CSL4050: Data Visualization

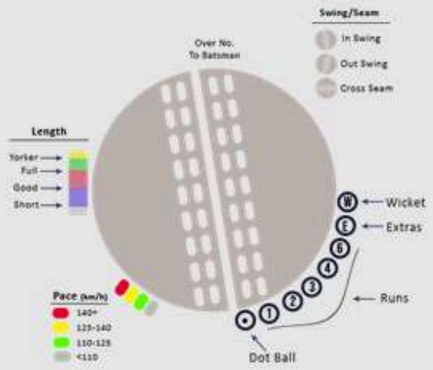
Introduction

Data for Graphics, Design principles, Value for visualization, Categorical, time series, and statistical data graphics, Introduction to Visualization Tools

2ND T20I | 9th July, 2022 | Edgbaston

ENGLAND *vs* INDIA

Analysis of Bhuvneshwar Kumar's bowling spell



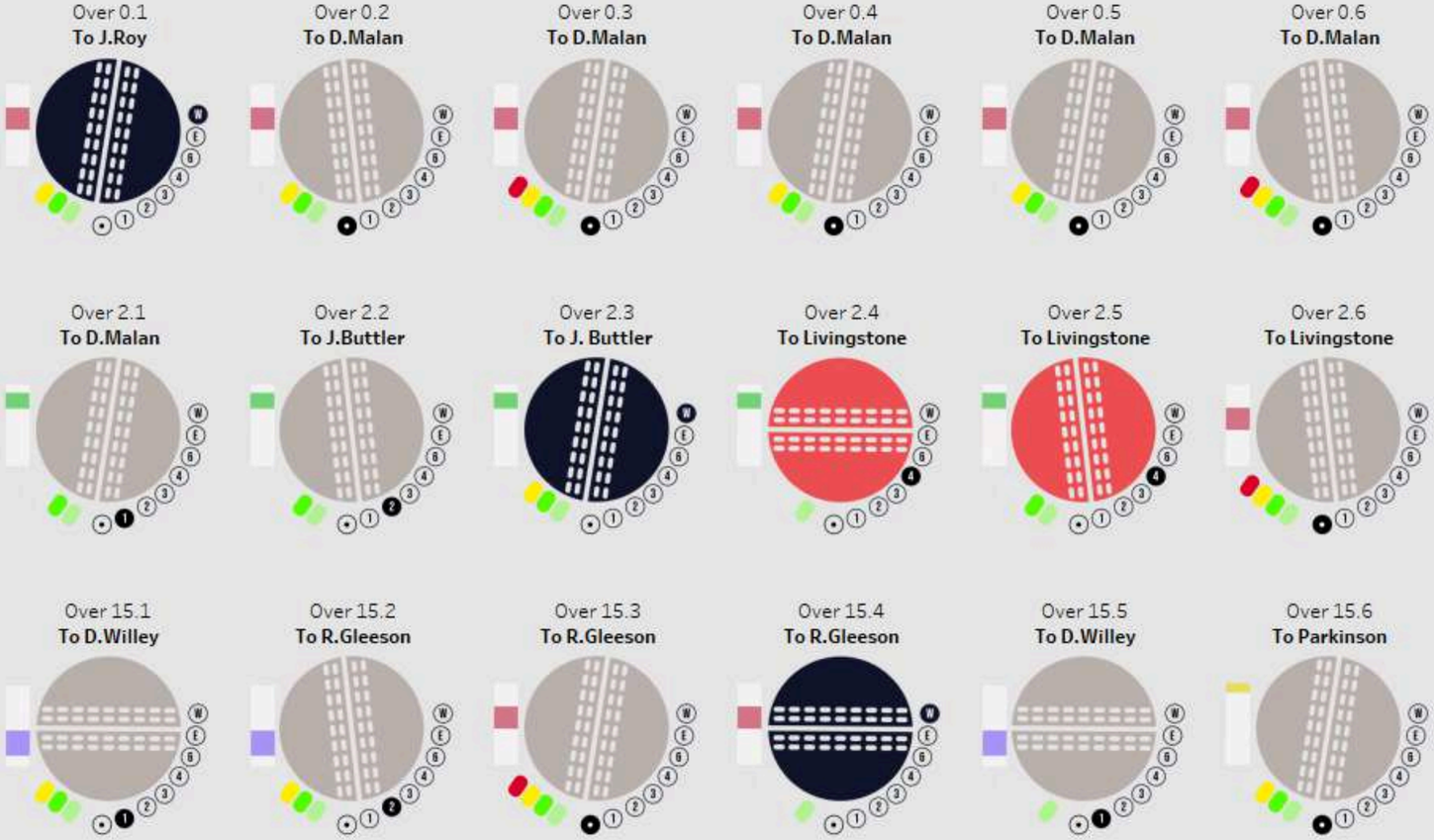
HOW TO READ

Bhuvneshwar Kumar

Right Arm Fast Medium

INDIA

OVERS	3
MAIDENS	1
RUNS	15
WICKETS	3
RUNS PER OVER	5
4S & 6S	2

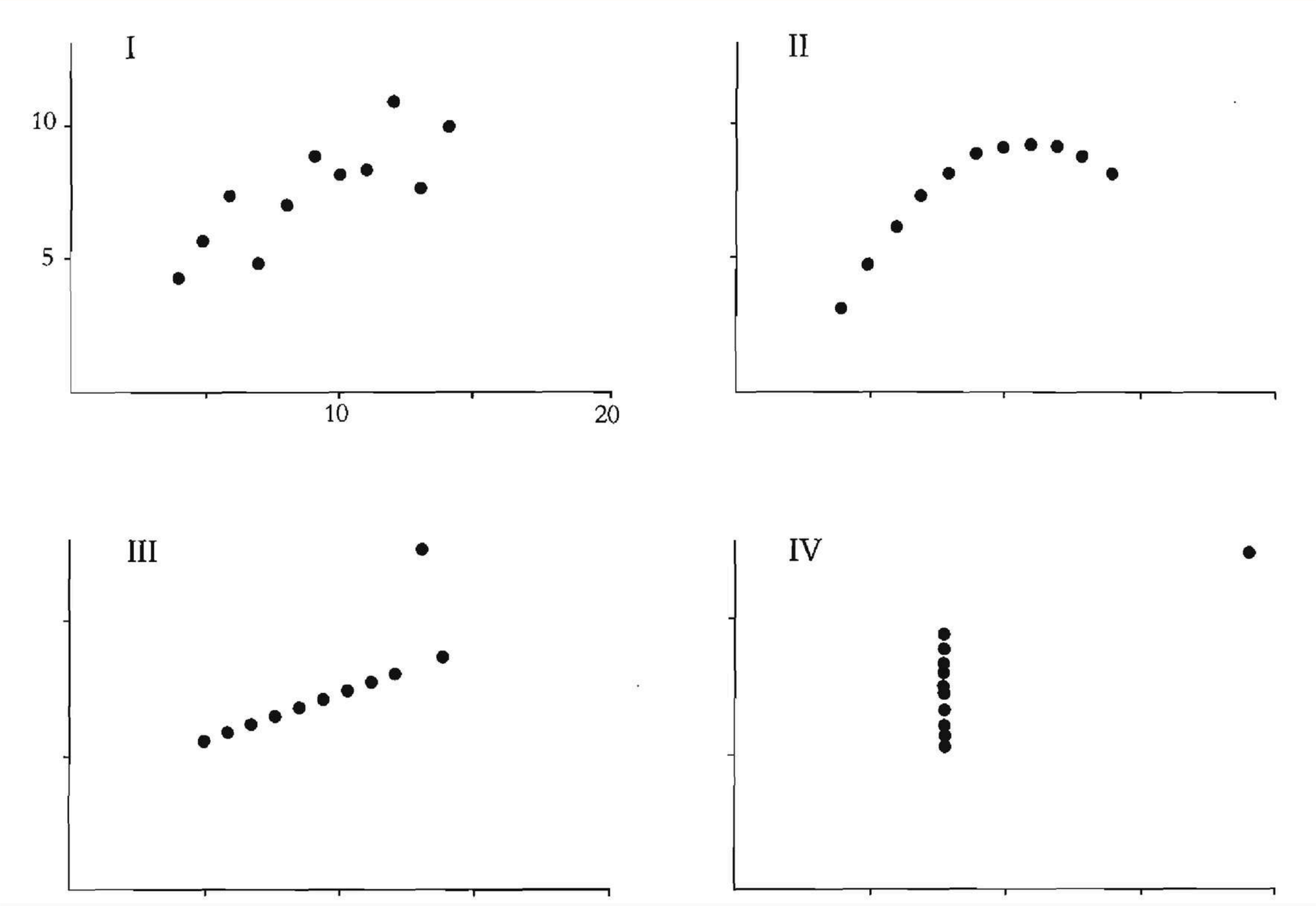


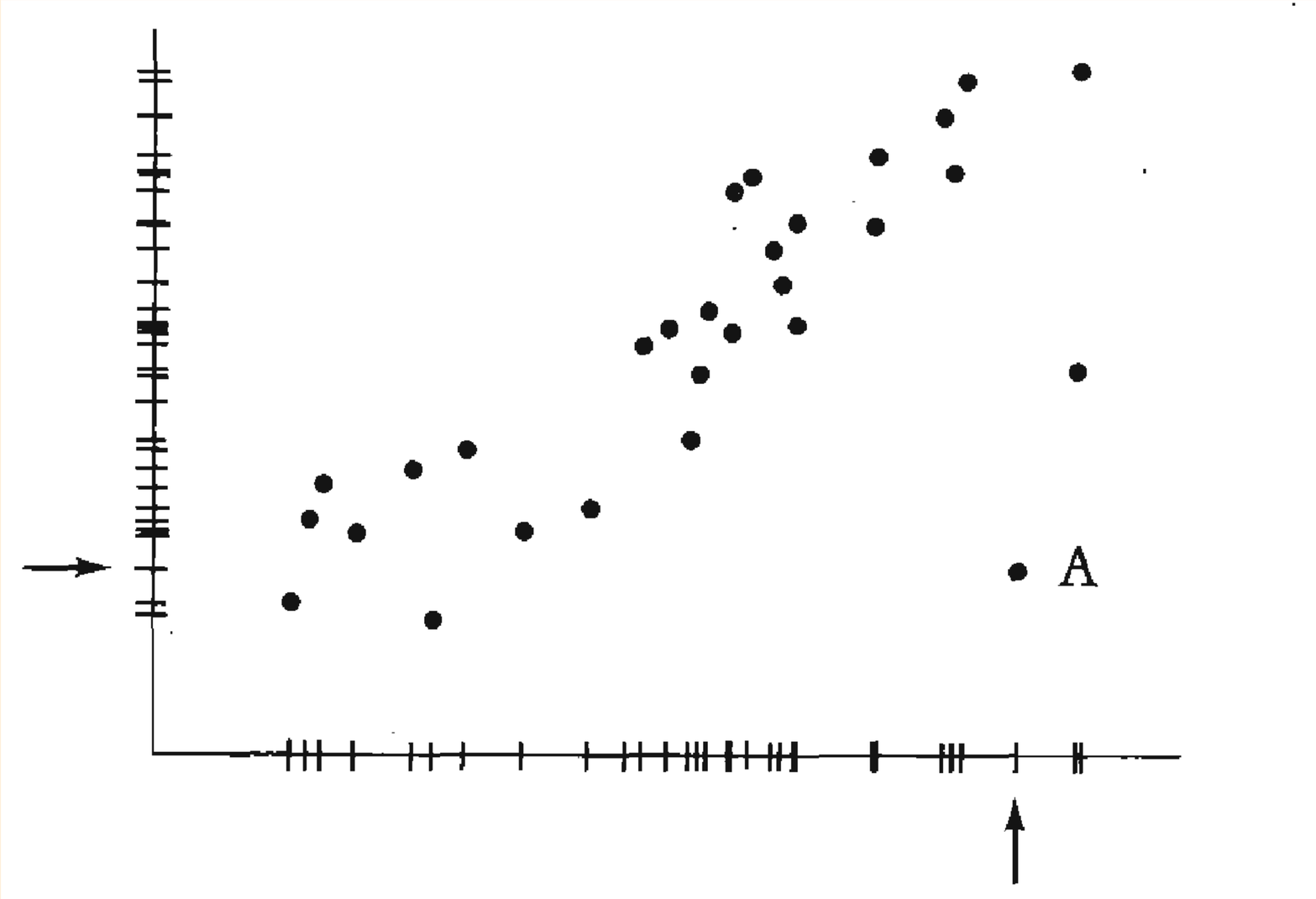
Design: Yash Sakhuja | Data: Designed using Cricbuzz Commentary | Icon Source: Flaticon

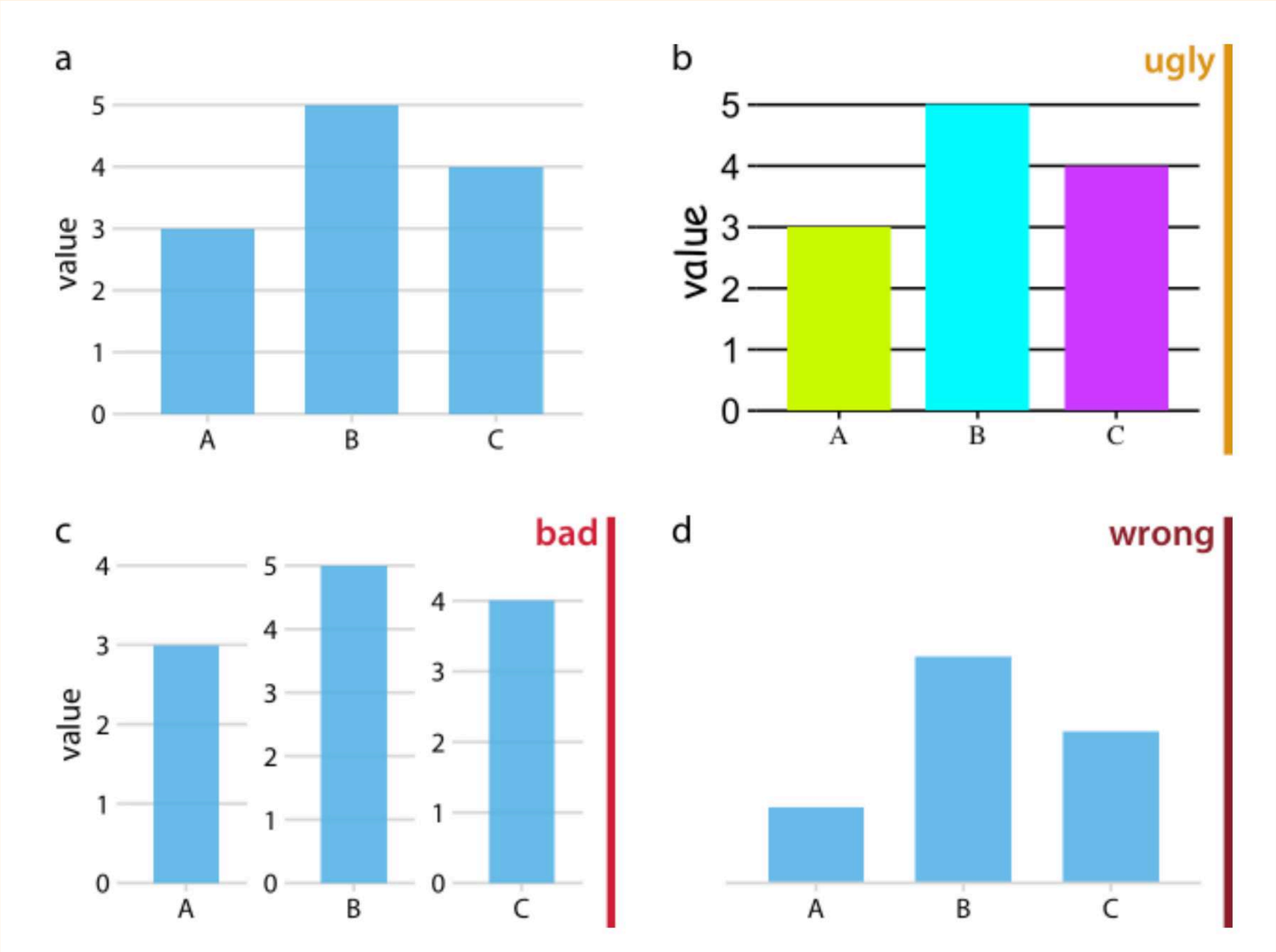
4 Different Data

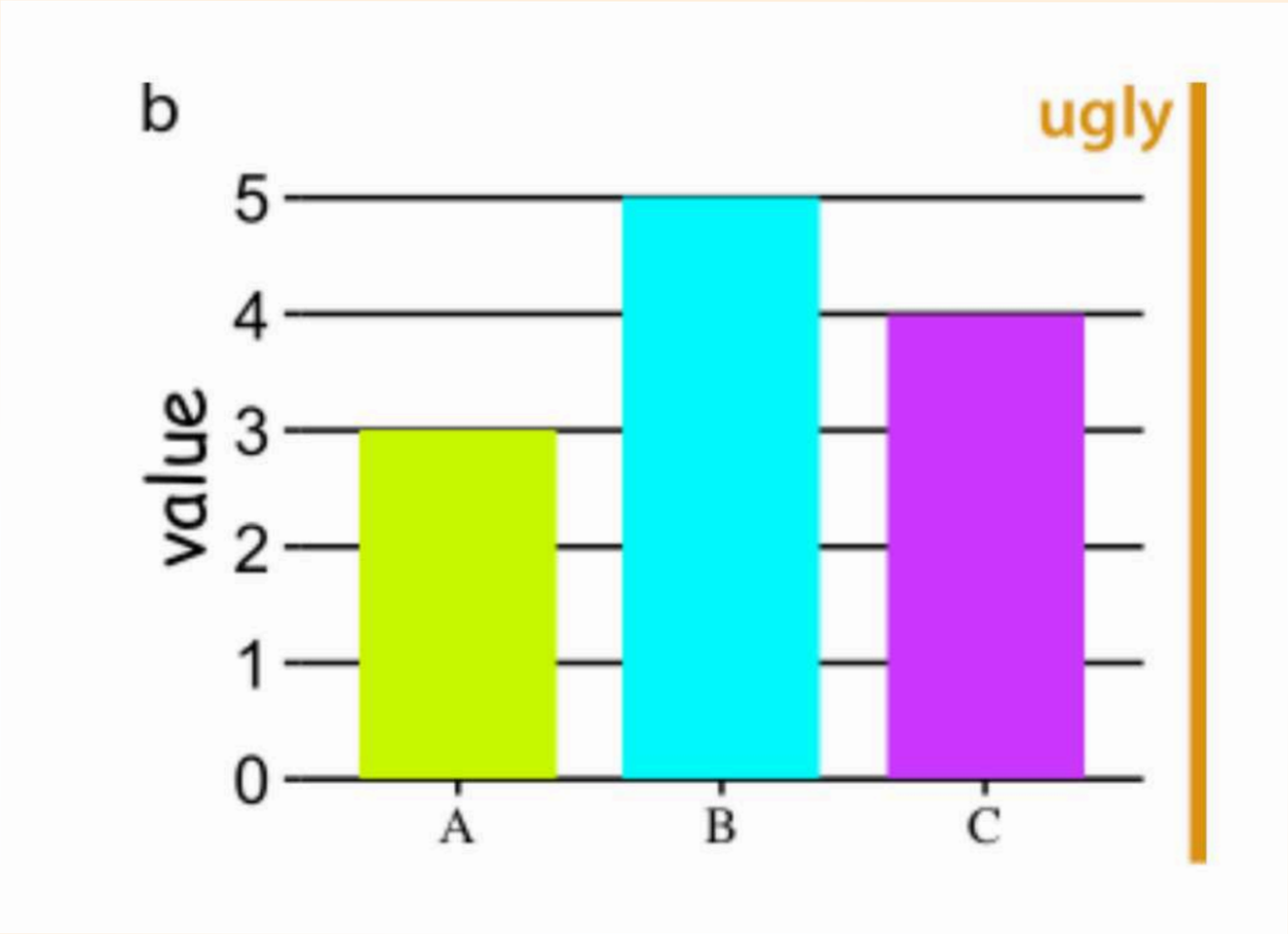
I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

$N = 11$
mean of X's $= 9.0$
mean of Y's $= 7.5$
equation of regression line: $Y = 3 + 0.5X$
standard error of estimate of slope $= 0.118$
$t = 4.24$
sum of squares $X - \overline{X} = 110.0$
regression sum of squares $= 27.50$
residual sum of squares of $Y = 13.75$
correlation coefficient $= .82$
$r^2 = .67$

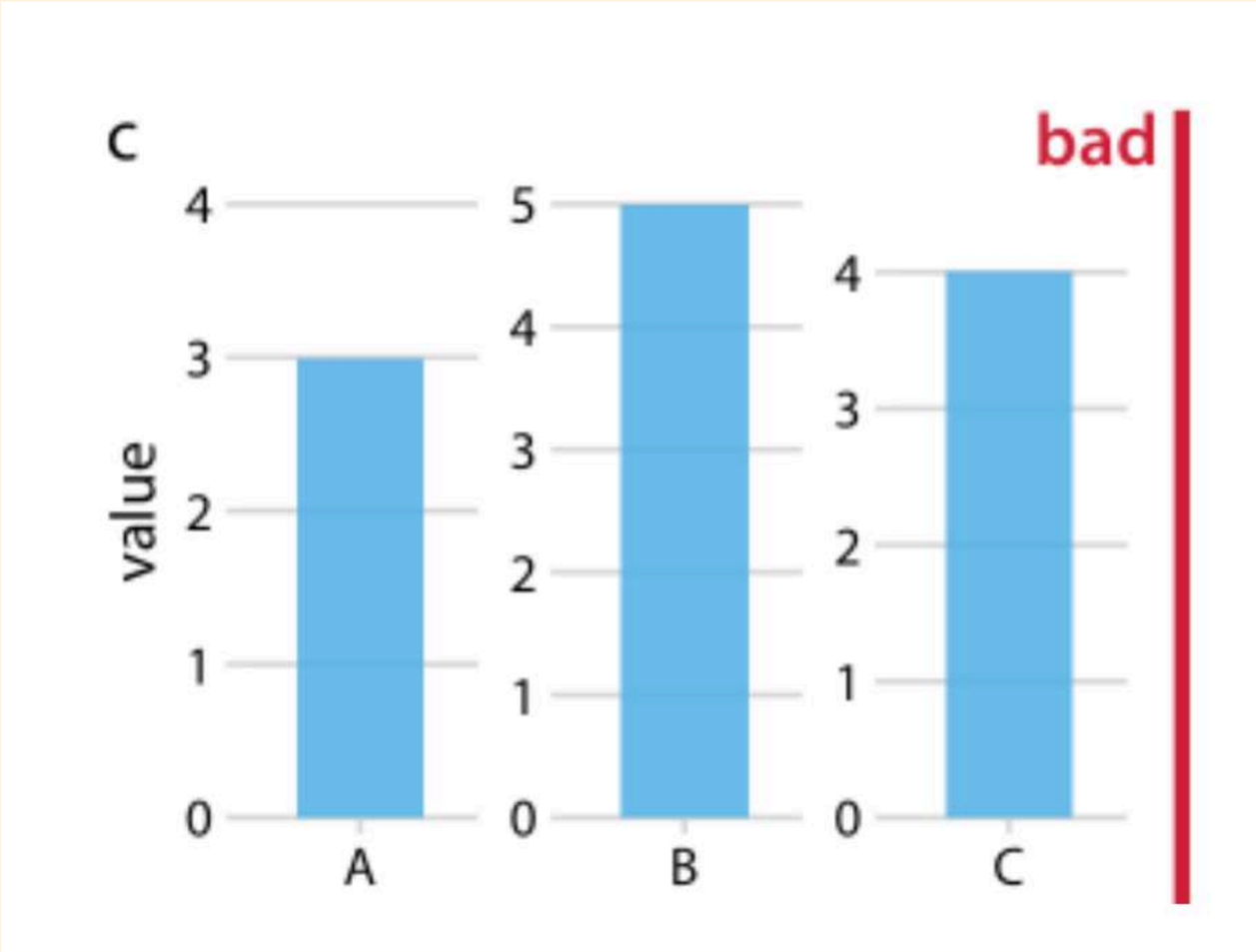

I
10
5
10
20
II
III
IV


A


a
value
5
4
3
2
1
0
A
B
C
b
ugly
value
5
4
3
2
1
0
A
B
C
c
bad
value
4
3
2
1
0
A
5
4
3
2
1
0
B
4
3
2
1
0
C
d
wrong
A
B
C

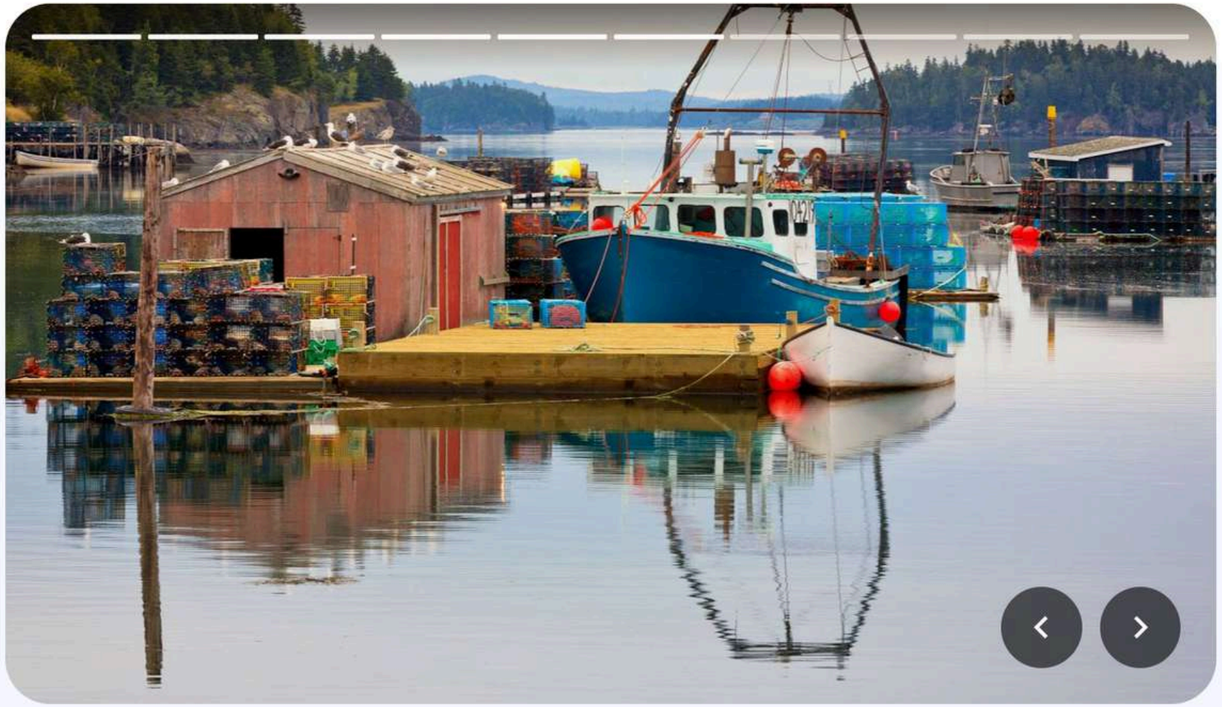


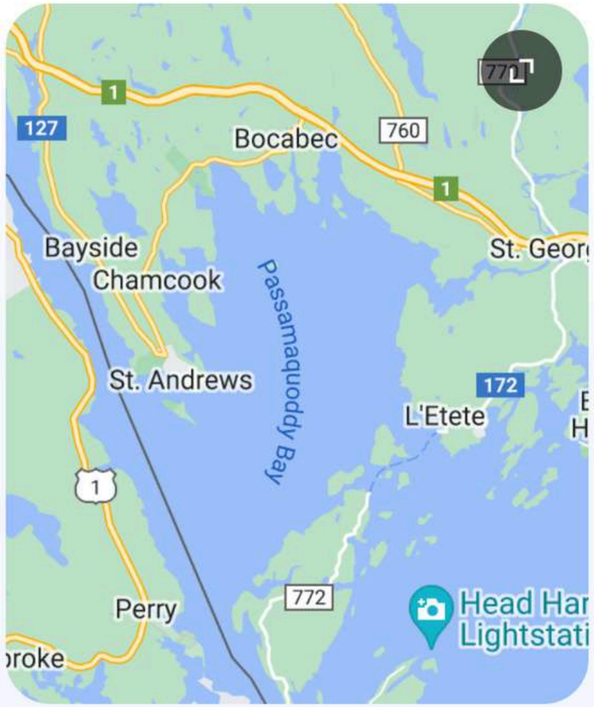
A figure that has aesthetic problems but otherwise is clear and informative



A figure that has problems related to perception; it may be unclear, confusing, overly complicated, or deceiving

A figure that has problems related to mathematics; it is objectively incorrect

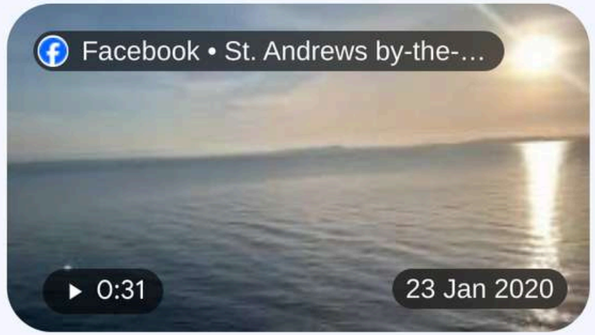




Weather

Sun	Mon	Tue
-5°	-3°	-4°

Weather data



See detailed insights & Compare multiple related Papers for :
“Passamaquoddy Bay”

Compare insights

Wikipedia
https://en.wikipedia.org › wiki › Passamaquoddy_Bay

Passamaquoddy Bay

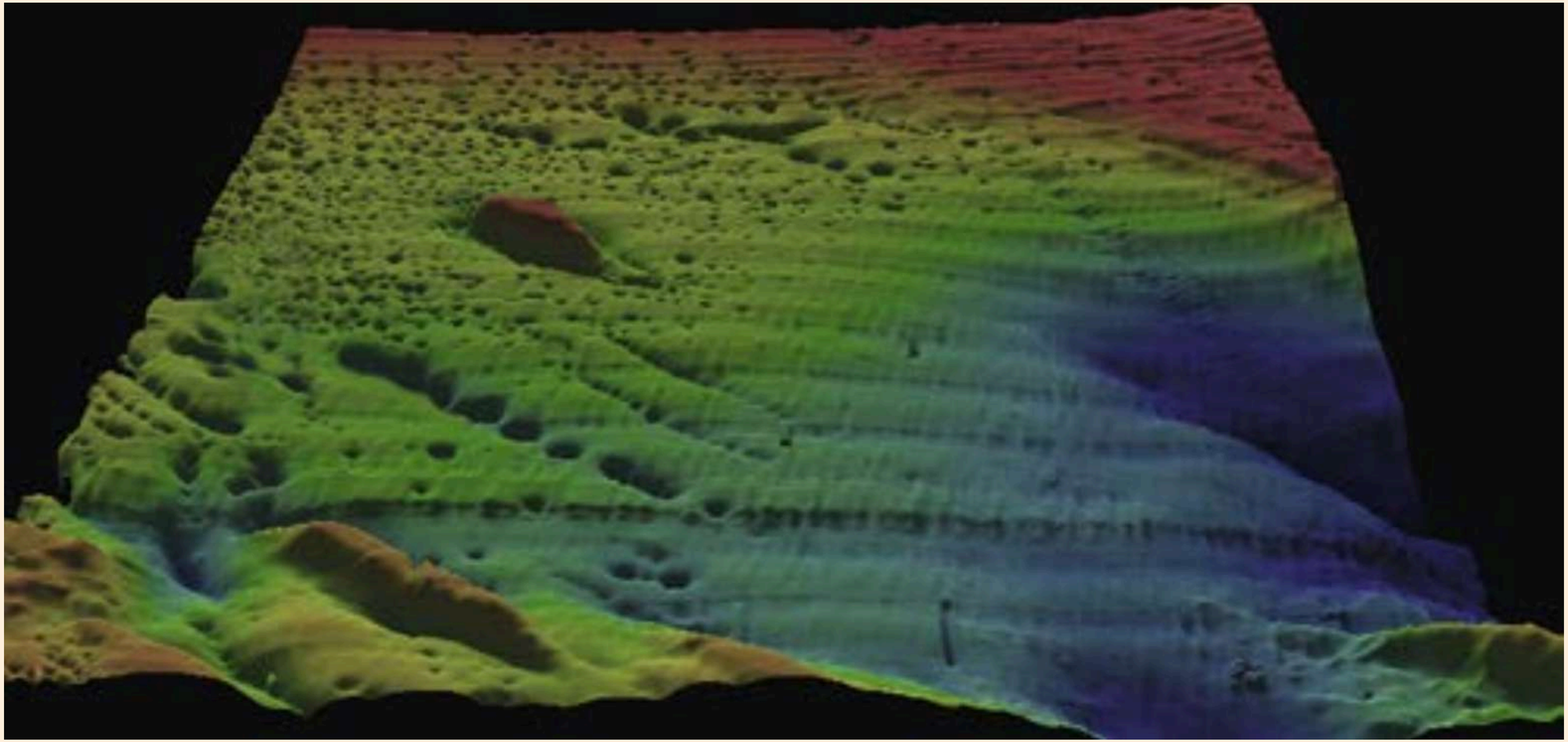
Passamaquoddy Bay (French: Baie de Passamaquoddy) is an inlet of the Bay of Fundy, between the U.S. state of Maine and the Canadian province of New ...

Overview · Smuggling

About

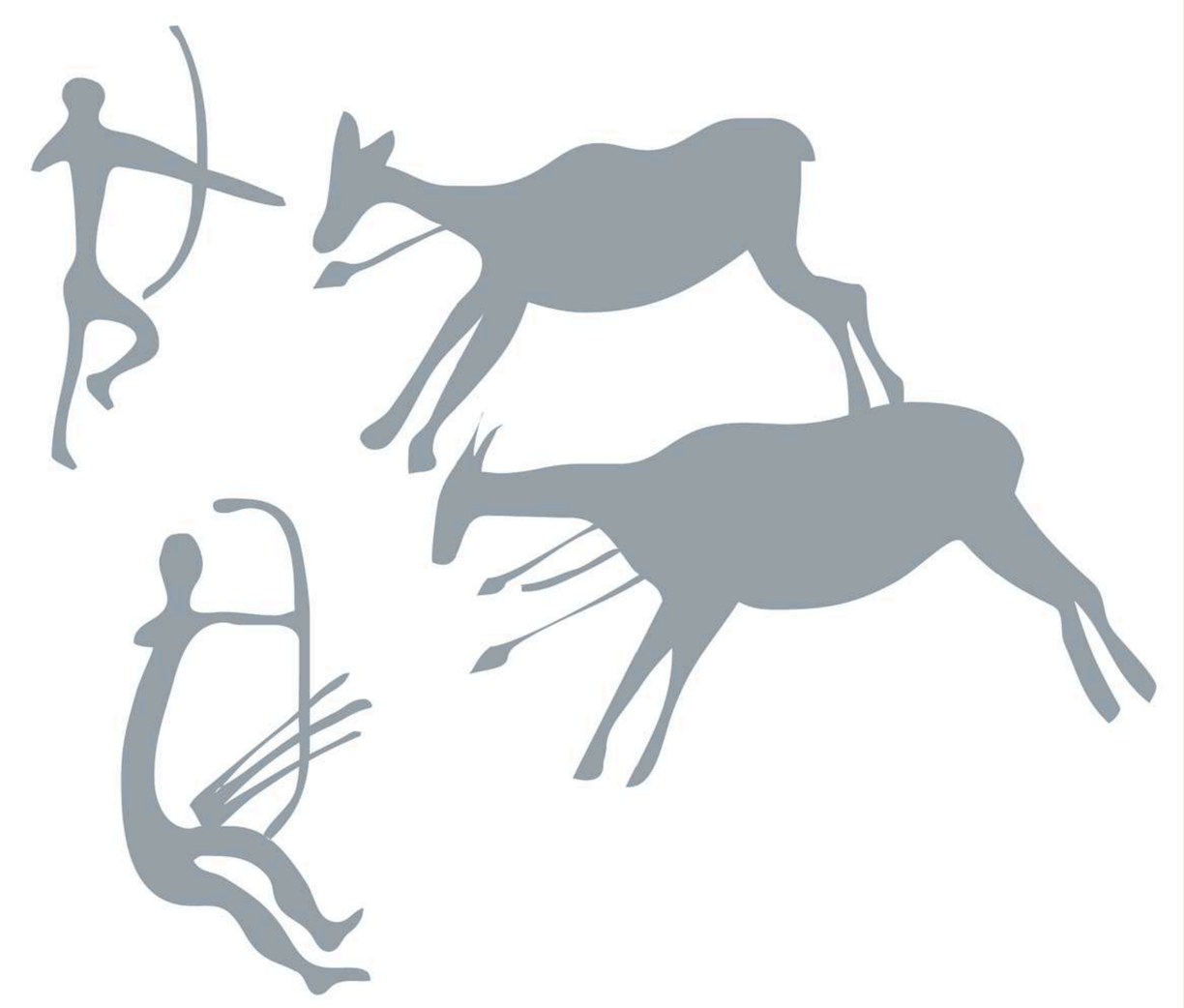
Passamaquoddy Bay is an inlet of the Bay of Fundy, between the U.S. state of Maine and the Canadian province of New Brunswick, at the mouth of the St. Croix River. Most of the bay lies within Canada, with its western shore bounded by Washington County, Maine. Wikipedia

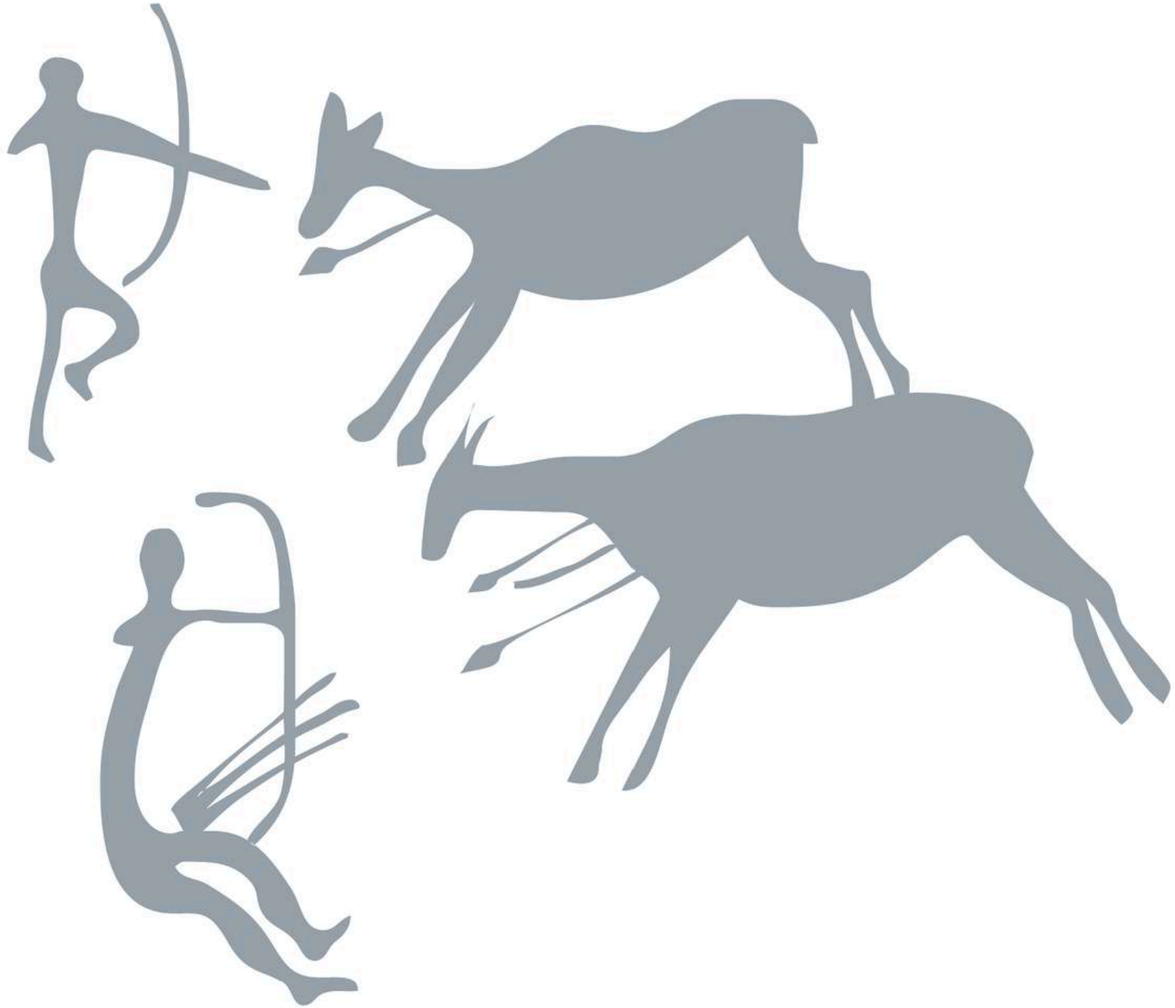
Feedback

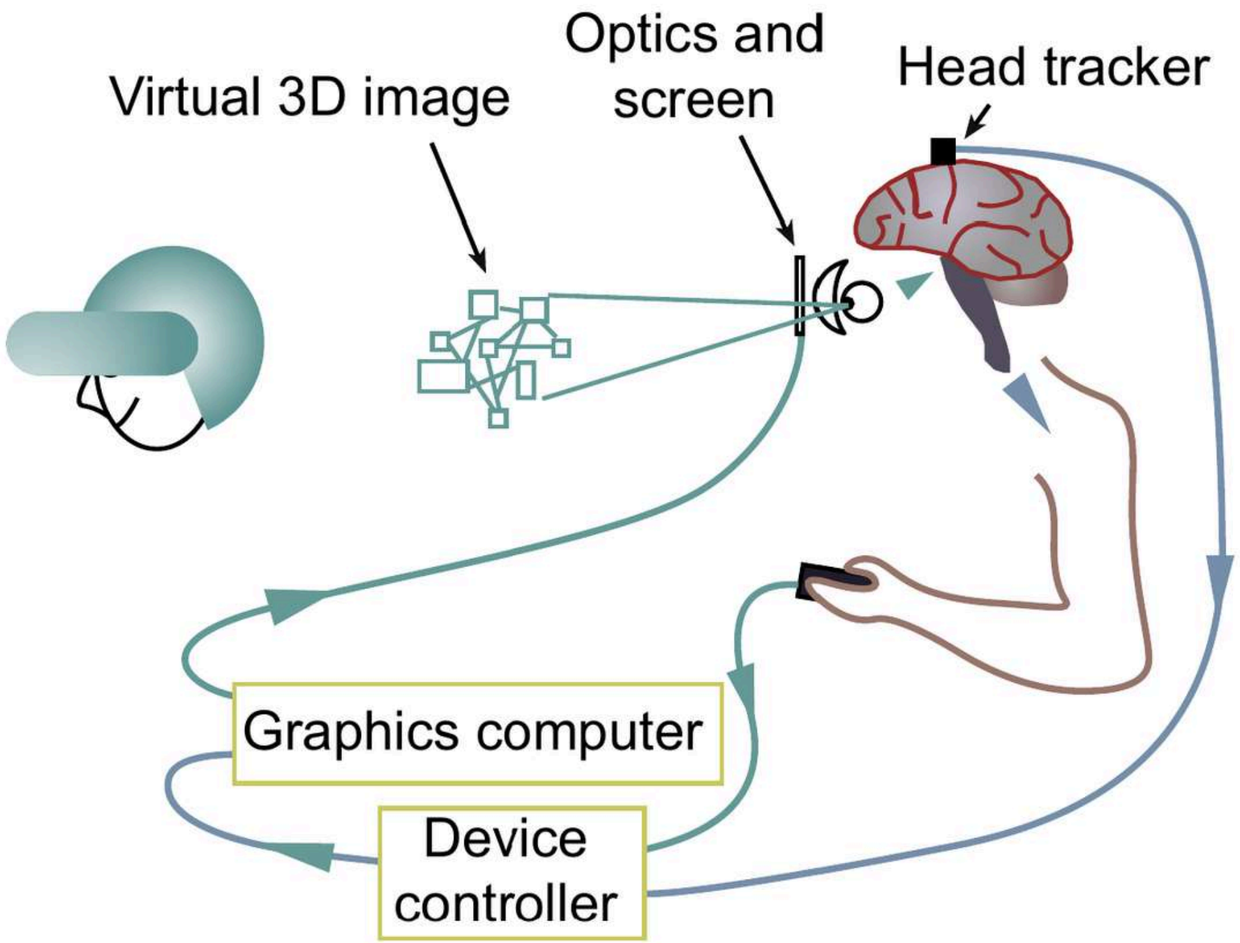


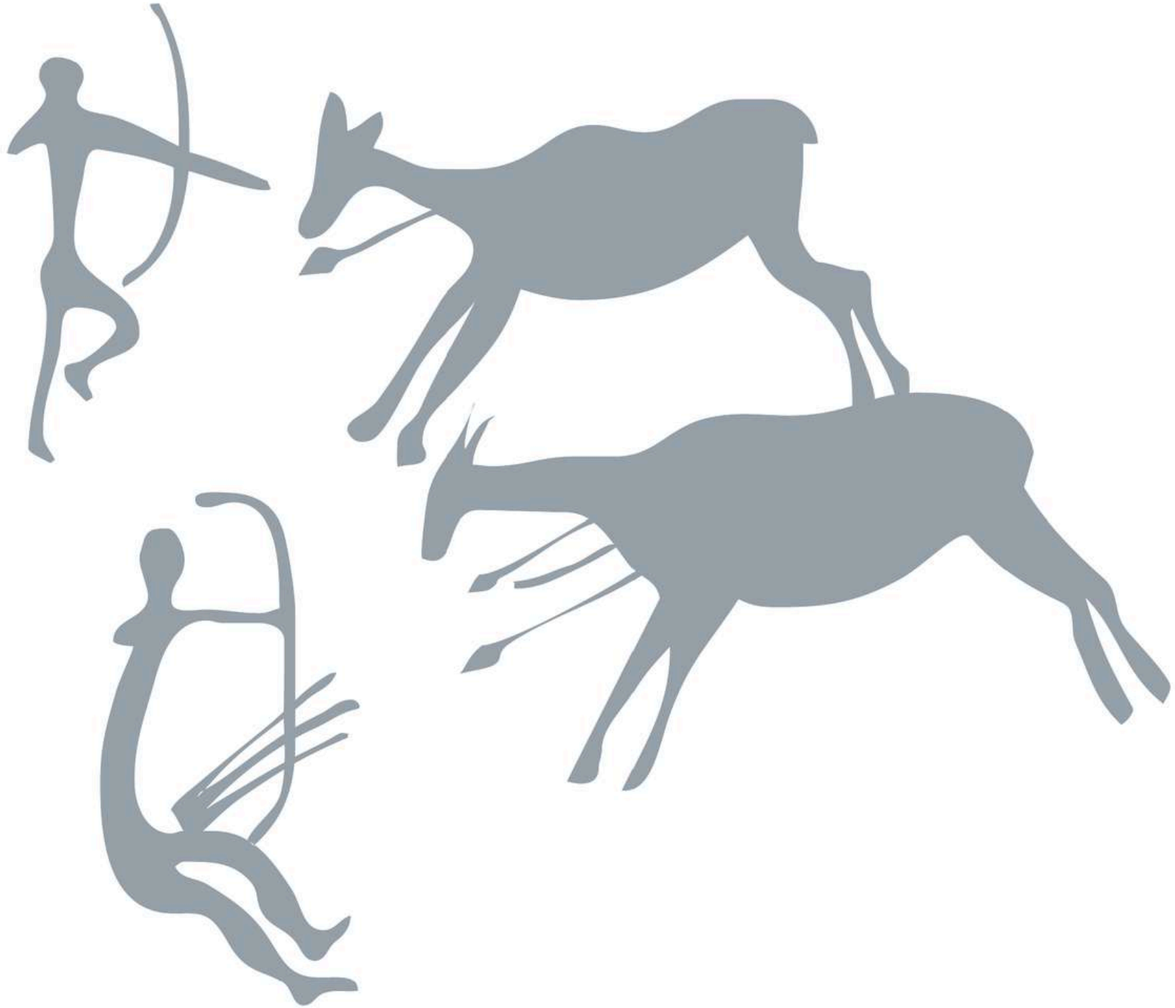
Passamoquoddy Bay visualization. (Data courtesy of the Canadian Hydrographic Service.)

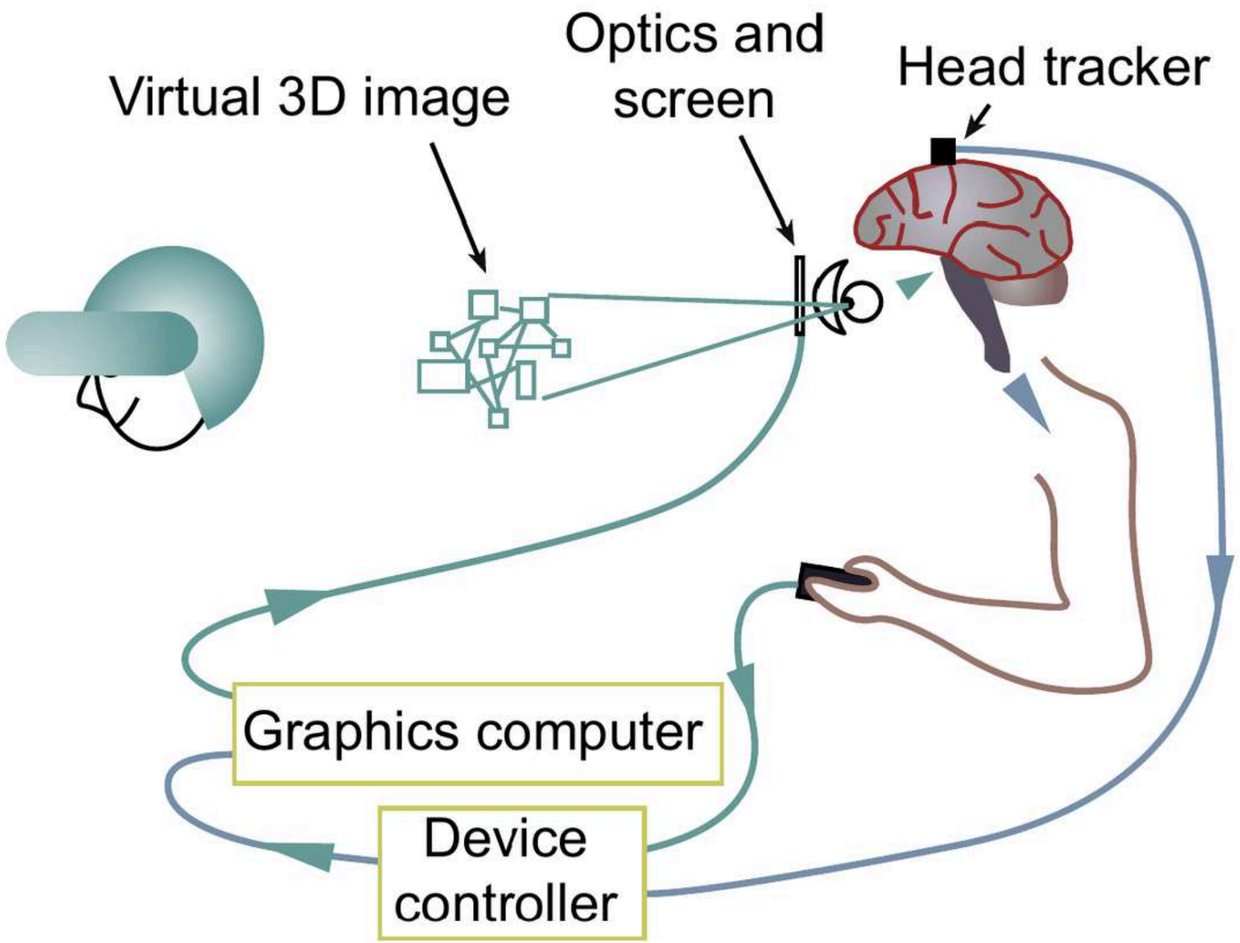
Visual Language





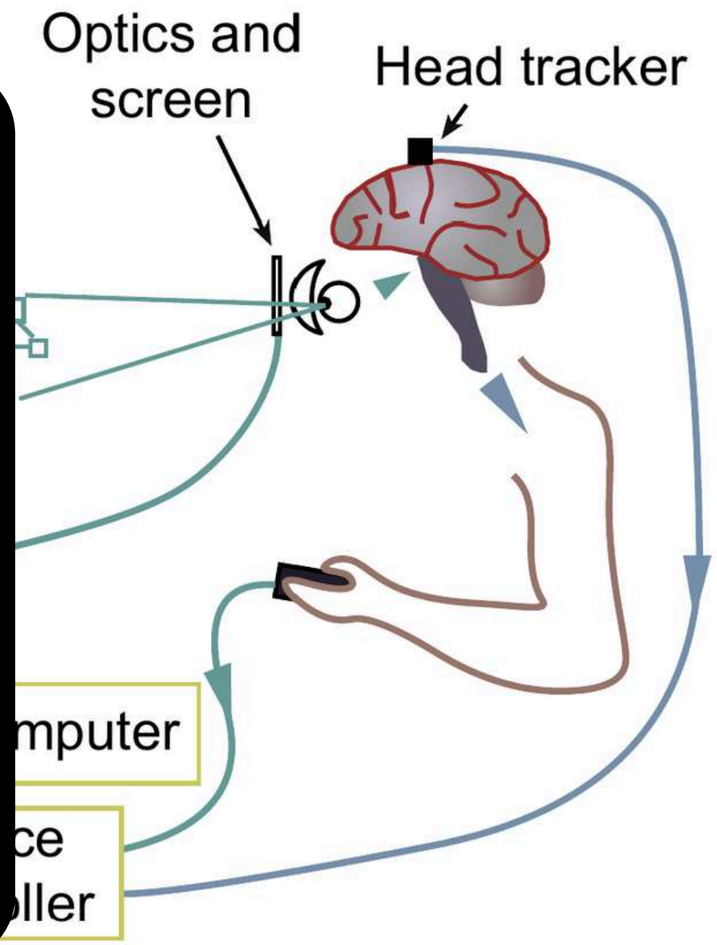

Virtual 3D image
Optics and screen
Head tracker
Graphics computer
Device controller





Interaction between a person and a computer using VR

$$x \propto \int_{1}^{\infty} \varpi \oint_{\lambda} \left\| \frac{\prod \lambda_i}{\oint \Psi} \right\|$$



Interaction between a person and a computer using VR

The value of a good visualization is that it lets us see patterns in data

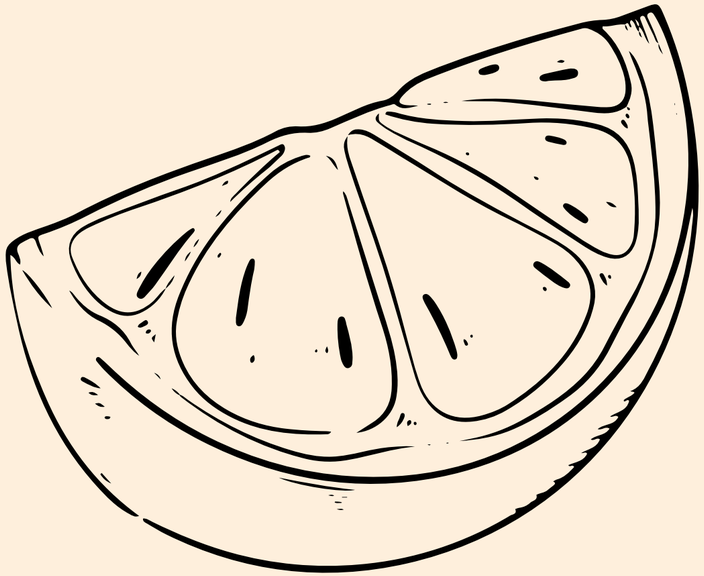
Therefore the science of pattern perception can provide a basis for design decisions, but the claim that visualization can be based on science may be disputed.

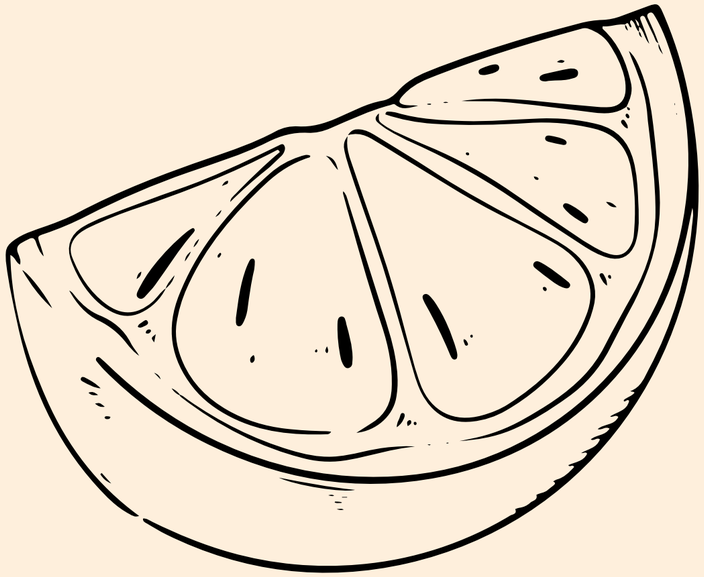
“

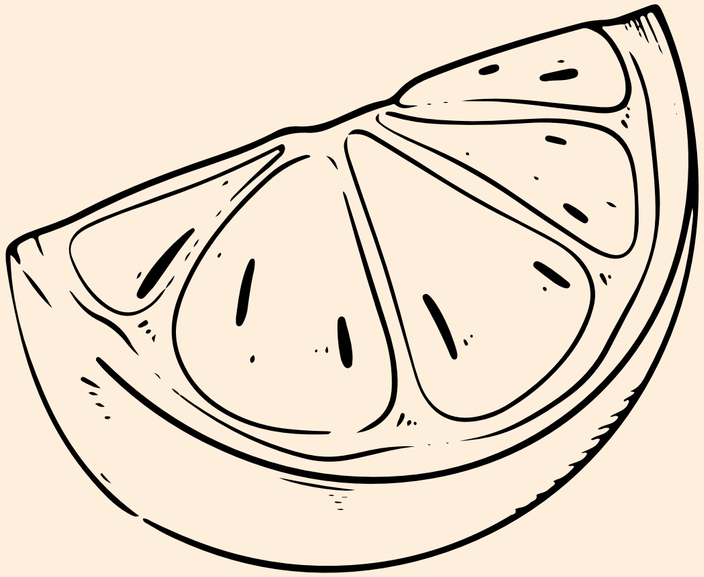
A Bush Negro woman turned a photograph this way and that, in attempting to make sense out of the shadings of gray on the piece of paper she held

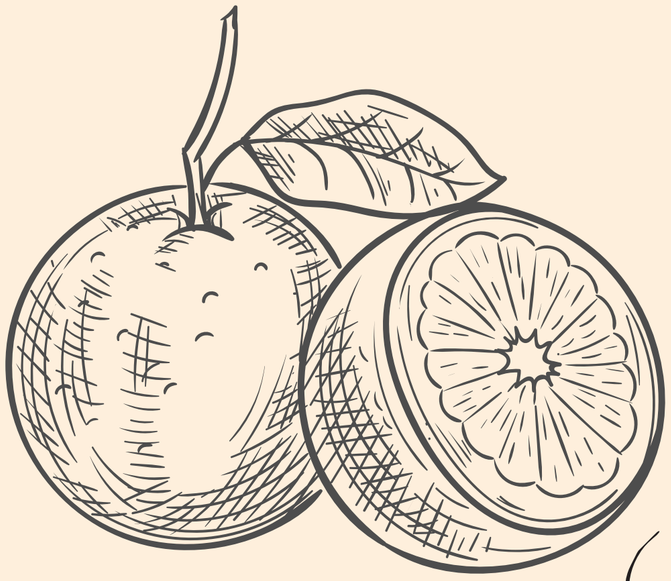
Herskovits, 1948

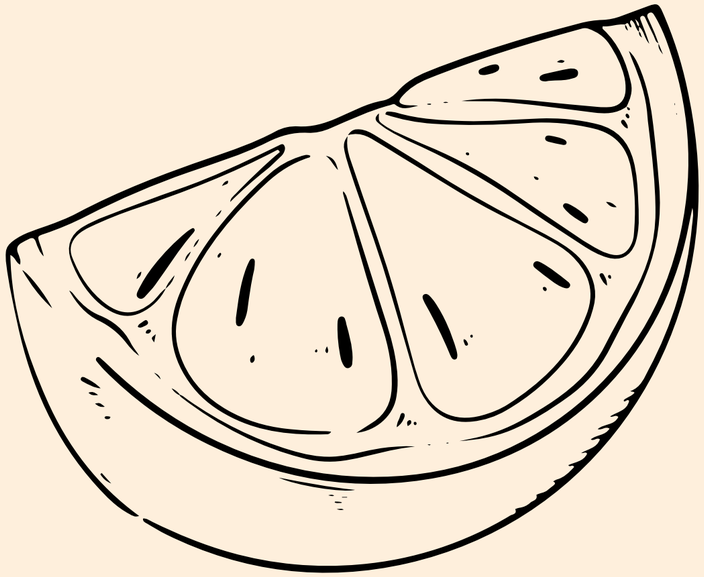
Kennedy (1974) argued that these people are amazed by the technology rather than unable to interpret the picture. After all, a photograph is a remarkable artifact.

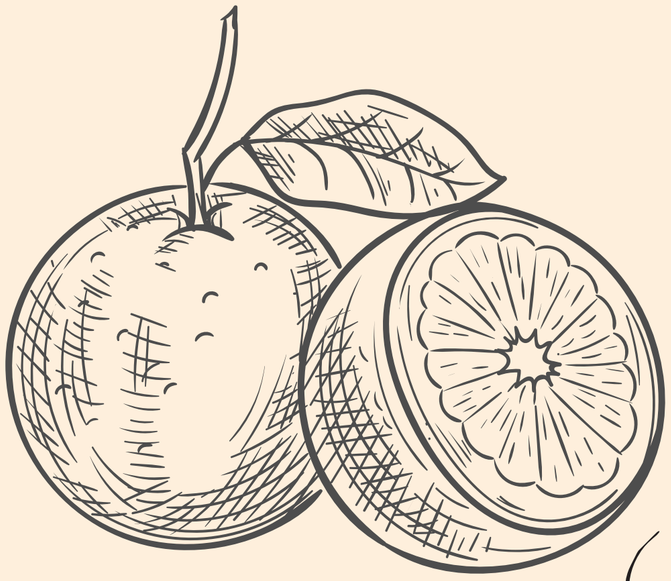


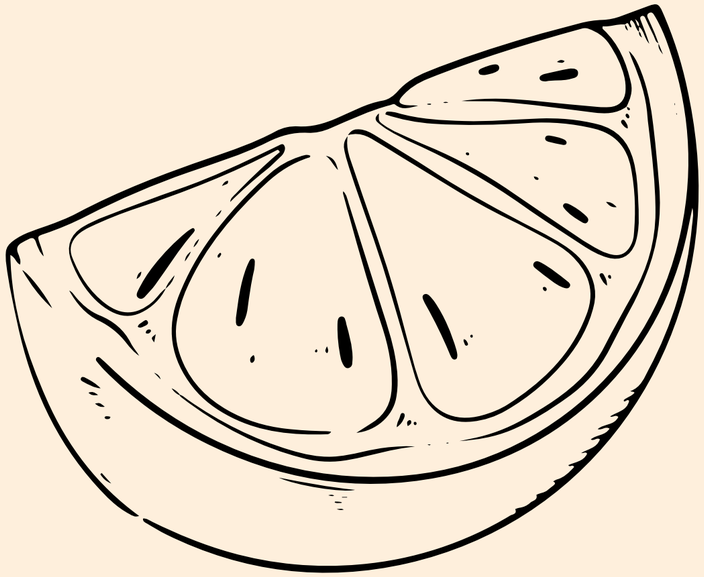


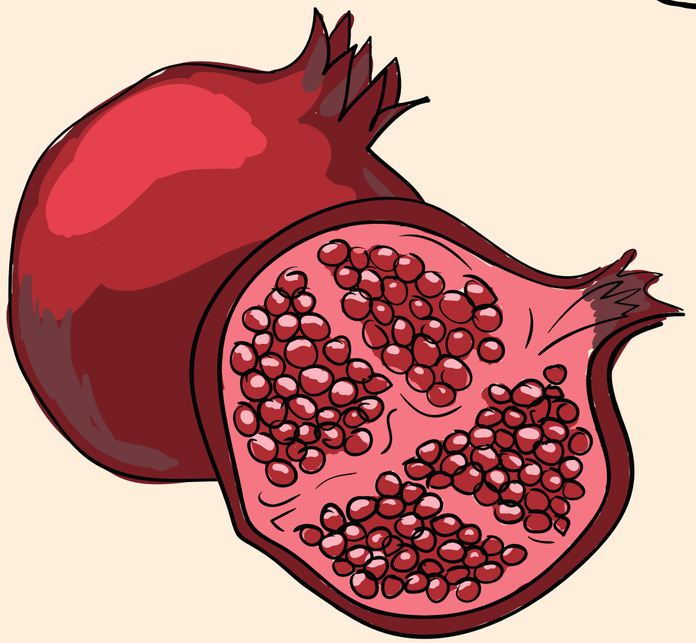


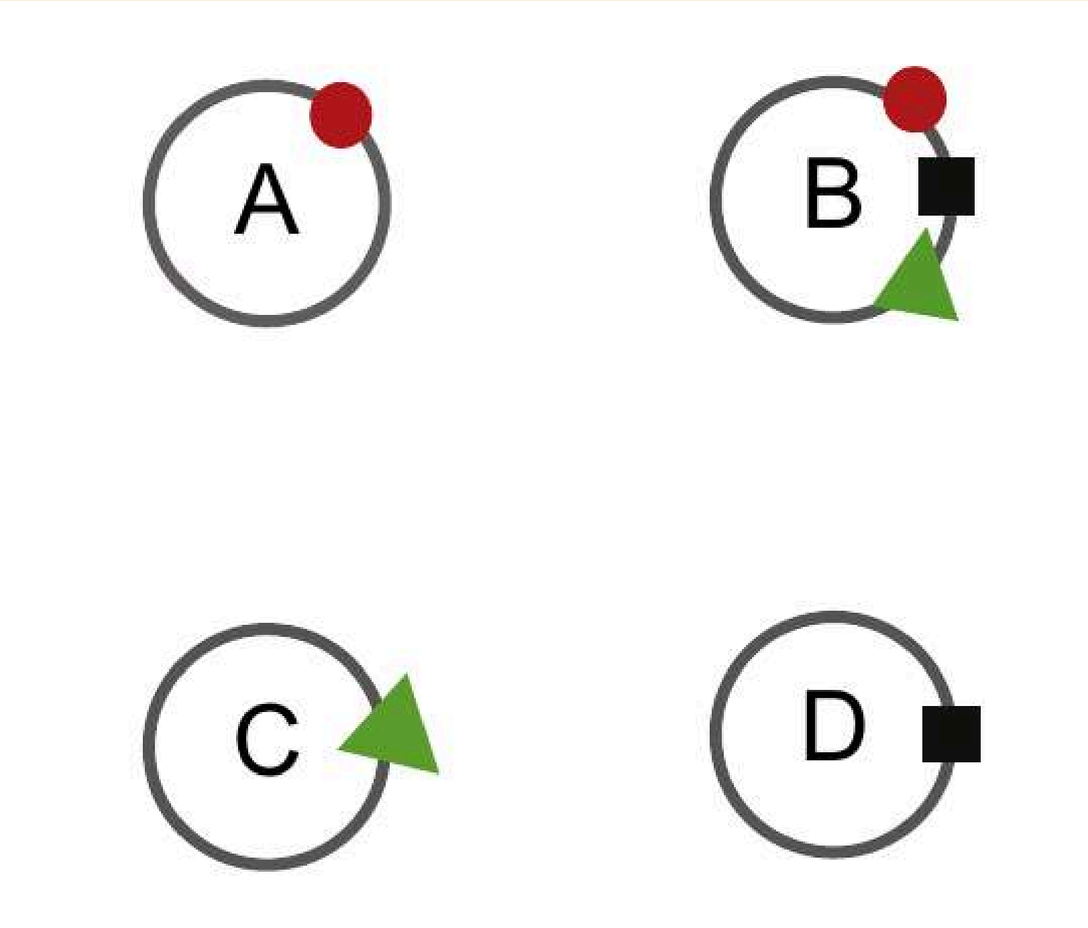


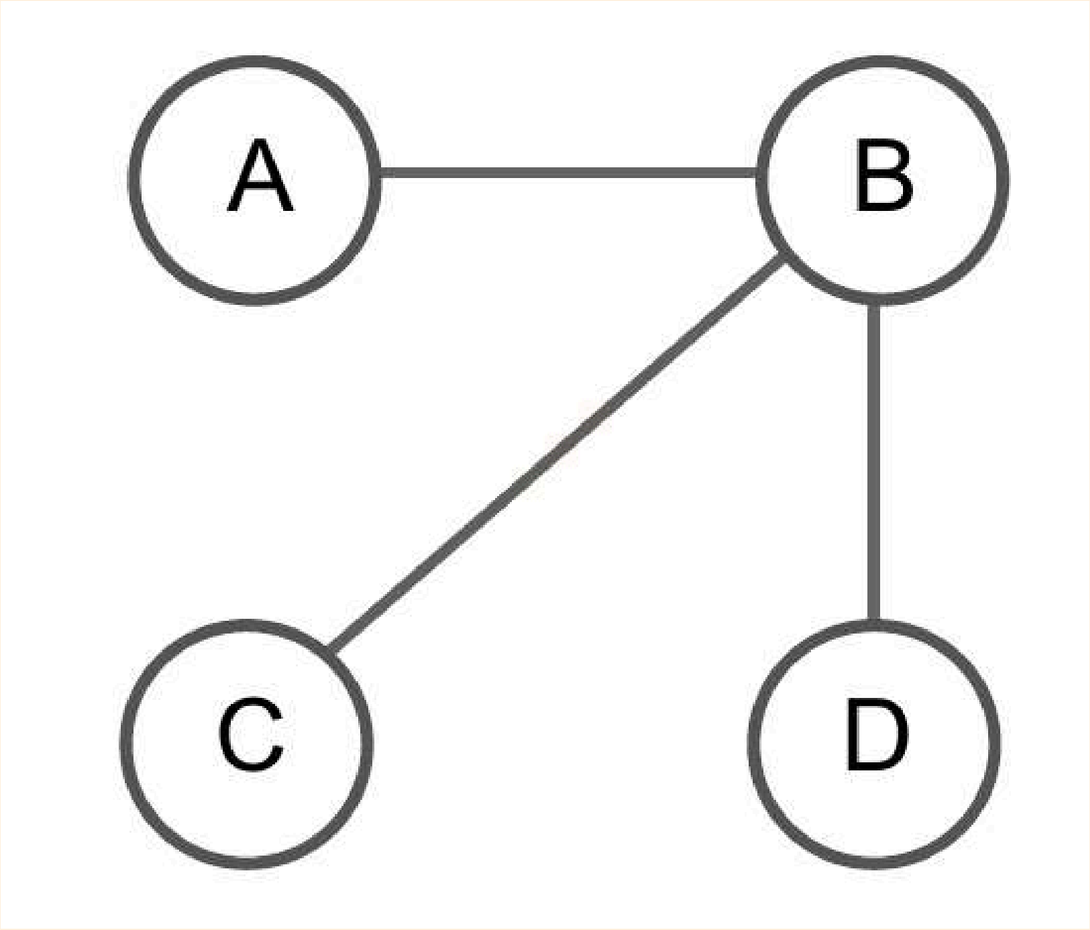




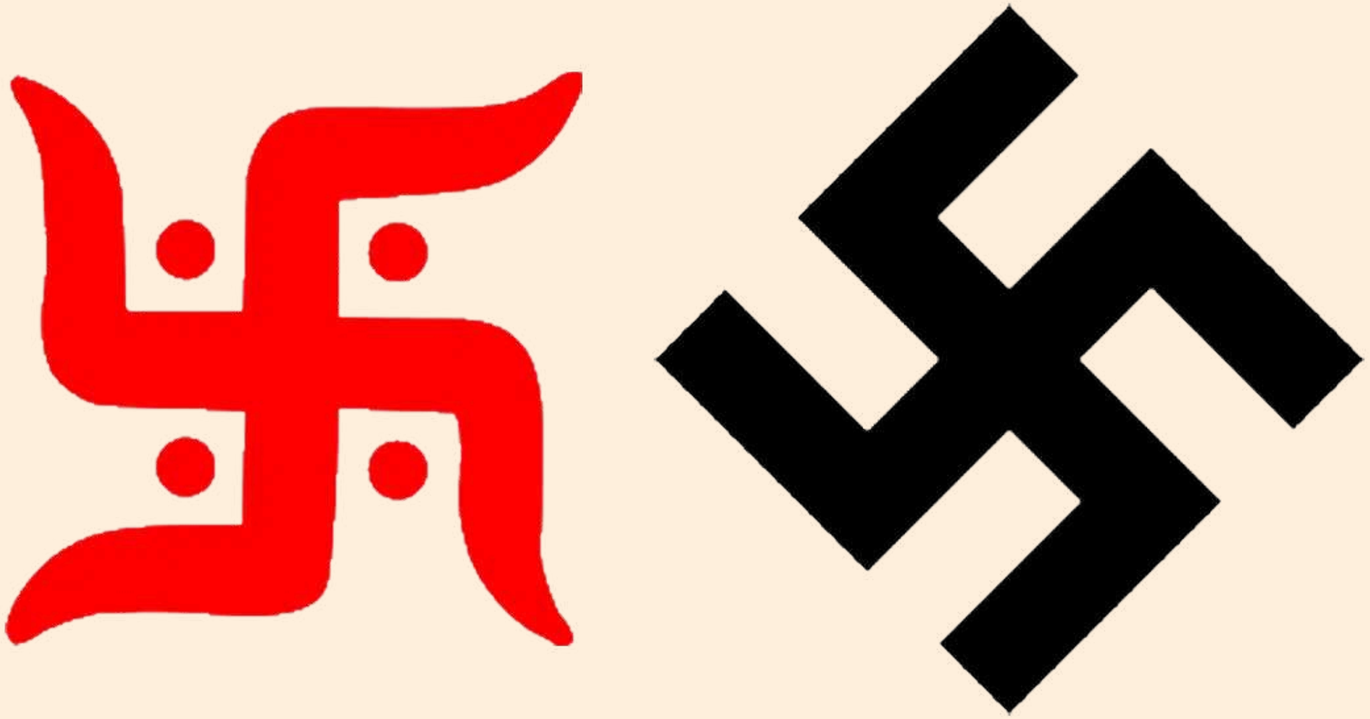





A
B
C
D



Line Drawing is Intuitive to Represent relationship



개

Пас

Inja

개 **Korean**

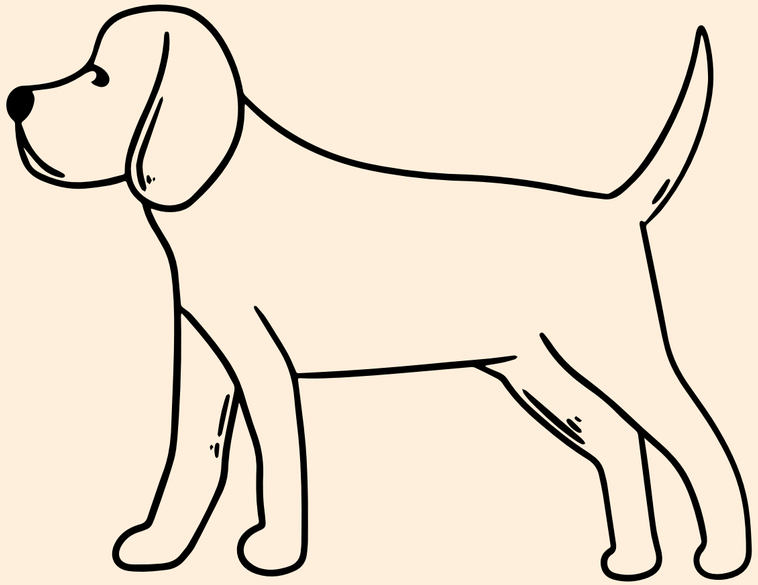
Пас **Serbian**

Inja **Zulu**

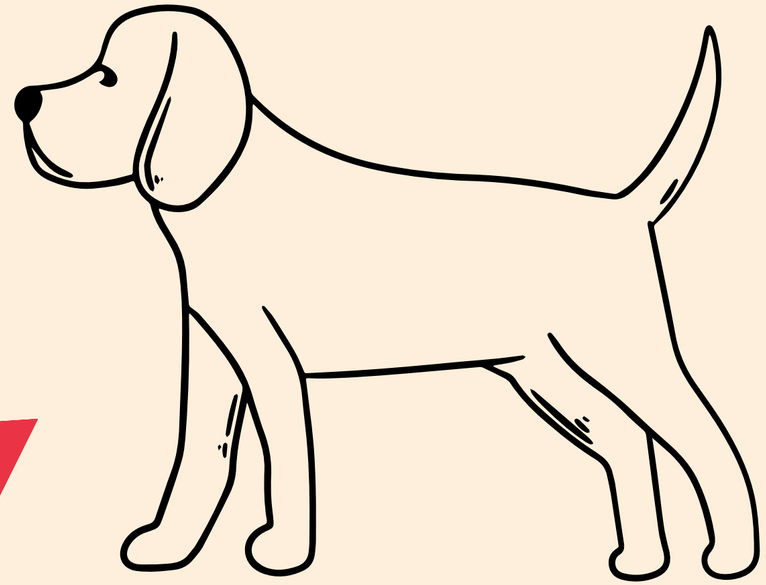
개 **Korean**

Пас **Serbian**

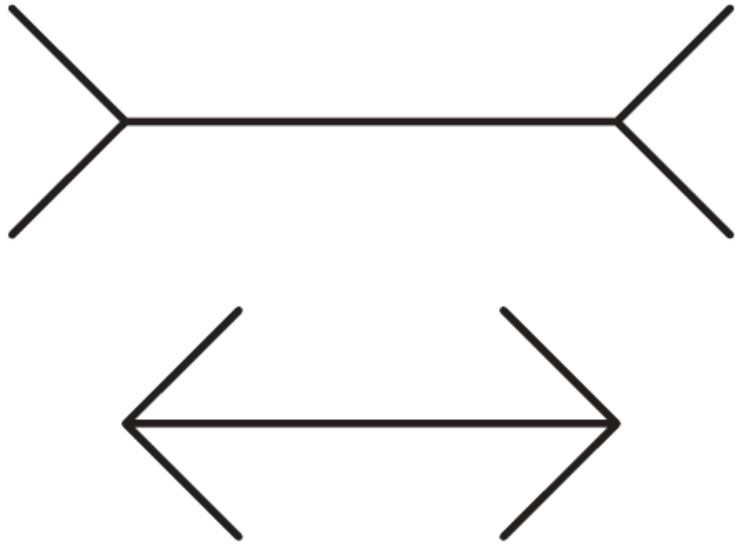
Inja **Zulu**

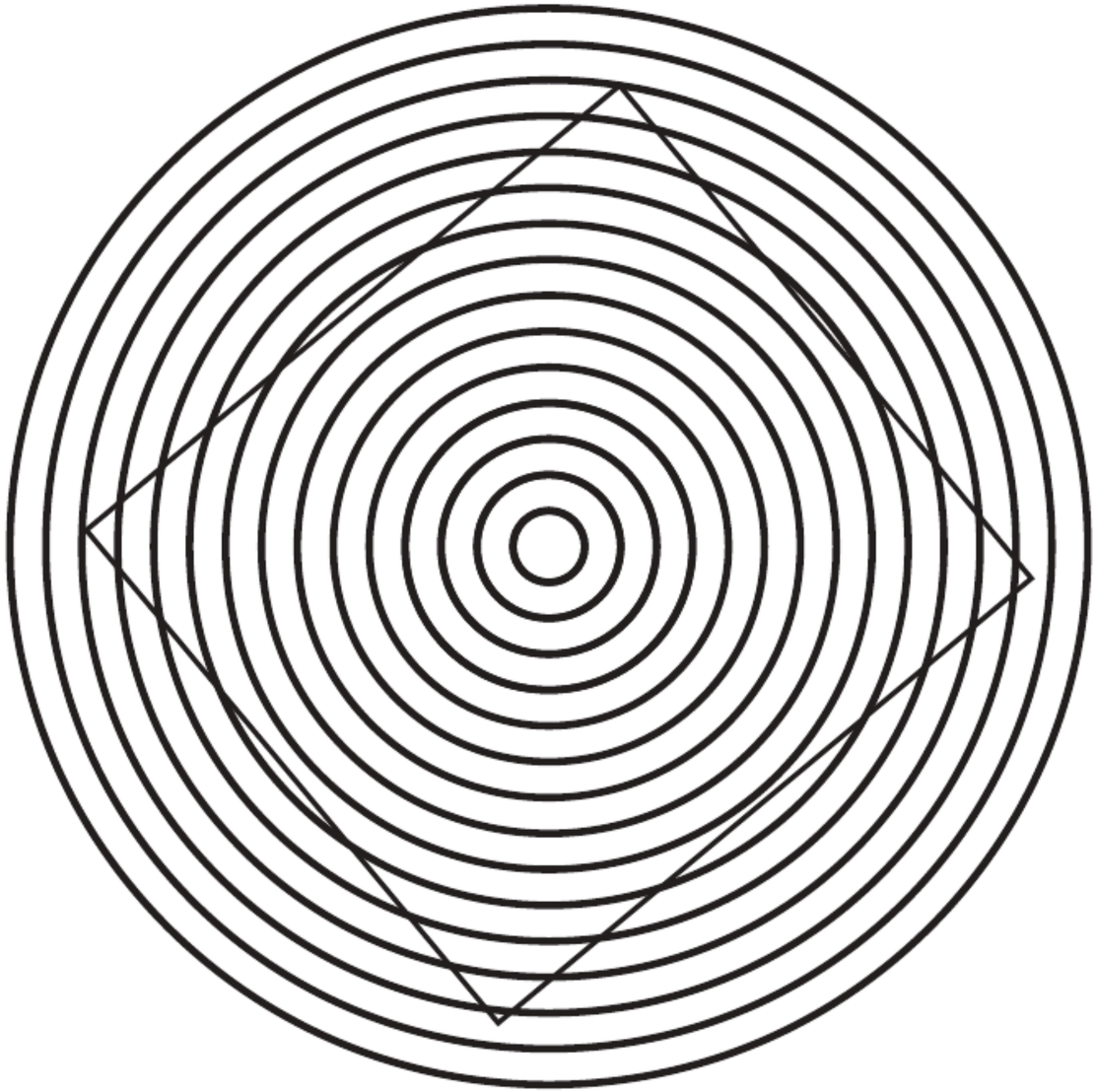


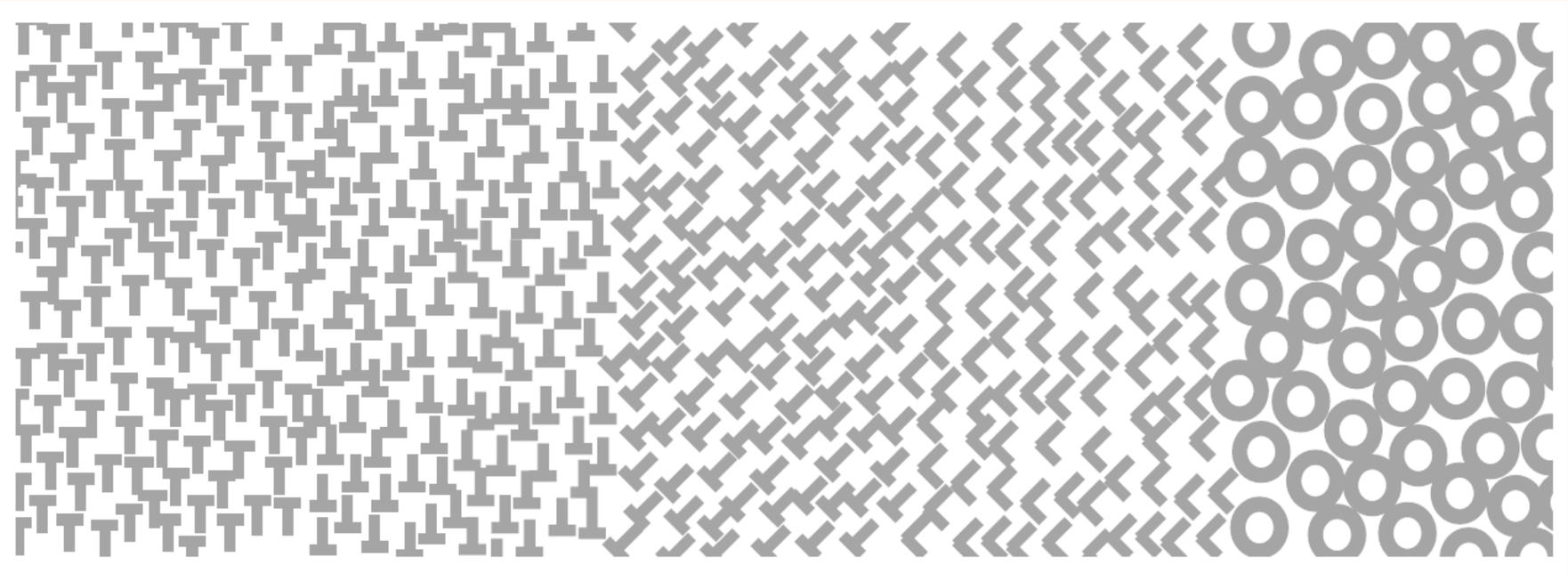
	개	Korean
Arbitrary	Пас	Serbian
	Inja	Zulu



Sensory



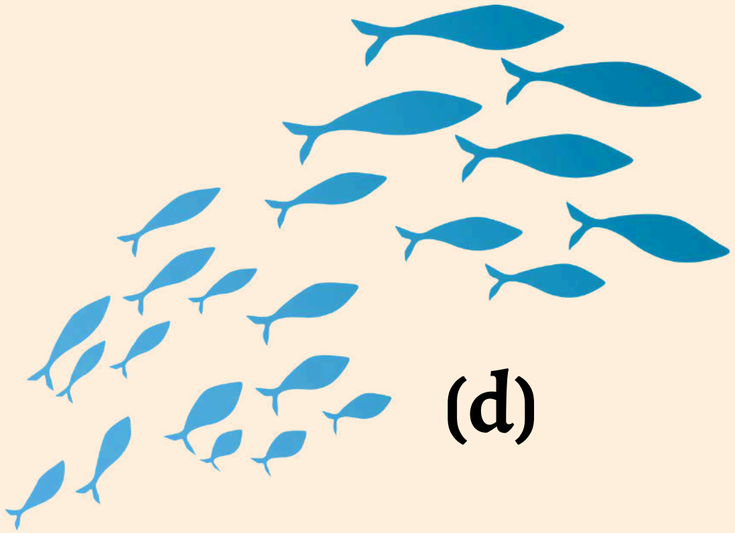




(a)

(b)

(c)

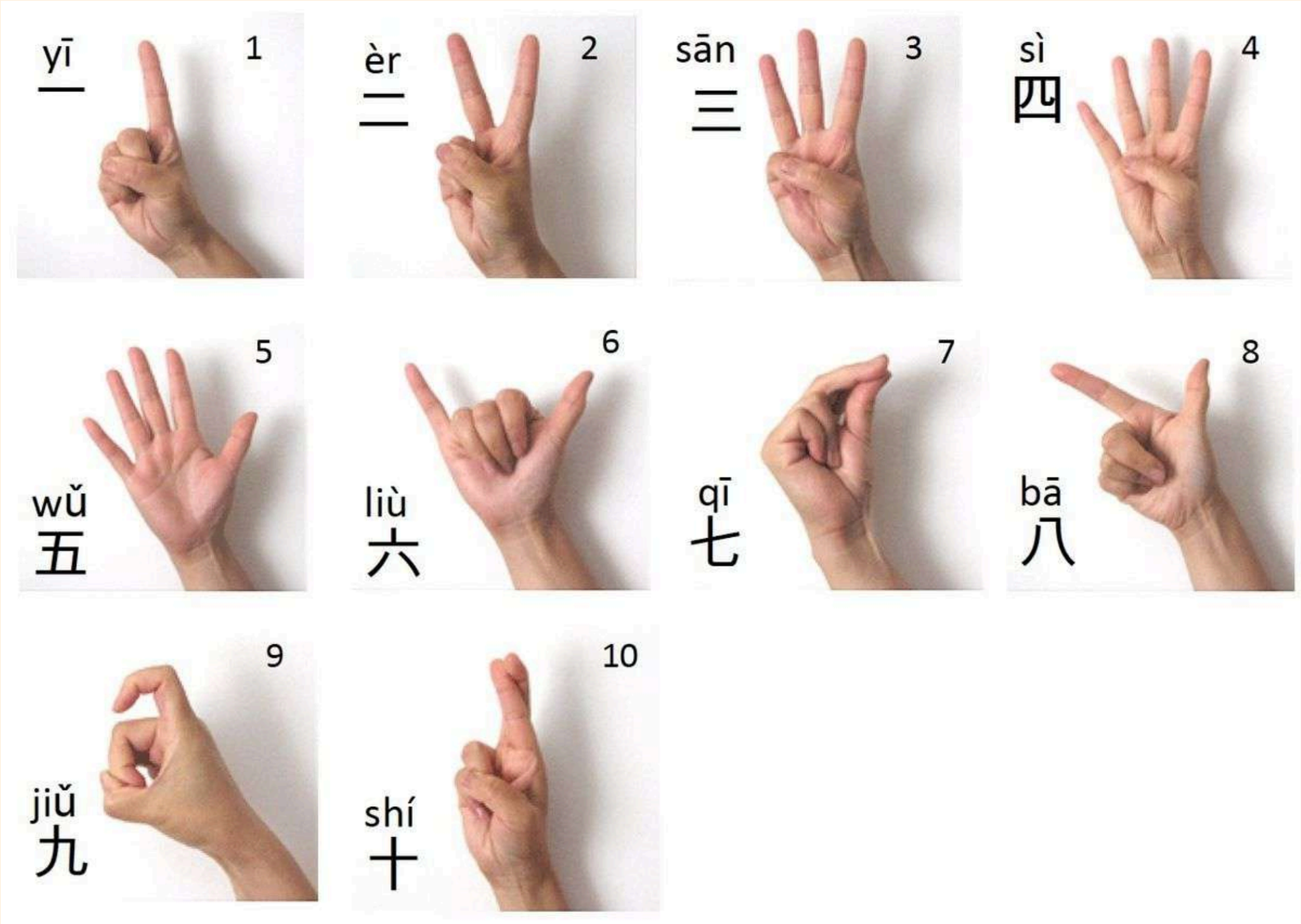


(d)

Join at

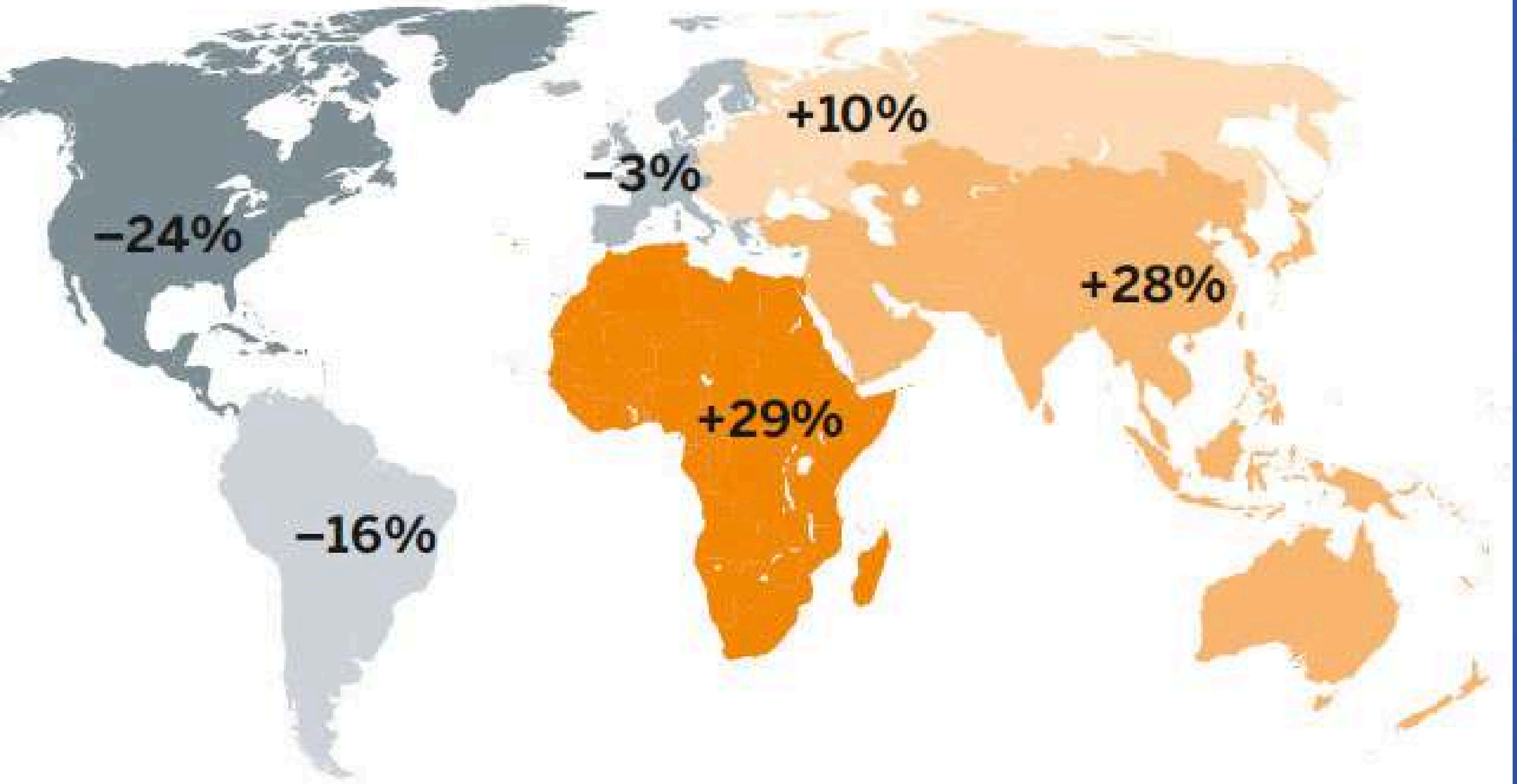
slido.com

#405 003


yī
一
1
èr
二
2
sān
三
3
sì
四
4
wǔ
五
5
liù
六
6
qī
七
7
bā
八
8
jiǔ
九
9
shí
十
10

REGIONAL REVENUE TRENDS, Q1 '10–Q4 '13

PERCENTAGE CHANGE



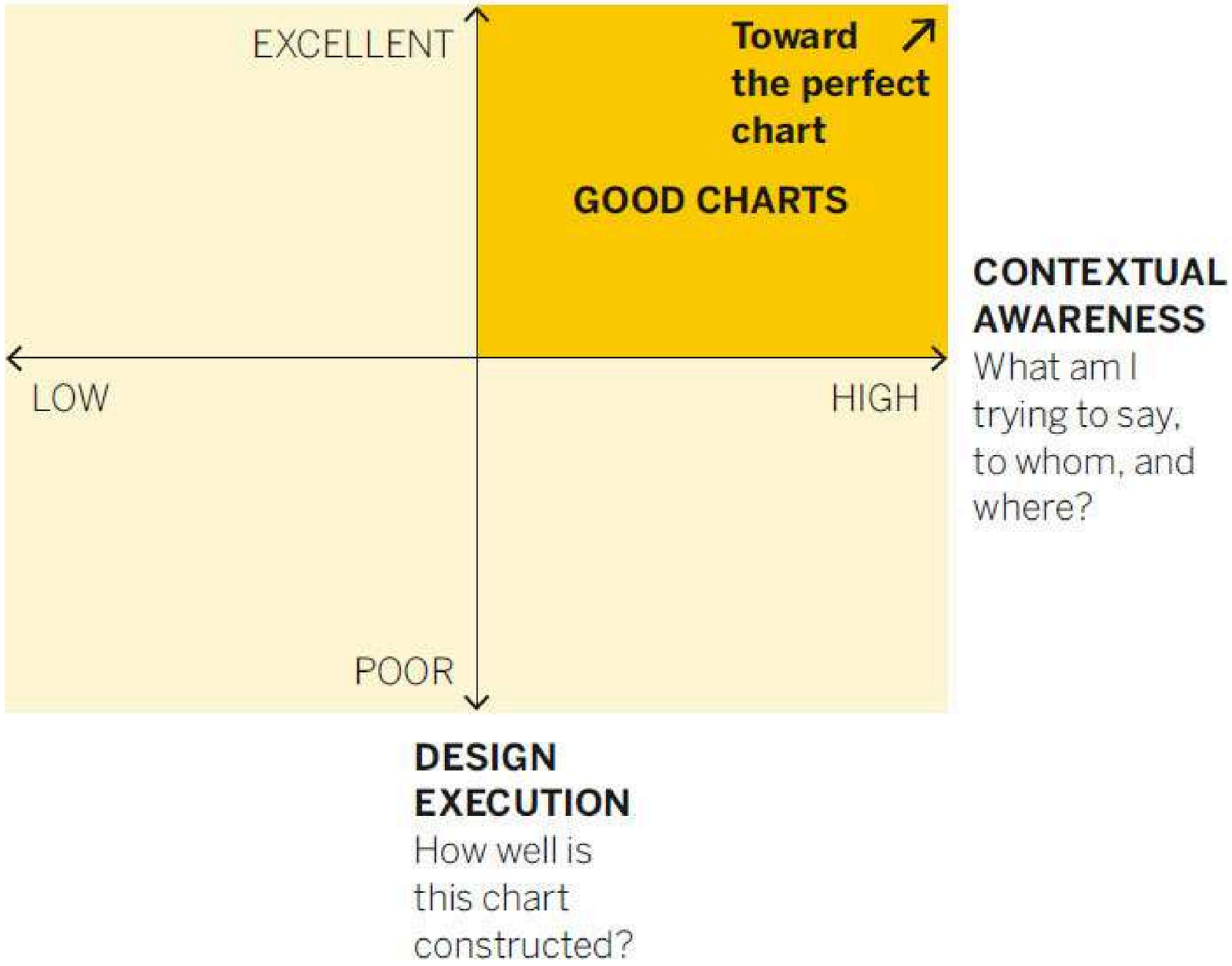
SOURCE: COMPANY RESEARCH

Join at

slido.com

#405 004

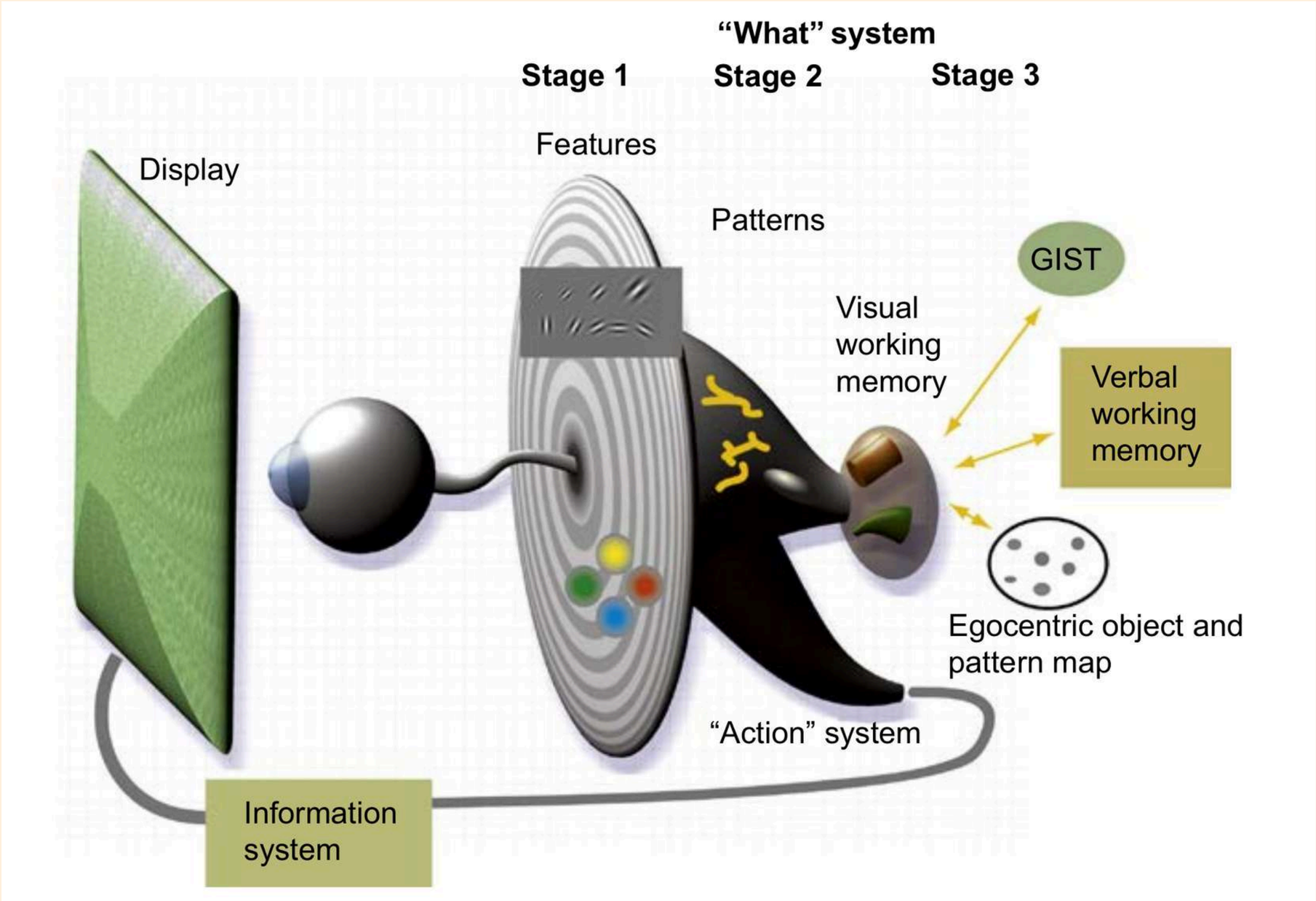
THE GOOD CHARTS MATRIX



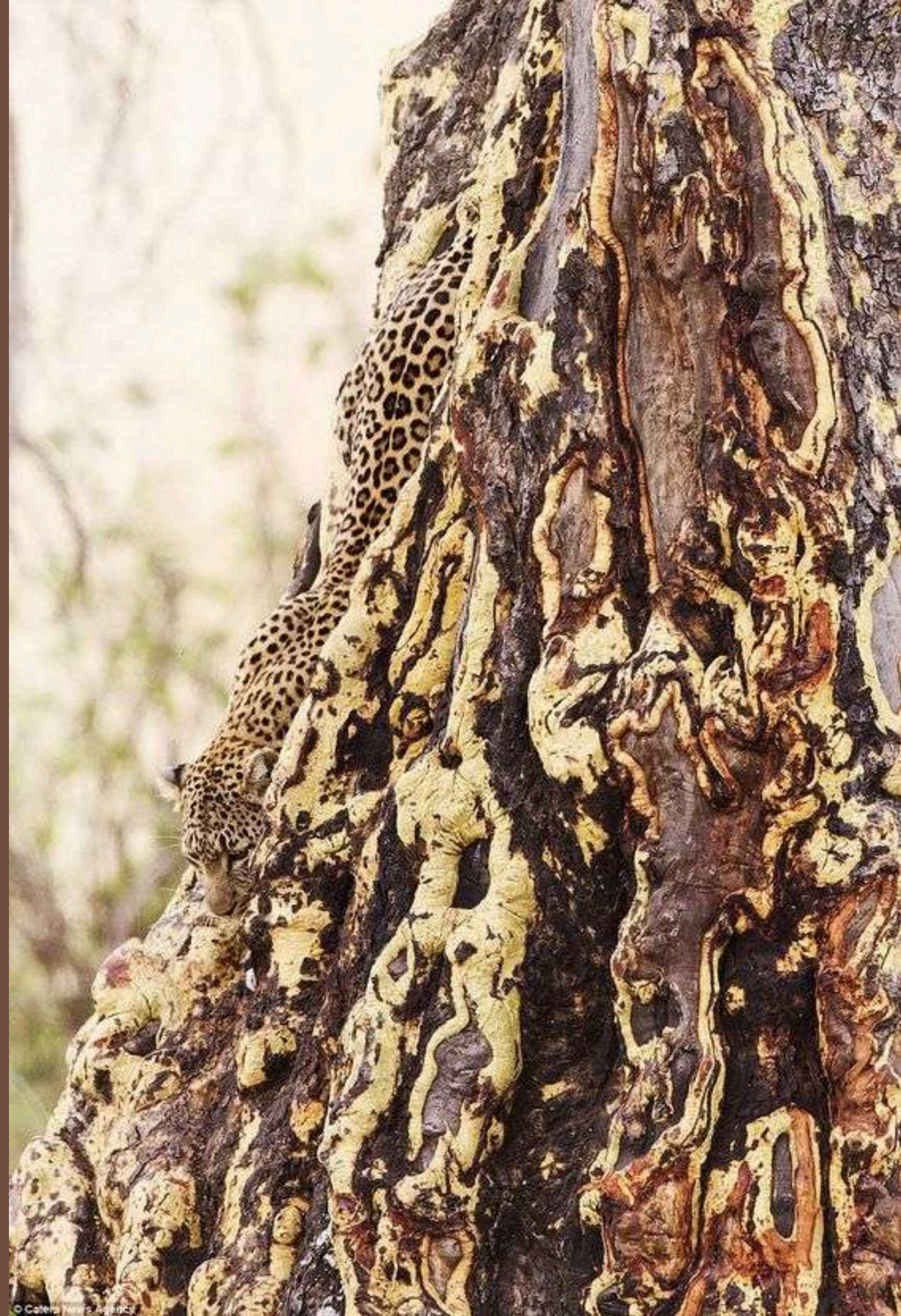
Join at

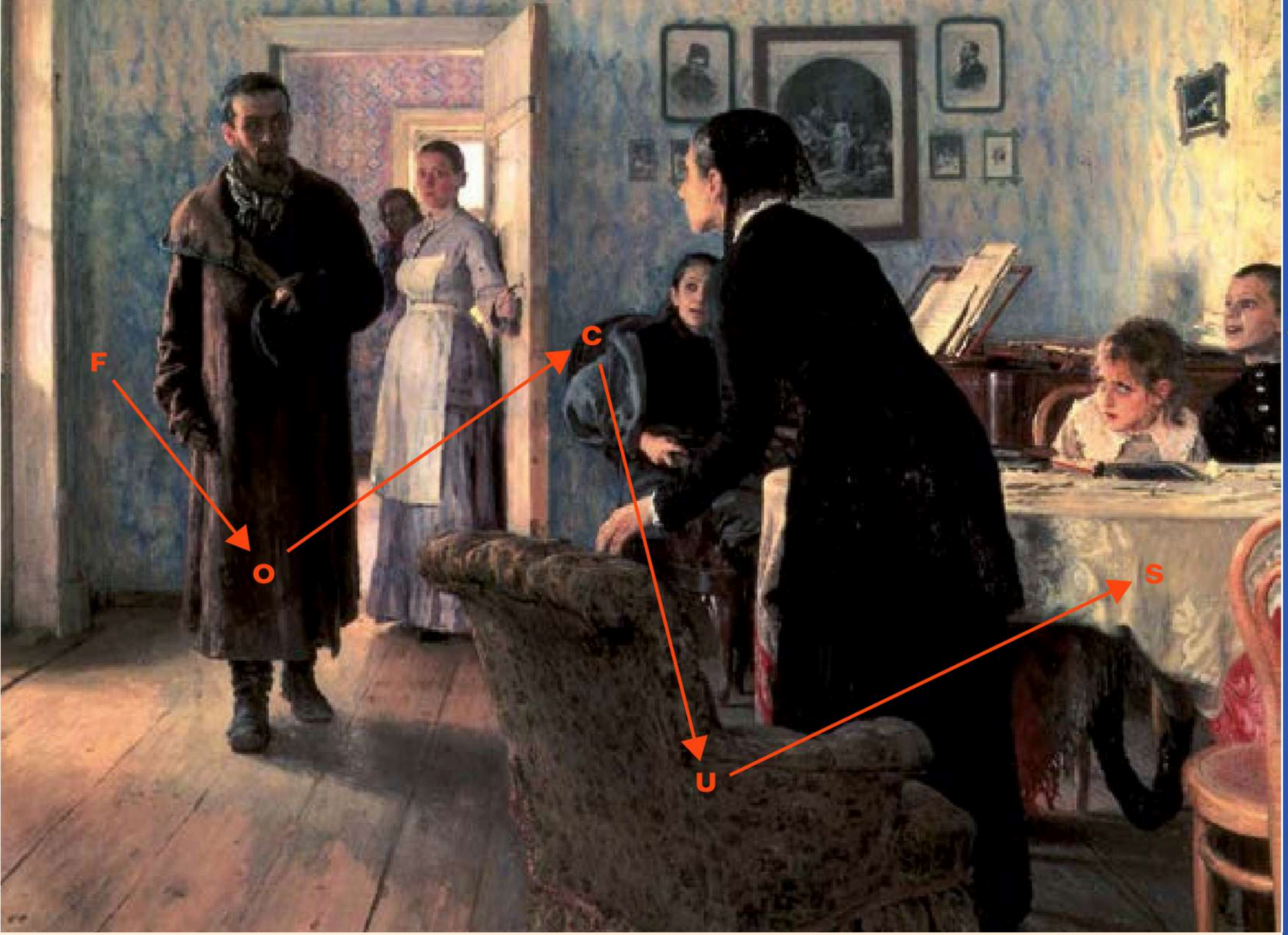
slido.com

#405 004


"What" system
Stage 1
Stage 2
Stage 3
Features
Display
Patterns
GIST
Visual
working
memory
Verbal
working
memory
Egocentric object and
pattern map
"Action" system
Information
system

What do you see here?





Join at

slido.com

#405 004

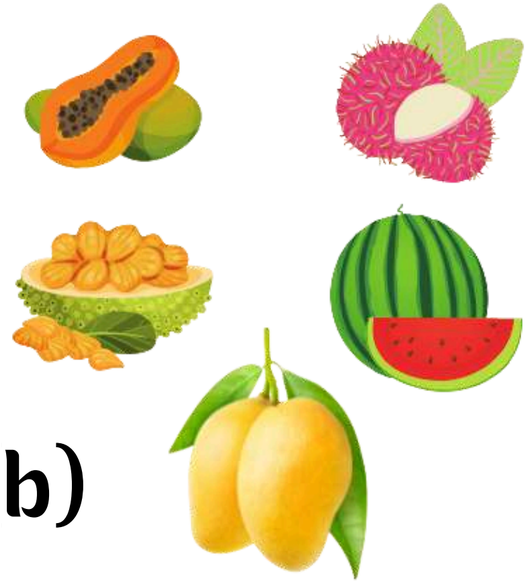
S. Kundu Ph.D.

Join at

slido.com

#40505

(a)



(b)

List of Indian Railway Trains and their Time Table

TRAIN NUMBER	TRAIN NAME	TRAIN TYPE	RUN DAYS	DEPARTURE TIME(HRS)	ARRIVAL TIME(HRS)
12267	MUMBAI CENTRAL - AHMEDABAD AC DURONTO EXP	DURONTO	M,T,W,T,F,S,S	23:25	05:55
12268	AHMEDABAD - MUMBAI CENT AC DURONTO EXP	DURONTO	M,T,W,T,F,S,S	23:40	06:00
22201	KOLKATA SEALDAH - PURI DURONTO EXPRESS	DURONTO	M,W,F	20:00	04:00
22204	SECUNDERABAD - VISAKHAPATNAM AC DURONTO EXPRESS	DURONTO	M,W,S	20:15	06:35
22206	MADURAI - CHENNAI CENTRAL AC DURONTO EXPRESS	DURONTO	TUE,THU	22:40	07:20
12426	JAMMU TAWI - NEW DELHI RAJDHANI EXPRESS	RAJDHANI	M,T,W,T,F,S,S	19:40	05:05
12430	NEW DELHI - LUCKNOW AC SF EXPRESS	RAJDHANI	M,T,F,S	20:50	06:40
12437	SECUNDERABAD - HAZRAT NIZAMUDDIN RAJDHANI EXPRESS	RAJDHANI	WED	12:45	10:25
12951	MUMBAI CENTRAL - NEW DELHI RAJDHANI EXPRESS	RAJDHANI	M,T,W,T,F,S,S	16:35	08:35

(c)

S. Kundu Ph.D.

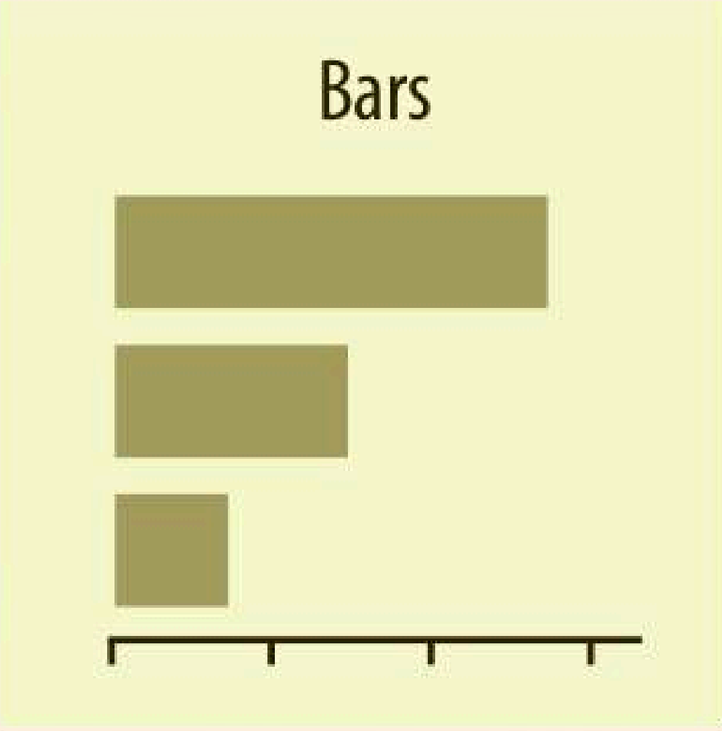
Join at

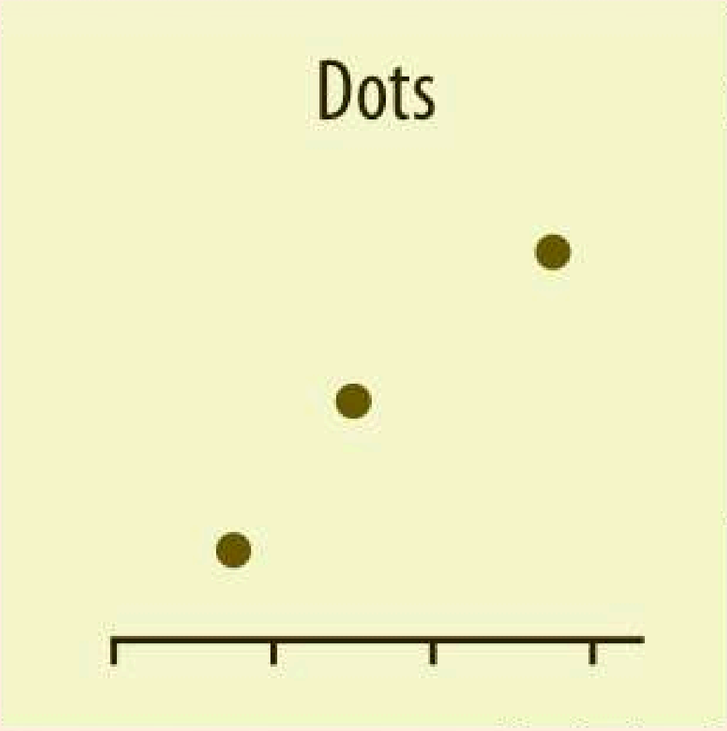
slido.com

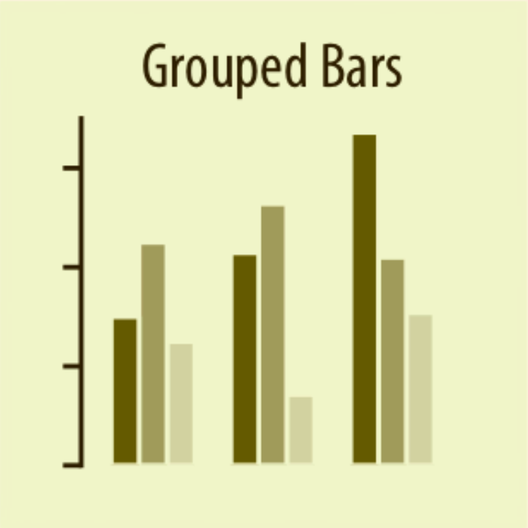
#40505

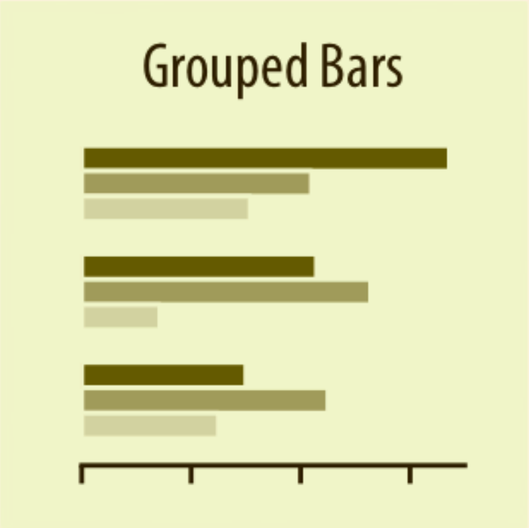
Catalog of Data Visualization

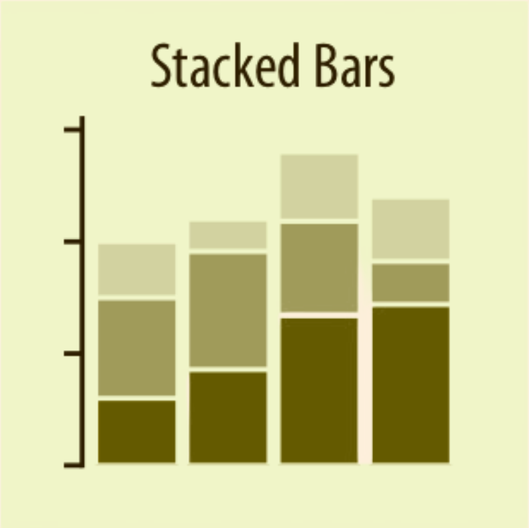
Amounts (Real values)







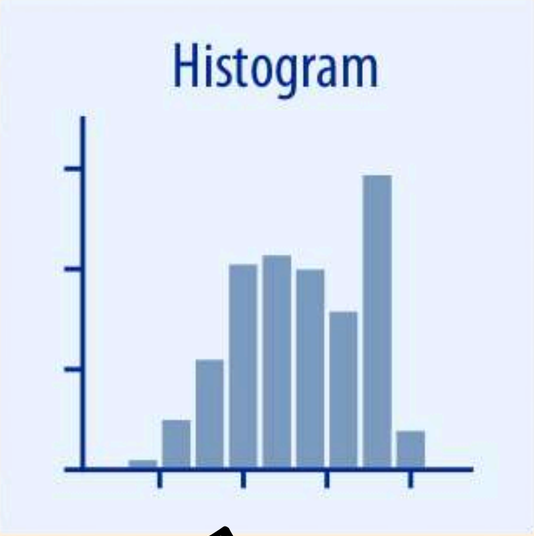


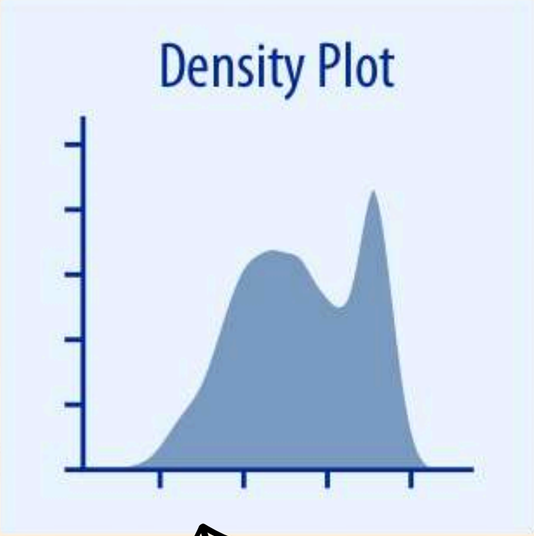


Amounts

Two or more sets of categories

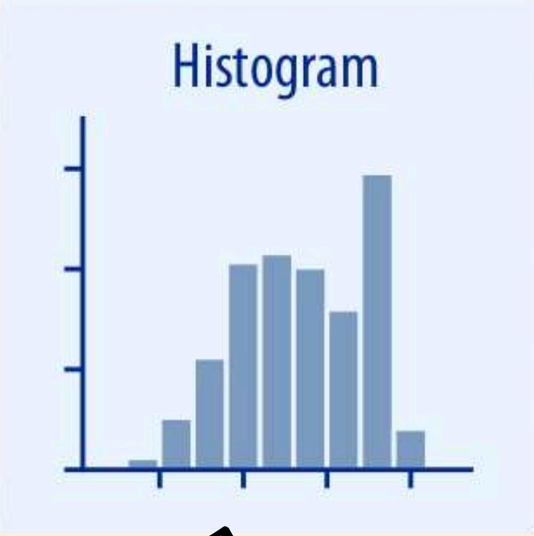
Frequency or Distribution

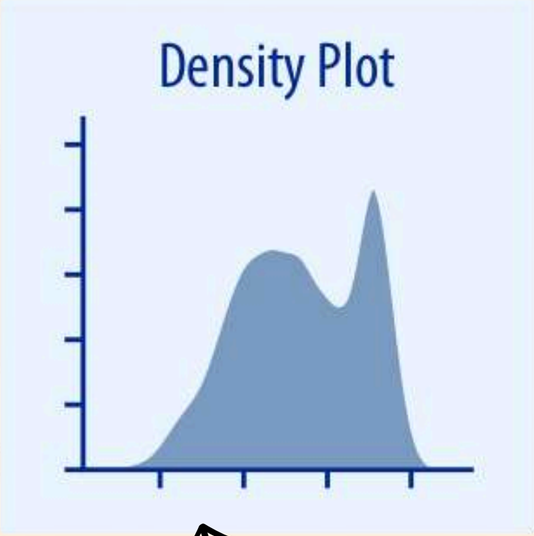


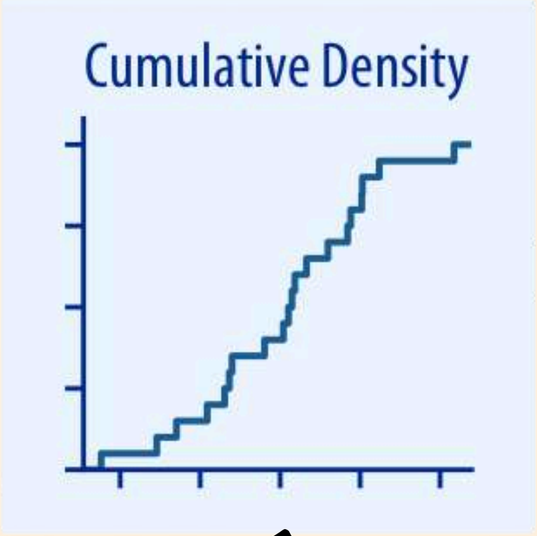


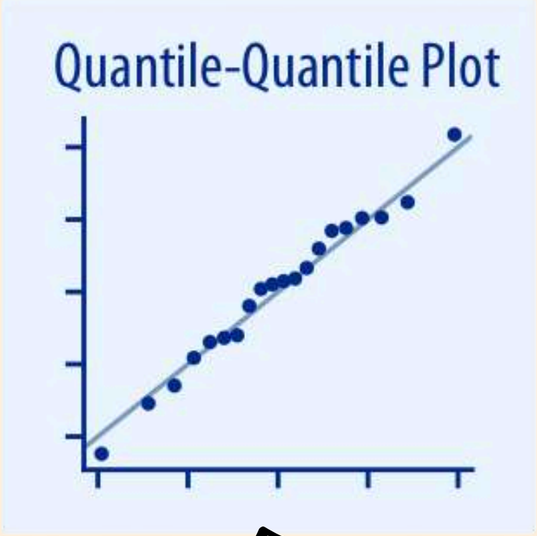
most intuitive visualizations
but both require arbitrary parameter choices
can be misleading

Frequency or Distribution



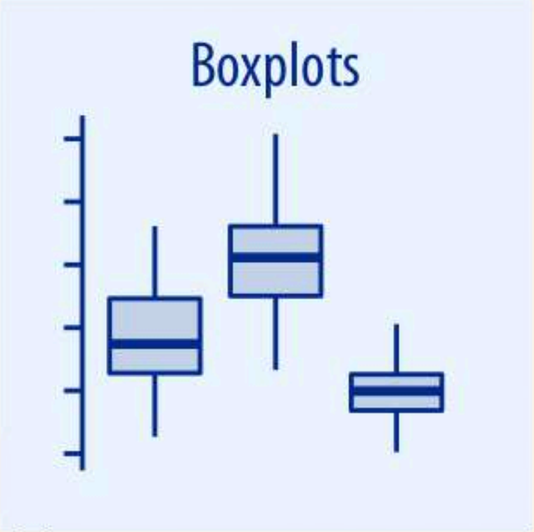


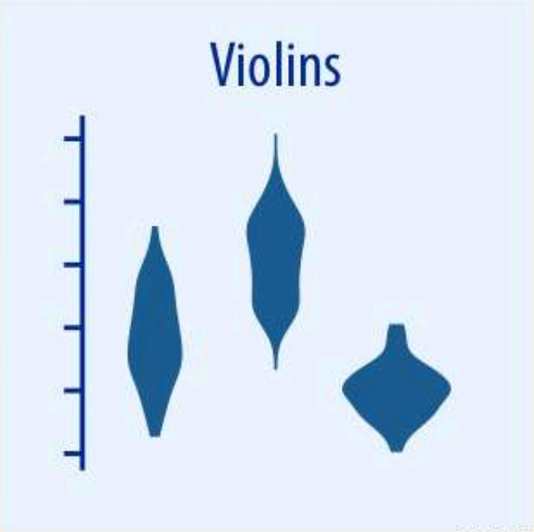


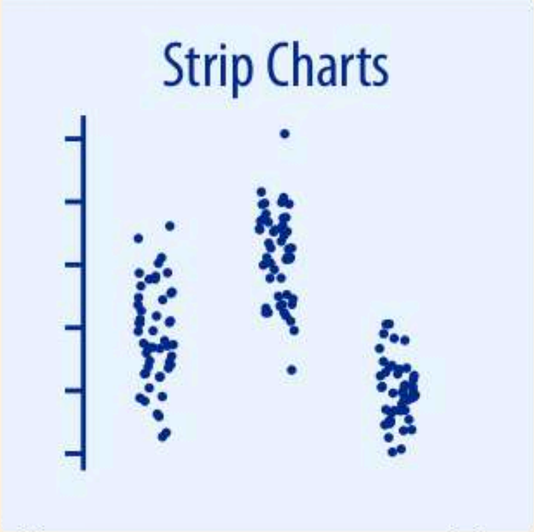


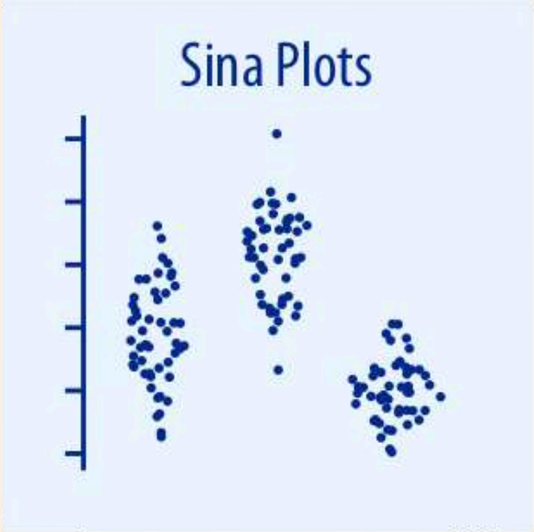
most intuitive visualizations
but require arbitrary parameter choices
can be misleading

represent the data faithfully
but can be more difficult to interpret

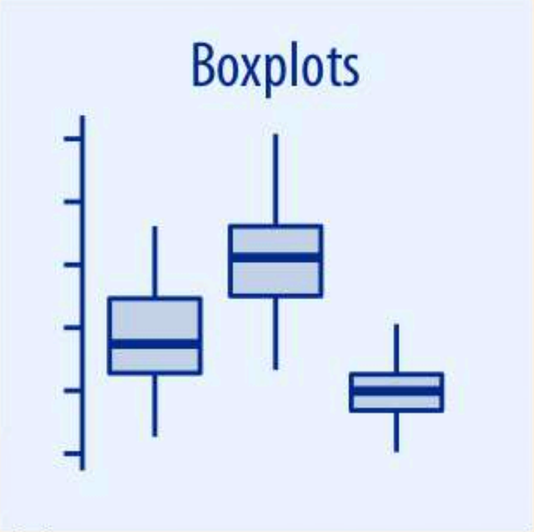


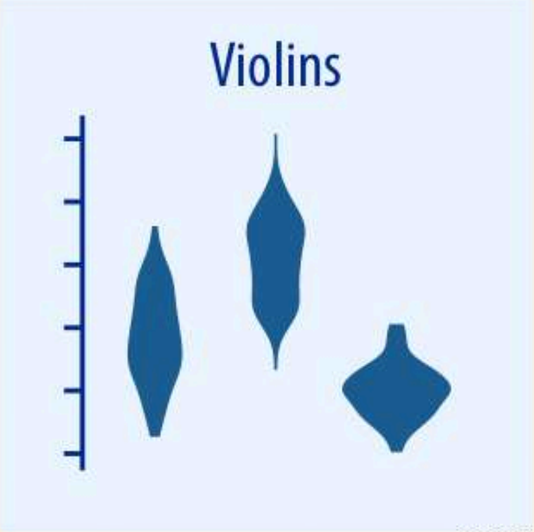


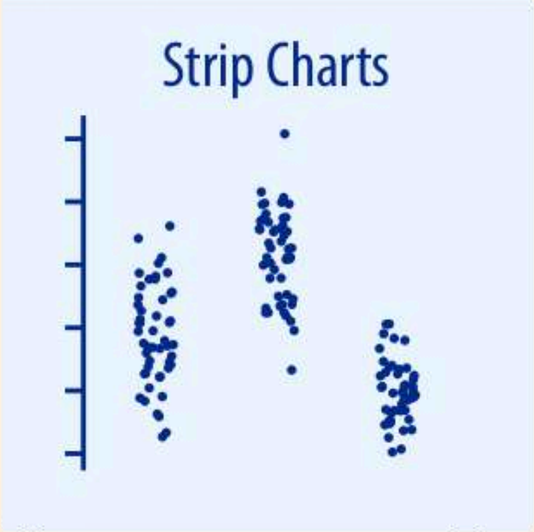


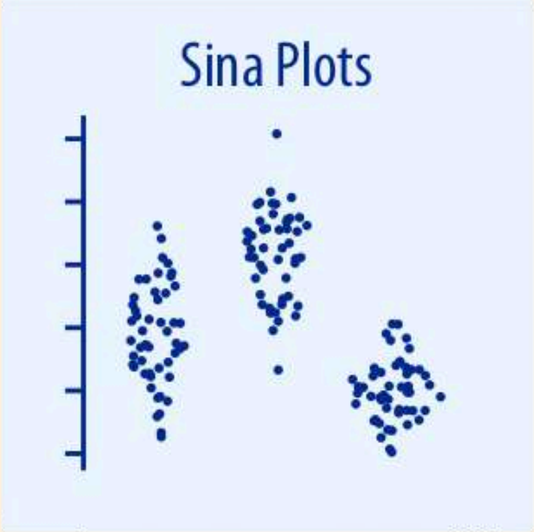


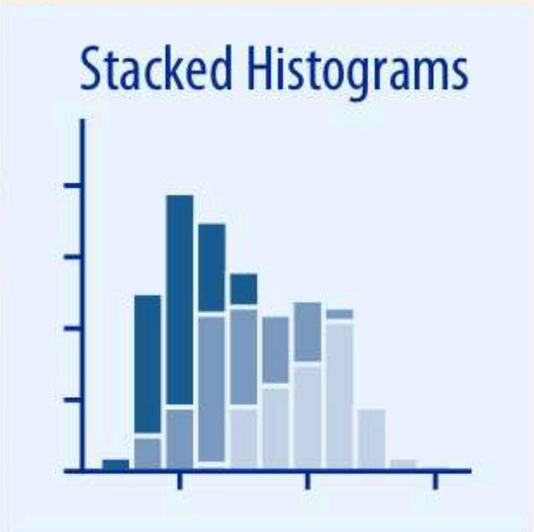
Multiple Distributions

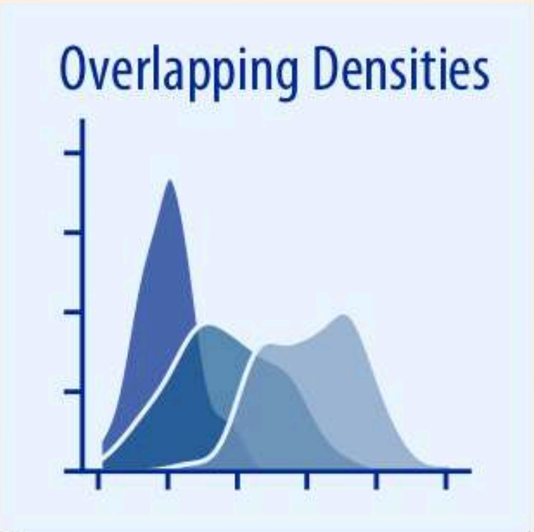


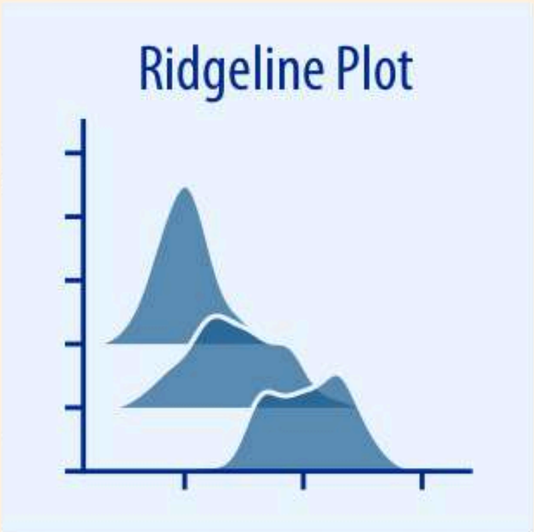








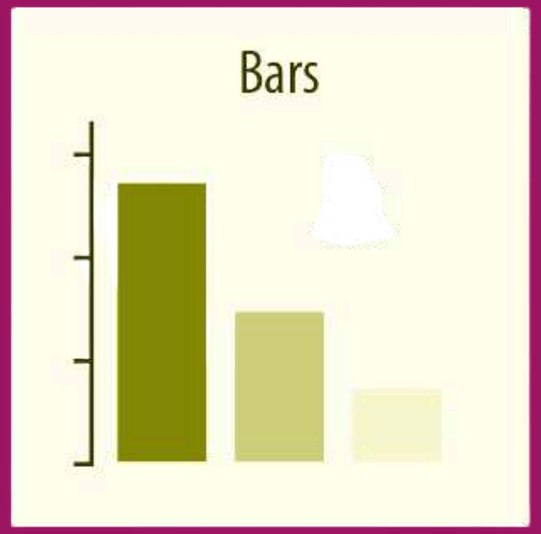


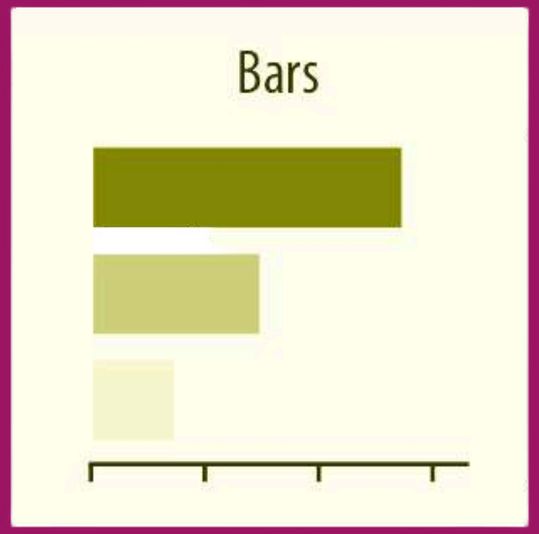


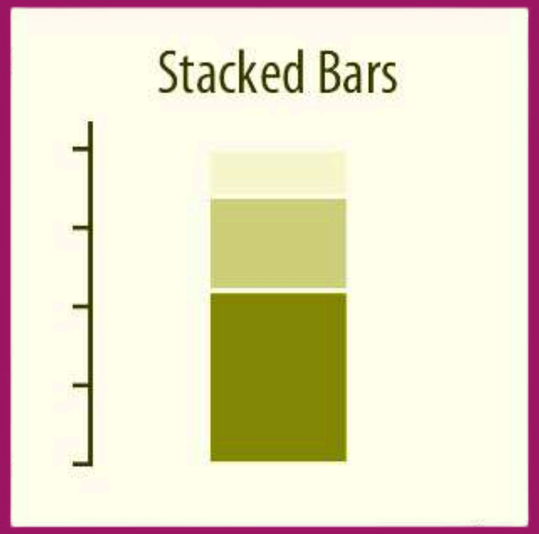
Multiple Distributions

more in-depth comparison of a smaller number of distributions

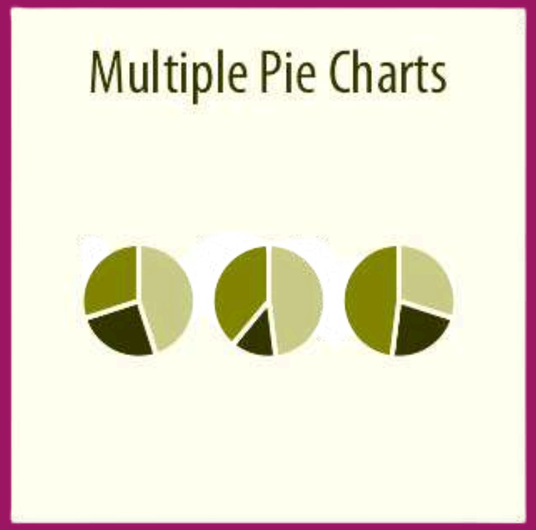
Proportions

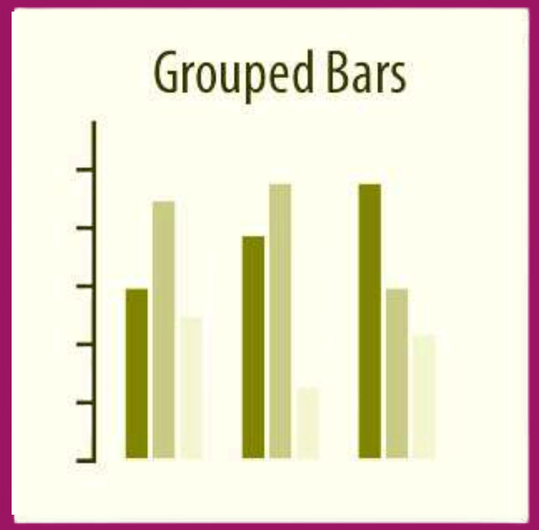


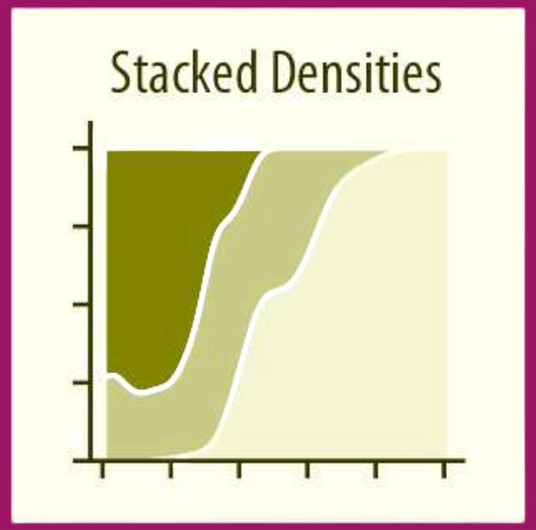




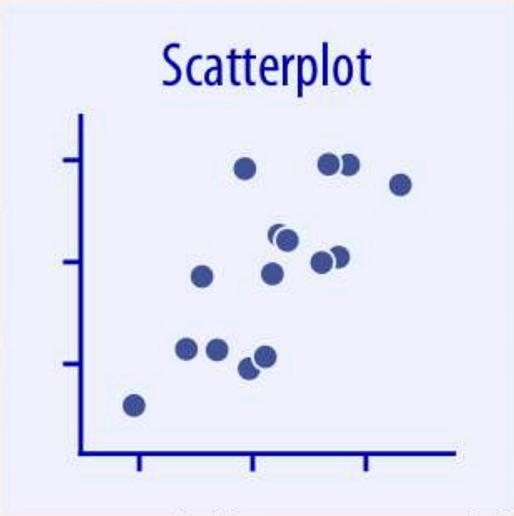
Multiple Proportions

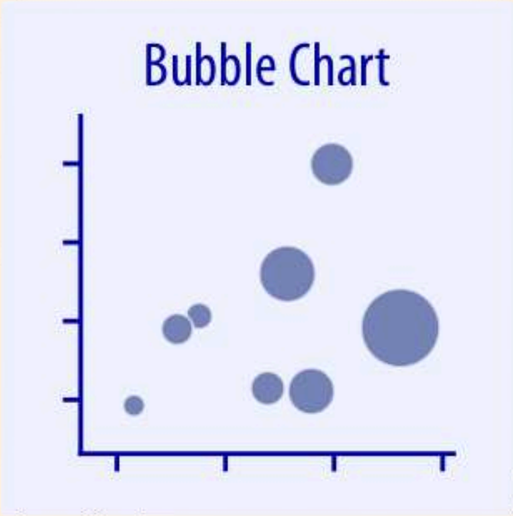


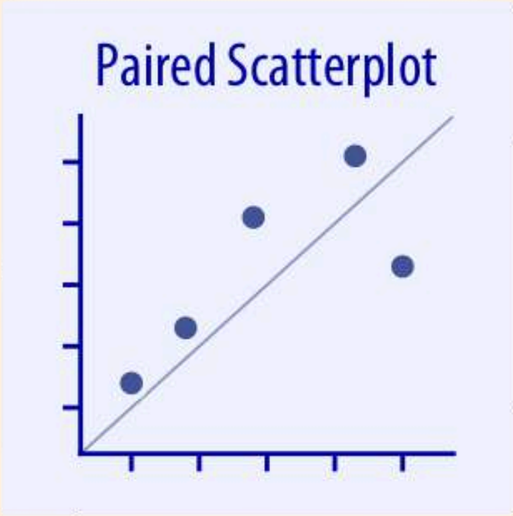


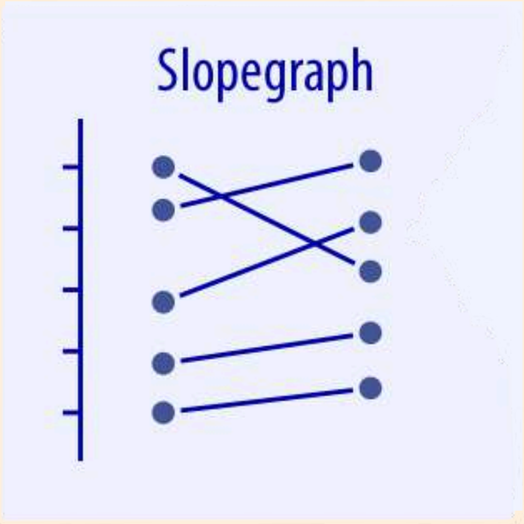


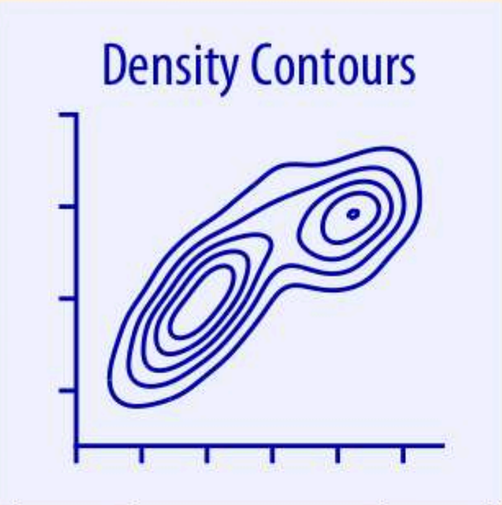
x-y relationship

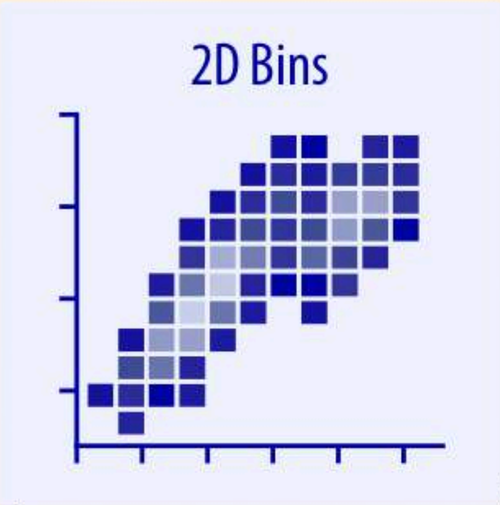


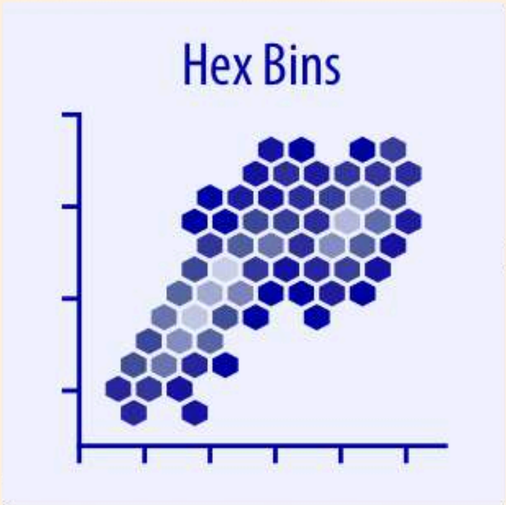


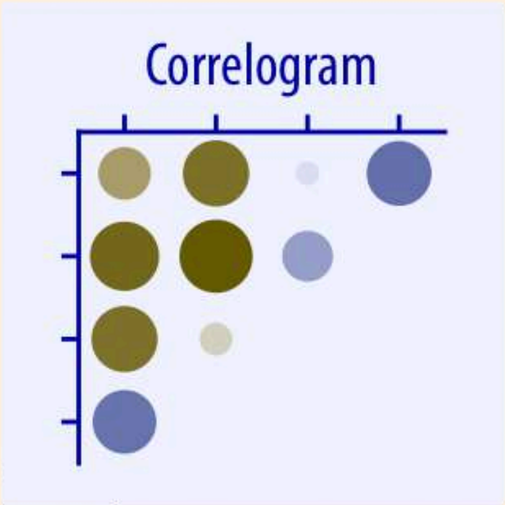




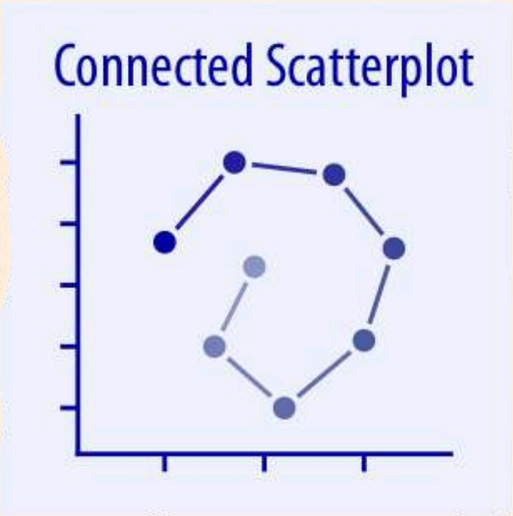


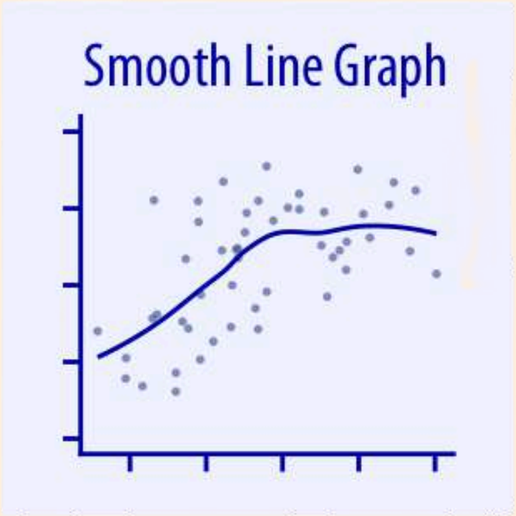






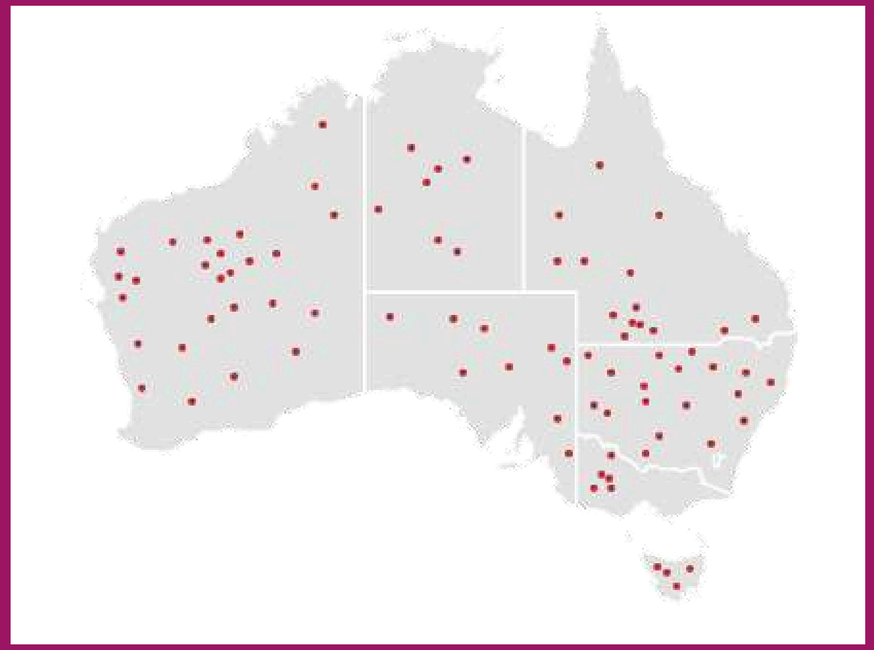
x-y relationship



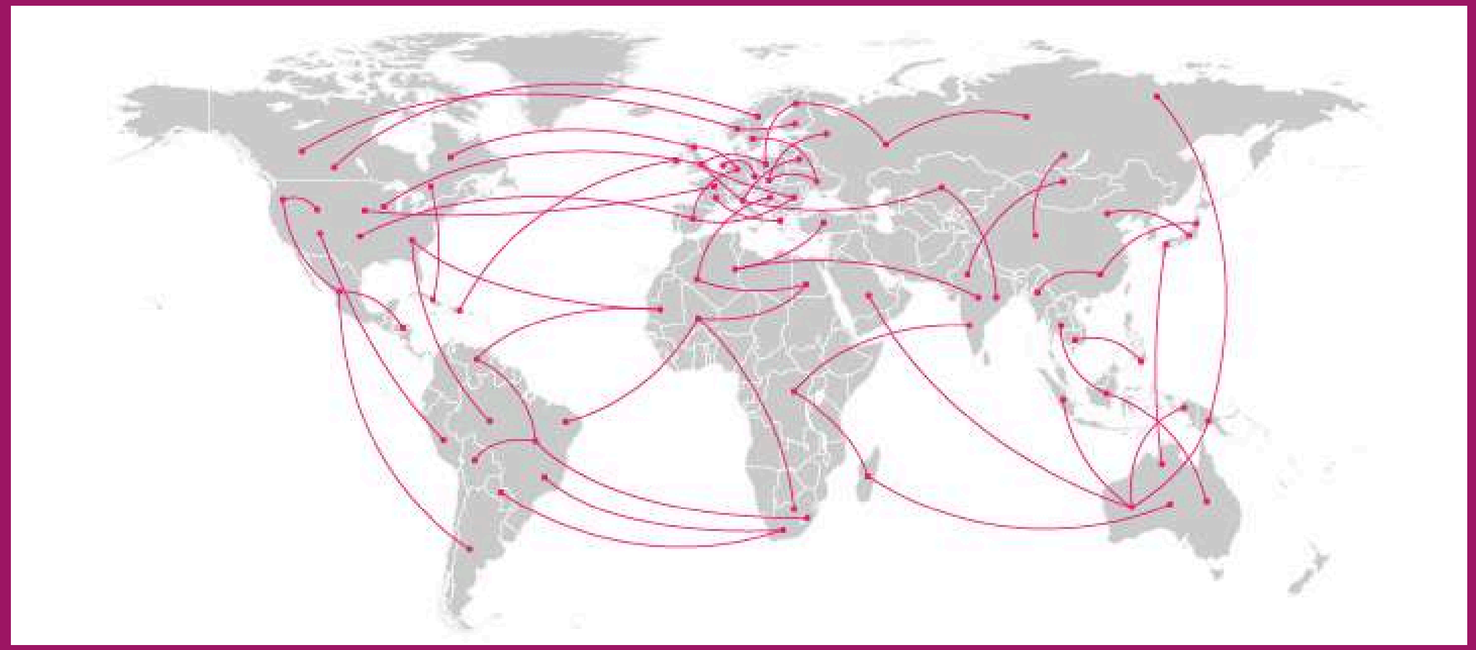


Map

Dot Map

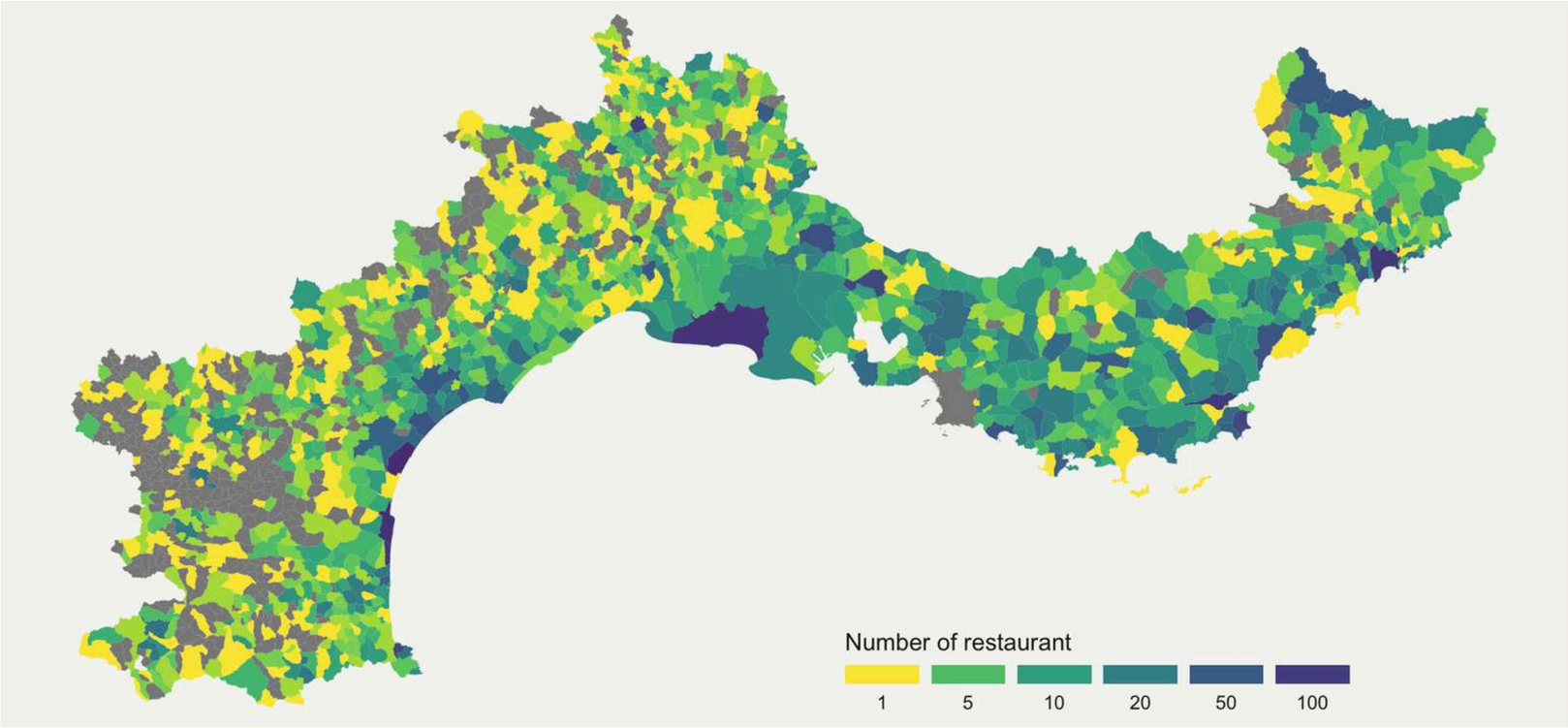


Connection Map



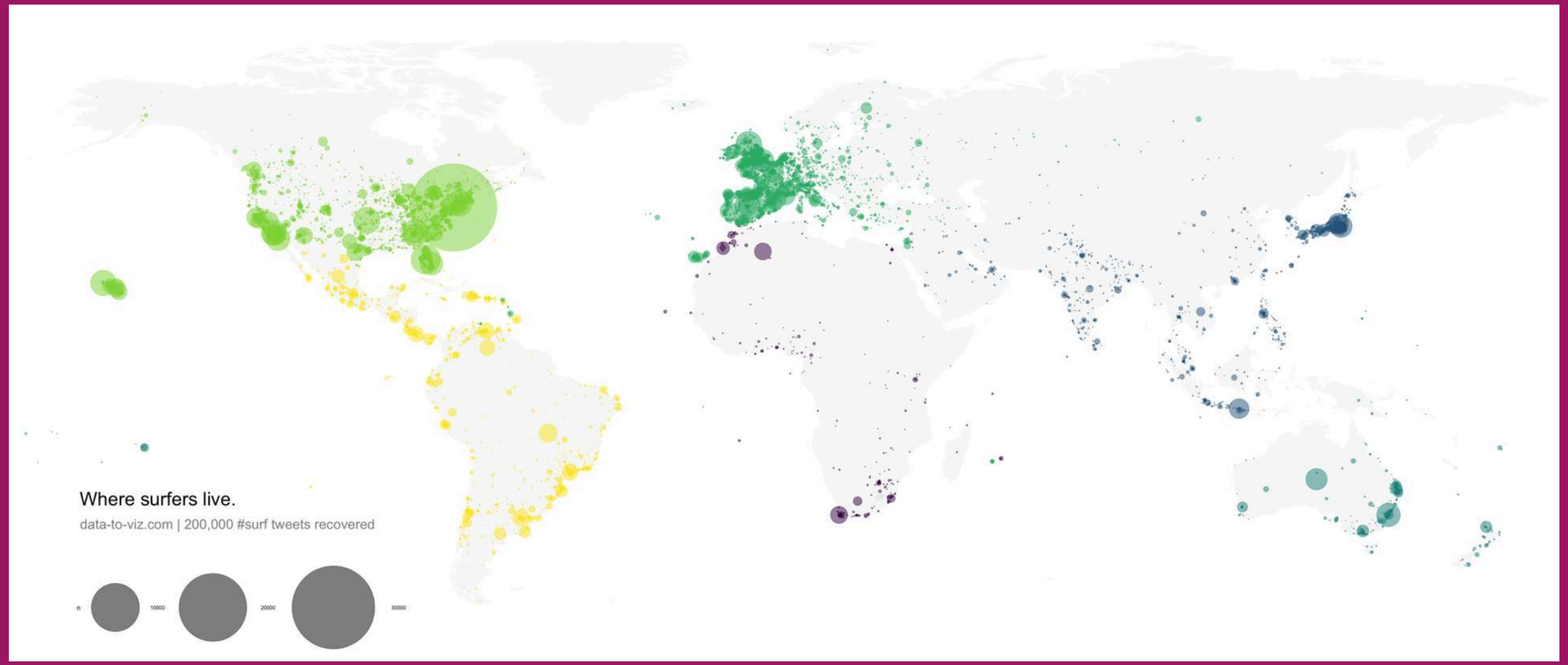
Map

Choropleth Map

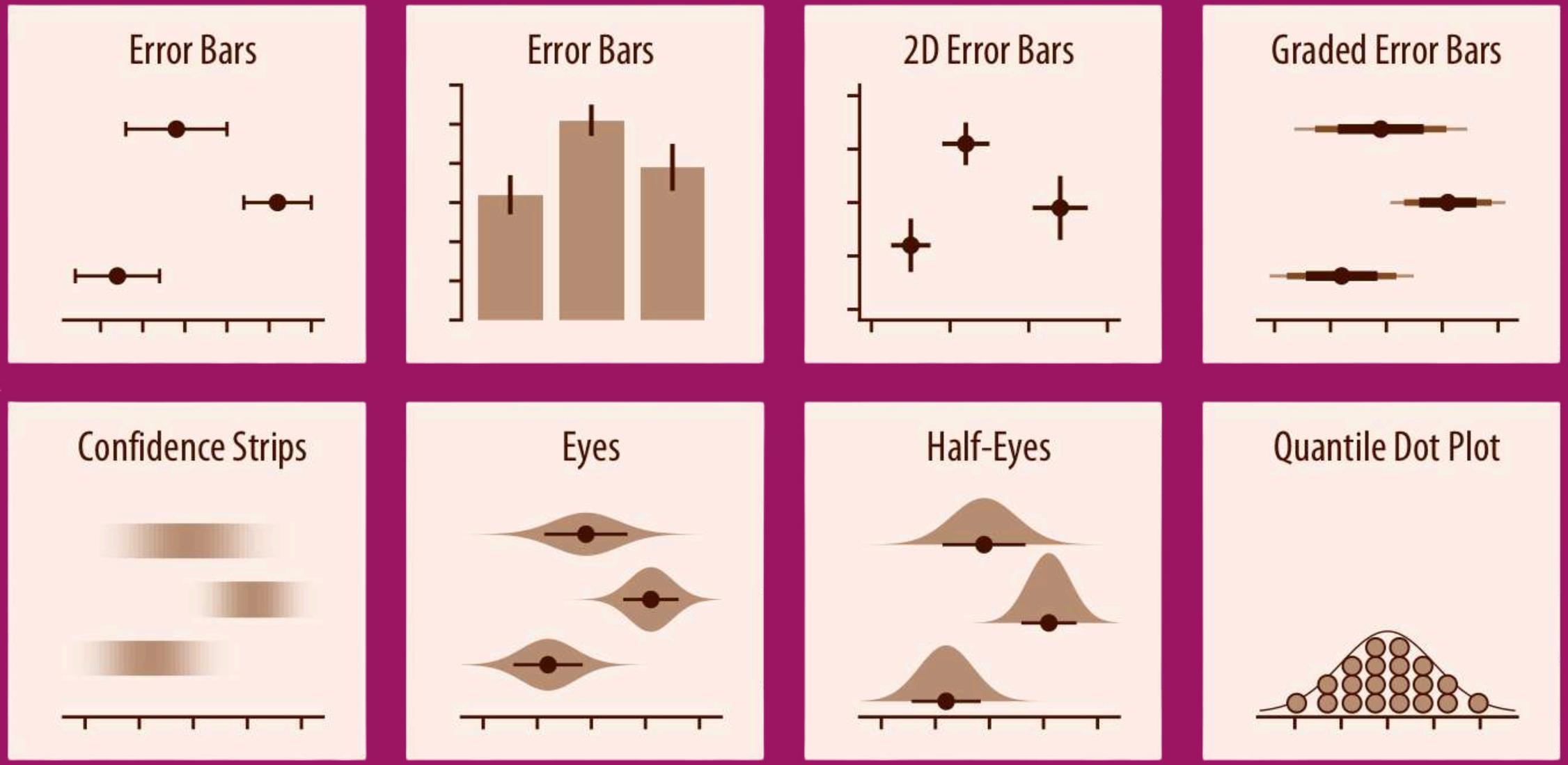


Map

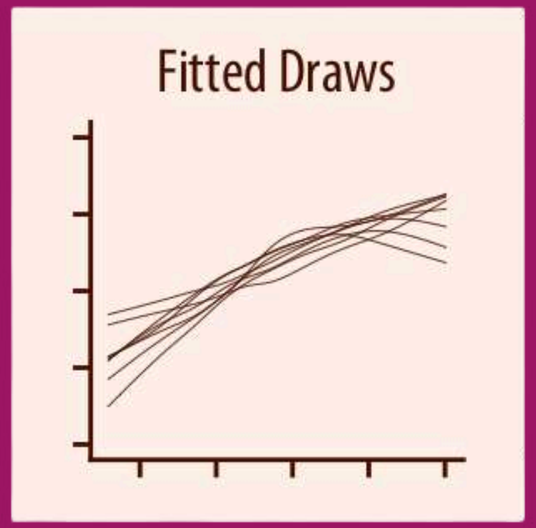
Bubble Map



Uncertainty

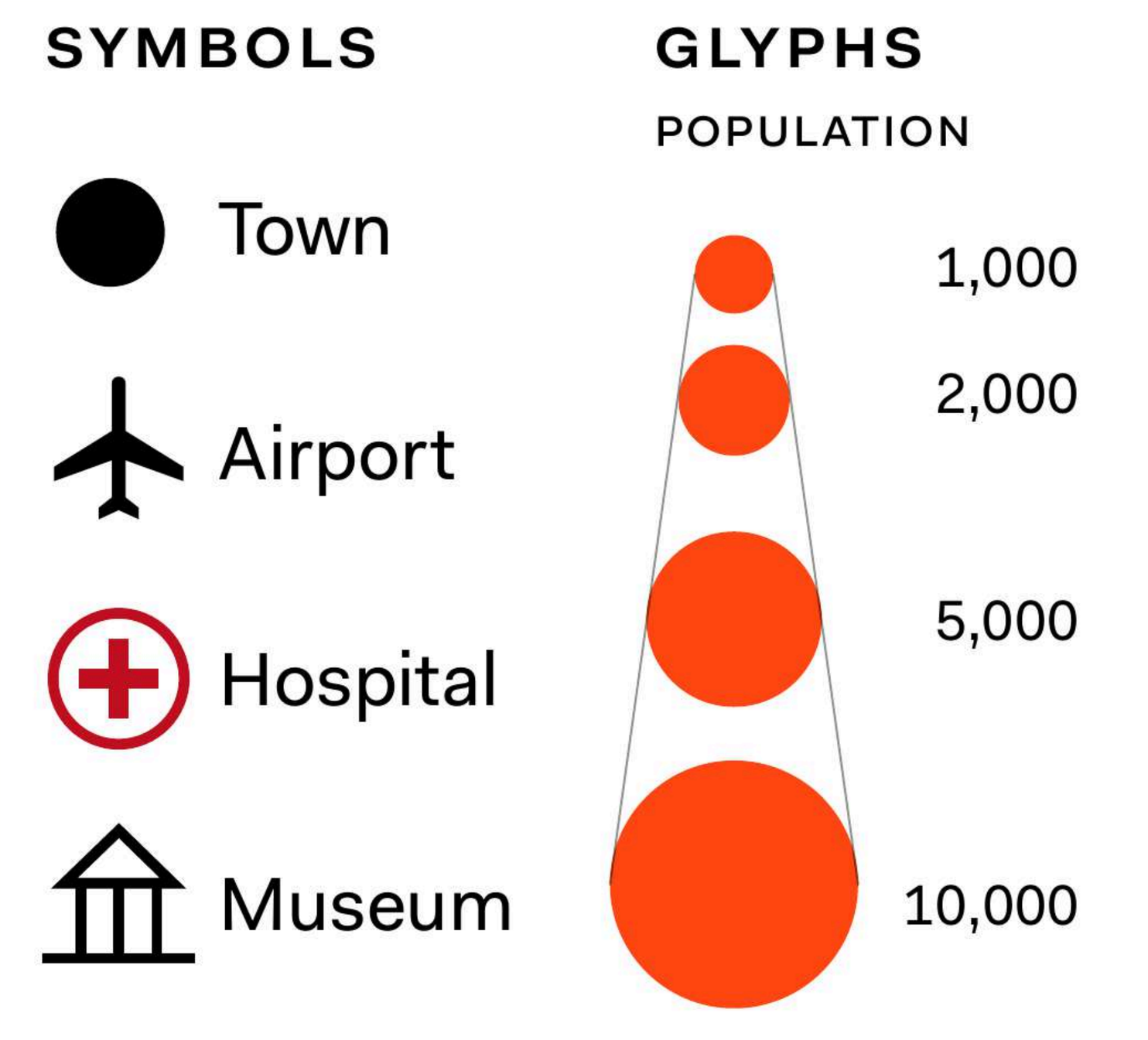


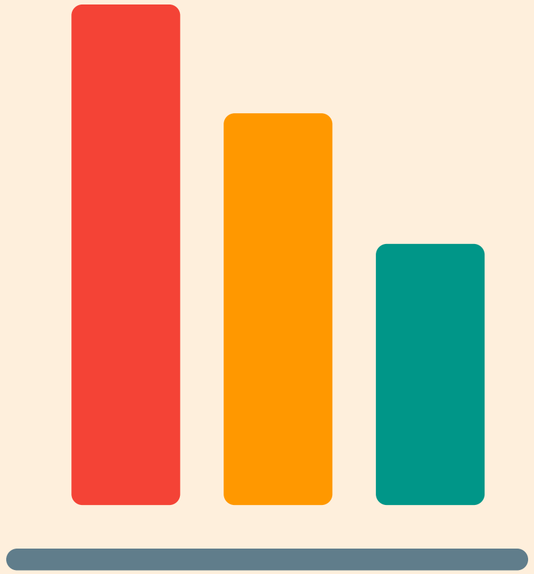
Uncertainty

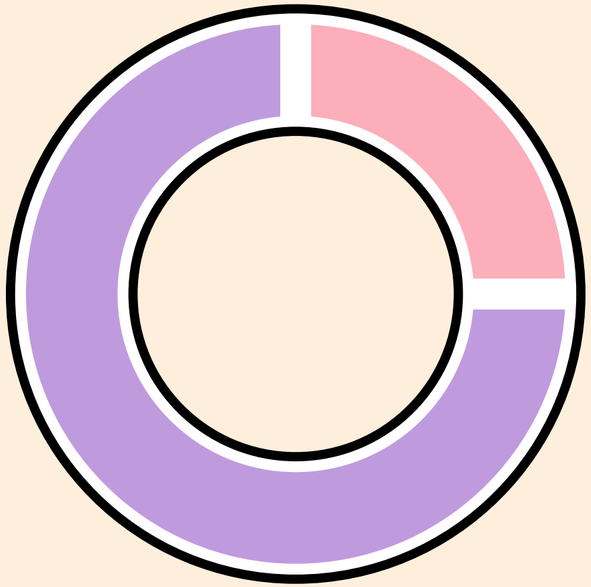


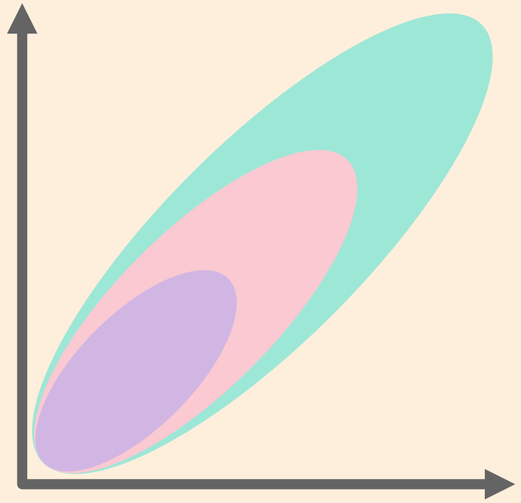
Aesthetics & Perception

Graphical perception Theory, Experimentation, and the Application, Graphical Integrity, Layering and Separation, Color and Information, Using Space Effectively


SYMBOLS
Town
Airport
Hospital
Museum
GLYPHS
POPULATION
1,000
2,000
5,000
10,000





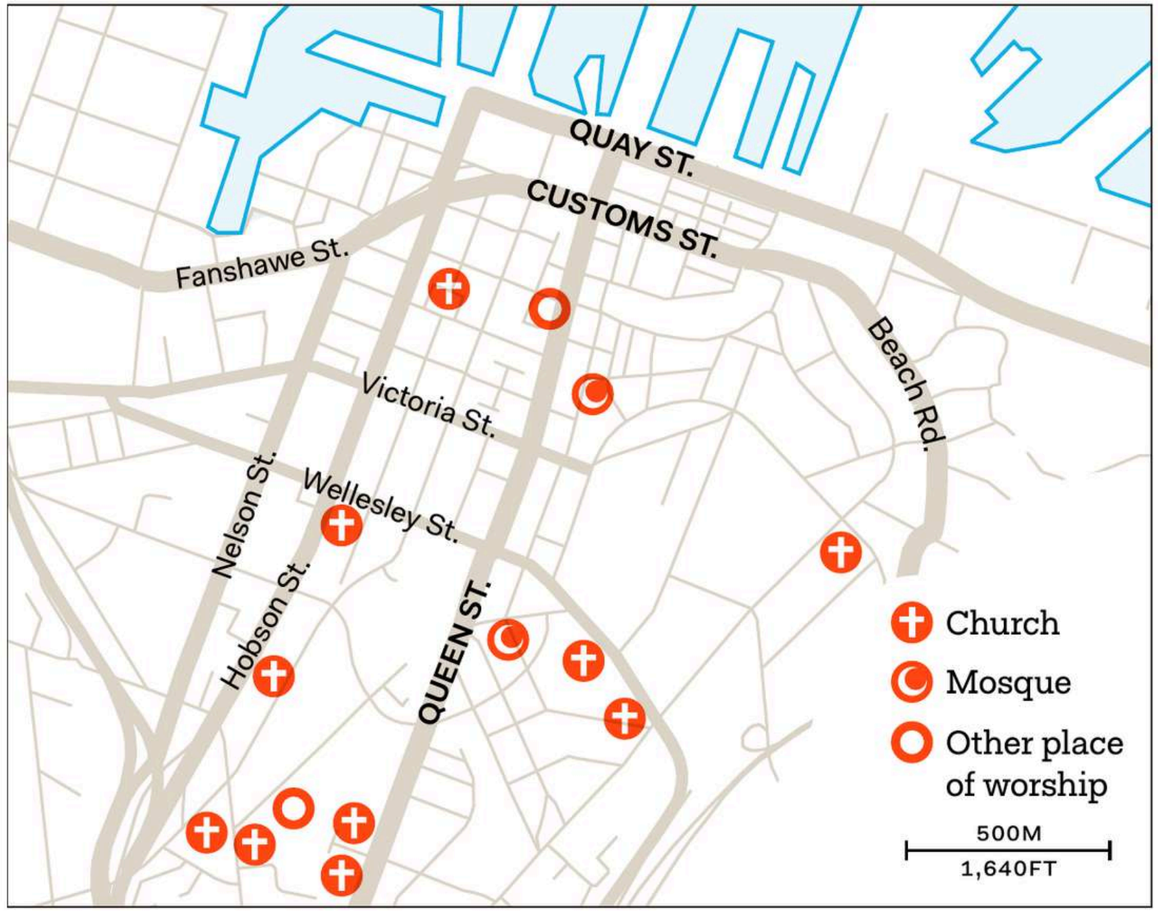


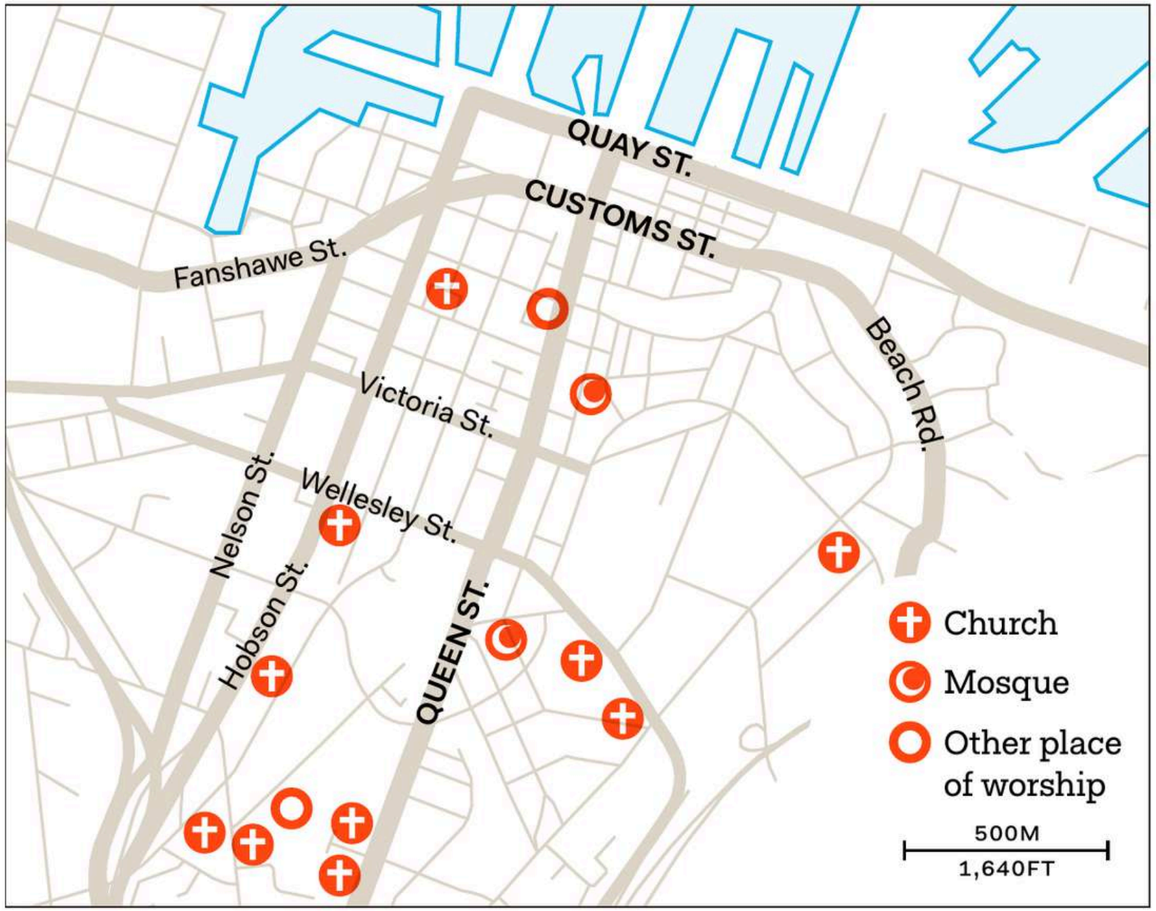
S. Kundu Ph.D.

Join at

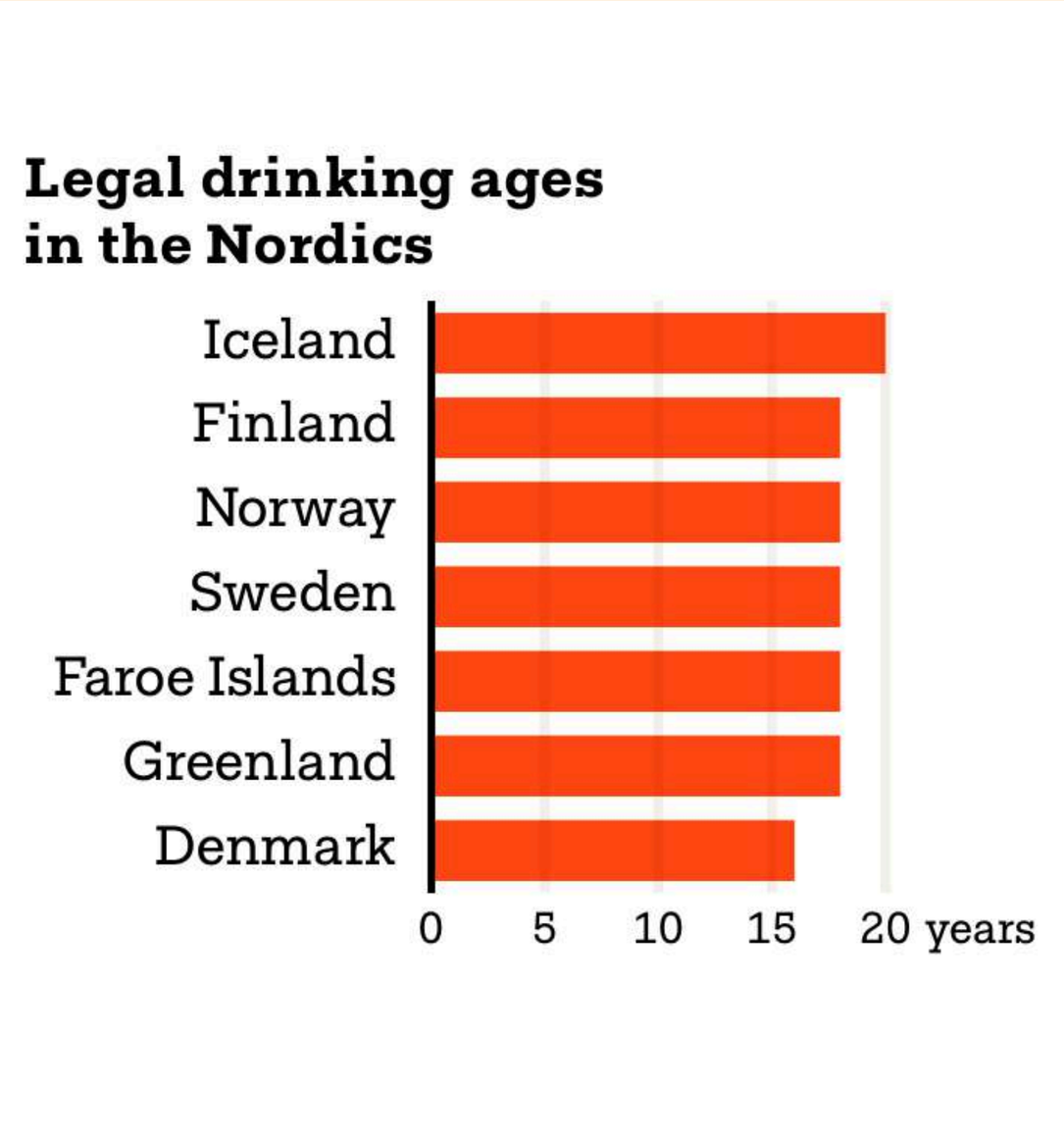
slido.com

#342


QUAY ST.
CUSTOMS ST.
Fanshawe St.
Victoria St.
Wellesley St.
Nelson St.
Hobson St.
QUEEN ST.
Beach Rd.
Church
Mosque
Other place of worship
500M
1,640FT



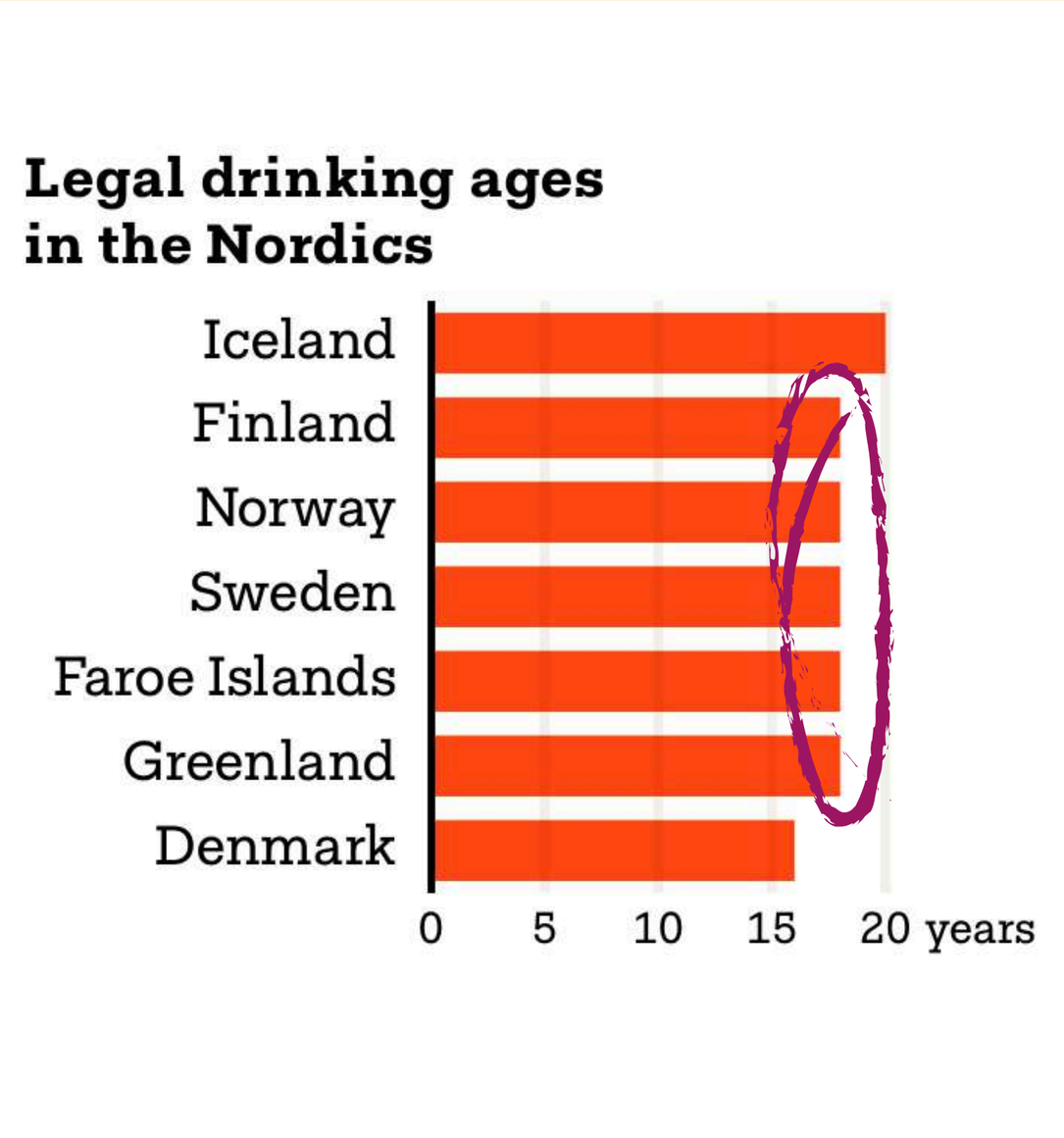
It is often said that the most important decision in making a map is not what to put in, but what to leave out



“

Tables usually outperform graphs in reporting small sets of 20 numbers or less

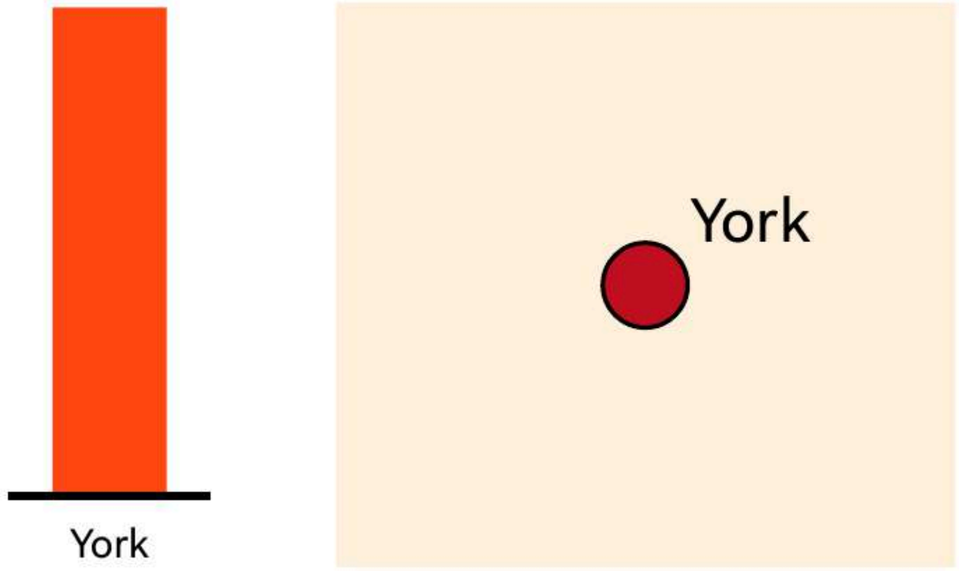
Edward Tufte, Professor Emeritus of statistics at Yale university



“

Tables usually outperform graphs in reporting small sets of 20 numbers or less

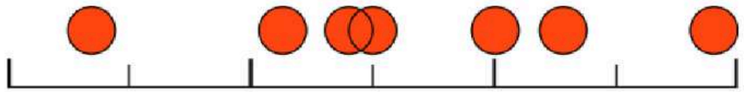
Edward Tufte, Professor Emeritus of statistics at Yale university

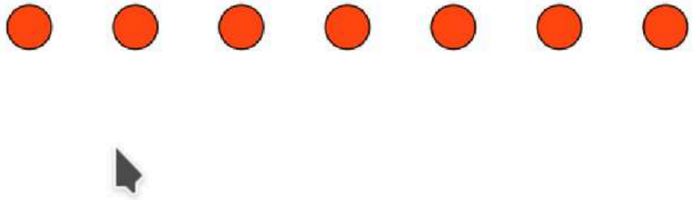


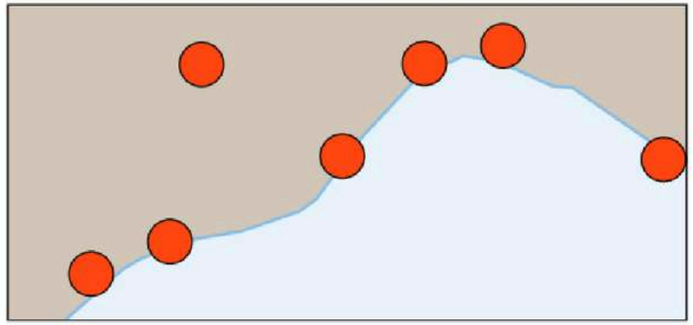
Meaningless bar chart and map.

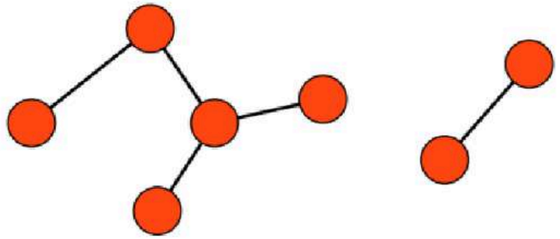
> At the heart of quantitative reasoning is a single question: Compared to what?

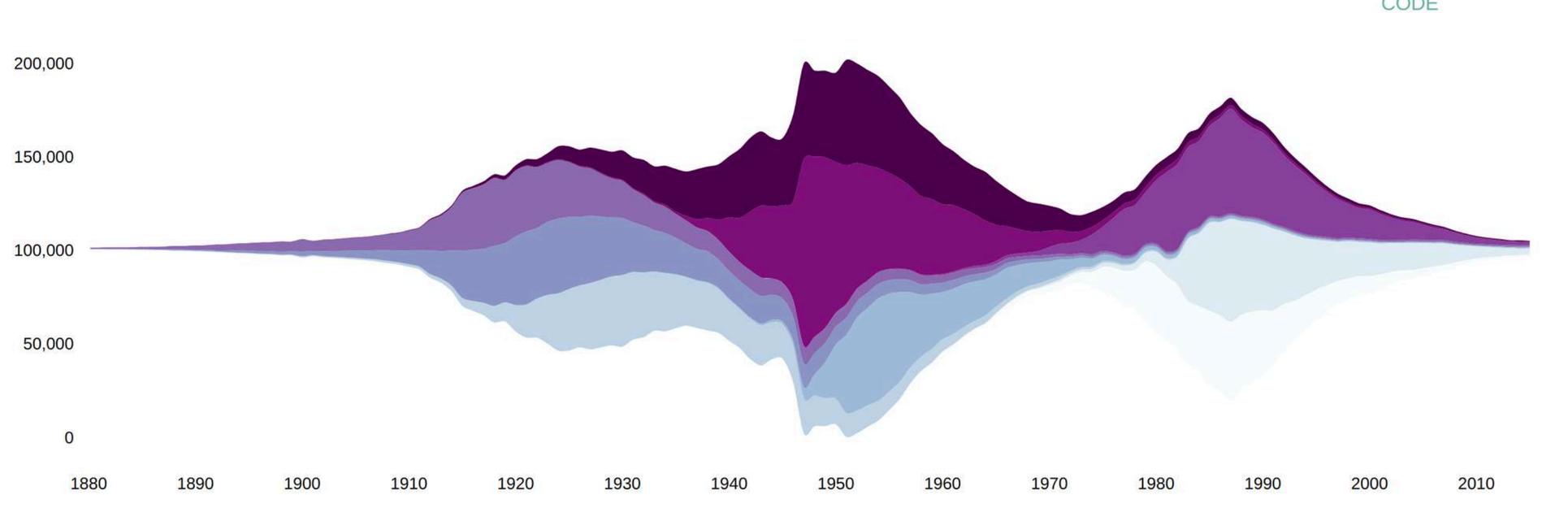
Edward Tufte, Professor Emeritus of statistics at Yale university







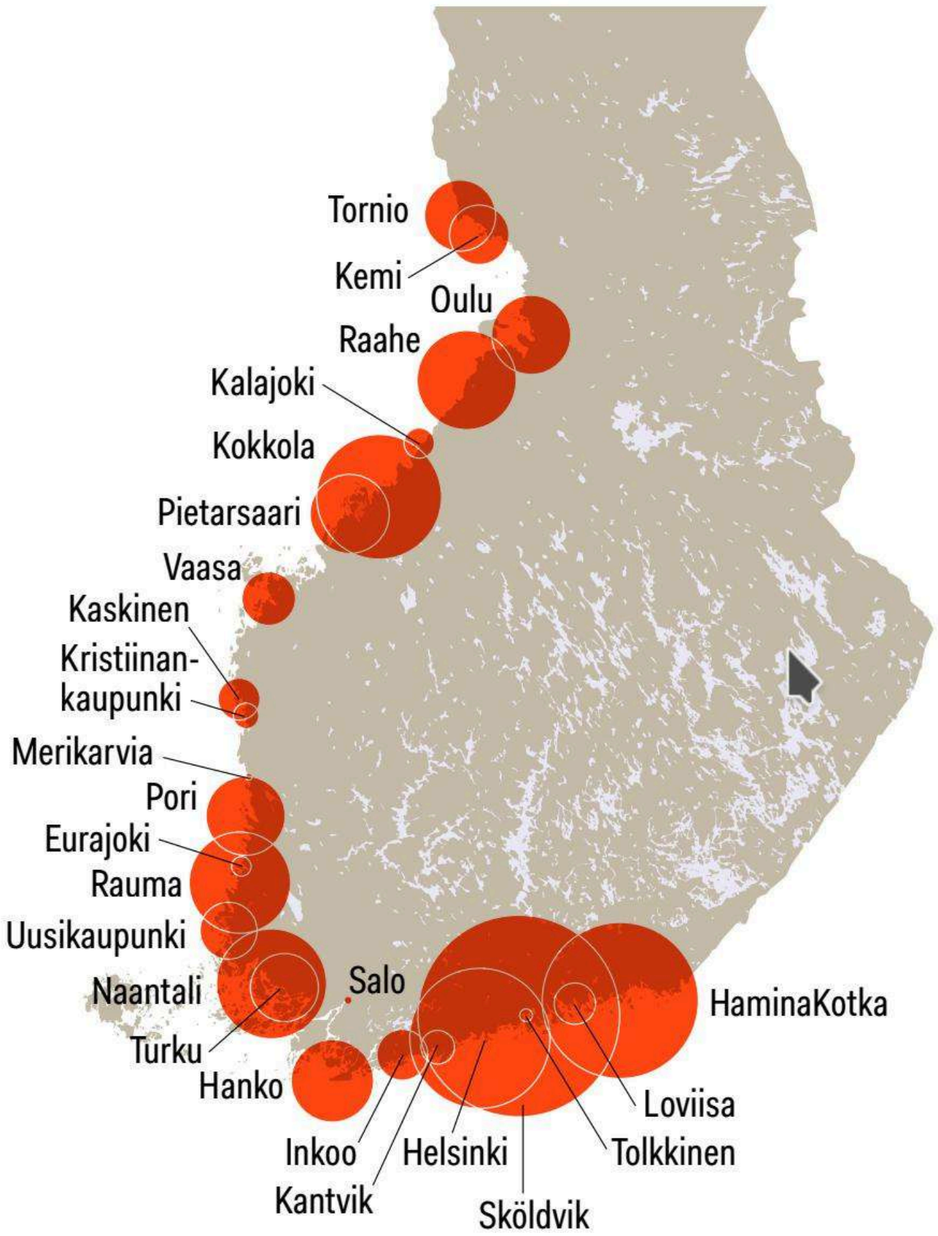



CODE
200,000
150,000
100,000
50,000
0
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010

MKhondkar published on TradingView.com, September 17, 2017 21:28 EDT
BATS:AAPL, W 159.88 ▲ +1.60 (+1.01%) O:160.50 H:163.96 L:157.91 C:159.88
AAPL
Vol
MA
MA
MA
1st iPhone Release
June 29 , 2007.
Yearly Return: 103%
iPhone 4 Release
June 7, 2010
Yearly Return: 53%
iPhone 5 Release
Sep 21, 2012
Yearly Return: 31%
iPhone 6 Release
Sep 9, 2014
Yearly Return: 38%
iPhone 7 Release
Sep 7, 2016
Yearly Return: 10%
iPhone 8 & X
Sep 12 , 2017.
159.88
156.26
120.36
111.91
221M
Volume MA
2006
2008
2010
2012
2014
2016
2018
2020
Created with TradingView

GRADUATED SYMBOL MAP

Ports on Finland's mainland coast, bubble size corresponds to annual cargo traffic volume



S. Kundu Ph.D.

Join at

slido.com

#342

S. Kundu Ph.D.

Join at

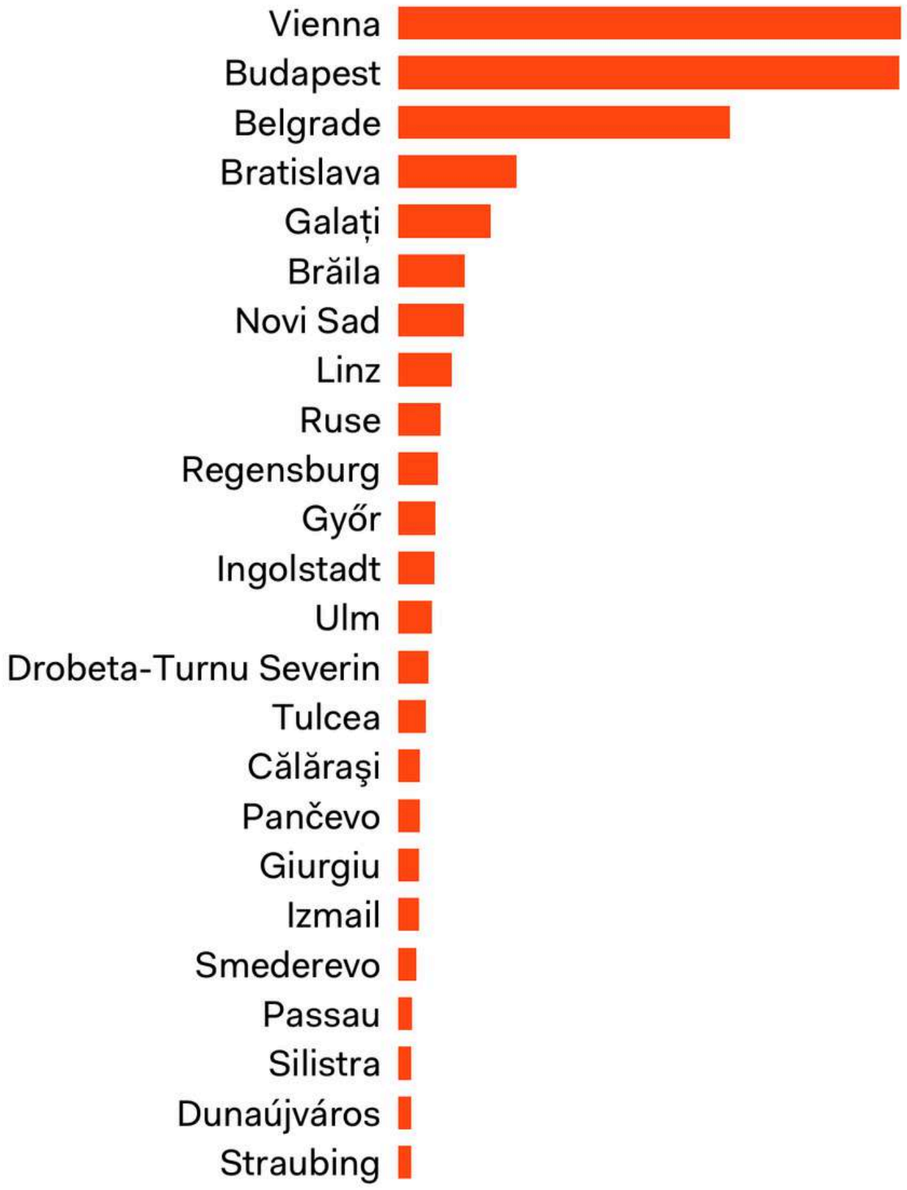
slido.com

#3520

CZECH REP.
SLOVAKIA
UKRAINE
MOLDOVA
Regensburg
Ingolstadt
Straubing
Bratislava
Passau
Győr
Ulm
GERMANY
Linz
HUNGARY
Vienna
Budapest
AUSTRIA
Dunaújváros
Galaţi
Izmail
ROMANIA
Brăila
Tulcea
SLOVENIA
Novi Sad
CROATIA
Pančevo
Drobeta-Turnu
Severin
Silistra
Milan
Zagreb
Călăraşi
Giurgiu
Belgrade
Smederevo
Ruse
BOSNIA
AND HERZ.
DANUBE
ITALY
SERBIA
BLACK SEA
BULGARIA
ADRIATIC SEA
MONTENEGRO
KOSOVO
Sofia
TURKEY
MACEDONIA
Rome
ALBANIA
GREECE
0
250KM
155MI

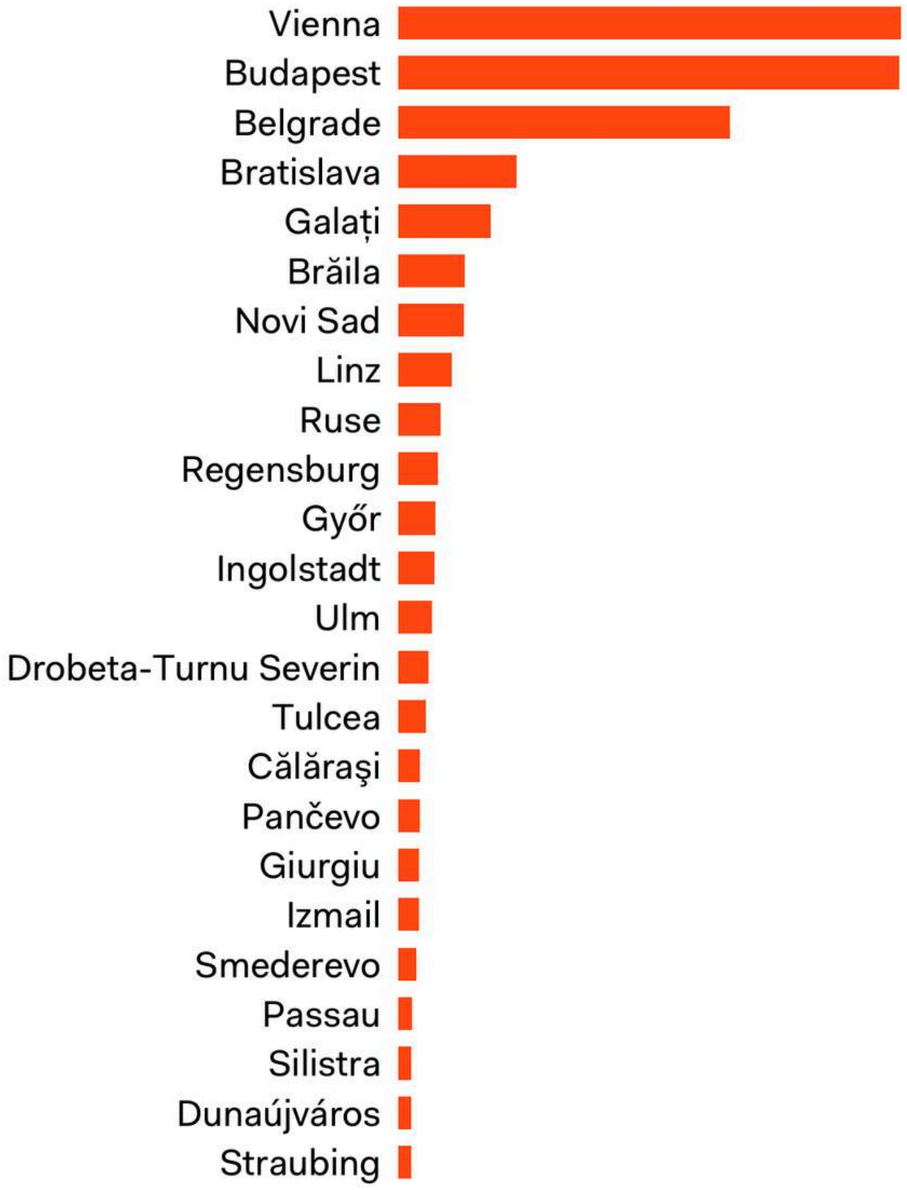
CONTINUUM BY MAGNITUDE

By population



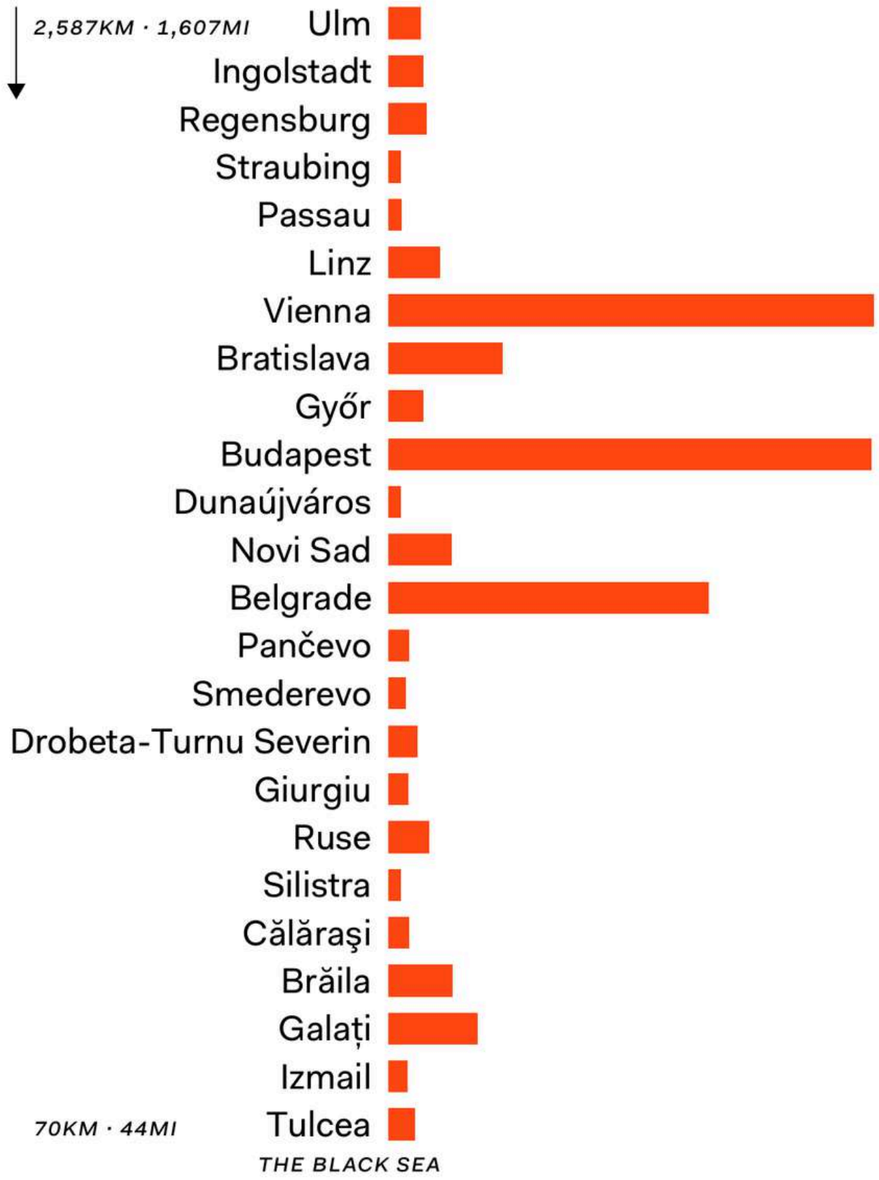
CONTINUUM BY MAGNITUDE

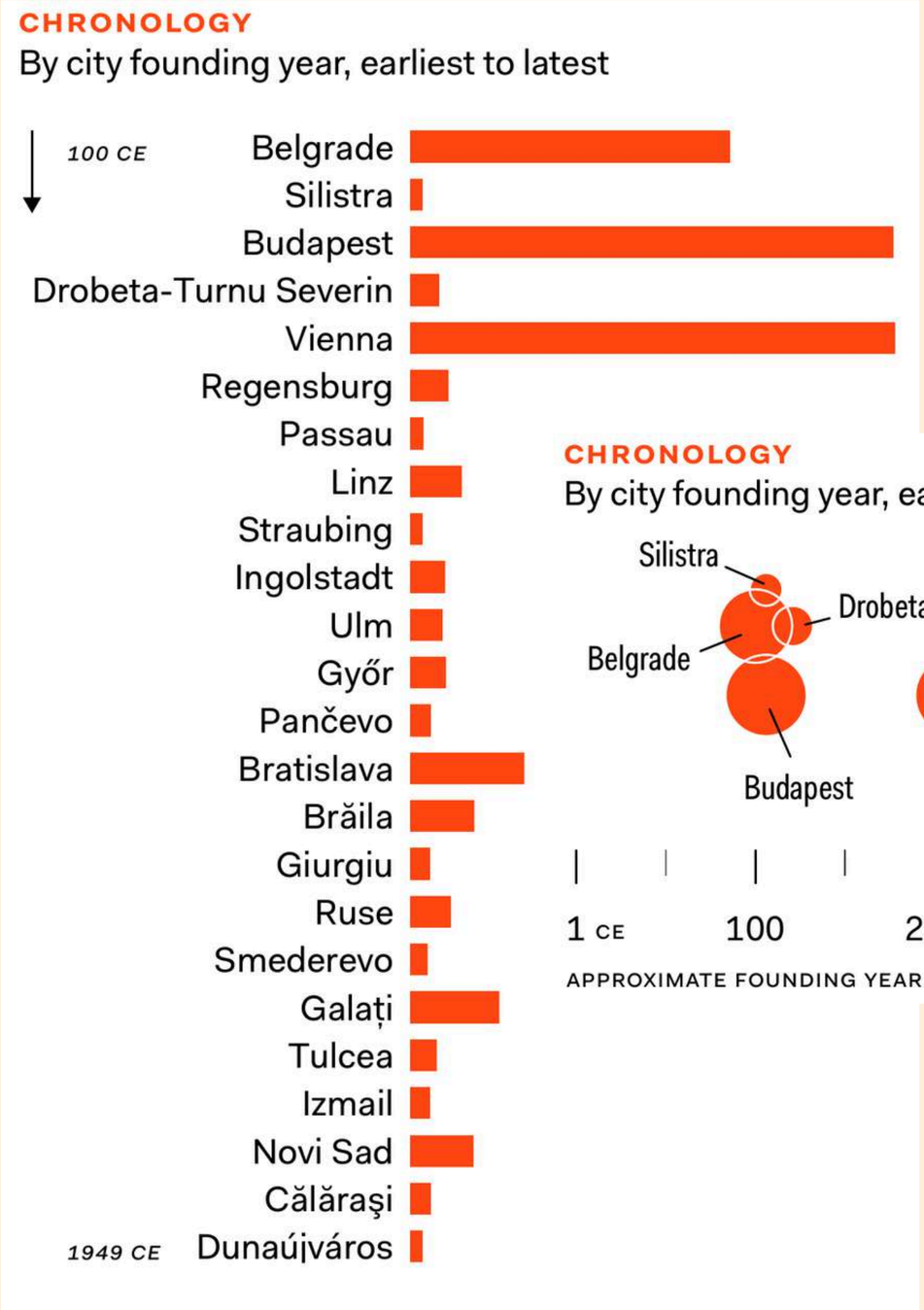
By population

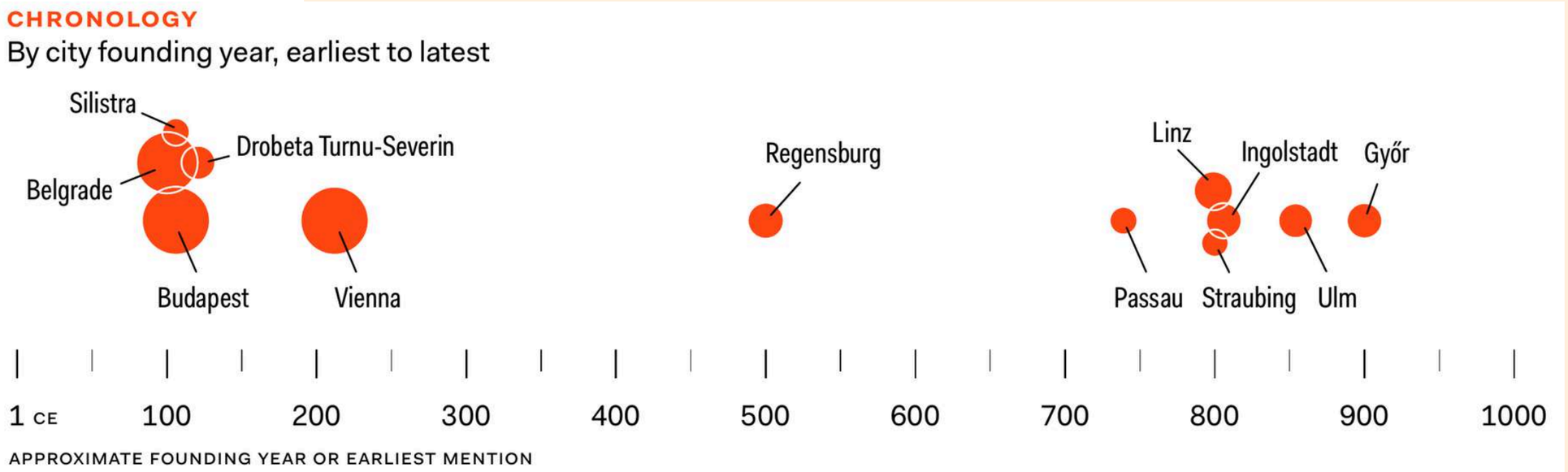


COORDINATES

By location along the Danube downstream from Ulm

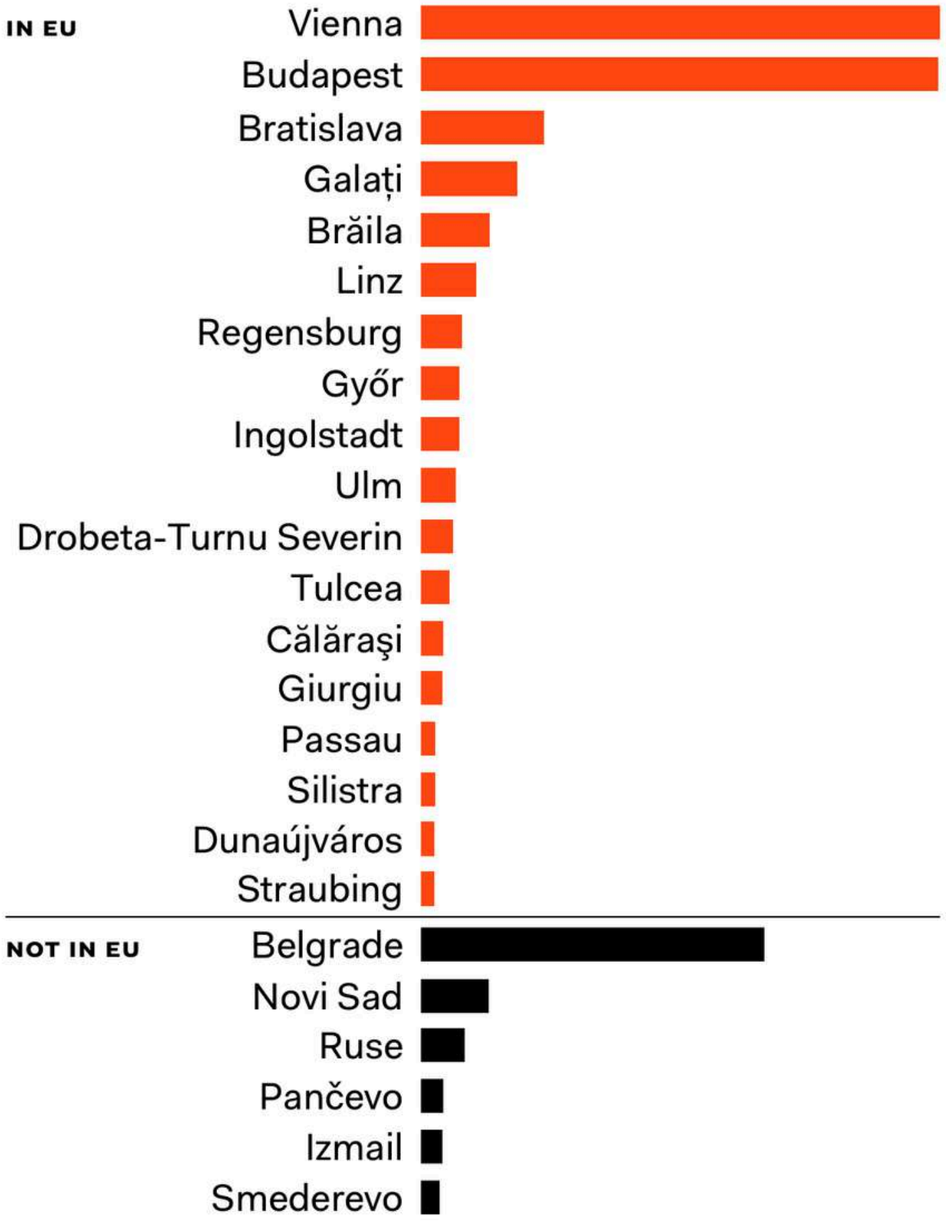



CHRONOLOGY
By city founding year, earliest to latest
100 CE
Belgrade
Silistra
Budapest
Drobeta-Turnu Severin
Vienna
Regensburg
Passau
Linz
Straubing
Ingolstadt
Ulm
Győr
Pančevo
Bratislava
Brăila
Giurgiu
Ruse
Smederevo
Galați
Tulcea
Izmail
Novi Sad
Călăraşi
1949 CE
Dunaújváros


CHRONOLOGY
By city founding year, earliest to latest
Silistra
Drobeta Turnu-Severin
Belgrade
Budapest
Vienna
Regensburg
Linz
Ingolstadt
Győr
Passau
Straubing
Ulm
1 CE
100
200
300
400
500
600
700
800
900
1000
APPROXIMATE FOUNDING YEAR OR EARLIEST MENTION

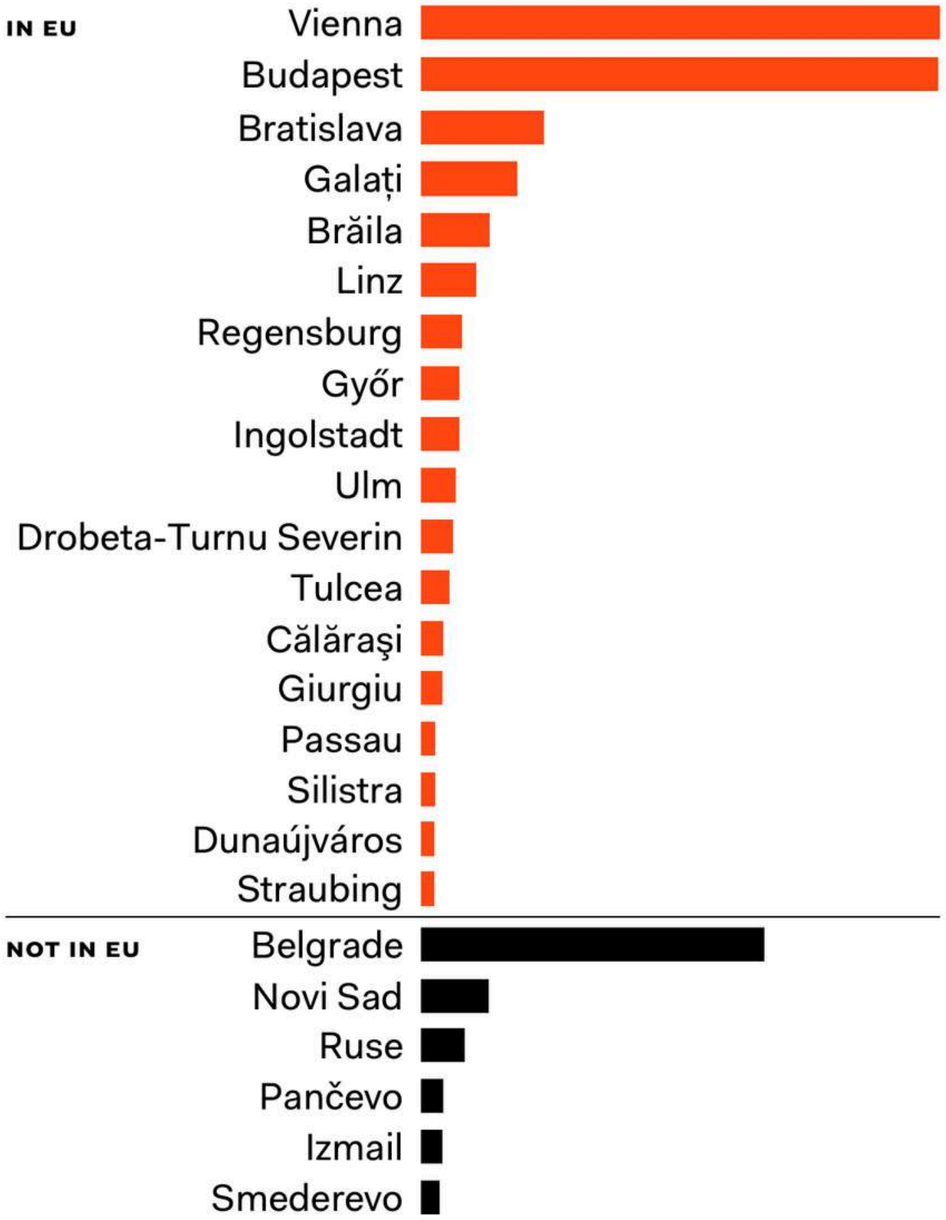
CATEGORIES

Cities in European Union member countries



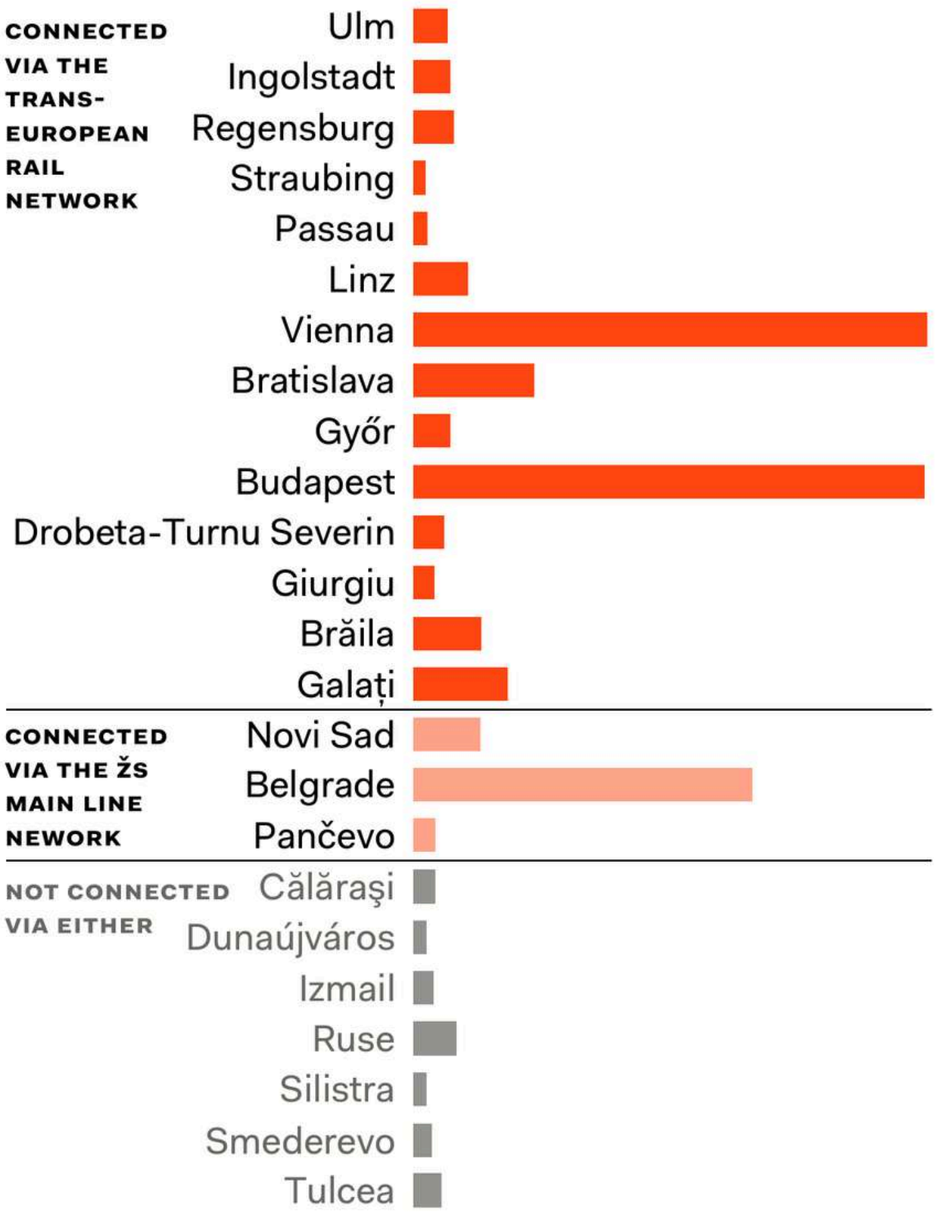
CATEGORIES

Cities in European Union member countries



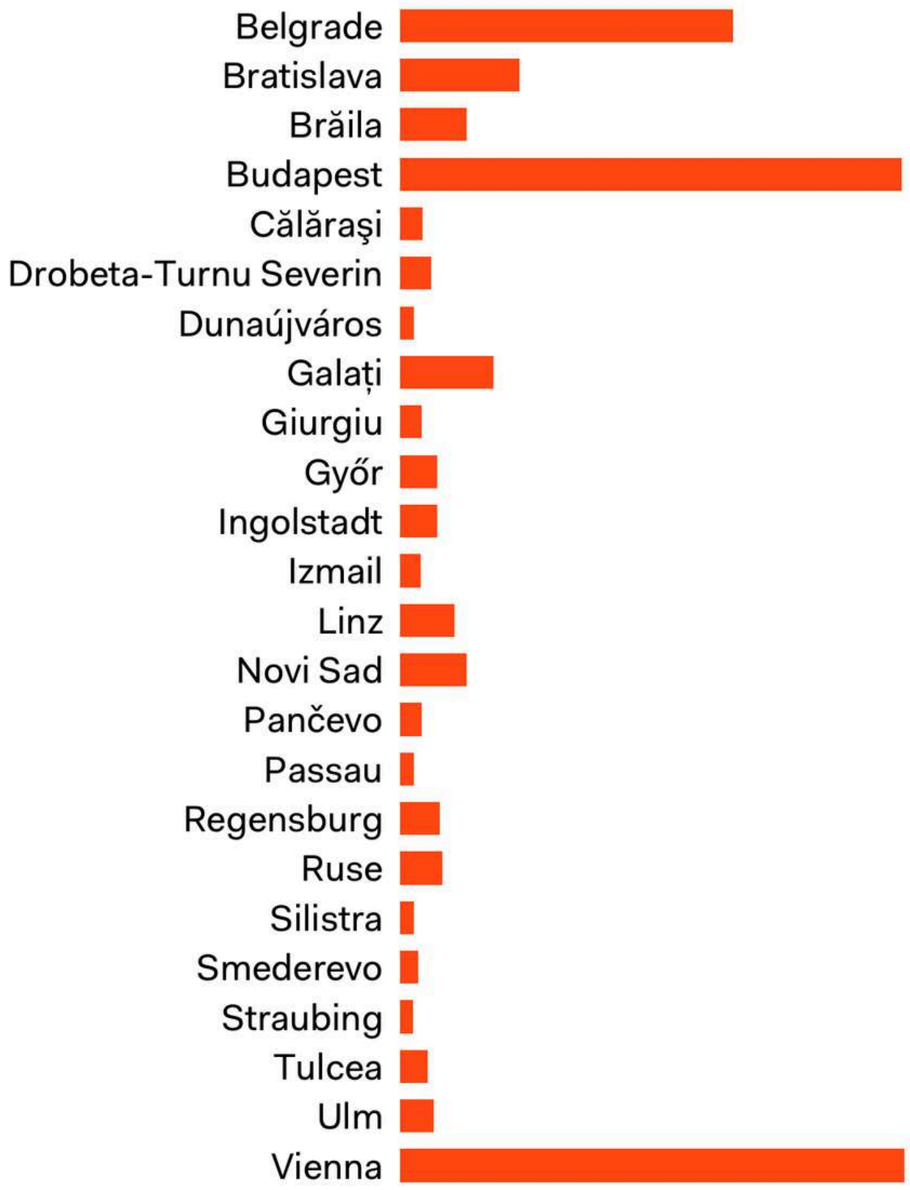
CONNECTION

By railroad connections

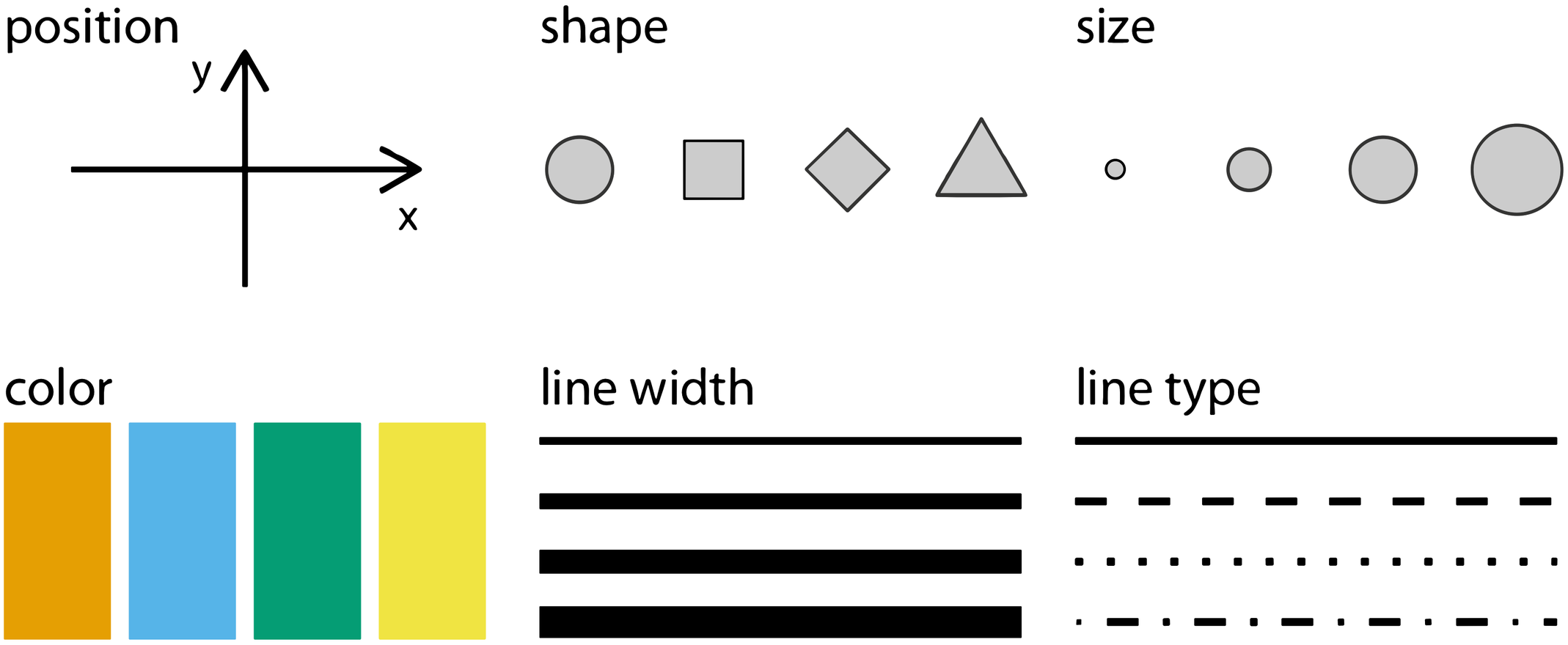


CONVENTION

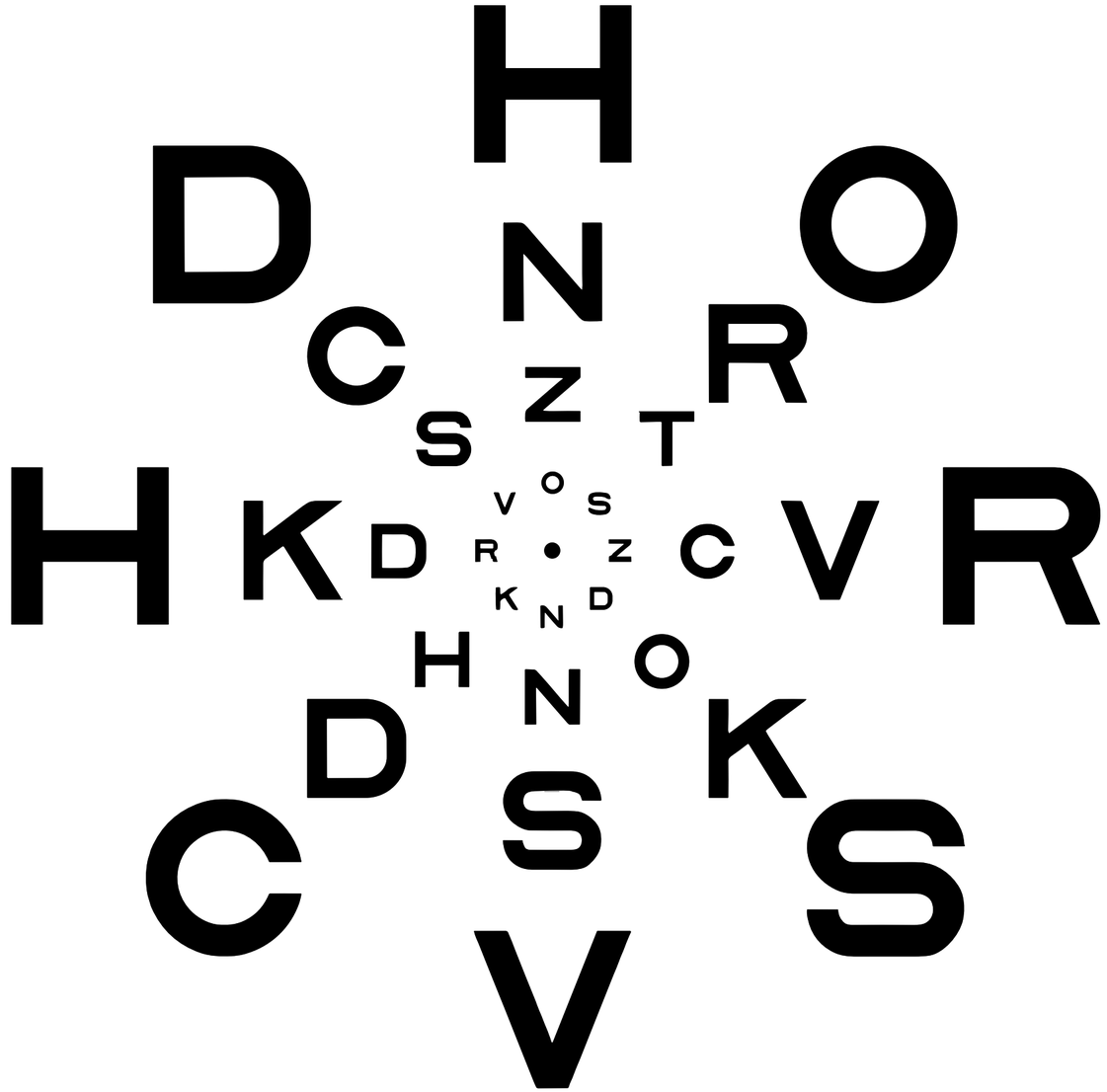
Alphabetical order by name

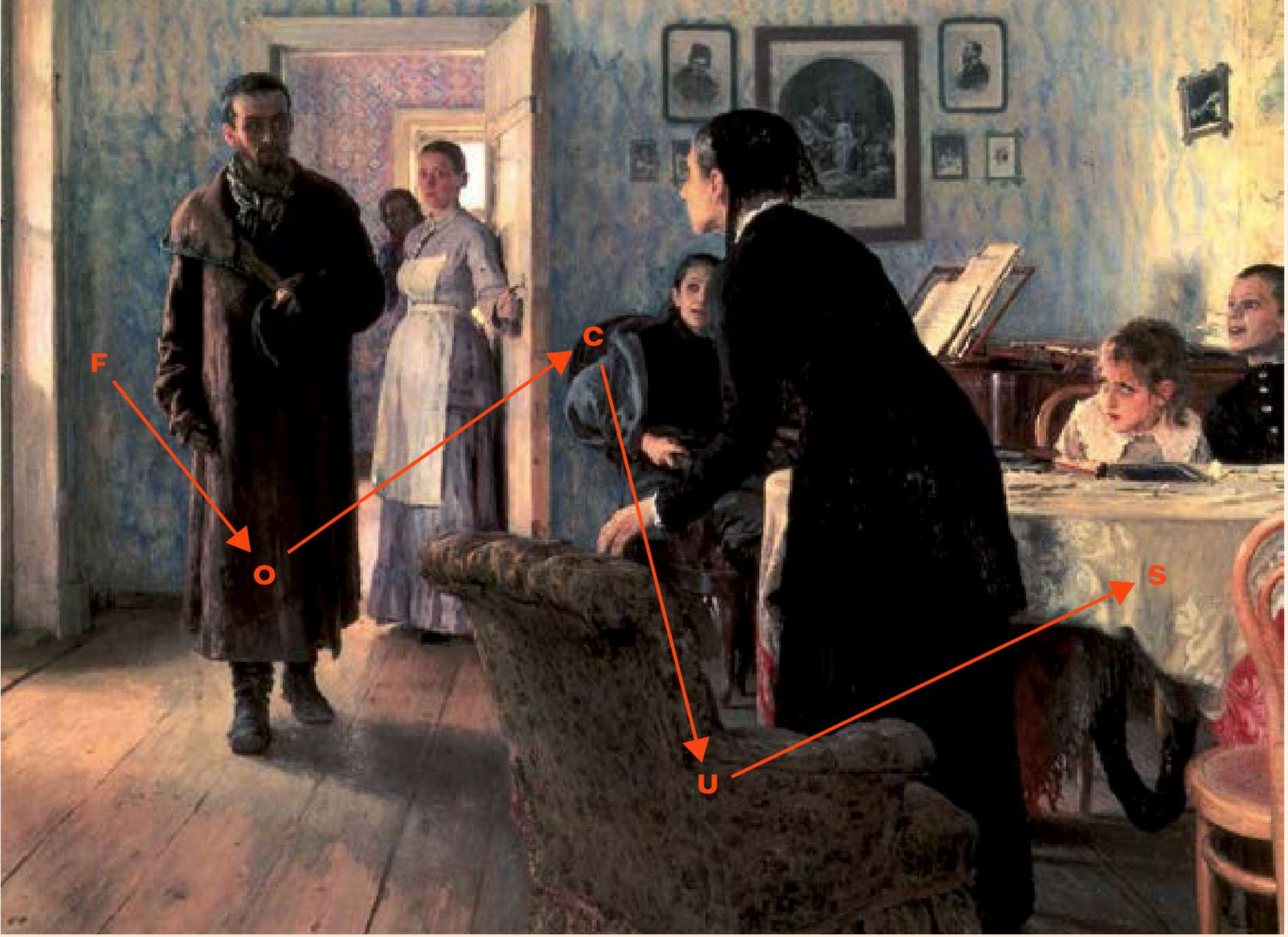


Aesthetics



Anstis' Peripheral Acuity Chart

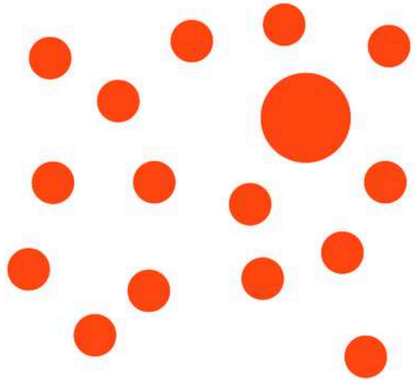




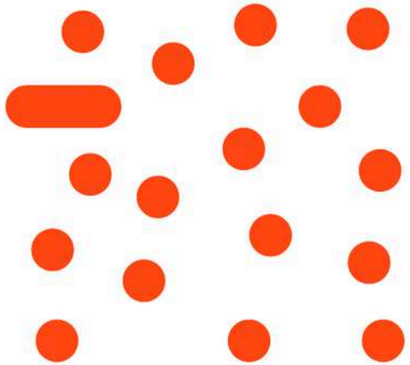
If the attention is focused on the letters drawn on top of the painting, the painting is "pushed back" from the view. If the attention is focused on the details of the painting the text becomes unreadable instead.

Shape

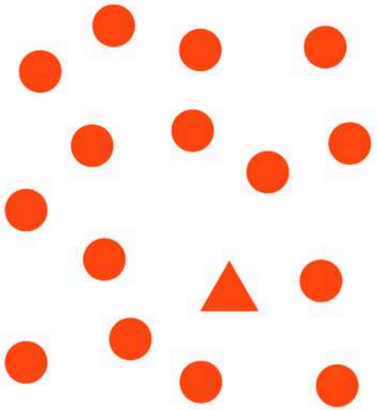
size



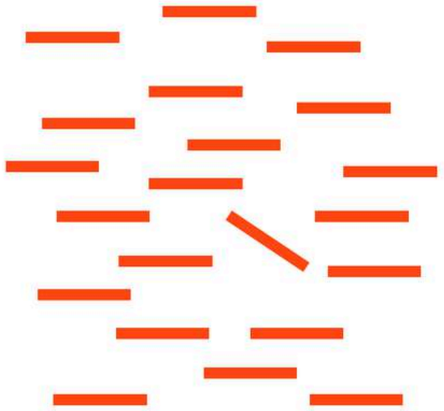
elongation



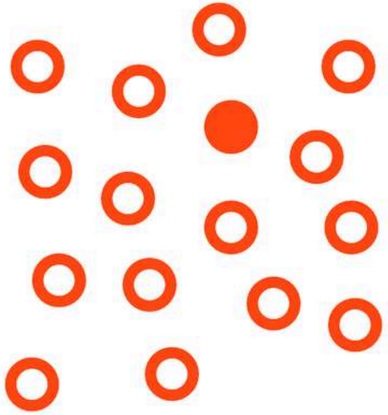
round vs. sharp



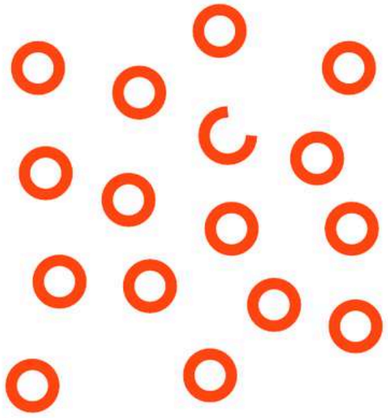
orientation



fill

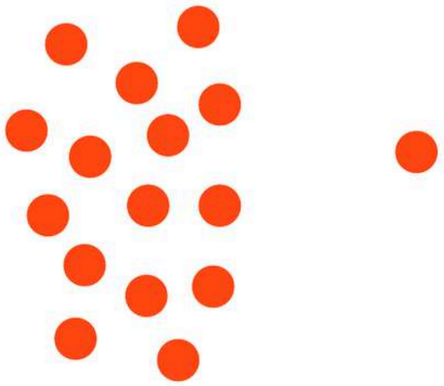


closure

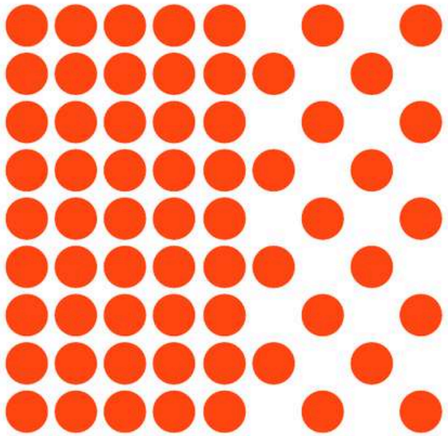


Position

grouping, quantity

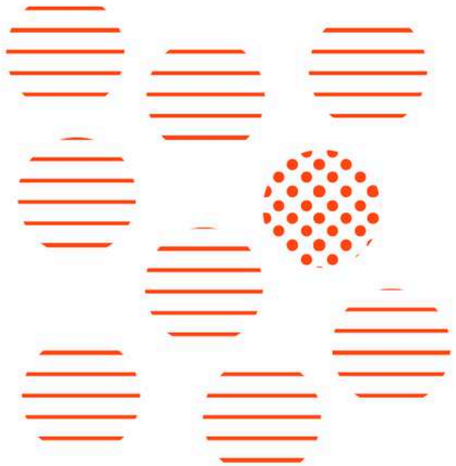


density

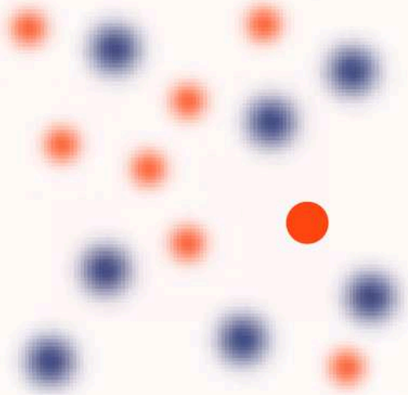


Other shape features

texture

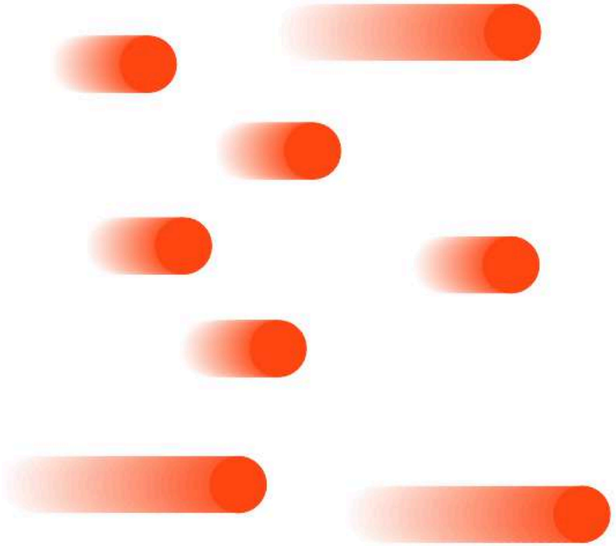


sharpness

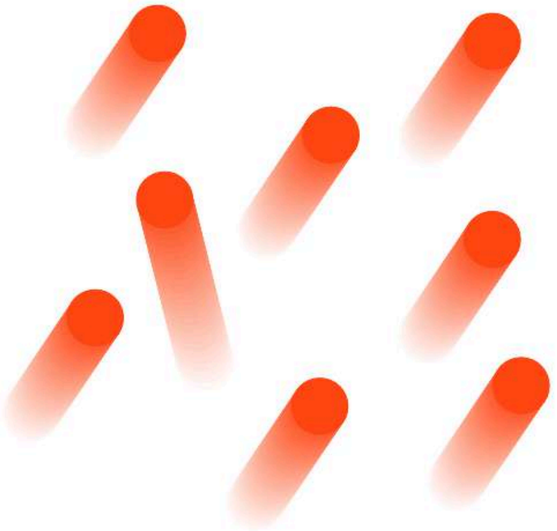


MOTION AND CHANGE

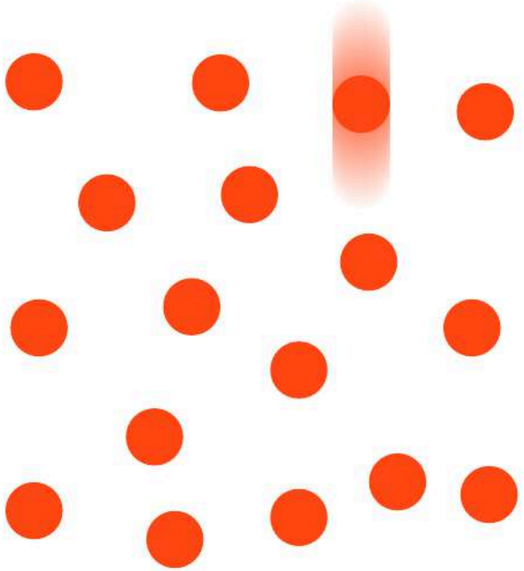
speed



direction

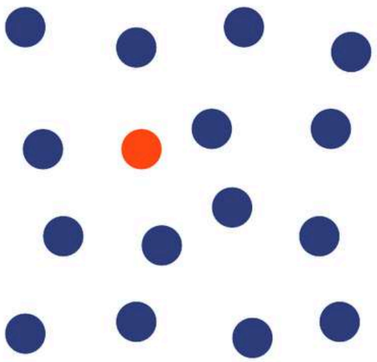


vibration

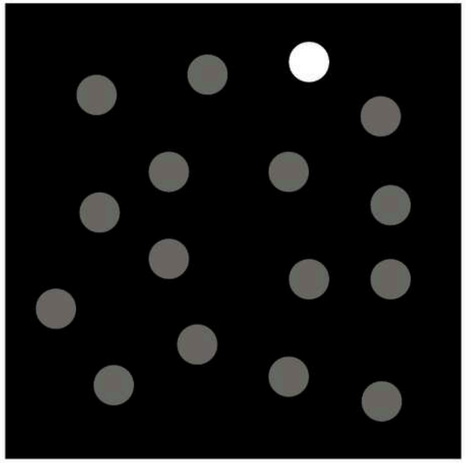


COLOR FEATURES

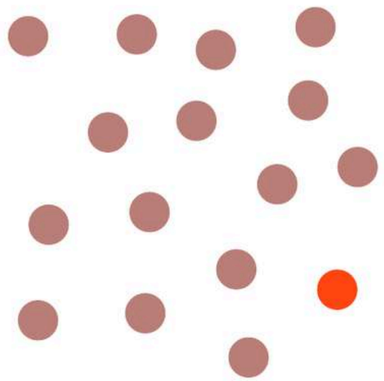
hue



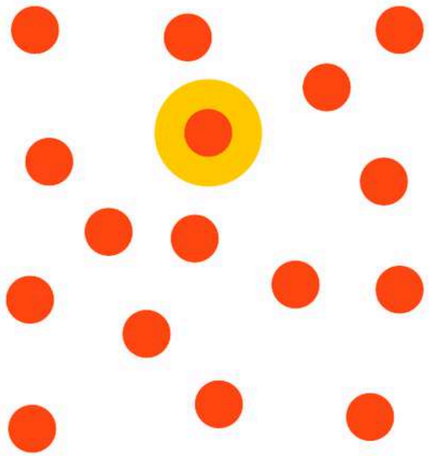
lightness



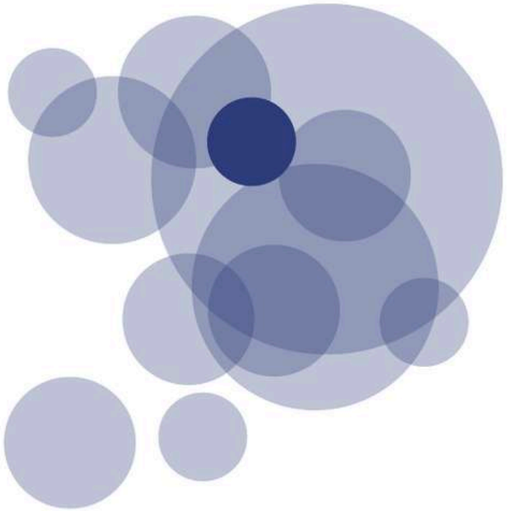
intensity



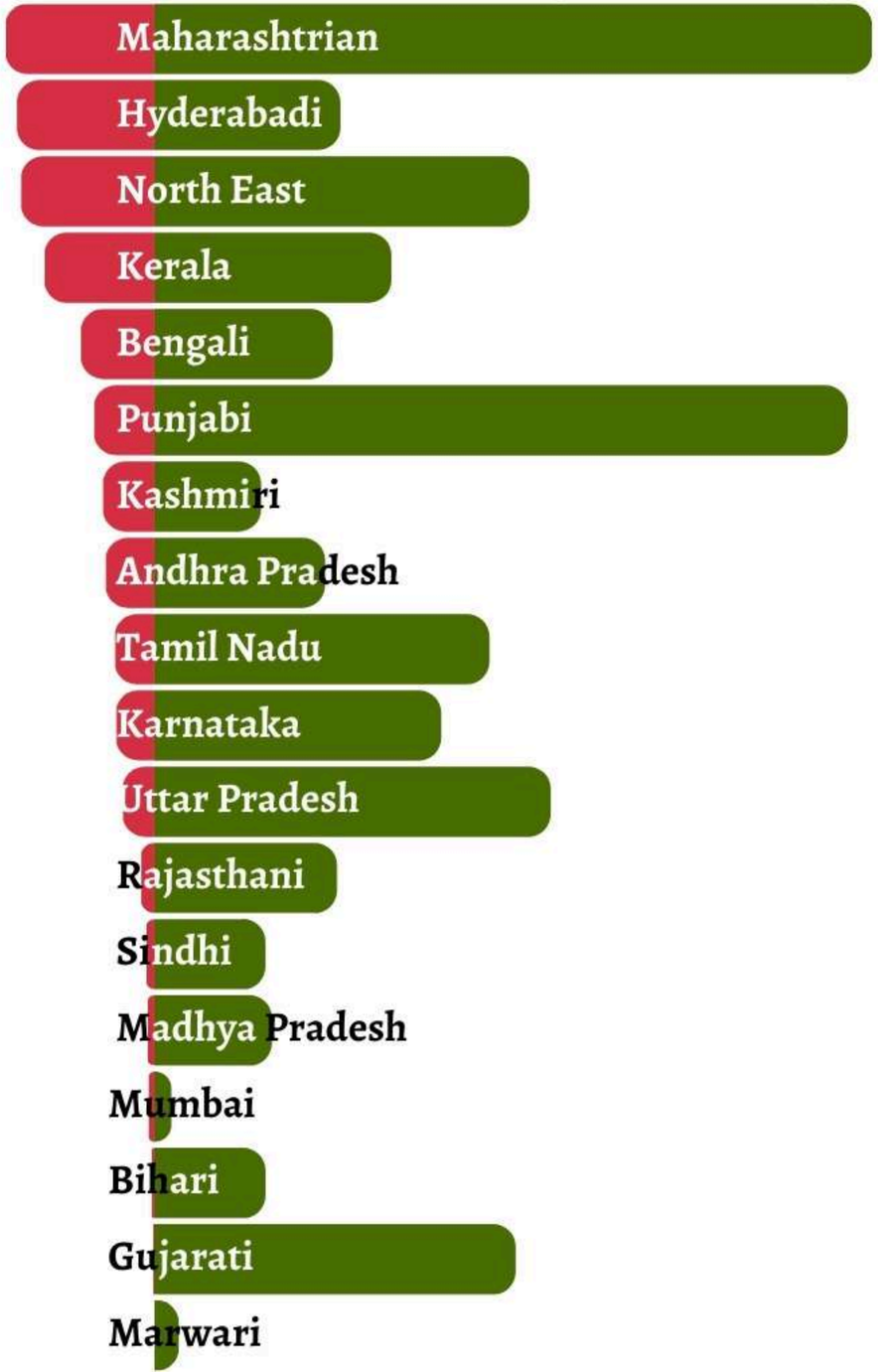
added surround color



opacity



Cuisines Distribution (Non-vegetarian vs Vegetarian)



S. Kundu Ph.D.

Join at

slido.com

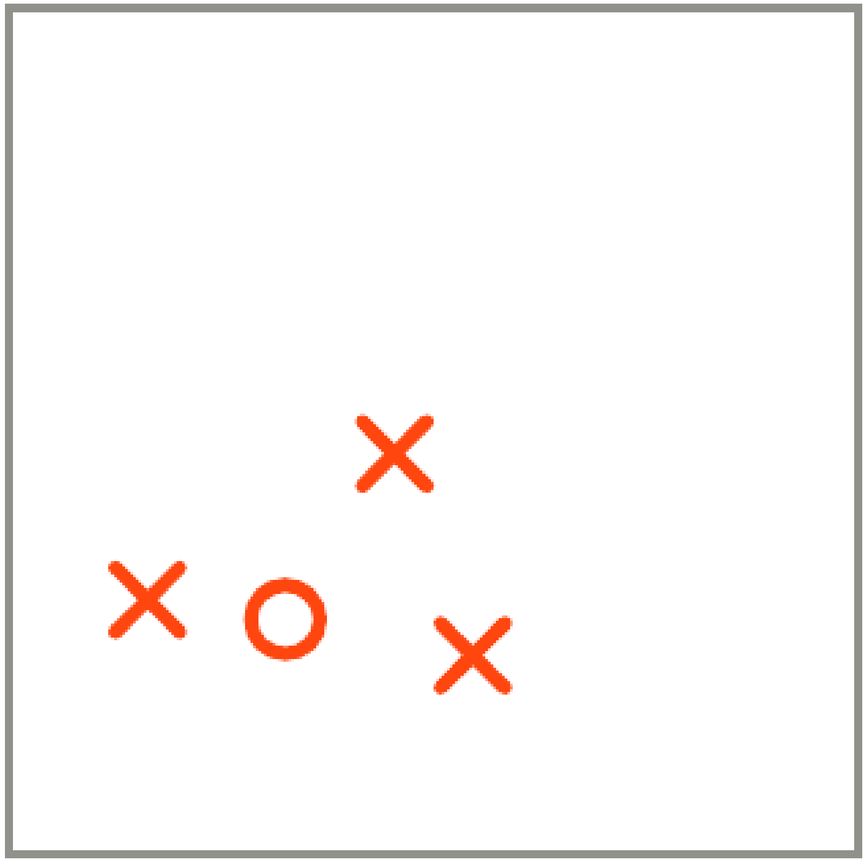
#3520

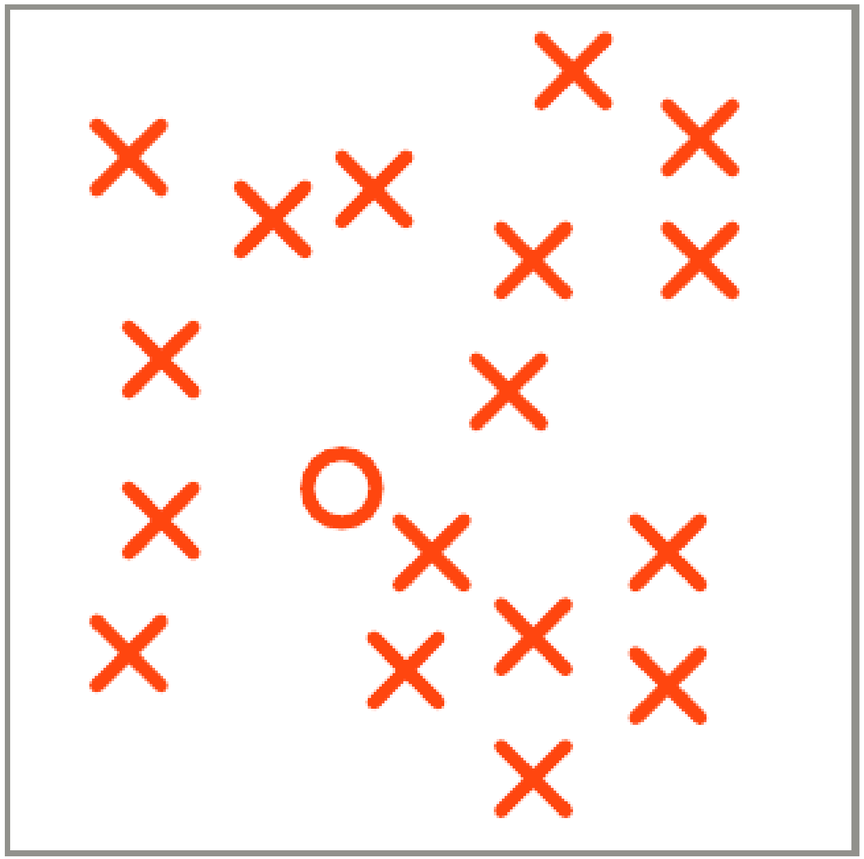
Although most people are good at recognizing other people's faces, spotting an individual in a crowd based on their facial features is difficult.

Whereas a person's face is a collection of features too complex to "pop out" from its surroundings, a person wearing unusually colored clothes stands out easily.

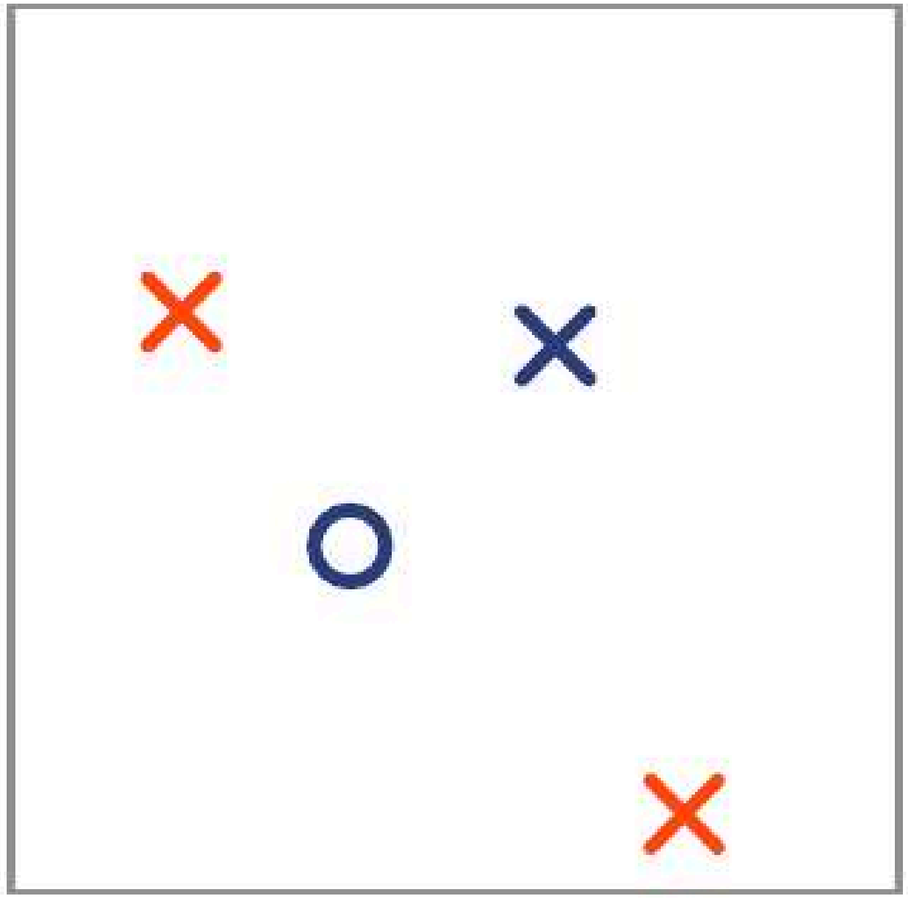
818889319180970813093 13	818889319180970813093 13
914947105734154510190 74	914947105734154510190 74
810875821635759064713 06	810875**2**16357590647130 6
609080081365694056114 58	609080081365694056114 58
072696361609030555309 56	07**2**69636160903055530 956
804557111073897096530 17	804557111073897096530 17

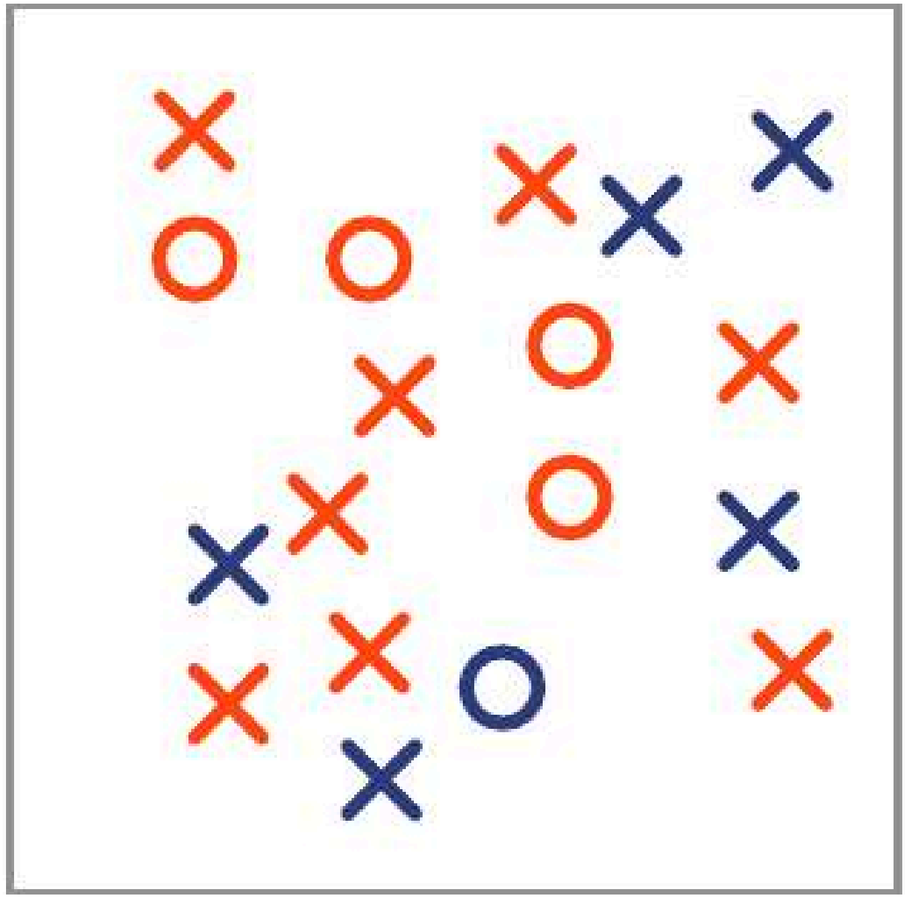
Feature search: find ○



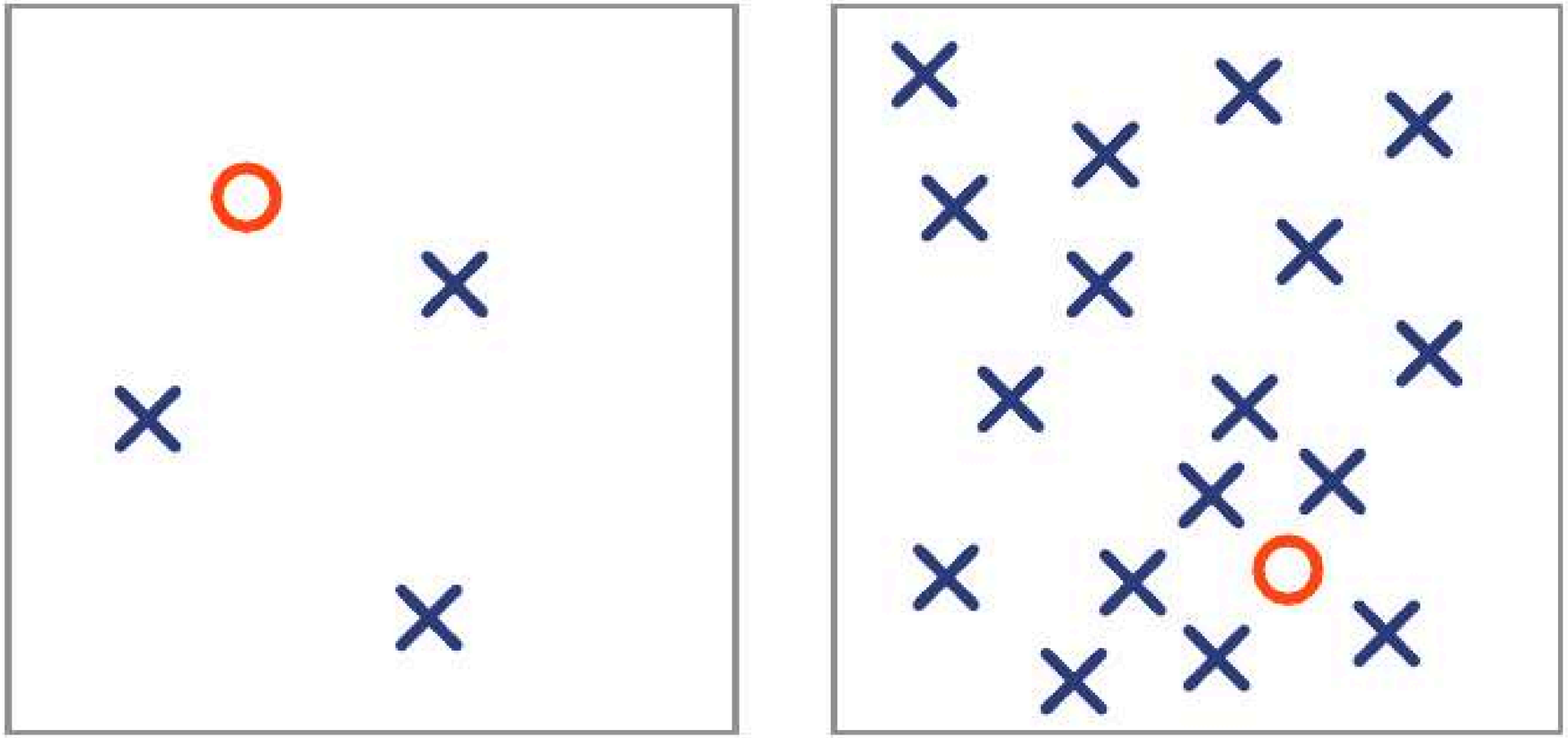


Conjunction search: find





Feature search with reinforcing features: find O



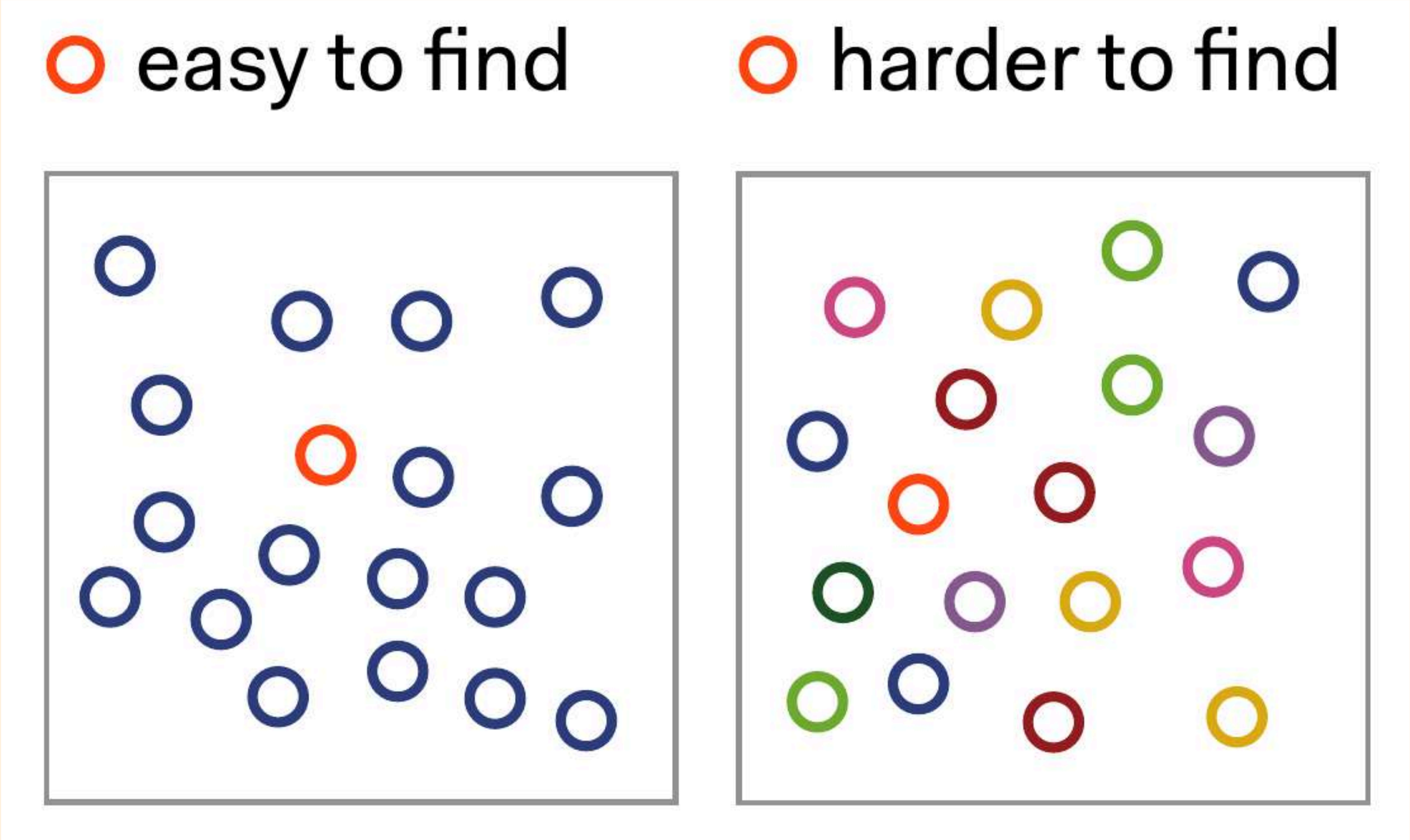

easy to find
harder to find

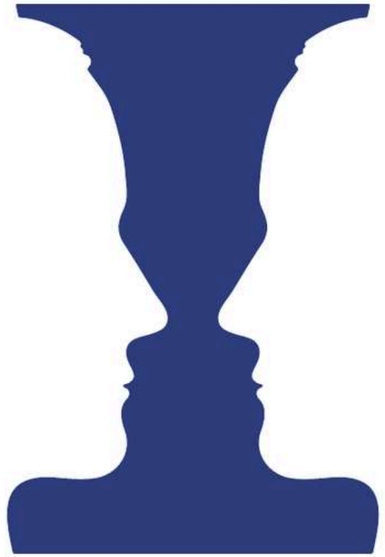
Figure-ground articulation

Clear figure-ground articulation

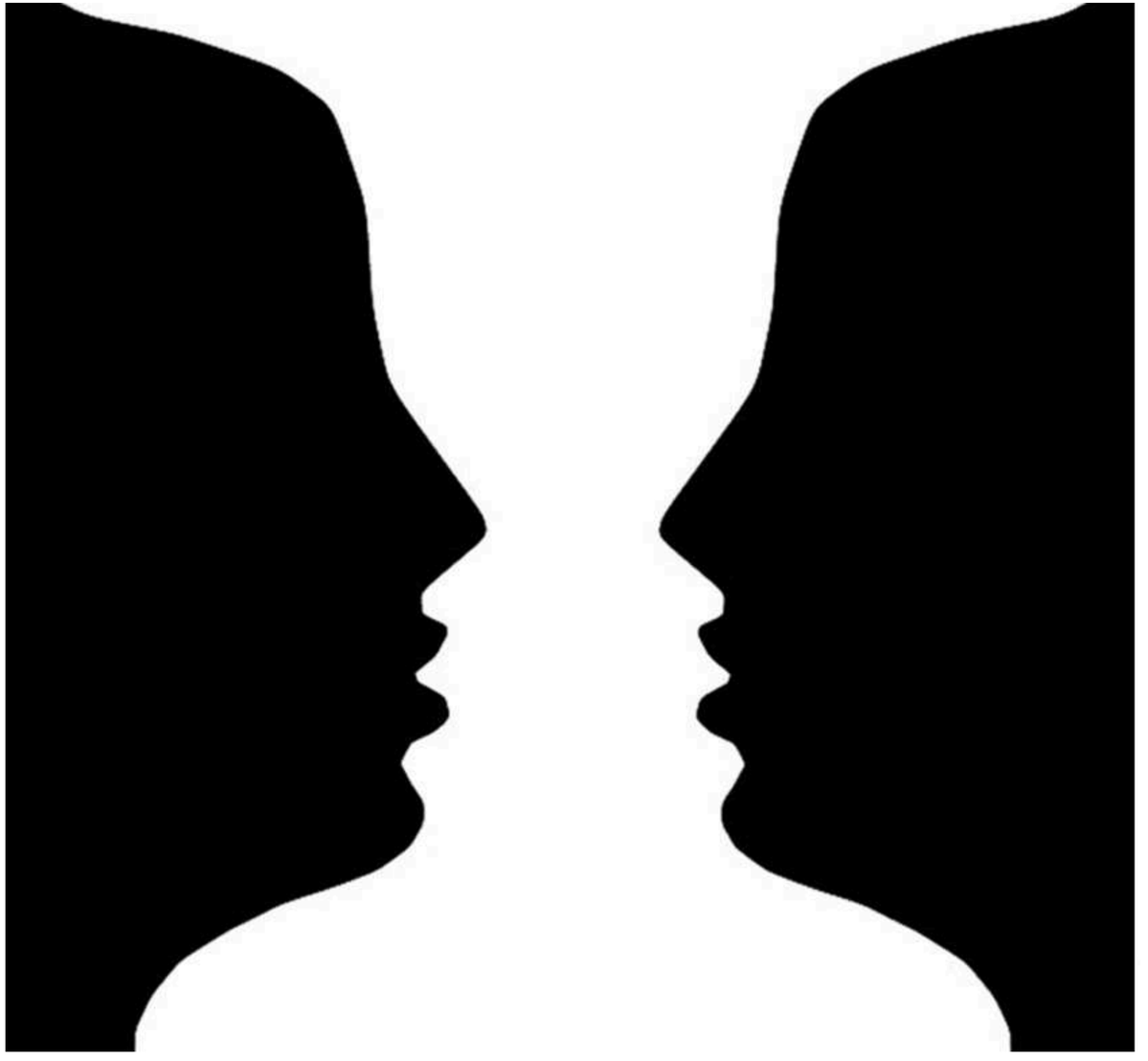
Recognizable object

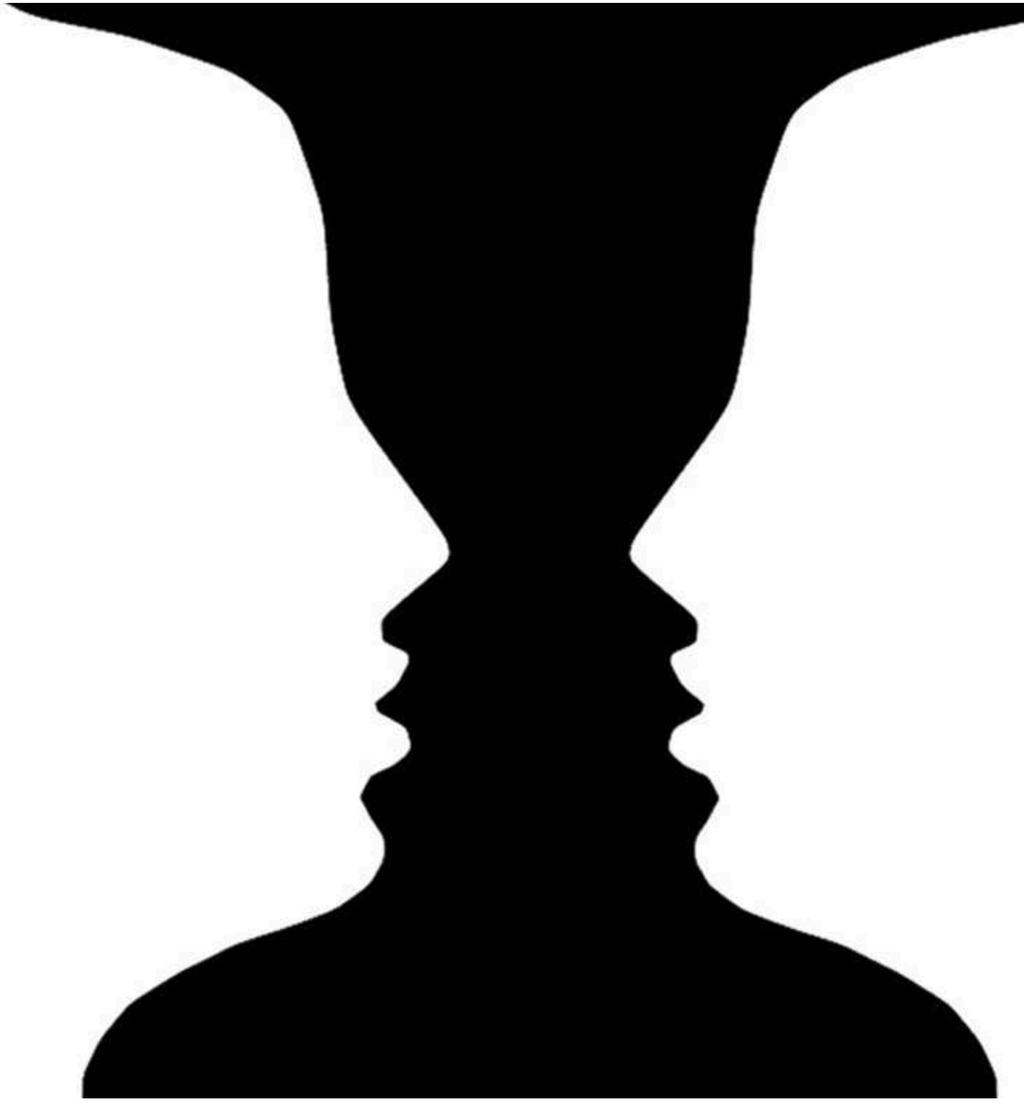
Evident lightness
and contrast hierarchy

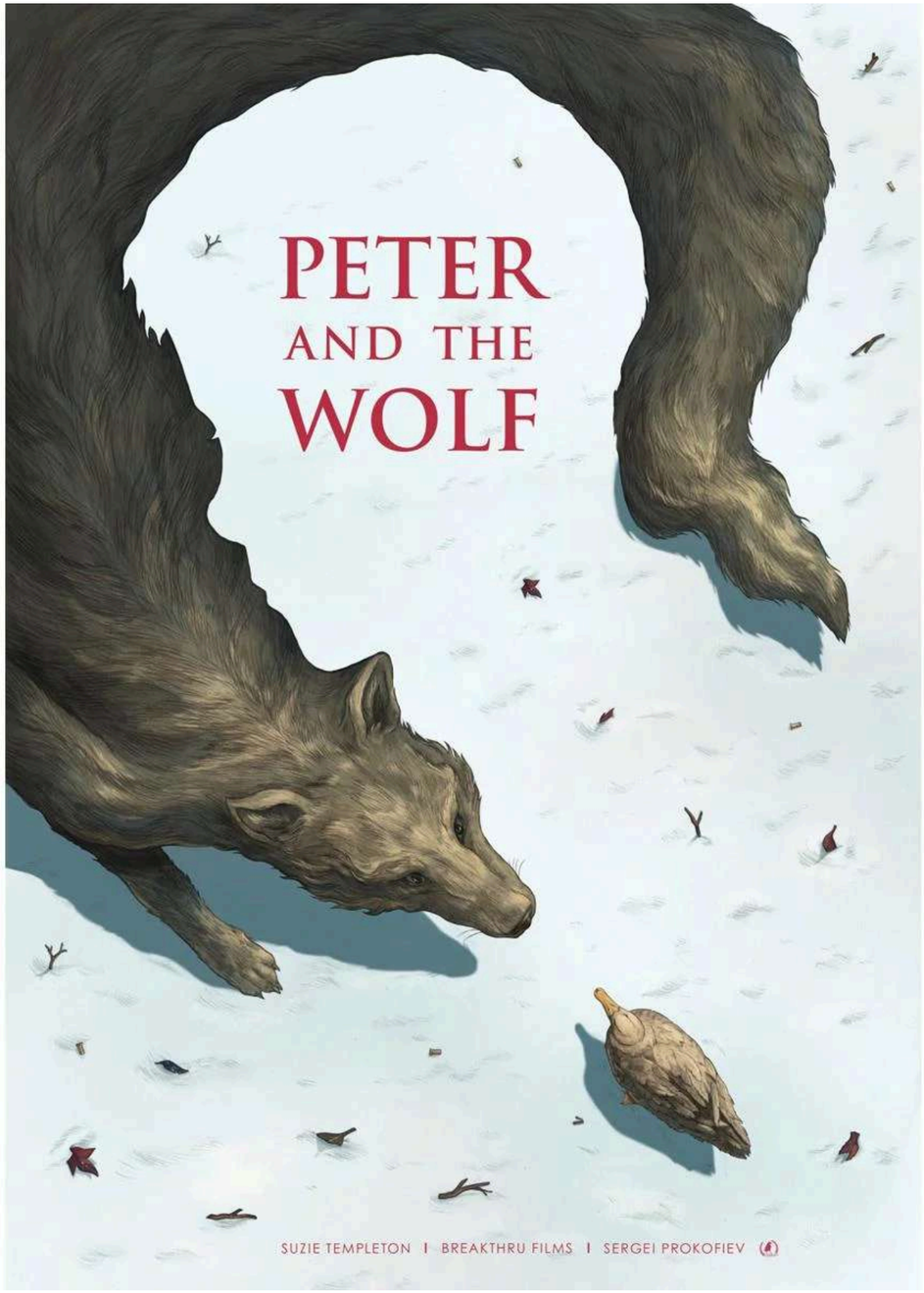
Multi-stable image



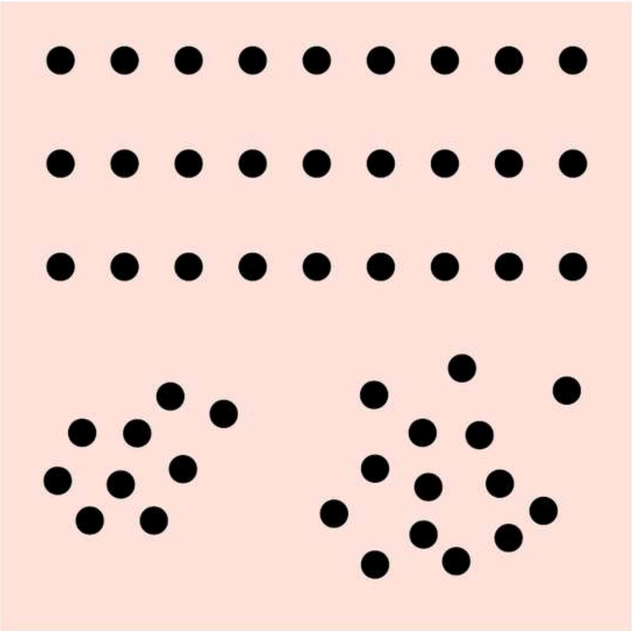
Rubin's vase

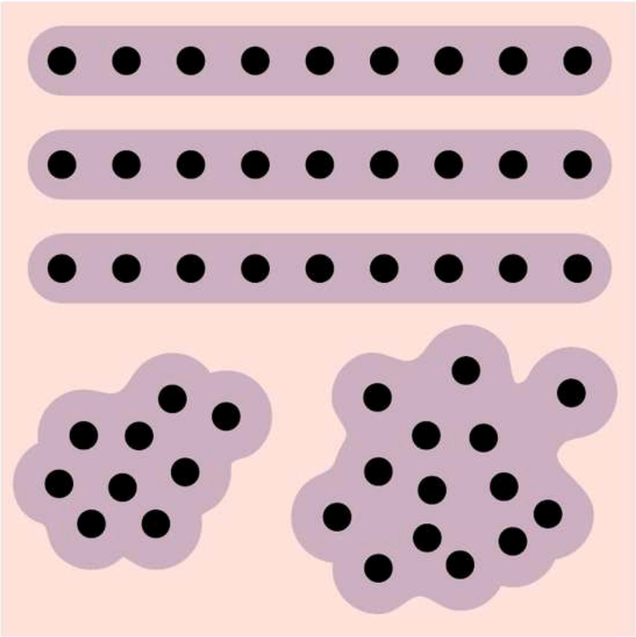




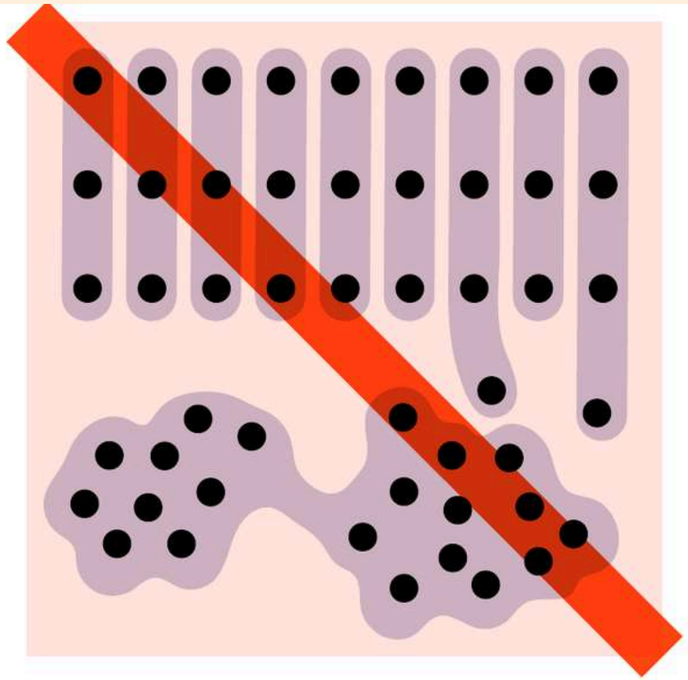

PETER
AND THE
WOLF
SUZIE TEMPLETON I BREAKTHRU FILMS I SERGEI PROKOFIEV

Law of proximity

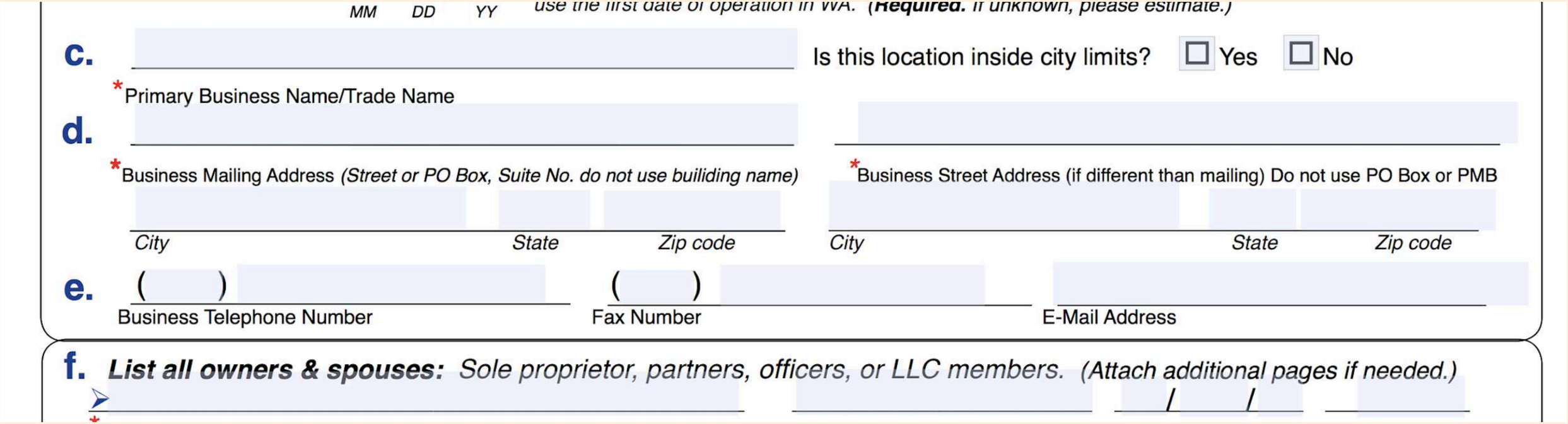




Perceptual decomposition based on the principle



Alternative decomposition violating the principle



MM DD YY use the first date of operation in WA. (**Required.** If unknown, please estimate.)

c. Is this location inside city limits? ☐ Yes ☐ No

*Primary Business Name/Trade Name

d.

*Business Mailing Address *(Street or PO Box, Suite No. do not use builiding name)*

*Business Street Address (if different than mailing) Do not use PO Box or PMB

City *State* *Zip code*

City *State* *Zip code*

e. ( ) Business Telephone Number

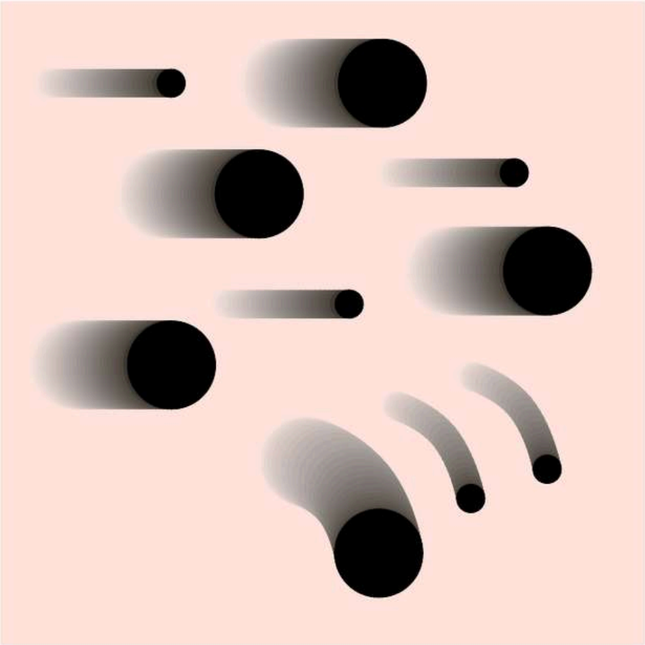
( ) Fax Number

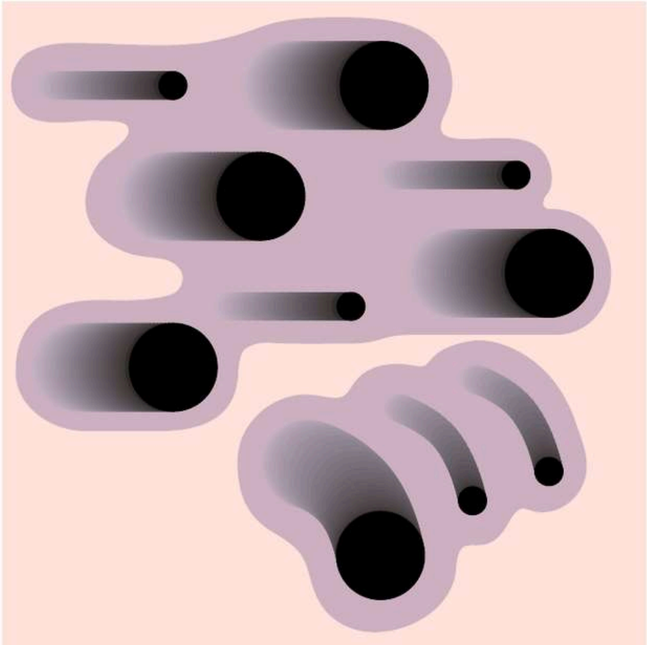
E-Mail Address

f. ***List all owners & spouses:*** *Sole proprietor, partners, officers, or LLC members. (Attach additional pages if needed.)*

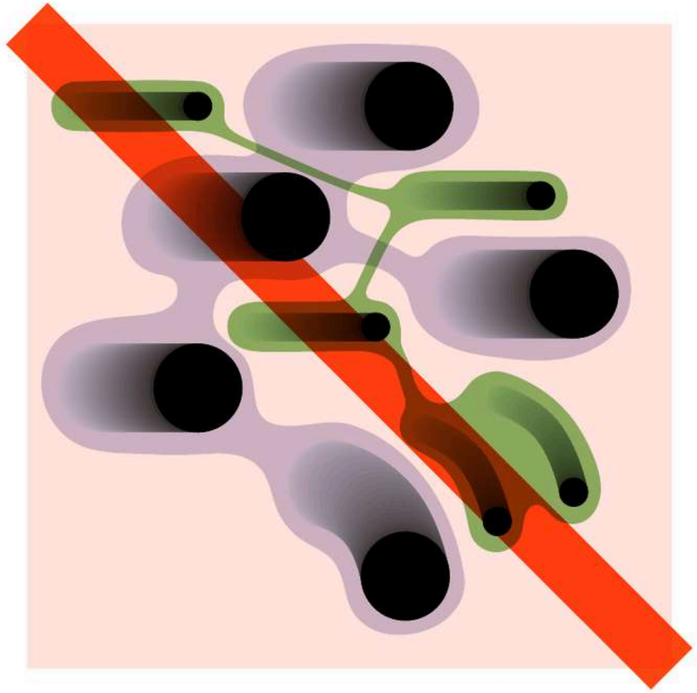
A USPS PDF form example where the proximity gestalt principle is impairing UX. Due to the field labels being closer to the fields under them, people would mistakenly believe that's where the information goes for "c." and "d." Yet, the information is supposed to be provided in the fields above the text labels.

Law of common fate



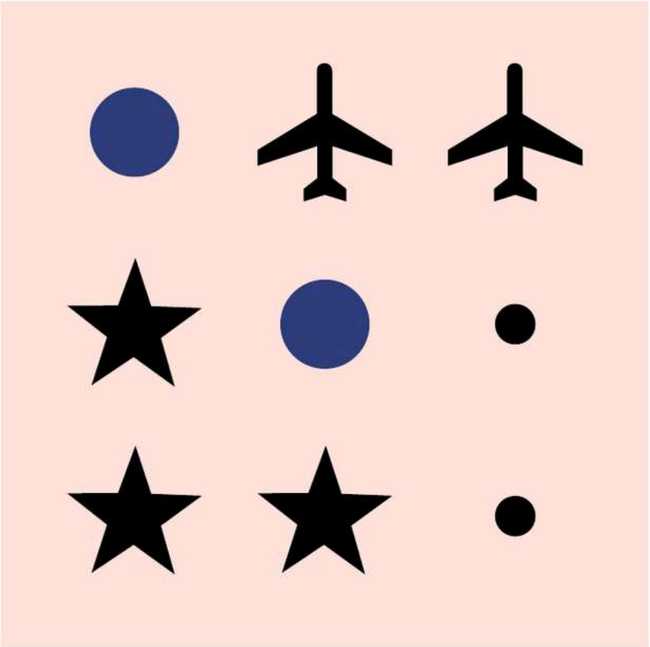


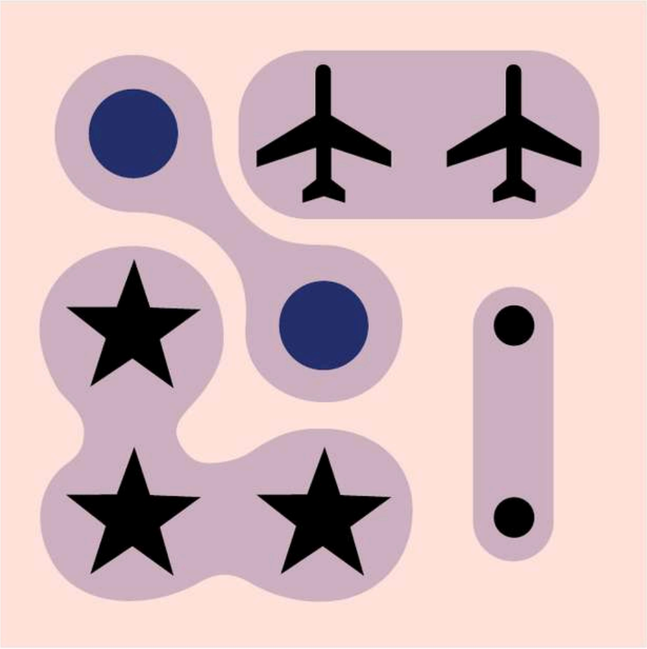
Perceptual decomposition based on the principle



Alternative decomposition violating the principle

Law of similarity

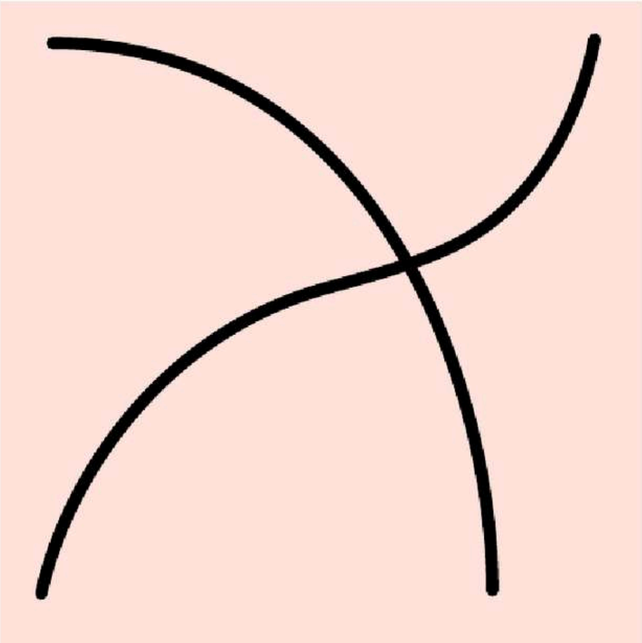


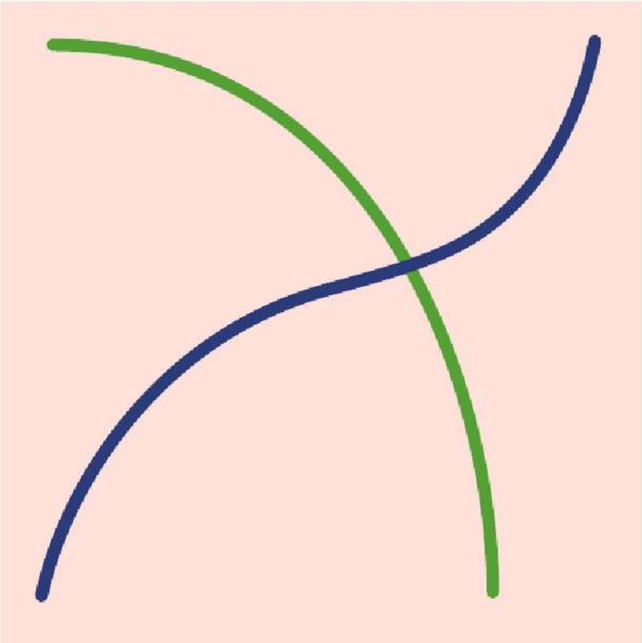


Perceptual decomposition based on the principle

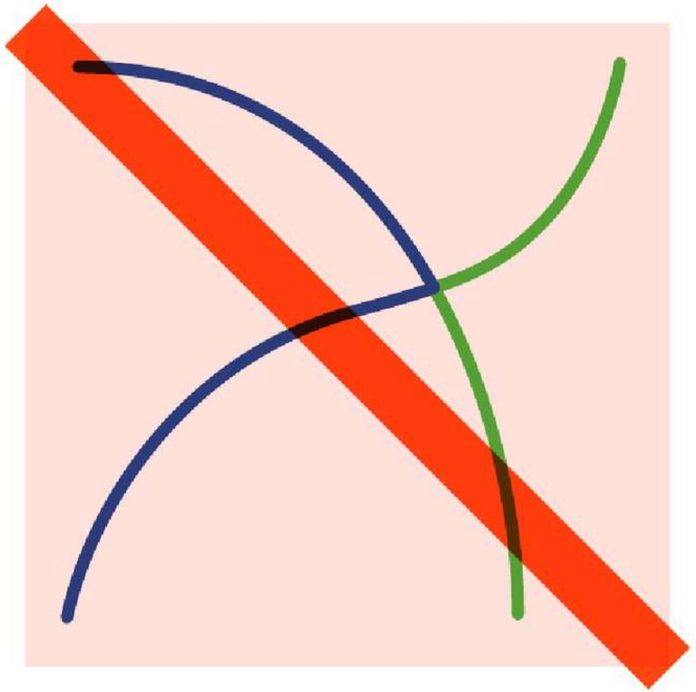
Alternative decomposition violating the principle

Law of continuity



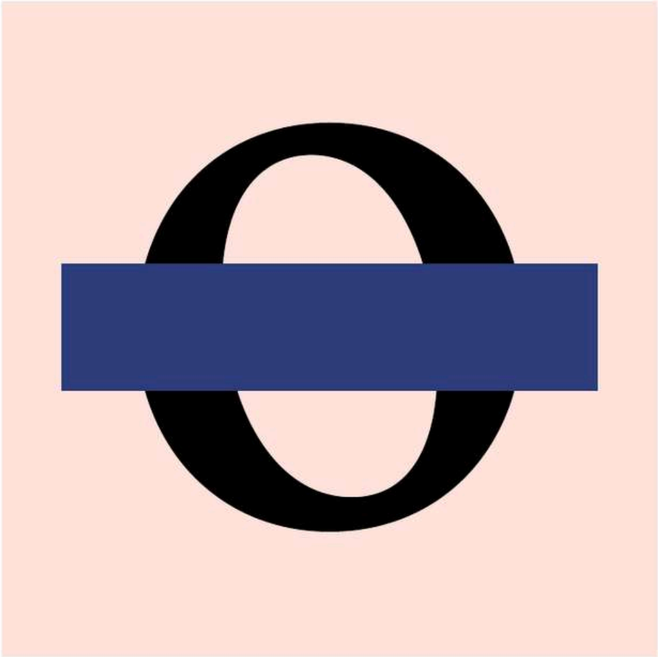


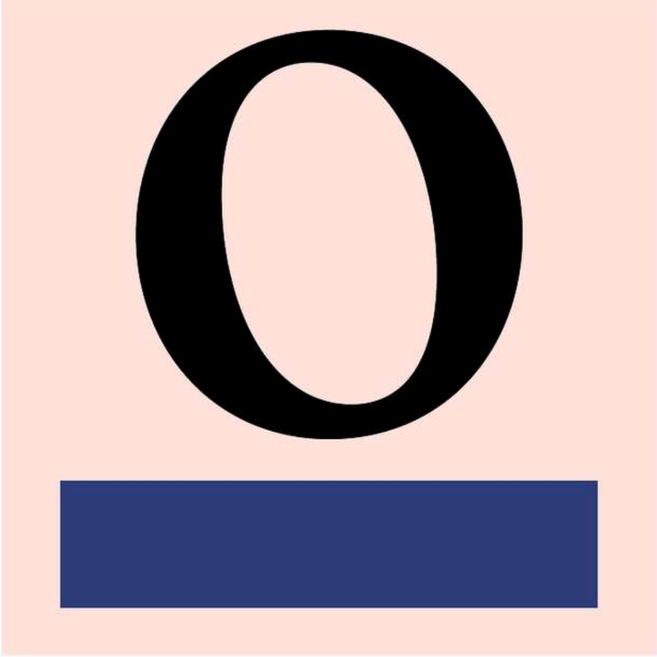
Perceptual decomposition based on the principle



Alternative decomposition violating the principle

Law of closure



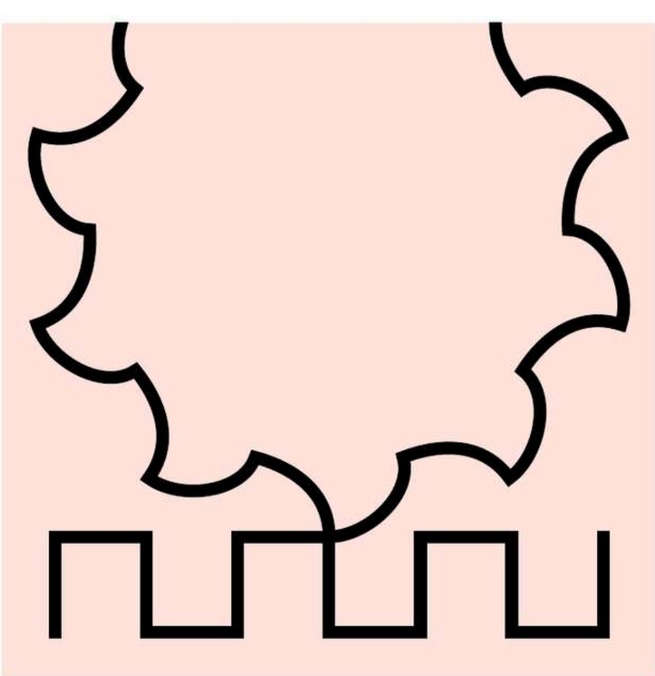


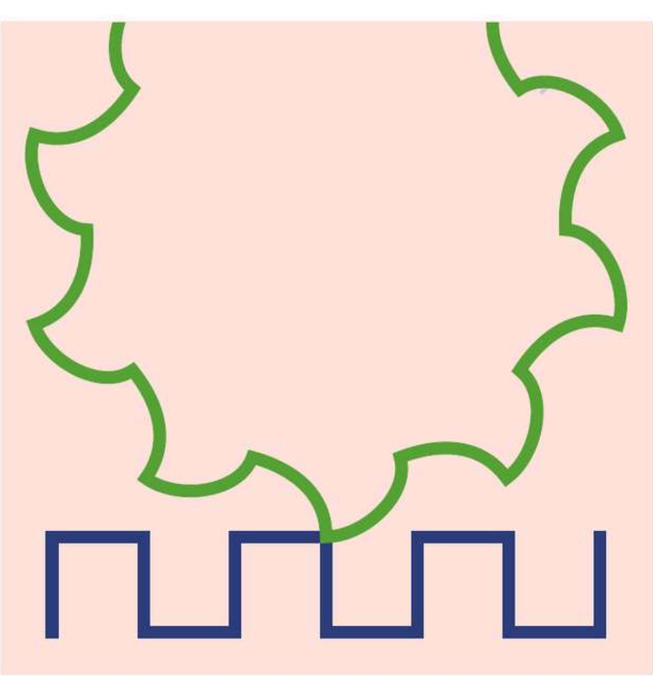
Perceptual decomposition based on the principle

Alternative decomposition violating the principle

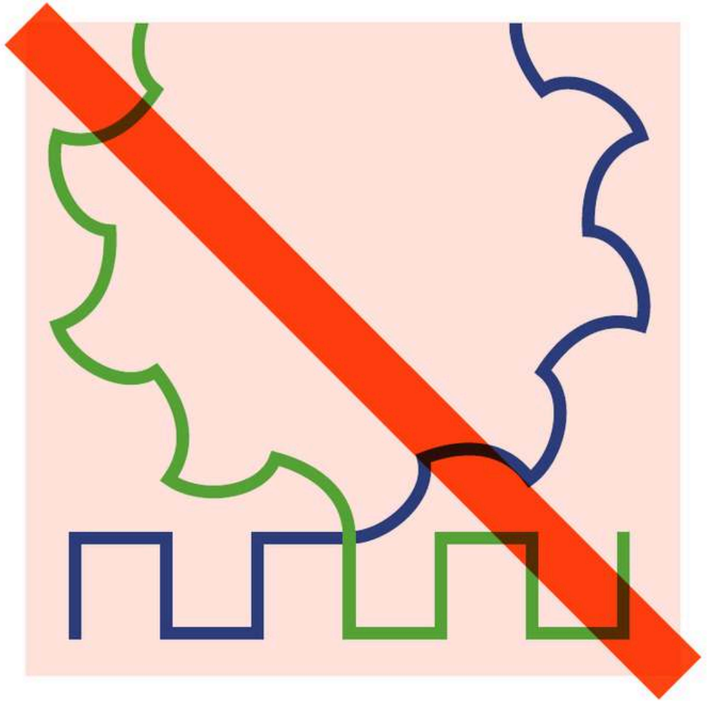
©
WWF®

Law of good Gestalt



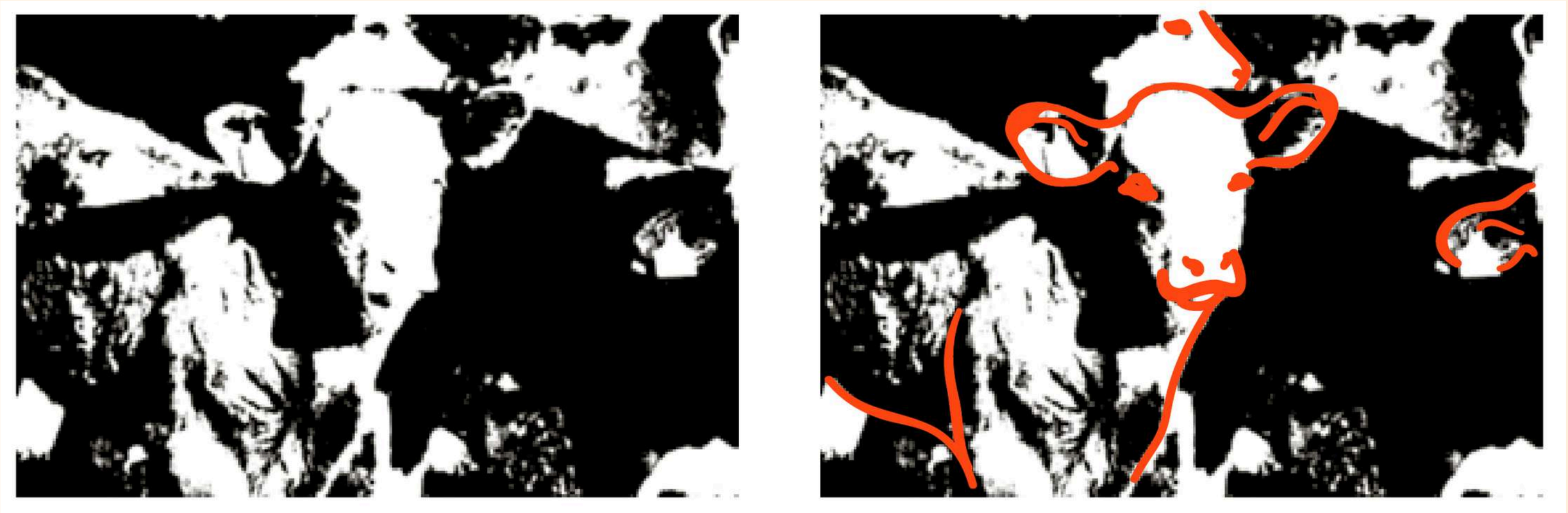


Perceptual decomposition based on the principle



Alternative decomposition violating the principle

Past experience principle



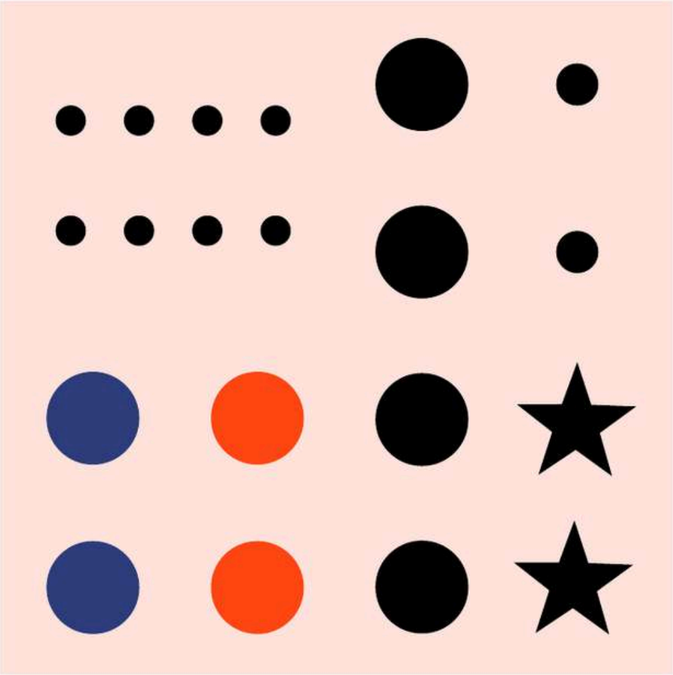
S. Kundu Ph.D.

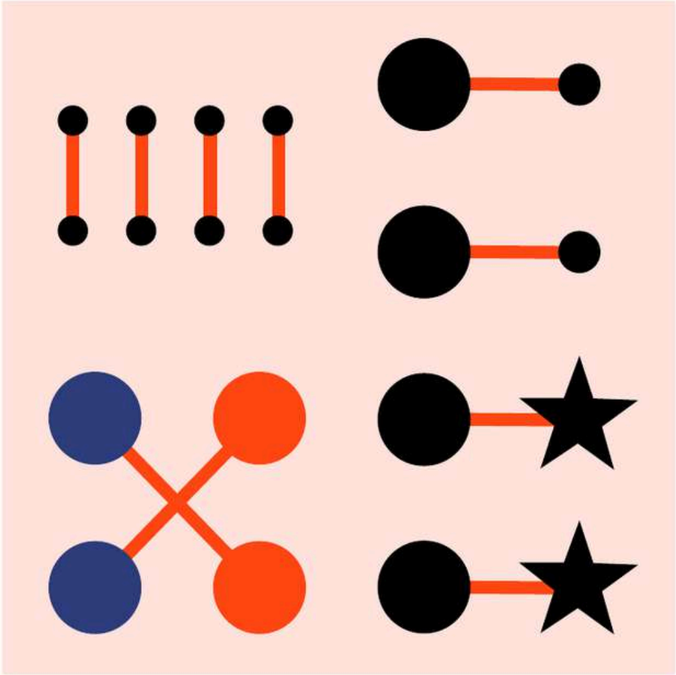
Join at

slido.com

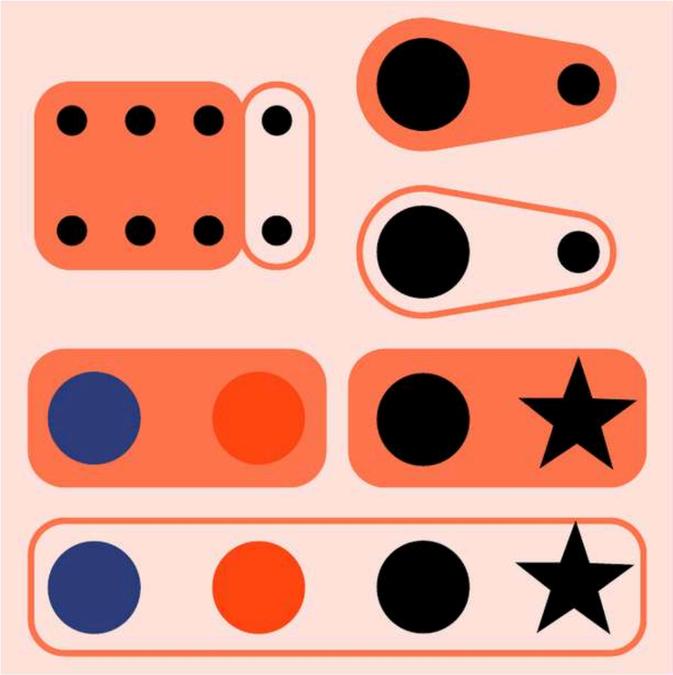
#3520

Connectedness and connecting regions or contours





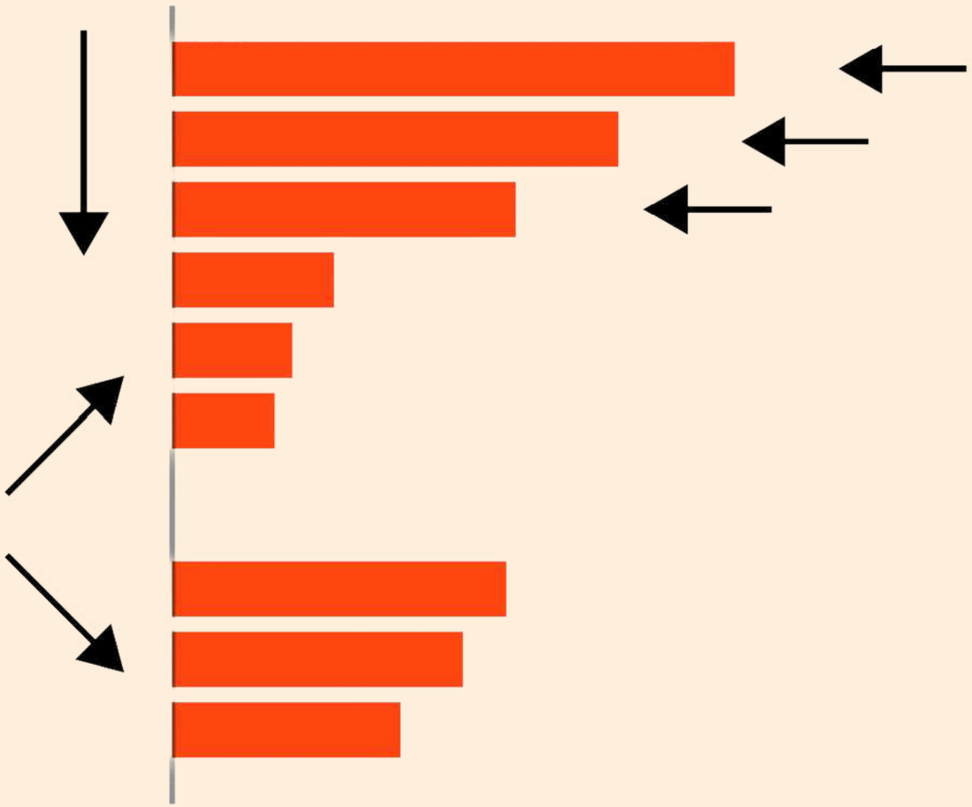
Connectedness



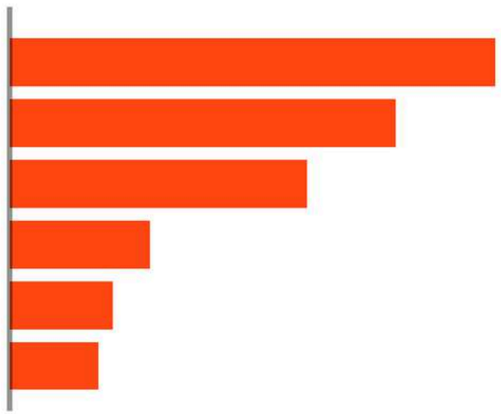
Connecting region or contour

Order
Position encodes ordinal data: the ordering of the bars

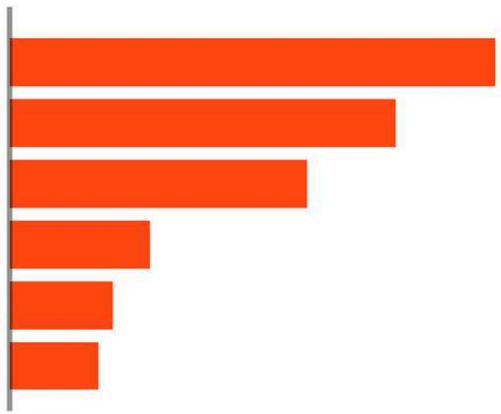
Category
Position encodes qualitative data: two visually separate groups of bars



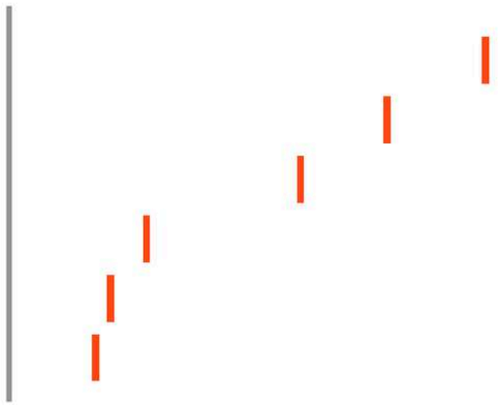
Numerical values
Position encodes numerical data: the locations of the ends of the bars



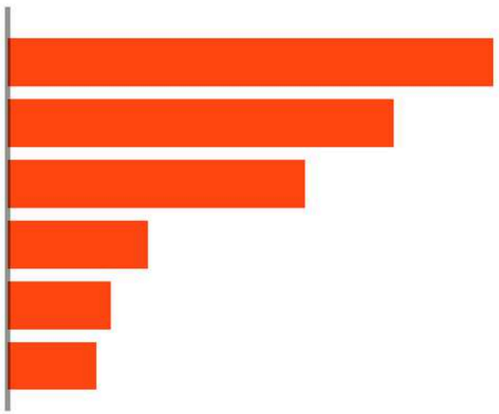
Position, length and area



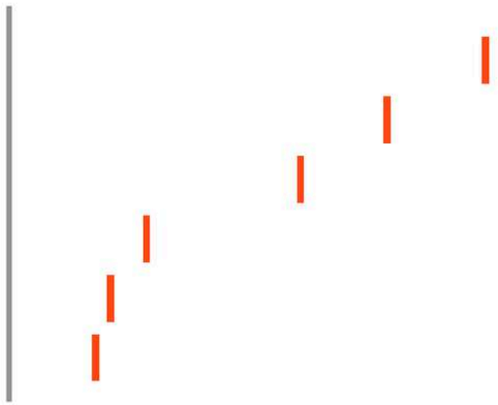
Position, length and area



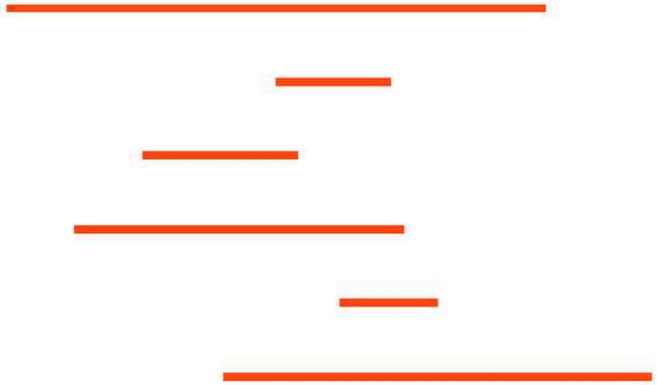
Position only



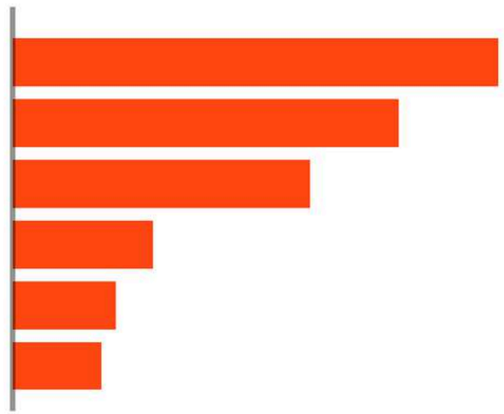
Position, length and area



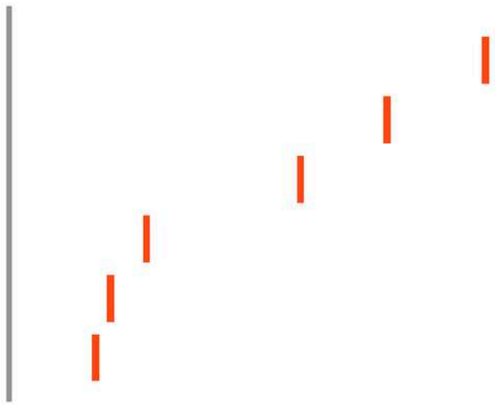
Position only



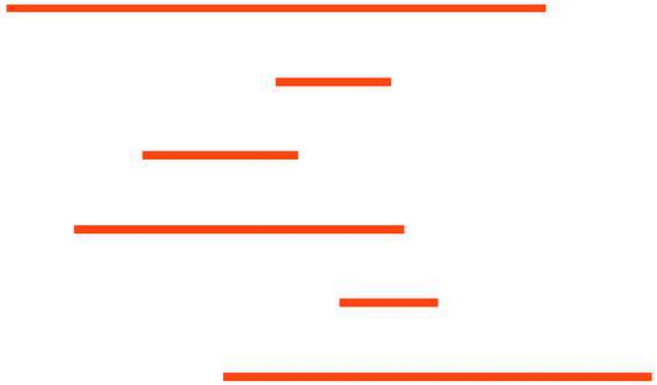
Length only



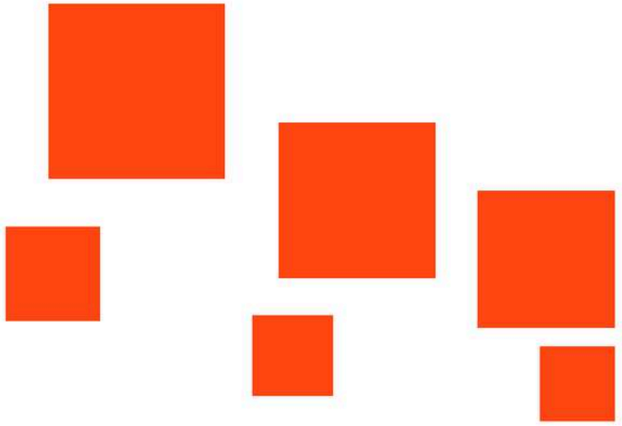
Position, length and area



Position only



Length only



Area only

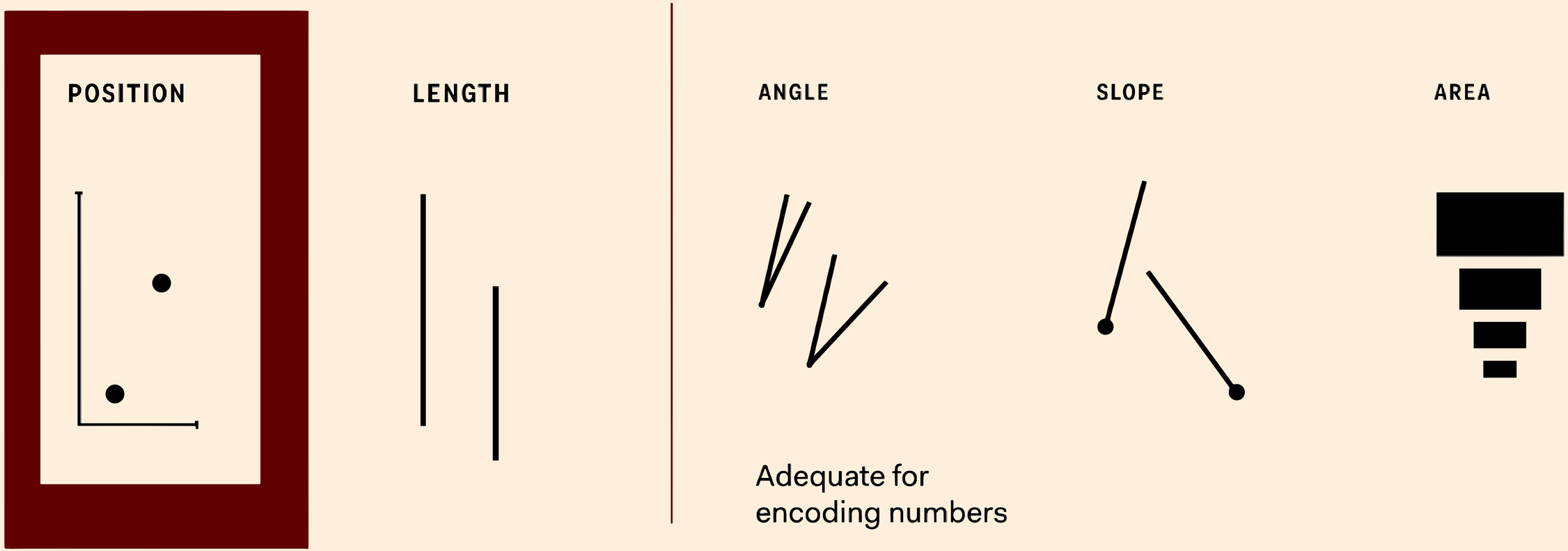
Join at Slido.com

#14837

Numbers (ratio or interval scale)

most precise →

→ least precise

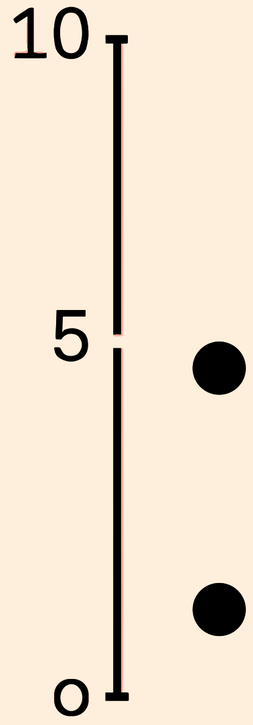


Adequate for
encoding numbers

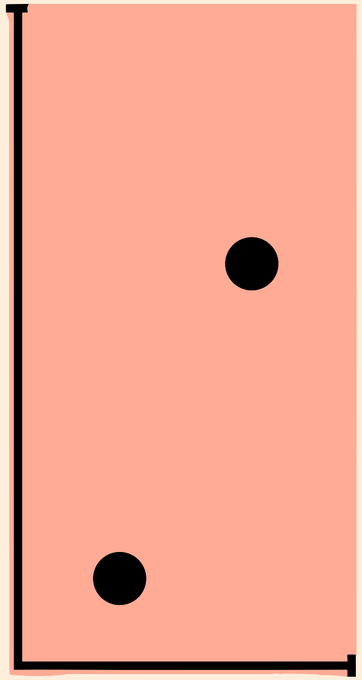
Using Position for Encoding Numbers

most precise →

Position along common explicit scale

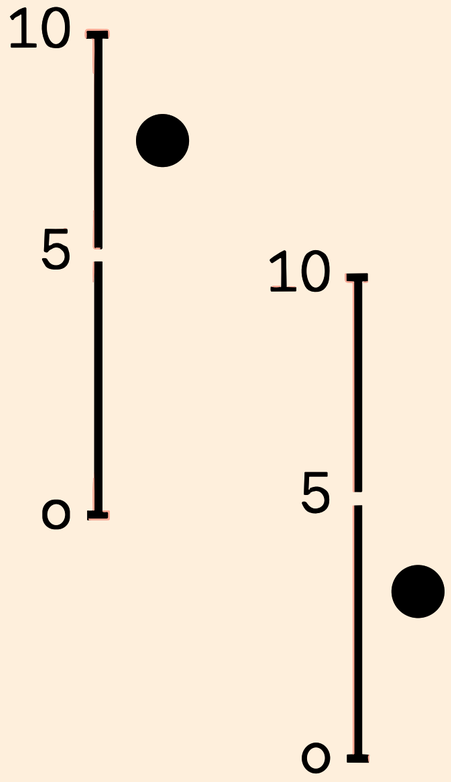


Position along common implicit scale

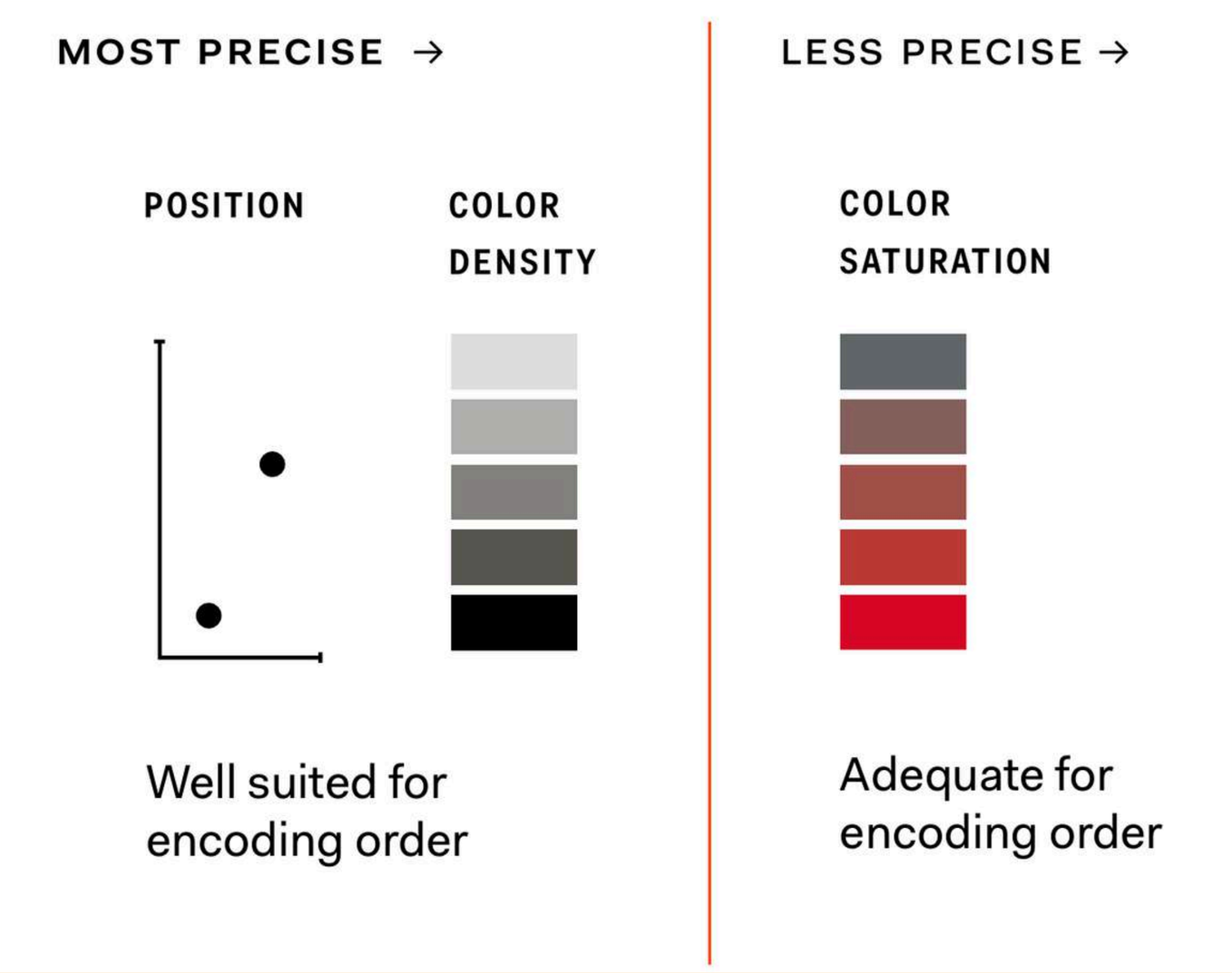


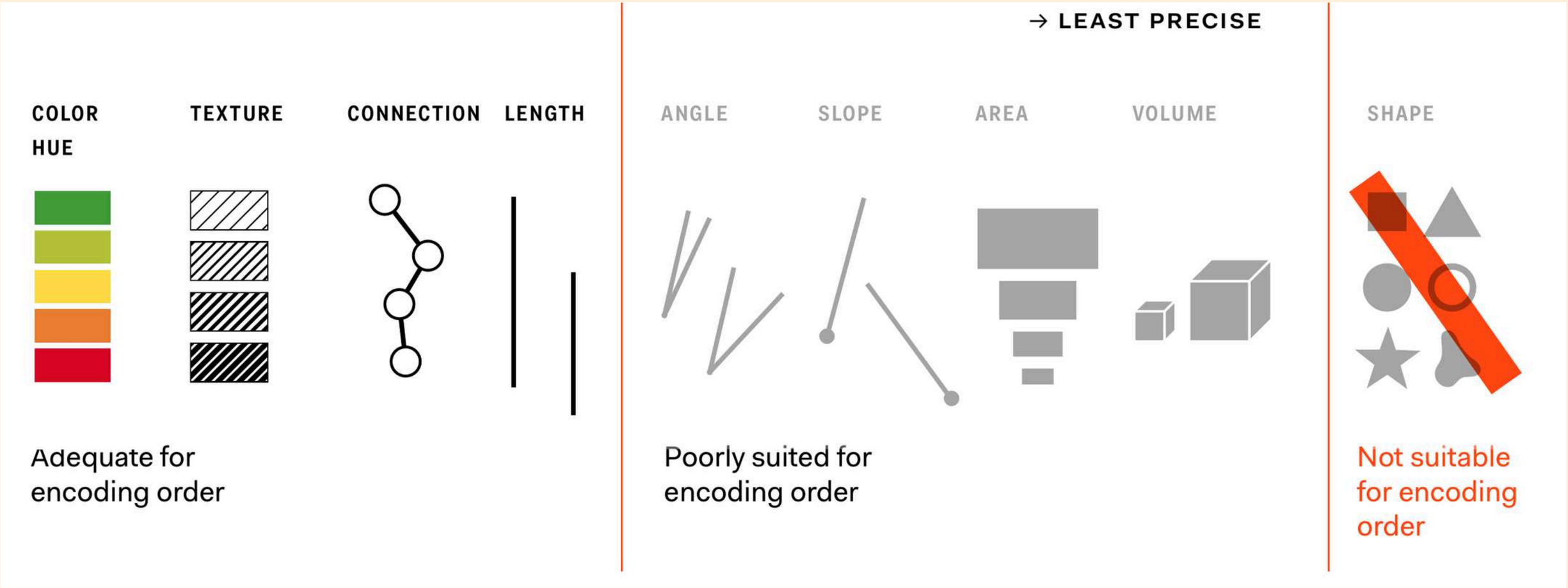
→ least precise

Position along non-aligned, but linked scales

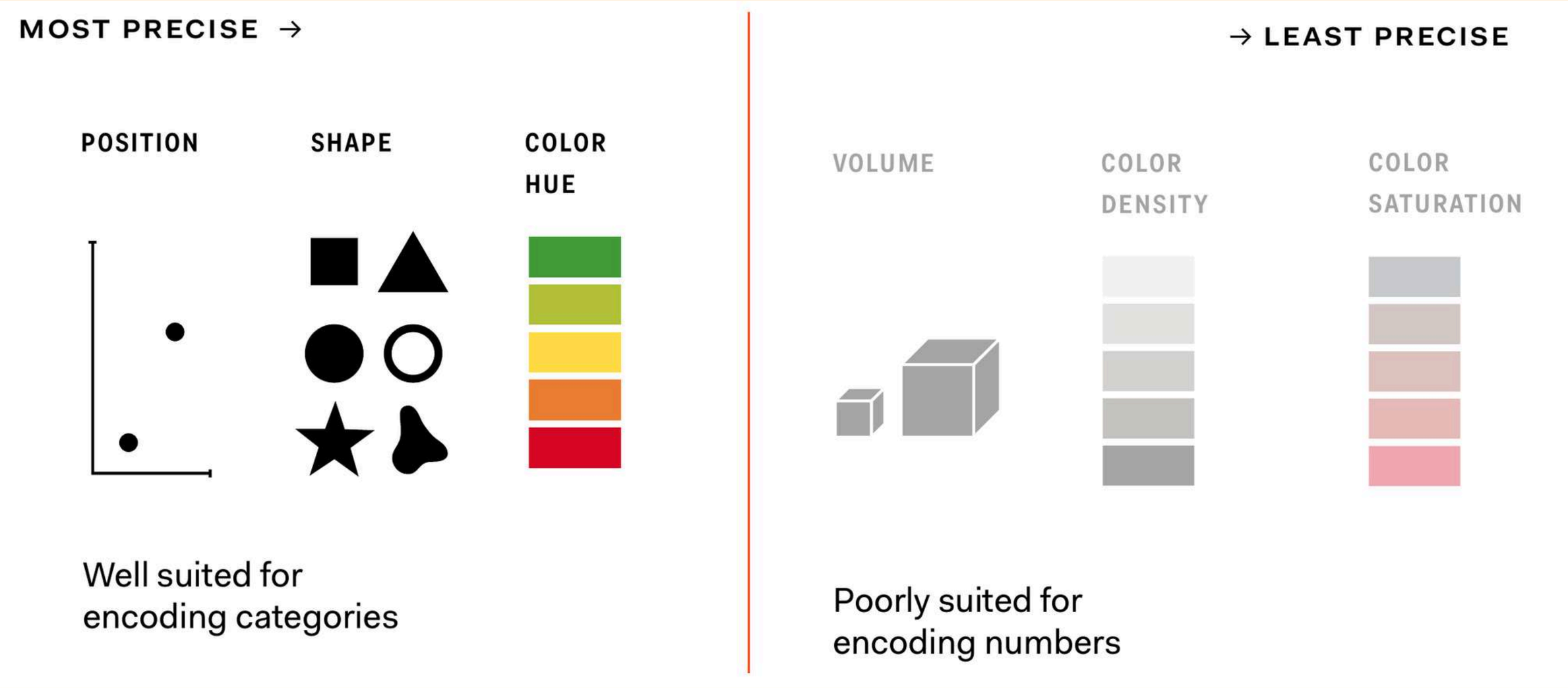


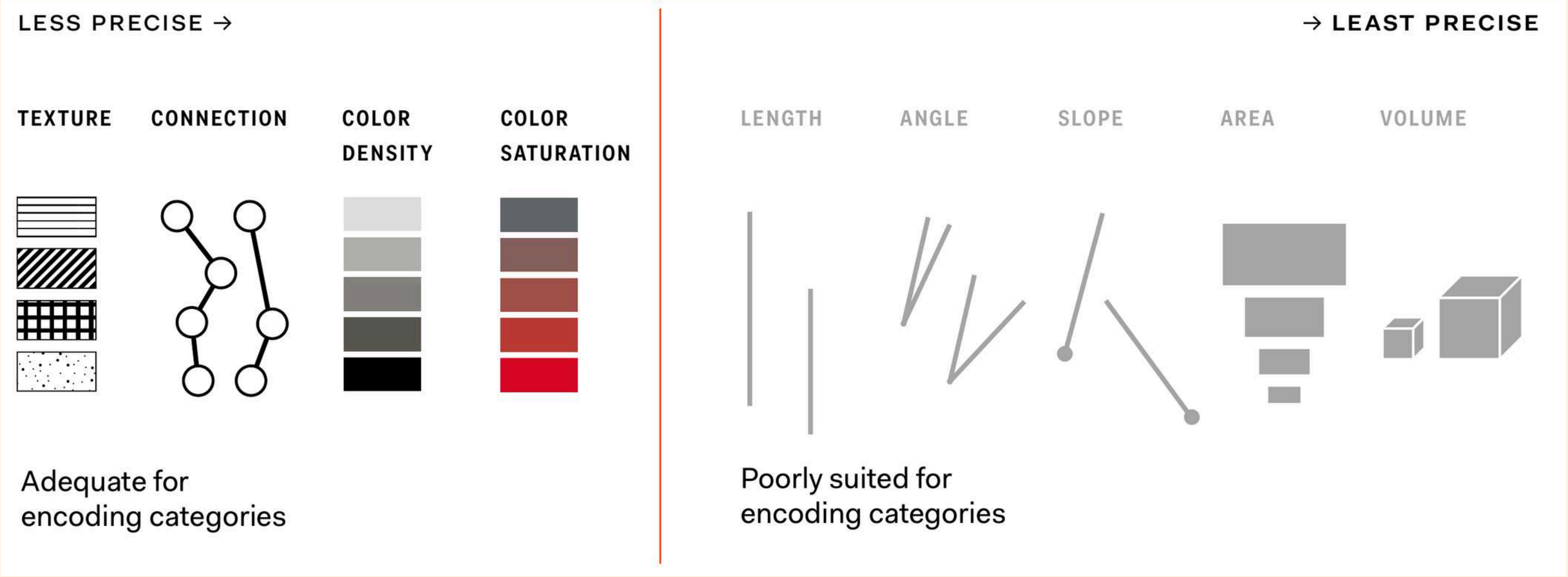
Order (ordinal scale)

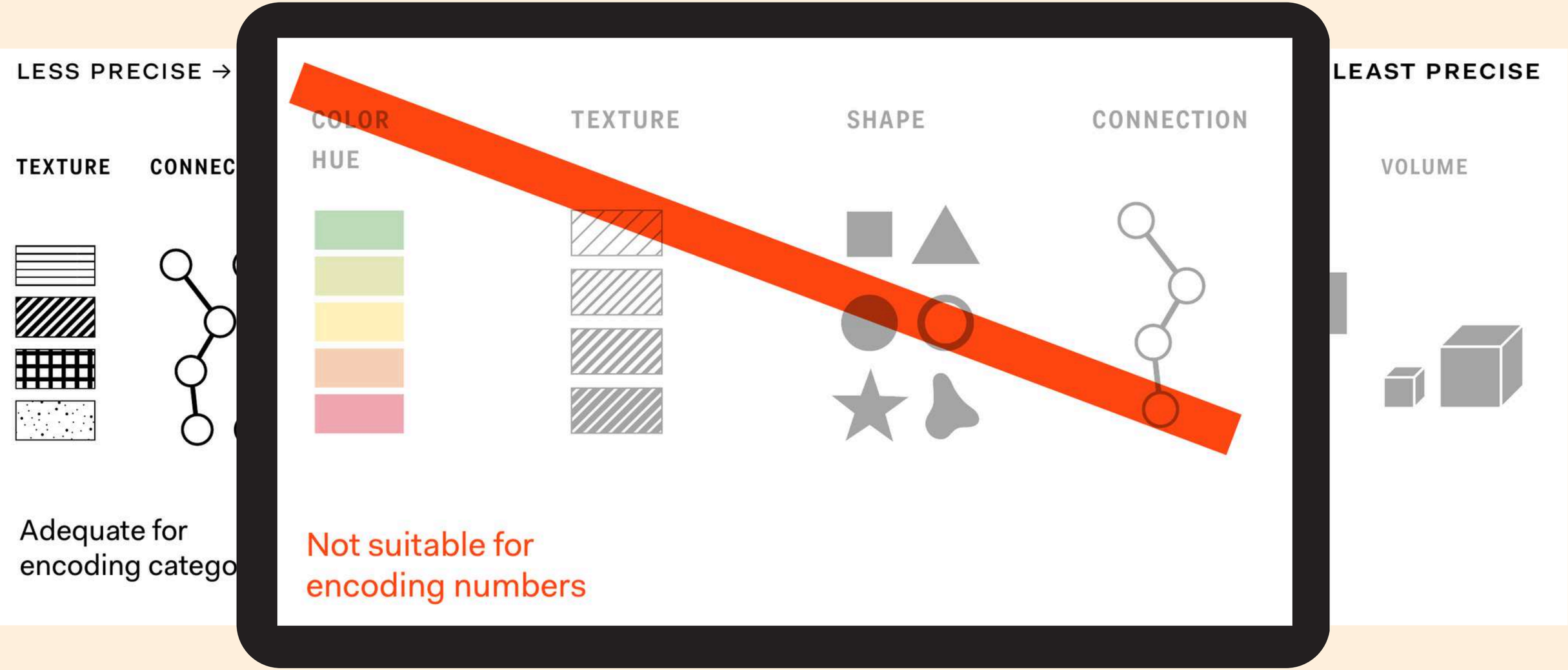



→ LEAST PRECISE
COLOR HUE
TEXTURE
CONNECTION
LENGTH
ANGLE
SLOPE
AREA
VOLUME
SHAPE
Adequate for encoding order
Poorly suited for encoding order
Not suitable for encoding order

Categories (nominal scale)




LESS PRECISE →
→ LEAST PRECISE
TEXTURE
CONNECTION
COLOR DENSITY
COLOR SATURATION
LENGTH
ANGLE
SLOPE
AREA
VOLUME
Adequate for encoding categories
Poorly suited for encoding categories

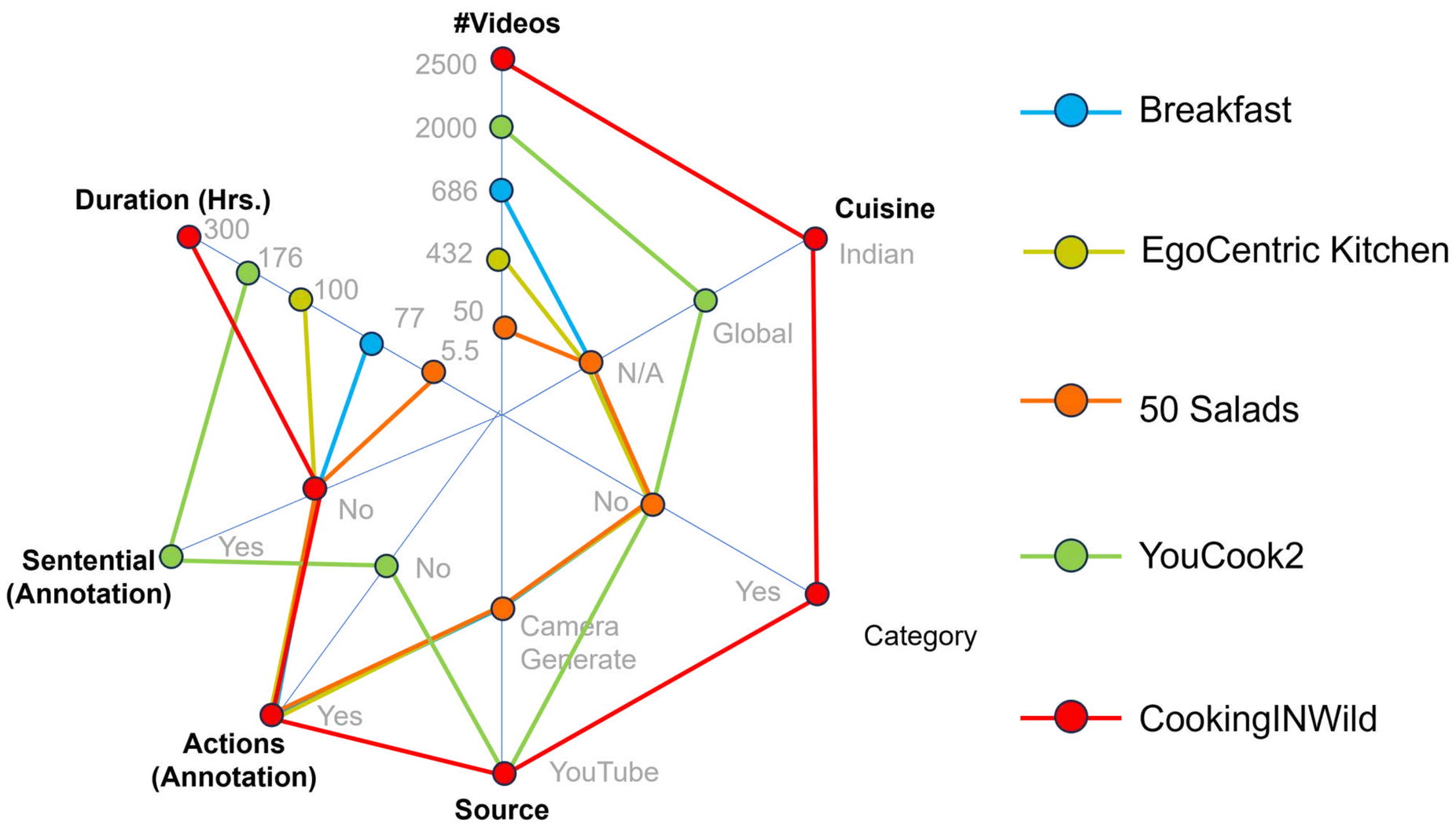

LESS PRECISE →
LEAST PRECISE
TEXTURE
CONNEC
COLOR
HUE
TEXTURE
SHAPE
CONNECTION
VOLUME
Adequate for
encoding catego
Not suitable for
encoding numbers

S. Kundu Ph.D.

Join at

slido.com

#14837



Summary

1. Seeing is an active process that largely takes place in the brain.
2. Visual perception is influenced by external visual stimuli, prior experience, and active focusing of attention.
3. The basic, pre-attentive characteristics of visual perception are called tunable features.
    a. These include shape and color features, and movement.
4. Gestalt laws are principles by which we perceive which individual features belong in the same group or entity.

1. The means by which numbers and other data can be turned into visual form are called visual variables.
    a. They have an inherent order of superiority arising from how human visual perception works.
    b. The best visual variables with which to encode quantitative data are position and length.
    c. Order is best encoded with position and color density (lightness).
    d. Categories are best encoded with position, shape, and color hue.

Join at Slido.com

#3748

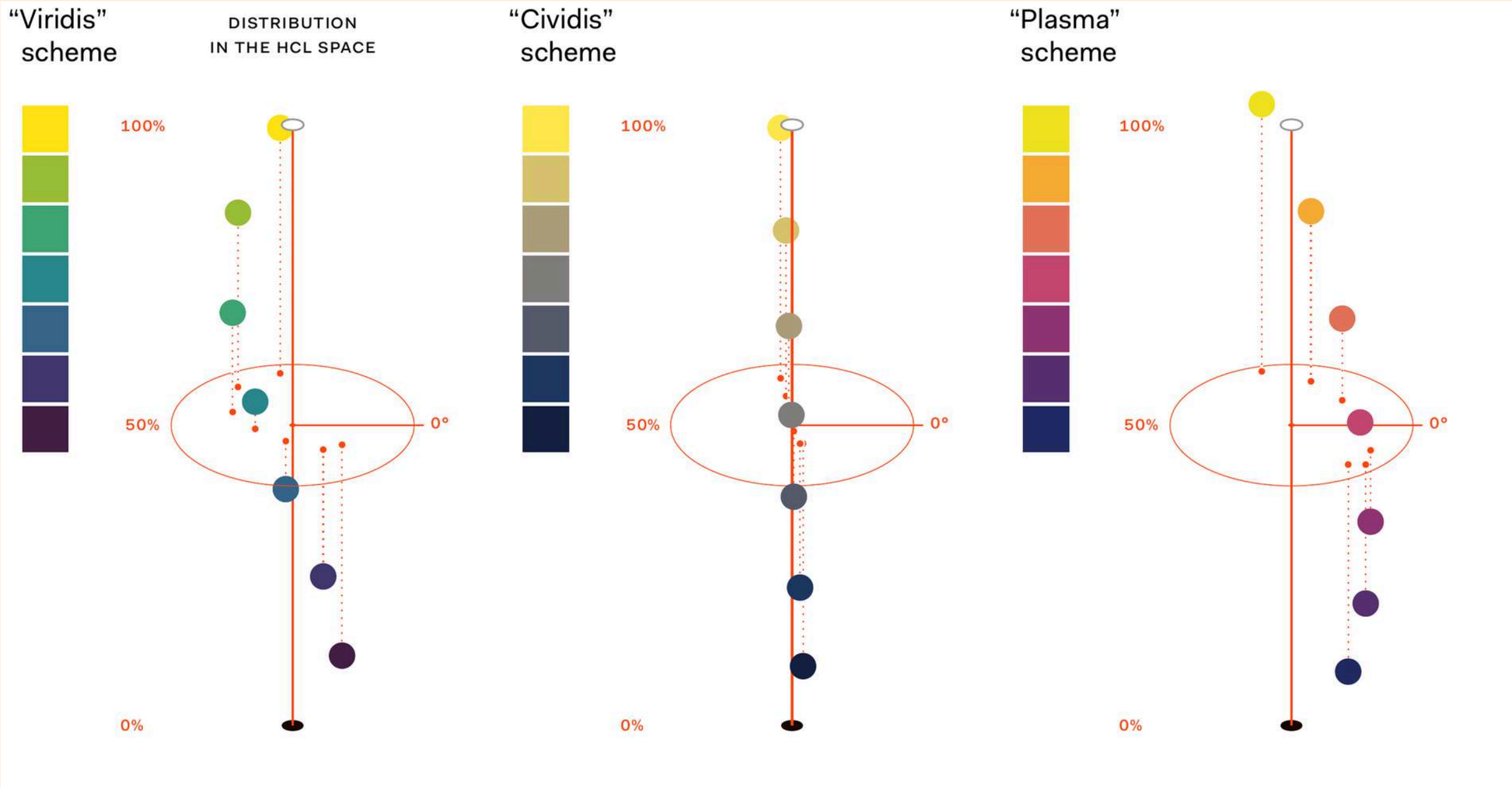
Elements of Color Perception

H - Hue

S/C - Saturation/Chroma

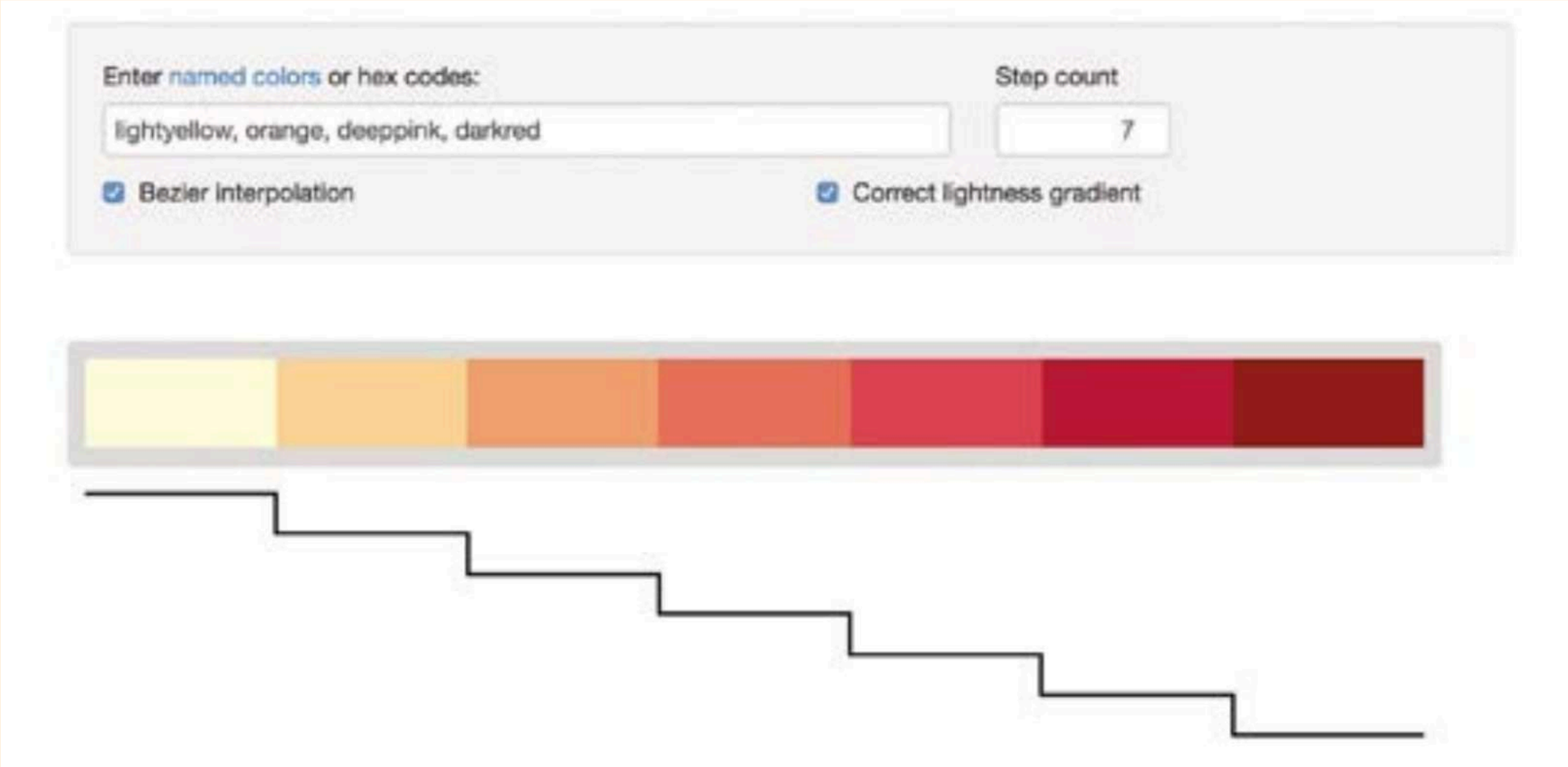
L - Lightness

Perceptually Uniform Color Schemes

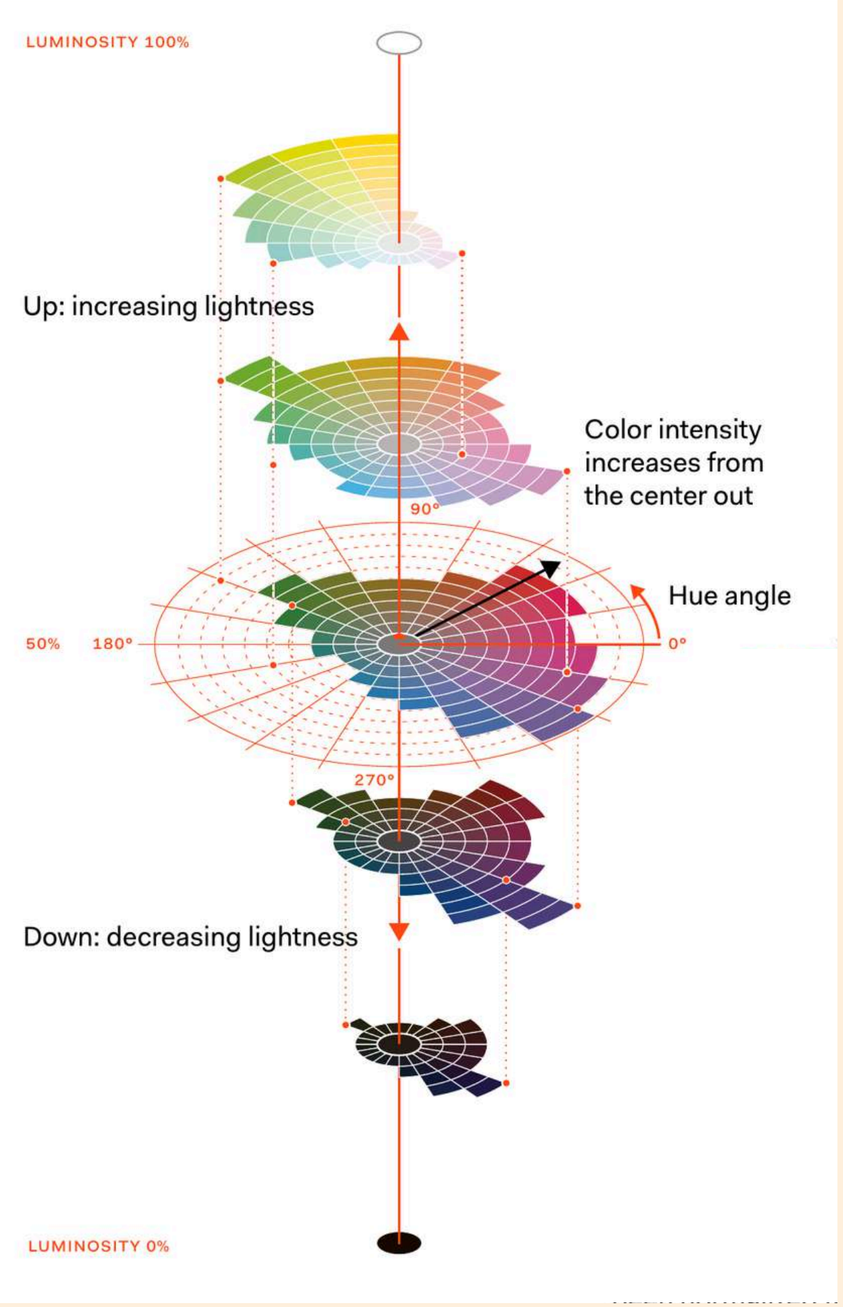


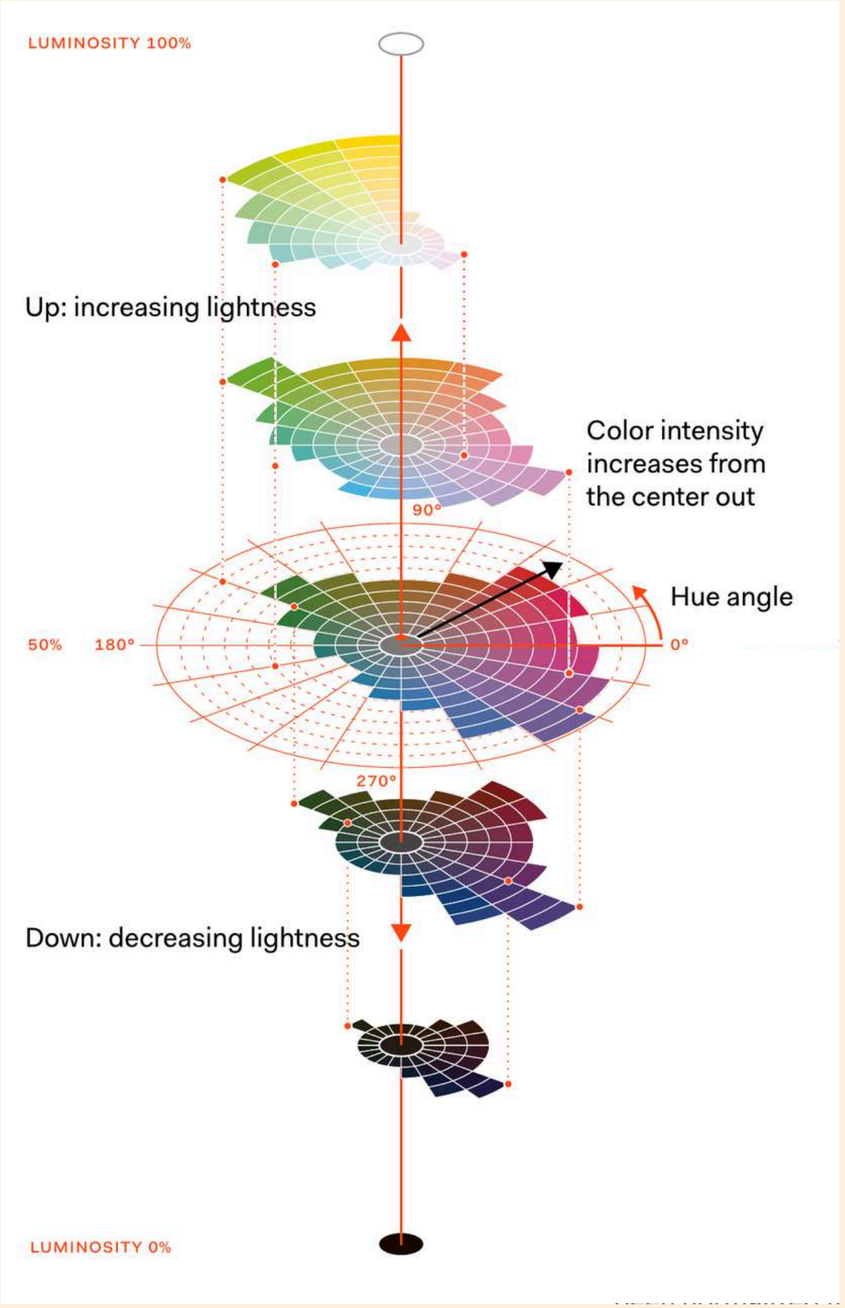
Numerically expressed distance between colors corresponds to their perceived visual difference

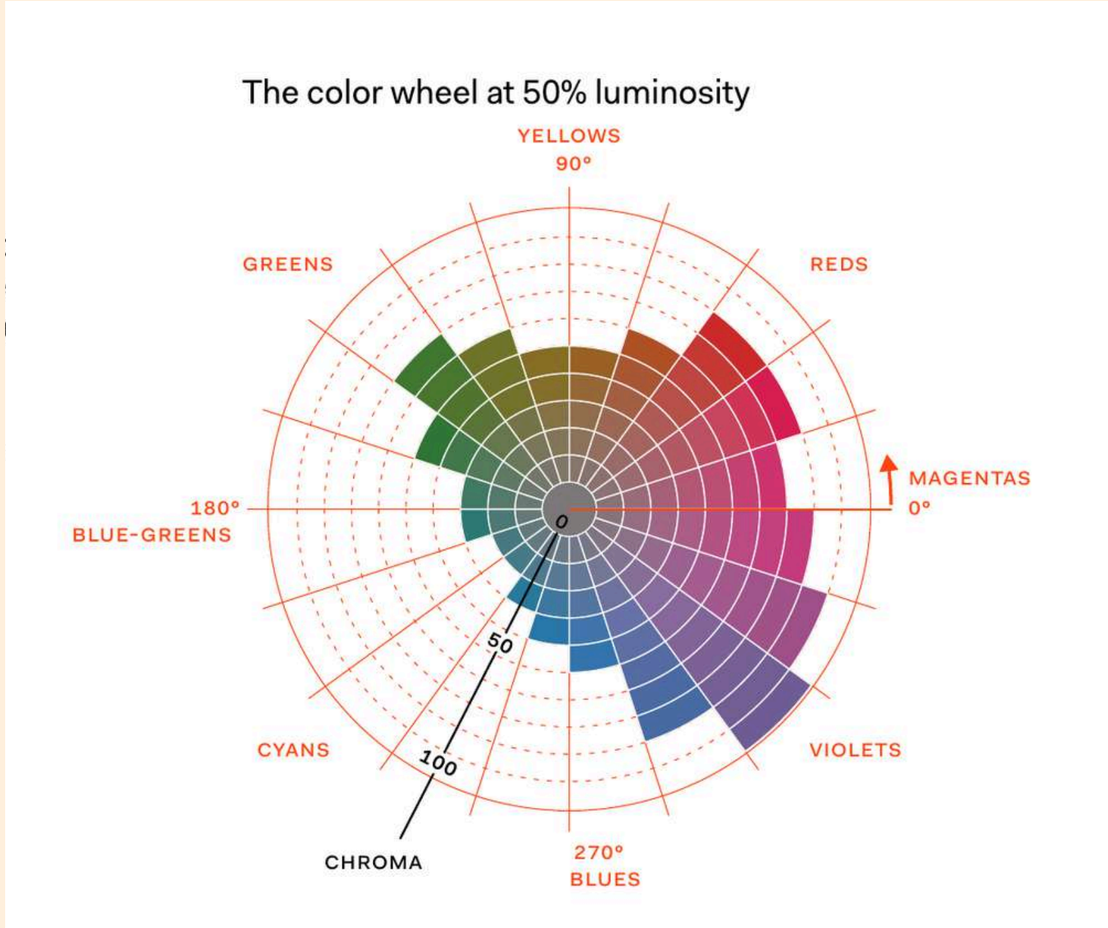
http://gka.github.io/palettes/

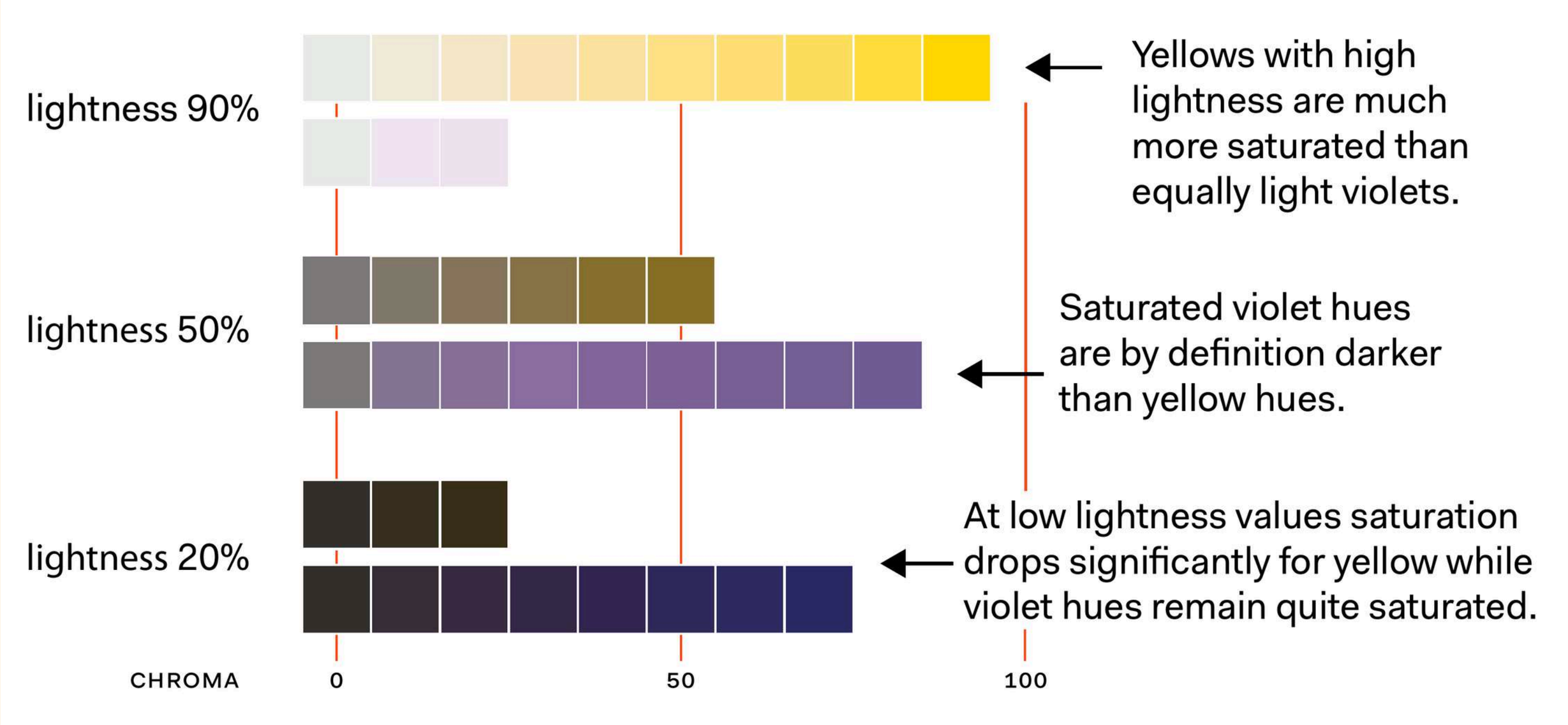


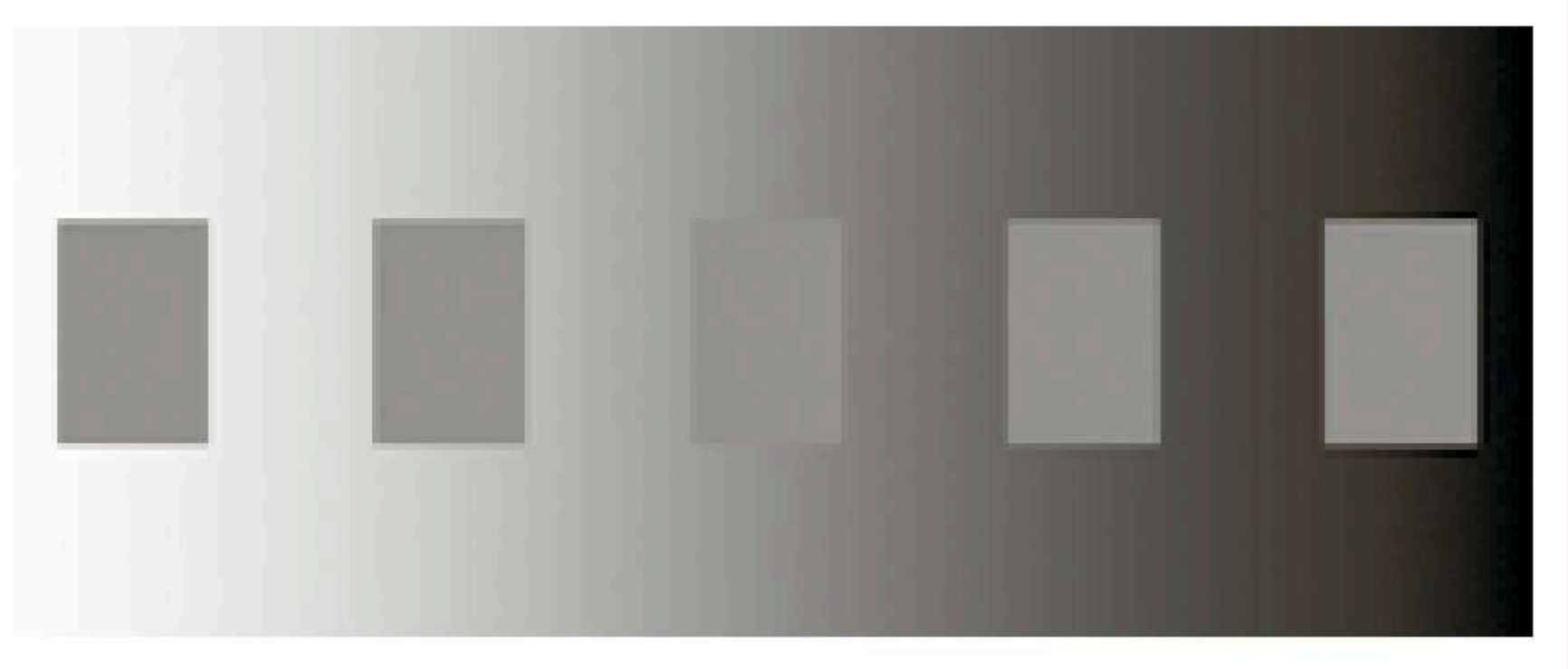
create perceptually uniform color scales with different parameters

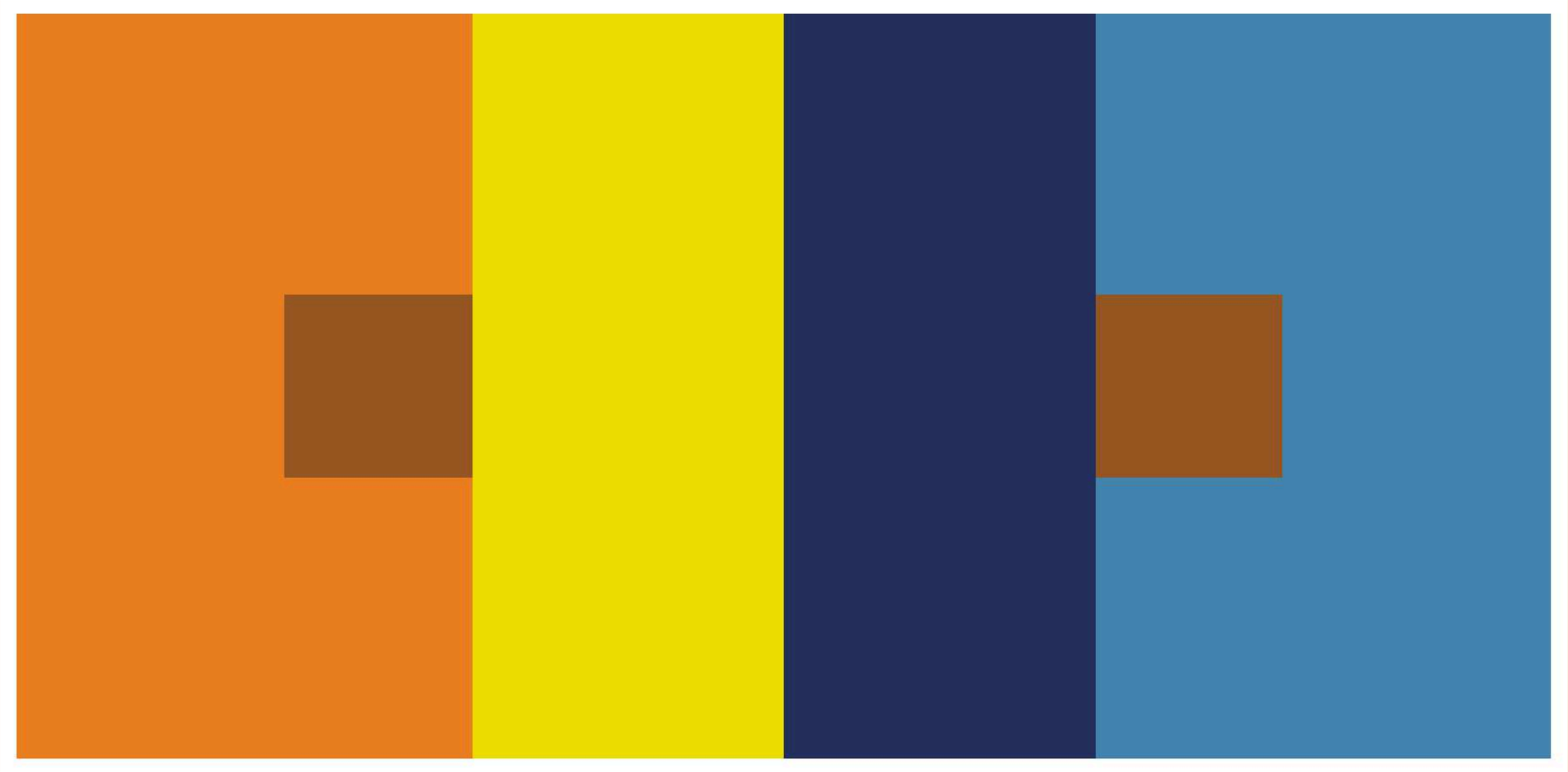

LUMINOSITY 100%
Up: increasing lightness
Color intensity increases from the center out
90°
Hue angle
50%
180°
0°
270°
Down: decreasing lightness
LUMINOSITY 0%

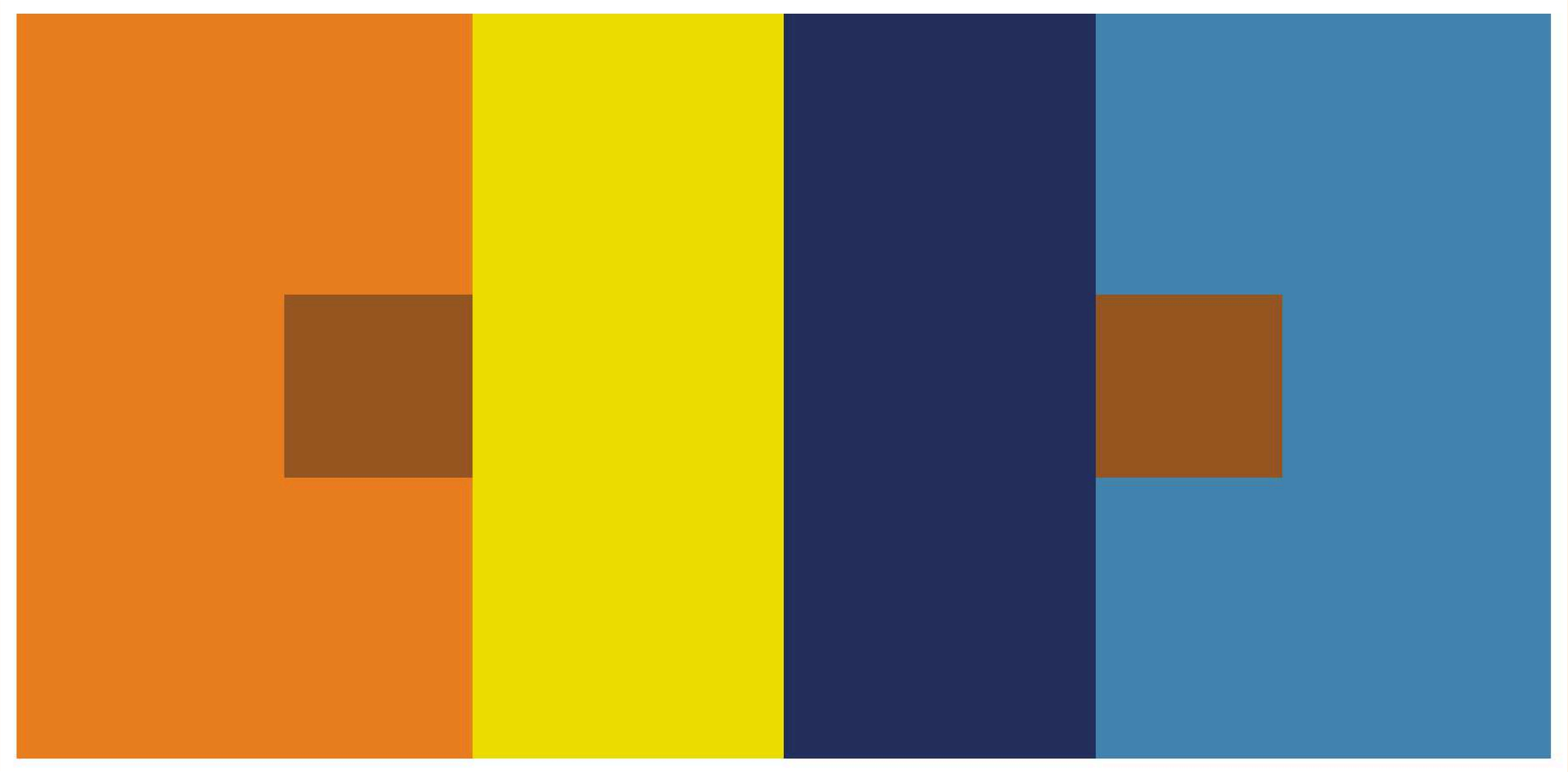

LUMINOSITY 100%
Up: increasing lightness
Color intensity increases from the center out
90°
Hue angle
50%
180°
0°
270°
Down: decreasing lightness
LUMINOSITY 0%


The color wheel at 50% luminosity
YELLOWS
90°
GREENS
REDS
MAGENTAS
0°
180°
BLUE-GREENS
0
50
100
CYANS
VIOLETS
CHROMA
270°
BLUES

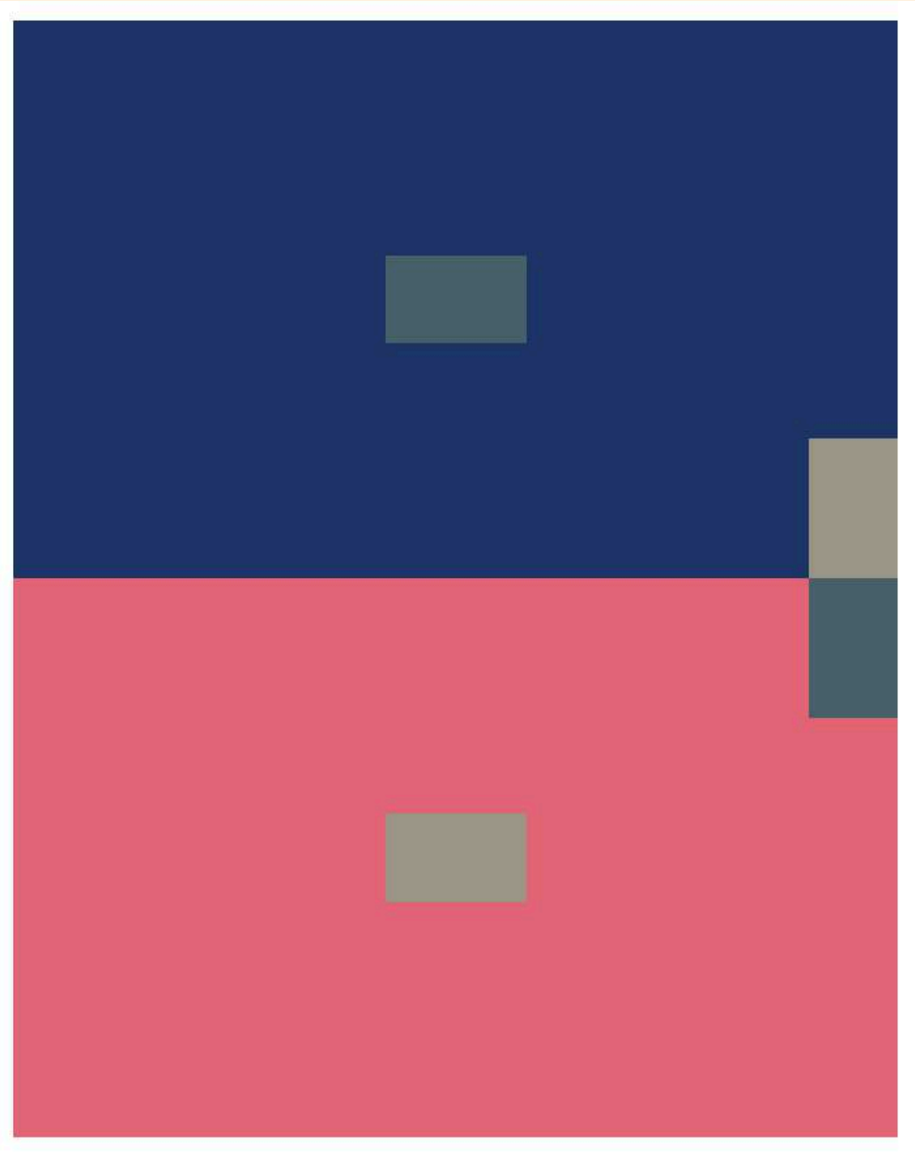

lightness 90%
lightness 50%
lightness 20%
CHROMA
0
50
100
Yellows with high lightness are much more saturated than equally light violets.
Saturated violet hues are by definition darker than yellow hues.
At low lightness values saturation drops significantly for yellow while violet hues remain quite saturated.

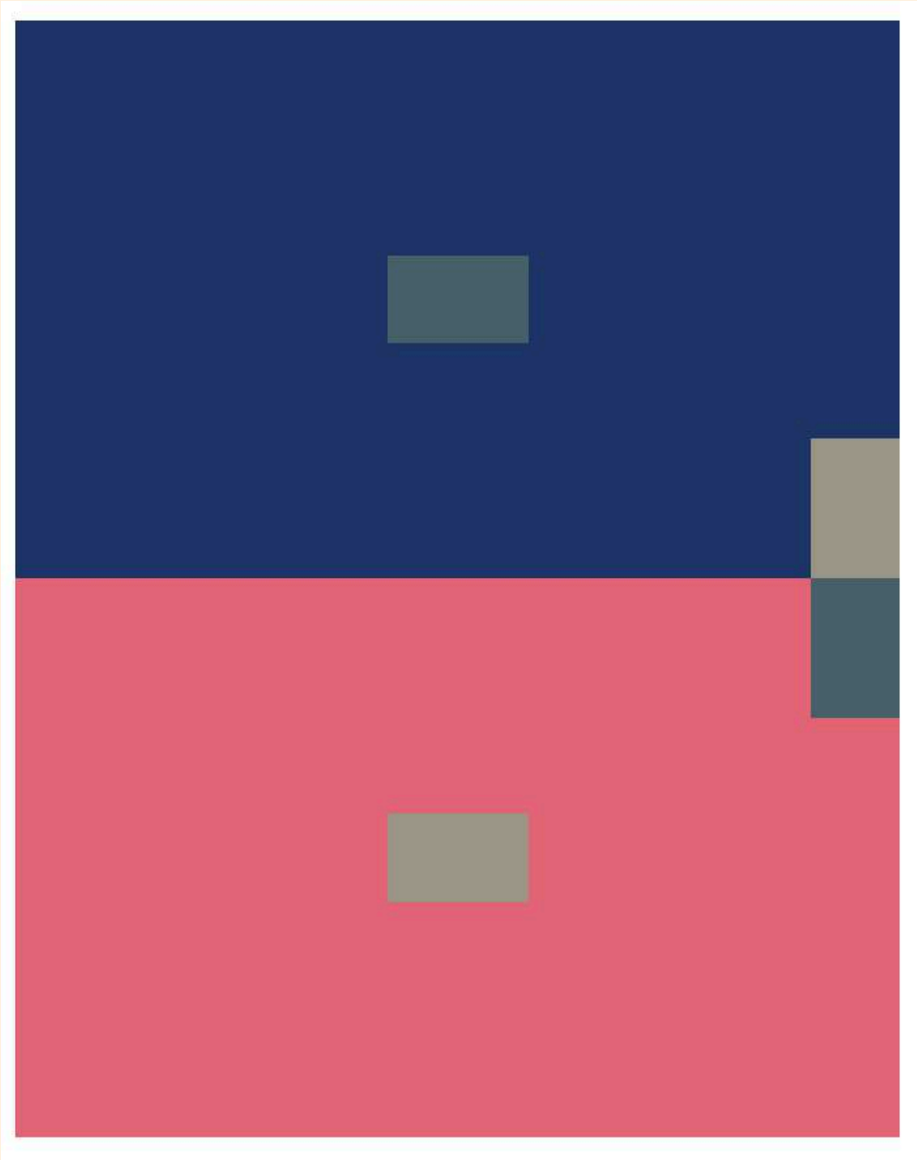






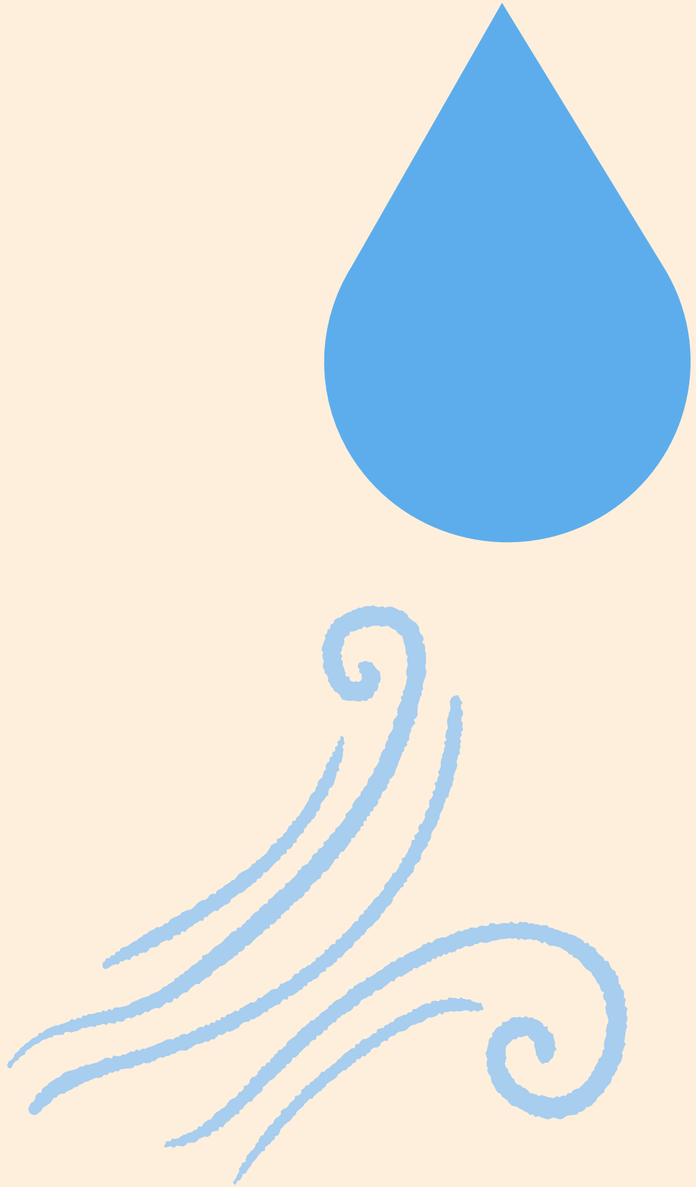
Both of the small rectangles are of the same color, but, due to simultaneous Here the small rectangles in the contrast, appear to differ in hue and lightness.

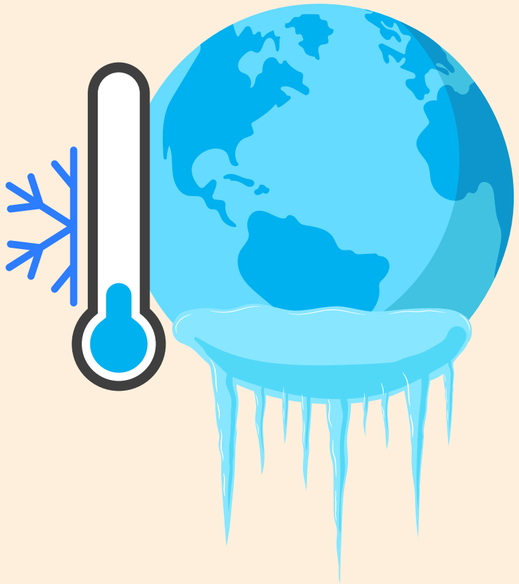


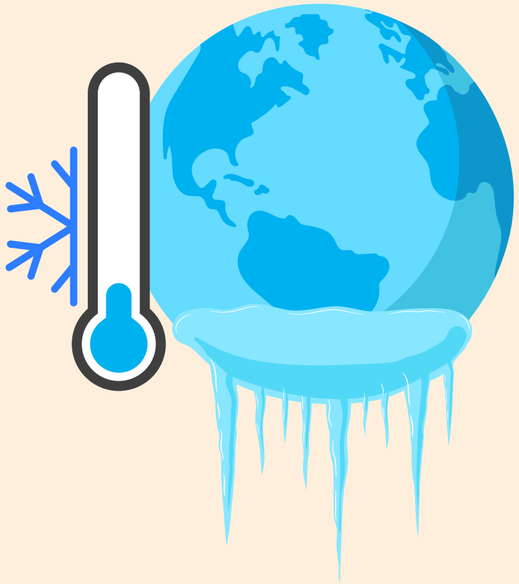


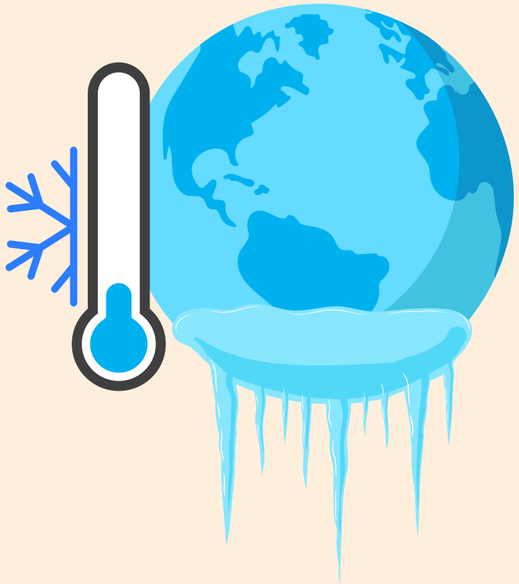
Here the small rectangles in the middle are of different color, but look similar because of color interaction with the surrounding large area of color.

China
Afghanistan
Pakistan
New Delhi
नई दिल्ली
Nepal
Bhutan
Bangladesh
India
Kolkata
কলকাতা
Myanmar
(Burma)
Mumbai
मुंबई
Laos
Thailand
Vietnam
Bangkok
กรุงเทพมหานคร
Cambodia
Bengaluru
ಬೆಂಗಳೂರು
Bay of Bengal
Arabian Sea
Andaman Sea
Gulf of
Thailand
South
China Se
Ho
香



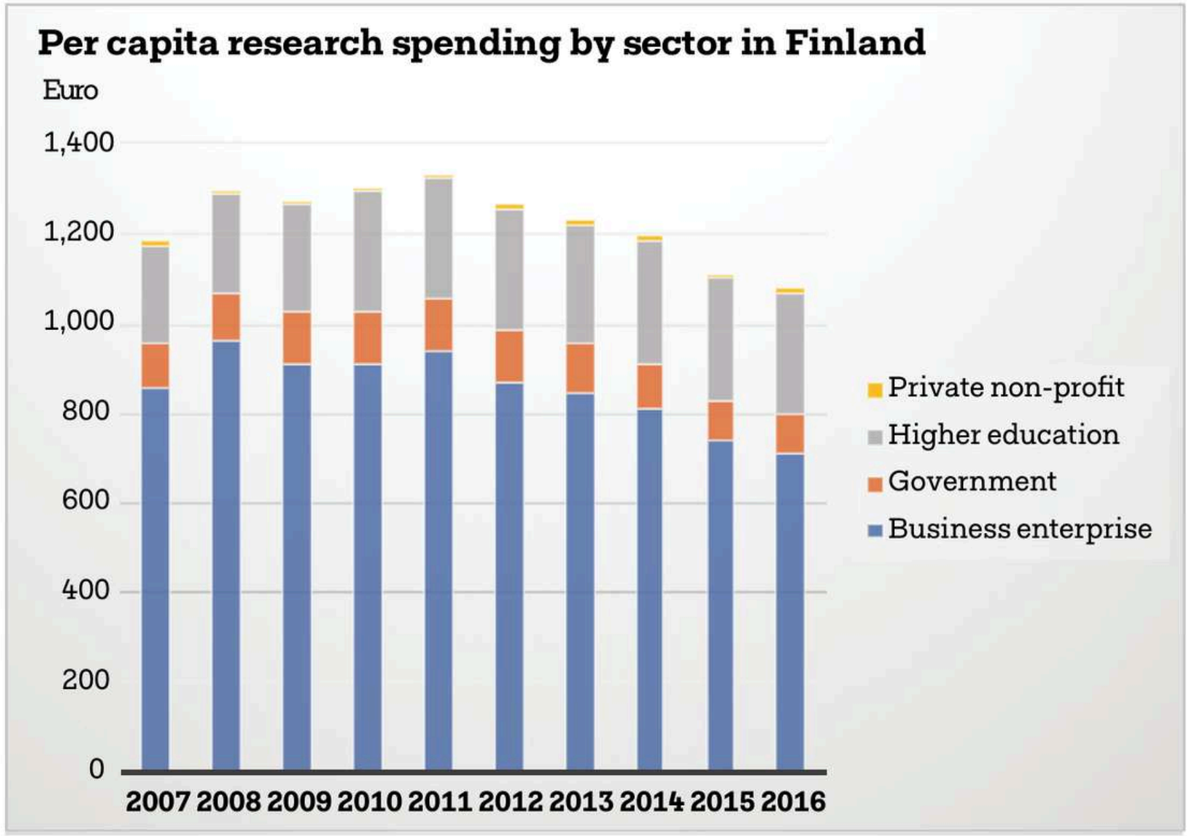




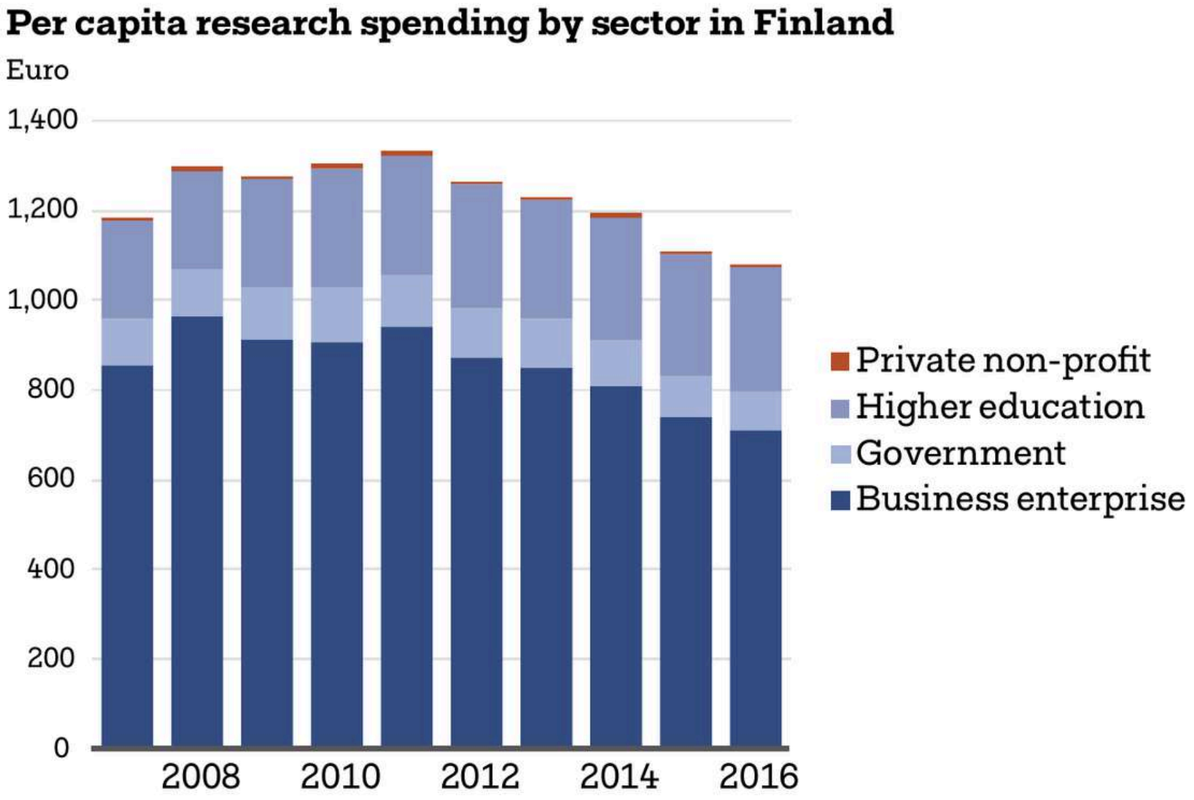


THE MOST EASILY DISTINGUISHABLE COLORS

Color	Approximation of appearance to a person with color vision deficiency	RGB value R 0–255	G 0–255	B 0–255	HCL value H 0–360	C 0–100	L 0–100
Vermilion		213	94	0	55.6	76.7	54.2
Bluish green		0	158	115	164.4	47.5	57.7
Blue		0	114	178	266.2	41.5	46.0
Yellow		240	228	66	100.0	76.6	89.1
Black		0	0	0	–	0.0	0.0
White		255	255	255	–	0.0	100.0
Orange		230	159	0	77.6	76.5	70.6
Sky blue		86	180	233	250.9	36.8	69.8
Reddish purple		204	121	167	344.6	40.0	61.0

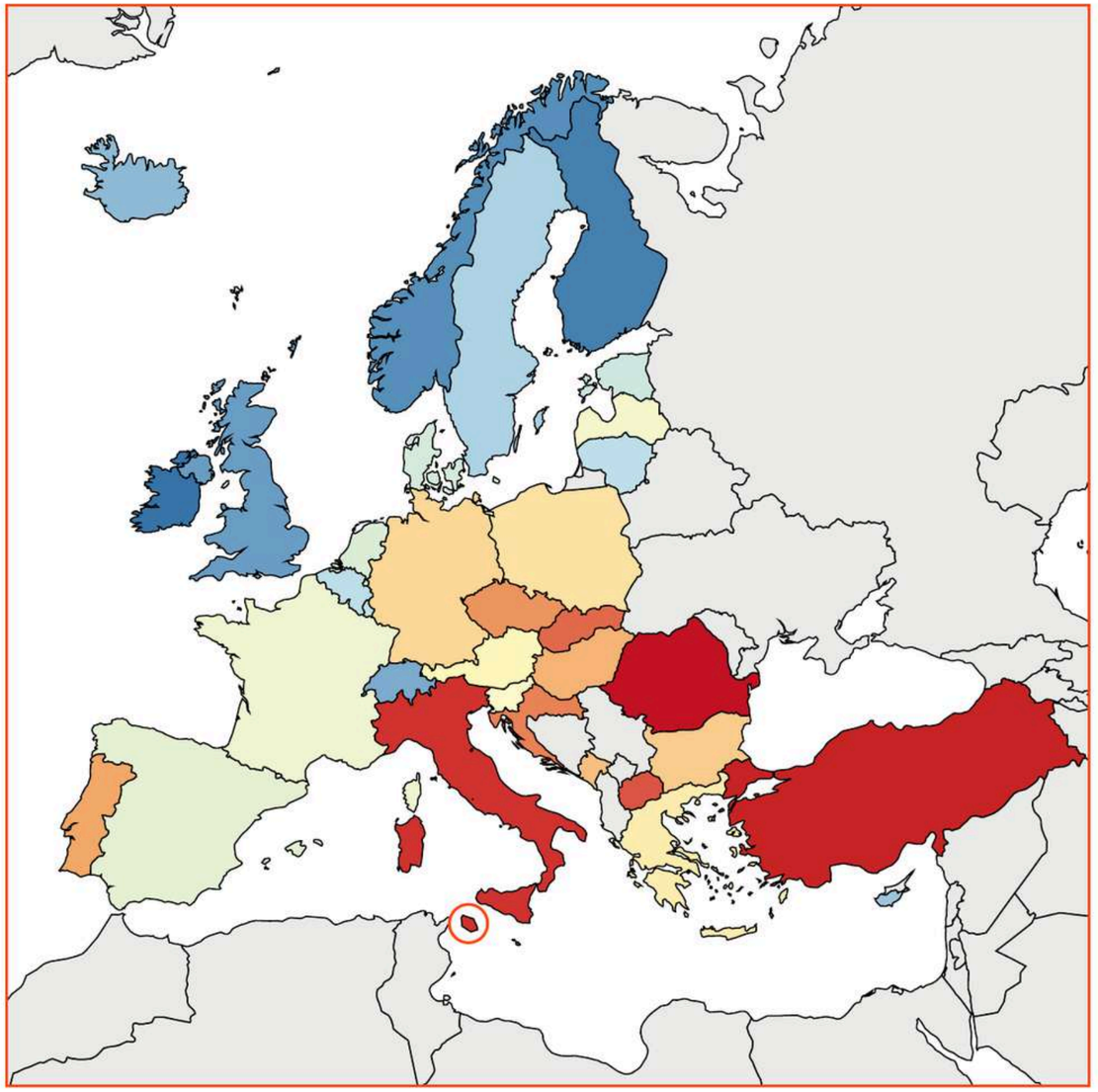


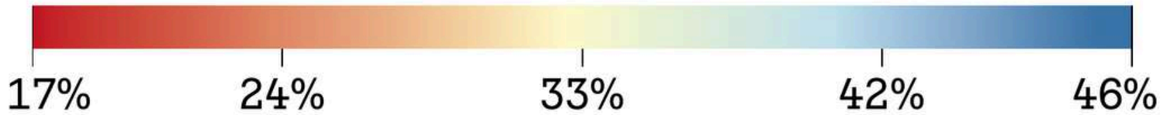
Office 365 default style with default colors



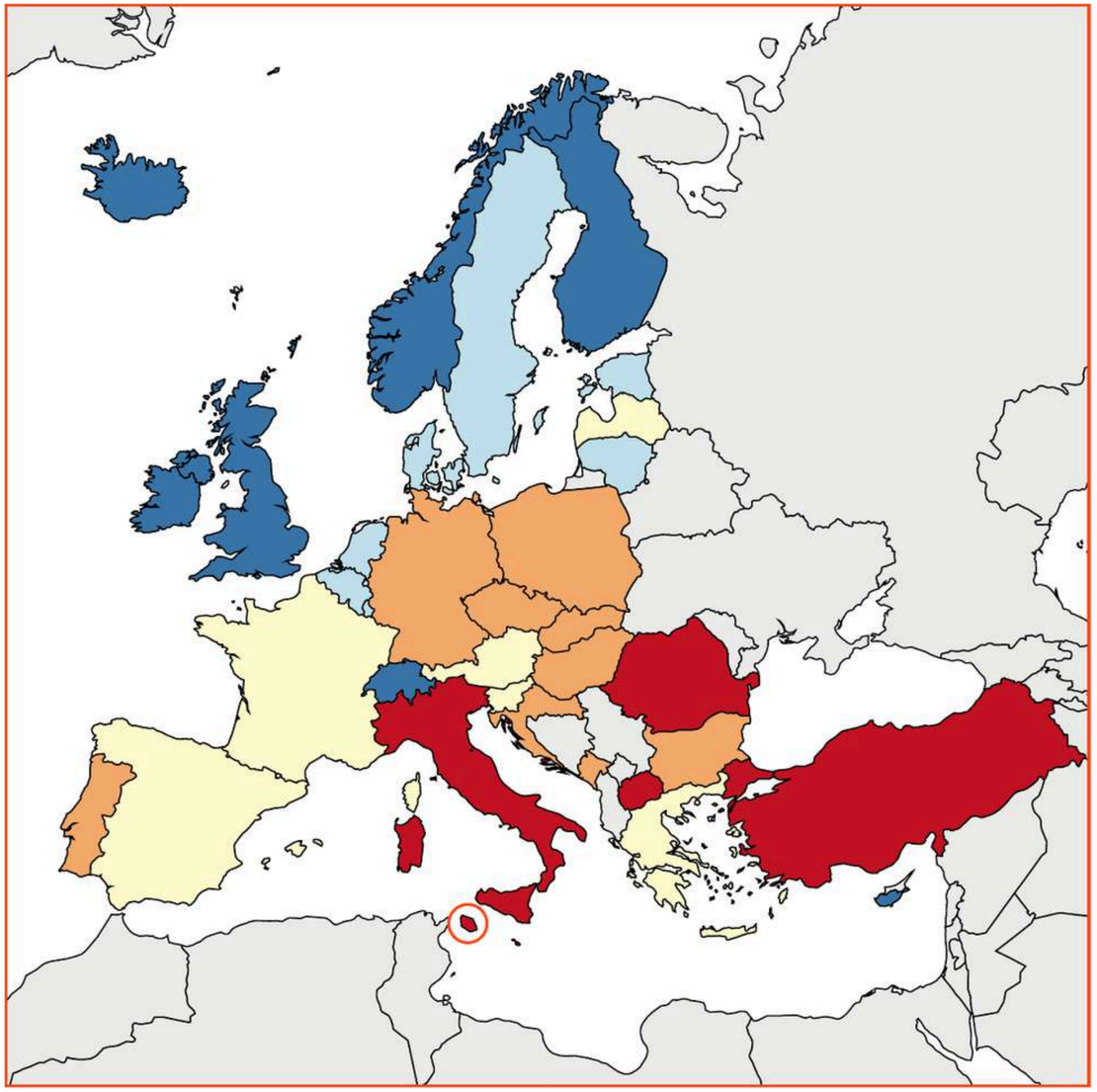
Modified monochrome blue palette

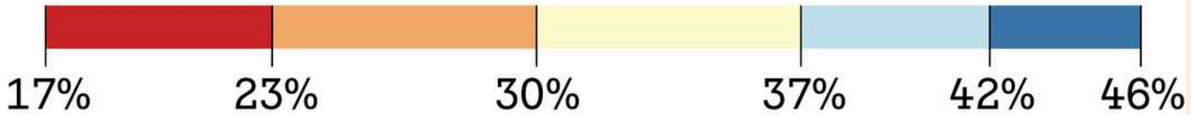
CONTINUOUS SCALE





CLASSIFIED SCALE





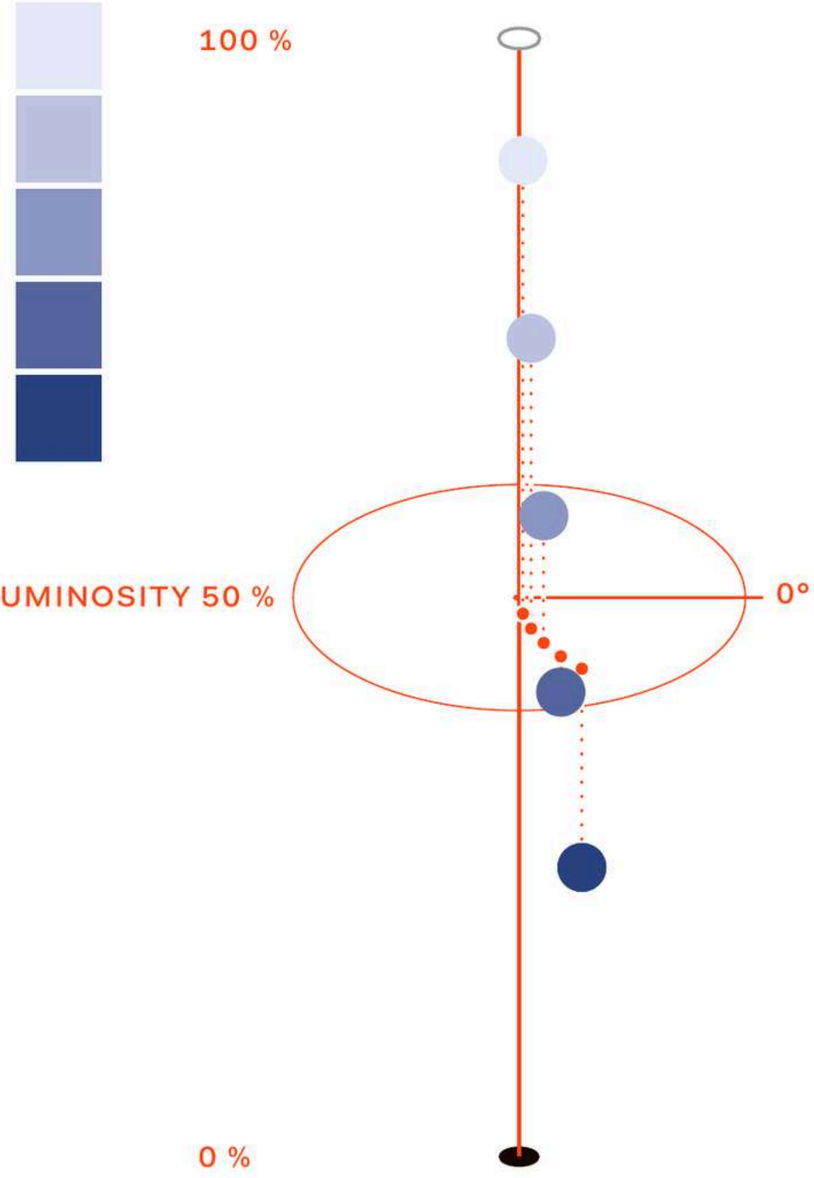
No data

Percentage of population with tertiary education in Europe, ages 25–64

Source: Eurostat 2018. *edat_lfse_04*

Single-hue scale

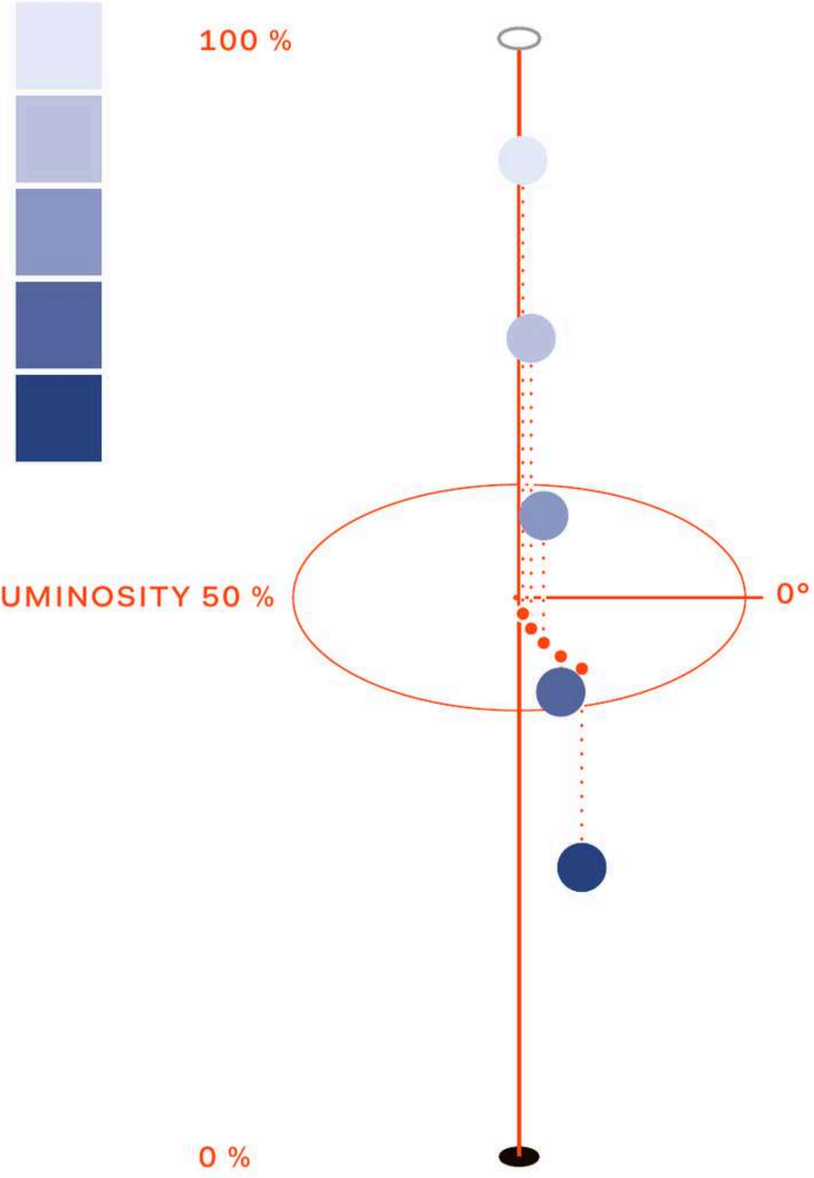
DISTRIBUTION
IN THE HCL SPACE

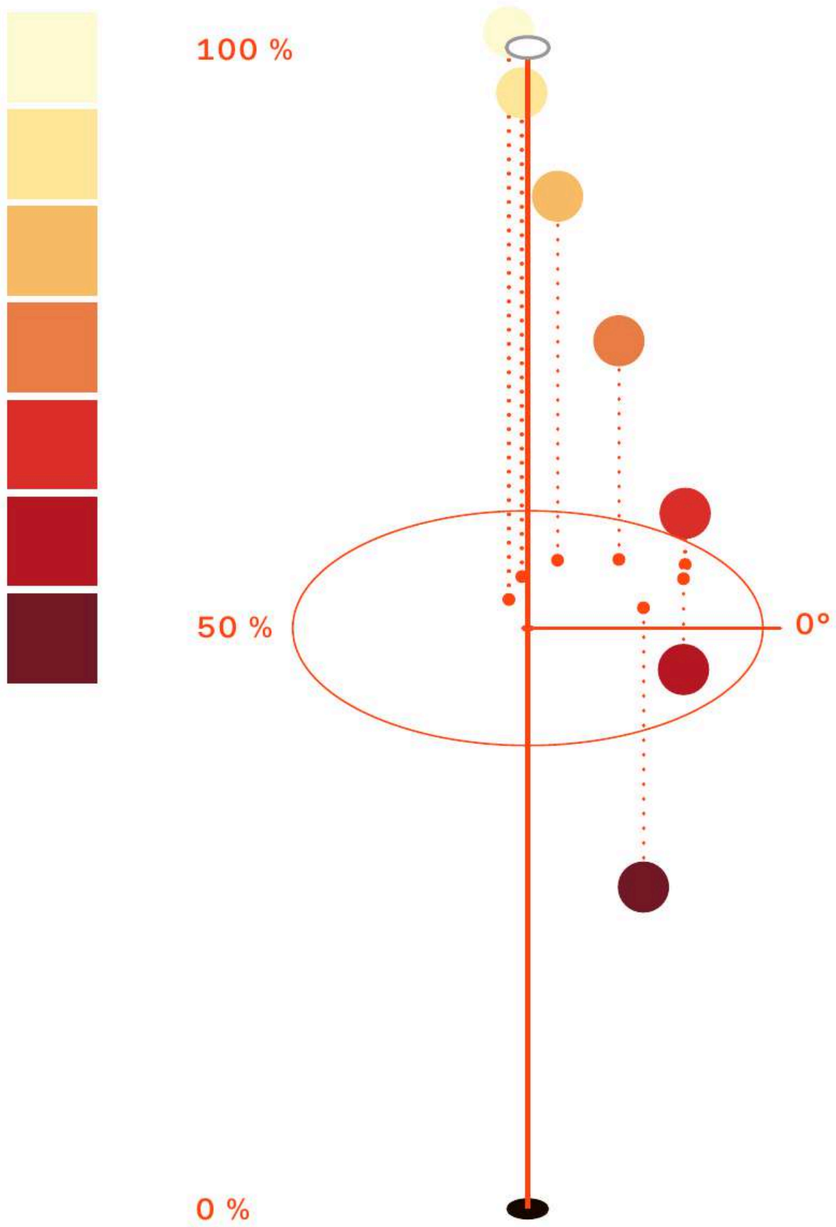


Single-hue scale

DISTRIBUTION IN THE HCL SPACE

Two-color sequential scale—light yellow to dark red



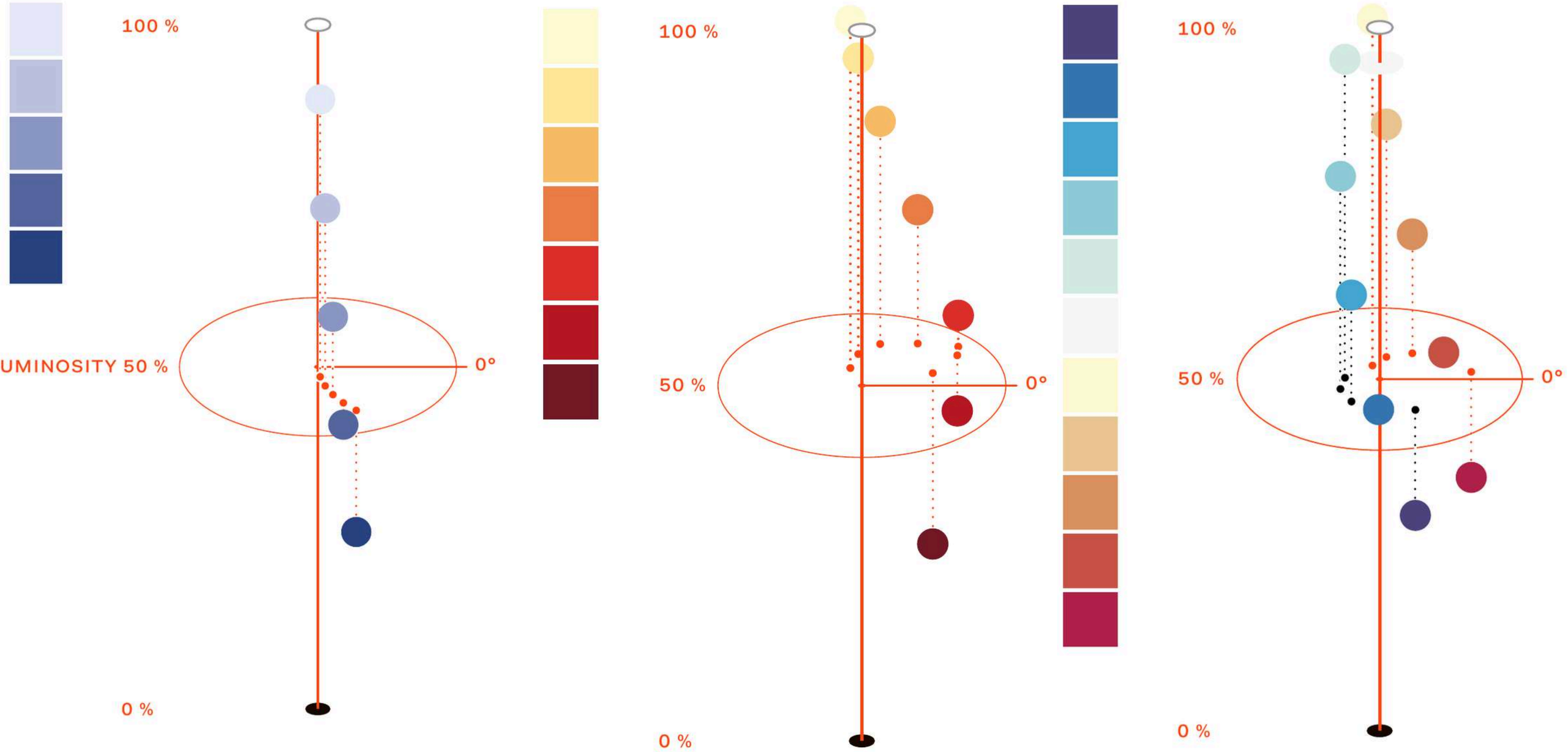


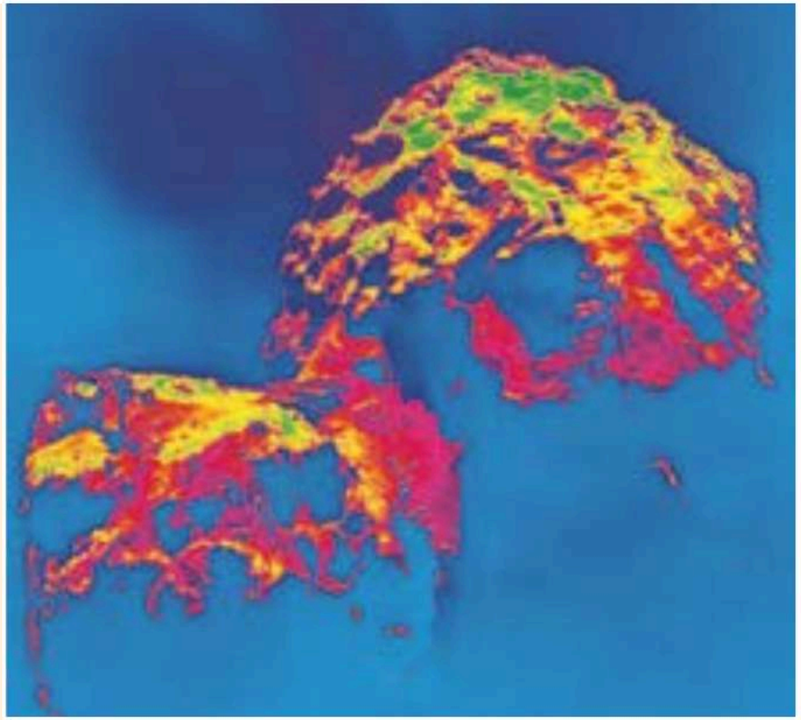
Single-hue scale

DISTRIBUTION IN THE HCL SPACE

Two-color sequential scale—light yellow to dark red

Bipolar (diverging) color scale—blue to red through gray



A rainbow scale and a corresponding grayscale image that shows the variations in lightness present in most rainbow scales.

A rainbow scale and a corresponding grayscale image that shows the variations in lightness present in most rainbow scales.