

# Lecture :

# Basic Principles of Visualization Design

DATA ANALYSIS & VISUALIZATION  
FALL 2021

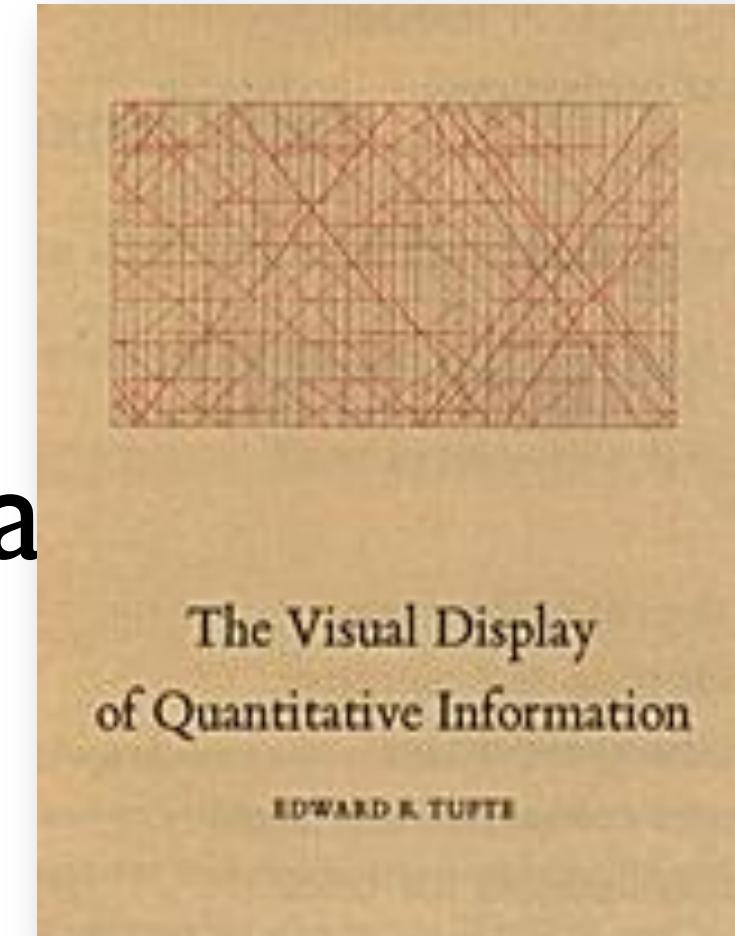
*Dr. Muhammad Faisal Cheema*  
FAST-NU

# GOALS FOR TODAY

- Learn basic “do’s and don’t’s of visualization design in order to be honest, have integrity, and be clear
- Learn Tufte’s “Graphical Integrity” principles
- *Be aware that there is a fuzzy gray area of interpretation and opinion on integrity*

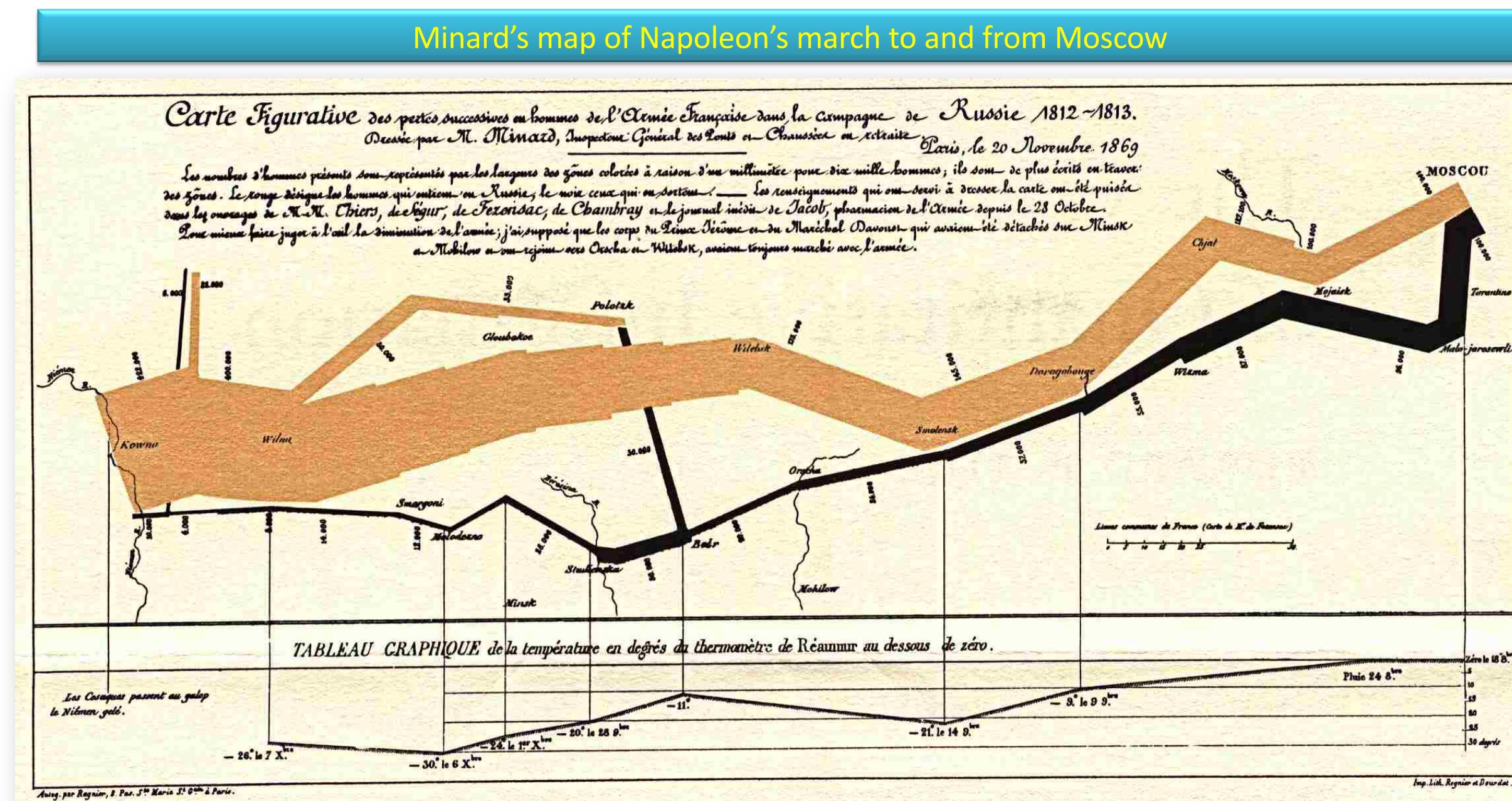
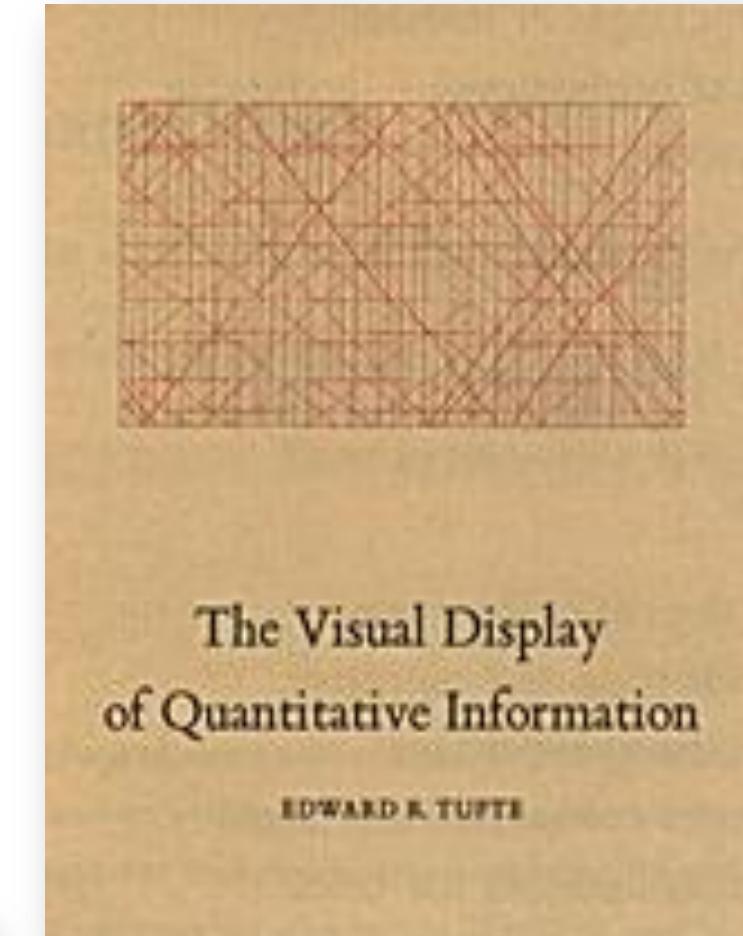
# Graphical Excellence

that which gives the viewer the greatest number of ideas  
in the shortest time  
with the least ink  
in the smallest space



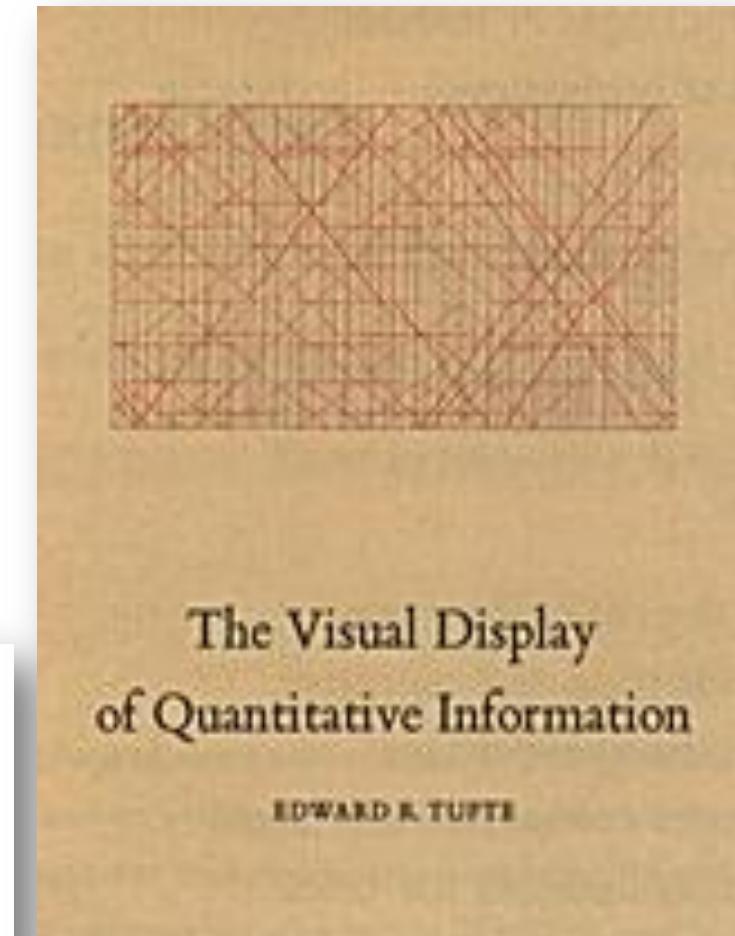
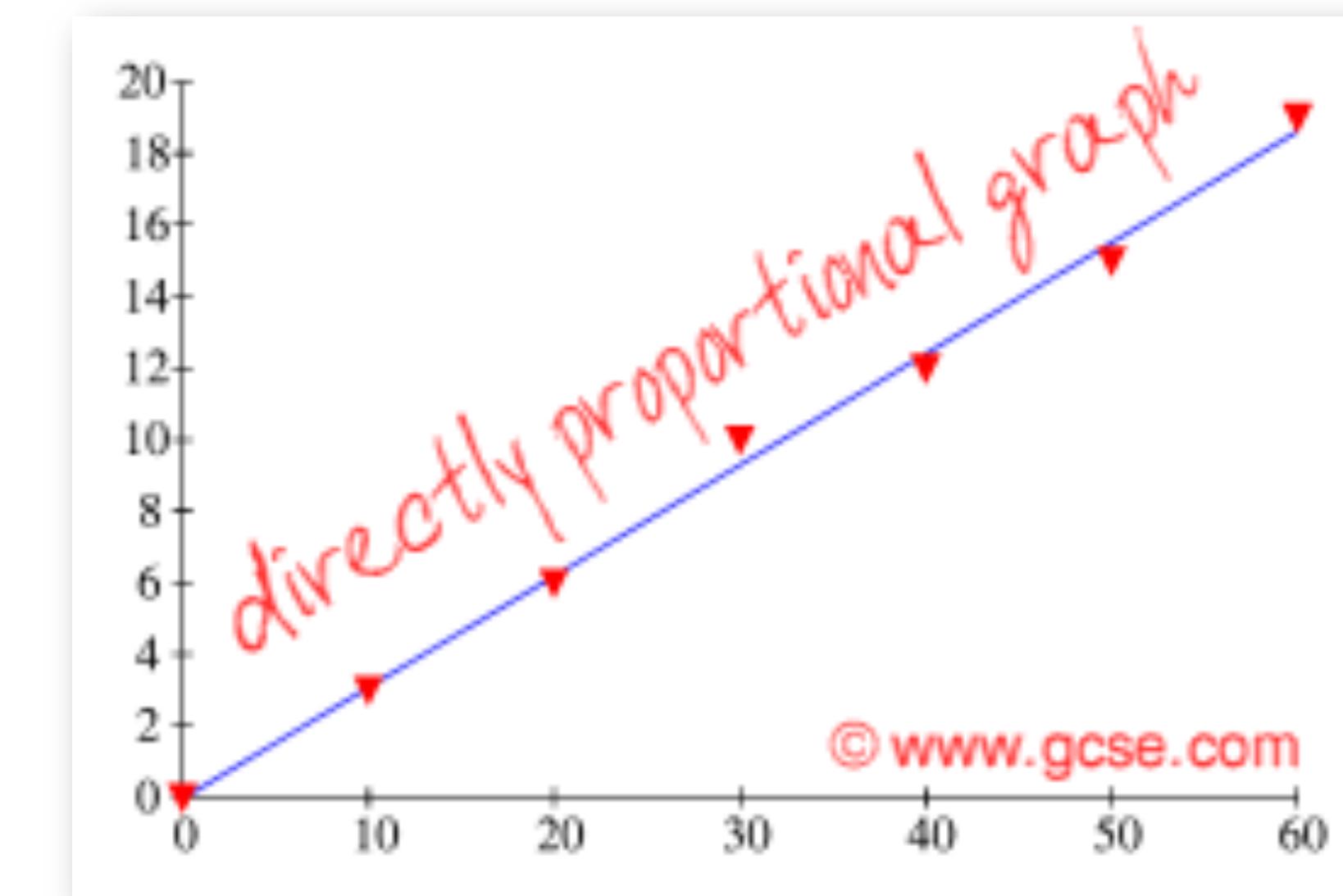
# Graphical Excellence

- that which gives the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space



# Graphical Integrity

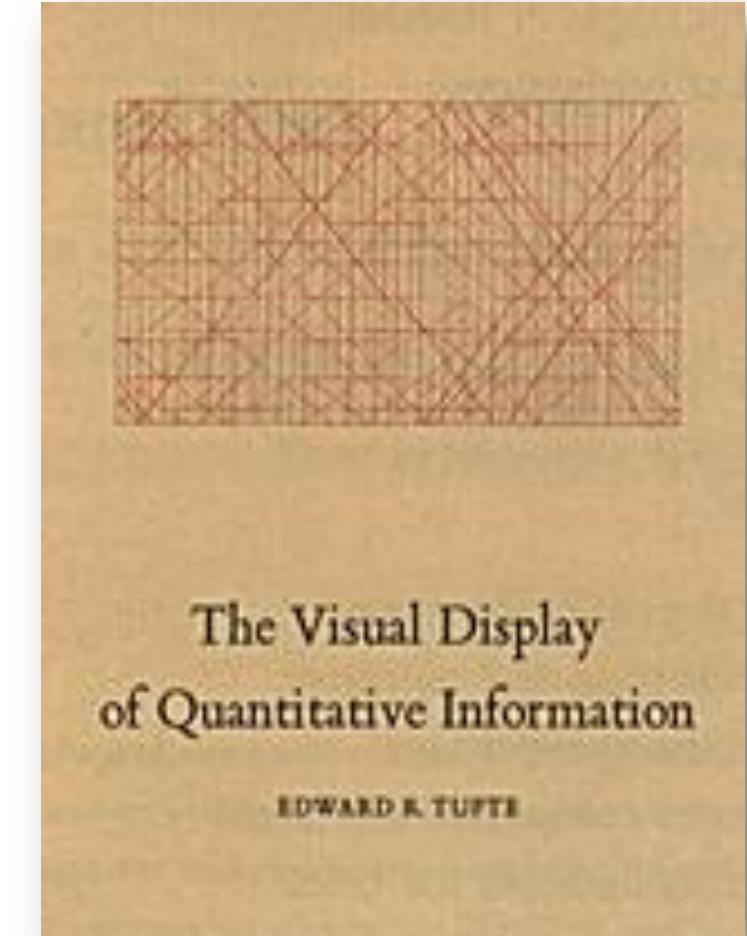
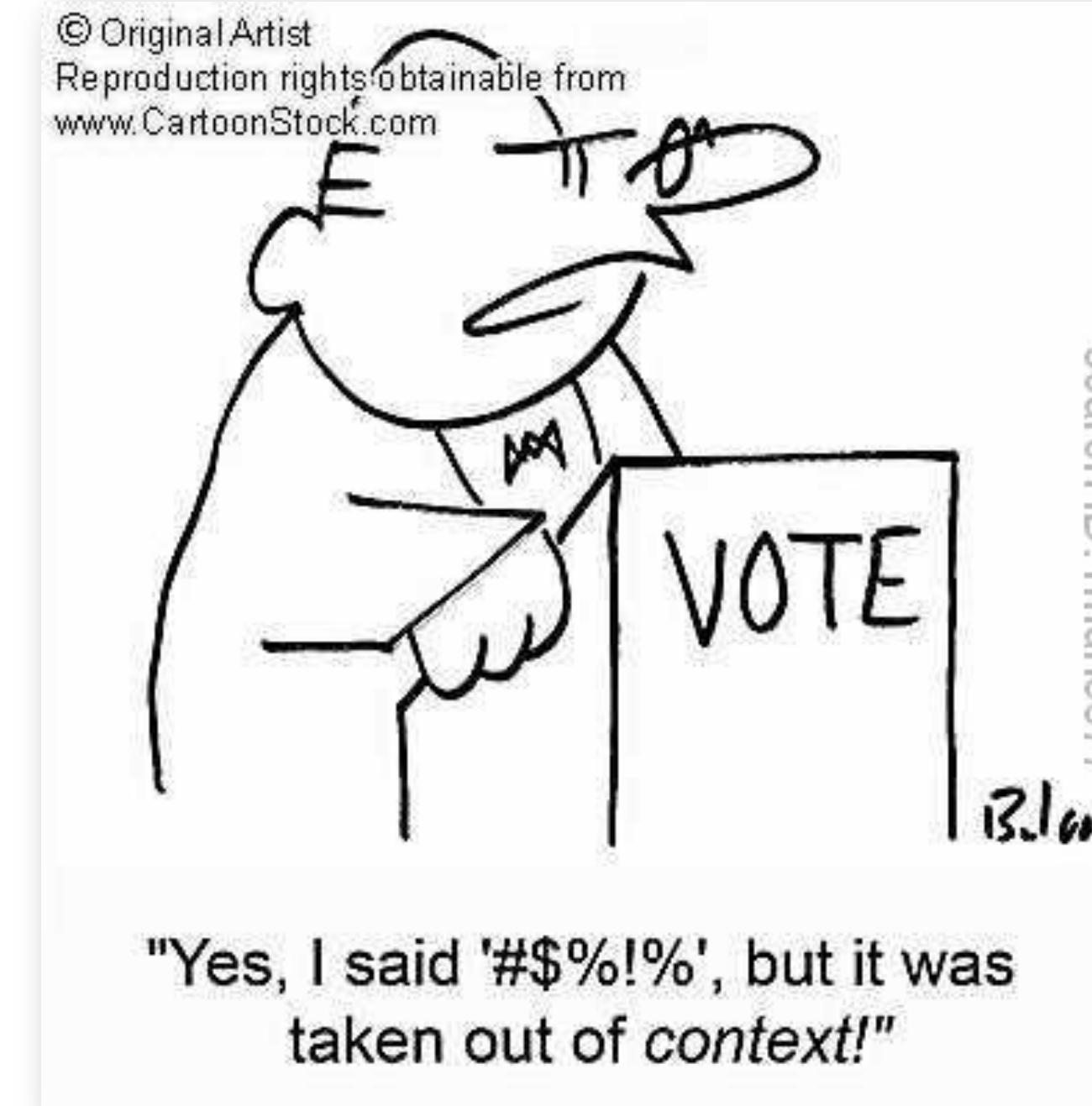
representation of numbers should be directly proportional to the numerical quantities represented



# Graphical Integrity

representation of numbers should be directly proportional to the numerical quantities represented

graphics must not quote out of context



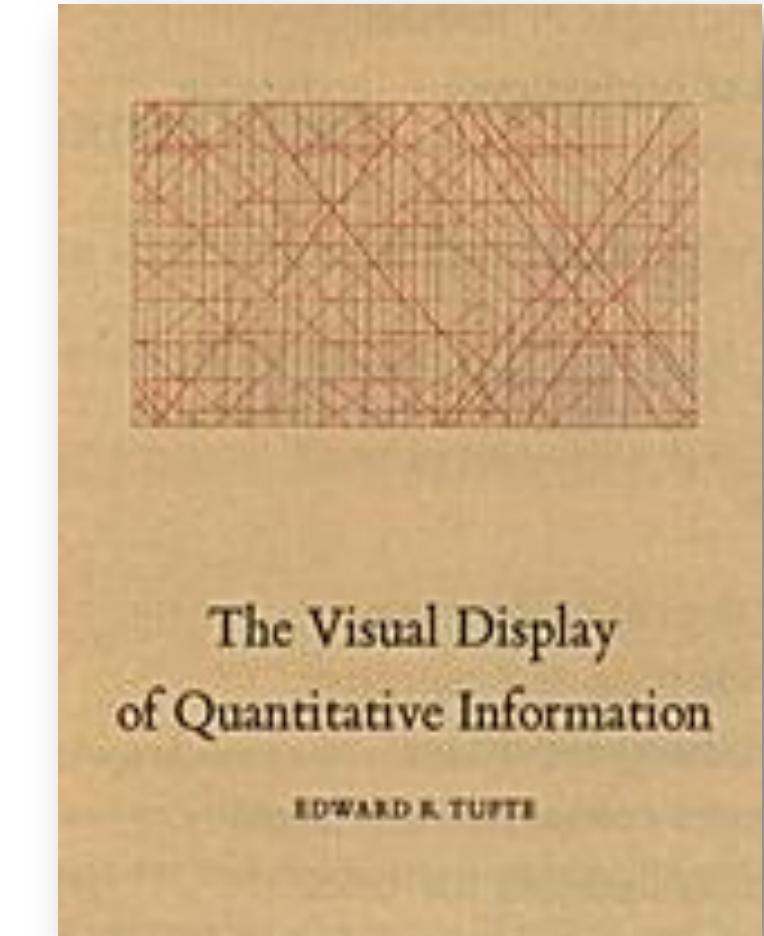
# Graphical Integrity

representation of numbers should be directly proportional to the numerical quantities represented

graphics must not quote out of context  
clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity

M Non-Laminated Tapes 26.2' long		
For use with PT-65, 70, 80, 85, 100, 110	3/8 "	1/2 "
BLACK ON CLEAR		M131
BLACK ON WHITE		M231
BLACK ON GOLD	M821	M831
BLACK ON SILVER	M921, M9213 (3 PC)	M931
BLACK ON GREEN	M721	M731
BLACK ON BLUE	M521	M531
BLACK ON PINK	ME21	ME31
RED ON WHITE		MK232
BLUE ON WHITE		MK233
BLACK ON YELLOW		MK631
M Tape Assortment Packs		
BLACK ON PINK/GREEN/SILVER‡	ME793	

\*1 tape of each color

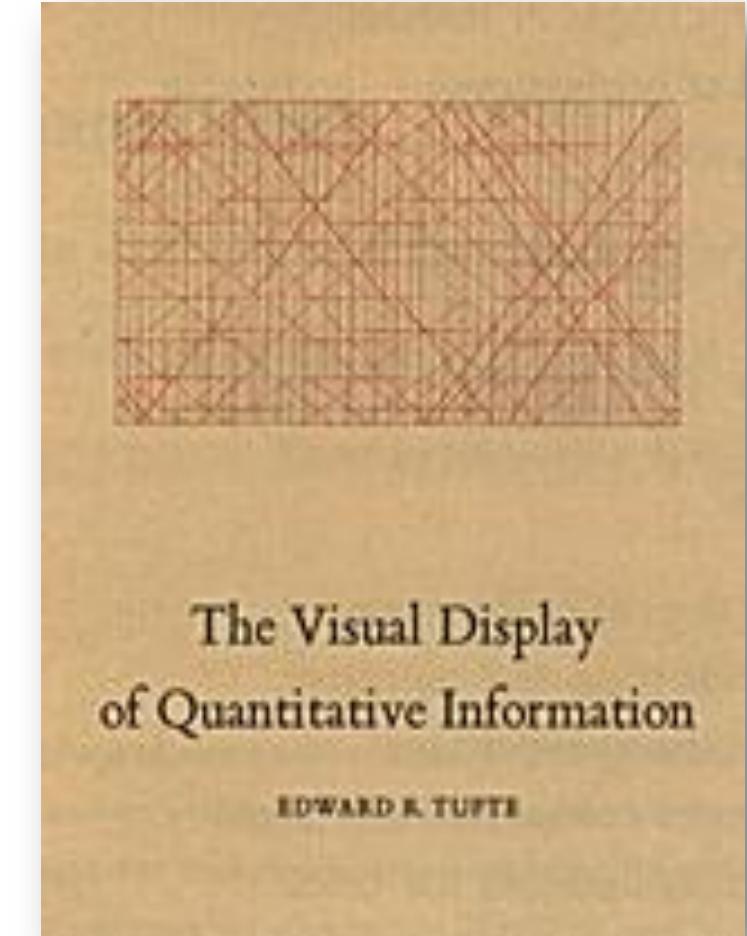
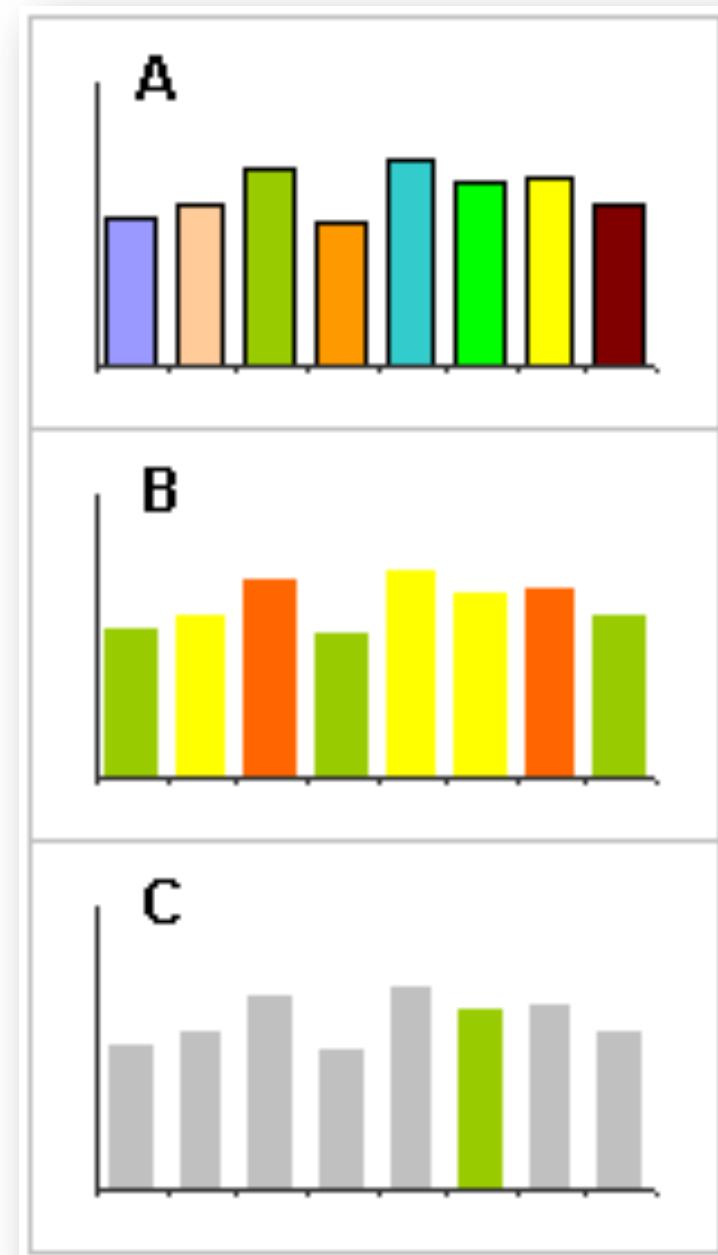


# Graphical Integrity

representation of numbers

should be directly proportional  
to the numerical quantities represented

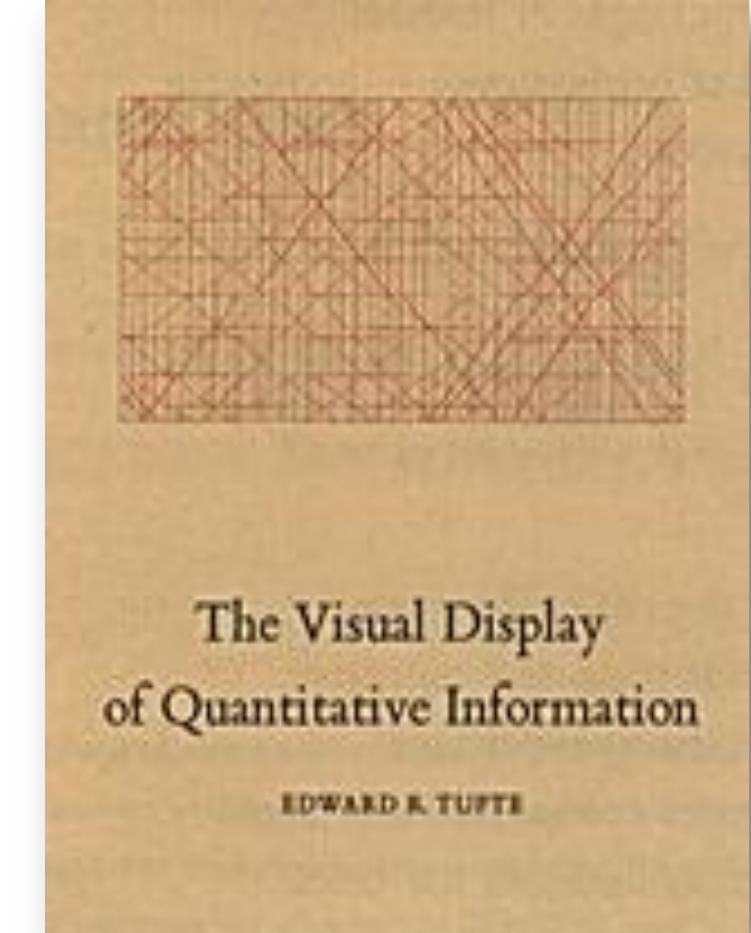
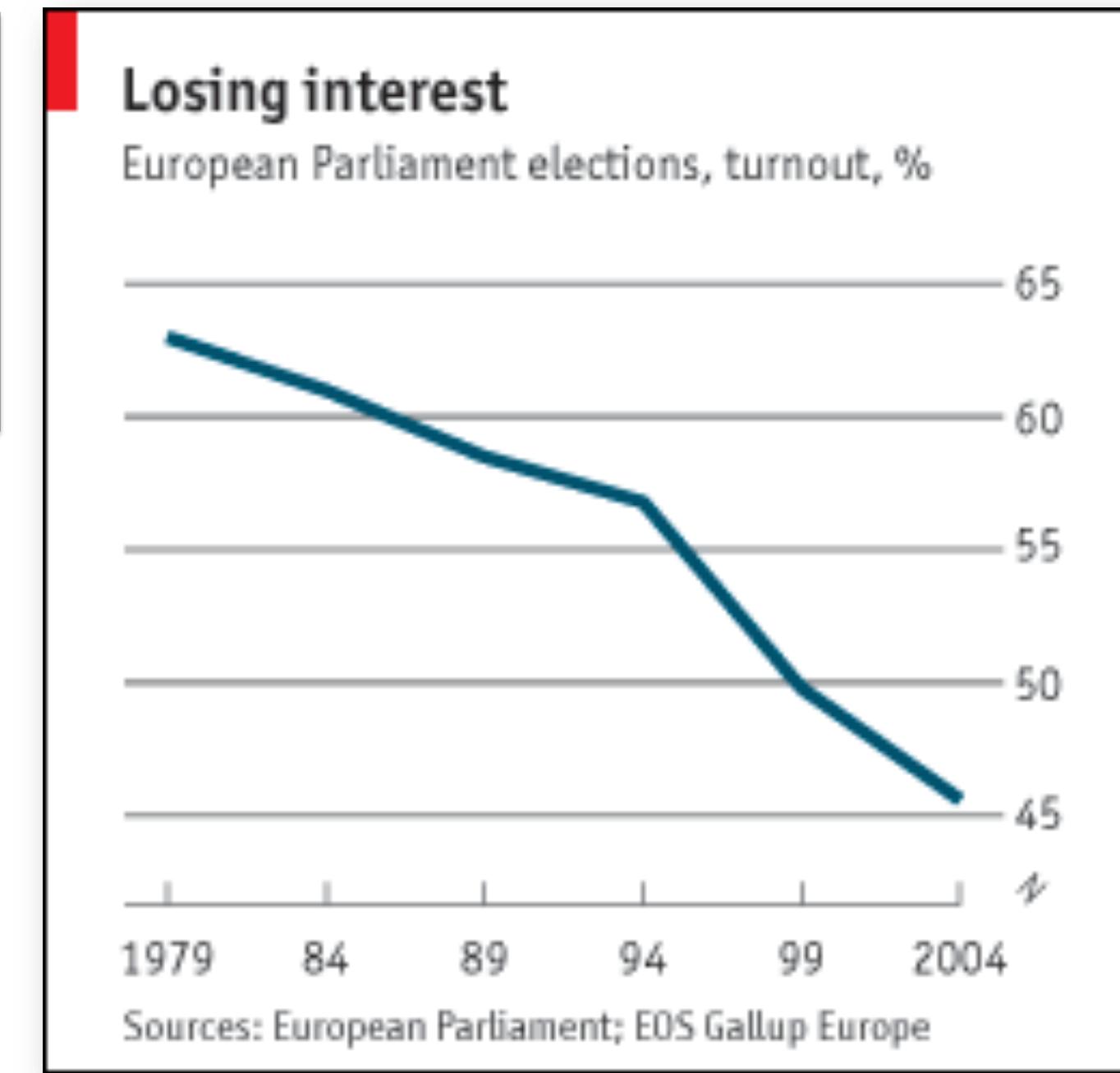
graphics must not quote out of context  
clear, detailed, and thorough labeling should be  
used to defeat graphical distortion and ambiguity  
**show data variation,  
not design variation**



# Graphical Integrity

representation of numbers should be directly proportional to the numerical quantities represented

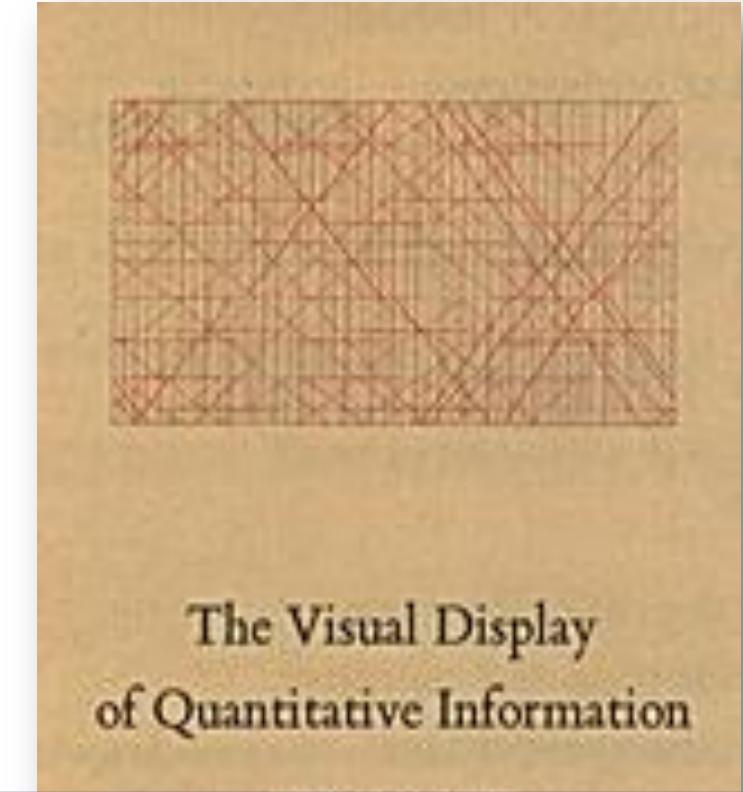
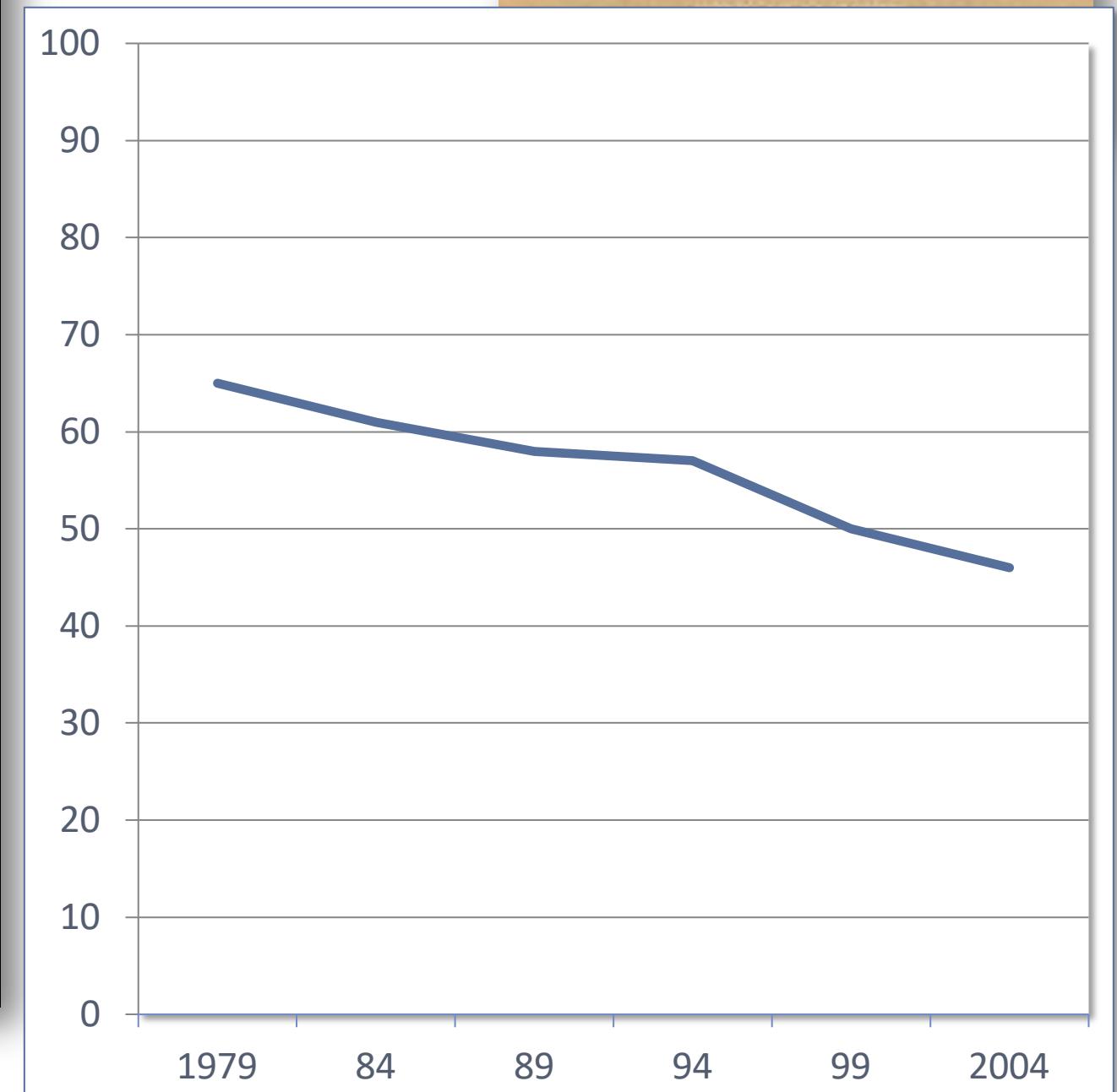
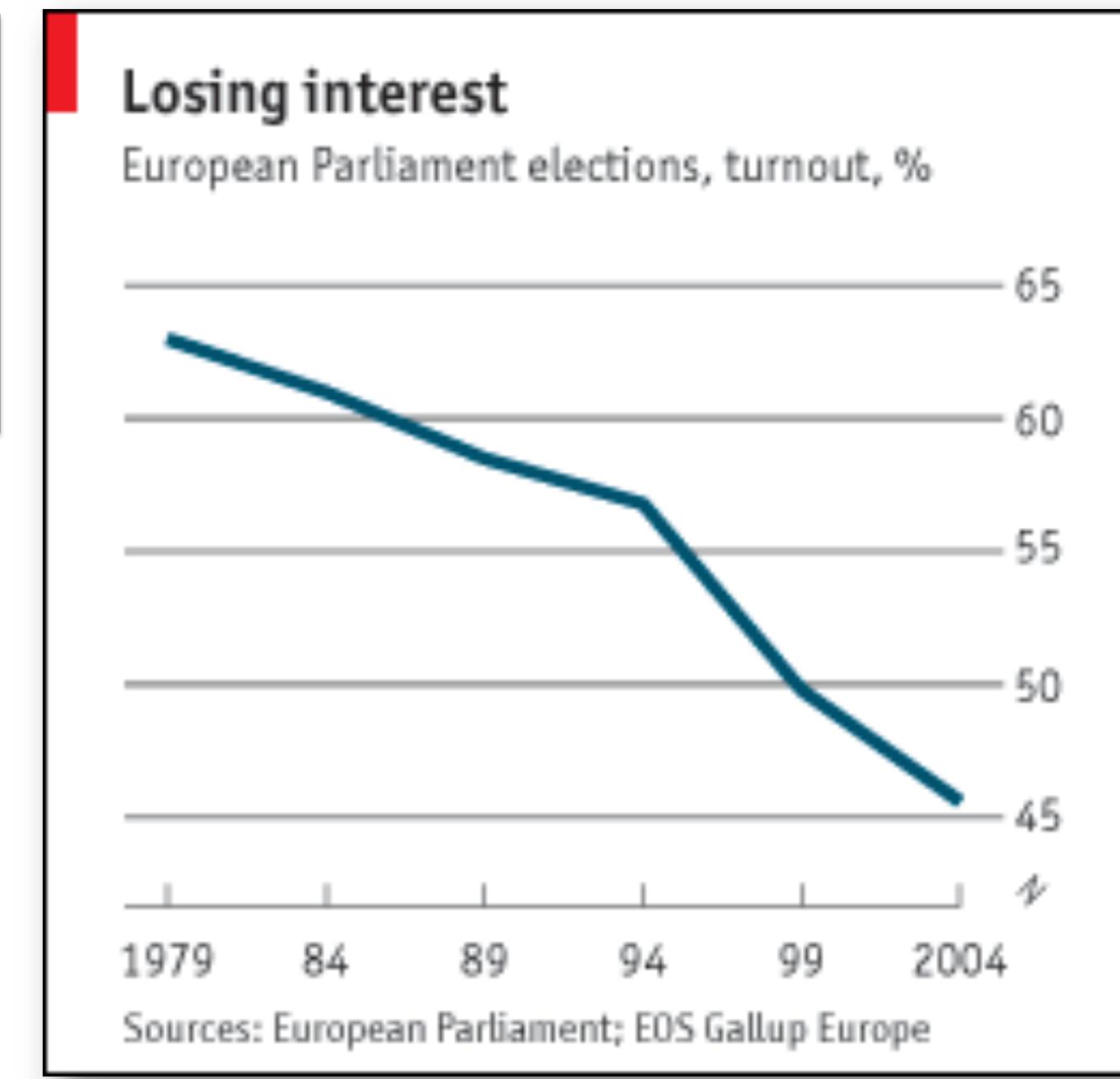
graphics must not quote out of context  
 clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity  
**show data variation, not design variation**



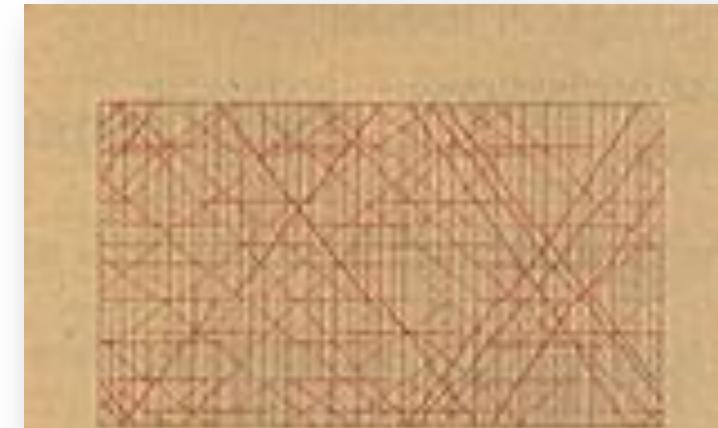
# Graphical Integrity

graphics must not quote out of context  
 clear, detailed, and thorough labeling should  
 be used to defeat graphical distortion and  
 ambiguity  
**show data variation,  
 not design variation**

representation of numbers should be directly proportional  
 to the numerical quantities represented



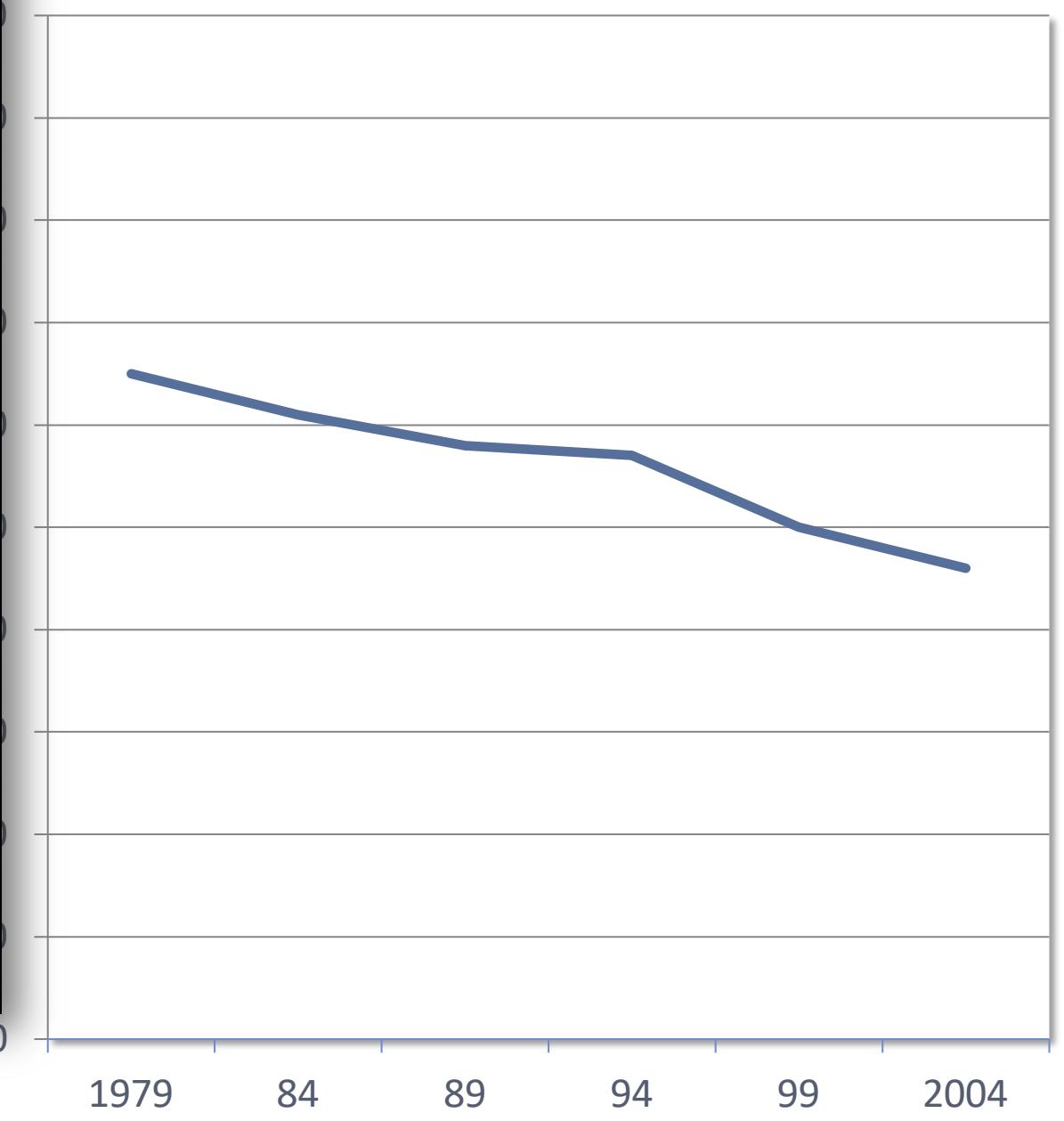
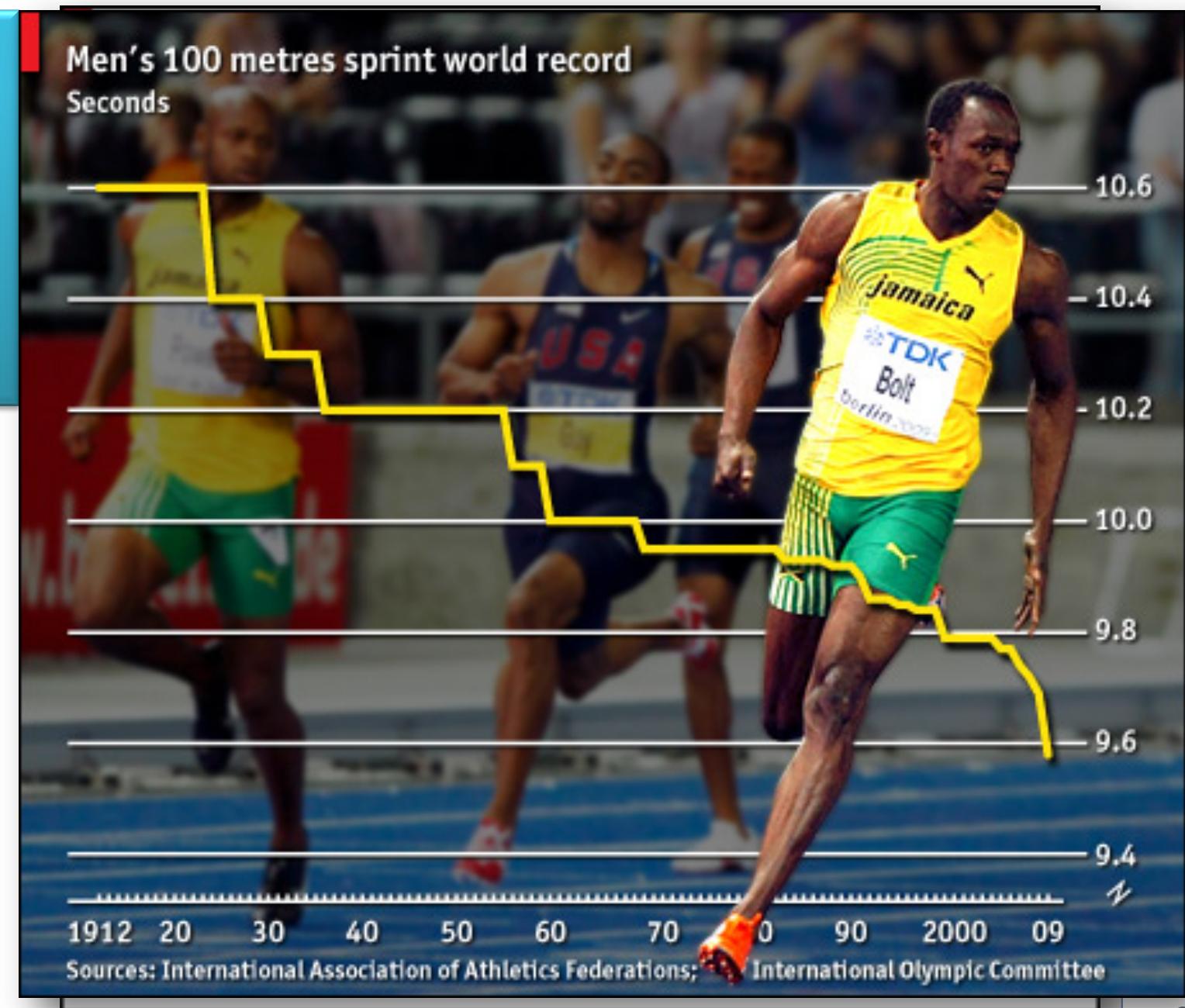
# Graphical Integrity



graphics must not quote out of context  
clear, detailed, and thorough labeling should  
be used to defeat graphical distortion and  
ambiguity

**show data variation,  
not design variation**

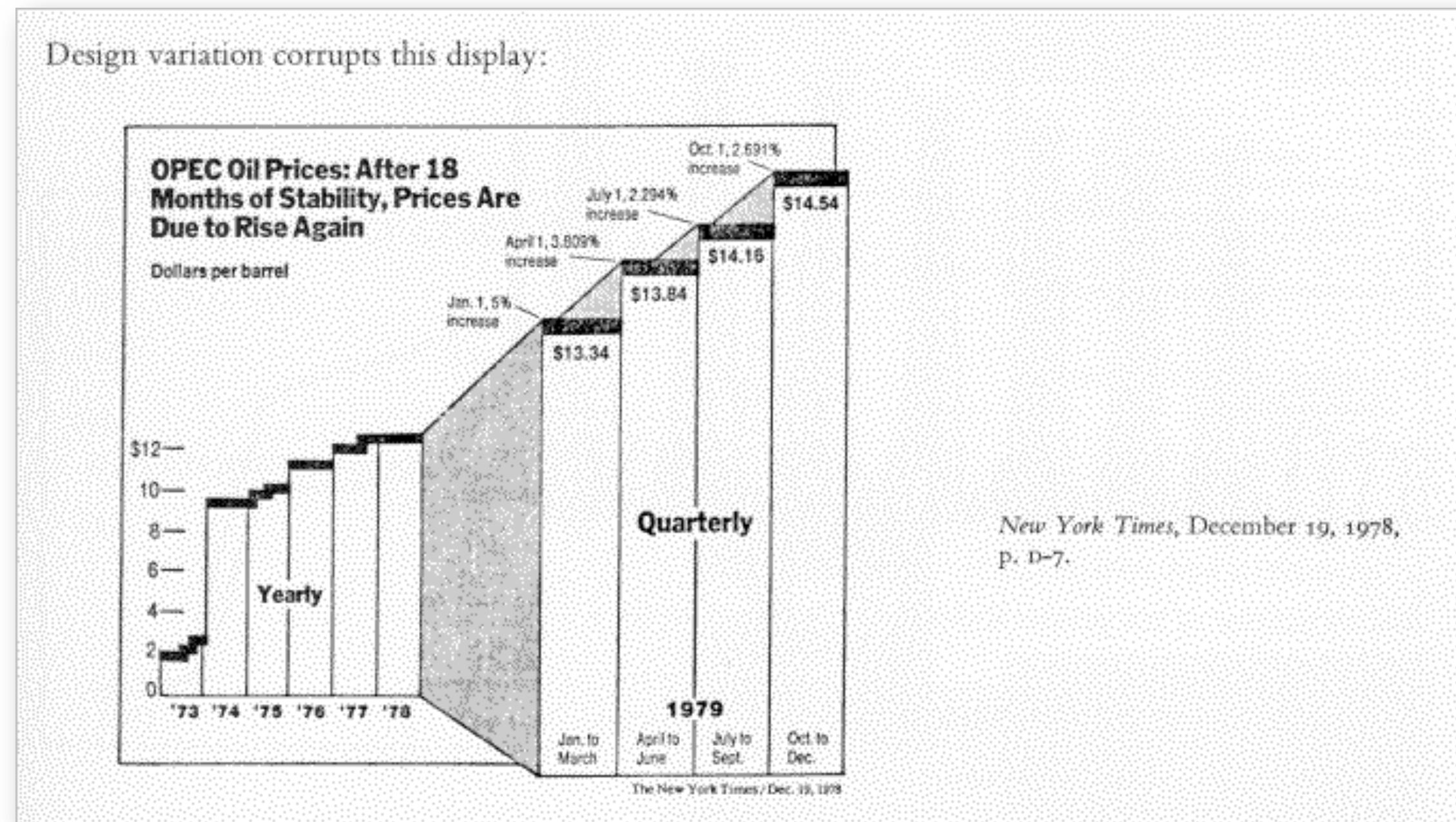
representation of numbers should be directly proportional  
to the numerical quantities represented



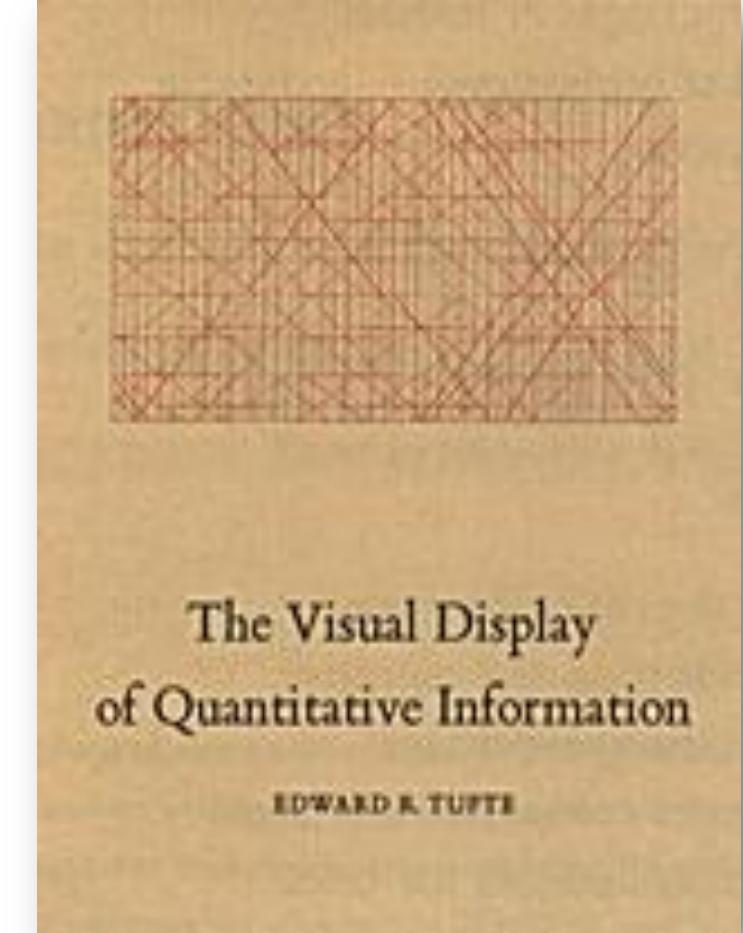
The Visual Display  
of Quantitative Information

# Graphical Integrity

show data variation, not design variation



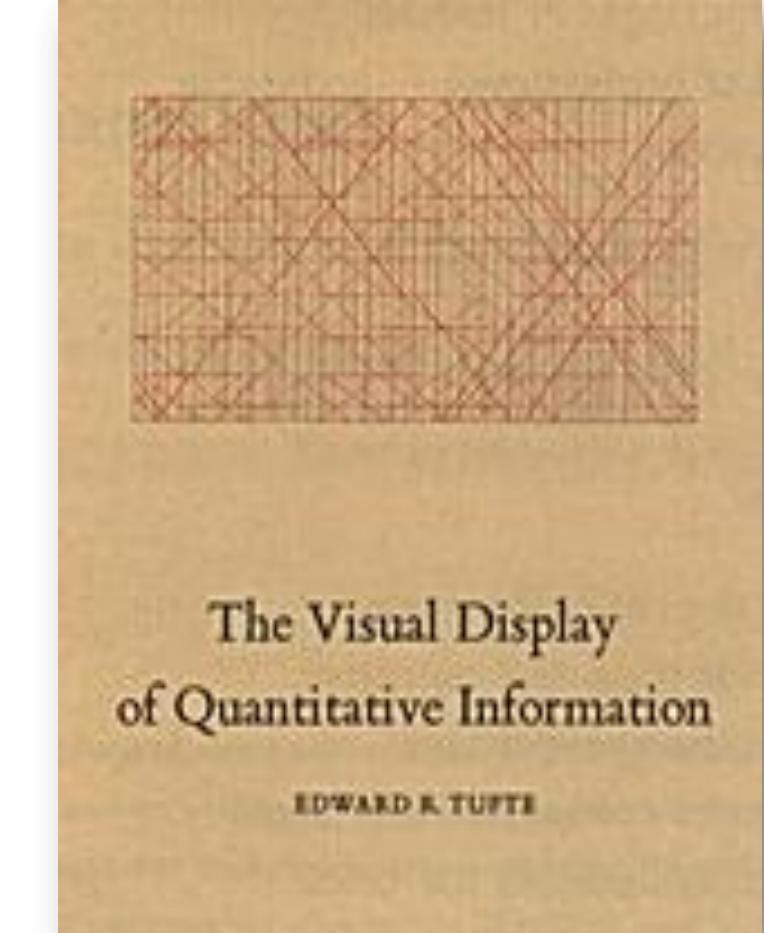
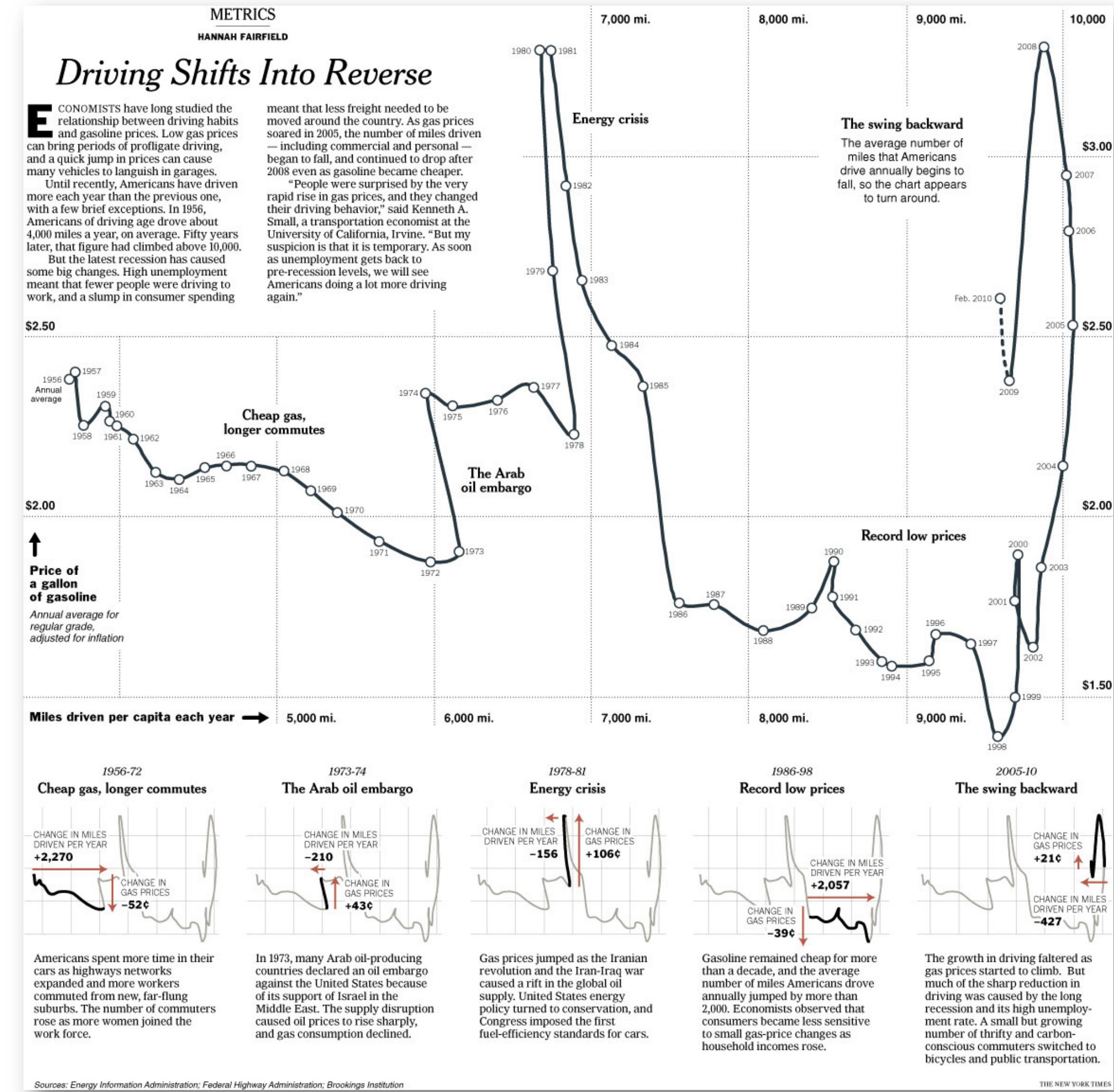
New York Times, December 19, 1978,  
p. D-7.



New York Times  
19 Dec 1978

# Graphical Integrity

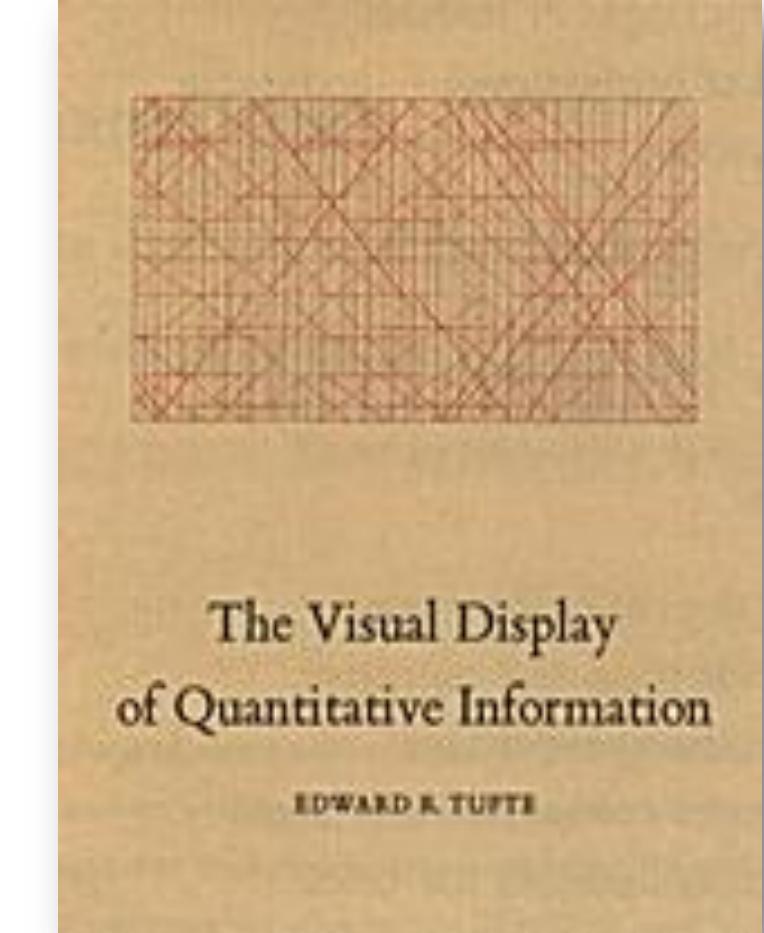
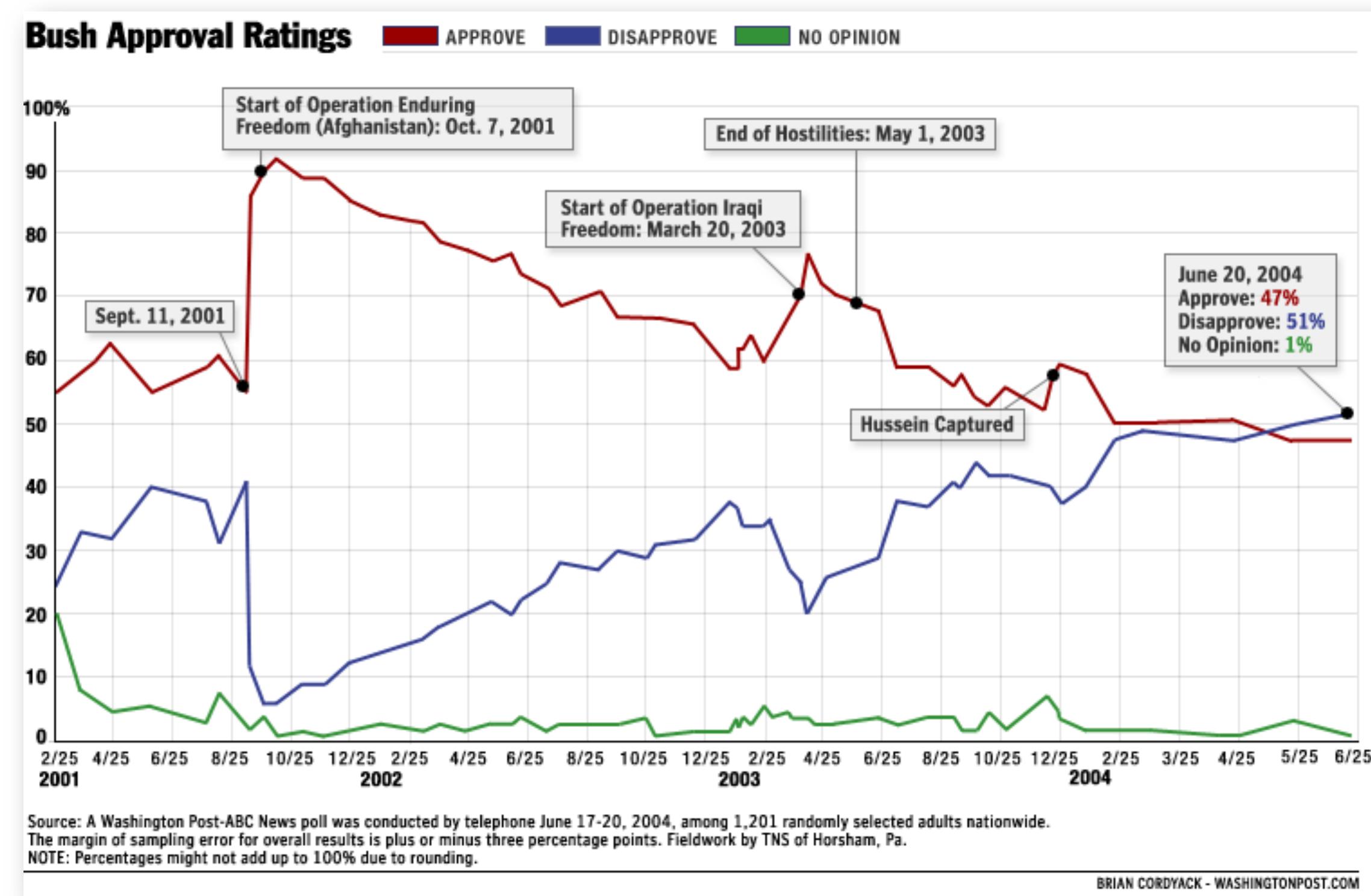
clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity



New York Times  
02 May 2010

# Graphical Integrity

clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity



Washington Post

# Some GRAPHICAL INTEGRITY Principles in detail

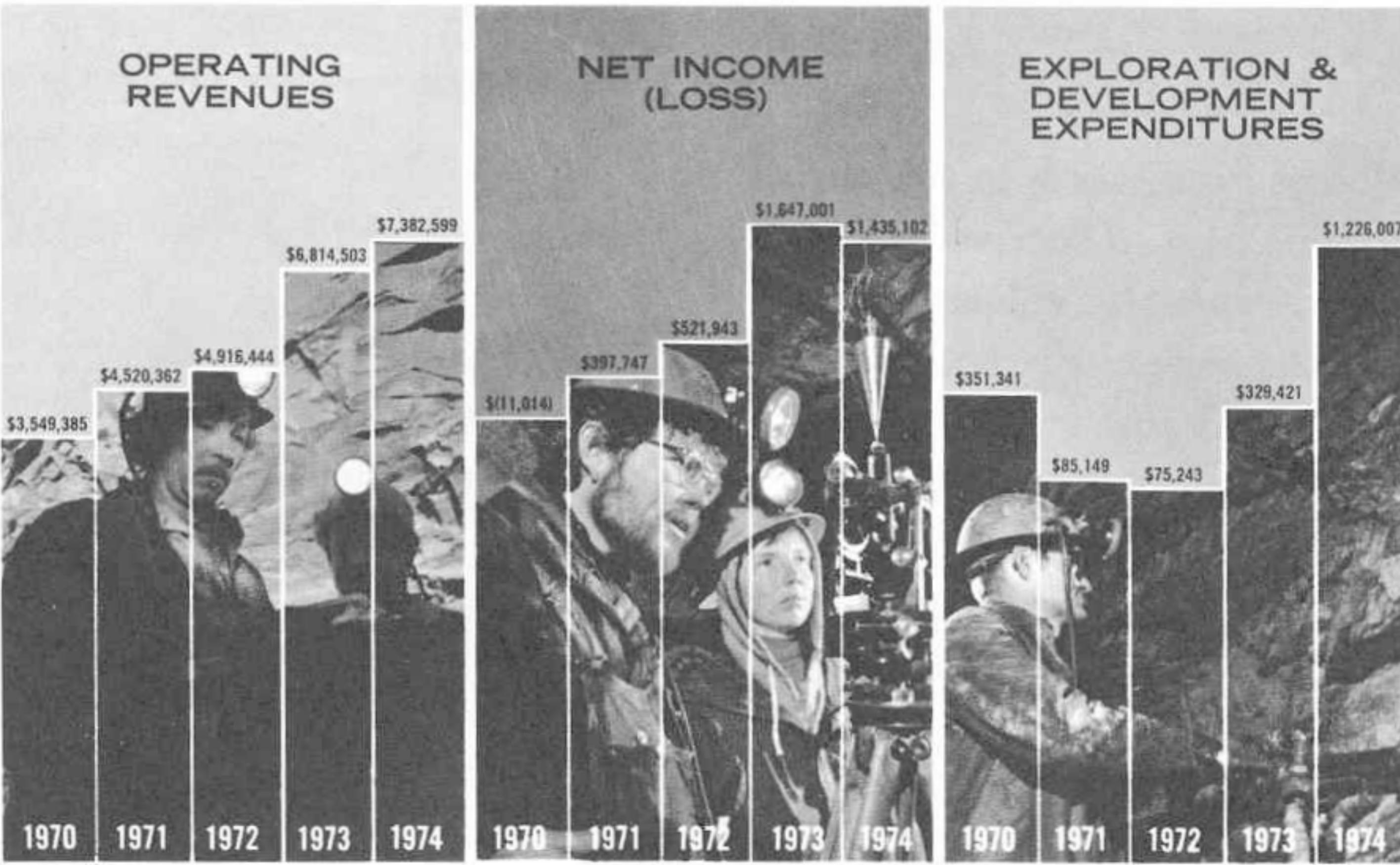
# “Graphical Integrity”

“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.”

# “Graphical Integrity”

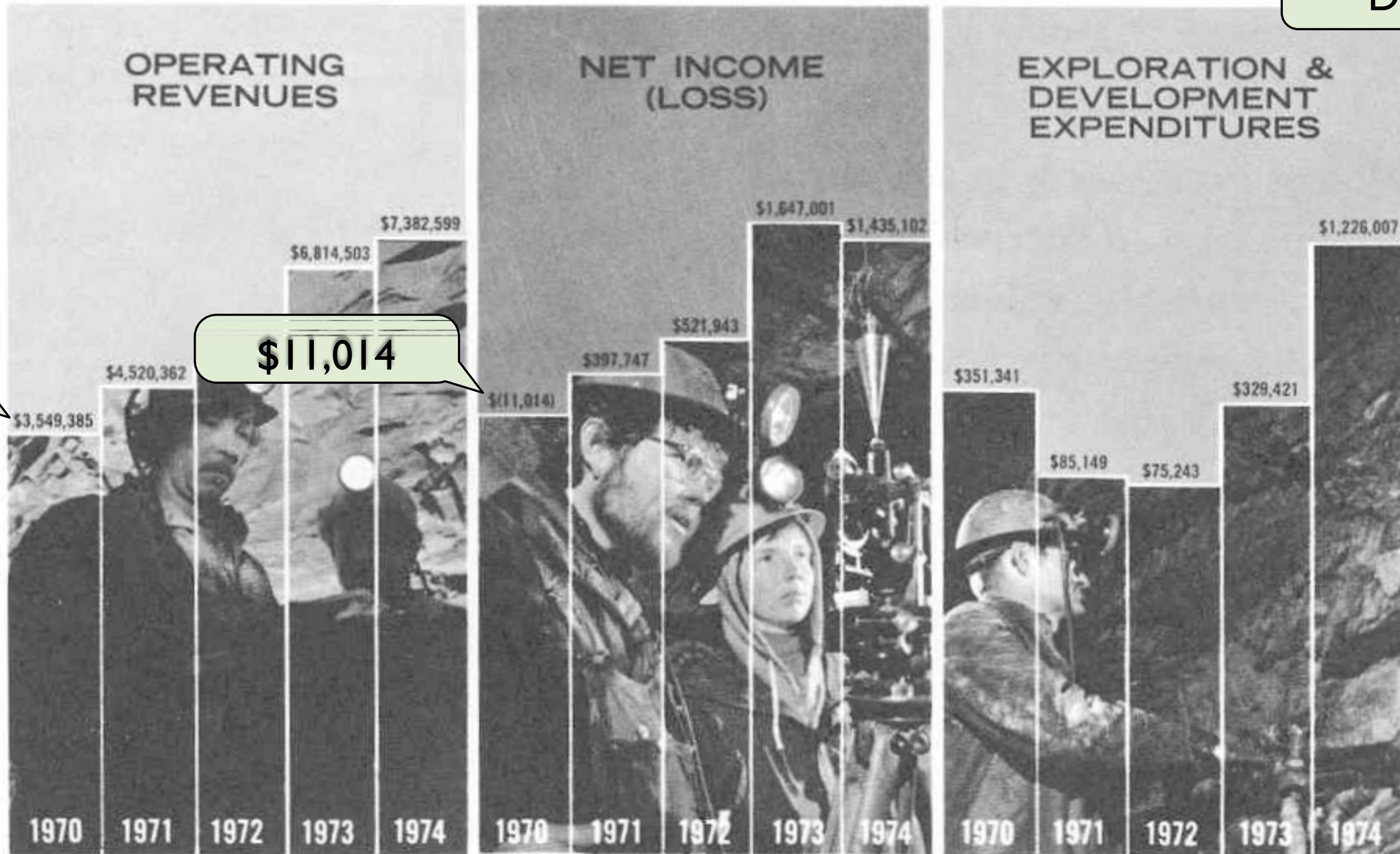
“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

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



“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)

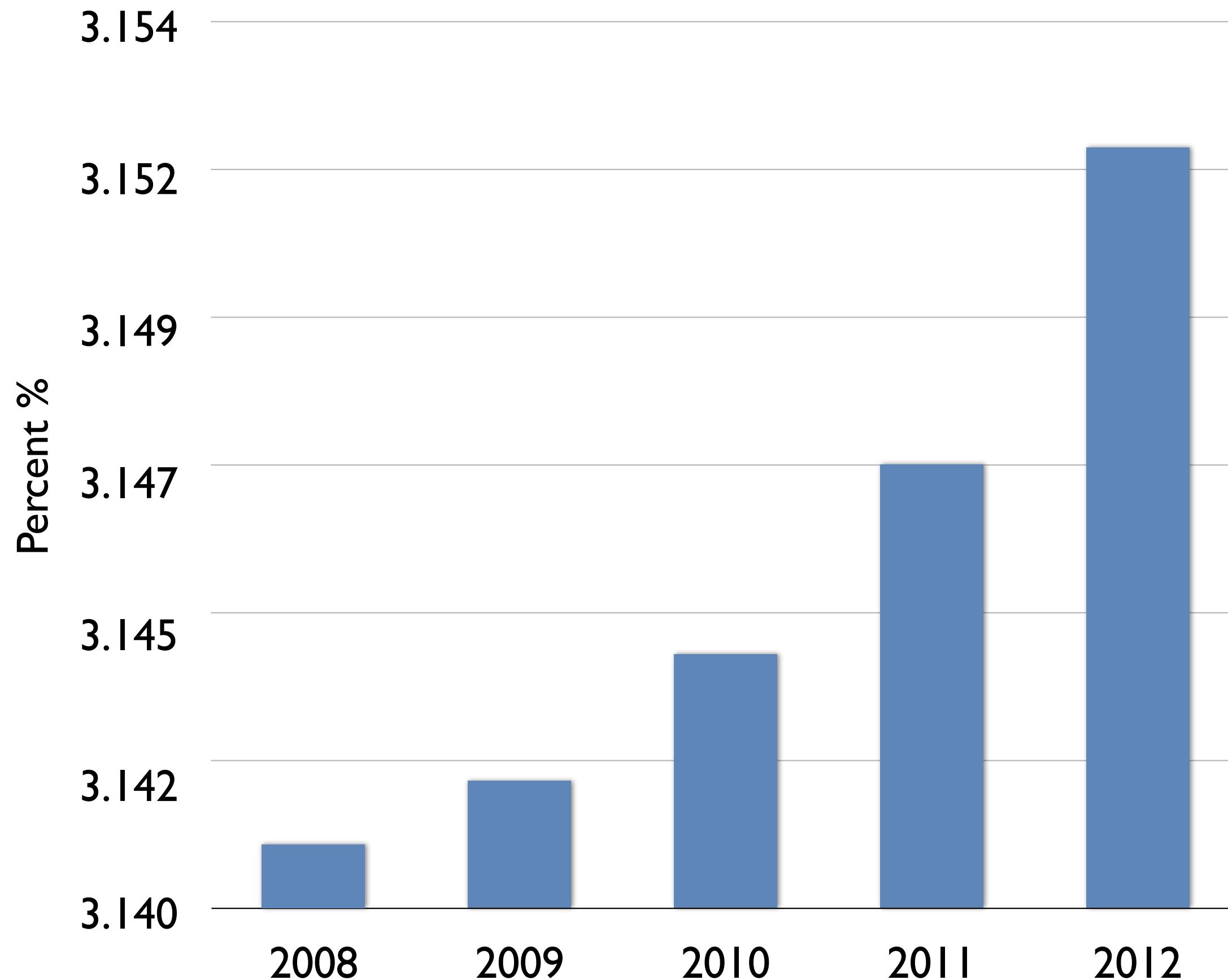


y-axis  
baseline?!

“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)

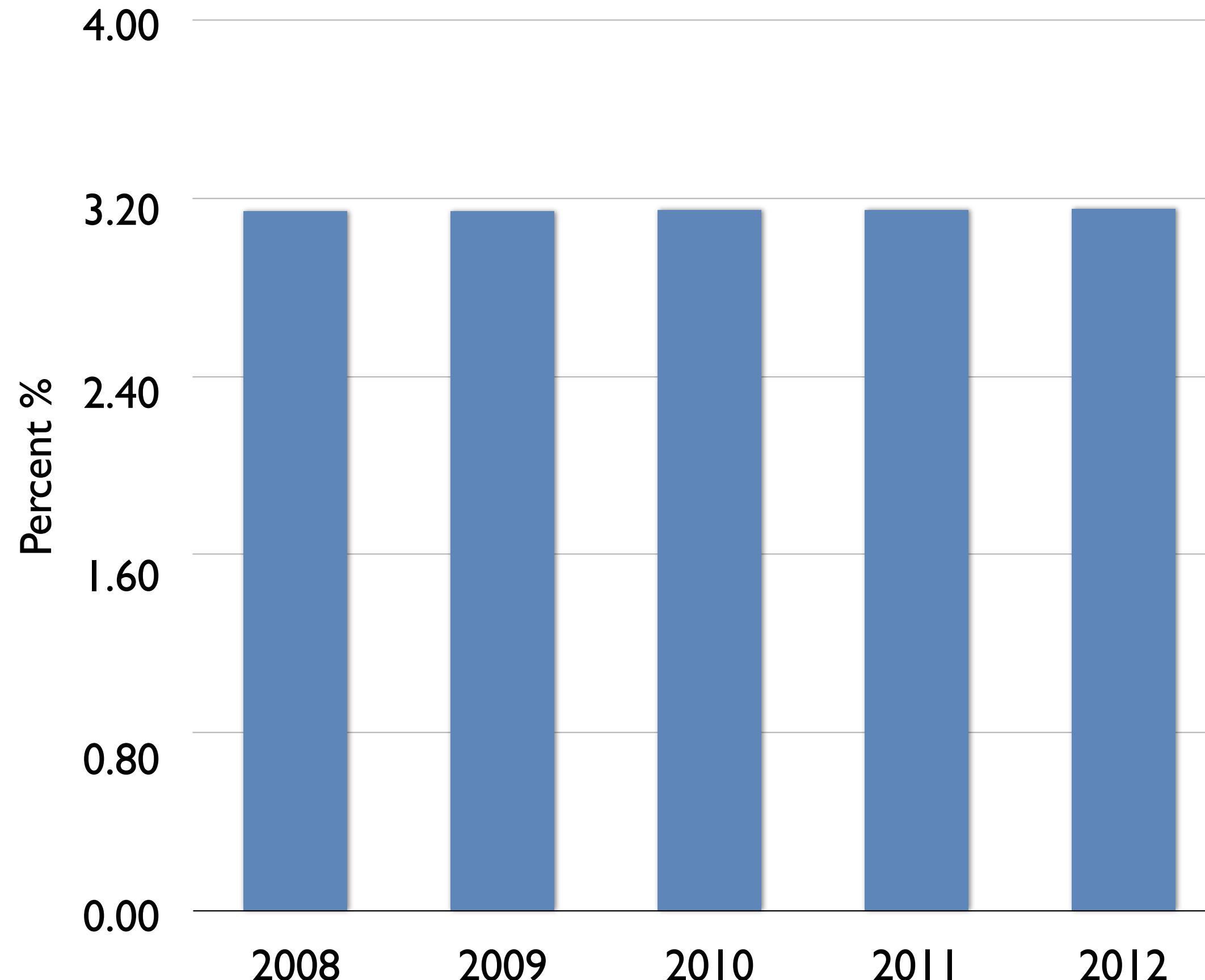
# Interest Rates



“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>

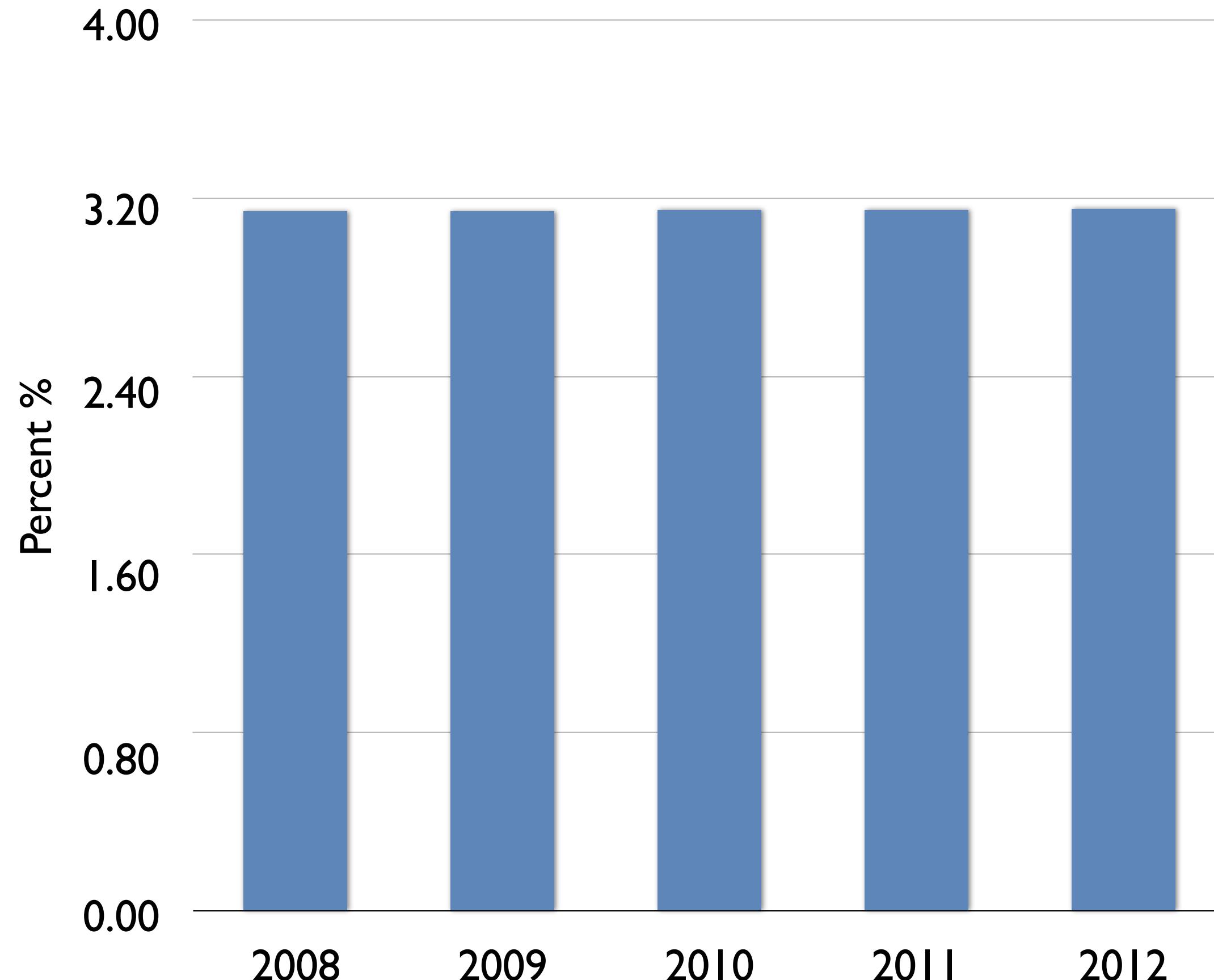
## Interest Rates



“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>

# Interest Rates

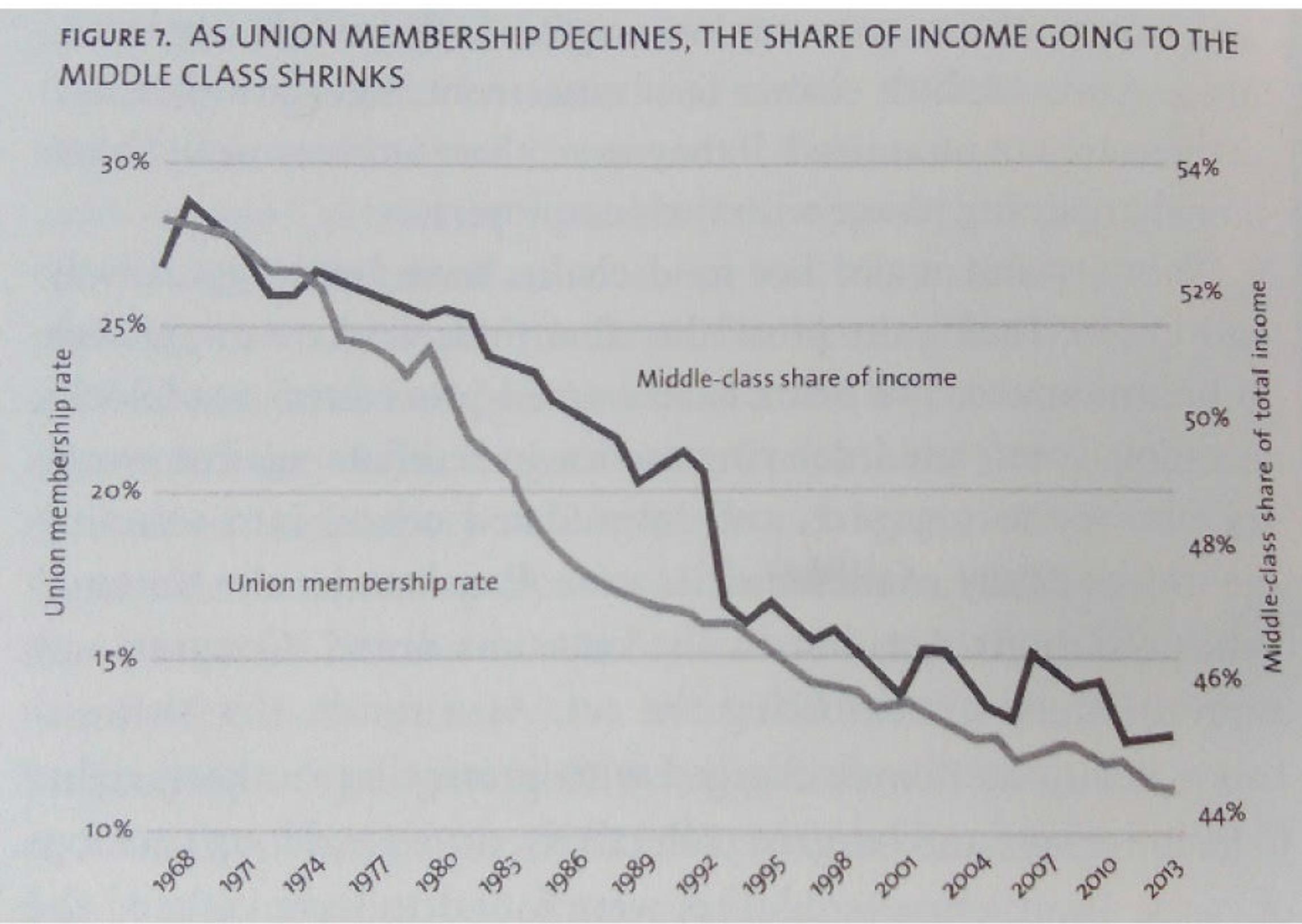


CONTEXT!

“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.”

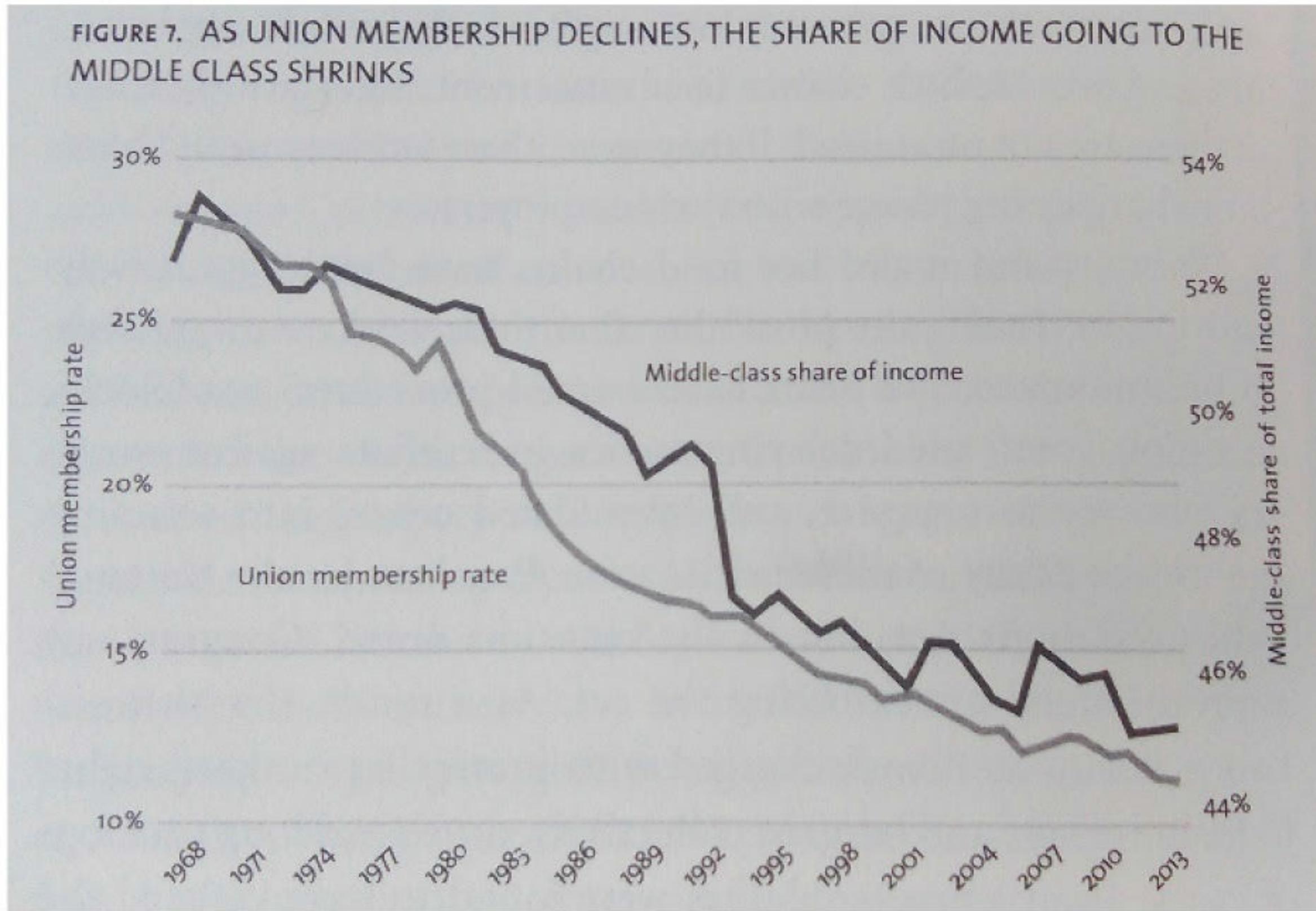
Based on <http://data.heapanalytics.com/how-to-lie-with-data-visualization>

FIGURE 7. AS UNION MEMBERSHIP DECLINES, THE SHARE OF INCOME GOING TO THE MIDDLE CLASS SHRINKS

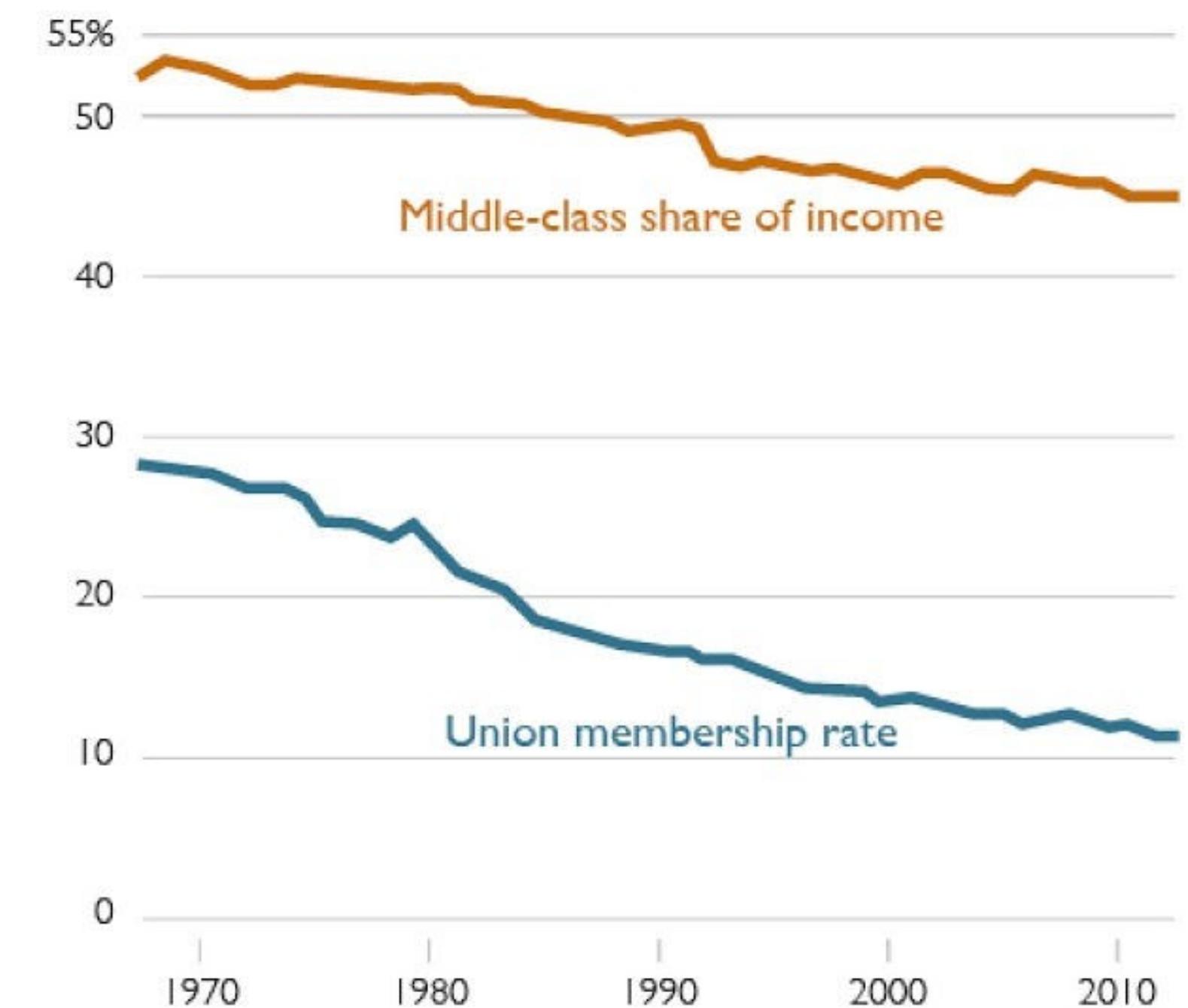


“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.” <http://www.thefunctionalart.com/2015/10/double-axes-double-mischief.html>

# ***“Double the axes, double the mischief”***



NEW VERSION



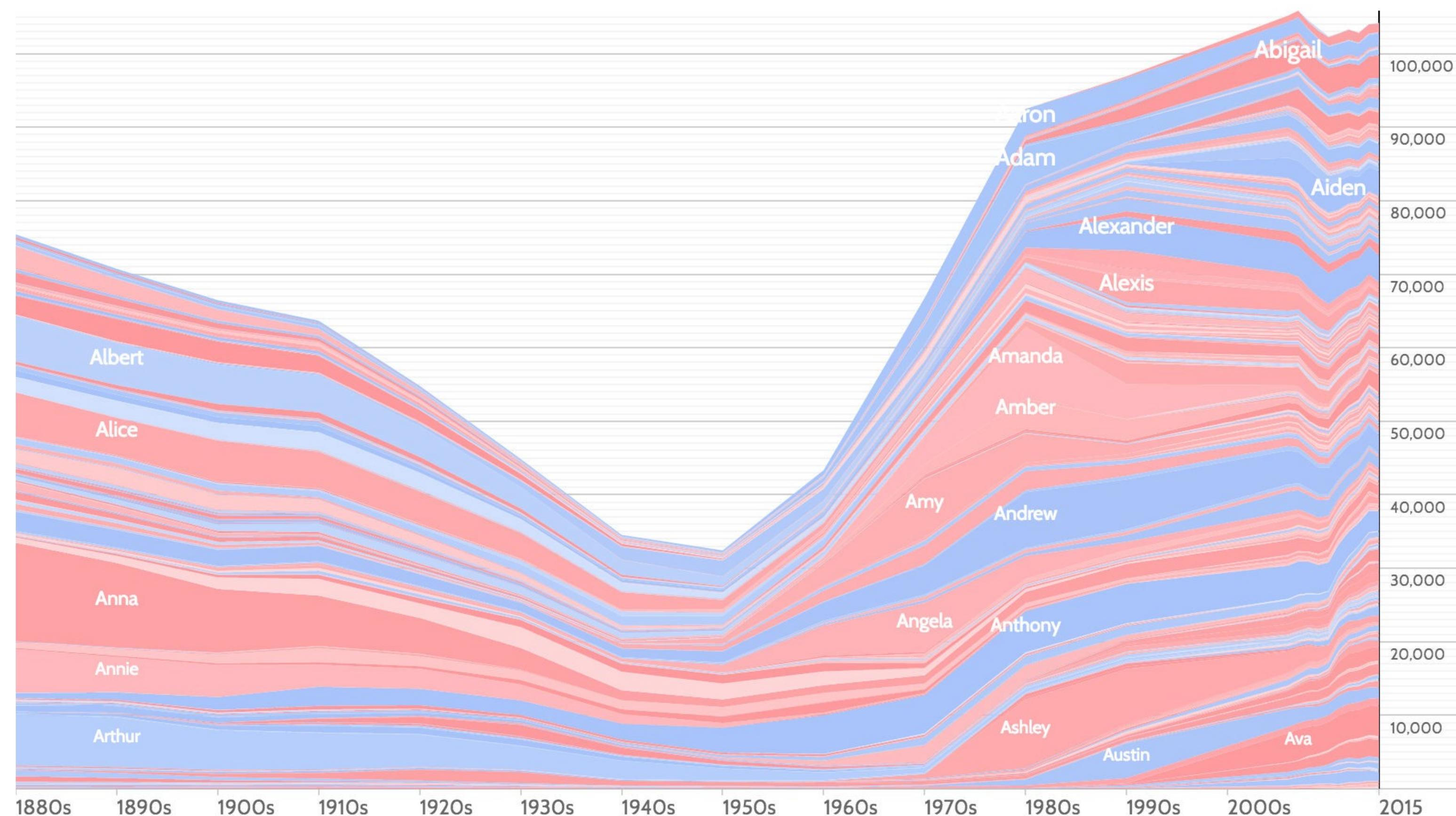
“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.” <http://www.thefunctionalart.com/2015/10/double-axes-double-mischief.html>

Baby Name > A  Both  Boys  Girls

boys	1000	500	100	25	1
girls	1000	500	100	25	1

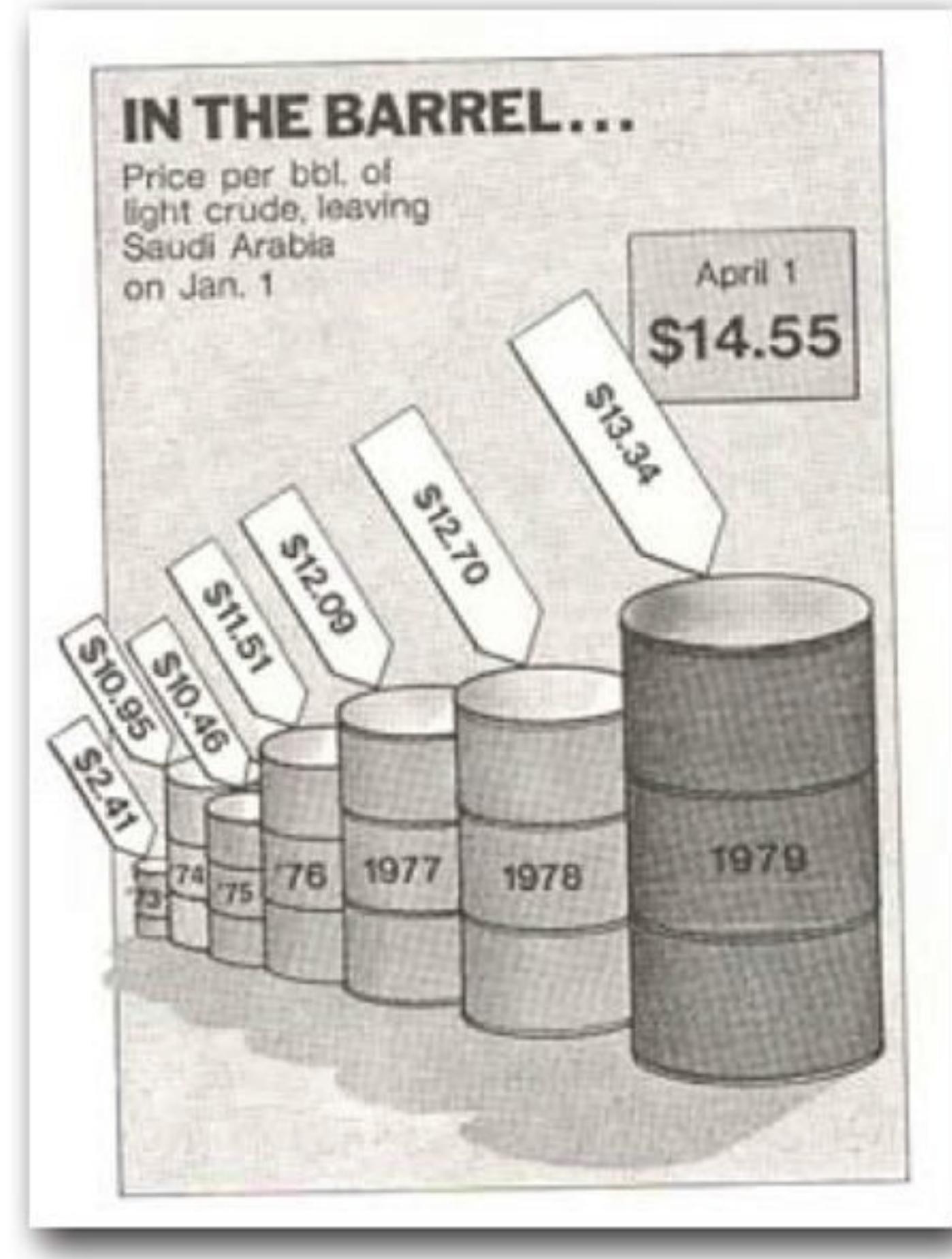
Current rank:

Names starting with 'A' per million babies



# “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.”



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

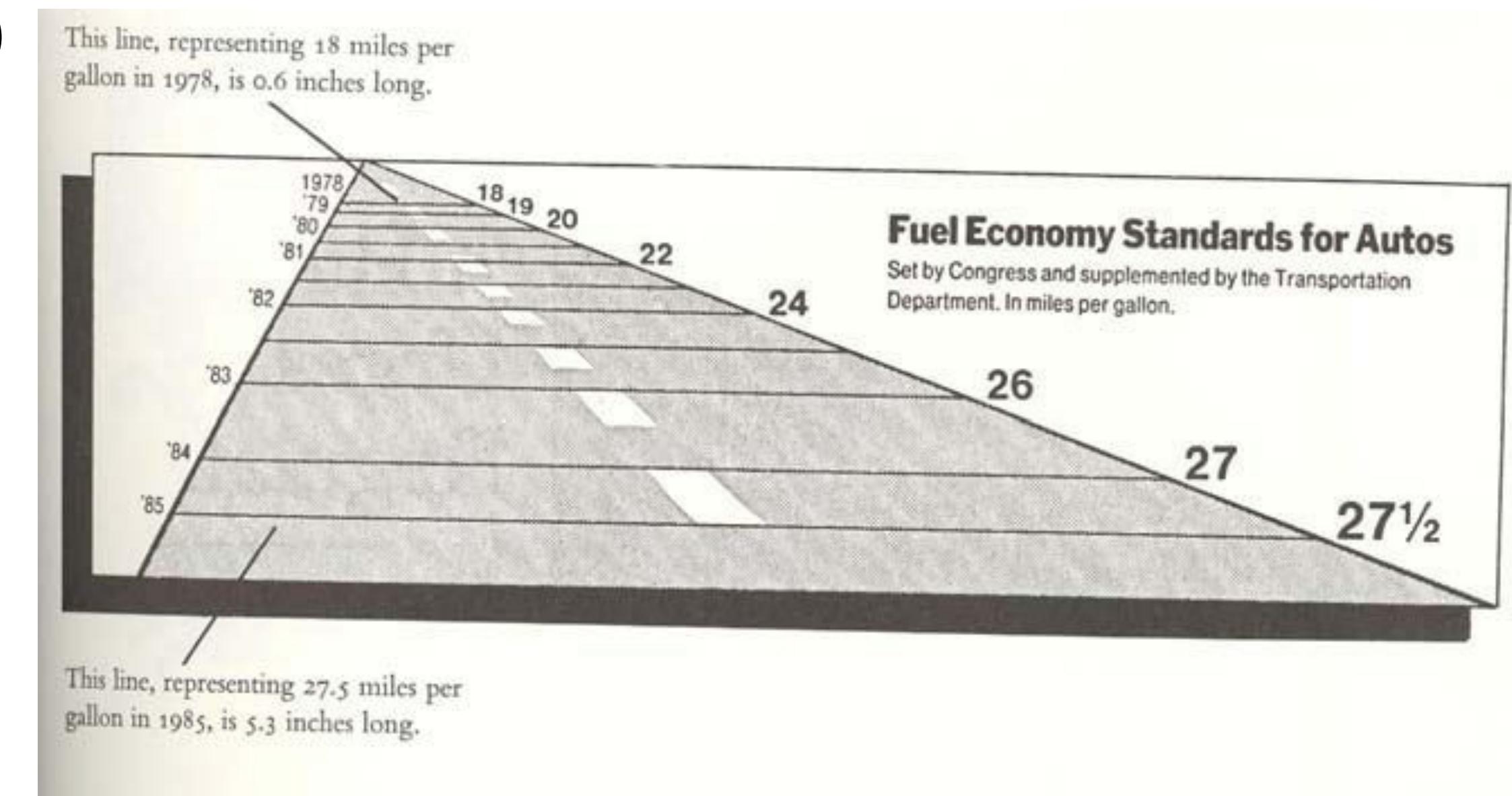
# Lie Factor

**Lie Factor = (Size of effect in graphic)  
(Size of effect in data)**

**Lie Factor =  $>1$ , overstating**

**Lie Factor =  $1$ , accurate :)**

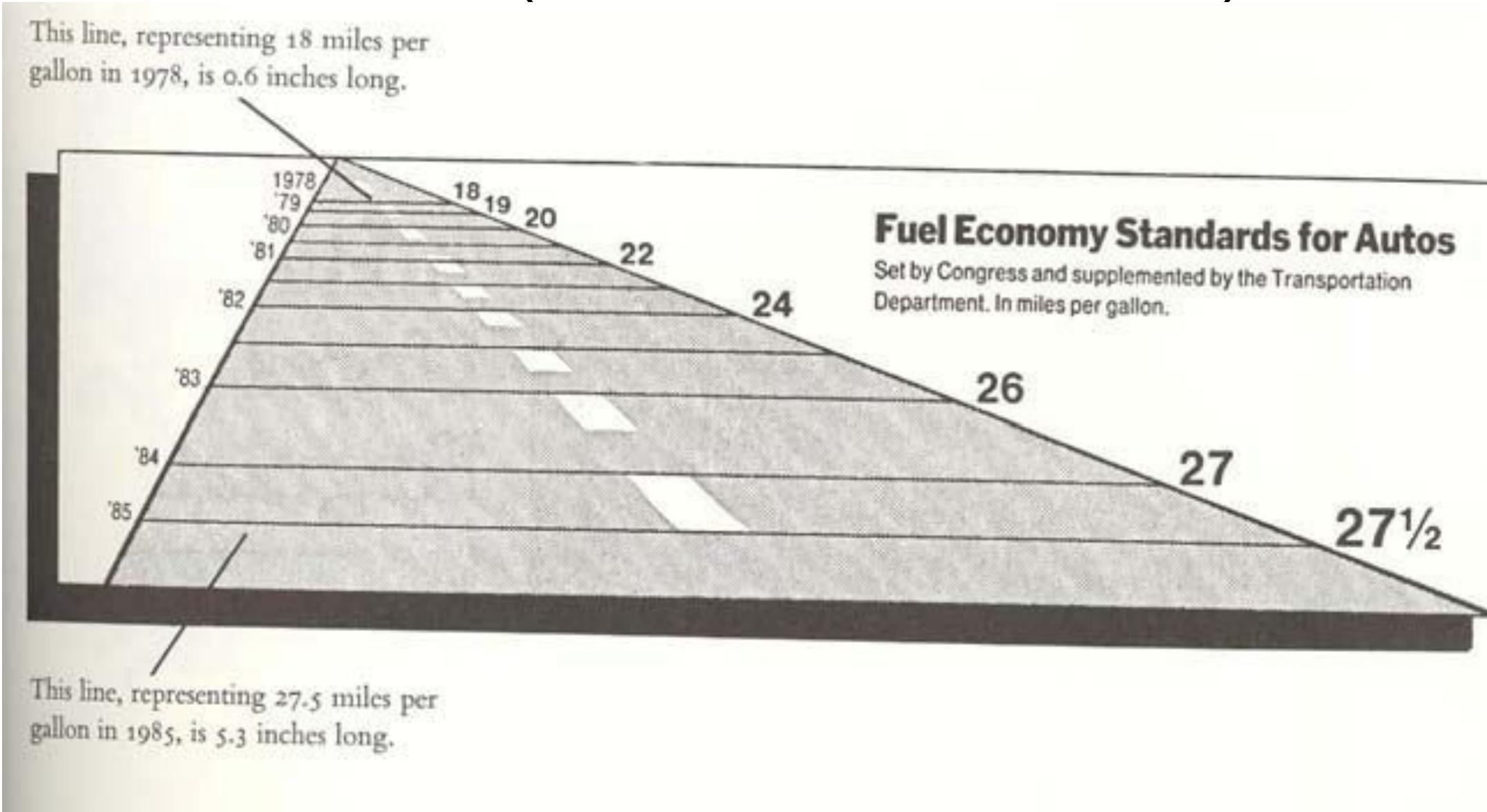
**Lie Factor =  $<1$ , understating**



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

# Lie Factor

**Lie Factor = (Size of effect in graphic)  
(Size of effect in data)**

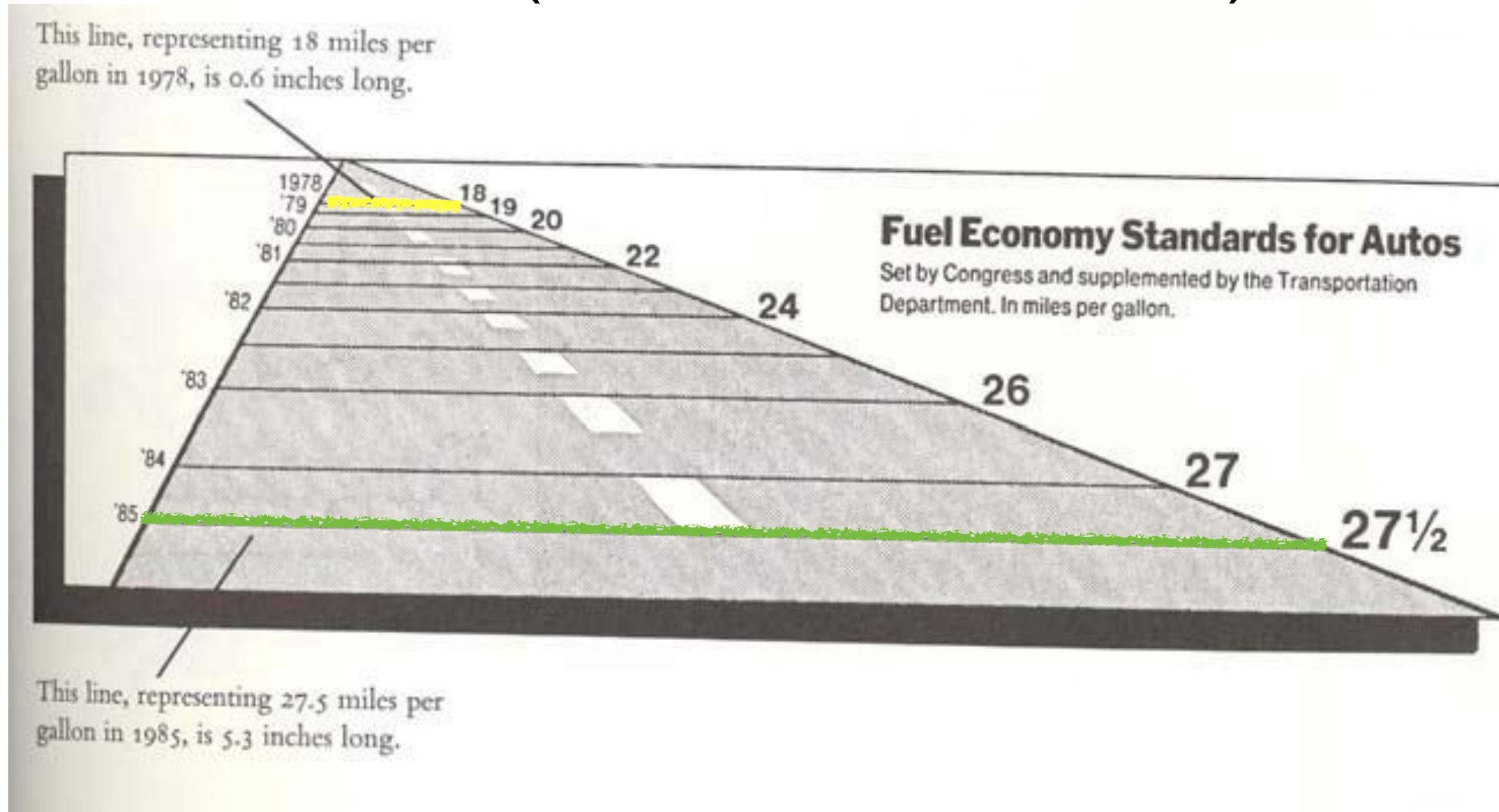


“The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured.”

Tufte, “Visual Display of Quantitative Information” (1983) 27

# Lie Factor

**Lie Factor = (Size of effect in graphic)  
(Size of effect in data)**



$$\text{Image} = \frac{5.3''}{0.6''} = 7.83 = 783\%$$

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

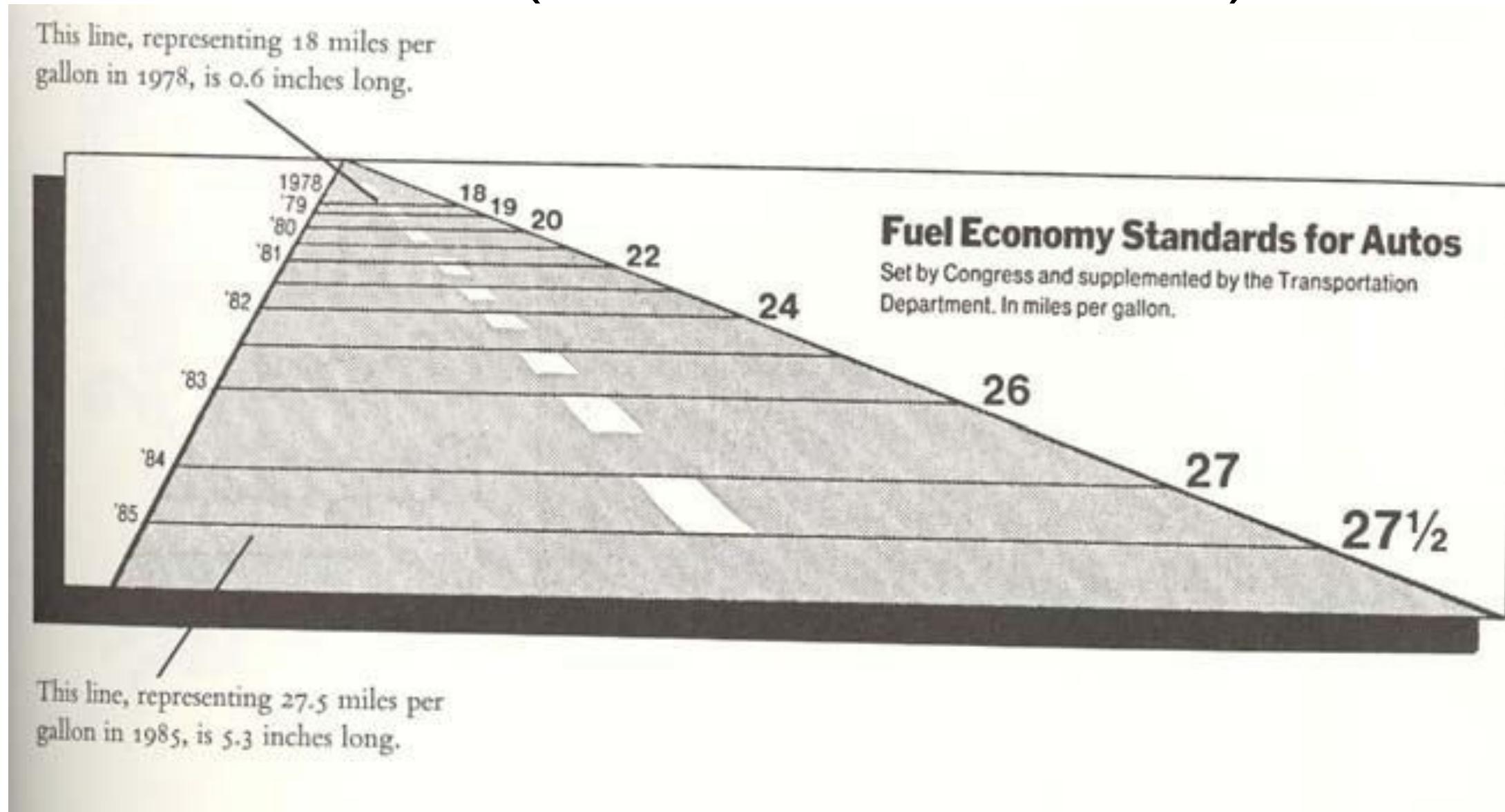
$$\text{Lie Factor} = \frac{783\%}{53\%} = 14.8$$

**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.”

# Lie Factor

**Lie Factor = (Size of effect in graphic)  
(Size of effect in data)**



$$\text{Image} = \frac{5.3''}{0.6''} = 7.83 = 783\%$$

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

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

**Lie Factor =  $> 1$ , overstating**



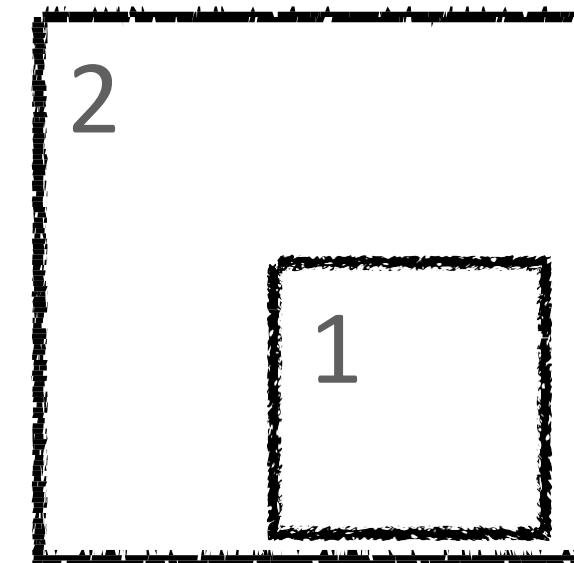
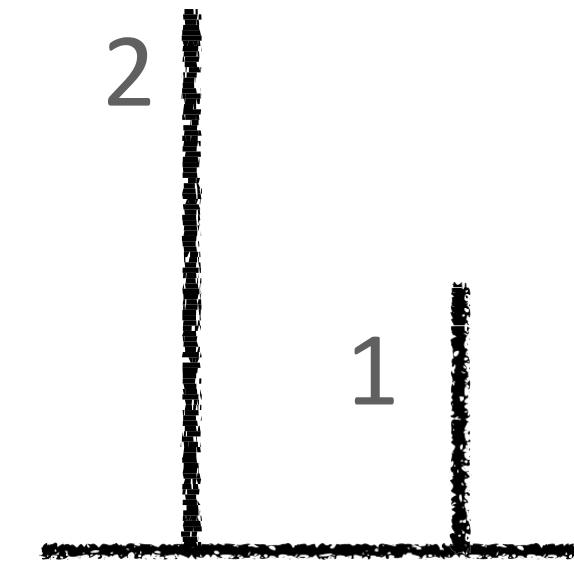
“The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured.”

IN-CLASS ACTIVITY:  
Calculate for yourself!

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

# Lie Factor

**Lie Factor = (Size of effect in graphic)  
(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.”

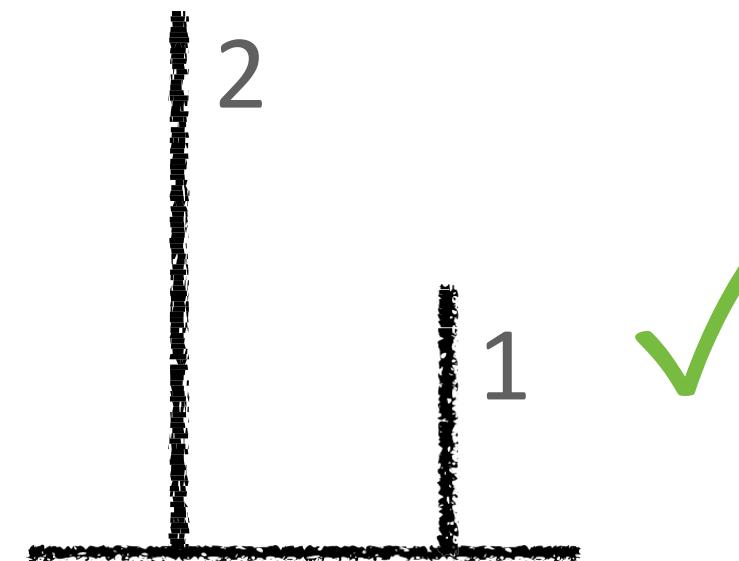
Tufte, “Visual Display of Quantitative Information” (1983)

IN-CLASS ACTIVITY:  
Calculate for yourself!

# Lie Factor

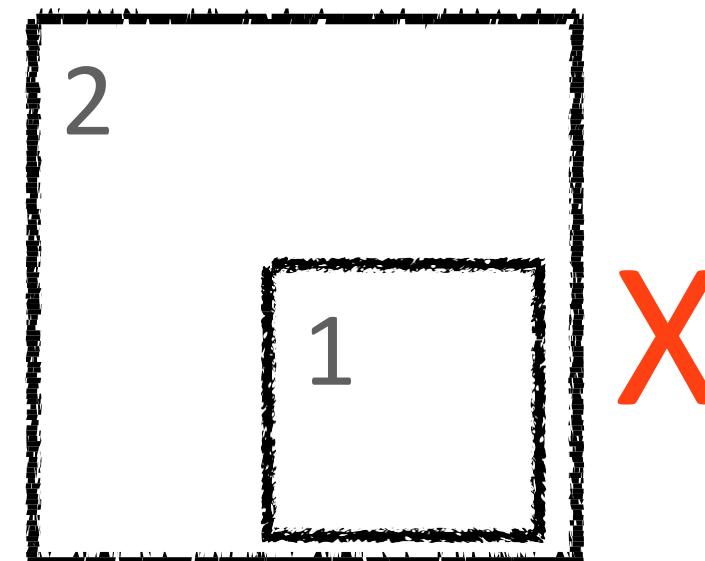
$$\text{Data} = \frac{2 - 1}{1} = 1 = 100\%$$

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



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

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



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

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

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

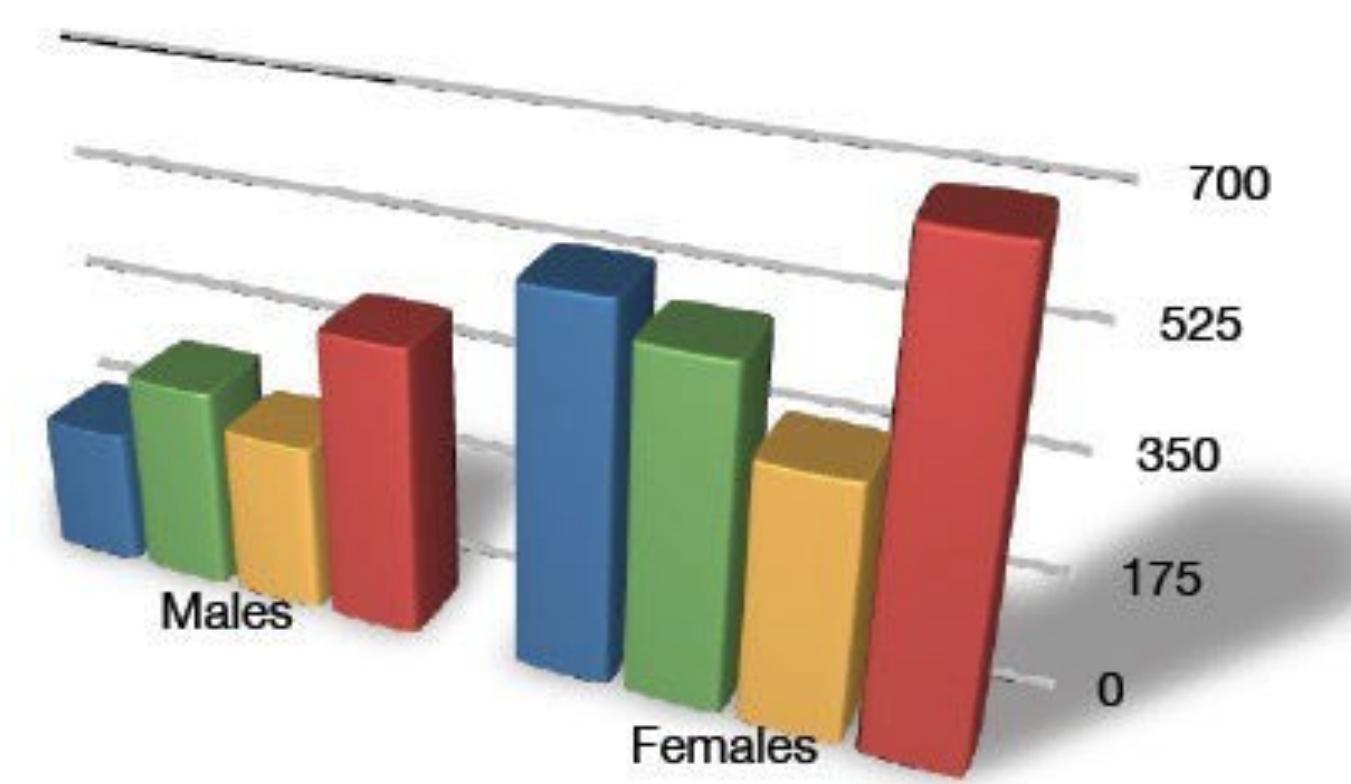
Tufte, “Visual Display of Quantitative Information” (1983)

# “Graphical Integrity”

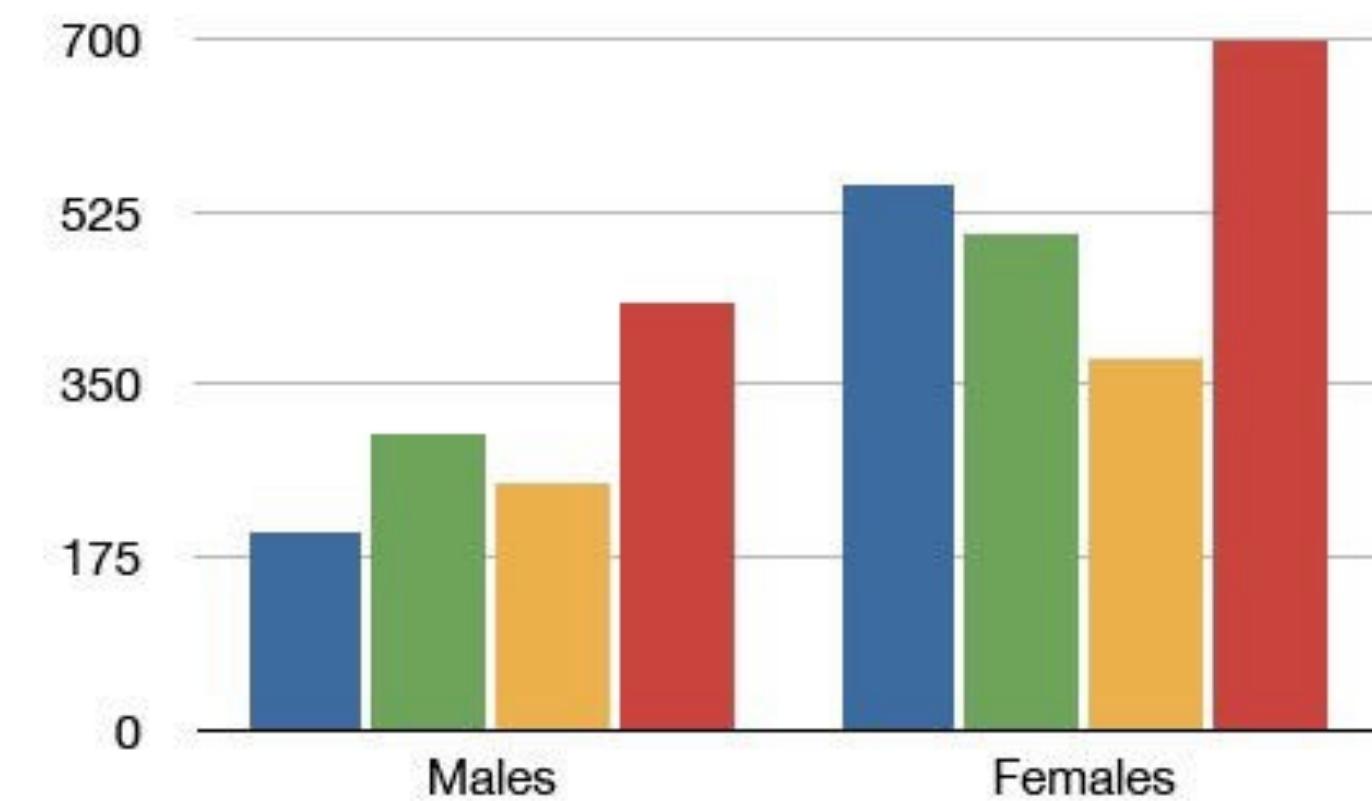
Data Ink = the ink used to show data

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

*Tufte: maximize the data  
ink ratio*



■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+



■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+

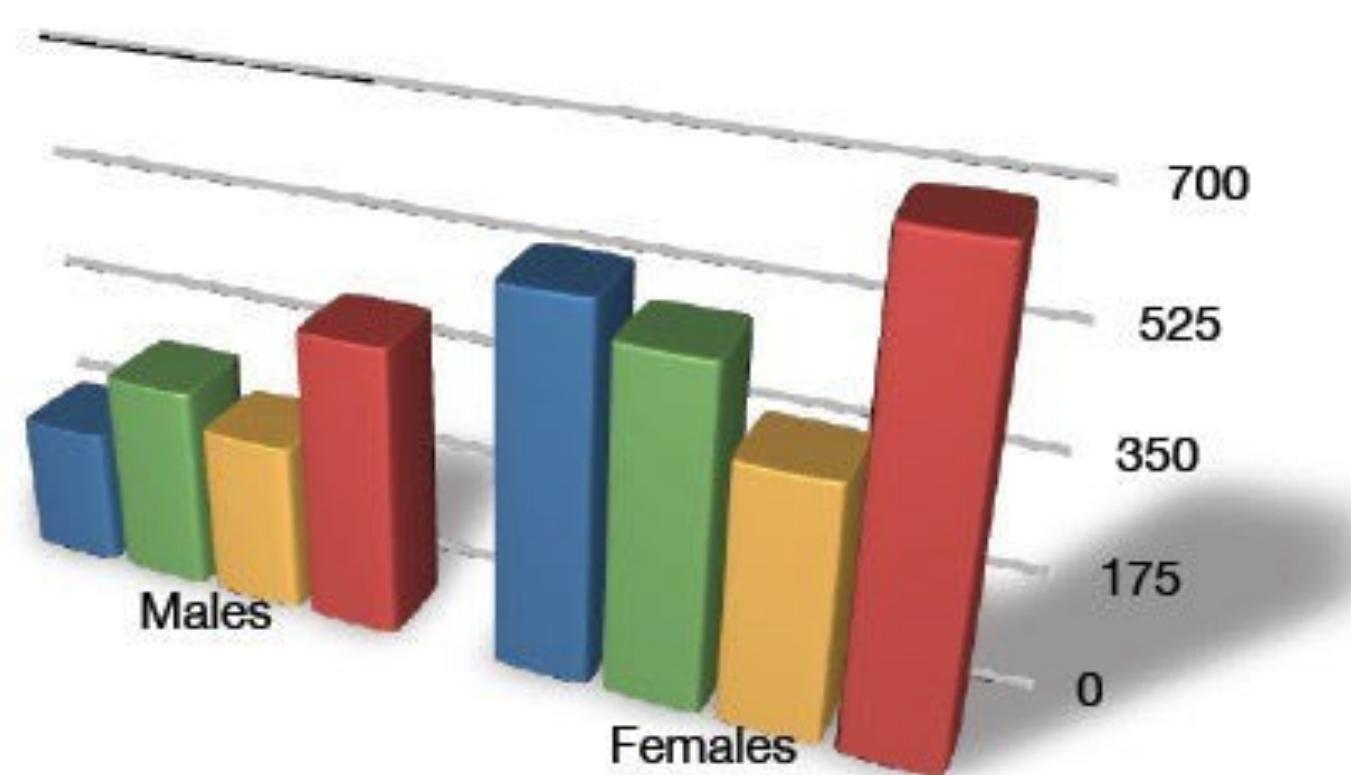
# “Graphical Integrity”

Data Ink = the ink used to show data

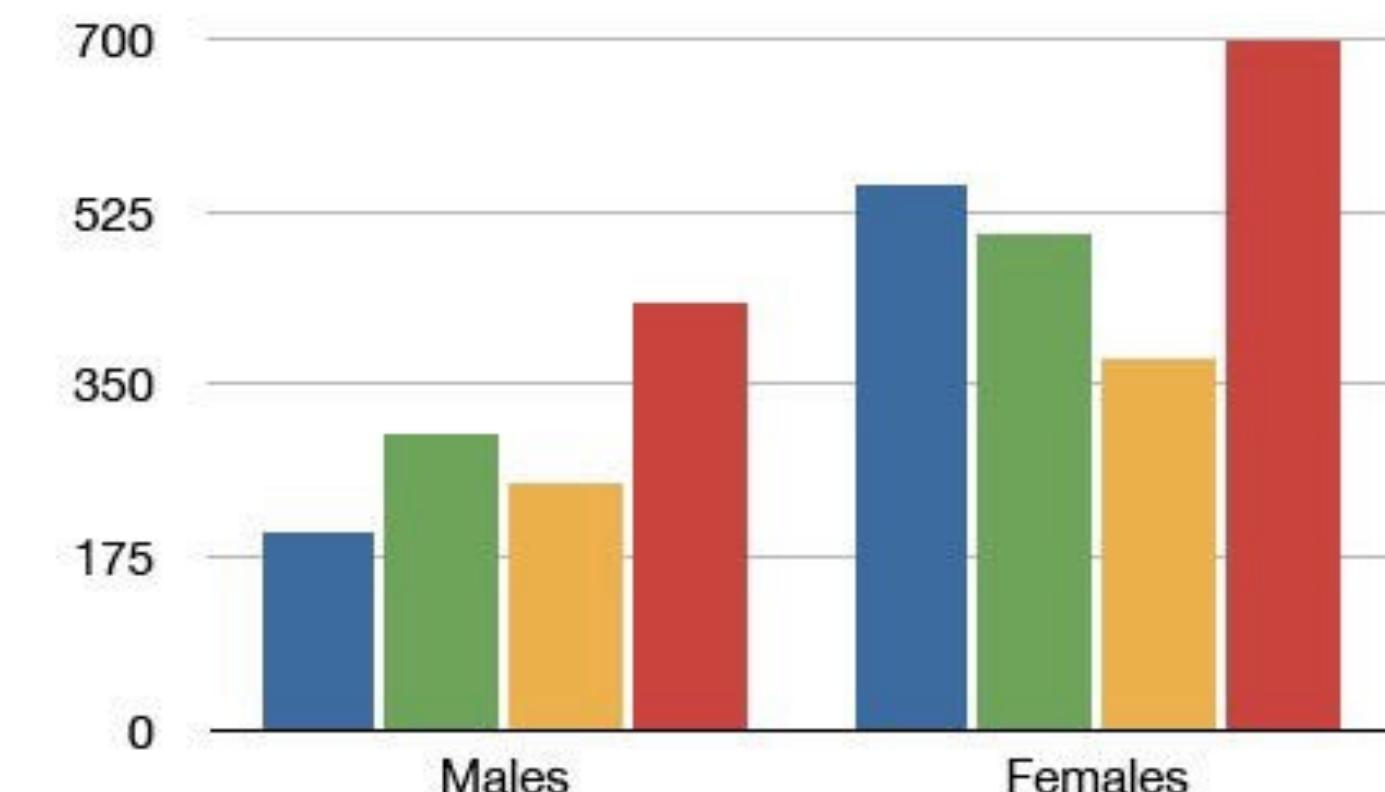
Data Ink Ratio =  $\frac{\text{data-ink}}{\text{total ink in graphic}}$

*Tufte: maximize the data  
ink ratio*

*Low Data Ink Ratio*



*High Data Ink Ratio*



■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+

■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+

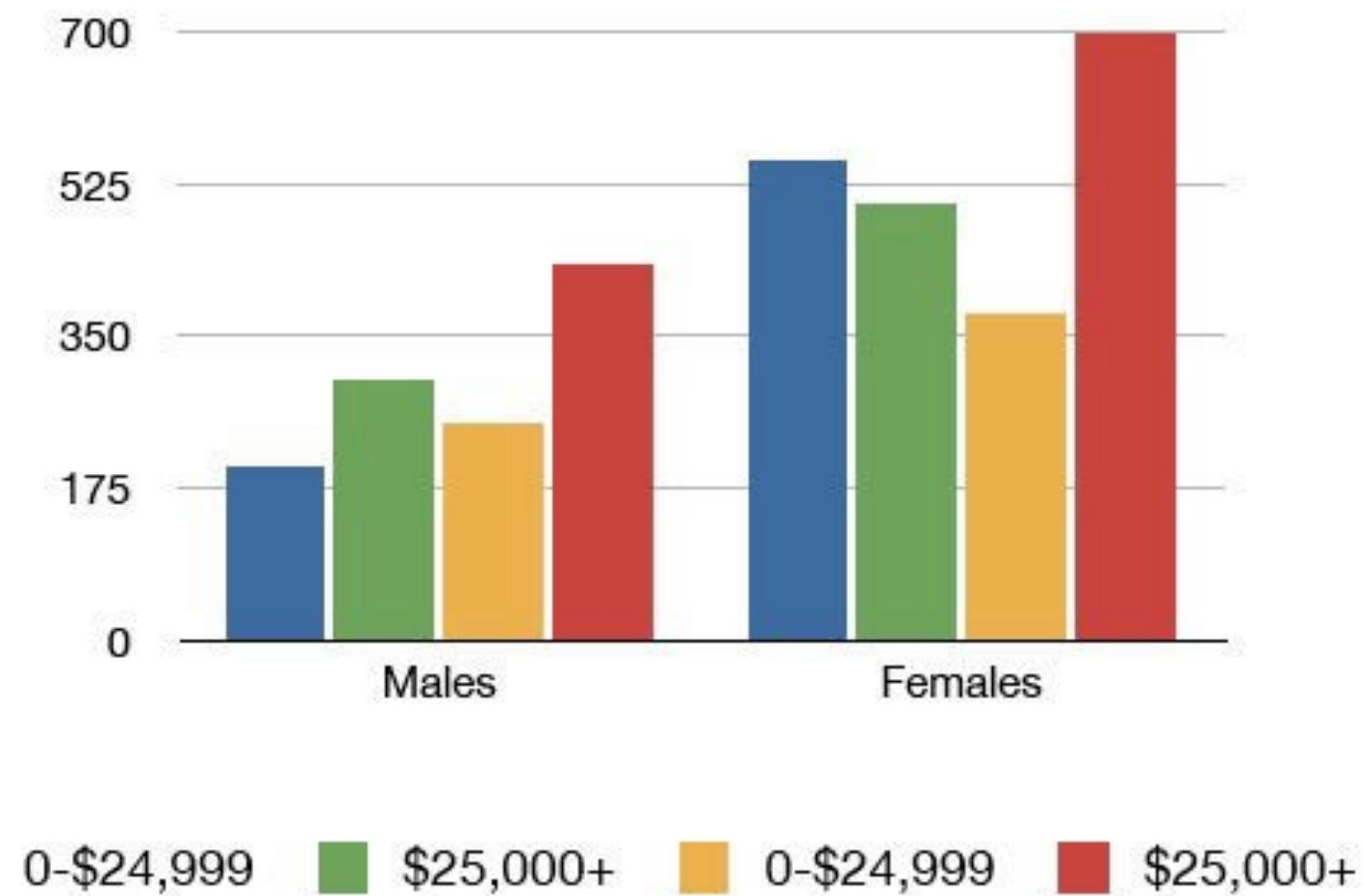
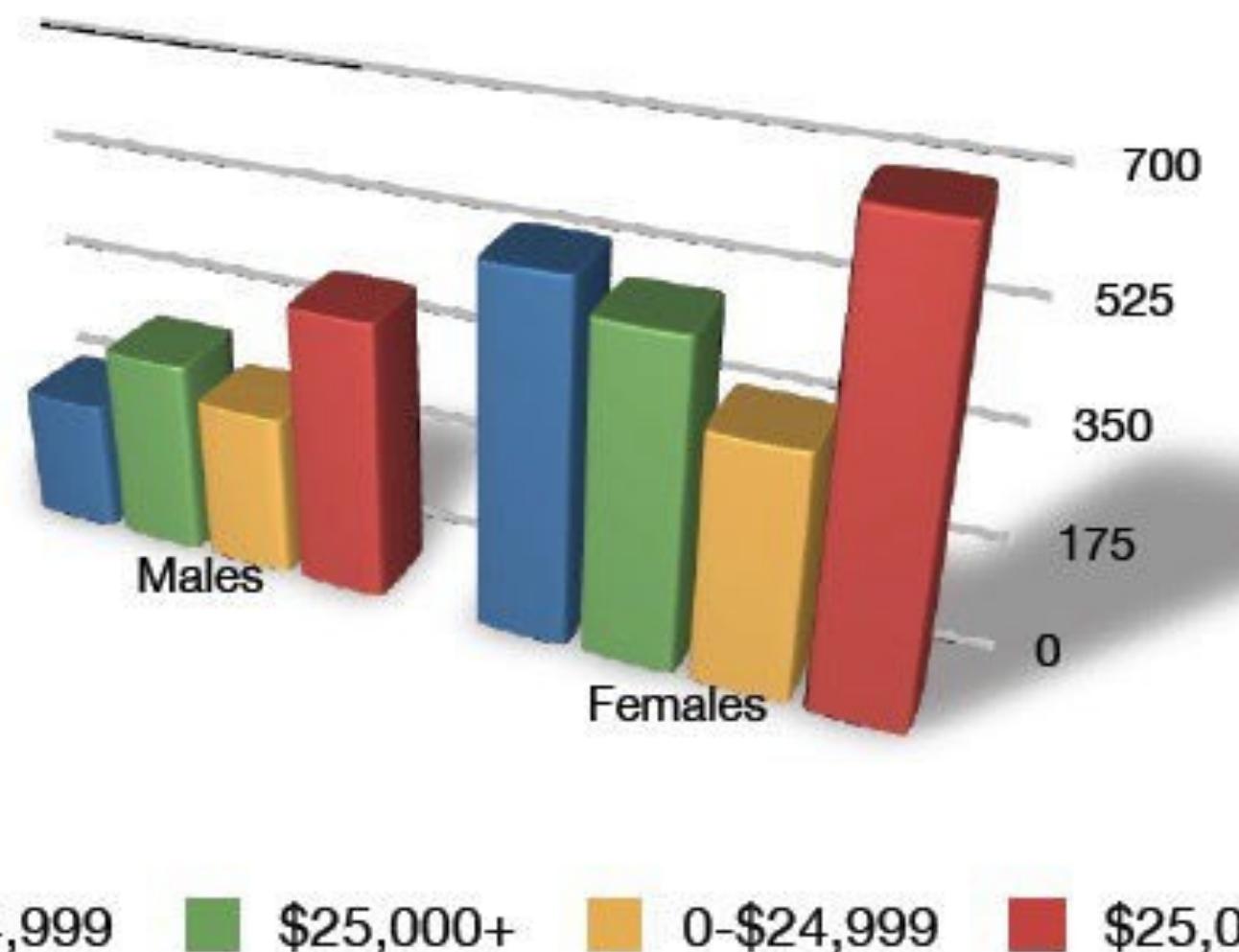
# “Graphical Integrity”

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

# “No Unjustified 3D”

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

# “No Unjustified 3D”



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

# “No Unjustified 3D”

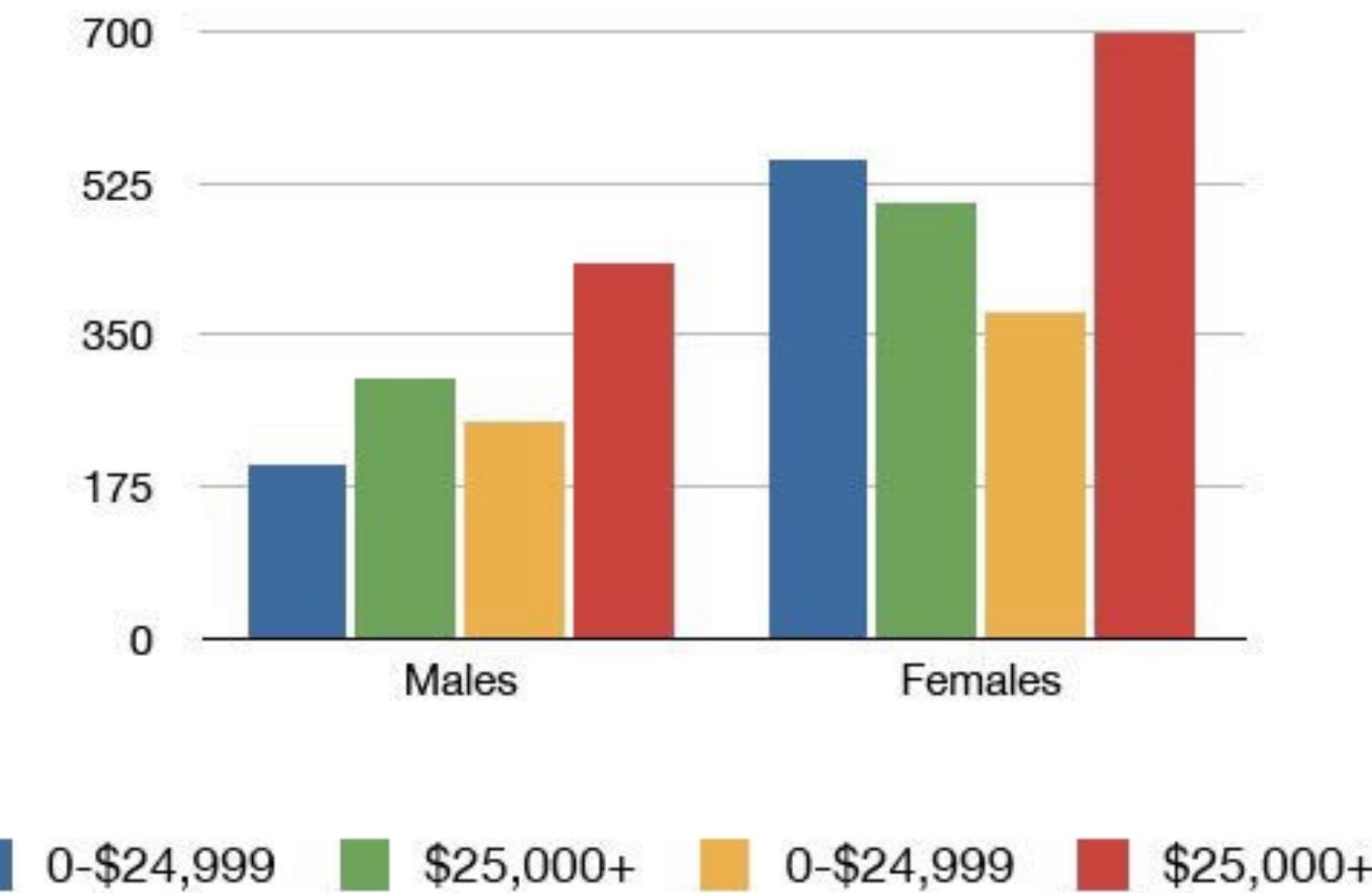
# Dimensions in data: 2

# Dimensions in plot: 3



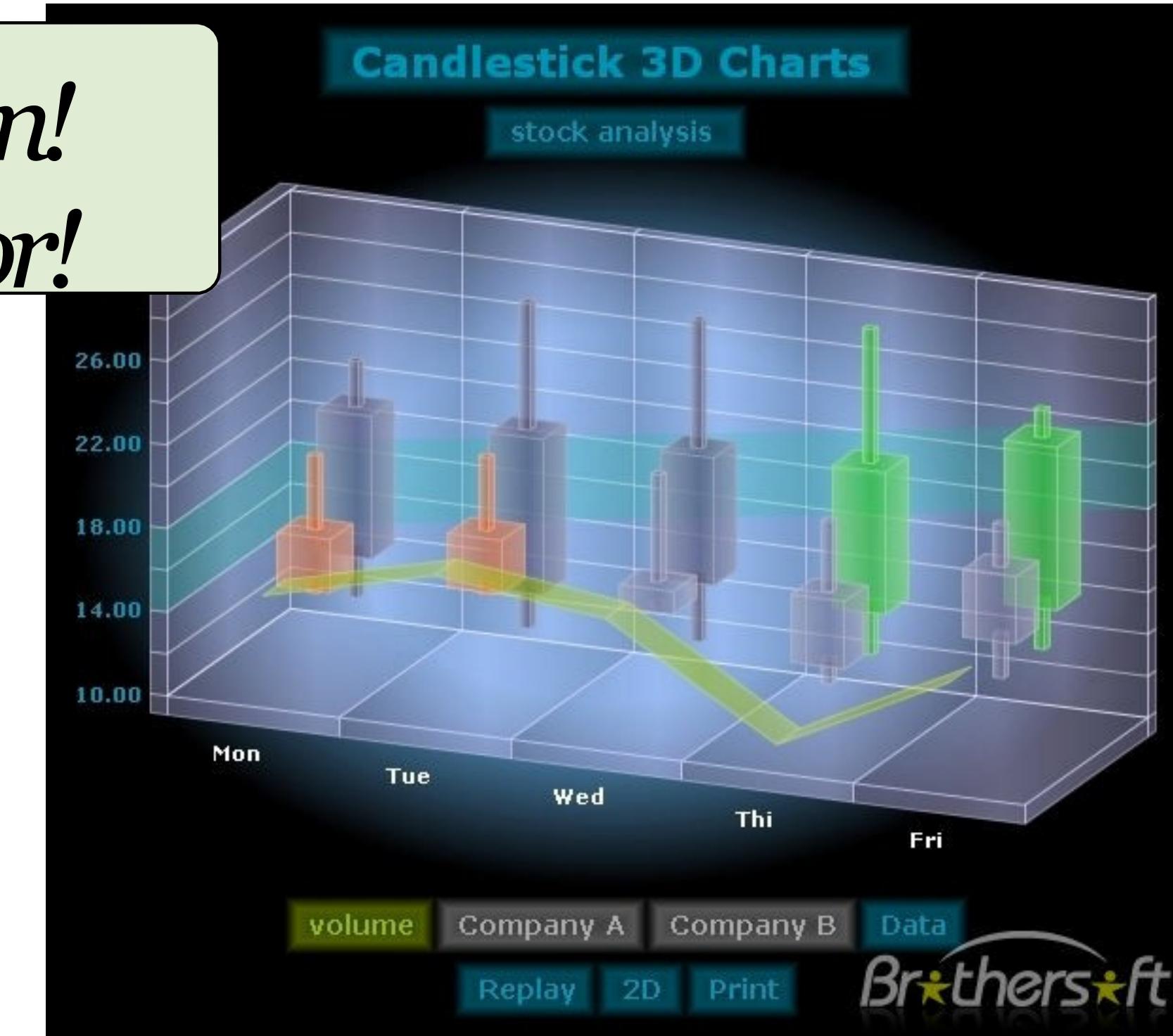
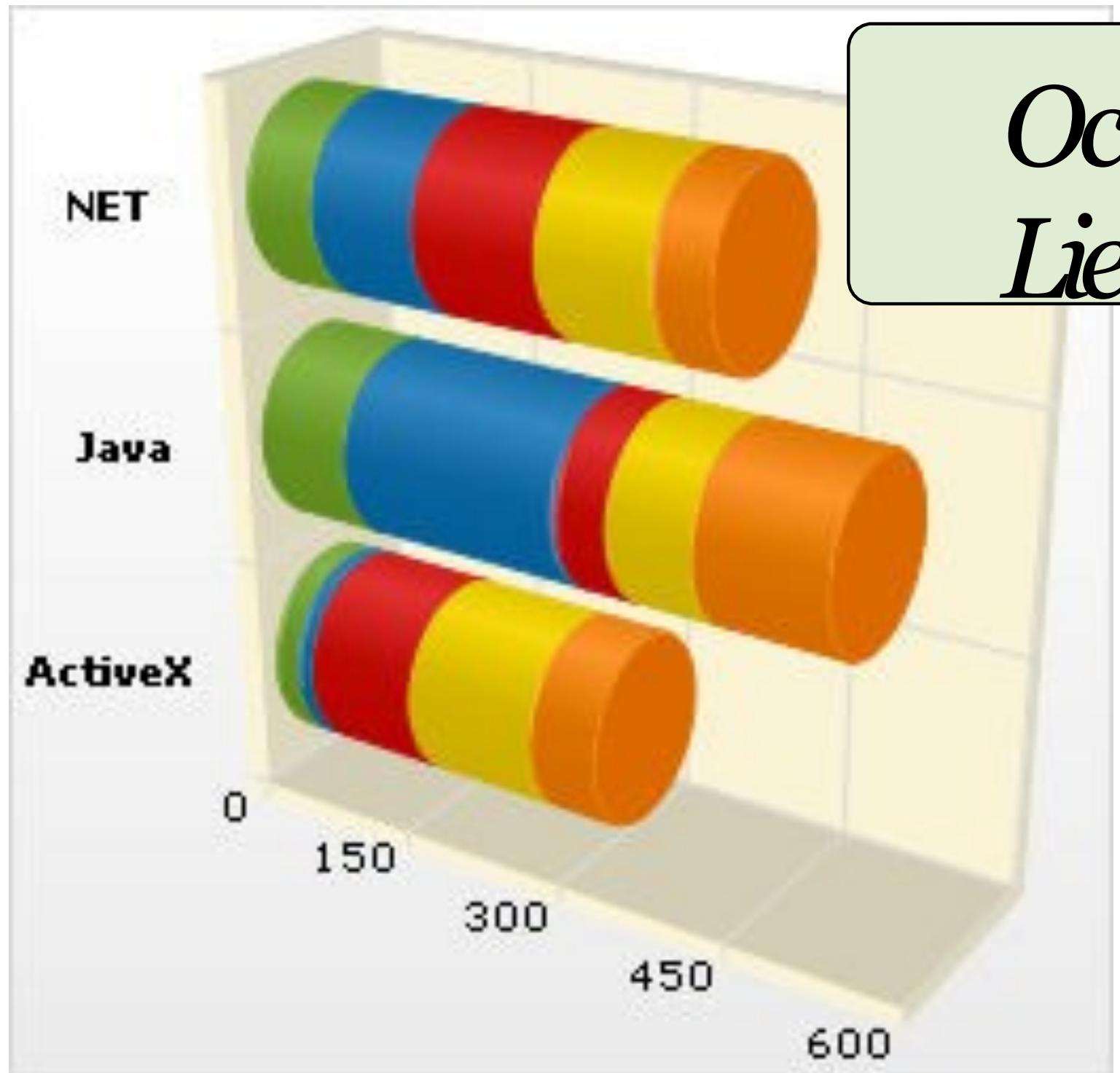
# Dimensions in data: 2

# 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”

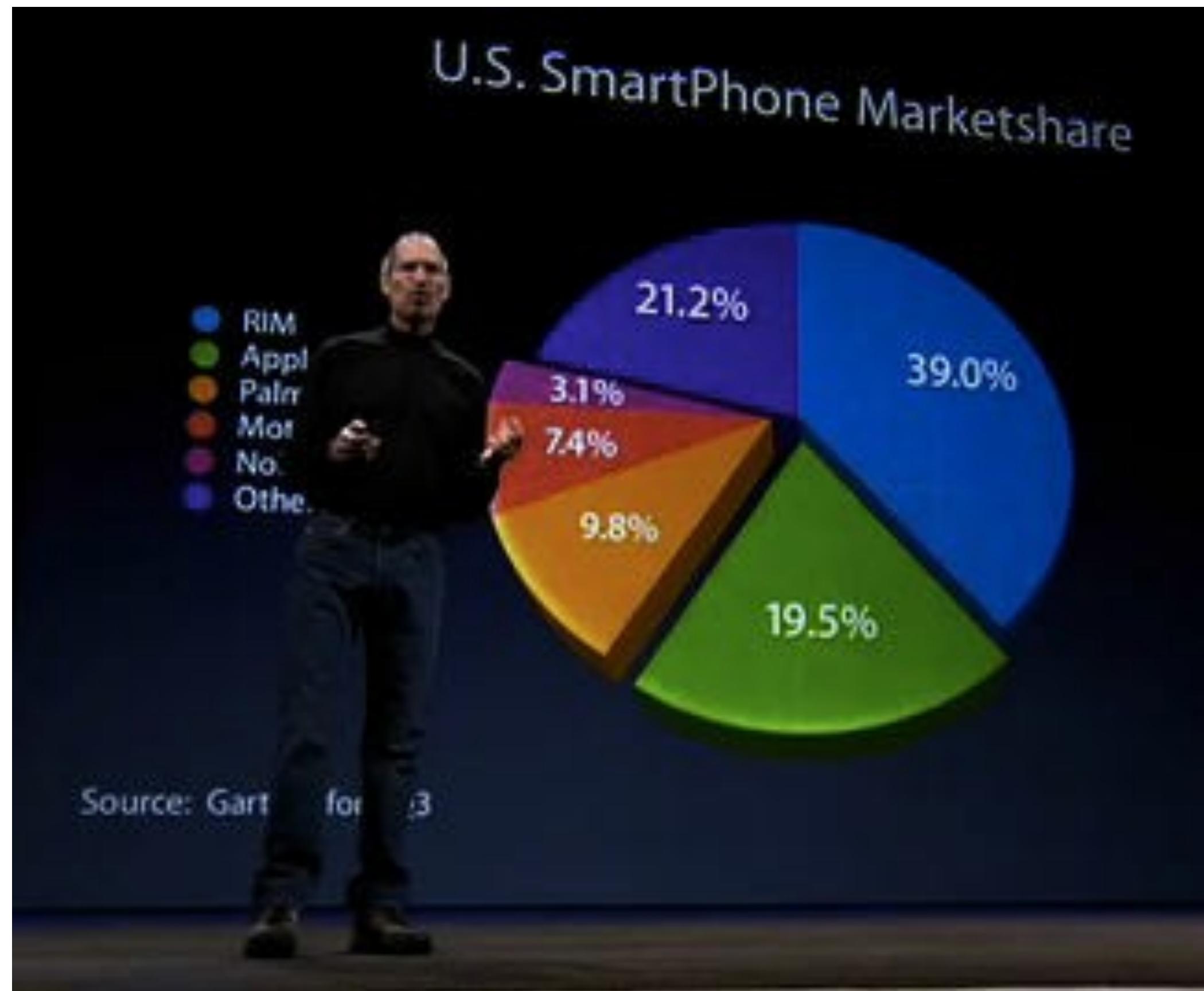


<http://help.infragistics.com/Help/Doc/WinForms/2014.2/CLR4.0/html/Images/Chart Bar Chart 03.png>

<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.”

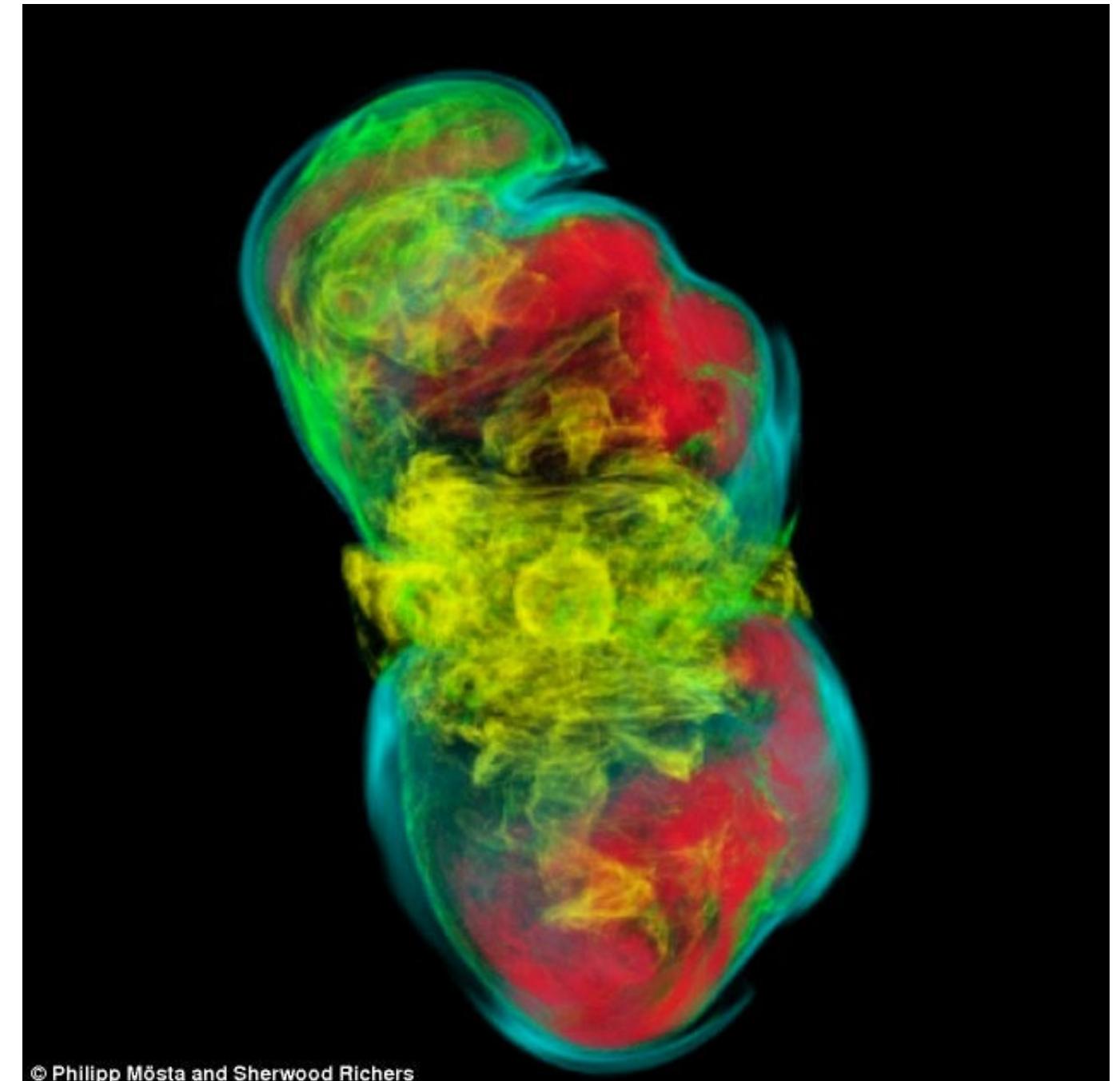
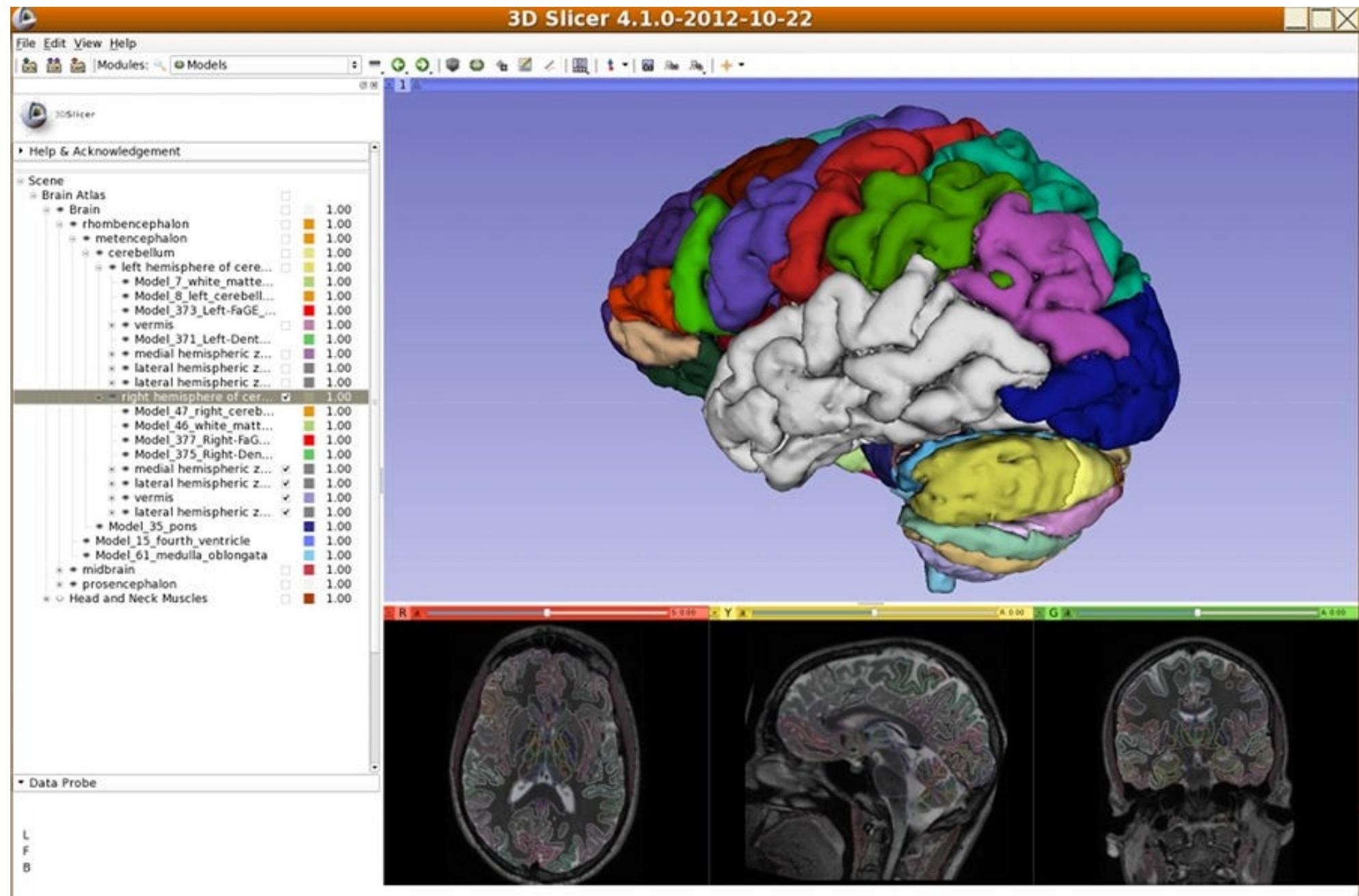
# “No Unjustified 3D”



Unjustified 3D!

Lie factor!

# “No Unjustified 3D”



© Philipp Mösta and Sherwood Richers

“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.”

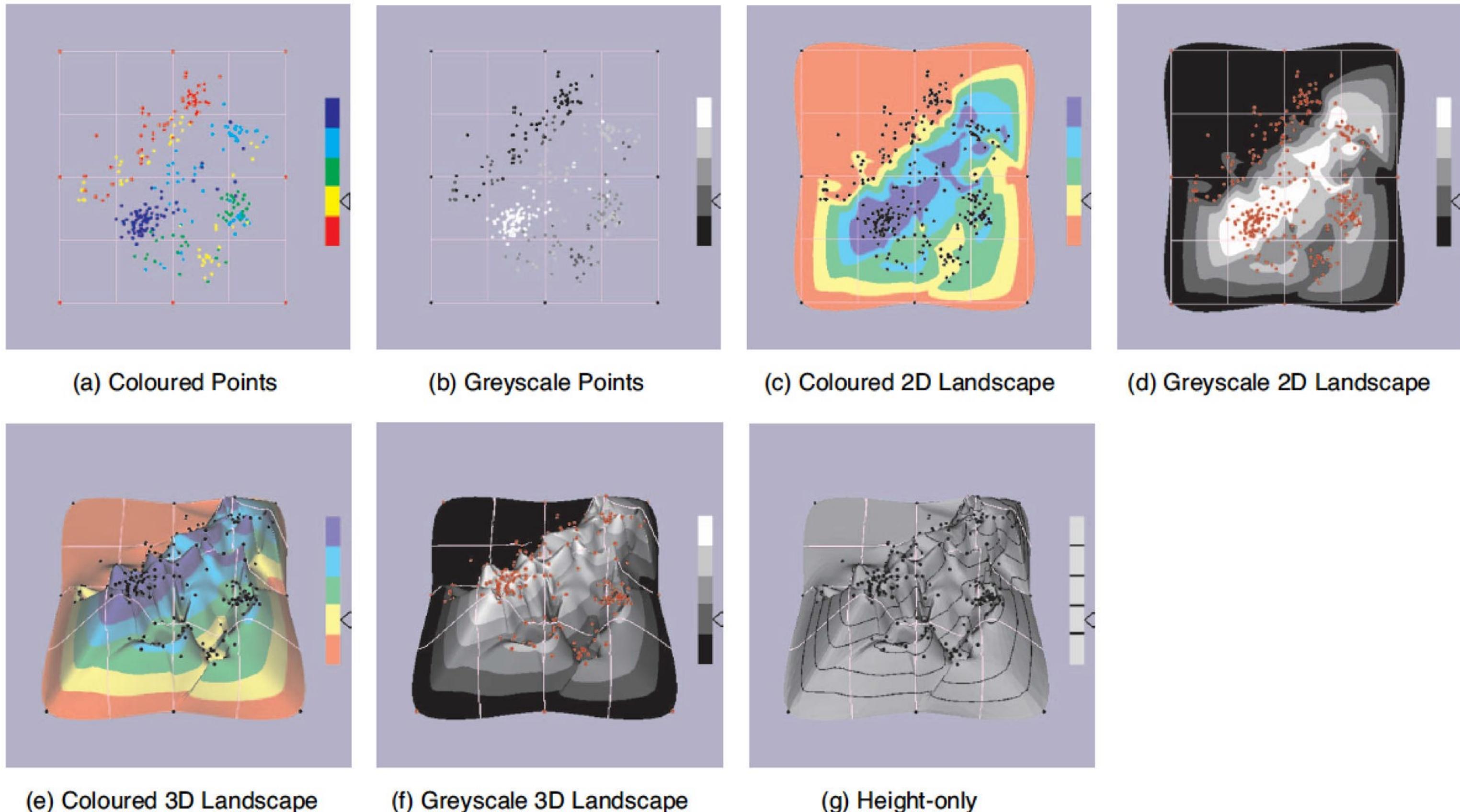


Fig. 1 Point-based displays and information landscapes used in our experiment. All displays show the same data.

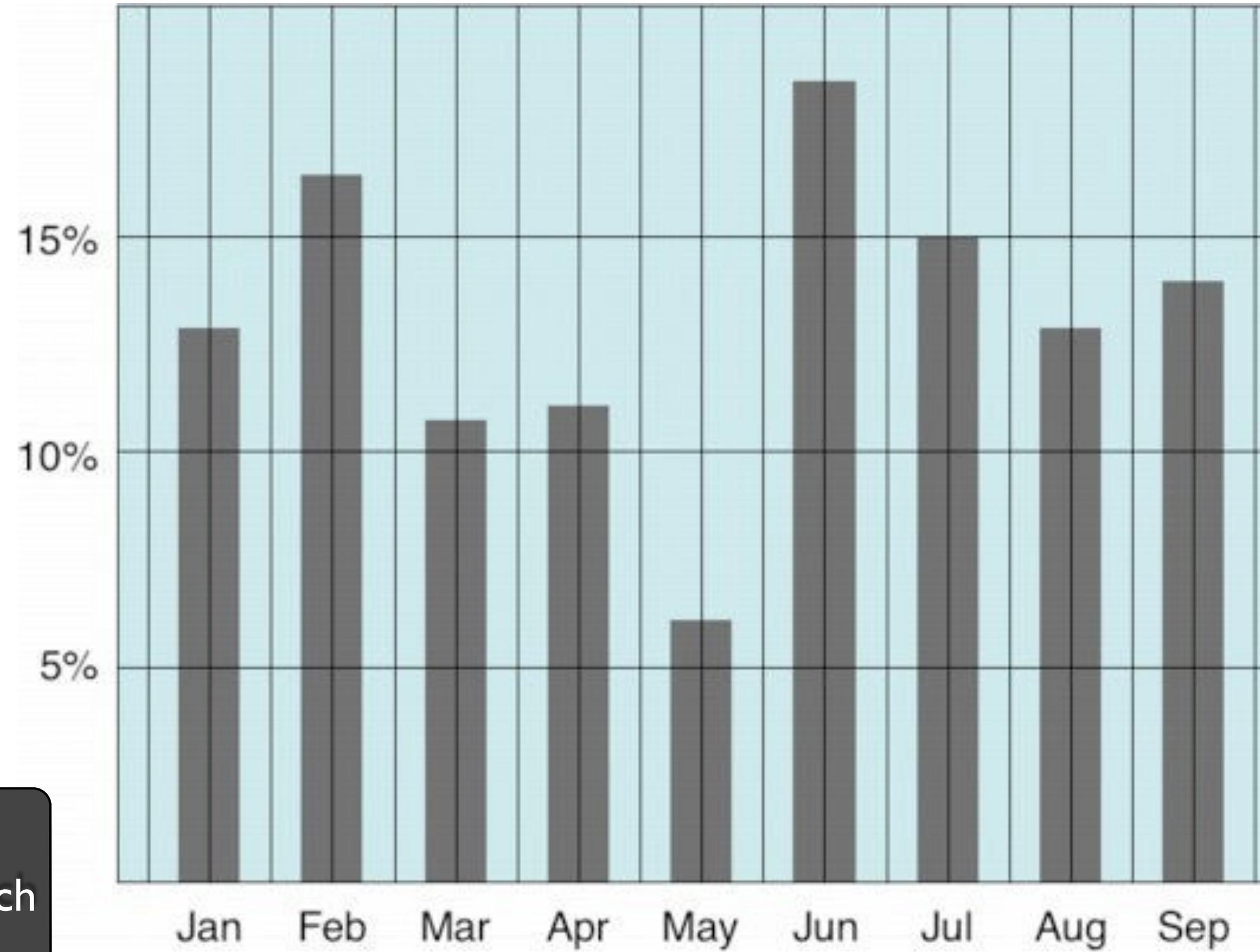
“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

# “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.

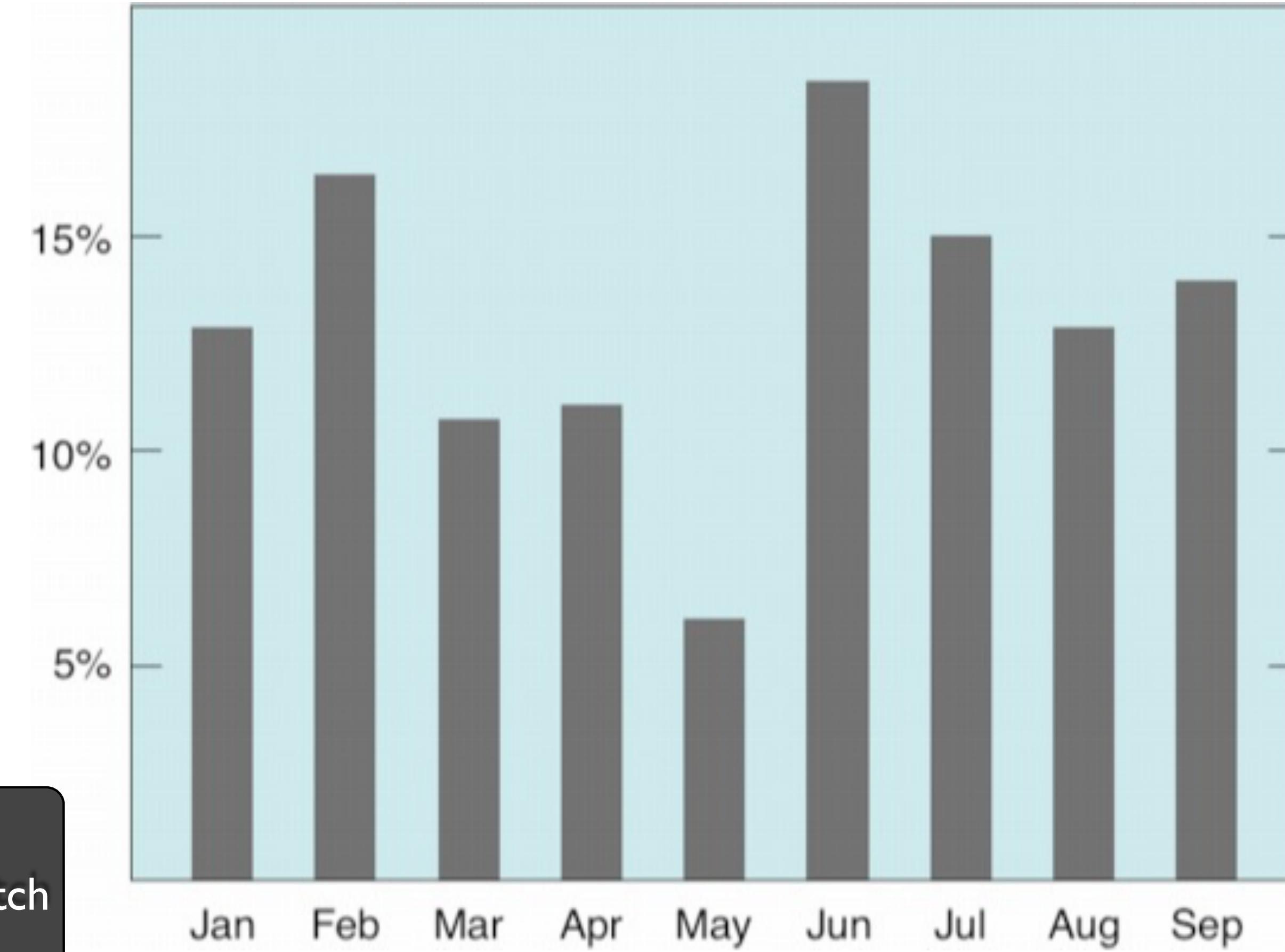
# “Graphical Integrity”



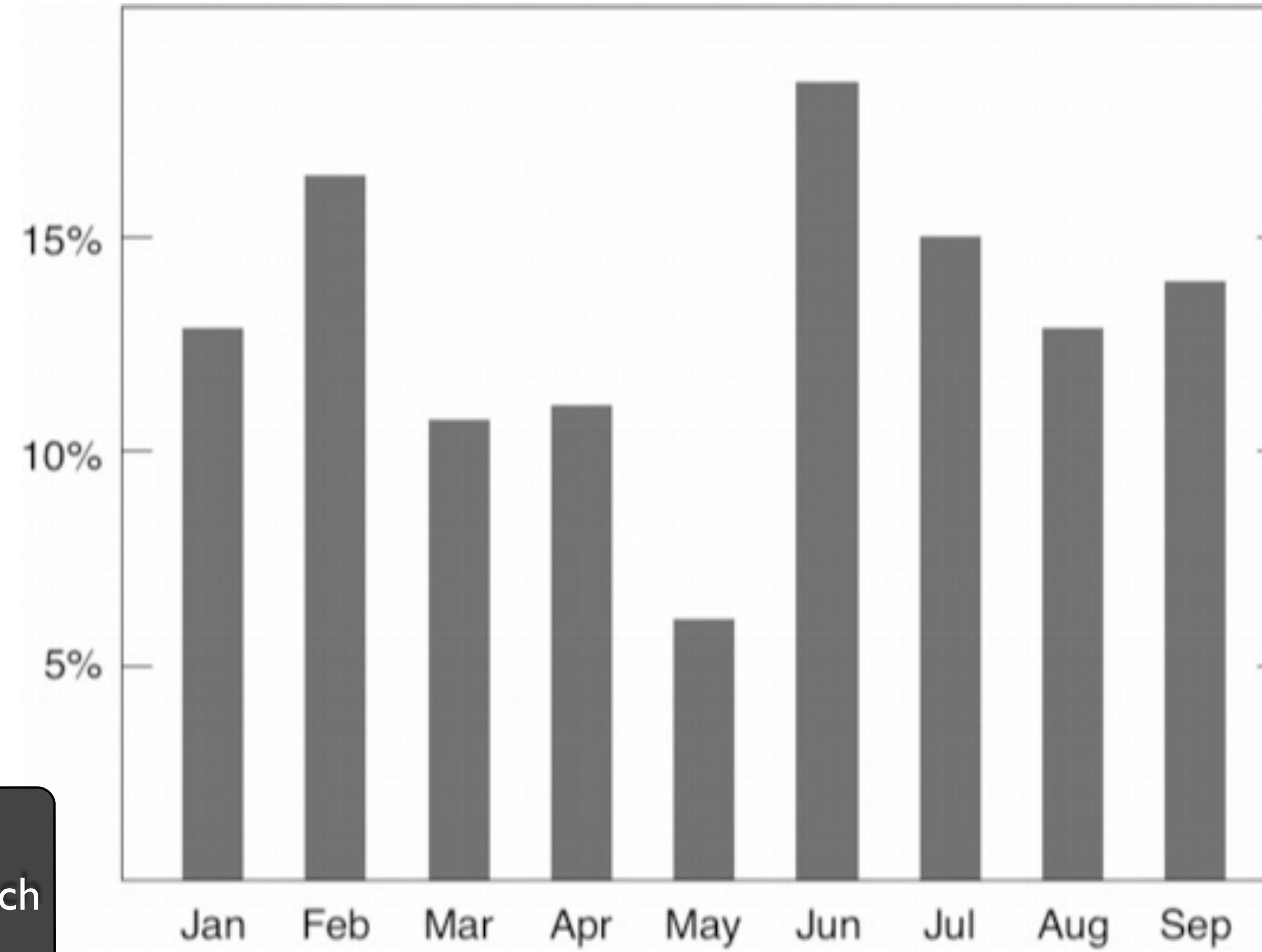
IN-CLASS ACTIVITY:  
Use paper/pen to sketch  
“Tufte” version!

# “Graphical Integrity”

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

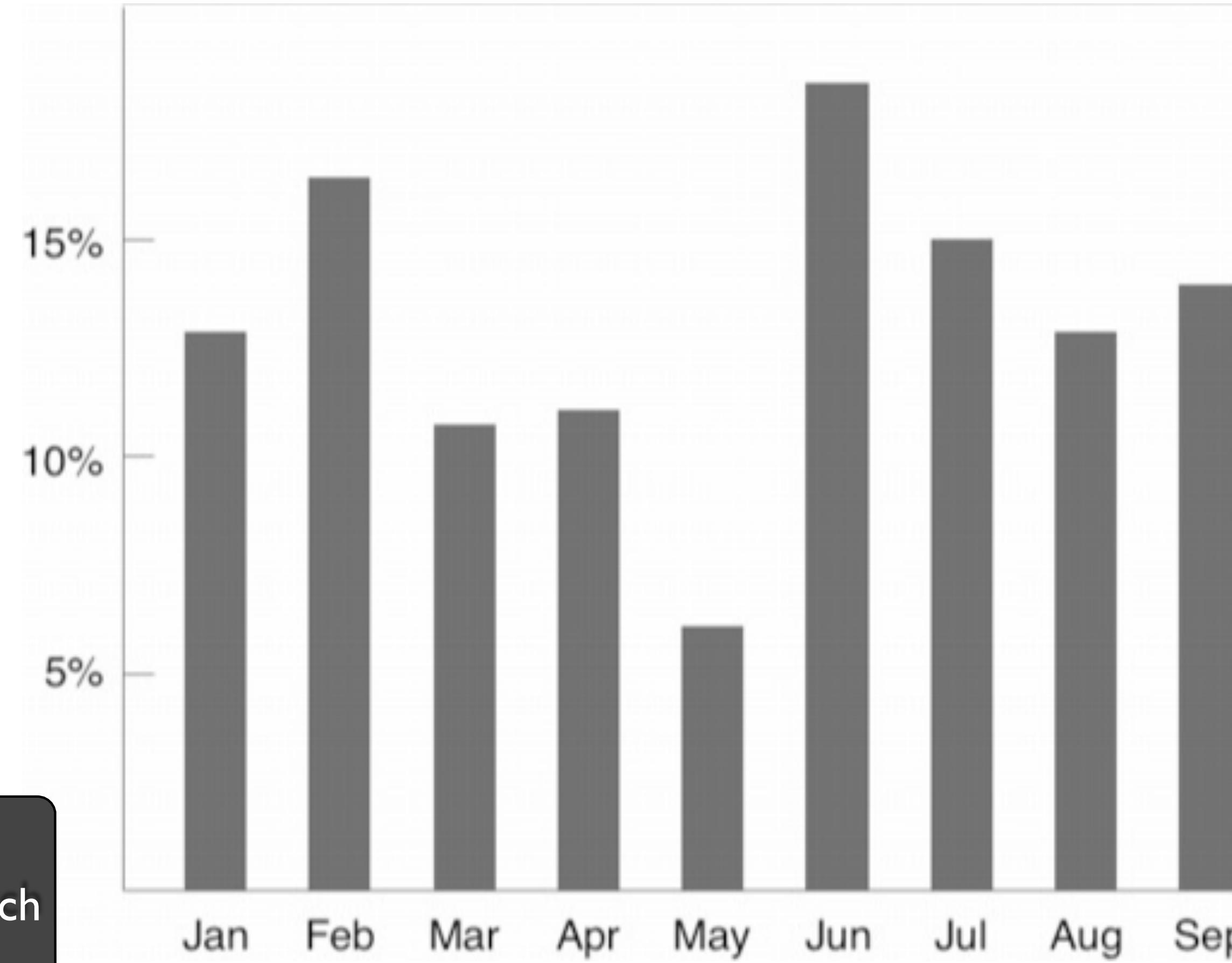


# “Graphical Integrity”



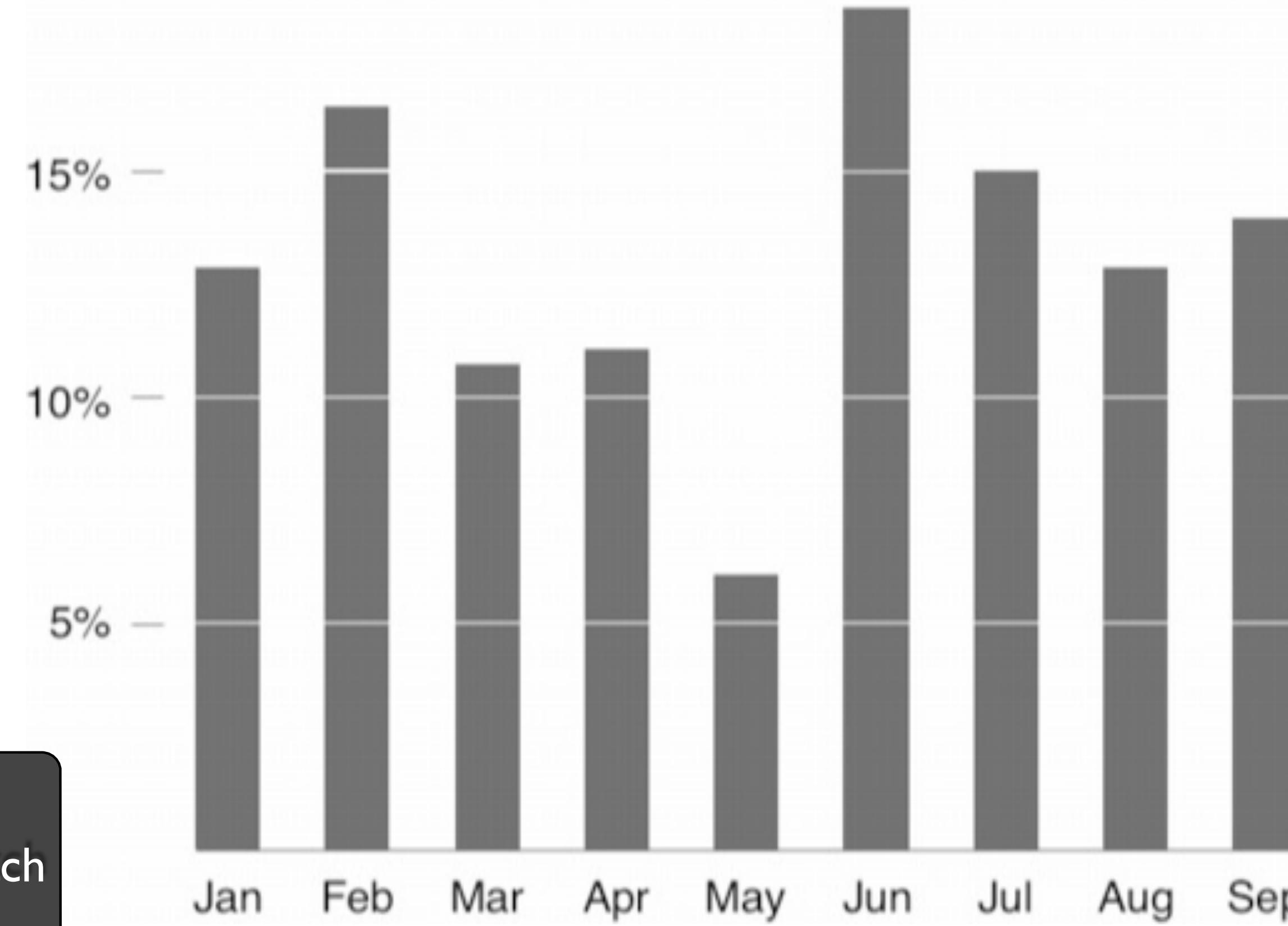
IN-CLASS ACTIVITY:  
Use paper/pen to sketch  
“Tufte” version!

# “Graphical Integrity”



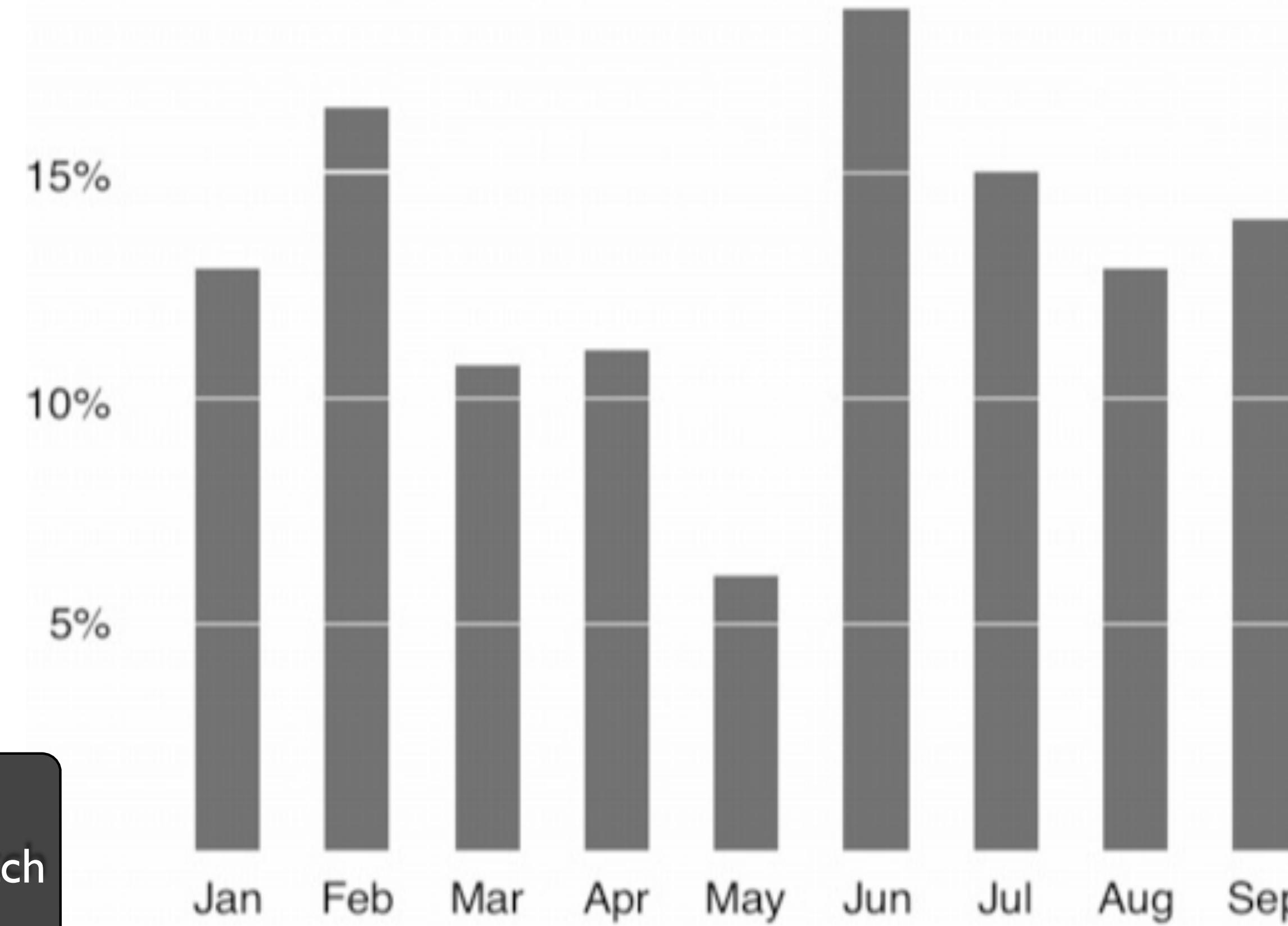
IN-CLASS ACTIVITY:  
Use paper/pen to sketch  
“Tufte” version!

# “Graphical Integrity”



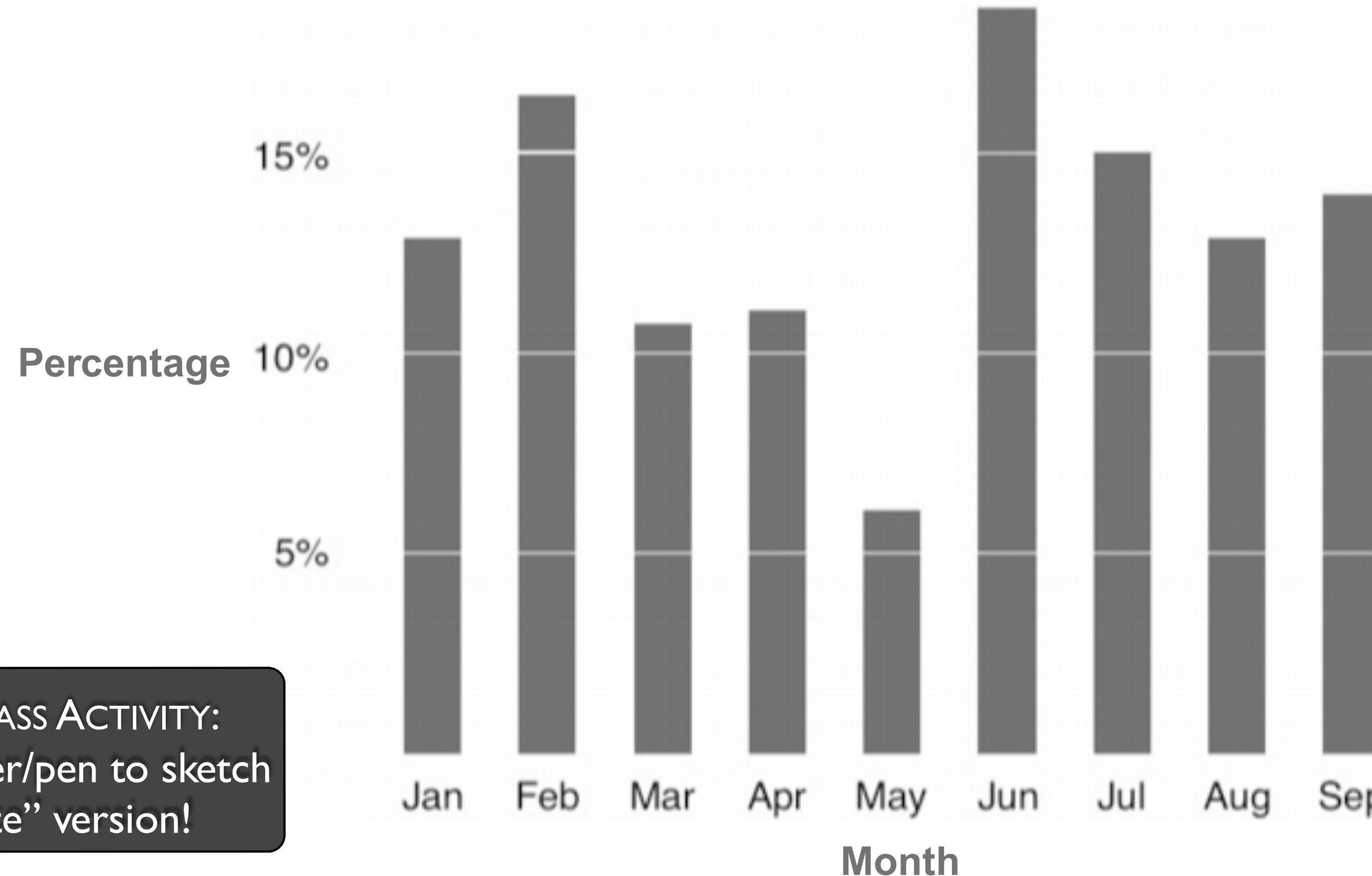
IN-CLASS ACTIVITY:  
Use paper/pen to sketch  
“Tufte” version!

# “Graphical Integrity”



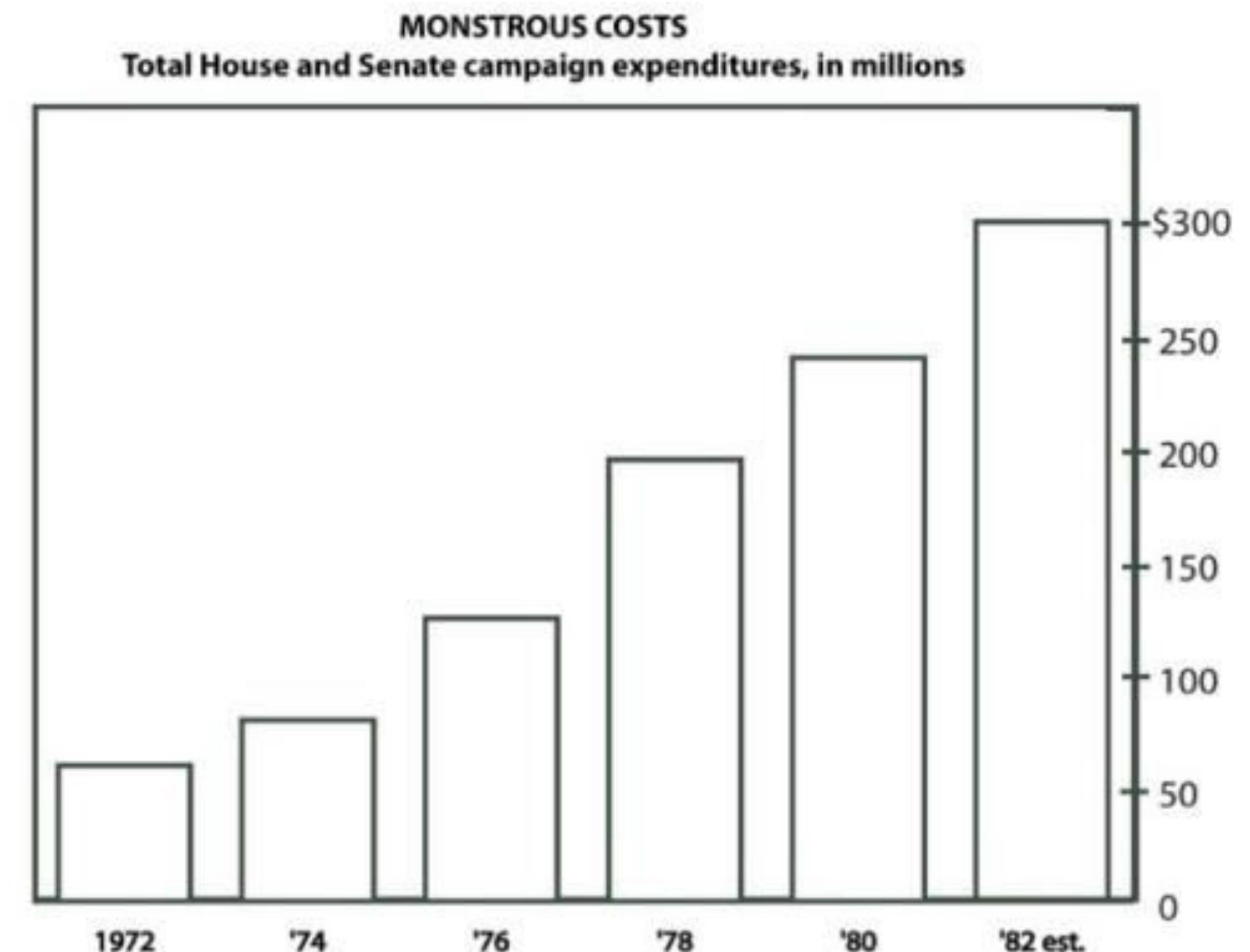
IN-CLASS ACTIVITY:  
Use paper/pen to sketch  
“Tufte” version!

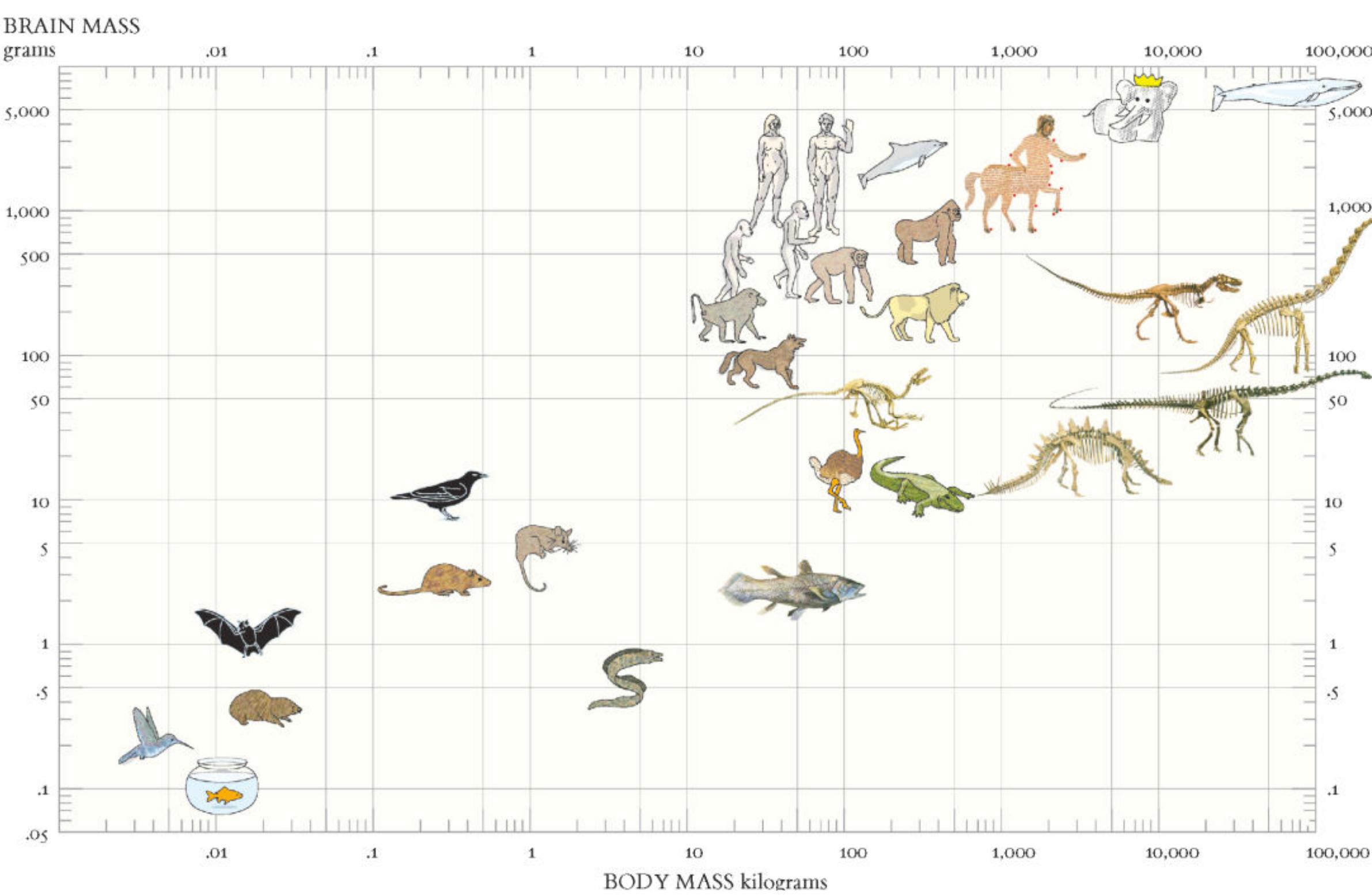
# “Graphical Integrity”



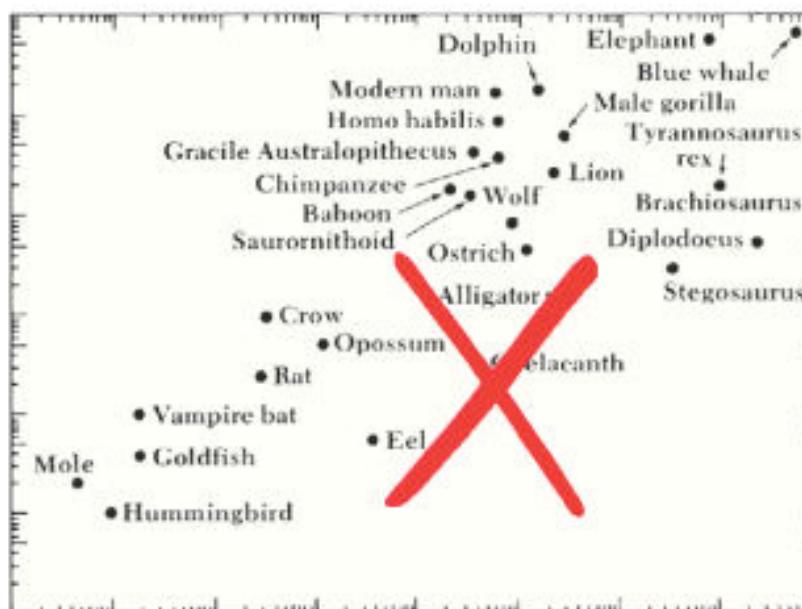
IN-CLASS ACTIVITY:  
Use paper/pen to sketch  
“Tufte” version!

# “Chart Junk”

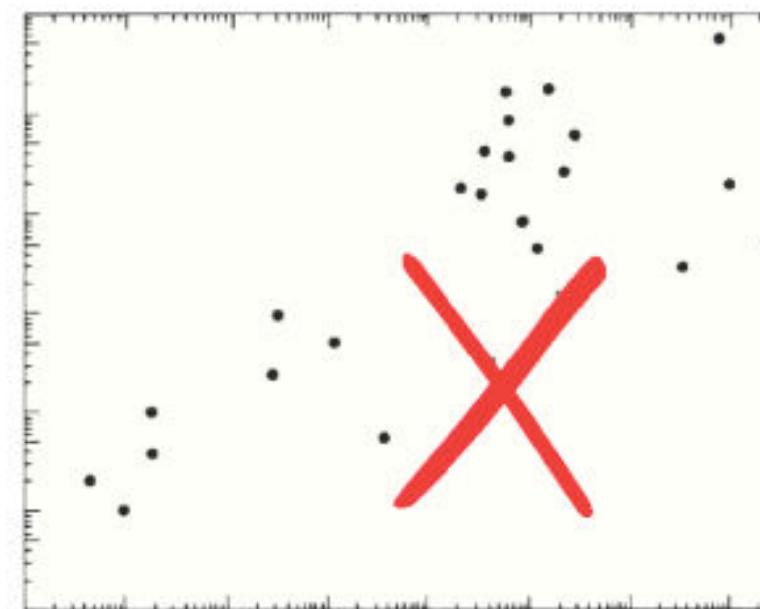




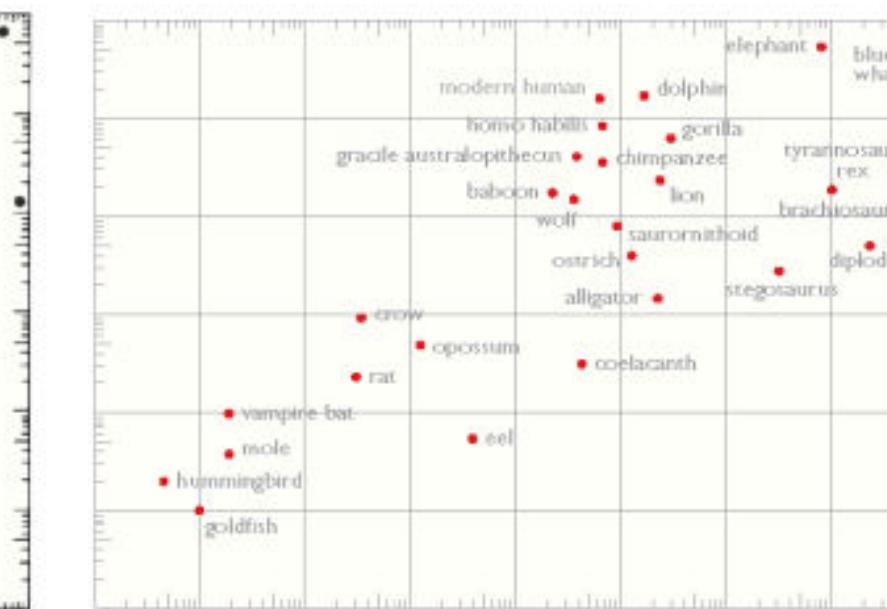
1



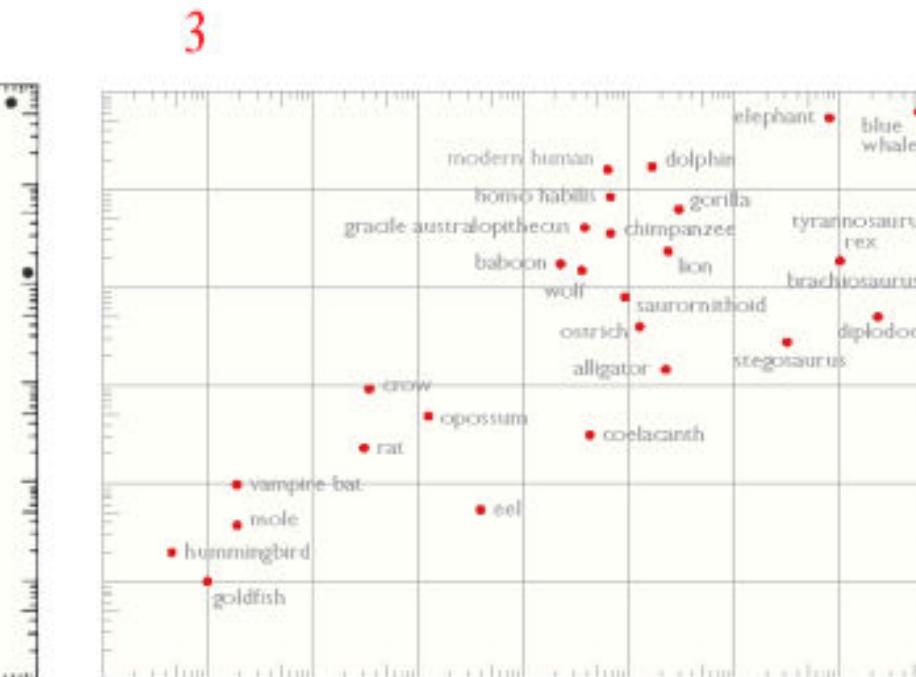
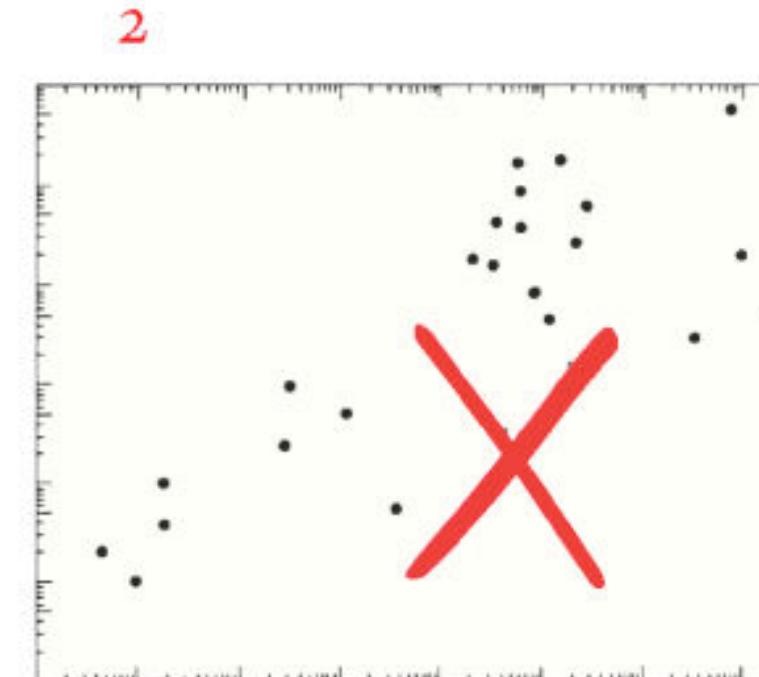
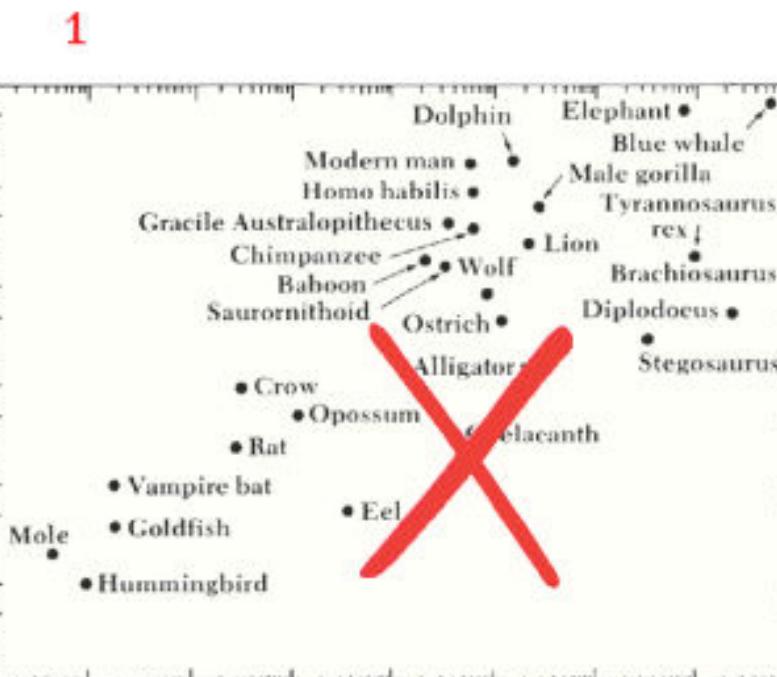
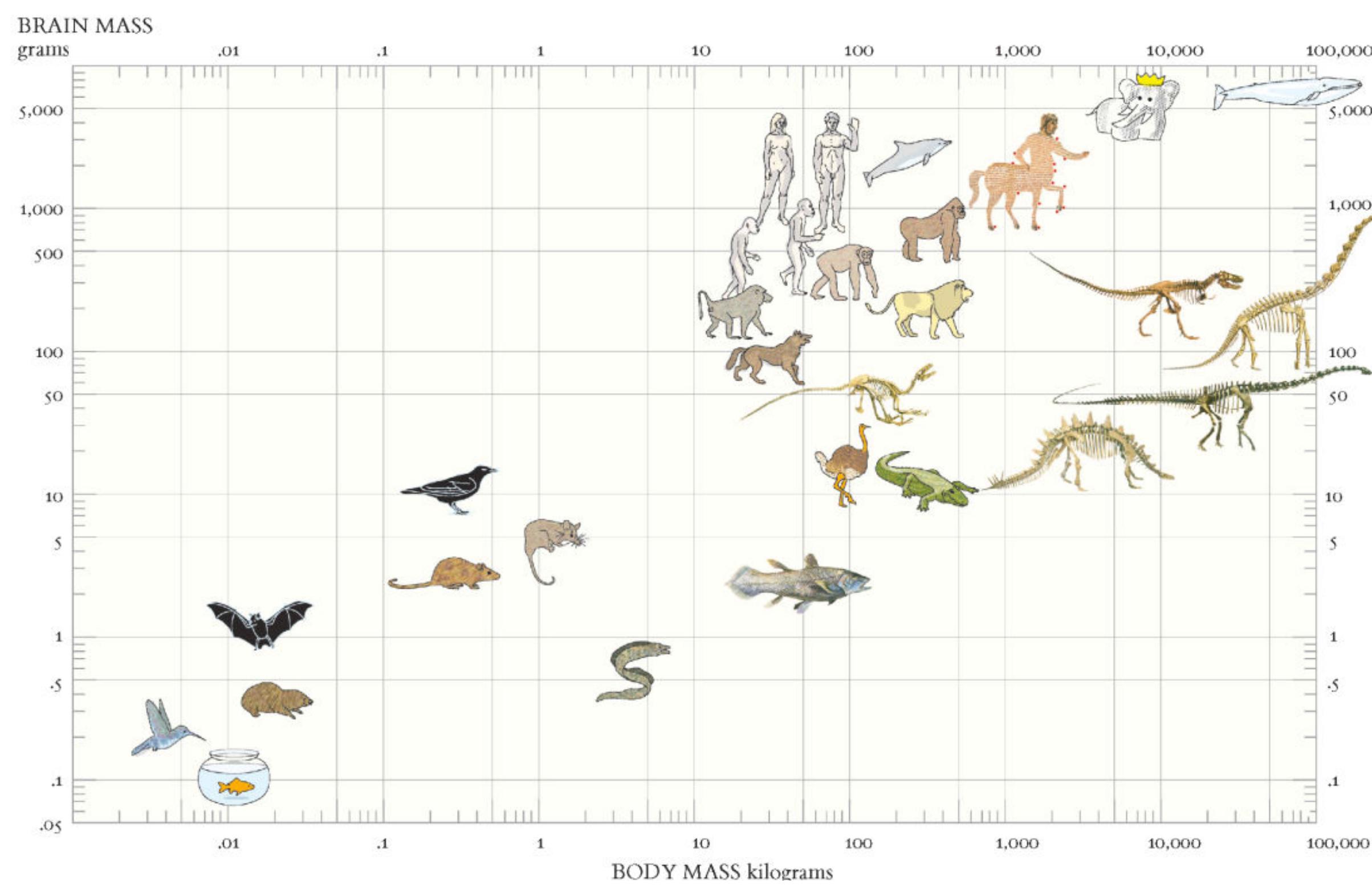
2



3



Not all “visual embellishments” are “chart junk”!



# “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...

# 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.