BEFORE WE START…. A LITTLE EXPERIMENT

37 75

37 75

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

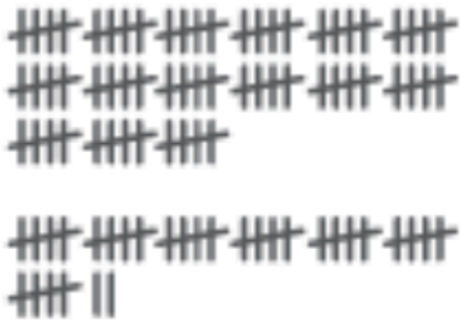
2. Try to come up with as many possible representations/encodings for the “data" above as you can, in the paper segments.

Feel free to be creative!

37 75

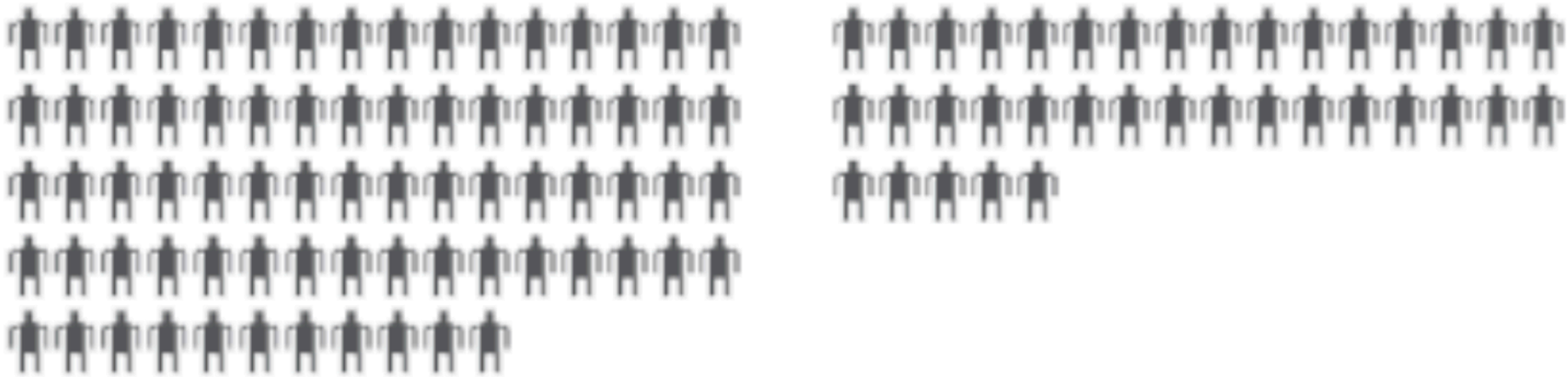
Thirty-seven
Seventy-five

XXXVII LXXV



Squares

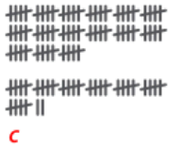
Isotypes



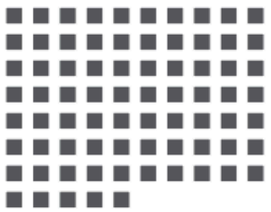
75, 37

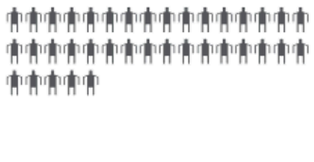
a

b



c

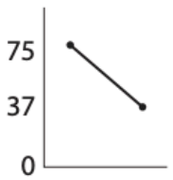




a

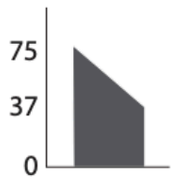
b

c



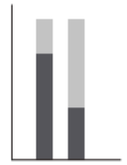
a

b



c

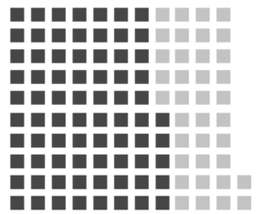
a



b

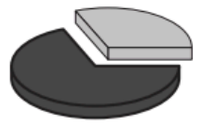
a

b



a

b



c

a

b

c

a

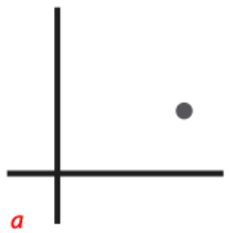
b

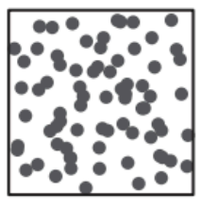
a

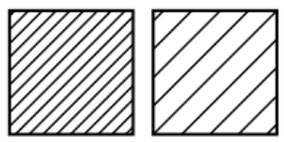
b

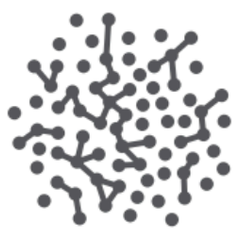
a

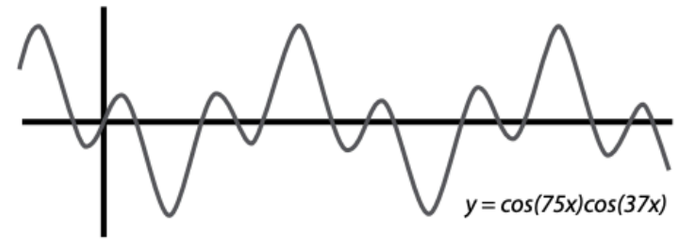
b

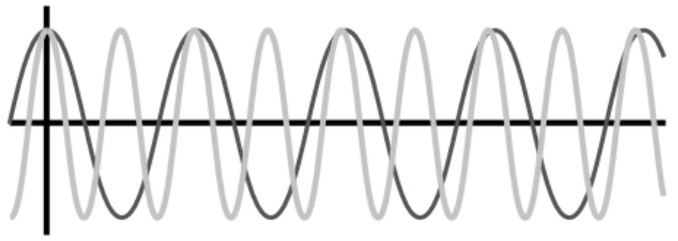


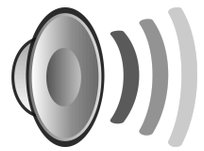






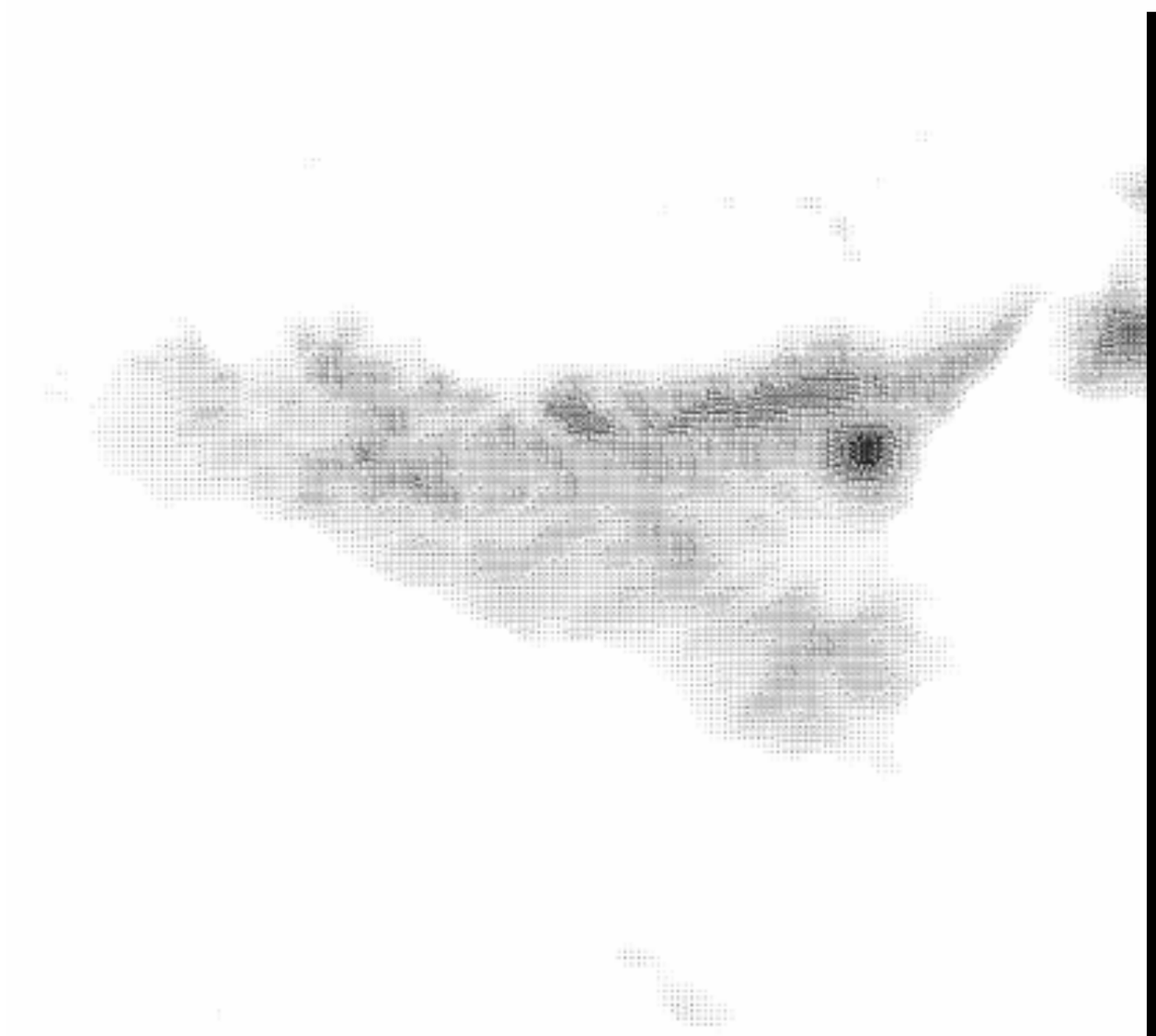


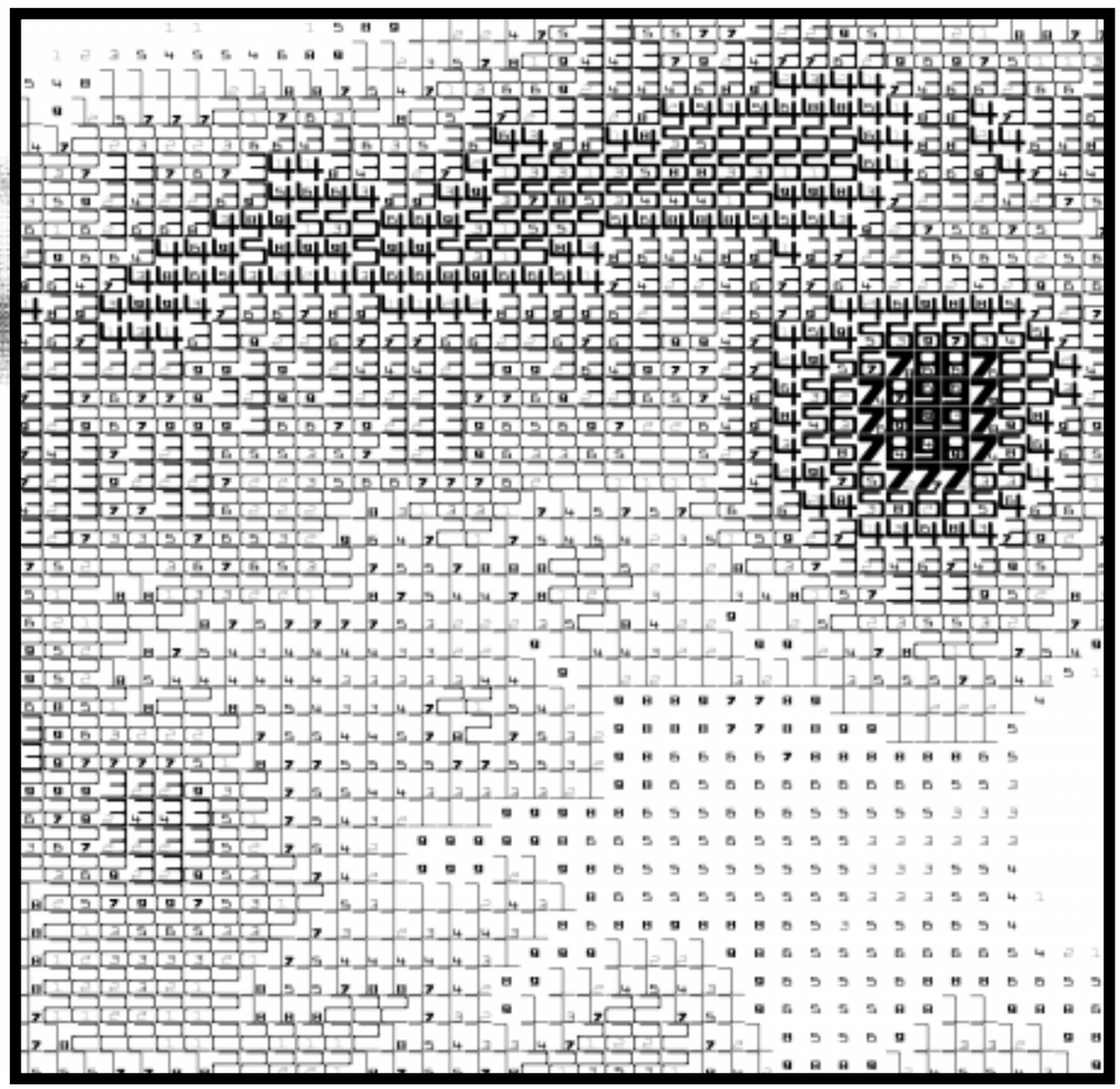




37Hz

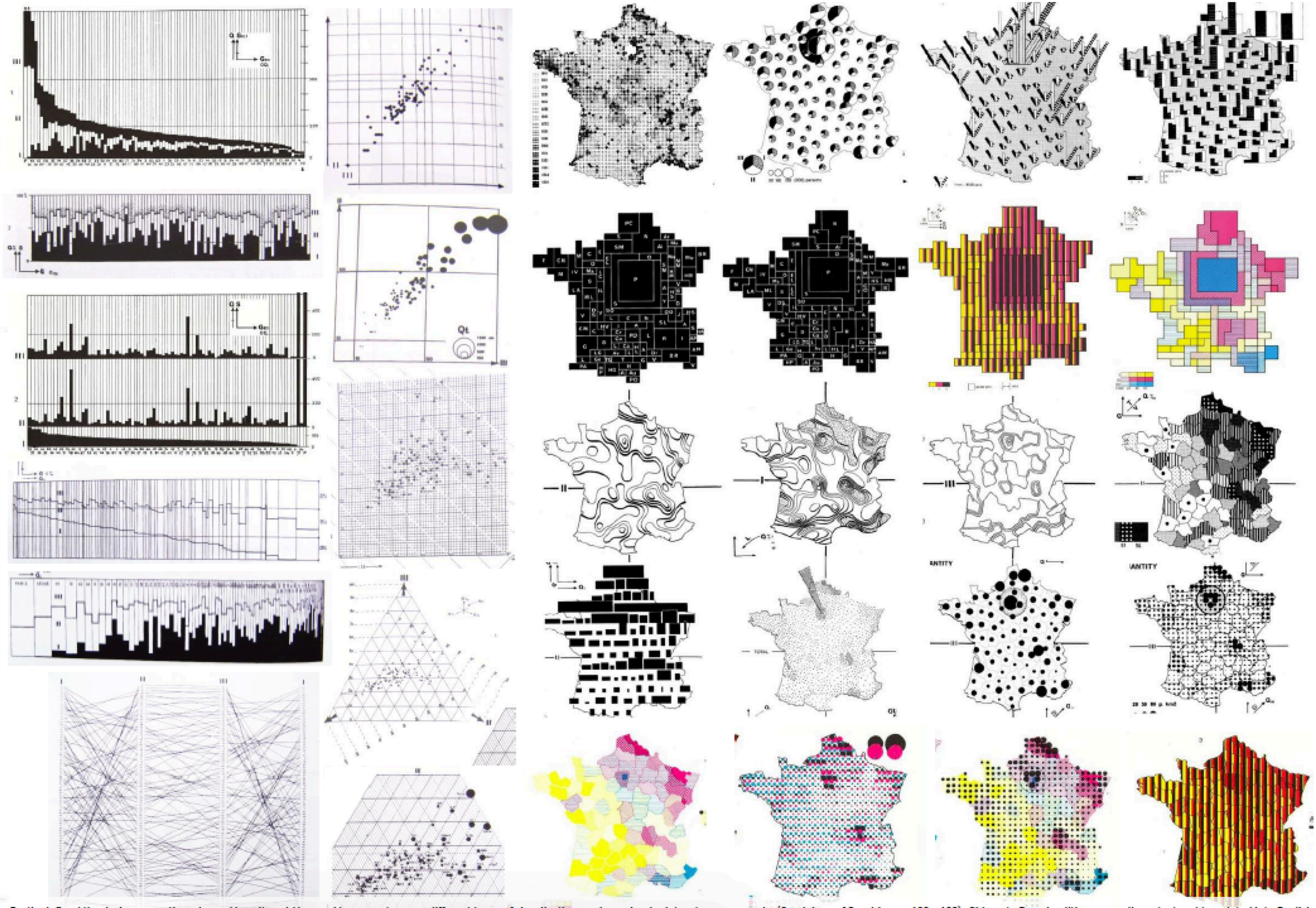
75Hz





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

Départements	Quantités (000)				Proportion %		
	I	II	III	Total	I	II	III
1 AIN	87	83	80	150	45	28	27
2 AISNE	96	74	66	193	29	37	34
3 ALLIER	65	45	57	167	38	27	34
4 Bses ALPES	15	8	13	35	43	24	33
5 Htes ALPES	16	6	13	35	44	21	35
6 ALPES Mmes	31	61	122	214	14	29	57
7 ARDECHE	48	32	25	105	45	31	24
8 ARDENNES	25	53	35	113	22	47	31
9 ARIEGE	33	17	14	64	52	26	22
10 AUBE	28	48	36	112	25	43	32
11 AUDE	50	20	32	102	49	19	32
12 AVEYRON	70	38	29	133	54	24	22
13 BOUCHES DU RH.	42	143	236	421	10	35	55
14 CALVADOS	70	55	69	194	36	28	36
15 CANTAL	45	13	20	78	58	16	26
16 CHARENTE	65	30	38	140	47	25	27
17 CHARENTE Mme	79	39	65	183	43	21	36
18 CHER	43	41	36	120	36	34	30
19 CORREZE	64	28	29	116	55	20	25
21 COTE D'OR	43	44	58	145	30	29	41
22 COTES DU NORD	131	30	67	228	58	15	27
23 CREUSE	58	13	17	88	66	15	19
24 DORDOGNE	104	34	41	179	58	19	23
25 DOUBS	35	67	39	142	25	47	28
26 DROME	46	38	35	119	39	31	30
27 EURE	48	52	45	145	33	36	31
28 EURE & LOIR	44	27	38	119	41	23	34
29 FINISTERE	184	26	89	329	56	23	27
30 GARD	40	31	32	144	28	26	36
31 HAUTE GARONNE	64	67	84	216	30	31	39
32 GERS	63	10	16	89	71	11	18
33 GIRONDE	115	107	170	322	36	27	43
34 HERAULT	62	69	71	173	36	23	41
35 ILLE & V.	137	68	82	279	49	24	29
36 INDRE	54	38	32	116	46	33	28
37 INDRE & L.	61	41	55	157	39	26	35
38 ISERE	68	136	78	282	24	48	28
39 JURA	39	34	27	100	39	34	27
40 LANDES	70	35	20	123	57	28	23
41 LOIR & CHER	51	27	30	108	47	25	28
42 LOIRE	56	160	82	298	19	54	27
43 Hte LOIRE	53	23	22	98	54	24	22
44 LOIRE INF.	101	108	105	314	32	34	34
45 LOIRET	51	51	34	136	38	38	25
46 LOT	41	10	16	67	61	15	24
47 LOT & GAR.	70	24	30	124	57	19	24
48 LOZERE	22	5	7	34	64	15	21
49 MAINE & L.	104	65	85	254	44	28	28
50 MANCHE	116	42	56	214	54	20	26
51 MARNE	44	37	67	148	30	34	46
52 Hte MARNE	25	28	28	81	31	35	34
53 MAYENNE	74	23	28	125	59	19	22
54 MEURTHE & M.	23	127	91	241	9	53	38
55 MEUSE	24	31	27	82	30	38	32
56 MORBIHAN	132	47	59	238	55	20	25
57 MOSELLE	36	173	94	303	12	57	31
58 NIEVRE	34	27	33	94	36	29	35
59 NORD	81	483	296	860	9	56	35
60 OISE	49	69	55	164	24	42	34
61 ORNE	65	30	35	130	50	23	27
62 P.D.C.	94	242	137	473	20	51	29
63 PUY DE DOME	89	79	63	232	38	34	28
64 Bses PYRENEES	89	49	62	191	47	25	32
65 Htes PYRENEES	37	27	28	92	40	29	31
66 PYRENEES ORIENT.	35	20	33	88	40	23	37
67 BAS-RHIN	36	122	114	312	24	39	37
68 Ht-RHIN	40	121	74	235	17	51	32
69 RHONE	44	215	194	453	10	47	43
70 Hte SAONE	34	32	23	89	38	36	26
71 SAONE & L.	84	77	62	233	61	33	26
72 SARTHE	87	45	58	190	46	24	30
73 SAVOIE	44	38	35	117	38	32	30
74 Hte SAVOIE	52	42	45	139	37	30	33
1 PARIS	2	575	940	1517	0	38	62
75 SEINE	8	574	550	1132	1	51	48
76 SEINE INF.	75	152	174	401	19	38	43
77 SEINE & M.	37	72	76	185	20	39	41
78 SEINE & O.	46	320	356	730	6	45	49
79 DEUX-SEVRES	71	29	33	133	53	22	25
80 SOMME	57	60	61	186	31	38	33
81 TARN	55	47	33	135	41	35	24
82 TARN & G.	44	13	18	75	59	17	24
83 VAR	53	50	81	184	29	31	44
84 VAUCLUSE	40	30	41	111	36	27	37
85 VENDEE	110	38	40	188	59	20	21
86 VIENNE	60	29	39	128	47	23	30
87 Hte VIENNE	64	47	45	156	41	30	29
88 VOSGES	36	85	43	174	21	54	25
89 YONNE	41	28	37	106	39	26	35
90 BELFORT	3	25	13	41	8	60	32
	5212	6705	6905	18825	28	35	37



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.

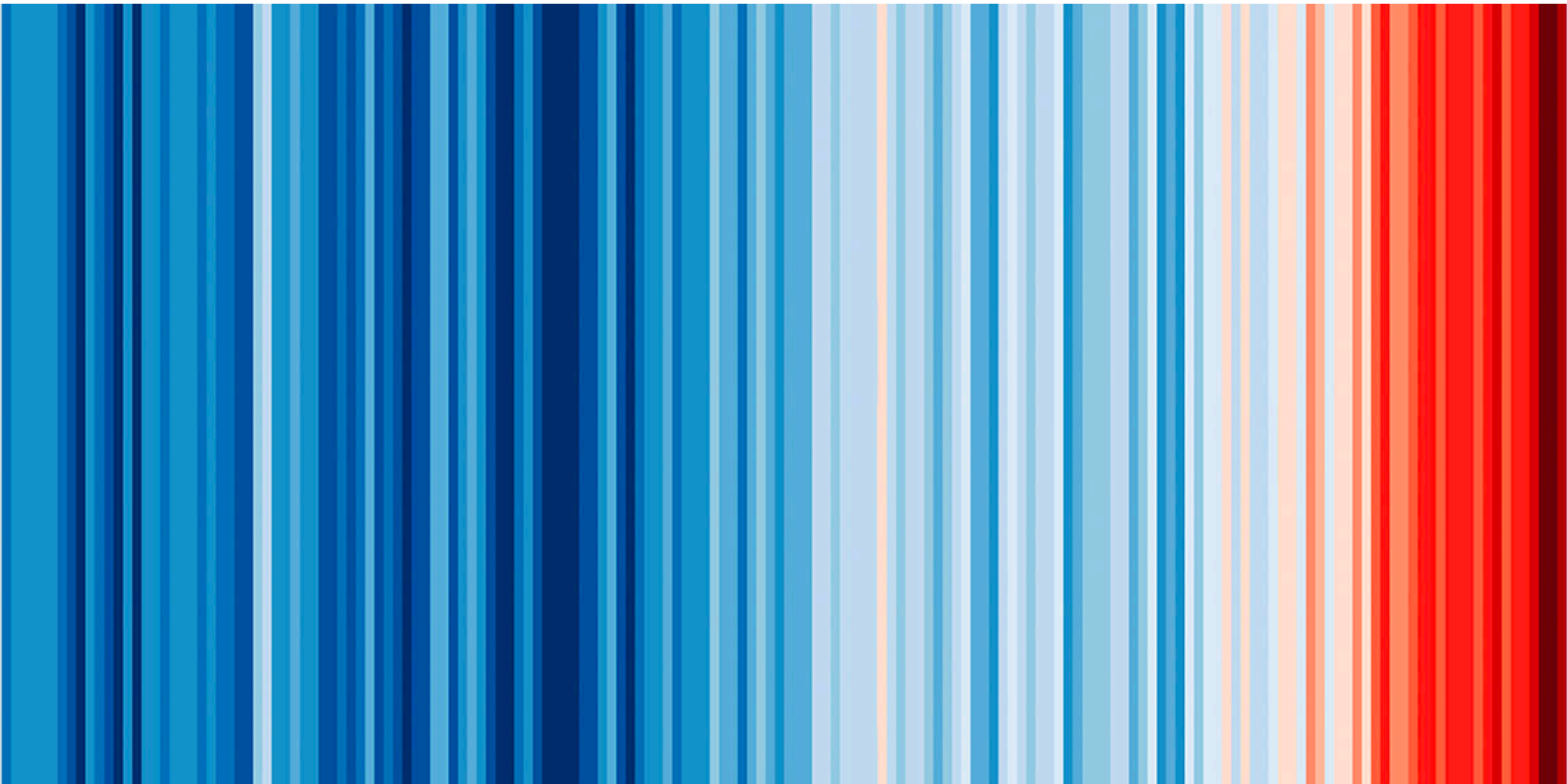
… ANOTHER LITTLE EXPERIMENT

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

JSC270 - Class #6 Prof. Chevalier

Communication

Visualization theory and principles, storytelling with data

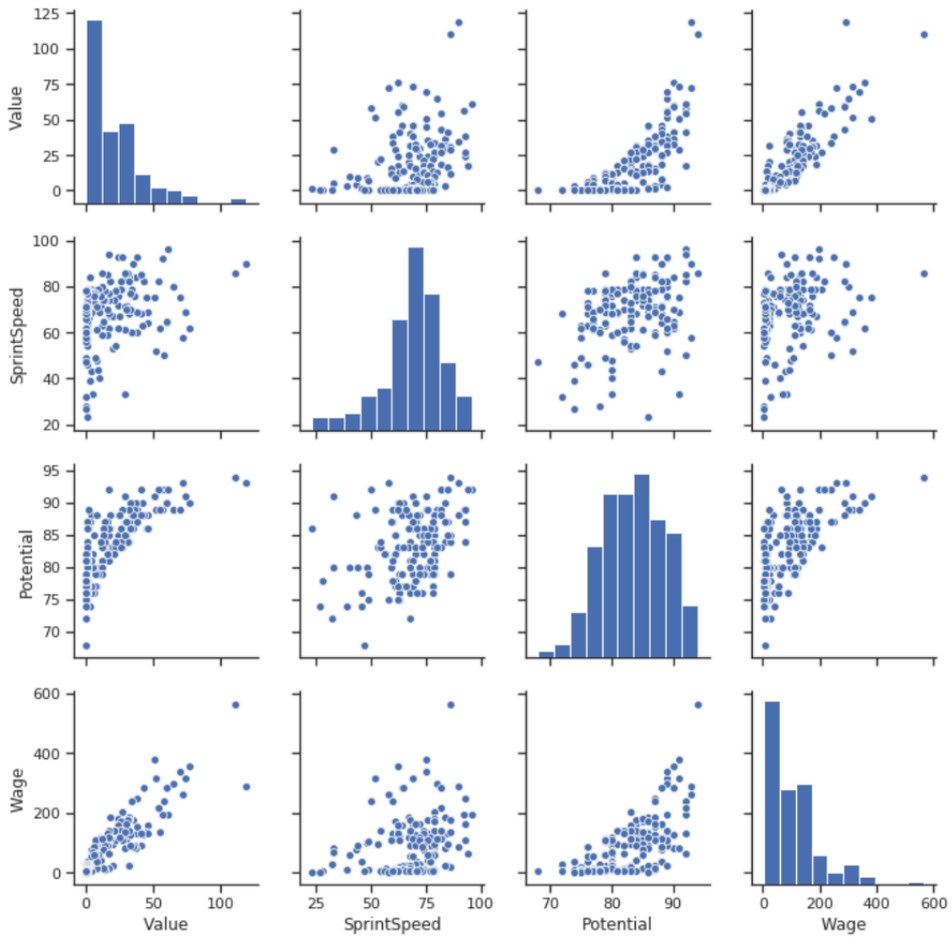


MAIN USES OF DATA VISUALIZATION

EXPLORATORY

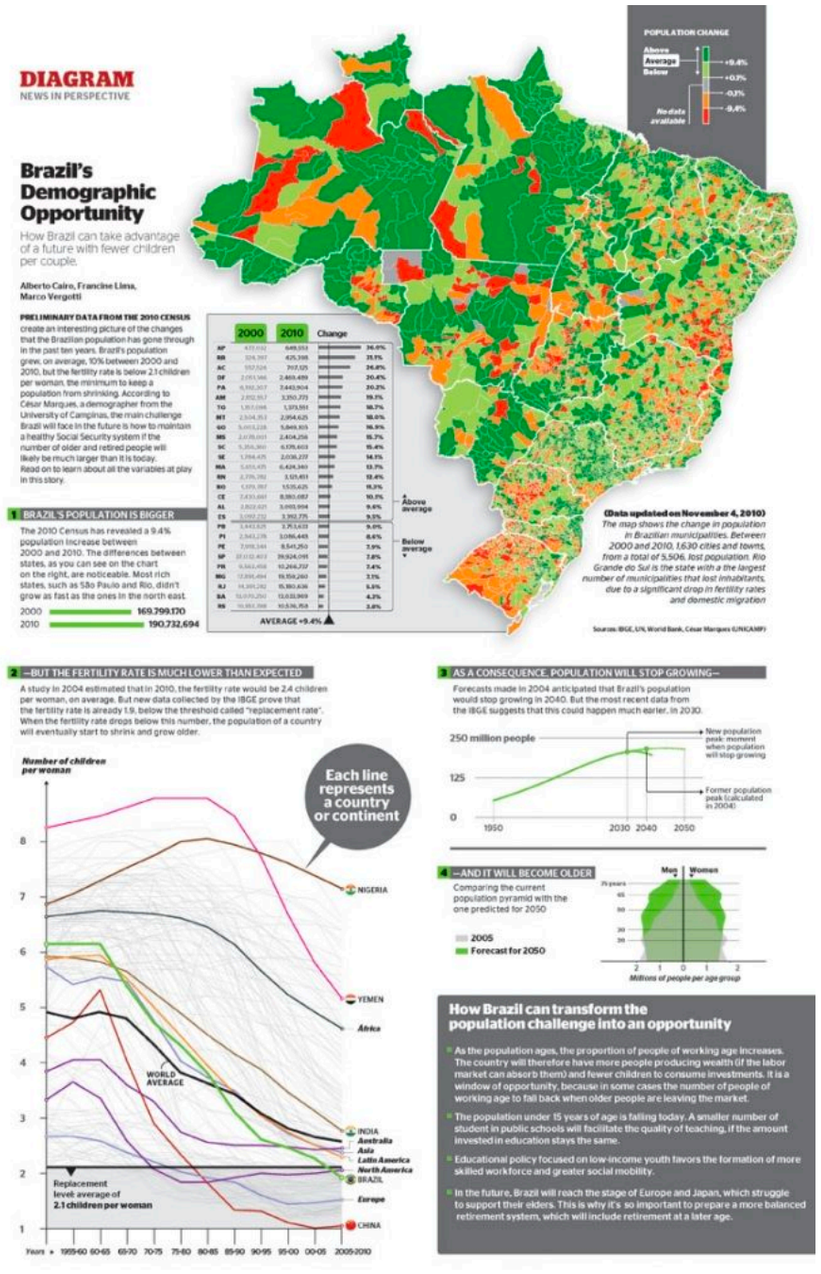
Acquiring insights

```
filtered_player_df = player_df[(player_df['Club'].isin(['FC Barcelona', 'Paris Saint-
Germain',
       'Manchester United', 'Manchester City', 'Chelsea', 'Real Madrid','FC Porto','FC
Bayern München'])) &
                     (player_df['Nationality'].isin(['England', 'Brazil', 'Argentina',
       'Brazil', 'Italy','Spain','Germany']))
                    ]
# Single line to create pairplot
g = sns.pairplot(filtered_player_df[['Value','SprintSpeed','Potential','Wage']])
```



EXPLANATORY

Storytelling



Encoding: designer

Decoding: user

1 Data

What is my data?
Which data type?
Ordinal / numerical / categorical?

2 Visual Mapping

What is my visual representation?
Which visual variables am I using?
How am I encoding my data?

3 Rendering

What is my medium?
monoscopic/stereoscopic?
Tangiblity?
Print / digical?

4 Perceiving

What does it show?
Where is big, medium, small?
How do things compare?
What relationships exist?

5 Interpreting

What does it mean?
What does color mean?
Wha does 'up 'mean?
What do these patterns show?

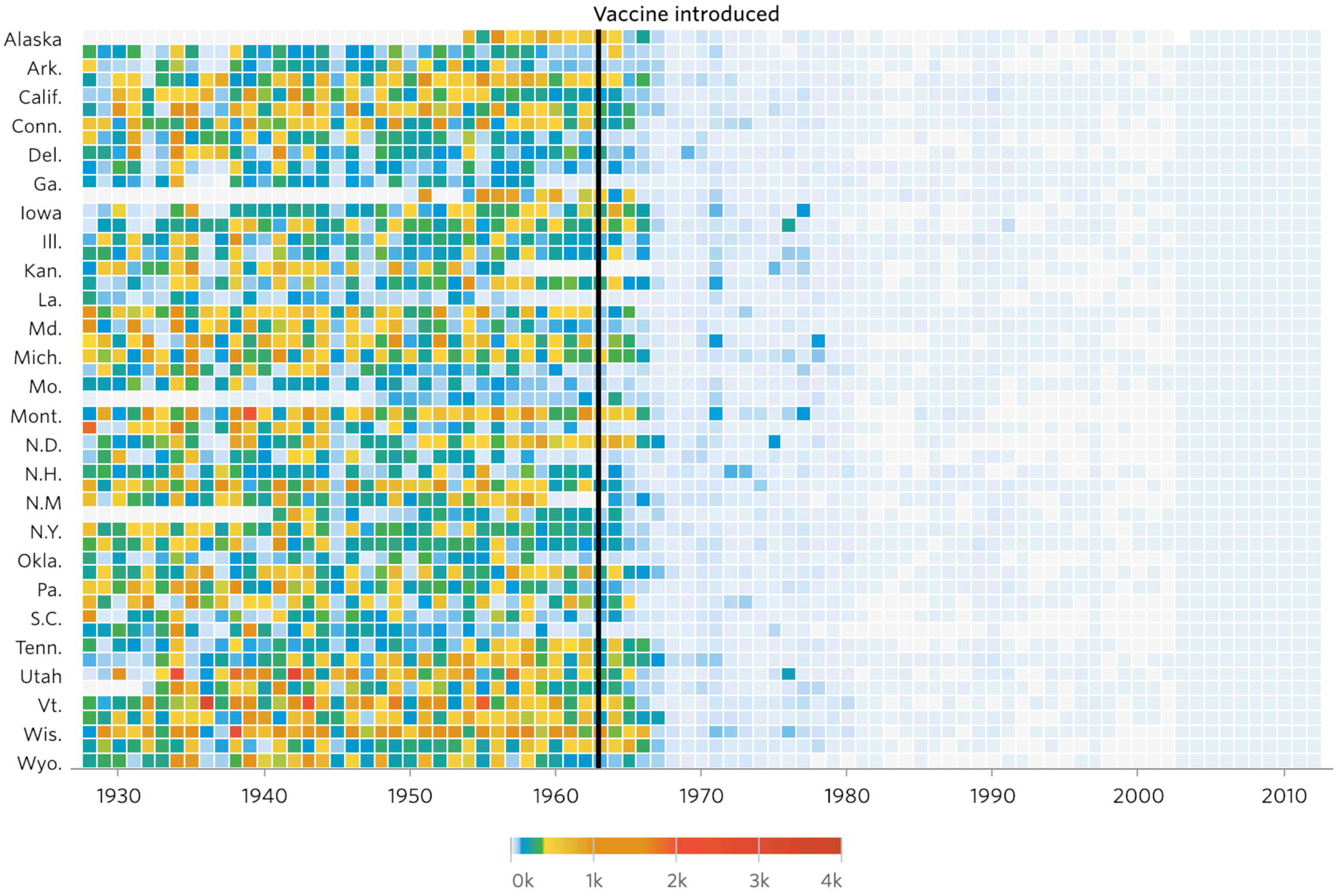
6 Comprehending

What does it mean for me?
What shall I do now?
Is this all true?
What do I learn?

View

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



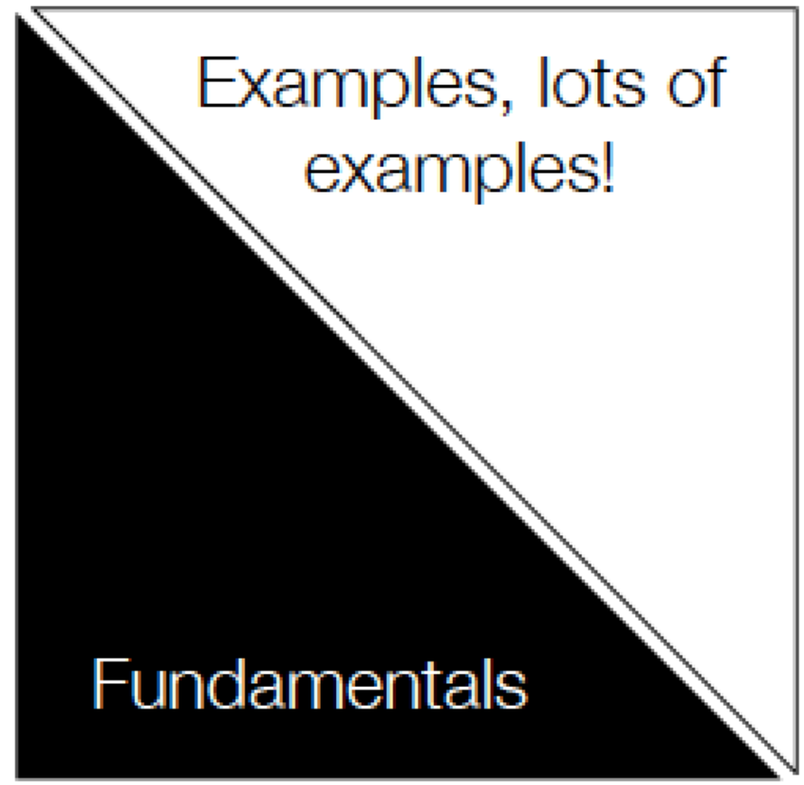
http://graphics.wsj.com/infectious-diseases-and-vaccines/

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!

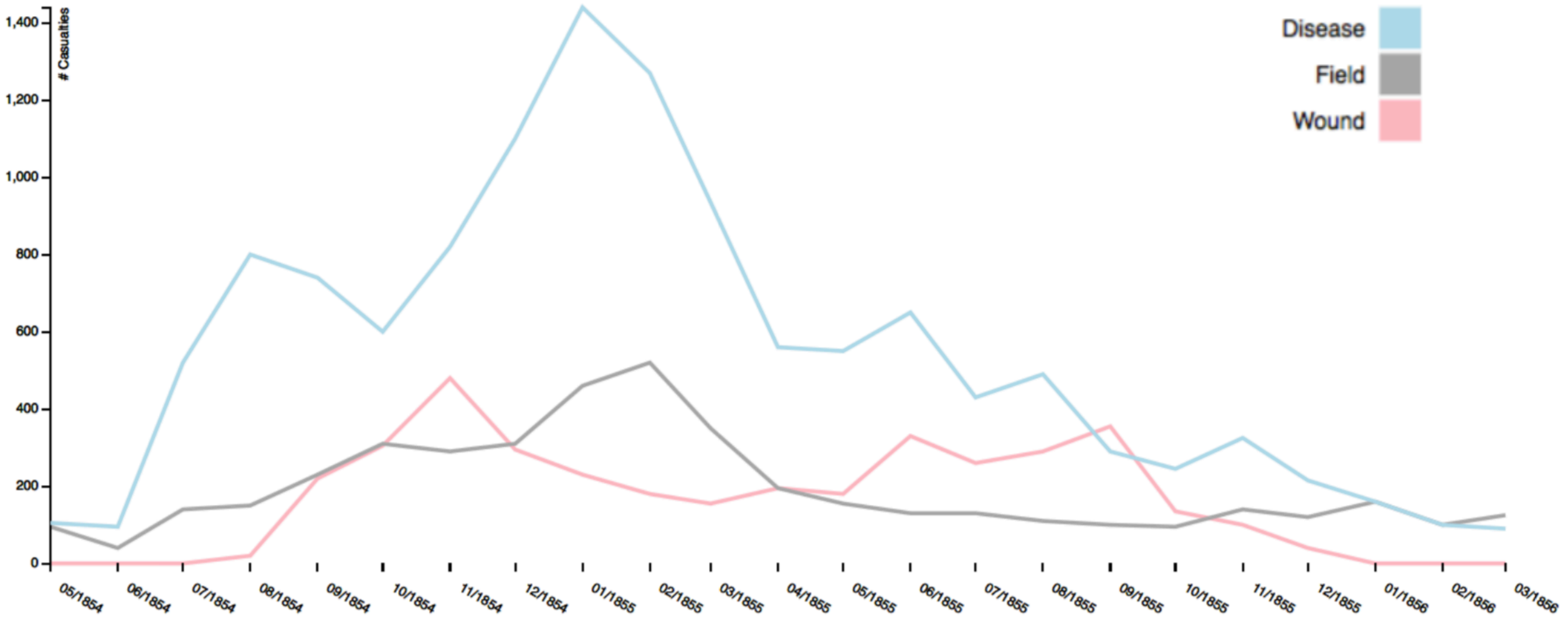


British Casualties in the Crimean War

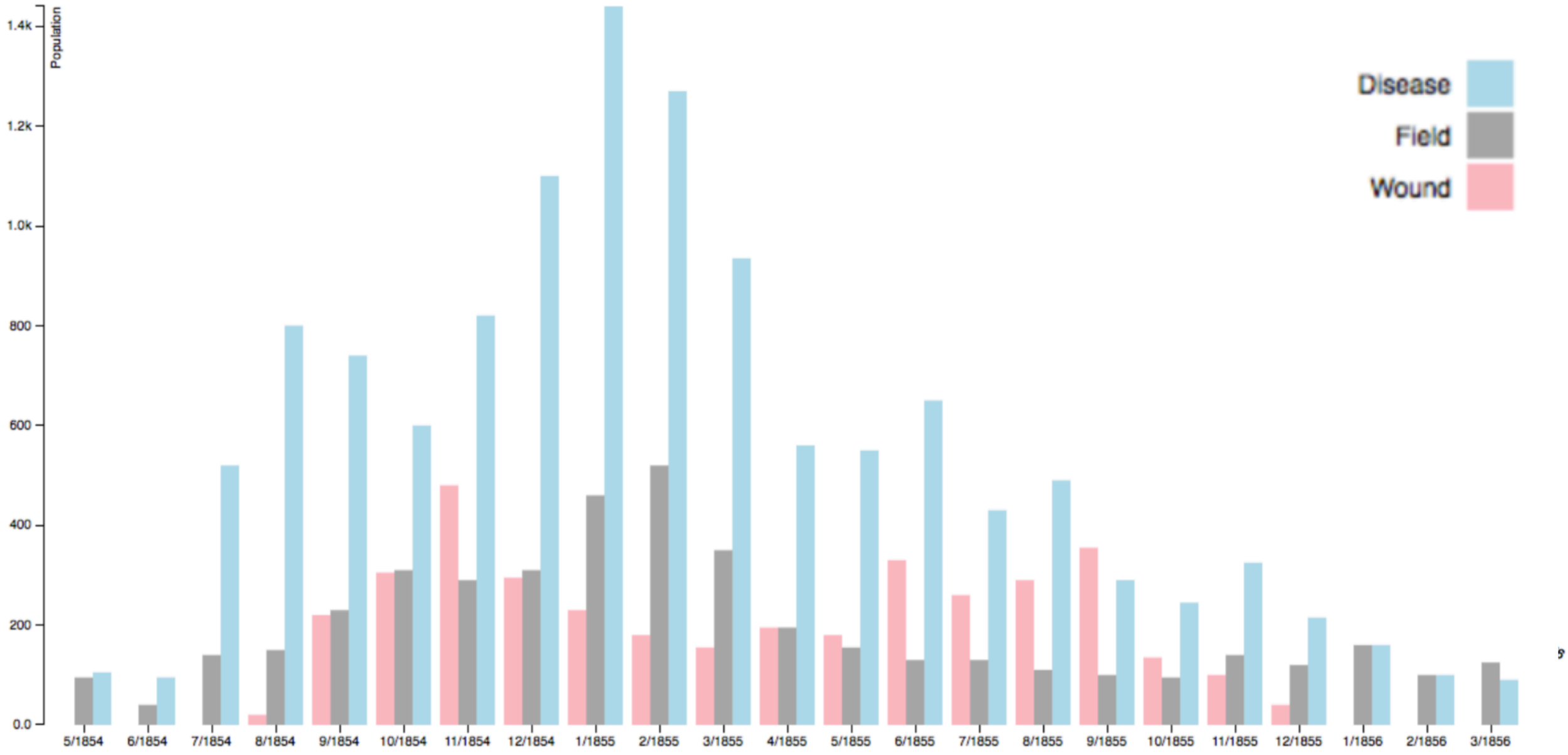
Data	Wound	Field	Disease
05/1854	0	95	105
06/1854	0	40	95
07/1854	0	140	520
08/1854	20	150	800
09/1854	220	230	740
10/1854	305	310	600
11/1854	480	290	820
12/1854	295	310	1100
01/1855	230	460	1440
02/1855	180	520	1270
03/1855	155	350	935
04/1855	195	195	560
05/1855	180	155	550
06/1855	330	130	650
07/1855	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%?

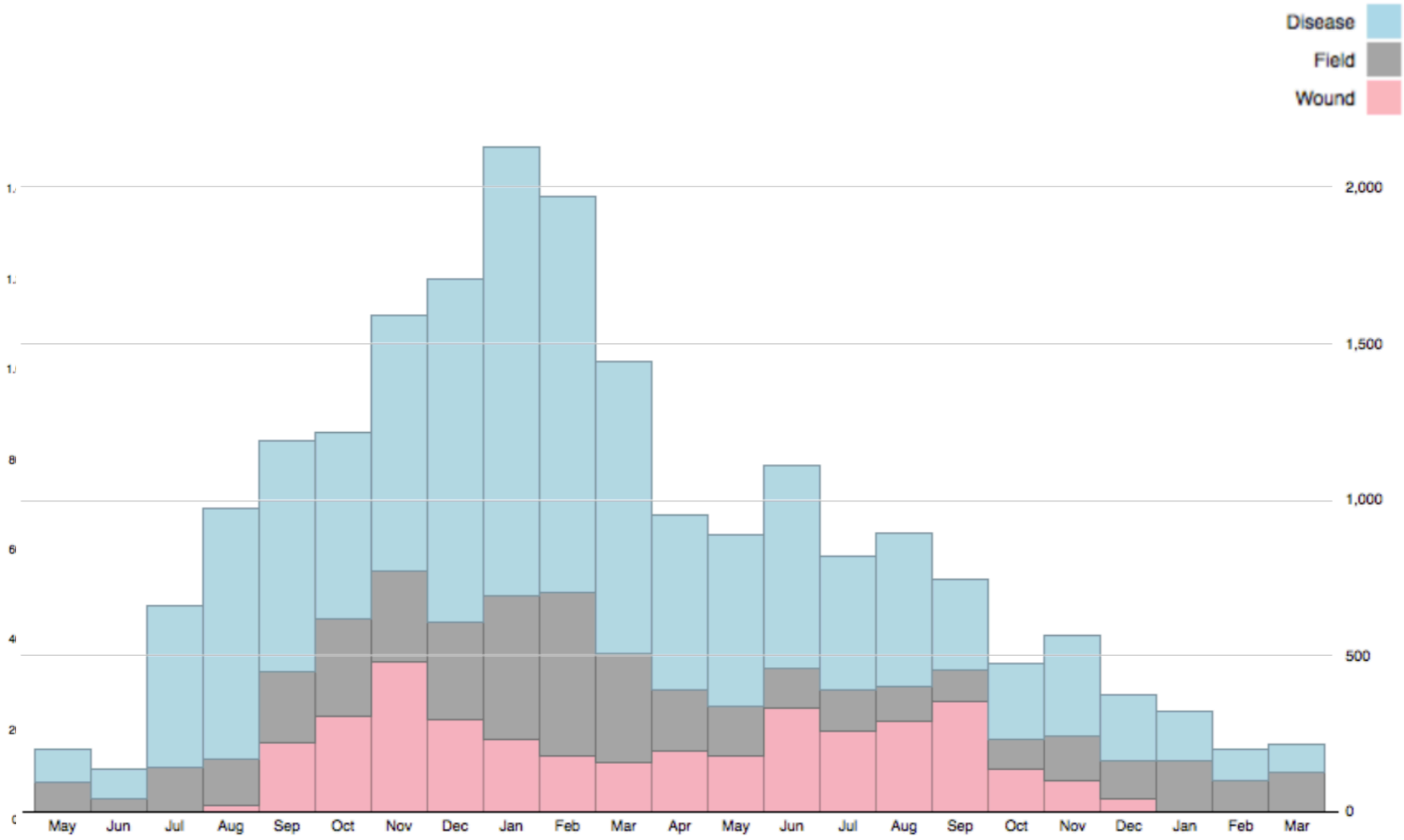
Month with highest casualty rates in the field?



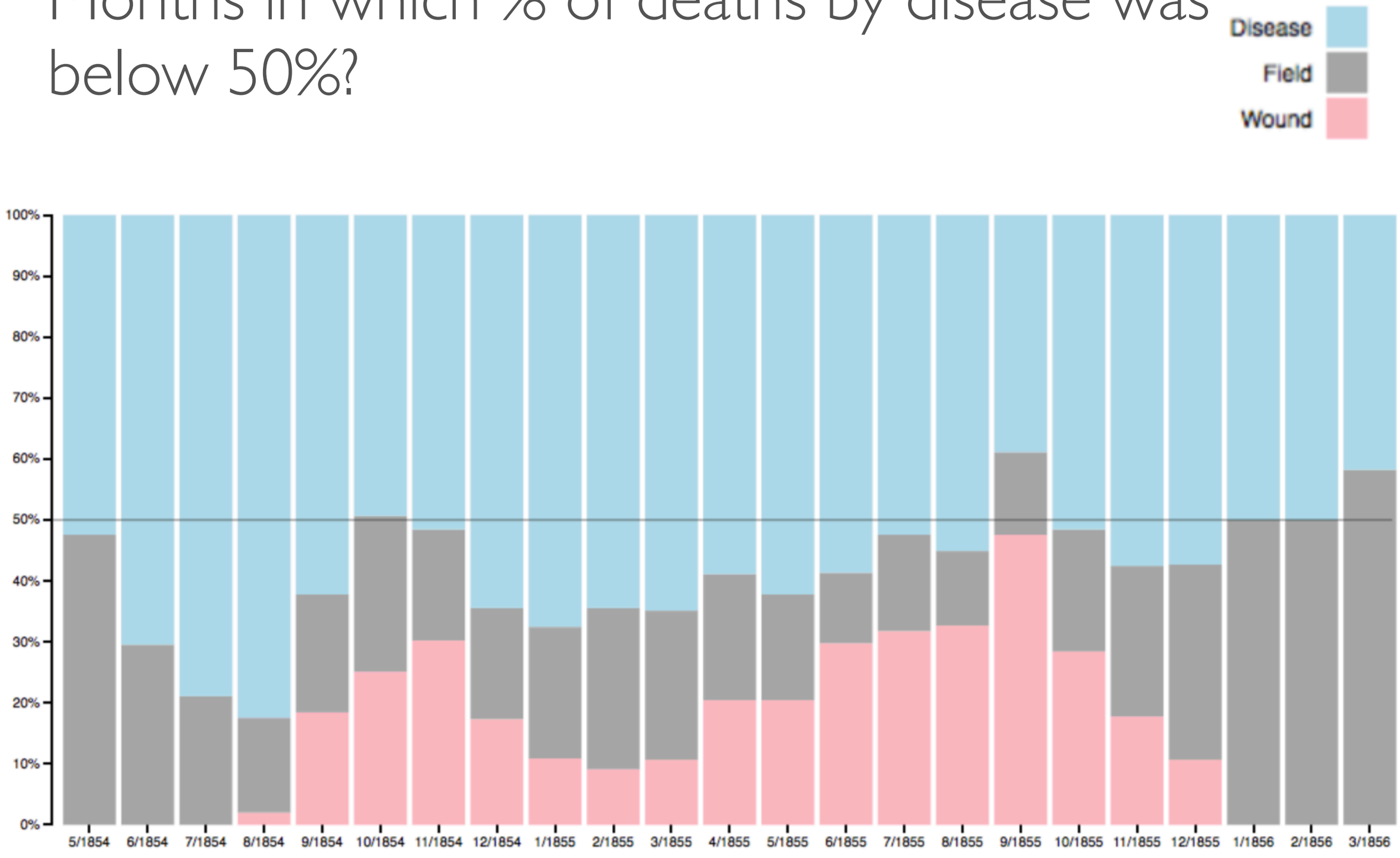
Months in which deaths by wound exceeds deaths in the field?



Month with highest total casualty rate?



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



WHY VISUALIZATION?

Temperature Anomalies by Country
Years 1880 - 2017

1880

-2.0 °C -1.0 °C +0.0 °C +1.0 °C +2.0 °C

Afghanistan Albania Algeria Andorra Angola Antarctica Argentina Armenia Australia Austria Azerbaijan Bahamas, The Bahrain Bangladesh Barbados Belarus

Belize Benin Bhutan Bolivia Bosnia and H. Botswana Brazil Brunei Bulgaria Burkina Faso Burundi Cabo Verde Cambodia Cameroon Canada Central African Rep.

Chad Chile China Colombia Comoros Congo, DR Congo, R Costa Rica Croatia Cuba Cyprus Czechia Côte d'Ivoire Denmark Djibouti Dominica

Dominican Republic Ecuador Egypt, Arab Rep. El Salvador Equatorial Guinea Eritrea Estonia eSwatini Ethiopia Fiji Finland France Gabon Gambia, The Georgia Germany

Ghana Greece Grenada Guatemala Guinea Guinea-Bissau Guyana Haiti Honduras Hungary Iceland India Indonesia Iran, Islamic Rep. Iraq Ireland

Israel Italy Jamaica Japan Jordan Kazakhstan Kenya Kiribati Korea, DPR Korea Kosovo Kuwait Kyrgyz Republic Lao PDR Latvia Lebanon

Lesotho Liberia Libya Liechtenstein Lithuania Luxembourg Macedonia, FYR Madagascar Malawi Malaysia Maldives Mali Malta Marshall Islands Mauritania Mauritius

Mexico Micronesia Moldova Monaco Mongolia Montenegro Morocco Mozambique Myanmar Namibia Nauru Nepal Netherlands New Zealand Nicaragua Niger

Nigeria Norway Oman Pakistan Palau Panama Papua New Guinea Paraguay Peru Philippines Poland Portugal Qatar Romania Russia Rwanda

Samoa San Marino Sao Tome and P. Saudi Arabia Senegal Seychelles Sierra Leone Singapore Slovak Republic Slovenia Solomon Islands Somalia South Africa South Sudan Spain Sri Lanka

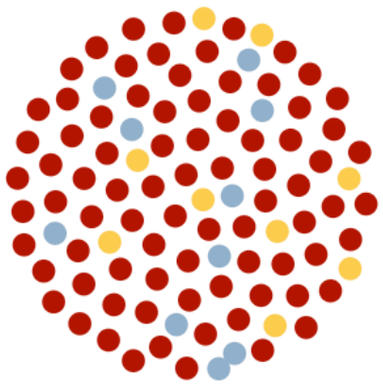
St. Kitts and Nevis St. Lucia St. V. and the G. Sudan Suriname Sweden Switzerland Syria Taiwan Tajikistan Thailand Timor-Leste Togo Tonga Trinidad and Tobago Tunisia

Turkey Turkmenistan Tuvalu Uganda Ukraine United Arab Emirates Uruguay USA Uzbekistan Vanuatu Venezuela, RB Vietnam Yemen, Rep. Zambia Zimbabwe

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.

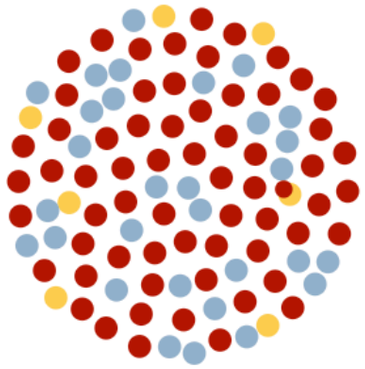
Watch how the measles outbreak spreads when kids get vaccinated – and when they don't

vaccinated · susceptible · vaccinated but susceptible · infected · contact with an infected person



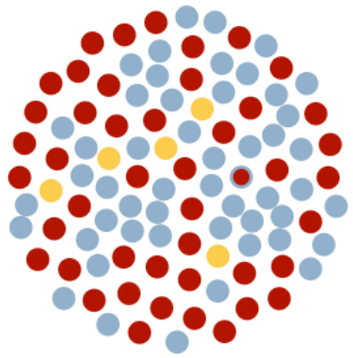
NOT PROTECTED

10.0% vax rate



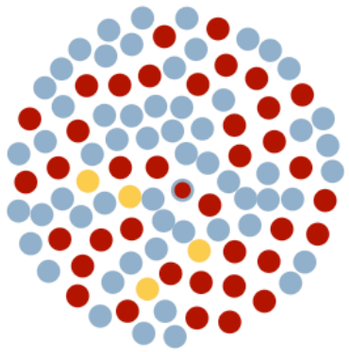
NOT PROTECTED

30.0% vax rate



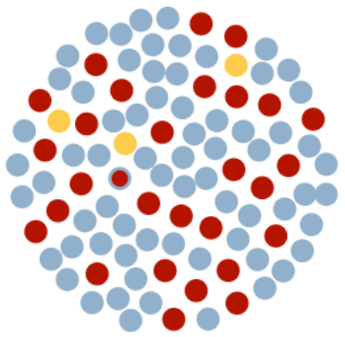
NOT PROTECTED

50.0% vax rate



NOT PROTECTED

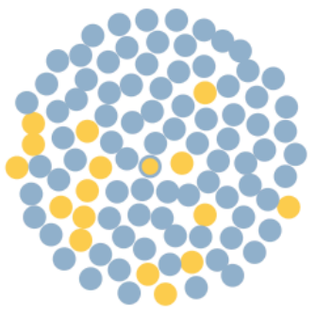
58.5% vax rate, similar to **Okanagan County, WA**



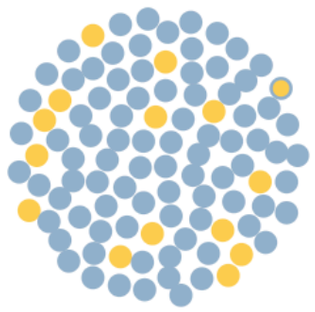
NOT PROTECTED

68.9% vax rate, similar to **Thurston County, WA**

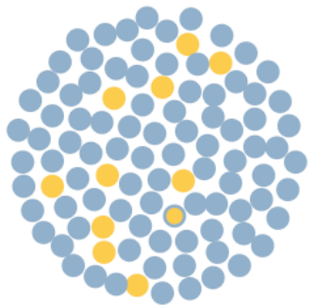
74.4% vax rate, similar to **Island County, WA**



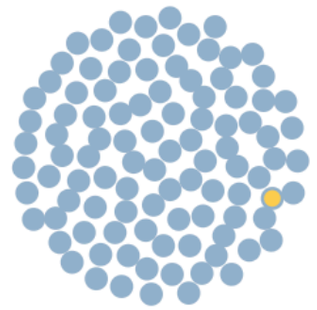
83.8% vax rate, similar to **Santa Cruz County, CA**



86.0% vax rate, similar to **Los Angeles County, CA**



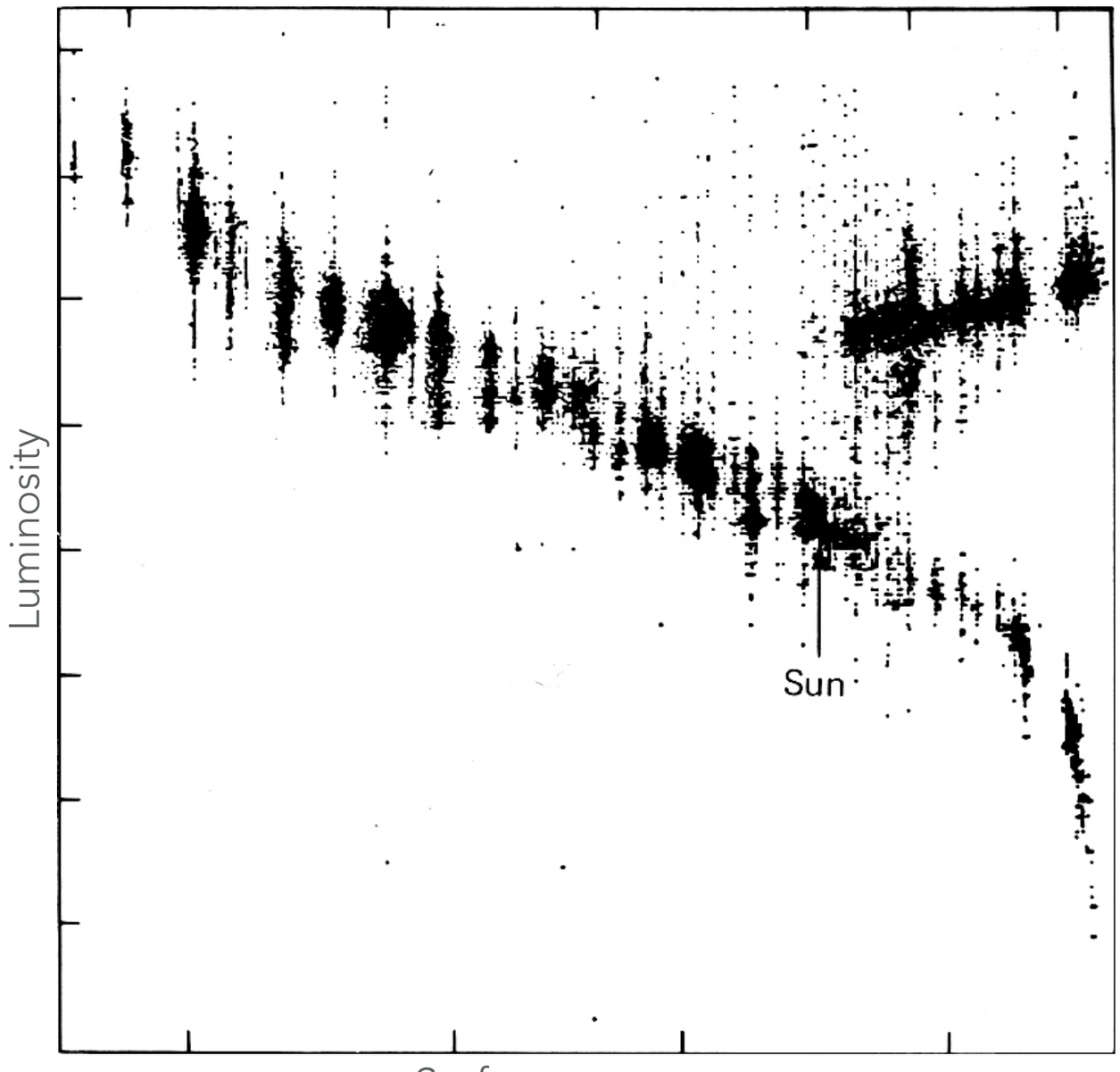
90.0% vax rate, similar to **Orange County, CA**

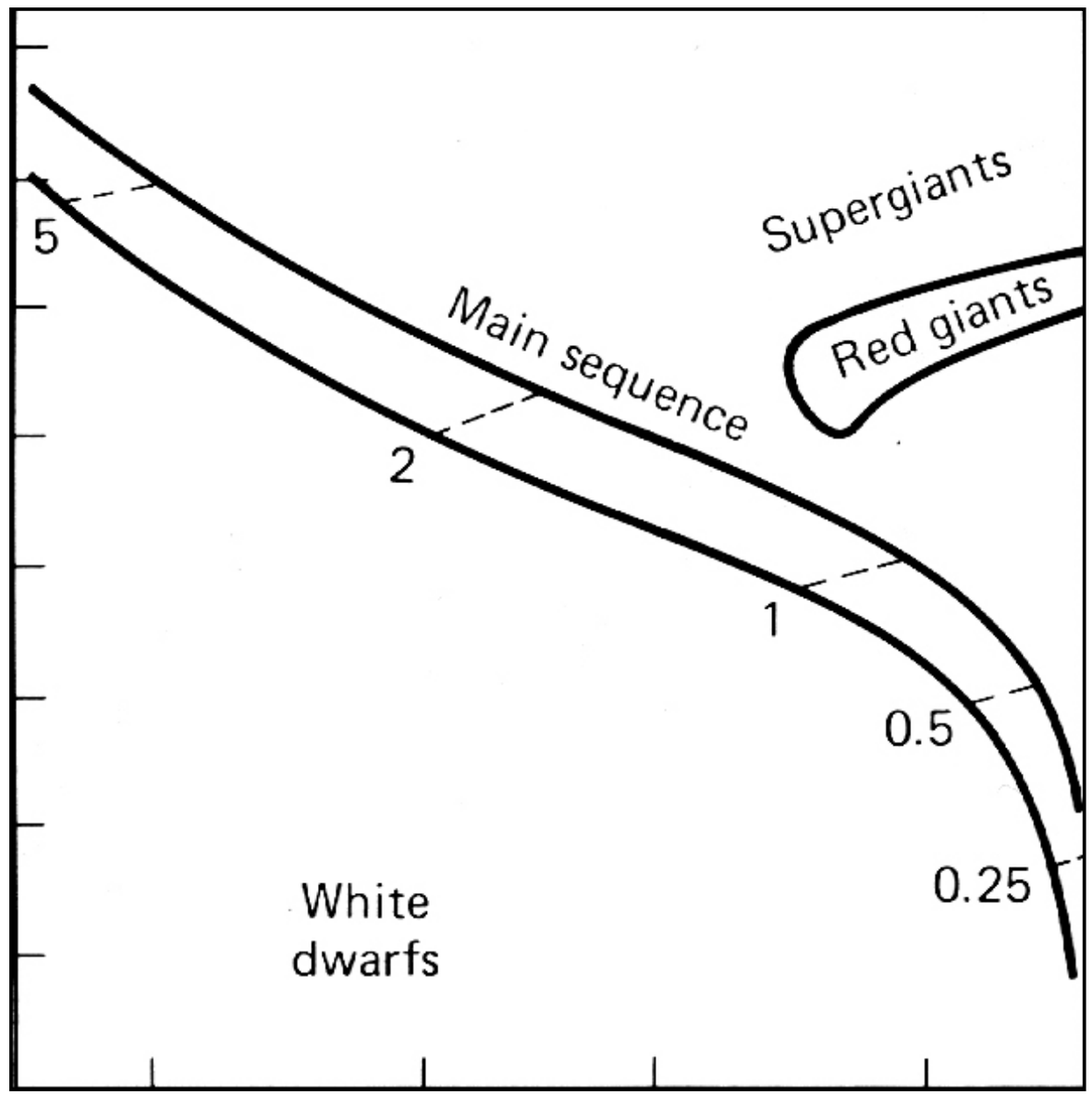


99.7% vax rate, similar to **Gadsden County, FL**

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

PERCEPTION AND COGNITION

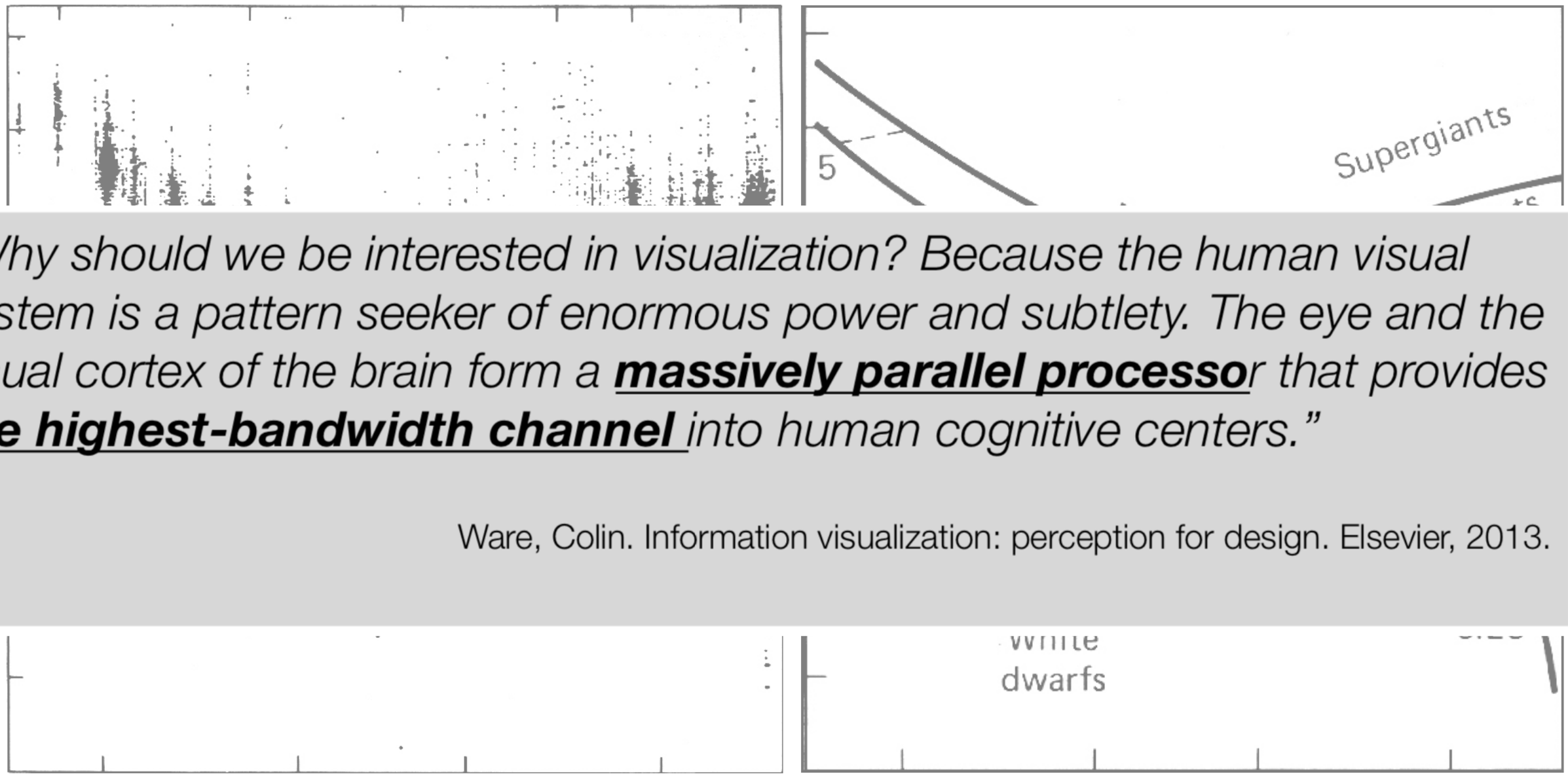




Hertzsprung Russell Diagram and its interpretation

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

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 ***<u>massively parallel processo</u>****r that provides* ***<u>the highest-bandwidth channel</u>*** *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]

1424164963575984759217659684748917284822859588198294509685048506958476121240440746748989851714959691245676596080208606083654164964575906439804792485769607812859607999187128452681014959691245677818742416496457576596081495969124567012859607991649645751278799187128452984969122359164964575958819825096357659608059600

14241649635759847592176596847489172848 2
28595881982945096850485069584761212404 4
07467489898517149596912456765960802086 0
60836541649645759064398047924857696078 1
28596079991871284526810149596912456778 1
87424164964575765960814959691245670128 5
96079916496457512787991871284529849691 2
22359164964575958819825096357659608059 6

This is your brain on visualization….

PERCEPTION FUNDAMENTALS

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

Data  **Graphics**

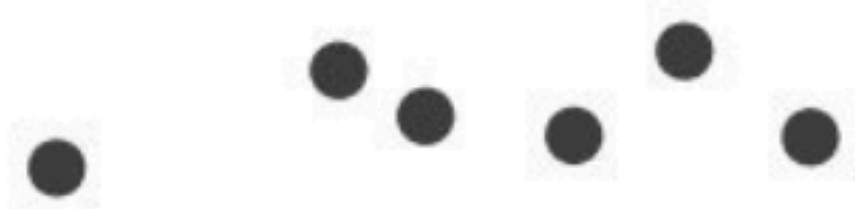
Which representations can I use to encode data?

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

MARKS

Geometric primitives

Points



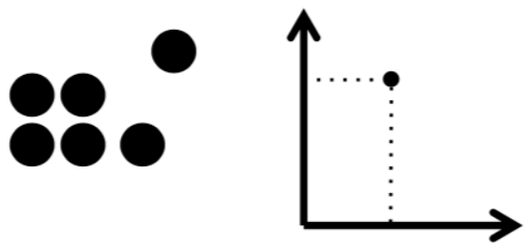
Lines

Areas

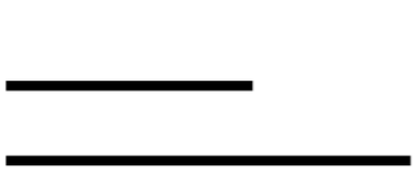
CHANNELS (VISUAL VARIABLES)

Control appearance of marks

Position



Size- Length

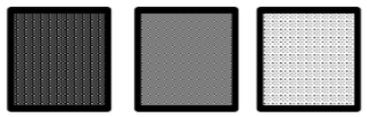


Size- Area

Size- Volume

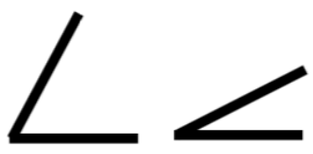
Format

Texture



Orientation/Direction

Angle



Color

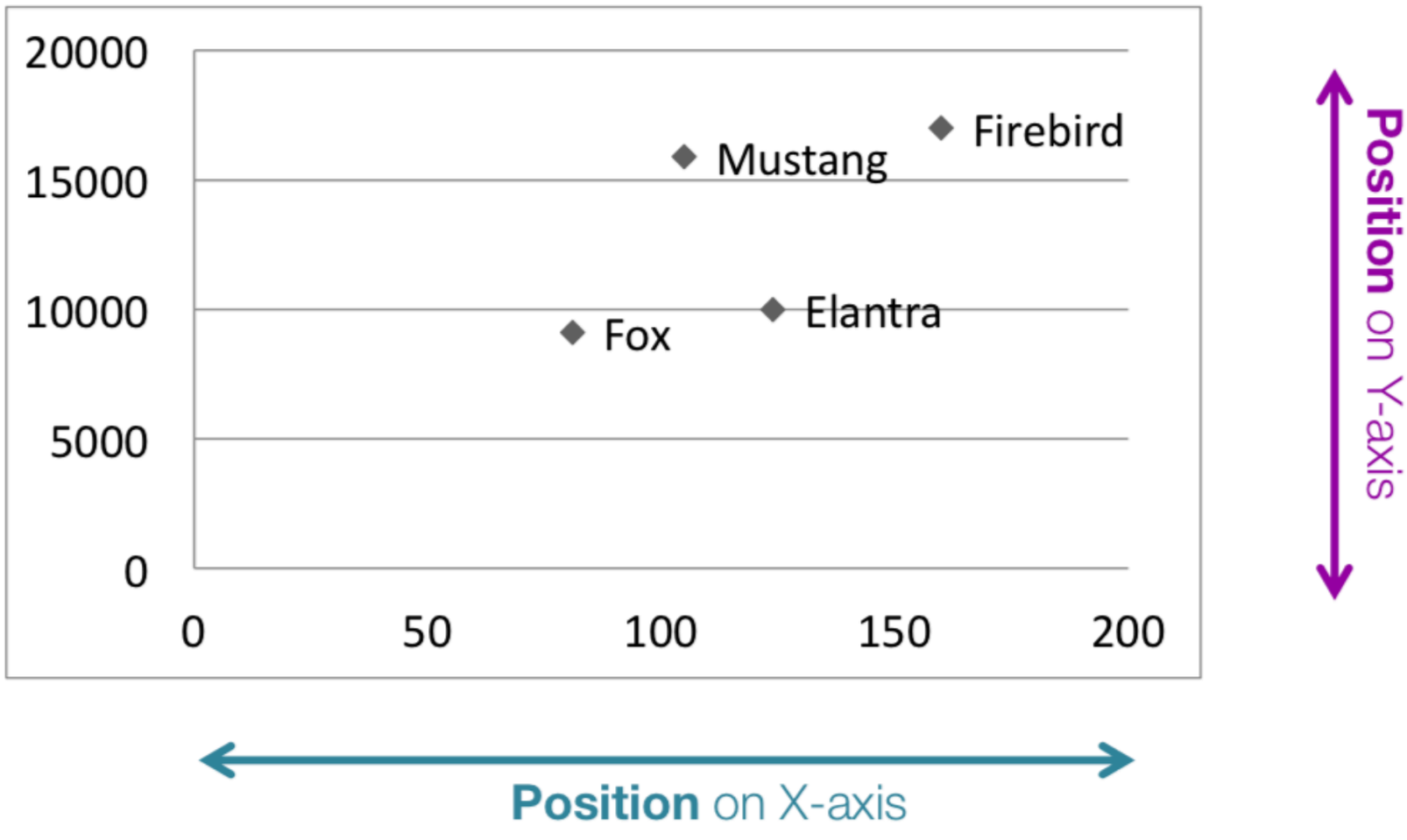
Color - Hue

Color - Saturation

Color - *Luminance*

MAPPING DATA TO VISUAL VARIABLES

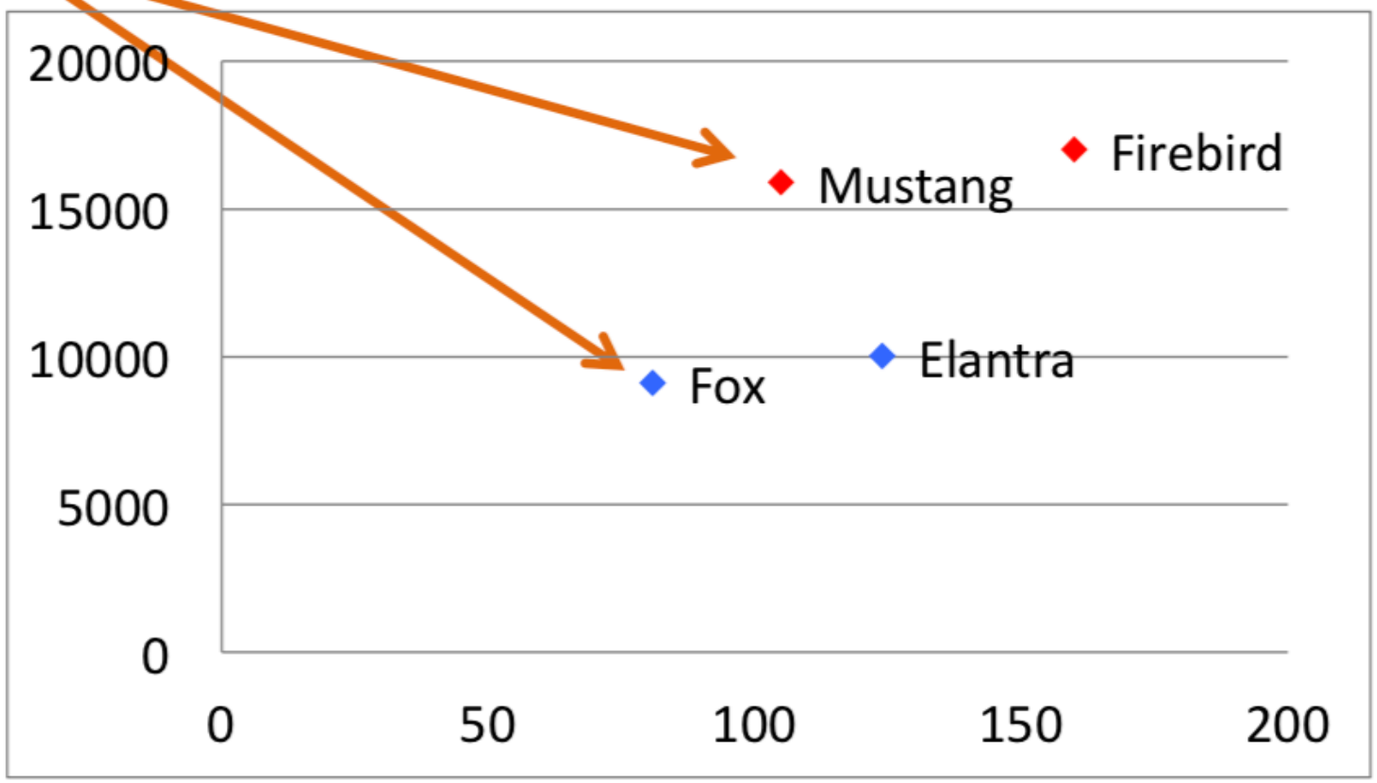
Model	Type	HP	Price	*Number Airbags*
Volkswagen Fox	Compact	81	9100	0 (no airbag)
Hyundai Elantra	Compact	124	10000	0 (no airbag)
Pontiac Firebird	Sport	160	17000	2 (driv. & pass.)
Ford Mustang	Sport	105	15900	1 (driver)



MAPPING DATA TO VISUAL VARIABLES

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)

Color (Hue)

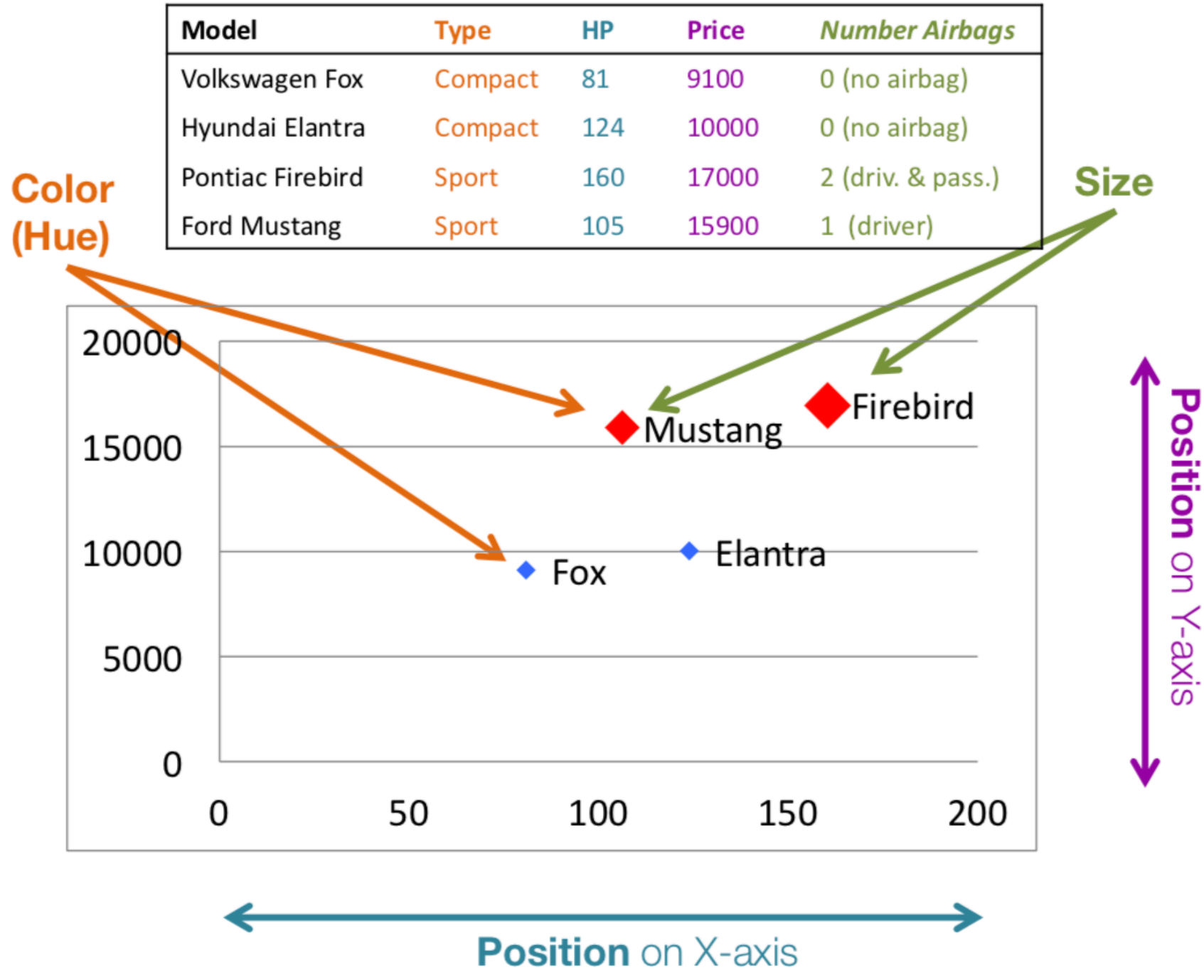


Position on Y-axis

Position on X-axis

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)



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.

Mackinlay, Jock. "Automating the design of graphical presentations of relational information." ACM Transactions on Graphics (TOG) 5.2 (1986): 110-141.

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

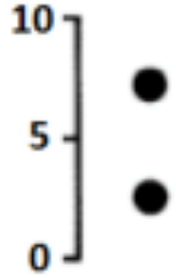
Effectiveness principle:
choosing the best representation to your data

Importance of attributes should match the "saliency" of the channel;

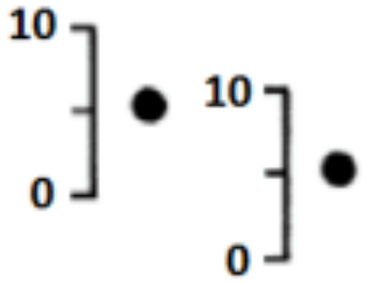
Most important attributes should be encoded using the most effective and noticeable channels.

Mackinlay, Jock. "Automating the design of graphical presentations of relational information." ACM Transactions on Graphics (TOG) 5.2 (1986): 110-141.

Position Common Scale



Position Non-Aligned Scale



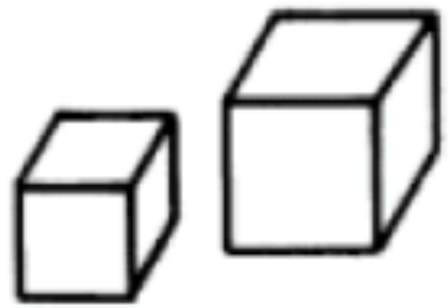
Length

Direction

Angle

Area

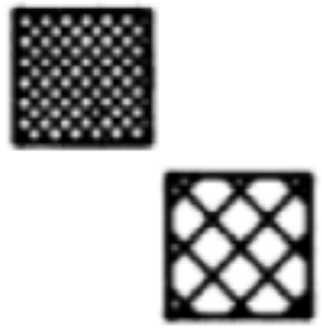
Volume



Curvature

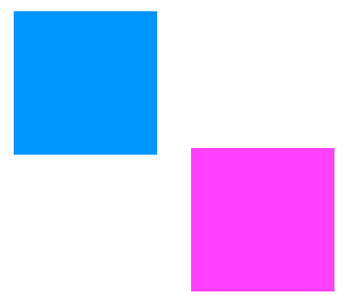
Shape

Shading



Color saturation

Hue



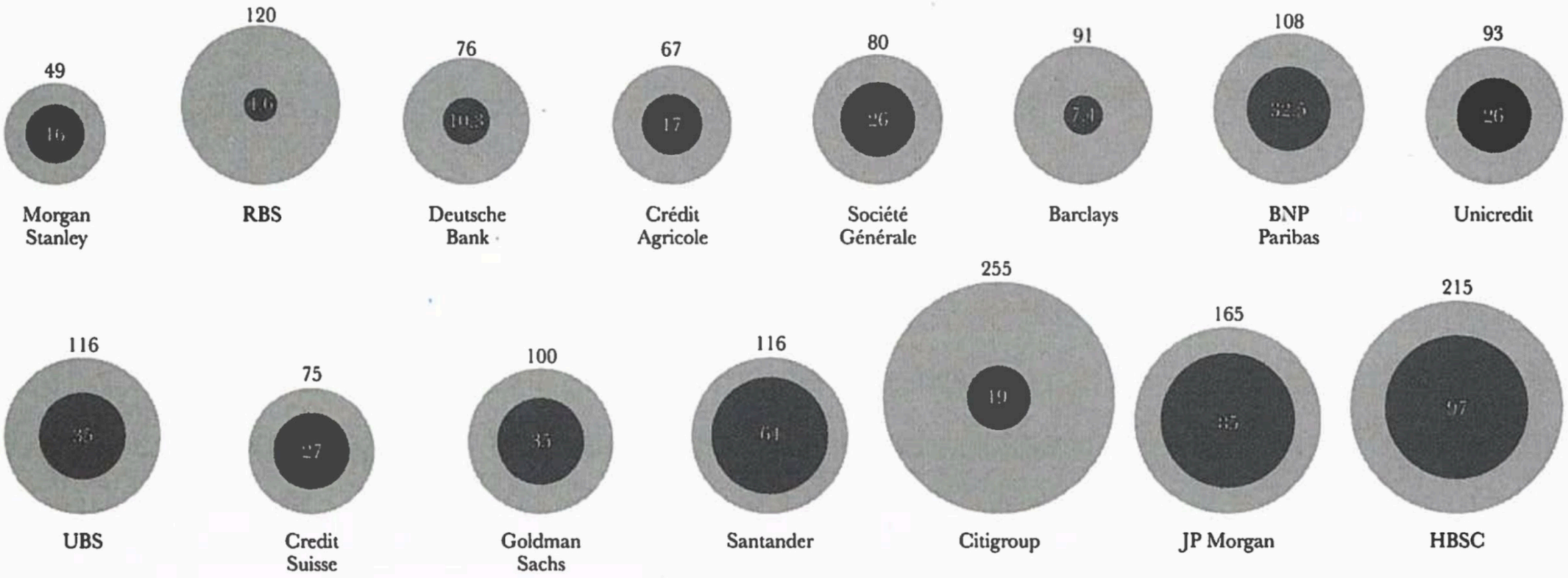
Market Capitalization of the World's Biggest Banks

In billions of dollars

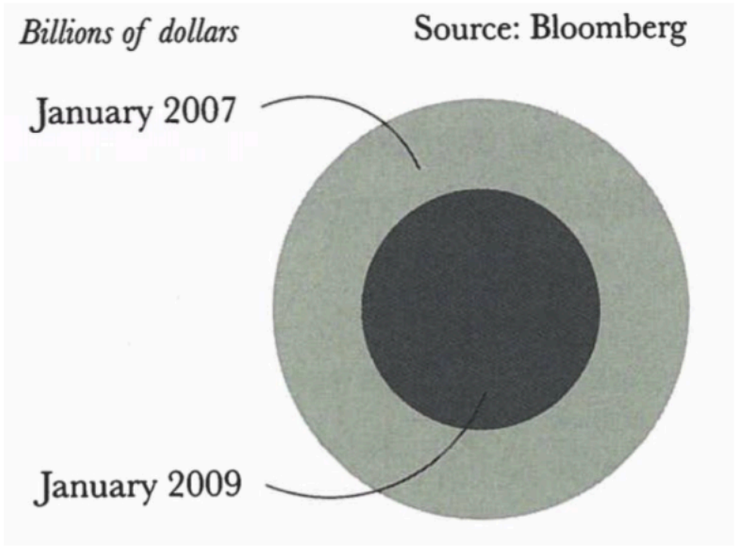
January 2007

January 2009

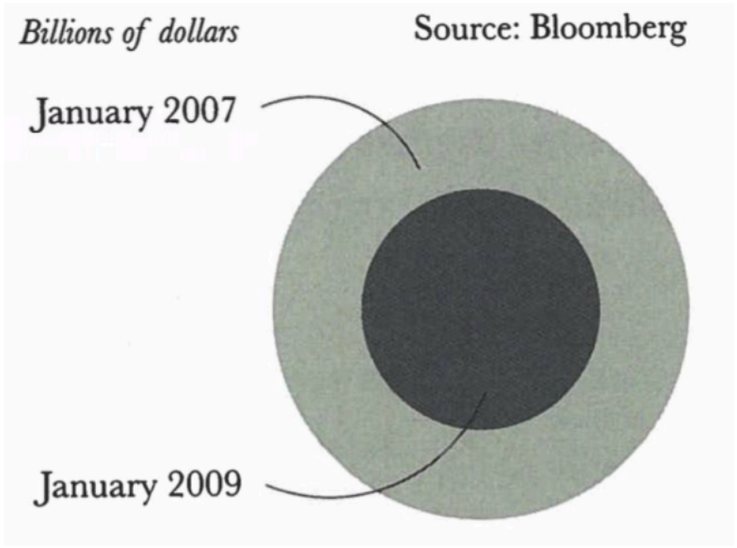
Source: Bloomberg

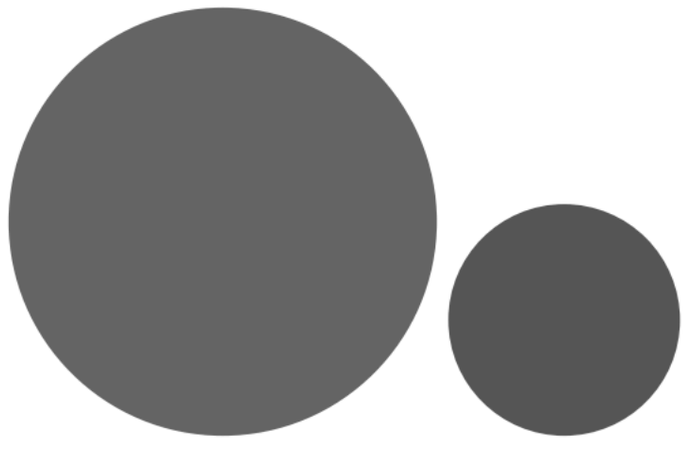


How to decode bubbles?

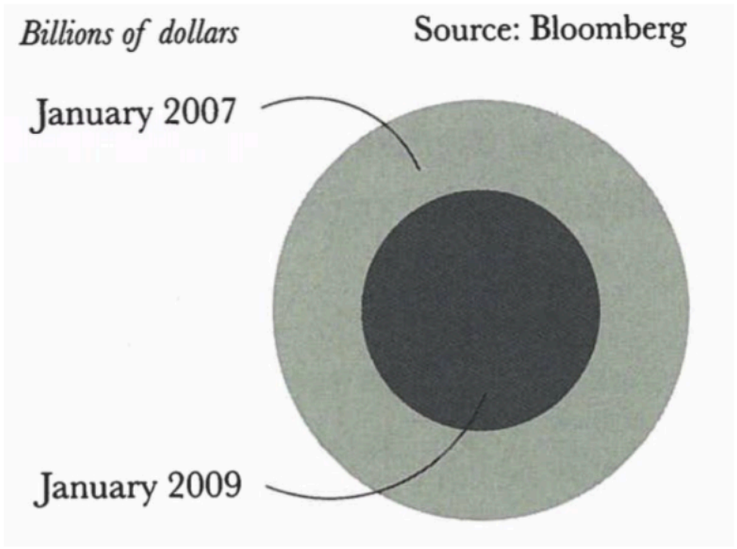


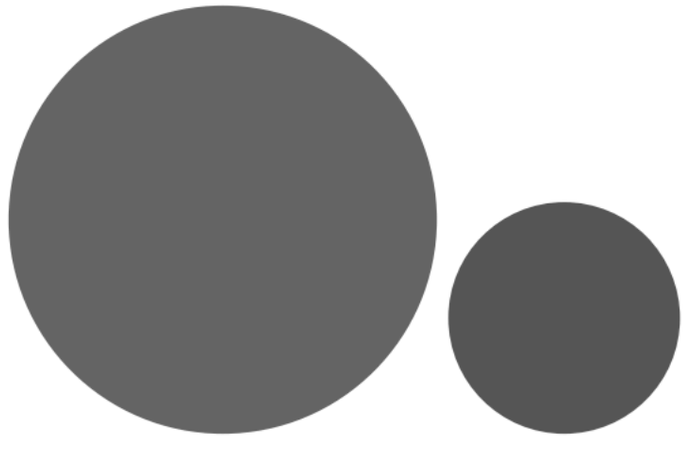
How to decode bubbles?





How to decode bubbles?





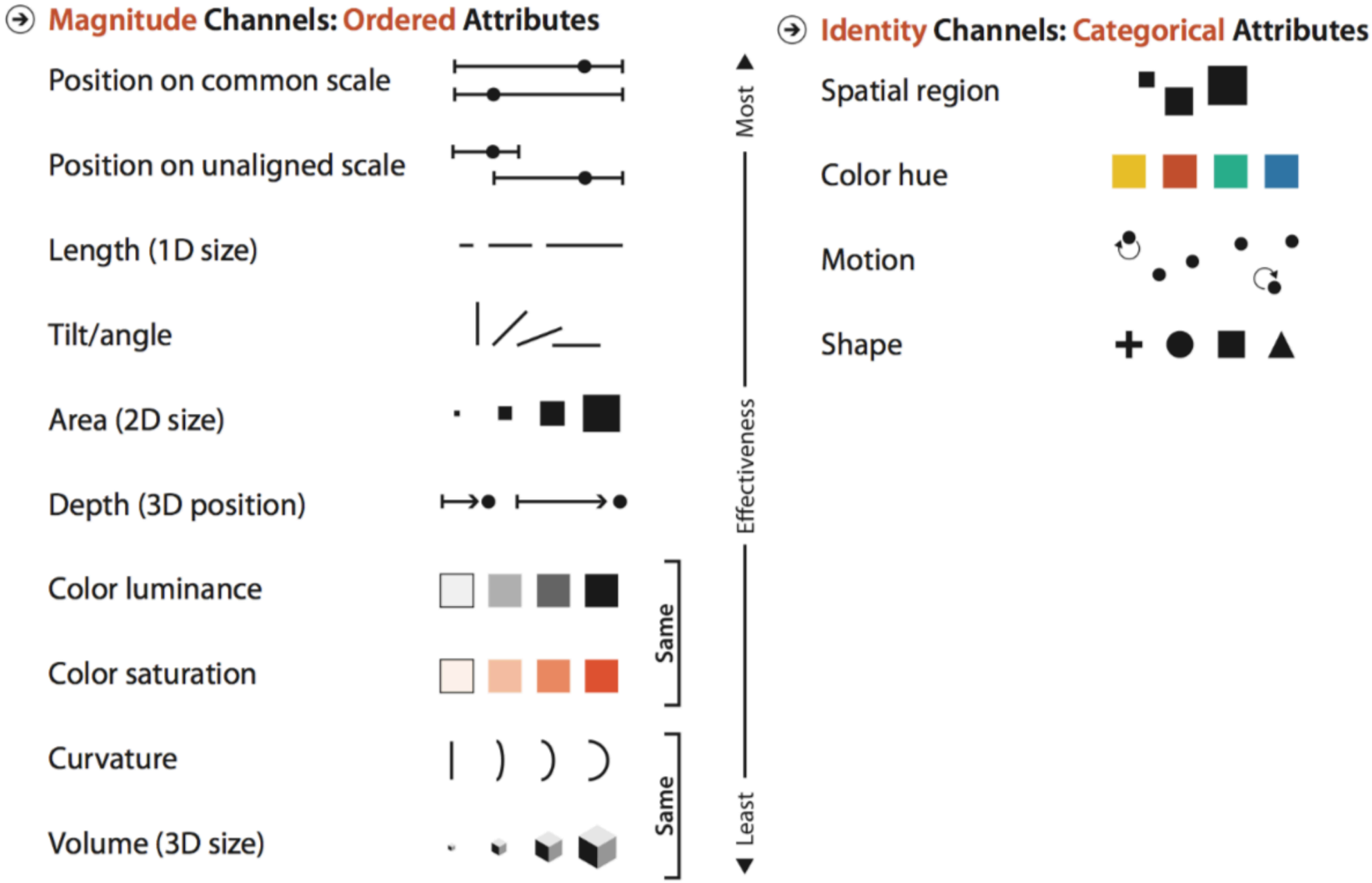
a)

b)

c)

EFFECTIVENESS PRINCIPLE

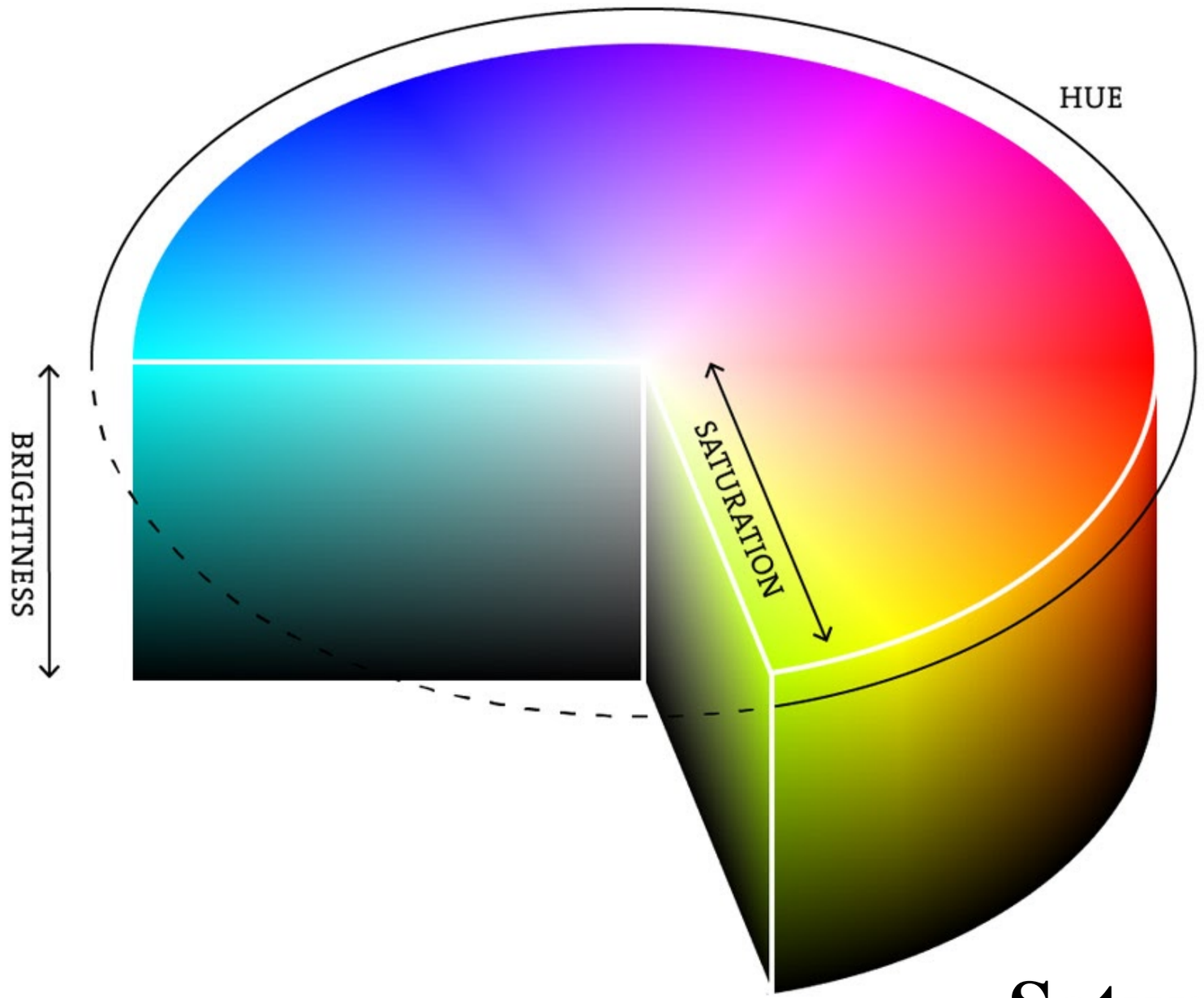
Some variables are perceptually better than others



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

HSB MODEL

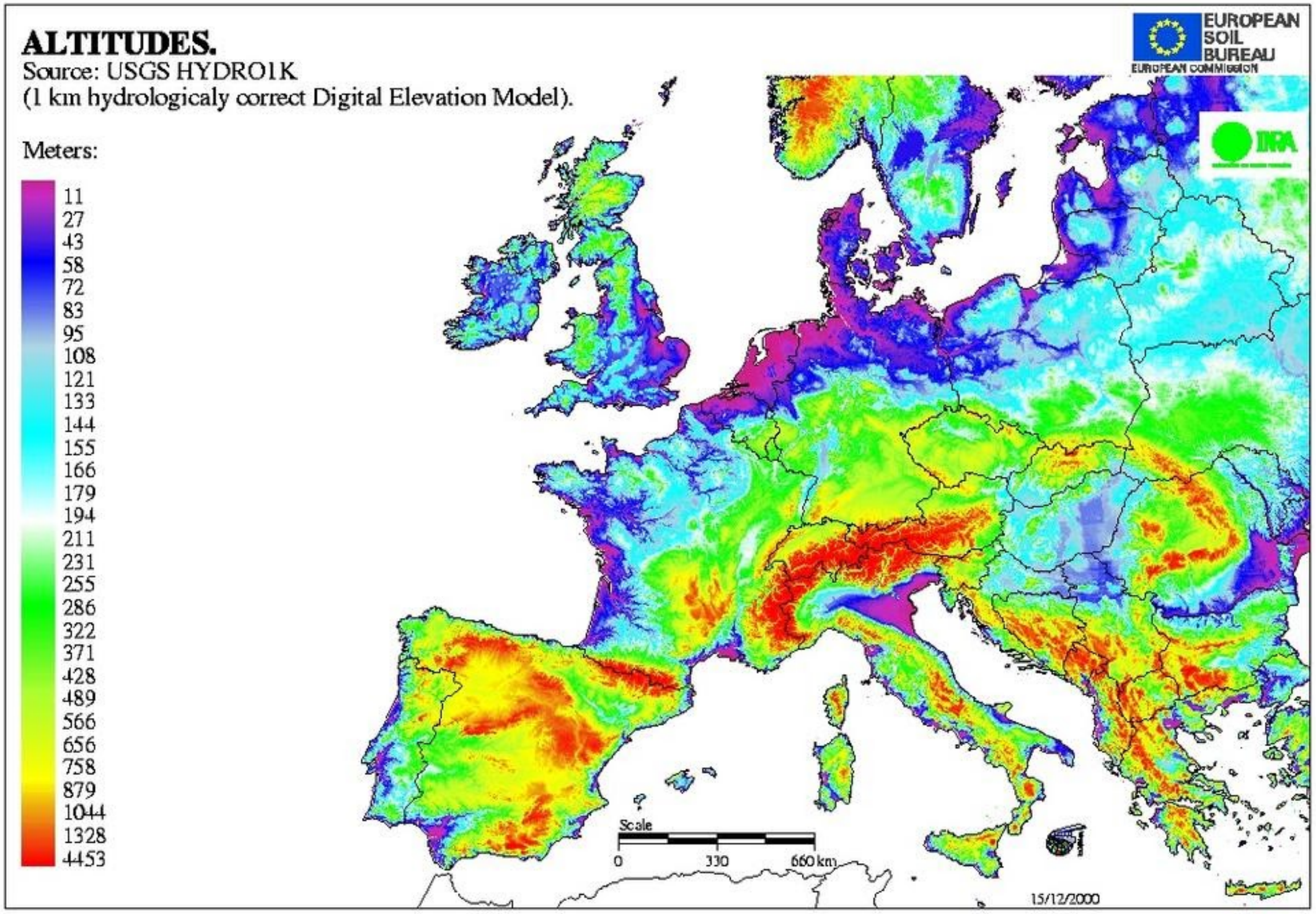
Hue



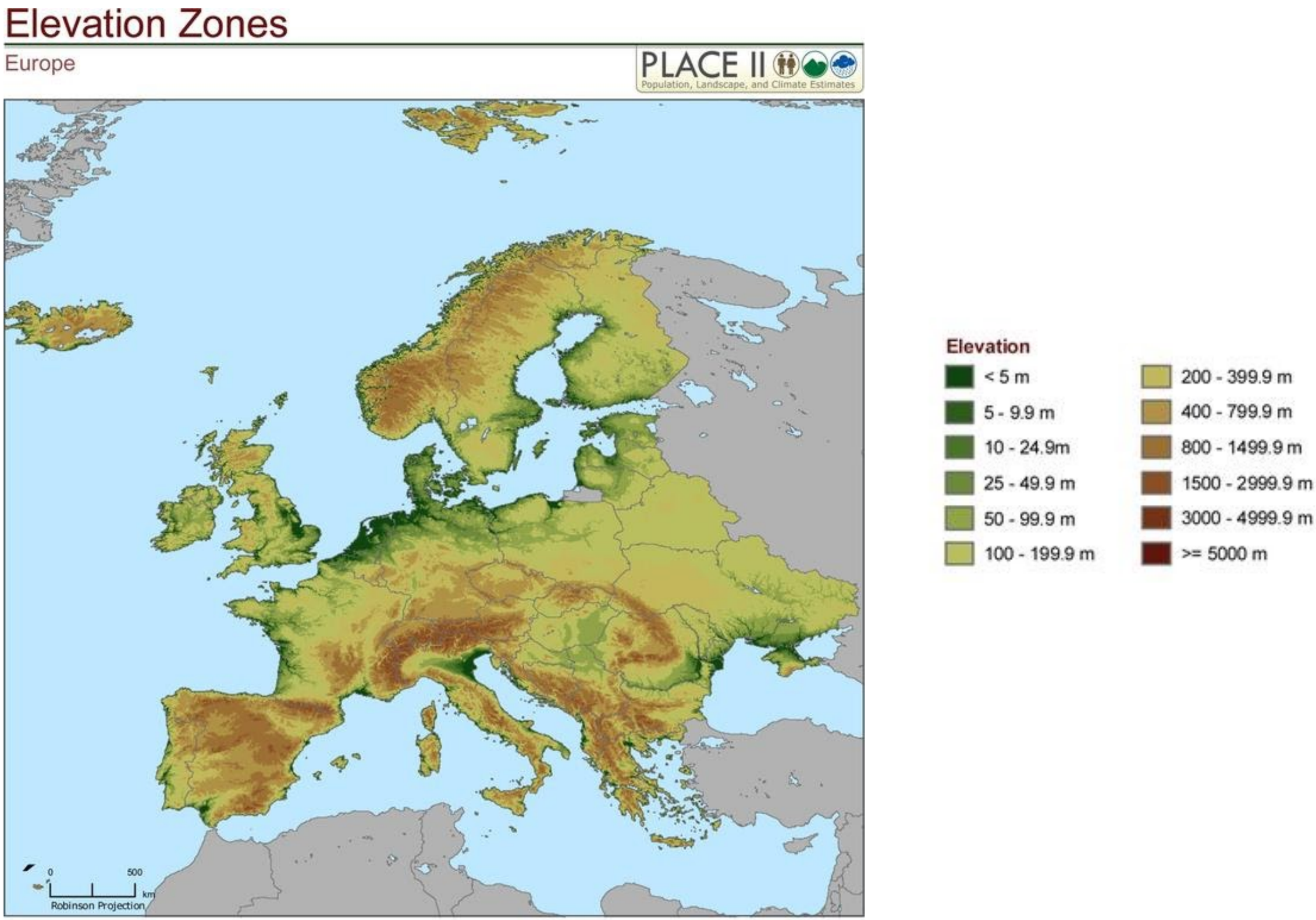
Brightness

Saturation

RAINBOW COLOR MAP



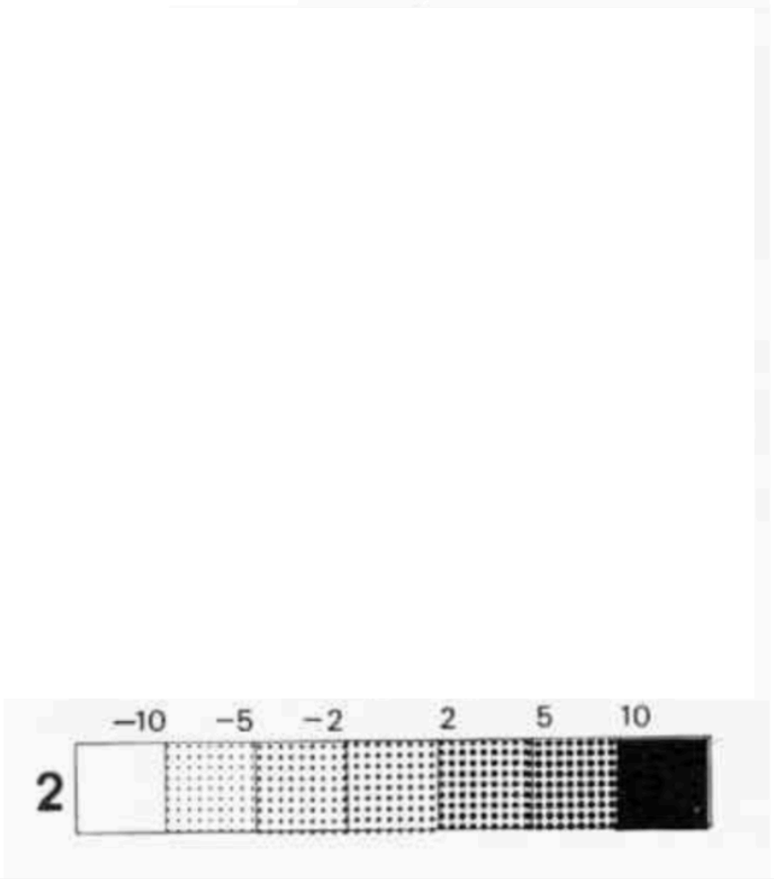
RAINBOW COLOR MAP



Rainbow Scale Considerations

Map 1

1

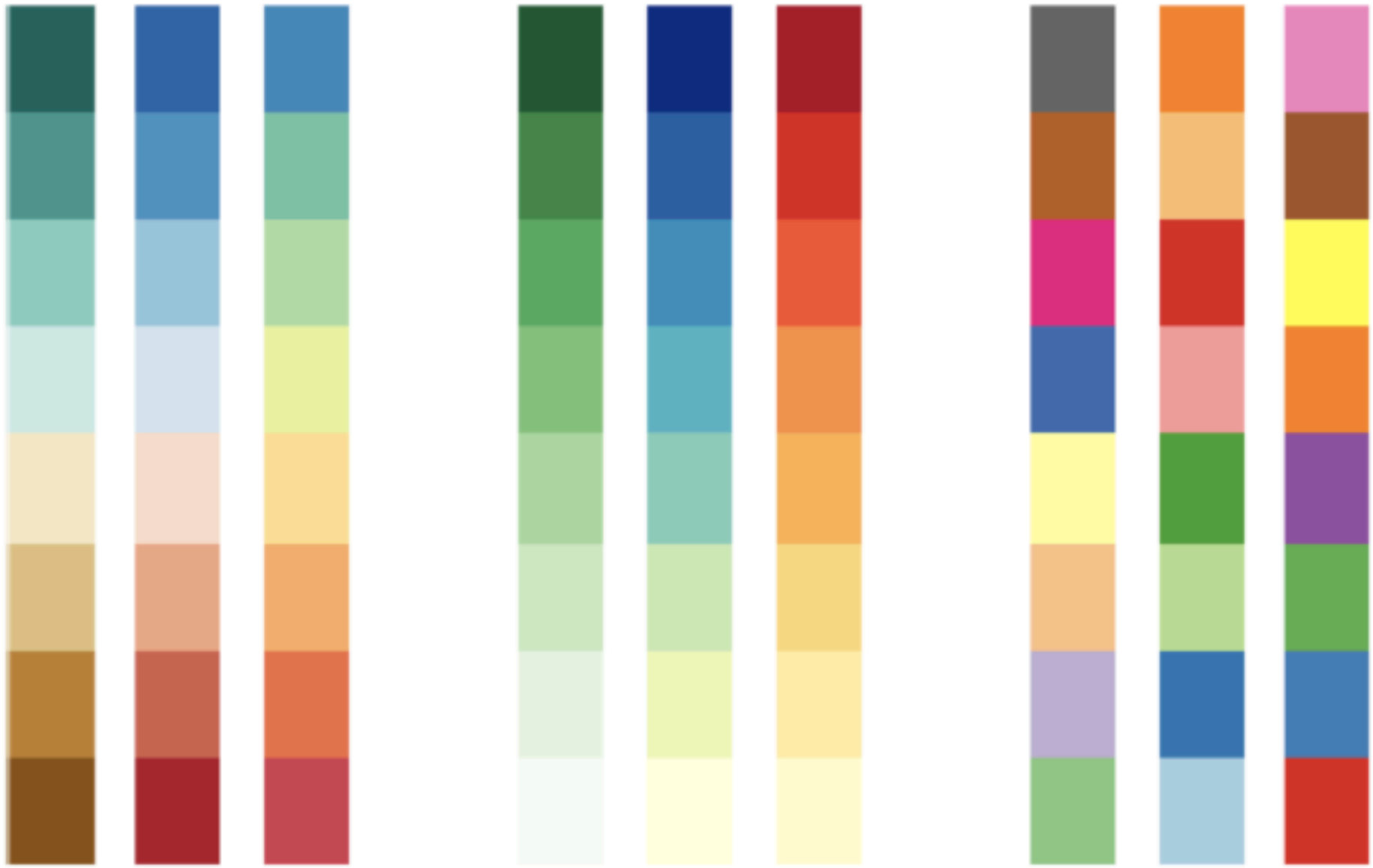


2

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

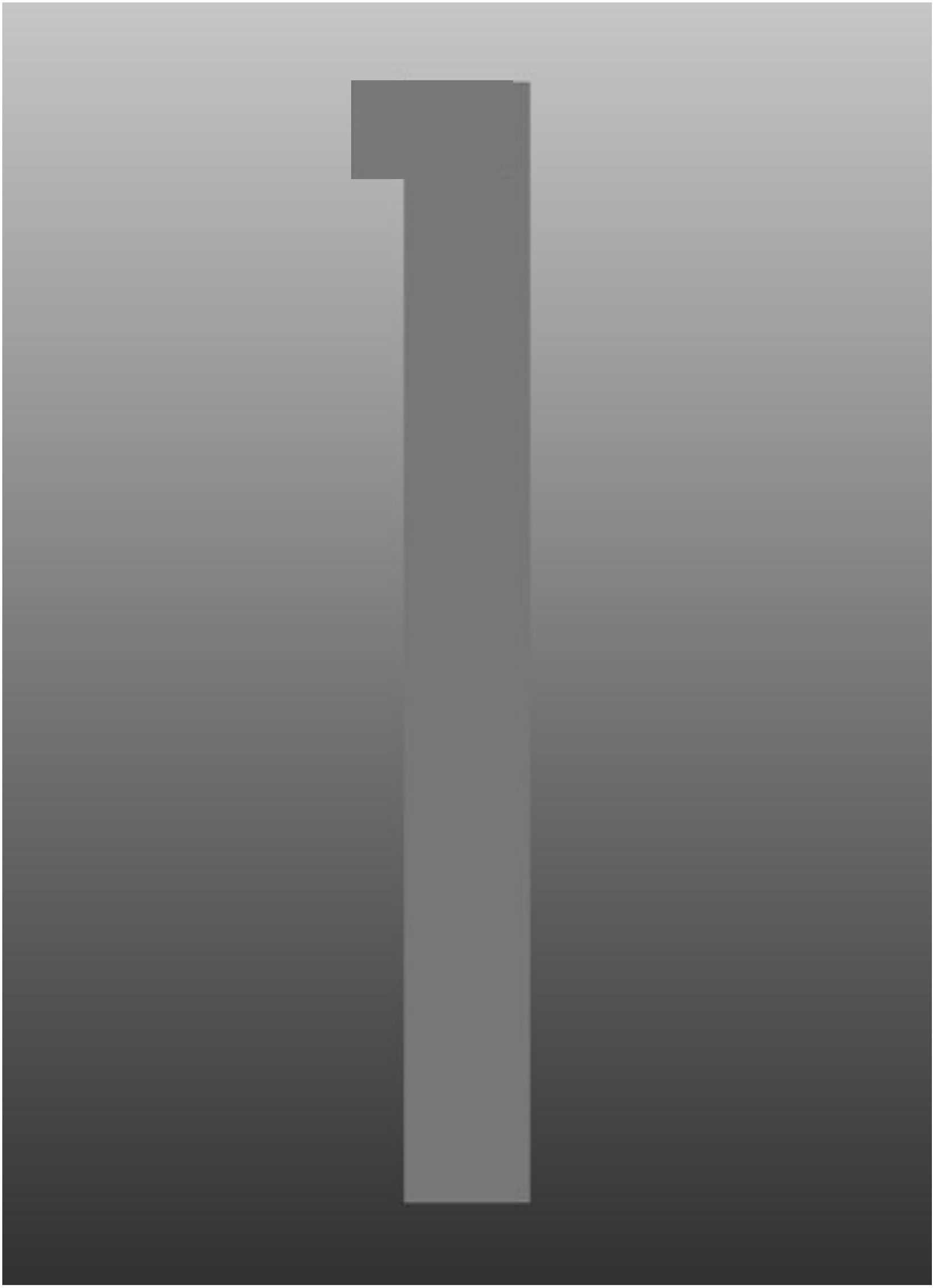
Top 5 drugs: country-level sales rank

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

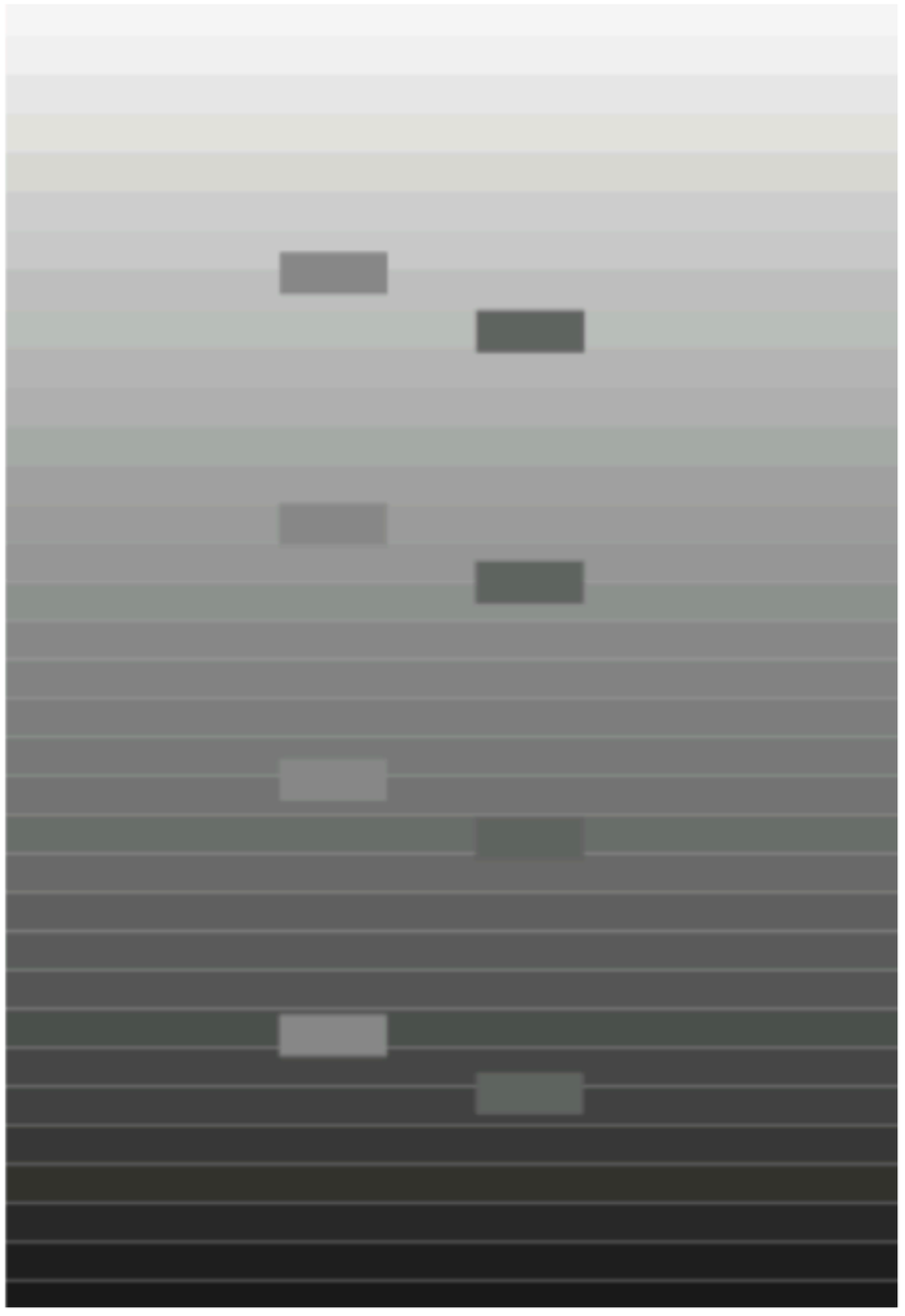


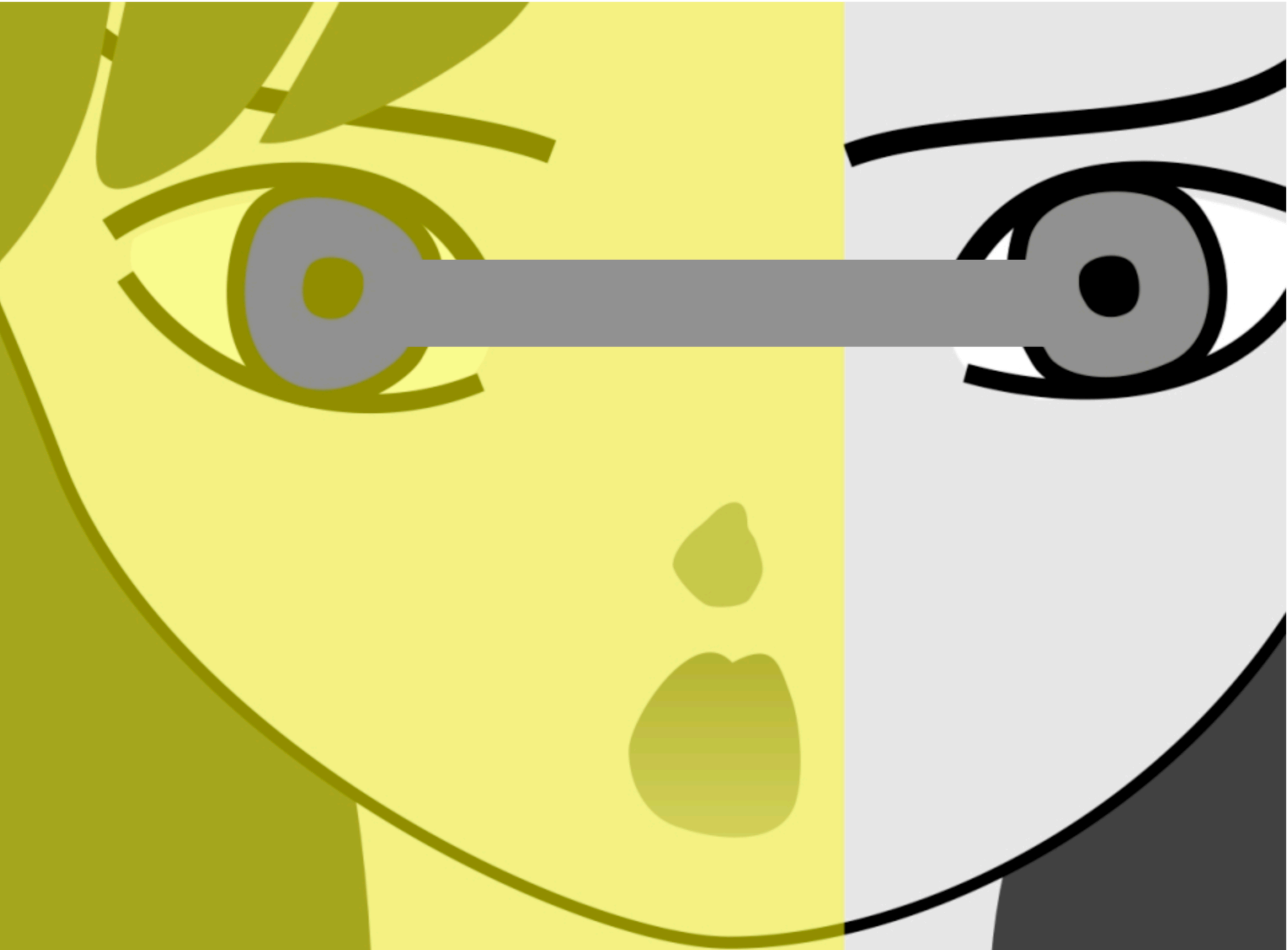
COLOR PERCEPTION IS LOCAL

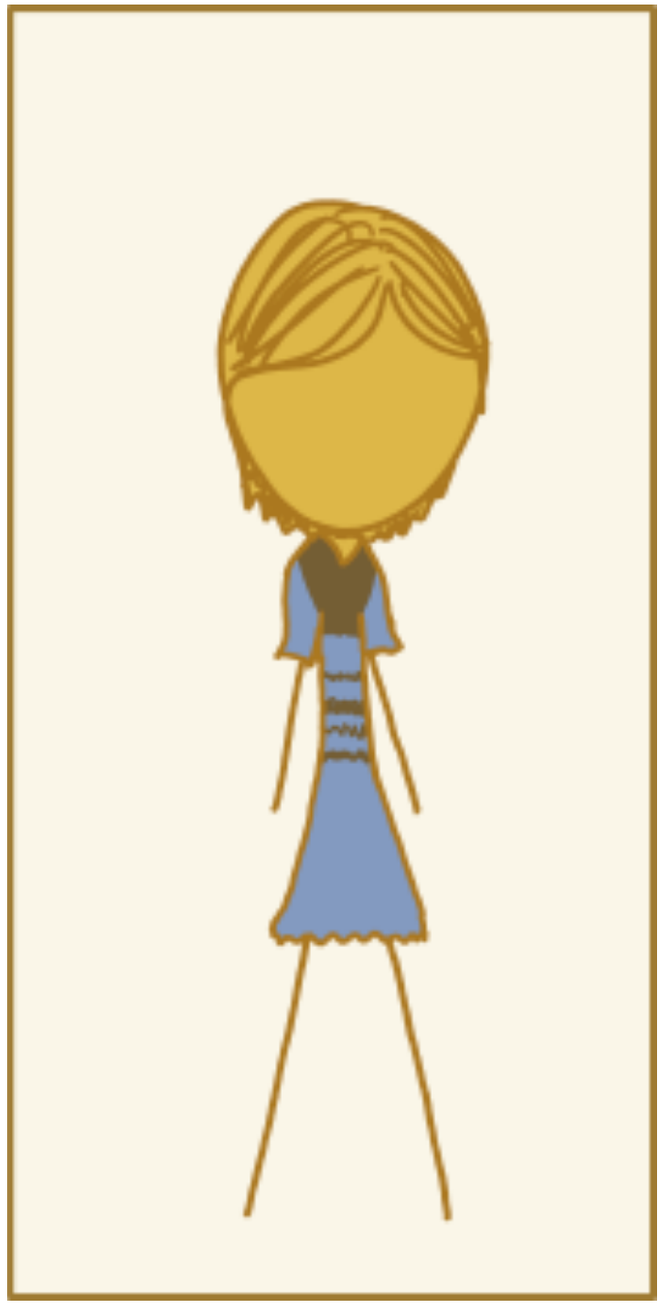
- Luminance perception is not absolute, but contrast relative ... and local!

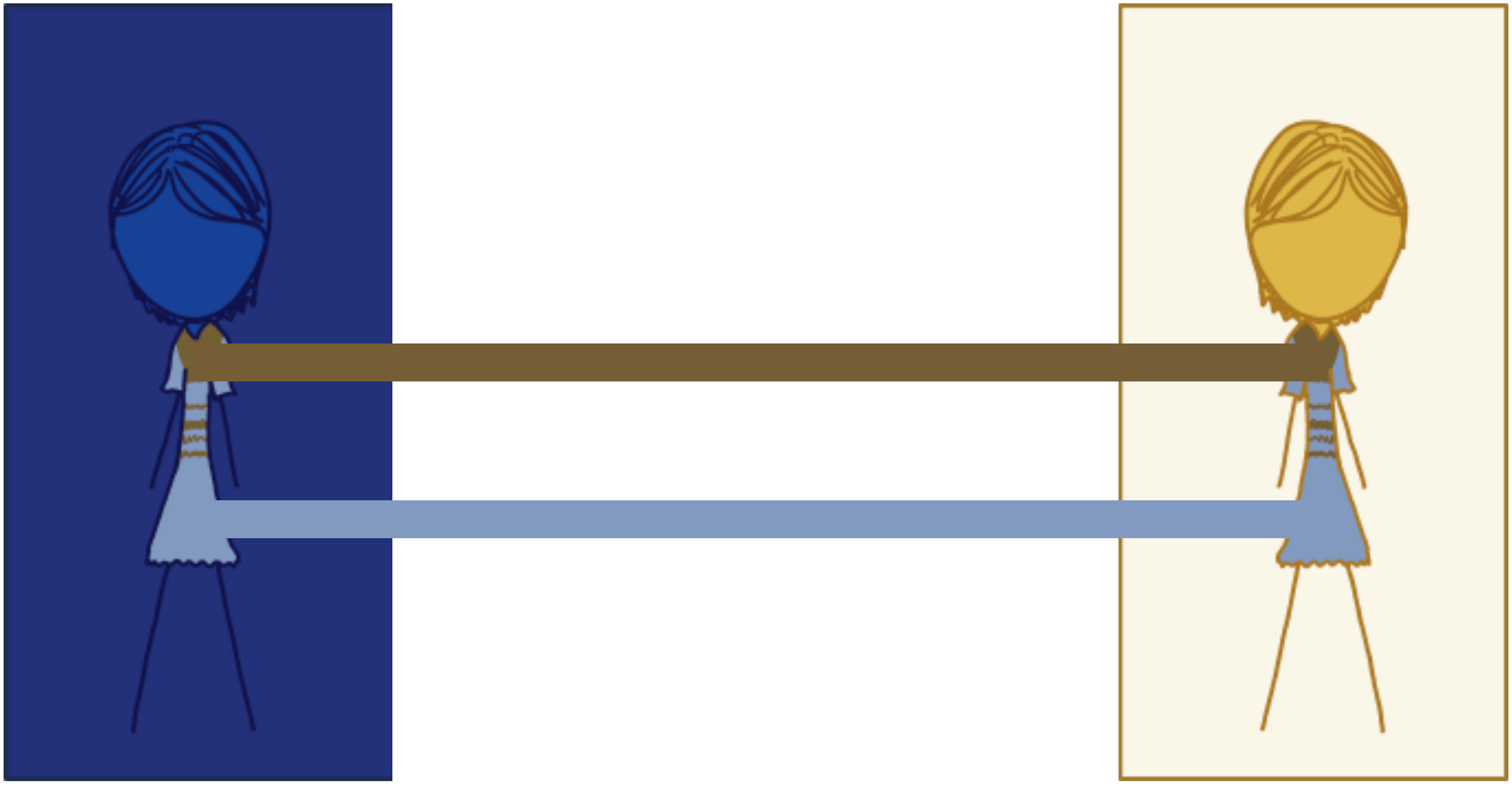
- Discern objects(foreground) from background

- Gray color scales are not recommended for more than 5 segments... too "imprecise" to convey quantitative data



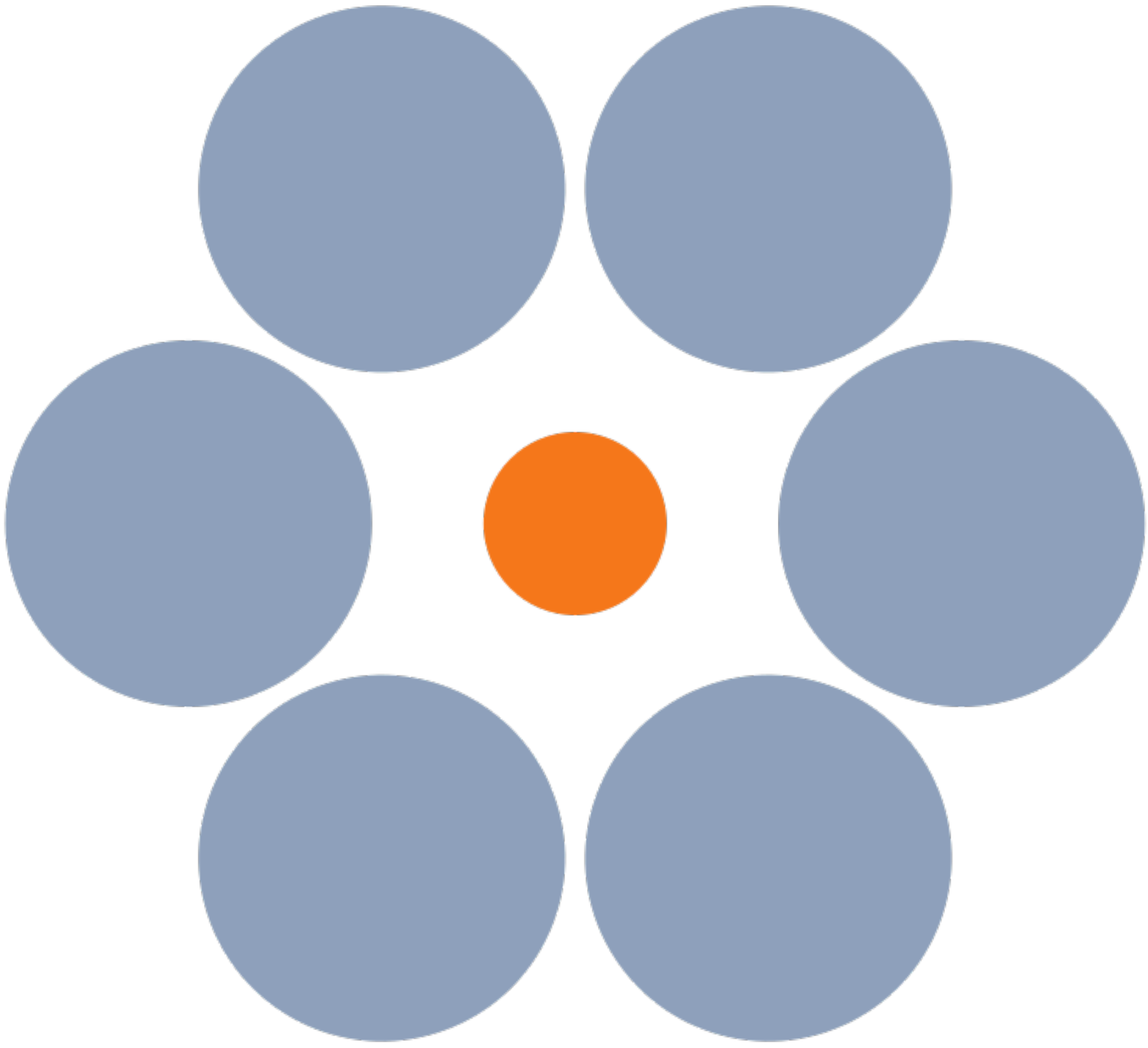


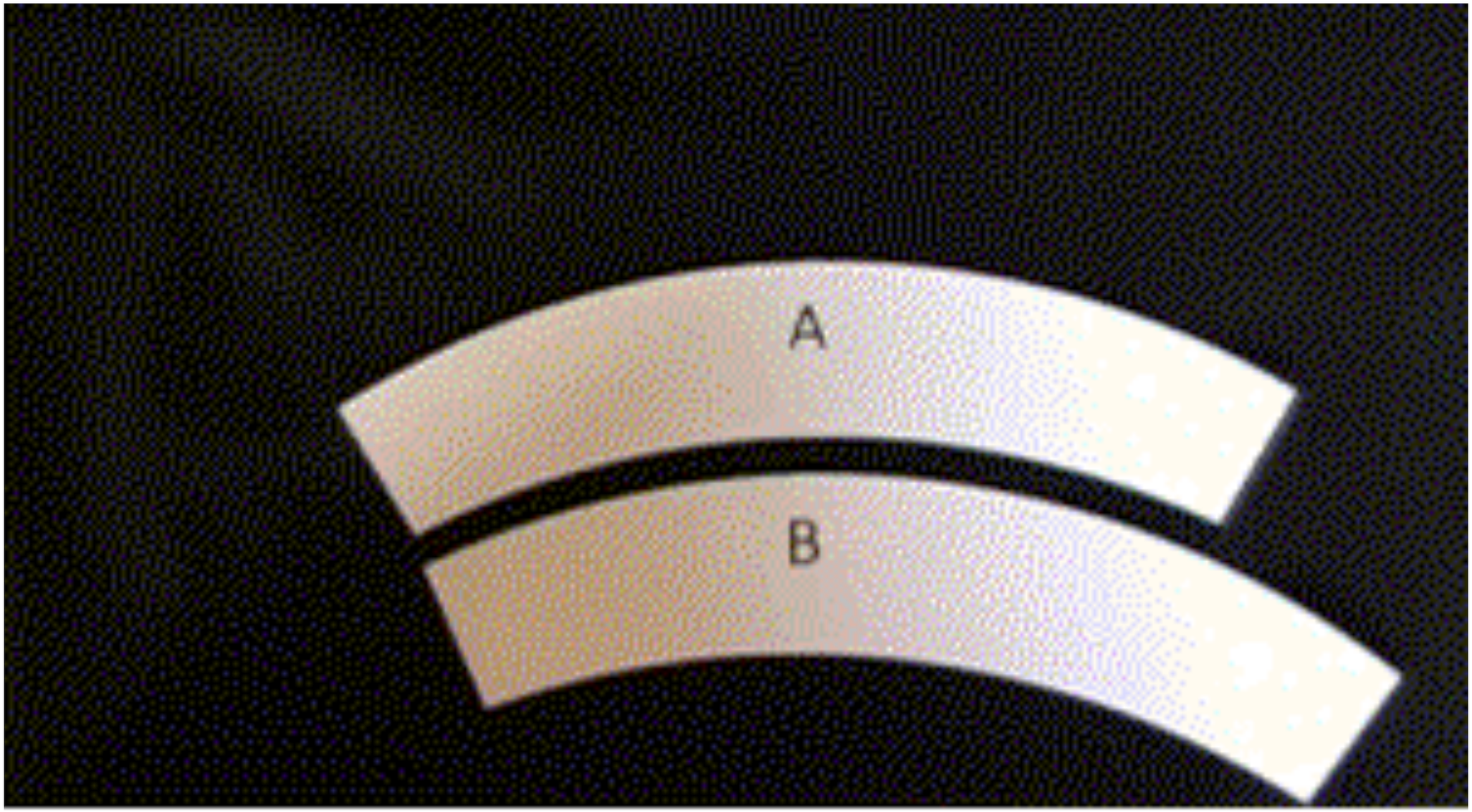


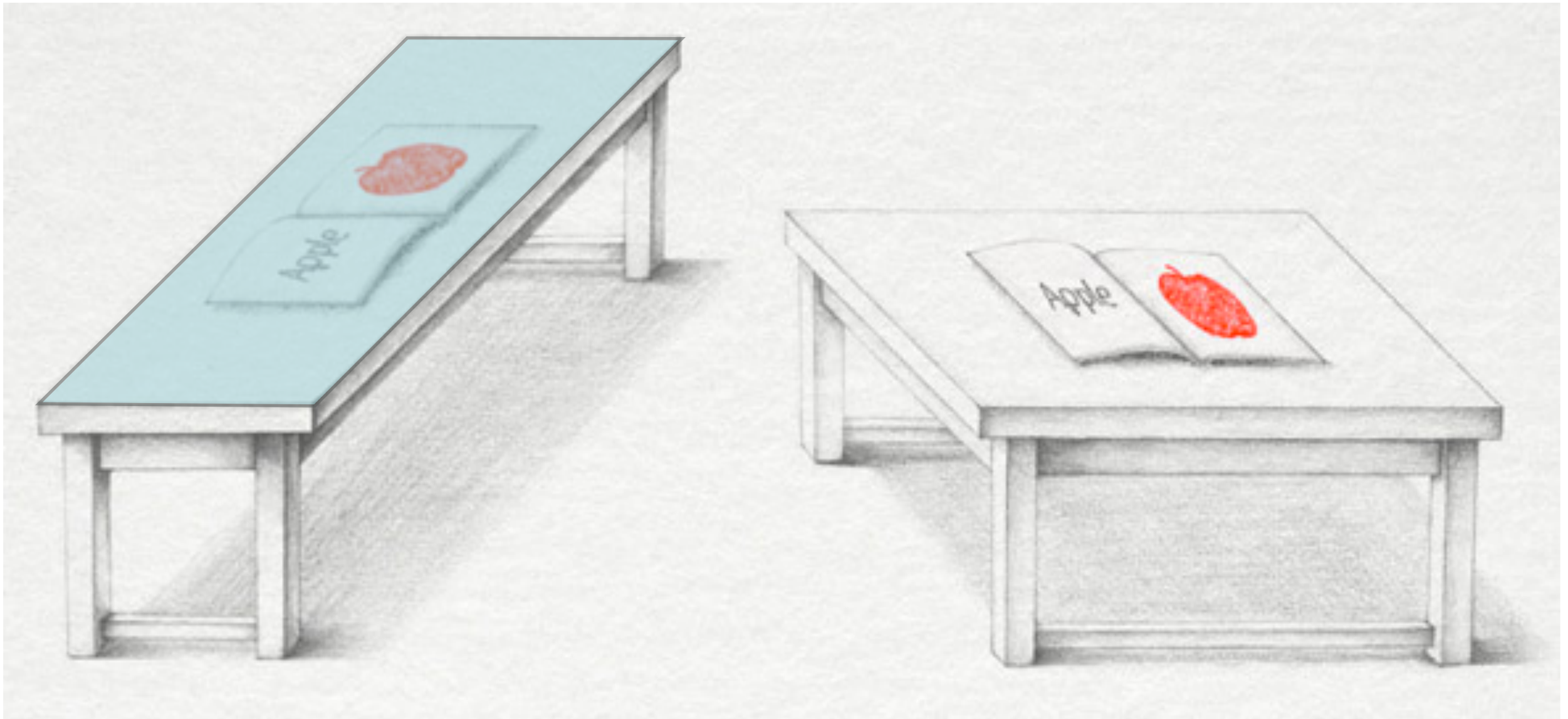


length?

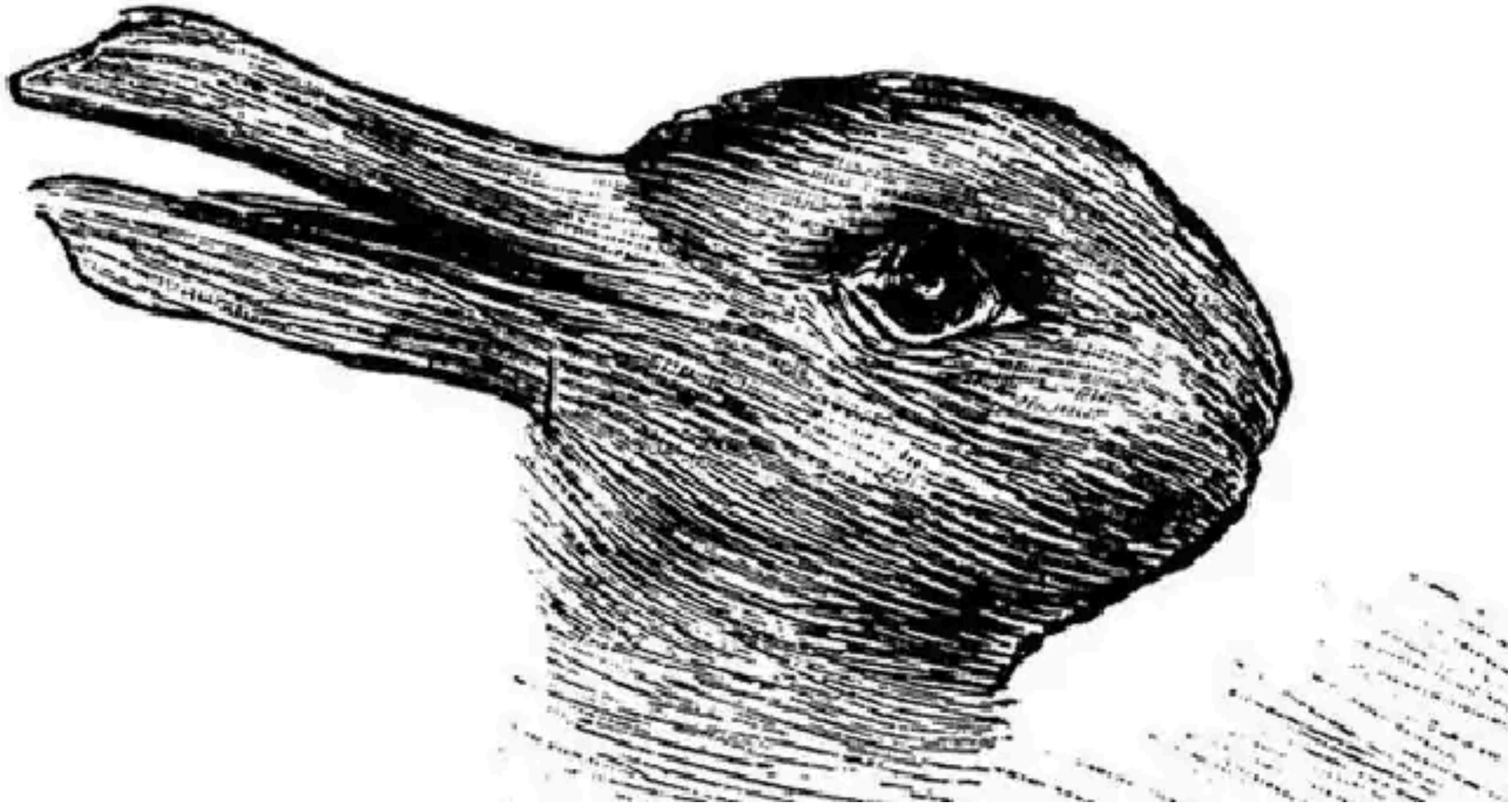
SIZE?

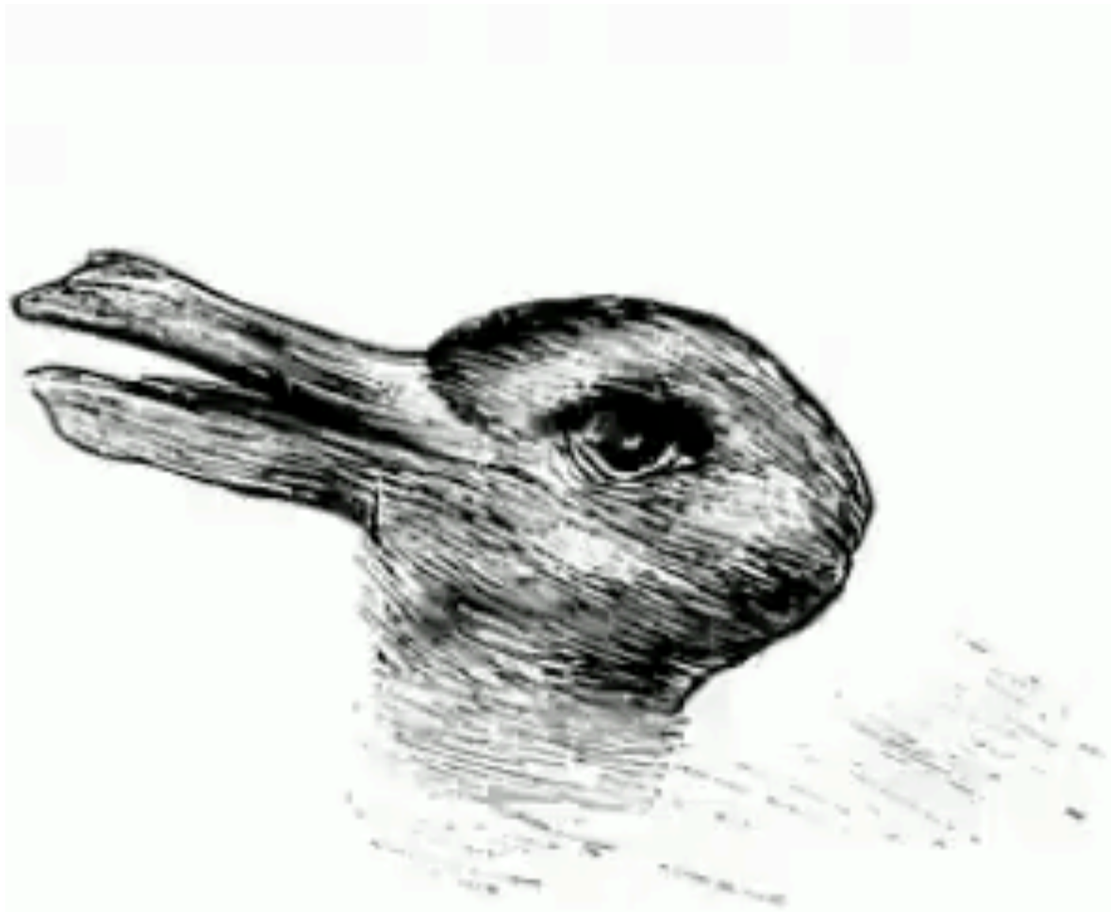


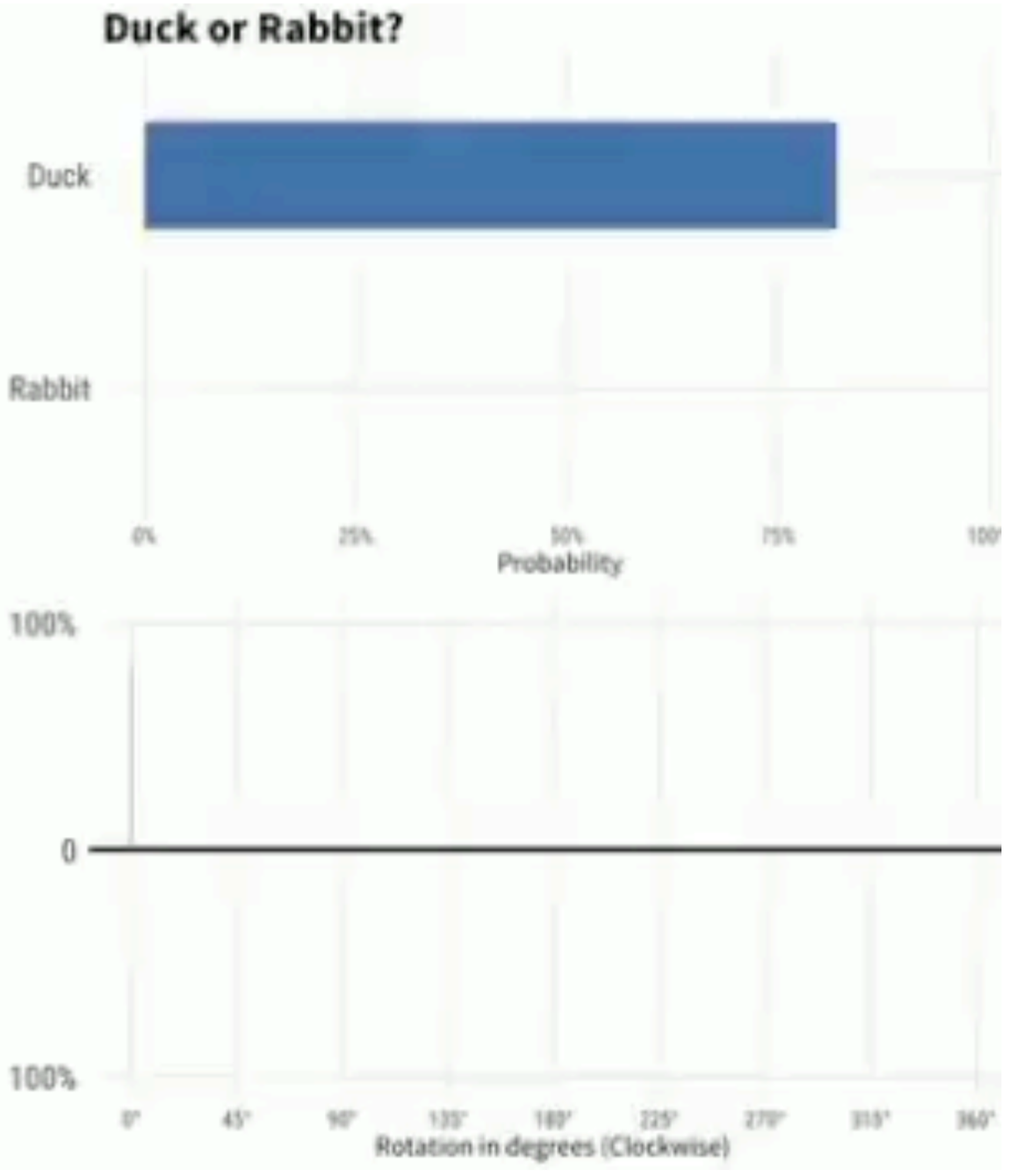

A
B


Apple
Apple

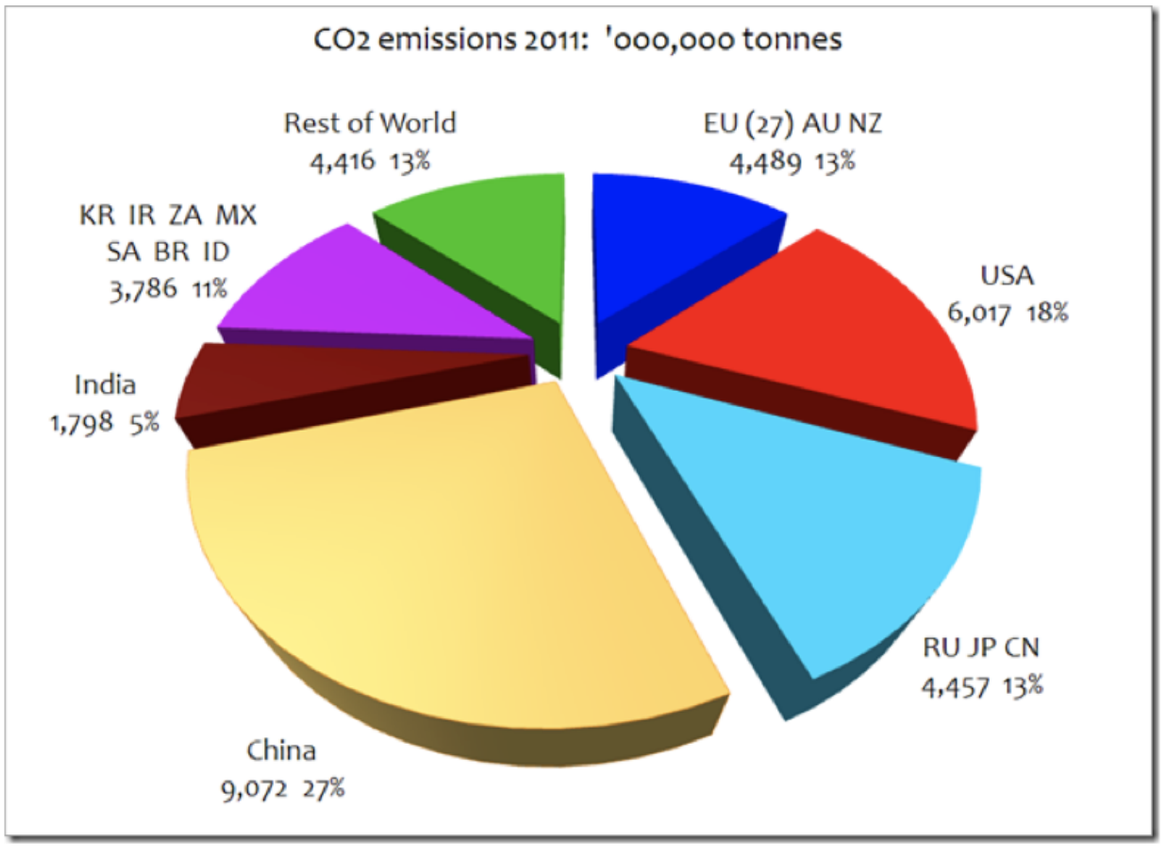
DECEPTION

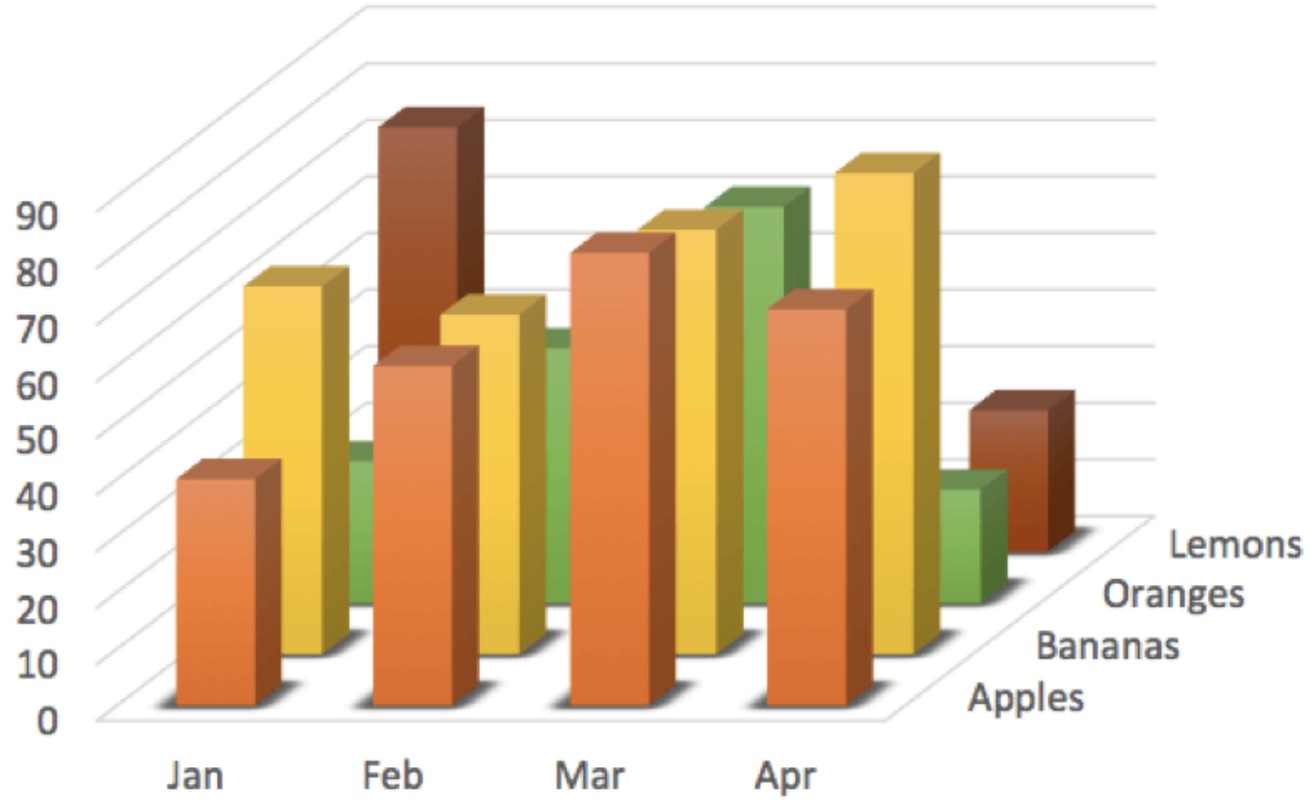


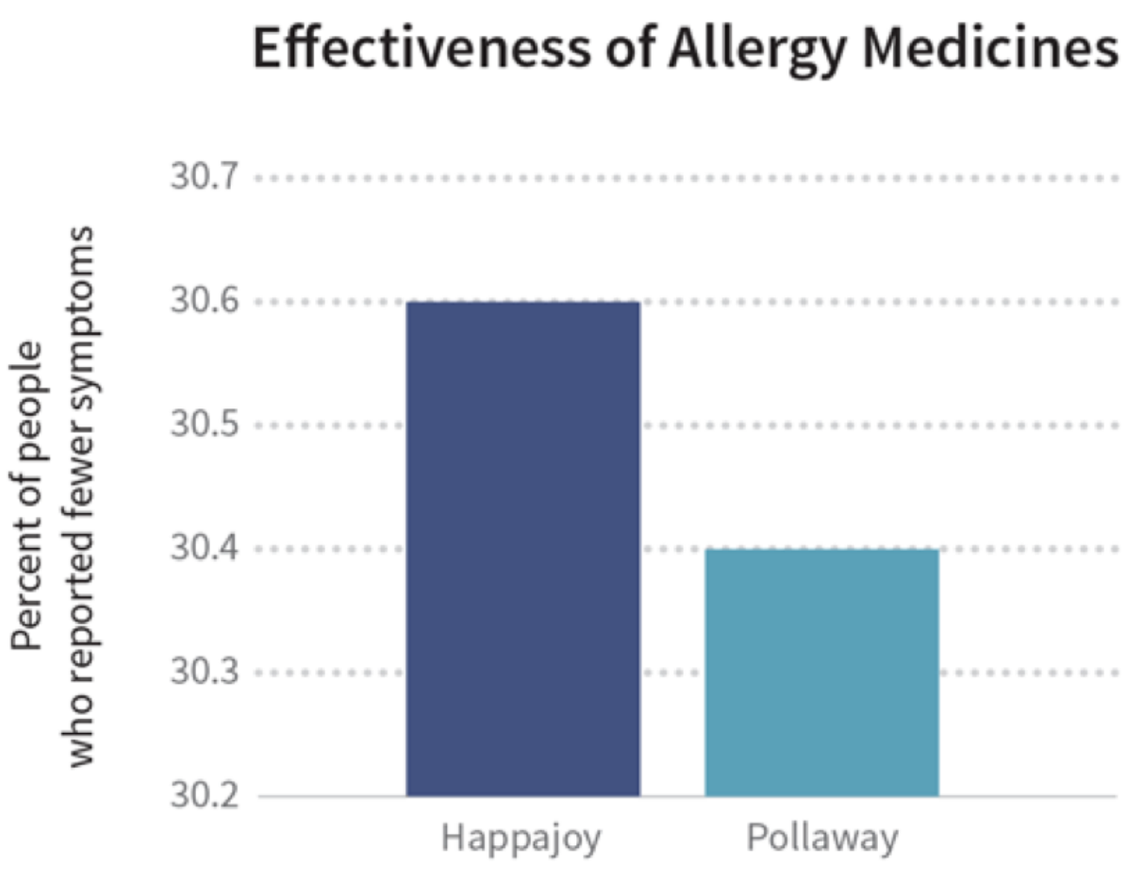


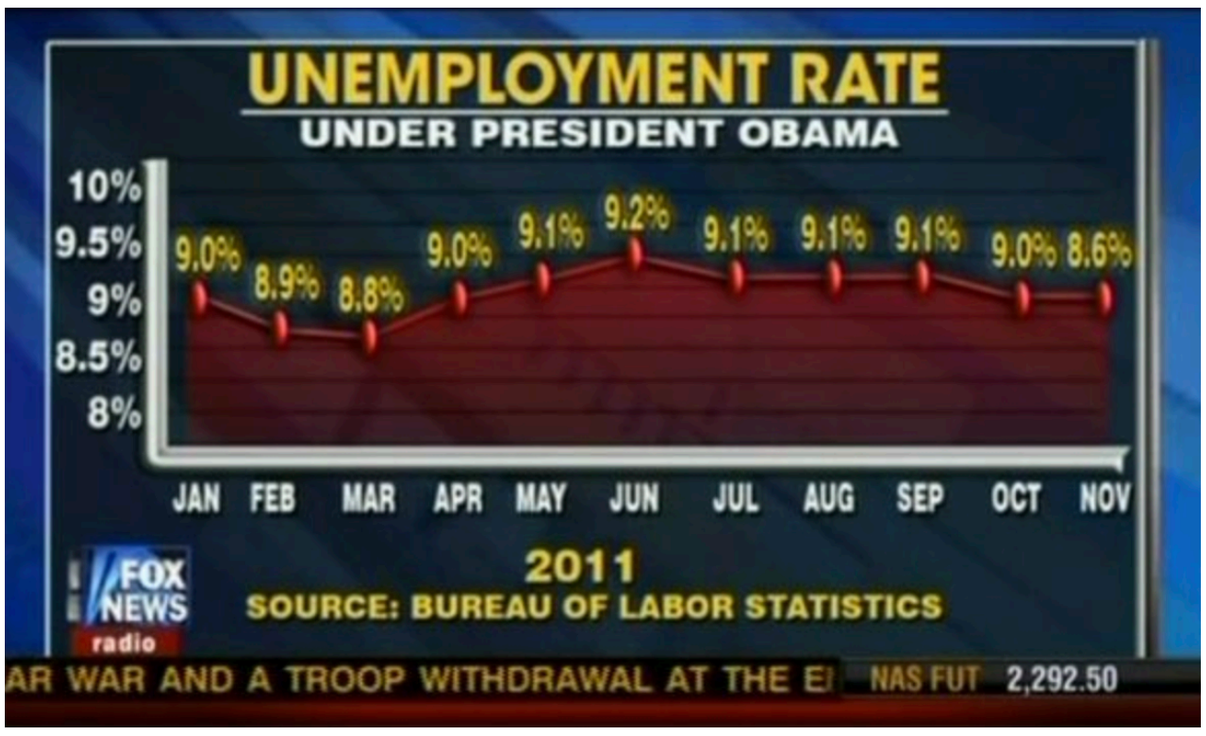


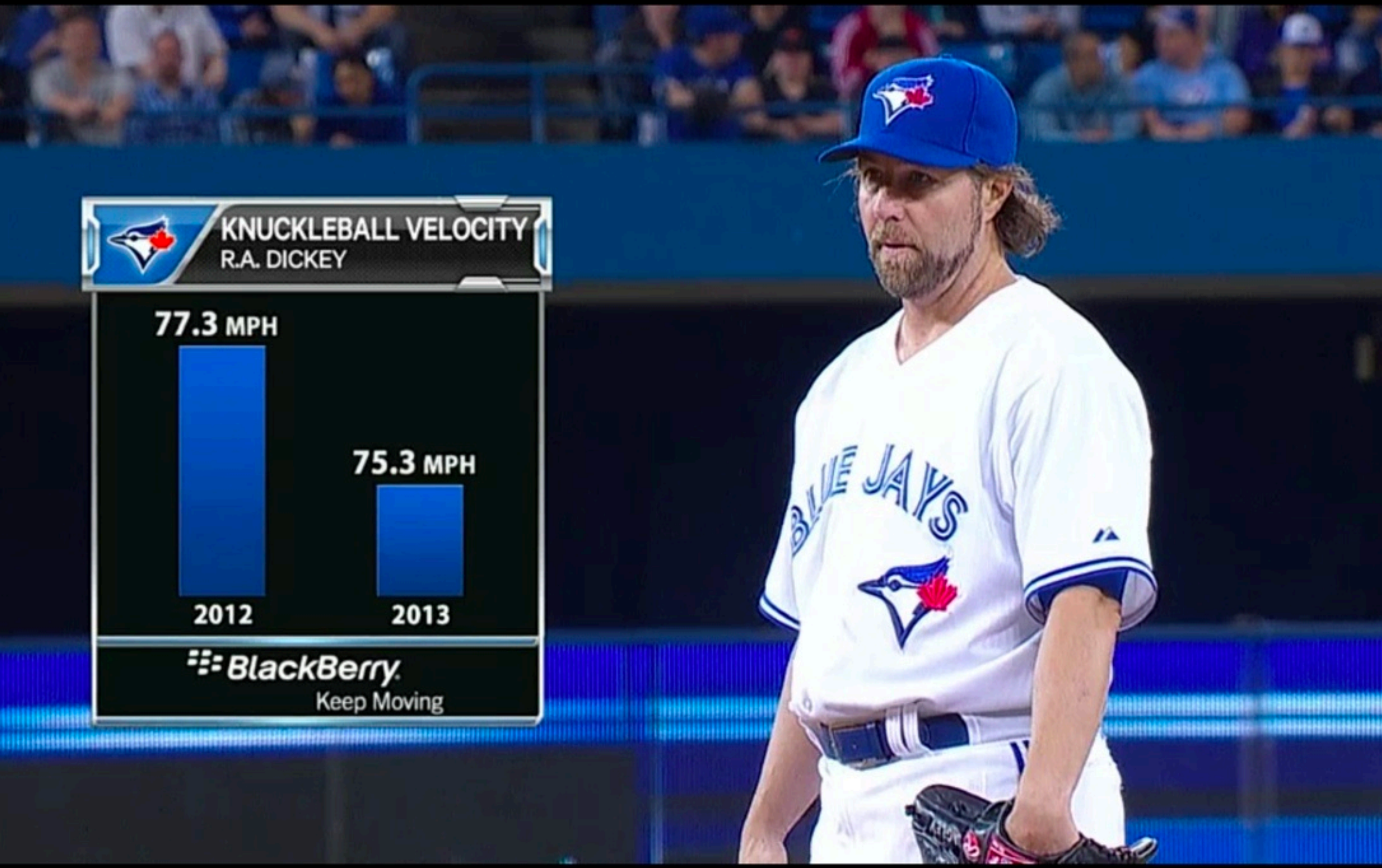
Predictions provided by the Google Cloud Vision API
Animation by Max Woolf (@minimaxir)


CO2 emissions 2011: '000,000 tonnes
Rest of World
4,416 13%
EU (27) AU NZ
4,489 13%
KR IR ZA MX
SA BR ID
3,786 11%
USA
6,017 18%
India
1,798 5%
RU JP CN
4,457 13%
China
9,072 27%


90
80
70
60
50
40
30
20
10
0
Jan
Feb
Mar
Apr
Lemons
Oranges
Bananas
Apples

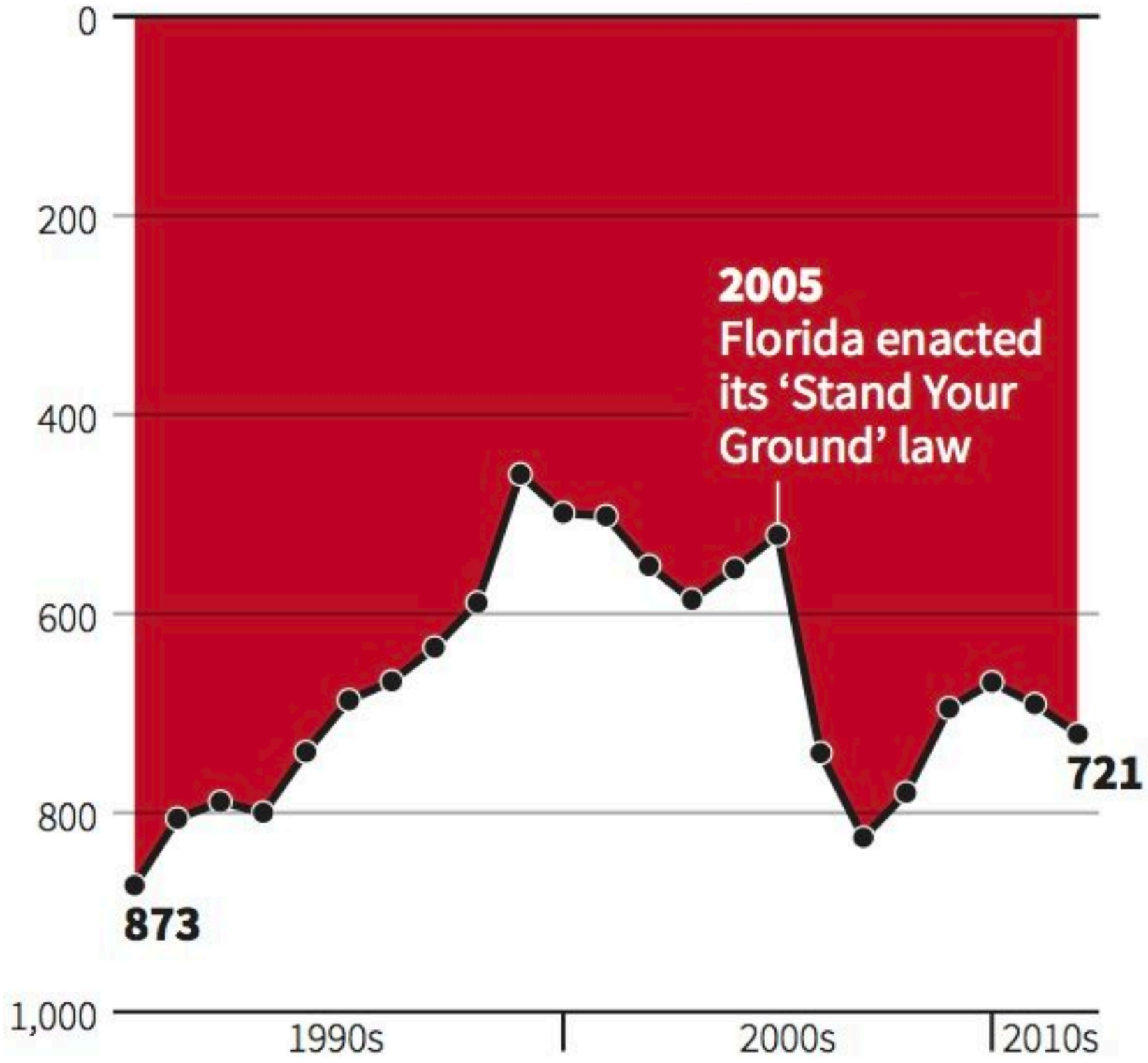

Effectiveness of Allergy Medicines
Percent of people
who reported fewer symptoms
30.7
30.6
30.5
30.4
30.3
30.2
Happajoy
Pollaway


UNEMPLOYMENT RATE
UNDER PRESIDENT OBAMA
10%
9.5%
9%
8.5%
8%
9.0%
8.9%
8.8%
9.0%
9.1%
9.2%
9.1%
9.1%
9.1%
9.0%
8.6%
JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV
2011
SOURCE: BUREAU OF LABOR STATISTICS
FOX NEWS radio
AR WAR AND A TROOP WITHDRAWAL AT THE E
NAS FUT 2,292.50

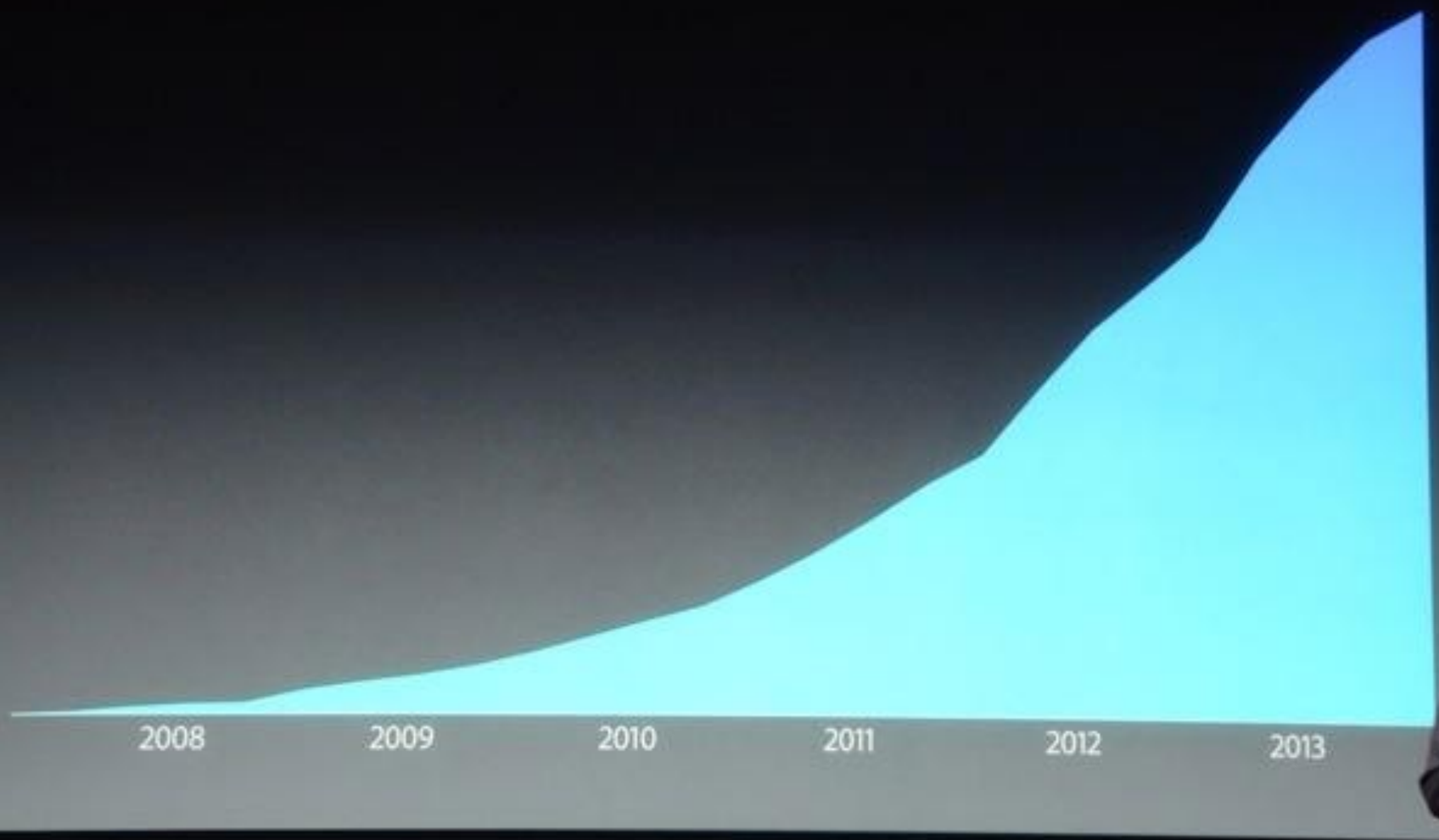

KNUCKLEBALL VELOCITY
R.A. DICKEY
77.3 MPH
75.3 MPH
2012
2013
BlackBerry
Keep Moving
BLUE JAYS

Gun deaths in Florida

Number of murders committed using firearms



Cumulative iPhone sales




Cumulative iPhone sales
400 million iPhones
300
200
100
0
All-time sales
Quarterly sales
2008
'08
2009
'09
2010
'10
2011
'11
2012
'12
2013
'13

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.

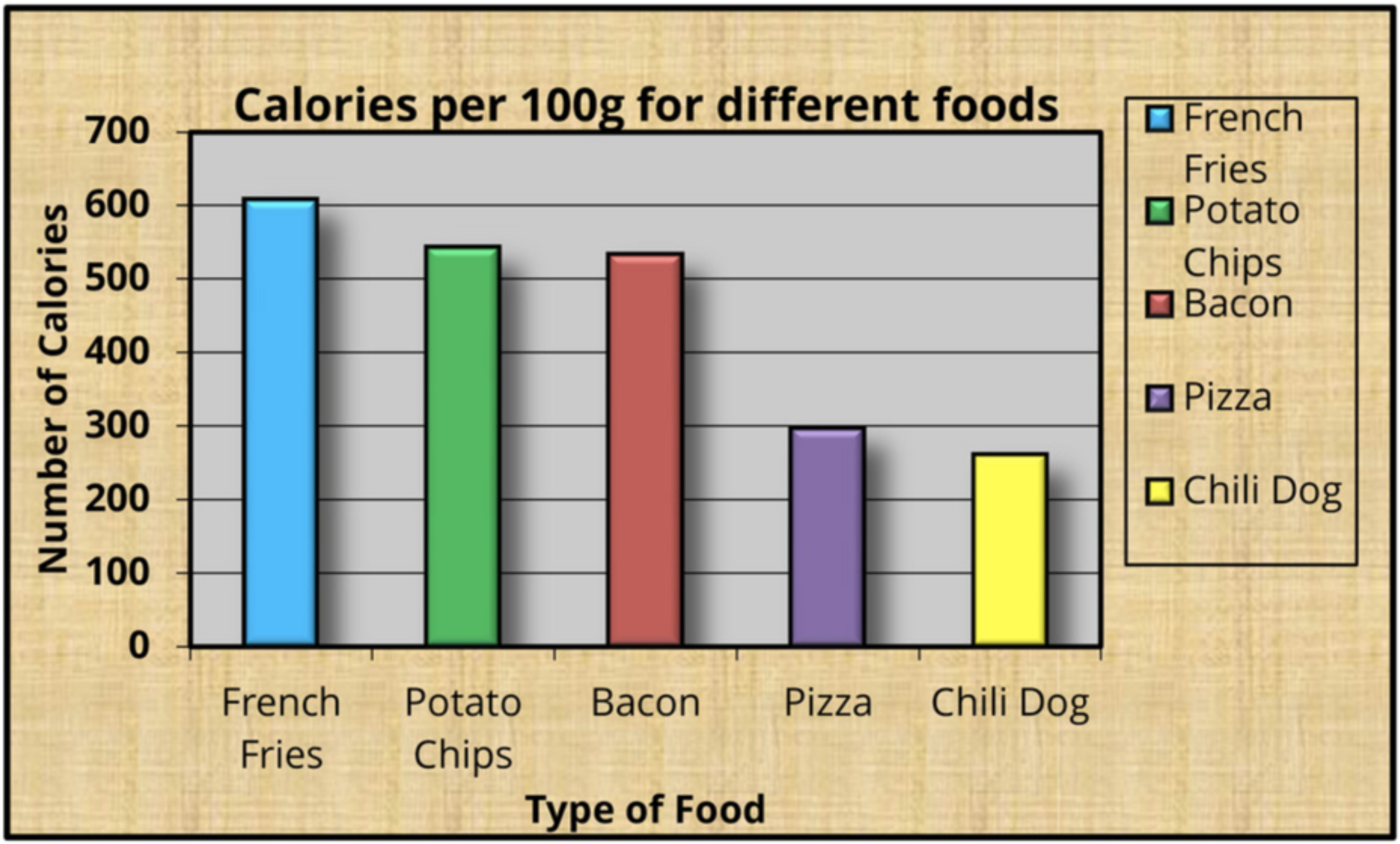
Jones, G. E. (2011). **How to lie with charts**. LaPuerta Books and Media.

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.

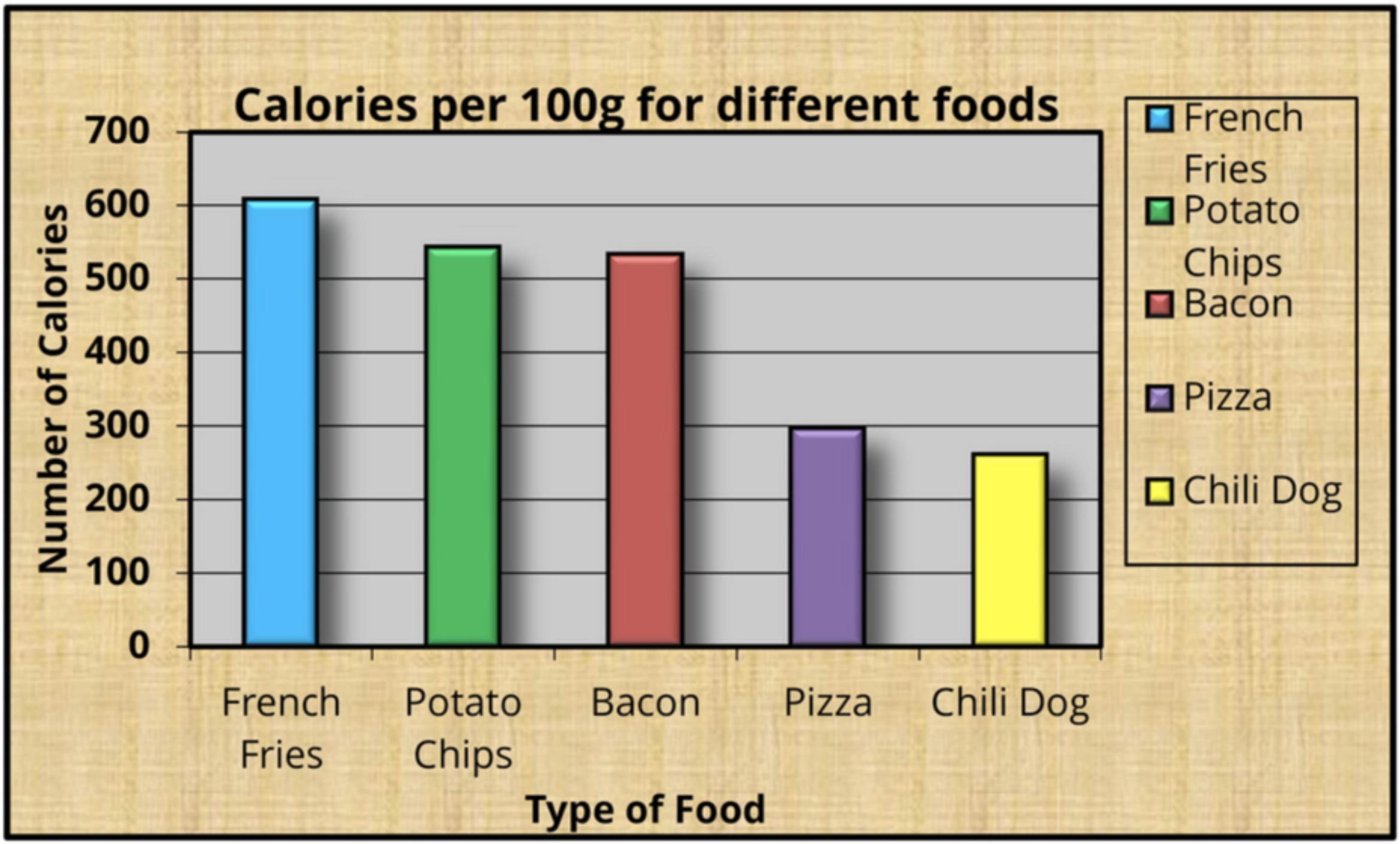
Pandey, A. V., Rall, K., Satterthwaite, M. L., Nov, O., & Bertini, E. (2015, April). **How deceptive are deceptive visualizations?: An empirical analysis of common distortion techniques**. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1469-1478). ACM.

CLARITY

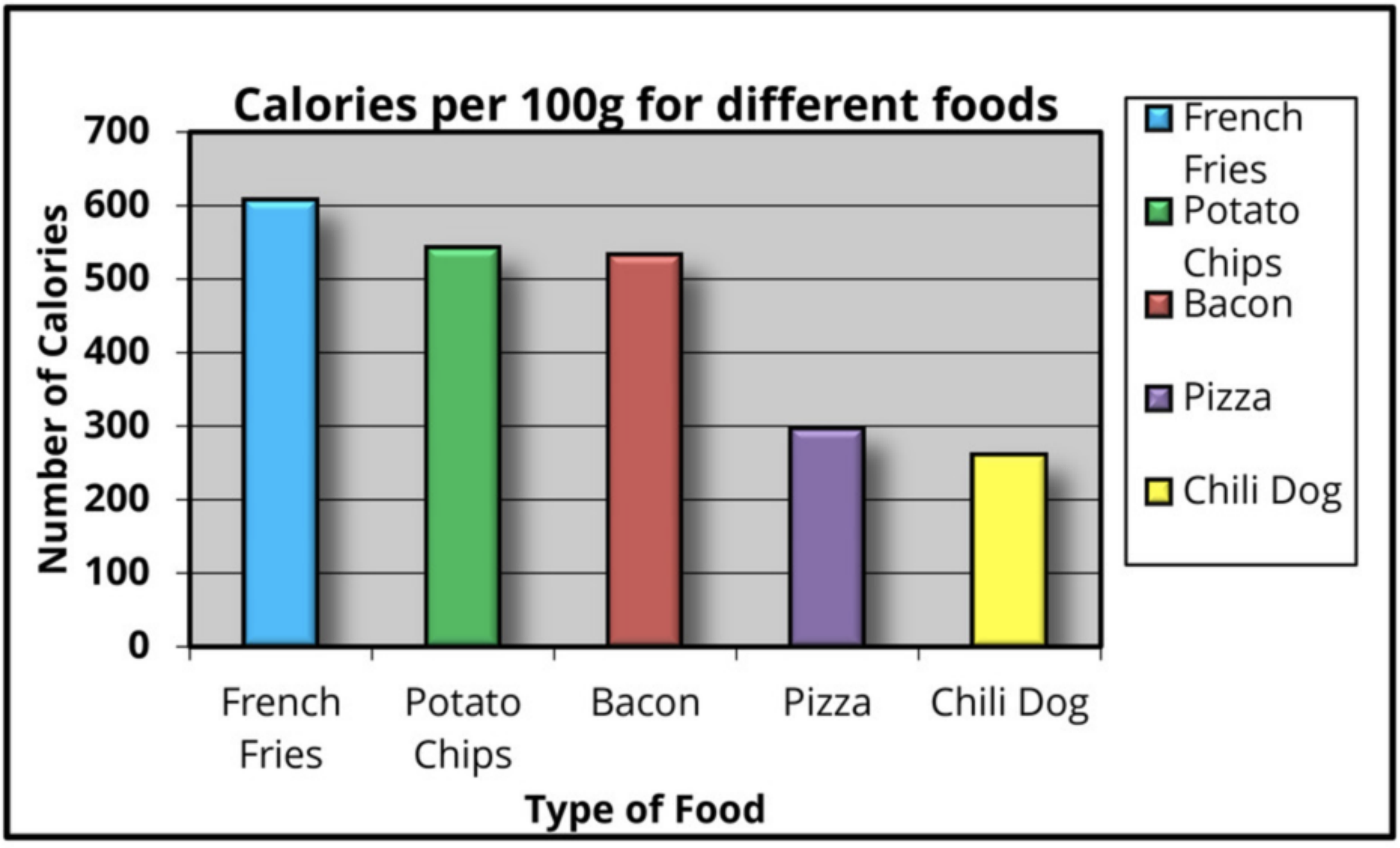

Calories per 100g for different foods
Number of Calories
700
600
500
400
300
200
100
0
French Fries
Potato Chips
Bacon
Pizza
Chili Dog
Type of Food
French Fries
Potato Chips
Bacon
Pizza
Chili Dog

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

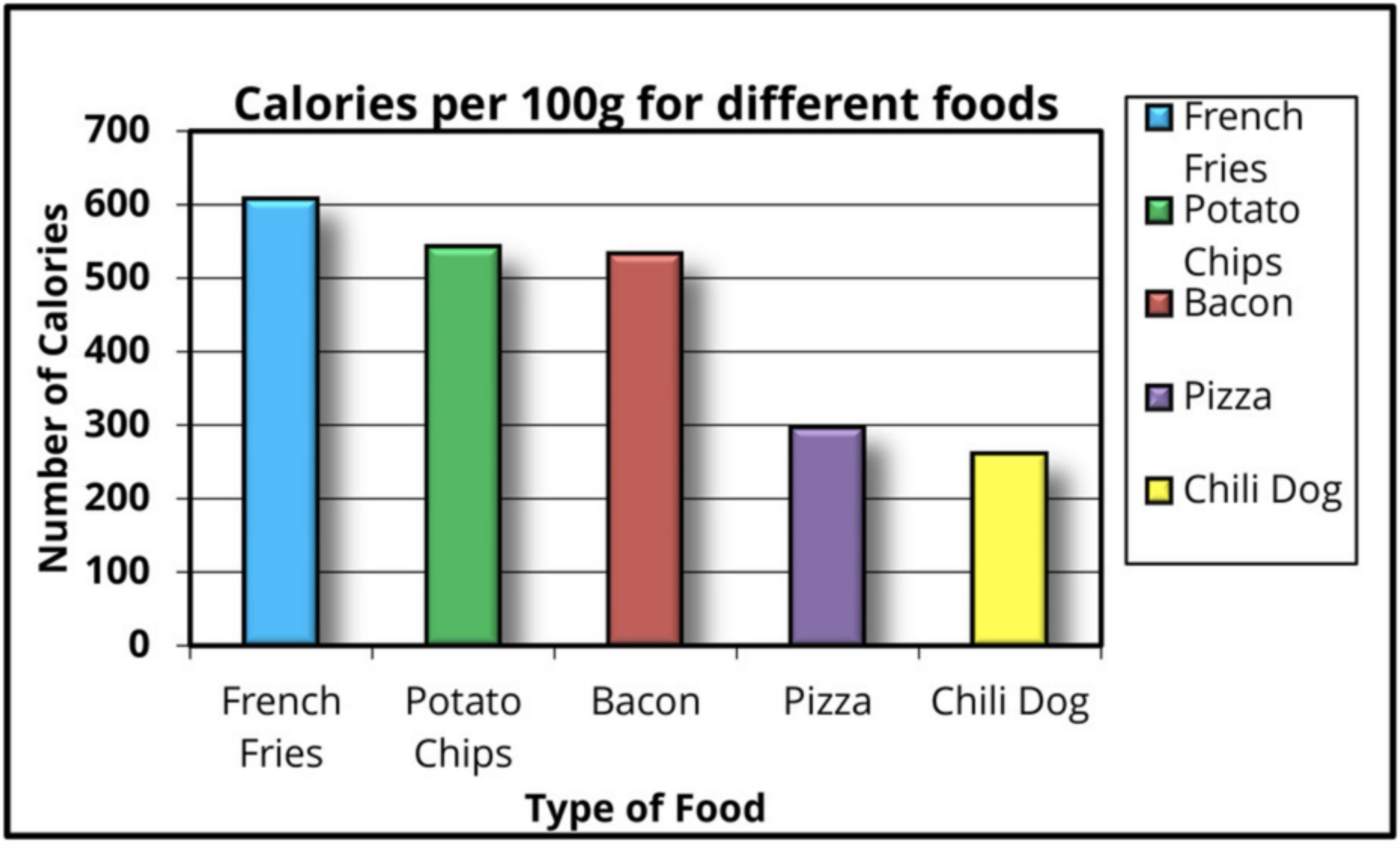
Remove backgrounds



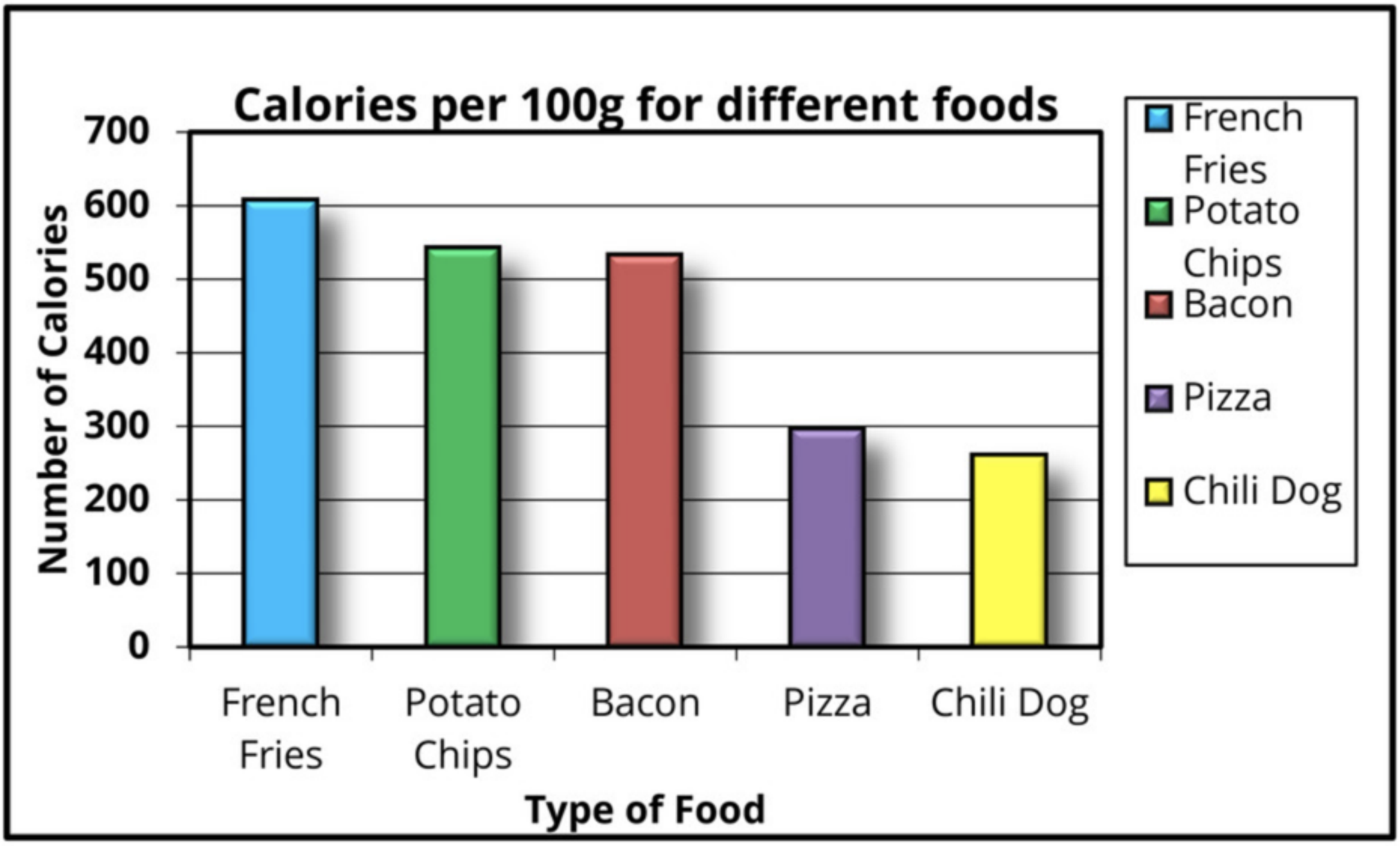
Remove backgrounds



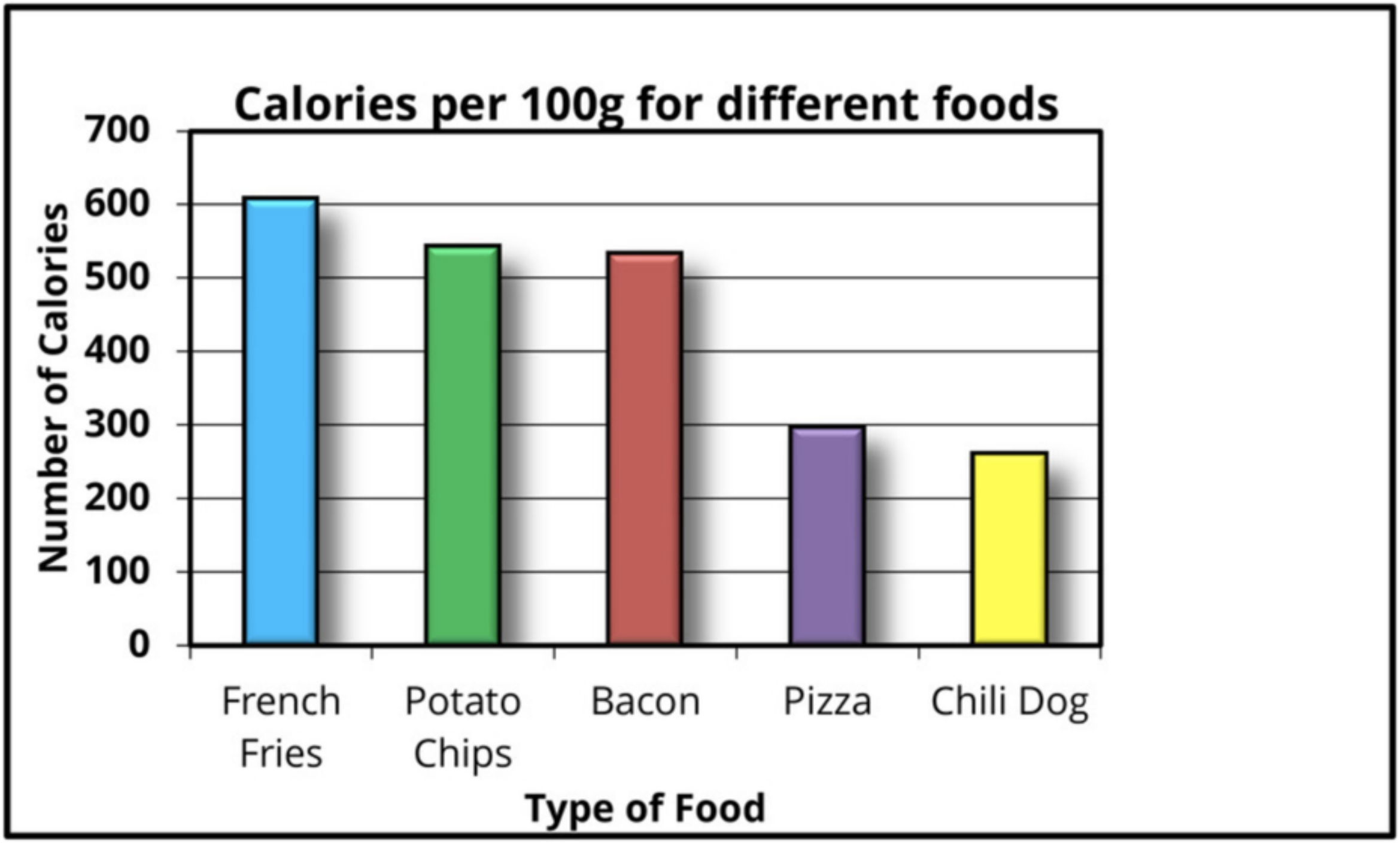
Remove backgrounds



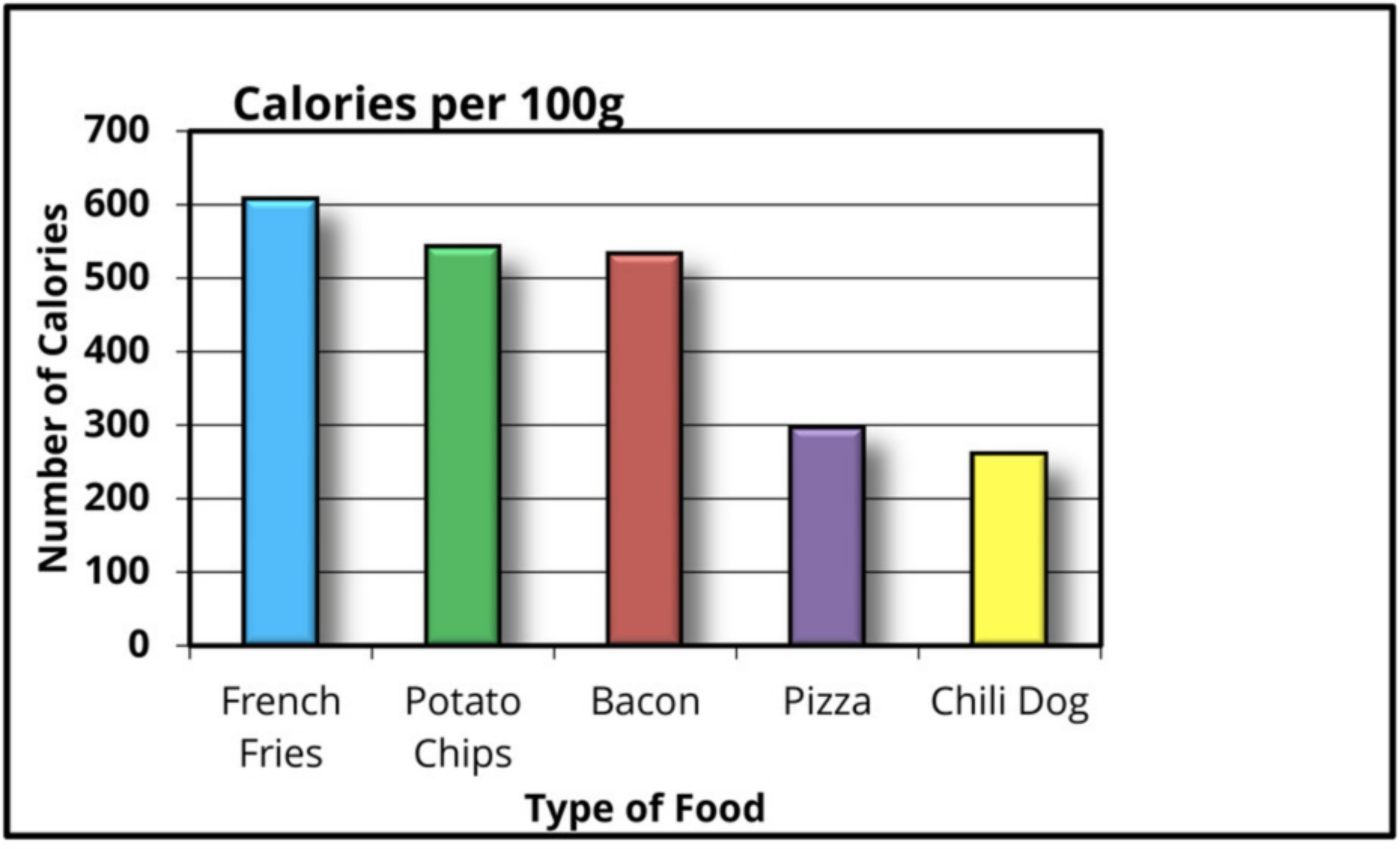
Remove redundant labels



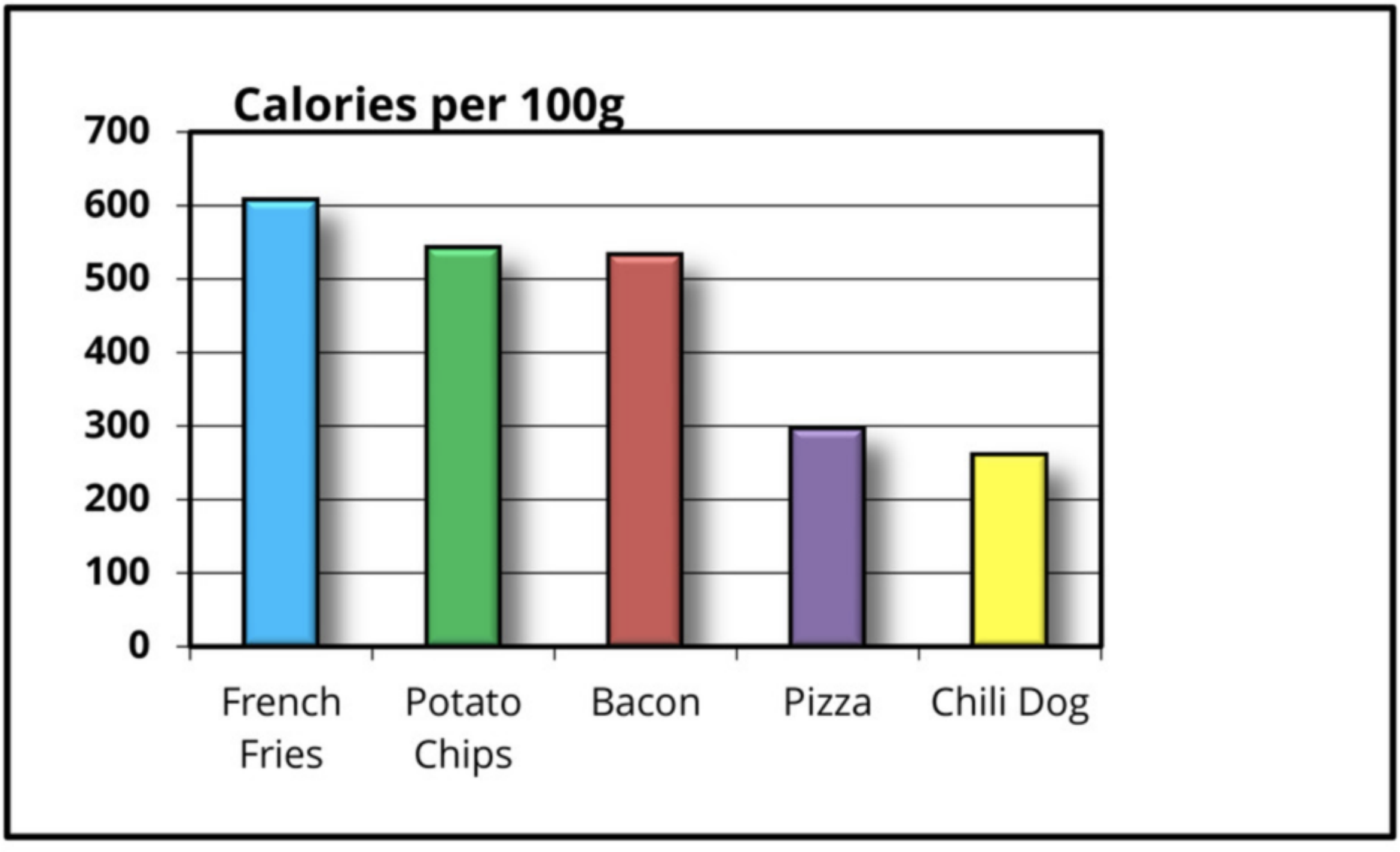
Remove redundant labels



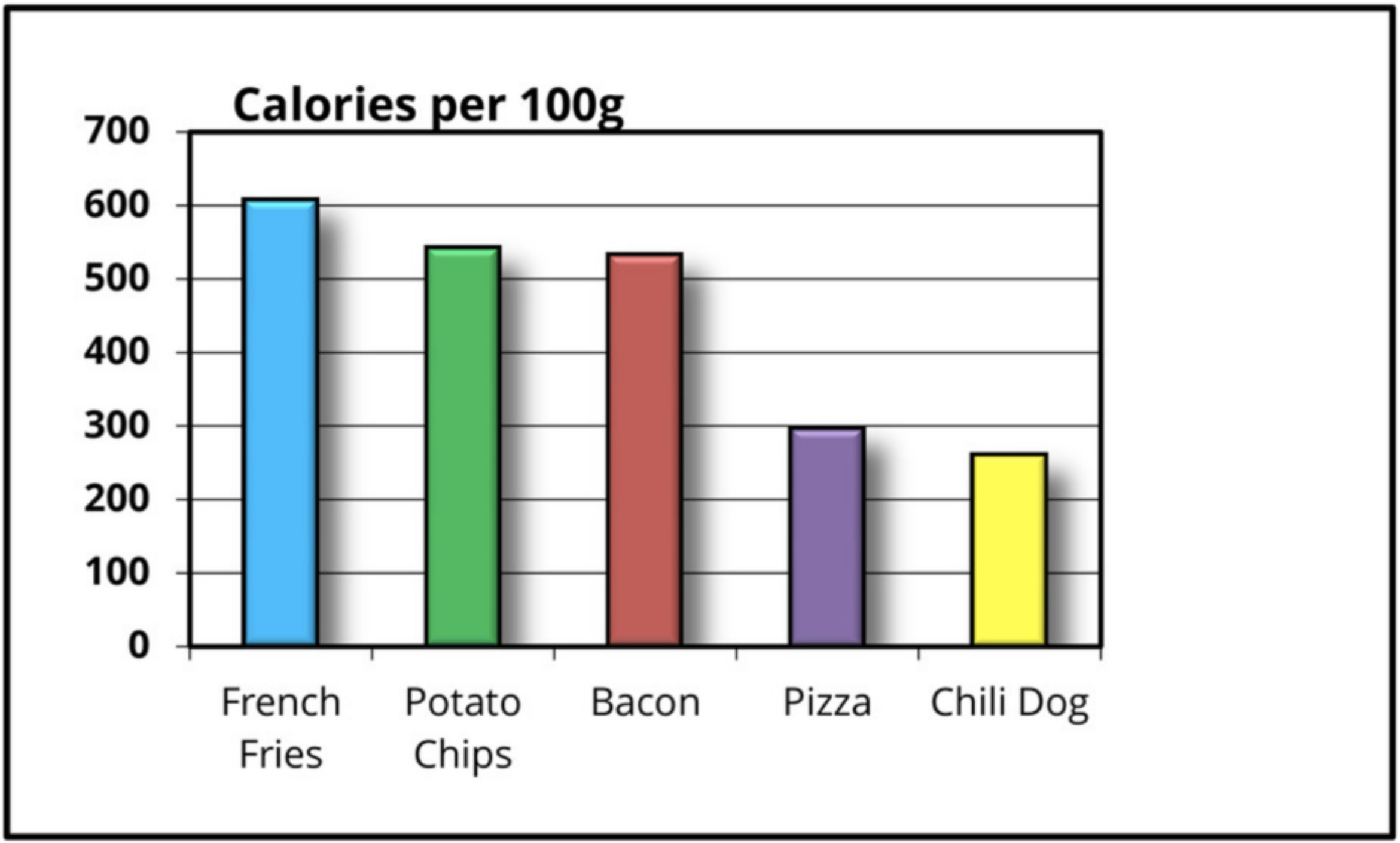
Remove redundant labels



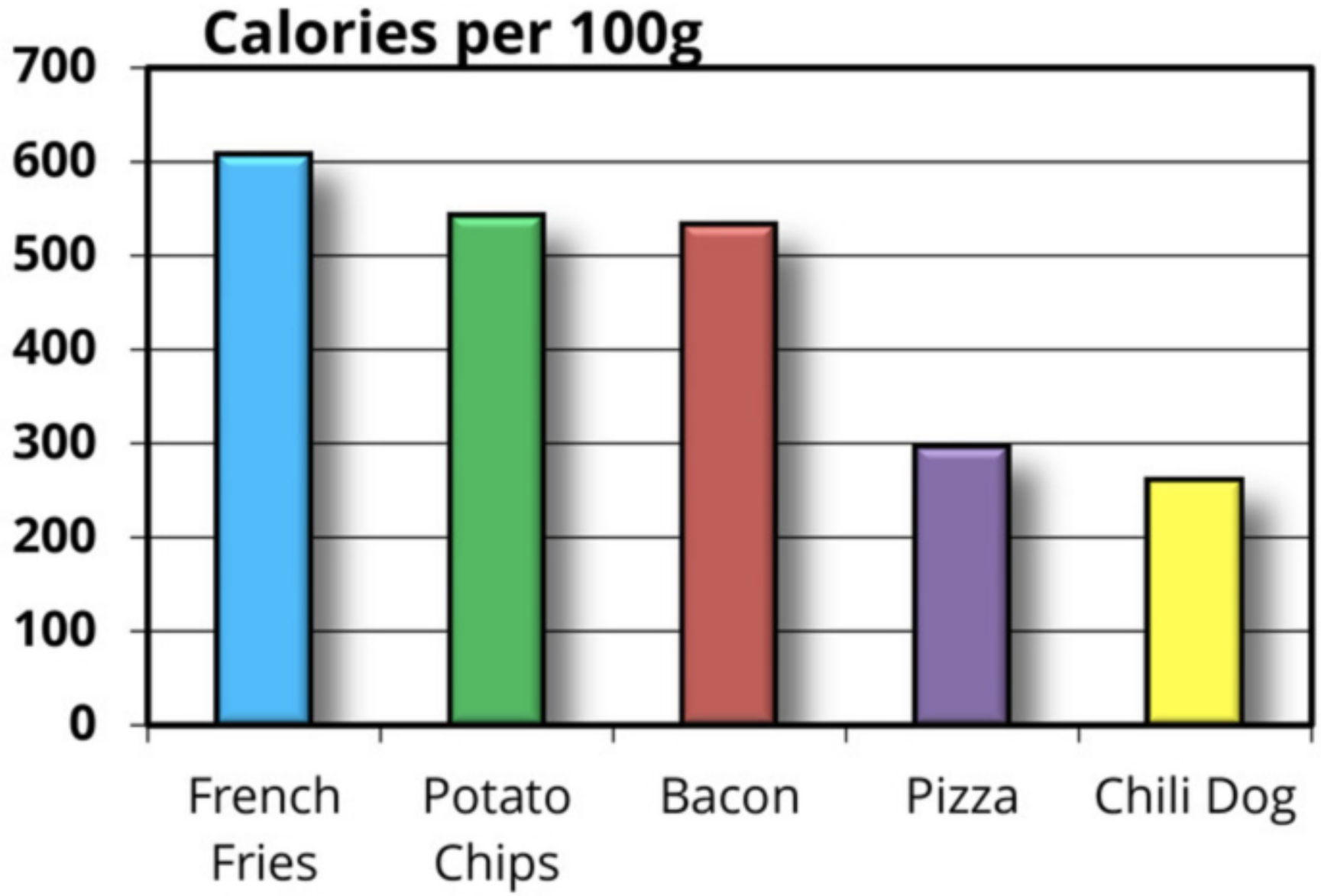
Remove redundant labels



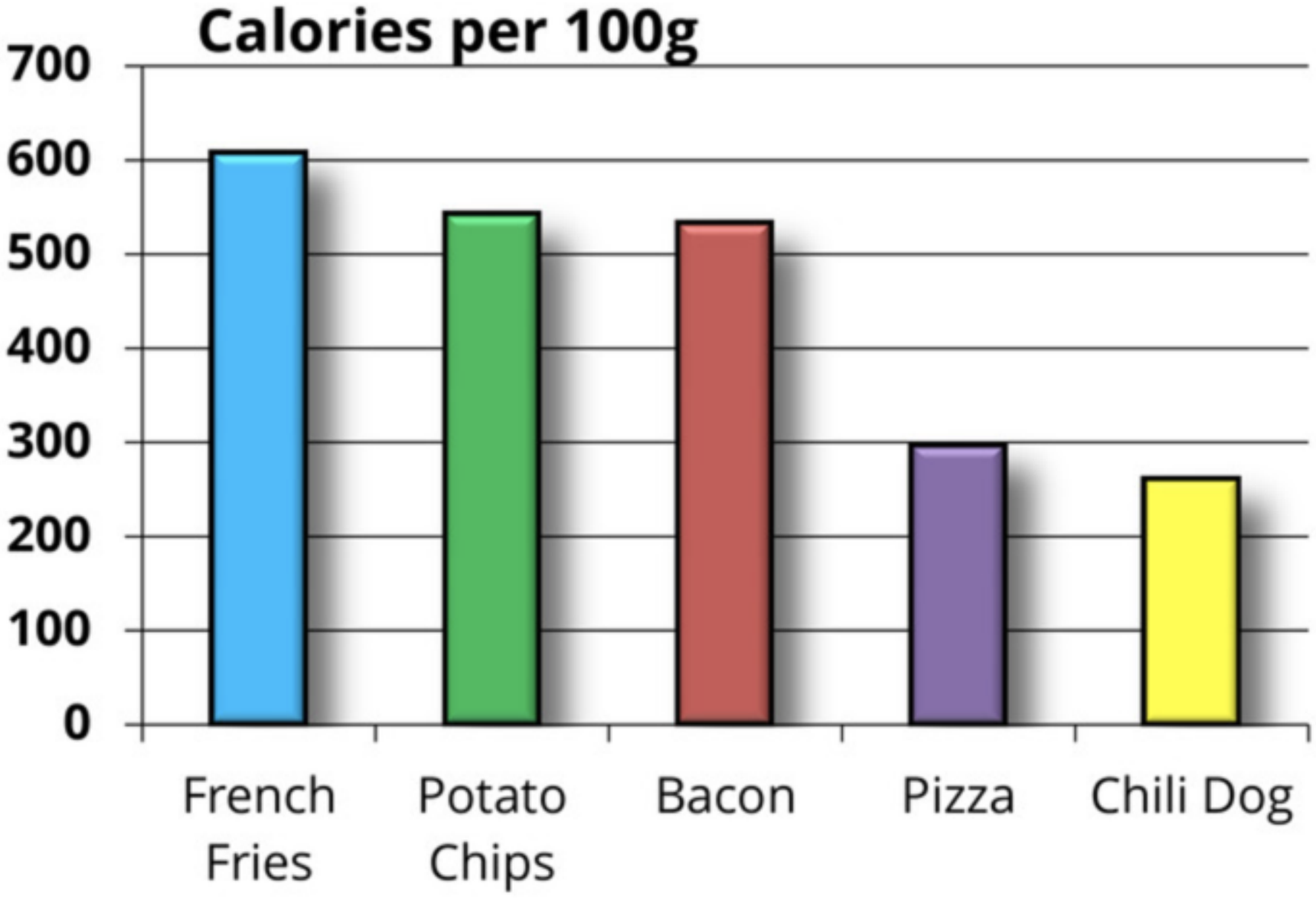
Remove borders



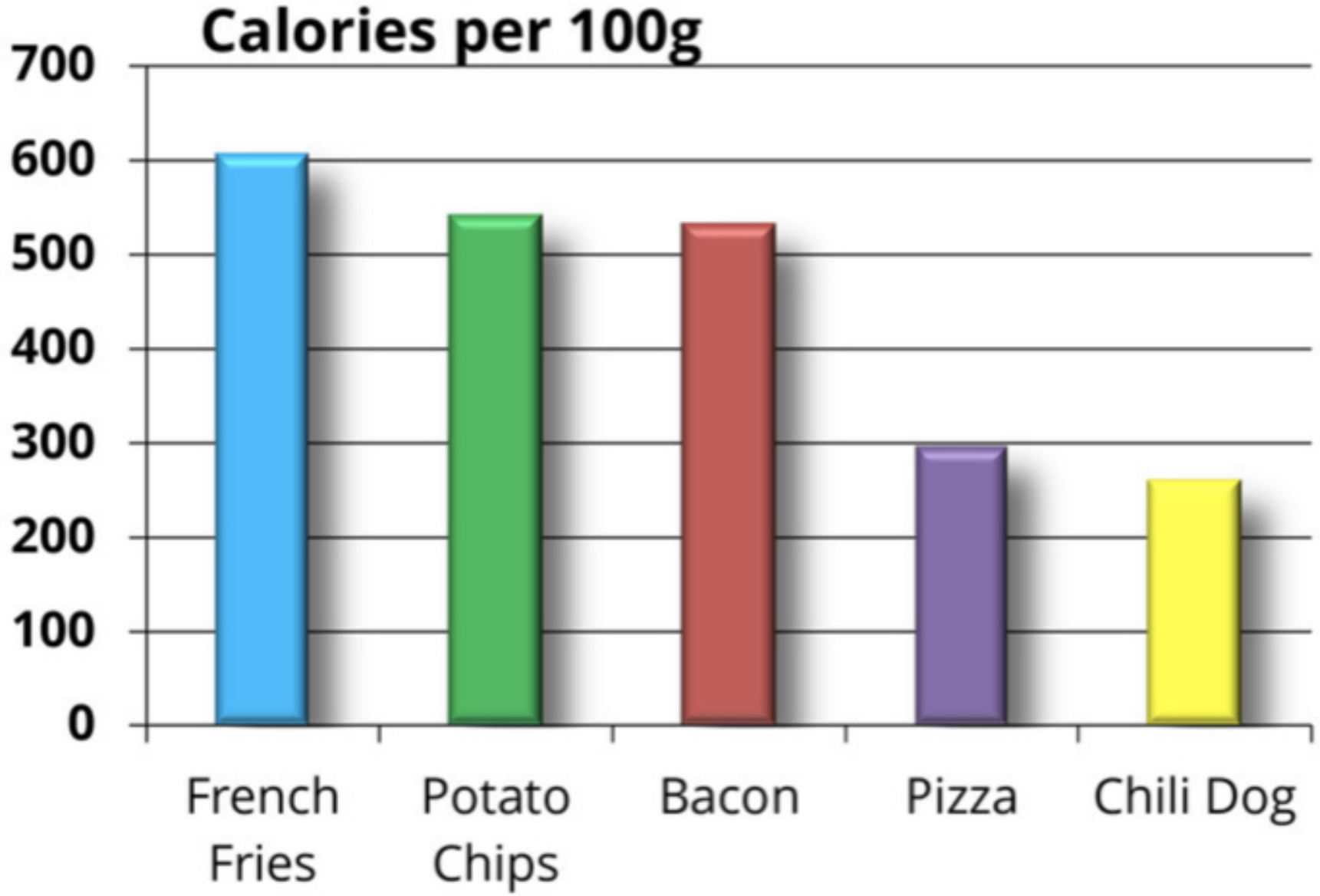
Remove borders



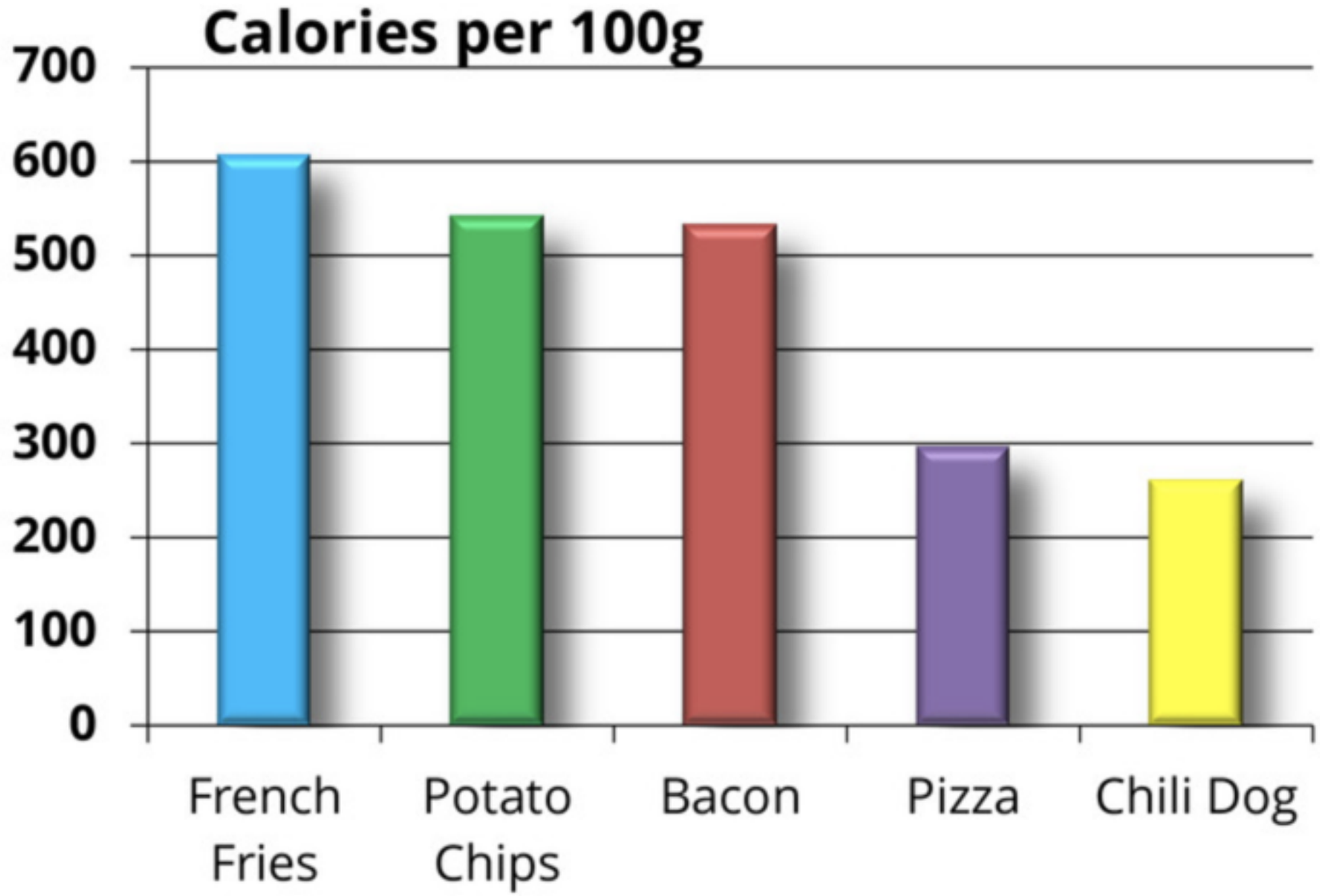
Remove borders



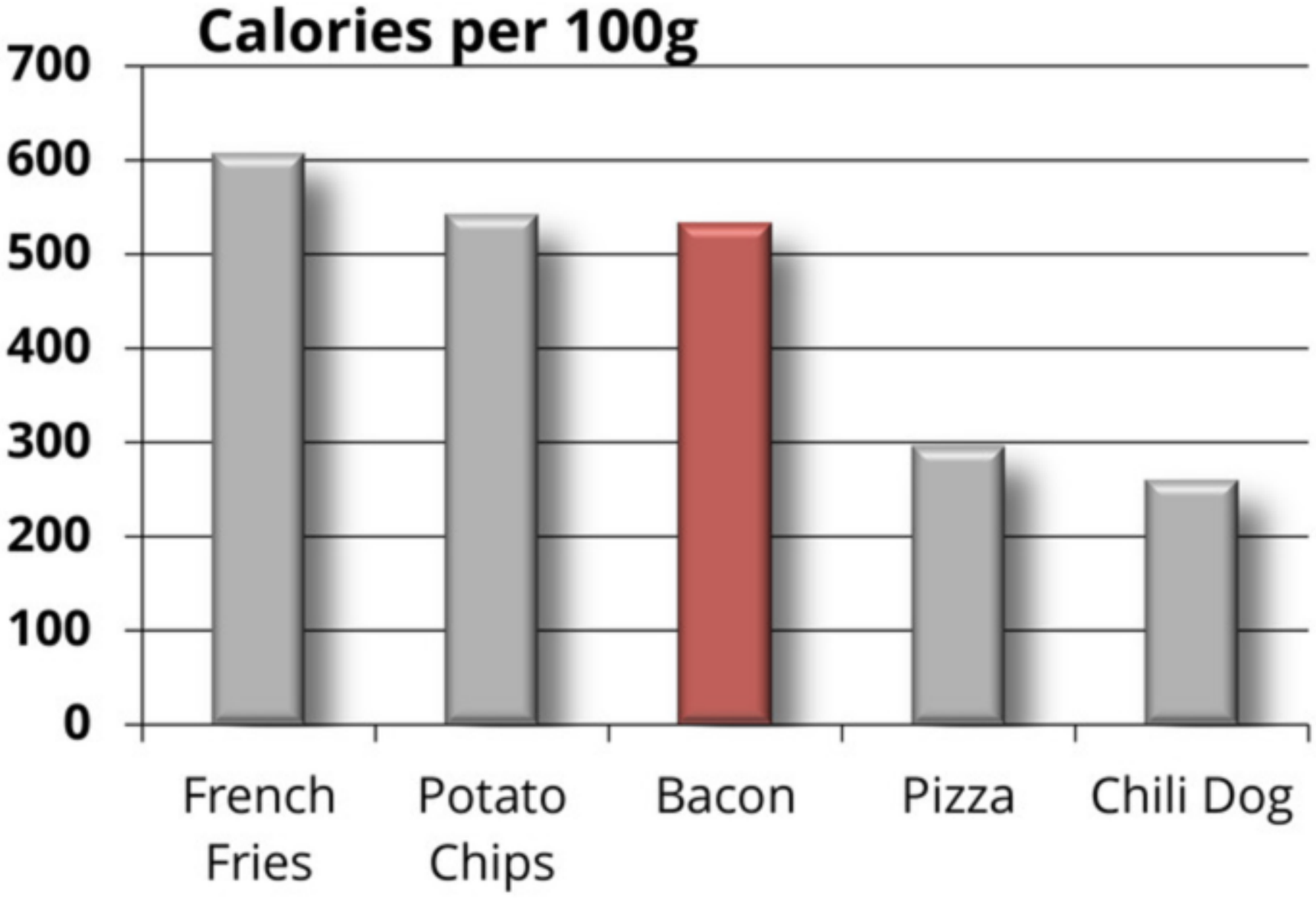
Remove borders



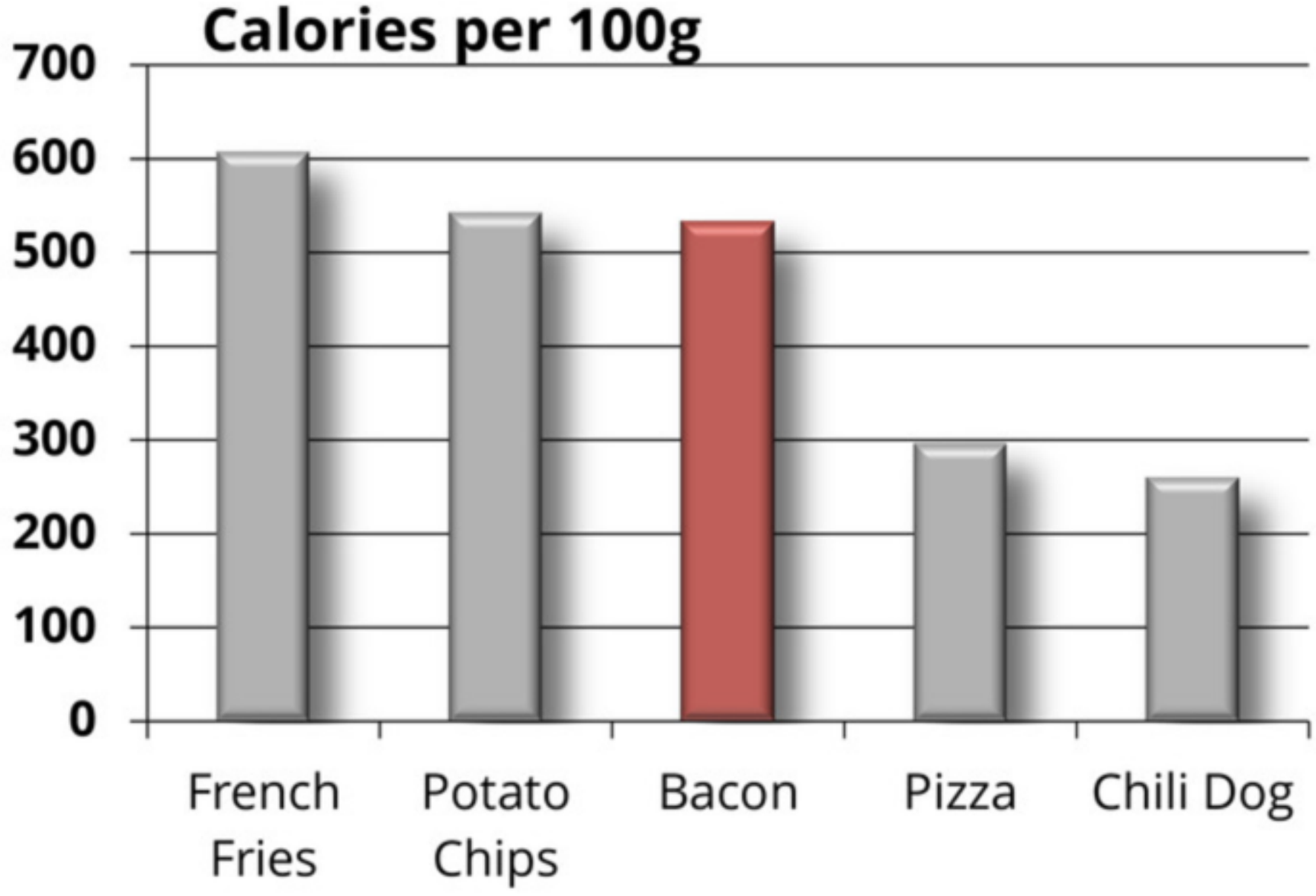
Reduce colors



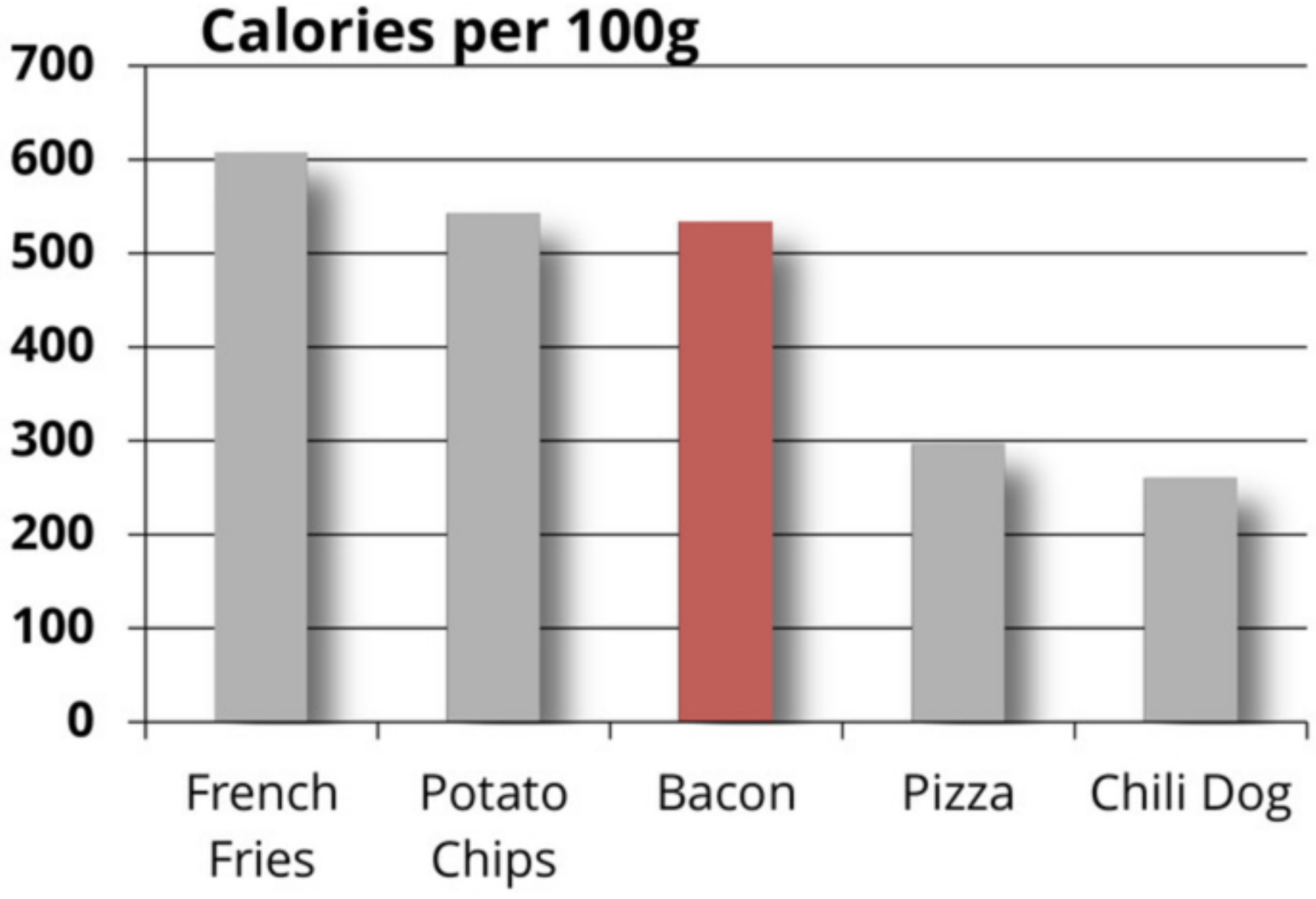
Reduce colors



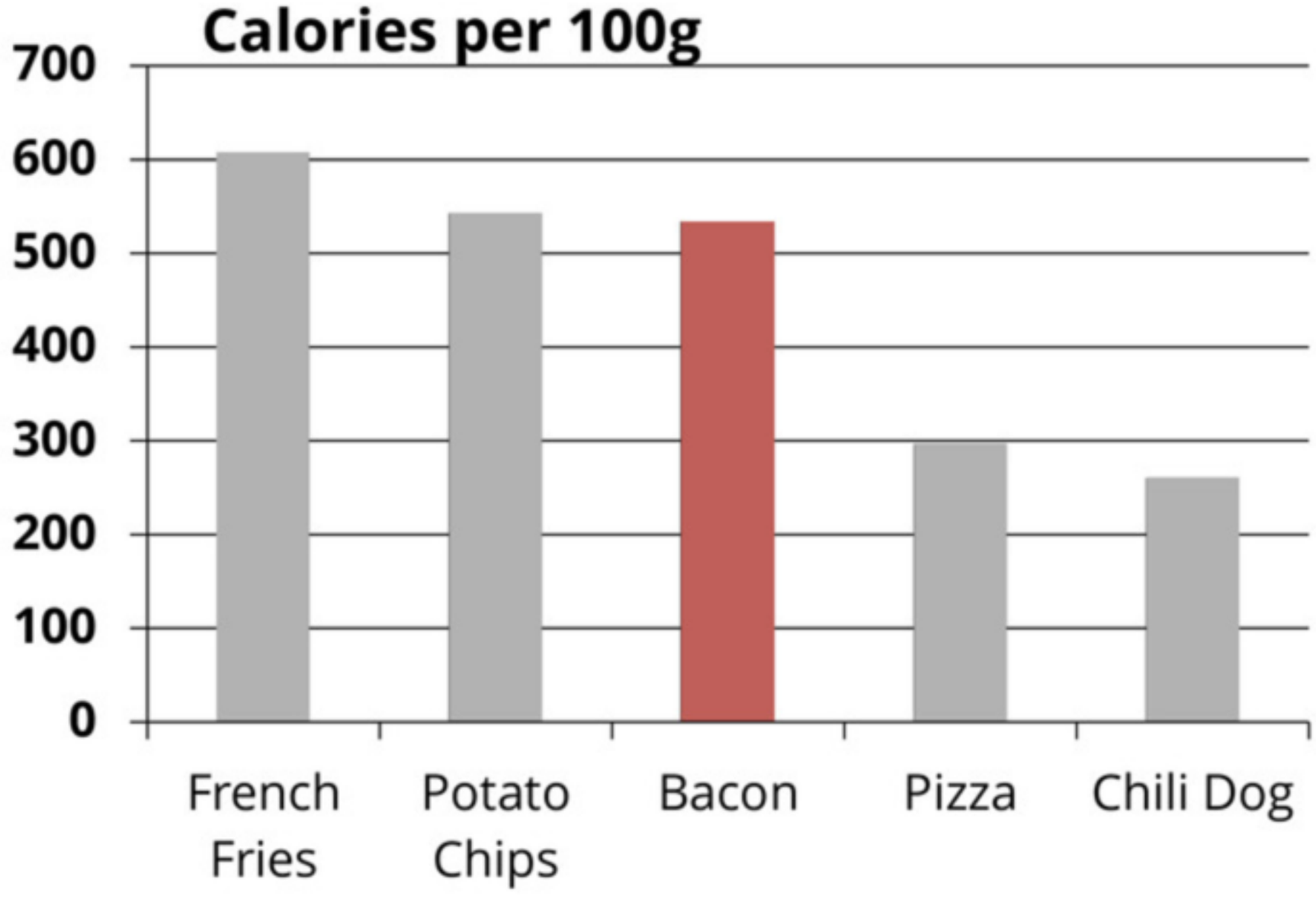
Remove special effects



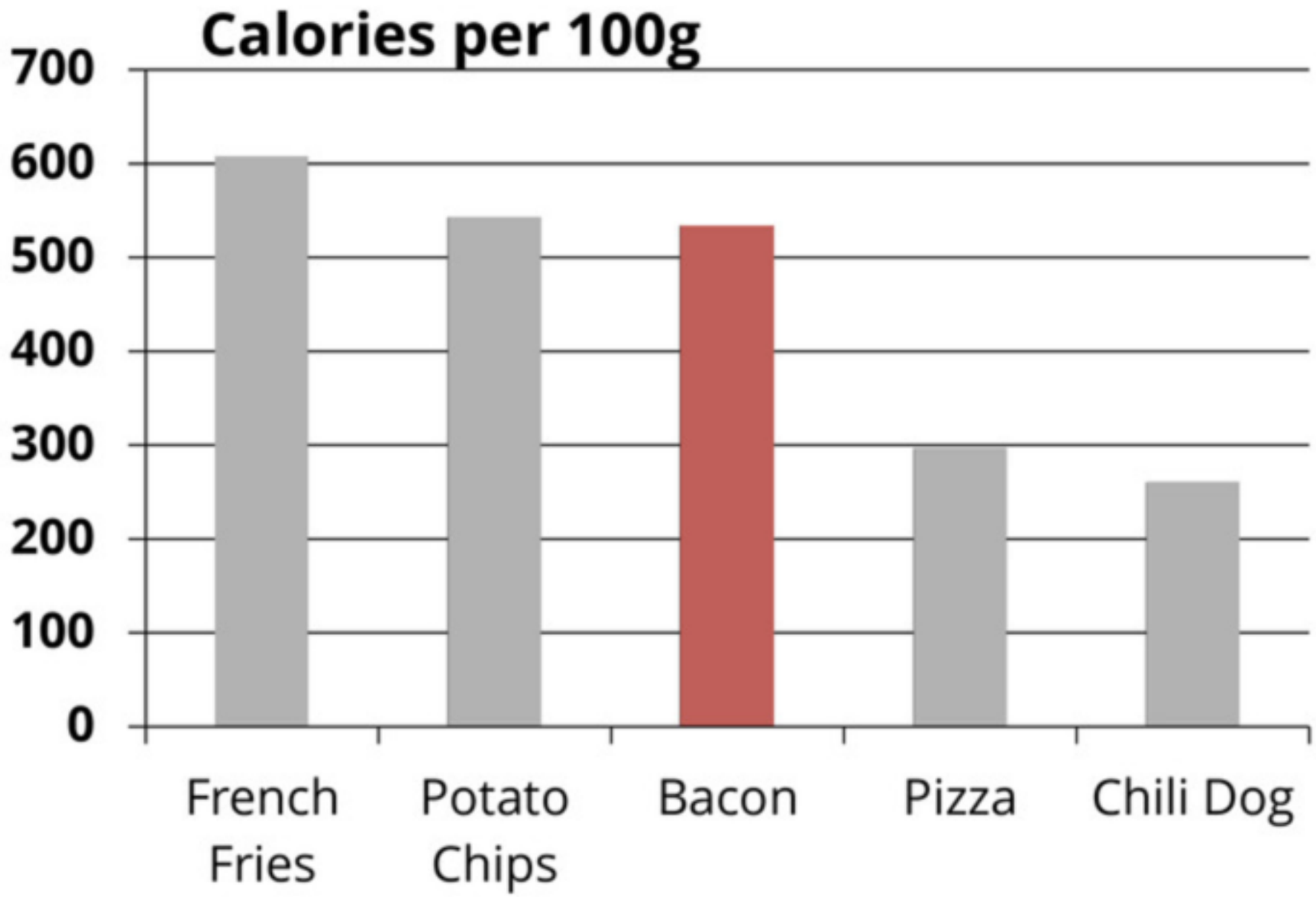
Remove special effects



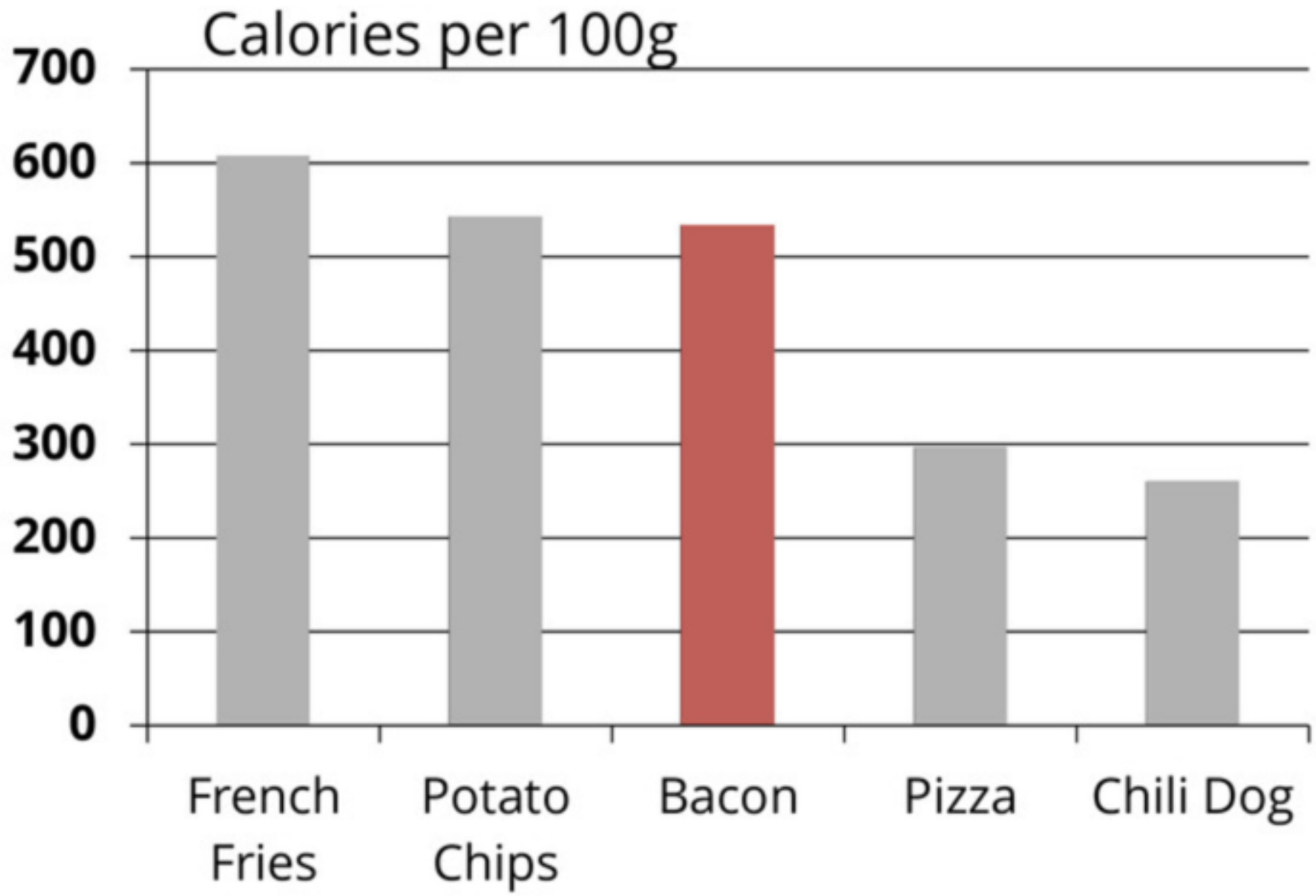
Remove special effects



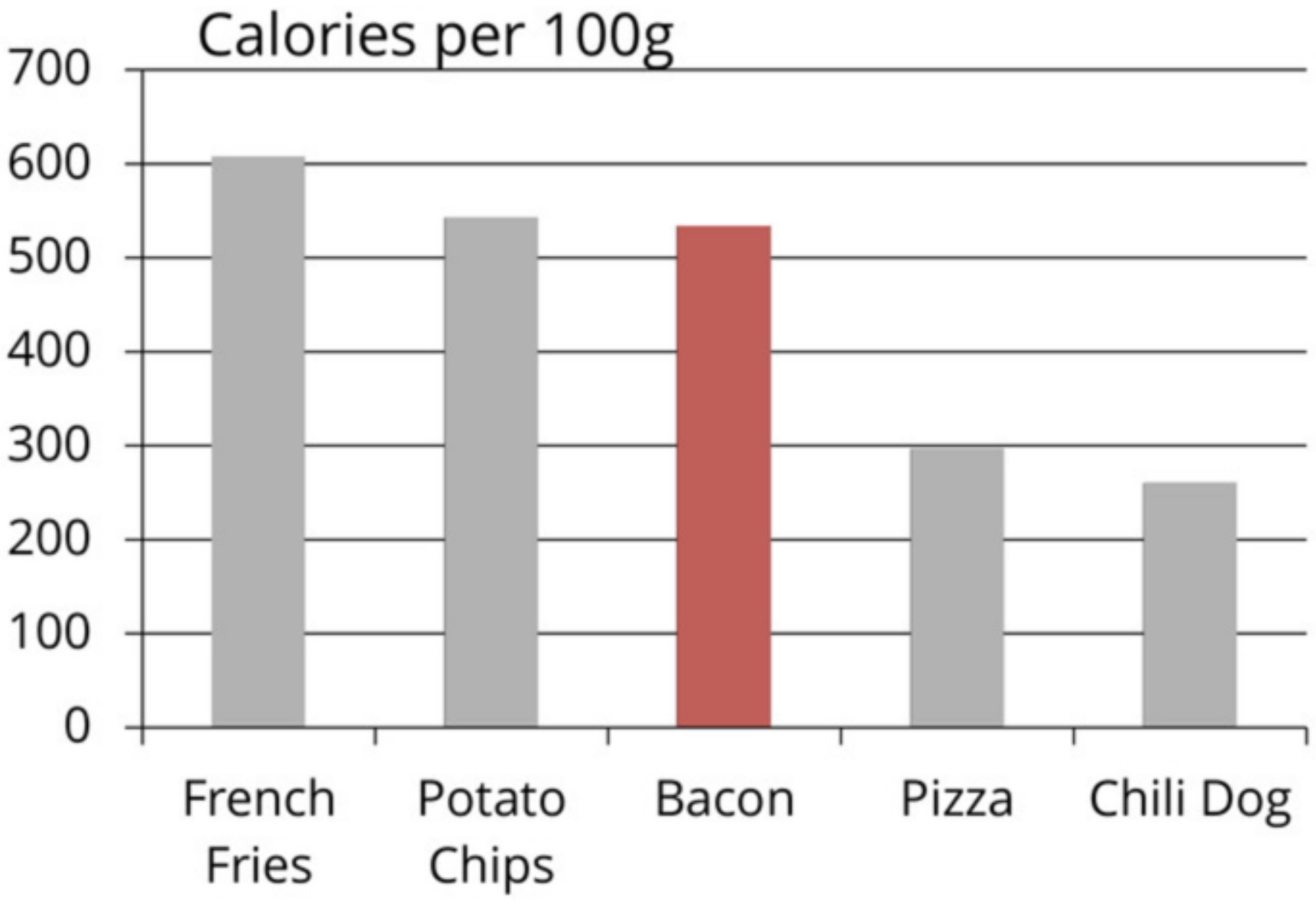
Remove bolding



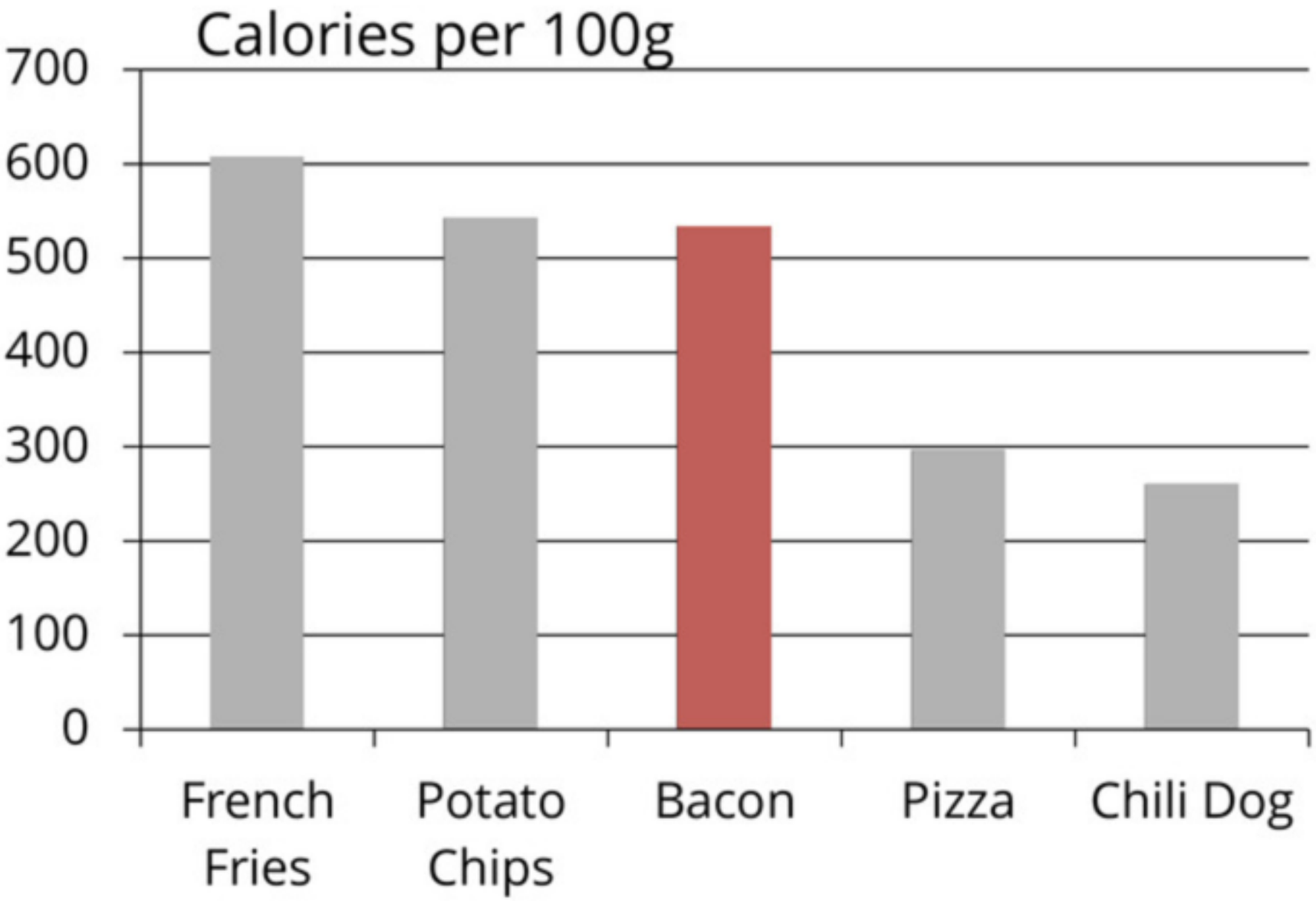
Remove bolding



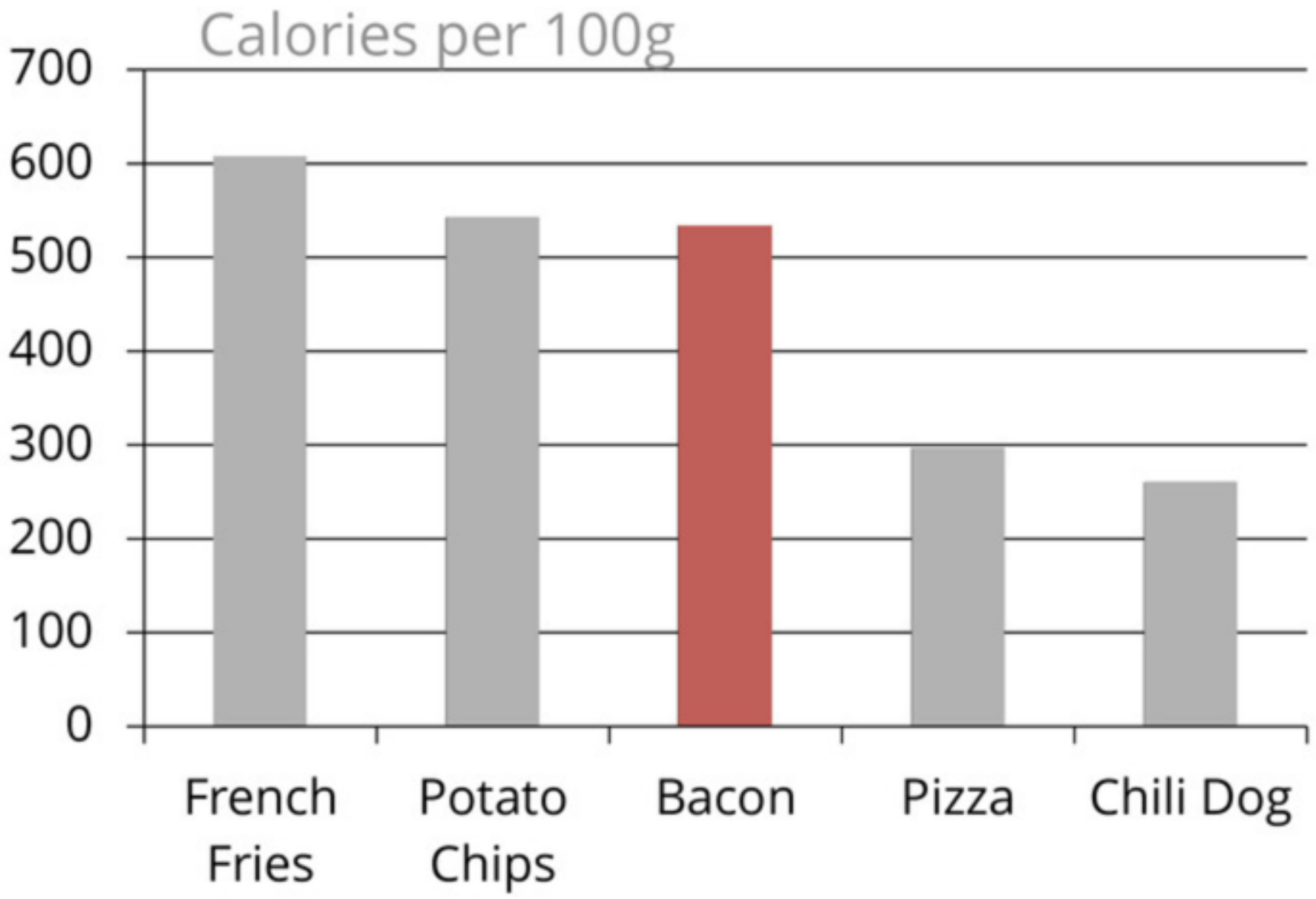
Remove bolding



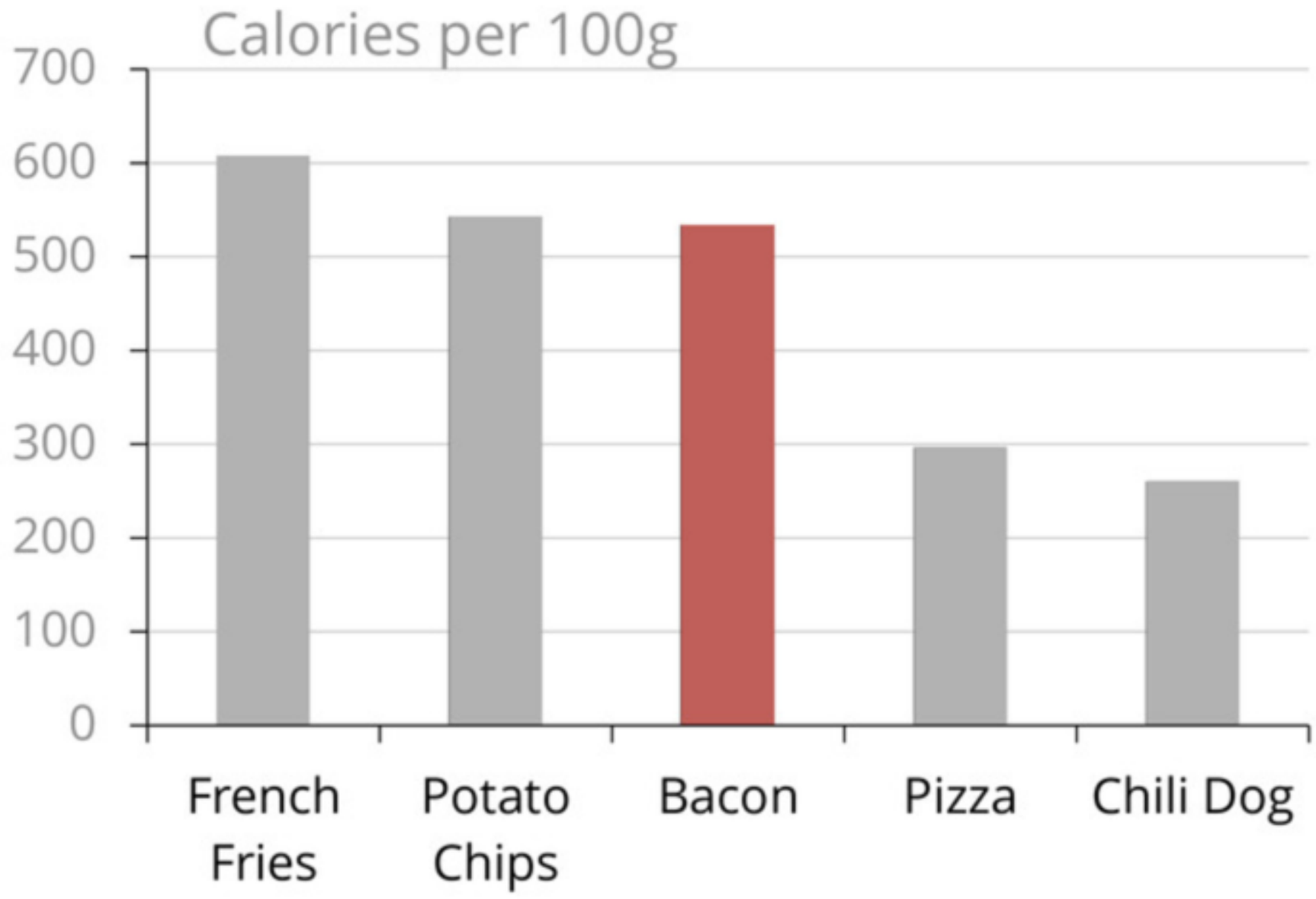
Lighten labels



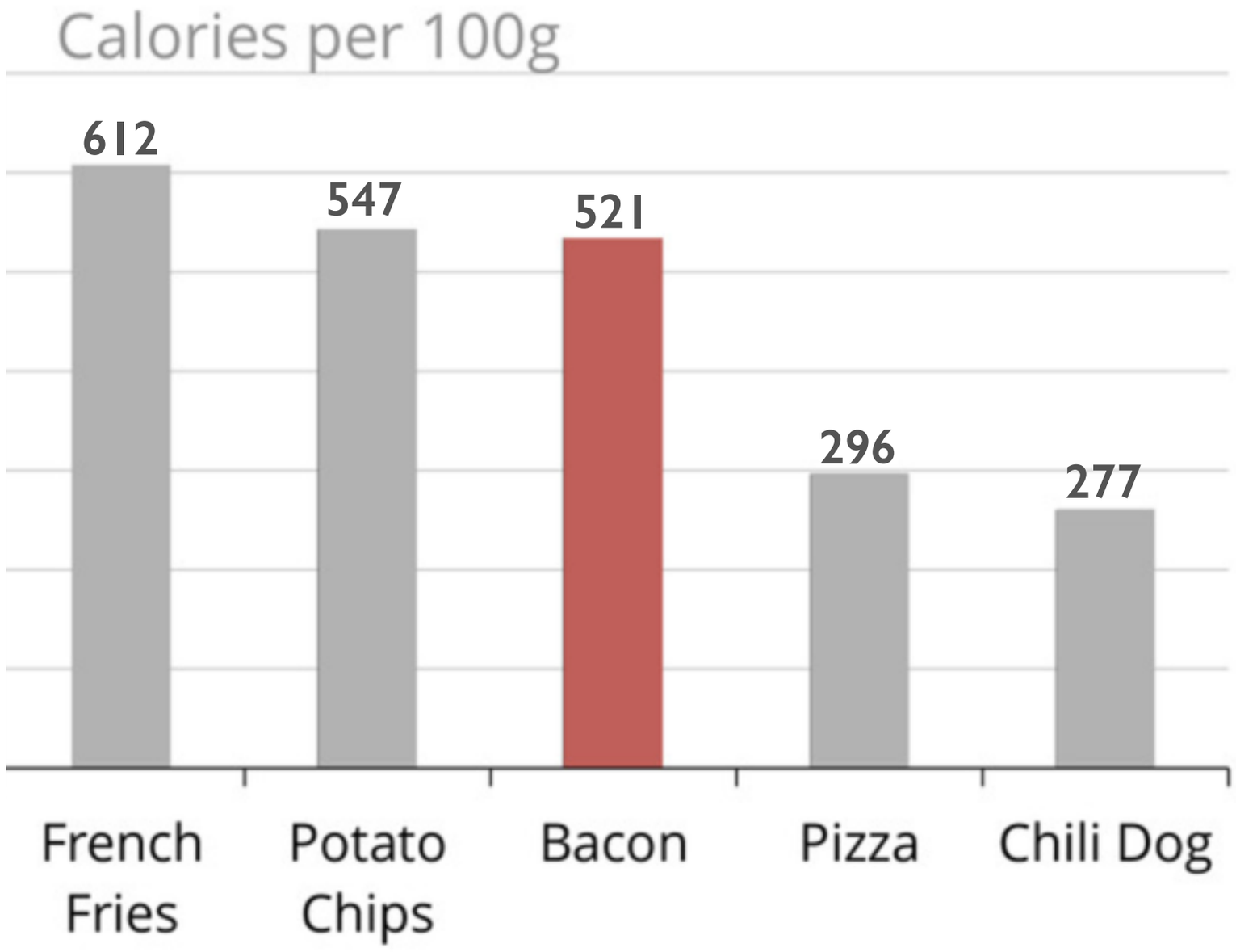
Lighten labels



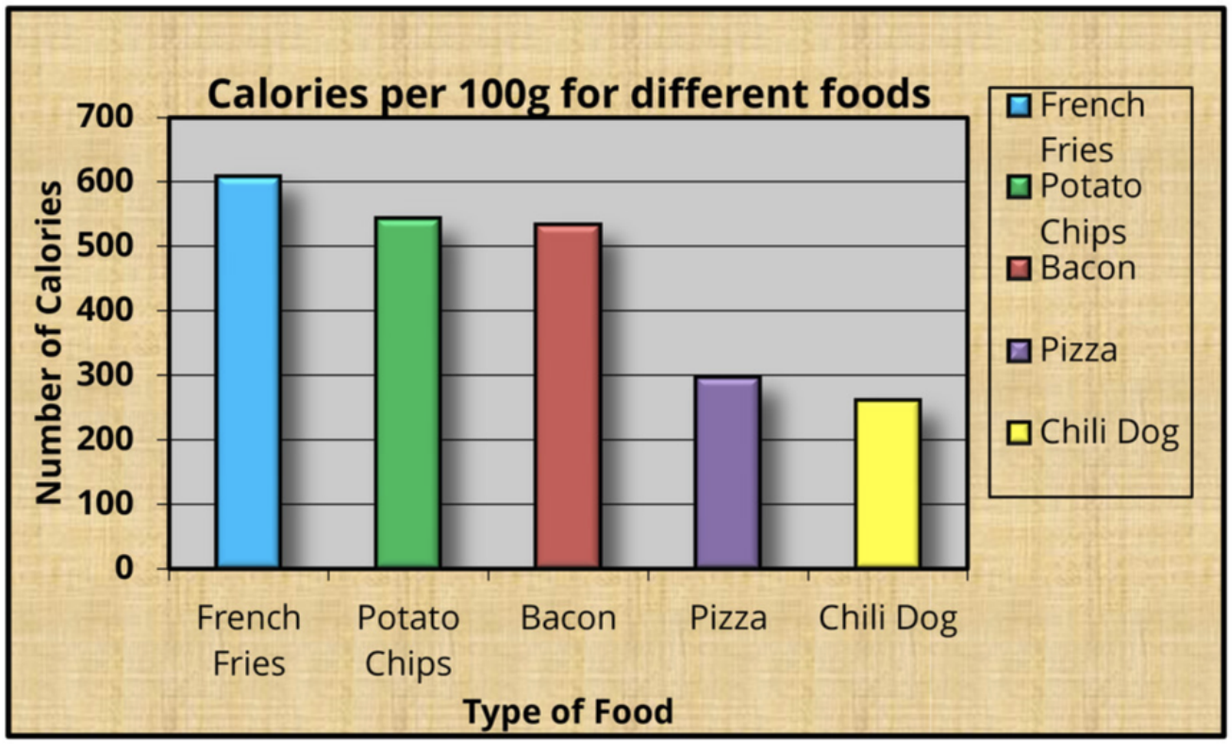
Lighten lines

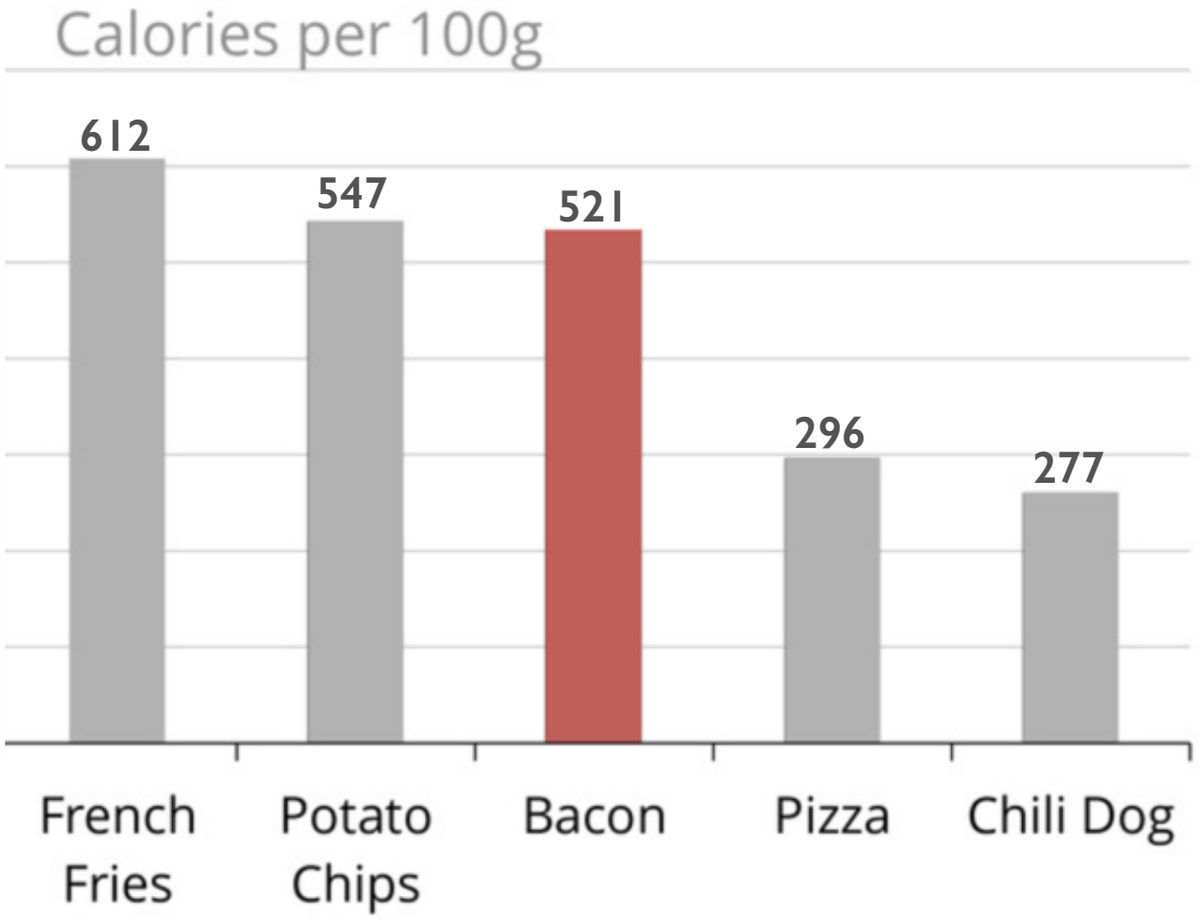


Direct label

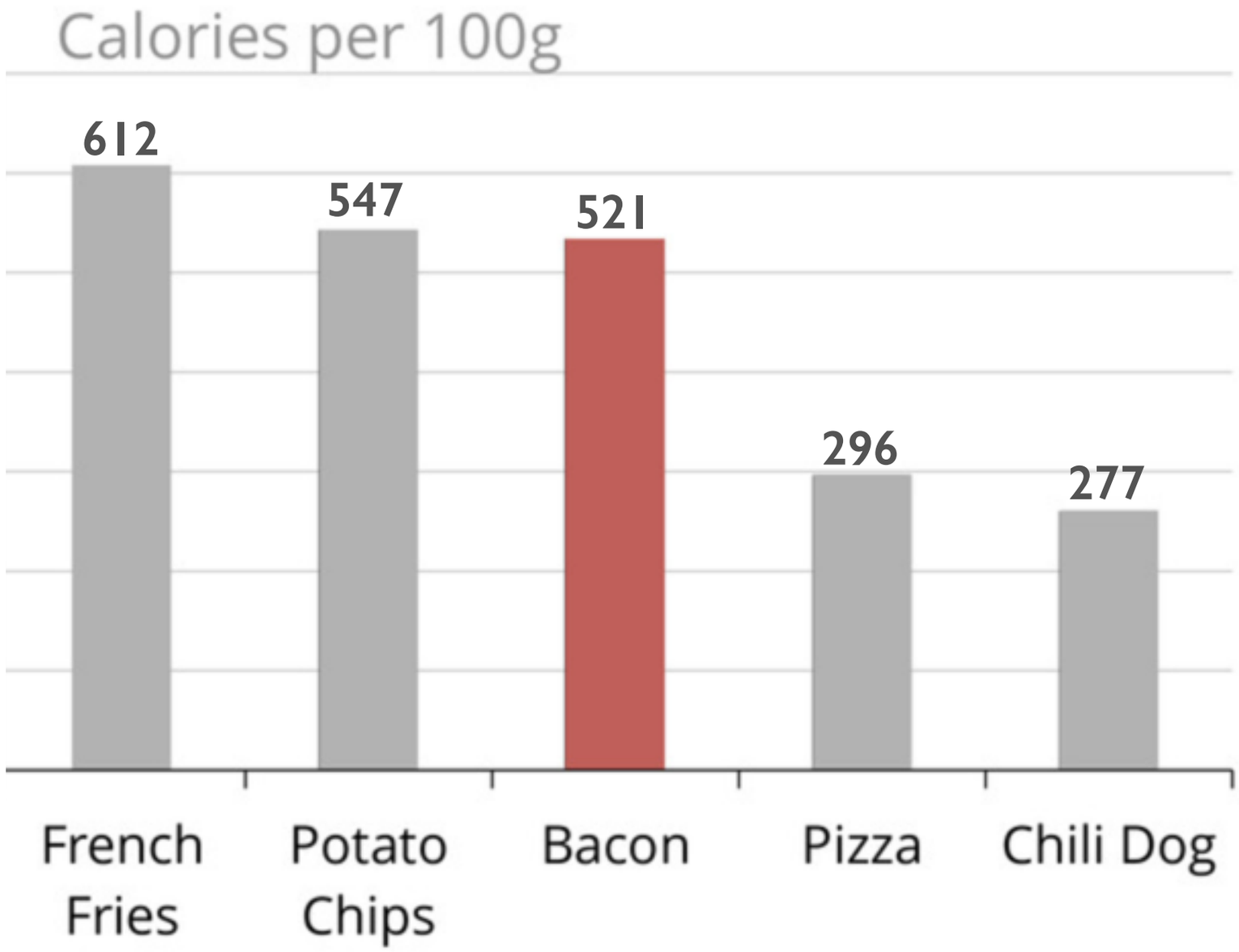


DATA-INK RATIO



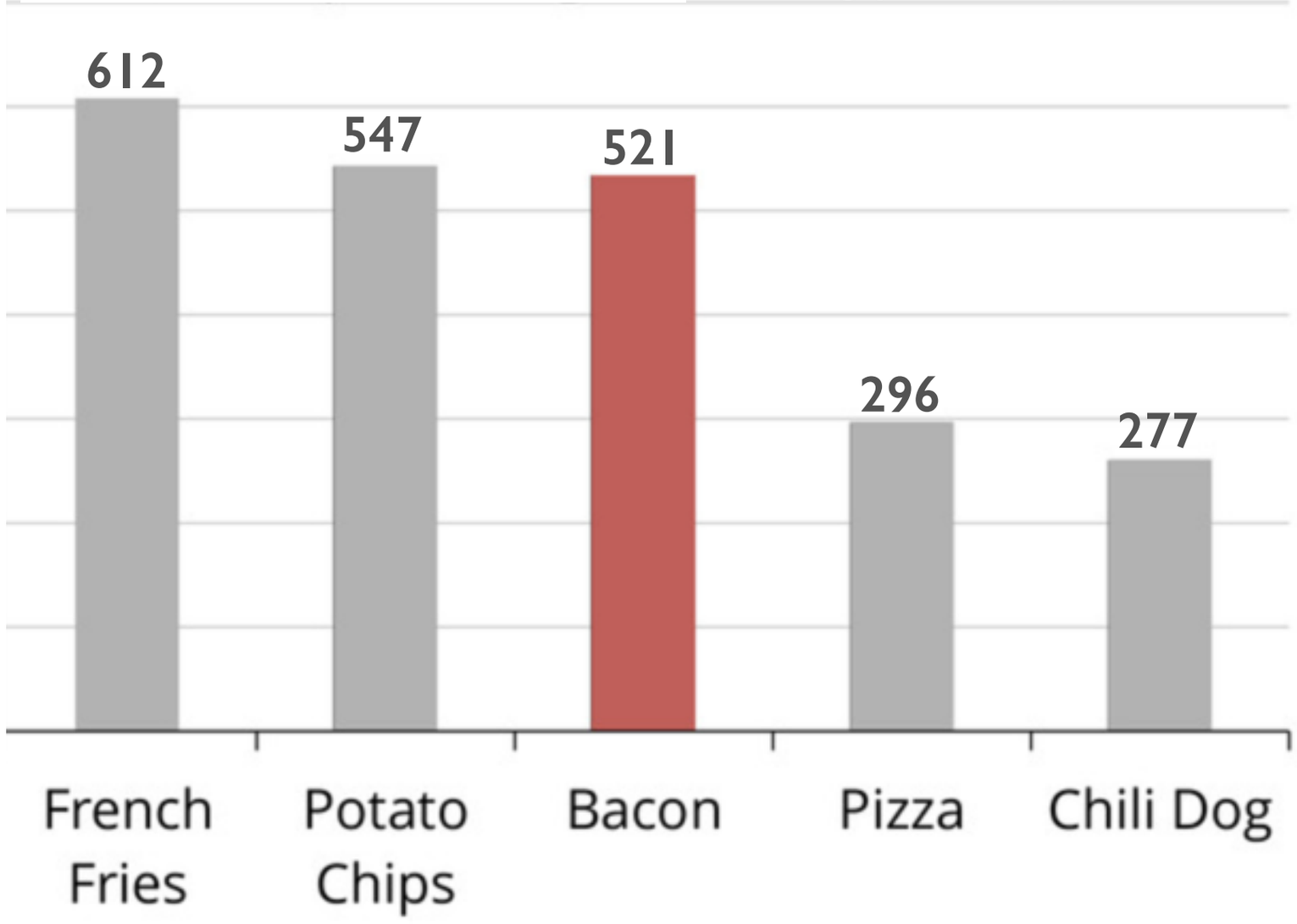


MEANINGFUL TITLE



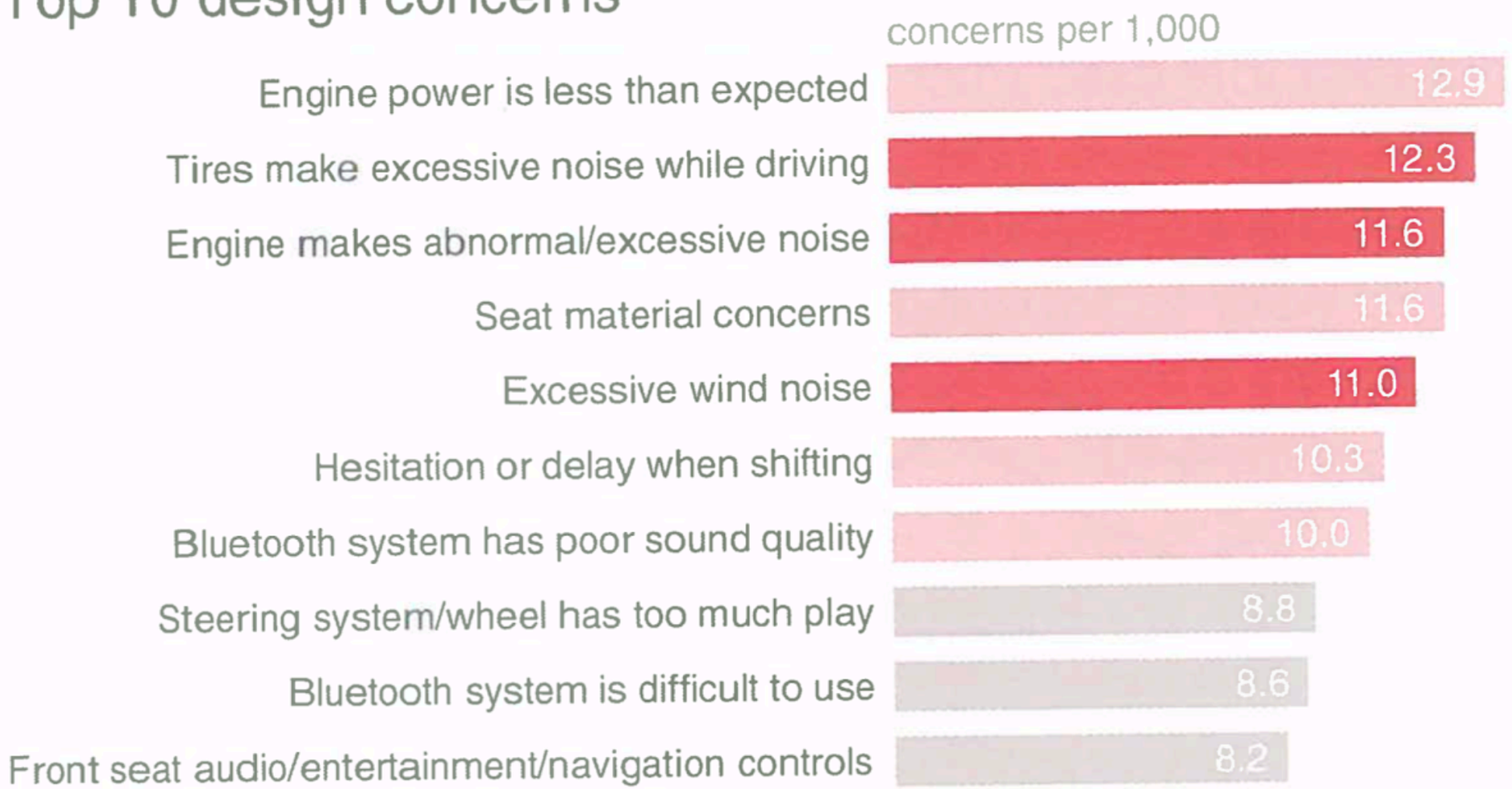
MEANINGFUL TITLE

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



Comments indicate that **noisy tire issues** are most apparent **in the rain**.

Complaints about **engine noise** commonly cited **after the car had not been driven for a while**.

Excessive **wind noise** is noted primarily in **freeway driving at high speeds**.

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)

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

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.

PUTTING IT ALL TOGETHER

Encoding: designer

Decoding: user

1 Data

What is my data?
Which data type?
Ordinal / numerical / categorical?

2 Visual Mapping

What is my visual representation?
Which visual variables am I using?
How am I encoding my data?

3 Rendering

What is my medium?
monoscopic/stereoscopic?
Tangiblity?
Print / digical?

4 Perceiving

What does it show?
Where is big, medium, small?
How do things compare?
What relationships exist?

5 Interpreting

What does it mean?
What does color mean?
Wha does 'up 'mean?
What do these patterns show?

6 Comprehending

What does it mean for me?
What shall I do now?
Is this all true?
What do I learn?

View

DATA-DRIVEN STORYTELLING

Provide **context**

Persons, collection, importance, questions, ...

Explain **processes**

Transformation, selection, aggregation, ...

Explain **insights**

Trends, outliers, groups, comparisons, ...

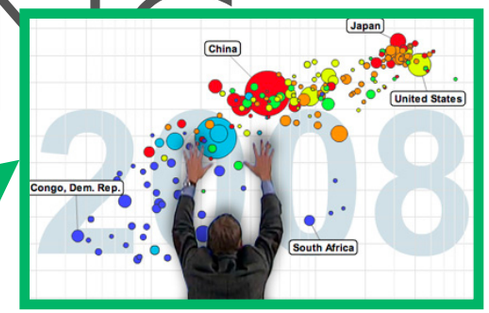
Explain **visualizations**

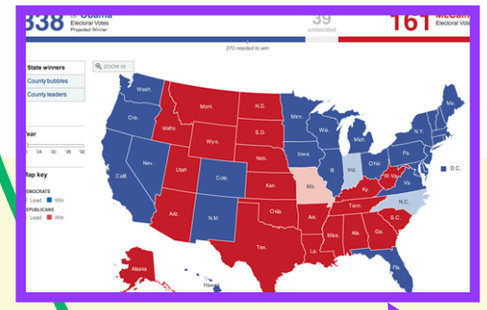
Talk to an **audience**

Conclude

Take home message

VISUALIZATION ENCODINGS





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, 6.6×10^6

Visual Attributes
- Position
- Size
- Hue
- Shape
- Brightness
- Etc

Marks
- Point
- Line
- Area

Layout
- Scatter
- Graph
- Line
- Tree map
- Etc

KNOW YOUR AUDIENCE'S TASK

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

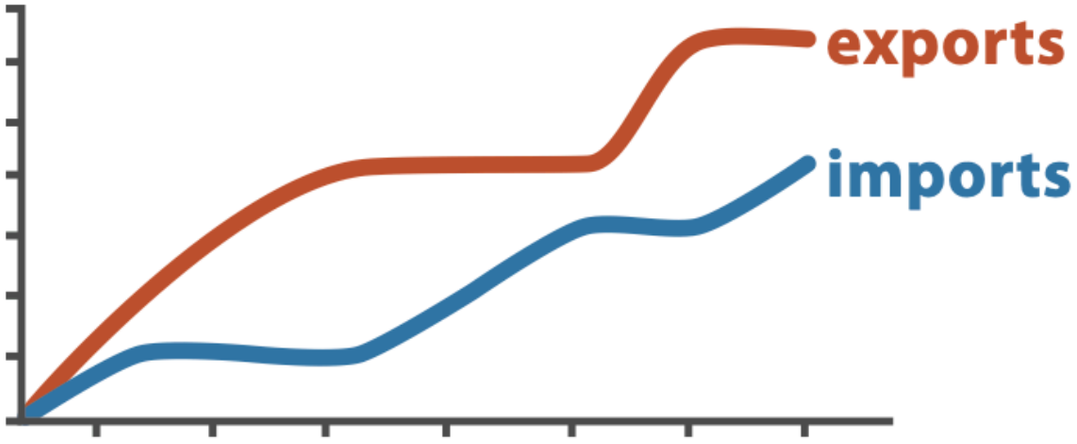
— Bill Buxton

"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"

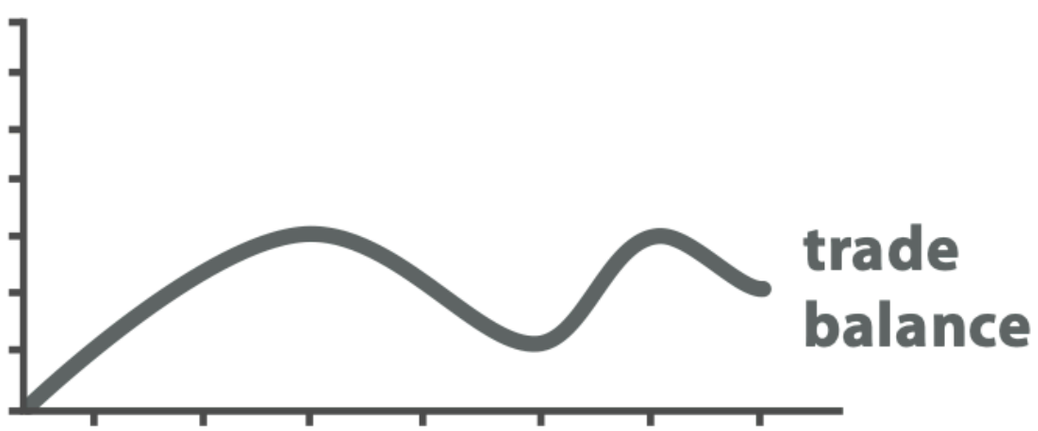
— Tamara Munzner

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

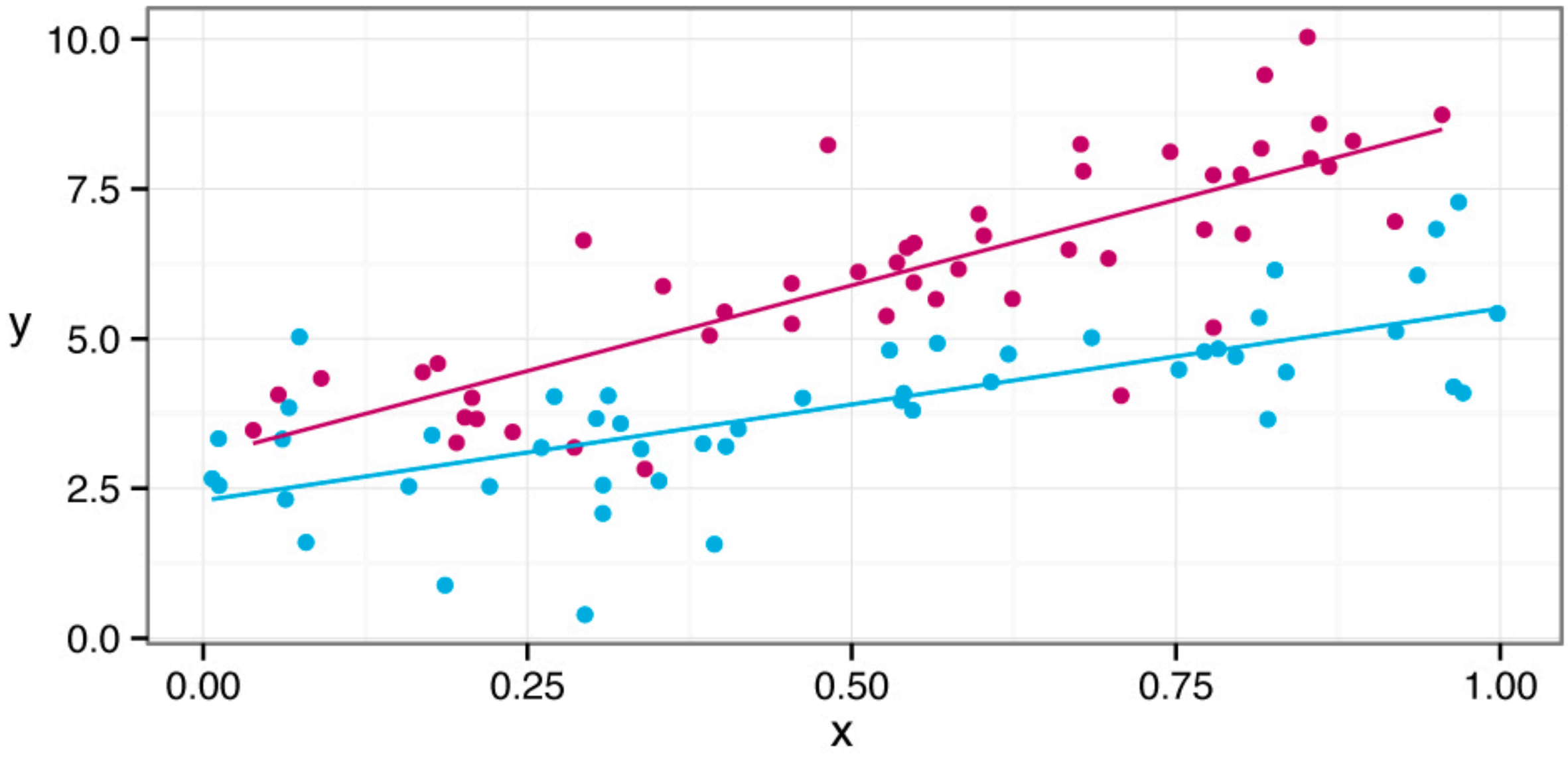


trade balance = exports – imports

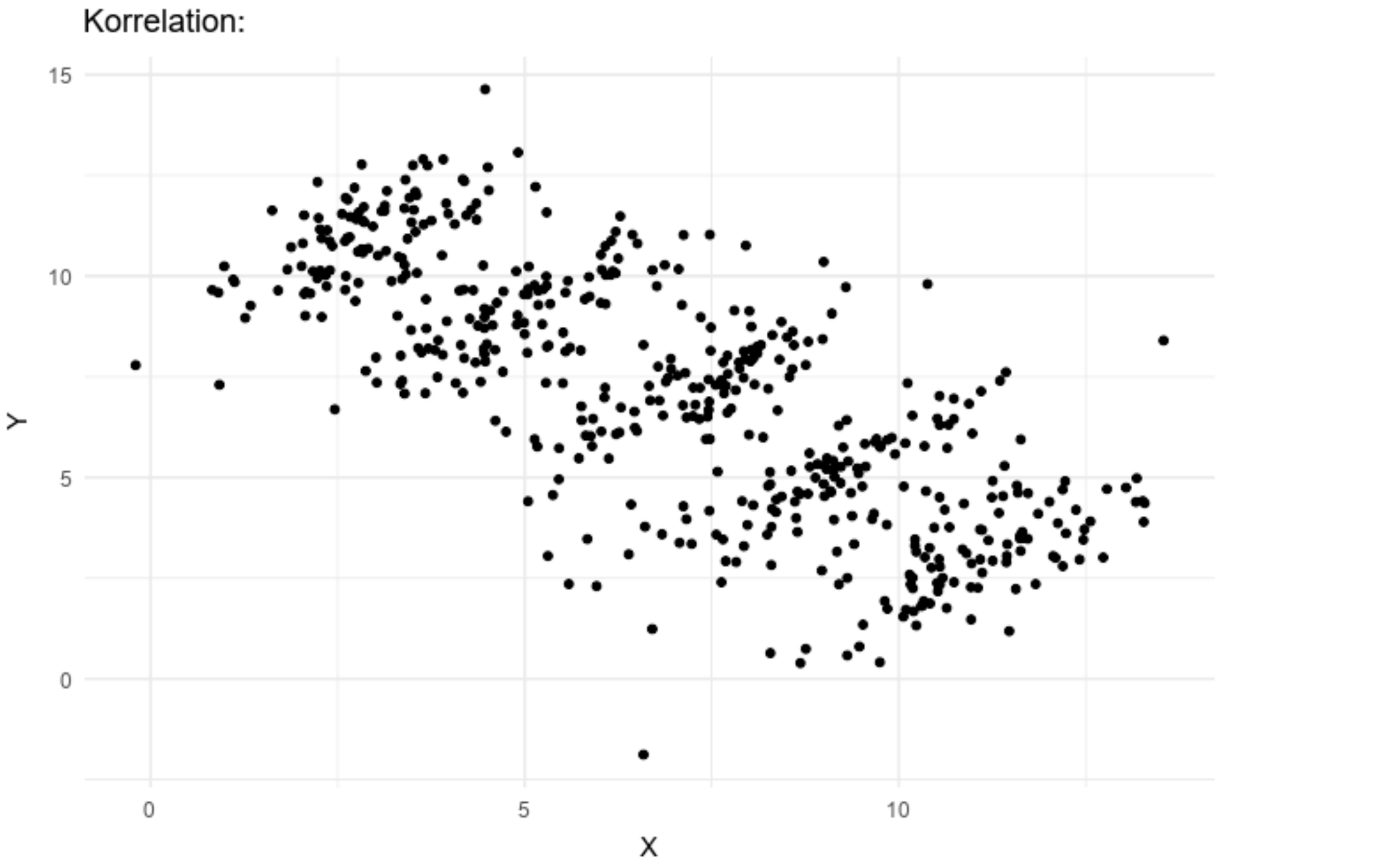
Derived Data

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?

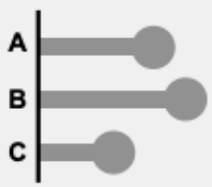
Categories — **Ordered Categories** — **Continuous Metrics**

Metric, split by 1 category

Bar (Row)

Rows allow readable labels, while columns awkwardly turn text

Lollipop



More focus on the positions of tops. Fun factor +1

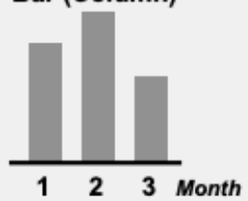
Dot Plot

A non-zero y-axis base may be less misleading here

Bar (Column)

Histogram. Boxes help convey the underlying bins

Bar (Column)



Time moves horizontally. So use Column, not Row

Area

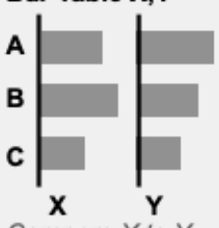
Adds continuity to x-axis

Line

A non-zero y-axis base may be less misleading here

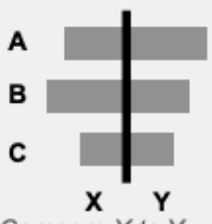
... by 2+ categories

Bar Table X,Y



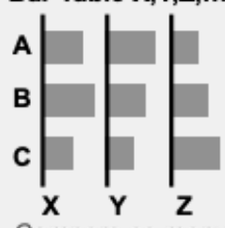
Compare X to Y, 'Small multiples'

Mirror Bar



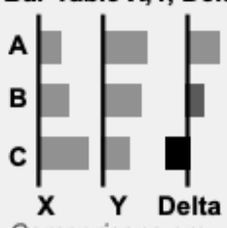
Compare X to Y

Bar Table X,Y,Z,...



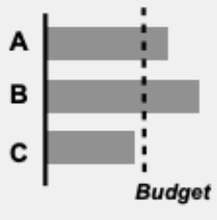
Compare as many as you like

Bar Table X,Y, Delta



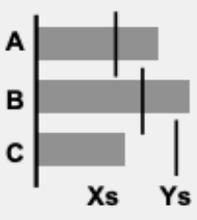
Comparisons are slow. Plot critical Deltas explicitly

Benchmark Bar



Compare X to a benchmark

Benchmarks Bar



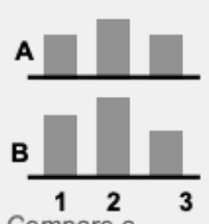
Compare X to Y. Fancier version called a 'Bullet graph'

Interleaved Bar

Compare X to Y (not recommended) ←

Interleaves two categories into one spatial dimension. Typically better to use Bar Table (above) instead

Bar Table

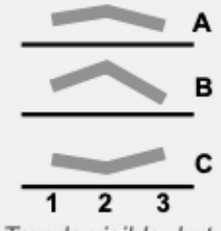


Compare a continuous metric across a category

Bar Line Table

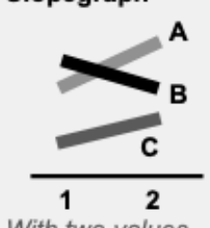
Compare two metrics

Line Table



Trends visible, but use Lines (below) if precision is key

Slopegraph



With two values, slope encodes delta

Dual Axis

Use (above) instead. Crossings here are salient, but meaningless

Lines

Compare many. Getting spaghetti? Use Line-Table (above)

WHICH VISUALIZATION?

… by 2+ categories — Part-to-Whole, split by 1 category

Pie

Screams 'Percentages!'

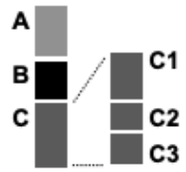
Stacked Bar (Row)

More precise and flex, but less screaming

Stacked Bar (Col)

Now I'm standing

Breakout Bar



Let's zoom in here. User different colors. Global starts left (or top)

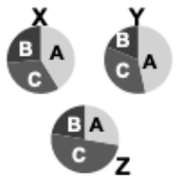
Waterfall

Waterfalls are vertical stacked bars that narrate financial values in a (typically) artificially imposed ordering

These lines are identical, with equal vertical separation at each X

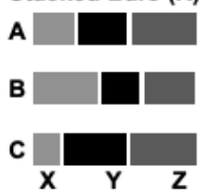
Beware of an illusion for these: seeing differences (lines), or category values (stacked area) can be difficult and even misleading

MultiPie



Please don't… (not recommended)

Stacked Bars (R)

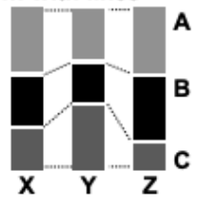


…Use this instead

Stacked Bars (C)

Horizontal flow implies an ordering

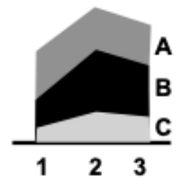
… with lines



Added lines suggest continuity, help depict changes

No legends. Instead, directly label actual values

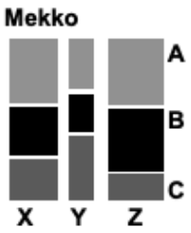
Stacked area



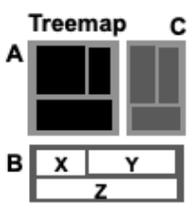
Now it's definitely continuous

Mekko stacks in X and Y simultaneously. Here XYZ might be absolute values of a market, and ABC are company market shares as %'s

Mekko



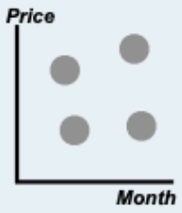
Treemap



Hierarchy, ~3 levels max. Size & color code different metrics. Often misused (info)

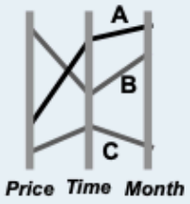
Metrics: relationships to other metrics

Scatter



A most powerful chart

Parallel Coordinates



Instead of Perpendicular Cartesian axes, Parallel format allows more axes

Look at this number
Just look at it.

Dot Array

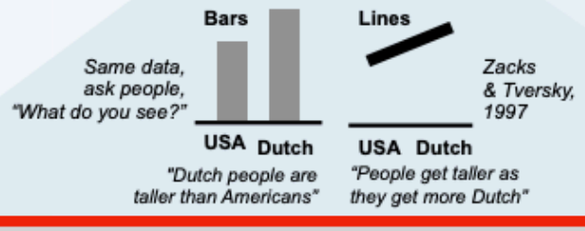
Dot Array %

43%
Huge Number

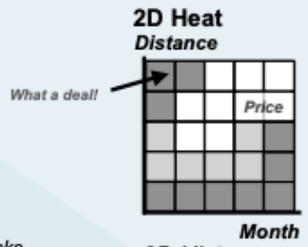
Icon Array (ISOTYPE, info)

Icon Array %

Connecting Lines imply continuous data

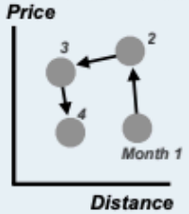


2D Heat



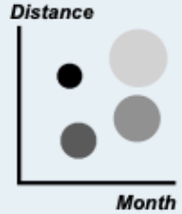
2D Histogram

Connected Scatter



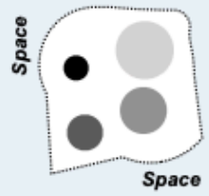
The power of a scatterplot, with points connected by time (info)

Rosling Scatter



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

Map



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

v1.2 Download the most recent version at ExperCeption.net

@SteveFranconeri

VISUALIZATION CHEAT SHEET

https://visualizationcheatsheets.github.io/

Cheat Sheets for Visualization Techniques

visualizationcheatcheets.github.io

Home
Download all (PDF+PNG), 76MB
Download guidelines (PDF), 8MB

Paper

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

By Type

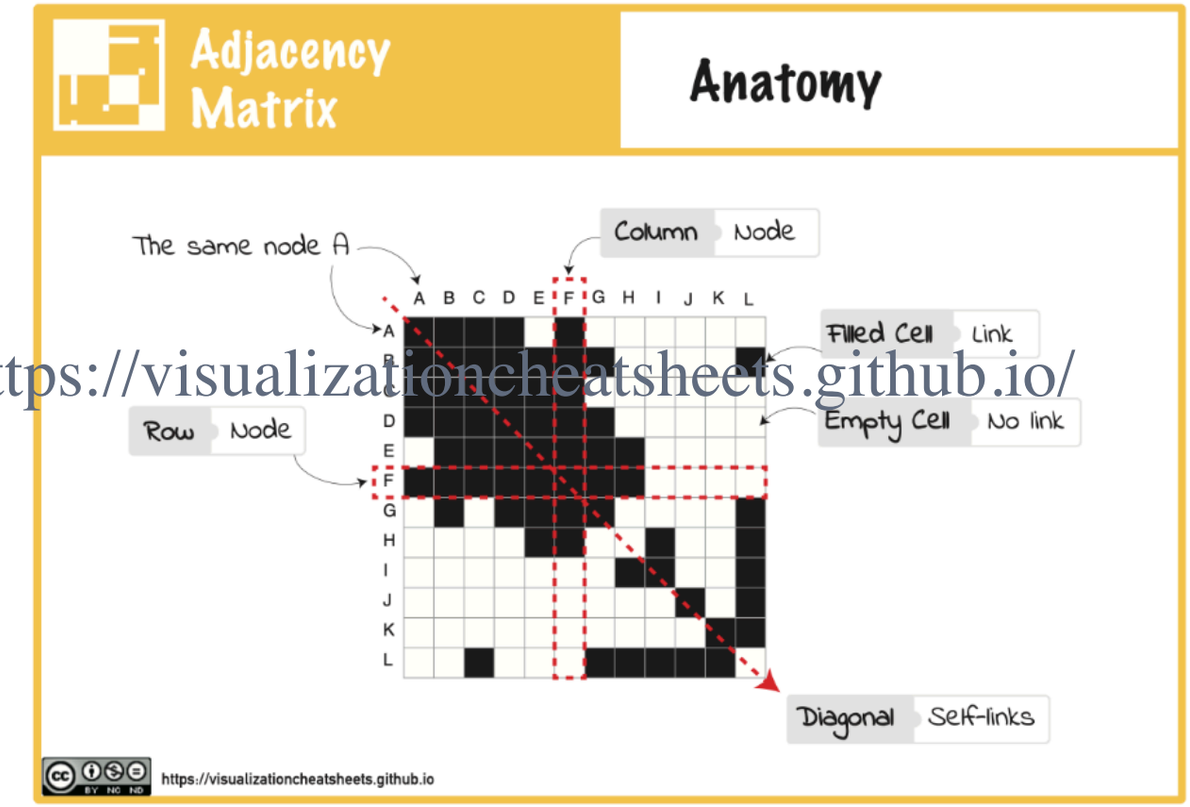
Anatomy | Introduction | Construction
Visual Pattern | Pitfalls
Well-known Relative | False Friends

By Visualization

Boxplots
Confluence Graphs
Adjacency Matrix
Parallel Coordinates

By Type

Anatomy



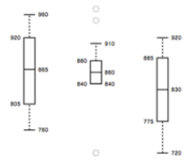
Anatomy explains the visual elements of a visualization technique, their composition, their specific terminology, and how they relate to the data. The visual components of a visualization can include individual visual marks as well as groups of marks, axes, locations in a visualization etc.

Construction

Examples, lots of examples!

Fundamentals

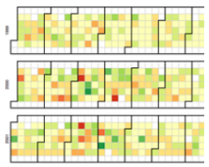
Box Plots



Bubble Chart

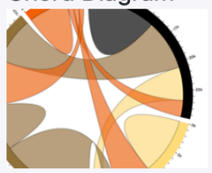
Bullet Charts

Calendar View



Non-contiguous Cartogram

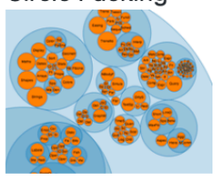
Chord Diagram



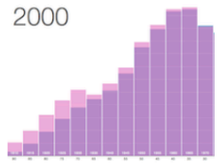
Dendrogram

Force-Directed Graph

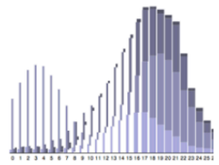
Circle Packing



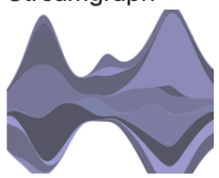
Population Pyramid



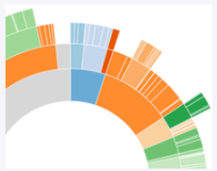
Stacked Bars



Streamgraph



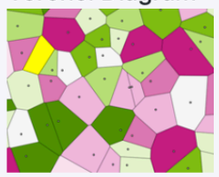
Sunburst

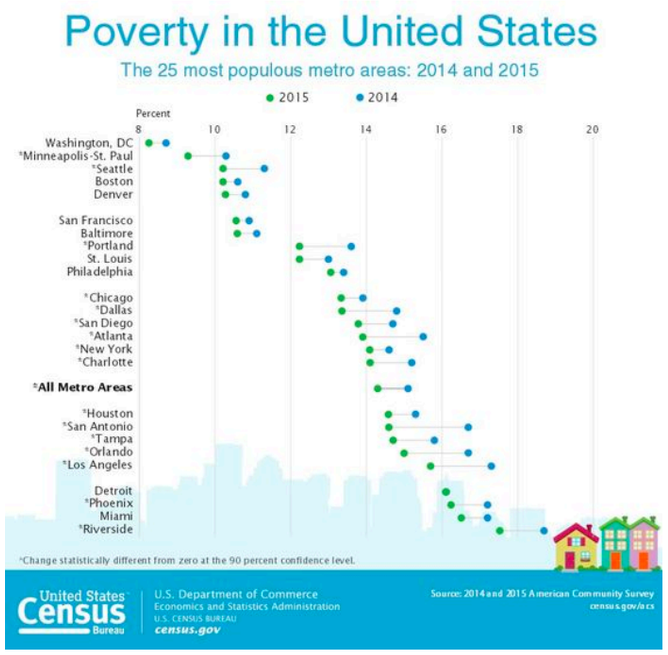


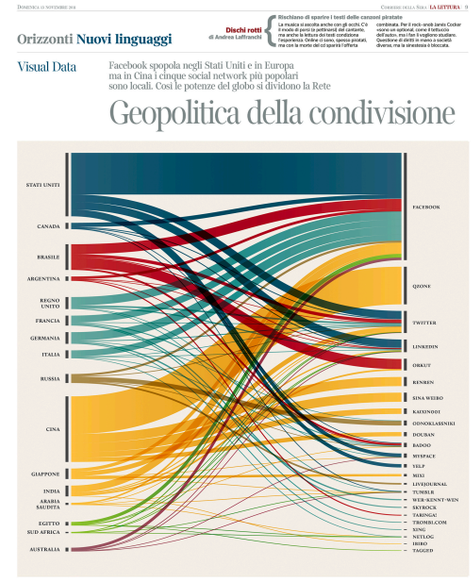
Node-Link Tree

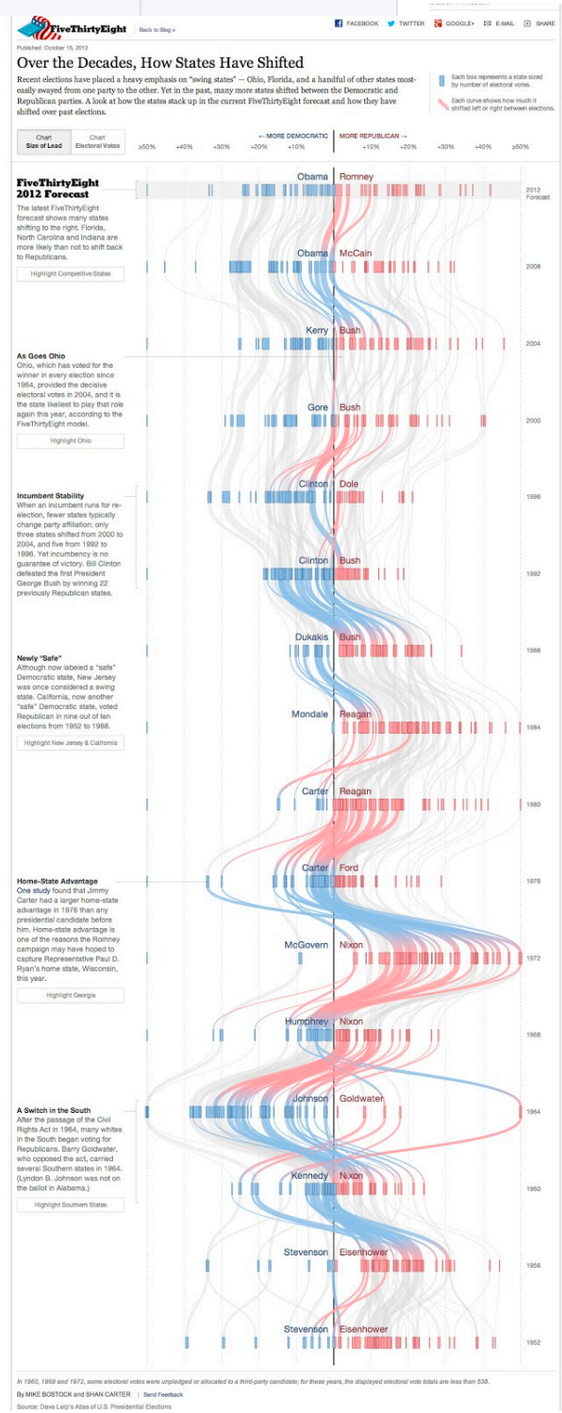
Treemap

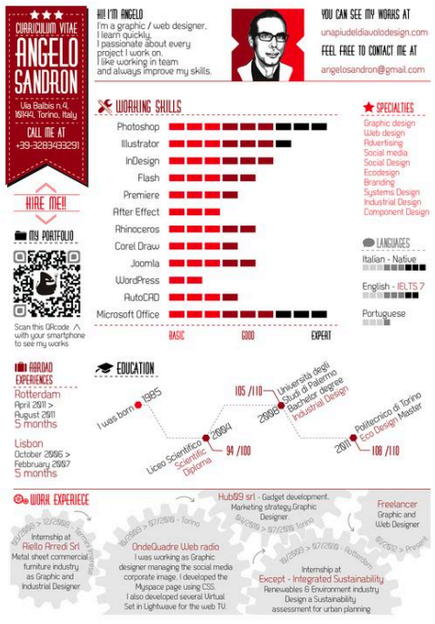
Voronoi Diagram

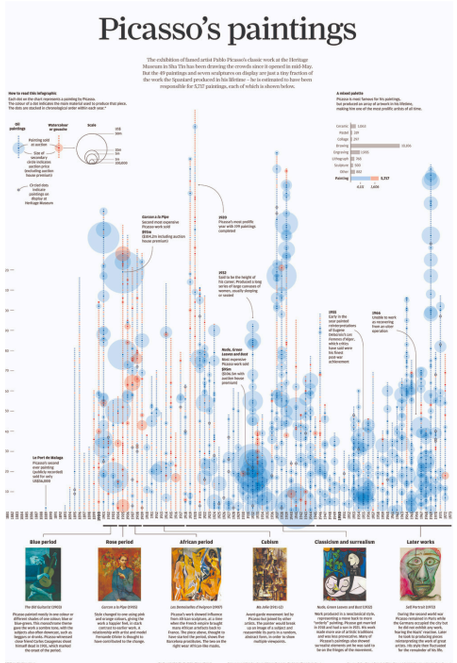


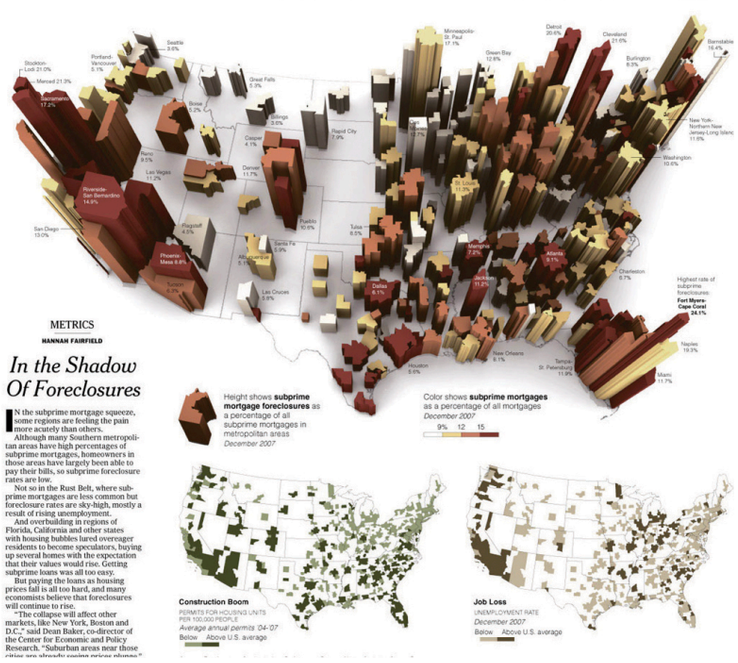










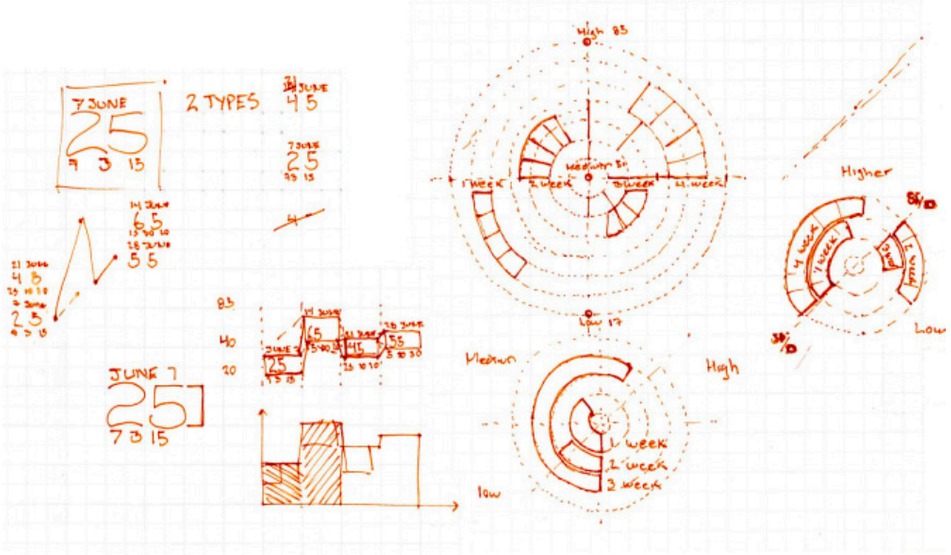


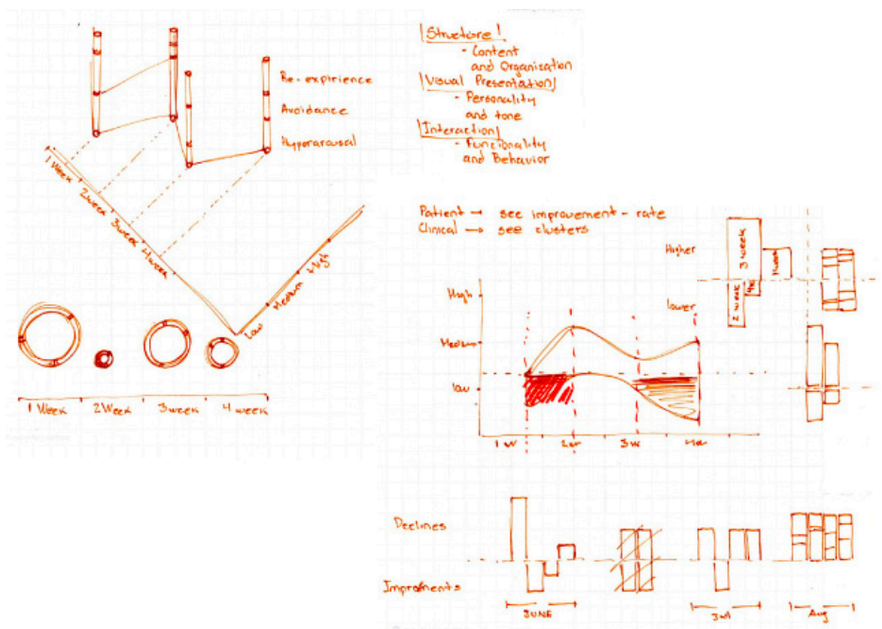
SKETCH!

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



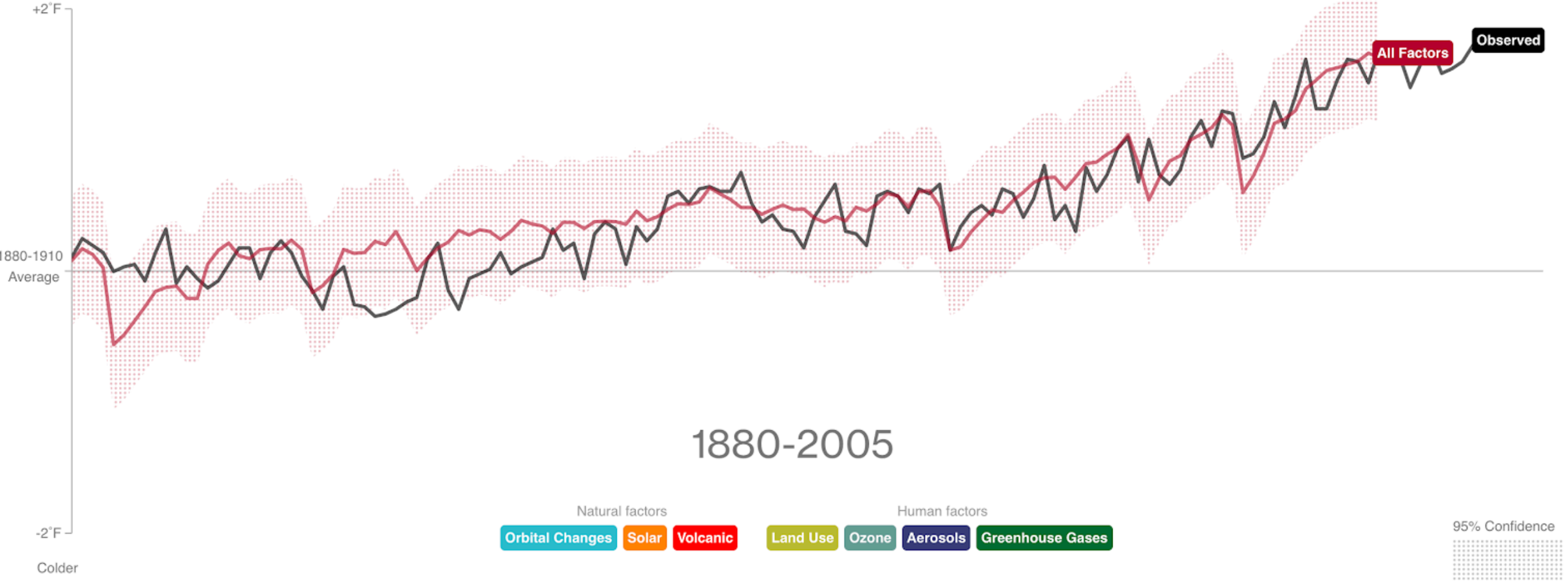


NARRATIVE PATTERNS

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?

Hotter



https://www.bloomberg.com/graphics/2015-whats-warming-the-world/

NARRATIVE PATTERNS

http://napa-cards.net

NAPA Cards *Preview*

Narrative Patterns for Data-Driven Storytelling

What is this?

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

Rhetorical question

HOW

Title the piece using a question that makes a point.

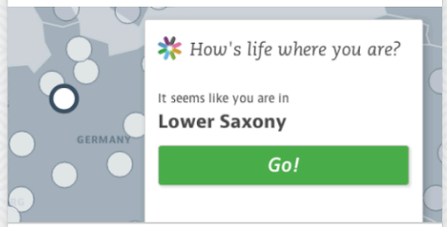
WHY

Implicitely 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.

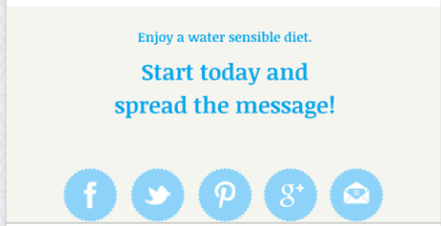
WHY

To lower the entry barrier and make information personally relatable.

EXAMPLE

OECD: Regional Well-Being

Call to action



HOW

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

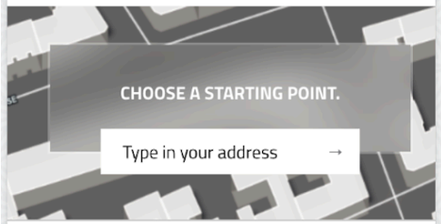
WHY

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.

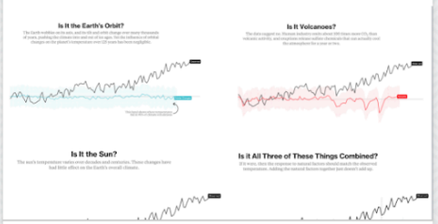
WHY

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.

EXAMPLE

Bloomberg: What's really warming the world?

Juxtaposition

HOW

Showing two or more visualizations (pictur

WHY

Compare cases or sa difference, absense c Allows for individual

EXAMPLE

David McCandless: Billio

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

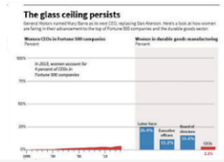
HOW

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



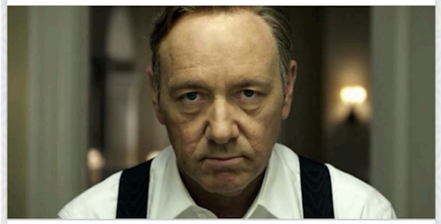
HOW

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

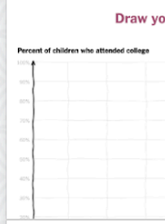
Show that there are concrete entities (e.g., a human being) behind, or at the heart of the data.

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 no interaction: intern directly addressed to after. If interaction: p

WHY

Stimulates the viewe game-lke experience reasoning about the

EXAMPLE

New York Times graphic

REFERENCES

EXPAND YOUR HORIZON

Lookup

Visualization Catalogs

www.datavis.ca

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

(historical!)

www.infovis-wiki.net

http://www.visual-literacy.org/periodic_table/periodic_table.html

Color

http://colorbrewer2.org/

Color Brewer

http://tools.medialab.sciences-po.fr/iwanthue/

I Want Hue

To Reflect:

http://hint.fm/

F. Viégas & M. Wattenberg

www.edwardtufte.com/bboard

E. Tufte

http://fellinlovewithdata.com/

Fell in love with data

E. Bertini

http://datastori.es/ (podcasts)

E. Bertini & M. Stefaner

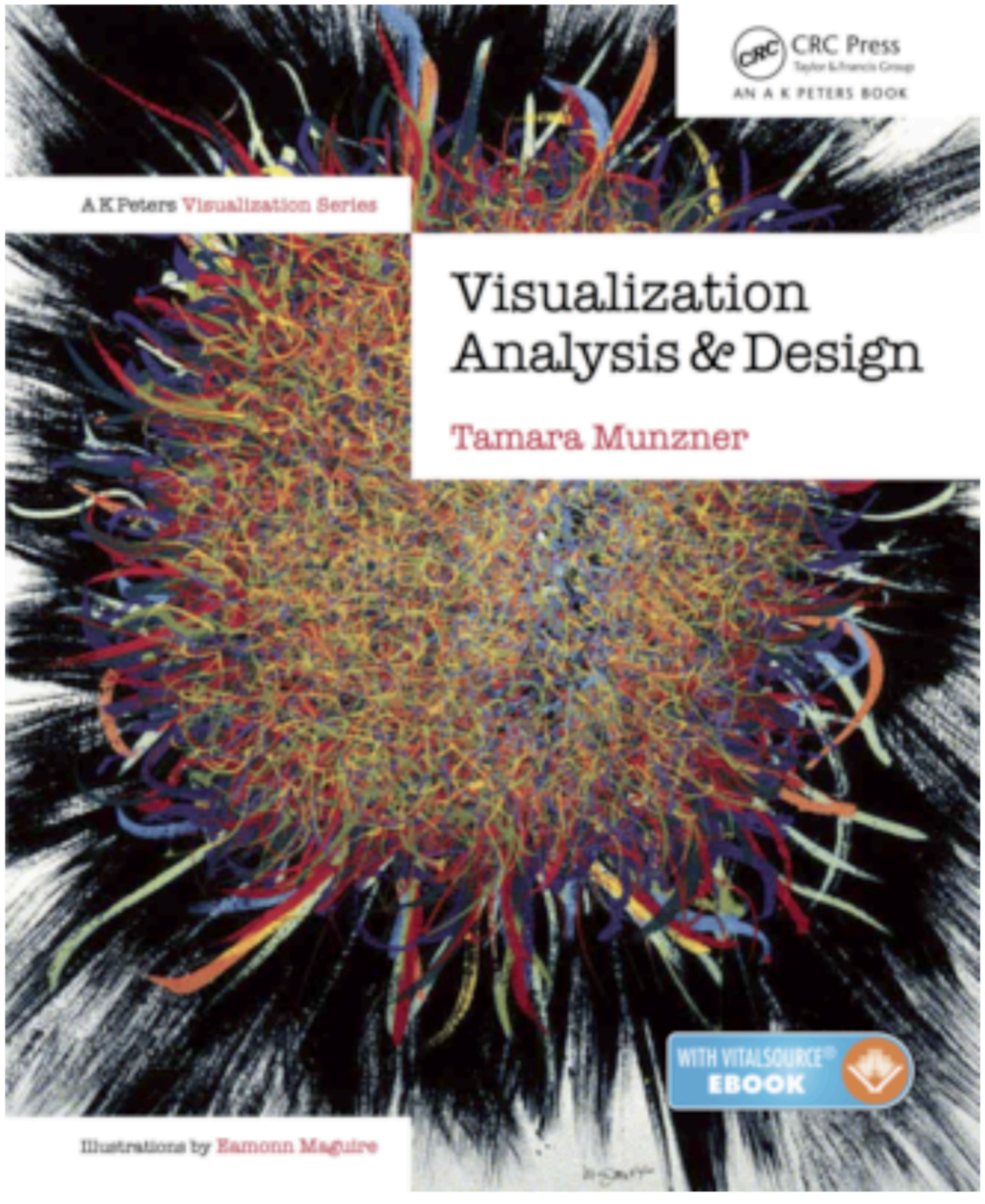
http://junkcharts.typepad.com/

Junk Charts: learn from others (mistakes)

Text-book on data visualization

(1st of its kind)

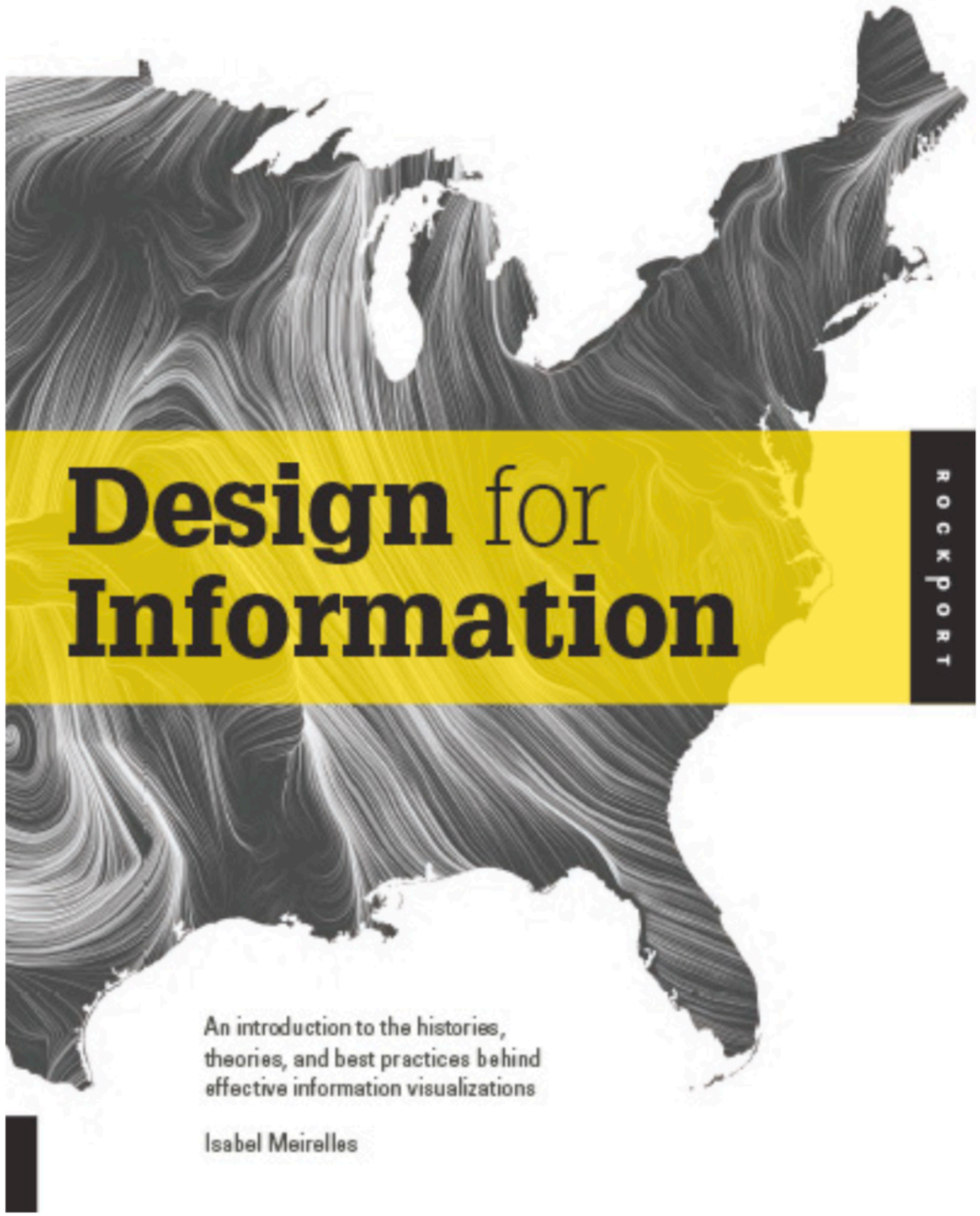
If you can only get one,
THIS is the one!



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

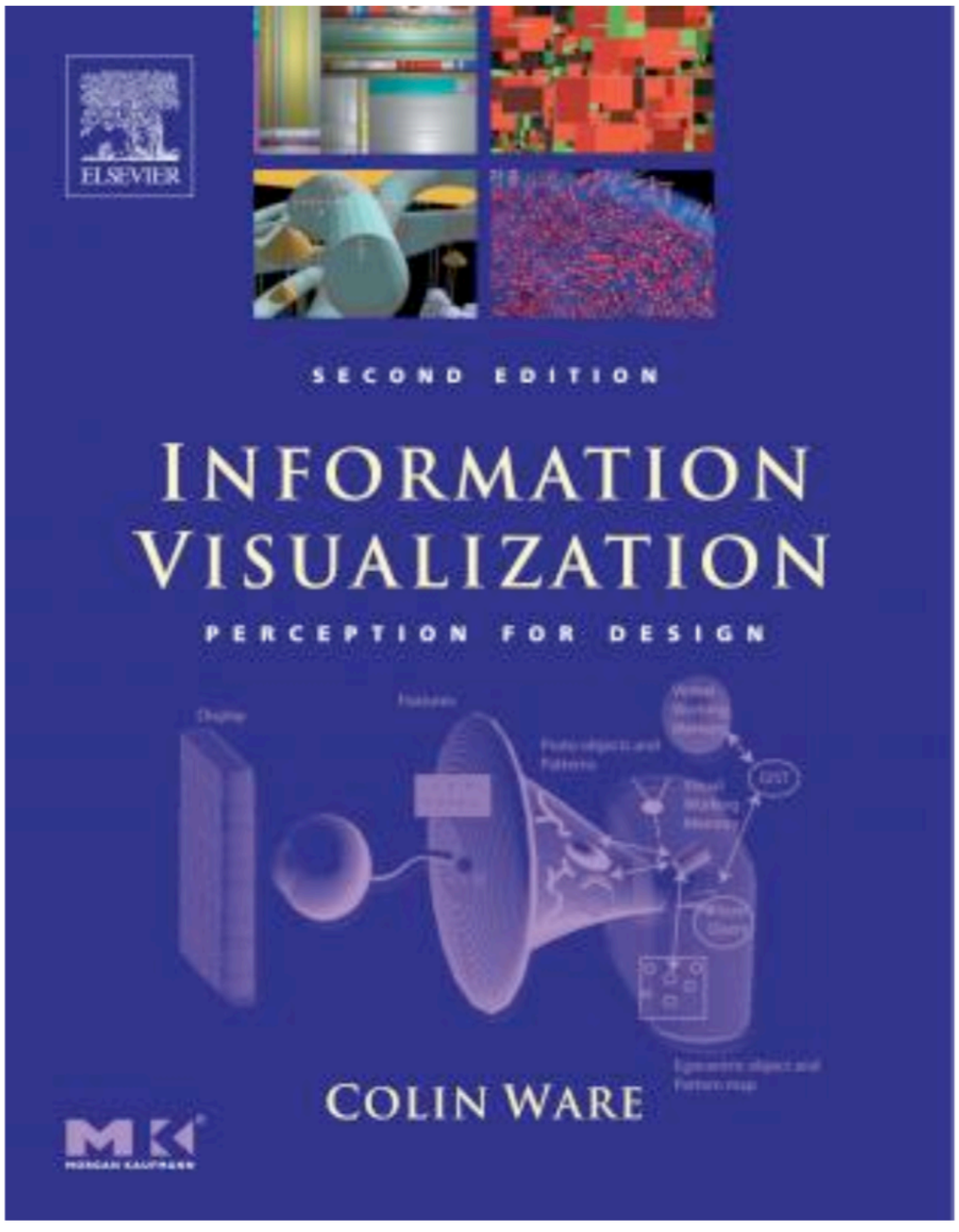
Infovis through examples, history and some free discussion. Very visual.

Beautiful!



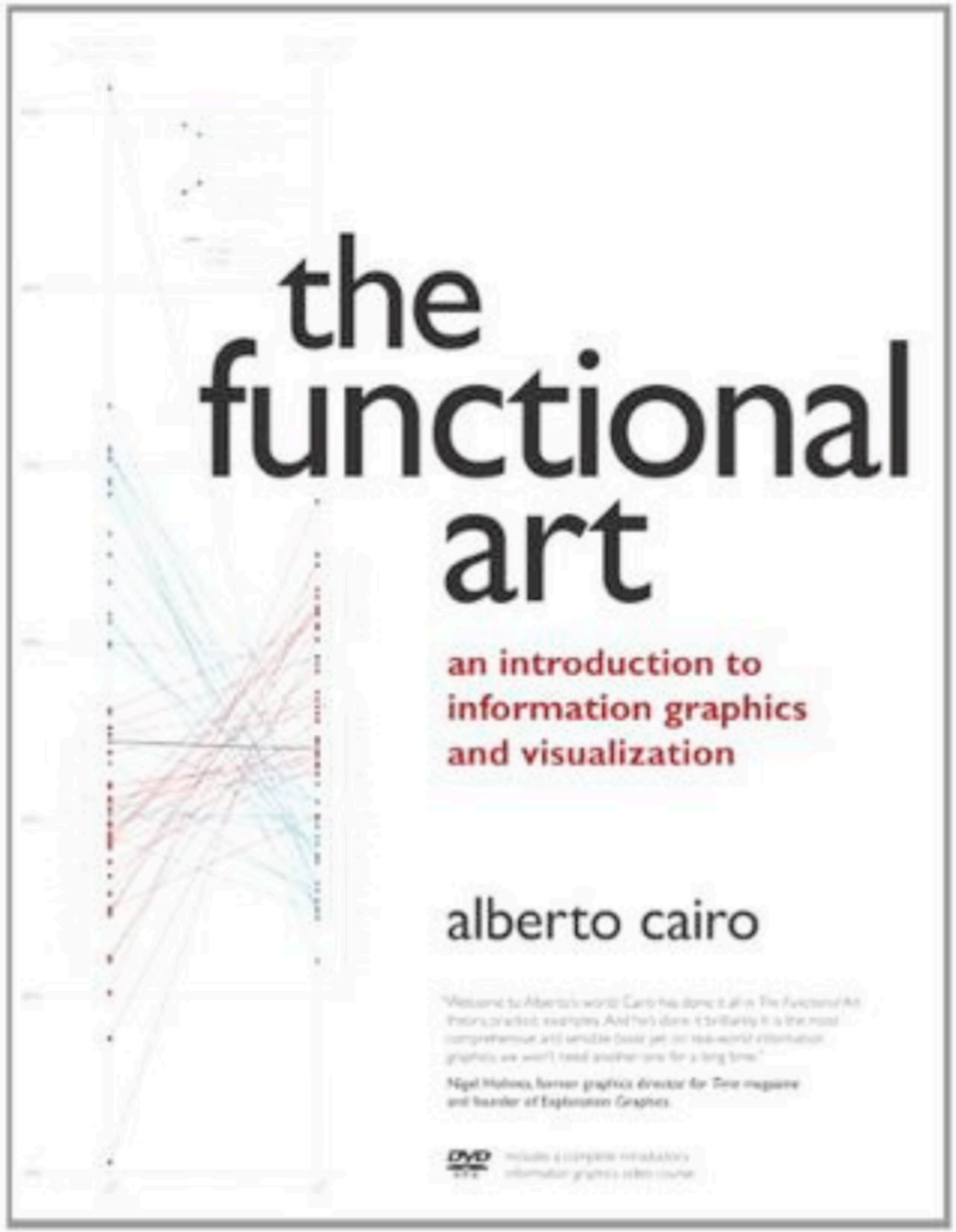
Meirelles, Isabel. Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations. Rockport publishers, 2013.

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



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

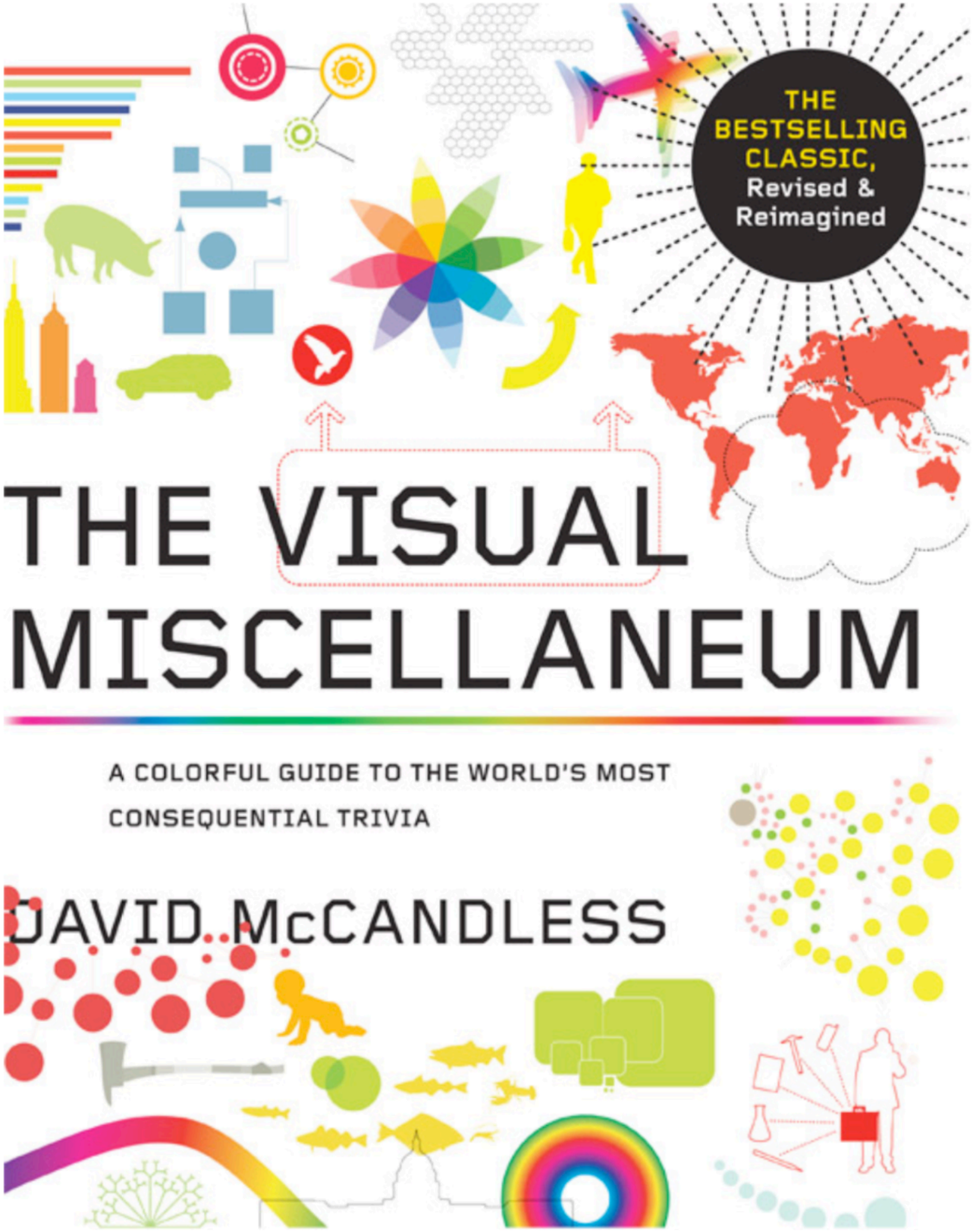
Focus on
infographics



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.