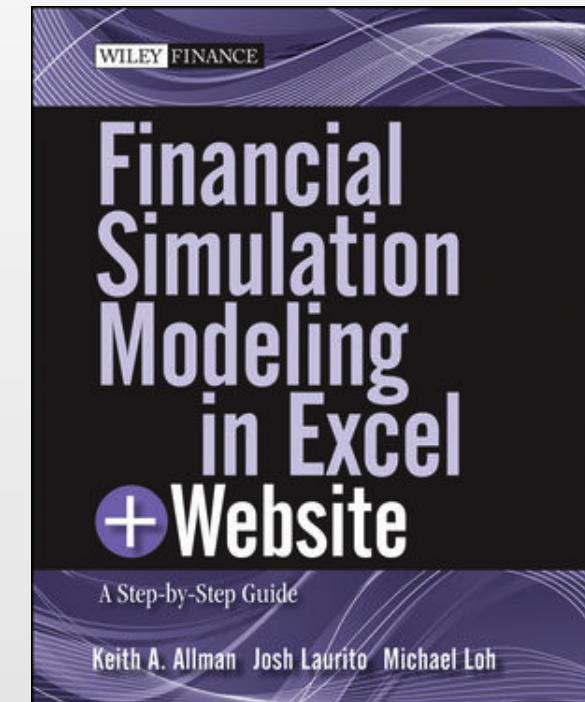


Knowledge & Visual Analytics

Lecture 1

Josh Laurito

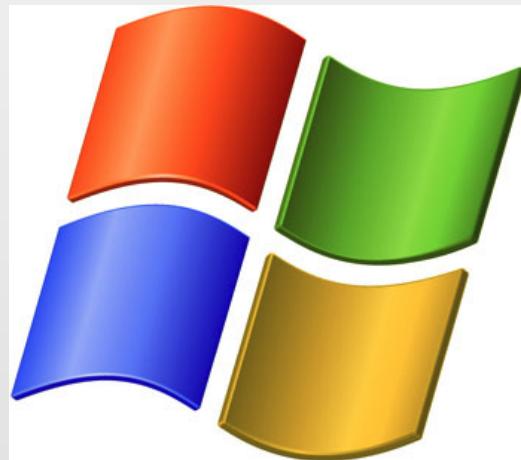
Who Am I?



What Do You Do?

Tech? Industries? Goals for this class?

Please fill out initial survey



images via wikipedia



What is this Class About?

How to present information

How to present quantitative information

How to present quantitative information graphically

How to use technology to present quantitative information graphically

General Class Info

All class dates are in the syllabus

Every class, a short quiz and a homework will be due

At the end of the semester, you will complete a larger, public project

If you have questions, email me at josh.laurito@gmail.com (or ping me on slack!)



Important Links

https://github.com/jlaurito/CUNY_IS608

<http://tinyletter.com/nycdatajobs/>

http://jlaurito.github.io/CUNY_IS608/

<https://cunymsda2015.slack.com/> (Thanks James!)

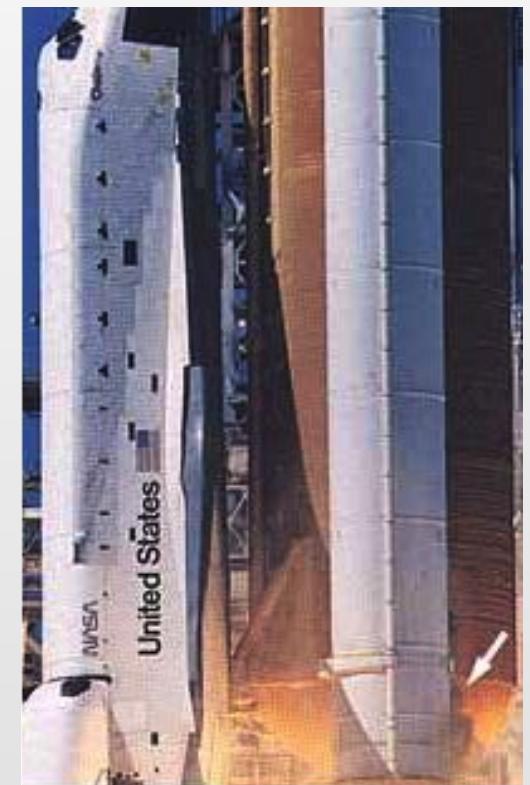
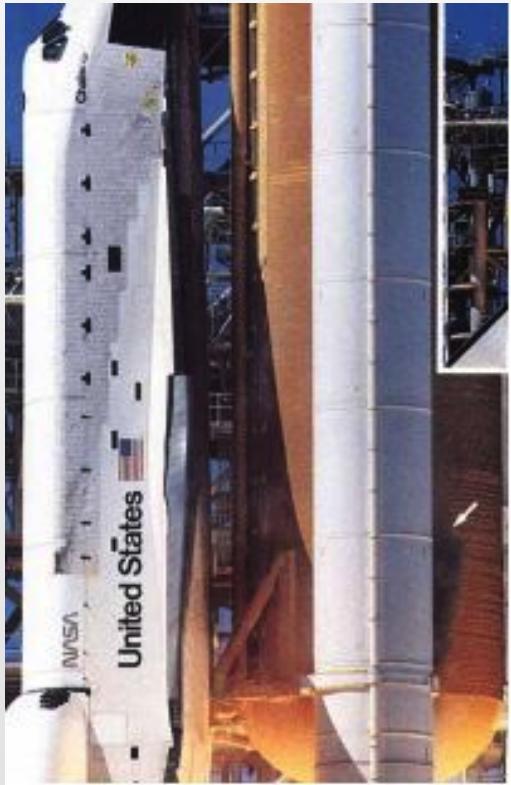
Visualization Technique & Theory

Importance of Data Visualization



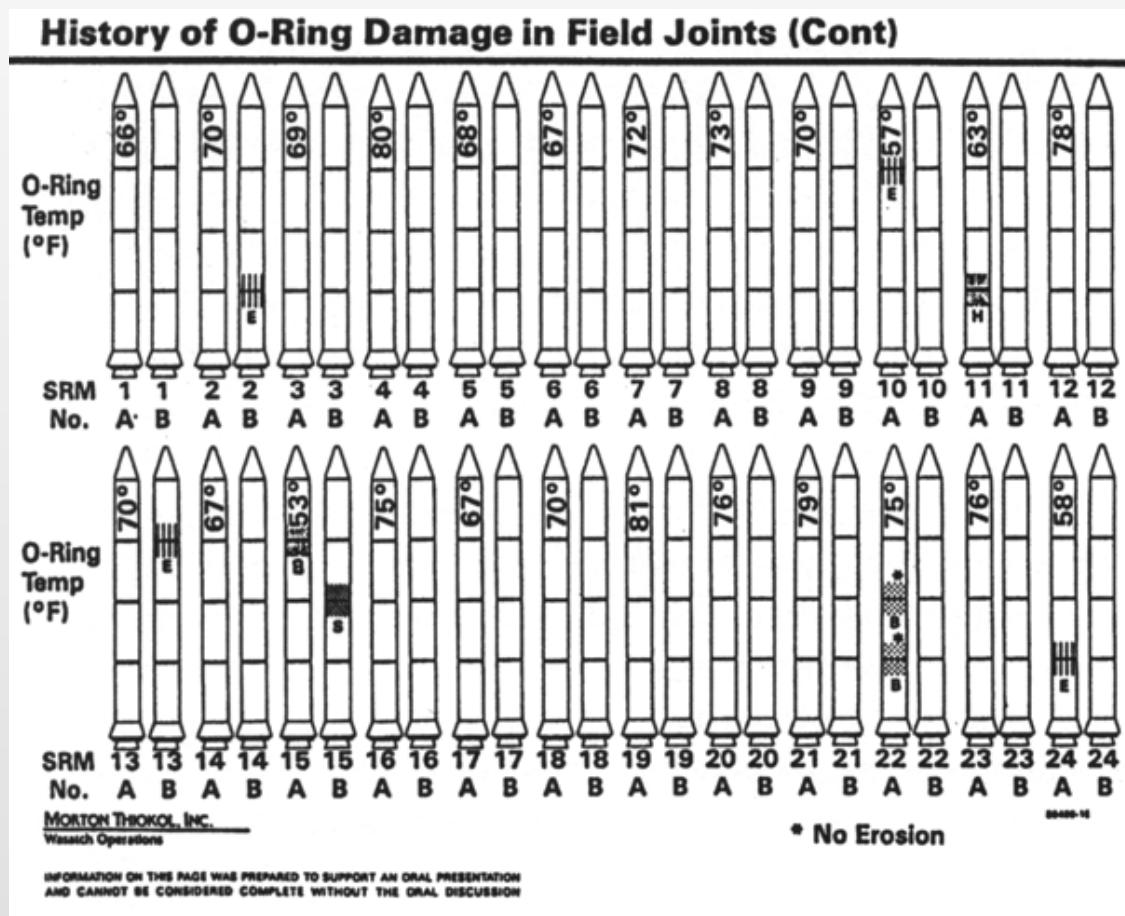
http://news.bbc.co.uk/2/shared/spl/hi/pop_ups/06/sci_nat_1986_challenger_disaster/html/1.stm

Importance of Data Visualization



<http://www.aerospaceweb.org/question/investigations/q0122.shtml>

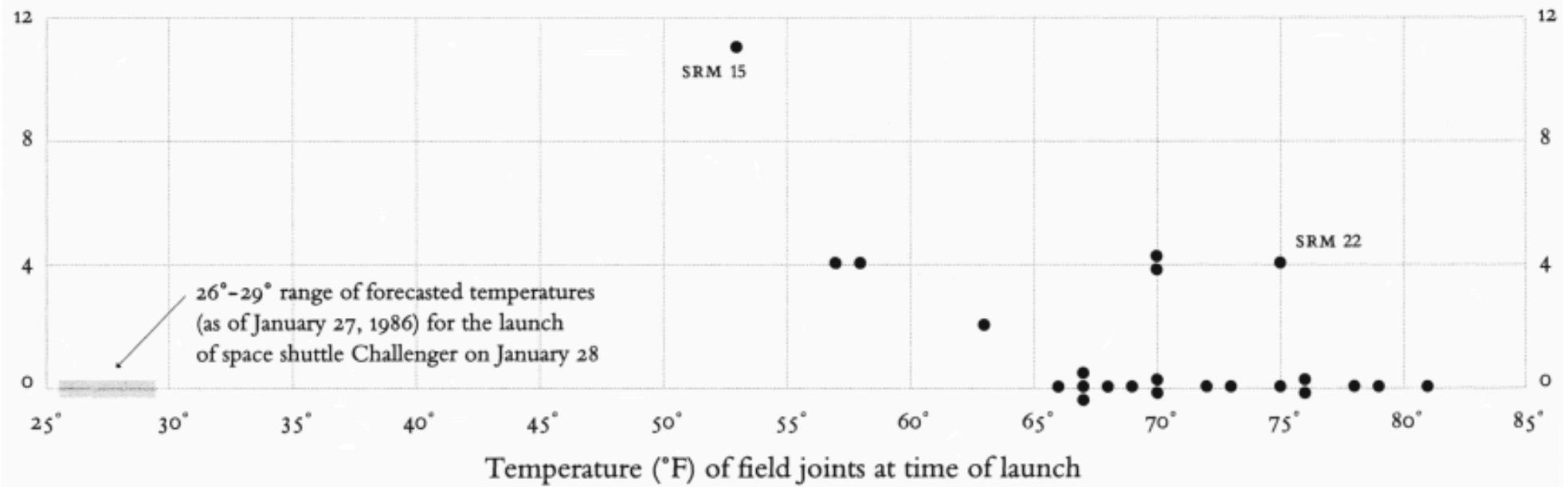
Importance of Data Visualization



From Tufte's *Visual Explanations*

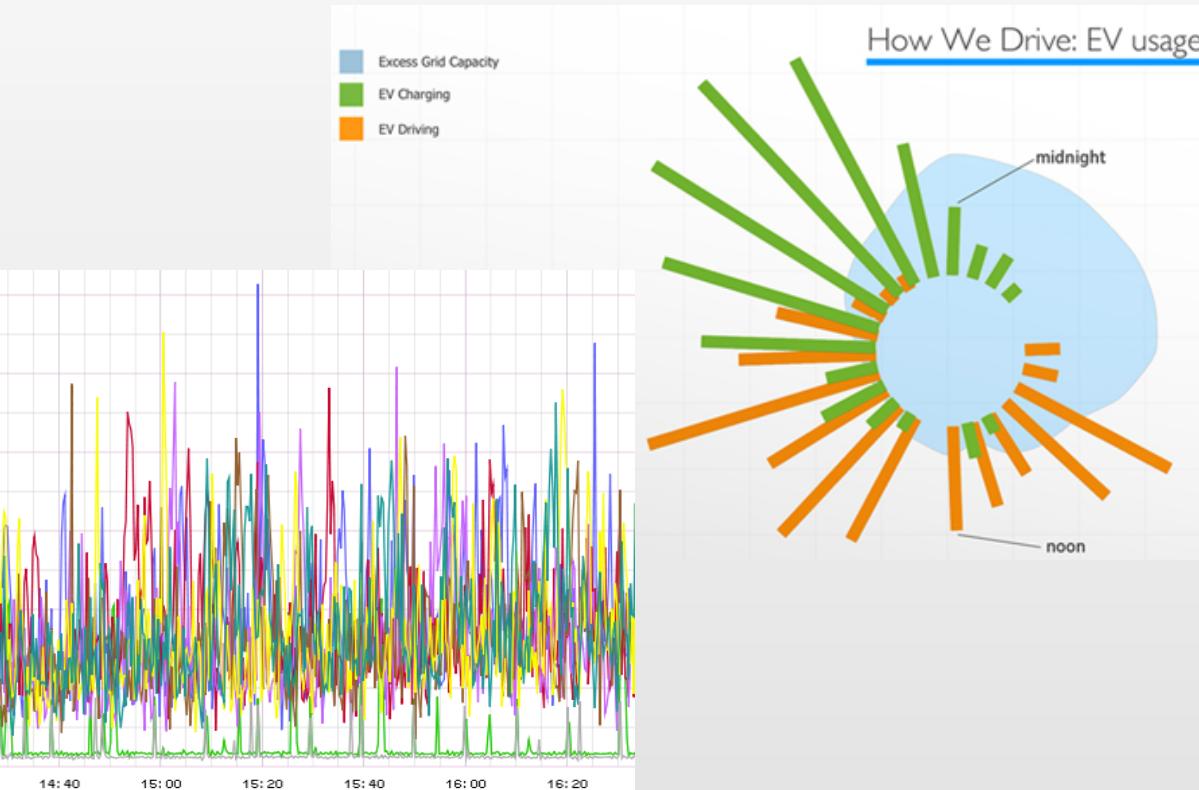
Importance of Data Visualization

O-ring damage
index, each launch



From Tufte's *Visual Explanations*

Other Bad Visualizations



[/blog.visual.ly/poor-visualization-can-do-more-harm-than-good/](http://blog.visual.ly/poor-visualization-can-do-more-harm-than-good/)
[/www.nbcchicago.com/news/local/FOX-News-Chart-Fails-Math-73711092.html](http://www.nbcchicago.com/news/local/FOX-News-Chart-Fails-Math-73711092.html)
[/themonkeycage.org/2013/01/08/how-2012-stacks-up-the-worst-graph-on-record/](http://themonkeycage.org/2013/01/08/how-2012-stacks-up-the-worst-graph-on-record/)

Goals of Data Visualization

Data Visualization has a few primary goals:

Minimize mental effort (images are more clear than numbers)

Make information easy to retain (memory is often visual)

Insight over beauty ('Aha!' is more important than 'wow')

Visualization Strategies

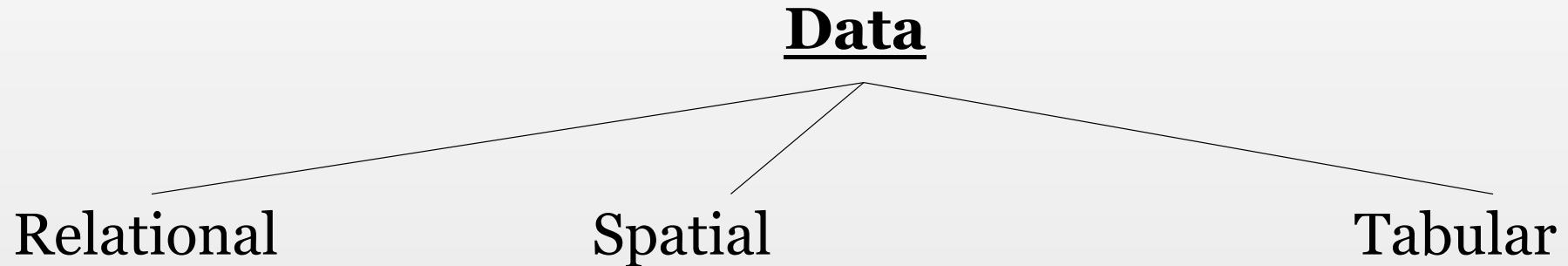
Ways to make visualization **clear, memorable and insightful:**

Match display to data and purpose

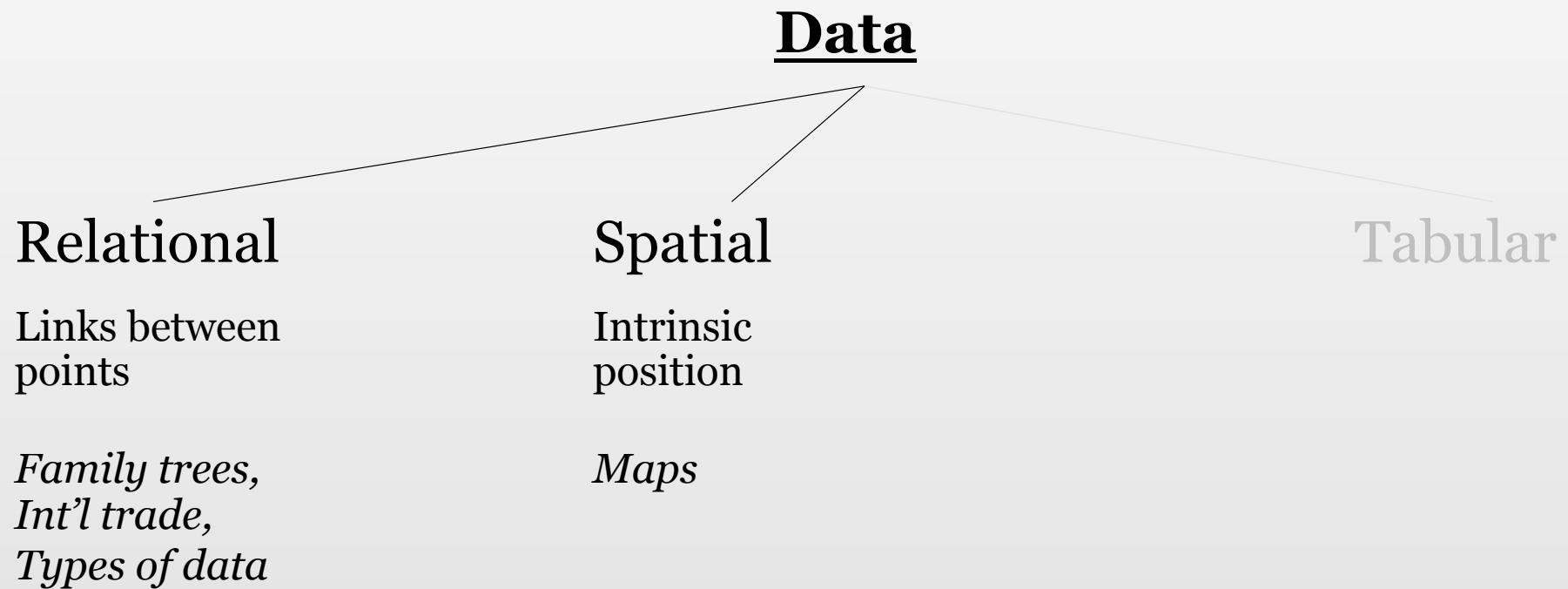
Maximize information-to-ink ratio

Principal of small multiples

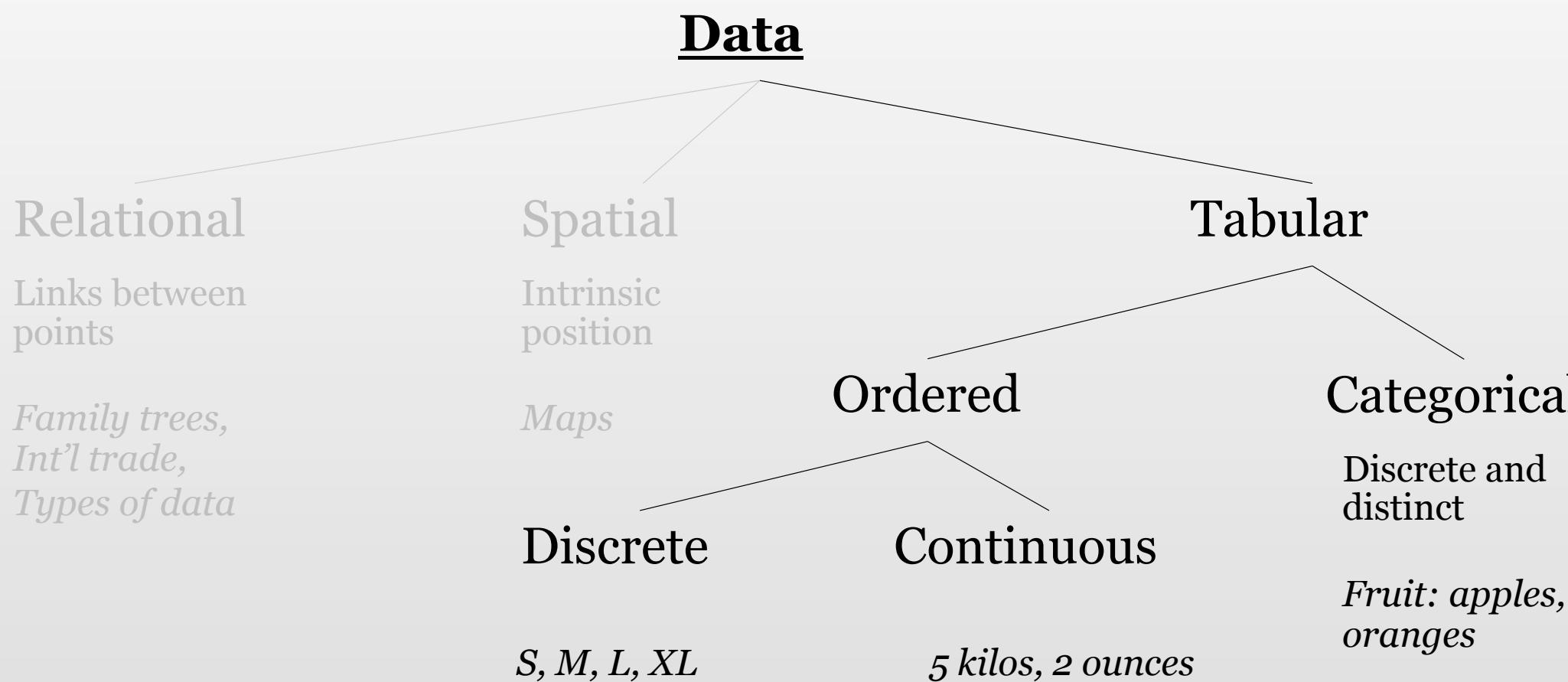
Match Display to Data and Purpose



Match Display to Data and Purpose



Match Display to Data and Purpose



Match Display to Data and Purpose

Relational

Spatial

Ordered
discrete/cont.

Categorical

Match Display to Data and Purpose

Relational

Spatial

Ordered
discrete/cont.

Categorical

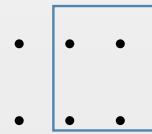
Maps



Match Display to Data and Purpose

Relational

Contain



Connect



Similarity

- 0 0
- 0 0

Proximity

- • •
- • •

Spatial

Maps



Ordered
discrete/cont.

Categorical

Match Display to Data and Purpose

Relational

Contain · · ·
 · · ·

Connect · ·
 · ·

Similarity · o o
 · o o

Proximity · · ·
 · · ·

Spatial

Maps


Ordered discrete/cont.

Position | · .

Length | —

Area ■ ■

Shade □ □

Saturation □ □

etc...

Categorical

Match Display to Data and Purpose

Relational

Contain	•	•	•	•
	•			
	•	•	•	•
Connect	•	•	•	•
	•			
	•	•	•	•
Similarity	•	○	○	•
	•	○	○	•
Proximity	•	•	•	•
	•	•	•	•

Spatial



Ordered discrete/cont.

Position		•	.
Length		—	—
Area	■	■	■
Shade	□	□	□
Saturation	□	□	□
etc...			

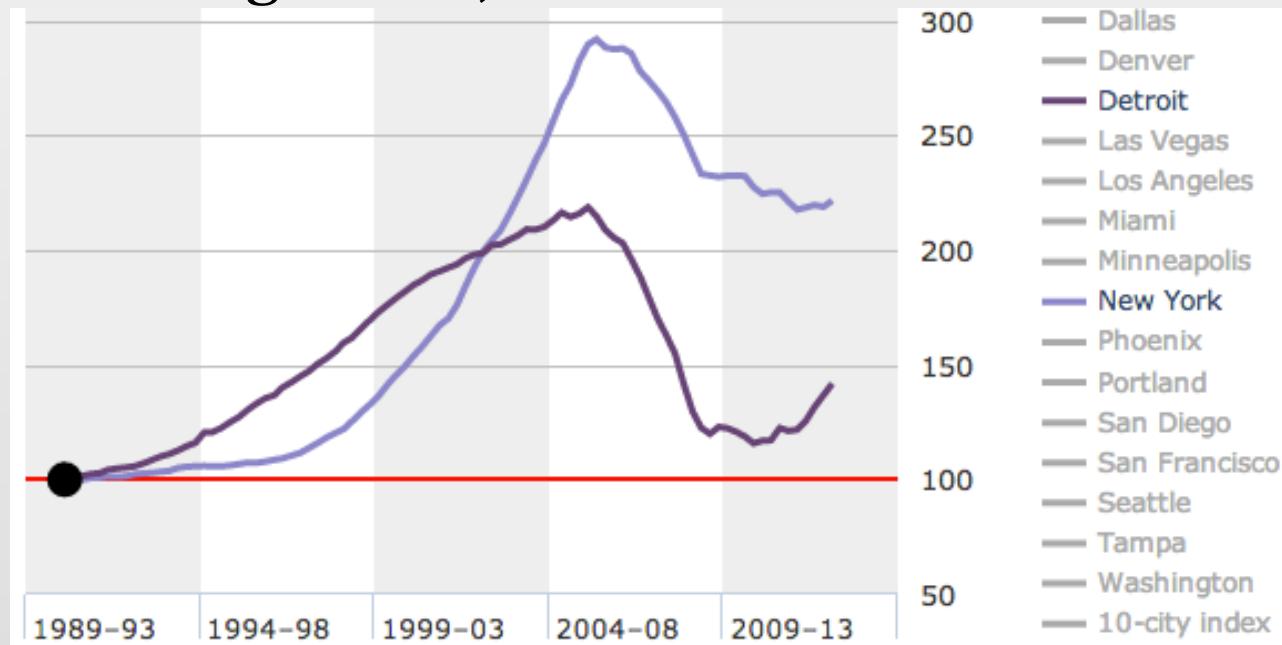
Categorical

Position	•	•	•
	•	•	•
	•	•	•
Color	•	•	•
	•	•	•
Shape	•	○	○
	•	○	○
Patterns	■	■	■
	■	■	■

Match Dimensions to Data

Use highly ranked channels for key info to keep memorable

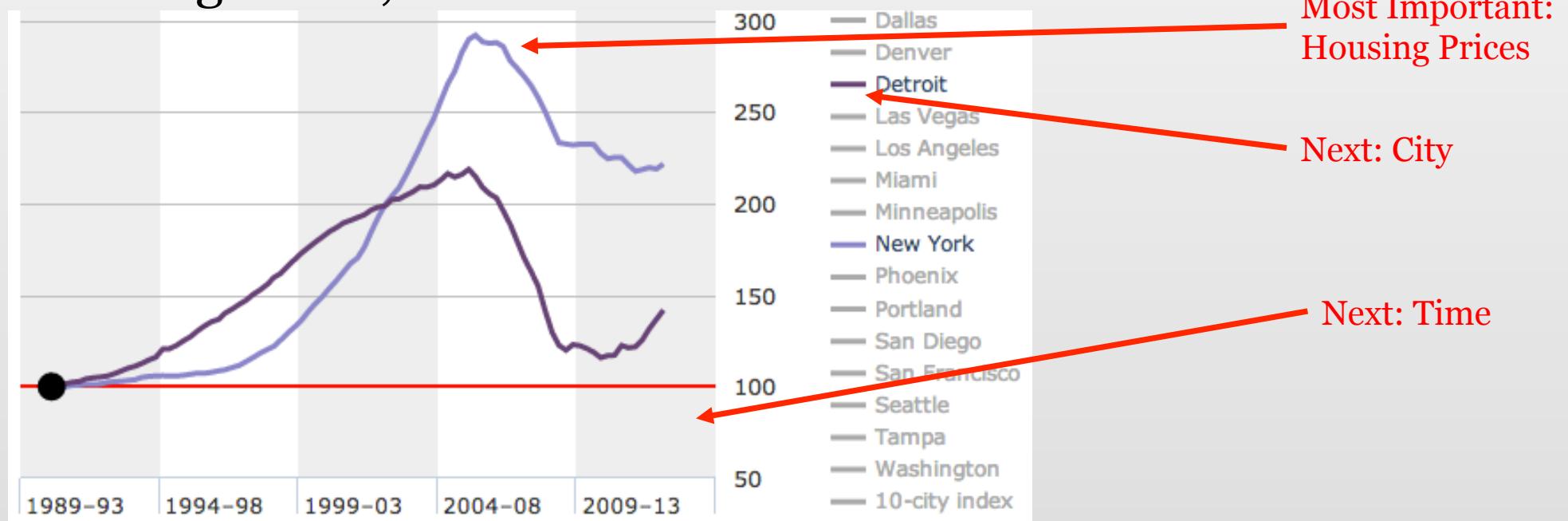
Housing Prices, *The Economist*



Match Dimensions to Data

Use highly ranked channels for key info to keep memorable

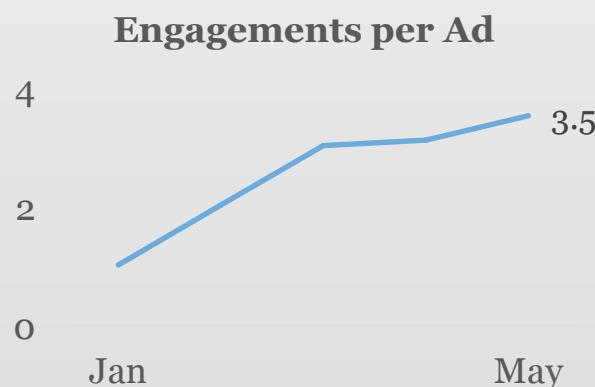
Housing Prices, *The Economist*



Maximize Information-to-Ink Ratio

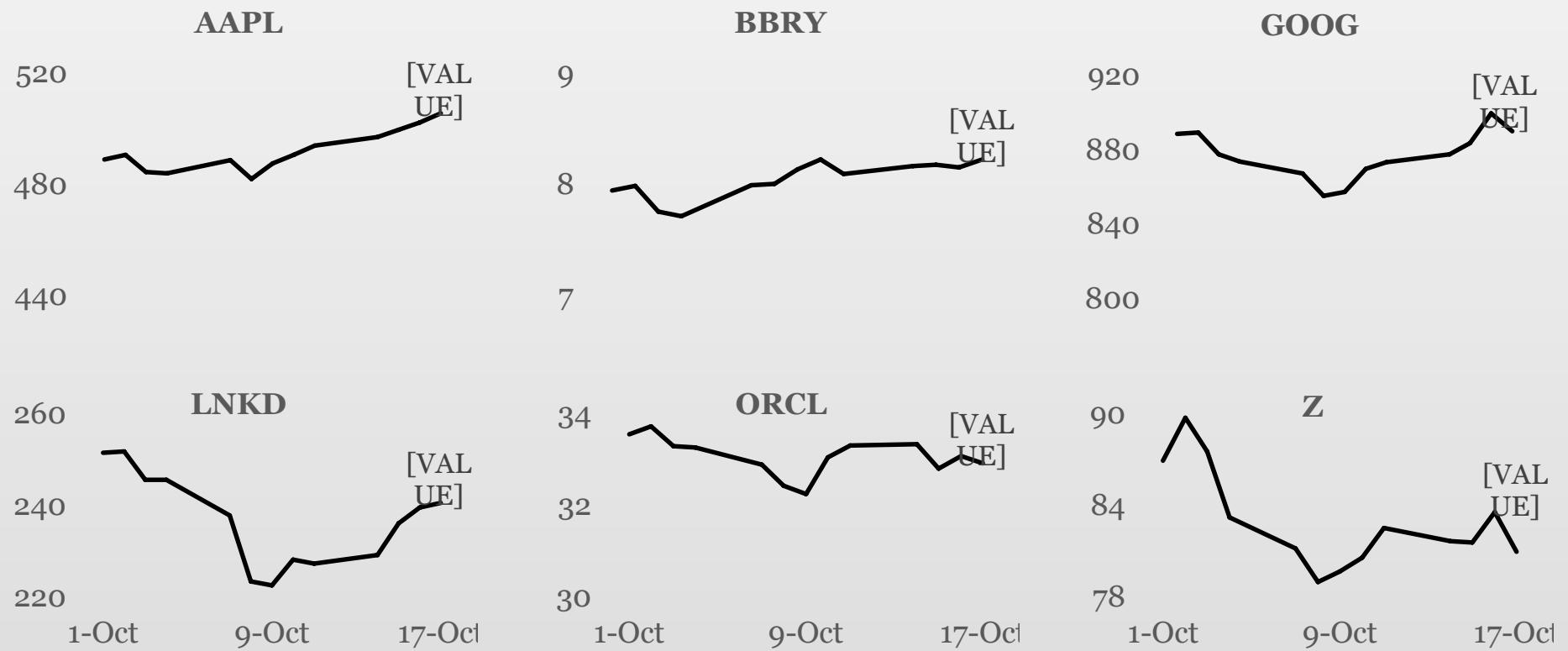
Achieve clarity by removing distractions & distracting elements

Taken to its logical endpoint: **Sparklines**

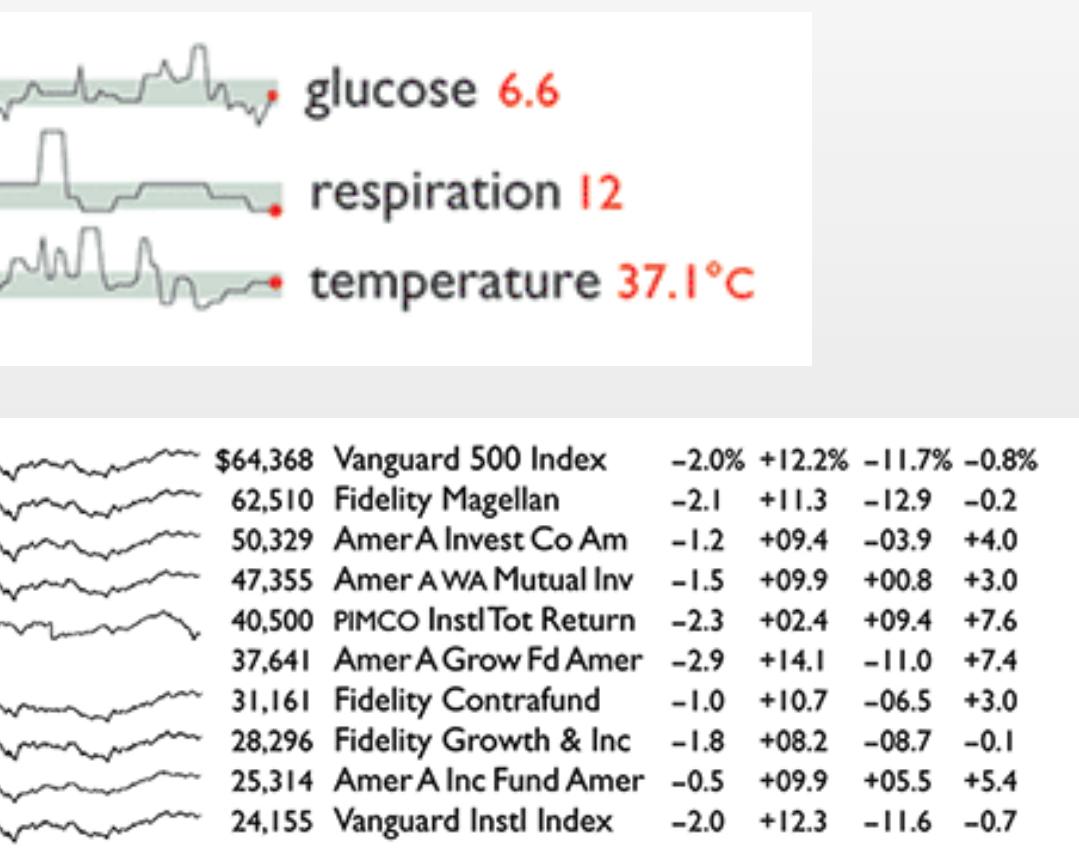


Principal of Small Multiples

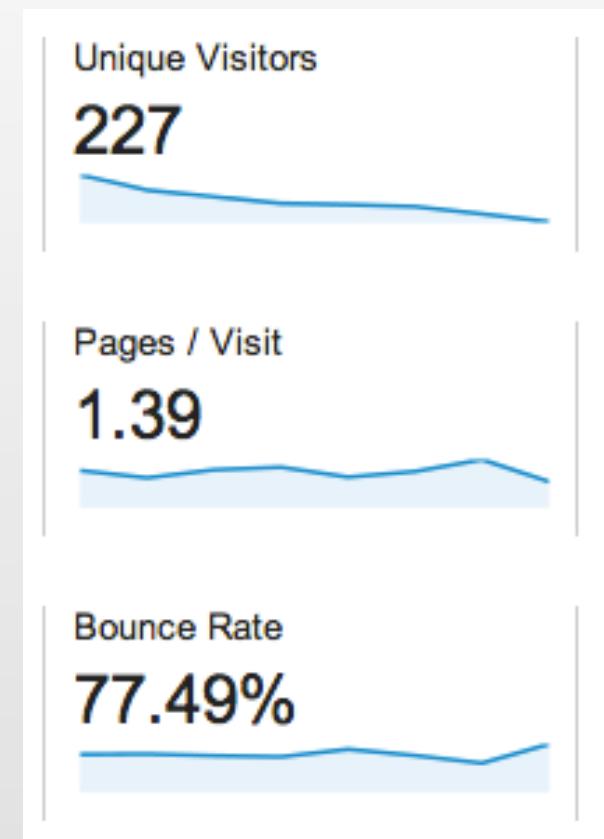
It is easier to compare items next to each other than over-plotted



Principal of Small Multiples In Practice

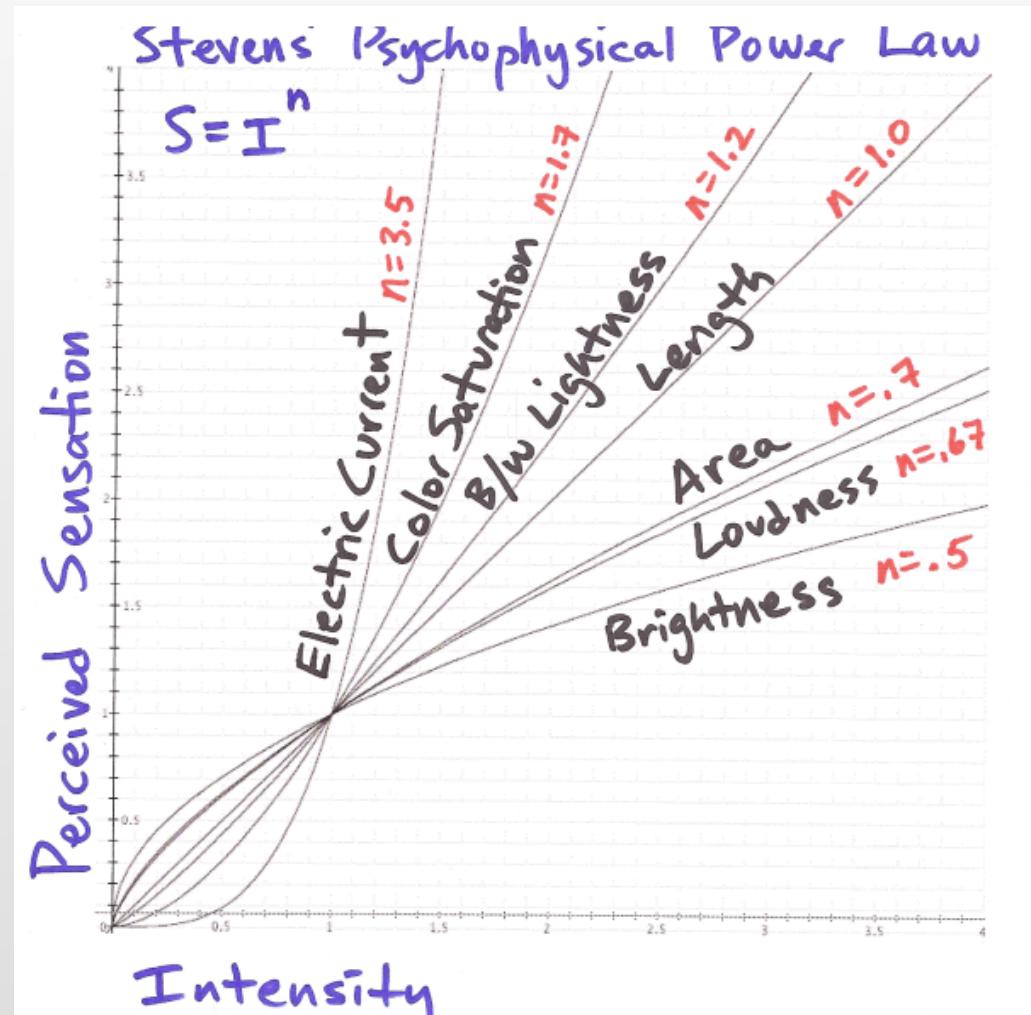


<http://www.edwardtufte.com/>



<https://www.google.com/analytics/>

Things To Keep in Mind - Dimensions



Things To Keep in Mind - Discrimination

~3 for Pie Charts



~7 for colors (not as a scale!)



~4 overlapping lines

Just the Basics

Data Visualization: seems small & cute, actually a huge topic!



image via <http://www.cnn.com>

Further Reading on Visualization

Pioneers:

Jacques Bertin



Edward Tufte



Current Leaders:

Tamara Munzner



Mike Bostock



favorites:
Bertin
Tufte
Munzner
Bostock

Assignment 1 – Business Data, ggplot2



Assignment 1 – Business Data, ggplot2

Read ggplot2 presentation

Submit code to me

Post images with some analysis

Contact me with Questions

josh.laurito@gmail.com is best