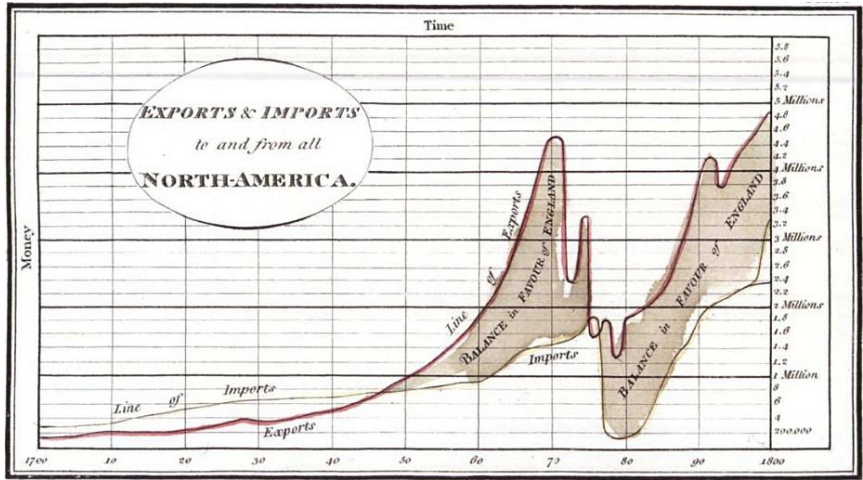


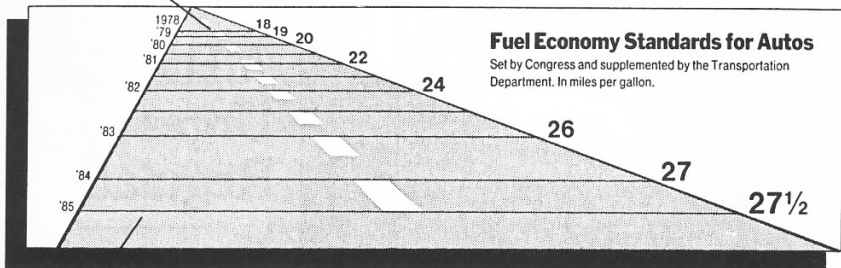
Are You Creating Chart Junk?

Yang Yang

University of Minnesota



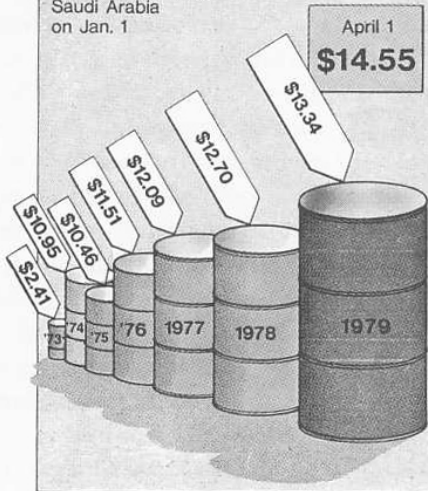
This line, representing 18 miles per gallon in 1978, is 0.6 inches long.

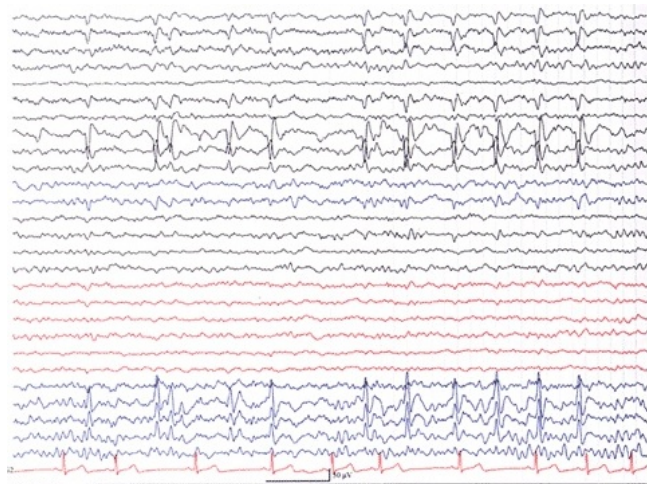


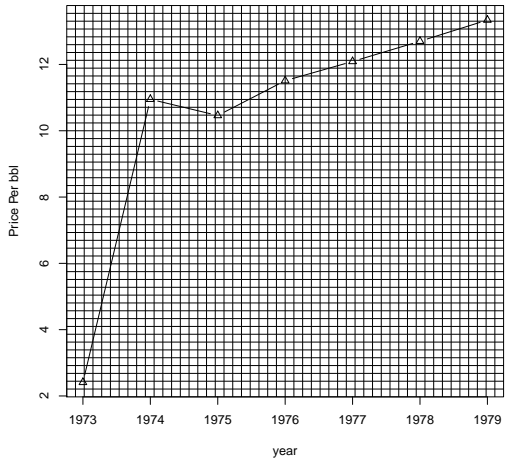
This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

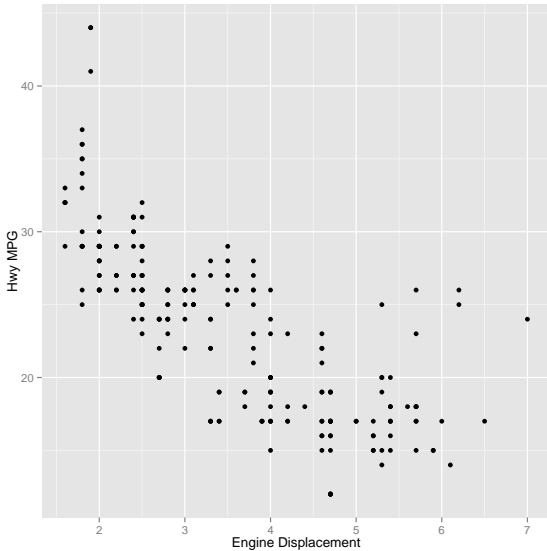
IN THE BARREL...

Price per bbl. of
light crude, leaving
Saudi Arabia
on Jan. 1



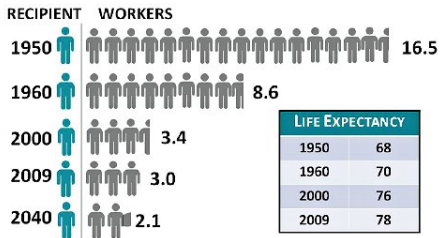




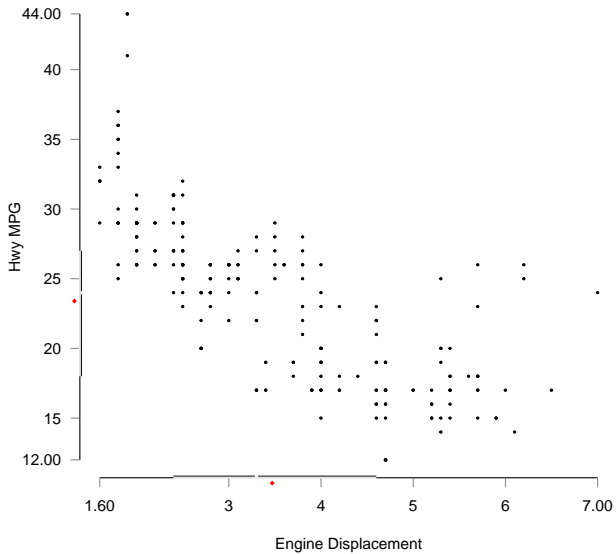




TAXPAYERS SUPPORTING EACH SOCIAL SECURITY RECIPIENT



SOURCE: SOCIAL SECURITY ADMINISTRATION, *THE 2010 ANNUAL REPORT OF THE BOARD OF TRUSTEES*;
CENTERS FOR DISEASE CONTROL, *U.S. LIFE TABLES*



Graphical Integrity

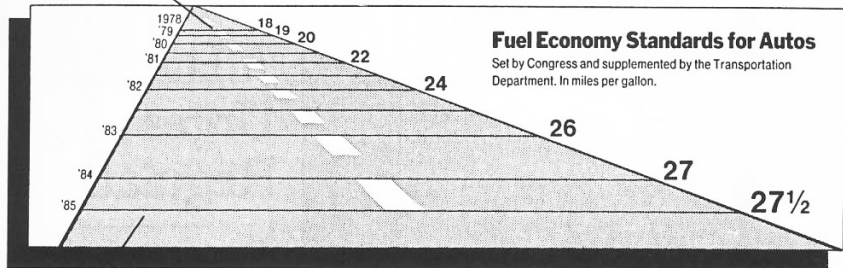
- Graphics are instruments for reasoning quantitative data
- Graphics shall not lie

Distortion in Graphics

- The physical representation of numbers \propto data represented
- Lie Factor = $\frac{\text{Size of effect shown in graphic}}{\text{size of effect in data}}$
- No distortion Lie Factor = 1

An Example

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.

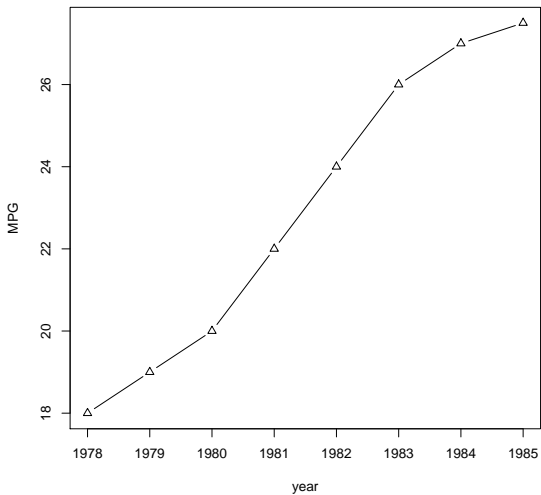


This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

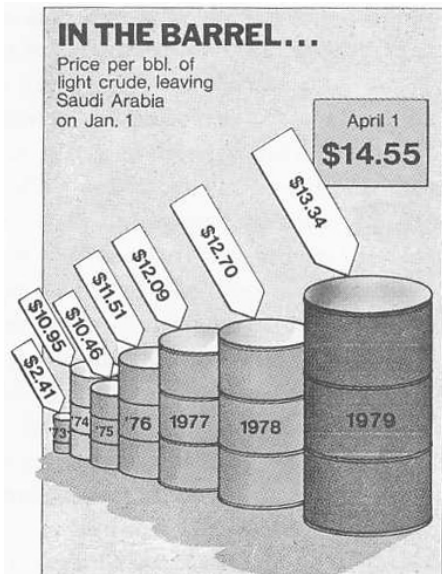
An Example: Cont

- Lie Factor = $(\frac{5.3-0.6}{0.6}) / (\frac{27.5-18}{18}) = 14.8421$
- $\gg 1$, BAD graphic!

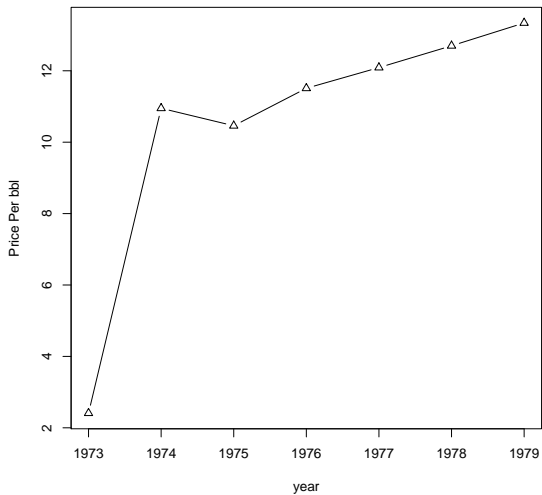
Improve the graphic



More Examples



Improve the graphic



Data Ink

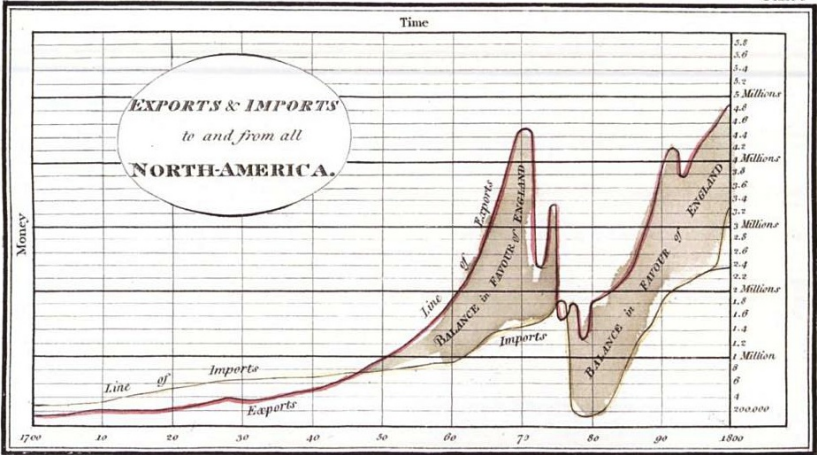
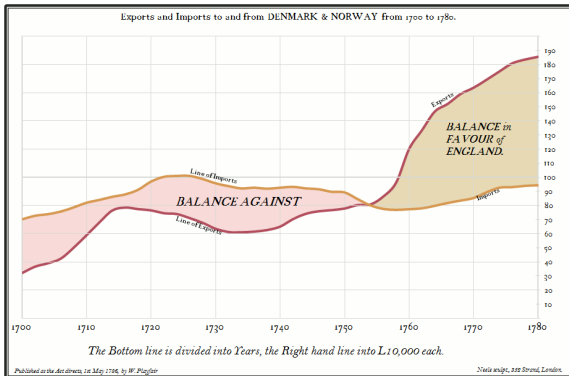


Figure : By William Playfair

“Above all else show the data”-Edward Tufte

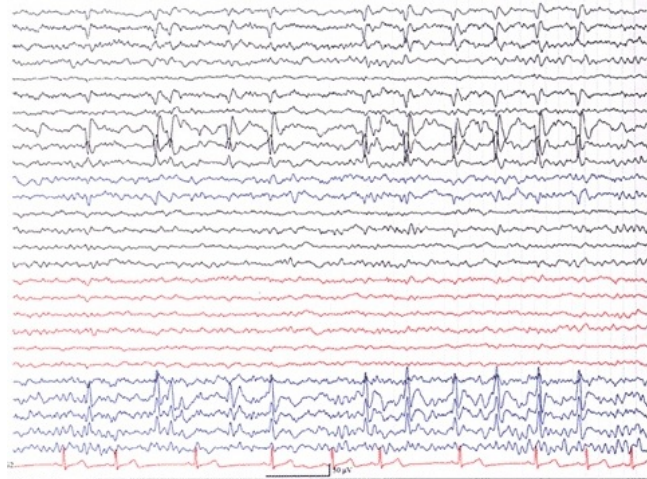
“Above all else show the data”-Edward Tufte



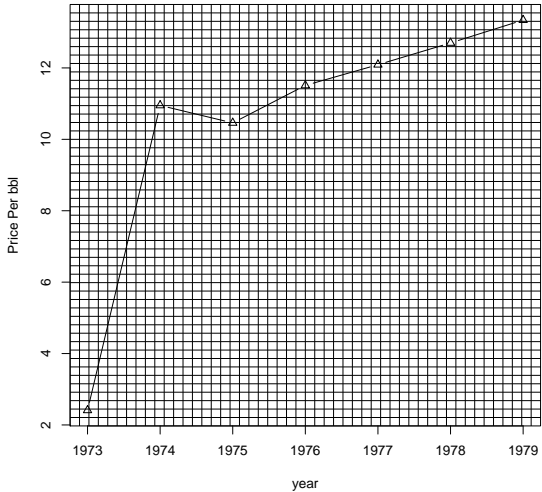
Data Ink

$$\begin{aligned} \text{Data-ink Ratio} &= \frac{\text{data-ink}}{\text{total ink used to print the graphic}} \\ &= 1.0 - \text{proportion of a graphic that} \\ &\quad \text{can be erased without loss of data-information} \end{aligned}$$

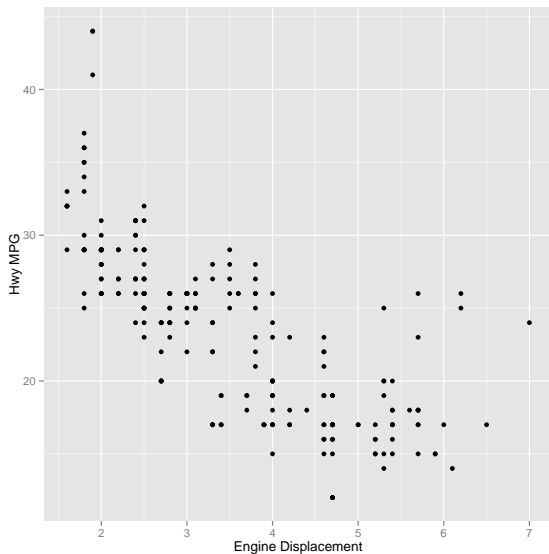
Examples: EEG



Examples: Grid

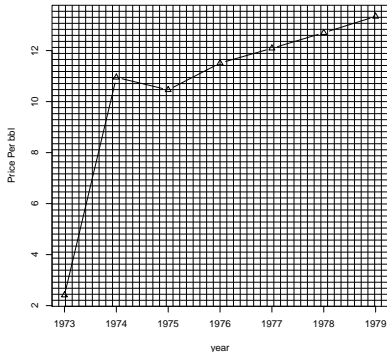


Examples: ggplot2



Grid = Chart Junk

- Grid lines contain no information
- Should be completely muted!

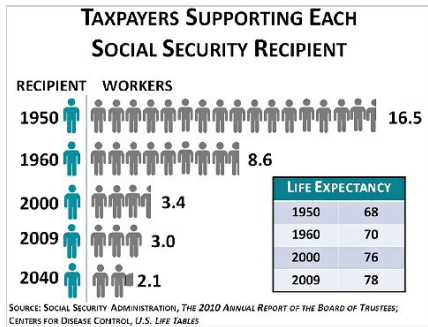


Duck = Chart Junk

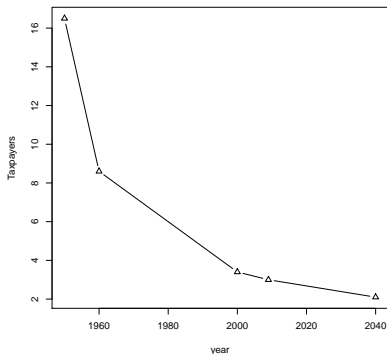
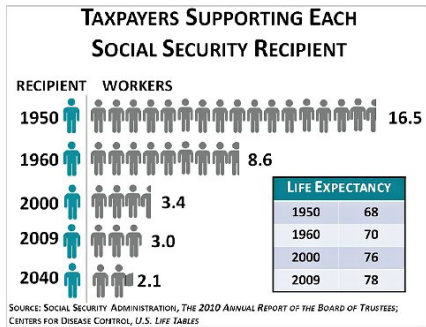
- Self promoting graphics: the duck
- "Modern Architects righteously abandoned ornament on buildings, but they unconsciously designed buildings that **were** ornament"
- It is OK to decorate construction, **NEVER** construct decoration



Example



Example

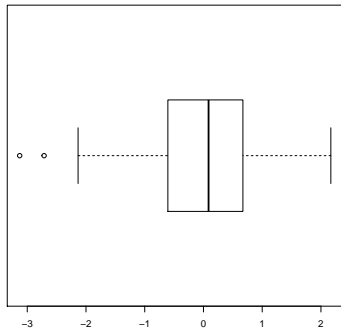


Maximize data-ink ratio

“Erase non-data-ink, within reason”-Edward Tufte

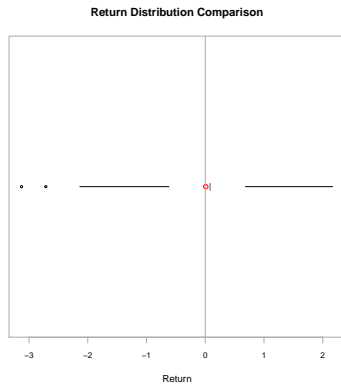
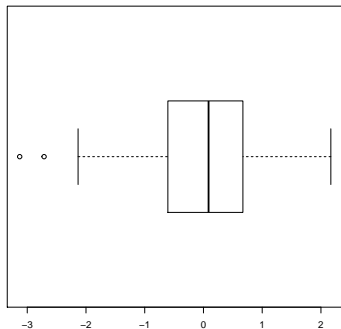
Maximize data-ink ratio

“Erase non-data-ink, within reason”-Edward Tufte

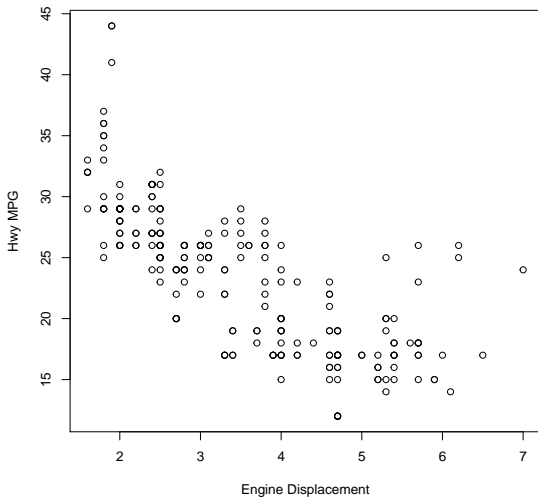


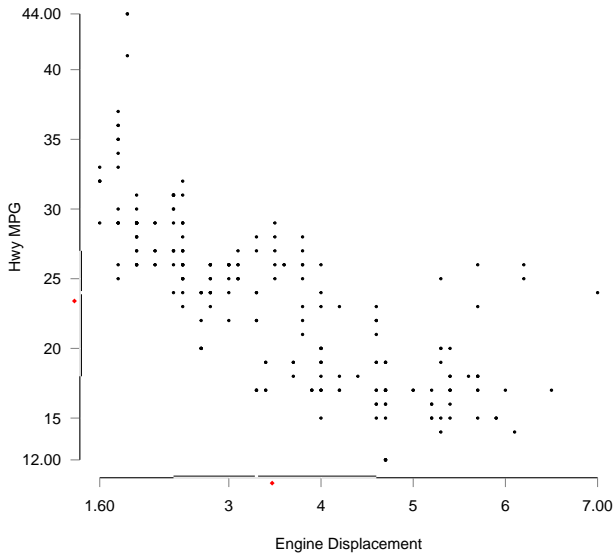
Maximize data-ink ratio

“Erase non-data-ink, within reason”-Edward Tufte



Scatterplot: MPG data





Thank you