



The Value of Visualization... and Why Interaction Matters

John Stasko

School of Interactive Computing

Georgia Institute of Technology

stasko@cc.gatech.edu





Pennard Golf Club



Data Visualization

Making pretty pictures?



Data Visualization

~~Making pretty pictures~~



Data Visualization

A cognitive process

Gain an understanding



Data Visualization

A cognitive process

Gain an **understanding**

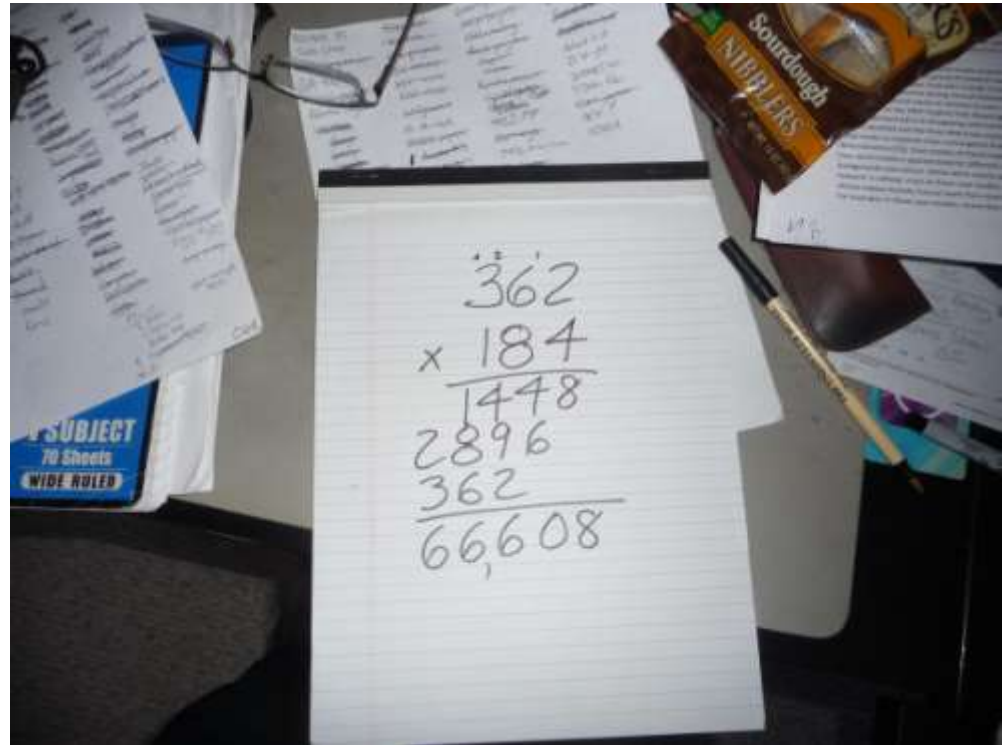


Visuals help us think

Provide a frame of reference, temporary storage area

Cognition → Perception

Pattern matching



Visualization's Value

Need to make the case better externally



The Value of Information Visualization

Jean-Daniel Fekete¹, Jarke J. van Wijk², John T. Stasko³, and Chris North⁴

¹ Université Paris-Sud, INRIA, Bât 490,
F-91405 Orsay Cedex, France,
Jean-Daniel.Fekete@inria.fr,
<http://www.aviz.fr/~fekete/>

² Department of Mathematics and Computing Science,
Eindhoven University of Technology, P.O. Box 513,
5600 MB EINDHOVEN, The Netherlands,
vanwijk@win.tue.nl,
<http://www.win.tue.nl/~vanwijk/>

³ School of Interactive Computing, College of Computing & GVU Center,
Georgia Institute of Technology, 85 5th St., NW,
Atlanta, GA 30332-0760, USA,
stasko@cc.gatech.edu,
<http://www.cc.gatech.edu/~john.stasko>

⁴ Dept of Computer Science, 2202 Kraft Drive,
Virginia Tech, Blacksburg, VA 24061-0106, USA,
north@vt.edu,
<http://people.cs.vt.edu/~north/>

Abstract. Researchers and users of Information Visualization are convinced that it has value. This value can easily be communicated to others in a face-to-face setting, such that this value is experienced in practice. To convince broader audiences, and also, to understand the intrinsic value of information visualization is more difficult however. In this paper we...

Information Visualization: Human-Centered Issues and Perspectives

Springer, 2008

Need for Visualization

Articulating value requires identifying purpose



Three stories...



People - Faculty | School of In...

www.ic.gatech.edu/people/faculty

Most Visited Getting Started Weather Zimbra T2 1331 7450 IH ihmga ESPN Golf gmail facebook JazzRadio

Georgia Tech College of Computing







School of Interactive Computing

About IC People Future Students Current Students Research News & Events Contact

Home > People > Faculty

People - Faculty

Faculty | Staff

	<p>Gregory Abowd Distinguished Professor; Regents Professor</p>		<p>Hua Ai Research Scientist II</p>		<p>Annie Antón Professor and Chair, School of Interactive Computing</p>
	<p>Alberto Apostolico Professor</p>		<p>Ronald Arkin Associate Dean for Research and Space Planning; Regents' Professor</p>		<p>Rosa Arriaga Senior Research Scientist</p>

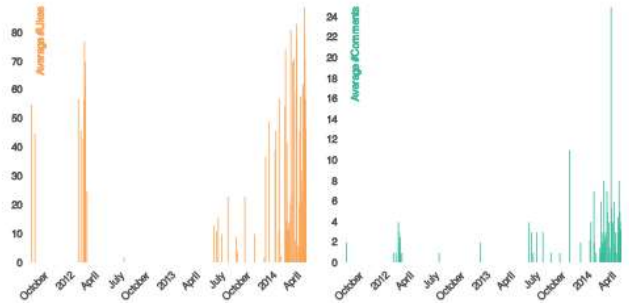
Instagram Visualization

Result for the hashtag # with #

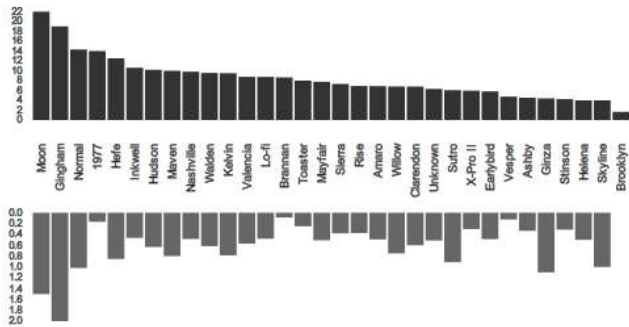
Options

Number of photos per day (photos get from API, not the actual photo counts on Instagram)

Enable Tooltip Enable Brush



Average numbers of likes per photos per filter



Average numbers of comments per photos per filter

Popular photos

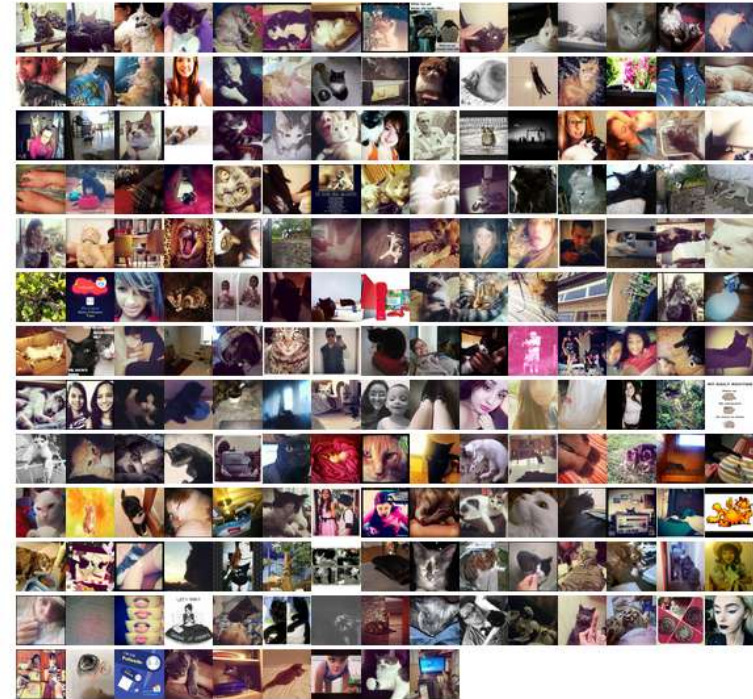


Photo Navigator On Off

Color schemes of the photo



Chloe Xie



Harry Callahan (Clint Eastwood)

“A man’s got to know his limitations.”
Magnum Force, 1973

Thought

If you can articulate very precisely what you're seeking, visualization likely isn't your best approach

OK, so what is visualization good for?

Applications of Visualization

Presentation

Analysis

Explanatory

Exploratory

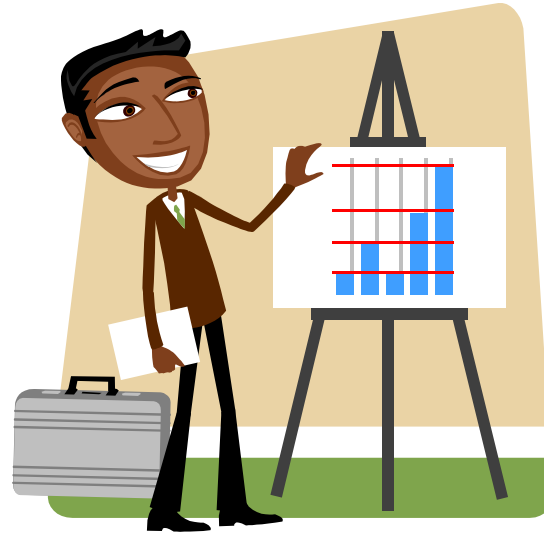
1. Presentation

Communicate data and ideas

Explain and inform

Provide evidence and support

Influence and persuade



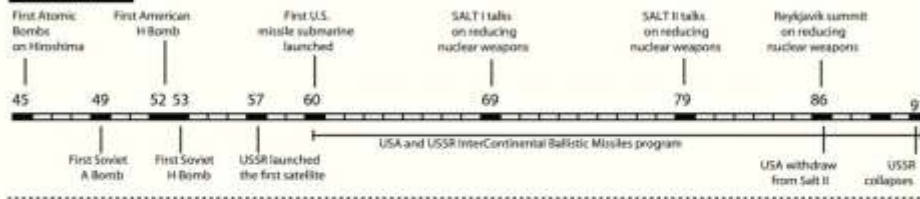
Infographics

All the rage...

THE NUCLEAR ARMS RACE

It was the main issue in the Cold War when both America and Russia challenging each other to increase their stockpiles of nuclear weapons.

TIMELINE



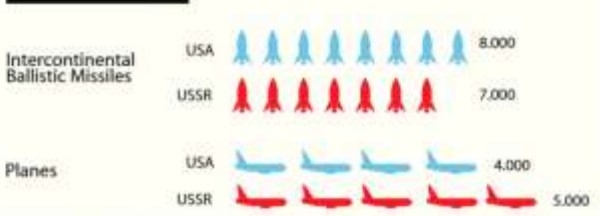
NUCLEAR STOCKPILE



ROCKET MODELS

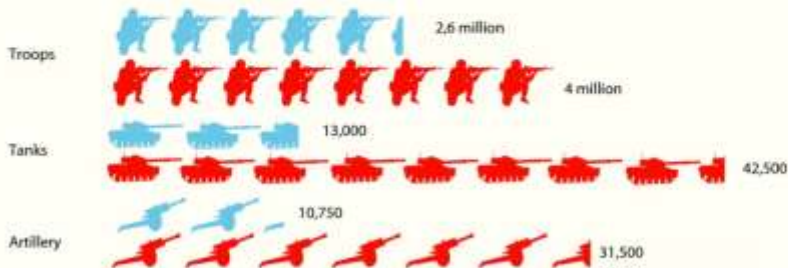


AIR CONTROL



A WORLD BREAK IN TWO

NATO and Warsaw Pact were both mutual defense treaties between states. The first one started on 1949 till today while the Warsaw Pact lasted from 1955 till 1991.



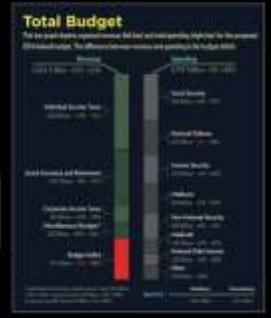
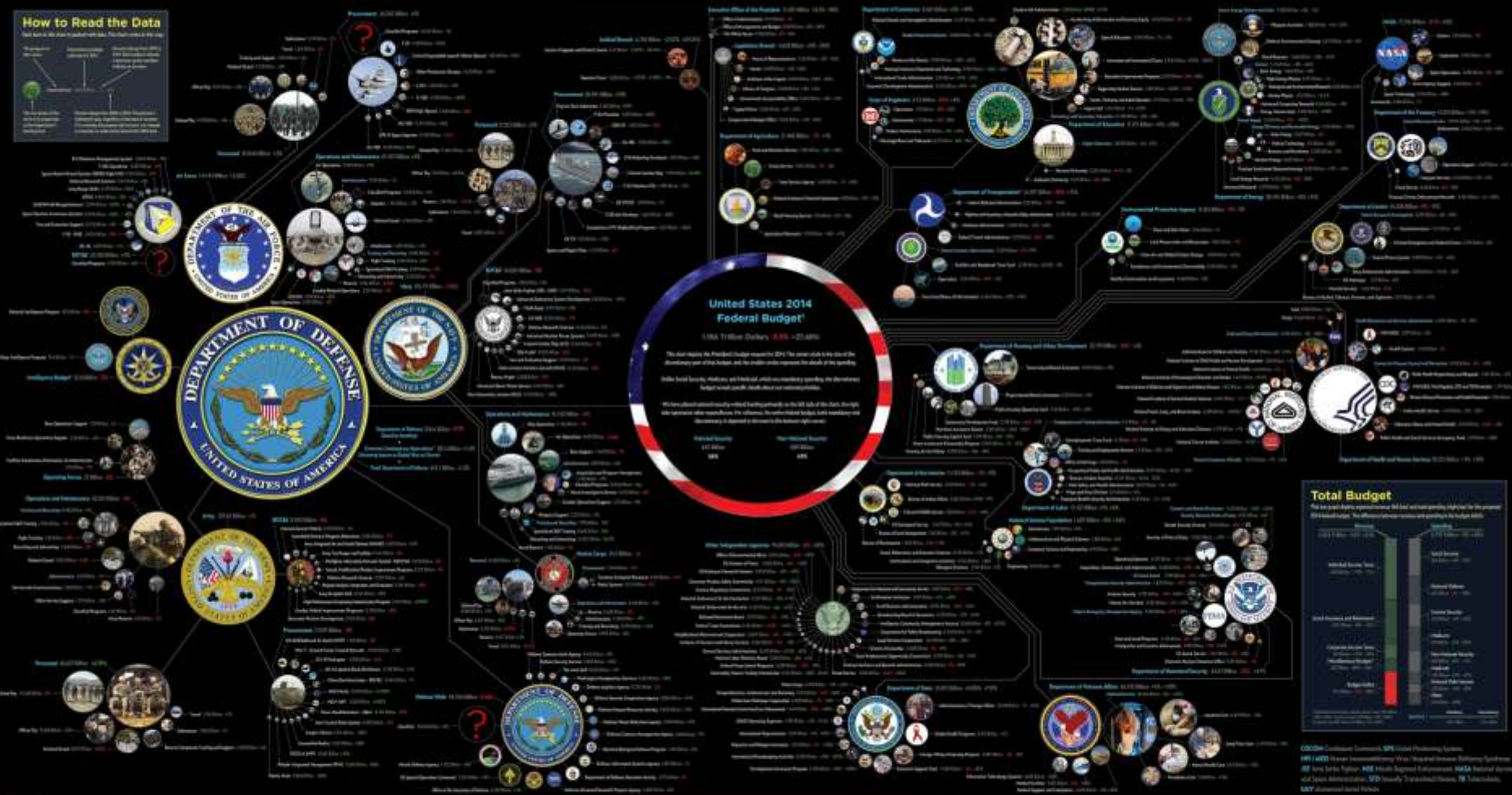
Death & Taxes

A Visual Guide to Where Your Federal Tax Dollars Go

United States 2014 Federal Budget

How to Read the Data

Each circle in the chart represents a different federal agency or program. The size of the circle is proportional to the amount of money that agency or program receives from the federal budget. The color of the circle corresponds to the agency or program's logo.

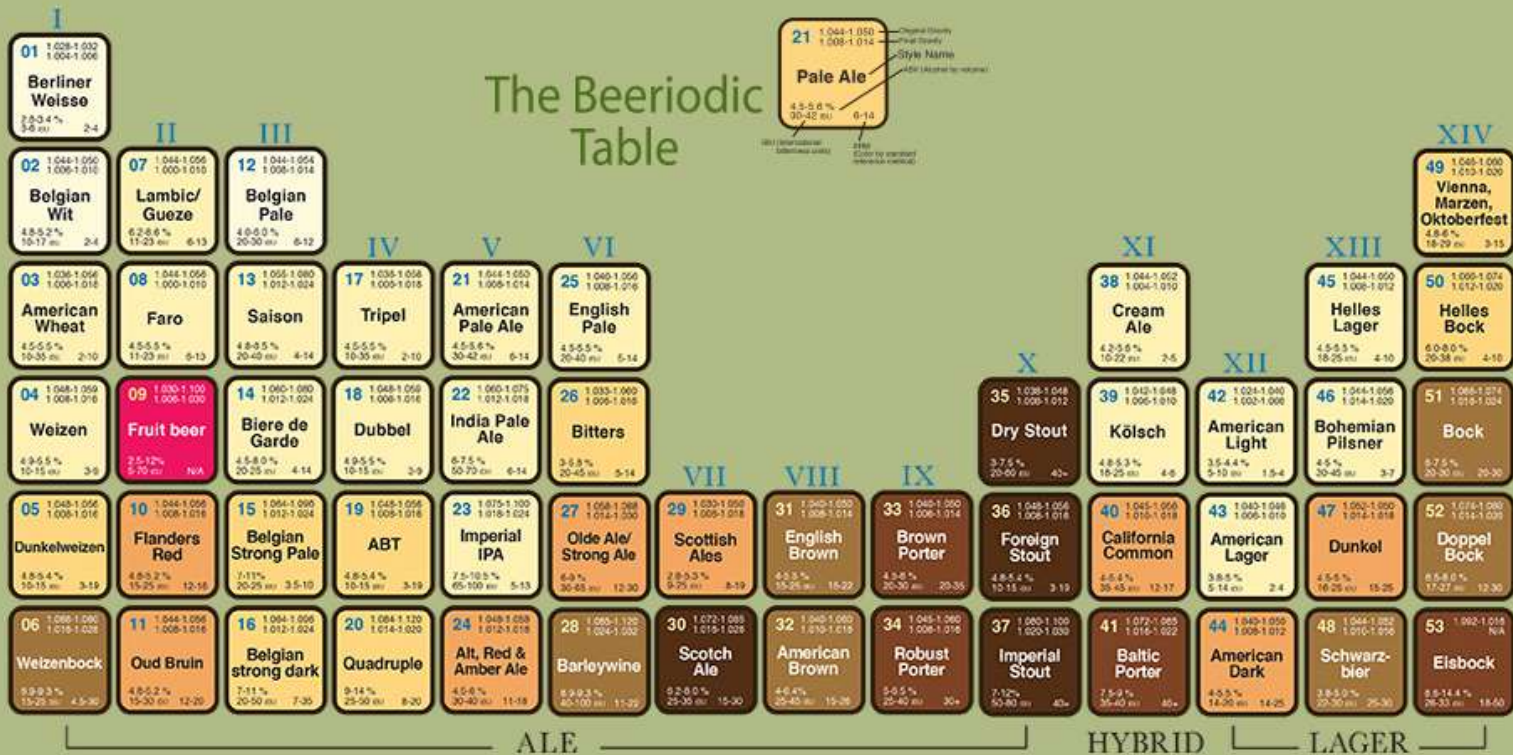


For Comparison



Source: U.S. Department of the Treasury, Office of Management and Budget, Budget of the United States Government, 2014. Data is preliminary and subject to change. All figures are in billions of dollars unless otherwise noted. © 2014 TIMEPLOTS. All rights reserved. TIMEPLOTS is a registered trademark of TIMEPLOTS, INC. All other trademarks are the property of their respective owners.

Click on any image below to see full description



<http://thebeermongers.com/beers/>

755

Steroids or Not, the Pursuit Is On

Barry Bonds is taking aim at the career home run record. He needs only six more to tie Babe Ruth and 47 to equal Hank Aaron.

Lines are cumulative home runs

Hank Aaron
755 homers
23 seasons



Babe Ruth
714 homers
22 seasons



Barry Bonds
708 homers
20 seasons

Bonds takes lead
Home runs
after 16 seasons
Bonds 567
Aaron 554
Ruth 516

600

714

755

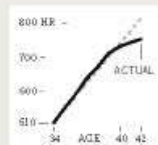
Bonds was injured last season. He played 14 games and hit 5 homers

Homer Pace After Age 34

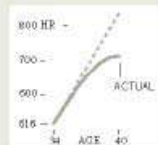
If the accusations are correct, Bonds was 34 in his first season on steroids. Here are projected home run paces for each player after age 34.

----- PROJECTED PACE BASED ON AVERAGE OF PREVIOUS FIVE SEASONS

Aaron
Actual homers slightly outpace projected homers for five seasons.



Ruth
Averaged 46.4 homers a season from age 30 to 34. Averaged 42.5 for next four seasons.



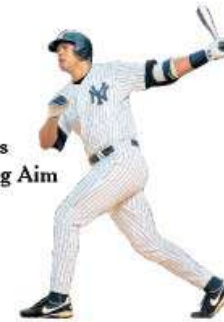
Bonds
From age 35 to 39, he averaged 14 more homers a season than projected.



Note: Ages as of July 1 of each season.

According to allegations in a book about Bonds, he began taking steroids before the 1999 season, his 14th in the league. Two seasons later, he hit 73 home runs, surpassing Aaron's career pace.

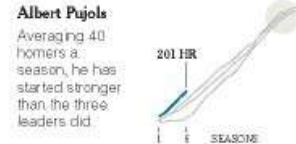
Others Taking Aim



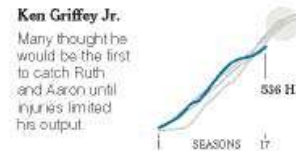
Alex Rodriguez
Is ahead of the pace set by all three home run leaders.



Albert Pujols
Averaging 40 homers a season, he has started stronger than the three leaders did.



Ken Griffey Jr.
Many thought he would be the first to catch Ruth and Aaron until injuries limited his output.



Differing Paths to the Top of the Charts

The top seven players on the career home run list, along with a look at Griffey (12th), Rodriguez (37th) and Pujols (tied 257th).



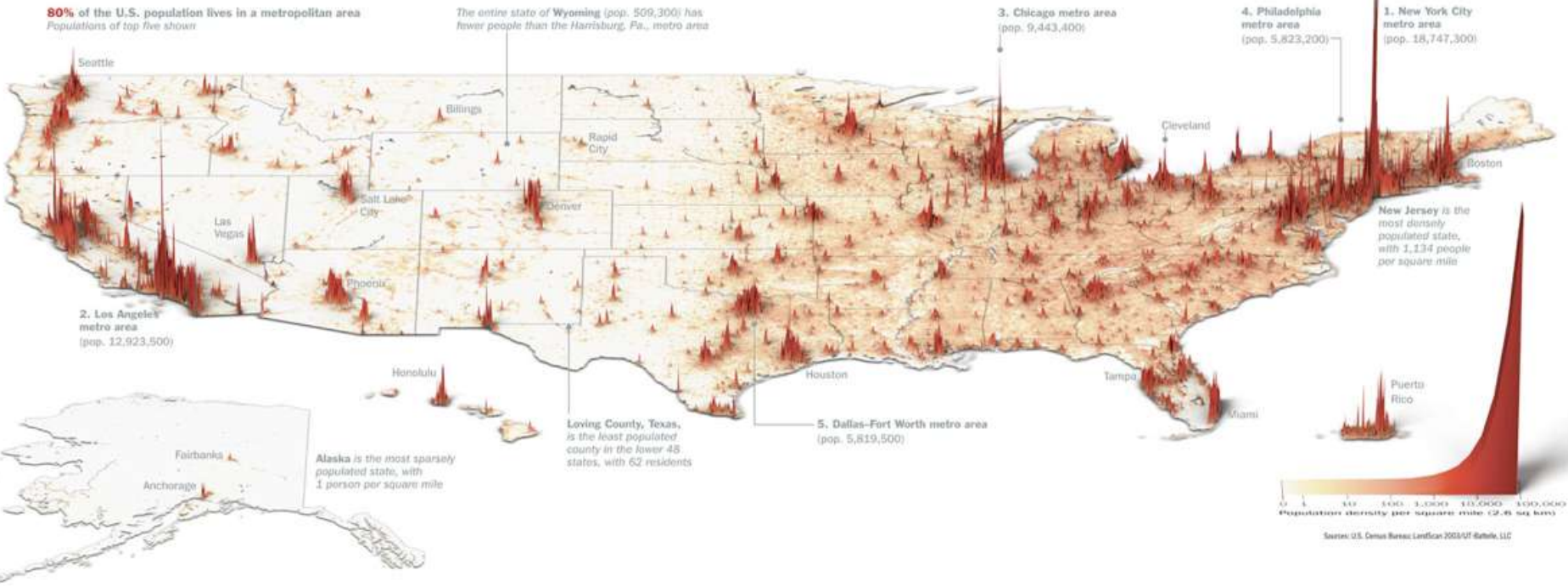
Illustration by Joe Ward/The New York Times

Where We Live...

Unlike many developed countries, the U.S. keeps growing. We are also moving south and west. But compared with China or India, the nation is a vast prairie

Our families are getting smaller—with one vital exception. Compared with those of Europe and Japan, the U.S. population is younger and more colorful because of the continued arrival of immigrants and their higher-than-average birthrates. Of the 100 million Americans who will join us in the next 37 years, half will be immigrants or their children. In the next few decades, 97% of the world's population growth will occur in the developing world; the U.S. is the largest developed country in the world that is still growing at a healthy clip. That matters, strategically, economical-

Ala.; Possum Trot, Ky.; or Lonelyville, N.Y. But they are all probably close to someone's idea of paradise. —By Nancy Gibbs



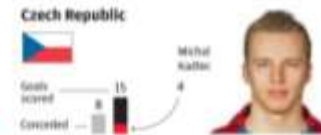
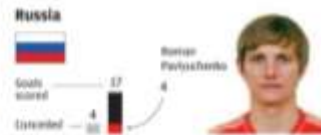
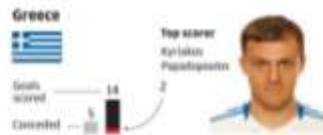
<http://infographicsnews.blogspot.com/2009/04/mantras-joe-lertolas-maps.html>

All of the goals

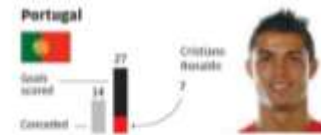
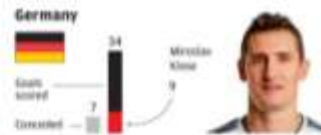
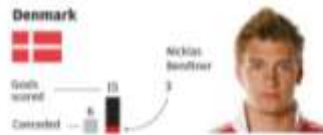
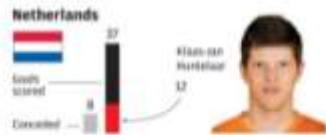
which took the teams to Euro 2012

Euro 2012 kicks off tomorrow night and promises to keep soccer fans glued to their television sets for the rest of the month. The Post takes a look at how the 16 finalists put the ball in the back of the net in qualifying, giving an idea of who is most likely to be a scoring threat in the finals, and from where.

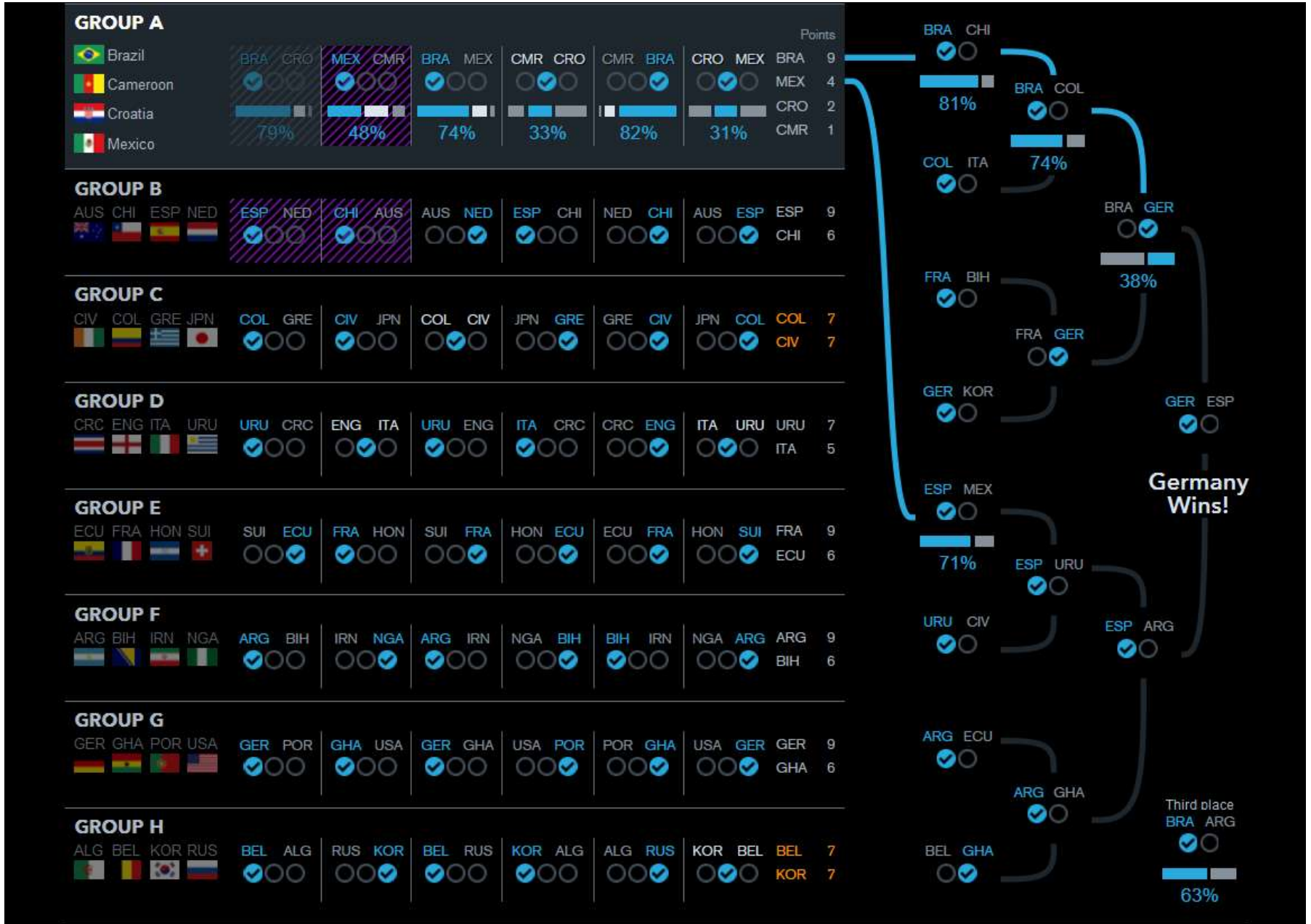
Group A



Group B



Simon Scarr
South China Morning Post
 June 7, 2012



the game • 2014 World Cup



TIMES READER PREDICTIONS

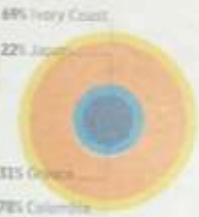
More than 3,000 Times readers have gone online to give their predictions about how the World Cup finals will develop. Starting with the groups, readers have worked their way through the competition and their thoughts are

collated below. Participants were asked which teams would qualify from the group stages and who would win the tournament overall. There was little disagreement, with only group E throwing up diverse viewpoints.

THE GROUP STAGES

Everyone thought Brazil would qualify from group A but were split over who join the hosts in the round of 16 by finishing second. Australia were the least fancied team to advance, a mere 1 per cent of people think they would get out of group B

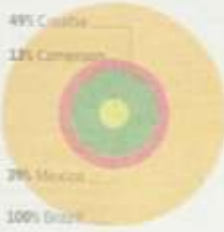
GROUP C



GROUP E



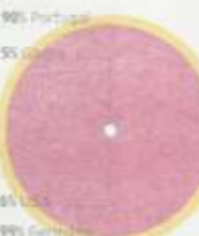
GROUP A



GROUP D



GROUP G



GROUP B



GROUP F

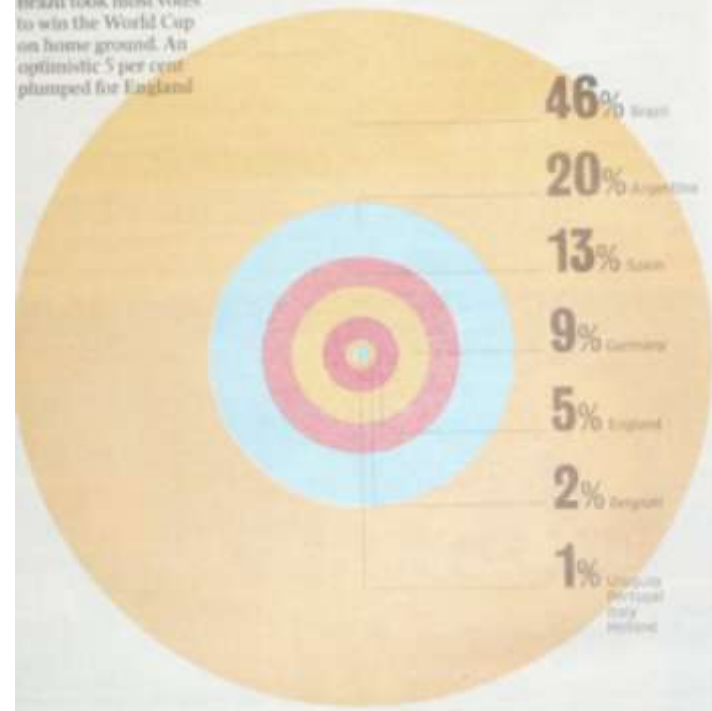


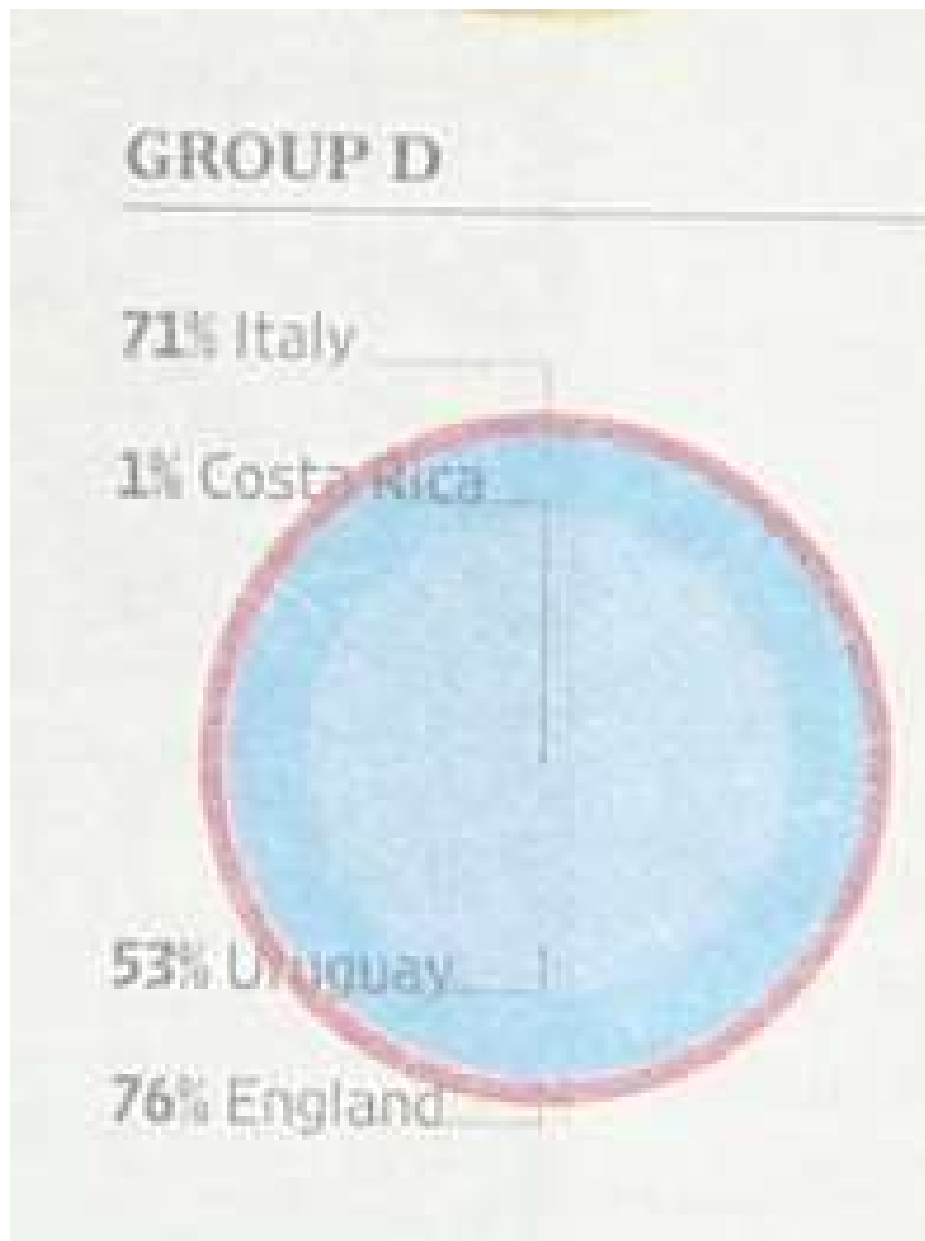
GROUP H



WHO WILL WIN THE WORLD CUP?

Brazil took most votes to win the World Cup on home ground. An optimistic 5 per cent plumped for England





GROUP G

90% Portugal

5% Ghana

6% USA

99% Germany



Frequent presentation goals

Clarify

Focus

Highlight

Simplify

Persuade

May just show a few variables
and/or a subset of the data cases

Simply presenting data ***visually*** can have a profound impact

My Class

<http://www.cc.gatech.edu/~stasko/7450>



CS 7450 - Information Visualization

Instructor: [John Stasko](#)
Fall 2013
Mon, Wed 3:00 - 4:30 pm
Whitaker Bldg. room 1103

Information visualization is a research area that focuses on the use of visualization techniques to help people understand and analyze data. While fields such as scientific visualization involve the presentation of data that has some physical or geometric correspondence, information visualization focuses on abstract data without such correspondences such as symbolic, tabular, networked, hierarchical, or textual information sources.

The objectives of the course are

- Learn the principles involved in information visualization
- Learn about the variety of existing techniques and systems in information visualization
- Develop skills in critiquing different visualization techniques as applied to particular tasks
- Learn how to evaluate visualization systems
- Gain a background that will aid the design of new, innovative visualizations

The course will follow a lecture/seminar style with much discussion of assigned readings, as well as viewing of videos and hands-on experience with research and commercial information visualization tools.

We will be reading recent research papers about the different course topics. In addition, we will be using one book for the course: [Now You See It](#) by Stephen Few, Analytics Press 2009. Also highly recommended is *Envisioning Information* by Edward Tufte, Graphics Press 1990.



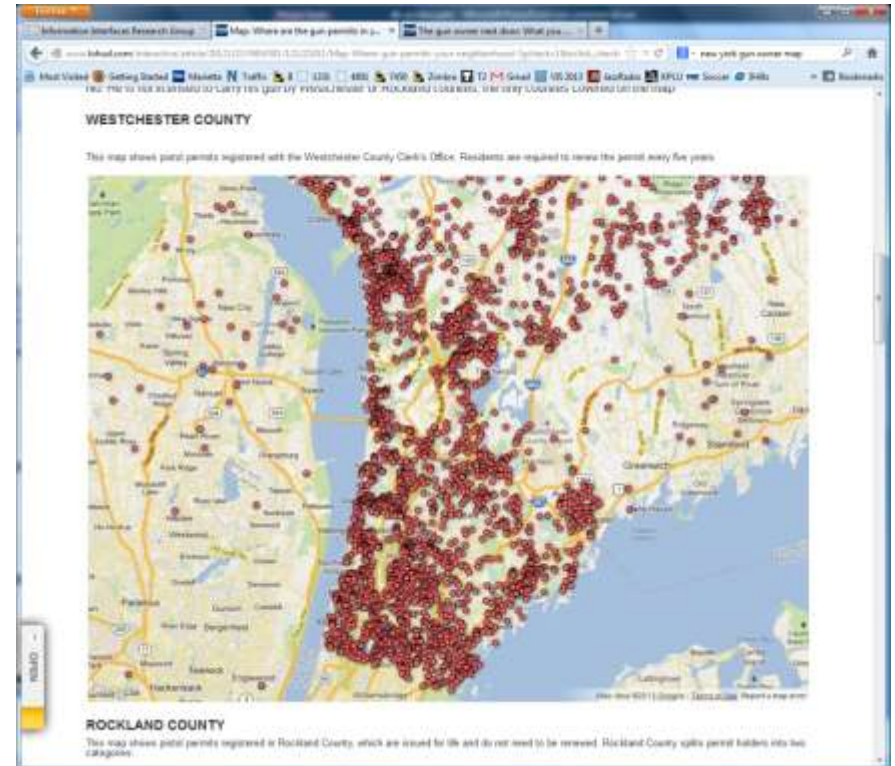
CS 7450
Fall '12
Design project



Nate Osborne
Nitya Noronha
Ameya Zambre
Pratik Zaveri

Gun ownership in New York counties

The screenshot shows a news article on the lohud.com website. The article title is "The gun owner next door: What you don't know about the weapons in your neighborhood". Below the title is a video player showing a man in a green jacket. To the right of the video is a sidebar with a "Green Team" logo and a list of "Most Popular" stories. The top of the page features a navigation bar and a search box.



http://www.lohud.com/apps/pbcs.dll/article?AID=2012312230056&nclick_check=1

http://www.lohud.com/interactive/article/20121223/NEWS01/121221011/Map-Where-gun-permits-your-neighborhood-?gcheck=1&nclick_check=1

TALKS
 Hans Rosling shows the best stats you've ever seen

TED2006 (Filmed Nov 2006, Posted Jun 2008)

About this talk [Open interactive transcript](#)

You've never seen data presented like this. With the drama and urgency of a sportscaster, statistic guru Hans Rosling debunks myths about the so-called "developing world."

About Hans Rosling

As a doctor and researcher, Hans Rosling identified a new parasite disease that led to fever in rural Africa. Now he looks at the bigger picture of social and economic development with his [TED talk](#) and [TEDx talk](#).

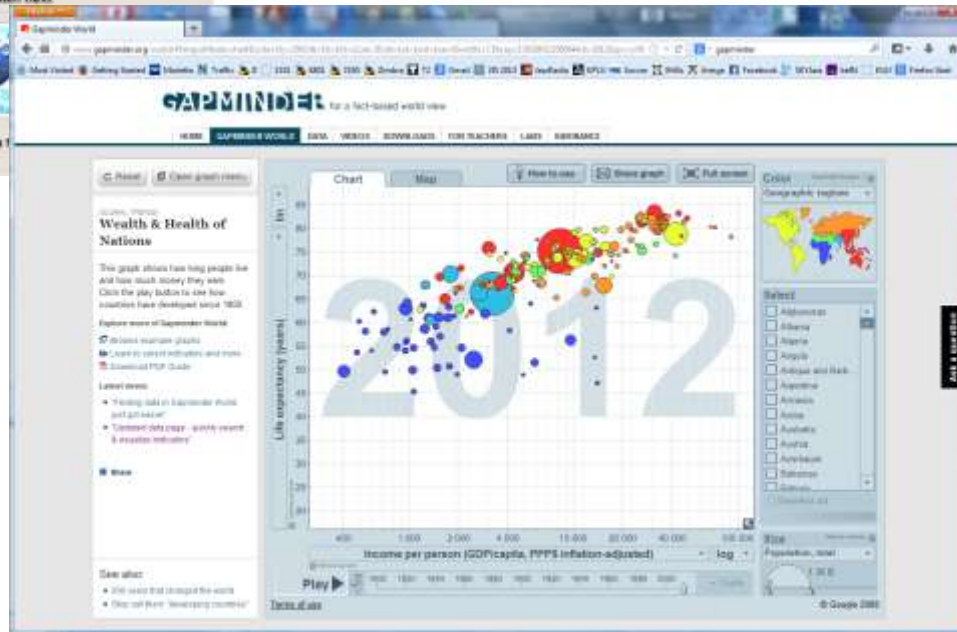
Thanks to our sponsor

fivebytwenty

What to watch next

Other talks "World"

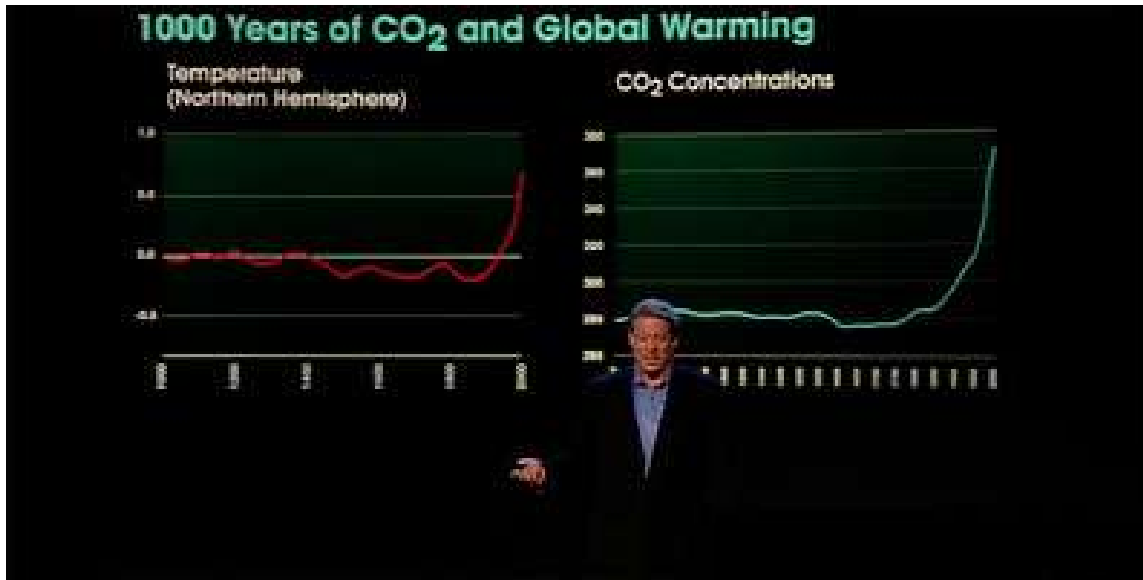
Hans Rosling
 Gapminder



2006

http://www.ted.com/index.php/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen.html

An Inconvenient Truth



Gore made extensive use of data graphics



2. Analysis

Explore the data

Assess a situation

Determine how to proceed

Decide what to do



Many Data Analysis Approaches

Statistics

Database & information retrieval

Data mining

Machine learning



“Contained within the data of any investigation is information that can yield conclusions to questions not even originally asked. That is, there can be surprises in the data...To regularly miss surprises by failing to probe thoroughly with visualization tools is terribly inefficient because the cost of intensive data analysis is typically very small compared with the cost of data collection.”

W. Cleveland

The Elements of Graphing Data

Frequent analysis goals

Show many variables

Illustrate overview and detail

Facilitate comparison

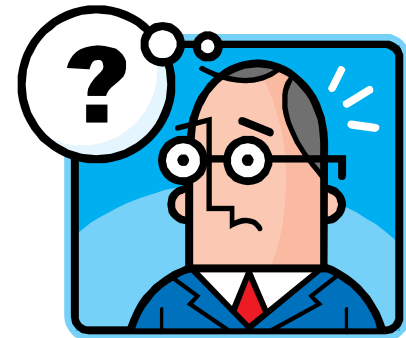
Display may not be easy to
interpret at first

Visualization most useful in **exploratory data analysis**

Don't know what you're looking for

Don't have a priori questions

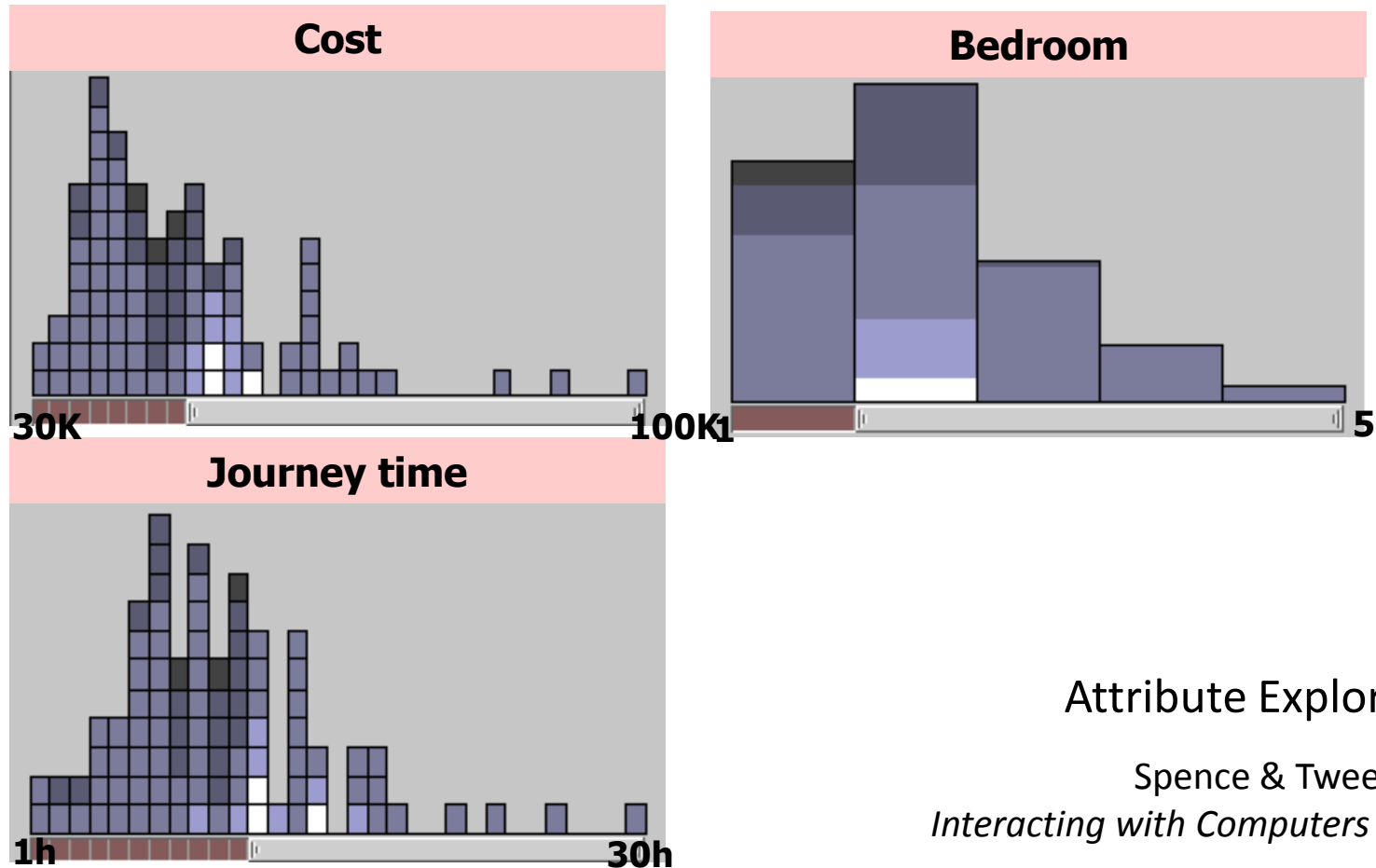
Want to know what questions to ask



At Its Heart: Uncertainty

Nothing is perfect

Humans trade off choices, compromise



Attribute Explorer

Spence & Tweedie
Interacting with Computers '98

Thought

Even analysis is about effective communication

Visualization's Value?

Visualization's Value?

$$V_{\text{alue}} = T + I + E + C$$

Visualization's Value?

$$V_{\text{alue}} = T + I + E + C$$

Ability to minimize the total **time** needed to answer a wide variety of questions about the data

(Without formal queries, Interaction really helps)

What kinds of questions?

“Low-level” tasks

Retrieve value

Filter

Compute derived value

Find extremum

Sort

Determine range

Characterize distribution

Find anomalies

Cluster

Correlate

Amar, Eagan & Stasko
InfoVis '95

Visualization's Value?

$$V_{\text{alue}} = T + \mathbf{I} + E + C$$

Ability to spur and discover **insights** or insightful questions about the data

(Would be very difficult with only the data)

What is Insight?

An individual observation about the data by the participant, a unit of discovery

Complex

Deep

Qualitative

Relevant

Unexpected

Saraiya, North, & Duca
TVCG '05

North
IEEE CG&A '06

What is Insight?

Sudden grasp of new relationships that are necessary to solve a problem and that were not learned in the past

Bernstein, Penner, Clarke-Stewart & Roy
Psychology, 6th edition

What is Insight?

Is not spontaneous “aha!” moments (eg, in cognitive science)

Is knowledge-building and model-confirmation

Like a substance that people acquire with the aid of systems

Chang, Ziemkiewicz, Green, & Ribarsky
IEEE CG&A '09

Visualization's Value?

$$V_{\text{alue}} = T + I + E + C$$

Ability to convey an overall **essence**
or take-away sense of the data

(The big picture: Whole is greater than
the sum of the parts)

Overview and detail

Focus + context

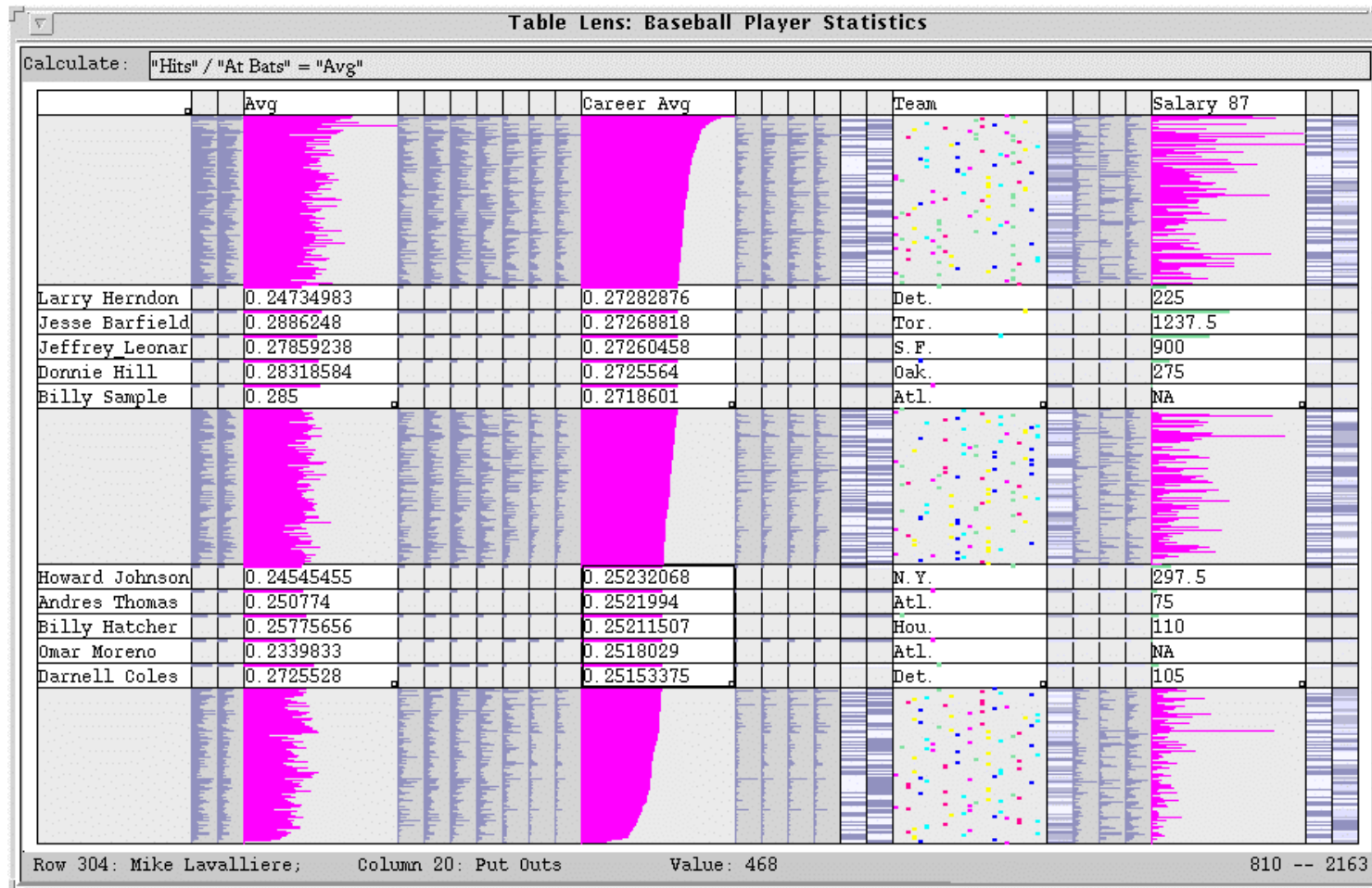


Table Lens

Rao & Card
CHI '94

Visualization's Value?

$$V_{\text{alue}} = T + I + E + C$$

Ability to generate **confidence** and trust about the data, its domain and context

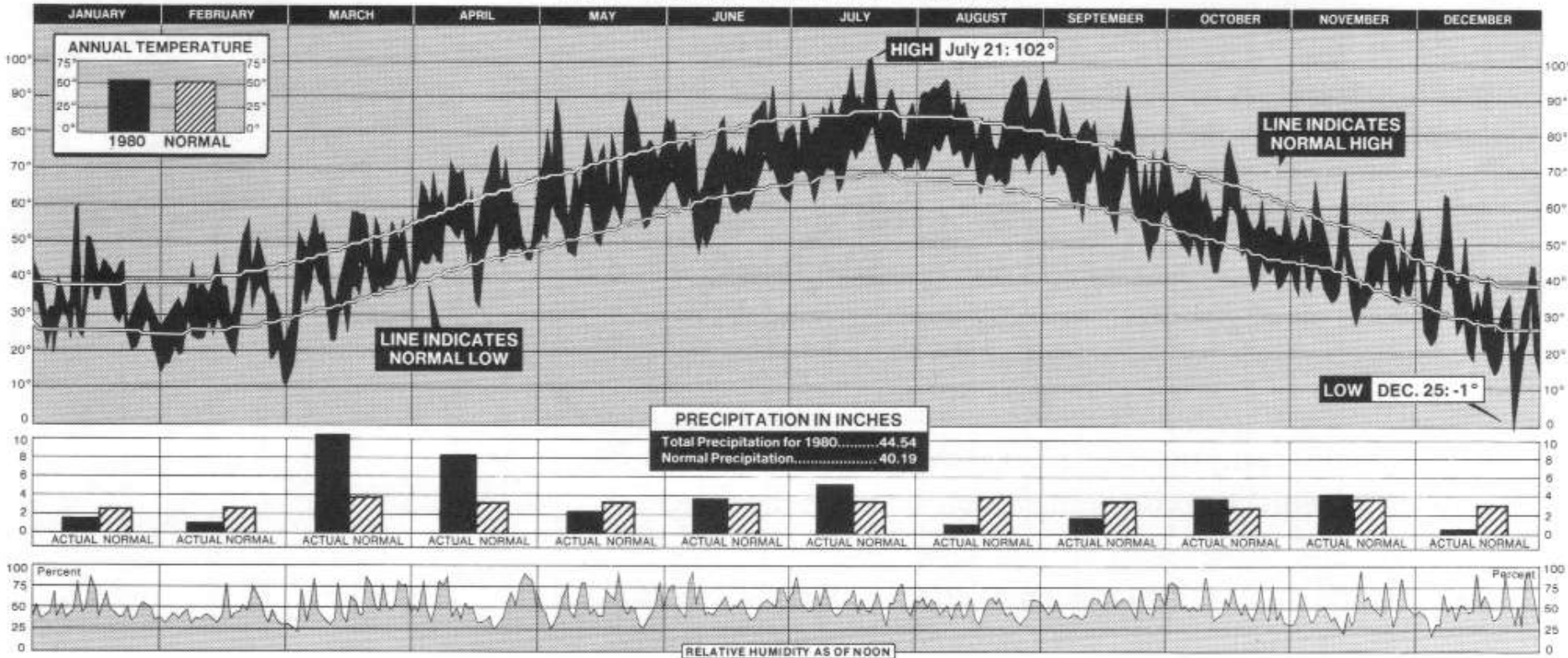
(Beneficial data analysis process side effects)

Some examples

1.

2220 numbers

NEW YORK CITY'S WEATHER FOR 1980



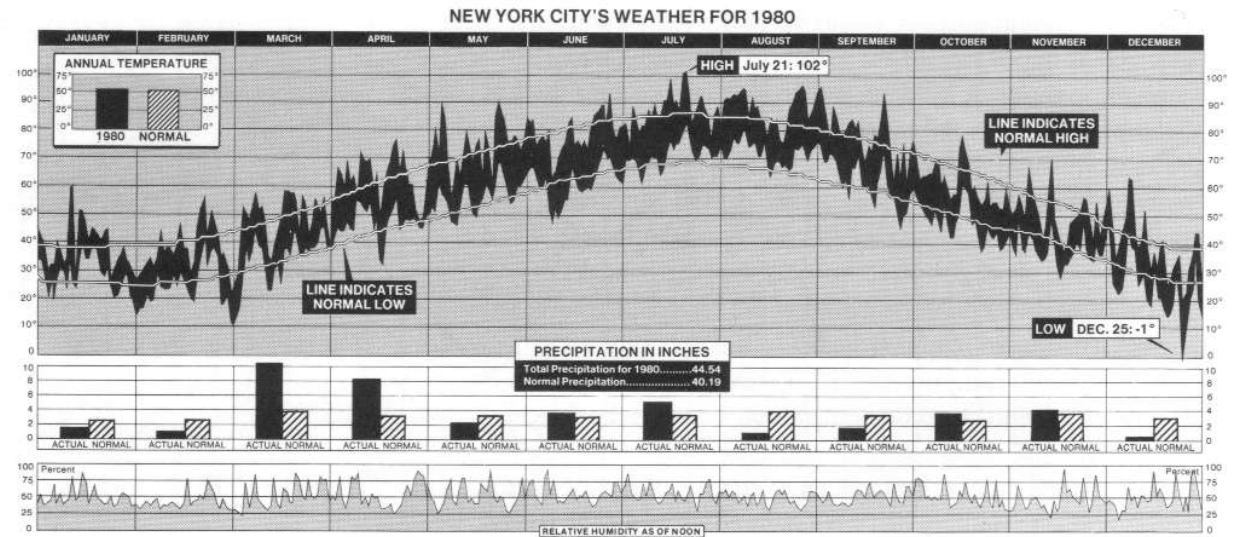
New York Times, January 11, 1981, p. 32.

E. Tufte
The Visual Display of Quantitative Information
1983

Data Values

- 365 High temp for each day
- 365 Low temp for each day
- 365 Avg high temp for each day
- 365 Avg low temp for each day
- 365 Precipitation for each day
- 365 Humidity for each day
- 12 Precipitation for each month
- 12 Avg precipitation for each month
- 1 Precipitation for the year
- 1 Avg precipitation per year
- 1 Highest temp (& day) for the year
- 1 Lowest temp (&day) for the year
- 1 Avg daily temp for the year
- 1 Avg daily temp per year

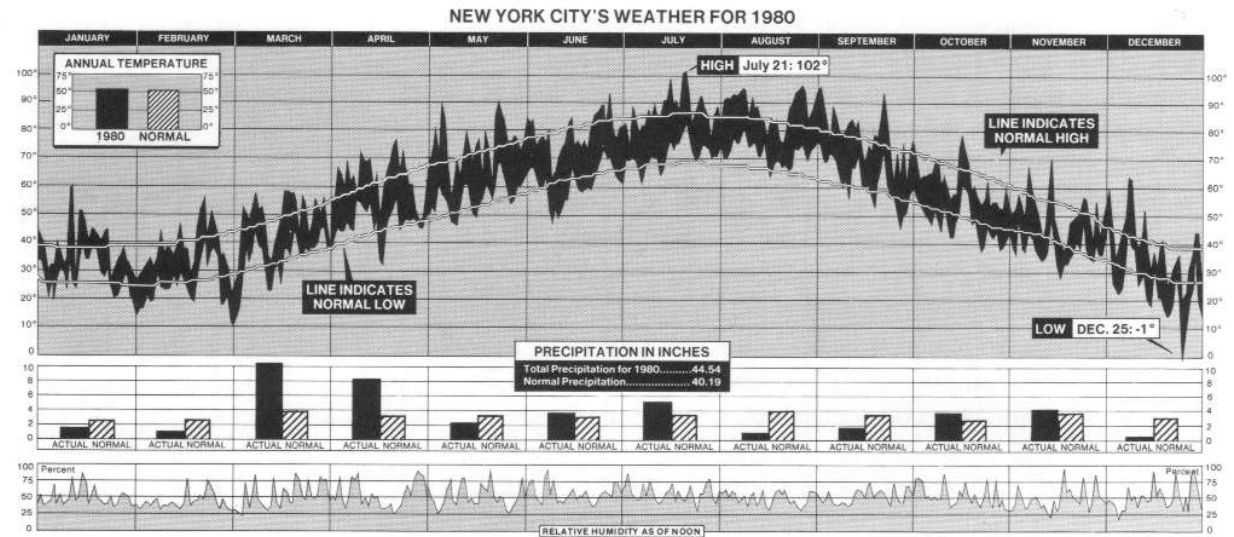
Visualization's Value



New York Times, January 11, 1981, p. 32.

$$V_{\text{alue}} = T + I + E + C$$

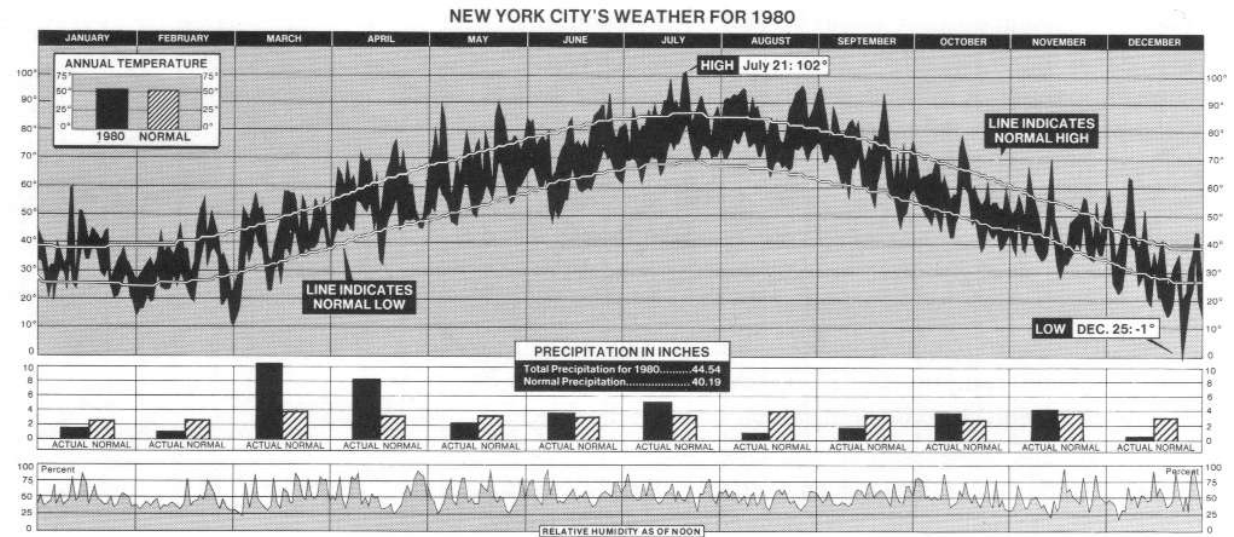
Visualization's Value



New York Times, January 11, 1981, p. 32.

$$V_{\text{alue}} = T + I + E + C$$

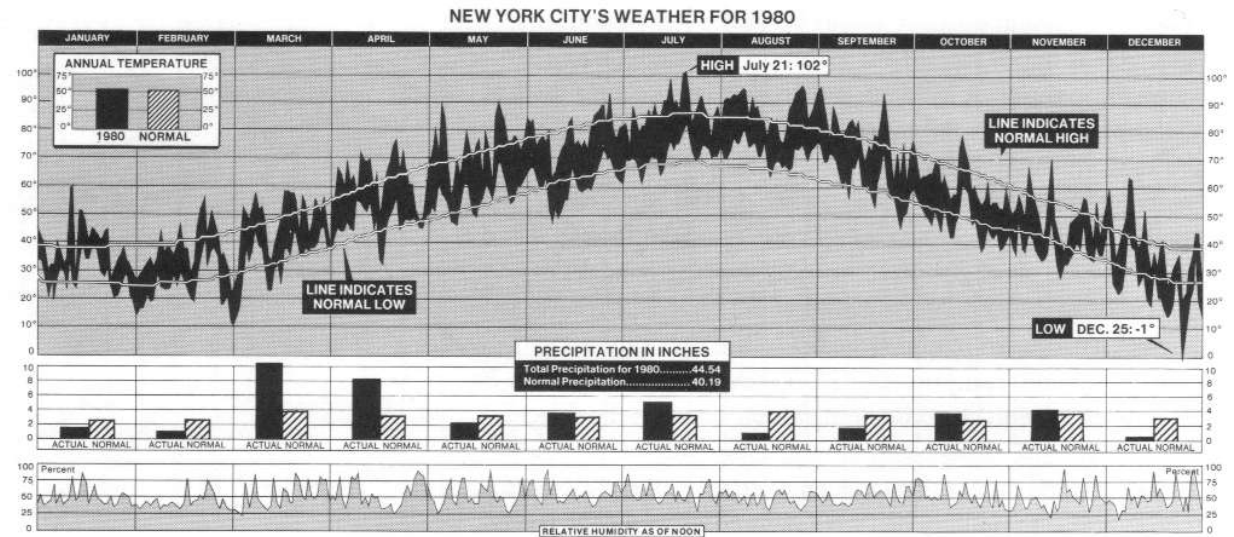
Visualization's Value



New York Times, January 11, 1981, p. 32.

$$V_{\text{alue}} = T + I + E + C$$

Visualization's Value



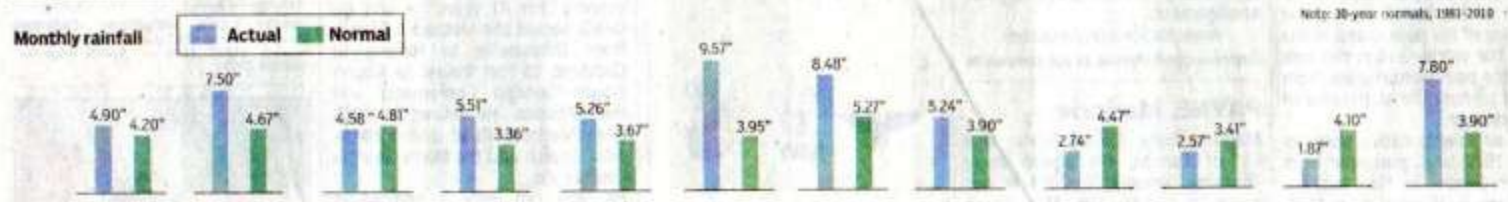
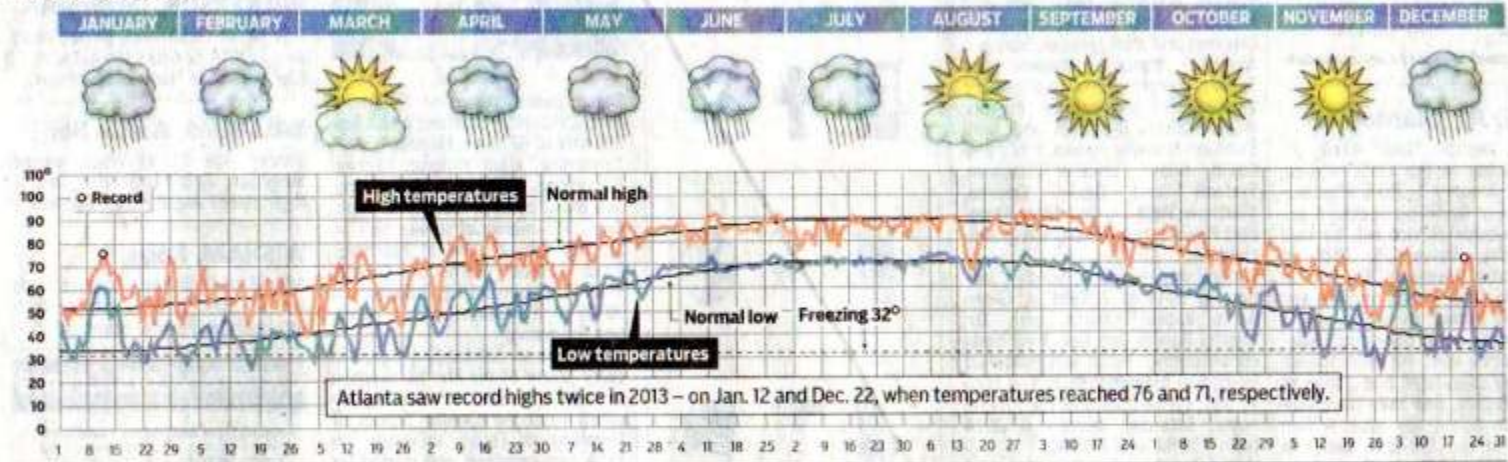
New York Times, January 11, 1981, p. 32.

$$V_{\text{alue}} = T + I + E + C$$

A month-by-month look at Atlanta's weather in 2013

Atlanta's 2013 weather can be summed up in one word – soggy. The city ended the year 16.31 inches above normal in rainfall, with much of that rain falling during the waterlogged summer months, which had about twice as many rainy days as dry ones. Rain fell into the city's official gauge at Hartsfield-Jackson International Airport on 17 days in June, 19 days in July and 24 days in August. The rainfall total for those three months was 23.29 inches, making the period the city's fourth wettest summer on record. Daily rainfall records were set on April 28 (1.73 inches), May 4 (3.40 inches), June 5 (4.14 inches), Aug. 8 (2.60 inches) and Oct. 6 (2.12 inches). Summer 2013 will also be remembered for a lack of extreme heat, particularly compared to the sweltering summer of 2012. Temperatures climbed into the 90s only 20 times, topping out at 92 on June 28 and Aug. 29. That's 32 less days of 90+ heat than in 2012, when the all-time record high of 106 was reached.

Daily temperatures



Sources: National Weather Service, National Drought Mitigation Center

Georgia drought fades during 2013

- Abnormally dry
- Moderate drought
- Severe drought
- Extreme drought
- Exceptional drought

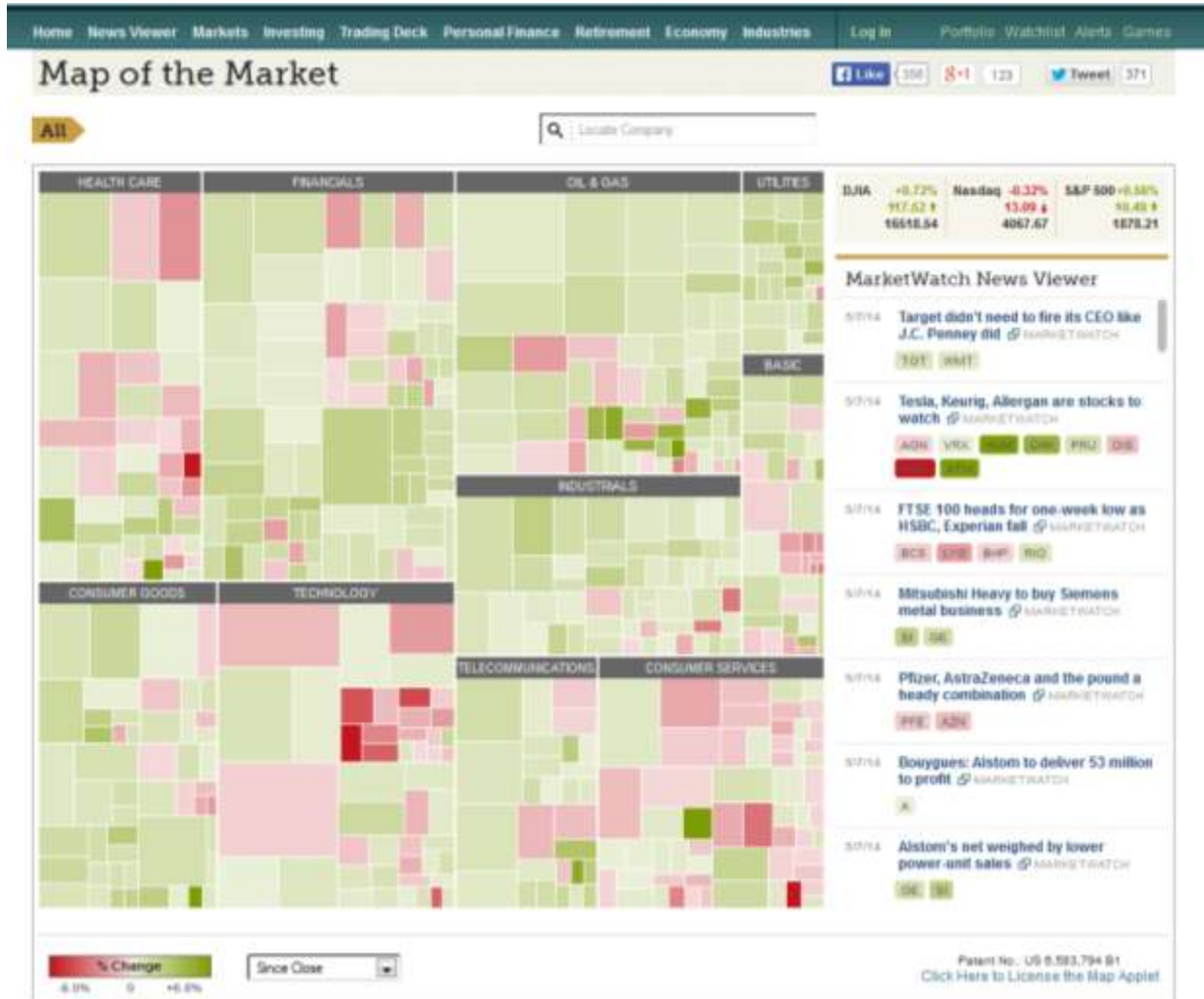


LINDA SCOTT, MICHAEL MORRIS / STAFF

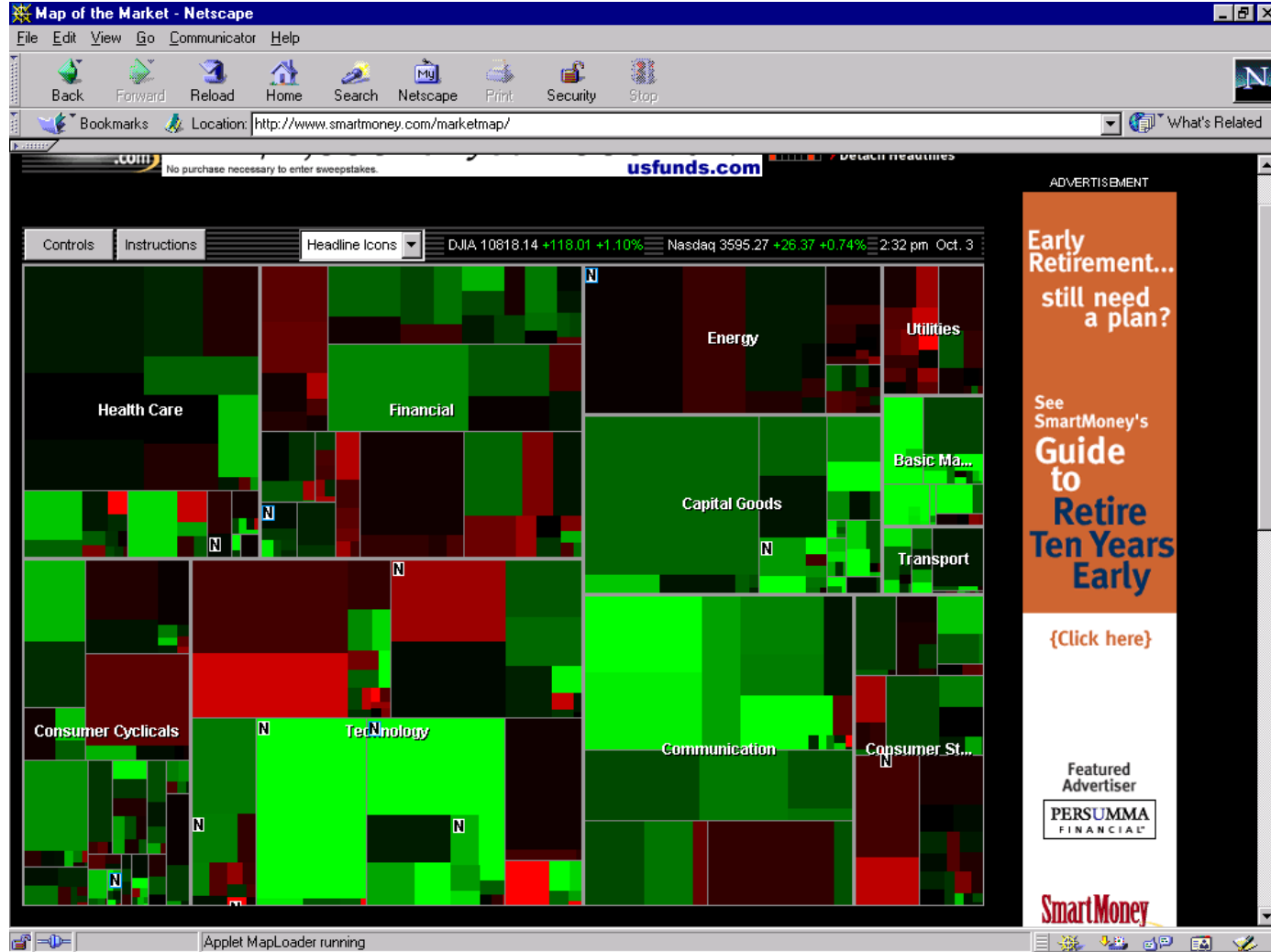
Atlanta Journal Constitution
Jan. 3, 2014

2.

Map of the Market

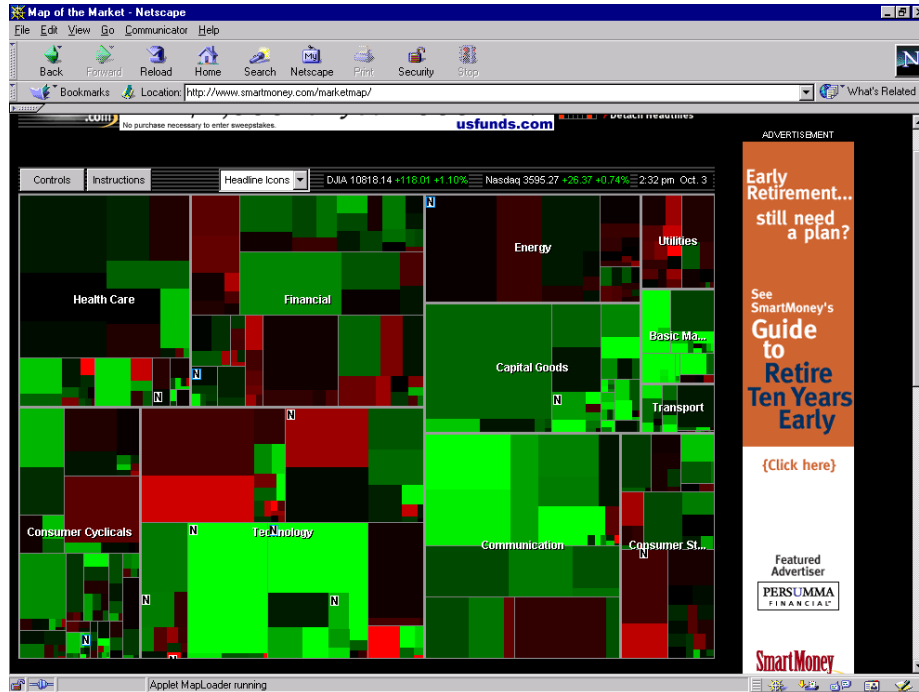


Map of the Market (old school)



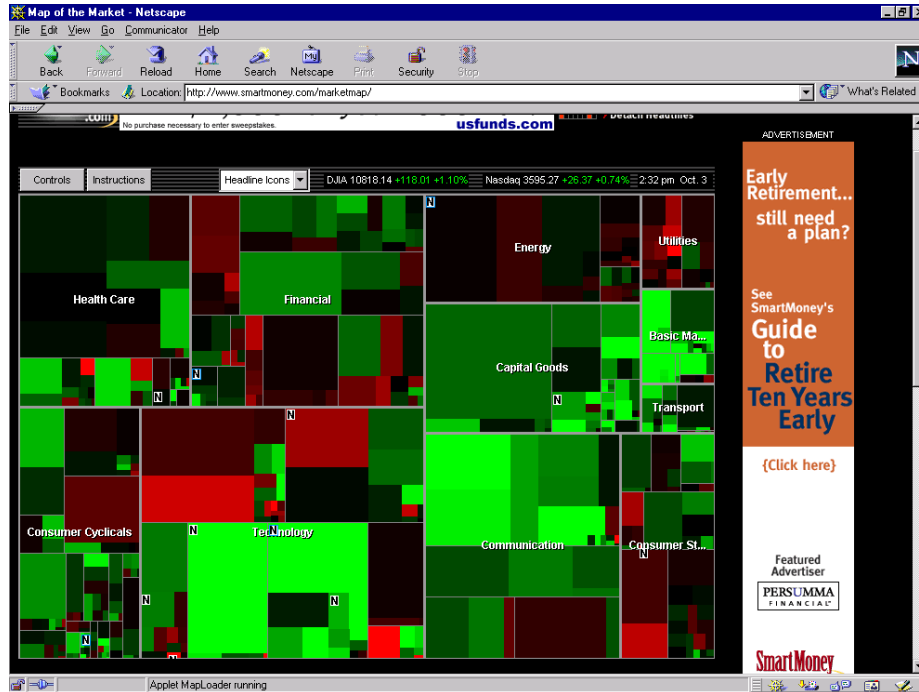
Wattenberg
CHI '99 EA

Visualization's Value



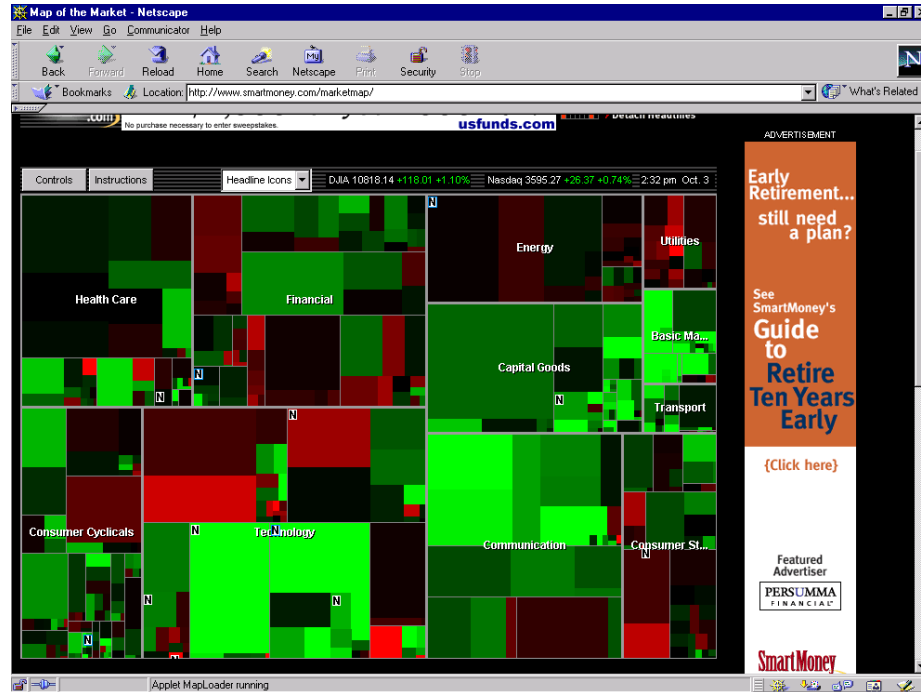
$$V_{\text{alue}} = T + I + E + C$$

Visualization's Value

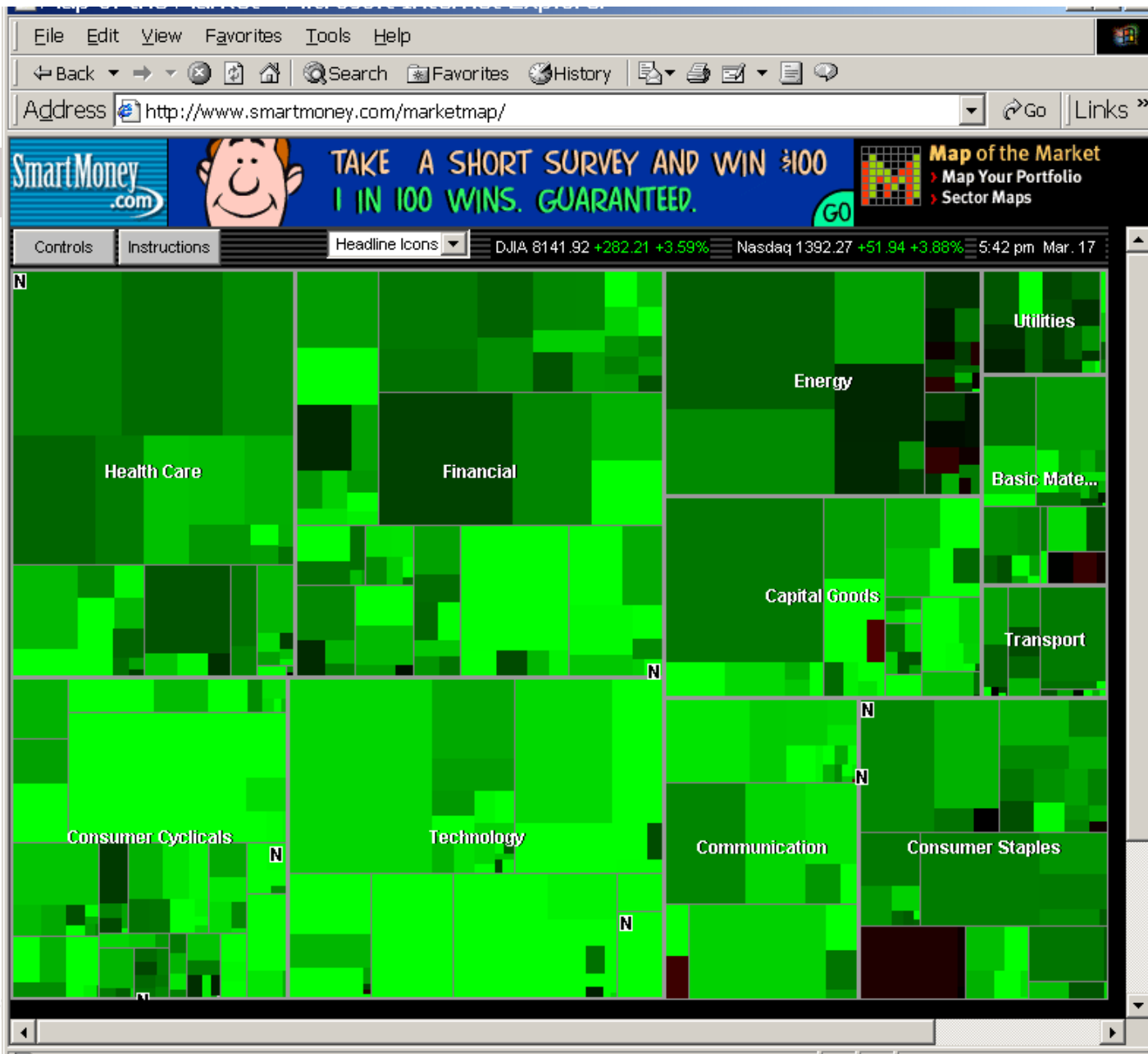


$$V_{\text{alue}} = T + I + E + C$$

Visualization's Value

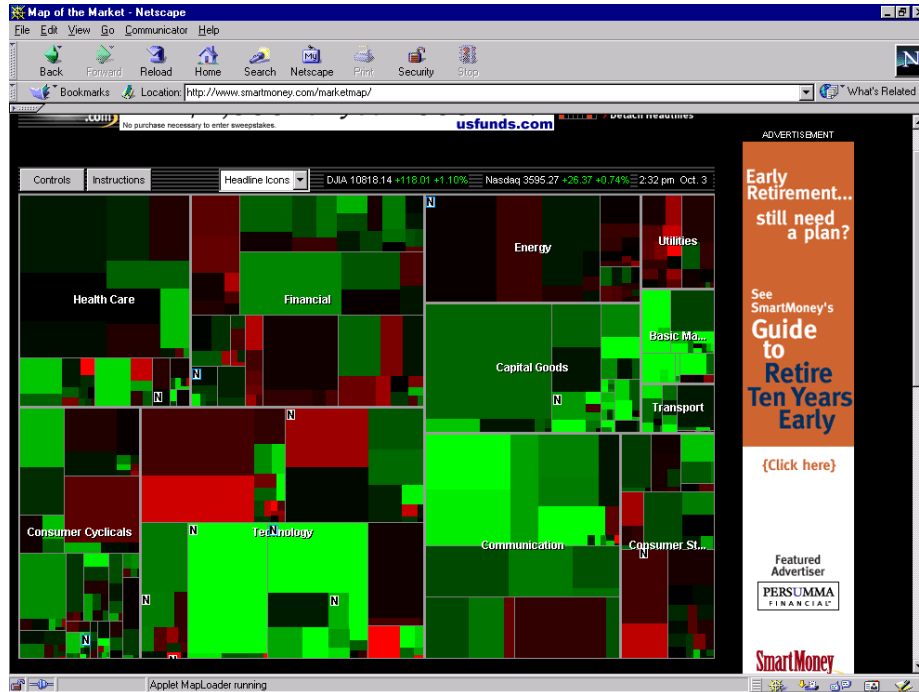


$$V_{\text{alue}} = T + I + E + C$$





Visualization's Value

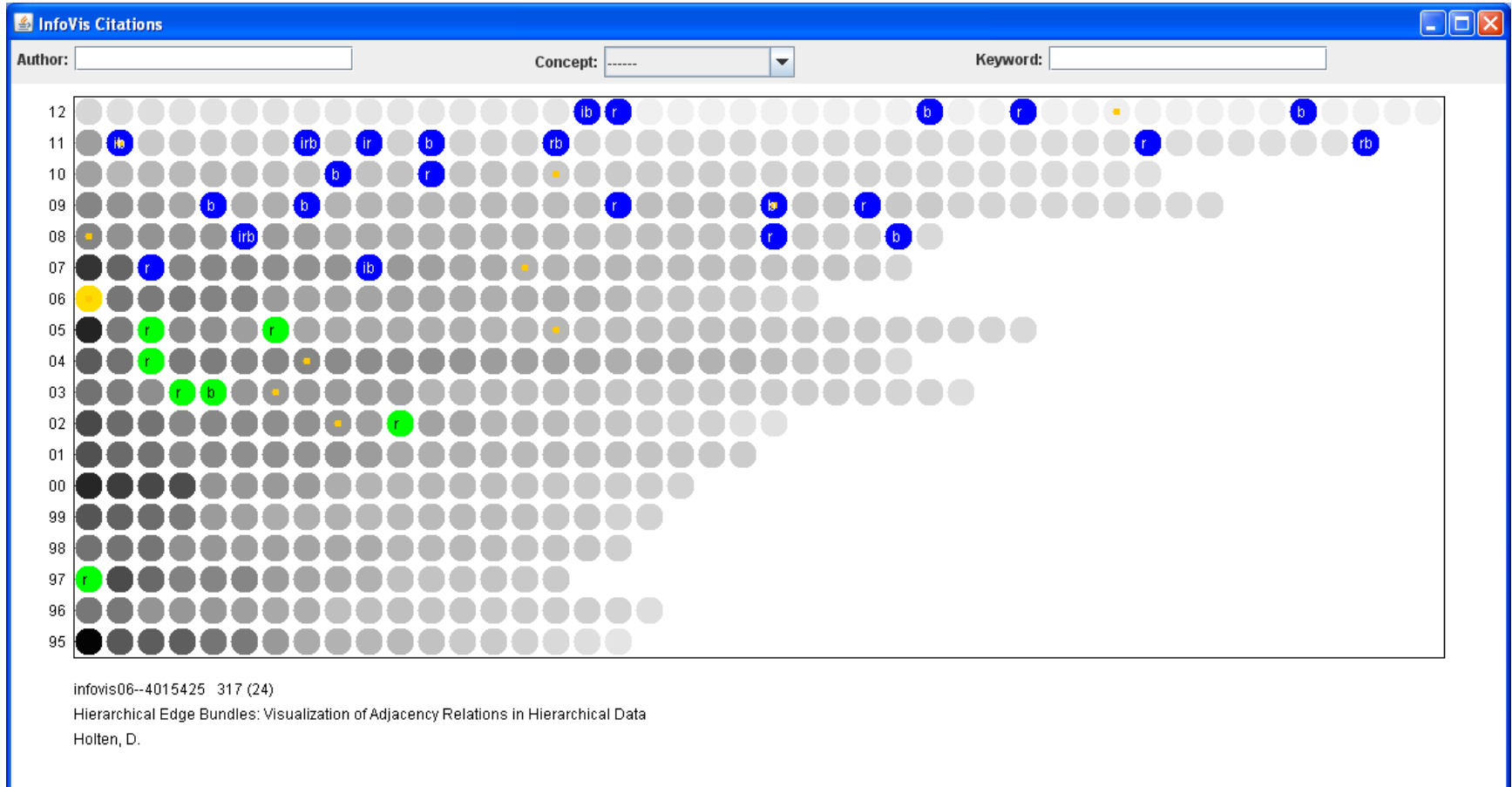


$$V_{\text{alue}} = T + I + E + C$$

3.

<http://www.cc.gatech.edu/gvu/ii/citevis>

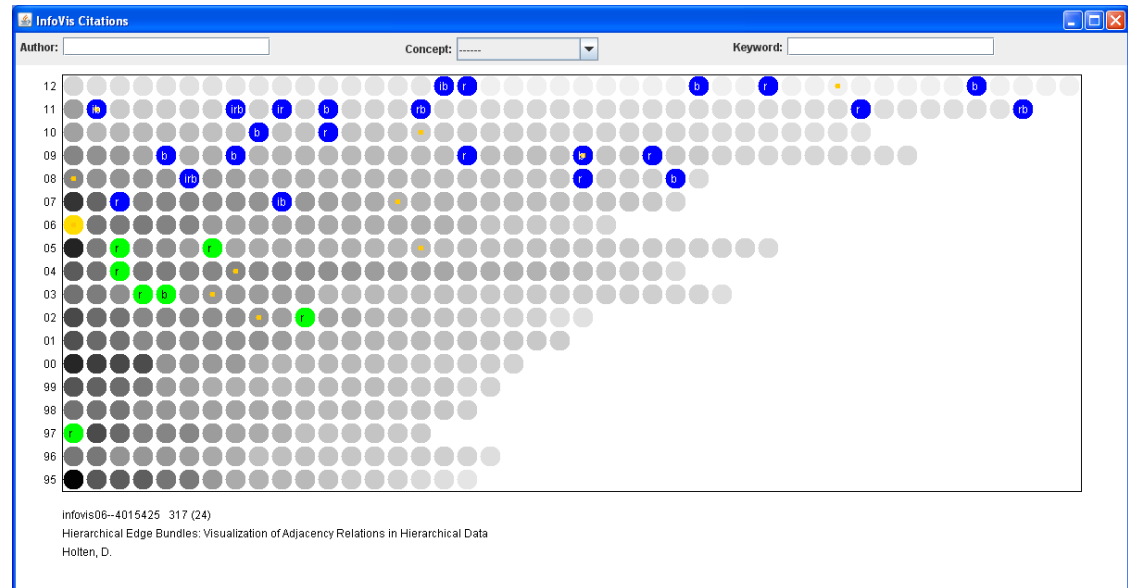
CiteVis



Demo

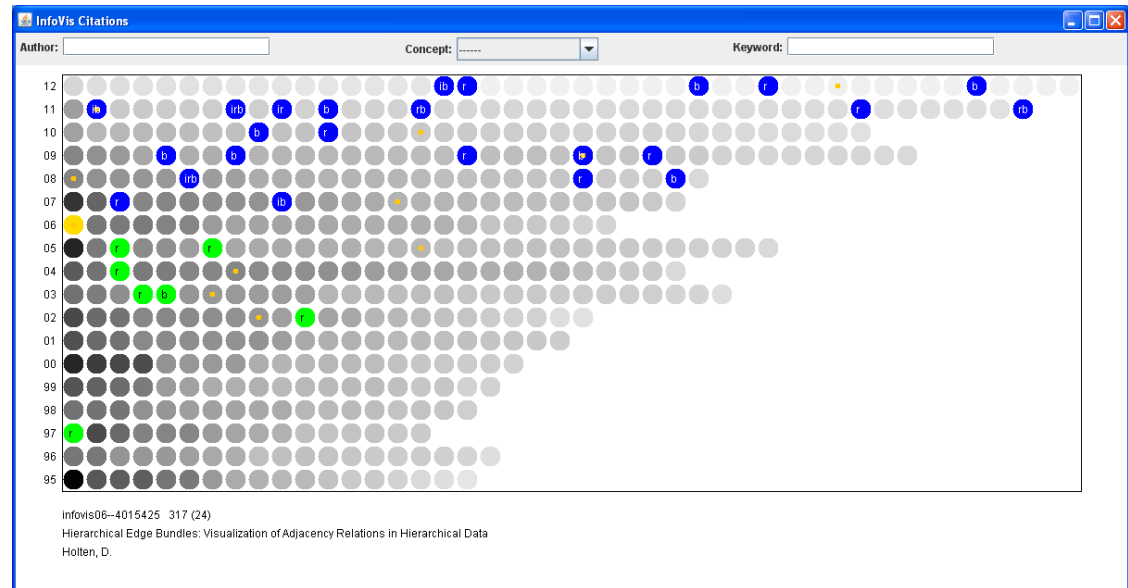
Stasko, Choo, Han, Hu, Pileggi, Sadana & Stolper
InfoVis poster '13

Visualization's Value



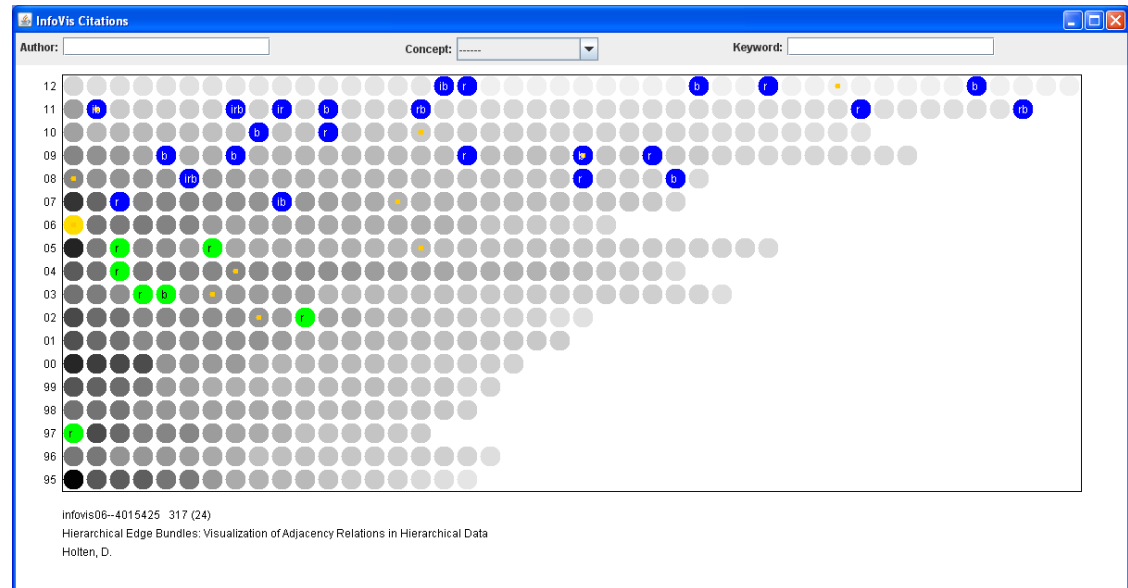
$$V_{\text{alue}} = T + I + E + C$$

Visualization's Value



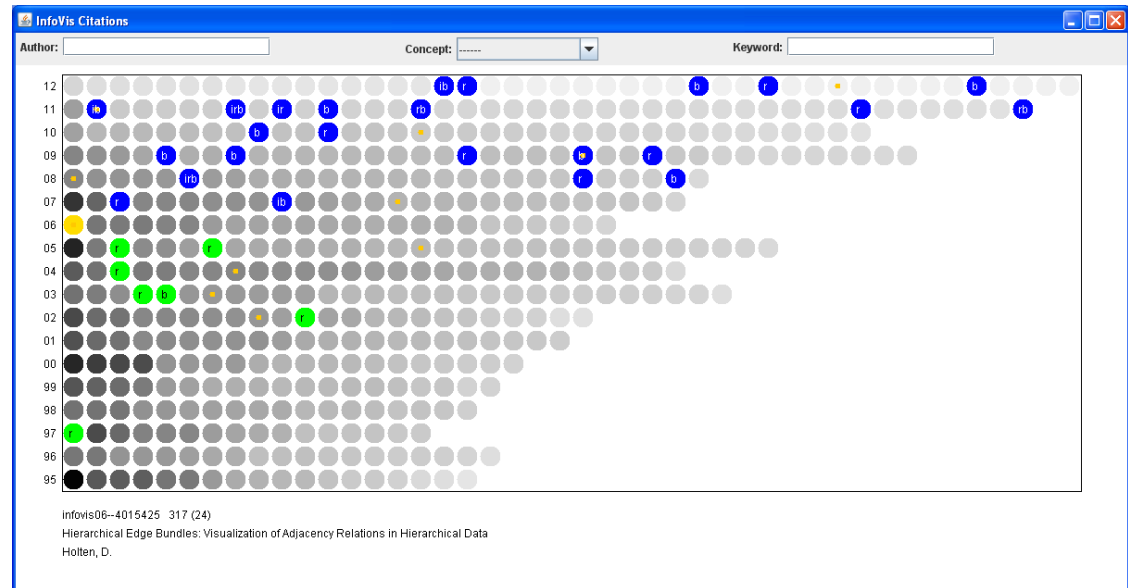
$$V_{\text{alue}} = T + I + E + C$$

Visualization's Value



$$V_{\text{alue}} = T + I + E + C$$

Visualization's Value



$$V_{\text{alue}} = T + I + E + C$$

Value

Not evaluation in traditional sense

HCI: Benchmark tasks

Visualization: Vagueness or absence of specific tasks
More about exploration and understanding



Pffft on evaluation.
Just build cool stuff.

STAR Report: Set Visualization

Elements		Element Attributes	
	Find elements belonging to a specific set		Find the attribute values of a certain element
	Find sets containing a specific element		Find the distribution of an attribute in a certain set or subset
	Find elements based on their set memberships		Compare the attribute values of sets or set intersections
	Find elements with a specific set membership degree		Analyze the set memberships for elements having certain attribute values
	Filter out elements based on their set memberships		Create a new set out of elements having certain attribute values
	Filter out elements based on their set membership degrees		
	Create a new set that contains certain elements		

Sets and Set Relations

	Find the # of sets in the set family		Find set intersections of a specific set
	Analyze inclusion relations		Find the set with largest pairwise set intersections
	Analyze inclusion hierarchies		Analyze set & set intersection cardinalities
	Analyze exclusion relations		Analyze and compare set similarities
	Analyze intersection relations		Analyze and compare set exclusiveness
	Find intersections between k sets		Highlight specific sets, subsets, or set relations
	Find sets involved in certain intersection		Create a new set using set-theoretical operators

SetViz.net
Visualizing Sets and Set-typed Data

Alsallakh, Micallef, Aigner, Hauser, Miksch & Rodgers
EuroVis '14

Data Visualization 101



Problem:

You have a lot of data (& attributes) to understand

Do you?

Pack all the data into one complex representation

Spread the data into multiple coordinated views

Use interaction to reveal different subsets of the data



Constituents

Two key aspects of data visualization

Representation

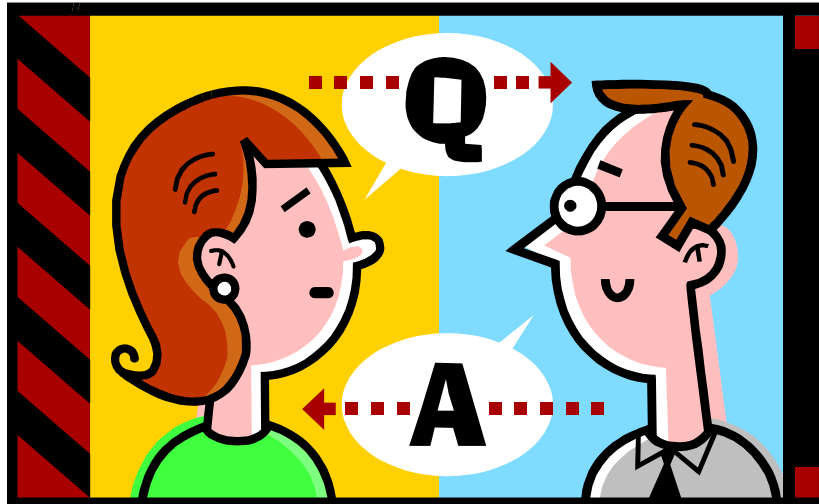
Interaction

“The effectiveness of information visualization hinges on two things: its ability to clearly and accurately represent information and our ability to interact with it to figure out what the information means.”

S. Few, Now you see it

Interaction is Vital

Engage in a dialog with your data



Fundamental nature: Equal sibling with representation or subordinate facilitator?

Interaction

Why interact?

Intent

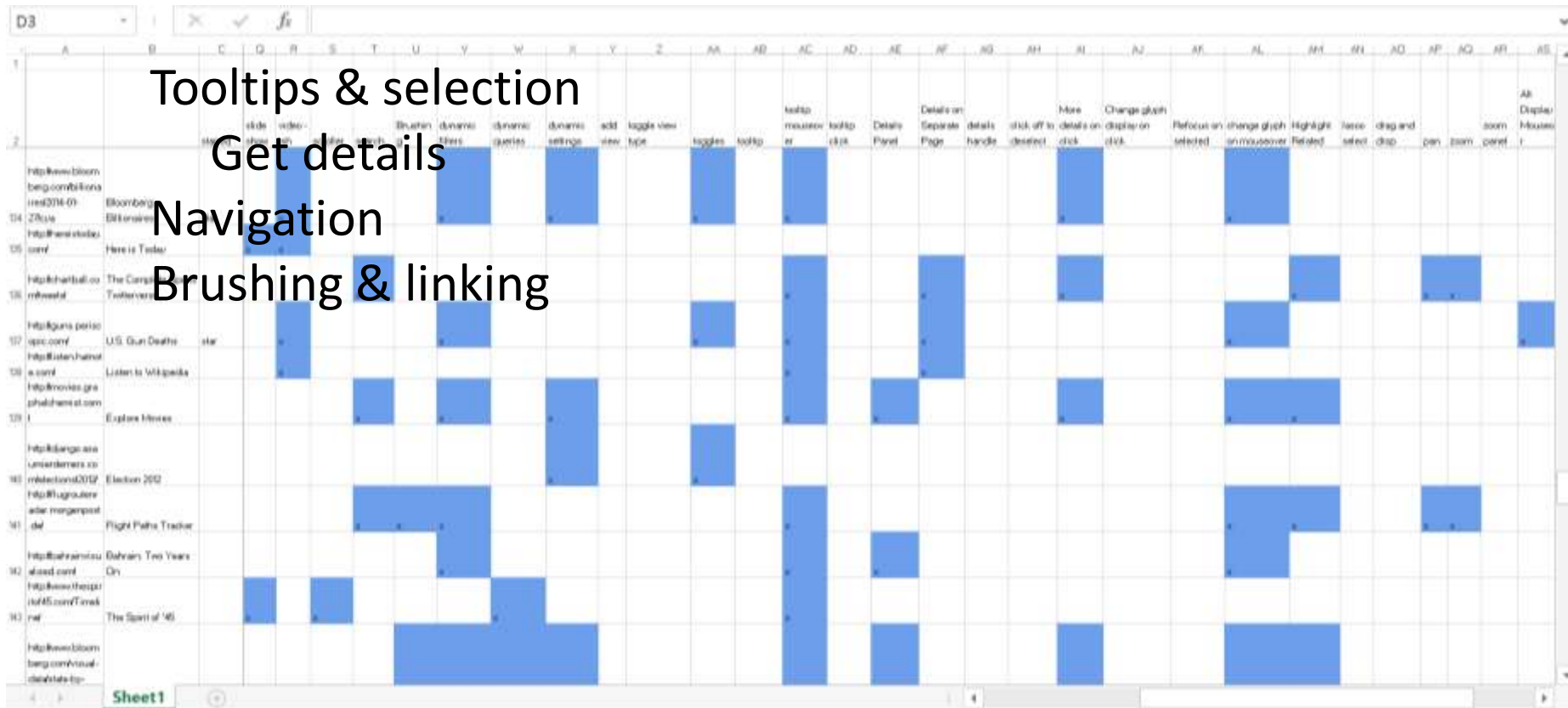
1. Select
2. Explore
3. Reconfigure
4. Encode
5. Abstract/Elaborate
6. Filter
7. Connect

Yi, Kang, Stasko & Jacko
TVCG (InfoVis) '07



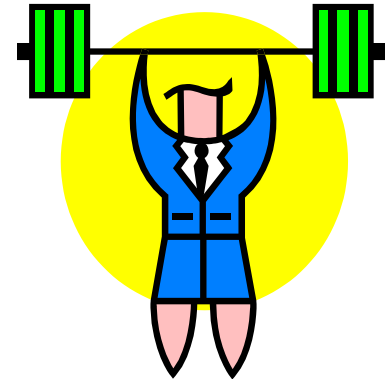
Interaction

How manifested today?

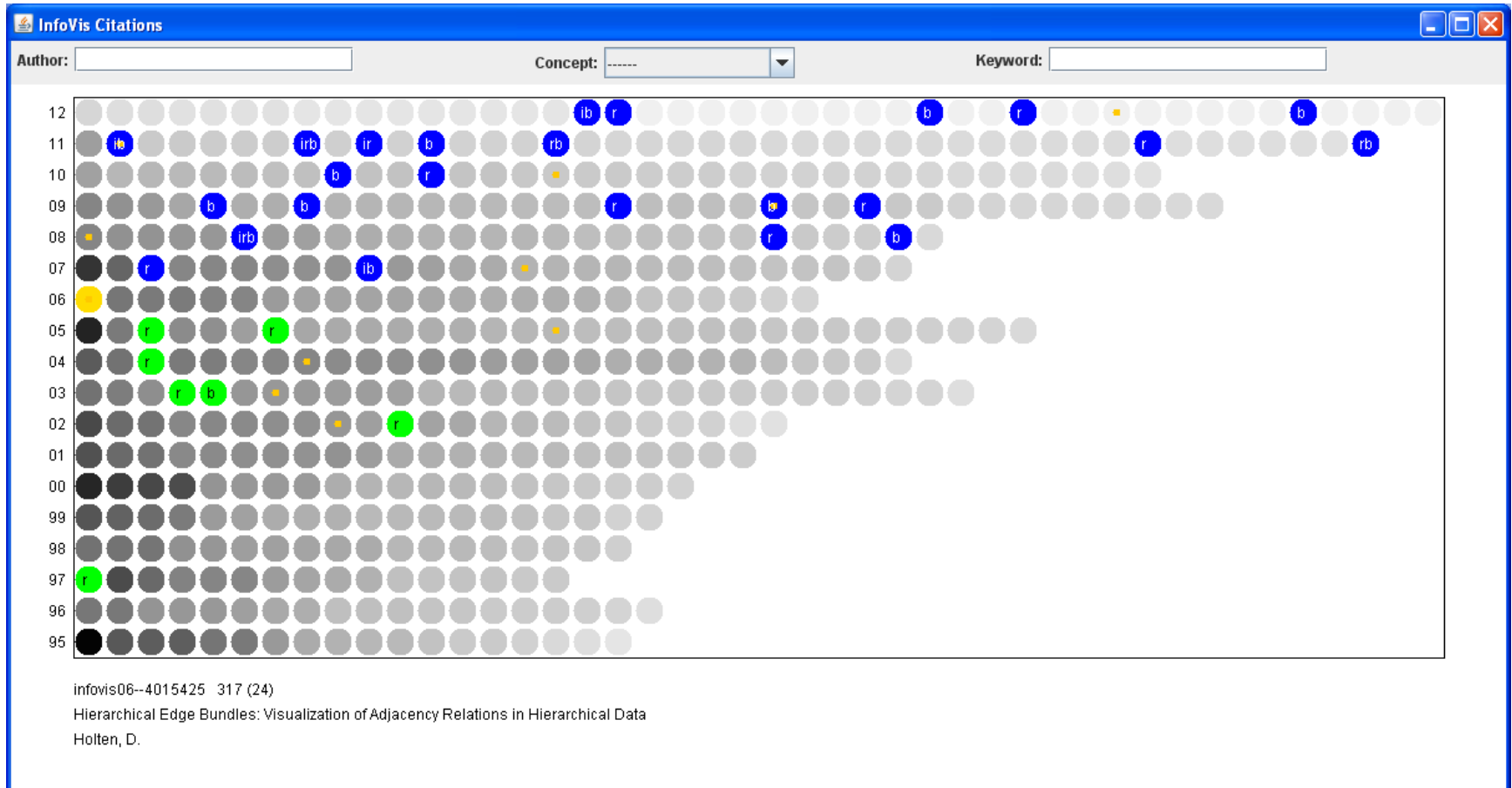


Interaction

Can we do more?

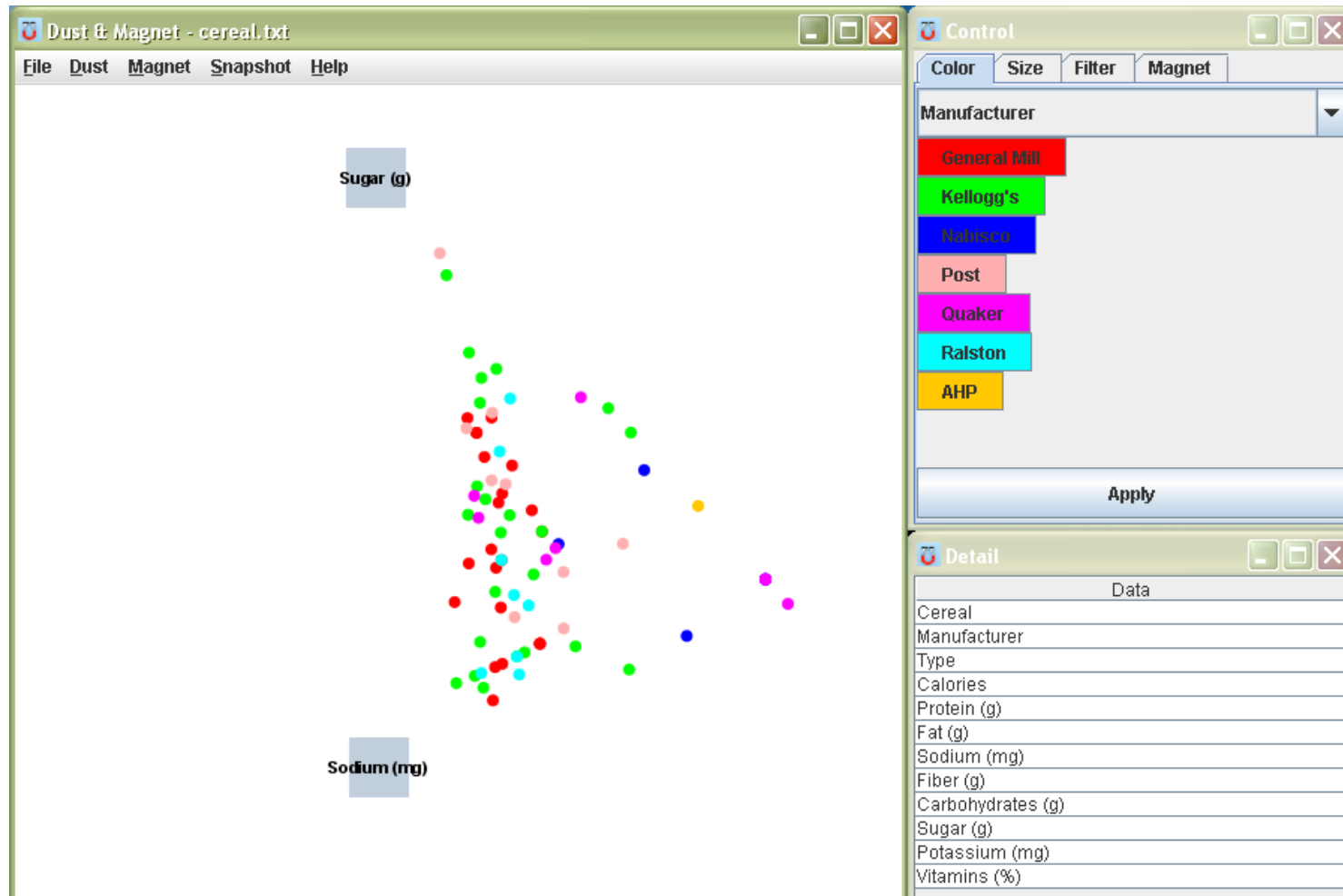


Employ interaction in a more fundamental manner to strengthen the power of visualization



Stasko, Choo, Han, Hu, Pileggi, Sadana & Stolper
InfoVis poster '13

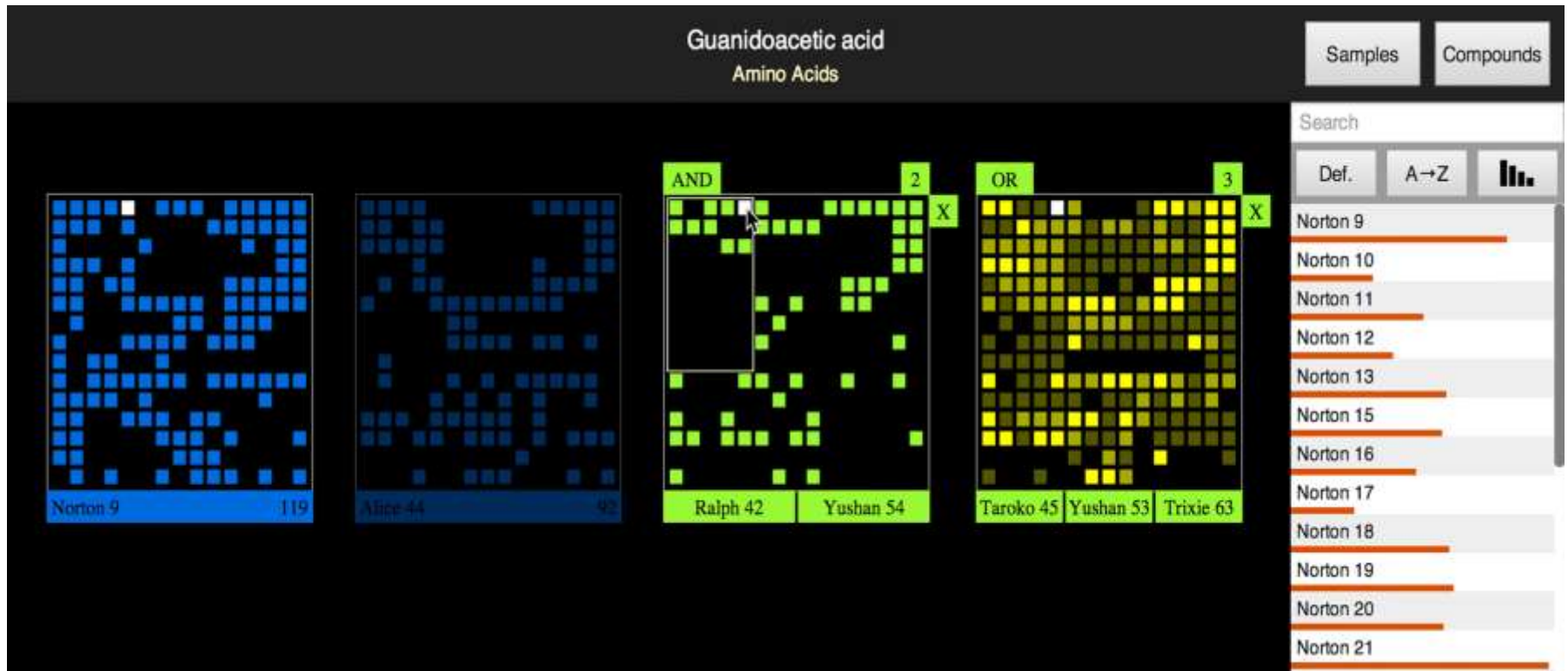
Dust and Magnet



Yi, Melton, Stasko & Jacko
Information Visualization '05

OnSet

<http://www.cc.gatech.edu/gvu/ii/setvis/>



Demo

Sadana, Major, Dove, & Stasko

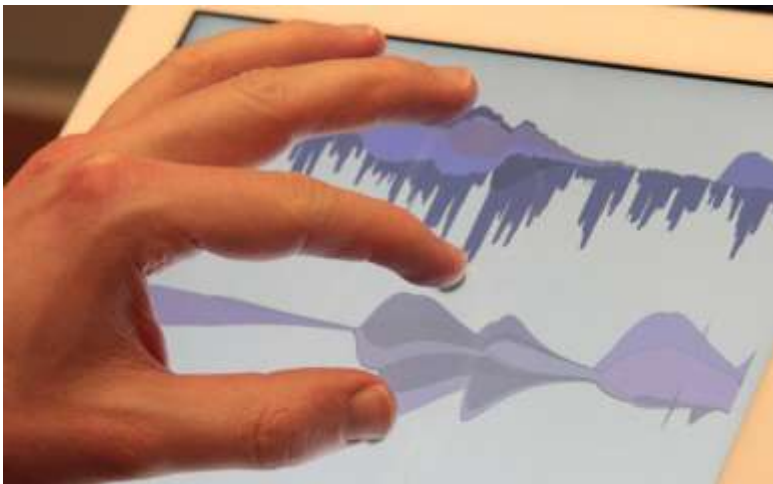
Interaction

What are the tools of interaction?

Traditional – Desktop: keyboard, mouse

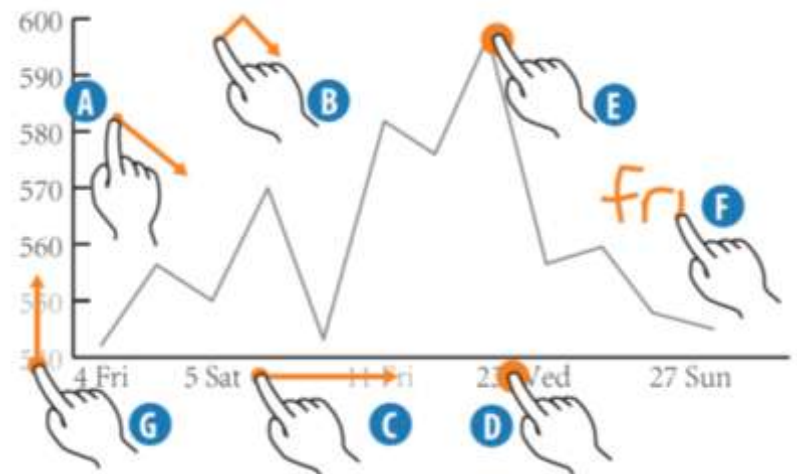
New – Tablet: fingers and multitouch

TouchWave



Baur, Lee & Carpendale
ITS '12

Selection Gestures

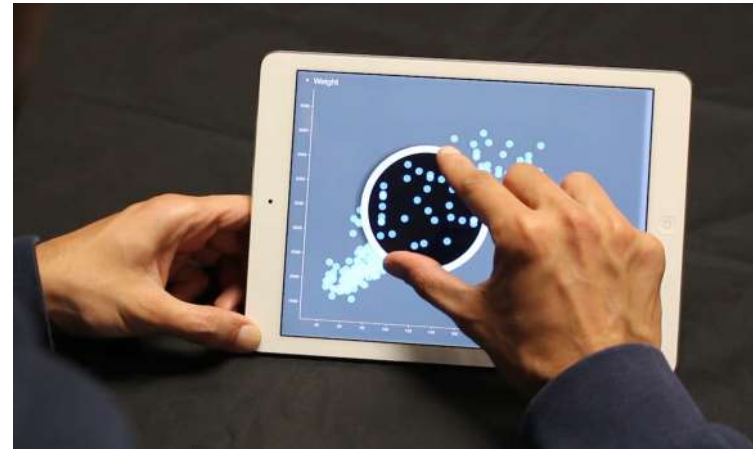
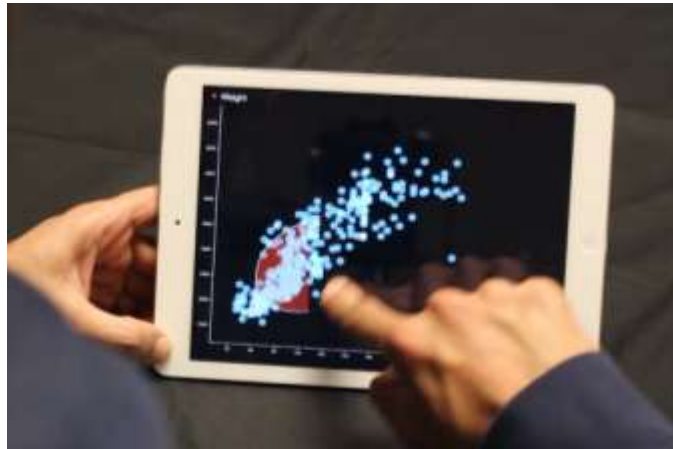
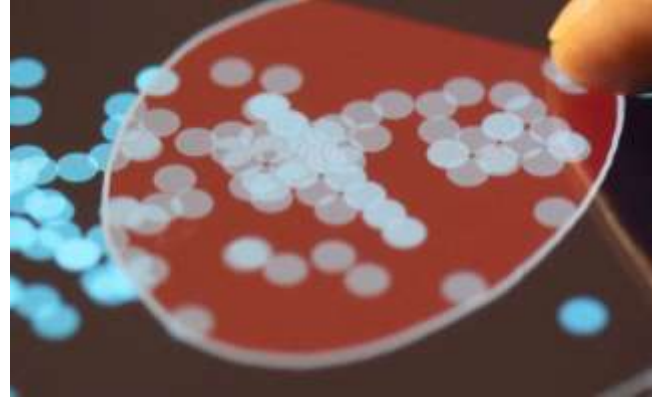


Willett, Lan & Isenberg
EuroVis '14



Moving to Tablets

Scatterplot



Video

Sadana & Stasko
AVI '14

Key Open Problems in Visualization

Key Open Problems in Visualization

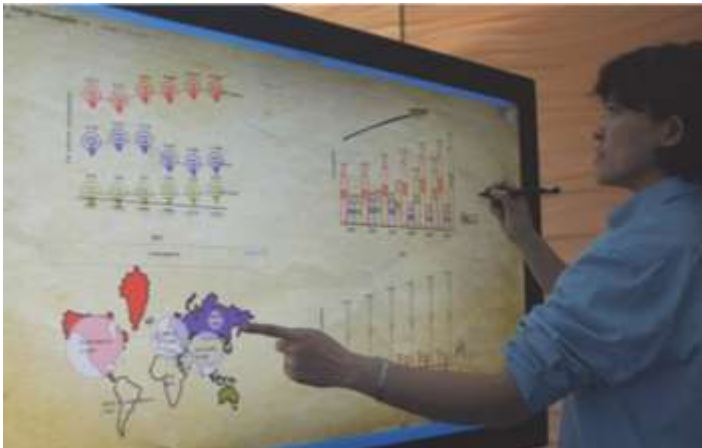
1. Assess and communicate **value**

$$V_{\text{alue}} = T + I + E + C$$

Key Open Problems in Visualization

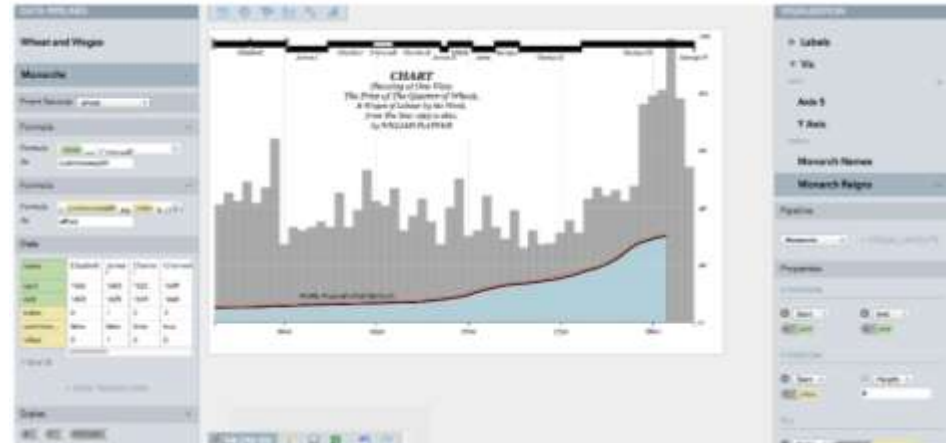
1. Assess and communicate value
2. Make the construction of visualizations easier

SketchStory



Lee, Kazi & Smith
TVCG (InfoVis) '13

Lyra



Satyanarayan & Heer
EuroVis '14

Key Open Problems in Visualization

1. Assess and communicate value
2. Make the construction of visualizations easier
3. Address real world “big” problems





Take Aways

Visualization is not for a few precise, concrete tasks
Exploratory data analysis

Presentation & analysis, related but different

Value = T + I + E + C

Interaction provides power, use it

And finally...



Thank you

Acknowledgments

- Supported by the DHS Center of Excellence in Command, Control & Interoperability (VACCINE Center)



- Supported by CCF-0808863 (FODAVA lead), NSF IIS-0915788, NSF IIS-1320537



National Science Foundation
WHERE DISCOVERIES BEGIN

- Supported by DARPA's XDATA program