BEFORE WE START…. A LITTLE EXPERIMENT

37 75

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

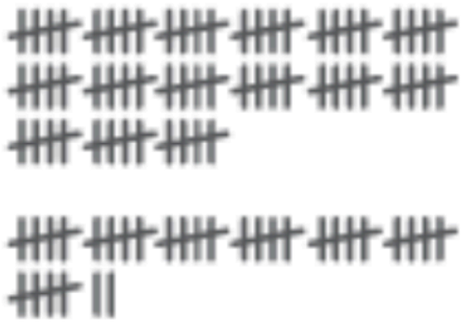
2. Try to come up with as many possible representations/encodings for the “data" above as you can, in the paper segments.

Feel free to be creative!

37 75

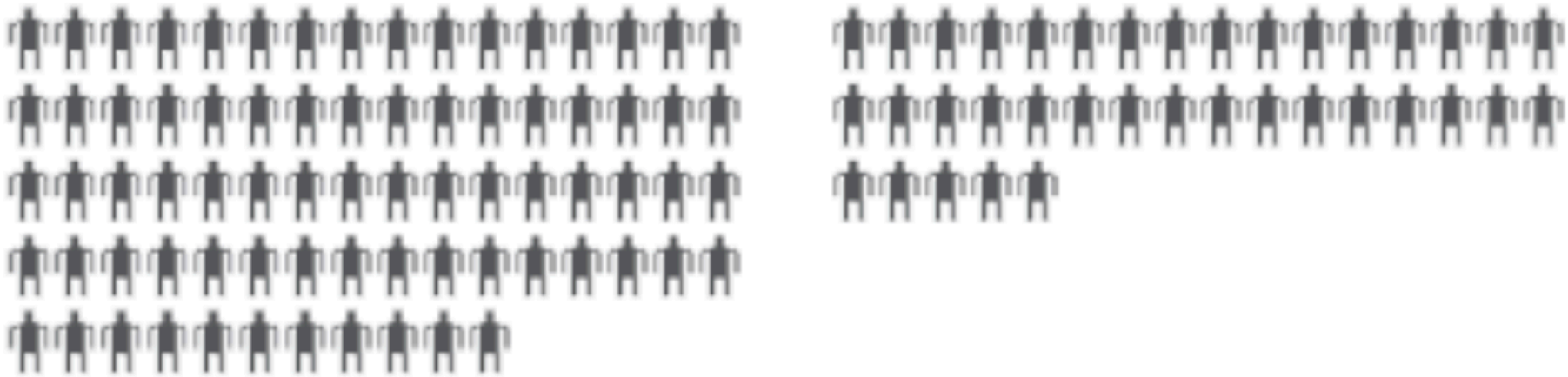
Thirty-seven
Seventy-five

XXXVII LXXV



Squares

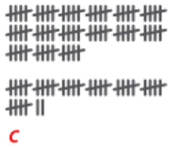
Isotypes



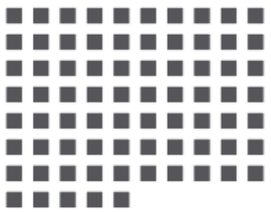
75, 37

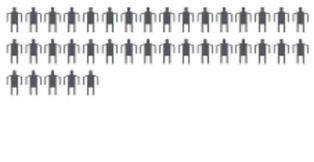
a

b



c

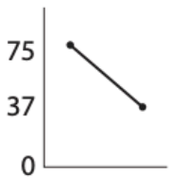




a

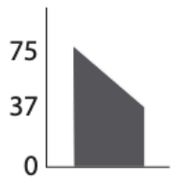
b

c



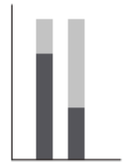
a

b



c

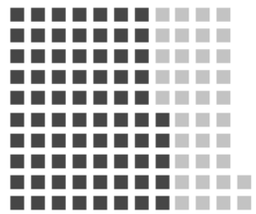
a



b

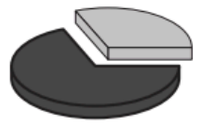
a

b



a

b



c

a

b

c

a

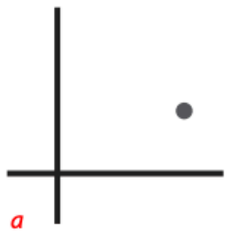
b

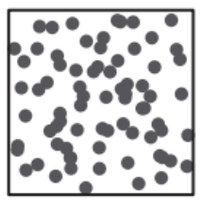
a

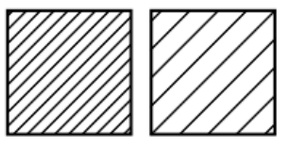
b

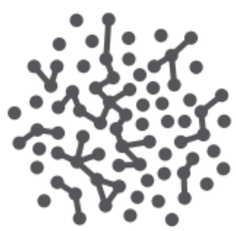
a

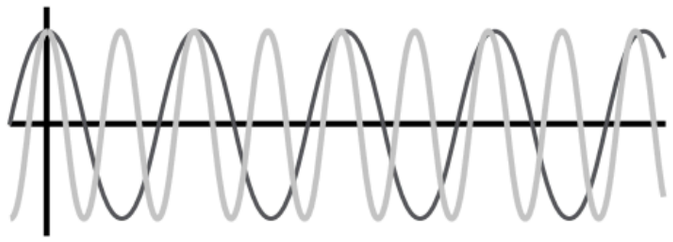
b

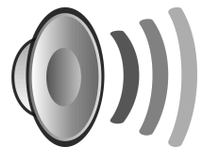






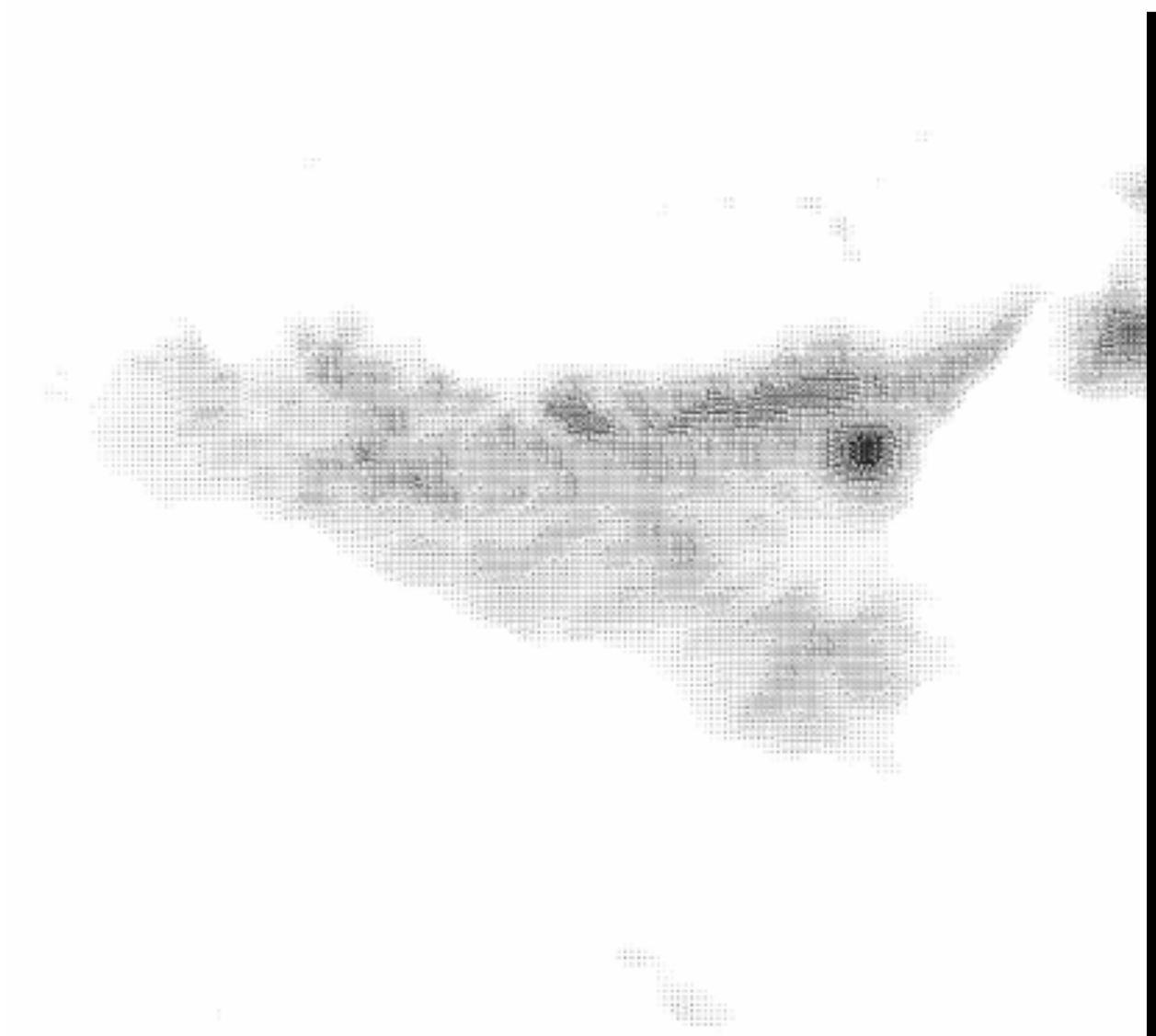


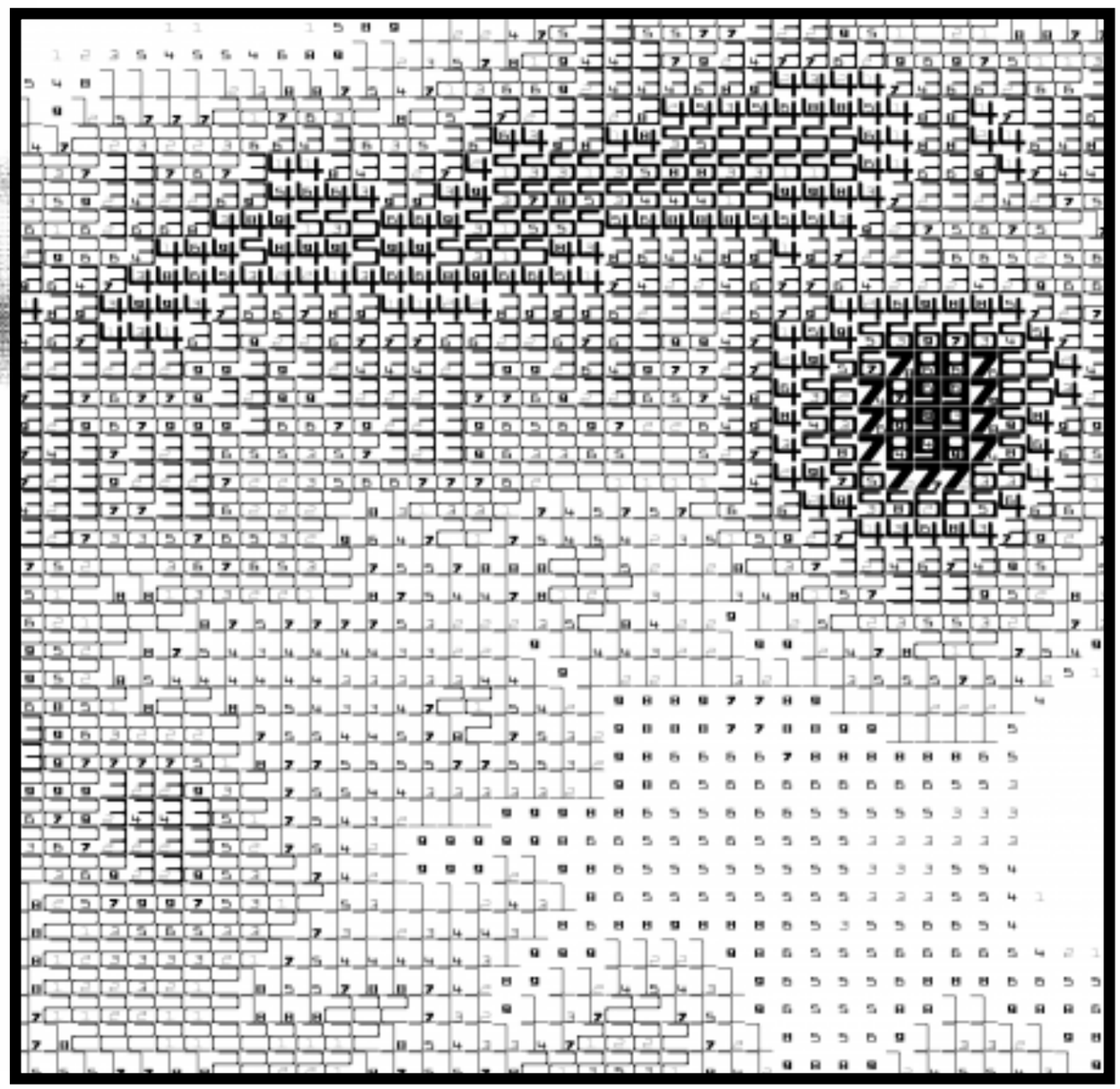




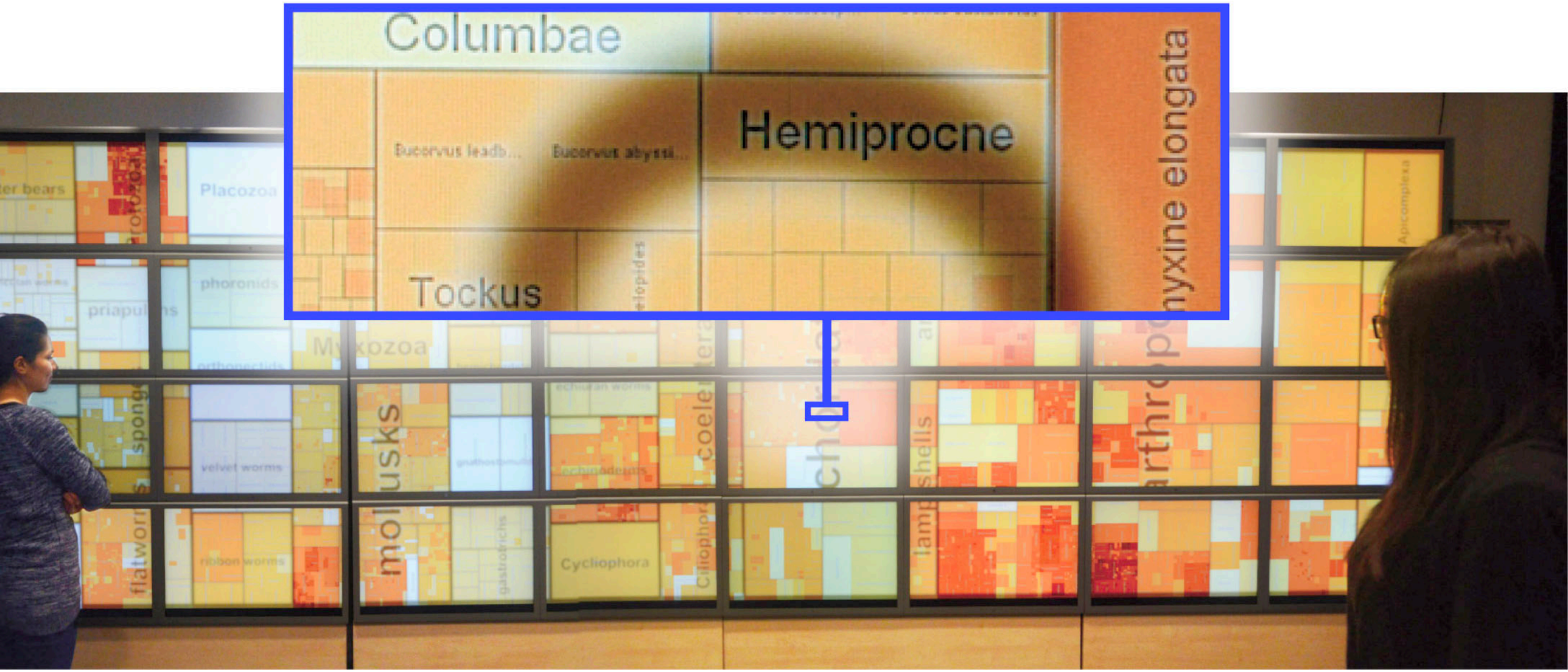
37Hz

75Hz



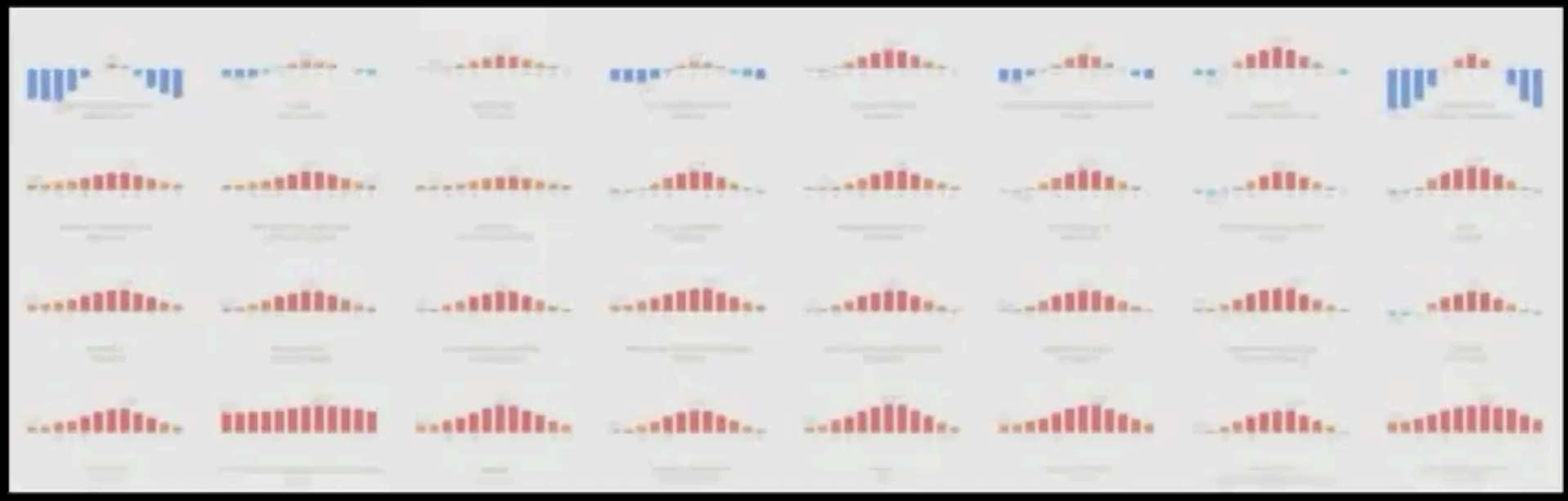


[Source: M. Nacenta, U. Hinrich, S. Carpendale - http://fatfonts.org]



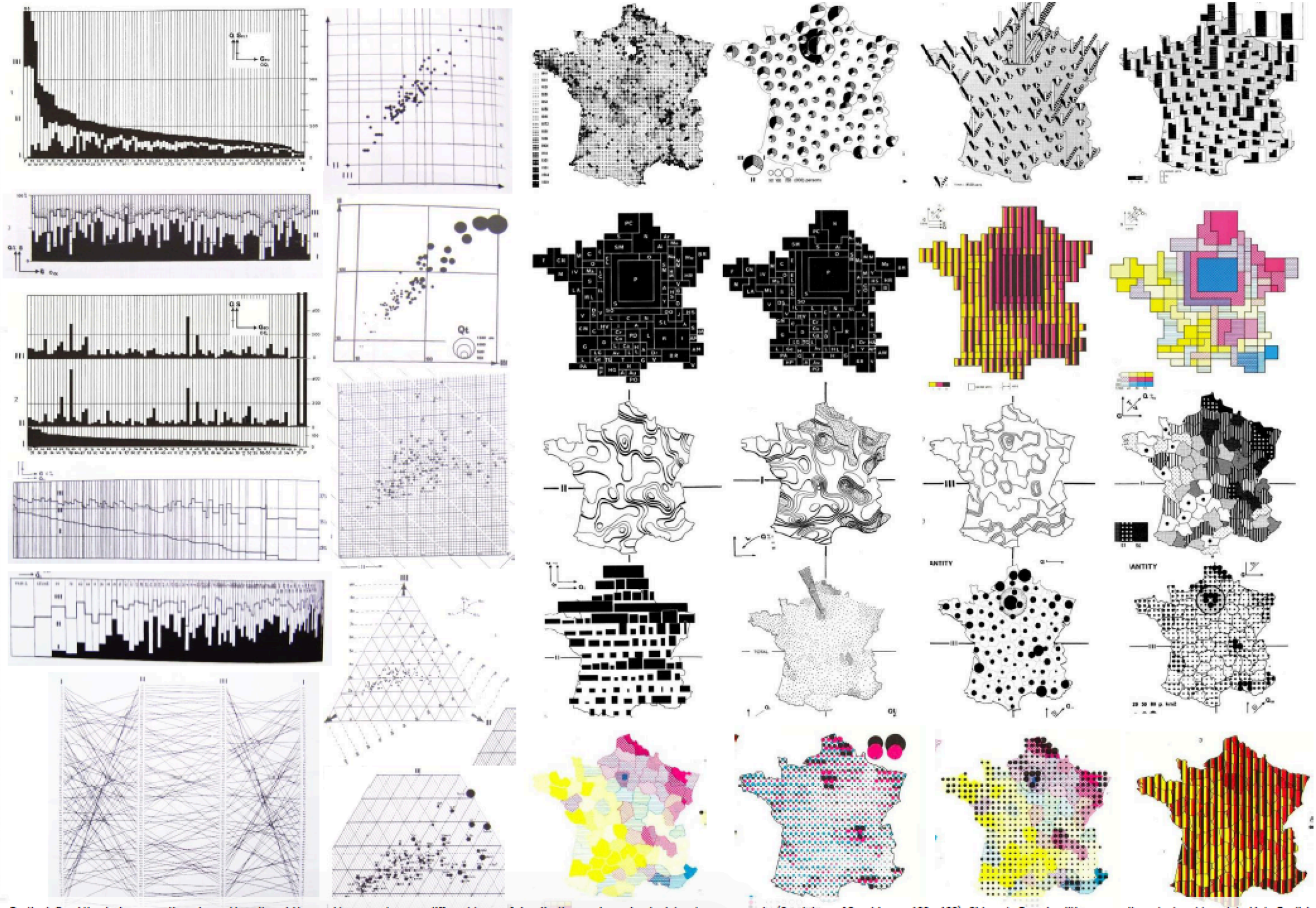
[Hybrid-Image Visualization for Large Viewing Environments Petra Isenberg, Pierre Dragicevic, Wesley Willett, Anastasia Bezerianos, Jean-Daniel Fekete IEEE Transactions on Visualization and Computer Graphics, Institute of Electrical and Electronics Engineers, 2013, 19 (12), pp.2346--2355.]

www.fraps.com



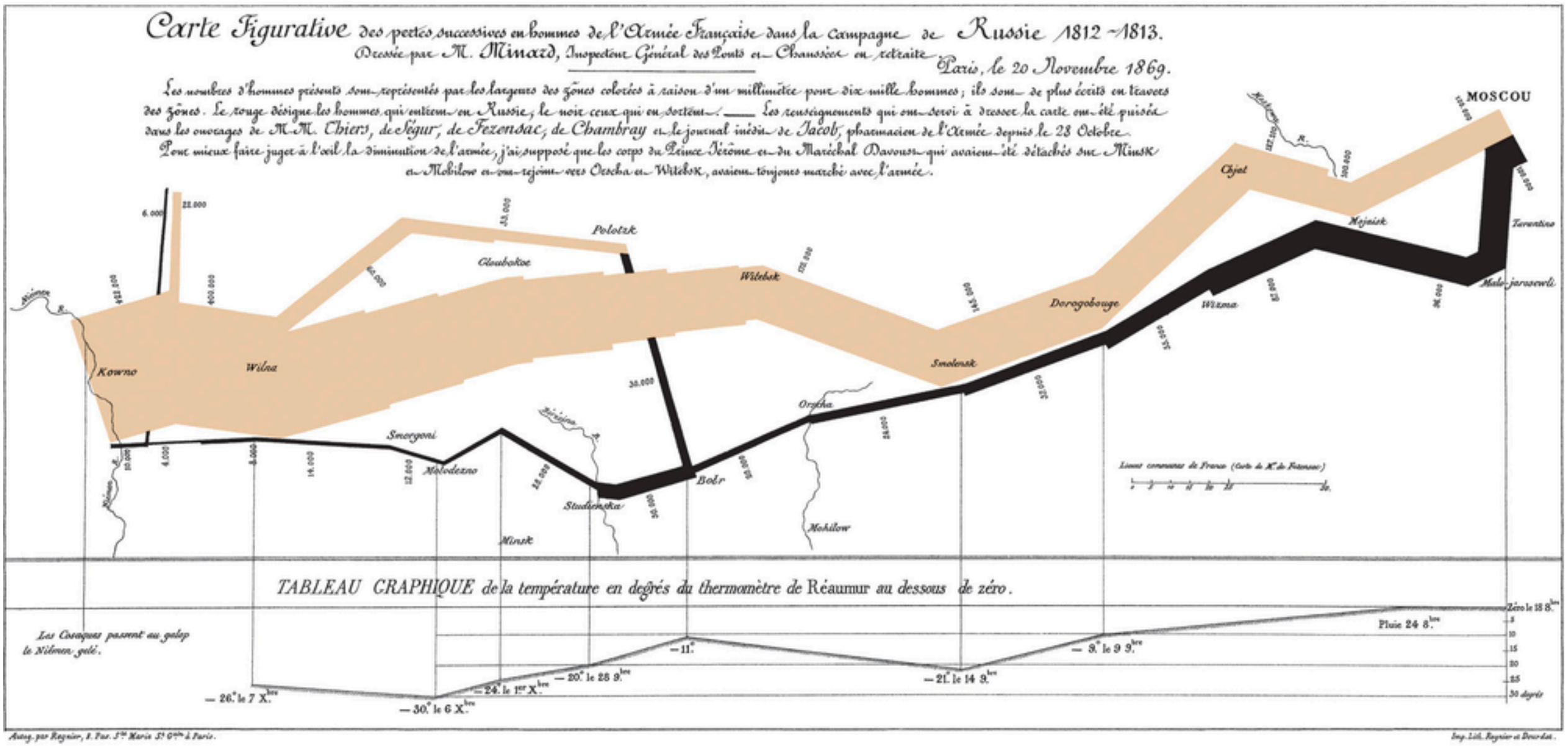
[Hybrid-Image Visualization for Large Viewing Environments Petra Isenberg, Pierre Dragicevic, Wesley Willett, Anastasia Bezerianos, Jean-Daniel Fekete IEEE Transactions on Visualization and Computer Graphics, Institute of Electrical and Electronics Engineers, 2013, 19 (12), pp.2346--2355.]

Départements	Quantités (000)				Proportion %		
	I	II	III	Total	I	II	III
1 AIN	67	43	40	150	45	28	27
2 AISNE	56	74	66	193	29	37	34
3 ALLIER	65	45	57	167	39	27	34
4 Bses ALPES	15	8	13	35	43	24	33
5 Htes ALPES	16	8	13	37	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 AUDE	38	48	36	132	25	43	32
11 AUBE	50	20	32	102	49	19	32
12 AVEYRON	70	32	29	131	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	36	38	140	47	26	27
17 CHARENTE Mme	79	39	65	183	43	21	36
18 CHER	43	41	36	120	36	34	30
19 CORREZE	64	23	29	116	55	20	25
21 COTE D'OR	43	41	59	143	30	29	41
22 COTES DU NORD	131	35	62	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	32	30
27 EURE	48	52	45	145	33	36	31
28 EURE & LOIR	44	27	38	109	41	25	34
29 FINISTERE	164	76	89	329	50	23	27
30 GARD	40	51	52	144	28	36	36
31 HAUTE GARONNE	64	67	84	216	30	31	39
32 GERS	63	10	16	89	71	11	18
33 GIRONDE	115	107	179	401	30	27	43
34 HERAULT	62	40	71	173	36	23	41
35 ILLE & V.	137	60	82	279	49	21	30
36 INDRE	54	30	32	116	46	26	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	25	28	123	57	20	23
41 LOIR & CHER	51	27	30	108	47	25	28
42 LOIRE	56	150	82	288	19	54	27
43 Hte LOIRE	52	23	22	97	54	24	22
44 LOIRE INF.	101	108	105	314	32	34	34
45 LOIRET	51	51	34	136	37	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	65	234	44	28	28
50 MANCHE	118	42	56	216	54	20	26
51 MARNE	44	57	67	168	26	34	40
52 Hte MARNE	25	28	28	81	31	35	34
53 MAYENNE	74	23	26	123	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	40	69	55	164	24	42	[illegible]
61 ORNE	65	30	35	130	50	23	[illegible]
62 P.D.C.	94	242	137	473	20	51	[illegible]
63 PUY DE DOME	80	79	63	222	36	36	[illegible]
64 Bses PYRENEES	80	49	62	191	42	25	[illegible]
65 Htes PYRENEES	37	27	28	92	40	29	[illegible]
66 PYRENEES ORIENT.	35	20	33	88	40	23	[illegible]
67 BAS-RHIN	76	122	114	312	24	39	[illegible]
68 Ht-RHIN	40	121	74	235	17	51	[illegible]
69 RHONE	44	215	194	453	10	47	[illegible]
70 Hte SAONE	34	32	23	89	38	36	[illegible]
71 SAONE & L.	94	77	62	233	41	33	[illegible]
72 SARTHE	87	45	58	190	46	24	[illegible]
73 SAVOIE	44	38	35	117	38	32	[illegible]
74 Hte SAVOIE	52	42	45	139	37	30	[illegible]
P PARIS	2	575	940	1517	0	38	[illegible]
75 SEINE	8	574	550	1132	1	51	[illegible]
76 SEINE INF.	75	152	174	401	19	38	[illegible]
77 SEINE & M.	37	72	76	185	20	39	[illegible]
78 SEINE & O.	46	328	356	730	6	45	[illegible]
79 DEUX-SEVRES	71	29	33	133	53	22	[illegible]
80 SOMME	57	68	61	186	31	36	[illegible]
81 TARN	55	47	33	135	41	35	[illegible]
82 TARN & G.	44	13	18	75	59	17	[illegible]
83 VAR	33	50	81	164	20	31	[illegible]
84 VAUCLUSE	40	30	41	111	36	27	[illegible]
85 VENDEE	110	38	40	188	59	20	[illegible]
86 VIENNE	60	29	39	128	47	23	[illegible]
87 Hte VIENNE	64	47	45	156	41	30	[illegible]
88 VOSGES	36	95	43	174	21	54	[illegible]
89 YONNE	41	28	37	106	39	26	[illegible]
90 BELFORT	3	25	13	41	8	60	[illegible]
	5212	6795	6995	18825	28	35	[illegible]



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.

Visualization theory and principles

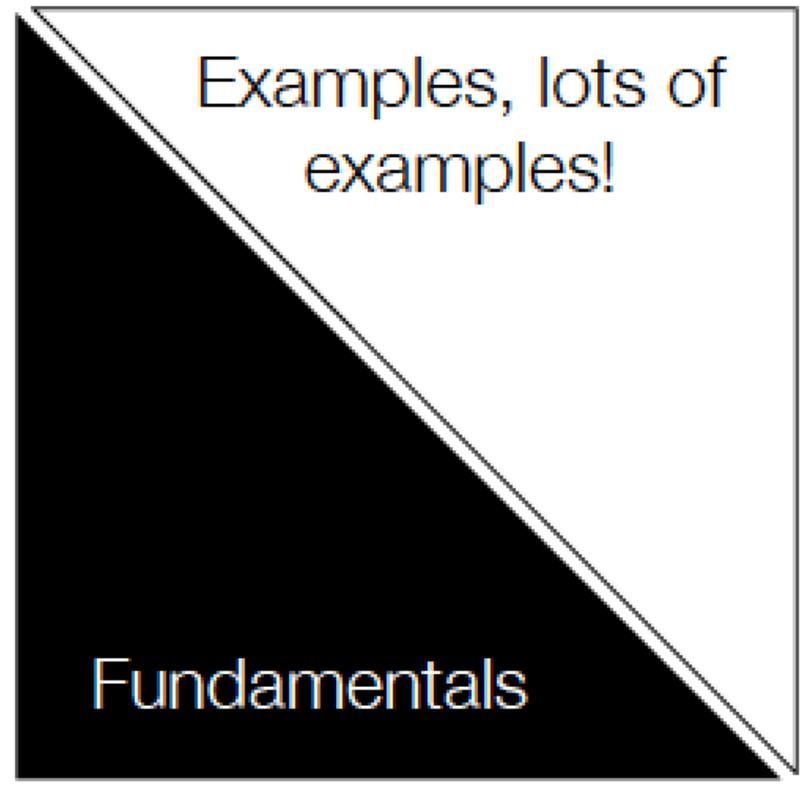


TAKE AWAY MESSAGE…

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!

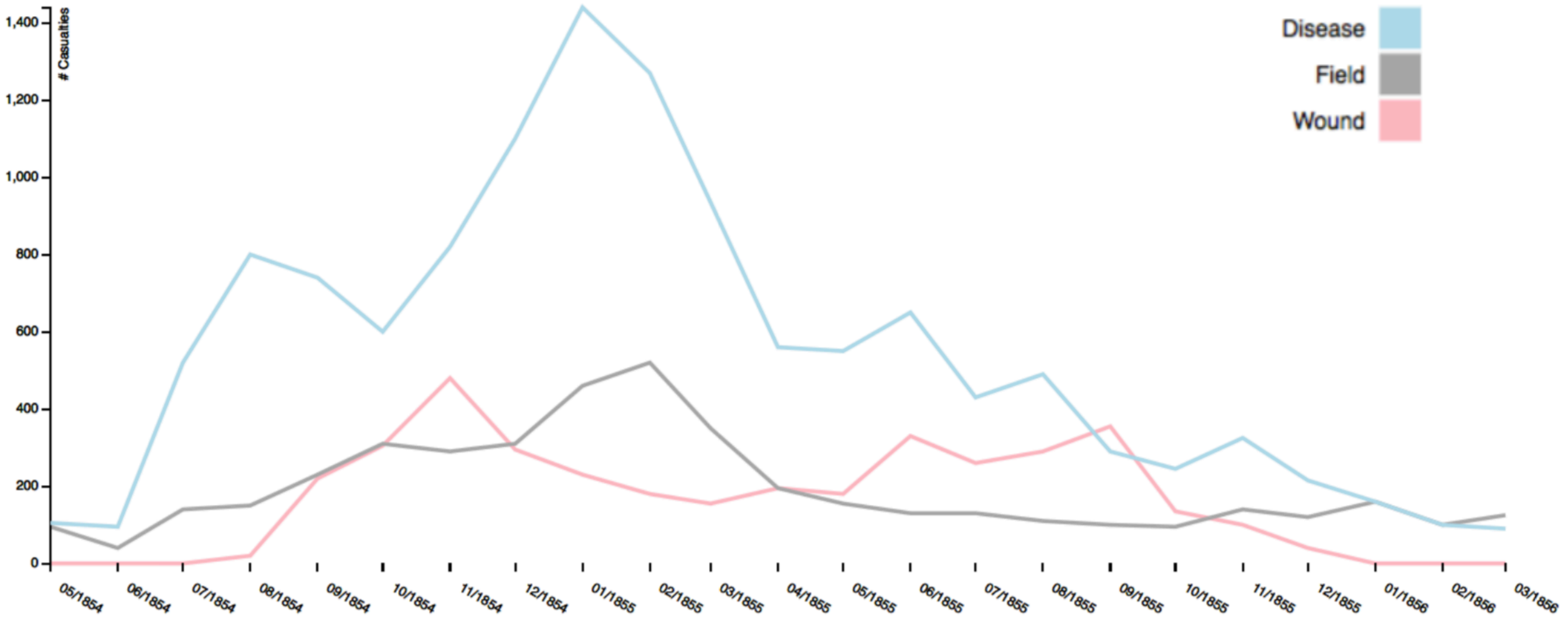


British Casualties in the Crimean War

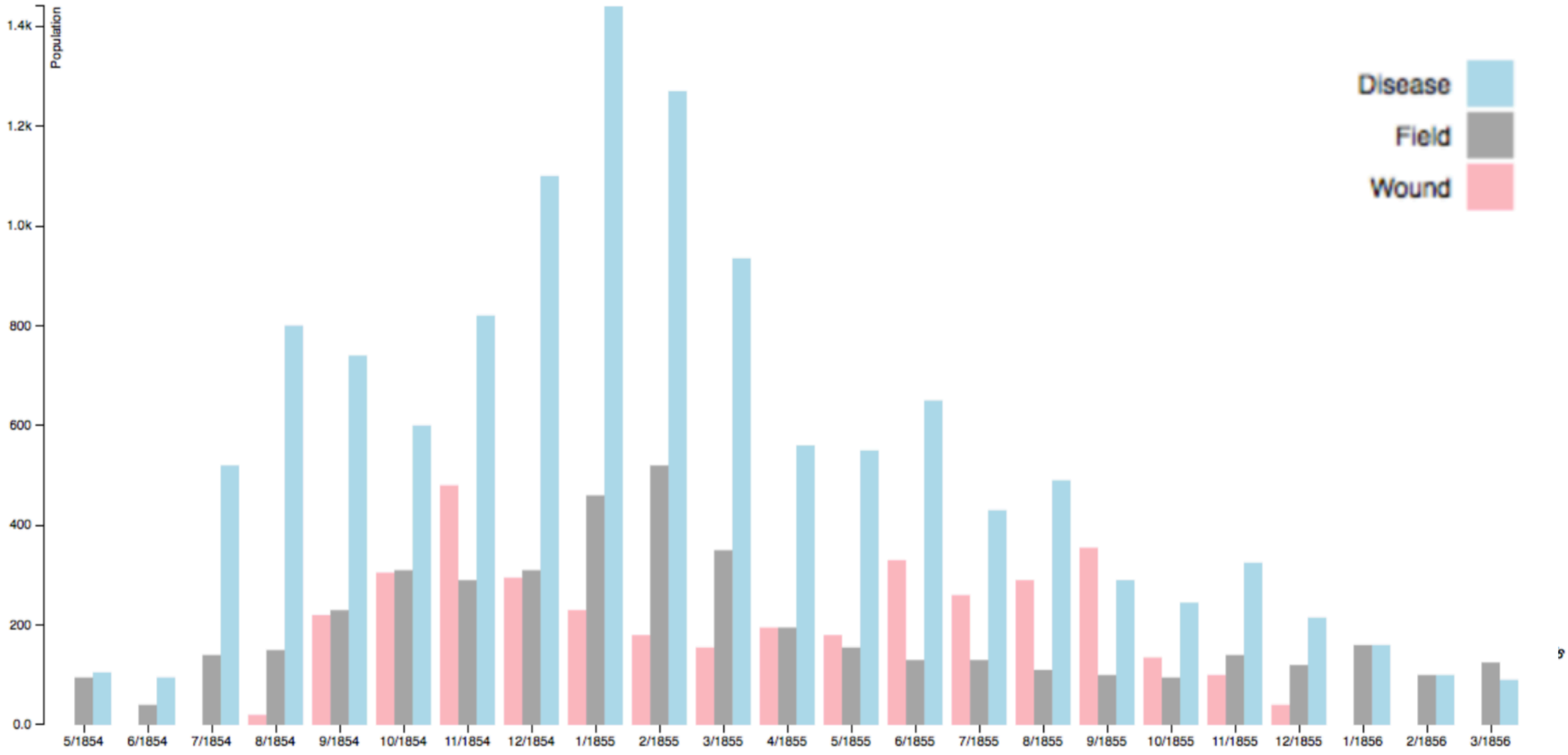
Data	Wound	Field	Disease
05/1854	0	95	105
06/1854	0	40	95
07/1854	0	140	520
08/1854	20	150	800
09/1854	220	230	740
10/1854	305	310	600
11/1854	480	290	820
12/1854	295	310	1100
01/1855	230	460	1440
02/1855	180	520	1270
03/1855	155	350	935
04/1855	195	195	560
05/1855	180	155	550
06/1855	330	130	650
07/1855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

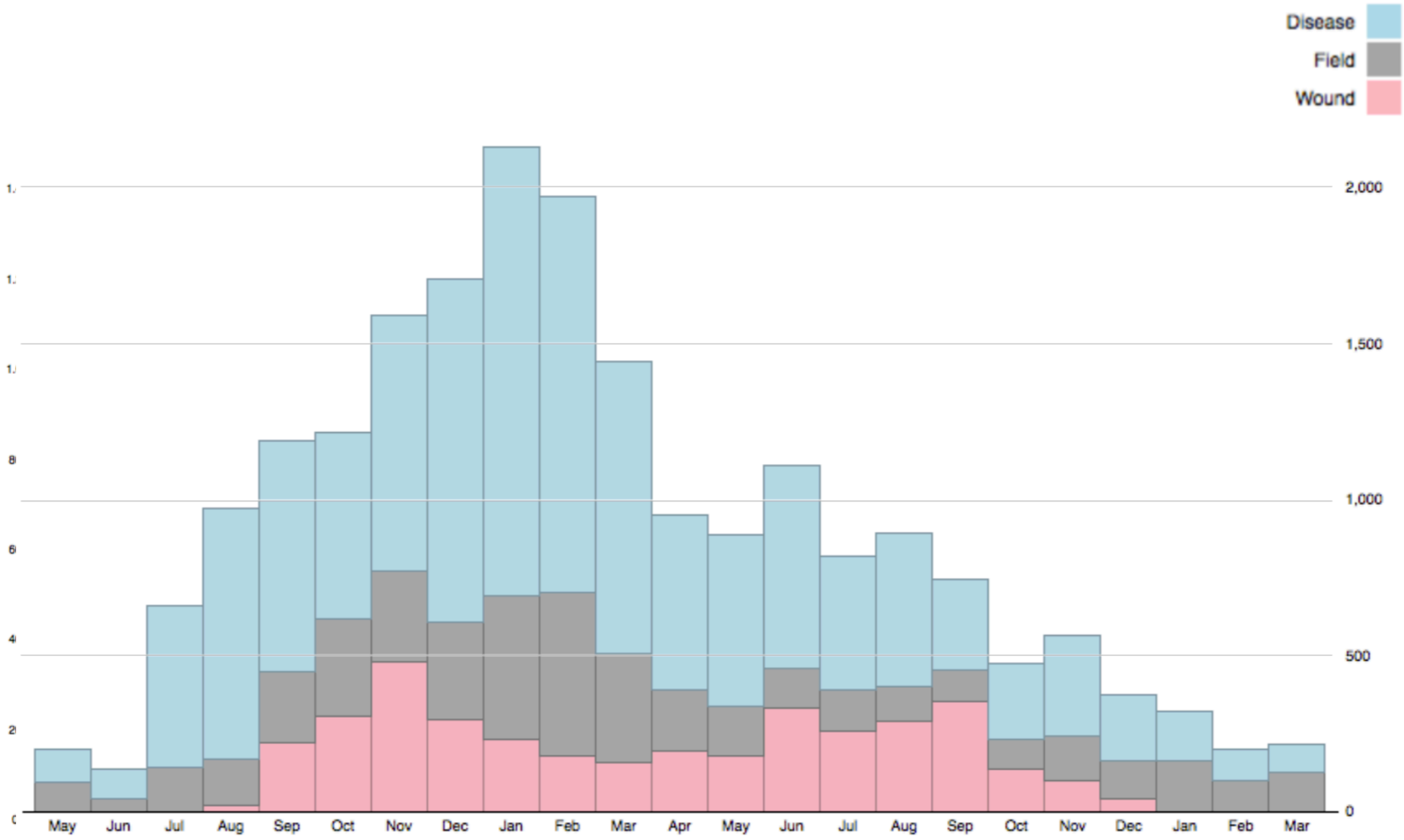
Month with highest casualty rates in the field?



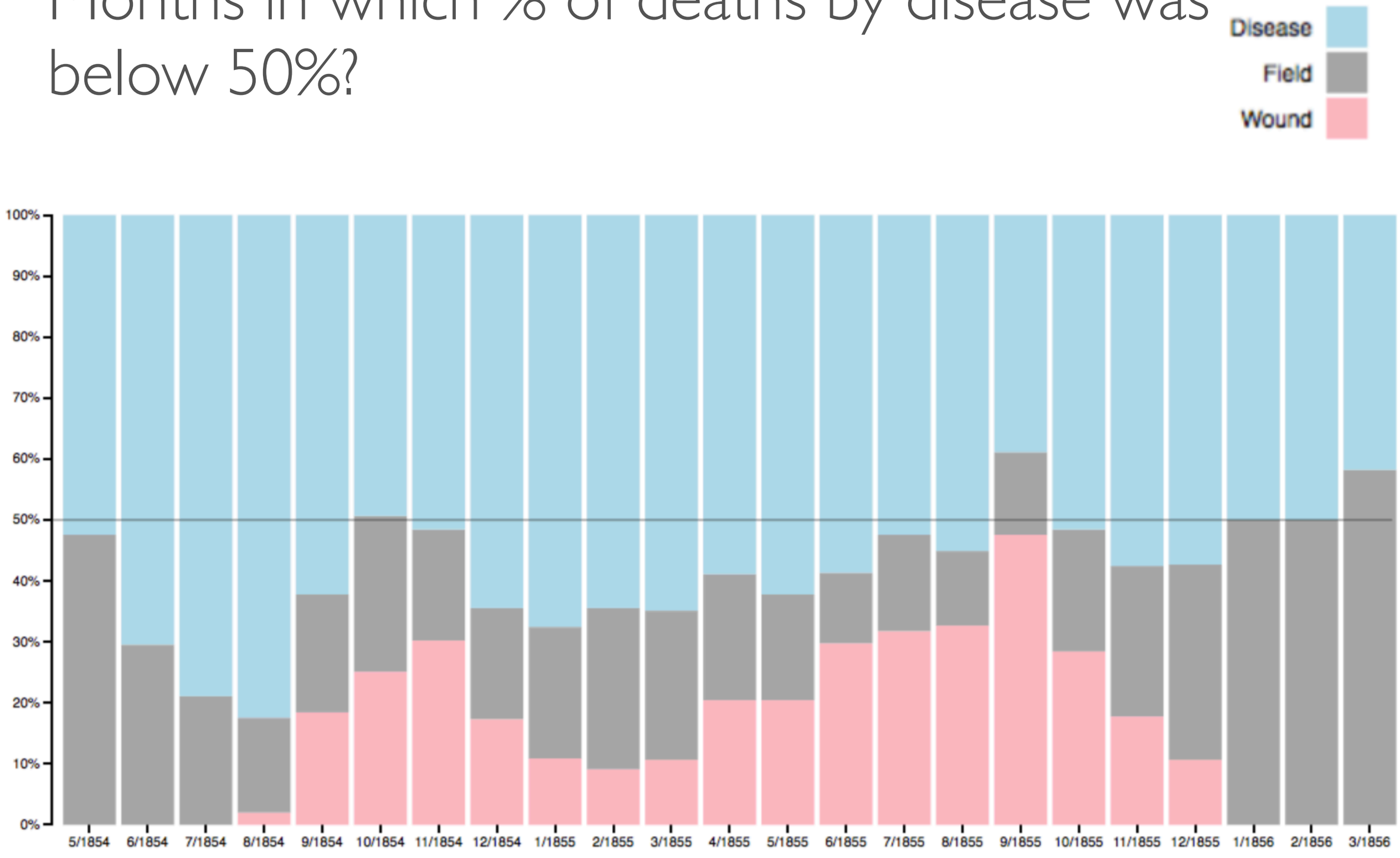
Months in which deaths by wound exceeds deaths in the field?



Month with highest total casualty rate?

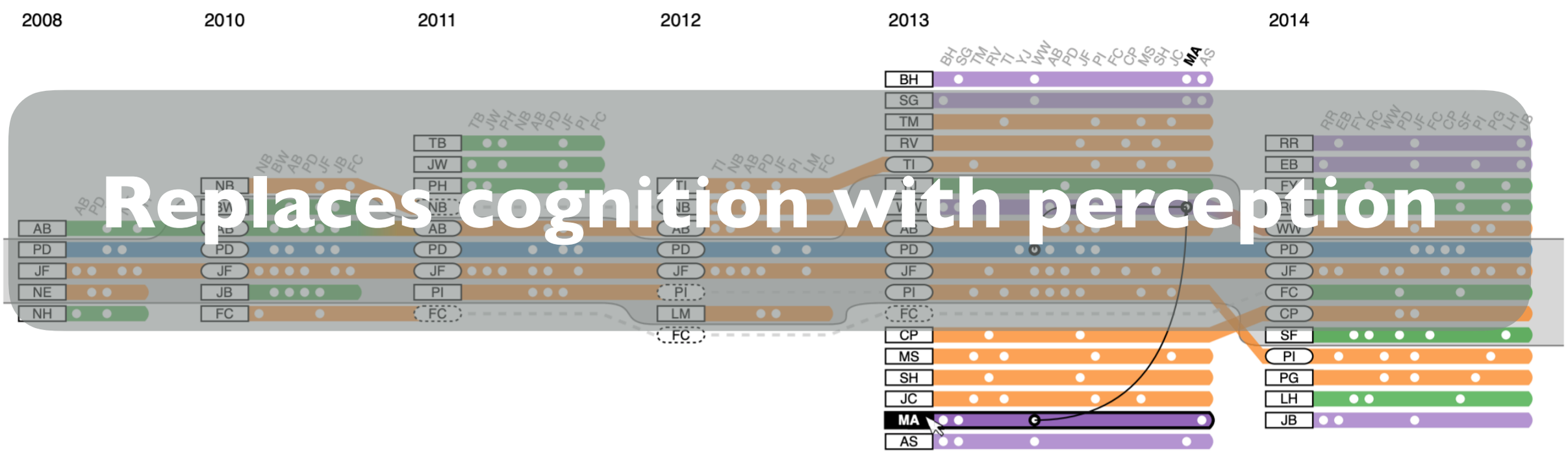


Months in which % of deaths by disease was below 50%?



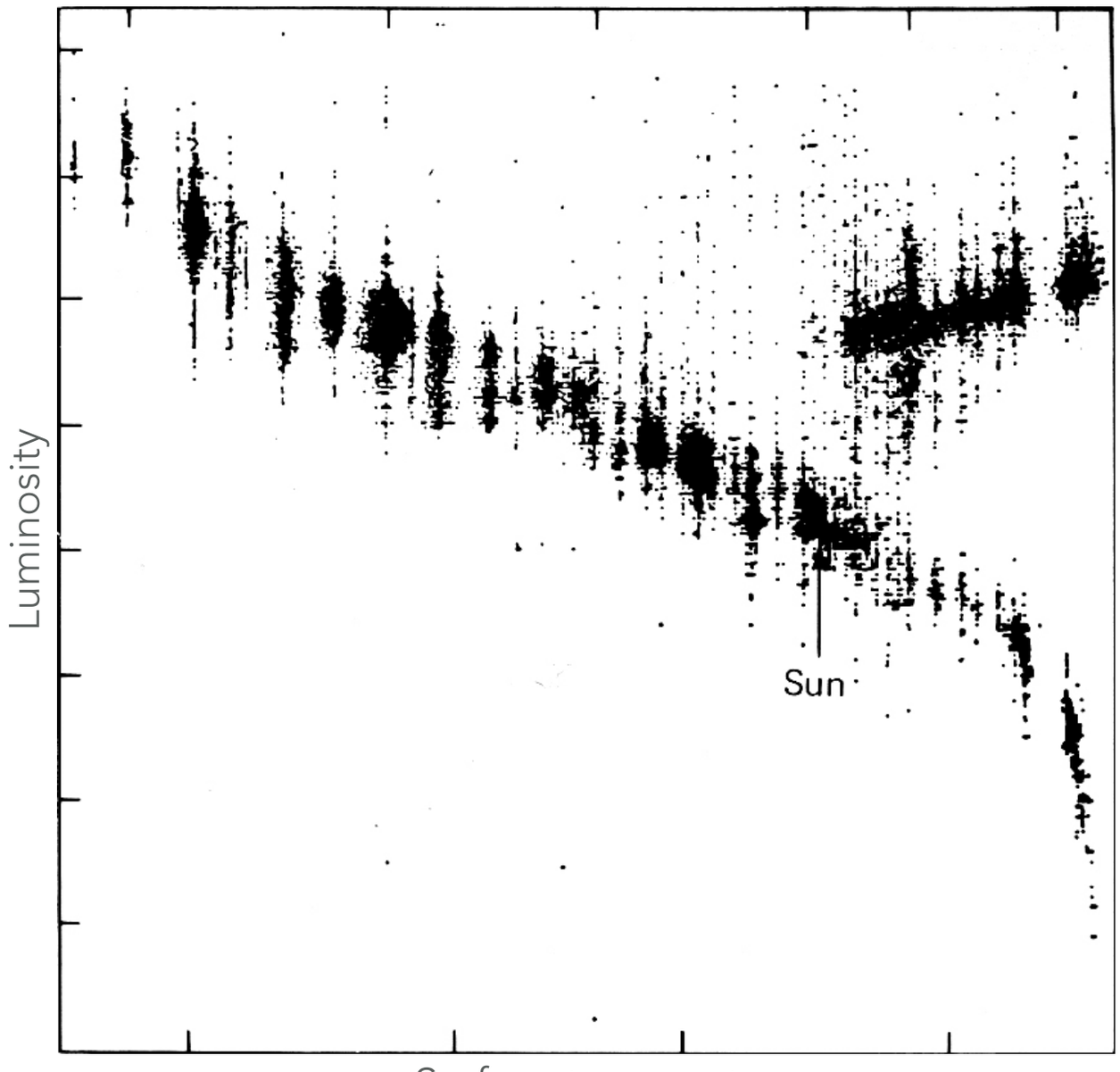
WHY VISUALIZATION?

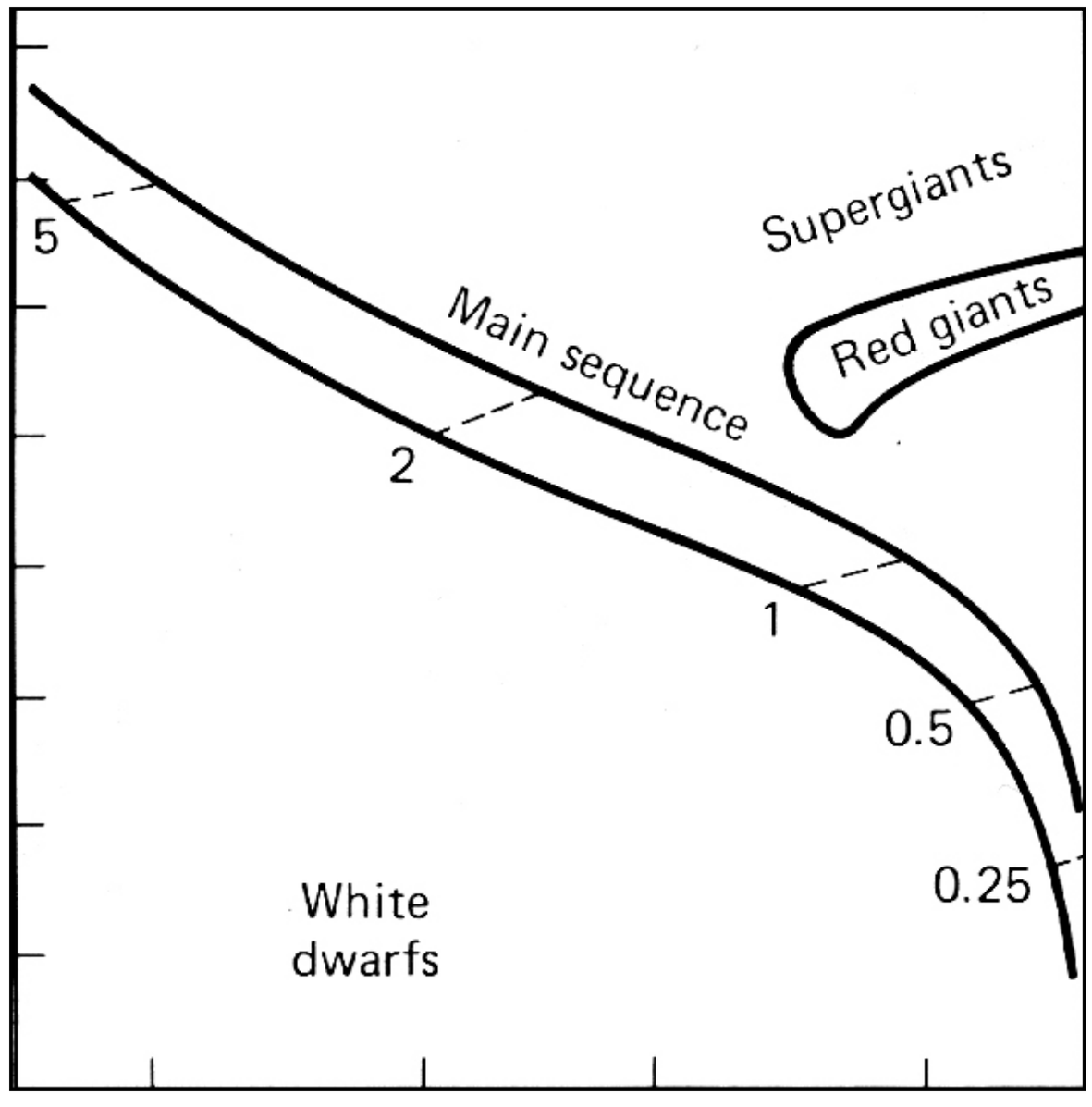
Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.



https://jeffjianzhao.bitbucket.io/webapp/EgoLines/egolines-demo.html

PERCEPTION AND COGNITION

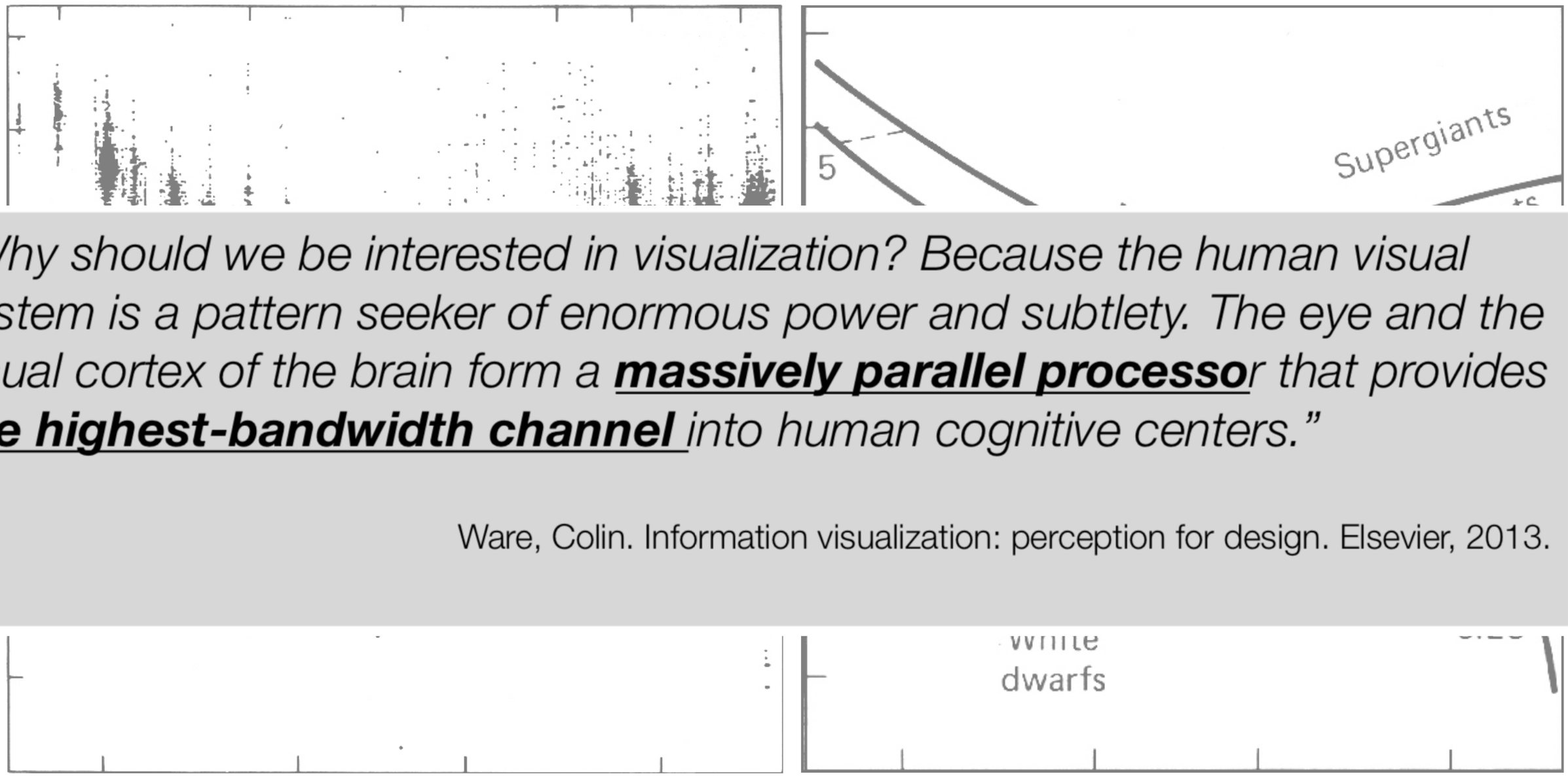




Hertzsprung Russell Diagram and its interpretation

[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

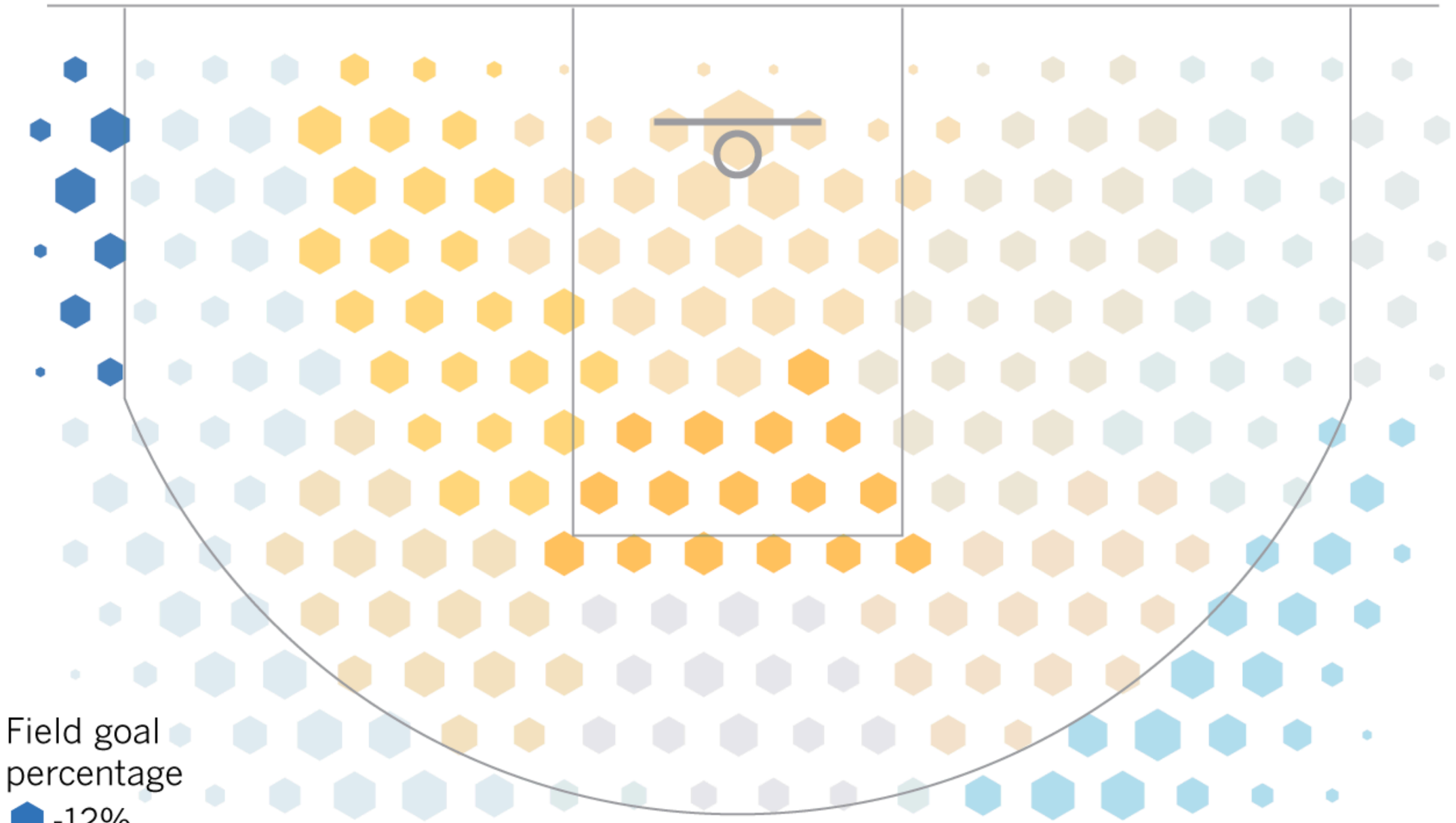
[Source: J.D. Fekete et al. "The Value of Information Visualization", 2008]

1424164963575984759217659684748917284822859588198294509685048506958476121240440746748989851714959691245676596080208606083654164964575906439804792485769607812859607999187128452681014959691245677818742416496457576596081495969124567012859607991649645751278799187128452984969122359164964575958819825096357659608059600

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

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

Kobe Bryant vs. league average (1996-2016)



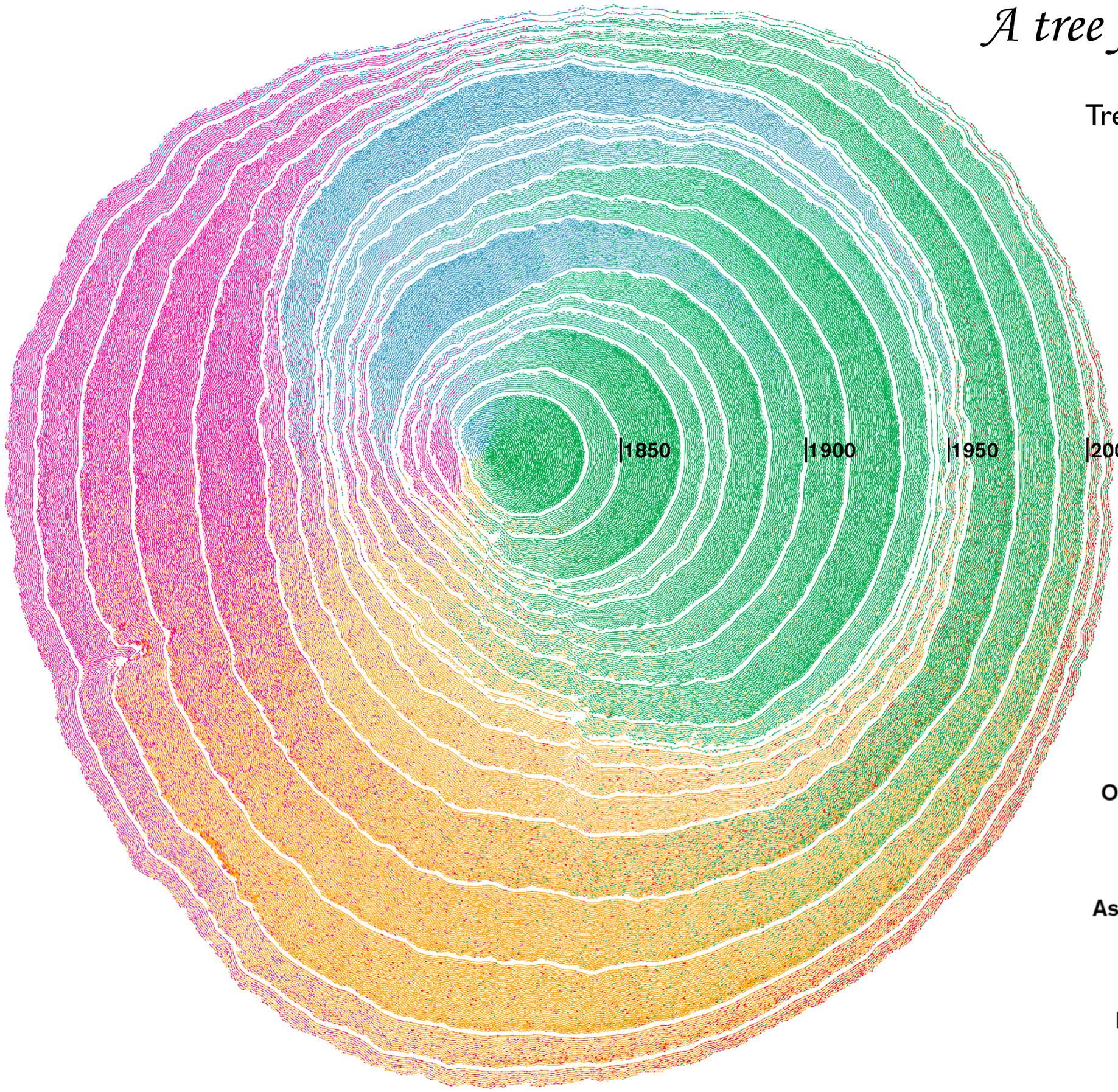
Field goal percentage
-12%
-8%
-4%
average
+4%
+8%
+12%

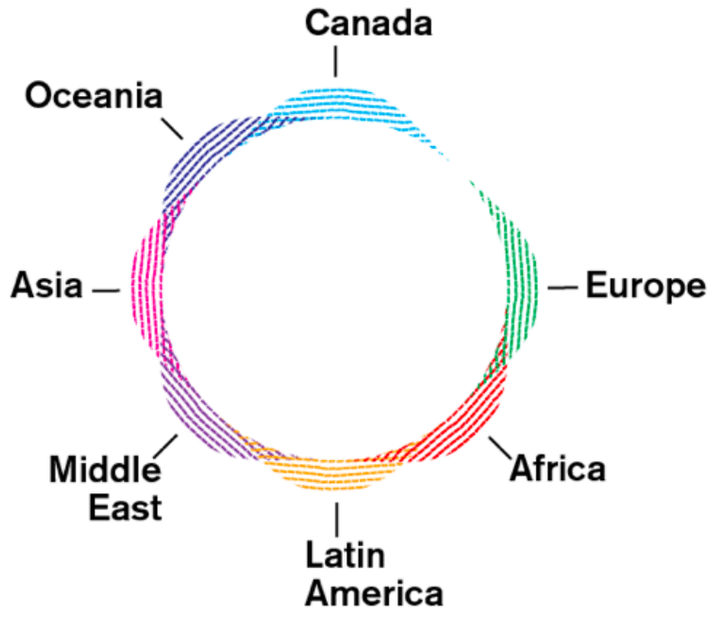
Source: stats.nba.com, toddwschneider.com

Joe Fox / @latimesgraphics

A tree for U.S. immigration

Tree rings showing immigration for
1830-2016
Each dot corresponds to
100 immigrants



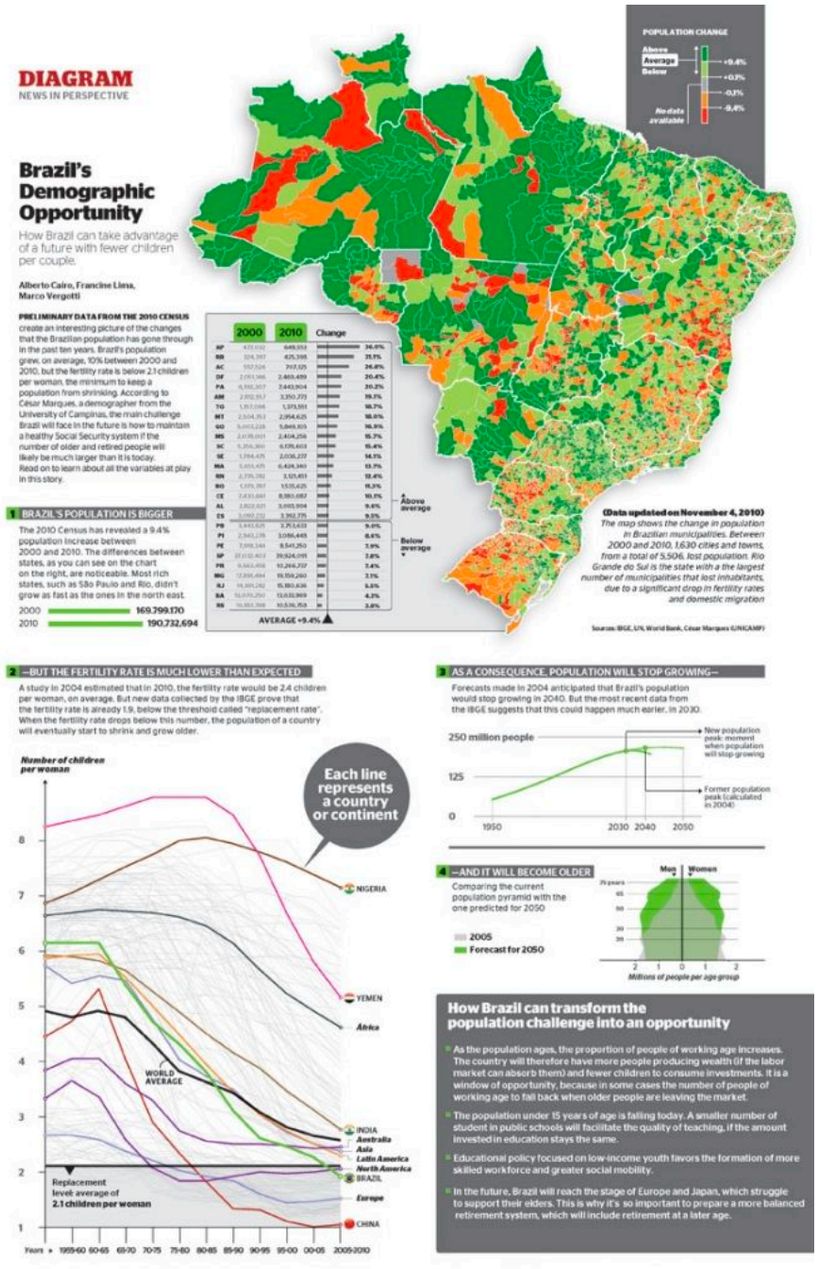


Pedro Cruz. Simulated Dendrochronology of U.S. immigration

MAIN USES OF DATA VISUALIZATION

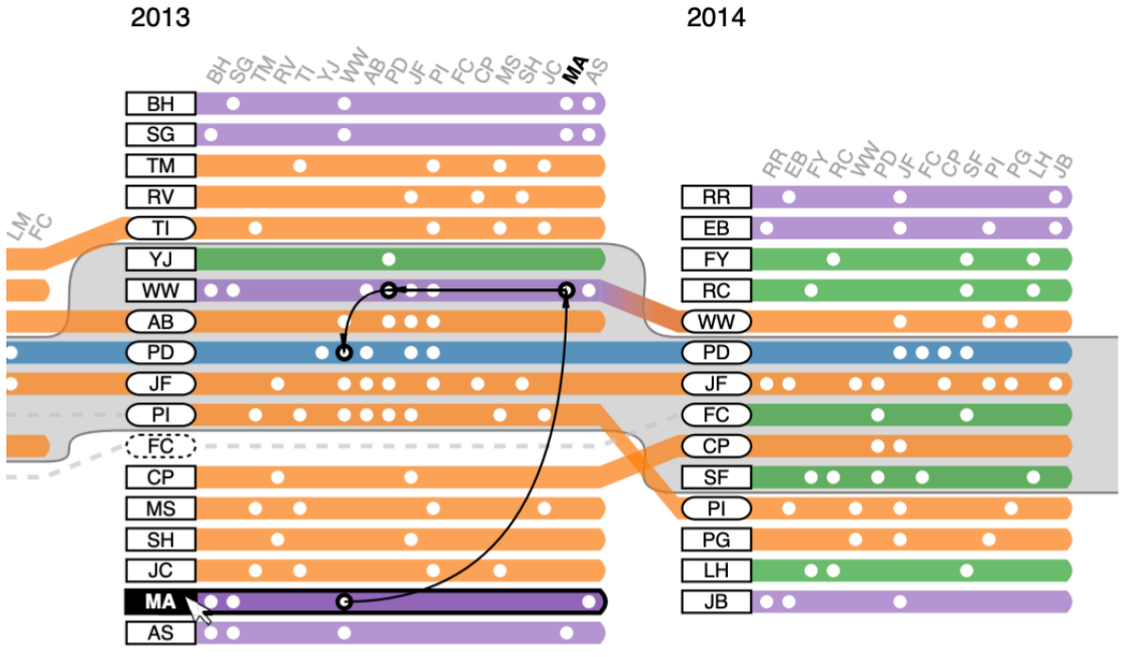
EXPLANATORY

Storytelling



EXPLORATORY

Acquiring insights



This is your brain on visualization….

PERCEPTION FUNDAMENTALS

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

Data  **Graphics**

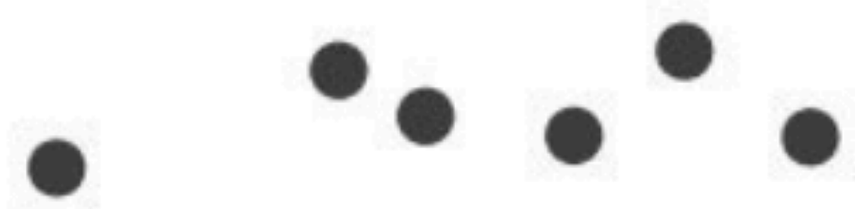
Which representations can I use to encode data?

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

MARKS

Geometric primitives

Points



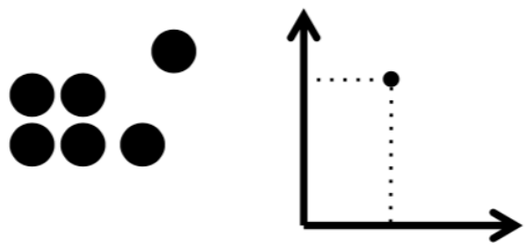
Lines

Areas

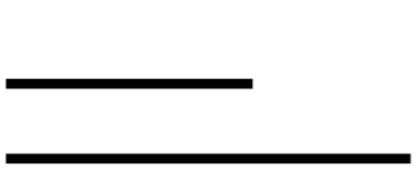
CHANNELS (VISUAL VARIABLES)

Control appearance of marks

Position



Size- Length

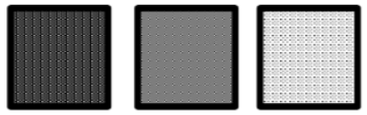


Size- Area

Size- Volume

Format

Texture



Orientation/Direction

Angle

Color

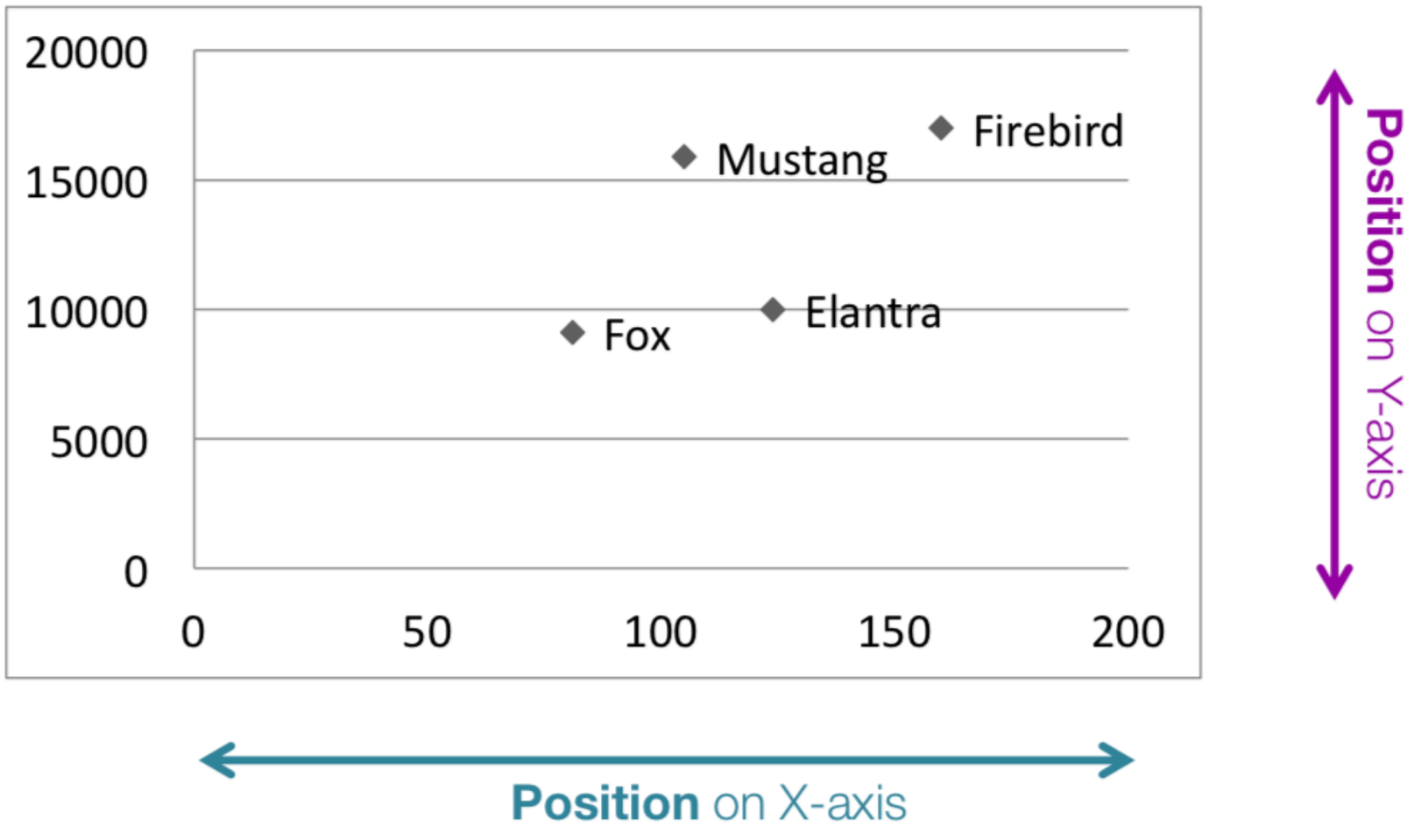
Color - Hue

Color - Saturation

Color - *Luminance*

MAPPING DATA TO VISUAL VARIABLES

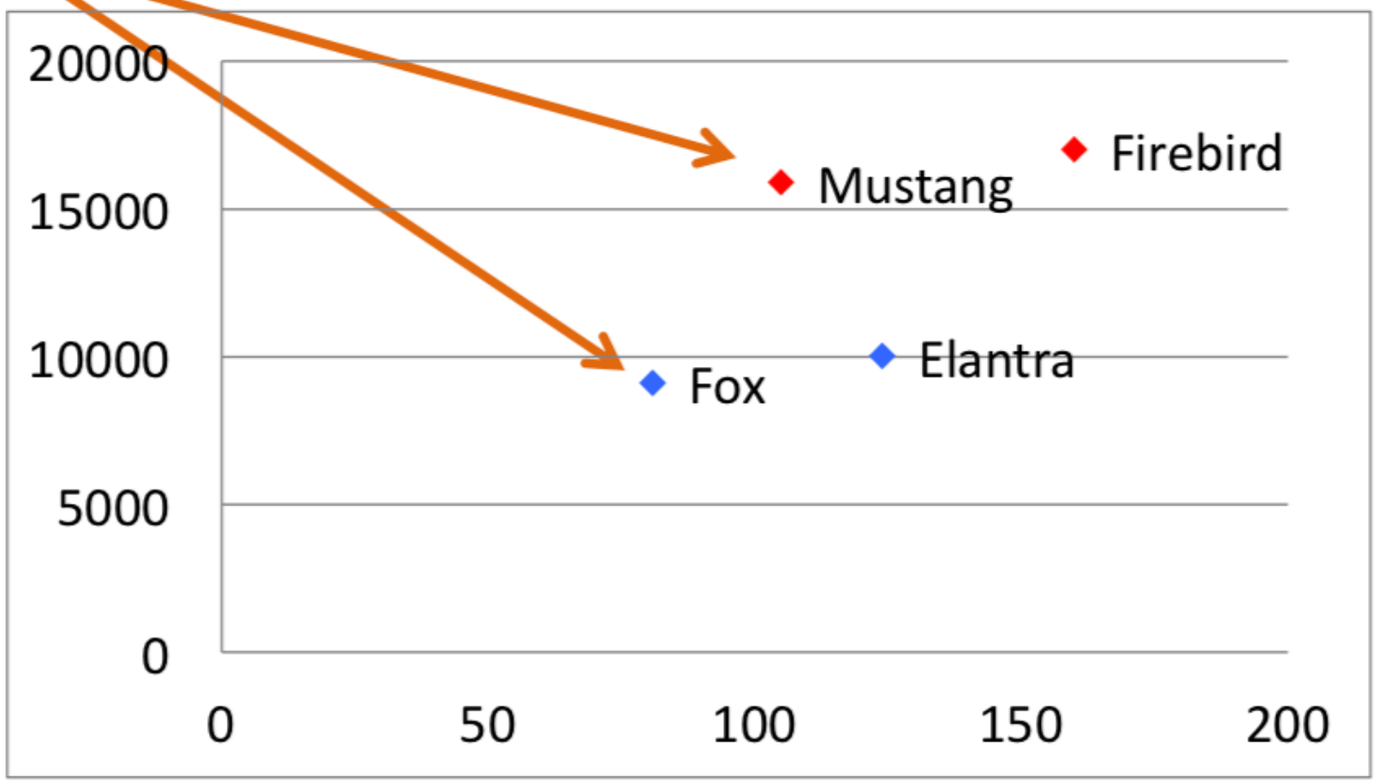
Model	Type	HP	Price	*Number Airbags*
Volkswagen Fox	Compact	81	9100	0 (no airbag)
Hyundai Elantra	Compact	124	10000	0 (no airbag)
Pontiac Firebird	Sport	160	17000	2 (driv. & pass.)
Ford Mustang	Sport	105	15900	1 (driver)



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)

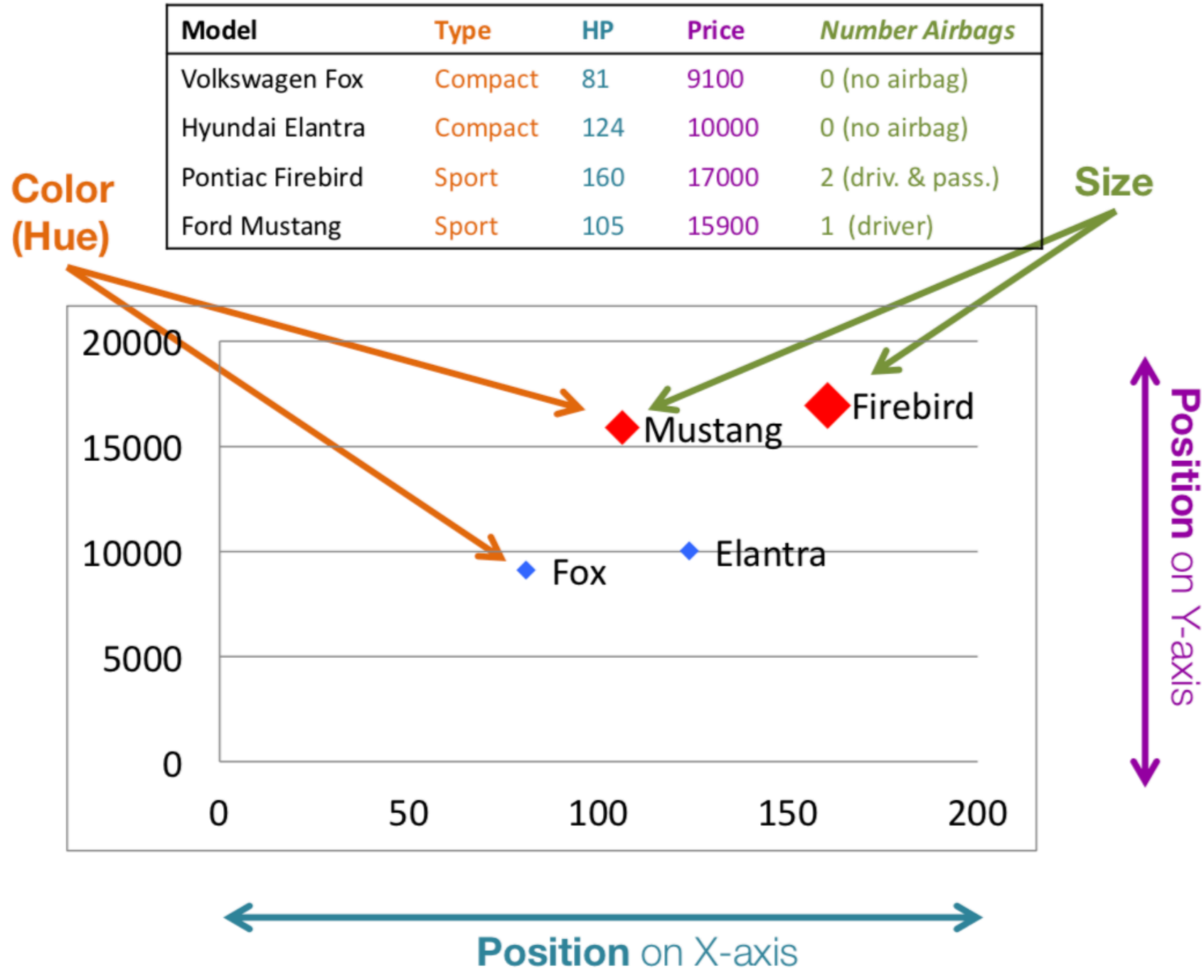


Position on Y-axis

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?

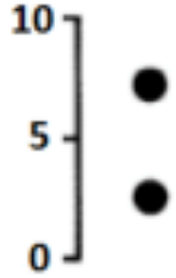
Effectiveness principle:
choosing the best representation to your data

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

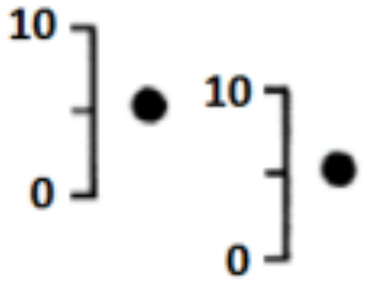
Most important attributes should be encoded using the most effective and noticeable channels.

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



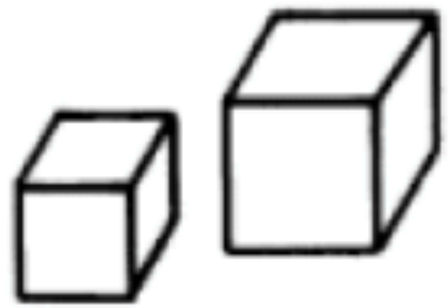
Length

Direction

Angle

Area

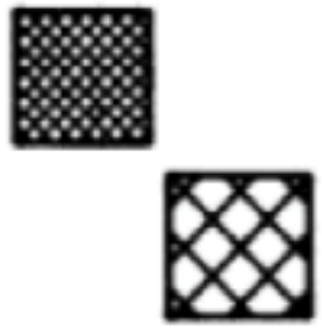
Volume



Curvature

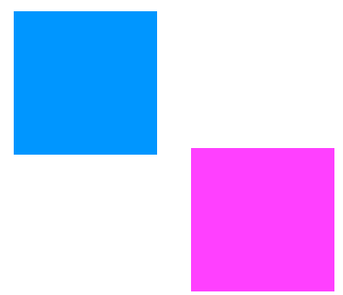
Shape

Shading



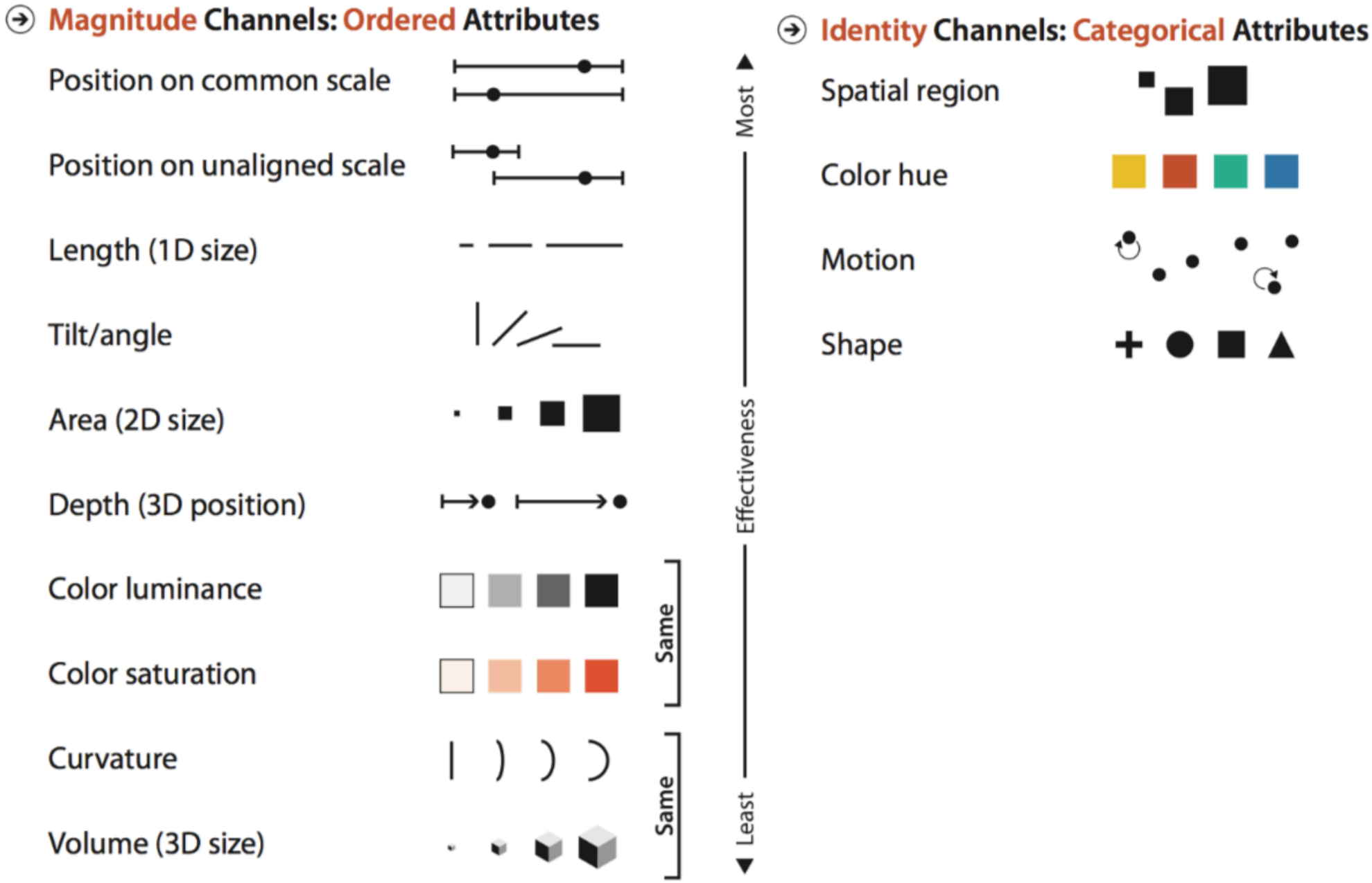
Color saturation

Hue



EFFECTIVENESS PRINCIPLE

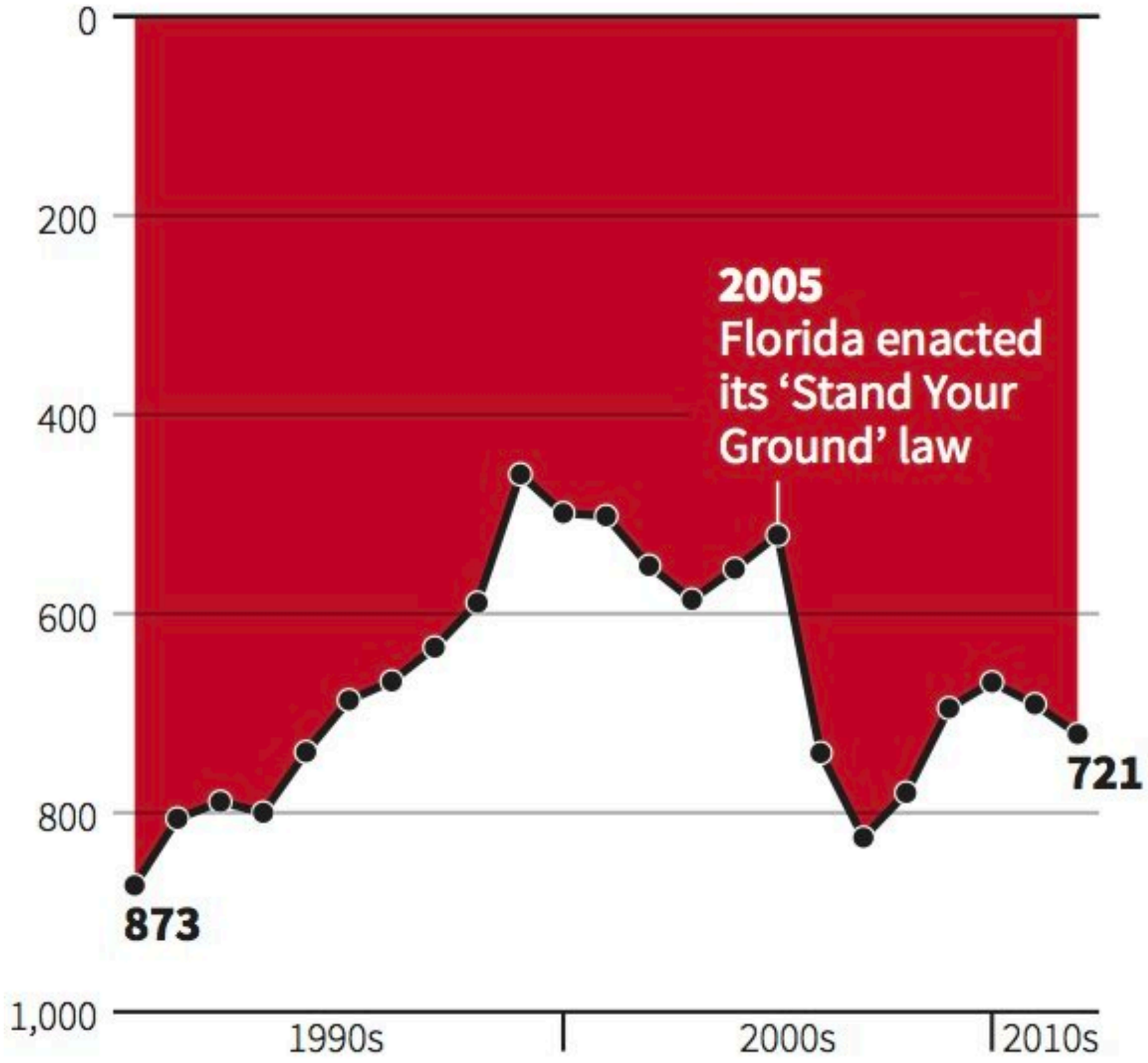
Some variables are perceptually better than others

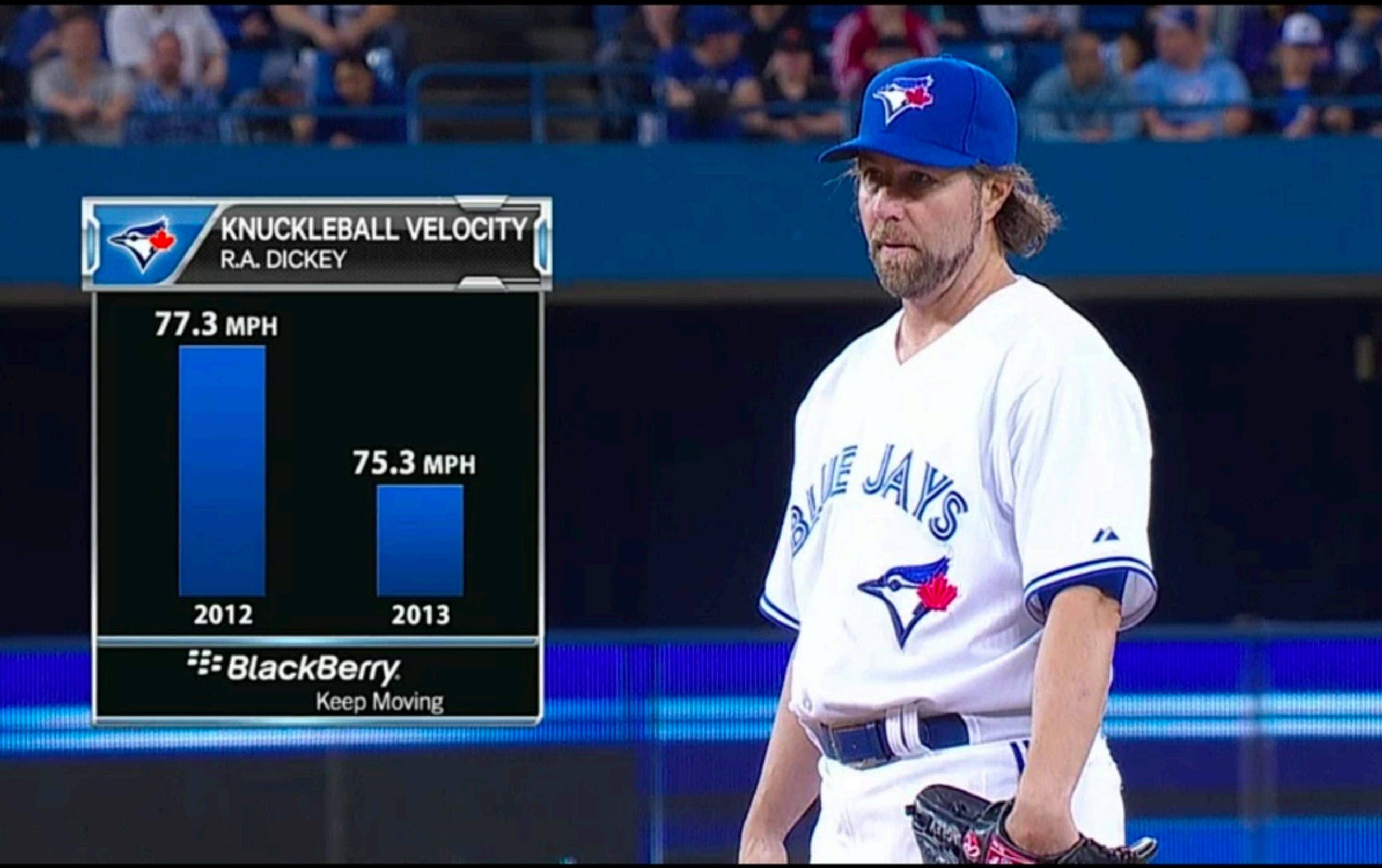


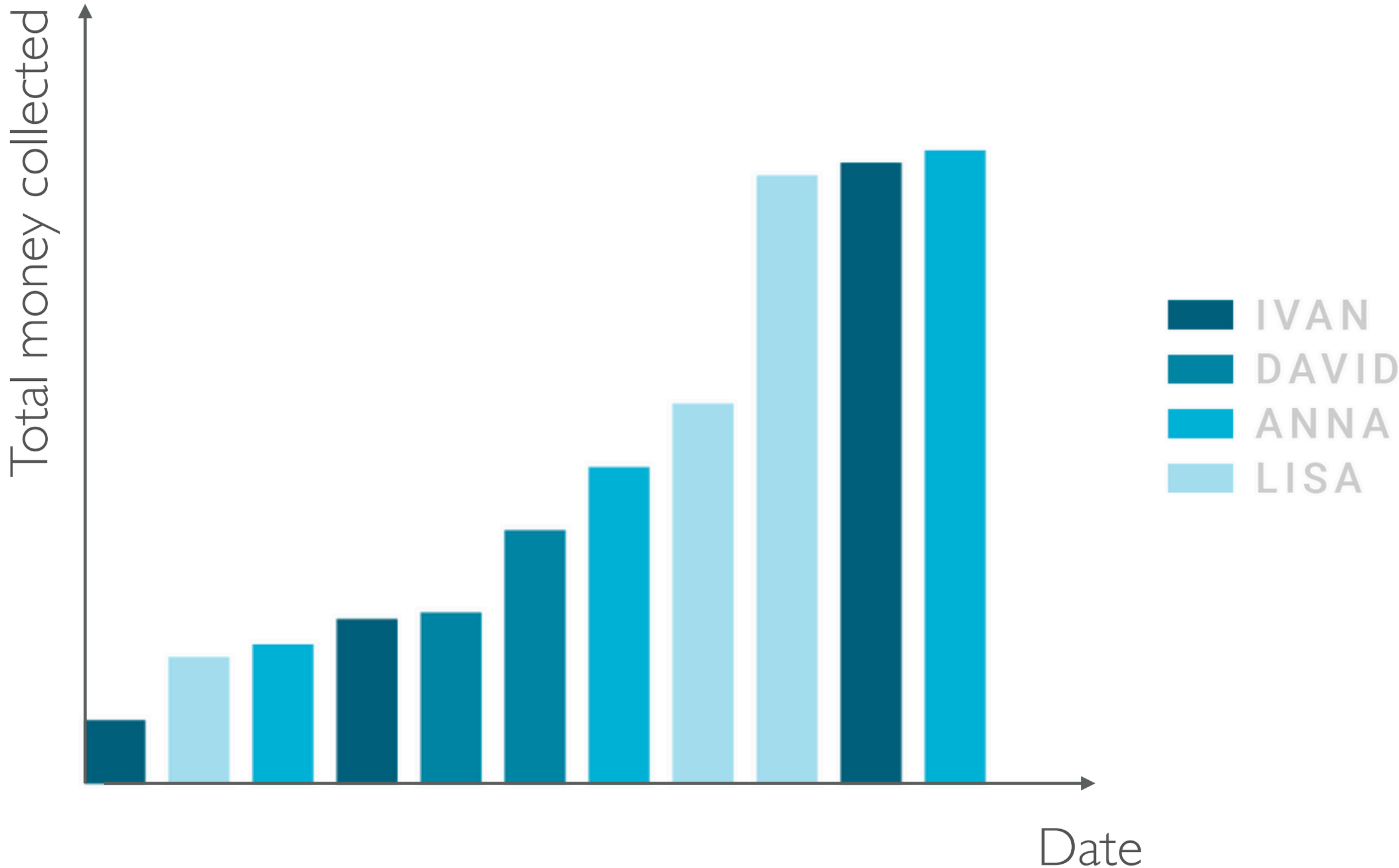
Munzner, Tamara. Visualization Analysis and Design. CRC Press, 2014.

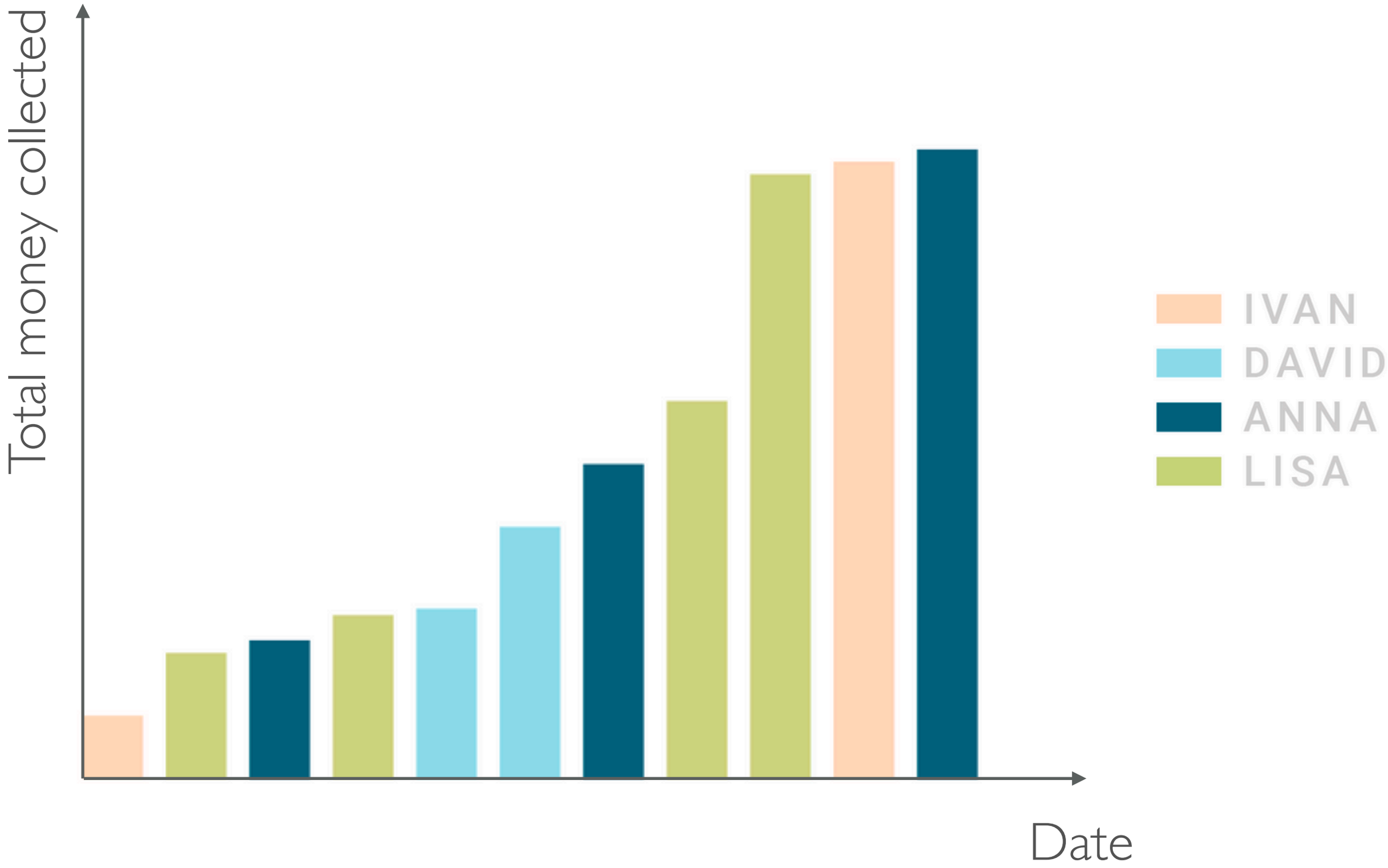
Gun deaths in Florida

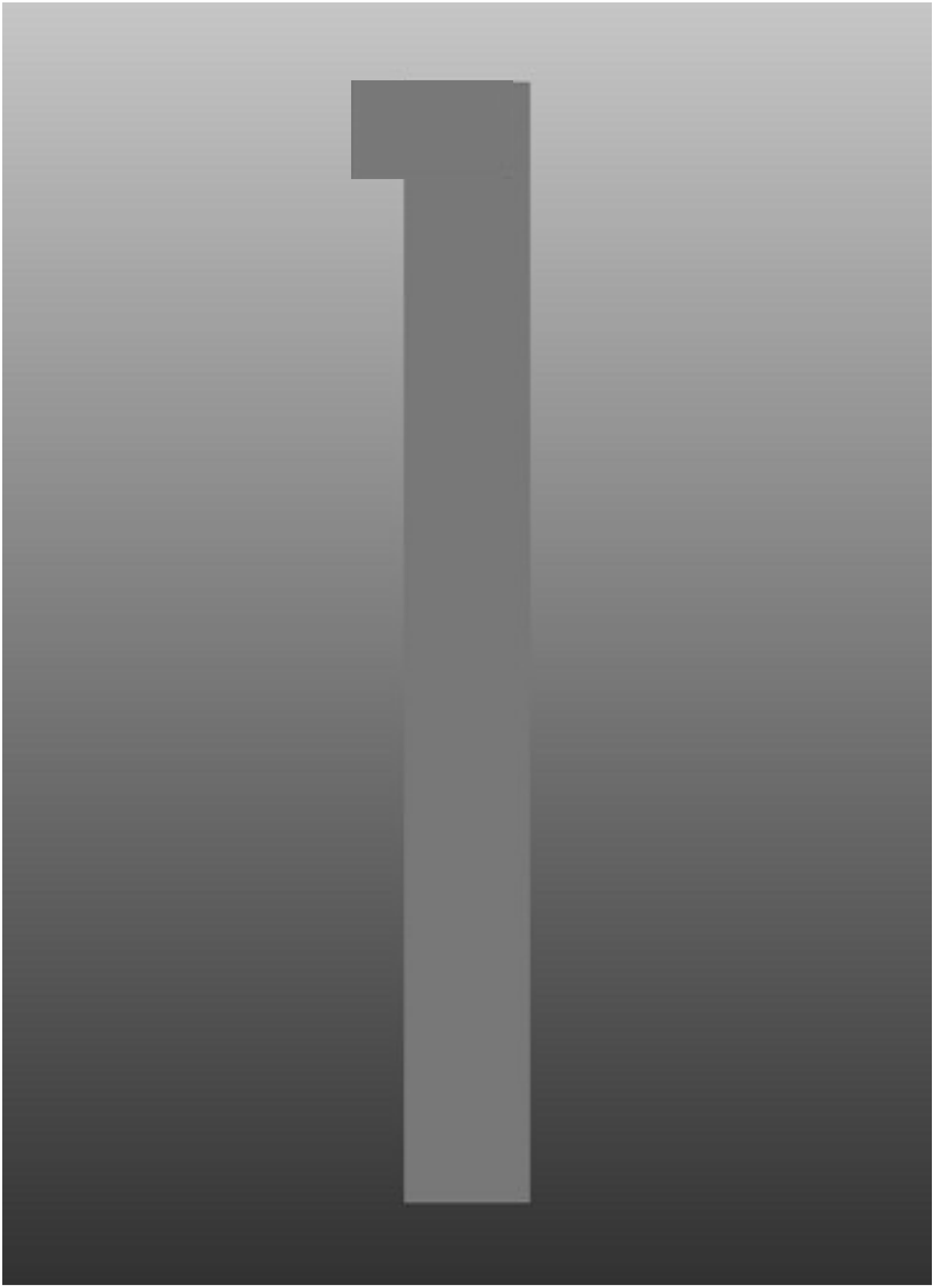
Number of murders committed using firearms




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


Total money collected
Date
IVAN
DAVID
ANNA
LISA


Total money collected
Date
IVAN
DAVID
ANNA
LISA

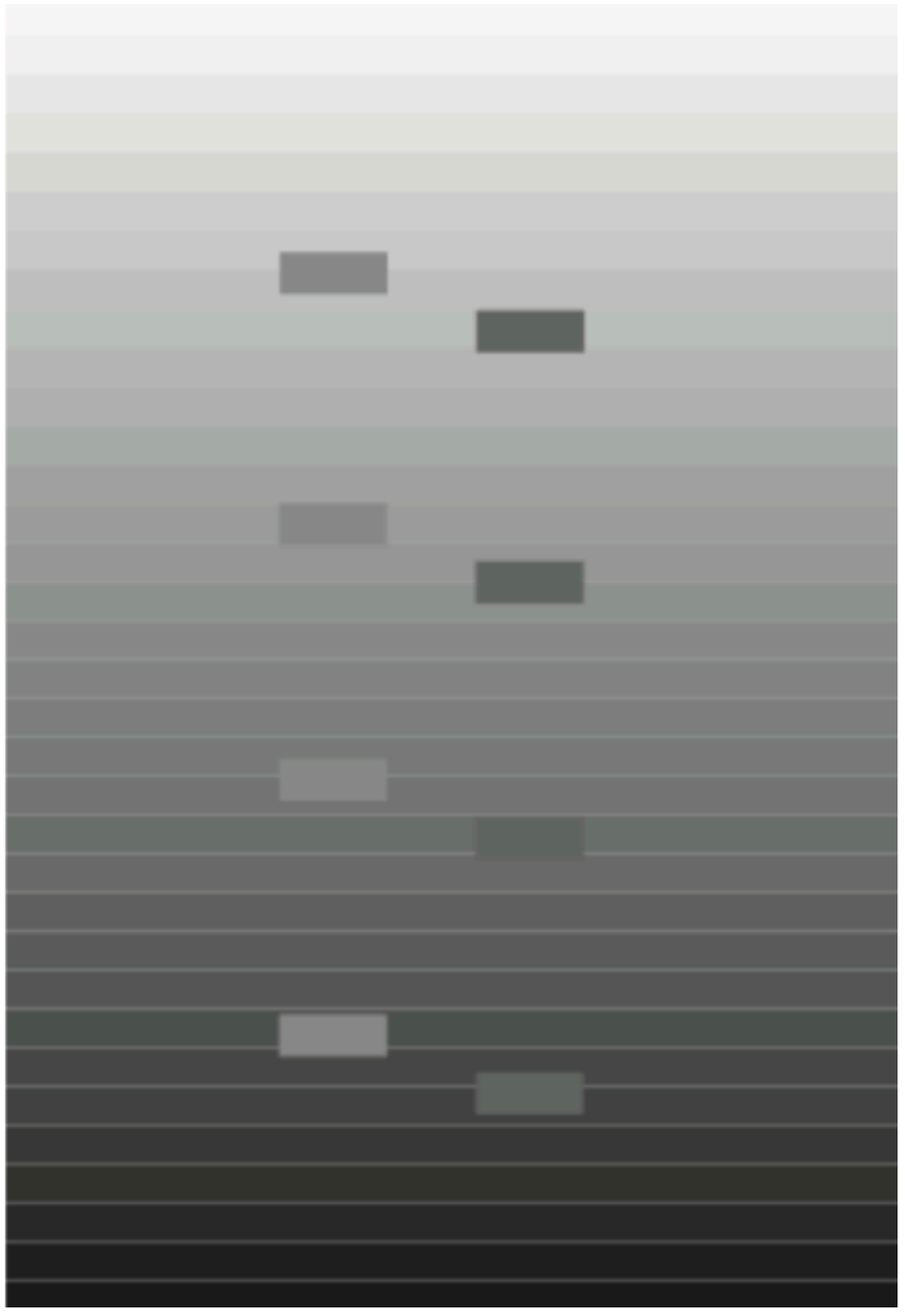


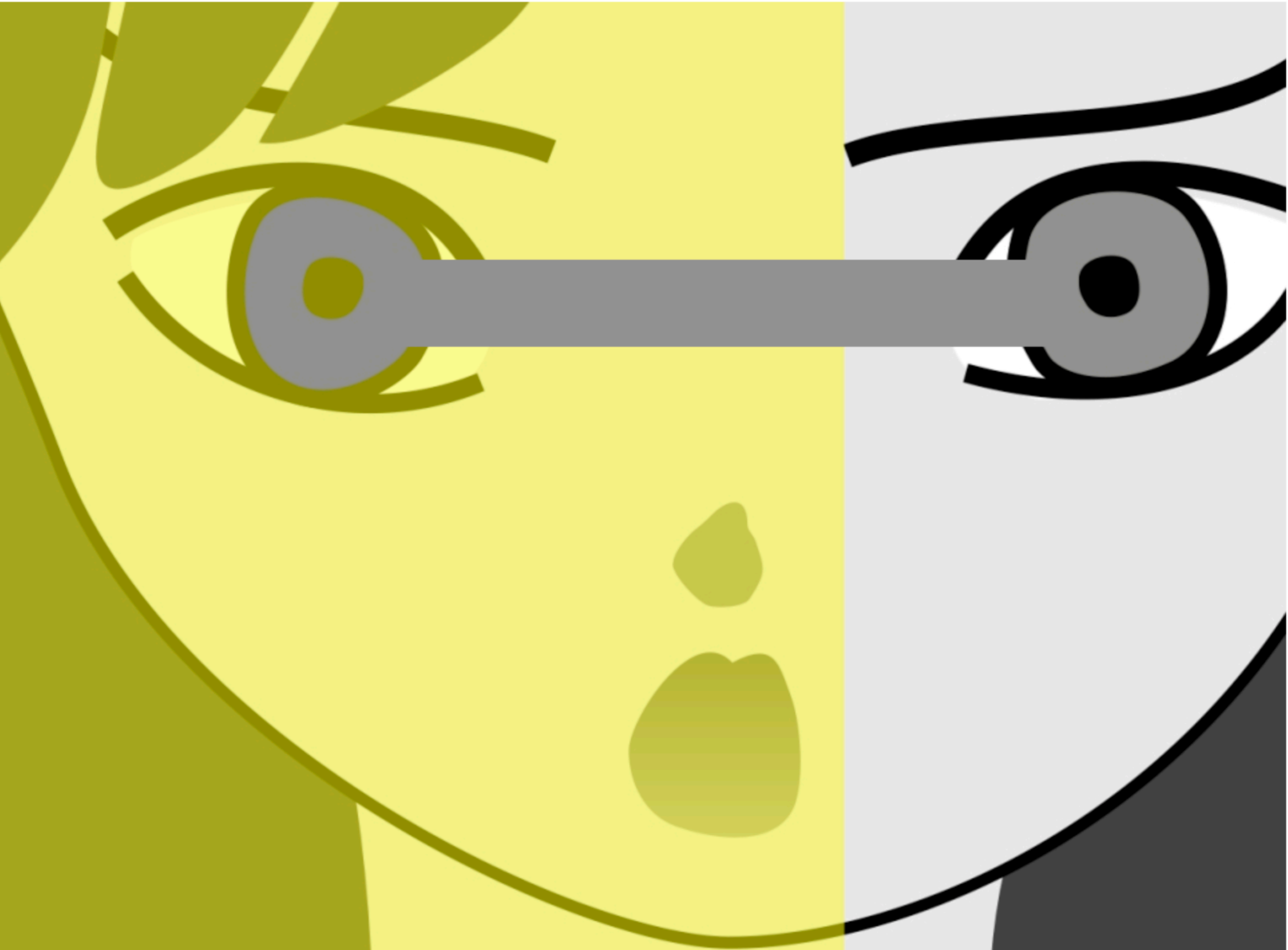
COLOR PERCEPTION IS LOCAL

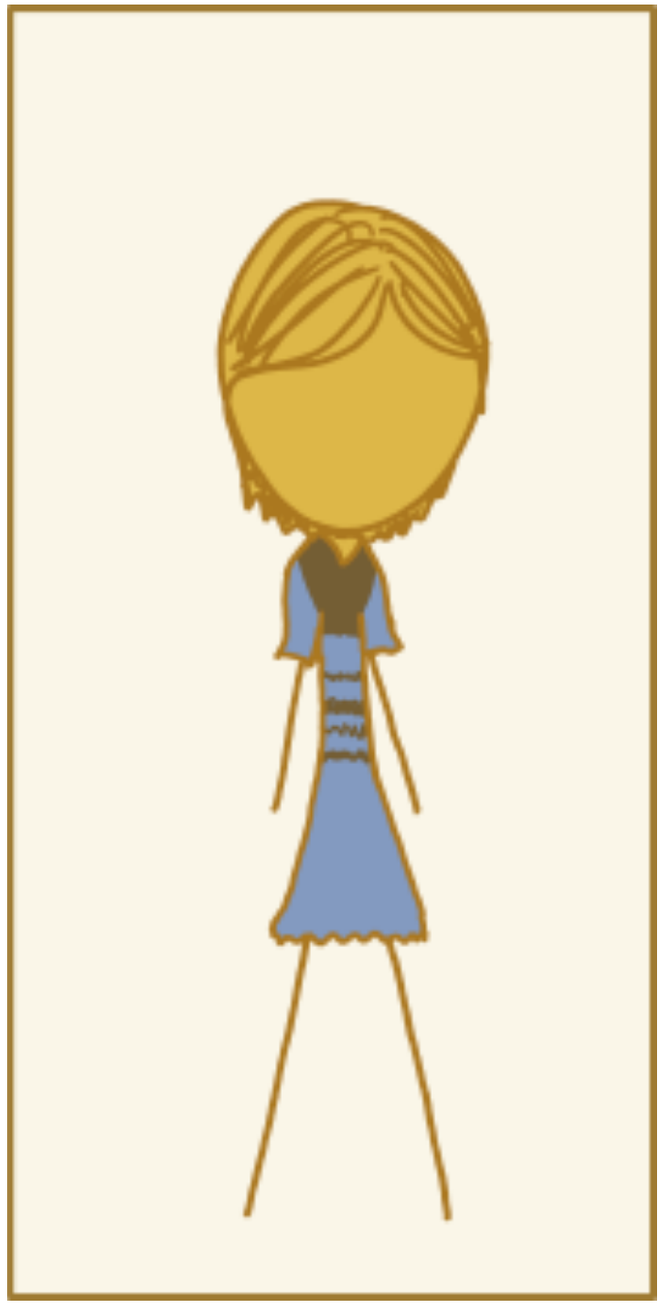
- Luminance perception is not absolute, but contrast relative ... and local!

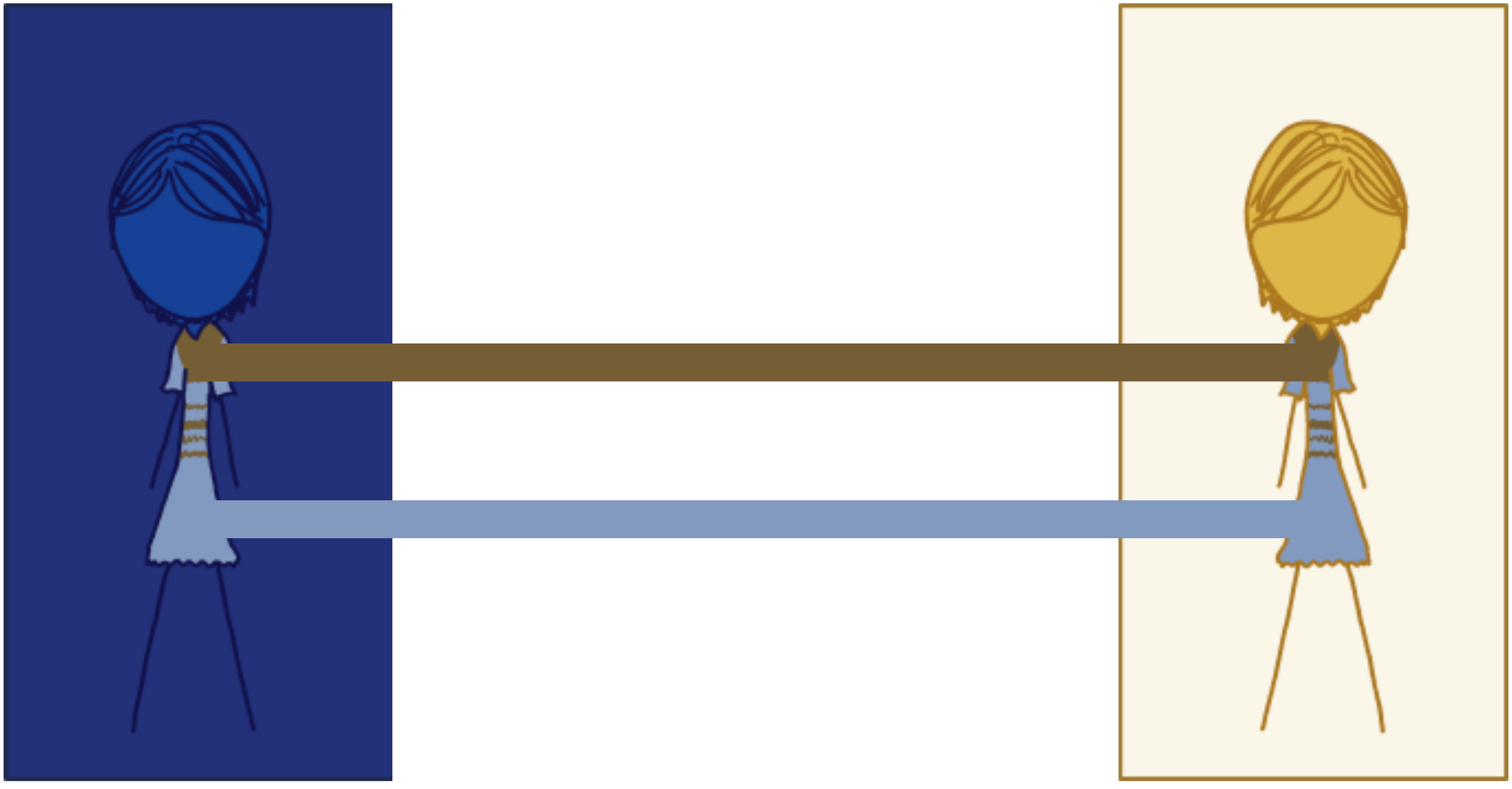
- Discern objects(foreground) from background

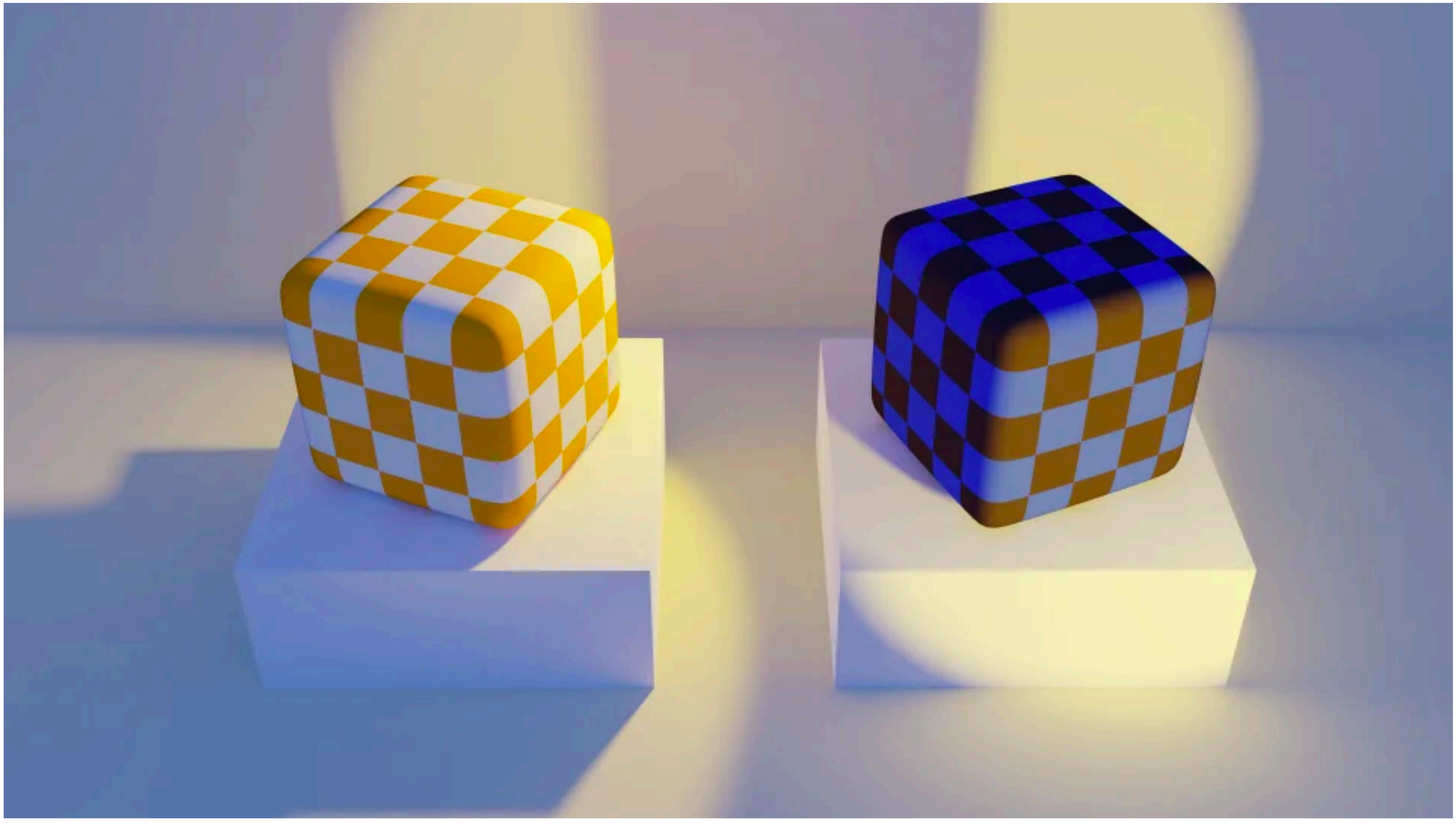
- Gray color scales are not recommended for more than 5 segments... too "imprecise" to convey quantitative data





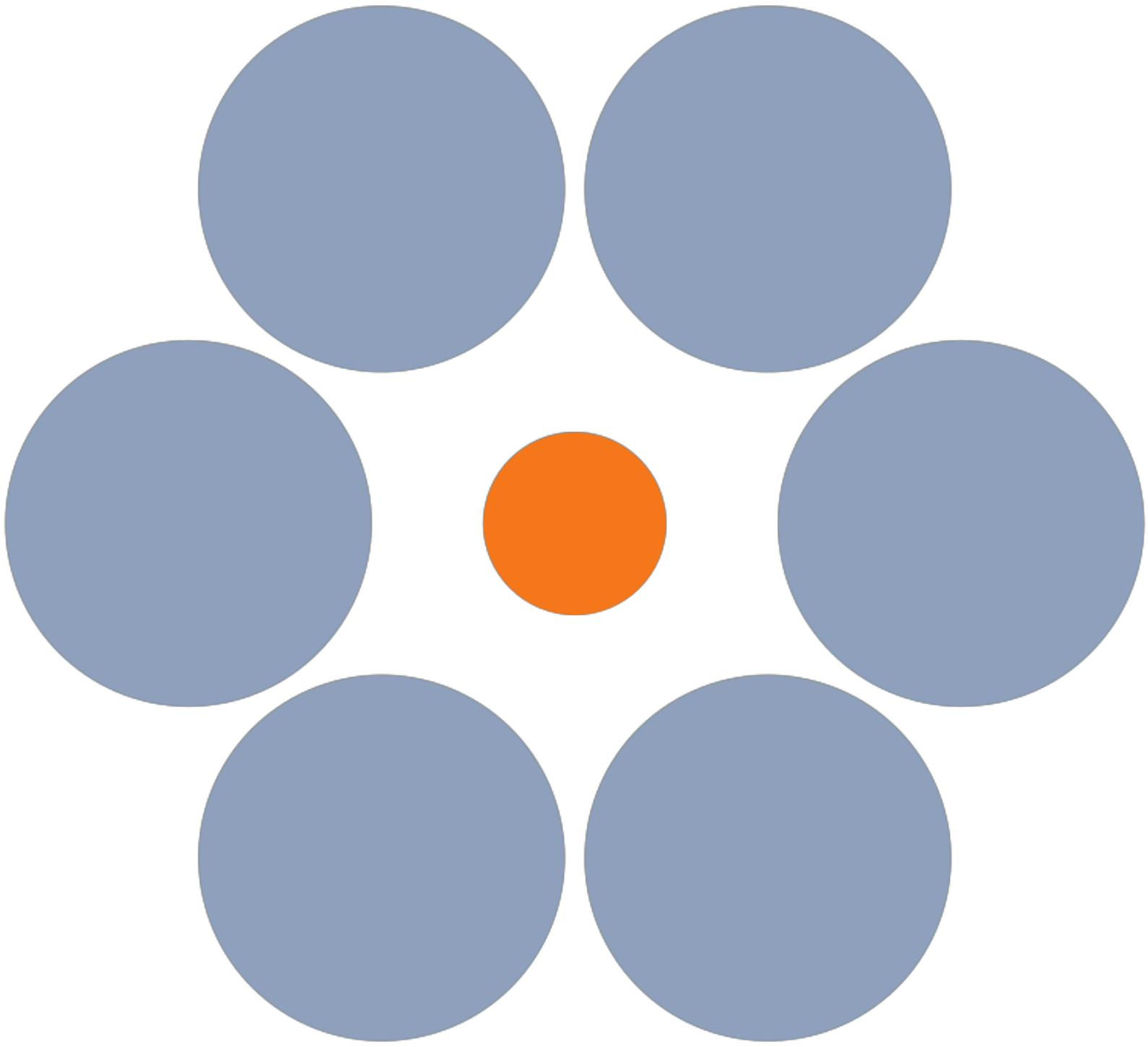


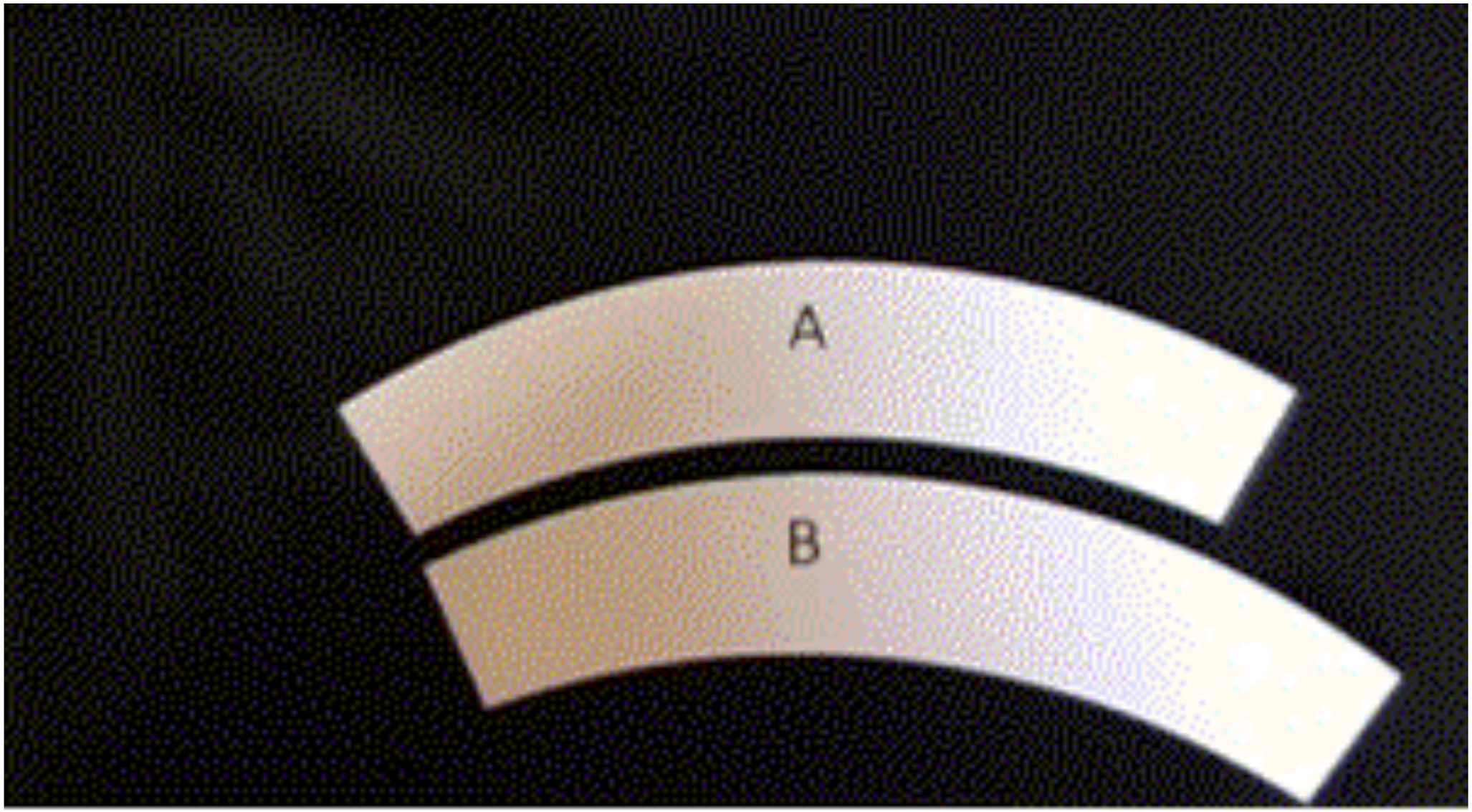


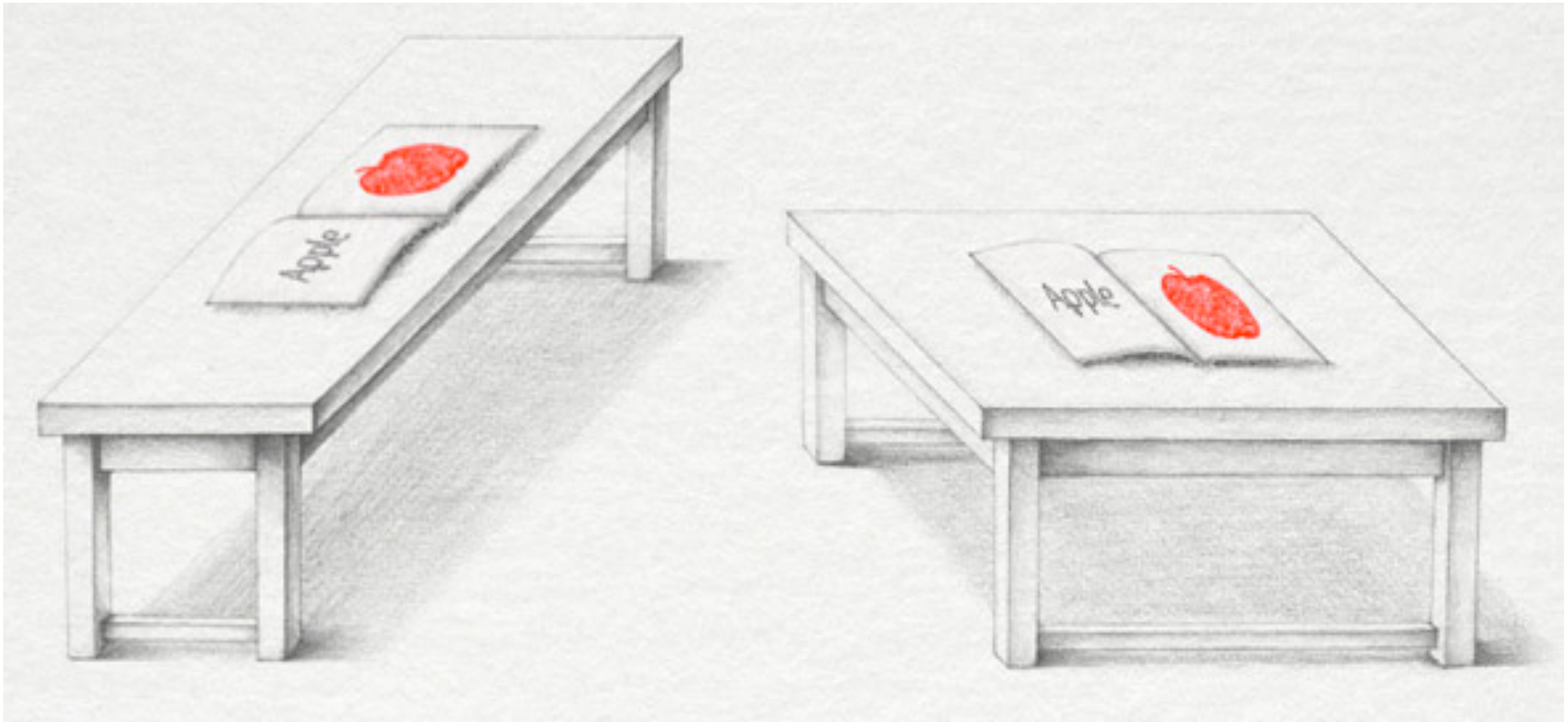


length?

SIZE?

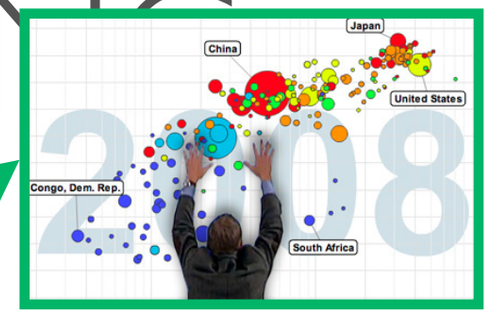


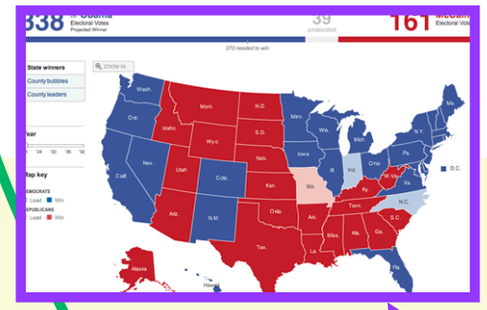

A
B


Apple
Apple

PUTTING IT ALL TOGETHER

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 USER'S TASK

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

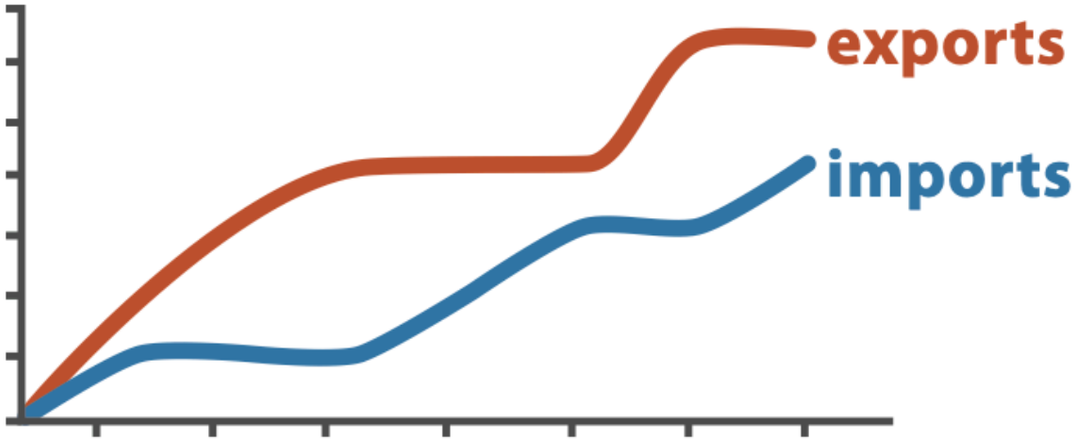
— Bill Buxton

"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"

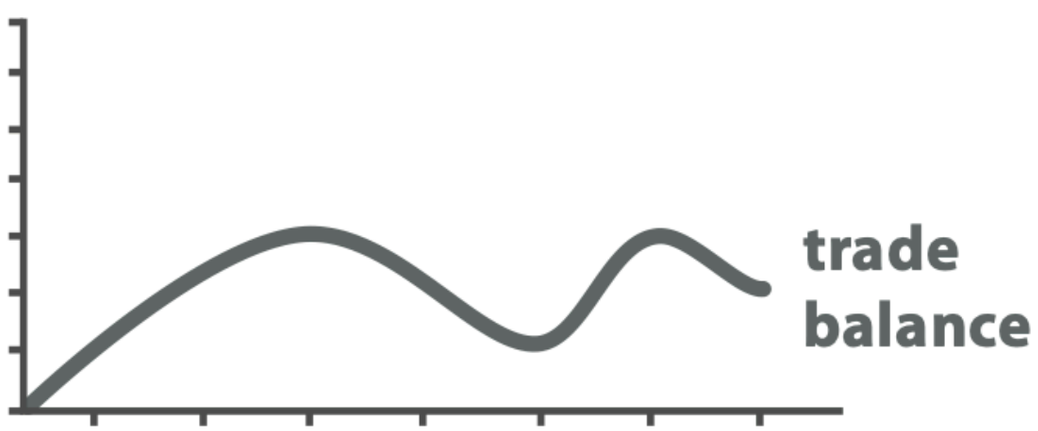
— Tamara Munzner

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

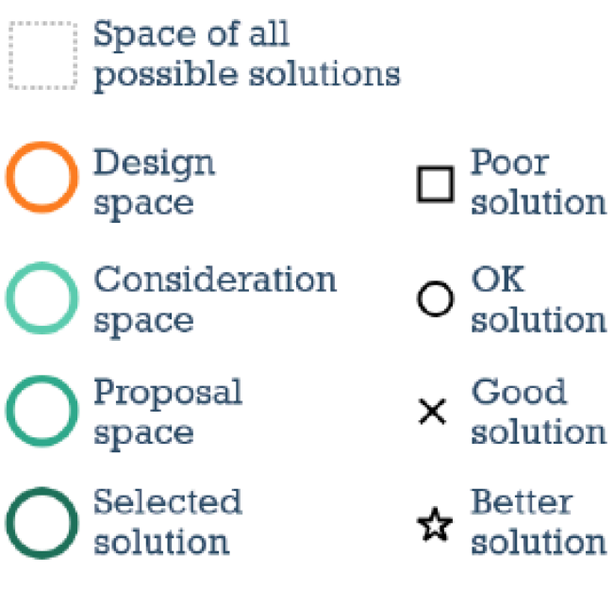
Examples, lots of examples!

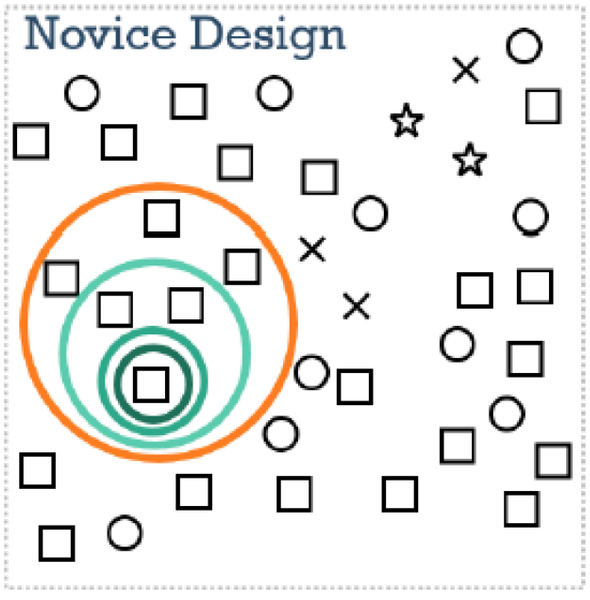
Fundamentals

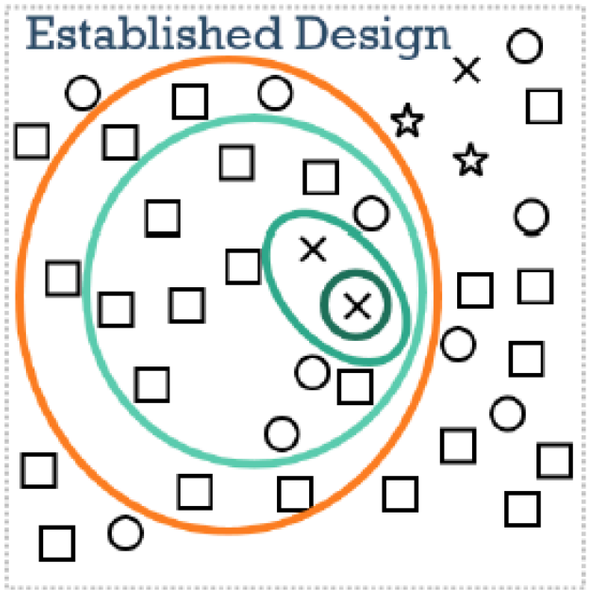
DESIGN SPACE?

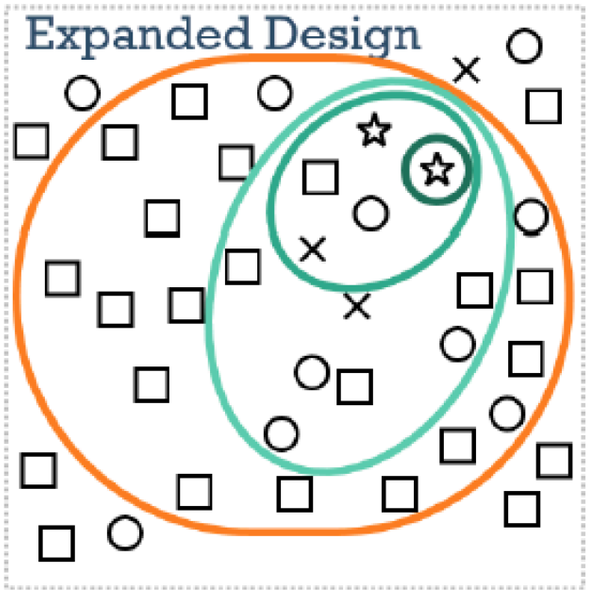
The set of possible design elements, parameters, and configurations that meet the specific application objectives.

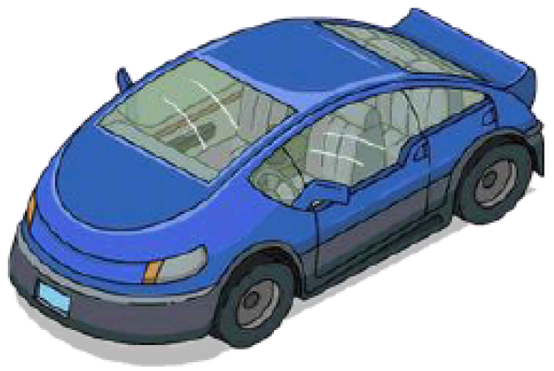
WHY DOES THE DESIGN SPACE MATTER?



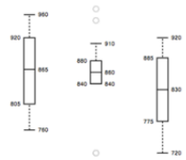








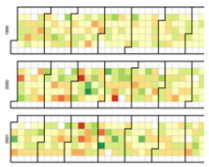
Box Plots



Bubble Chart

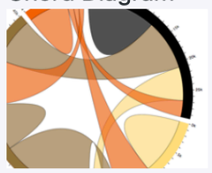
Bullet Charts

Calendar View



Non-contiguous Cartogram

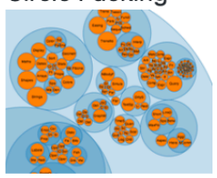
Chord Diagram



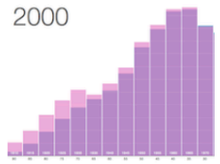
Dendrogram

Force-Directed Graph

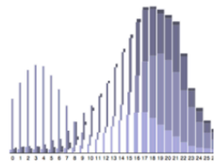
Circle Packing



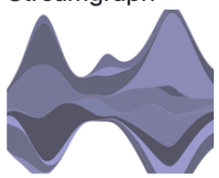
Population Pyramid



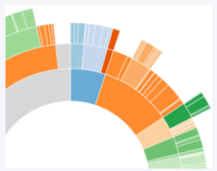
Stacked Bars



Streamgraph



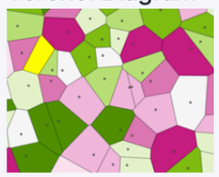
Sunburst

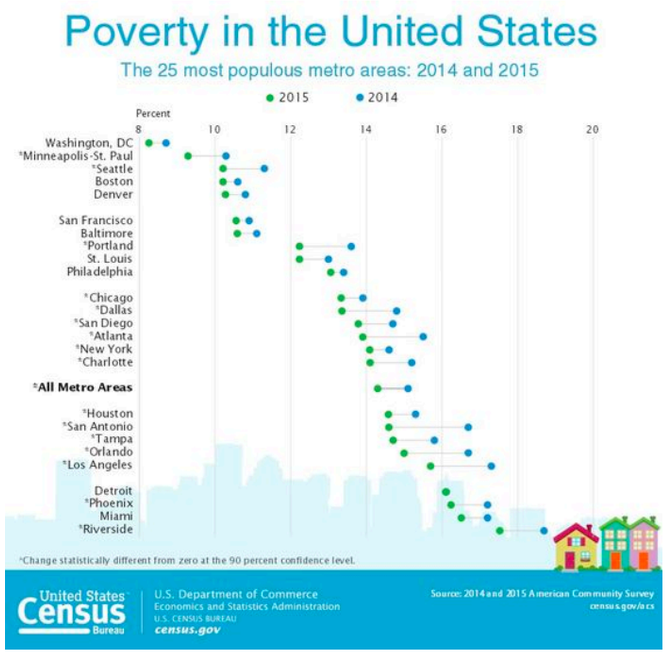


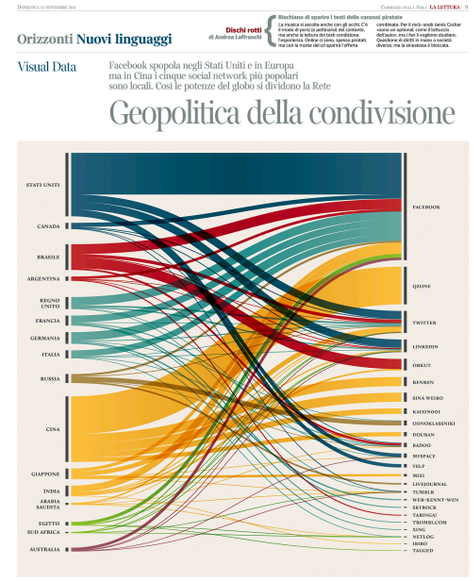
Node-Link Tree

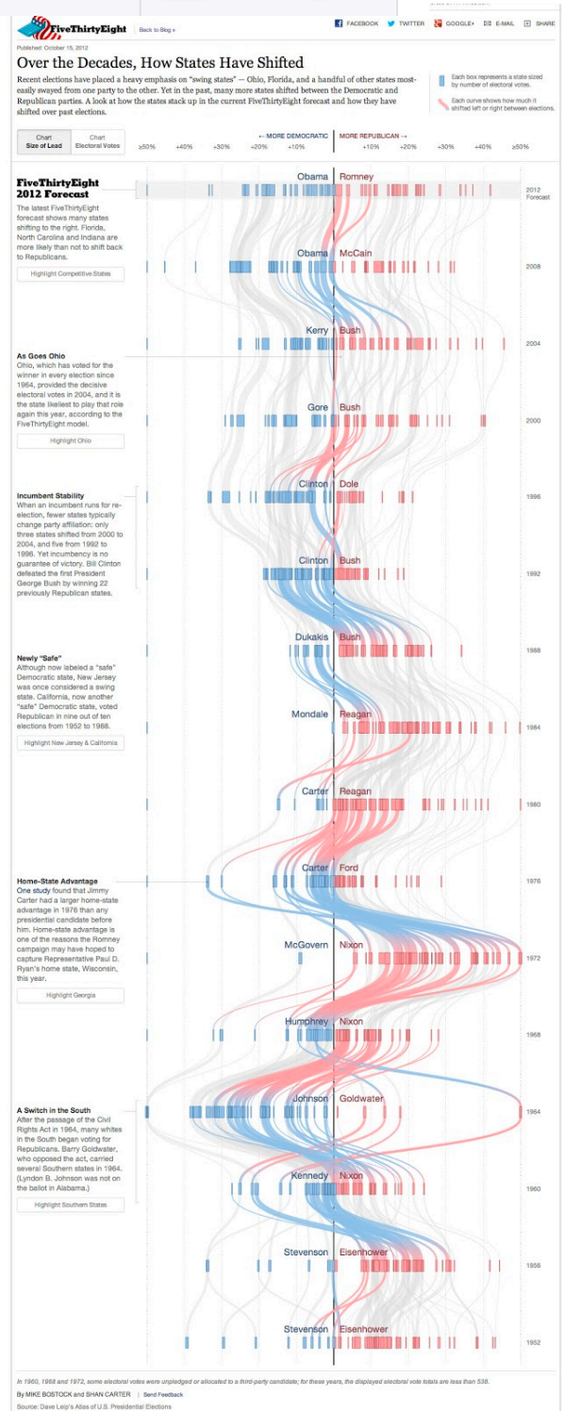
Treemap

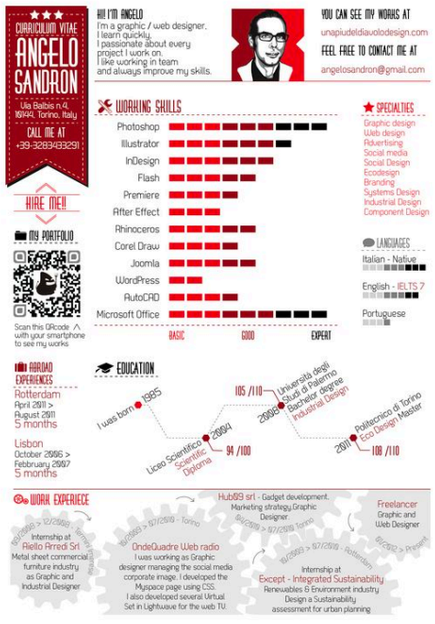
Voronoi Diagram

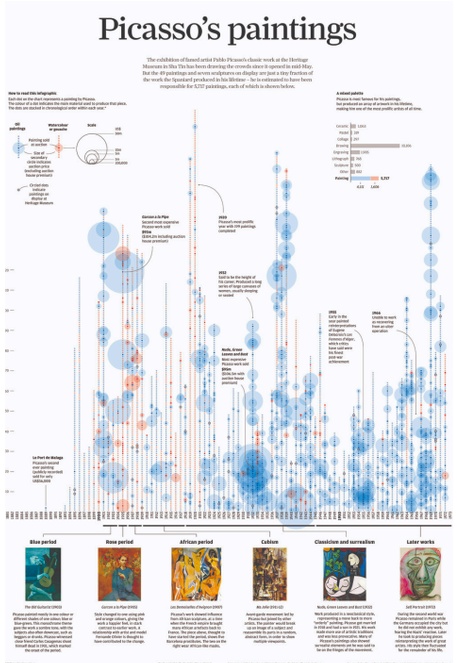


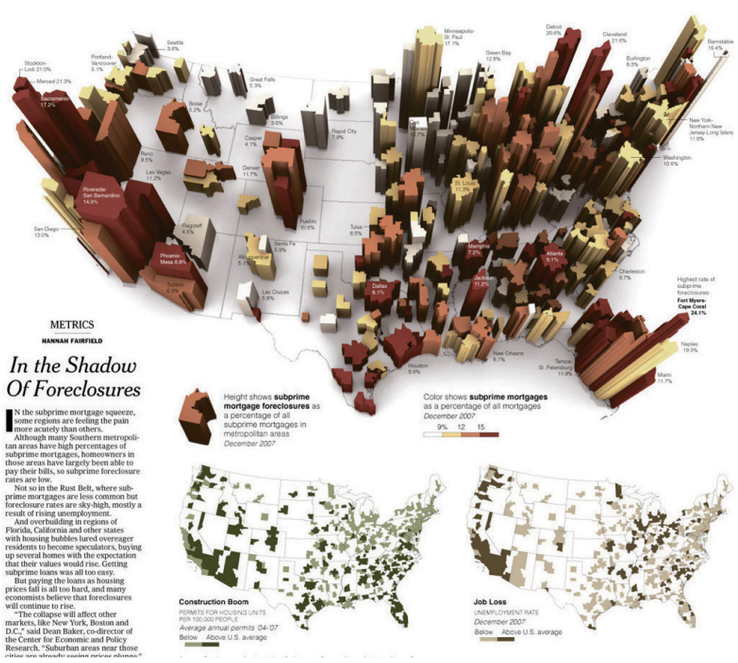










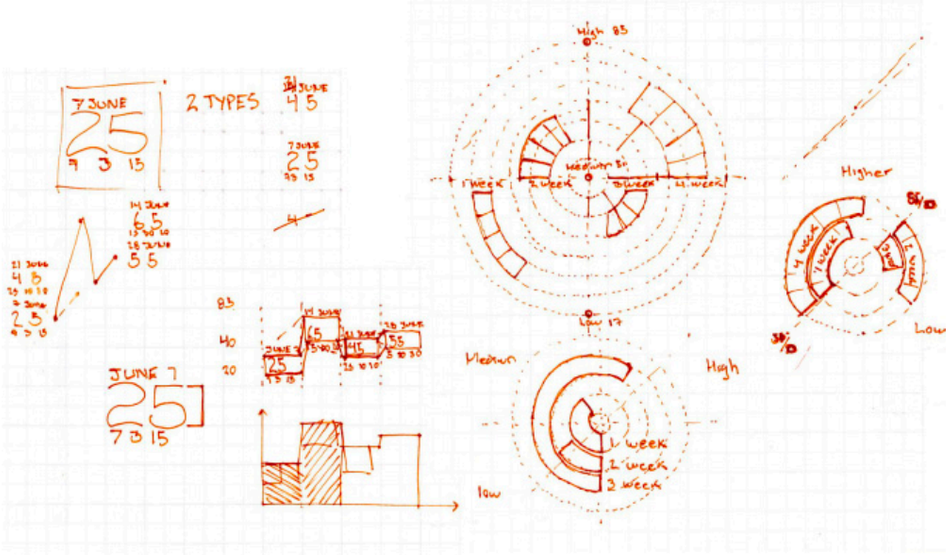


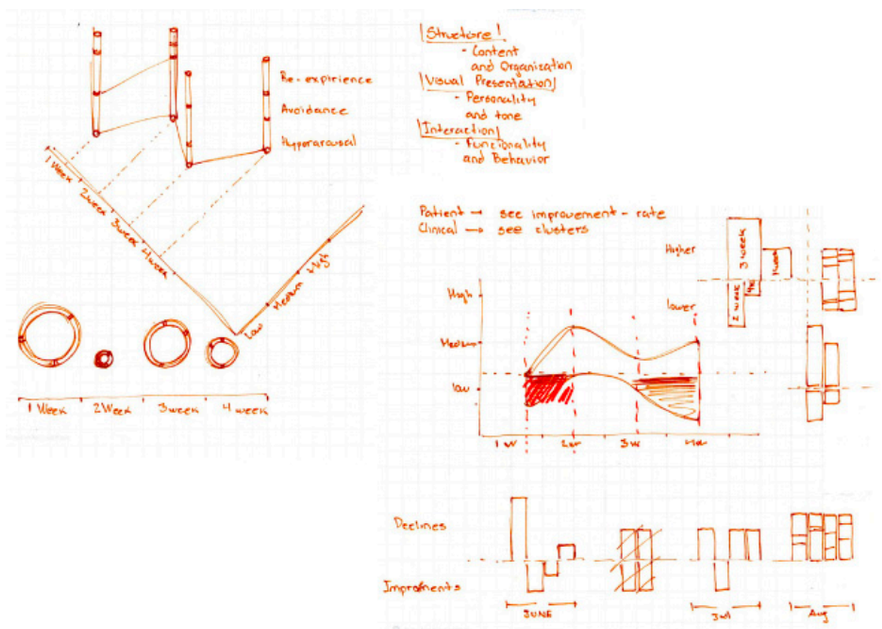
SKETCH!

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

Prototype ideas in early stages

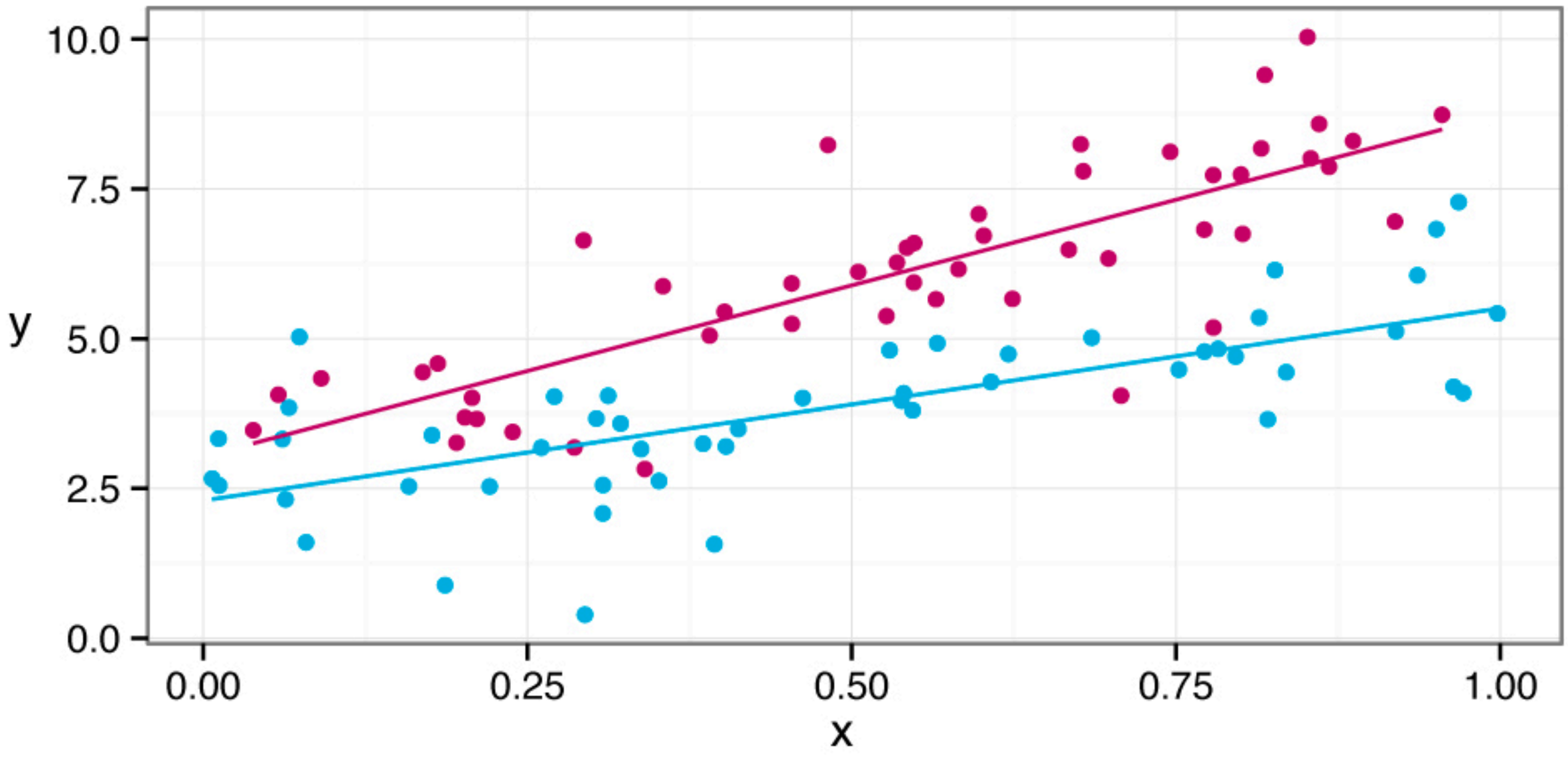
Share and discuss your designs with other people

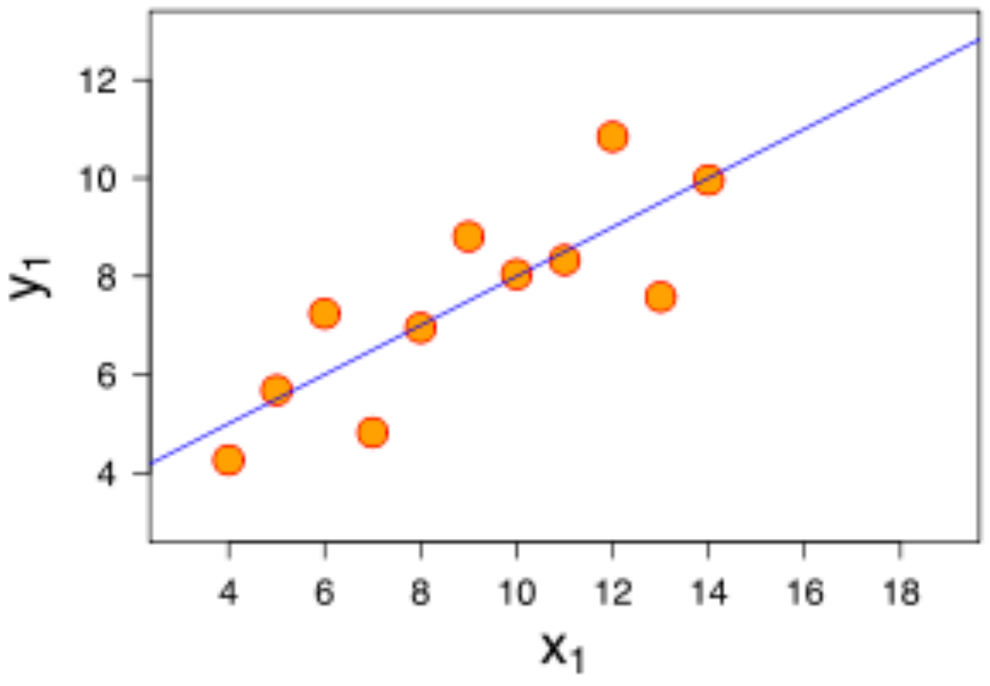


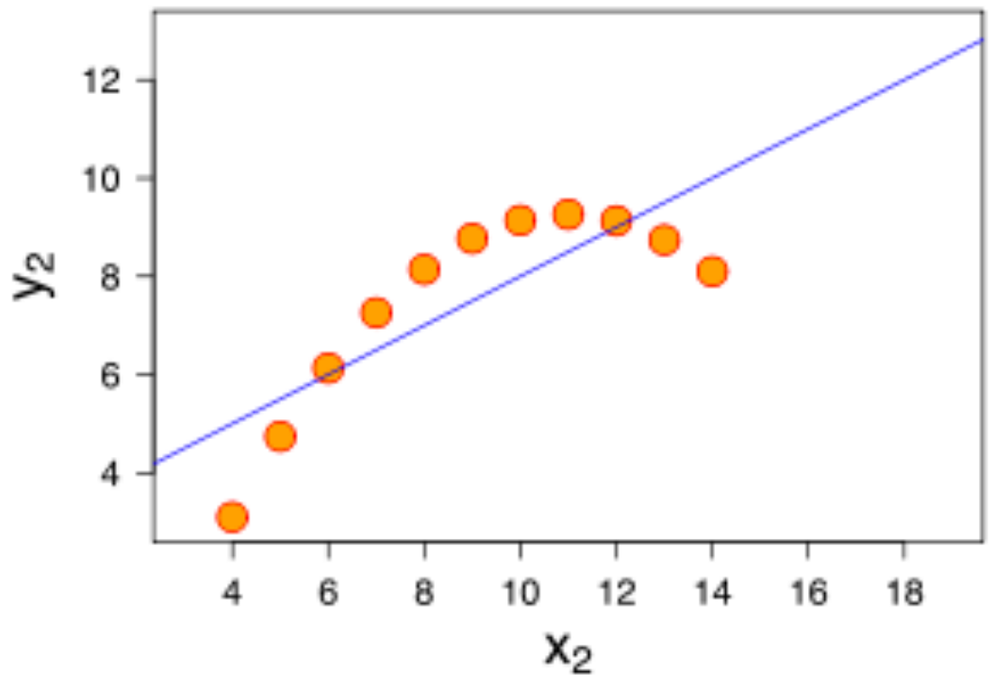


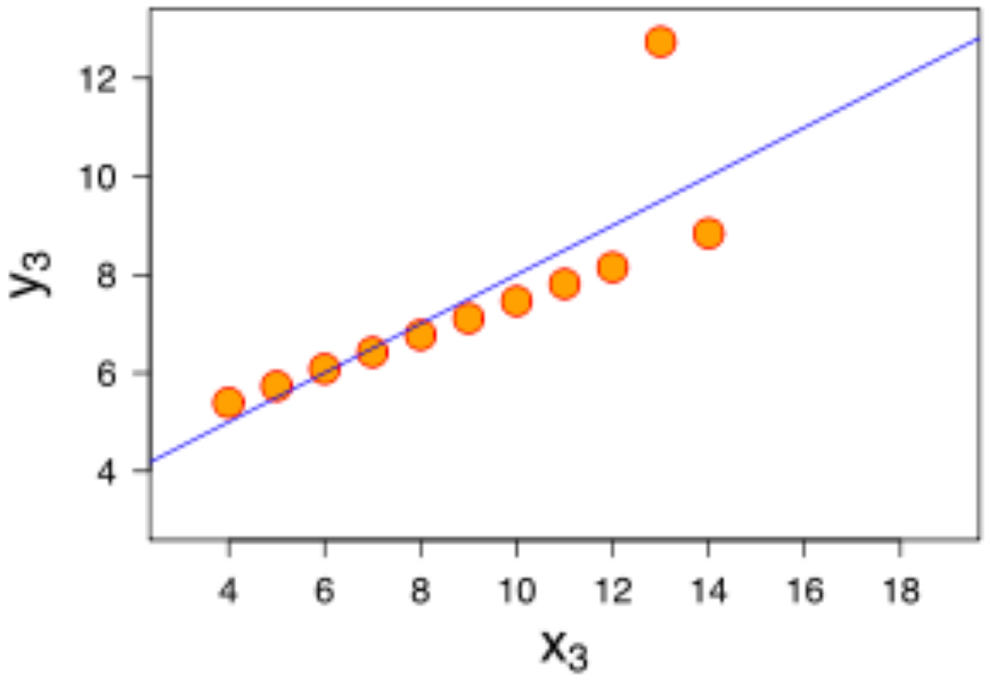
DON'T JUST GRAPH IT, FIT IT!

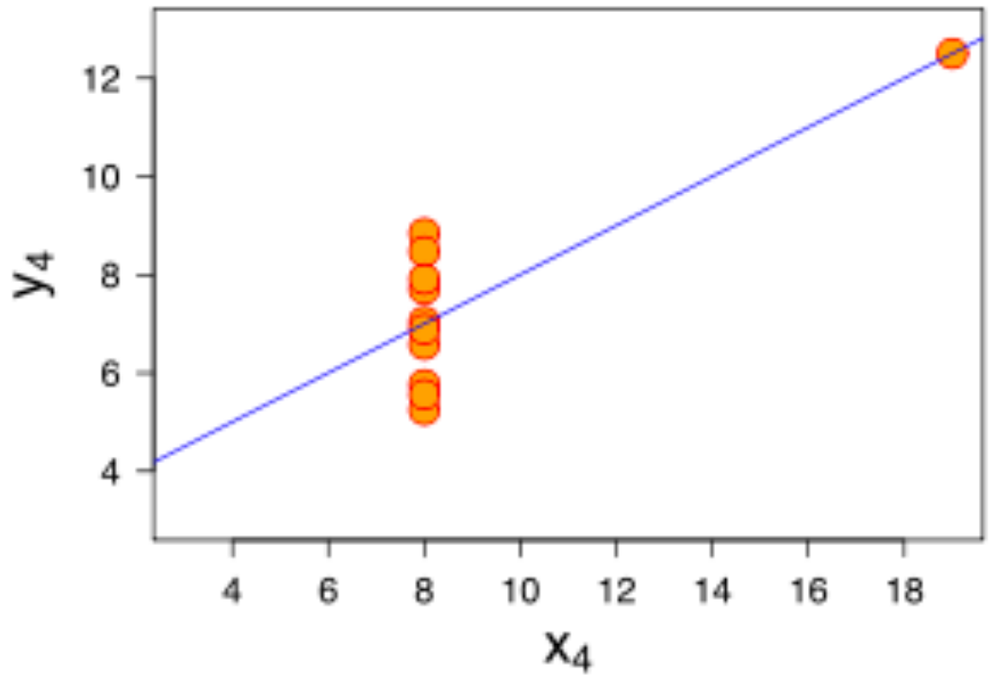
Graphing is needed, to see the details of the raw data
Fitting data to mathematical functions is needed too!
You can display both in the same graph…



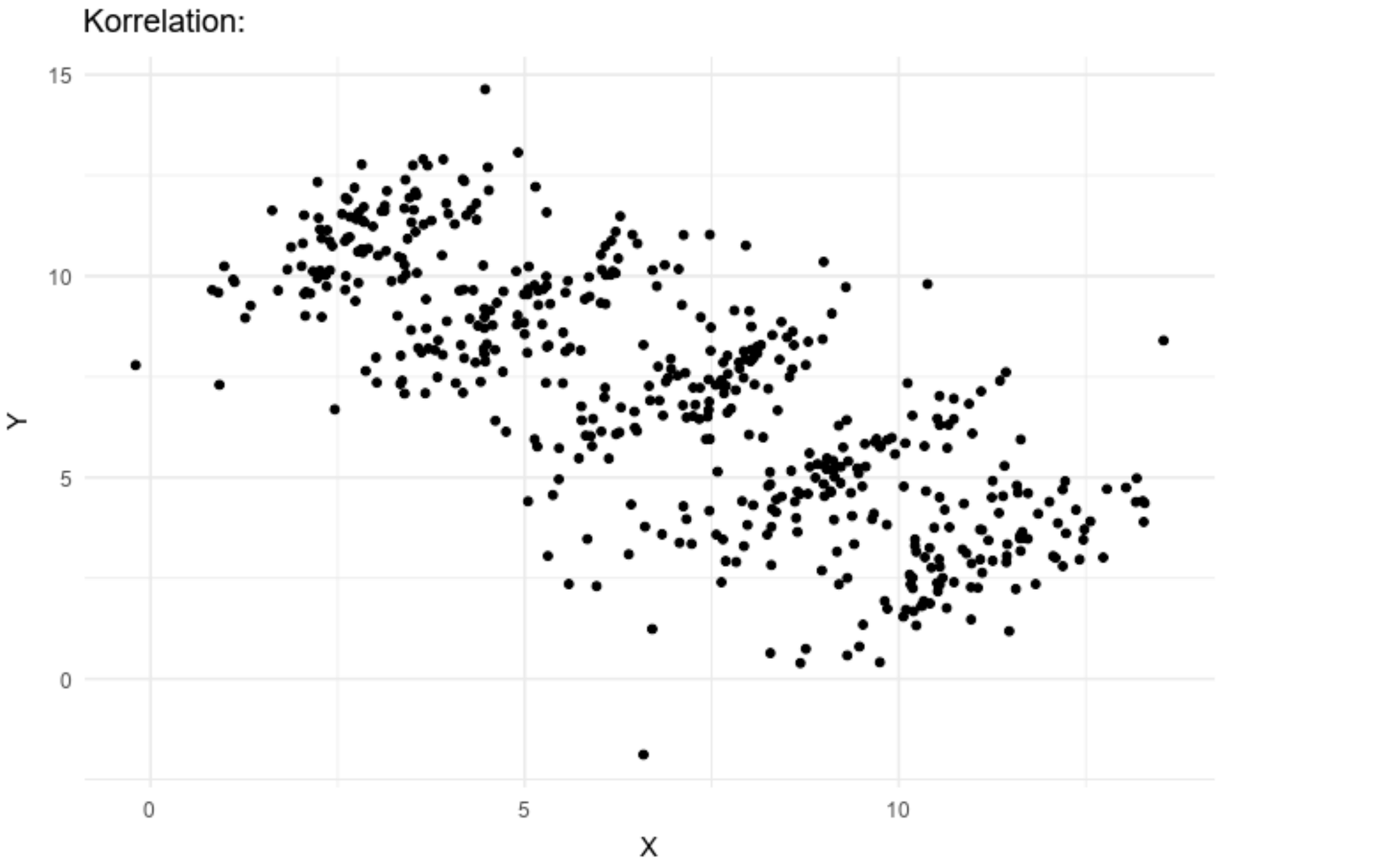

y1
x1
12
10
8
6
4
4
6
8
10
12
14
16
18

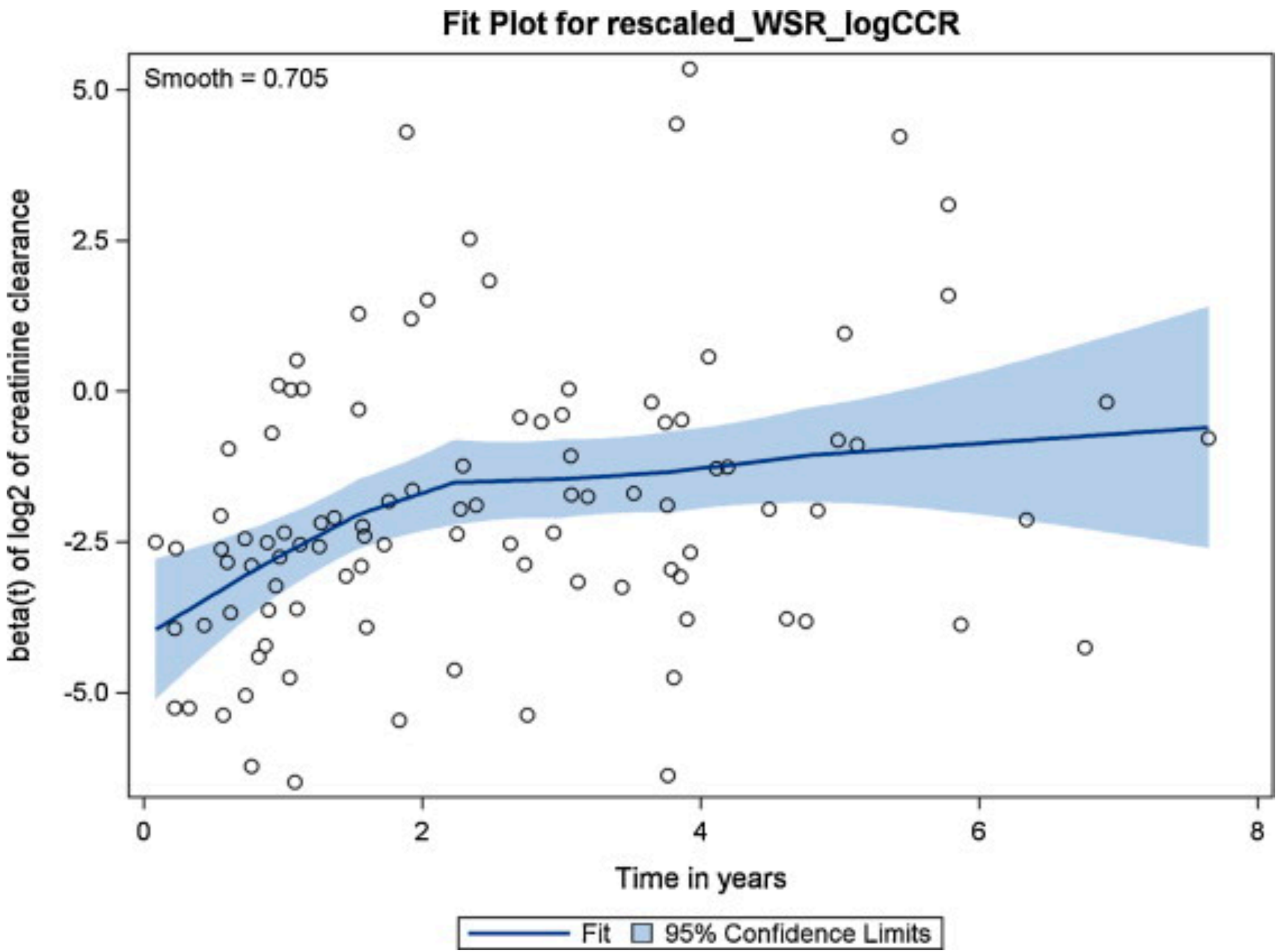

y2
x2
12
10
8
6
4
4
6
8
10
12
14
16
18


y3
x3
12
10
8
6
4
4
6
8
10
12
14
16
18

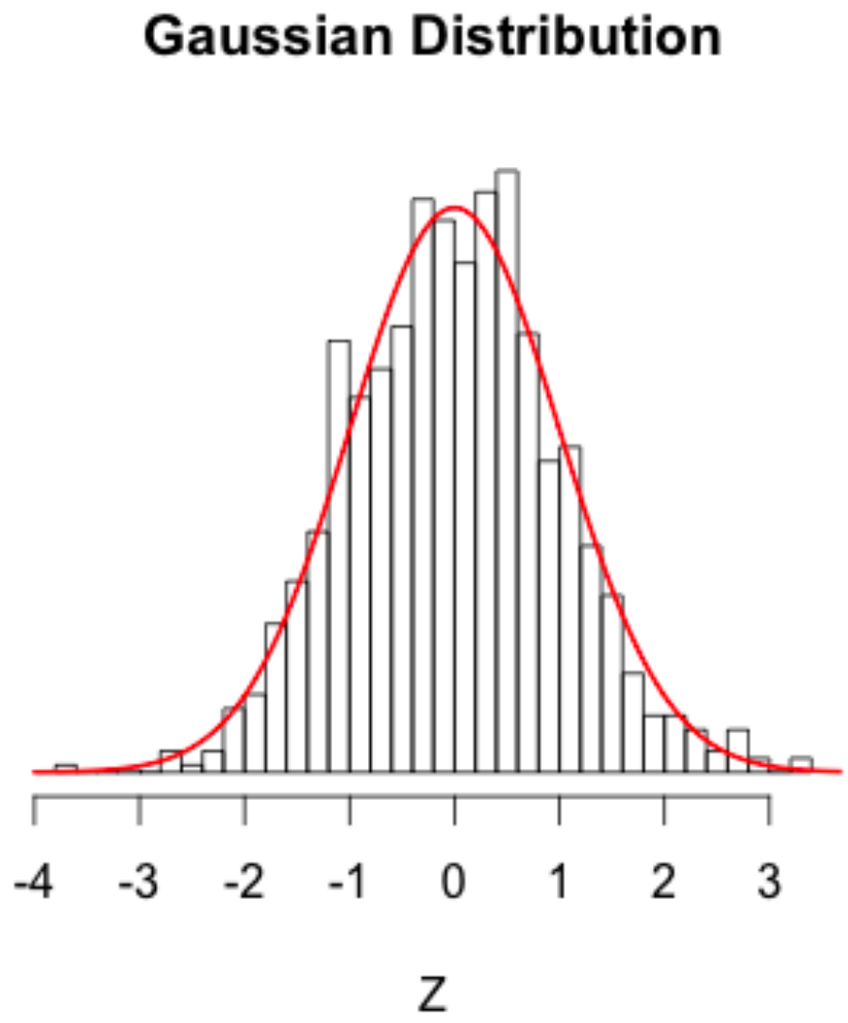

y4
x4
12
10
8
6
4
4
6
8
10
12
14
16
18

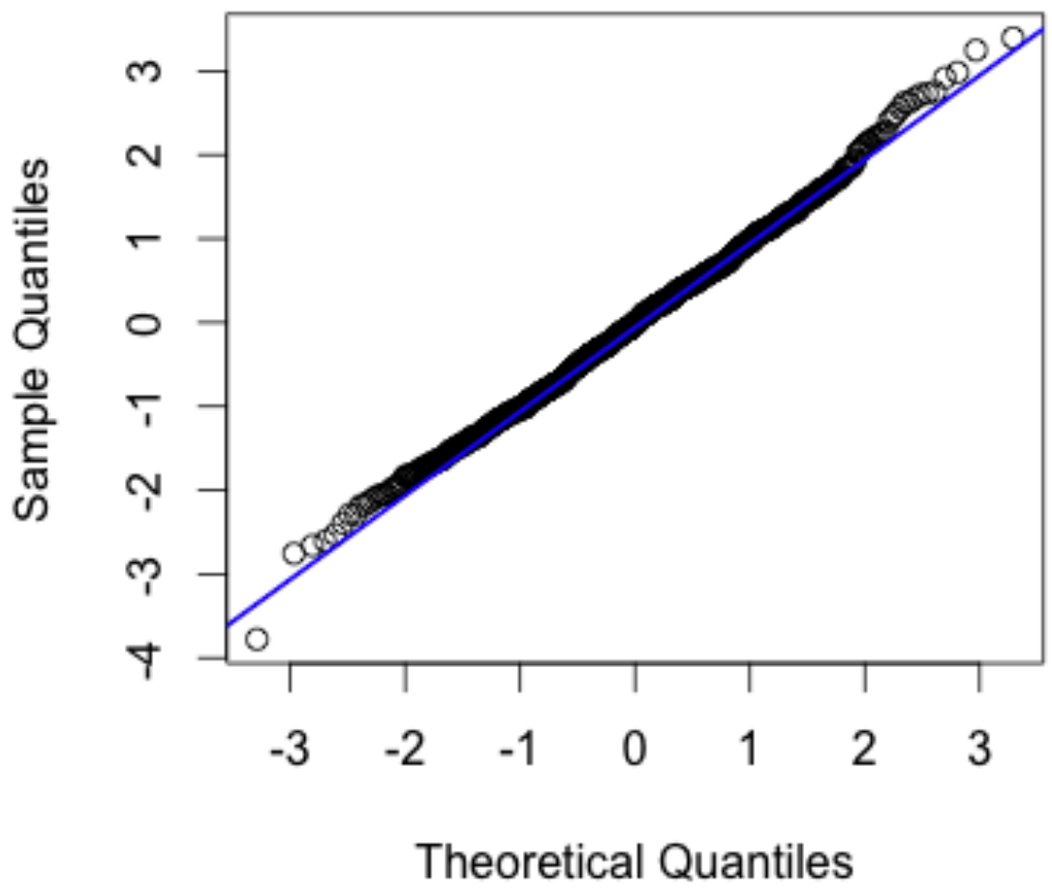
ITERATIVE FITTING



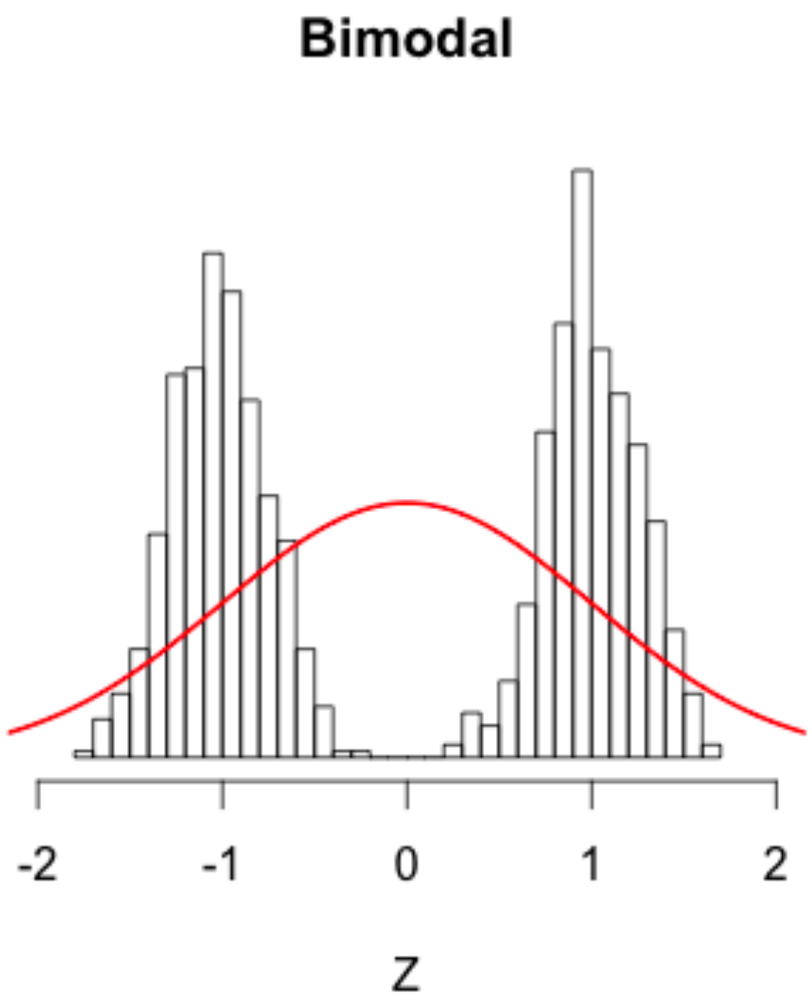

Fit Plot for rescaled_WSR_logCCR
Smooth = 0.705
beta(t) of log2 of creatinine clearance
5.0
2.5
0.0
-2.5
-5.0
0
2
4
6
8
Time in years
Fit
95% Confidence Limits

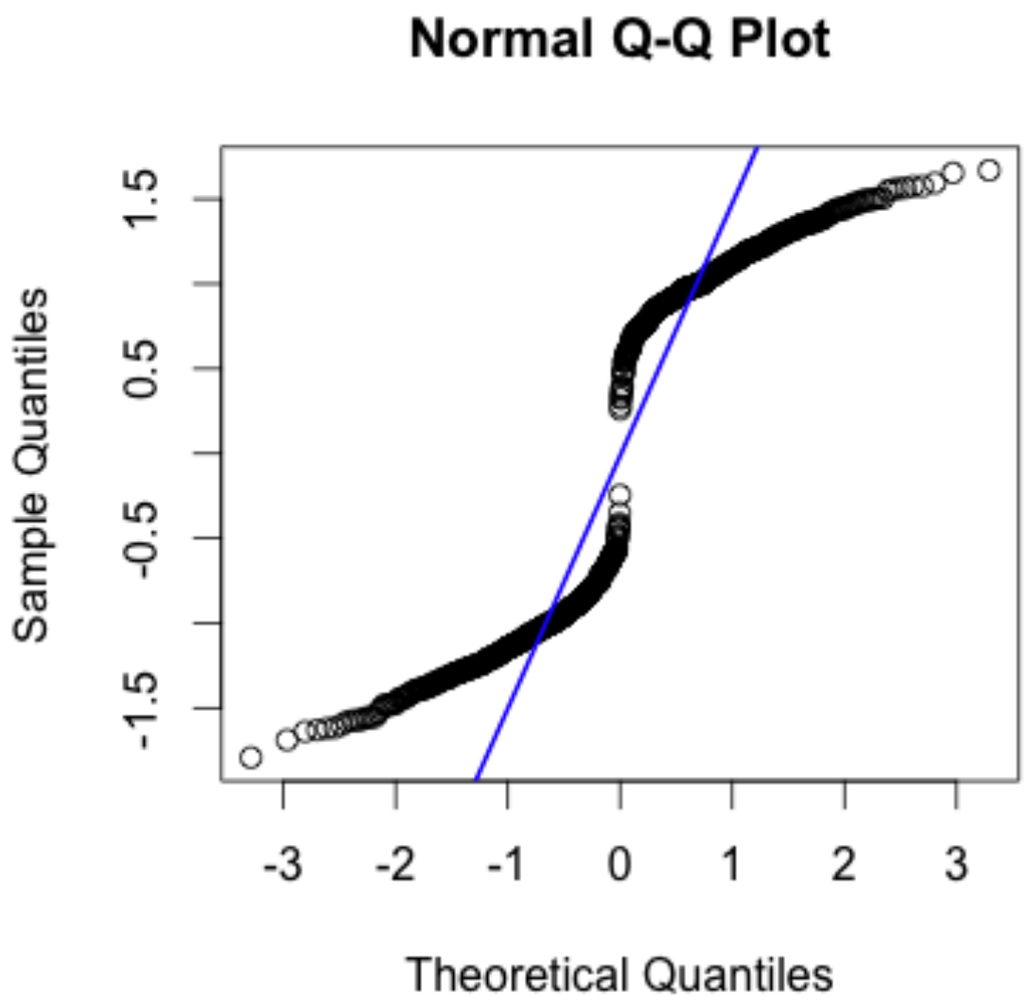
QUANTILE-QUANTILE PLOT (Q-Q PLOT)



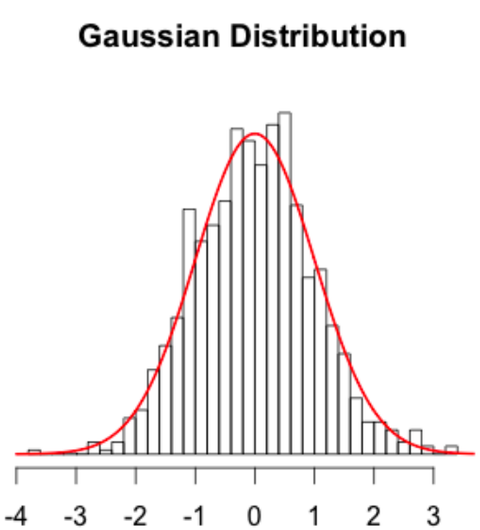


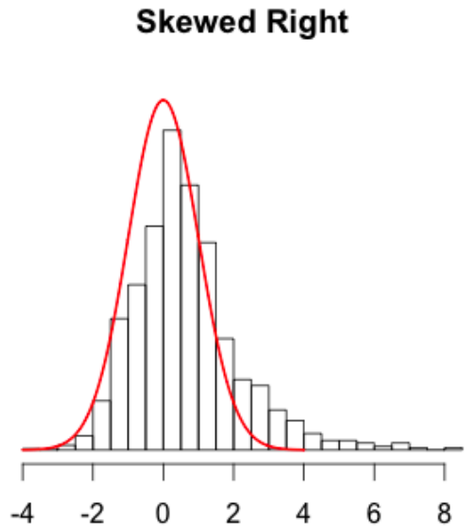
QUANTILE-QUANTILE PLOT (Q-Q PLOT)

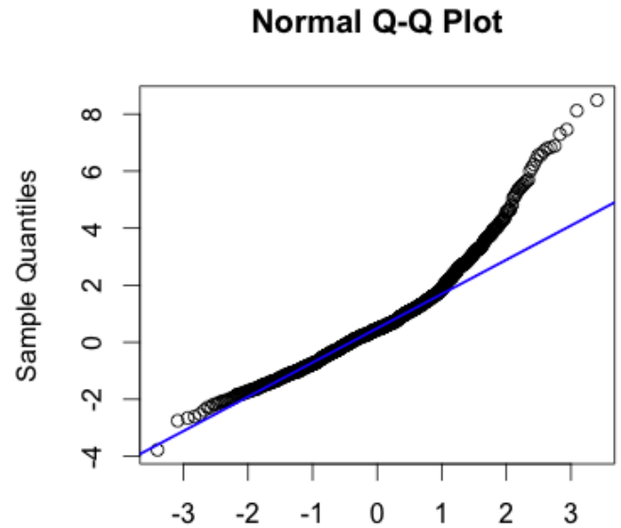


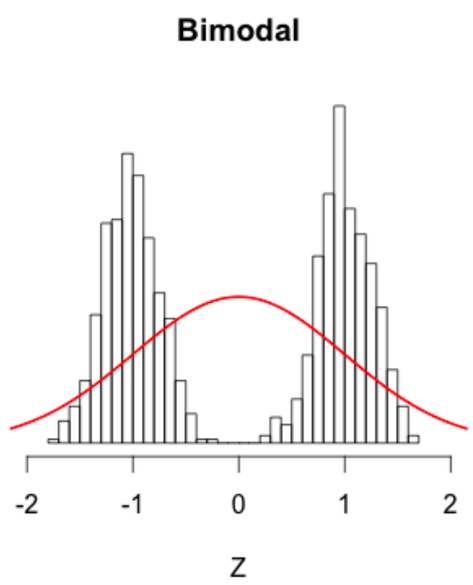


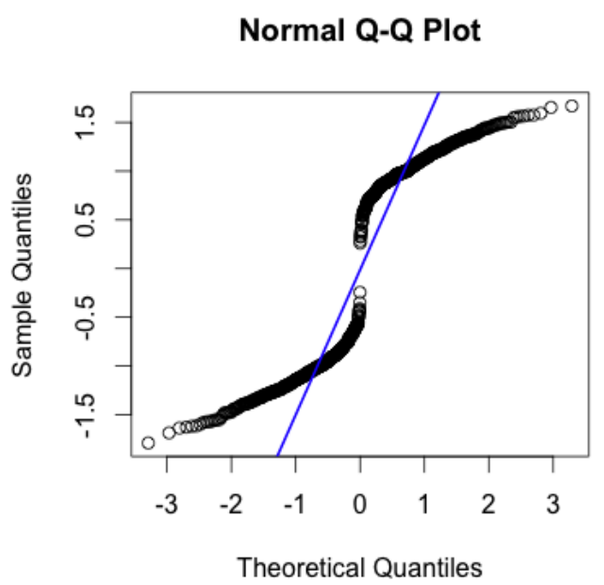
QUANTILE-QUANTILE PLOT (Q-Q PLOT)

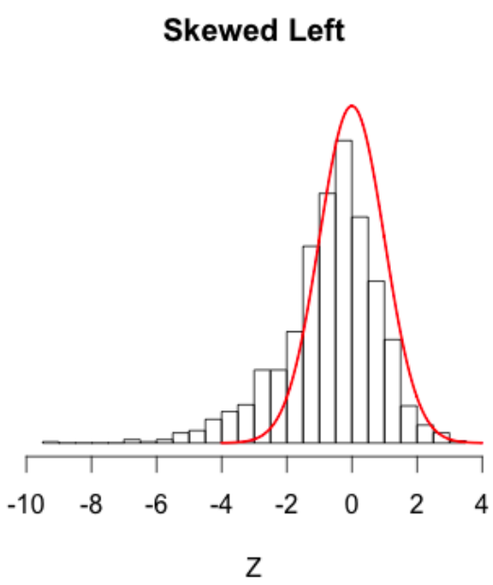


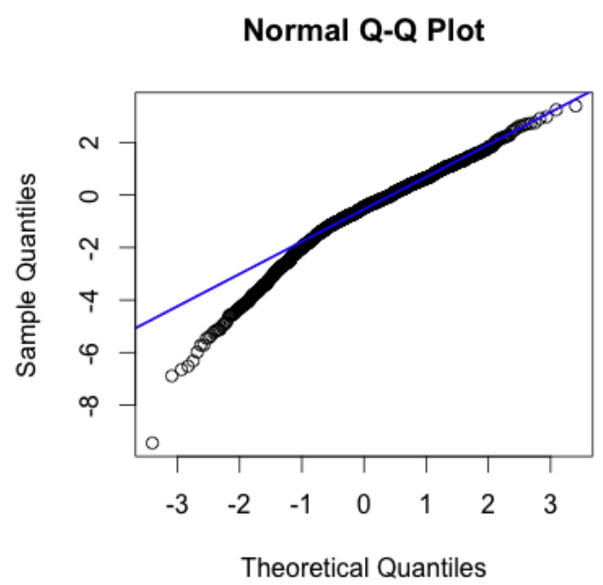




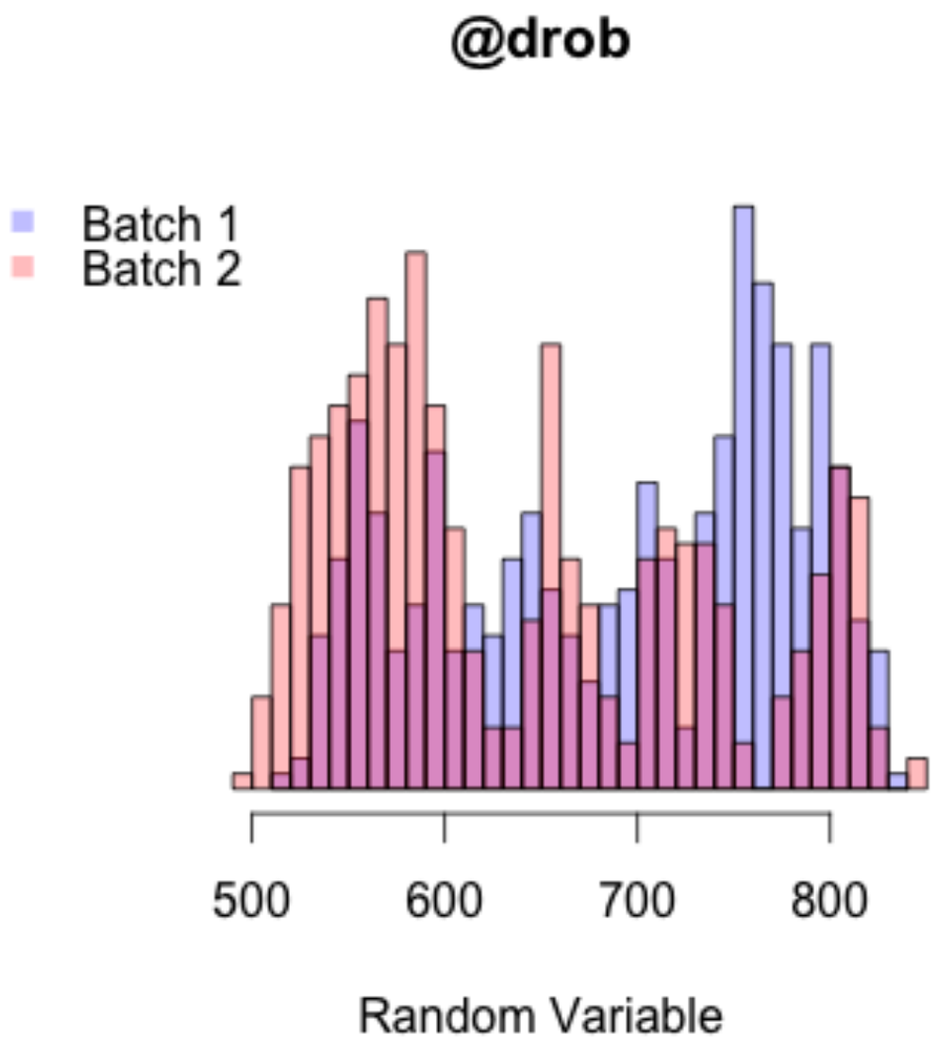


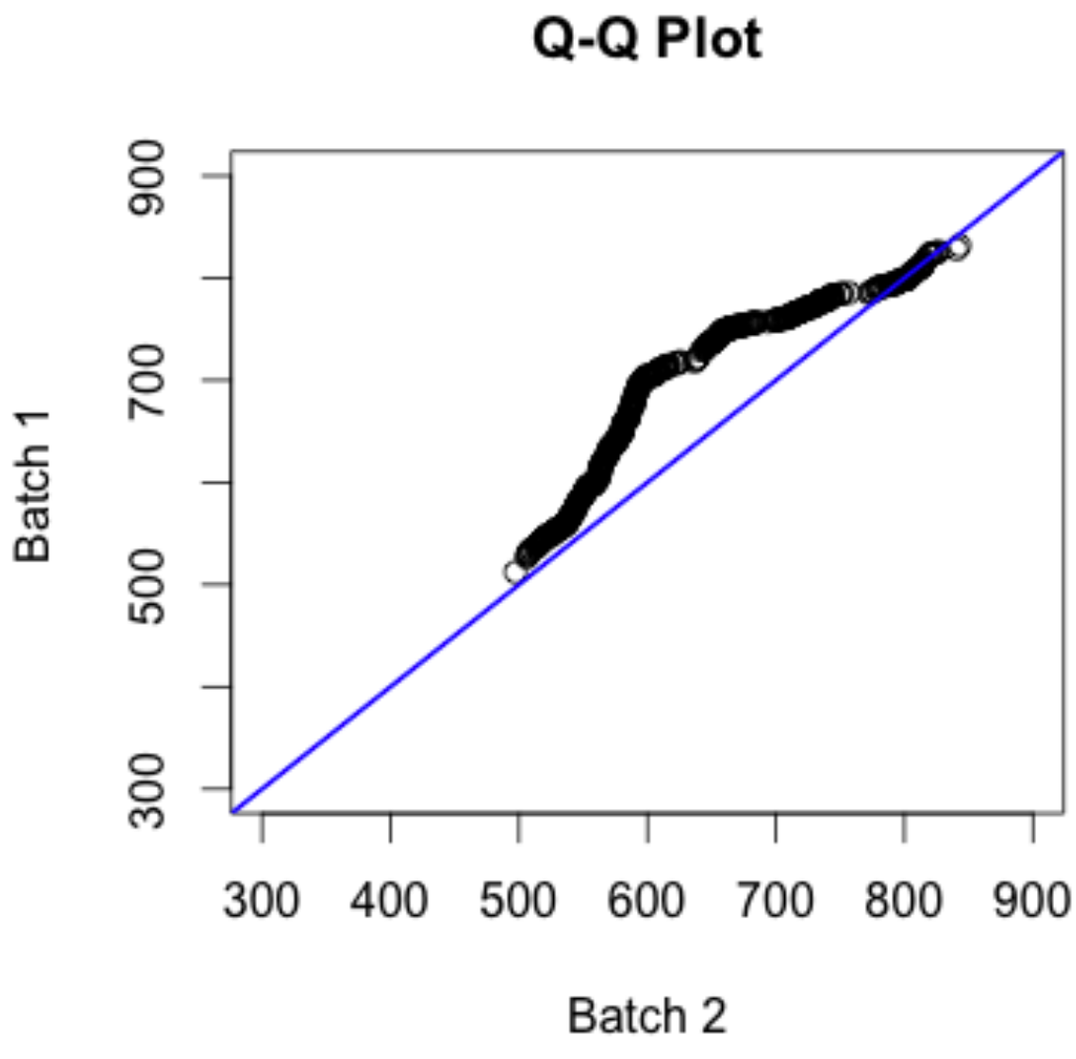






QUANTILE-QUANTILE PLOT (Q-Q PLOT)





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

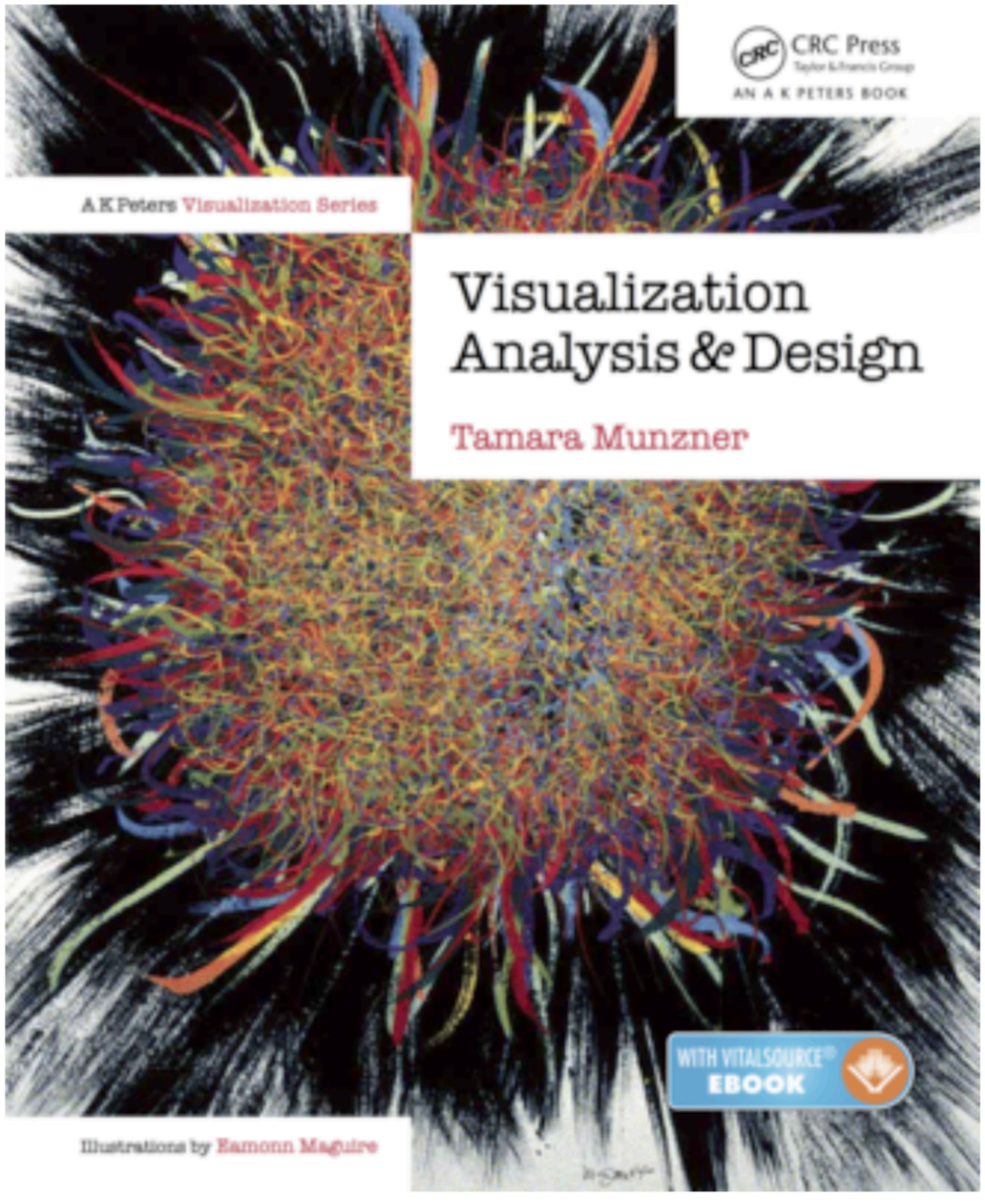
http://junkcharts.typepad.com/

Junk Charts: learn from others (mistakes)

Text-book on data visualization

(1st of its kind)

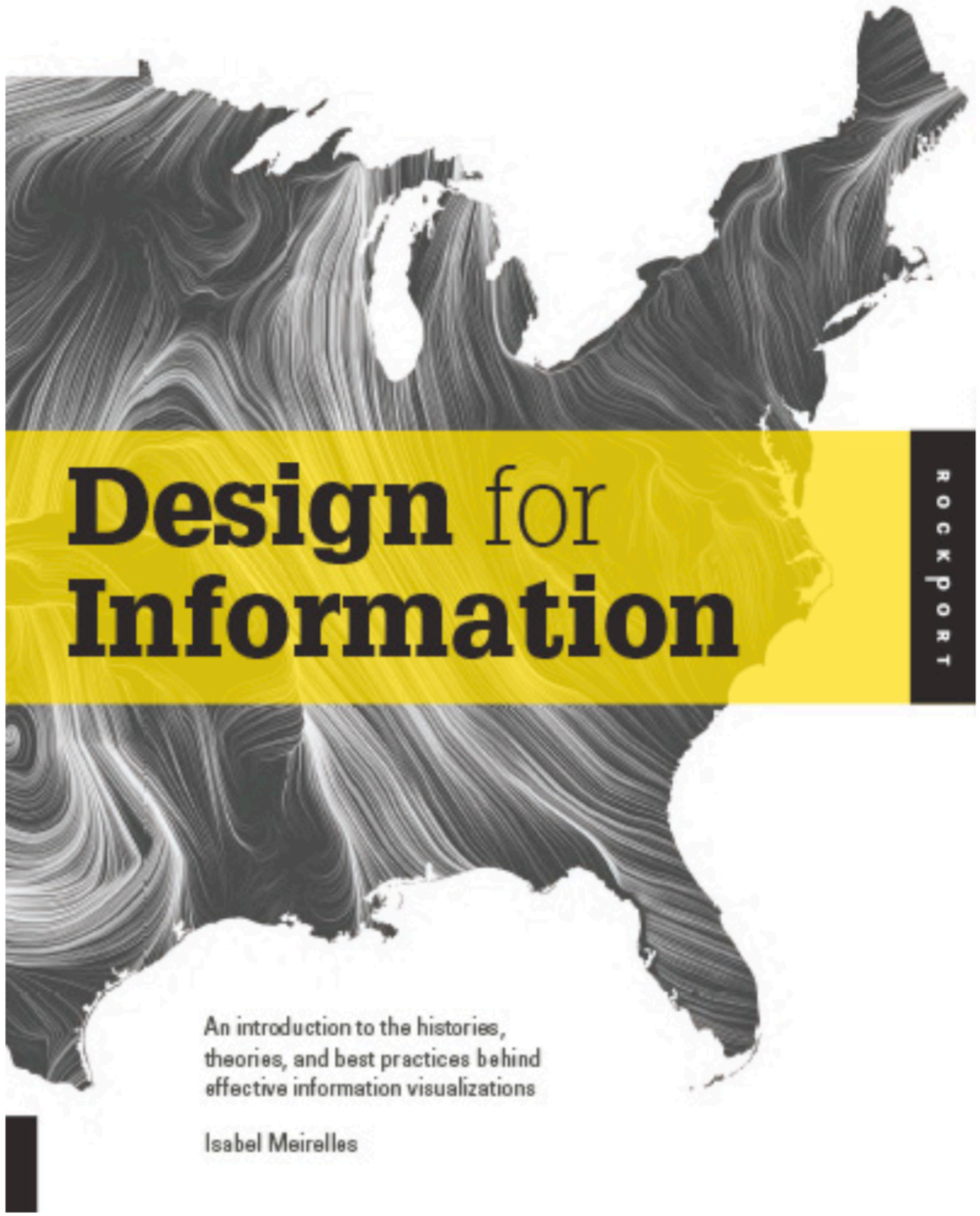
If you can only get one,
THIS is the one!



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

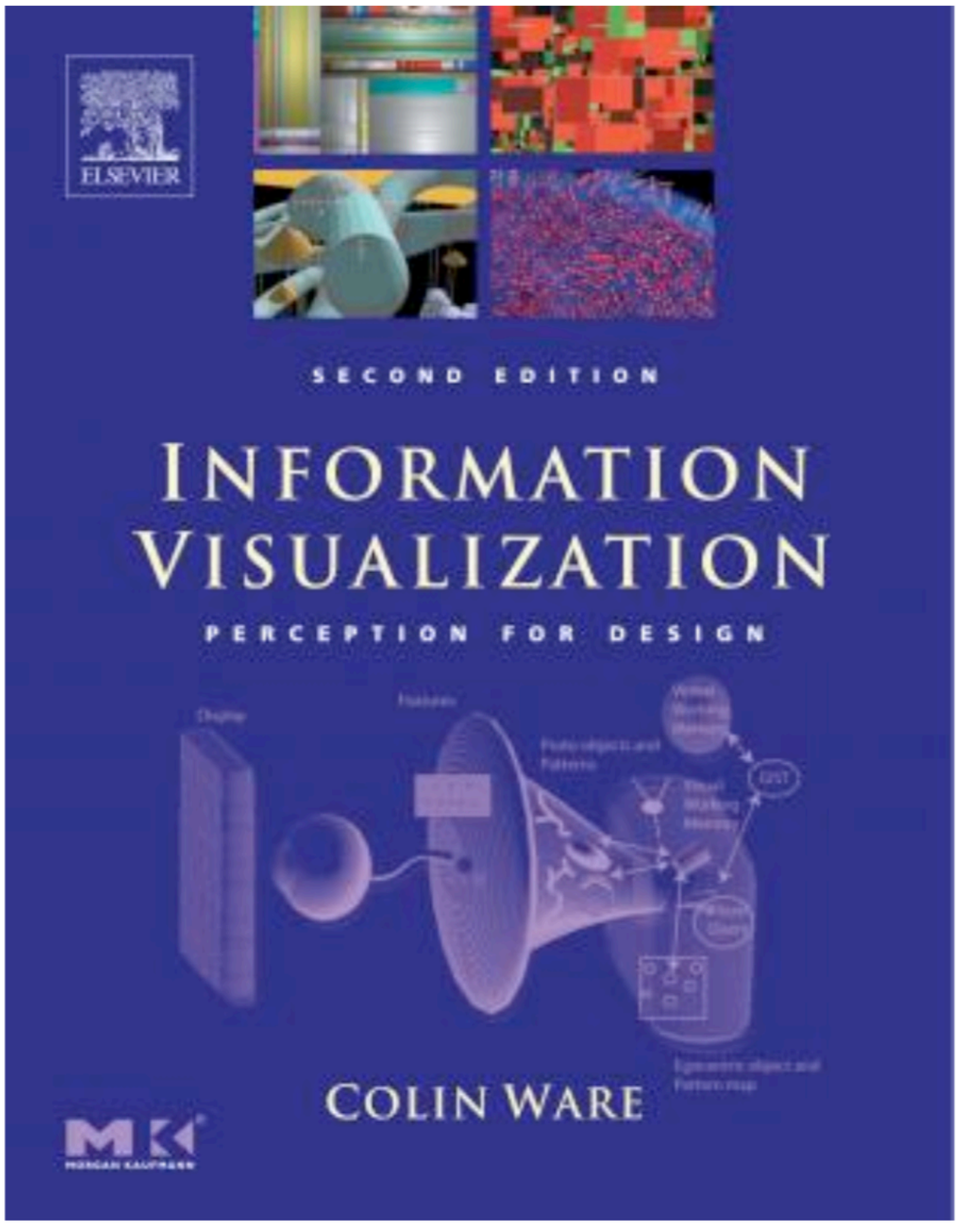
Infovis through examples, history and some free discussion. Very visual.

Beautiful!



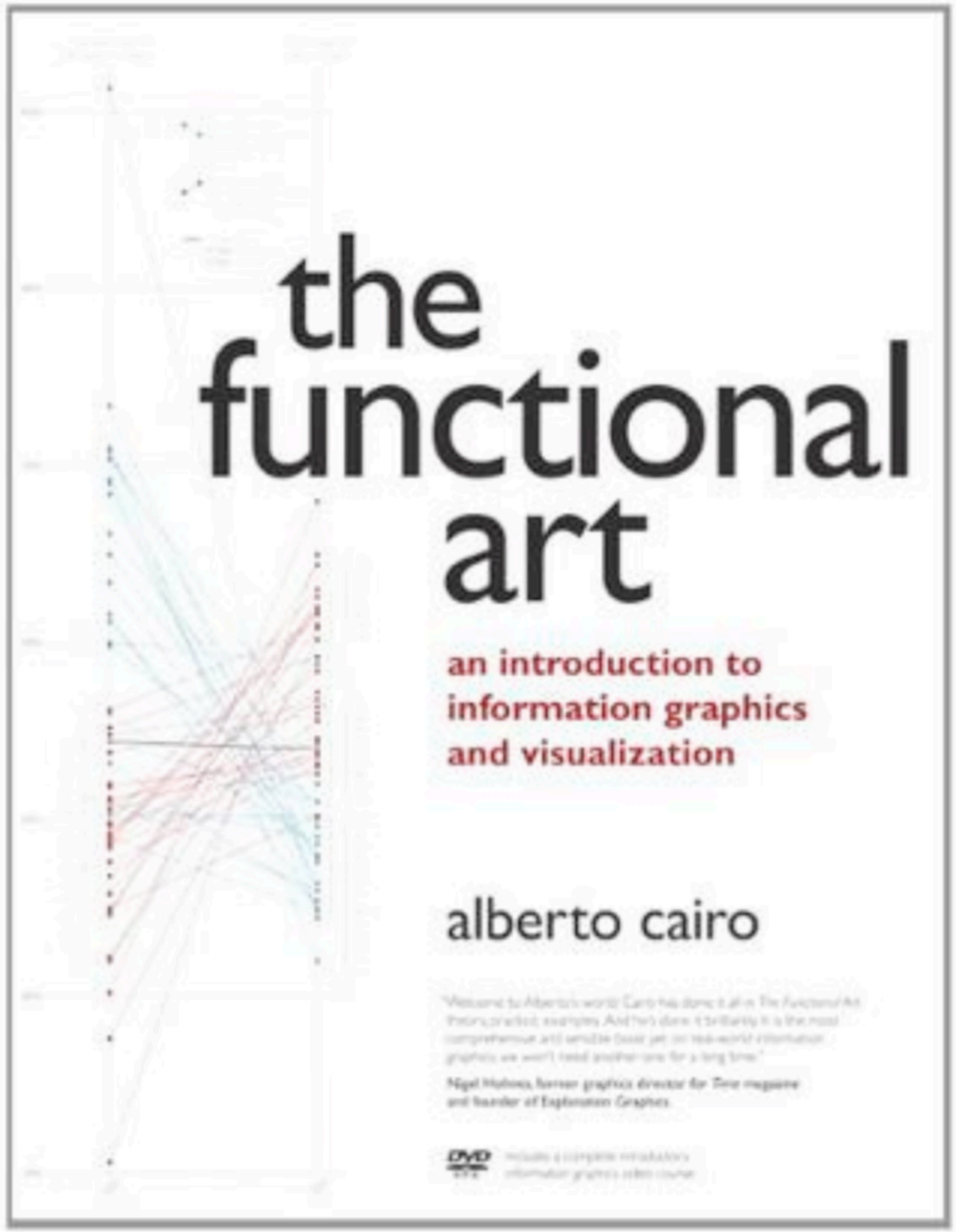
Meirelles, Isabel. Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations. Rockport publishers, 2013.

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



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

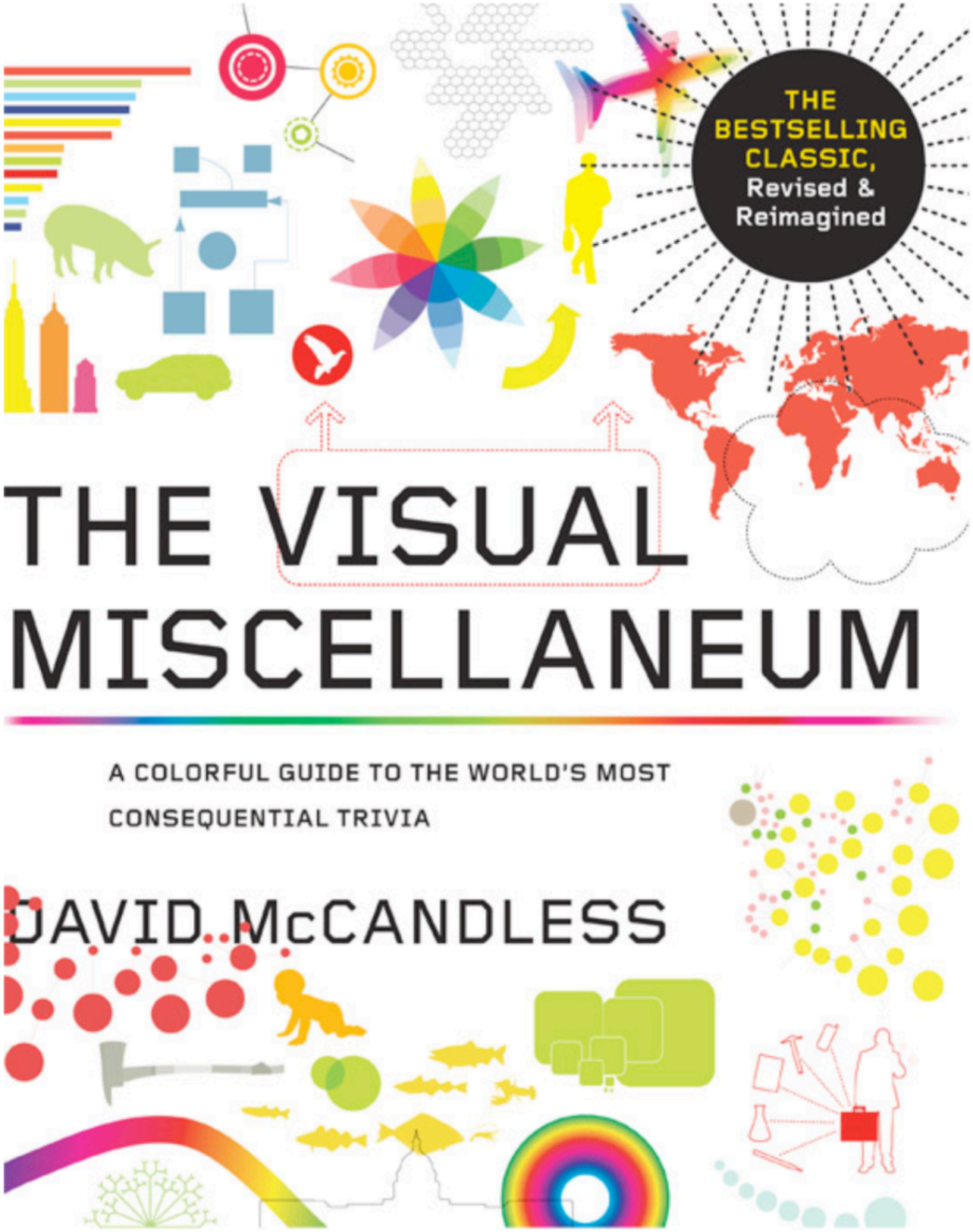
Focus on
infographics



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.

LET'S DECODE VISUALIZATIONS

MARKS

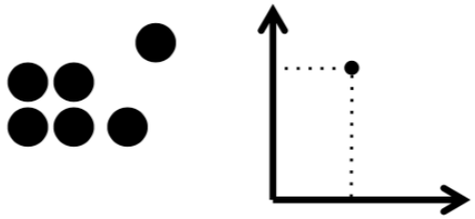
Points

Lines

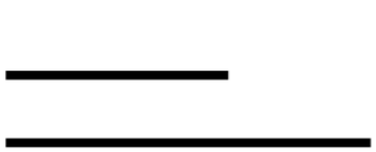
Areas

CHANNELS

Position



Size- Length



Size- Area

Size- Volume

Format

Texture

Orientation/Direction

Angle

Color

Color - Hue

Color - Saturation

Color - *Luminance*

DECODING EXERCISE

Marks used?

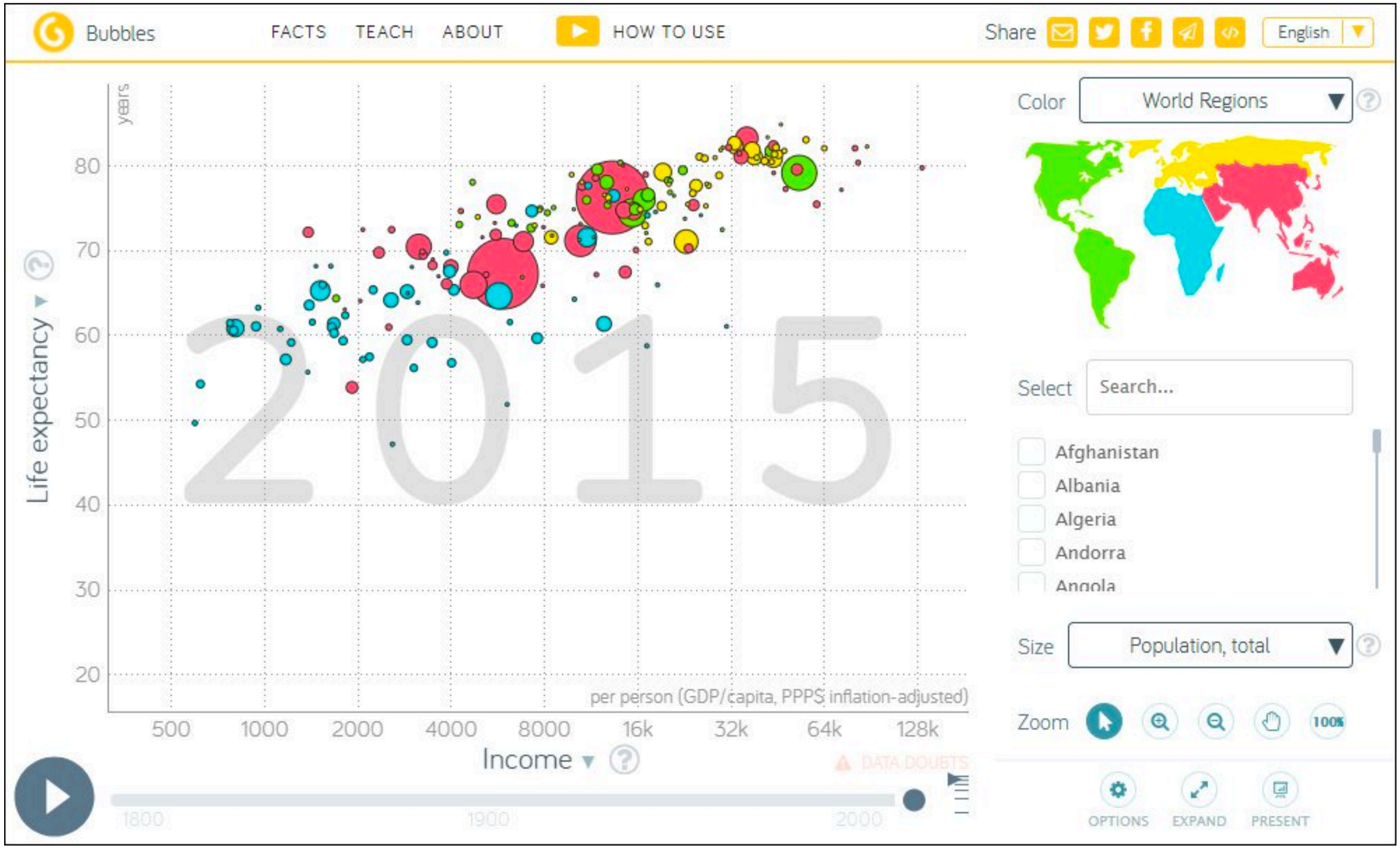
Mark of type …… encodes ….

Visual channels used?

Channel … encodes …

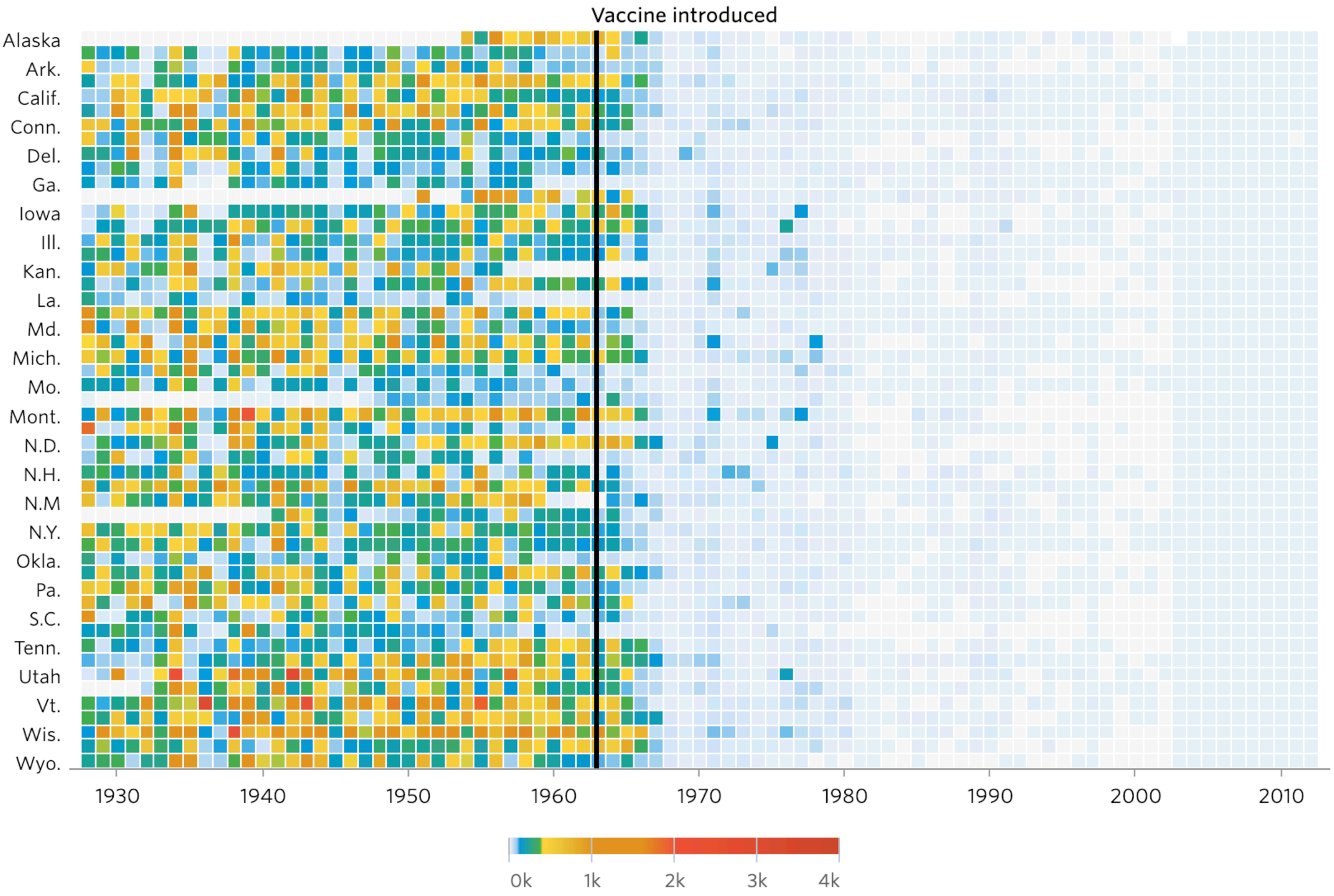
Channel … encodes …

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)



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

Measles

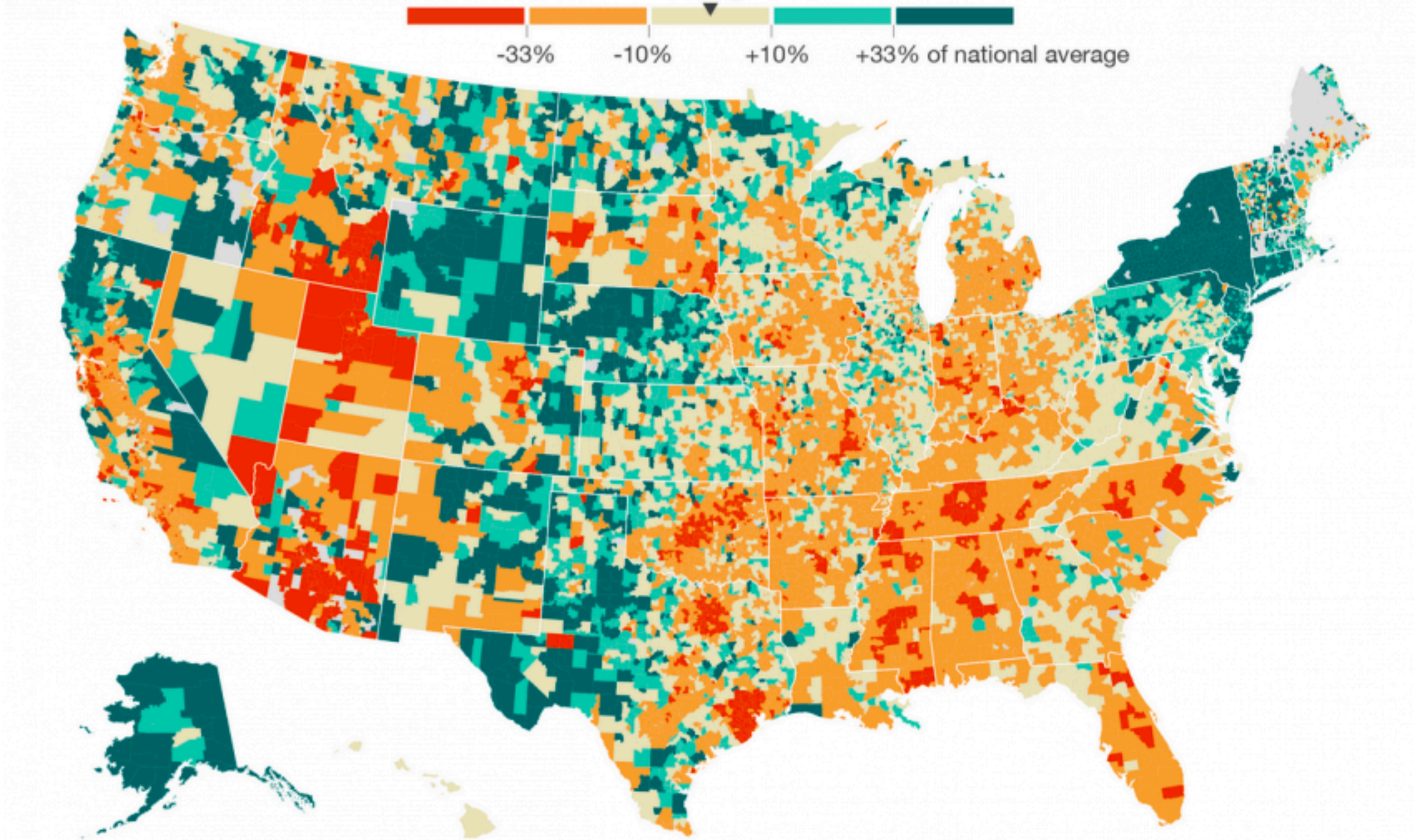


http://graphics.wsj.com/infectious-diseases-and-vaccines/

SPENDING PER STUDENT, BY SCHOOL DISTRICT

Adjusted for regional differences, for primary and unified school districts

National average: $11,841

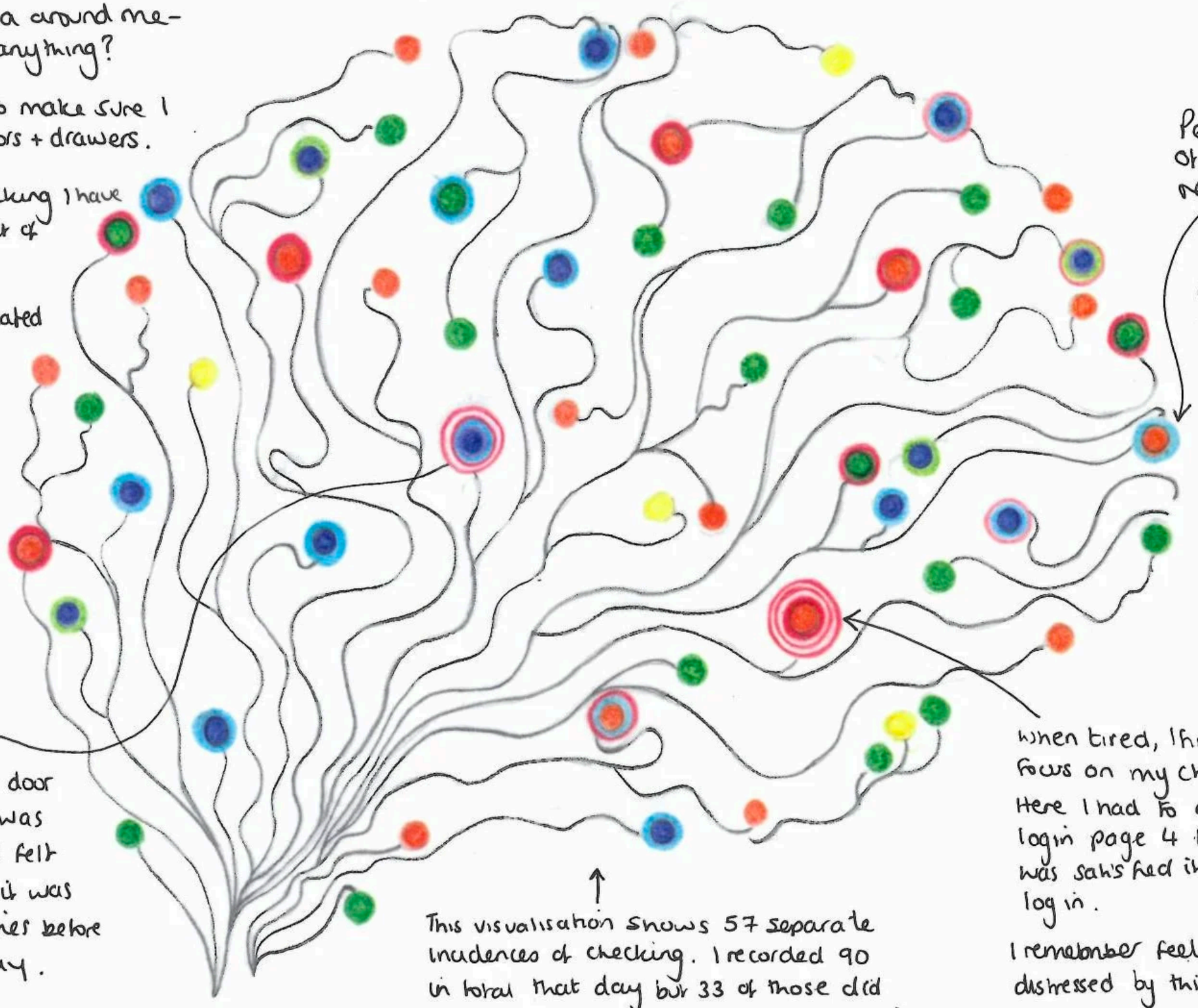


A Day of OCD - Conscious Acts of checking

I am afraid of losing material and digital information.
Each circle represents one act of checking to ensure I don't lose information.

Key:

- check clothes inside + out
- check floor area around me - have I dropped anything?
- try the handle to make sure I have locked doors + drawers.
- IT systems - checking I have logged in and out of secure systems.
- Action ritually repeated 4 times
- Action ritually repeated 2 times.
- Action repeated as I do not trust myself that I checked properly first time round.



People with OCD often ritually repeat actions.

Here I am checking my computer is turned off, 4 times.

I checked my front door handle 4 times. I was feeling anxious and felt compelled to check it was locked another 8 times before I could walk away.

This visualisation shows 57 separate incidences of checking. I recorded 90 in total that day but 33 of those did not relate to my fear of losing information.

When tired, I find it hard to focus on my checks.
Here I had to open my email login page 4 times before I was satisfied it was secure to log in.

I remember feeling frustrated and distressed by this.

ARAB SPRING

Revolution –

Disturbances are suppressed –

The government have concessions –

Number of victims of the Arab revolution

about 10 | about 100 | about 1 000 | about 10 000 | more 50 000