

# DSBA 5122: Visual Analytics

## Class 2: Visual Representations Basics I

Ryan Wesslen

January 28, 2019

# The Truth Continuum (Chapter 3)

"Any visualization is a model."

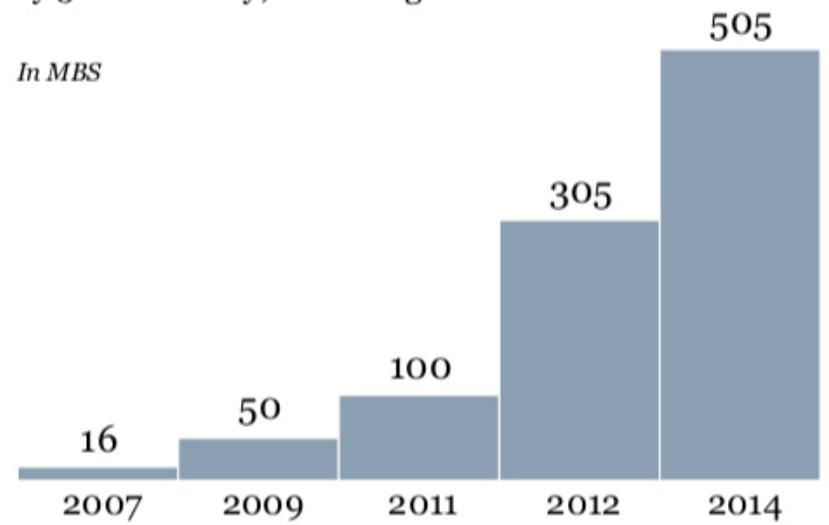
# The Truth Continuum (Chapter 3)

"Any visualization is a model."

"The more adequately a model fits ... without being needlessly complex, and the easier ... to interpret it correctly, the better it will be."

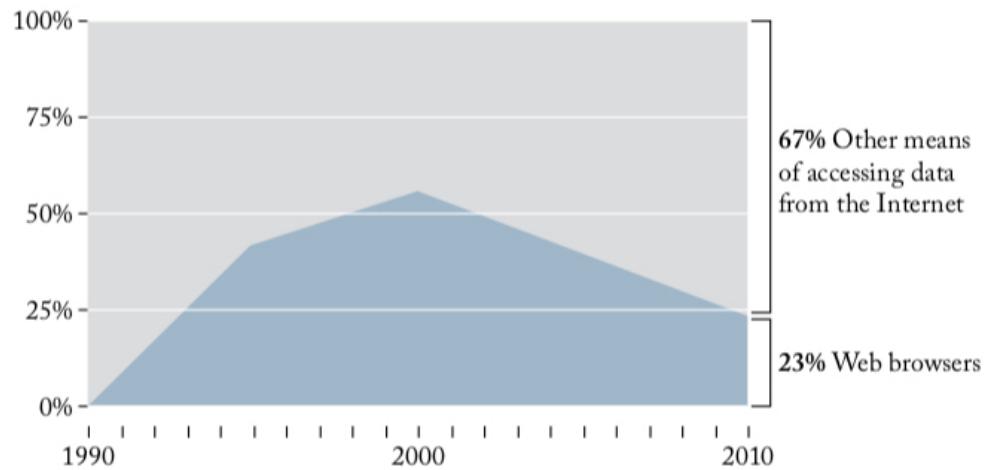
## Cable's Top Internet Speeds

Cable providers have improved broadband speeds by 50% annually, on average

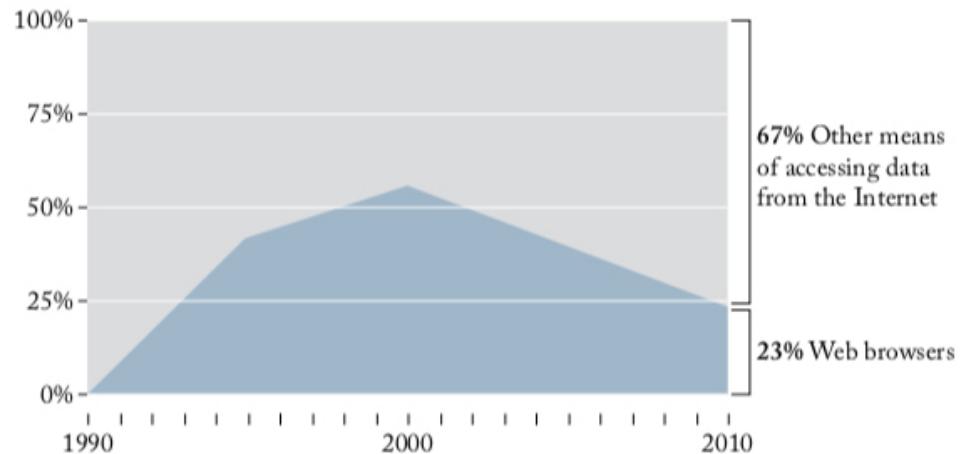


**Figure 3.1** Misleading your audience may yield benefits in the short term. In the long term, however, it may destroy your credibility.

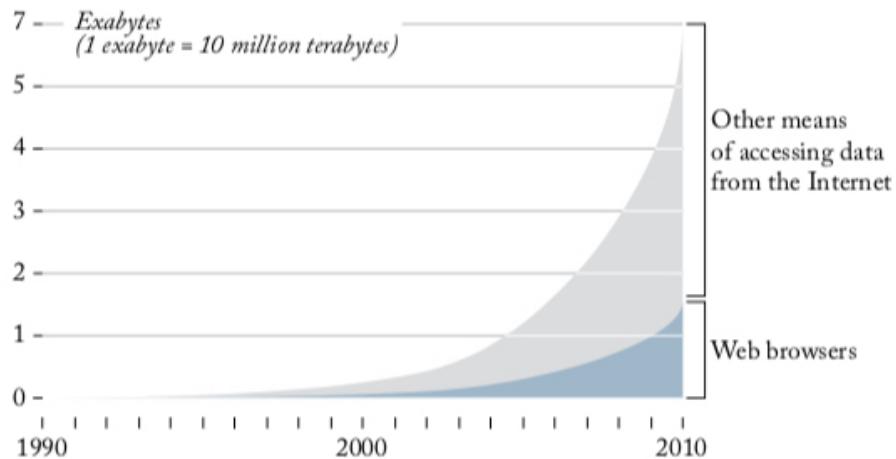
### Internet Traffic in the U.S.

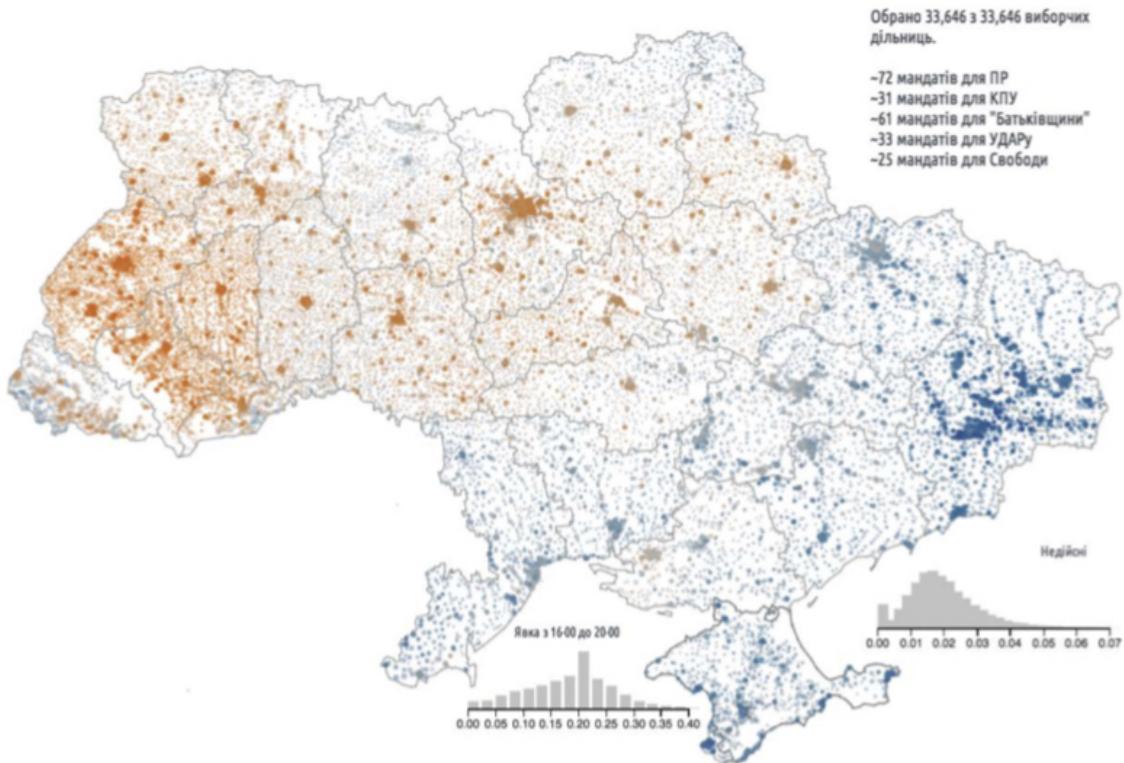


### Internet Traffic in the U.S.



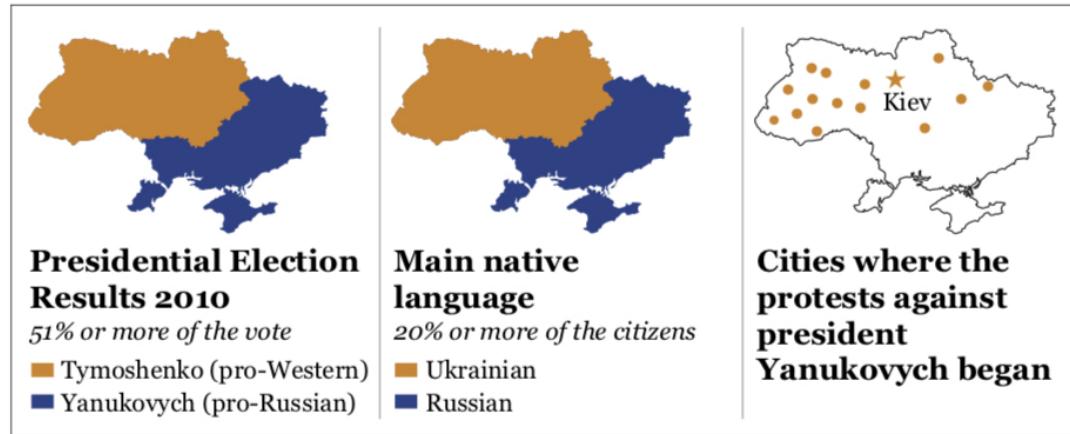
### Internet Traffic in the U.S.





**Figure 3.5** Results of the Ukrainian parliamentary elections, 2012. (Visualization by Texty: <http://texty.org.ua/mod/datavis/apps/elections2012/>.)

# The Mind, That Clumsy Modeler



**Figure 3.6** Ukraine seems to be a completely divided country.

# The Mind, That Clumsy Modeler

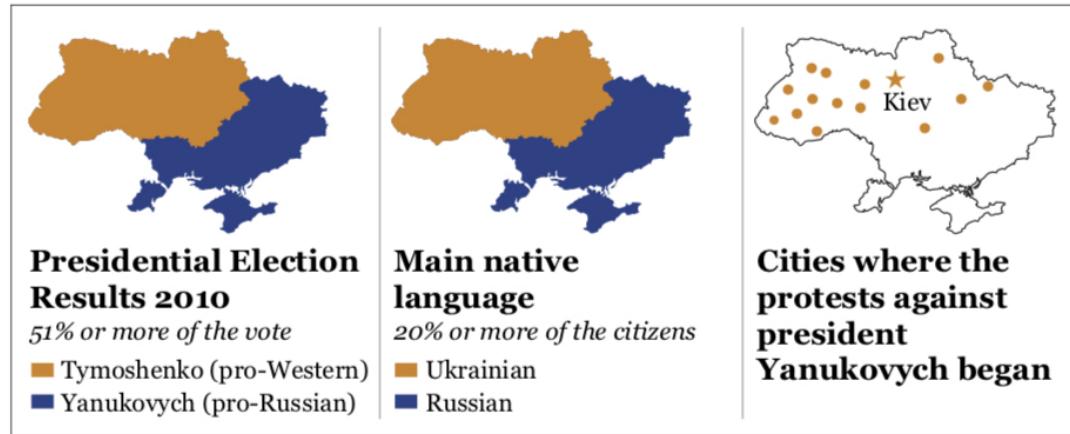


Figure 3.6 Ukraine seems to be a completely divided country.

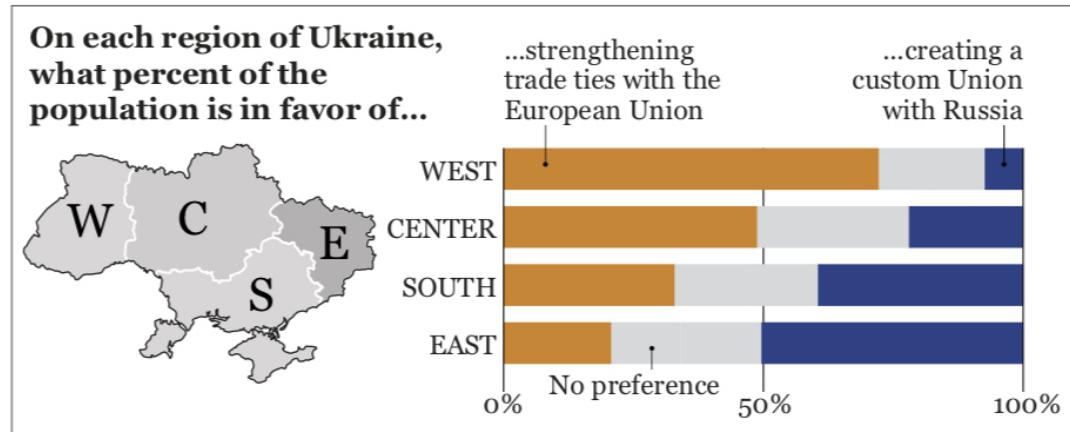


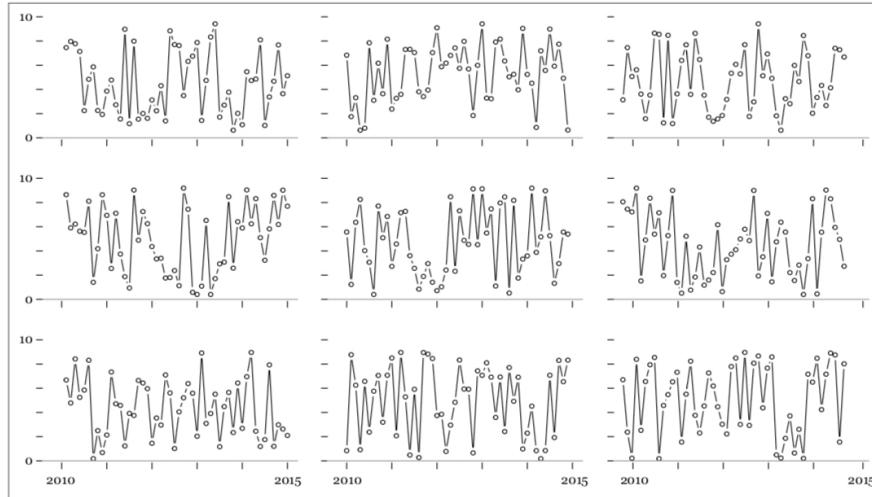
Figure 3.7 What Ukrainians want.

# Why Are We So Often Mistaken?

Cairo's three "bugs" in thinking (Ch. 3)

# Patternicity Bug

Humans have an innate ability to detect **patterns**, especially visually. This is why visualizations can be so important!



However, many patterns that your eyes and brain detect in data are the result of pure **coincidences** and **noise**, or "mental overfitting".

# Storytelling Bug

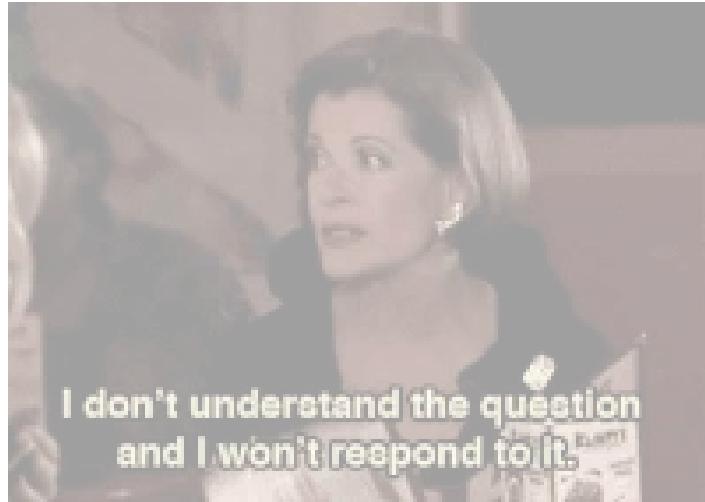
Once we detect patterns, we naturally try to find a **cause-effect explanation** for them.

Storytelling can be a potent tool for communicating effectively, but it is **dangerous** if it blinds us toward evidence that should compel us to tweak or discard our models.



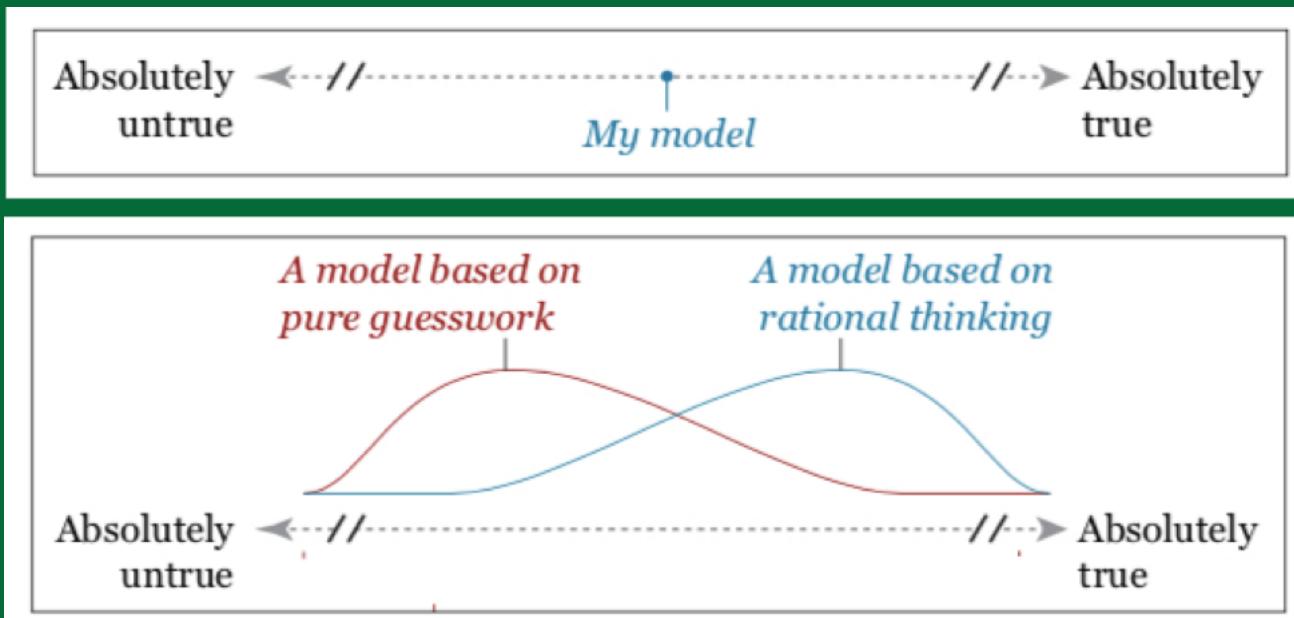
# Confirmation Bug

Once a good story takes over our understanding of something, we'll attach to it like leeches to warm, plump flesh. **An attack on our beliefs will be seen as personal.**



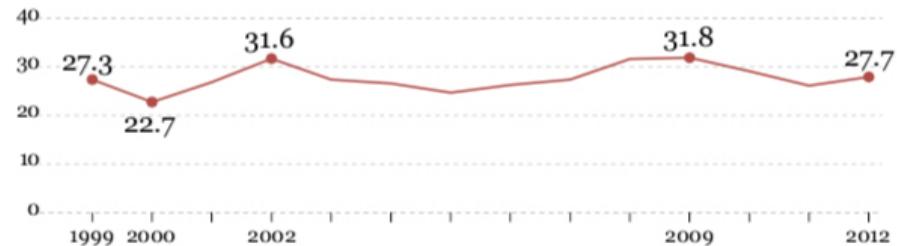
"The way we present information matters as much as the soundness of the information itself."

# Truth Is Neither Absolute, Nor Relative (Chapter 3)

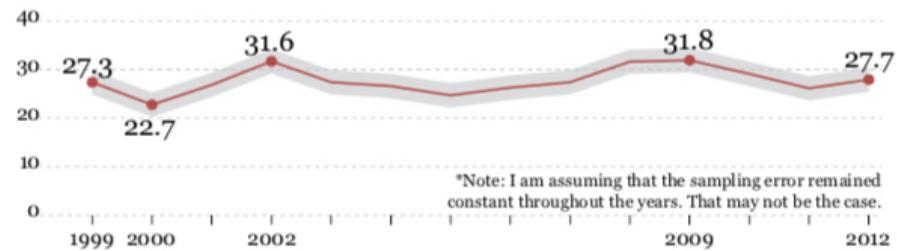


Bachelor's degree in Journalism and Mass Communication recipients who wish they had chosen another career

*Percentage*



*With sampling error (+/- 2.3)\**



**Figure 3.9** How many journalism and mass communication grads say they aren't happy with their choice of major? I am assuming that the error was the same in all years, which may not be the case.

Absolutely  
untrue

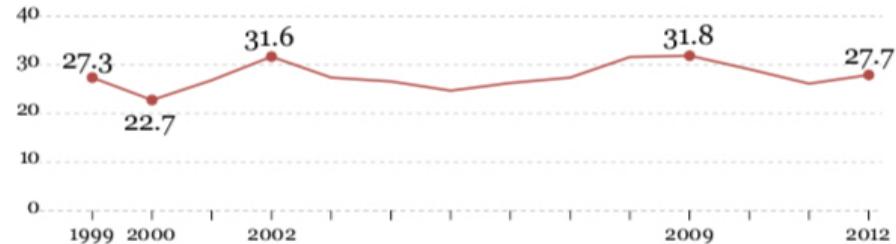
**Model 1**

28% wished that they'd  
chosen a different field

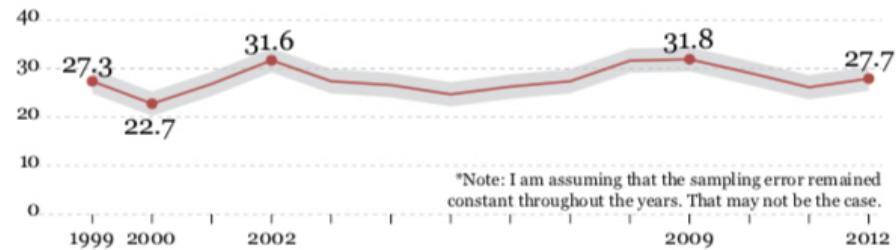
Absolutely  
true

Bachelor's degree in Journalism and Mass Communication recipients who wish they had chosen another career

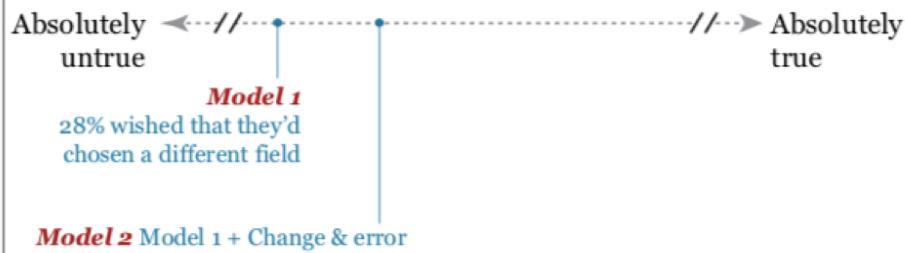
*Percentage*



*With sampling error (+/- 2.3)\**



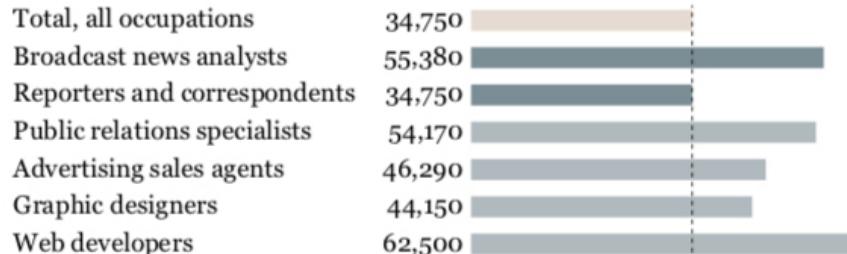
**Figure 3.9** How many journalism and mass communication grads say they aren't happy with their choice of major? I am assuming that the error was the same in all years, which may not be the case.



### **Percent change in employment, projected 2012-2022**

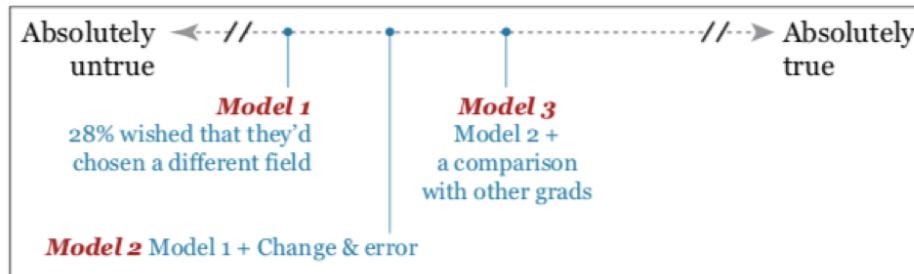


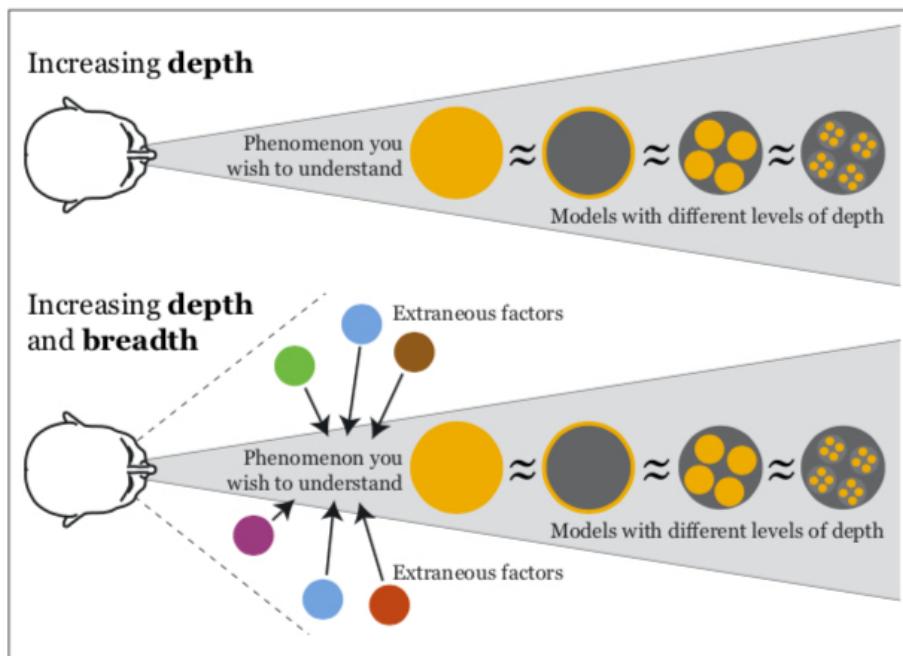
### **Median annual wages, May 2012. In US \$**



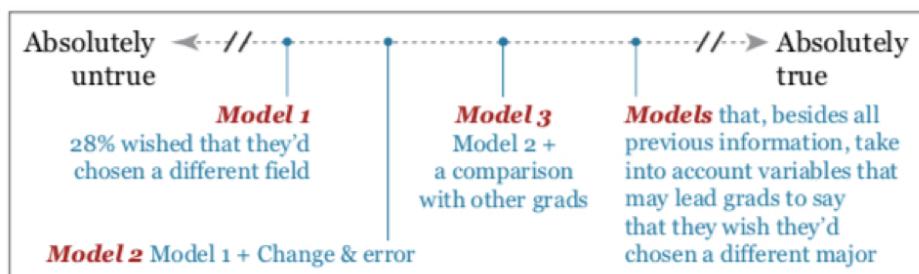
SOURCE: U.S. Bureau of Labor Statistics, Occupational Employment Statistics

**Figure 3.10** Comparison of several occupations in media.





**Figure 3.11** Levels of depth and breadth.



# Basic Principles of Visualization (Chapter 5)

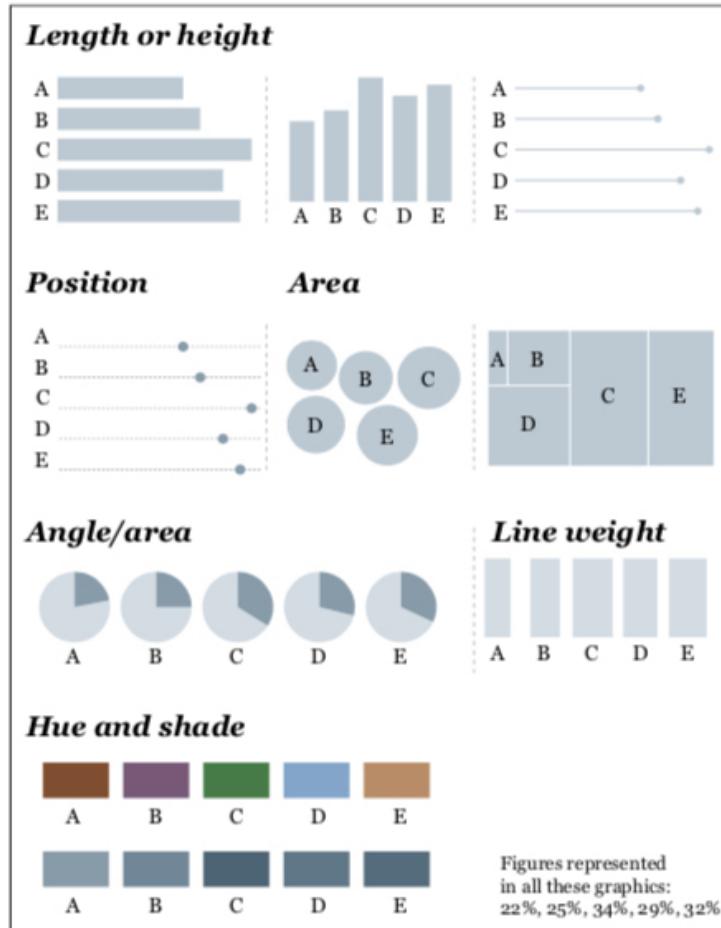
## Cairo's design suggestions for visualizations:

1. Think about the task(s) you want to enable
2. Try different graphic forms
3. Arrange the components of the graphic
4. Test the outcomes

# 1. Think about the task(s) you want to enable

- What is your problem?
- What action (task) can someone do to solve the problem?
- What data do you have?
- What are the most important data elements?

## 2. Try different graphic forms



**Figure 5.2** Different methods of encoding the same small data set. Remember that, perhaps because our client requested it, countries are organized alphabetically. Otherwise, it'd make more sense to arrange the figures from largest to smallest.

# 3. Arrange the graphic components

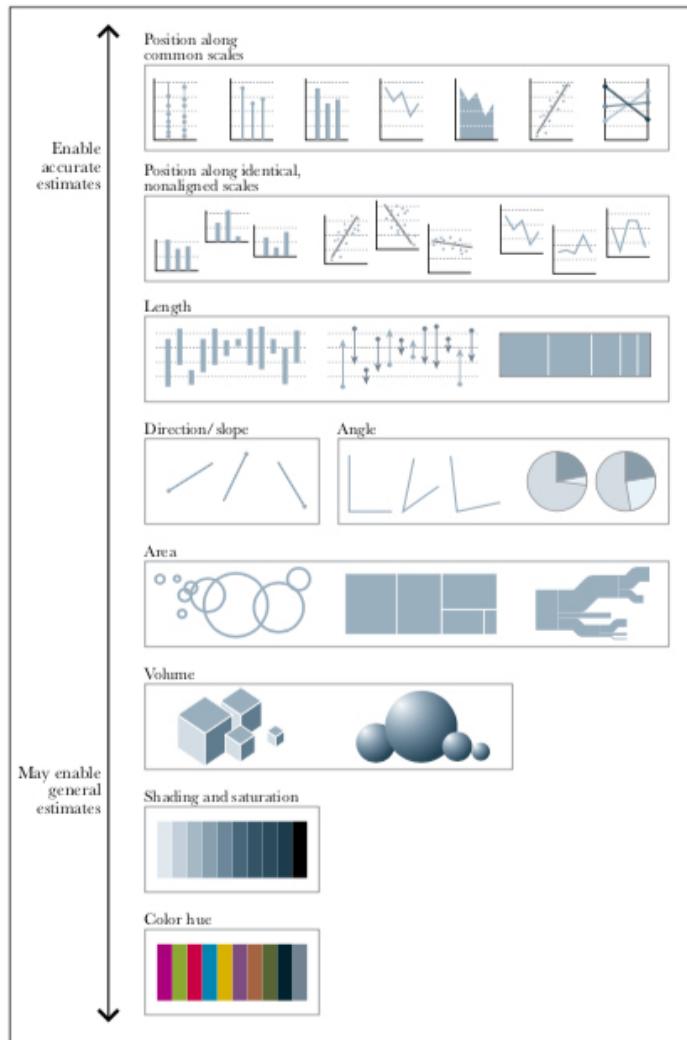


Figure 5.5 Scale of elementary perceptual tasks, inspired by William Cleveland and Robert McGill.

## 4. Test the outcomes

### *Distribution of one-year returns*

#### Fund A: High risk, high return



#### Fund B: Low risk, low return



### *Distribution of average annual returns over 30 years*

#### Fund A: High risk, high return



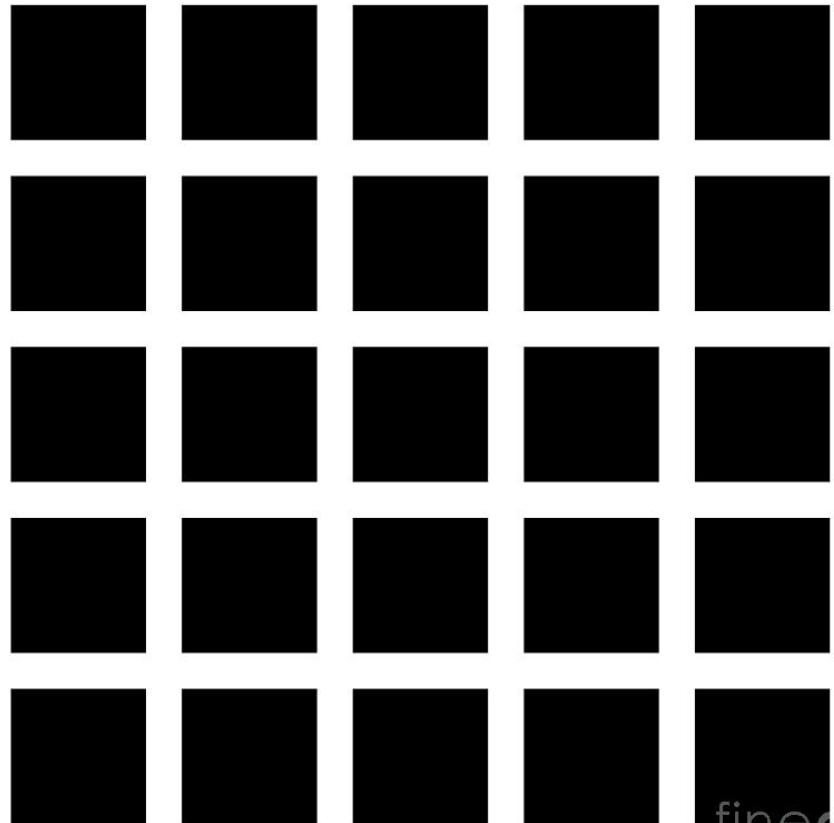
#### Fund B: Low risk, low return



**Figure 5.22** Charts based on Richard H. Thaler's *Misbehaving: The Making of Behavioral Economics* (2015).

Look at Data (Chapter 1, Healy)

# Hermann Grid Effect

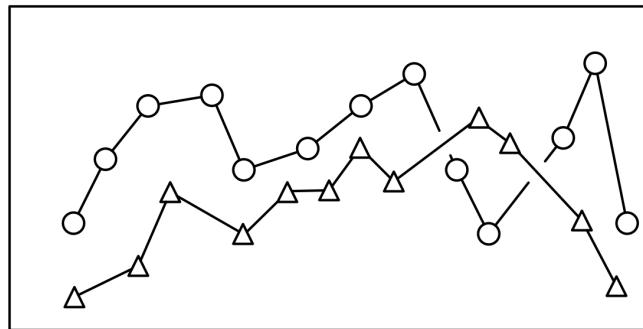
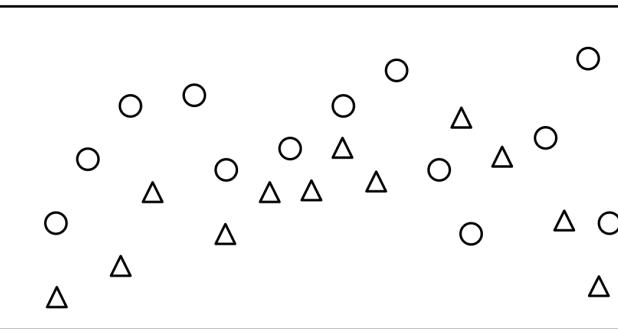
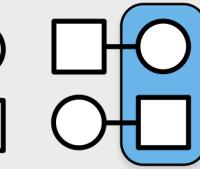
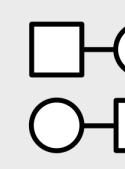
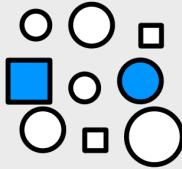
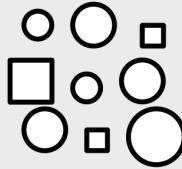
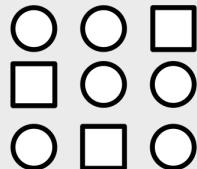
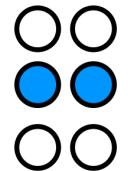
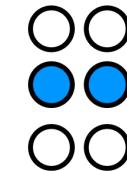
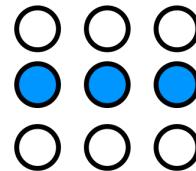
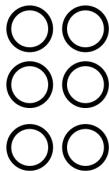
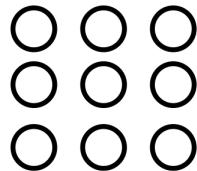


finde  
einfach

## Mach bands

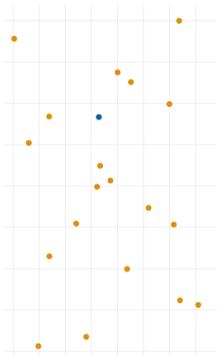


# Gestalt

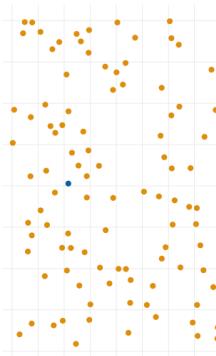


# Pre-attentive

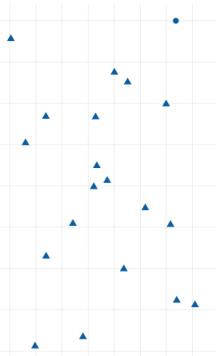
Color Only, N=20



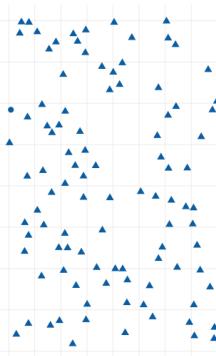
Color Only, N=100



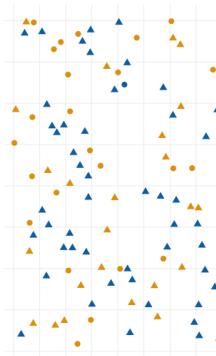
Shape Only, N=20



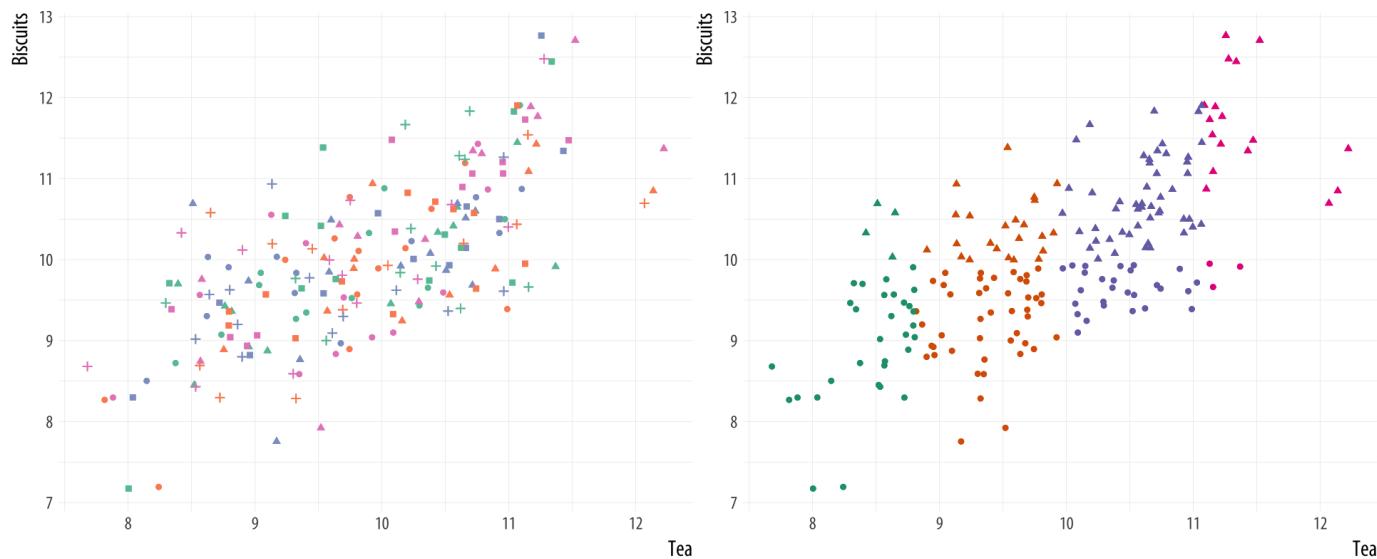
Shape Only, N=100

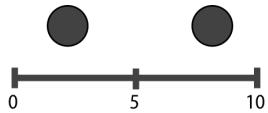


Color & Shape, N=100

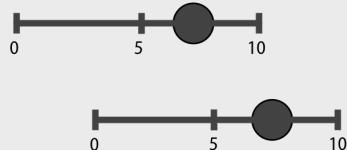


# Multichannel





Position on  
a common scale



Position on  
unaligned  
scales



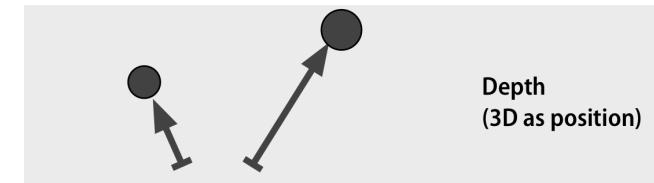
Length



Tilt or Angle



Area (2D as size)



Depth  
(3D as position)



Color luminance  
or brightness



Color saturation  
or intensity



Curvature



Volume  
(3D as size)

# Cleveland and McGill, 1984

## Position and length

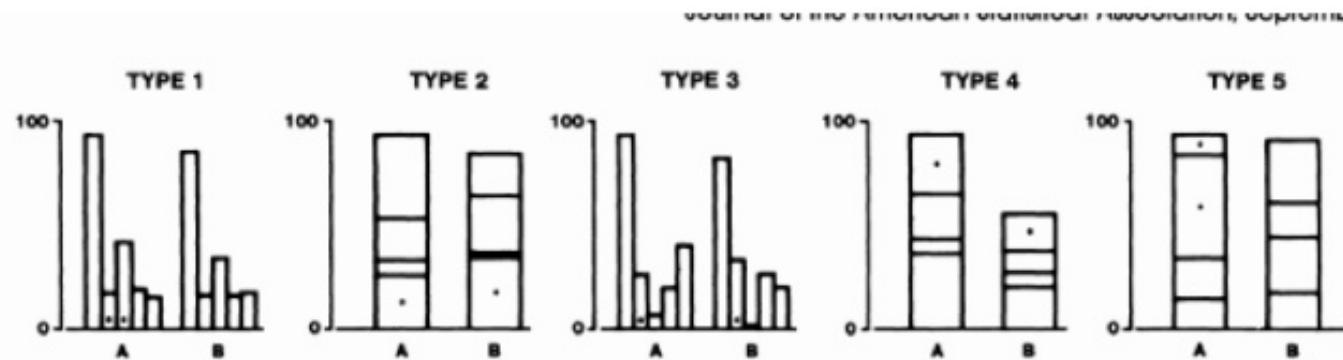
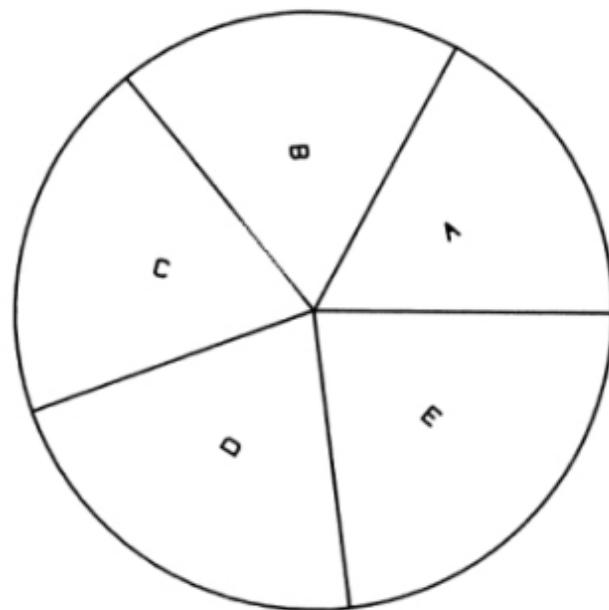
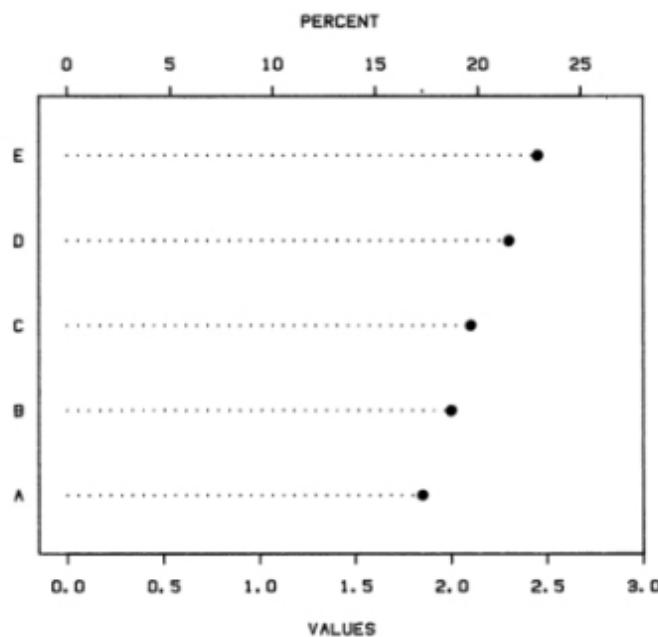


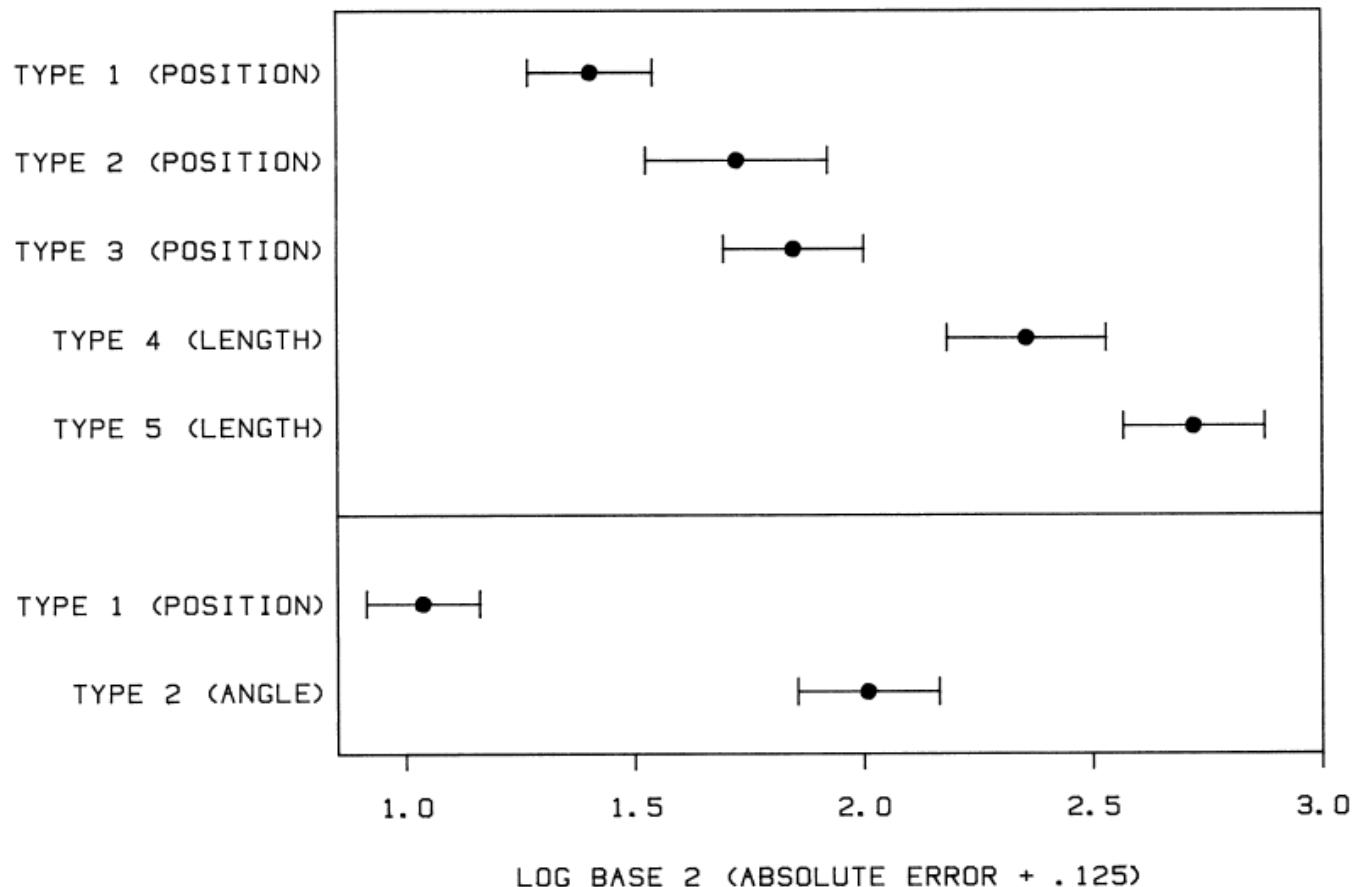
Figure 4. Graphs from position-length experiment.

Cleveland W and McGill R. Graphical Perception: Theory, experimentation and applications to the development of graphical methods. JASA 79 (387): 531-554; 1984.

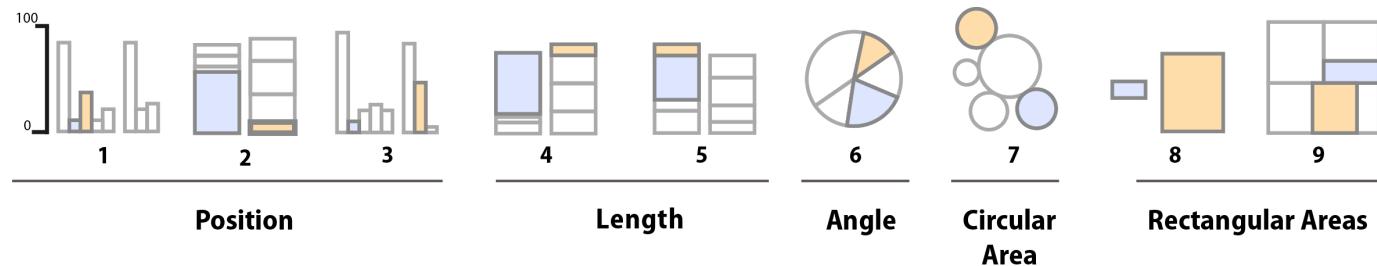
# Cleveland and McGill, 1984



# Cleveland and McGill, 1984



# Heer and Bostock, 2010





# Question: How do we implement?



# A (Layered) Grammar of Graphics!

`ggplot2` and its `tidyverse` friends