







Evaluating Graphics

Approach to creating criteria: Begin by examining the guidelines for constructing graphics and work these into criteria for evaluating graphics.

-  Tufte's rules
-  Cleveland's principles
-  Carr's modifications
-  Wainer's analysis
-  Friendly's gallery
-  Our criteria

Dr. Cook and Dr. Hofmann, Iowa State University

Tufte's Rules



Show the data



Induce the viewer to think about the data



Avoid distorting what the data have to say



Present many numbers in a small space



Make large data sets coherent



Reveal the data at several levels of detail



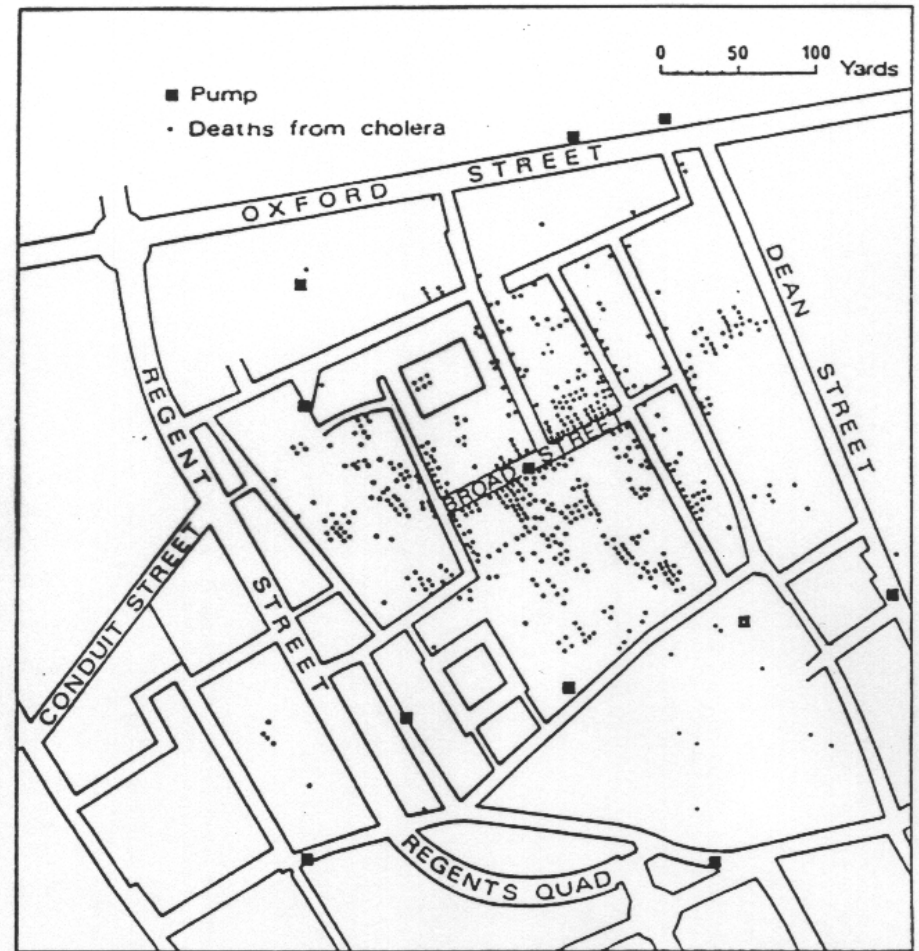
Serve a reasonably clear purpose



Be closely integrated with the statistical and verbal descriptions of the data

Show the data

What is the data in this picture?



Vignette Figure 1-1. Snow's map of cholera. The affected well is clearly identified by the concentration of cases in its vicinity. Reprinted from Howe, G. M. (1972). *Man, environment, and disease in Britain*. New York: Barnes and Noble Books, p. 178. Copyright 1972. Reproduced by Permission. Original source: Snow, J. (1855). *On the mode of communication of cholera*. London, 1855.

Source: http://xxi.ac-reims.fr/fig-st-die/actes/actes_2000/thouez/tl3.gif

Data

 Address of people who died from cholera

 Location of water pumps

Support

Map of London

.....Improvements

De-emphasize the map to raise the focus on the data

Data-ink ratio

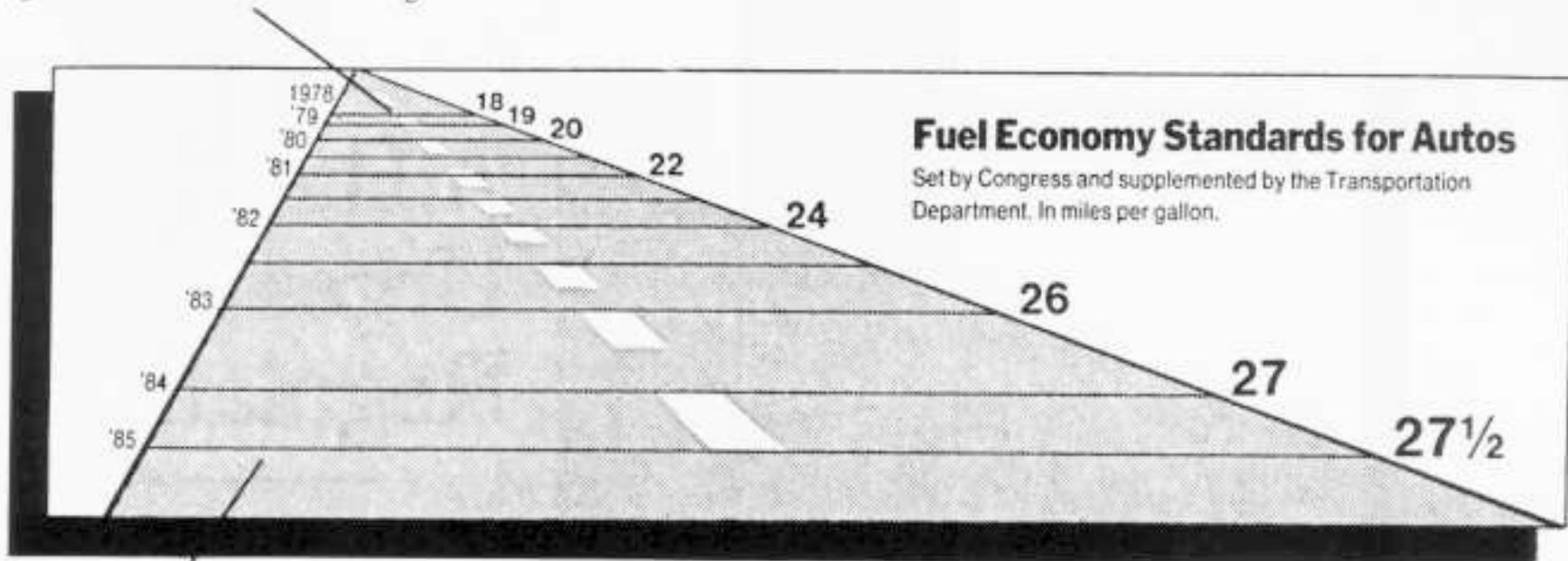
Divide the total ink used to draw the data by the total ink used to draw the graphic.

How do you calculate this? Not easily!

Avoid distorting the data

What's the data?	How is it represented?
Year	Time line
Fuel economy standard	<i>Line segment</i>

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.

Lie Factor (Tufte)

*Should be
close to 1*

$$\frac{\text{Size of effect shown in graphic}}{\text{Size of effect in data}}$$

Fuel economy example

$$\text{Data} \quad \frac{27.5-18.0}{18.0} \times 100 = 53\%$$

$$\text{Graphic} \quad \frac{5.3-0.6}{0.6} \times 100 = 783\%$$

$$\text{Lie factor} = \frac{783}{53} = 14.8 \gg 1 \quad \text{Huge!}$$

Avoid distorting the data

This is clearly a distorted view of the famous London Underground (The Tube)!

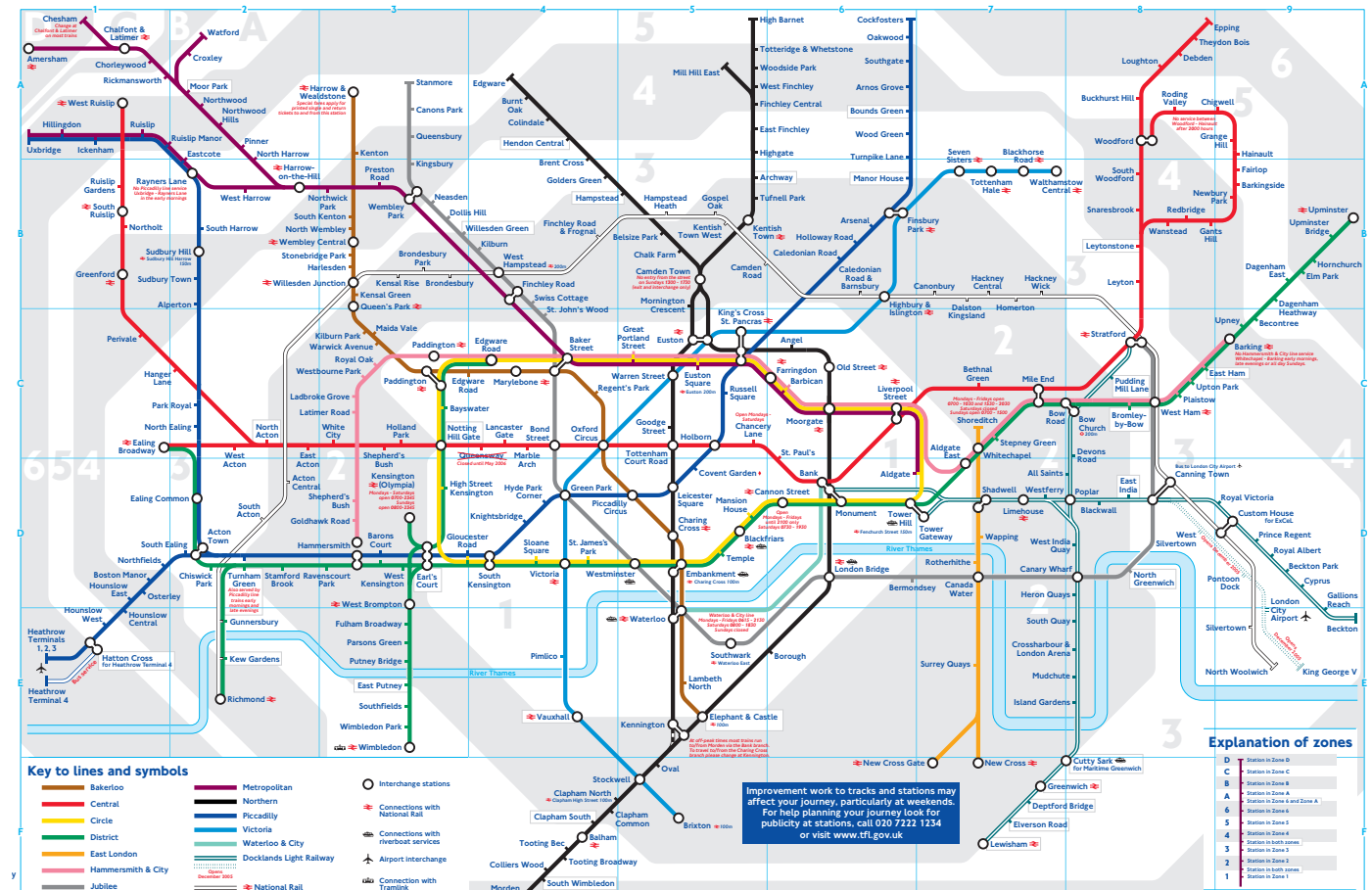
BUT... What's the data?



Sequence of stations



Interchanges between stations



Originally designed by Harry Beck, 1931

Chart junk

Interior decoration of graphics that generates a lot of ink but does not tell the viewer anything new.

3 numbers, decorated by the “Rising Sun” in front and a “Star-spangled banner” at back. Add a 3D effect - from Wainer (1997).

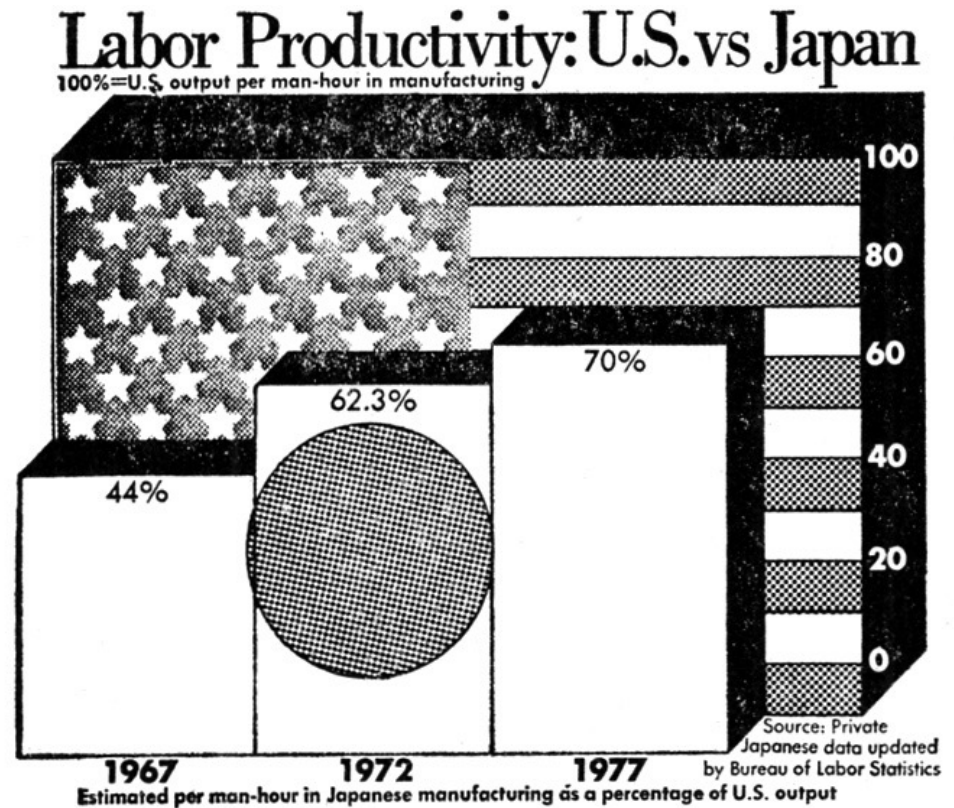


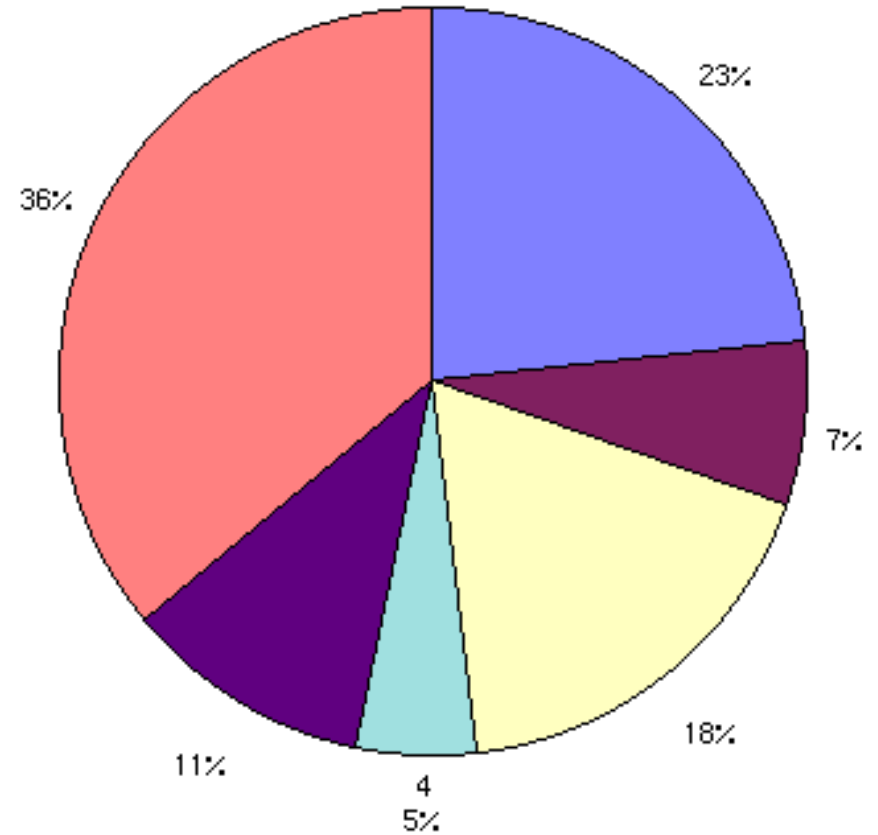
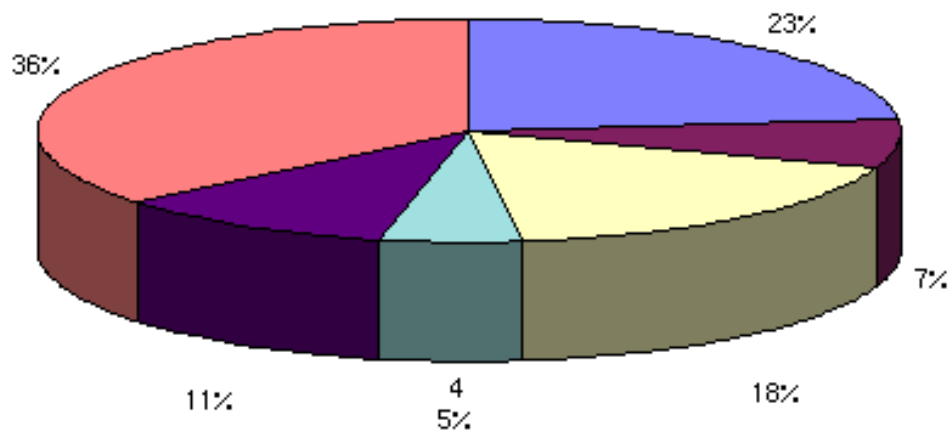
Figure 3. A low density graph (© 1978, The Washington Post) with chart-junk to fill in the space ($ddi = .2$).

Chart junk

... was the assumption that data graphics were mainly devices for showing the obvious to the ignorant. It is hard to imagine any doctrine more likely to stifle intellectual progress in a field. The assumption led down two fruitless paths in the graphically barren years from 1930 to 1970: First, that graphics had to be 'alive', 'communicatively dynamic', over-decorated and exaggerated (otherwise all the dullards in the audience would fall asleep in the face of these boring statistics)....

More chart junk

Why use 3D? To distort the perception of the pie slices?



Beyond chart junk: Induce the viewer to think about the data

How do you do this?

Cleveland's principles of graphical construction

Cleveland's graphical construction concerns primarily statistical plots of data, for a scientific audience. The two over-reaching principles are:



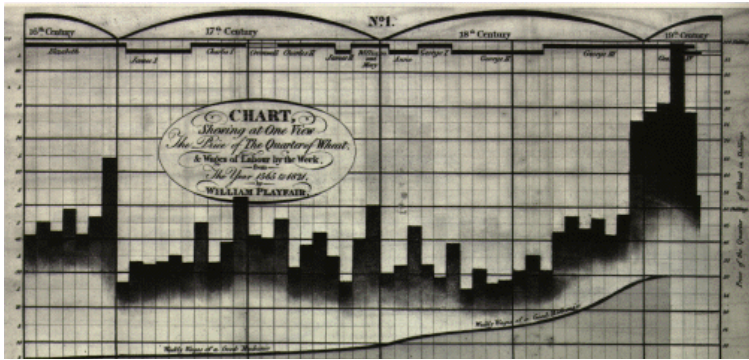
Make the data stand out



Avoid superfluity

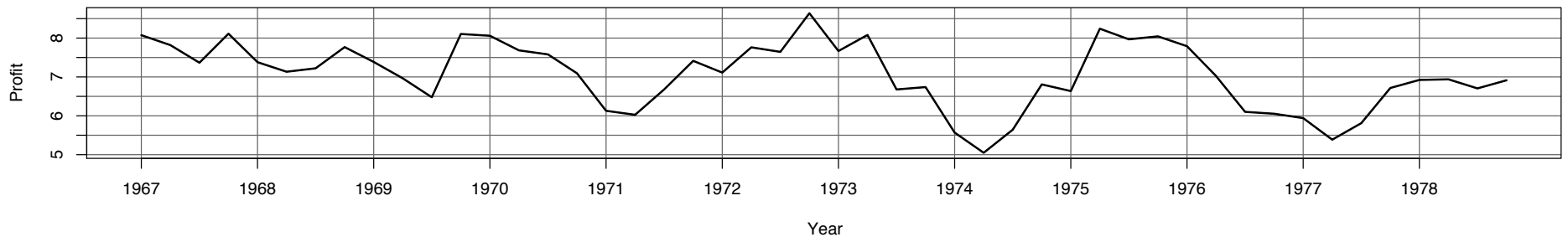
Cleveland's principles

- ✿ **Terminology:** title, caption, legend/key, data label, reference line, data rectangle, ...
- ✿ **Clear** vision/understanding
- ✿ Use of **guides:** scales, axes, tick marks, grid lines, legends
- ✿ **Aspect ratio:** scale of horizontal to vertical
- ✿ Extensive **captions**
- ✿ **Resolution:** clarity under re-scaling

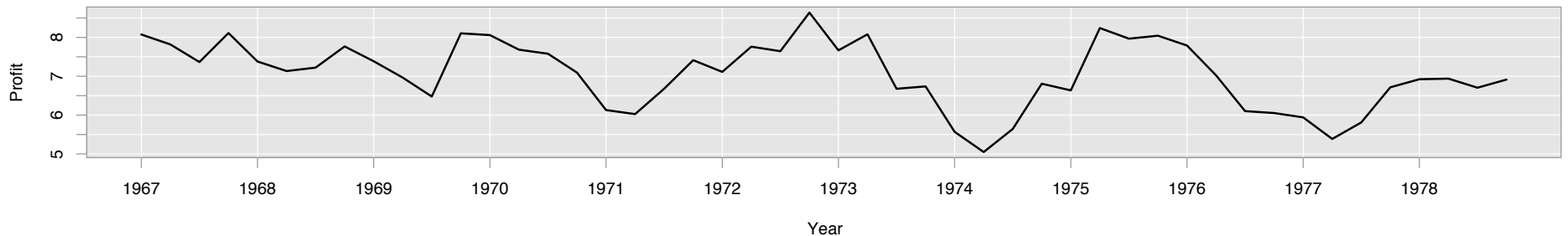


De-emphasize grids

UK Pig Production



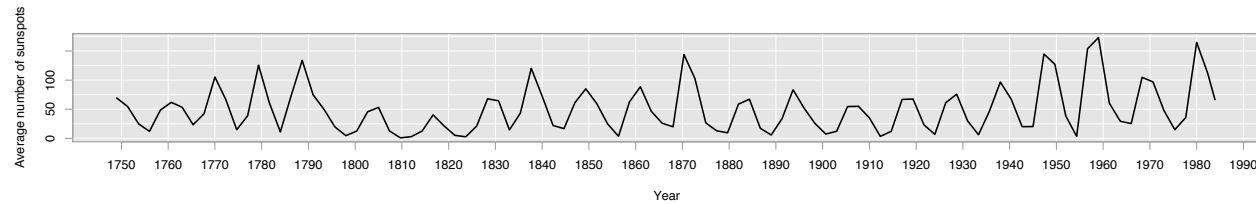
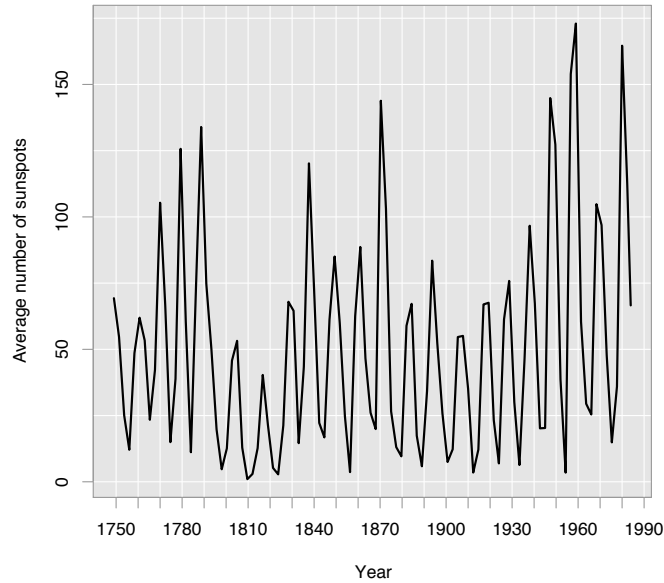
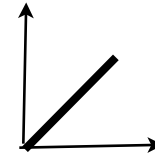
UK Pig Production



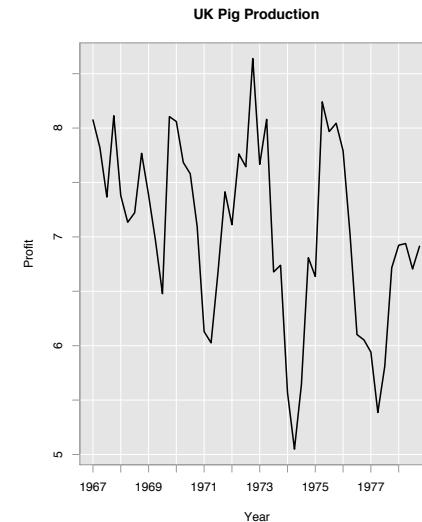
Dan Carr: background+pale grid+dark data marks
Sets plot off from page, and makes grey scale equivalent to text block of same size

Aspect ratio

Banking to 45° - Lines on average are at



Banking criterion
difficult to calculate



Familiarity and surprise



Use of conventions (Tufte, Wainer):
familiarity, ease of reading, for example,

Common types of plots - barchart,
pie chart, map, ...

Time on the horizontal axis, map
with north at top of page, ...



Bending conventions - elicit interest

Cognate Strategies (Kostelnick & Roberts)

- ✿ Arrangement (genre, spatial orientation)
- ✿ Emphasis (color, figure-ground contrast)
- ✿ Clarity (perceptual principles)
- ✿ Conciseness (“chartjunk” and other clutter)
- ✿ Tone (technical, subdued, loud)
- ✿ Ethos (credible, relevant to readers)

Mapping data to elements

- 🌸 Numeric: points (along a line), area, size, angle
- 🌸 Categorical: Color, location on a line
- 🌸 Time: lines connecting consecutive points
- 🌸 Geography: map coordinates, points along line in special coordinate system

Criteria for Evaluating Graphics

Context (Rhetoric)

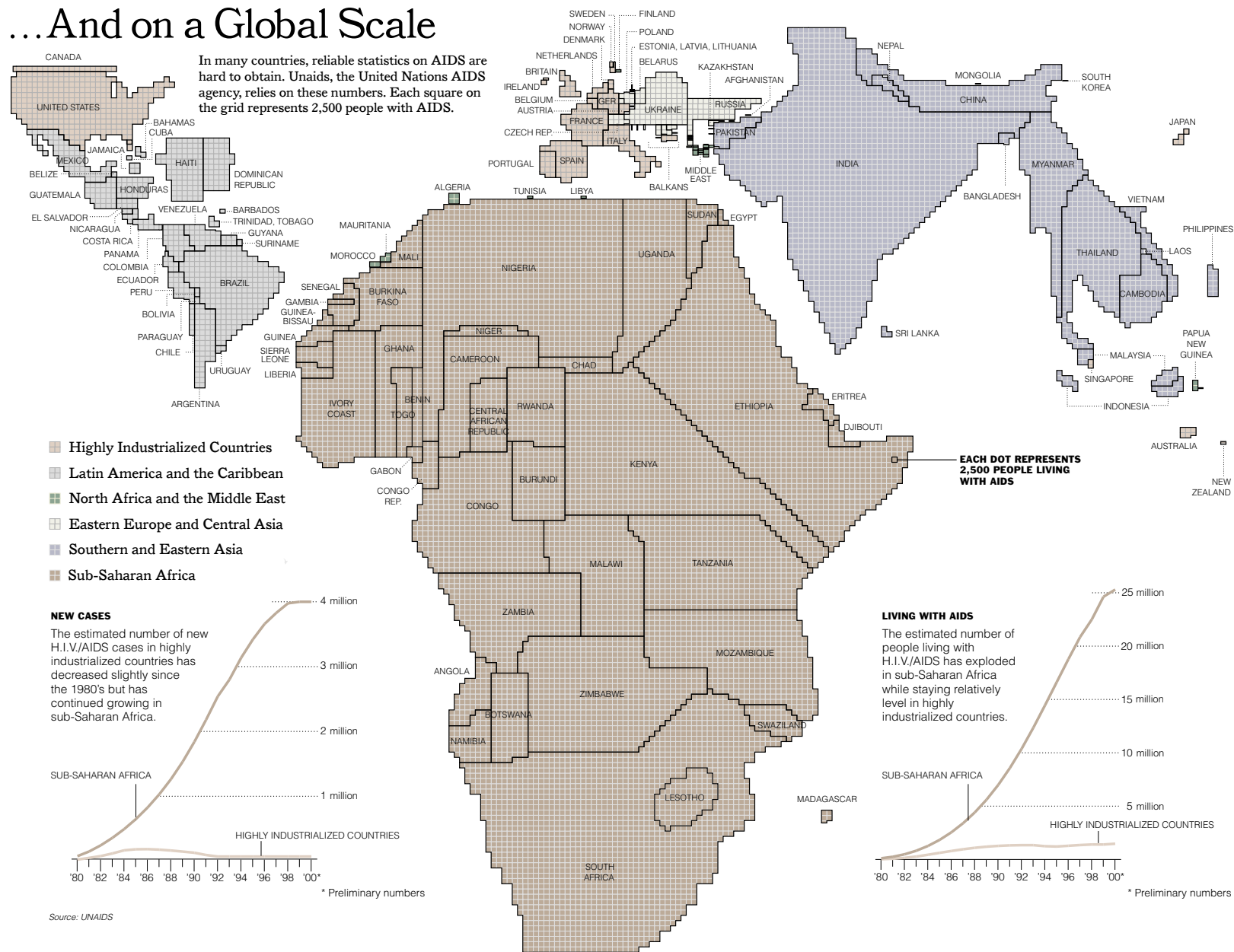
1. What is the main message? Sub-messages? Story.
2. Why/when was it produced? Does it have Kairos?
3. Who's the audience?
4. What are the pieces of information?
5. How is the information coded into the graphic?
6. What conventions are used? What is unconventional?
7. Is the data accurately represented? Lie factor, trustworthiness.
8. What is the ratio of data to ink in the plot? High, medium, low.
9. What's missing?

Content (Aesthetic)

Perception (Perceptual)

10. How clearly is the information represented?
11. What is emphasized, de-emphasized?
12. How is the viewer drawn in?
13. What is your overall impression, opinion?

...And on a Global Scale



Context (Rhetoric)

1. *Main message: AIDS is a big problem in the continent of Africa.*

Sub-messages:

- New AIDS cases in the highly industrialized world are diminishing.
- Number of people living with AIDS in highly industrialized countries is fairly flat now.
- Numbers for China and Russia are suspicious.

2,3. Produced by the NY Times for its readership, fairly educated audience. Probably not kairos right now.

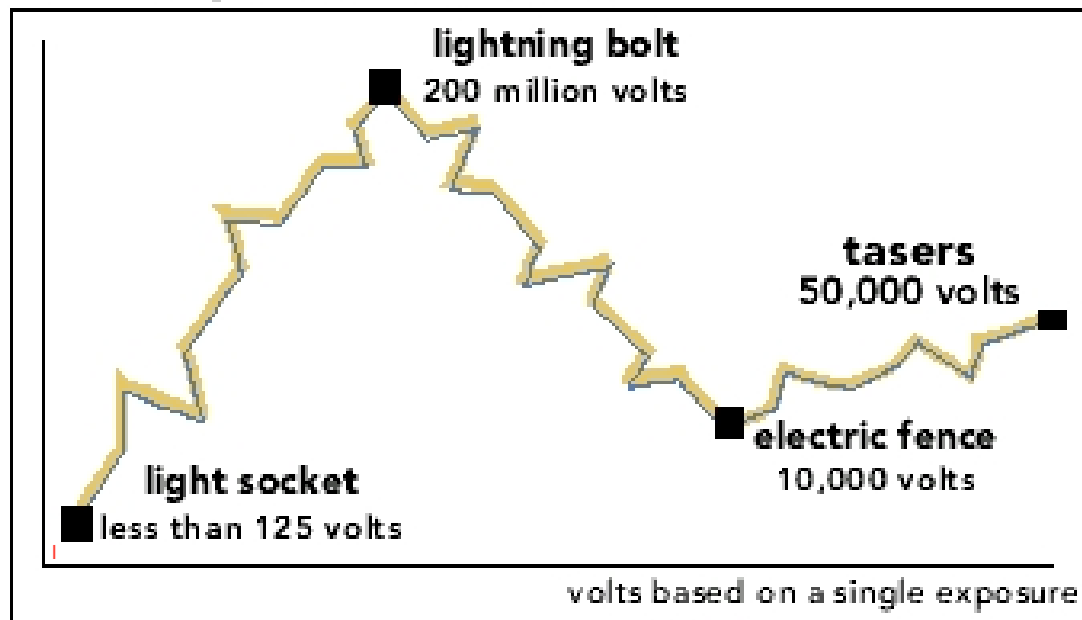
Content (Aesthetic)

- 4,5. Geographic location - xy position in rough map coords
Number of AIDS cases - as area
- 6. Map is familiar, sizes are all wrong!
- 7. Has a guide to number=area, accuracy high, although it is in blocks of 2500 people.
- 8. Almost all of the ink is data. There is a lot of text, and subsidiary information.
- 9. Perhaps population totals of country, to see proportions or rates. Availability of drugs, health care.

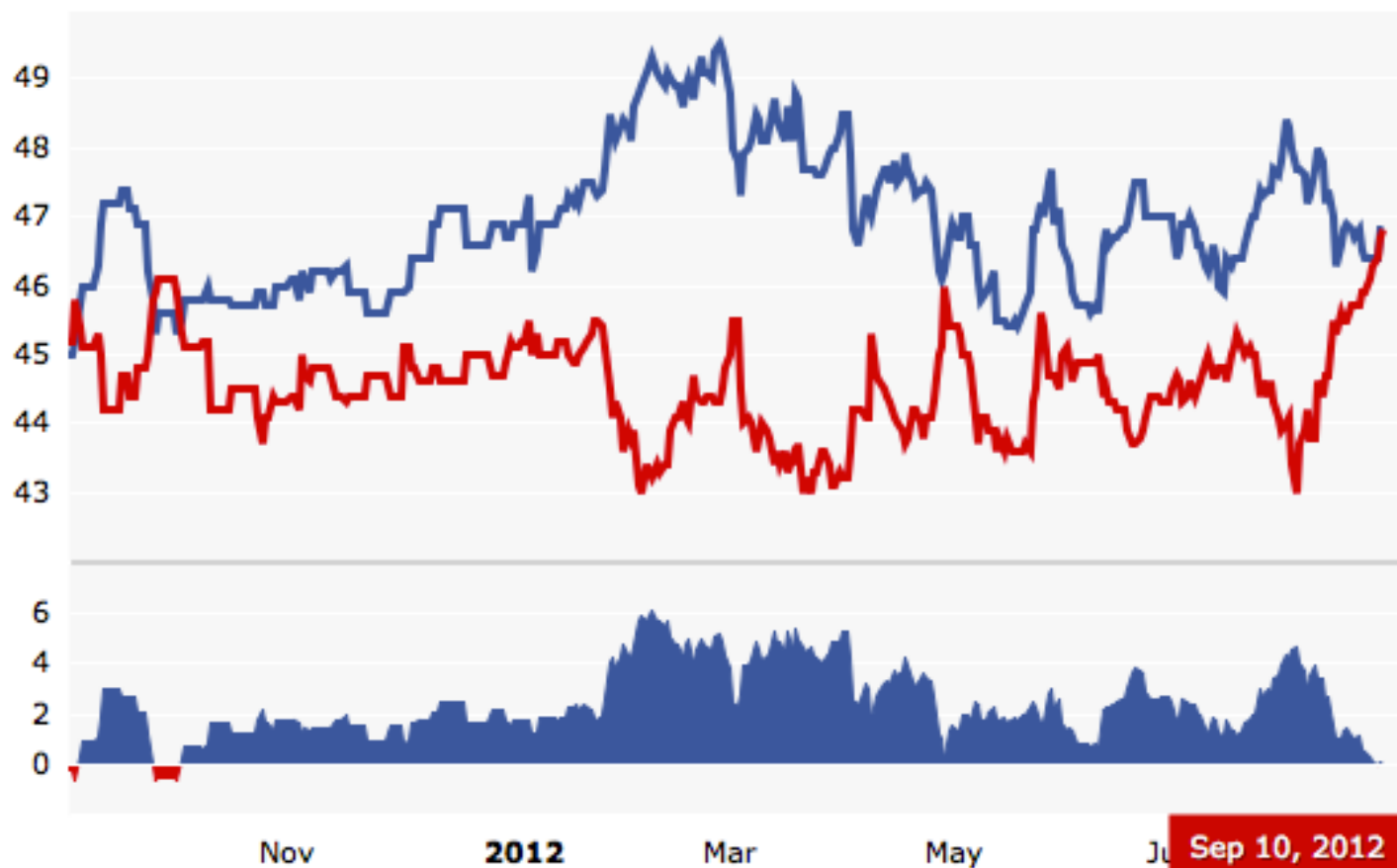
Perception

- I 0. Mapping is made clear by legends, and explanations that don't dominate the graphical elements.
- I 1. Counts of AIDS cases is emphasized.
- I 2. Familiar view of shapes of countries draws us in, but wrong sizes elicits interest.
- I 3. I love it! There is a huge message and there are multiple levels of detail.

How dangerous are tasers?



Graphic: Megan Steenson/Iowa State Daily



FROM: 09/12/2011

TO: 08/5/2012

APPLY

RESET

7D

14D

30D

6M

1Y

MAX