

# Creating effective visualisations

**Hadley Wickham**

Assistant Professor / Dobelman Family Junior Chair  
Department of Statistics / Rice University

June 2012



**HELLO**

my name is

**Hadley**

**Critique**

**Graphics are like  
pumpkin pie**

The four **C**'s of critiquing a graphic

**Content**



**Construction**



**Context**



**Consumption**



# Content

# Content

What data (variables) does the graph display?

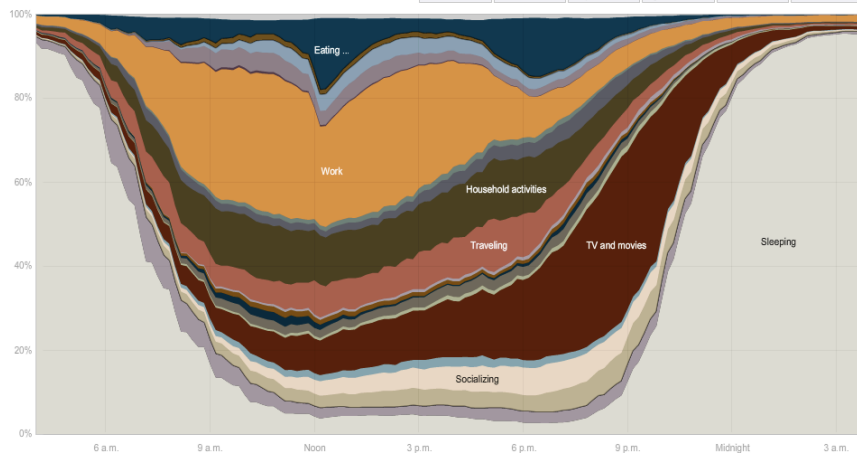
What non-data is present?

What is **pumpkin** (essence of the graphic) vs what is **spice** (useful additional info)?

## Everyone

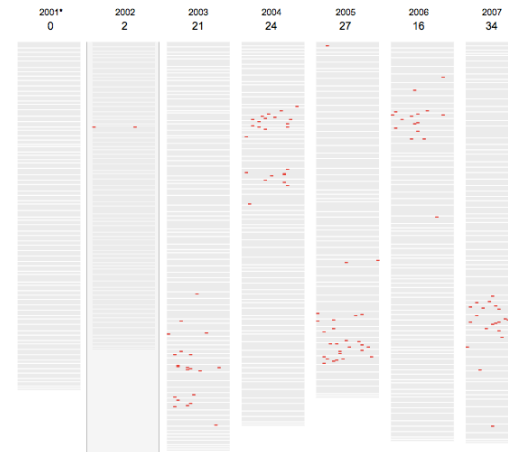
Sleeping, eating, working and watching television take up about two-thirds of the average day.

<b>Everyone</b>	Employed	White	Age 15-24	H.S. grads	No children
Men	Unemployed	Black	Age 25-54	Bachelor's	One child
Women	Not in lab...	Hispanic	Age 65+	Advanced	Two+ children



<http://nyti.ms/np29Yk>

## Use of the phrase "Iraq" in past State of the Union Addresses

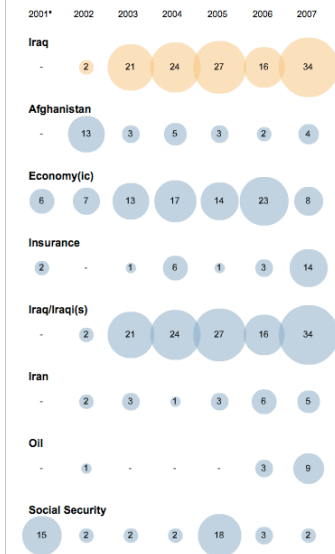


### The word in context

IRAQ continues to flaunt its hostility toward America and to support terror. The Iraqi regime has plotted to develop anthrax, and nerve gas, and nuclear weapons for over a decade. This is a regime that has already used poison gas to murder thousands of its own citizens -- leaving the bodies of mothers huddled over their dead children. This is a regime that agreed to international inspections -- then kicked out the inspectors. This is a regime that has something to hide from the civilized world.

~ 2002 (Paragraph 20 of 67)

## Compared with other words



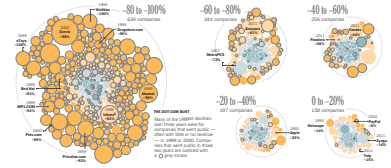
<http://nyti.ms/r8KdvU>

### What Happens After the I.P.O.?

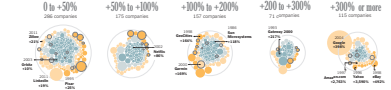
There have been about 2,400 technology, Internet and software IPO's since 1980. On the first day of trading, the average stock rose 32 percent above its offer price. But in the three years after that, most companies had negative returns, according to statistics compiled by Jay Ritter, a professor of finance at the University of Florida. Companies with higher valuations compared with their revenue before the IPO had lower returns, generally poorly.



#### Returns three years after the IPO: The decliners ...

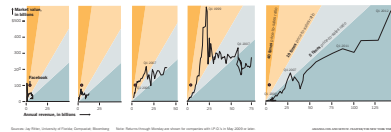


#### ... and the gainers.



#### Over the Long Haul

Performance over the long haul (between 1980 and 1999) indicates a correlation between IPO returns and market returns. Values are correlated with the S&P 500. A list of five technology IPOs and their market returns is provided.



# Your turn

In small groups, identify the data and non-data in each of the three plots. Which features are the most important? Which are just useful background information?

# Construction

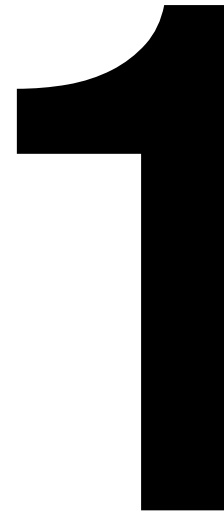
# Construction

How many layers are on the plot?  
What data does each layer display? What sort of geometric object does it use? Is it a summary of the raw data? How are variables mapped to aesthetics?



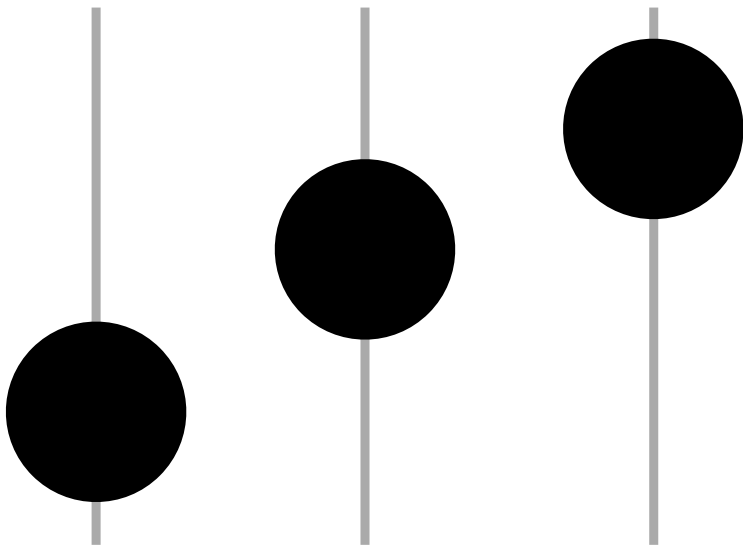


<http://www.dreamsystemsmedia.com/blog/index.php/social-media-statistics-of-the-day/>

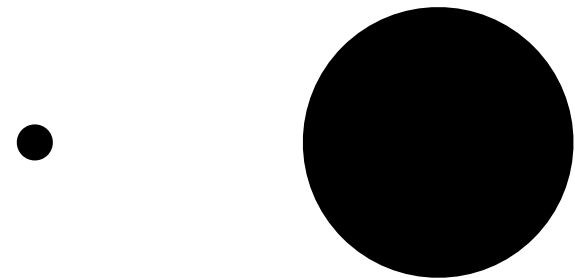


Match perceptual and data topology

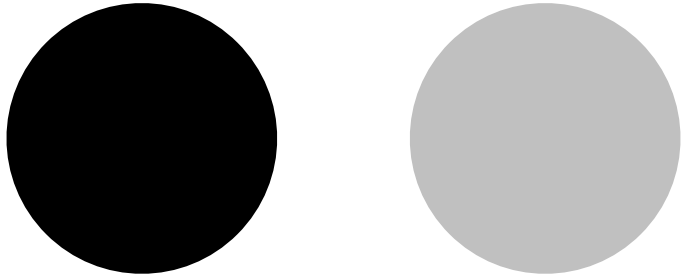
Which represents the larger value?



Which represents the larger value?



Which represents the larger value?



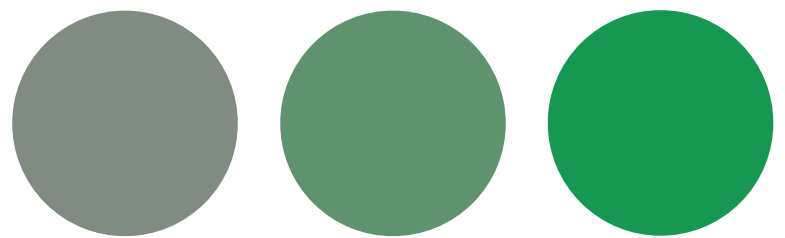
Which represents the larger value?



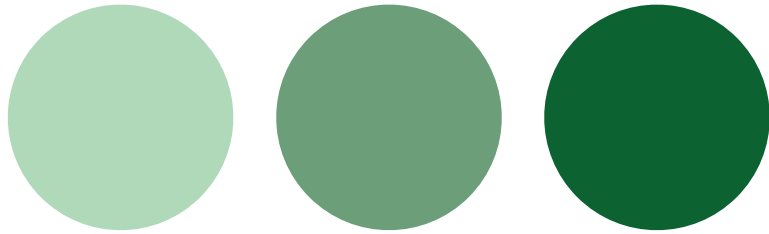
Which represents the larger value?



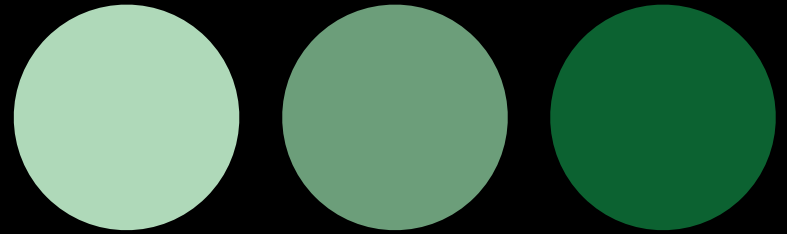
Which represents the larger value?



Which represents the larger value?



Which represents the larger value?

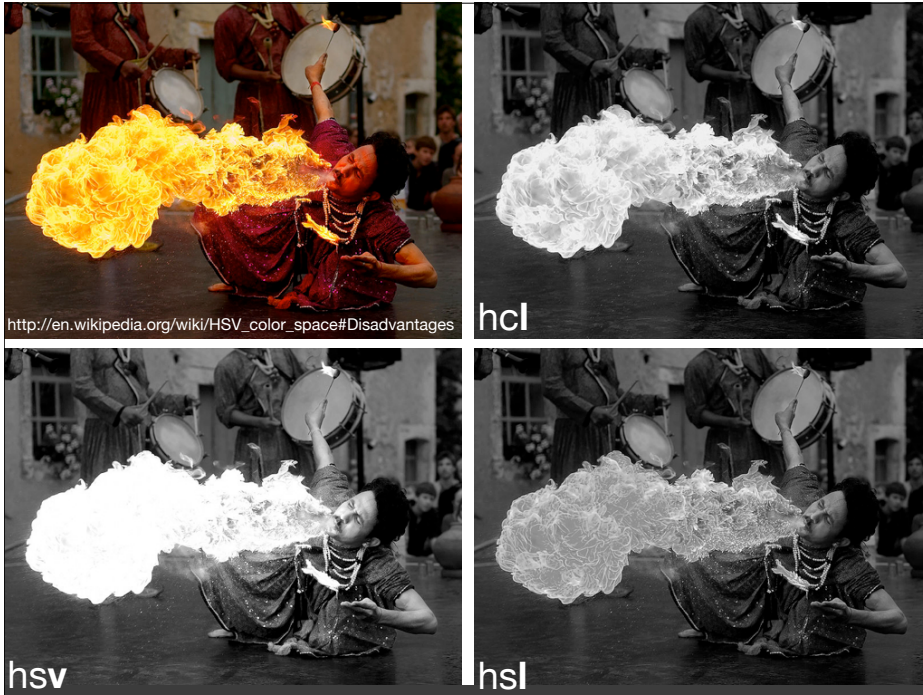


What are  
the three  
important  
components  
of colour?

**RGB**

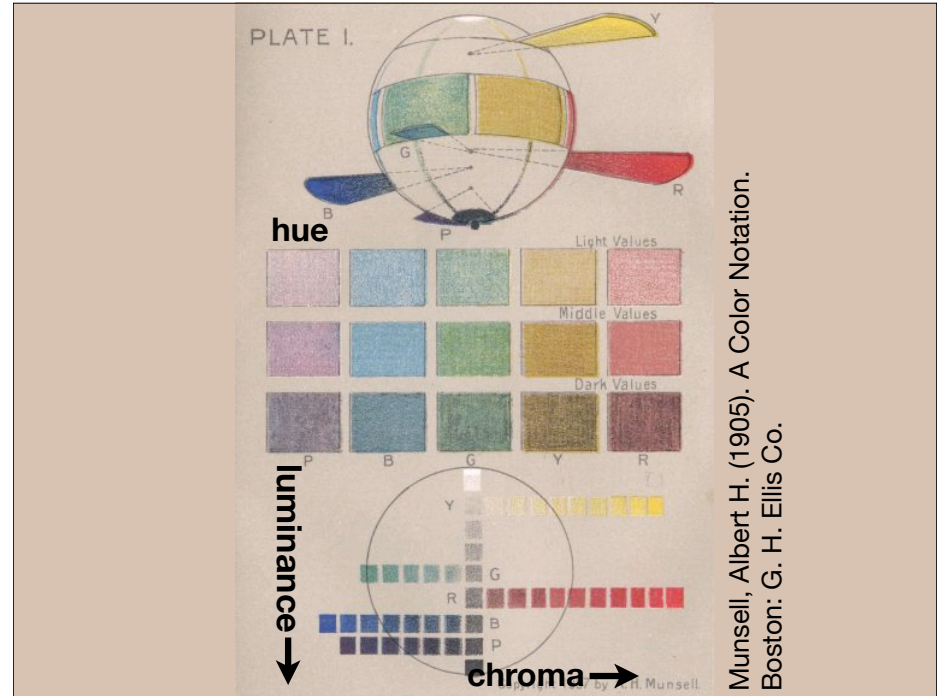
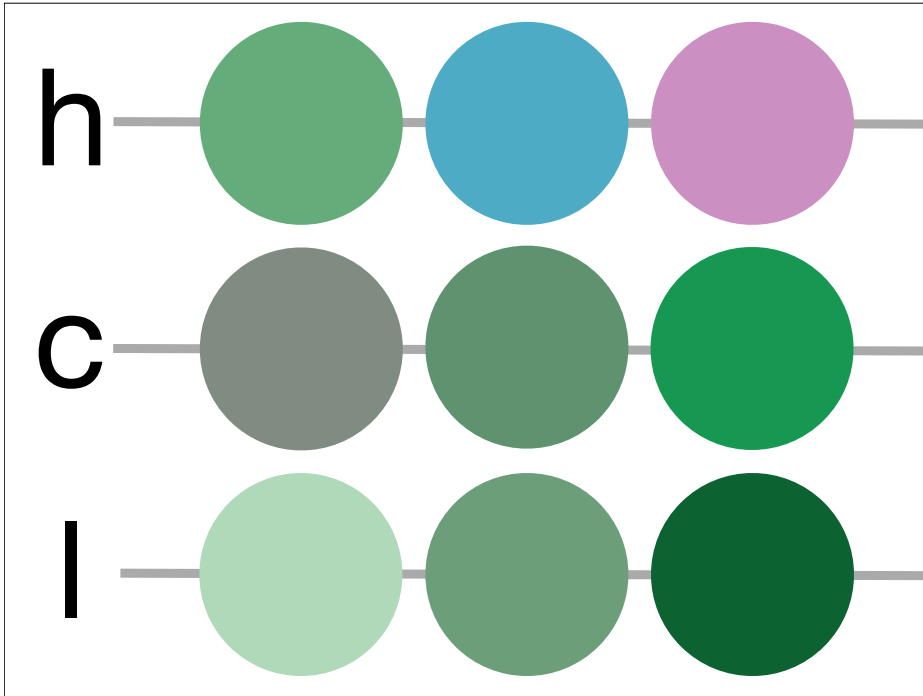


~~RGB~~  
 HSV HSL  
 HCL  
 (aka polar LUV)



~~RGB~~  
~~HSV~~ ~~HSL~~  
 HCL  
 (aka polar LUV)





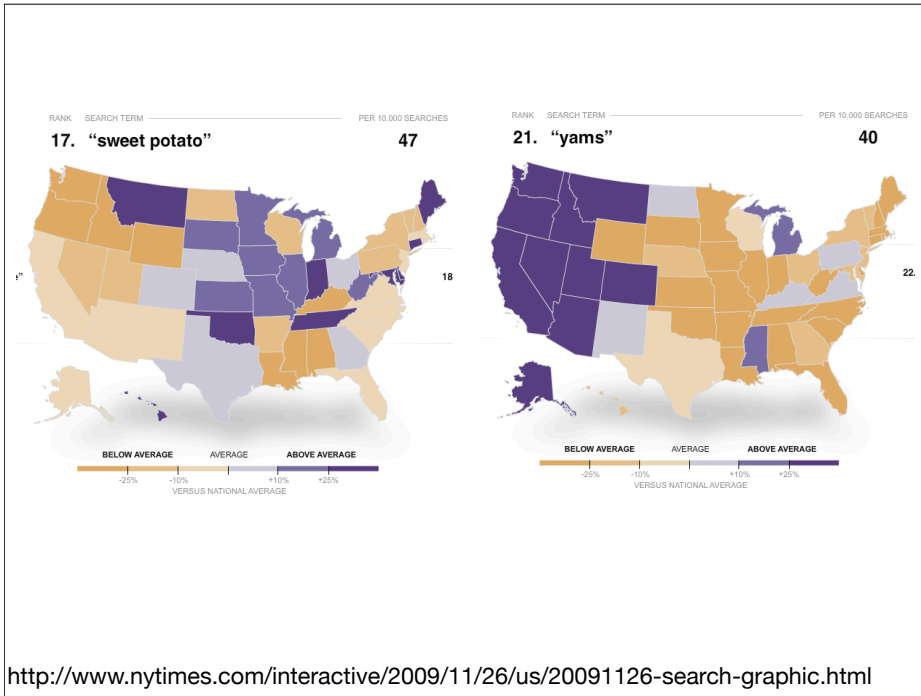
## Why care?

### Perceptually uniform

**Hue** is unordered. Use evenly spaced hues with equal chroma and luminance to make aesthetically pleasing discrete palettes.

**Chroma** and **luminance** are ordered. Easy to make perceptually uniform gradients by varying either (or both). Never use rainbow scales again!

Aesthetic	Topology
Position	Ordered
Size	Ordered
Luminance	Ordered
Chroma	Ordered
Shape	Unordered
Hue	Unordered

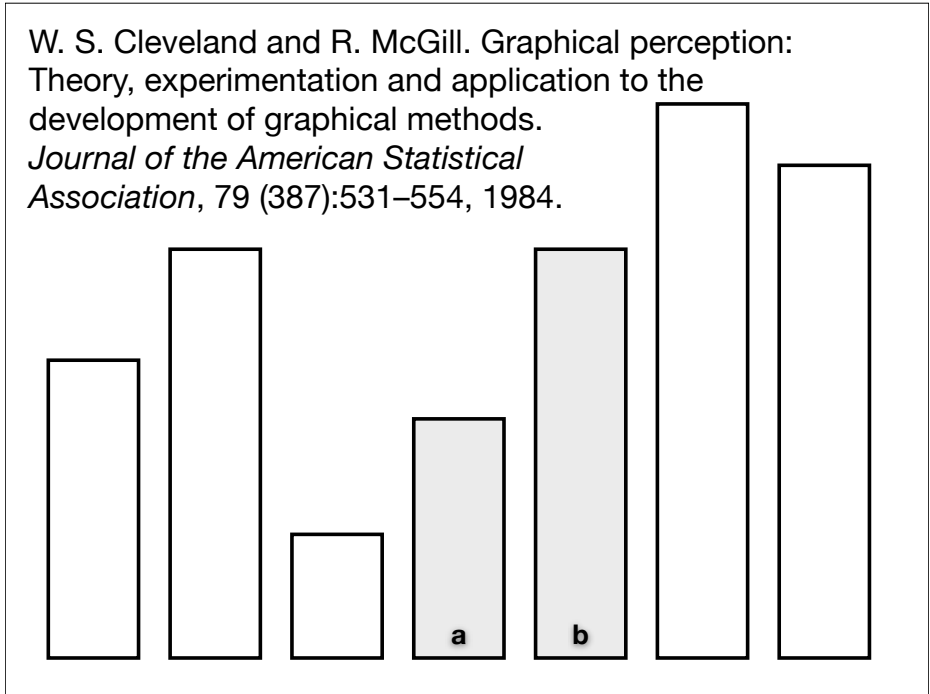


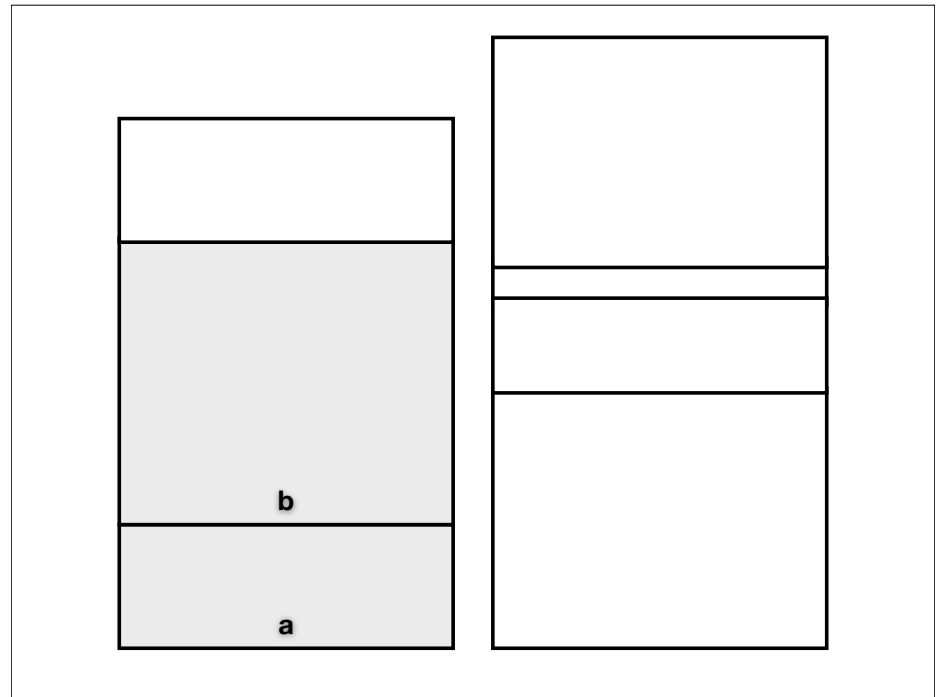
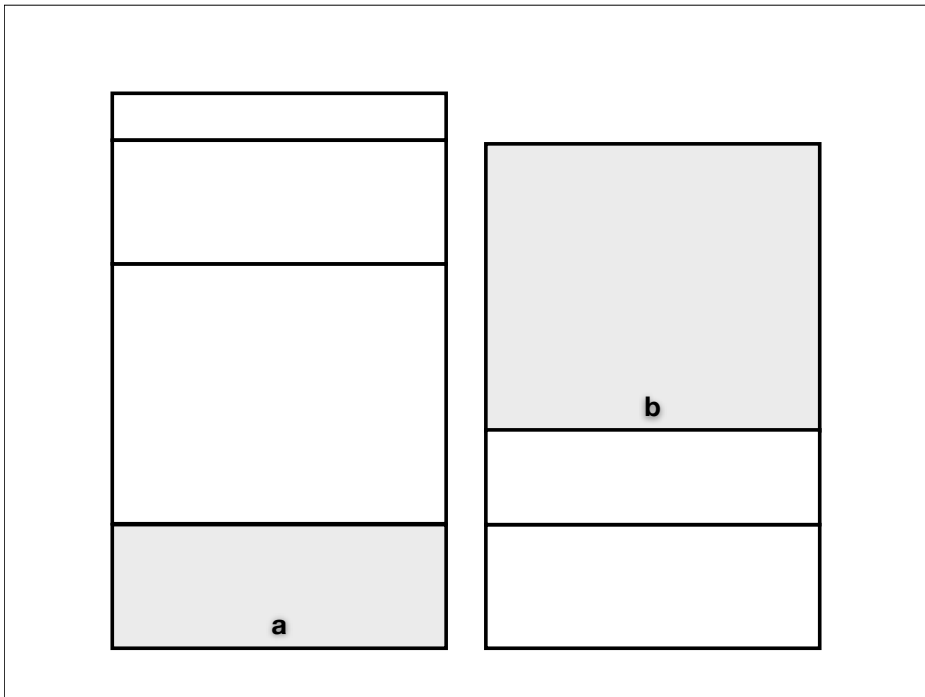
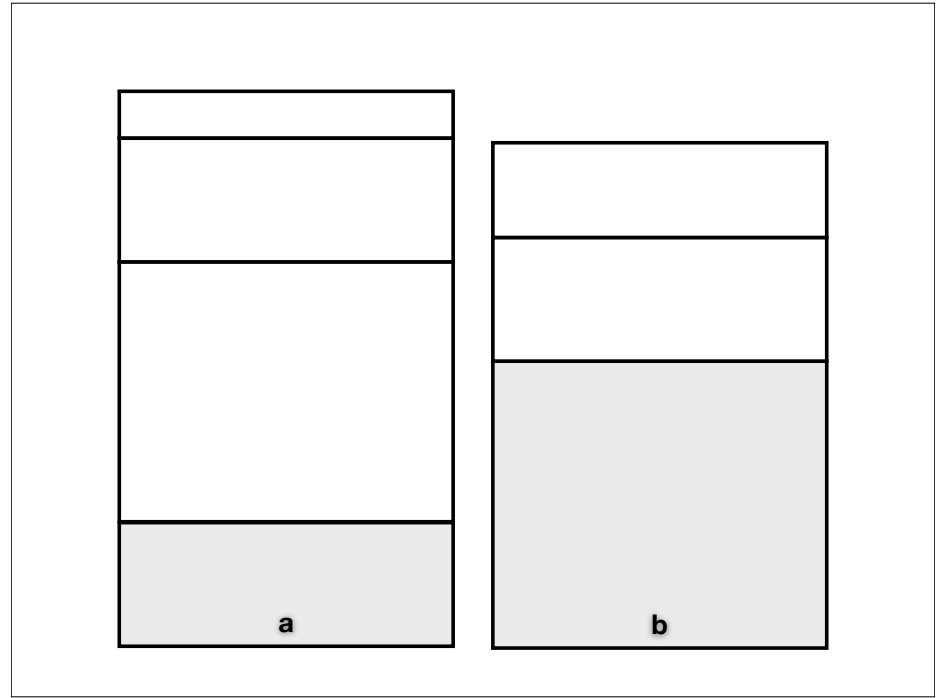
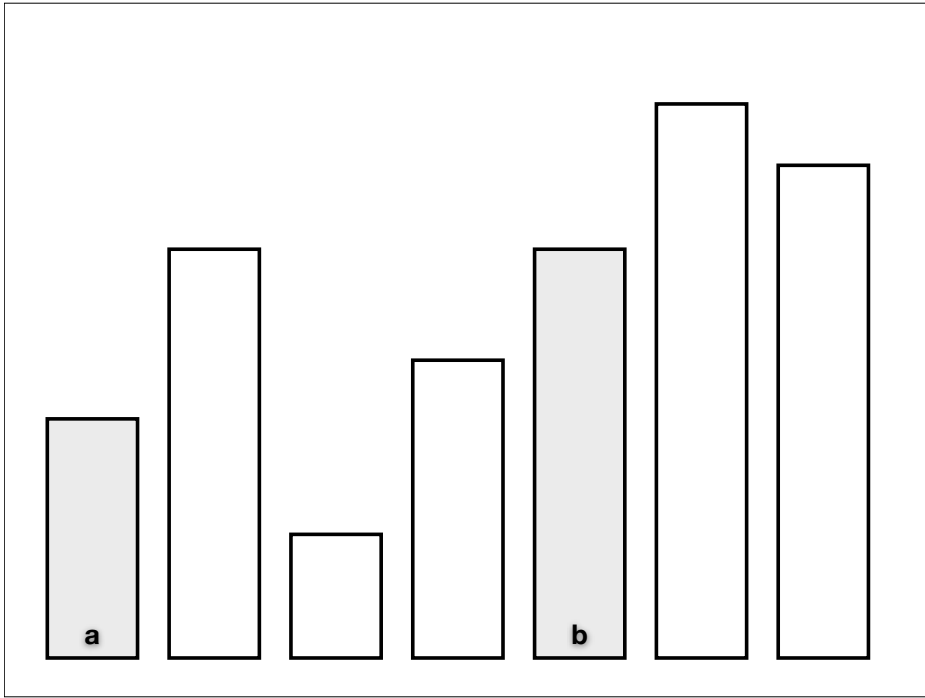
# Your turn

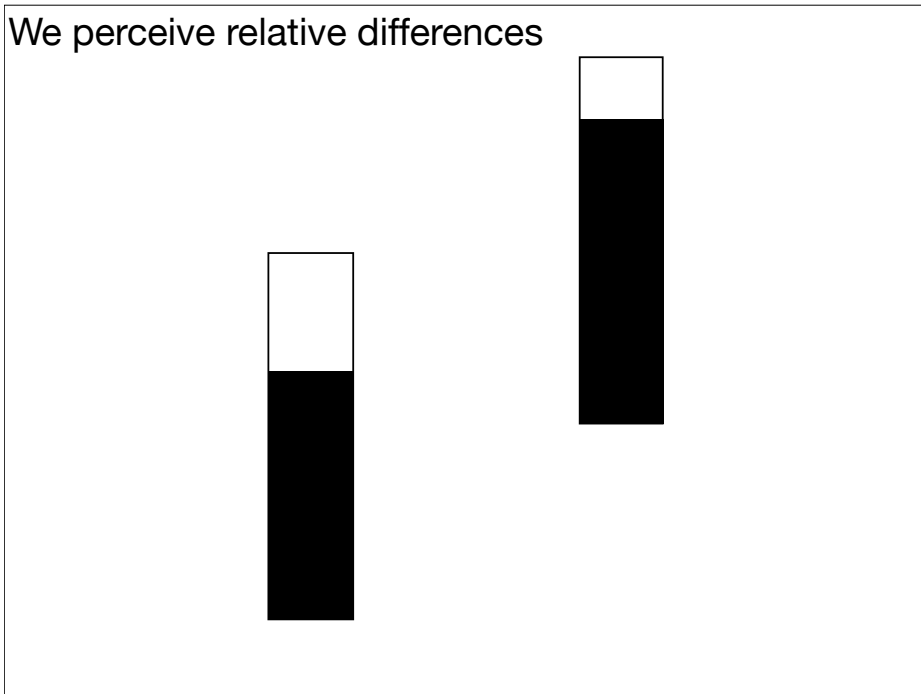
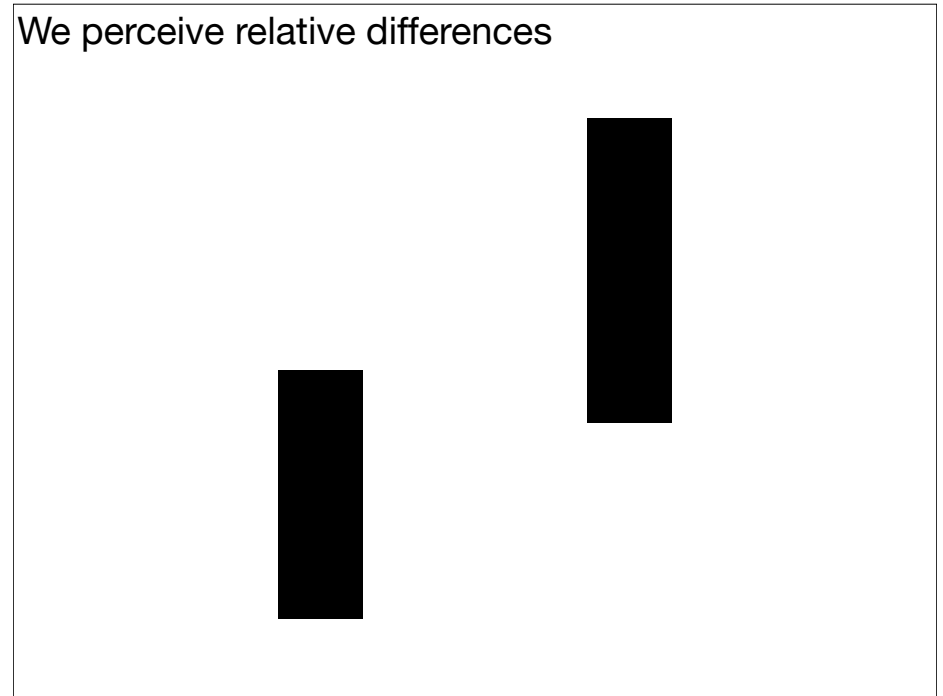
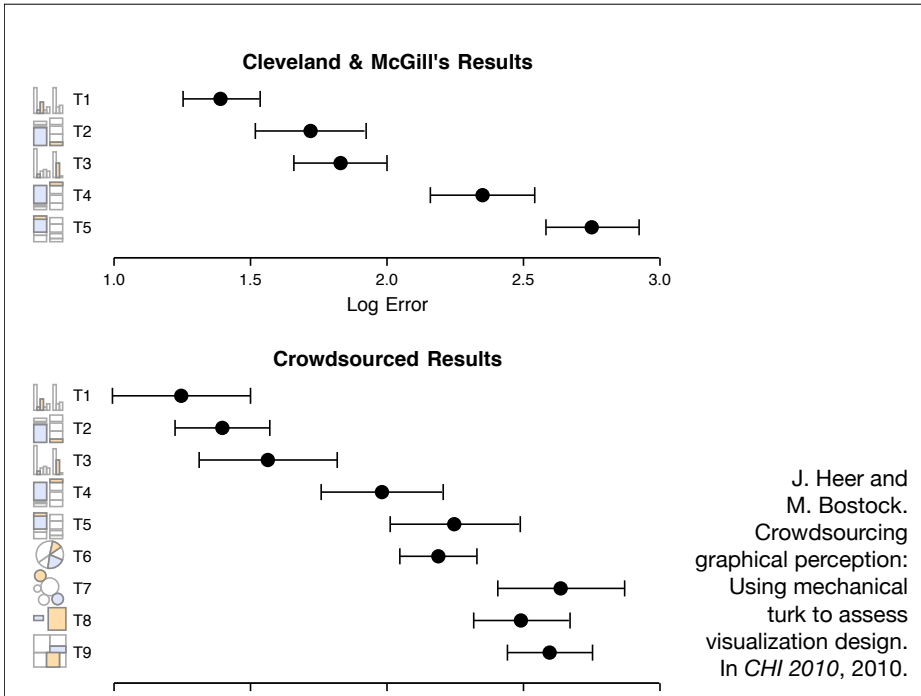
In small groups, work through each of the three graphics. Does the data topology match the perceptual topology?

# 2

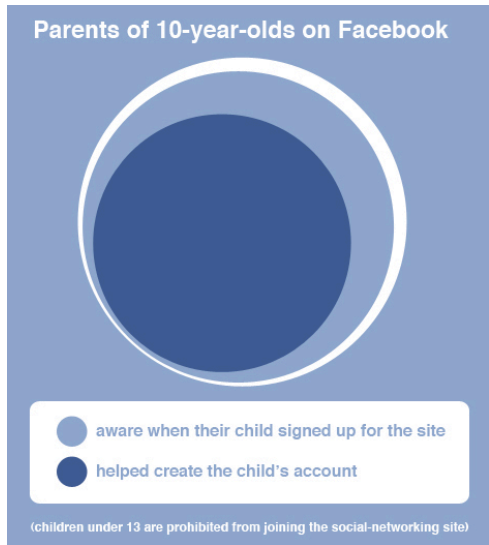
Make important comparisons easy



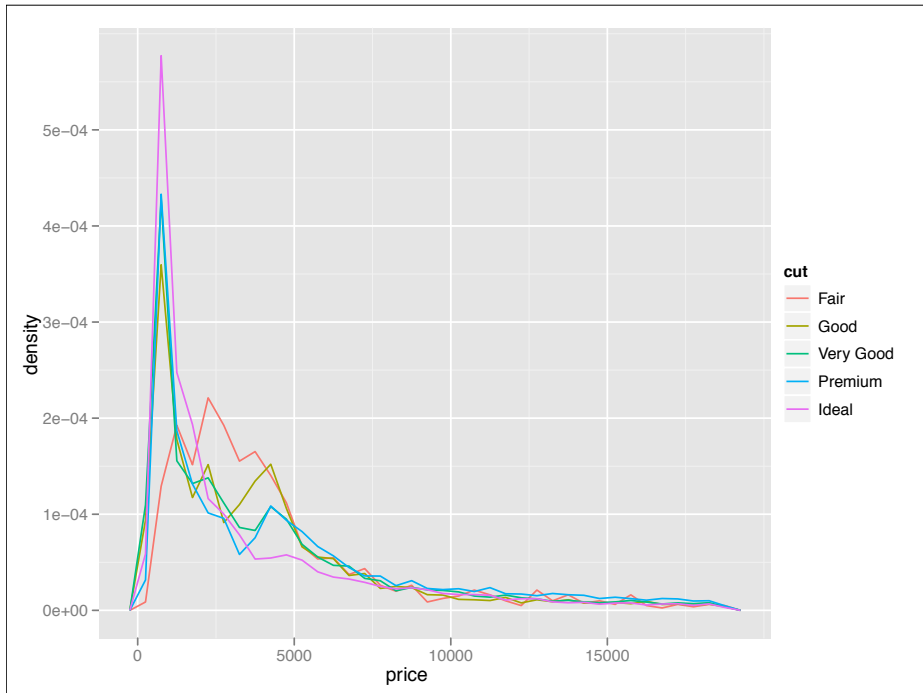
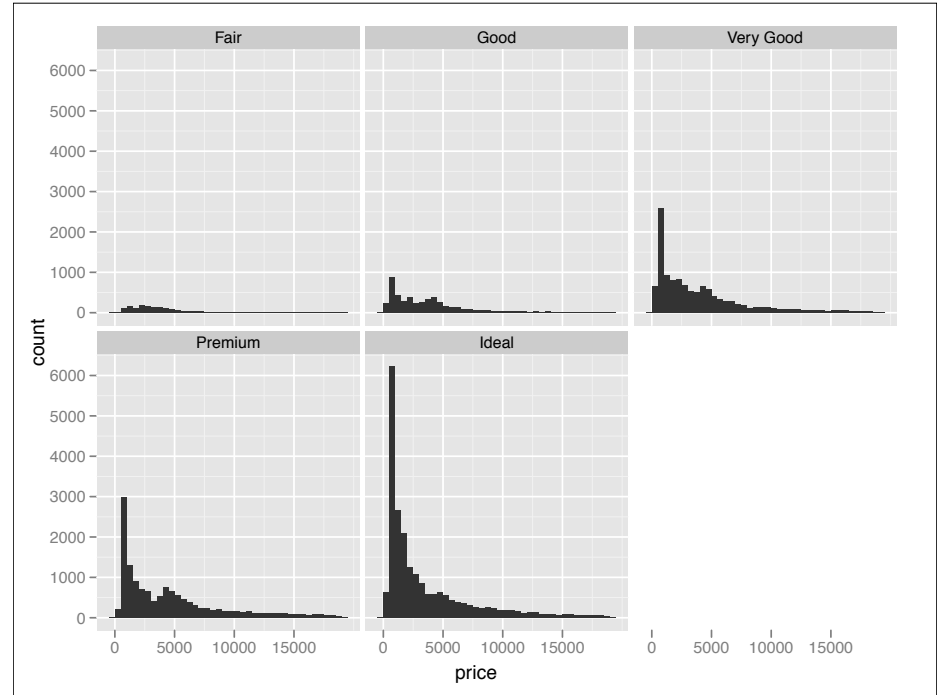




- Position
- Length / Angle
- Area
- Volume / Chroma / Luminance
- ×
- Close objects are easier to compare than distant objects
- ×
- Perception is relative



<http://www.dreamsystemsmedia.com/blog/index.php/social-media-statistics-of-the-day/>



## Common misunderstanding

Pie charts are bad! **Die pie chart, DIE**

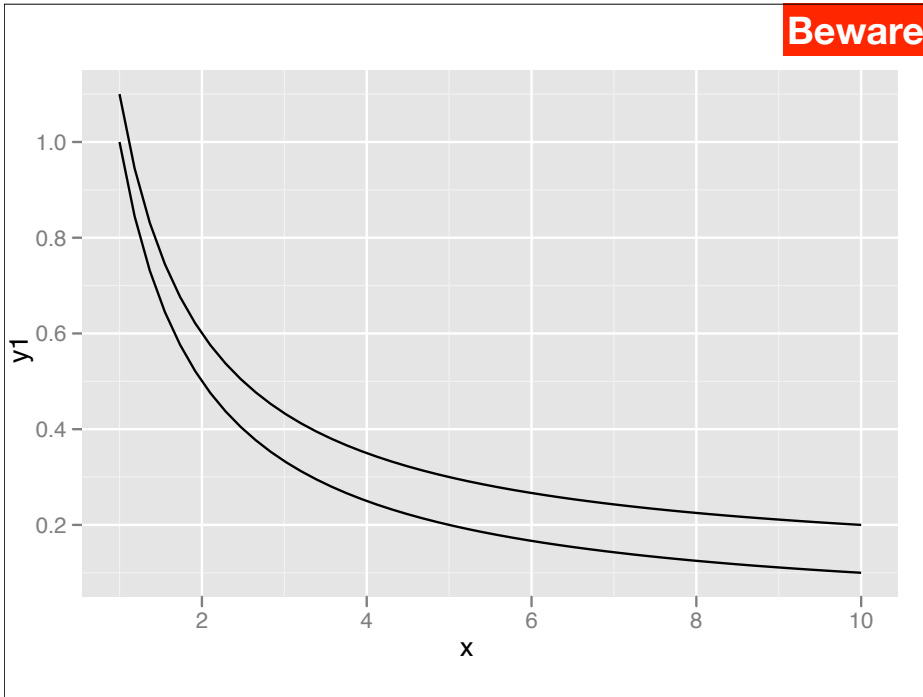
Pie charts are bad when you want to accurately compare two numbers

**But:**

As good as bars for estimating percentage of whole.  
Better than bars for comparing compound proportions (A + B vs C + D)

I. Spence. No Humble Pie: The Origins and Usage of a Statistical Chart. *Journal of Educational and Behavioral Statistics*, 30:353–368, 2005.

Beware



Beware

[http://www.michaelbach.de/ot/sze\\_sinellusion/index.html](http://www.michaelbach.de/ot/sze_sinellusion/index.html)

Beware

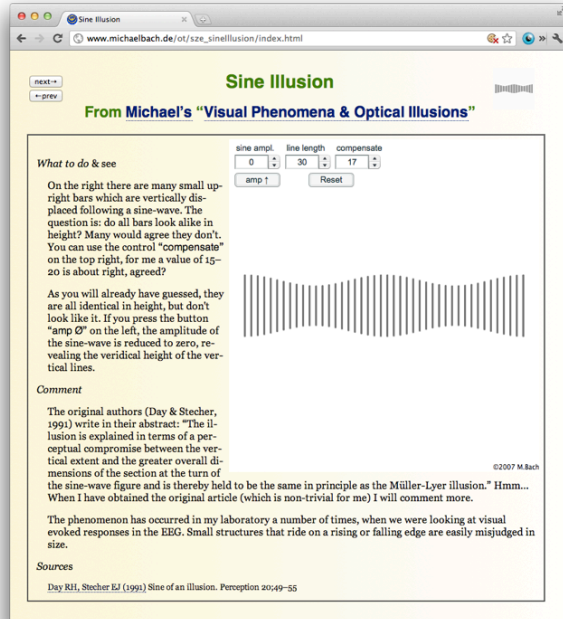
[http://www.michaelbach.de/ot/sze\\_sinellusion/index.html](http://www.michaelbach.de/ot/sze_sinellusion/index.html)

Beware

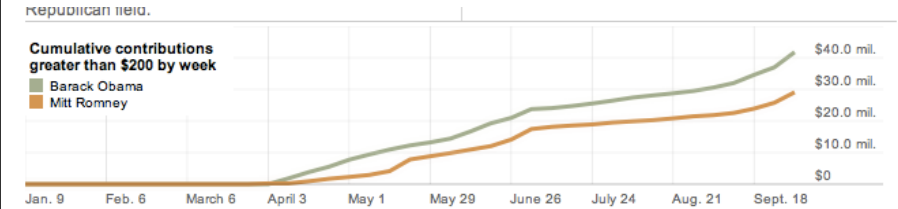
[http://www.michaelbach.de/ot/sze\\_sinellusion/index.html](http://www.michaelbach.de/ot/sze_sinellusion/index.html)



Beware



[http://www.michaelbach.de/ot/sze\\_sinellusion/index.html](http://www.michaelbach.de/ot/sze_sinellusion/index.html)



<http://elections.nytimes.com/2012/campaign-finance>

Ensure important comparisons are close

Use position, then length/area, then chroma/luminance

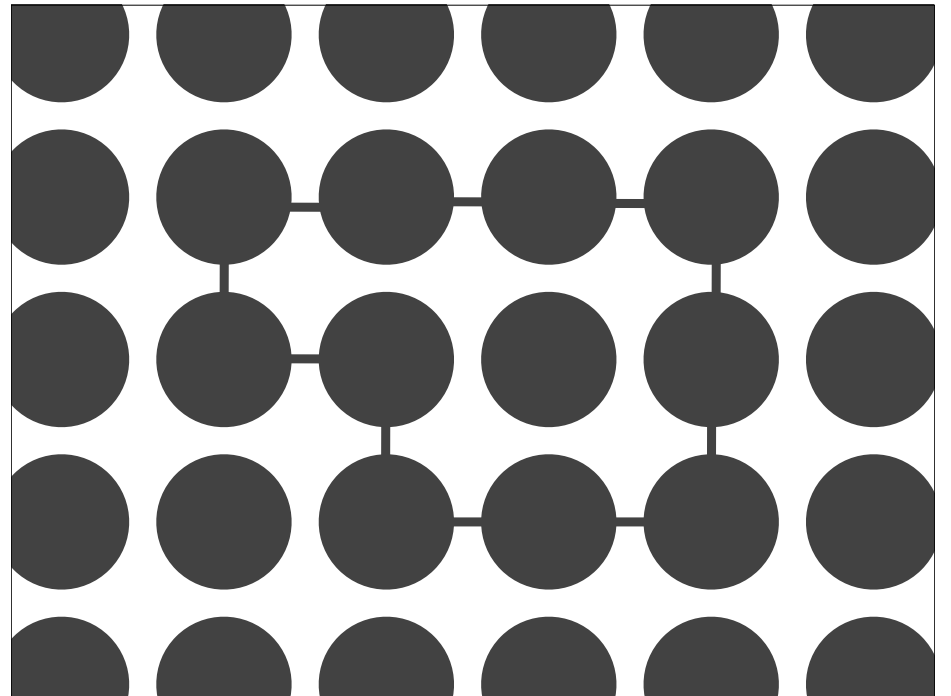
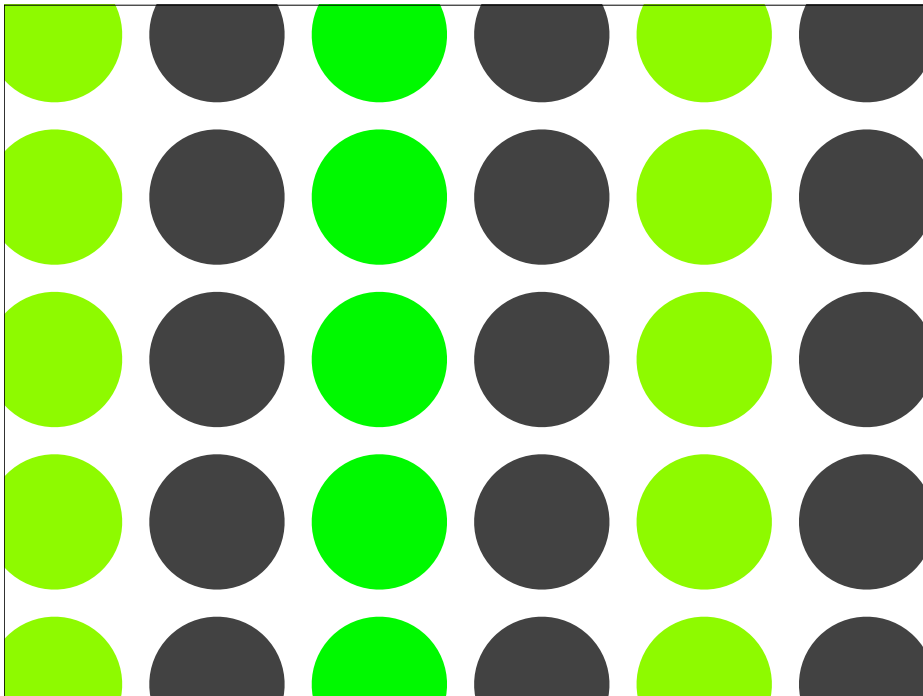
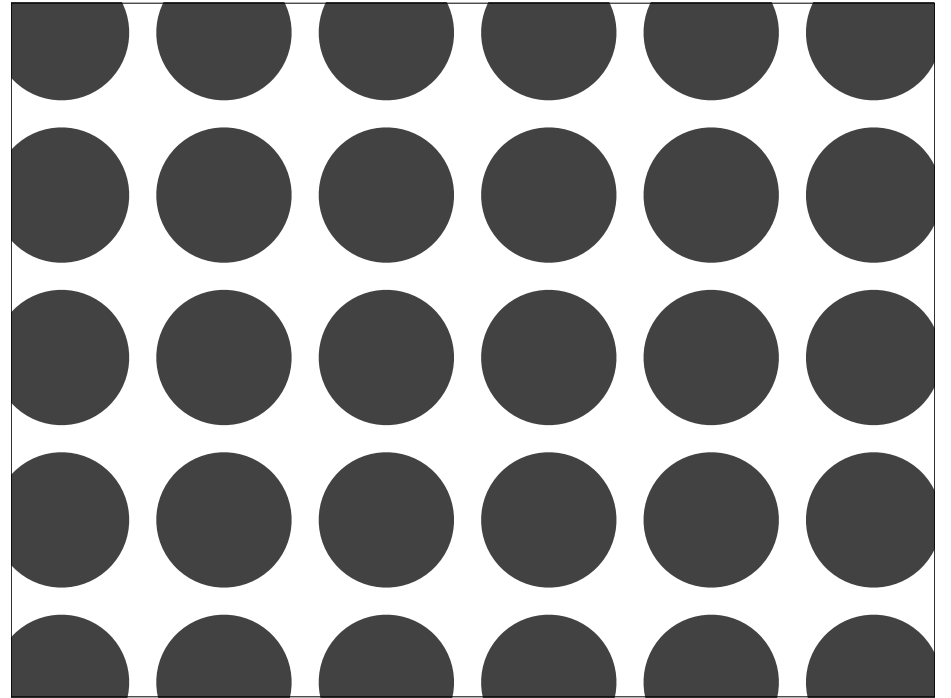
If possible, display comparisons directly

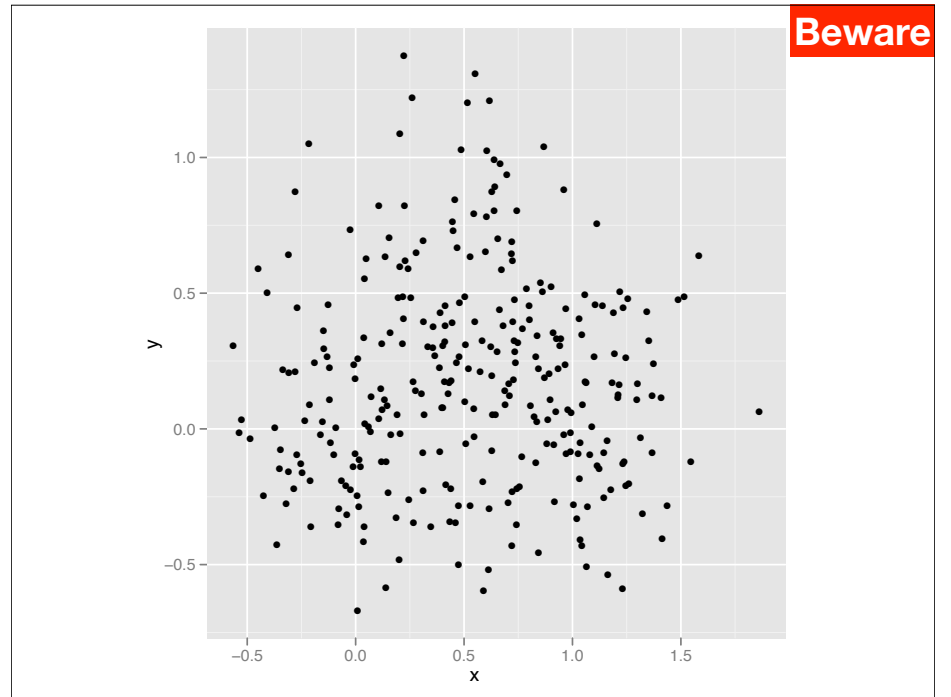
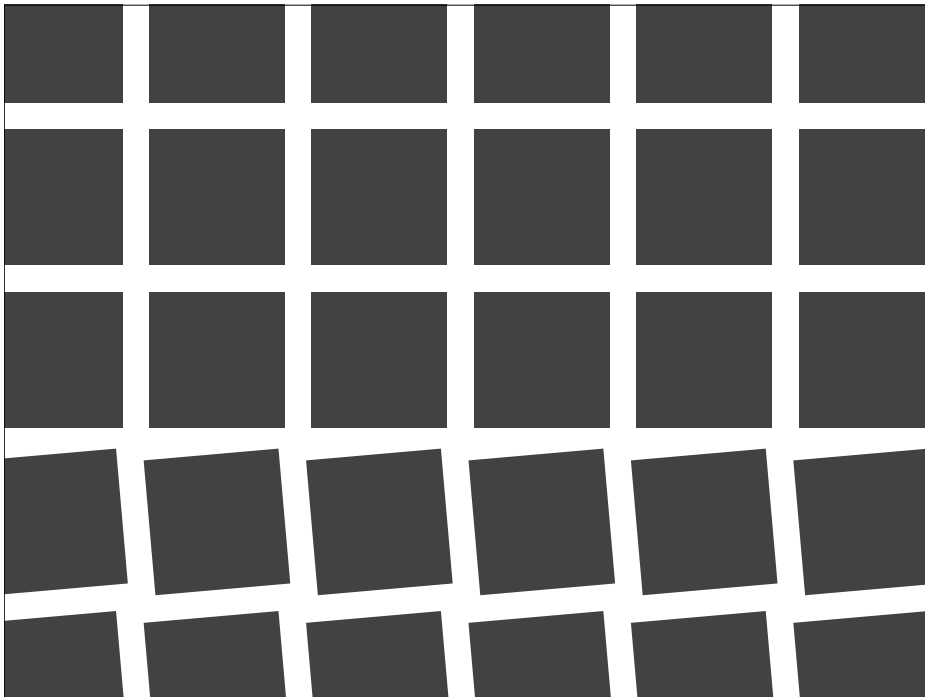
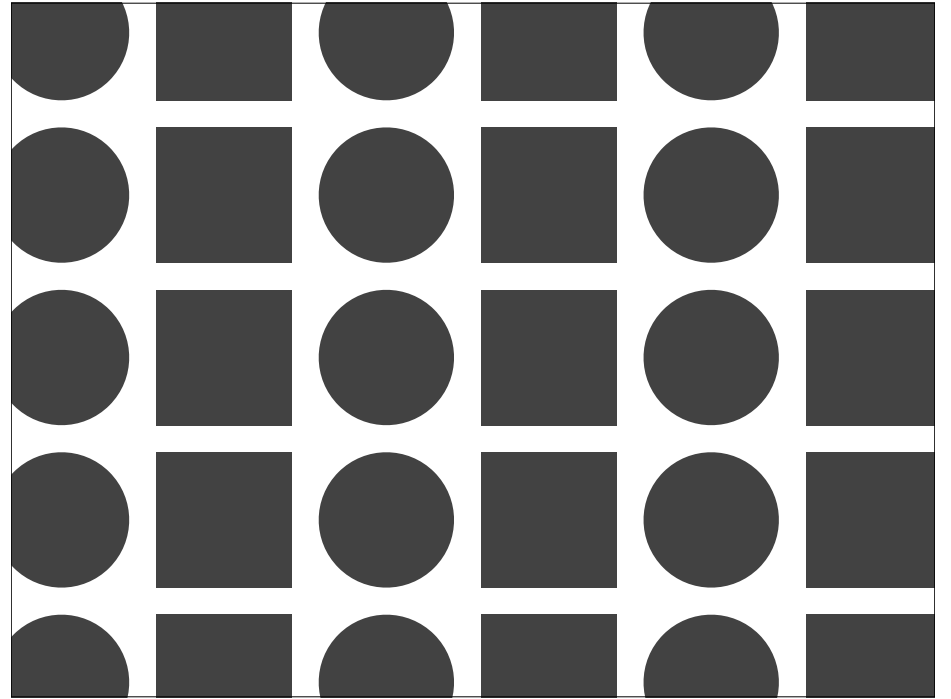
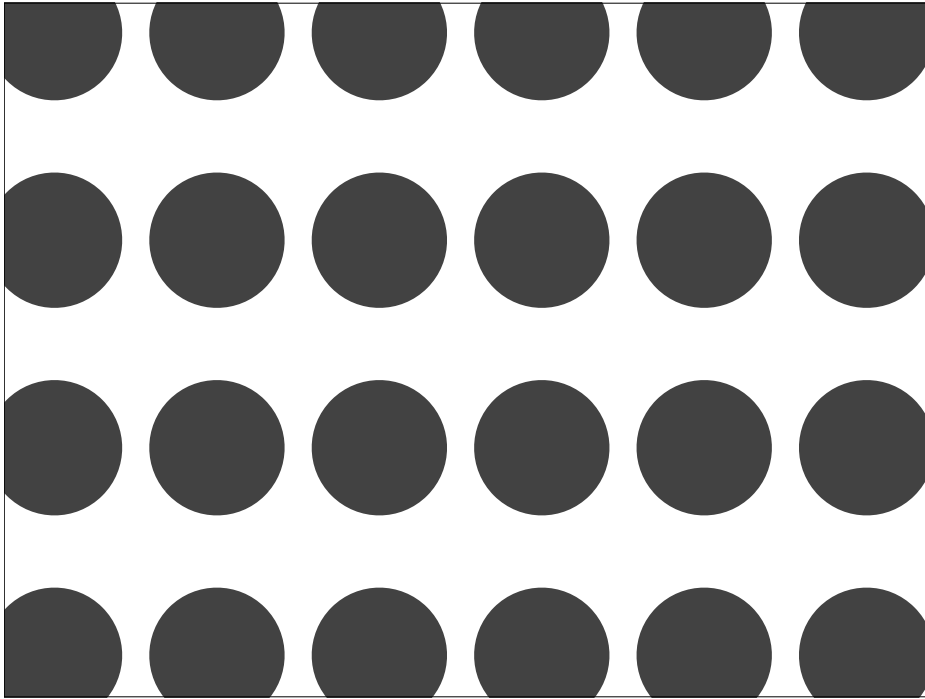
Your turn

In small groups, work through each of the three graphics. What are the important comparisons? What's easy to do and what's hard to do?

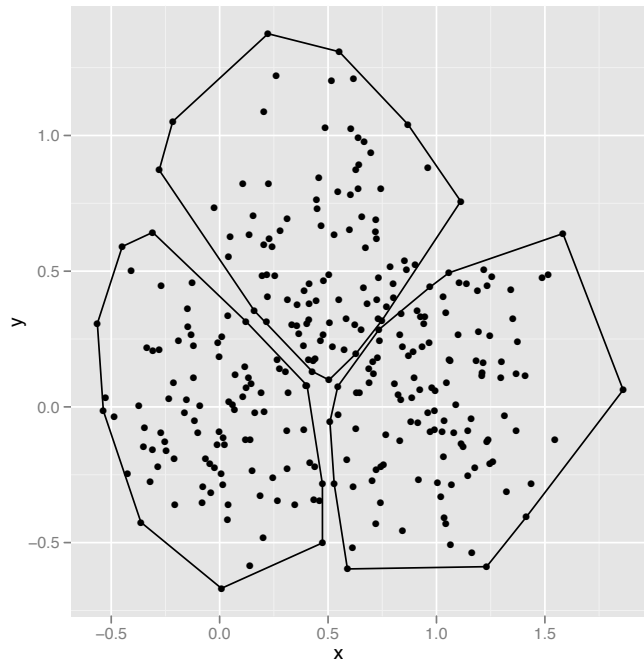
# 3

Visual connections  
should reflect real  
connections





**Beware**



## Your turn

In small groups, work through each of the three graphics. Are components of the graphics appropriately connected?

# 4

Beware of animation!  
(Compare in space, not time)

We often don't notice abrupt changes <http://youtu.be/FWSxSQspiQ?t=0m12s>

We often miss gradual changes too <http://youtu.be/1nL5ulsWMYc>

And movement makes us miss other changes

<http://visionlab.harvard.edu/silencing/>

## Your turn

In small groups, work through the three graphics. (Use the online version of the facebook graphic at <http://nyti.ms/NEglDh>) How has animation been used? Is it effective or ineffective?

A large, bold, black number '1' is positioned on the left side of the box.

Match perceptual  
and data topology

A large, bold, black number '2' is positioned on the left side of the box.

Make important  
comparisons easy

A large, bold, black number '3' is positioned on the left side of the box.

Visual connections  
should reflect real  
connections

# 4

Beware of animation!

# 5\*

Visualisation is only one part of data analysis

# Your turn

## Your turn

In your groups, discuss some of the graphics you've bought along. What works well? What could you do better? I'll circulate and help you out.



## Other sources

<http://projects.nytimes.com/census/2010/map>

<http://kevinquealy.com/>

<http://flowingdata.com/category/visualization/infographics/>

<http://flowingdata.com/category/visualization/statistical-visualization/>

**More  
resources**

## More

<http://chartsnthings.tumblr.com/>

<http://junkcharts.typepad.com/>

<http://flowingdata.com/2012/04/27/data-and-visualization-blogs-worth-following/>