Does being Ethical make Good Visual Sense?

AUSTIN E. COFRIN
SCHOOL of BUSINESS
UNIVERSITY of WISCONSIN-GREEN BAY

Dr. Gaurav Bansal
UW-Green Bay

Frederick E. Baer Professor in Business

Learning Objectives

- Good visualizations = Ethical visualizations
- Design principles
- How to Critique our own visualizations

The four datasets represented here probably have different….
(select all that apply)

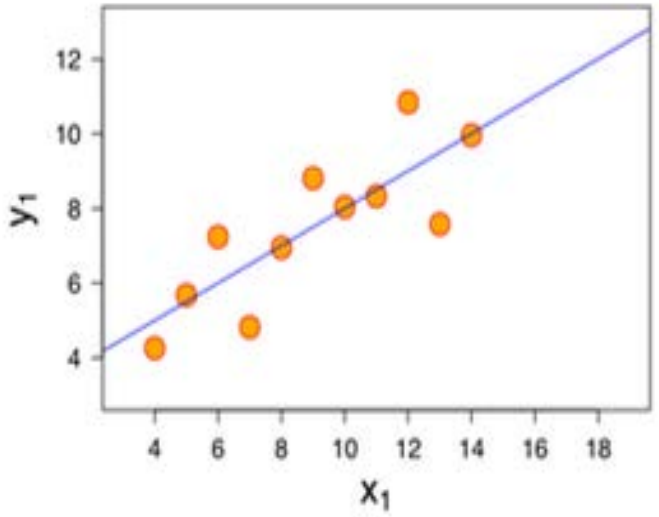
A. Mean

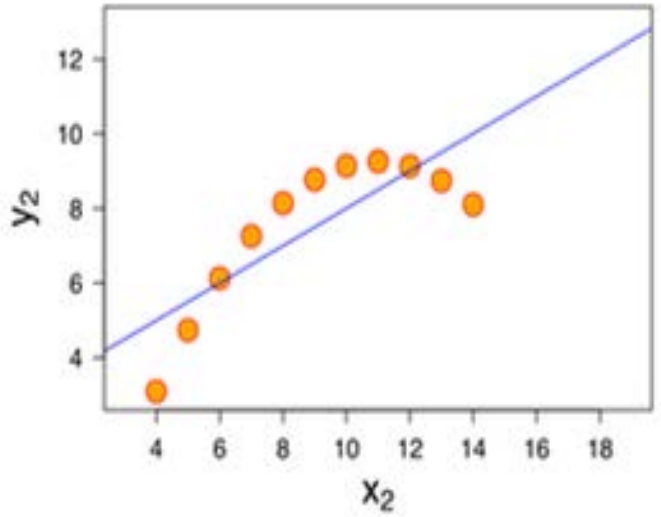
B. Standard deviation

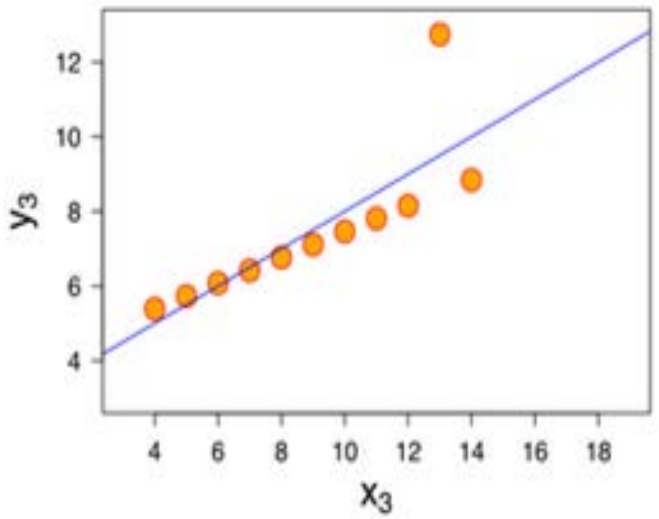
C. Correlations (x and y)

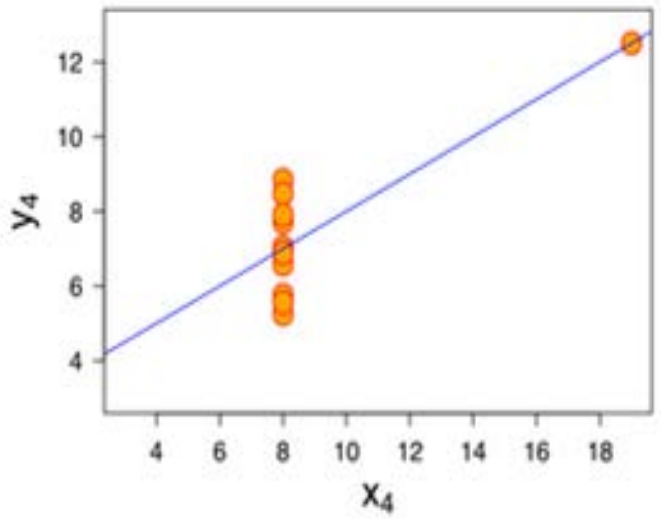
D. All of the above

E. None of the above









Anscombe's Quartet

4 different sets of data with identical correlation coefficients (.81).
What does this say about each data set? Not much.

data set 1		data set 2		data set 3		data set 4	
X1	Y1	X2	Y2	X3	Y3	X4	Y4
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

9.00	7.50		9.00	7.50		9.00	7.50		9.00	7.50	<< Mean
3.32	2.03		3.32	2.03		3.32	2.03		3.32	2.03	<< Standard Deviation
0.82			0.82			0.82			0.82		<< Correlation between x and y

All four datasets have same

Mean

Standard deviation, and

Correlation Values!

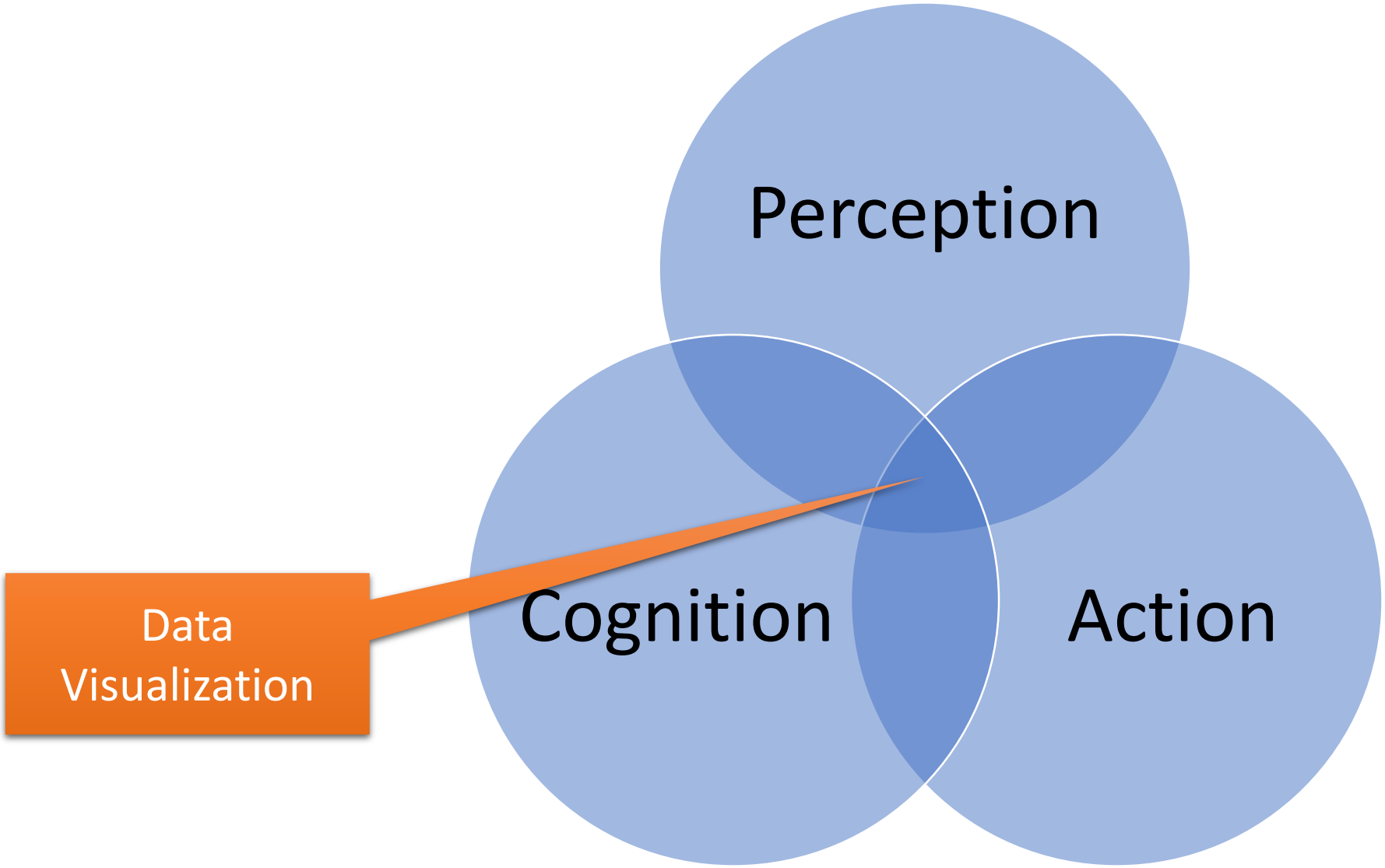
Edward Tufte

Cleveland

Visualization Excellence!

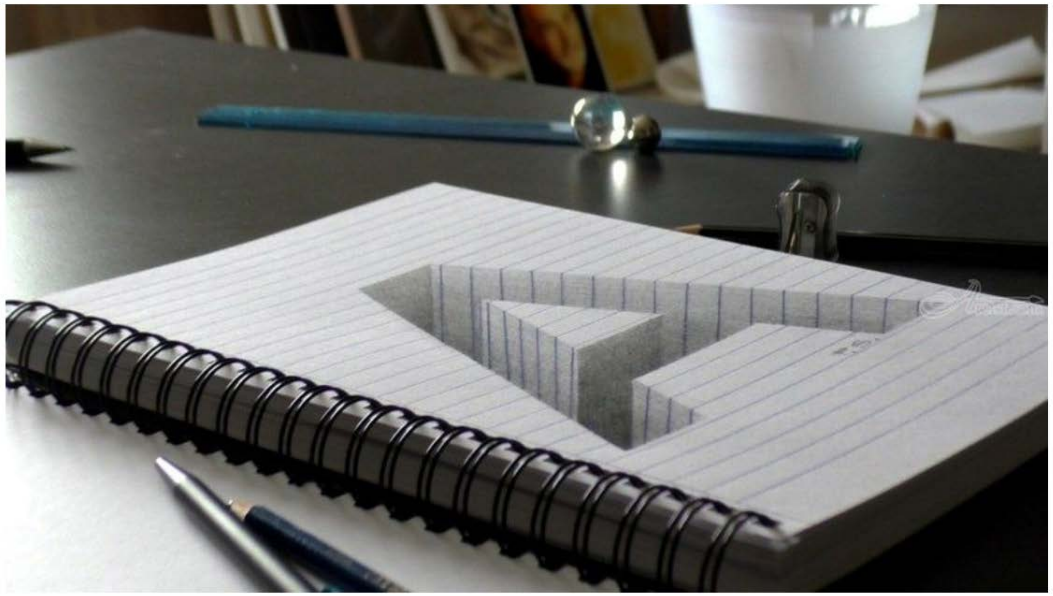
Is that which gives to the viewer the

- greatest number of ideas in the
- shortest time with the
- least ink in the
- smallest space.


Perception
Cognition
Action
Data Visualization

Ever wonder why the words ‘understanding’ and ‘seeing’ are synonymous

Drawing of a hole letter A in line paper/ 3D Trick Art Optical Illusion



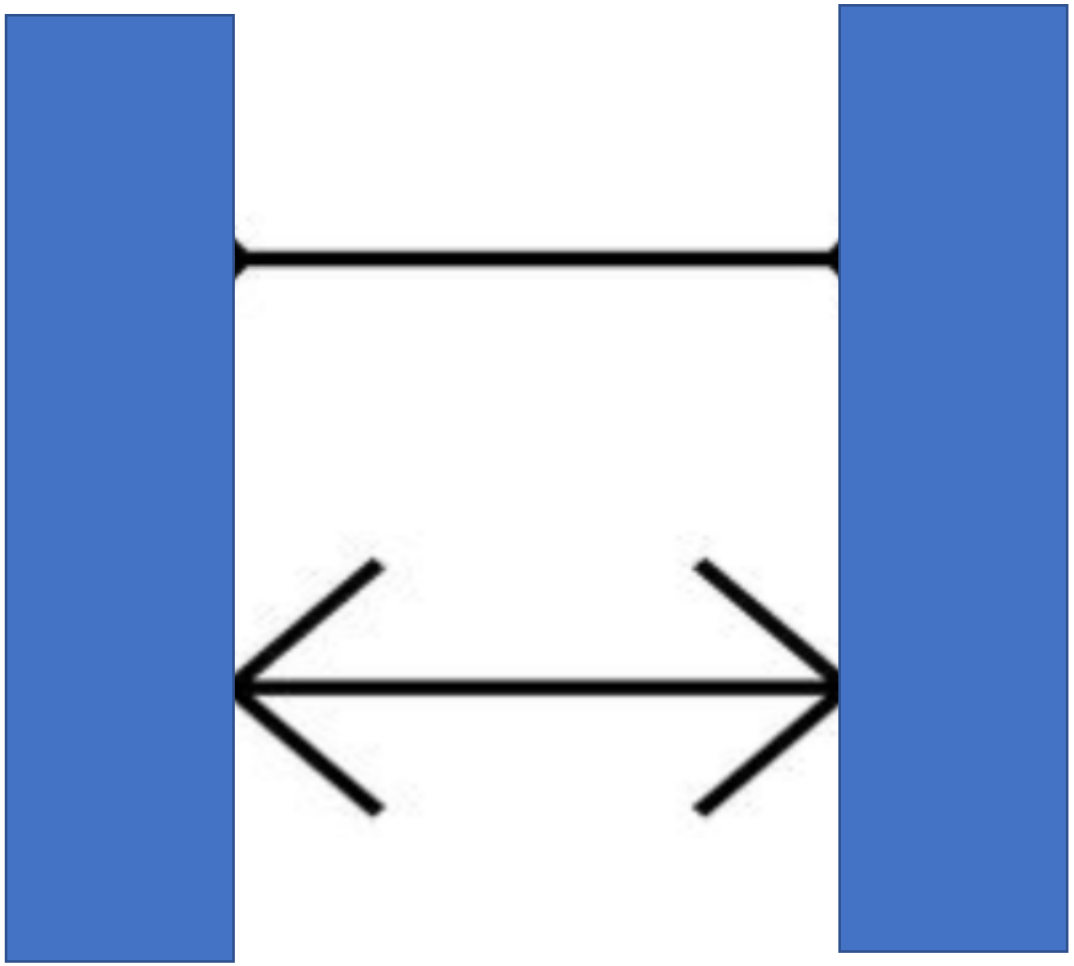
https://artist.com/stefan-pabst/drawing-of-a-hole-letter-a-in-line-paper-3d-trick-art-optical-illusion/art-details-91-7563/

Perception!

Which one is longer?

A

B



How much darker?

A

B

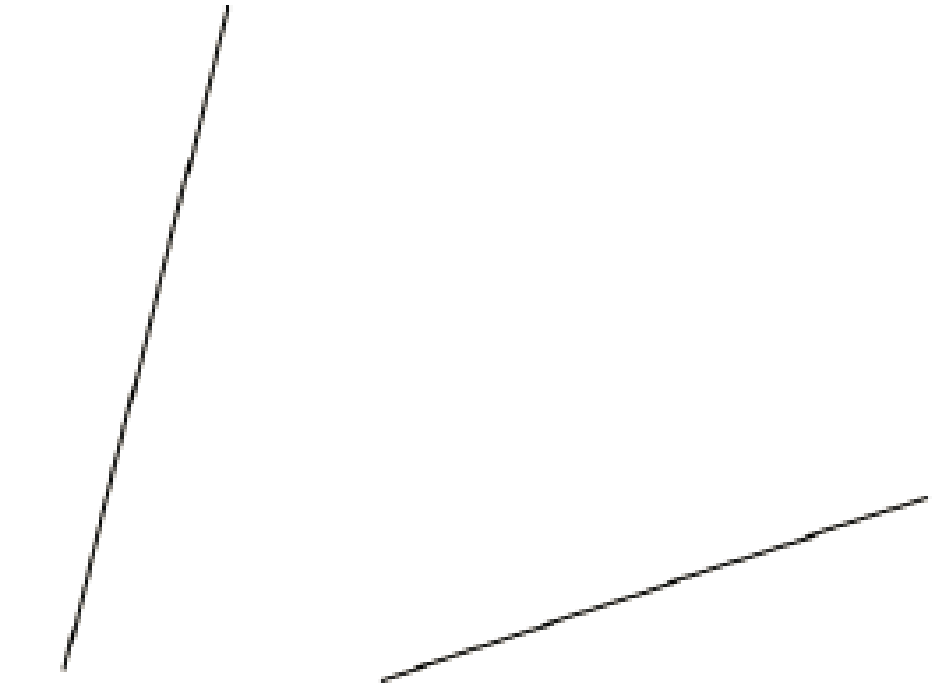
McKinlay 1986; Stephen's Power Law 1957

How much larger (area)?

A B

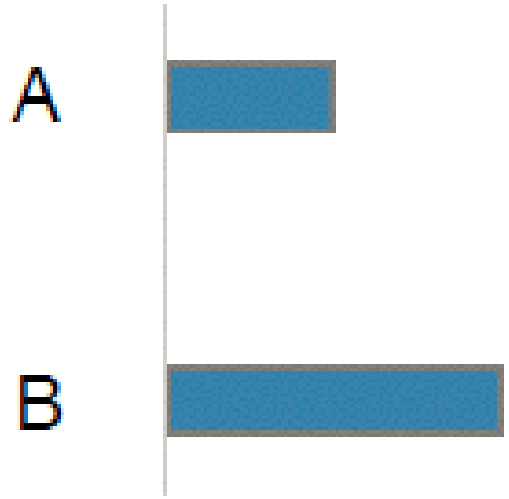
Cleveland and McGill 1984

How much steeper?



Cleveland and McGill 1984

How much longer?



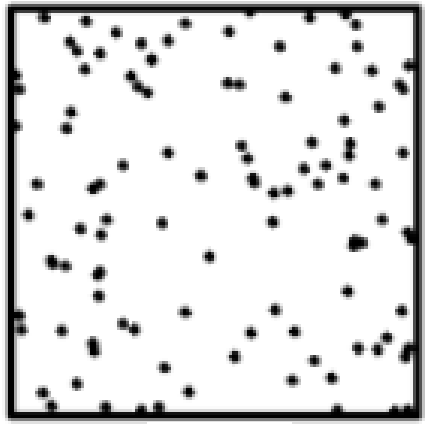
Cleveland and McGill 1984

Following perception based rules, we can present our data in such a way that the important and informative patterns stand out. If we disobey the rules, our data will be incomprehensible or misleading.

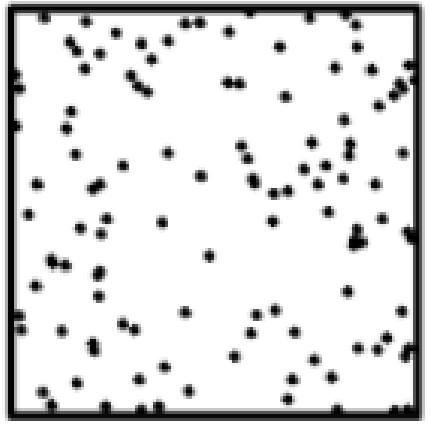
(*Information Visualization, Second Edition,* Colin Ware, Morgan Kaufmann Publishers, 2004, page xxi)

Experiment1

A

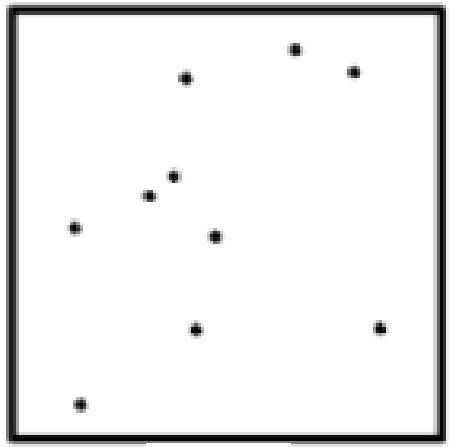


B

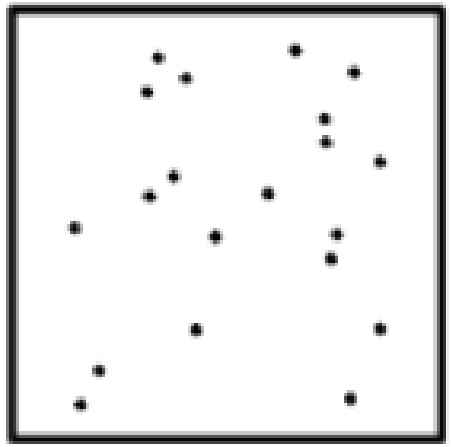


Experiment2

A

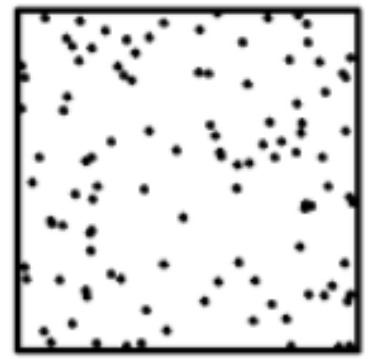


B



10

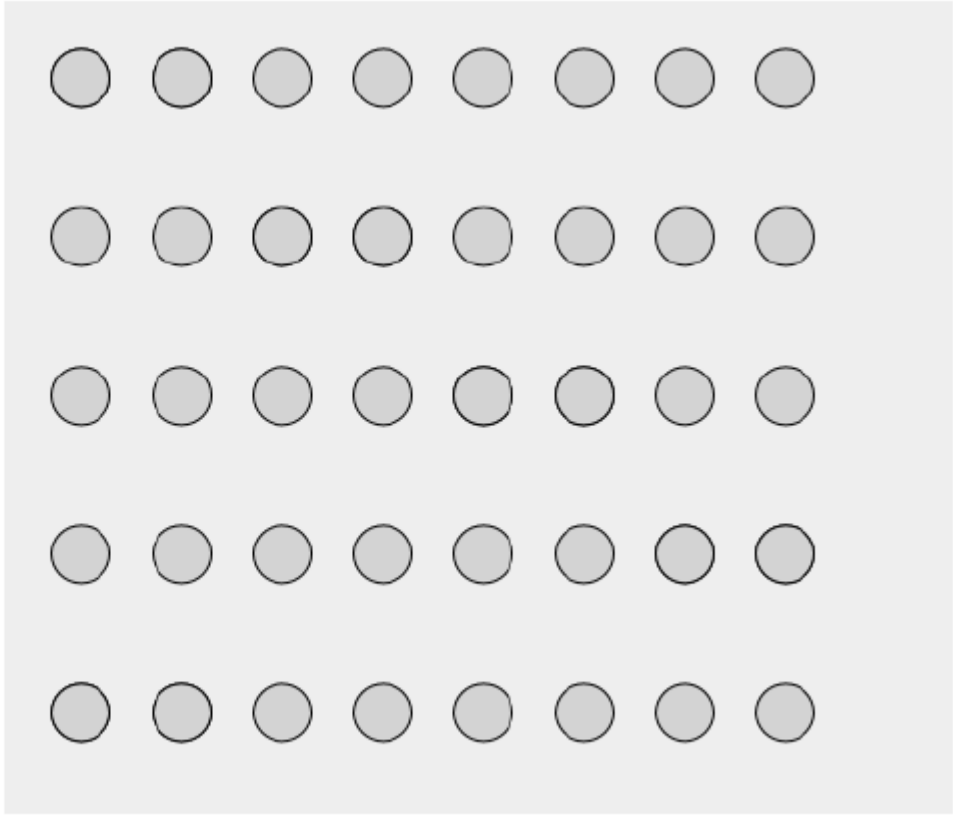
20



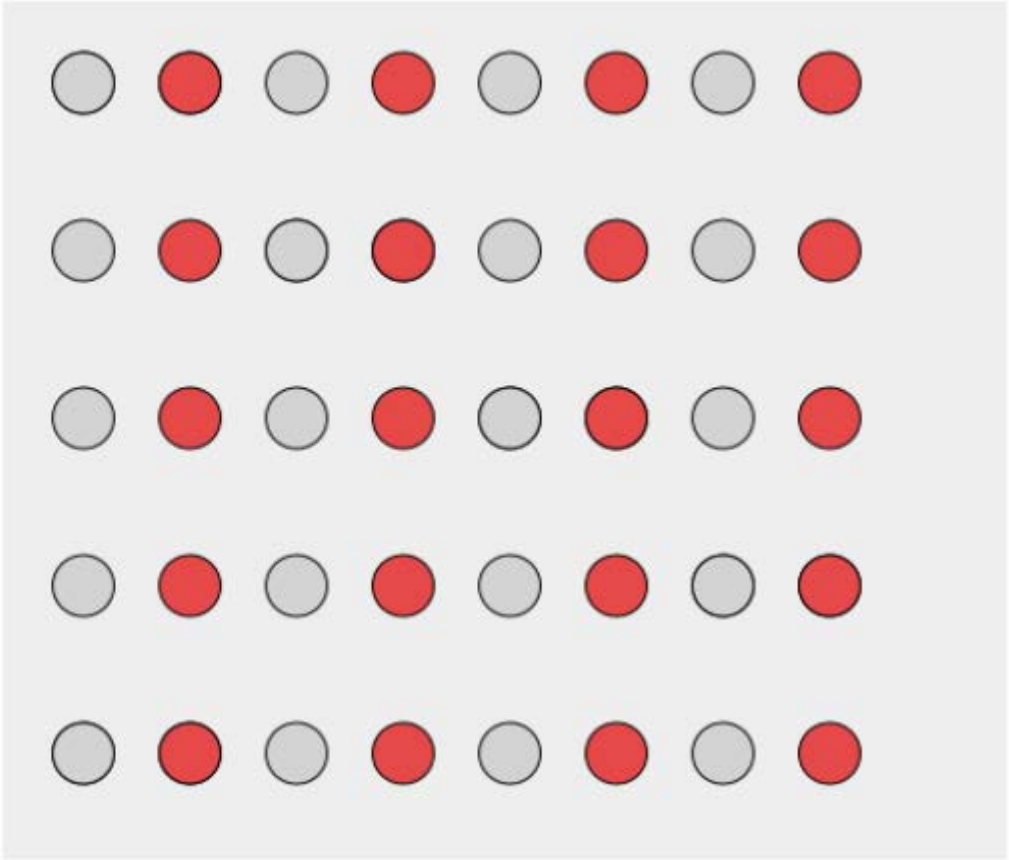
110

120

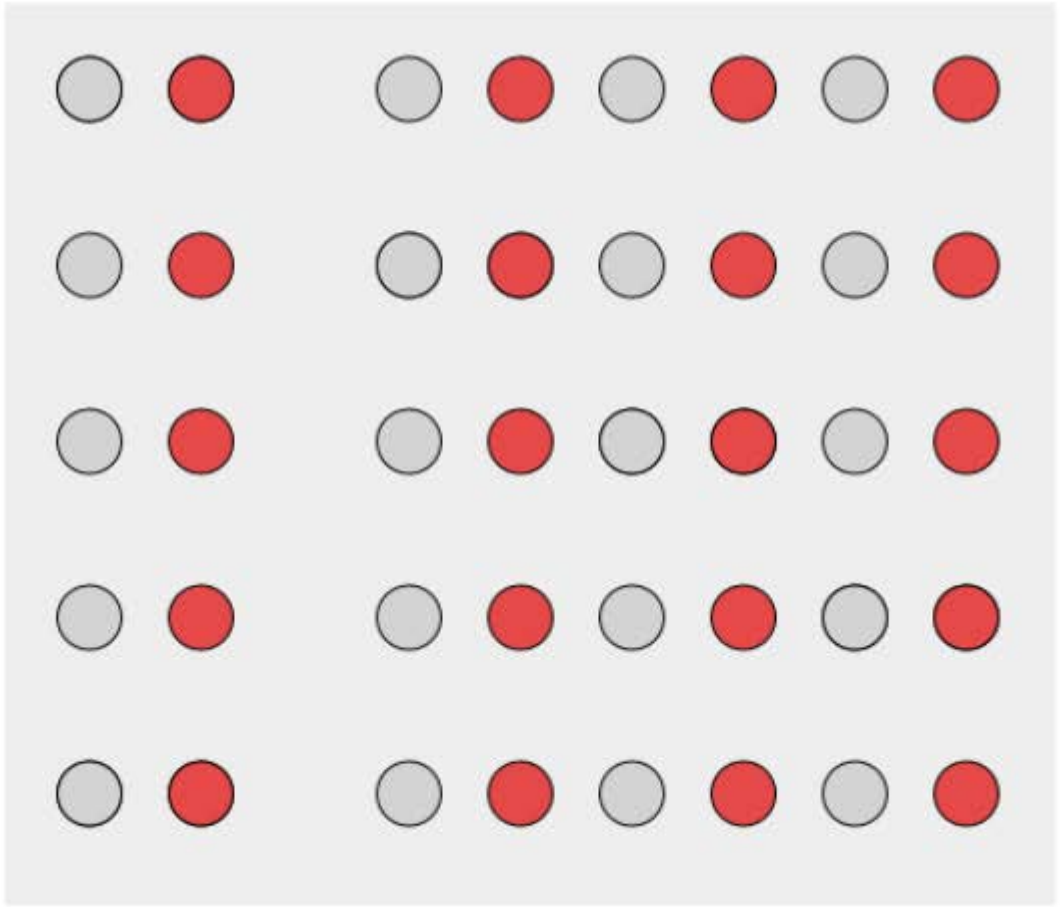
Do we see Rows or Columns?



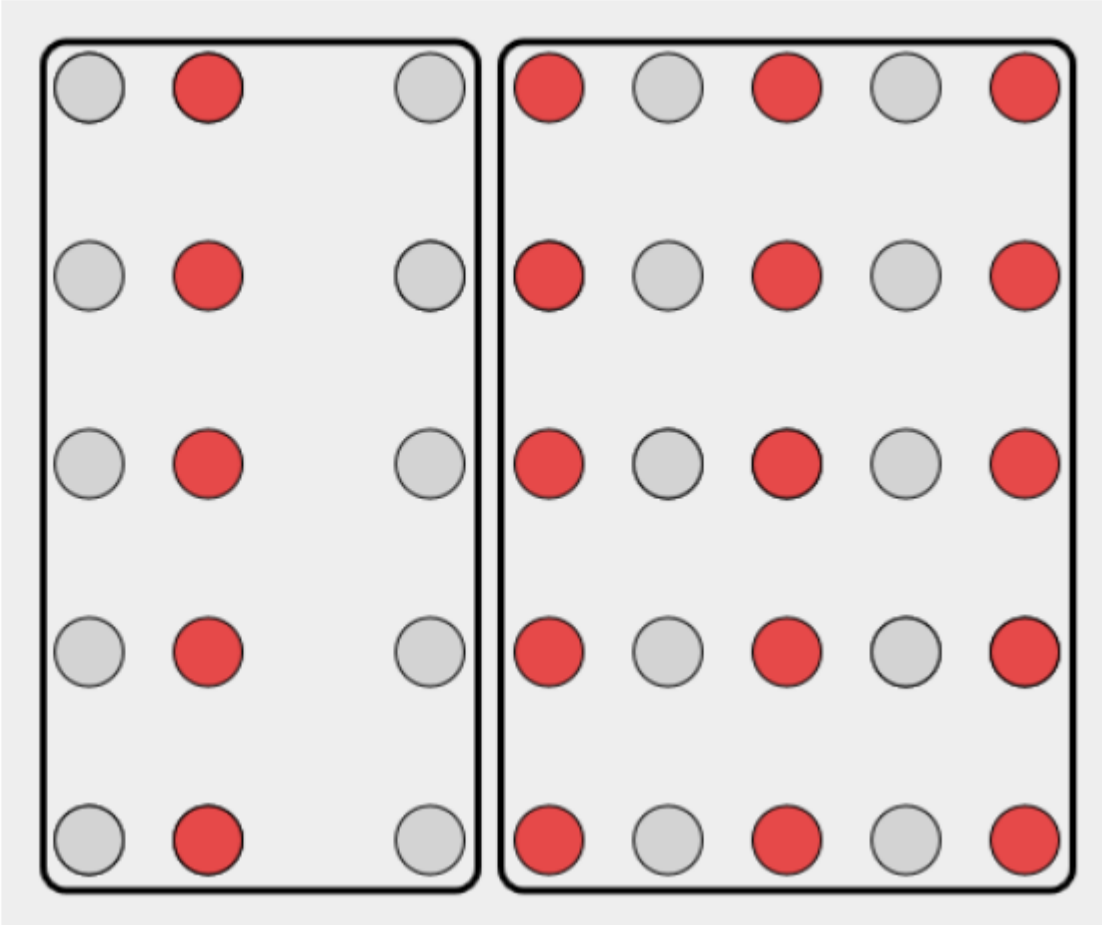
Do we see Rows or Columns?



What do we see now?



Well……What do we see now?



Red vs. Blue?

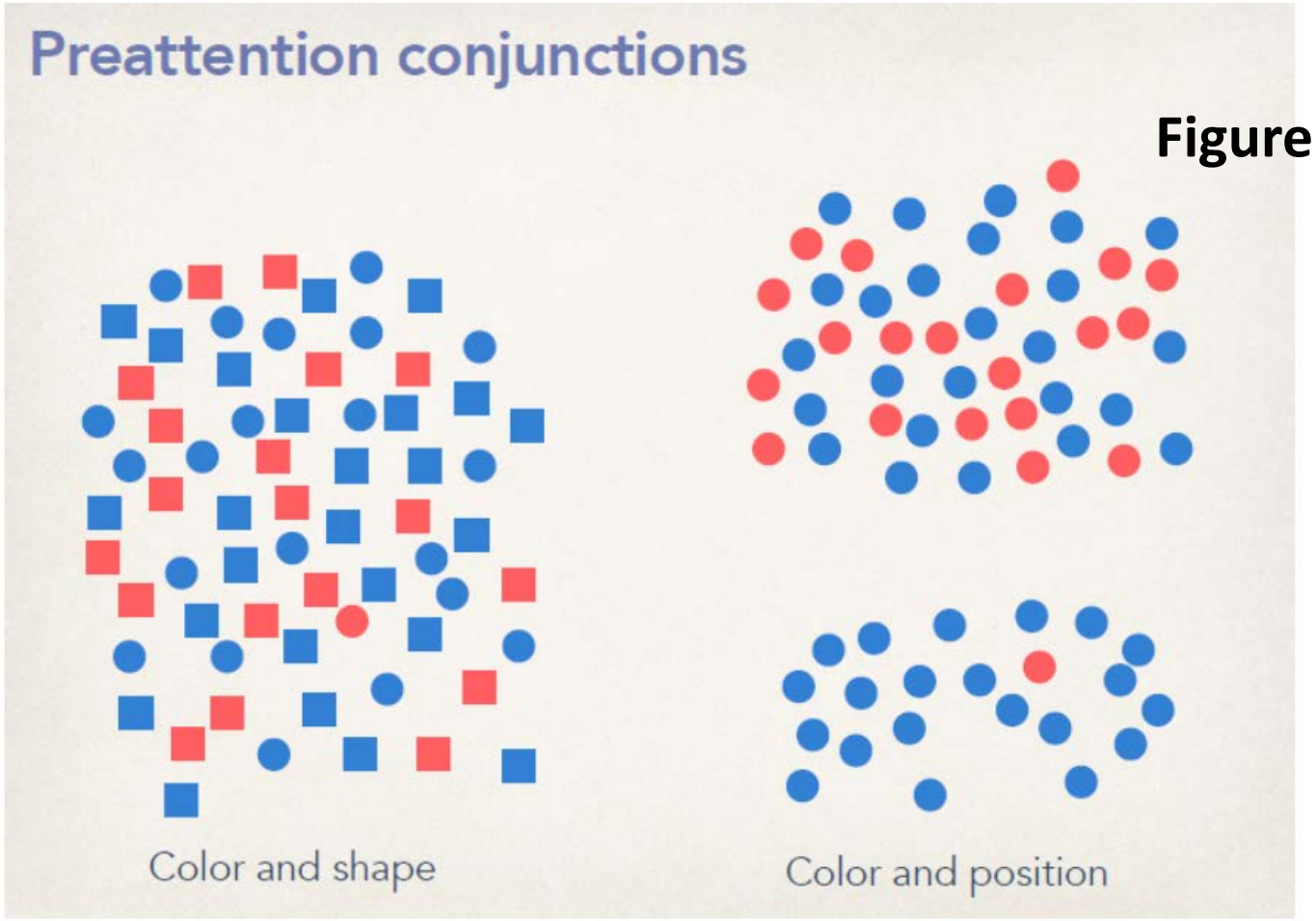
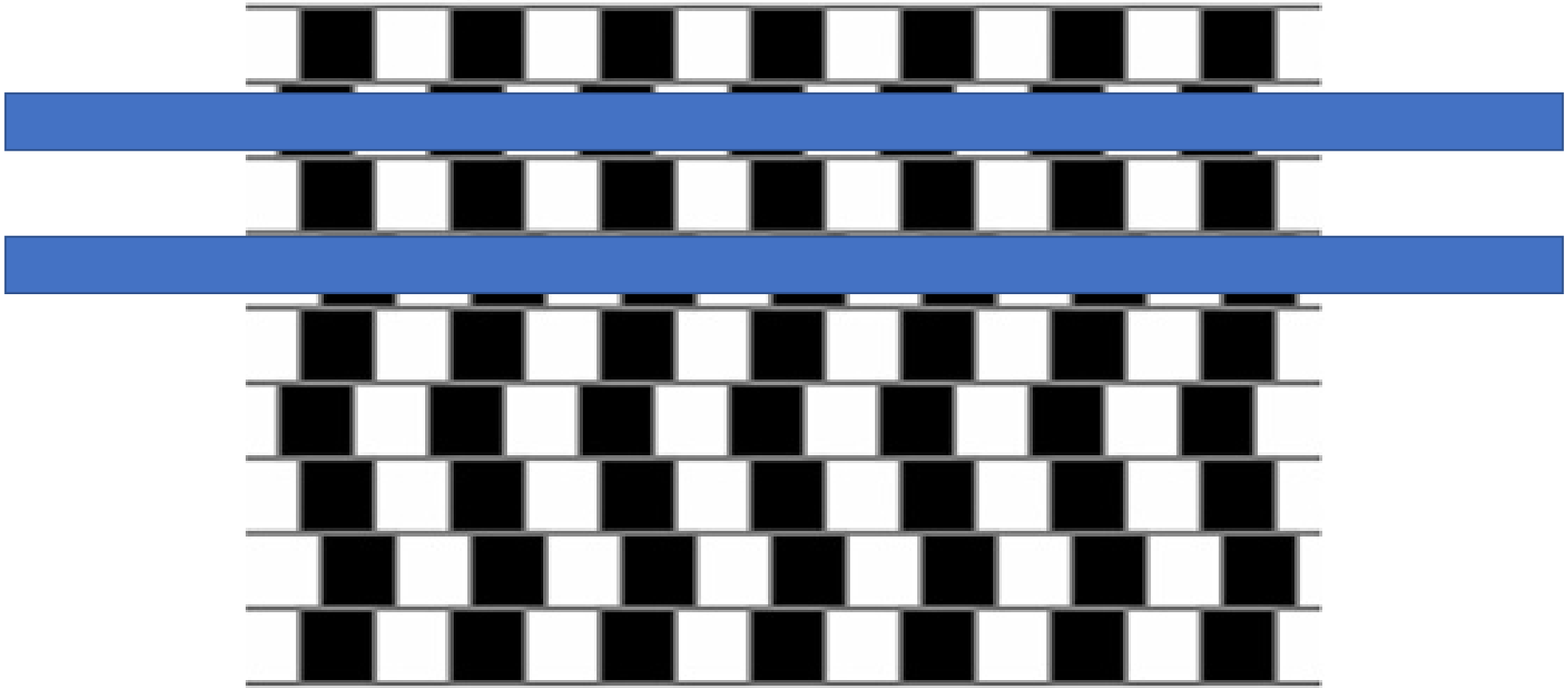


Figure C

Figure A

Figure B

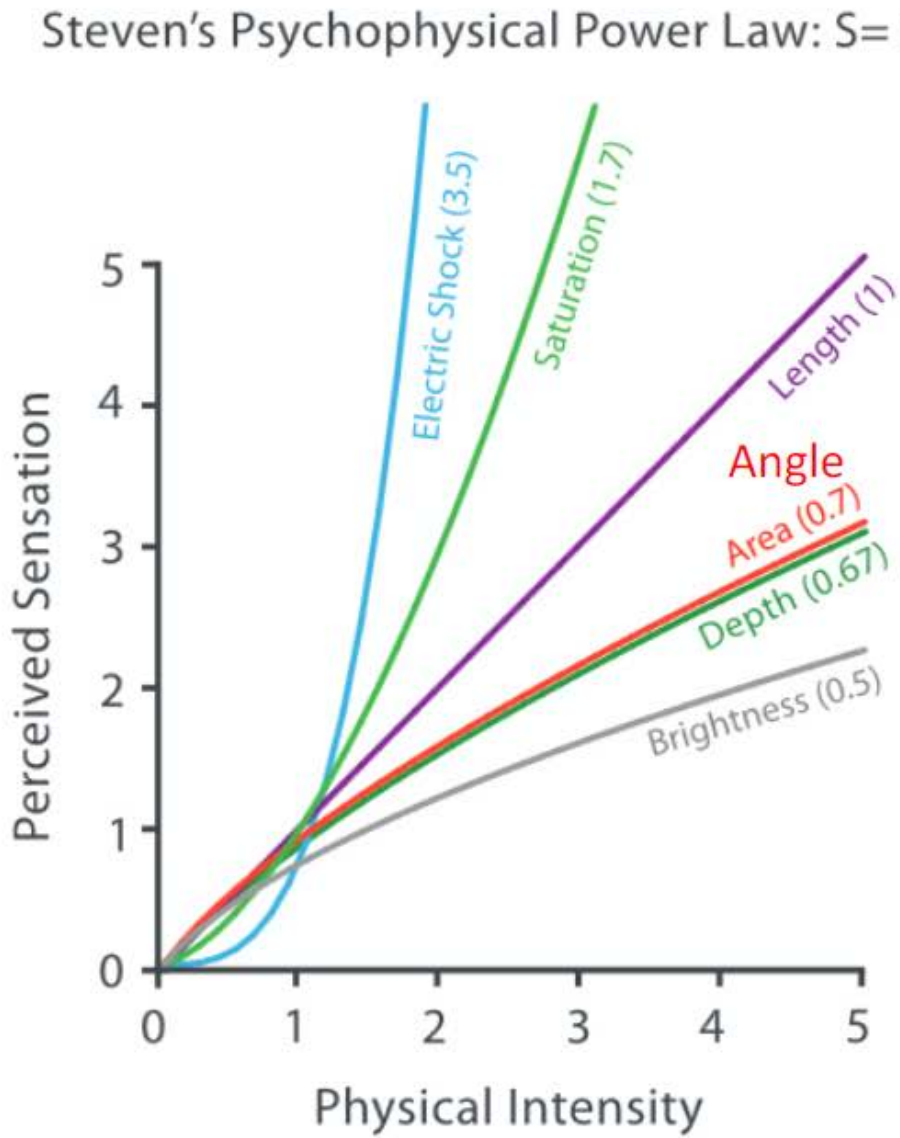


How many threes?

45929078059772098775972655665110049836645
27107462144654207079014738109743897010971
43907097349266847858715819048630901889074
25747072354745666142018774072849875310665

Perception applied to Graphs!

All Channels are not Created Equal



Steven's law suggests that we accurately perceive one unit change in length.

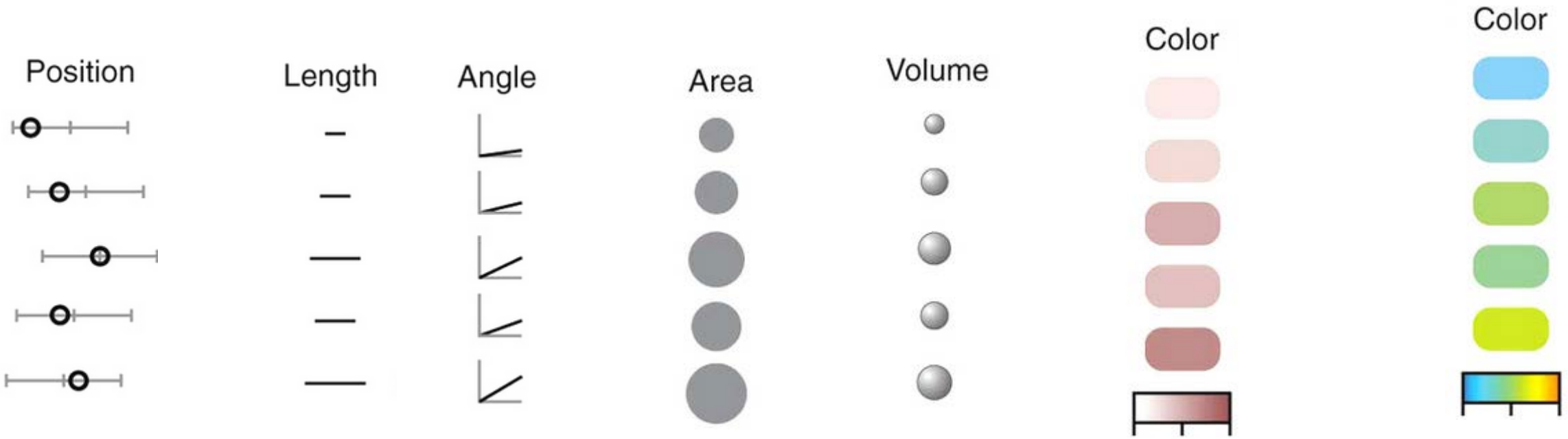
For other channels either we overestimate (e.g., N>1; electric shock) or we underestimate (e.g., N<1; brightness).

If area is doubled, we perceive the change to be little less than double. Similarly, when brightness is doubled, we perceive the change to be little less than double as well. (Our senses are better at judging changes in *area* than they are at judging changes in *brightness.*)

However, we perceive one unit change in electric shock as more than one unit change .(this is not at all surprising to me!!)

Note: Angle (slope) is also underestimated!

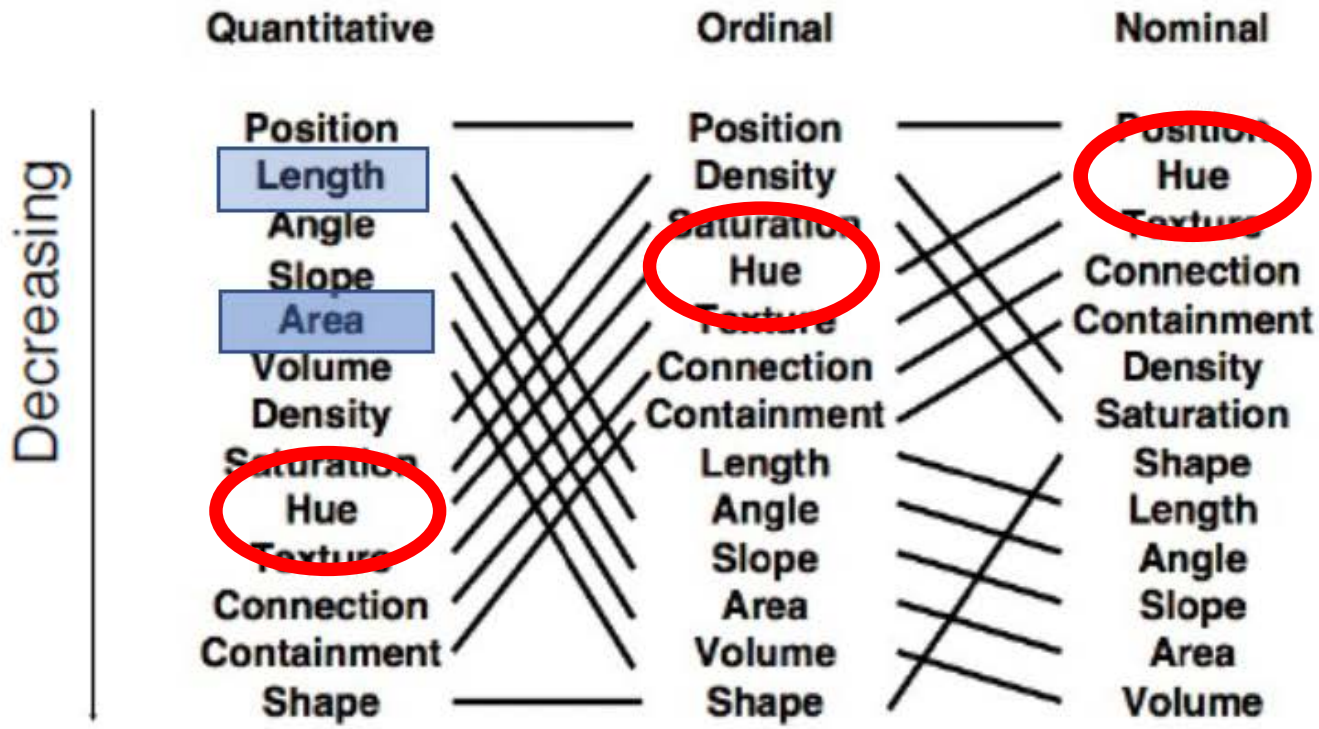
Effectiveness



Cleveland and McGill 1984
Mackinlay 1986

Channel Rankings

Jock Mackinlay, 1986

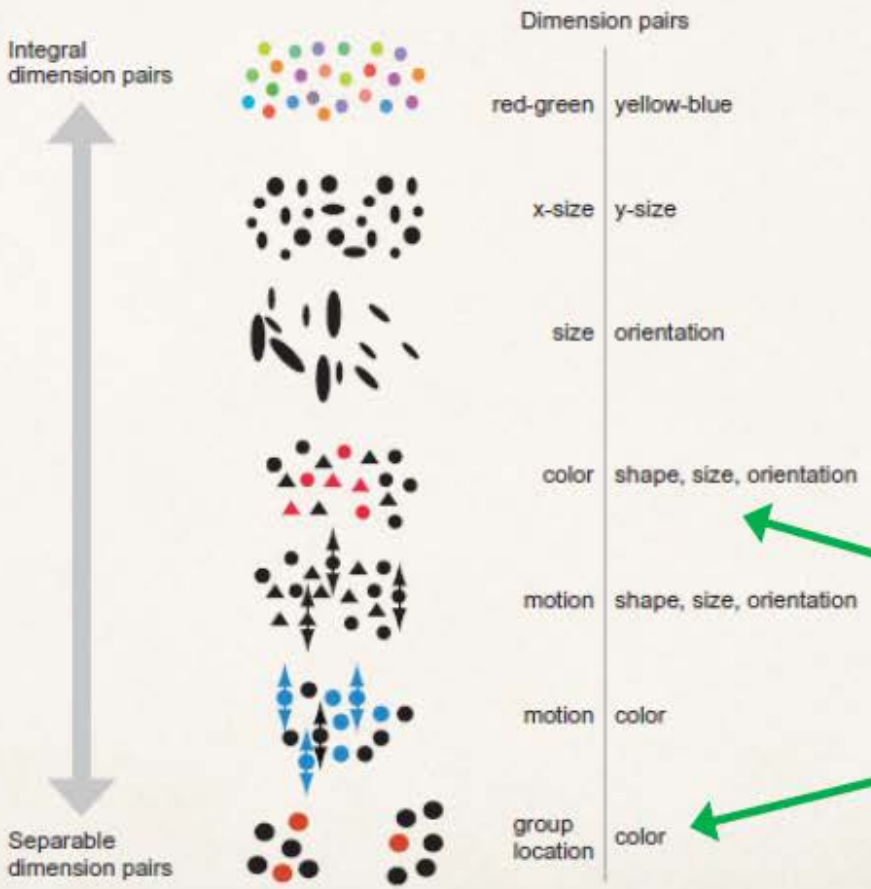


Not all channels are equally effective all the time!

For instance, for quantitative variables length (of a bar in a bar chart) is better than area of the pie chart!

[Mackinlay, Automating the Design of Graphical Presentations of Relational Information, 1986]

Integral - Separable attributes



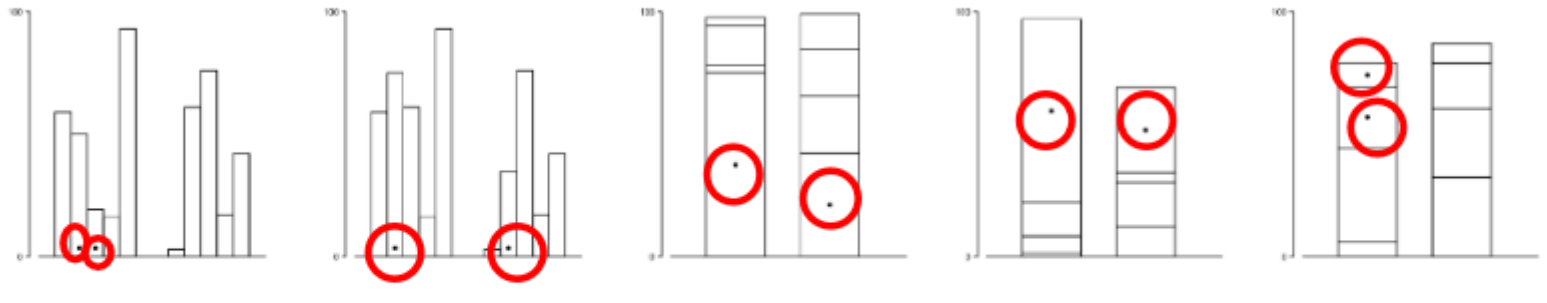
Although early research suggested that conjunction searches were never pre-attentive, it has emerged that there are a number of pre-attentive dimension pairs that do allow for conjunctive search.

Example: **color and position** are easier to separate than **color and shape**.

Weber law suggests that we can sense values visually. More specifically it states that size of Just Noticeable Difference (JND) is proportional to the intensity of the stimulus.

$$\frac{\Delta I}{I} = K$$

The Weber Fraction



(a) Adjacent Bars (Type 1) (b) Separated Bars (Type 3) (c) Aligned Stacked Bars (Type 2) (d) Unaligned Stacked Bars (Type 4) (e) Divided Bar (Type 5)

Fig. 1: The five bar chart tasks studied in Cleveland & McGill [1]. Study participants were asked to estimate the height of the shorter marked bar as a percent of the taller marked bar. Cleveland & McGill's ranked these tasks from lowest error (Type 1) to highest error (Type 5).

In their seminal paper, Cleveland and McGill (1984) showed that comparisons made using two adjacent bars are more accurate than those made using separated bars; refer to the dots that are being compared in figure **(a)** to the dots that are being compared in figure **(b)**.

Figures **(a)** through **(e)** have been ranked in the descending order of accuracy. When comparing the areas associated with the "two dots" (as shown in each figure respectively), **(a)** is preferred to figure **(b)**, and so on.

Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods

William S. Cleveland; Robert McGill

Journal of the American Statistical Association, Vol. 79, No. 387 (Sep., 1984), 531-554.

Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28198409%2979%3A387%3C531%3AGPTEAA%3E2.0.CO%3B2-Y

Journal of the American Statistical Association is currently published by American Statistical Association.

I encourage you to read the Cleveland and McGill paper. It is the same paper reference by Edward Tufte and Stephen Few.

This paper is available as an optional reading in the online course.

Design Principles

Tufte!

- "Get your viewers out of the puzzle-solving business."
- "Good design cannot rescue bad content."
- "Good design is clear thinking made visible. Bad design is stupidity made visible."
- "We must study disinformation and trickery and then reverse those techniques for the good of truth telling."
- “Tell the truth about the data. “
- “Above all, do no harm”

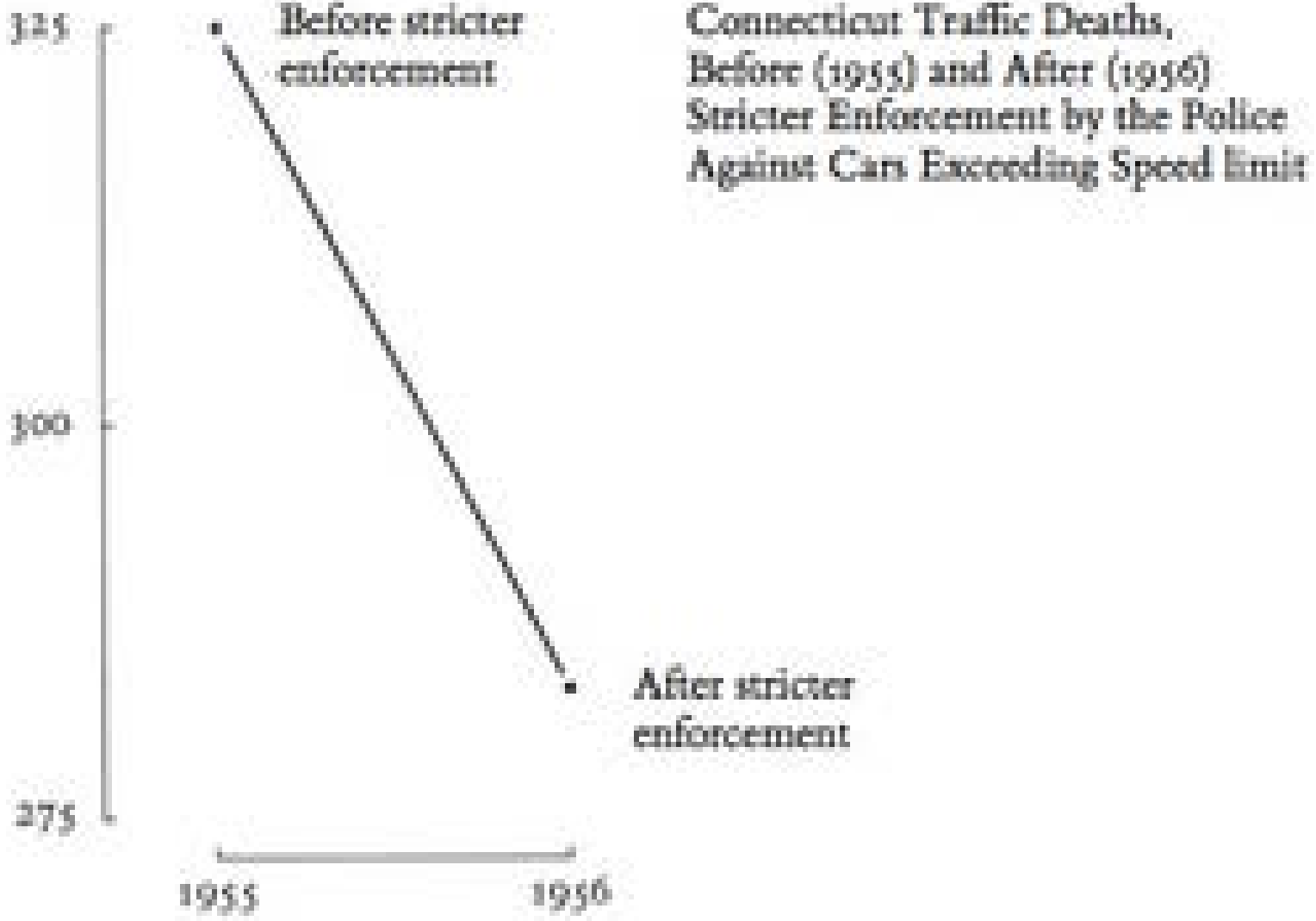
Tufte!

- Facilitate Comparison
  - Context
  - Small Multiples
  - Multivariate – more information for easy comparison
  - Quantitative Correlations
- Do not distract
  - High Data Ink Ratio
  - Lie Factor Close to 1: Show Data Variations and not Design Variations
  - No Chart Junk

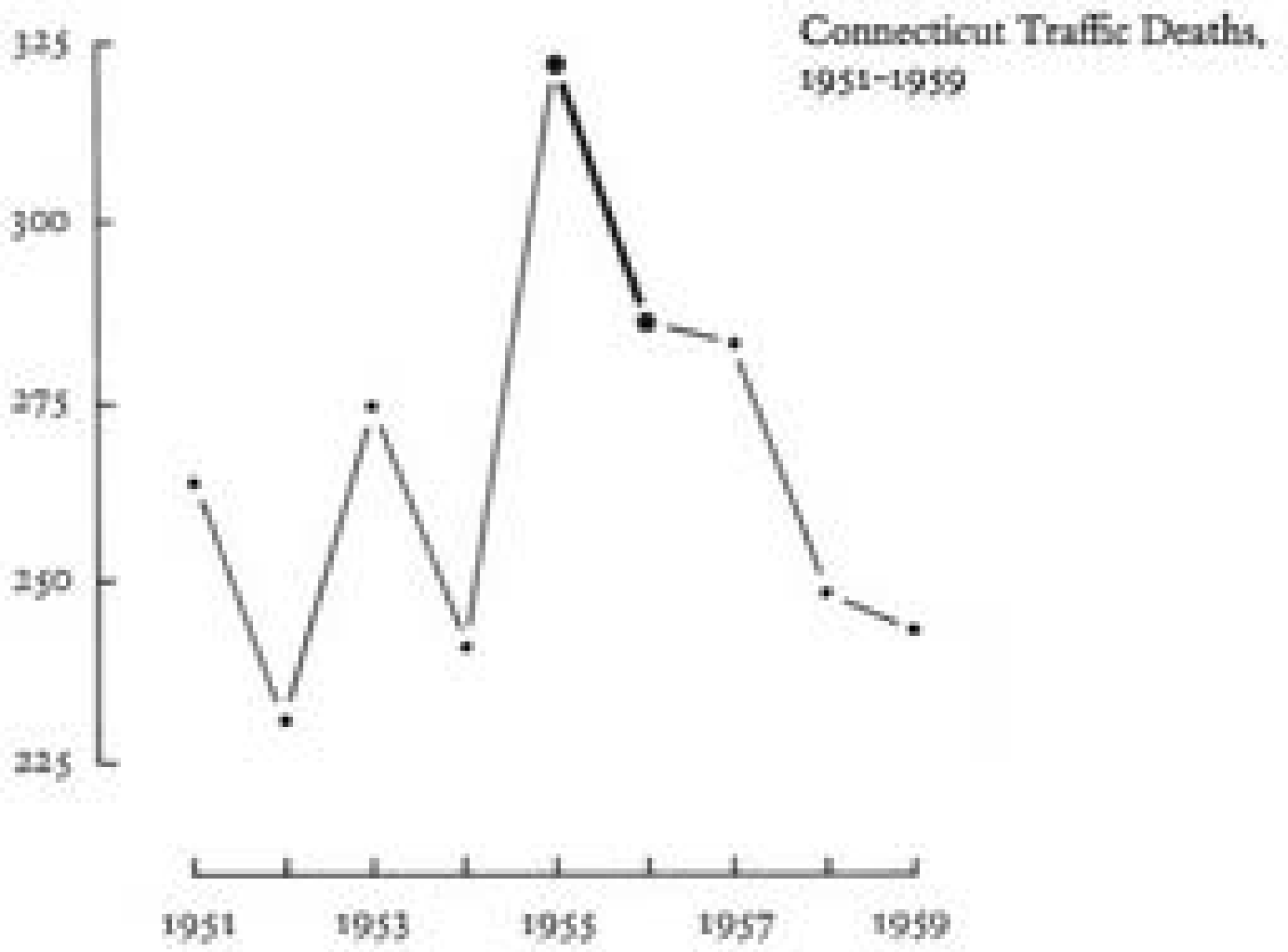
Aim of good data graphics is to show data

- Clearly
- Accurately
- Both (CORRECT)

Which one is better and why? (A)

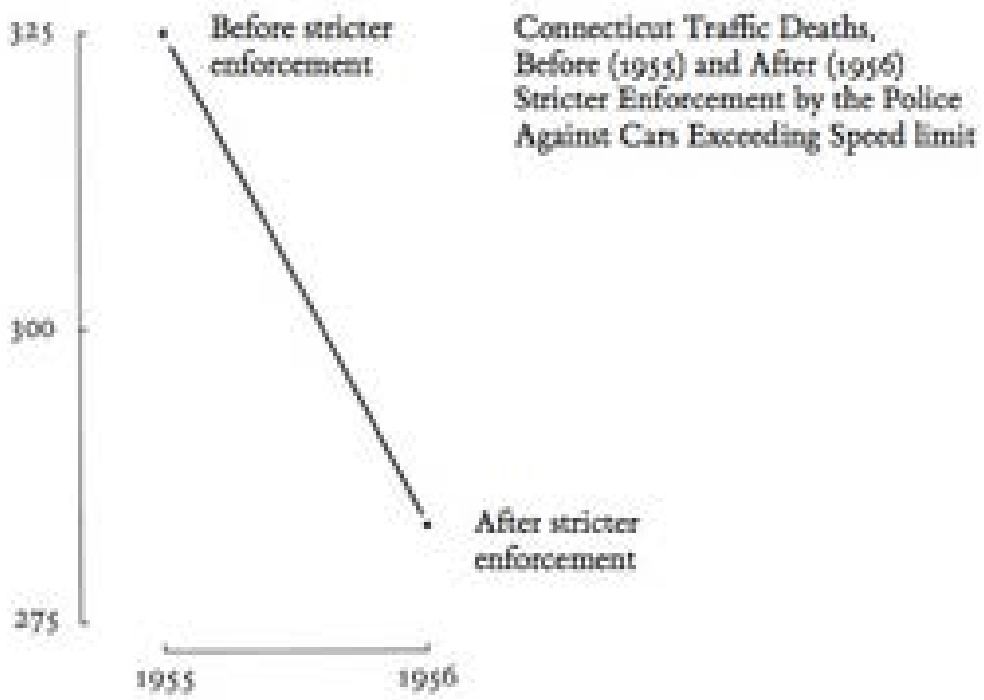


Which one is better and why? (B)

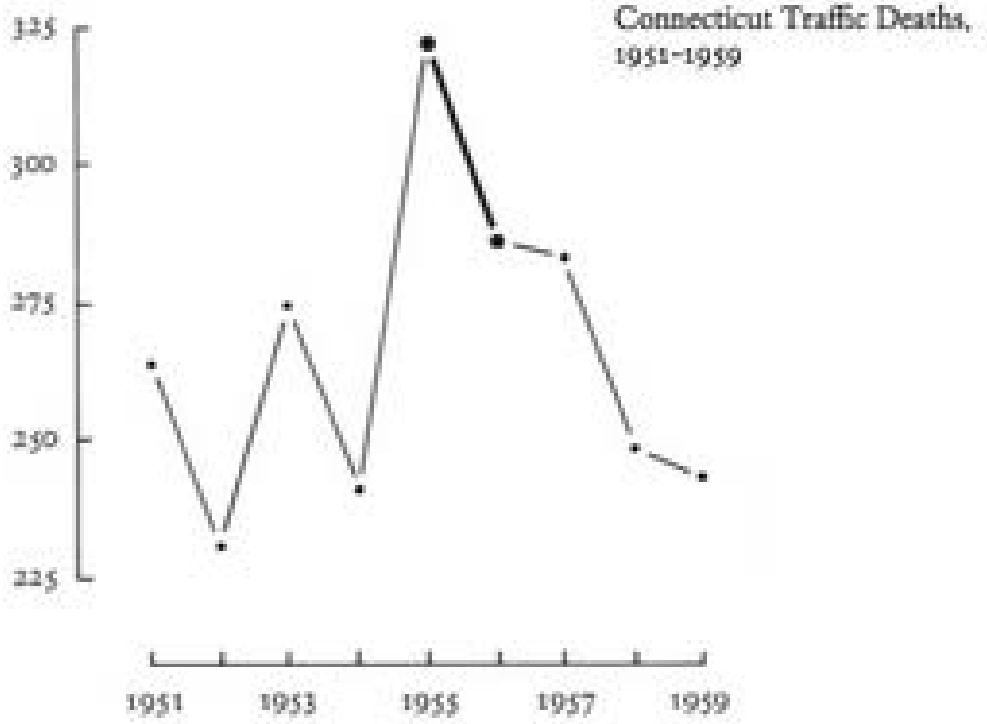


Design Principle	Context

A

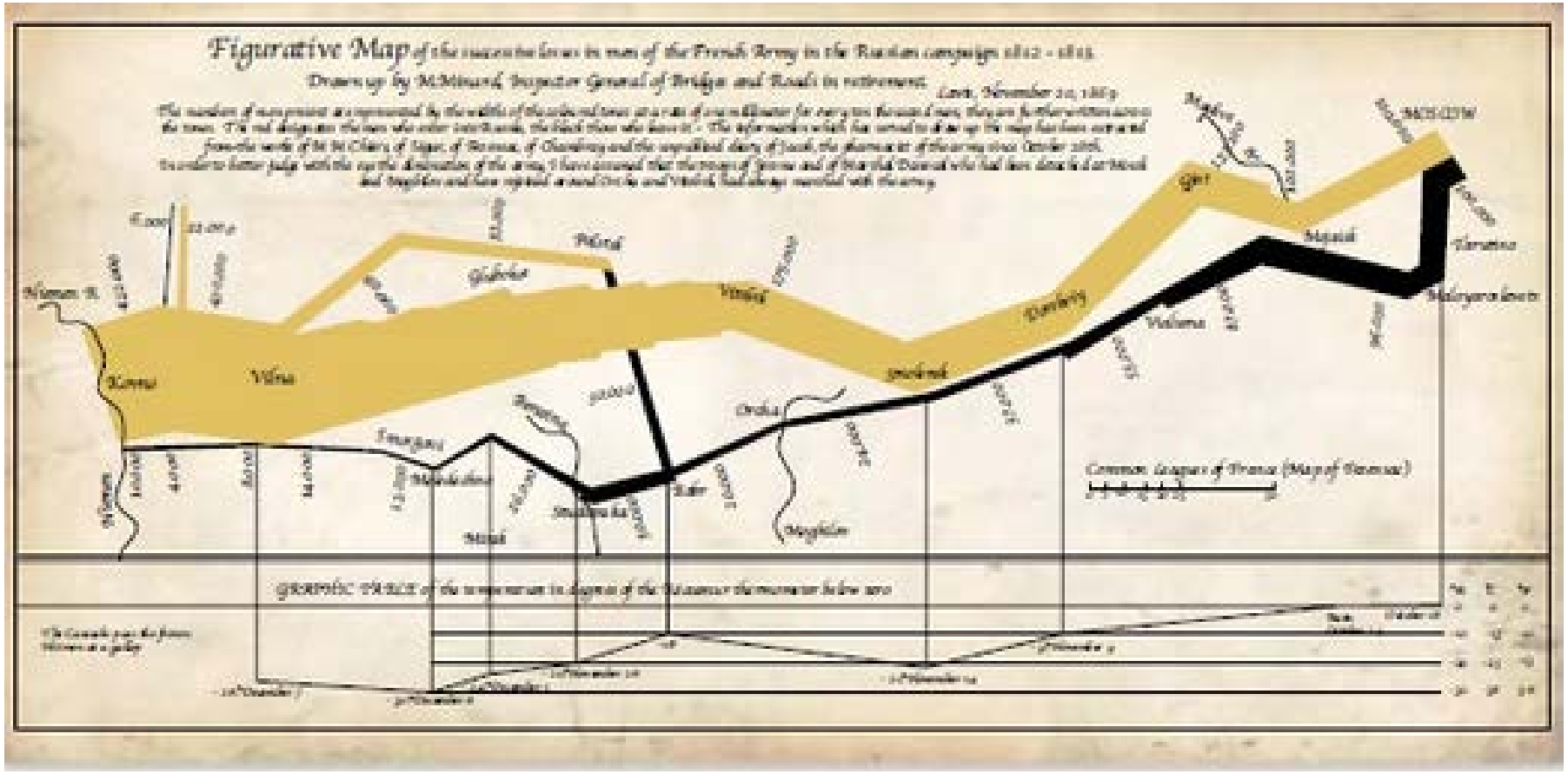


B



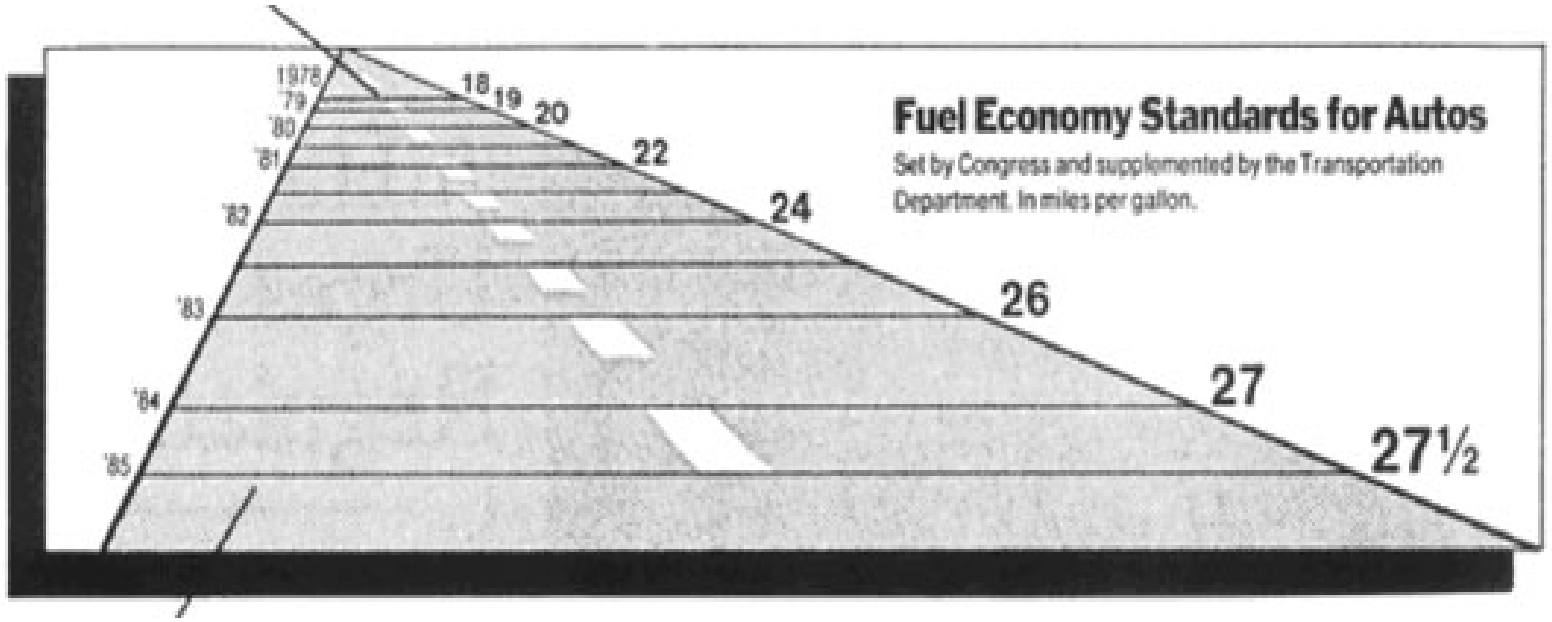
Design Principle | Multivariate

This graph appears in at least two of the Tufte's books. Tufte classifies it as "best statistical graphic ever made".

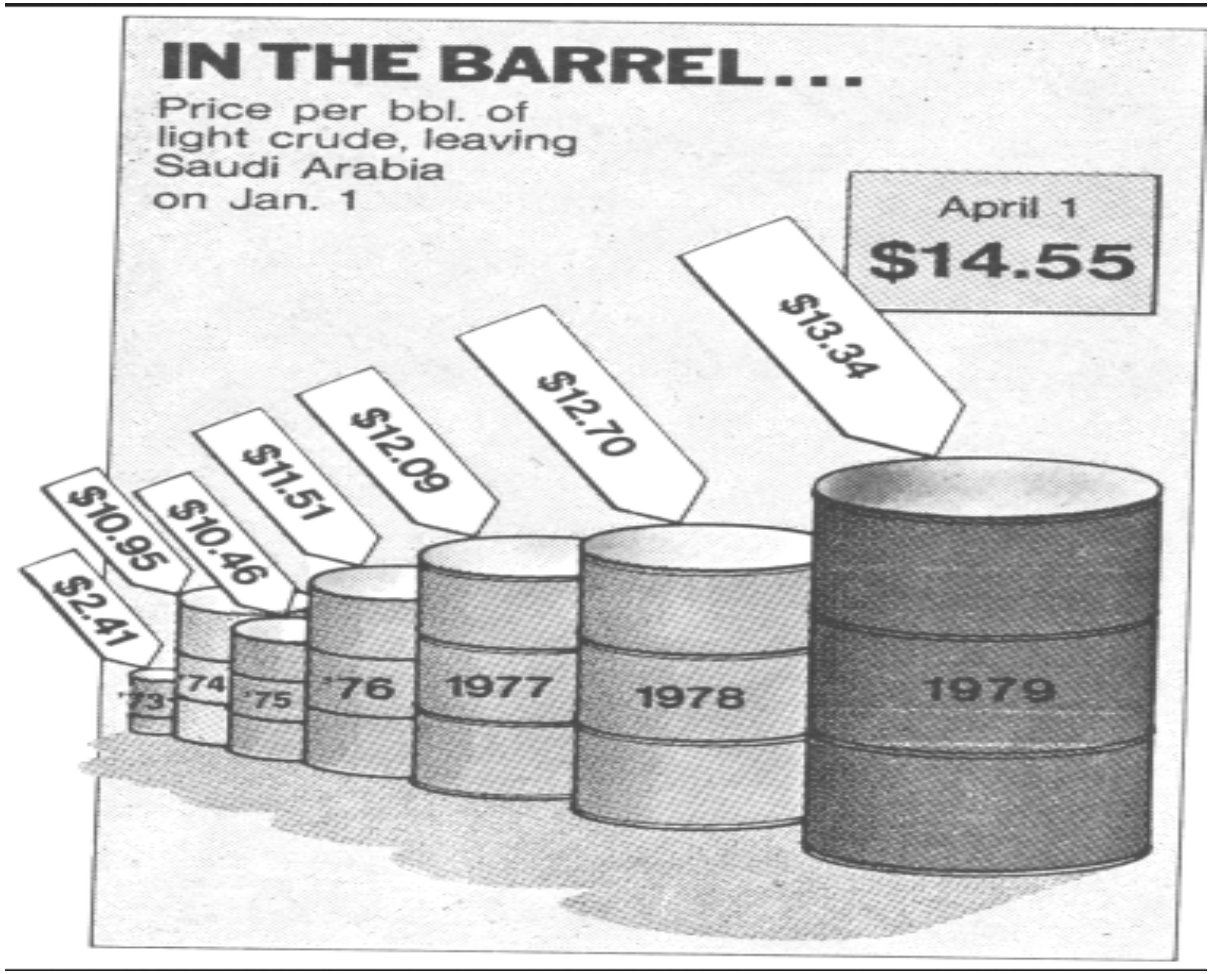


Design Principle | Lie Factor

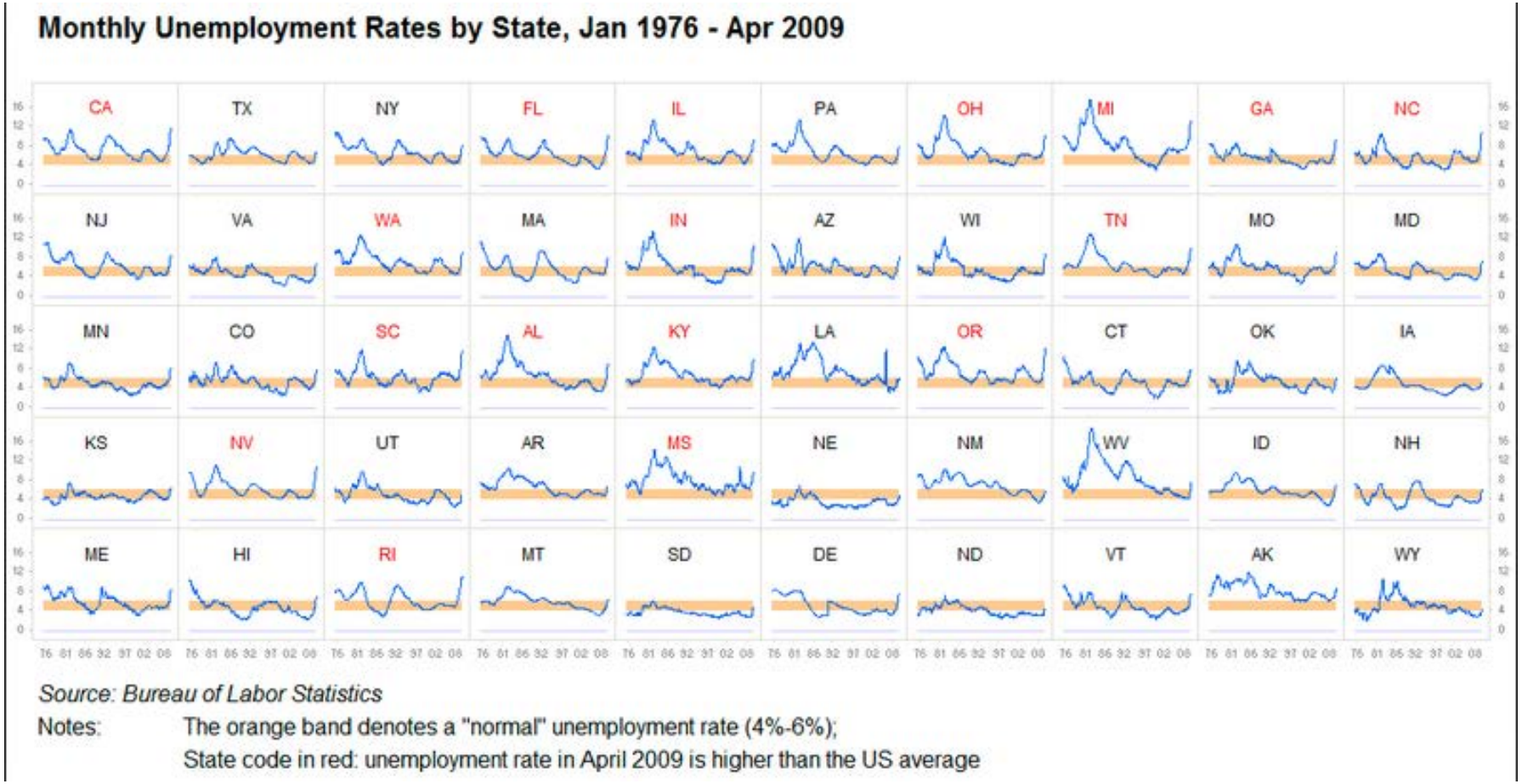
What is wrong with the following graph?



Is lie factor more than 1?



Small Multiples: What are they good for?

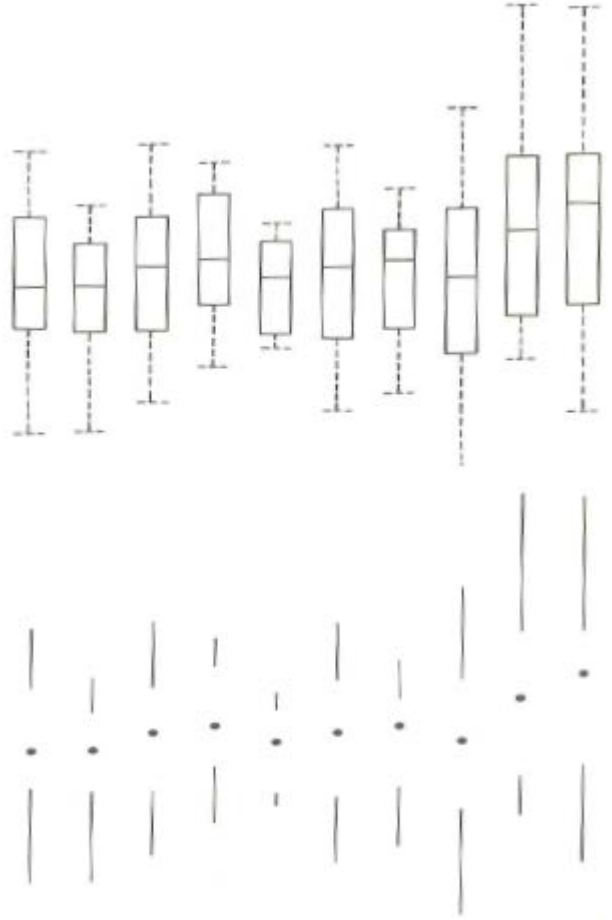


Design Principle	Small Multiples

Limited Short Term Memory!

To maximize data-ink ratio, one can

a. erase non-data ink
b. erase redundant data-ink



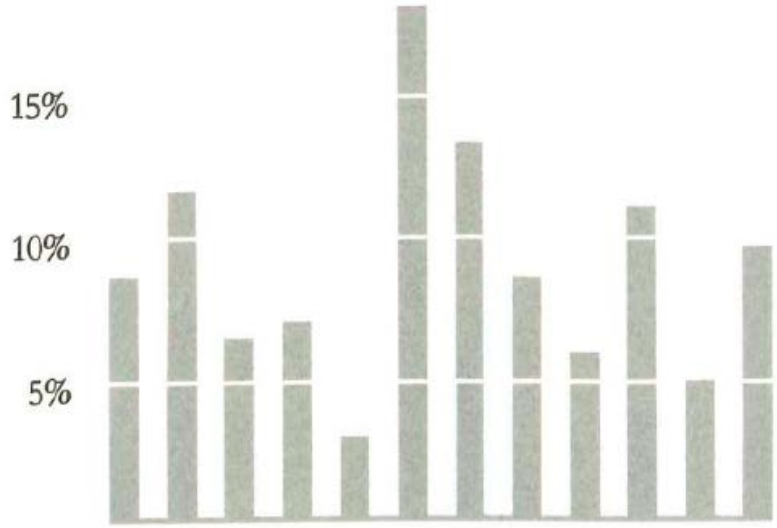
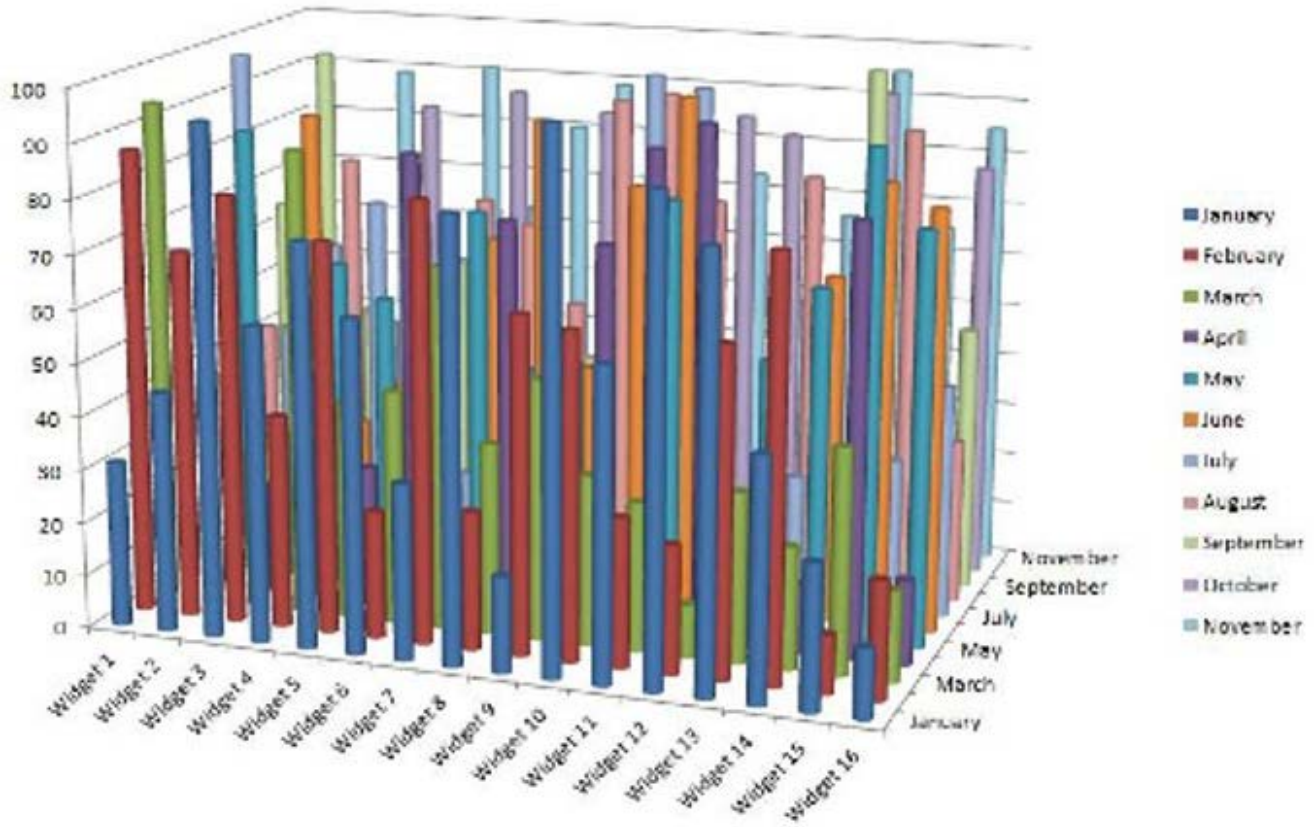


Chart junk!



Visual Display of Quantitative Information

Pick the Best One!

A

B

C

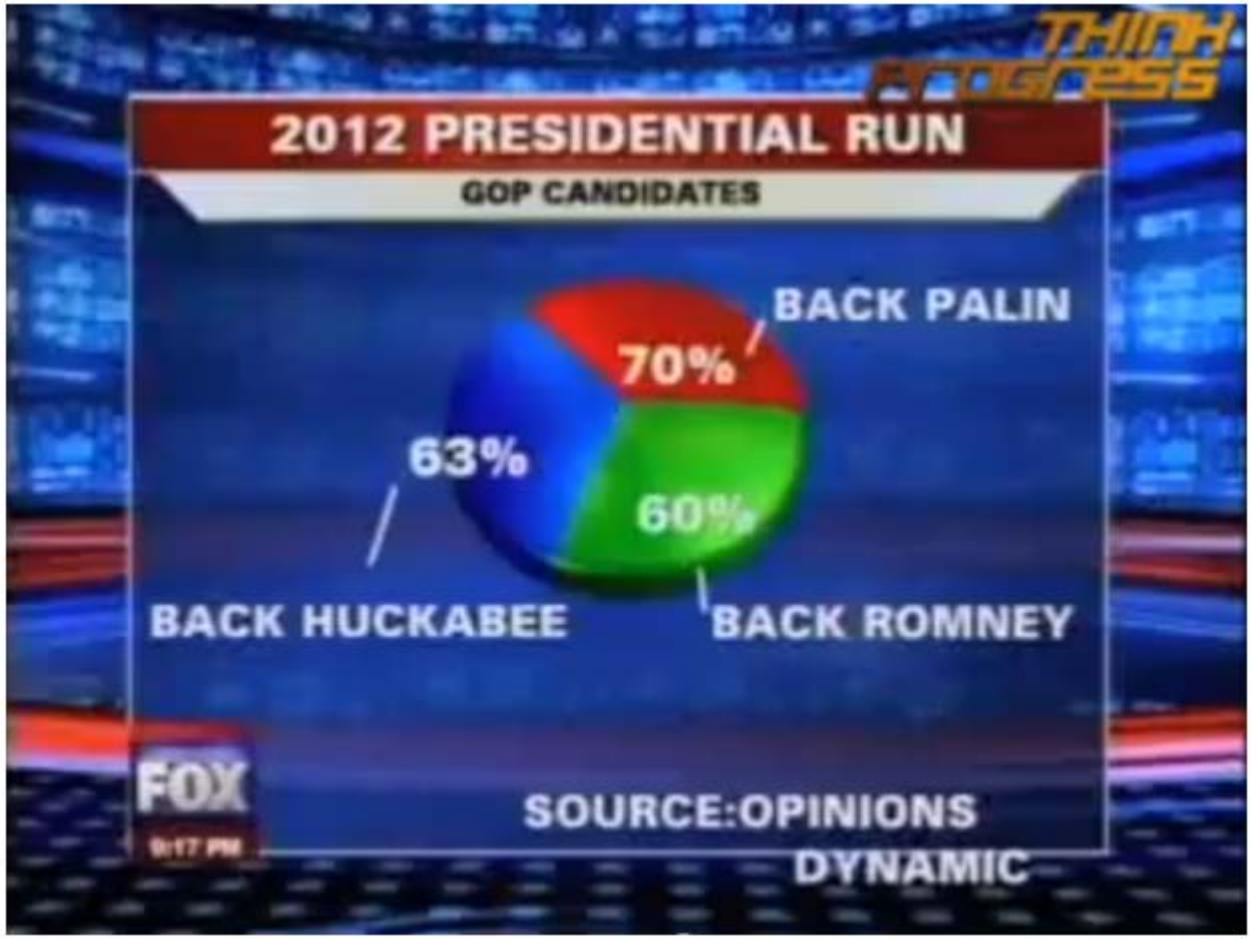
A

B

C

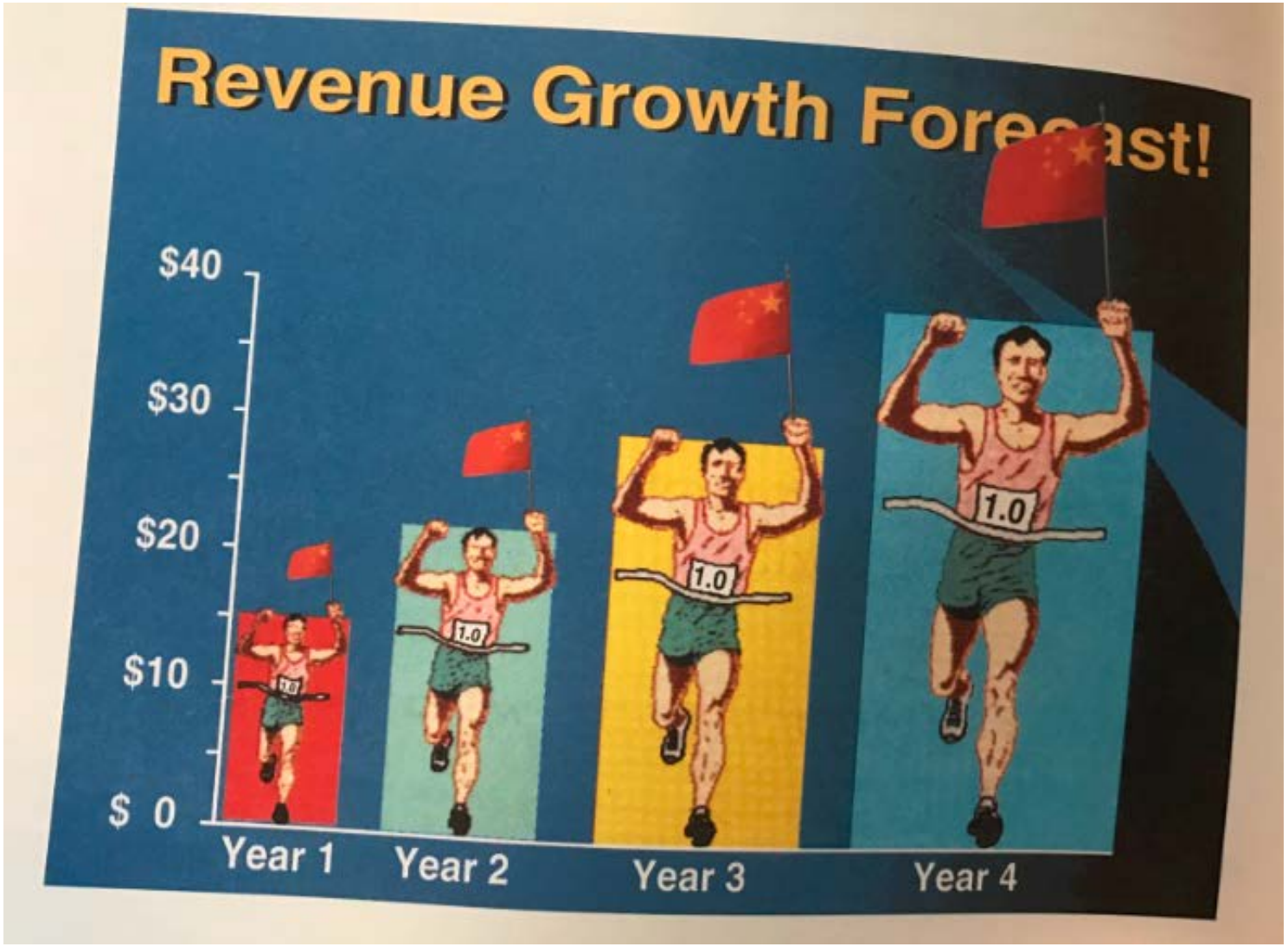
Design Principle | Avoid Areas

Pie Charts Should Add up to 100



Design Principle | Avoid Areas

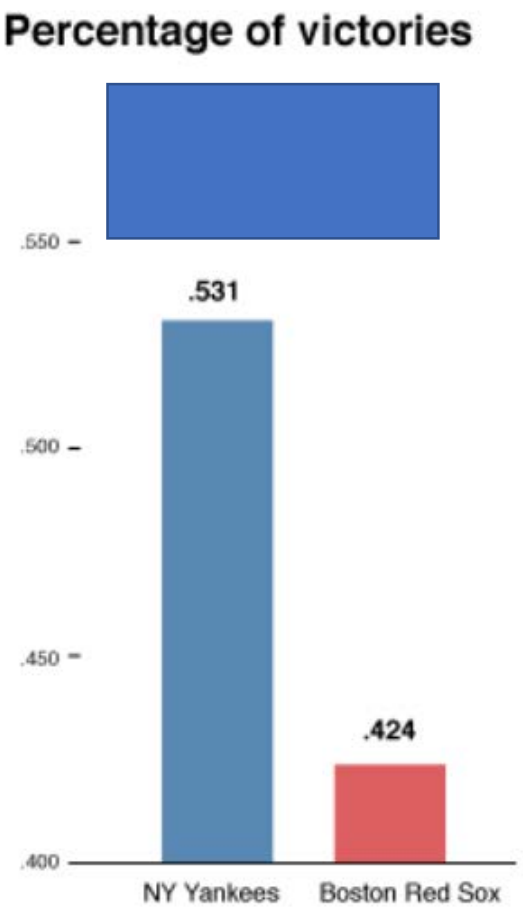
Areas and Volumes Create Misrepresentation

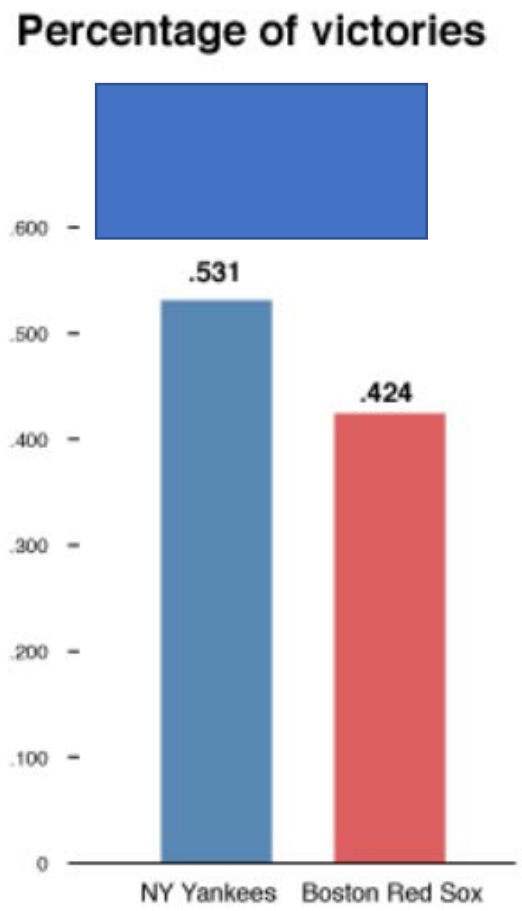


Normalization

Design Principle | Y-Axis

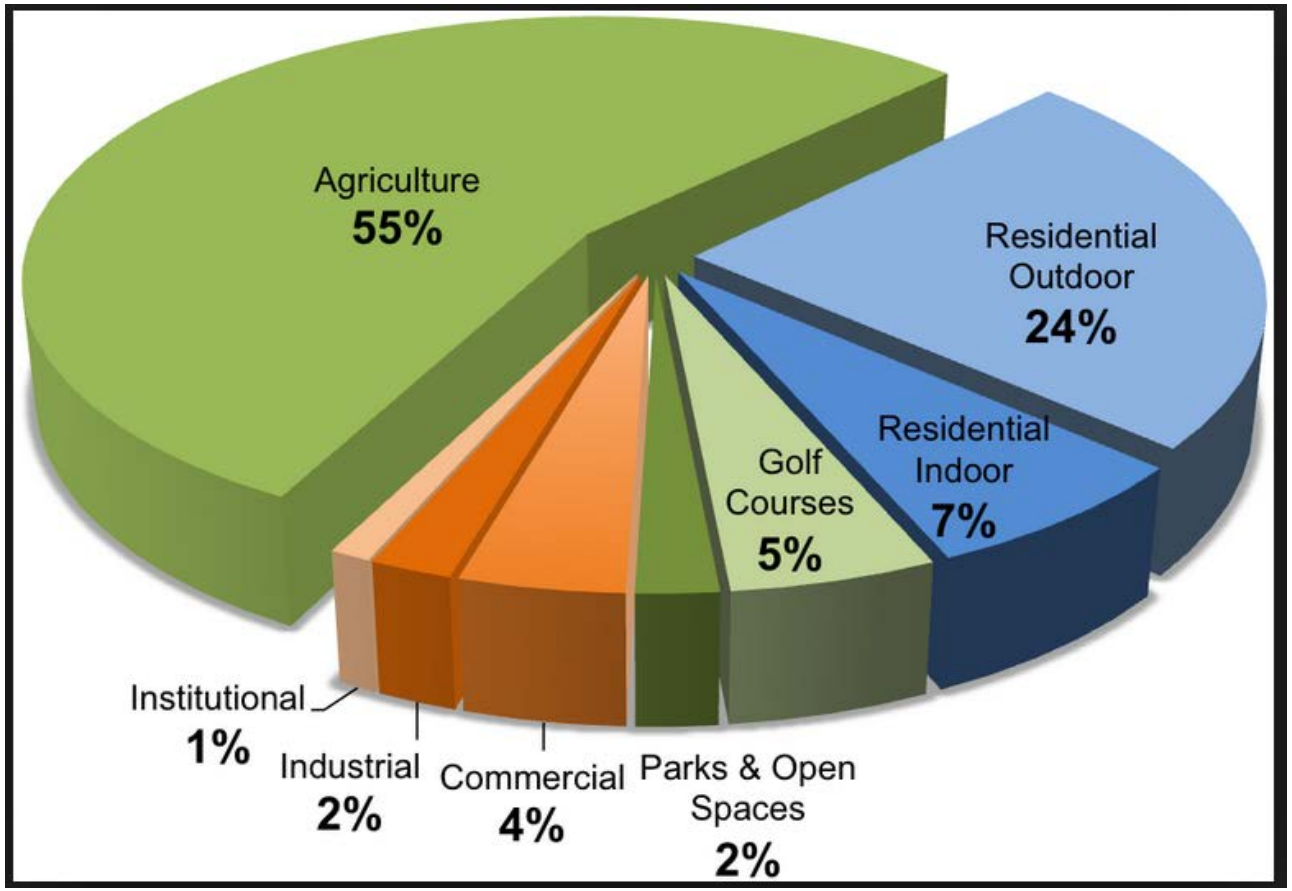
Have the Y axis right!

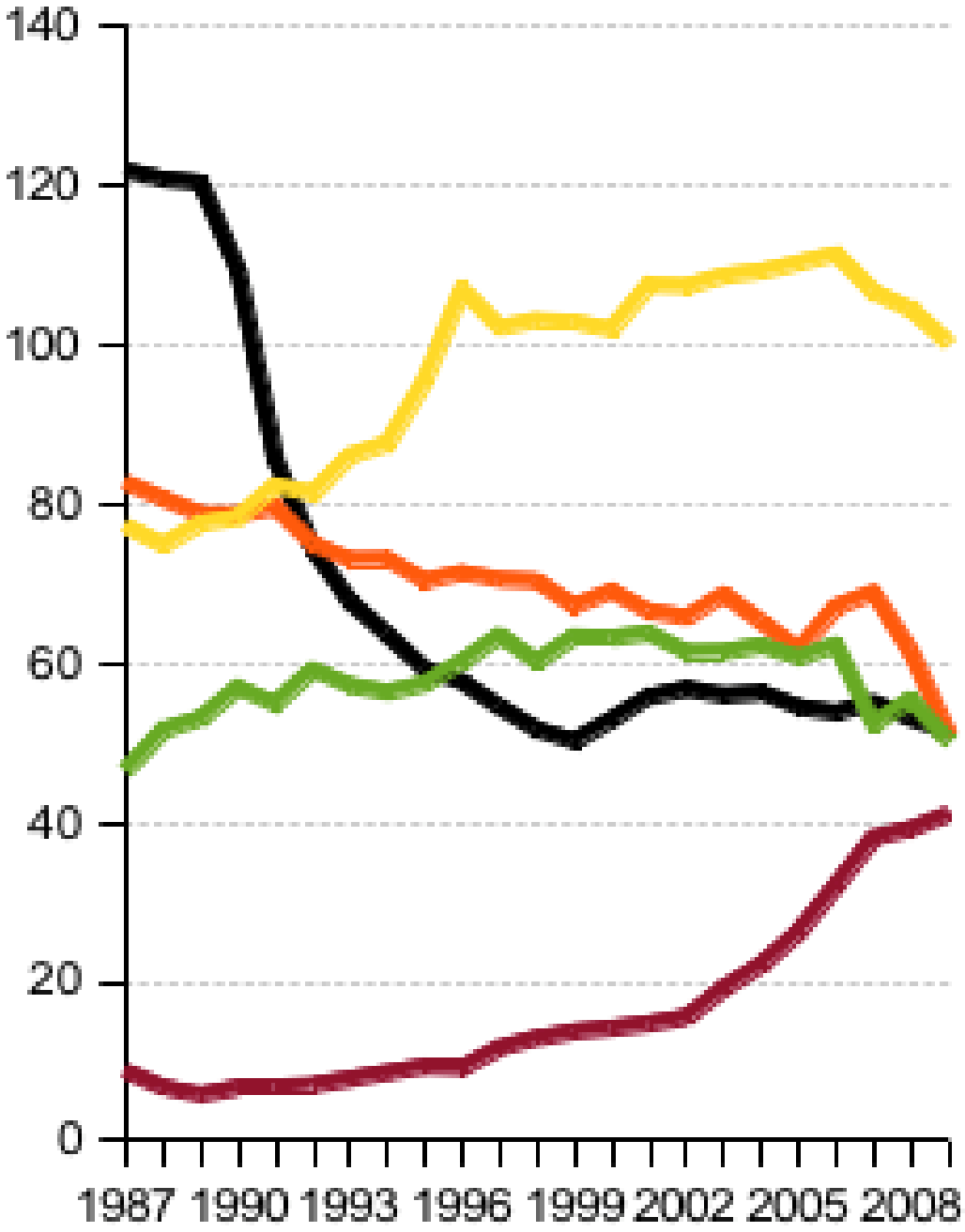




Short Assignment!

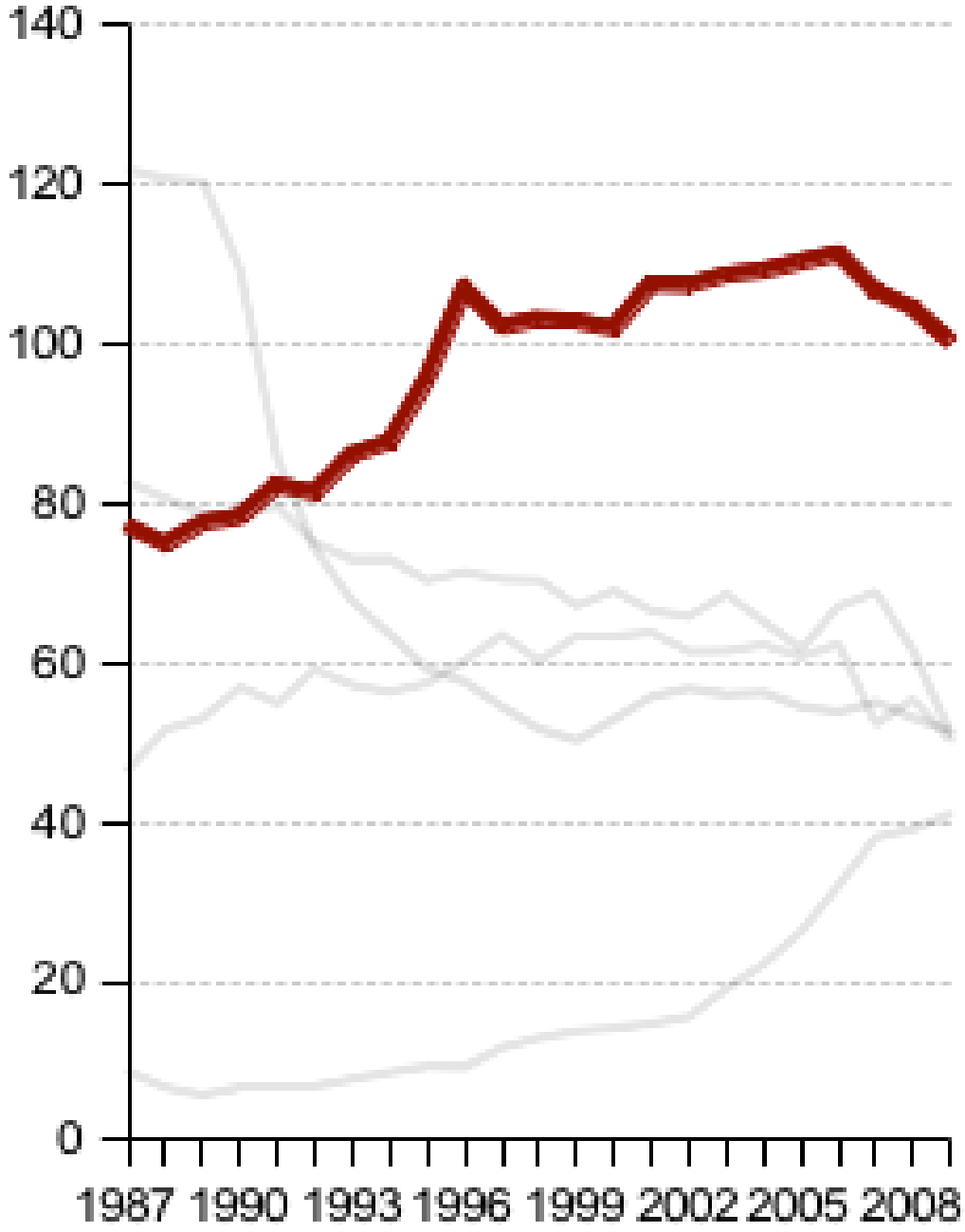
What is wrong with this graphic?





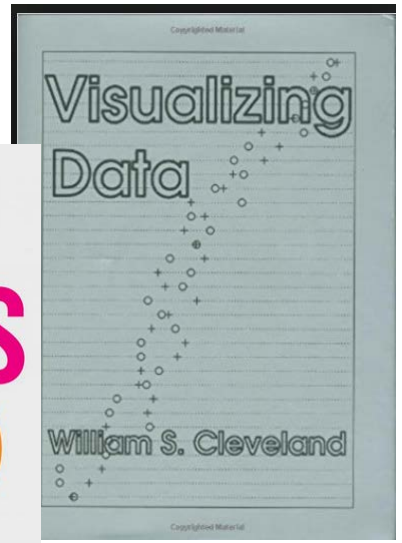
Assignment:

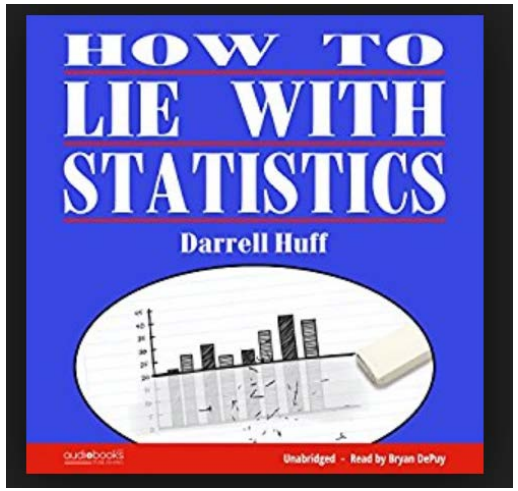
Emphasize
the performance of
“Yellow county”?

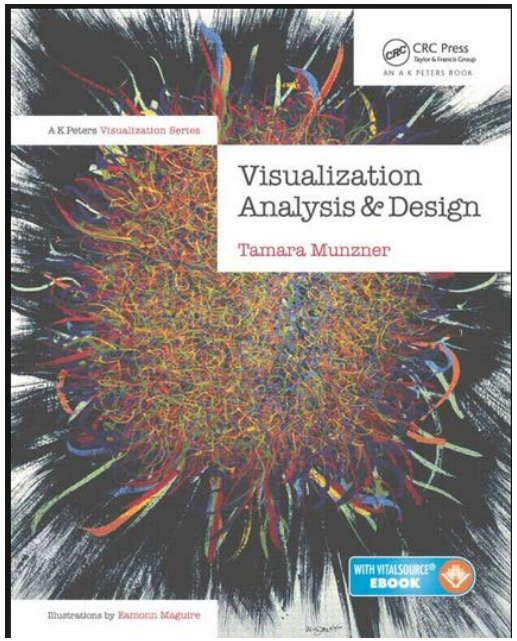

140
120
100
80
60
40
20
0
1987
1990
1993
1996
1999
2002
2005
2008

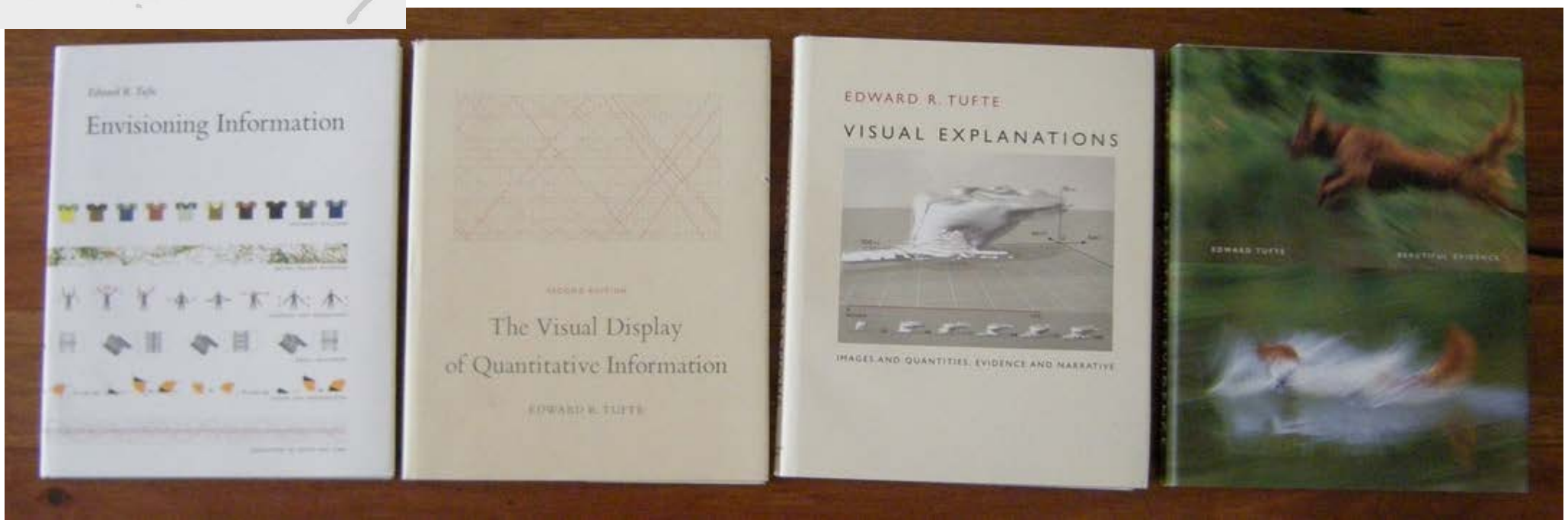
References


GOOD
CHARTS
SCOTT
BERINATO
The HBR Guide
to Making
Smarter, More
Persuasive Data
Visualizations
HARVARD
BUSINESS
REVIEW
PRESS


Visualizing
Data
William S. Cleveland


HOW TO
LIE WITH
STATISTICS
Darrell Huff
audiobooks
Unabridged - Read by Bryan DePuy


CRC Press
AN A K PETERS BOOK
A K Peters Visualization Series
Visualization
Analysis & Design
Tamara Munzner
WITH VITALSOURCE®
EBOOK
Illustrations by Eamonn Maguire


Envisioning Information
The Visual Display
of Quantitative Information
EDWARD R. TUFTE
VISUAL EXPLANATIONS
IMAGES AND QUANTITIES, EVIDENCE AND NARRATIVE

Thankyou!