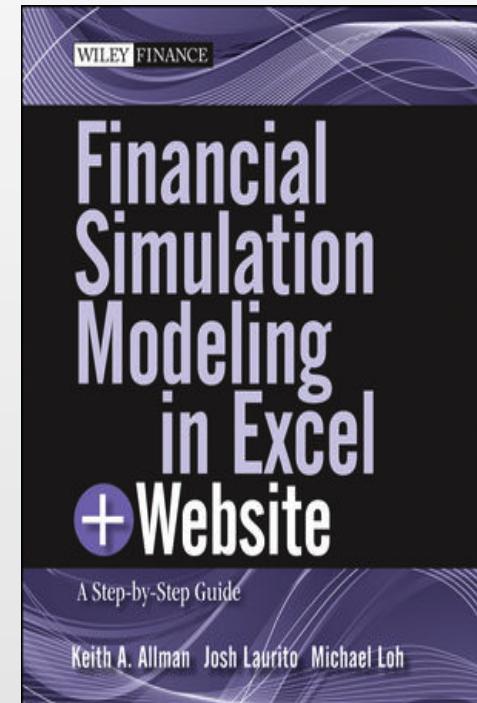


Knowledge & Visual Analytics

Lecture 1

Josh Laurito

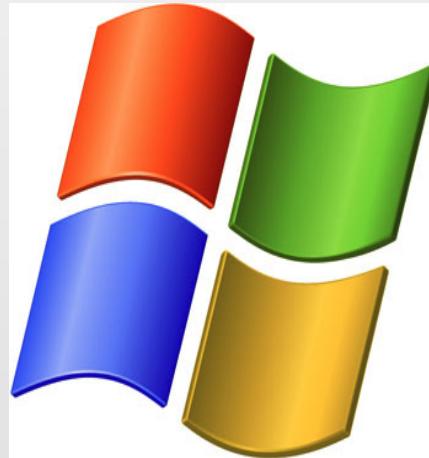
Who Am I?



What Do You Do?

Tech? Industries? Goals for this class?

Please fill out initial survey



images via wikipedia



Mac

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 homework will be due (sometimes a quiz as well)

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/>

[https://www.youtube.com/playlist?
list=PLOqYZ_RIA4vzvZFmHGAF2BGmF4YdXbfgE](https://www.youtube.com/playlist?list=PLOqYZ_RIA4vzvZFmHGAF2BGmF4YdXbfgE)

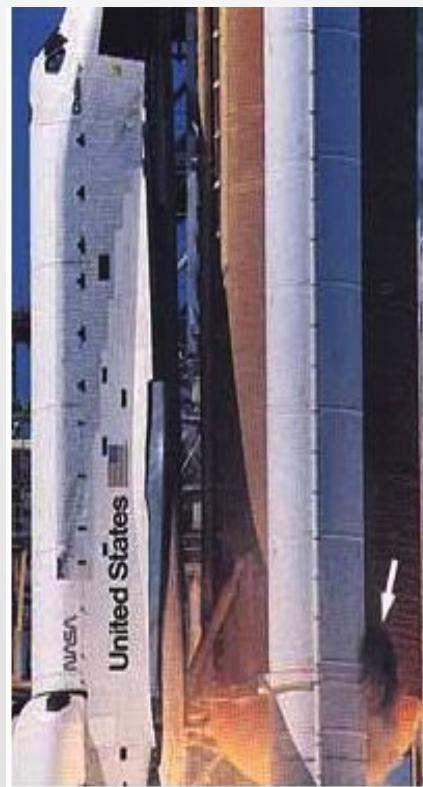
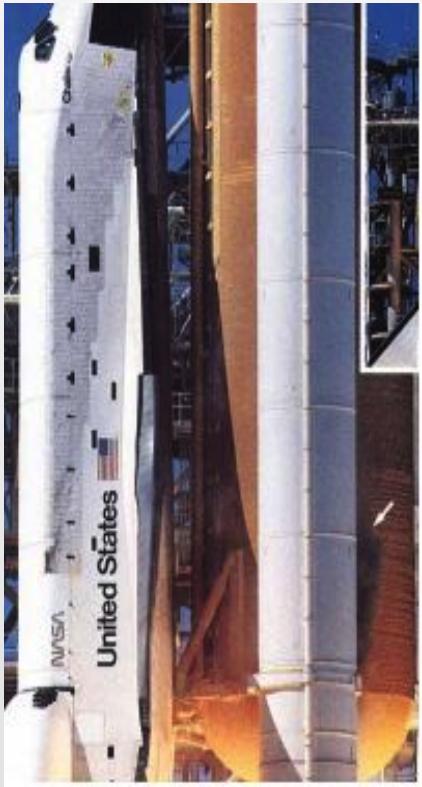
Visualization Technique & Theory

Importance of Data Visualization



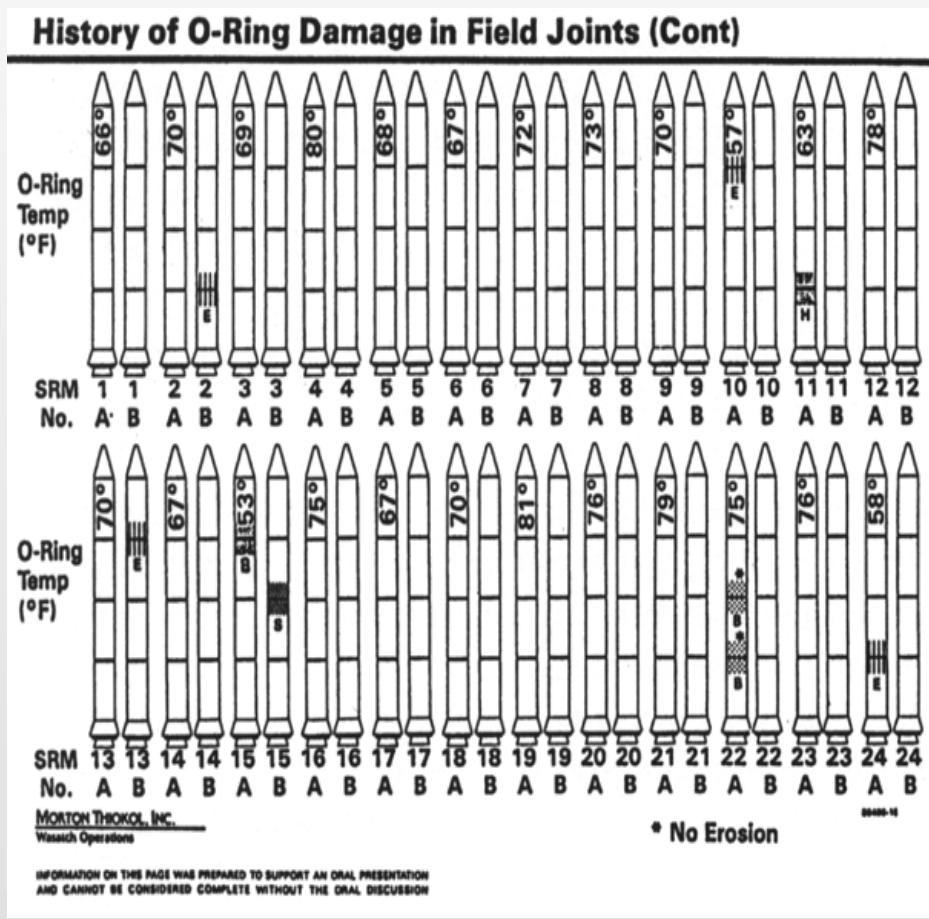
[http://news.bbc.co.uk/2/shared/spl/hi/pop_ups/06/
sci_nat_1986_challenger_disaster/html/1.stm](http://news.bbc.co.uk/2/shared/spl/hi/pop_ups/06/sci_nat_1986_challenger_disaster/html/1.stm)

Importance of Data Visualization



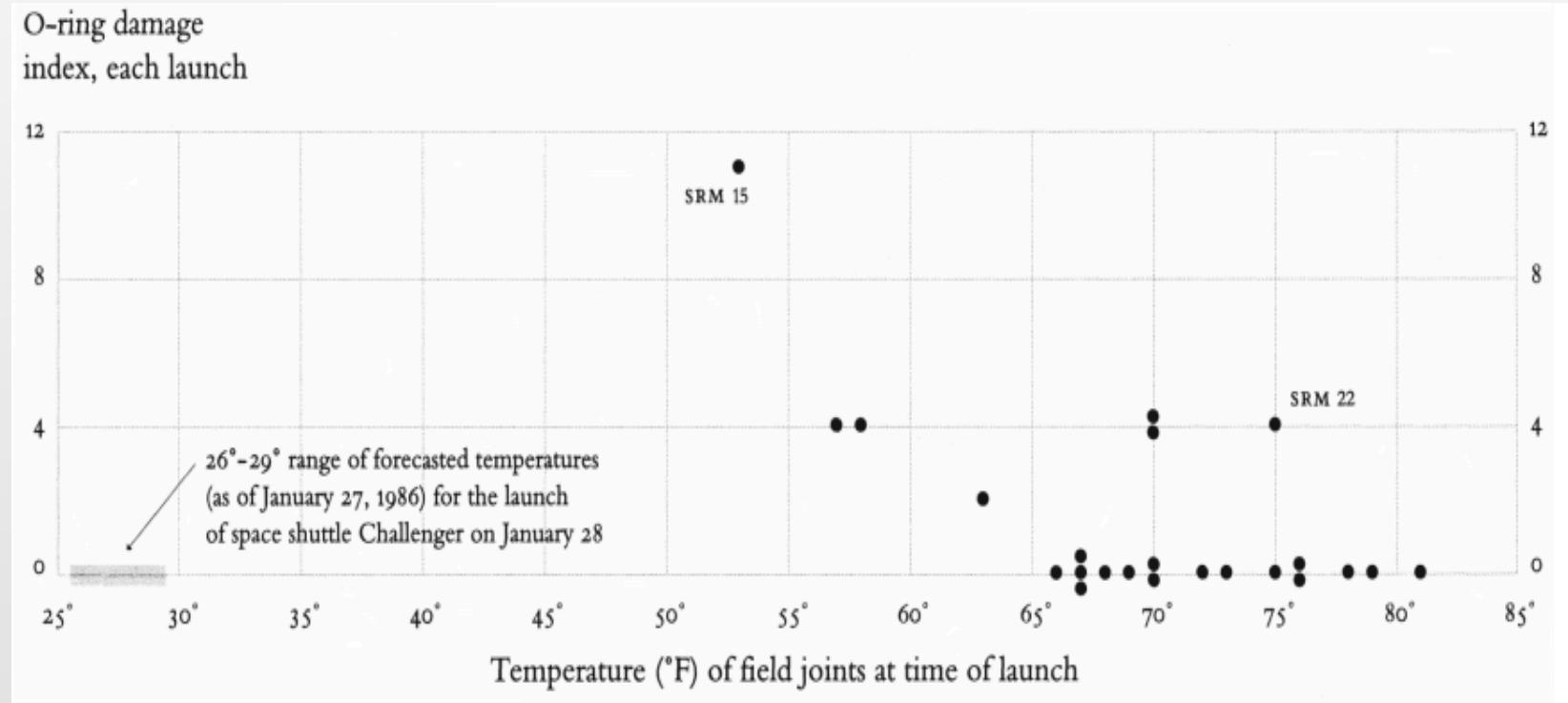
[http://www.aerospaceweb.org/question/investigations/
q0122.shtml](http://www.aerospaceweb.org/question/investigations/q0122.shtml)

Importance of Data Visualization



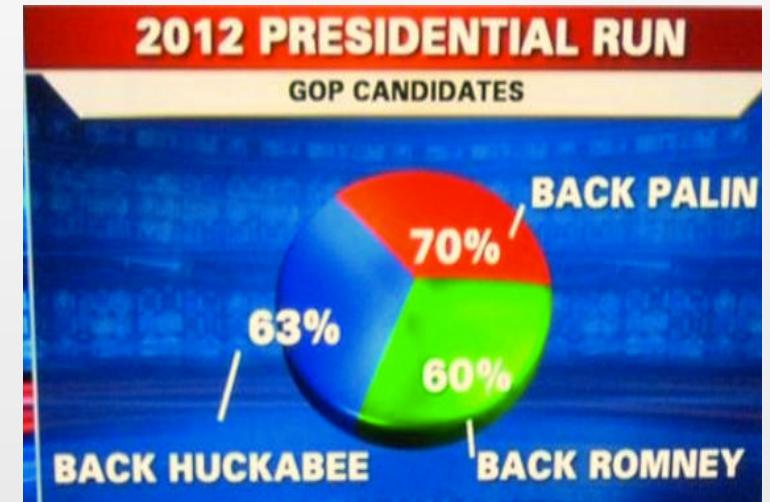
From Tufte's *Visual Explanations*

Importance of Data Visualization



From Tufte's *Visual Explanations*

Other Bad Visualizations



<http://blog.visualization-can-do-more-harm-than-good/>
<http://www.nbcchicago.com/news/local/FOX-News-Chart-Fails-Math-73711092.html>

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

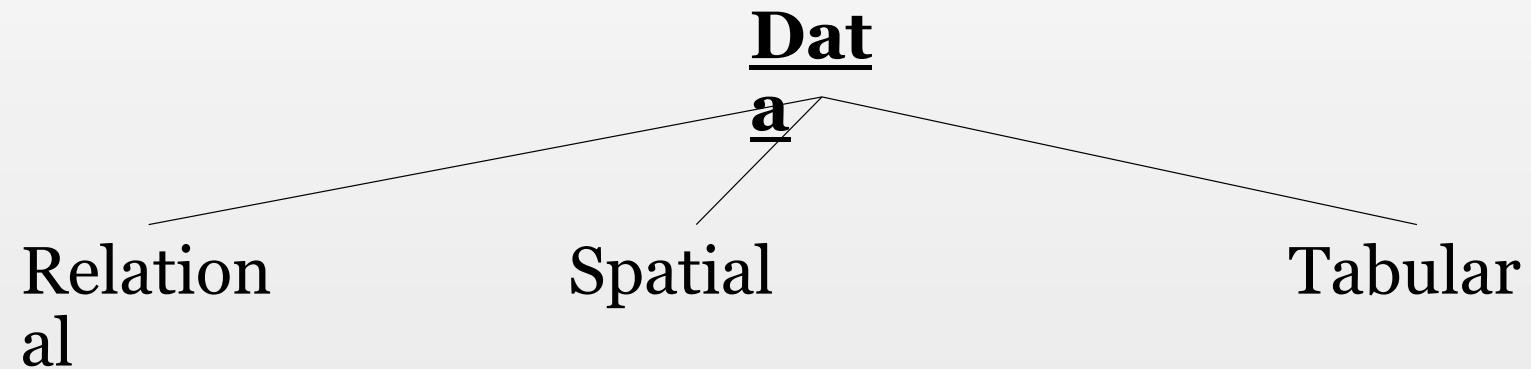
Ways to make visualization **clear, memorable** and **insightful**:

Match display to data and purpose

Maximize information-to-ink ratio

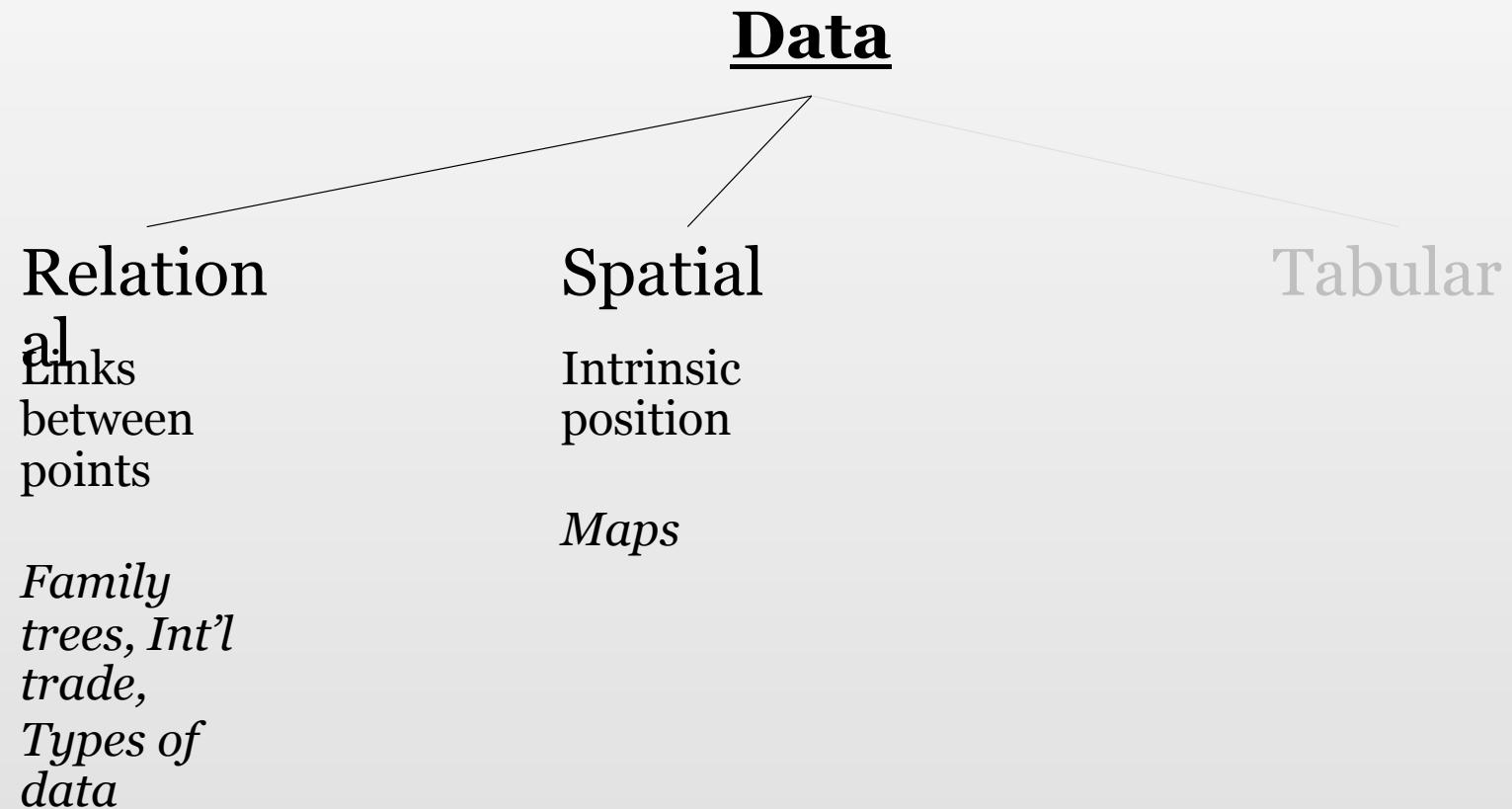
Principal of small multiples

Match Display to Data and Purpose

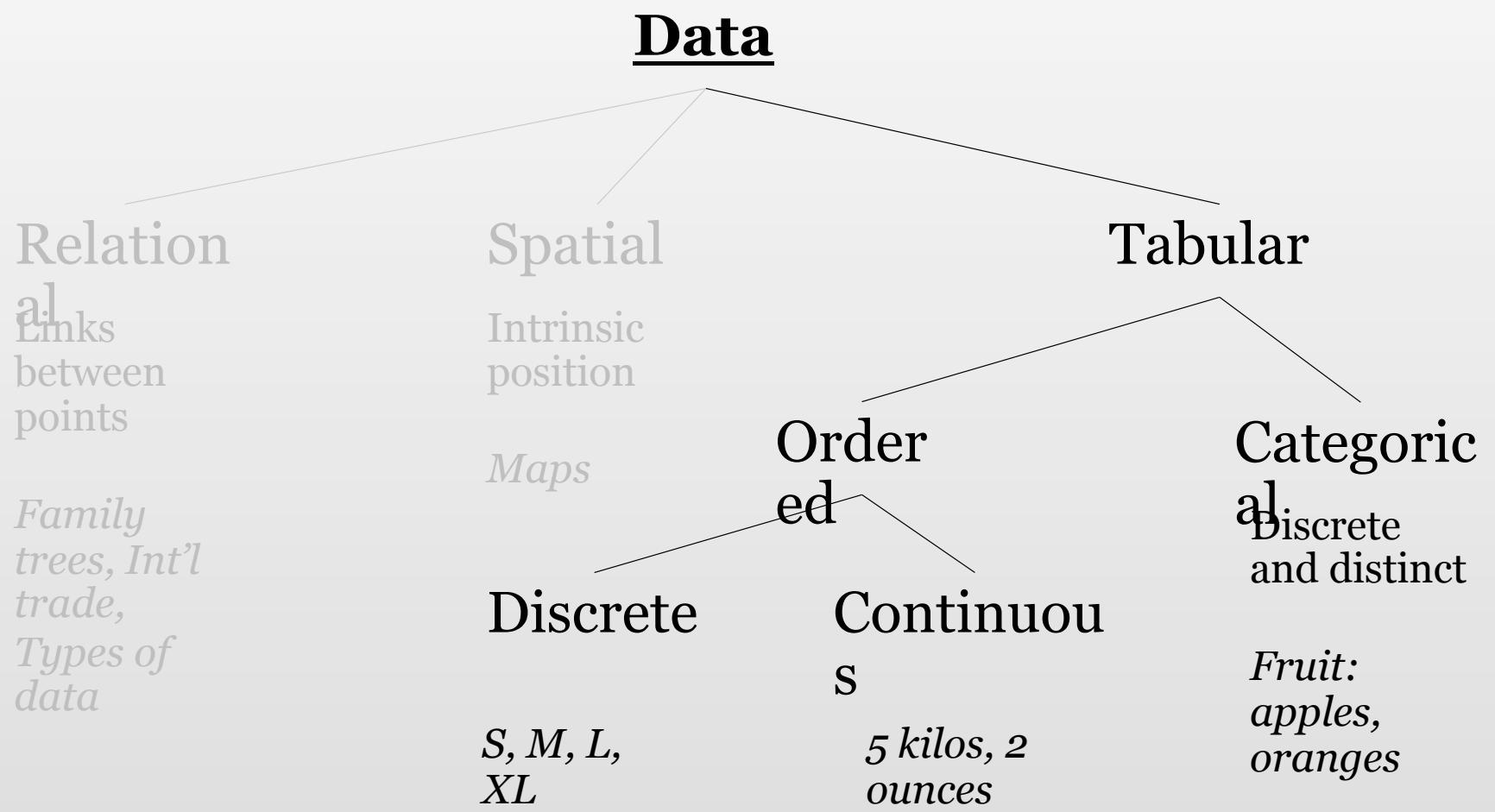


This and the next few slides paraphrase Tamara Munzner <http://www.cs.ubc.ca/~tmm/talks/vizbi11/>

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
Maps

Ordered
discrete/
cont.

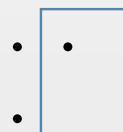
Categorical



Match Display to Data and Purpose

Relational

Contain



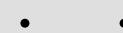
Connect



Similarity



Proximity



Spatial

Maps



Ordered
discrete/
cont.

Categoric
al

Match Display to Data and Purpose

Relational

Contain
• •

Connect
• •

Similarity
• 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
• o o

Proximity
• •

Spatial

Maps



Ordered
discrete/
cont.

Position
| . .

Length
| - -

Area
■ ■

Shade
□ □

Saturation
□ □

etc...

Categorical

• •

Position
• •

• •

Color
• o o

• o o

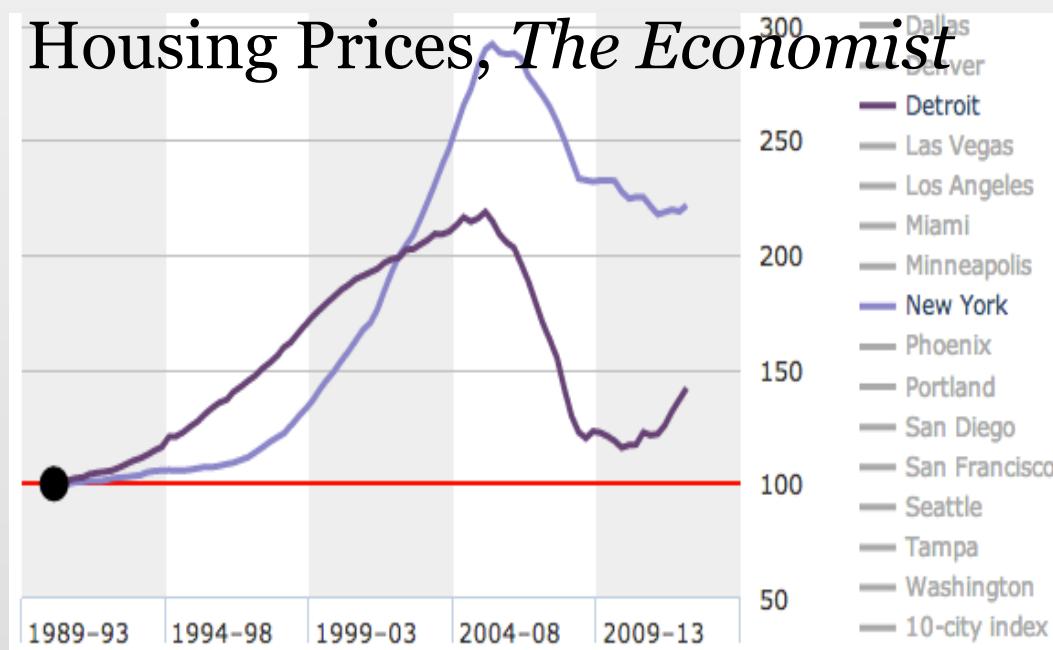
Shape
□ □

□ □

Patterns
□ □

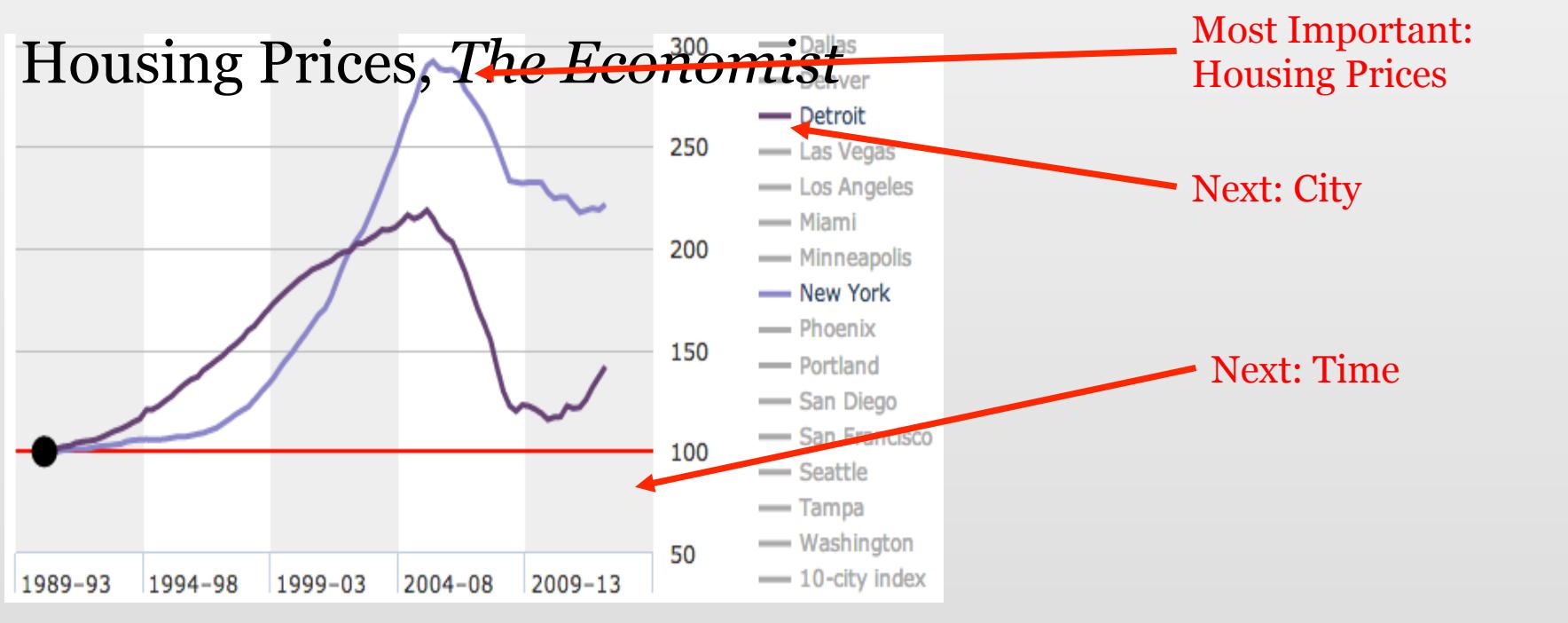
Match Dimensions to Data

Use highly ranked channels for key info to keep memorable



Match Dimensions to Data

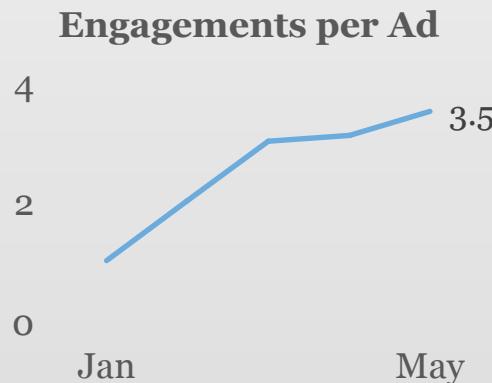
Use highly ranked channels for key info to keep memorable



Maximize Information-to-Ink Ratio

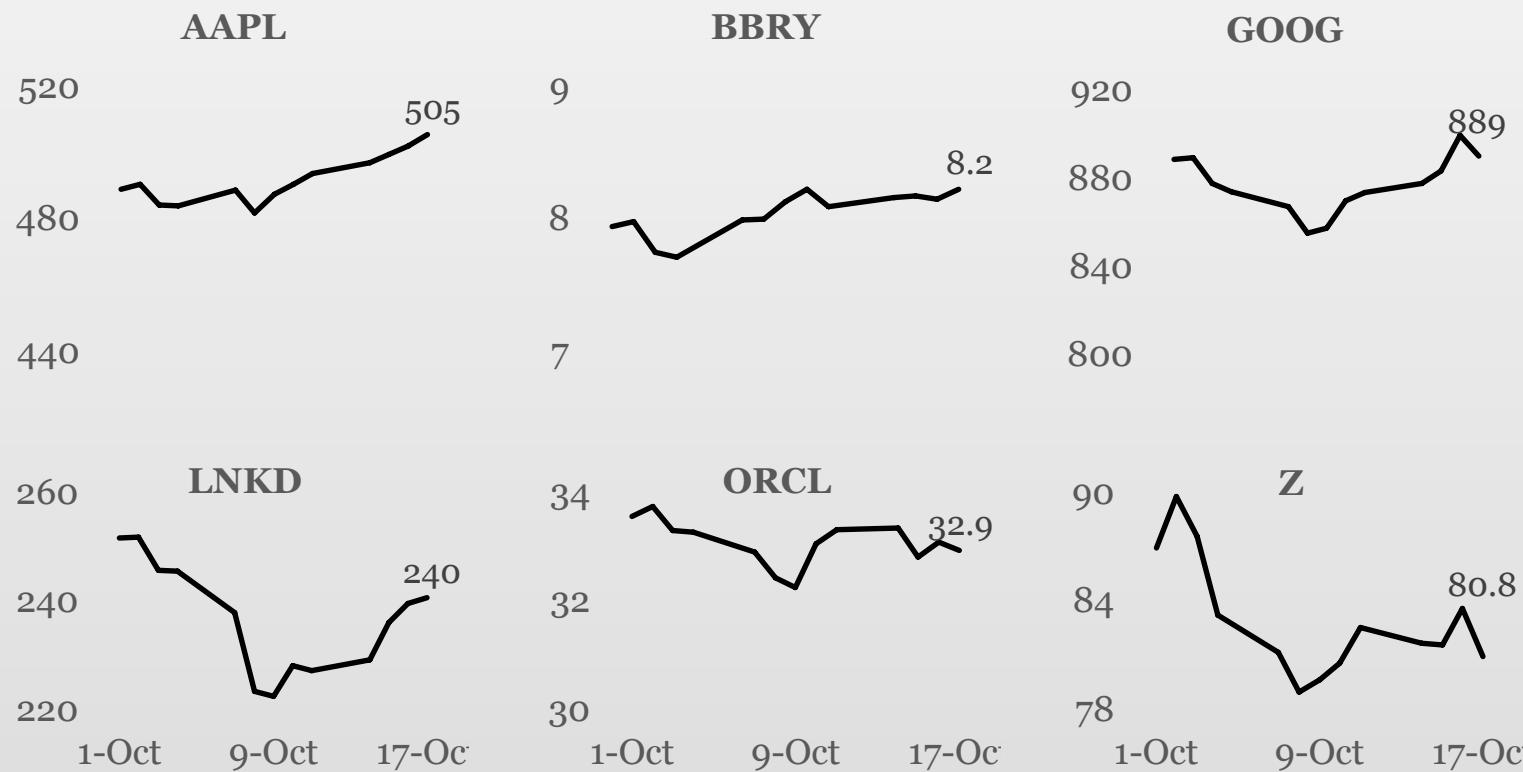
Achieve clarity by removing distractions & distracting elements

Taken to its logical endpoint: **Sparklines**

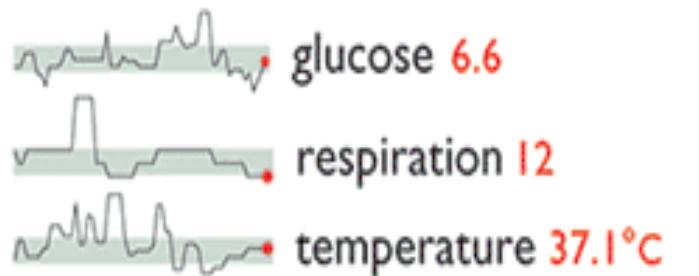


Principal of Small Multiples

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



Principal of Small Multiples In Practice



| | | | | | |
|----------|-----------------------|-------|--------|--------|-------|
| \$64,368 | Vanguard 500 Index | -2.0% | +12.2% | -11.7% | -0.8% |
| 62,510 | Fidelity Magellan | -2.1 | +11.3 | -12.9 | -0.2 |
| 50,329 | AmerA Invest Co Am | -1.2 | +09.4 | -03.9 | +4.0 |
| 47,355 | Amer AWA Mutual Inv | -1.5 | +09.9 | +00.8 | +3.0 |
| 40,500 | PIMCO InstlTot Return | -2.3 | +02.4 | +09.4 | +7.6 |
| 37,641 | Amer A Grow Fd Amer | -2.9 | +14.1 | -11.0 | +7.4 |
| 31,161 | Fidelity Contrafund | -1.0 | +10.7 | -06.5 | +3.0 |
| 28,296 | Fidelity Growth & Inc | -1.8 | +08.2 | -08.7 | -0.1 |
| 25,314 | AmerA Inc Fund Amer | -0.5 | +09.9 | +05.5 | +5.4 |
| 24,155 | Vanguard Instl Index | -2.0 | +12.3 | -11.6 | -0.7 |

Unique Visitors

227

Pages / Visit

1.39

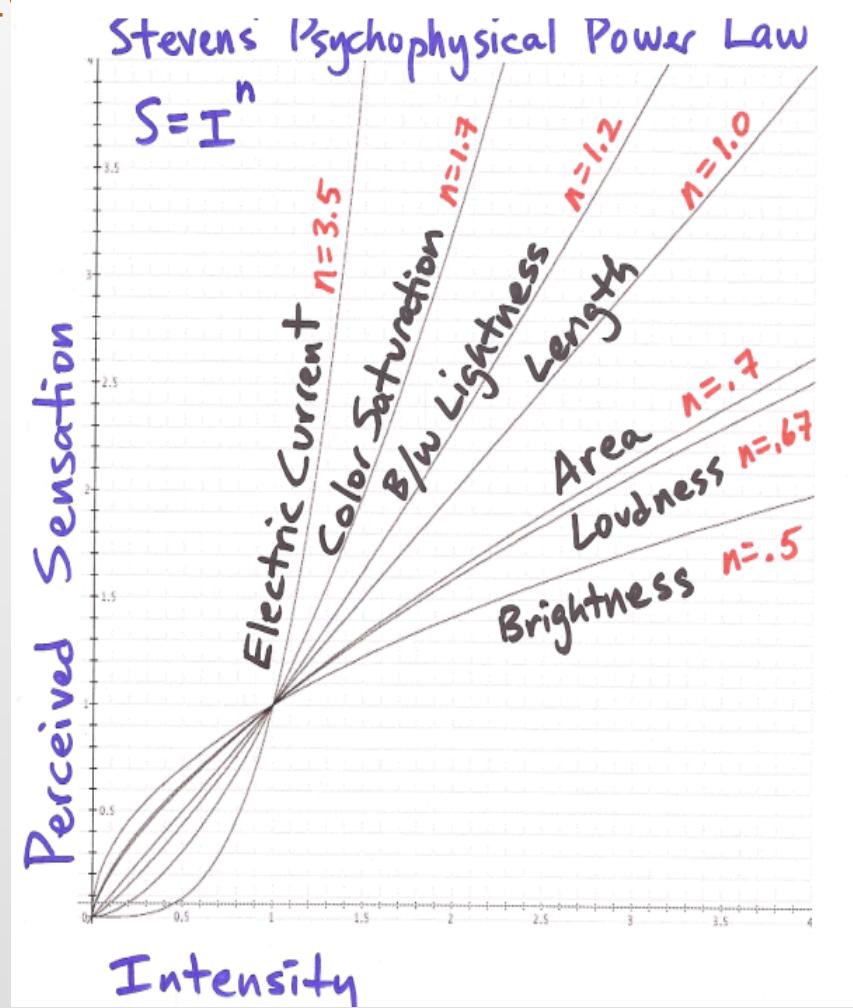
Bounce Rate

77.49%

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

<https://www.google.com>

Things To Keep in Mind - Dimensions

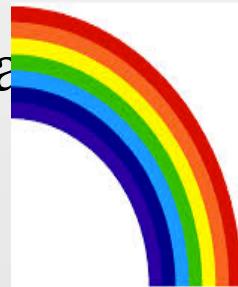


Things To Keep in Mind - Discrimination

~3 for Pie Charts



~7 for colors (not as a scale)



~4 overlapping lines

Just the Basics

Data Visualization: seems small & cute, actually a



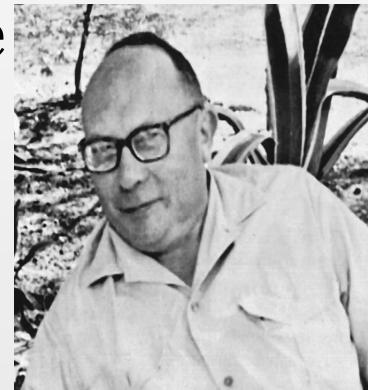
image via <http://www.cnn.com>

Further Reading on Visualization

Pioneers:

Edward Tufte

Jacques Bertin



Current Leaders:

Favorites:

[Bertin](#)

[Tufte](#)

[Munzner](#)

[Bostock](#)



Assignment 1 – Business Data, ggplot2



Assignment 1 – Business Data, ggplot2

Read ggplot2 docs

Submit code to me

Contact me with Questions

josh.laurito@gmail.com is best