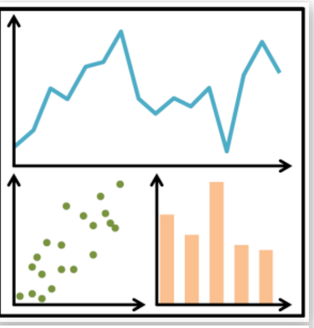
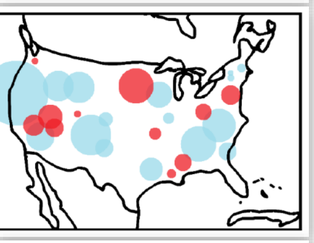
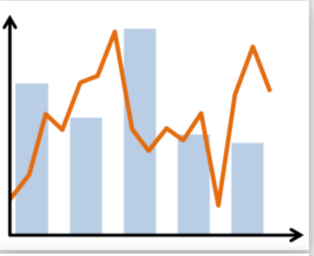
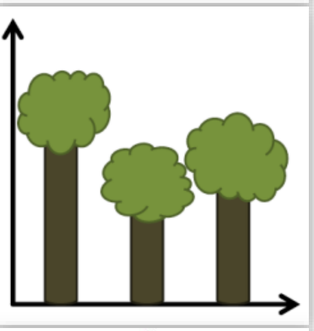
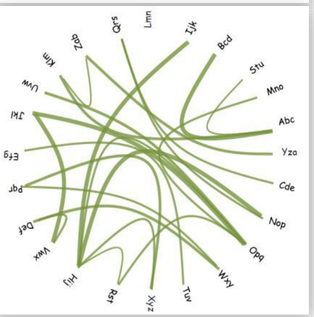






Lecture 6:
Design rules of thumb, continued…

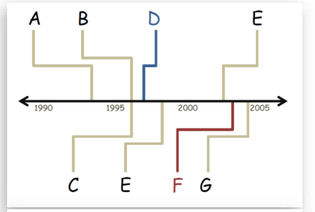
DS 4200

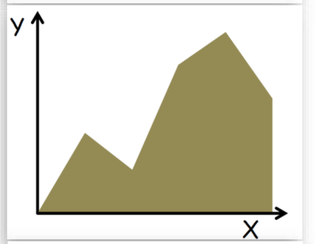
FALL 2020

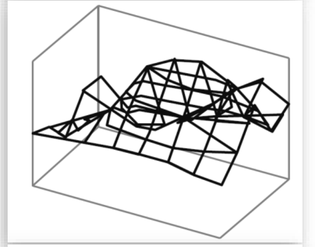
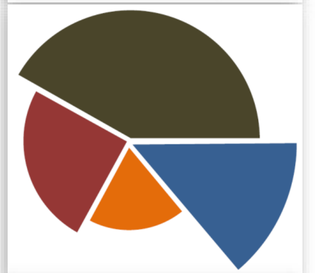
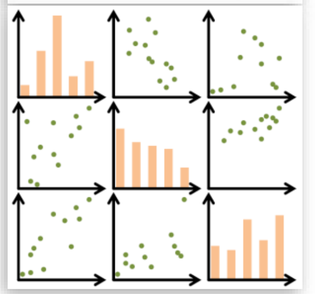
Prof. Cody Dunne

NORTHEASTERN UNIVERSITY

Slides and inspiration from Michelle Borkin, Krzysztof Gajos, Hanspeter Pfister, Miriah Meyer, Jonathan Schwabish, and David Sprague





Check-in

READING QUIZ

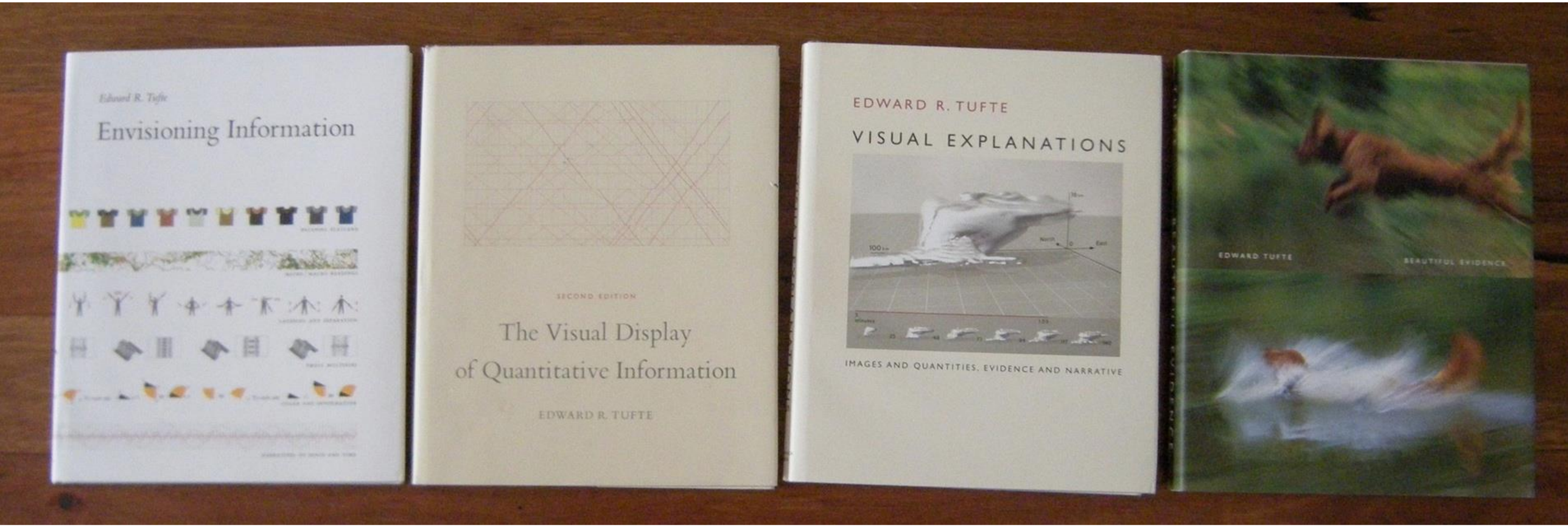
Quiz — Data Types & Tasks

Password: ??????

Now, on DS 4200…

DESIGN & RULES OF THUMB

Edward Tufte



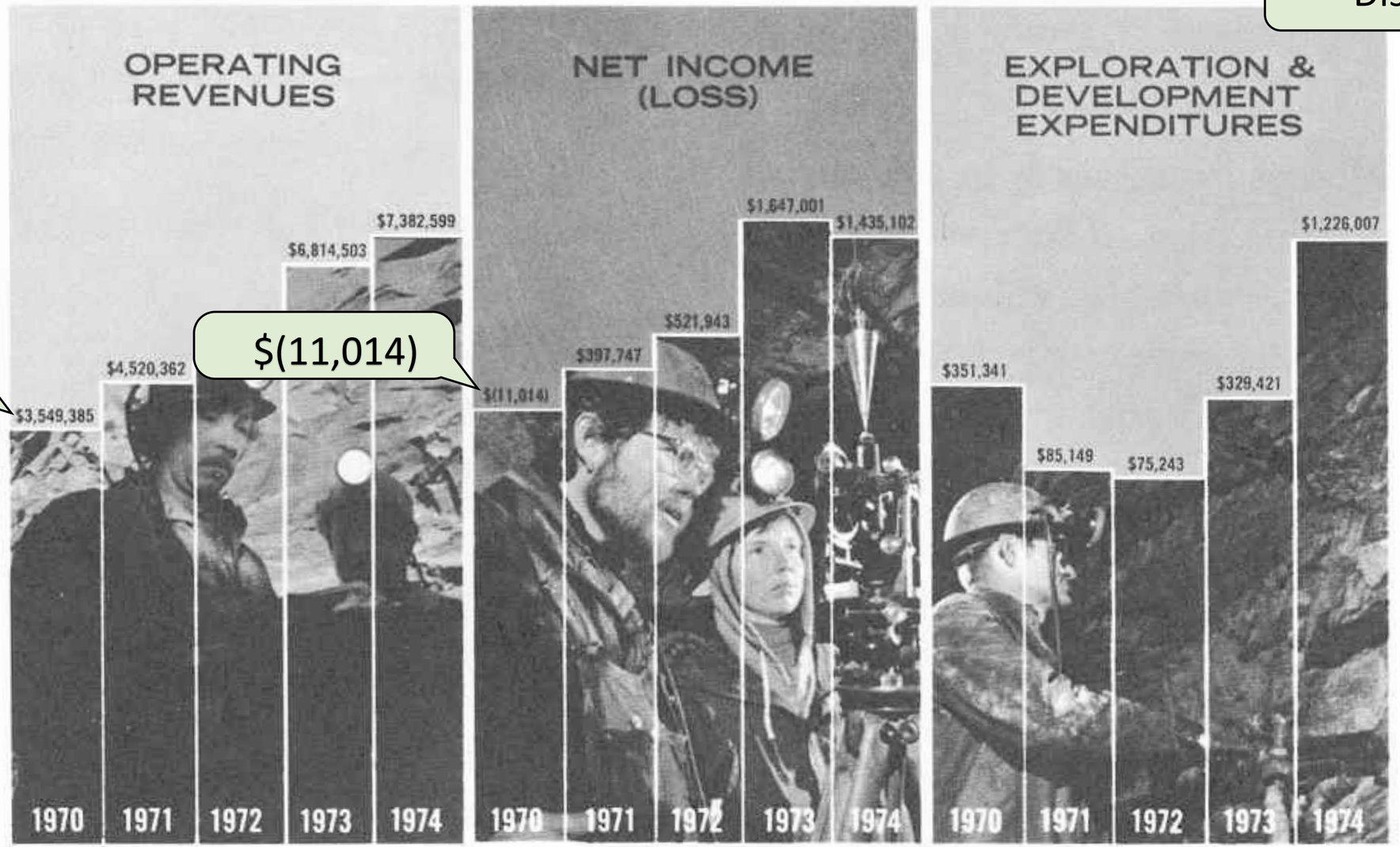
“Graphical Integrity”

“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)

“Distorted Scales”



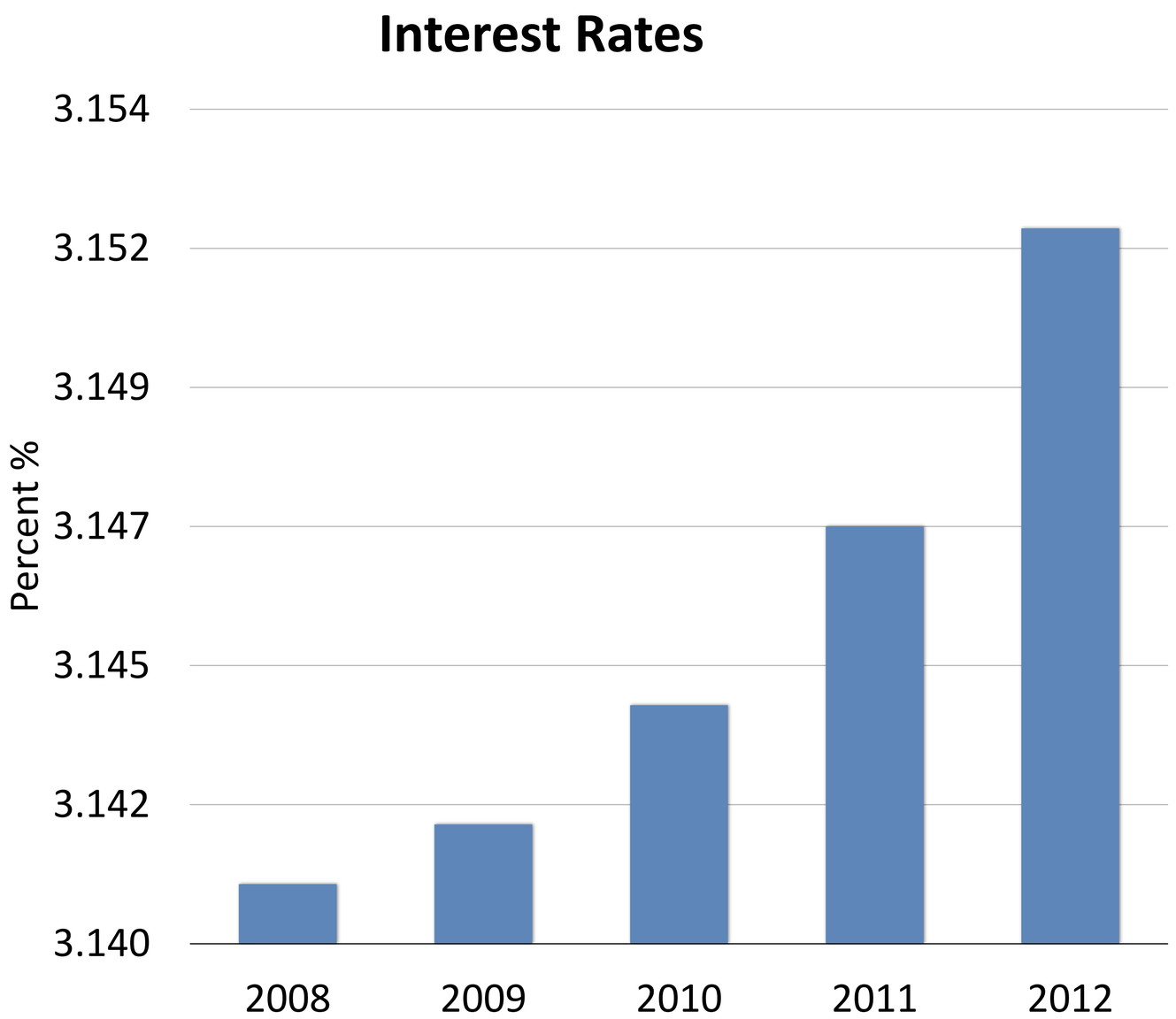
$3,549,385

$(11,014)

y-axis baseline?!

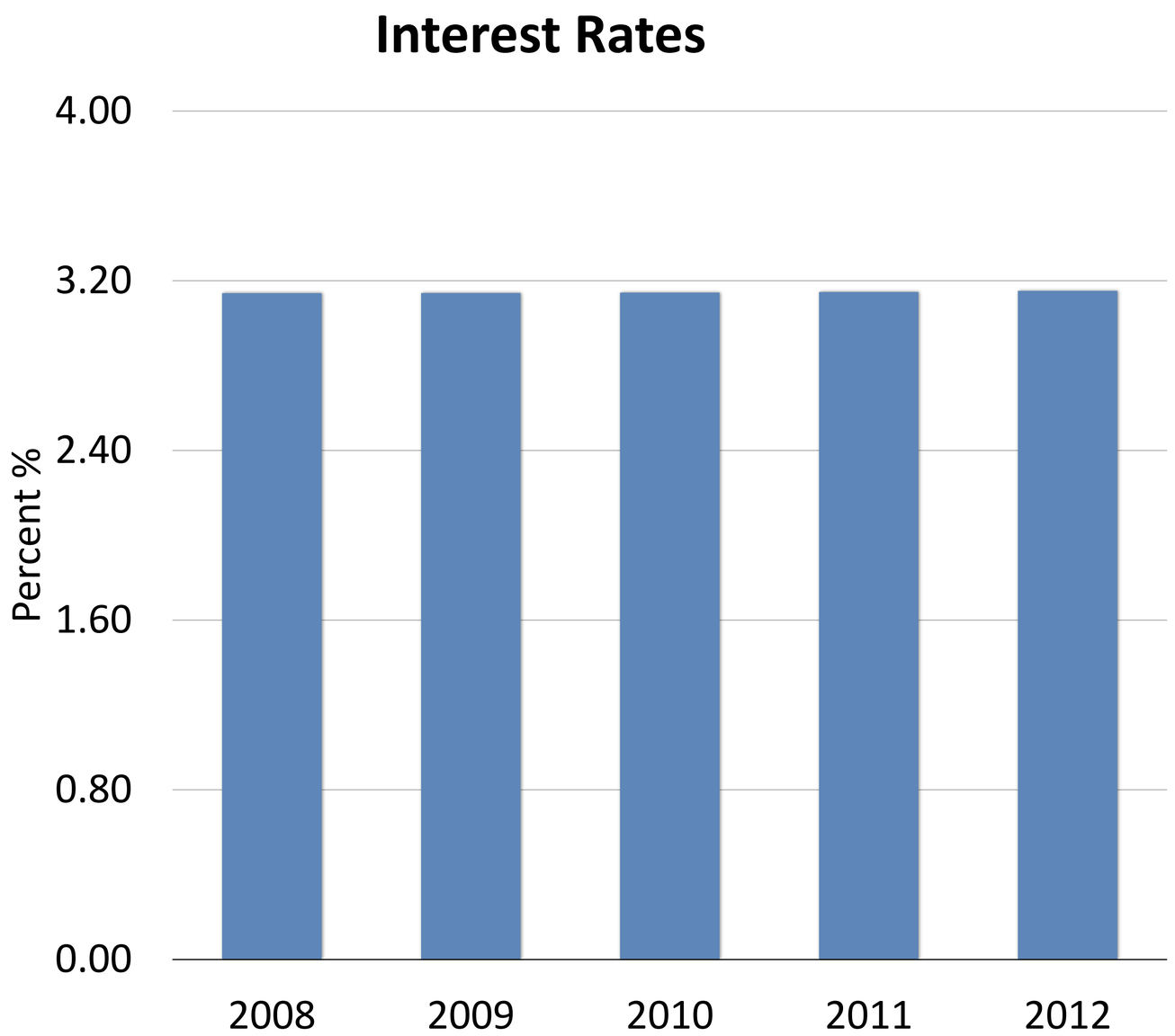
“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

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



"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."

Based on http://data.heapanalytics.com/how-to-lie-with-data-visualization

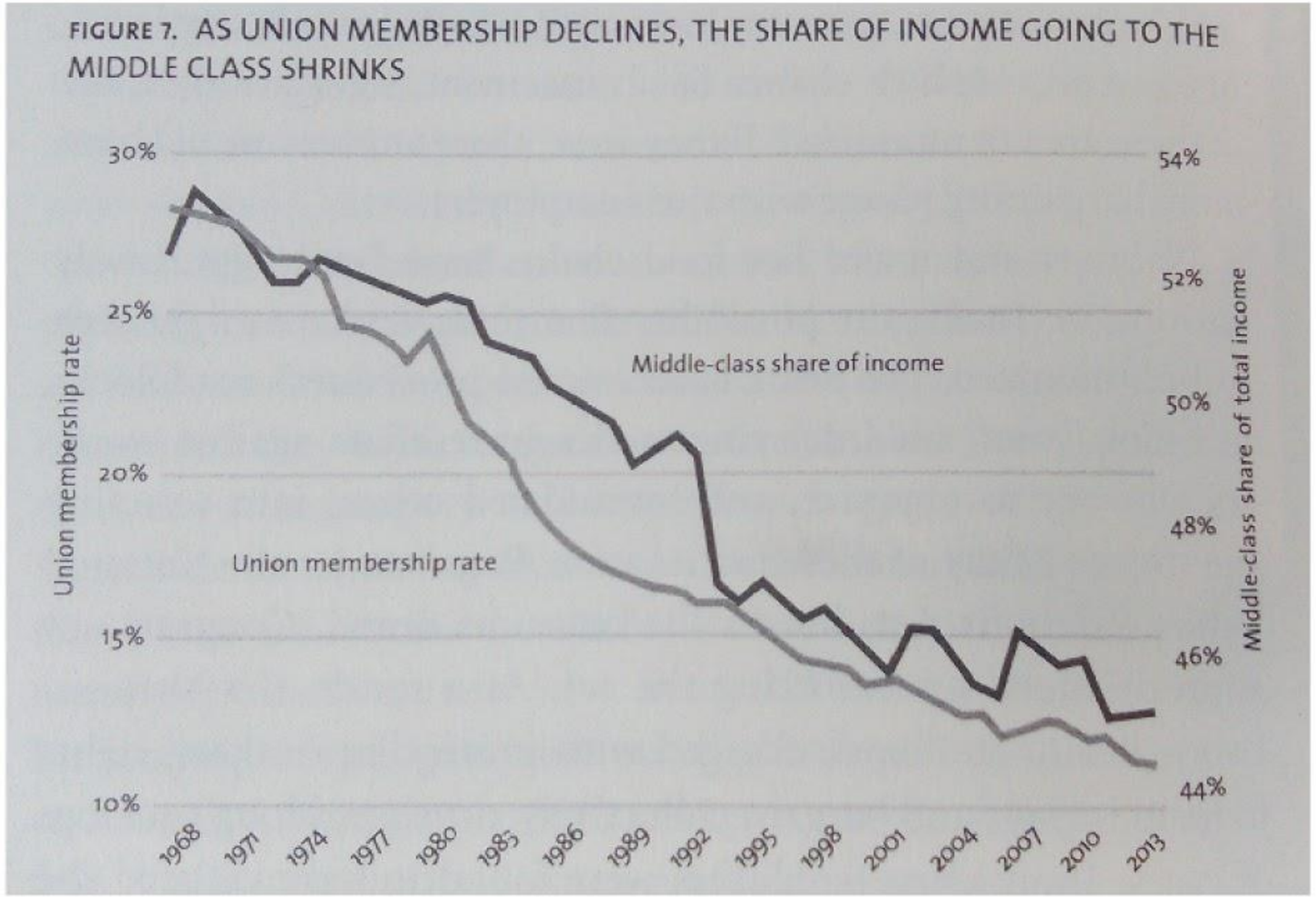


CONTEXT!

"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."

Based on http://data.heapanalytics.com/how-to-lie-with-data-visualization

"Double the axes, double the mischief"

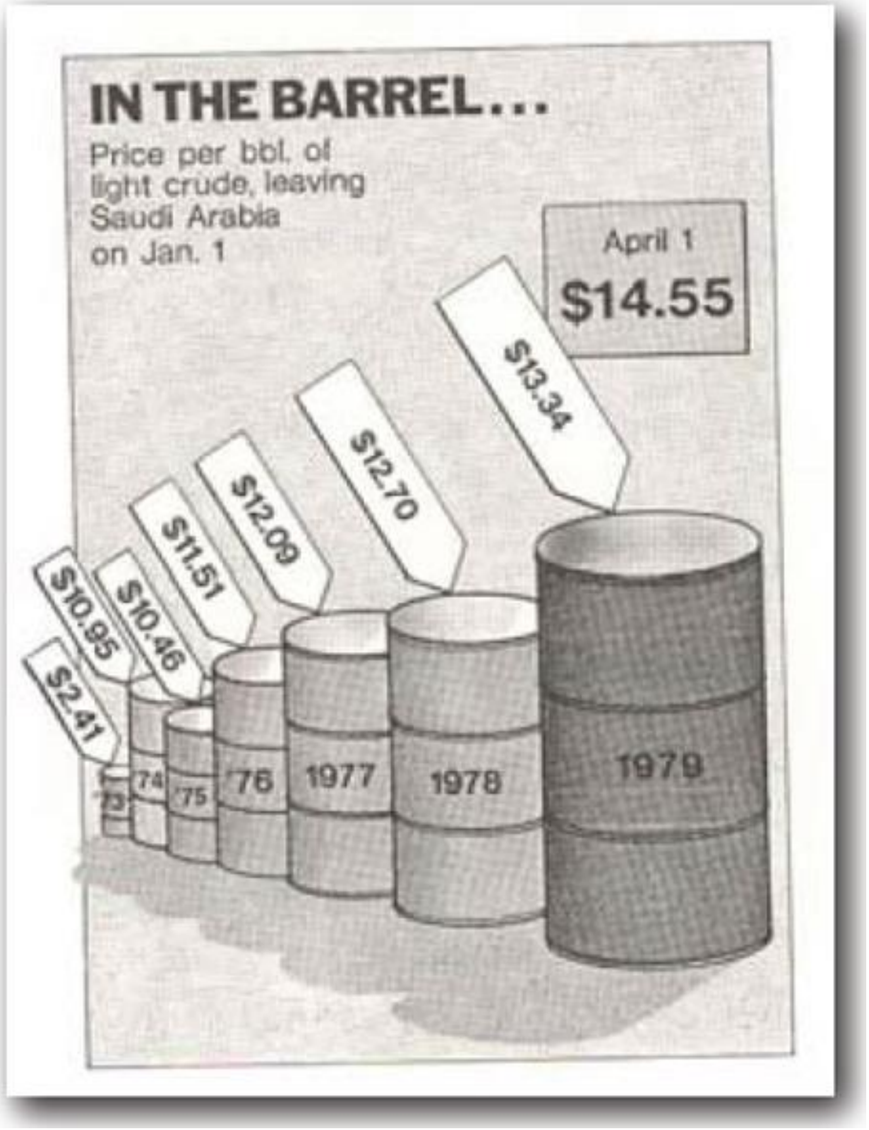


"Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data." http://www.thefunctionalart.com/2015/10/double-axes-double-mischief.html

“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)



"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)

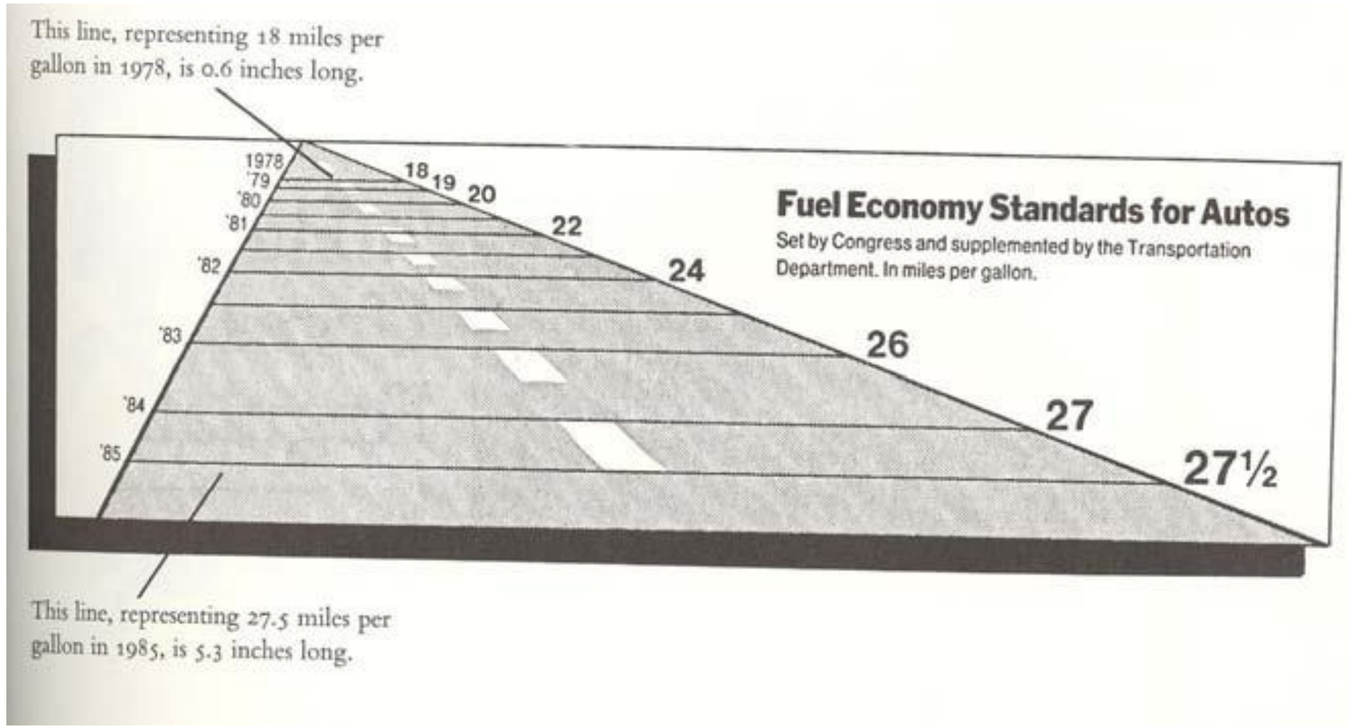
Lie Factor

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

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

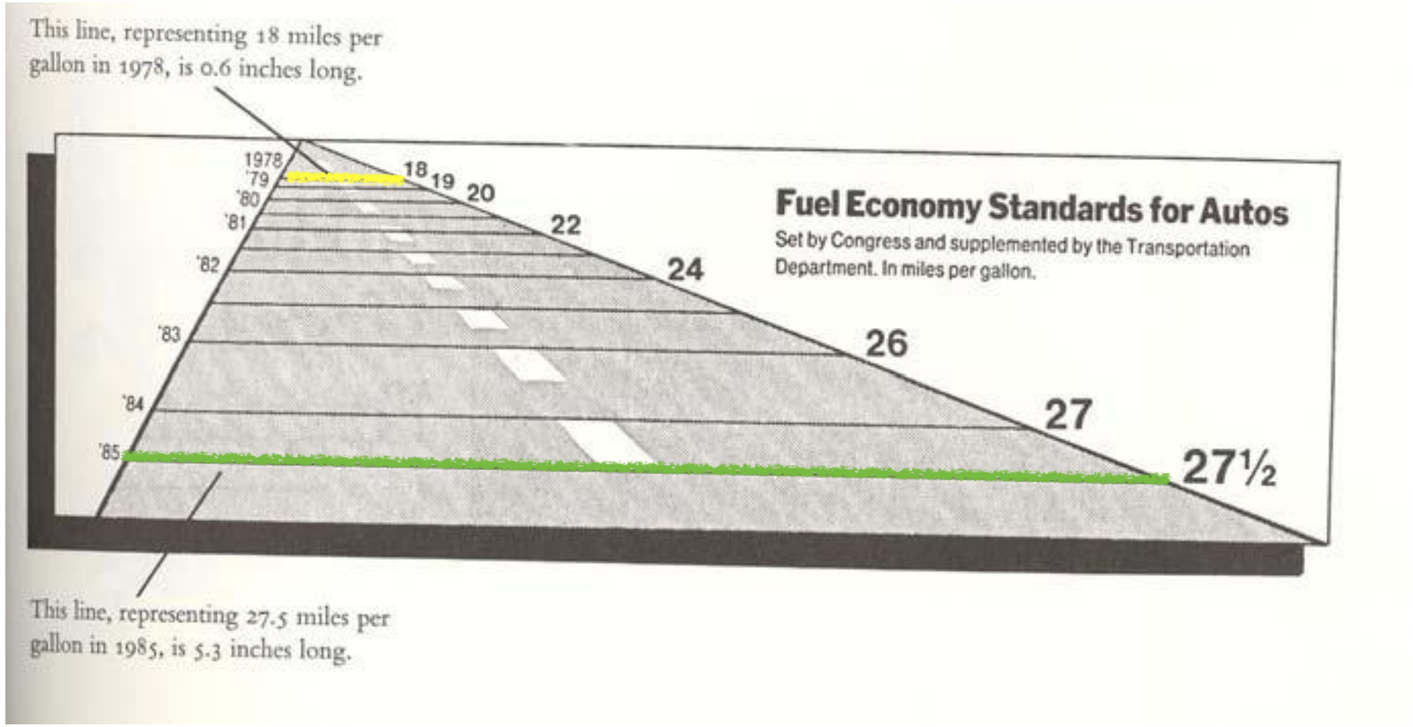


Image = $\frac{5.3" - 0.6"}{0.6"}$ = 7.83 = 783%

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

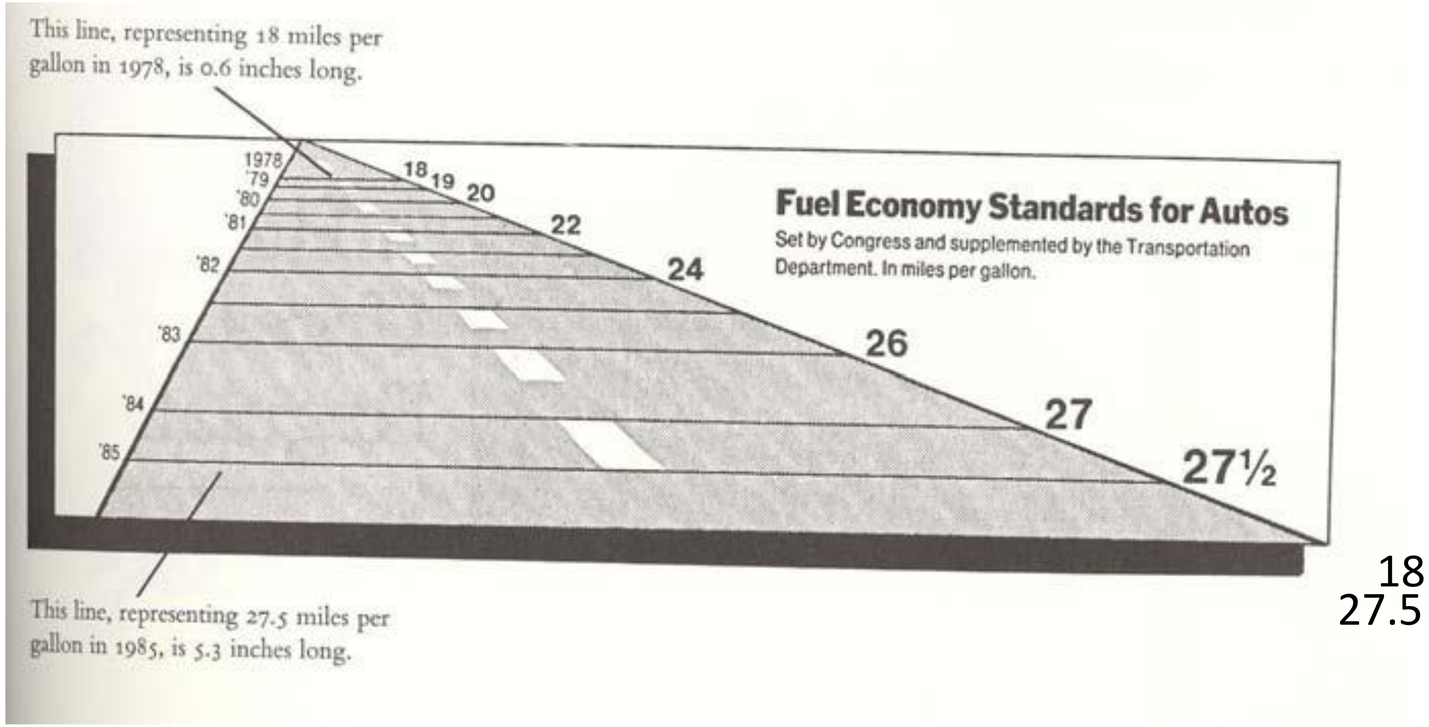


Image = $\frac{5.3" - 0.6"}{0.6"}$ = 7.83 = 783%

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

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

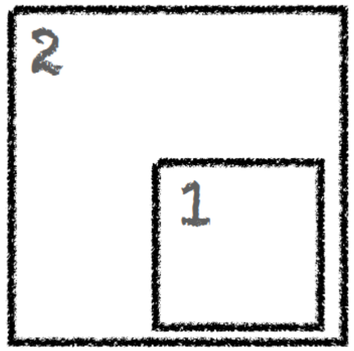
$$\text{Lie Factor} = \frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$$

✓

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

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

Make sure *area* is proportional to data!

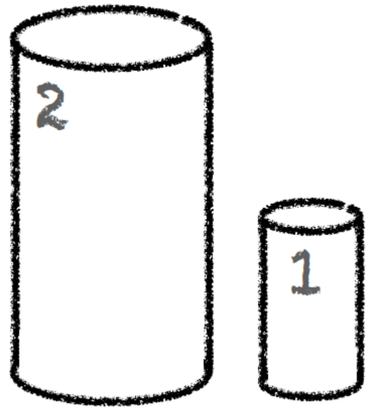


X

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

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

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)

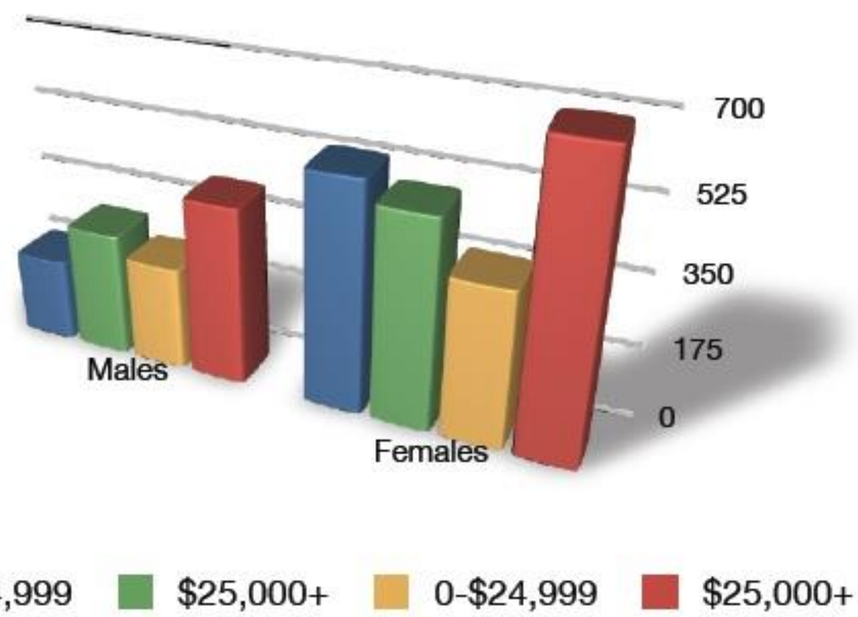
“Graphical Integrity”

Data Ink = the ink used to show data

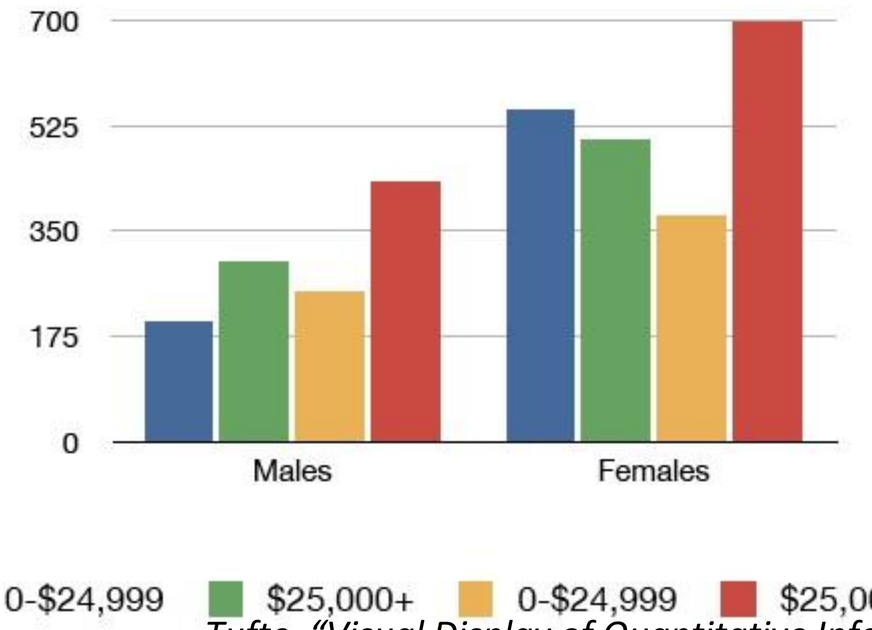
Data Ink Ratio = $\frac{\text{data-ink}}{\text{total ink in graphic}}$

Tufte: maximize the data ink ratio

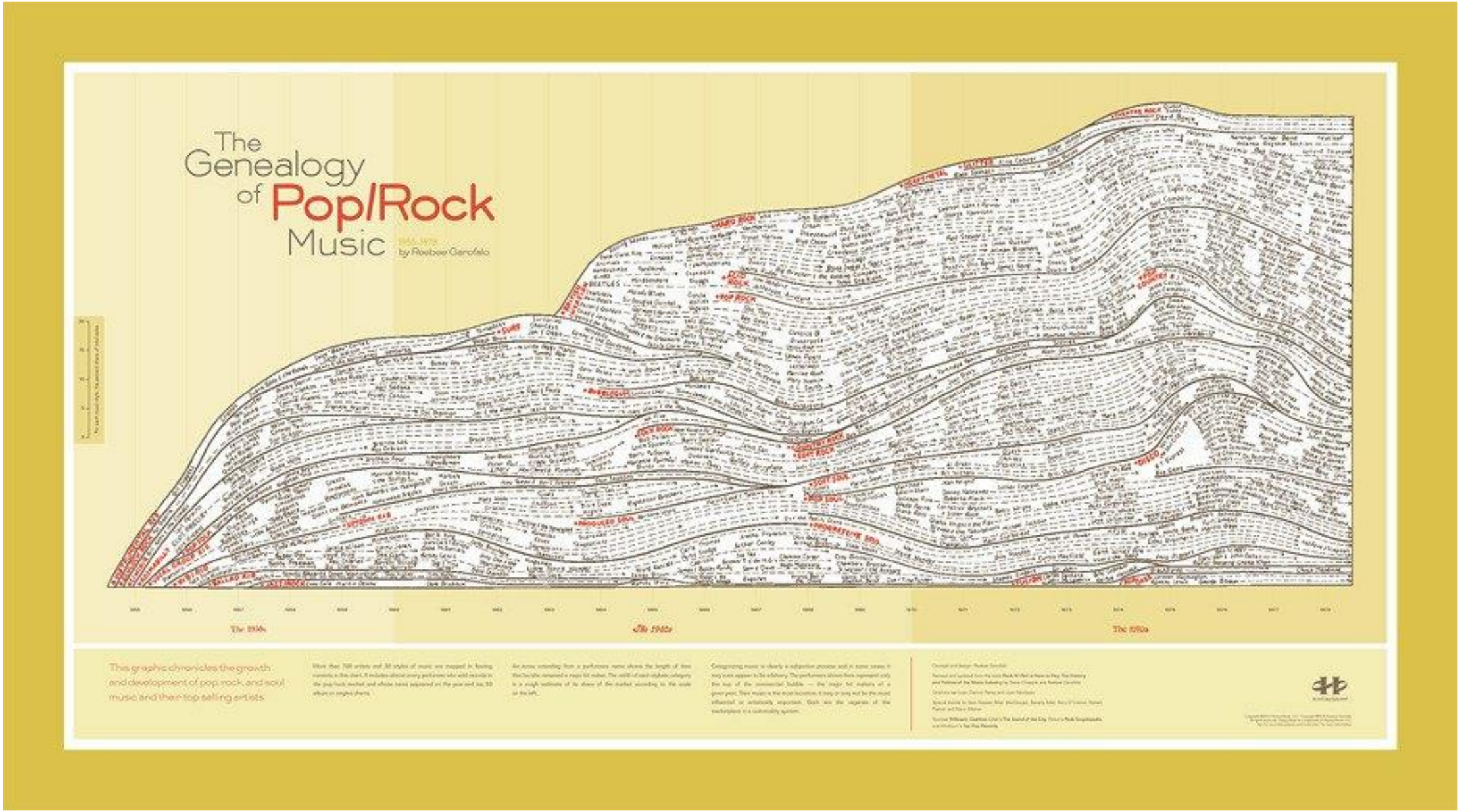
Low Data Ink Ratio



High Data Ink Ratio



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



High Data Ink Ratio

[Reebee Garofalo, Genealogy of Pop/Rock Music]

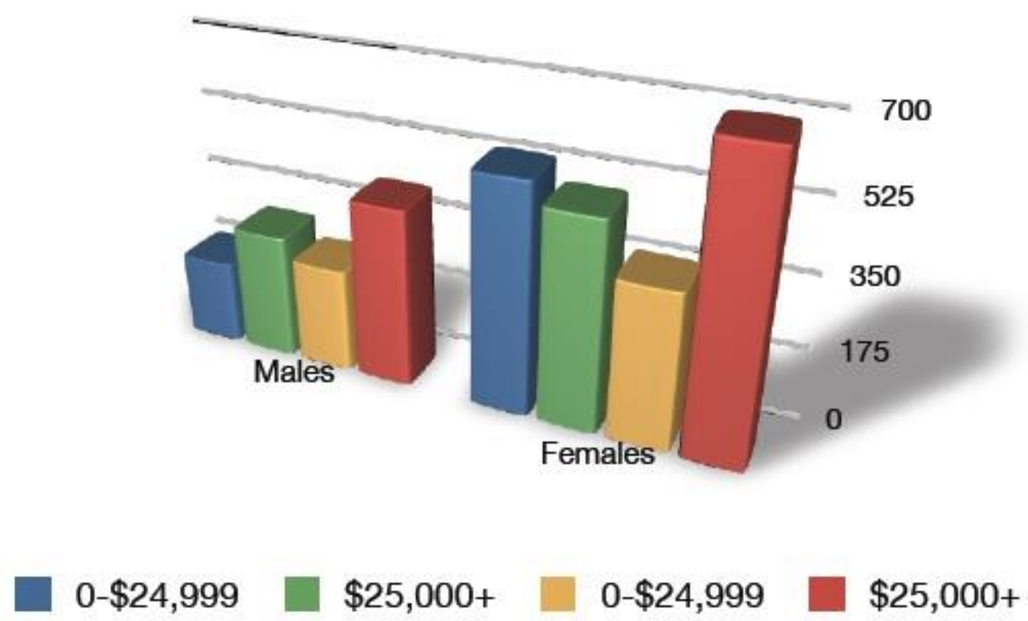
“Graphical Integrity”

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

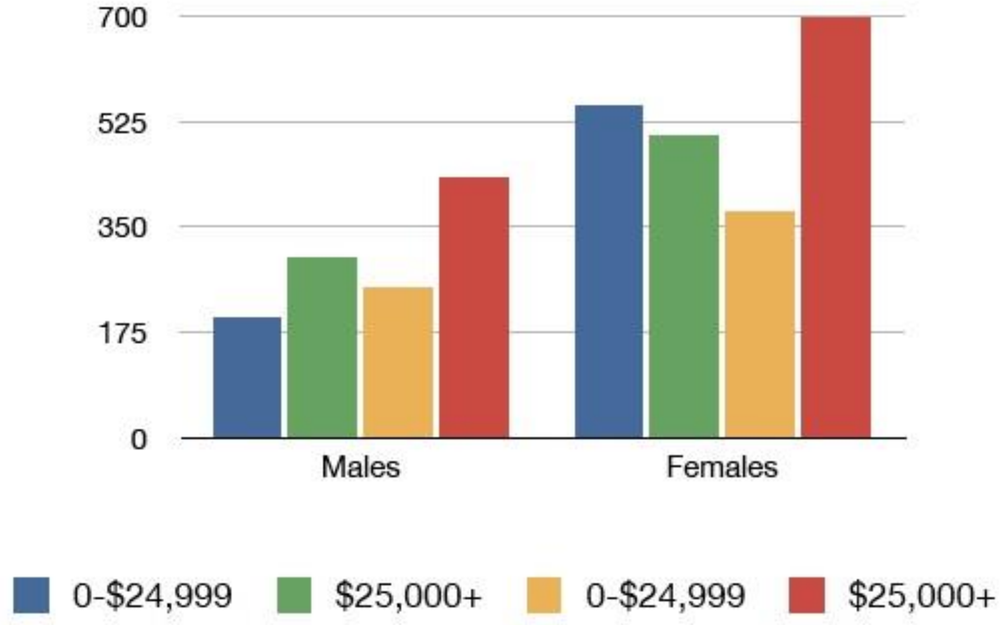
Tufte, “Visual Display of Quantitative Information” (1983)

“No Unjustified 3D”

Dimensions in data: 3
Dimensions in plot: 4



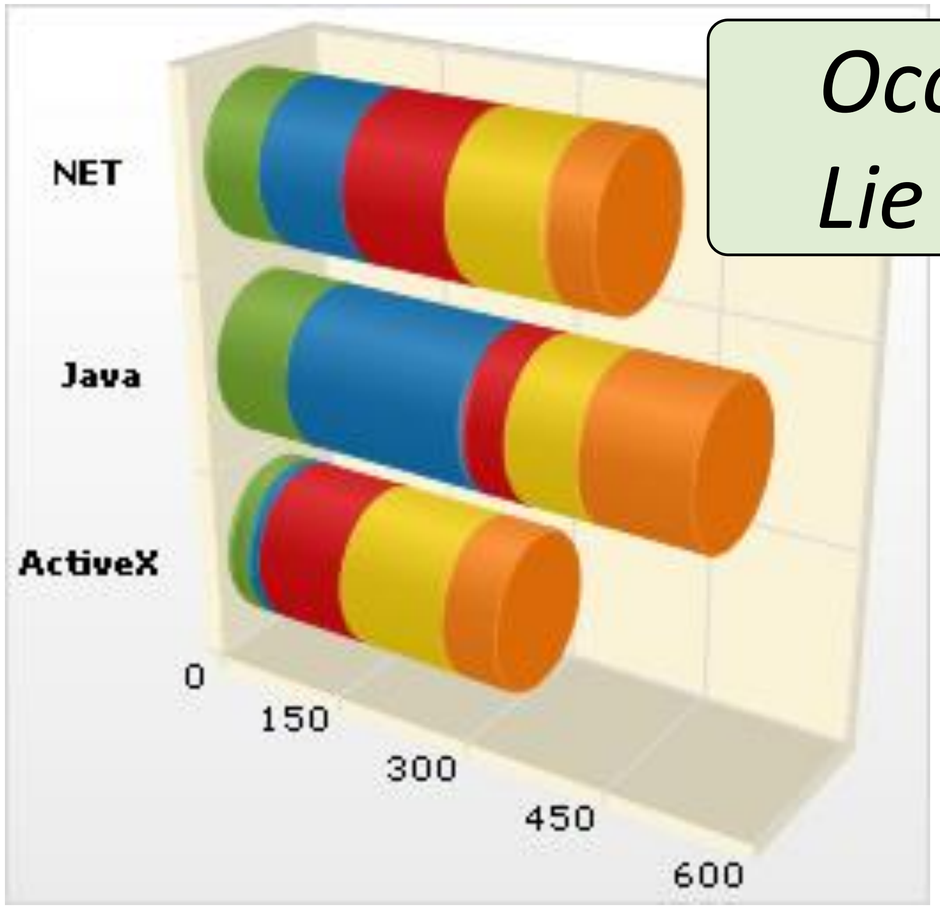
Dimensions in data: 3
Dimensions in plot: 3



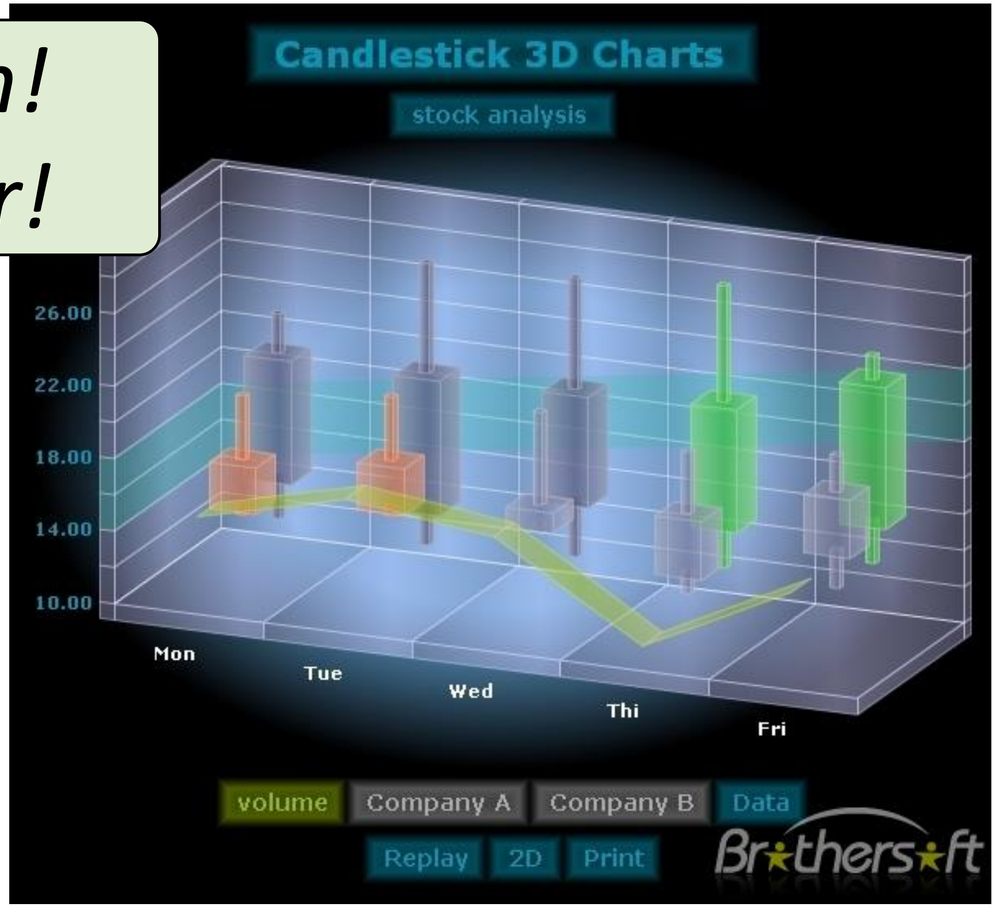
“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

“No Unjustified 3D”

Occlusion!
Lie Factor!



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

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

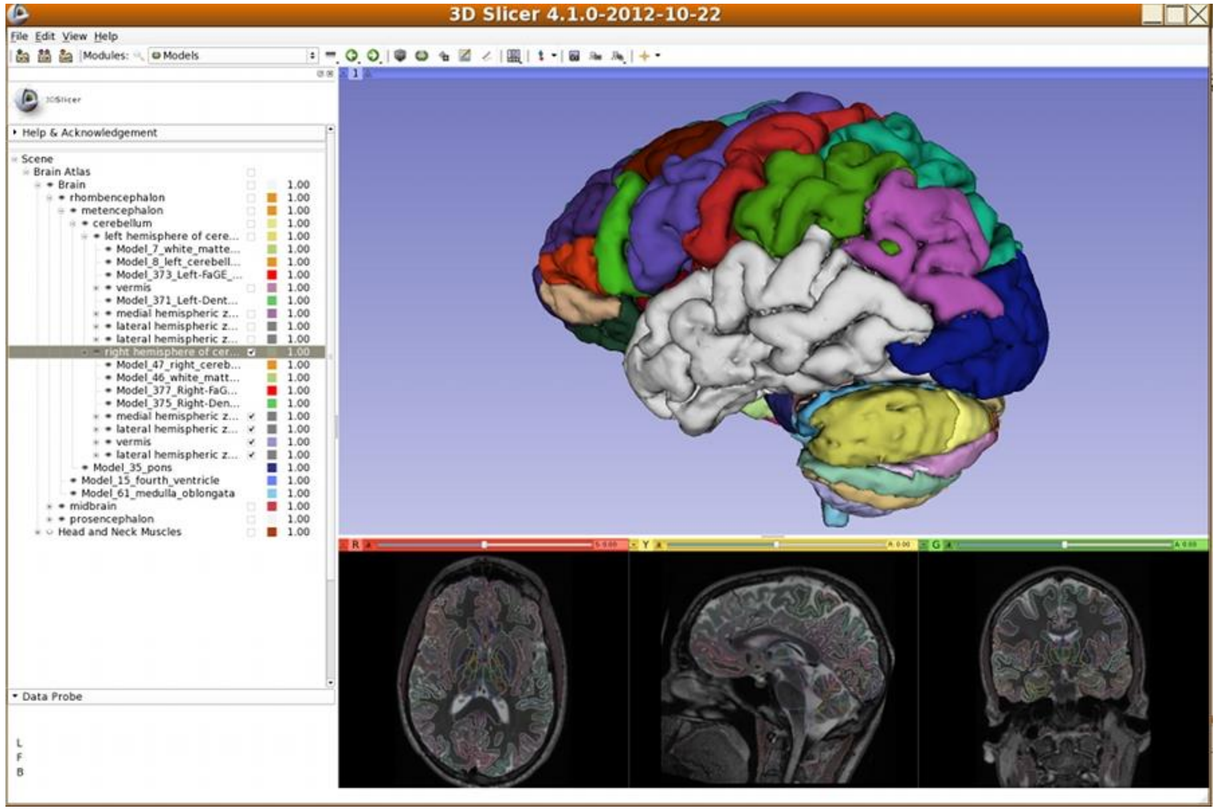
“No Unjustified 3D”

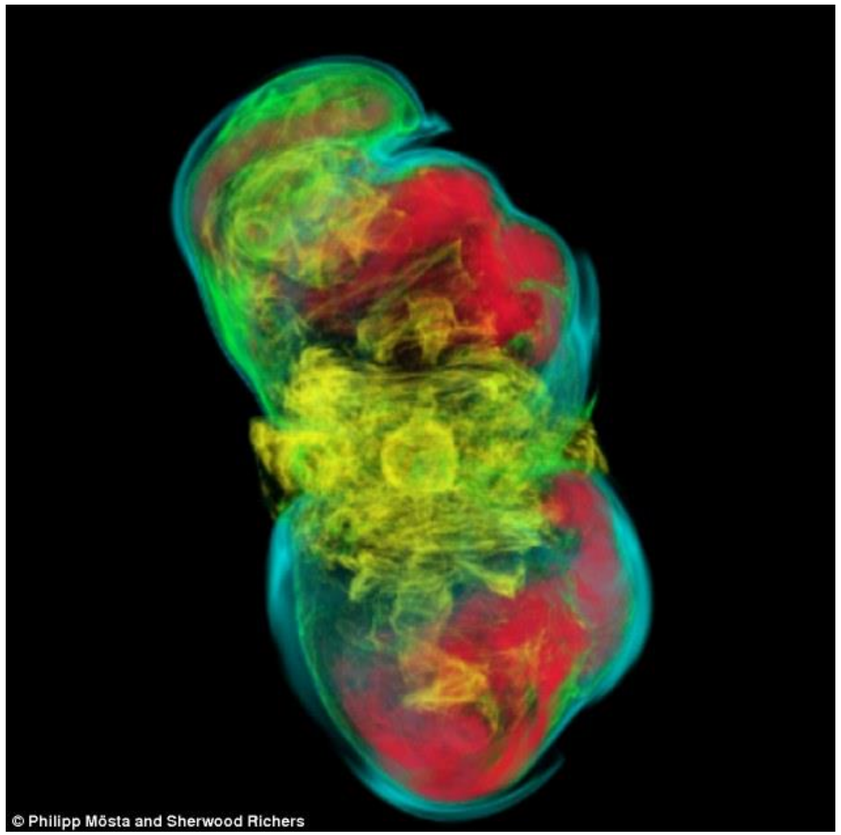
Unjustified 3D!

Lie factor!

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

“No Unjustified 3D”



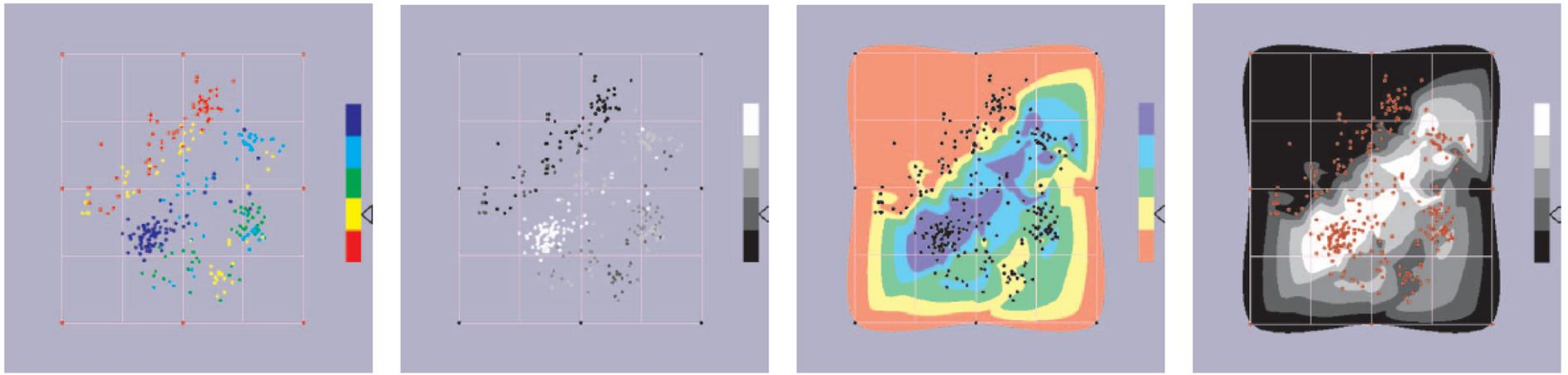


“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

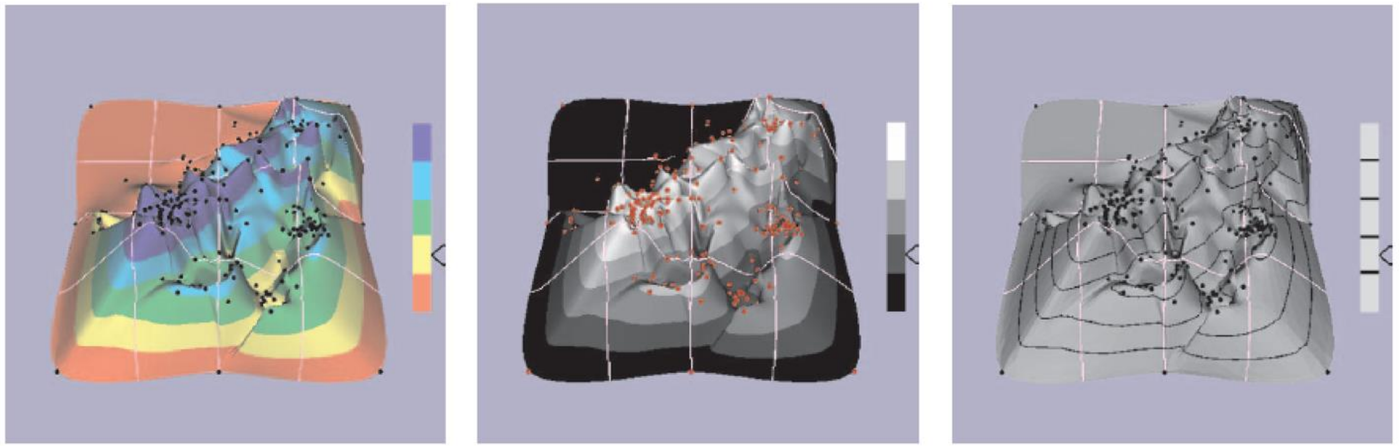
“No Unjustified 3D”

This is not just a design principle, it has lots of experimental and quantitative data to back it up!

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



(a) Coloured Points (b) Greyscale Points (c) Coloured 2D Landscape (d) Greyscale 2D Landscape



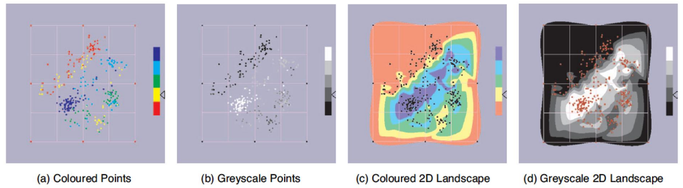
(e) Coloured 3D Landscape (f) Greyscale 3D Landscape (g) Height-only

Fig. 1 Point-based displays and information landscapes used in our experiment. All displays show the same data.

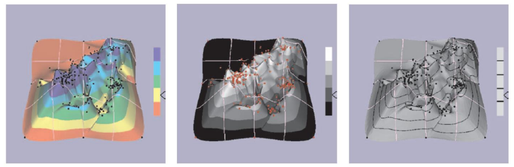
Dr. David Sprague (Former Lecturer, Khoury)

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

Tory, et al. (2007)



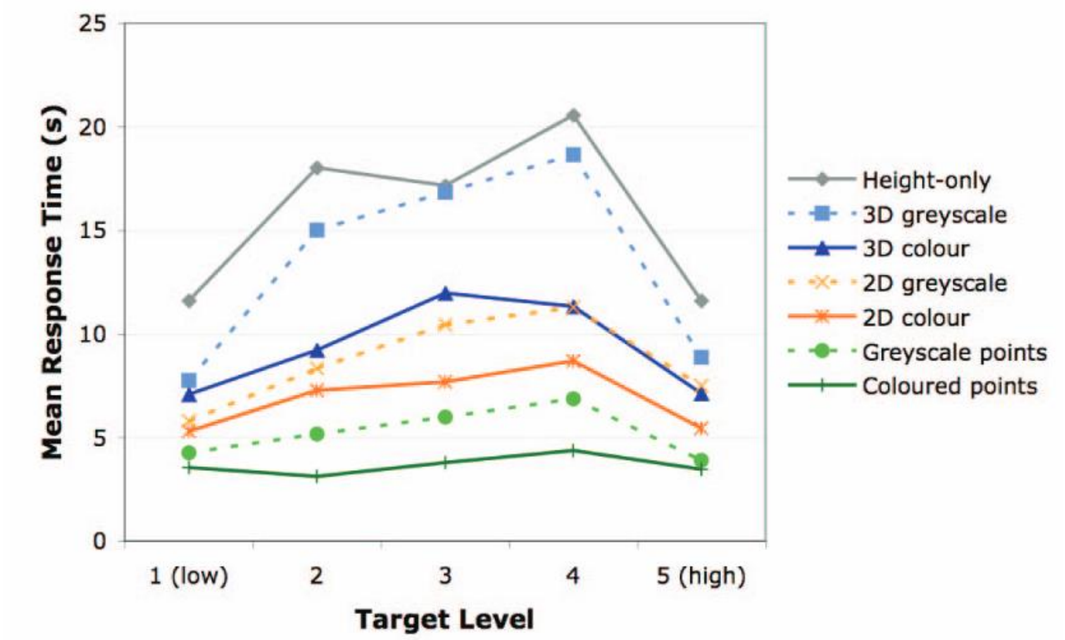
(a) Coloured Points (b) Greyscale Points (c) Coloured 2D Landscape (d) Greyscale 2D Landscape

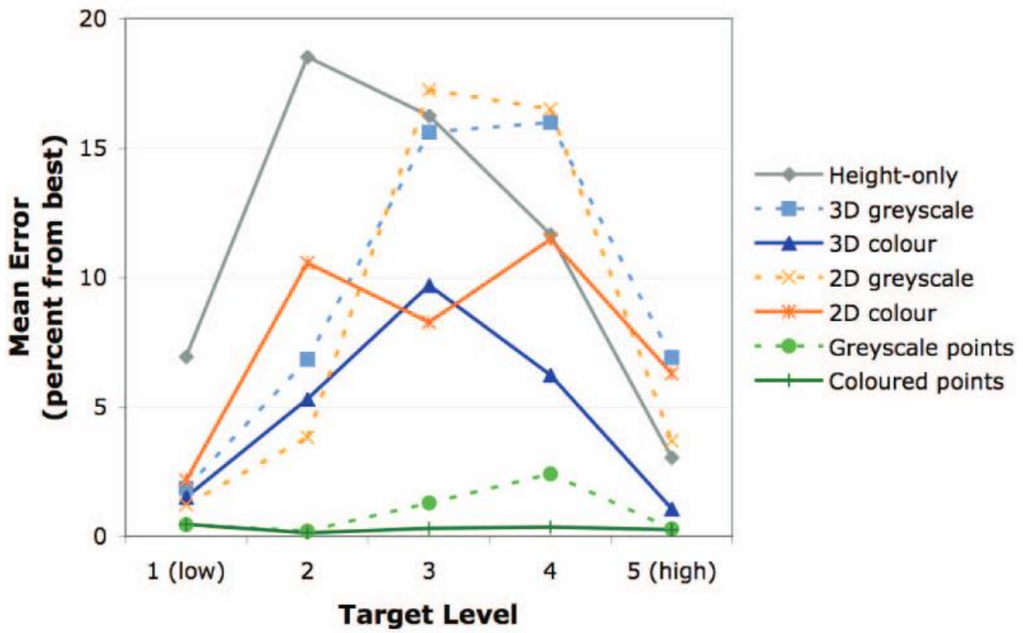


(e) Coloured 3D Landscape (f) Greyscale 3D Landscape (g) Height-only

Fig. 1 Point-based displays and information landscapes used in our experiment. All displays show the same data.

“Which spatial area contained the most points of a specified target value range?”





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

Tory, et al. (2007)

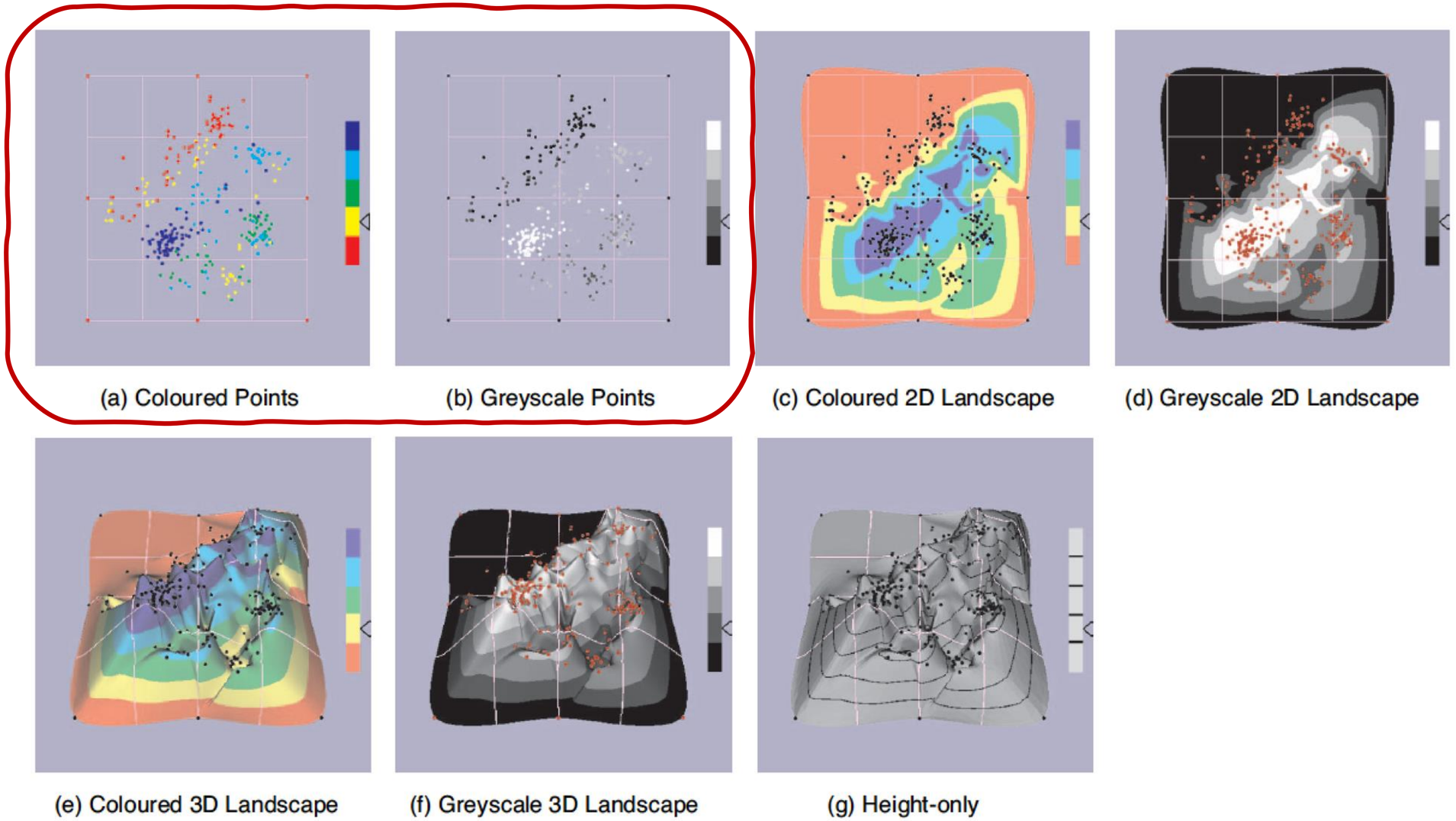
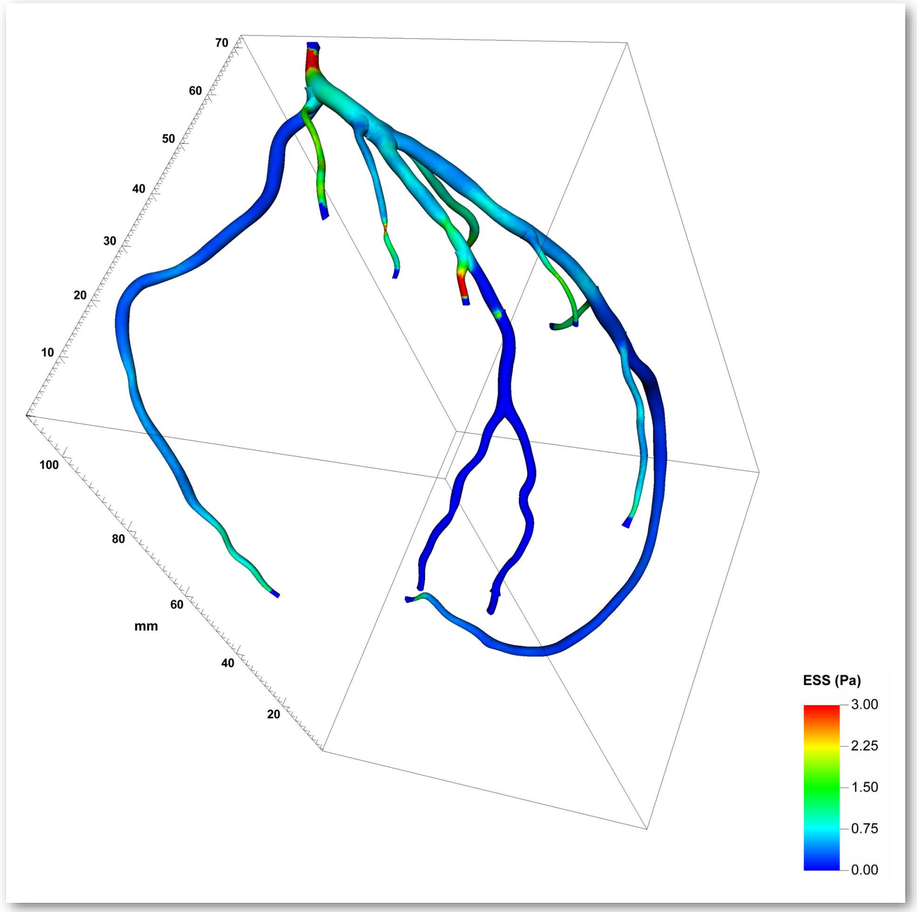


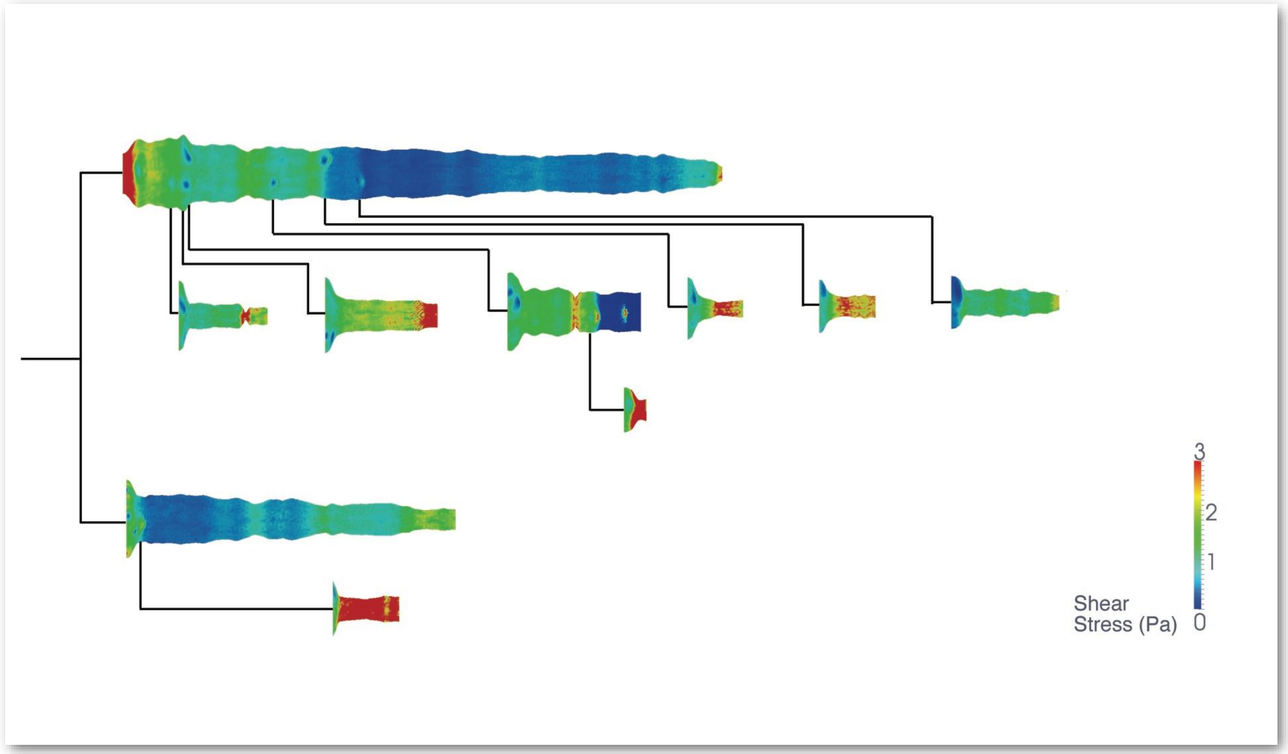
Fig. 1 Point-based displays and information landscapes used in our experiment. All displays show the same data.

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

Tory, et al. (2007)

“No Unjustified 3D”





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

Borkin, et al. (2011)

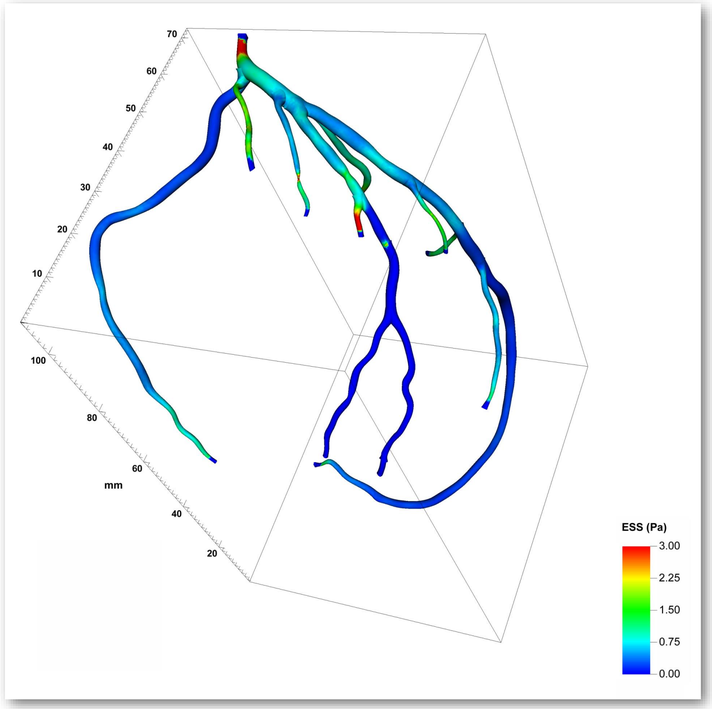
ACCURACY

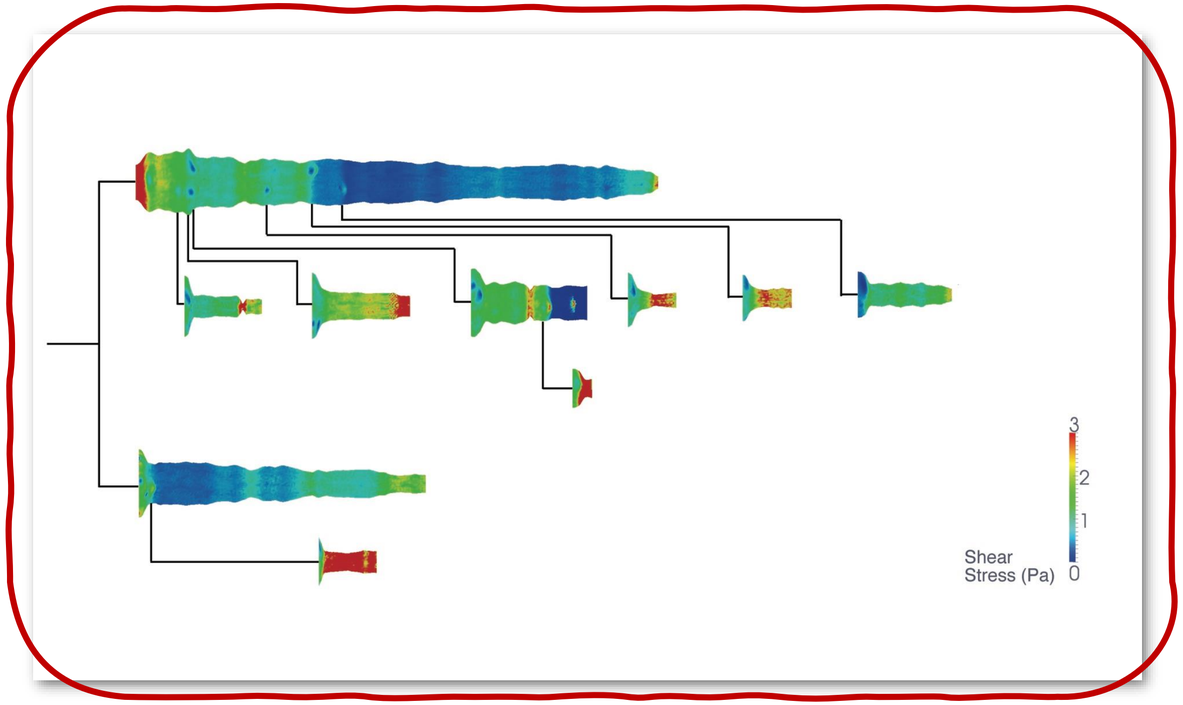
Strong effect of **dimensionality** on accuracy

39%

How many diseased regions found?

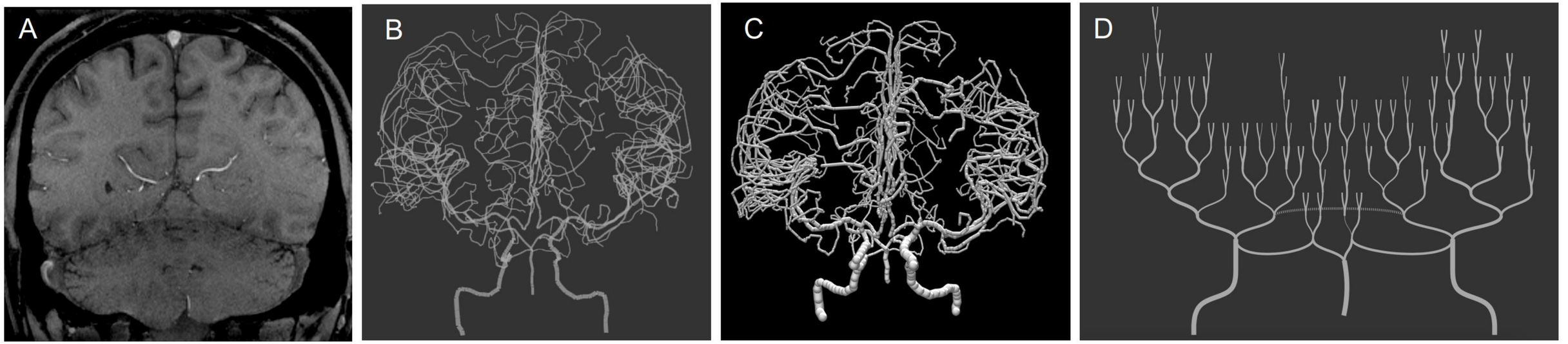
62%





Borkin, et al. (2011)

“No Unjustified 3D”



Data Source

MIPS: 2D projection of 3D Arteries

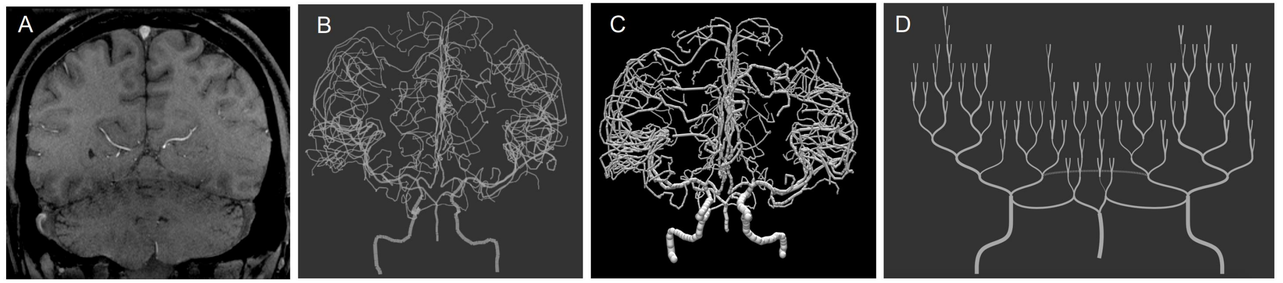
3D Rendering

CerebroVis: Novel Network Visualization

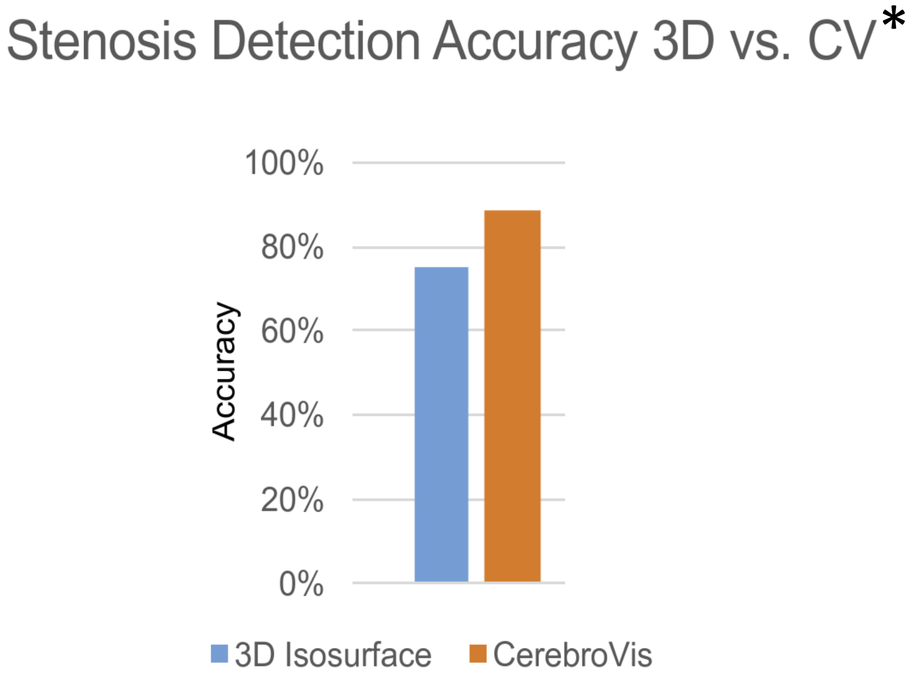
“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

Pandey et al. VIS 2019

“No Unjustified 3D”



Data Source | **MIPS: 2D projection of 3D Arteries** | **3D Rendering** | **CerebroVis: Novel Network Visualization**

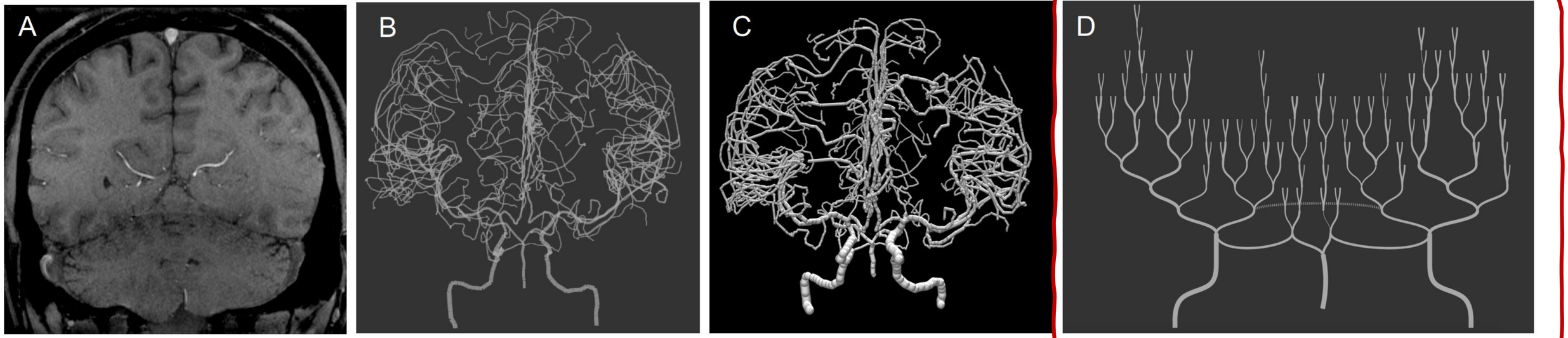


** Only 3 neuroradiologists tested, but also iterative design with feedback at each step.*

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

Pandey et al. VIS 2019

“No Unjustified 3D”



Data Source

MIPS: 2D projection of 3D Arteries

3D Rendering

CerebroVis: Novel Network Visualization

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

Pandey et al. VIS 2019

“Graphical Integrity”

To achieve graphical “excellence” according to Tufte:

1. Above all else show the data.
2. Maximize the data-ink ratio.
3. Erase non-data ink.
4. Erase redundant data ink.
5. Revise and edit.

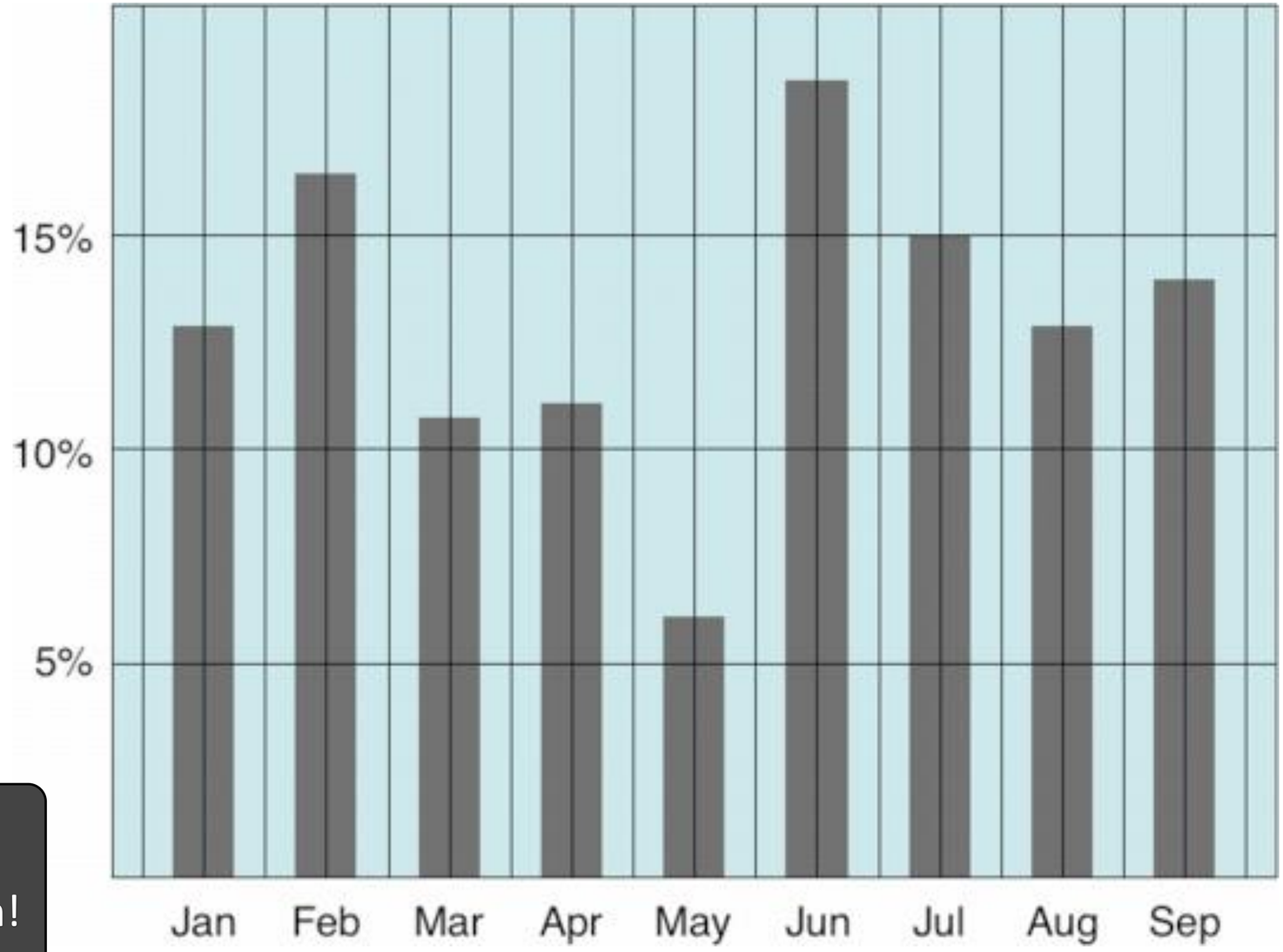
Tufte, “Visual Display of Quantitative Information” (1983)

IN-CLASS EXERCISE

(No submission)

In-Class Sketching — “Graphical Integrity”

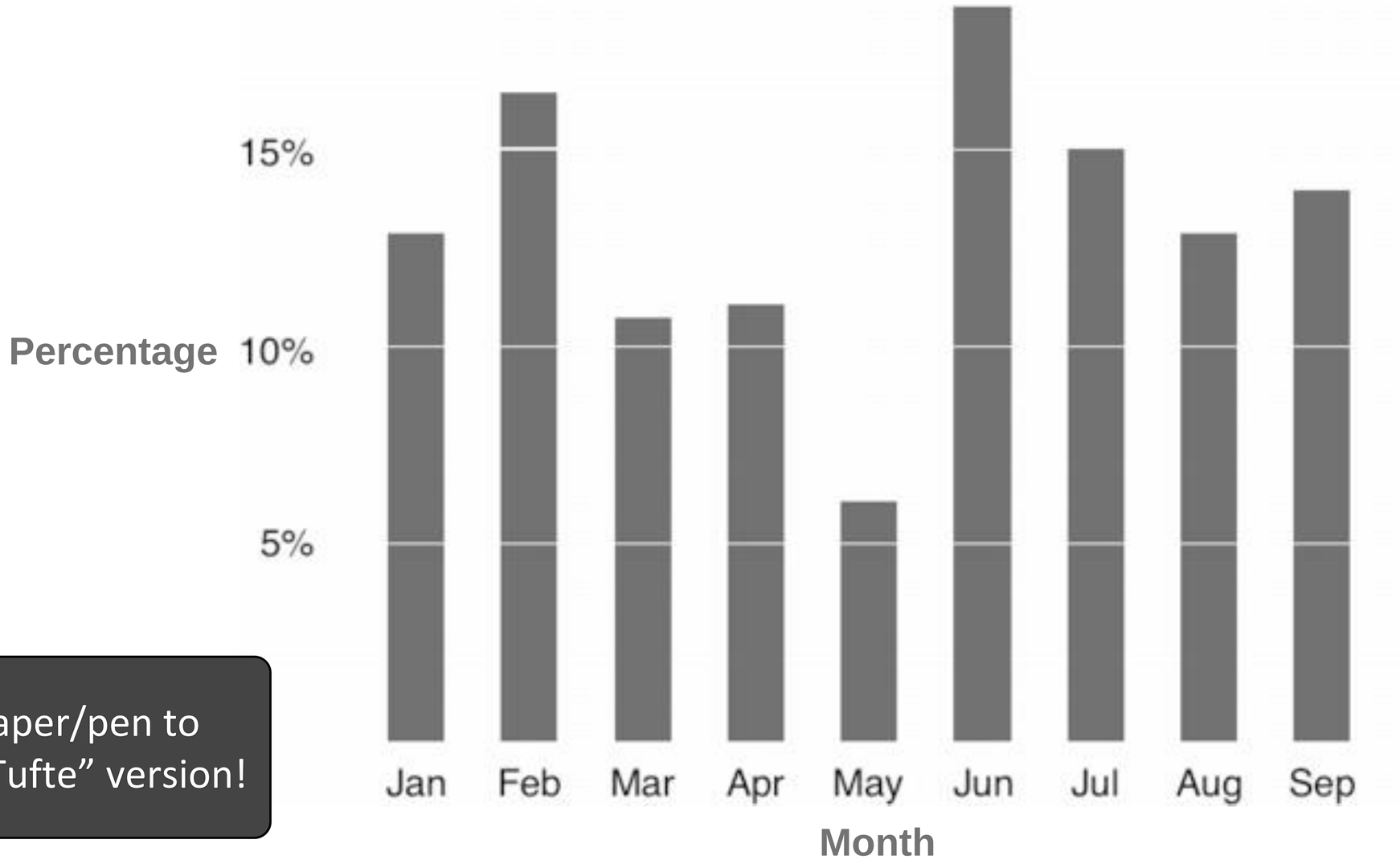
~8 min



Use paper/pen to sketch “Tufte” version!

In-Class Sketching — “Graphical Integrity”

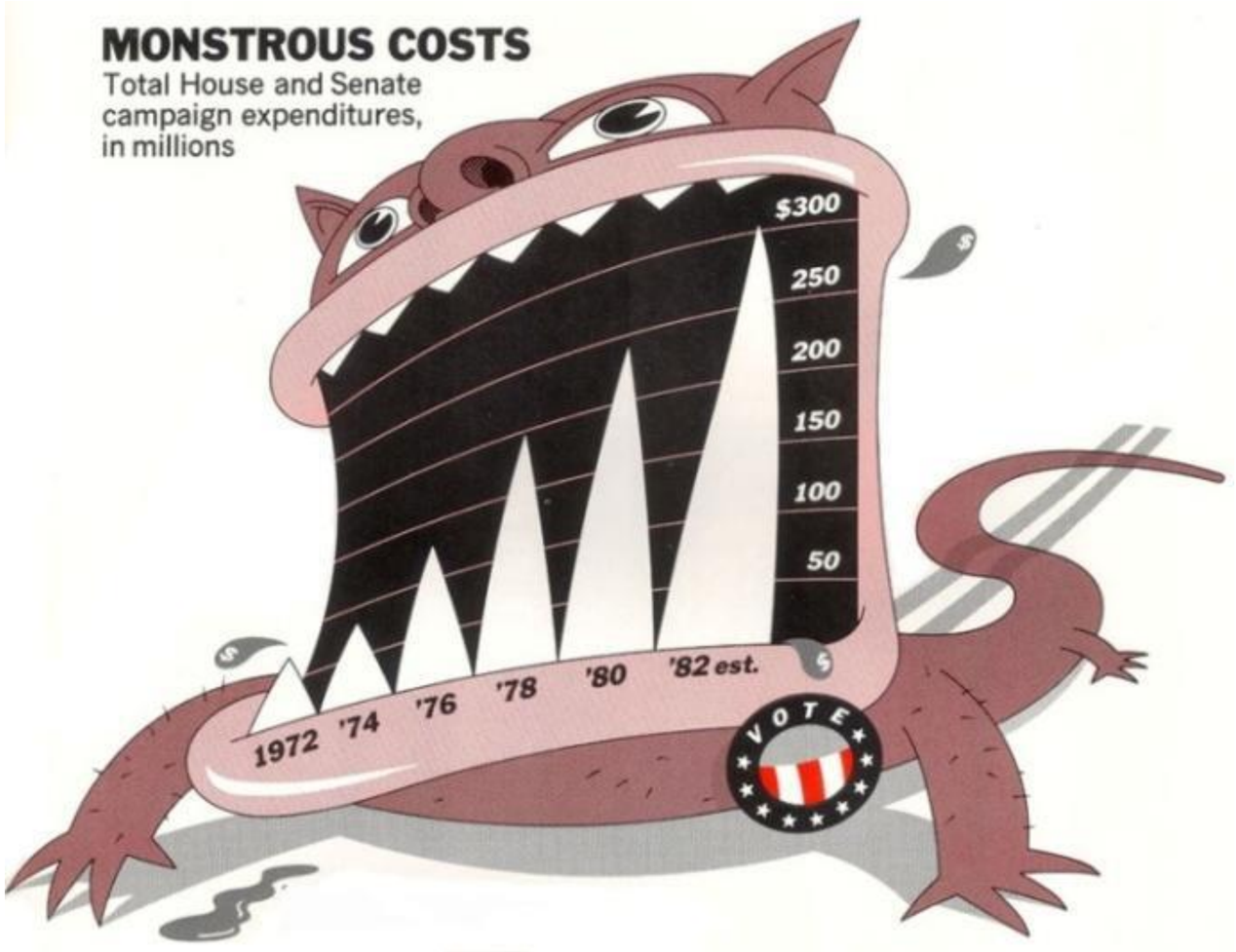
~8 min

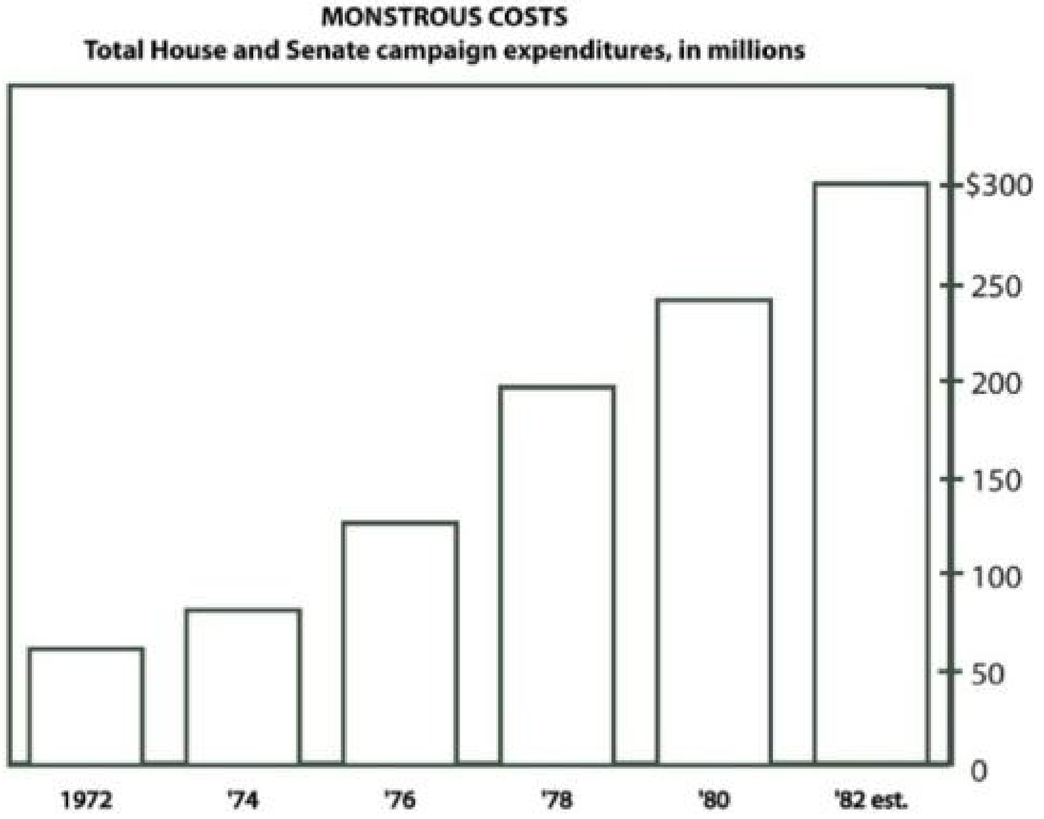


Use paper/pen to sketch “Tufte” version!

CHART JUNK

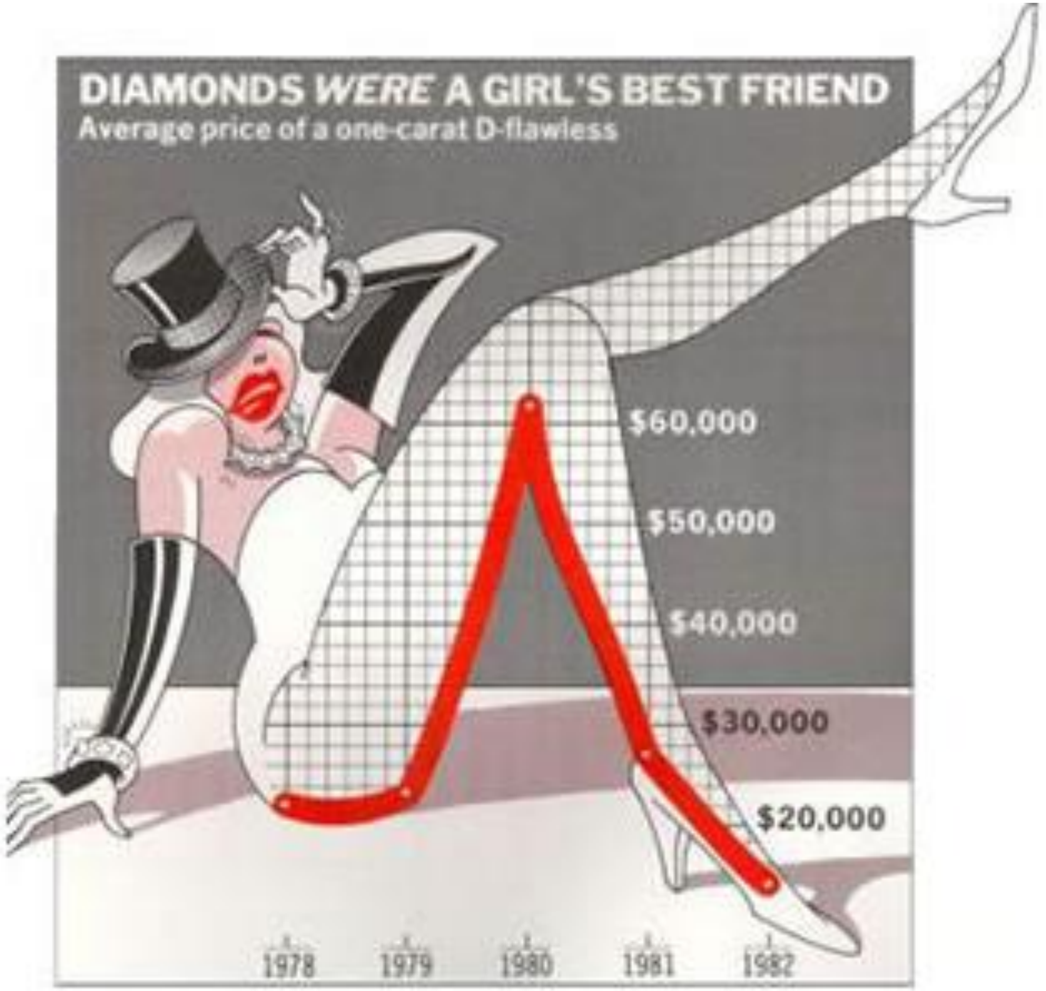
“Chart Junk”

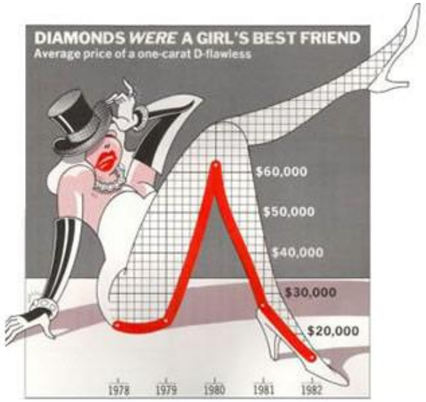


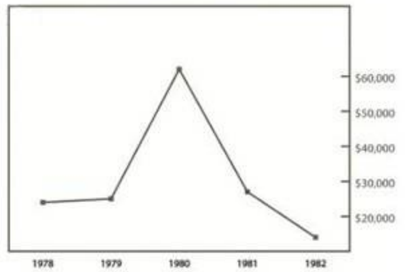


Bateman, et al. (2010)

“Chart Junk”



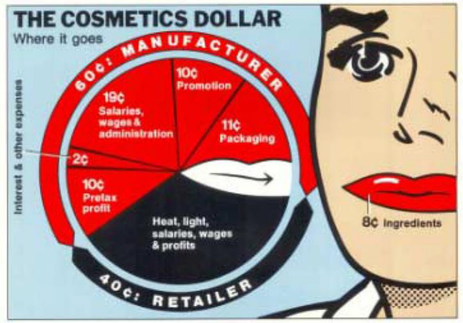


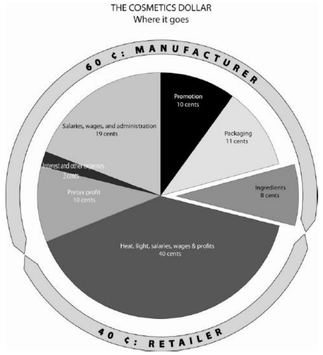


Bateman, et al. (2005)

“Chart Junk Debate”

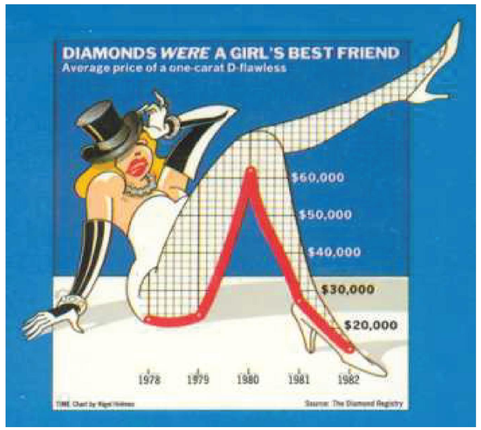
Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts





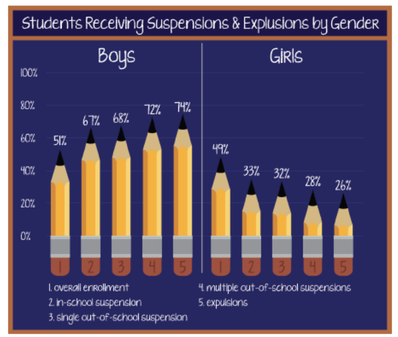
Bateman, et al. (2010)

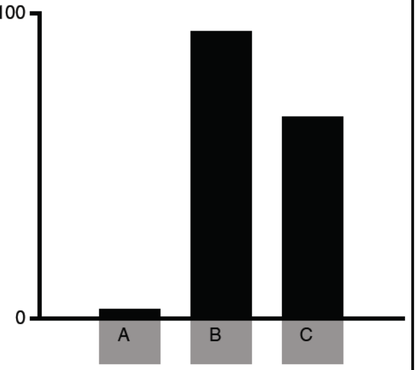
Benefitting InfoVis with Visual Difficulties



Hullman, et al. (2011)

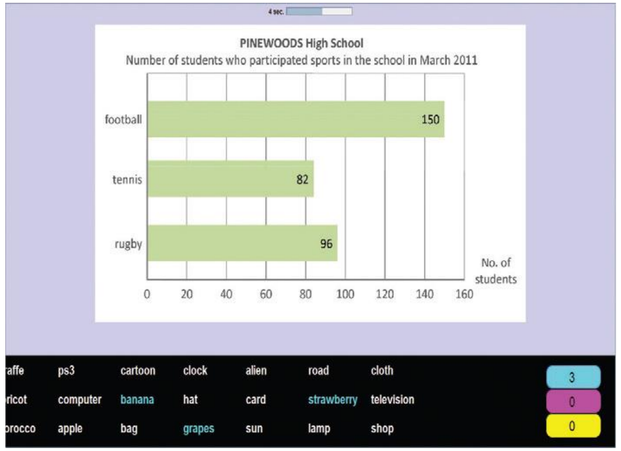
An Evaluation of the Impact of Visual Embellishments in Bar Charts





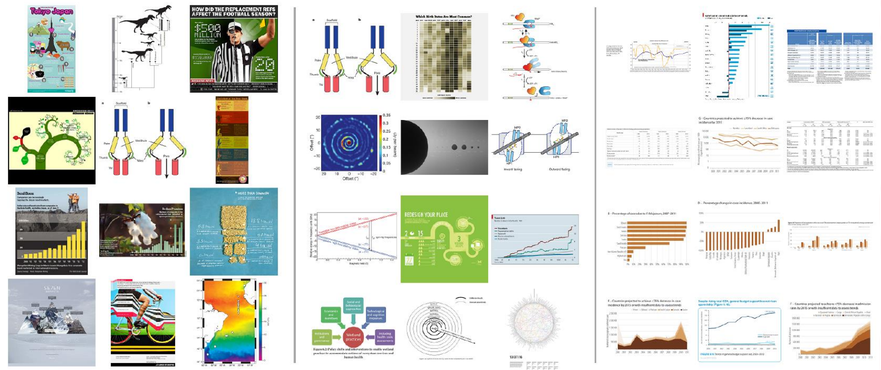
Skau, et al. (2015)

An Empirical Study on Using Visual Embellishments in Visualization



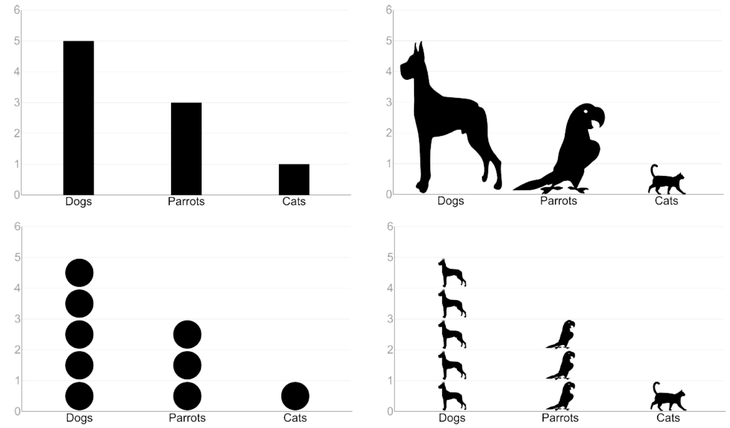
Borgo, et al. (2012)

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)

Upcoming Assignments & Communication

A look at the upcoming assignments and deadlines

- Textbook, Readings & Reading Quizzes

- 2020-09-29 (*tomorrow 11:59pm*)
  Project 1 — Initial Idea Pitches & Related Work
  (In-Class Project Pitches W)
  Assignment 3 — Critique "39 studies in 30 minutes“

- 2020-10-06
  Assignment 4a — D3 Basic Charts
  Assignment 4b — Altair & JupyterLab Setup
  (Altair & Jupyter Lab Tutorial W)
  Assignment 4c — Register for IEEE VIS 2020
  Project 2 — Proposal, Related Work, & Group Charter

- 2020-10-13
  Assignment 5 — Altair Basic Plots (available soon)

- 2020-10-20
  Assignment 6 — D3 Event Handling (available soon)
  Project 3 — Interview & Task Analysis

https://c.dunne.dev/ds4200f20

Everyday Required Supplies:

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

Use **Canvas Discussions** for general questions,
email the **instructor & TAs** for questions **specific to you**.