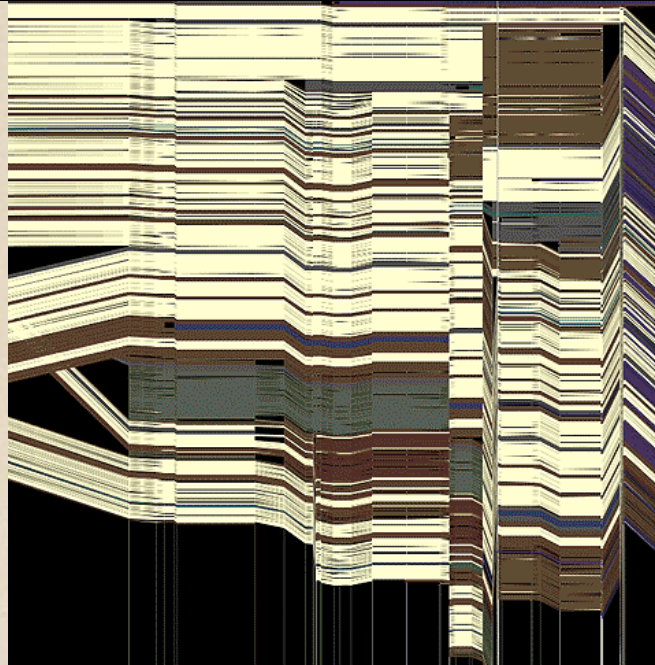
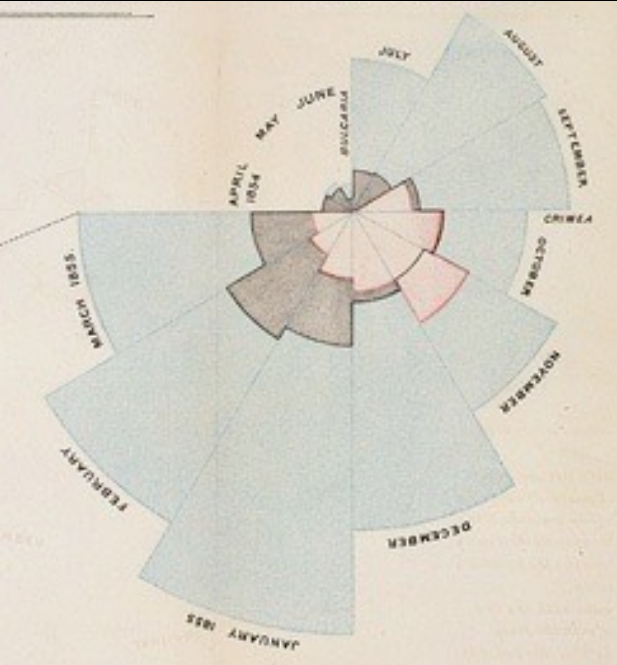


# CSE 512 - Data Visualization

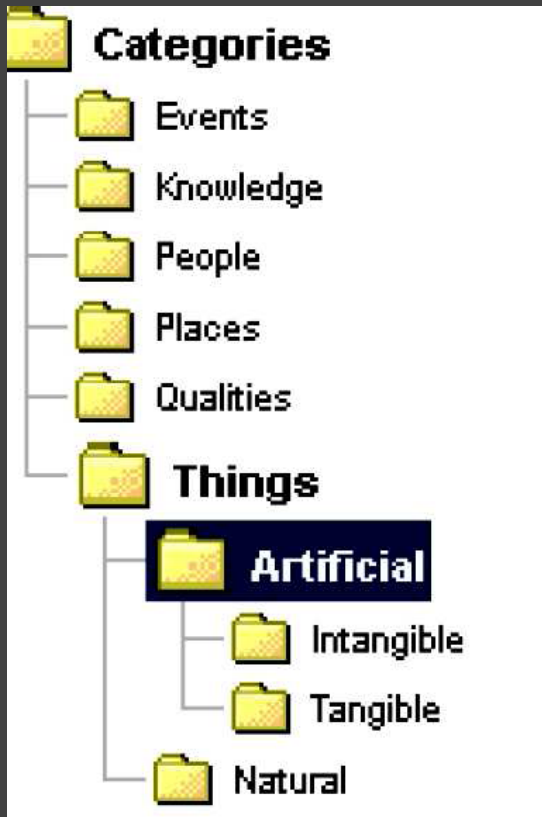
# Evaluation



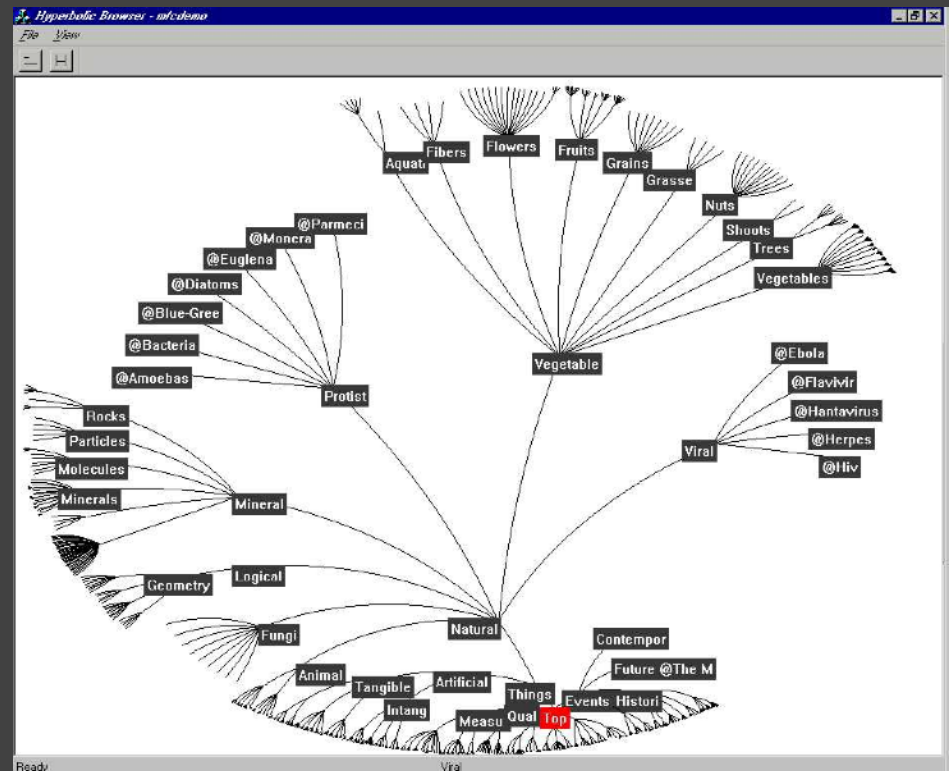
Jeffrey Heer University of Washington

How do we determine if a  
visualization is *effective*?

# Example: Tree Browsers



VS.



# Evaluation Methods

## Inspection or Principled Rationale

Apply design heuristics, perceptual principles

## Informal User Study

Have people use visualization, observe results

## Controlled Experiment

Choose appropriate tasks / users to compare

Choose metrics (time, error, **what else?**)

# Evaluation Methods

## Field Deployment or Case Studies

Observation and Interview

Document effects on work practices

## Theoretical Analysis

Algorithm time and space complexity

## Benchmarks

Performance (e.g., interactive frame rates)

Scalability to larger data sets

# Topics

Focus+Context (Trees, Spatial Navigation)

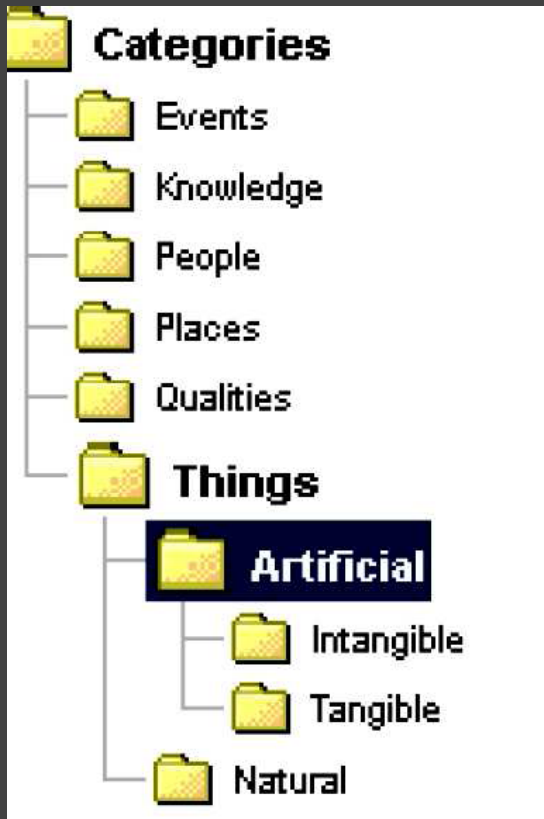
Data Density of Time Series

Perceptual Organization of Graphs

Discussion and Course Evaluation

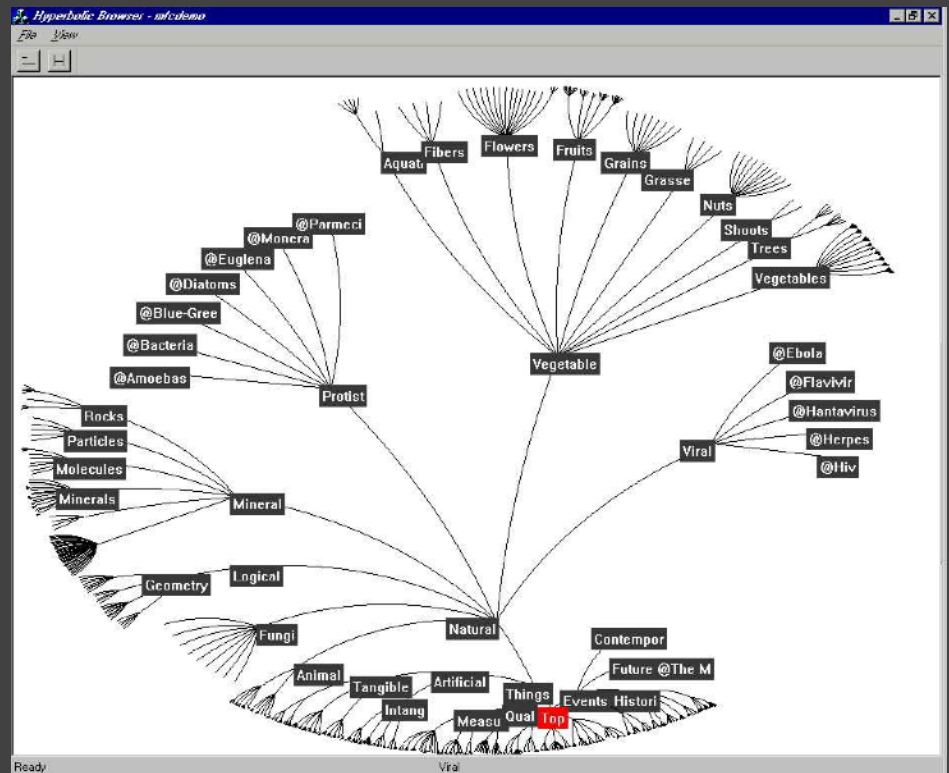
**Trees**

# The Great Browse-Off! [CHI 97]



Microsoft File Explorer

VS.



Xerox PARC Hyperbolic Tree



**Which visualization is better?**

# Which visualization is better?

Xerox PARC researchers ran eye-tracking studies to investigate... [Pirolli et al 00]

# Which visualization is better?

Xerox PARC researchers ran eye-tracking studies to investigate... [Pirolli et al 00]

Subjects performed both retrieval and comparison tasks of varying complexity.

# Which visualization is better?

Xerox PARC researchers ran eye-tracking studies to investigate... [Pirolli et al 00]

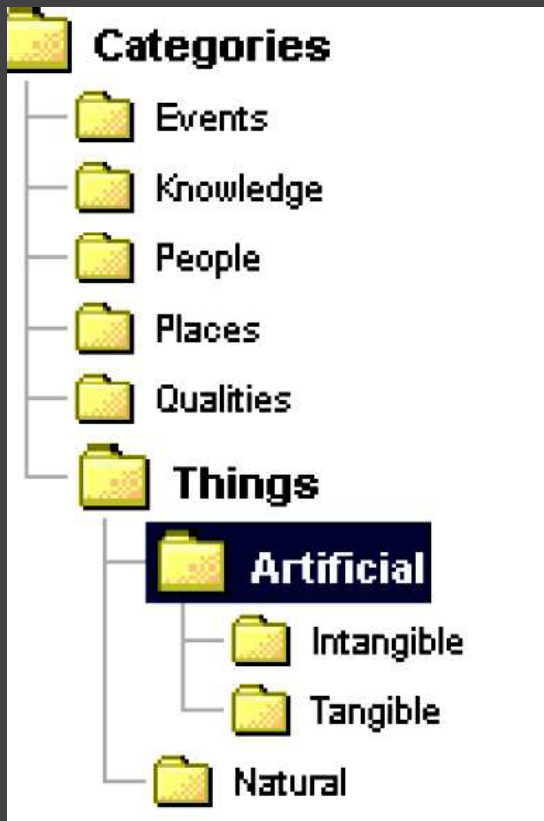
Subjects performed both retrieval and comparison tasks of varying complexity.

**No significant performance differences** were found across task conditions.

**How do users navigate the tree?**

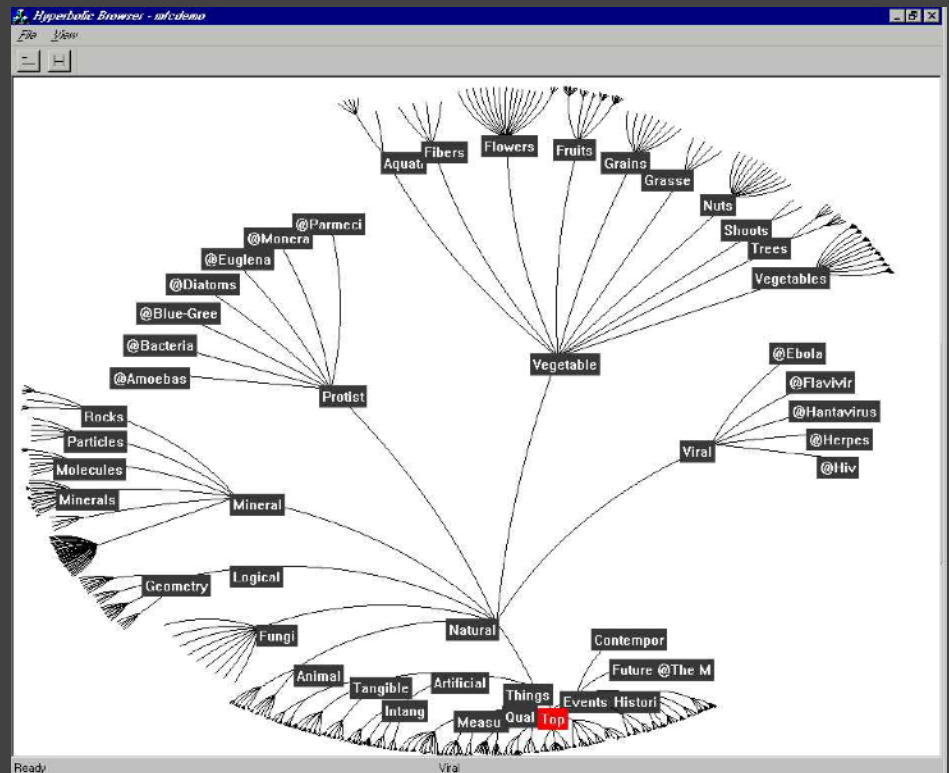
# How do users navigate the tree?

They read the labels!



Microsoft File Explorer

VS.



Xerox PARC Hyperbolic Tree

# How do users navigate the tree?

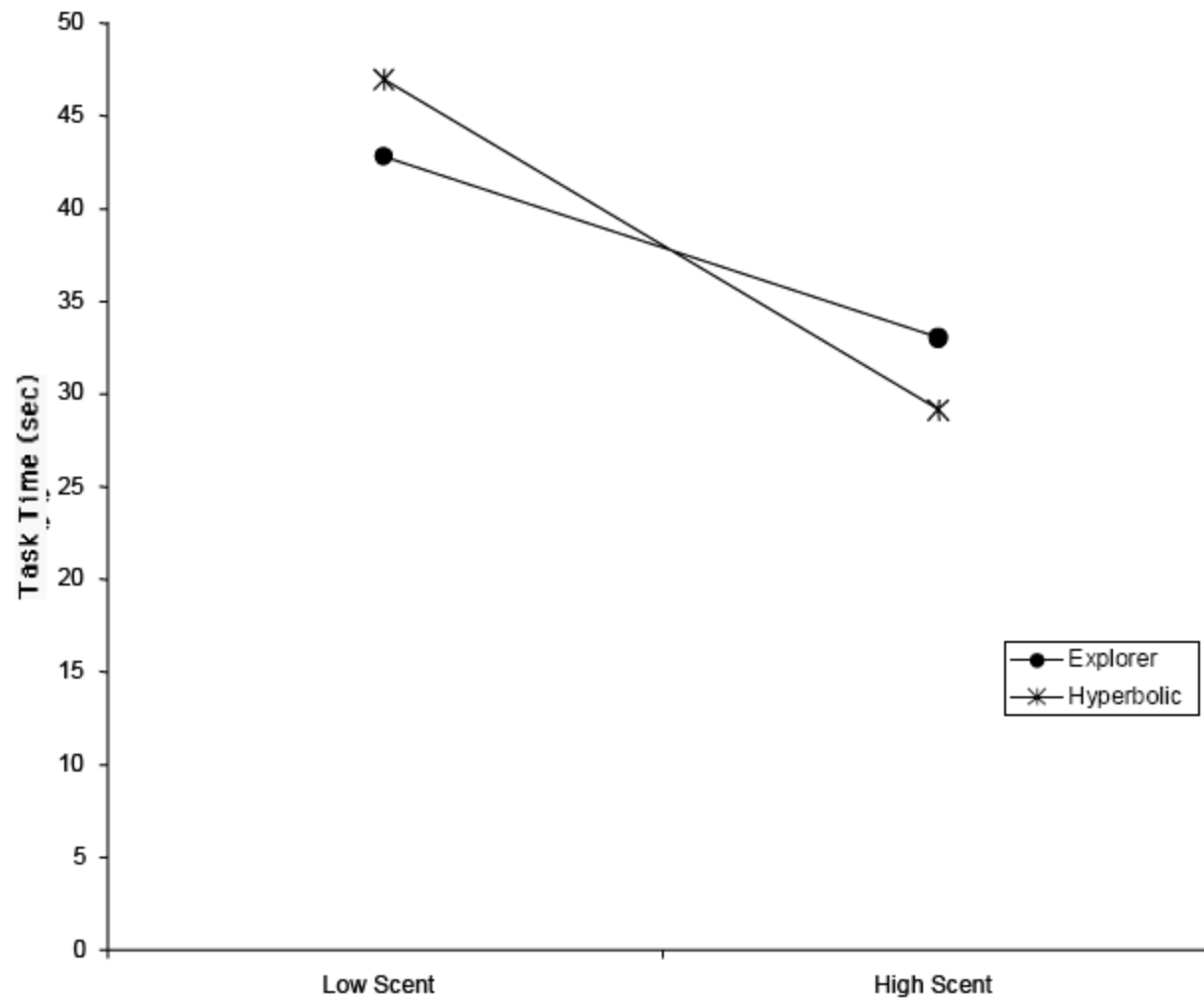
**Information Scent:** A user's (imperfect) perception of the value, cost, or access path of information sources obtained from proximal cues. [Pirolli & Card 99]

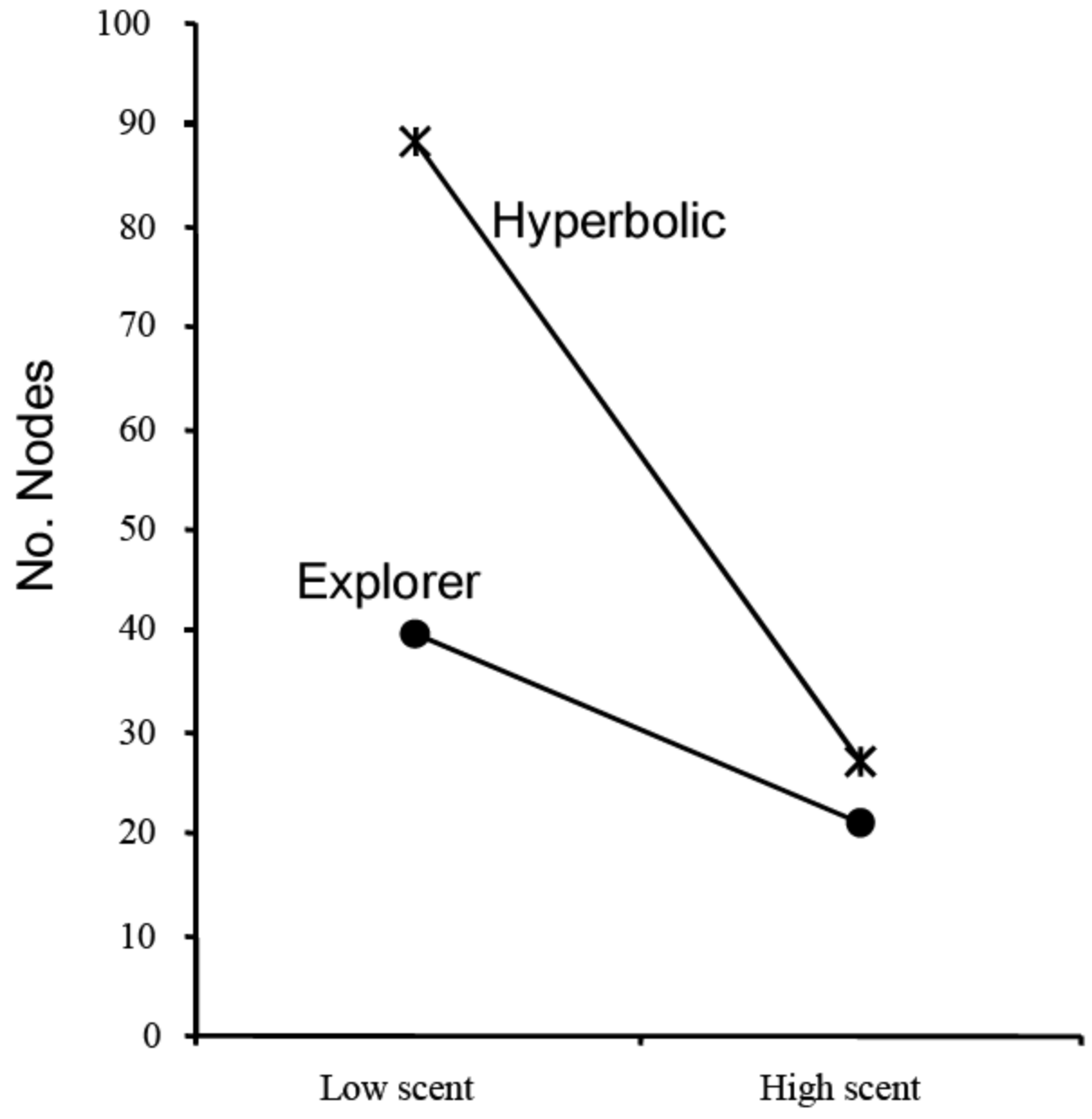
# How do users navigate the tree?

**Information Scent:** A user's (imperfect) perception of the value, cost, or access path of information sources obtained from proximal cues. [Pirolli & Card 99]

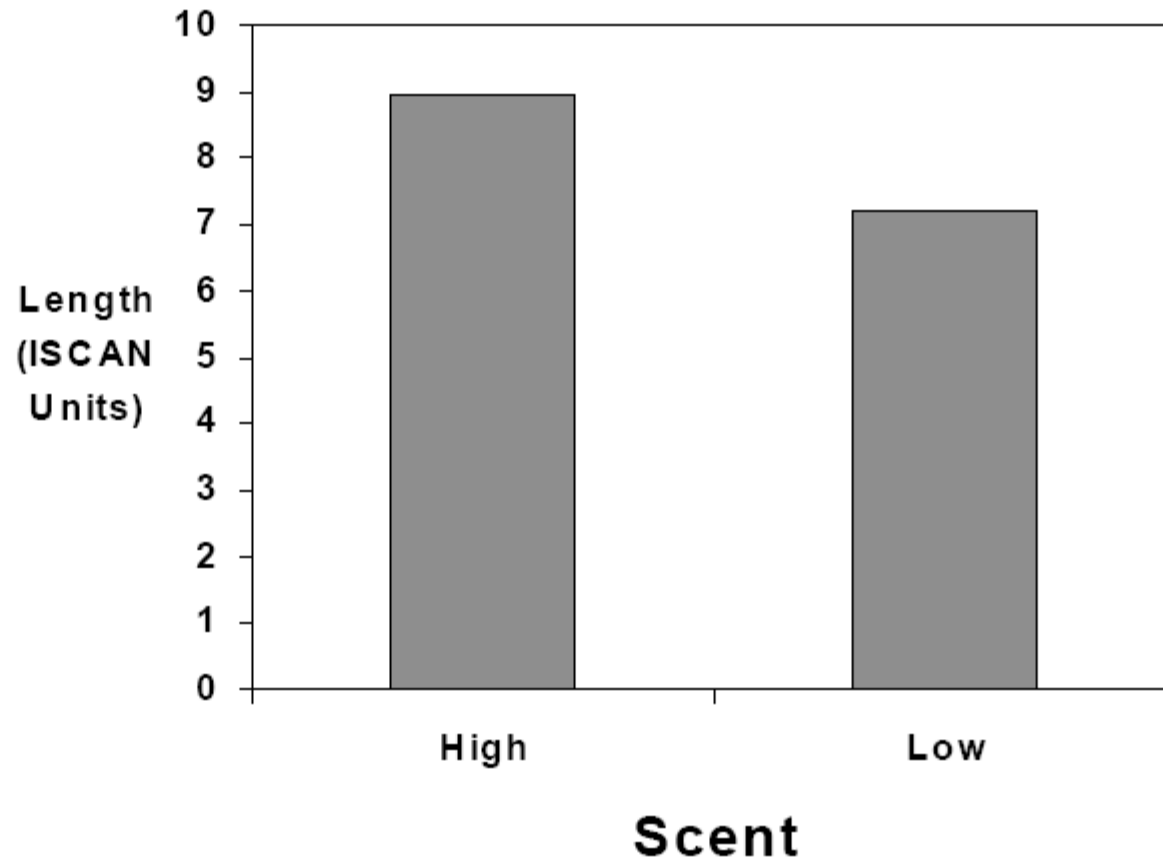
**Operationalize as:** the proportion of participants who correctly identified the location of the task answer from looking at upper branches in the tree.



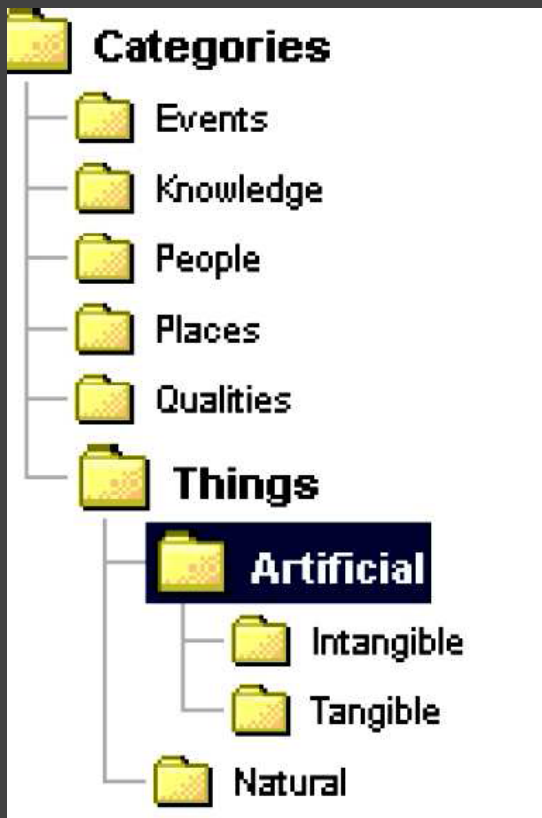




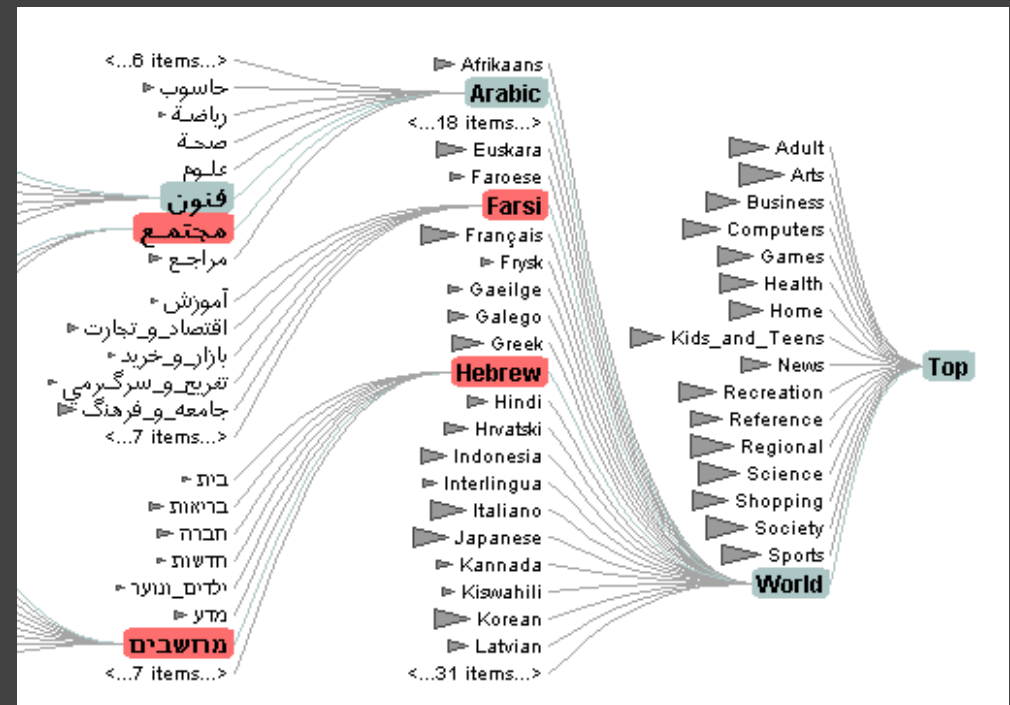
# Length of Eye Movements



# More Evaluations



vs.



# Evaluation of DOI Trees

## DOITree vs. Windows Explorer [Budiu, AVI 06]

Nodes visited (avg)	DOI:83	Exp:53	$p < .005$
Revisitation (avg)	DOI:6.6	Exp:8.2	$p < .005$
Divergence (avg)	DOI:4.6	Exp:3.9	$p < .001$

DOITree more forgiving to navigation errors

**BUT** no significant difference in task time

## DOITree vs. Google Directory [Pirolli, CHI 06]

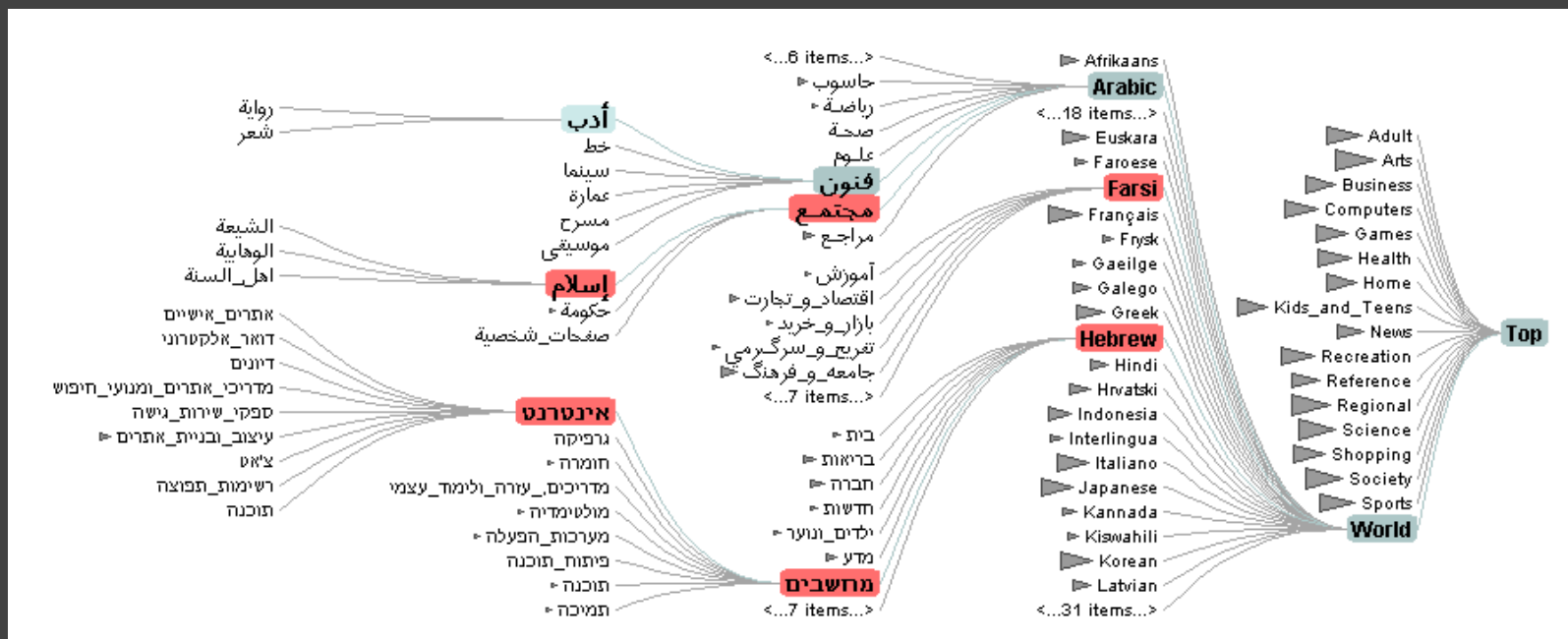
DOITree has superior task knowledge transfer

# Design Guidelines

# Design Guidelines

Support rapid visual scanning

Most people don't read in circles!



# Design Guidelines

People don't read in circles!

**Showing more is not always better**

**Distractors** can decrease task performance

Interaction with quality of **information scent**



# Design Guidelines

People don't read in circles!

Showing more is not always better

**Navigation cues critical to search**

**Informative labels** or landmarks needed

Poor **information scent** undermines search

# Lessons Learned

Both **task** and **data properties** (here, *information scent*) may interact with the visualization type in unexpected ways.

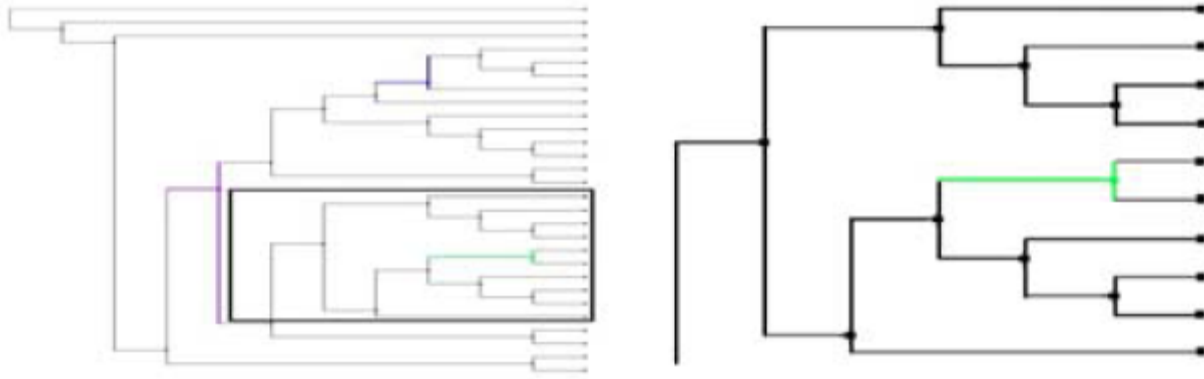
Equal **performance** in terms of accuracy or response time is **not the whole picture**.  
We often require more detailed study!

# Spatial Navigation

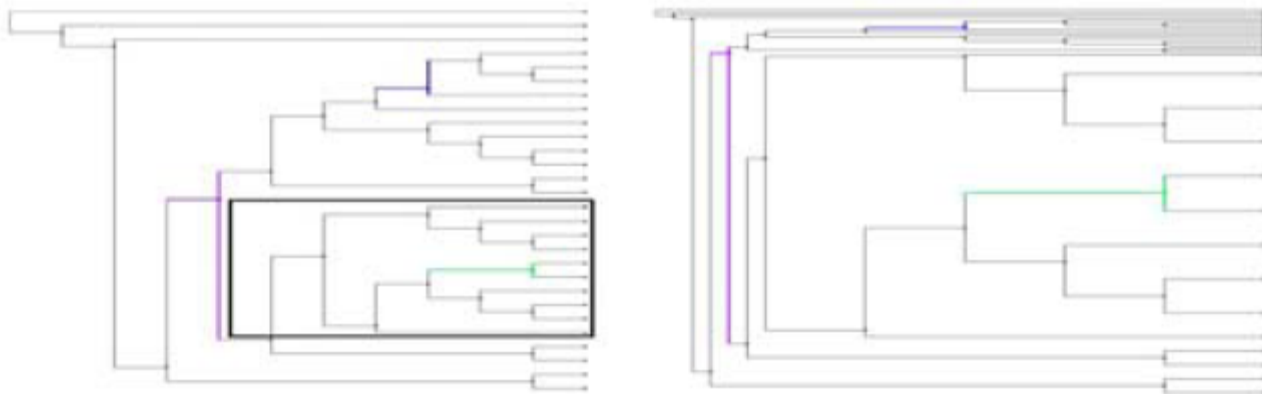
# **An Evaluation of Pan & Zoom and Rubber Sheet Navigation with and without an Overview**

Dmitry Nekrasovski, Adam Bodnar, Joanna McGrenere,  
François Guimbretière, Tamara Munzner

# Pan & Zoom vs. Rubber Sheet



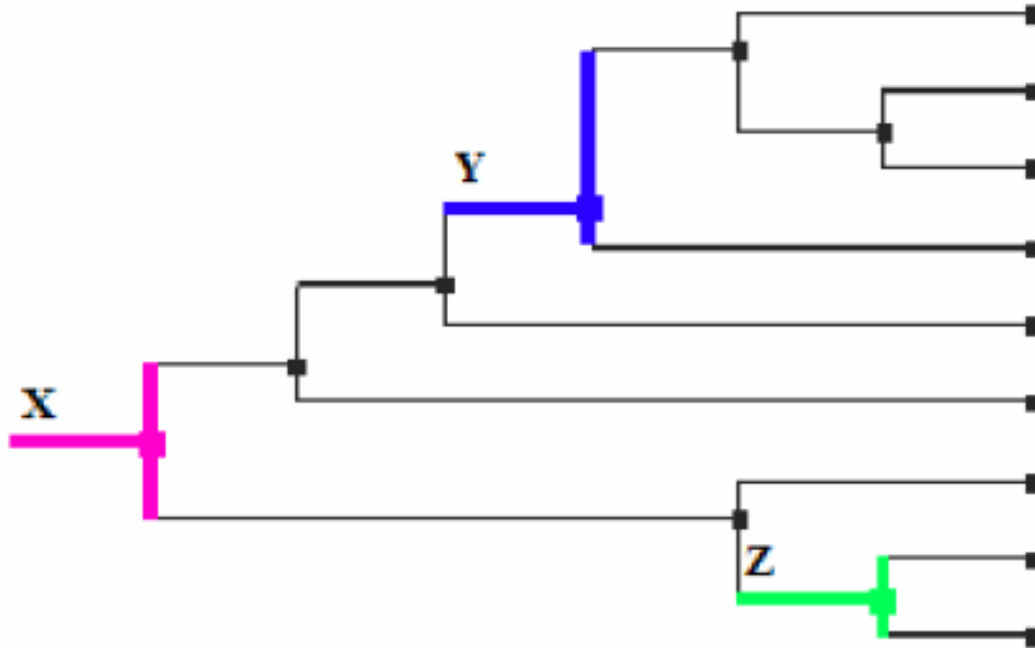
(i) PZN



(ii) RSN

# Experimental Task

Compare topological distance between nodes in a dendrogram.



# Experiment

Compare performance in 4 conditions:

1. Pan & Zoom (no overview)
2. Pan & Zoom (with overview)
3. Rubber Sheet (no overview)
4. Rubber Sheet (with overview)

40 subjects (24F/16M), between 18-39 years old.

Right-handed, normal vision.

Between-subjects design.

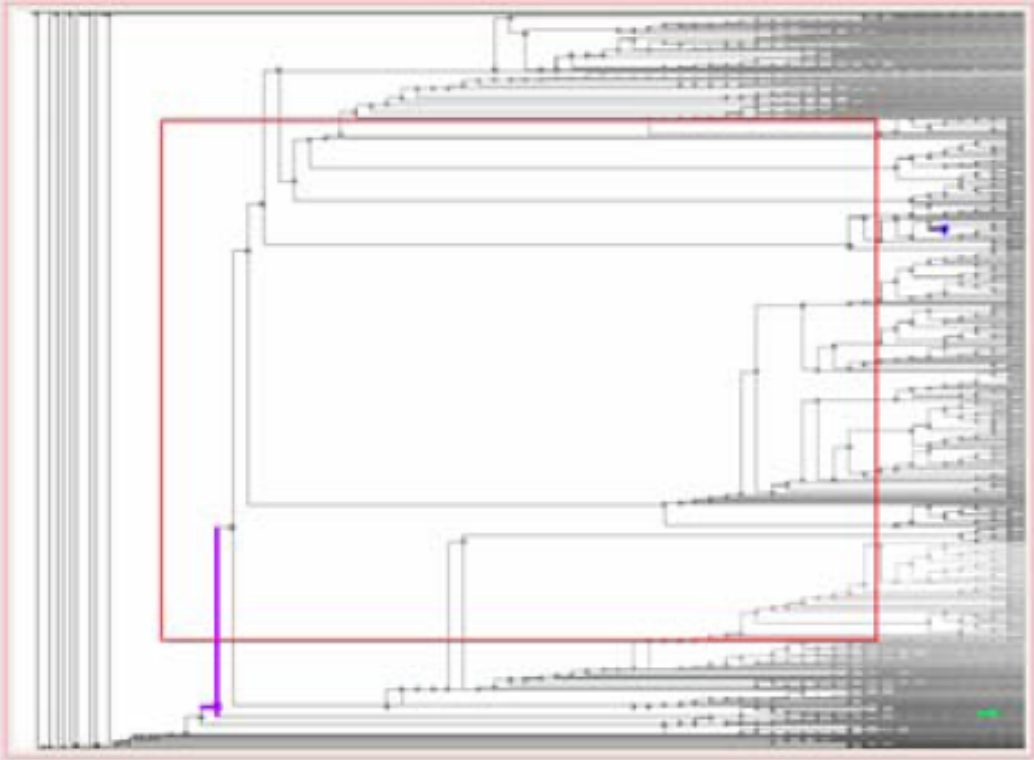
# 1. Rubber Sheet / No Overview

Evaluation13.gib, CL, level = 8  
File Edit Tools Help

Which node is the purple node closer to in terms of topological distance?

Blue  Green

Drag with LEFT mouse button to ZOOM IN  
Drag with RIGHT mouse button to PAN  
Press R to RESET the visualization  
Press ESCAPE to CLEAR the current mouse drag



The image shows a complex network graph visualization within a software window. The graph consists of numerous nodes and edges, forming a dense, interconnected structure. A red rectangular bounding box highlights a specific region of the graph, which is the focus of the question. Within this red box, a purple node is visible at the bottom left. The graph is rendered in a grayscale style with some color highlights (purple and green). The interface includes a menu bar at the top with 'File', 'Edit', 'Tools', and 'Help'. Below the menu bar, there is a title bar indicating the file name 'Evaluation13.gib, CL, level = 8'. The main area of the window contains the graph visualization. To the left of the graph, there are instructions for interacting with the visualization: 'Drag with LEFT mouse button to ZOOM IN', 'Drag with RIGHT mouse button to PAN', 'Press R to RESET the visualization', and 'Press ESCAPE to CLEAR the current mouse drag'. Above the graph, there is a question: 'Which node is the purple node closer to in terms of topological distance?' and two radio buttons labeled 'Blue' and 'Green', followed by a 'Submit' button.



# 2. Pan & Zoom / No Overview

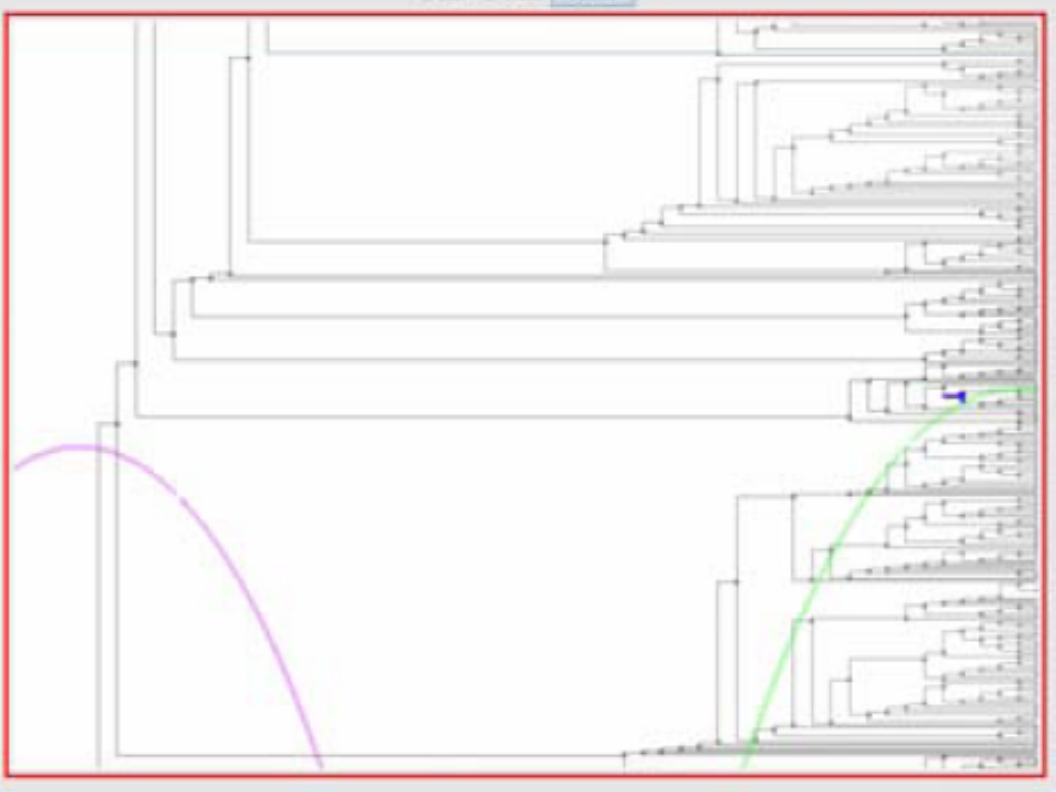
EvolutionT3.gd, C2, level = 8

File Edit View Help

Which node is the purple node closer to in terms of topological distance?

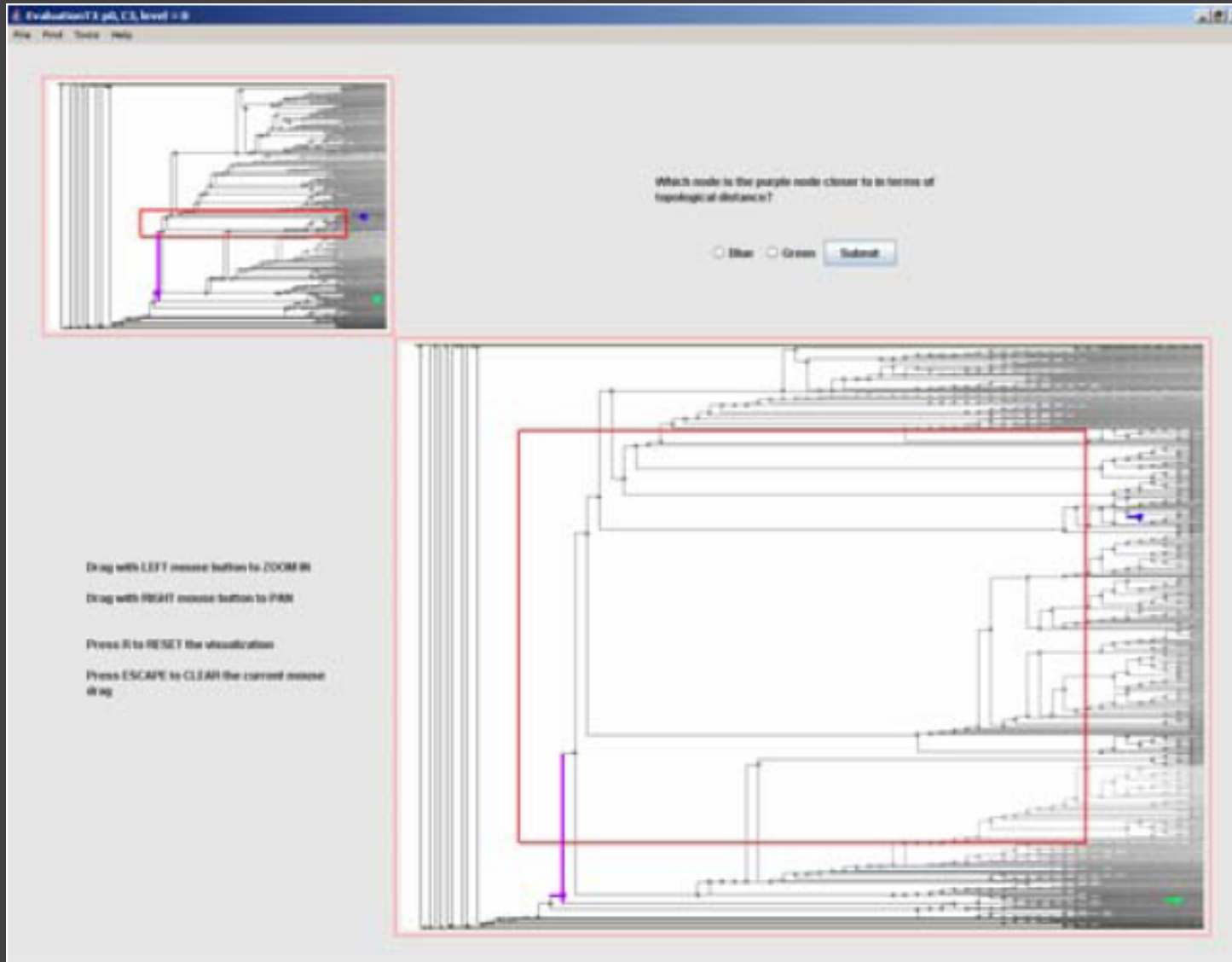
Blue  Green

Drag with LEFT mouse button to ZOOM IN  
Drag with MIDDLE mouse button to ZOOM OUT  
Drag with RIGHT mouse button to PAN  
Press R to RESET the visualization  
Press ESCAPE to CLEAR the current mouse drag



# 3. Rubber Sheet / Overview

Evaluation 1.1 pt, C3, level = 3  
File Edit View Help



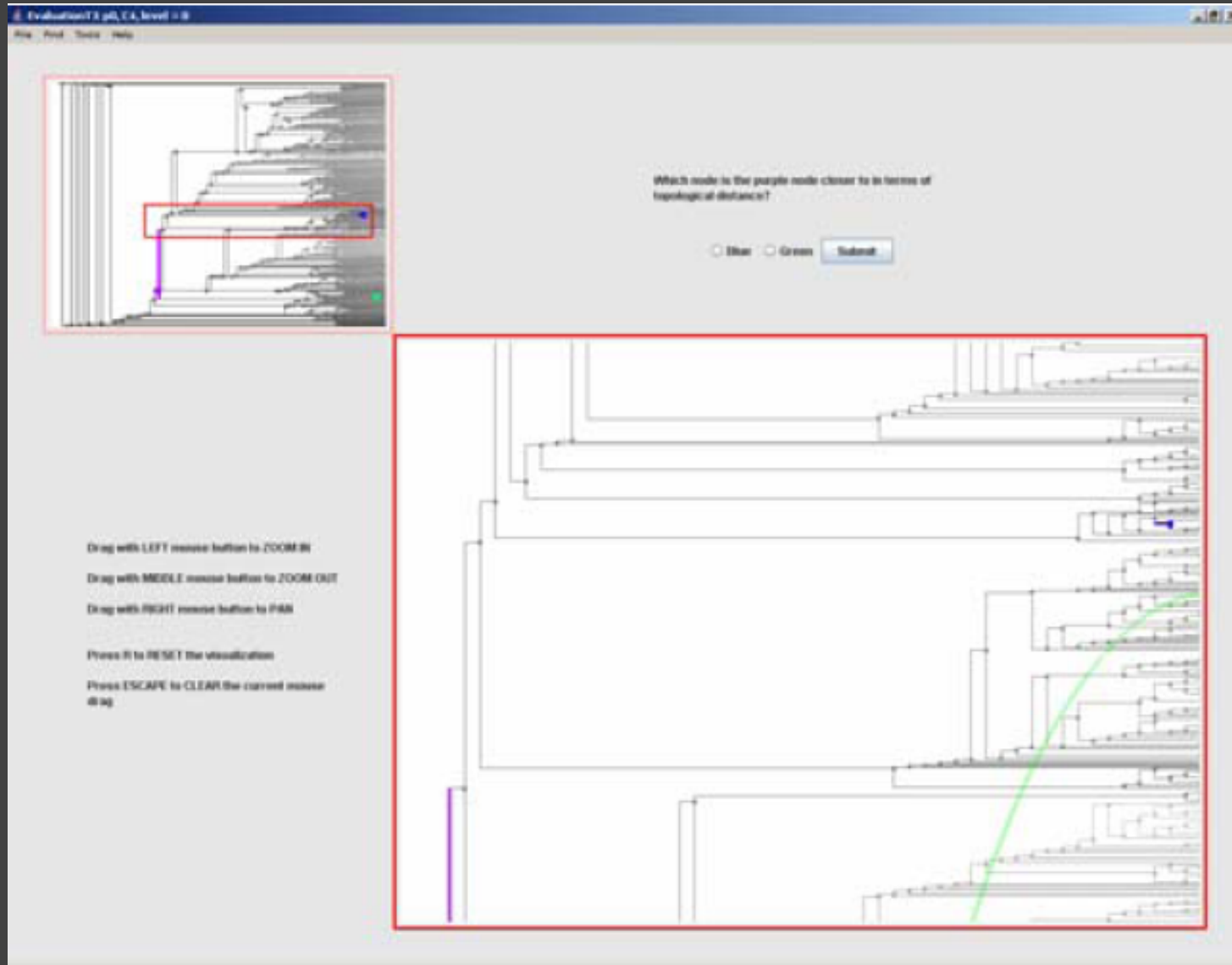
Which node is the purple node closer to in terms of topology or distance?

Blue  Green

Drag with LEFT mouse button to ZOOM IN  
Drag with RIGHT mouse button to PAN  
Press R to RESET the visualization  
Press ESCAPE to CLEAR the current mouse drag

# 4. Pan & Zoom / Overview

Evolution 1.3 pt, CA, level = 9  
File Edit View Help



Which node is the purple node closer to in terms of topology or distance?

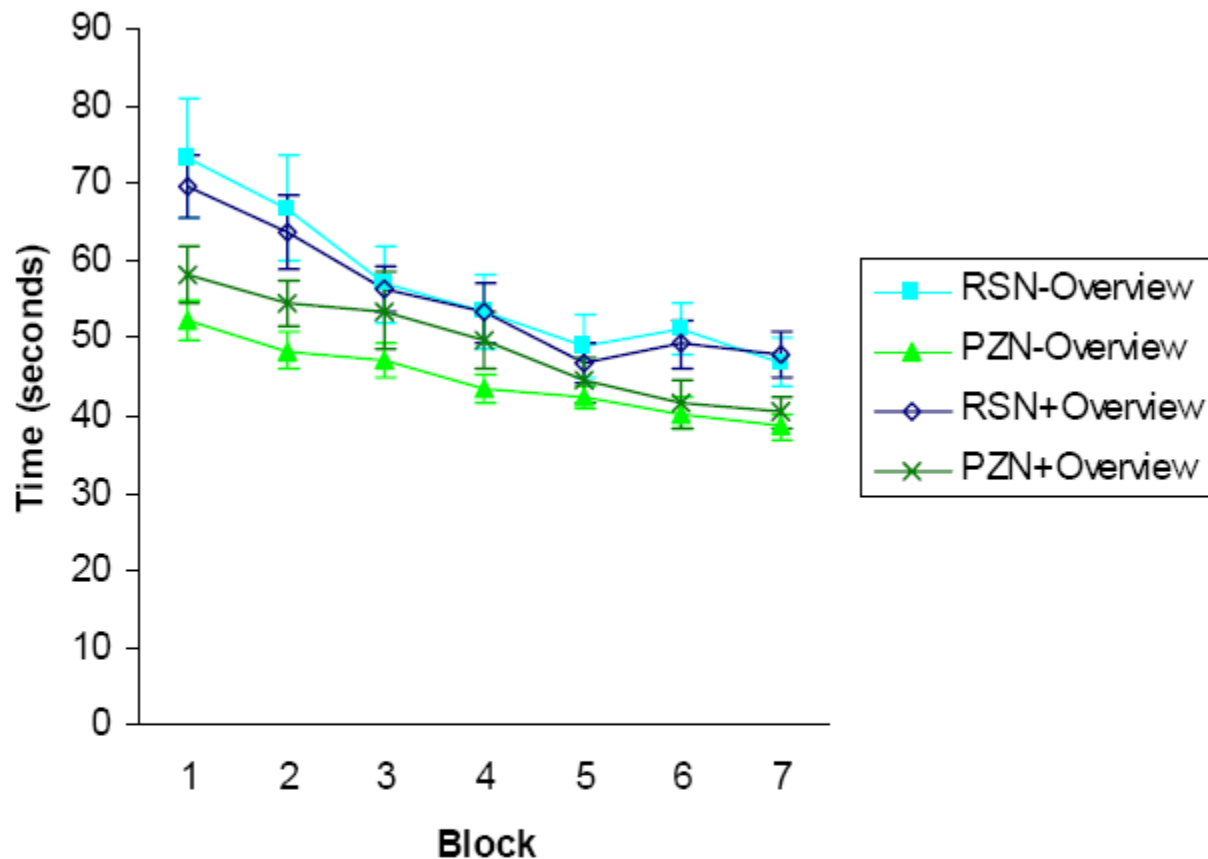
Blue  Green

Drag with LEFT mouse button to ZOOM IN  
Drag with MIDDLE mouse button to ZOOM OUT  
Drag with RIGHT mouse button to PAN  
Press F to RESET the visualization  
Press ESCAPE to CLEAR the current mouse drag

# Hypotheses

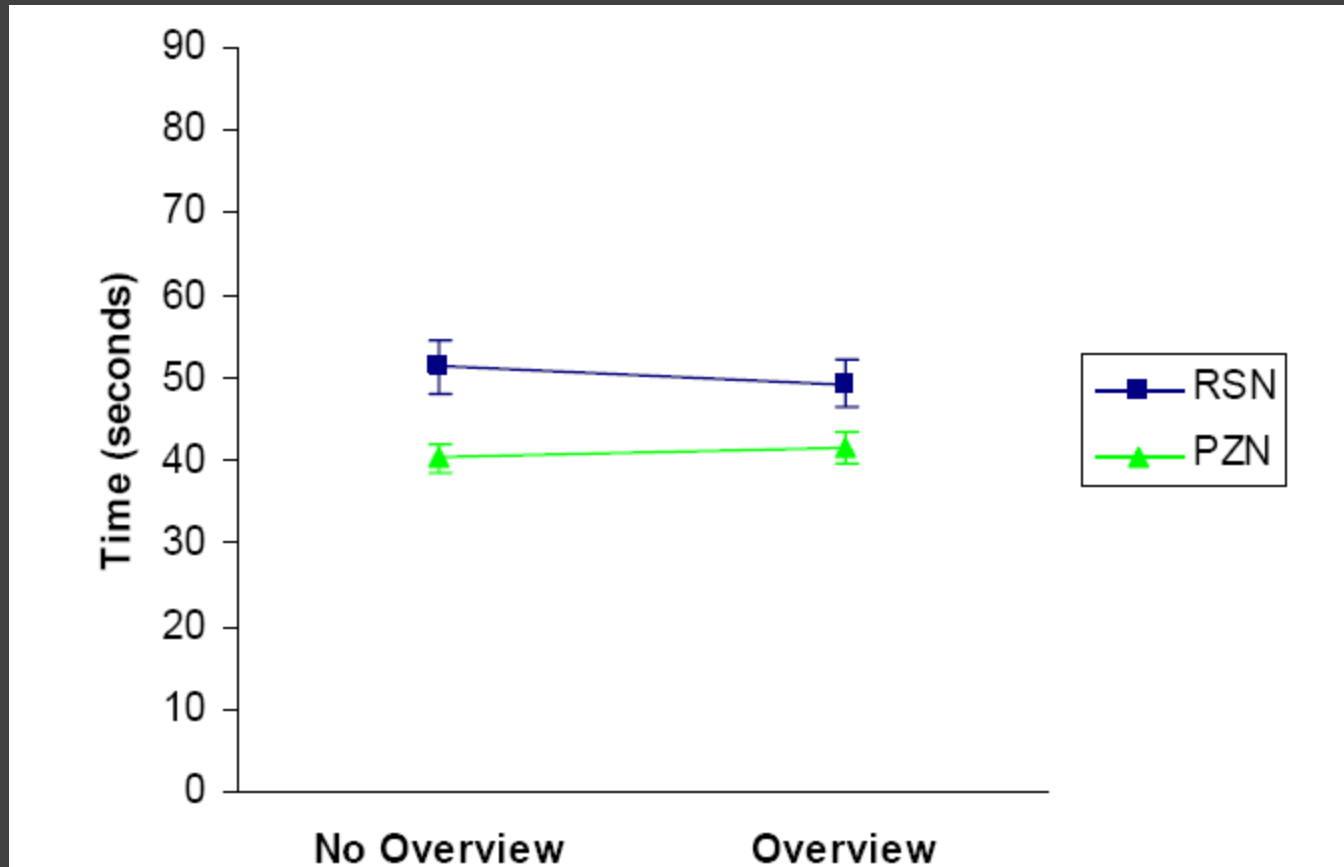
1. RSN interfaces perform better than PZN interfaces independently of the presence or absence of an overview.
2. For RSN, the presence of an overview does not result in better performance.
3. For PZN, the presence of an overview results in better performance.

# Results: H1 False



**Figure 7: Mean completion times per trial for each interface by block in seconds (N=40).**

# Results: H2 True, H3 False



**Figure 9: Block 7 mean per-trial completion times in seconds by navigation technique with and without an overview.**

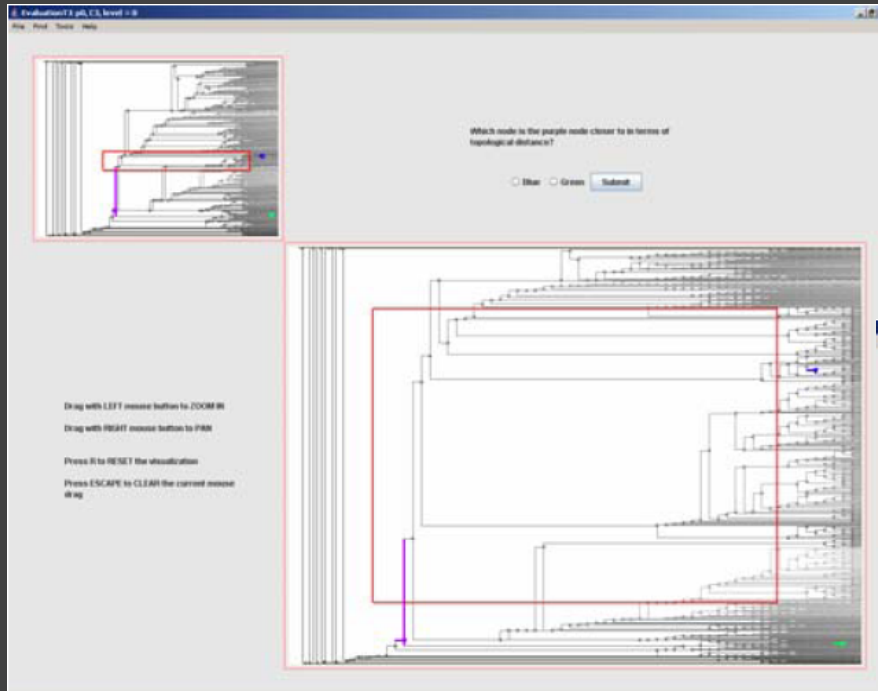
# Results

R1. Pan & Zoom had lower completion times, navigation actions, resets, and reported mental demand.

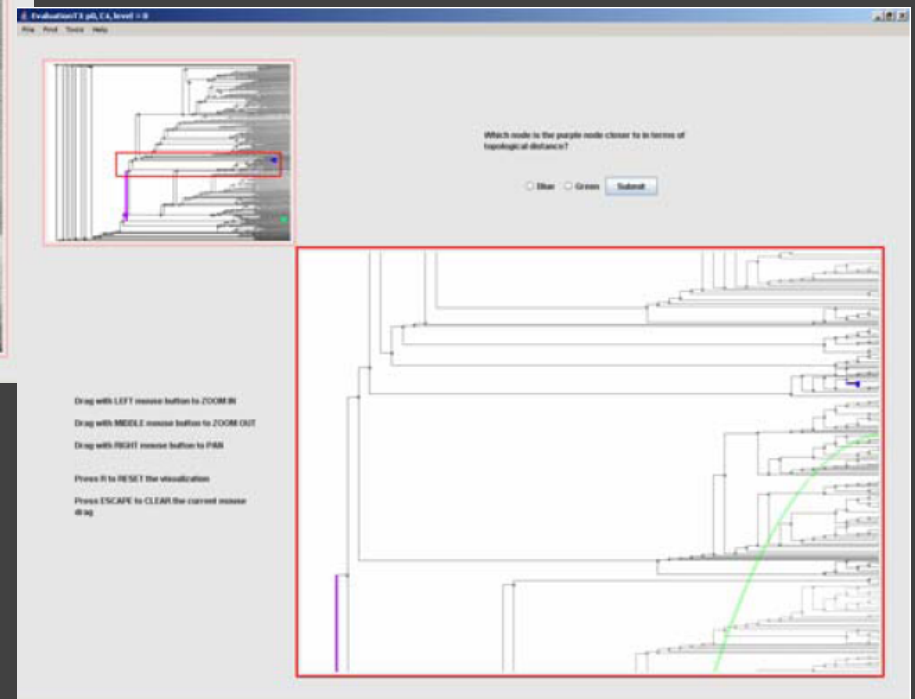
R2. Overview has no significant impact on rubber sheet navigation, though it was reported to reduce physical demand.

R3. Overview has no significant impact on pan & zoom navigation, though it was reported to reduce physical demand.

# Thoughts?

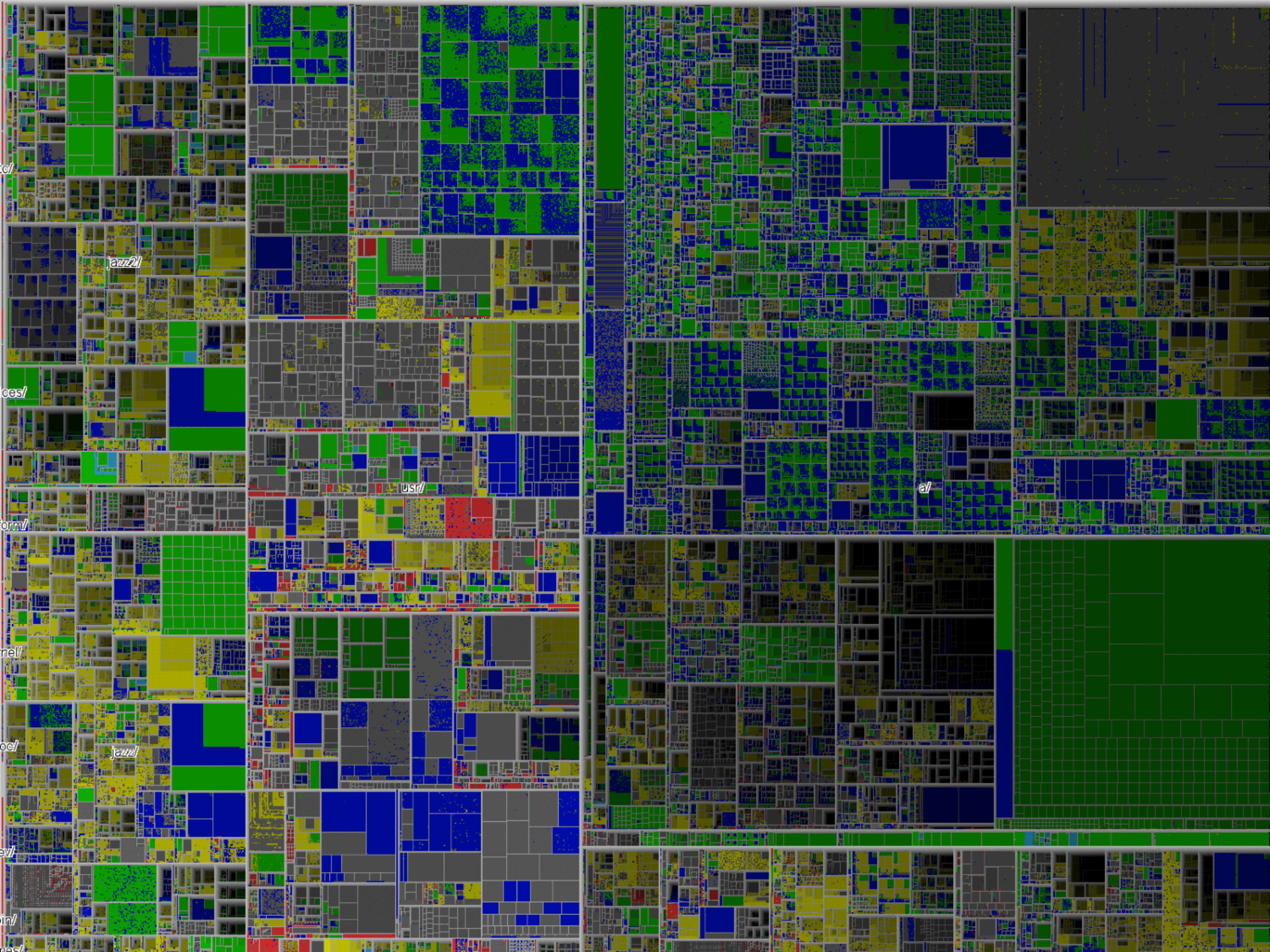


Does this generalize  
for overview displays?





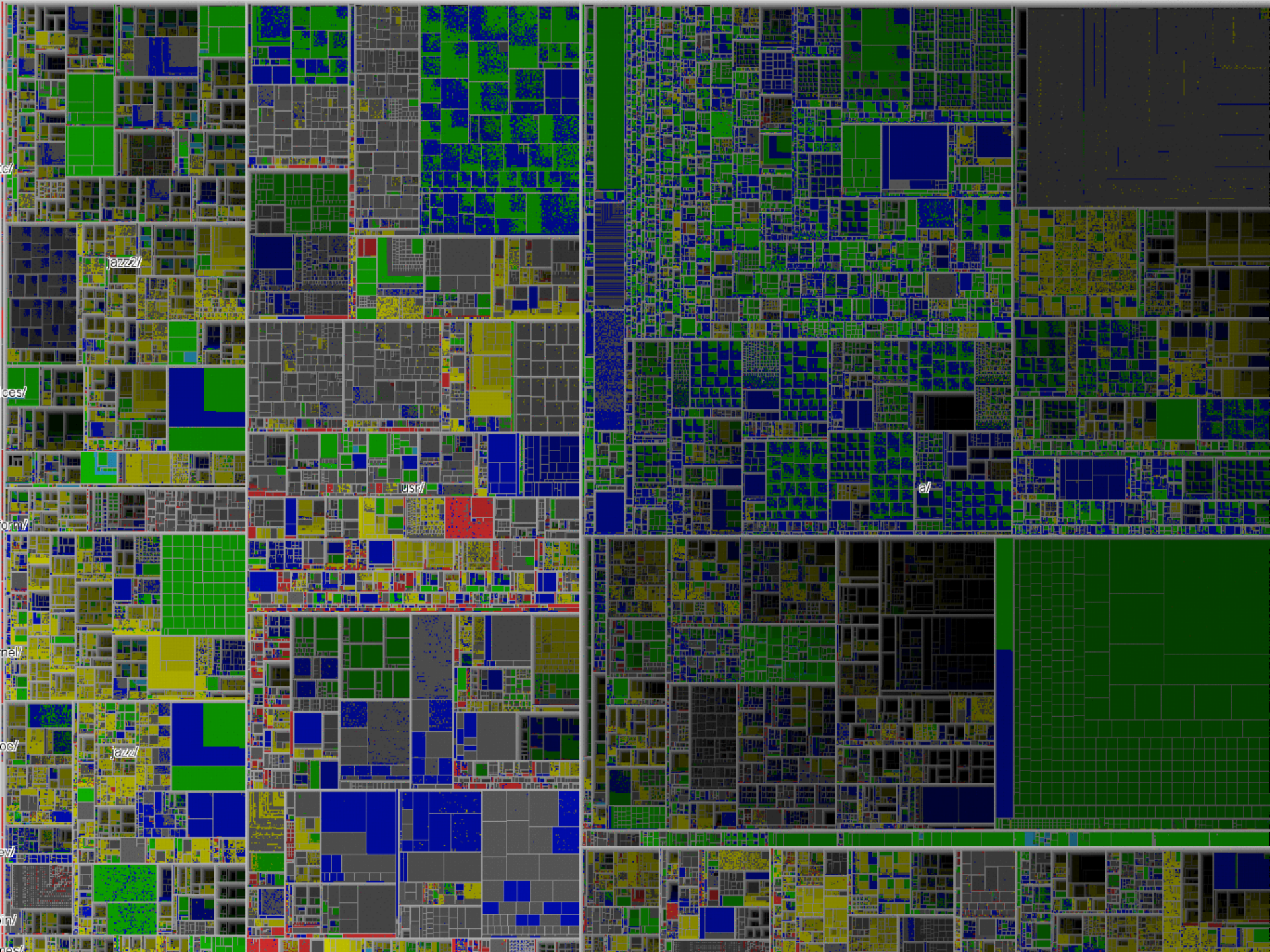
# Data Density

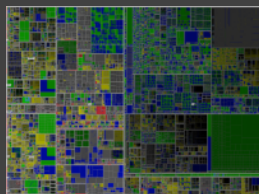


$$\text{Data Density} = \frac{(\# \text{ entries in data})}{(\text{area of graphic})}$$

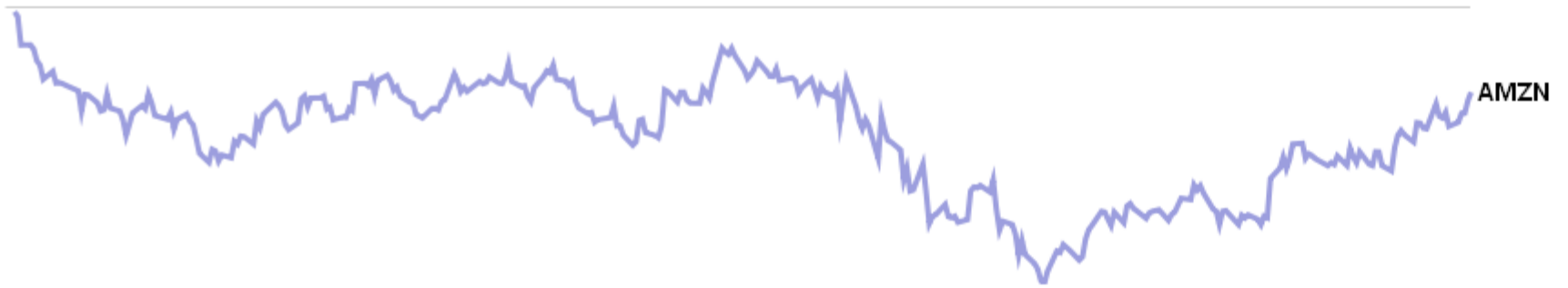
“Graphical excellence... gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space”

[Tufte 83]

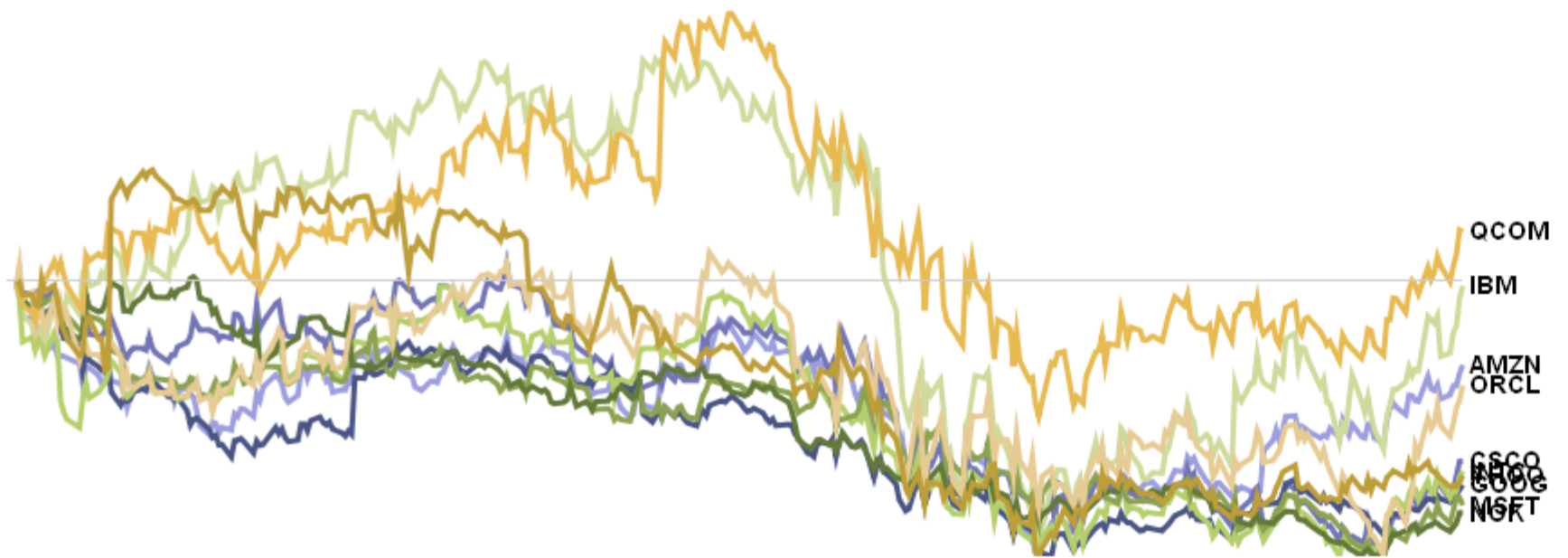




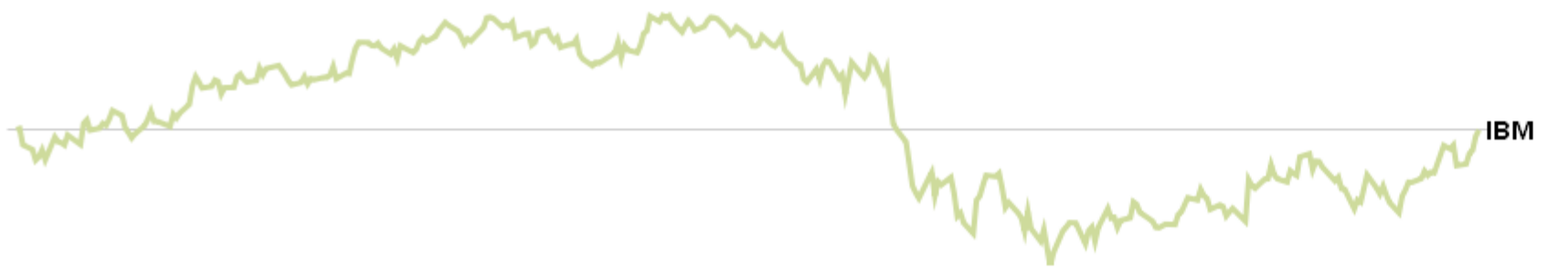
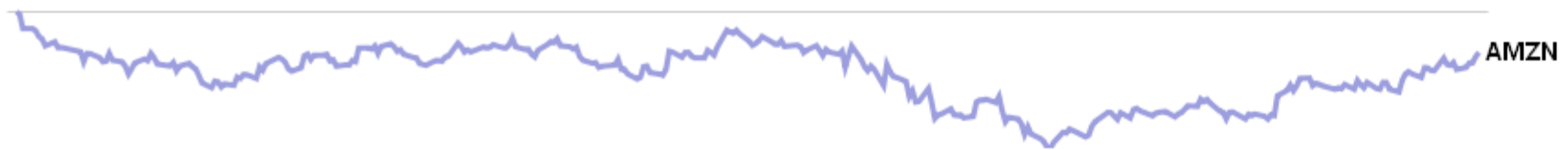
## Relative Technology Stock Performance: Jan 2008 - Present



## Relative Technology Stock Performance: Jan 2008 - Present



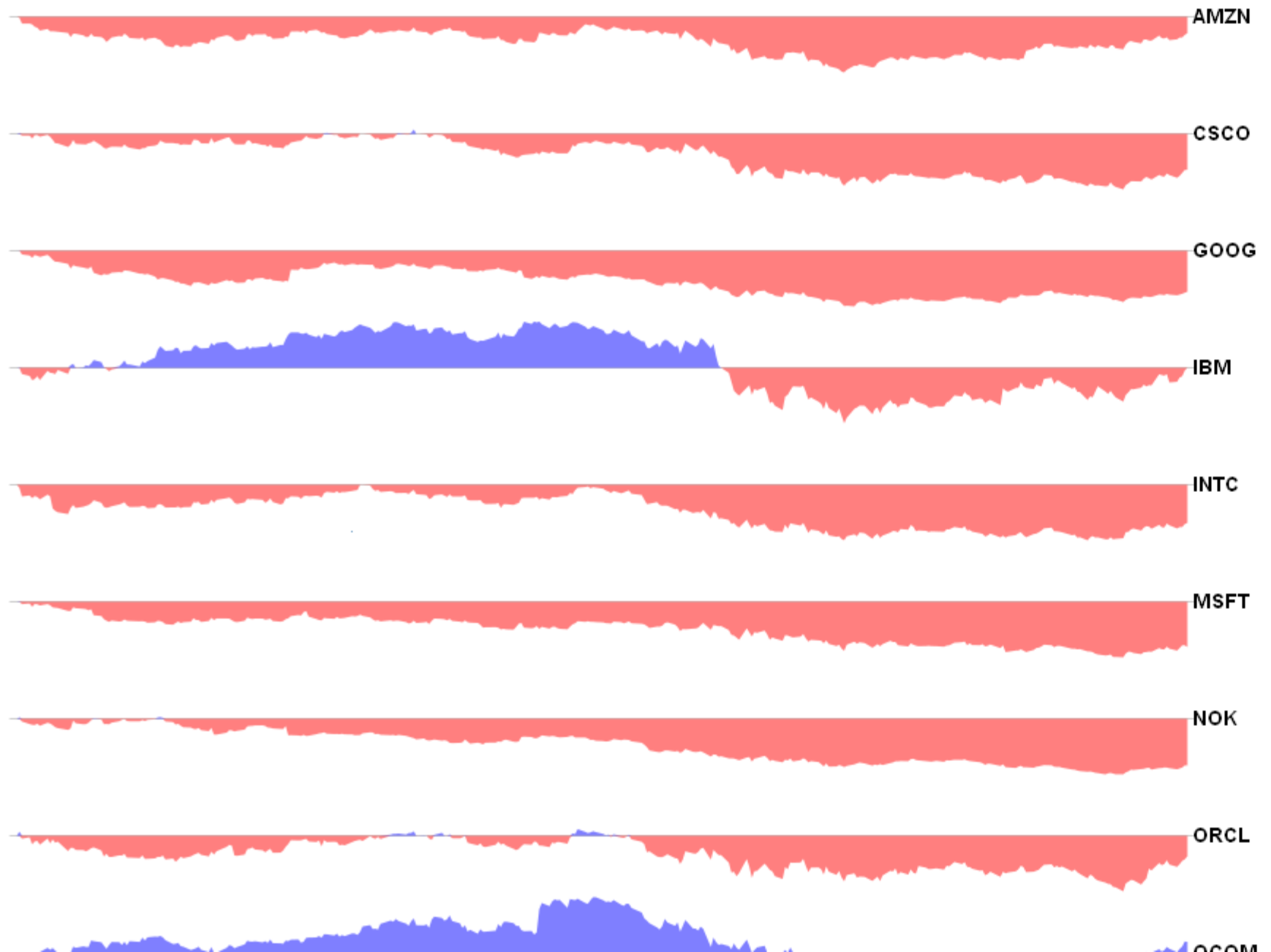
# Relative Technology Stock Performance: Jan 2008 - Present



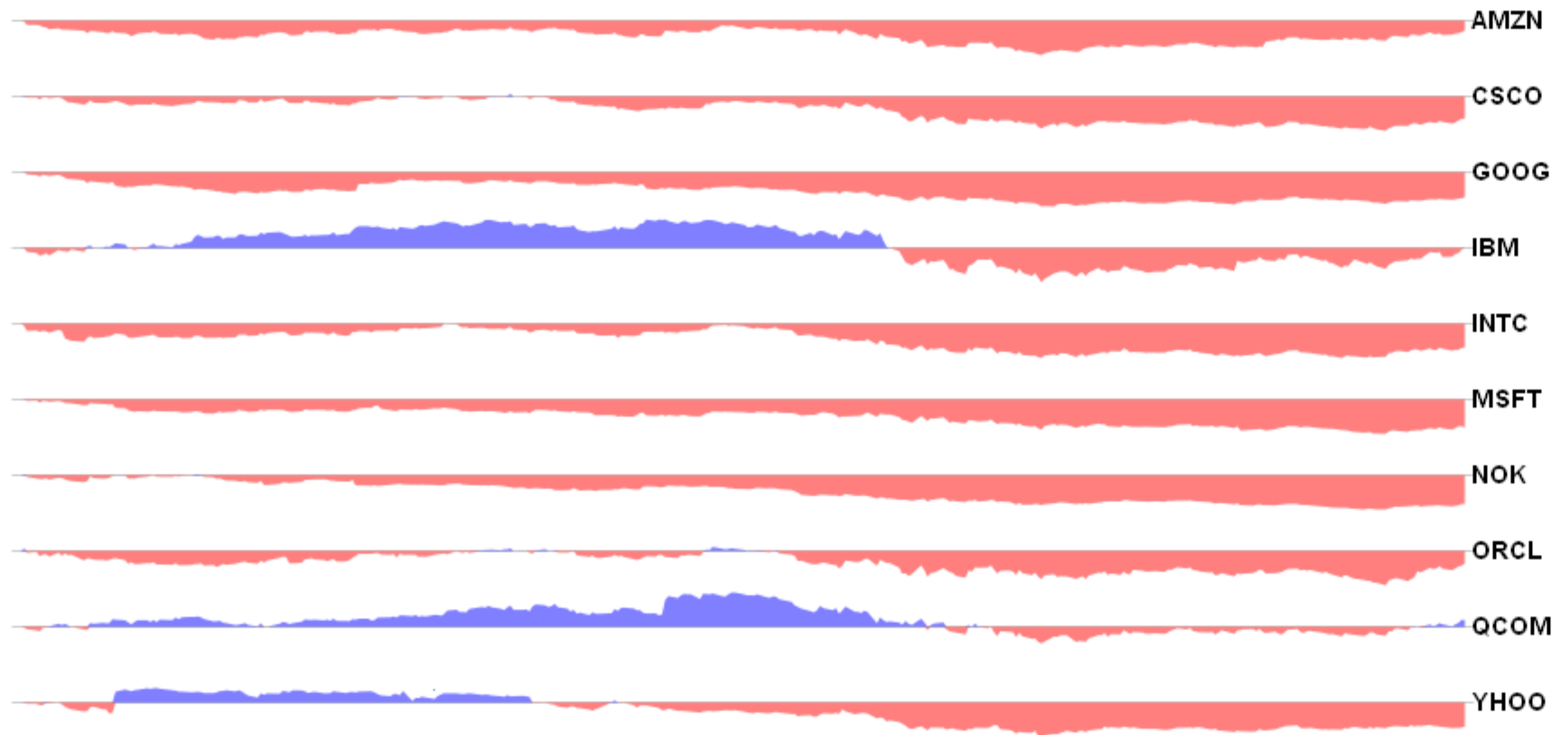




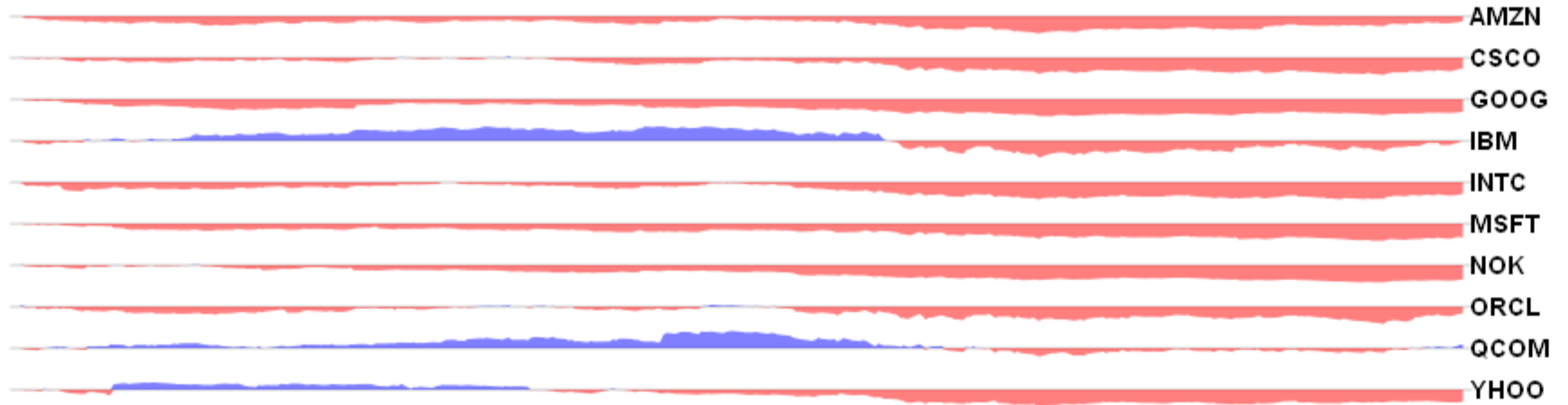
# Relative Technology Stock Performance: Jan 2008 - Present



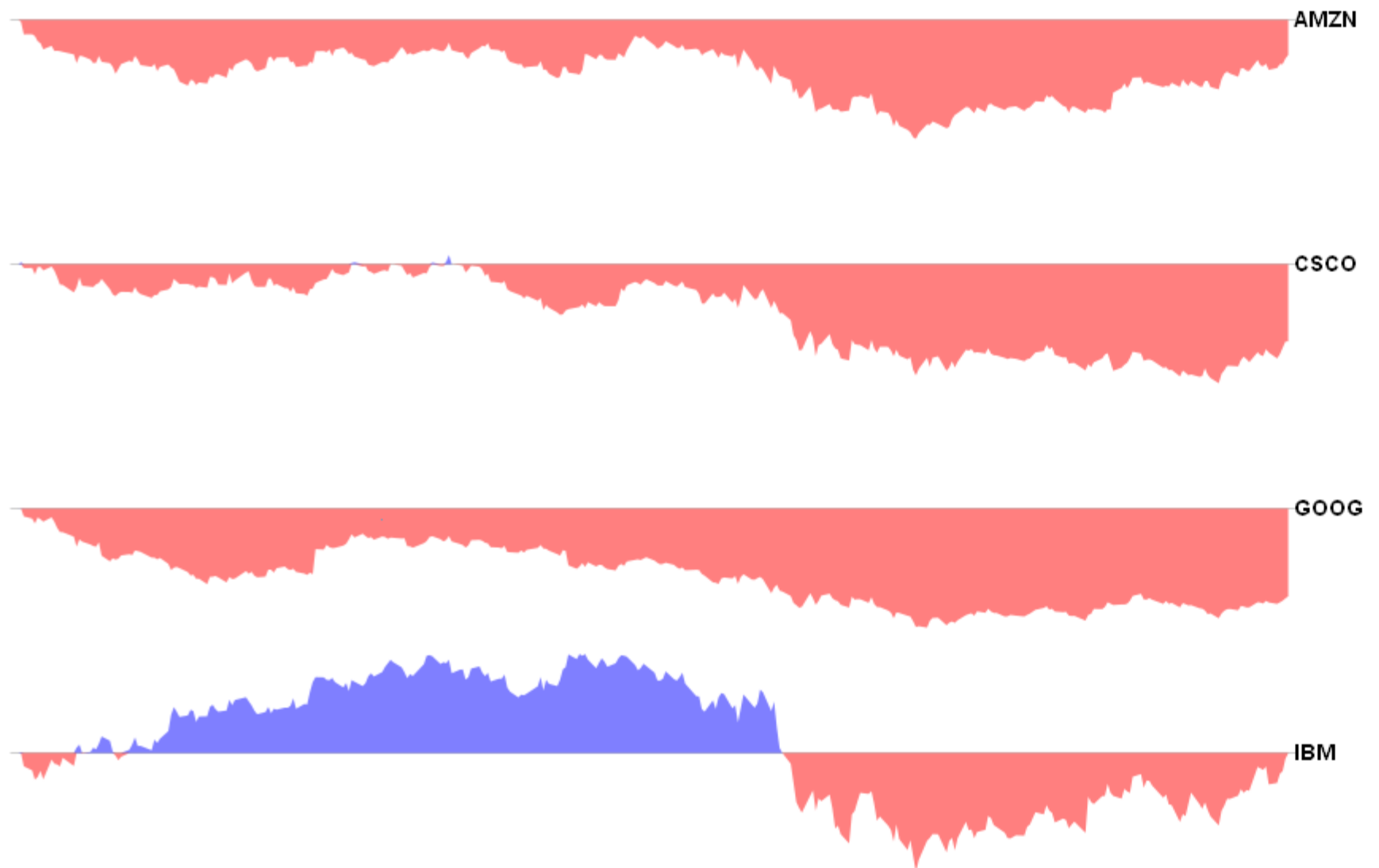
## Relative Technology Stock Performance: Jan 2008 - Present



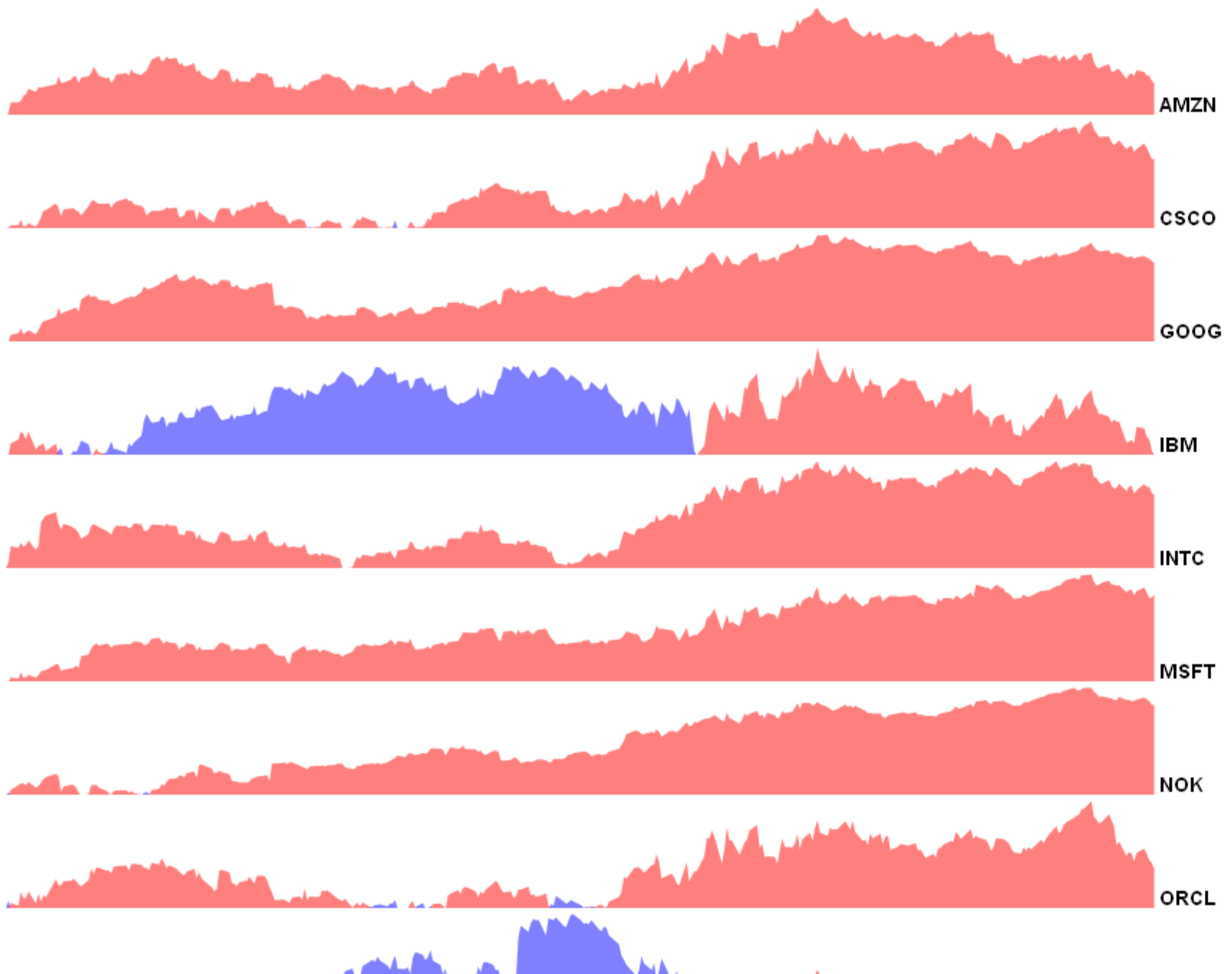
## Relative Technology Stock Performance: Jan 2008 - Present



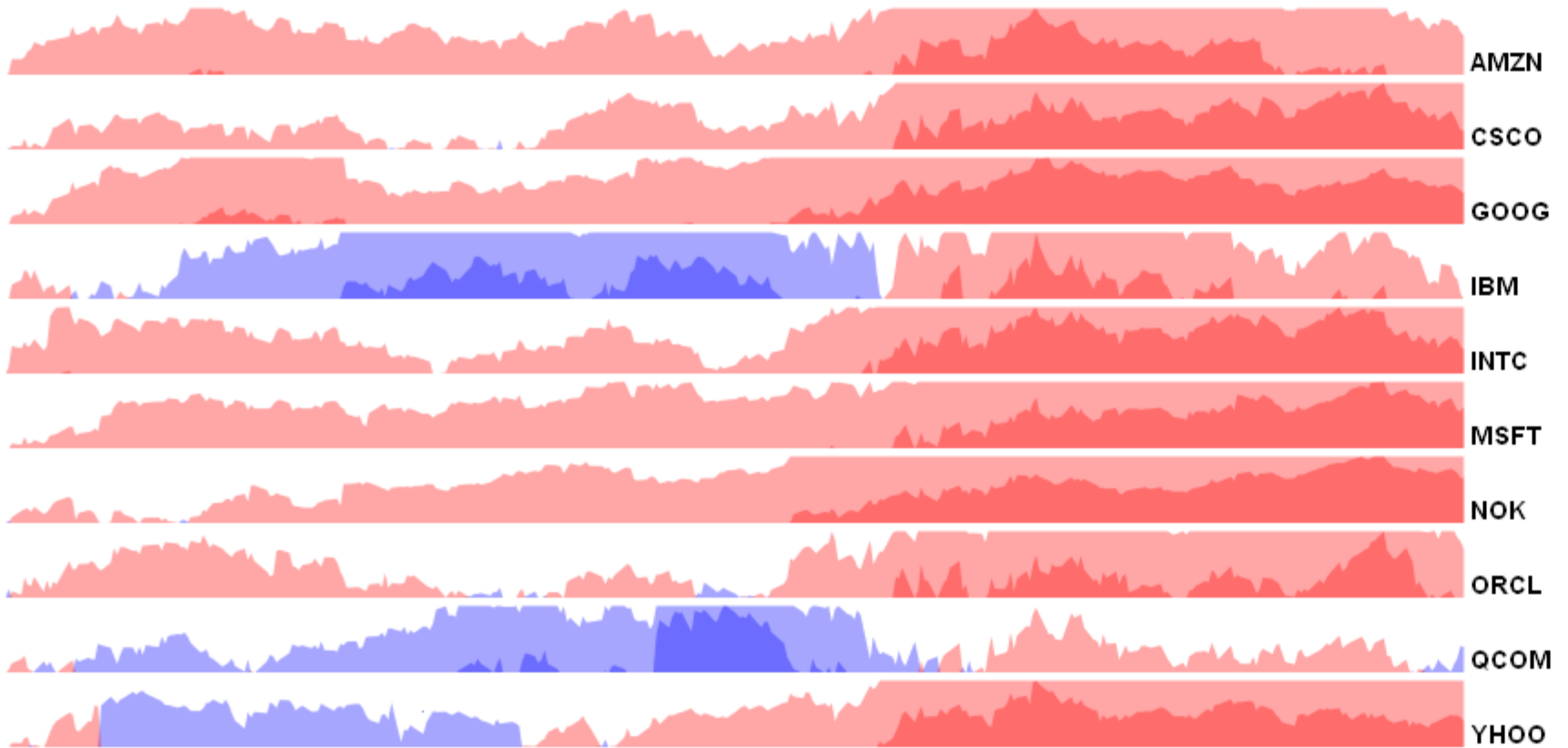
# Relative Technology Stock Performance: Jan 2008 - Present



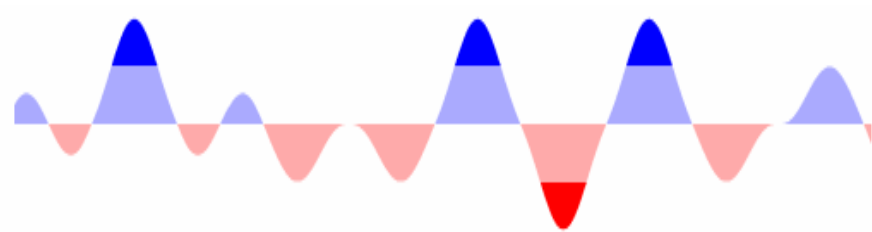
# Relative Technology Stock Performance: Jan 2008 - Present



## Relative Technology Stock Performance: Jan 2008 - Present



# Horizon Graphs



**Segment** Peaks



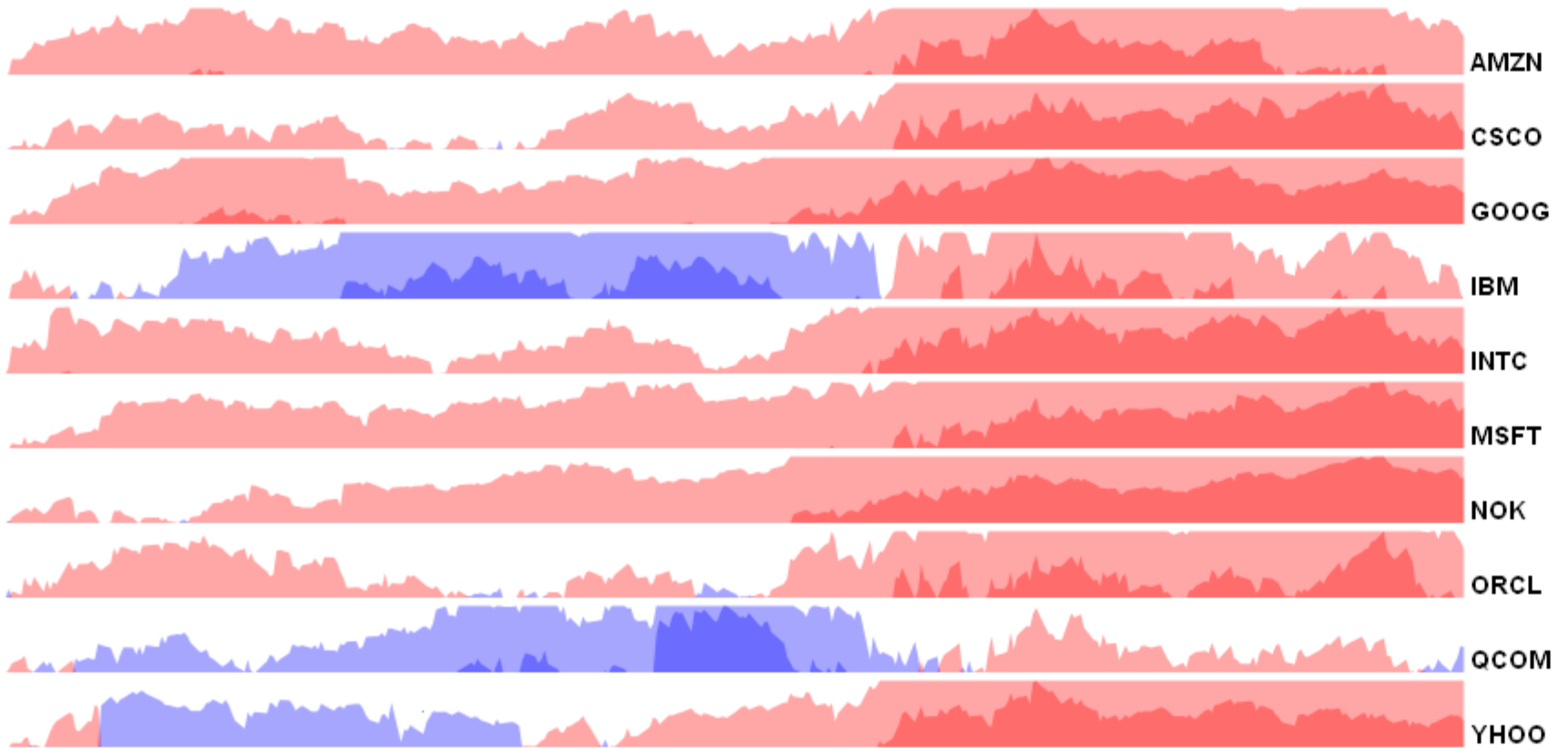
**Layer** Segments



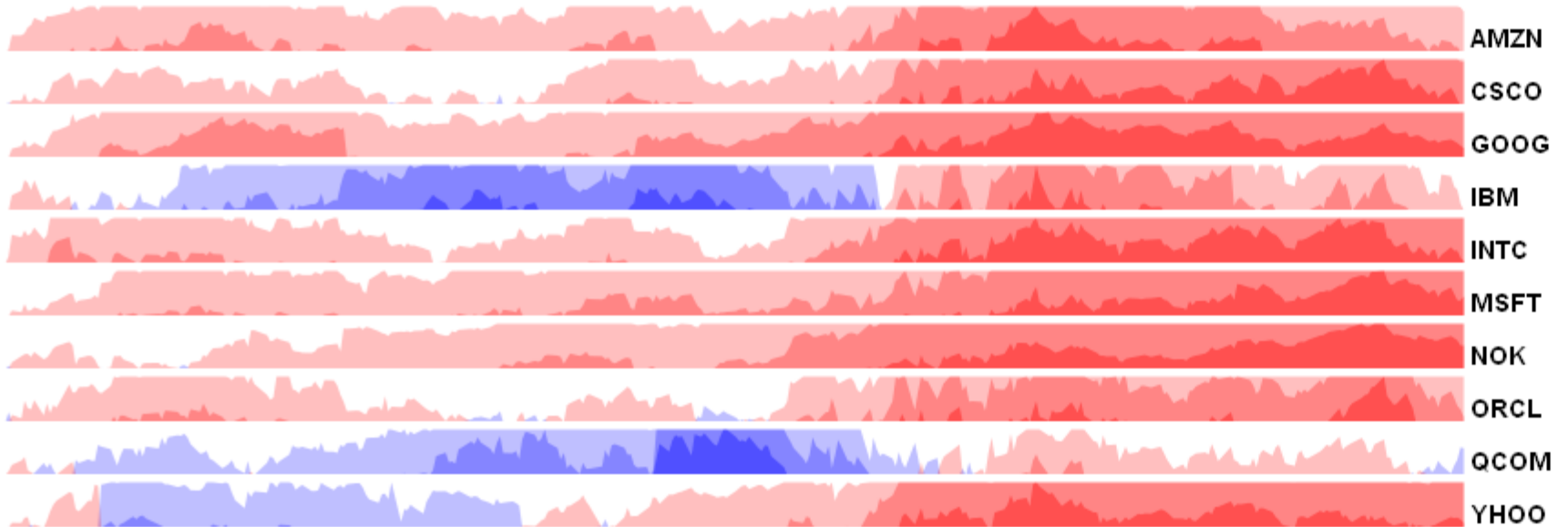
**Mirror** Negative Values



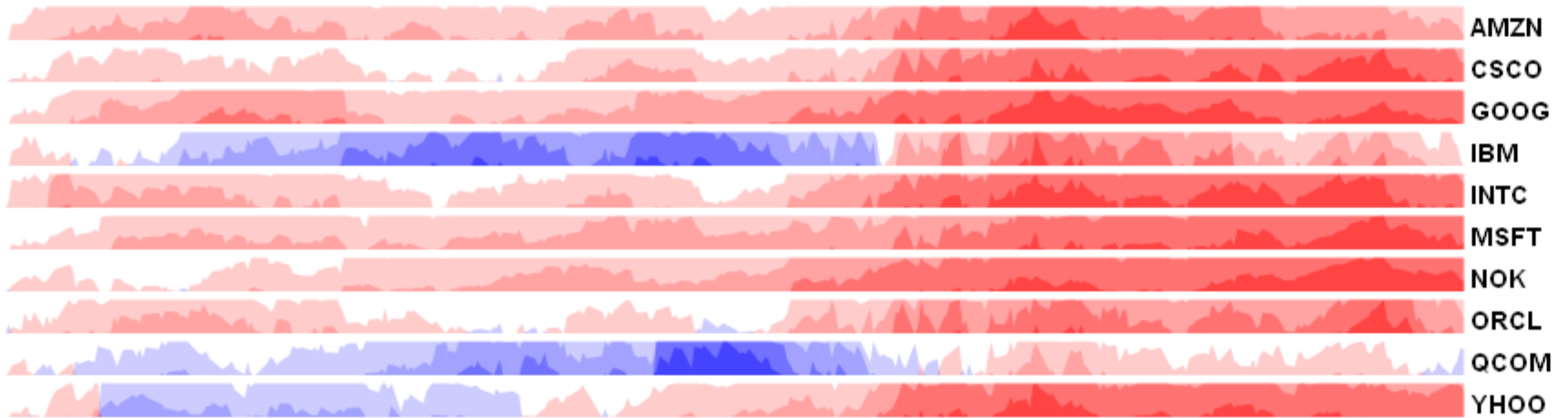
## Relative Technology Stock Performance: Jan 2008 - Present



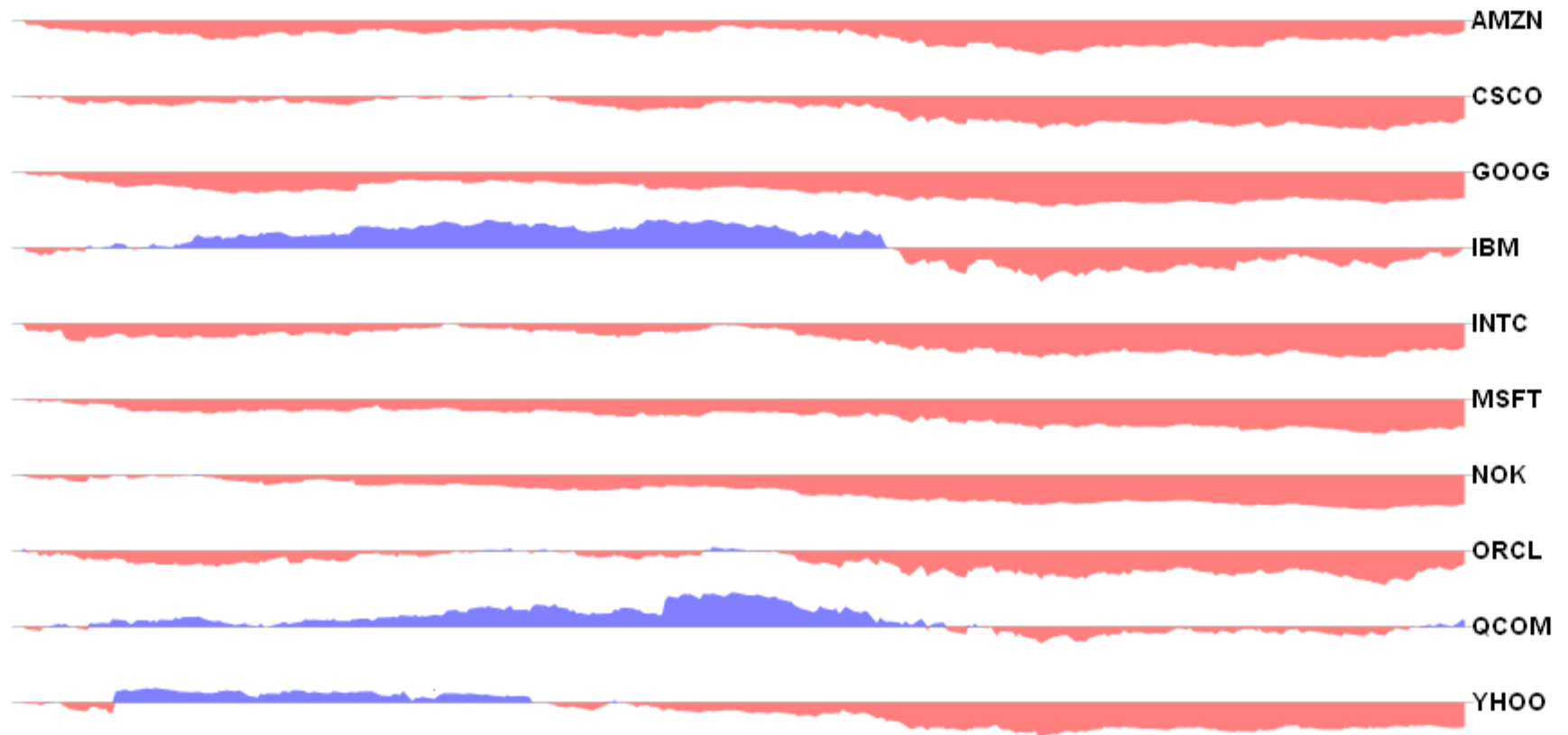
## Relative Technology Stock Performance: Jan 2008 - Present



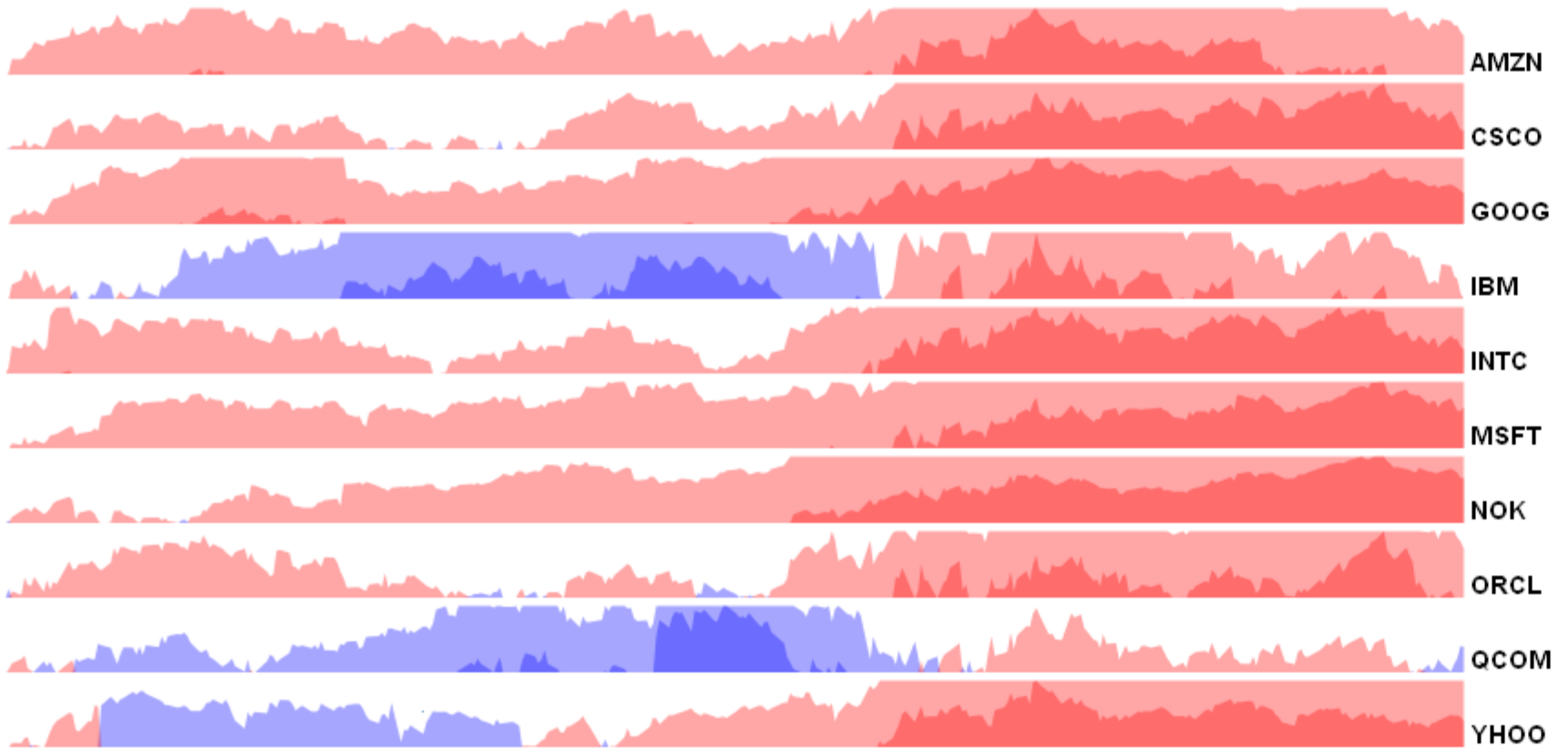
## Relative Technology Stock Performance: Jan 2008 - Present



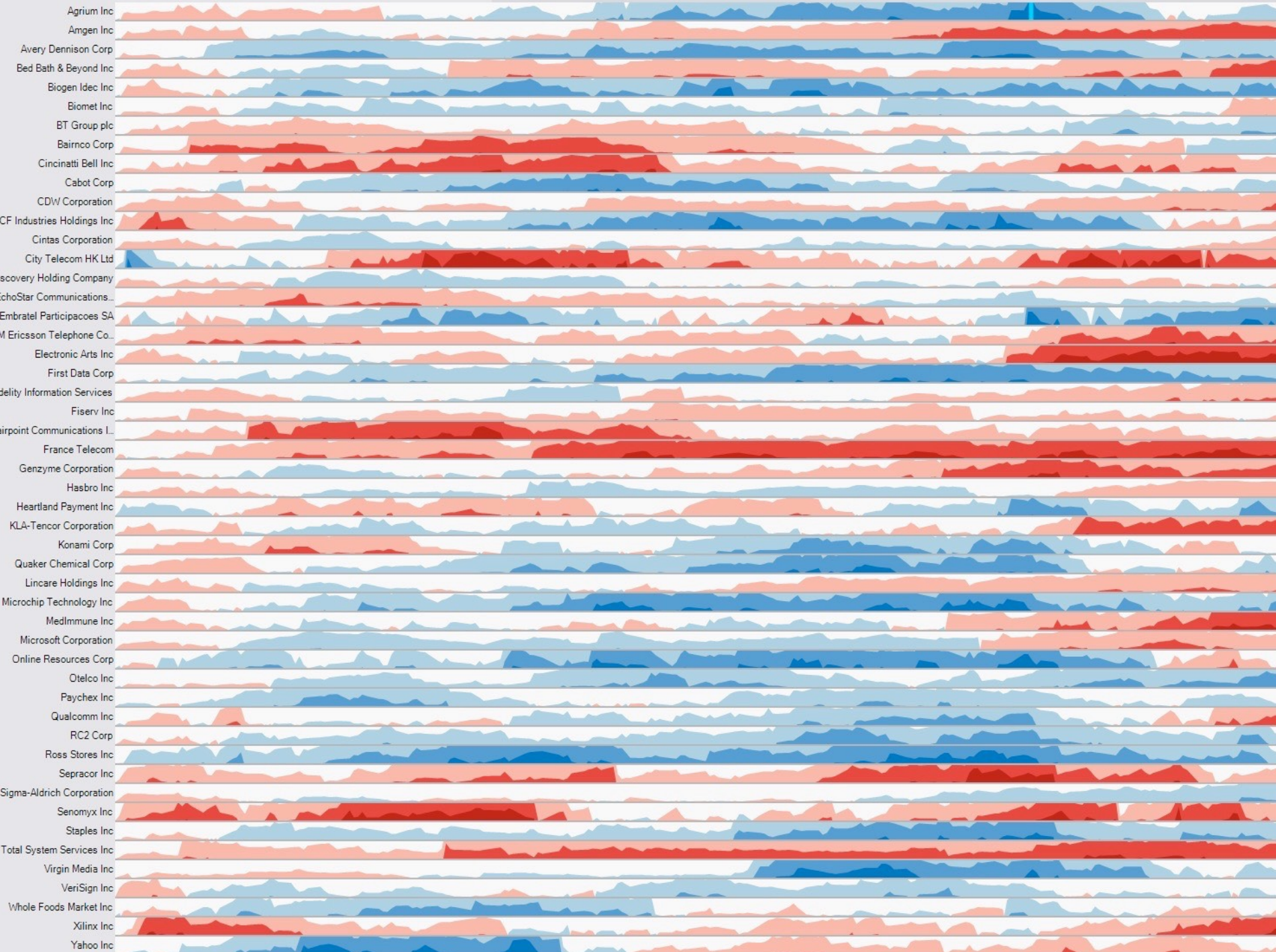
## Relative Technology Stock Performance: Jan 2008 - Present



## Relative Technology Stock Performance: Jan 2008 - Present



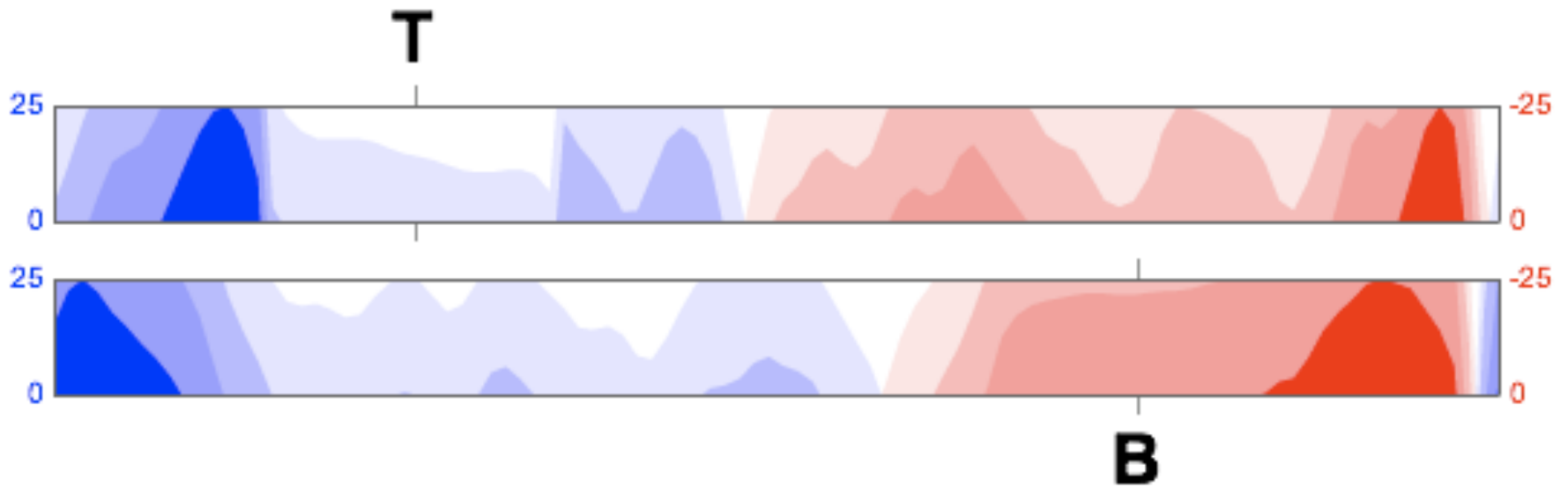
10/03/2005



# Experiment: Chart Type & Size

**Q1:** How do mirroring and layering affect estimation time and accuracy compared to line charts?

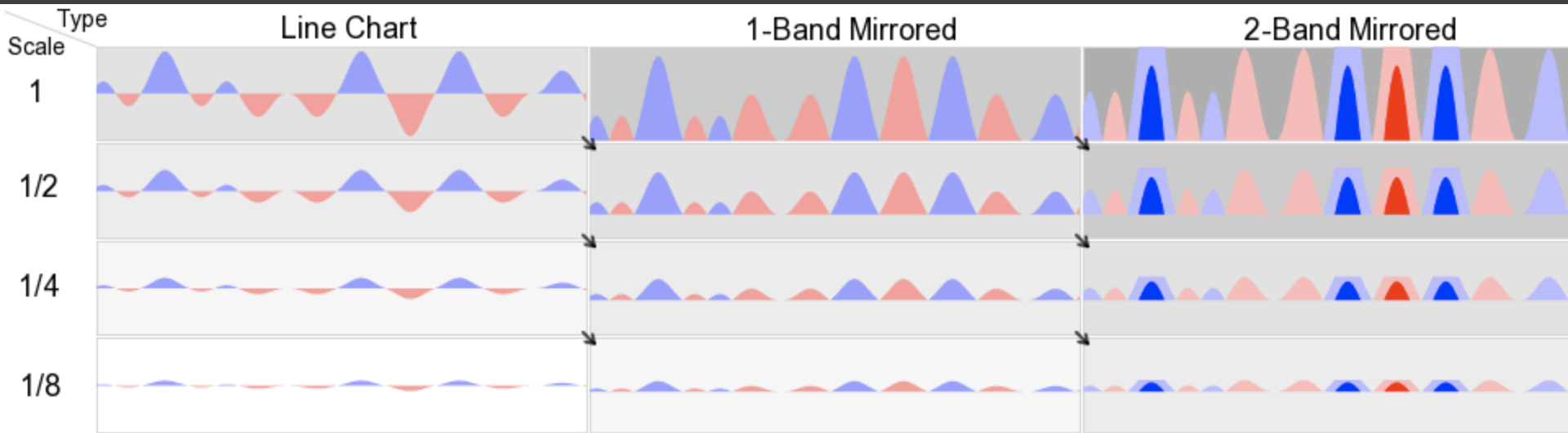
**Q2:** How does chart size affect estimation time and accuracy?



**Estimate the difference between T and B (0-200) to within 5 values.**

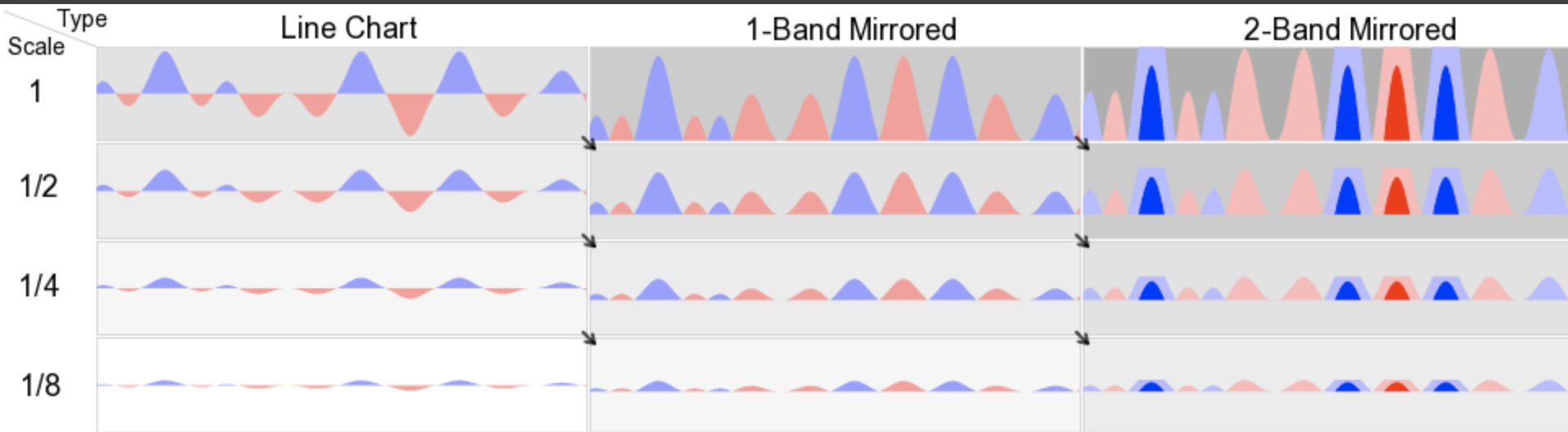


# Experiment Design



- 3 (chart type) x 4 (size) within-subjects design
- N = 30 (17 male, 13 female), undergrads
  - 14.1 inch LCD display, 1024 x 768 resolution
  - At scale = 1, chart is 13.9 x 1.35 cm (48 px)

# Experiment Design

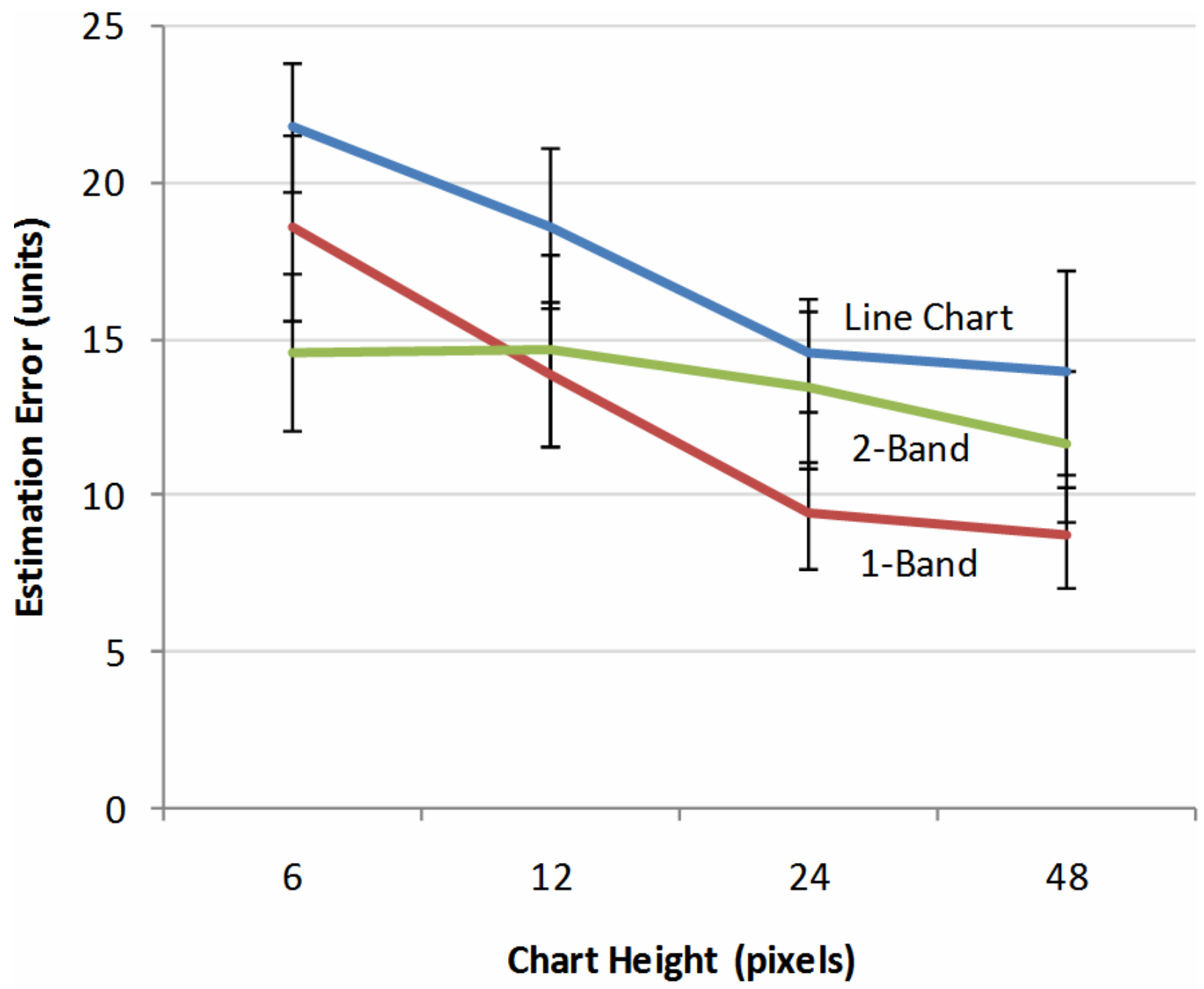


3 (type) x 4 (size) within-subjects design

N = 30 (17 male, 13 female), undergrads

2 (type) x 3 (size: 1/8, 1/12, 1/24) follow-up

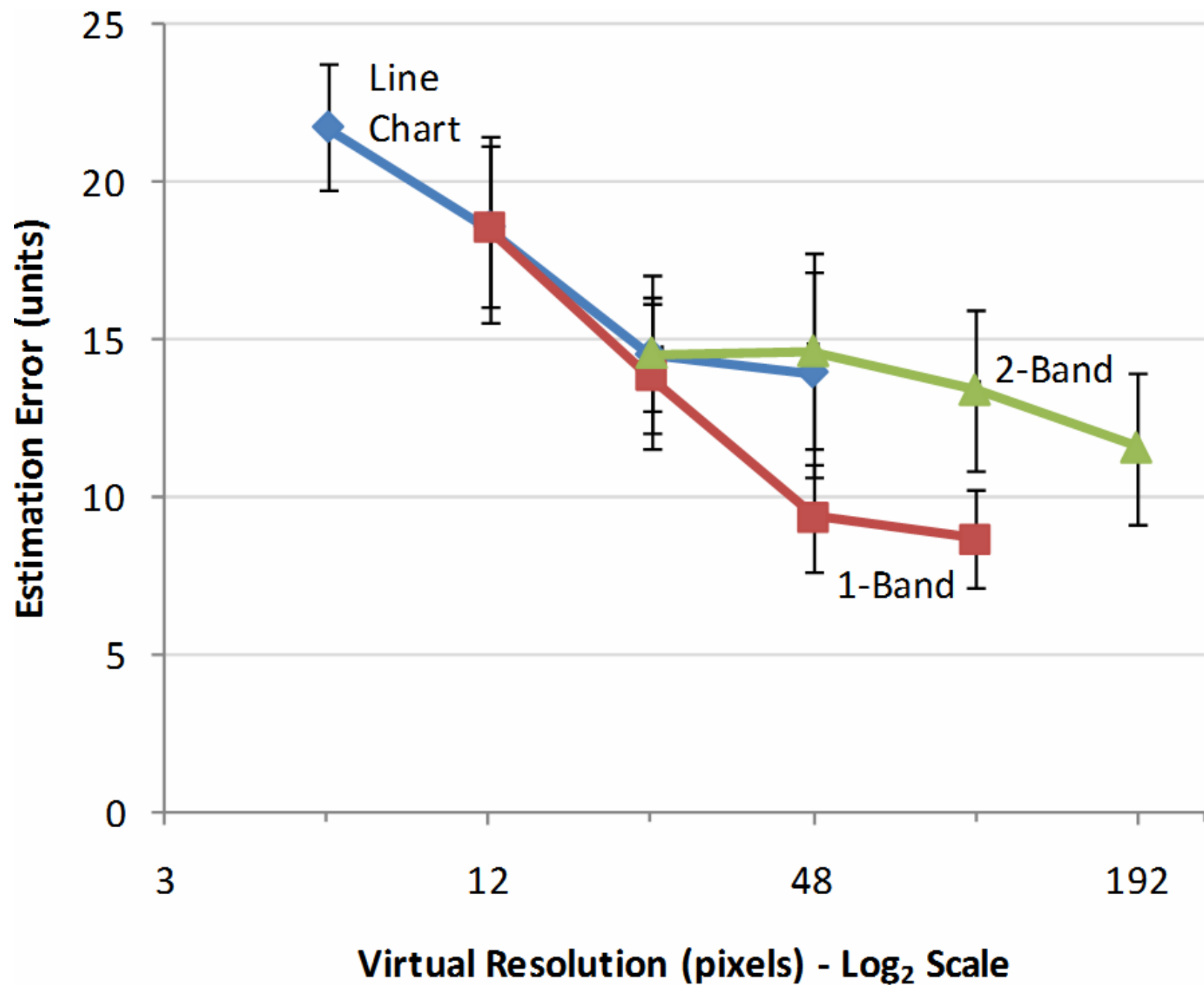
N = 8 (6 male, 2 female), engineering grads

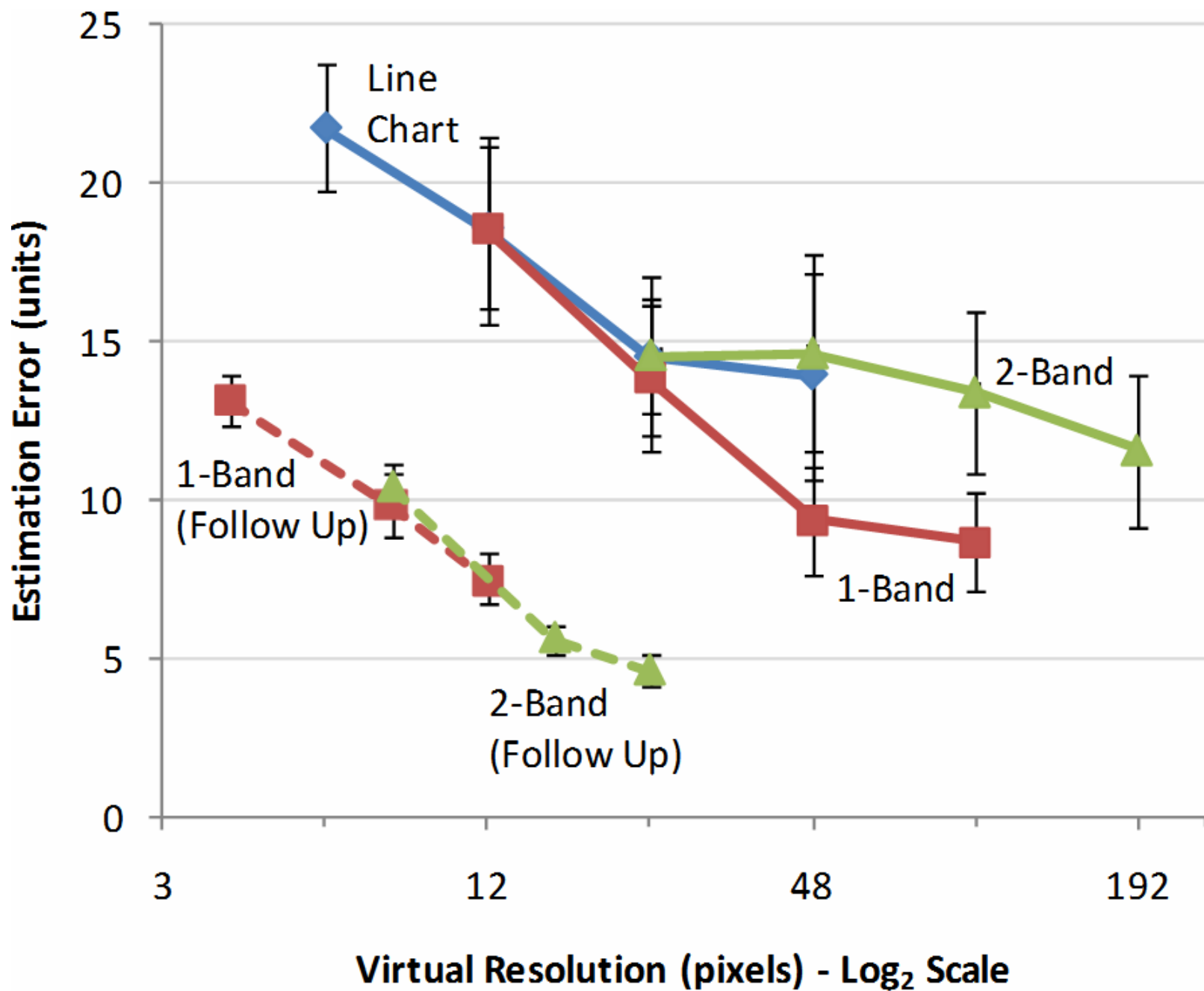


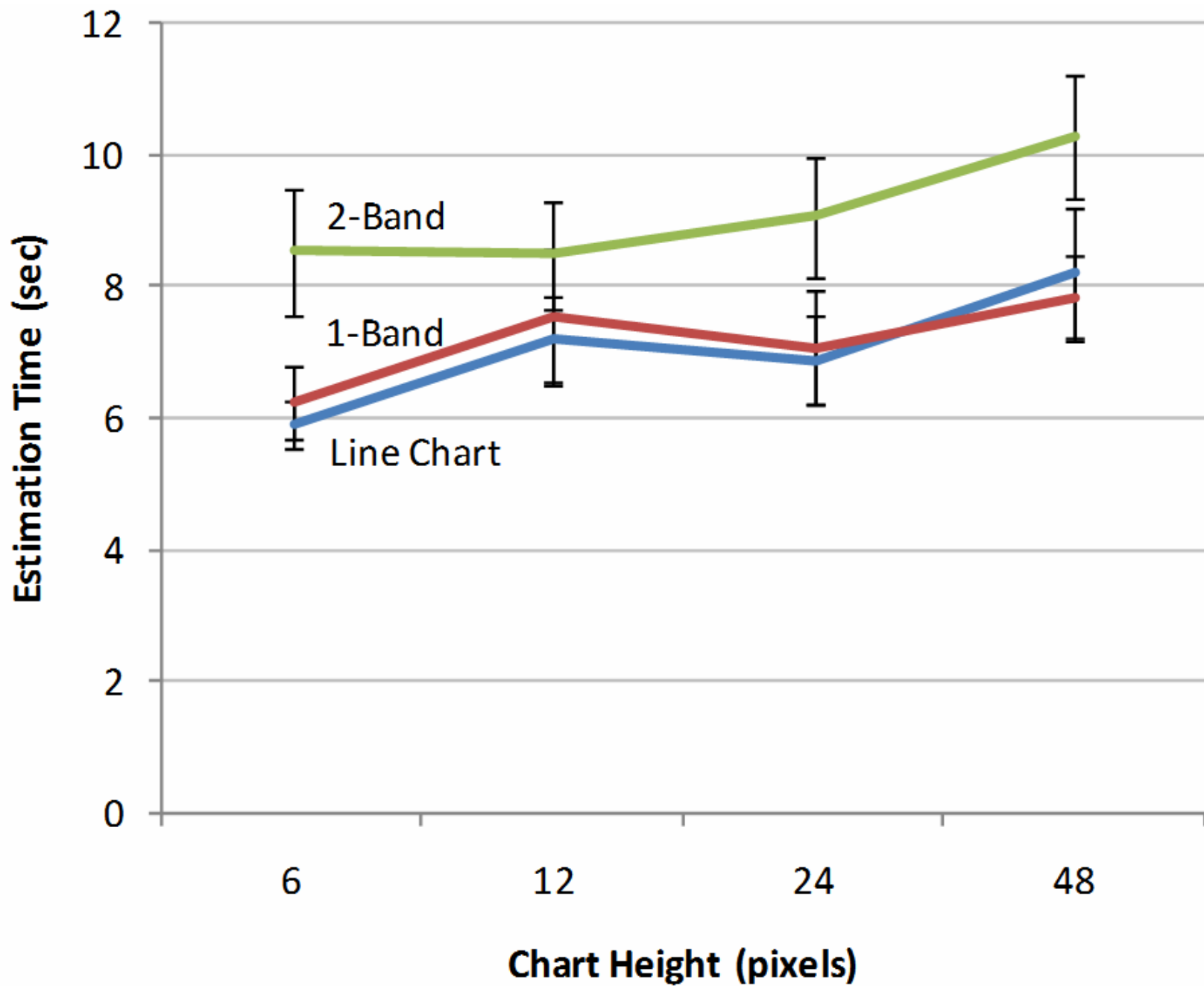
# Virtual Resolution (VR)

The un-mirrored, un-layered height of a chart









# Experiment Results

**Q1:** 2-band horizon graph (but not mirrored graph) has higher baseline estimation time and error.

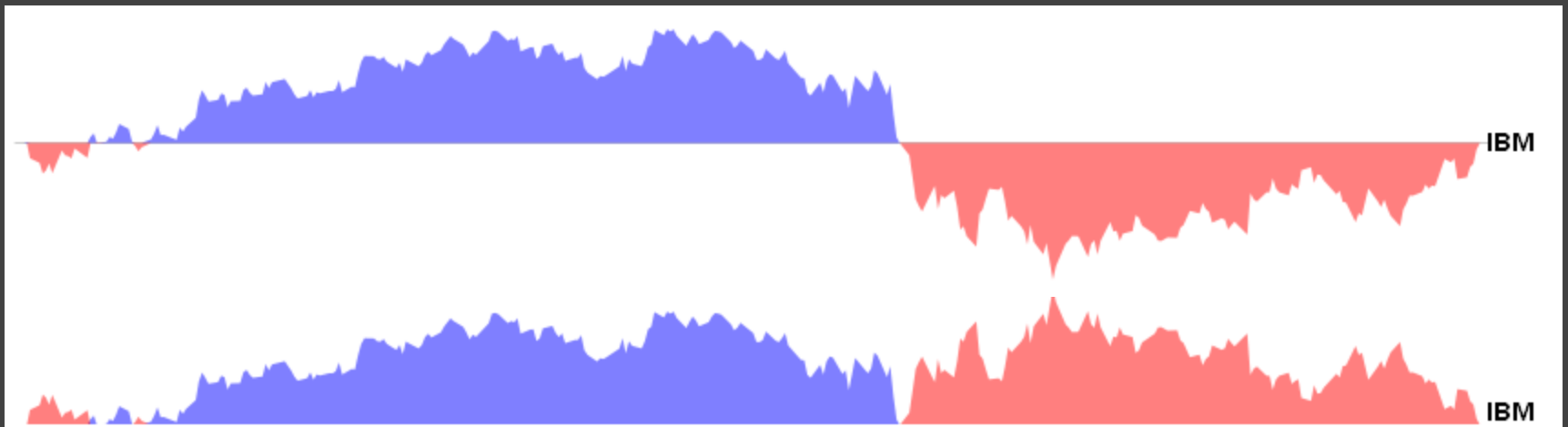
**Q2:** Estimation error increases as the *virtual resolution* decreases.

Estimation time decreases as the *physical height* decreases.



# Design Guidelines

Mirroring does not hamper perception



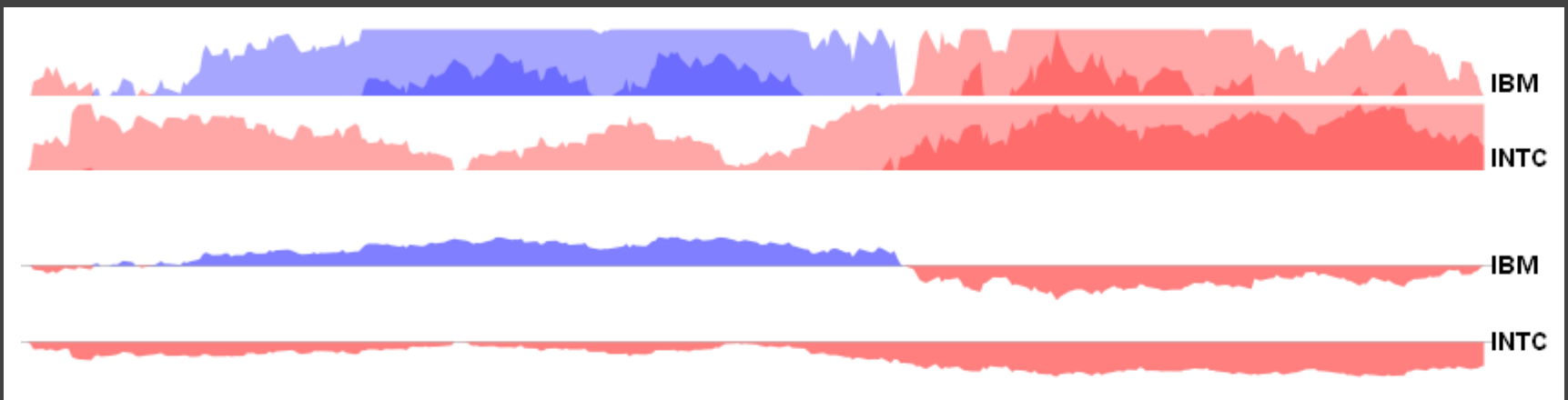
# Design Guidelines

Mirroring does not hamper perception

**Layered bands beneficial for smaller charts**

**2-band mirror charts** more accurate for heights under 6.8mm (24 pixels @ 1024x768)

Predict benefits for 3 bands under 1.7mm (6 px)



# Design Guidelines

Mirroring does not hamper perception

Layered bands beneficial for smaller charts

## Optimal chart sizing

**Sweet spots** in time/error curves

6.8mm (24 px) for line chart & mirrored chart

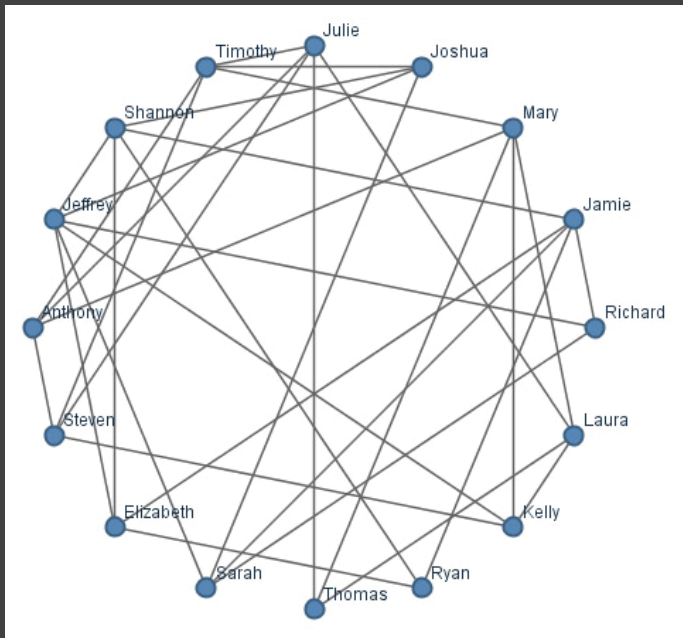
3.4mm (12 px) for 2-band horizon graph

FOLLOW-UP QUESTION:

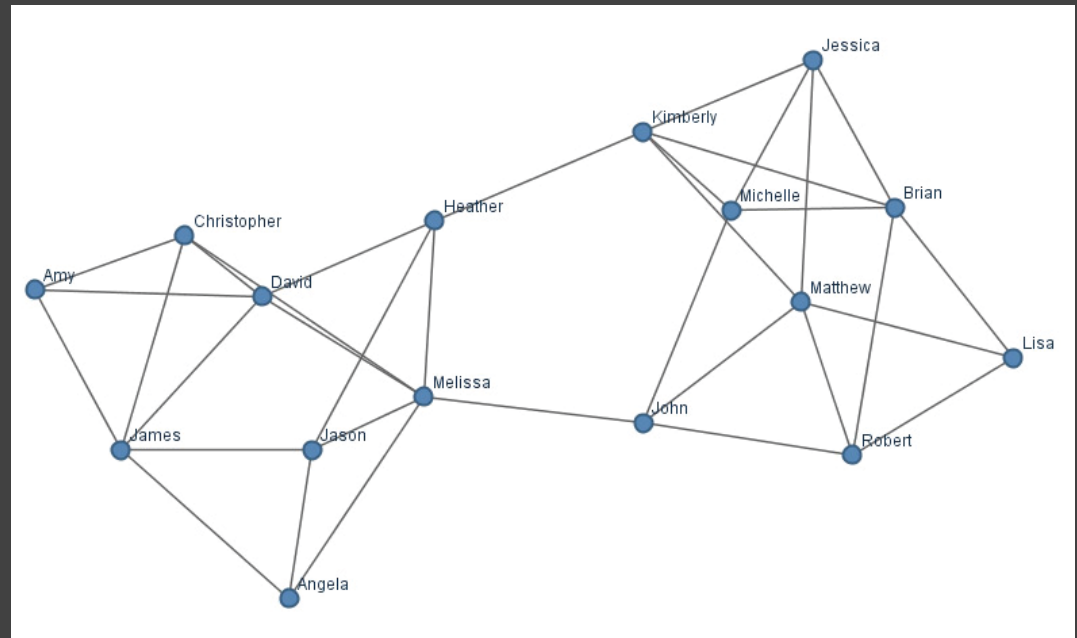
What other **tasks** and  
**performance measures**  
should one test?

# Perceptual Organization of Node-Link Diagrams

# Perceptual Organization of Graphs



Circular



Force-Directed

# Experiment Design

## Factors

Circular or Force-Directed Seed Layout

# of Between-Cluster Edges (“masking”)

All graphs had two primary clusters

## Measures

# of Edge Crossings

Average Edge Length

Average Node Distance

within or between clusters





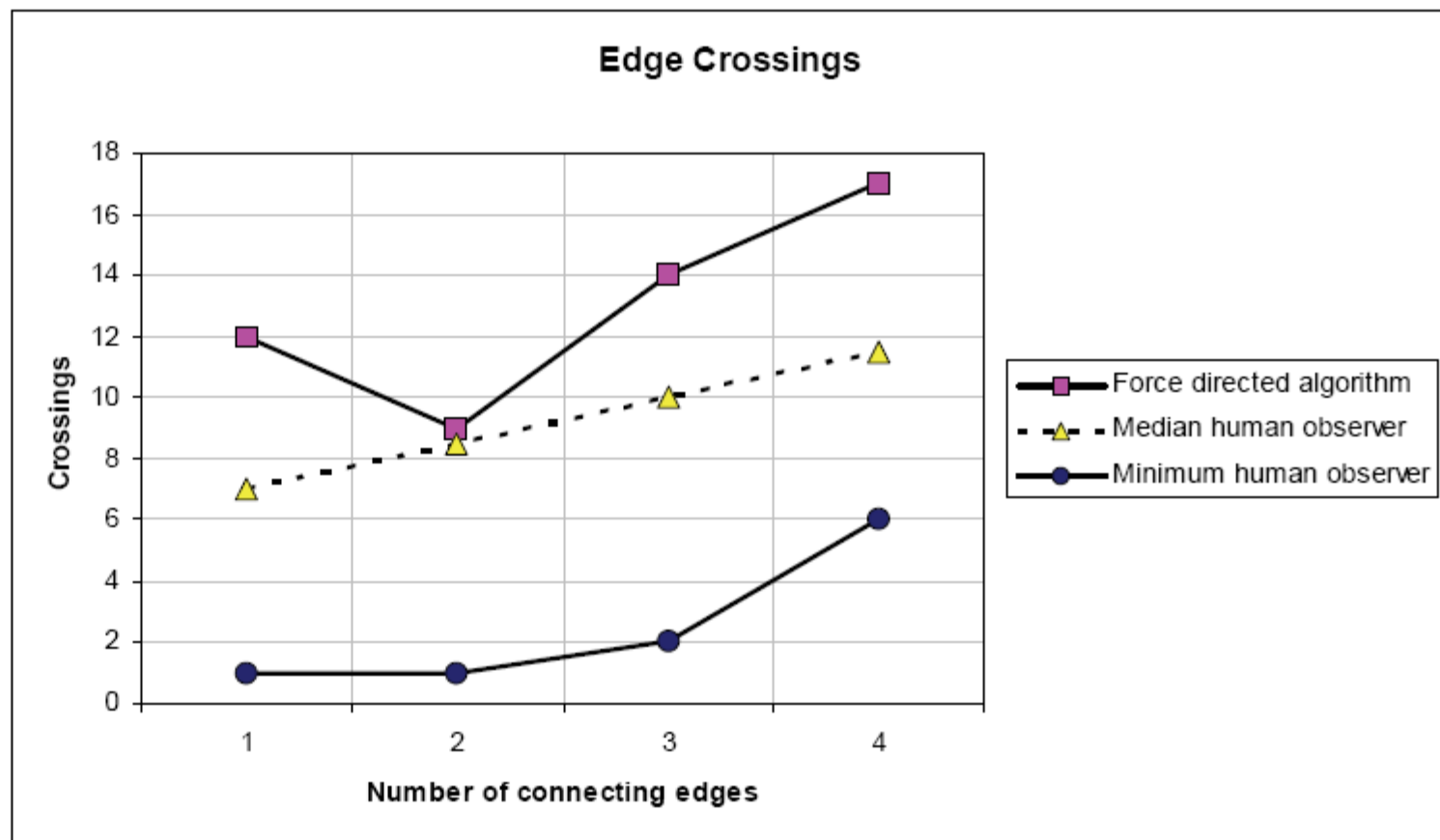


Figure 4. Edge Crossings. Human observers produced graph layouts with fewer edge crossings than the force-directed graph algorithm.

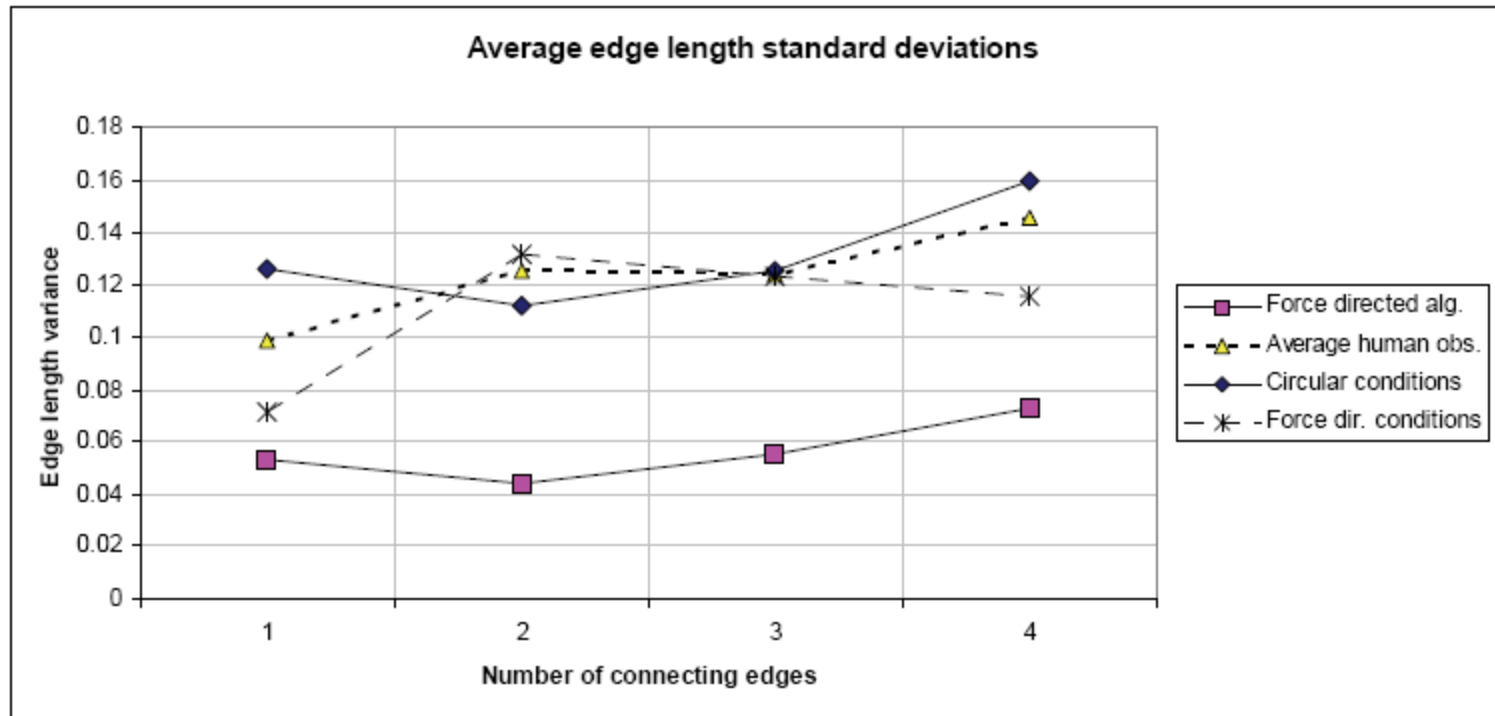


Figure 5. Edge Length Distribution. Human observers did not focus on maintaining equal edge length as much as the force directed algorithm.

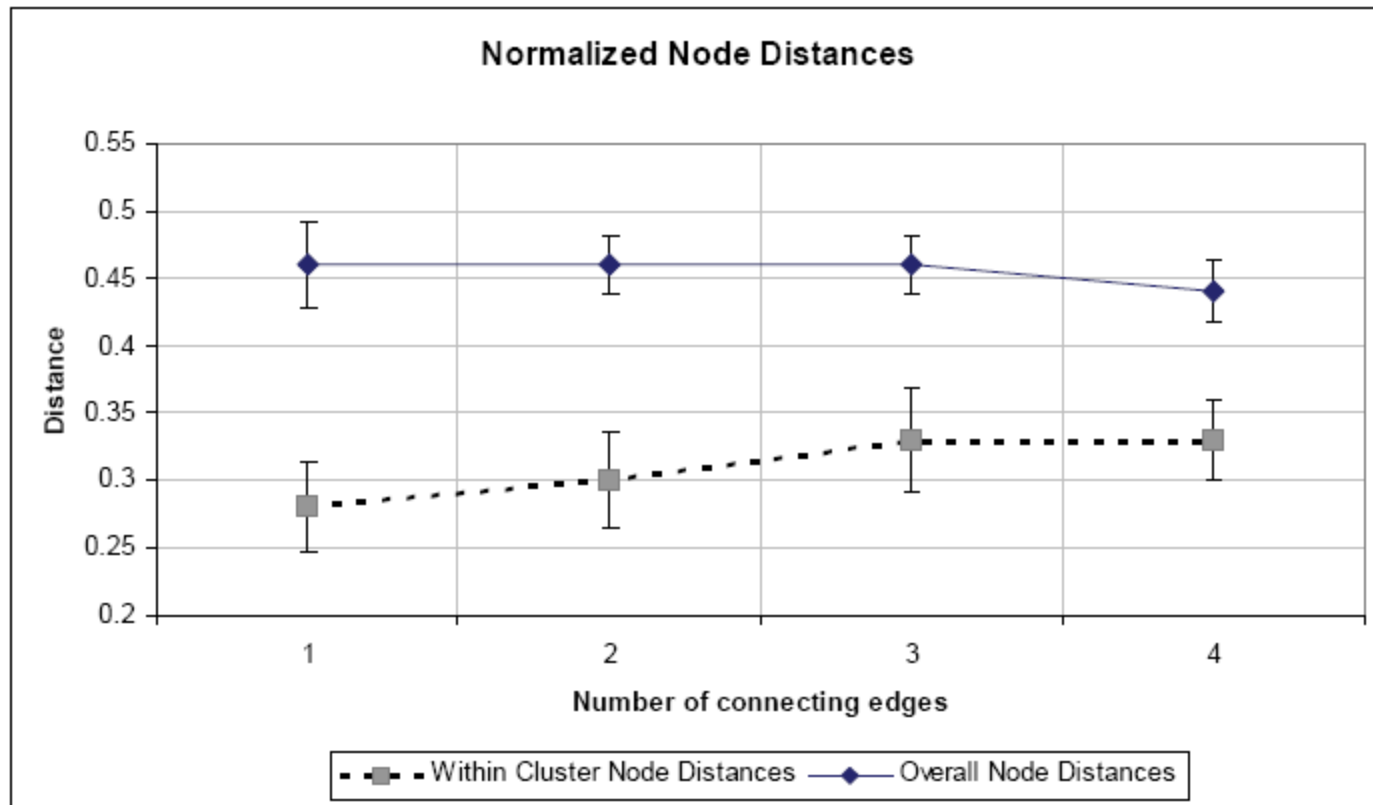


Figure 7. Cluster Extraction. For all levels of masking, the distance between nodes within a cluster is significantly smaller than the overall inter-node distance, demonstrating perceptual grouping. Error bars show 95% confidence intervals

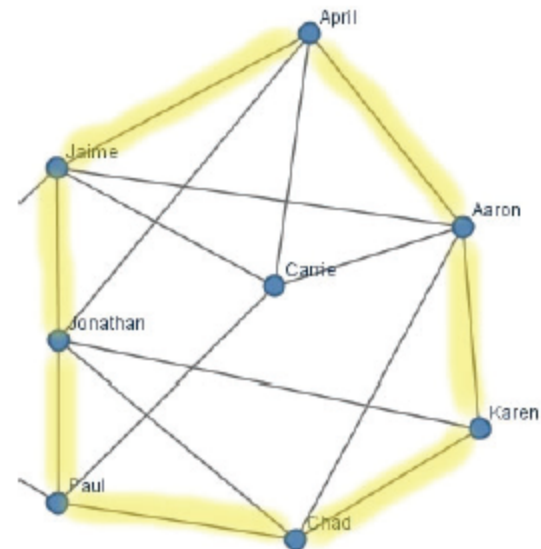
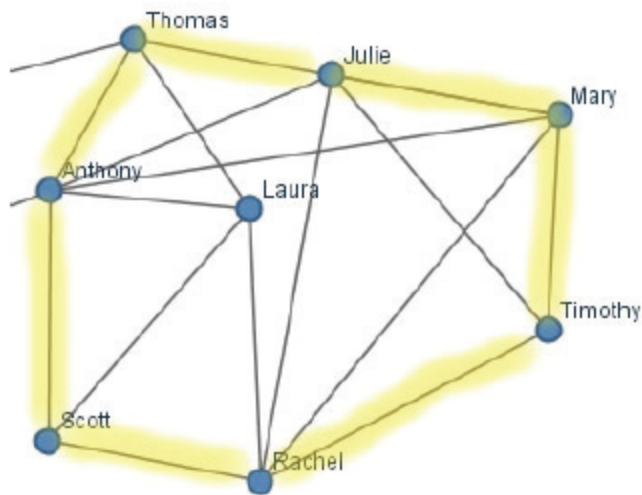


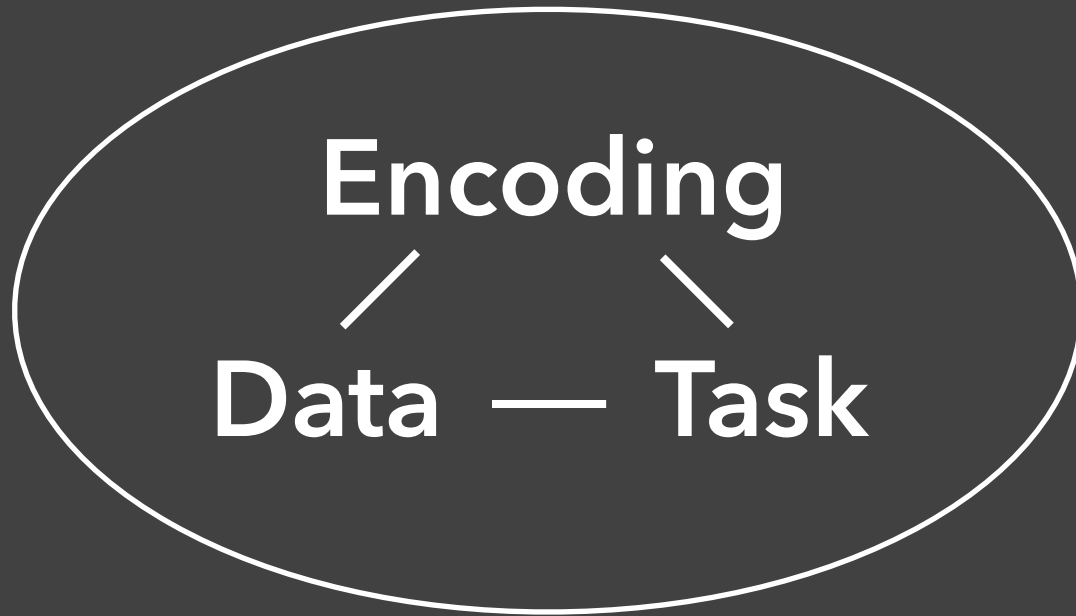
Figure 9. Cluster Hulls. Two examples of user-generated layouts where cluster edges formed a hull enclosing the cluster, organizing it into a single perceptual group.

# Summary

Design and analyze visualization techniques in context of real-world use.

Time/error analyses can be insightful, but they don't provide a complete picture.

Performance measures may be more suited to serious analysis than casual use?



**Users & Domain**

# Administrivia

# Final Project Deliverables

## **Poster Printing**

Non-CSE teams should have sent printing requests to us already. Others can print on their own!

## **Final Deliverables**

Post to GitHub repo by 11:59pm, Wed 5/30.

## **Final Project Showcase**

CSE Atrium, Thur 5/31. Open to public 12-2pm.  
*Arrive at 11:45am! Charge your laptop battery!*

[Read assignment description for more!](#)



# Course Evaluation

**Official course evaluation, due by 6/3**

*Counts as part of course participation.*

Your opinion is valued!

<https://uw.iasystem.org/survey/192876>

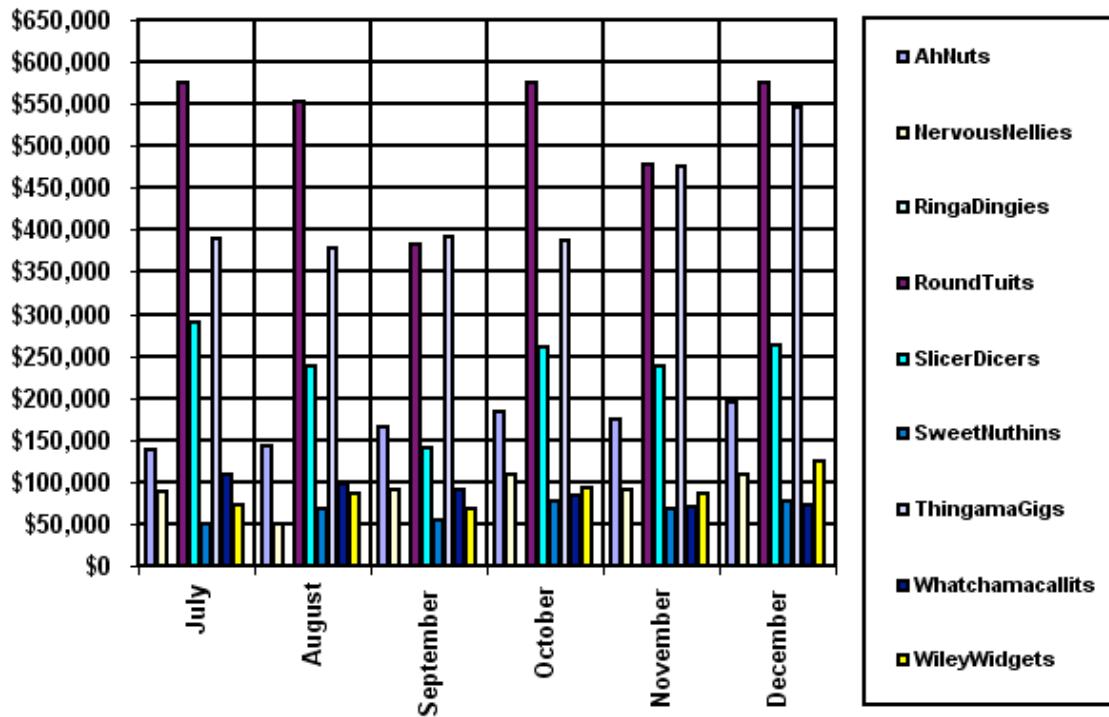
# Course Summary

# Data and Image Models

		LES VARIABLES DE L'IMAGE									
		POINTS			LIGNES			ZONES		12	14
Z	XY 2 DIMENSIONS DU PLAN										
	TAILLE										
	VALEUR										
		LES VARIABLES DE SÉPARATION DES IMAGES								13	
	GRAIN										
	COULEUR										
	ORIENTATION										

# Visualization Design

## SlicerDicers' Sales Compared to Other Products

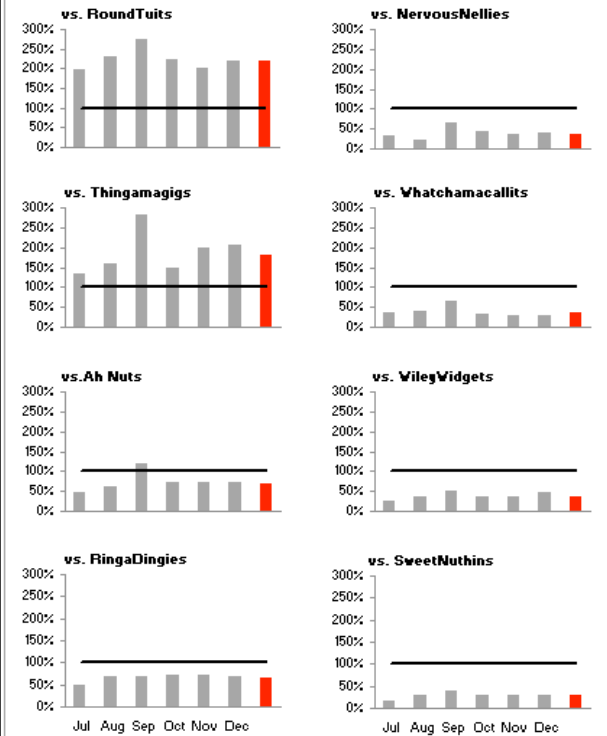


Problematic design

## Sales of SlicersDicers Compared to Other Products

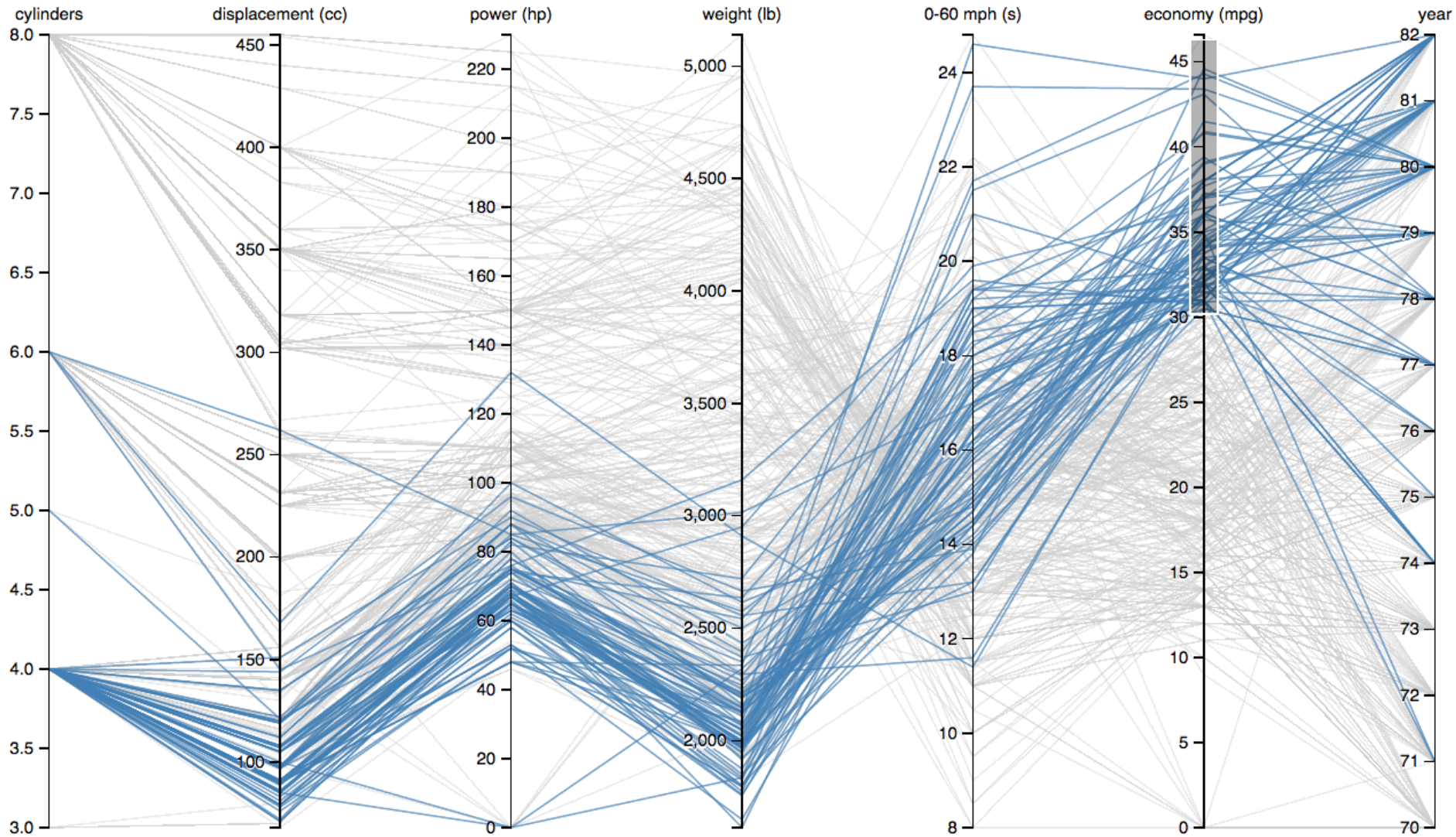
July - December, 2003

(SlicersDicers' sales are displayed as black reference lines of 100%; the red bars represent the average monthly sales percentage for July through December.)

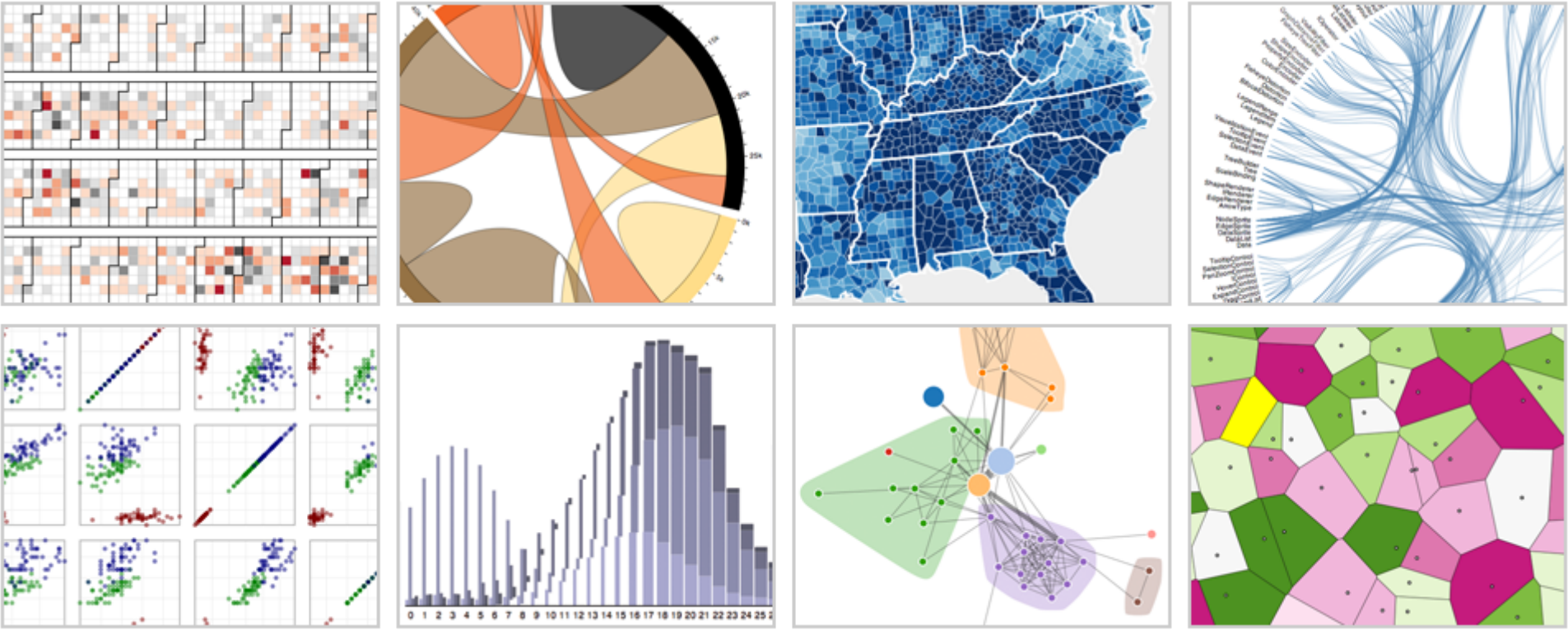


Redesign

# Exploratory Data Analysis

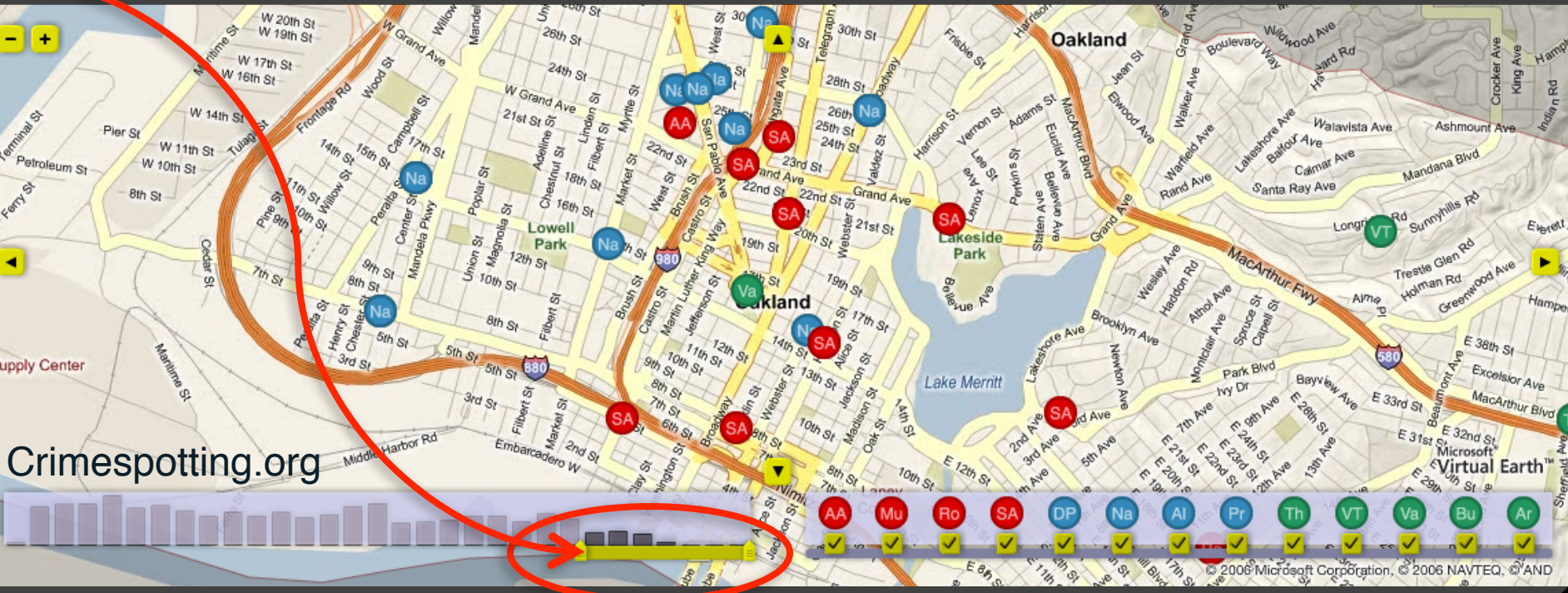
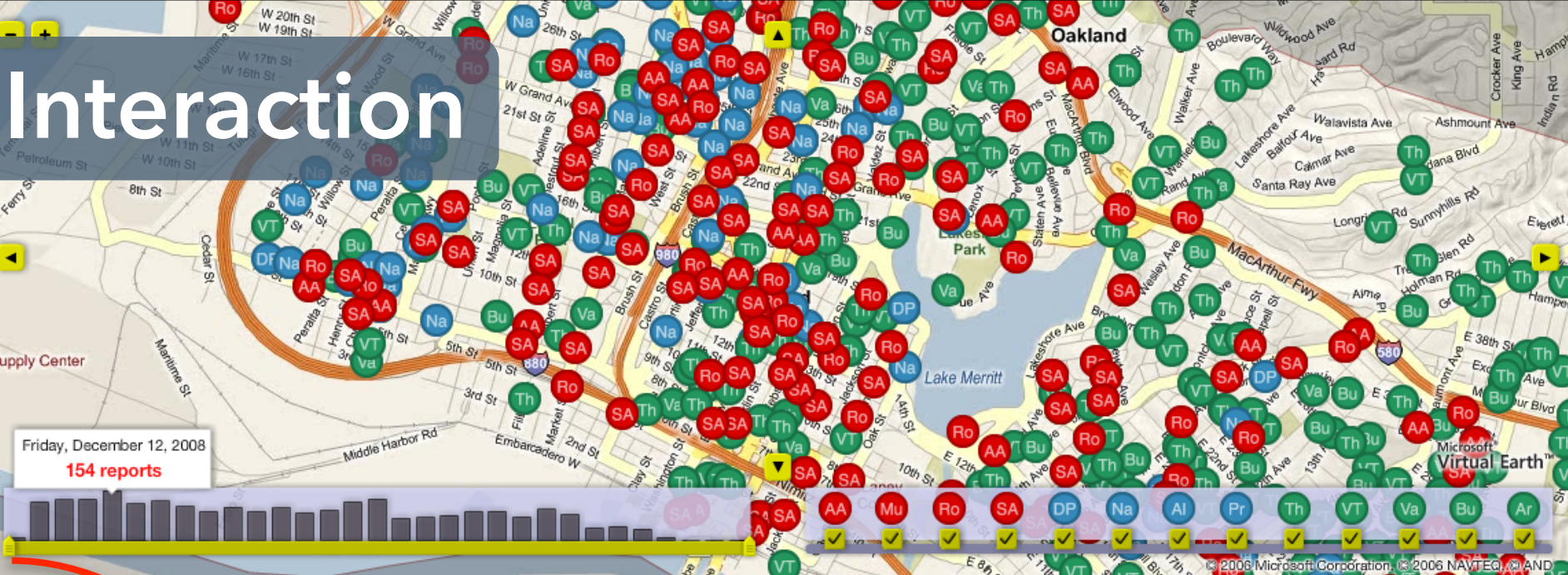


# Visualization Software

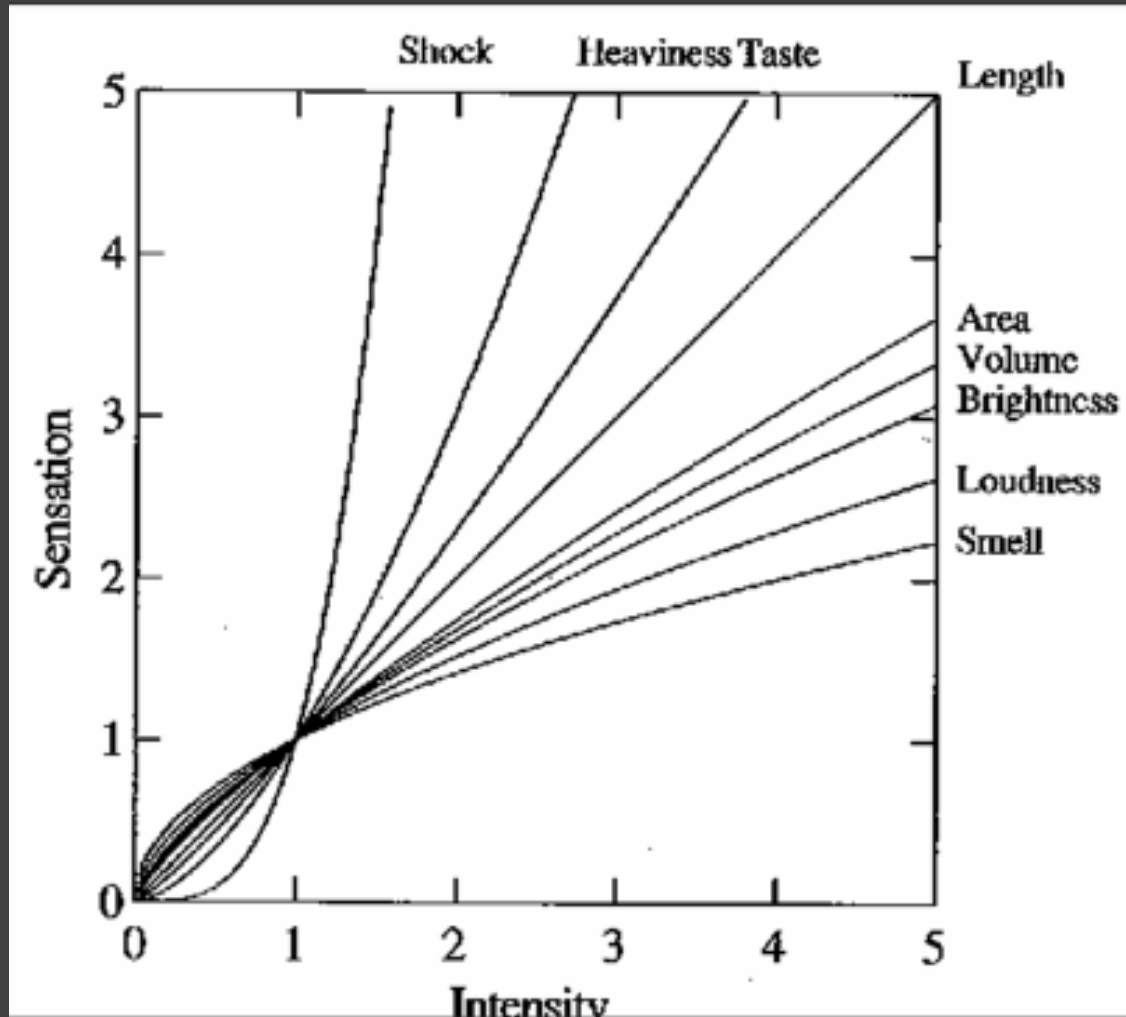


D3: Data-Driven Documents

# Interaction



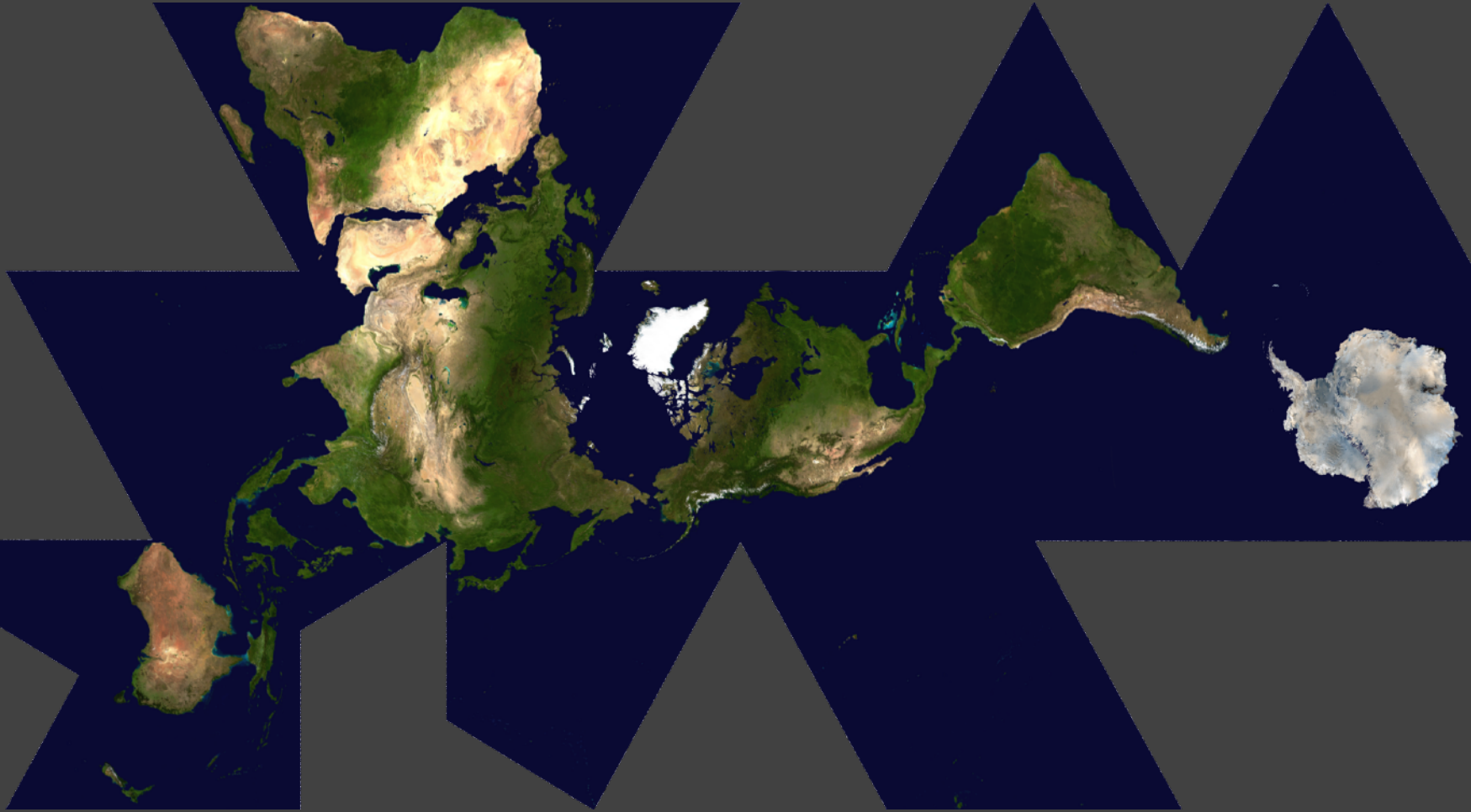
# Graphical Perception



The psychophysics of sensory function [Stevens 61]



# Maps

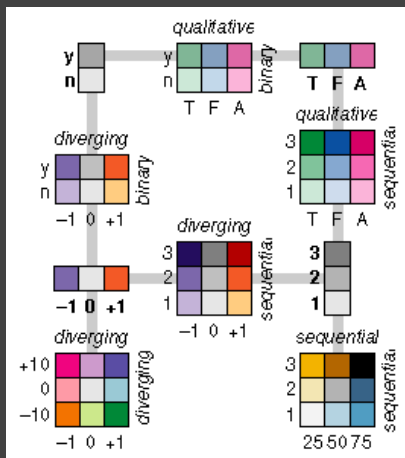
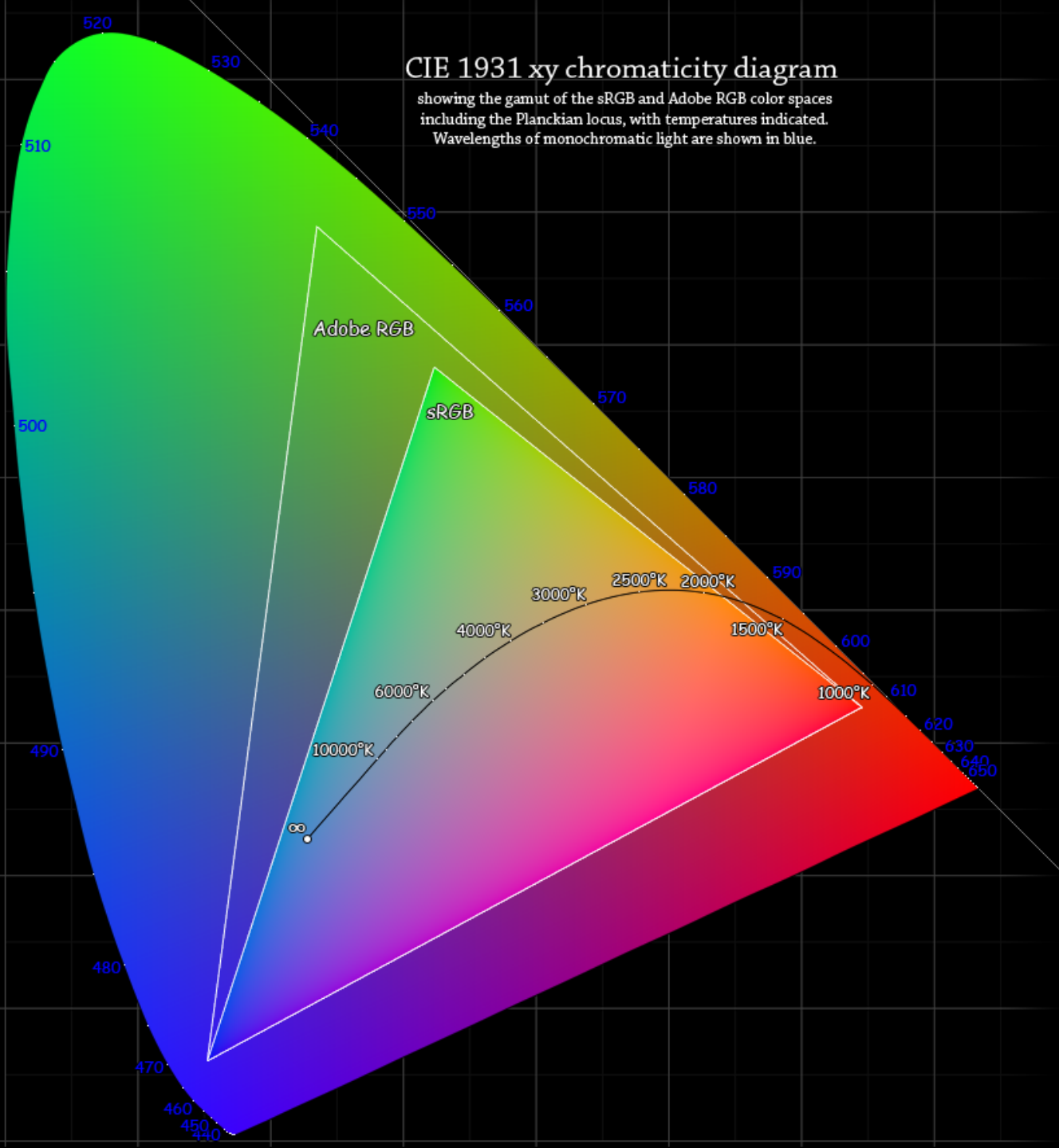


Dymaxion Maps [Fuller 46]

# Color

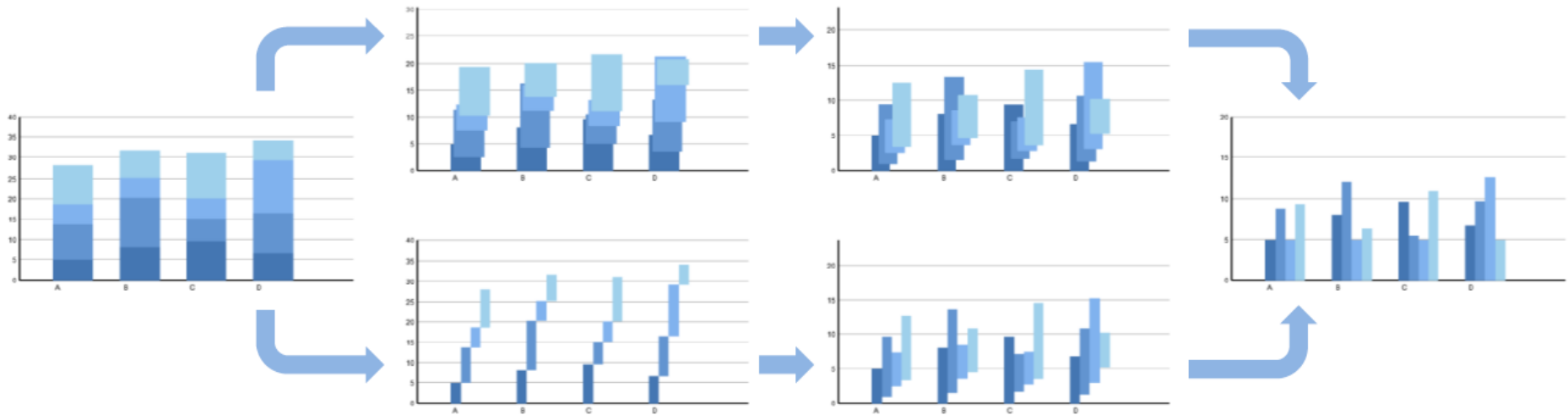
## CIE 1931 xy chromaticity diagram

showing the gamut of the sRGB and Adobe RGB color spaces including the Planckian locus, with temperatures indicated. Wavelengths of monochromatic light are shown in blue.



Color Brewer

# Animation



Animated transitions in statistical data graphics [Heer & Robertson 07]

# Narrative

Recent elections have placed a heavy emphasis on “swing states” — Ohio, Florida and the other competitive states. Yet in the past, many more states shifted between the Democratic and Republican parties. A look at how the states stuck or shifted over time and how they have shifted over past elections.

- Each box represents a state sized by number of electoral votes.
- Each curve shows how much it shifted left or right between elections.

Chart Size of Lead | Chart Electoral Votes

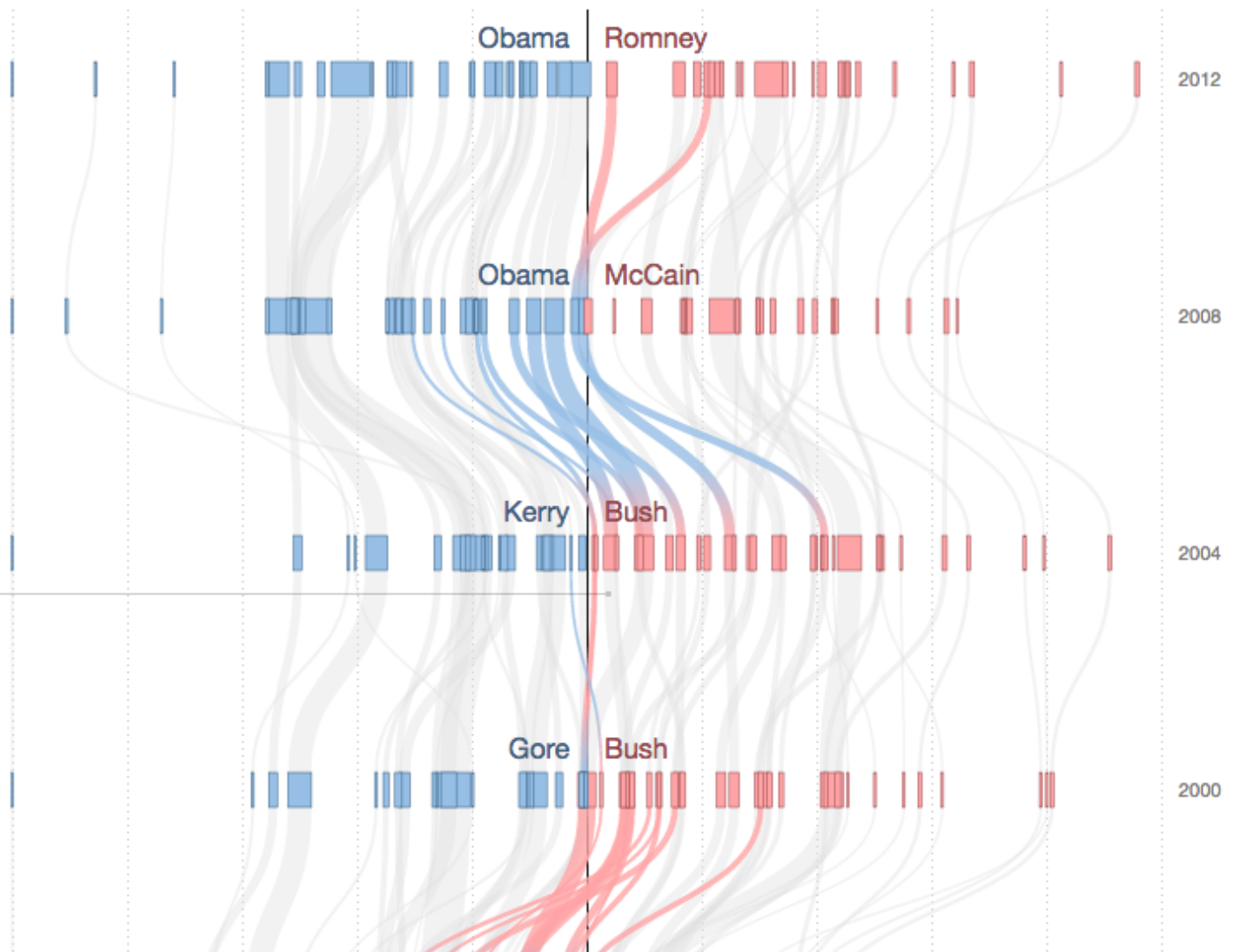
← MORE DEMOCRATIC | MORE REPUBLICAN →  
≥50% +40% +30% +20% +10% | +10% +20% +30% +40% ≥50%

**Obama Re-elected**  
The country voted about 5 percentage points more Republican in 2012 than in 2008. Obama lost North Carolina and Indiana, but won every tossup except Florida, which remains too close to call.

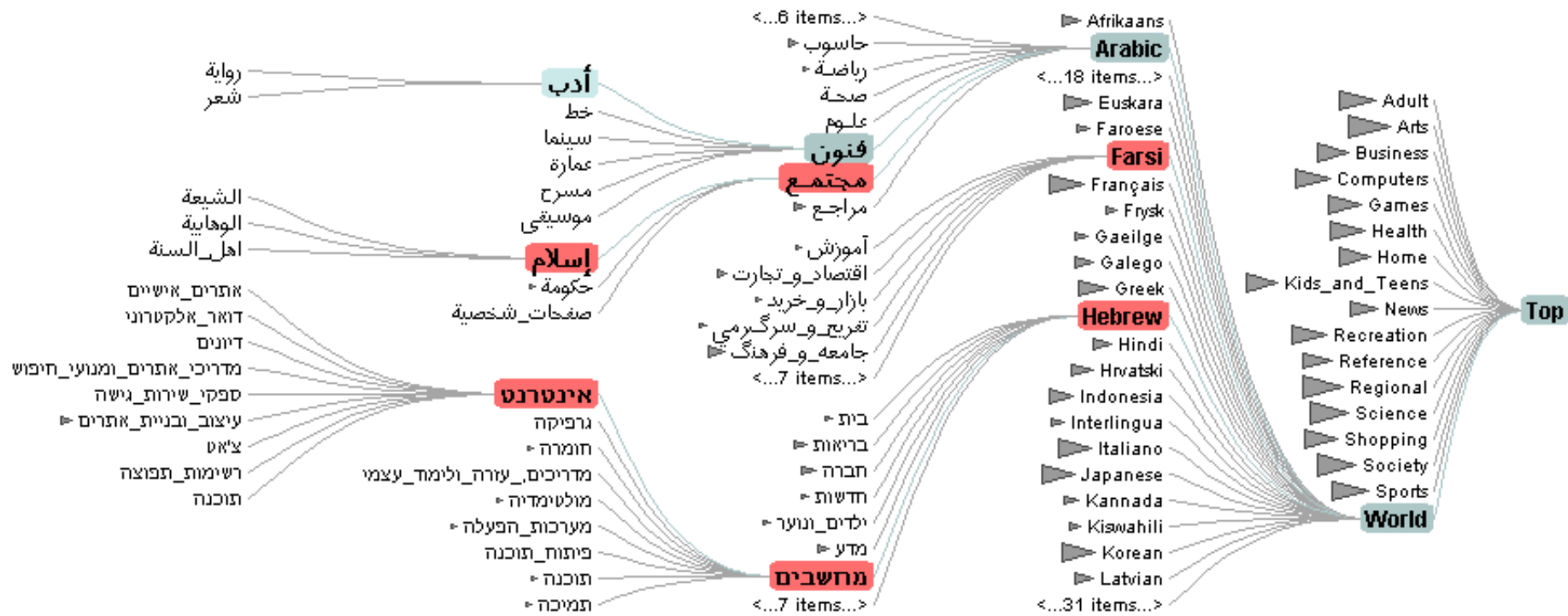
Highlight Tossups

**As Goes Ohio**  
Ohio, which has voted for the winner in every election since 1964, provided the decisive electoral votes in 2004, and it is the state likeliest to play that role again this year, according to the FiveThirtyEight model.

Highlight Ohio

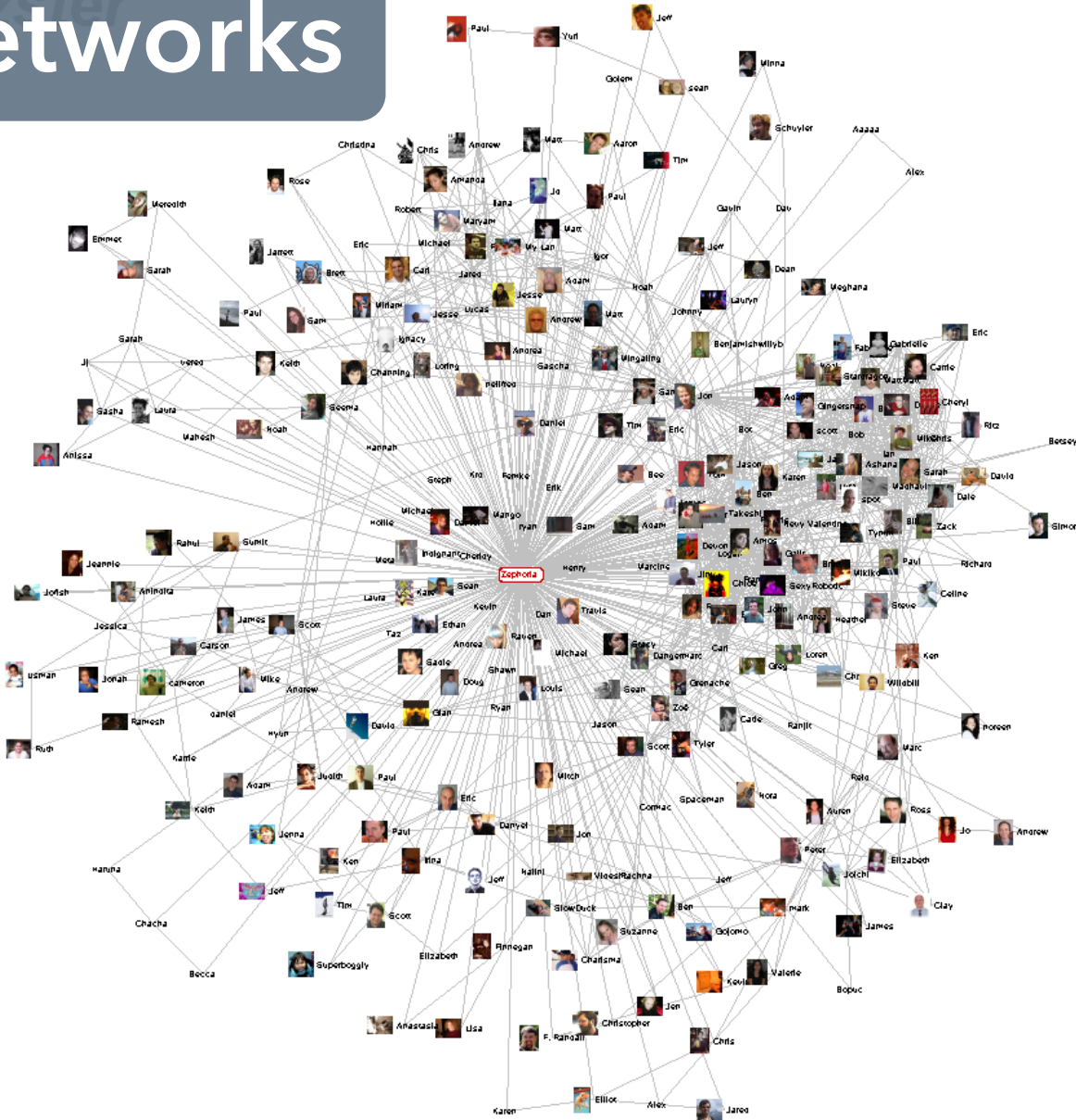


# Hierarchies



Degree-Of-Interest Trees [Heer & Card 04]

# Networks



community >>

Enable

search >>

## Zephoria

User ID	21721
Friends	<input type="checkbox"/> 266
Age	??
Gender	<input type="checkbox"/> Female
Status	<input type="checkbox"/> Single
Location	San Francisco, CA
Hometown	Lancaster, PA
Occupation	researcher: social networks, identity, context
Interests	apophenia, observing people, culture, questioning power, reading, buddhism, ipseity, computer-mediated communication, social networks, technology, anthropology, stomping
Music	psytrance/goa/trance [Infected Mushroom, Son Kite... Iboga/Digital Structures], Ani Difranco, downtempo, Thievery Corporation, Beth Orton, Morcheeba, Ween, White Stripes
Books	Authors: Erving Goffman, Stanley Milgram, Jeanette Winterson, Eric Schlosser, Leslie Feinberg, Dorothy Allison, Italo Calvino, Hermann Hesse
TV Shows	??
Movies	Koyaanisqatsi, Amelie, Waking Life, Tank Girl, The Matrix, Clockwork Orange, American Beauty, Fight Club, Boys Don't Cry
Member Since	??
Last Login	2003-10-21
Last Updated	2003-10-21
About	[Some know me as danah...]

I'm a geek, an activist and an academic, fascinated by people and society. I see life as a very large playground and enjoy exploring its intricacies. I revel in life's chaos, while simultaneously providing my own insane element.

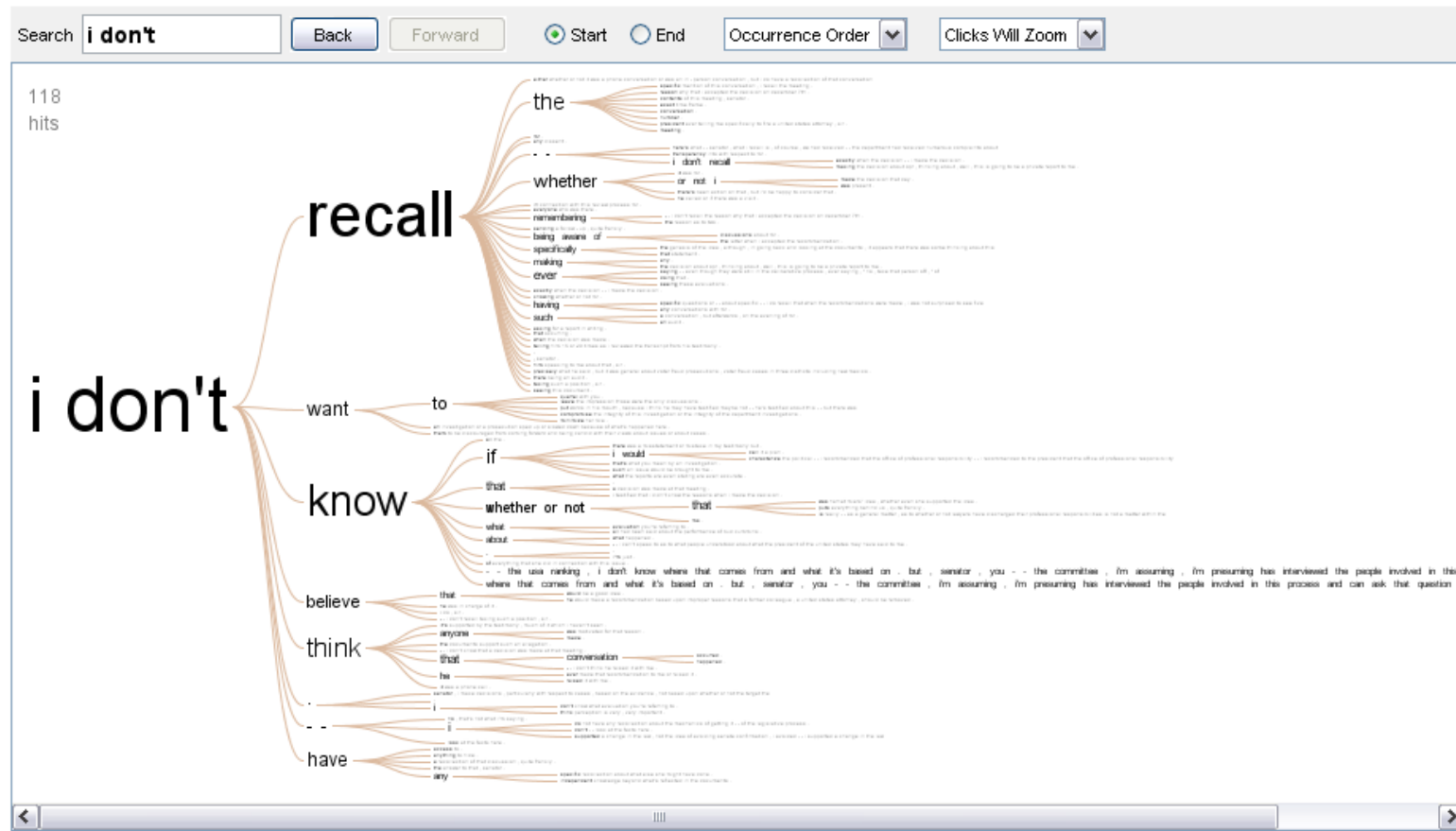
My musings:  
<http://www.zephoria.org/thoughts/>

**Want to Meet** Someone who makes life's complexities seem simply elegant.

# Text

## Visualizations : Word tree / Alberto Gonzales

Creator: Martin Wattenberg  
Tags:



Data file: [Word in testimony from Gonzales, 4/19/2007](#) Data source: CQ Transcript Wire via the Washington Post This data set has not yet been rated

Comments (4)

currently showing

This visualization has 4 positive and 0 negative

- explore**
- visualizations
  - data sets
  - comments
  - topic hubs
- participate**
- create visualization
  - upload data set
  - create topic hub
  - register
- learn more**
- quick start
  - visualization types
  - data format & style
  - about Many Eyes
  - FAQ
  - blog

**contact Us**

- contact
- report a bug

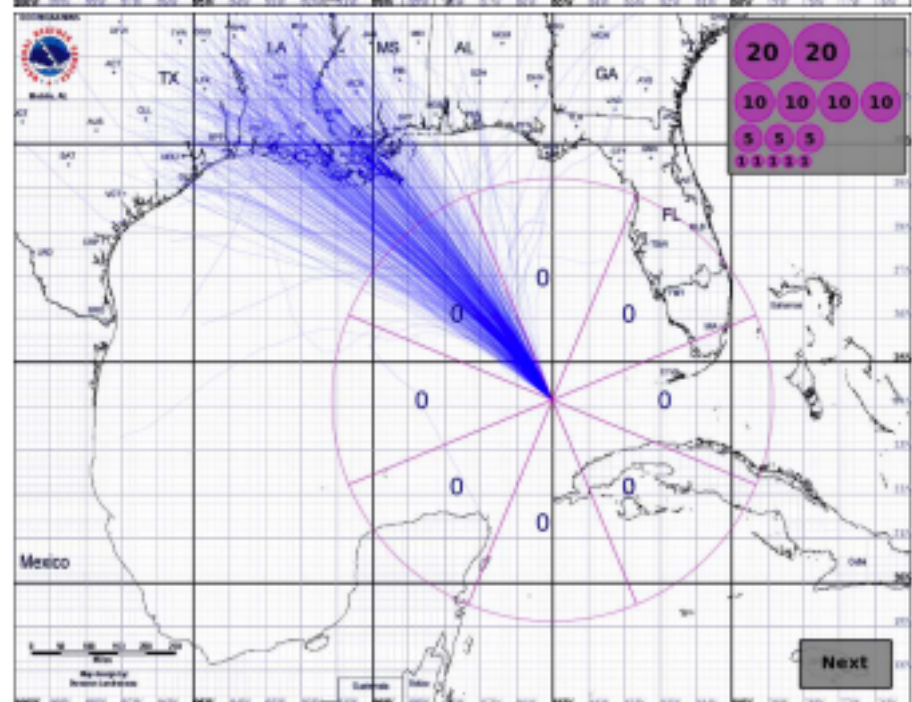
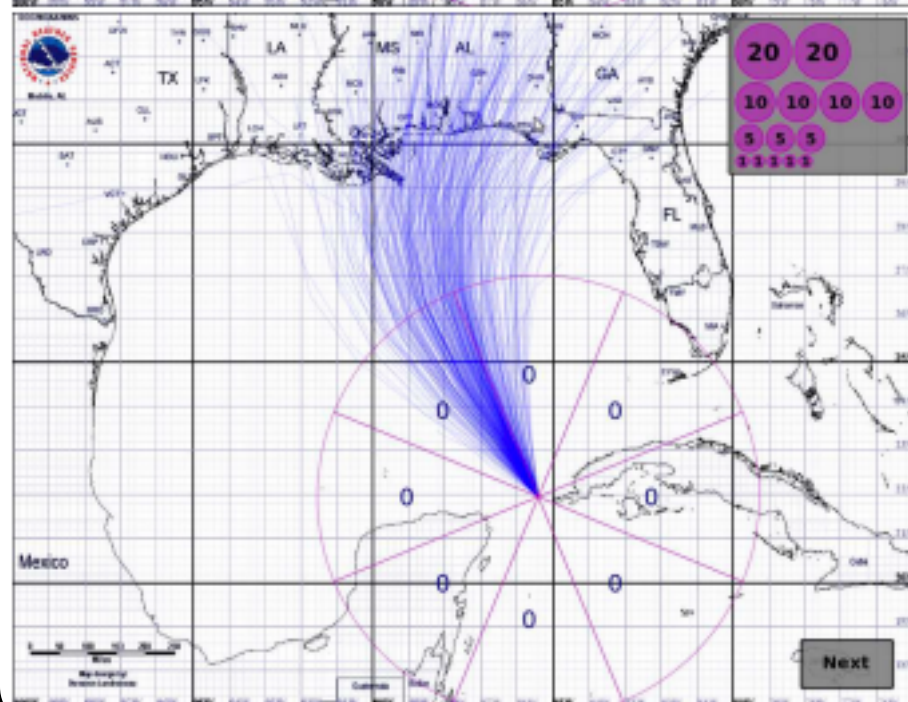
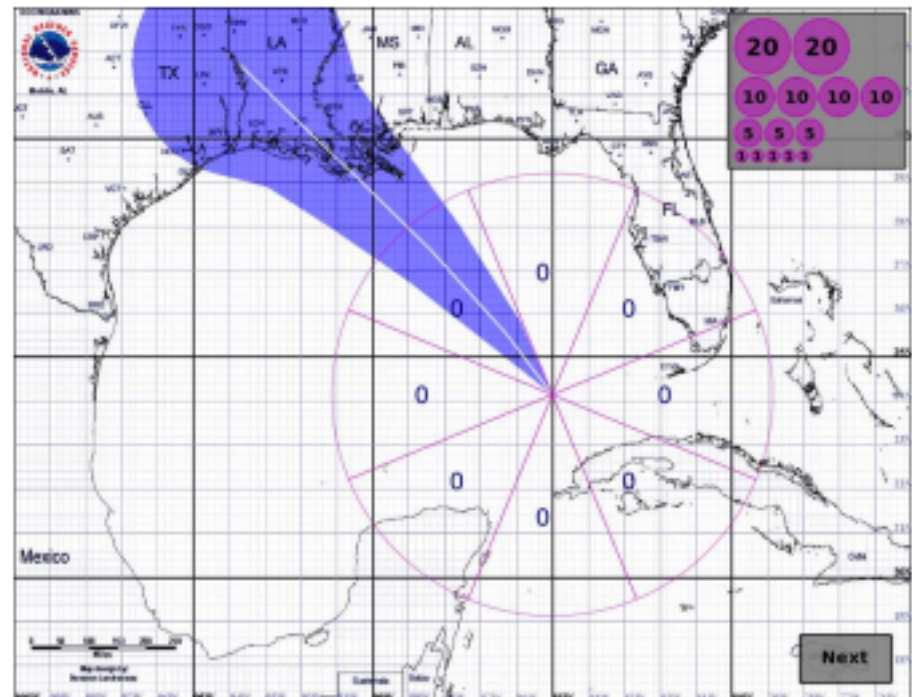
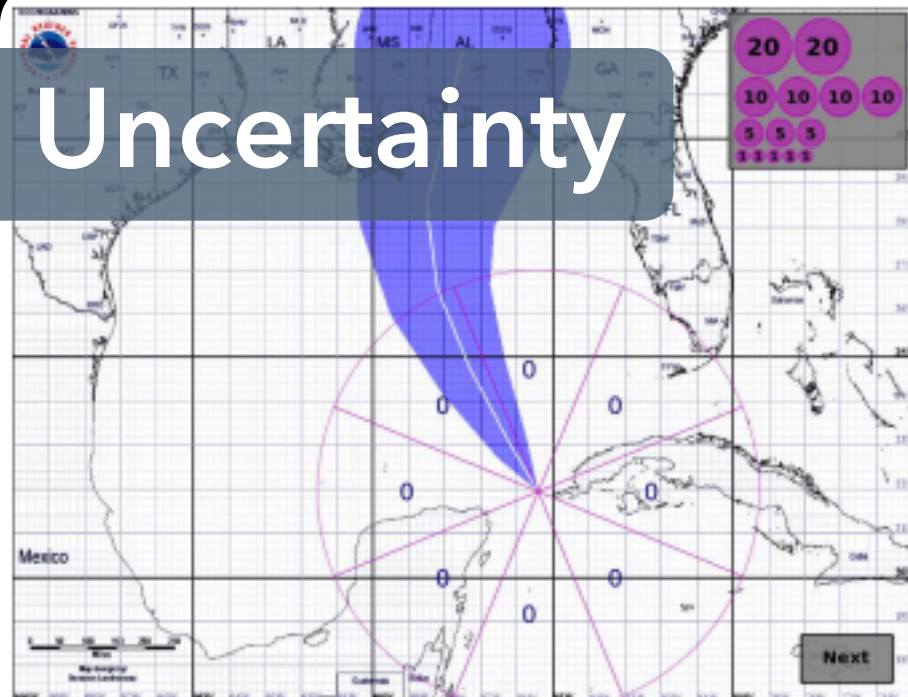
**legal**

- terms of use

**Popular Dataset Tags**

2007 2008 bible blog books census crime education eharmony election energy food health inauguration internet ireland literature lyrics media music network obama people politics population president prices religion social

# Uncertainty





# The Future of Visualization

Where is more work required?

What emerging technologies and societal trends will impact visualization design?

What did you find most difficult in creating visualizations and designing techniques?

# Thank You!

