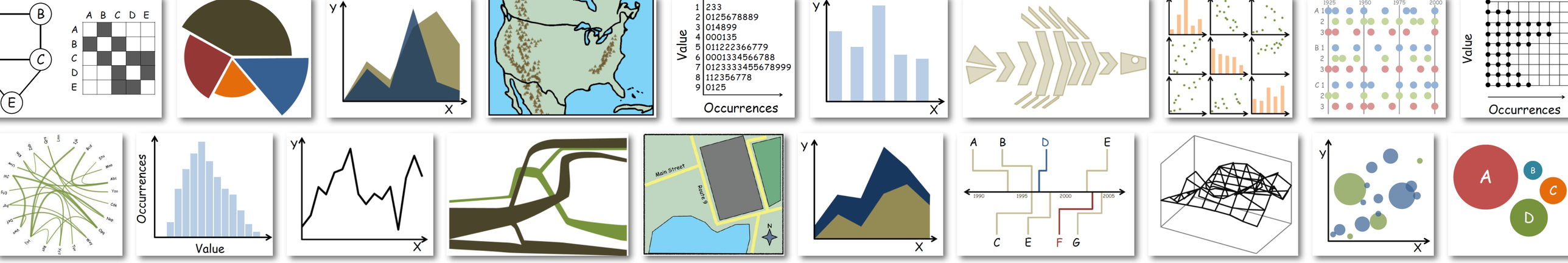
Cody Dunne
Northeastern University

Design Rules of Thumb

Feel free to interrupt with questions!

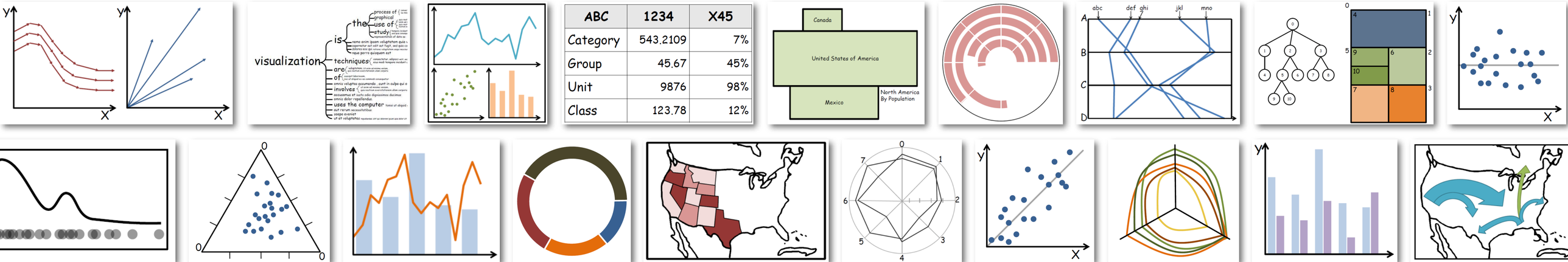
Checking In

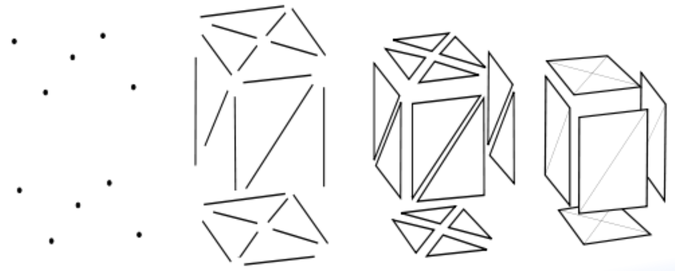
PREVIOUSLY, ON DS 4200…



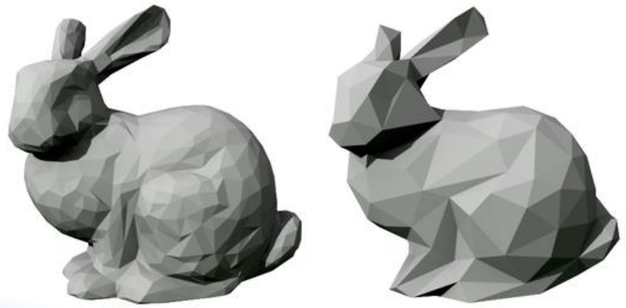
(static or interactive) (abstract or spatial)

visualization: the visual representation of data to reinforce human cognition





computer graphics



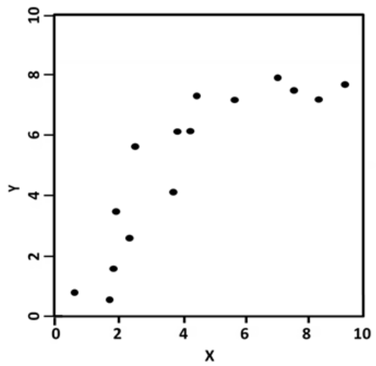
HCI

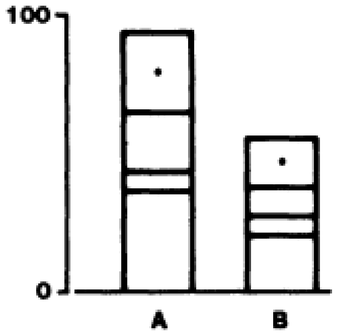
design

visualization

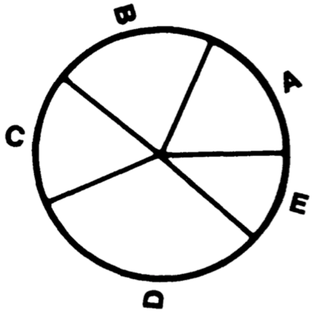
psychology

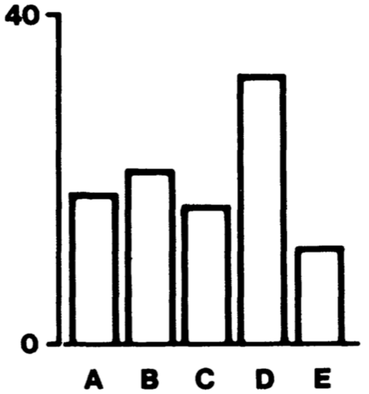
art





statistics





No catalogue of techniques can convey a willingness to look for what can be seen, whether or not anticipated. Yet this is at the heart of exploratory data analysis. ... the picture-examining eye is the best finder we have of the wholly unanticipated.

– Tukey, 1980

“change blindness”

https://www.youtube.com/watch?v=FWSxSQsspiQ

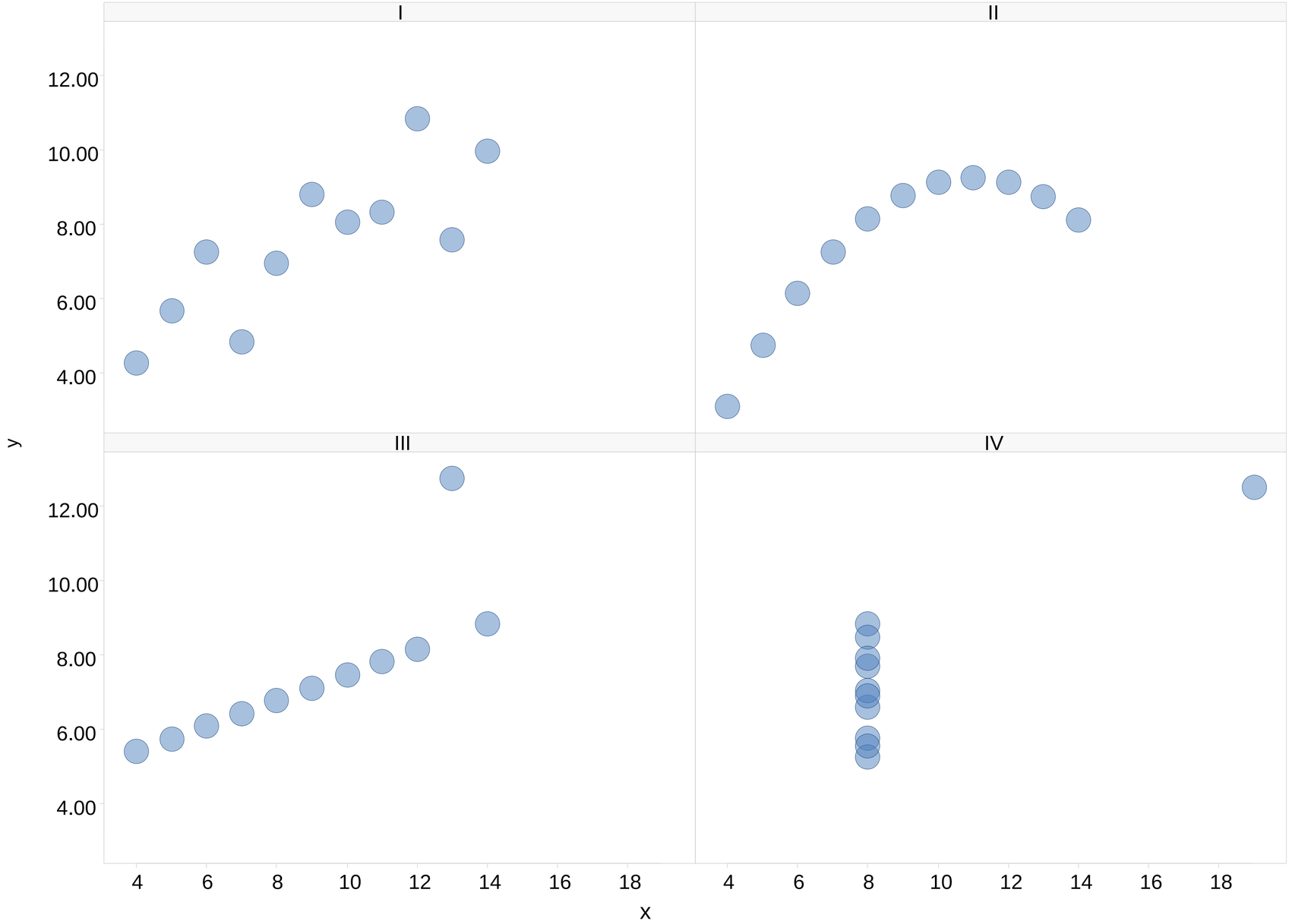
Plan for Today

- Discuss more on what visualization is & why we care
- Discuss some basic design rules of thumb

Again, why do we need visualization?

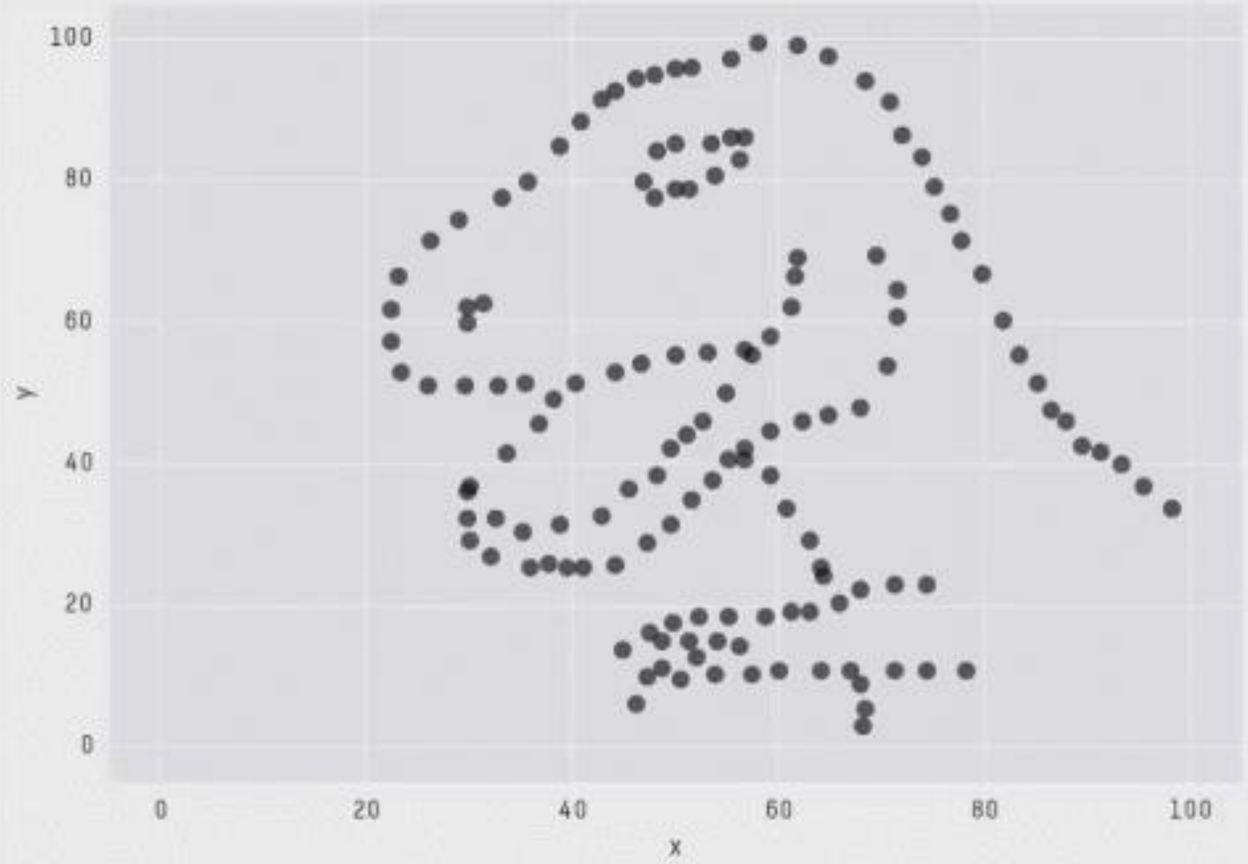
I		II		III		IV	
x	y	x	y	x	y	x	y
10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
6.00	7.24	6.00	6.13	6.00	6.08	8.00	5.25
4.00	4.26	4.00	3.10	4.00	5.39	19.00	12.50
12.00	10.84	12.00	9.13	12.00	8.15	8.00	5.56
7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91
5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89

	Value	Equality
X Mean	9	=
Y Mean	7.50	.00
X Variance	11	=
Y Variance	4.12	.00
Correlation	0.816	.000
Linear regression line	y = 3.00 + 0.500x	.00 and .000


I
II
III
IV
y
x
12.00
10.00
8.00
6.00
4.00
4
6
8
10
12
14
16
18

There are three types of lies: lies, damned lies, and statistics

Unknown author, popularized by Mark Twain

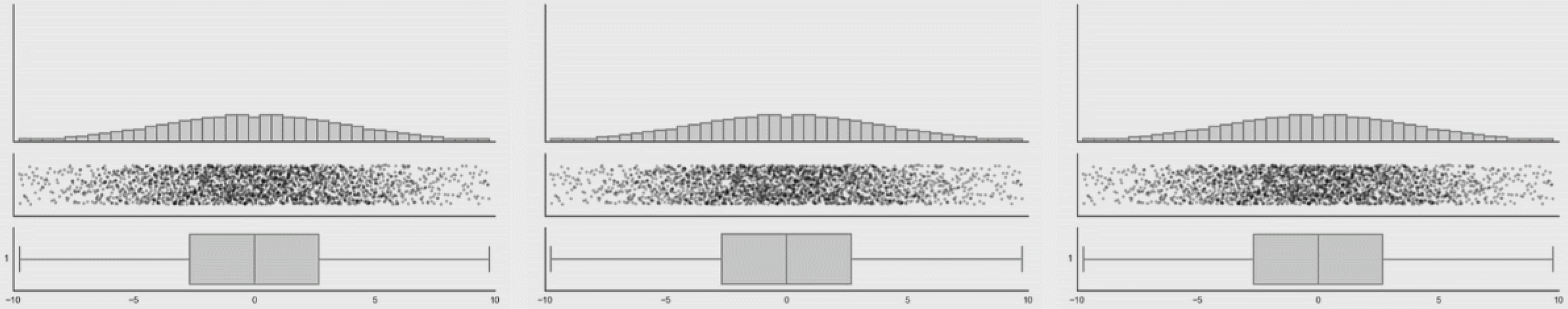


X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526

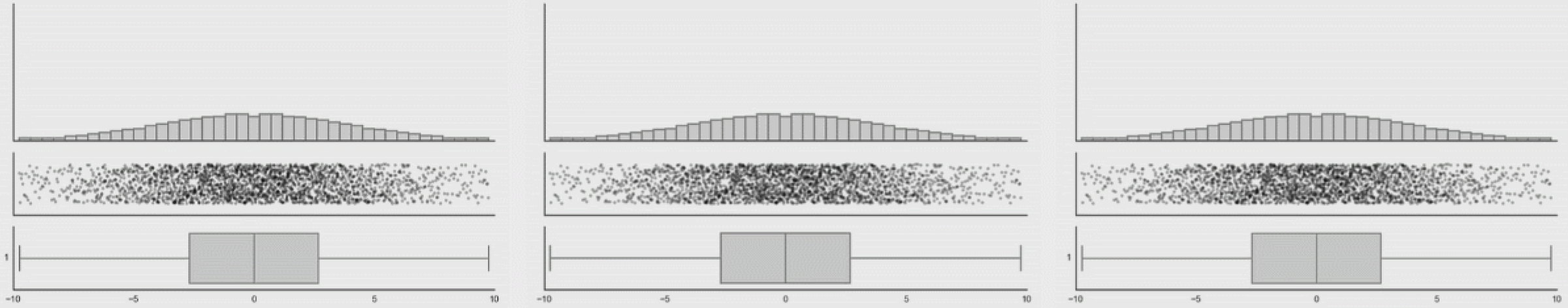
Matejka &Fitzmaurice, 2017

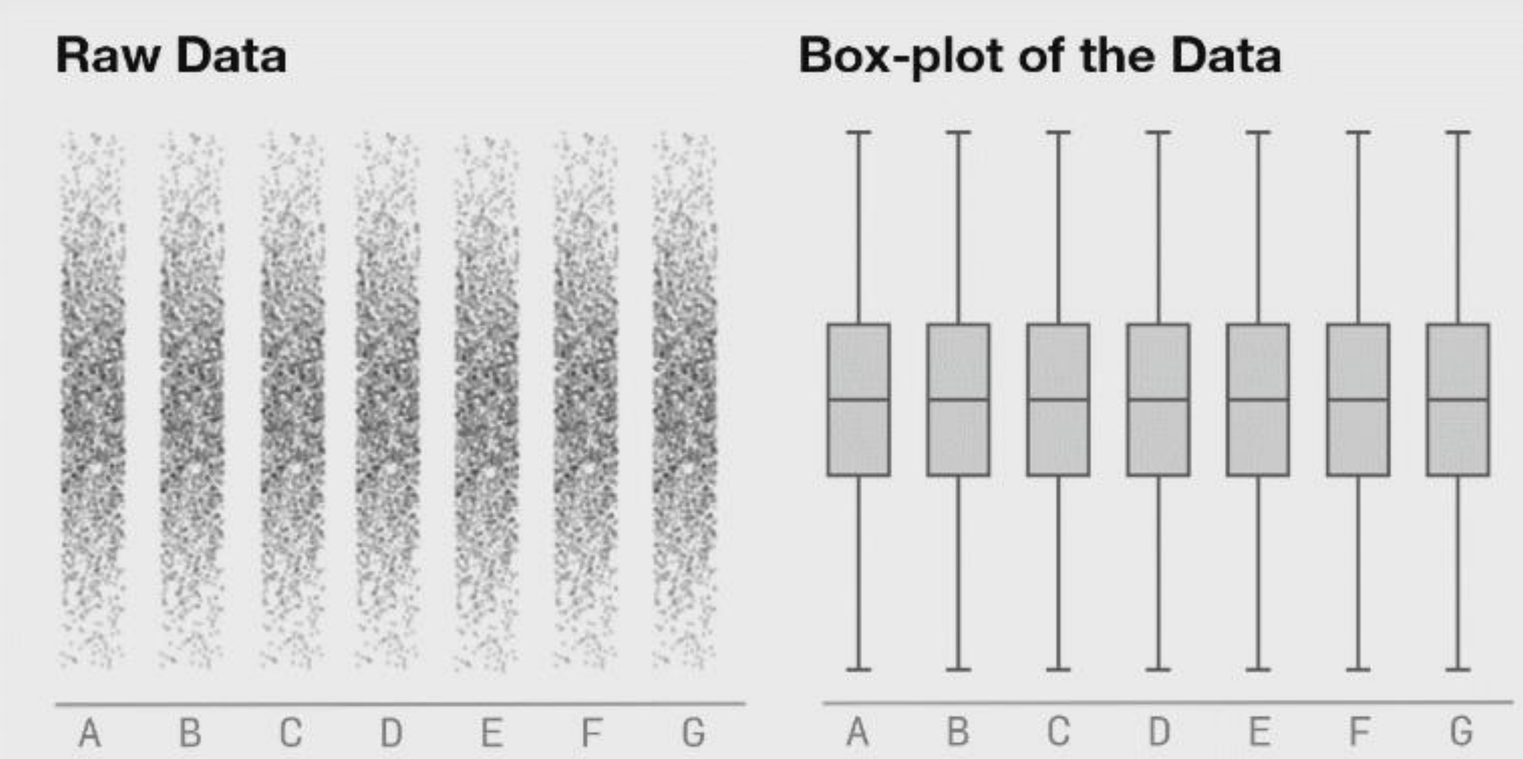
No catalogue of techniques can convey a willingness to look for what can be seen, whether or not anticipated. Yet this is at the heart of exploratory data analysis. ... the picture-examining eye is the best finder we have of the wholly unanticipated.

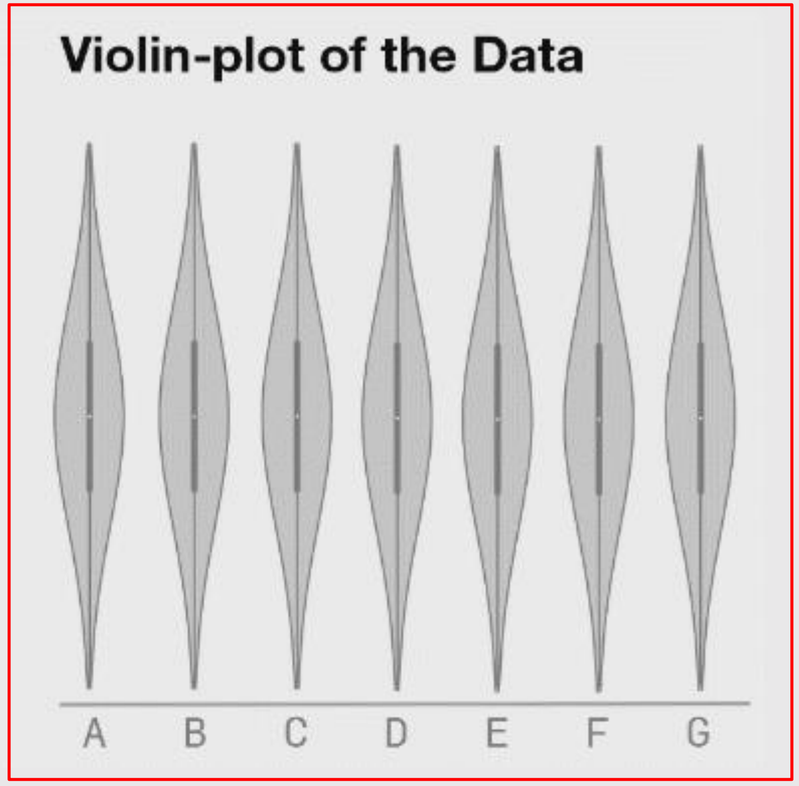
– Tukey, 1980



Matejka &Fitzmaurice, 2017







Matejka &Fitzmaurice, 2017

Ok, but why do we need visualization?

Why visualize your data?

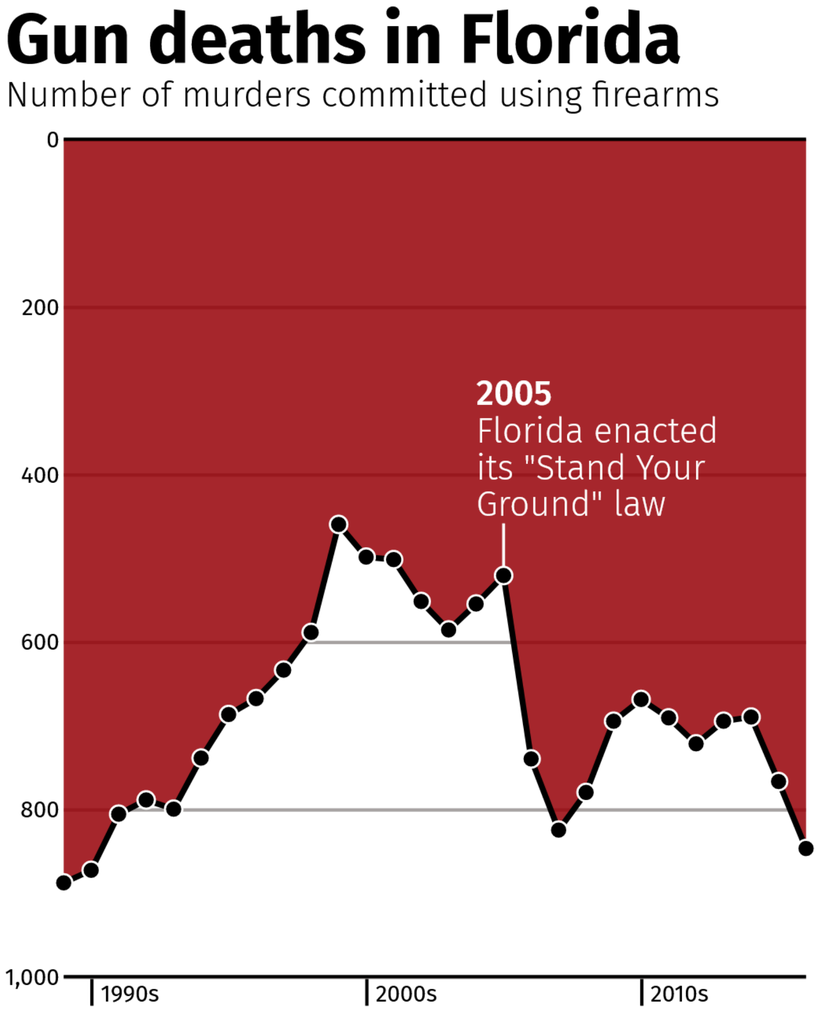
- Help cognition
- Expand memory
- Generate hypotheses
- Answer questions
- Make decisions
- Find patterns
- Record
- Clarify
- Communicate
- Inspire

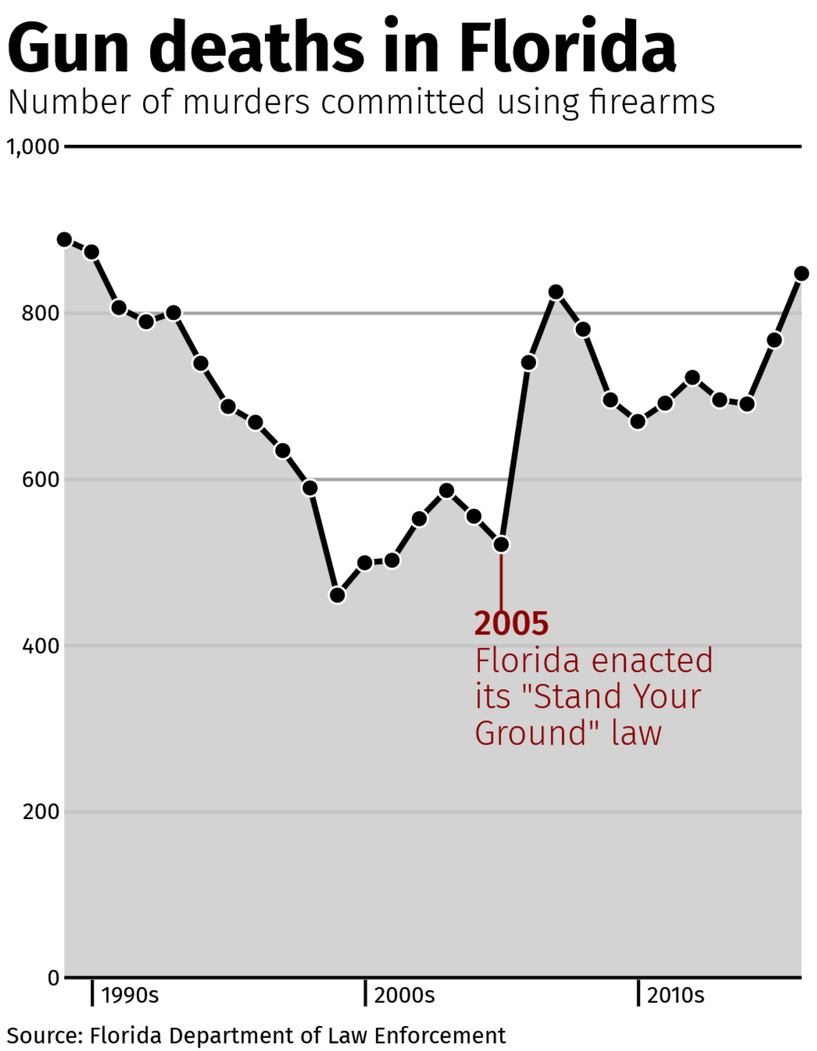
Design Rules of Thumb

Design Rules of Thumb

1. Function first, form next

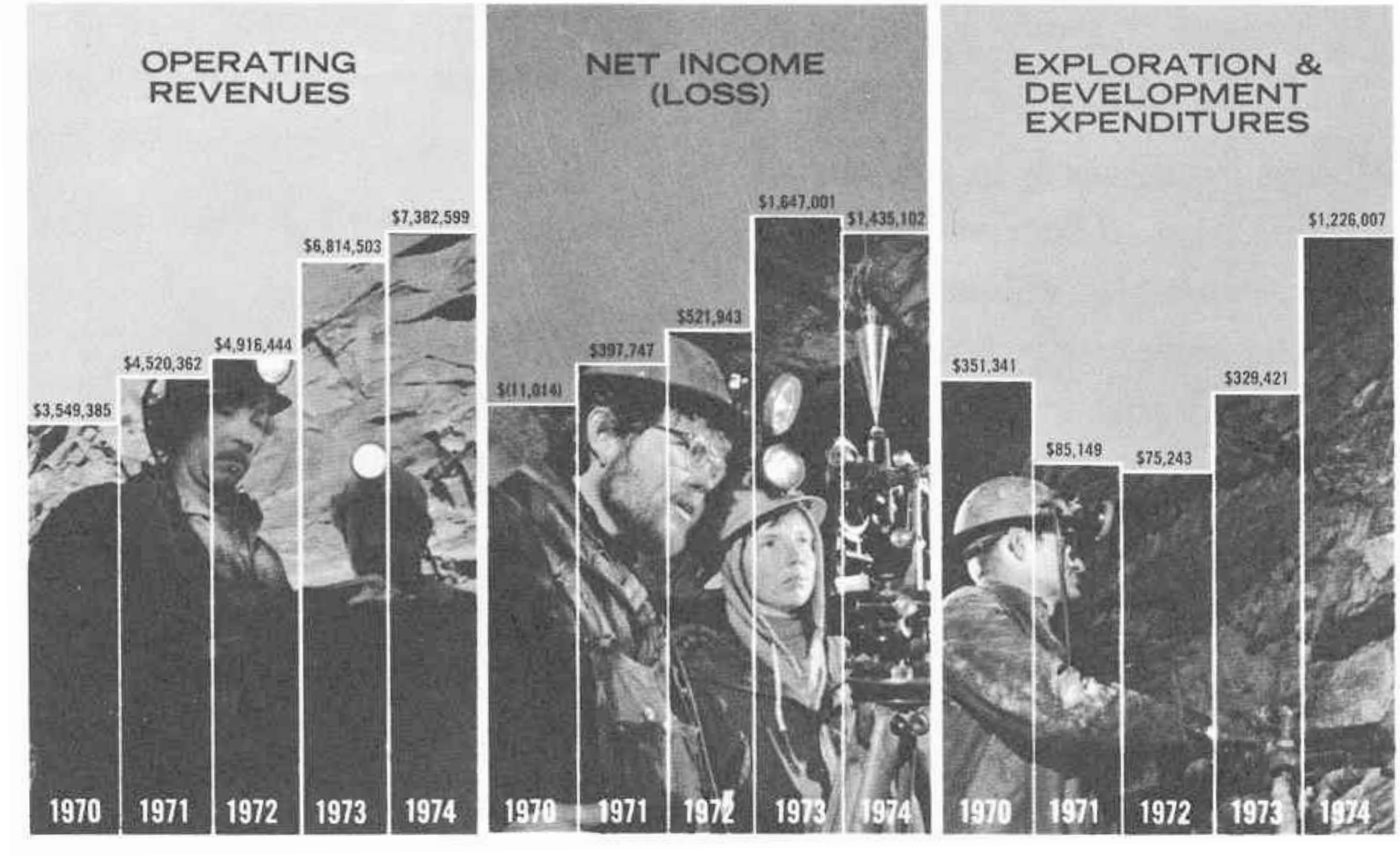
“Function first, form next





Christine Chen (2014), Sophie Engle (2020), Cody (2022)

“Function first, form next



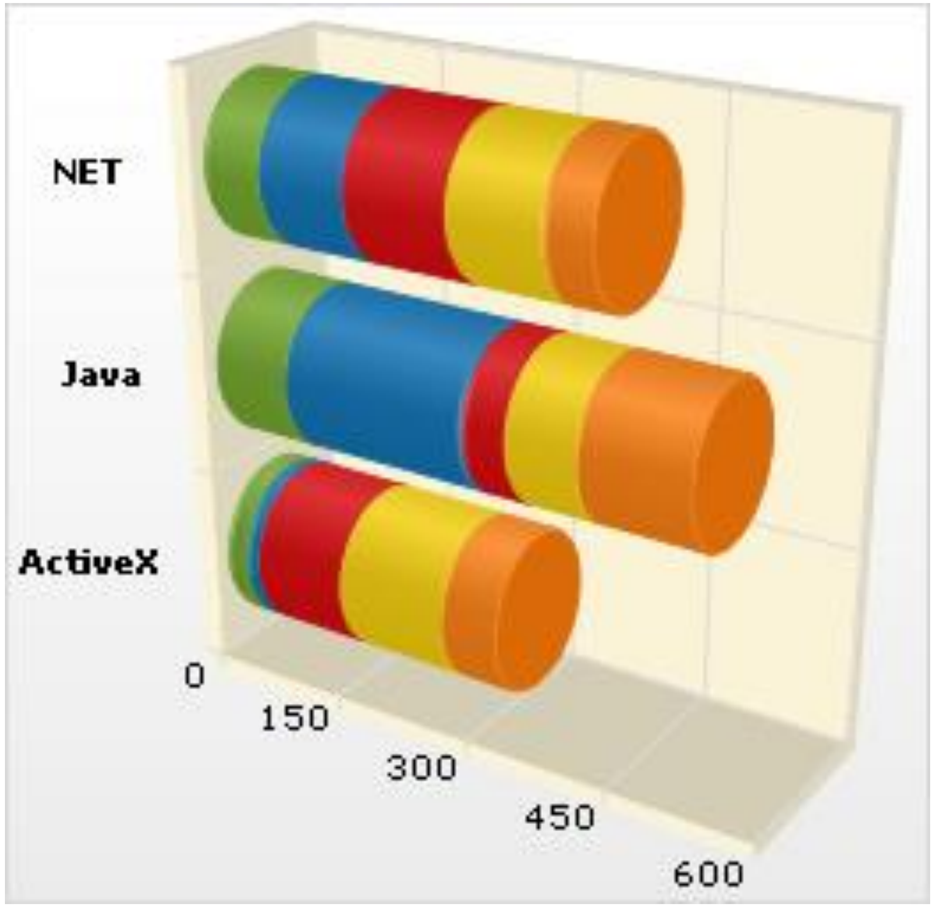
“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

Tufte (1983)

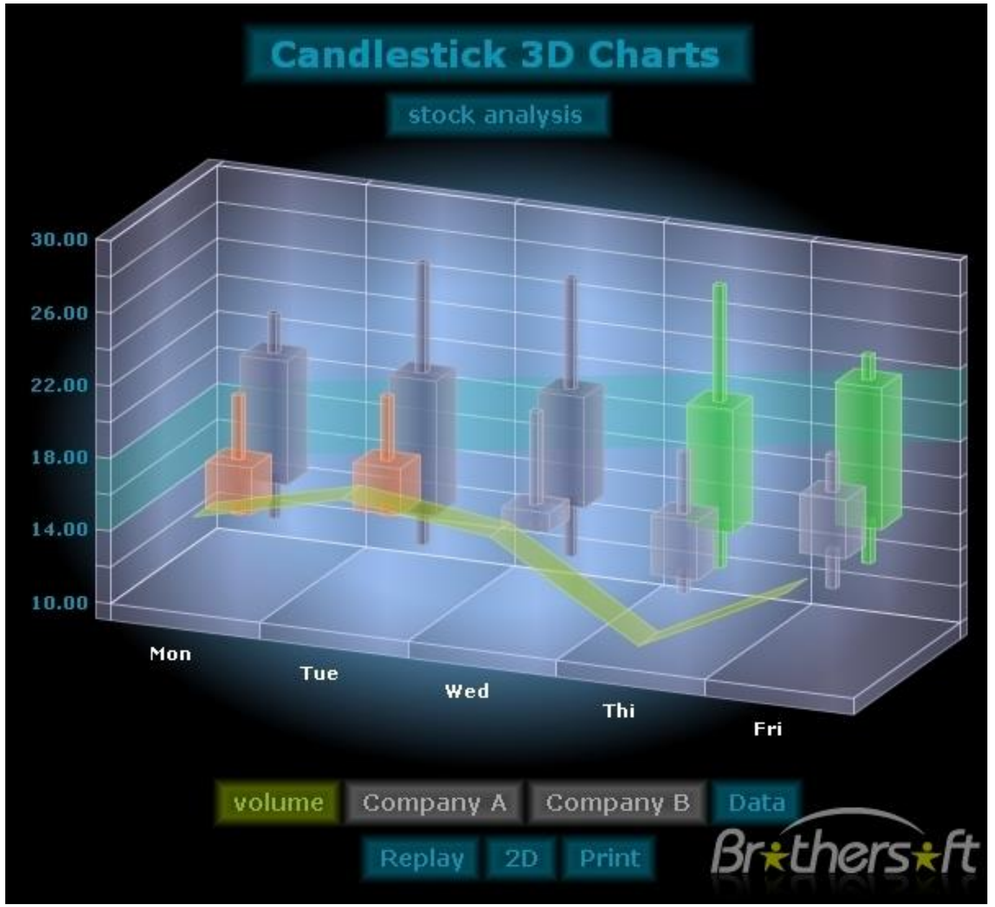
Design Rules of Thumb

1. Function first, form next
2. No unjustified 3D

“No Unjustified 3D”



http://help.infragistics.com/Help/Doc/WinForms/2014.2/CLR4.0/html/Images/Chart_Bar_Chart_03.png

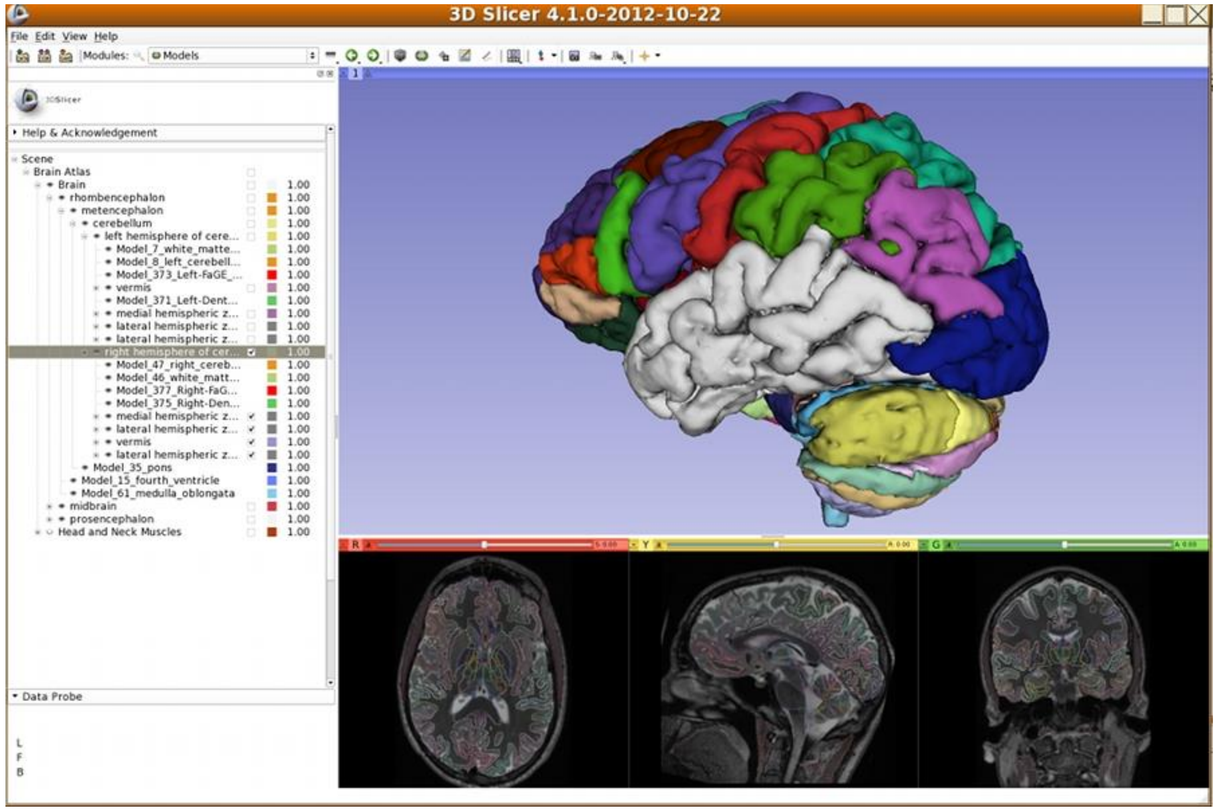


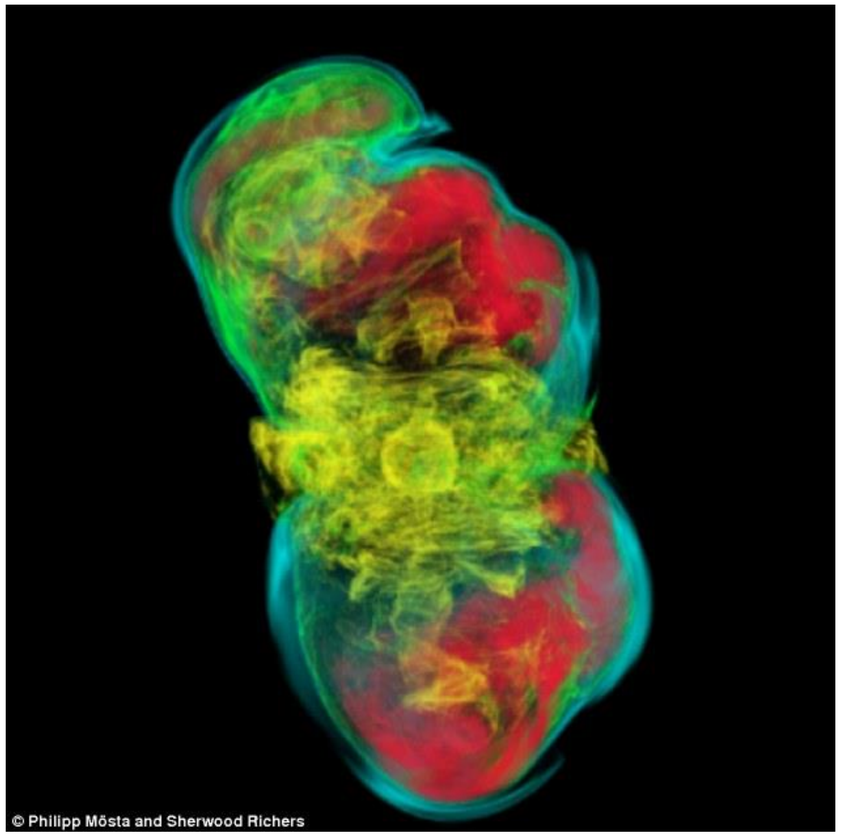
http://img.brothersoft.com/screenshots/softimage/0/3d_charts-171418-1269568478.jpeg

“No Unjustified 3D”

http://stats.stackexchange.com/questions/109076/what-is-your-favorite-statistical-graph/109080

“No Unjustified 3D”

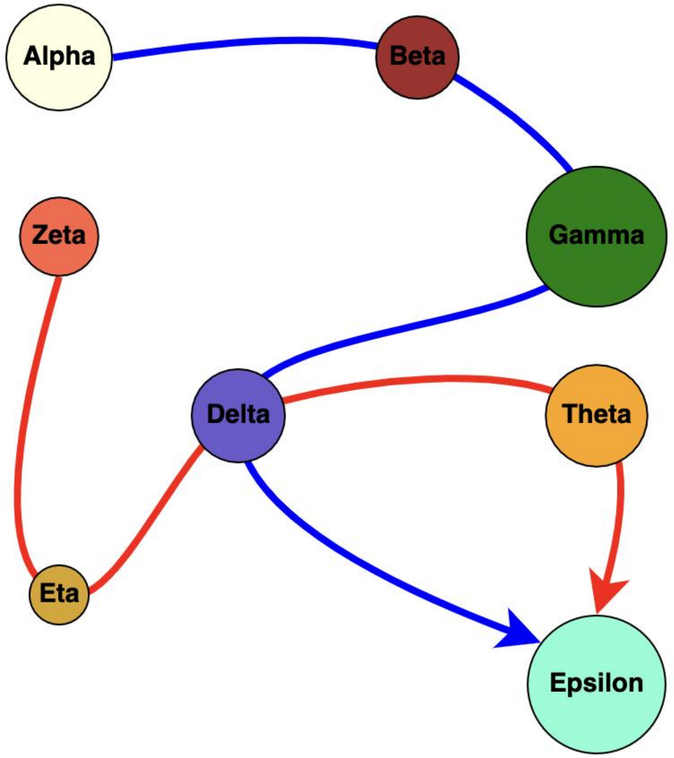




Design Rules of Thumb

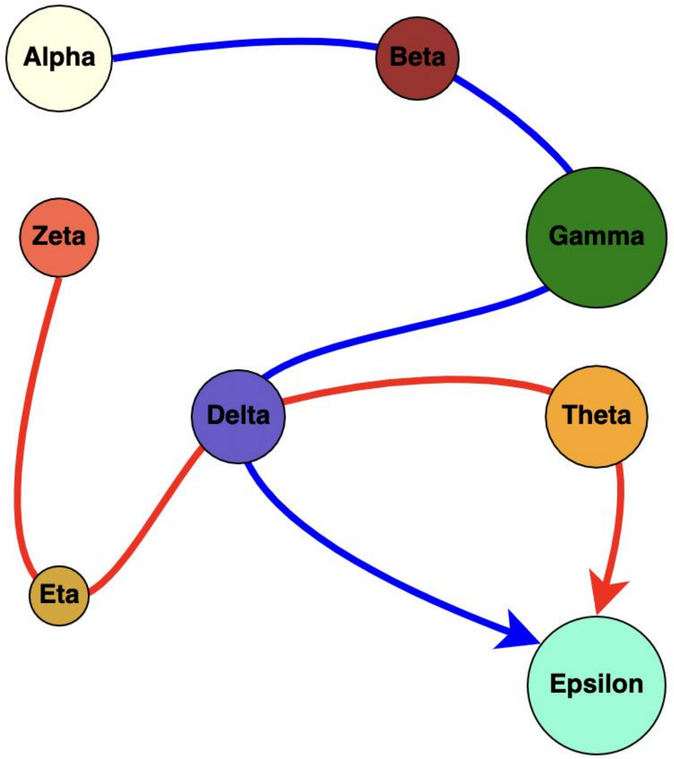
1. Function first, form next
2. No unjustified 3D
3. No unjustified 2D

“No Unjustified 2D”



Task: What color is Delta?

“No Unjustified 2D”



Task: What color is Delta?

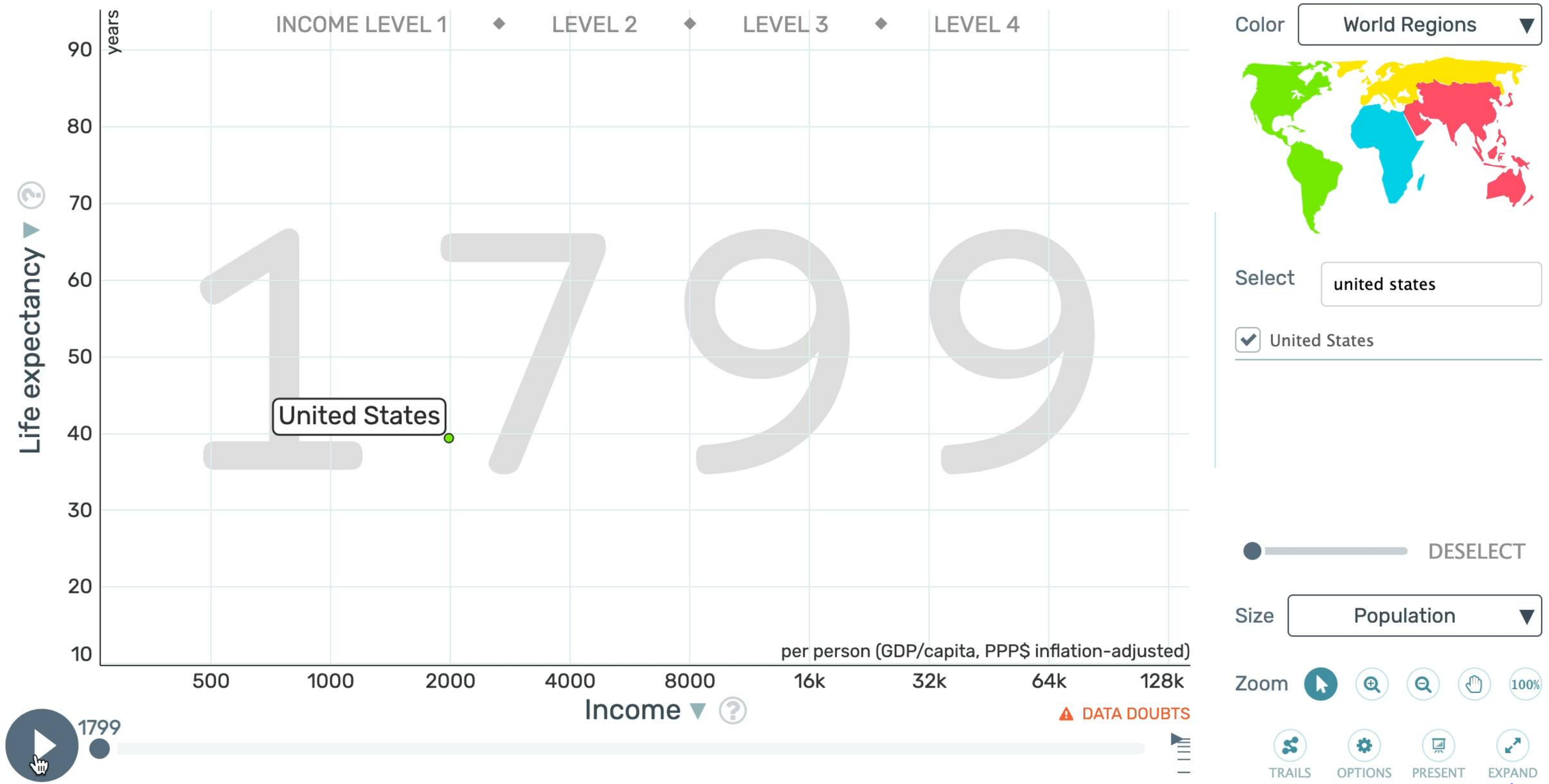
Node	Color
Alpha	White
Beta	Maroon
Delta	Purple
Epsilon	Teal
Eta	Mustard Yellow
Gamma	Green
Theta	Orange
Zeta	Pink

If the task doesn’t need a 2D visualization, then don’t use one.

Design Rules of Thumb

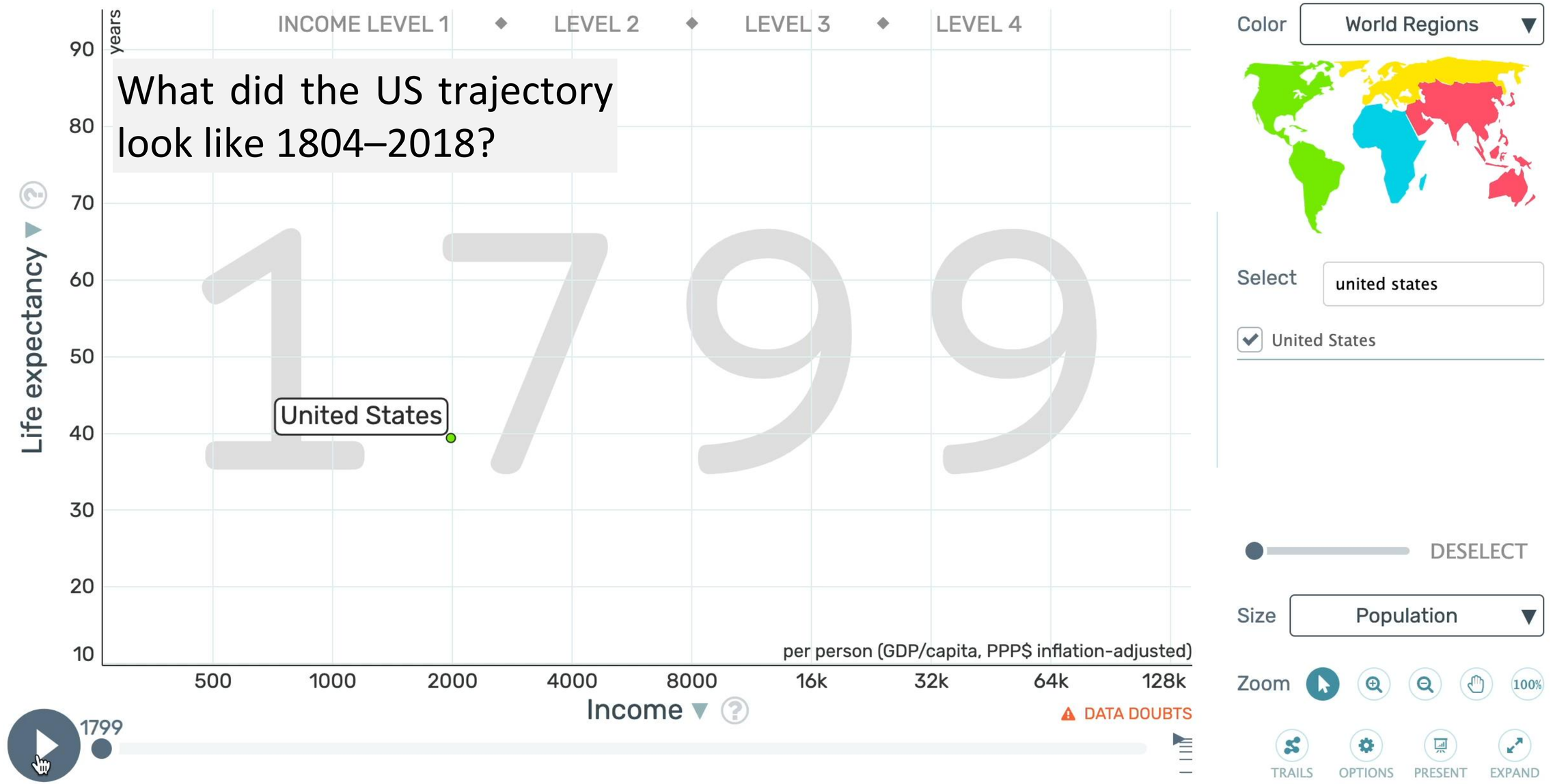
1. Function first, form next
2. No unjustified 3D
3. No unjustified 2D
4. Eyes beat memory

“Eyes Beat Memory”

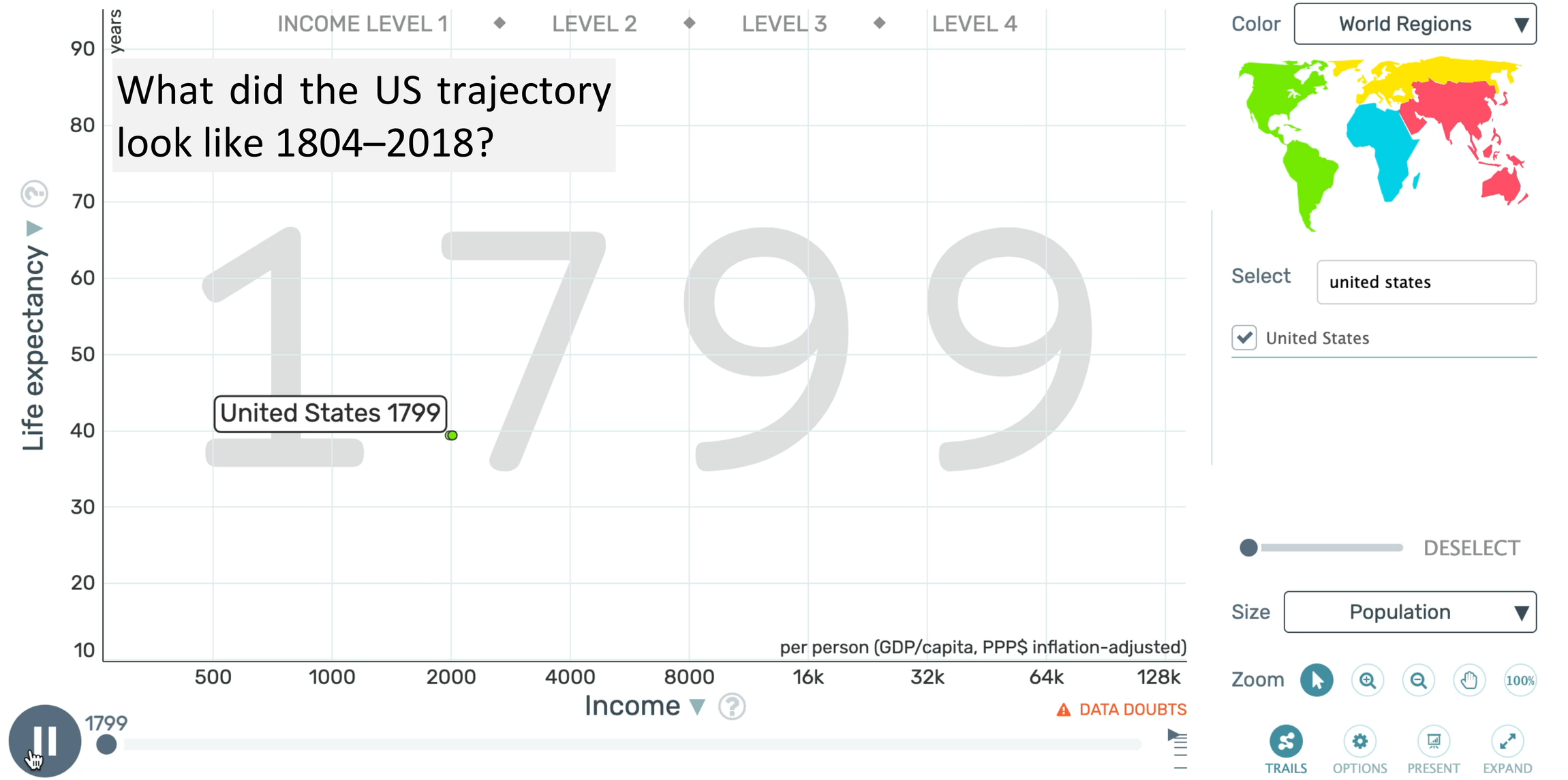


Gapminder, 2018

“Eyes Beat Memory”



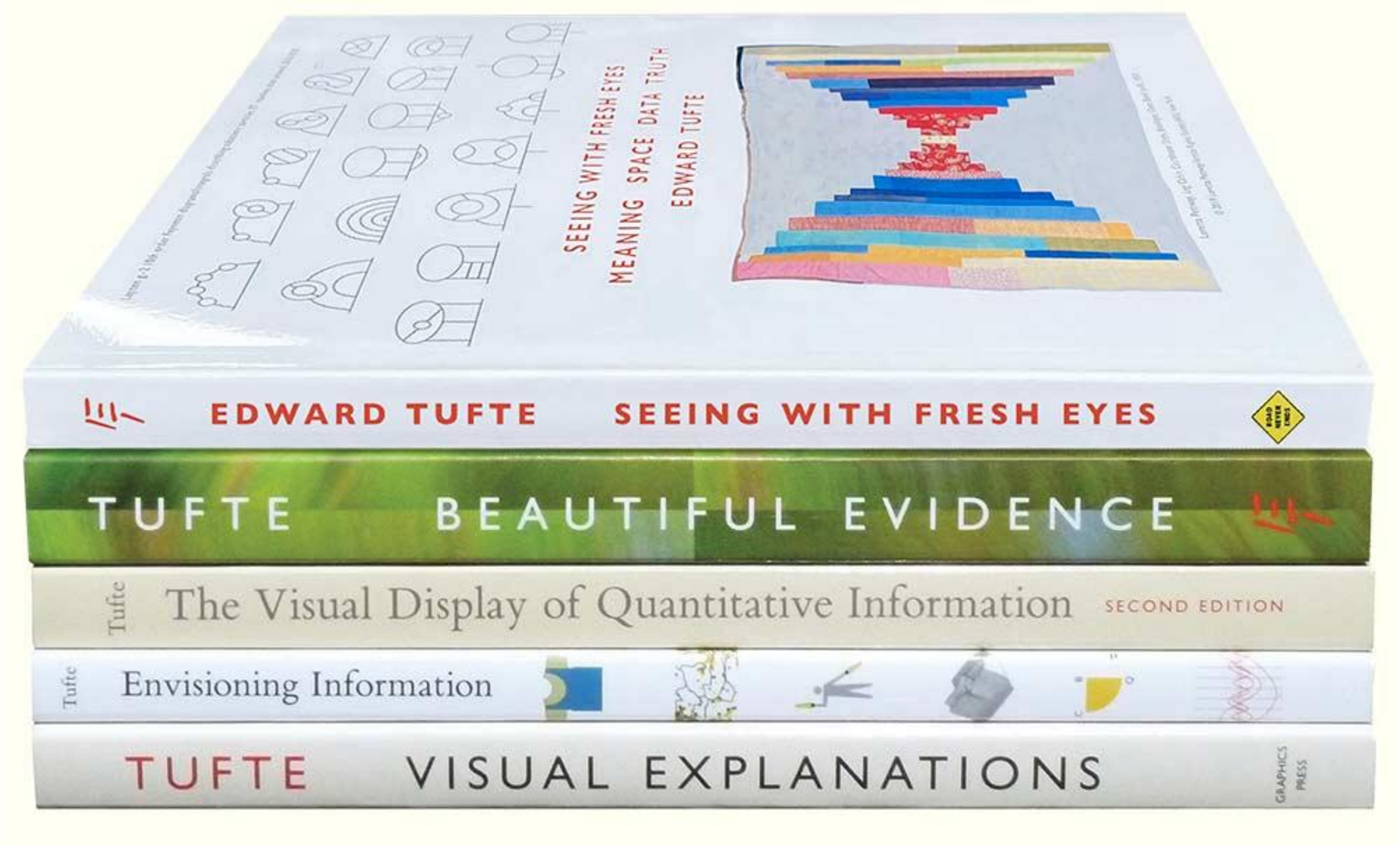
“Eyes Beat Memory”



Design Rules of Thumb

1. Function first, form next
2. No unjustified 3D
3. No unjustified 2D
4. Eyes beat memory

Edward Tufte



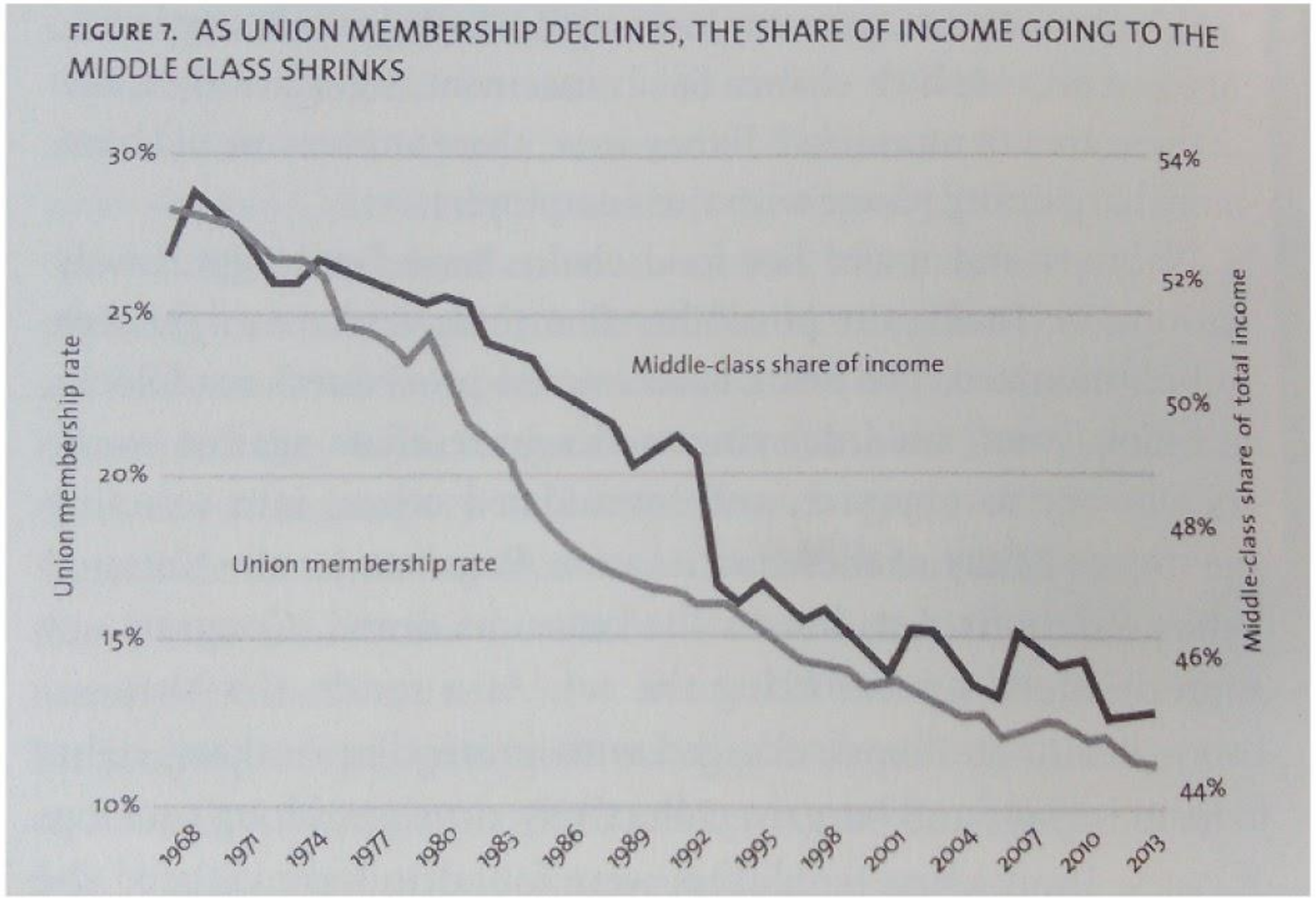
“Graphical Integrity”

“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

(Axes and axis labels, titles, annotations, legends, etc.)

Tufte, “Visual Display of Quantitative Information” (1983)

"Double the axes, double the mischief"



"Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data." http://www.thefunctionalart.com/2015/10/double-axes-double-mischief.html

“Graphical Integrity”

“The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured.”

Tufte, “Visual Display of Quantitative Information” (1983)

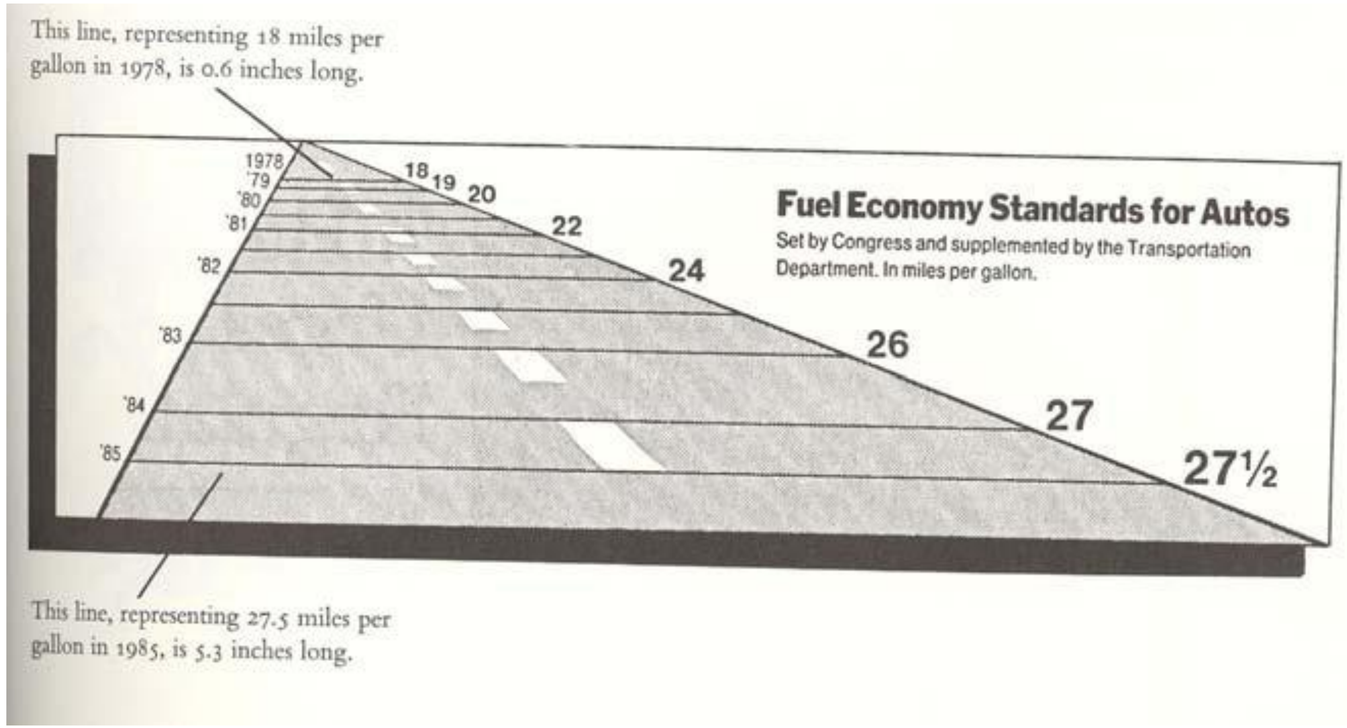
Lie Factor

$$\text{Lie Factor} = \frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$$

Lie Factor = >1, overstating

Lie Factor = 1, accurate :-)

Lie Factor = <1, understating



"The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured."

Tufte, "Visual Display of Quantitative Information" (1983)

Lie Factor

Lie Factor = $\frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$

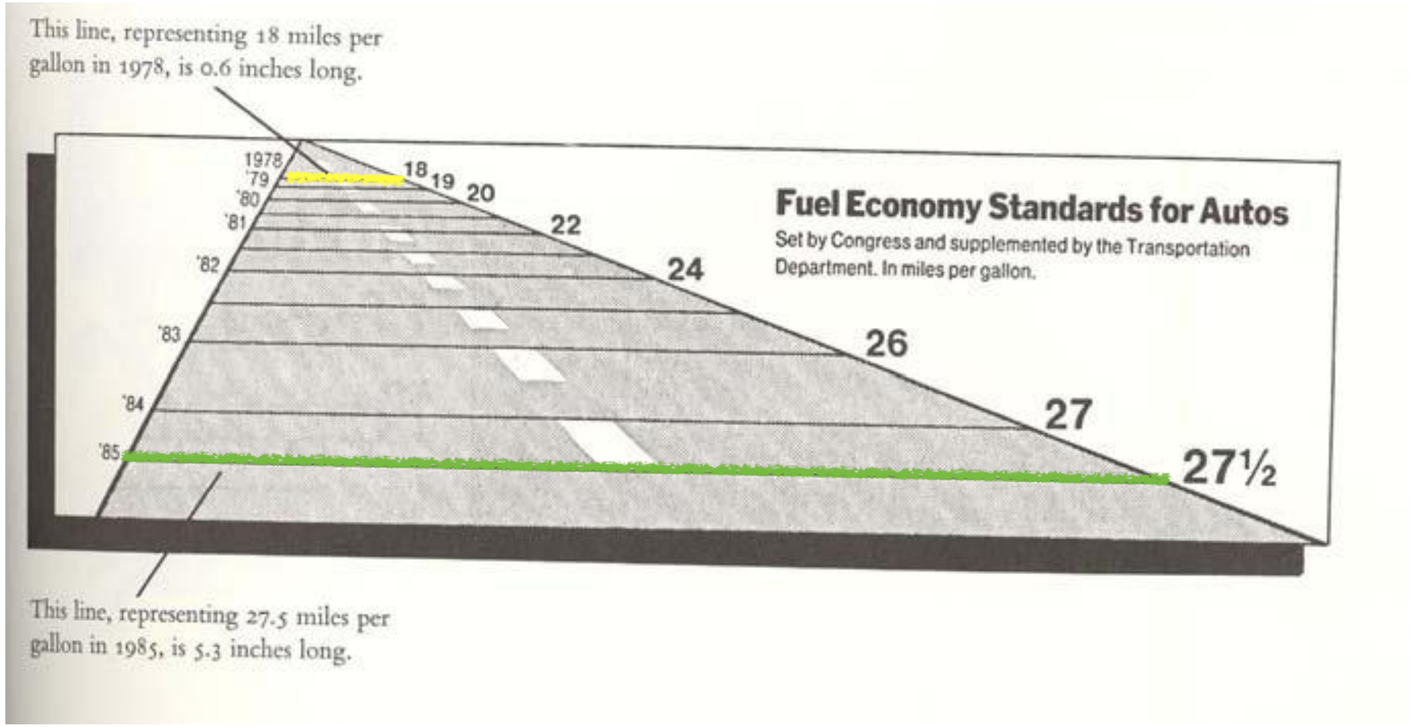


Image = $\frac{5.3" - 0.6"}{0.6"}$ = 7.83 = 783%

Data = $\frac{27.5 - 18}{18}$ = 0.53 = 53%

Lie Factor = $\frac{783\%}{53\%}$ = 14.8

Lie Factor = >1, overstating

"The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured."

Tufte, "Visual Display of Quantitative Information" (1983)

Lie Factor

Lie Factor = $\frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$

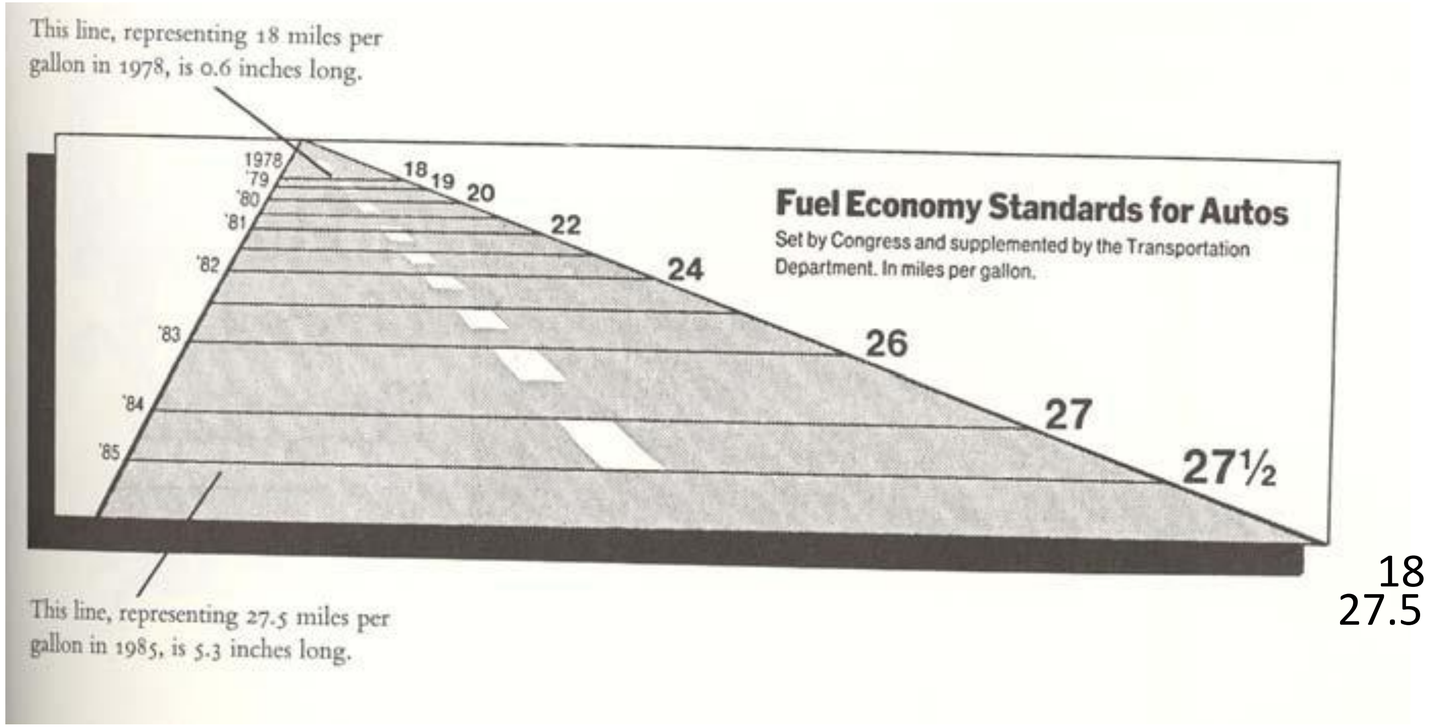


Image = $\frac{5.3" - 0.6"}{0.6"}$ = 7.83 = 783%

Data = $\frac{27.5 - 18}{18}$ = 0.53 = 53%

Lie Factor = $\frac{783\%}{53\%}$ = 14.8

Lie Factor = >1, overstating

18
27.5

"The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured."

Tufte, "Visual Display of Quantitative Information" (1983)

IN-CLASS ACTIVITY:
Calculate for yourself!

Lie Factor

$$\text{Data} = \frac{2 - 1}{1} = 1 = 100\%$$

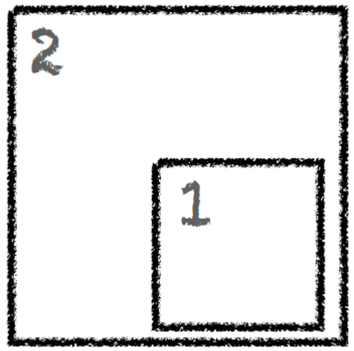
$$\text{Lie Factor} = \frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$$

✓

$$\text{Image} = \frac{2 - 1}{1} = 1 = 100\%$$

$$\text{Lie Factor} = \frac{100\%}{100\%} = 1$$

Make sure *area* is proportional to data!

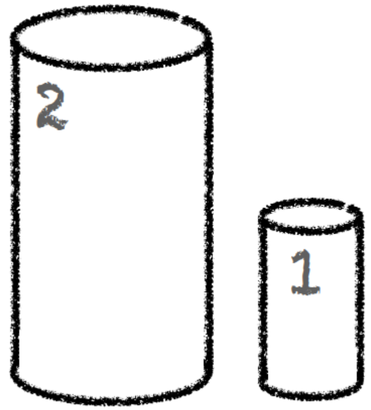


X

$$\text{Image} = \frac{2^2 - 1^2}{1^2} = 3 = 300\%$$

$$\text{Lie Factor} = \frac{300\%}{100\%} = 3$$

Don't use 3D bar charts!



X!!!

$$\text{Image} = \frac{2*\pi 1^2 - 1*\pi 0.5^2}{1*\pi 0.5^2} = 7 = 700\%$$

$$\text{Lie Factor} = \frac{700\%}{100\%} = 7$$

"The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured."

Tufte, "Visual Display of Quantitative Information" (1983)

“Graphical Integrity”

“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

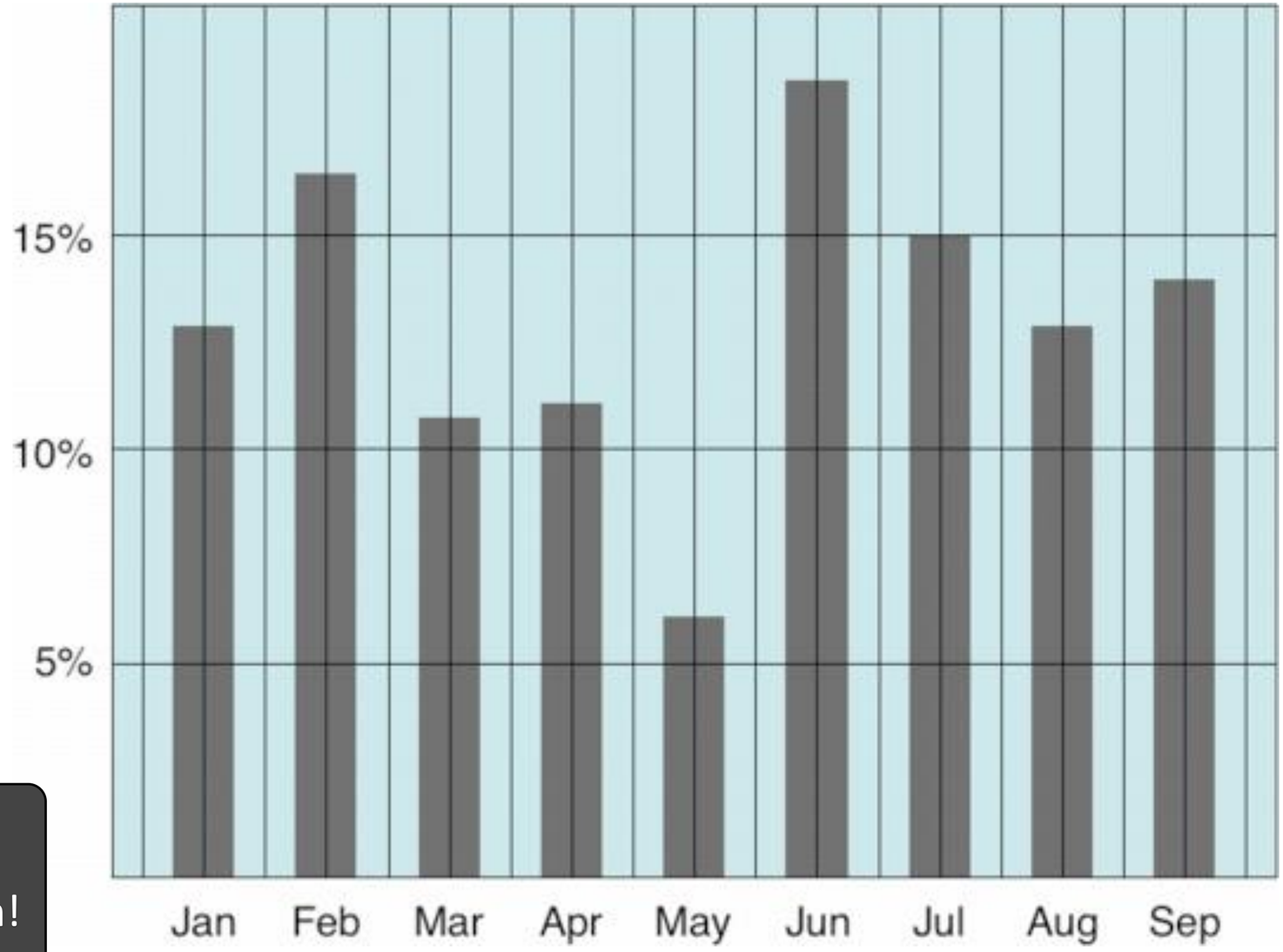
Tufte, “Visual Display of Quantitative Information” (1983)

IN-CLASS EXERCISE

- (No submission)

In-Class Sketching — “Graphical Integrity”

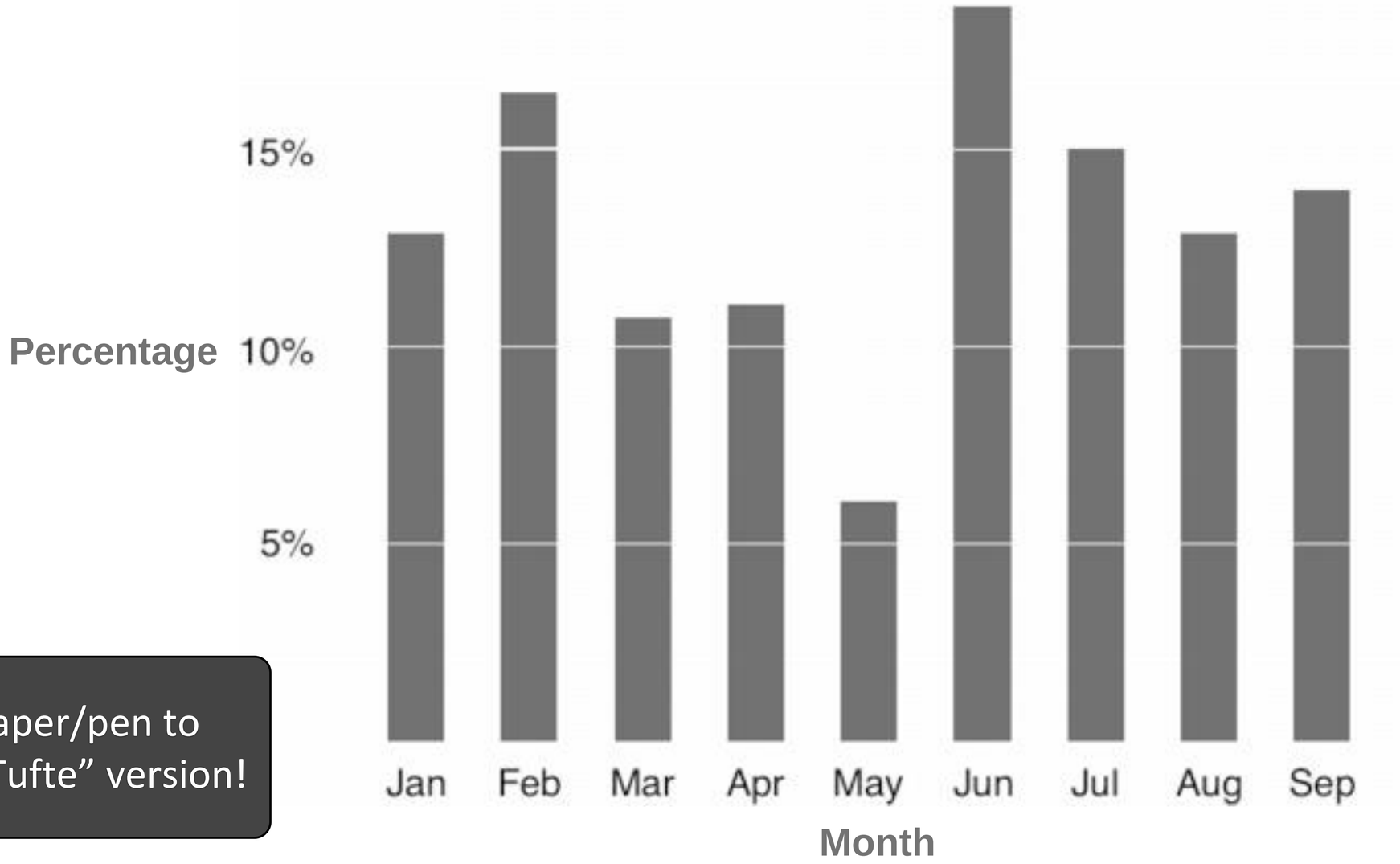
~8 min



Use paper/pen to sketch “Tufte” version!

In-Class Sketching — “Graphical Integrity”

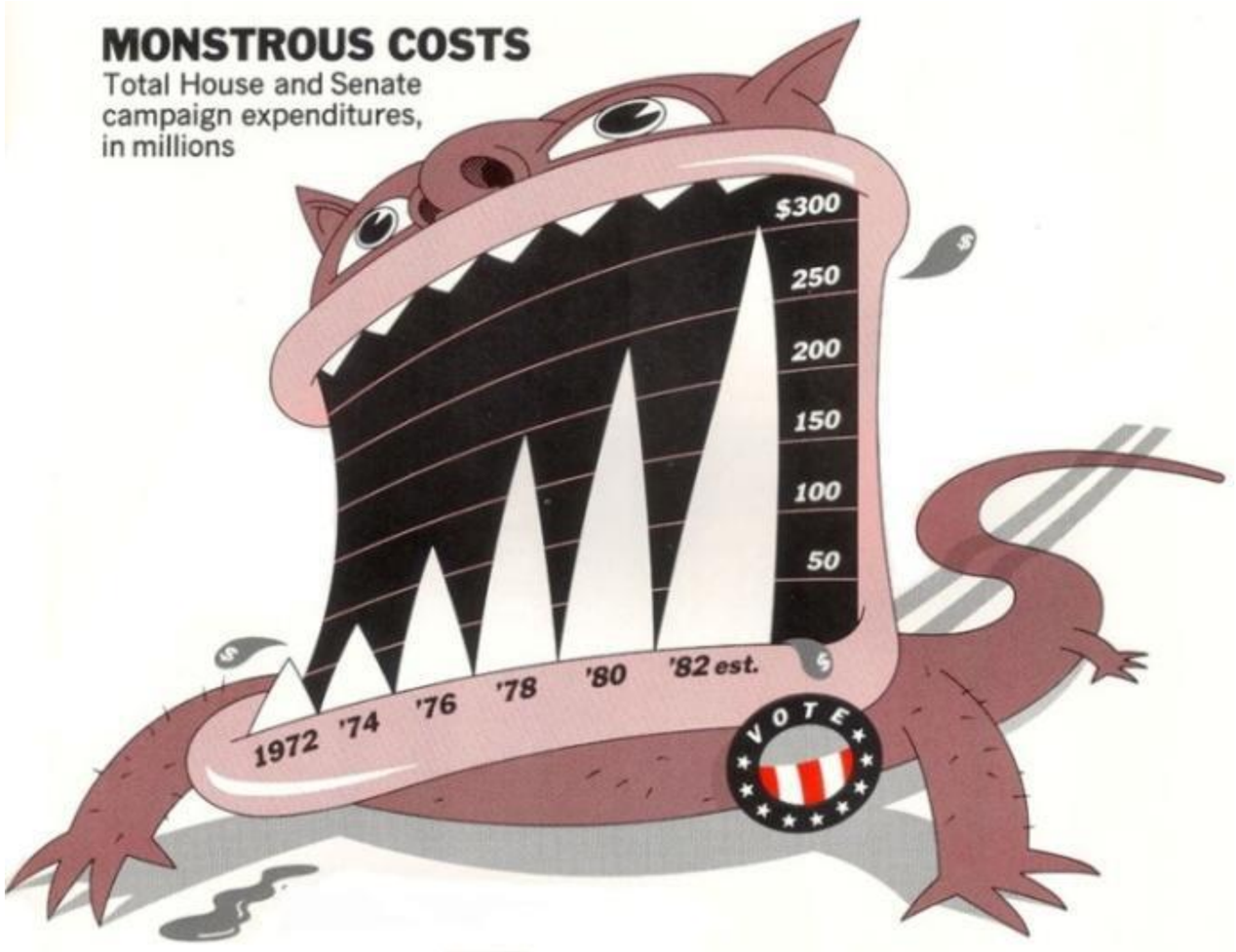
~8 min

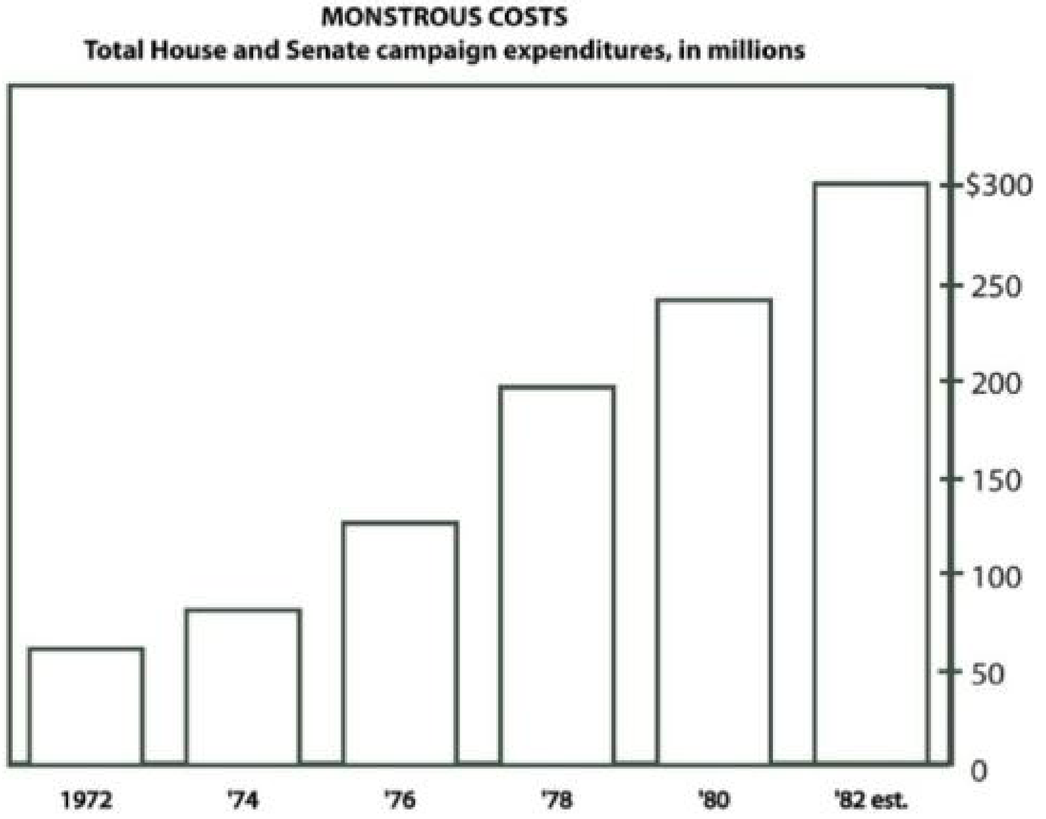


Use paper/pen to sketch “Tufte” version!

“CHART JUNK”

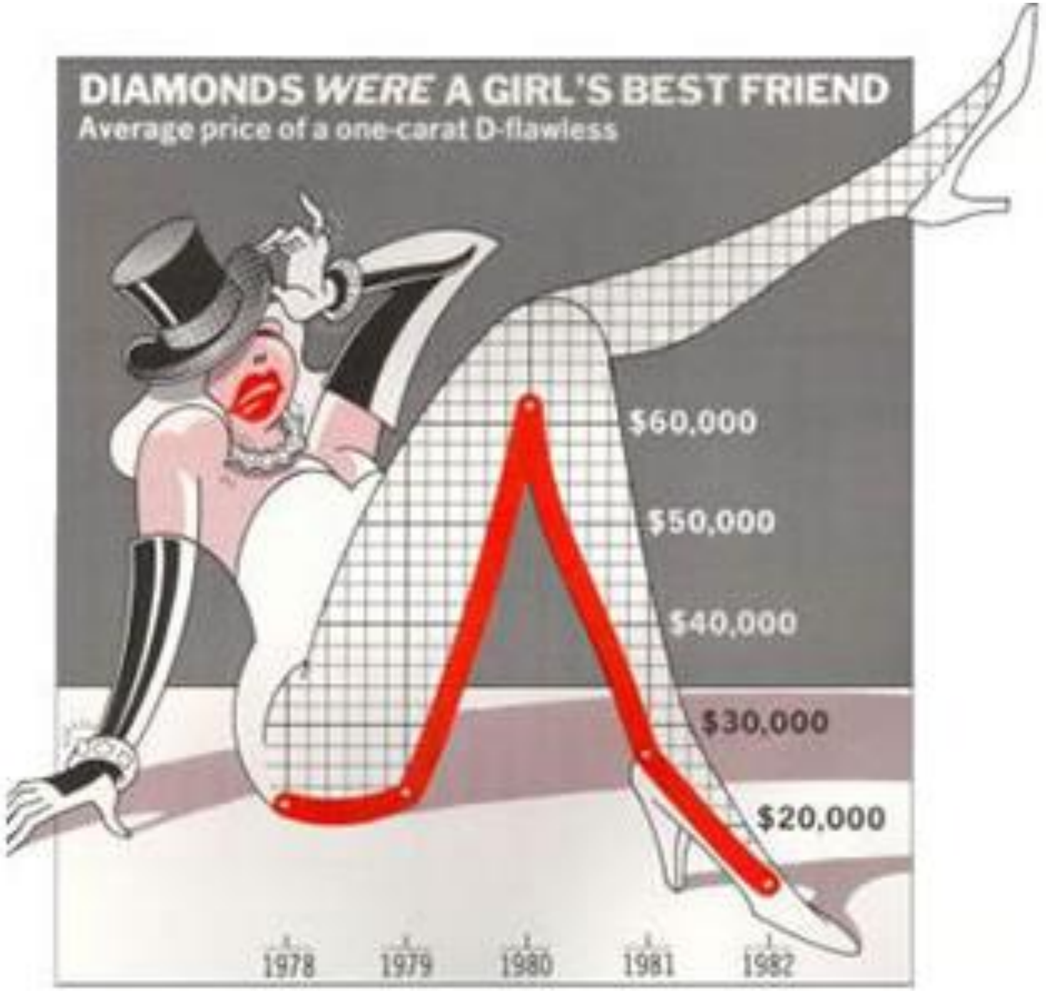
“Chart Junk”

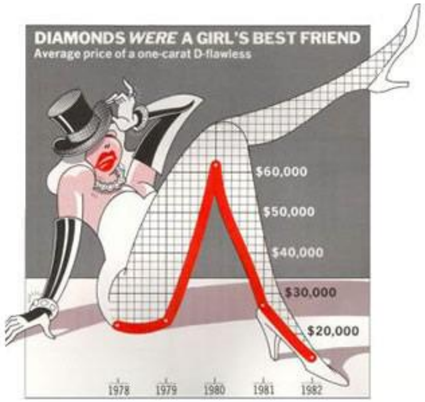


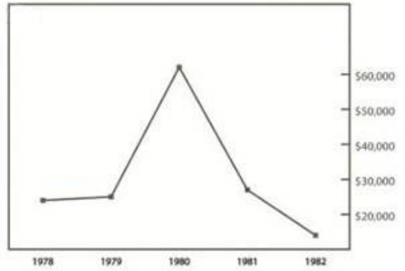


Bateman, et al. (2010)

“Chart Junk”



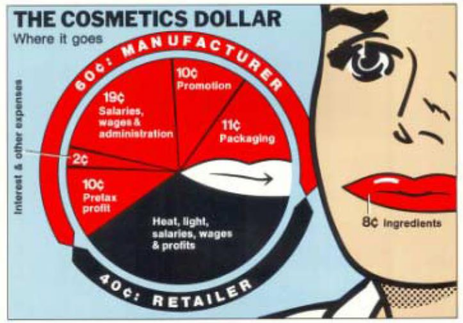


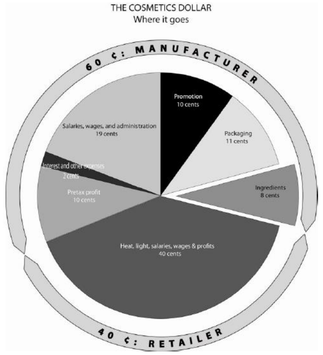


Bateman, et al. (2005)

“Chart Junk Debate”

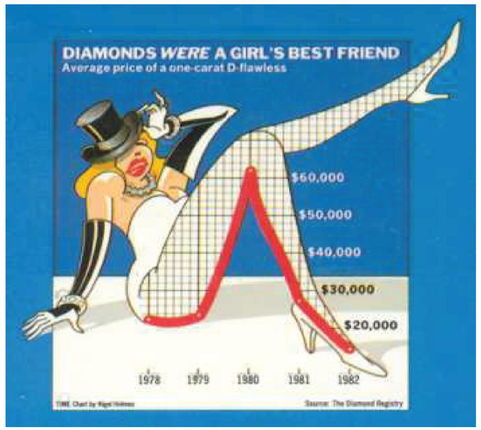
Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts





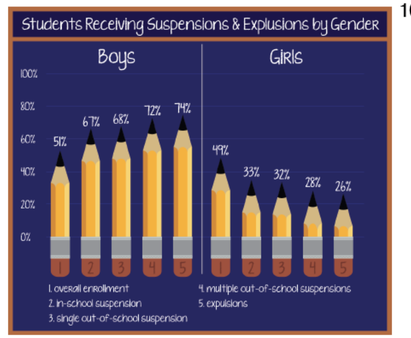
Bateman, et al. (2010)

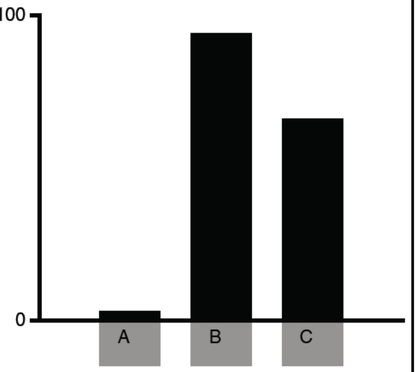
Benefitting InfoVis with Visual Difficulties



Hullman, et al. (2011)

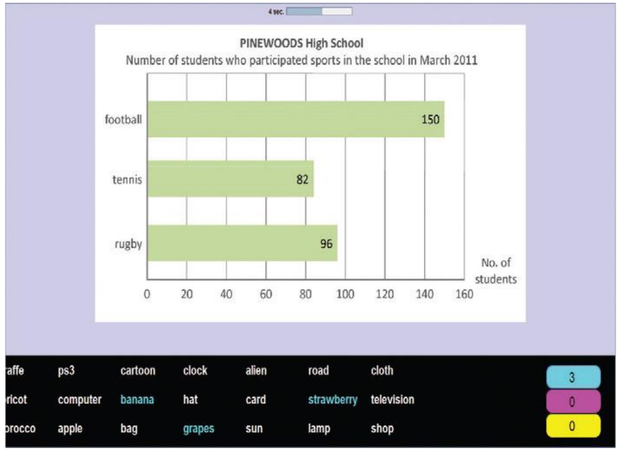
An Evaluation of the Impact of Visual Embellishments in Bar Charts





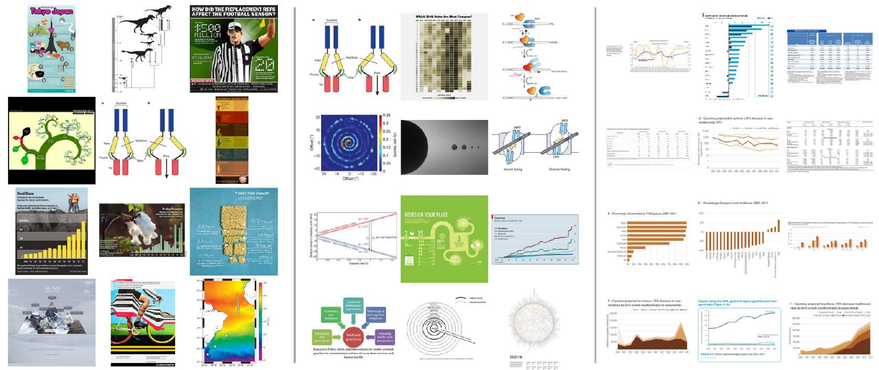
Skau, et al. (2015)

An Empirical Study on Using Visual Embellishments in Visualization



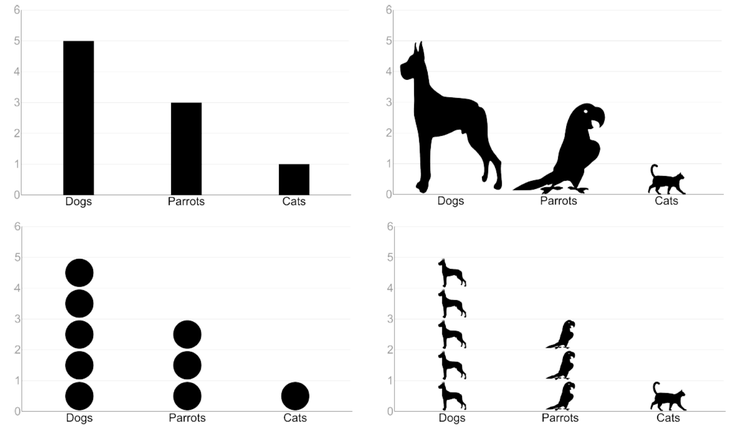
Borgo, et al. (2012)

What makes a visualization memorable?



Borkin, et al. (2013)
Borkin, et al. (2015)

ISOTYPE Visualization – Working Memory, Performance, and Engagement with Pictographs



Haroz, et al. (2015)

“Chart Junk”

Chart junk can... persuade, help with memorability, engage

... bias, limit data-ink ratio, clutter, lower trust

<u>Take-away</u>: it depends on your audience, task, and context...

For Next Time

neu-ds-4200-s22.github.io/schedule

Look at the upcoming assignments and deadlines

- Textbook, Readings, & Reading Quizzes—Variable days
- In-Class Activities—If due, they are due 11:59pm the same day as class

Everyday Required Supplies:

- 5+ colors of pen/pencil
- White paper
- Laptop and charger

Use Canvas Discussions for general questions, email codydunne-and-tas@ccs.neu.edu for questions specific to you.

Week	Topics	Assignments
#1: Jan 17–21	What is visualization Design rules of thumb	A1—Setting up
#2: Jan 24–28	JS development, projects Marks & channels	A2—Encodings & xenographics
#3: Jan 31–Feb 04	Data types and tasks, Tableau D3 tutorial 1/2	P1—Pitches★
#4: Feb 07–11	In-class group formation D3 tutorial 2/2	A3—Tableau analysis P2—Proposal★
#5: Feb 14–18	Altair and JupyterLab Arrange tables	A4—D3 basic charts
#6: Feb 21–25	Color Pop-out, illusions	A5—Altair basic charts P3—Interview & tasks
#7: Feb 28–Mar 04	Interaction & animation (2)	A6—D3 event handling P4—Data, Initial sketches
#8: Mar 07–11	Trees & networks (2)	P5—Final sketches & plan★