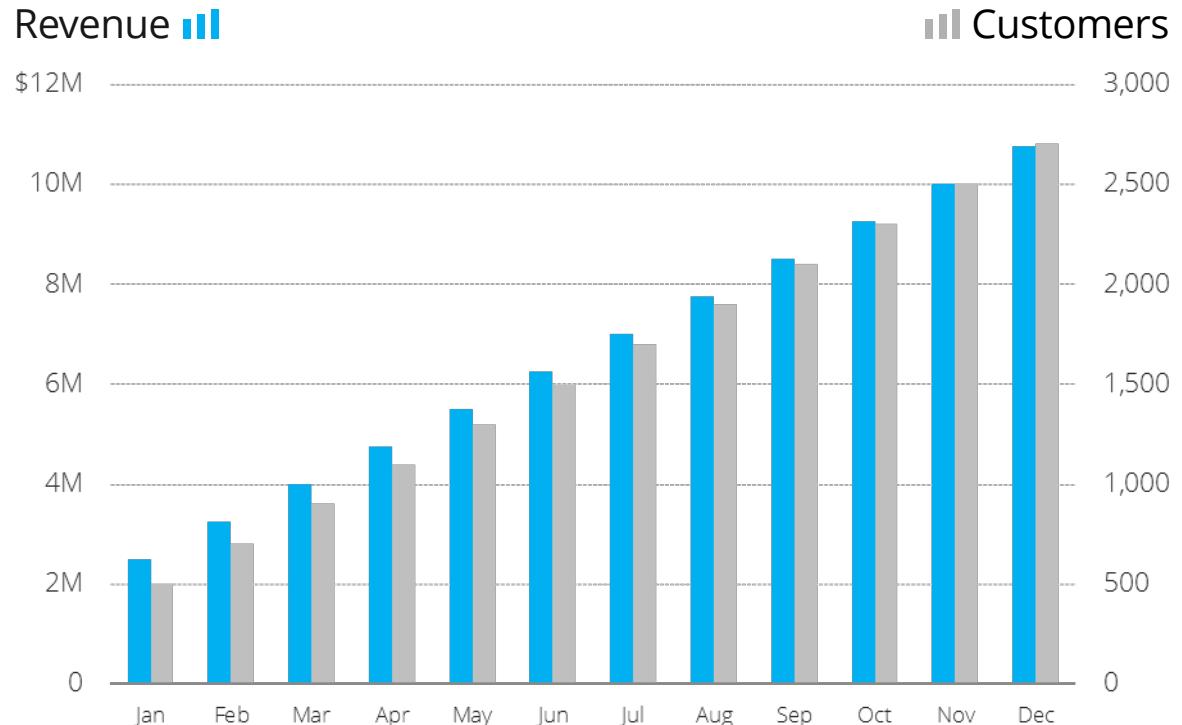
Running Total Revenues



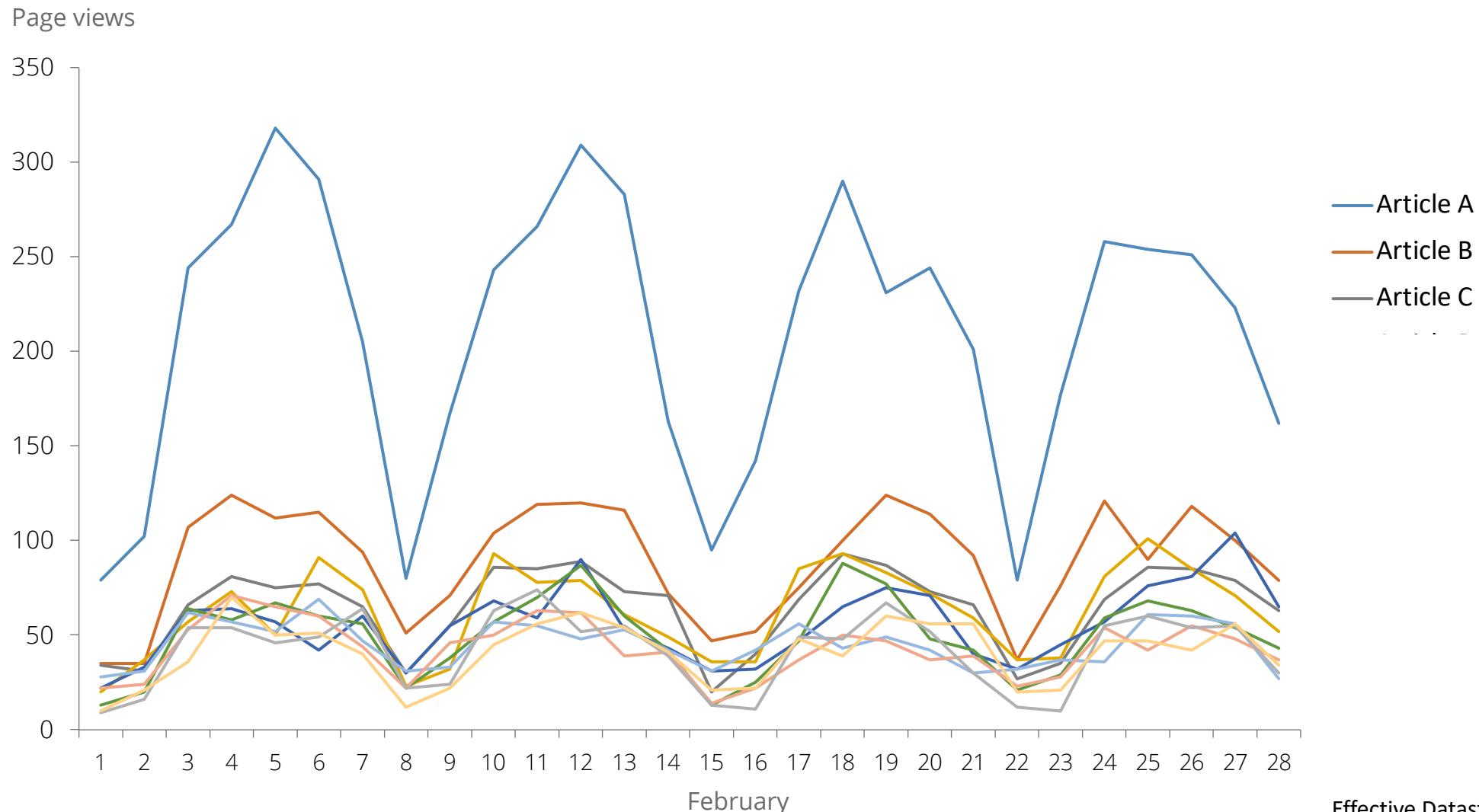
Revenue Variance between
FY2018 and FY2017



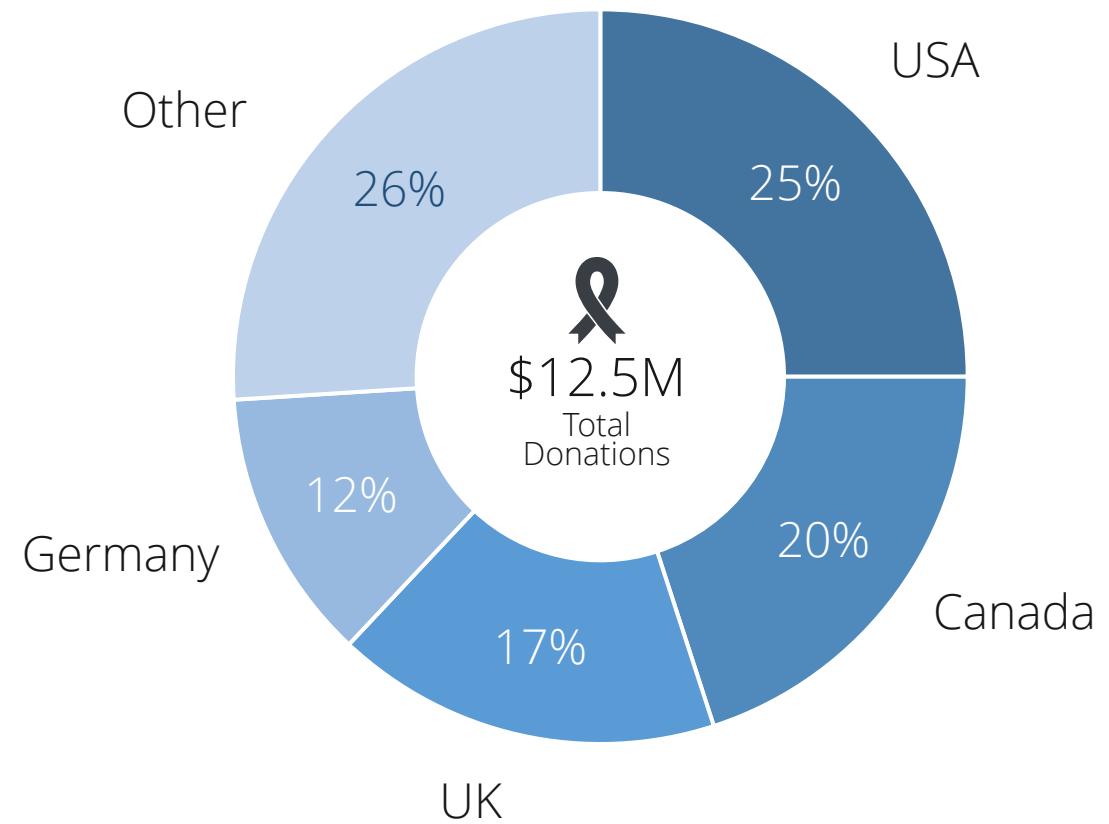
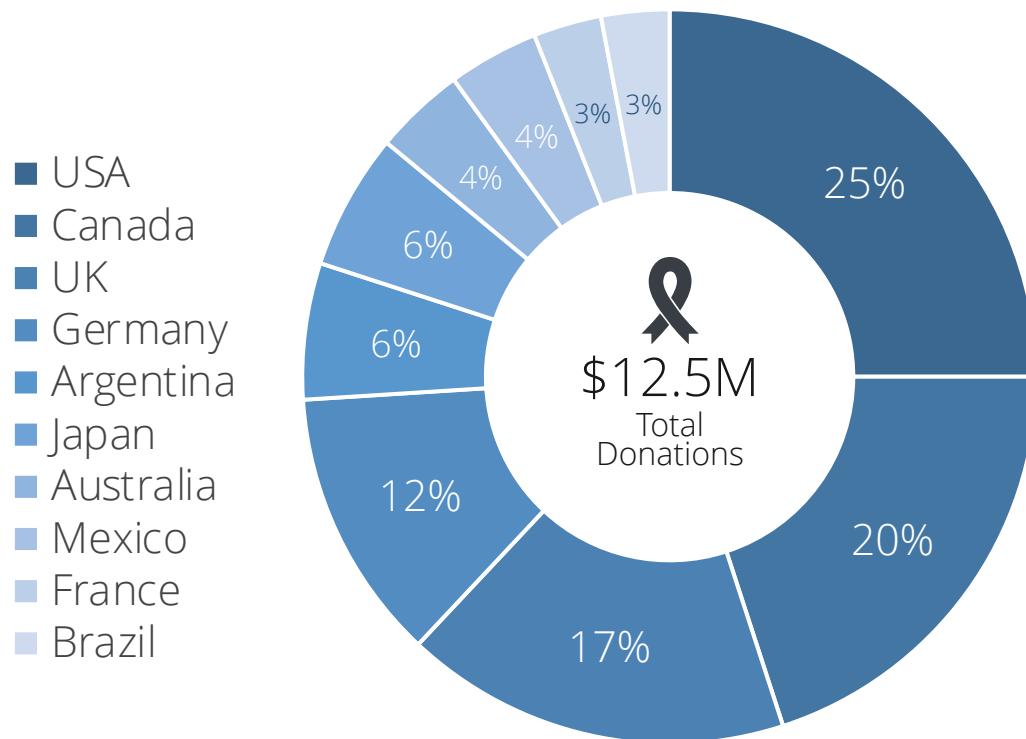
Clarify the problem using a calculated metric



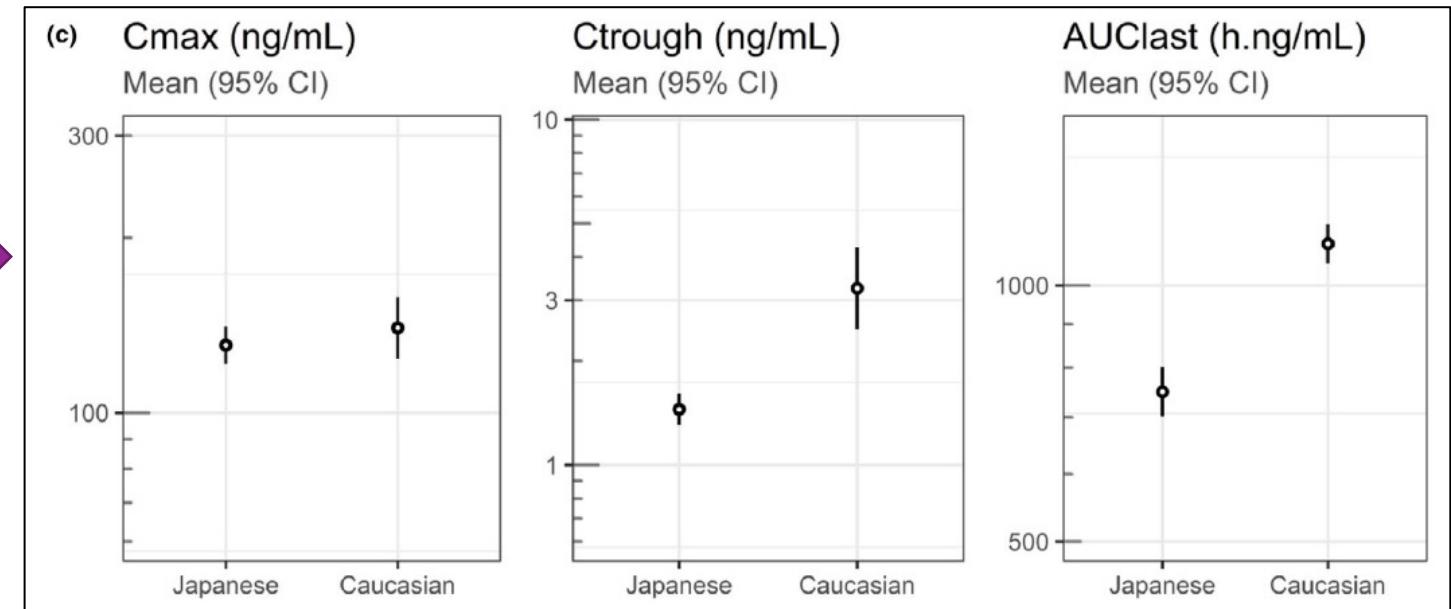
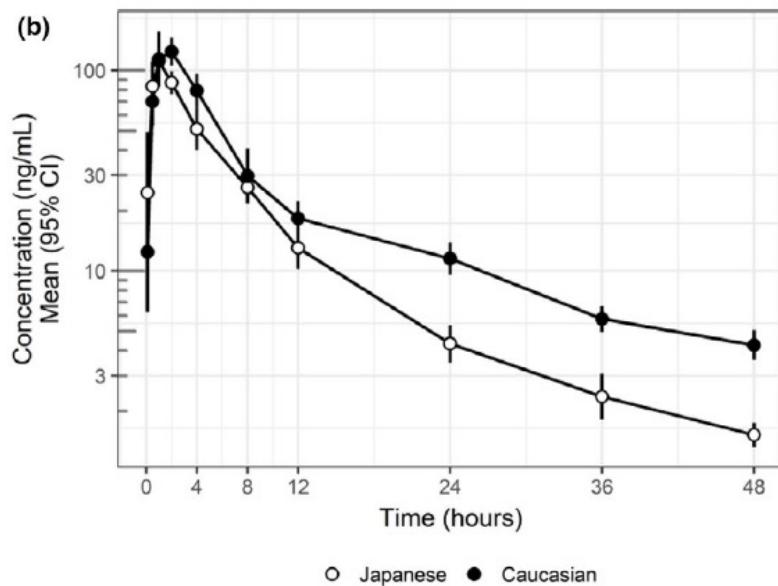
Reduce information overload: Deprioritize unnecessary data



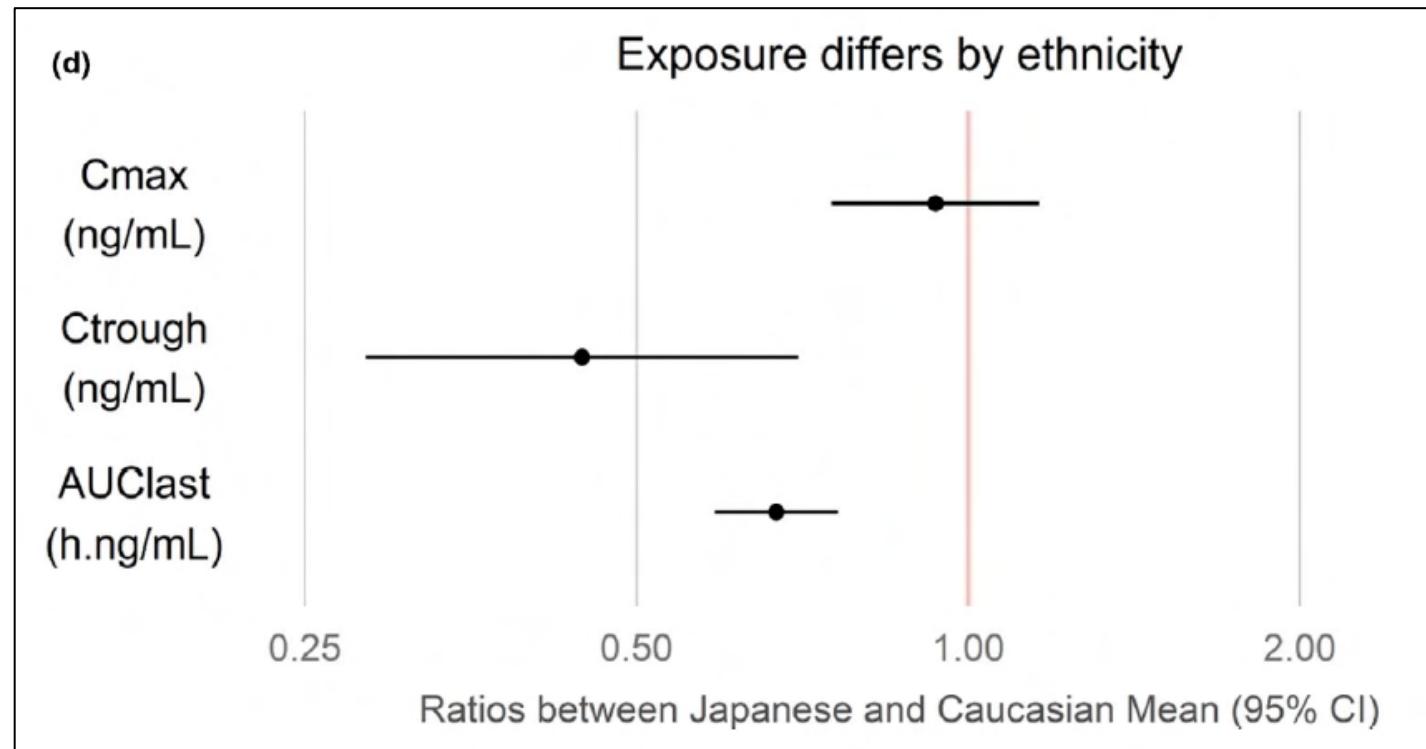
Reduce information overload: Group less important data

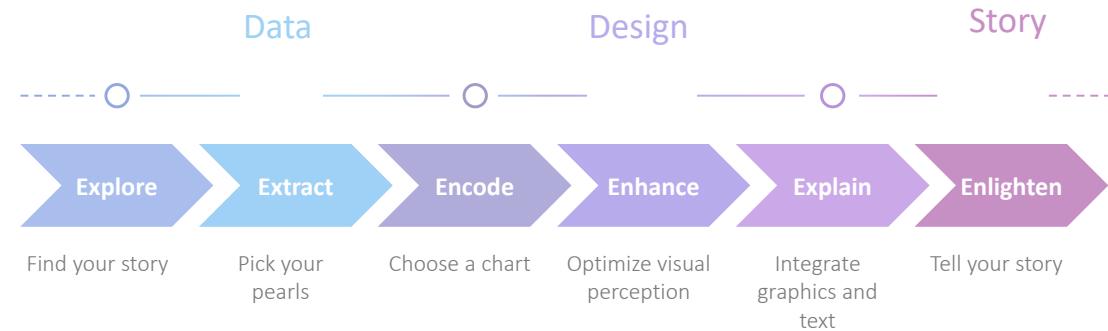


Reduce information overload: which value is most relevant?



Reduce information overload: which value is most relevant?

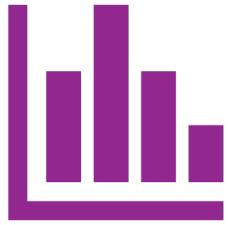




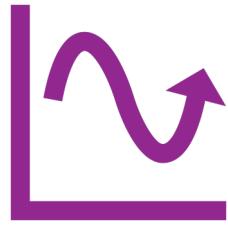
Encode: choose a chart

A picture is worth a thousand words

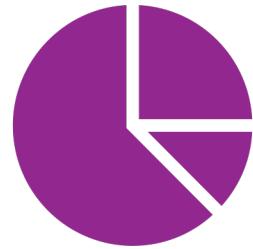
What do you want to show?



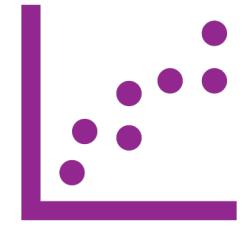
Comparison



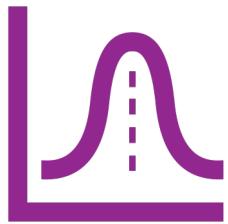
Trend



Part of a whole



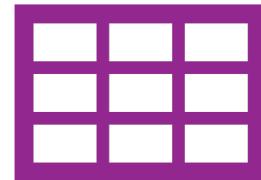
Relationship



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

relevant ranking

target or benchmark

Support easy comparison

small multiple

gridlines

Visual hierarchy

less is more

emphasize

Show and tell

descriptive title

annotation

don'ts

Misleading design

cutting of Y axes

multiple Y axis

Beauty over accuracy

3D effect

fluid lines

Too much

more than 4 series

too much or meaningless color

Hard to read

too much details

align text other than horizontally

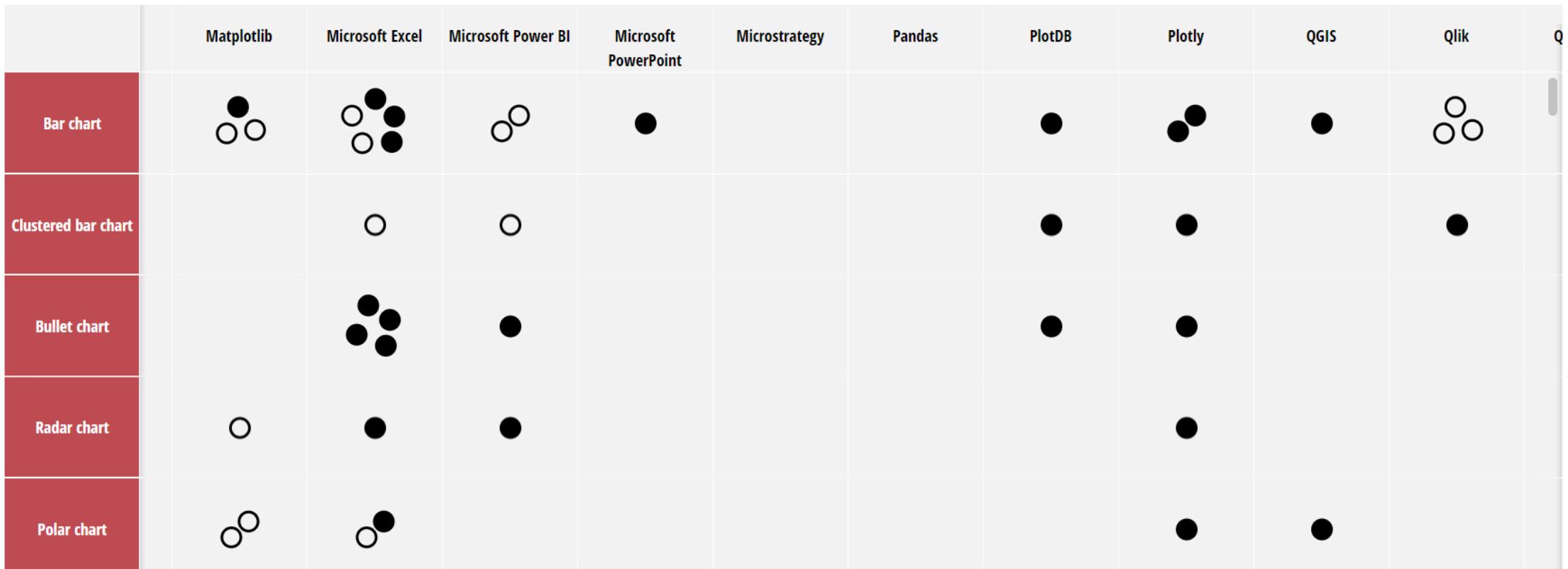
The Chartmaker directory

⊕ THE CHARTMAKER DIRECTORY

ABOUT

Filter by chart name or AKA

Reference Type: ● Example ● Solution | Chart Families: ● Categorical ● Hierarchical ● Relational ● Temporal ● Spatial





from Data to Viz

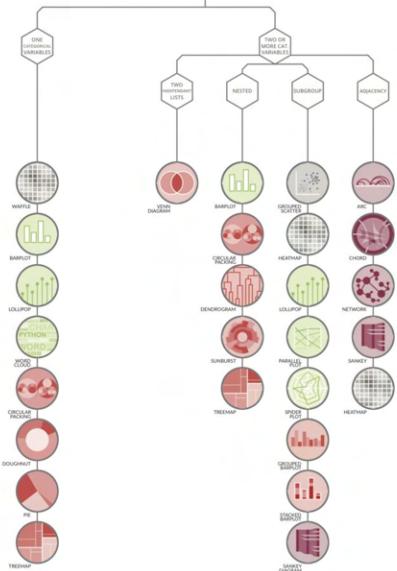
'From Data to Viz' is a classification of chart types based on input data format. It will help you find the perfect chart in three simple steps :

- 1** Identify what type of data you have.
- 2** Go to the corresponding decision tree and follow it down to a set of possible charts.
- 3** Choose the chart from the set that will suit your data and your needs best.

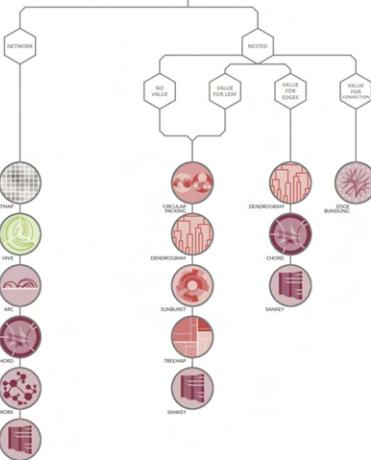
Dataviz is a world with endless possibilities and this project does not claim to be exhaustive. However it should provide you with a good starting point. For an interactive version and much more, visit:

data-to-viz.com

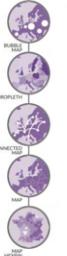
CATEGORIC



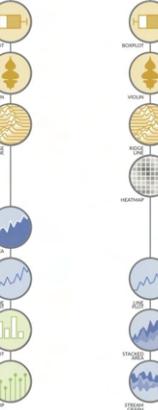
RELATIONAL



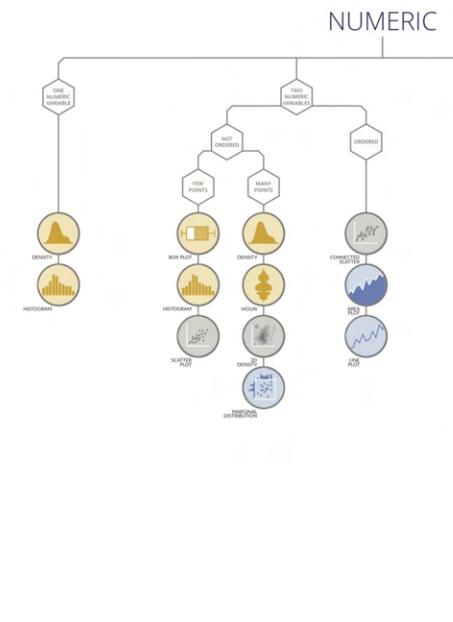
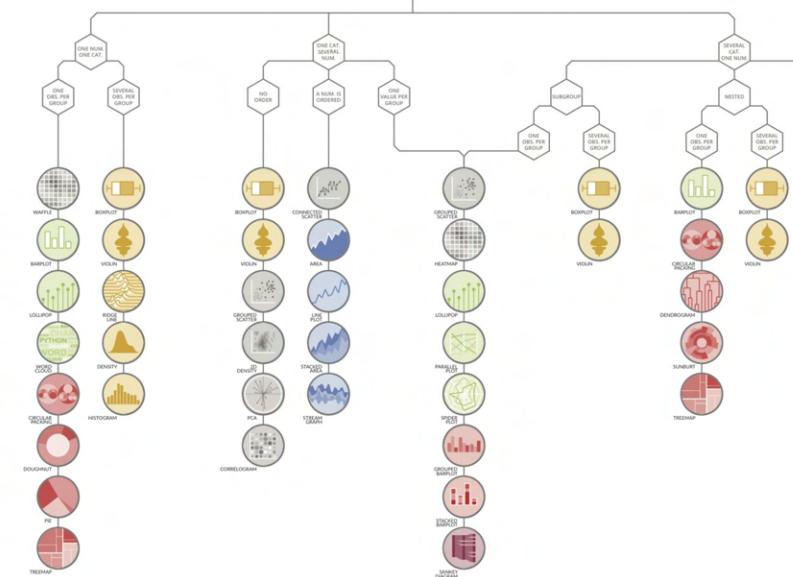
MAP



TIME SERIES



CATEGORIC AND NUMERIC





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. Does the distance between cities affect the number of infections?

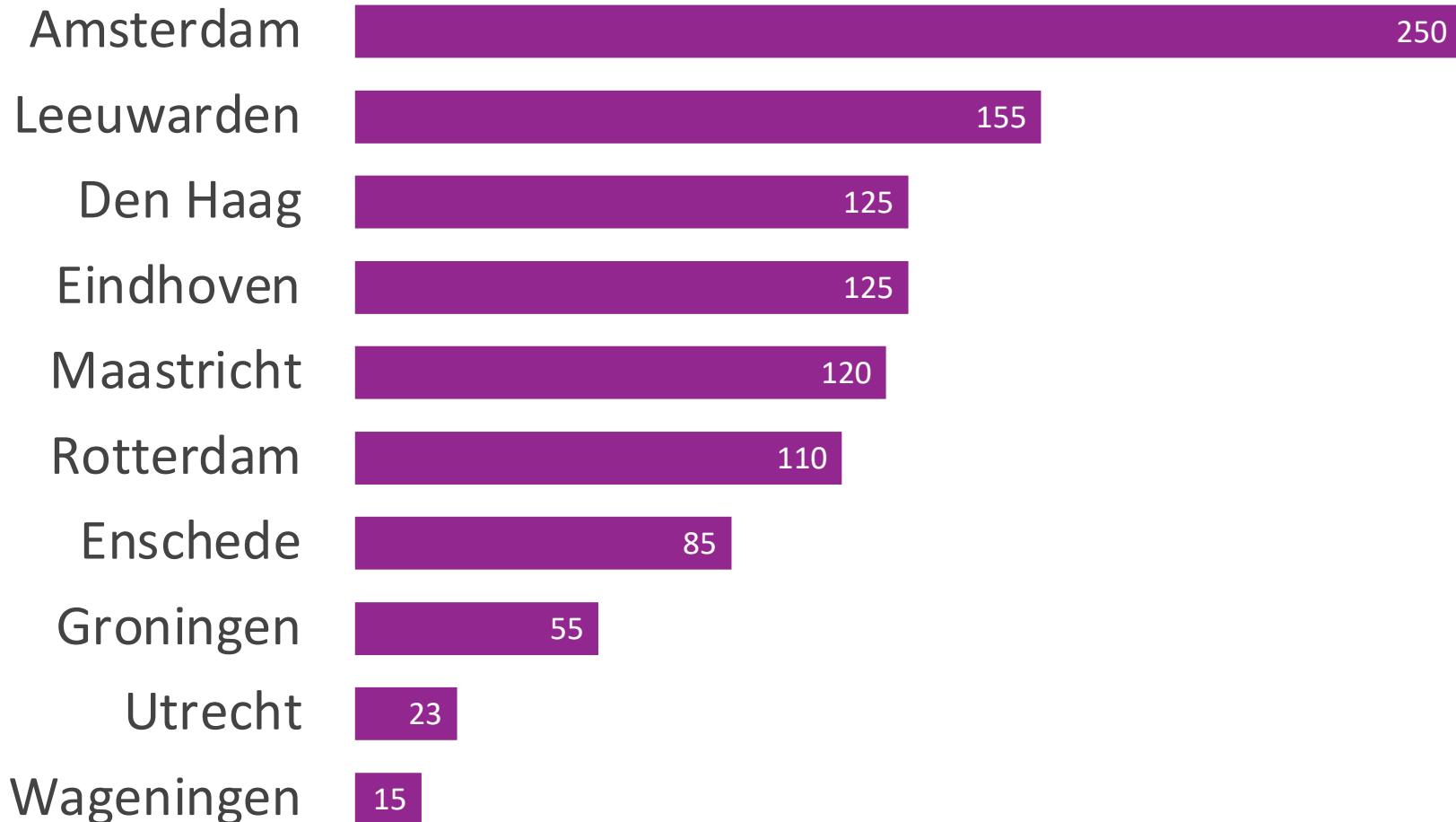
- **1) Create a visual tailored to the reader's question**
- **2) How can you change the visual for another audience?**
 - More accurate / engaging
 - More focused / detailed

Use your imagination!

How many infections per city?

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

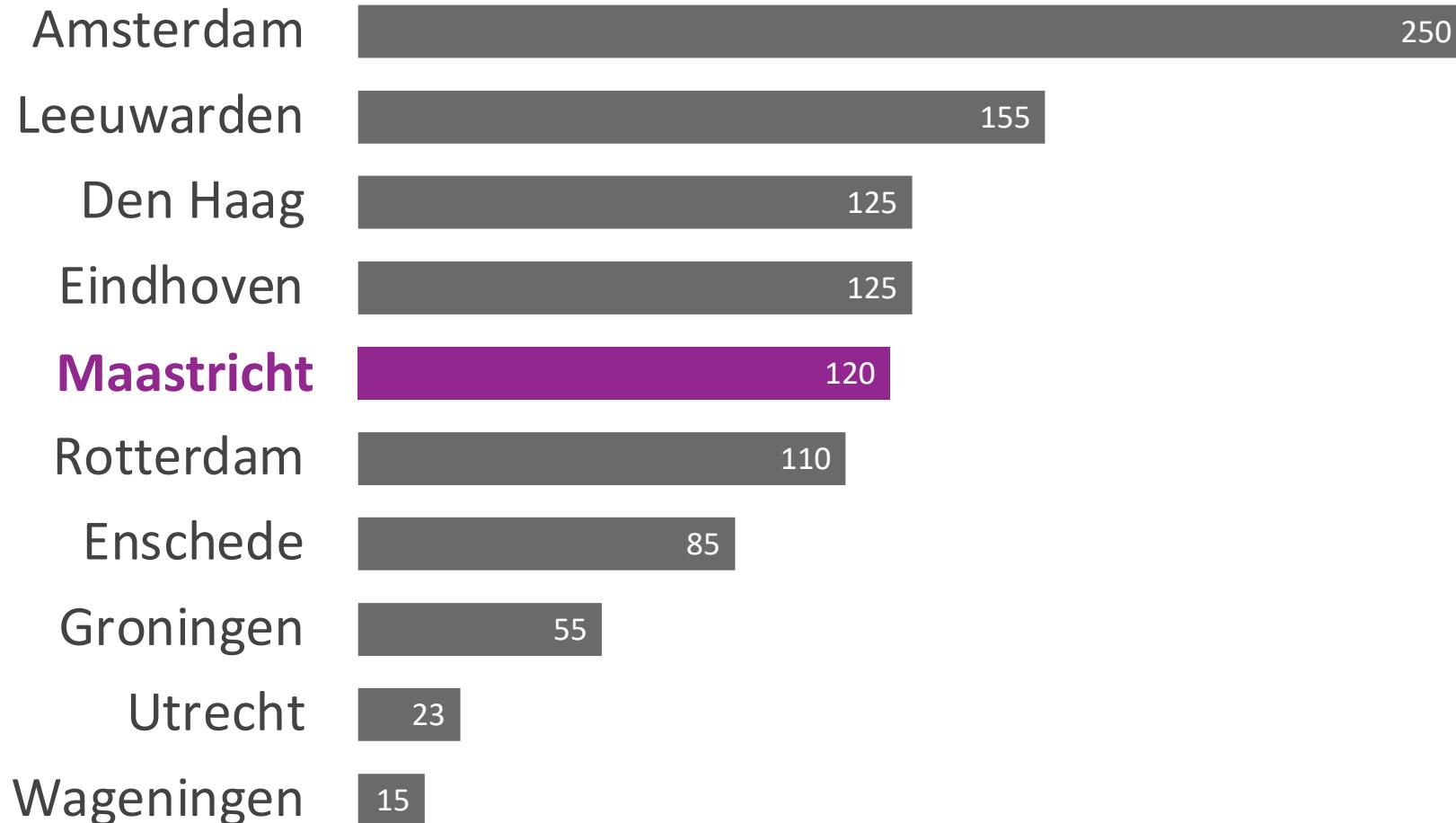
How many infections per city?



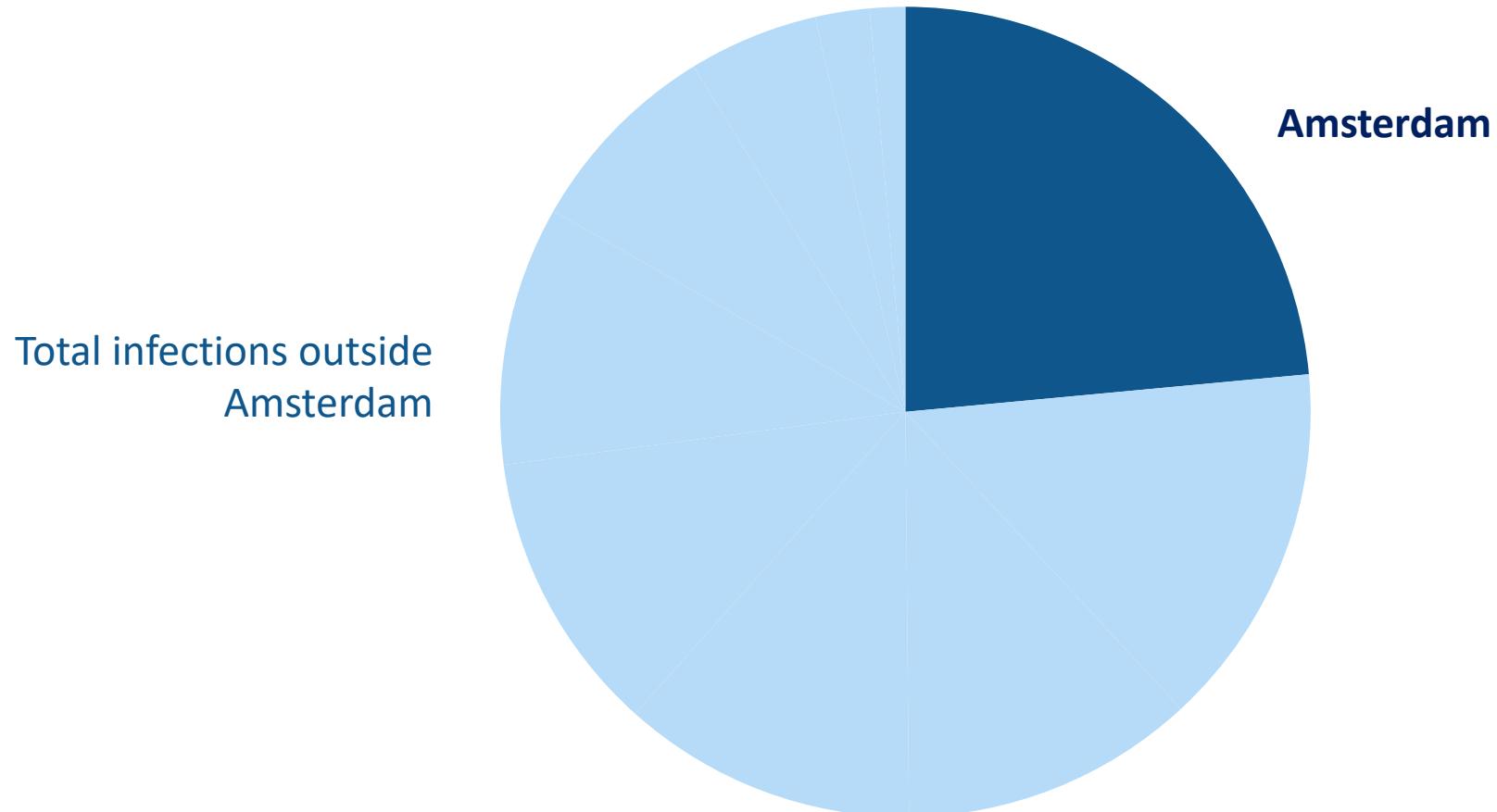
How does Maastricht rank?

	City	No. infections
1.	Amsterdam	250
2.	Leeuwarden	155
3.	Den Haag	125
4.	Eindhoven	125
5.	Maastricht	120
6.	Rotterdam	110
7.	Enschede	85
8.	Groningen	55
9.	Utrecht	23
10.	Wageningen	15

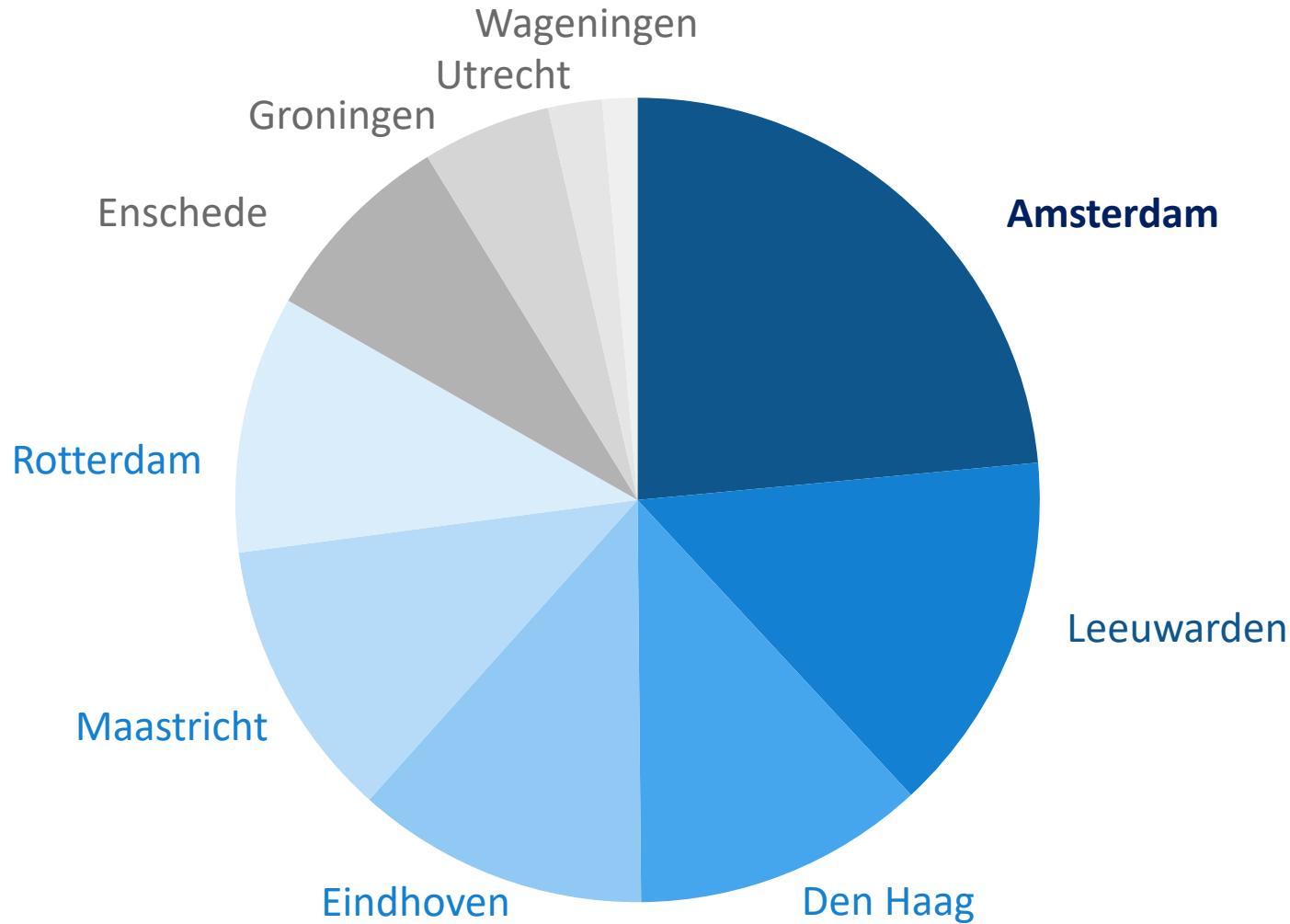
How does Maastricht rank?



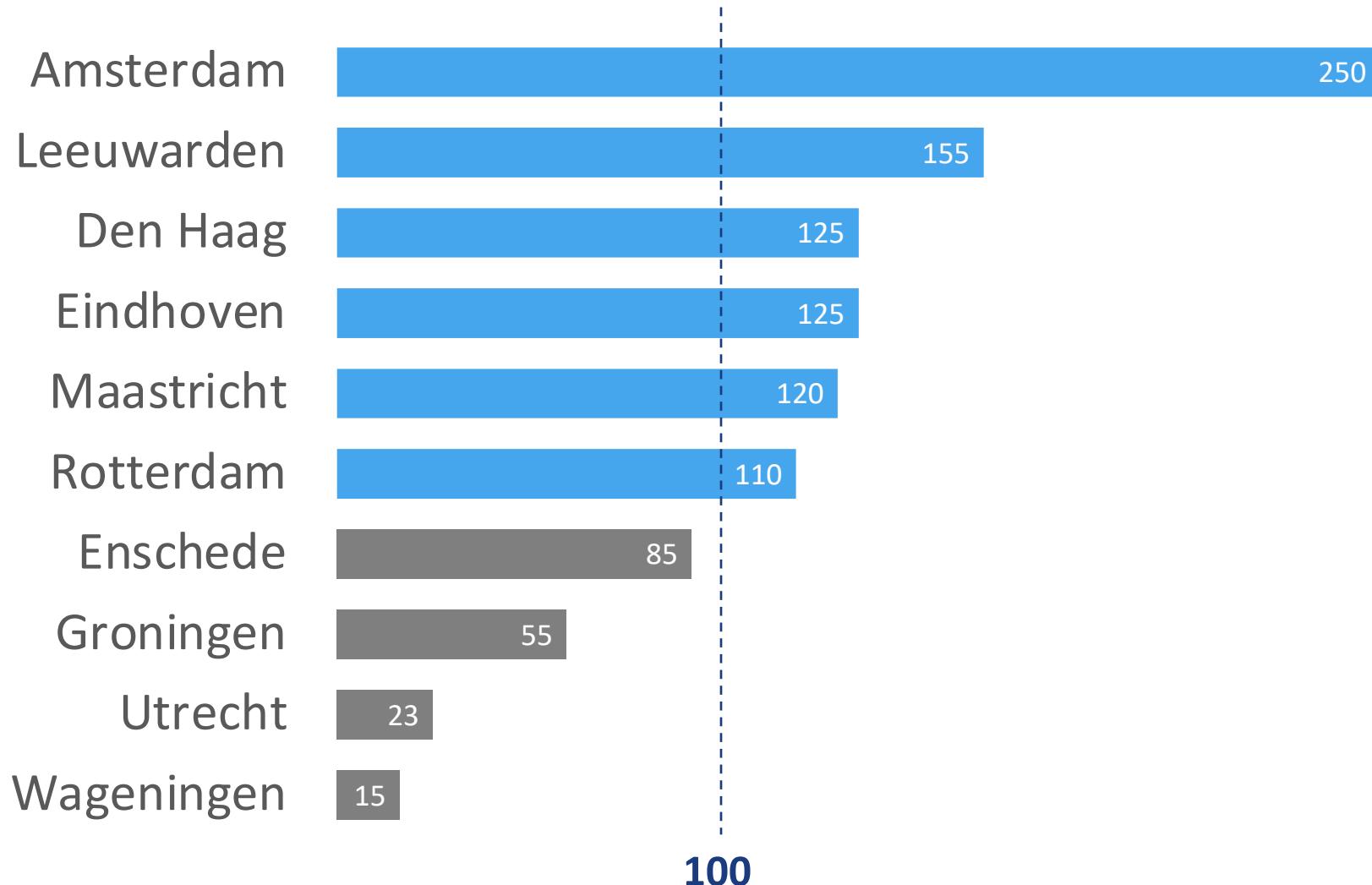
How does the number of infections in Amsterdam related to other cities?



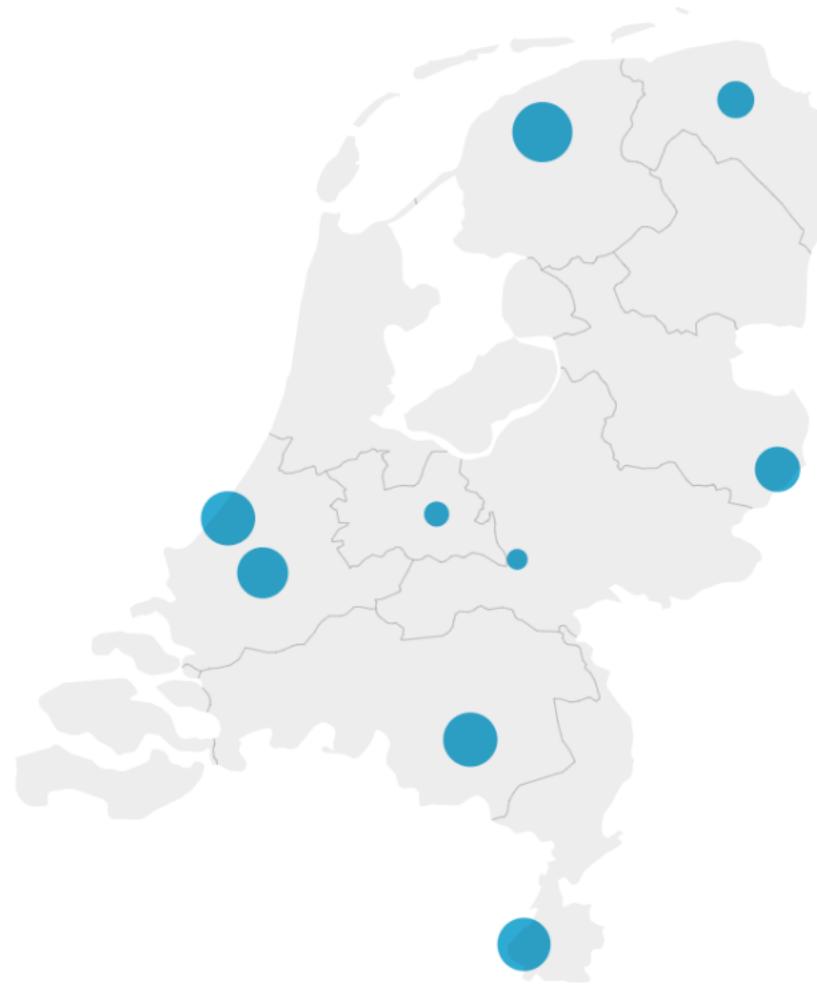
How does the number of infections in Amsterdam related to other cities?

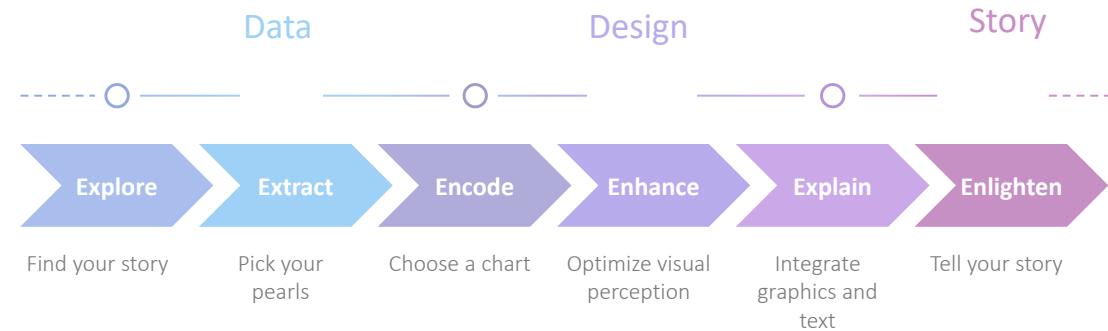


Which cities have more than 100 infections?



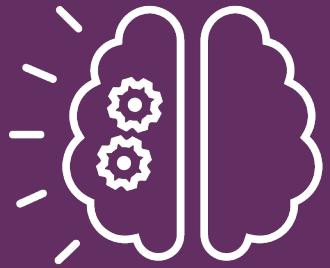
Does the distance between cities affect the number of infections?





Enhance: optimize visual perception

The eyes only see what the mind is prepared to comprehend



Visual perception

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!

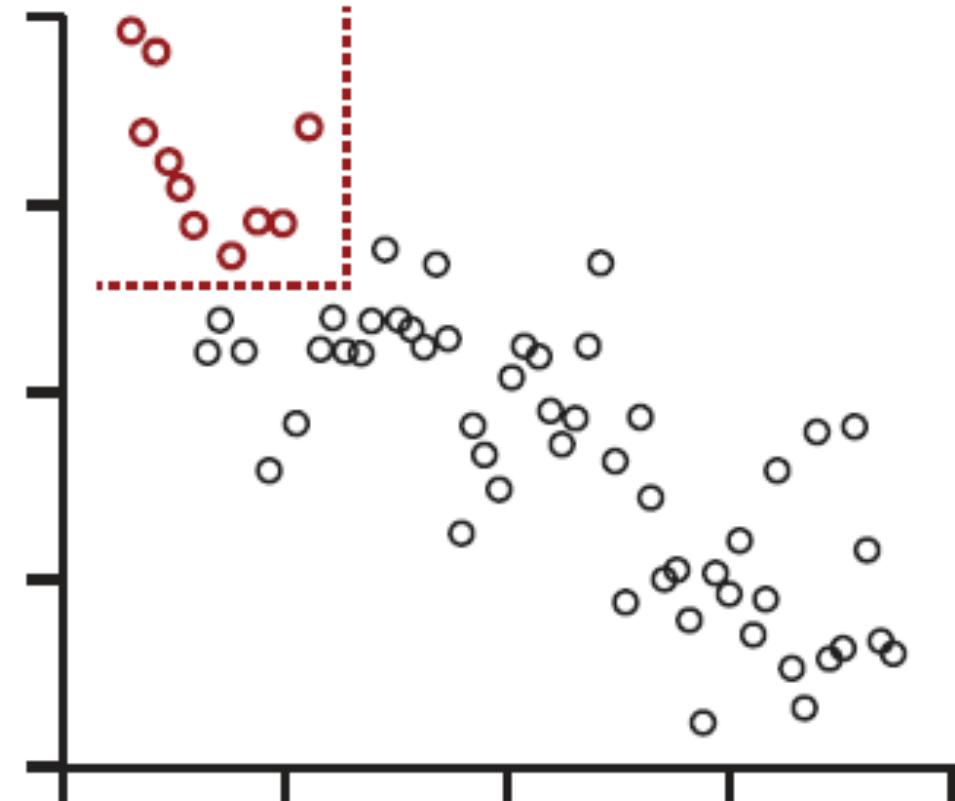
321654643216465132168461321
646432136748651349687463123
668432616266984892536496874
651687964649843616984616957
465167412323161689312439873
421432875943869234987596387

You can use pre-attentive attributes in tables

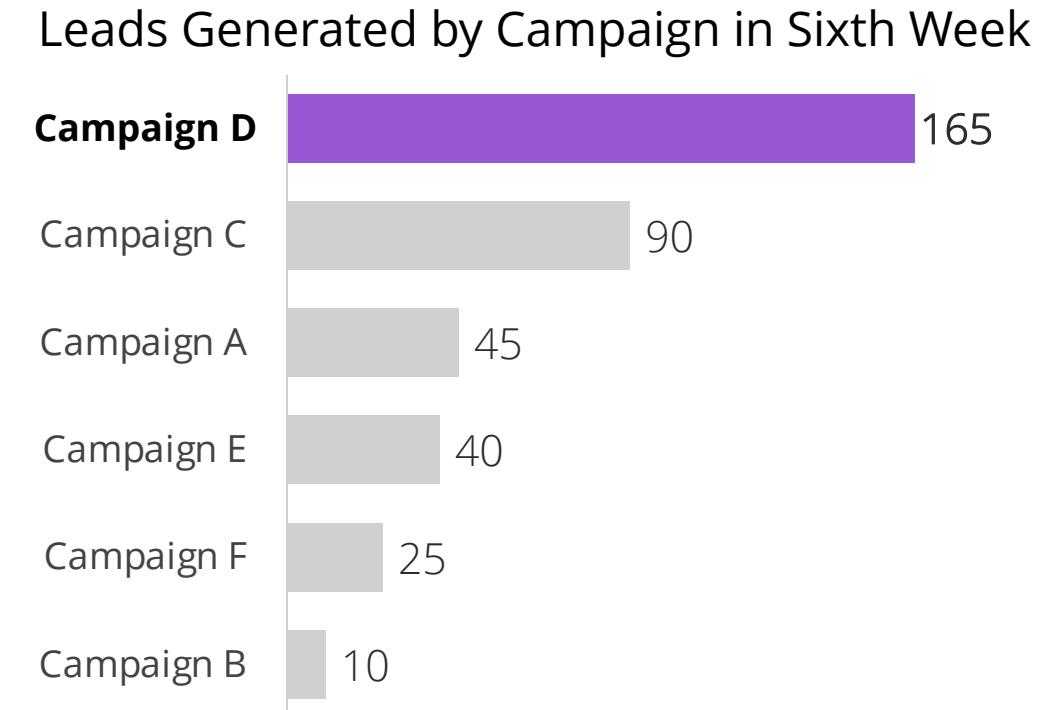
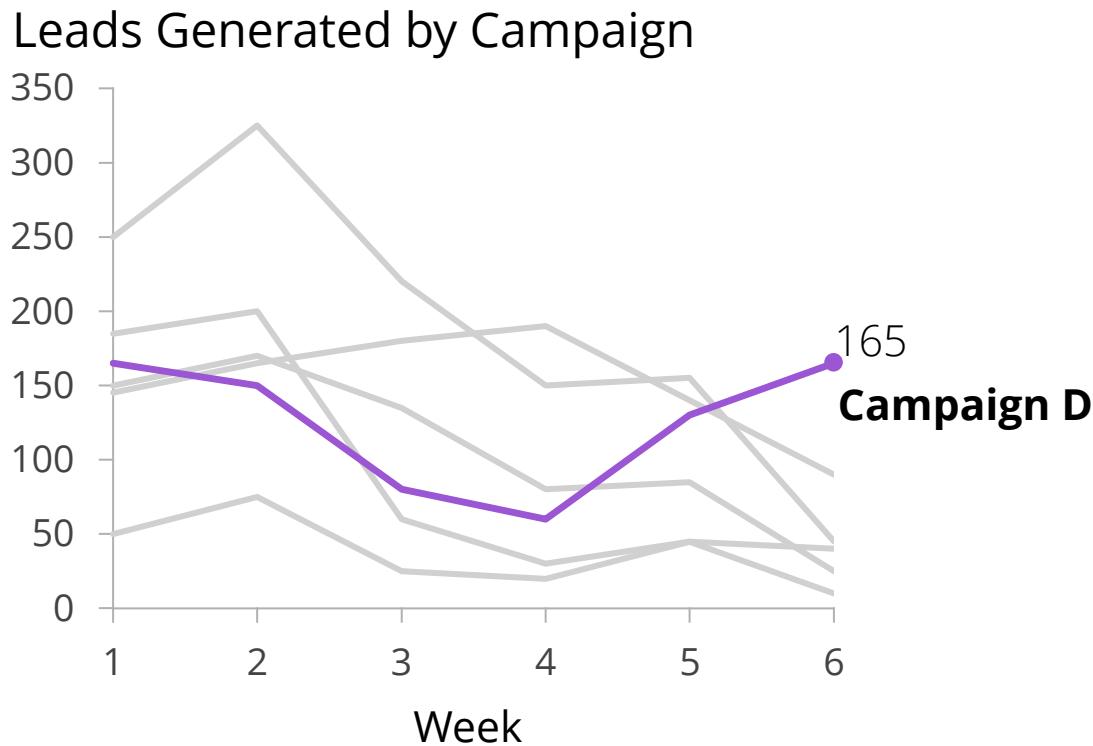
Color name	RGB (1-255)
Black	0, 0, 0
Orange	230, 159, 0
Sky blue	86, 180, 233
Bluish green	0, 158, 115
Blue	0, 114, 178
Vermillion	213, 94, 0

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



To draw attention to what's important



You can also use pre-attentive attributes in text.

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

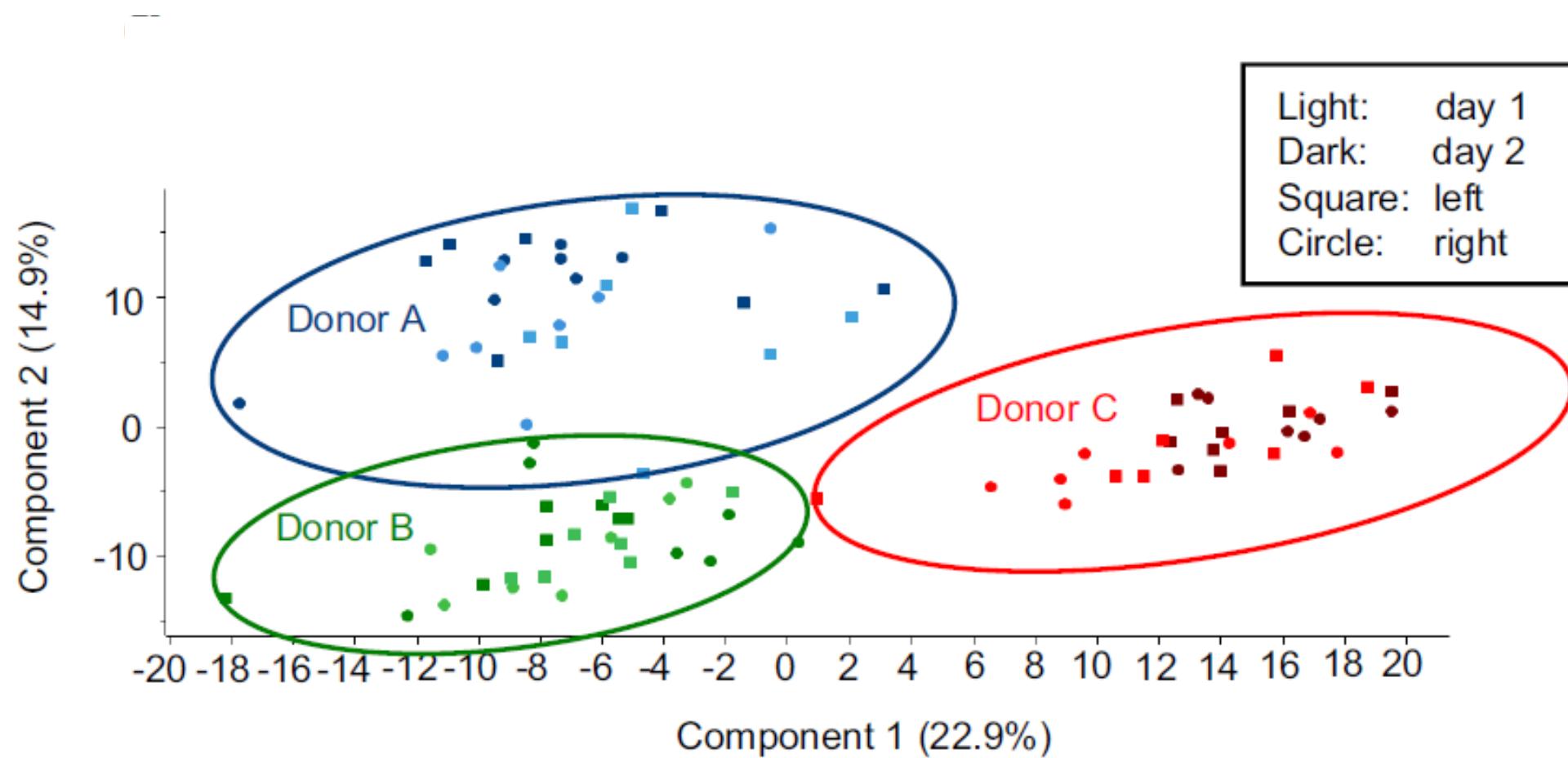
Without pre-attentive attributes you just have plain text. Or a plain figure. There is no guidance for the reader or viewer.

With pre-attentive attributes, you can:

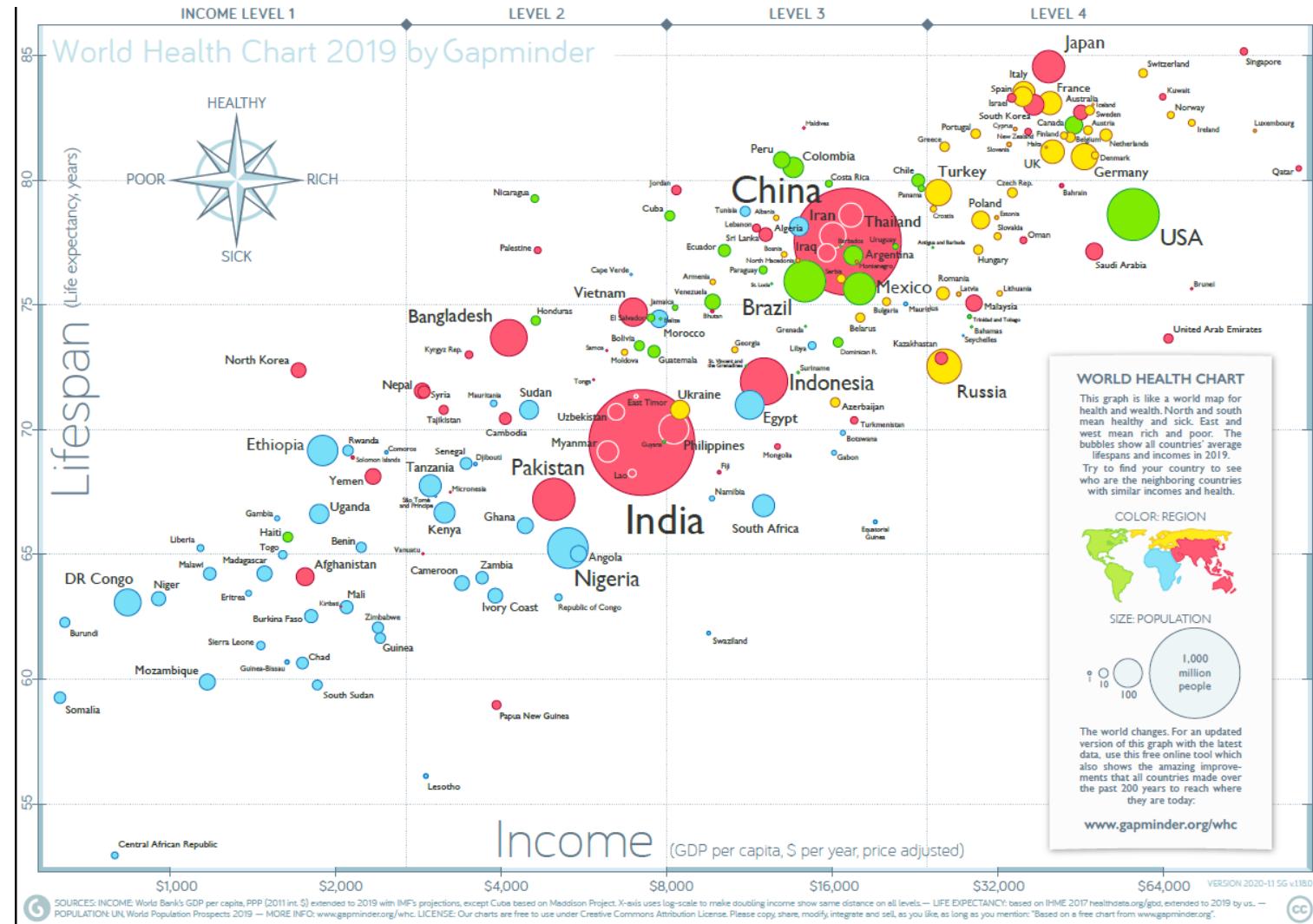
- (1) *Draw your audience attention to where you want them to look*, and
- (2) *Create a visual hierarchy of information*

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

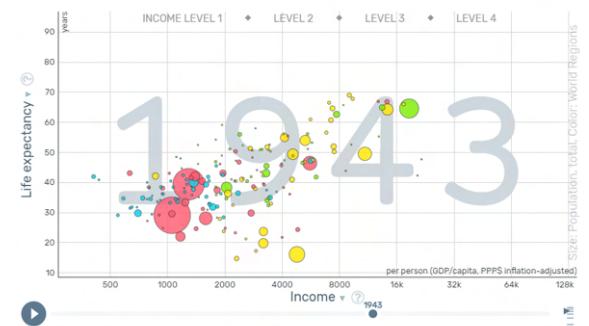
Pre-attentive attributes can create a visual hierarchy



Hans Rosling's Bubble chart



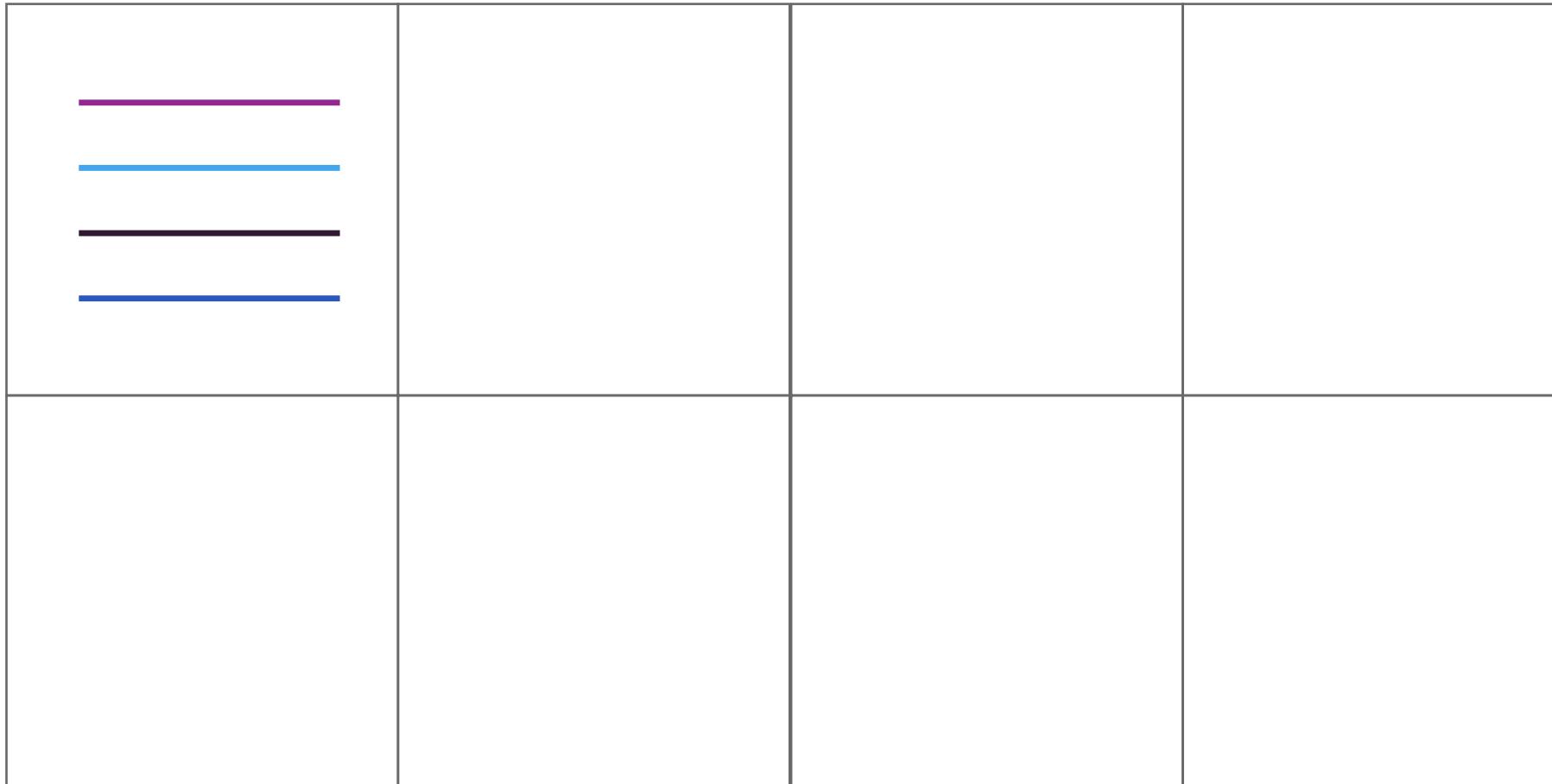
Positie
Grootte
Kleur
Beweging



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

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

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

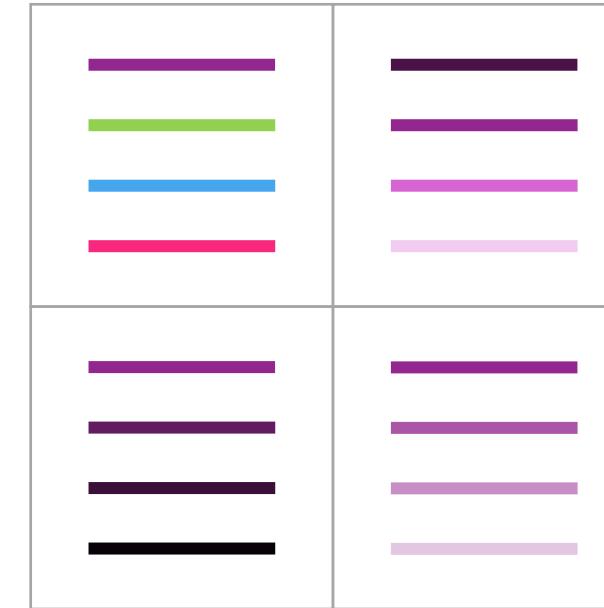
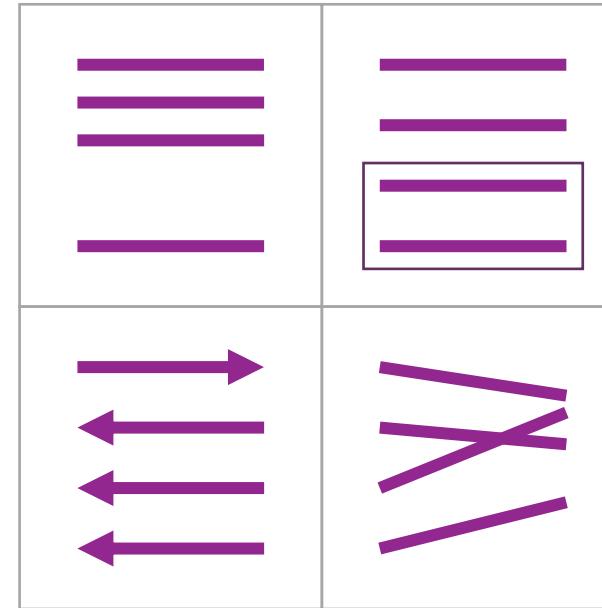
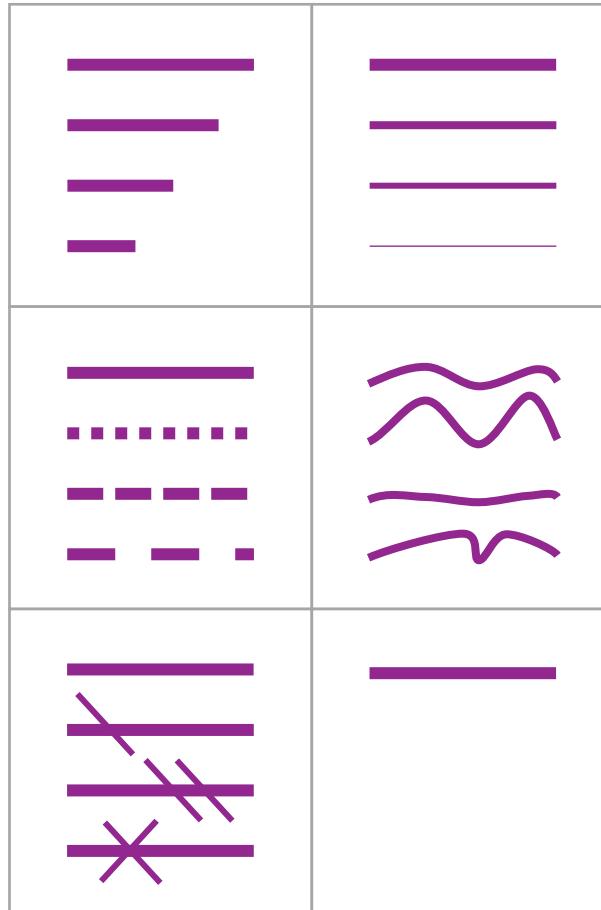


* From each other or at least one should be seen as different from the others

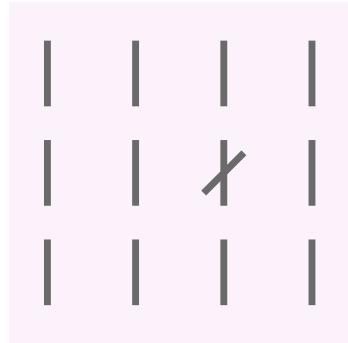


5 min

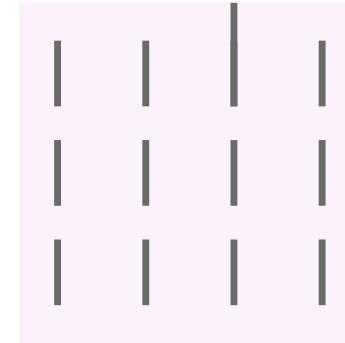
Exercise: draw 4 lines using 1 visual attribute at a time to differentiate them*



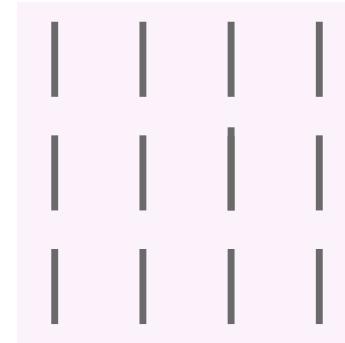
Pre-attentive attributes make your audience see what you want them to see before even seeing it!



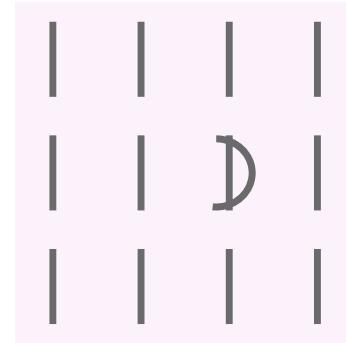
Orientation



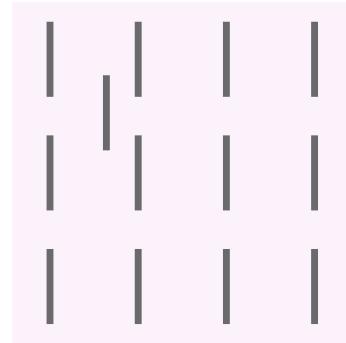
Length



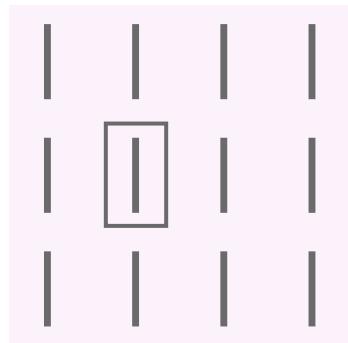
Width



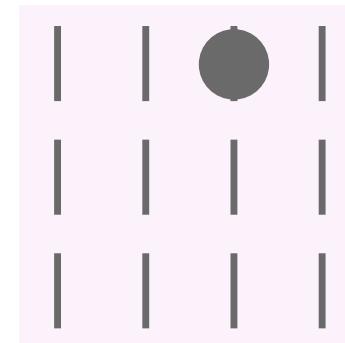
Curvature



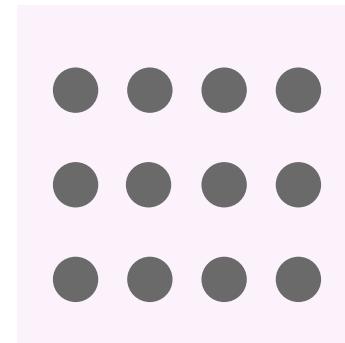
Position



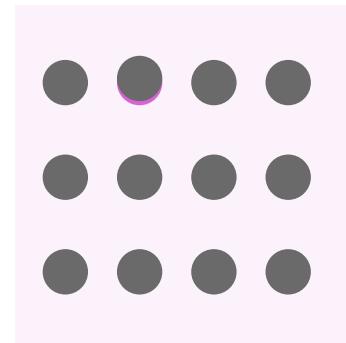
Enclosure



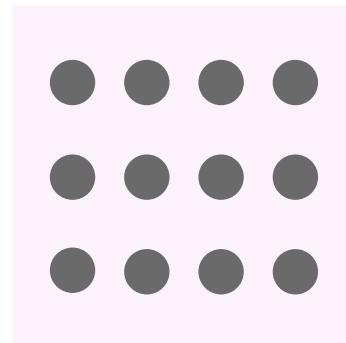
Shape



Size



Color



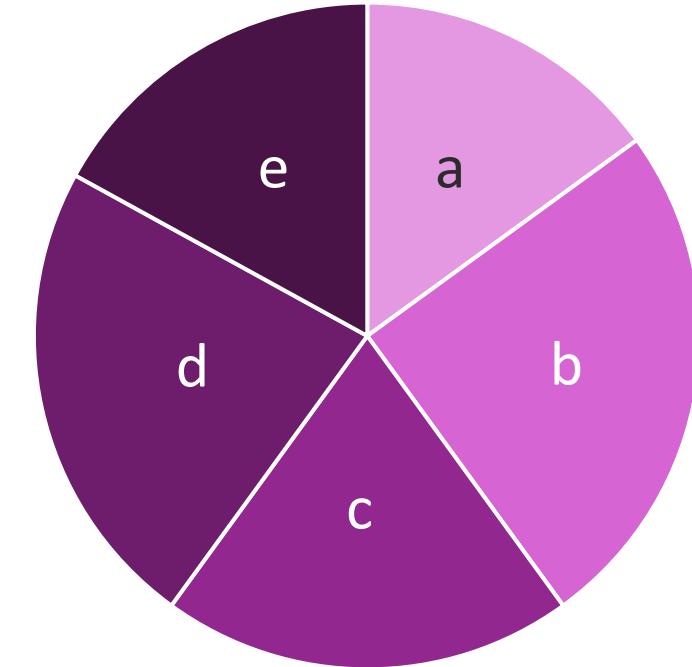
Motion

Pre-attentive attributes can be used to map quantities, e.g. by length or by angle

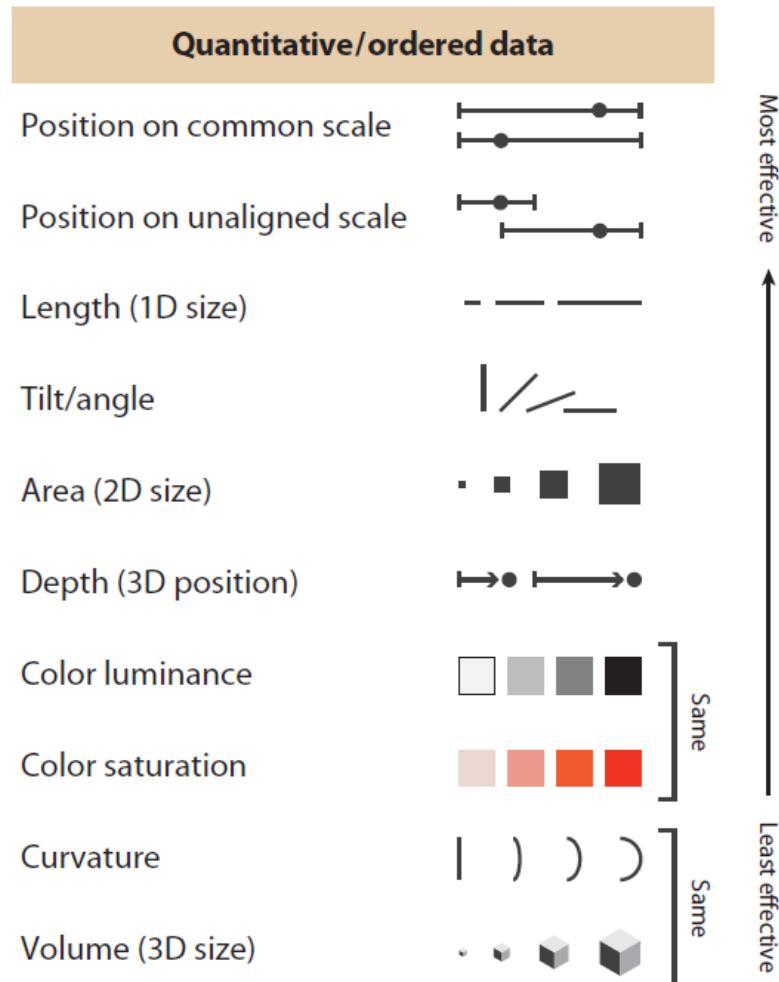
Which bar is highest?



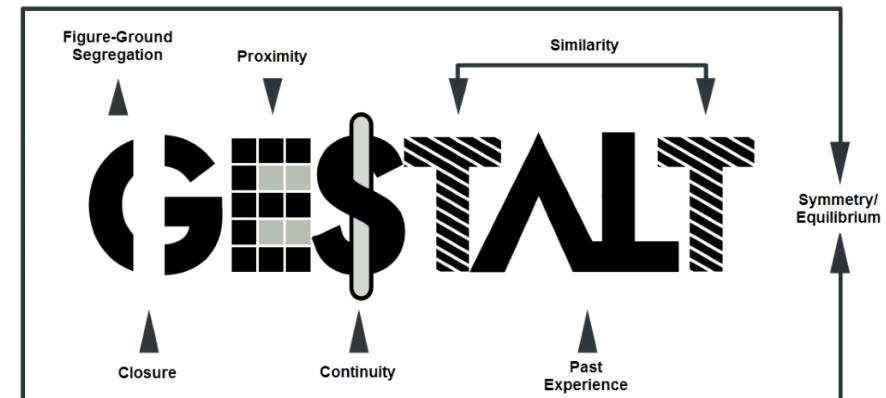
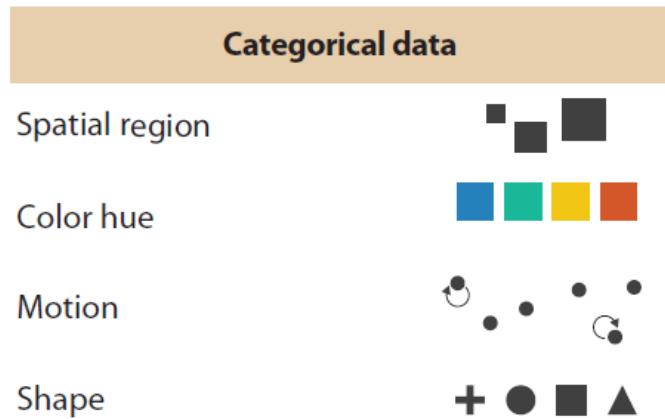
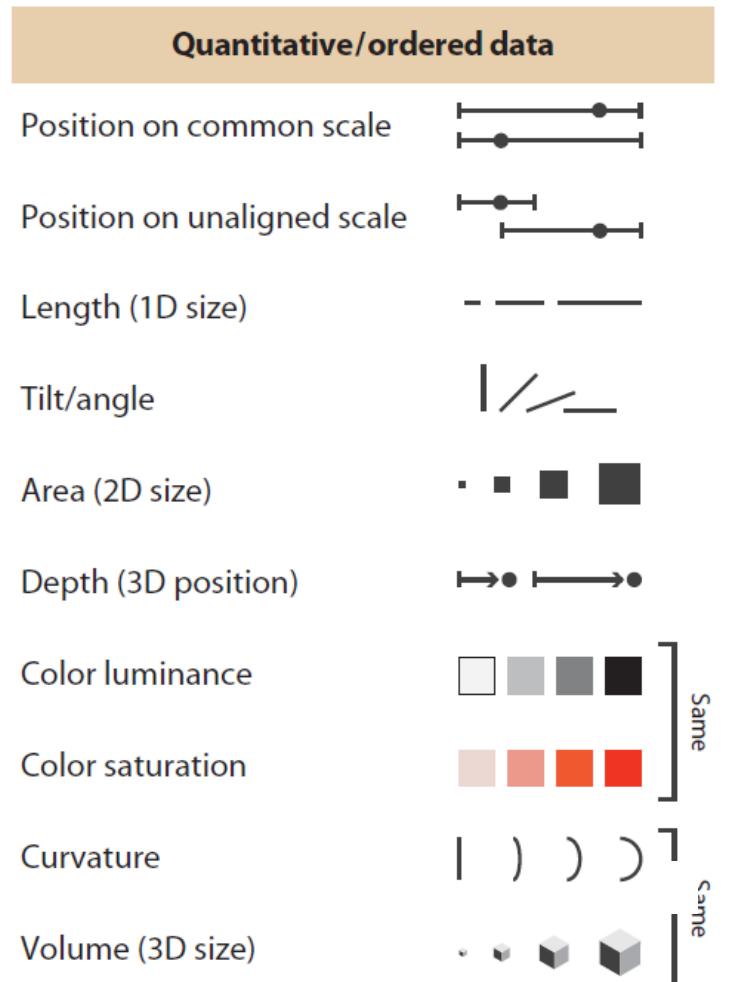
Which slice is biggest?



Some attributes are more effective to differentiate **quantities**



Some attributes are more effective to differentiate **quantities** or **categories**

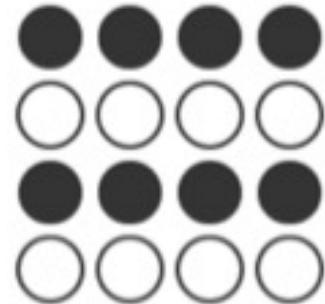


Gestalt principles: The whole is other than the sum of the parts.

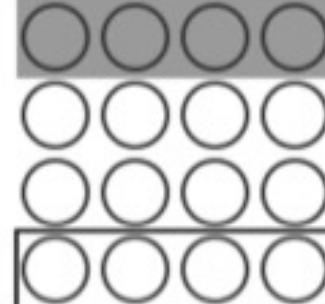
Proximity



Similarity



Enclosure



Symmetry

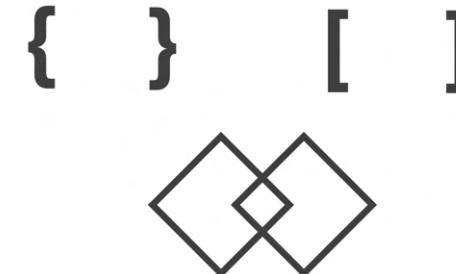
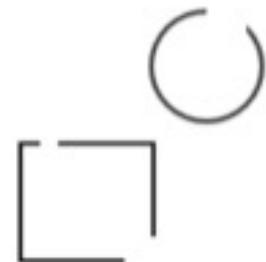


Figure-ground



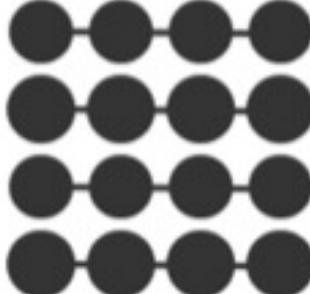
Closure



Continuity



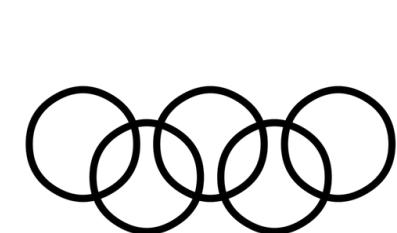
Connection



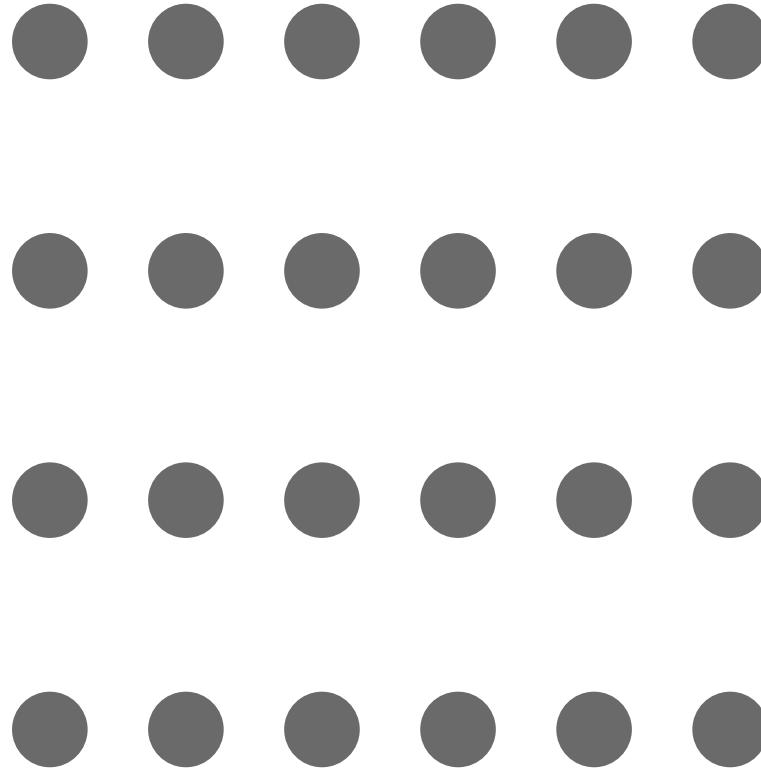
Common Fate



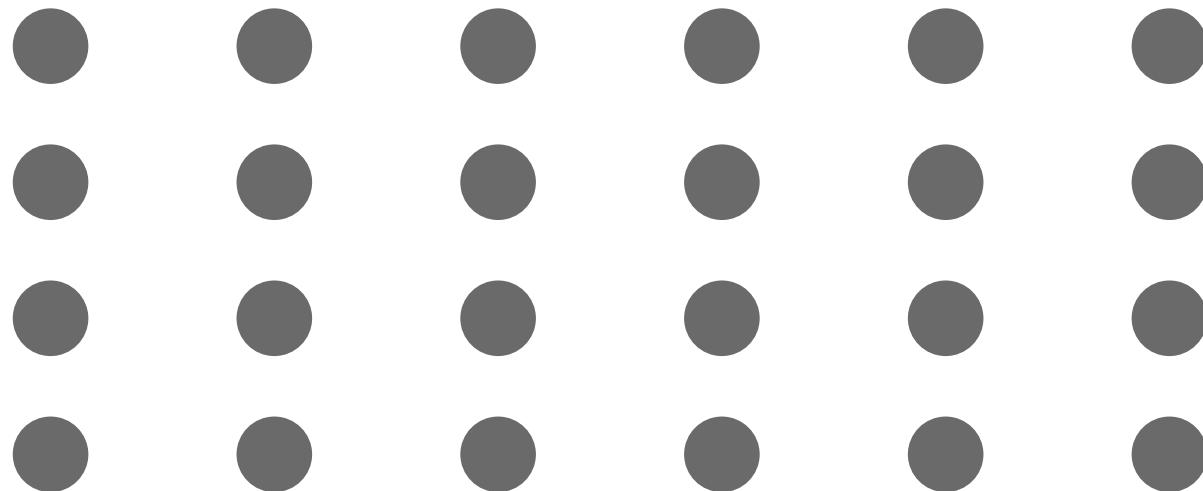
Simplicity



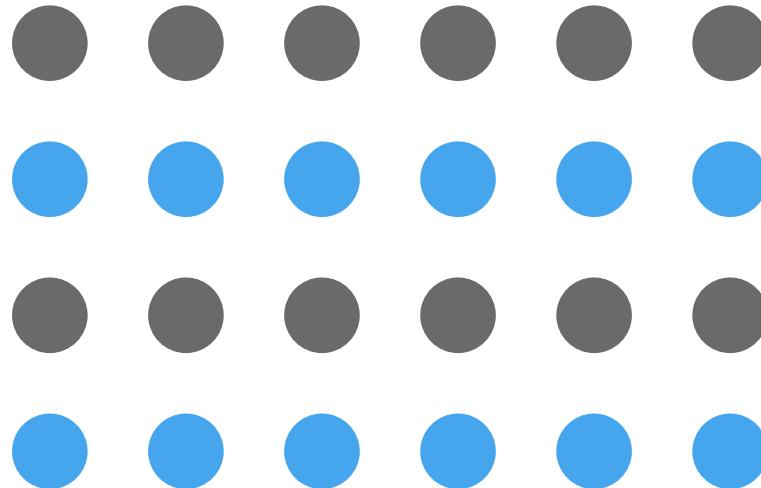
Do you see 4 rows or 6 columns?



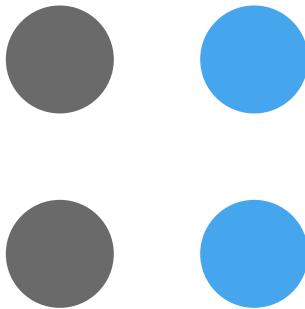
And now?



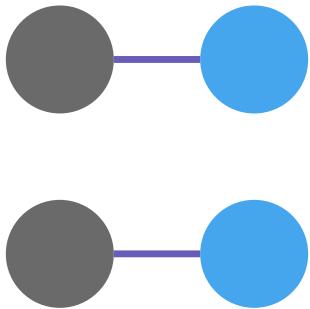
Do you see 4 rows or 6 columns?



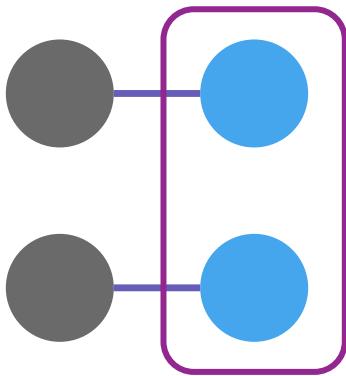
Which dots belong together?



And now...?



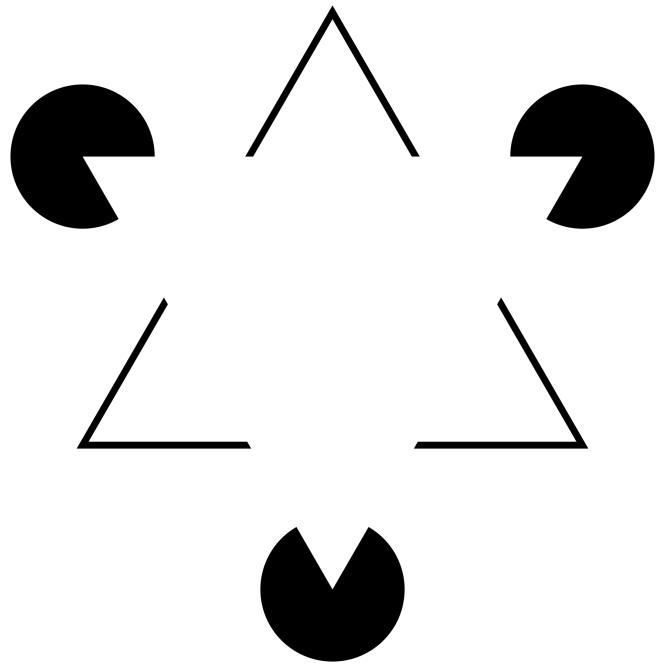
And now...?



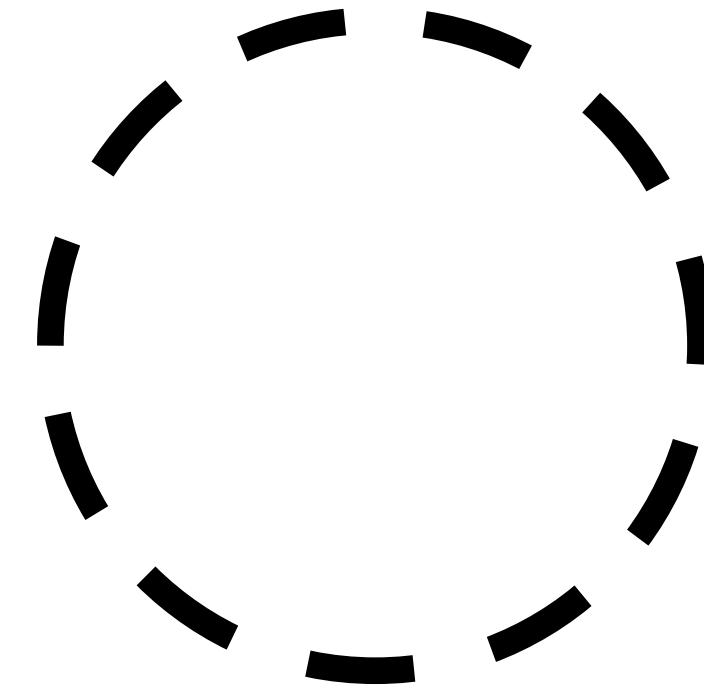
Can you read this?

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!

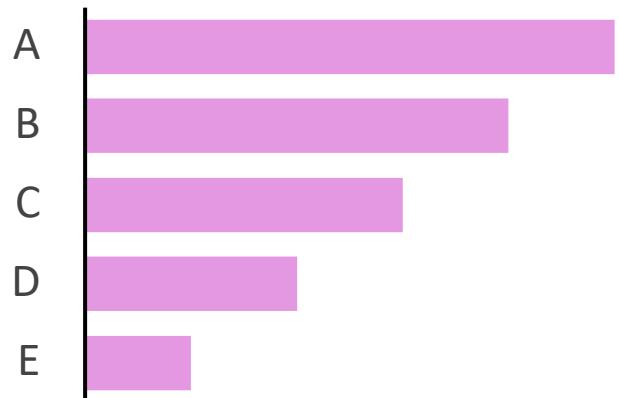
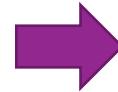
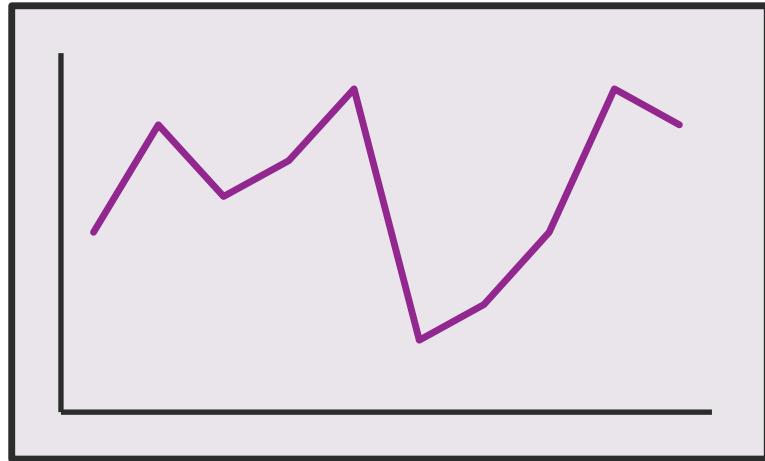


Continuity



Closure

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

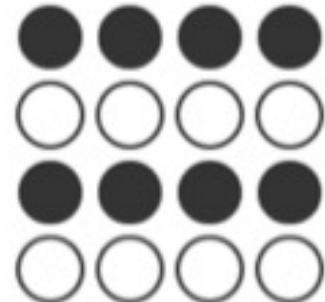


Gestalt principles: The whole is other than the sum of the parts.

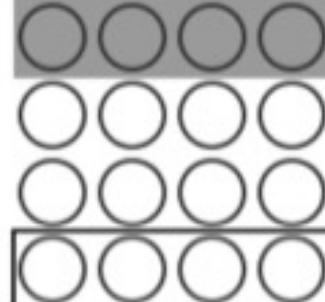
Proximity



Similarity



Enclosure



Symmetry

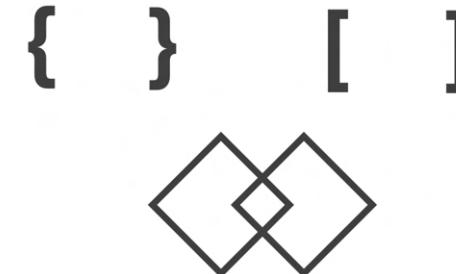
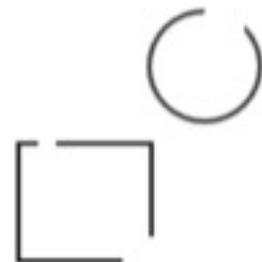


Figure-ground



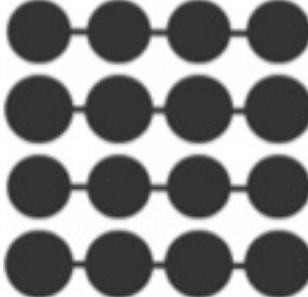
Closure



Continuity



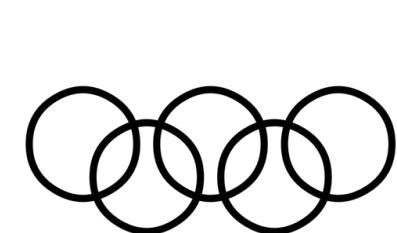
Connection



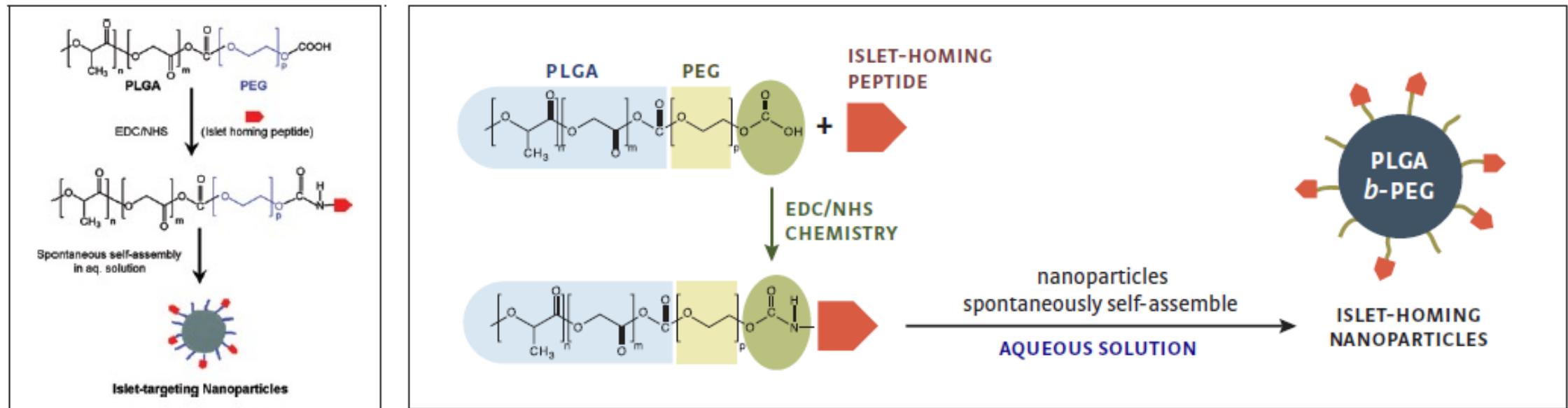
Common Fate



Simplicity



The Gestalt principles also apply to graphical abstracts



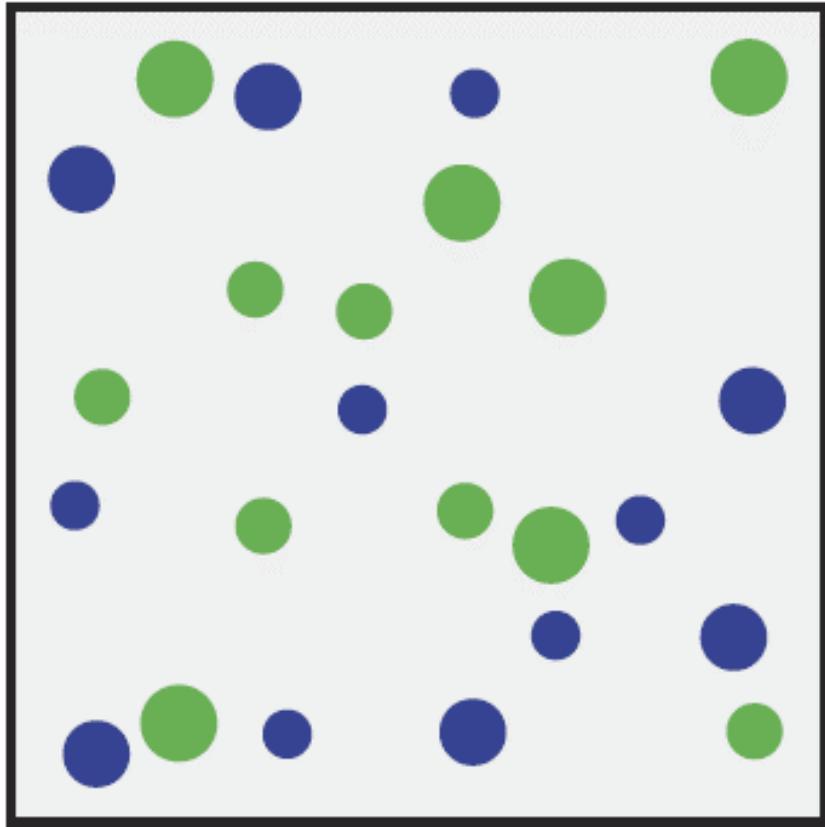
Shapes and colors used to identify similar structures.

But...

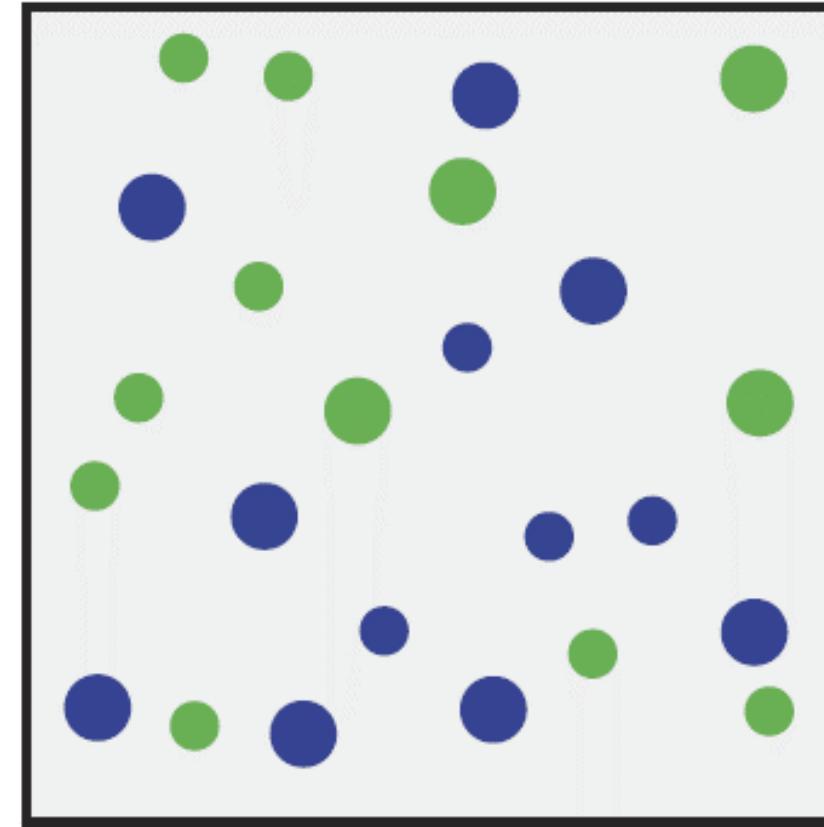


Which panel has the biggest **green** dots?

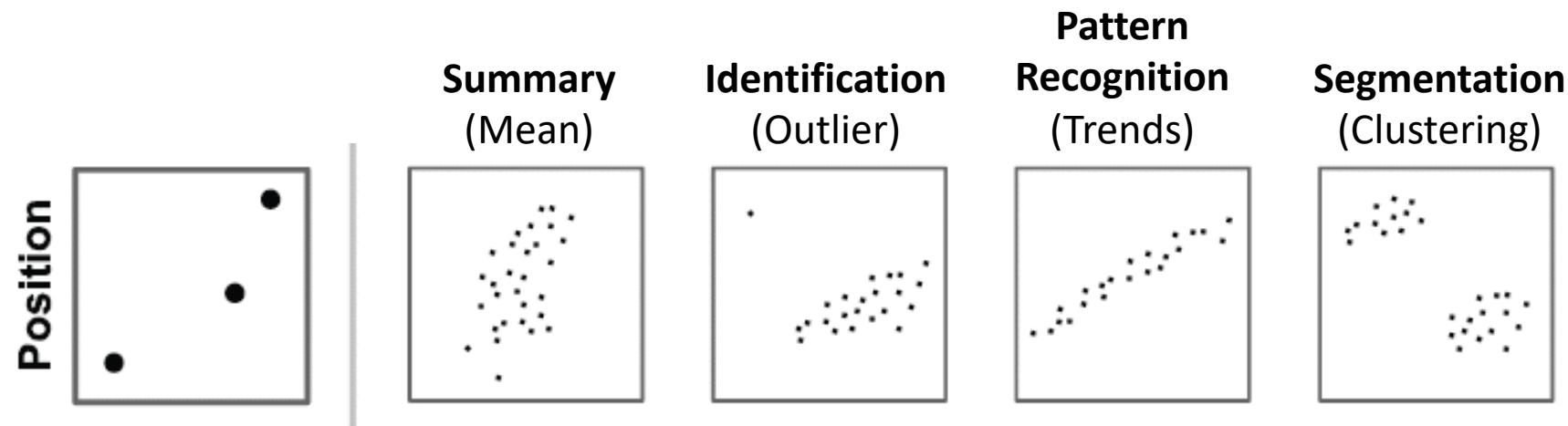
A



B

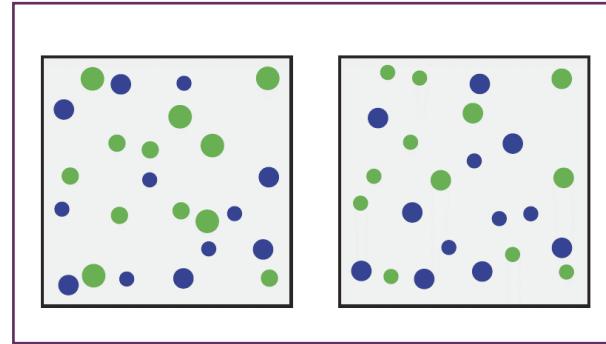
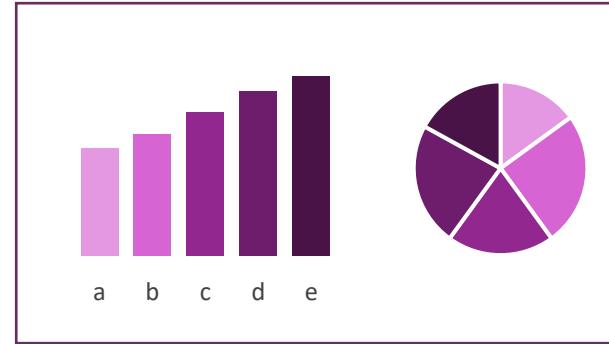
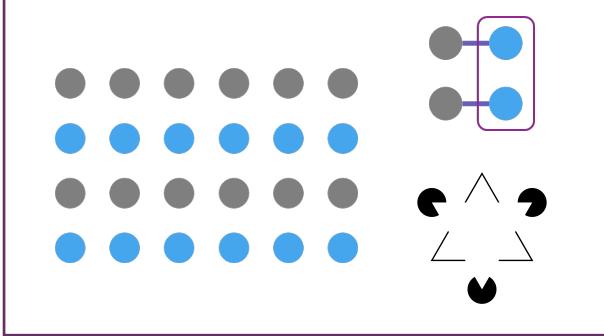


We can quickly deduce basic descriptive statistics from visuals!



The **strengths** of our brain - summary

321654643216465132168461321
646432136748651349687463123
668432616266984892536496874
651687964649843616984616957
465167412323161689312439873
421432875943869234987596387



1.

See before seeing it
(Pre-attentive attributes)

2.

See what belongs
together
(Gestalt Principles)

3.

Deduce quantitative
information from visual
shapes

4.

Perform basal descriptive
statistics

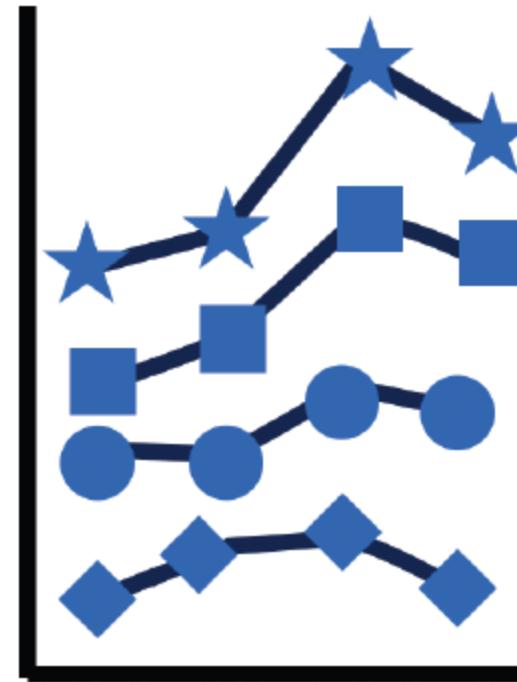
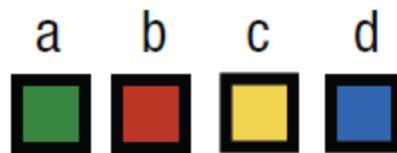
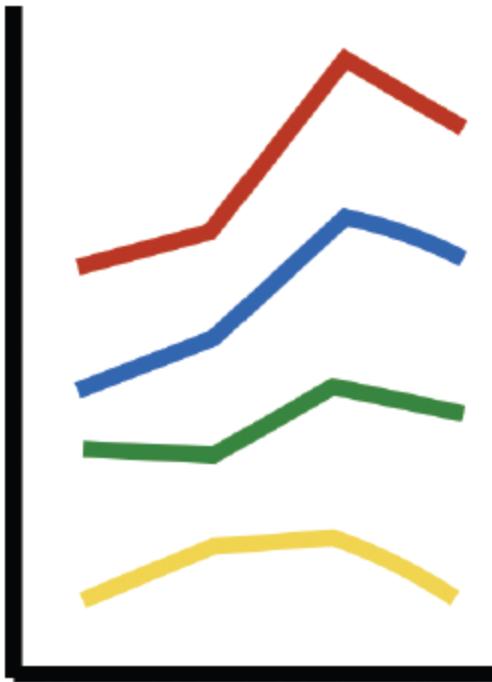
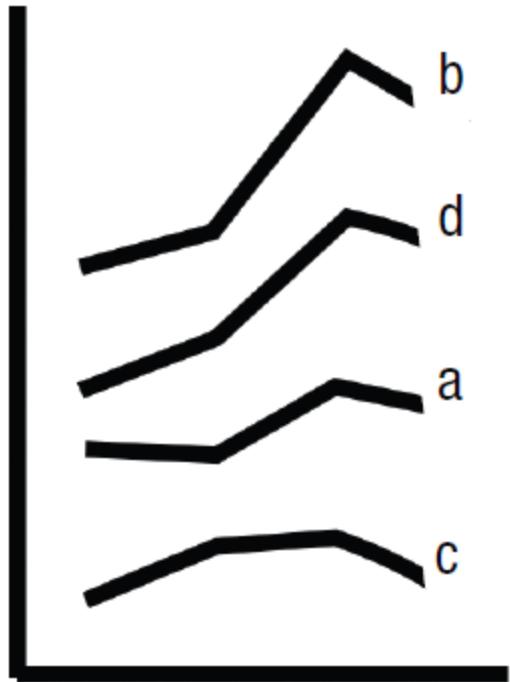


Visual perception

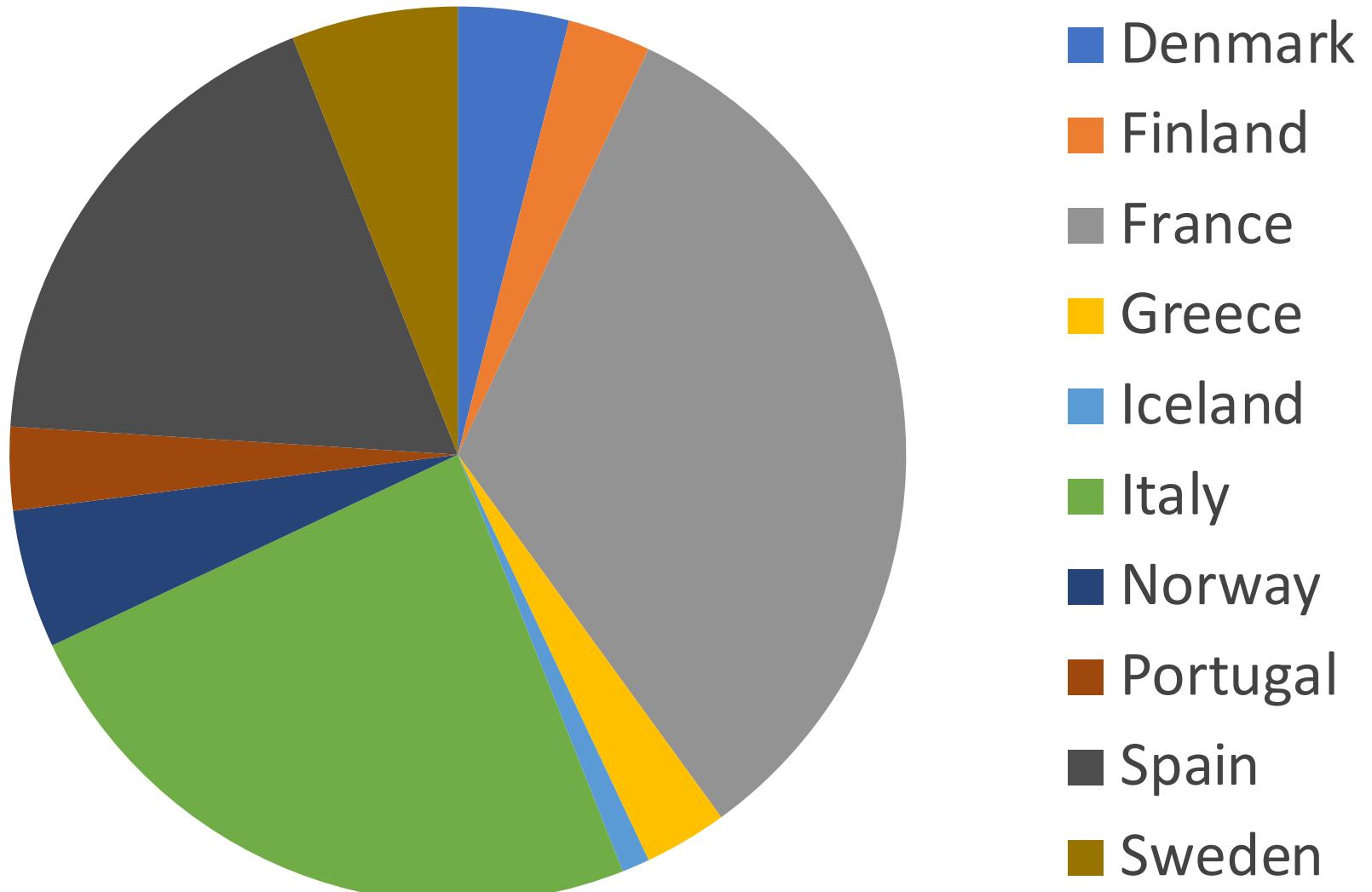
The weaknesses of our brain

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

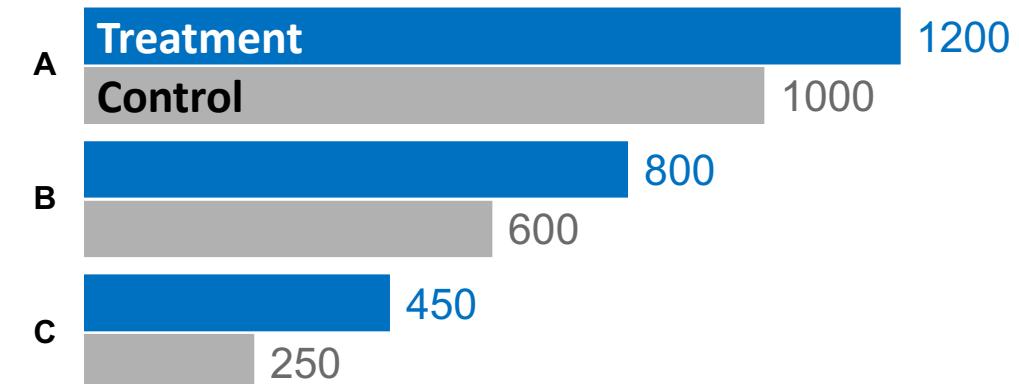
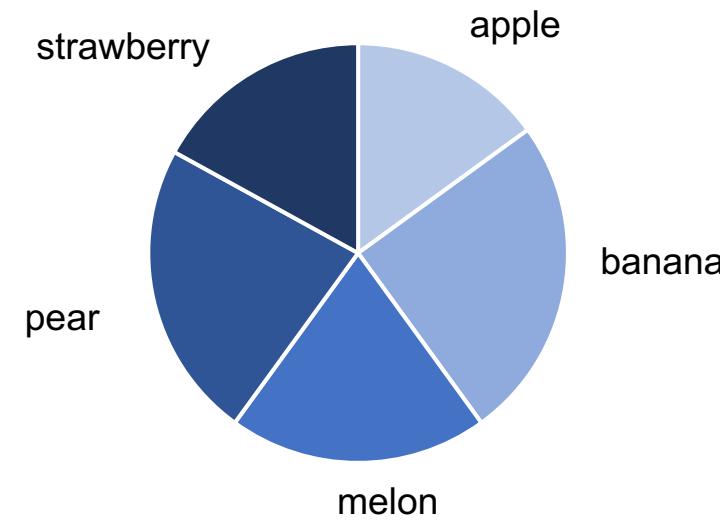
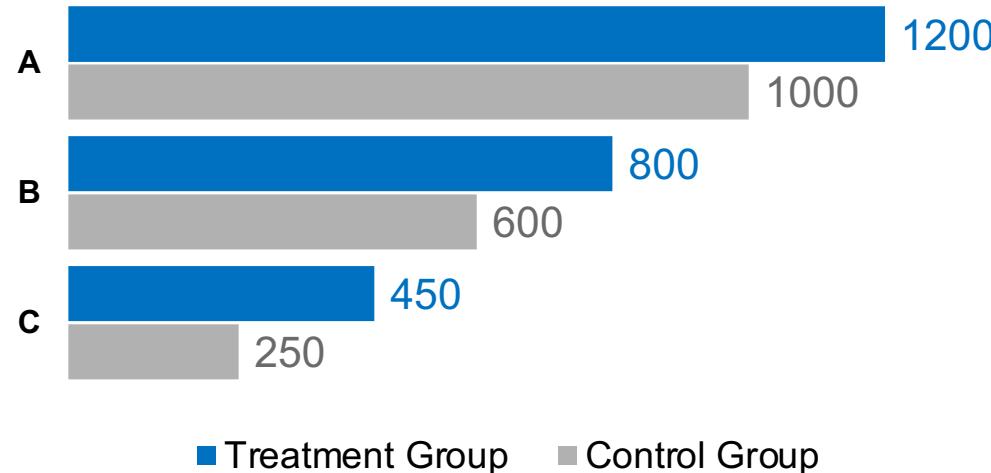
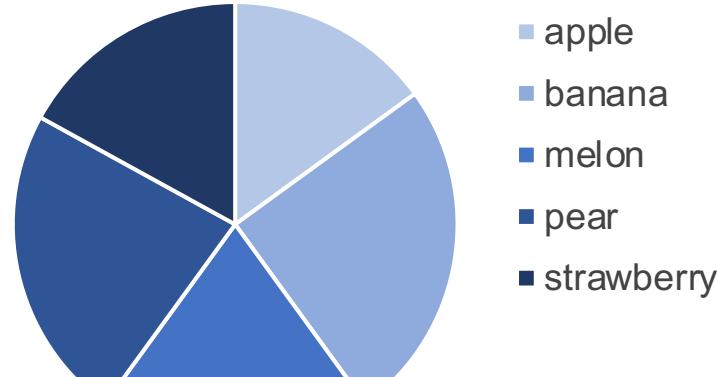
What is the order from top-to-bottom?



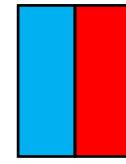
Norway has more widgets than Finland?



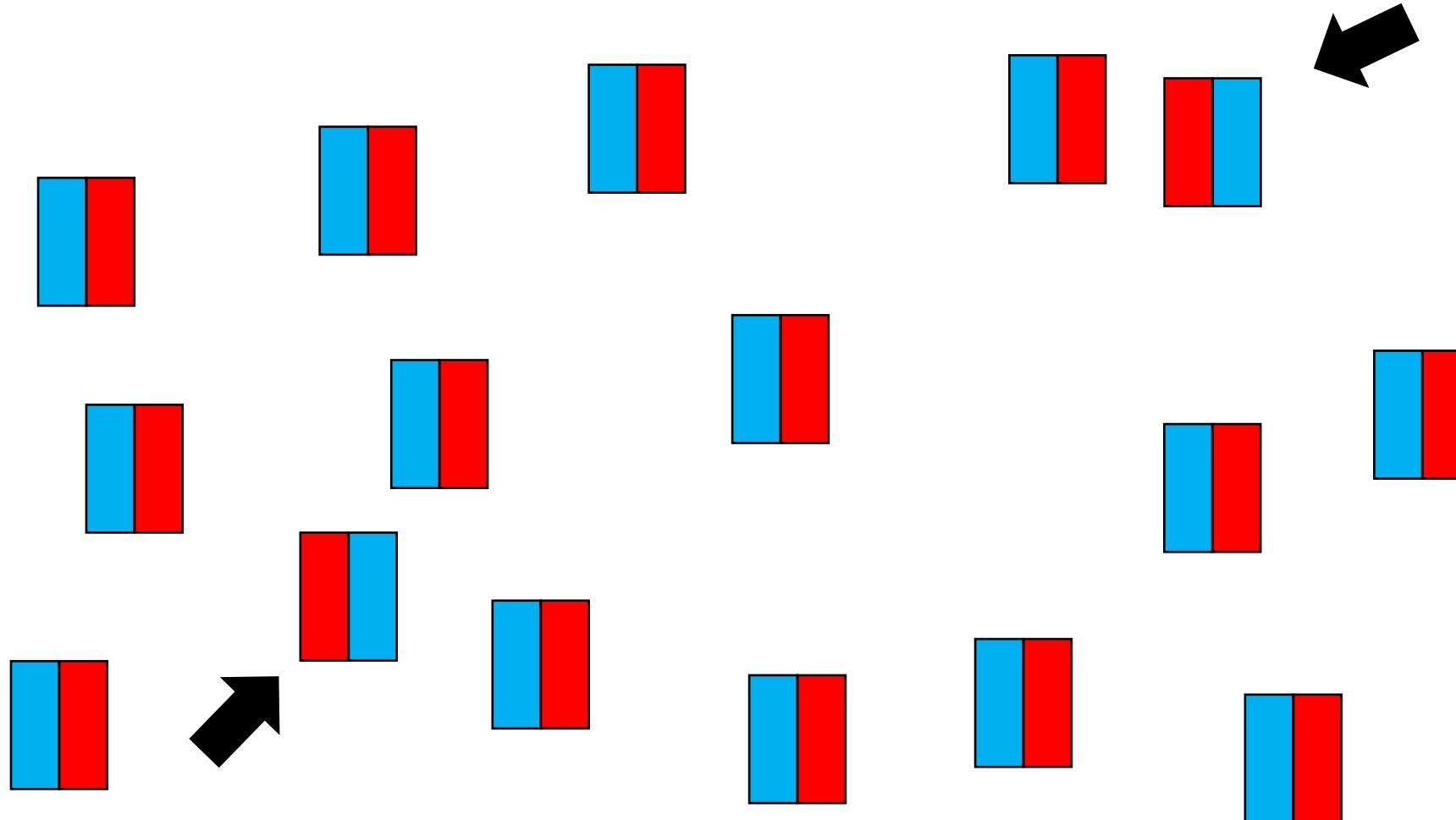
Avoid legends and label directly



Comparing one set of colors is easy...



But: which one is not like the others?



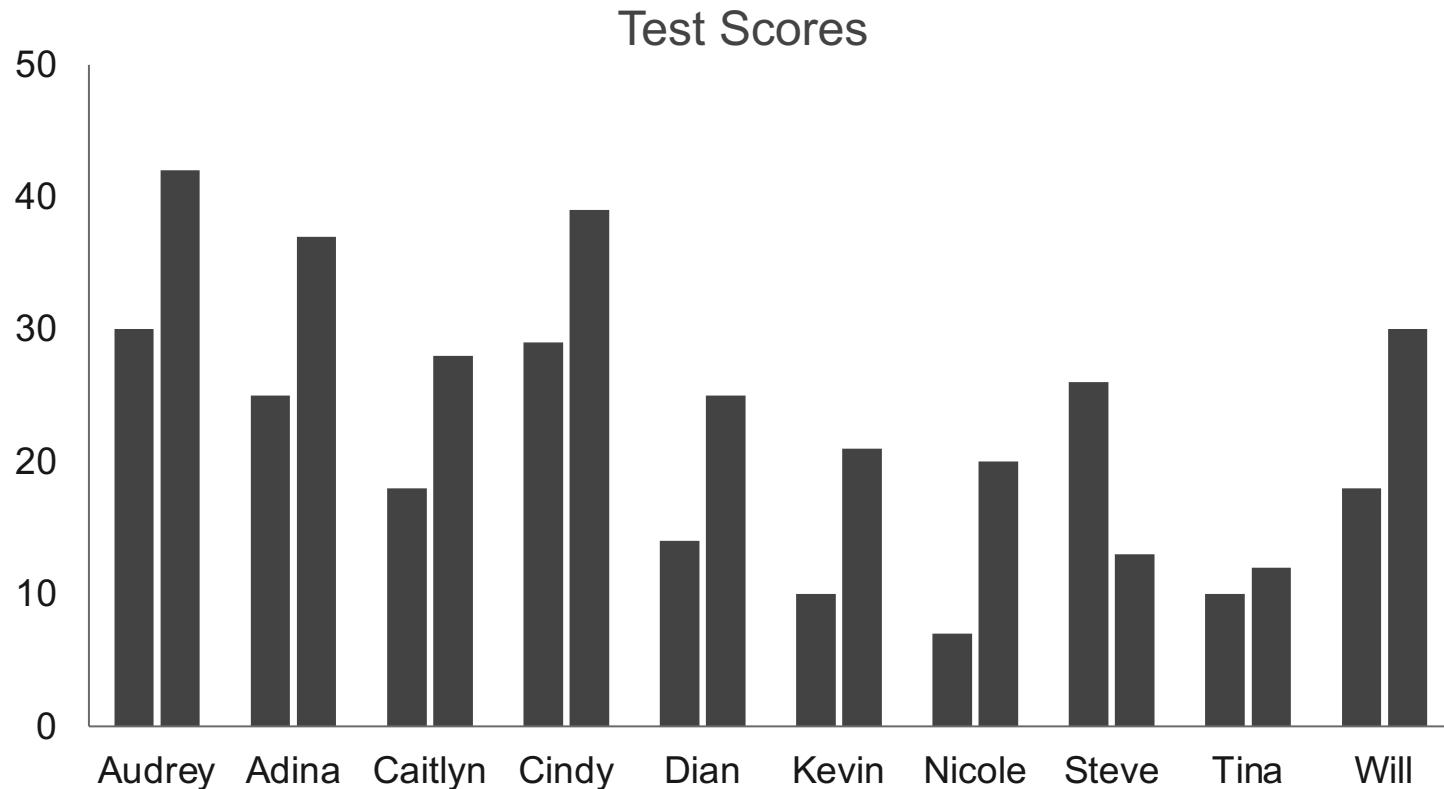
Do you see small-big or big-small?



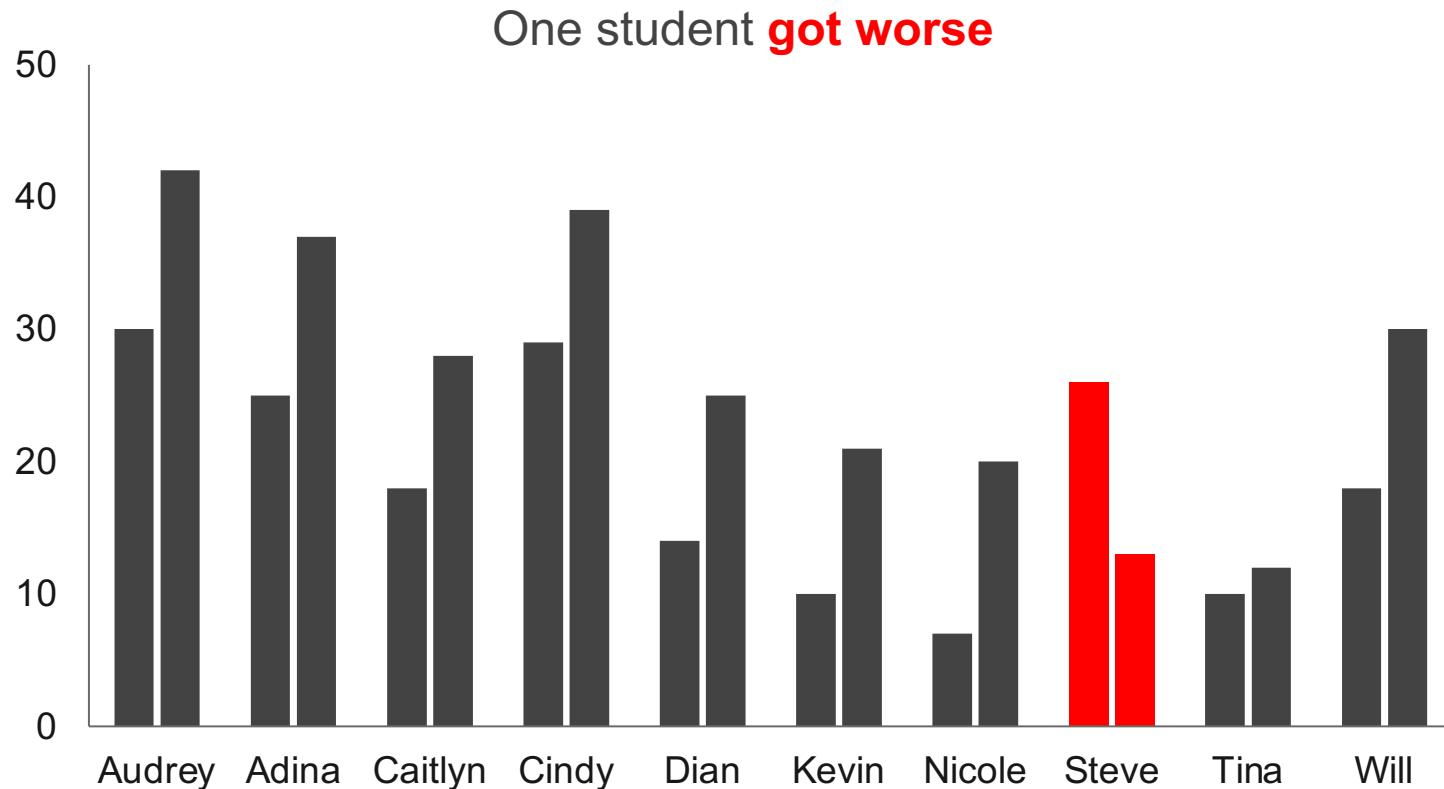
Which one is not like the others?



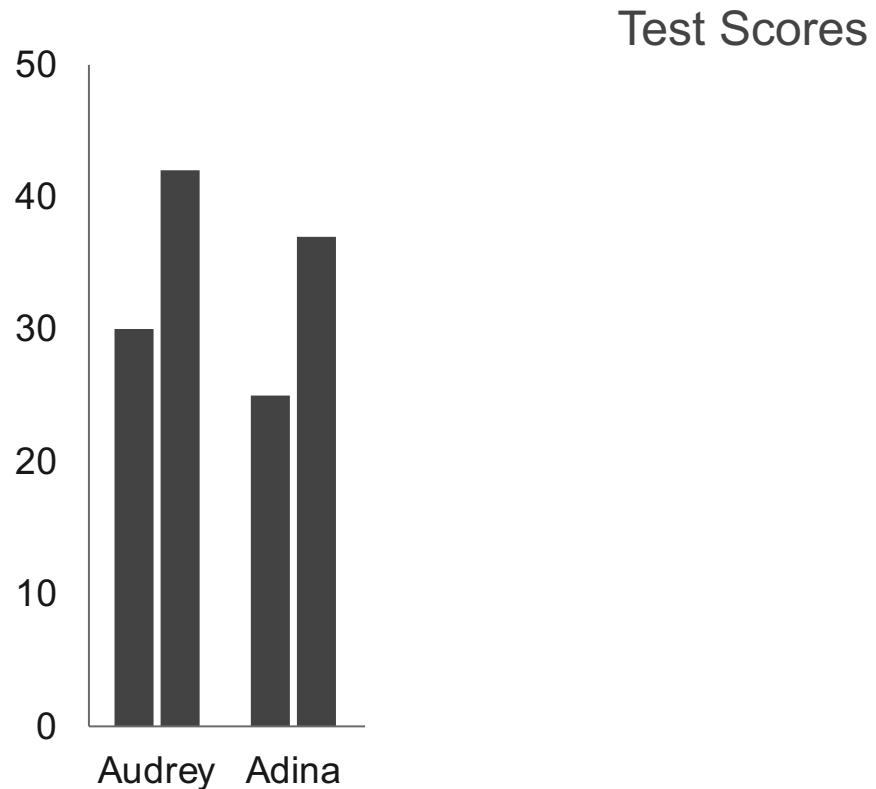
Which student got worse?



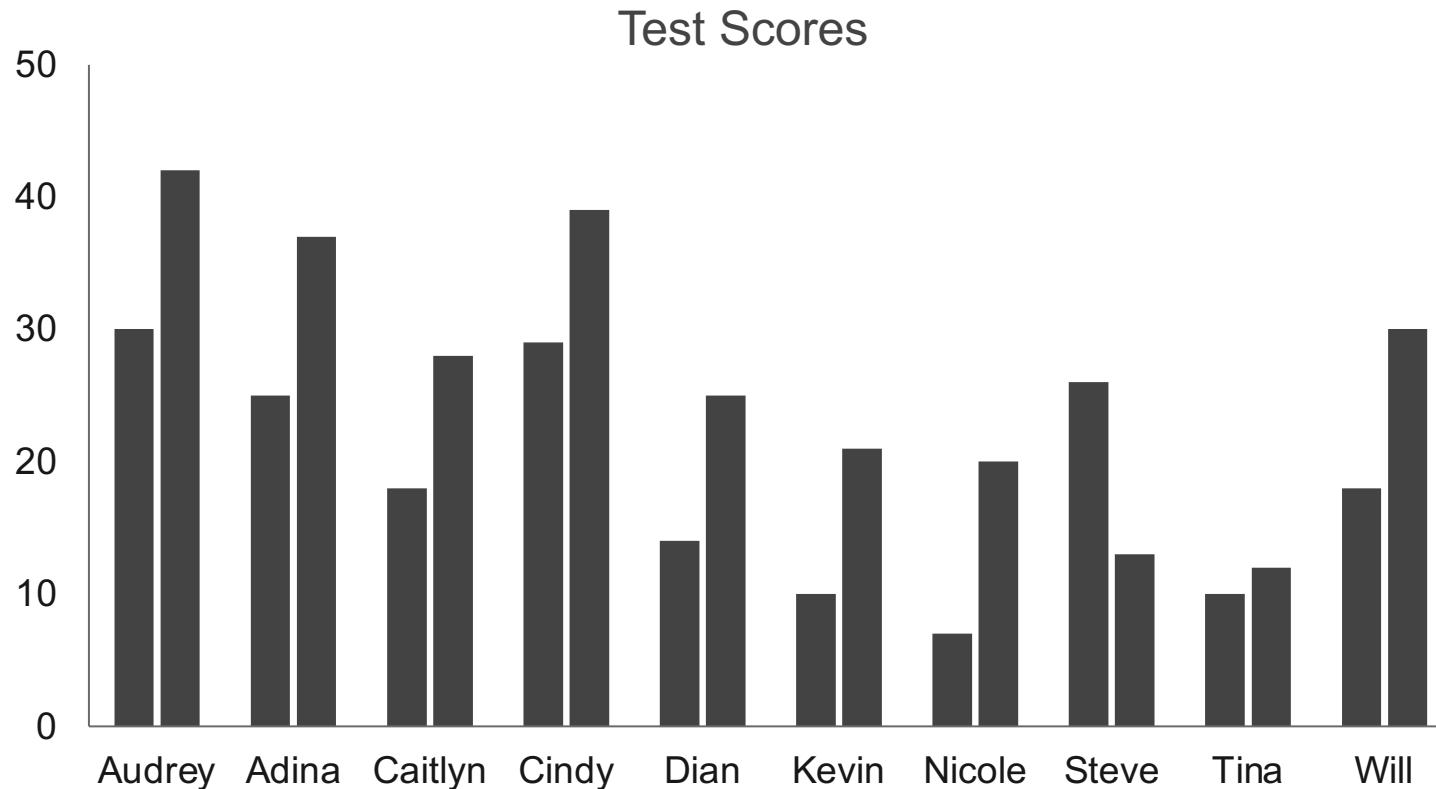
Which student got worse?



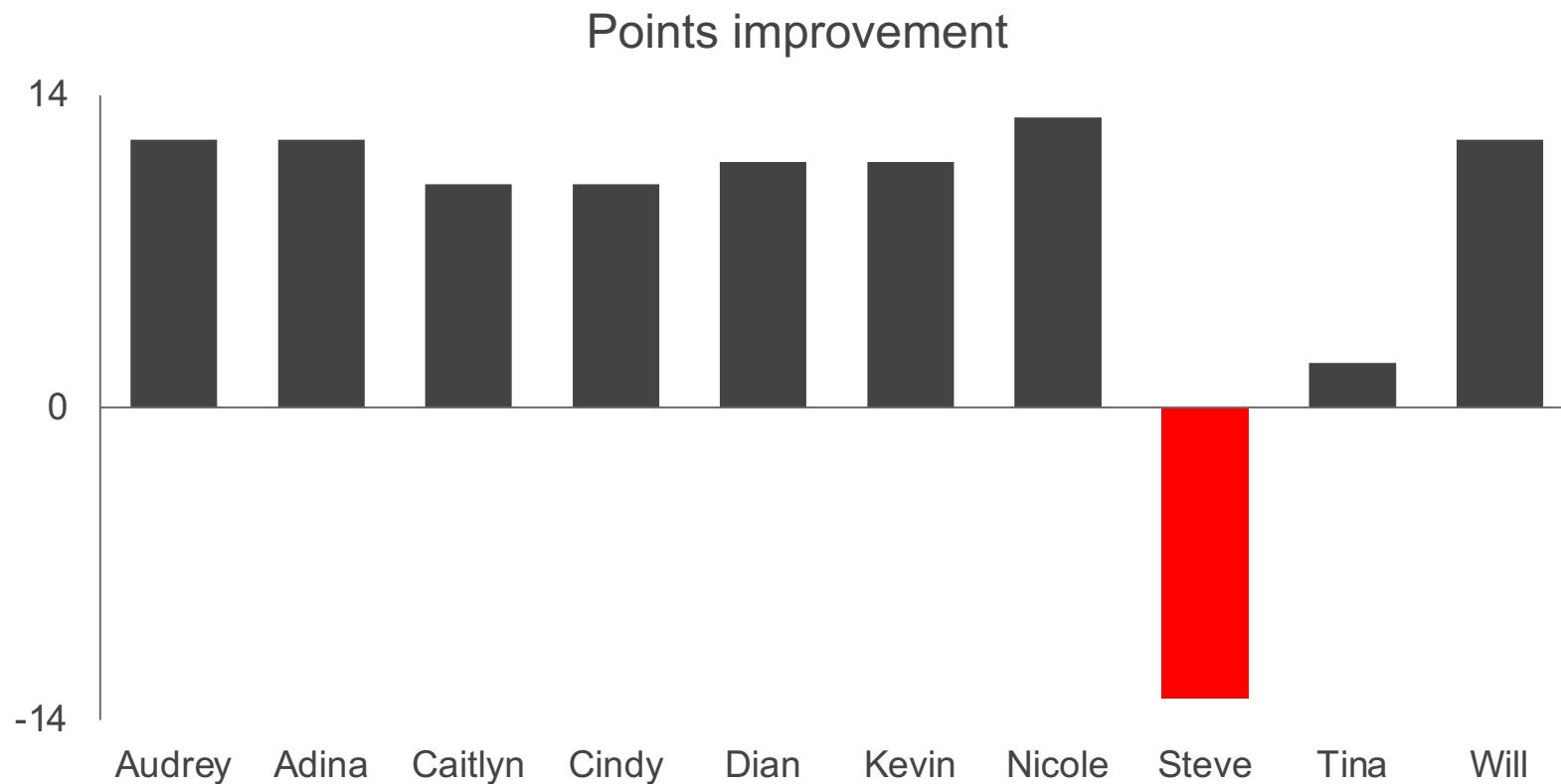
What is the average improvement?



What is the average improvement?



Enable direct comparisons...



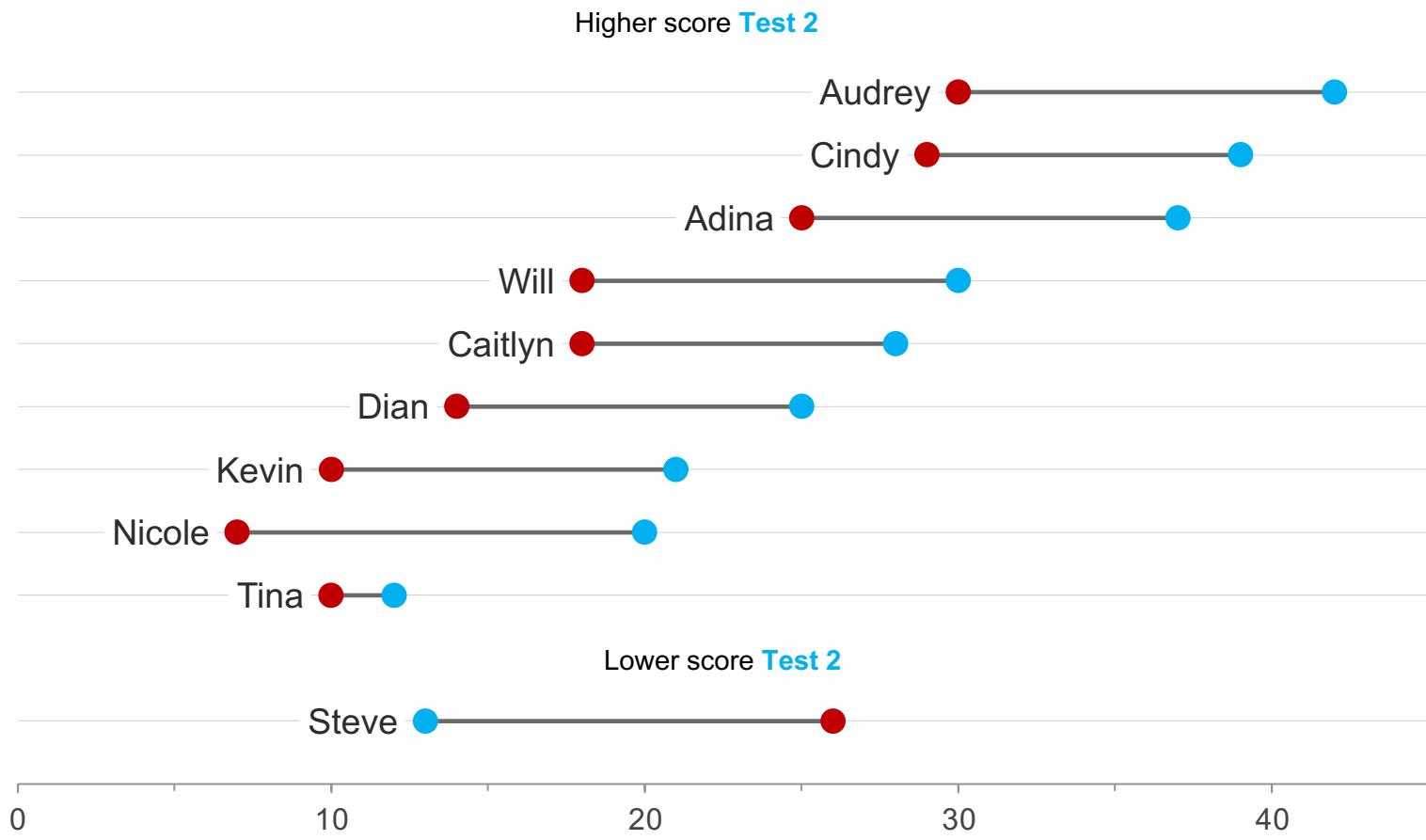
Enable direct comparisons...



Enable direct comparisons

Scores for **Test1** and **Test2**

Which student got worse?



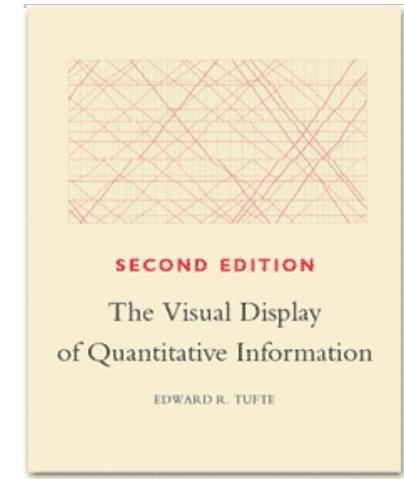
Where to look?



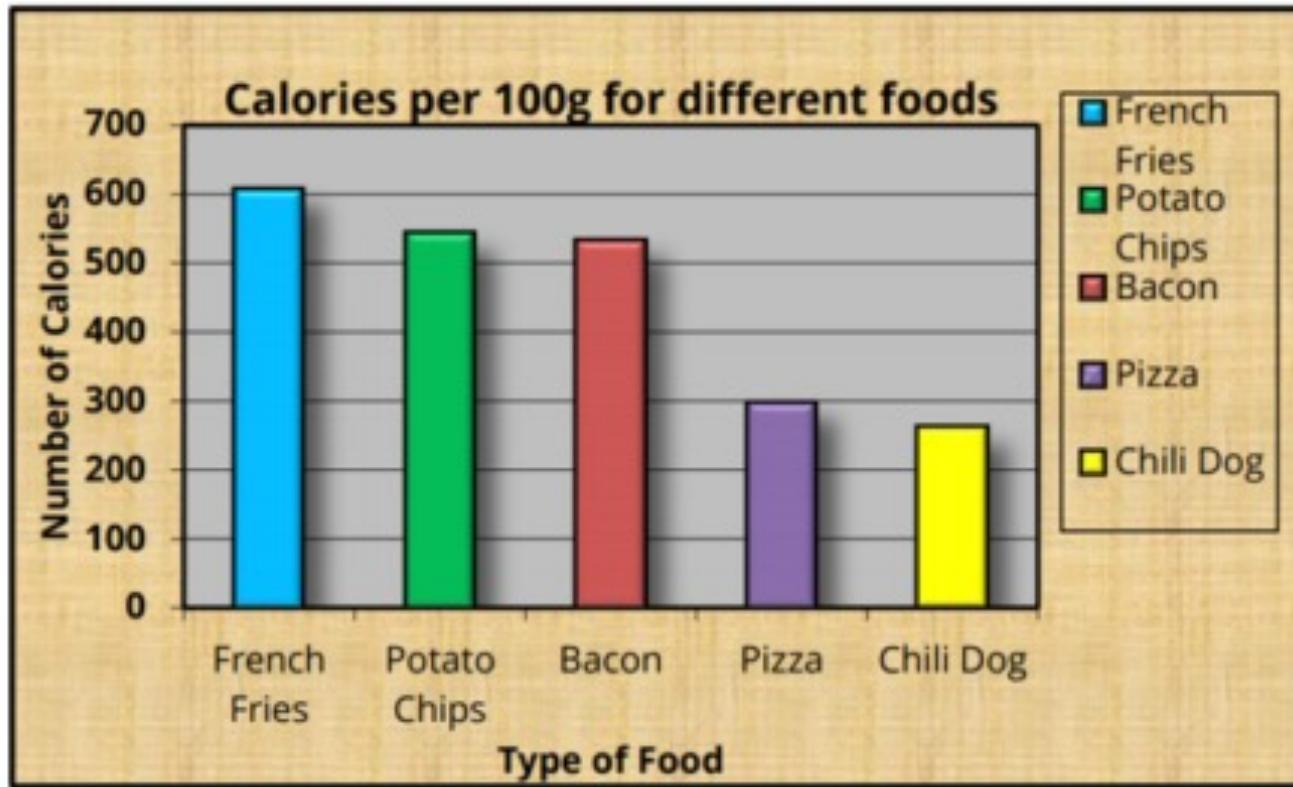
The data-ink ratio

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



De data-ink ratio: remove to improve

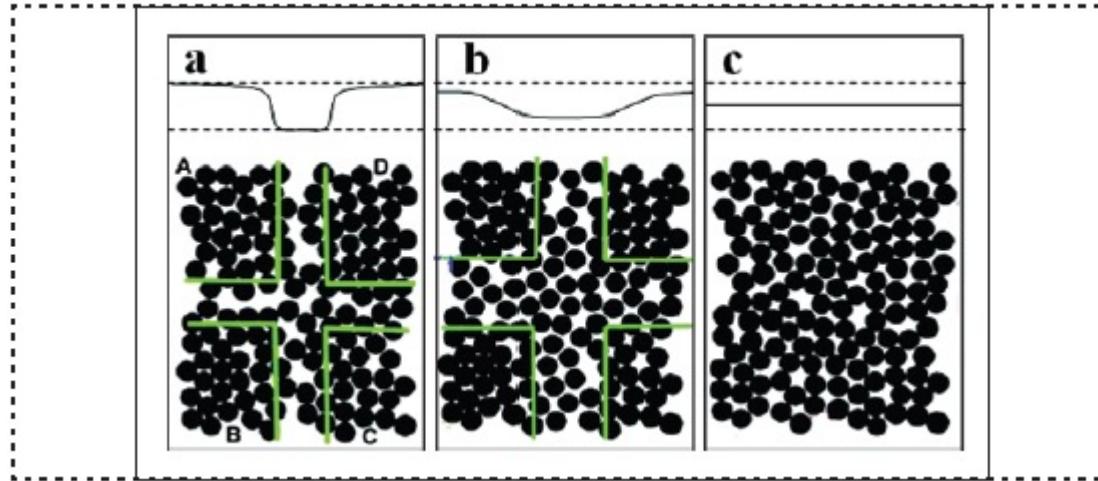
Remove
to improve
(the **data-ink** ratio)

Data-ink ratio also applies to tables

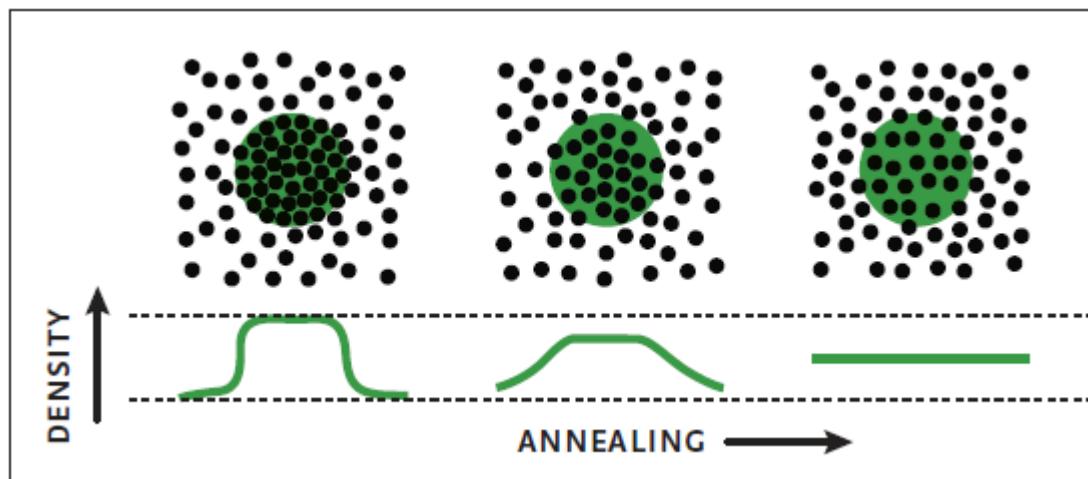
Student	Test1	Test2	Difference
Audrey	30.00	42.00	12.00 pts
Adina	25.00	37.00	12.00 pts
Caitlyn	18.00	28.00	10.00 pts
Cindy	29.00	39.00	10.00 pts
Dian	14.00	25.00	11.00 pts
Kevin	10.00	21.00	11.00 pts
Nicole	7.00	20.00	13.00 pts
Steve	26.00	13.00	-13.00 pts
Tina	10.00	12.00	2.00 pts
Will	18.00	30.00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

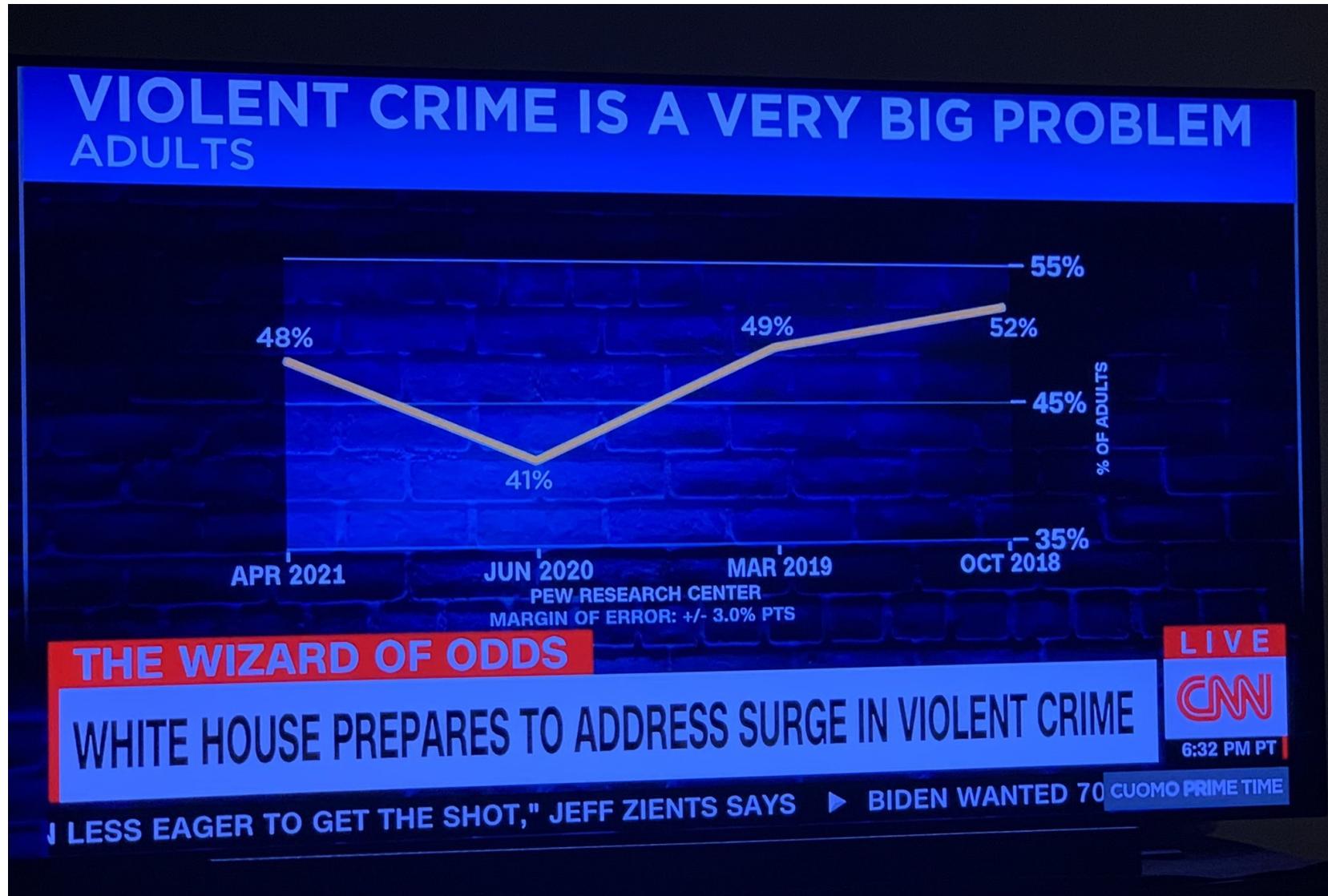


Too many interior boxes clutter the layout.



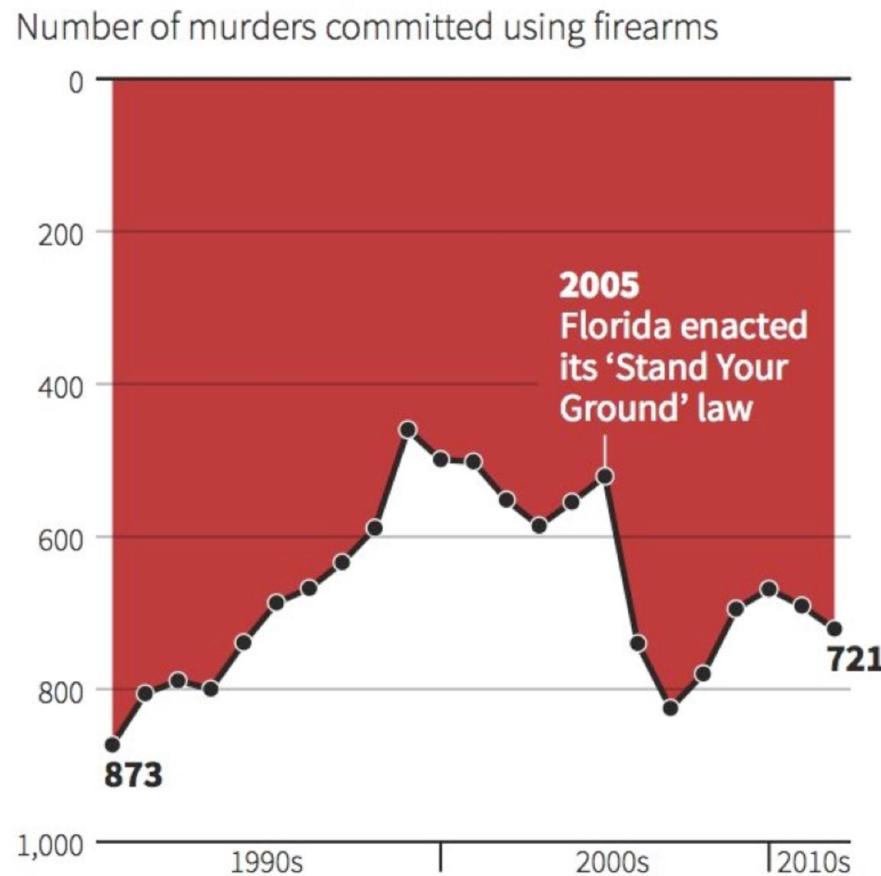
Boxes removed and arrows simplified.

Is violent crime at its highest point?



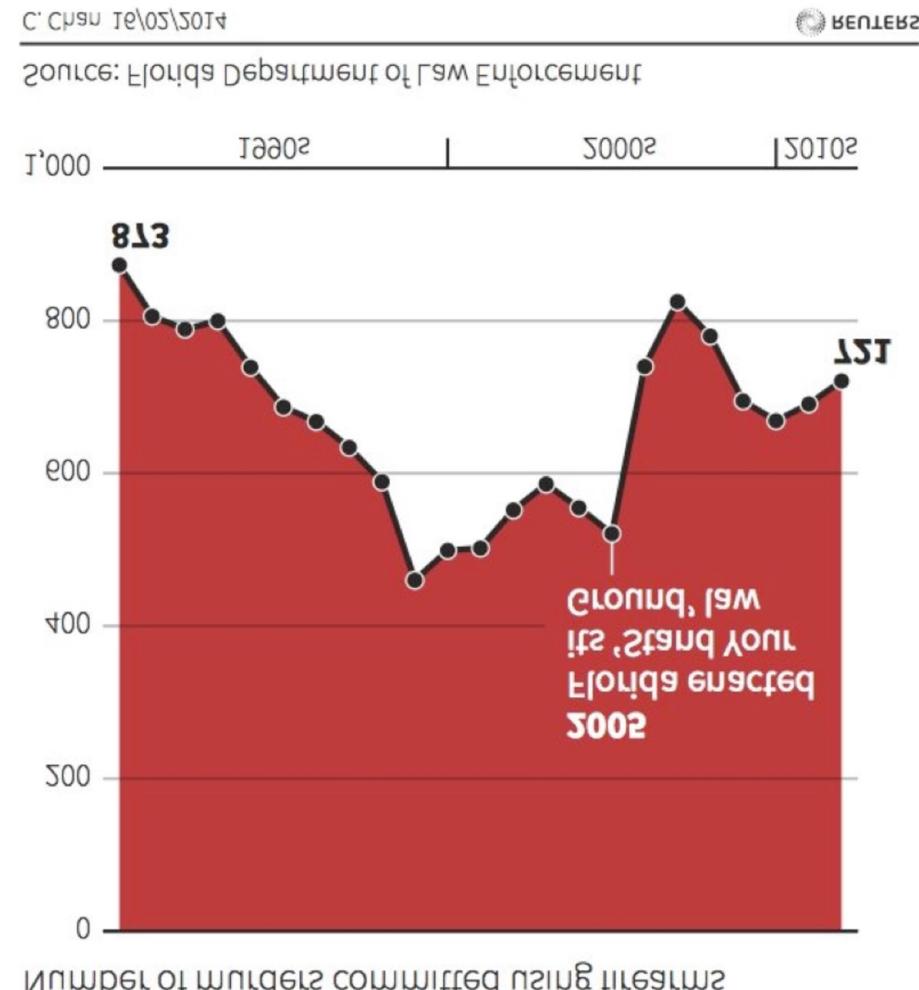
Do gun deaths increase or decrease?

Gun deaths in Florida



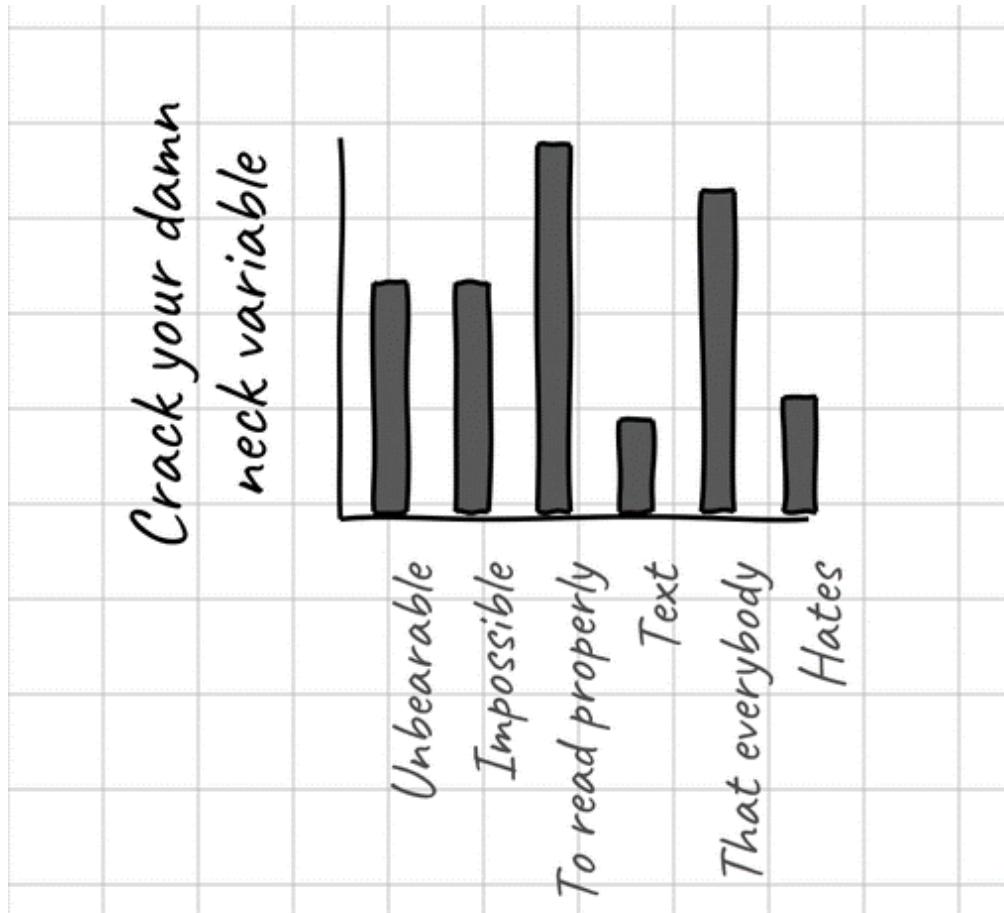
Source: Florida Department of Law Enforcement

Do gun deaths increase or decrease?

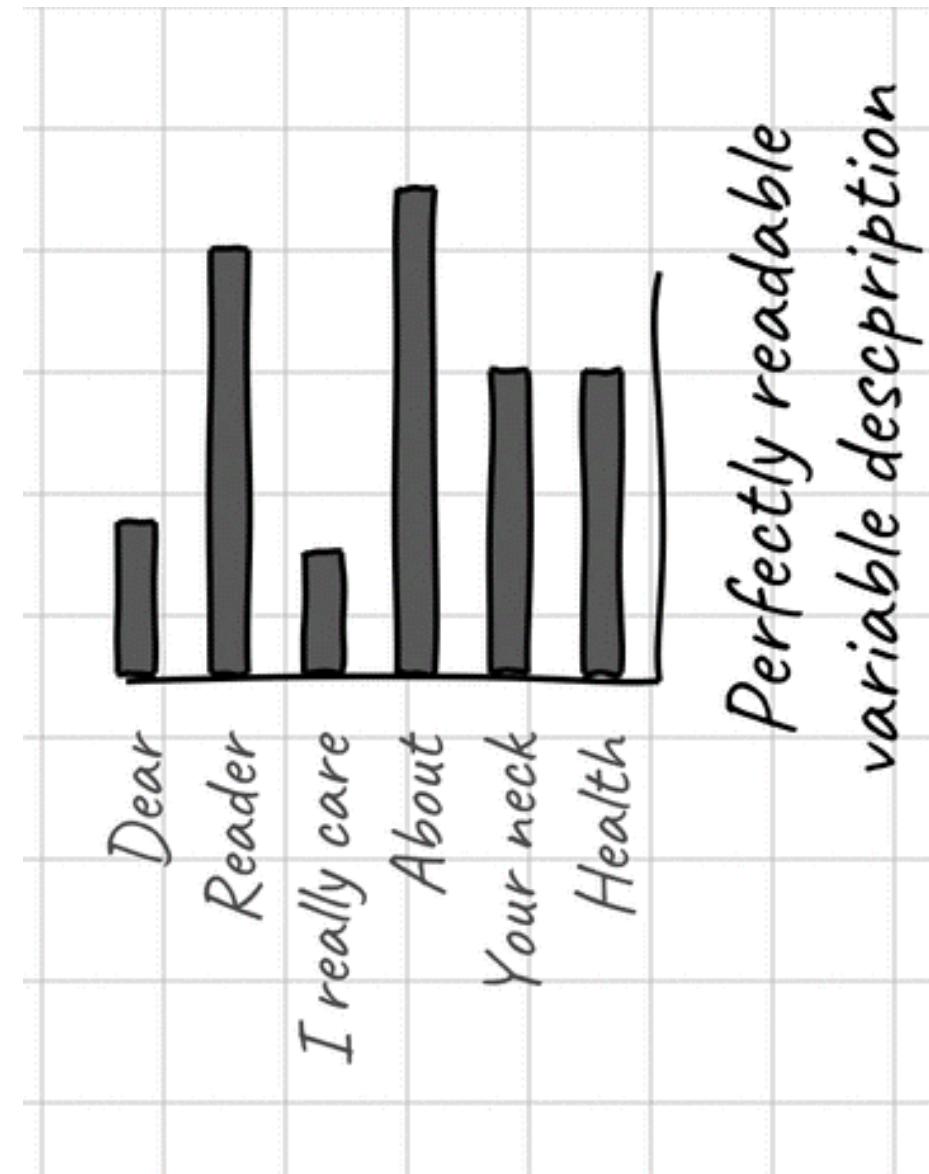
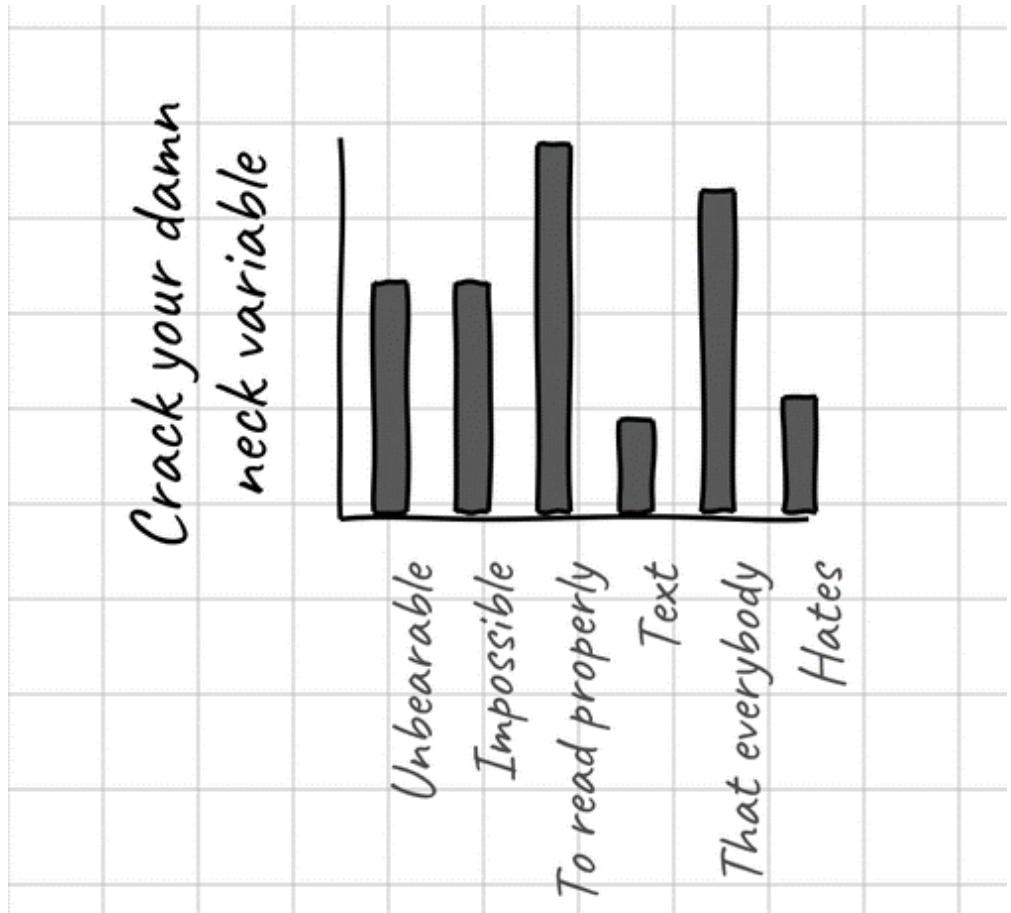


Our goals in Florida

How do you feel about reading this?



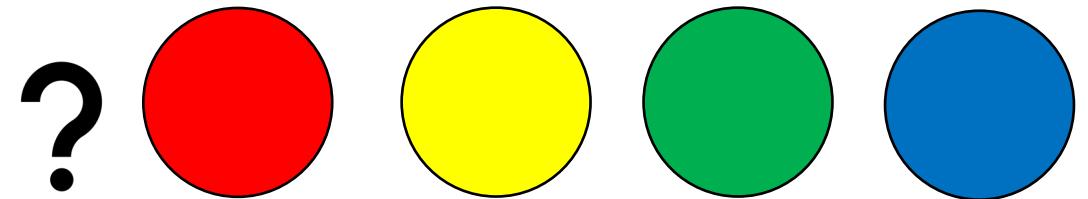
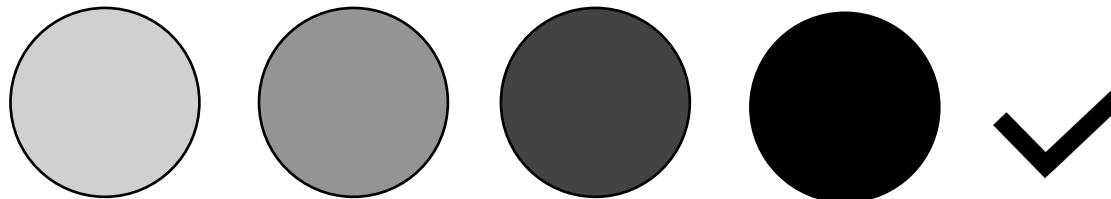
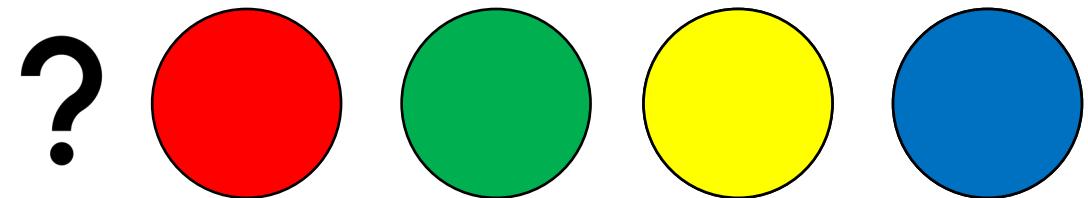
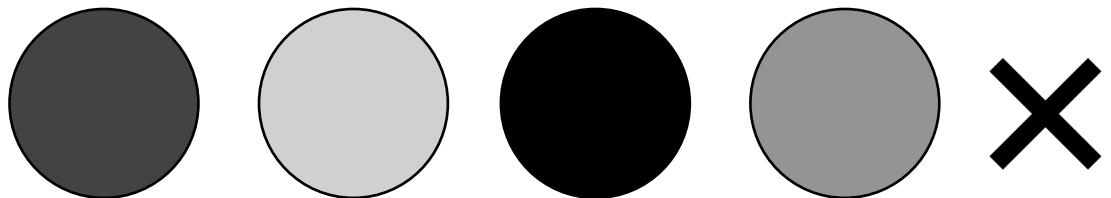
How do you feel about reading this?



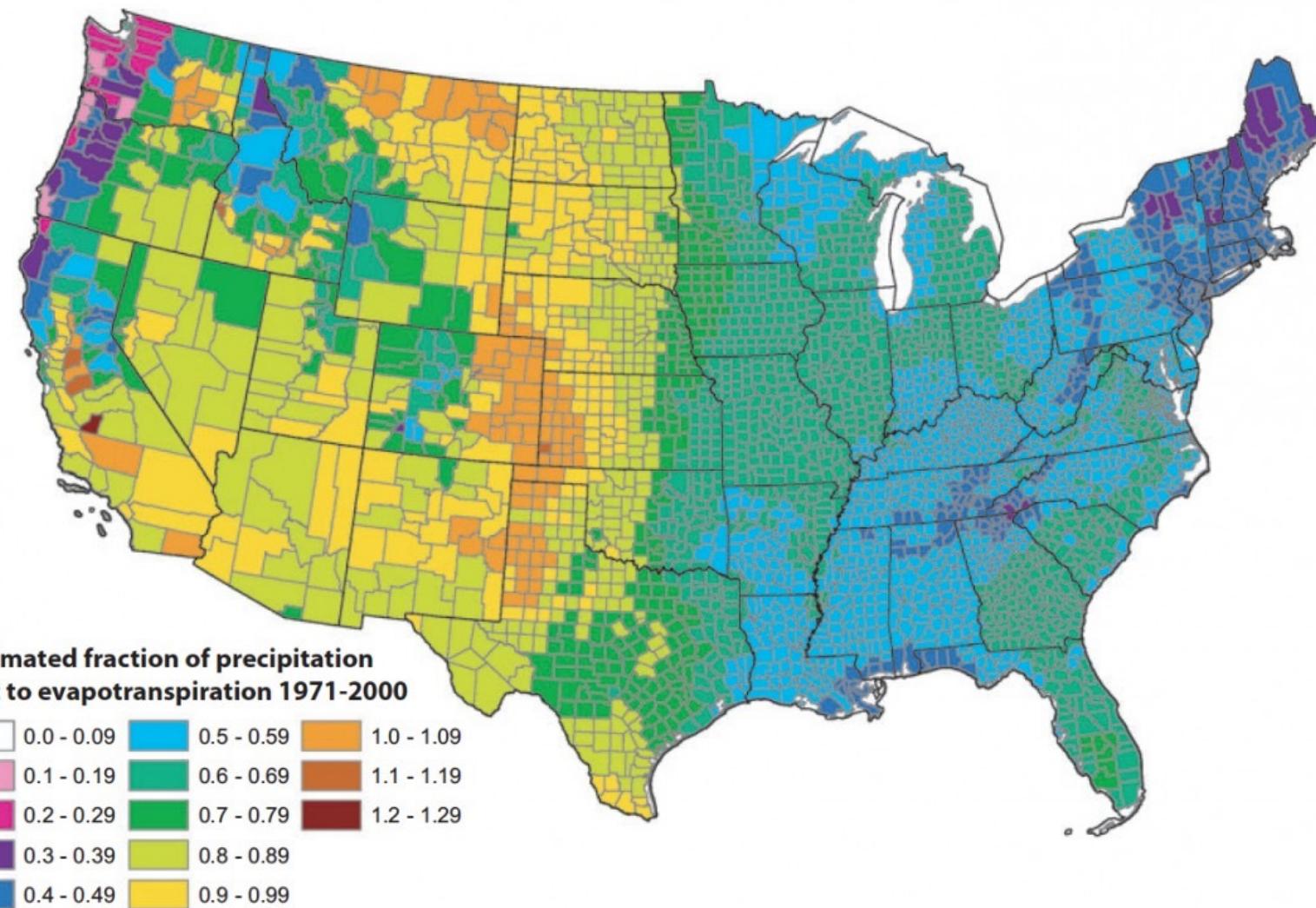
How to rank those circles from low to high?



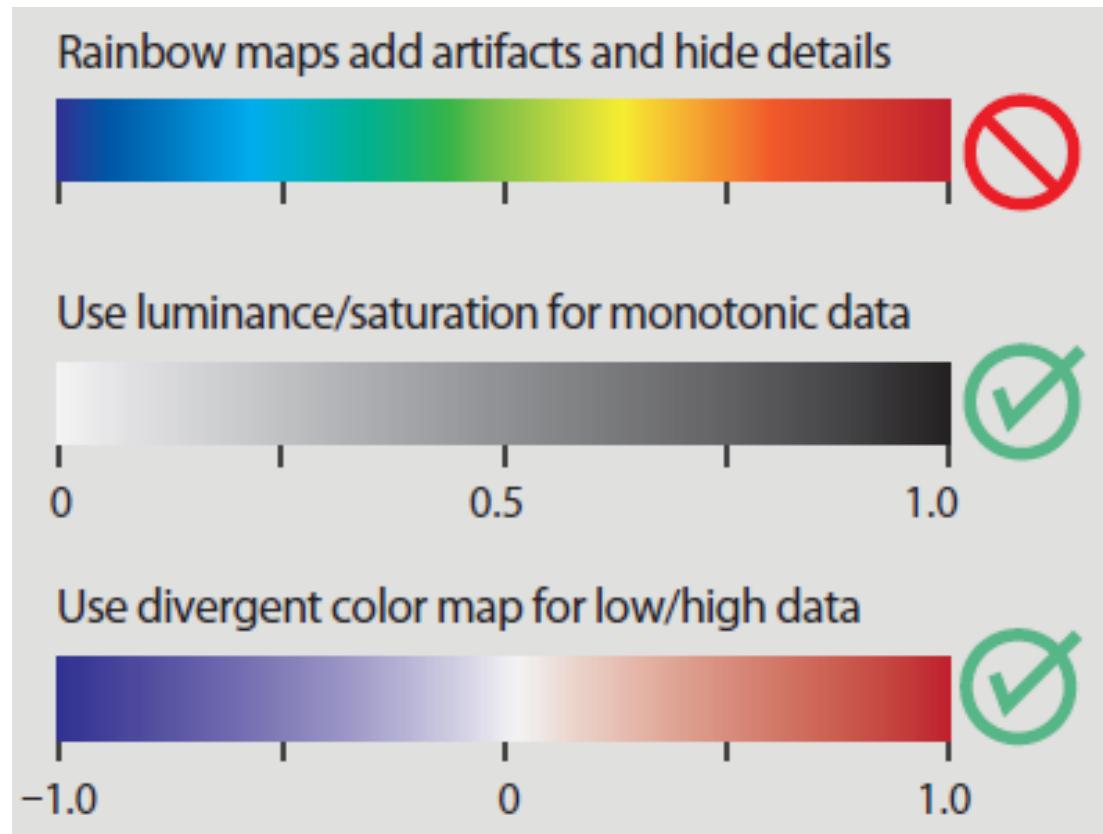
How to rank those circles from low to high?



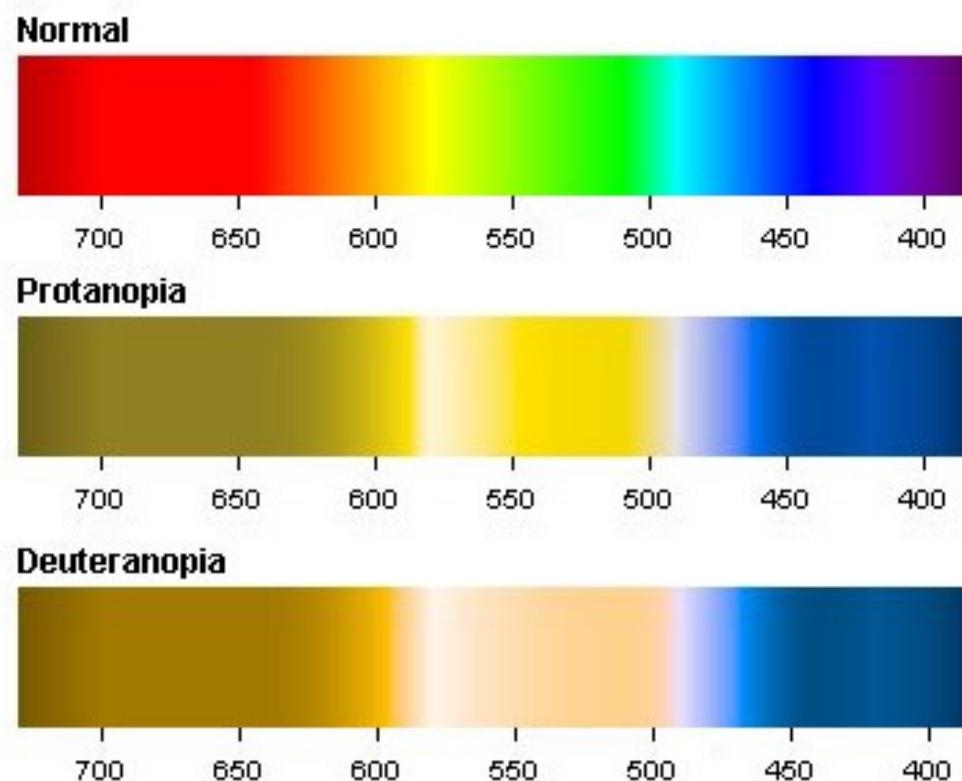
Avoid rainbow color maps for quantities!



Use a monochromatic or a diverging color scale instead



Another issue with color: color vision deficiency



~2% male population

~6% male population

Use a color blindness simulator or colorblindness-friendly color scales instead



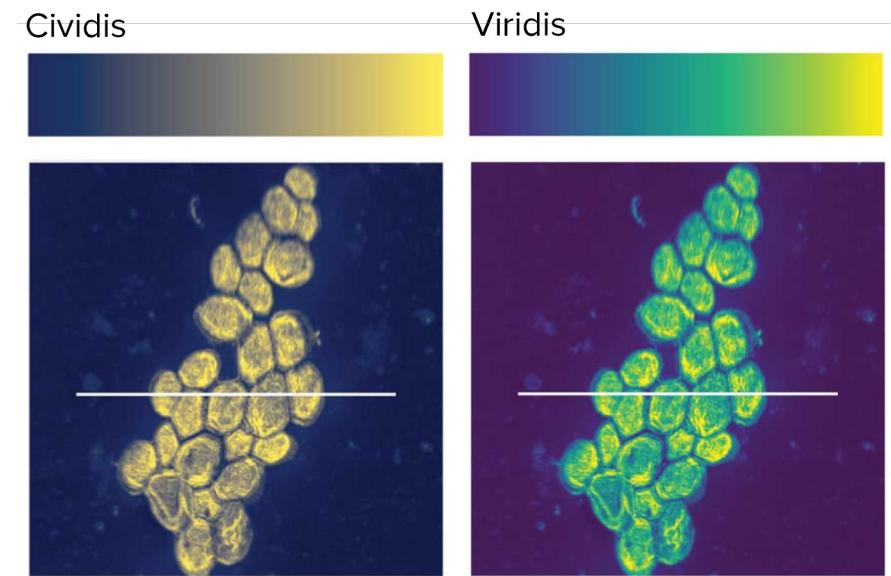
Color Oracle

Design for the Color Impaired

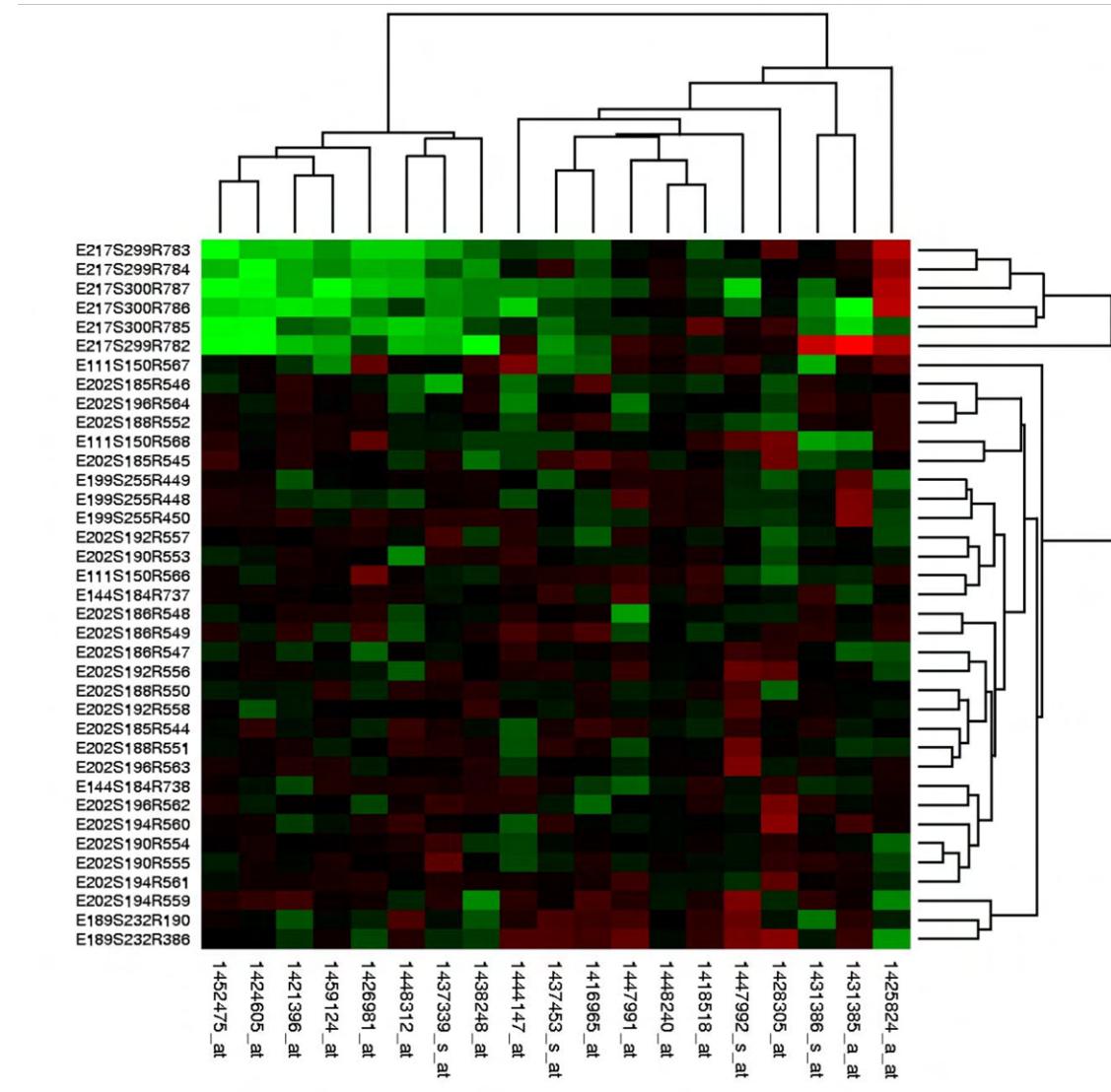


Color Oracle is a free color blindness simulator for Windows, Mac and Linux. It takes the guesswork out of designing for color blindness by showing you in real time what people with common color vision impairments will see.

Color Oracle applies a full screen color filter to art you are designing, independently of the software in use. Eight percent of all males are affected by color vision impairment – make sure that your graphical work is readable by the widest possible audience.



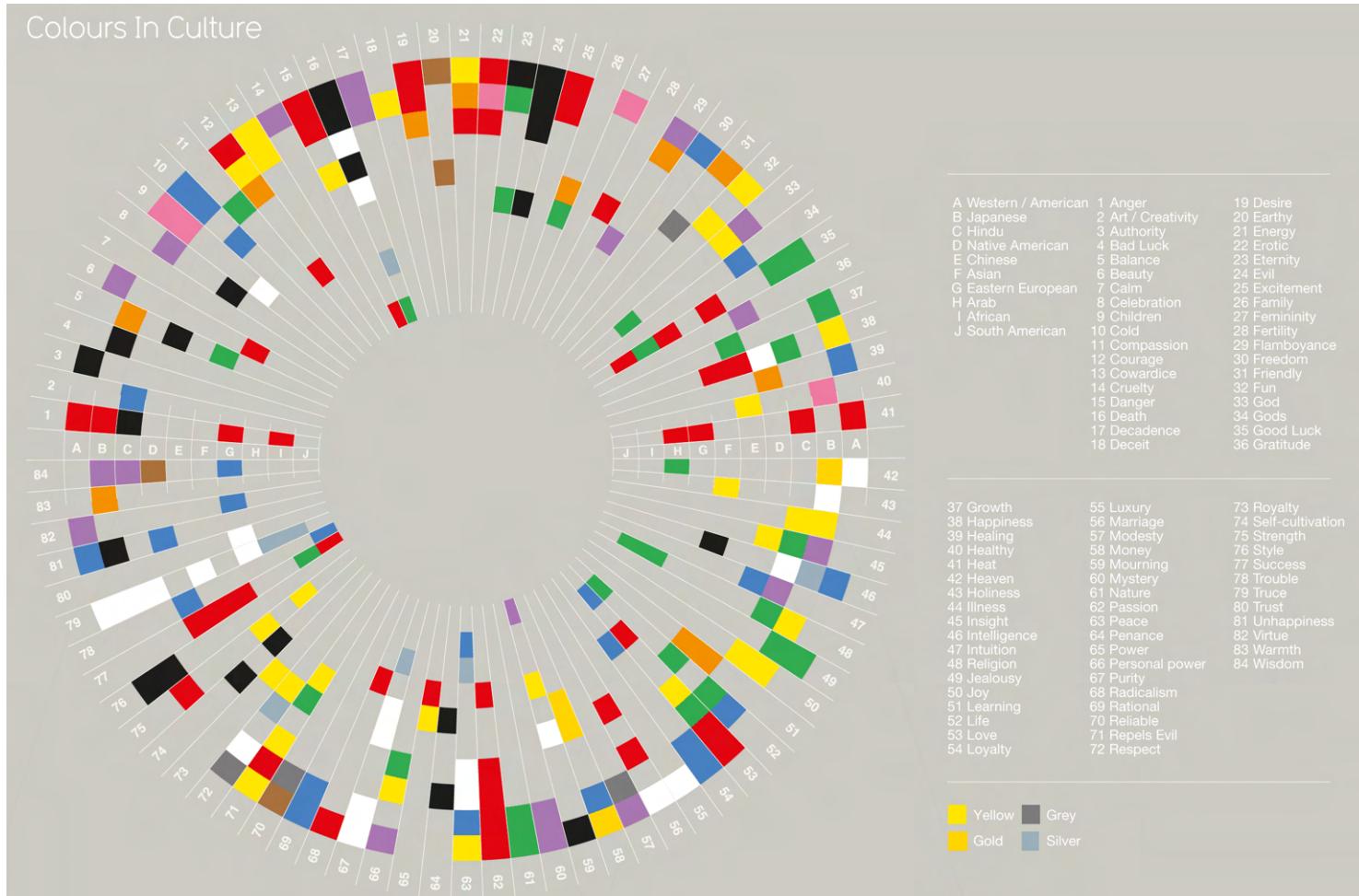
Be aware: contrast can create illusions!



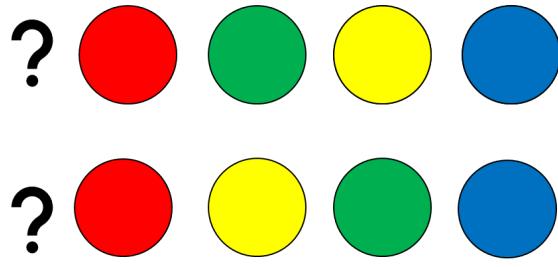
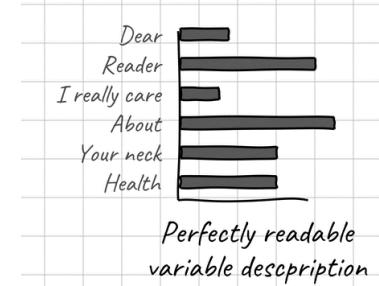
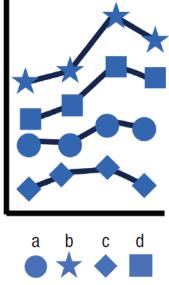
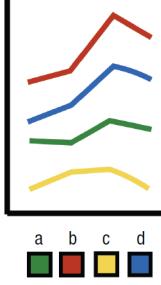
Lastly, color is associated to feelings...



...and differs between cultures



The **weaknesses** of our brain - Summary



1.

Use our short-term
memory

2.

Process too much
information at once

3.

Go against the natural
flow

4.

Correct quantification of
color

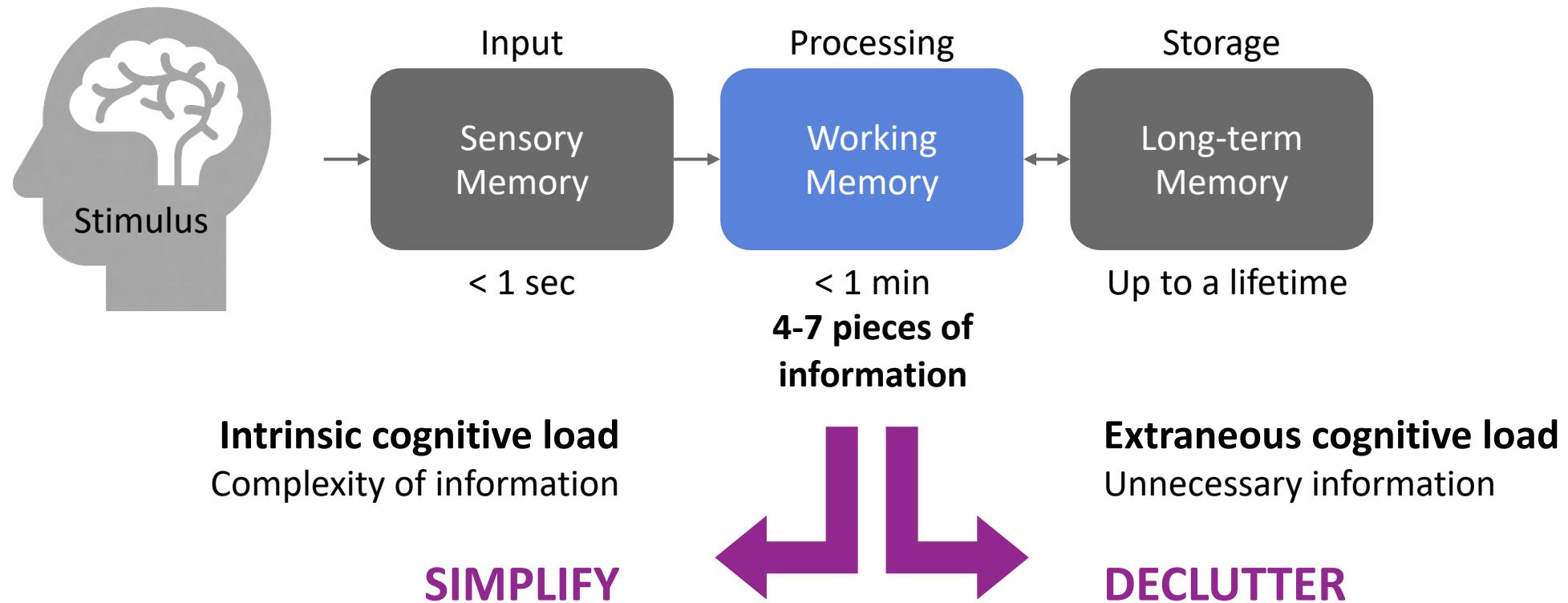


<https://declutterme.london>

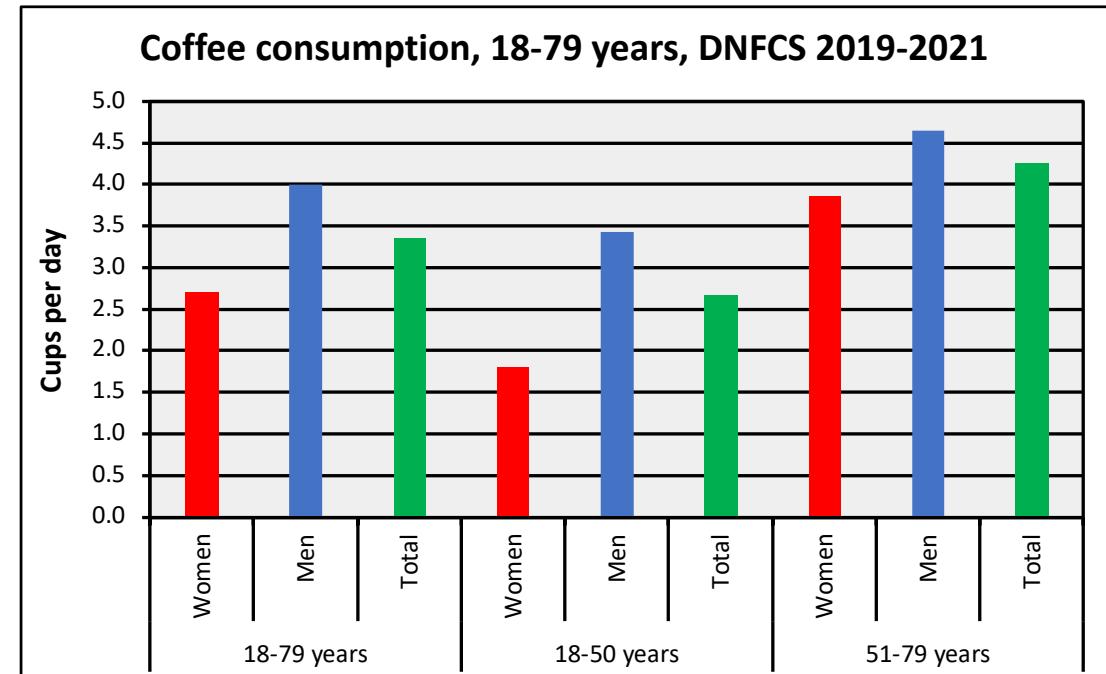
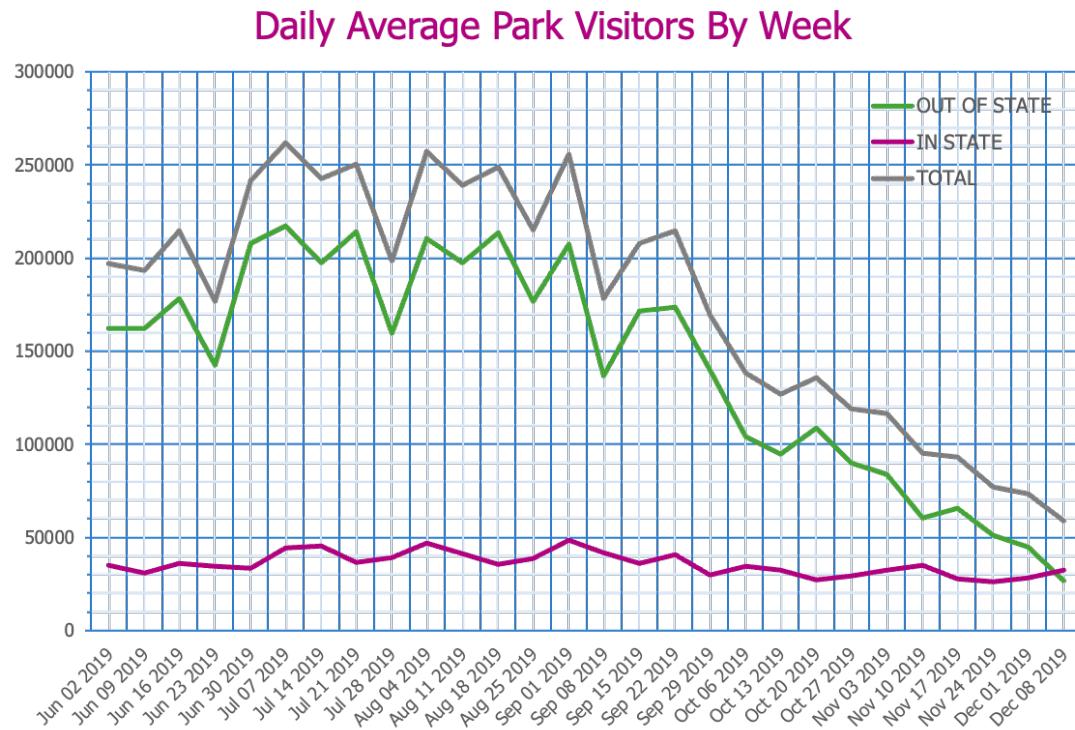
Exercise: Simplify and Declutter



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



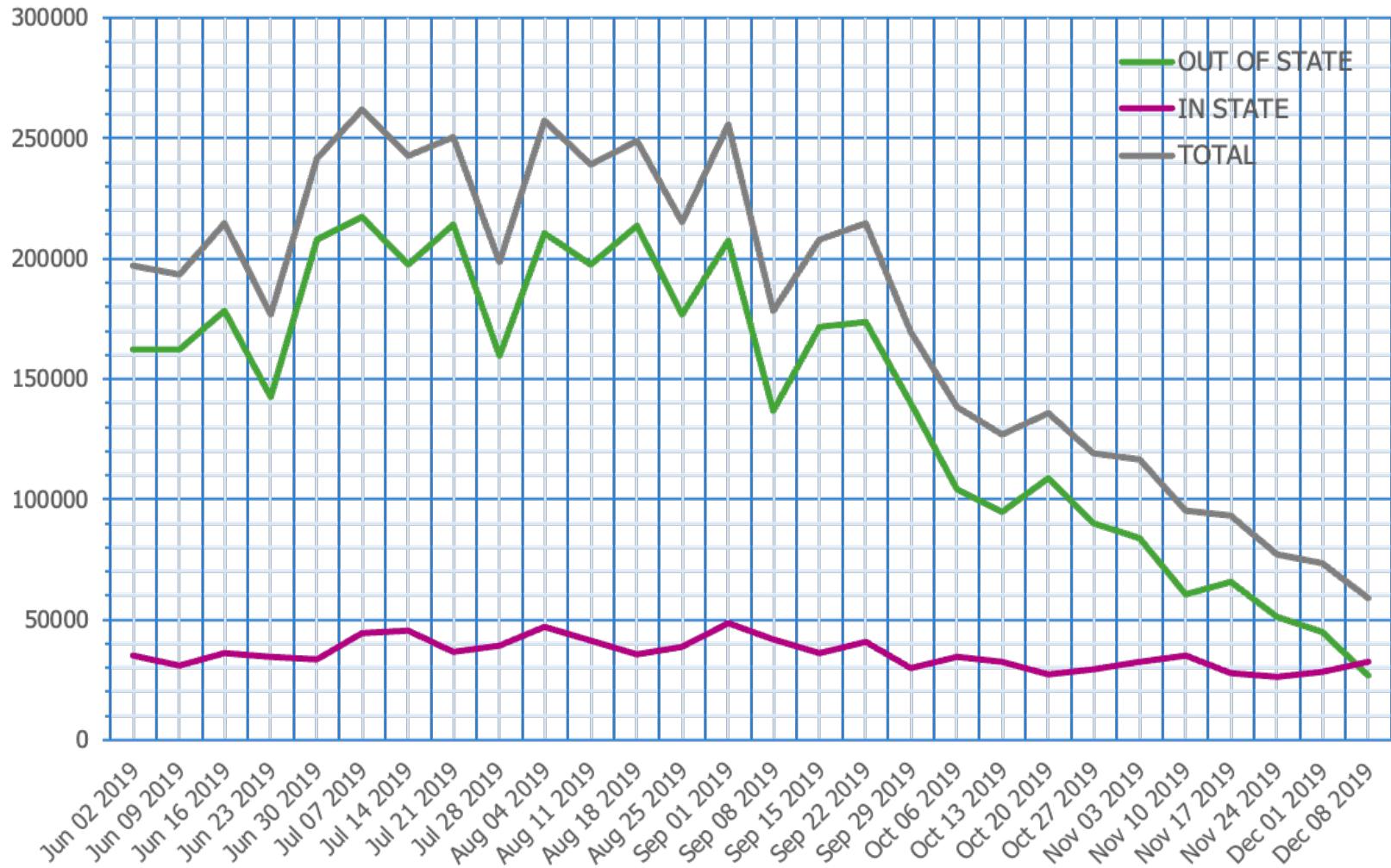
Exercise: Simplify and Declutter



10 min

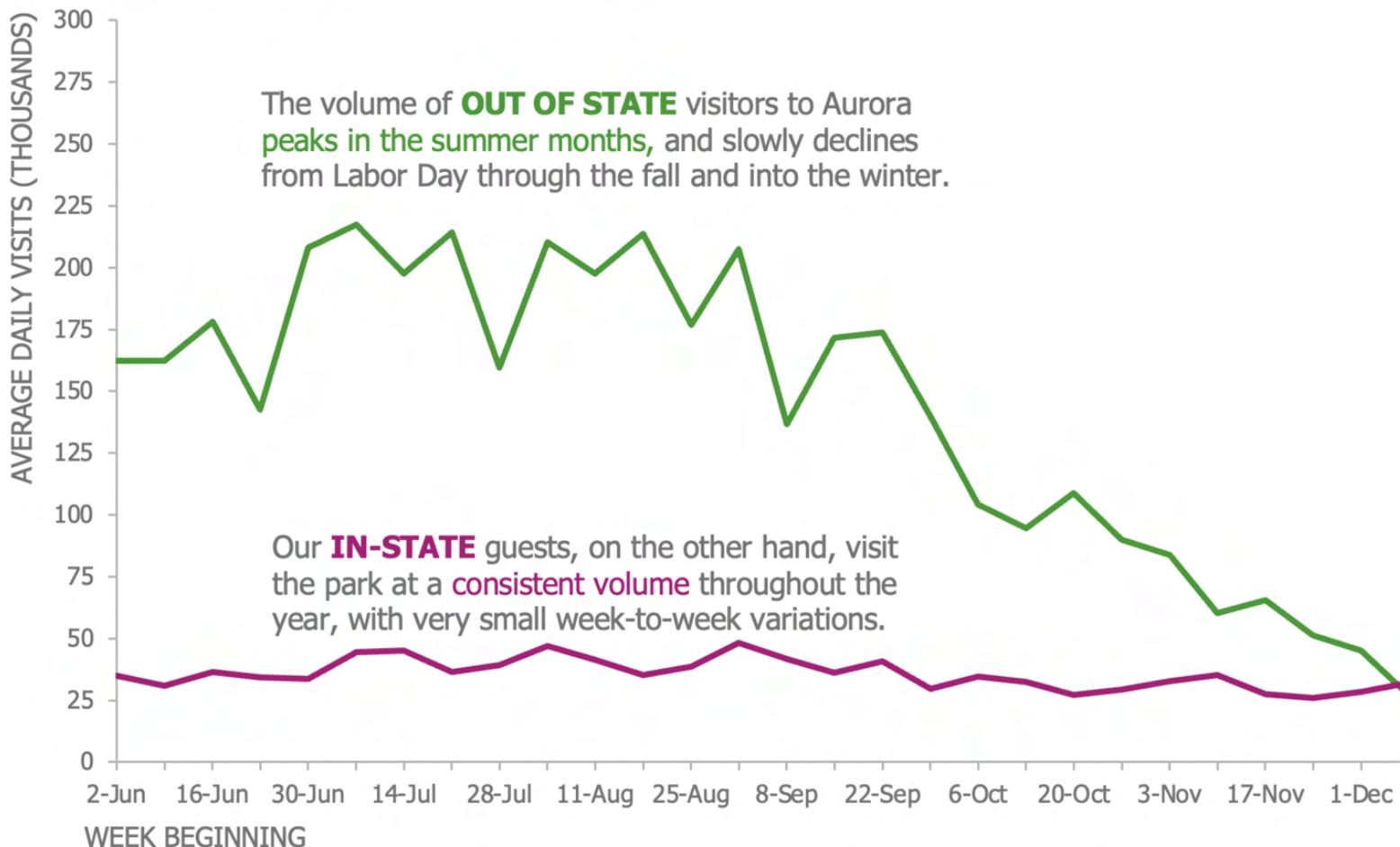


Daily Average Park Visitors By Week



Daily visitors to Aurora Park in summer/fall 2019

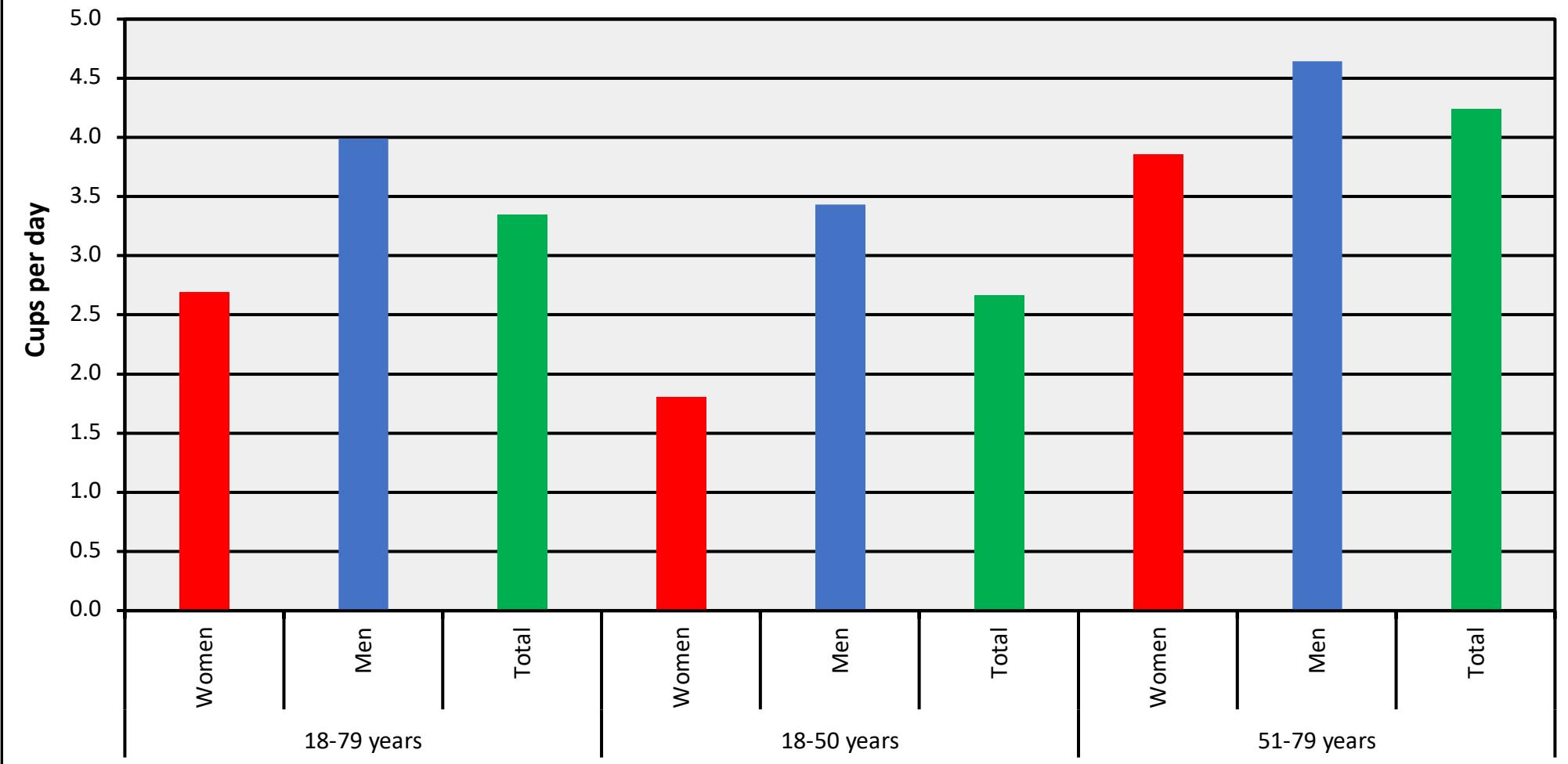
VALUES ARE CALCULATED WEEKLY AS A 7-DAY AVERAGE



The volume of **OUT OF STATE** visitors to Aurora peaks in the summer months, and slowly declines from Labor Day through the fall and into the winter.

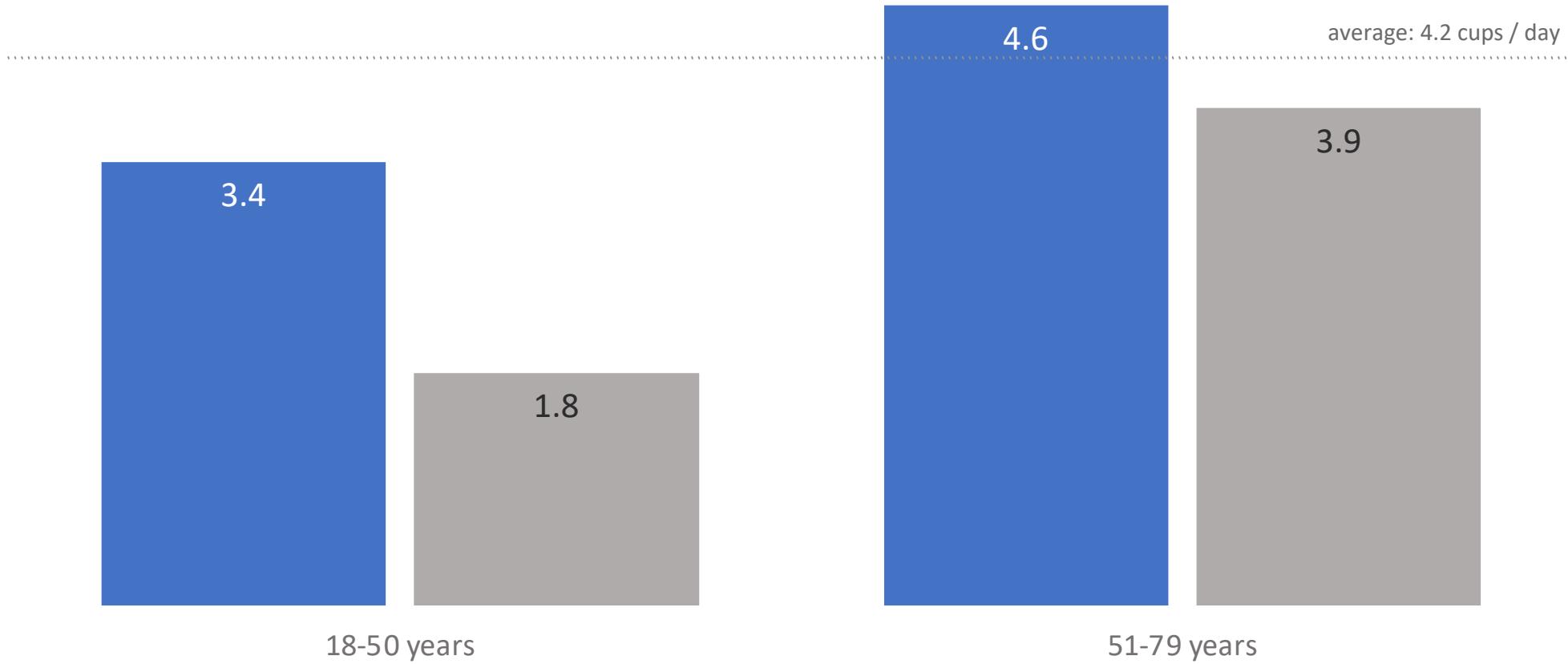
Our **IN-STATE** guests, on the other hand, visit the park at a **consistent volume** throughout the year, with very small week-to-week variations.

Coffee consumption, 18-79 years, DNFCS 2019-2021



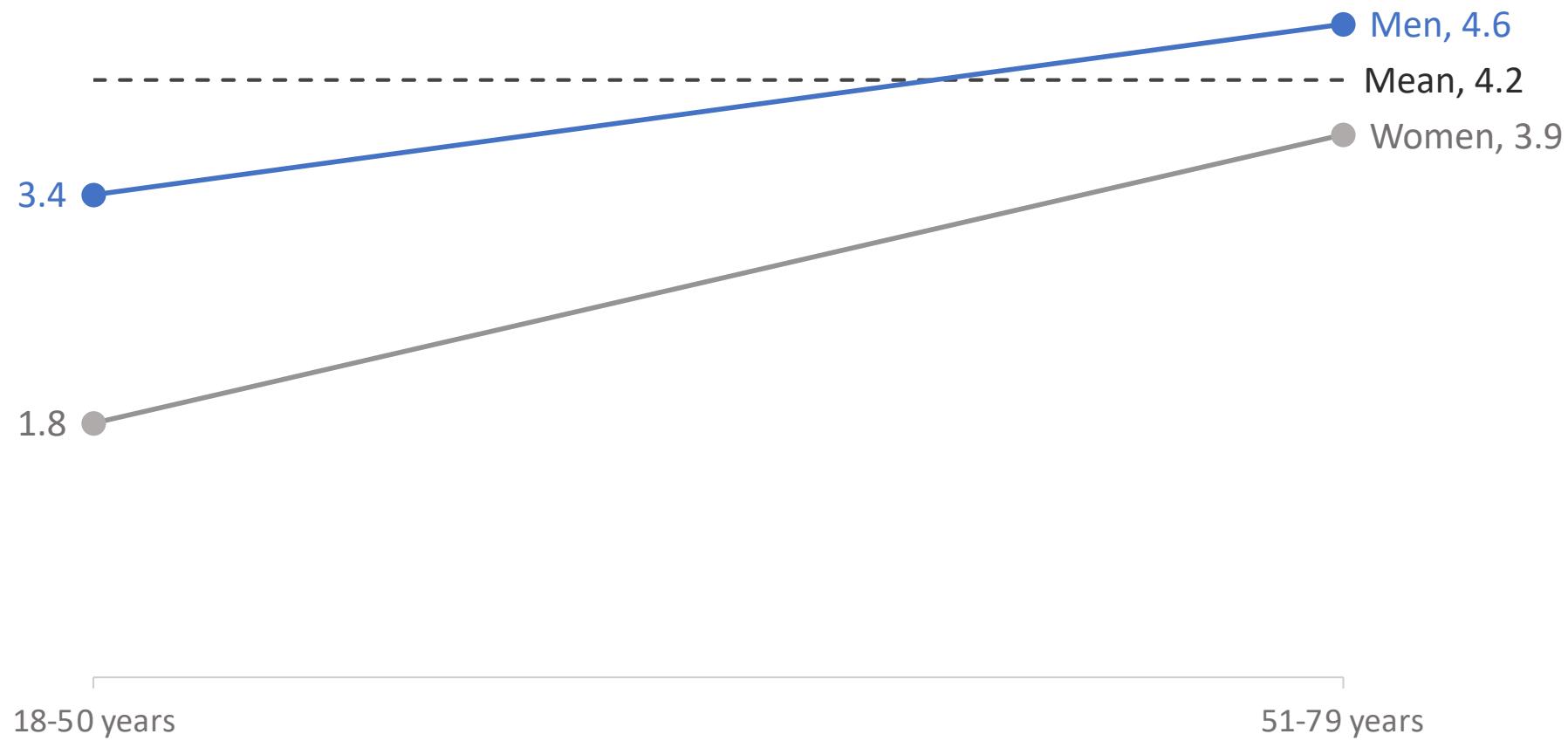
Men drink more coffee than women

No. of cups per day increases with age (Dutch National Food Survey 2019-2021)



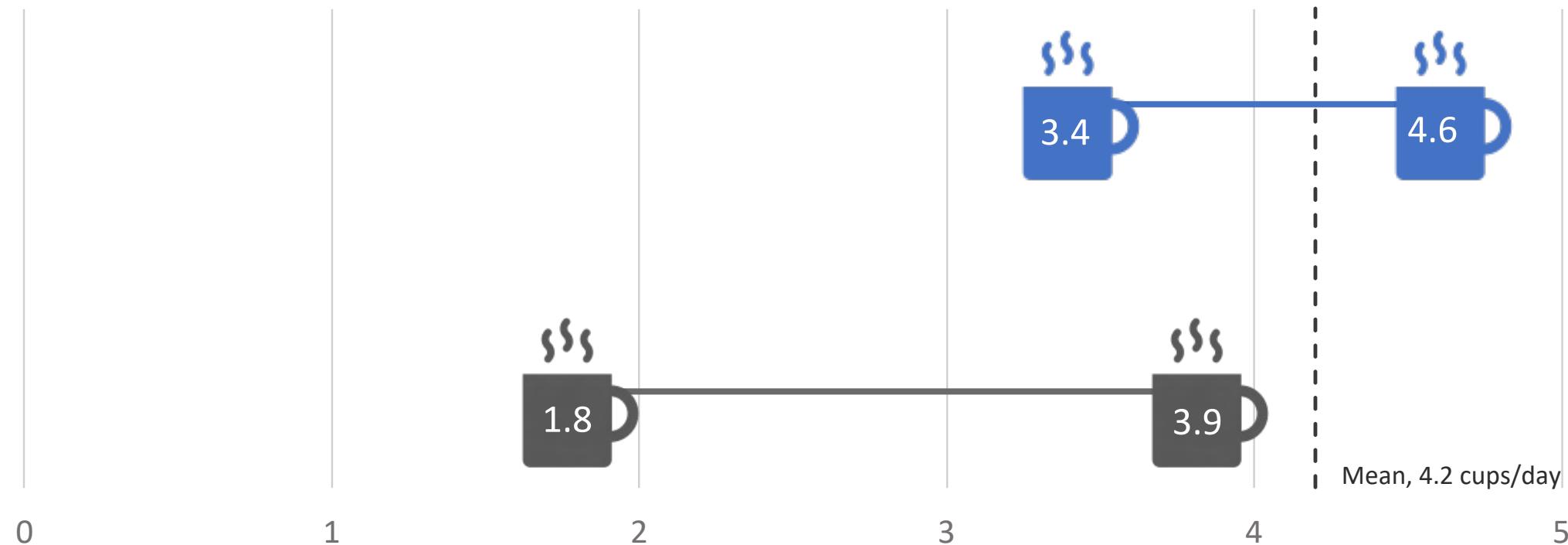
Men drink more coffee than women

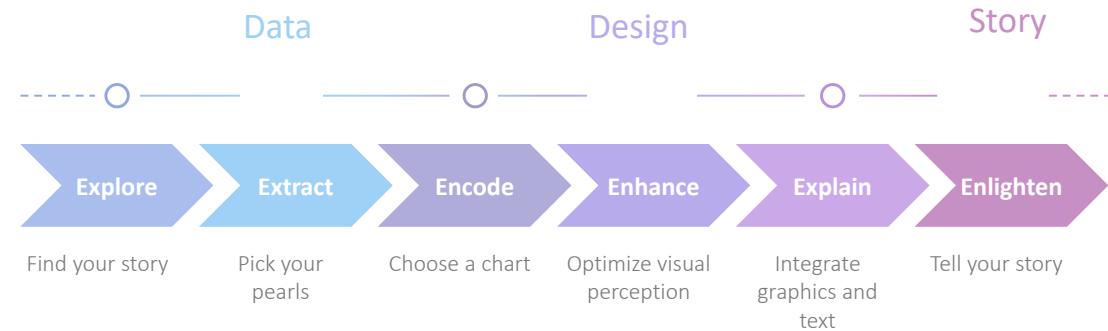
No. of cups per day increases with age (Dutch National Food Survey 2019-2021)



Men drink more coffee than women

No. of cups per day increases with age (Dutch National Food Survey 2019-2021)

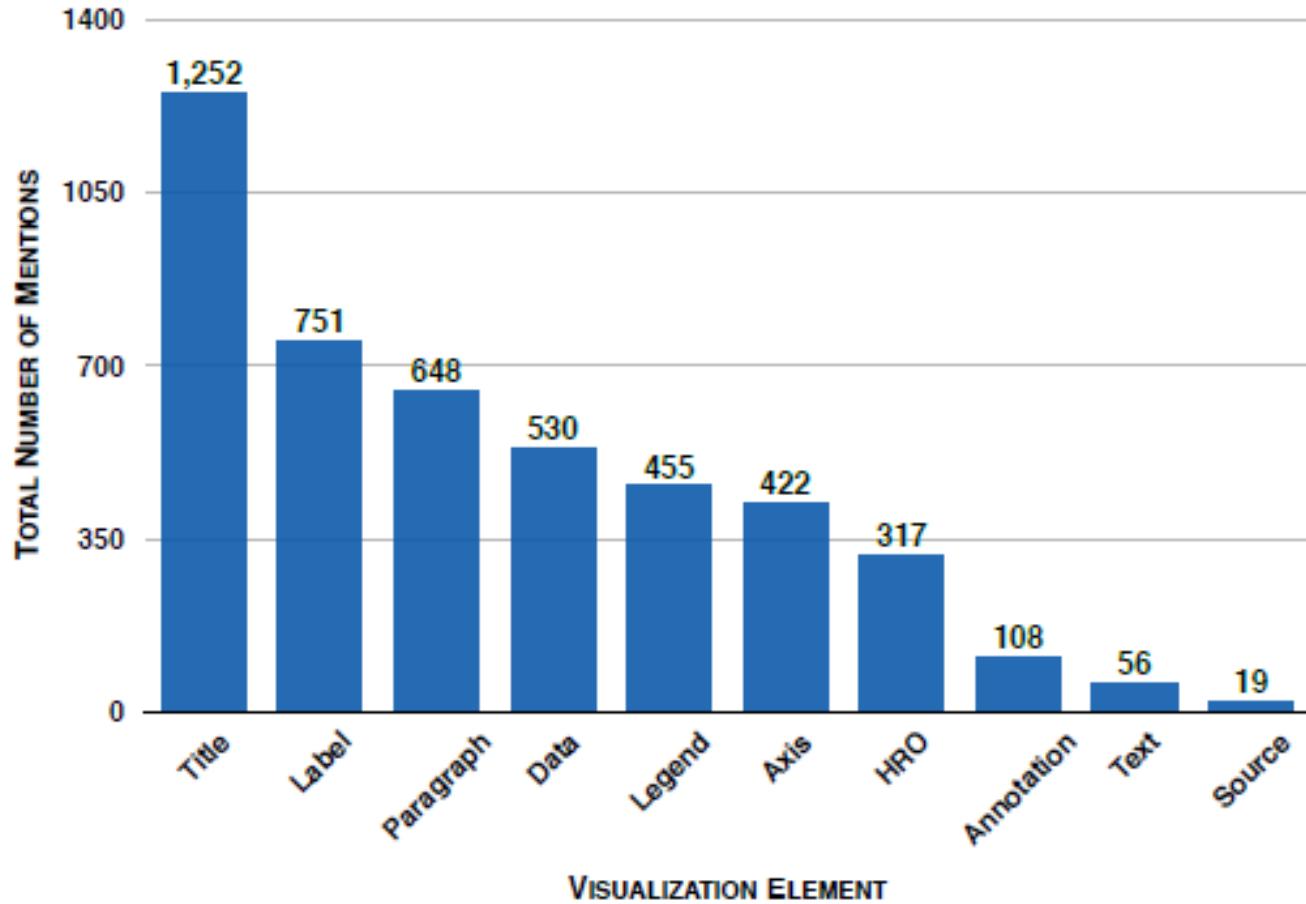




Explain: integrate graphics and text

If you can't explain it simply, you don't understand it well enough

What makes a visual memorable?



Rank:	1st	2nd	3rd
Overall	Title (46%)		
Infographics	Title (72%)	Label (27%)	Paragraph (24%)
News	Paragraph (45%)	Title (43%)	Paragraph (24%)
Government	Title (55%)	Legend (26%)	Label (33%)
Science	Label (27%)	Axis (14%)	Data (21%)
			Legend (13%)

The title is the reader's first impression

Weak title



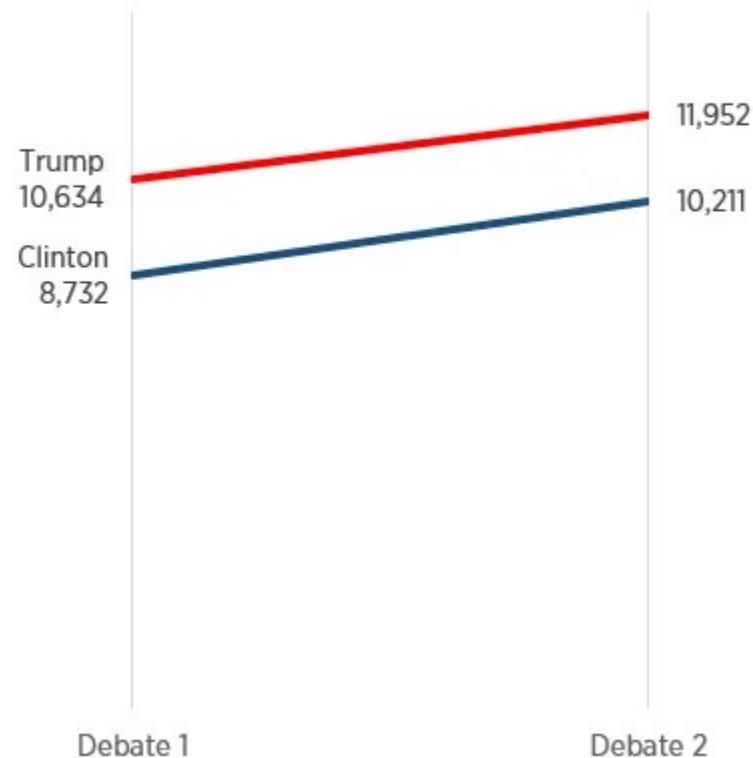
Strong title

Average **pretest** & **posttest** scores show the parenting program made the biggest impact on helping participants handle their own stress.
Scale was 0-5, with 5 representing "Excellent."

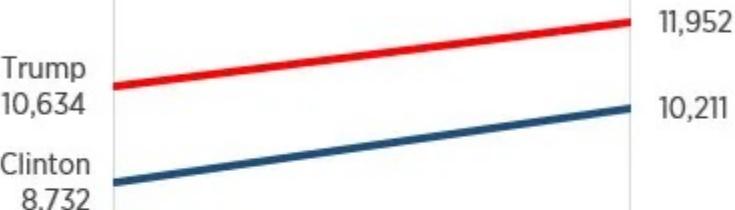


Describe ≠ Judge

Words Spoken at Debate

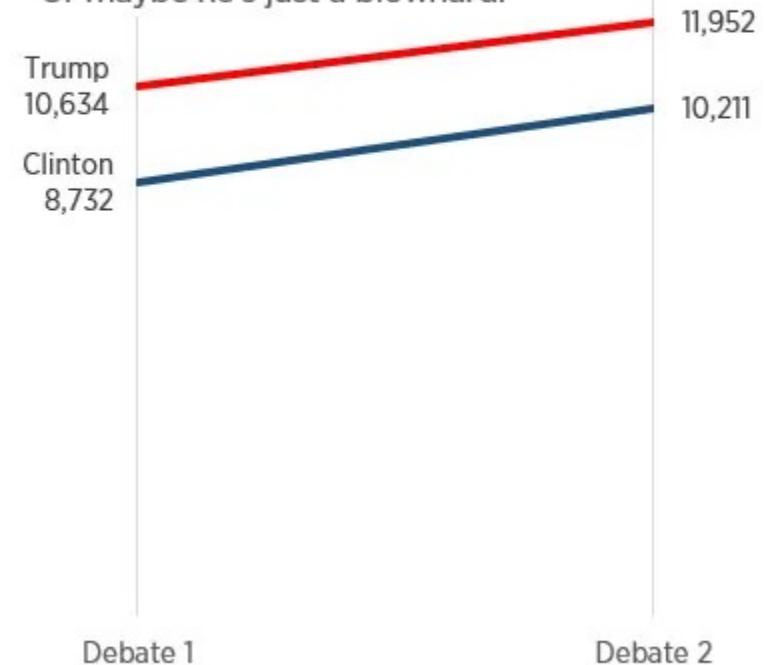


Trump spoke more words during both debates.

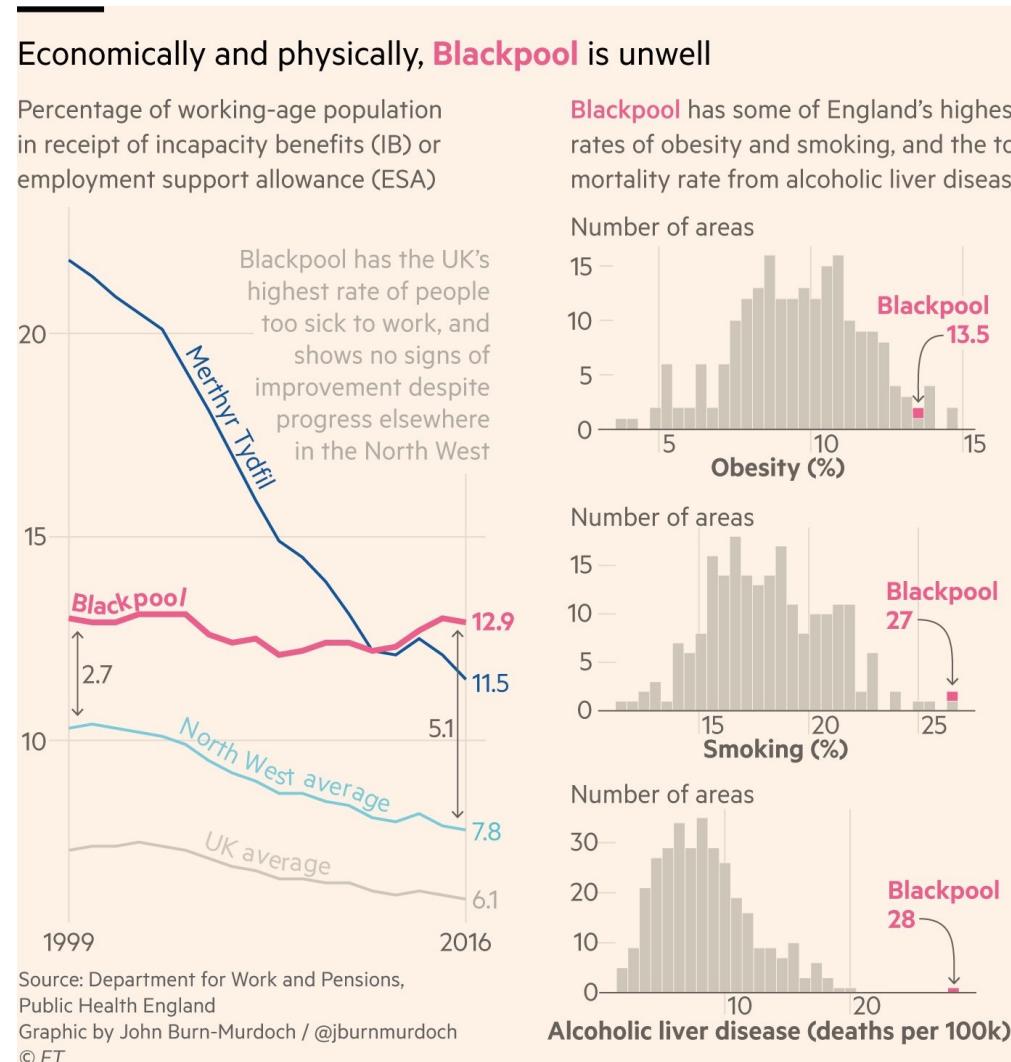


Trump spoke more words during both debates.

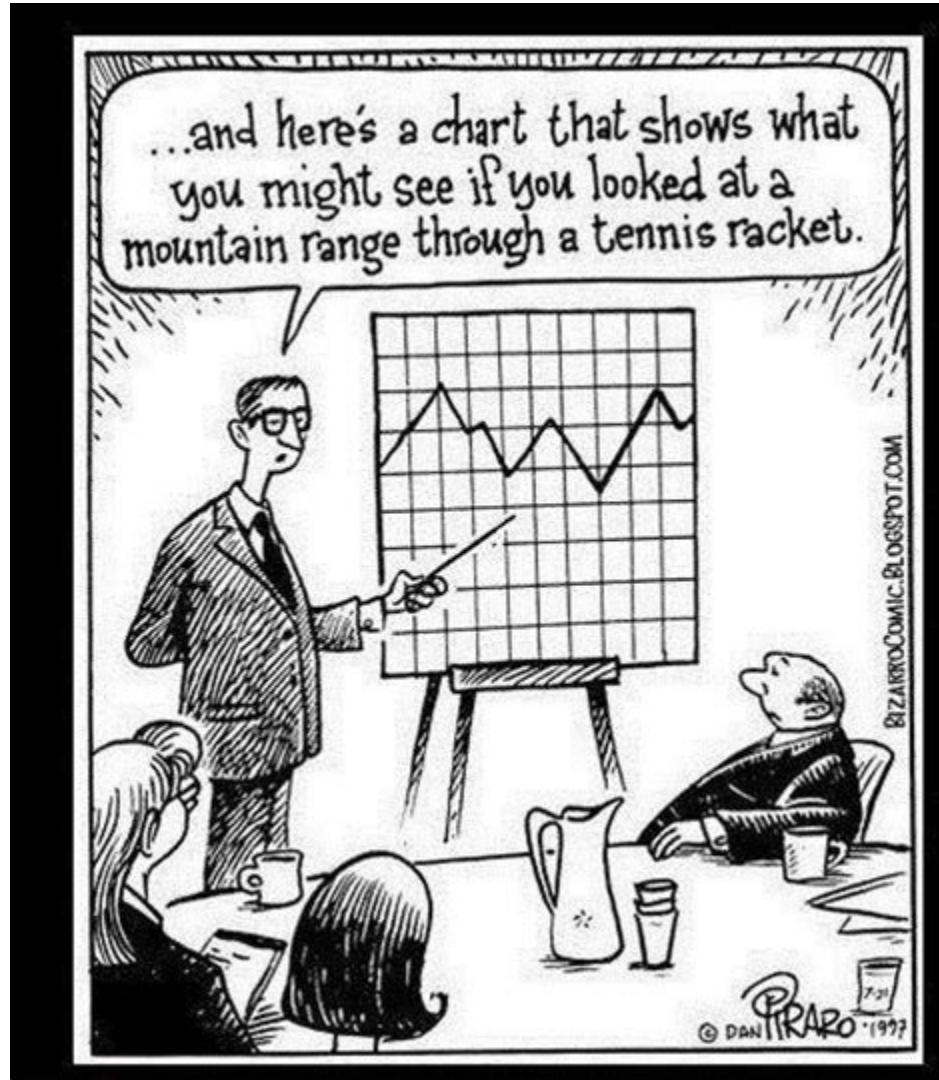
Therefore he deserves to be president.
Or maybe he's just a blowhard.

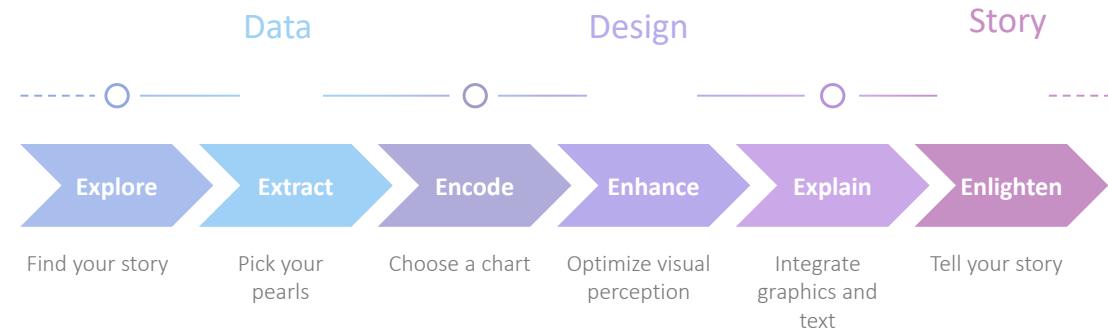


Annotate: describe and clarify what's important



Explain: much easier when presenting

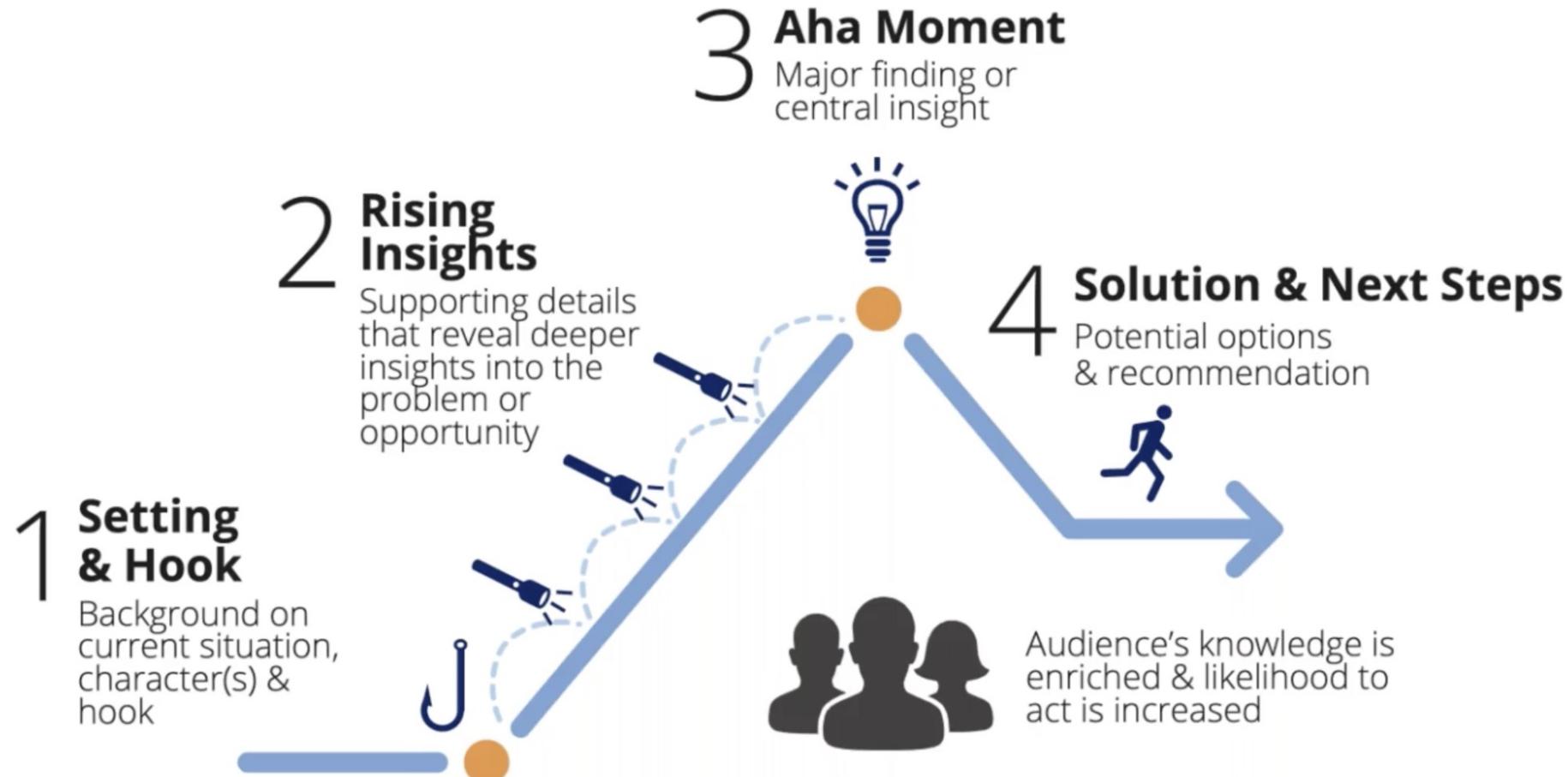




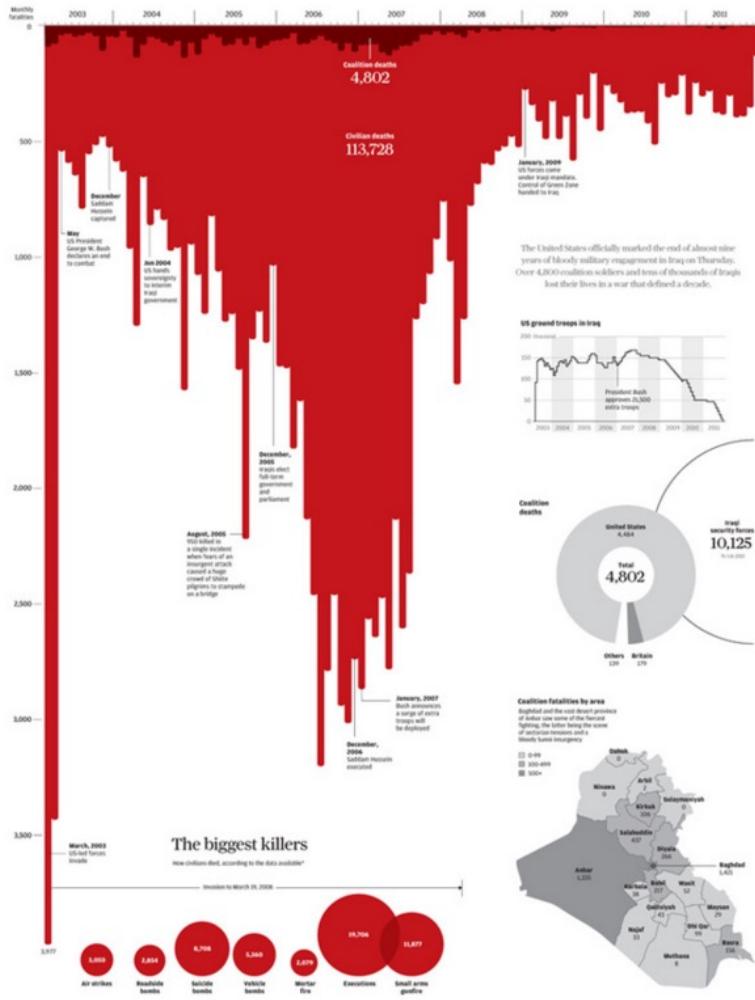
Enlighten: tell your story

Words are how we think, stories are how we link

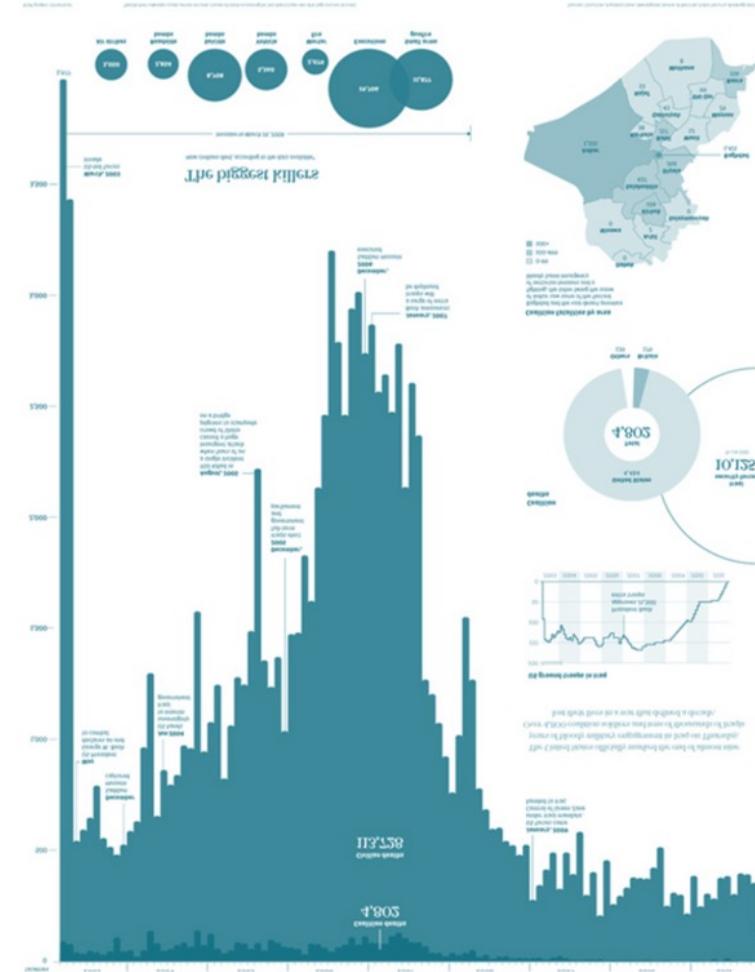
Enlighten: Add the Story to your data



Same data, different story



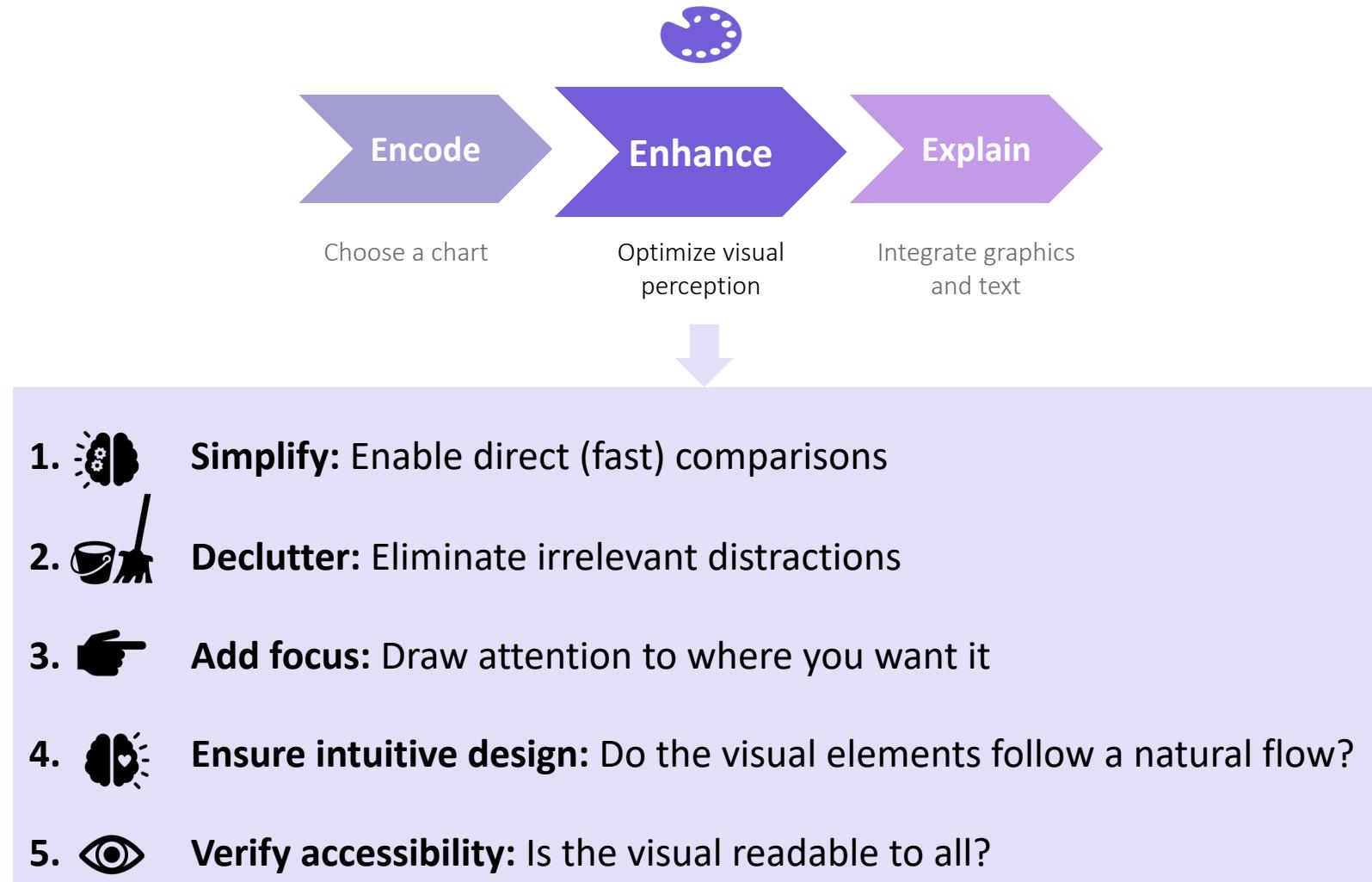
Iraq: Deaths on the decline



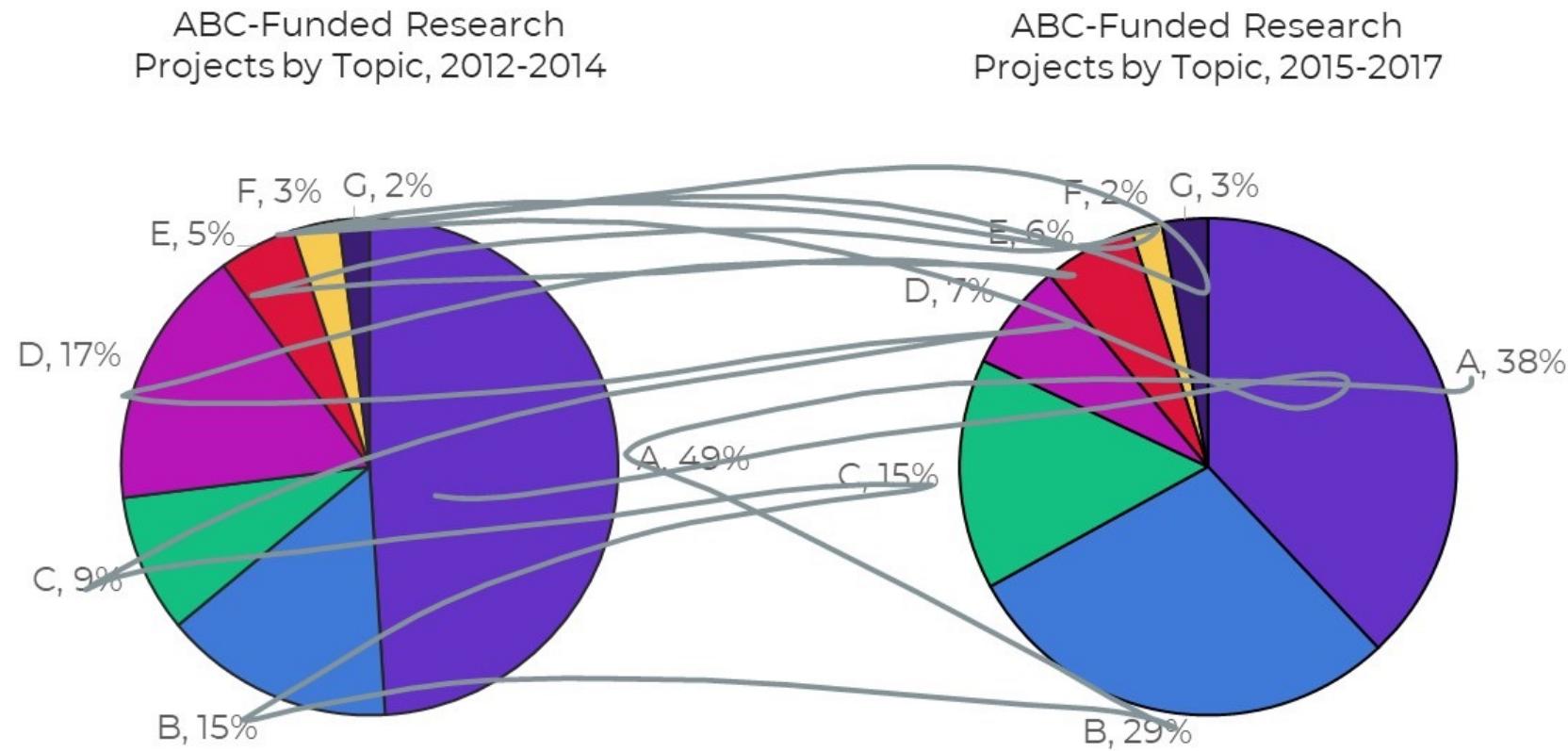
Putting it all together

The 5 key steps in visual design

Pulling it all together: **5 steps** to optimize visual perception



Simplify: Enable direct comparisons



Declutter: Minimize distractions



Effective. Not optimal.



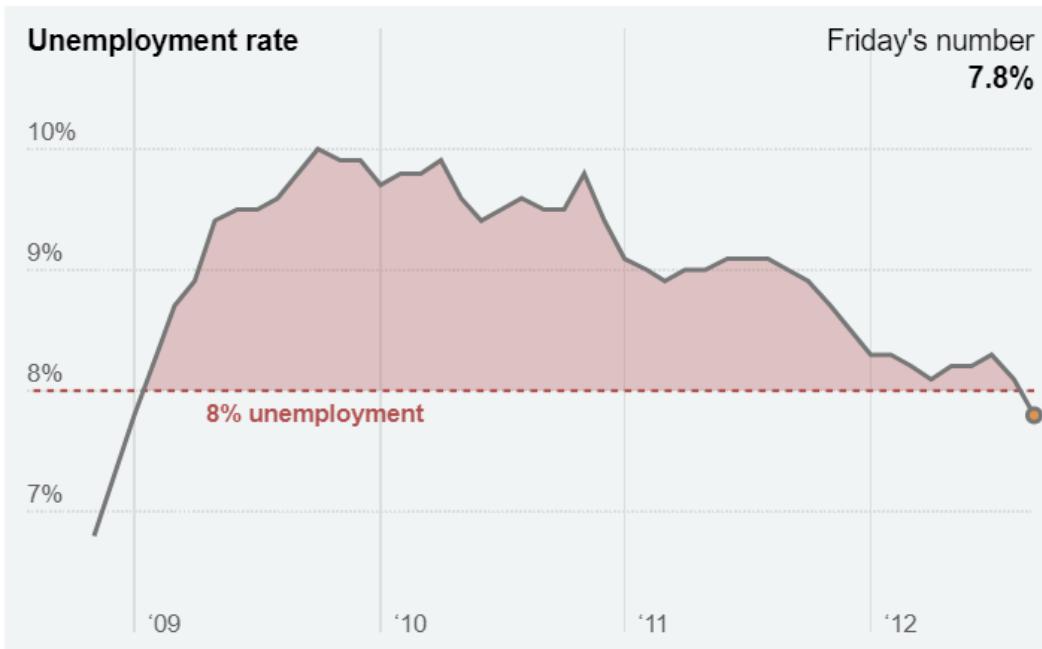
Simple, effective, optimal.



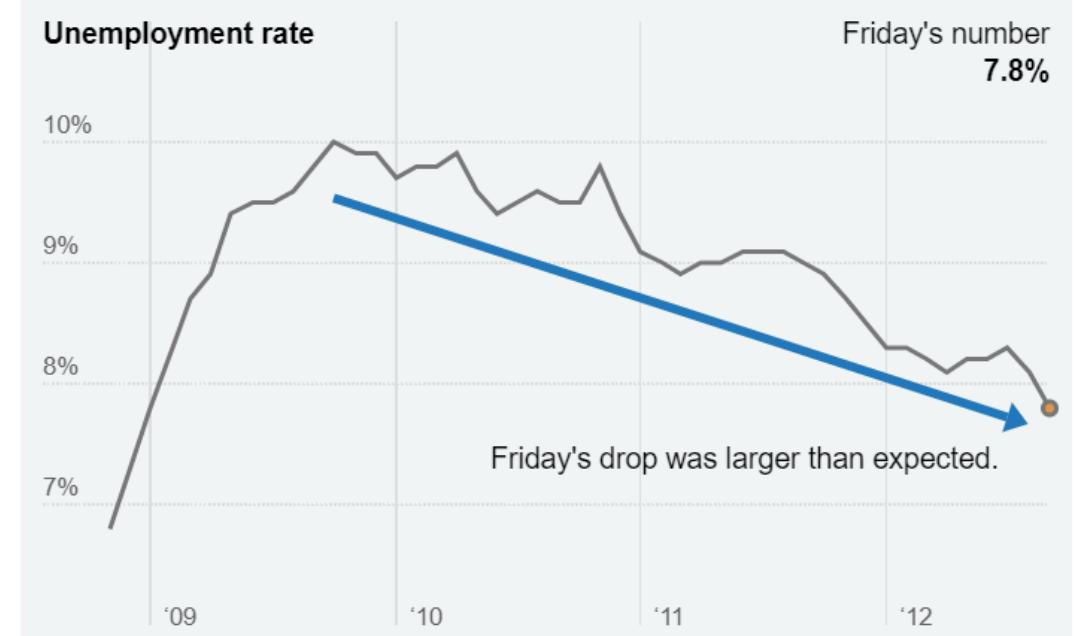
Bad.

Add focus: Highlight what's important

The rate was above 8 percent for 43 months.



The rate has fallen more than 2 points since its recent peak.



How a republican might see things...



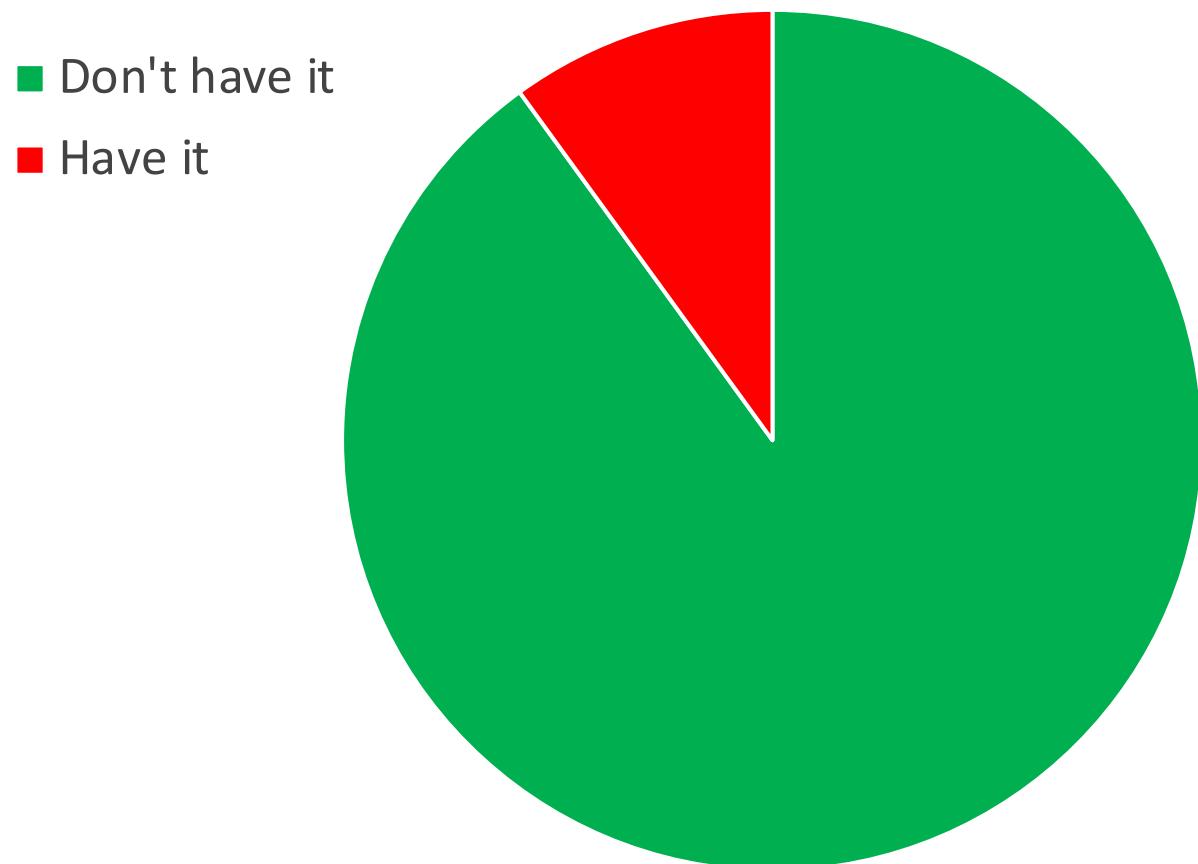
How a democrat might see things...

Intuitive: Follow the reading direction



Accessible: Be mindful of your color choices

Red-green color blindness in men



Exercise

A look at your work.



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
4.  **Ensure intuitive design:** Do the visual elements follow a natural flow?
5.  **Verify accessibility:** Is the visual readable to all?

- Is the most appropriate chart type used to present the data/message?
- What takes up 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?

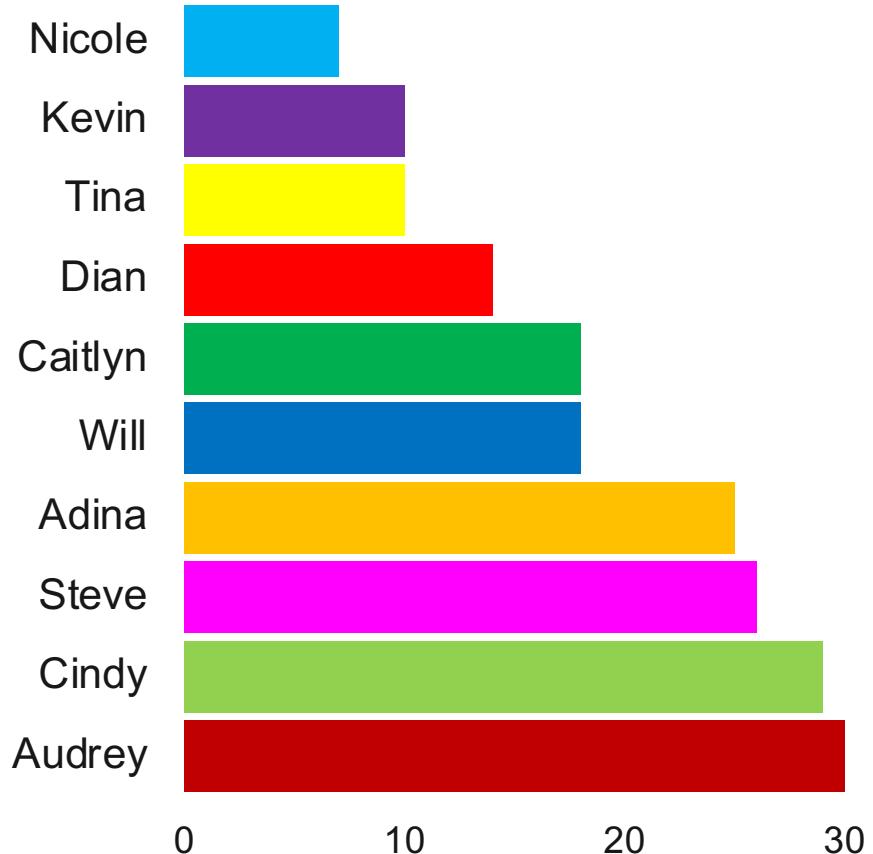


Do's and Don'ts

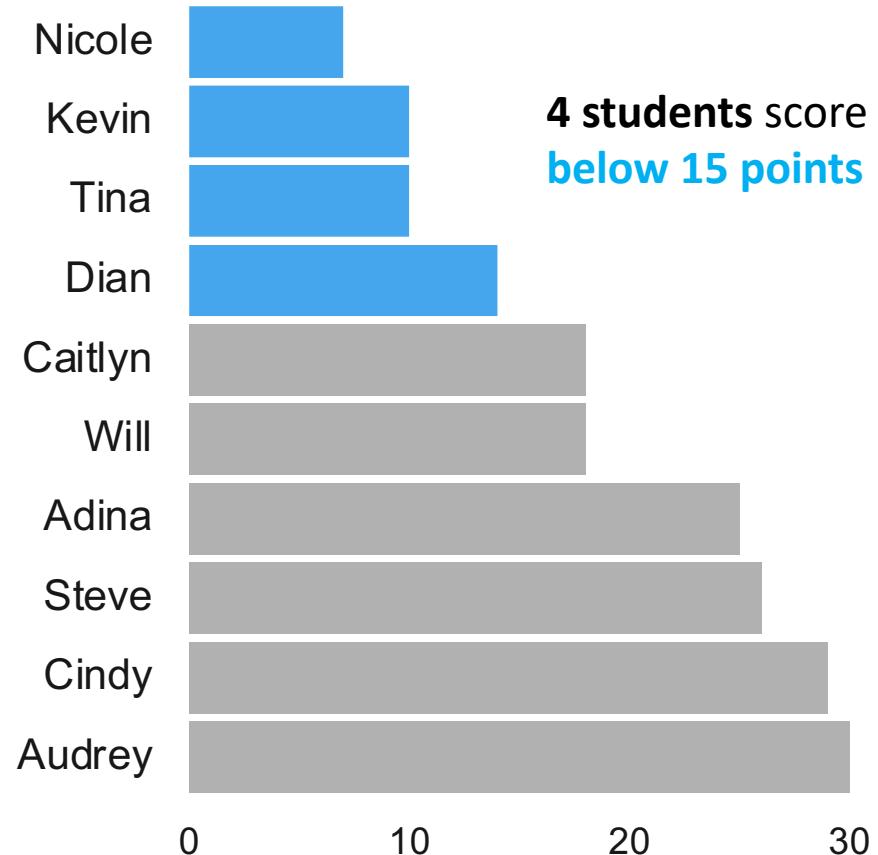
Show the *right* data, show the *data* right!



Don't: Use too many (useless) colors

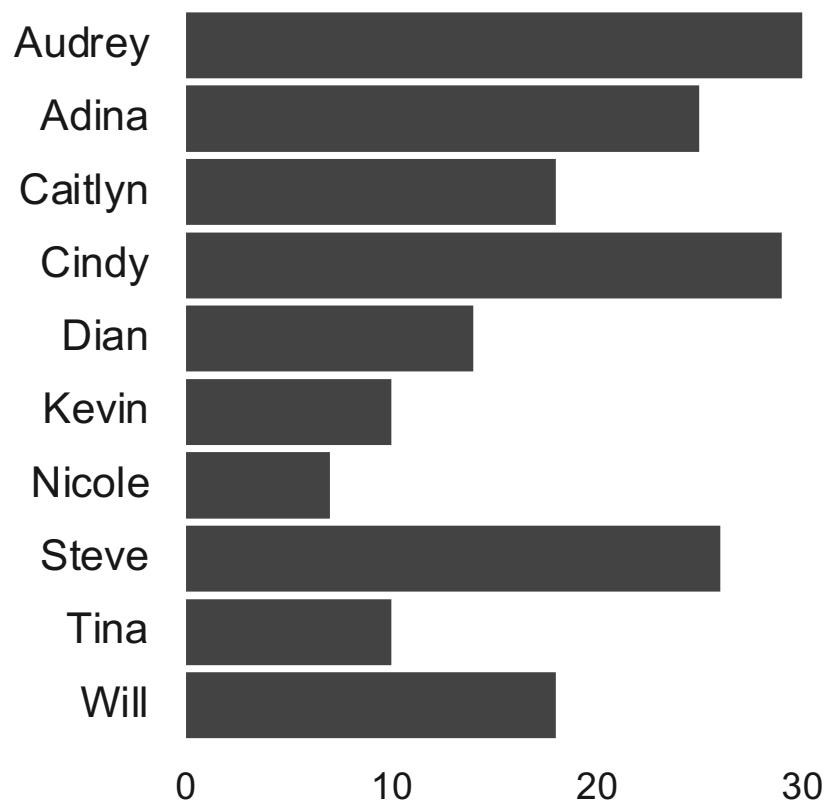


Do: Highlight and annotate your key findings

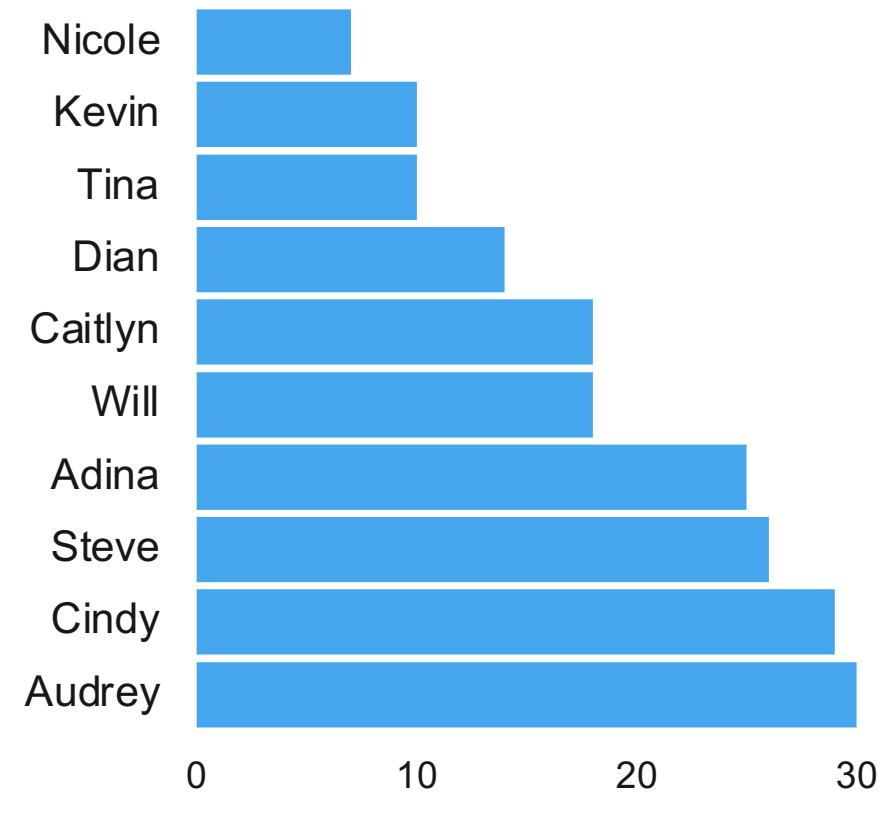


Do: Order your data logically, e.g., alphabetically or by number

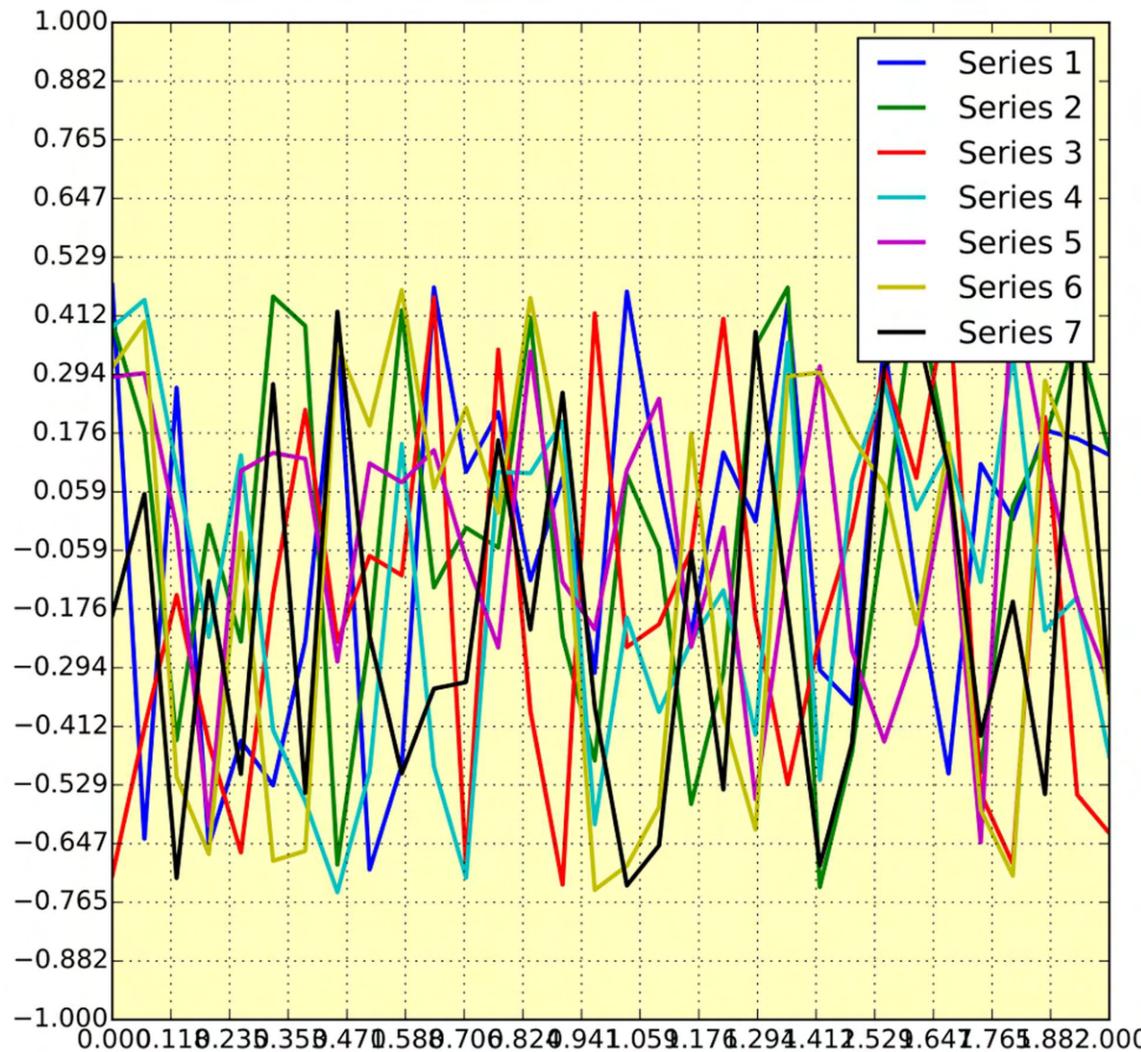
Student's scores for test 1



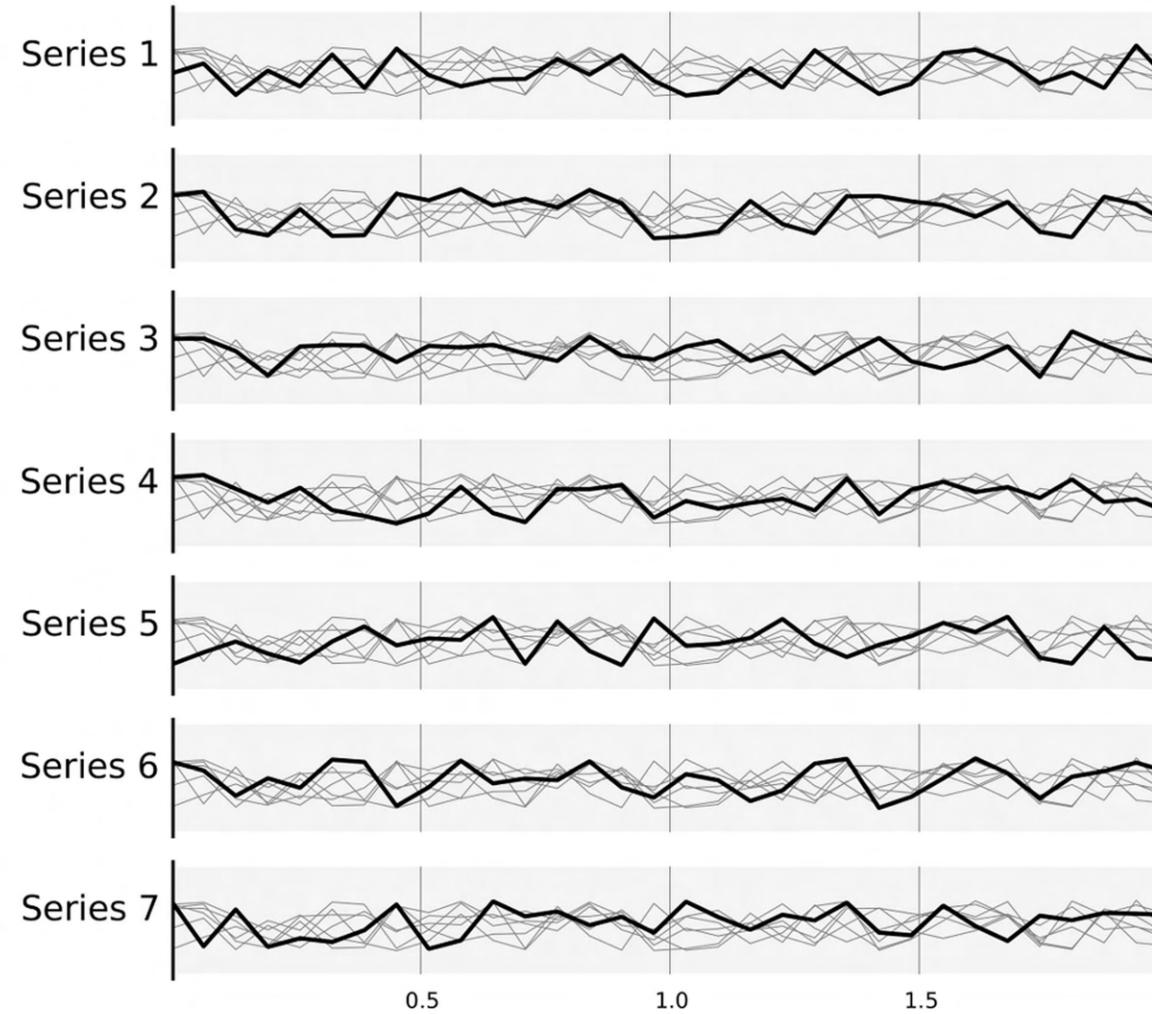
Student's scores for test 1



Don't: Use a spaghetti chart

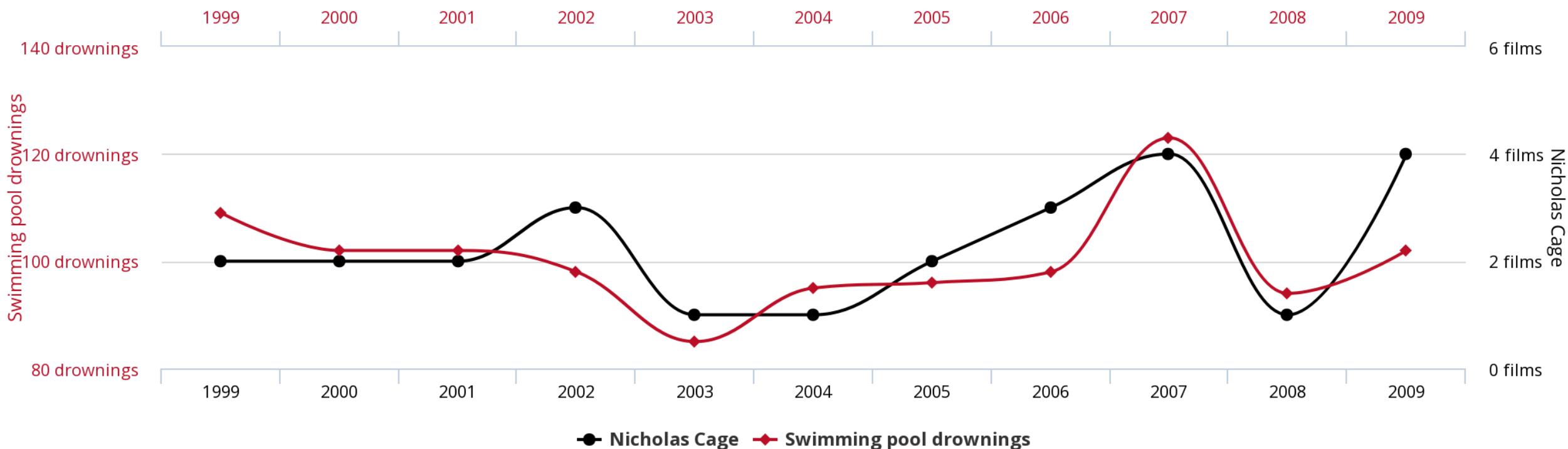


Do: Use small multiples instead

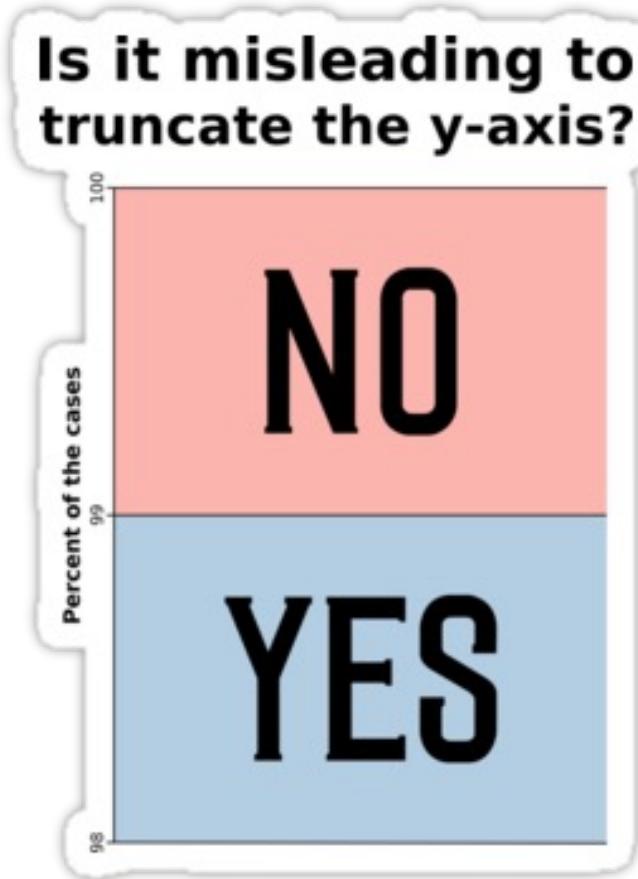


Don't: use double axes

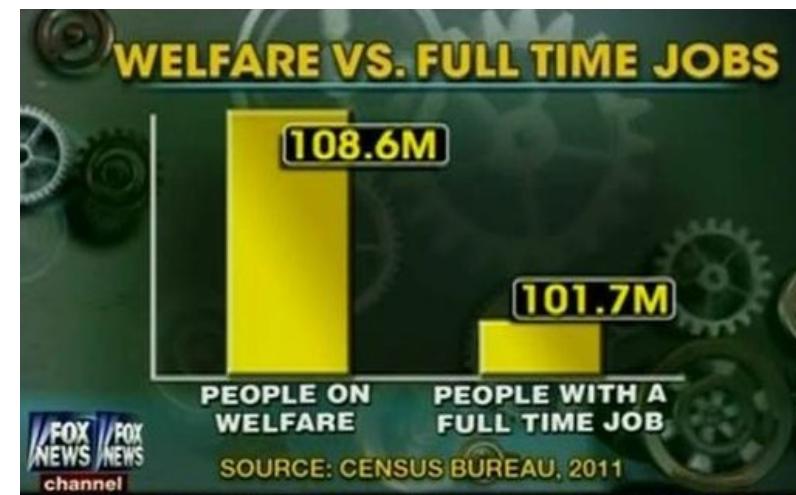
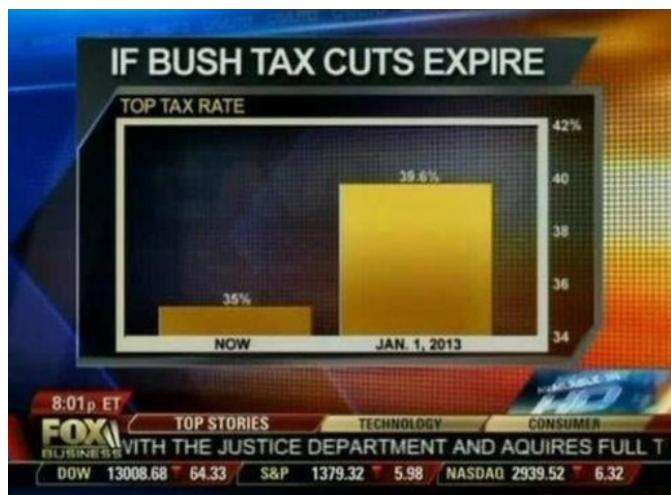
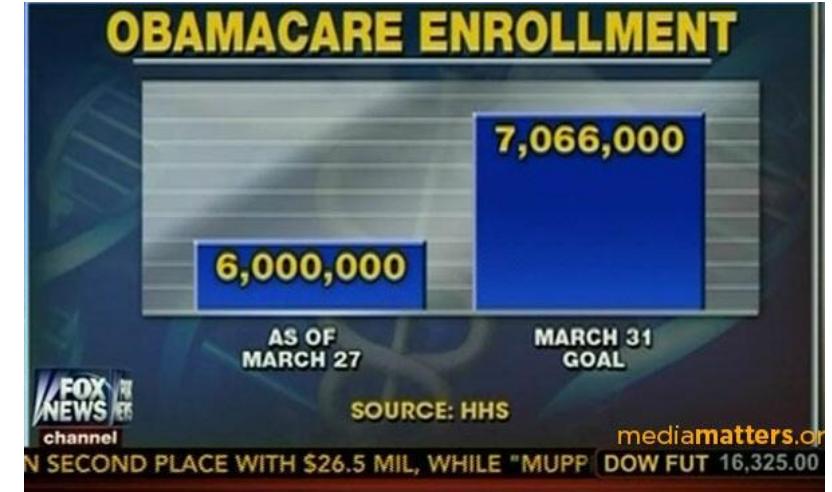
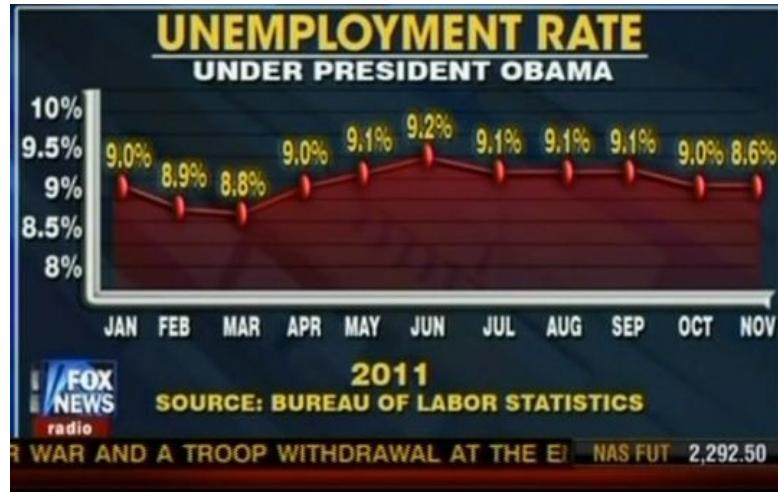
Number of people who drowned by falling into a pool
correlates with
Films Nicolas Cage appeared in



Don't: Truncate the y-axis

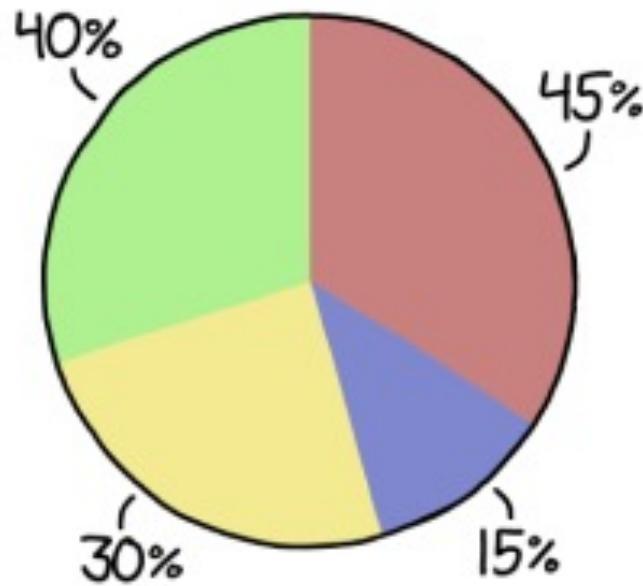


Who does that, you say?

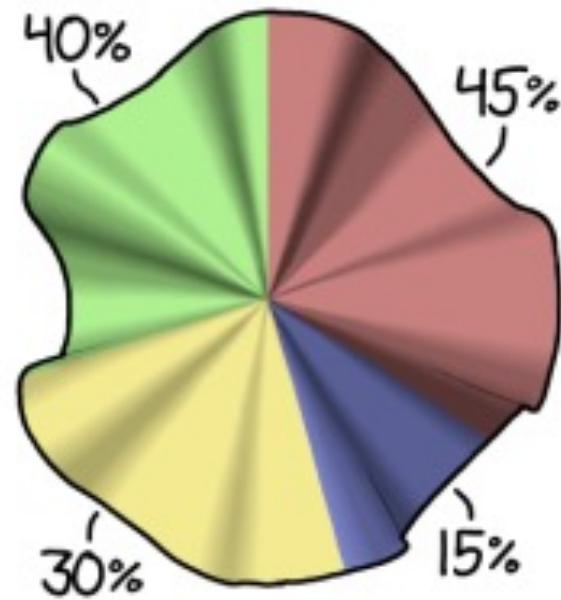


Don't: Use pie charts where the numbers don't add up

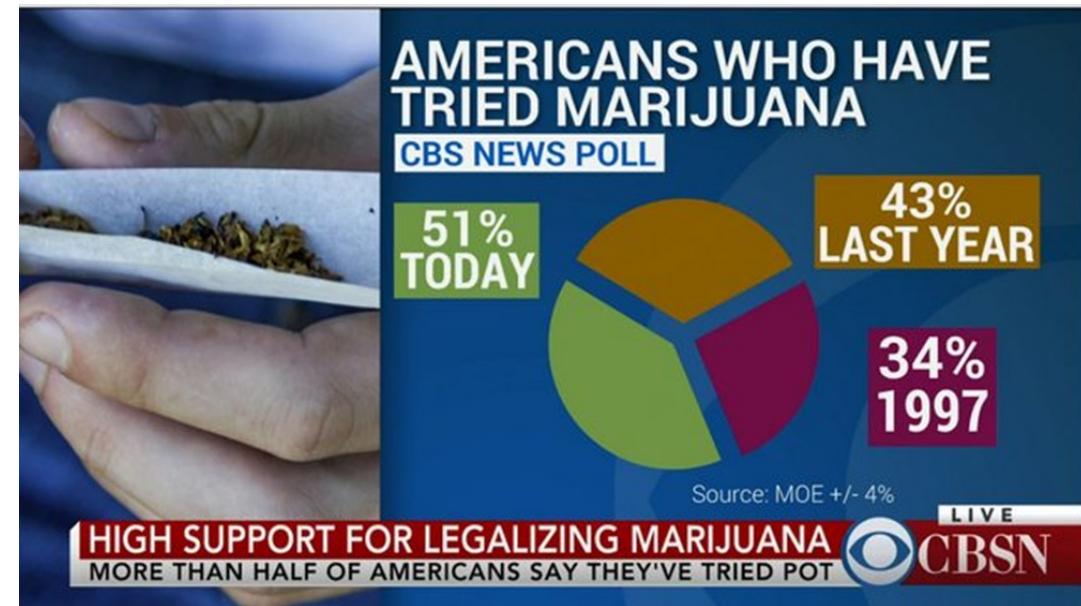
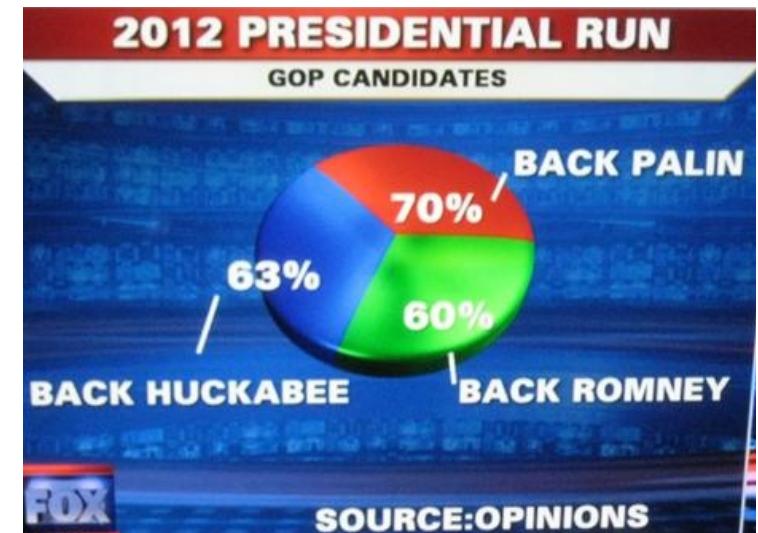
WRONG:



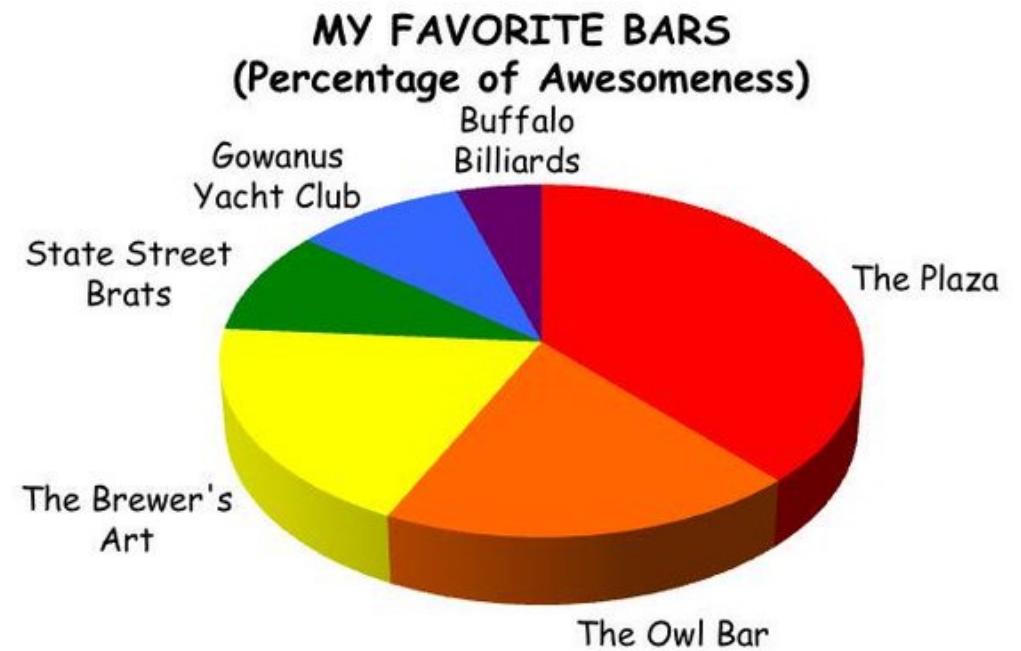
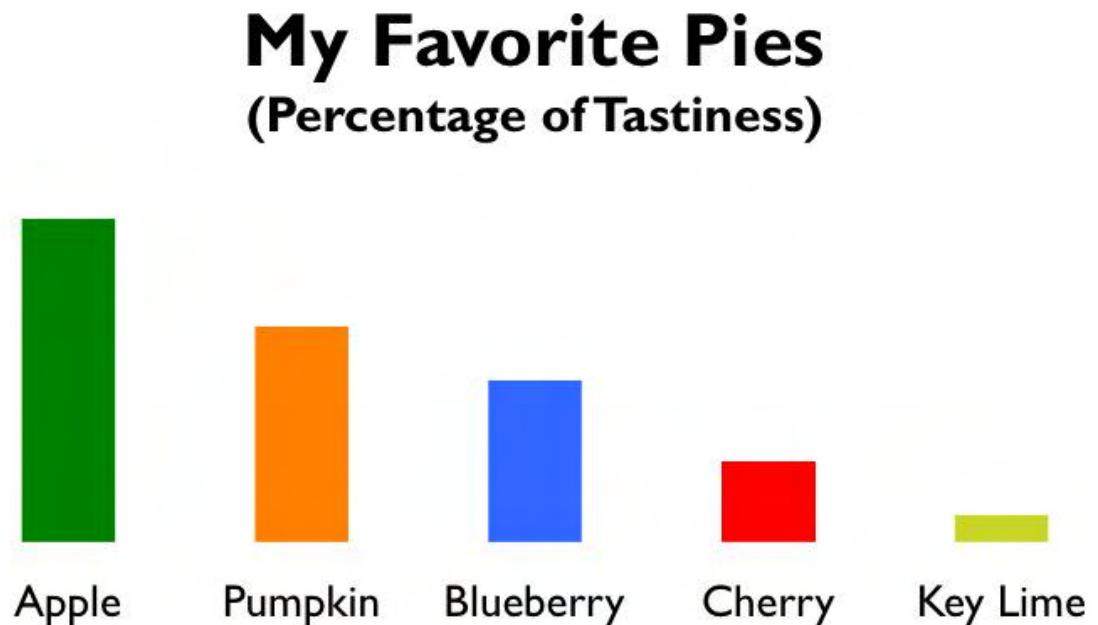
RIGHT:



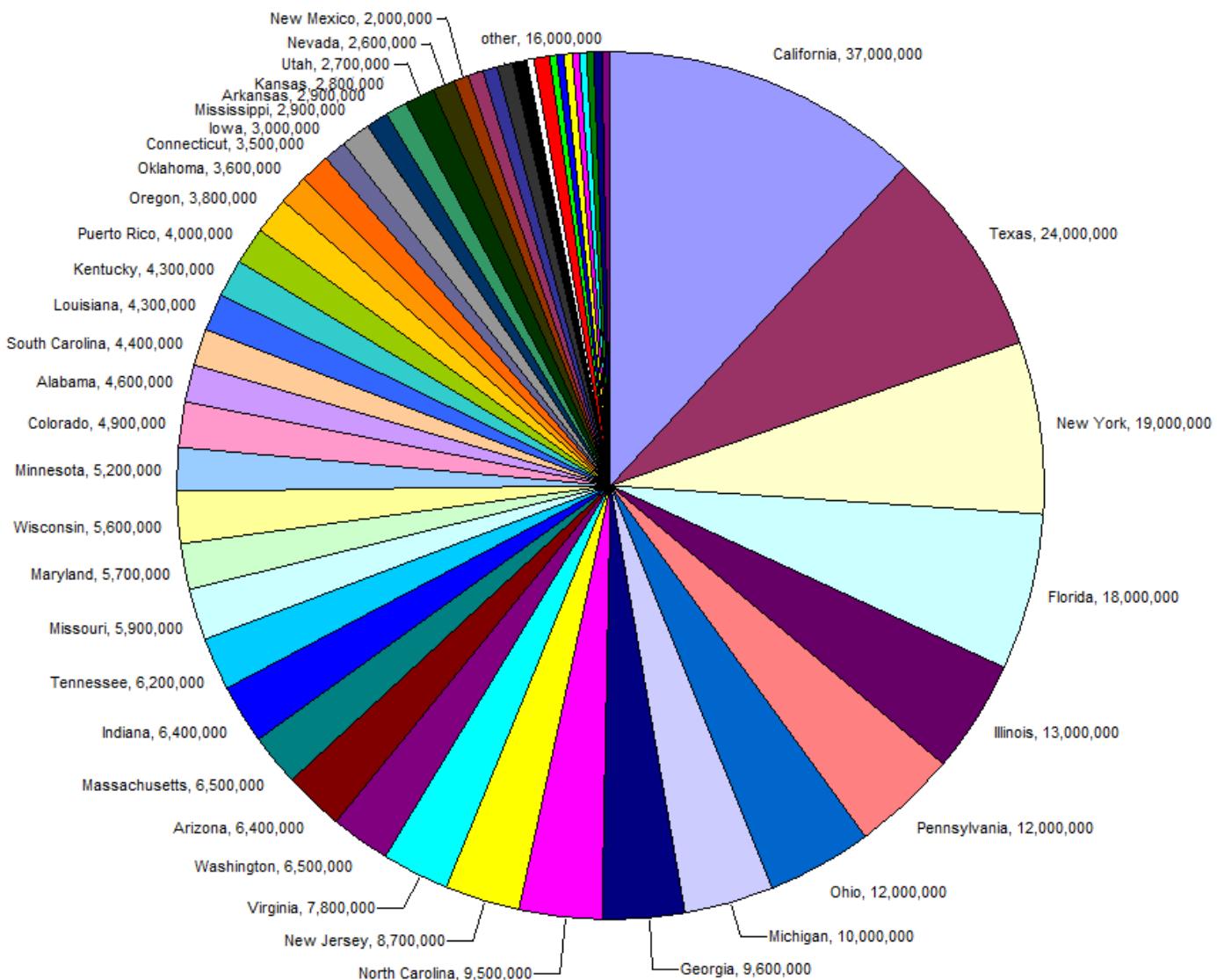
Who does that, you say?



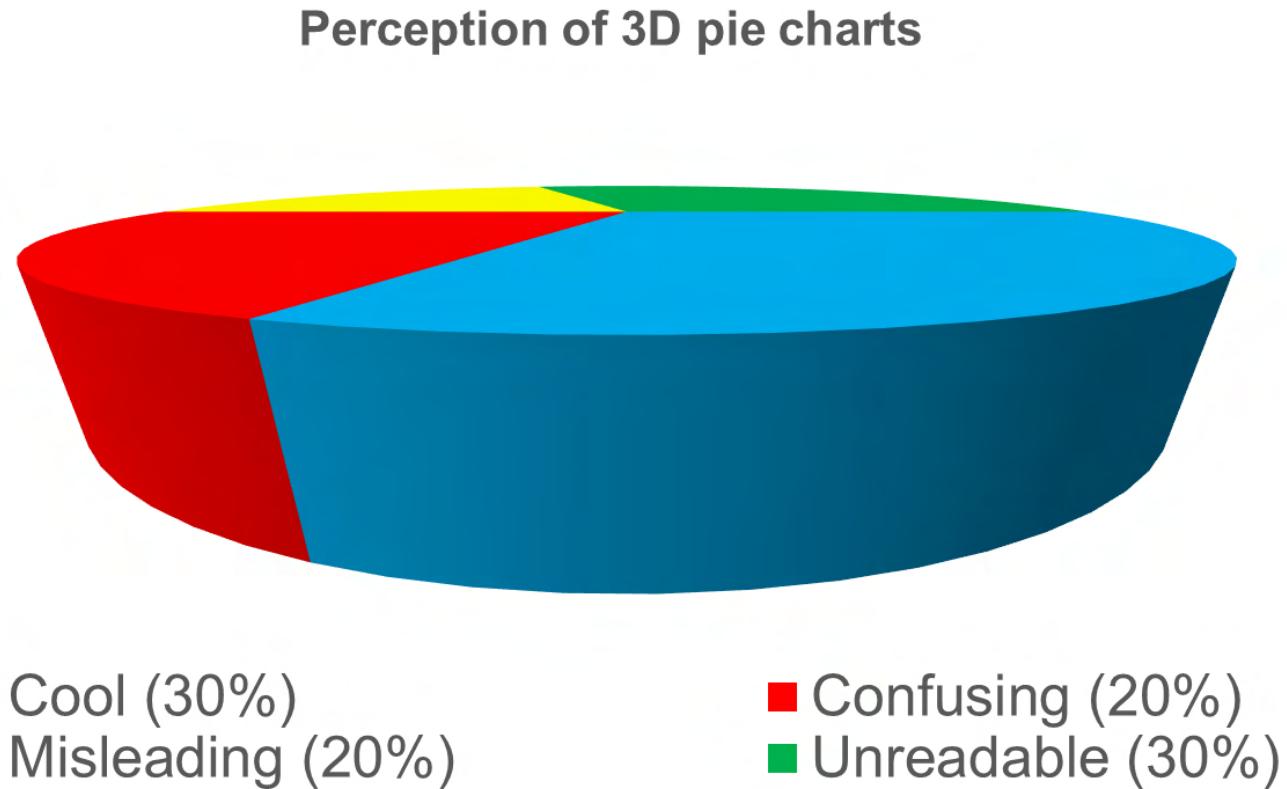
Do: Know when to choose for bars or pies



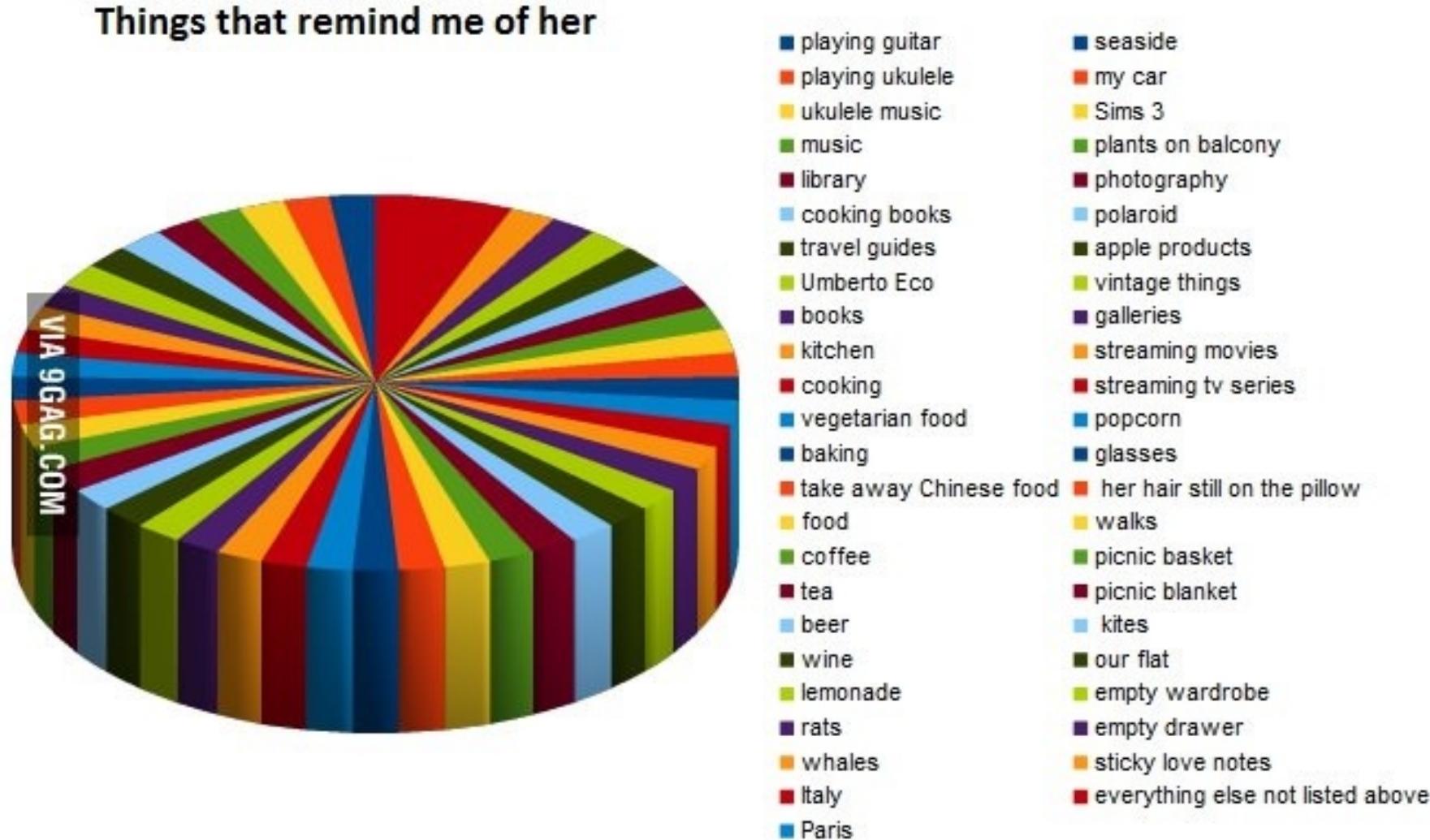
Don't: Use pie charts with >5 categories



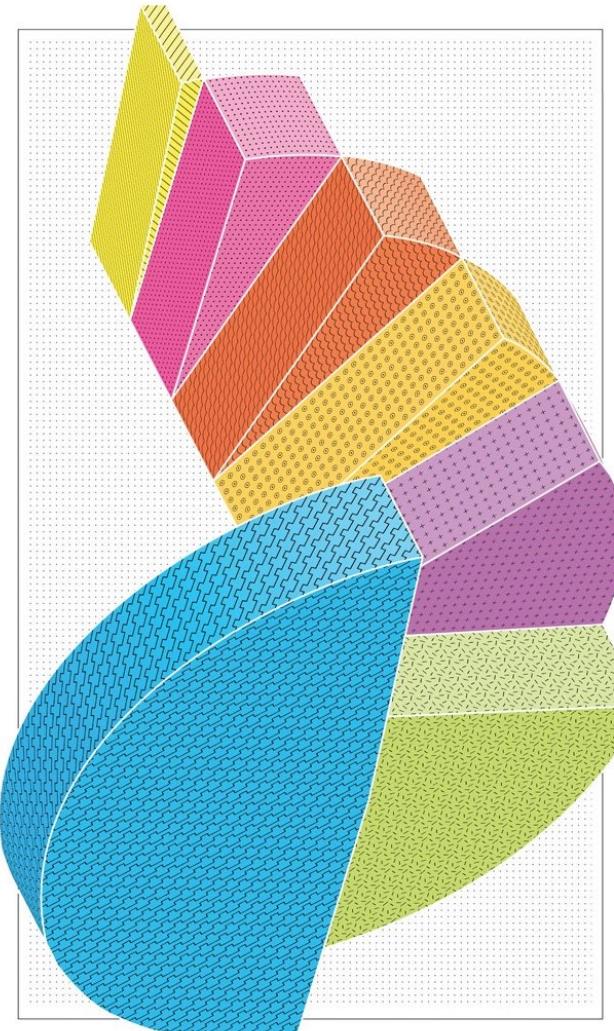
Don't: Use 3D pie charts



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



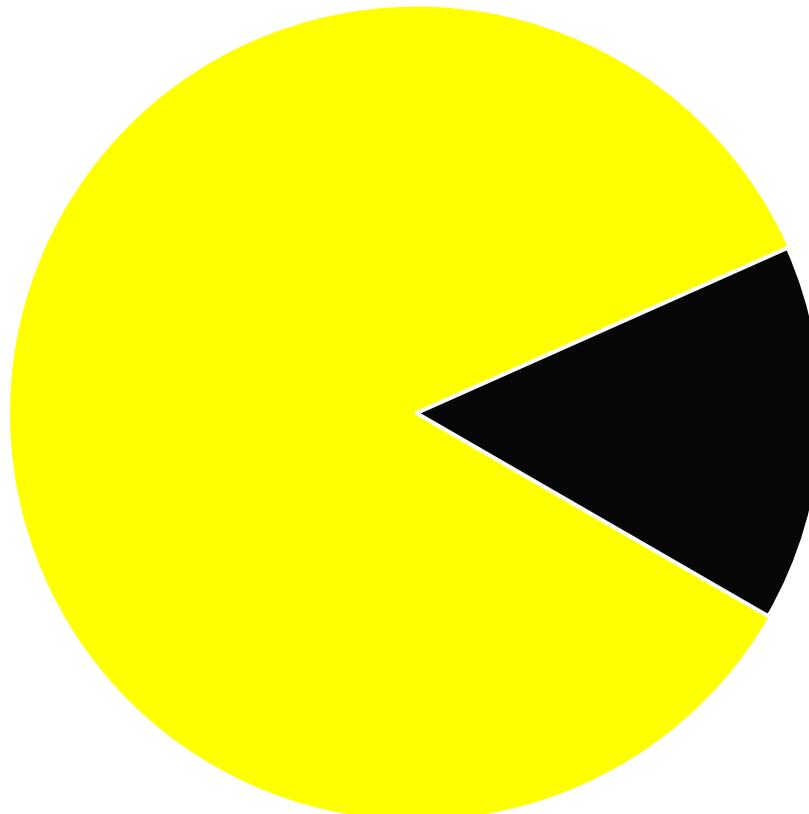
Don't: Use patterned 3D winding staircase lying on its side pie diagrams with >5 categories!



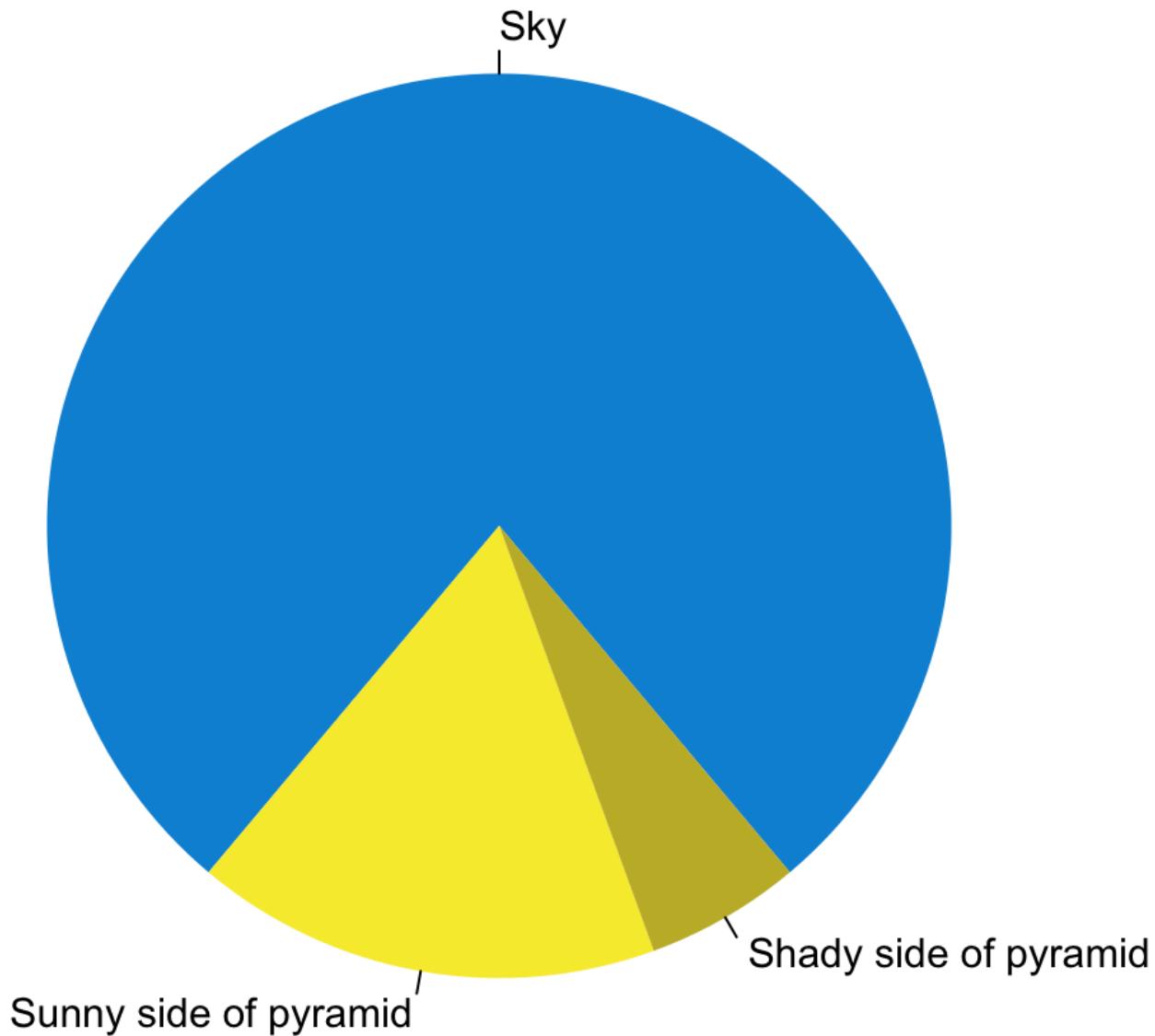
Do: Use pie charts to show the part-to-whole relation of 2 categories

Percentage of chart that looks like Pacman

- Pacman
- Not Pacman



Do: Pie charts for 3 categories is okay too



Do: Sometimes >5 groups work ☺



Do: Know when pie charts are *the best!*



Thank you!



Irene van den Broek, PhD

✉ Irenevandenbroek@gmail.com

LinkedIn icon @IrenevdBroek

Twitter icon @JeBentWatJeMeet

Happy Viz-ing!



<https://irenevdb.rbind.io>

Find the slides and resources here:

<https://irenevdb.github.io/VLAG/Presentation.pdf>

<https://irenevdb.github.io/VLAG/ReferenceList.pdf>

