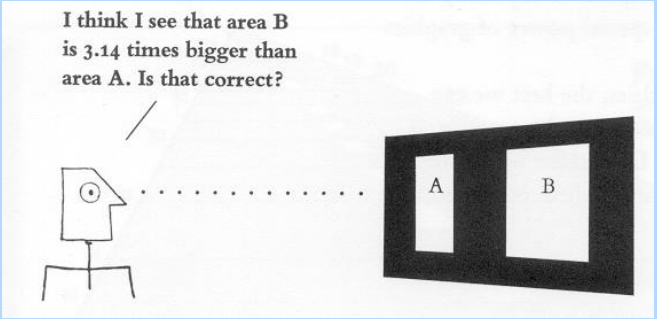
Achieving graphical Integrity

A graphic does not distort if the visual representation is consistent with the numerical representation.

- Is the magnitude of ‘visual representations’ as physically measured on the graphic?
- Or the perceived magnitude?



Approach

Conduct a study of visual perception of the graphics.

Circles – perceived area grows more slowly than measured area

reported perceived area = (actual area)^{x}, where x = 0.8+/-0.3

Lines -

Lie Factors

Given perceptual difficulties – strive for uniformity (predictability) in graphics (p56)

- ‘the representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.’

- ‘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.’

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

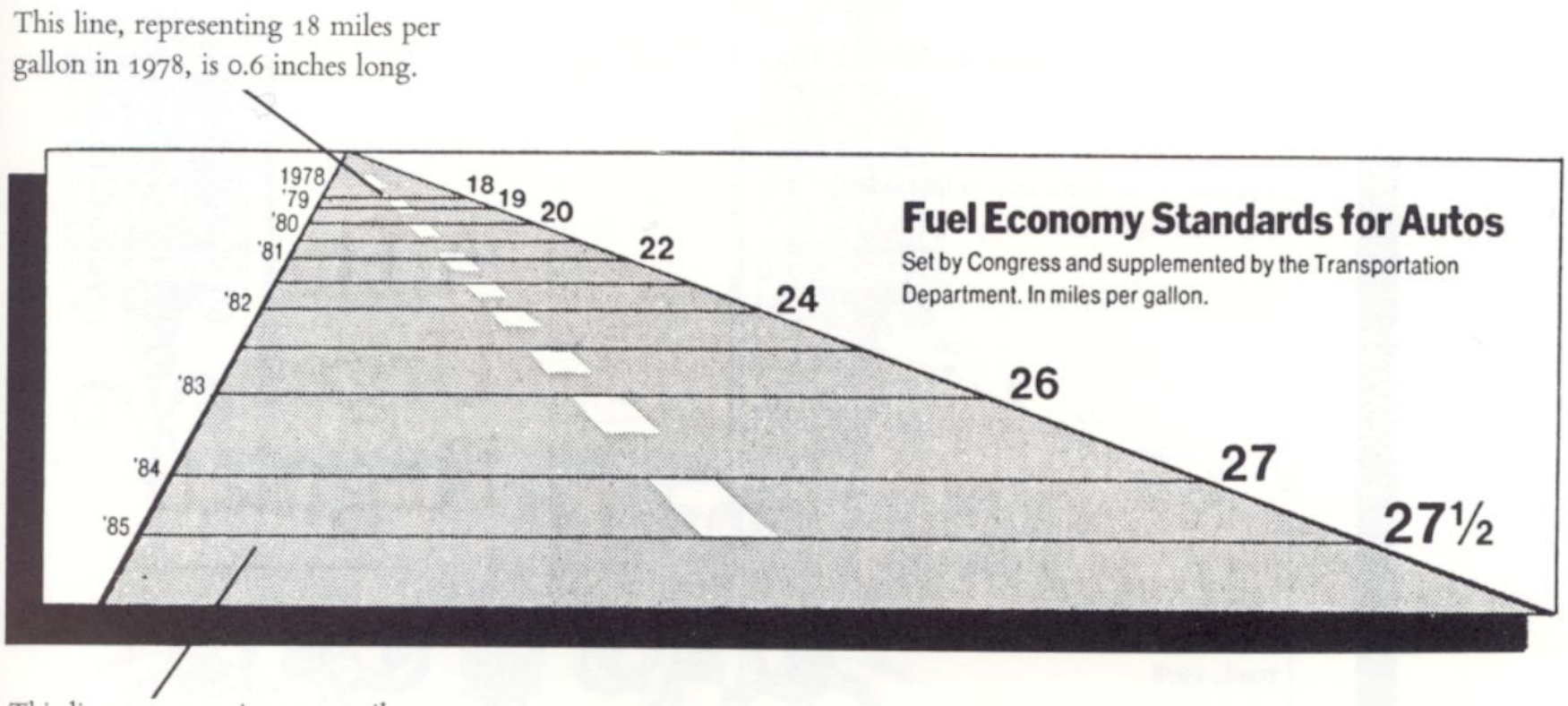
Lie factor of 1 – is desirable – lie factors > 1.05 or < 0.95 go beyond plotting errors

Extreme example

Fuel economy standards for automobiles

18 miles/gallon in 1978 to 27.5 miles/gallon in 1985

Increase of 53% = (27.5 – 18.0)/(18.0) x 100

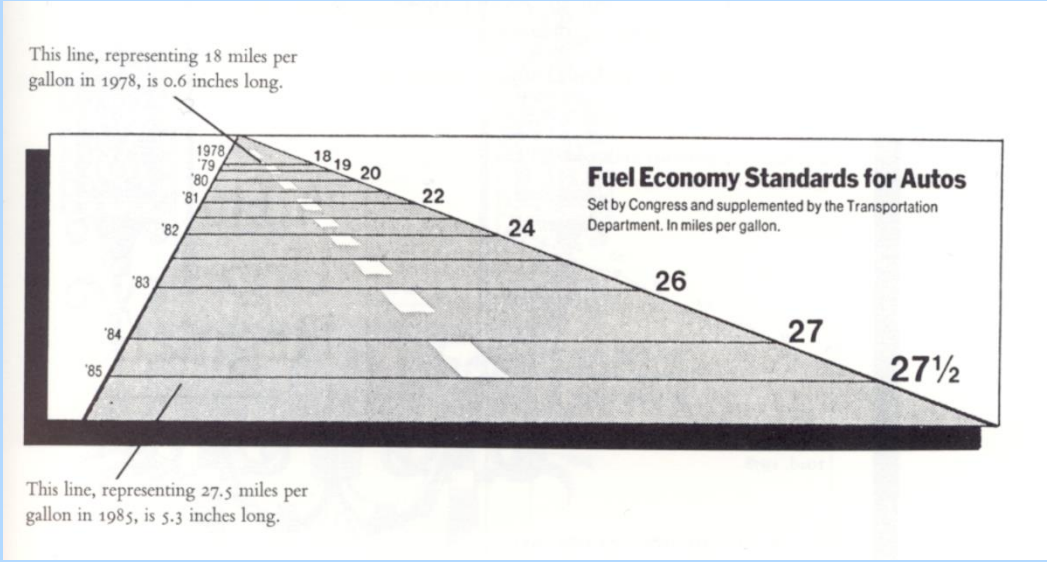


Extreme example

Graphic increase

783% = (5.3 – 0.6)/(0.6) x 100

Lie Factor = 783/53 = 14.8

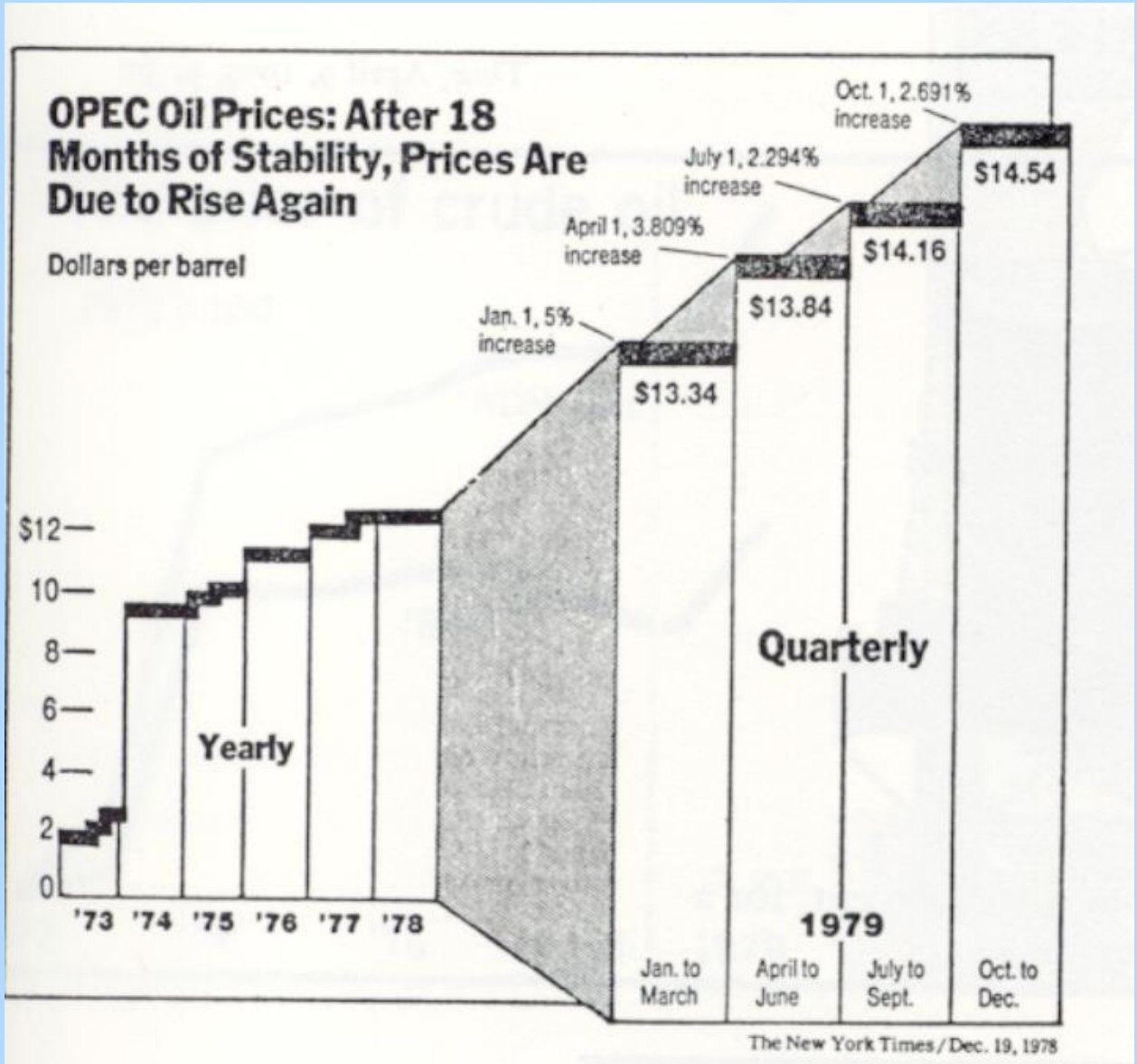


Additional confounding factors

- Usually the future is in front of us
- Dates remain same size and fuel factors increase
- Includes perspective distortion – how to read change in perspective

Design Variation vs Data Variation

New York Times, Dec. 19, 1978, p.D-7 (Tufte, 1983, p61)



Sheelagh Carpendale

Design Variation vs Data Variation

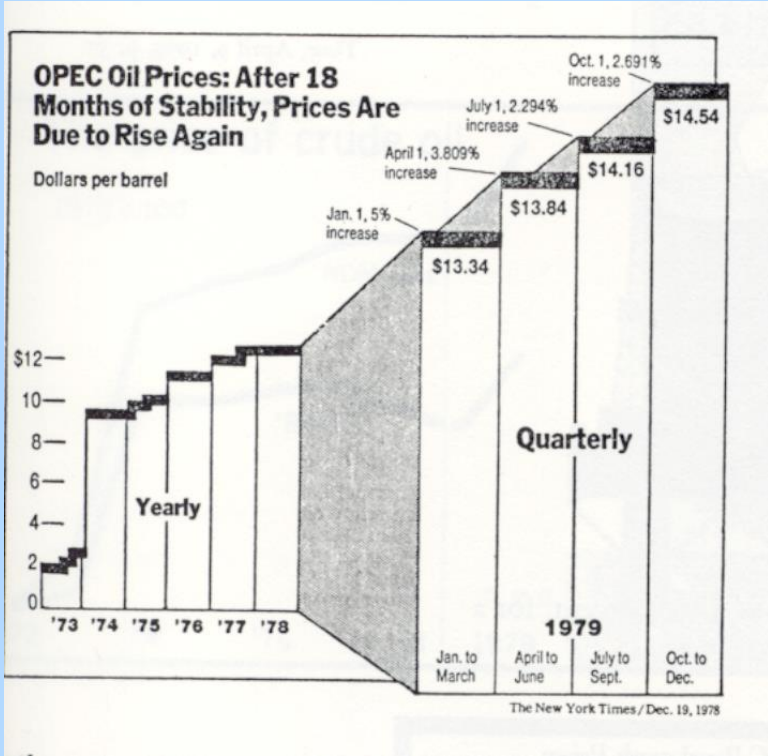
5 different vertical scales show price

1973 -1978	$8.00
Jan. – Mar. 1979	$4.73
Apr. – June 1979	$4.37
Jul. – Sept. 1979	$4.16
Oct. – Dec. 1979	$3.92

2 different horizontal scales show passage of time

1973-1978	3.8 years
1979	0.57 years

With both scales shifting the distortion is multiplicative



National Science Foundation, Science Indicators, 1974 (Washington D.C., 1976), p.15, (Tufte, 1983, p60)

Visual Area and Numerical measure

Use of area to portray 1D data can be confusing

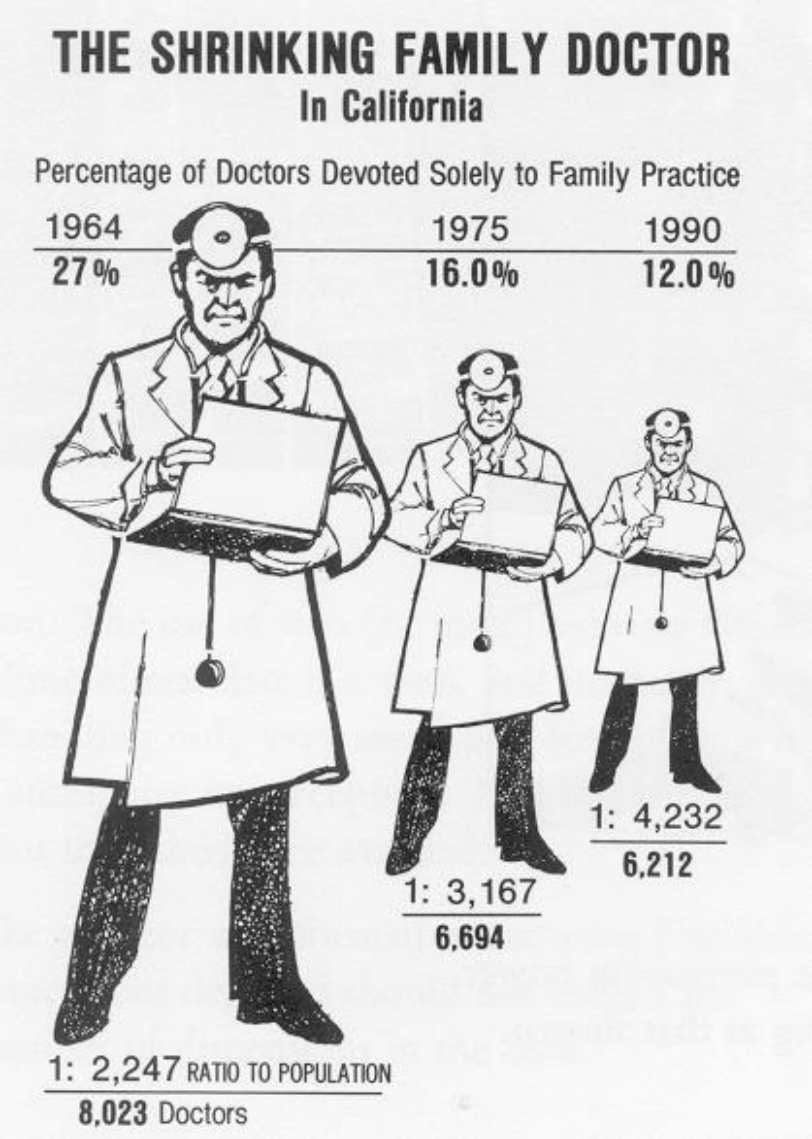
- Area has 2 dimensions

The ‘incredible’ shrinking family doctor

Lie factor of 2.8

Plus perspective distortion

Plus incorrect horizontal spacing



Los Angeles Times, August 5, 1979 p.3, (Tufte, 1983, p69)

Visual Area and Numerical Measure

Use of area to portray 1D data can be confusing

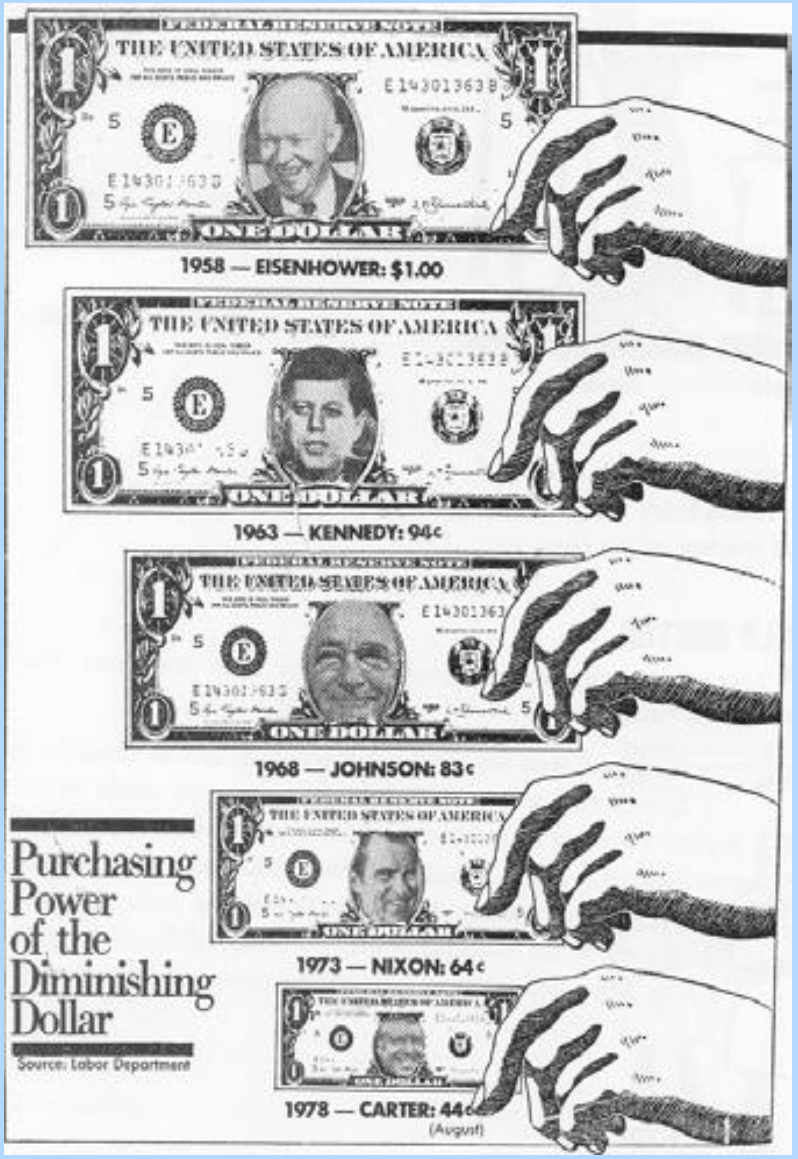
- Area has 2 dimensions

The ‘incredible’ shrinking dollar

The size of the dollar is adjusted in both height and width

The affect on the area in multiplicative

A more accurate representation of the 1978 dollar would be about twice the size of the one in this chart



Washington Post, Oct. 25, 1978 p.1, (Tufte, 1983, p70)

Visual Area and Numerical Measure

Use of area to portray 1D data can be confusing

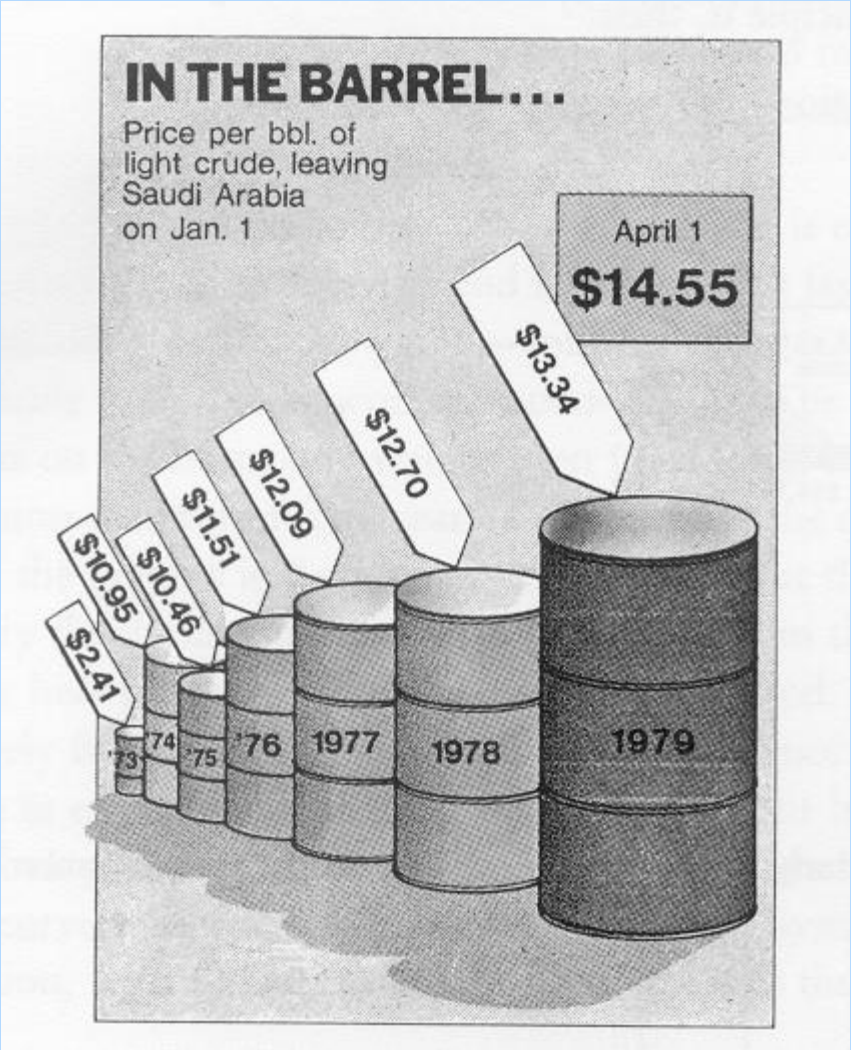
-Area has 2 dimensions

-There are considerable ambiguities of how people perceive area (2D) and then convert that to 1D data

-Even more so with volumes (3D)

By surface area lie factor is 9.4

By volume – lie factor is 59.4 – probably a record



Time, Apr. 9, 1979 p.1, (Tufte, 1983, p62)

Graphical Integrity - Summary

- ‘The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.
- Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the graphic itself. Label important events.
- Show data variation, not design variation.
- In time-series displays of money, deflated and standardized units of monetary measurement are nearly always better than nominal units.
- The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.
- Graphics must not quote data out of context.’

(Tufte, 1983, p77)