

**BEFORE WE START.... A LITTLE  
EXPERIMENT**

**37 75**

37

75

1. Grab a set of post-it and a pen; gather up in pairs



2. Try to come up with as many possible representations/encodings for the “data” above as you can, in the paper segments.

**Feel free to be creative!**

37

75

Thirty-seven  
Seventy-five

XXXVII

LXXV

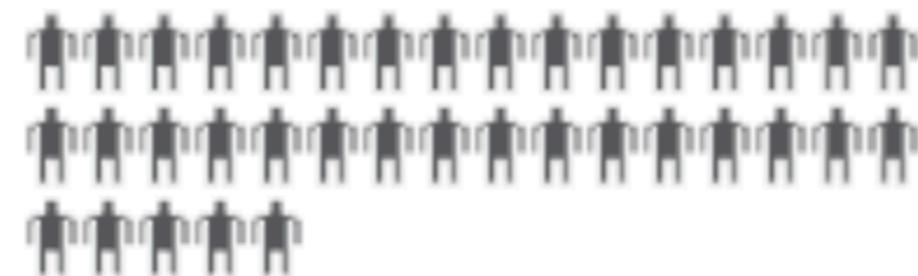
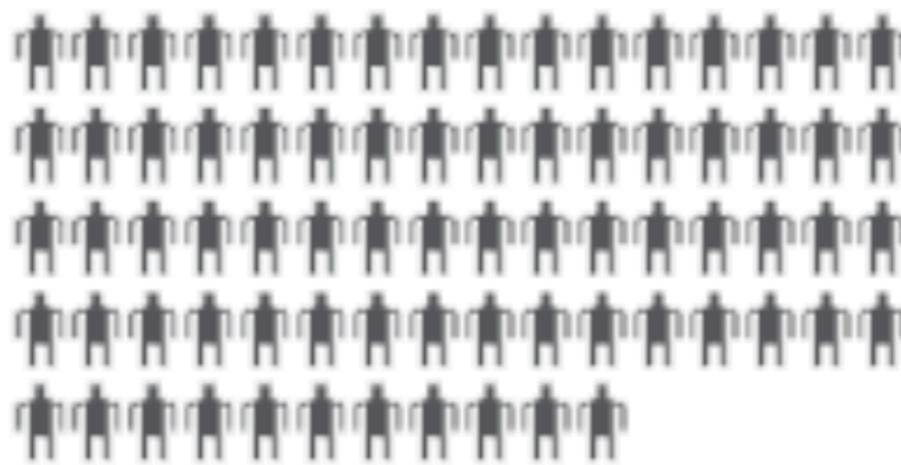


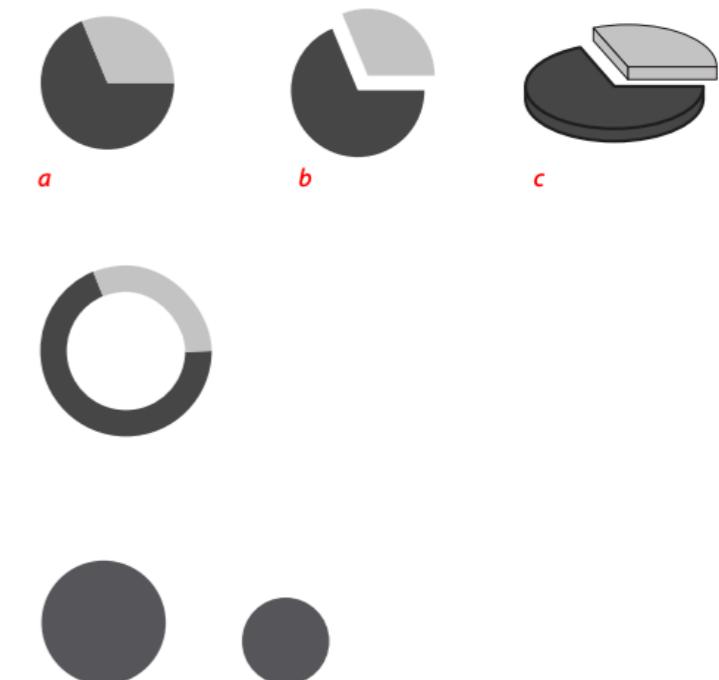
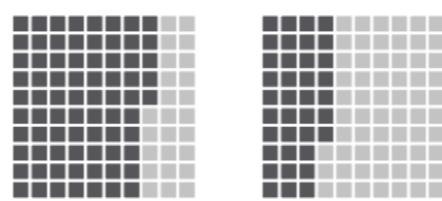
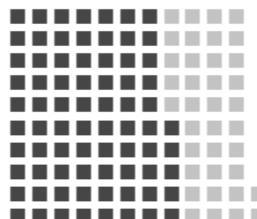
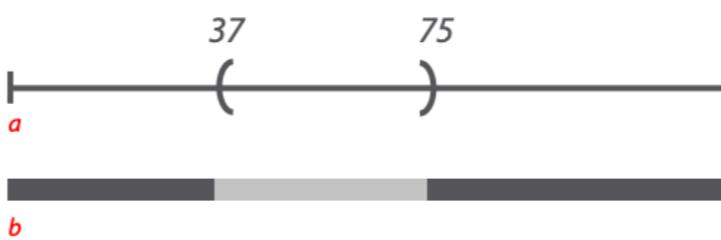
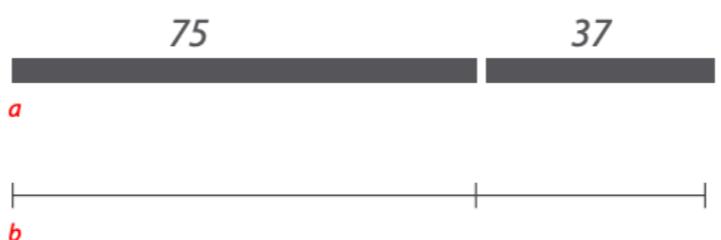
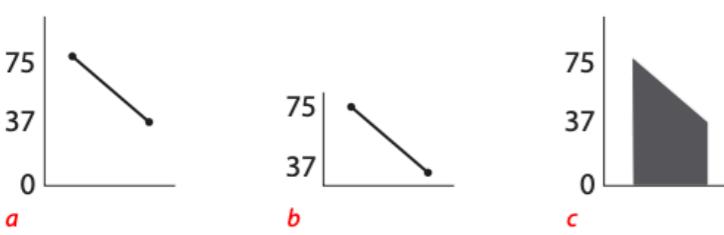
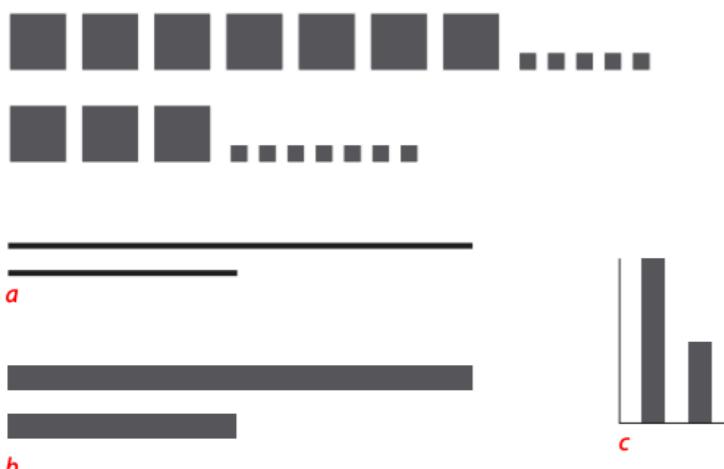
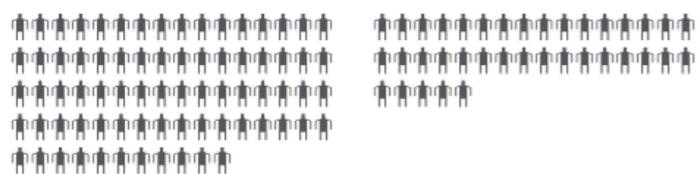
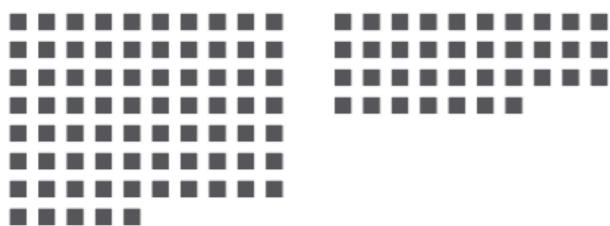
串 串 串 串 串  
串 串 串 串 串  
串 串 串 串 串  
= =

# Squares



# Isotypes







*a*

*b*

*c*



*a*



*b*



*a*



*b*

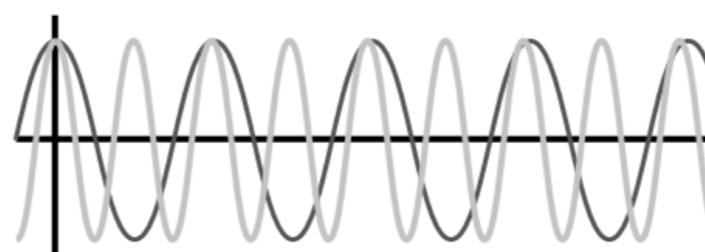
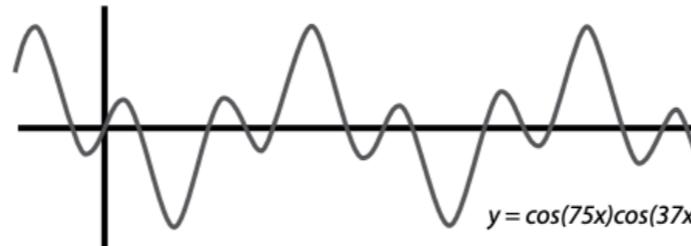
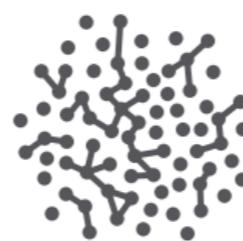
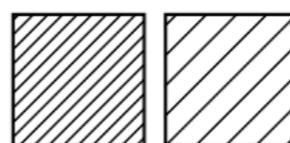
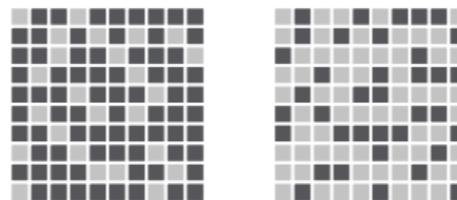
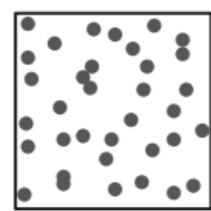
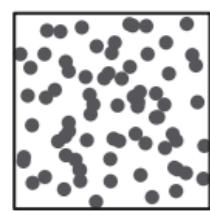
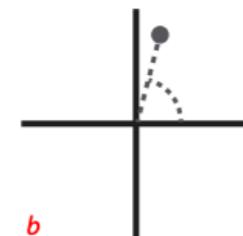
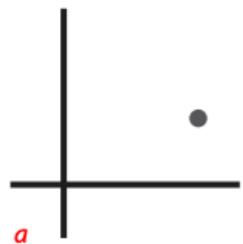


*a*



*b*

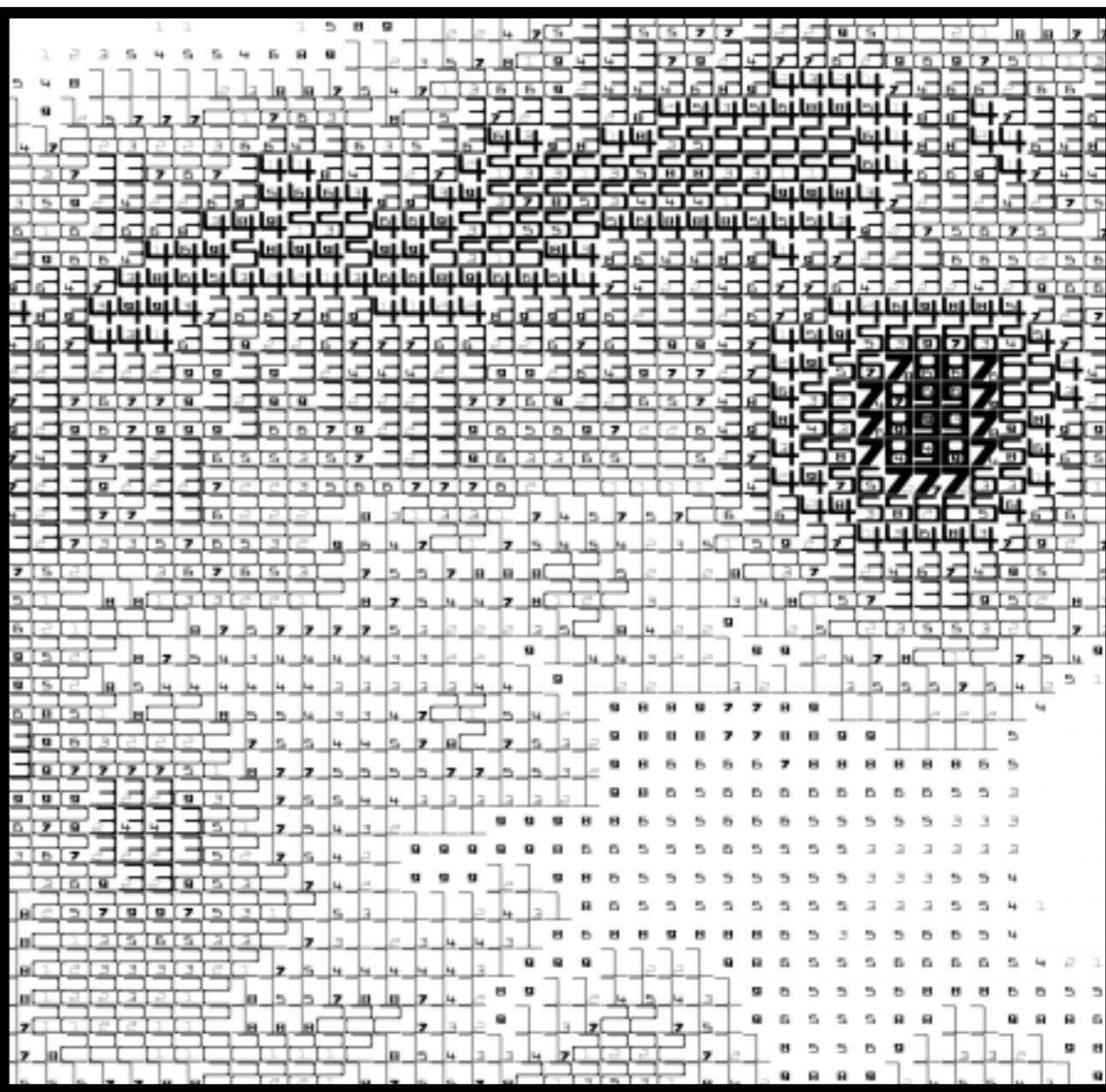
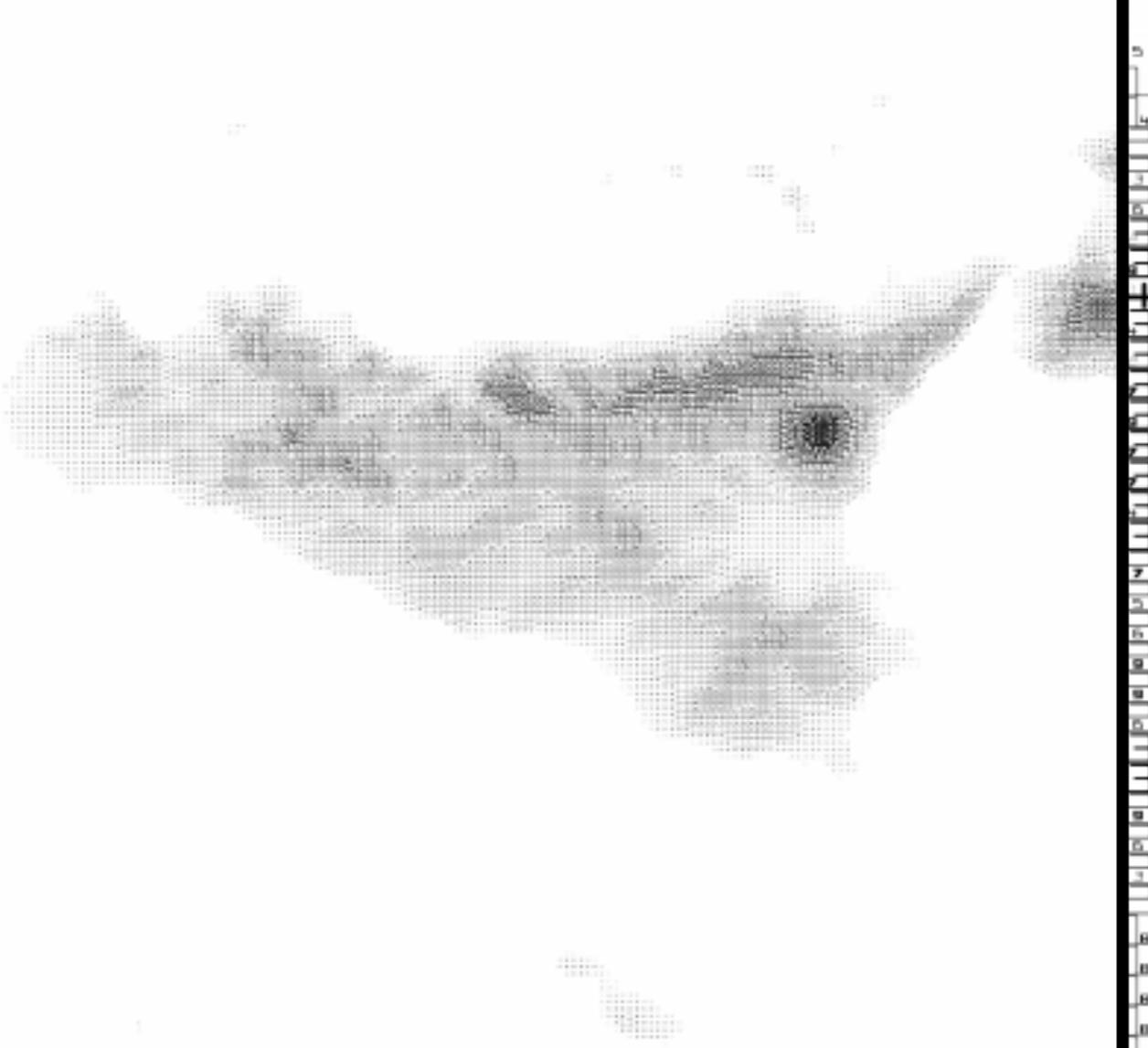




7<sub>s</sub> 3<sub>s</sub>

1 2 3 4 5 6 7 8 9

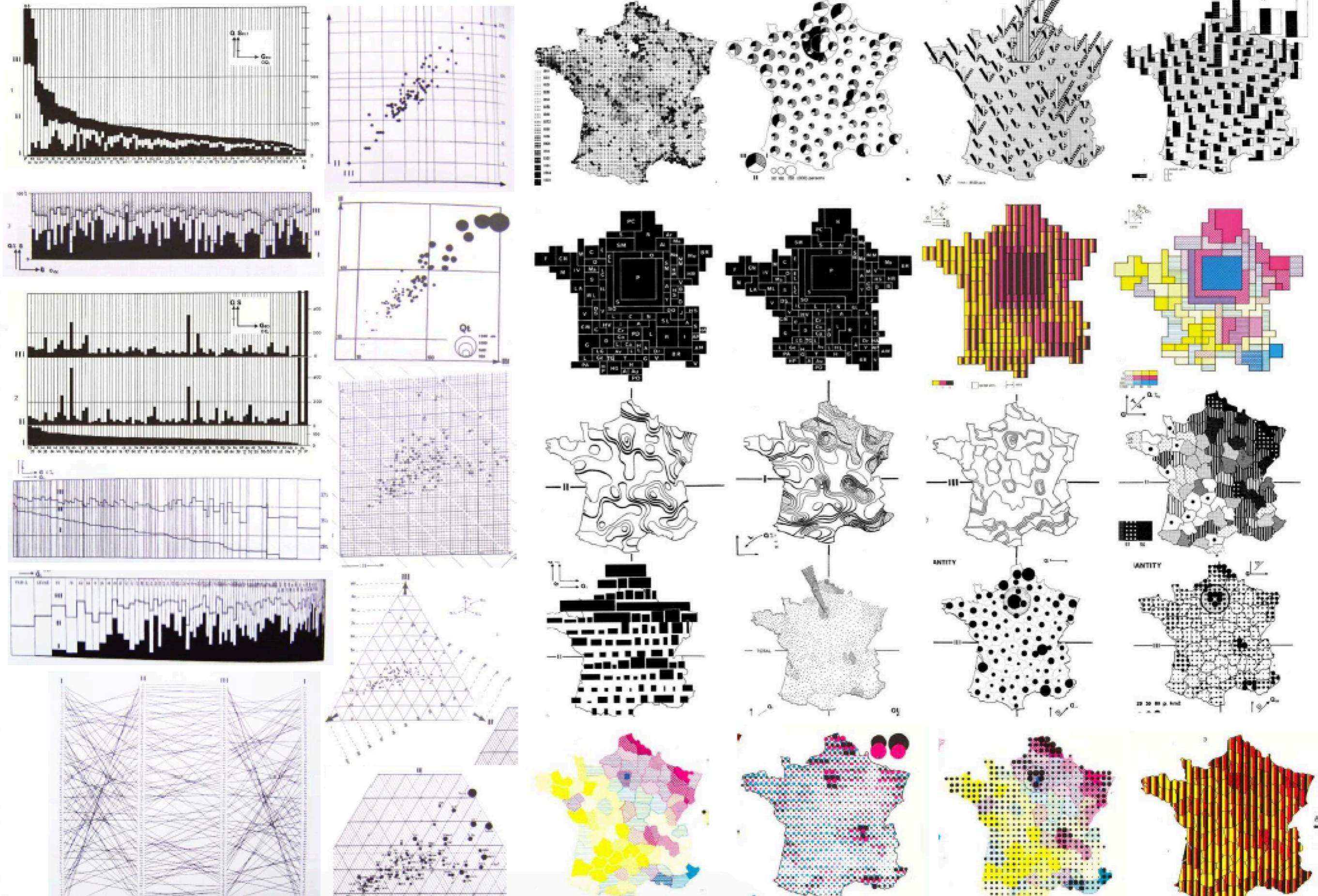
1 9 2 3 4 8 5 6 7 8 9



[Source: M. Nacenta, U. Hinrich, S. Carpendale - <http://fatfonts.org>]

## Départements

	Quantités (000)			Proportion %					
	I	II	III	Total	I	II	III		
1 AIN	87	13	40	140	45	38	27		
2 AISNE	94	24	64	182	28	37	34		
3 ALLIER	65	45	57	167	38	37	34		
4 Hautes ALPES	15	8	12	35	43	34	33		
5 Basses ALPES	16	8	13	37	44	21	35		
6 ALPES MARITIMES	31	61	122	214	14	39	57		
7 ARDÈCHE	48	32	25	105	45	21	24		
8 ARDENNES	25	53	21	119	22	47	31		
9 ARIège	35	17	14	66	52	29	24		
10 AUDE	28	48	26	112	25	43	32		
11 AVEYRON	50	20	32	102	49	19	33		
12 BOUCHES DU RHÔNE	79	92	29	199	54	34	21		
13 CALVADOS	79	55	69	193	36	38	36		
14 CANTAL	45	13	29	83	56	16	21		
15 CHARENTE	65	26	39	140	47	28	25		
17 CHARENTE MARITIME	79	39	63	182	43	21	36		
18 CHIERS	43	41	36	120	26	34	39		
19 COISE	64	23	29	116	25	28	23		
21 CÔTE D'OR	43	44	29	112	30	29	41		
22 CÔTES-DU-NORD	131	33	62	226	58	15	57		
23 CRÈUSE	59	13	17	89	60	15	19		
24 DORDOGNE	104	34	42	180	58	19	23		
25 DOLUS	25	67	39	132	25	47	39		
26 DRÔME	46	38	35	119	39	38	39		
27 EURE	48	52	45	145	28	36	31		
28 EURE & LOIR	64	27	38	129	41	25	34		
29 FINISTERE	154	26	89	269	50	22	27		
30 GARD	40	31	32	104	29	26	36		
31 HAUTE-GARONNE	64	67	64	215	39	31	39		
32 GERS	43	19	16	78	71	11	19		
33 GIROUD	115	197	179	392	30	27	43		
34 HERAULT	92	69	71	172	26	23	41		
35 ILLE & V.	137	68	82	279	49	23	39		
36 ISÈRE	94	39	32	126	45	26	29		
37 INDRE & LO.	61	61	65	187	39	26	35		
38 ISÈRE	68	126	78	222	24	48	28		
39 JURA	39	24	27	100	39	34	27		
40 LANDS	79	25	29	122	53	20	29		
41 LOIRE & CHER	51	27	29	108	47	25	26		
42 LOIRE	56	169	82	219	19	54	37		
43 Ille LOIRE	81	23	22	97	54	24	22		
44 LOIRE INF.	101	109	105	324	29	34	34		
45 LOIR ET CHER	51	51	34	136	32	28	35		
46 LOT	41	10	16	67	61	15	24		
47 LOT & GAR.	79	24	20	124	37	19	24		
48 LOZÈRE	22	5	7	34	64	13	21		
49 MAINES & LO.	104	65	65	234	44	28	28		
50 MANCE	116	43	56	214	34	20	36		
51 MARNE	64	57	67	188	39	35	34		
52 Haute MARNE	29	26	28	83	31	26	31		
53 MAYENNE	74	23	29	126	23	28	39		
54 MONTBELLÉGNY	23	127	91	241	8	53	57		
55 MOSSE	24	31	27	82	24	27	39		
56 MORBIAN	132	47	59	238	55	20	55		
57 MUSSELLE	36	179	94	309	94	94	93		
58 NIÈVRE	34	21	23	70	34	39	29		
59 NOIR	81	493	298	809	9	56	39		
60 ORNE	49	63	55	168	49	35	46		
61 OISE	65	39	35	139	42	27	39		
62 P. I. C.	94	242	127	473	20	51	51		
63 Pyrénées ORIENTALES	60	79	69	228	35	35	36		
64 Pyrénées OCCIDENTALES	60	69	68	197	40	40	39		
65 Haute PYRÉNÉES	37	27	29	90	40	29	39		
66 PYRÉNÉES ORIENTALES	25	29	23	76	40	29	29		
67 HAUTE-SAÔNE	76	122	114	312	24	39	39		
68 HAUTE-SAÔNE	49	121	74	295	17	38	47		
69 SAÔNE	44	213	194	455	10	47	39		
70 SAÔNE & LO.	34	32	29	89	39	38	32		
71 SAÔNE	87	77	62	226	41	33	33		
72 SAÔNE	87	45	58	199	40	24	24		
73 SAÔNE	44	39	35	117	38	32	32		
74 Haute SAÔNE	52	42	45	138	27	36	36		
75 SAÔNE	7	275	949	1327	0	38	31		
76 SAÔNE INF.	8	274	550	1122	1	51	51		
77 SAÔNE & LO.	37	72	36	185	20	38	34		
78 SAÔNE & LOIR	46	329	356	830	0	45	45		
79 SAÔNE-MARNE	71	29	33	133	53	28	28		
80 SAÔNE	57	60	61	186	21	38	38		
81 TARN	55	47	33	125	41	35	35		
82 TARN & G.	48	18	16	59	35	39	37		
83 VAR	53	59	61	184	20	36	31		
84 VAUCLUSE	40	30	41	111	36	37	37		
85 VENDÉE	110	38	40	182	59	20	20		
86 VIENNE	60	29	39	128	47	33	33		
87 Haute VIENNE	64	47	43	156	41	38	38		
88 VORGES	36	95	63	174	21	34	34		
89 YONNE	41	28	37	106	39	26	26		
90 BULLENT	3	25	13	41	0	60	60		
91	444	444	444	1332	6795	6926	16225	28	35



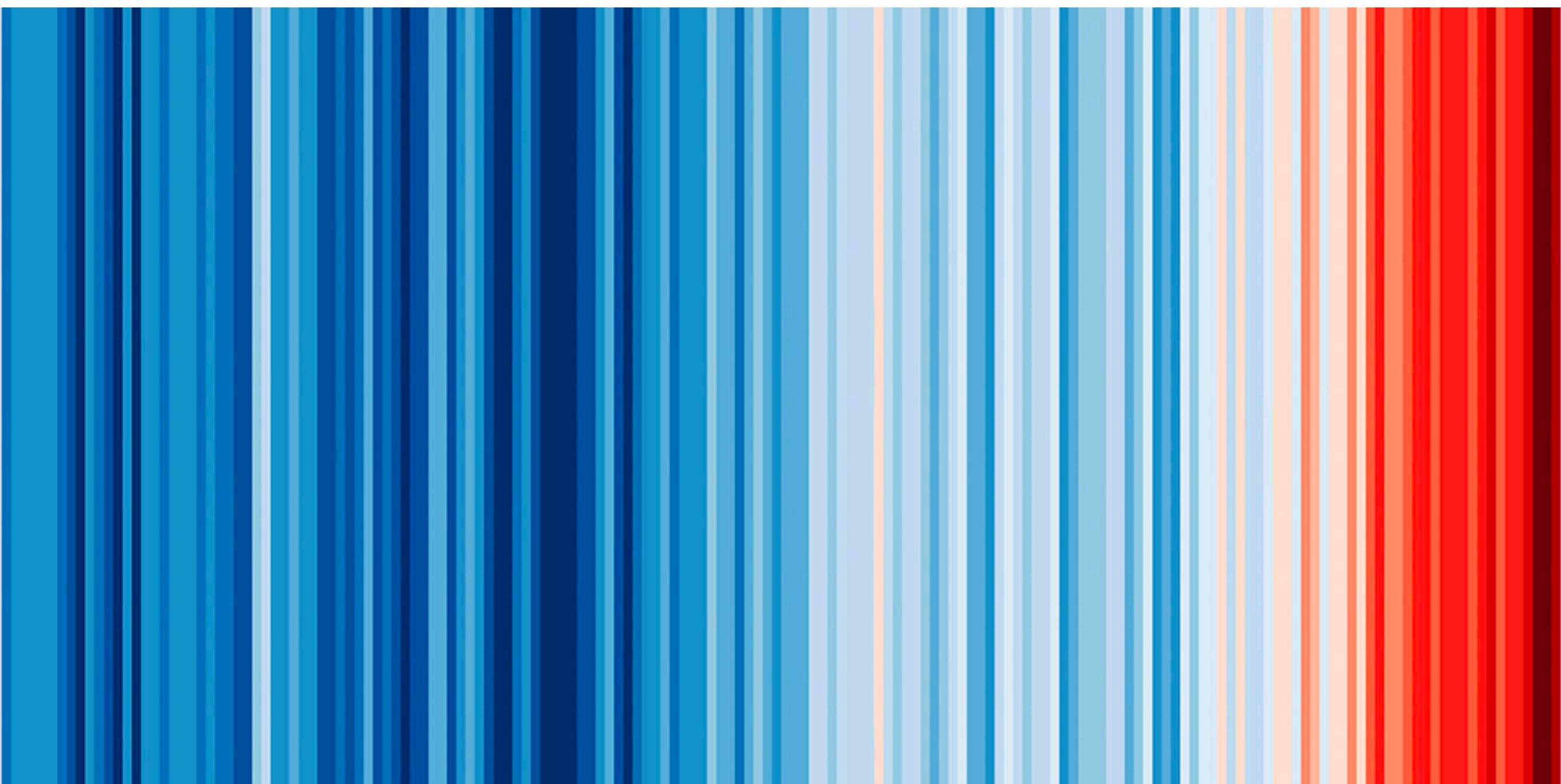
Bertin defined the design space then showed how it could be used to generate many different types of visualizations; using a simple dataset as an example. (Semiology of Graphics, pg 100 ~ 138). Side note French edition appendix on text, not translated into English.

**... ANOTHER LITTLE EXPERIMENT**

**PURPLE** **YELLOW** **RED**  
**BLACK** **RED** **GREEN**  
**RED** **YELLOW** **ORANGE**  
**BLUE** **PURPLE** **BLACK**  
**RED** **GREEN** **ORANGE**

# Communication

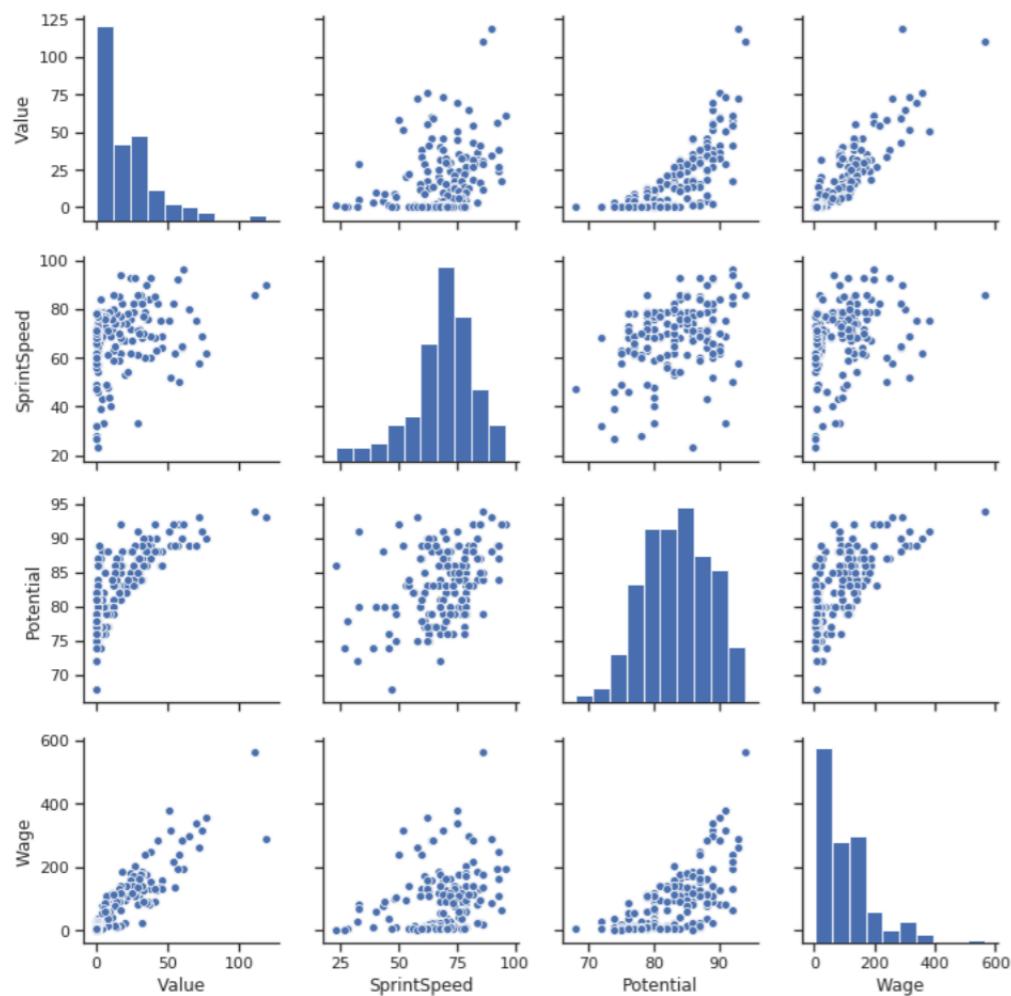
Visualization theory and principles, storytelling with data



# MAIN USES OF DATA VISUALIZATION

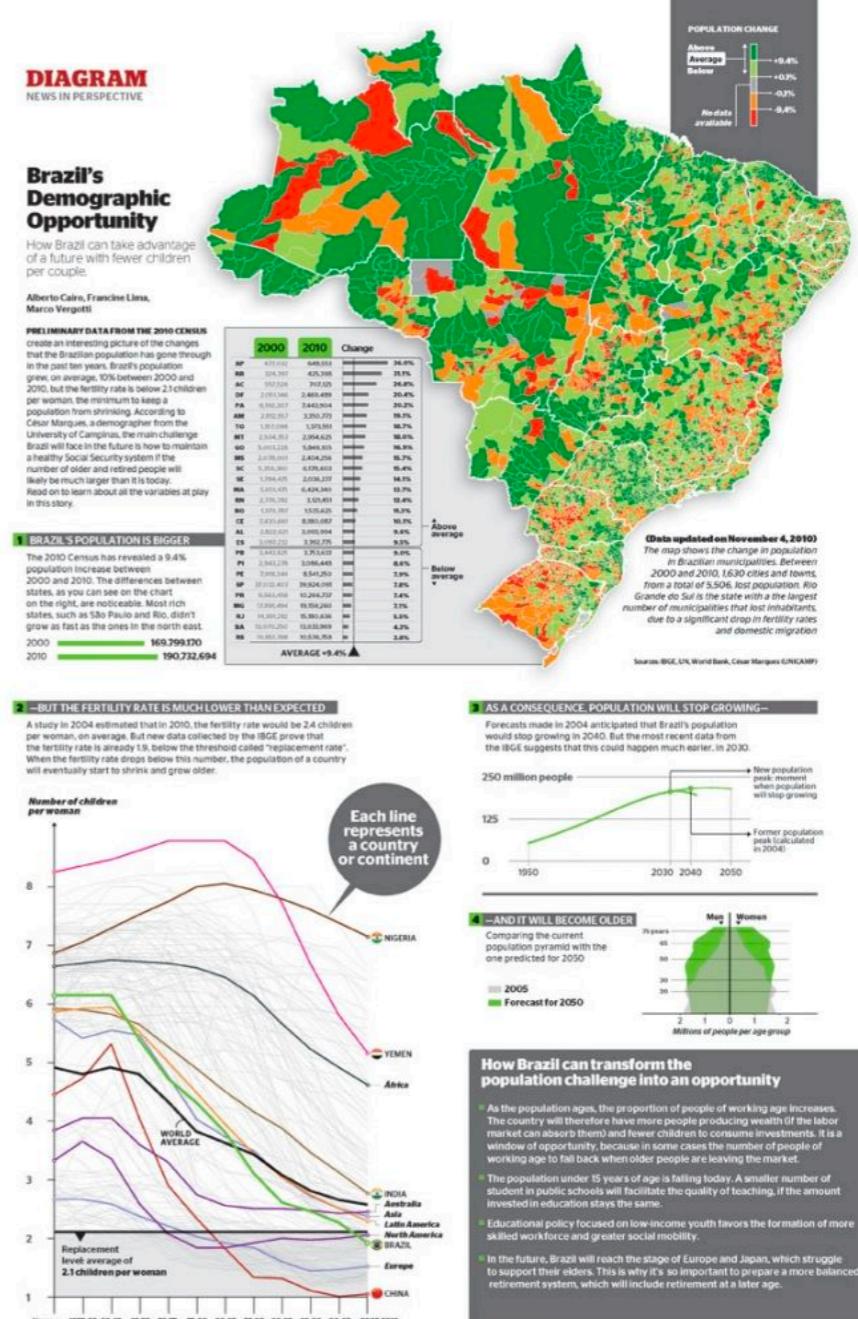
## EXPLORATORY

Acquiring insights



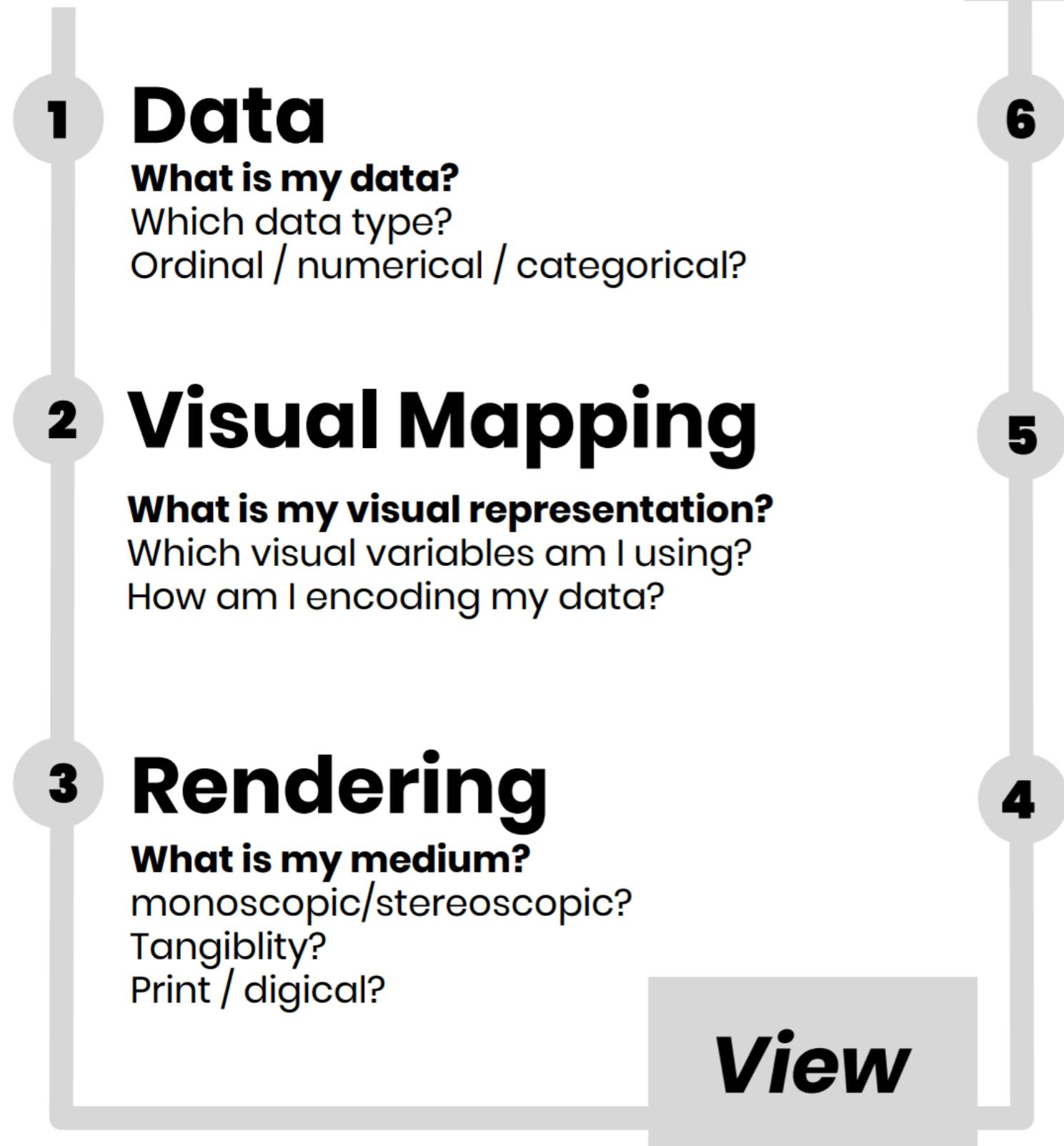
## EXPLANATORY

Storytelling



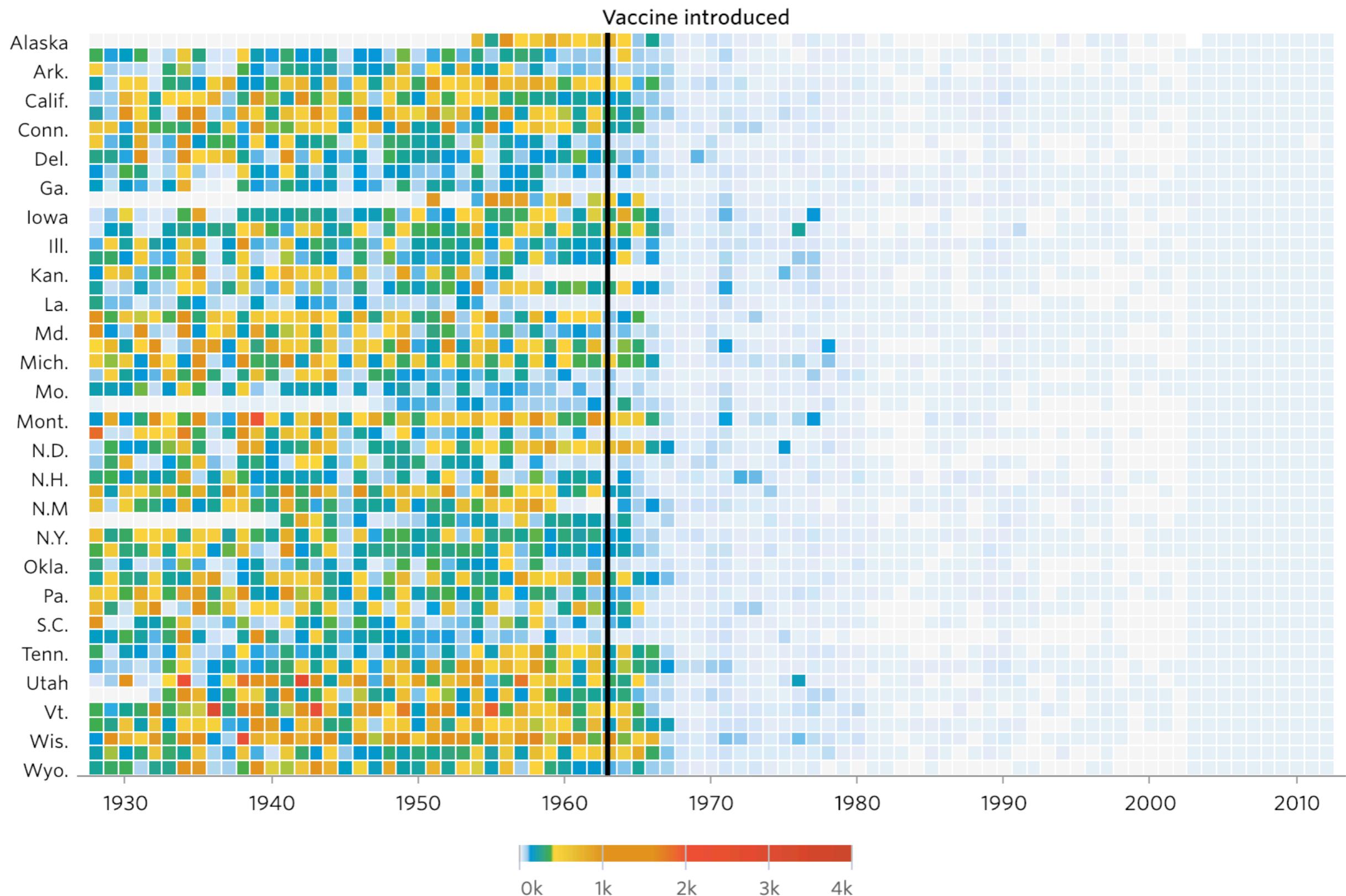
# Encoding: designer

# Decoding: user



# The number of people infected with measles declines after vaccines were introduced

Infected people per 100,000 people, measured over 70-some years and across all 50 states and the District of Columbia

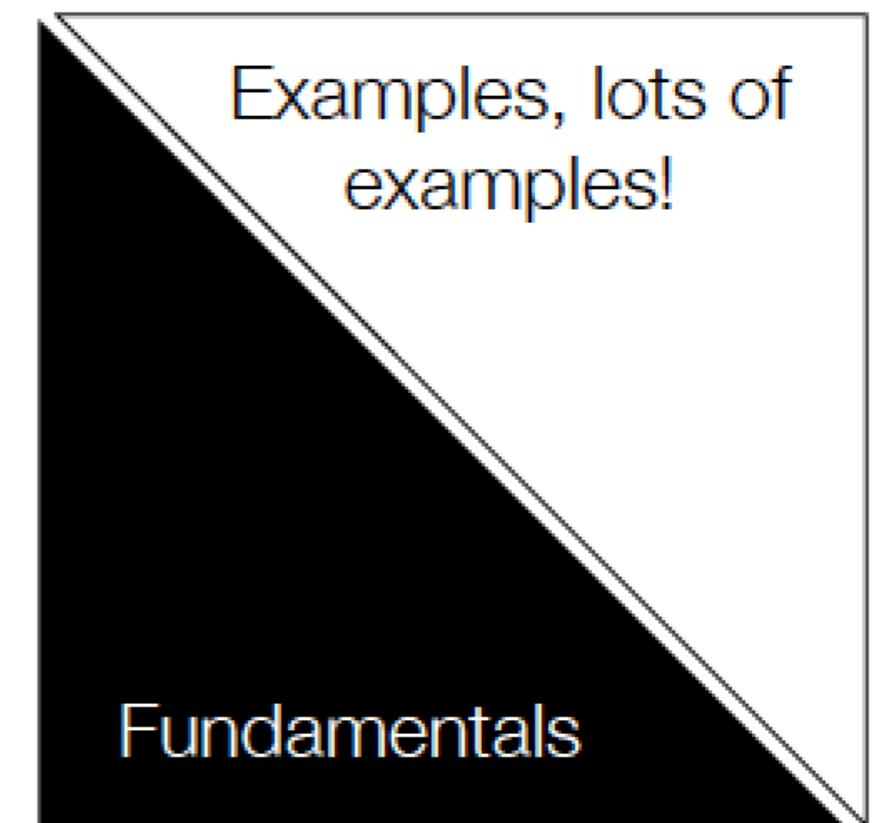


# TAKE AWAY MESSAGE....

It takes **experience!**

But despite of that, we all have the “necessary prerequisites” to be a good visualization designer...

.... we just need to learn how to **see** things a bit differently!

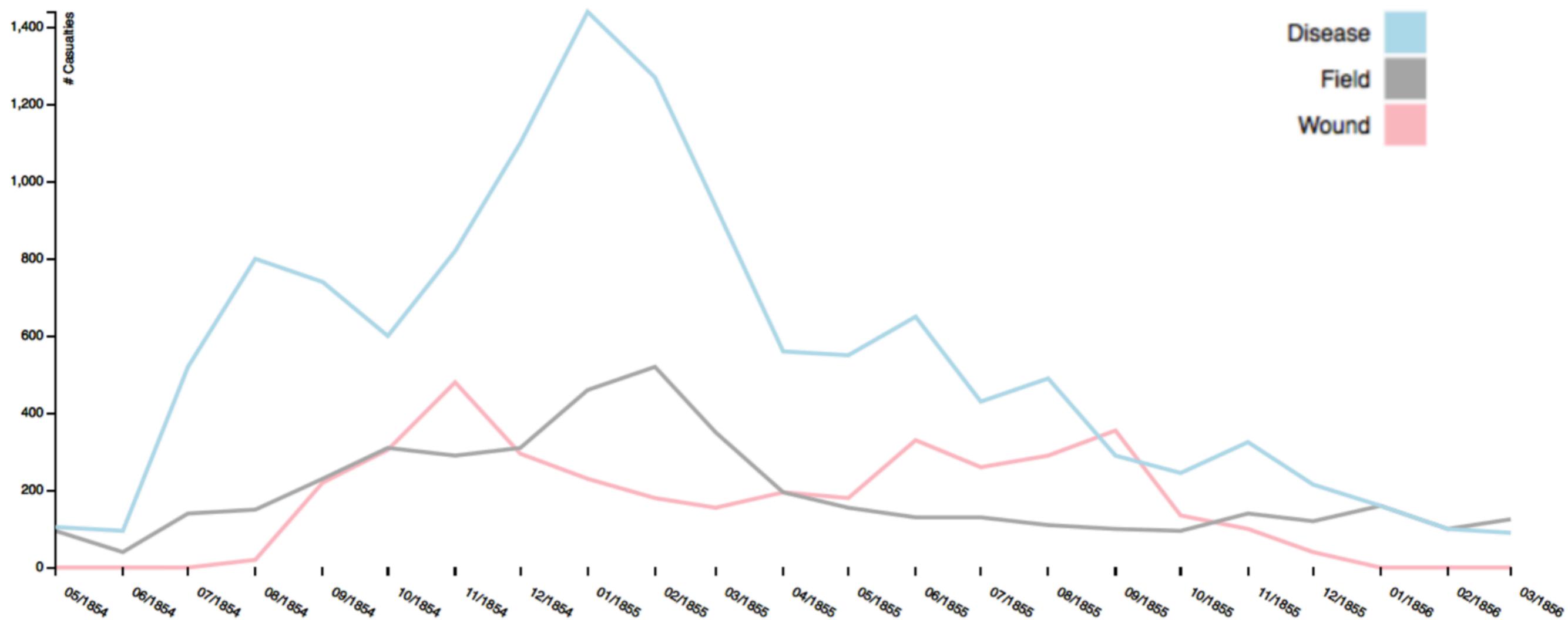


## British Casualties in the Crimean War

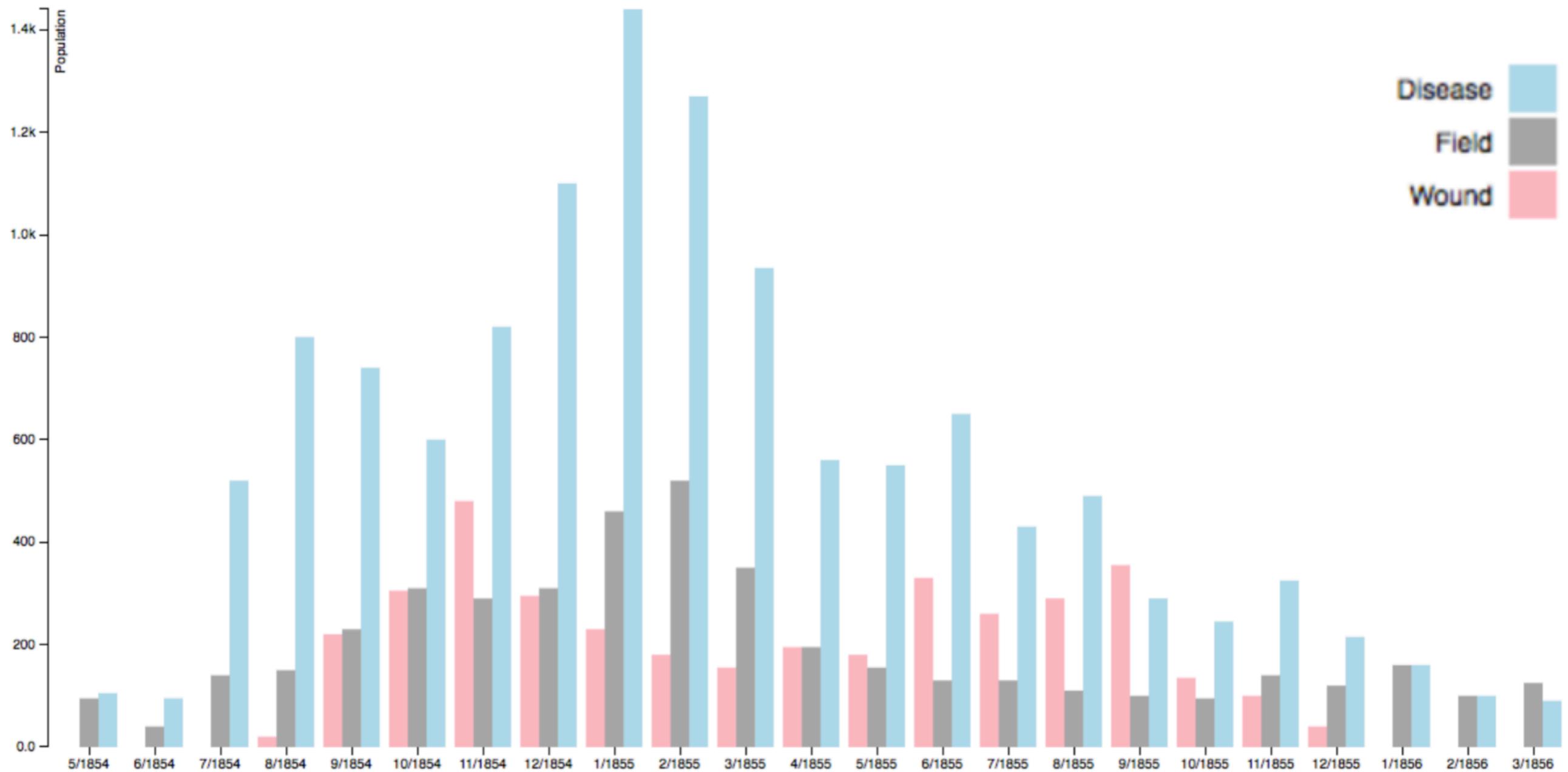
Data	Wound	Field	Disease
05/1854	0	95	105
06/1854	0	40	95
07/1854	0	140	520
08/1854	20	150	800
09/1854	220	230	740
10/1854	305	310	600
11/1854	480	290	820
12/1854	295	310	1100
01/1855	230	460	1440
02/1855	180	520	1270
03/1855	155	350	935
04/1855	195	195	560
05/1855	180	155	550
06/1855	330	130	650
07/1855	260	130	430
08/1855	290	110	490
09/1855	355	100	290
10/1855	135	95	245
11/1855	100	140	325
12/1855	40	120	215
01/1856	0	160	160
02/1856	0	100	100
03/1856	0	125	90

- ▶ Month with highest casualty rates in the field?
- ▶ Months in which deaths by wound exceeds deaths in the field?
- ▶ Month with highest total casualty rate?
- ▶ Months in which % of deaths by disease was below 50%?

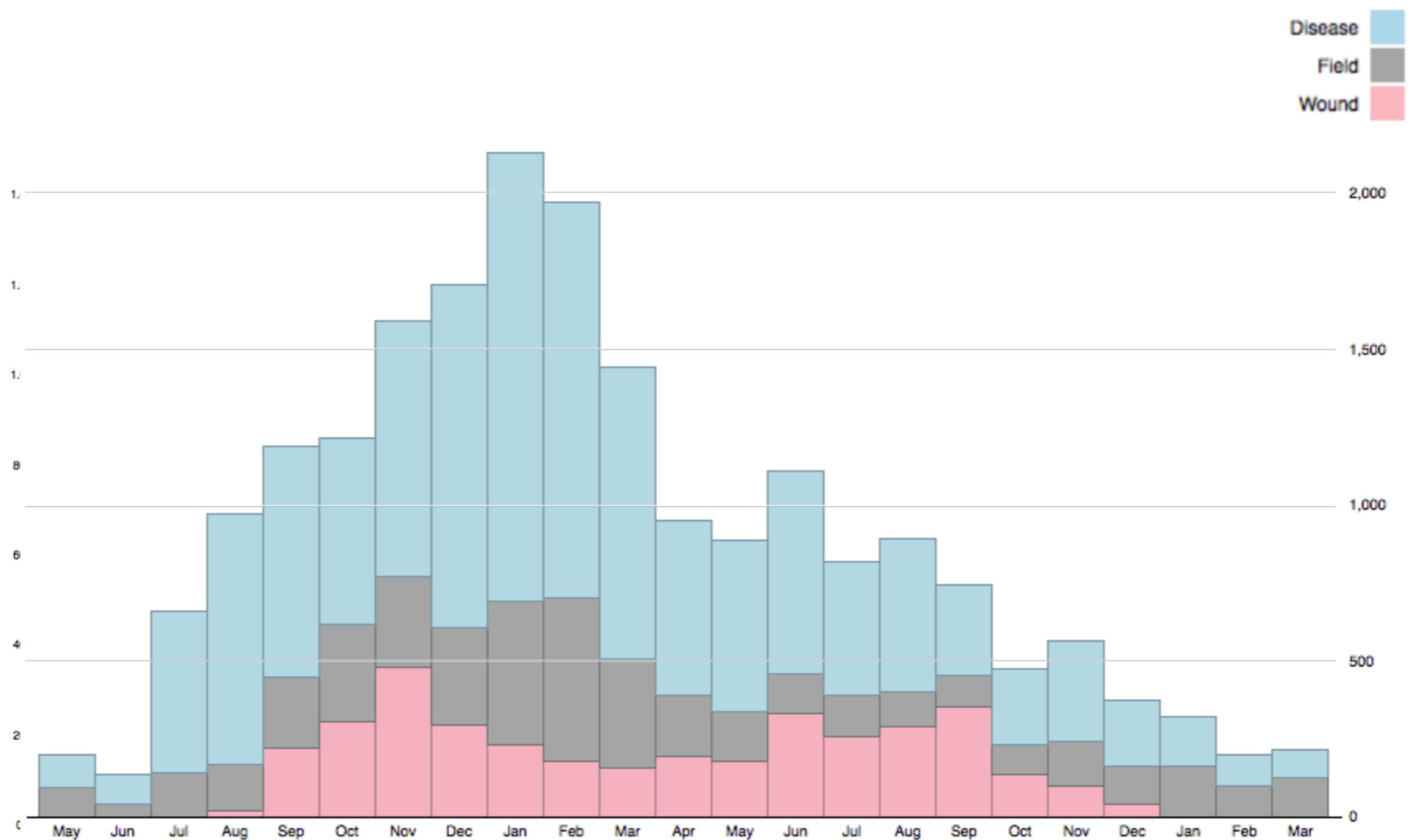
# Month with highest casualty rates in the field?



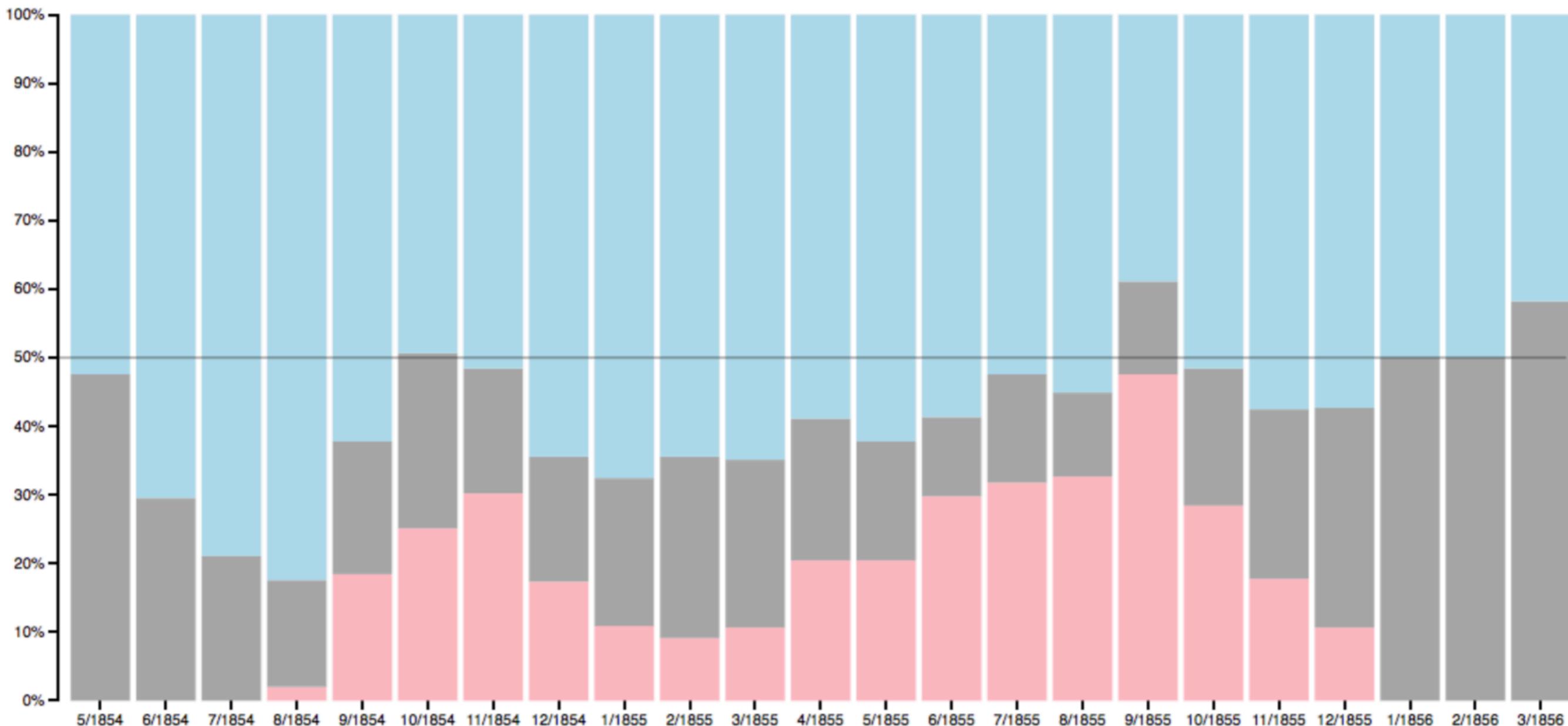
# Months in which deaths by wound exceeds deaths in the field?



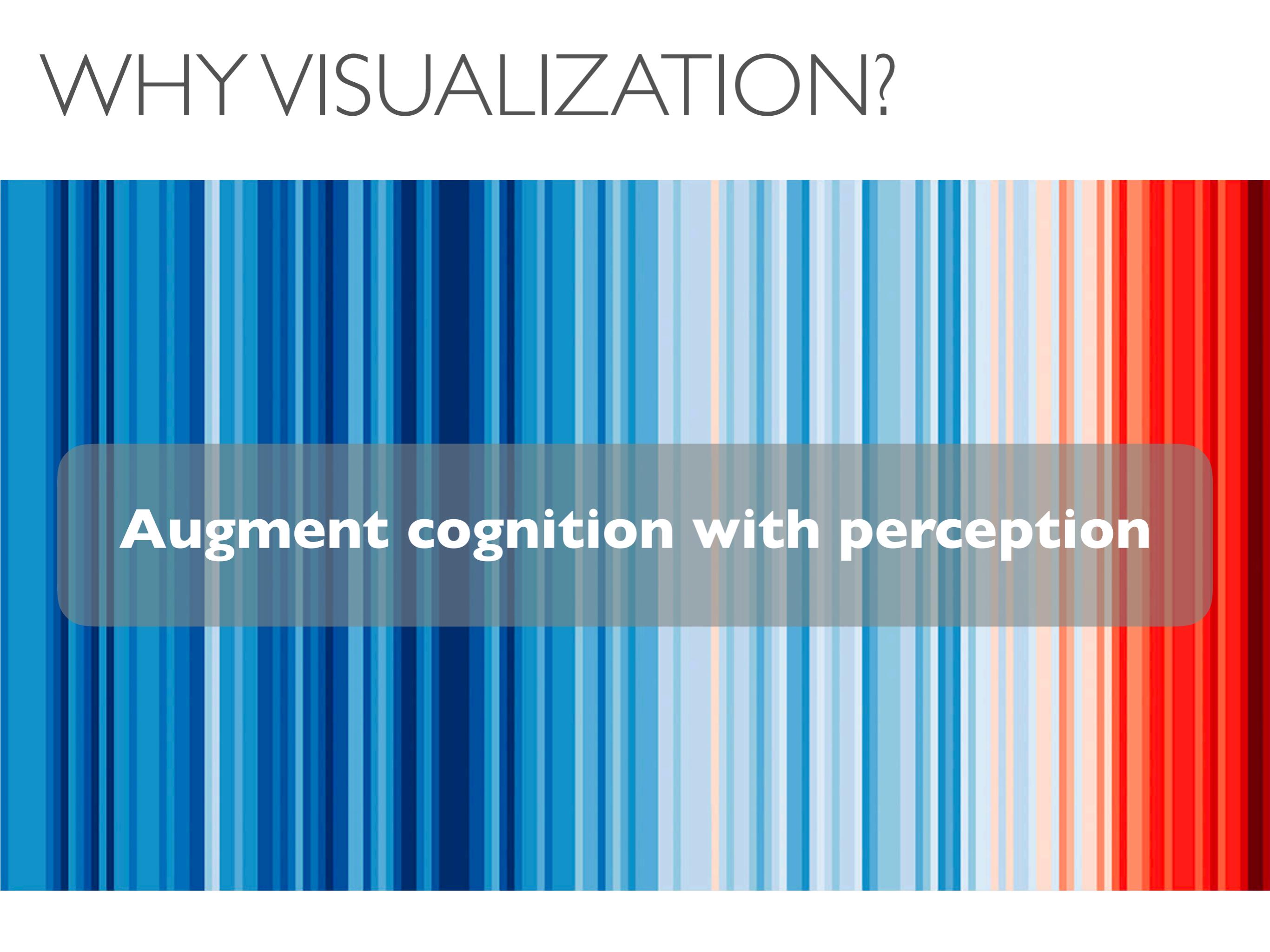
# Month with highest total casualty rate?



# Months in which % of deaths by disease was below 50%?



# WHY VISUALIZATION?



**Augment cognition with perception**

# Temperature Anomalies by Country Years 1880 - 2017

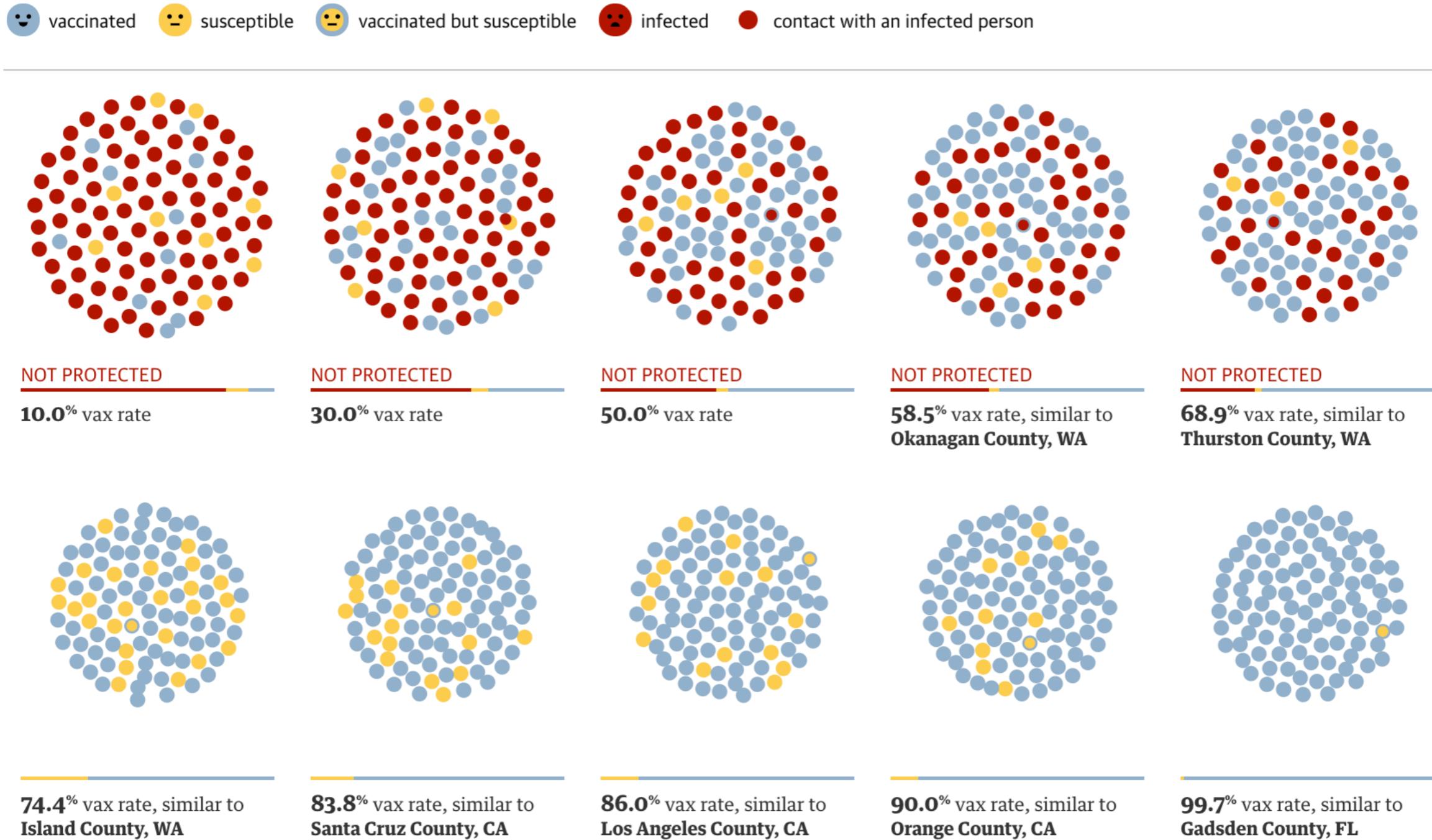
1880



Data Source:  
 NASA GISS, GISTEMP Land-Ocean Temperature Index (LOTI), ERSSTv5, 1200km smoothing  
<https://data.giss.nasa.gov/gistemp/>  
 Average of monthly temperature anomalies. GISTEMP base period 1951–1980.

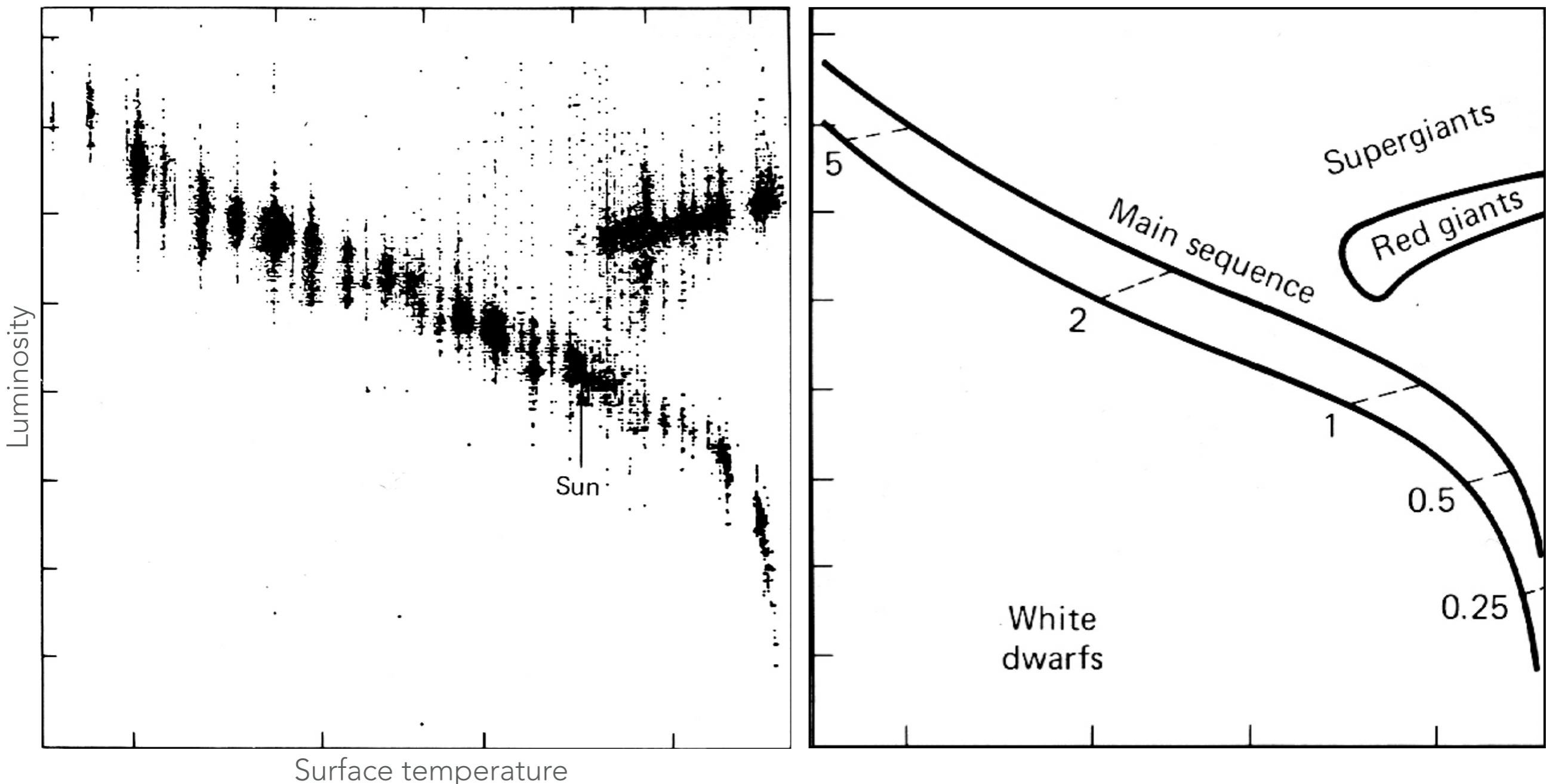
Video license: CC-BY-4.0  
 Antti Lipponen (@anttilip)

# Watch how the measles outbreak spreads when kids get vaccinated - and when they don't



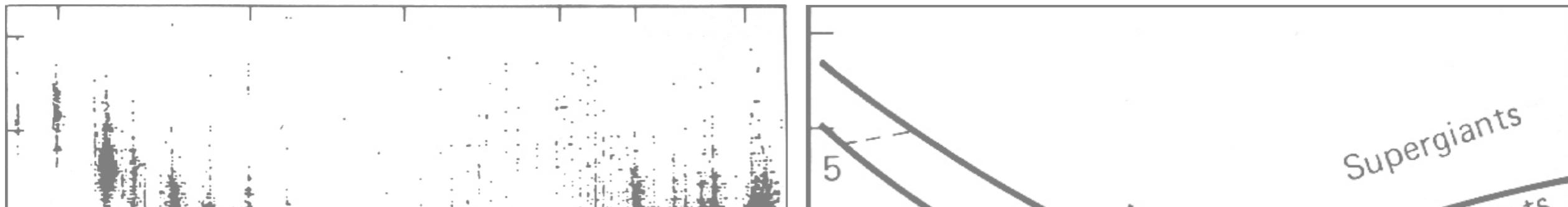
<https://www.theguardian.com/society/ng-interactive/2015/feb/05/-sp-watch-how-measles-outbreak-spreads-when-kids-get-vaccinated>

# PERCEPTION AND COGNITION



Hertzsprung Russell Diagram and its interpretation

# PERCEPTION AND COGNITION



*“Why should we be interested in visualization? Because the human visual system is a pattern seeker of enormous power and subtlety. The eye and the visual cortex of the brain form a **massively parallel processor** that provides **the highest-bandwidth channel** into human cognitive centers.”*

Ware, Colin. Information visualization: perception for design. Elsevier, 2013.



## Hertzsprung Russell Diagram and its interpretation

[Source: J.D. Fekete et al. “The Value of Information Visualization”, 2008]

142416496357598475921765968474891728482  
285958819829450968504850695847612124044  
074674898985171495969124567659608020860  
608365416496457590643980479248576960781  
285960799918712845268101495969124567781  
874241649645757659608149596912456701285  
960799164964575127879918712845298496912  
223591649645759588198250963576596080596

142416496357598475921765968474891728482  
285958819829450968504850695847612124044  
074674898985171495969124567659608020860  
608**3**6541649645759064**3**980479248576960781  
285960799918712845268101495969124567781  
874241649645757659608149596912456701285  
960799164964575127879918712845298496912  
**2235**9164964575958819825096**3**576596080596

*This is your brain on visualization....*

## **PERCEPTION FUNDAMENTALS**

Visualization design is based on the mapping between data and visual representation

**Data**



**Graphics**

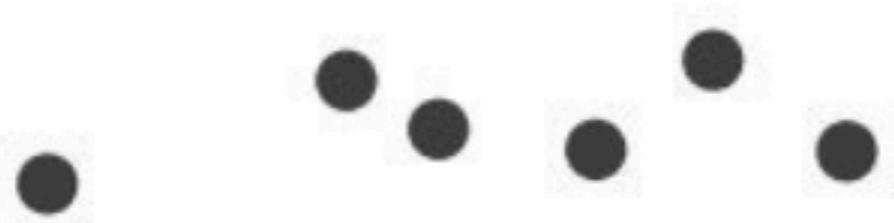
Which representations can I use to encode data?

Which representations are more suitable to ensure I'm conveying the right message?

# MARKS

Geometric primitives

→ Points



→ Lines

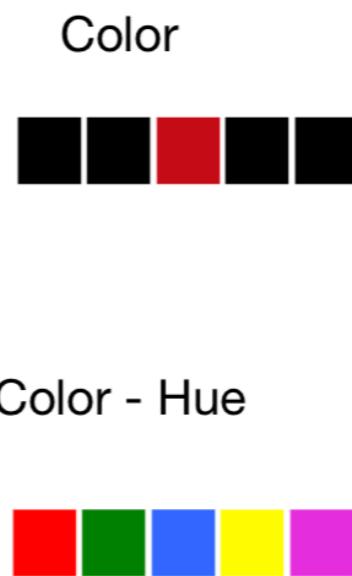
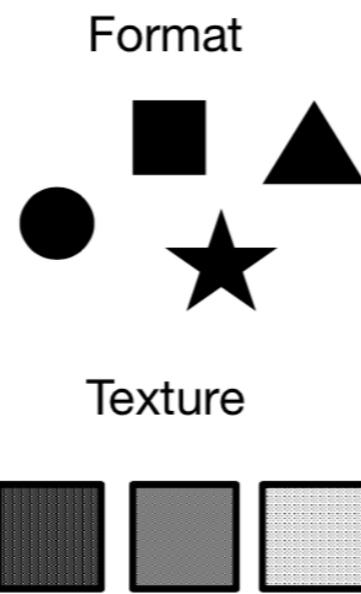
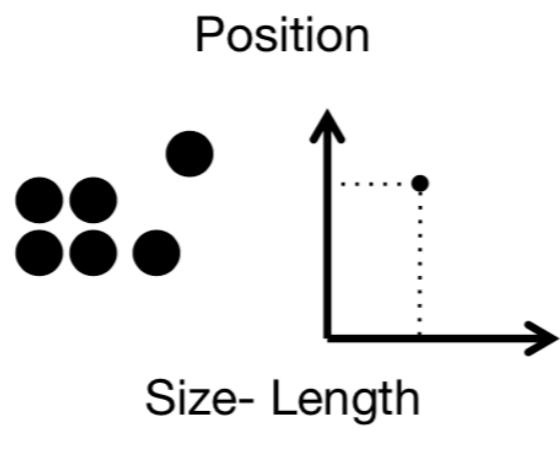


→ Areas



# CHANNELS (VISUAL VARIABLES)

Control appearance of marks



Size- Area



Orientation/Direction



Color - Saturation



Size- Volume



Angle

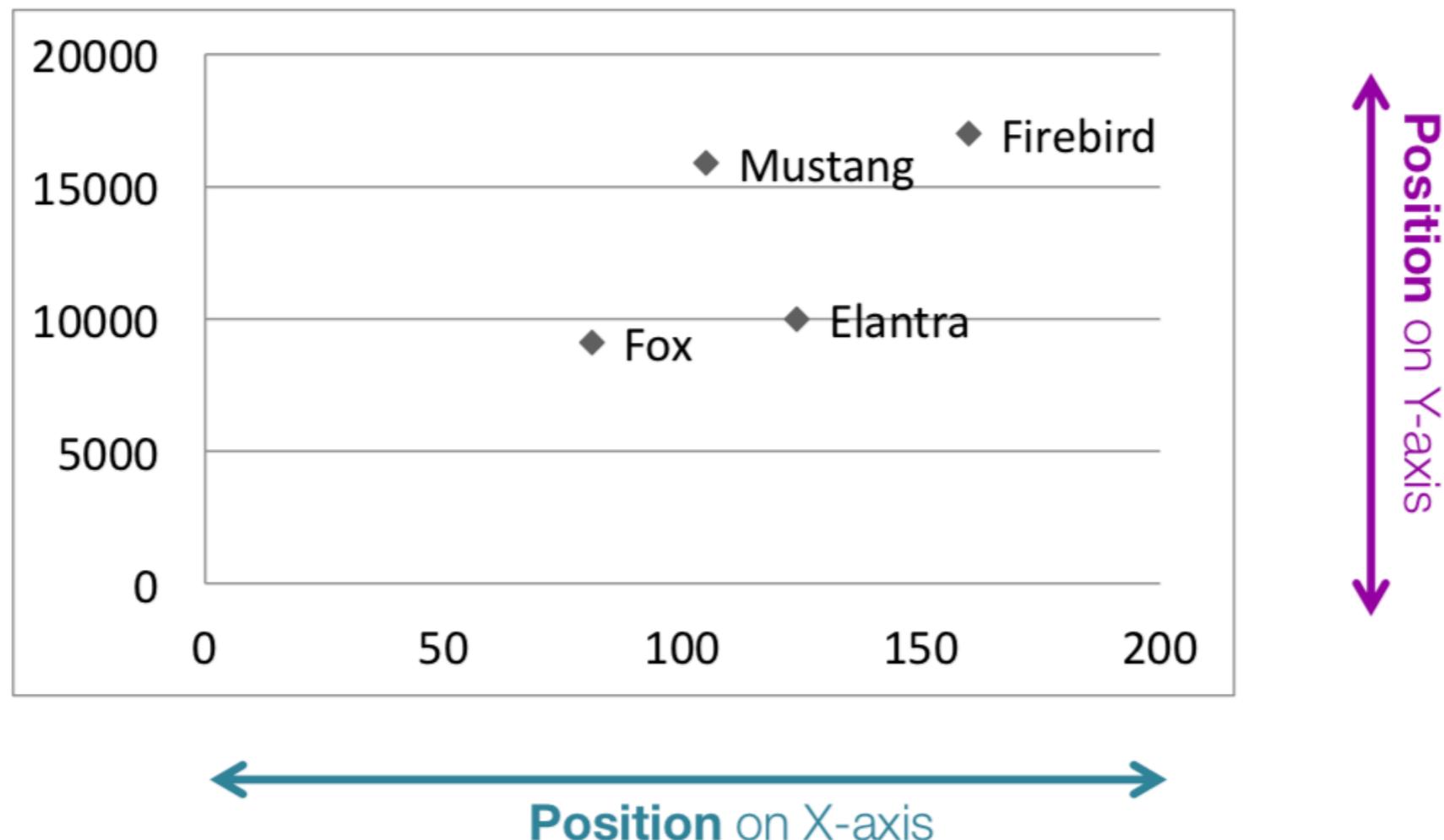


Color - Luminance

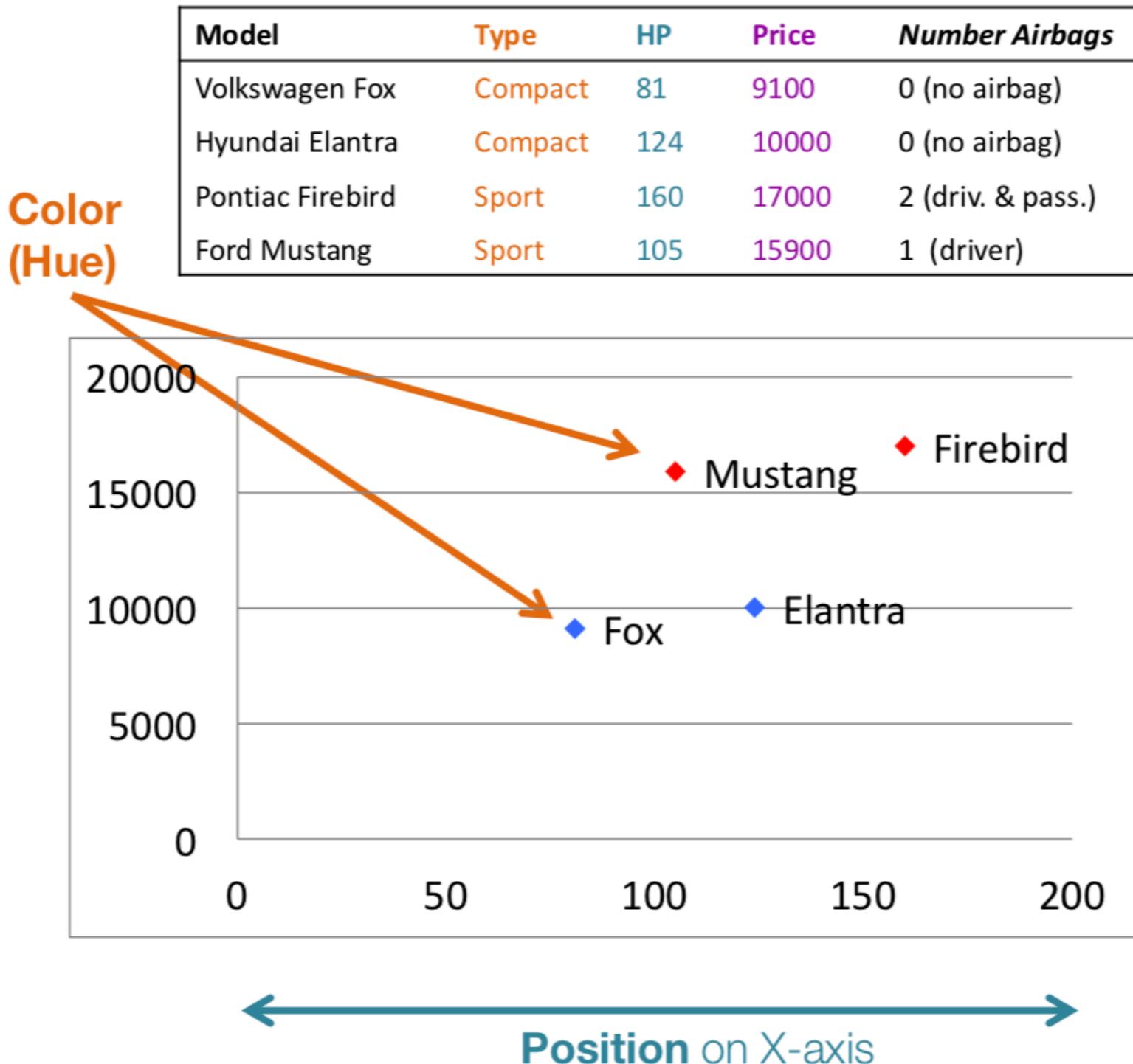


# MAPPING DATA TO VISUAL VARIABLES

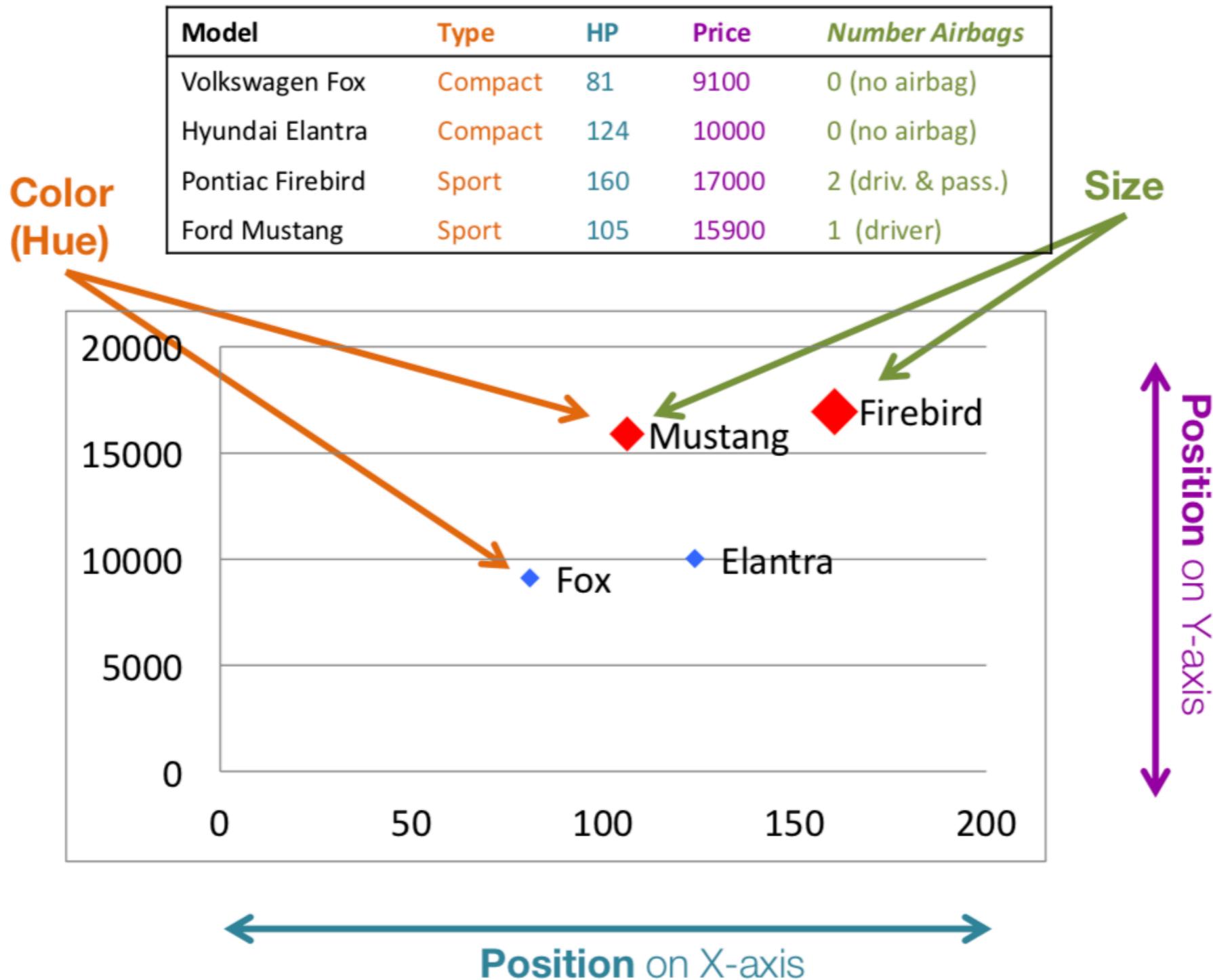
Model	Type	HP	Price	Number Airbags
Volkswagen Fox	Compact	81	9100	0 (no airbag)
Hyundai Elantra	Compact	124	10000	0 (no airbag)
Pontiac Firebird	Sport	160	17000	2 (driv. & pass.)
Ford Mustang	Sport	105	15900	1 (driver)



# MAPPING DATA TO VISUAL VARIABLES



# MAPPING DATA TO VISUAL VARIABLES



## Which representations can I use to encode data?

### **Expressiveness principle: use adequate/suitable data representations**

Encodings should convey all, and only, the information of associated attributes.

e.g. Ordinal data representation should convey “order”; similarly, “categorical data” should not be shown in a way that implies order.

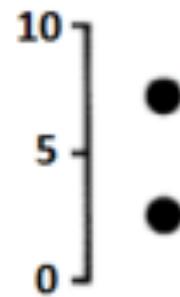
**Which representations are more suitable to ensure  
I'm conveying the right message?**

**Effectiveness principle:  
choosing the best representation to your data**

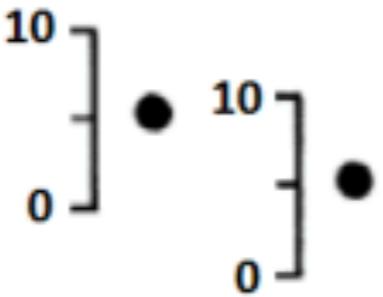
Importance of attributes should match the “saliency” of the channel;

Most important attributes should be encoded using the most effective and noticeable channels.

**Position Common Scale**



**Position Non-Aligned Scale**



**Length**



**Direction**



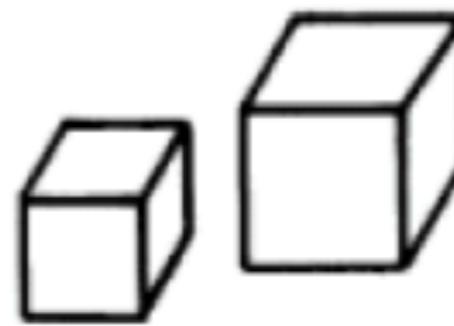
**Angle**



**Area**



**Volume**



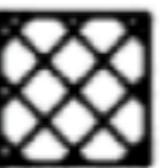
**Curvature**



**Shape**



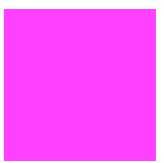
**Shading**



**Color saturation**



**Hue**



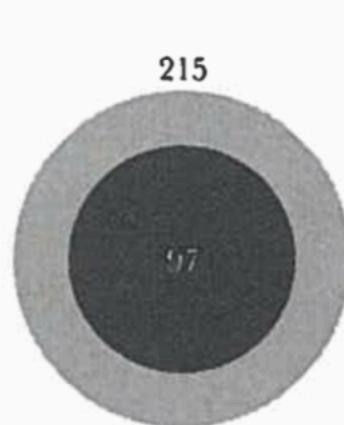
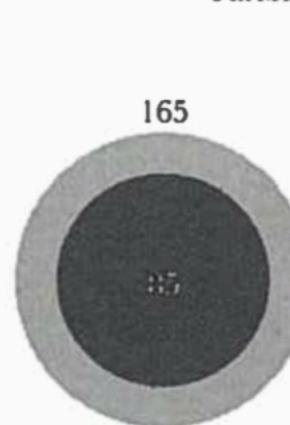
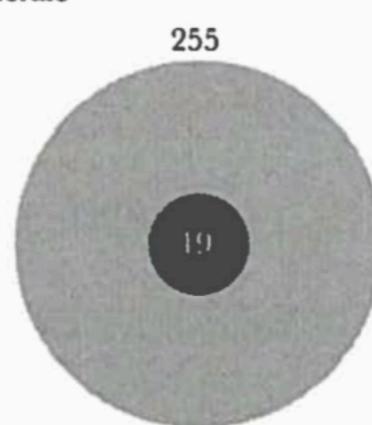
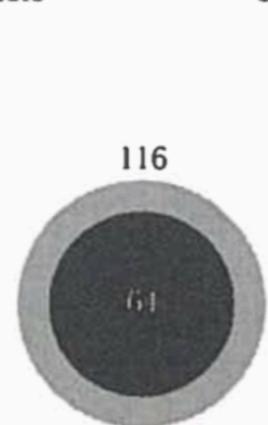
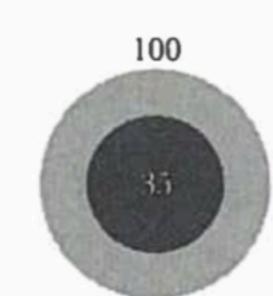
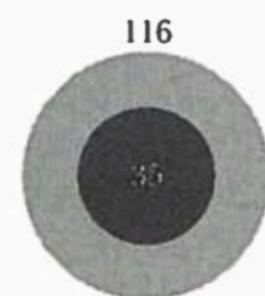
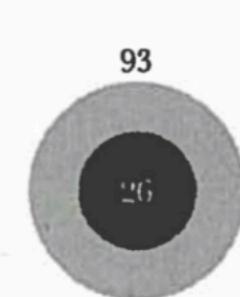
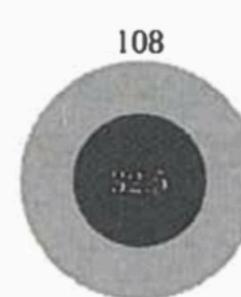
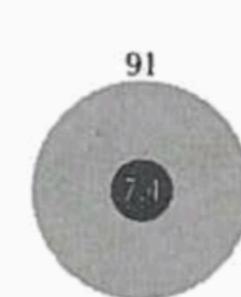
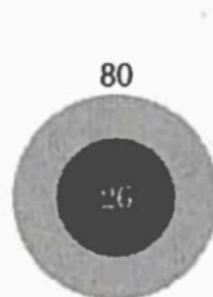
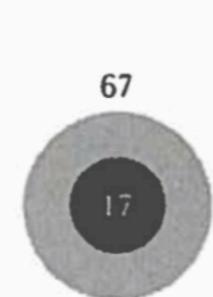
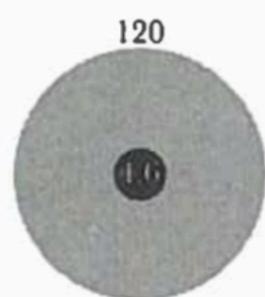
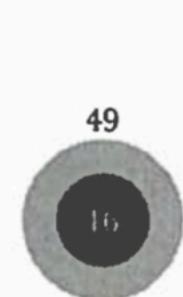
## Market Capitalization of the World's Biggest Banks

In billions of dollars

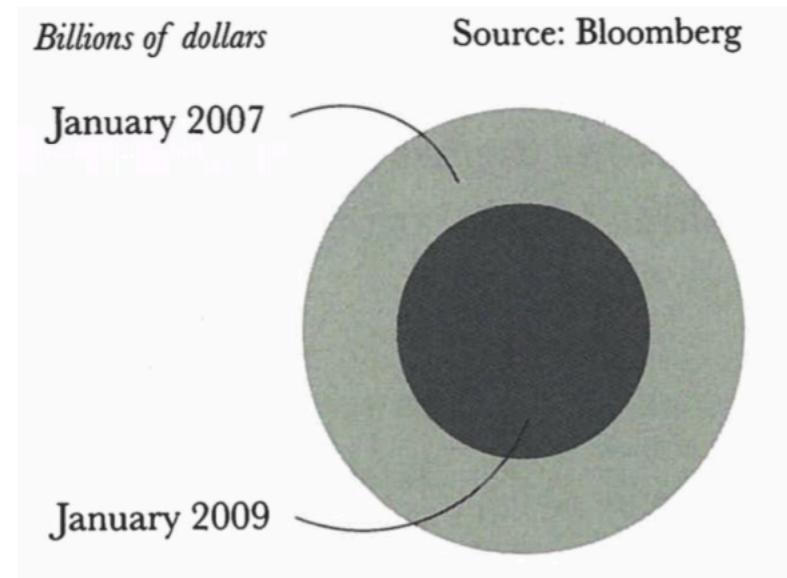
January 2007

January 2009

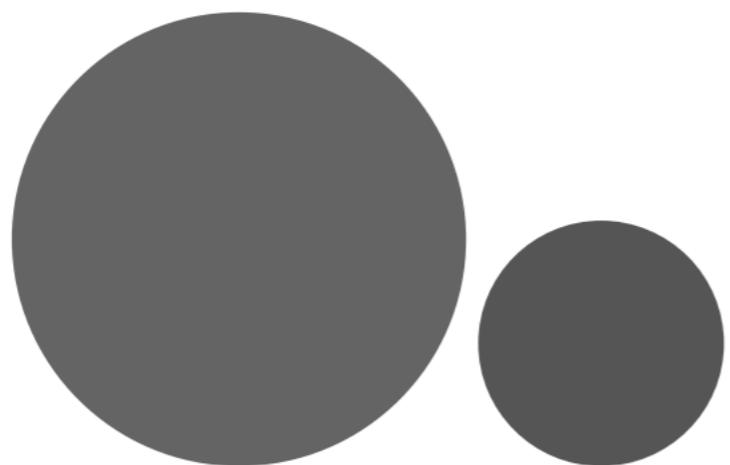
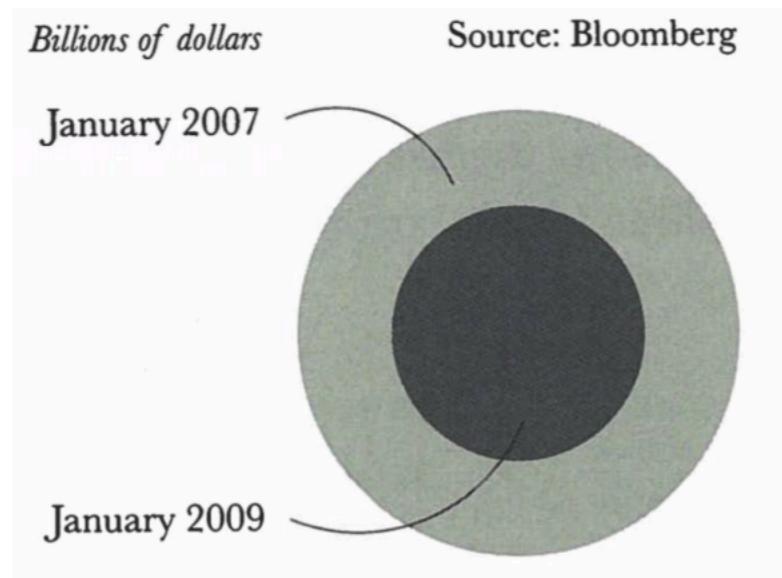
Source: Bloomberg



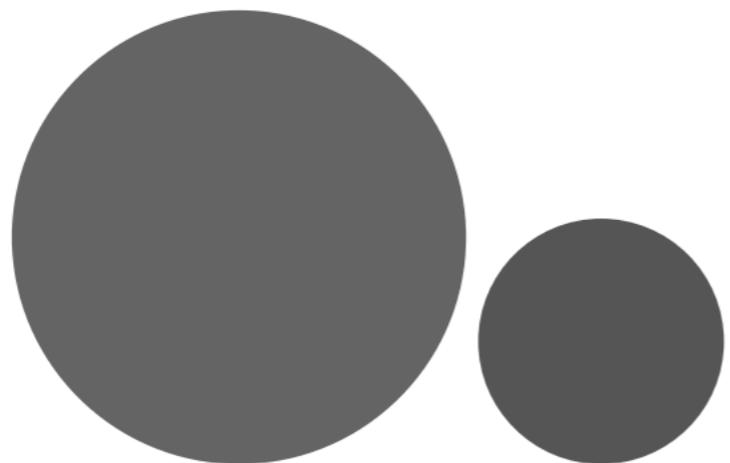
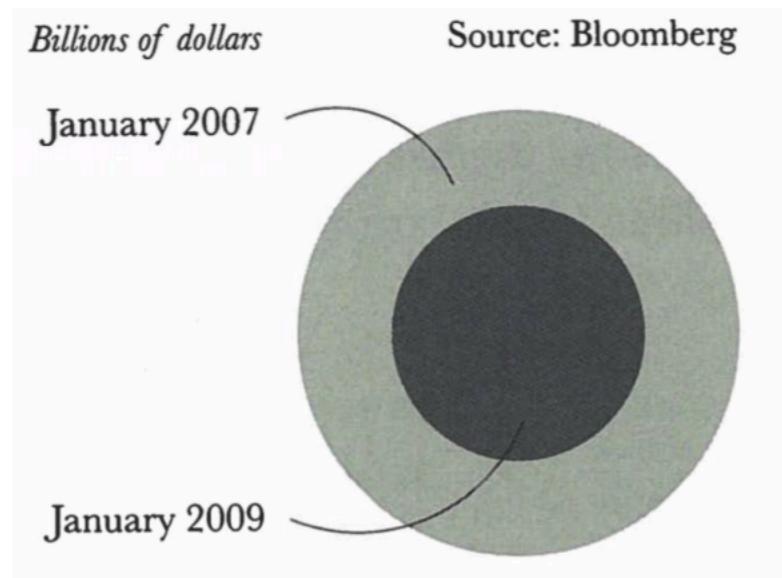
# How to decode bubbles?



# How to decode bubbles?



# How to decode bubbles?



a)



b)



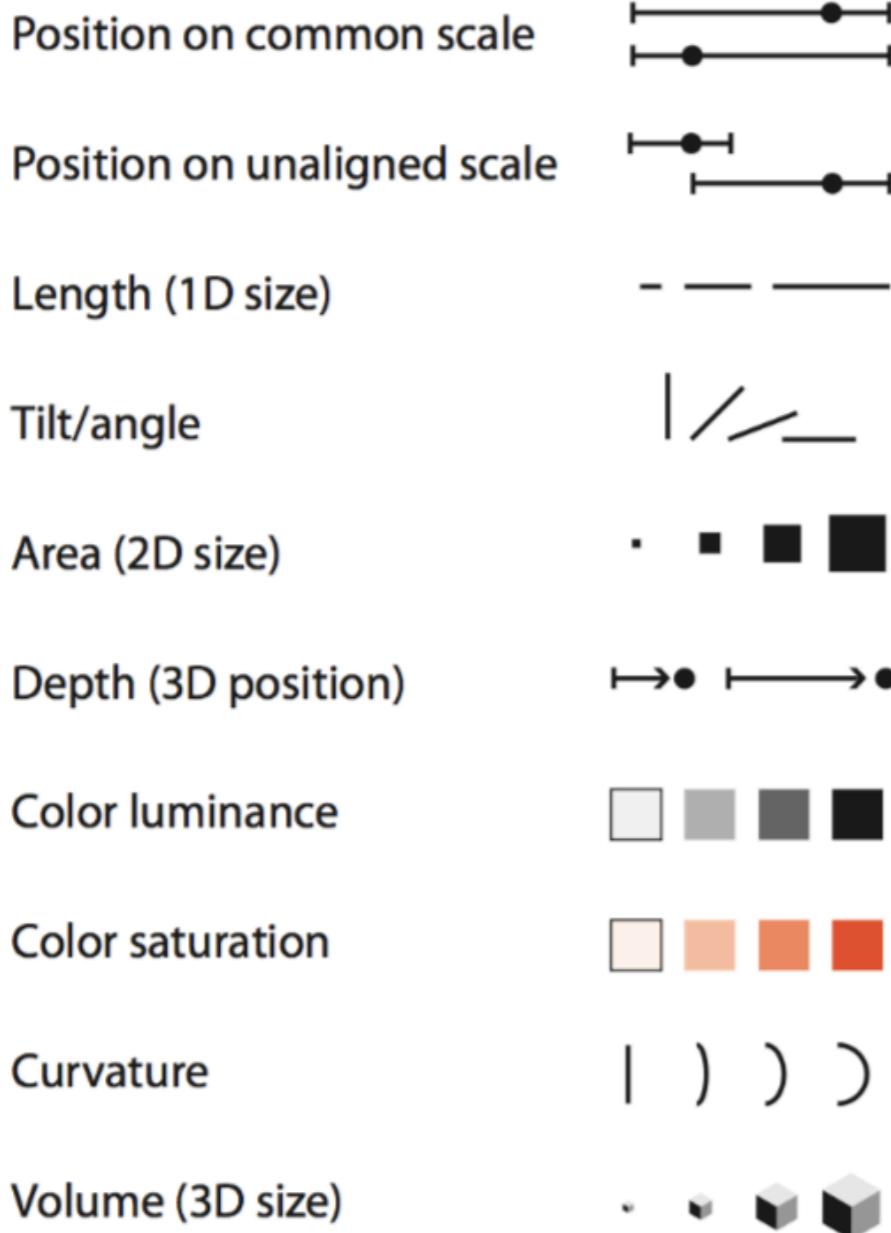
c)



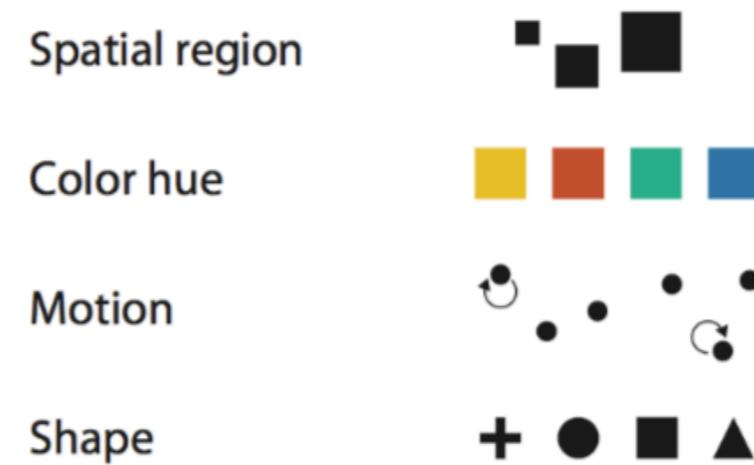
# EFFECTIVENESS PRINCIPLE

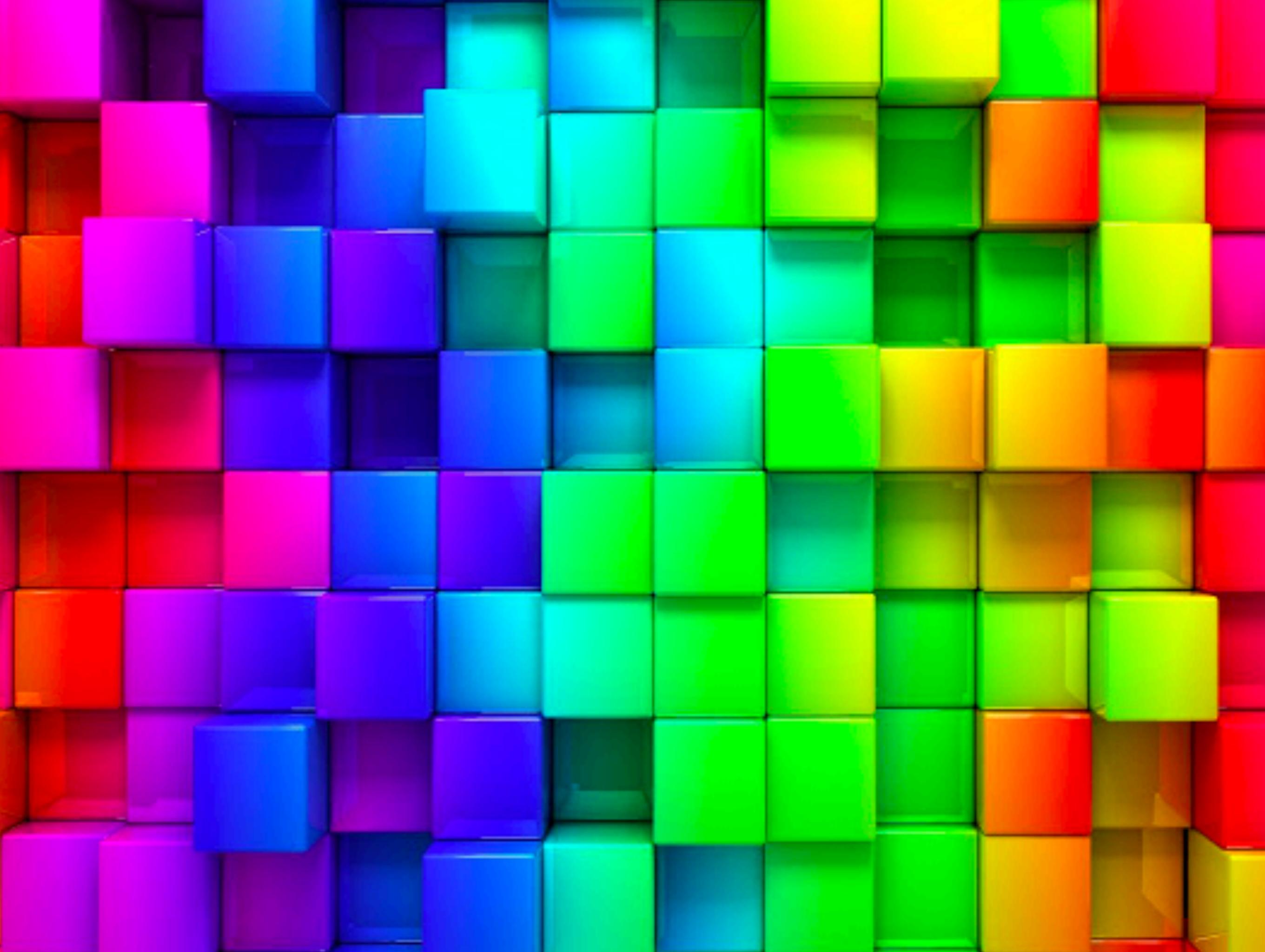
Some variables are perceptually better than others

## ④ Magnitude Channels: Ordered Attributes

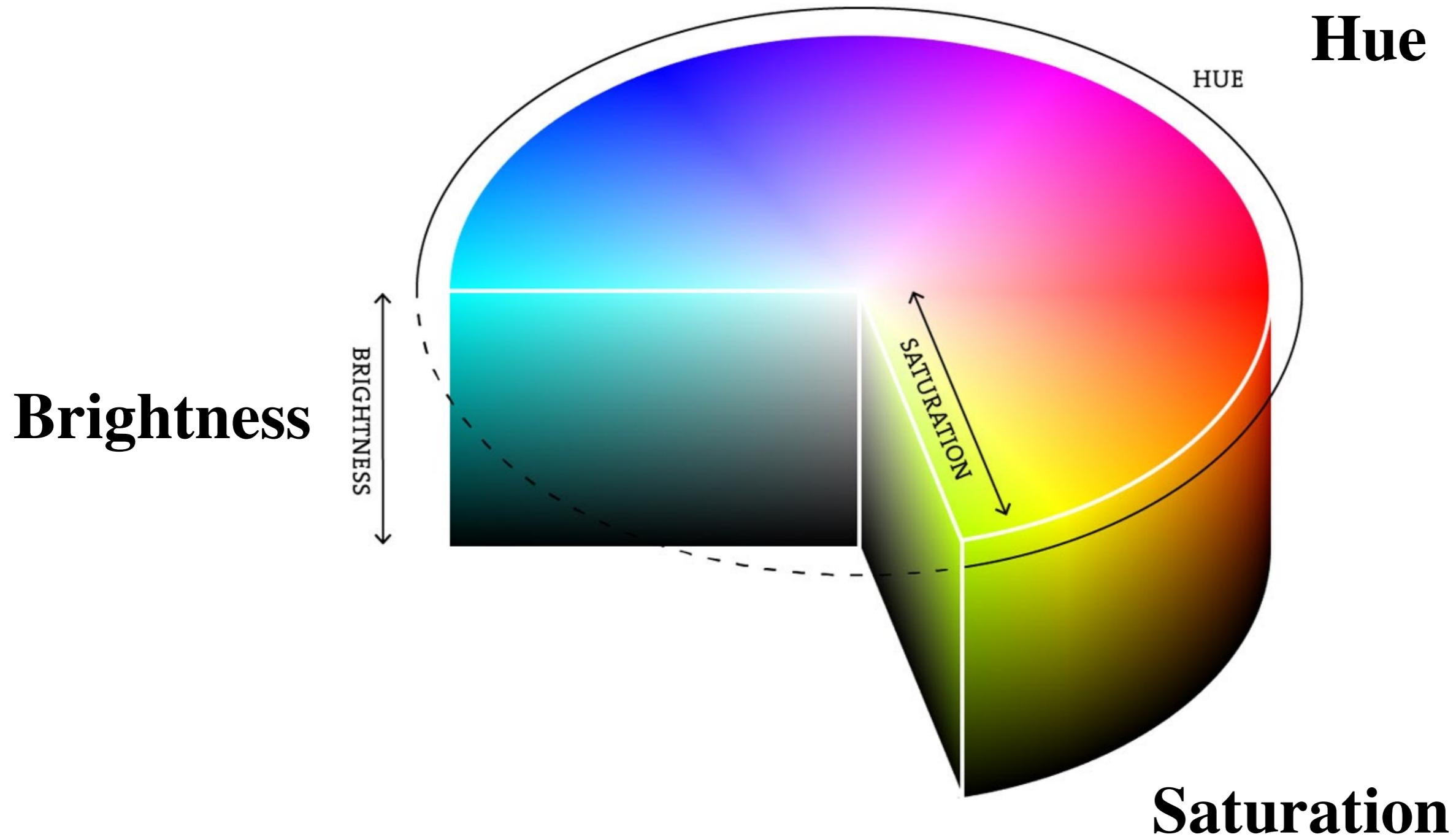


## ④ Identity Channels: Categorical Attributes

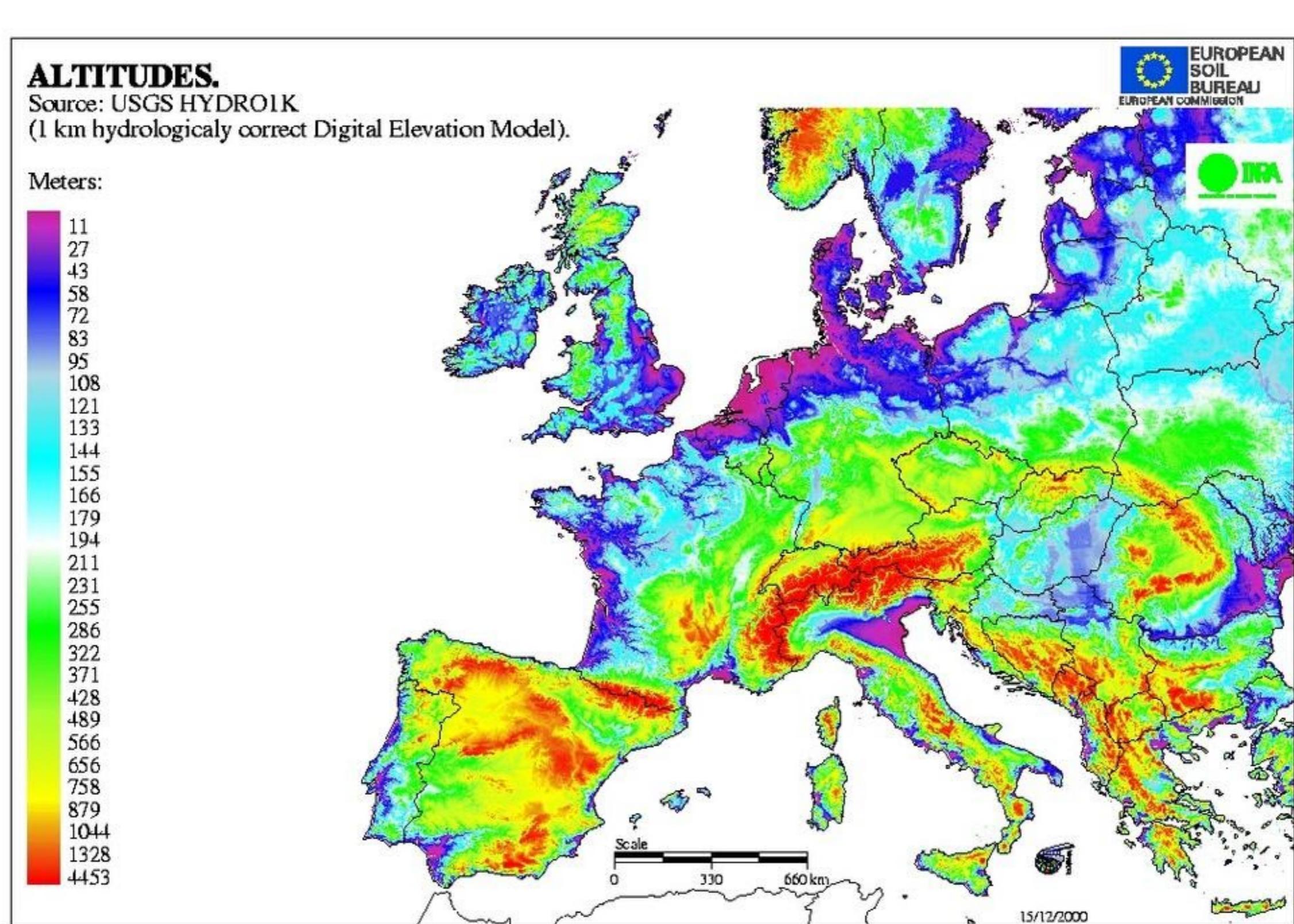




# HSB MODEL



# RAINBOW COLOR MAP

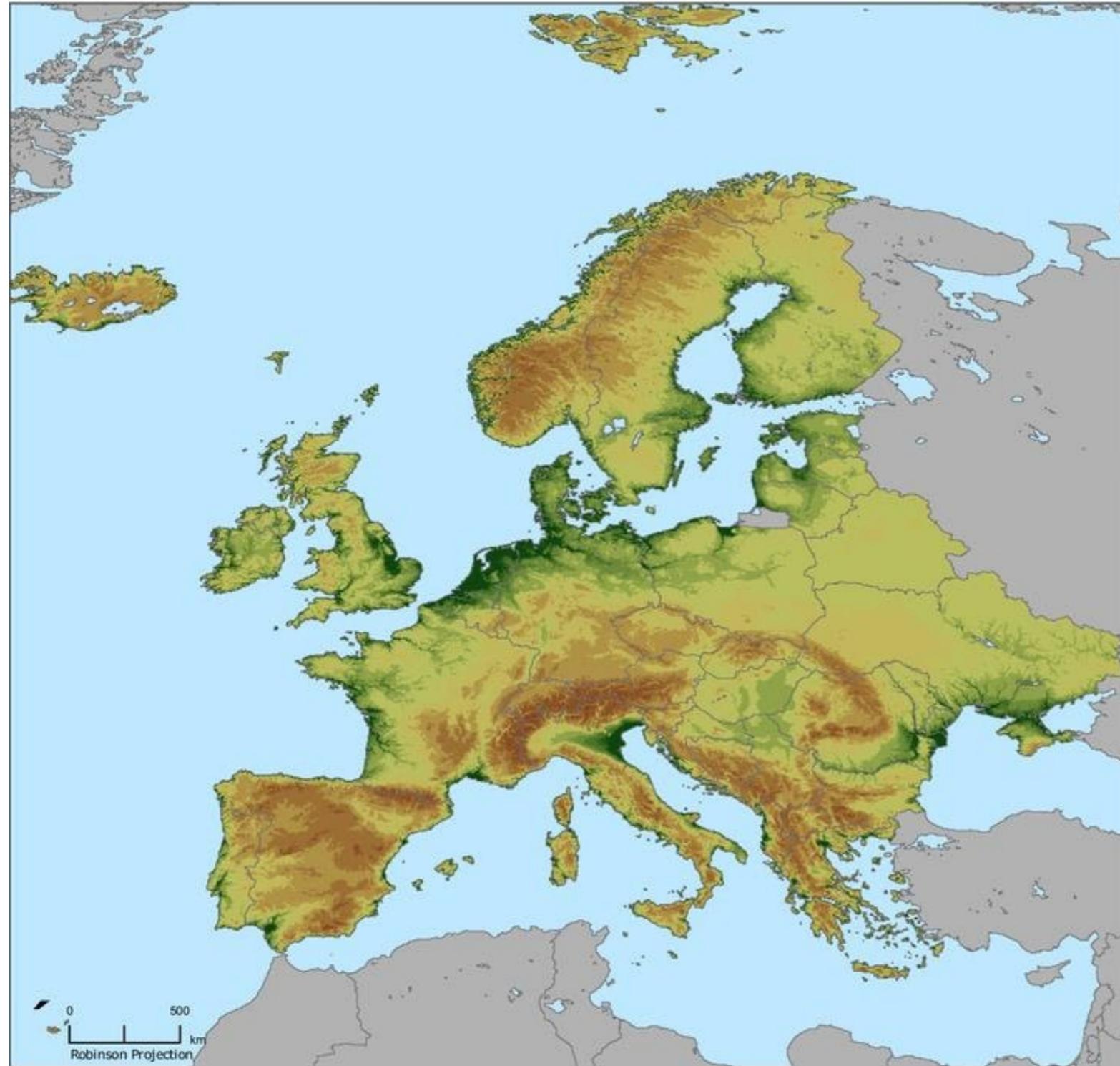


# RAINBOW COLOR MAP

## Elevation Zones

Europe

PLACE II  
Population, Landscape, and Climate Estimates



## Rainbow Scale Considerations

Map 1

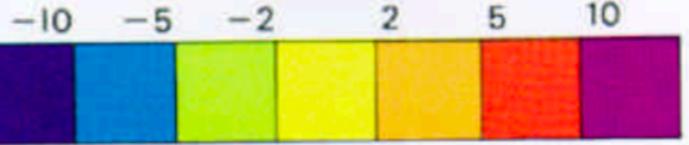


1

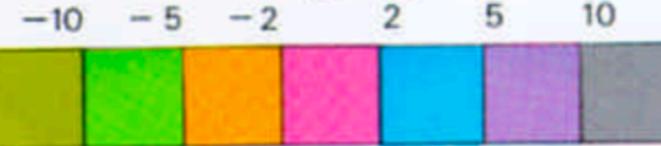
2



4



10



# COLOR SCALES



diverging



sequential



qualitative

# COLOR (MIS)USE

Country Level Sales Rank Top 5 Drugs

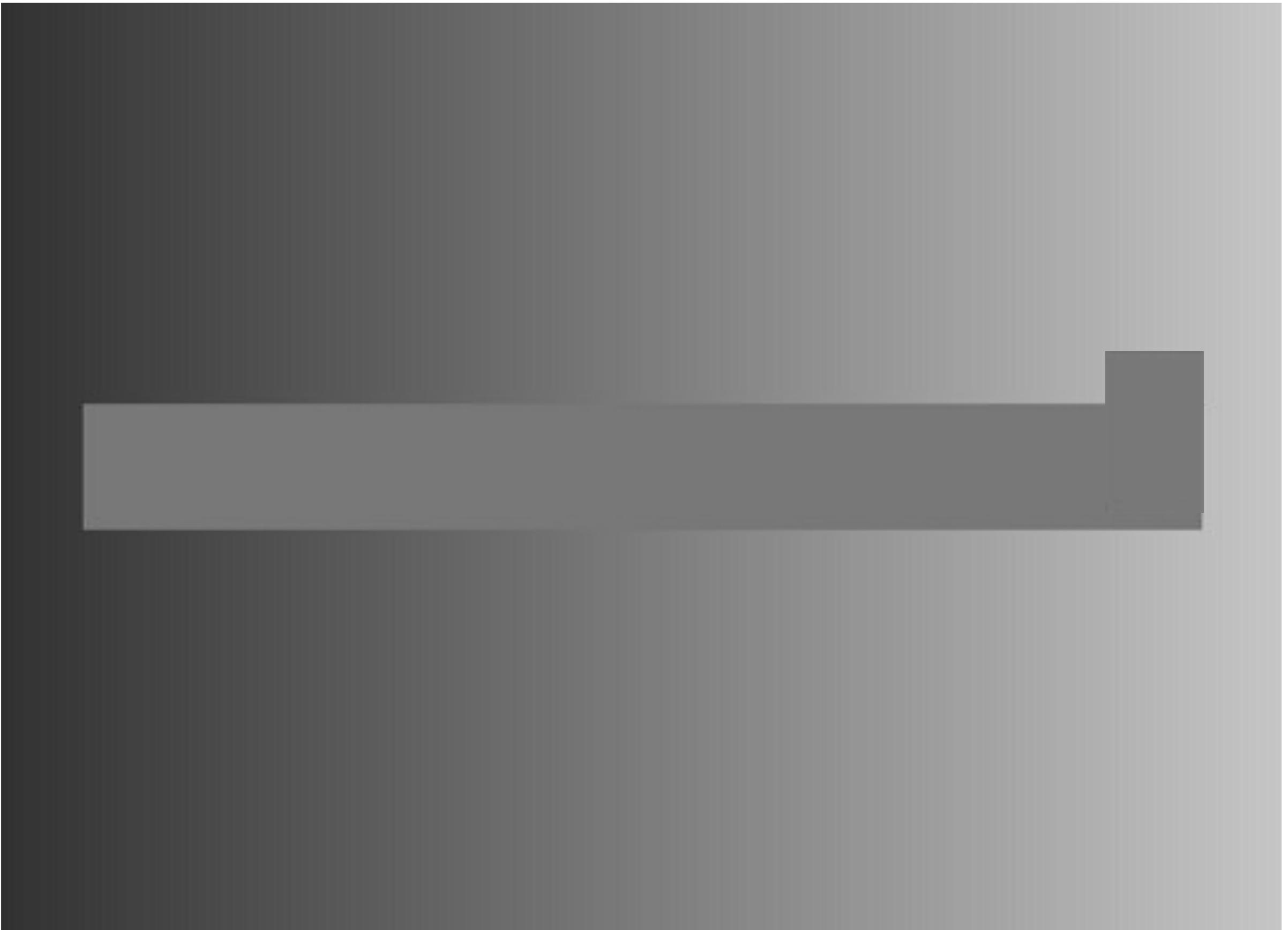
Rainbow distribution in color indicates sales rank in given country from #1 (red) to #10 or higher (dark purple)

Country	A	B	C	D	E
AUS	1	2	3	6	7
BRA	1	3	4	5	6
CAN	2	3	6	12	8
CHI	1	2	8	4	7
FRA	3	2	4	8	10
GER	3	1	6	5	4
IND	4	1	8	10	5
ITA	2	4	10	9	8
MEX	1	5	4	6	3
RUS	4	3	7	9	12
SPA	2	3	4	5	11
TUR	7	2	3	4	8
UK	1	2	3	6	7
US	1	2	4	3	5

Top 5 drugs: country-level sales rank

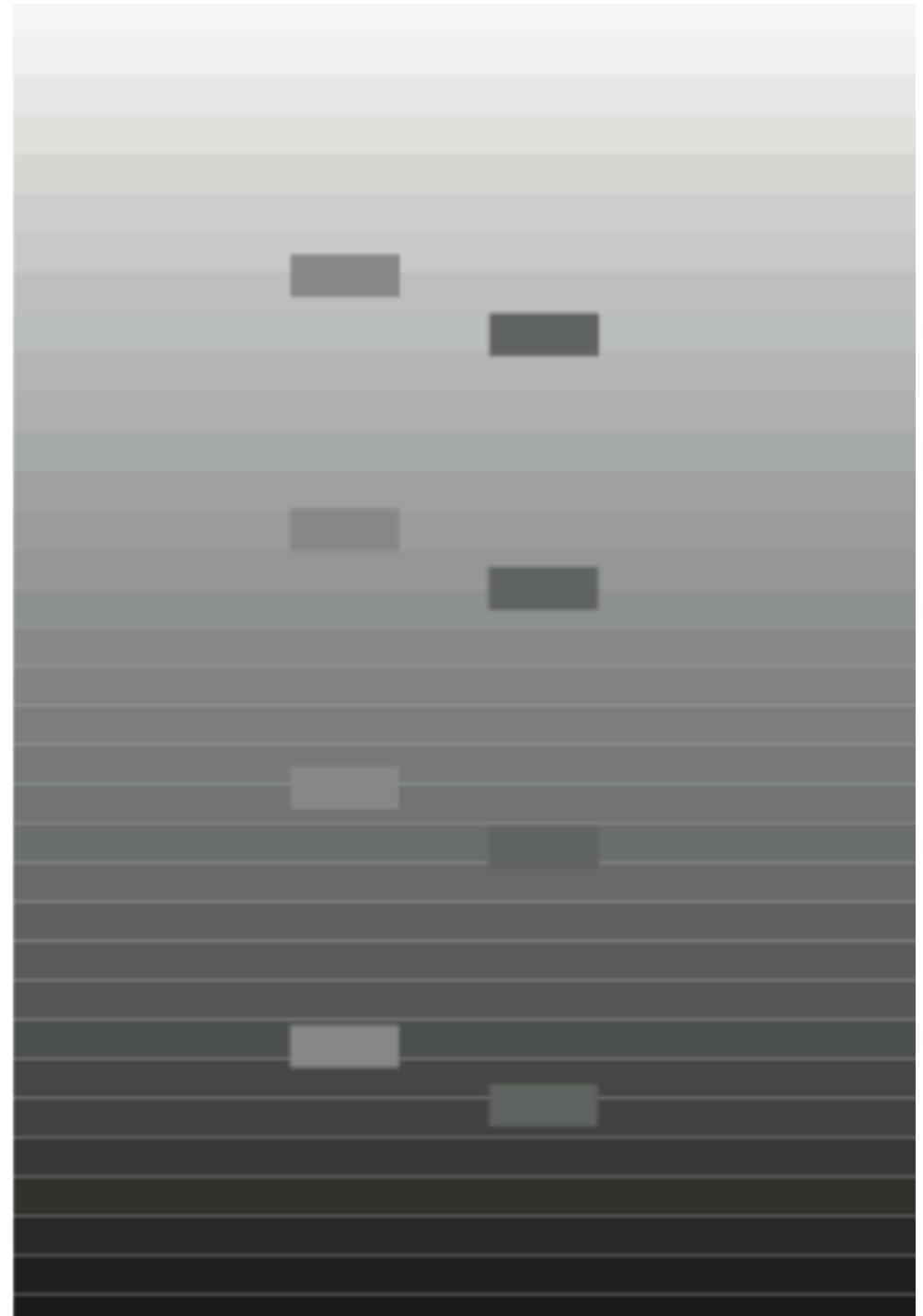
RANK	1	2	3	4	5+
COUNTRY   DRUG	A	B	C	D	E
Australia	1	2	3	6	7
Brazil	1	3	4	5	6
Canada	2	3	6	12	8
China	1	2	8	4	7
France	3	2	4	8	10
Germany	3	1	6	5	4
India	4	1	8	10	5
Italy	2	4	10	9	8
Mexico	1	5	4	6	3
Russia	4	3	7	9	12
Spain	2	3	4	5	11
Turkey	7	2	3	4	8
United Kingdom	1	2	3	6	7
United States	1	2	4	3	5

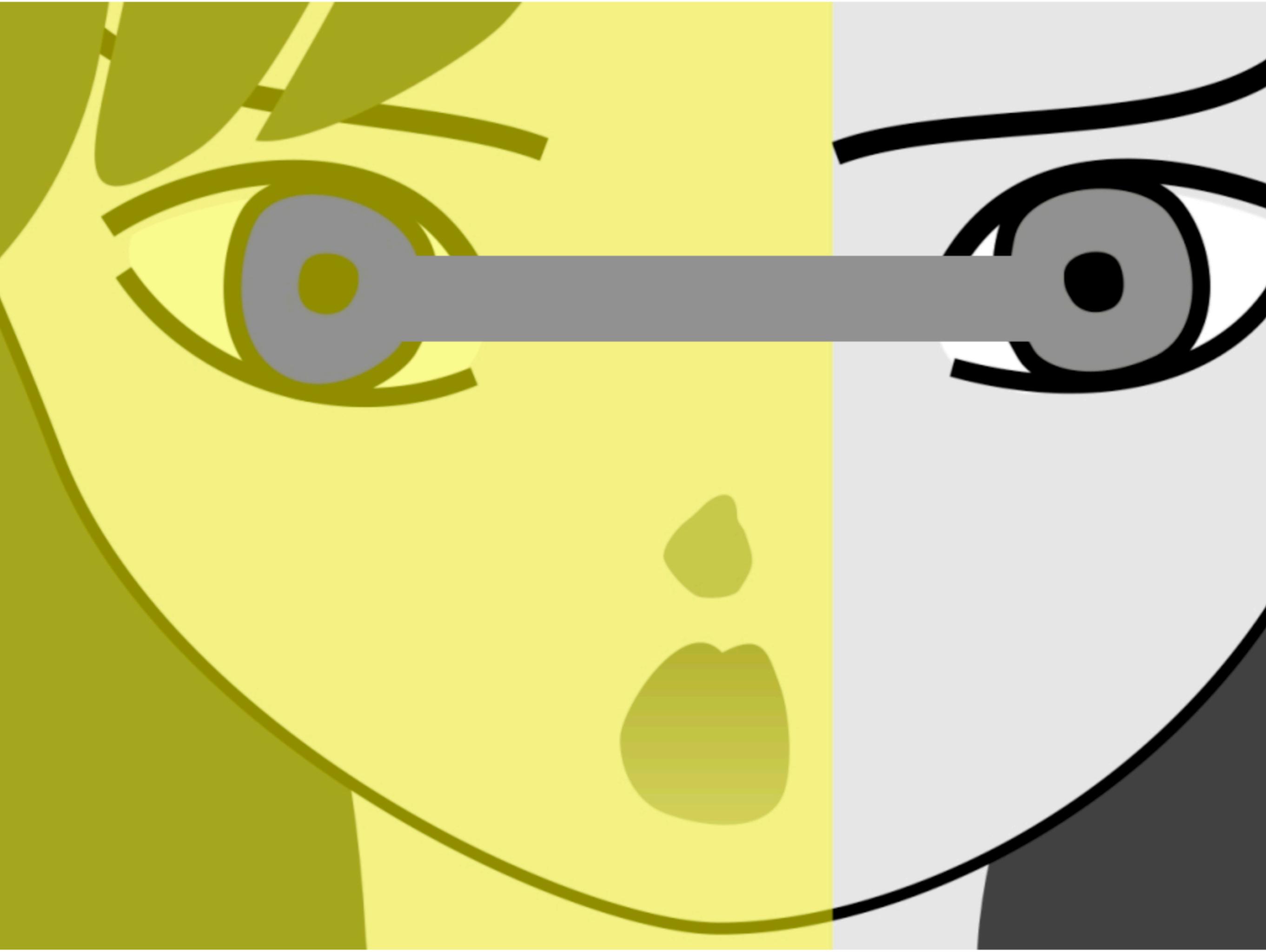
FIGURE 4.15 Use color sparingly



# COLOR PERCEPTION IS LOCAL

- Luminance perception is not absolute, but contrast relative ... and local!
- Discern objects(foreground) from background
- Gray color scales are not recommended for more than 5 segments... too “imprecise” to convey quantitative data



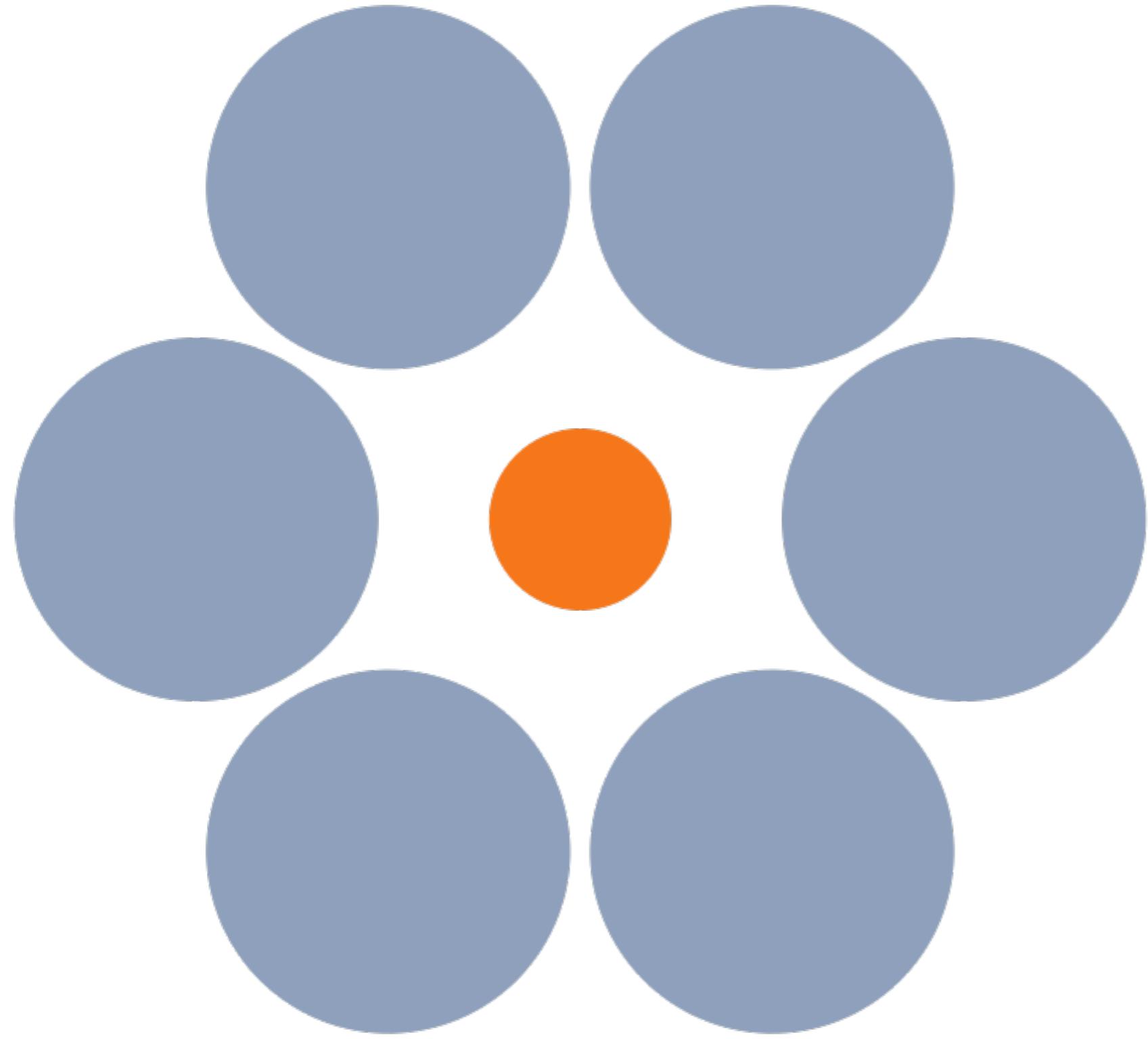


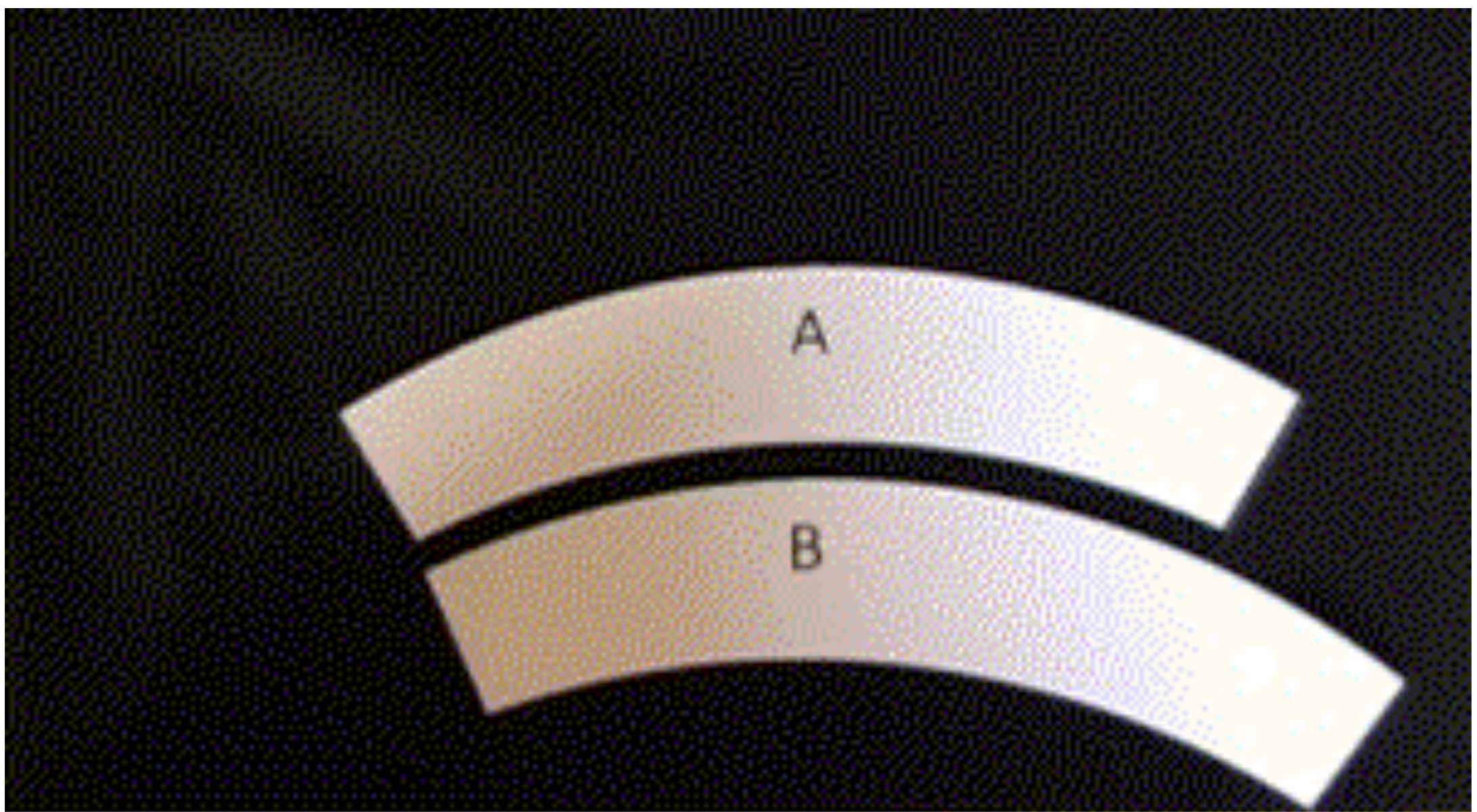


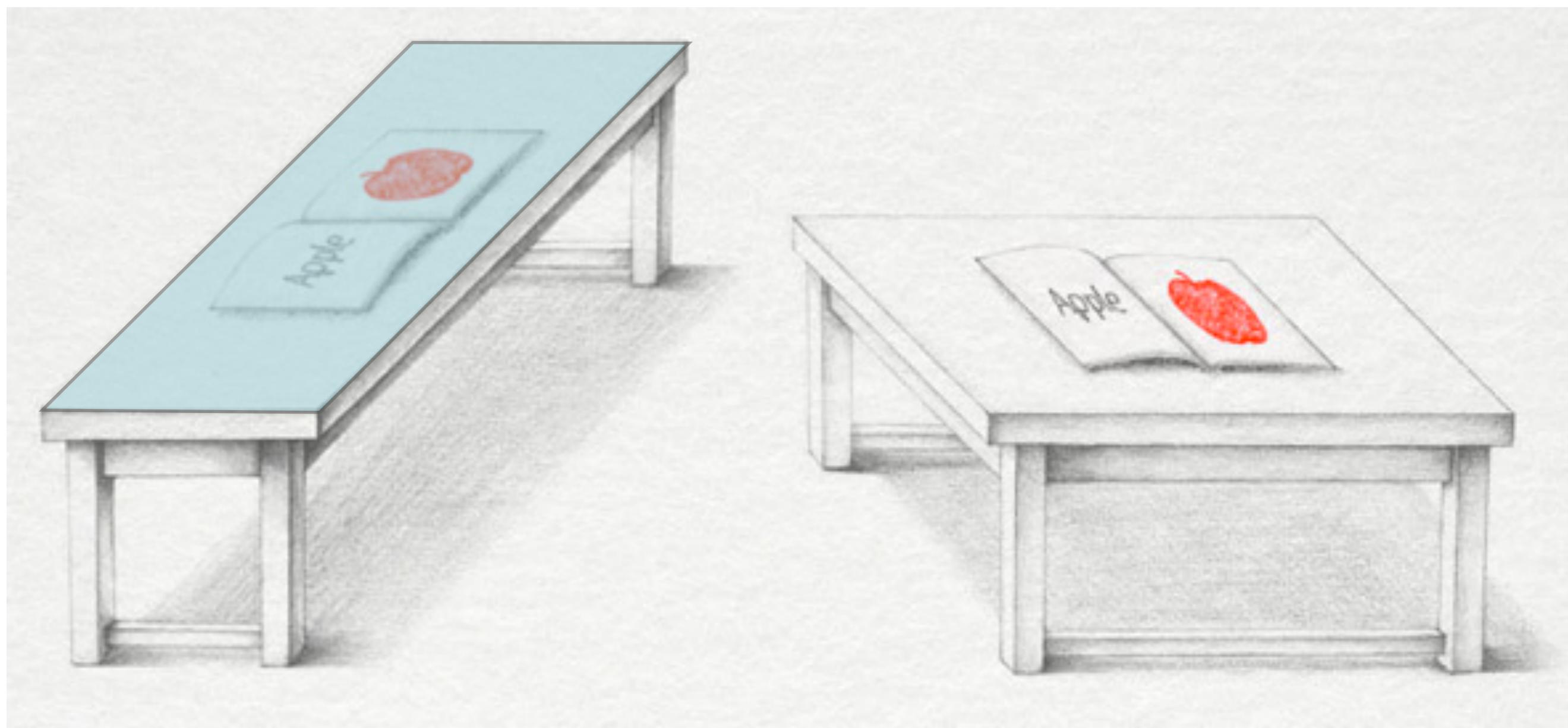


length?

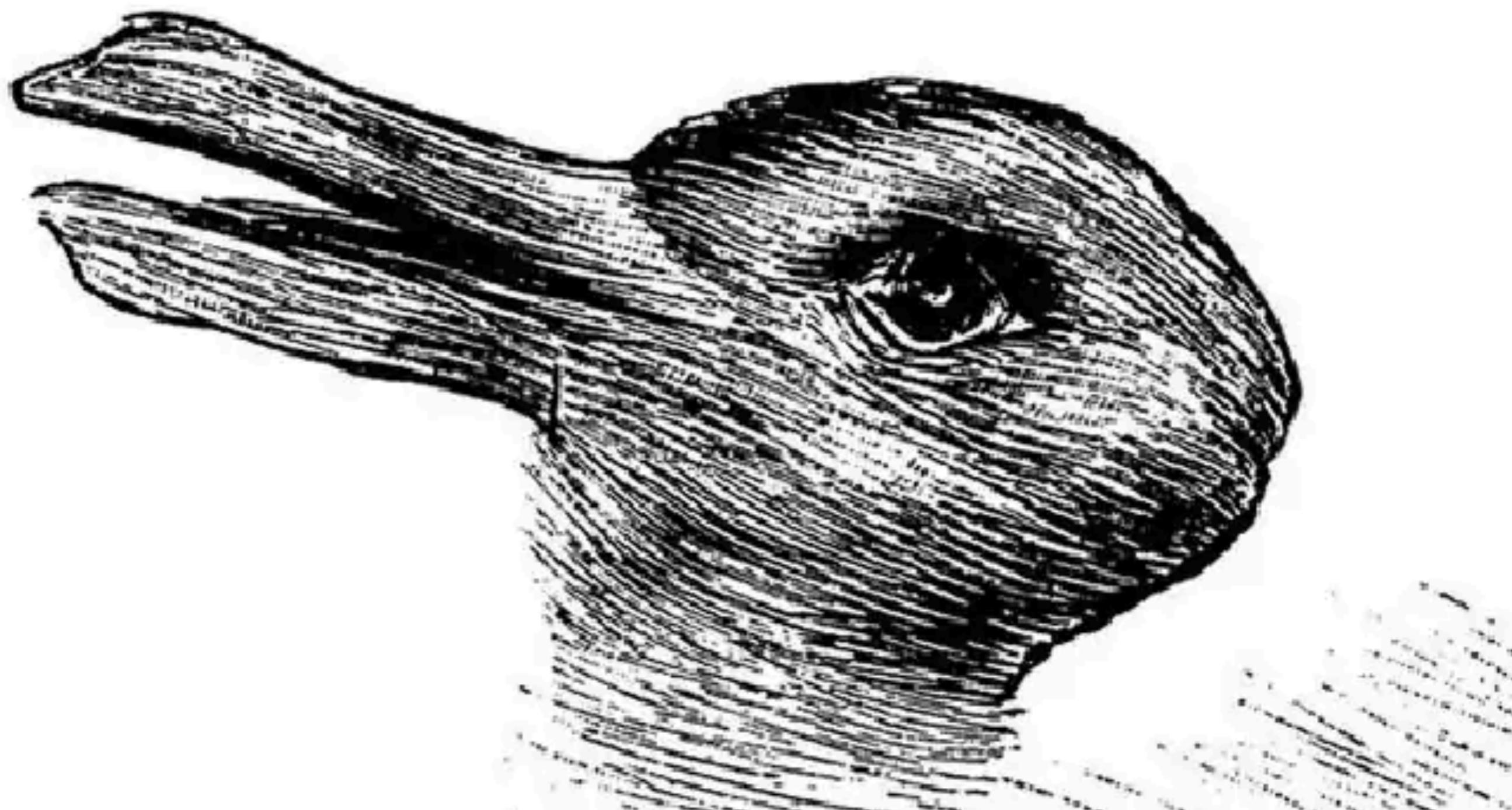
SIZE?

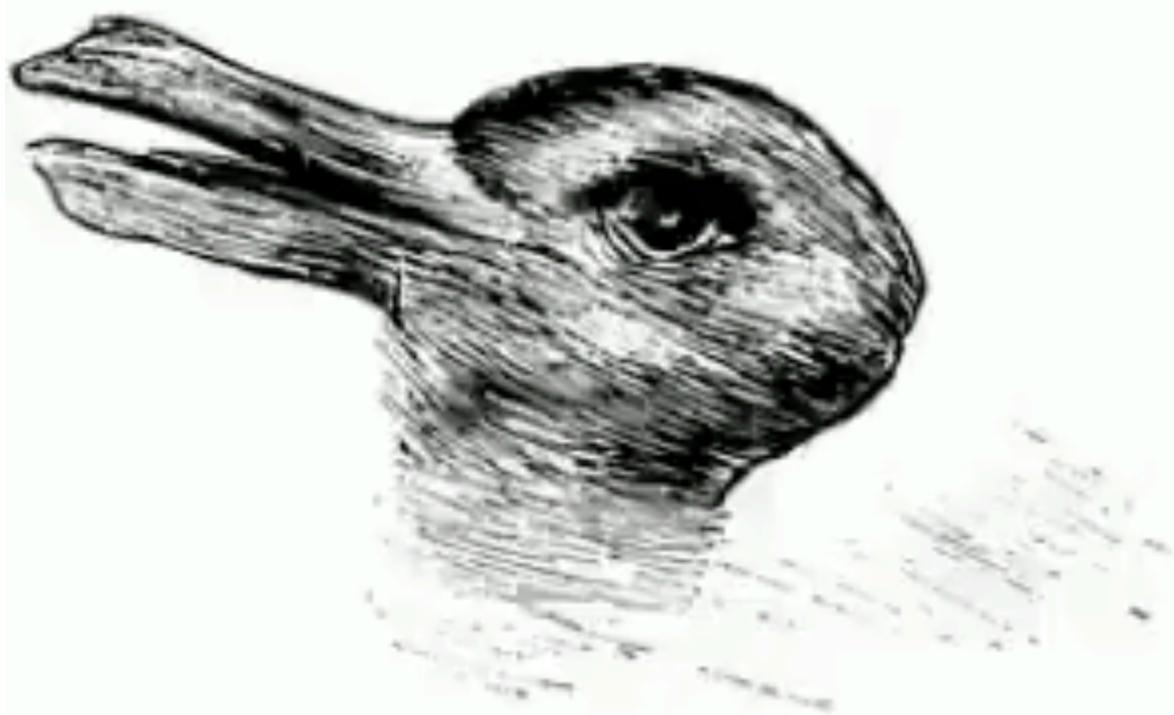




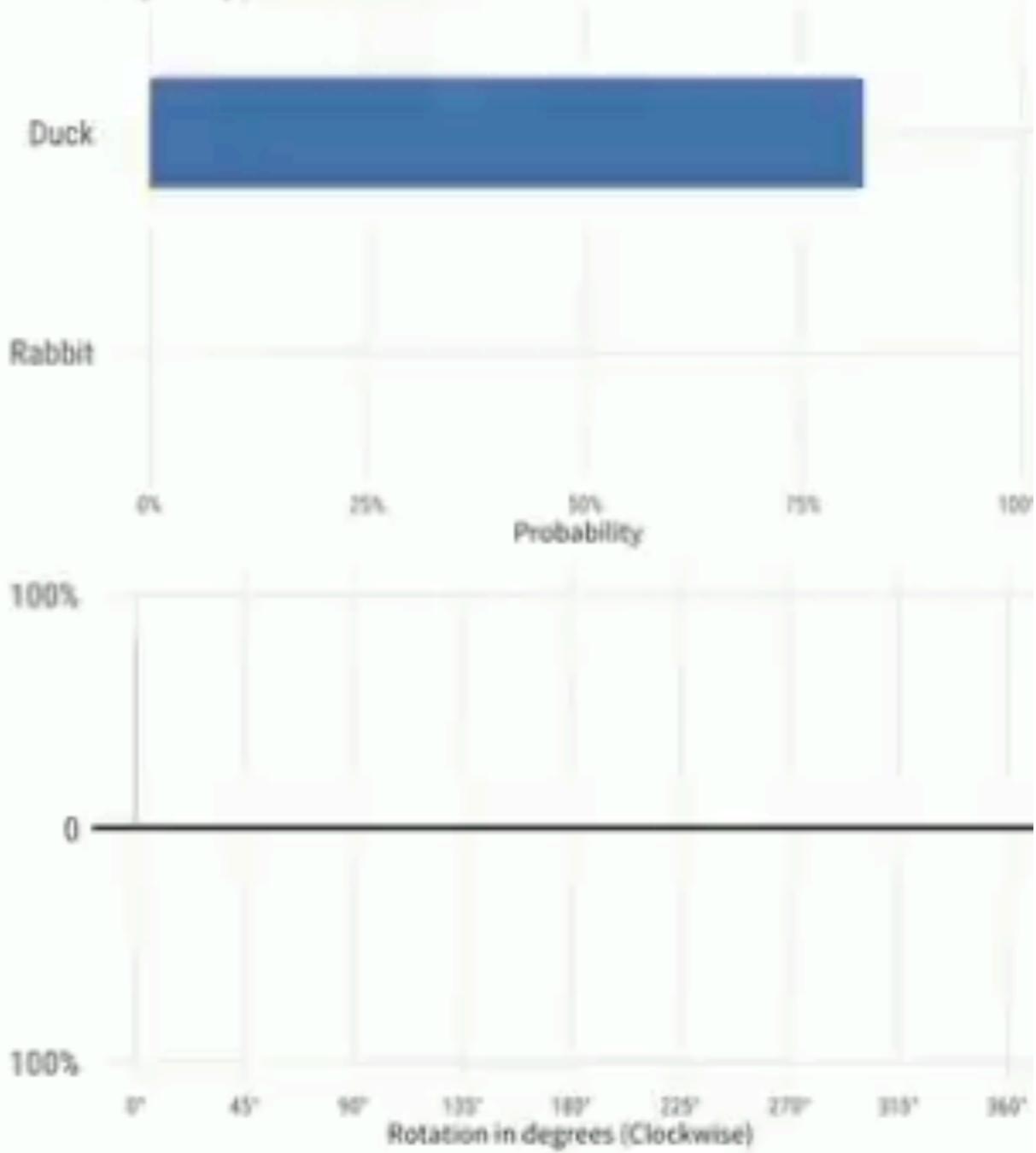


# DECEPTION

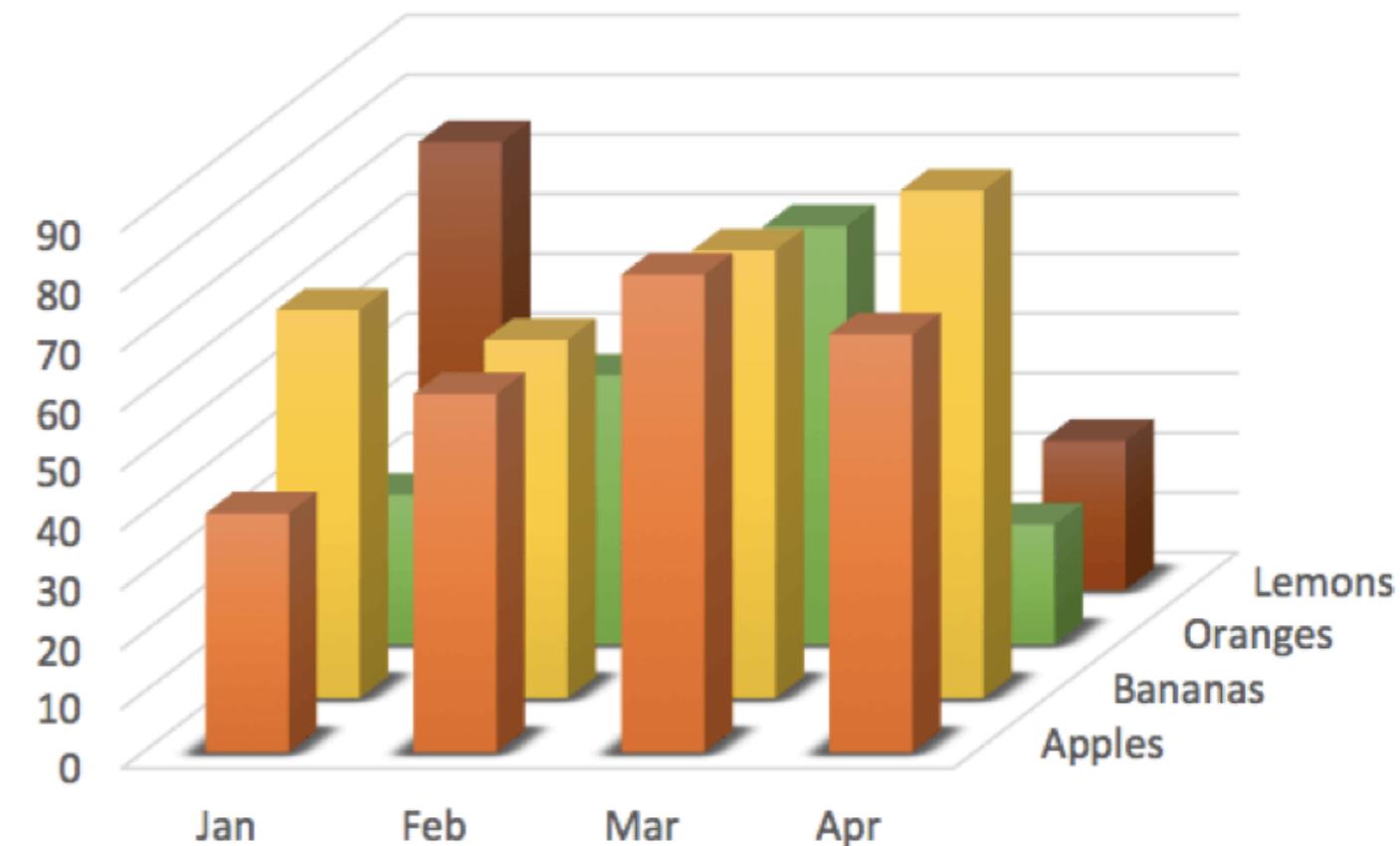
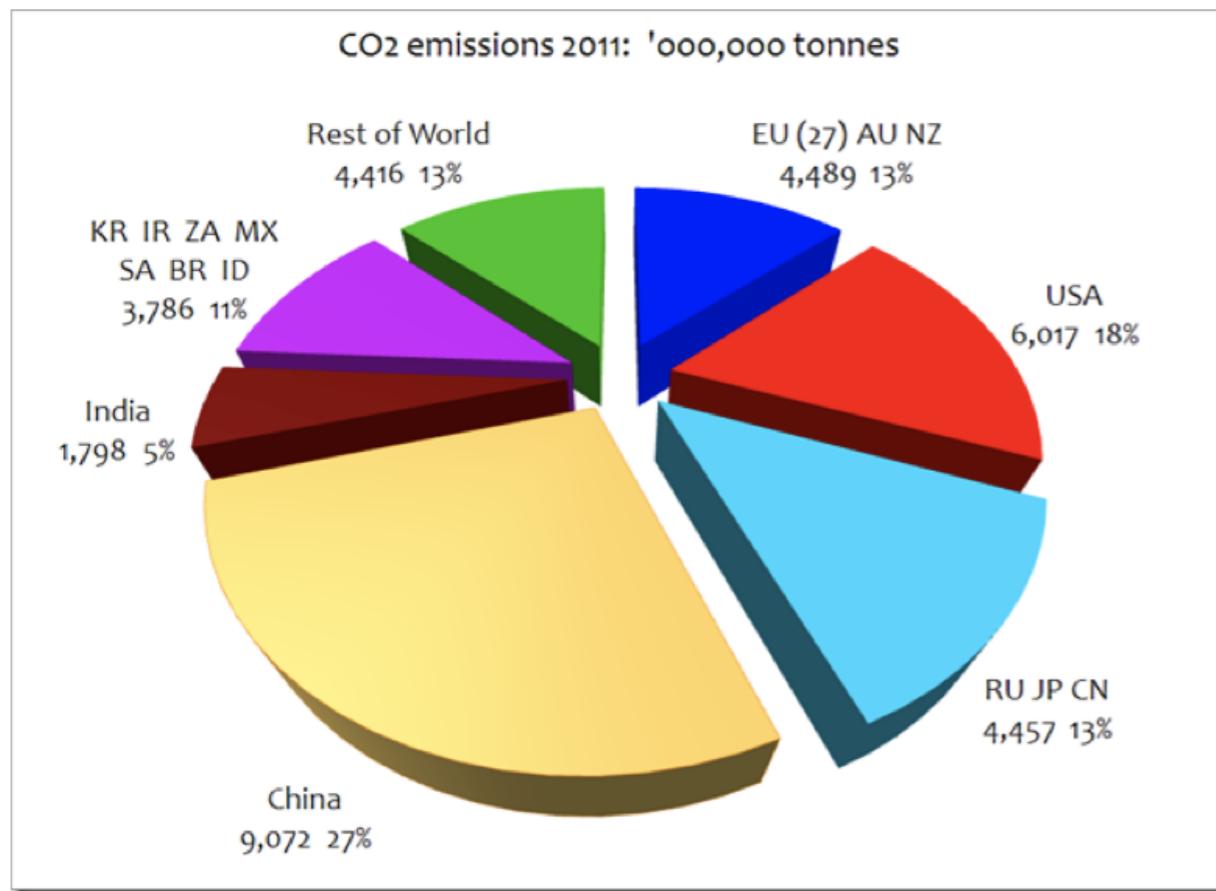




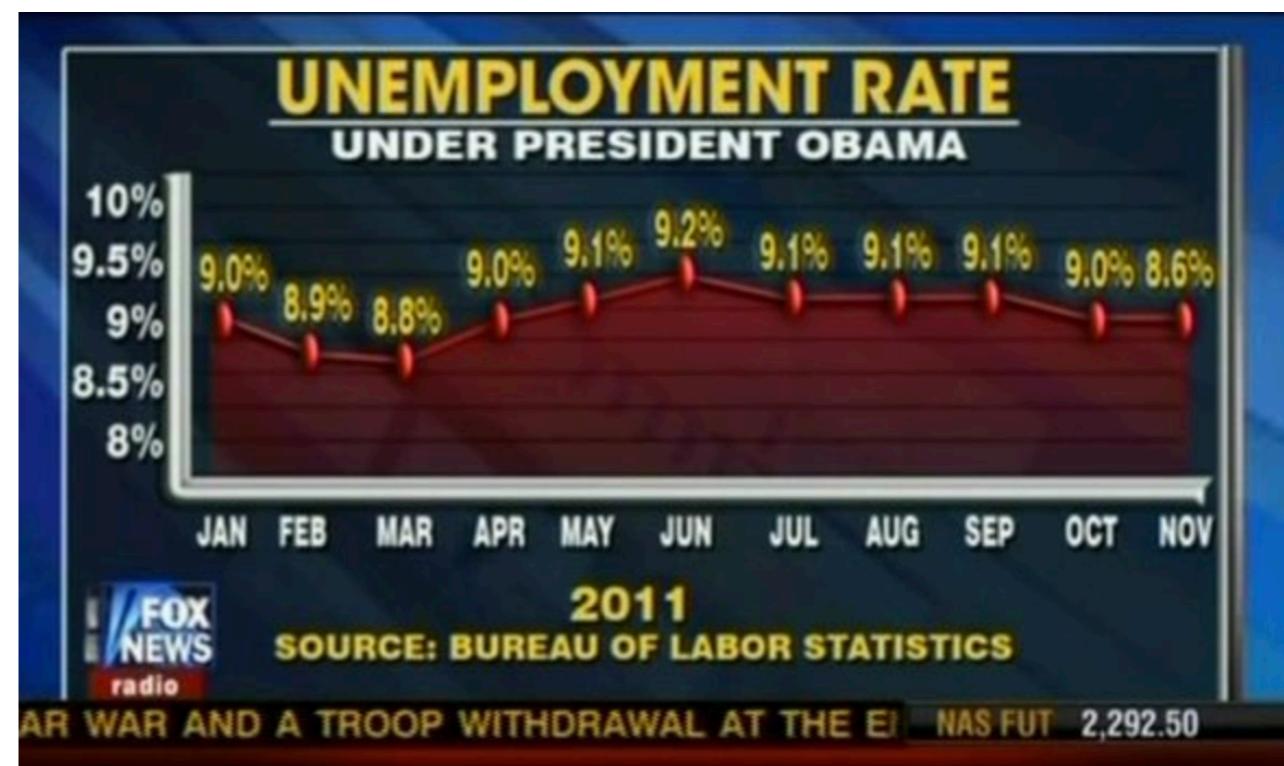
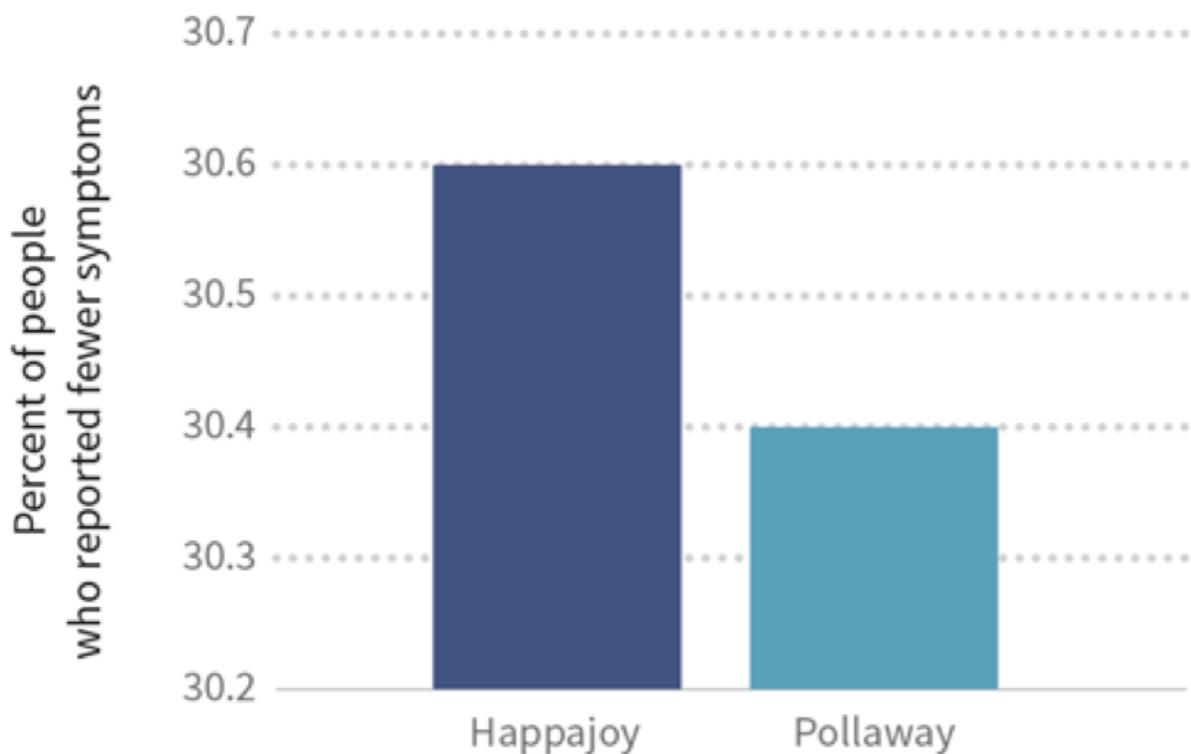
## Duck or Rabbit?

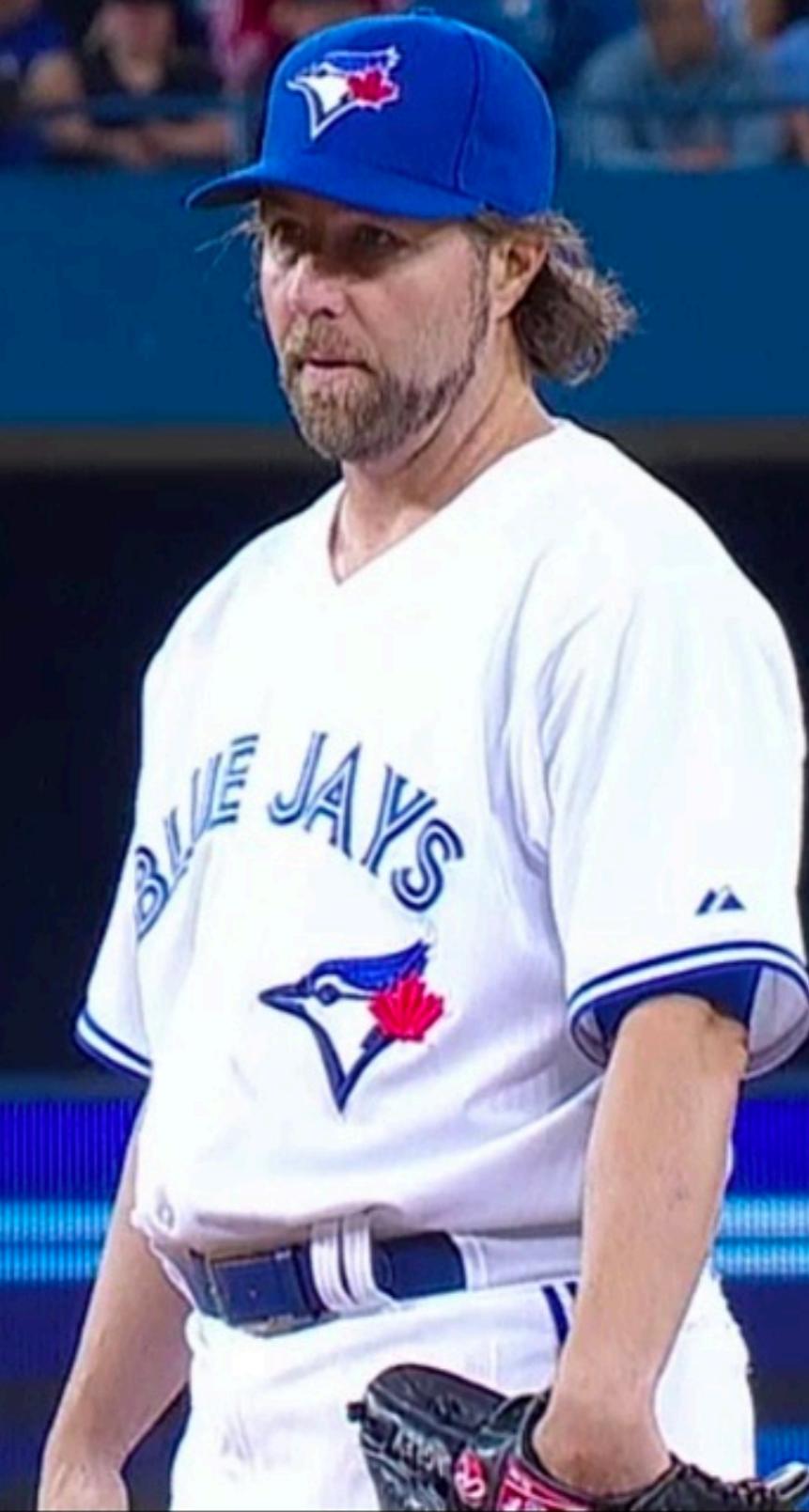
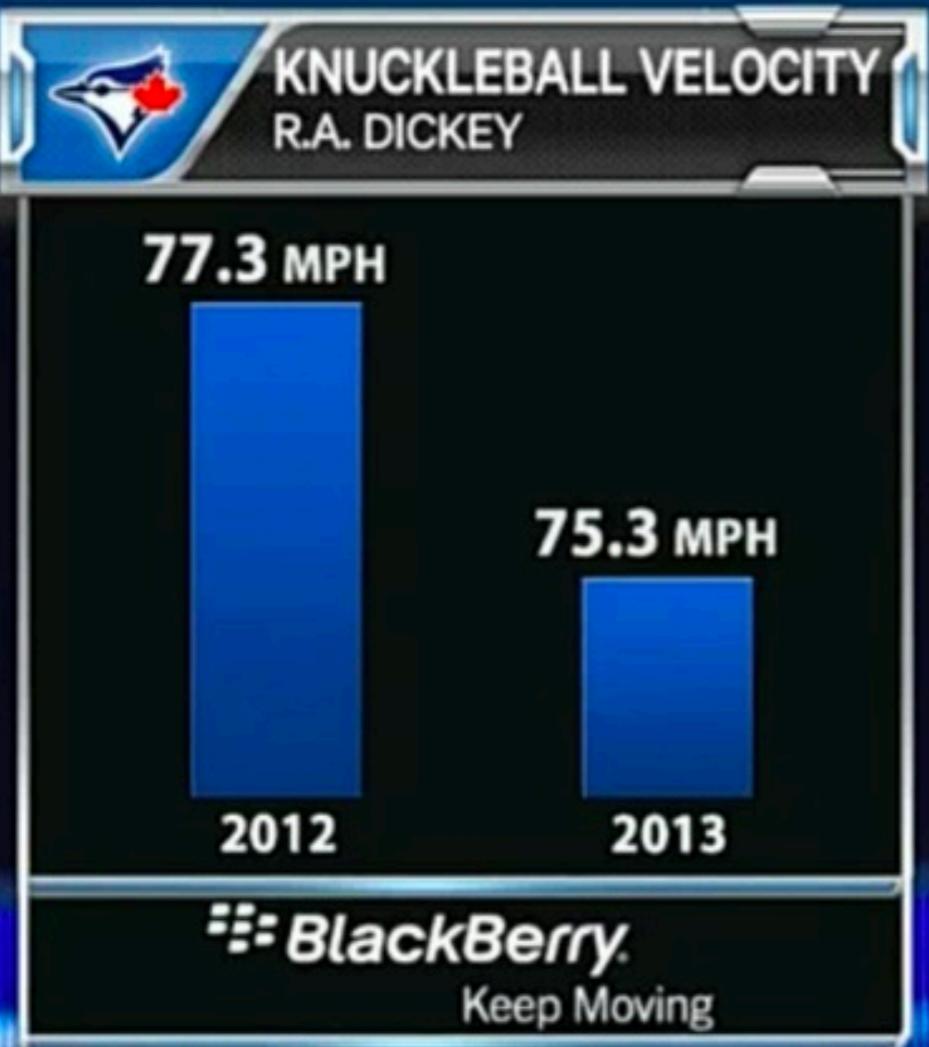


Predictions provided by the Google Cloud Vision API  
Animation by Max Woolf (@minimaxir)



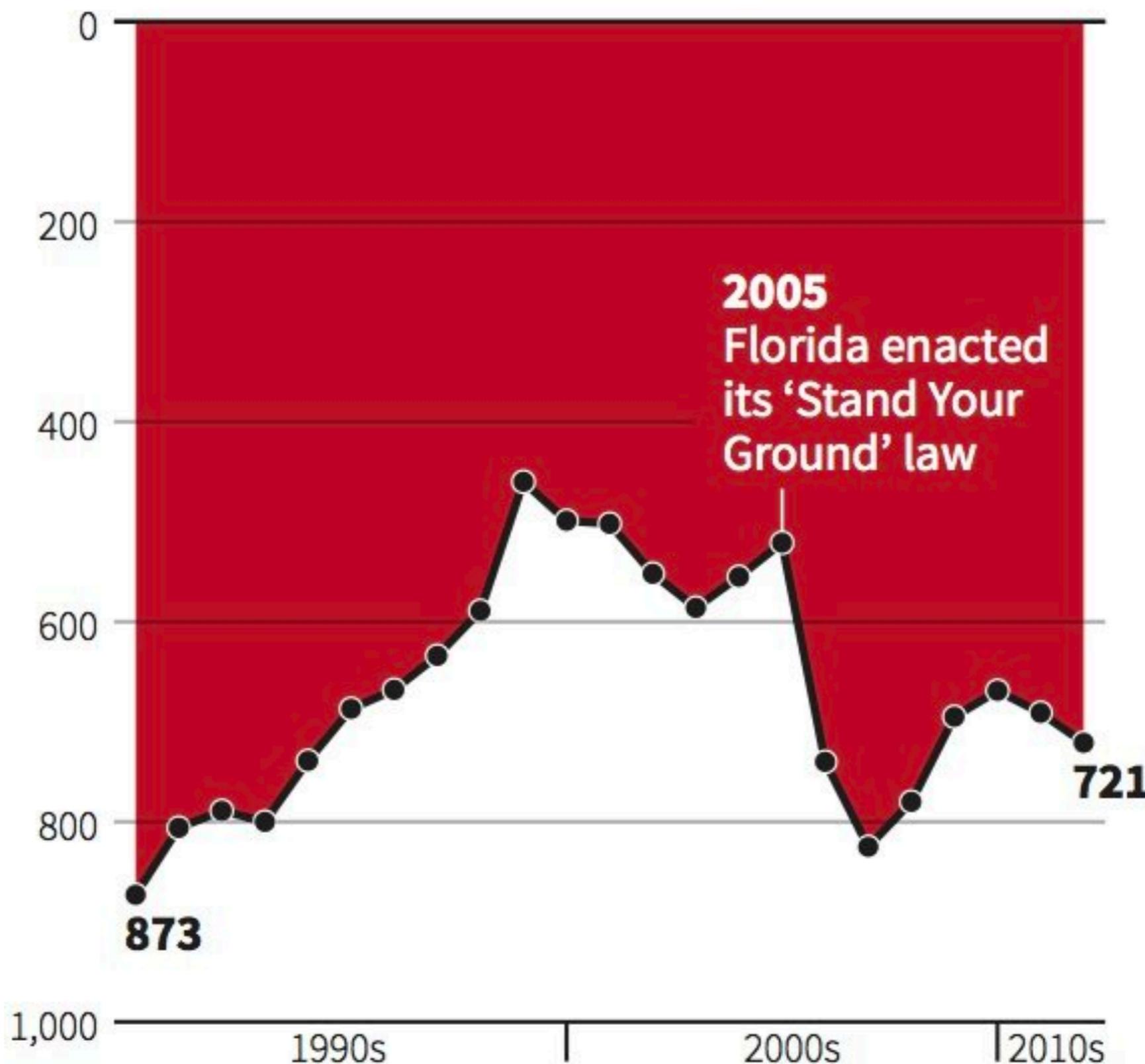
## Effectiveness of Allergy Medicines



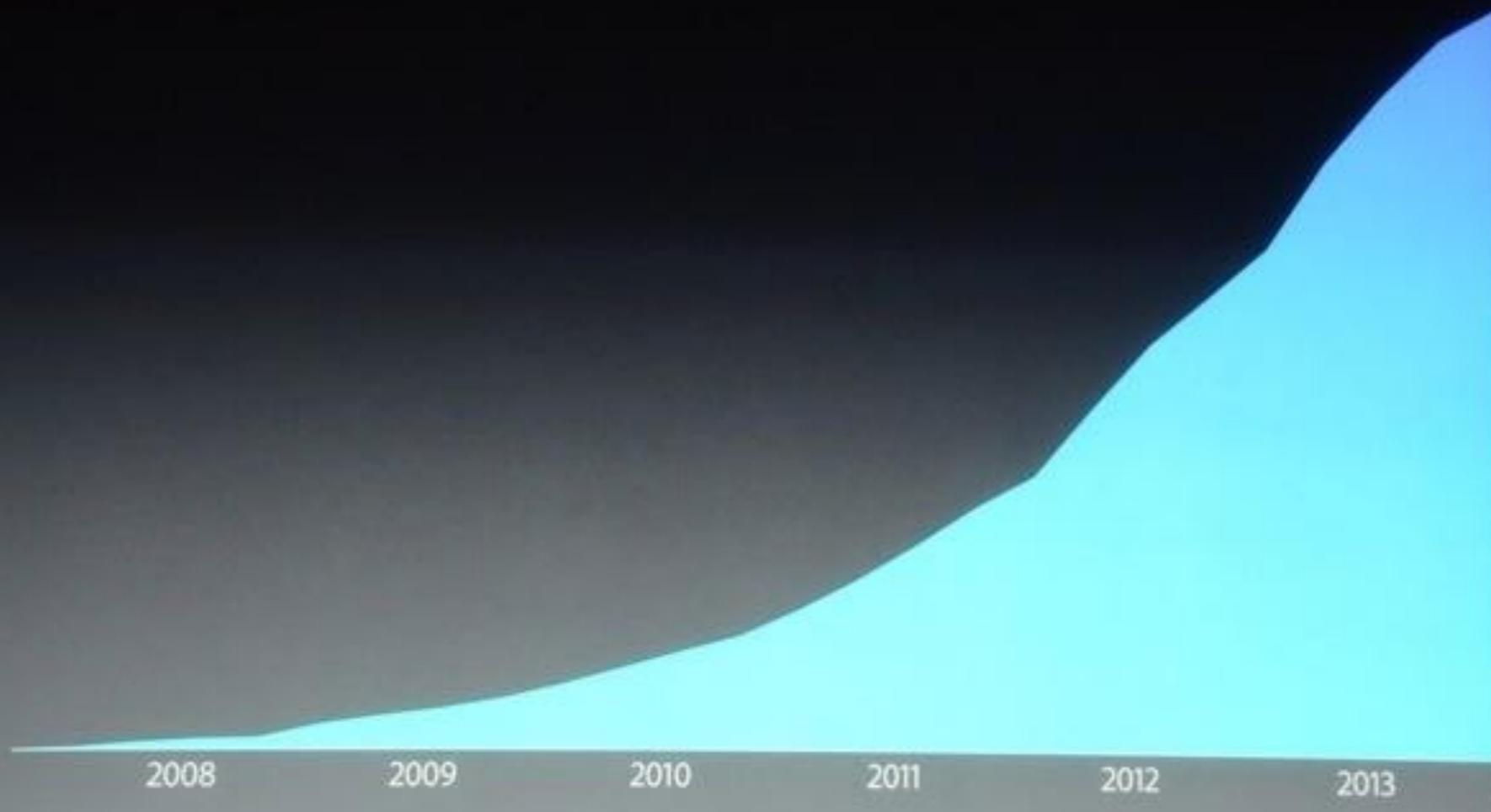


# Gun deaths in Florida

Number of murders committed using firearms

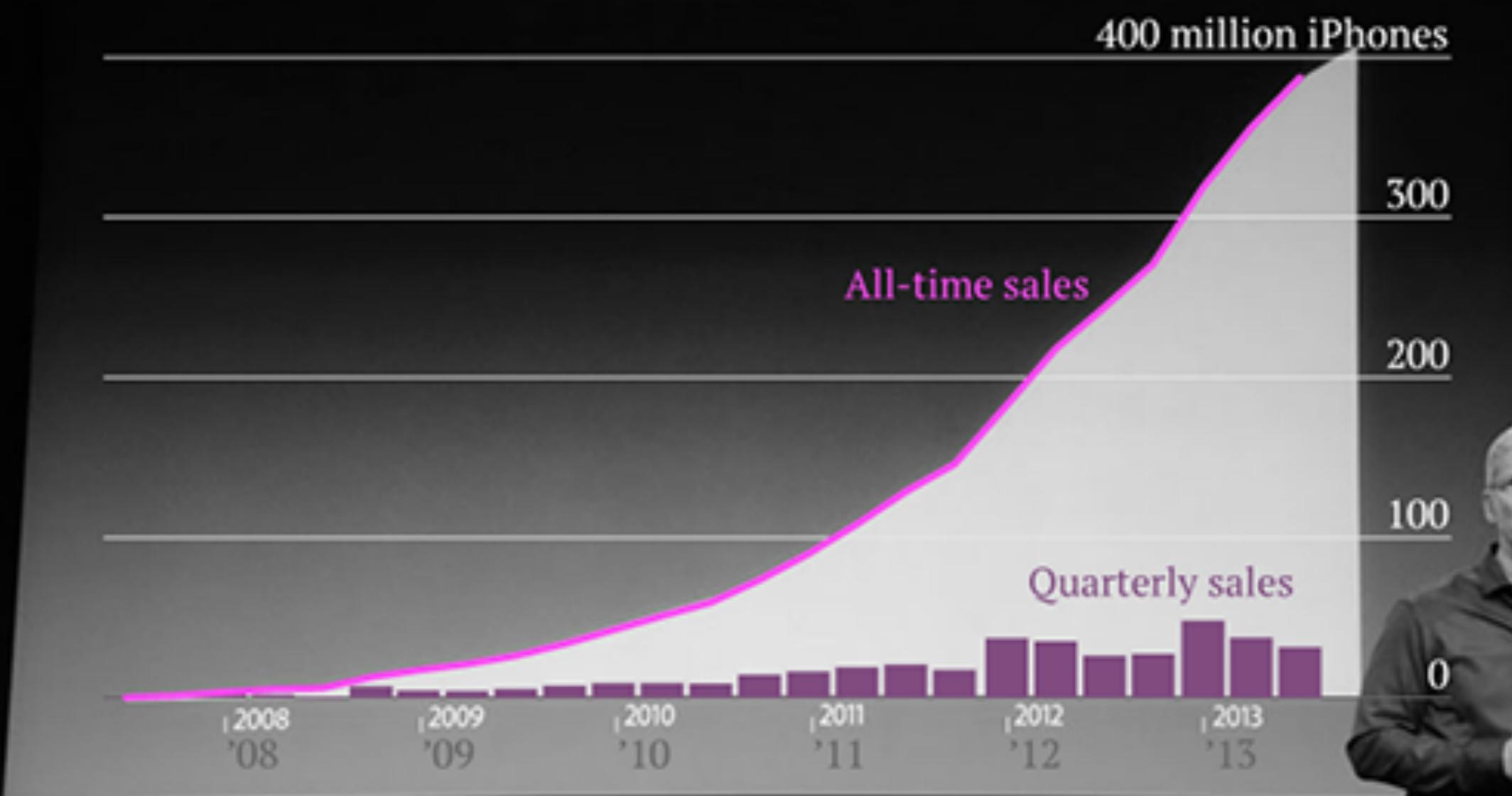


## Cumulative iPhone sales



THE NEW YORK TIMES

## Cumulative iPhone sales



Quartz  
Da  
Photo: T

THE VERGE

# HOW TO (NOT) LIE WITH VISUALIZATION

Huff, D. (1993). **How to lie with statistics.** WW Norton & Company.

King, G. (1986). **How not to lie with statistics: Avoiding common mistakes in quantitative political science.** American Journal of Political Science, 666-687.

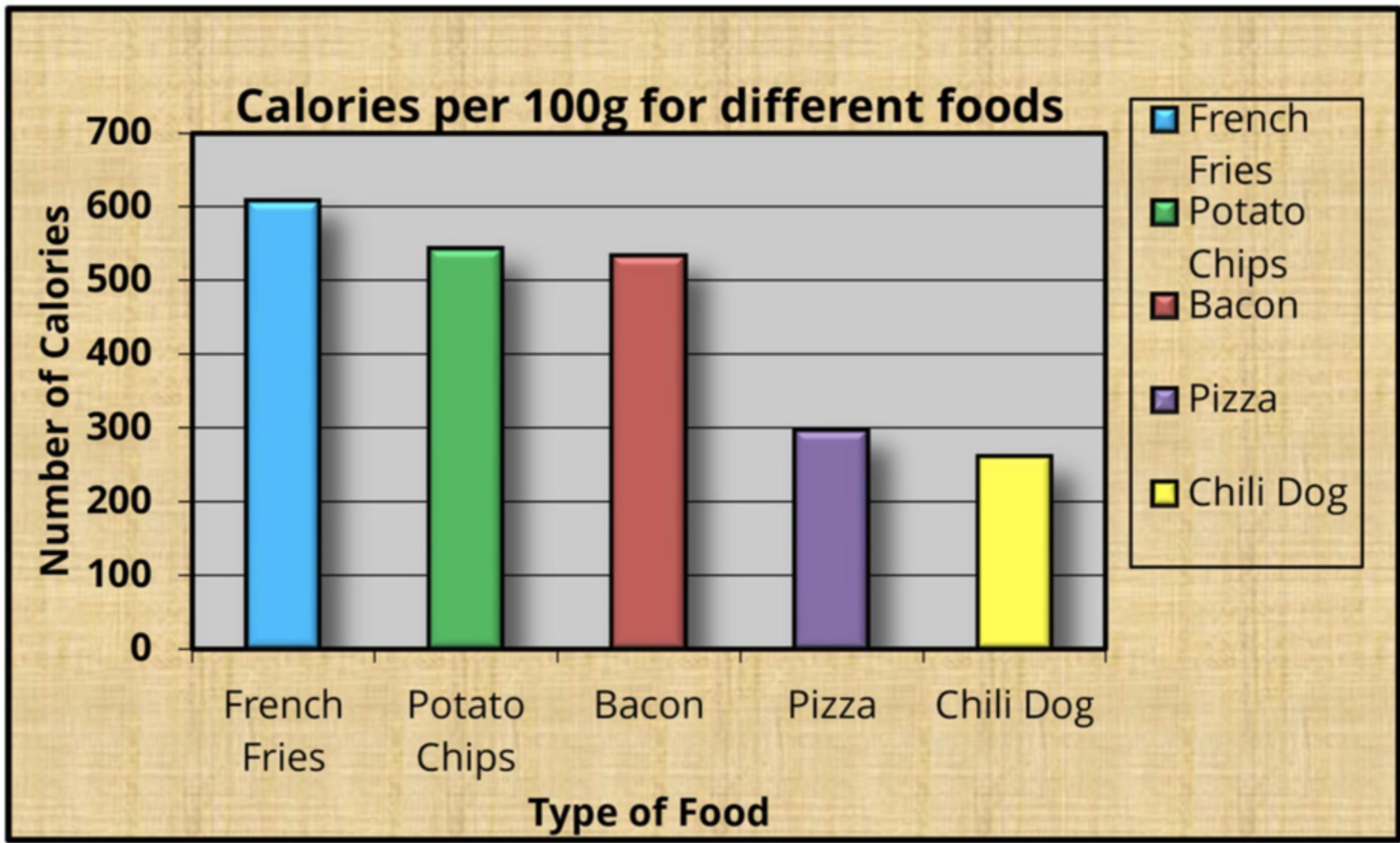
Jones, G. E. (2011). **How to lie with charts.** LaPuerta Books and Media.

Monmonier, M. (2018). **How to lie with maps.** University of Chicago Press.

Rogowitz, B. E., Treinish, L. A., & Bryson, S. (1996). **How not to lie with visualization.** Computers in Physics, 10(3), 268-273.

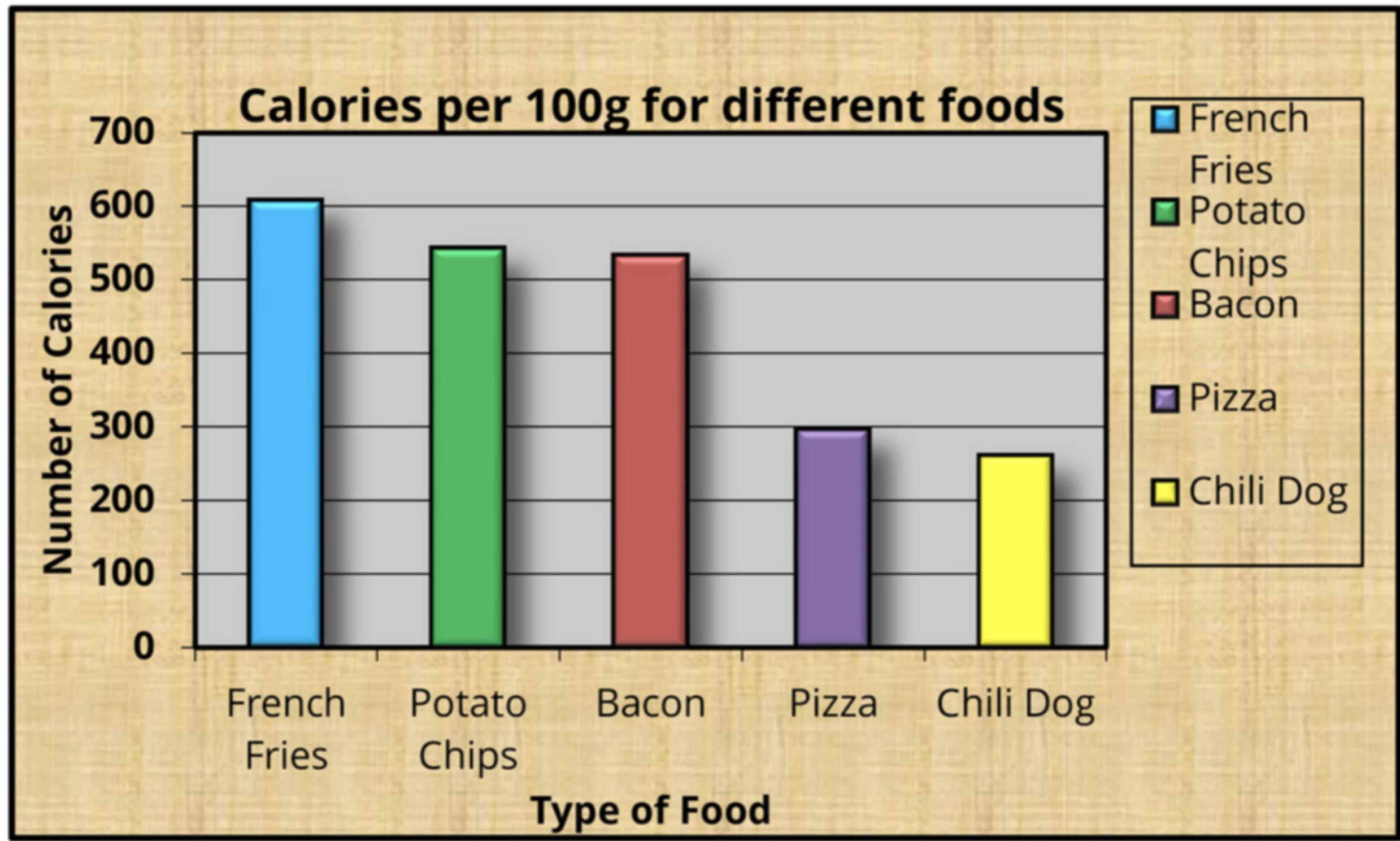
Pandey, A. V., Rall, K., Satterthwaite, M. L., Nov, O., & Bertini, E. (2015, April). **How deceptive are deceptive visualizations?: An empirical analysis of common distortion techniques.** In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1469-1478). ACM.

# CLARITY

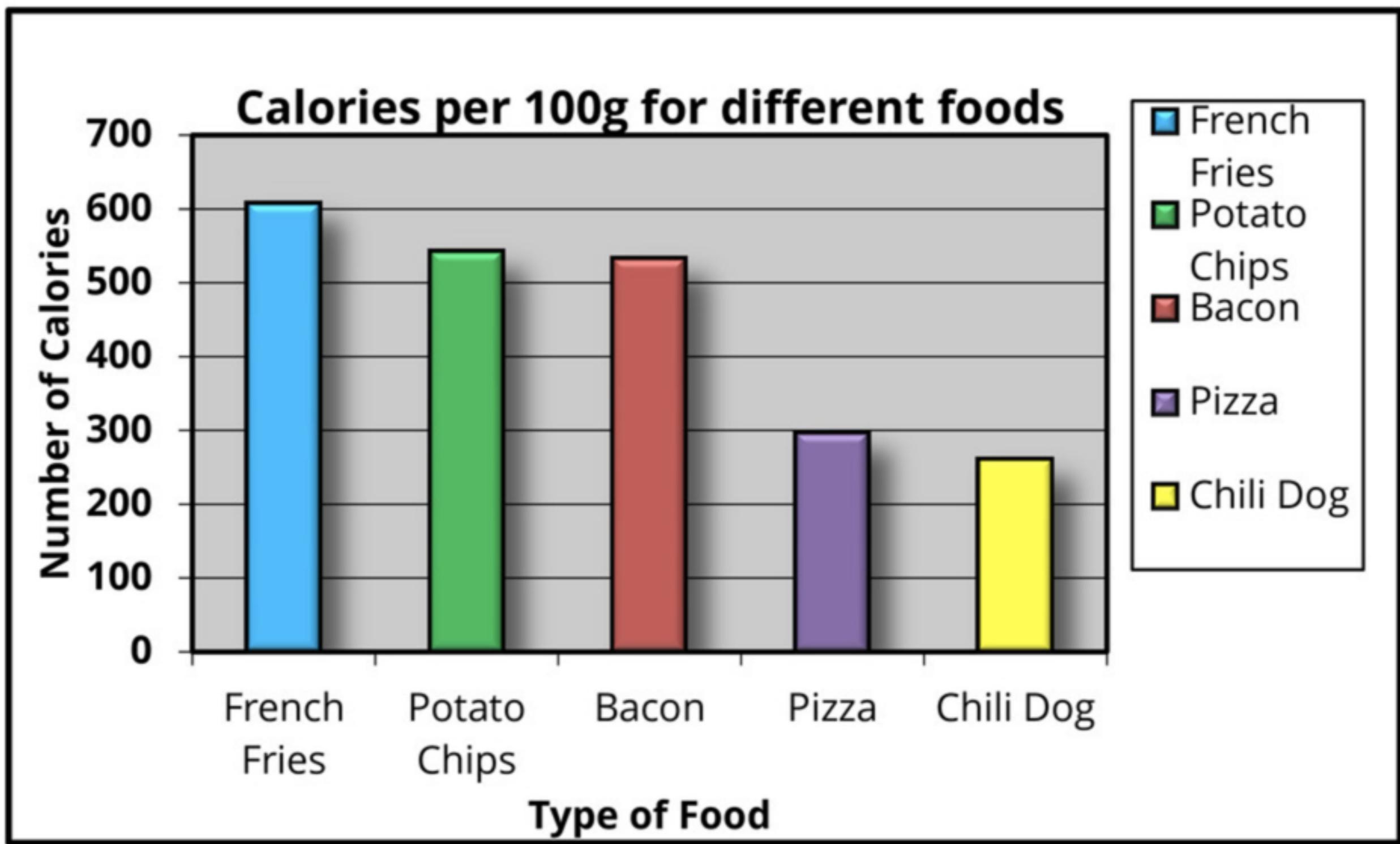


$$\text{Data-Ink Ratio} = \frac{\text{Data-Ink}}{\text{Total Ink used to represent the graphic}}$$

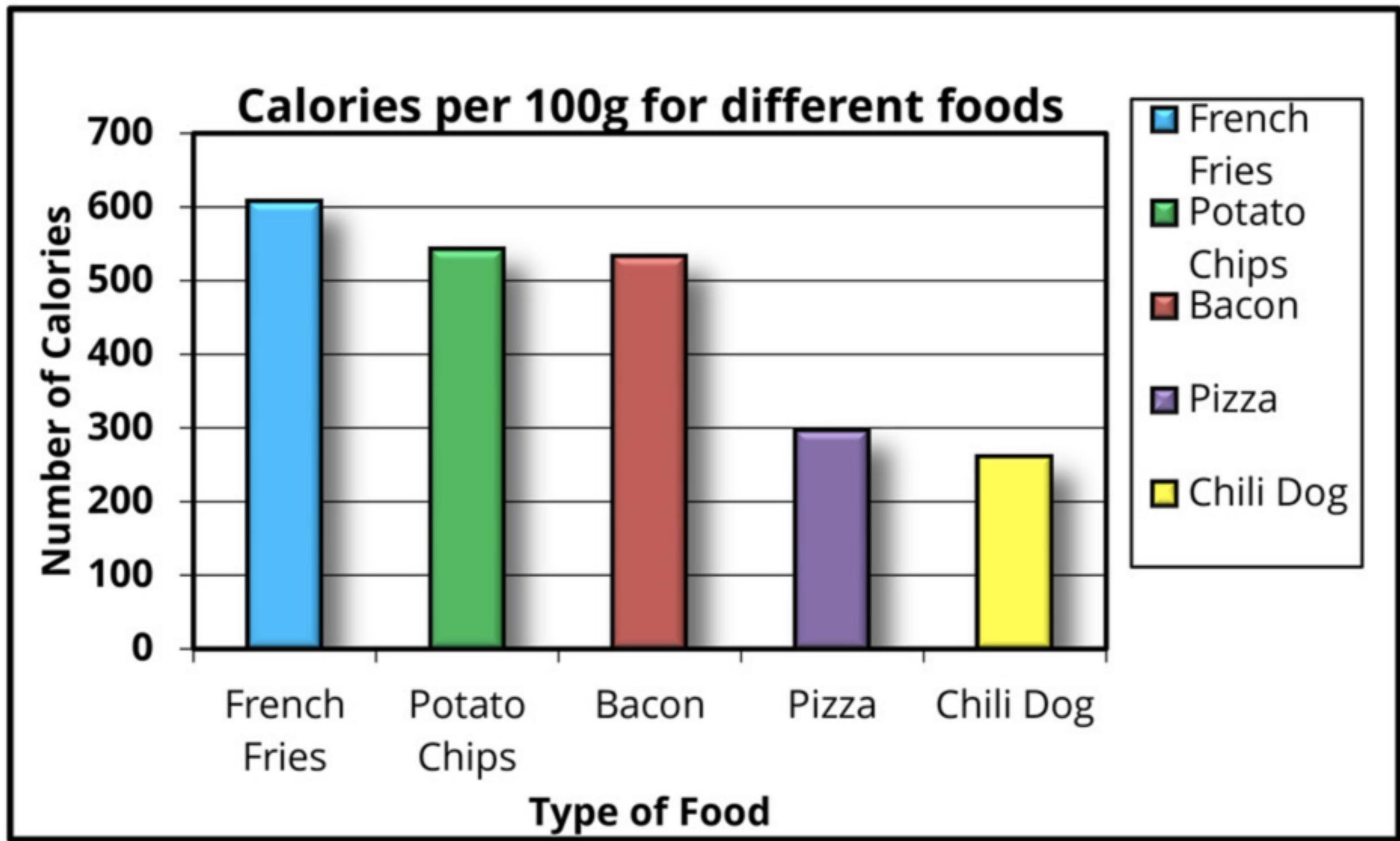
# Remove backgrounds



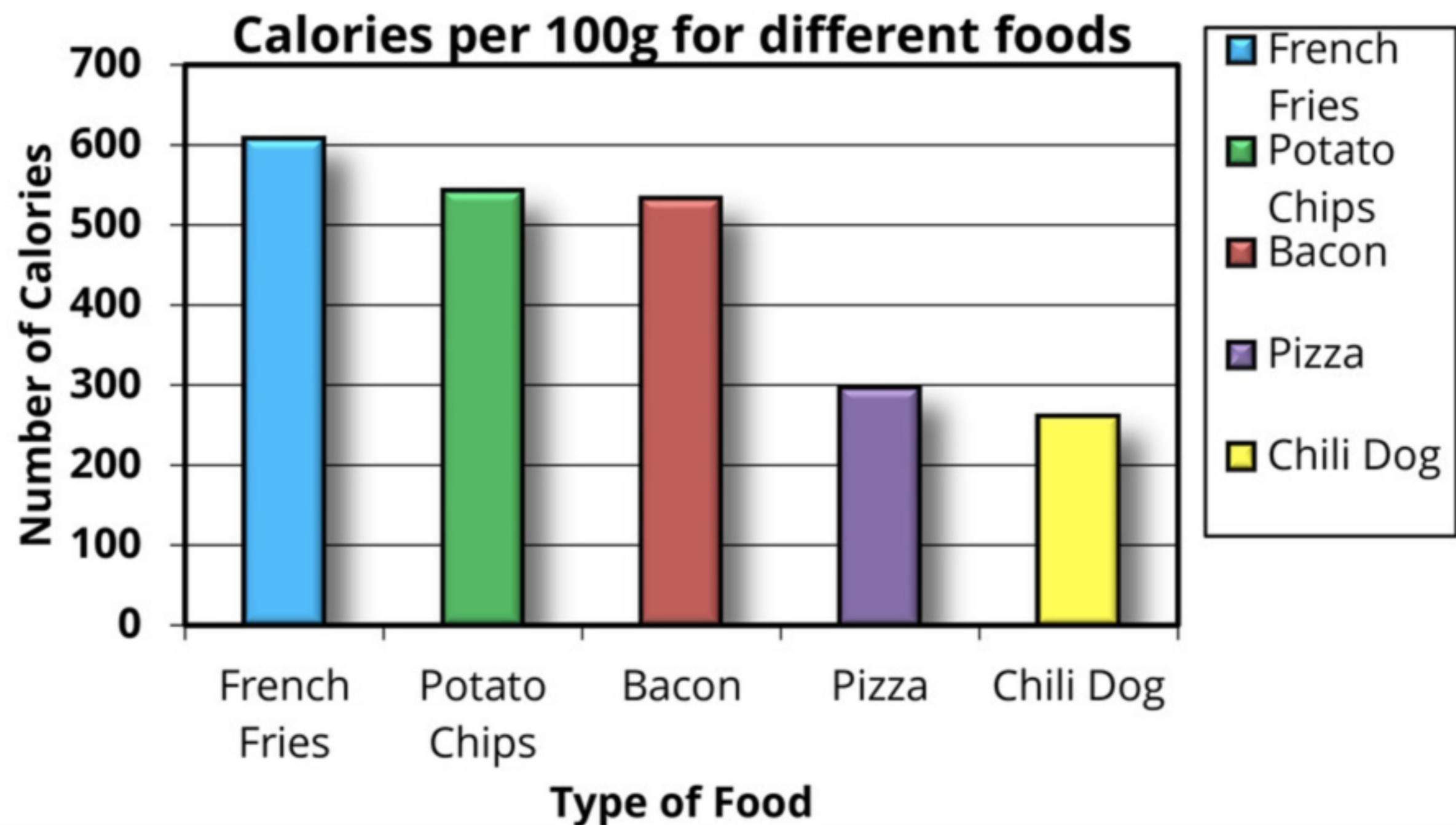
# Remove backgrounds



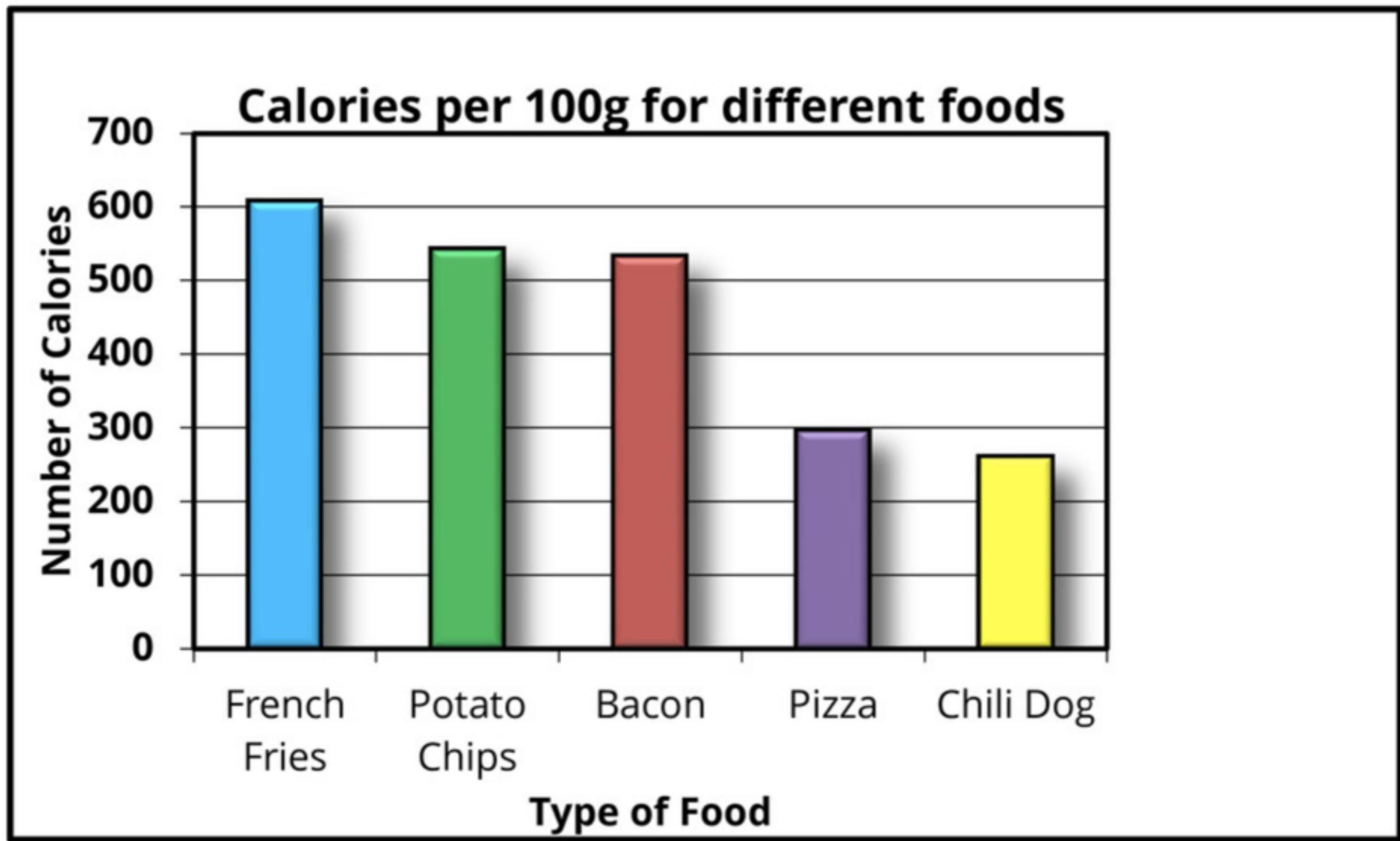
# Remove backgrounds



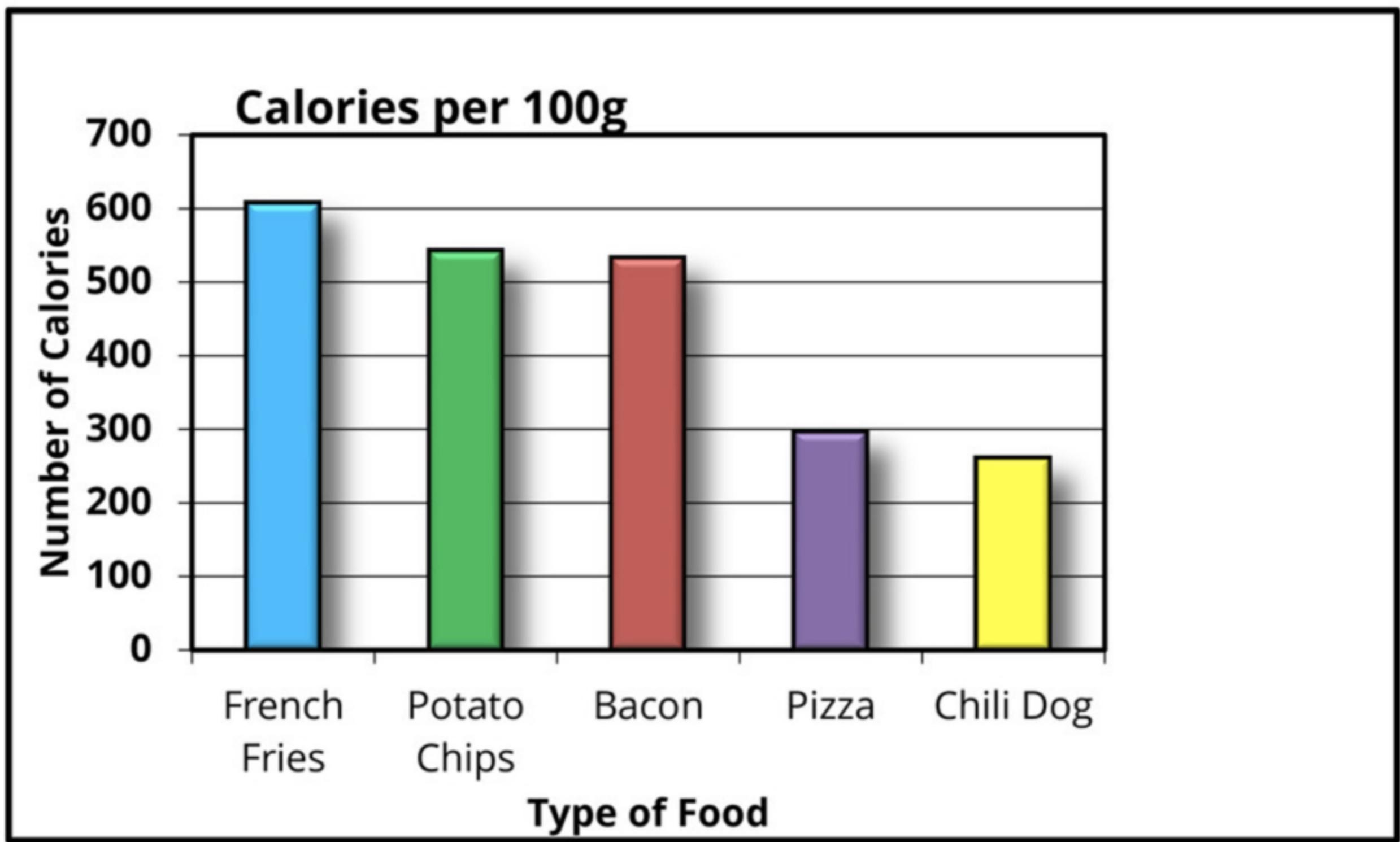
# Remove redundant labels



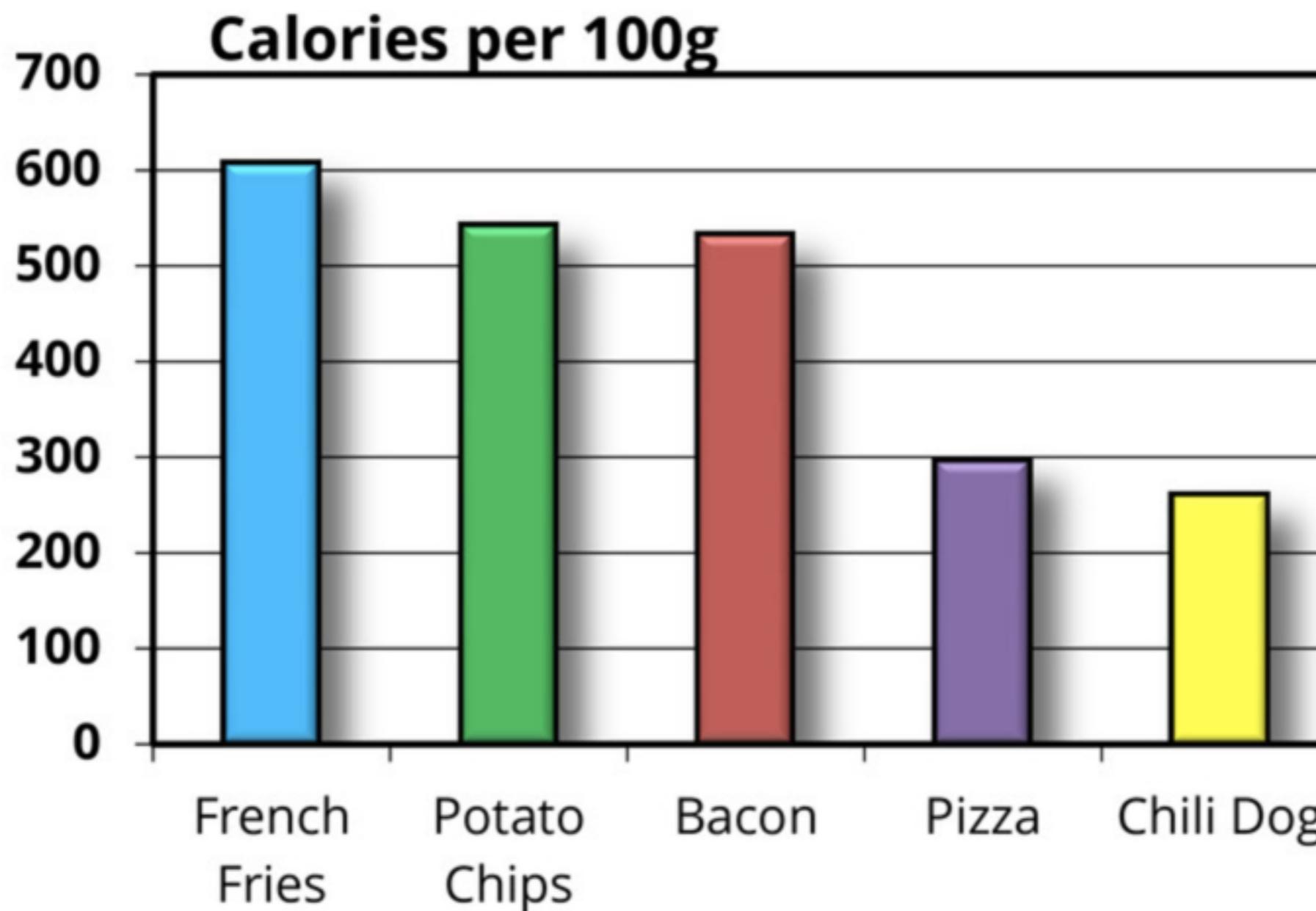
# Remove redundant labels



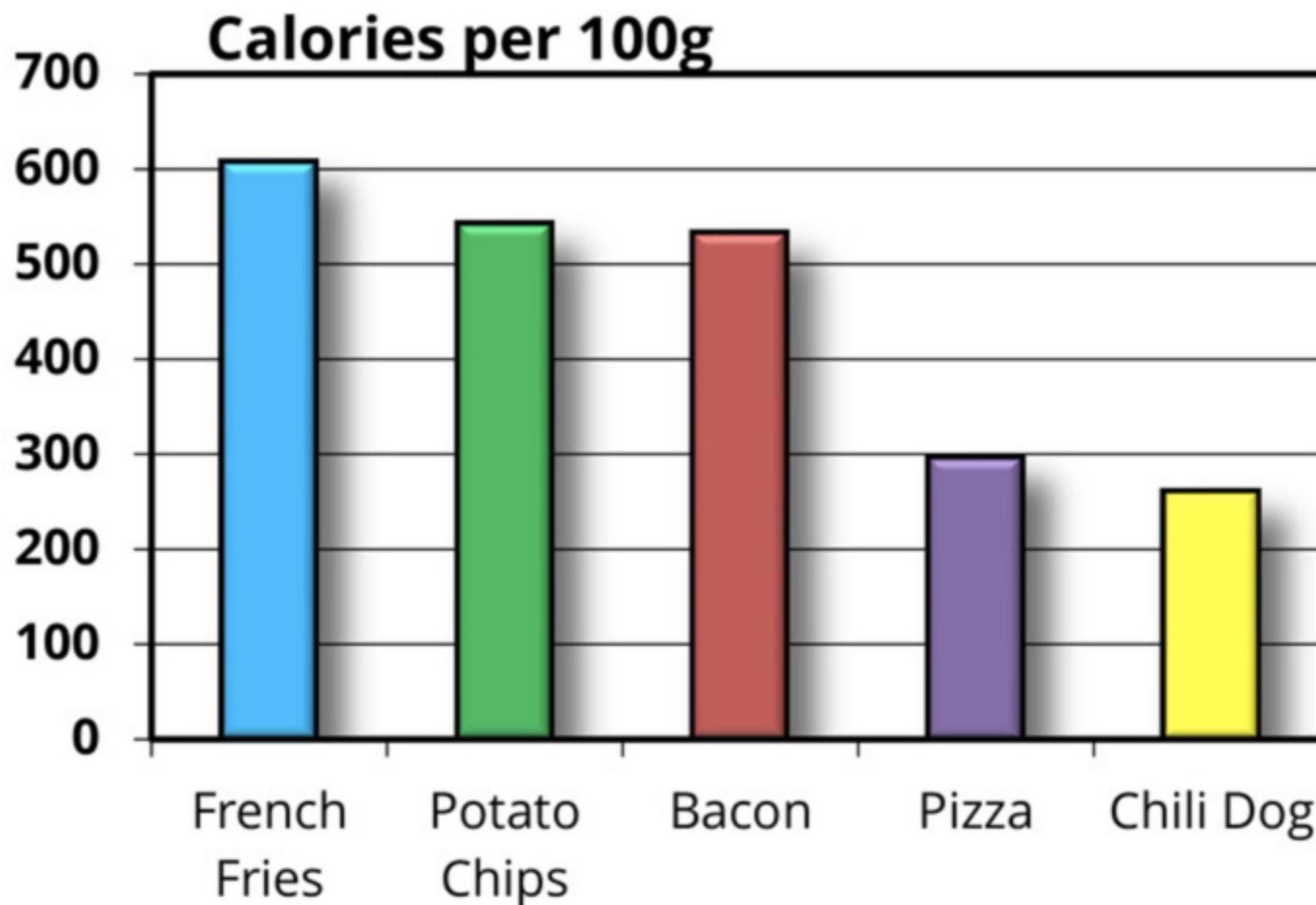
# Remove redundant labels



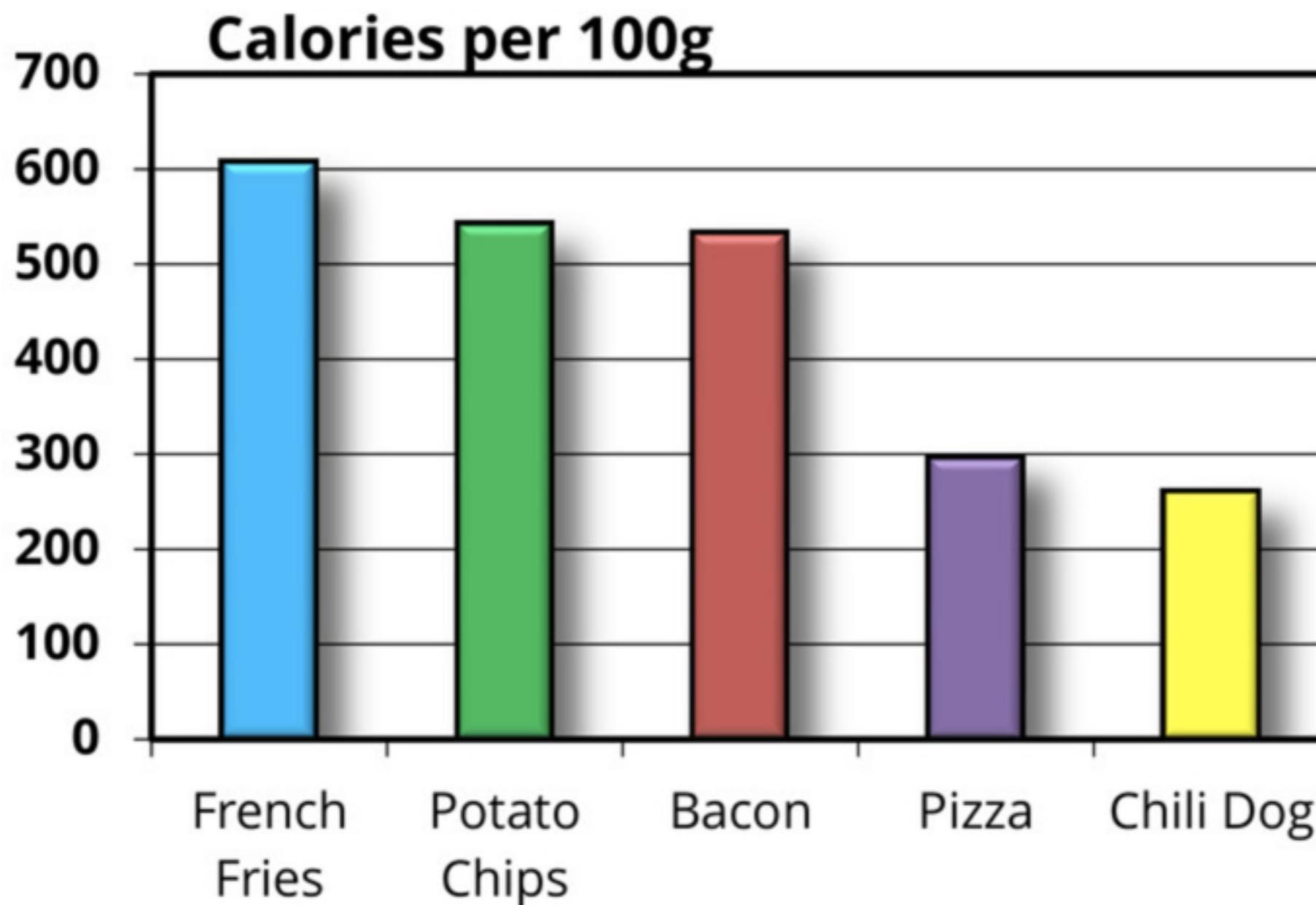
# Remove redundant labels



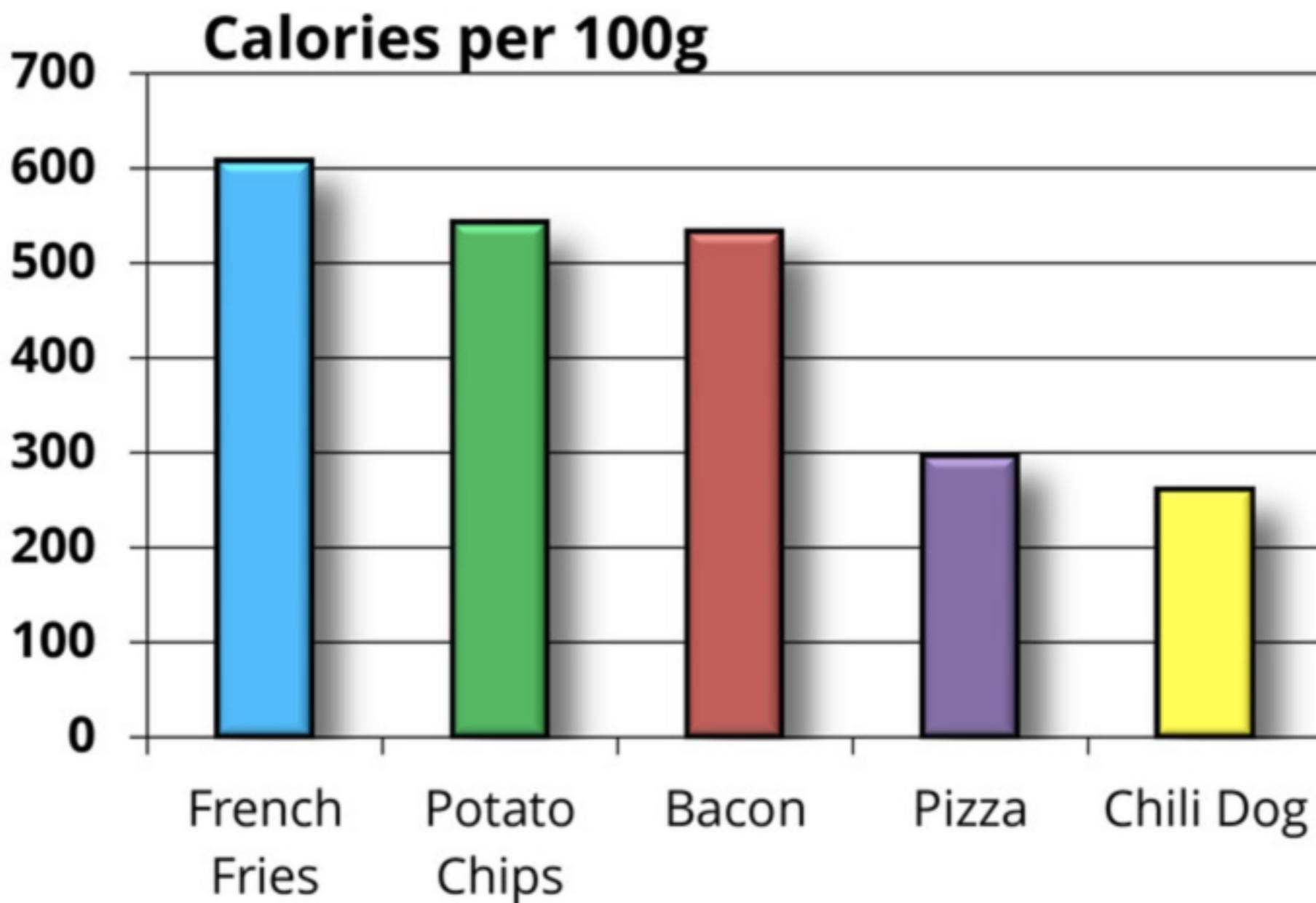
# Remove borders



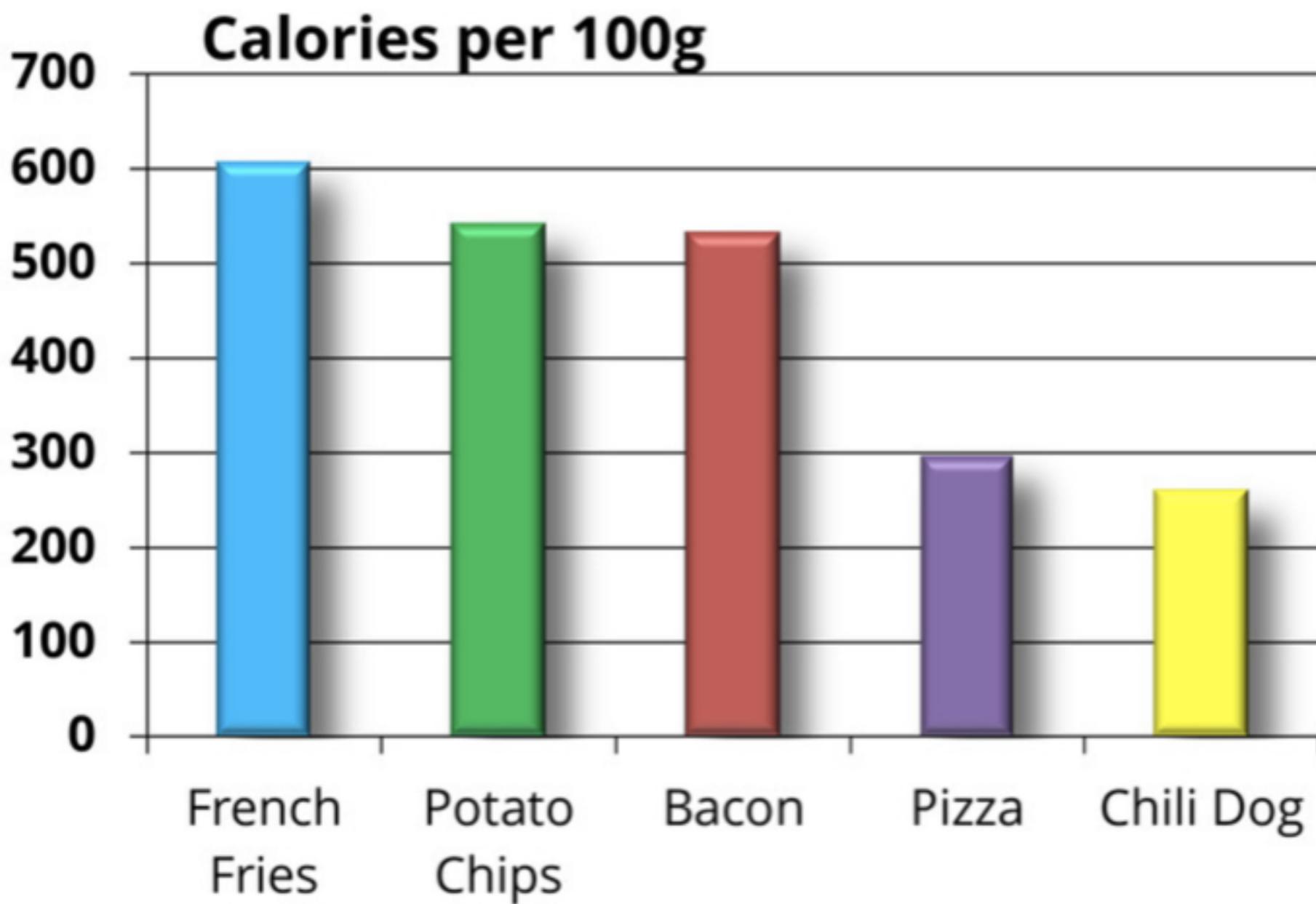
# Remove borders



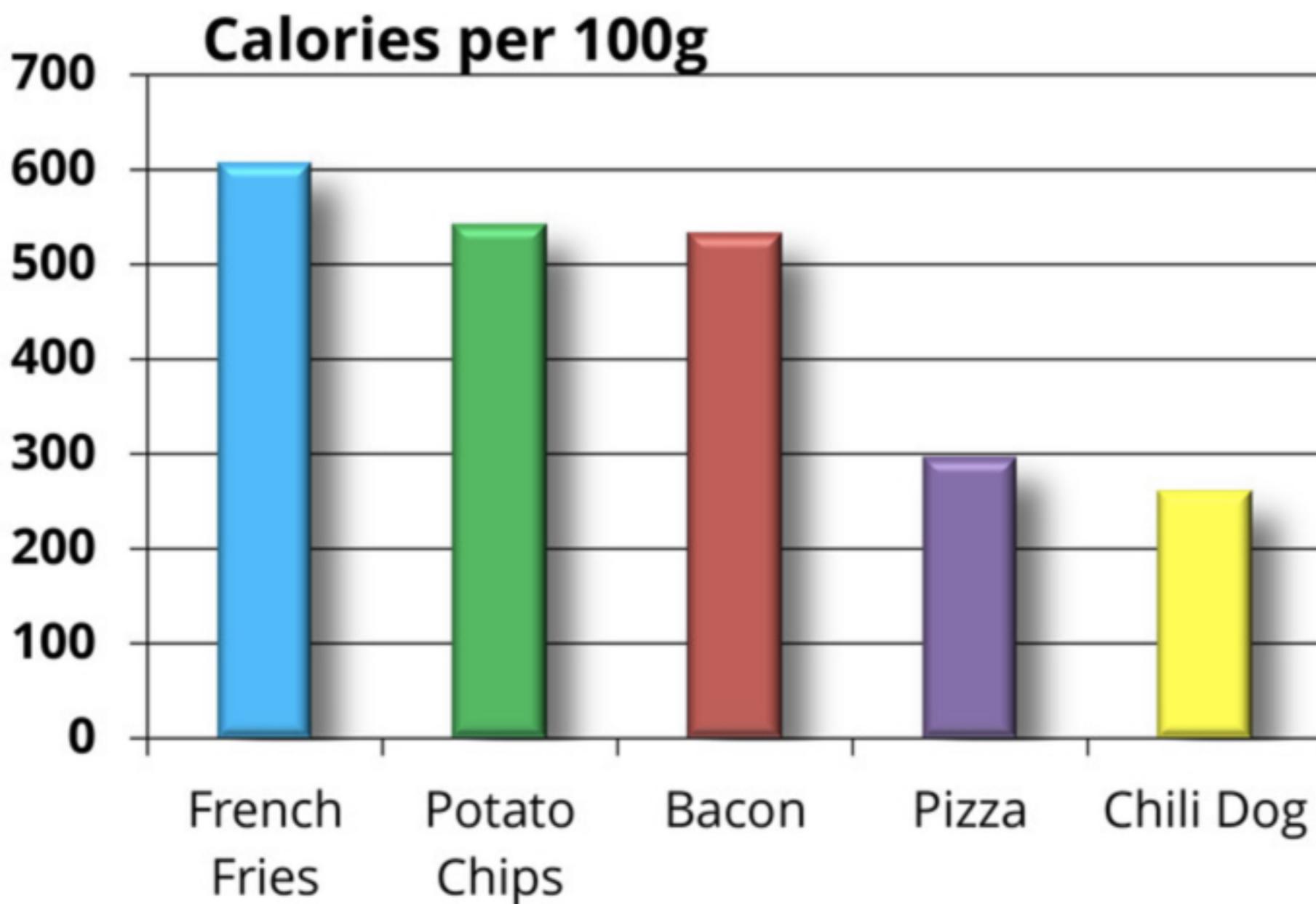
# Remove borders



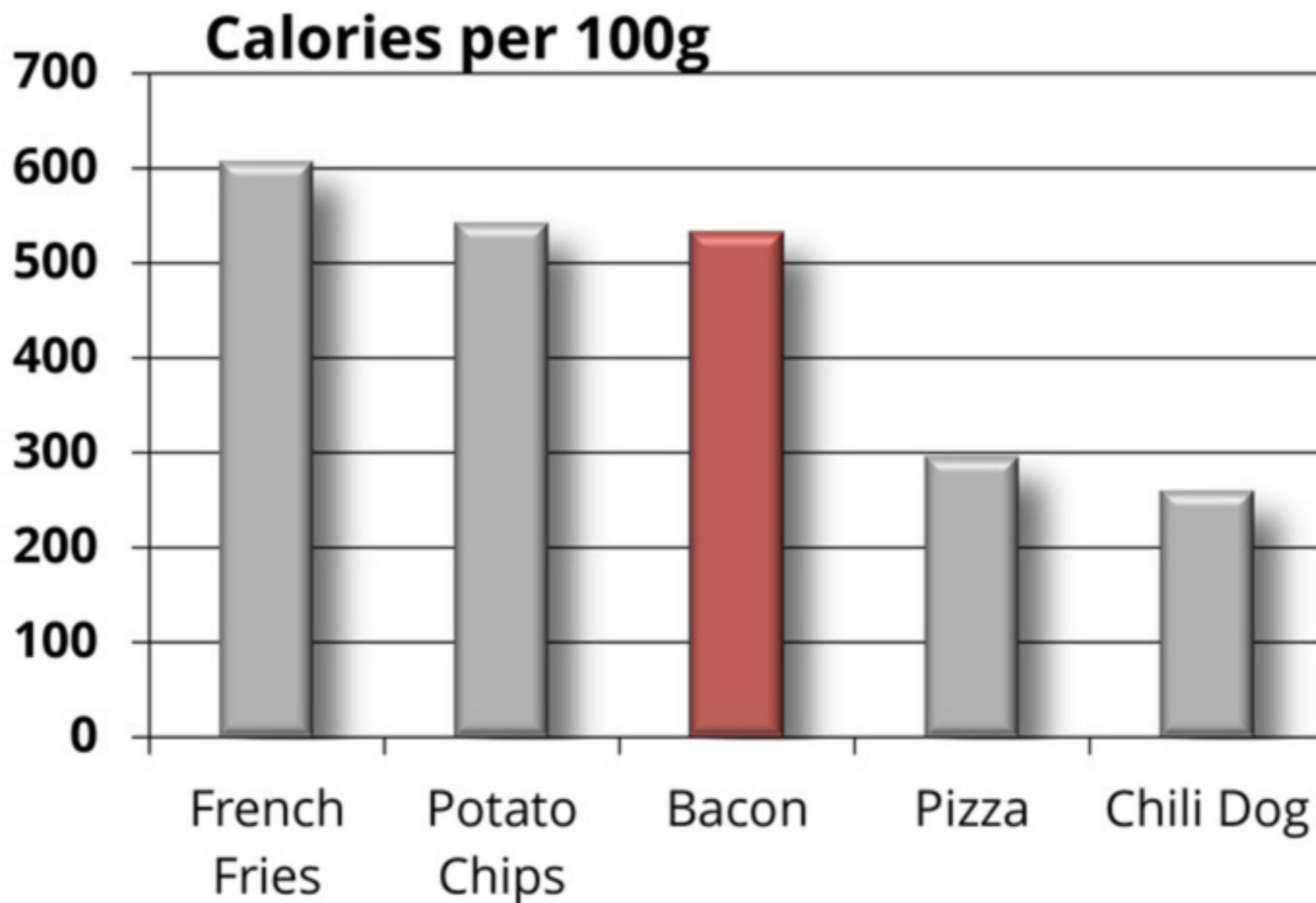
# Remove borders



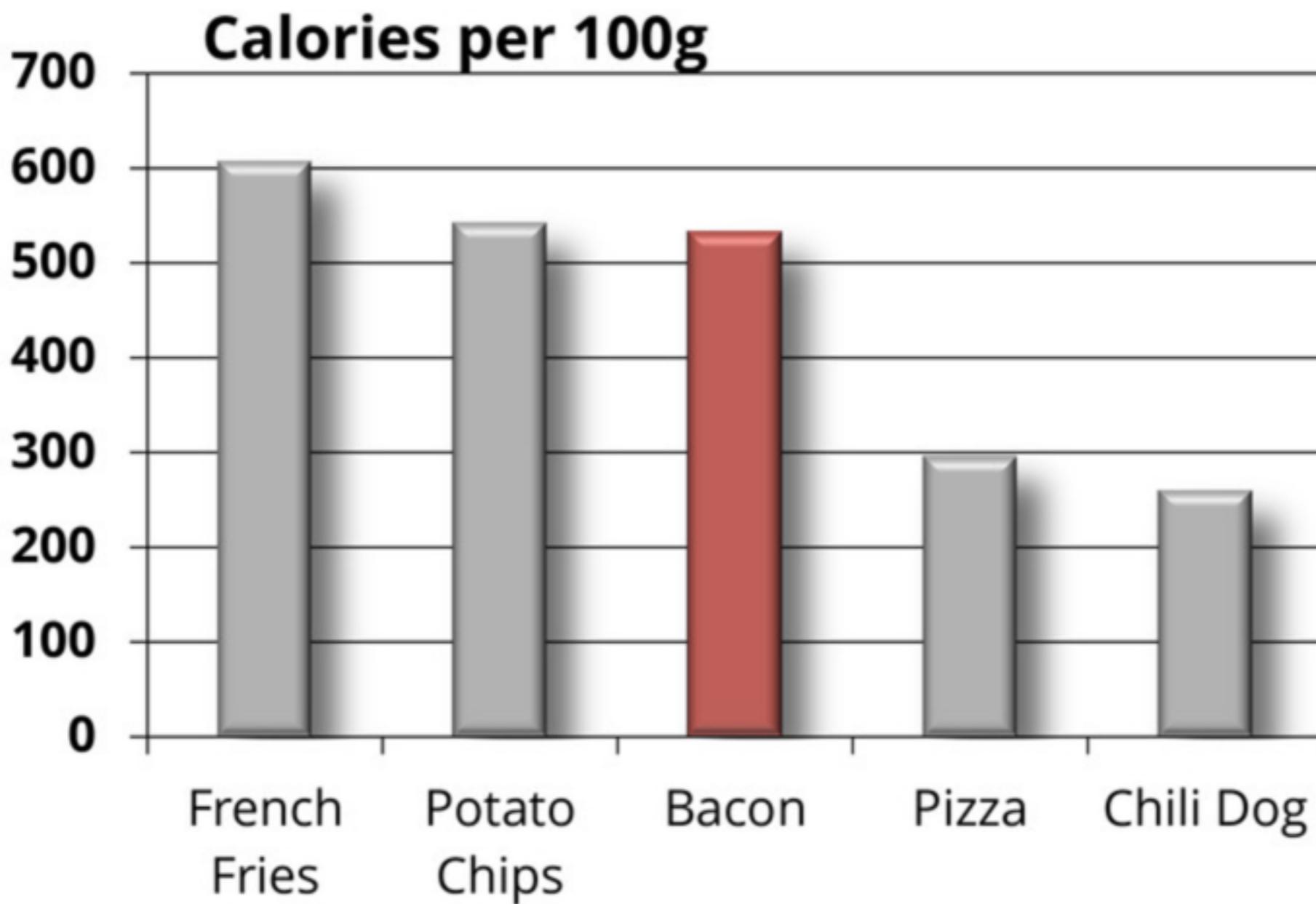
# Reduce colors



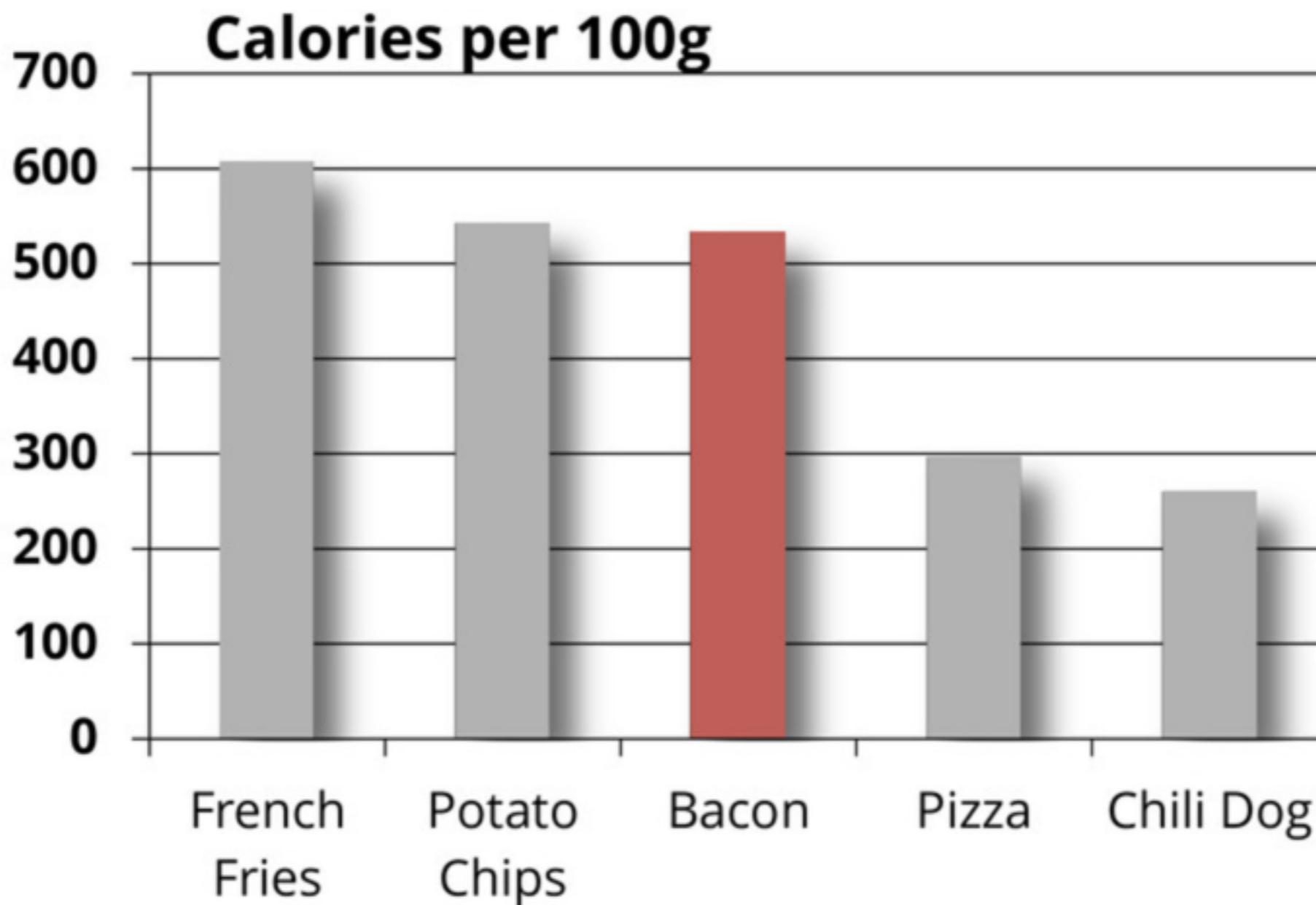
# Reduce colors



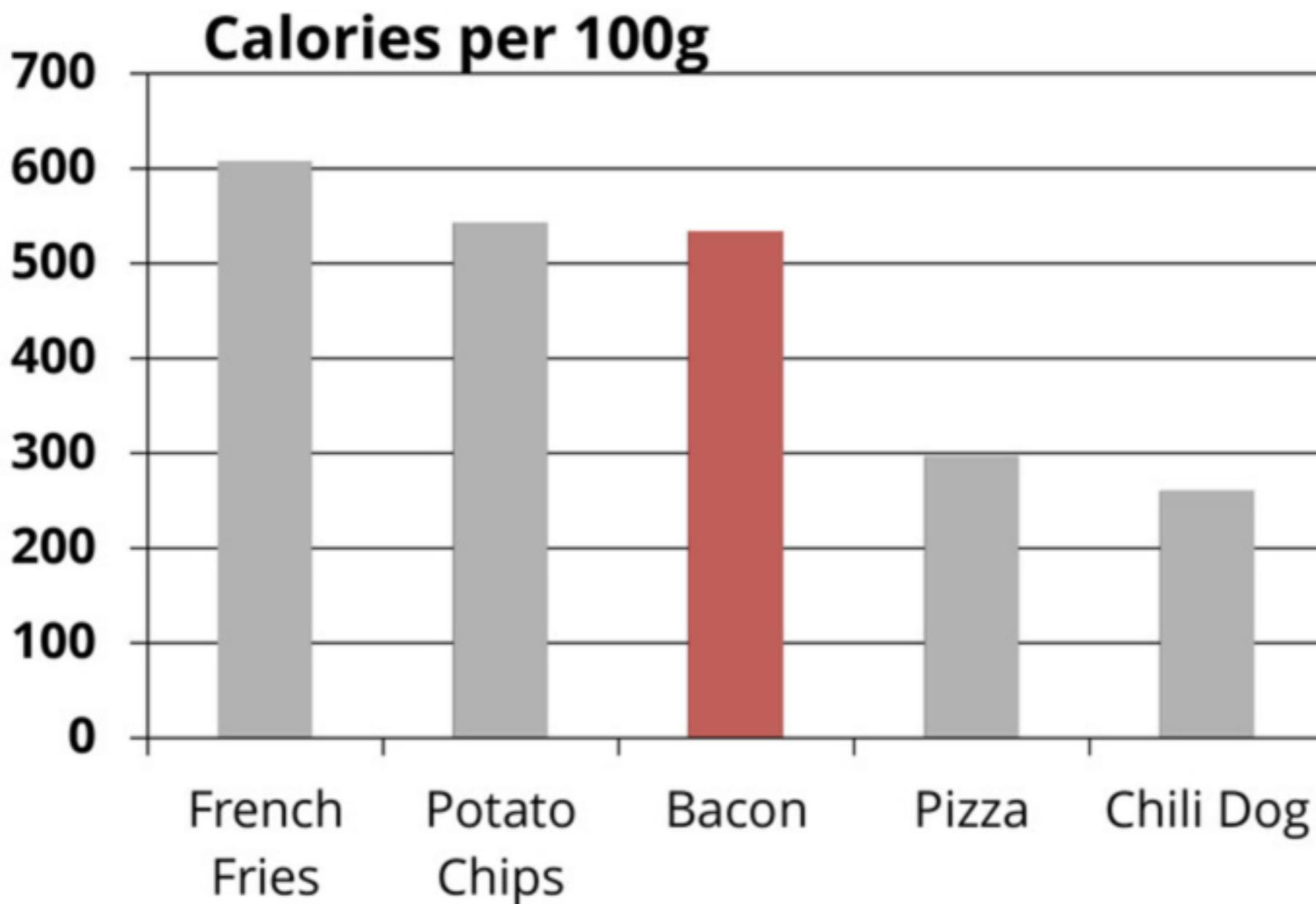
# Remove special effects



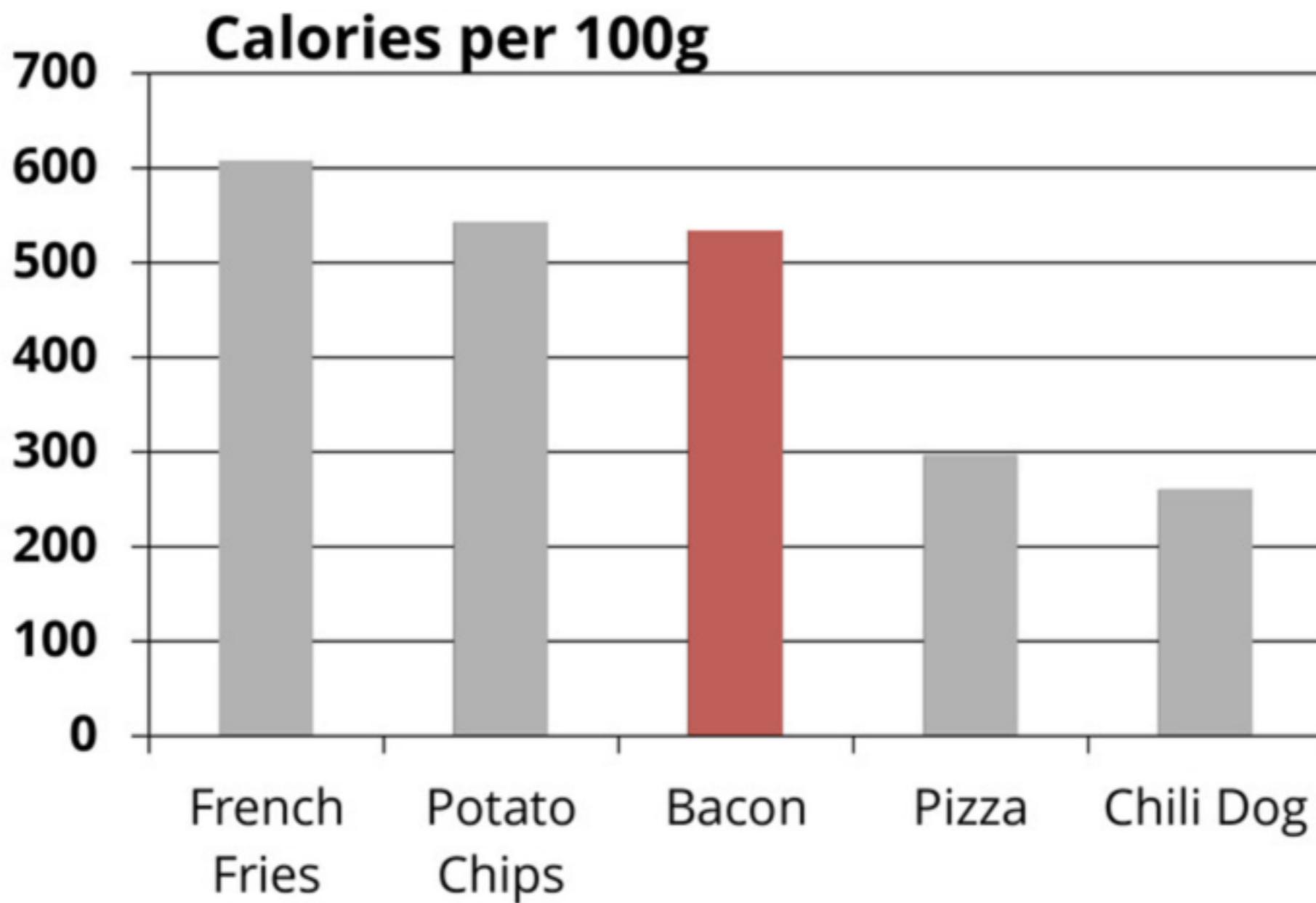
# Remove special effects



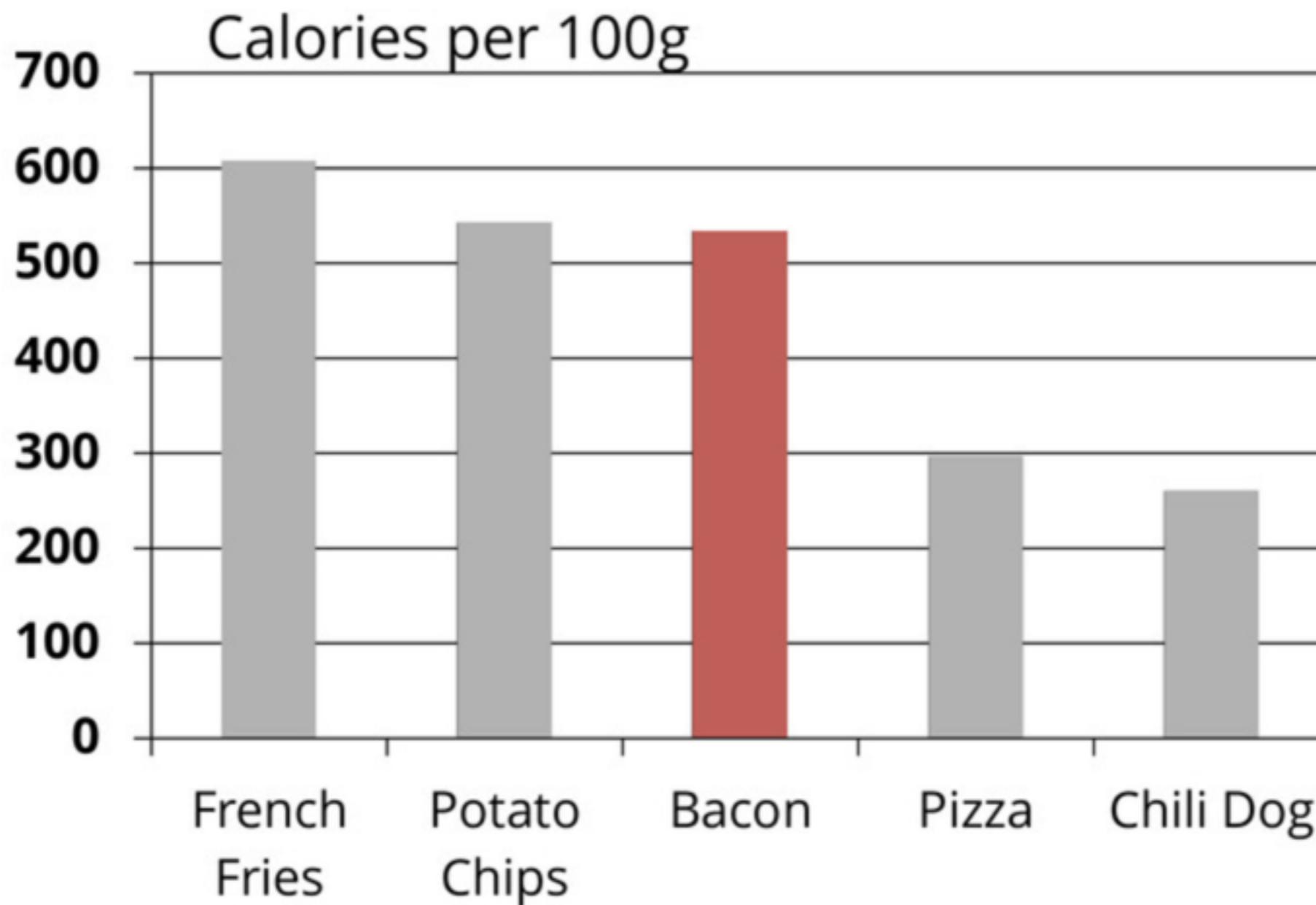
# Remove special effects



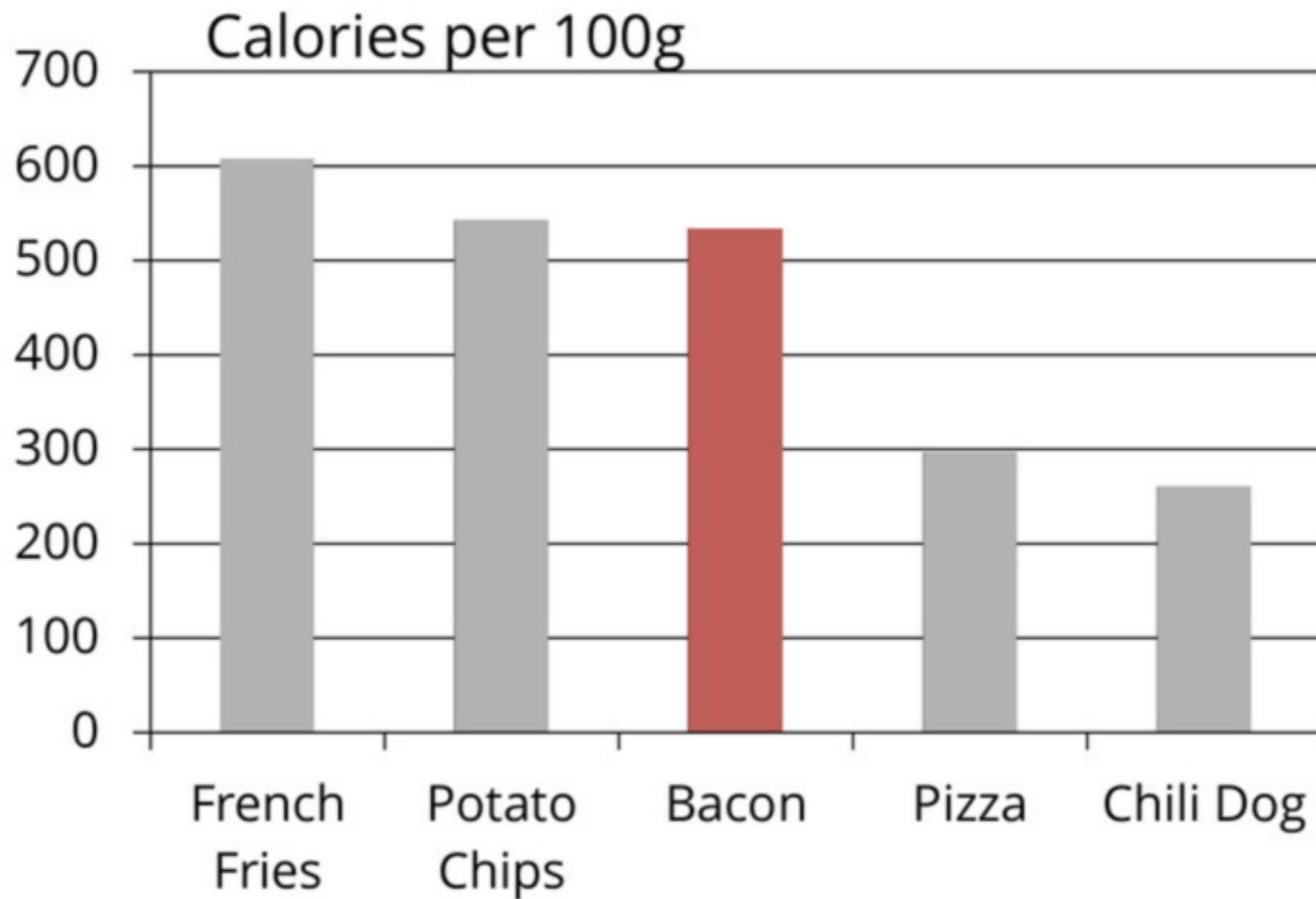
Remove bolding



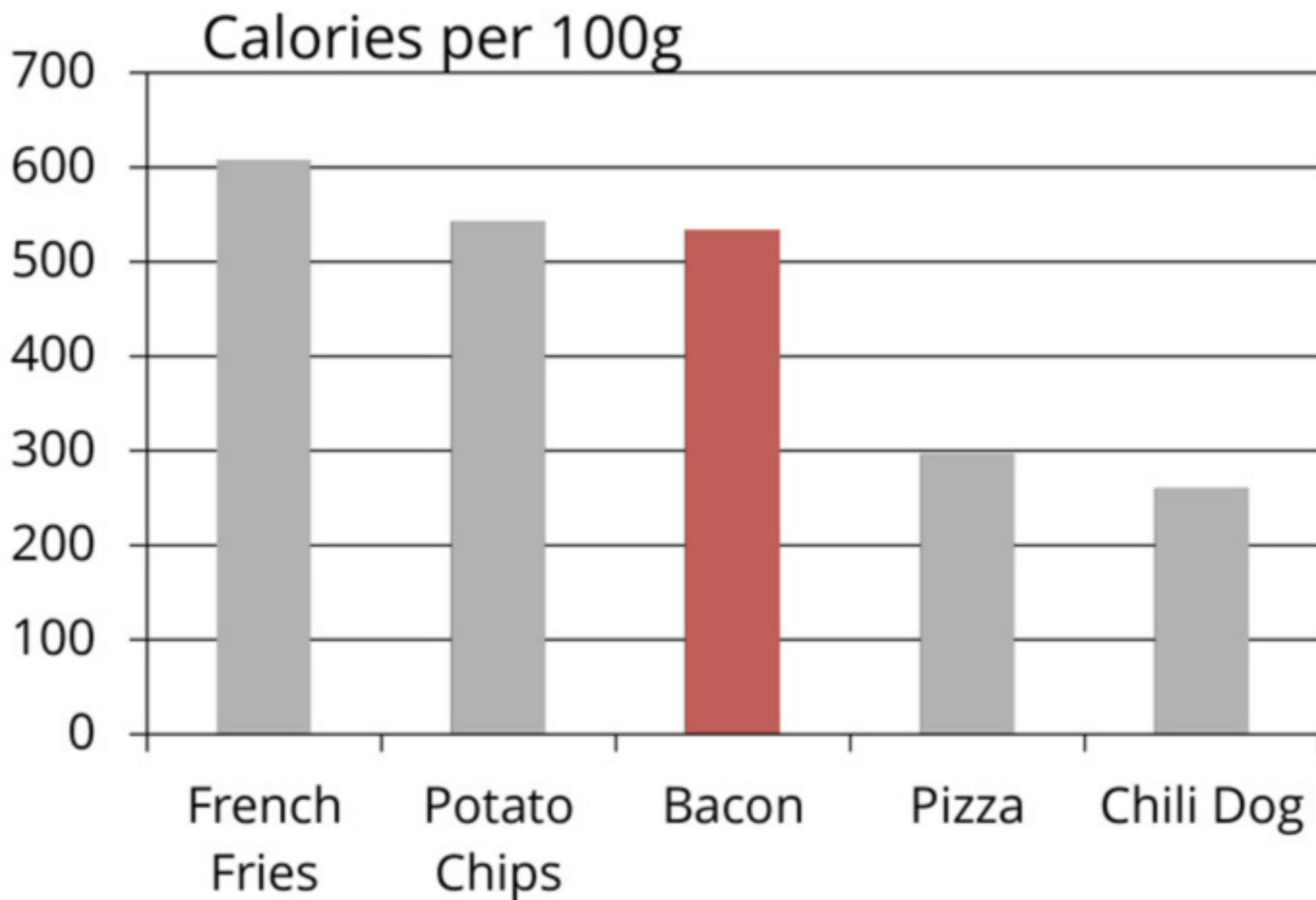
# Remove bolding



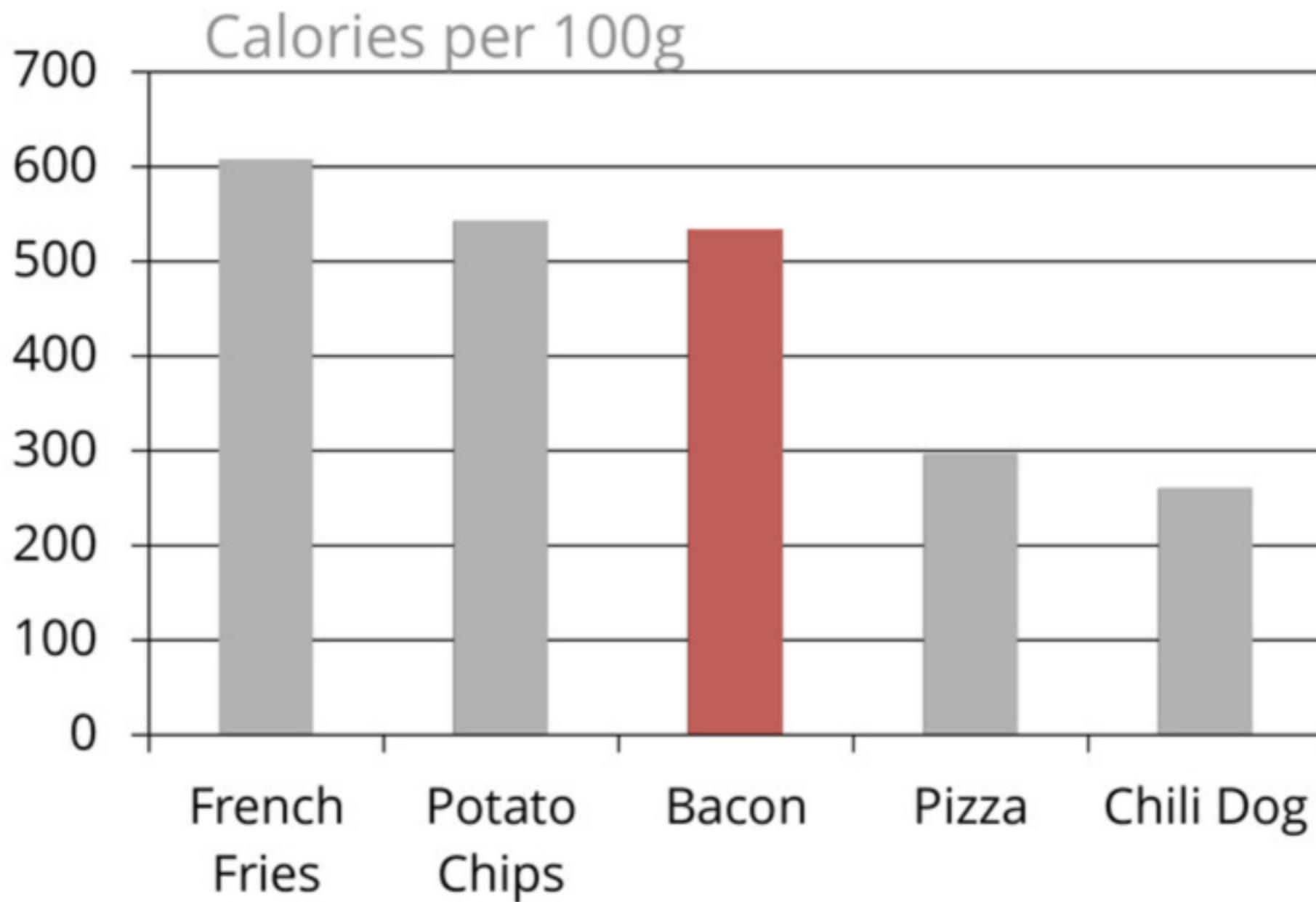
# Remove bolding



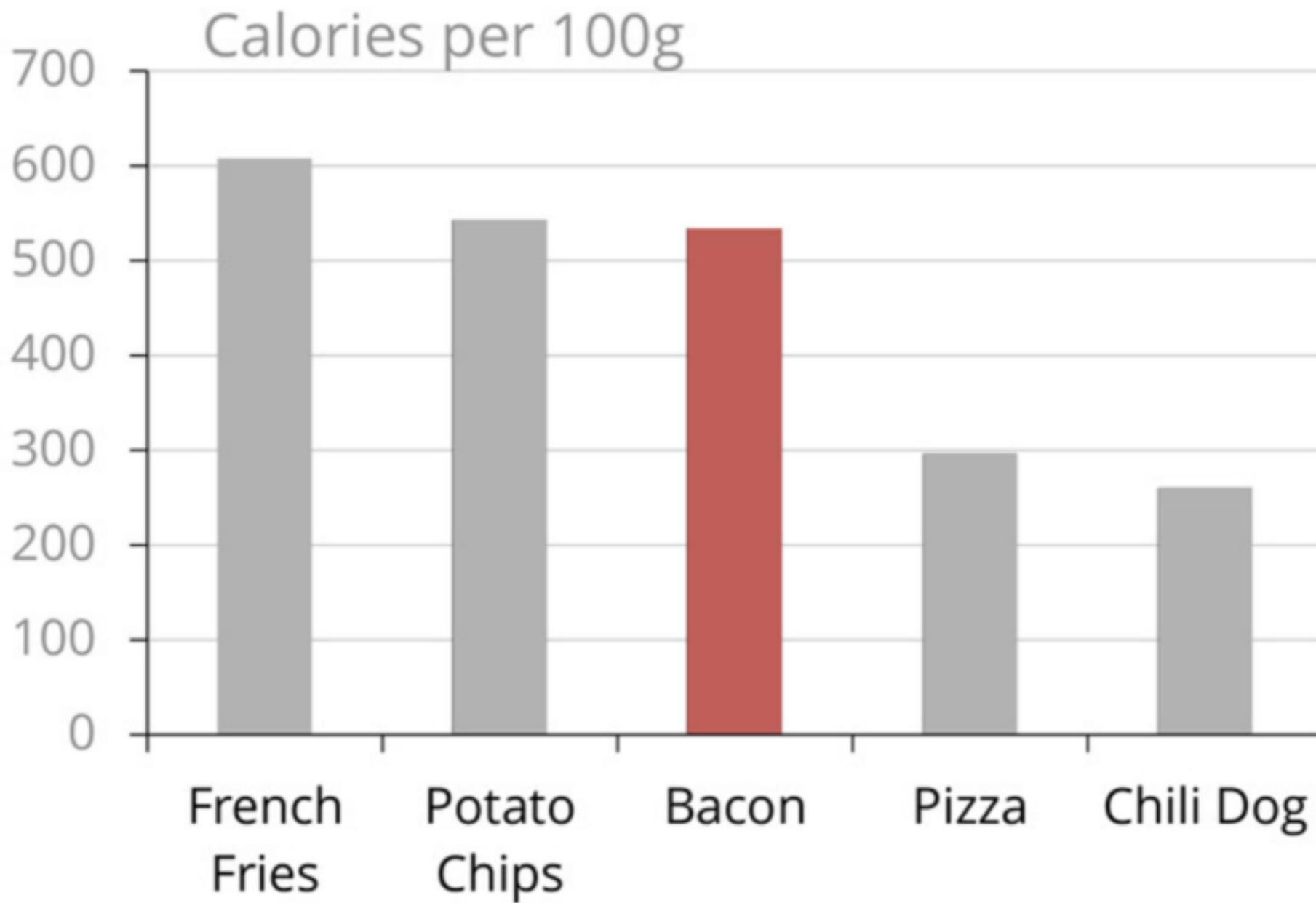
# Lighten labels



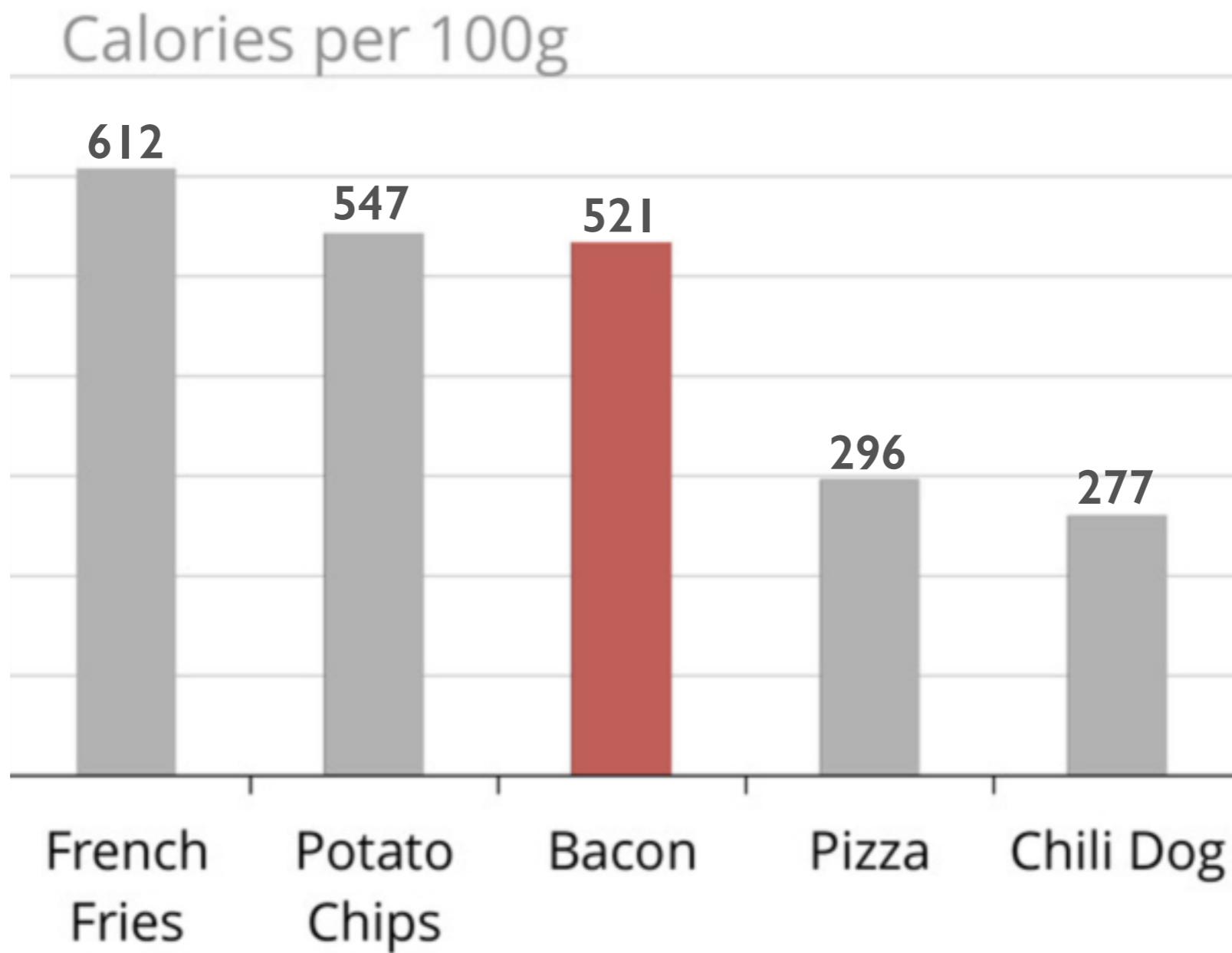
# Lighten labels



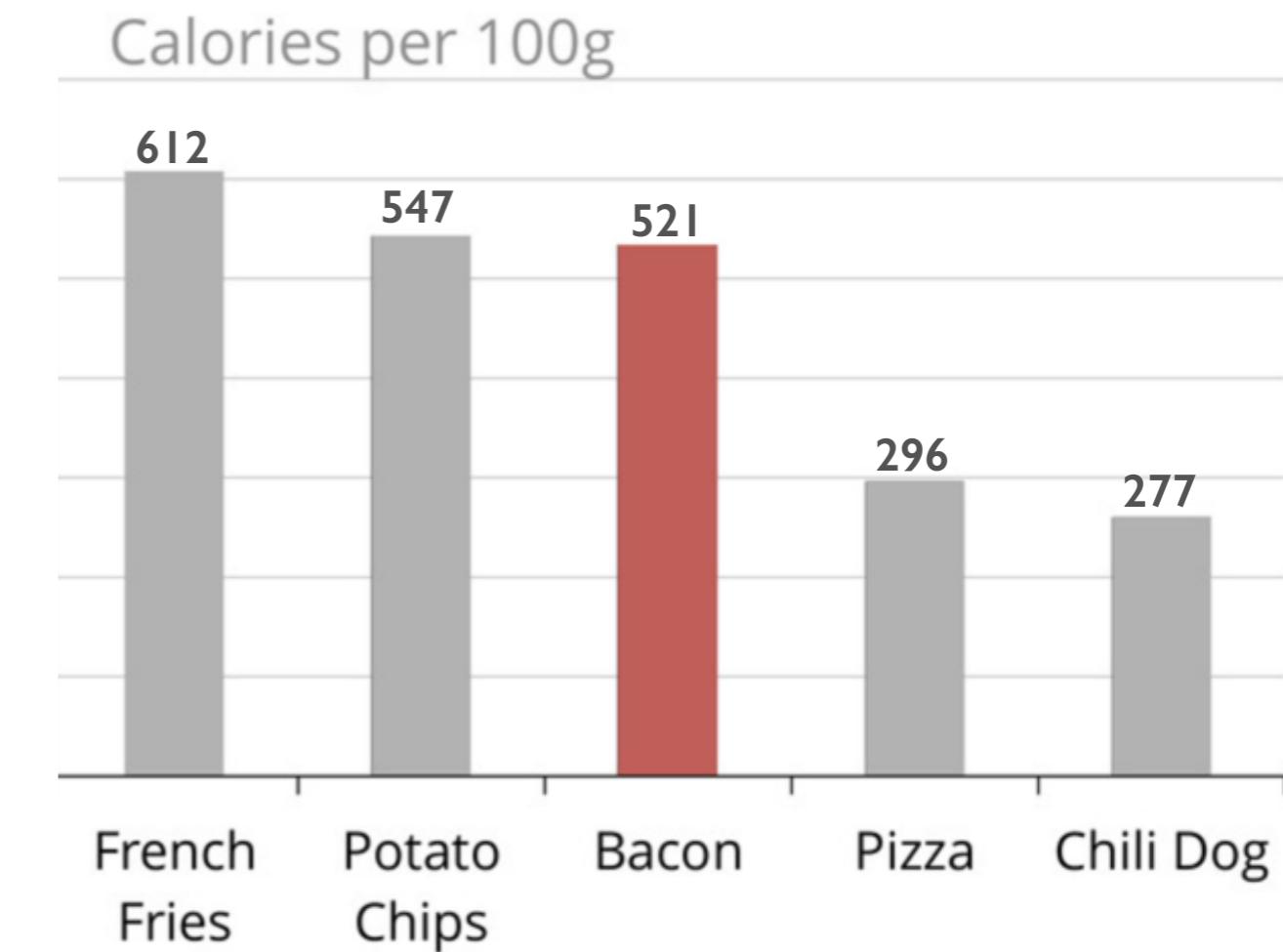
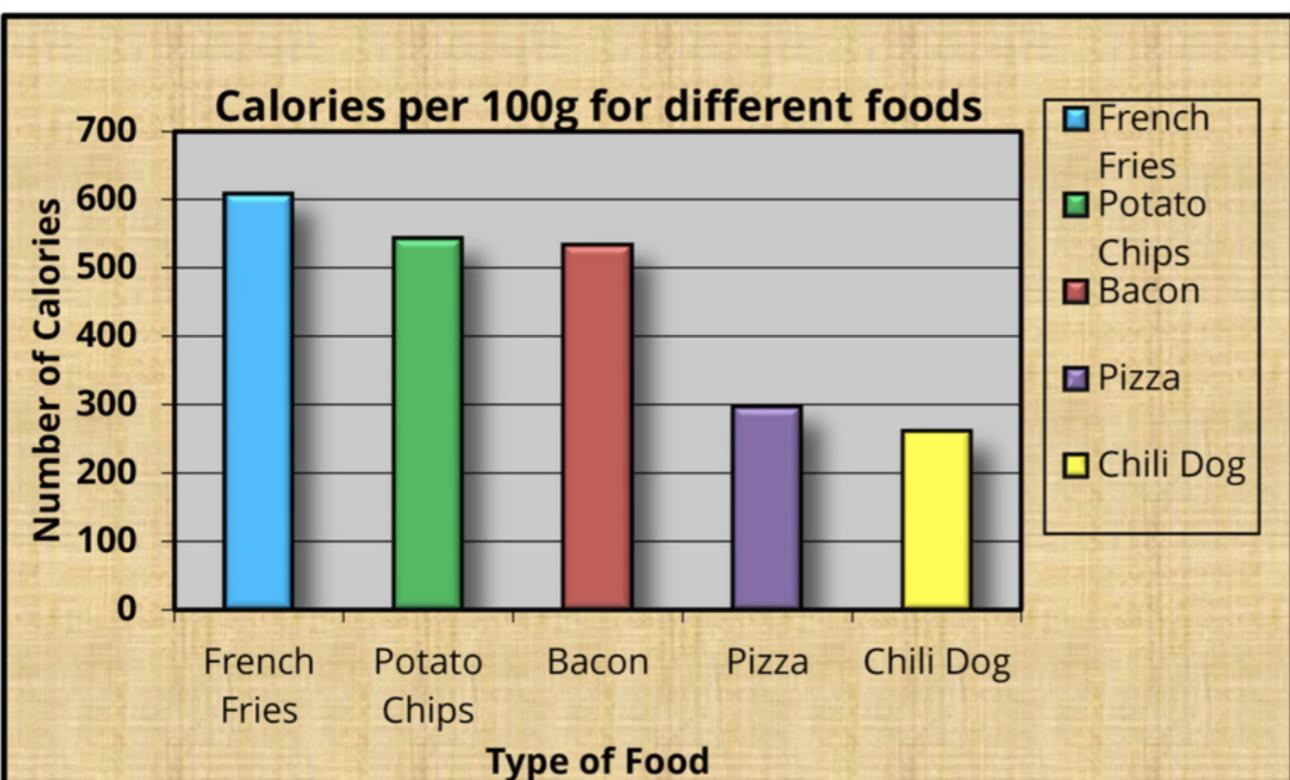
# Lighten lines



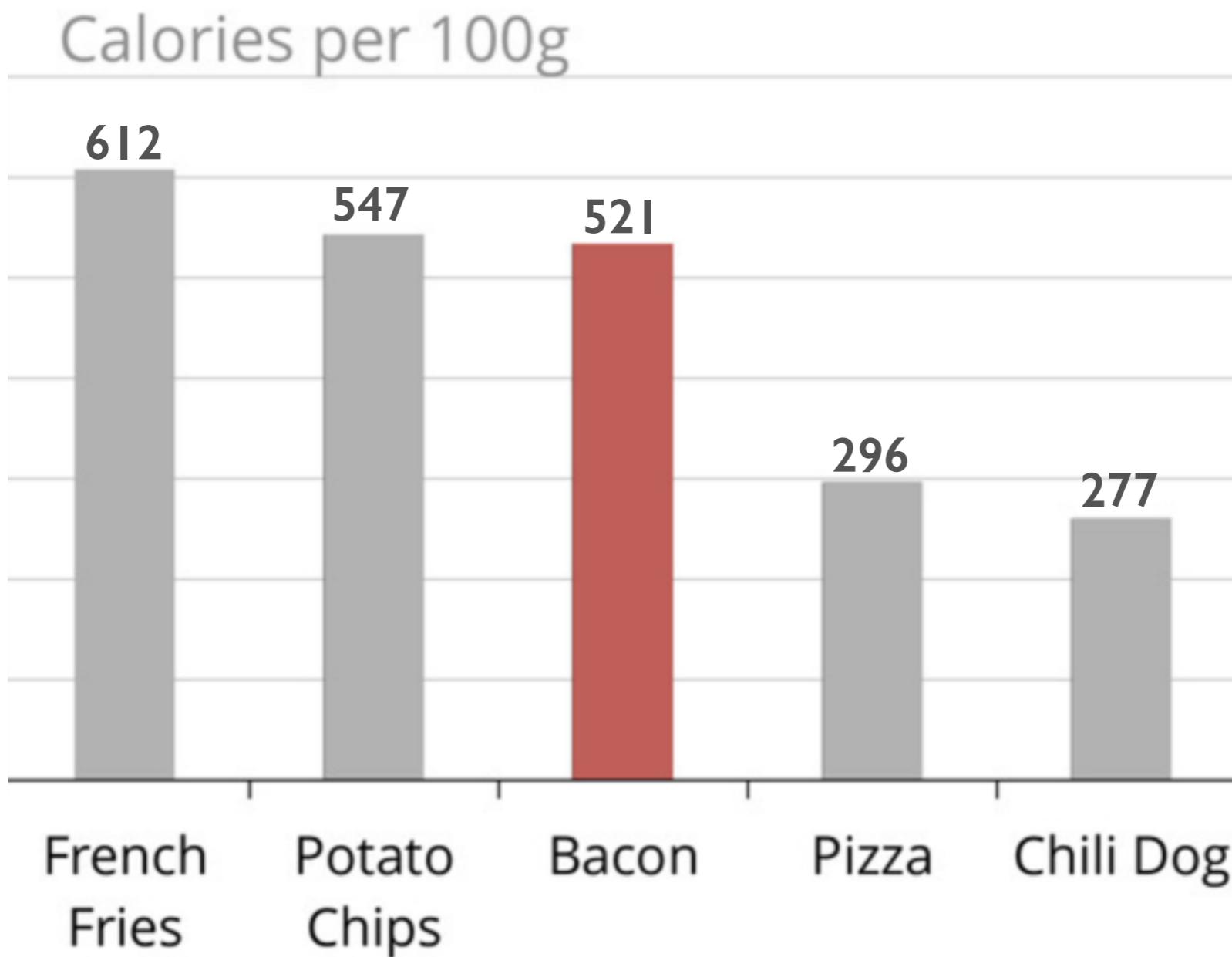
# Direct label



# DATA-INK RATIO



# MEANINGFUL TITLE

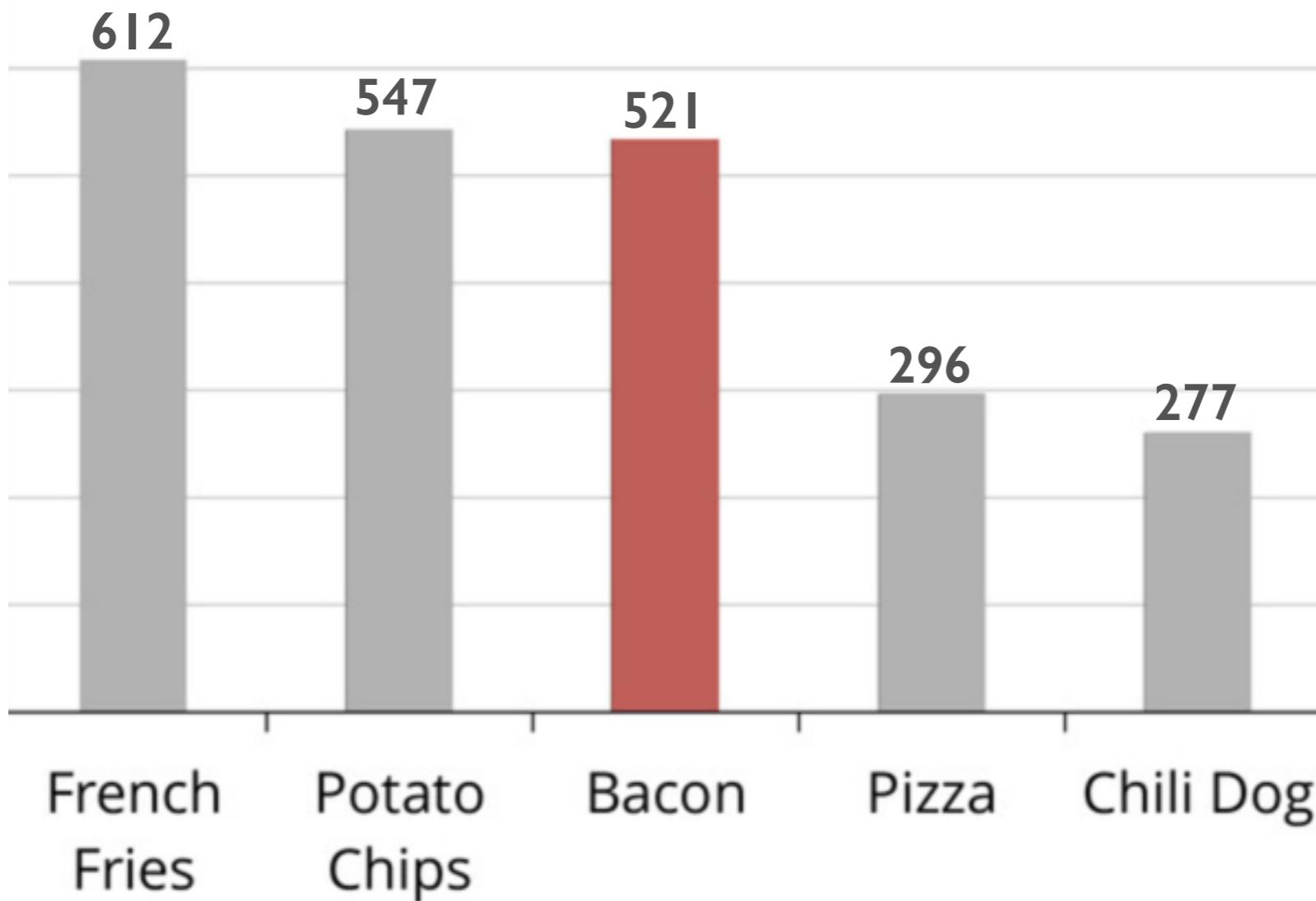


# MEANINGFUL TITLE

Calories per 100g.

Eating 100g of **bacon** is more than 1/4 of the recommended energy daily intake.

---



# VISUAL HIERARCHY OF INFORMATION

## Top 10 design concerns



# CONVEY THE MESSAGE

**Demonstrating effectiveness** is most important consideration when selecting a provider

In general, **what attributes are the most important** to you in selecting a service provider?

(Choose up to 3)



Survey shows that **demonstration of results** is the single most important dimension when choosing a service provider.

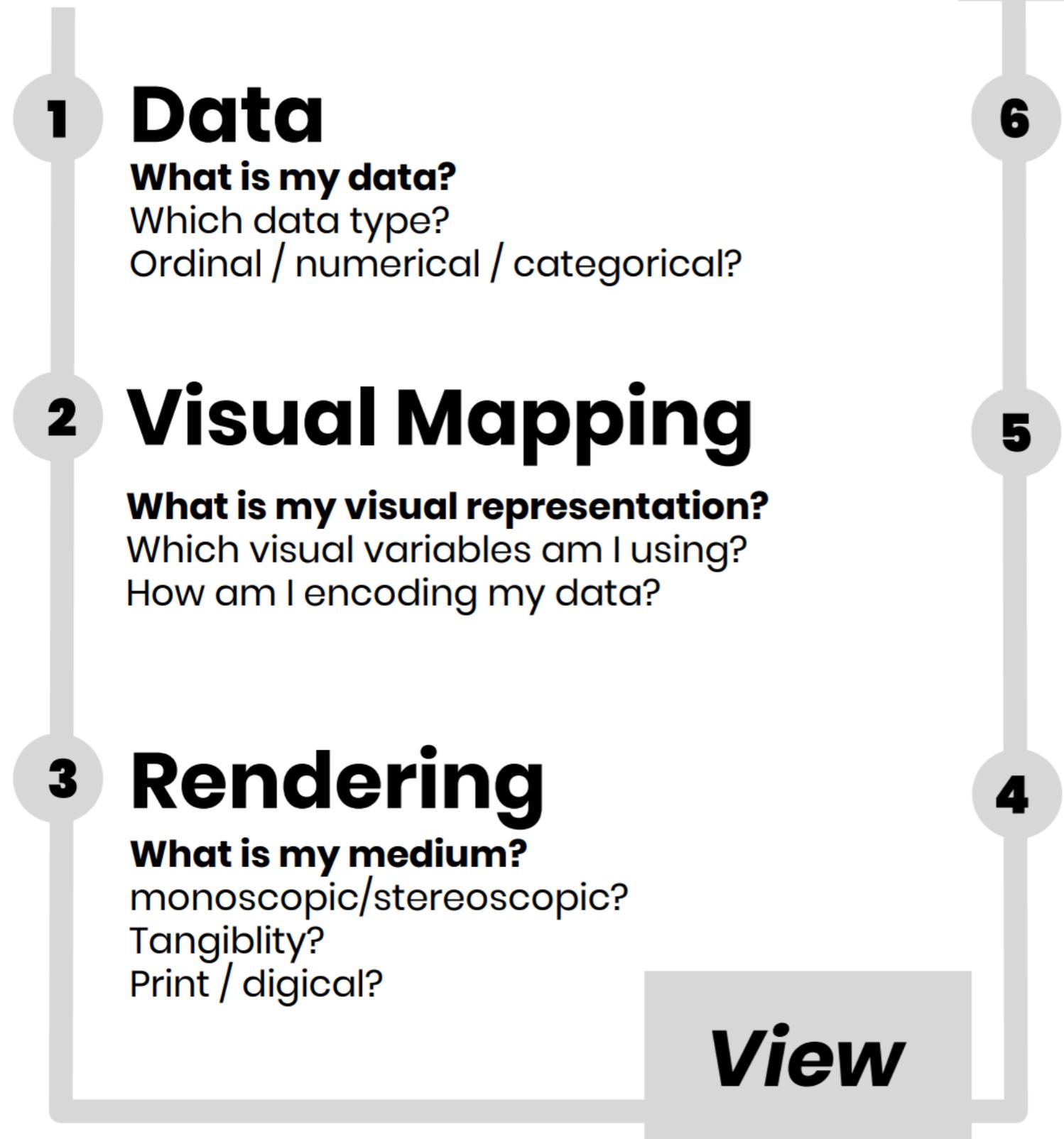
**Affordability** and **experience working together previously**, which were hypothesized to be very important in the decision making process, were both cited less frequently as important attributes.

Data source: xyz; includes N number of survey respondents.  
Note that respondents were able to choose up to 3 options.

**PUTTING IT ALL TOGETHER**

# Encoding: designer

# Decoding: user



# DATA-DRIVEN STORYTELLING

Provide **context**

Persons, collection, importance, questions, ...

Explain **processes**

Transformation, selection, aggregation, ...

Explain **insights**

Trends, outliers, groups, comparisons, ...

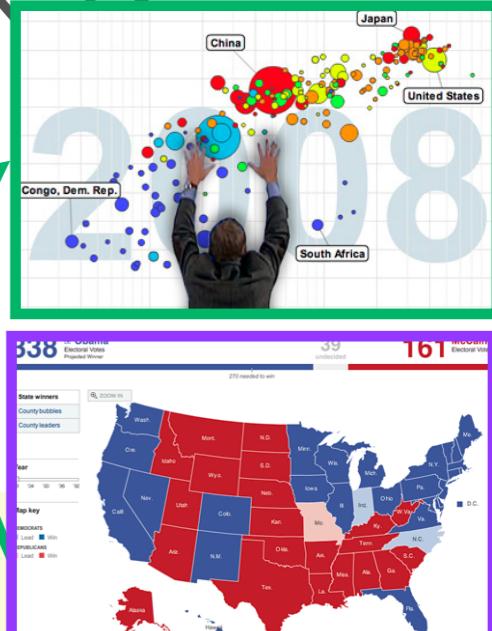
Explain **visualizations**

Talk to an **audience**

**Conclude**

Take home message

# VISUALIZATION ENCODING



## Visualization Encoding Pipeline

Structured Data map to Visual Attributes draw as Marks plot on a Layout

Nominal

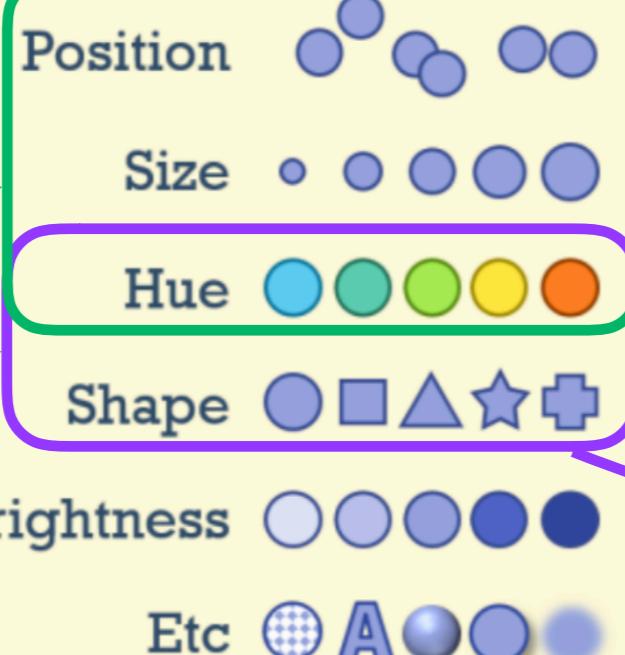
Apple, Banana, Pear

Ordered

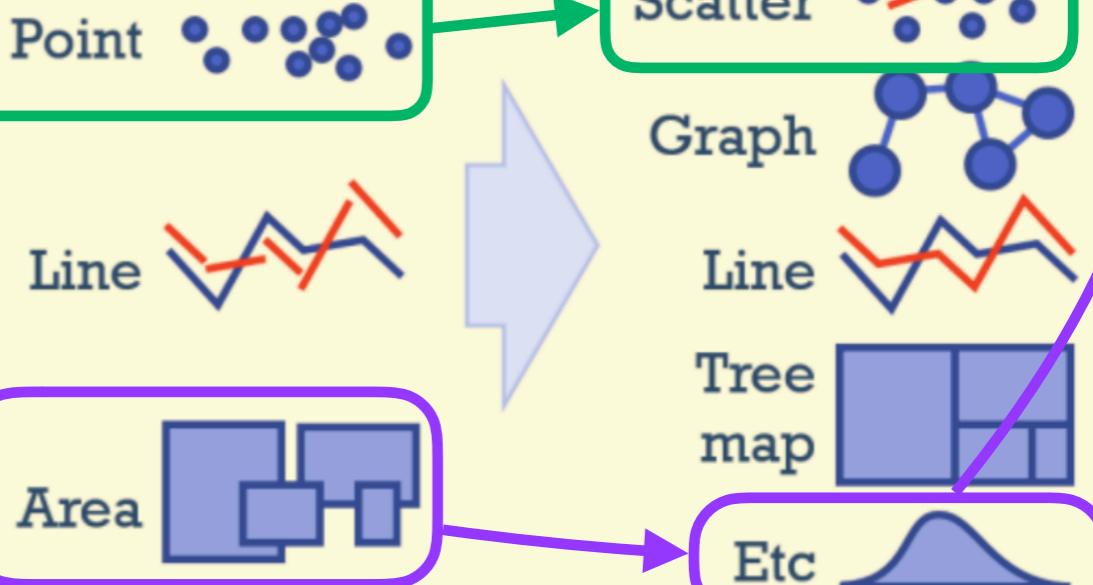
Mint, Good, Fair, Poor

Quantitative

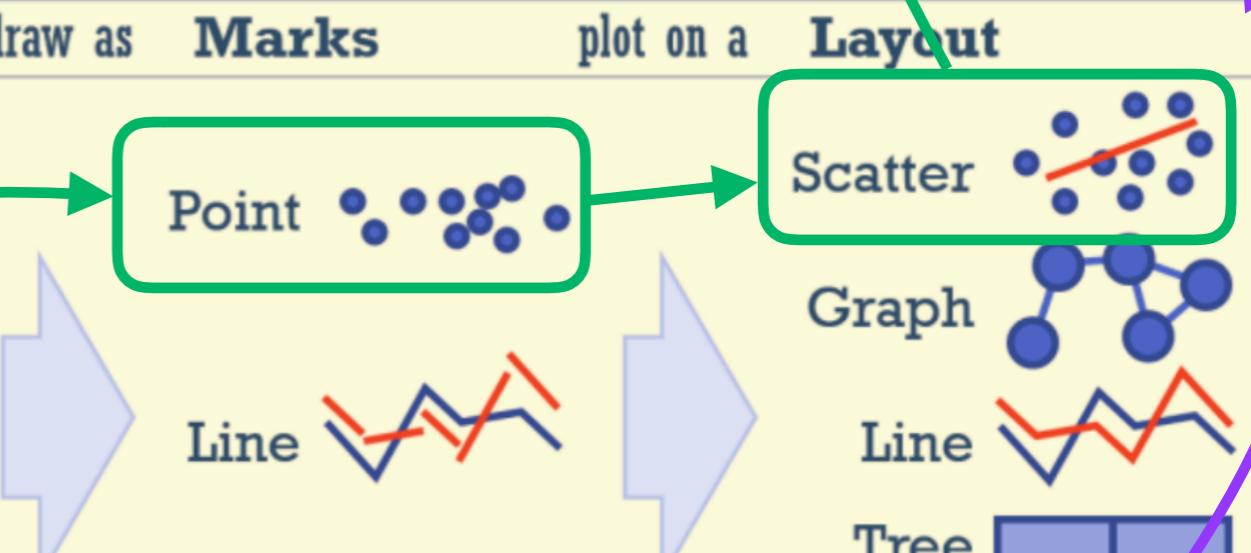
0, 3, 4.2, -31.2,  $6.6 \times 10^6$



Point



Marks



plot on a Layout

# KNOW YOUR AUDIENCE'S TASK

“Every [designed] thing is best at something, and worst at something else”

— Bill Buxton

"A tool that serves well for one task can be poorly suited for another, for exactly the same dataset. The task of the users is an equally important constraint for a vis designer as the kind of data that the users have"

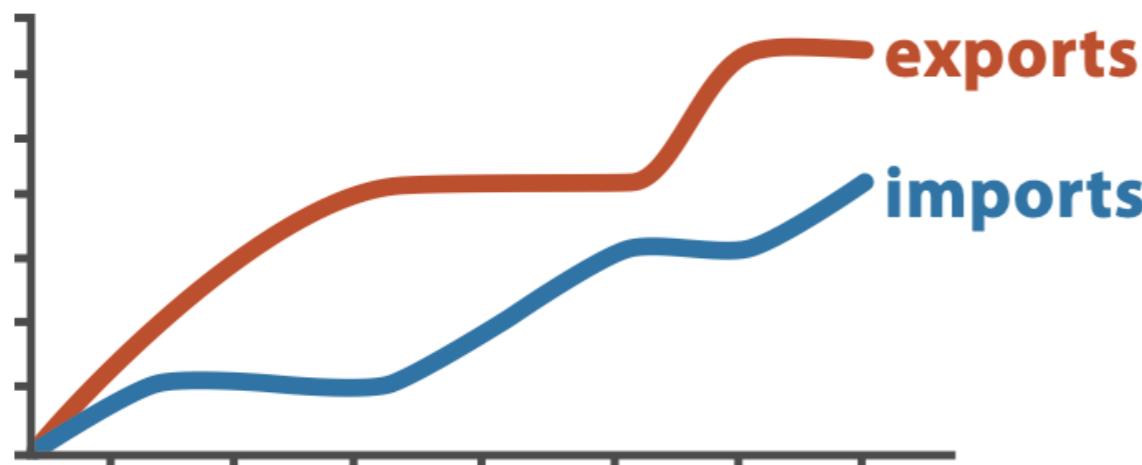
— Tamara Munzner

# DON'T JUST DRAW WHAT YOU'RE GIVEN!

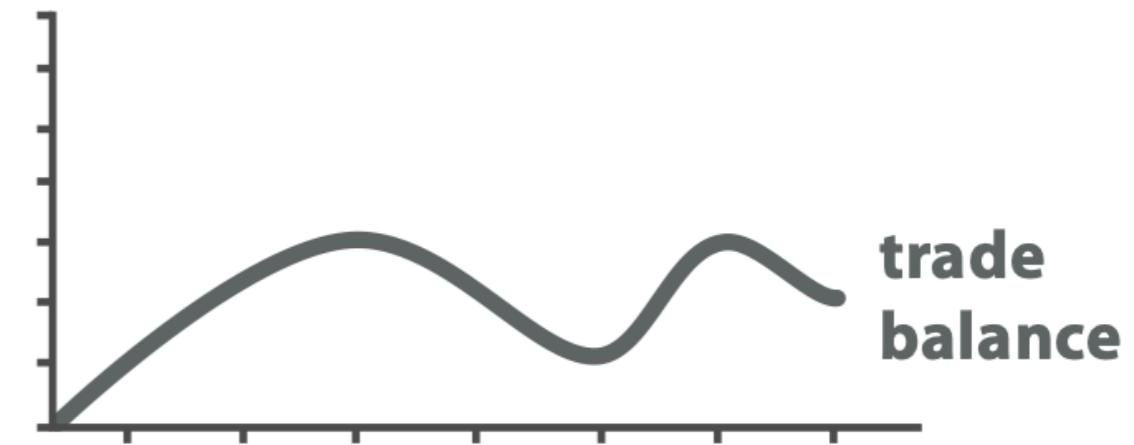
Decide what the right thing to show is

Create it with a series of transformations from the original dataset

Draw that



Original Data



$$\text{trade balance} = \text{exports} - \text{imports}$$

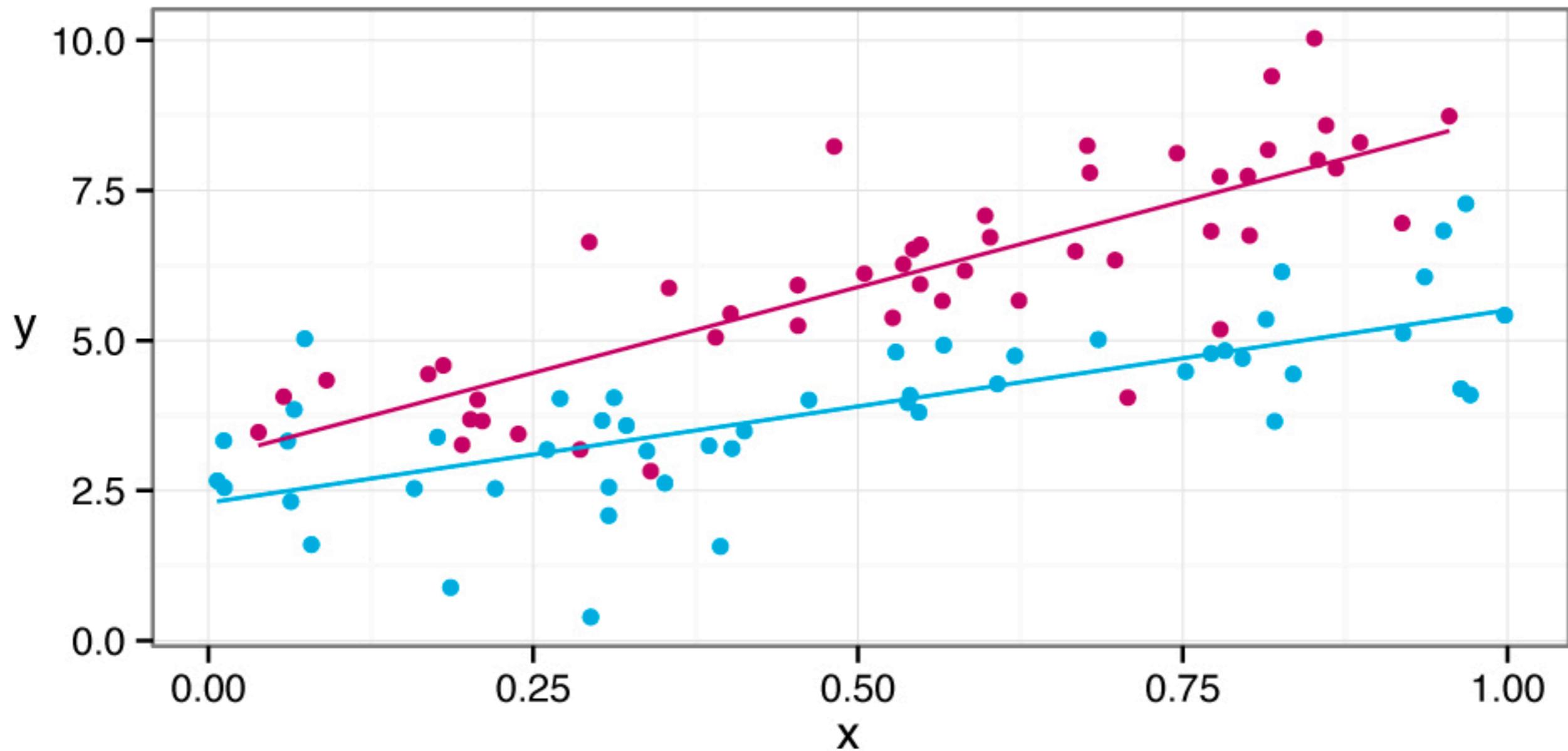
Derived Data

# DON'T JUST GRAPH IT, FIT IT!

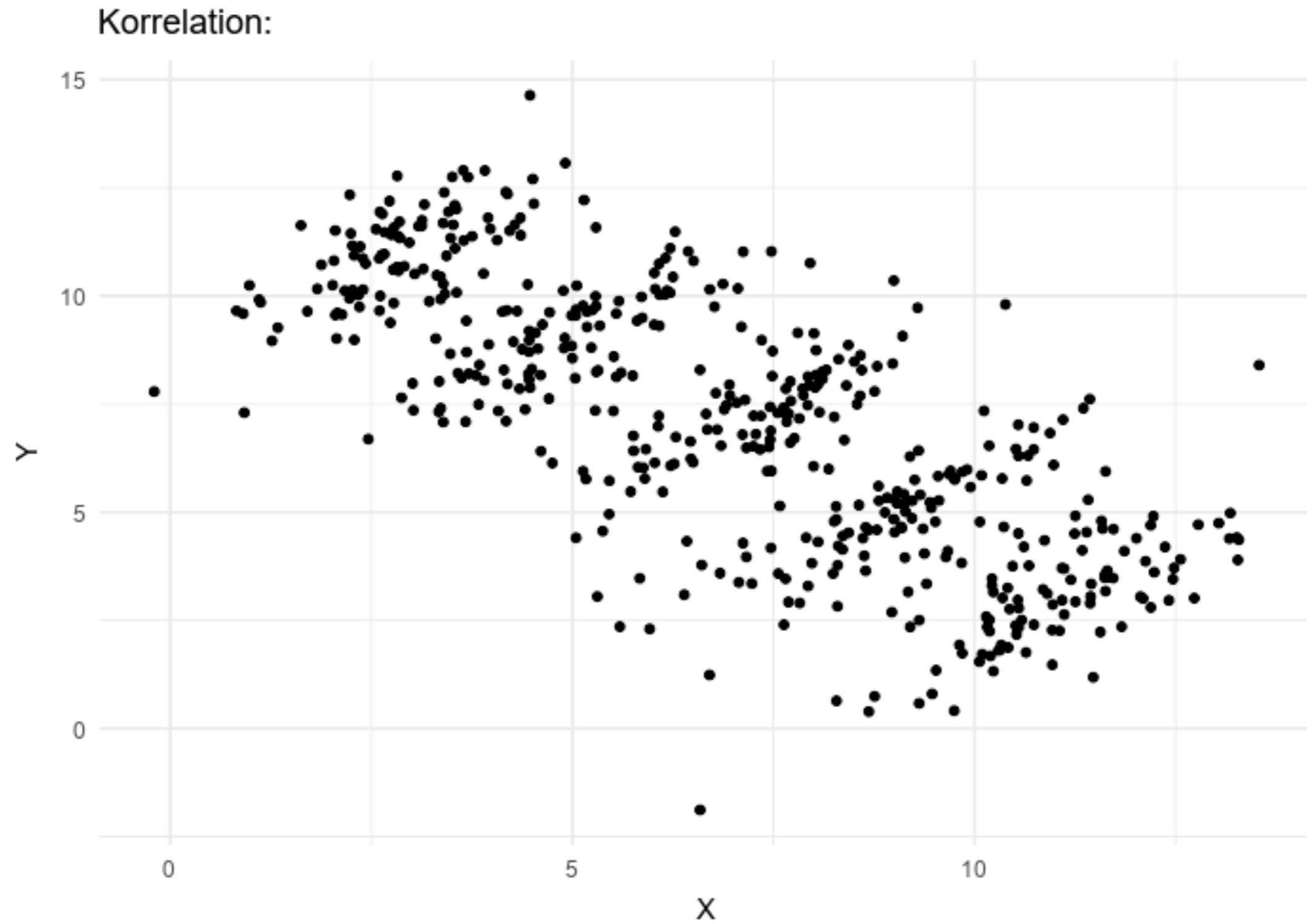
Graphing is needed, to see the details of the raw data

Fitting data to mathematical functions is needed too!

You can display both in the same graph...



# ITERATIVE FITTING

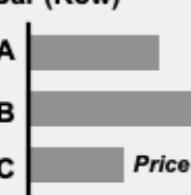
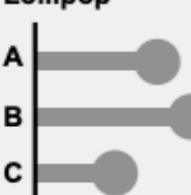
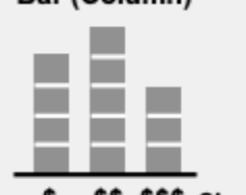
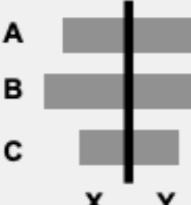
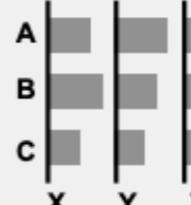
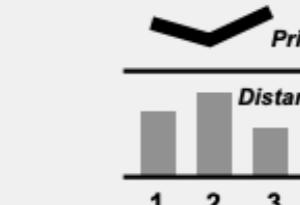
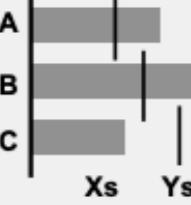
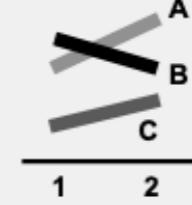




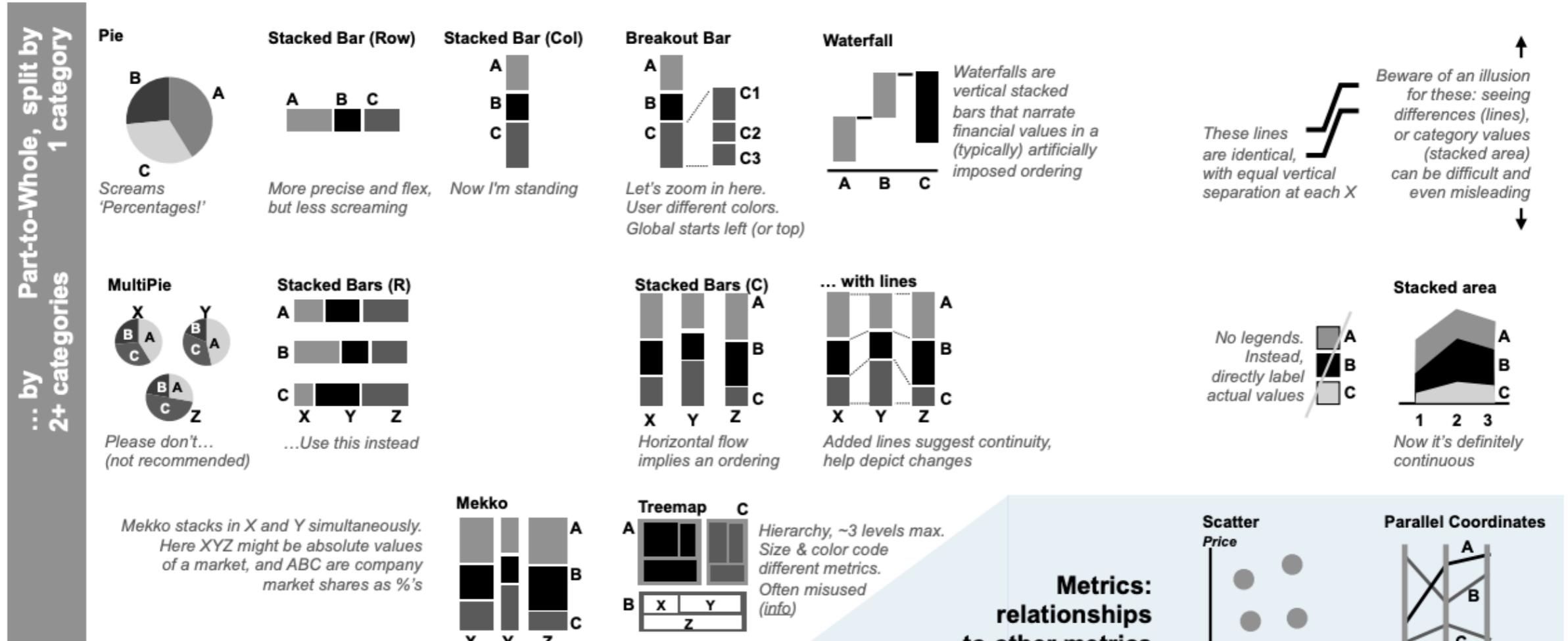
Examples, lots of  
examples!

Fundamentals

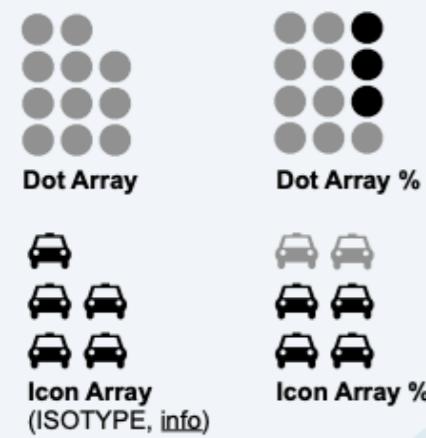
# WHICH VISUALIZATION?

	<b>Categories</b>	<b>Ordered Categories</b>			<b>Continuous Metrics</b>	
Metric, split by 1 category	Bar (Row)	Lollipop	Dot Plot	Bar (Column)	Bar (Column)	Area
						
... by 2+ categories	Bar Table X,Y	Mirror Bar	Bar Table X,Y,Z,...	Bar Table X,Y, Delta	Bar Table	Bar Line Table
						
	Compare X to Y, 'Small multiples'	Compare X to Y	Compare as many as you like	Comparisons are slow. Plot critical Deltas explicitly	Compare a continuous metric across a category	Compare two metrics
	Benchmark Bar	Benchmarks Bar	Interleaved Bar		Slopegraph	Dual Axis
						
	Compare X to a benchmark	Compare X to Y. Fancier version called a 'Bullet graph'	Compare X to Y (not recommended)	Interleaves two categories into one spatial dimension. Typically better to use Bar Table (above) instead	With two values, slope encodes delta	Use (above) instead. Crossings here are salient, but meaningless
						Compare many. Getting spaghetti? Use Line-Table (above)

# WHICH VISUALIZATION?



**Look at this number  
Just look at it.**

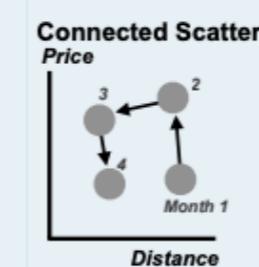
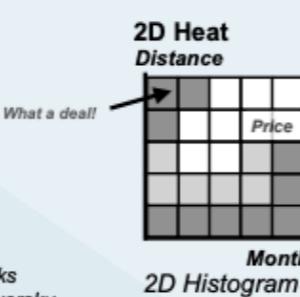


**43%**

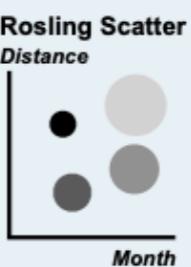
Huge Number

Connecting Lines imply continuous data  
Same data, ask people, "What do you see?"  
Bars Lines  
USA Dutch  
"Dutch people are taller than Americans"

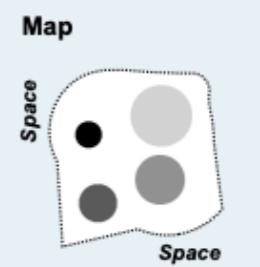
2D Heat Distance  
What a deal!  
Zacks & Tversky, 1997  
2D Histogram  
Distance



The power of a scatterplot, with points connected by time (info)



Watch Hans Rosling's famous TED talk. He adds two more metrics, and moves it over time



Maps and scatterplots share the same DNA  
Color = Flight Time  
Size = Price

# VISUALIZATION CHEAT SHEET

<https://visualizationcheatsheets.github.io/>



## Cheat Sheets for Visualization Techniques

[visualizationcheatcheats.github.io](https://visualizationcheatcheats.github.io)

[Home](#)

[Download all \(PDF+PNG\), 76MB](#)

[Download guidelines \(PDF\), 8MB](#)

## Paper

**Cheat Sheets for Data Visualization Techniques:** Zehzhong Wang, Lovisa Sundin, Dave Murray-Rust, Benjamin Bach, *ACM Conference on Human Factors in Computing Systems (CHI)*, 2020

## By Type

[Anatomy](#) | [Introduction](#) | [Construction](#)

[Visual Pattern](#) | [Pitfalls](#)

[Well-known Relative](#) | [False Friends](#)

## By Visualization



Boxplots



Confluence Graphs



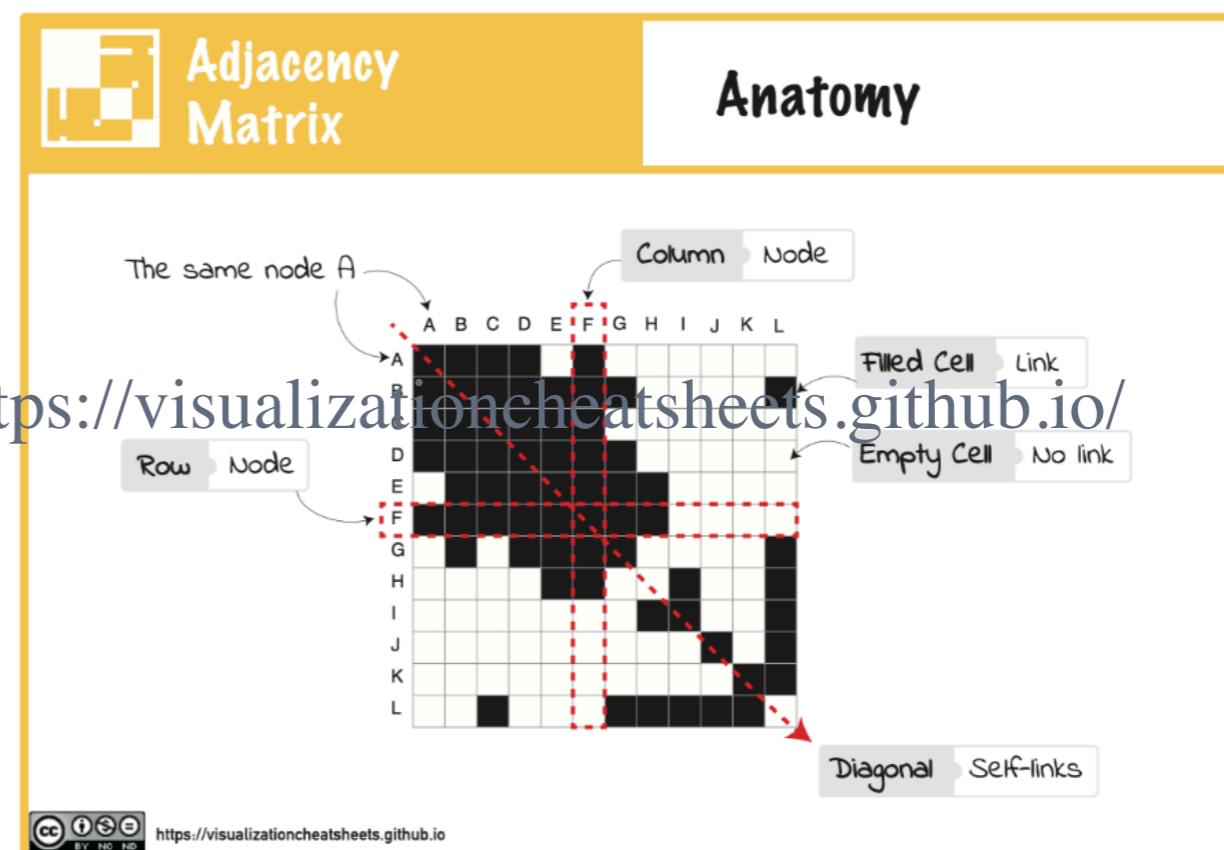
Adjacency Matrix



Parallel Coordinates

## By Type

### Anatomy



Anatomy explains the visual elements of a visualization technique, their composition, their specific terminology, and how they relate to the data. The visual components of a visualization can include individual visual marks as well as groups of marks, axes, locations in a visualization etc.

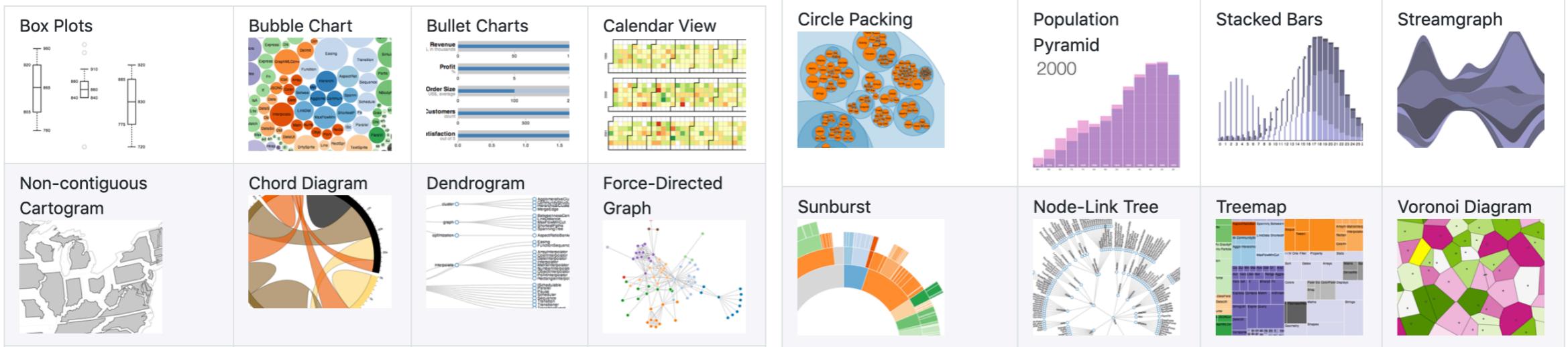
### Construction





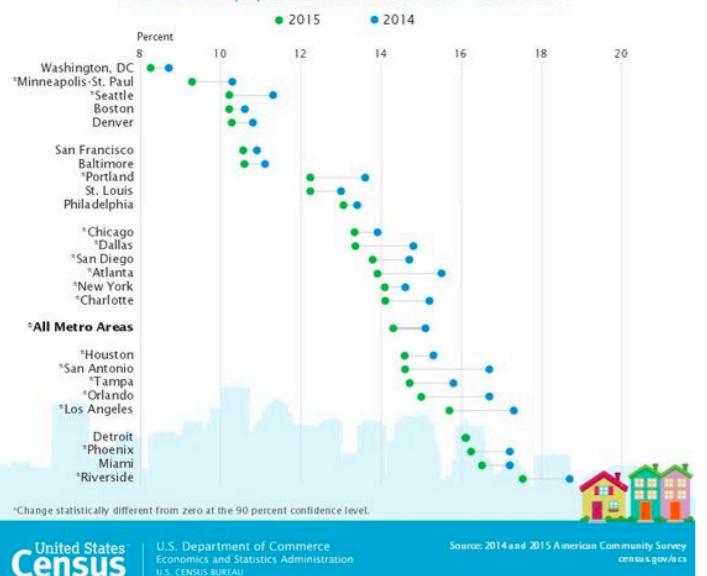
Examples, lots of  
examples!

Fundamentals

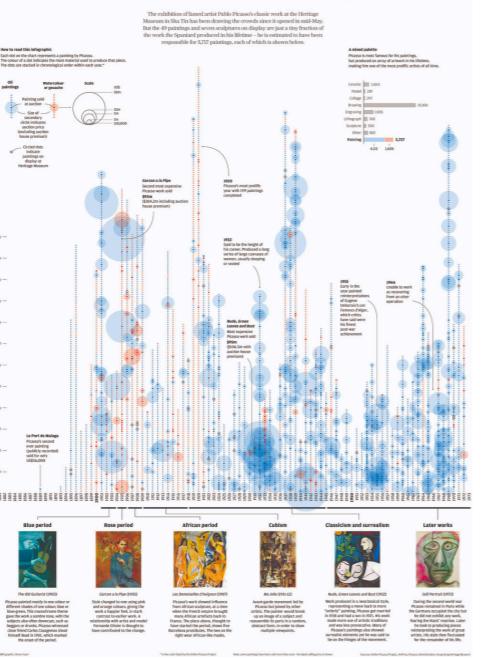


## Poverty in the United States

The 25 most populous metro areas: 2014 and 2015



## Picasso's paintings



## In the Shadow Of Foreclosures

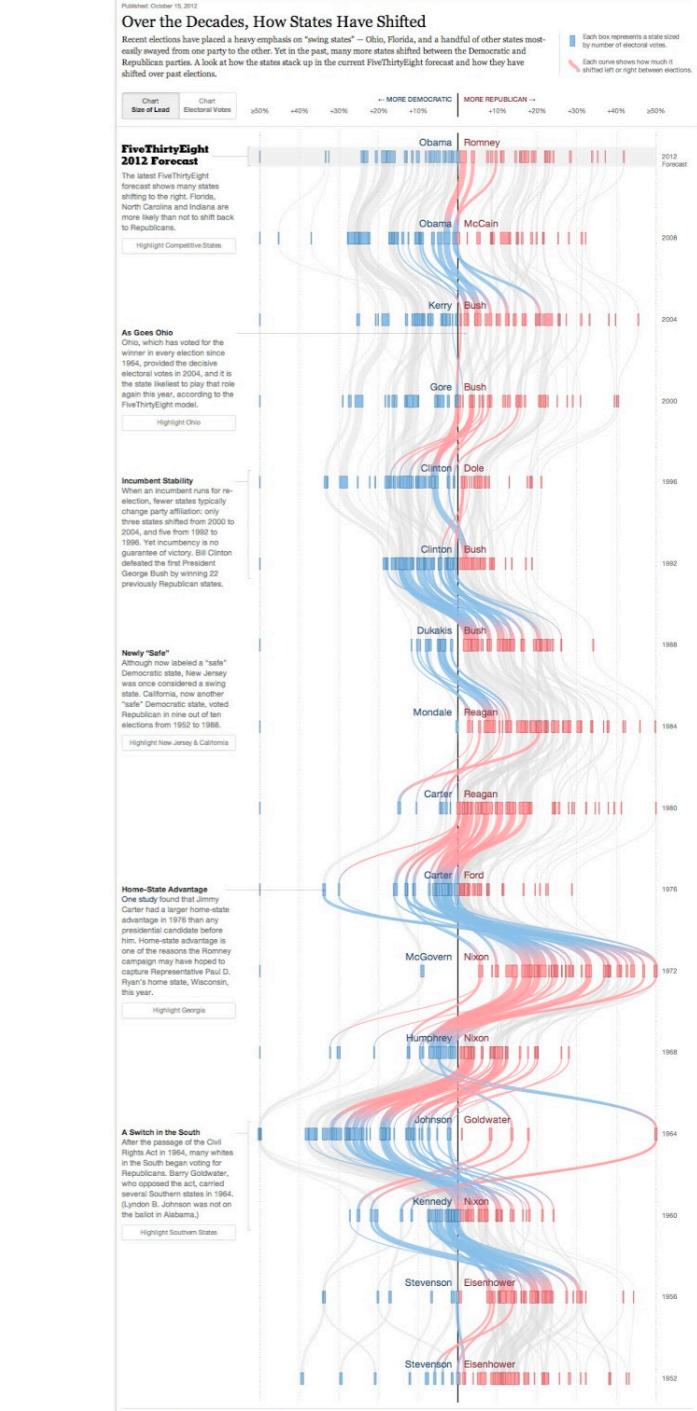
In the subprime mortgage squeeze, some regions are feeling the pain more acutely than others. Although most Southern metropolitan areas have high percentages of subprime mortgages, homeowners in those areas have been forced to pay their bills, so subprime foreclosure rates are low.

Not so in the Rust Belt, where subprime mortgages are less common but foreclosure rates are sky-high, mostly a result of overbuilding.

And overbuilding in regions of Florida, California and other states with large real estate bubbles has led to speculators buying up several homes at once, hoping that their values would rise. Getting second loans was all too easy.

But as housing prices fall, foreclosing prices are all too hard, and many economists believe that foreclosures will continue.

"The collapse will affect other markets," said New York City and D.C.-based David Baker, co-director of the Center for Economic and Policy Research. "Suburban areas near those

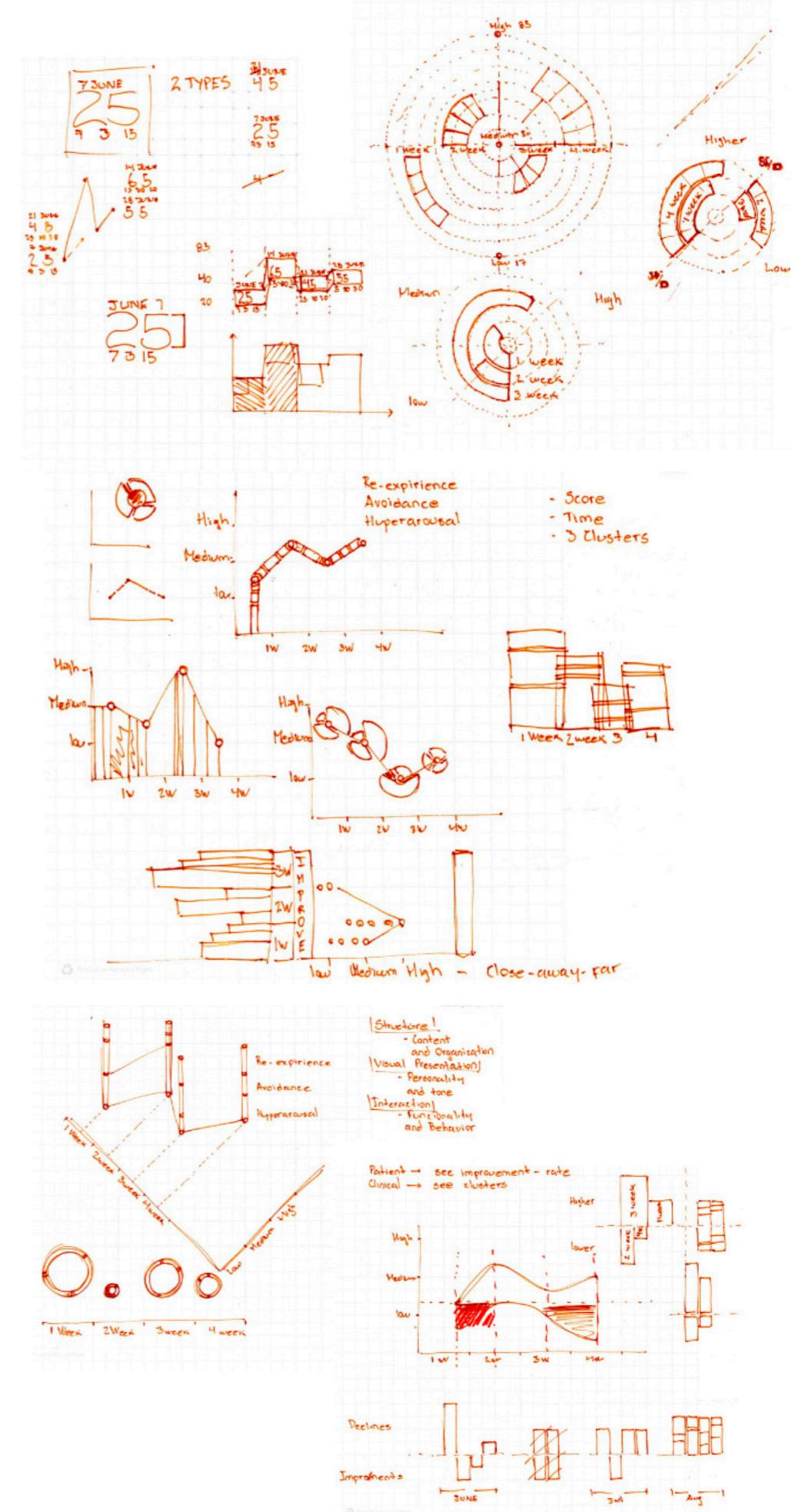


# SKETCH!

Don't commit to a solution immediately; (easy to make, easy to discard)

Prototype ideas in early stages

Share and discuss your designs with other people

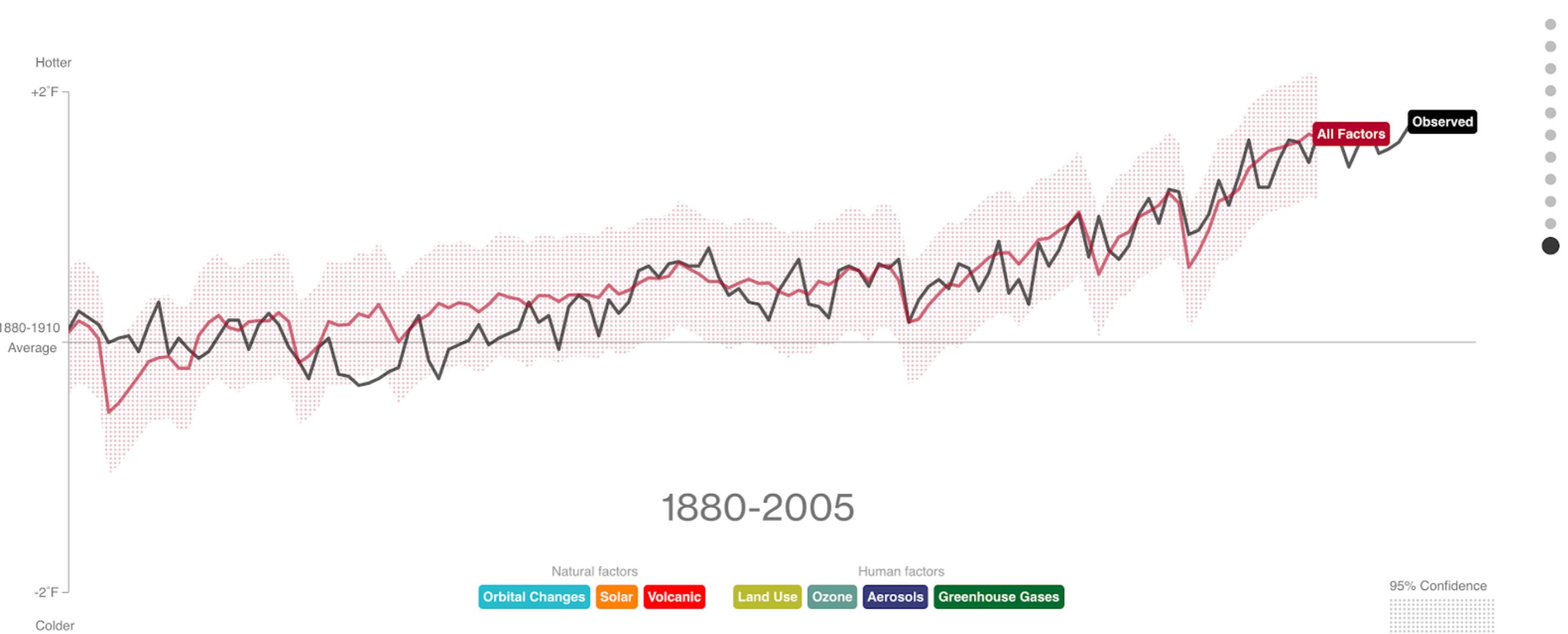


# NARRATIVE PATTERNS

## Compare and Contrast

Putting the possible natural and human causes of climate change alongside one another makes the dominant role of greenhouse gases even more plainly visible.

The only real question is: What are we going to do about it?



# NARRATIVE PATTERNS

<http://napa-cards.net>

## NAPA Cards Preview

What is this?

EXAMPLE: [Gannett Iraq war maps for CNN](#)

EMPATHY [4] ENGAGEMENT [3] FRAMING [11] FLOW [3] ARGUMENT [3]

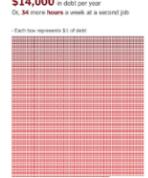
children's college chances

Narrative Patterns for Data-Driven Storytelling

### Rhetorical question

#### Can You Live on the Minimum Wage?

By ANDREW ABBAS / JUL 14, 2014



#### HOW

Title the piece using a question that makes a point.

#### WHY

Implicitly drives the viewer to the takeaway of a more open ended/exploratory piece.

#### EXAMPLE

New York Times graphics: Can you live on the minimum wage?

### Familiar Setting



#### HOW

Embed the entry point of a narrative in an environment which the viewer is familiar with.

#### WHY

To lower the entry barrier and make information personally relatable.

#### EXAMPLE

OECD: Regional Well-Being

### Call to action

#### Enjoy a water sensible diet.

Start today and spread the message!



#### HOW

Start with presentation of problem that needs solving or situation that needs changing. End with call to action.

#### WHY

Provides motivation; offers outline for narrative; gives sense of purpose.

#### EXAMPLE

Angela Morelli: Virtual water

### Incorporating the audience



#### HOW

Audience becomes part of the narrative.

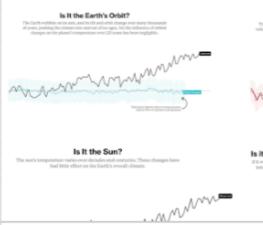
#### WHY

Increase empathy, facilitate personal comparisons. Helps make the data and information more personally relevant. To allow reader to use herself as measuring device.

#### EXAMPLE

Ubisoft: Collapse

### Repetition



#### HOW

A phenomenon is re-presented multiple times with changes to the main variable/dimension (the data) or the time frame or the animation pattern (see the example) while other elements of the story don't change. Note: this could be like running multiple simulations with the data to see which is most effective (and in this case, the story would be more reader driven)

#### WHY

To show that the same phenomenon happens over and over again. It also strengthens the narrative through rhythm. Establishing a constant reference frame can help to emphasize change and differences.

Bloomberg: What's really warming the world?

### Juxtaposition



#### HOW

Showing two or more visualizations (picture-in-picture).

#### WHY

Compare cases or situations, difference, absence or contrast. Allows for individual interpretation.

#### EXAMPLE

David McCandless: Billion

### Gradual visual reveal



#### HOW

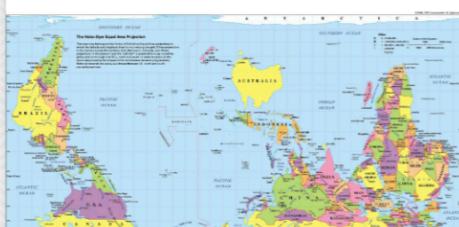
Use visual hierarchy to control sequence of perception of individual elements

#### WHY

Lets the story unfold in the viewer's mind while they read the graphic, to chunk the material to make it easier to absorb

#### EXAMPLE

### Defamiliarization



#### HOW

Present something known and familiar in a novel, unexpected way

#### WHY

To confront and challenge viewers' expectations; to highlight and question implicit assumptions. To force thinking differently about a well-known case.

### Convention breaking



#### HOW

Use or establish a graphical convention, then break it.

#### WHY

To engender surprise; to show extra-ordinary nature of data; to question the use of the convention; challenge assumptions.

### Breaking the fourth wall



#### HOW

Subject unexpectedly addresses the viewer.

#### WHY

To make a direct connection with viewer; to demonstrate the artificiality of the presentation; to challenge the objectivity of the observation. To make it clear that this is one interpretation. Surprise, so attention-getting.

#### EXAMPLE

### Humans behind the dots



#### HOW

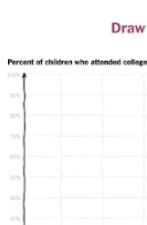
Show that there are concrete entities (e.g., a human being) behind, or at the heart of the data.

#### WHY

Make abstract data more relatable, and possibly establish an emotional connection between the viewer and the fate of the entities.

#### EXAMPLE

### Make a guess



#### HOW

Enable the viewer to make a guess before interacting with the data. No interaction: introduce the data. Directly addressed to the viewer: ask a question after. If interaction: provide feedback.

#### WHY

Stimulates the viewer to think and reason about the data. Encourages active participation.

#### EXAMPLE

New York Times graphic

## **REFERENCES**

# EXPAND YOUR HORIZON

## Lookup

Visualization Catalogs

[www.datavis.ca](http://www.datavis.ca)

<http://www.datavis.ca/milestones/>

(historical!)

[www.infovis-wiki.net](http://www.infovis-wiki.net)

[http://www.visual-literacy.org/periodic\\_table/periodic\\_table.html](http://www.visual-literacy.org/periodic_table/periodic_table.html)

Color

<http://colorbrewer2.org/>

Color Brewer

<http://tools.medialab.sciences-po.fr/iwanhue/>

I Want Hue

## To Reflect:

<http://hint.fm/>

F. Viégas & M. Wattenberg

[www.edwardtufte.com/bboard](http://www.edwardtufte.com/bboard)

E. Tufte

<http://fellinlovewithdata.com/>

*Fell in love with data*

E. Bertini

<http://datastori.es/> (podcasts)

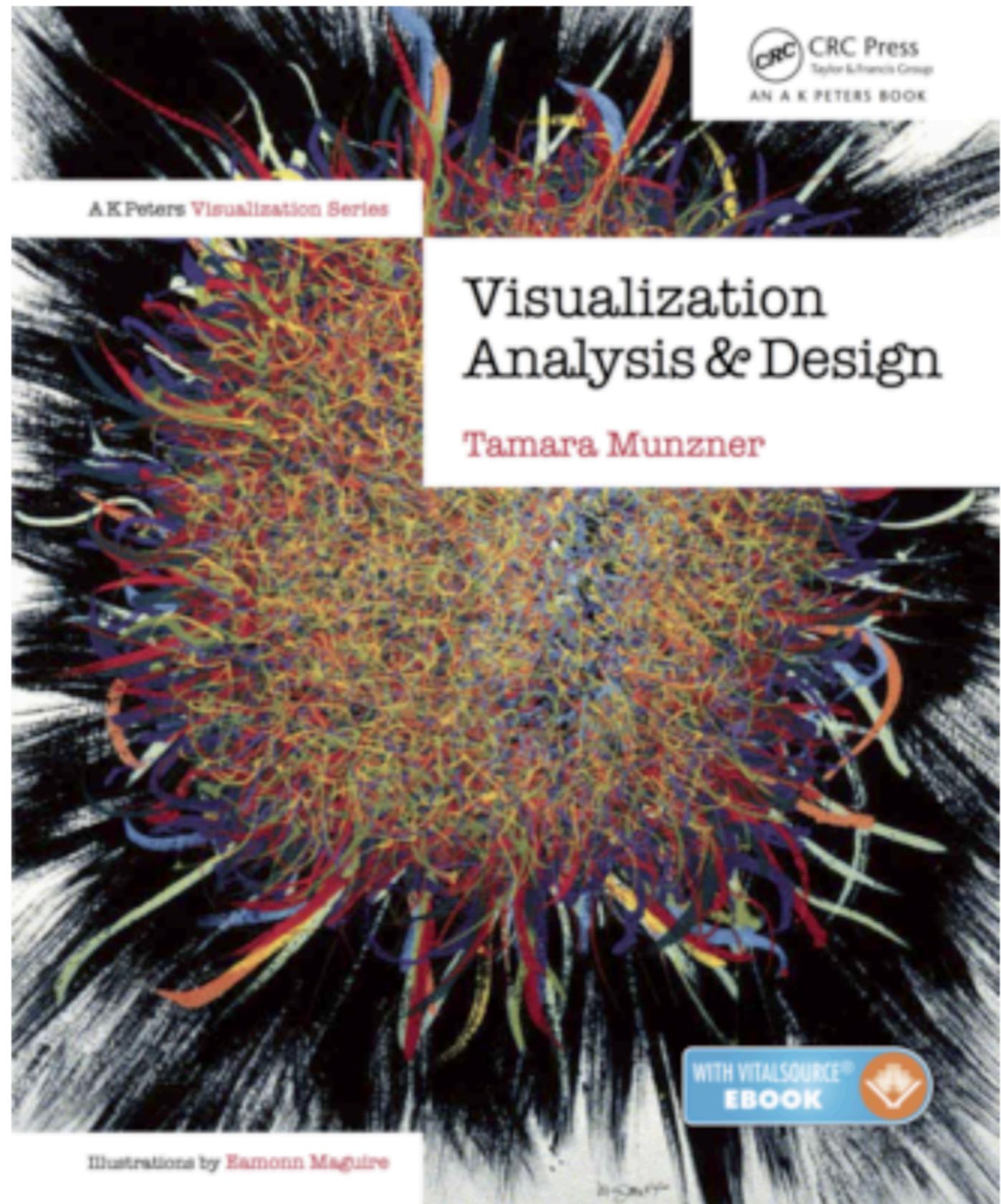
E. Bertini & M. Stefaner

<http://junkcharts.typepad.com/>

Junk Charts: learn from others (mistakes)

Text-book on data visualization  
(1<sup>st</sup> of its kind)

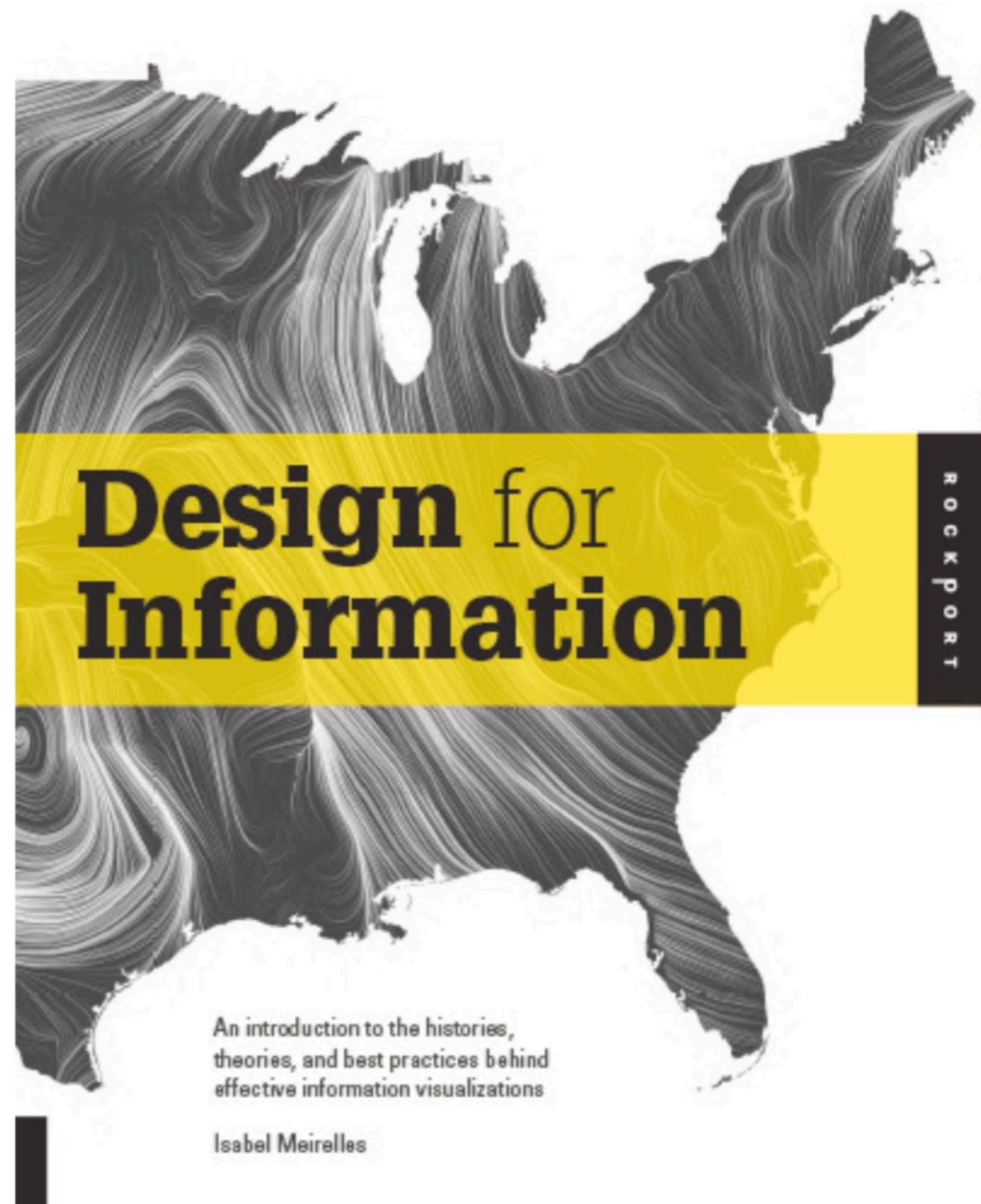
If you can only get one,  
THIS is the one!



Munzner, Tamara. Visualization Analysis and Design. CRC Press, 2014.

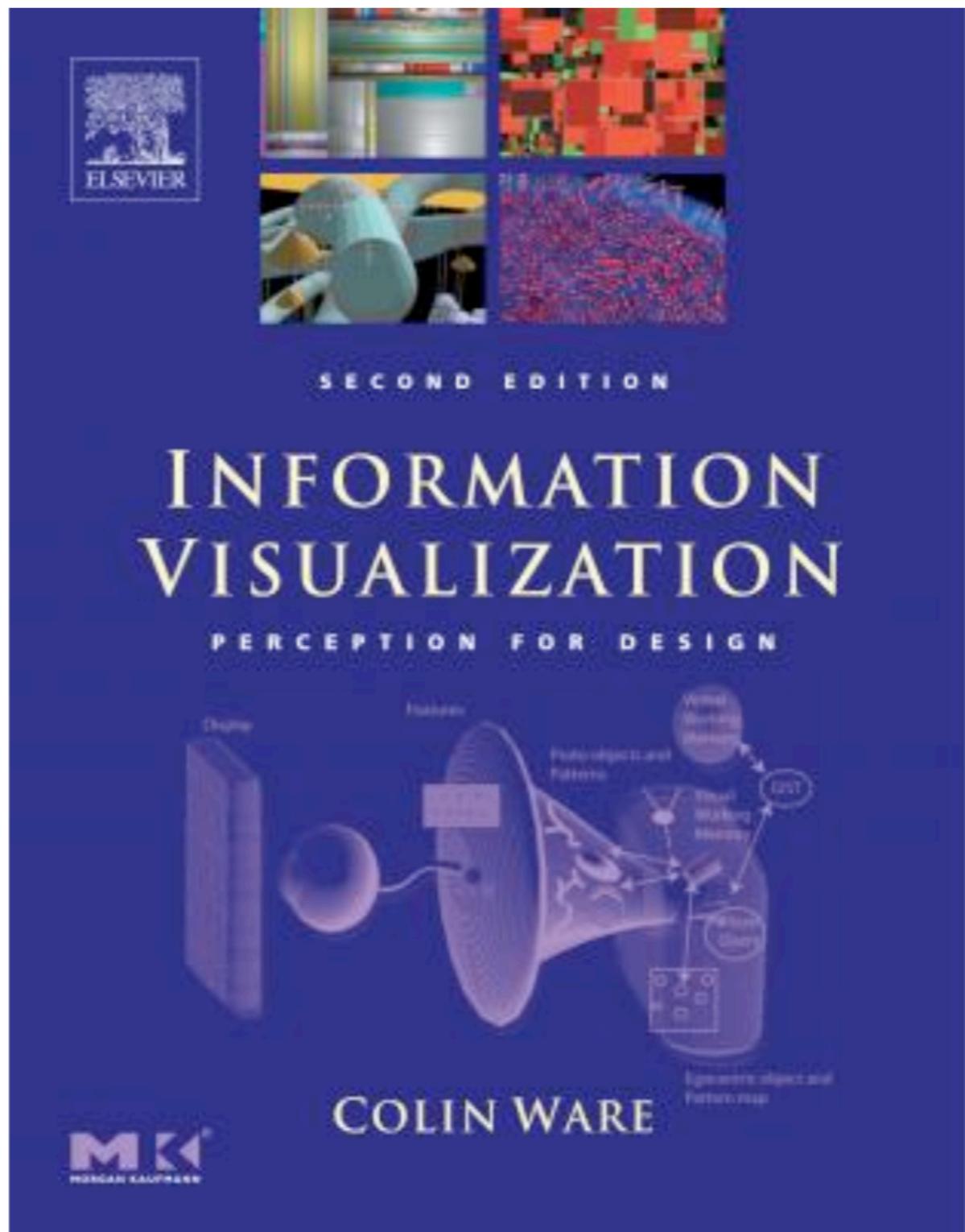
Infovis through examples, history and some free discussion.  
Very visual.

***Beautiful!***



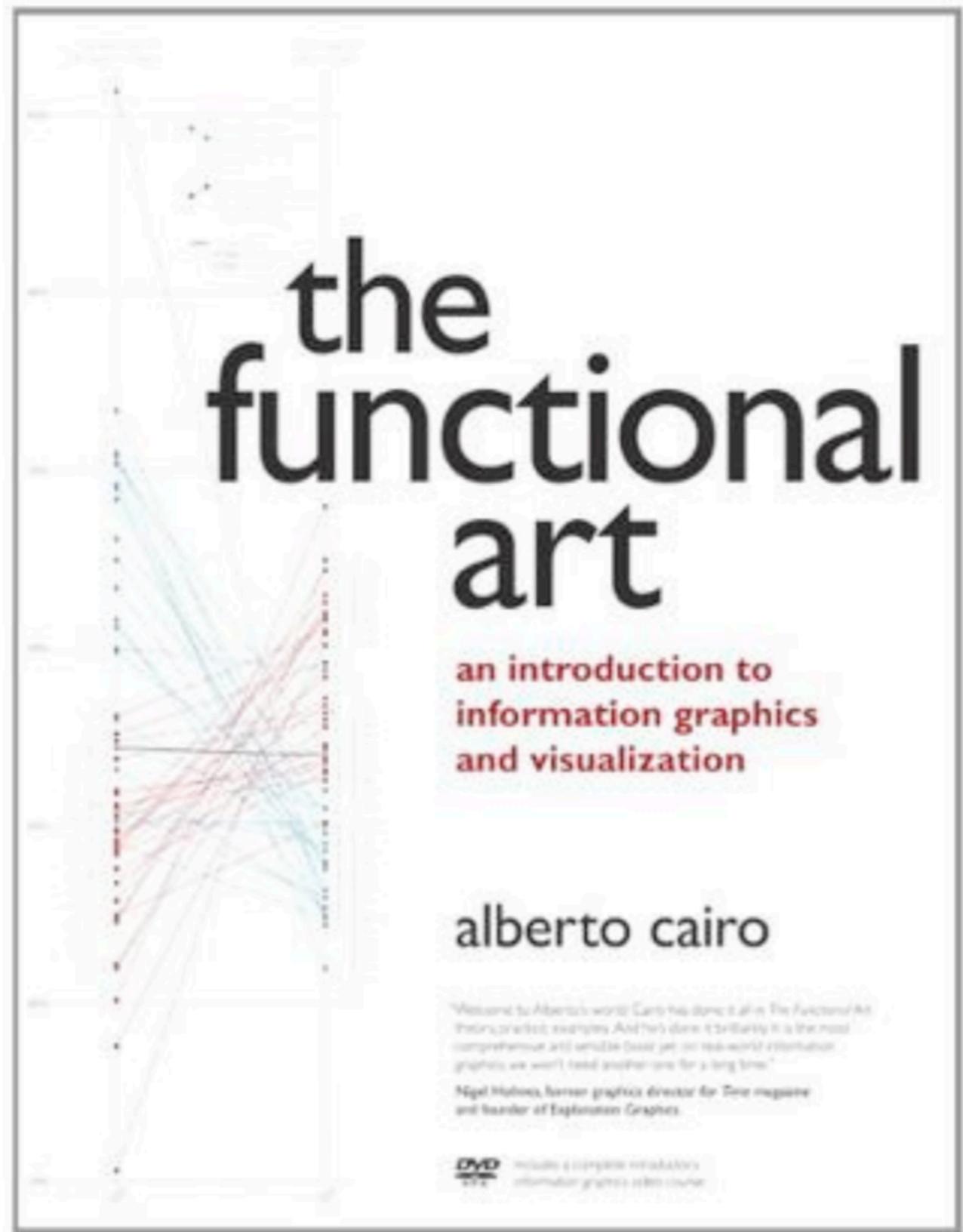
Meirelles, Isabel. Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations. Rockport publishers, 2013.

All about perception,  
and the funny tricks our  
brain plays on us.



Ware, Colin. Information visualization: perception for design. Elsevier, 2013.

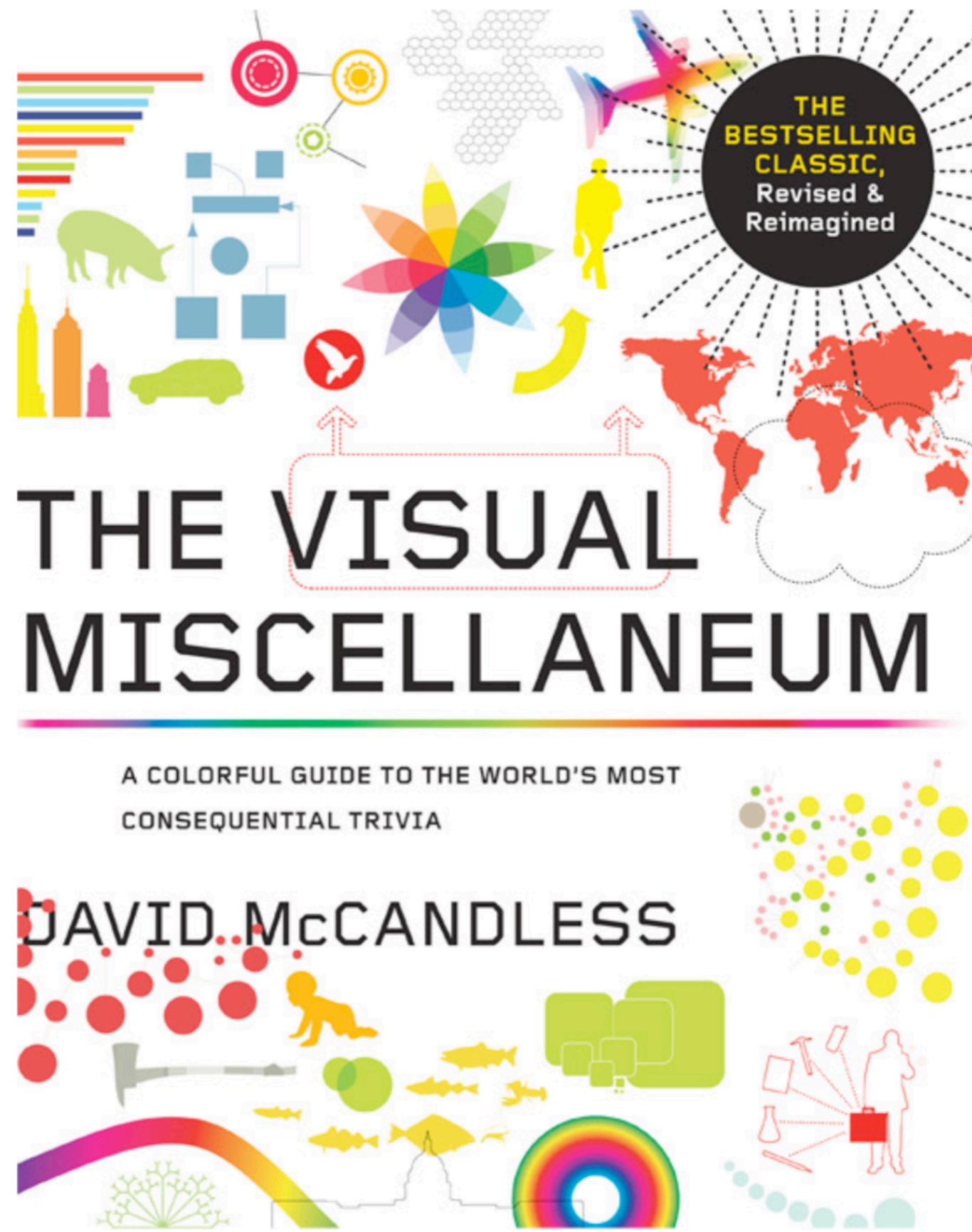
Focus on  
*infographics*



Cairo, Alberto. *The Functional Art: An introduction to information graphics and visualization*. New Riders, 2012.

For inspiration!

Some excellent visualizations! And some so-so...



McCandless, David. *The visual miscellaneum: a colorful guide to the world's most consequential trivia*. Collins Design, 2009.