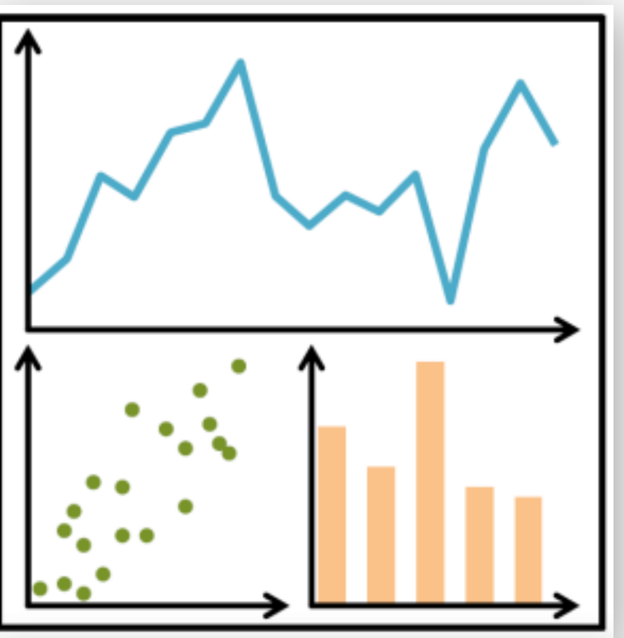
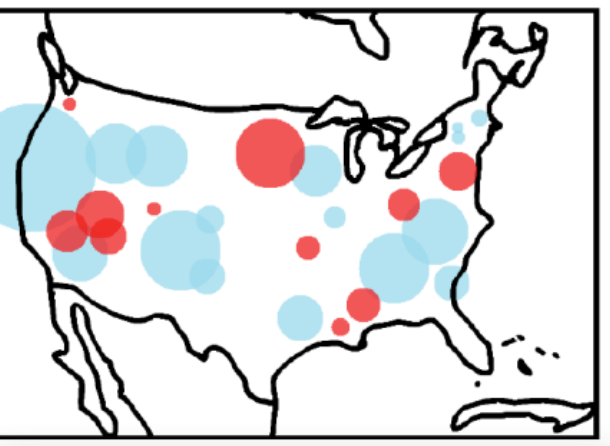
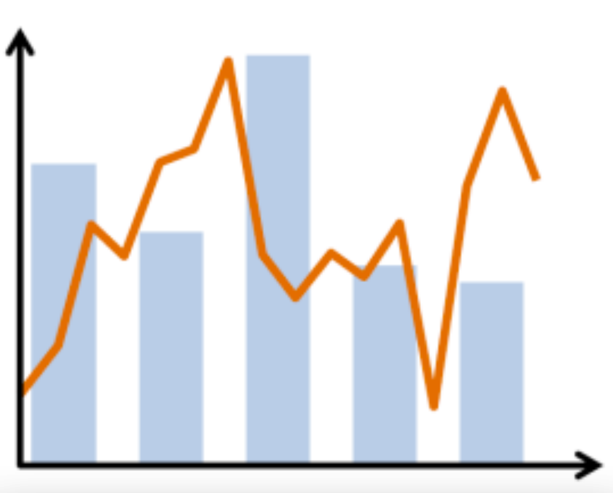
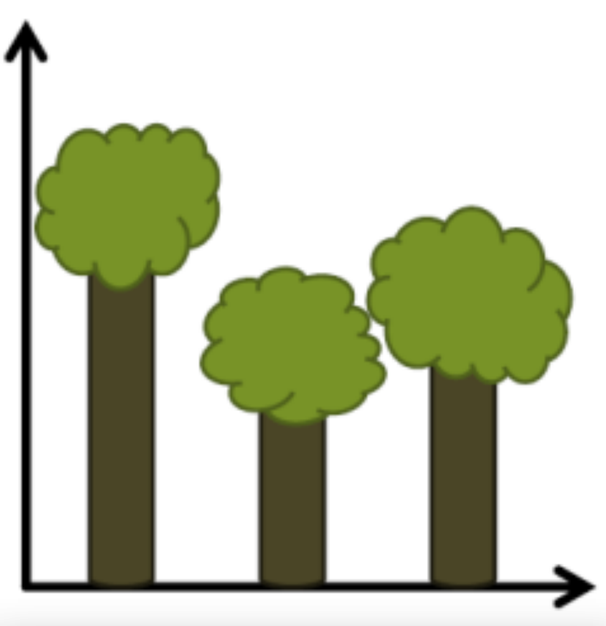
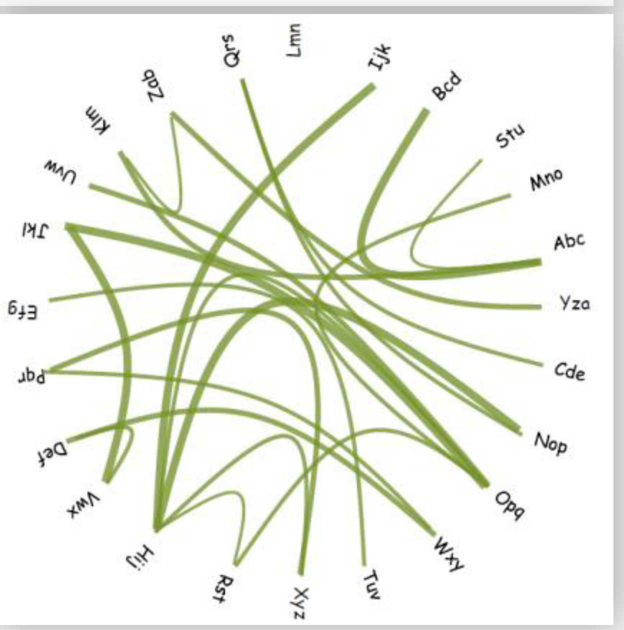
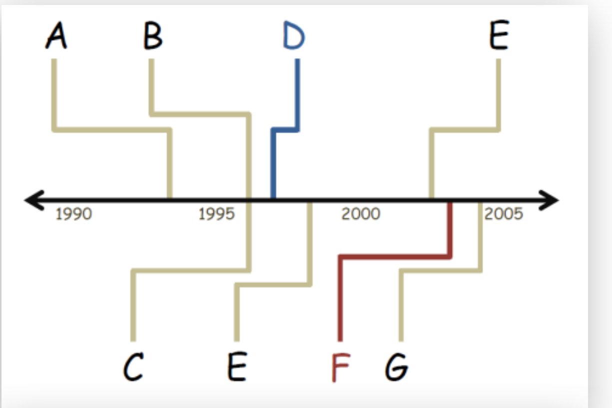
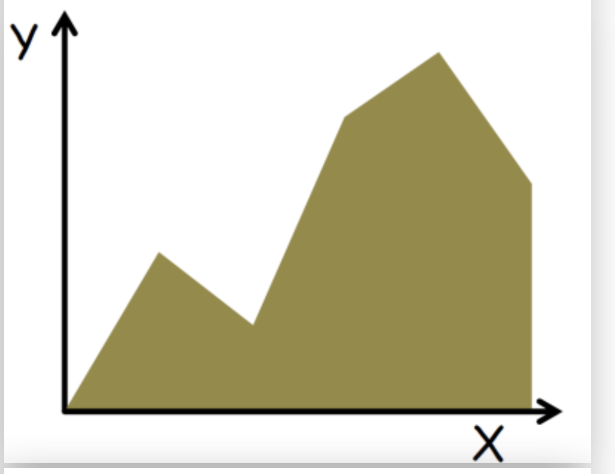






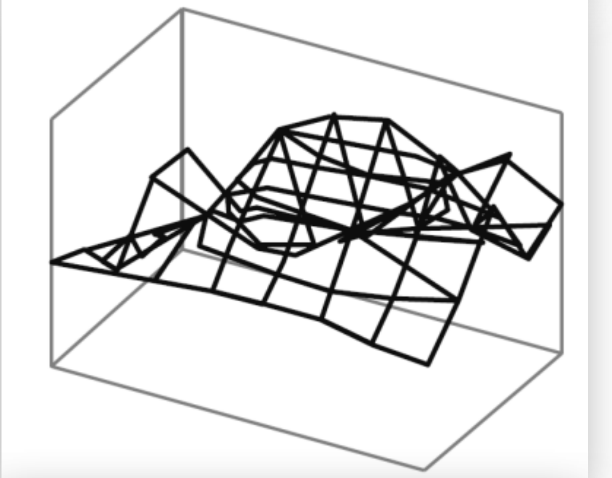
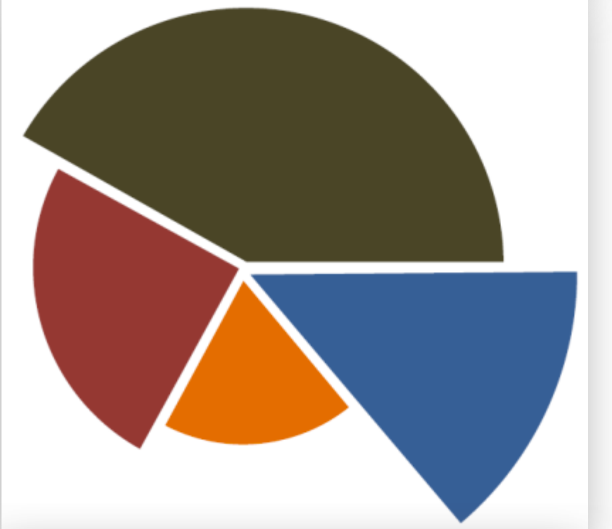
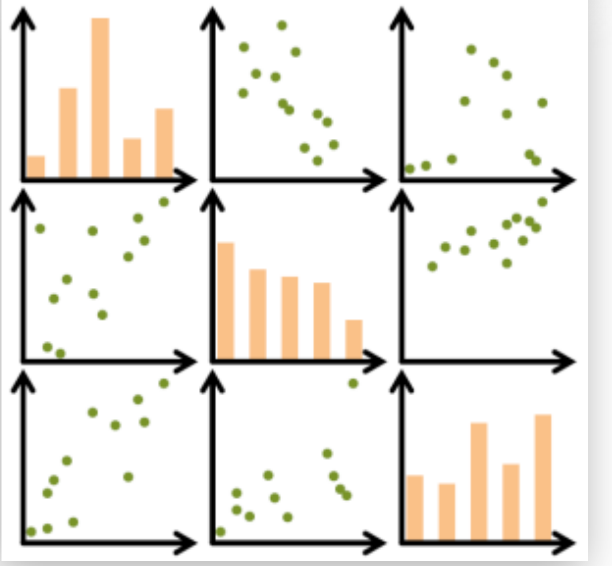
Rules of Thumb & Design

SPECIAL TOPICS ON VISUALIZATION IN NETWORK SCIENCE
OCTOBER 10, 2017 - NORTHEASTERN UNIVERSITY

Prof. Michelle Borkin





Today's Outline

- Who am I?
- In-class exercise
- Design Principles ("Rules of Thumb")
- Case Study: Artery Visualization for Heart Disease Diagnostics
- In-class exercise

WHO IS LECTURING TODAY?

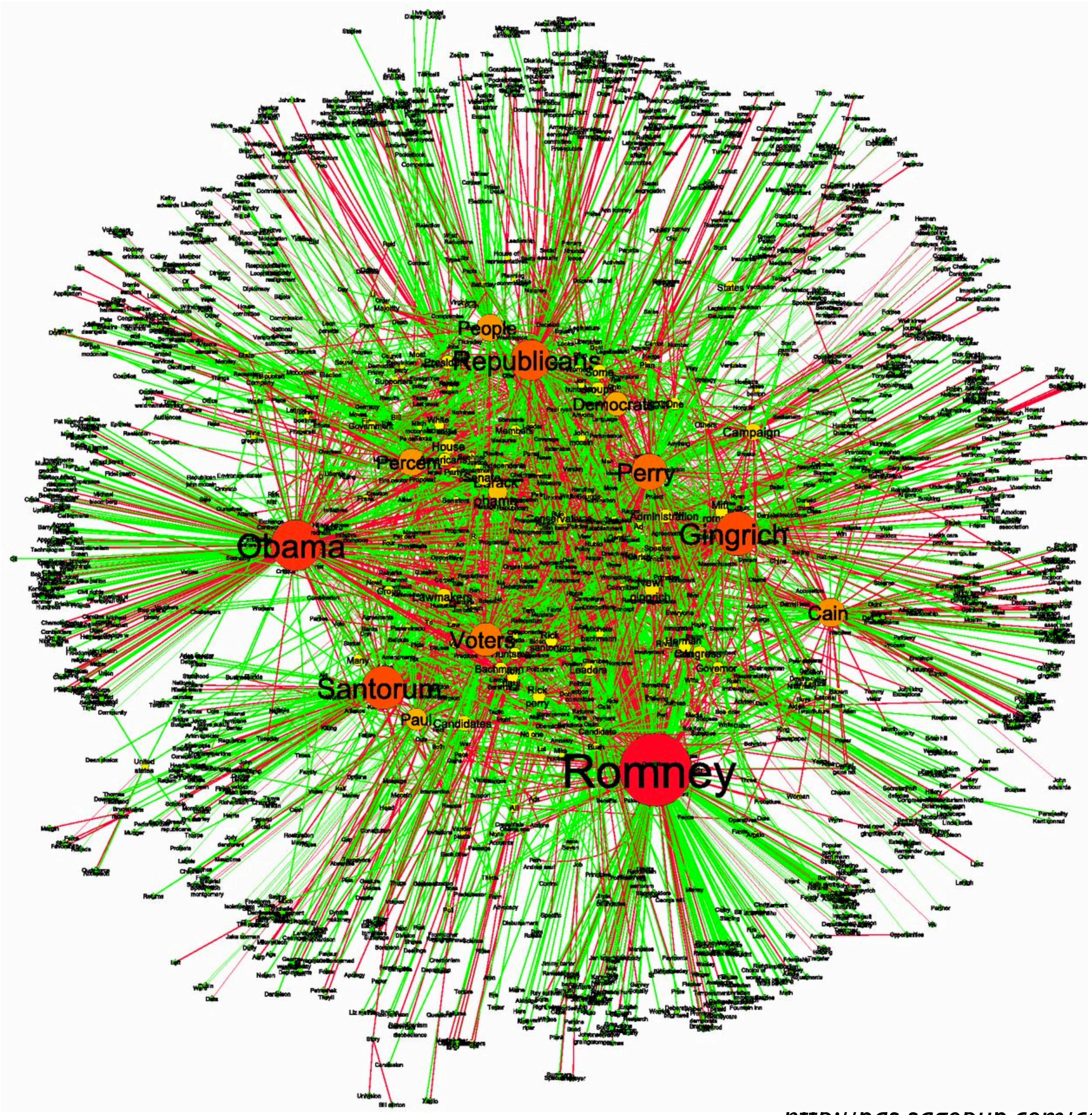
Prof. Michelle Borkin
West Village H, 310D
m.borkin@northeastern.edu

Research interests: visualization (information and scientific), human-computer interaction (HCI), medical imaging and radiology, astrophysics, and cognition and perception.

astronomy & physics ⟶ *applied physics* ⟶ *computer science* ⟶

In-class exercise

Hall of FAME or Hall of Shame



http://bds.sagepub.com/content/2/1/2053951715572916

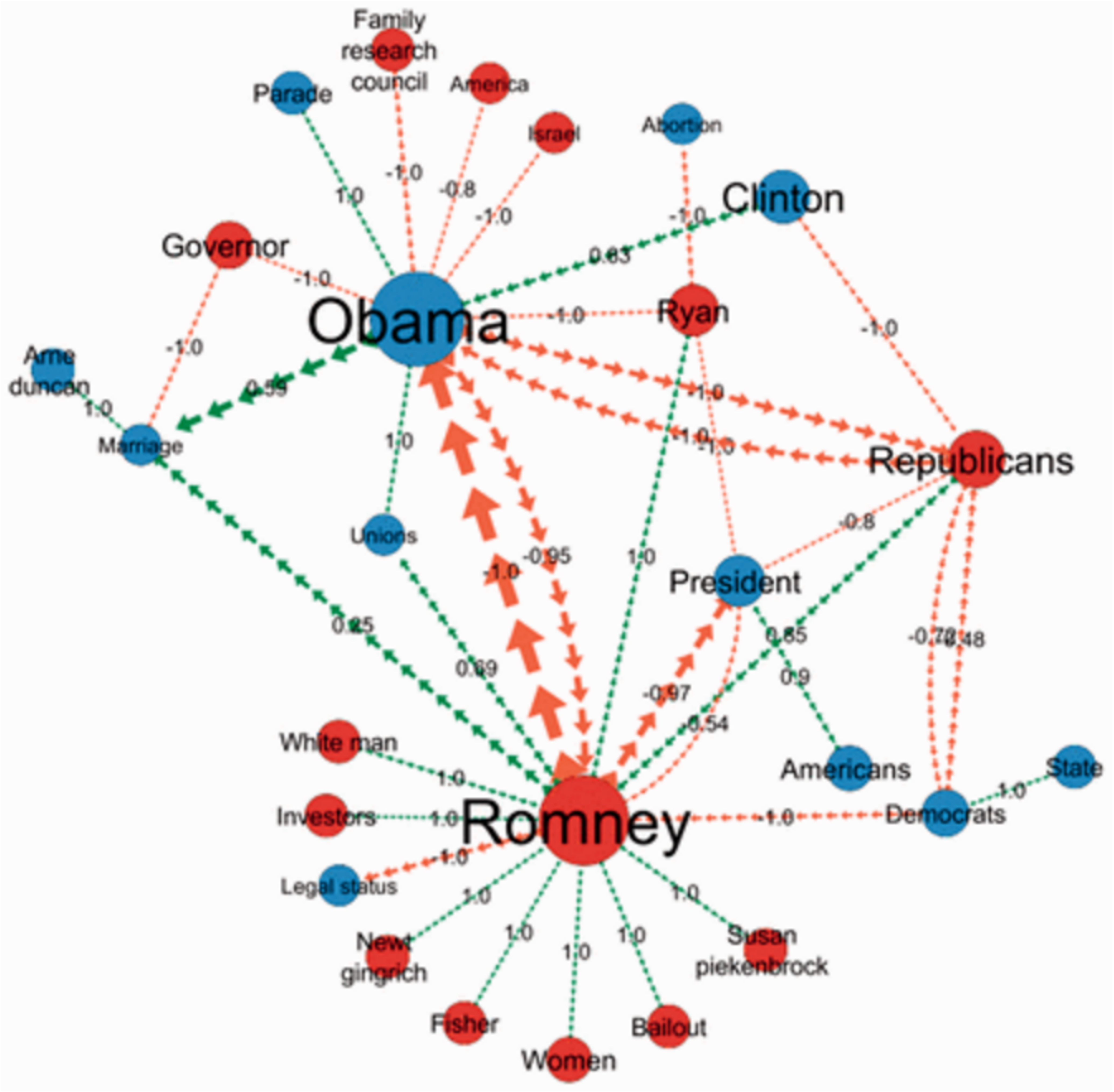
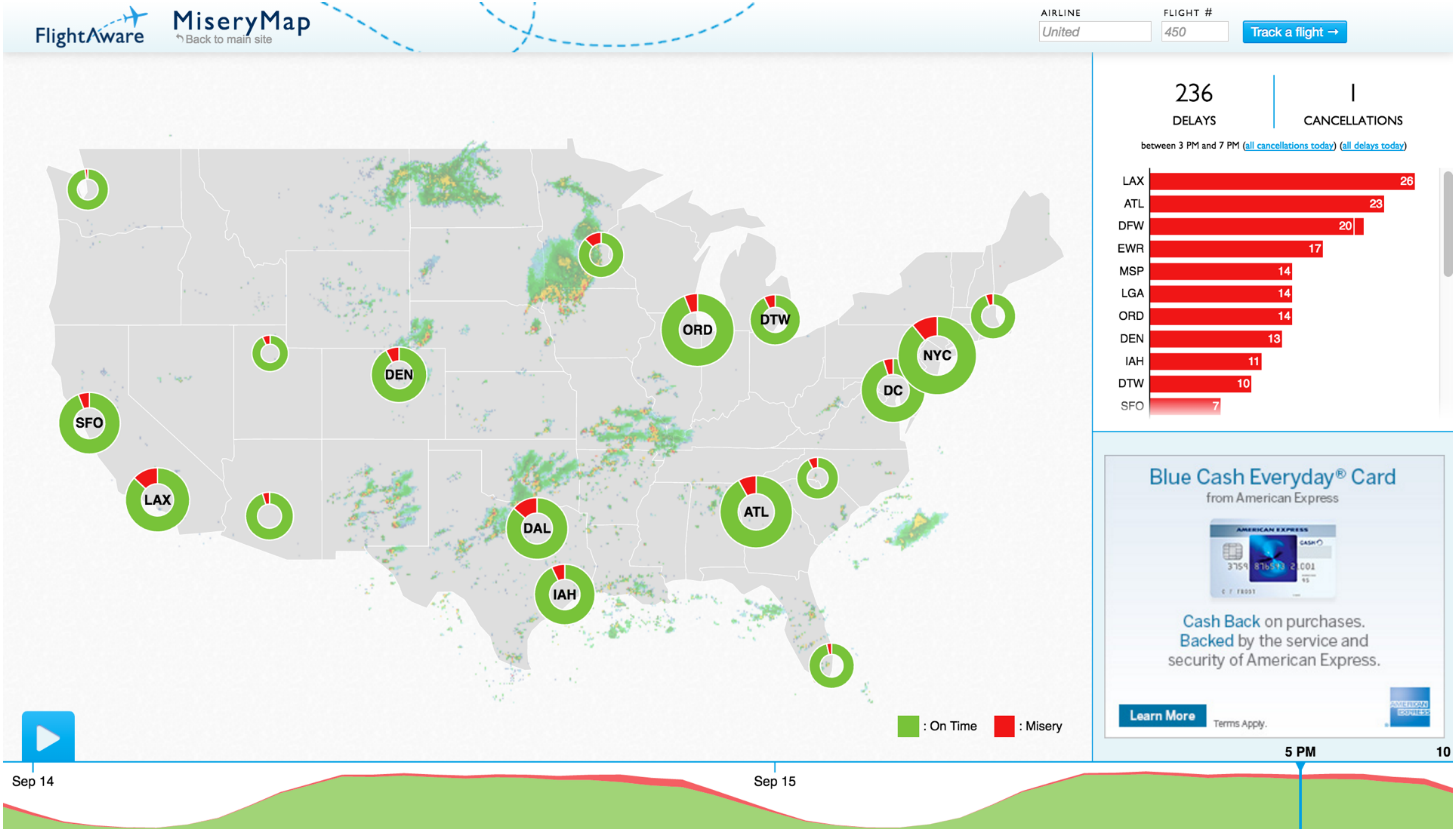


Figure 4. Download figure | Open in new tab | Download powerpoint

A subset of the election network, coloured by partitioning it via the first eigenvalue of the symmetrised adjacency matrix (see Appendix A8). Note that the split captures well the expected distinction between the Republican (red) and Democratic (blue) camps. The orange and green links show negative and positive relations between entities.

http://bds.sagepub.com/content/2/1/2053951715572916

http://hangtime.blogs.nba.com/2014/03/19/analytics-art-nba-passing/



http://flightaware.com/miserymap/all/

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.”

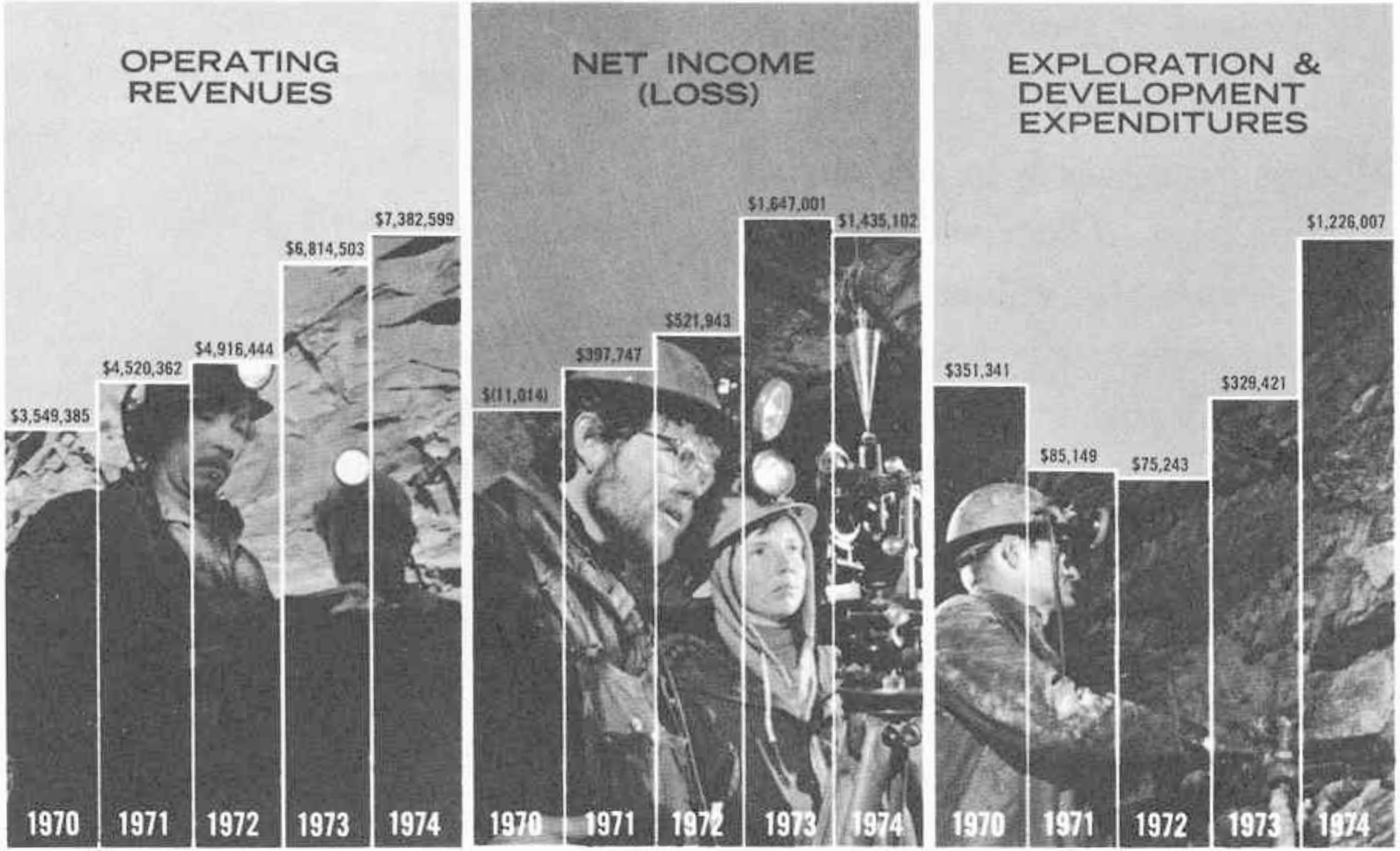
Tufte, “Visual Display of Quantitative Information” (1983)

“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.)

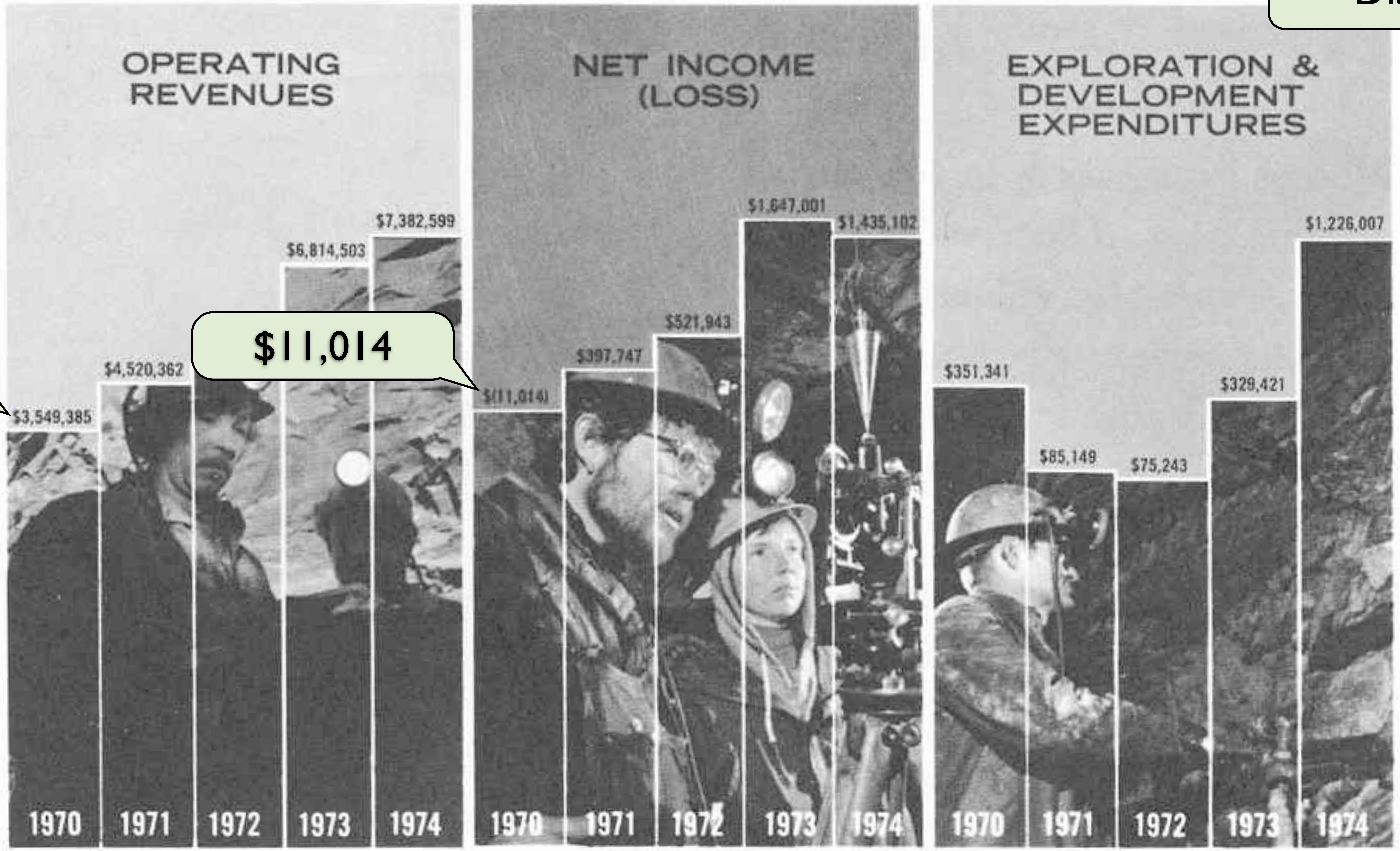
Tufte, “Visual Display of Quantitative Information” (1983)



"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)

“Distorted Scales”



$3,549,385

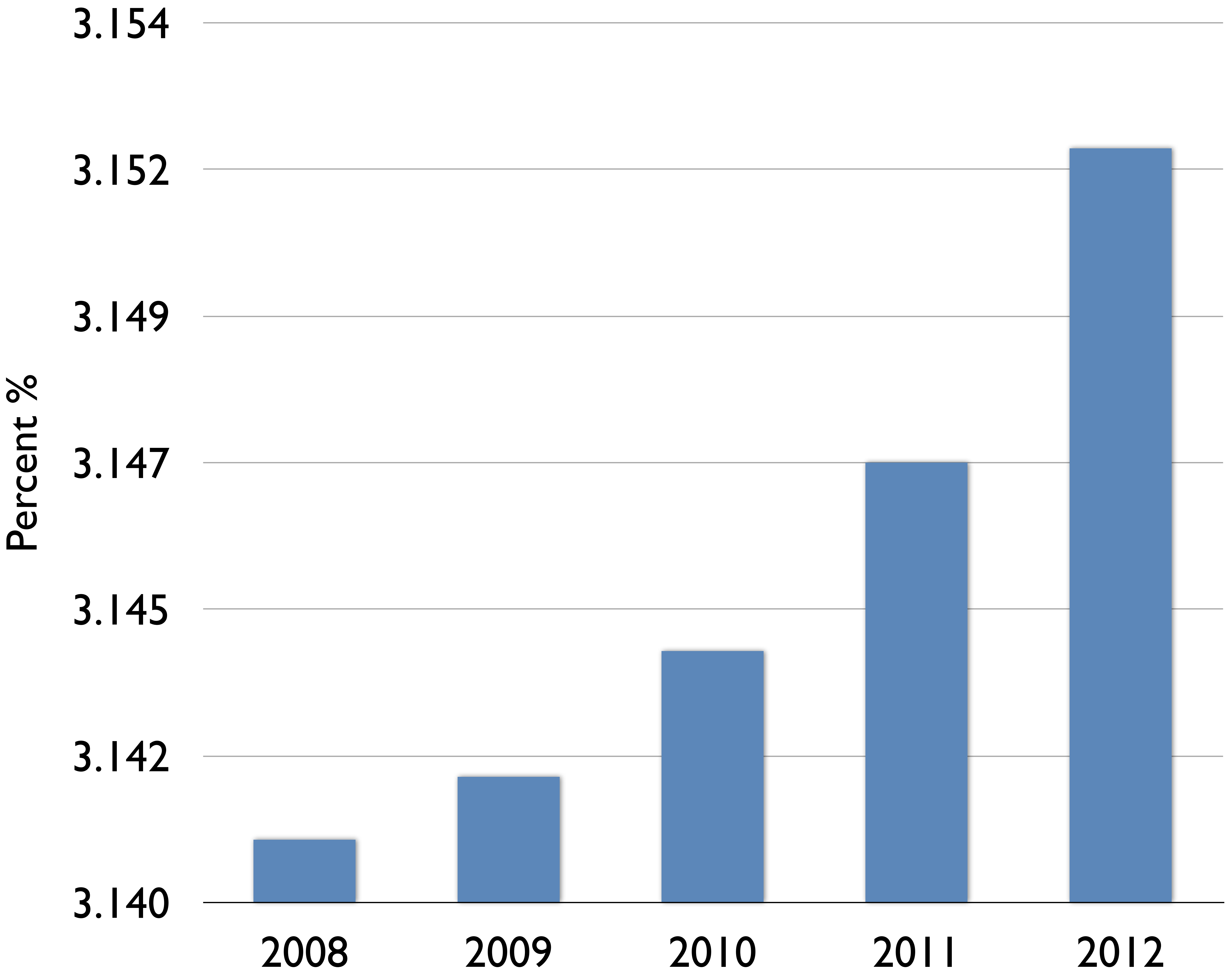
$11,014

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.”

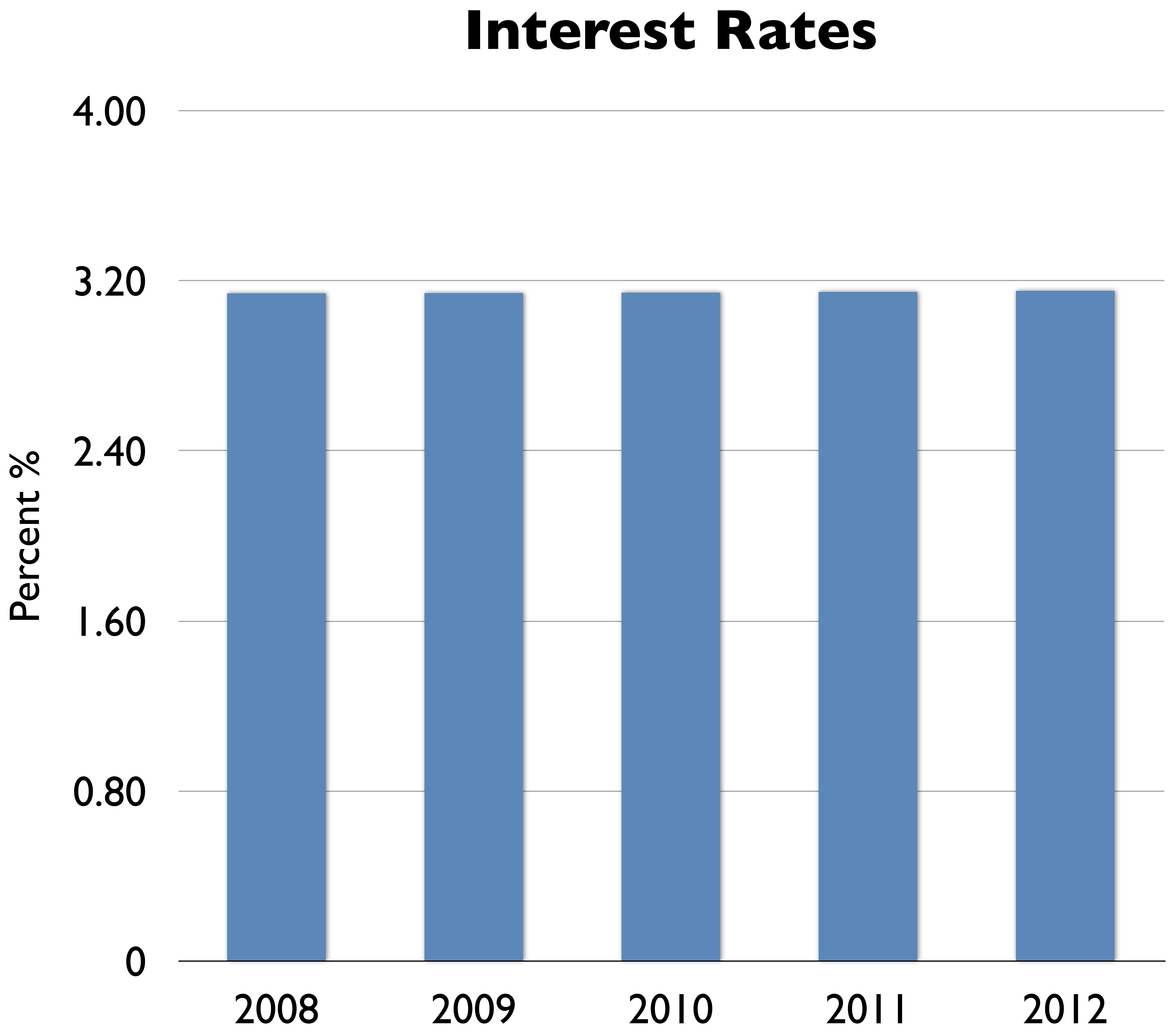
Tufte, “Visual Display of Quantitative Information” (1983)

Interest Rates



"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

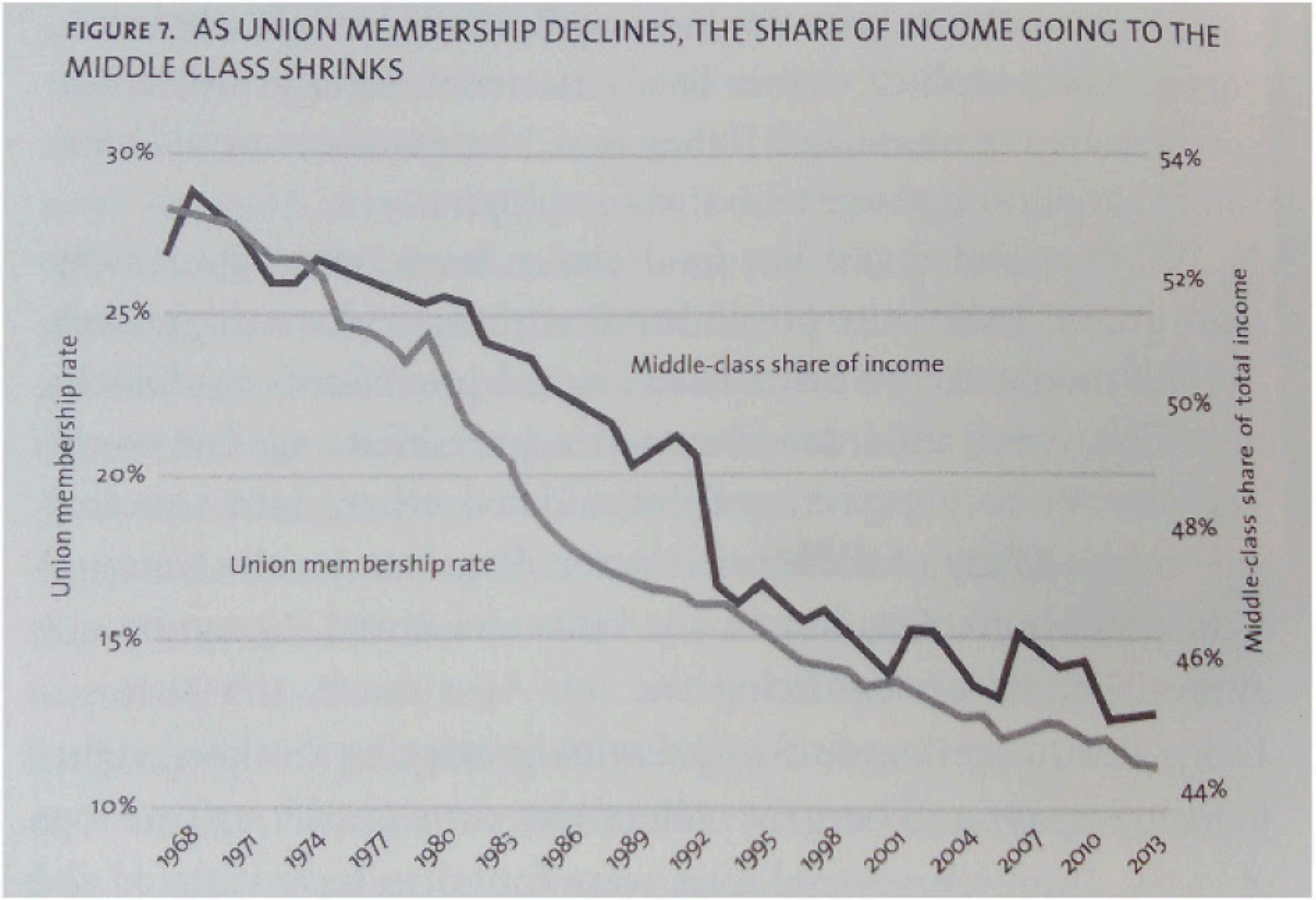


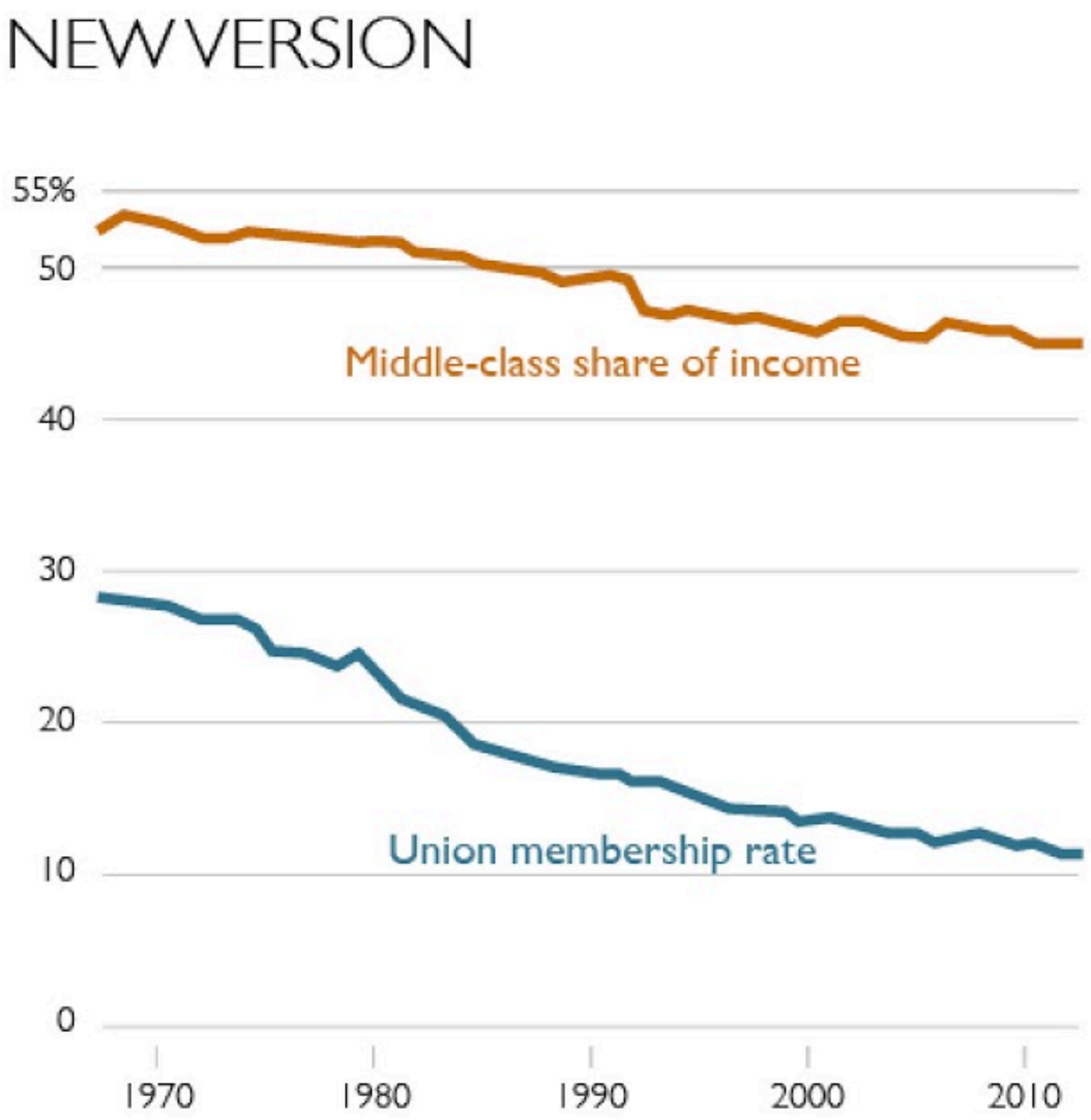
CONTEXT!

"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”



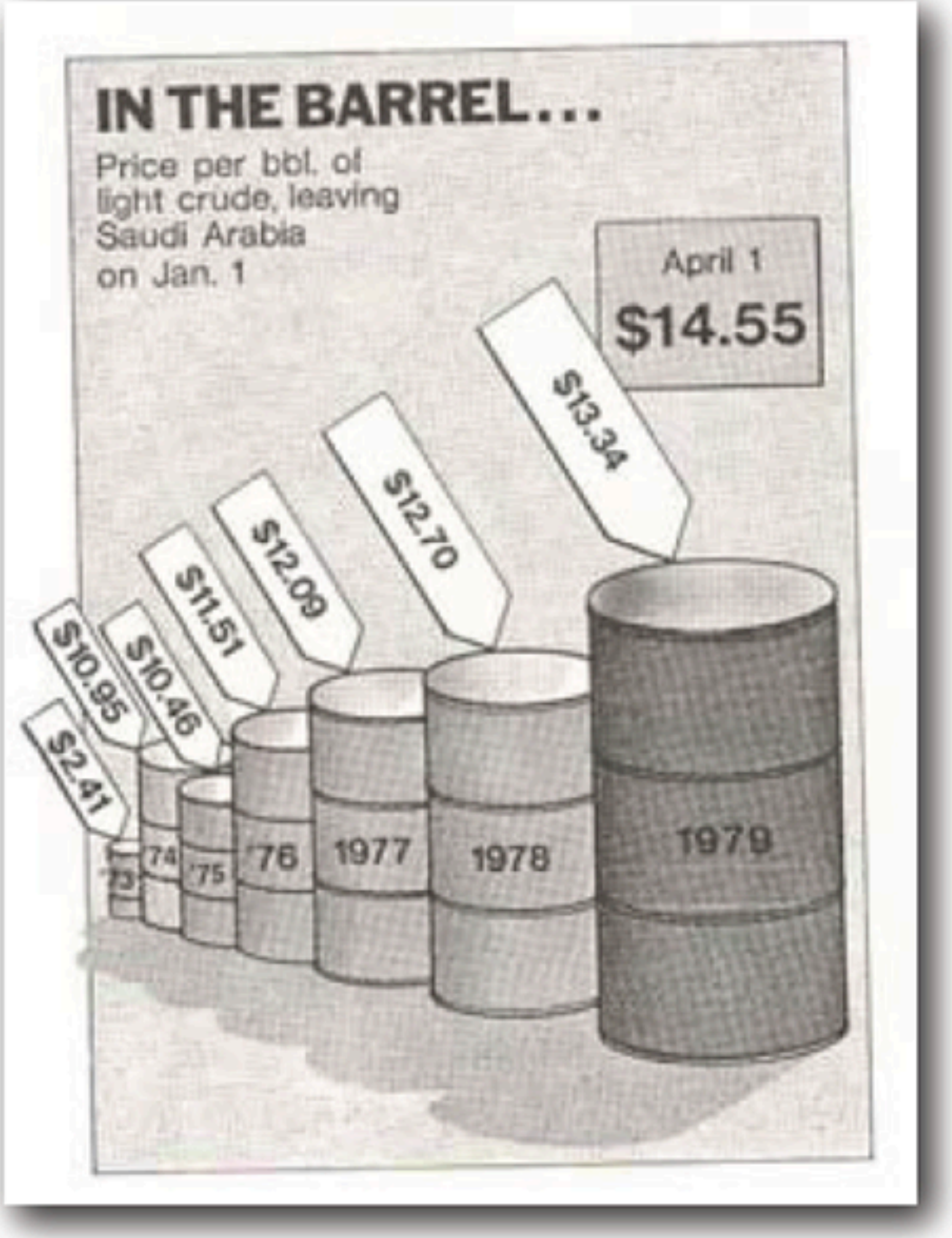


“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.” http://www.thefunctionalart.com/2015/10/double-axes-double-mischief.html

“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)

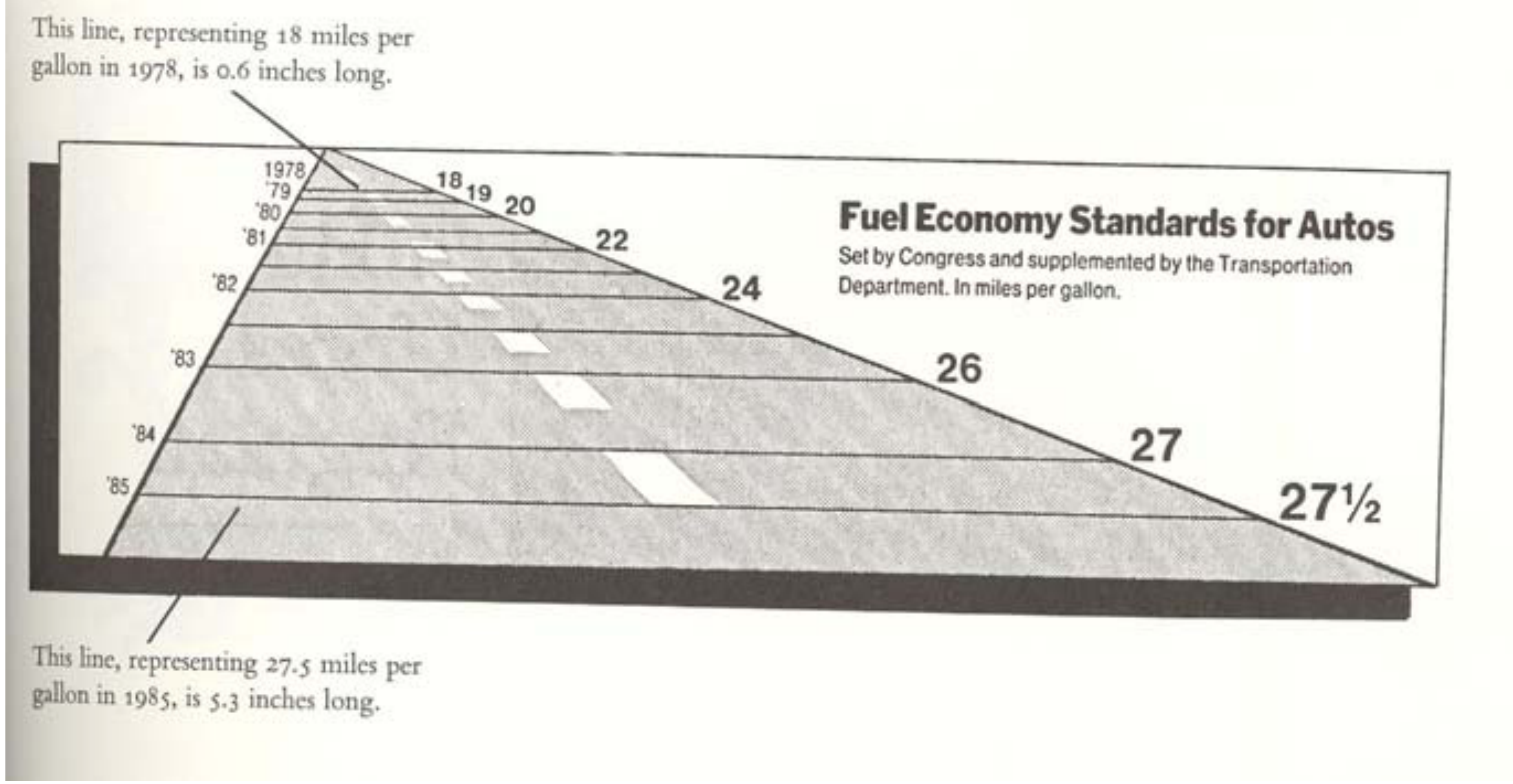


"The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured."

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

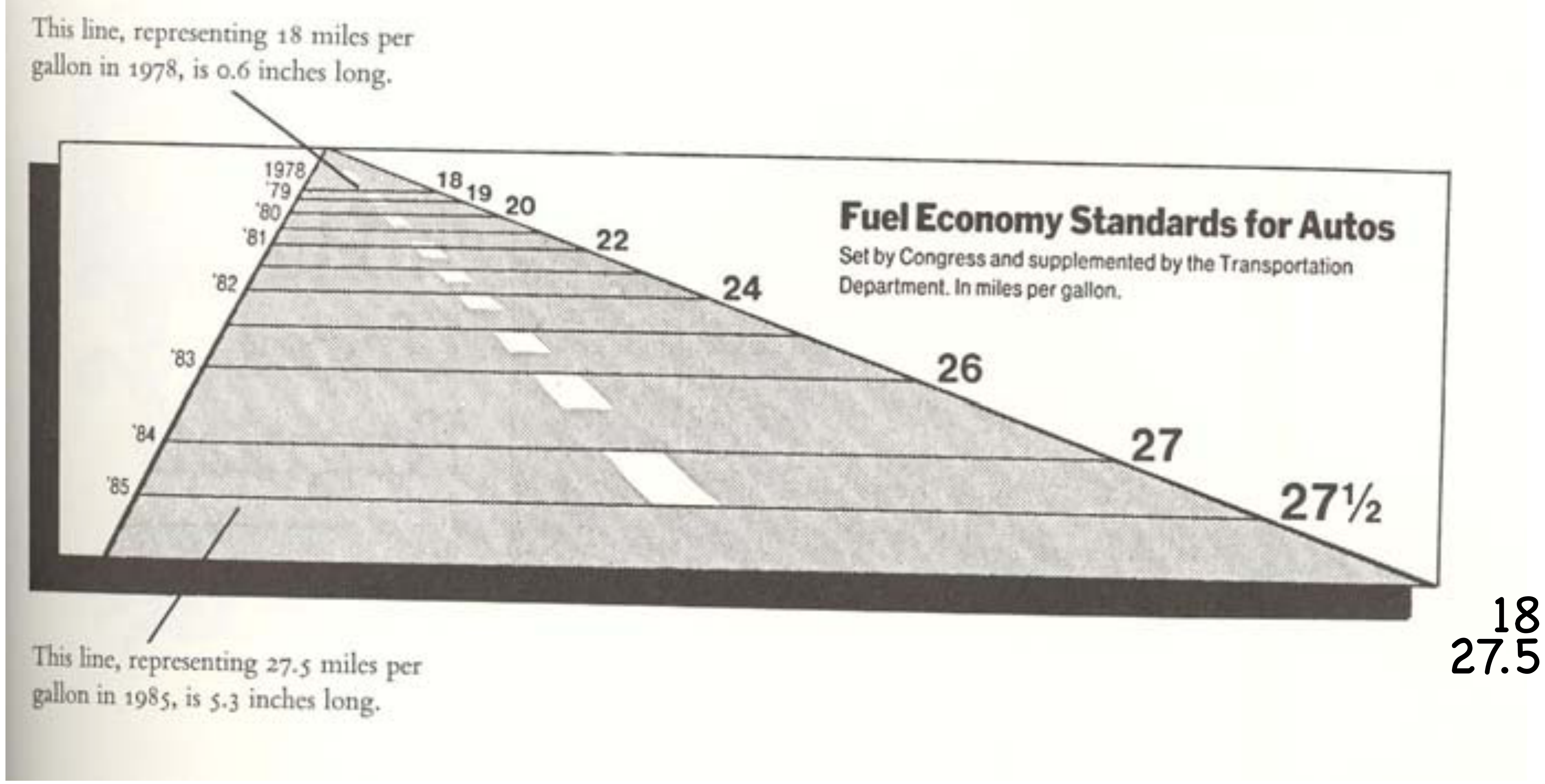


Image = $\frac{5.3" - 0.6"}{0.6"}$ = 7.83 = 783%

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."

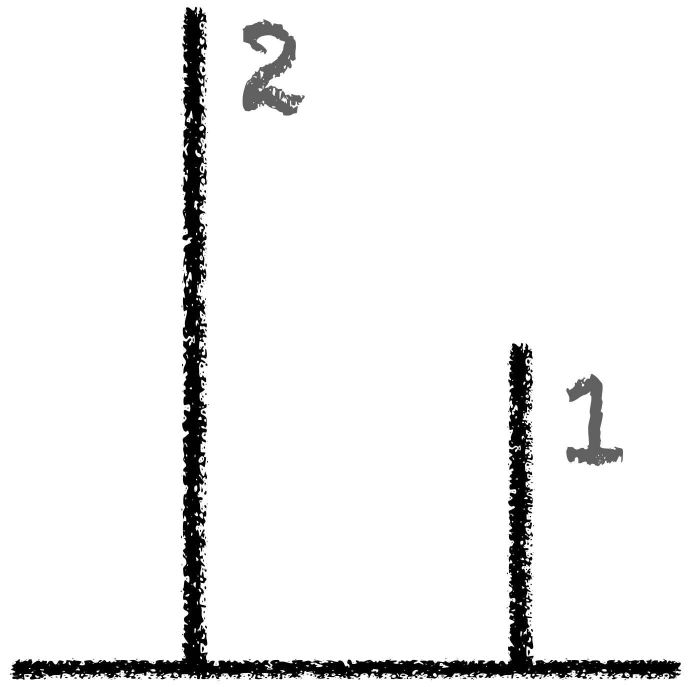
Tufte, "Visual Display of Quantitative Information" (1983)

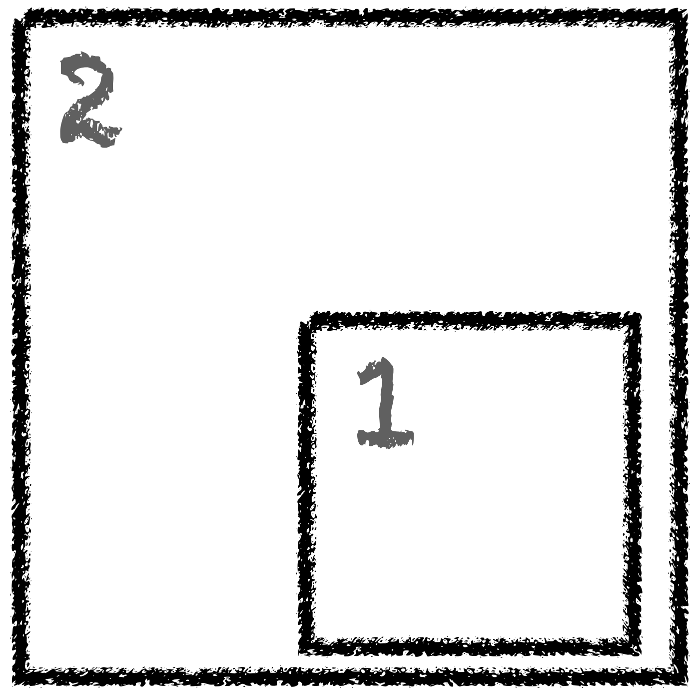
IN-CLASS ACTIVITY:
Calculate for yourself!

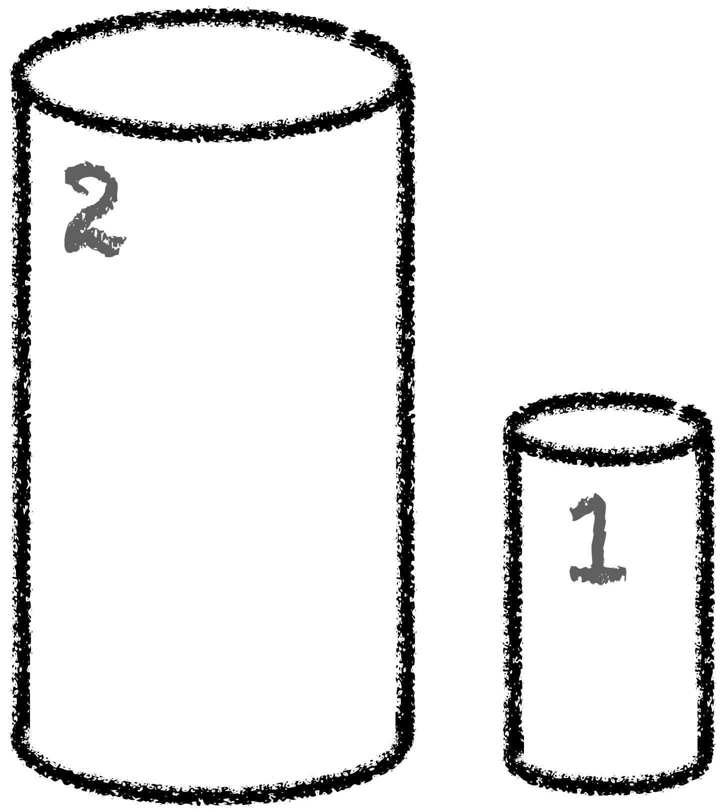
Lie Factor

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

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







"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!

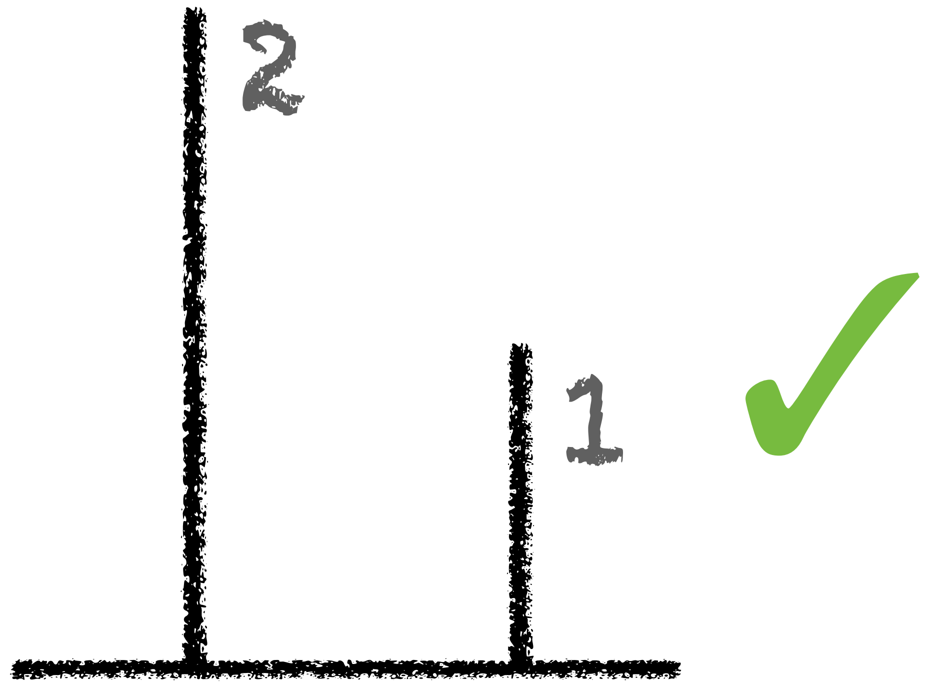
Lie Factor

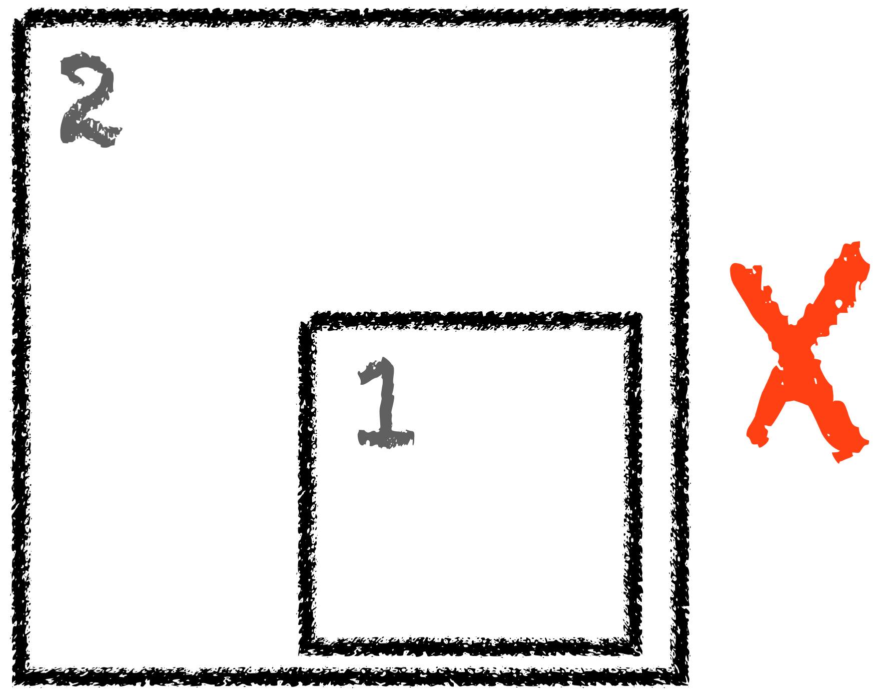
$$\text{Data} = \frac{2 - 1}{1} = 1 = 100\%$$

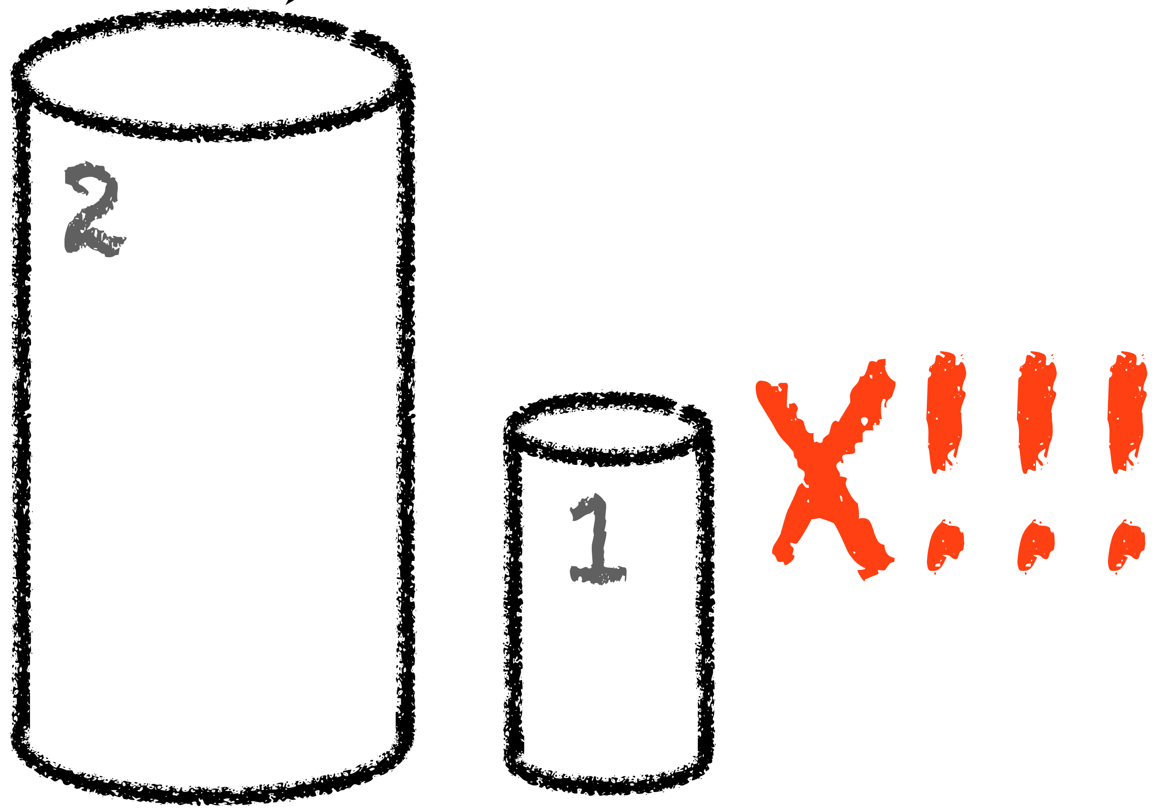
$$\text{Lie Factor} = \frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$$

Make sure *area* is proportional to data!

Don't use 3D bar charts!







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

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

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

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

$$\text{Image} = \frac{2{*}\pi 1^2 - 1{*}\pi 0.5^2}{1{*}\pi 0.5^2} = 3 = 700\%$$

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

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

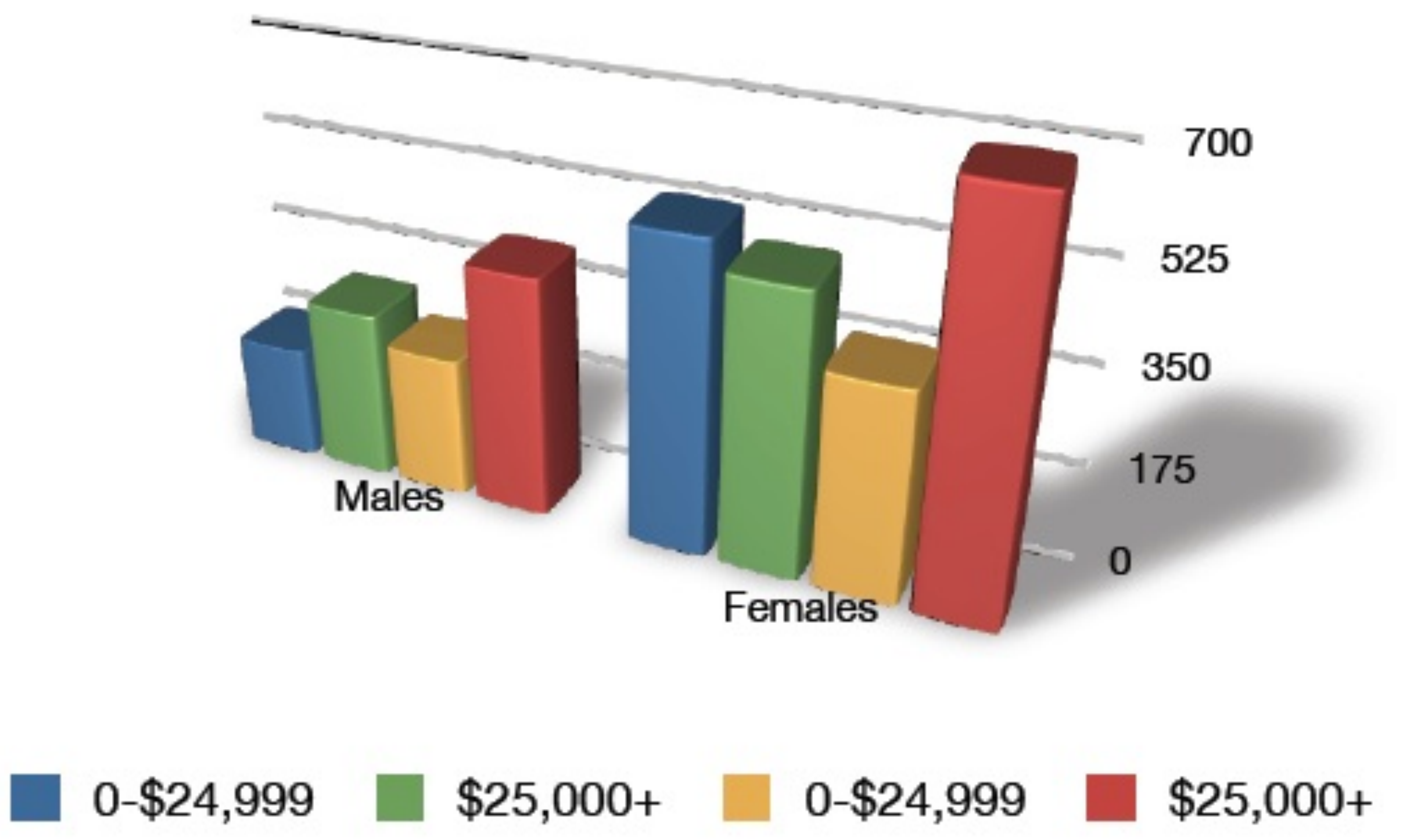
Tufte, "Visual Display of Quantitative Information" (1983)

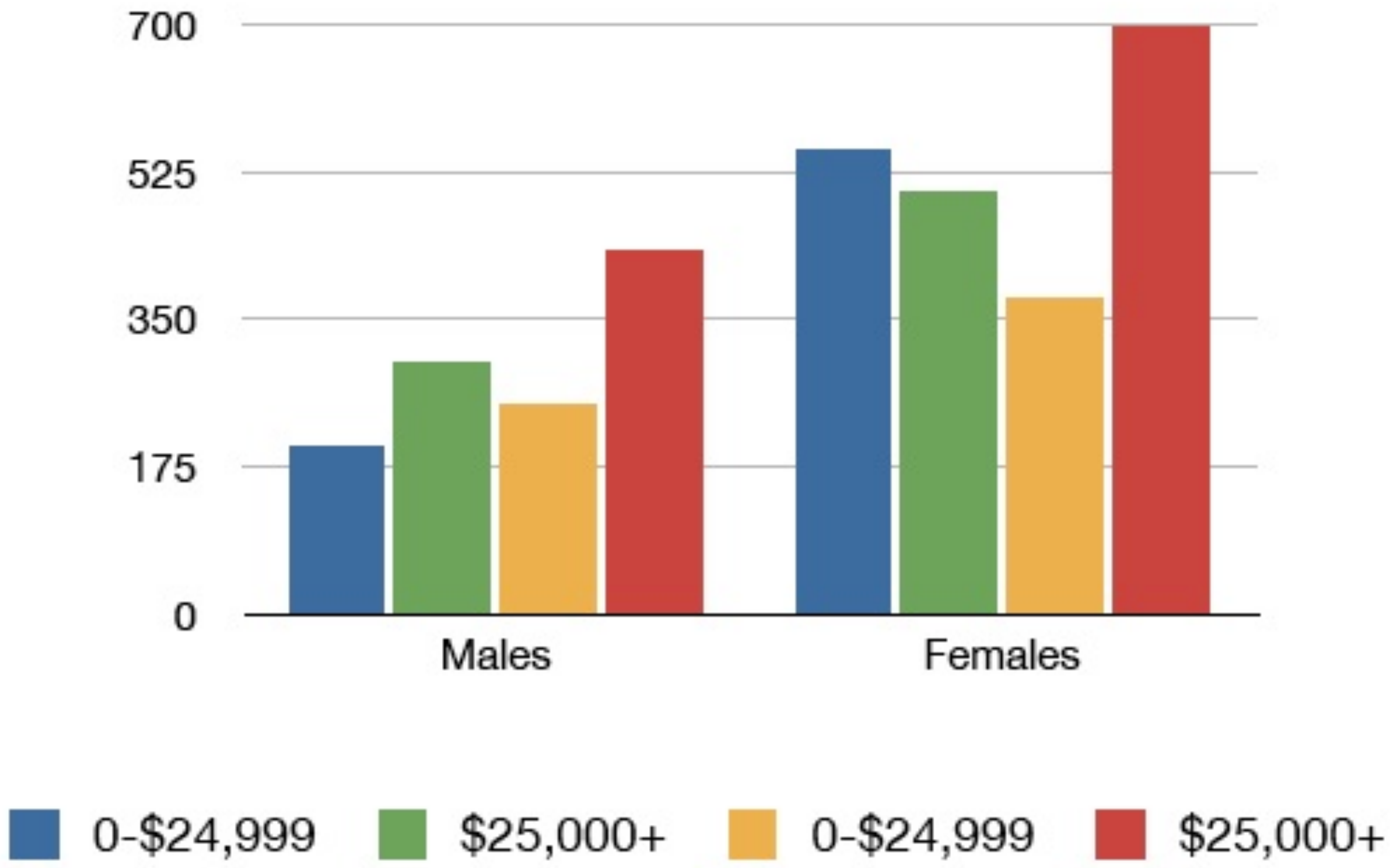
“Graphical Integrity”

Data Ink = the ink used to show data

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

Tufte: maximize the data ink ratio





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

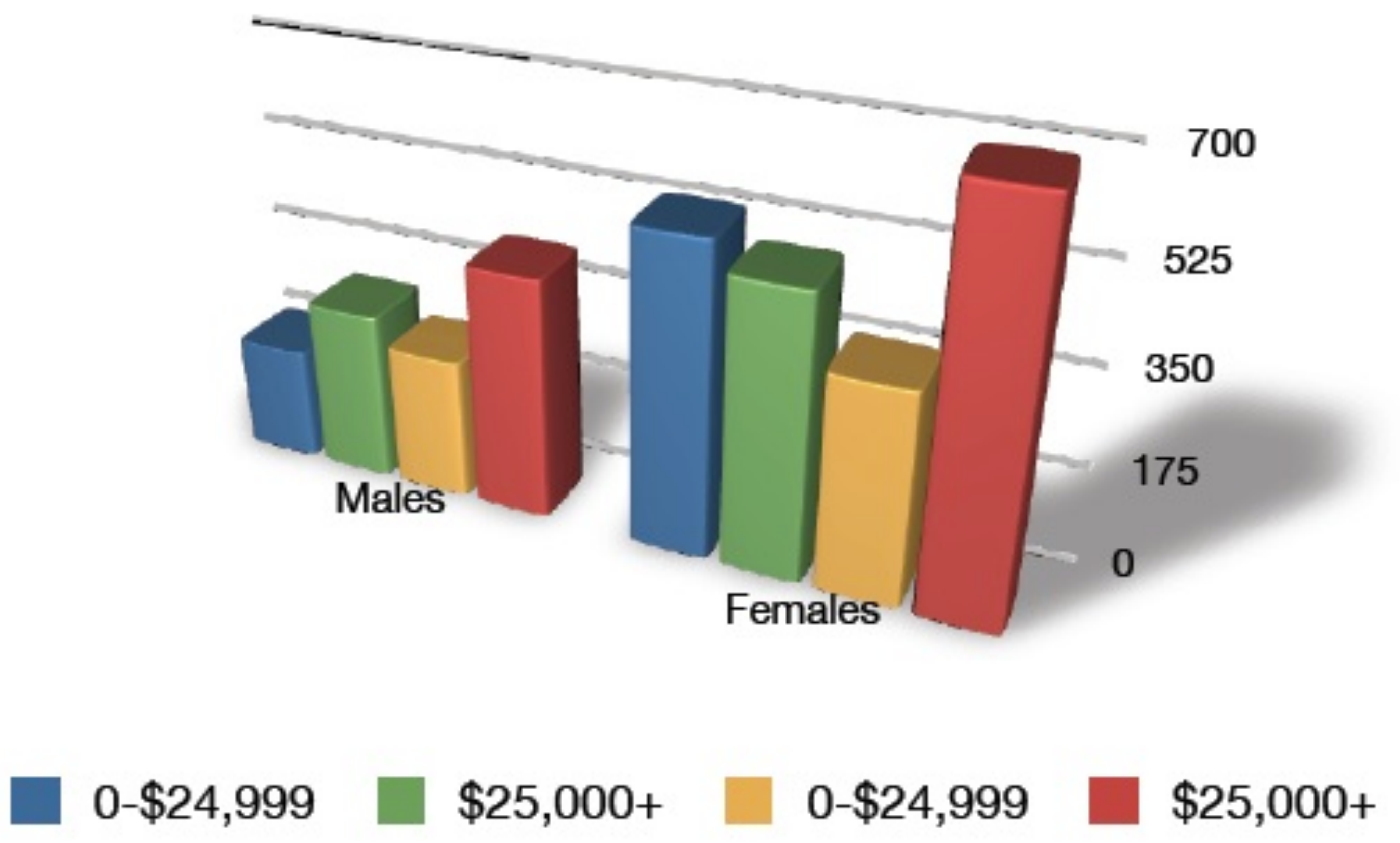
"Graphical Integrity"

Data Ink = the ink used to show data

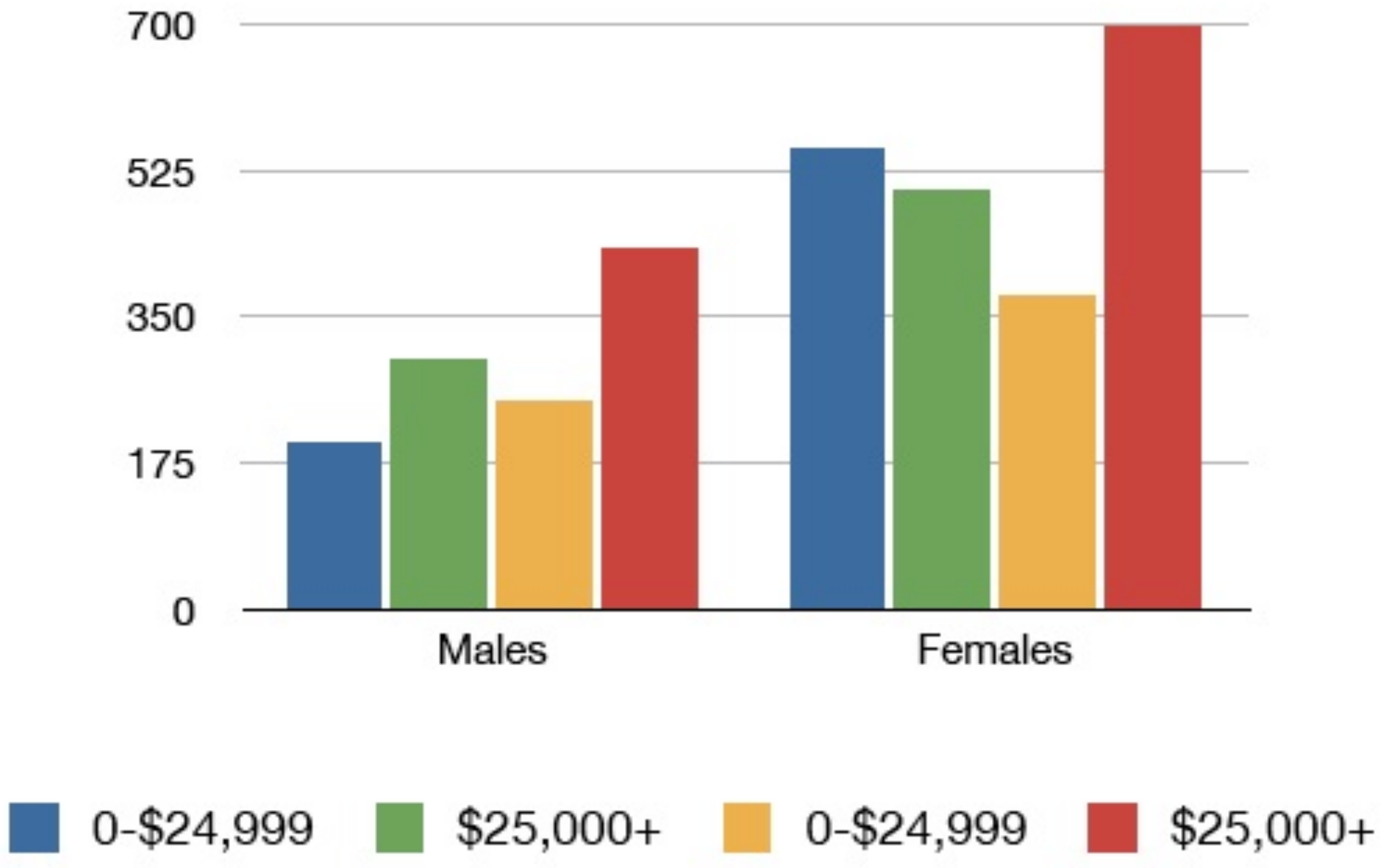
Data Ink Ratio = $\frac{\text{data-ink}}{\text{total ink in graphic}}$

Tufte: maximize the data ink ratio

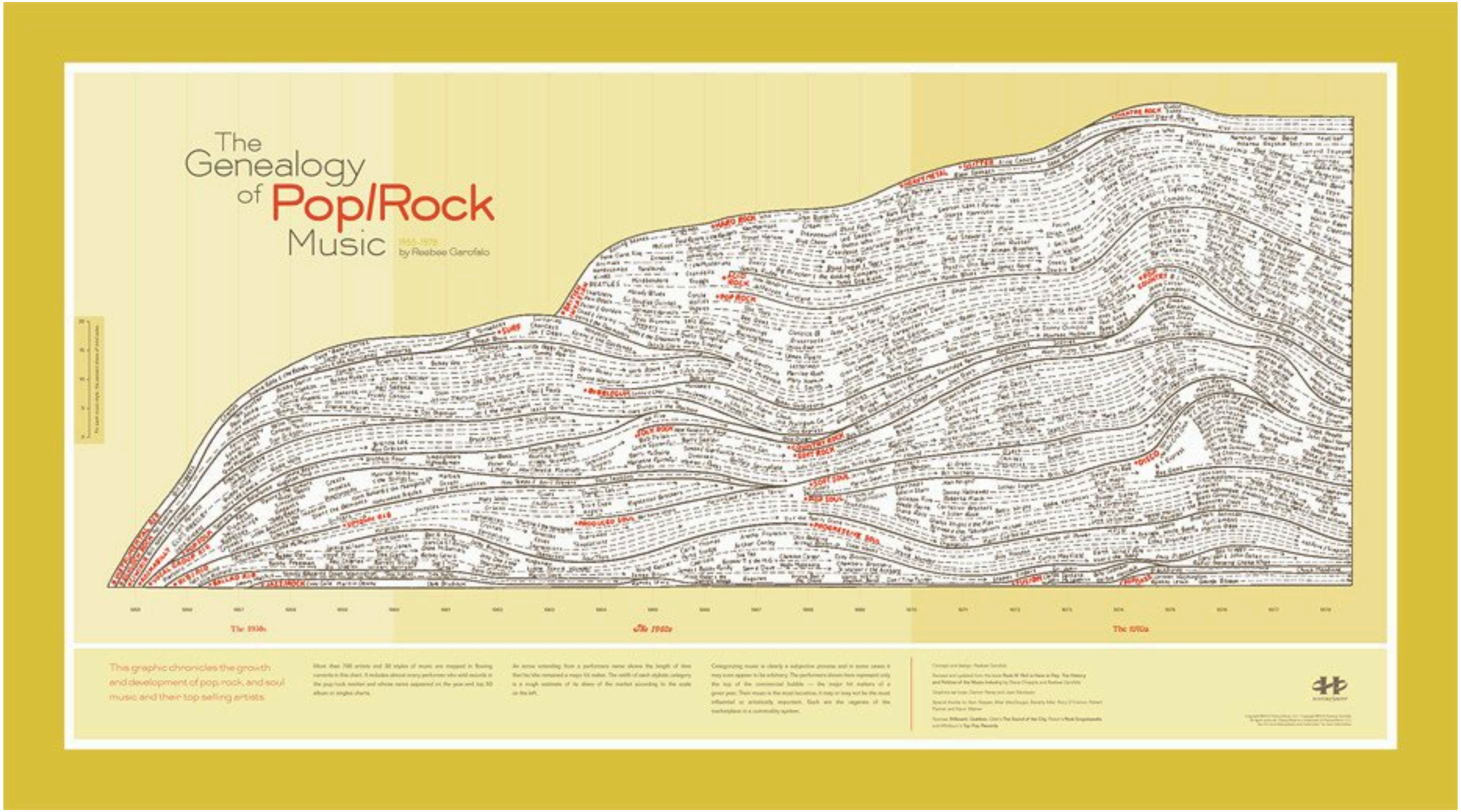
Low Data Ink Ratio



High Data Ink Ratio



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



High Data Ink Ratio

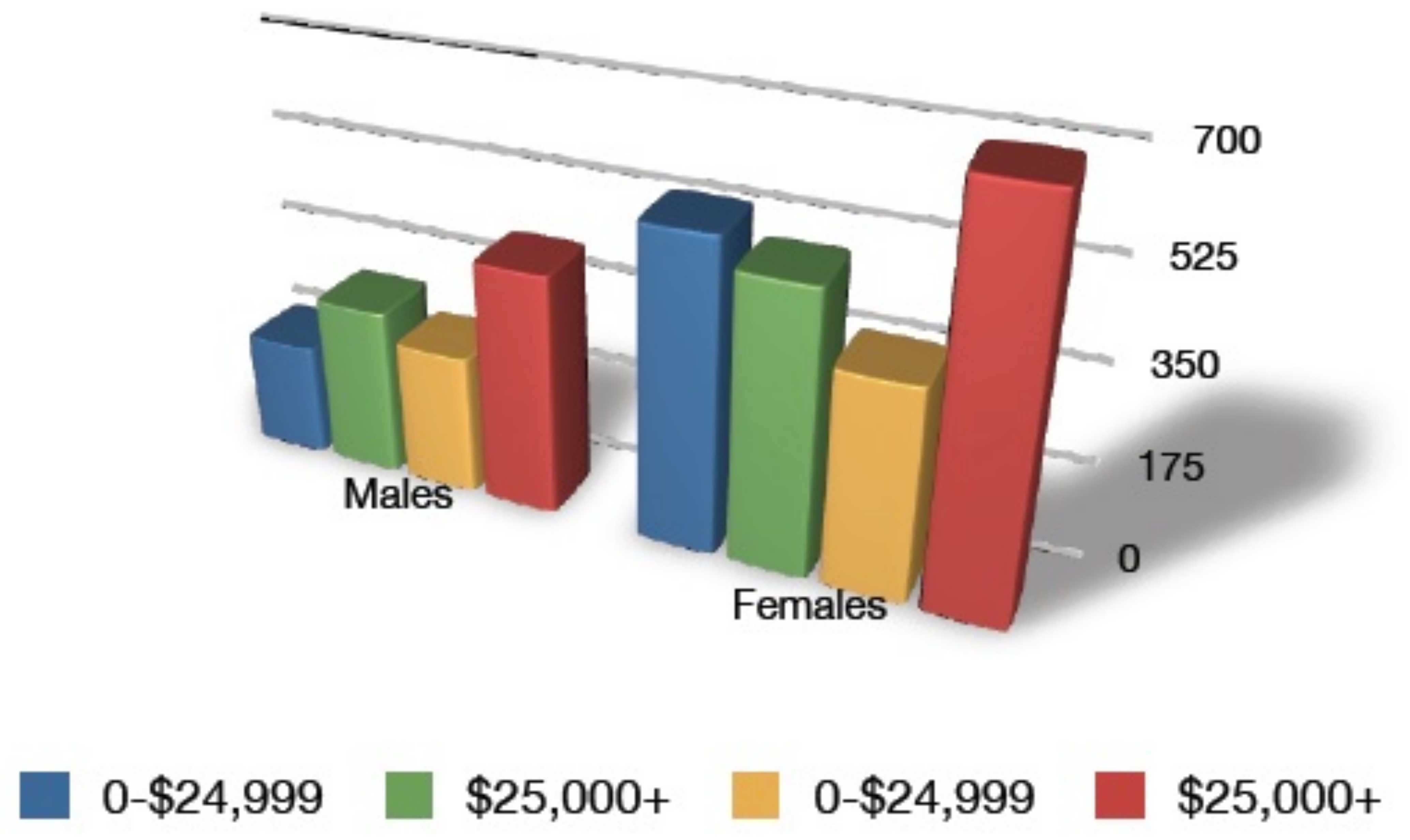
[Reebee Garofalo, Genealogy of Pop/Rock Music]

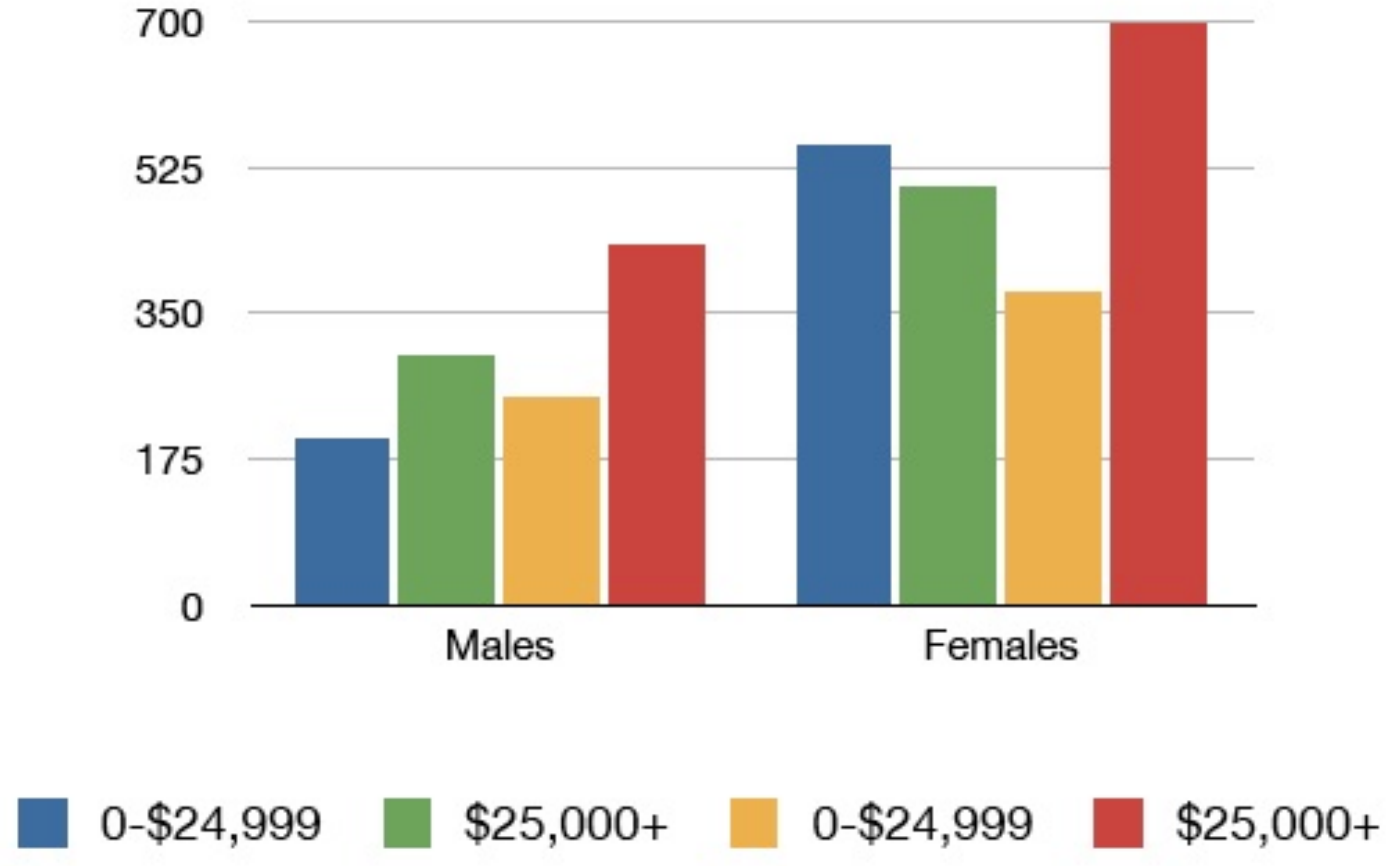
“Graphical Integrity”

“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”

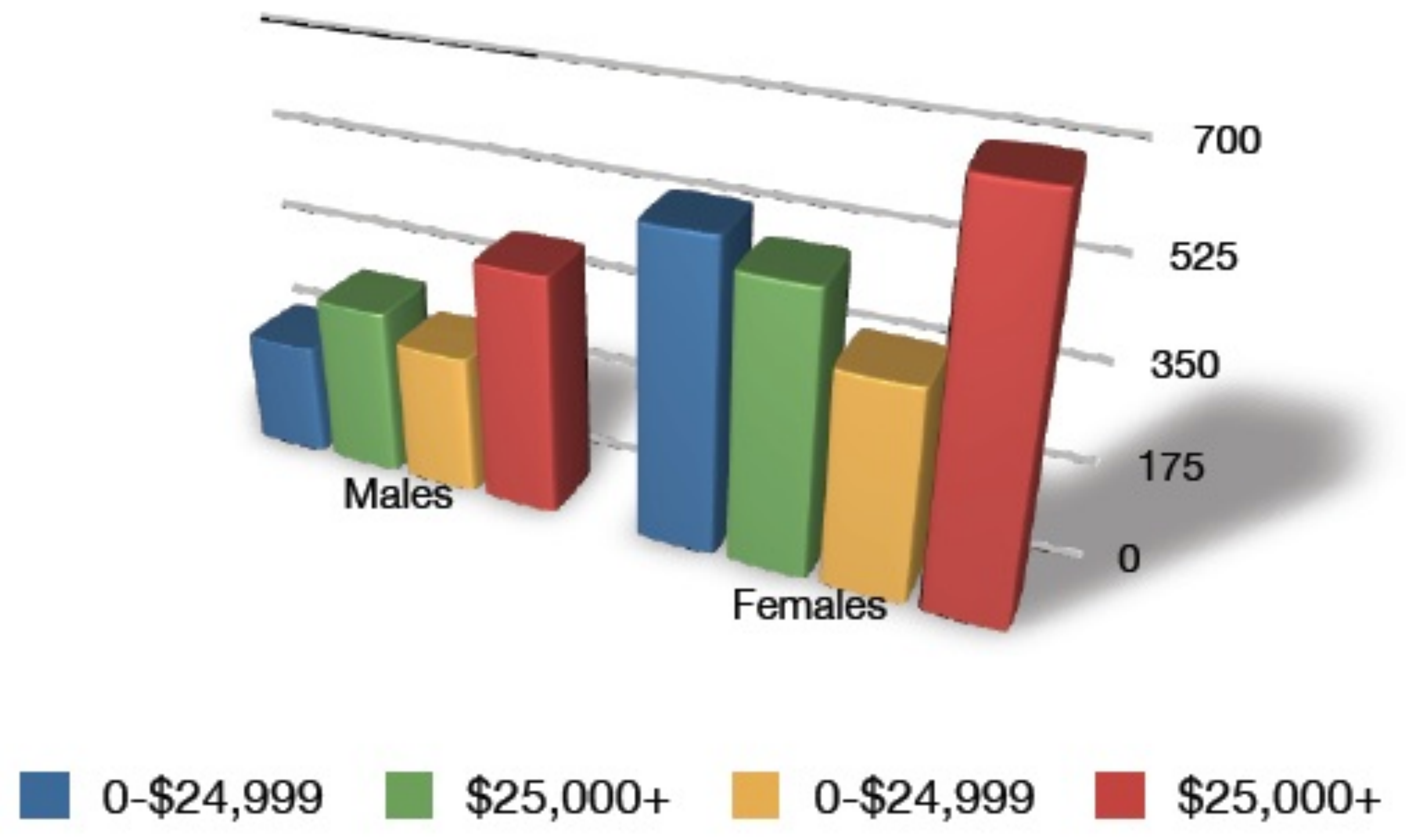




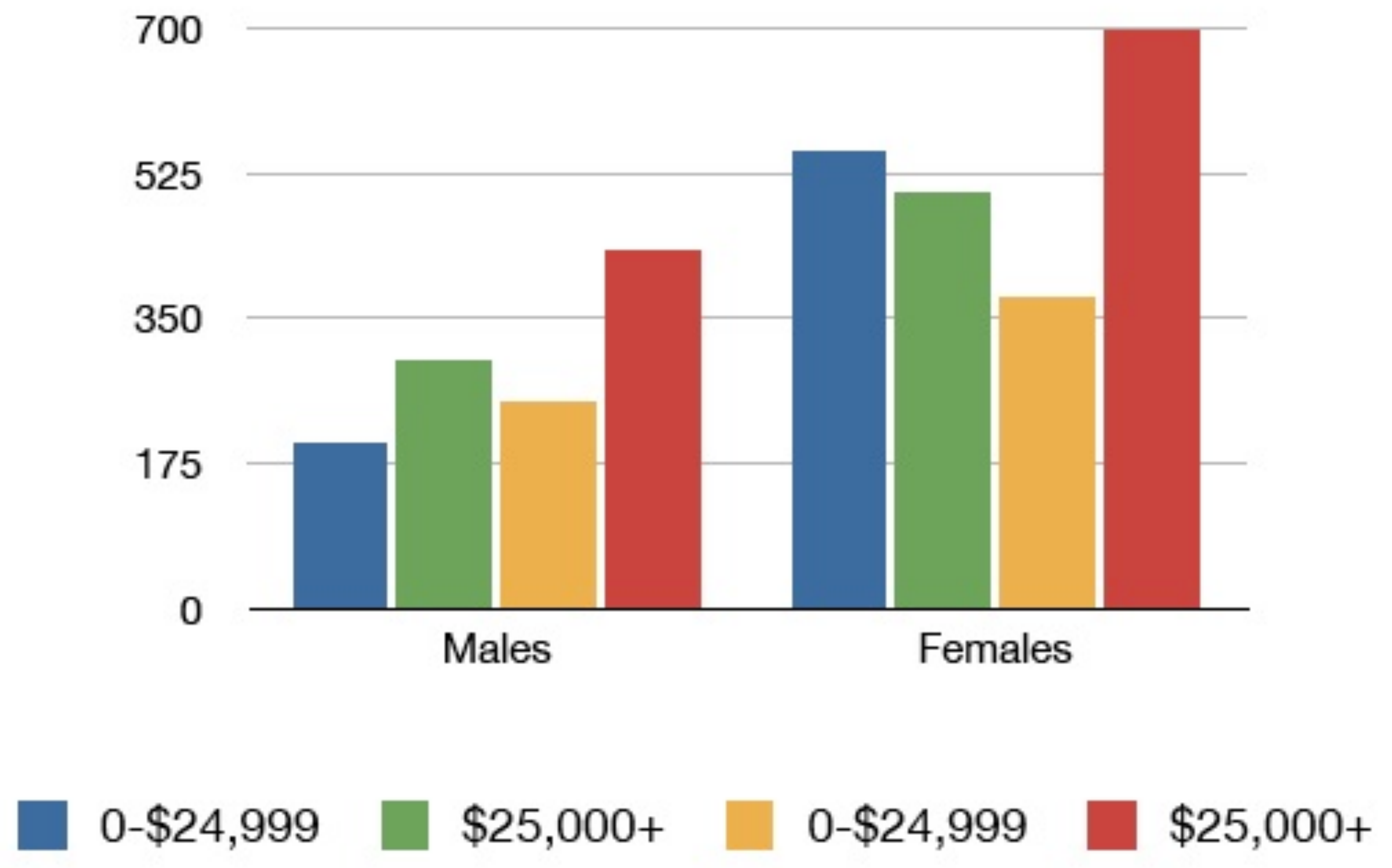
“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

"No Unjustified 3D"

Dimensions in data: 2
Dimensions in plot: 3



Dimensions in data: 2
Dimensions in plot: 2

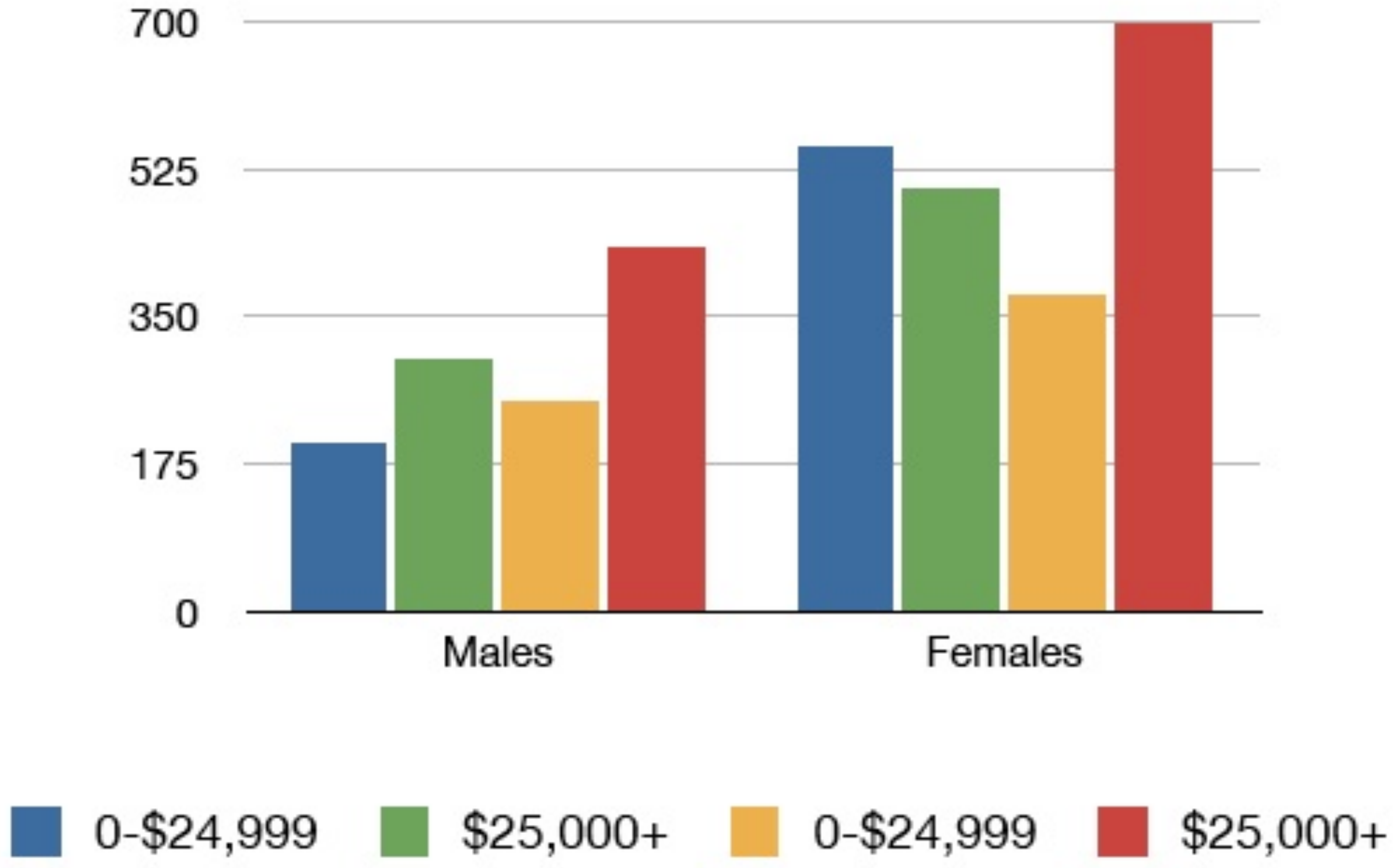


"The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data."

“No Unjustified 3D”

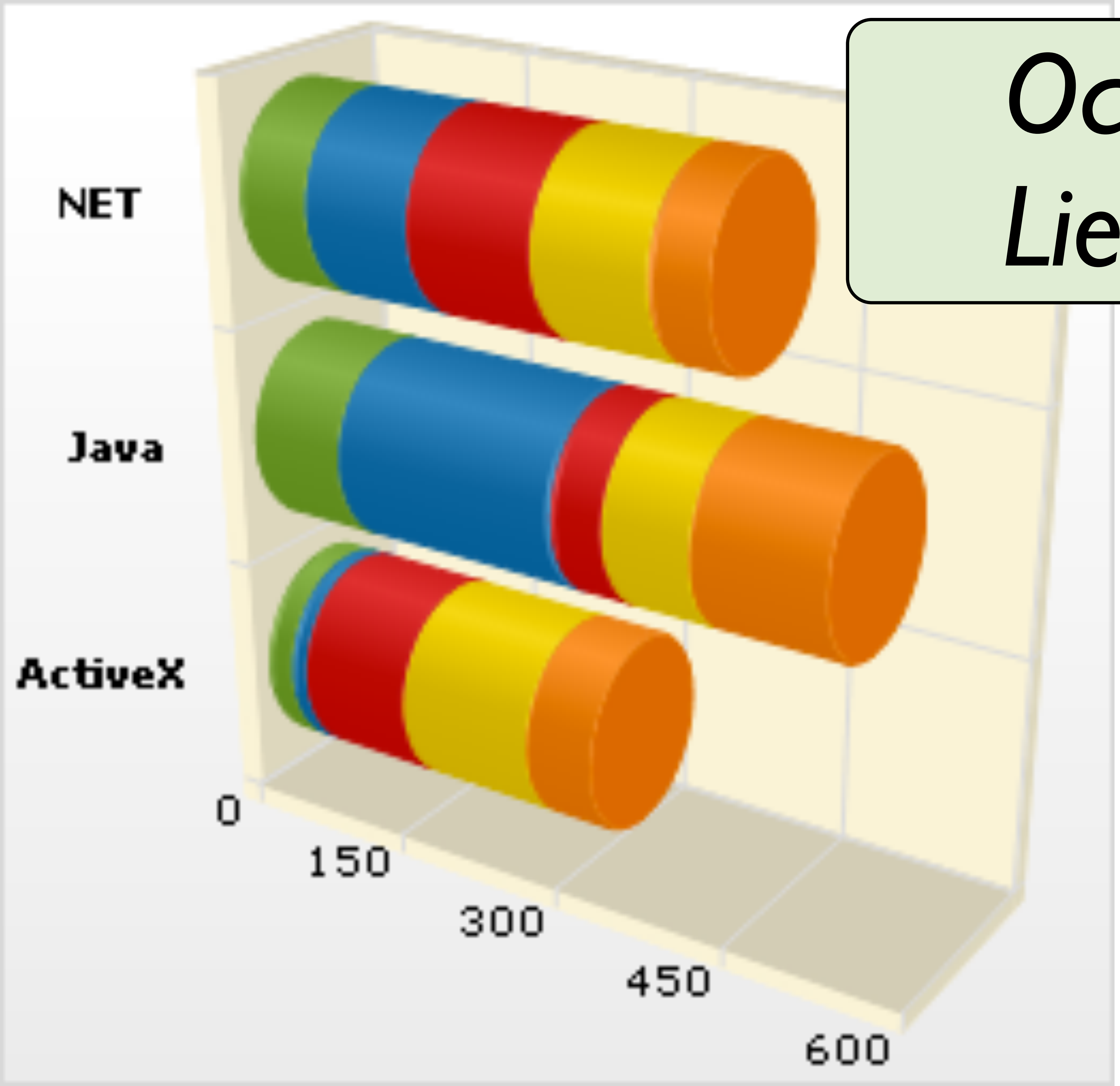
Dimensions in data: 2
Dimensions in plot: 3

Dimensions in data: 2
Dimensions in plot: 2

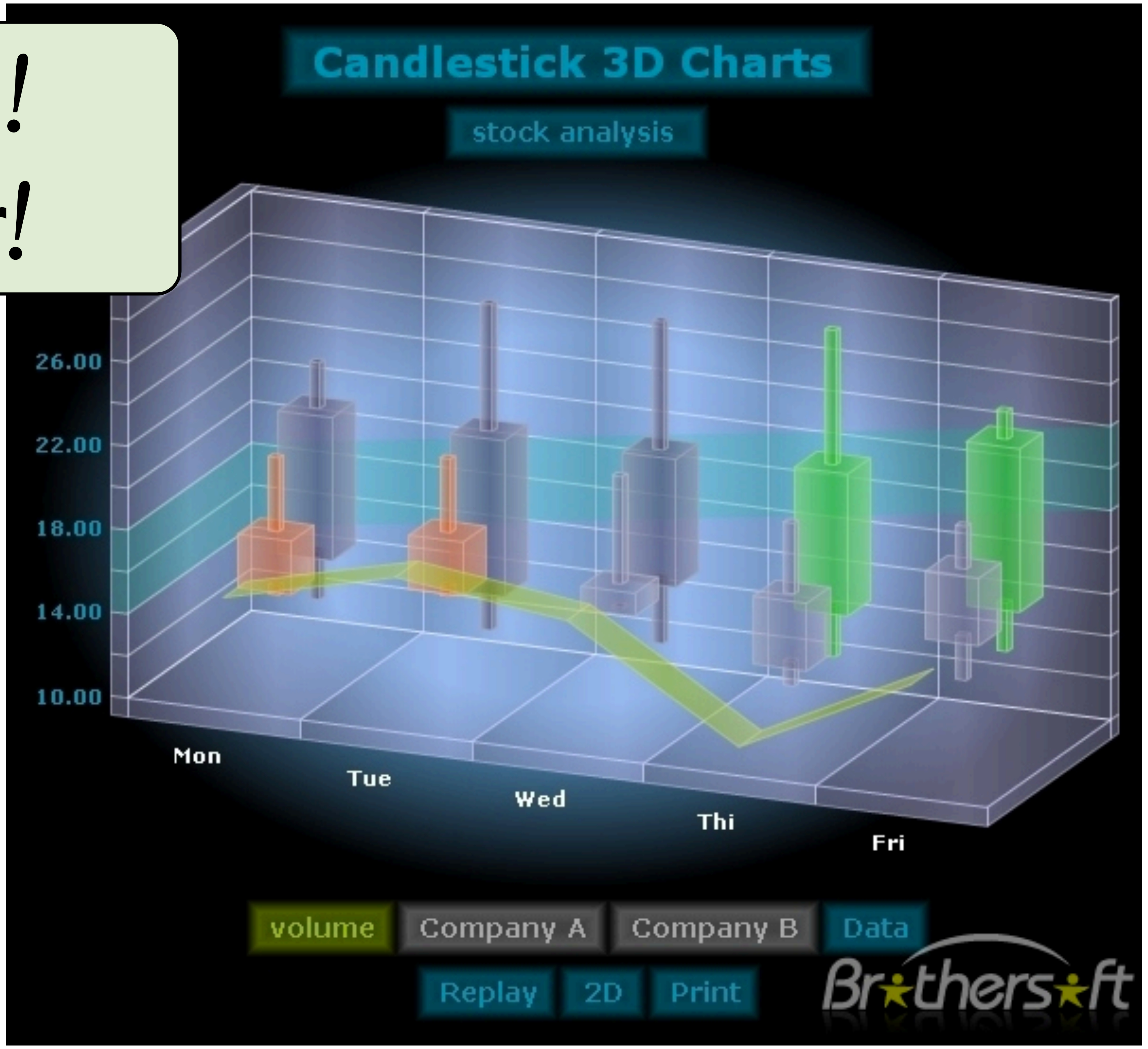


“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

"No Unjustified 3D"



Occlusion!
Lie Factor!

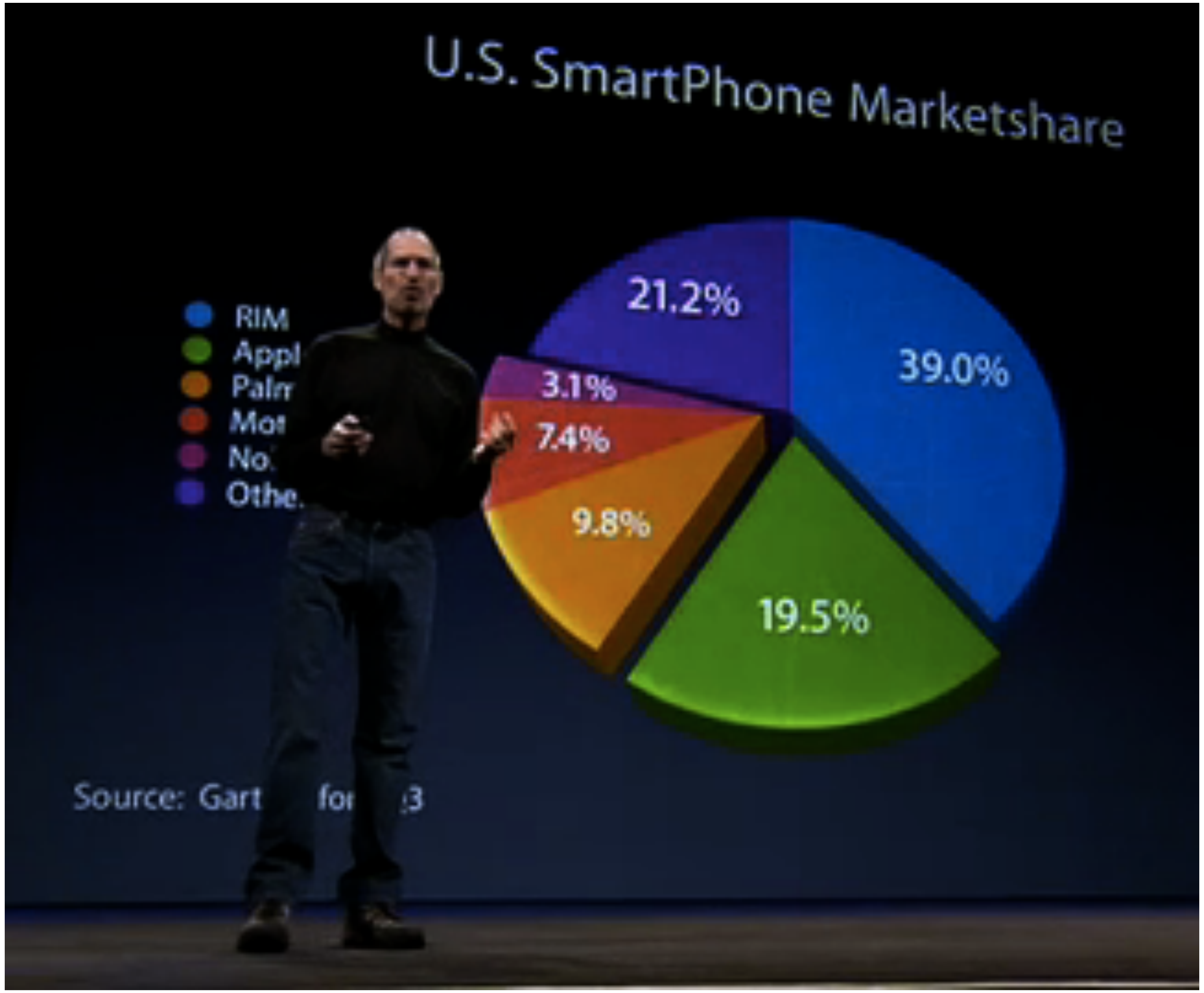


http://help.infragistics.com/Help/Doc/WinForms/2014.2/CLR4.0/html/Images/Chart_Bar_Chart_03.png

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"

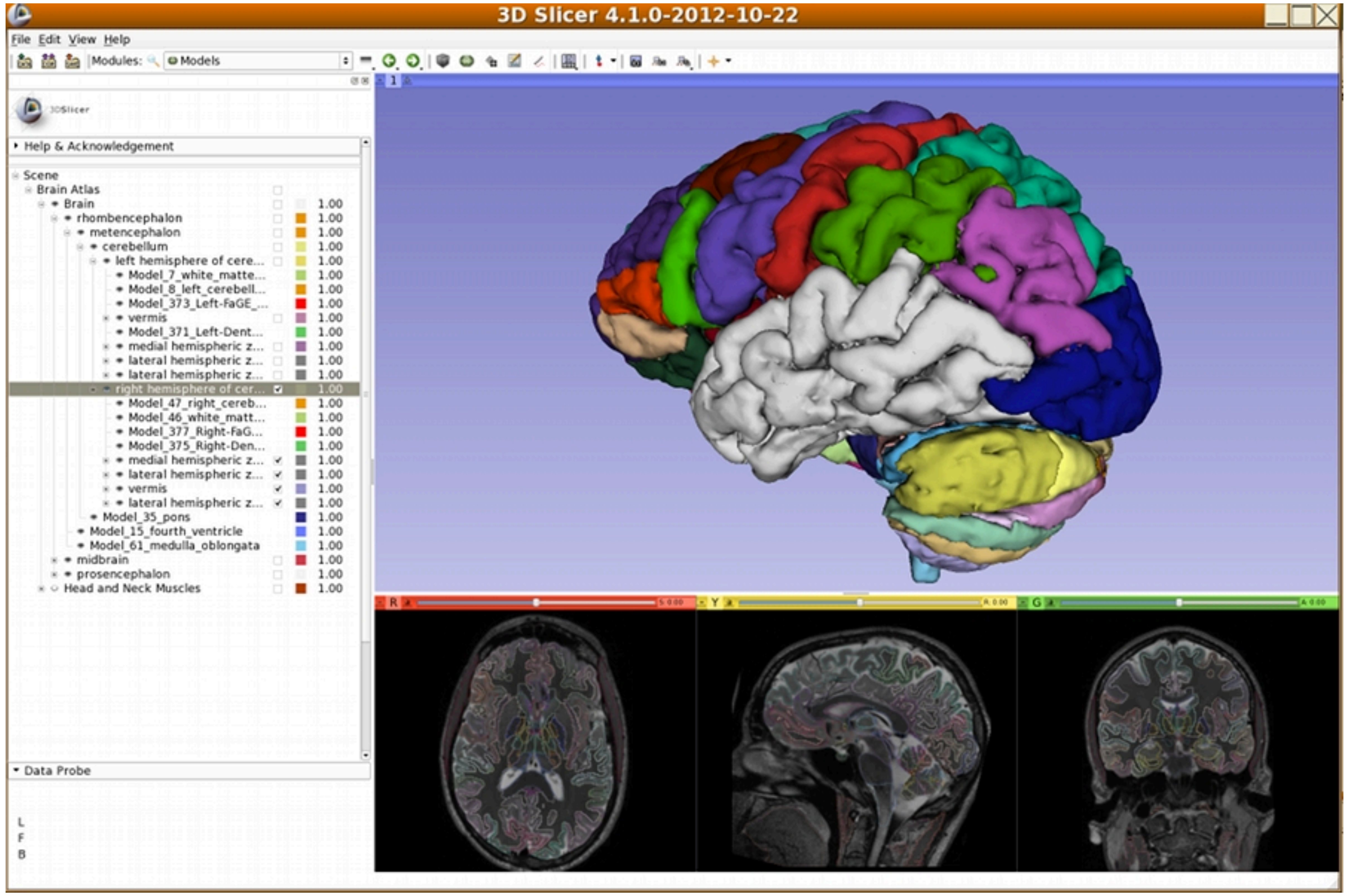


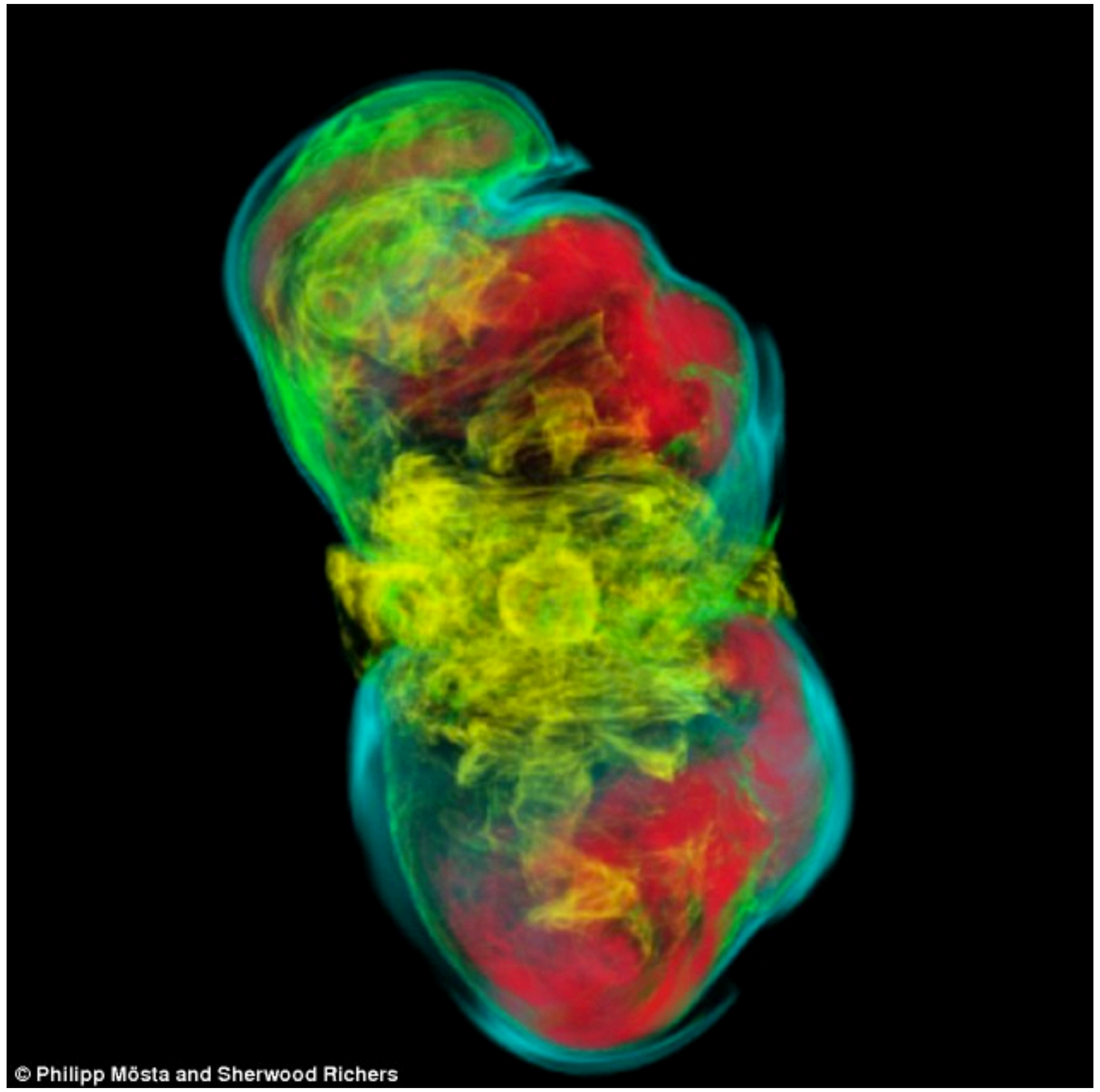
Unjustified 3D!

Lie factor!

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

“No <u>Unjustified</u> 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.”

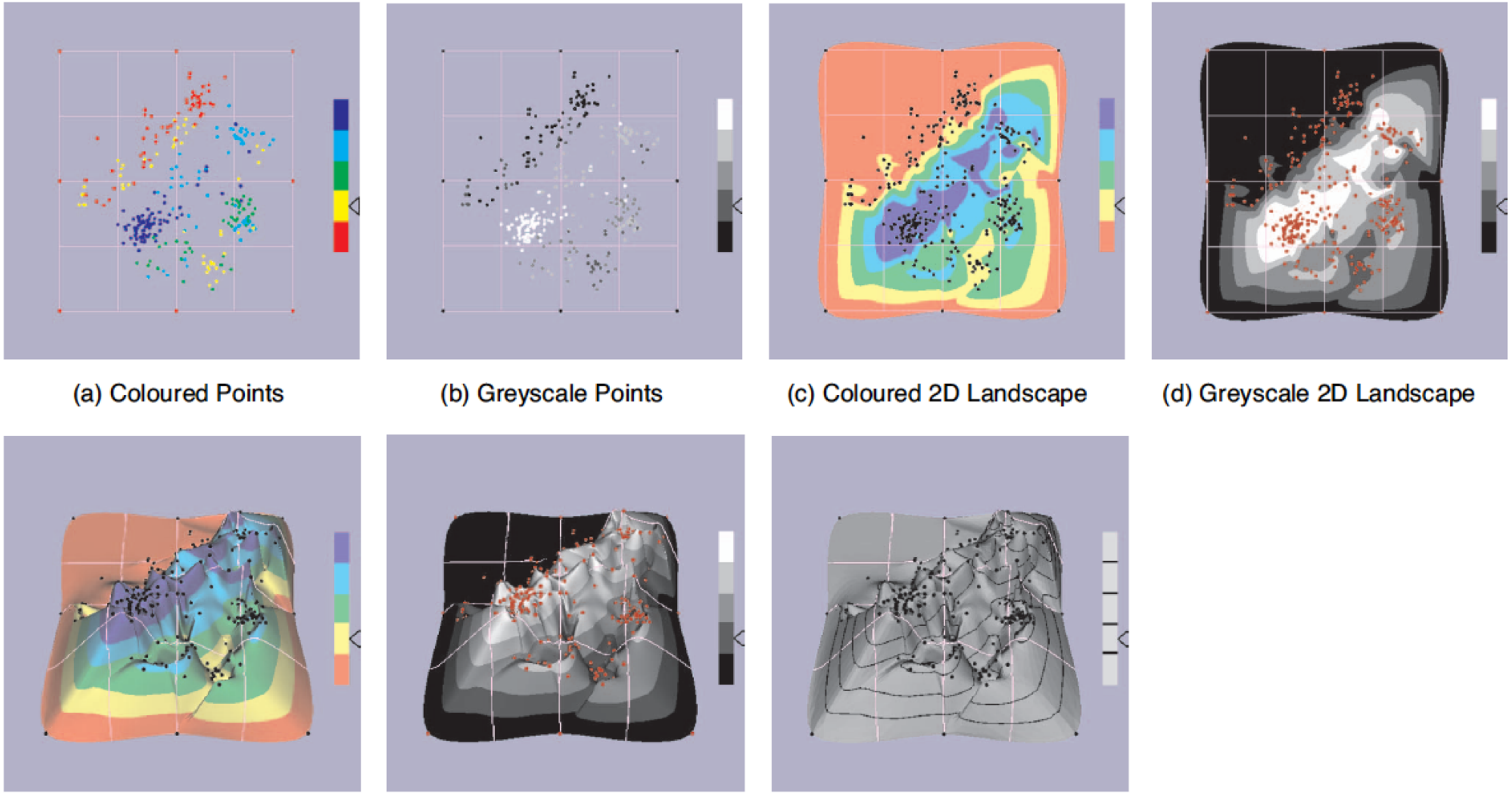
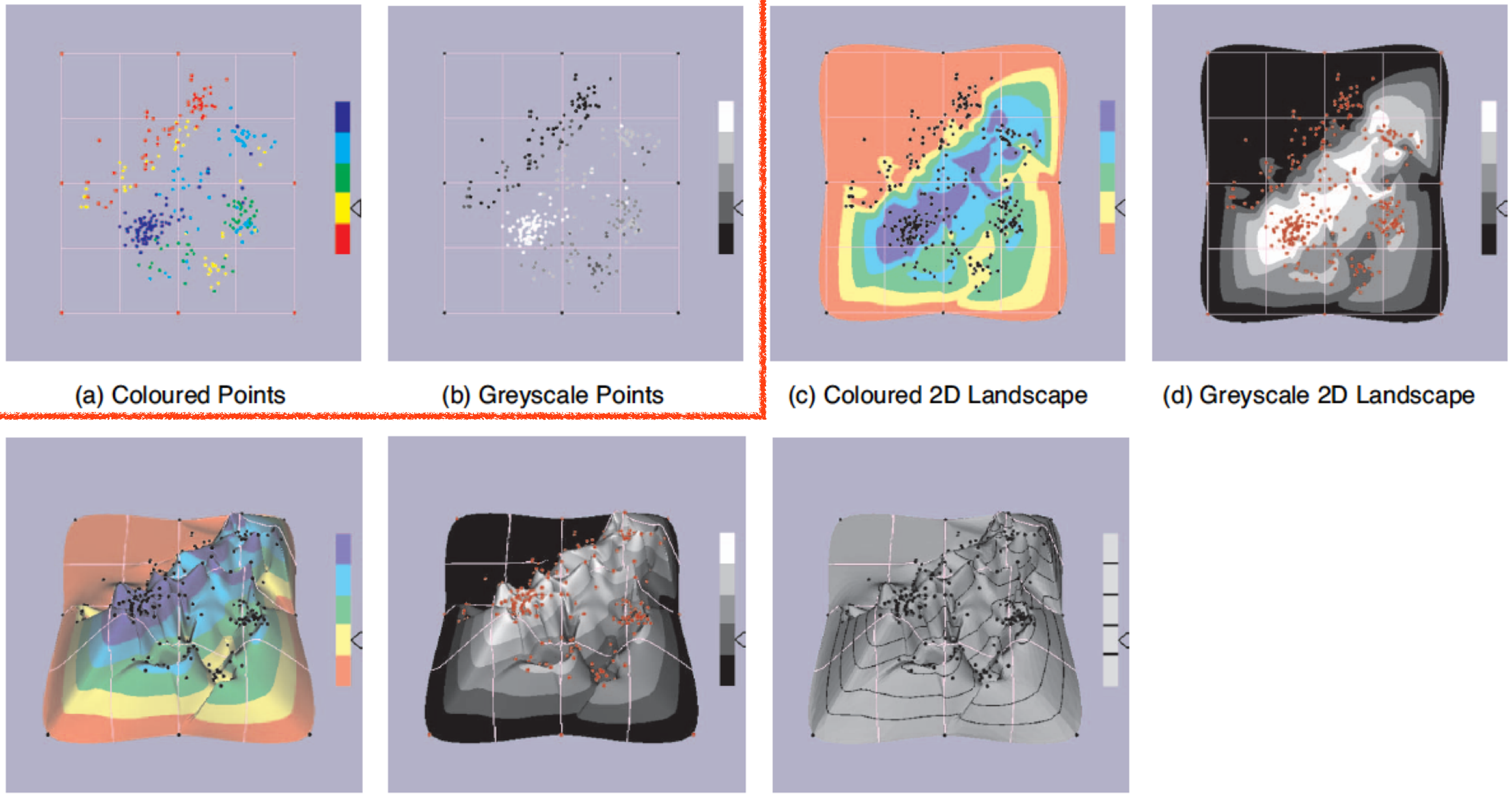

(a) Coloured Points
(b) Greyscale Points
(c) Coloured 2D Landscape
(d) Greyscale 2D Landscape
(e) Coloured 3D Landscape
(f) Greyscale 3D Landscape
(g) Height-only

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

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

Tory, et al. (2007)



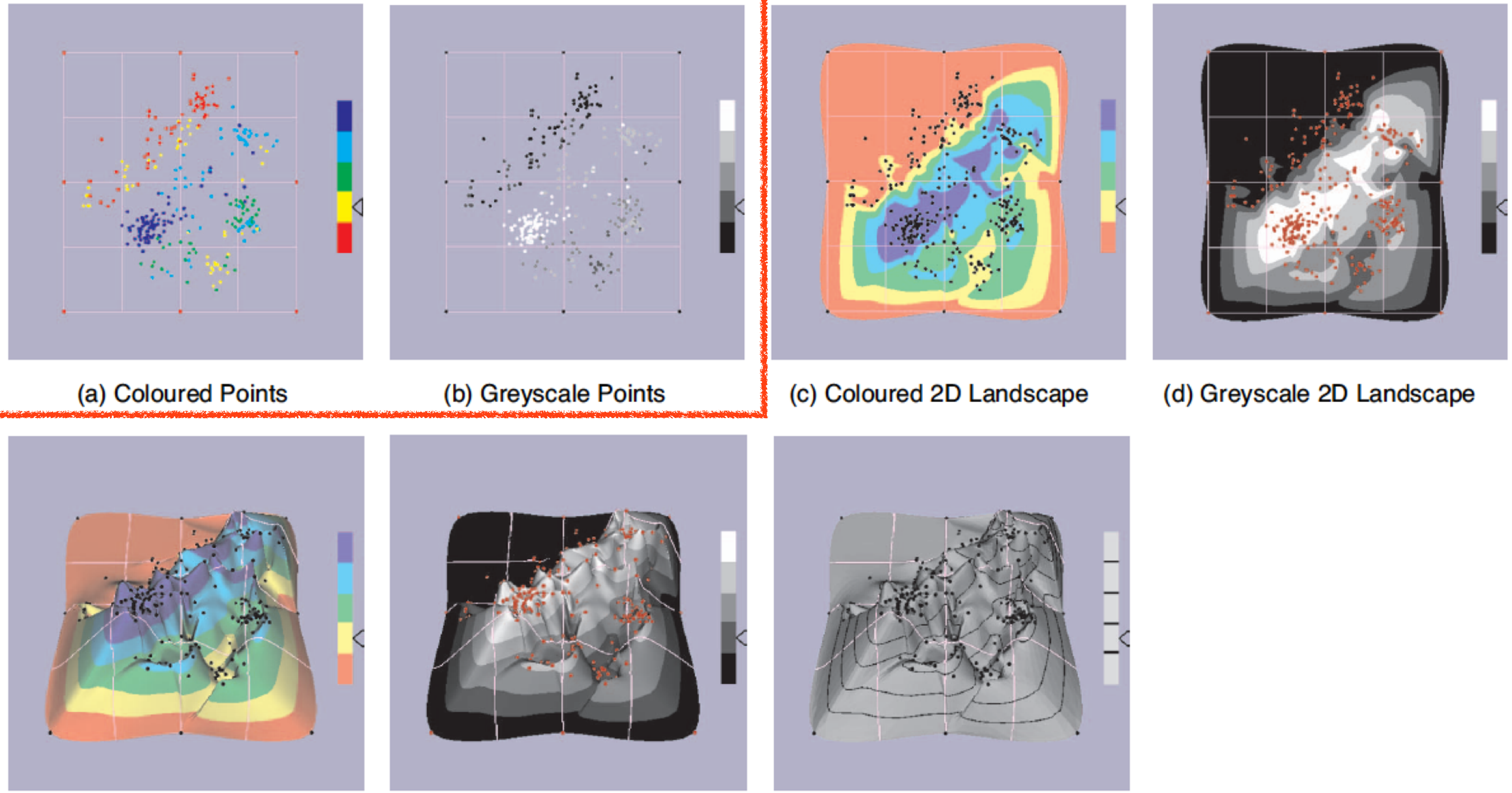
(a) Coloured Points (b) Greyscale Points (c) Coloured 2D Landscape (d) Greyscale 2D Landscape

(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.

"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

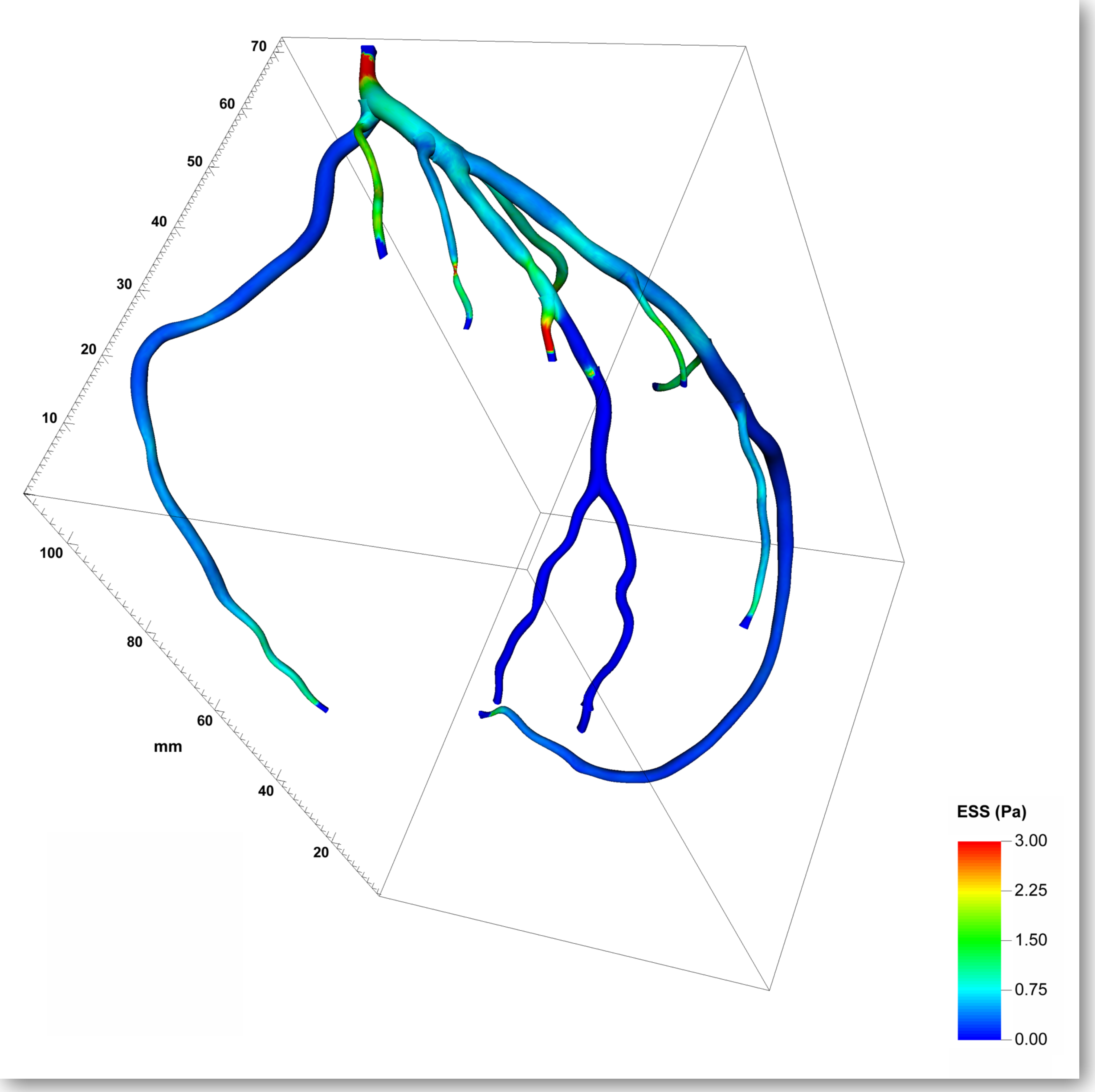
Fig. 1 Point-based displays and information landscapes used in our experiment. All displays show the same data.

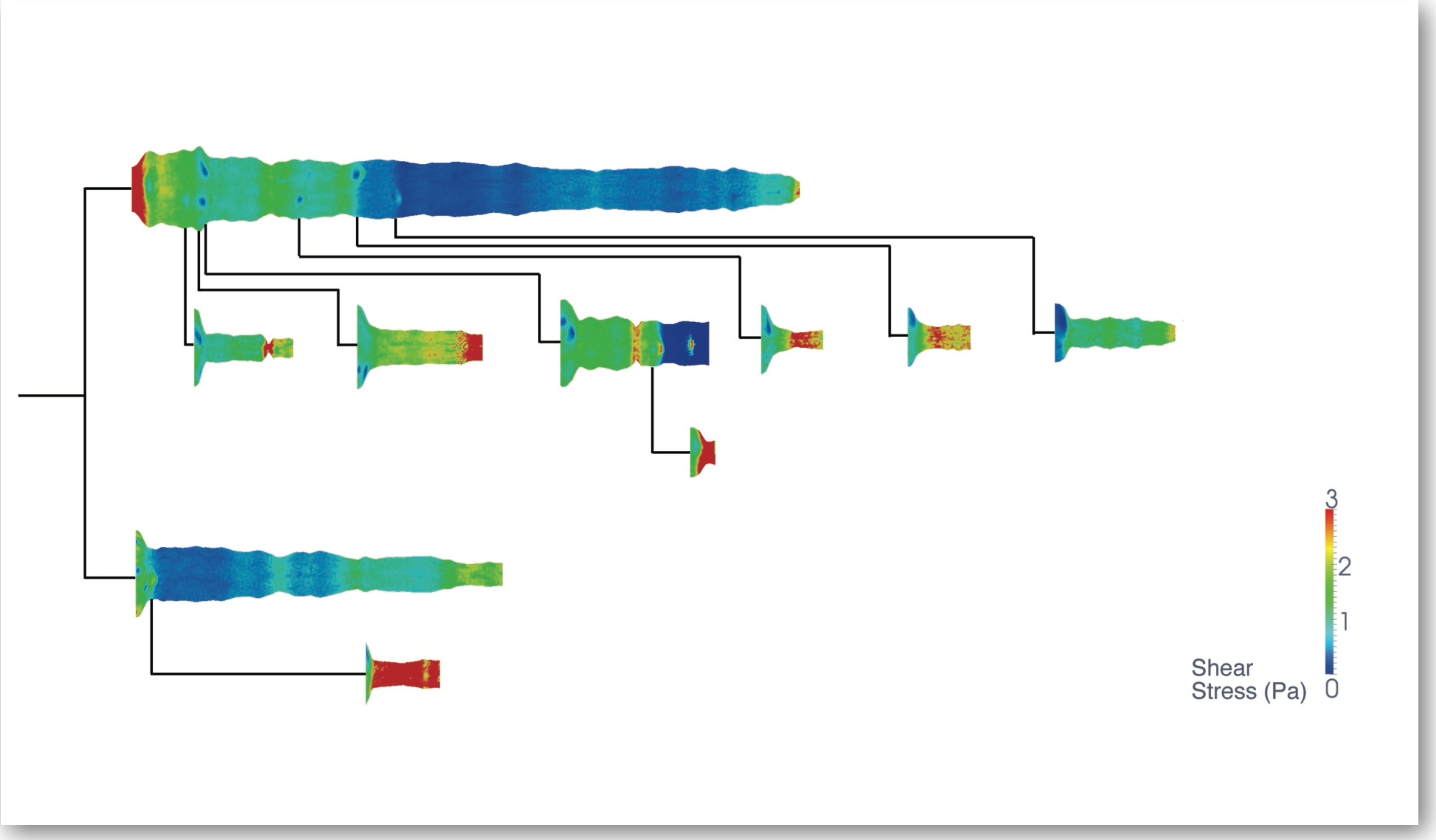
Dr. David Sprague
(Lecturer, CCIS)

"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)

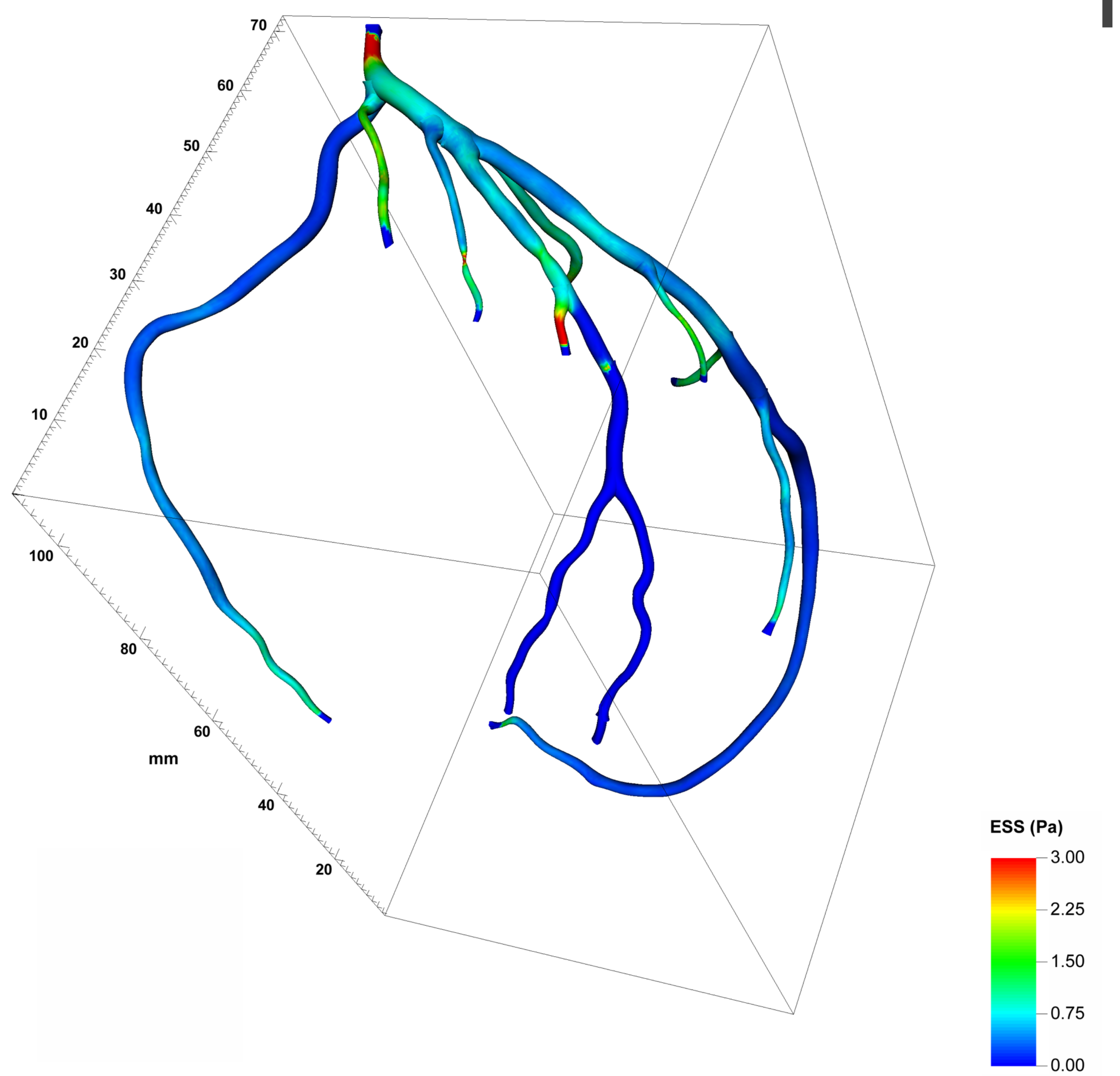
ACCURACY

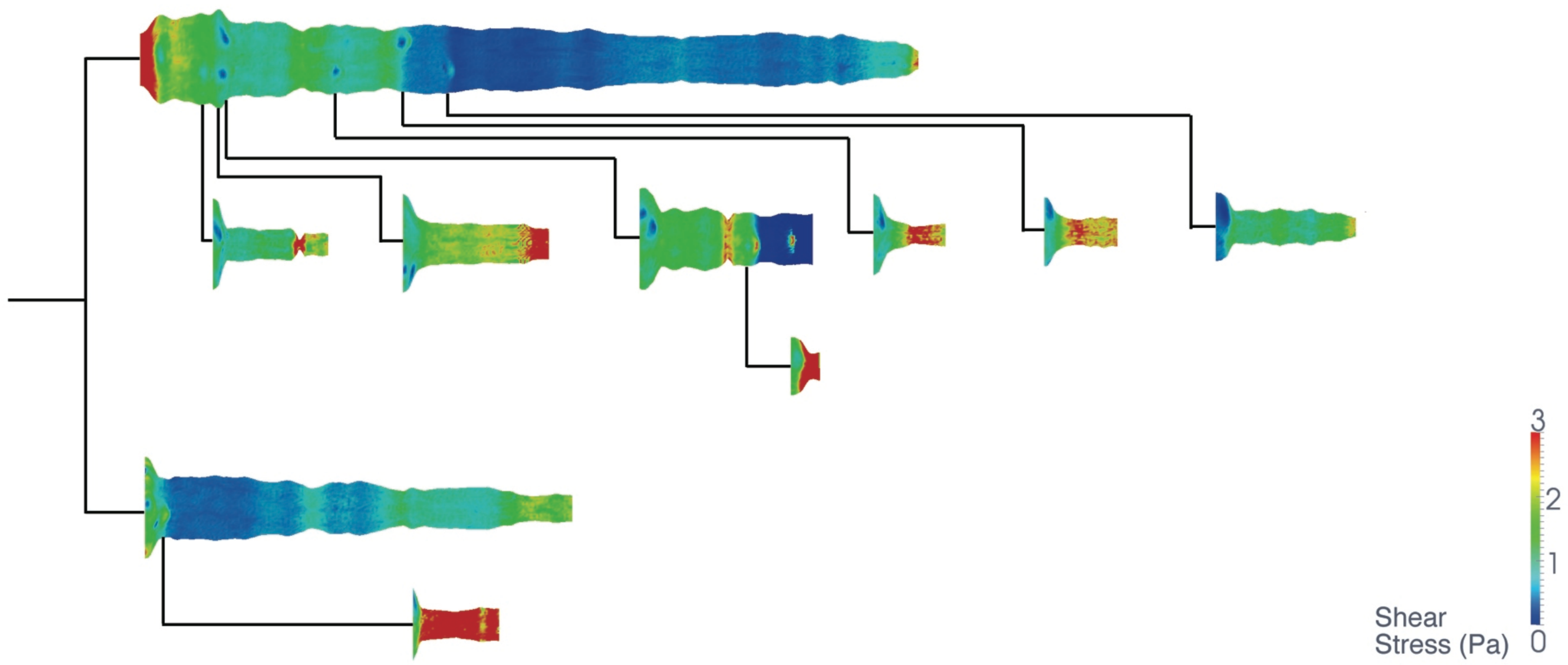
Strong effect of **dimensionality** on accuracy

39%

How many low ESS regions found?

62%





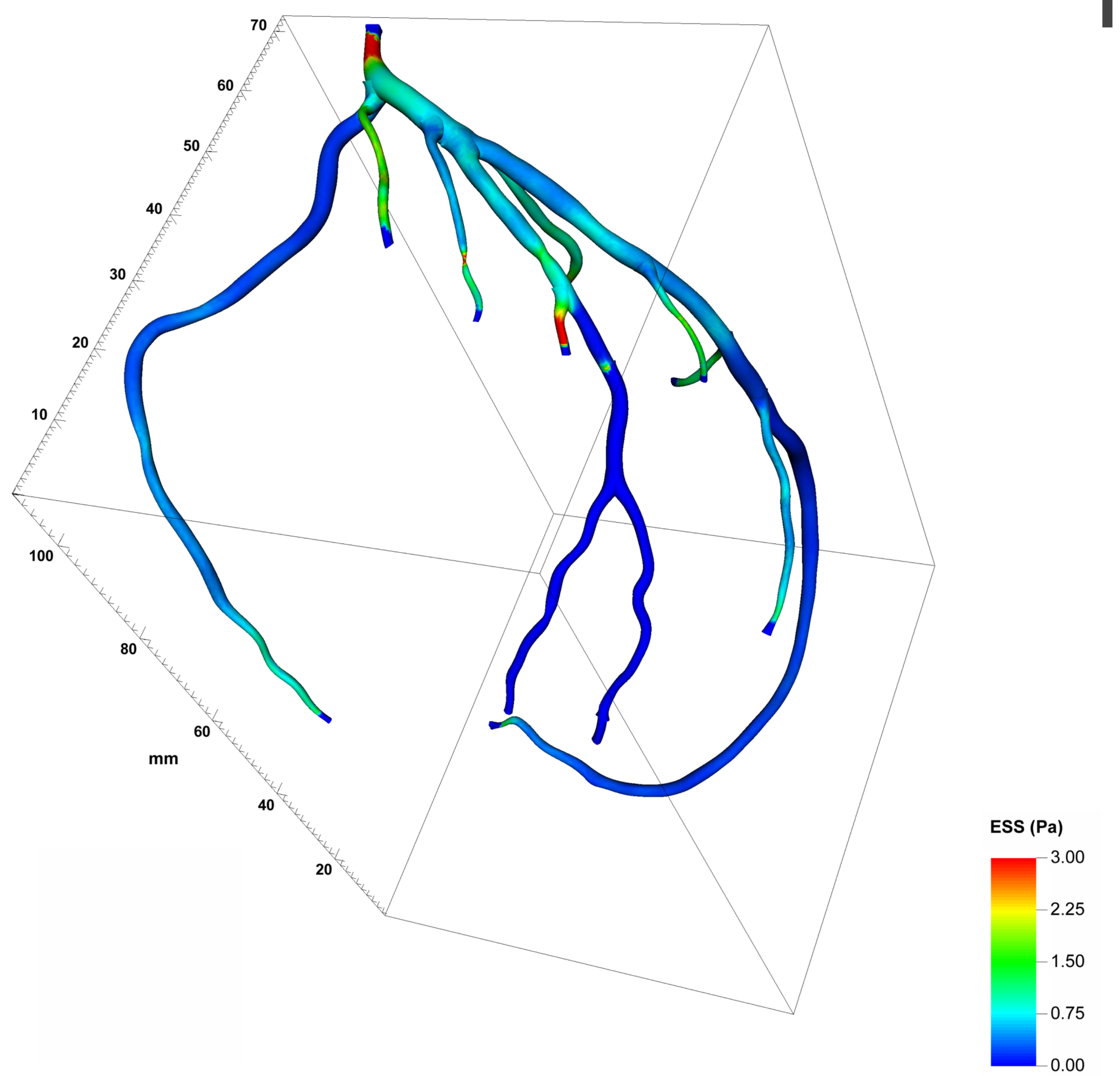
Accuracy

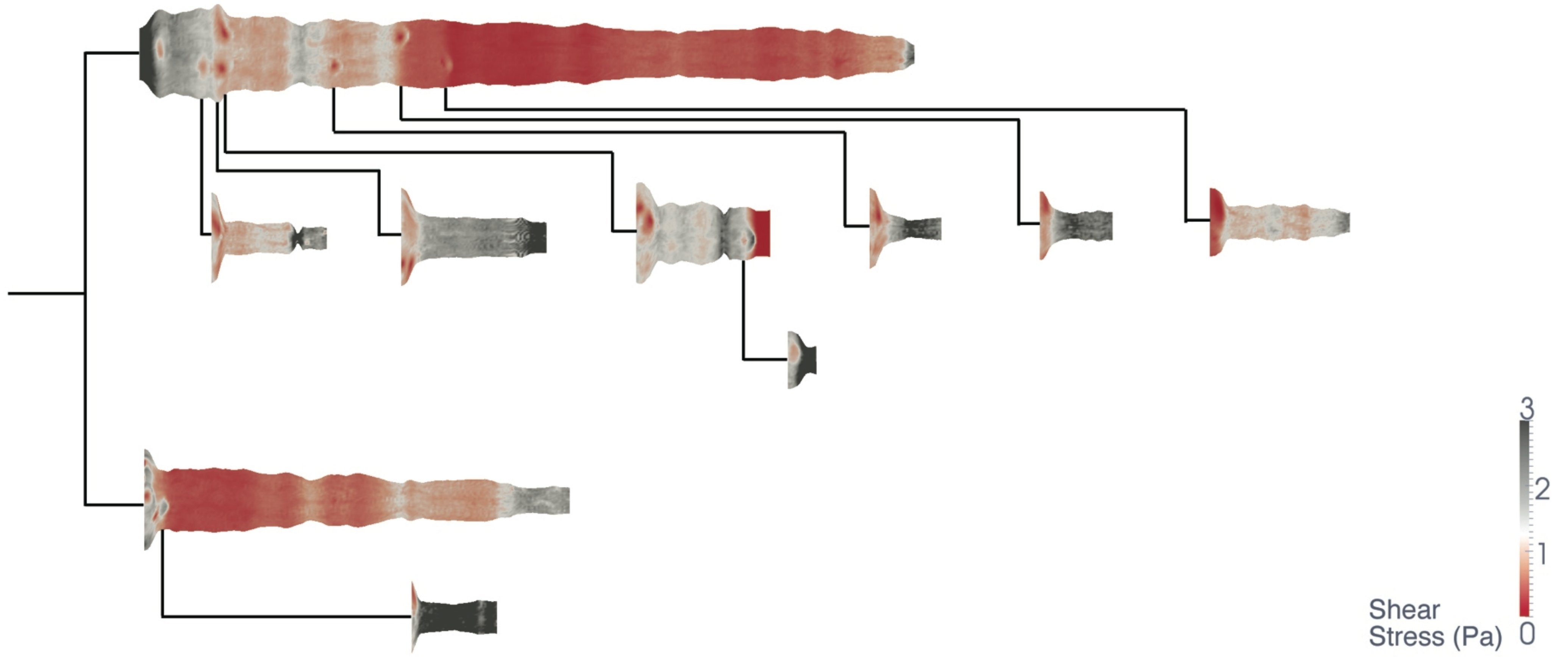
Strong effect of **dimensionality** on accuracy ...as well as **color**

39%

How many low ESS regions found?

91%





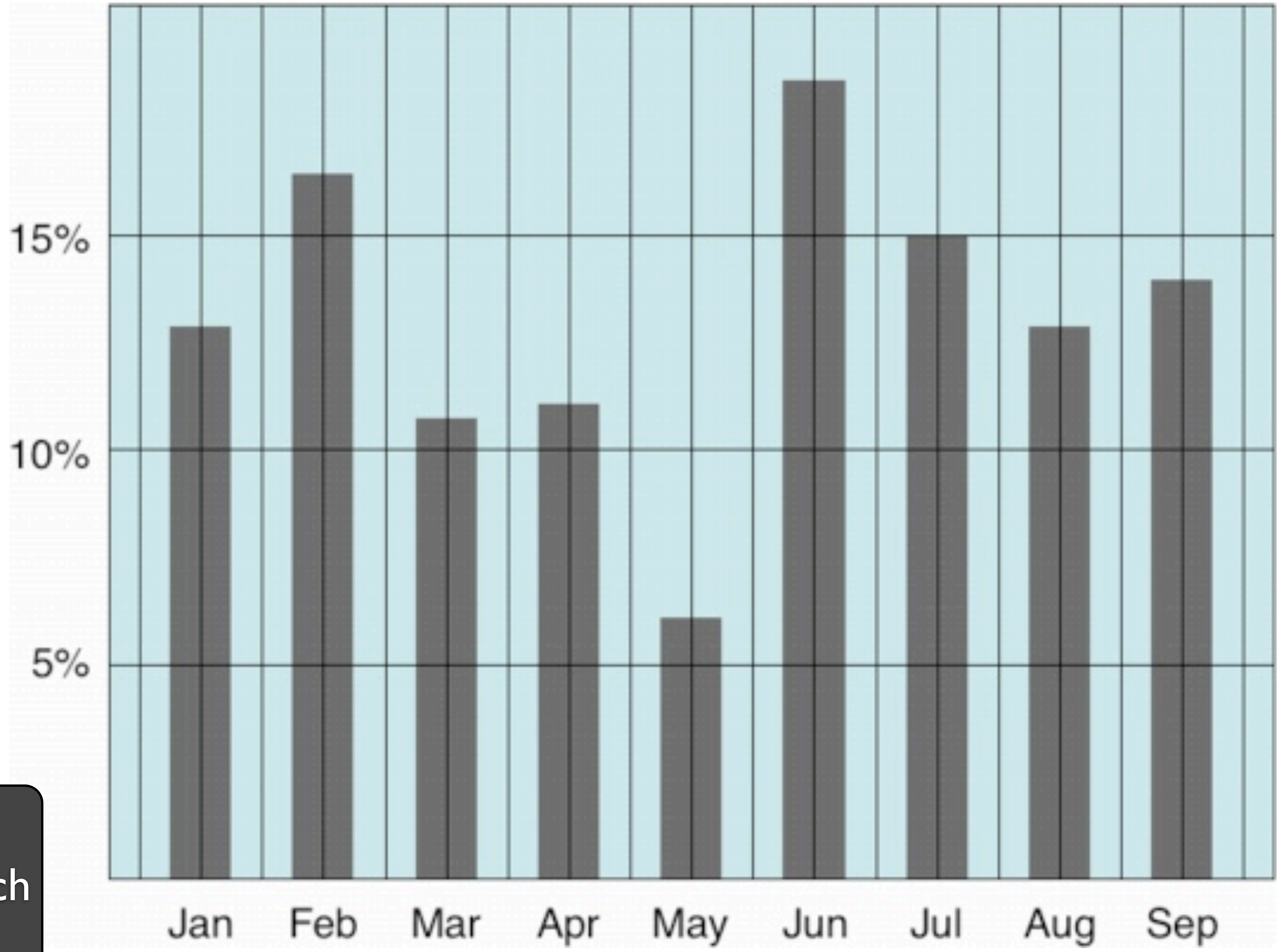
“Graphical Integrity”

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.

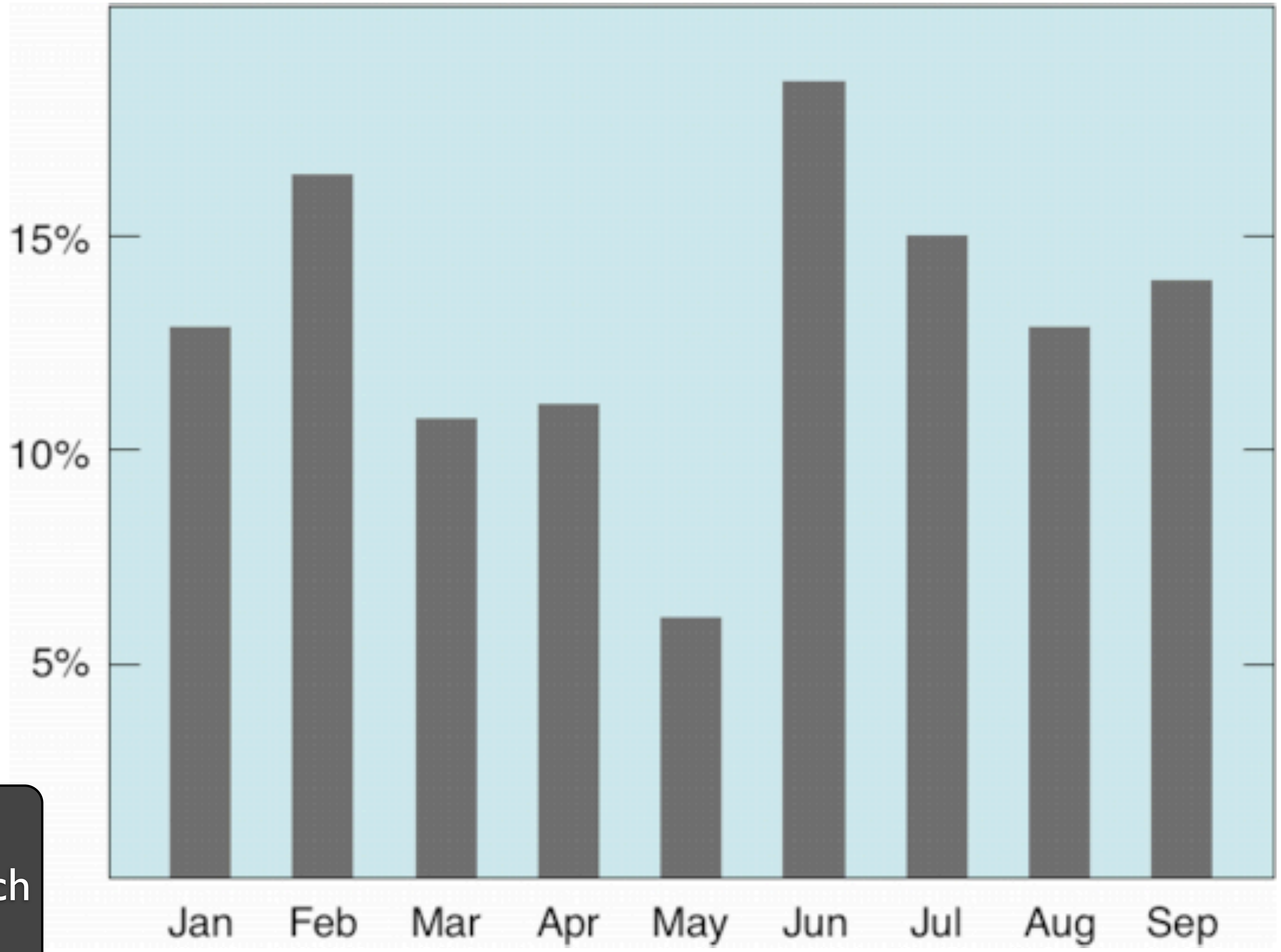
Tufte, “Visual Display of Quantitative Information” (1983)

“Graphical Integrity”



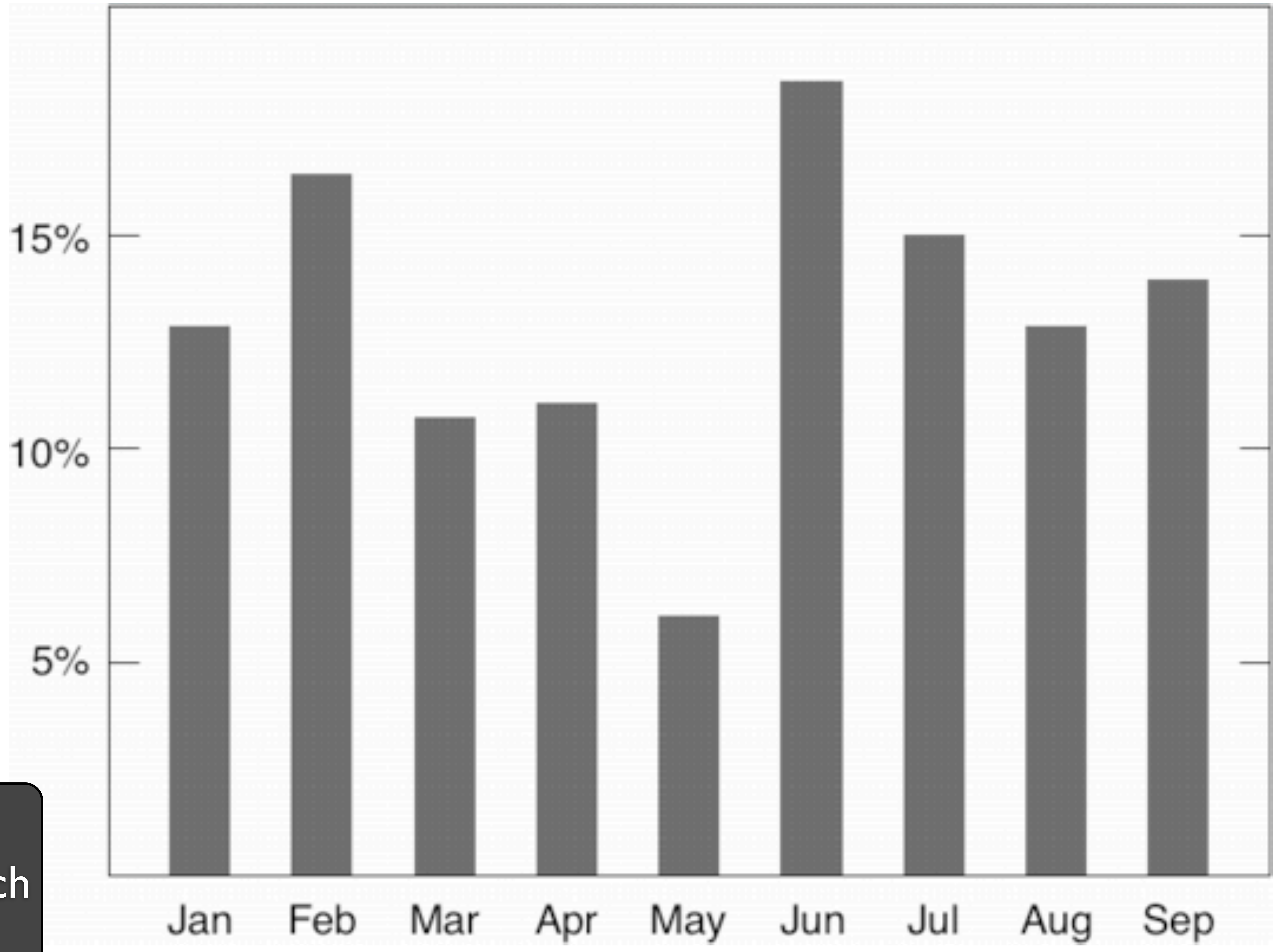
IN-CLASS ACTIVITY:
Use paper/pen to sketch
“Tufte” version!

“Graphical Integrity”



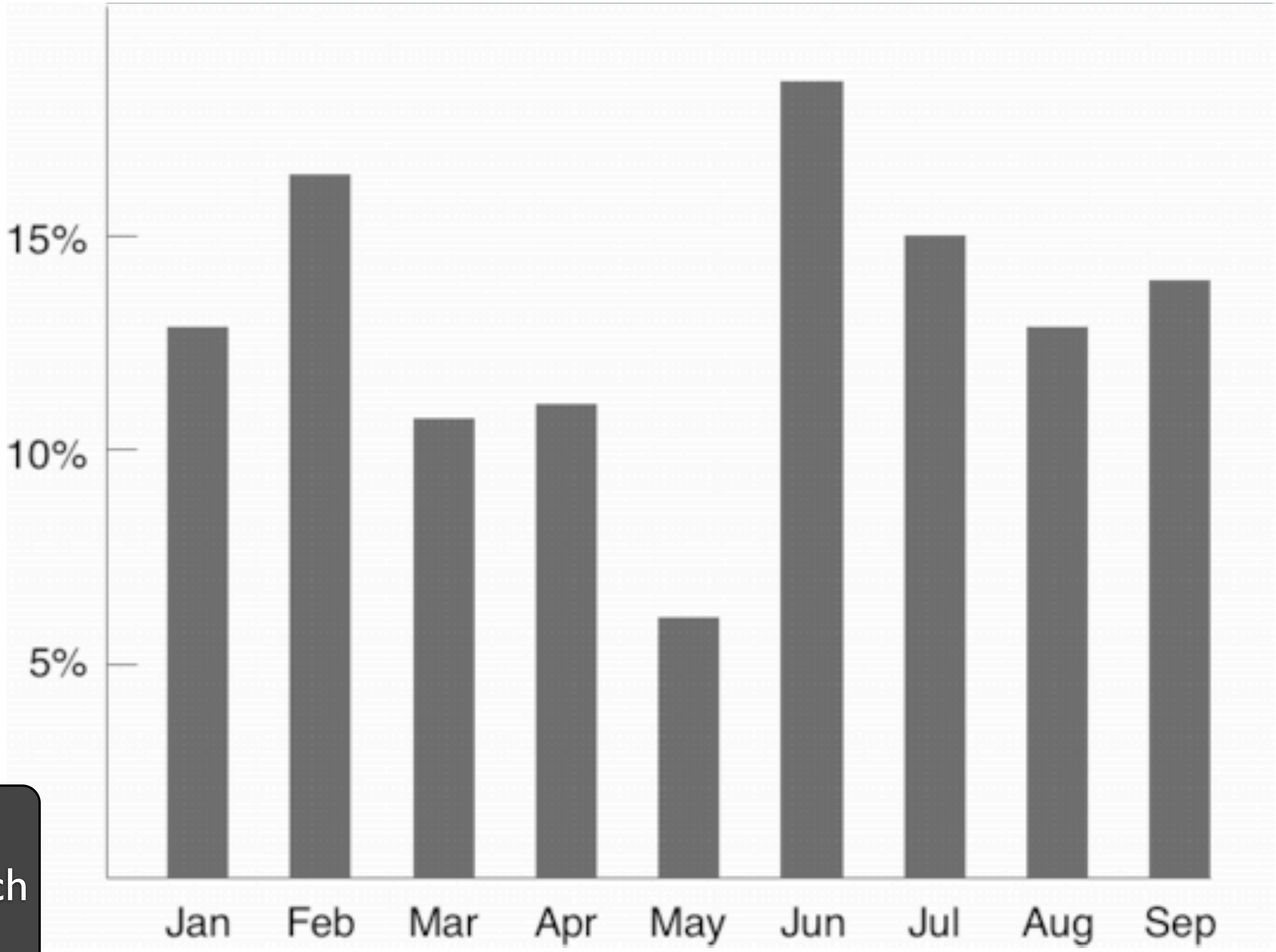
IN-CLASS ACTIVITY:
Use paper/pen to sketch
“Tufte” version!

“Graphical Integrity”



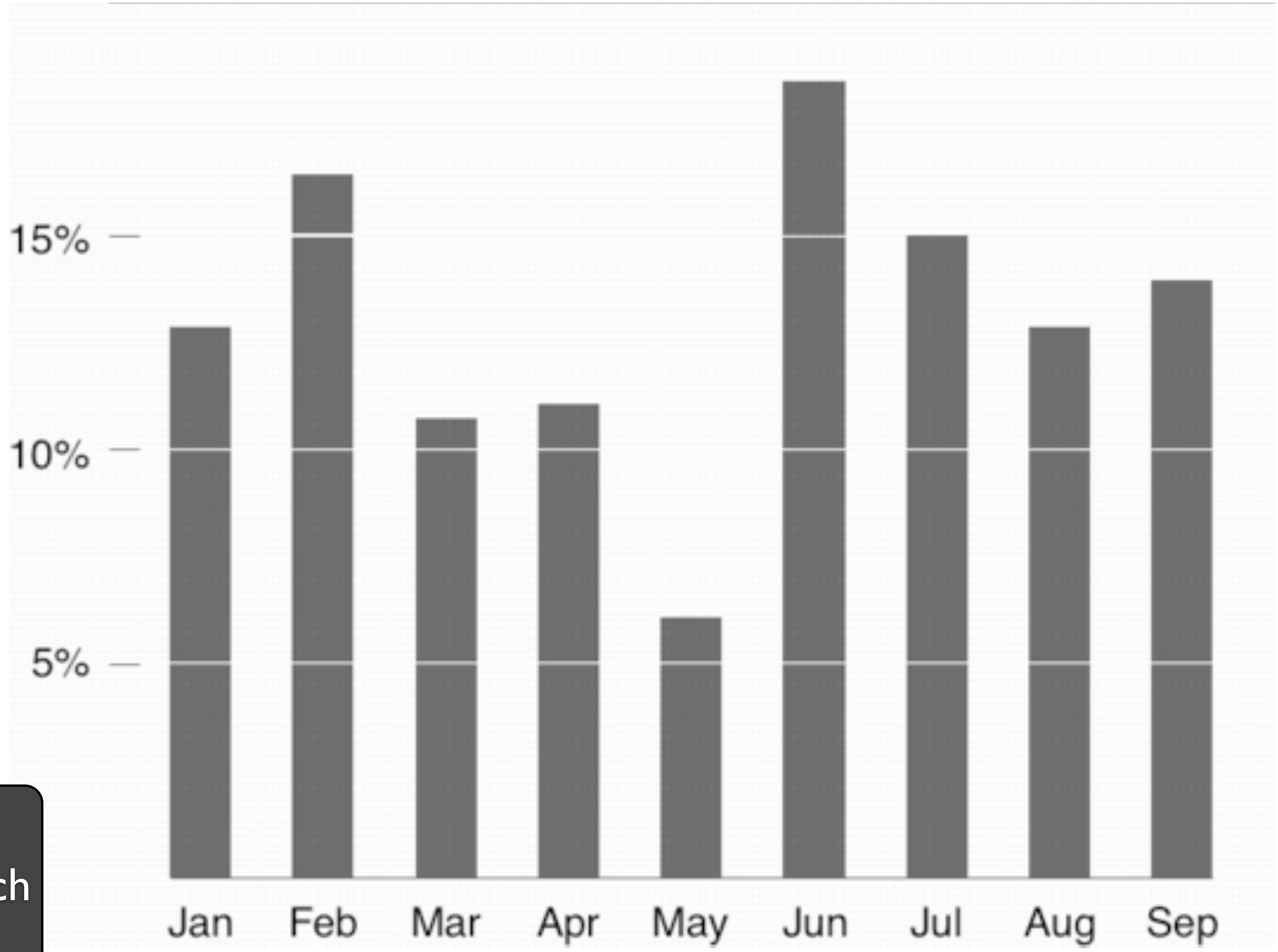
IN-CLASS ACTIVITY:
Use paper/pen to sketch
“Tufte” version!

"Graphical Integrity"



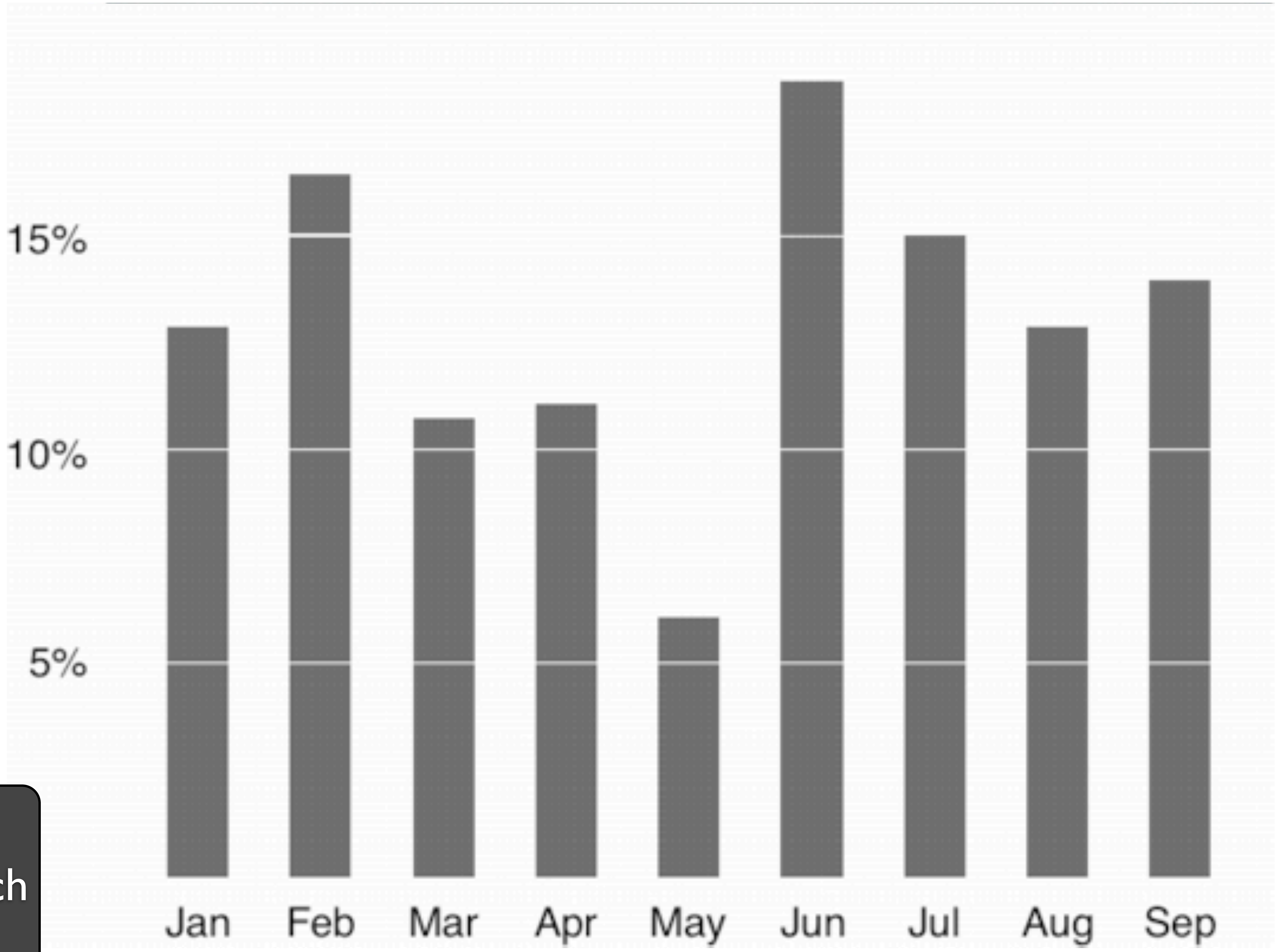
IN-CLASS ACTIVITY:
Use paper/pen to sketch
"Tufte" version!

“Graphical Integrity”



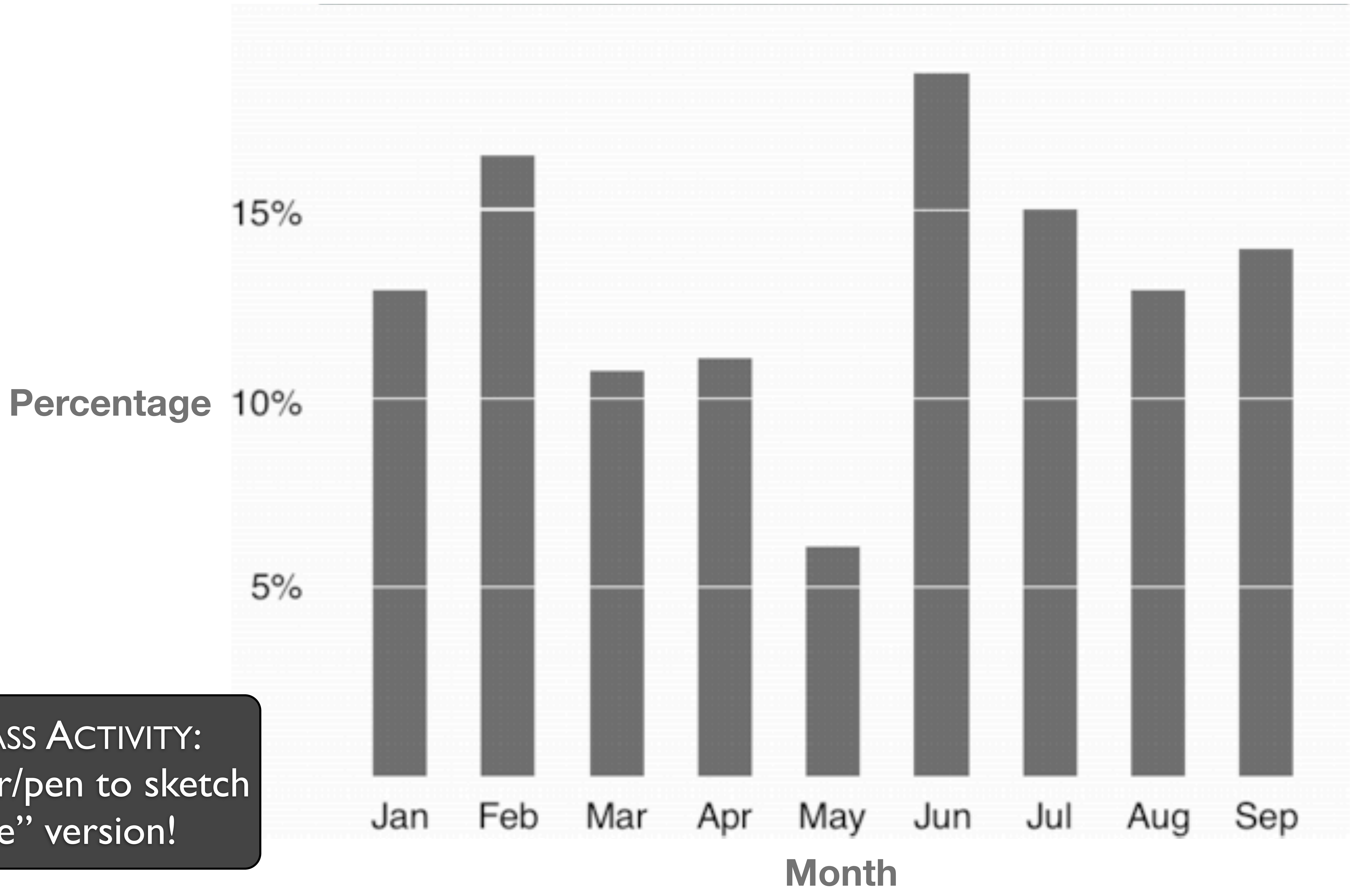
IN-CLASS ACTIVITY:
Use paper/pen to sketch
“Tufte” version!

"Graphical Integrity"



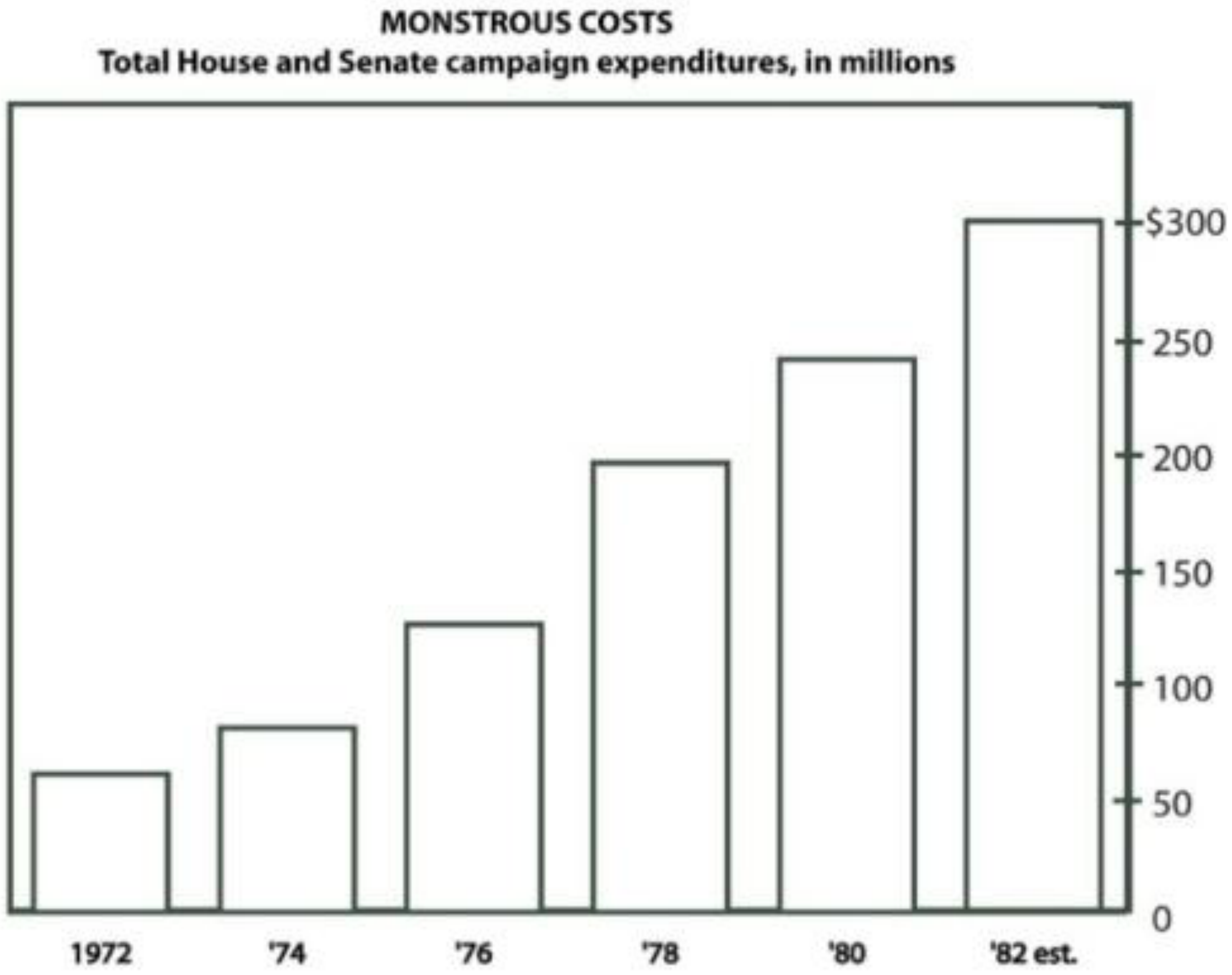
IN-CLASS ACTIVITY:
Use paper/pen to sketch "Tufte" version!

“Graphical Integrity”



IN-CLASS ACTIVITY:
Use paper/pen to sketch
“Tufte” version!

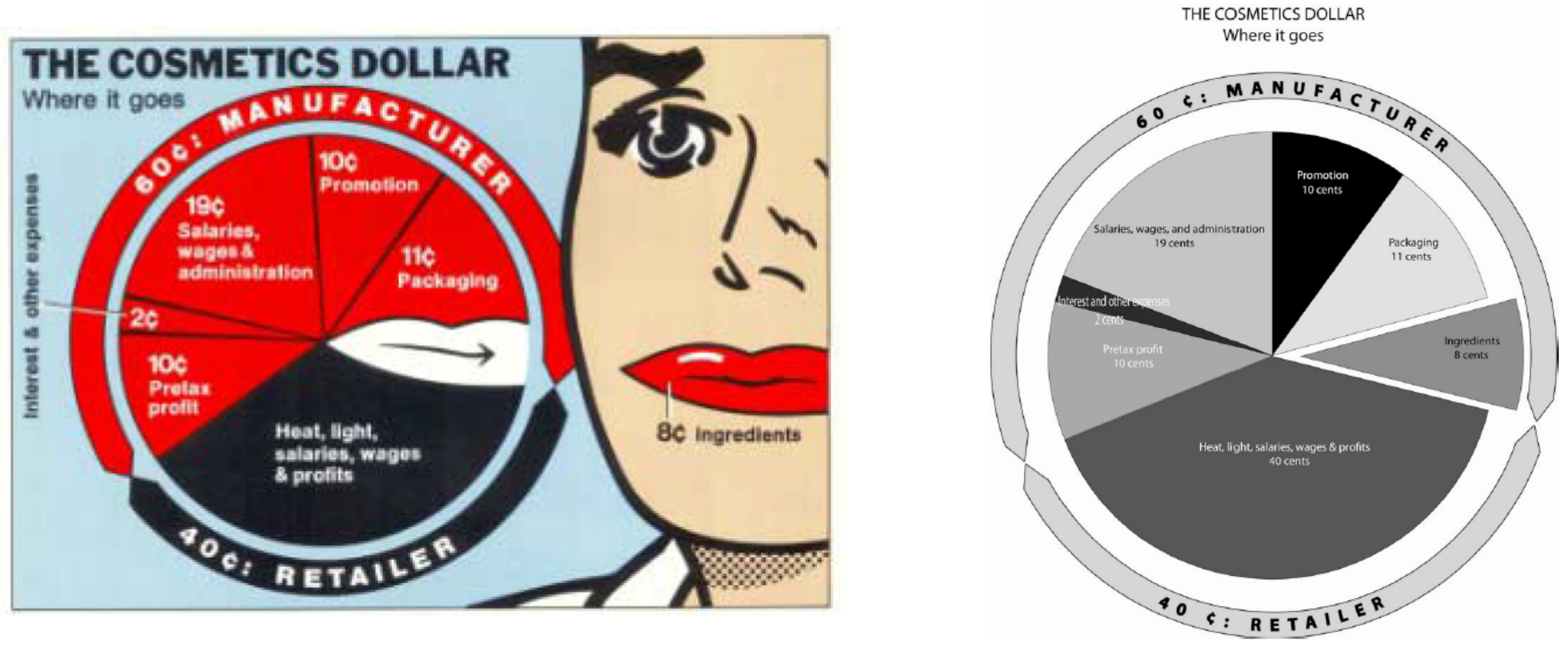
"Chart Junk"



Bateman, et al. (2010)

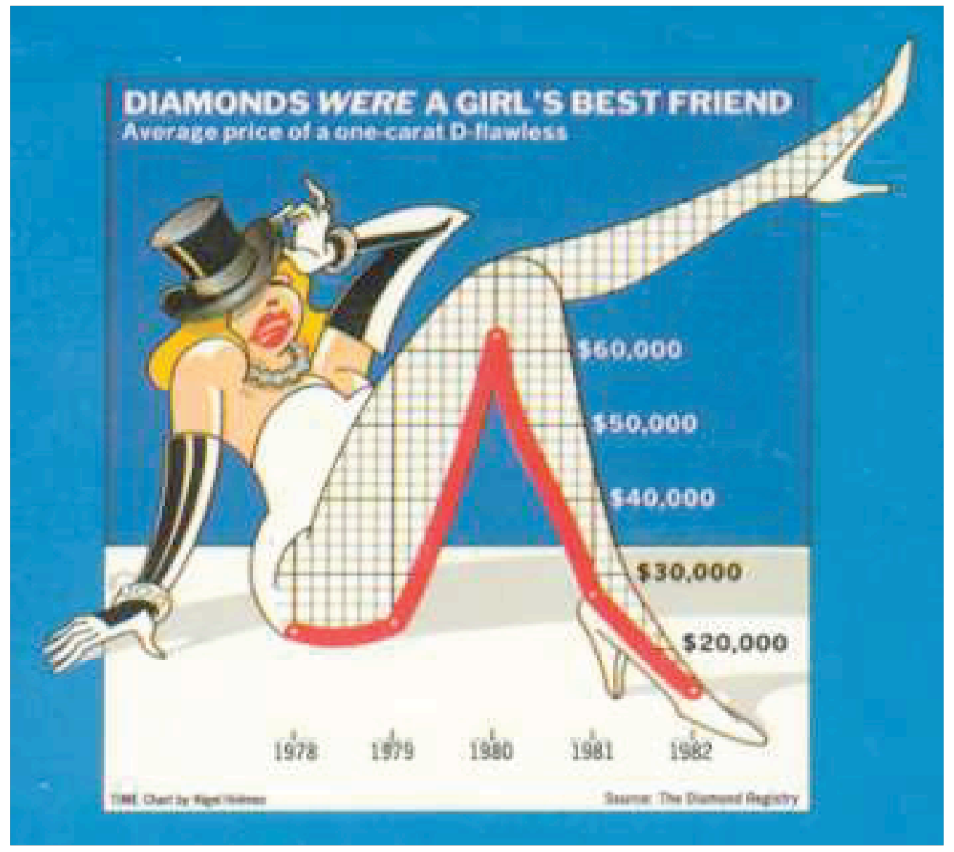
“Chart Junk Debate”

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



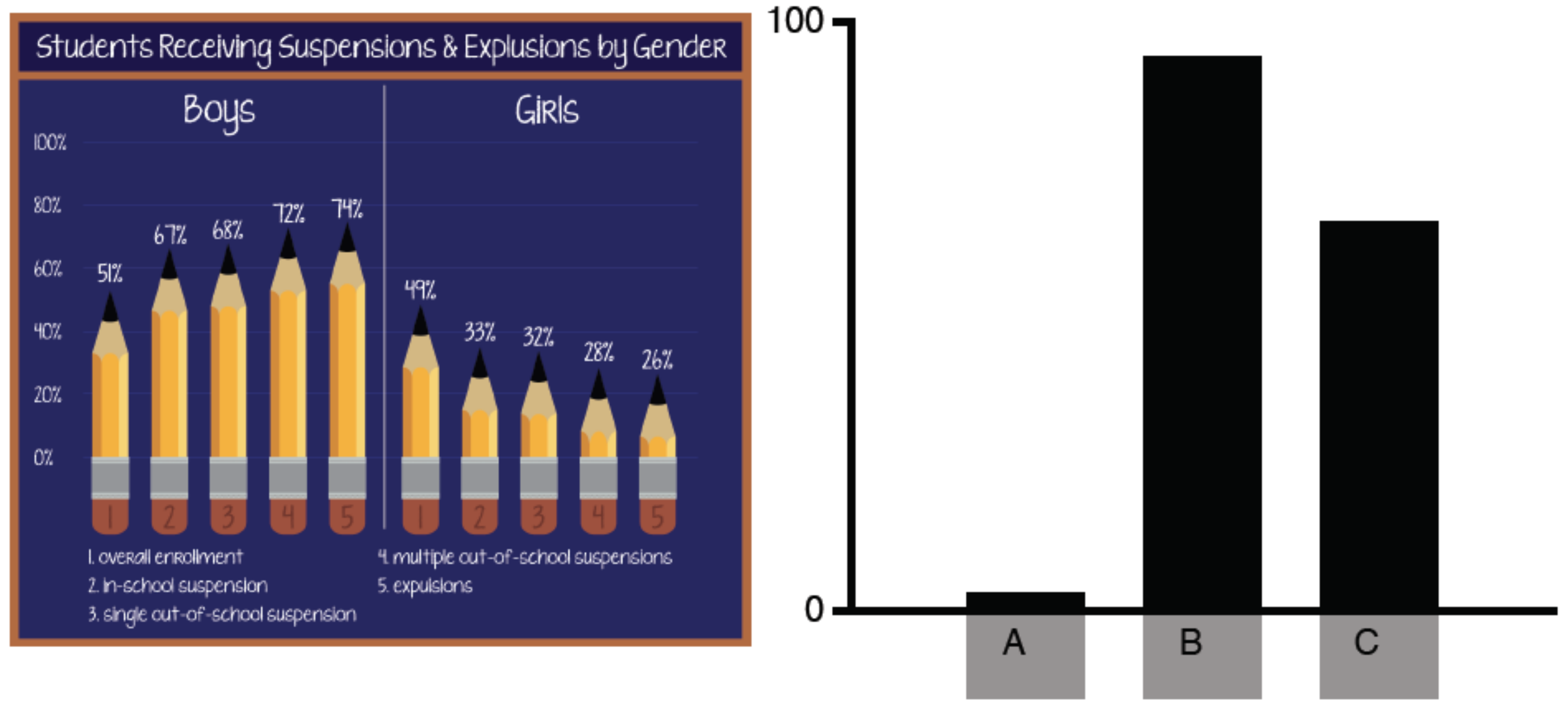
Bateman, et al. (2010)

Benefitting InfoVis with Visual Difficulties



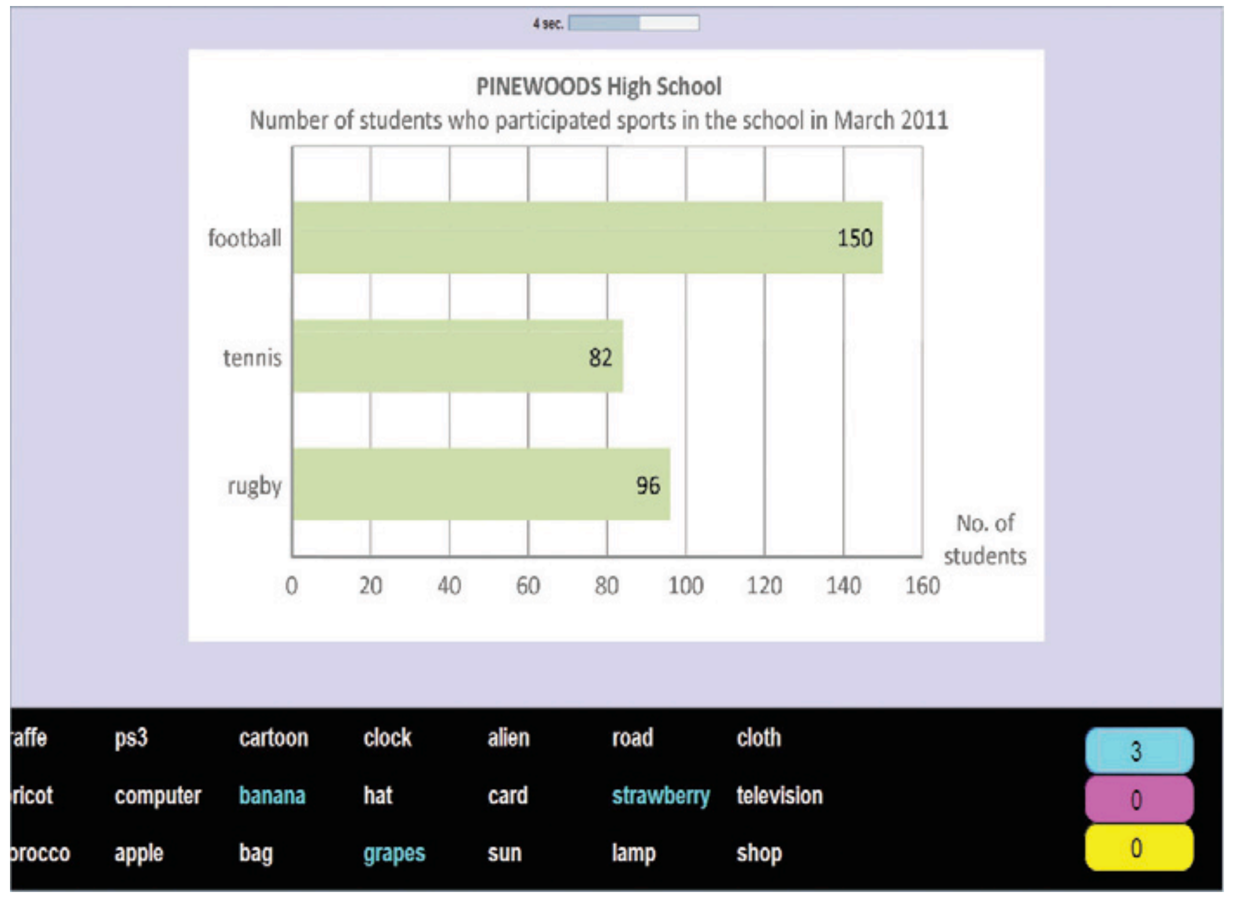
Hullman, et al. (2011)

An Evaluation of the Impact of Visual Embellishments in Bar Charts



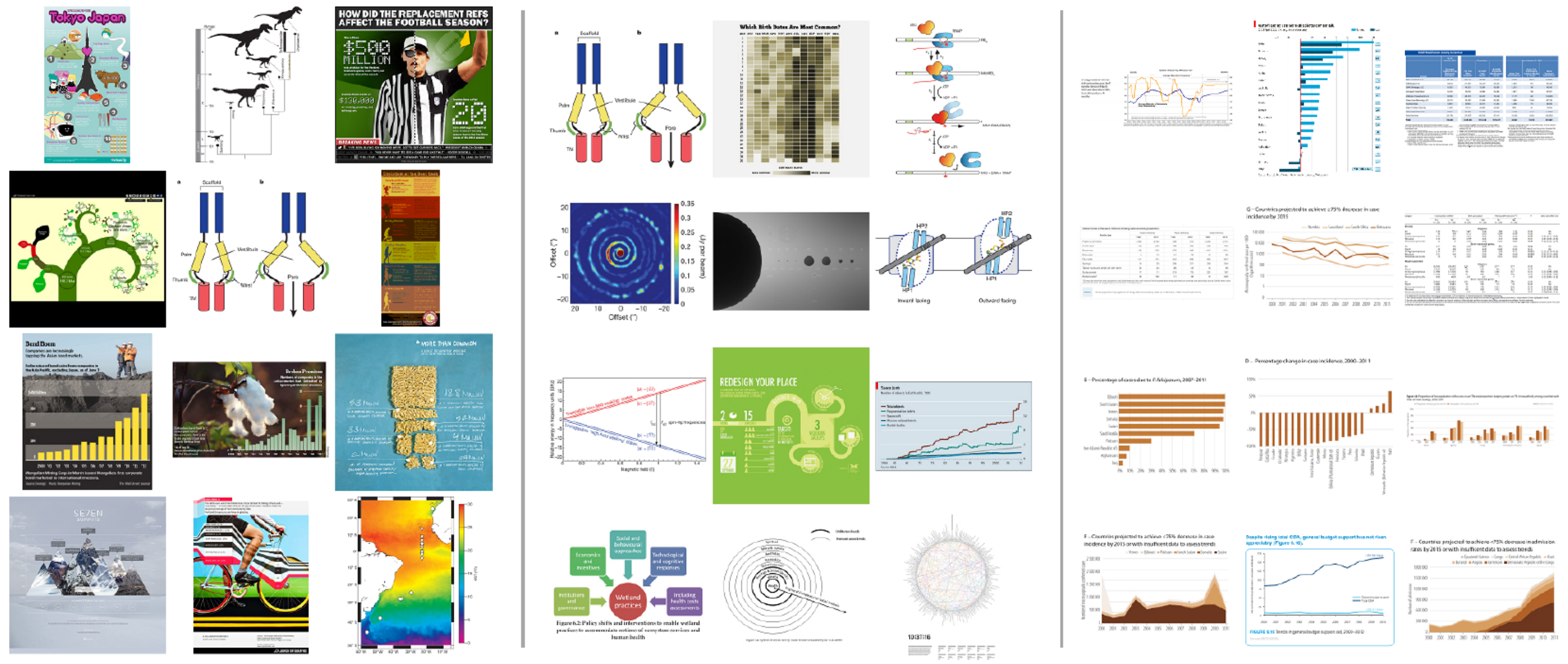
Skau, et al. (2015)

An Empirical Study on Using Visual Embellishments in Visualization



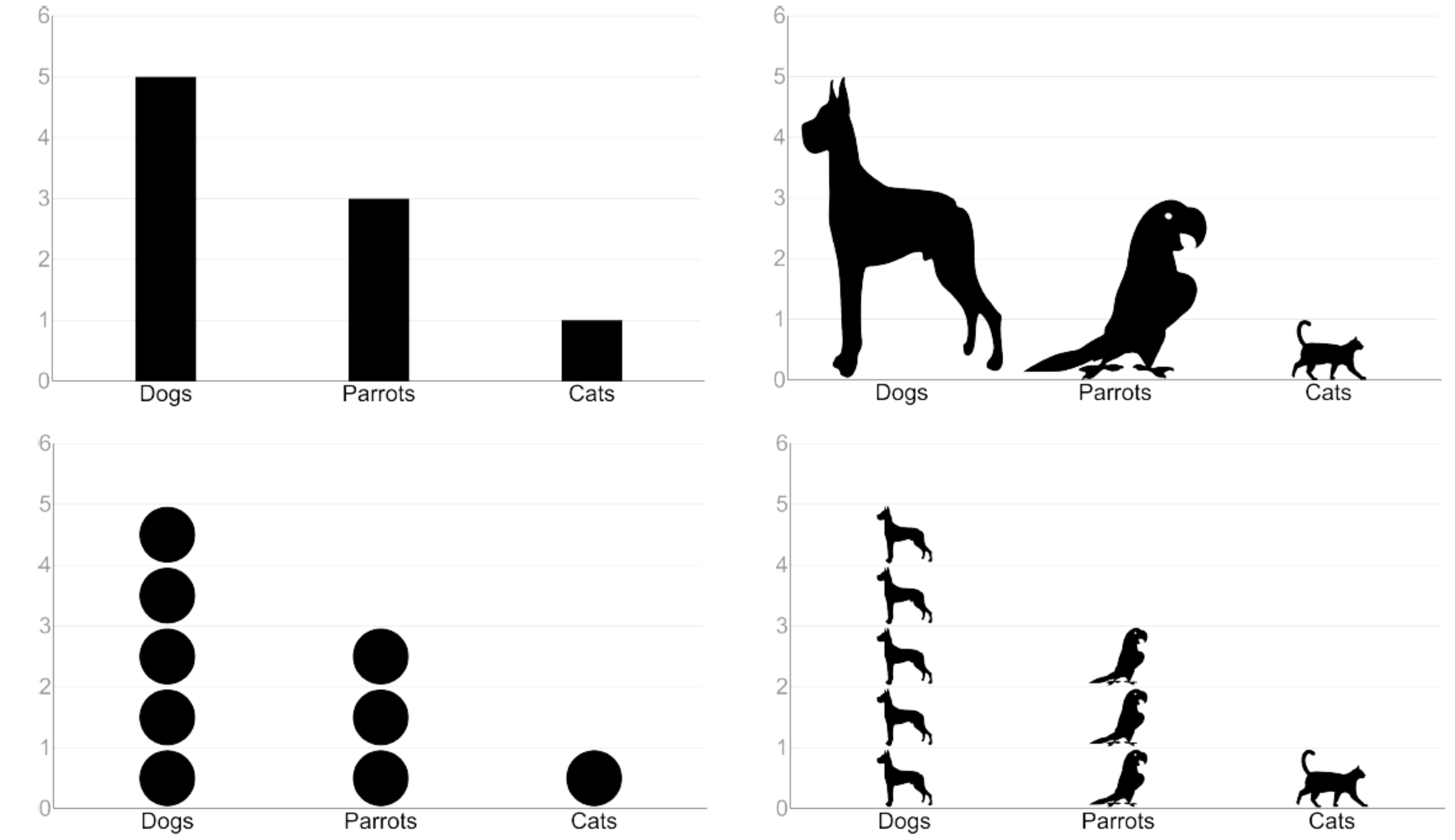
Borgo, et al. (2012)

What makes a visualization memorable?



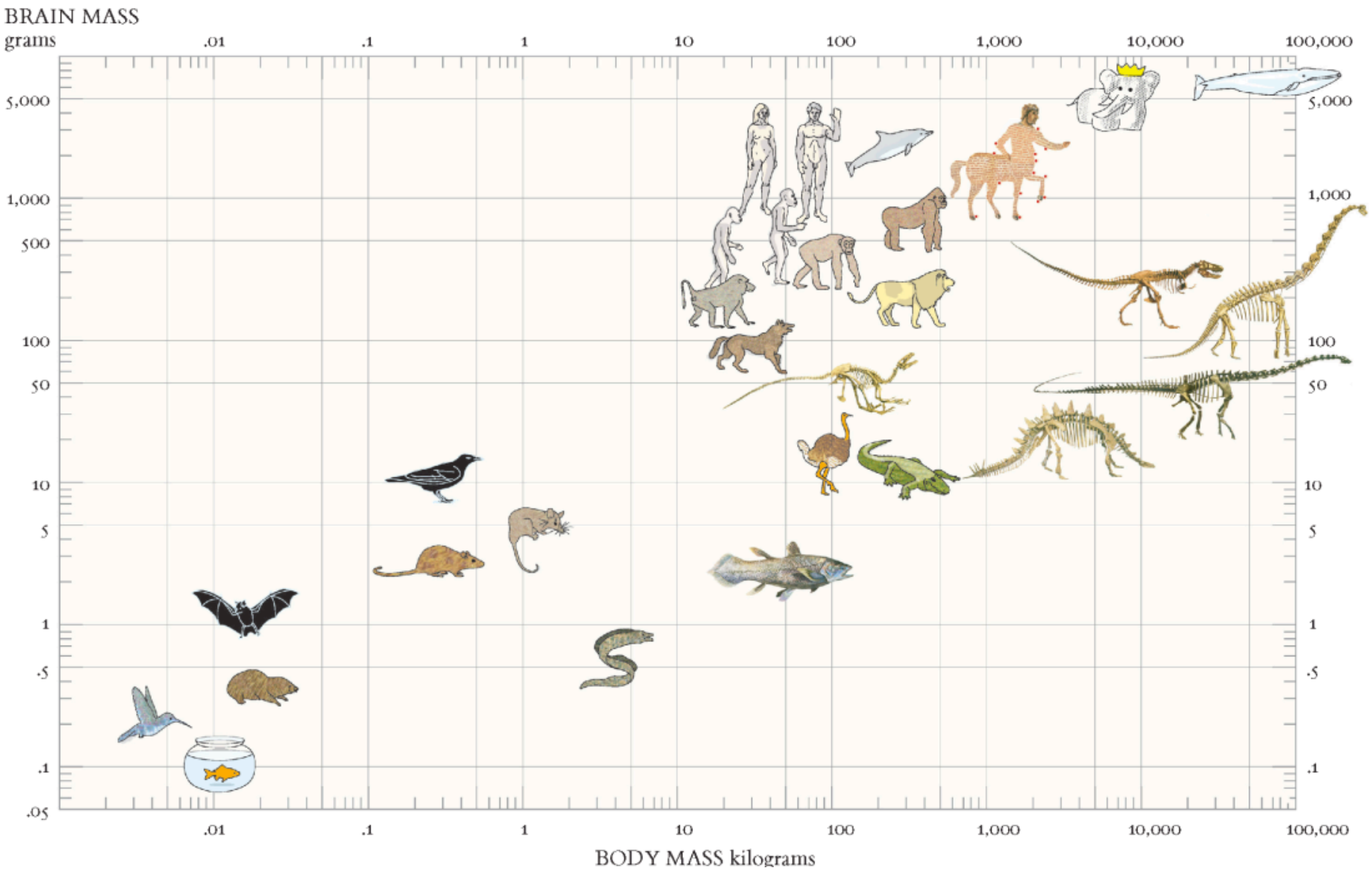
Borkin, et al. (2013)
Borkin, et al. (2015)

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



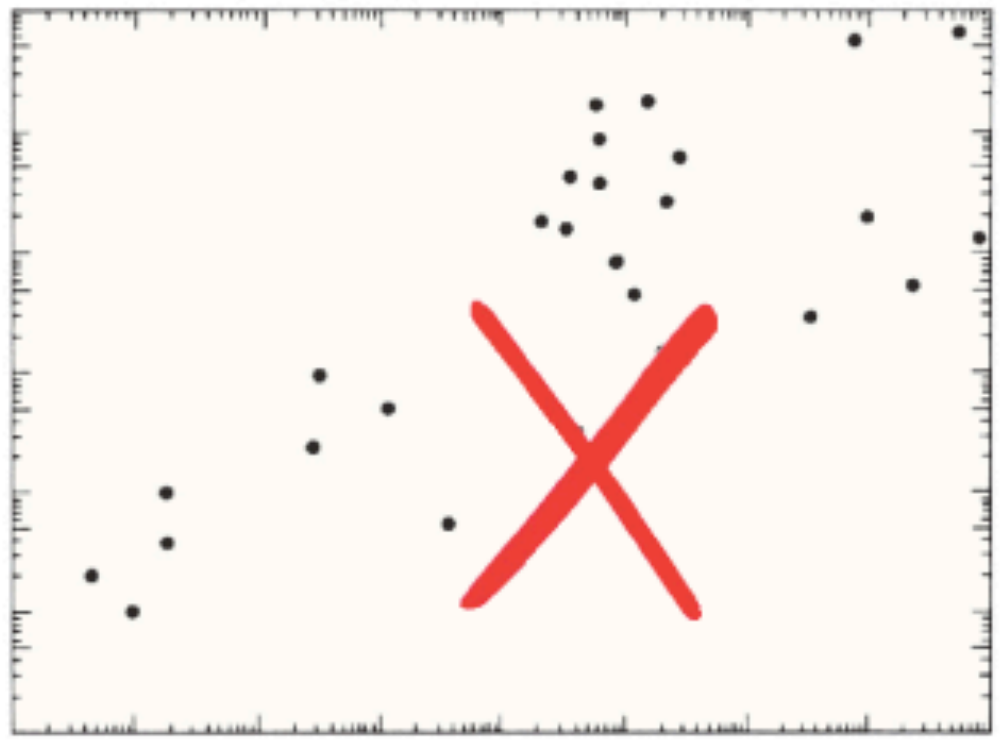
Haroz, et al. (2015)

Not all "visual embellishments" are "chart junk"!



1

2



3

Tufte, "Beautiful Evidence" (2006)

“Chart Junk”

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

... bias, reduce data-ink ratio, clutter, degrade trust

Take-away: it depends on your audience, task, and context...

“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)

Similar advice of William Cleveland

(*The Elements of Graphing Data,* 1985)

- CLEAR VISION: Make clear visualizations, and ensure that the data stands out.
- CLEAR UNDERSTANDING: Ensure that main points and conclusions are graphically clear and represented.
- SCALES: Pick appropriate axes and tick-mark scales, and ensure all the data is represented.
- GENERAL STRATEGY: Ensure all the data is represented. Design your visualizations carefully and allow time to proofread.

Tufte: graphical displays should...

1. Show the data
2. Avoid distorting what the data have to say
3. Encourage comparisons
4. Reveal the data at several levels of detail
5. Serve a reasonably clear purpose
6. Be closely integrated with the statistical and verbal descriptions

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

In-class exercise

Design Critique & Redesign

Visualization Critique Rules

NEUTRAL VOICE: Criticism must be expressed in a way that states the facts and presents them in the most balanced possible light. Inflammatory remarks or *ad hominem* attacks are unacceptable.

FACTS: Every statement must be backed up by facts that can be independently checked. Mere opinions or anecdotal evidence are not enough, unless accompanied by more reliable evidence that points in the same direction.

CLEAR GOAL: A critique must serve a goal. Simply criticizing a work for its shortcomings is not enough. The critic must state an alternative solution in a way that is clear and complete enough to provide the basis for further research.

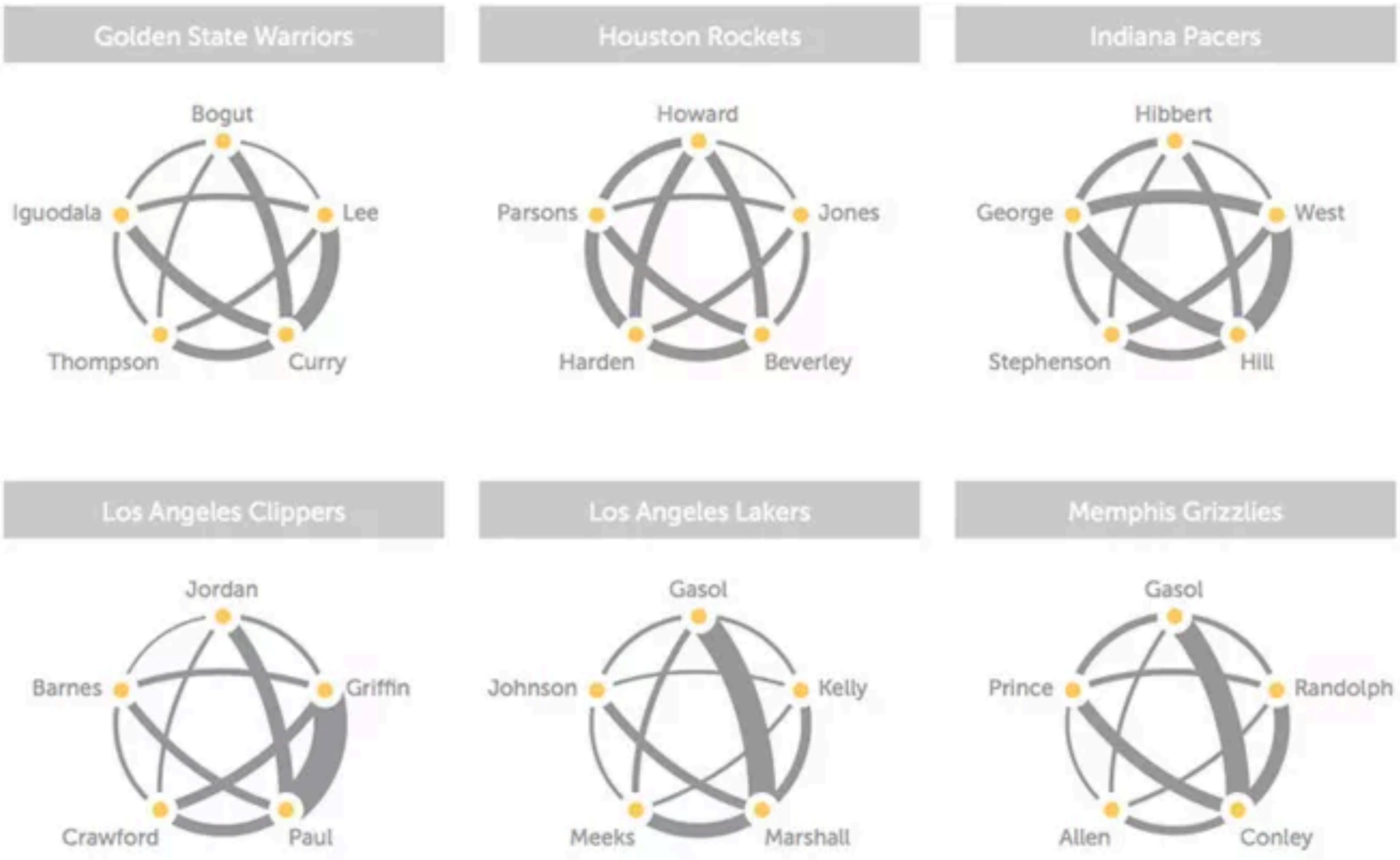
In-class exercise: Critique & Redesign

INSTRUCTIONS:

- Break-out into groups of ~5 people
- Discuss/critique visualization within your group
- Use the pens/paper to brainstorm redesigns of the visualization, and pick one to refine/sketch as a group
- Each group will have a representative present their sketch/idea

In-class exercise: Critique & Redesign

1. Who is the intended audience?
2. What information does this visualization represent?
3. How many data dimensions does it encode?
4. List several tasks, comparisons or evaluations it enables
5. What principles of excellence best describe why it is good / bad?
6. Can you suggest any improvements?
7. Why do you like / dislike this visualization?



http://www.theatlantic.com/past/docs/images/issues/200709/win.jpg

WHO IS LECTURING TODAY?

Prof. Michelle Borkin
West Village H, 310D
m.borkin@northeastern.edu

Research interests: visualization (information and scientific), human-computer interaction (HCI), medical imaging and radiology, astrophysics, and cognition and perception.

astronomy & physics ⟶ *applied physics* ⟶ *computer science* ⟶