

# Design Rules of Thumb — Continued 2

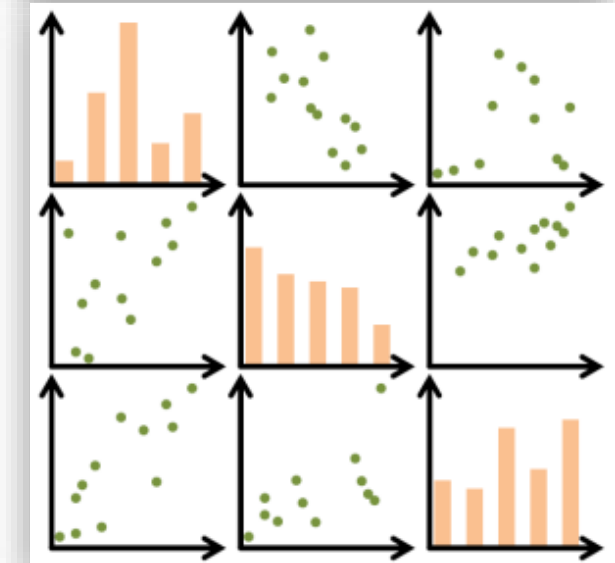
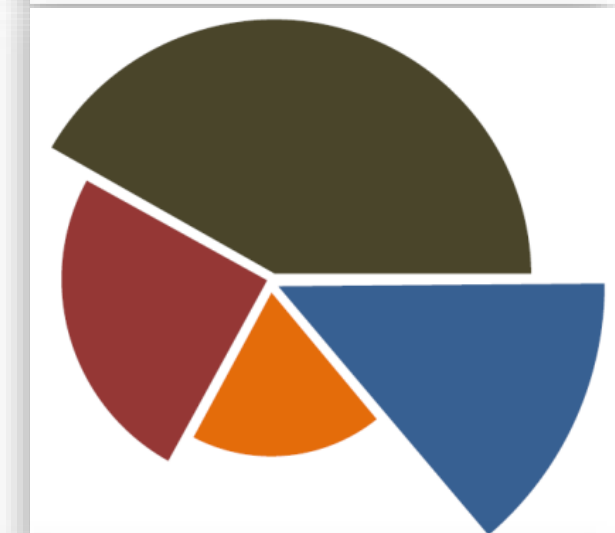
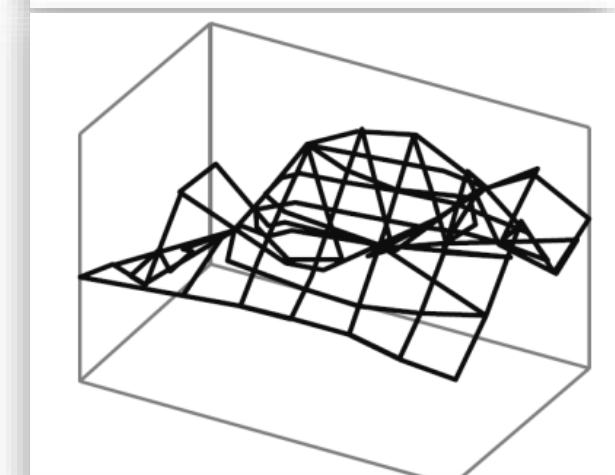
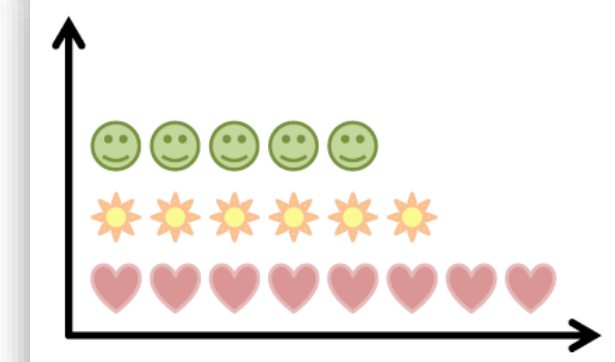
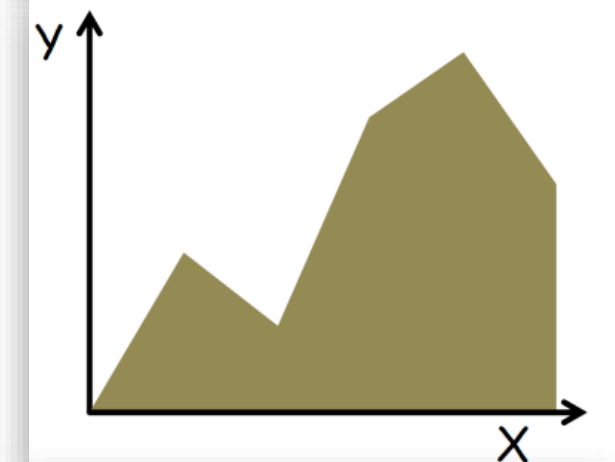
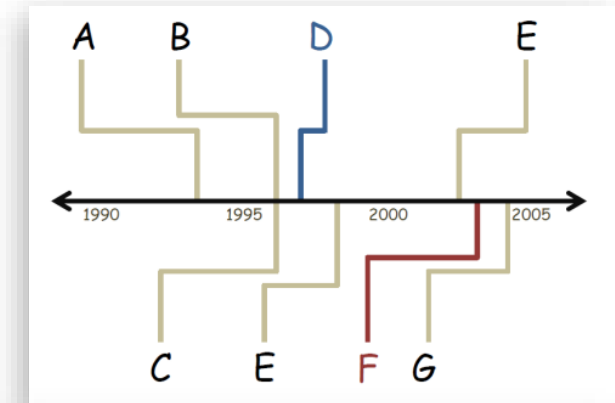
CS 7250

SPRING 2020

*Prof. Cody Dunne*

*NORTHEASTERN UNIVERSITY*

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



PREVIOUSLY, ON CS 7250...

JS, D3

**IN-CLASS PROGRAMMING —**

---

**D3 LINE CHART (PART 1)**

Now, ON CS 7250...

**IN-CLASS PROGRAMMING —**

---

**D3 LINE CHART (PART 2)**

PREVIOUSLY, ON CS 7250...

# “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.)*



# “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.”

Now, ON CS 7250...

# DESIGN & RULES OF THUMB

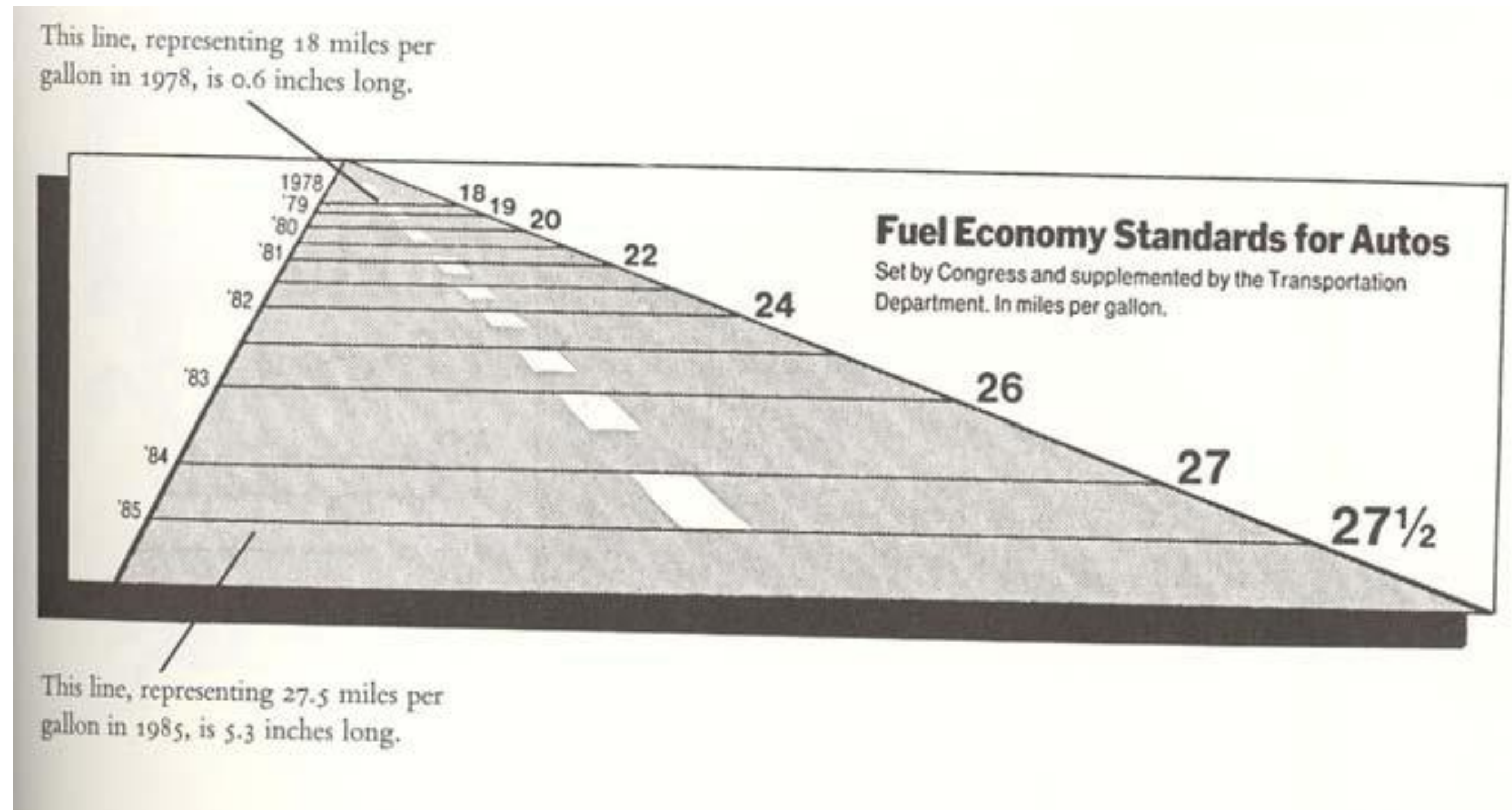
# Lie Factor

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

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.”

# Lie Factor

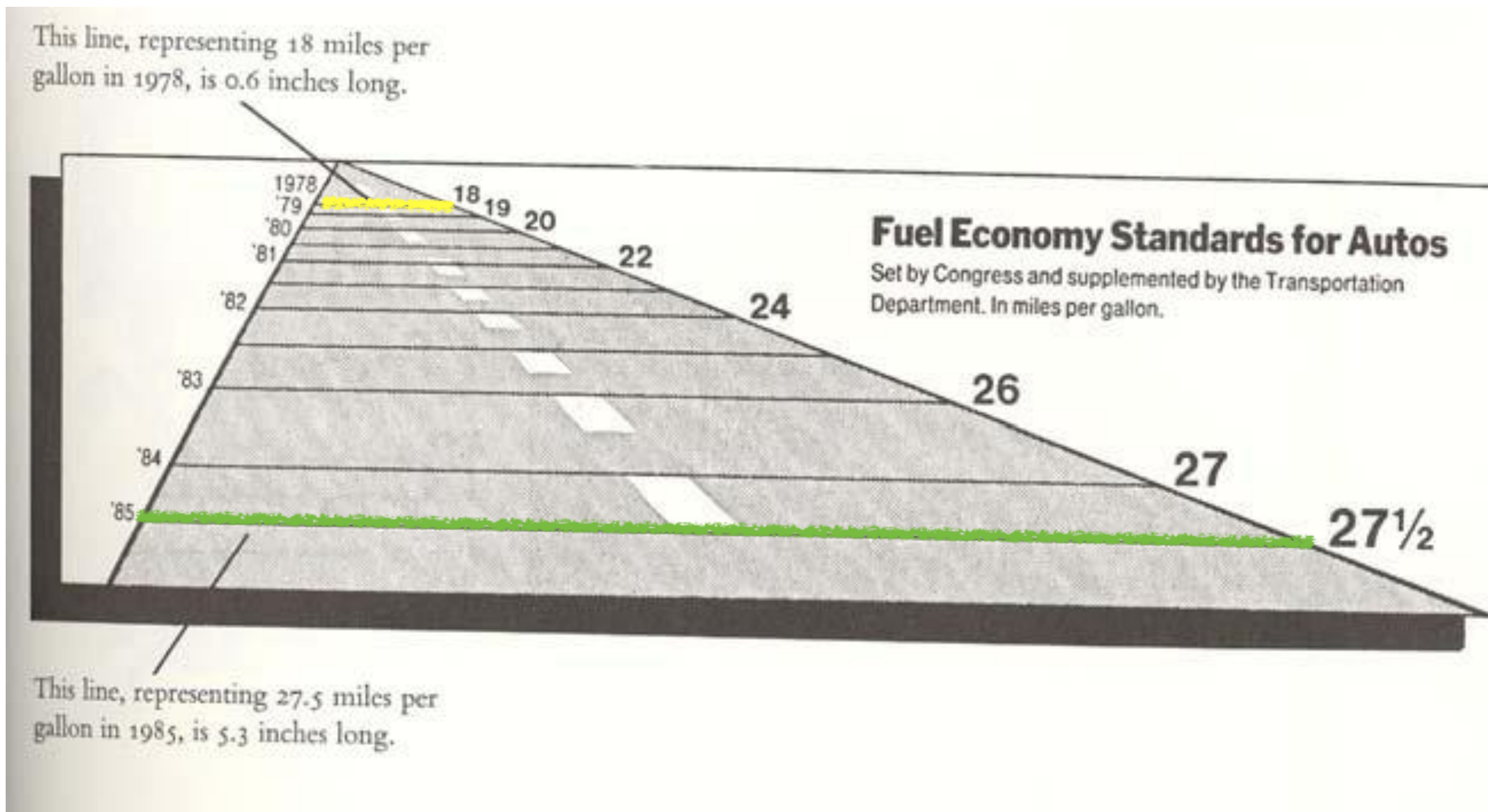
$$\text{Lie Factor} = \frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$$

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

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

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

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

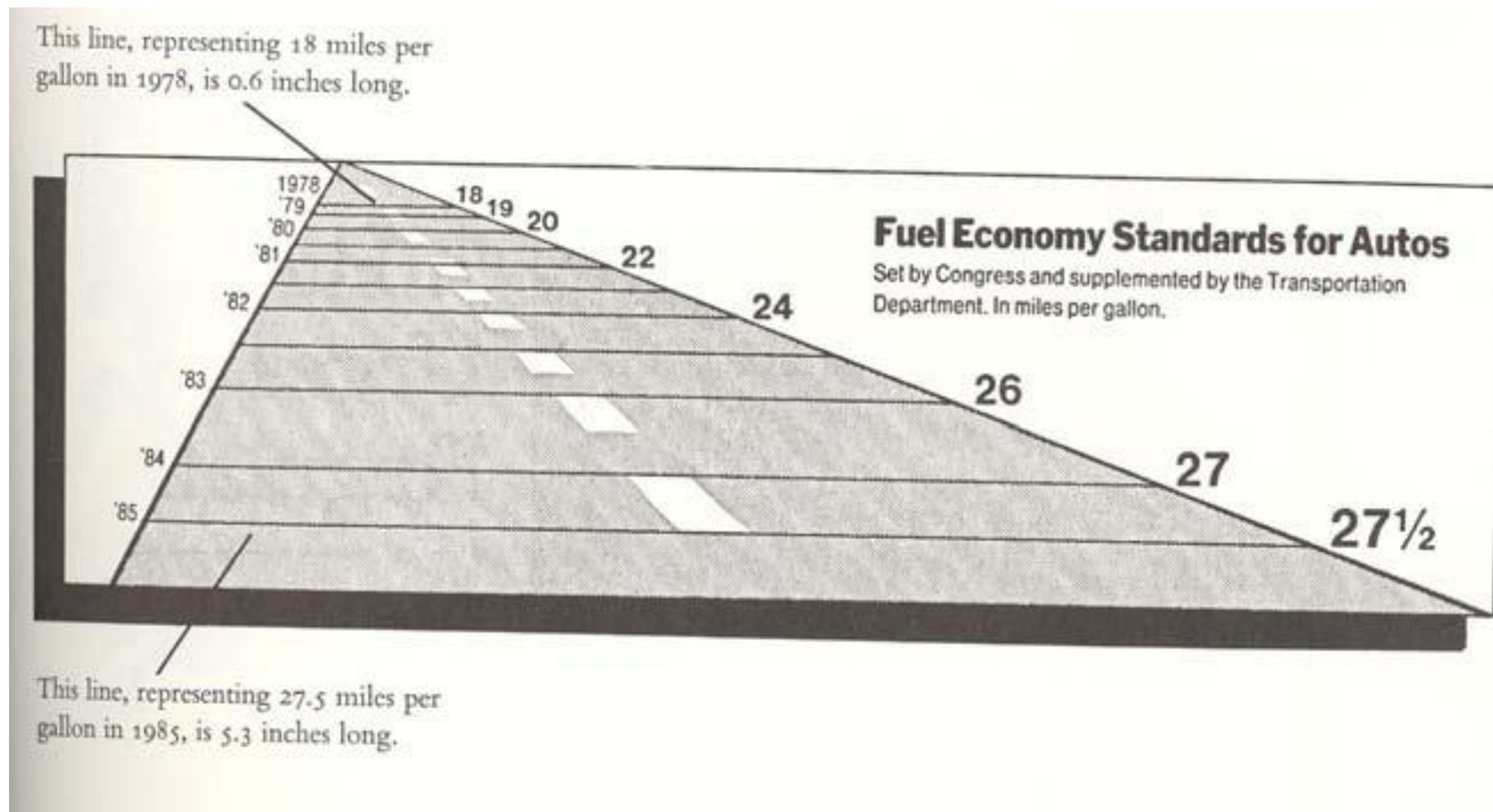
$$\text{Lie Factor} = \frac{\text{Size of effect in graphic}}{\text{Size of effect in data}}$$

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

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

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

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.”

IN-CLASS ACTIVITY:  
Calculate for yourself!

# Lie Factor

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

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

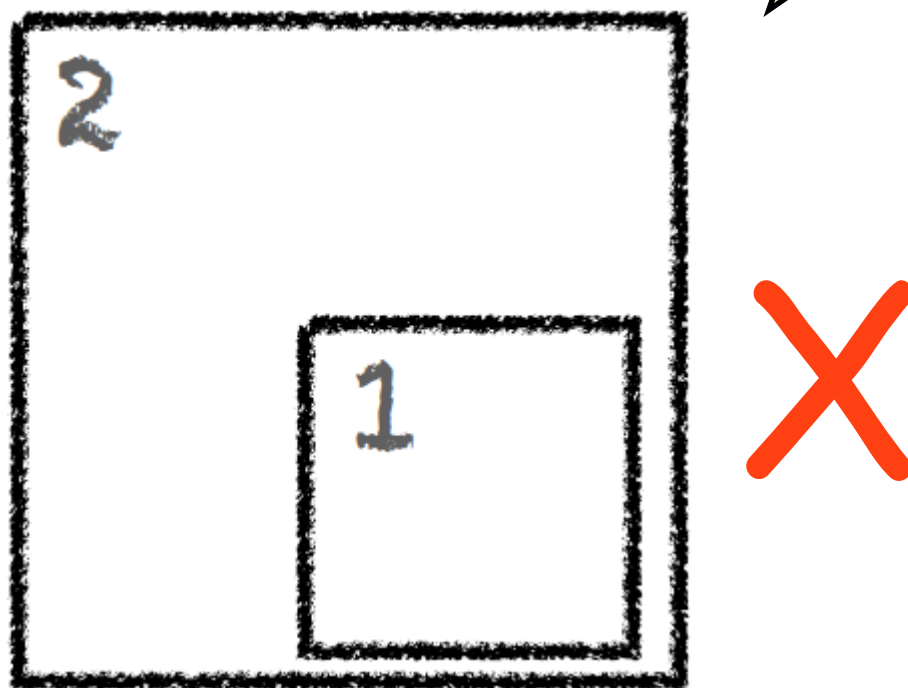
Make sure *area* is proportional to data!

Don't use 3D bar charts!



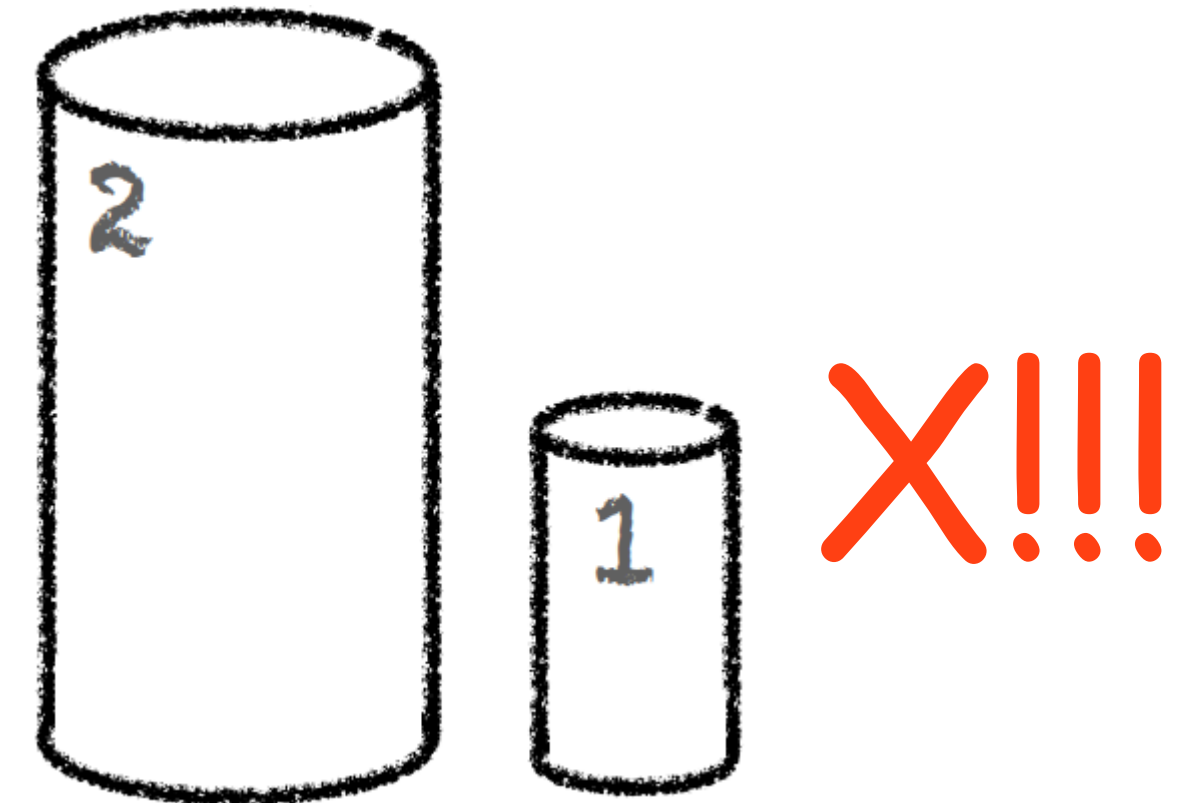
$$\text{Image} = \frac{2 - 1}{1} = 1 = 100\%$$

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



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

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



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

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

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

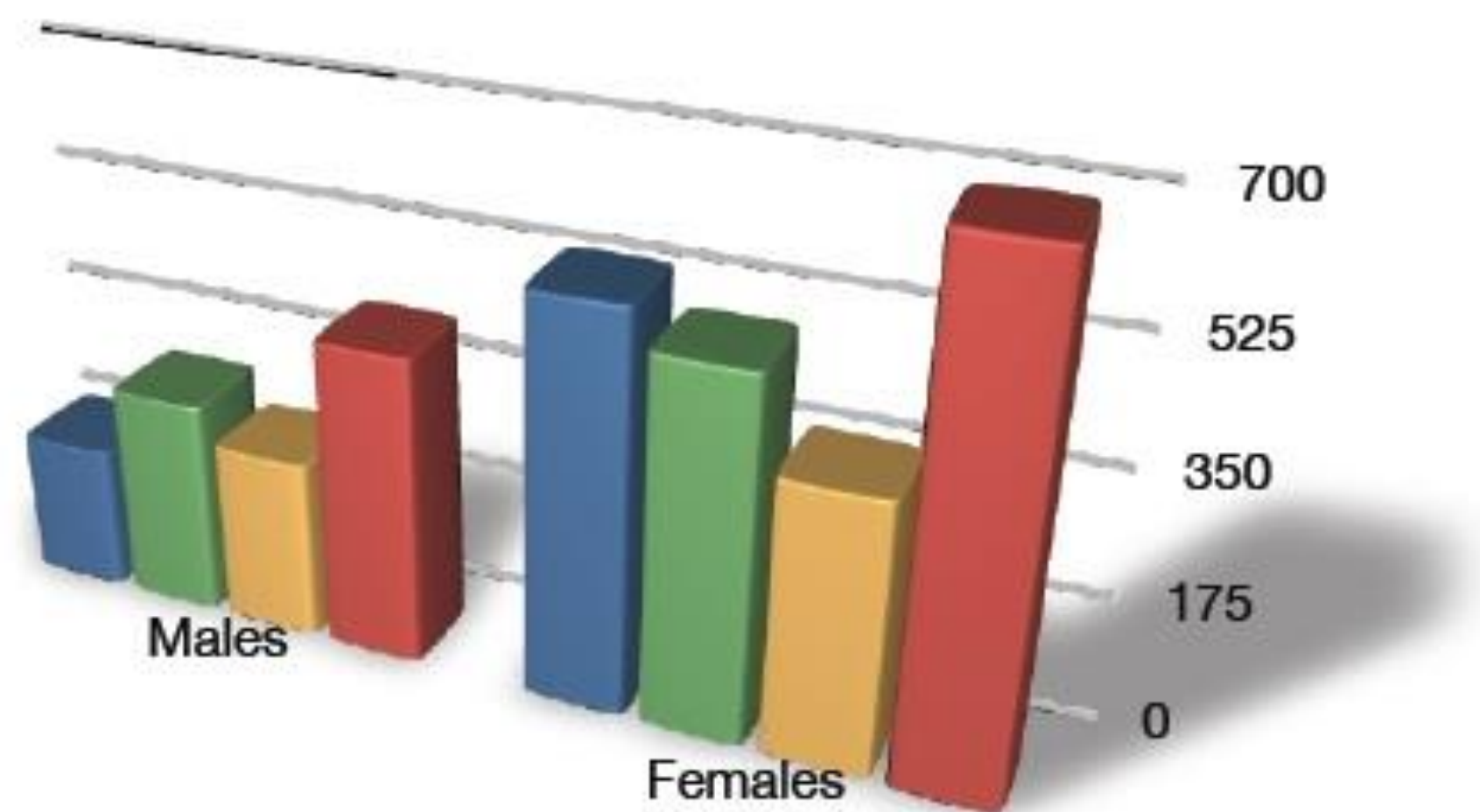
# “Graphical Integrity”

Data Ink = the ink used to show data

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

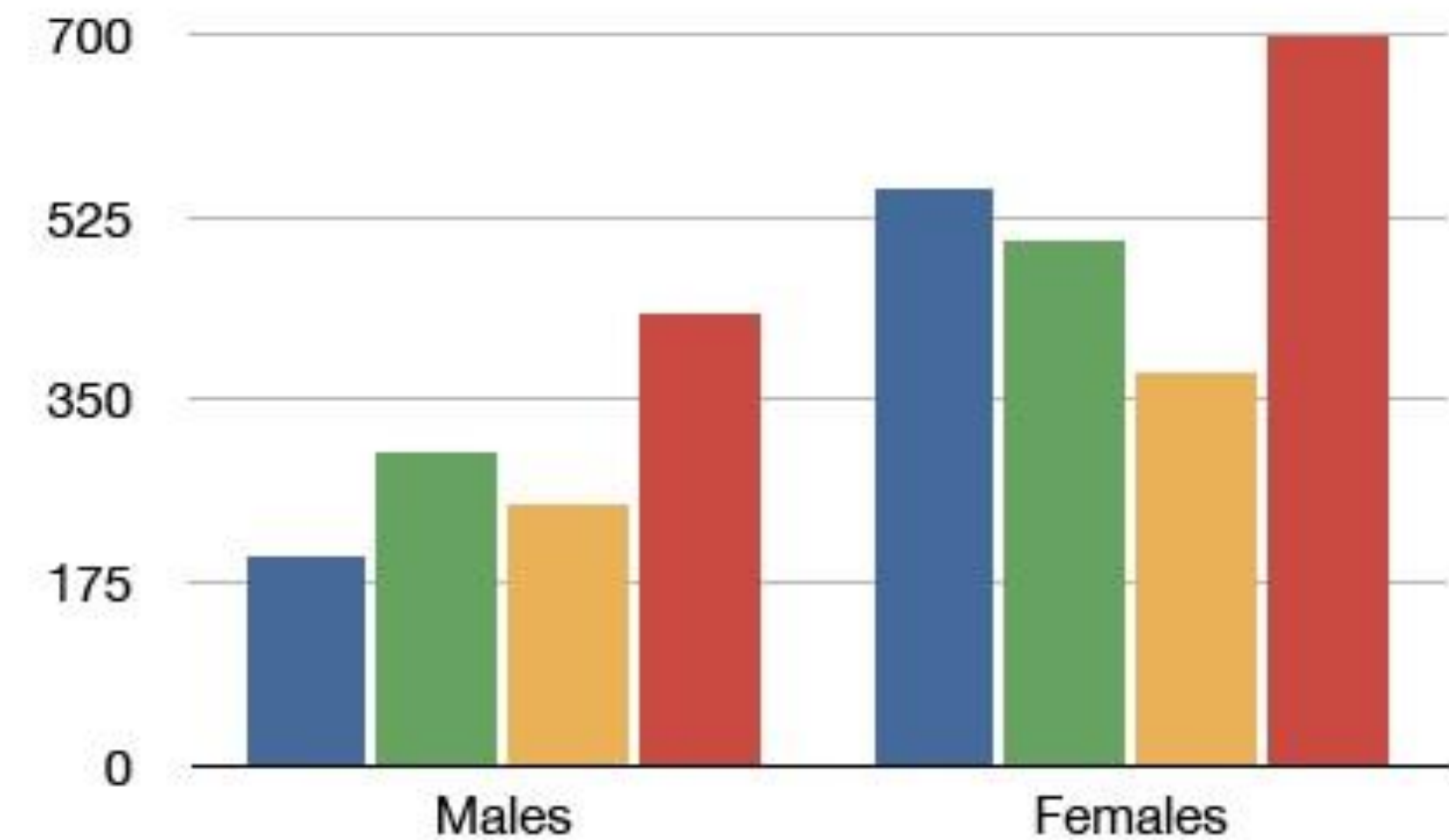
*Tufte: maximize the data ink ratio*

*Low Data Ink Ratio*



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

*High Data Ink Ratio*



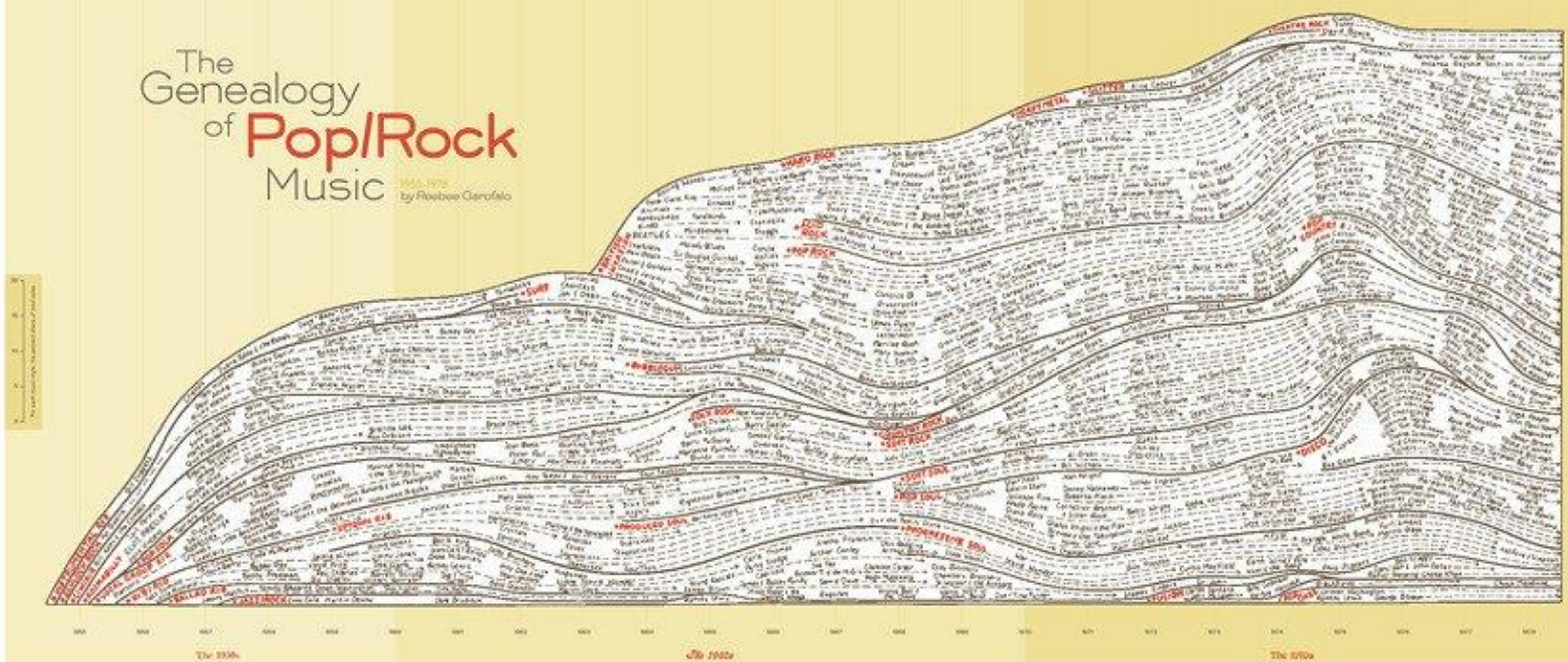
■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+

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



# The Genealogy of Pop/Rock Music 1955-1975

by Reebee Garofalo



This graphic chronicles the growth and development of pop, rock, and soul music and their top-selling artists.

More than 700 artists and 30 styles of music are mapped in flowing patterns in this chart. It includes almost every performer who sold records in the pop/rock market and whose names appeared on the year-end top 50 when it begins there.

An arrow starting from a performer's name shows the length of time that he/she remained a major hit maker. The width of each industry category is a rough estimate of its share of the market according to the scale on the left.

Categorizing music is clearly a subjective process and in some cases it may seem appear to be arbitrary. The performers shown here represent only the tip of the "commercial" bubble — the major hit makers of a given year. Their music is the most accessible, it may or may not be the most influential or artistically important. Each has the approval of the marketplace in a commercially system.

Copyright © 1995, Reebee Garofalo  
 Revised and updated from the book *Rock: A Social History: The History and Politics of the Music Industry* by David Gaughran and Reebee Garofalo  
 Original text © 1993, David Gaughran and Reebee Garofalo  
 Special thanks to Neil Young, Bill Graham, Bruce Springsteen, Henry Diltz, Billy Bragg, John Mellencamp, and Bruce Springsteen.  
 Thanks to William Gramling for the Sound of the City Photo and Book Services, and to the Rock Hall.



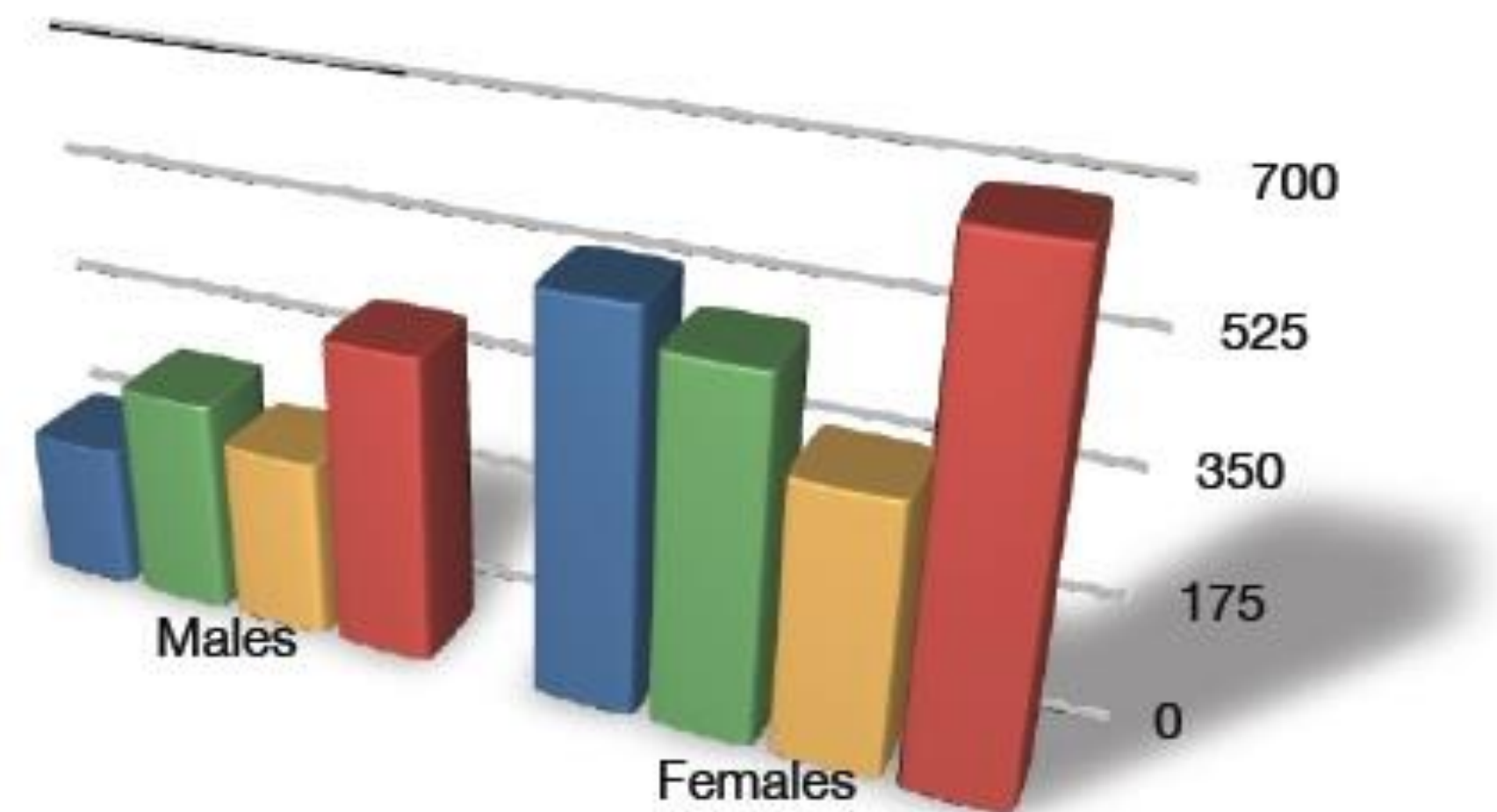
# “Graphical Integrity”

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

# “No Unjustified 3D”

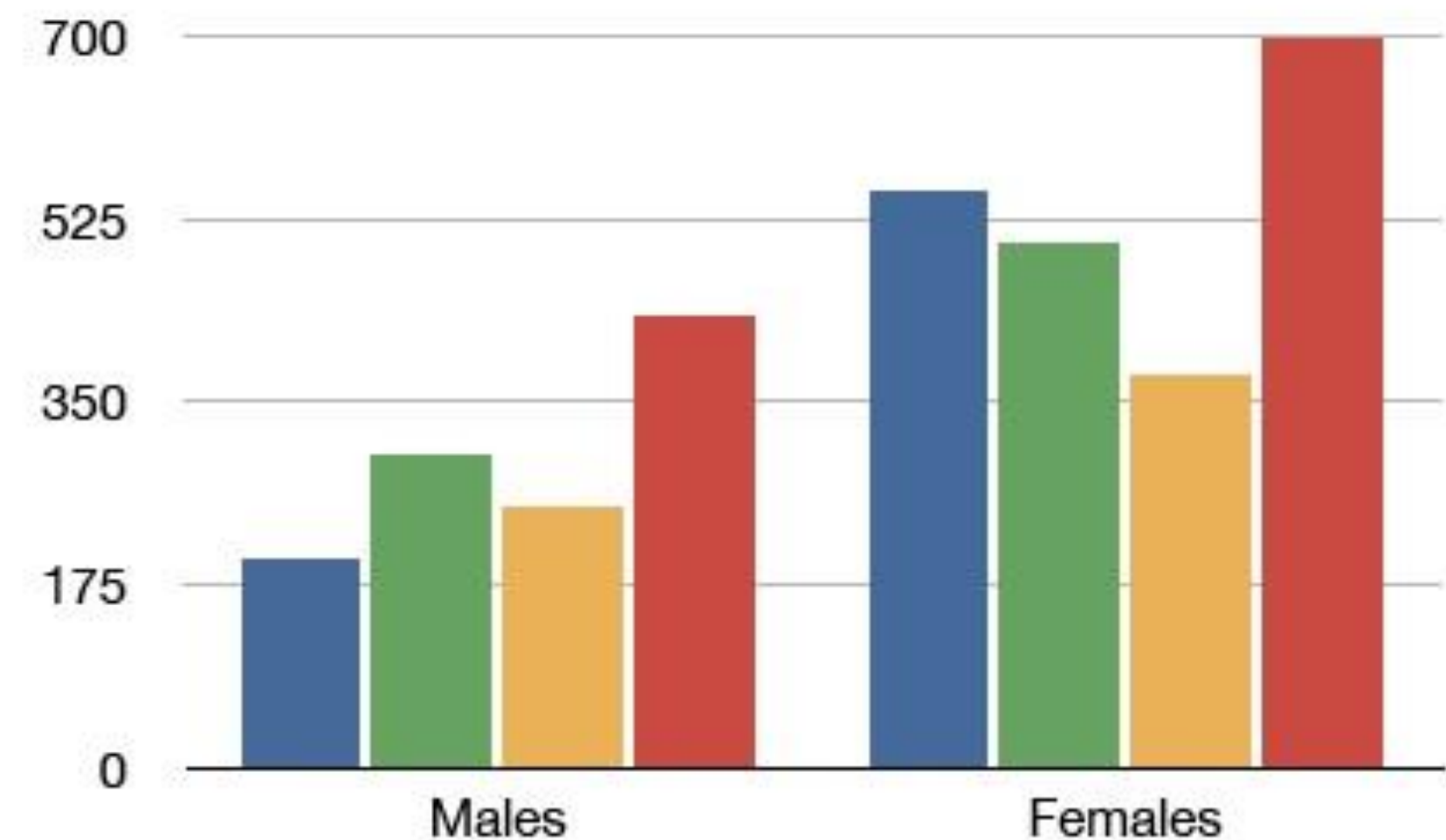
# Dimensions in data: 3

# Dimensions in plot: 4



# Dimensions in data: 3

# Dimensions in plot: 3

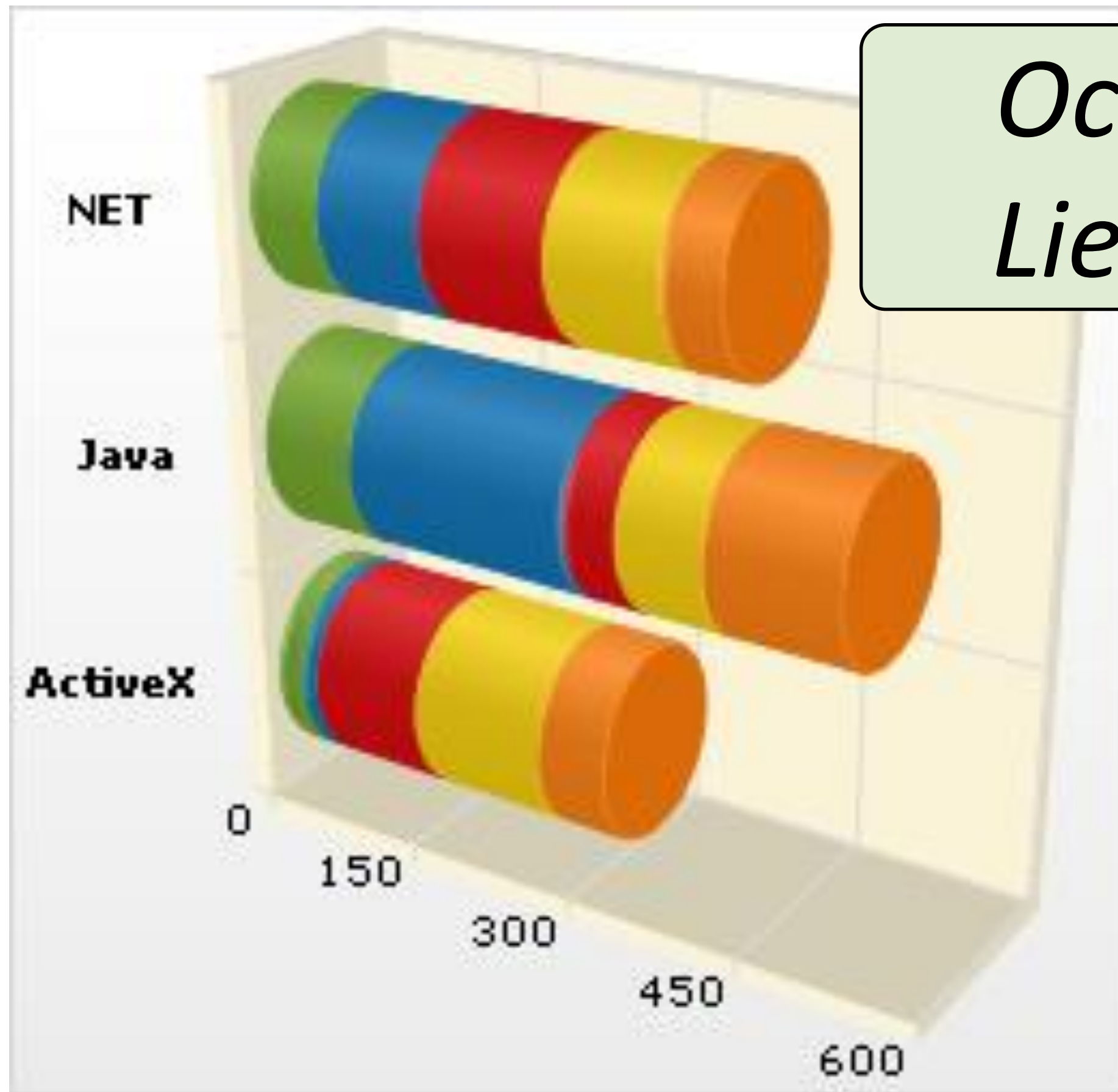


■ 0-\$24,999 ■ \$25,000+ ■ 0-\$24,999 ■ \$25,000+

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

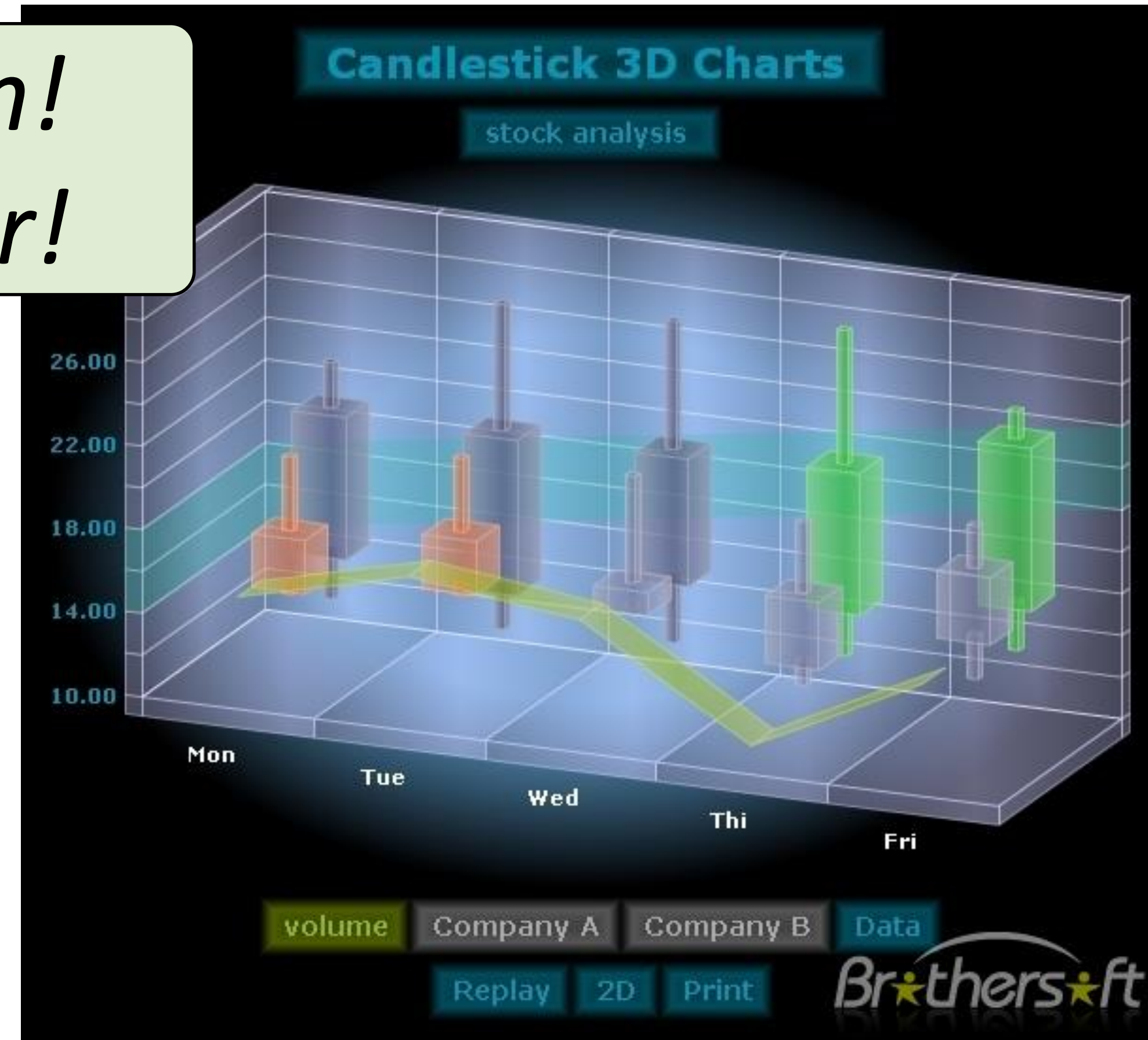
“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

# “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://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](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.”

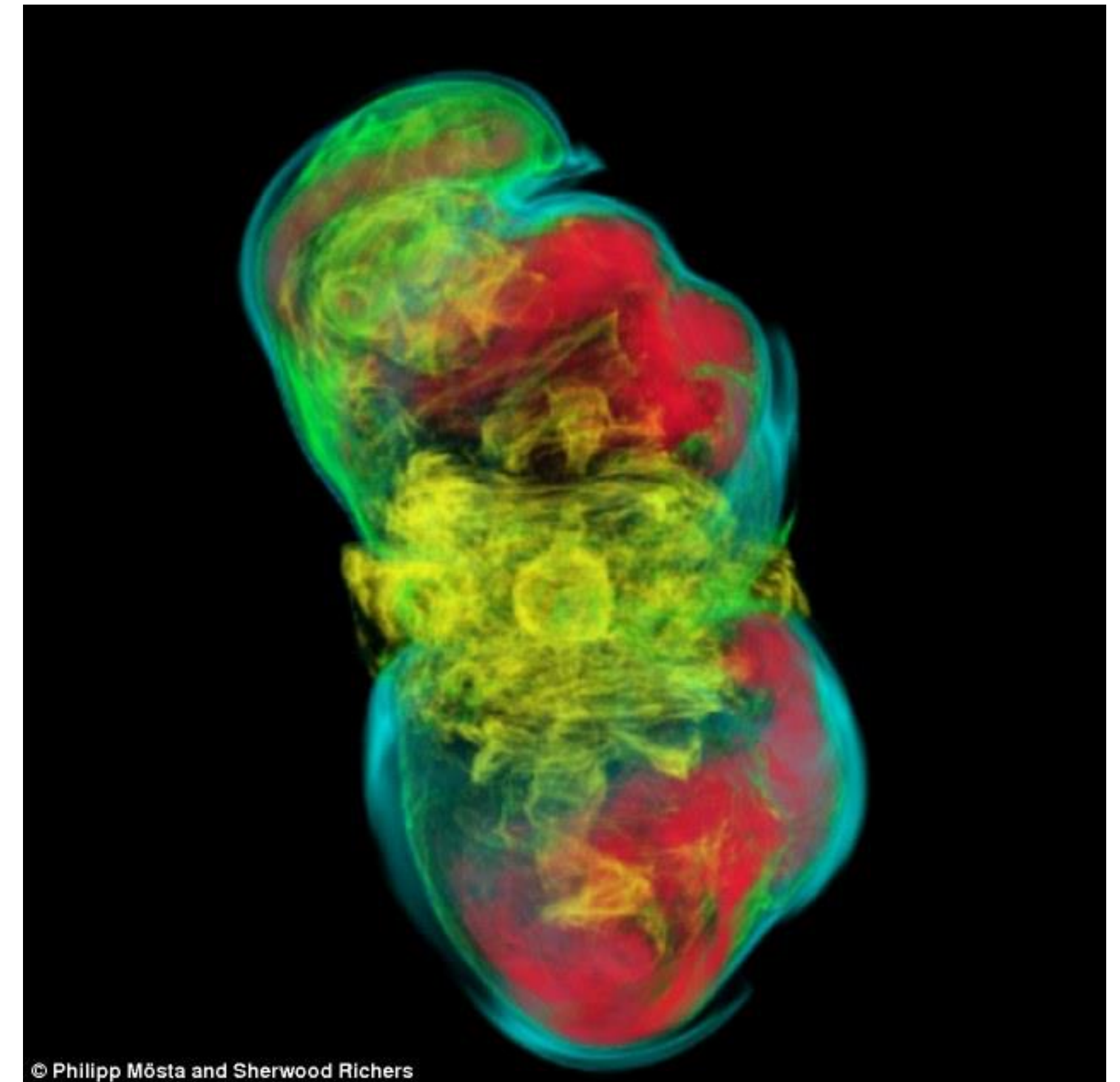
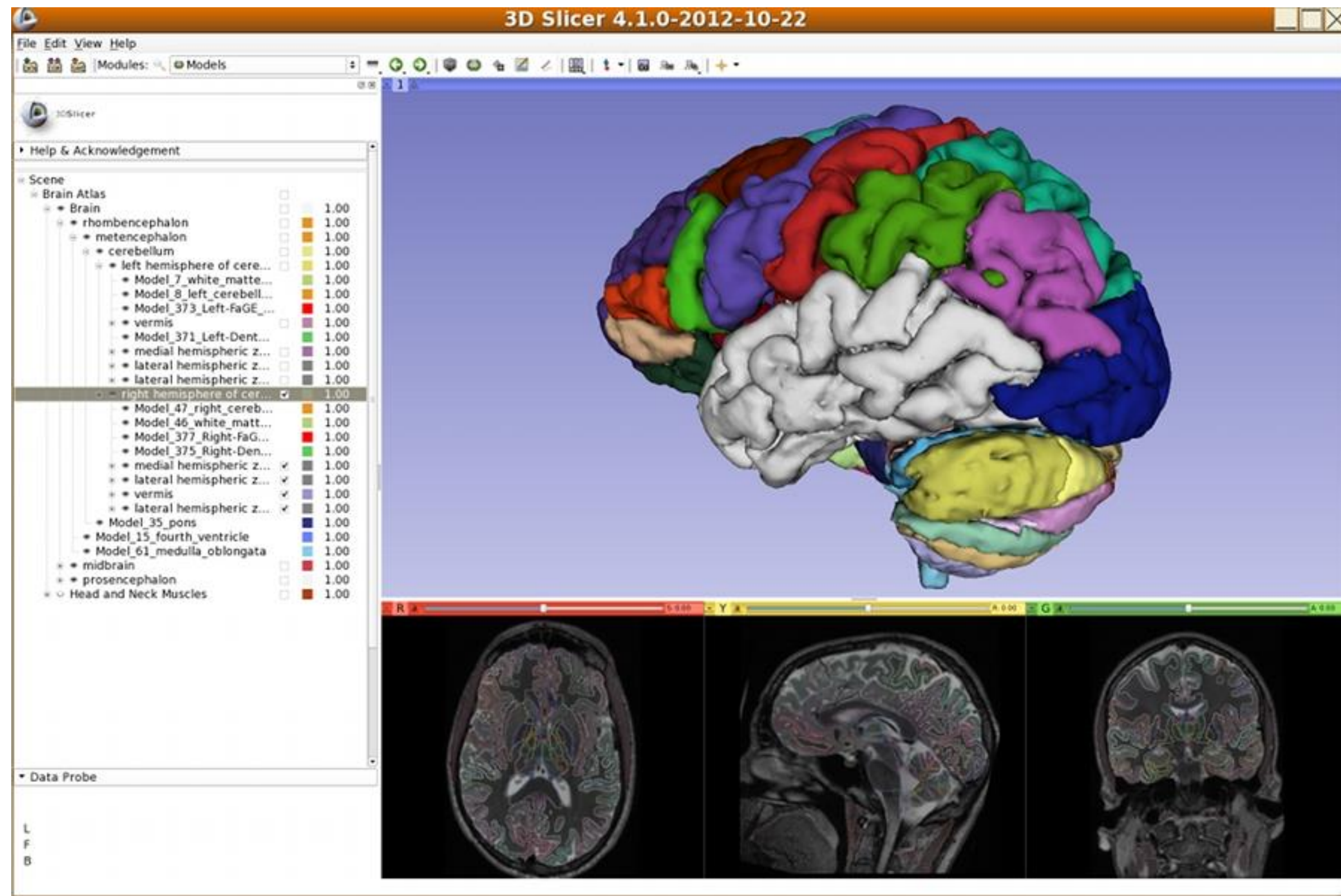
# “No Unjustified 3D”



Unjustified 3D!

Lie factor!

# “No Unjustified 3D”



“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.”

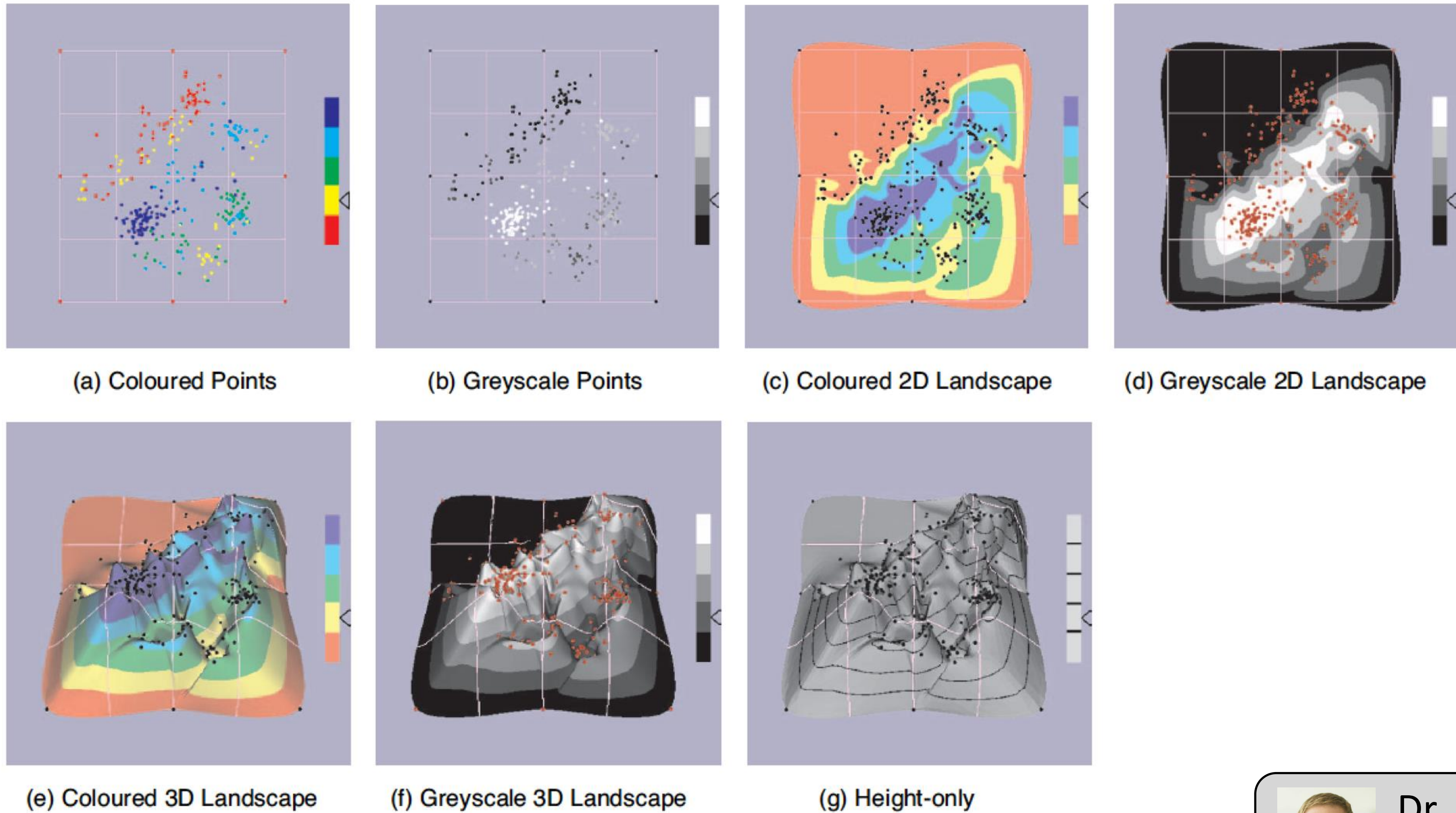
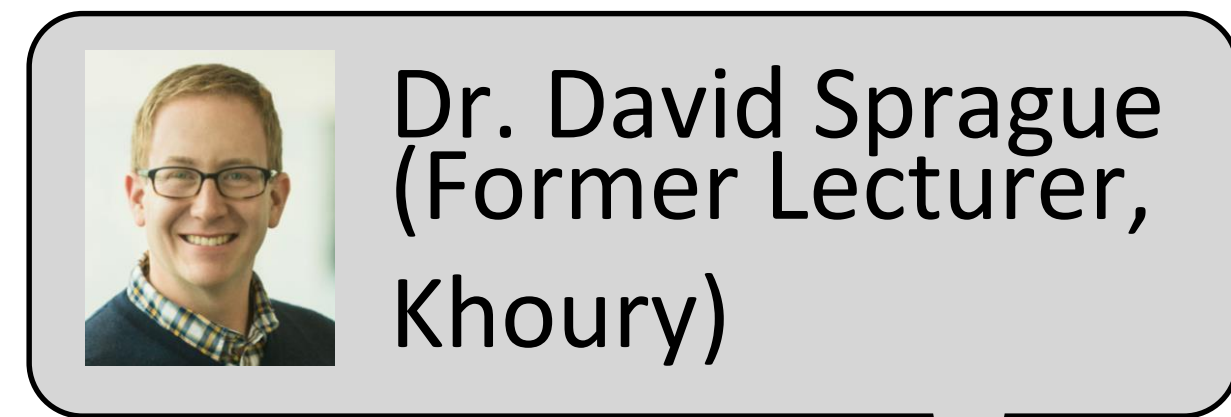
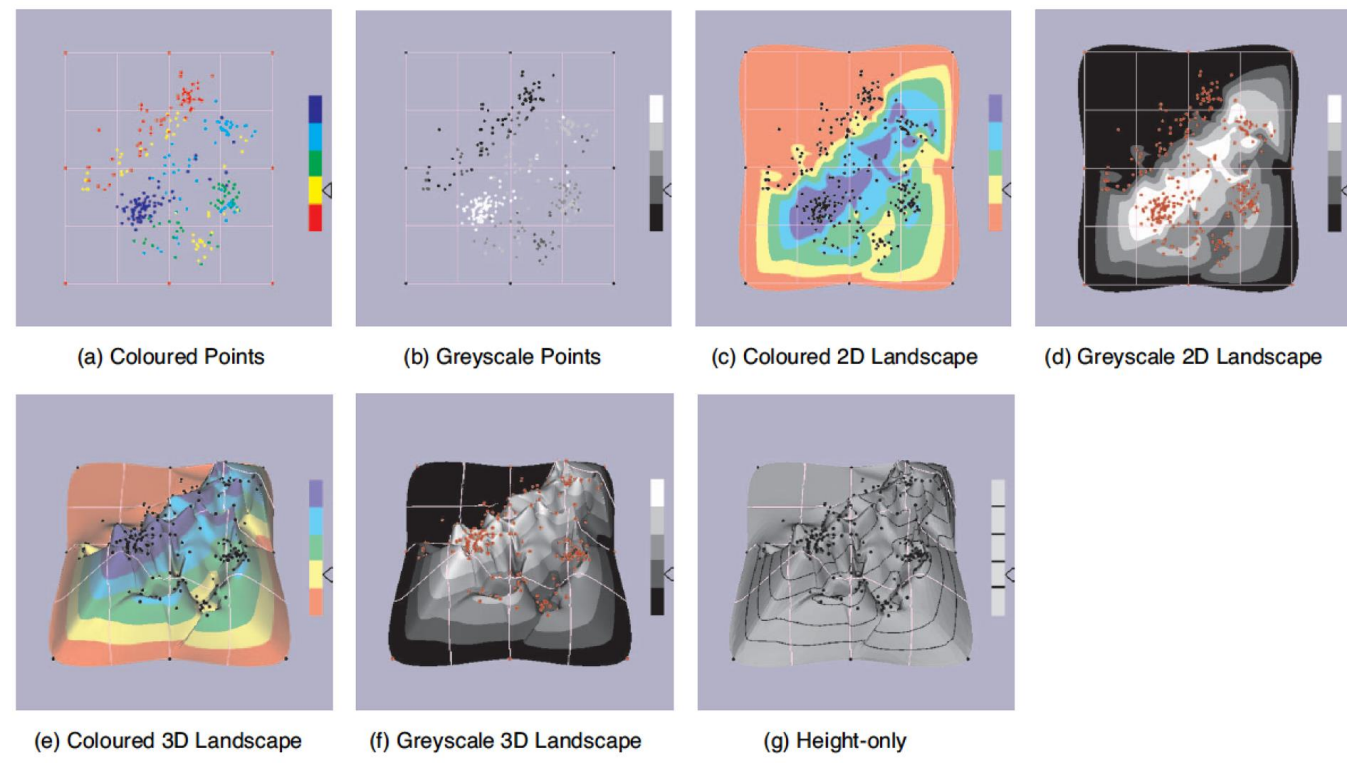


Fig. 1 Point-based displays and information landscapes used in our experiment. All displays show the same data.



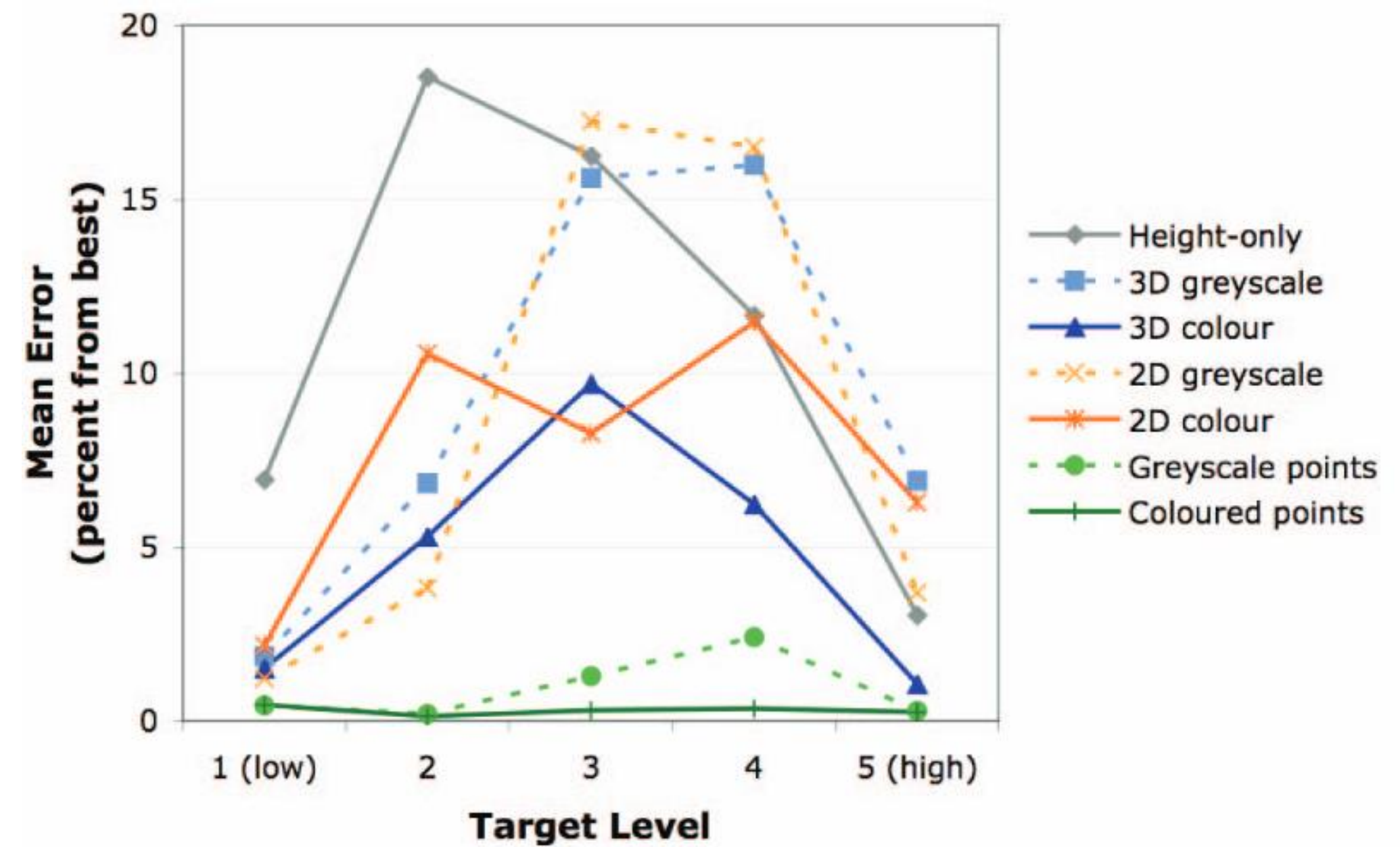
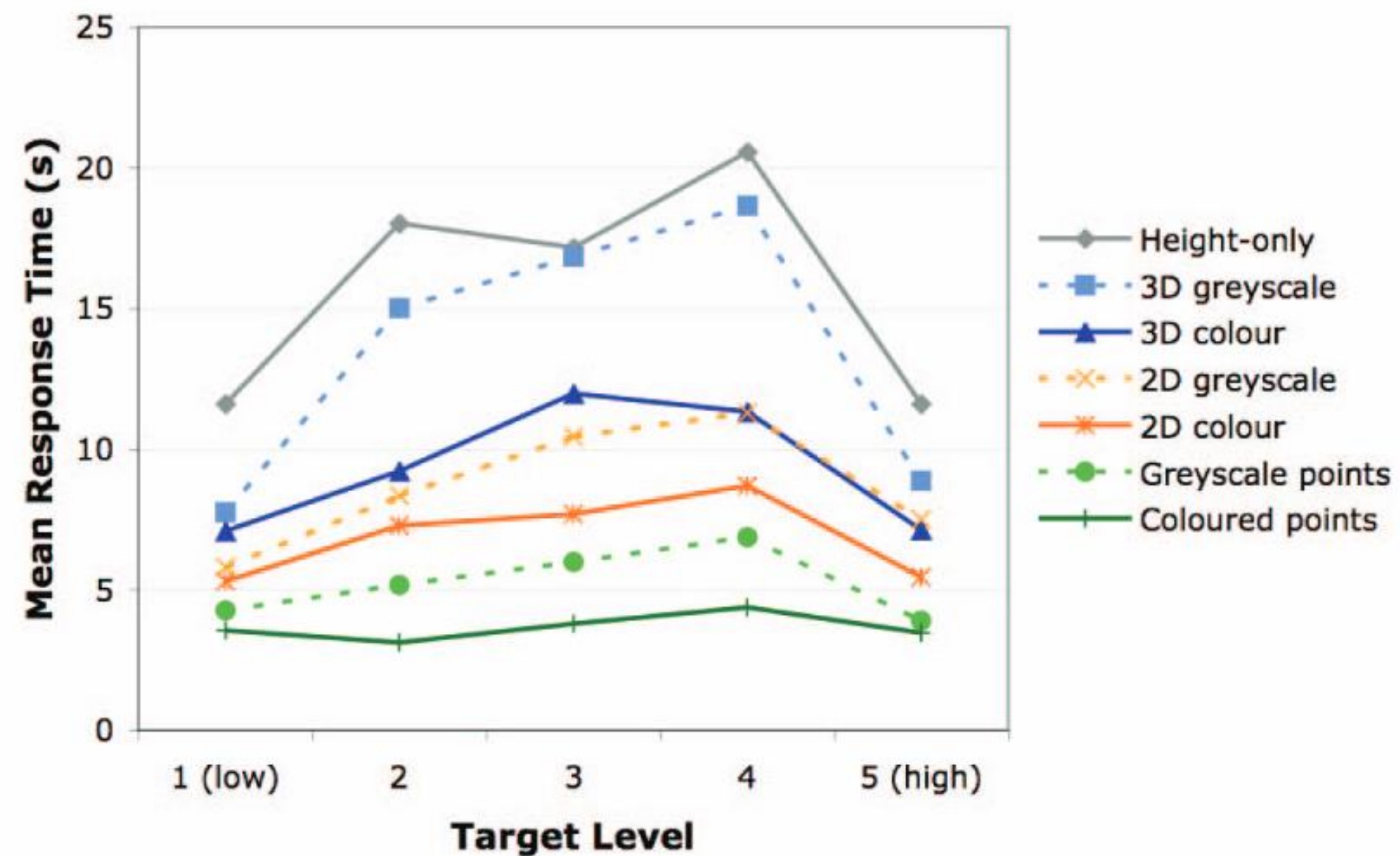
“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”





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

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.”

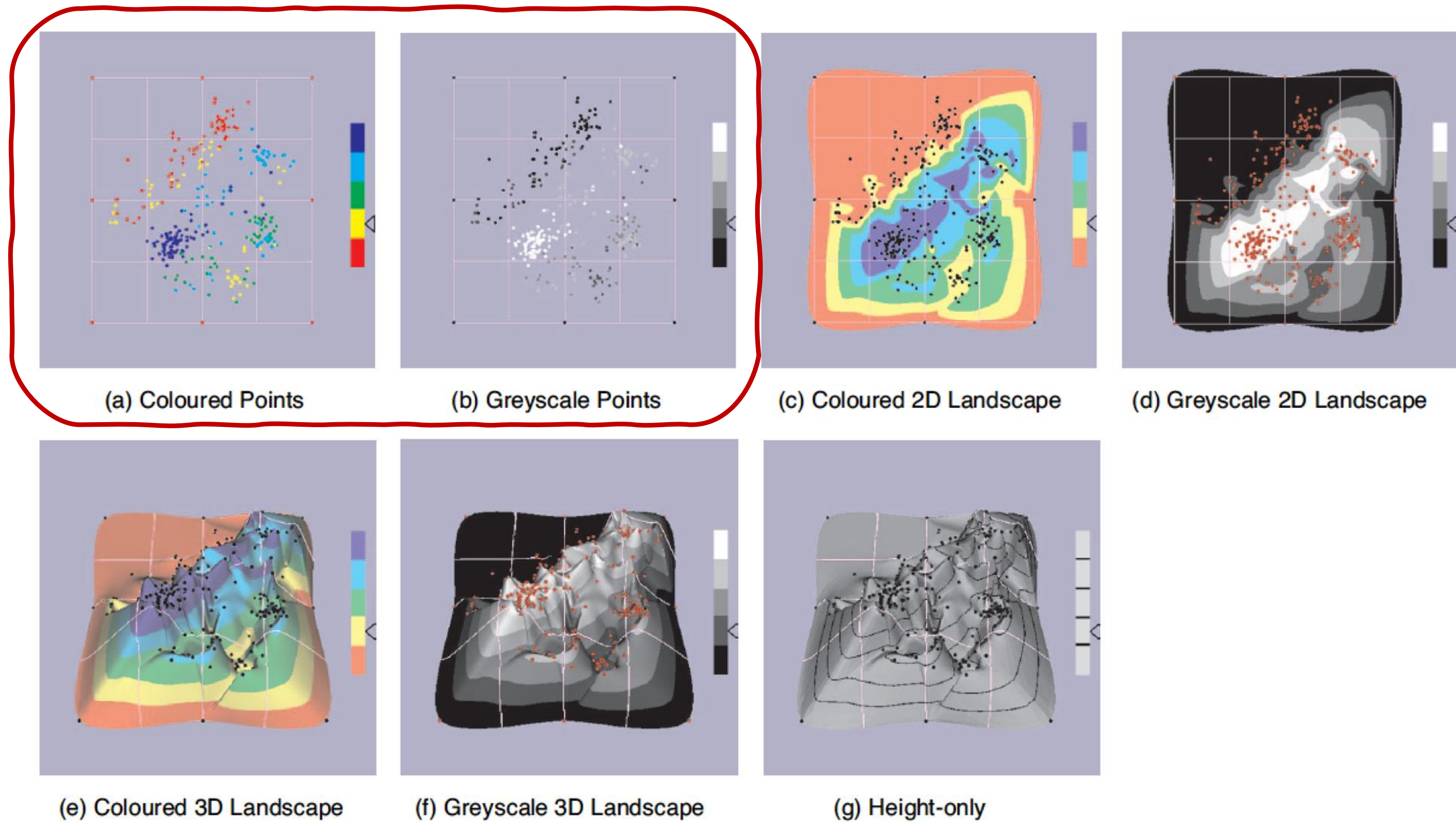
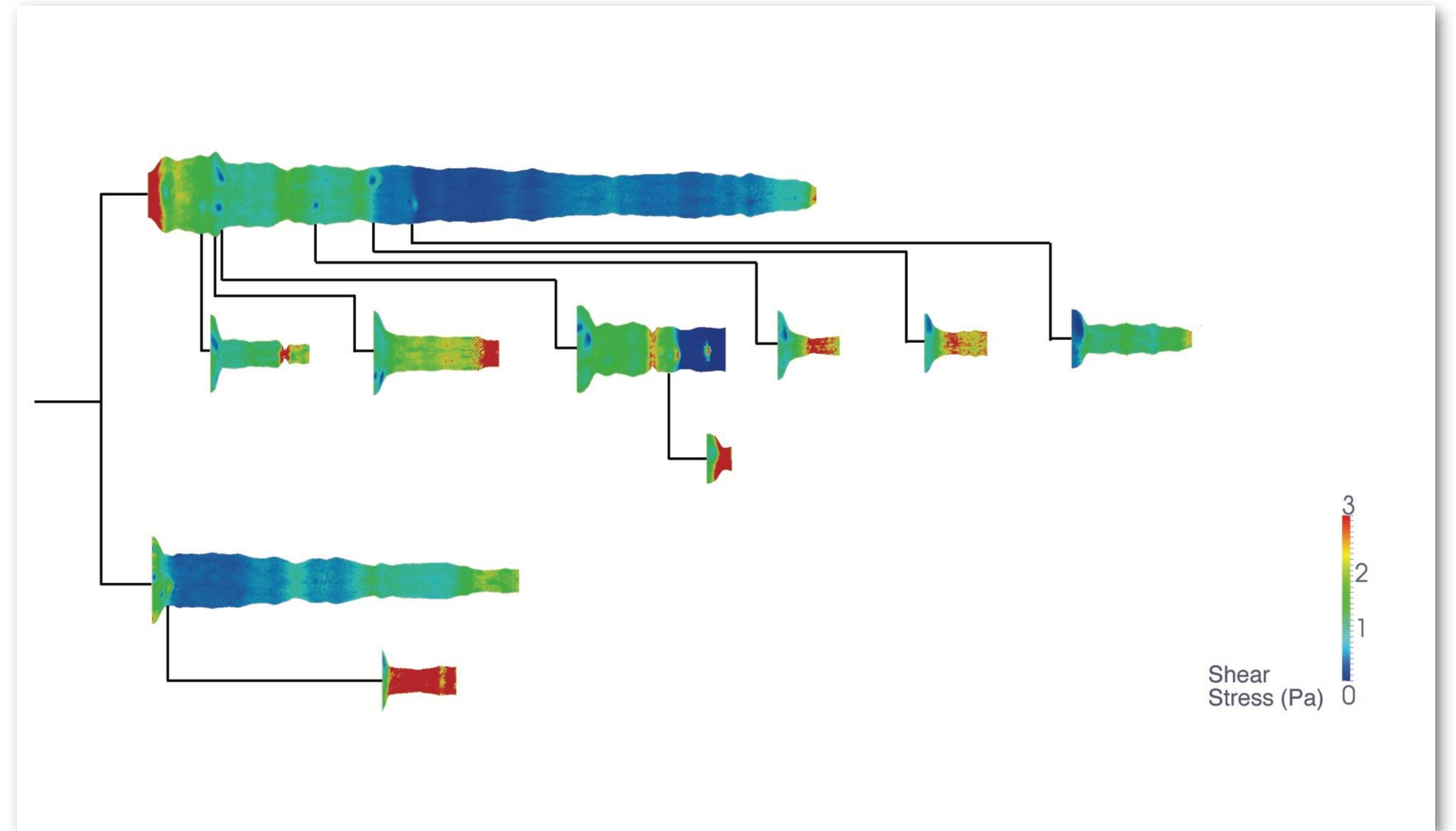
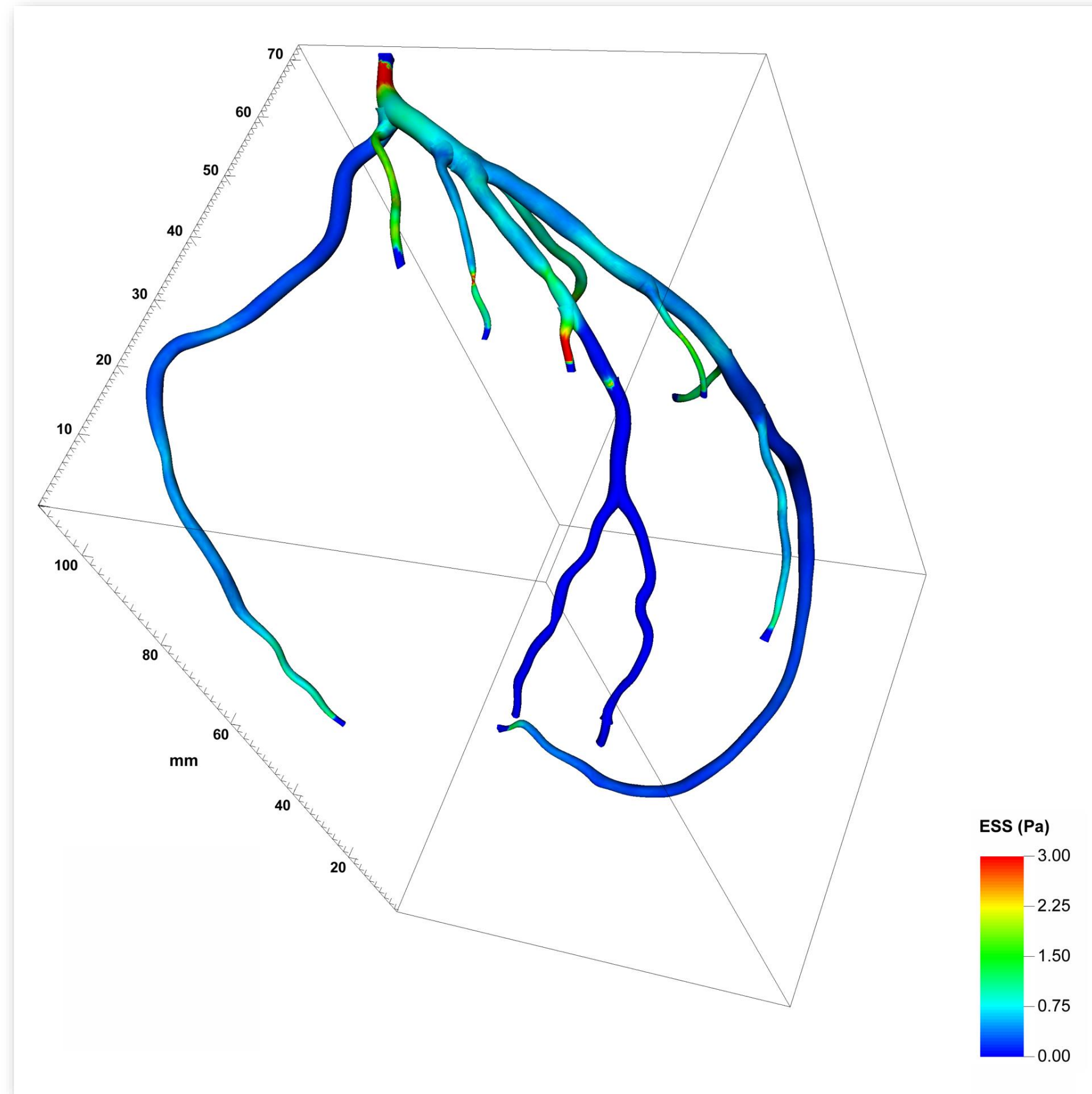


Fig. 1 Point-based displays and information landscapes used in our experiment. All displays show the same data.

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

# “No Unjustified 3D”



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

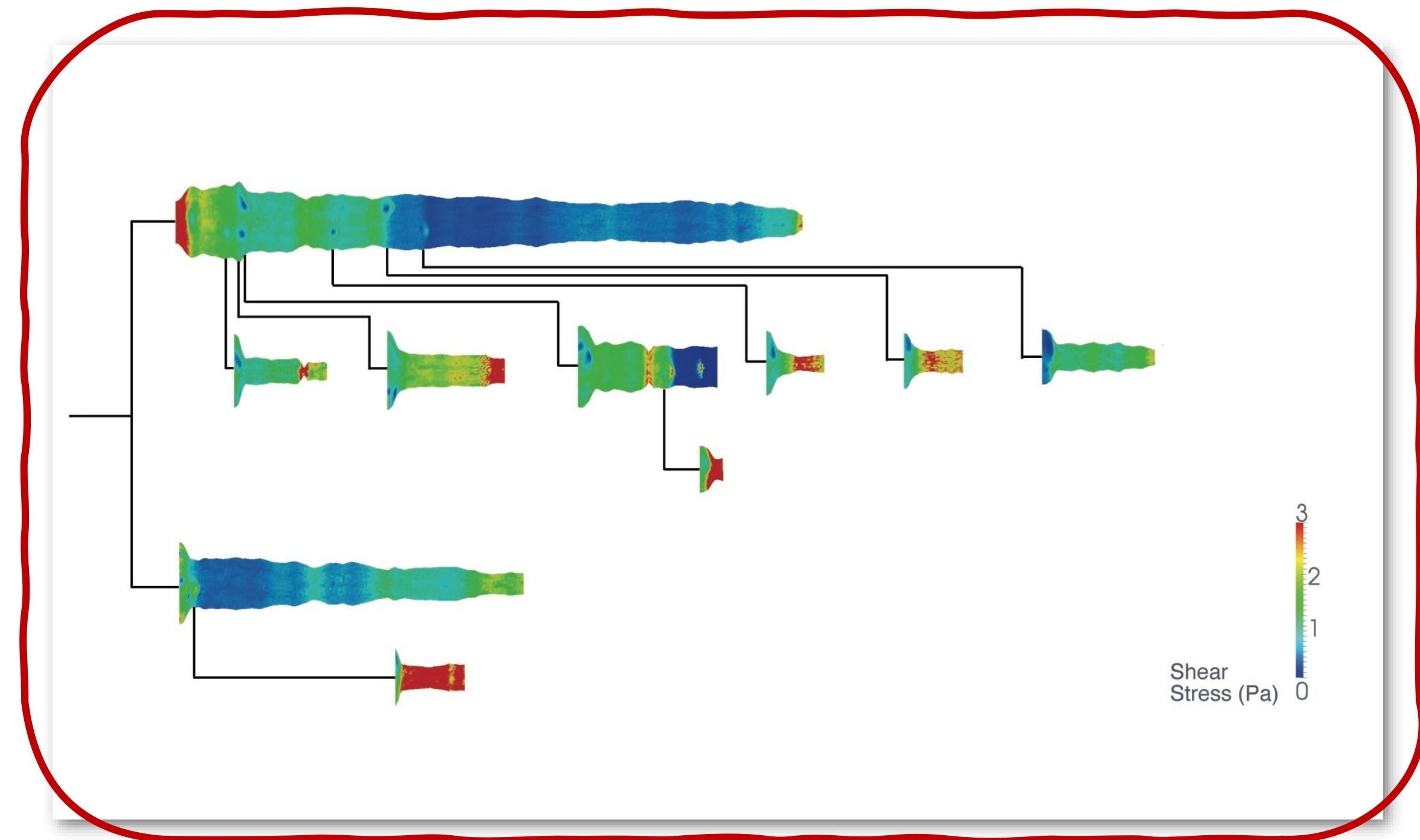
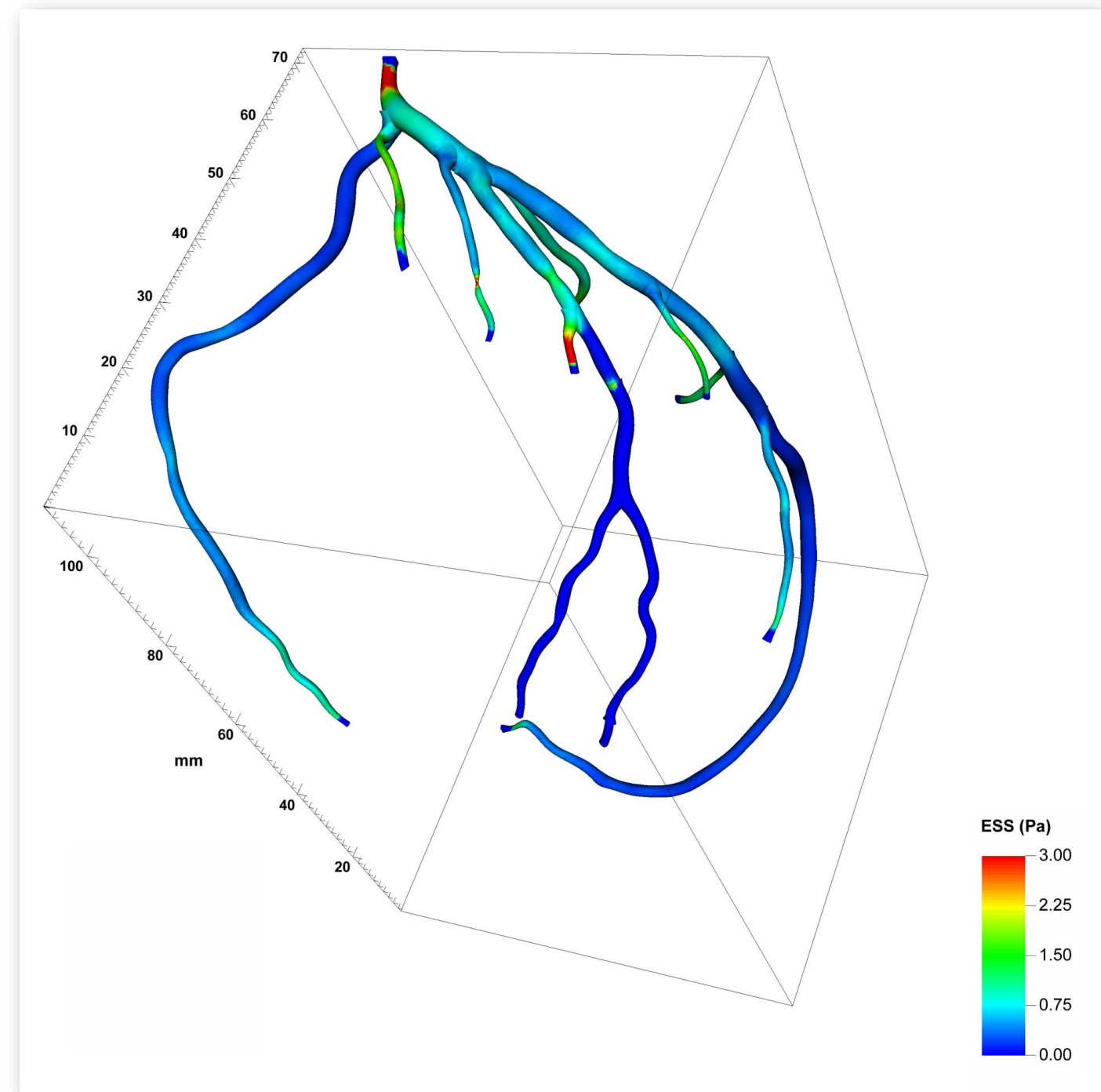
# ACCURACY

Strong effect of **dimensionality** on accuracy

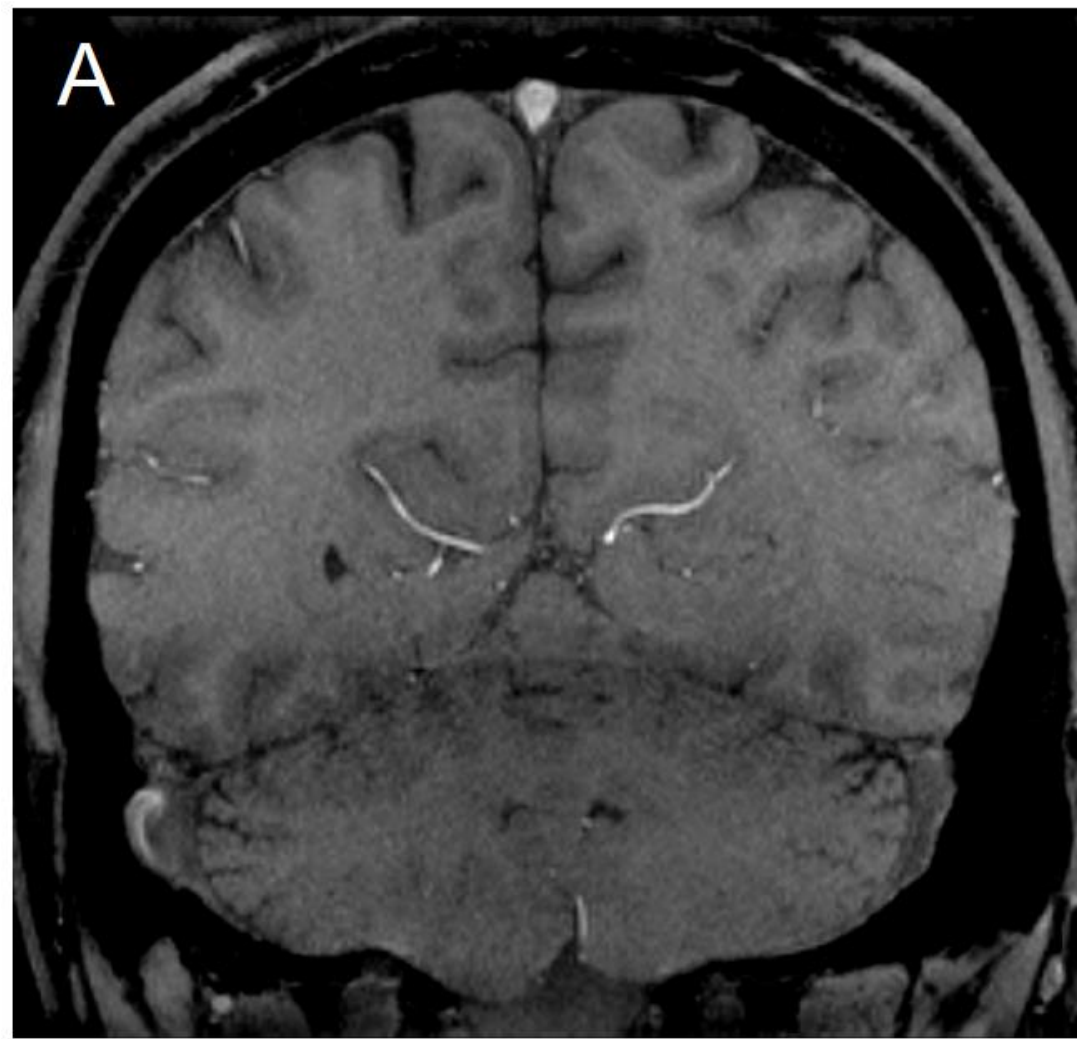
**39%**

How many diseased  
regions found?

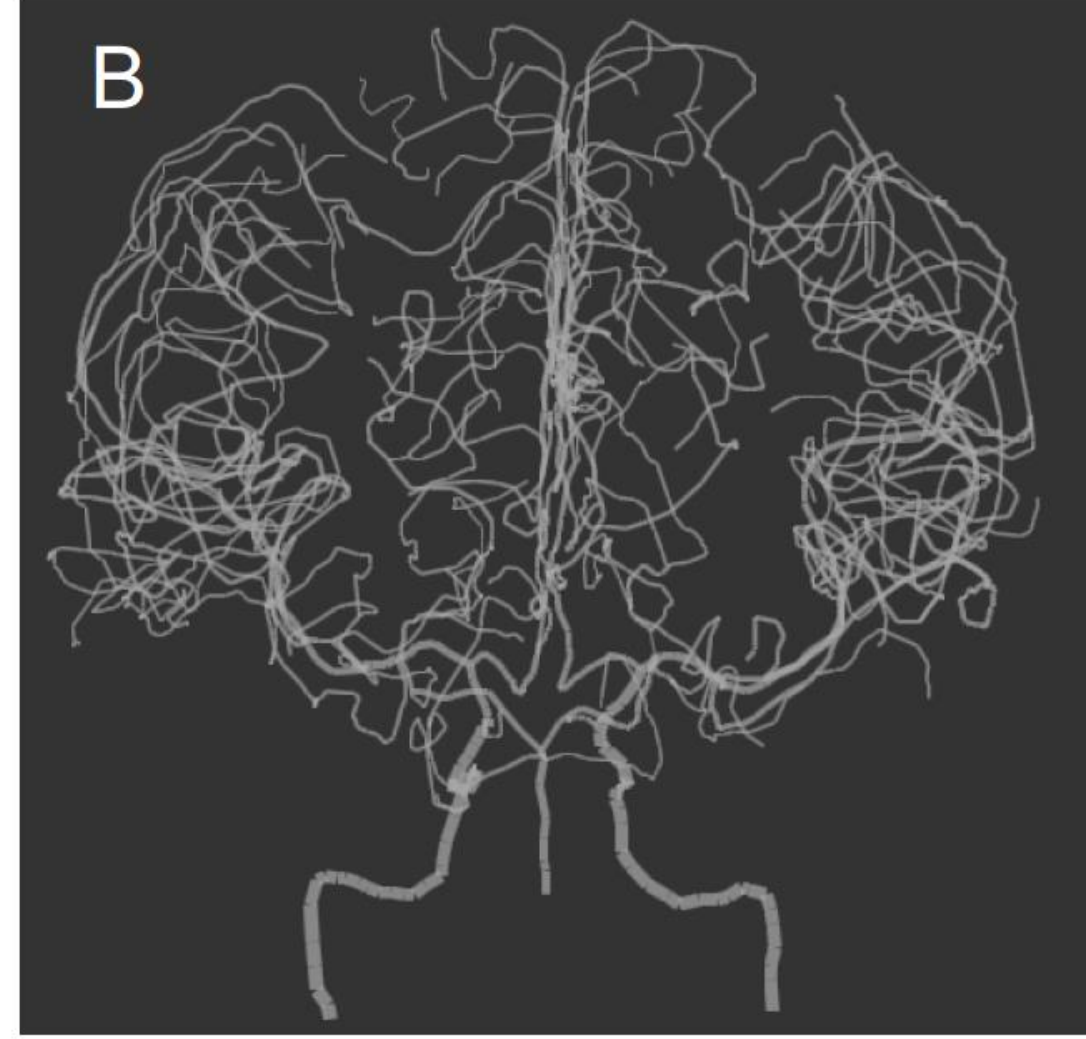
**62%**



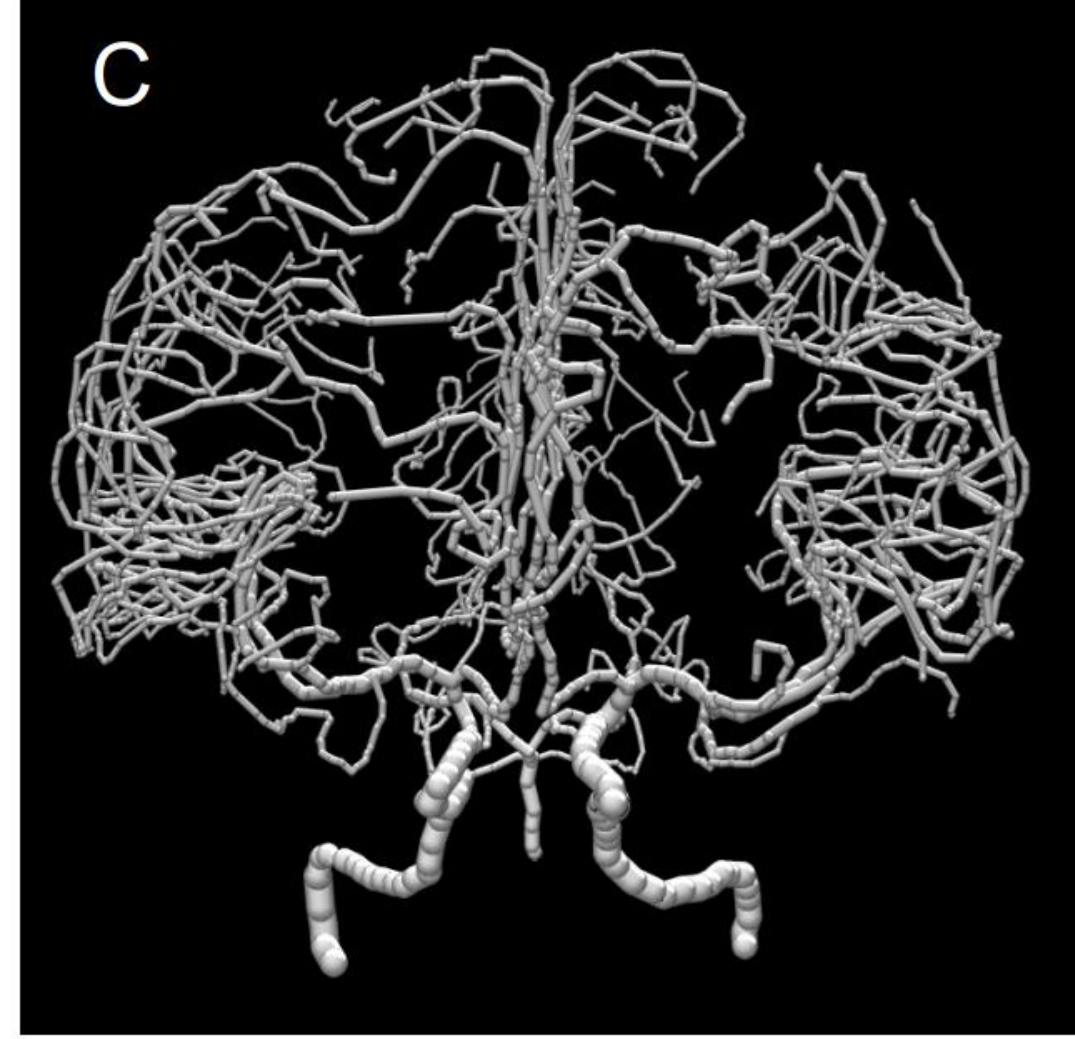
# “No Unjustified 3D”



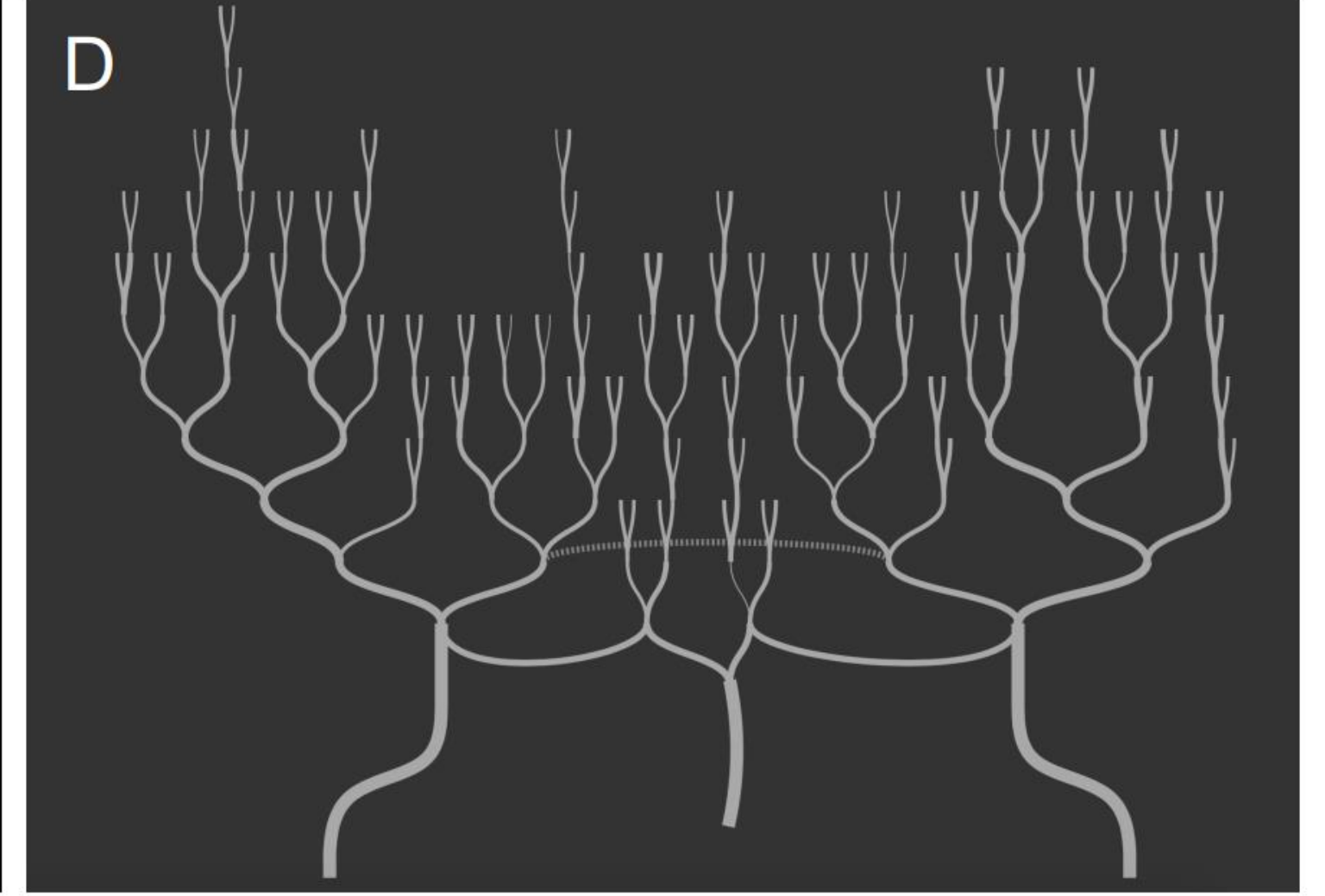
**Data Source**



**MIPS: 2D projection  
of 3D Arteries**



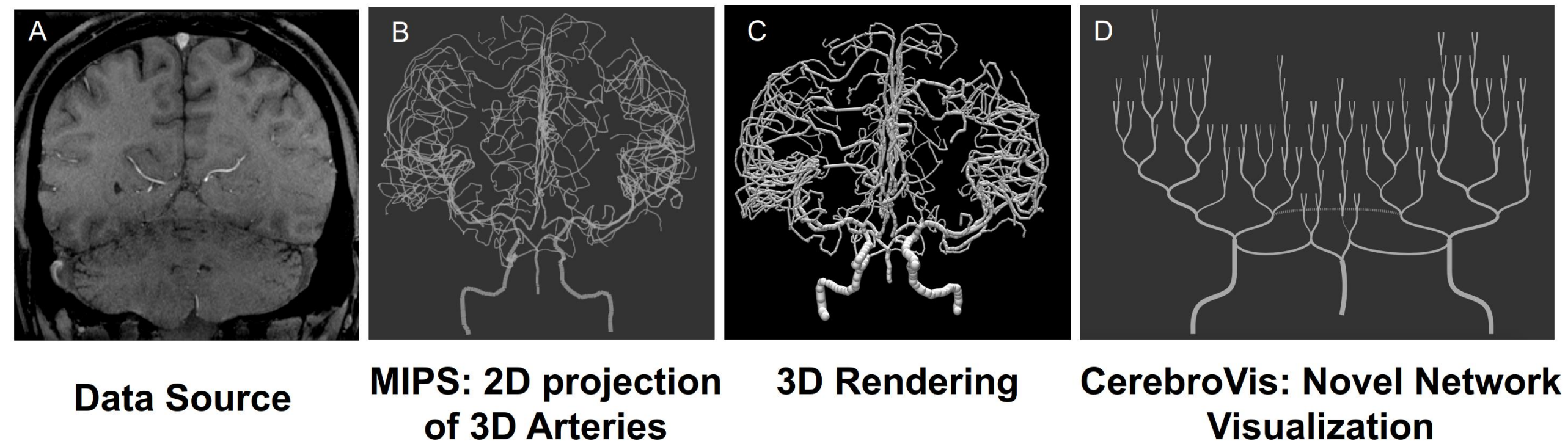
**3D Rendering**



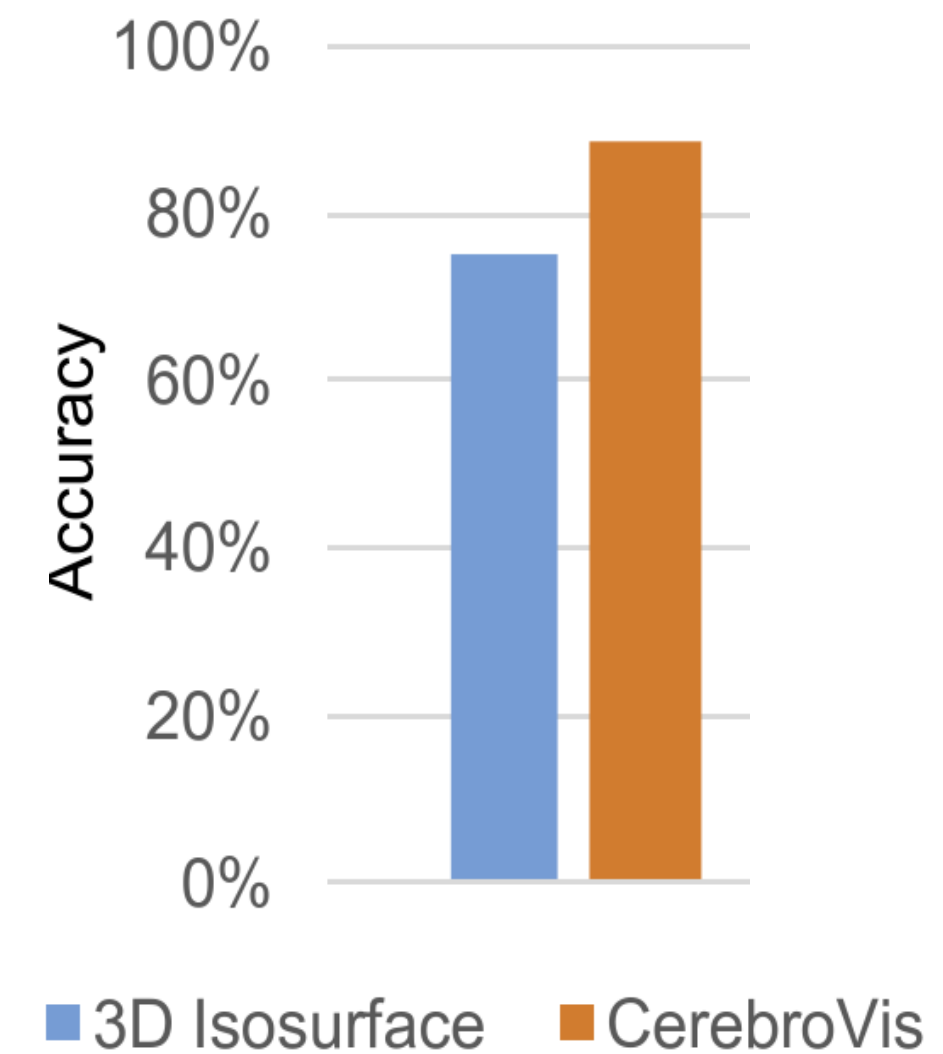
**CerebroVis: Novel Network  
Visualization**

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

# “No Unjustified 3D”

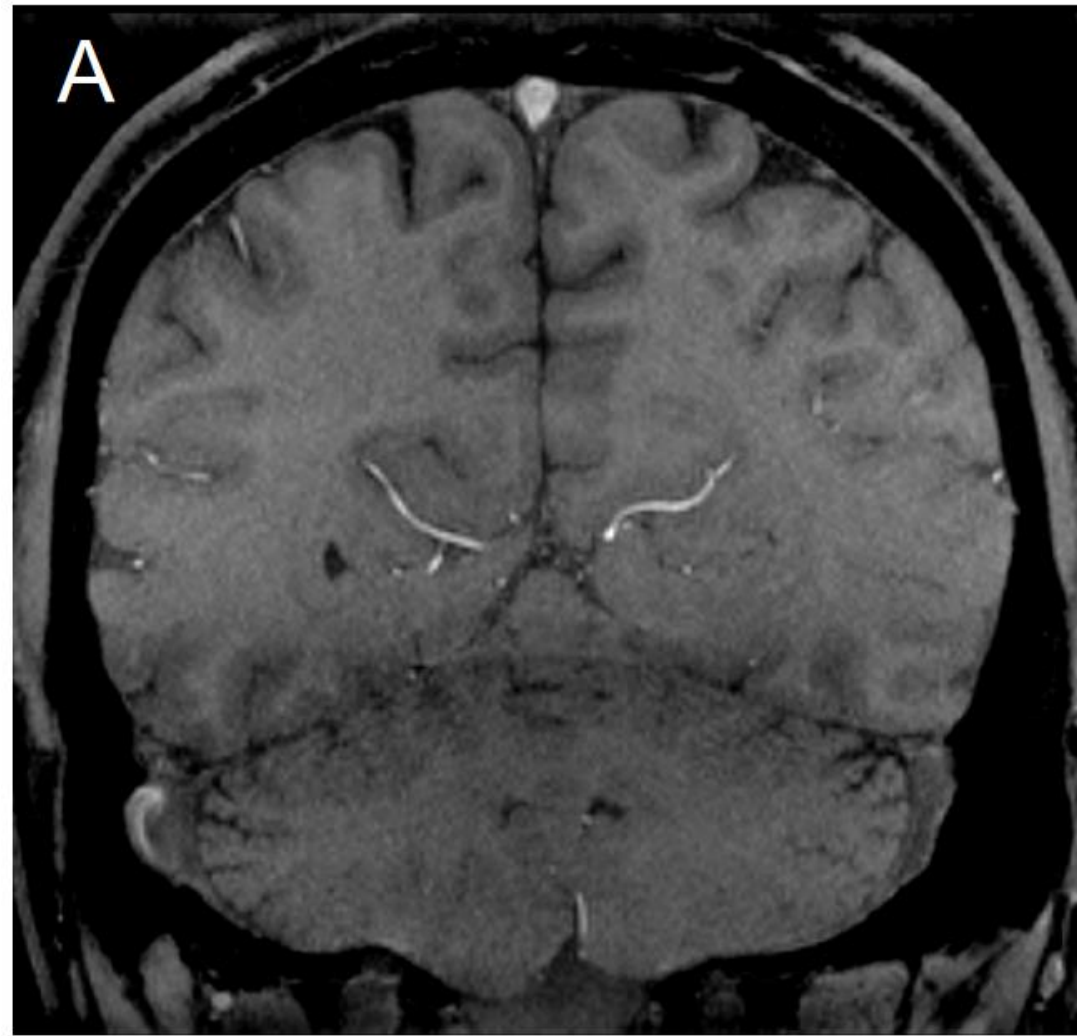


Stenosis Detection Accuracy 3D vs. CV

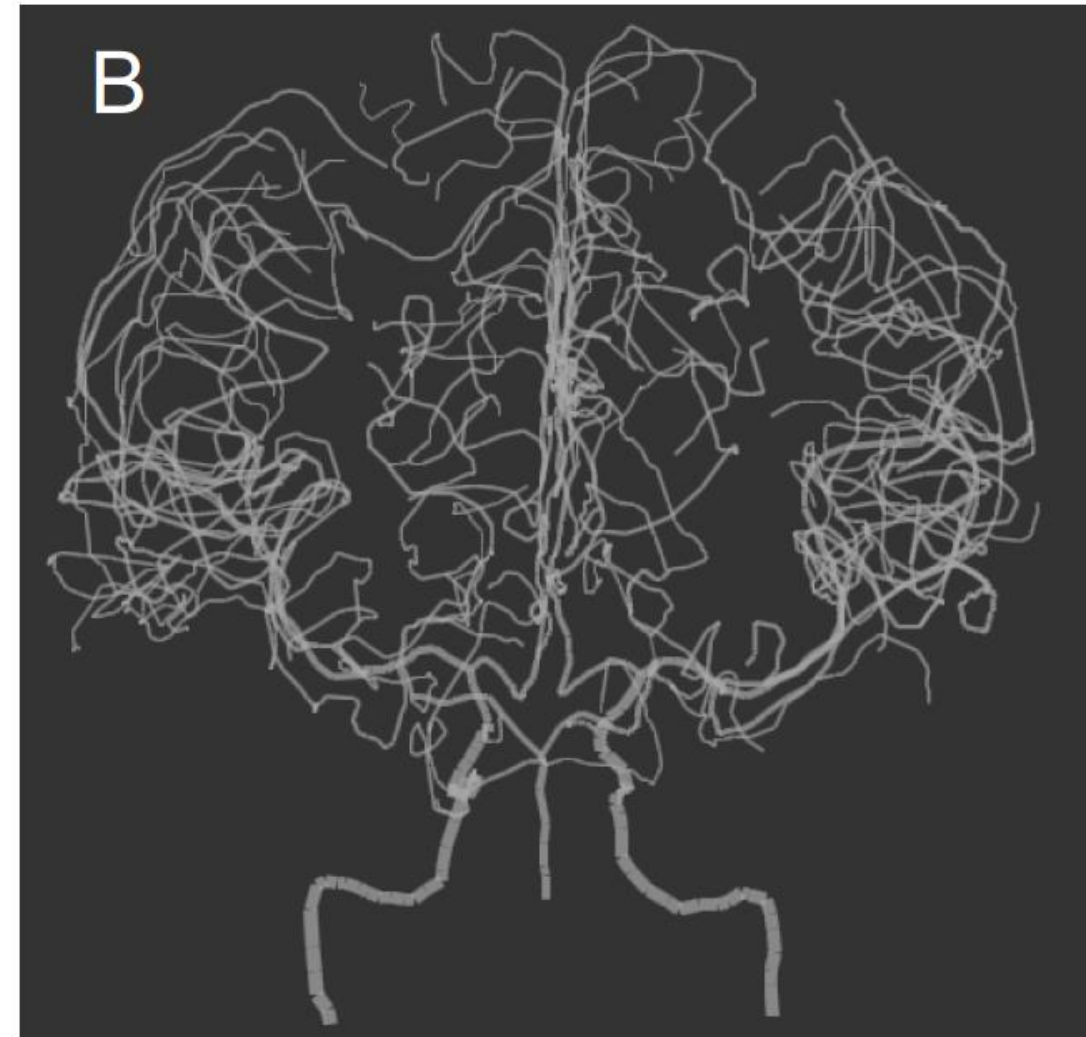


“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

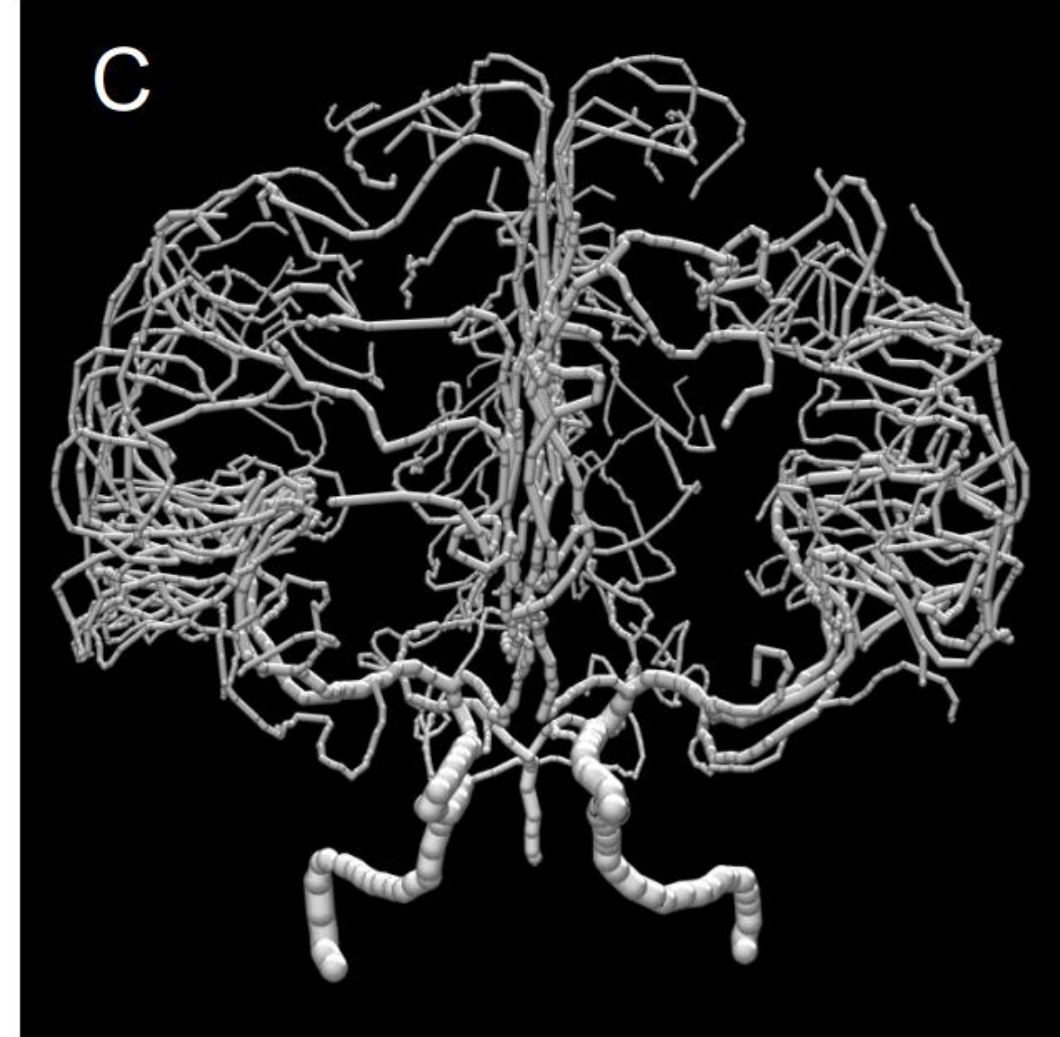
# “No Unjustified 3D”



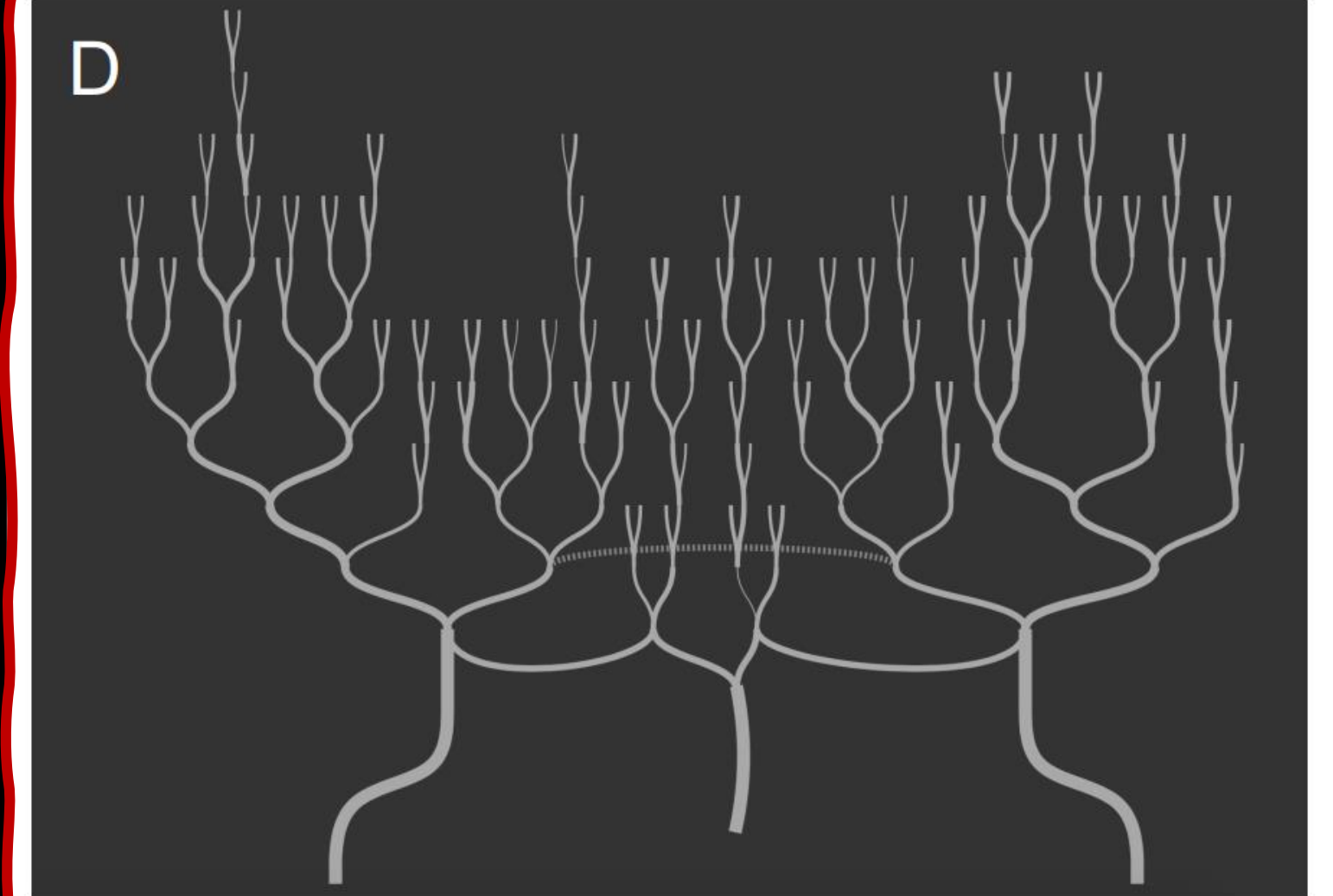
**Data Source**



**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.”