http://poloclub.gatech.edu/cse6242

CSE6242: Data & Visual Analytics

Information Visualization Crash Course

(AKA Information Visualization 101)

Chad Stolper
Google
(graduated from Georgia Tech CS PhD)

What is Infovis?

Why is it Important?

Human Perception

Chart Basics

(If Time, Some Color Theory)

The Shneiderman Mantra

Where to Learn More

What is Information Visualization?

Information Visualization

"The use of **computer**-supported, **interactive**, **visual** representations of abstract data to **amplify cognition**."

Card, Mackinlay, and Shneiderman 1999

Communication

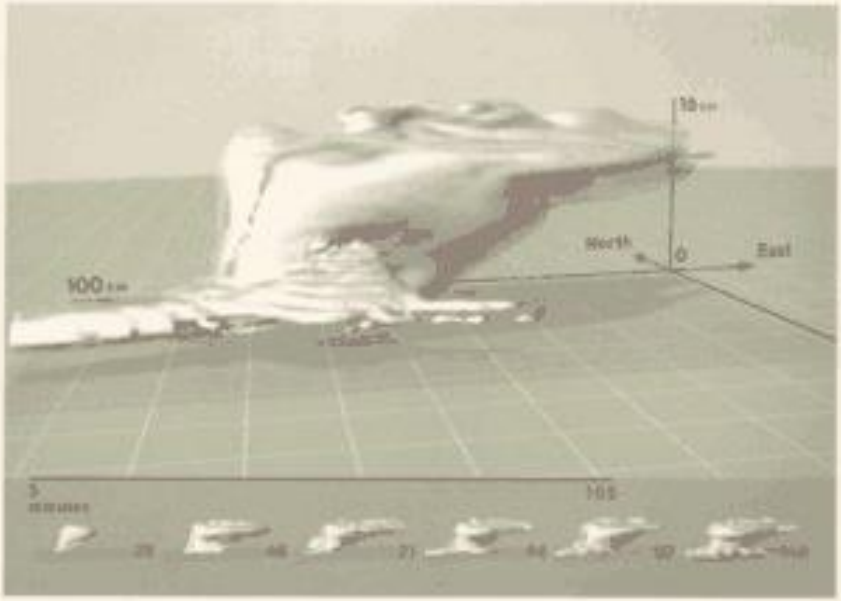
Exploratory Data Analysis (EDA)

Communication

(gone wrong)

EDWARD R. TUFTE

VISUAL EXPLANATIONS



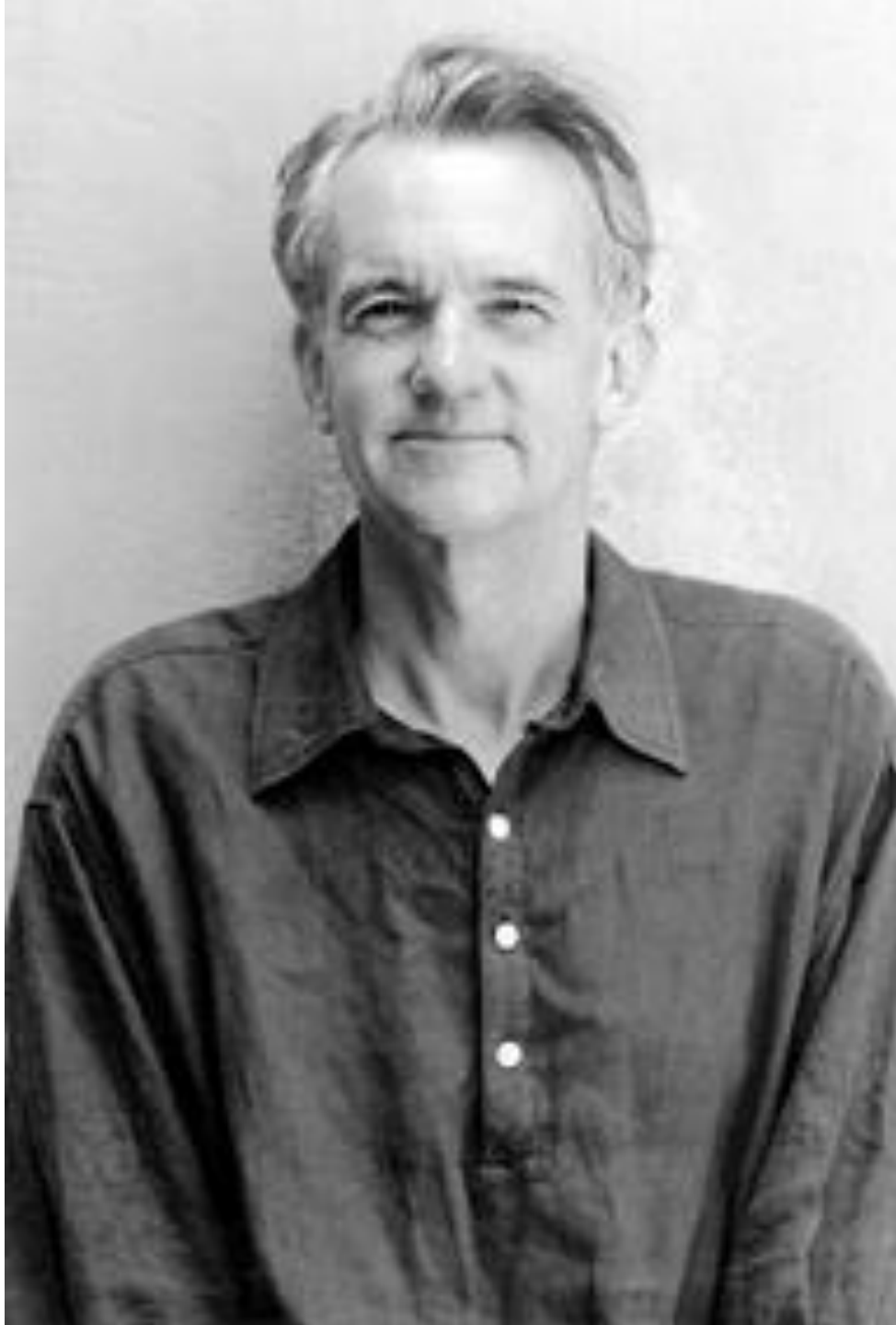
IMAGES AND QUANTITIES, EVIDENCE AND NARRATIVE

Edward Tufte

An American statistician and professor emeritus of political science, statistics, and computer science at Yale University.

He is noted for his writings on information design and as a pioneer in the field of data visualization.

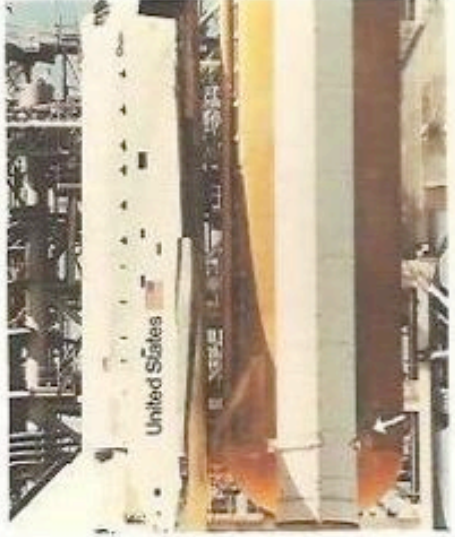
-Wikipedia



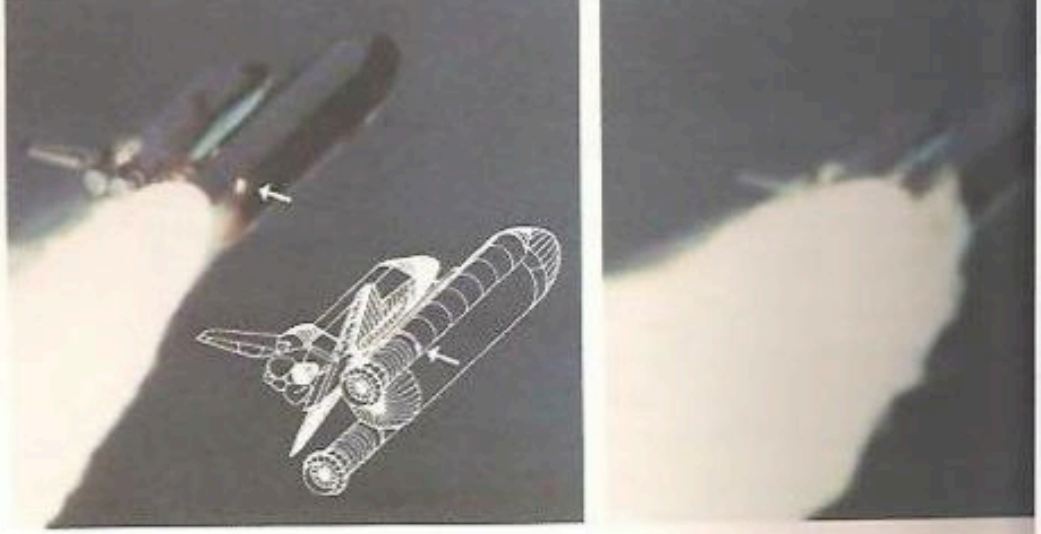
Space Shuttle Challenger

January 28, 1986

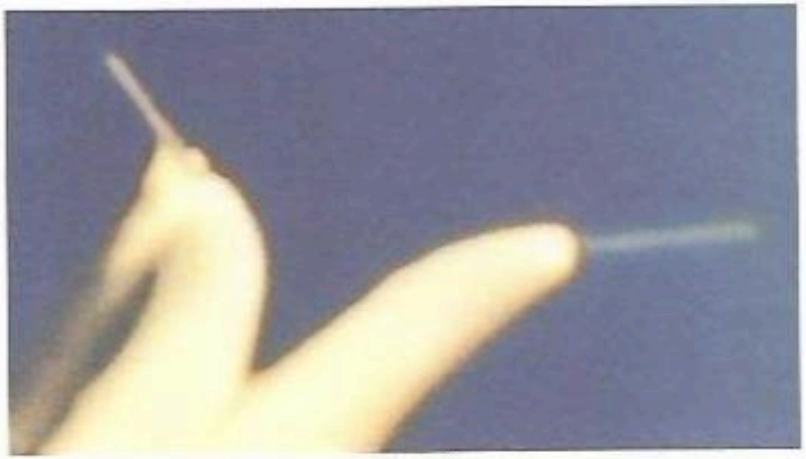
Morning Temperature: 31°F



Less than 1 second after ignition, a puff of smoke appeared at the aft joint of the right booster, indicating that the O-rings burned through and failed to seal. At this point, all was lost.

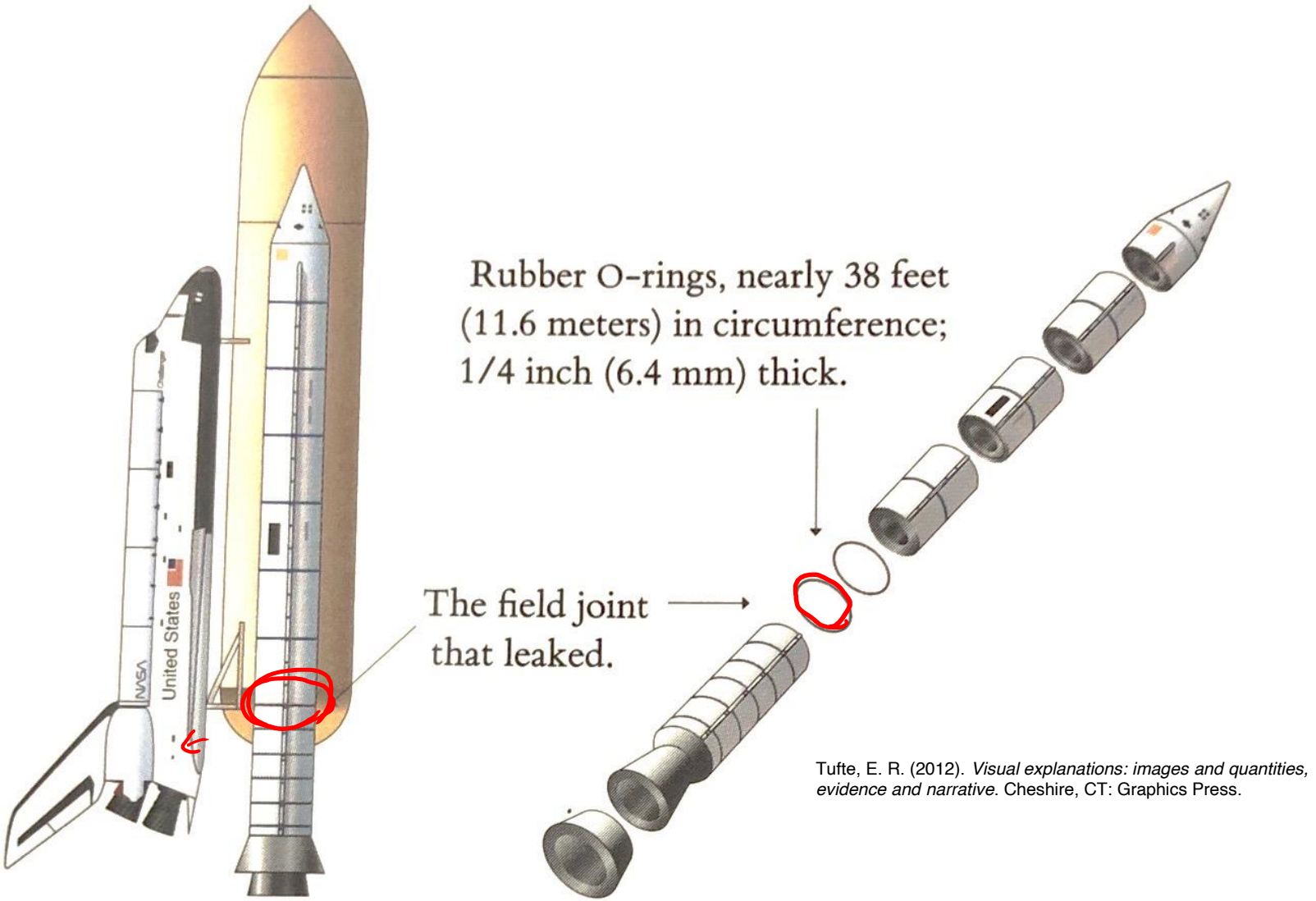


On the launch pad, the leak lasted only about 2 seconds and then apparently was plugged by [illegible] and insulation as the shuttle rose, flying through rather strong cross-winds. Then 58.788 seconds [illegible] ignition, when the Challenger was 6 miles up, a flicker of flame emerged from the leaky joint. [illegible] seconds, the flame grew and engulfed the fuel tank (containing liquid hydrogen and liquid [illegible]). That tank ruptured and exploded, destroying the shuttle.



As the shuttle exploded and broke up at approximately 73 seconds after launch, the two booster rockets crisscrossed and continued flying wildly. The right booster, identifiable by its failure plume, is now to the left of its non-defective counterpart.

The flight crew of Challenger 51-L. Front row, left to right: [illegible] Smith, pilot; Francis R. (Dick) Scobee, commander; Ronald [illegible]. Back row: Ellison S. Onizuka, S. Christa McAuliffe, George [illegible] Judith A. Resnik.



Tufte, E. R. (2012). *Visual explanations: images and quantities, evidence and narrative*. Cheshire, CT: Graphics Press.

Most Watched Science Experiment

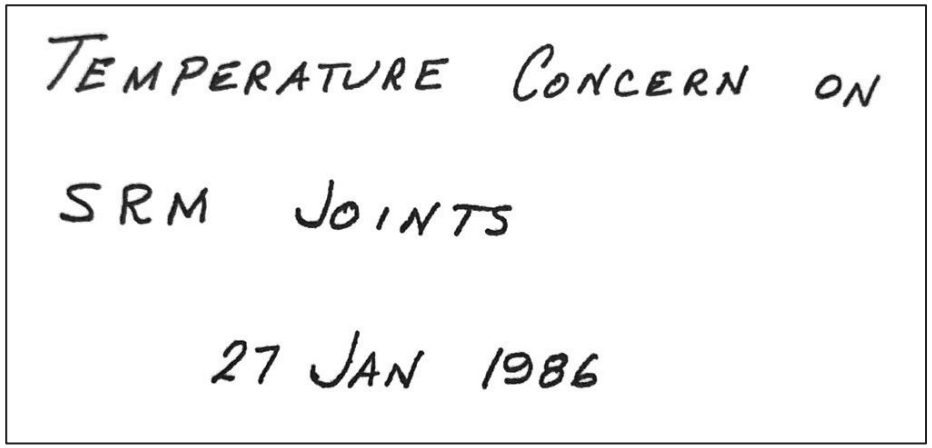
Richard Feynman, Physics Nobel laureate explained how rubber became rigid in cold temperate

YouTube video: https://youtu.be/6Rwcbsn19c0

Video originally from: http://www.FeynmanPhysicsLectures.com

How did this happen?

Engineers at **Morton Thiokol**, the rocket maker, presented on the day before and recommended not to launch.



Tufte, E. R. (2012). *Visual explanations: images and quantities, evidence and narrative*. Cheshire, CT: Graphics Press.

Conclusions:

- Temperature of O-ring is not only parameter controlling Blow-By

  SRM 15 with Blow-By had an O-ring temp at 53°F
  SRM 22 with Blow-By had an O-ring temp at 75°F
  Four Development Motors with no Blow-By were tested at O-ring temp of 47° to 52°F

  Development Motors had putty packing which resulted in better performance

- At about 50°F Blow-By could be experienced in case joints

- Temp for SRM 25 on 1-28-86 launch will be 29°F 9 AM
  38°F 2 PM

- Have no data that would indicate SRM 25 is different than SRM 15 other than temp

Recommendations:

- O-ring temp must be ≥ 53°F at launch

  Development Motors at 47° to 52°F with putty packing had no Blow-By
  SRM 15 (the best simulation) worked at 53°F

- Project ambient conditions (temp & wind) to determine launch time

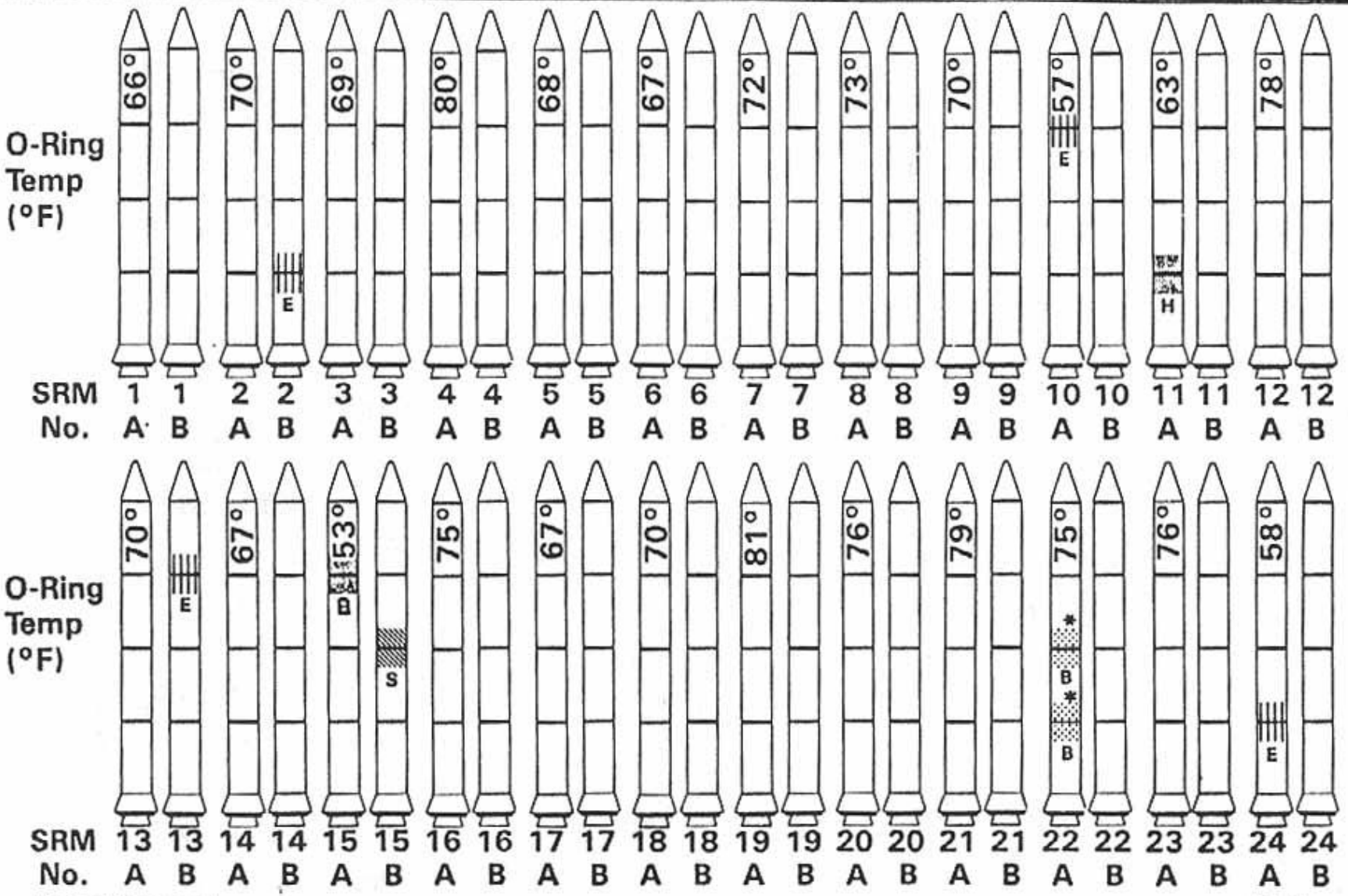
Recommendations:

- O-RING TEMP MUST BE ≥ 53°F AT LAUNCH

  DEVELOPMENT MOTORS AT 47° TO 52°F WITH PUTTY PACKING HAD NO BLOW-BY
  SRM 15 (THE BEST SIMULATION) WORKED AT 53°F

- PROJECT AMBIENT CONDITIONS (TEMP & WIND) TO DETERMINE LAUNCH TIME

History of O-Ring Damage in Field Joints (Cont)

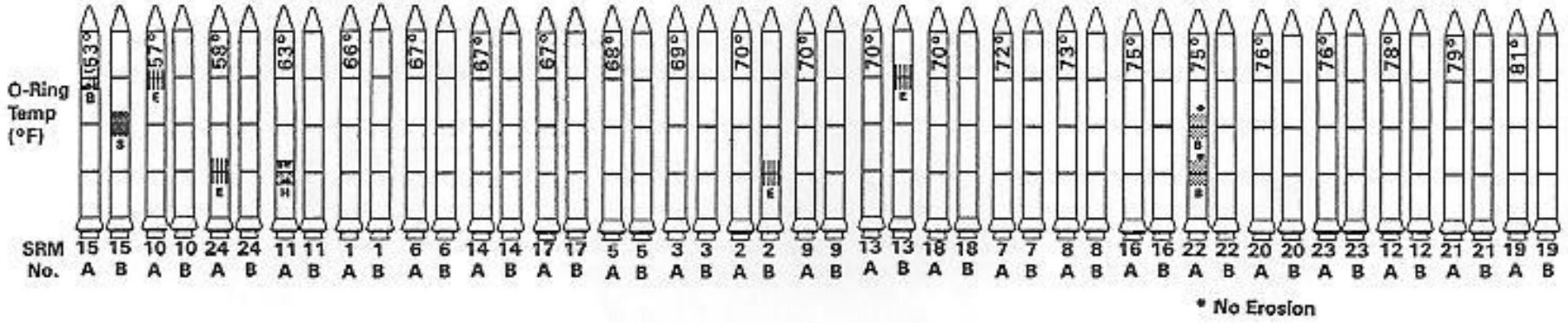


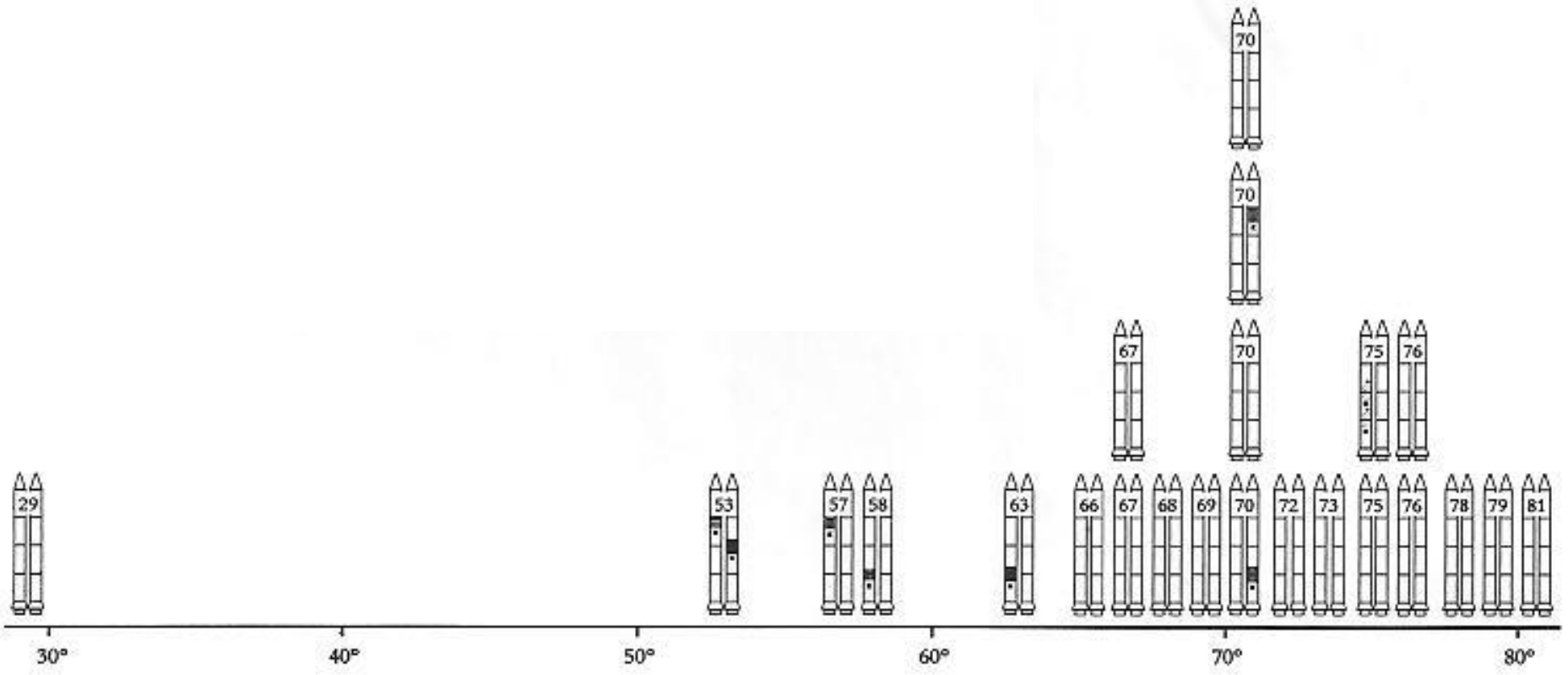
MORTON THIOKOL, INC.
Wasatch Operations

86486-1E

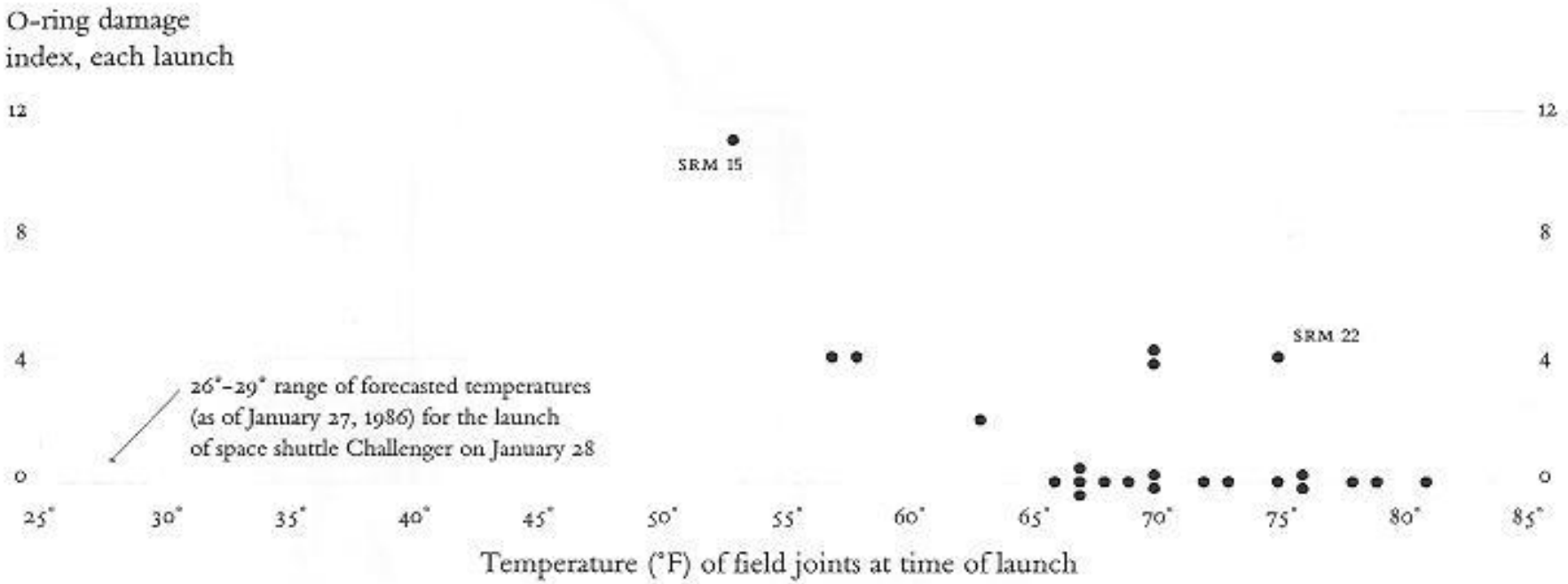
* No Erosion

INFORMATION ON THIS PAGE WAS PREPARED TO SUPPORT AN ORAL PRESENTATION AND CANNOT BE CONSIDERED COMPLETE WITHOUT THE ORAL DISCUSSION


O-Ring Temp (°F)
53°
57°
58°
63°
66°
67°
67°
67°
67°
68°
69°
70°
70°
70°
70°
72°
73°
75°
75°
76°
76°
78°
79°
81°
SRM No.
15 A
15 B
10 A
10 B
24 A
24 B
11 A
11 B
1 A
1 B
6 A
6 B
14 A
14 B
17 A
17 B
5 A
5 B
3 A
3 B
2 A
2 B
9 A
9 B
13 A
13 B
18 A
18 B
7 A
7 B
8 A
8 B
16 A
16 B
22 A
22 B
20 A
20 B
23 A
23 B
12 A
12 B
21 A
21 B
19 A
19 B
* No Erosion


70
70
67
70
75
76
29
53
57
58
63
66
67
68
69
70
72
73
75
76
78
79
81
30°
40°
50°
60°
70°
80°

Flight	Date	Temperature °F	Erosion incidents	Blow-by incidents	Damage index	Comments
51-C	01.24.85	53°	3	2	11	Most erosion any flight; blow-by; back-up rings heated.
41-B	02.03.84	57°	1		4	Deep, extensive erosion.
61-C	01.12.86	58°	1		4	O-ring erosion on launch two weeks before Challenger.
41-C	04.06.84	63°	1		2	O-rings showed signs of heating, but no damage.
1	04.12.81	66°			0	Coolest (66°) launch without O-ring problems.
6	04.04.83	67°			0	
51-A	11.08.84	67°			0	
51-D	04.12.85	67°			0	
5	11.11.82	68°			0	
3	03.22.82	69°			0	
2	11.12.81	70°	1		4	Extent of erosion not fully known.
9	11.28.83	70°			0	
41-D	08.30.84	70°	1		4	
51-G	06.17.85	70°			0	
7	06.18.83	72°			0	
8	08.30.83	73°			0	
51-B	04.29.85	75°			0	
61-A	10.30.85	75°		2	4	No erosion. Soot found behind two primary O-rings.
51-I	08.27.85	76°			0	
61-B	11.26.85	76°			0	
41-G	10.05.84	78°			0	
51-J	10.03.85	79°			0	
4	06.27.82	80°			?	O-ring condition unknown; rocket casing lost at sea.
51-F	07.29.85	81°			0	


O-ring damage
index, each launch
12
8
4
0
SRM 15
SRM 22
26°–29° range of forecasted temperatures
(as of January 27, 1986) for the launch
of space shuttle Challenger on January 28
25°
30°
35°
40°
45°
50°
55°
60°
65°
70°
75°
80°
85°
Temperature (°F) of field joints at time of launch

So, communication is **extremely important**.

Visualization can help with that – **communicate ideas and insights**.

http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html

Visualization can also help with **Exploratory Data Analysis** (EDA)

But why do you need to explore data at all???

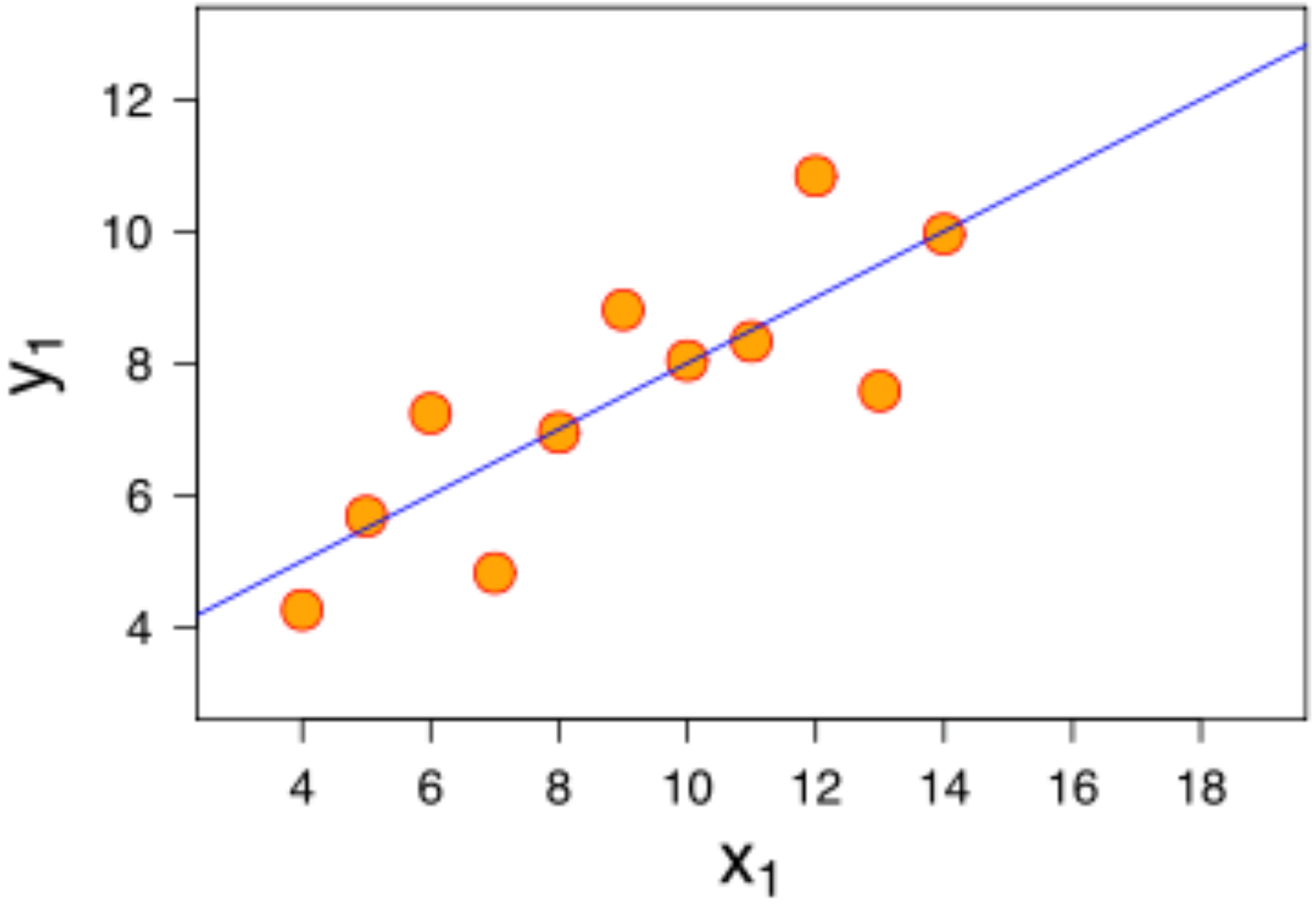
“There are three kinds of lies: lies, damned lies, and statistics.”

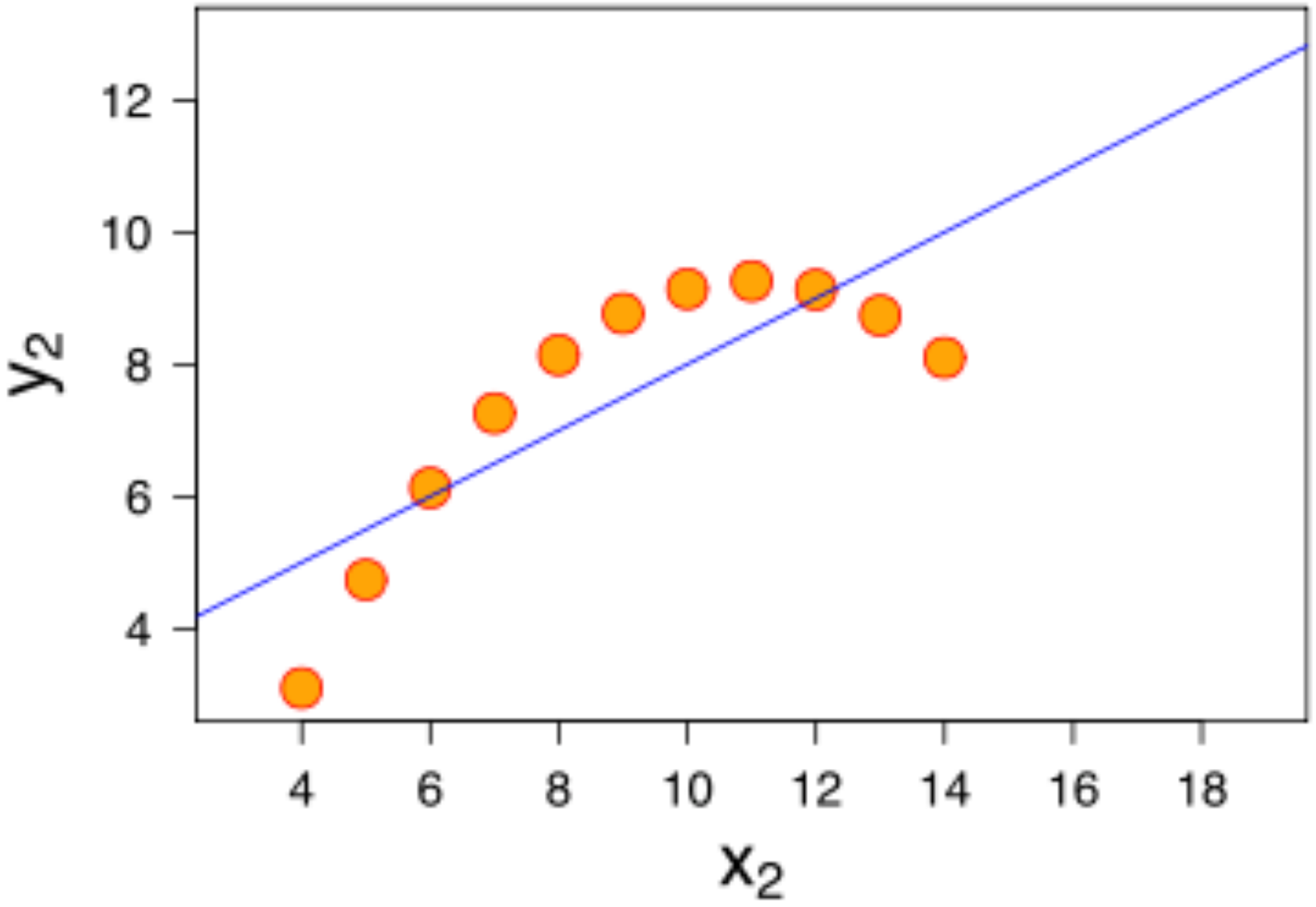
https://en.wikipedia.org/wiki/Lies,_damned_lies,_and_statistics

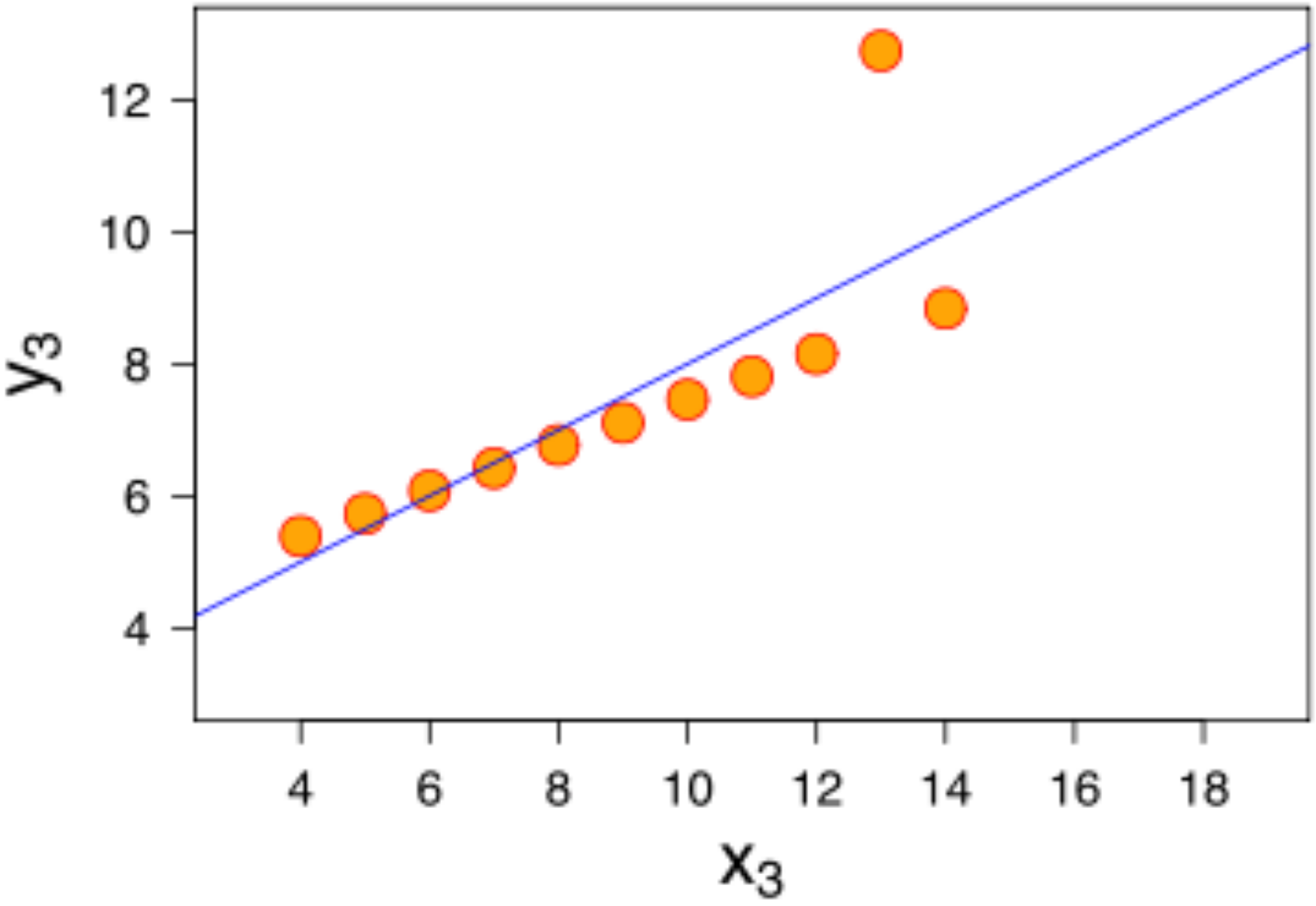
Mystery Data Set

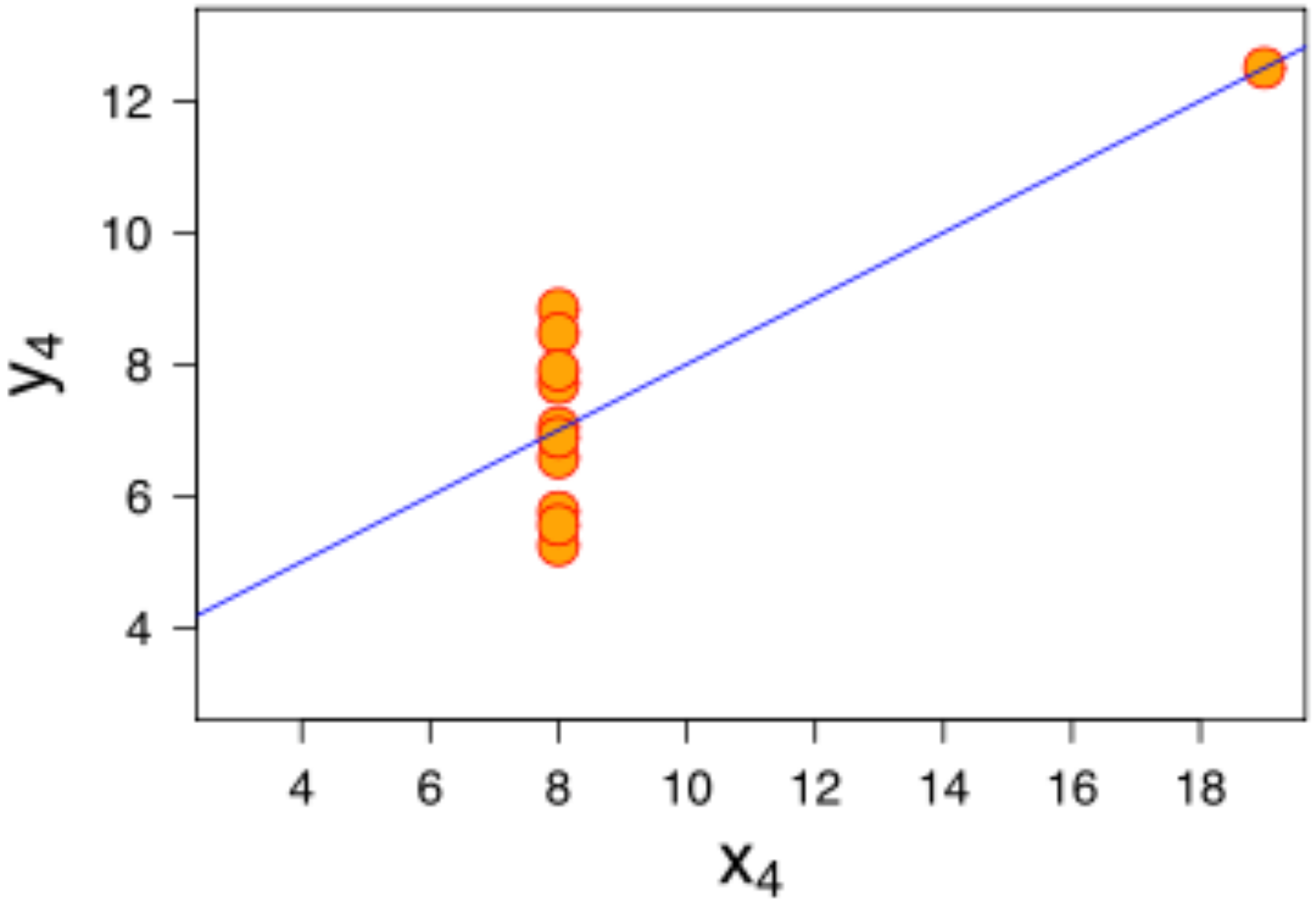
Mystery Data Set

Property	Value
mean(x)	9
variance (x)	11
mean(y)	7.5
variance (y)	4.122
correlation (x,y)	0.816
Linear Regression Line	y = 3 + 0.5x

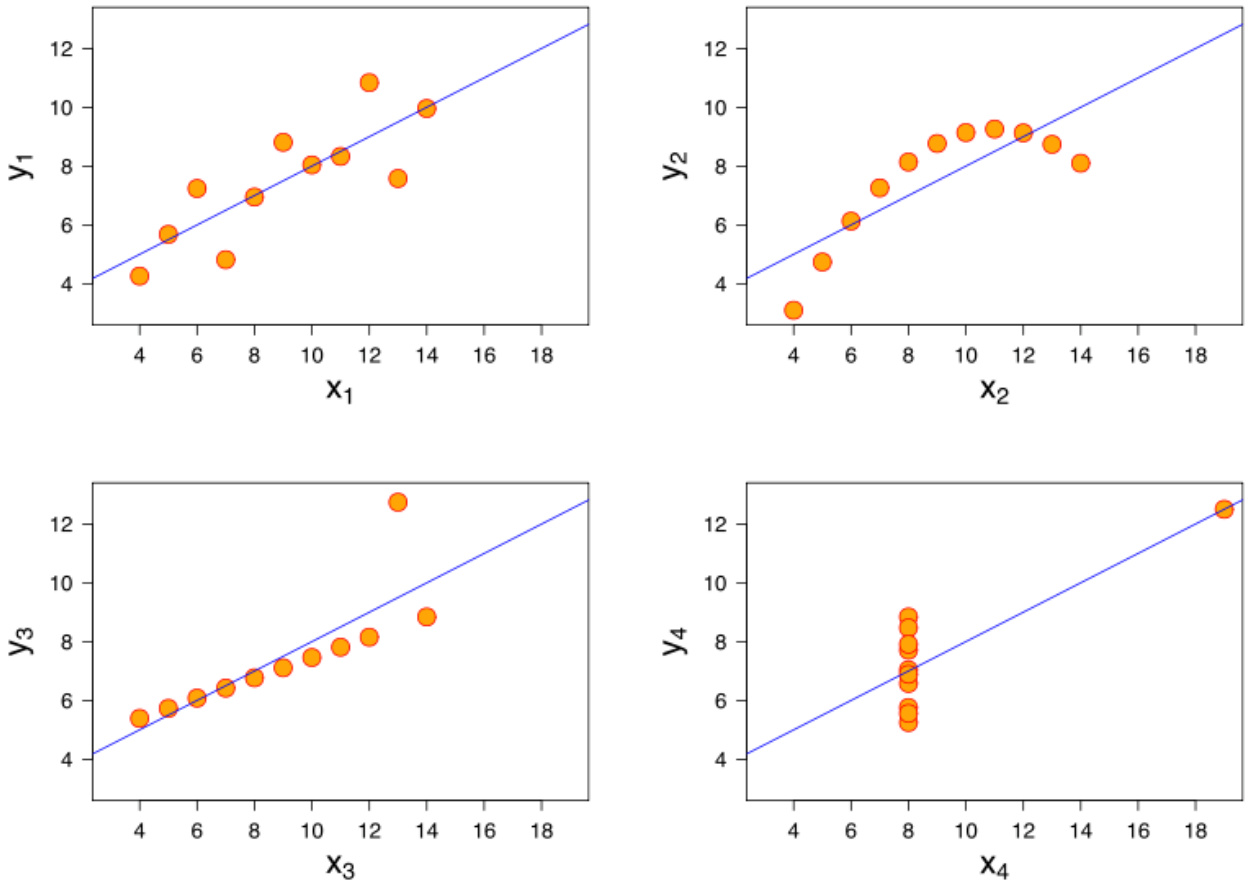

12
10
8
6
4
y_1
4
6
8
10
12
14
16
18
x_1


12
10
8
6
4
y_2
4
6
8
10
12
14
16
18
x_2


12
10
8
6
4
y_3
4
6
8
10
12
14
16
18
x_3

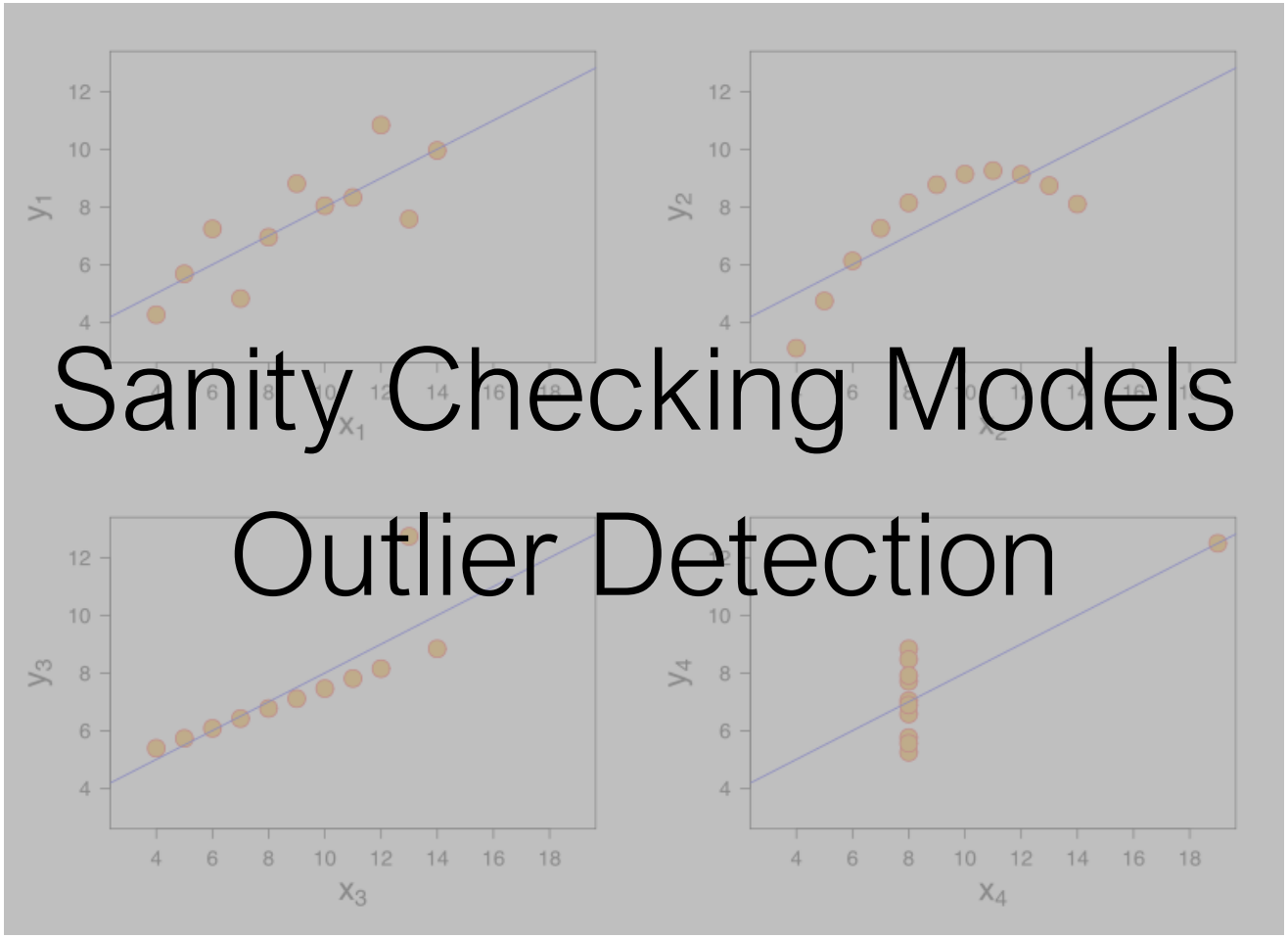

12
10
8
6
4
y_4
4
6
8
10
12
14
16
18
x_4

Anscombe's Quartet



https://en.wikipedia.org/wiki/Anscombe%27s_quartet

Anscombe's Quartet



Sanity Checking Models

Outlier Detection

Data visualization leverages **human perception**

Name the five senses.

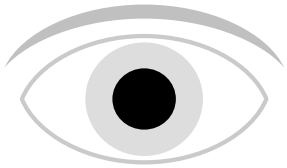
$$I(x) = \log\left(\frac{1}{p(x)}\right) \leftarrow p(x) \approx 0$$

Sense	Bandwidth (bits/sec)
Sight	10,000,000
Touch	1,000,000
Hearing	100,000
Smell	100,000
Taste	1,000

http://www.britannica.com/EBchecked/topic/287907/information-theory/214958/Physiology

A (Simple) Model of Human Visual Perception

A (Simple) Model of Human Perception



Stage 1

Parallel detection of basic features into an iconic store

Stage 2

Serial processing of object identification and spatial layout

Stage 1: Pre-Attentive Processing

Rapid

Parallel

Automatic

(Fleeting = lasting for a short time)

Stage 2: Serial Processing

Relatively Slow

(Incorporates Memory)

Manual

Stage 1: Pre-Attentive Processing

The eye moves every 200ms

(so this processing occurs every 200ms-250ms)

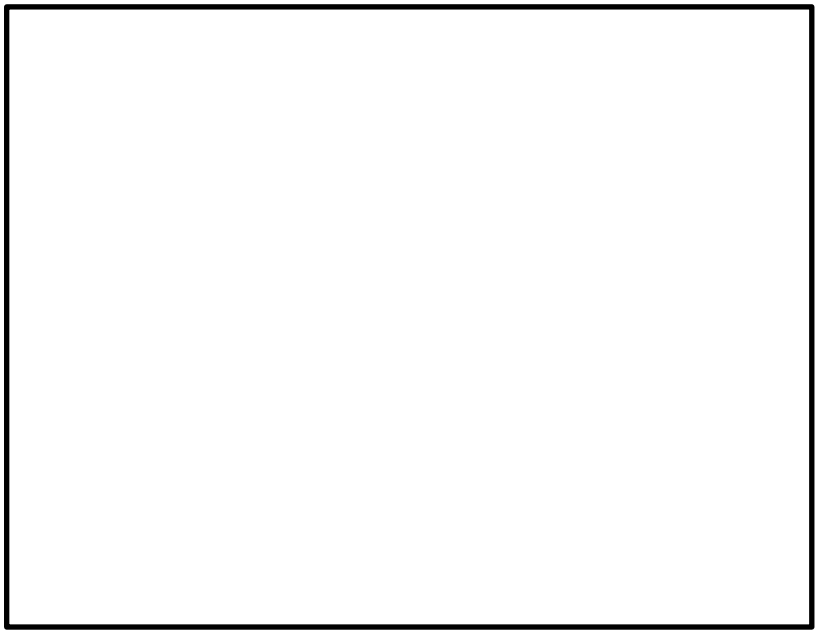
Example

128176875613897654698450698560498282676298098584582245098564589450984509809435859091030209905959595772564675050678904567884578980982167765487636490856091294968 6

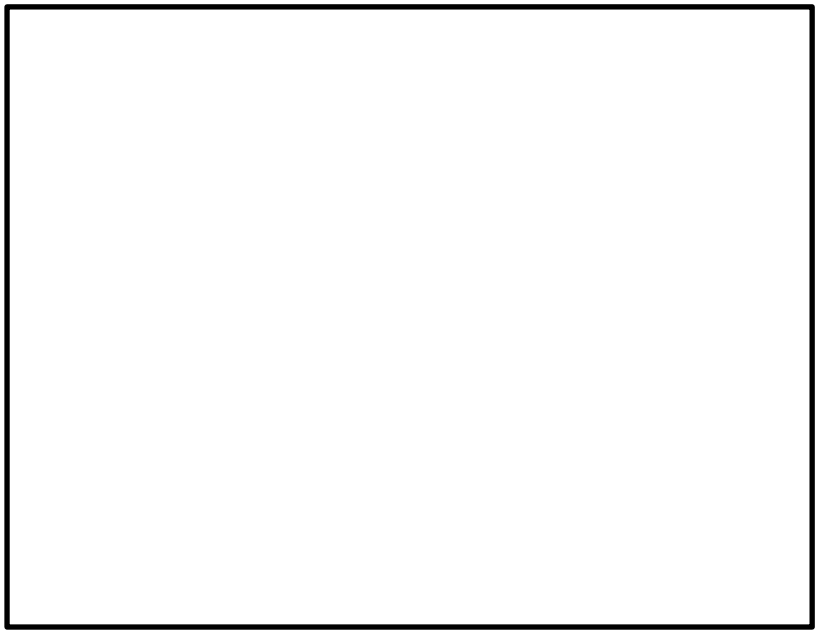
Example

128176875613897654698450698560498282676298098584582245098564589450984509809435859091030209905959595772564675050678904567884578980982167765487636490856091294968 6

A few more examples from Prof. Chris Healy at NC State



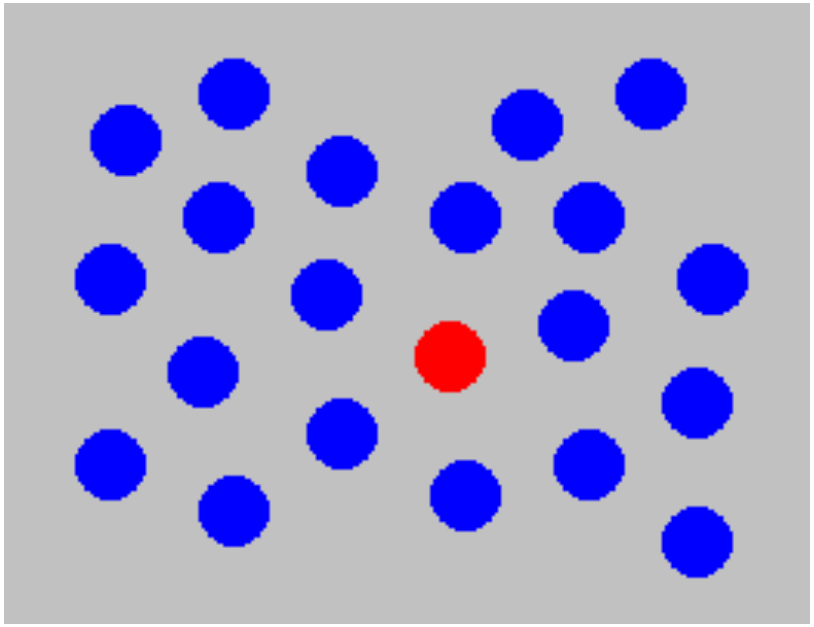
Left Side

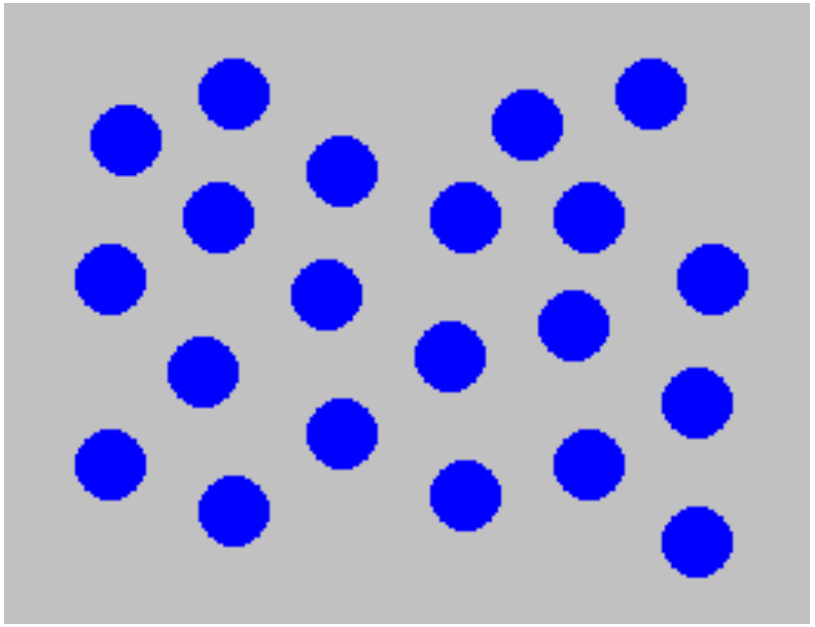


Right Side

Raise your hand if a RED DOT is present…

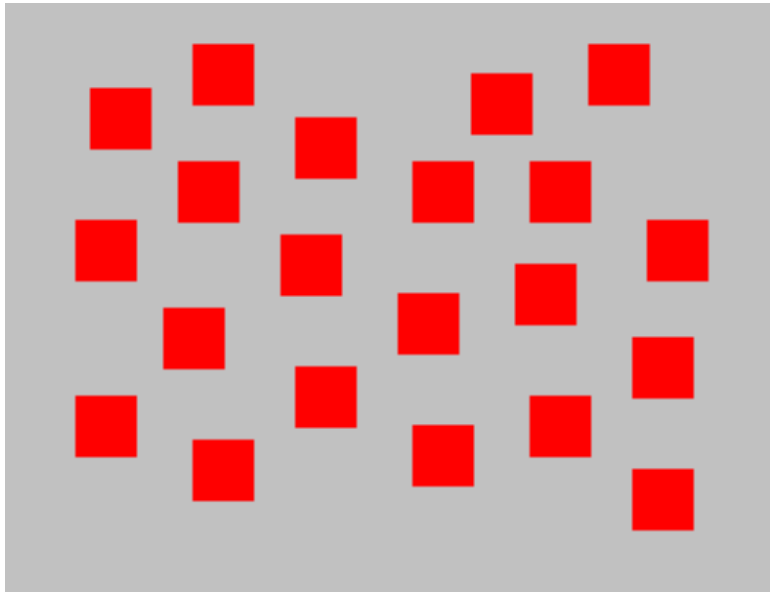
(On the left or on the right?)

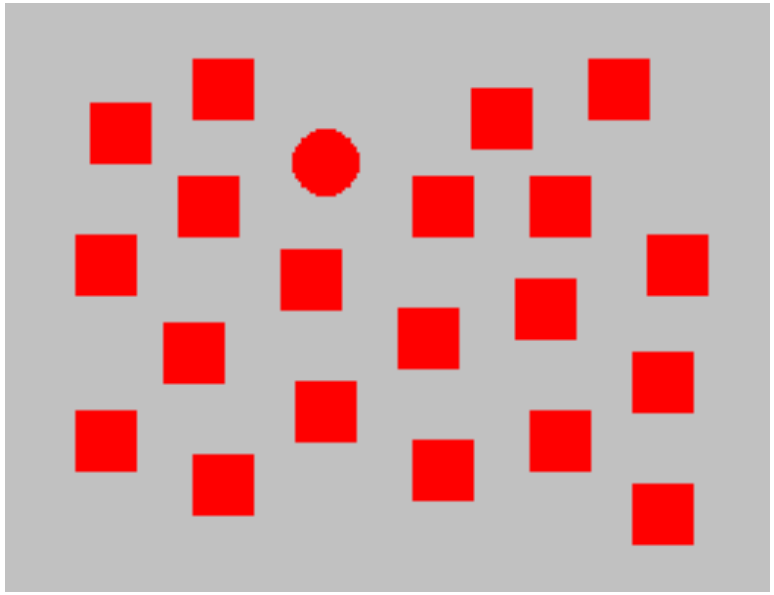




Color (hue) is pre-attentively processed.

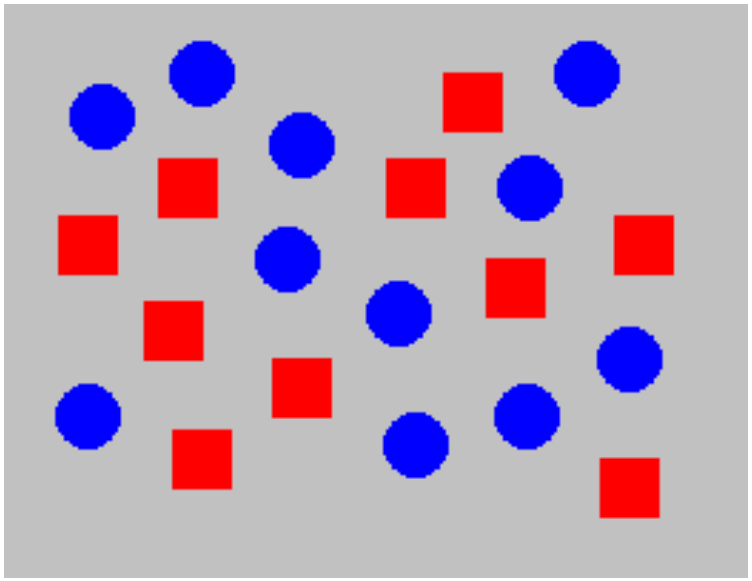
Raise your hand if a RED DOT is present…

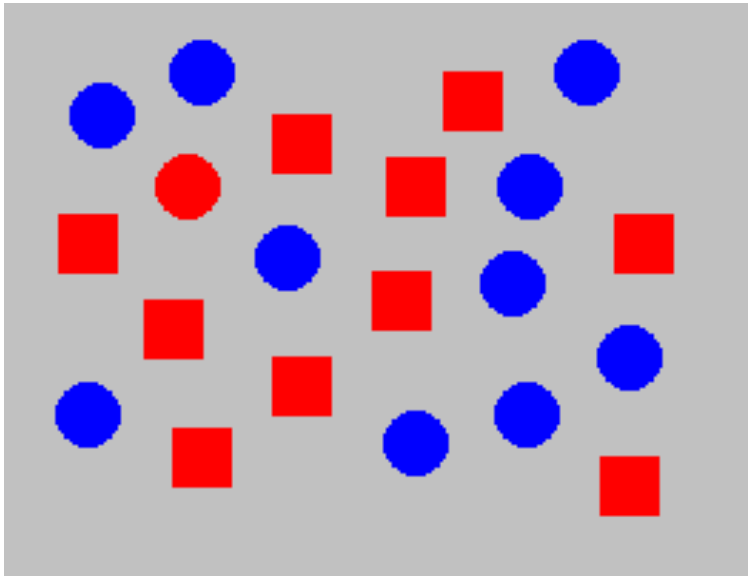




Shape is pre-attentively processed.

Determine if a RED DOT is present…



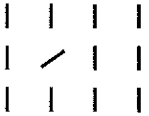
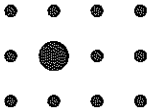
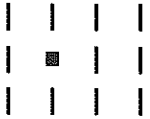
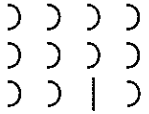
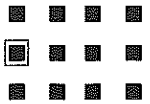
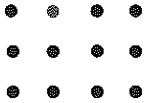
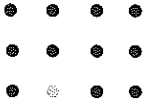


Hue and shape together are NOT pre-attentively processed.

Pre-Attentive Processing

- length
- width
- size
- curvature
- number
- terminators
- intersection
- closure

- hue
- lightness
- flicker
- direction of motion
- binocular lustre
- stereoscopic depth
- 3-D depth cues
- lighting direction

Group	Attribute		
Form	Length	Width	Orientation
			
	Size	Shape	Curvature
			
	Enclosure	Blur	
			
Color	Hue	Intensity	
			
Spatial Position	2-D Position	Spatial Grouping	
			
Motion	Direction		
			

Stephen Few
"Now You See It"
pg. 39

Pre-Attentive → Cognitive

Gestalt Psychology

Berlin, Early 1900s

Gestalt Psychology

Goal was to understand pattern perception

Gestalt (German) = “seeing the whole picture all at once” instead of a collection of parts

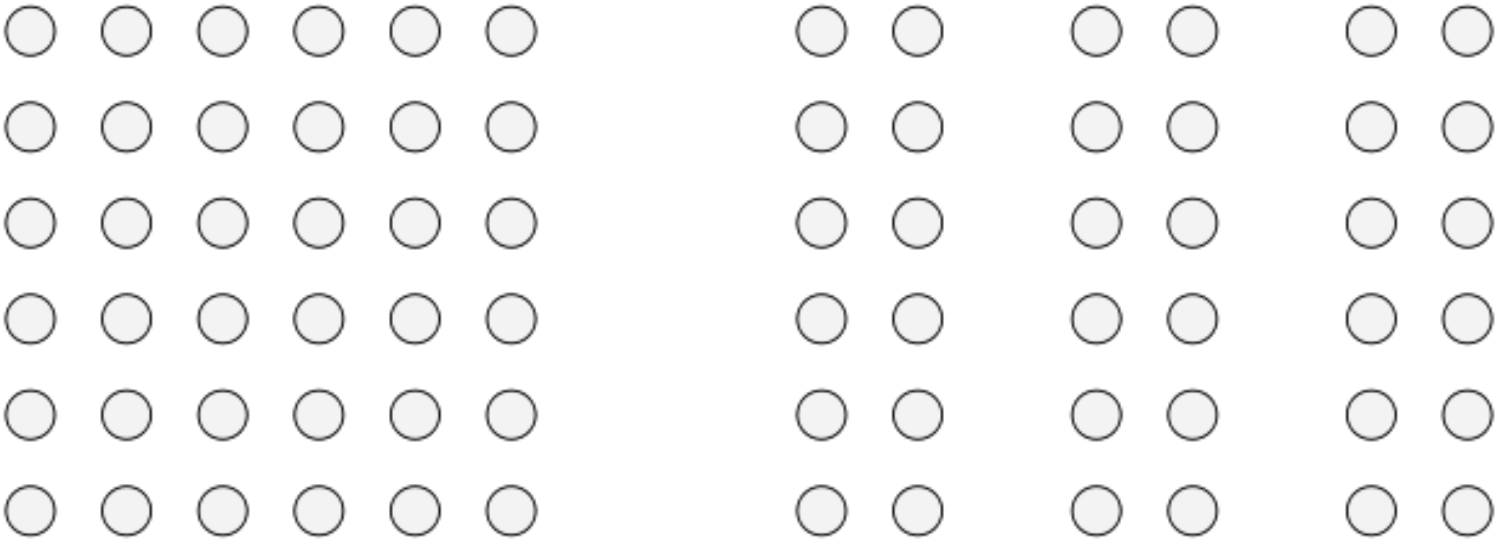
Identified 8 “Laws of Grouping”

http://study.com/academy/lesson/gestalt-psychology-definition-principles-quiz.html

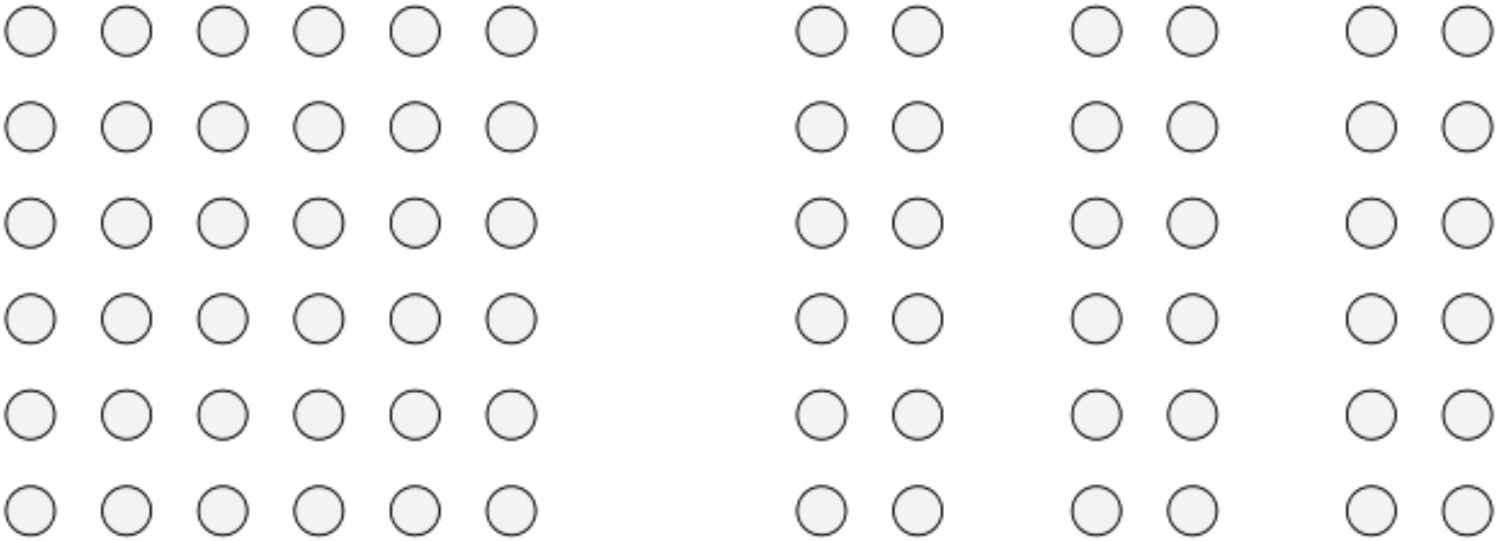
Gestalt Psychology

1. Proximity
2. Similarity
3. Closure
4. Symmetry
5. Common Fate
6. Continuity
7. Good Gestalt
8. Past Experience

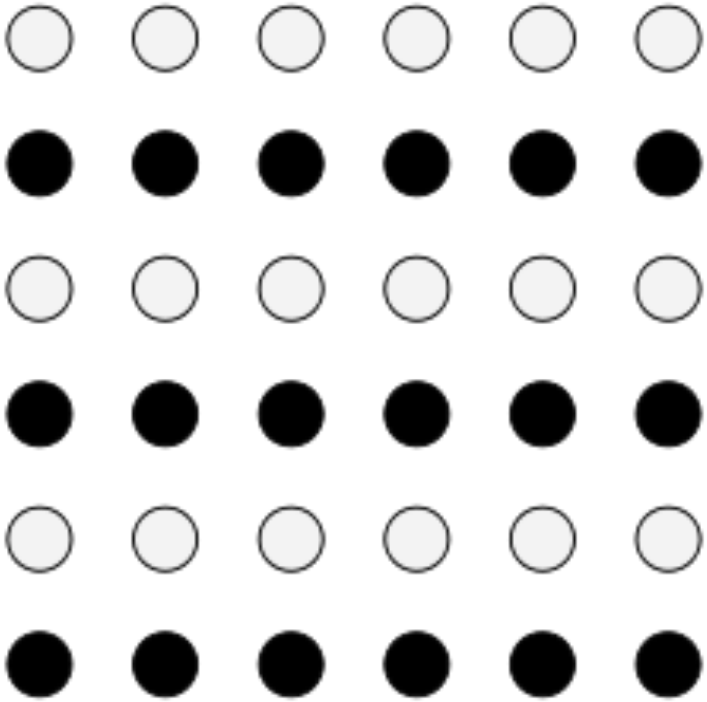
How many groups are there?



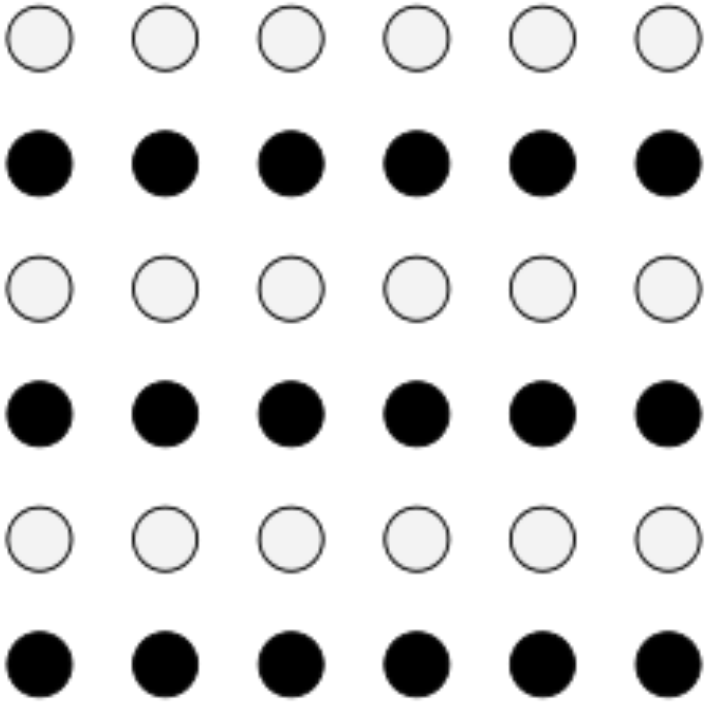
Proximity



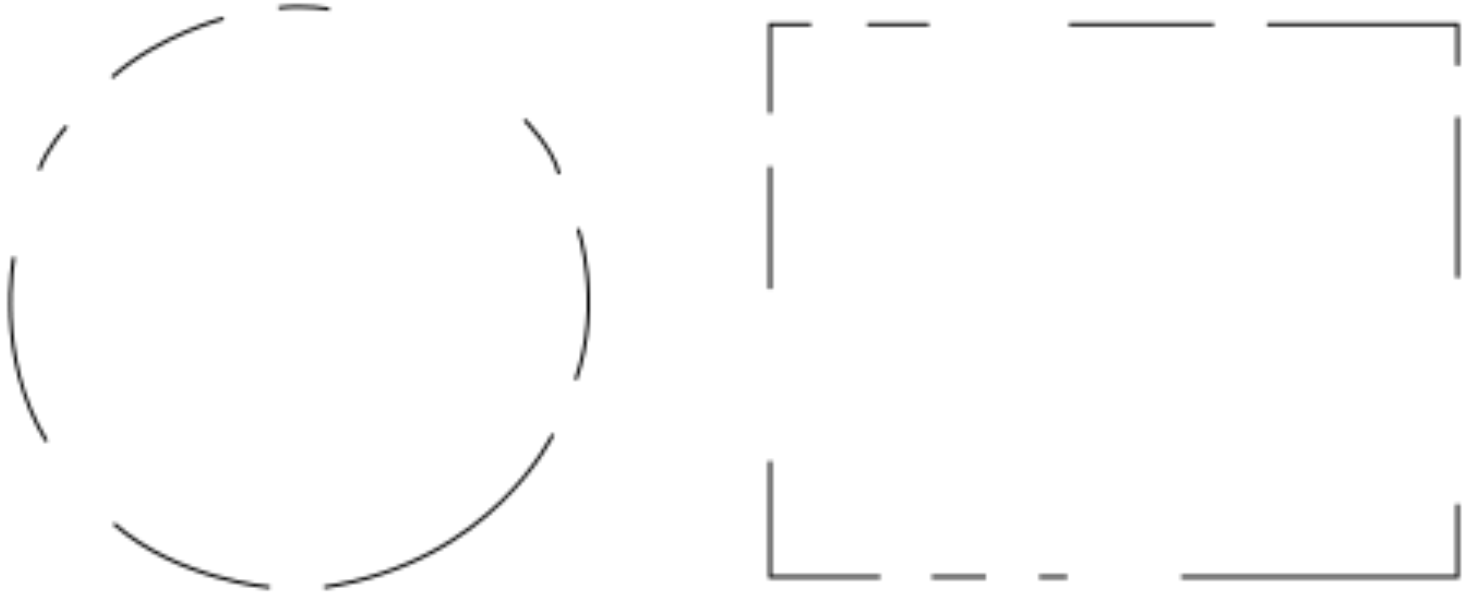
How many groups are there?



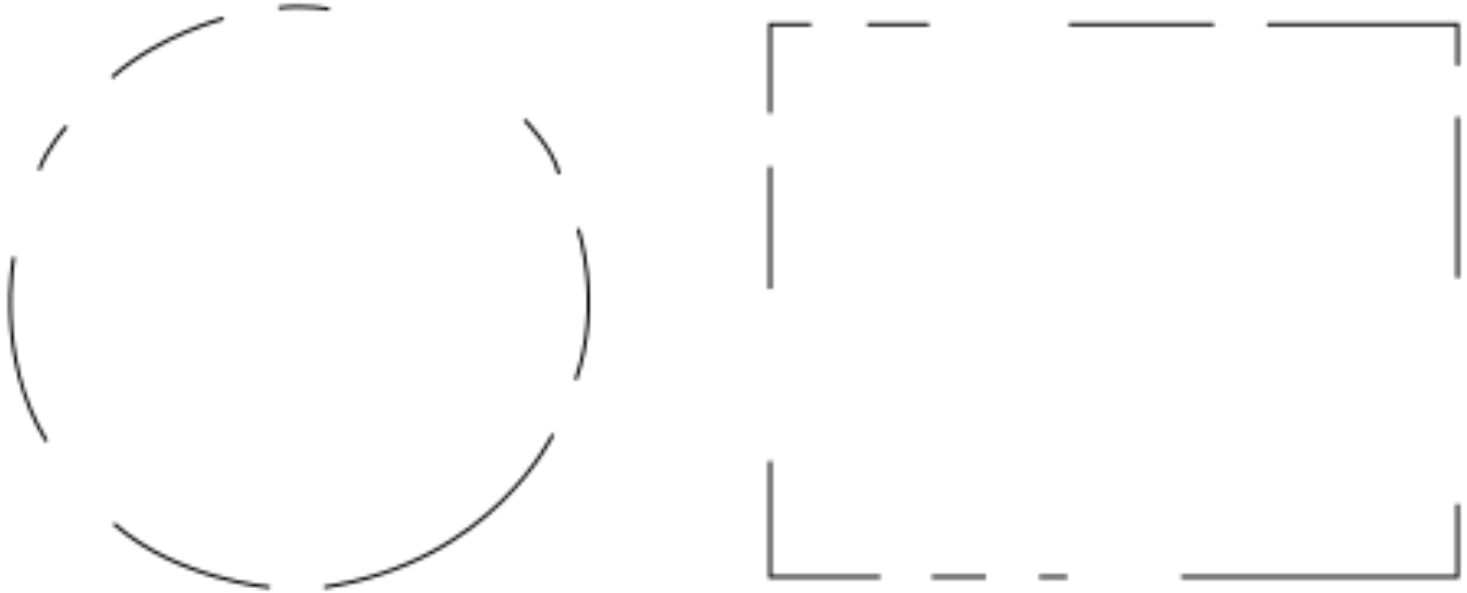
Similarity



How many shapes are there?



Closure



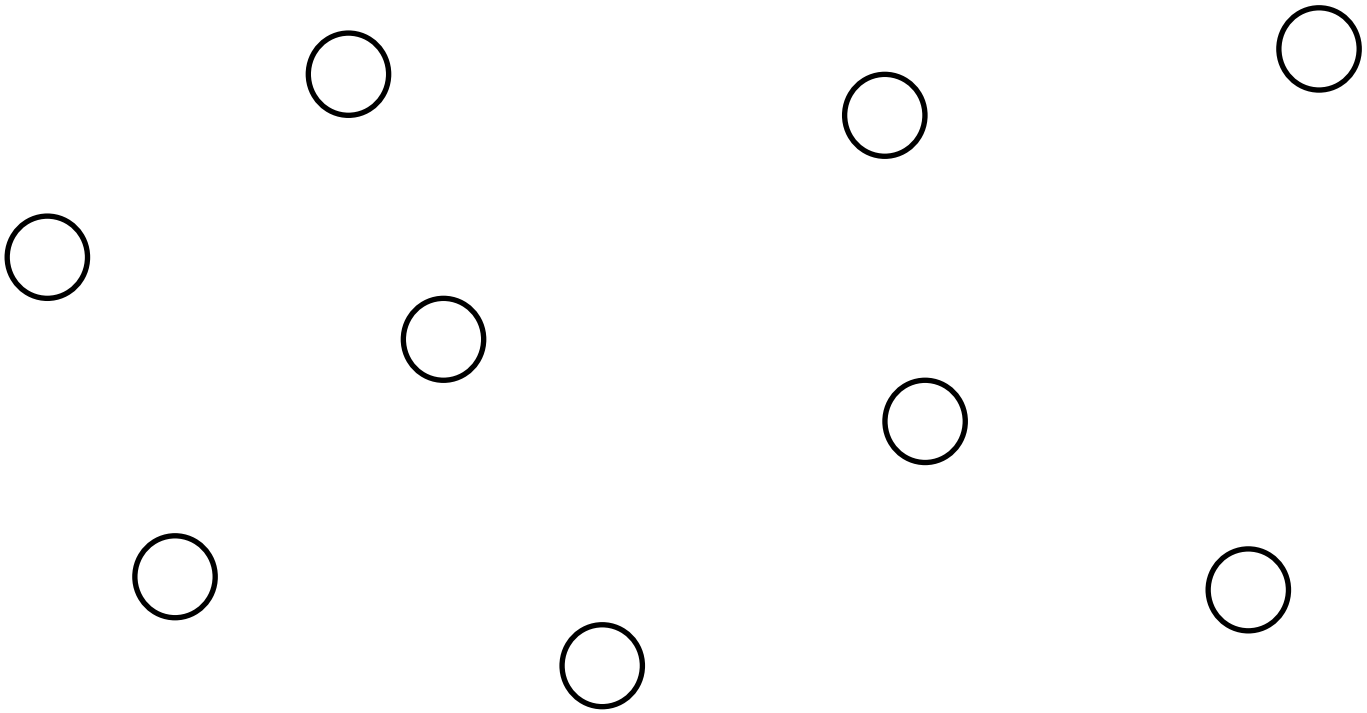
How many items are there?

( ) { } [ ]

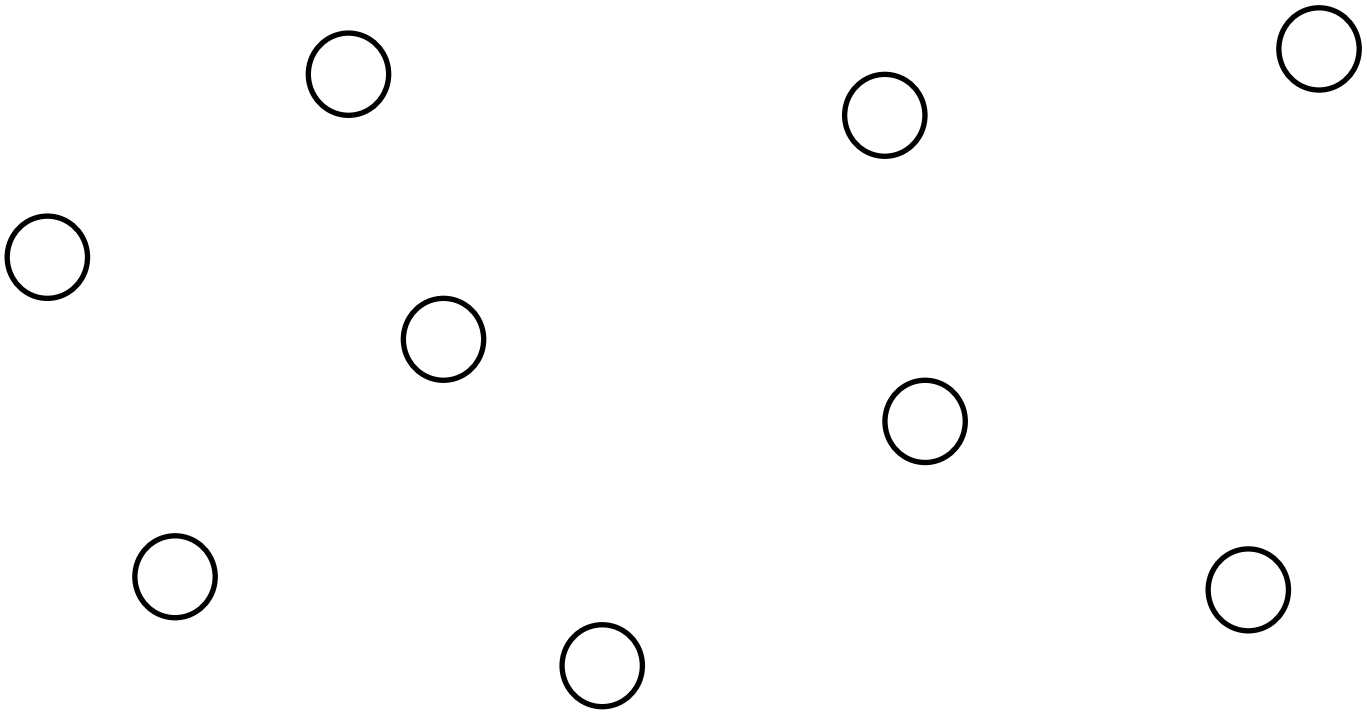
Symmetry

( ) { } [ ]

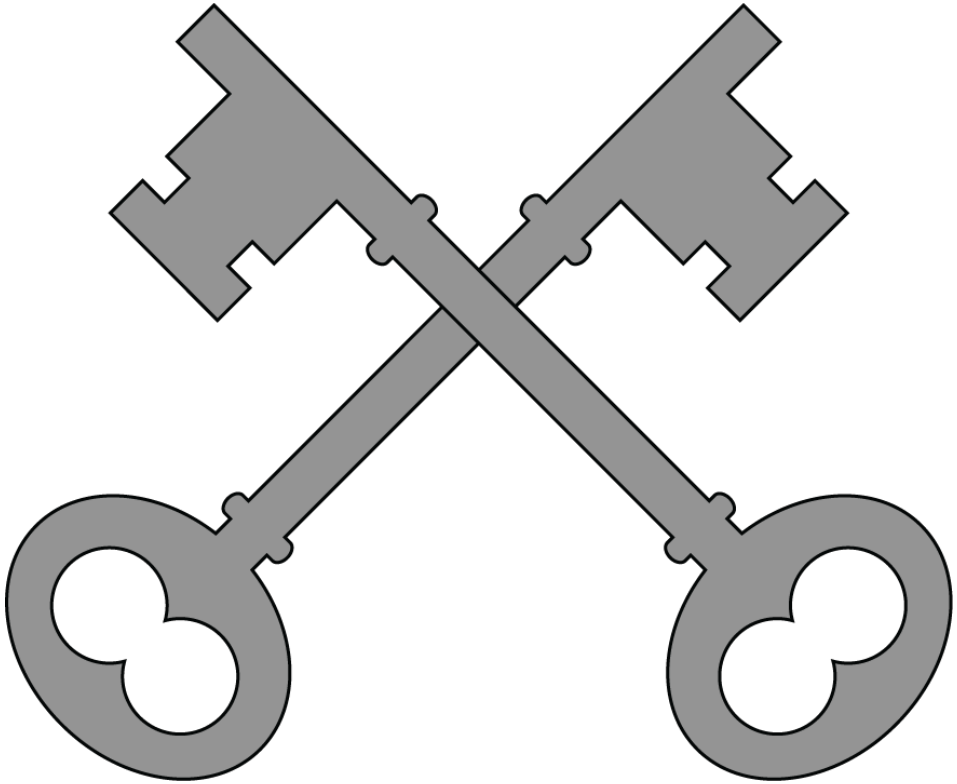
How many sets are there?



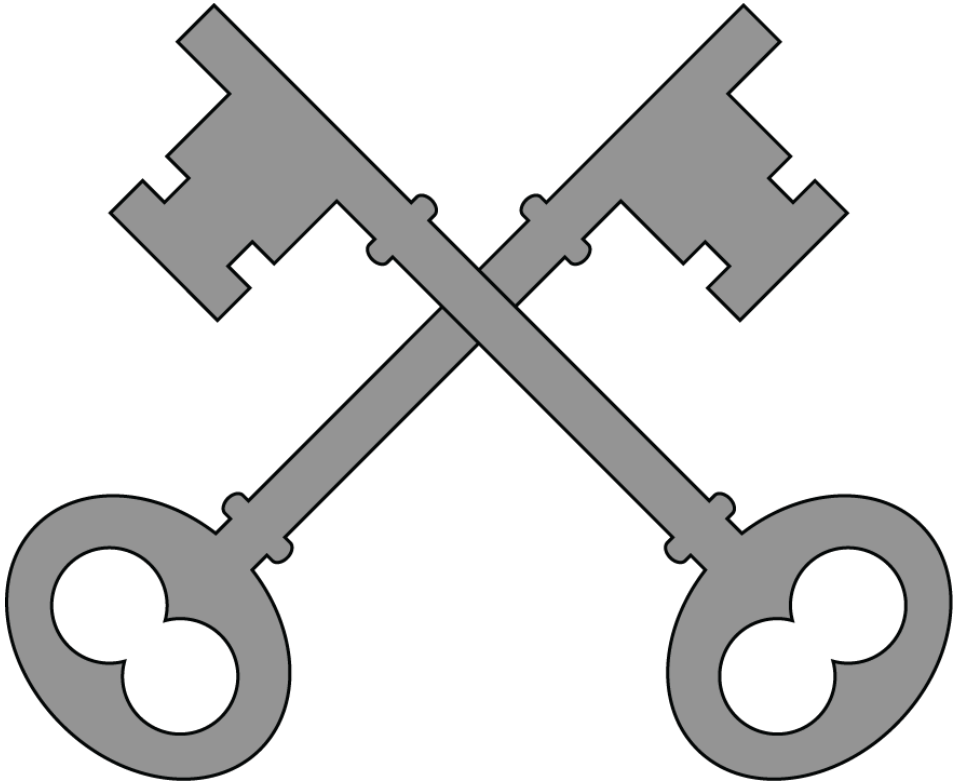
Common Fate



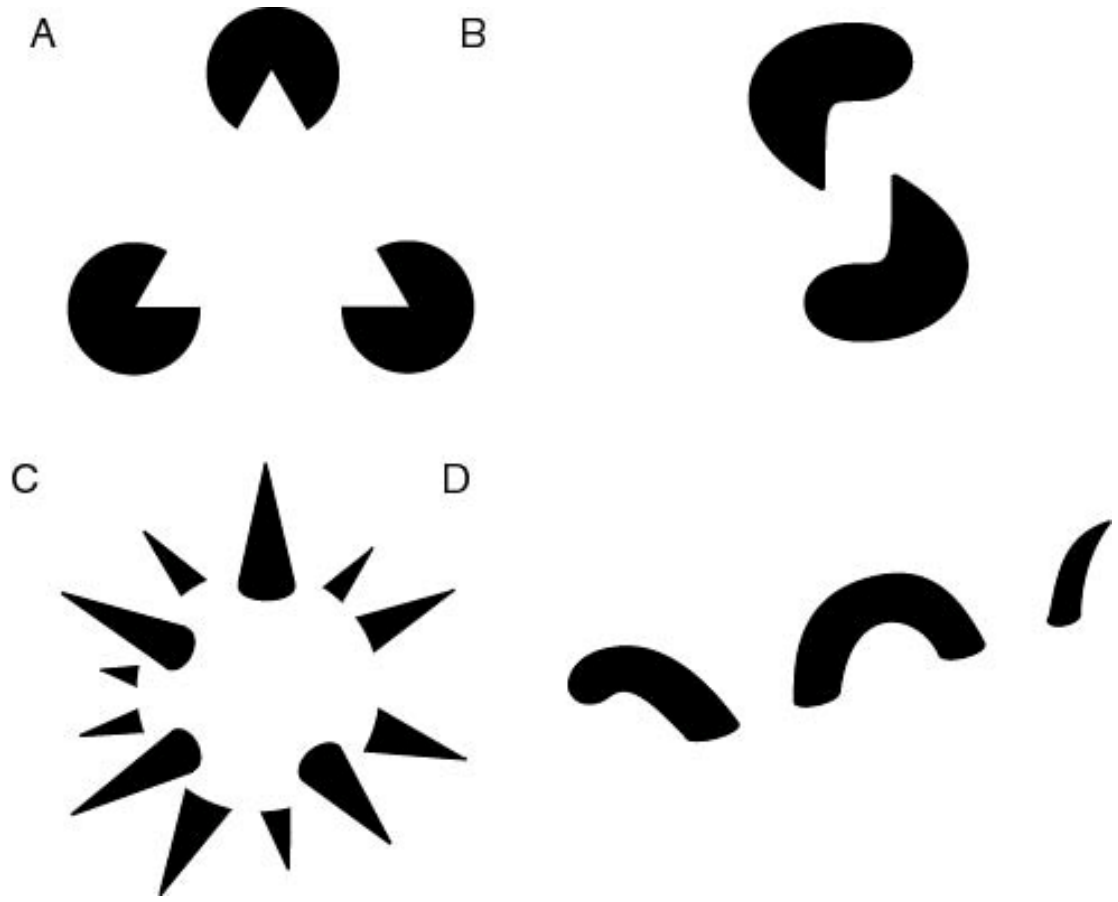
How many objects are there?



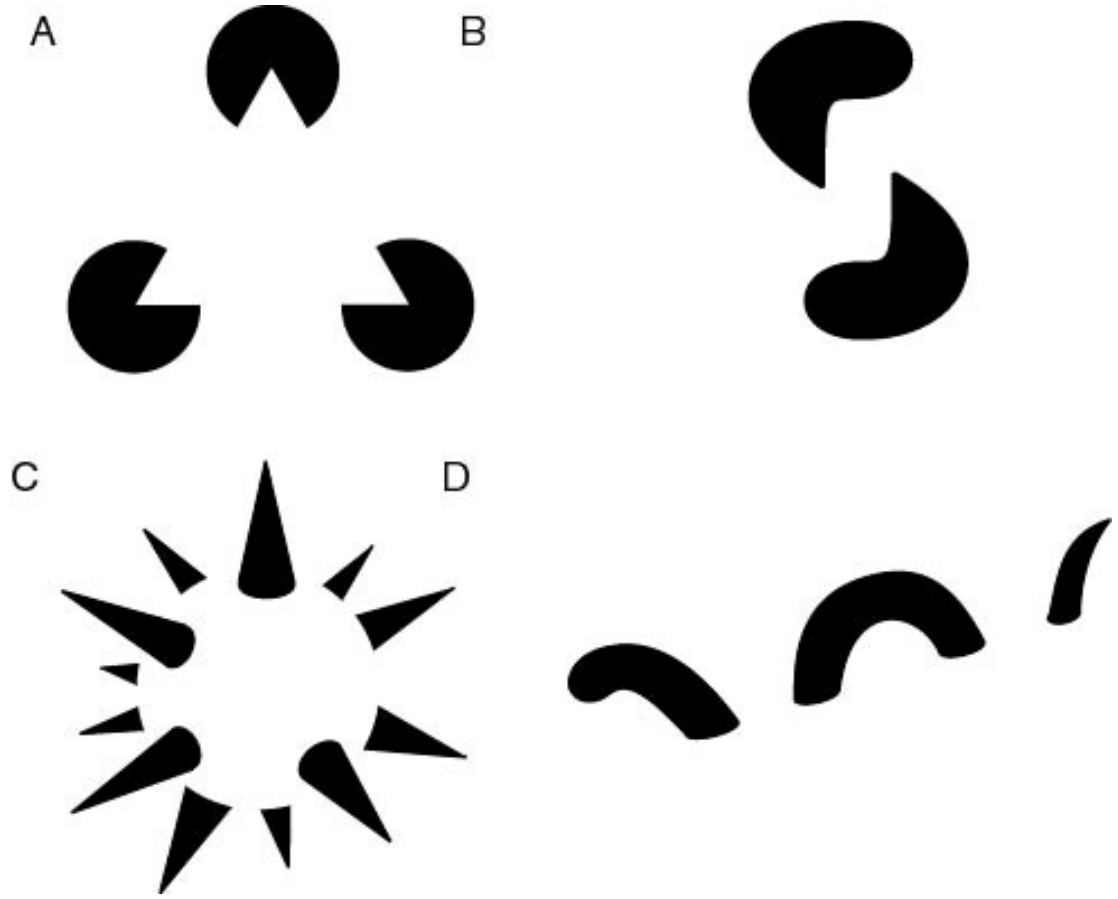
Continuity



How many objects are there?


A
B
C
D

Good Gestalt



What is this word?

CLIP

Past Experience

CLIP

Pre-Attentive Processing

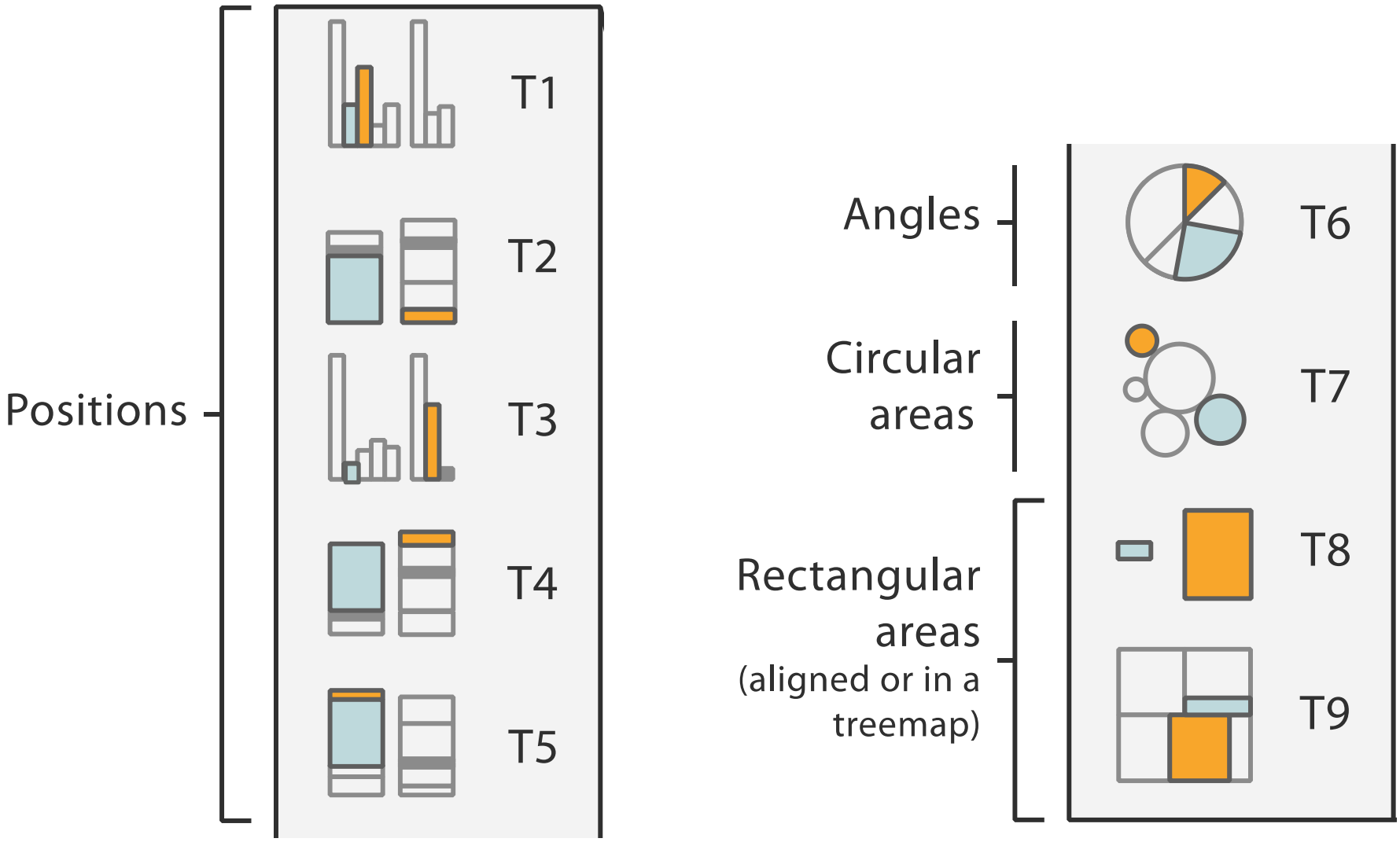
Gestalt Laws

Detect Quickly

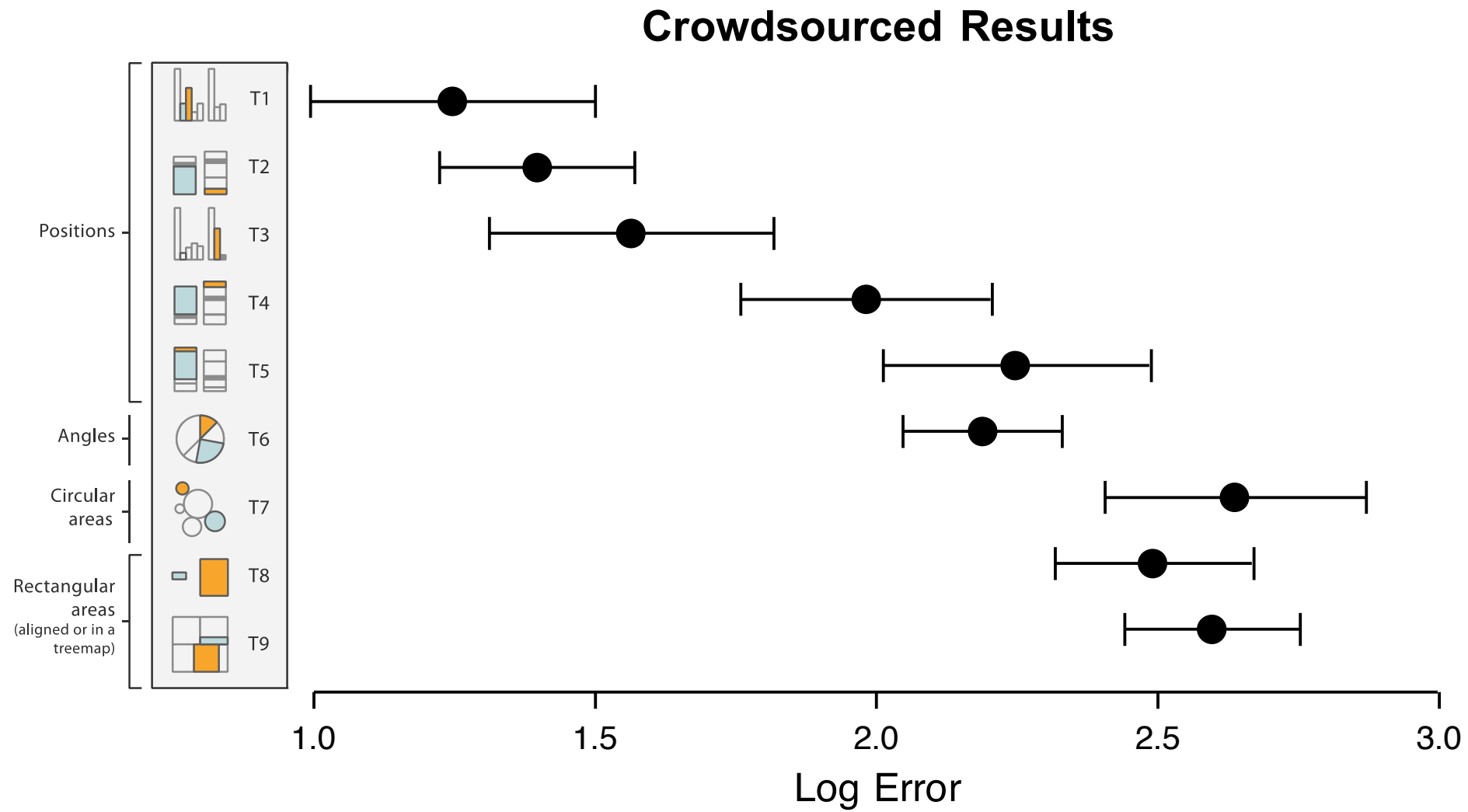
Detect quickly does NOT mean

detect accurately

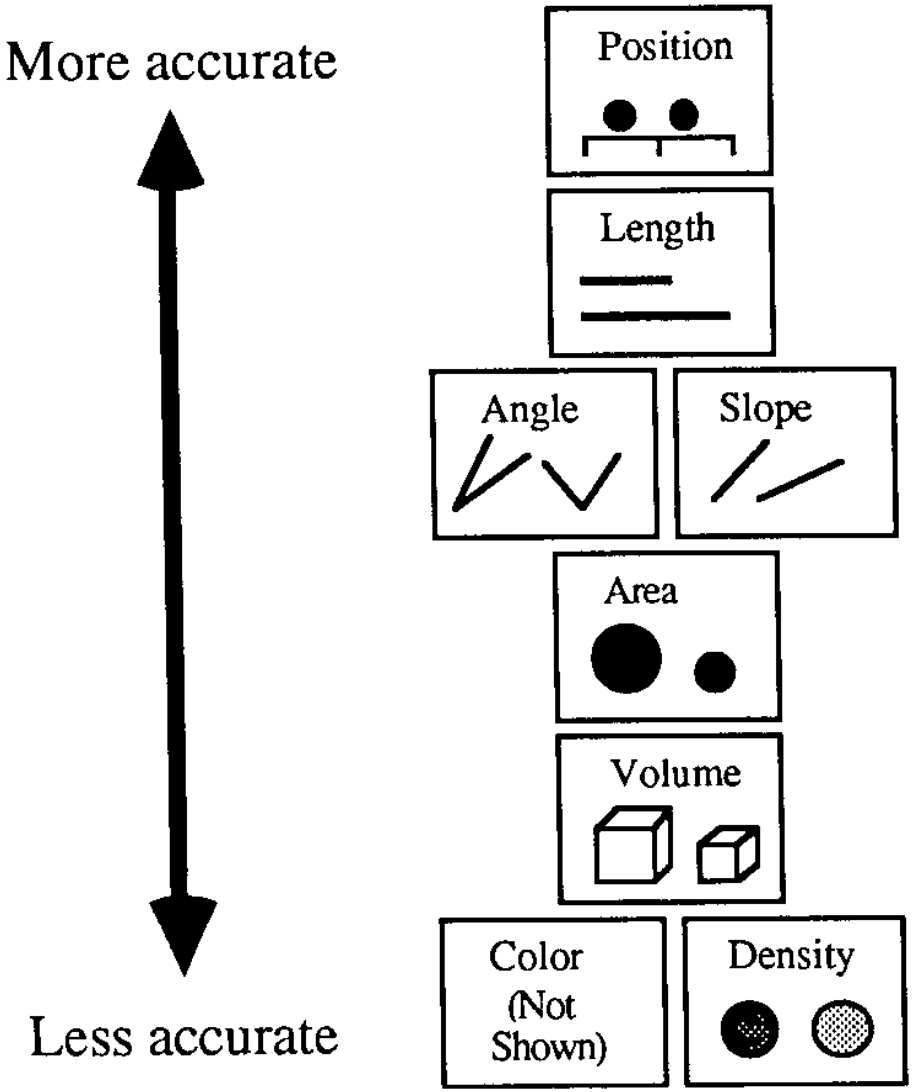
Ideally you want both.



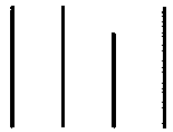
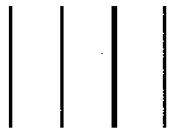
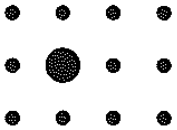
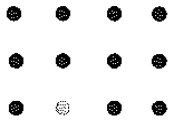
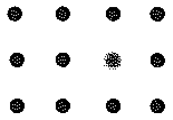
Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design.Heer and Bostock. Proc ACM Conf. Human Factors in Computing Systems (CHI) 2010, p. 203–212.



Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design.Heer and Bostock. Proc ACM Conf. Human Factors in Computing Systems (CHI) 2010, p. 203–212.



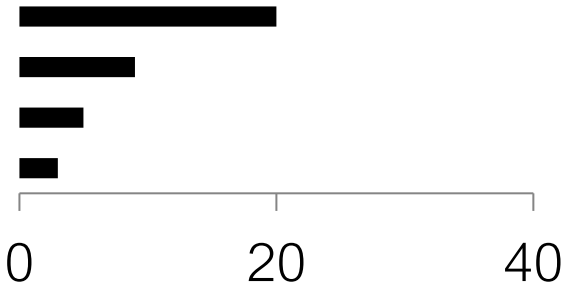
Mackinlay, 1986

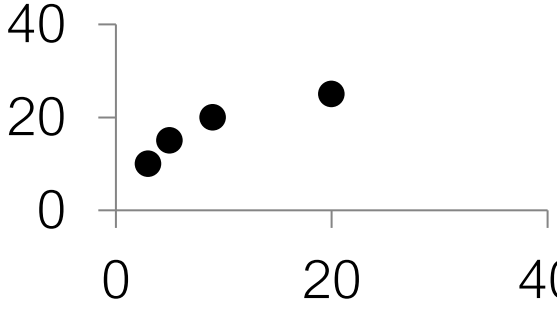
Precision of Quantitative Perception	Attribute	Example	Description
Very precise	Length		Longer = greater
	2-D Position		Higher or farther to the right = greater
Not very precise	Width		Wider = greater
	Size		Bigger = greater
	Intensity		Darker = greater
	Blur		Clearer = greater

Stephen Few
"Now You See It"
pg. 41

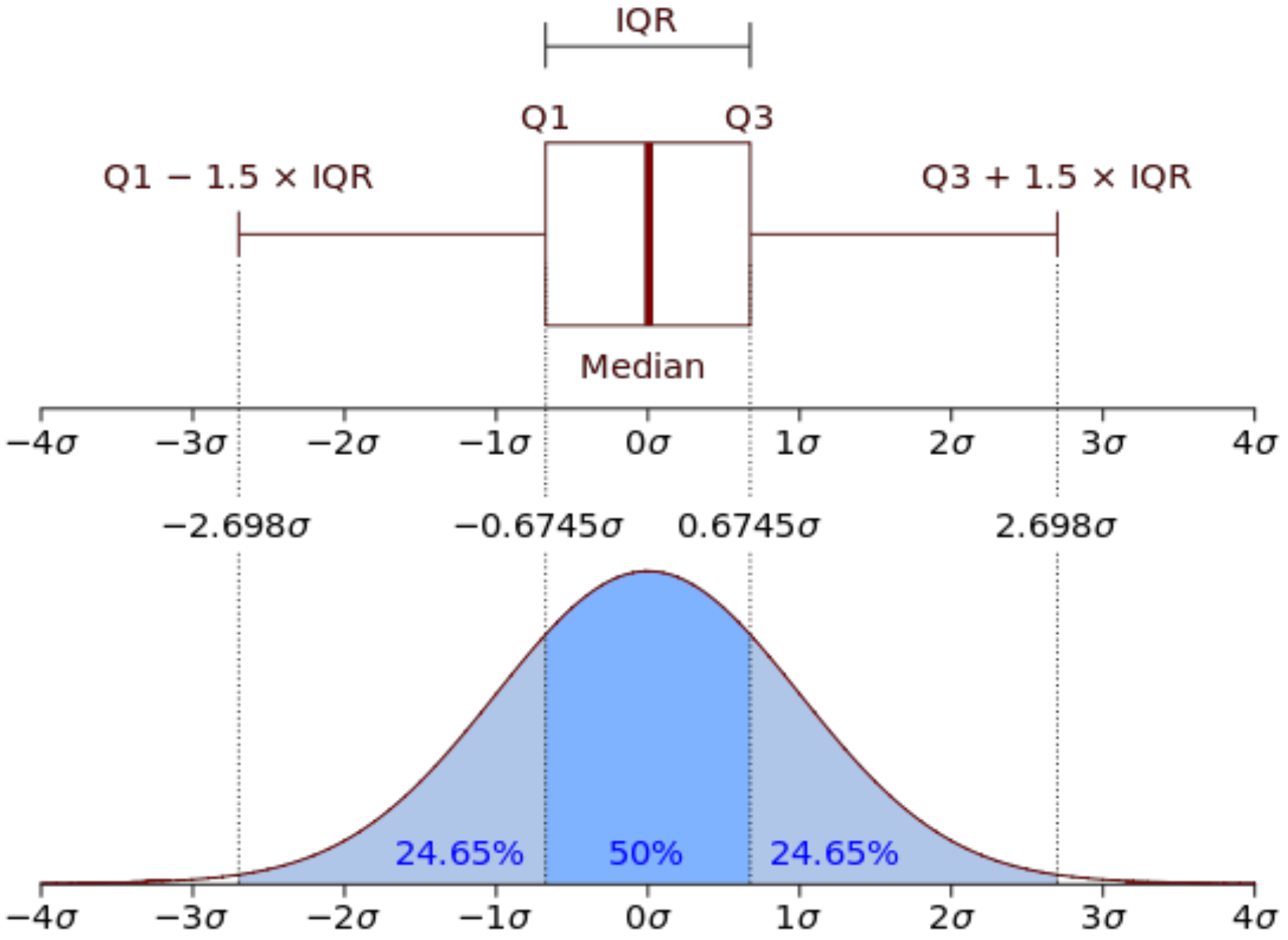
What does this tell us?

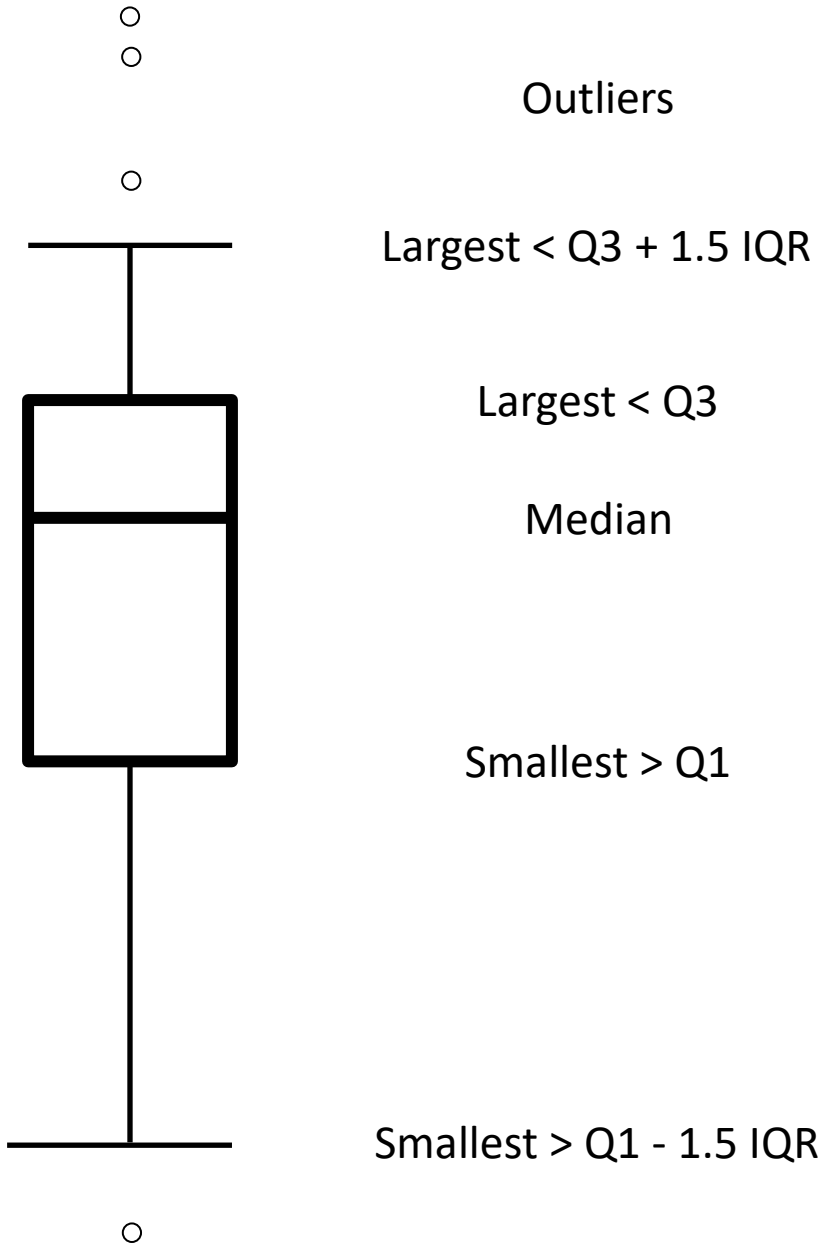
Barcharts, scatterplots, and line charts are *really* effective for quantitative data





(and for statistical distributions)
Tukey Box Plots


IQR
Q1
Q3
Q1 − 1.5 × IQR
Q3 + 1.5 × IQR
Median
−4σ
−3σ
−2σ
−1σ
0σ
1σ
2σ
3σ
4σ
−2.698σ
−0.6745σ
0.6745σ
2.698σ
24.65%
50%
24.65%


Outliers
Largest < Q3 + 1.5 IQR
Largest < Q3
Median
Smallest > Q1
Smallest > Q1 - 1.5 IQR

Tufte's Chart Principles

Edward Tufte

Tufte's Chart Principles

DO NOT LIE!

Tufte's Chart Principles

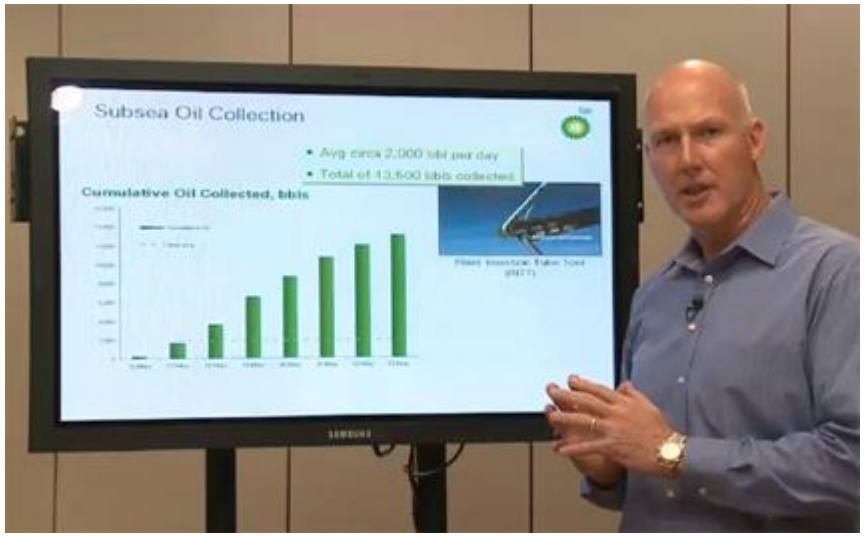
DO NOT LIE!

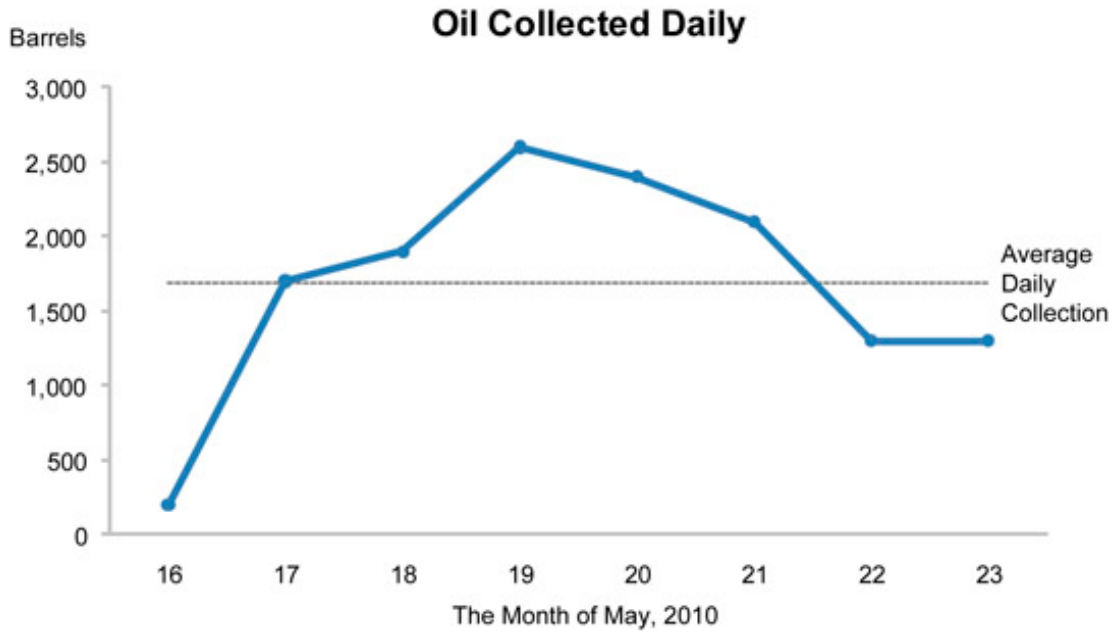
Maximize Data-Ink Ratio

Minimize Chart Junk


Subsea Oil Collection
• Avg circa 2,000 bbl per day
• Total of 13,500 bbls collected
Cumulative Oil Collected, bbls
SAMSUNG


“Cumulative”
• Avg circa 2,000 bbl per day
• Total of 13,500 bbls collected
Cumulative Oil Collected, bbls
SAMSUNG





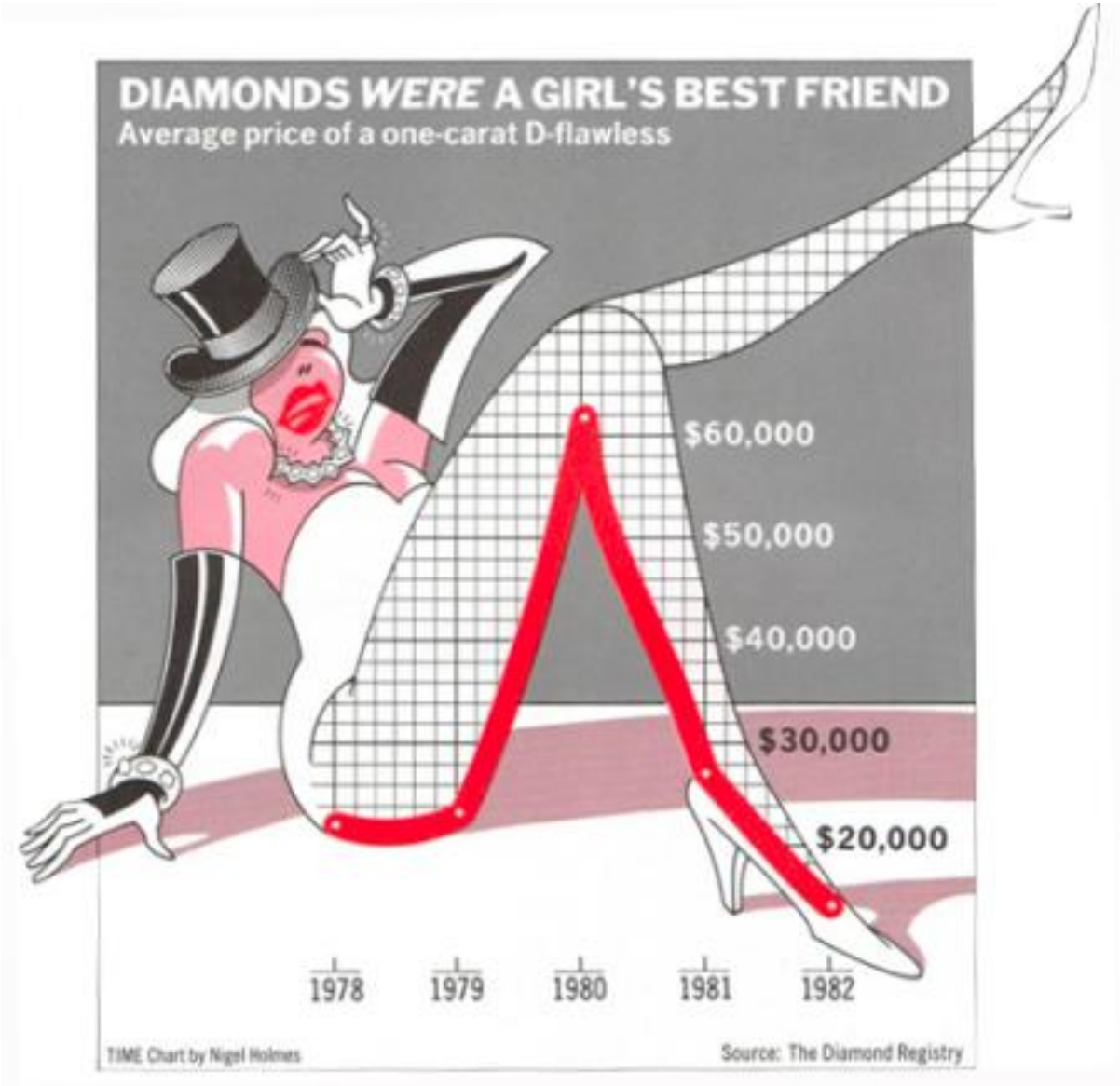
http://www.perceptualedge.com/blog/?p=790

Tufte's Chart Principles

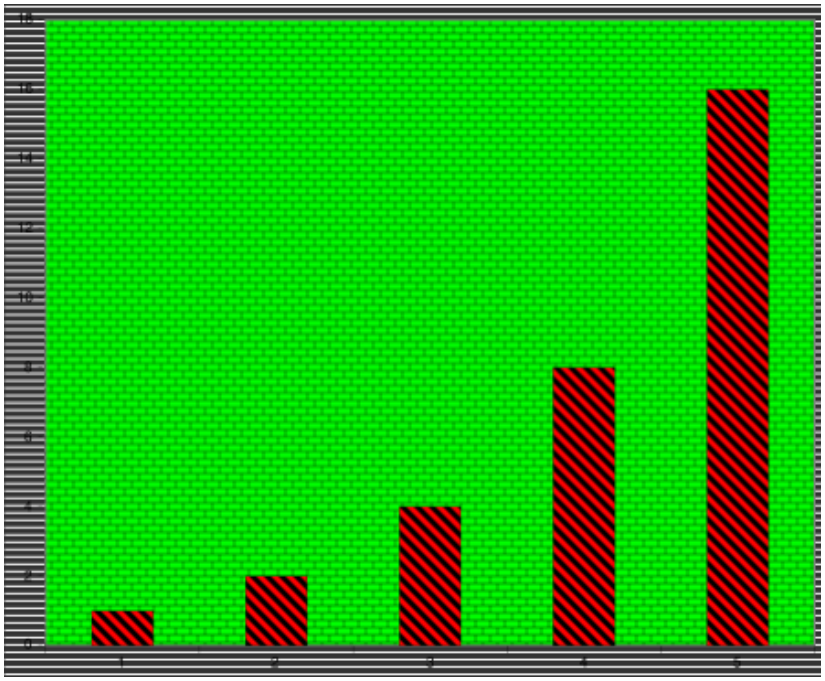
DO NOT LIE!

Maximize Data-Ink Ratio

Minimize Chart Junk



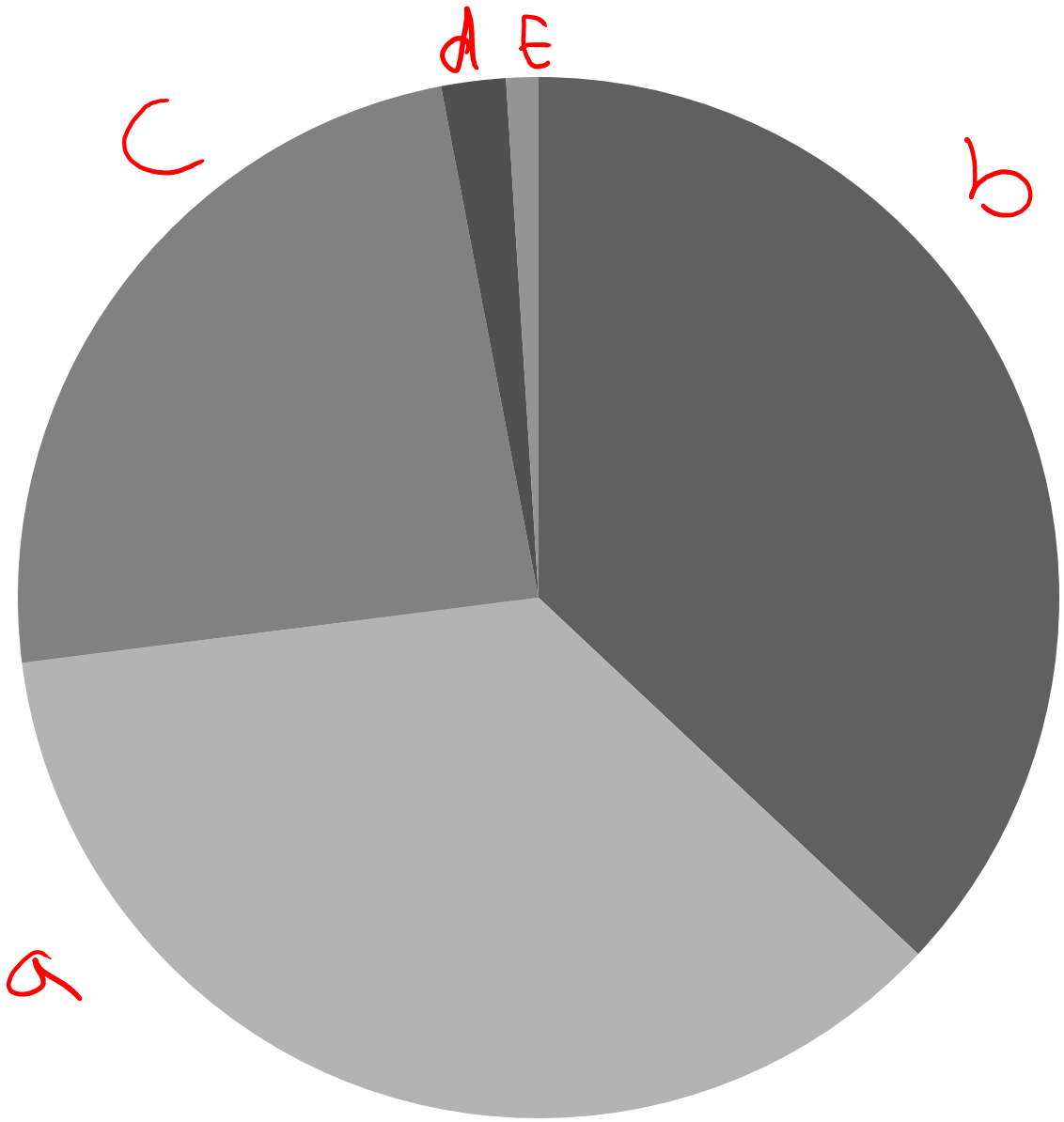
http://skilfulminds.com/2011/04/05/exploring-the-usefulness-of-chartjunk-at-stl-ux-2011/

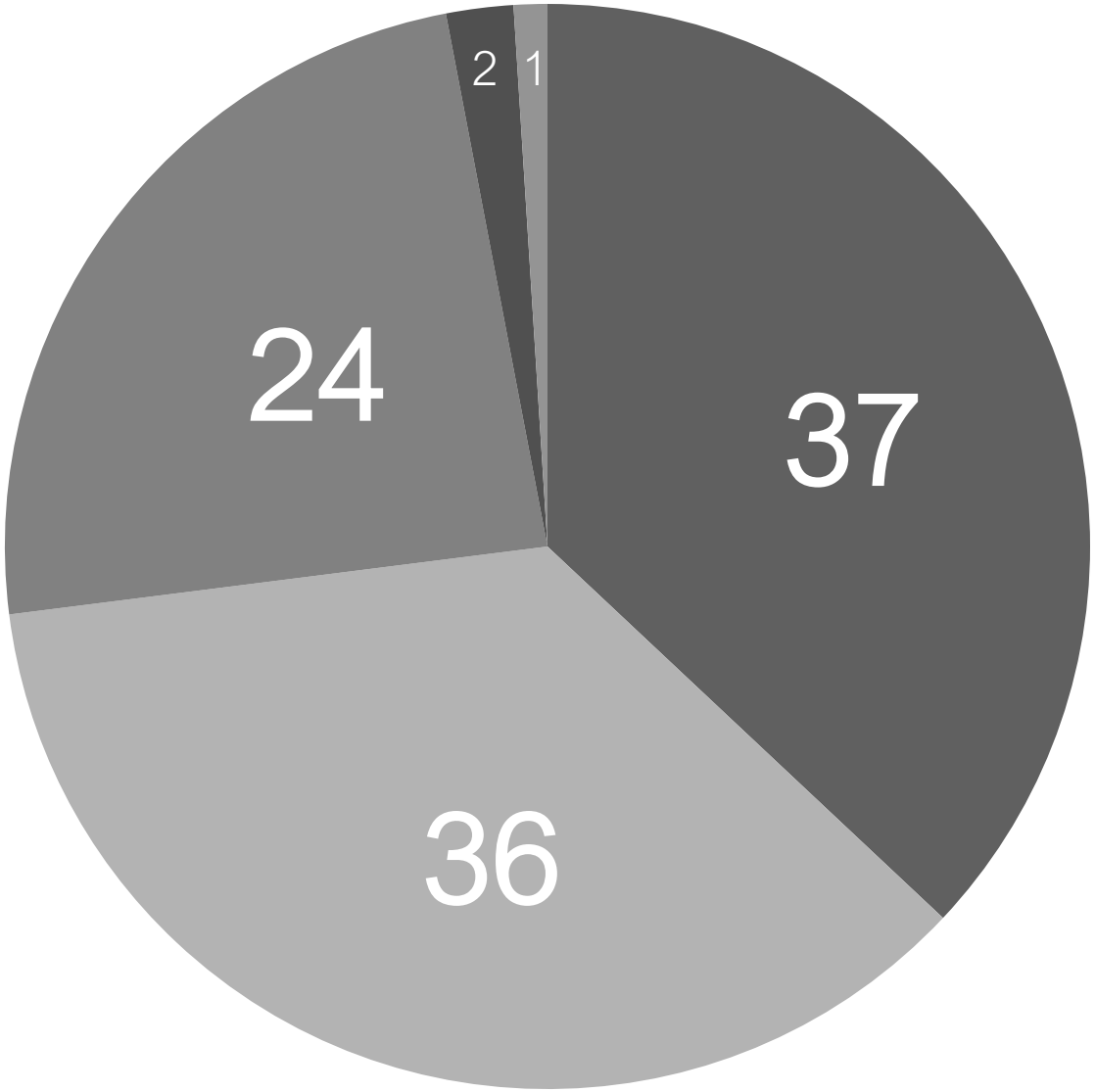


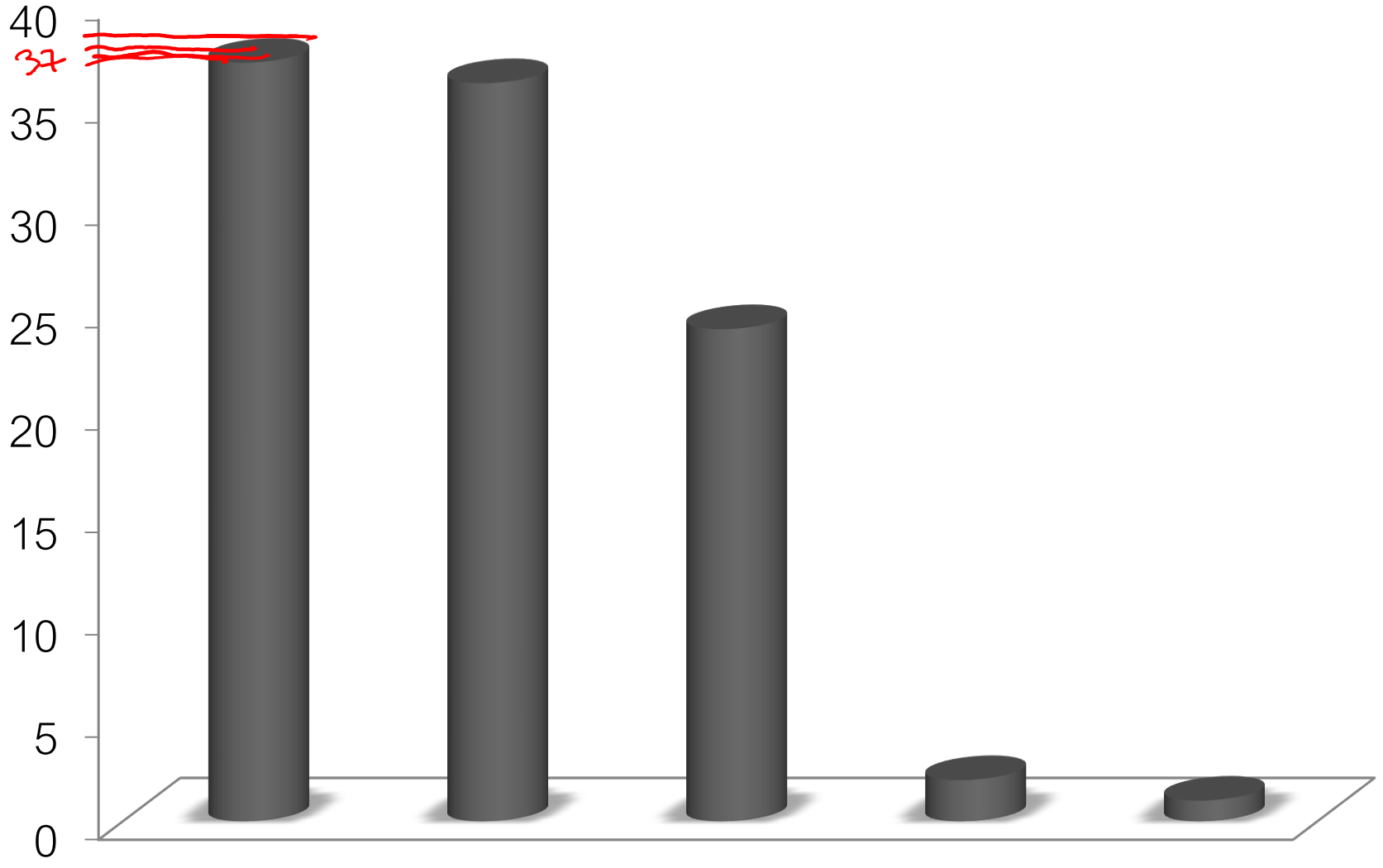
Chartjunk. (2017, October 05). Retrieved December 01, 2017, from https://en.wikipedia.org/wiki/Chartjunk

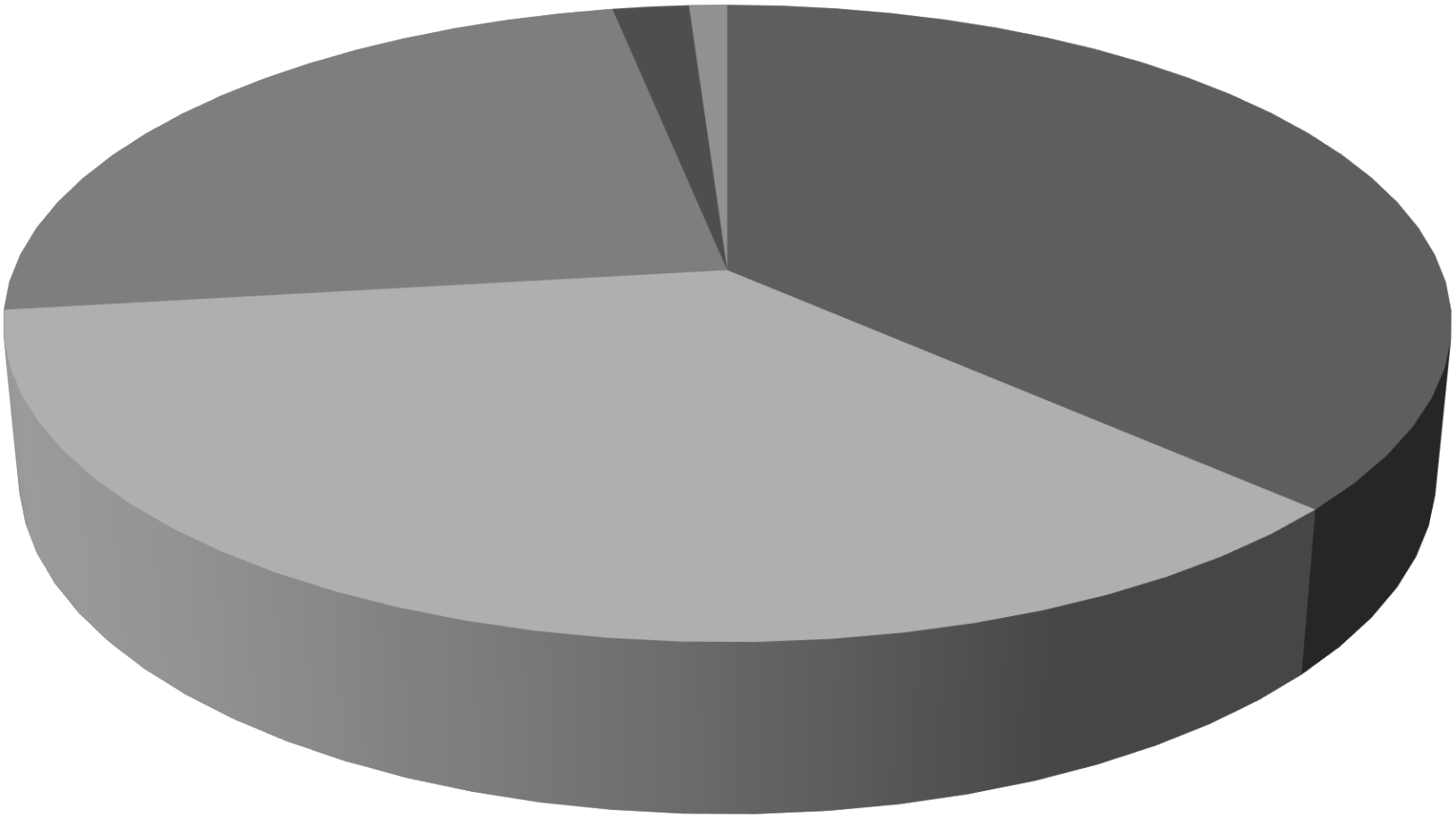
Please…

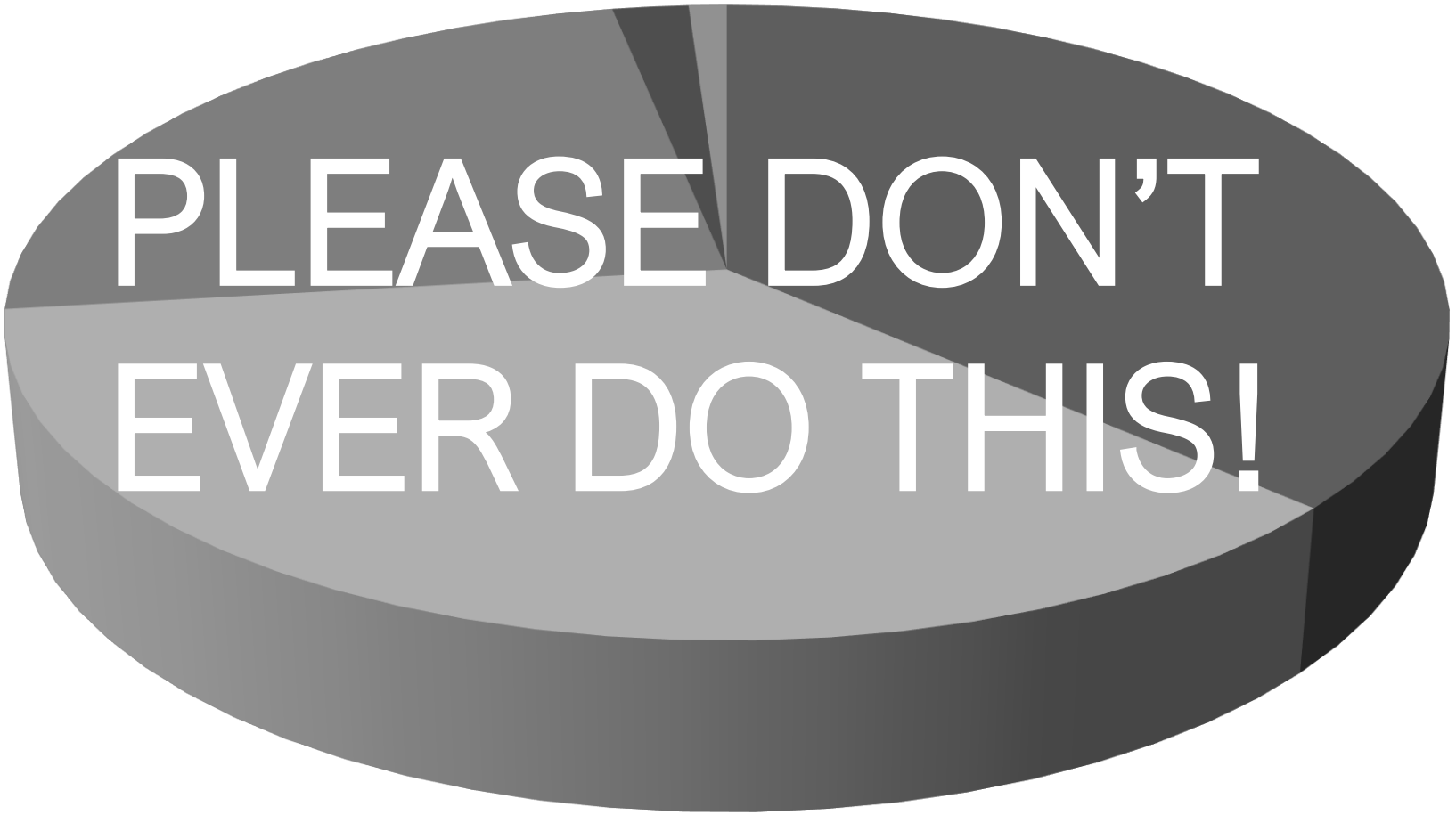
No pie charts.
No 2.5D charts.

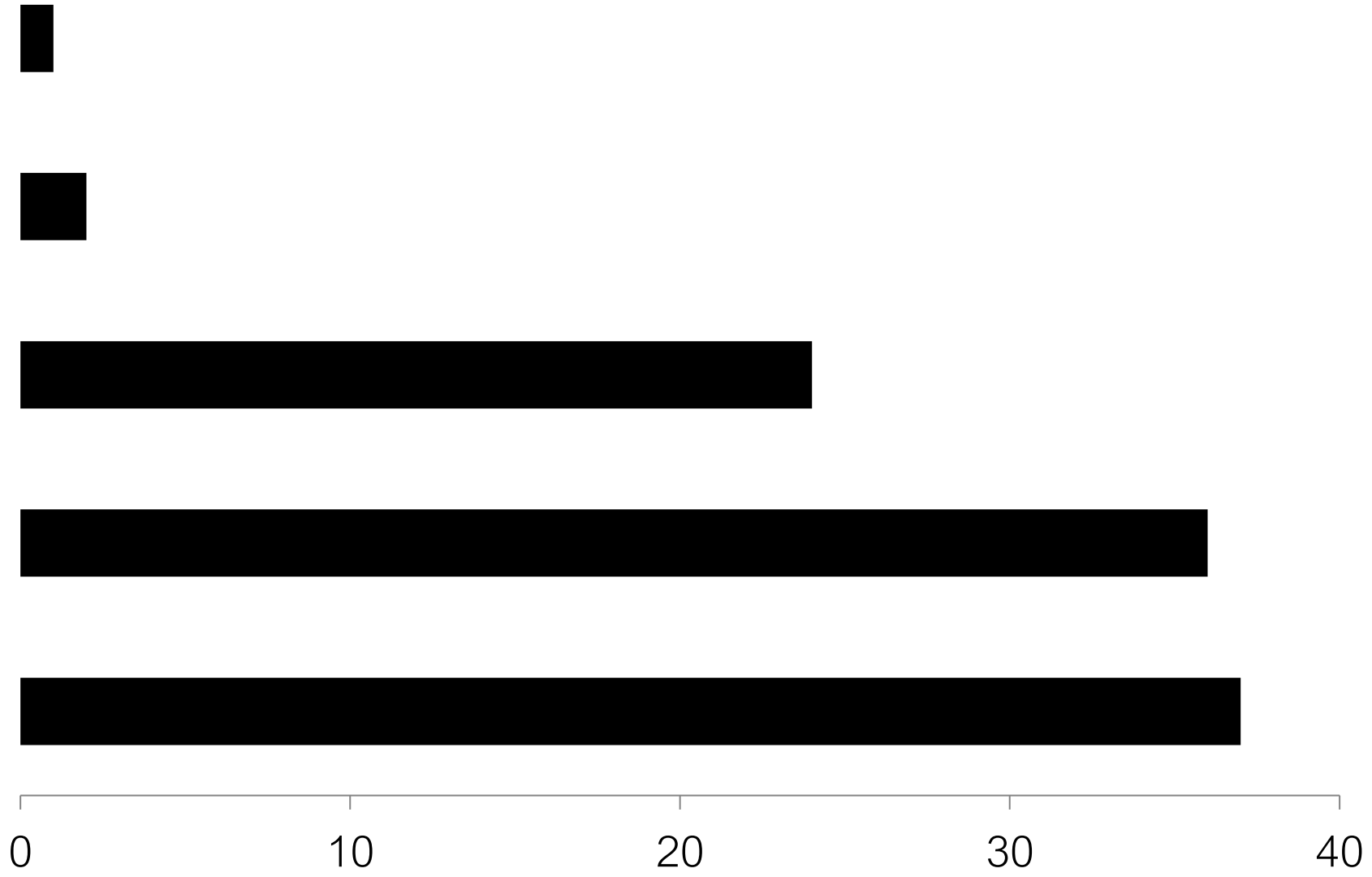

d
e
c
b
a


2
1
24
37
36


40
35
30
25
20
15
10
5
0

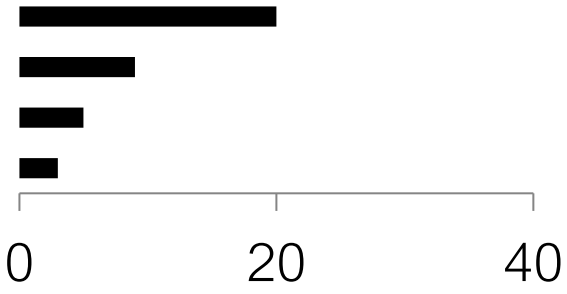


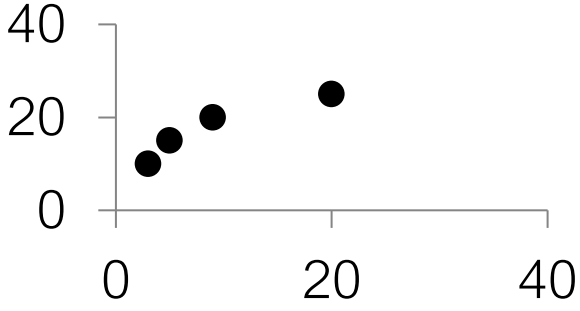

PLEASE DON'T
EVER DO THIS!

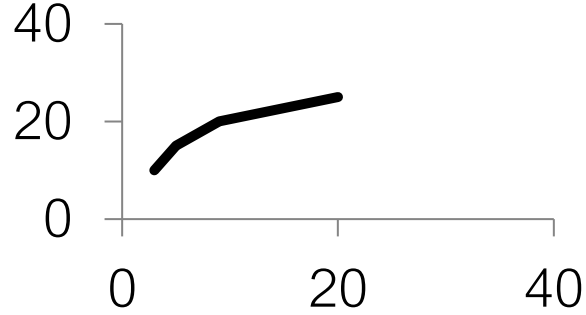

0
10
20
30
40

But otherwise…

Barcharts, scatterplots, and line charts are *really effective* for quantitative data







Anyone else bored
by my color choices?

In fact, grayscale can be risky…

In fact, grayscale can be risky…

Color is Powerful

Color

Call attention to information

Increase appeal

Increase memorability

Another dimension to work with

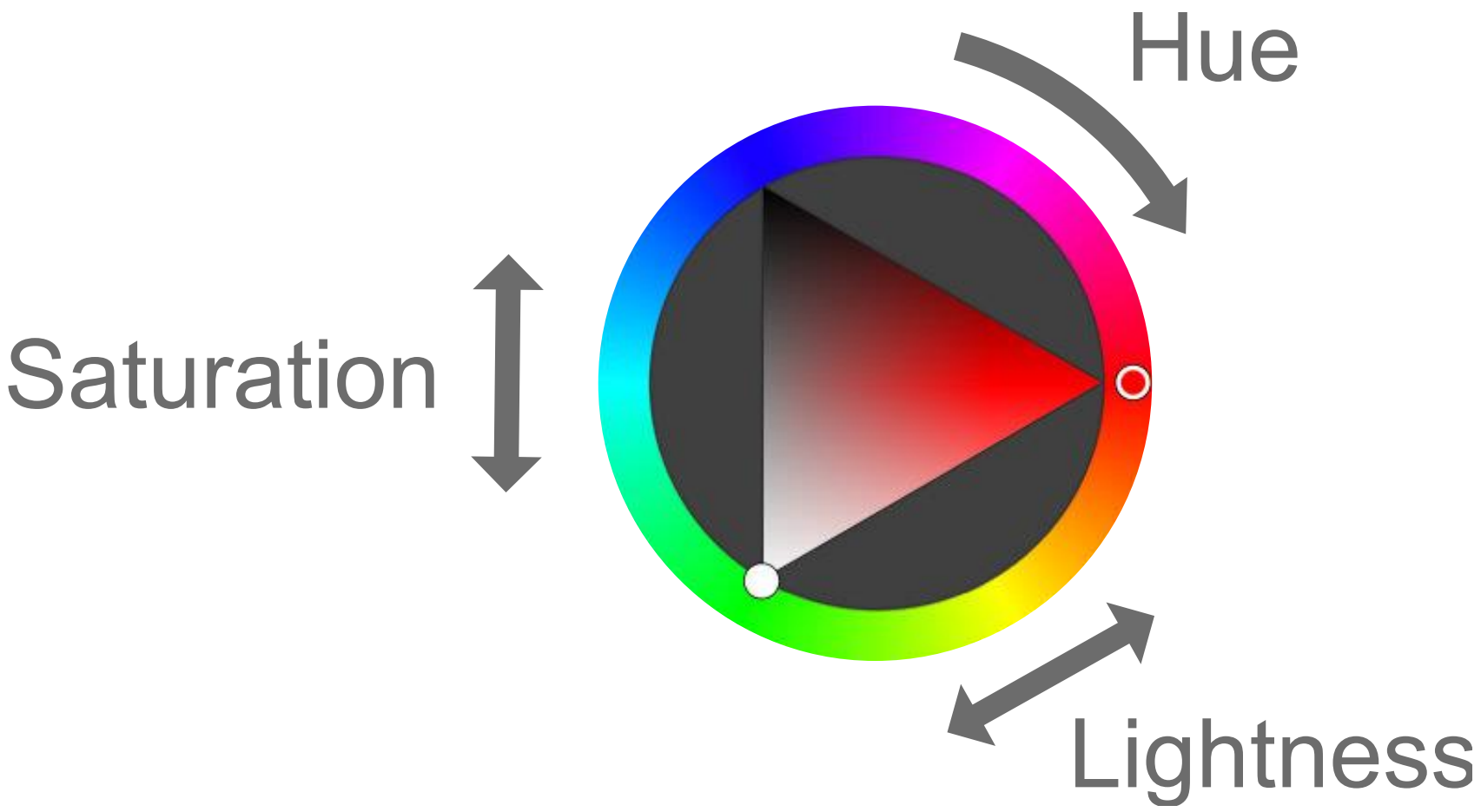
Have you heard of RGB?

Additive color model: colors create by mixing **red**, **green**, **blue** light

RGB color model. (2017, November 20). Retrieved December 01, 2017, from https://en.wikipedia.org/wiki/RGB_color_model

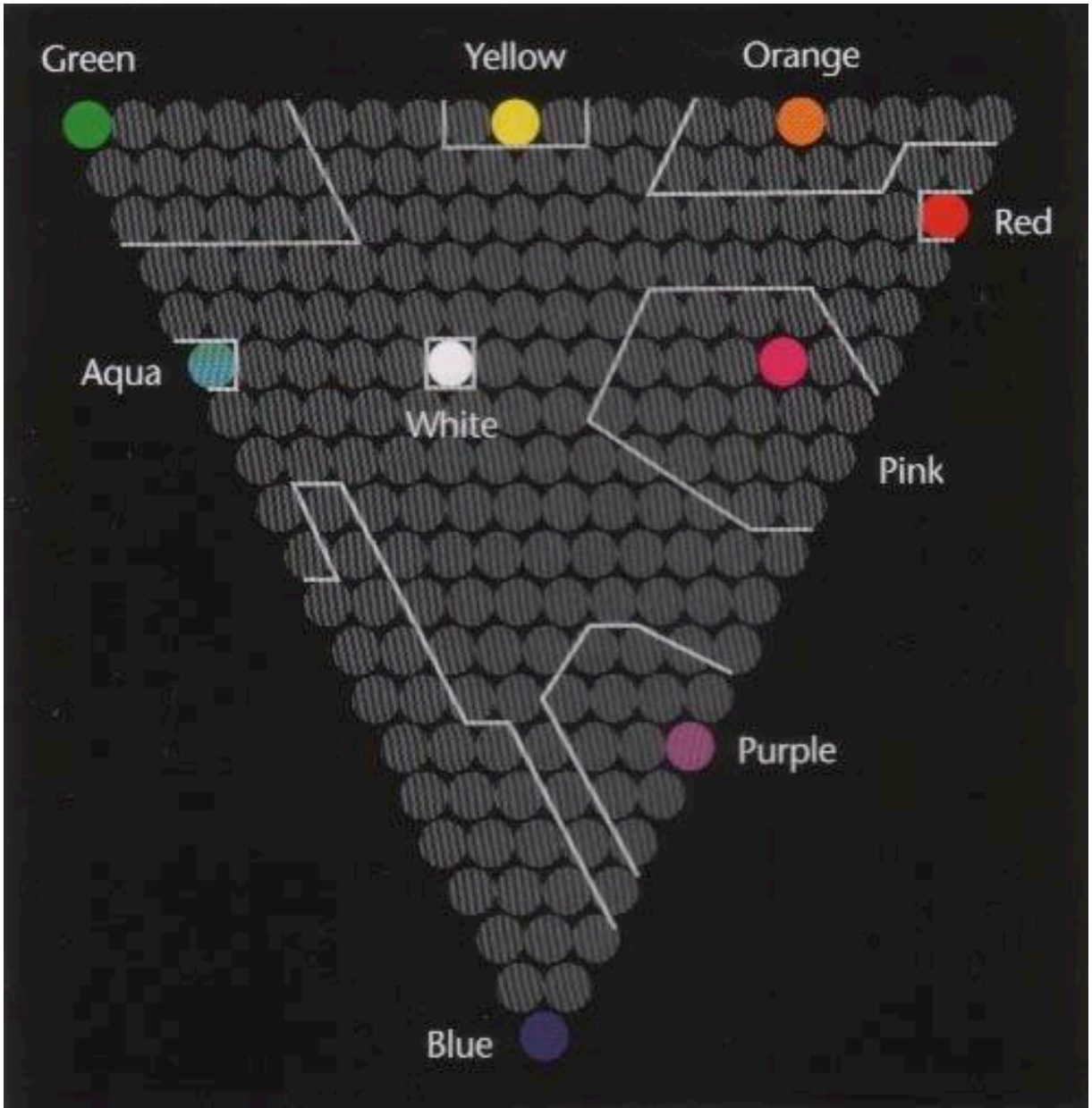
We see in RGB,
but we don't interpret in RGB…

HSV Color Model



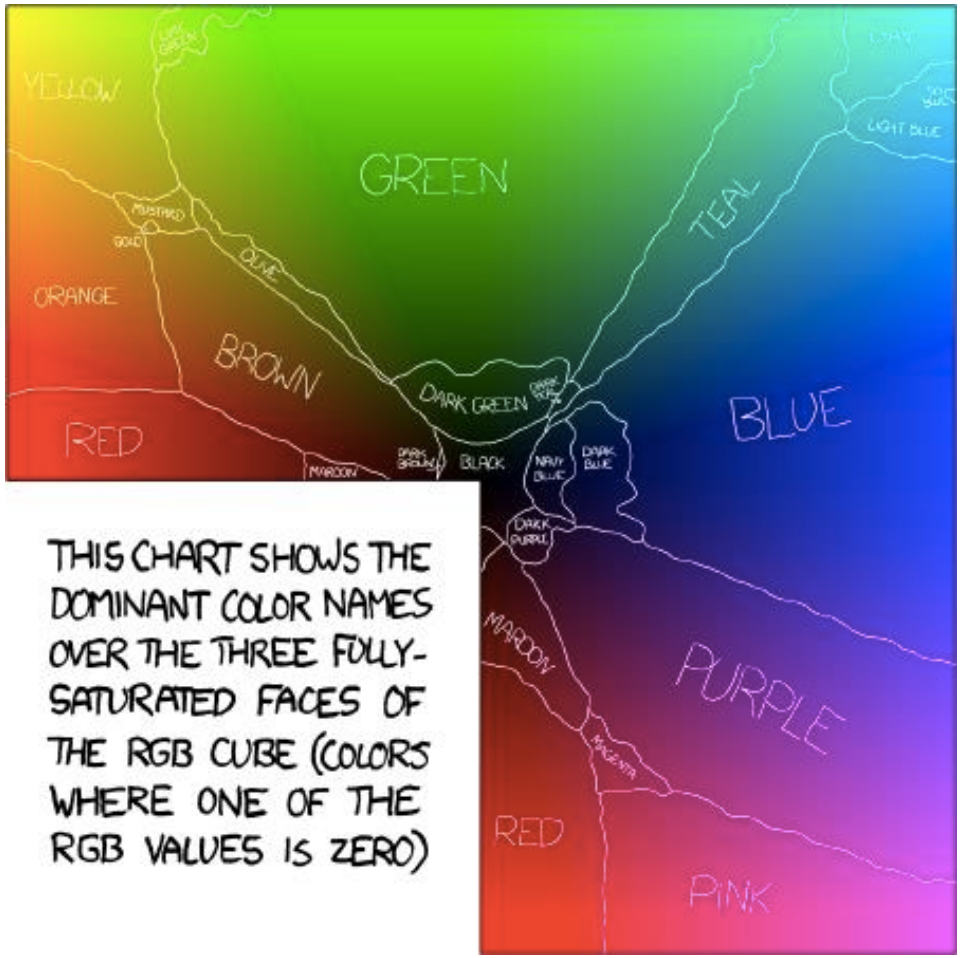
Source: color picker in Affinity Designer

Hue



Post & Greene, 1986

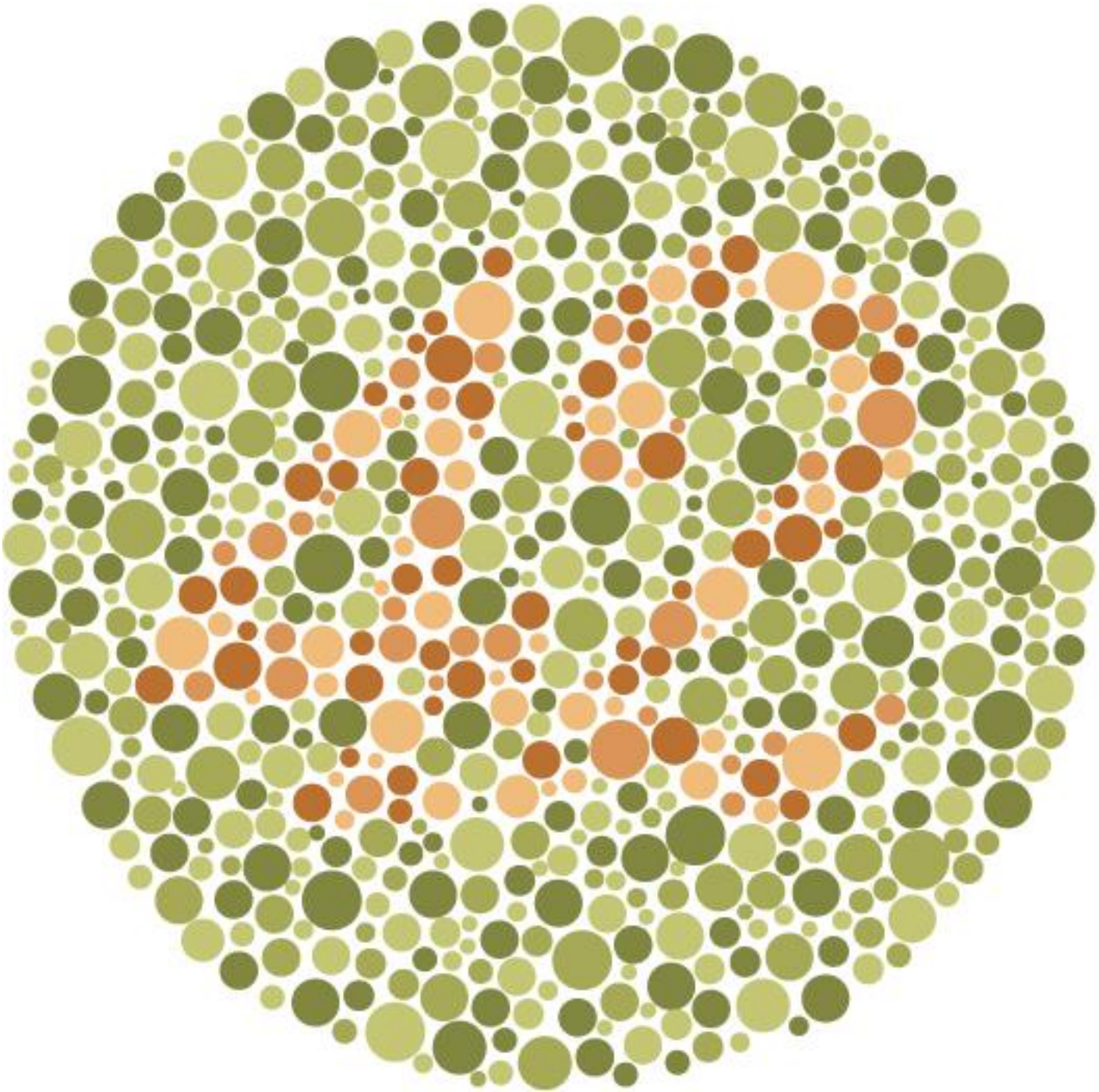
Hue

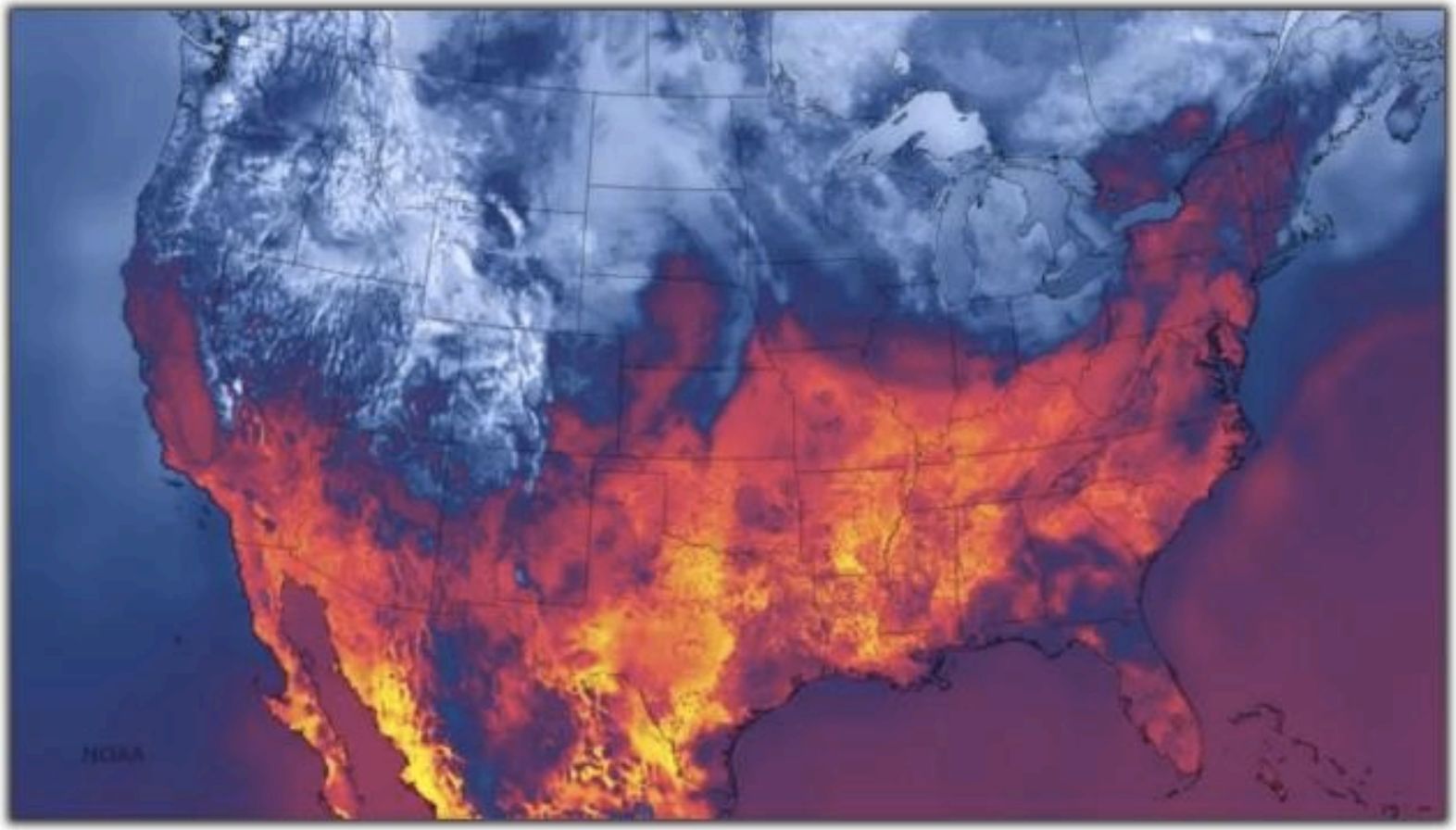


2010/05/03/color-survey-results/

Hue and Colorblindness

10% of males and 1% of females are Red-Green Colorblind

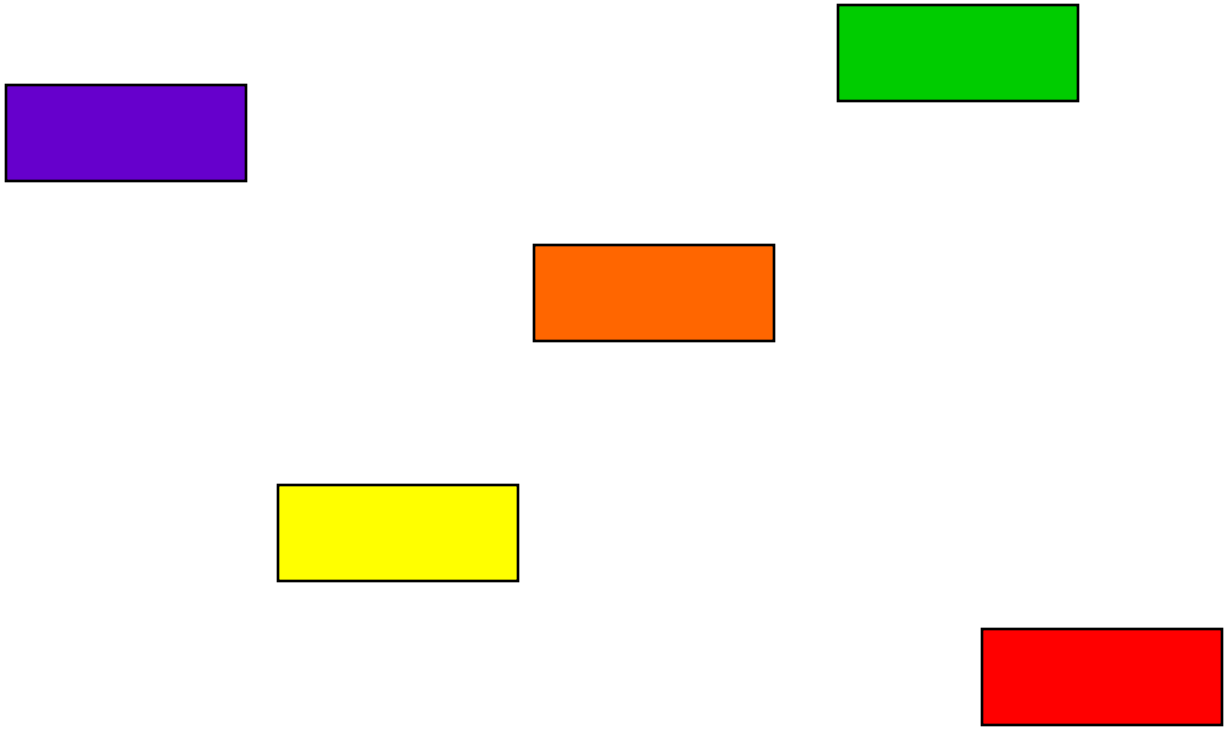


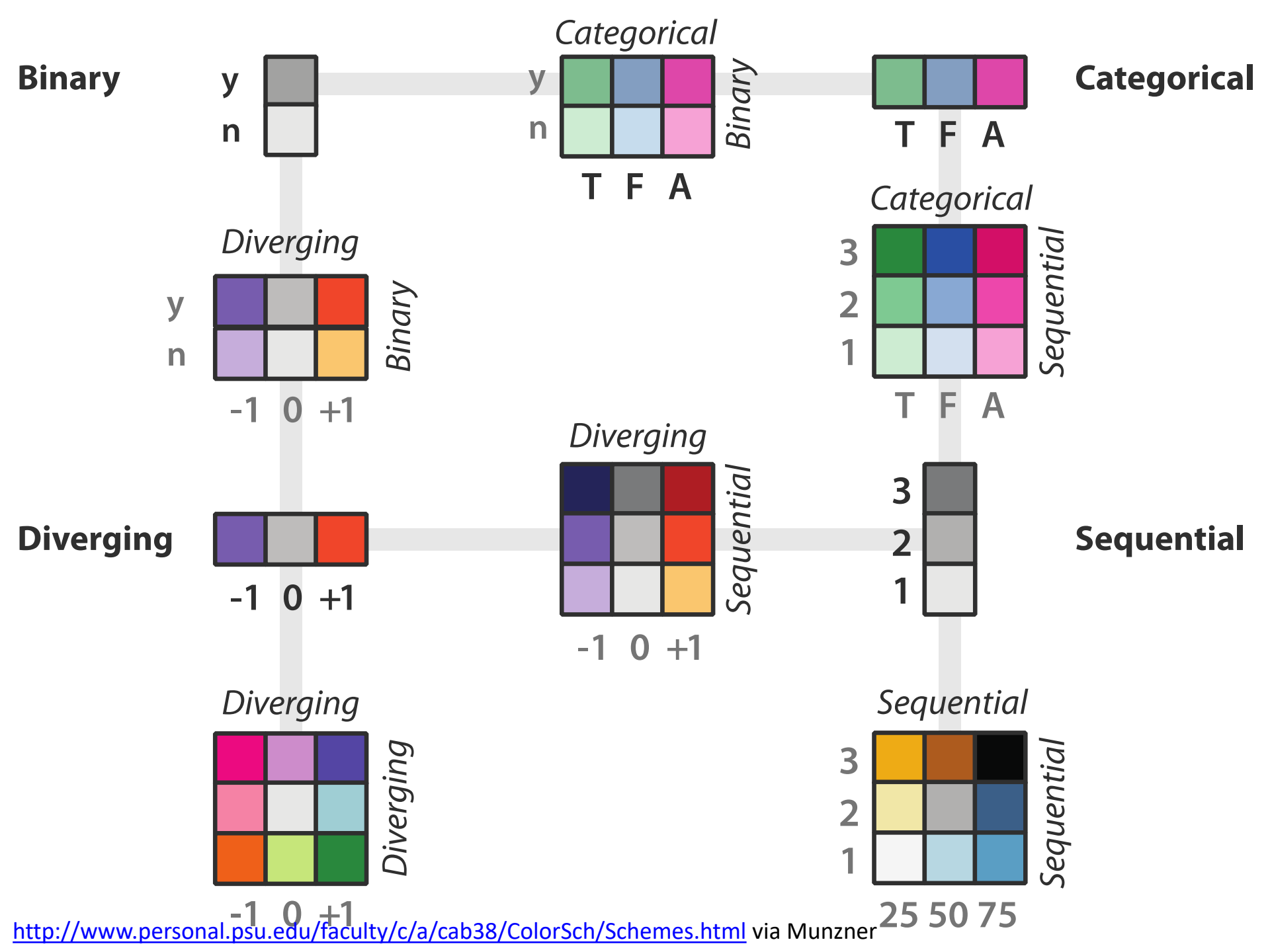


http://viz.wtf/post/98981561686/ht-matthewbgilmore-noaas-new-weather-modelling

Color and Quantitative Data

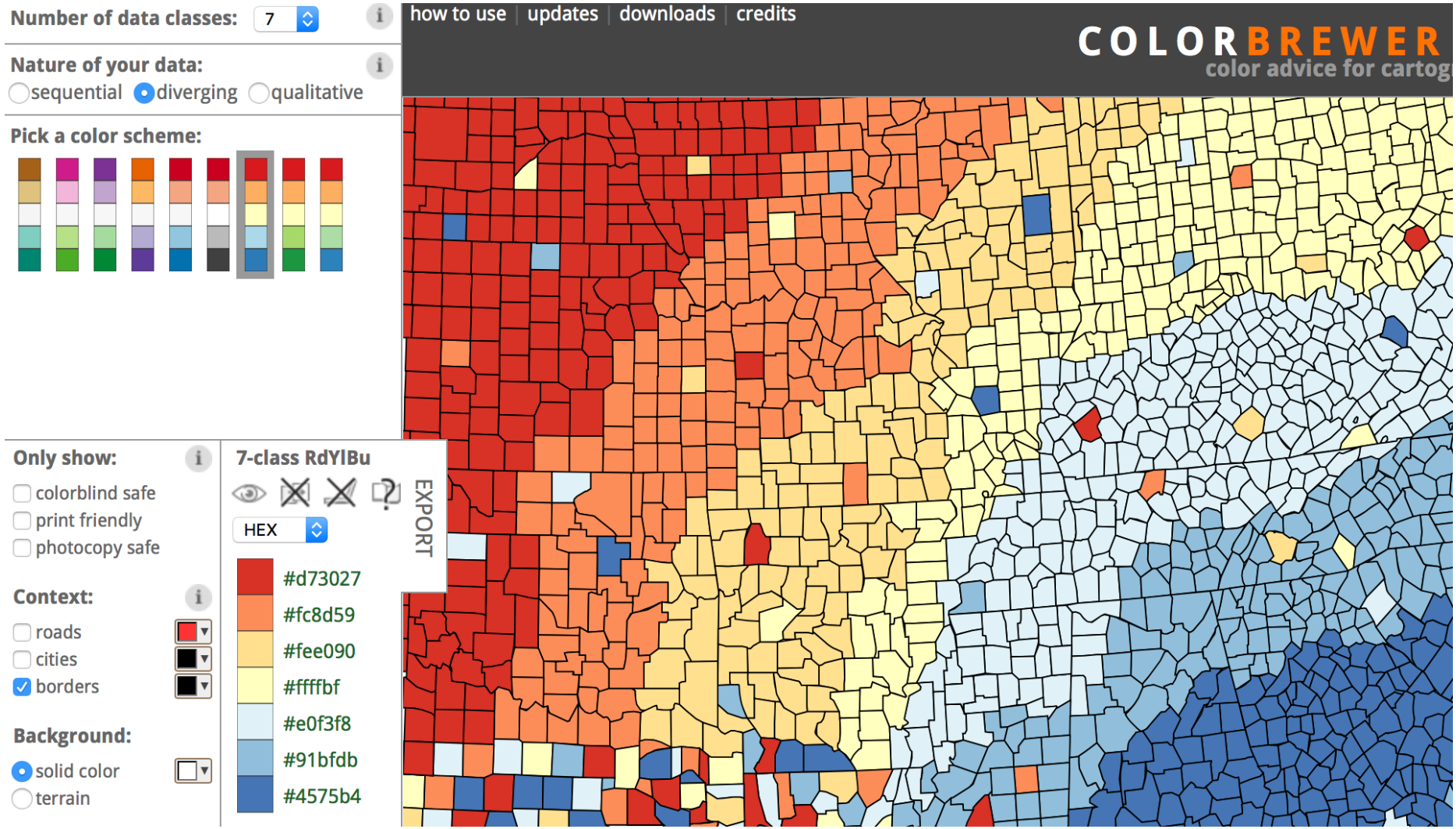
Can you order these (low→hi)?





http://www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html via Munzner

Color Brewer for Picking Color Scales



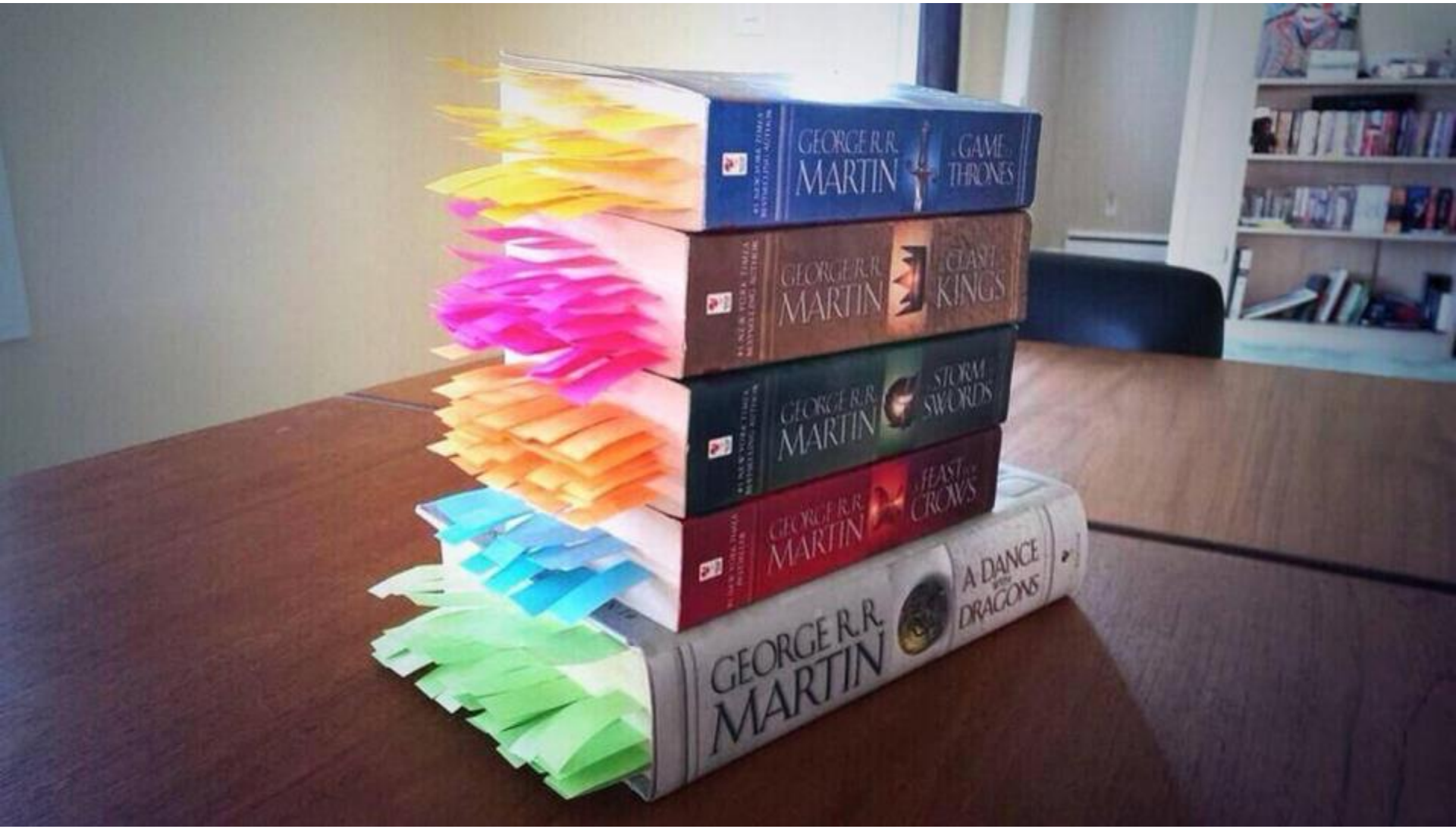
COLORBREWER 2.0. (n.d.). Retrieved December 01, 2017, from http://colorbrewer2.org/

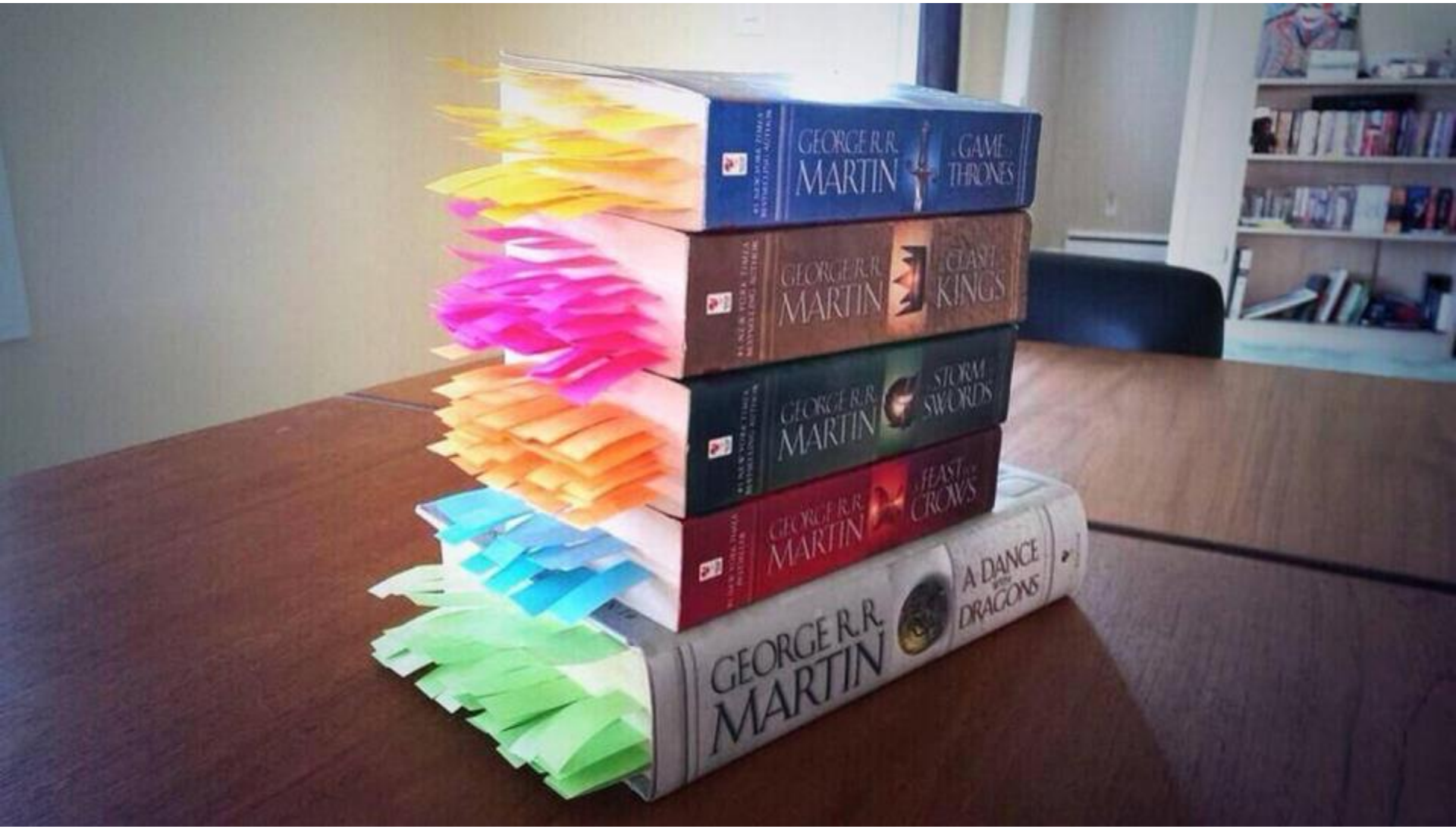
Overview
Zoom+Filter
Details on Demand

Shneiderman Mantra

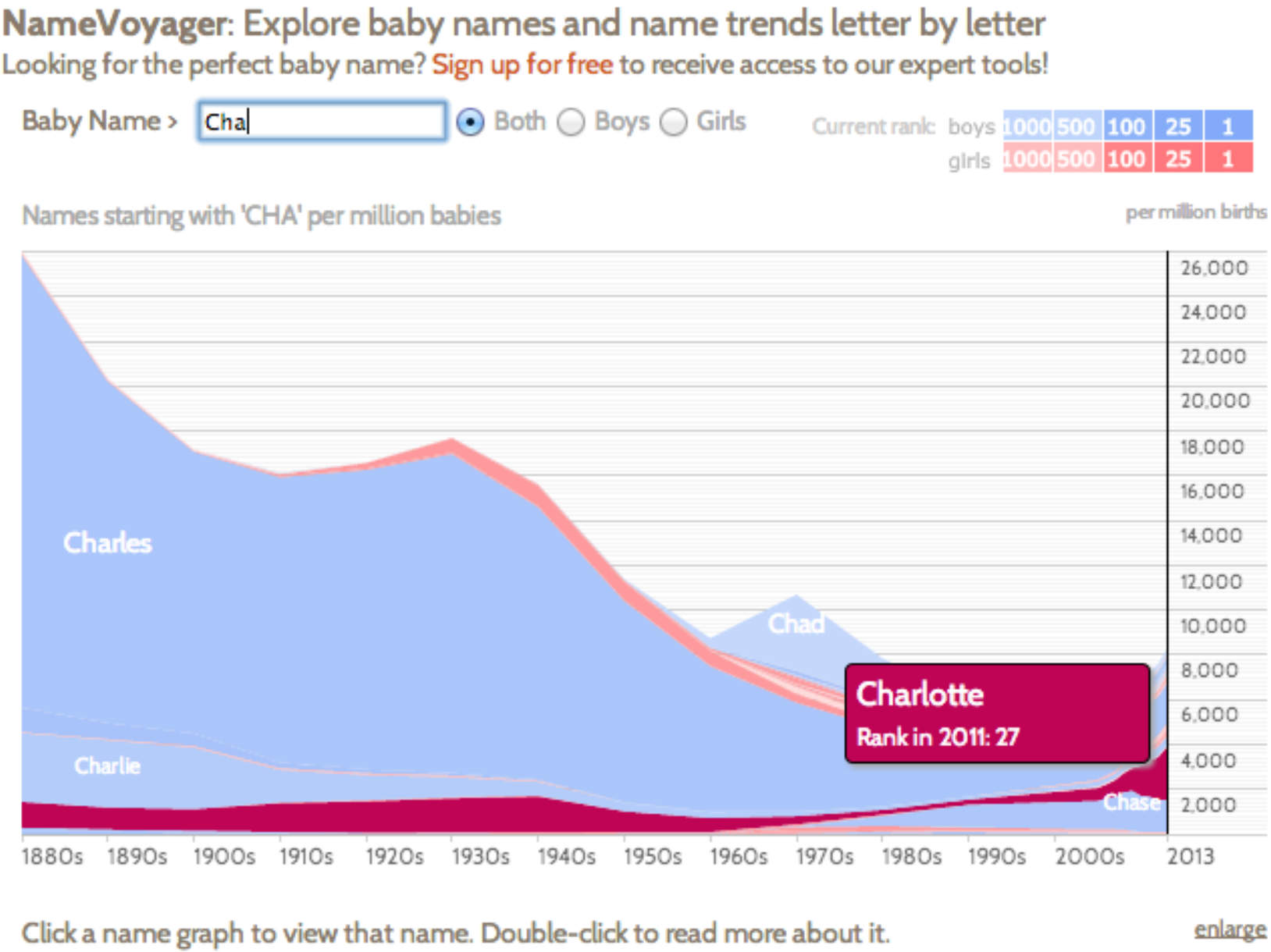
(Information-Seeking Mantra)

https://www.mat.ucsb.edu/g.legrady/academic/courses/11w259/schneiderman.pdf


GEORGE R.R. MARTIN
A GAME OF THRONES
GEORGE R.R. MARTIN
A CLASH OF KINGS
GEORGE R.R. MARTIN
A STORM OF SWORDS
GEORGE R.R. MARTIN
A FEAST FOR CROWS
GEORGE R.R. MARTIN
A DANCE WITH DRAGONS



http://visual.ly/every-single-death-game-thrones-series



http://www.babynamewizard.com/voyager

Where to learn more?

CS 7450
Information Visualization
Every Fall

Visualization @GeorgiaTech

vis.gatech.edu

Georgia Tech Visualization Lab

How to Make Good Charts

- Edward Tufte's One-Day Workshop
  - http://www.edwardtufte.com/tufte/courses
- Edward Tufte, *Visual Display of Quantitative Information*
  - http://www.edwardtufte.com/tufte/books_vdqi
- Stephen Few, *Show Me the Numbers: Designing Tables and Graphs to Enlighten*
  - http://www.amazon.com/Show-Me-Numbers-Designing-Enlighten/dp/0970601972/ref=la_B001H6IQ5M_1_2?s=books&ie=UTF8&qid=1385050724&sr=1-2

Visualization Theory "Books"

- Tamara Munzner VIS Tutorial and Book
  - http://www.cs.ubc.ca/~tmm/talks.html
  - http://www.cs.ubc.ca/~tmm/vadbook/
- Colin Ware, *Information Visualization: Perception for Design*
  - http://www.amazon.com/Information-Visualization-Perception-Interactive-Technologies/dp/1558605118
- Stephen Few, *Now You See It*
  - http://www.amazon.com/Now-You-See-Visualization-Quantitative/dp/0970601980/ref=pd_bxgy_b_img_z
- Edward Tufte, *Envisioning Information*
  - http://www.edwardtufte.com/tufte/books_ei
- Edward Tufte, *Visual Explanations*
  - http://www.edwardtufte.com/tufte/books_visex
- Edward Tufte, *Beautiful Evidence*
  - http://www.edwardtufte.com/tufte/books_be
- Tamara Munzner, Visualization Analysis & Design
  - http://www.amazon.com/Visualization-Analysis-Design-AK-Peters/dp/1466508914

Perception and Color Websites

- Chris Healy, NC State
  - http://www.csc.ncsu.edu/faculty/healey/PP/index.html
- Color Brewer
  - http://colorbrewer2.org/
- Maureen C. Stone (Color Links, Blog, Workshops)
  - http://www.stonesc.com/color/index.htm
- Subtleties of Color by Robert Simmon of NASA
  - http://blog.visual.ly/subtleties-of-color/

Visualization Blogs

- Flowing Data by Nathan Yau
  - http://flowingdata.com/
- Information Aesthetics by Andrew Vande Moere
  - http://infosthetics.com/
- Information is Beautiful by David McCandless
  - http://www.informationisbeautiful.net/
- Visual.ly Blog
  - http://blog.visual.ly/
- Indexed Comic by Jessica Hagy
  - http://thisisindexed.com/

Infographics

Visual.ly/view

(wtfviz.net)