

Collaborative Visual Analysis

Maneesh Agrawala

Jessica Hullman

CS 294-10: Visualization

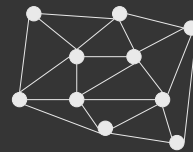
Fall 2014

Last Time: Graph Layout

Graphs and Trees

Graphs

Model relations among data
Nodes and edges

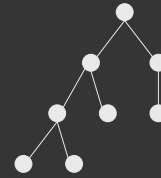


Trees

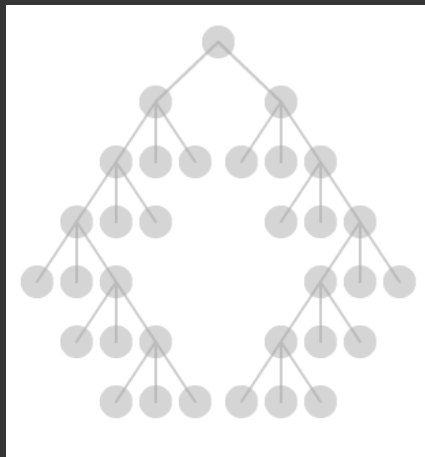
Graphs with hierarchical structure

- Connected graph with $N-1$ edges

Nodes as *parents* and *children*



Reingold & Tilford's Tidier Layout



Goal: maximize density and symmetry.

Originally for binary trees, extended by Walker to cover general case.

This extension was corrected by Buchheim et al to achieve a linear time algorithm.

Force-Directed Layout

Edges = springs

$$F = -k * (x - L)$$

Nodes = charged particles

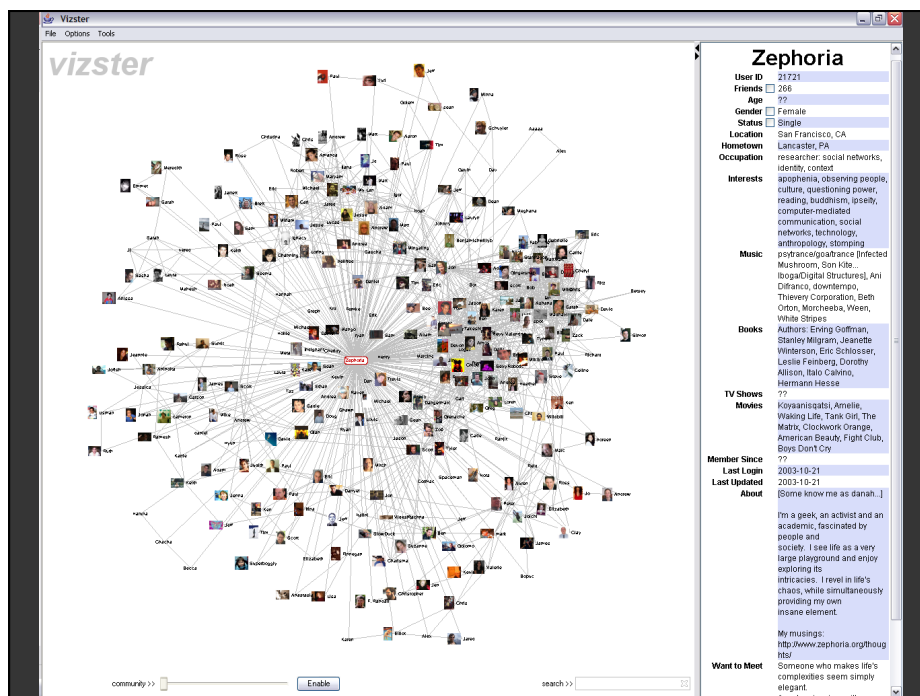
$$F = G * m_1 * m_2 / x^2$$

Repeatedly calculate forces, update node positions

Naïve approach $O(N^2)$

Speed up to $O(N \log N)$ using quadtree or k-d tree

Numerical integration of forces at each time step



Constrained Optimization Layout

Minimize stress function

$$\text{stress}(\mathbf{X}) = \sum_{i < j} w_{ij} (\|\mathbf{X}_i - \mathbf{X}_j\| - d_{ij})^2$$

- \mathbf{X} : node positions, d : optimal edge length,
- w : normalization constants
- Use global (*majorization*) or localized (*gradient descent*) optimization

→ Says: Try to place nodes d_{ij} apart

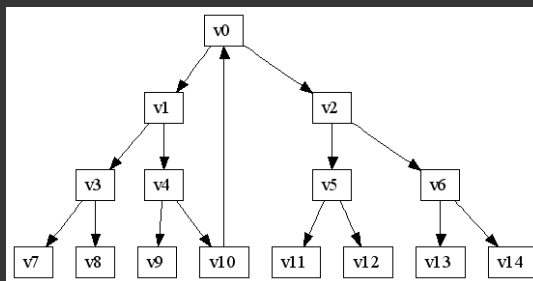
Add hierarchy ordering constraints

$$E_H(\mathbf{y}) = \sum_{(i,j) \in E} (y_i - y_j - \delta_{ij})^2$$

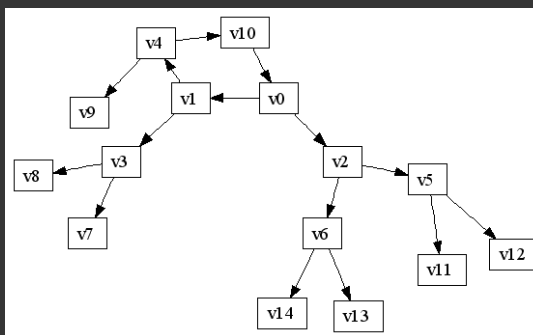
- y : node y-coordinates
- δ : edge direction (e.g., 1 for $i \rightarrow j$, 0 for undirected)

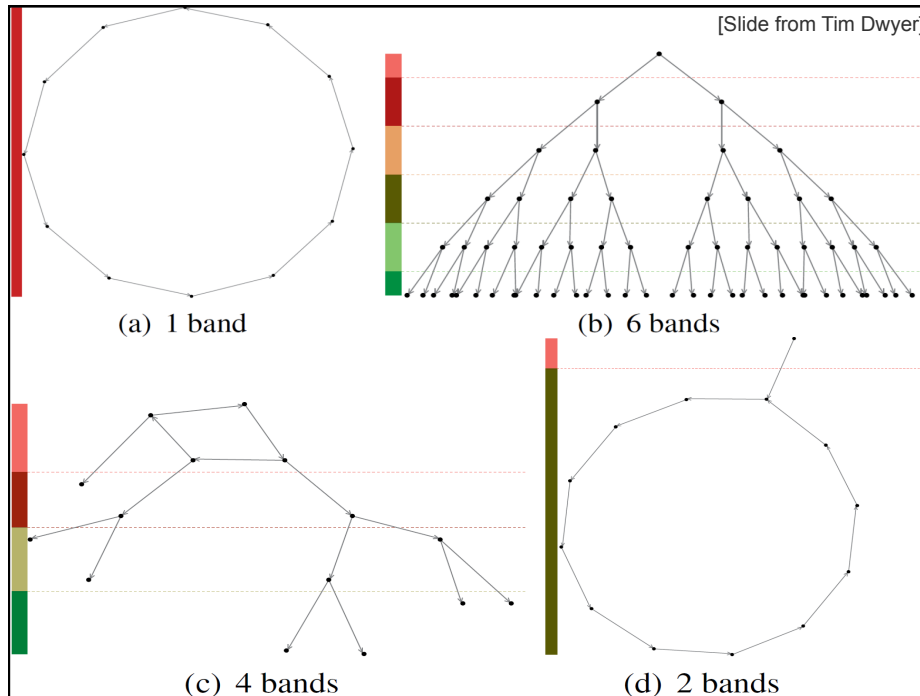
→ Says: If i points to j , it should have a lower y-value

Sugiyama layout (dot)
Preserve tree structure



DiG-CoLa method
Preserve edge lengths





Attribute-Driven Layout

Large node-link diagrams get messy!
Is there additional structure we can exploit?

Idea: Use data attributes to perform layout

- e.g., scatter plot based on node values

Dynamic queries and/or brushing can be used to explore connectivity

Attribute-Driven Layout

The "Skitter" Layout

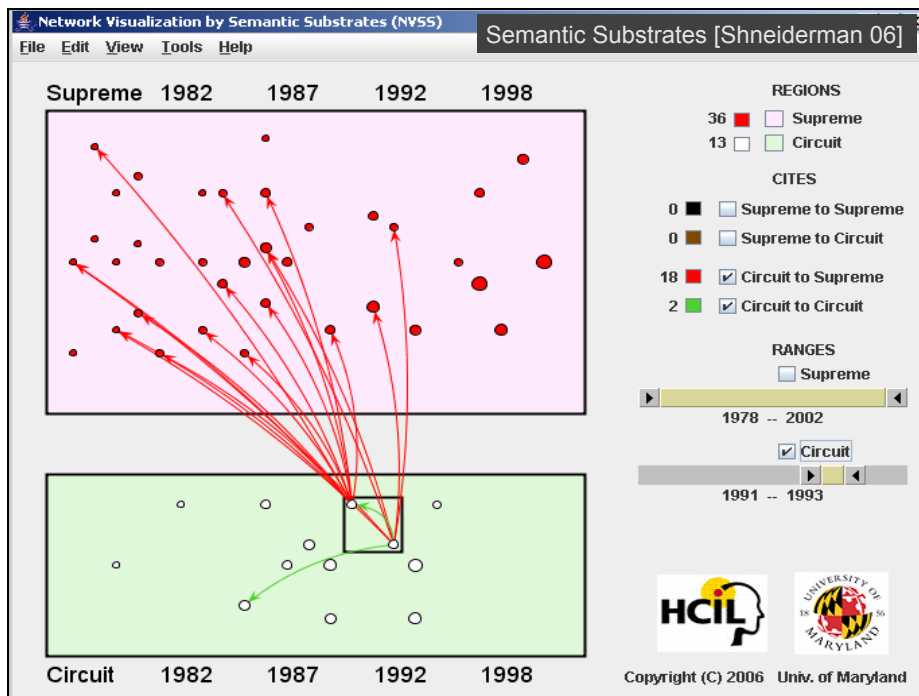
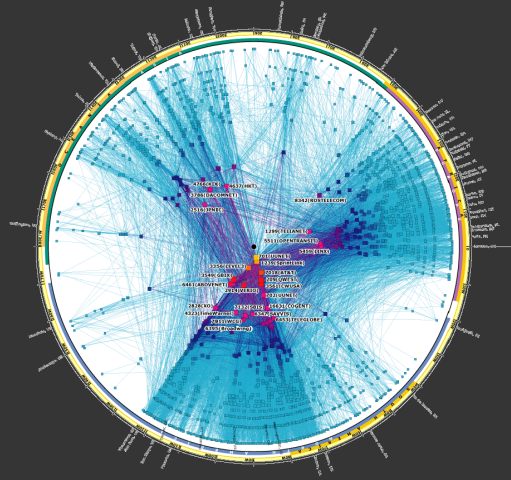
- Internet Connectivity
- Radial Scatterplot

Angle = Longitude

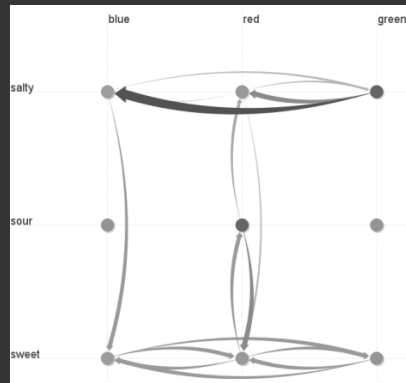
- Geography

Radius = Degree

- # of connections
- (a statistic of the nodes)

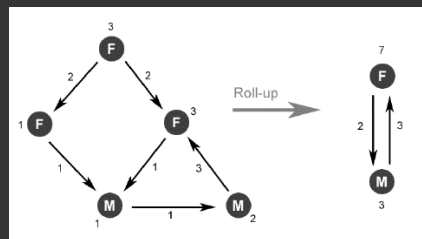


PivotGraph [Wattenberg 2006]



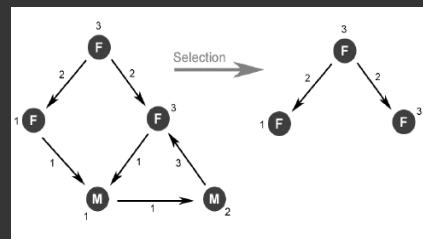
Layout aggregated graphs according to node attributes

Operators



Roll-Up

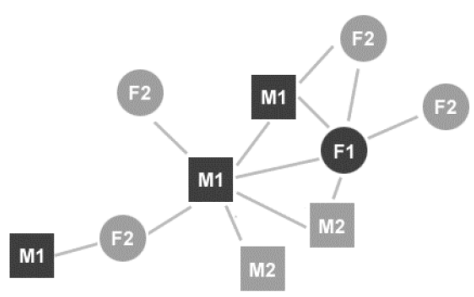
Aggregate items with matching data values



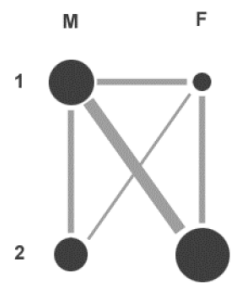
Selection

Filter on data values

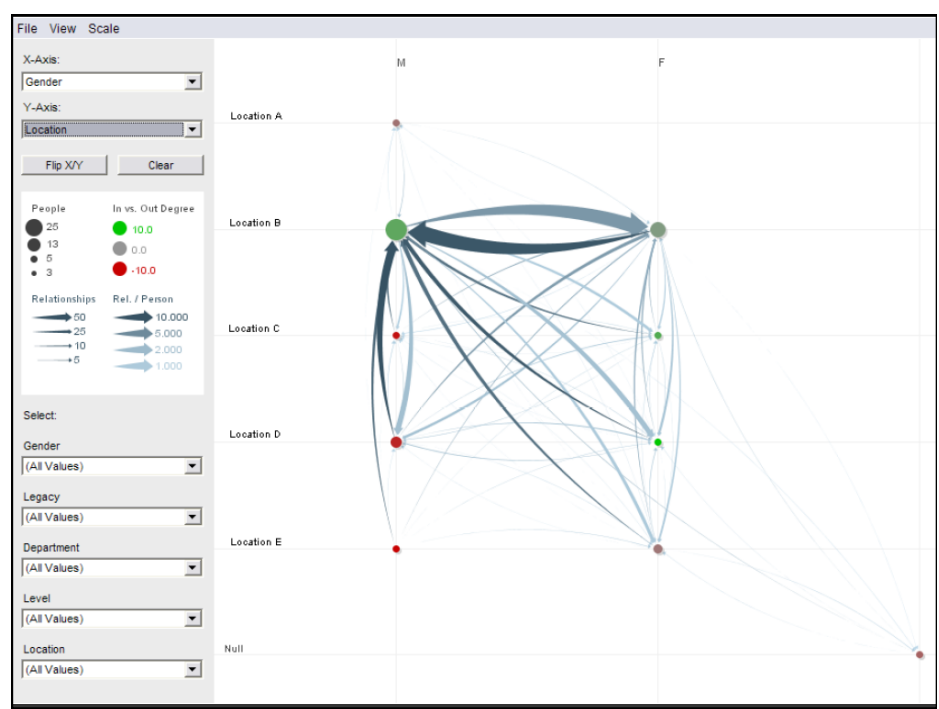
PivotGraph

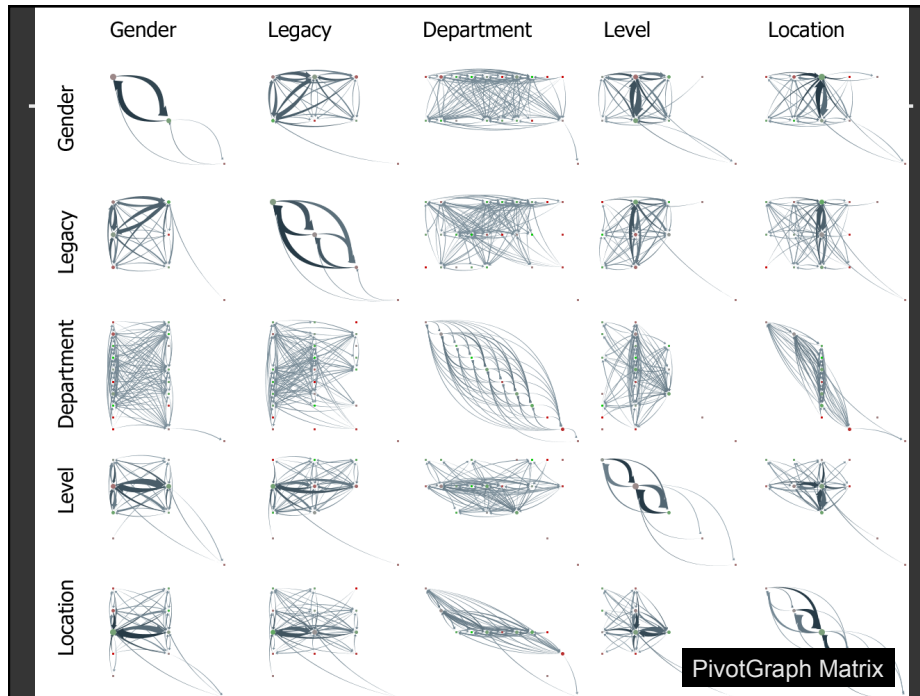


Node and Link Diagram



PivotGraph Roll-up



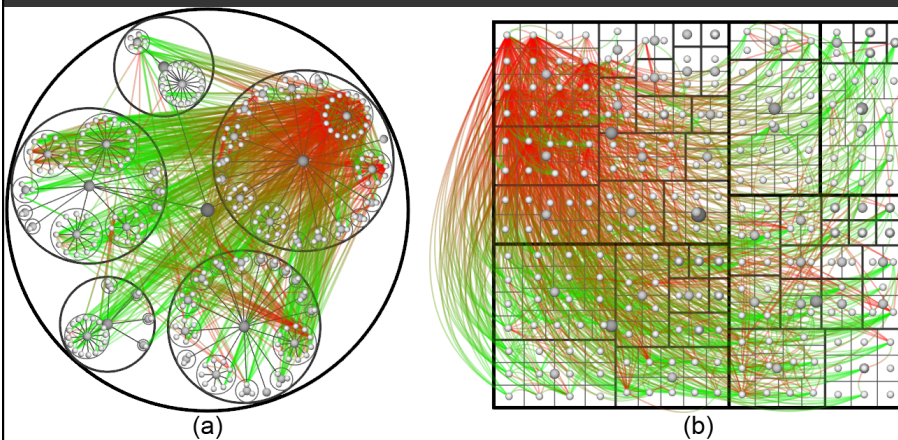


Limitations of PivotGraph

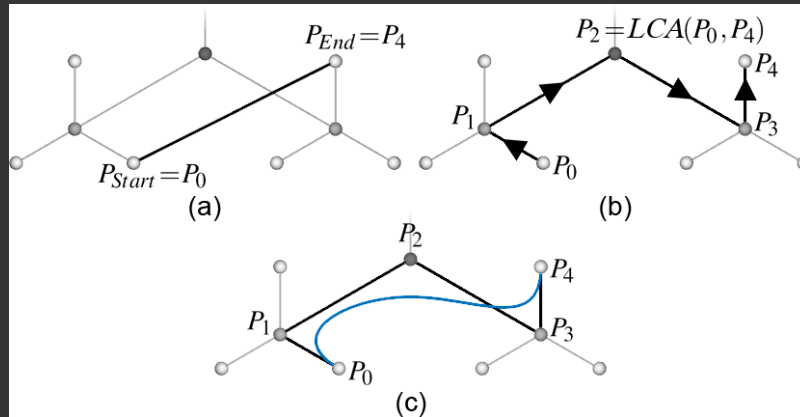
- Only 2 variables (no nesting as in Tableau)
- Doesn't support continuous variables
- Multivariate edges?

Hierarchical Edge Bundles

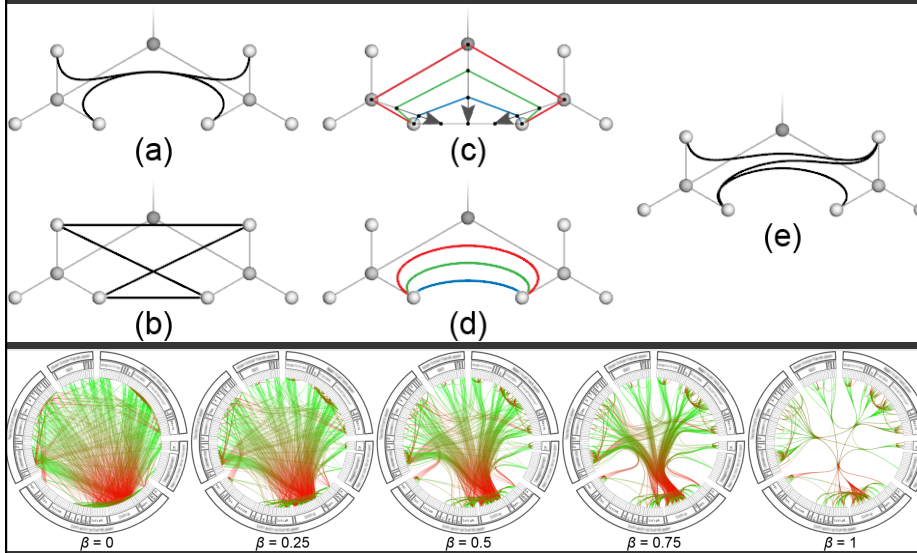
Trees with Adjacency Relations

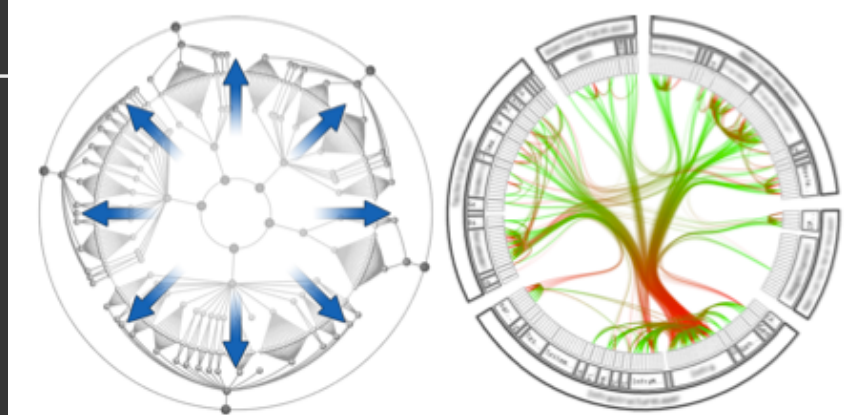


Bundle Edges along Hierarchy

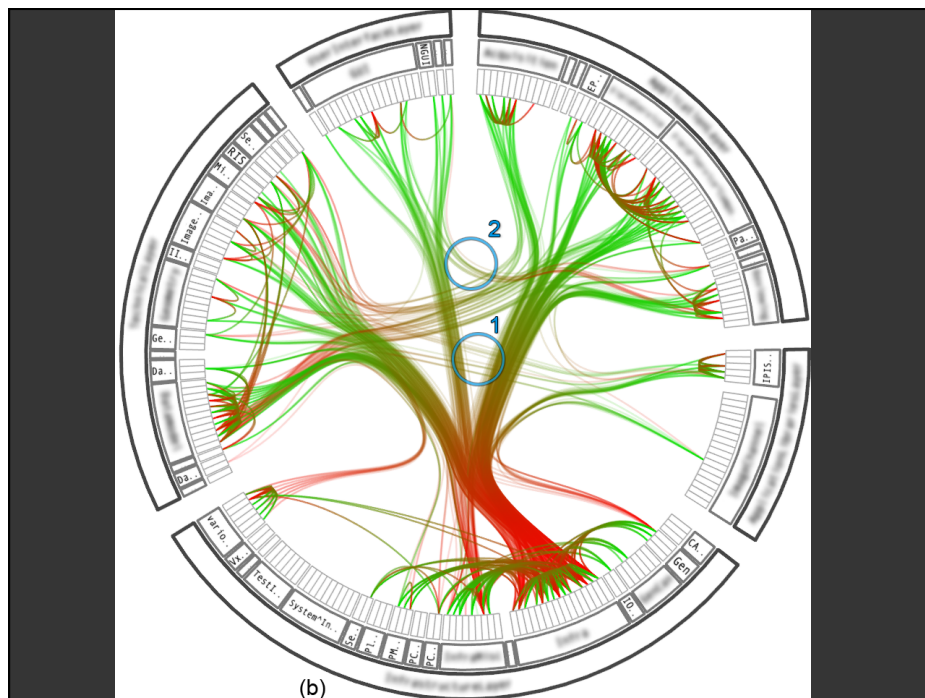


Configuring Edge Tension





Use radial tree layout for inner circle
 Mirror to outside
 Replace inner tree with hierarchical edge bundles



Summary



Tree Layout

Indented / Node-Link / Enclosure / Layers

How to address issues of scale?

- Filtering and Focus + Context techniques

Graph Layout

Tree layout over spanning tree

Hierarchical “Sugiyama” Layout

Optimization (Force-Directed Layout)

Attribute-Driven Layout

Announcements

Final project

Design new visualization method

- Pose problem, Implement creative solution

Deliverables

- Implementation of solution
- 8-12 page paper in format of conference paper submission
- 1 or 2 design discussion presentations

Schedule

- Project proposal: 10/27
- Project presentation: 11/10, 11/12
- Final paper and presentation: TBD, likely 12/1-12/5

Grading

- Groups of up to 3 people, graded individually
- Clearly report responsibilities of each member

In-Class Project Presentations

Dates: 11/10 and 11/12

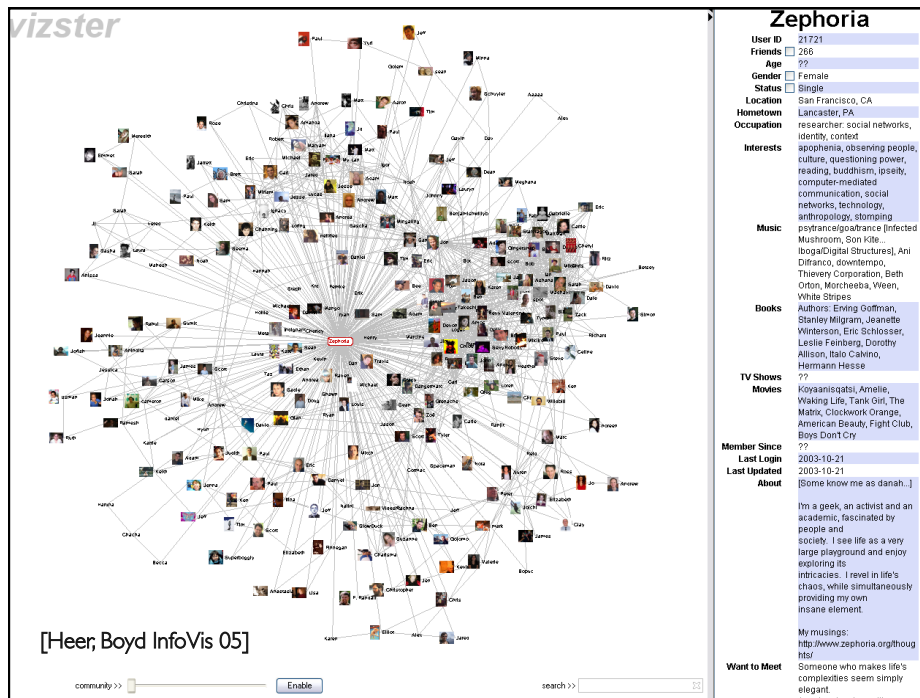
- Description of problem you are addressing
- Survey of related work
- Description/storyboard and demo of approach
- A list of milestones for finishing the project by the deadline

Scheduling

- Send me dates you cannot attend class by tomorrow
- Next class We will ask for volunteers to present on each day

Collaborative Visual Analysis

A Tale of Two Visualizations

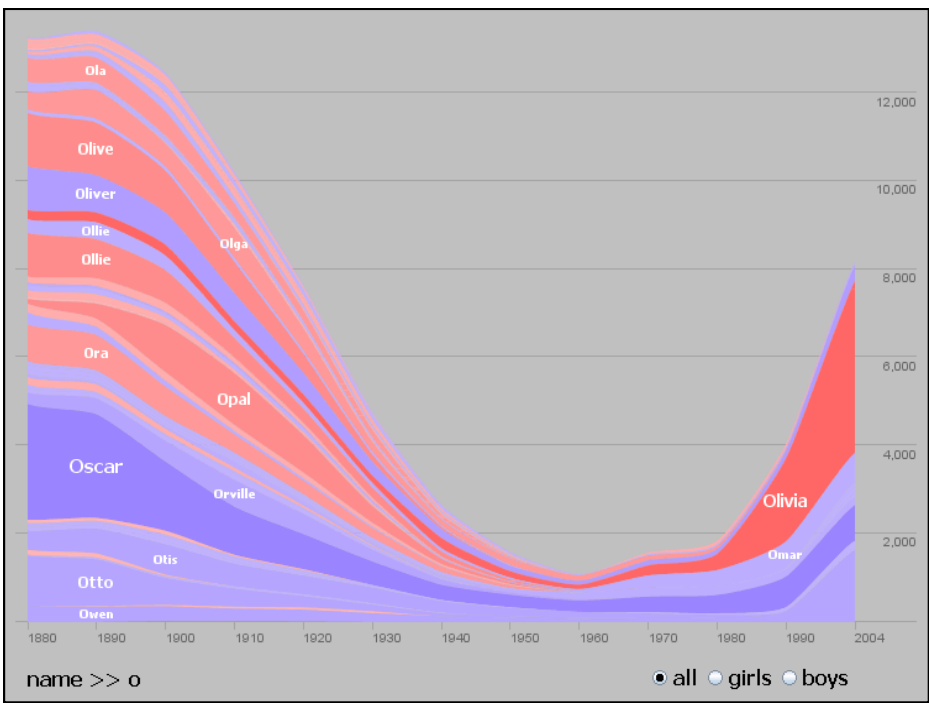
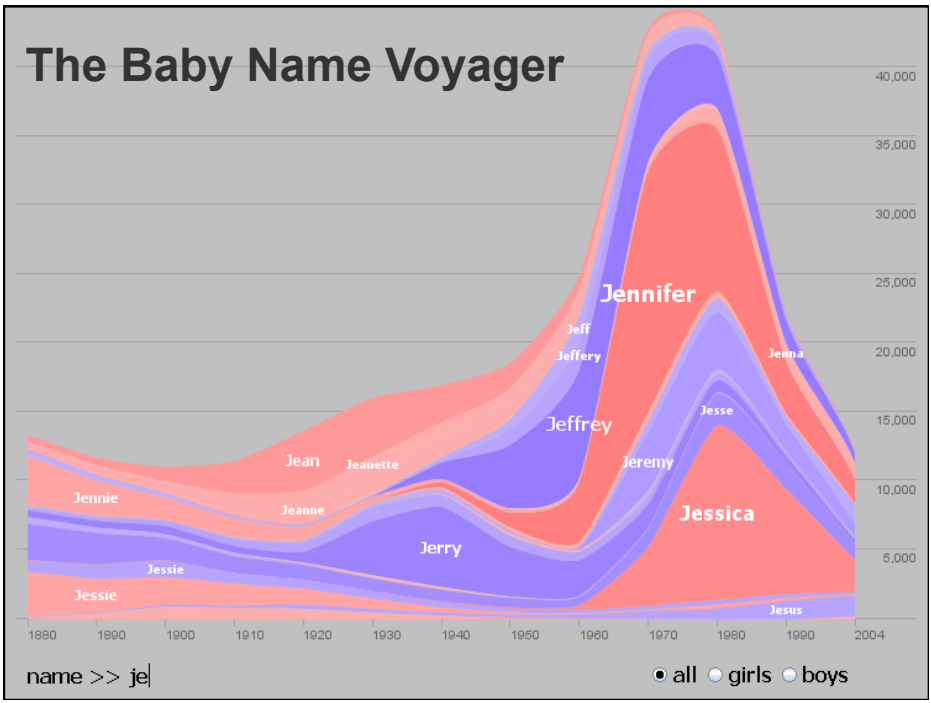


Observations

Groups spent more time in front of the visualization than individuals

Friends encouraged each other to unearth relationships, probe community boundaries, and challenge reported information

Social play resulted in informal analysis, often driven by story-telling of group histories



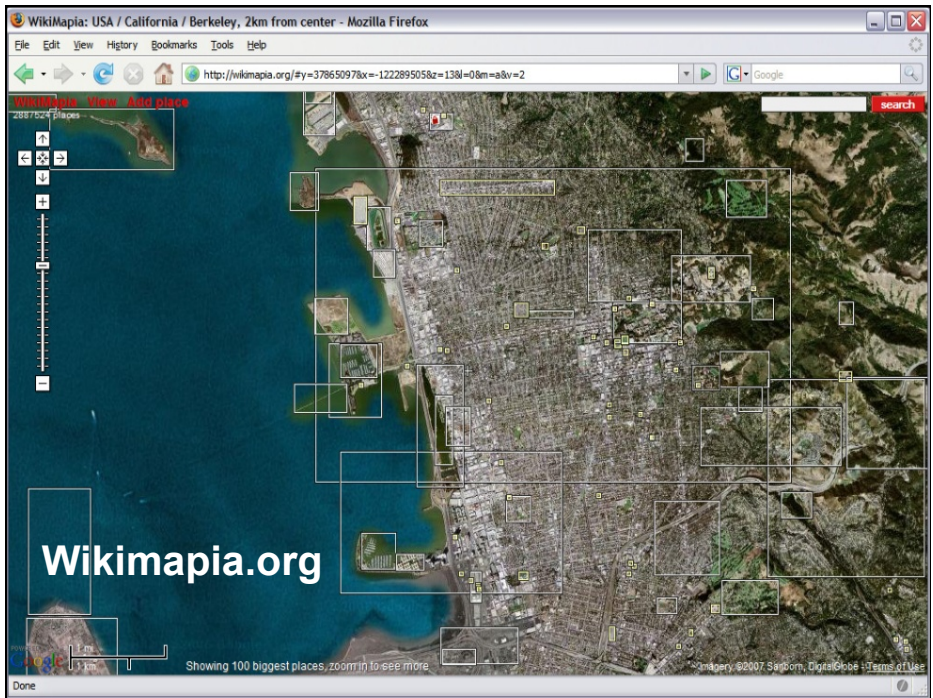
Social Data Analysis

Visual sensemaking can be **social** as well as cognitive

Analysis of data coupled with **social interpretation and deliberation**

How can user interfaces catalyze and support **collaborative visual analysis**?

Social Data Analysis



Spotfire Decision Site Posters

www-958.ibm.com/software/analytics/manyeyes/

Many Eyes Log in

Explore

- Visualizations
- Data sets
- Comments
- Topic centers

Participate


- Create a visualization
- Upload a data set
- Create a topic center
- Register

Learn more

- Quick start
- Visualization types
- About Many Eyes
- Privacy
- Blog


Try our featured visualizations

Top 10 Oil Producing Nations



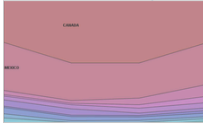
Barrels of crude per day
by evannk

Select Crime Stats in Lafayette, LA



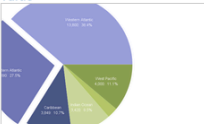
1999 to 2011
by rxr

US Brewed Malted Beer Exports




2008 to 2012
by crossborders1

Global Distribution of Leatherback Turtle



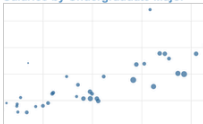
An internationally endangered species.
by Anonymous

Cosmetic Procedures



2012
by KatieCreative

Salaries by Undergraduate Major



Starting and mid-career salaries
by dsibx

Many-Eyes

[Viégas, Wattenberg, et al. CHI 07]

Try out the newest version of IBM Many Eyes! Try ManyEyes v2 >

Tableau Public

www.tableausoftware.com/public/gallery/taleof100

tableau public

SEARCH

HOW IT WORKS
see the magic happen

GALLERY
see what others visualized

COMMUNITY
join the conversation

DOWNLOAD
get the data

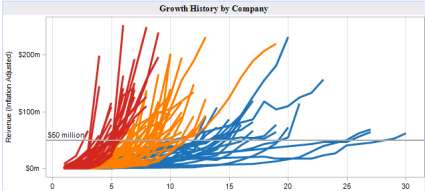
Tale of 100

How fast do successful tech companies grow? The Wall Street Journal posted this visualization that compares the performance of 100 of the largest publicly traded software companies across more than three decades.

Tale of 100 Entrepreneurs

Click to interact

■ Rocket Ship
 ■ Hot Company
 ■ Slow Burner



Growth History by Company

GALLERY

- Business and Real Estate
- Government and Public Data
- Health and Science
- Sports
- Travel and Lifestyle
- All

Vizes in the wild showcase

View Gallery

Tableau Public makes data beautiful and extremely useful. I was amazed by the interactivity, incredibly useful for my readers.

Stephen McDaniel
Freshalytics

20

Where can Collaborators Contribute?

Data Management

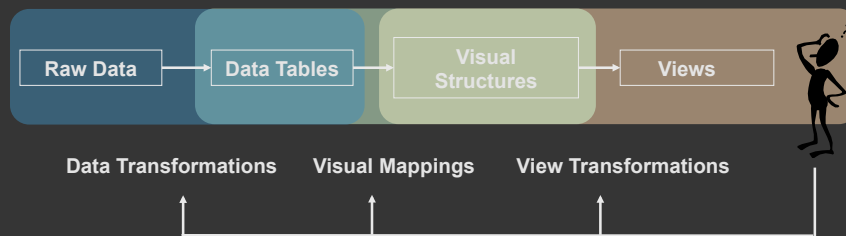
Contribute Data
Clean Data
Categorize Data
Moderate Data
Create Metadata

Visualization

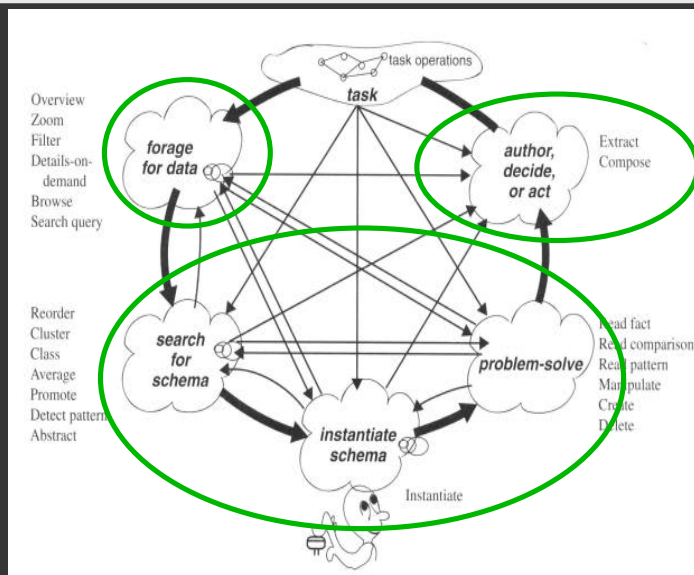
Select Data Sources
Apply Visual Encoding
Author Software

Visual Analytics

Observations
Hypotheses
Evidence (+/-)
Summarize
Report / Presentation



Sensemaking [Card, Mackinlay, & Schneiderman]



Design Considerations

[Heer & Agrawala VAST 07, IVS 08]

- 1. Division, allocation, and integration of work**
- 2. Common ground and awareness**
- 3. Reference and deixis (pointing)**
- 4. Identity, trust, and reputation**
- 5. Group formation and management**
- 6. Incentives and engagement**
- 8. Presentation and decision-making**

**Sense.us: Collaborative
Visualization of Demographic Data**

sense.us > social data visualization - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://sense.us/

sense.us Log in to add comments! IBM e-mail pass login help

Dogear Tags
Dogear Tag Usage, May 2005 to August 2006
source: IBM Dogear
10 comments

Dogear People
Dogear Bookmarking by Person, May 2005 to August 2006
source: IBM Dogear
1 comment

Job Voyager
Reported Occupations of U.S. Labor Force, 1850-2000
source: http://pums.org
139 comments

Birthplace Voyager
Reported Birthplace of U.S. Residents, 1850-2000
source: http://pums.org
10 comments

sense.us - social data visualization
sense.us is a prototype system for collaborative visualization.
- See the data. See what people have to say about it.
- Dive into the data and share your explorations.

The site requires Java 1.5+ and either Firefox or Internet Explorer.
Use your IBM W3/BluePages e-mail and password to login.
Use at least 1024x768 resolution for the best experience.

Check out the user's guide and privacy policy before getting started.

Having problems using Firefox with Java 1.4? Some users using Firefox and Java 1.4 have found that comments aren't loading properly. If you run into this problem, consider upgrading to Java 1.5 (Windows, Linux) or using Internet Explorer (on Windows systems). Sorry for any inconvenience!

U.S. Census State Map
State Map of 2000-2005 Census Data
source: U.S. Census Bureau
16 comments

Population Pyramid
U.S. Population Demographics, 1850-2000
source: http://pums.org
7 comments

Done

[Heer, Viégas, Wattenberg CHI 07]

Exploratory Design Rationale

Sharing within visualization and across the web

http://sense.us/jobs#scale=1&gender=All&query=mili

Reported Occupations of U.S. Labor Force, 1850-2000 (source: http://pums.org)

>> mili all men women % of Work Force

Annotations: Military, Military

comments (5) New Comment | View All (143)

▼ Is this military info right?

I would have expected a different pattern for the military, but then again maybe this is just the military industrial complex growing and growing.
by **Martin Sharp** on Fri Jul 21, 2006 10:15 AM

here are labels where I would have expected big jumps.
by **Martin Sharp** on Fri Jul 21, 2006 10:16 AM

well, there was also the cold war right after ww2, which might be part of the reason why there's such a huge jump after the 40s. It is also interesting that there is such a dip between the 70s and the 80s.
by **Julia Hernandez** on Fri Jul 21, 2006 11:01 AM

I guess a lot of it has turned to robots, and the industrial complex, as martin suggested, though it would be interesting to see the comparison of the fall in military personnel next to the rise in DOD funding for robots and industry.
by **Jesse O'Brien** on Fri Jul 21, 2006 11:51 AM

I think the jumps have more to do with the economy at large rather than any particular military conflict. Lots of money in conflict has already been spent before the conflict starts.
by **Fred Klein** on Wed Aug 2, 2006 10:24 AM

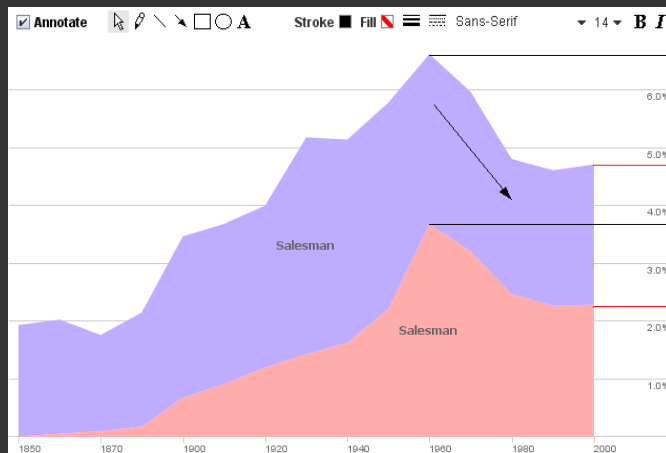
reply

bookmarks (1) Add View | Save | Load | View All

Exploratory Design Rationale

Sharing within visualization and across the web

Pointing at interesting trends, outliers

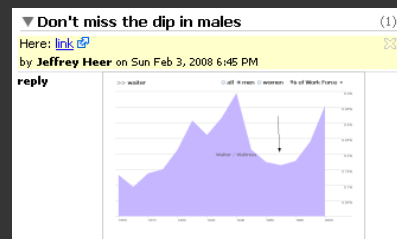
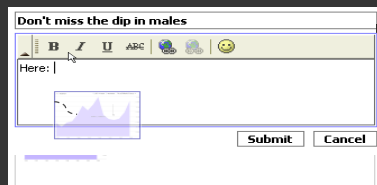


Exploratory Design Rationale

Sharing within visualization and across the web

Pointing at interesting trends, outliers

Collecting and linking related views



Exploratory Design Rationale

Sharing within visualization and across the web

Pointing at interesting trends, outliers

Collecting and linking related views

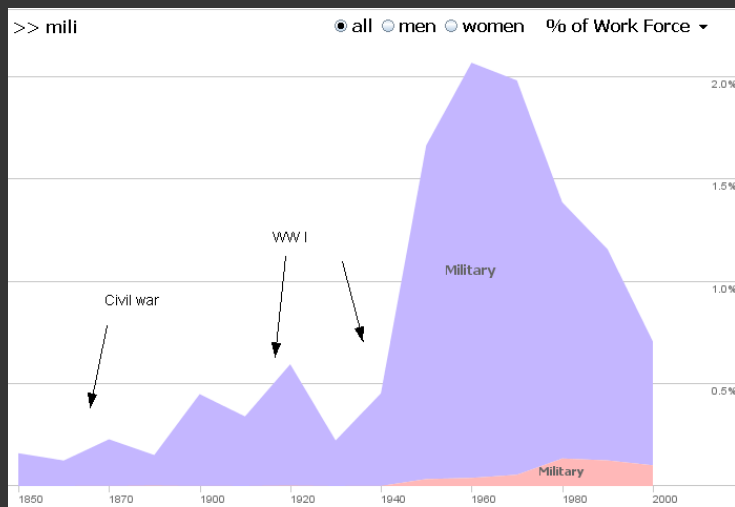
Awareness of social activity

Ushers?
Is this like the movie theatre kind of usher?
by **Martin Sharp** on Fri Jul 21, 2006 10:24 AM

Drop in household workers
This trend is almost the opposite of the "clerical worker" and "secretary" categories. It's amazing how much it drops.
by **Julia Hernandez** on Fri Jul 21, 2006 10:50 AM

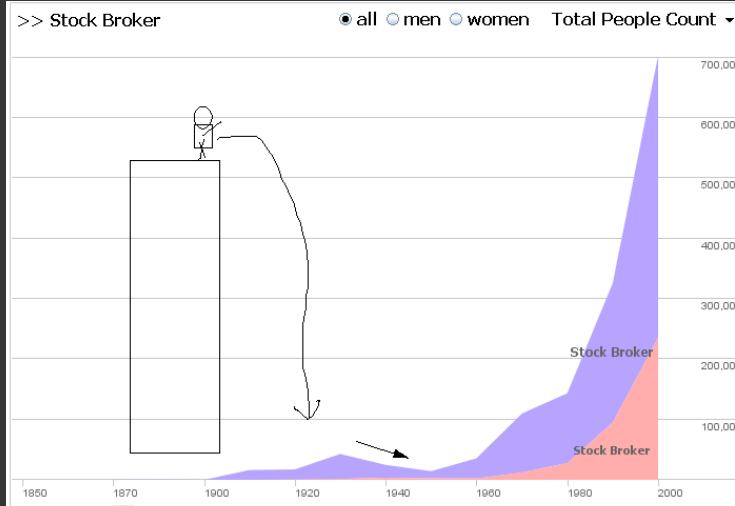
Disappearing professors???
Note how a bunch of different kinds of professors "disappear" in the year 2000, whereas the "professor" category take off.
Is this just a change in the way the census aggregates data?
by **Julia Hernandez** on Fri Jul 21, 2006 11:07 AM

Building Off of Others



Data Jokes

Great depression "killed" a lot of brokers



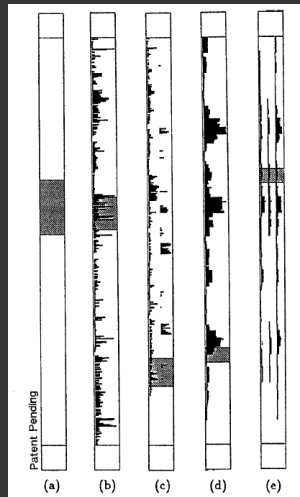
Sharing in External Media

Designing for Social Data Analysis

Social Data Analysis

How can users' activity traces be used to improve awareness in collaborative analysis?

Social Navigation



Read & Edit Wear, Hill et al 1992

del.icio.us / jheer /

your bookmarks | your network | subscriptions | links for you | post

All your items (189)

< earlier | later > page 1 of 19

2005 MicroGraph Winners [edit / delete](#)
to visualization photography science ... [saved by 5 other people](#) ... on sept 20

NEJM -- The Spread of Obesity in a Large Social Network over 32 Years [edit / delete](#)
to health networks research science social visualization psychology ... [saved by 38 other people](#) ... on sept 06

PetersProjection-over gif (GIF Image, 1203x600 pixels) [edit / delete](#)
to maps visualization ... [saved by 3 other people](#) ... on aug 30

Forbes Lists, World's Richest People, and more - Forbes.com [edit / delete](#)
to business data economics reference statistics ... [saved by 202 other people](#) ... on aug 25

Well-formed data | Elastic lists | Nobel prize winners demo [edit / delete](#)
to visualization vizgallery infovis flash sparklines design ... [saved by 112 other people](#) ... on aug 24

Data Visualization: Modern Approaches | Graphics [edit / delete](#)
to visualization web community infovis vizgallery design ... [saved by 10 other people](#) ... on aug 24

Christian Cantrell: ActionScript PNG and JPEG Encoders Updated [edit / delete](#)
... [saved by 14 other people](#) ... on aug 24

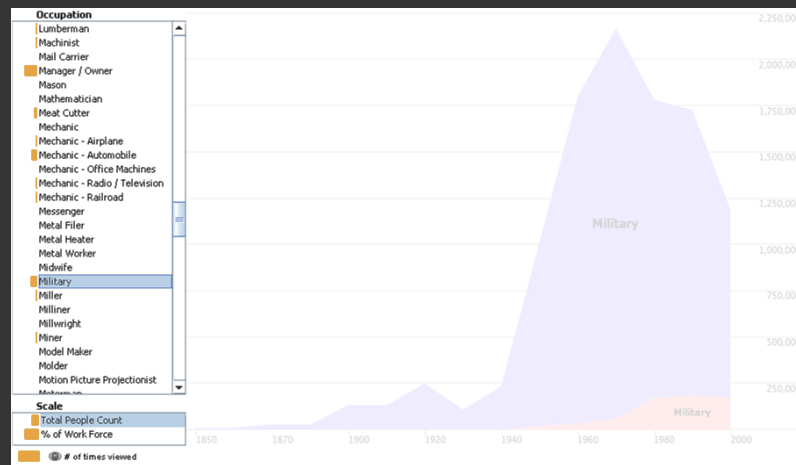
Acrobots [edit / delete](#)
to flash fun interaction animation ... [saved by 61 other people](#) ... on aug 23

Wikinvest | The Investing Wiki with Research about Companies, Investment Concepts, and more [edit / delete](#)
to collaboration web wiki reference finance economics business analysis visualization ... [saved by 10 other people](#) ... on aug 24

Los Angeles Times Homicide Map [edit / delete](#)
to maps statistics visualization infovis interaction news data ... [saved by 62 other people](#) ... on aug 14

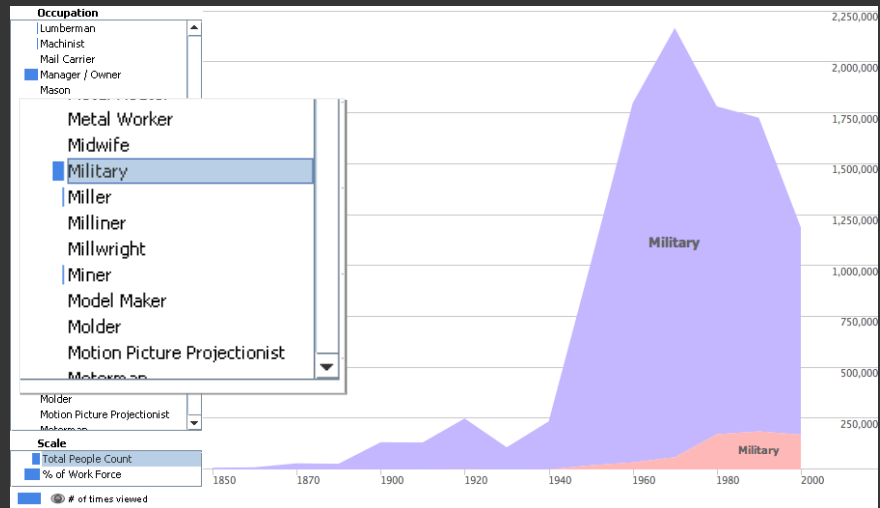
< earlier | later > page 1 of 19

Scented Widgets [InfoVis 07]

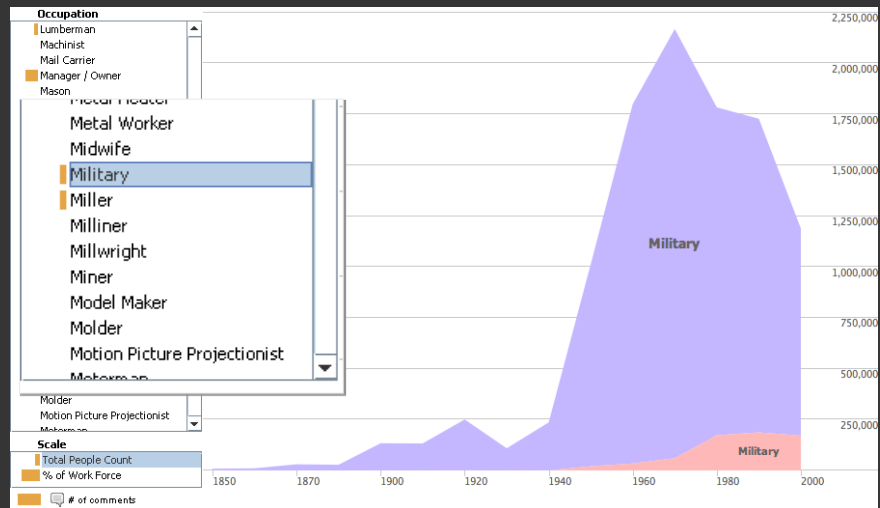


Visual navigation cues embedded in interface widgets

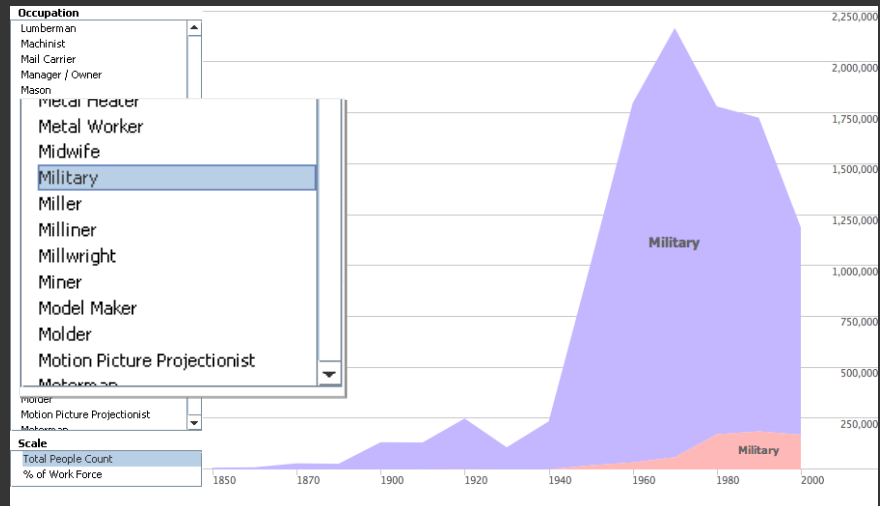
Visitation counts



Comment counts



No scent (baseline)



Do activity cues affect usage?

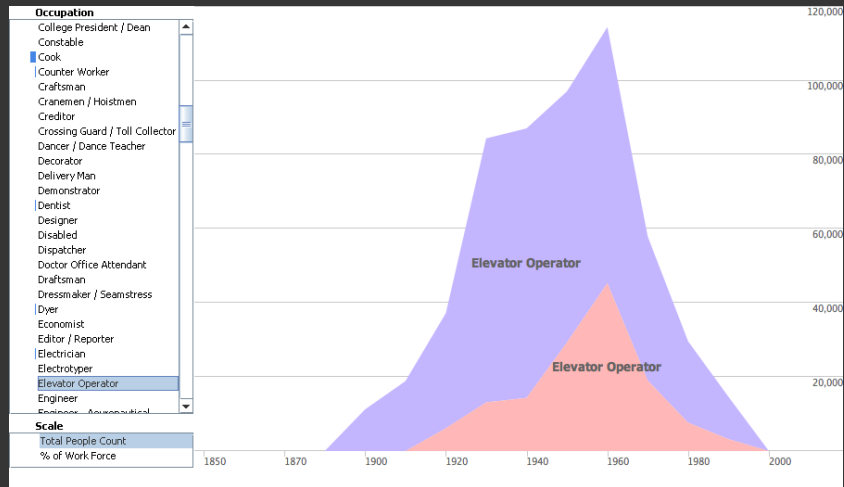
Hypotheses: With activity cues, subjects will

1. Exhibit more revisitation of popular views
2. Make more unique observations

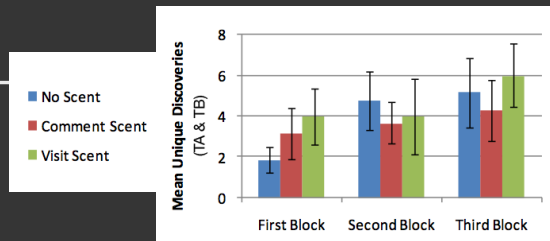
Controlled experiment with 28 subjects
Collect evidence for and against an assertion

Varied scent cues (3) and foraging task (3)

“Technology is costing jobs by making occupations obsolete.”



Results



Unique Discoveries

Visit scent had sig. higher rate of discoveries in first block
 Less reliance on scent when subjects were familiar with data and visualization

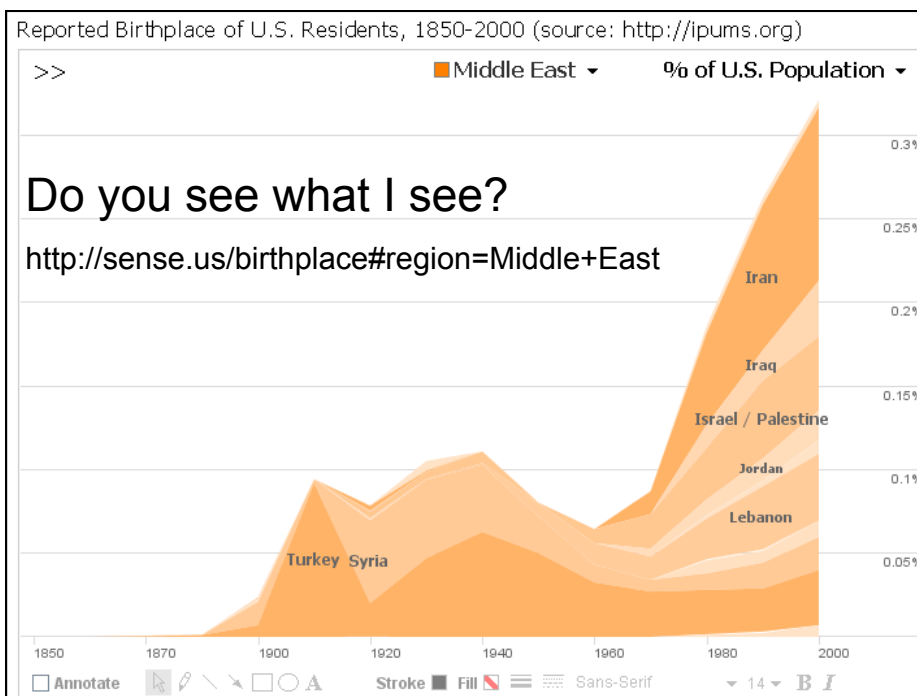
Revisitation

Visit and comment scent conditions correlate more highly with seed usage than no scent

Social Data Analysis

How can users' activity traces be used to improve collaborative analysis?

How should annotation techniques be designed to provide nuanced pointing behaviors?



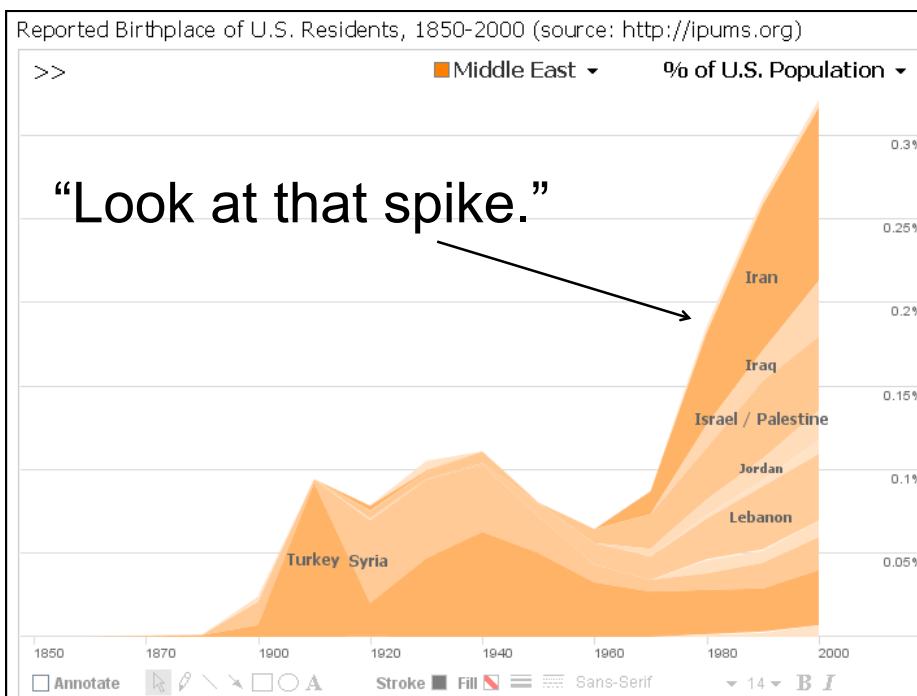
Common Ground

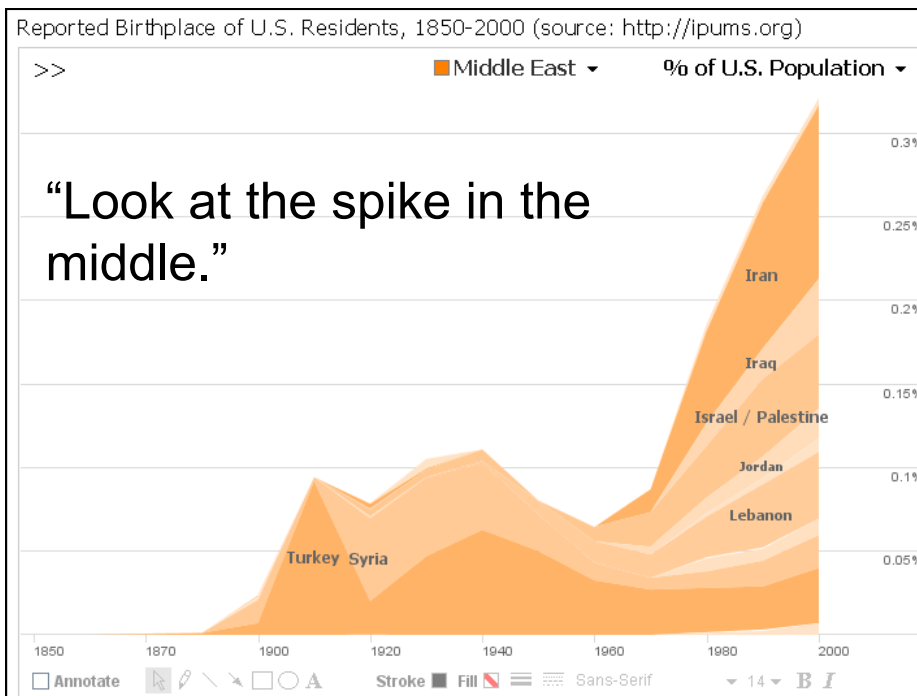
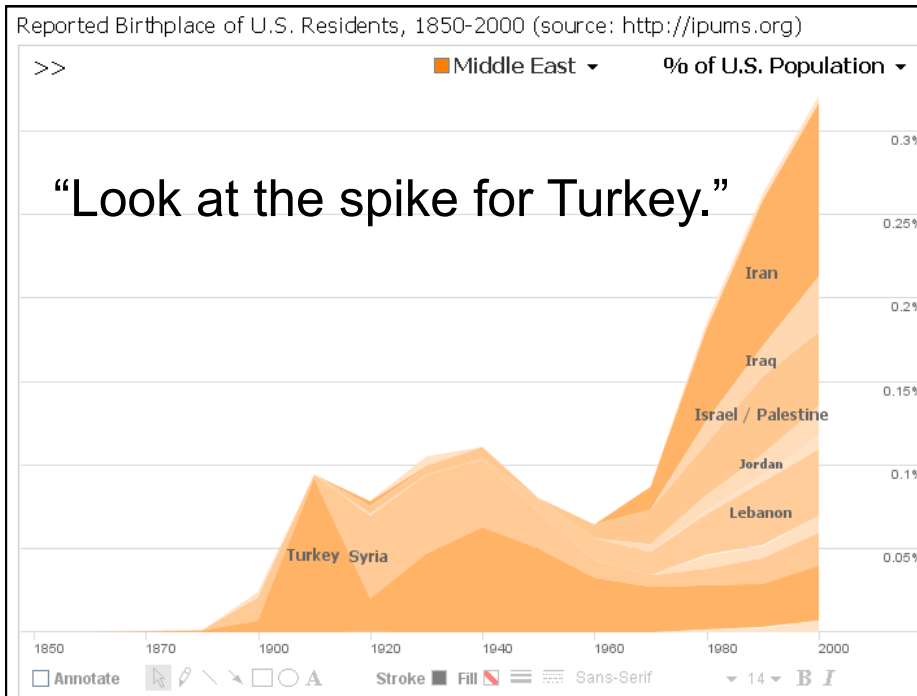
The shared understanding enabling conversation and collaborative action [Clark & Brennan '91]

Do you see what I see? → View sharing (URLs)

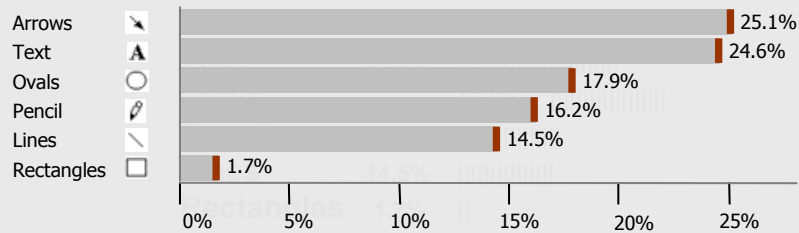
Principle of Least Collaborative Effort: participants exert *just enough* effort to successfully communicate

[Clark & Wilkes-Gibbs '86]





Use of Annotations



39.0% of comments included annotations

Pointing to specific points, trends, or regions (88.6%)

Drawing to socialize or tell jokes (11.4%)

Variety of subject responses

'Not always necessary', but 'surprisingly satisfying'

Some concern about professional look

**Many-Eyes: Social Data
Analysis at an Internet Scale**

www-958.ibm.com/software/analytics/manyeyes/

Many Eyes Log in

Explore

- Visualizations
- Data sets
- Comments
- Topic centers

Participate


- Create a visualization
- Upload a data set
- Create a topic center
- Register

Learn more

- Quick start
- Visualization types
- About Many Eyes
- Privacy
- Blog


Try our featured visualizations

Top 10 Oil Producing Nations



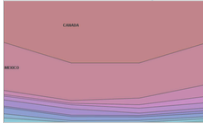
Barrels of crude per day
by evannk

Select Crime Stats in Lafayette, LA



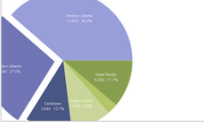
1999 to 2011
by rkr

US Brewed Malted Beer Exports




2008 to 2012
by crossborders1

Global Distribution of Leatherback Turtle



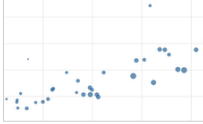
An internationally endangered species.
by Anonymous

Cosmetic Procedures



2012
by KatieCreative

Salaries by Undergraduate Major



Starting and mid-career salaries
by dsbx

Many-Eyes
[Viégas, Wattenberg, et al. CHI 07]

Try out the newest version of IBM Many Eyes! Try ManyEyes v2 >

Templates for Visualizing Data

Create a visualization in three easy steps

1

CHOOSE A DATA SET

First things first. Choose the data set that you would like to visualize.

2

CHOOSE A VISUALIZATION

Then choose the kind of visualization you would like to use.

3

CUSTOMIZE & PUBLISH

Customize what your visualization will look like when others visit. Then publish it.

Start with a data set

You have two choices:

- use one of the existing data sets on the site, or
- Upload your own data set

WORD NET WEB

Bar chart

Network graph

Pie chart

me and

Bar chart

Scatter plot

Treemap

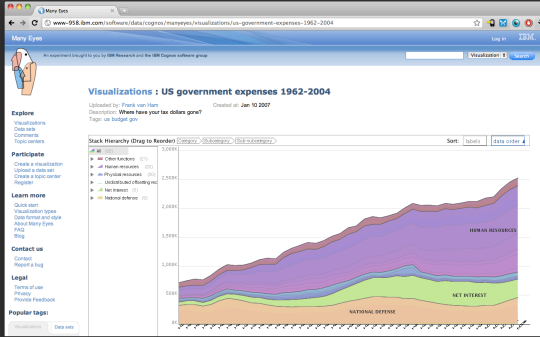
music news TEXT woods words

Bubble chart

Bubble chart

Treemap

Commenting



Mark Waterberg says:
Veterans benefits are going down, percentage-wise.
Posted Jan 11 2007
See view for this comment



Anonymous says:
Health care costs exploded in 1990, and then just kept rising. (Compare veterans benefit hospitals, which grow linearly.)
Posted Jan 12 2007
See view for this comment



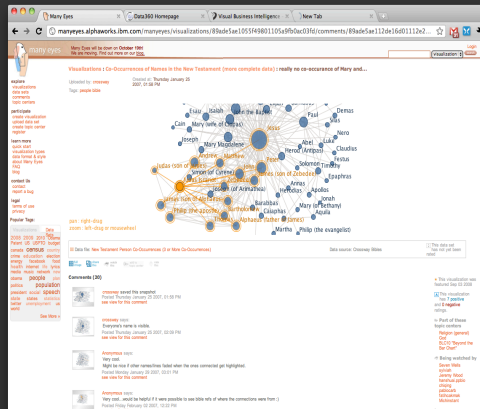
Anonymous says:
Spaceflight funding peaks long before the Apollo landings and never recovers. (Guess someone didn't spin that "one small step" well enough?)
Posted Jan 12 2007
See view for this comment

Josh Peters says:
Re: Spaceflight funding. Had the Soviet Union continued its space expansion to the extent that seemed to compete with the USA's, you can bet our spending would have skyrocketed.
Posted Jan 23 2007



Anonymous says:
The national defense expenses in 1968 and 2004 are roughly equal, but the difference in percentage is enormous. Once wonder if this leads to less resentment for spending money?
Posted Jan 10 2007
See view for this comment

Outside Communities



ESV ENGLISH STANDARD VERSION® BIBLE



The screenshot shows the Swivel.com interface. The main content is a line graph titled "The Status of George Bush's War on Terrorism" showing "Incidents, Injuries, and Fatalities" from 1970 to 2007. The graph shows a significant spike in activity starting around 2001. The legend includes:

- Global Terrorist Incidents, 1968-2007
- Incidents
- Injuries
- Fatalities

 The page also features a "Comments" section with two entries:

- Anonymous says:** "Go BUSH!" posted 11 months ago.
- Tyler says:** "What a completely disingenuous graph, and comments to follow. You have absolutely no reason to think the line would stay flat without Bush. You probably thought graphs like these (and your site more widely) would help people overcome preconceived notions by showing hard data. Instead, you're just amplifying people's preconceived notions -- adding legitimacy."

 The right side of the page has a dark background with the text "Swivel.com, and others..." in white.

A graveyard of "YouTubes for Data"

Emergent analysis and discussion isn't very good

Visualizations : Domestic Net Migration 2004-2008
Uploaded by: Anonymous Created at: Sunday March 07 2010, 08:38 PM
Description:
Tags: states united co

Domestic Net Migration

Anonymous says:
Very impressive blog. I will share it with my friends as well.
Posted Monday October 11 2010, 02:02 AM
[see view for this comment](#)

neshvox says:
Thanks for the link to this resource.
[hardtall pants](#) | [mikimoto necklace](#)
Posted Friday April 16 2010, 02:00 PM
[see view for this comment](#)

Search>>
Color No selection

To highlight or find to click or ctrl-click

Many Eyes – circa 4/2012

128,478 Visualizations

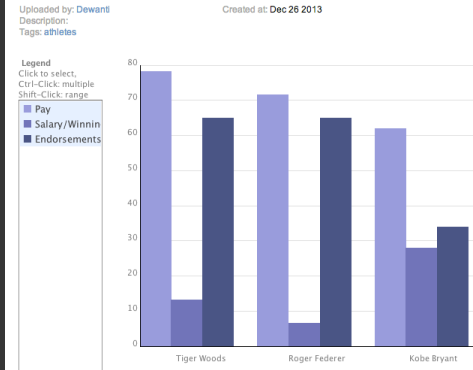
17,340 Comments

only ~11% of comments provided a plausible hypothesis or explanation for the data in the chart

**Can having access to social information
be harmful to analysis?**

**The Impact of Social Information
on Visual Judgments**

Visualizations : Highest Paid Athletes of 2013 3 VARIABLES



Can evidence of others' graphical judgments (accurate or not) impact subsequent users' judgments?

Comments

Currently showing



Highest Paid Athletes of 2013 3 VARIABLES

Woods and Federer are equal in endorsements. Bryant barely gets half as many.

Posted Dec 26 2013
See view for this comment

Social Influence



Asch 1951

Social Influence

Normative versus Informational

Using social signals as evidence when we're choosing:

- Restaurants
- Music (Salganik et al. 2009)
- Tags (Golder et al. 2007)

Controlled Experiment

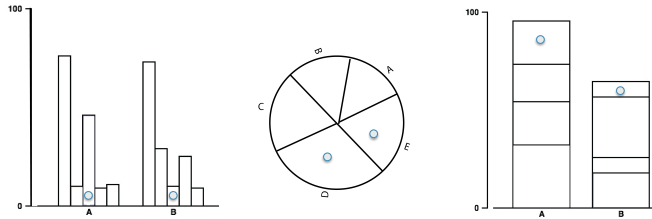
Interested in impact on accuracy of graphical judgments (given task where subjects are incentivized to be correct)

Two conditions:

Non-social: Visual judgment task with no social information

Social: Same task, presented with summary of judgments from non-social group

Visual judgment task

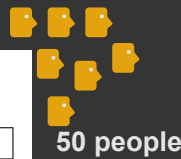
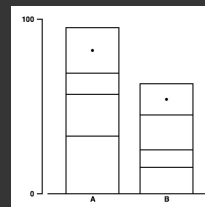


Cleveland & McGill 1984

1. Which of the two marked sections is larger?
2. Make a quick visual judgment of what proportion the smaller section is of the larger.

AMT study: Heer & Bostock 2010

Non-social



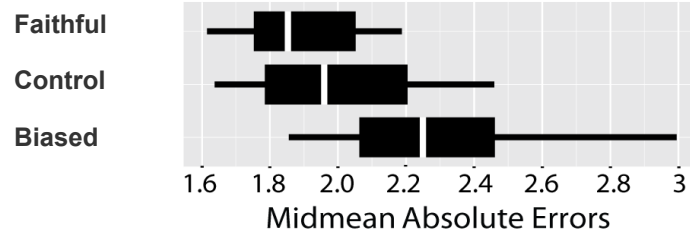
Social

Mean (56%) 1SD (78%)



Results

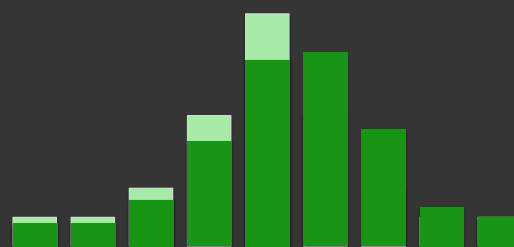
Can having access to others' graphical judgments (accurate or not) impact subsequent users' judgments?



Does the amount of social information matter?

Implications

No social information?



Mean \neq true value
(the reality)

Issues when analysis remains shallow ...

Can we augment social data analysis to support deeper analysis and synthesis?