

CS4460

InfoVis Overview

Alex Endert
Spring 2019

recall from last time

Data Overload

- Challenge: How to make use of the data
 - How do we make sense of the data?
 - How do we harness this data in decision-making processes?
 - How do we avoid being overwhelmed?



The Challenge

- Transform the data into information (understanding, insight) thus making it useful to people
- How do we let people understand phenomena, given the data?



Visualization

- Definition

- “The use of computer-supported, interactive visual representations of data to **amplify cognition**.”

From [Card, Mackinlay Shneiderman '98]

today, we'll cover

- more in-depth overview of InfoVis
 - some more definitions
 - examples
- 2 primary reasons people use InfoVises
 - Purposes/Uses of InfoVis
- some examples
 - we will dive into details of these in later lectures
- a case study of why good InfoVis design is important

Visualization

- Often thought of as **process of making a graphic** or an image
 - “show me the data!”
 - “I’ve got all this data, visualize it!”
- Really is a **cognitive process**
 - Form a mental image of something
 - Internalize an understanding
- “The purpose of visualization is **insight, not pictures**”
 - Insight: discovery, decision making, explanation

Main Idea

- Visuals help us think
 - Provide a frame of reference, a temporary storage area (think: long division)
- Cognition → Perception
- Pattern matching
- External cognition aid
 - Role of external world in thinking and reason (think: organizing items in your kitchen)
 - We use it everyday!

Larkin & Simon '87
Card, Mackinlay, Shneiderman '98

Expressed Well

“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

Part of our Culture

- “I see what you’re saying”
- “Seeing is believing”
- “A picture is worth a thousand words”

Purposes of InfoVis

- Two main uses of InfoVis
 - **Analysis** – Understand your data better and act upon that understanding
 - **Presentation** – Communicate and inform others more effectively

1. Analysis

- Given all the data, then
 - understand, compare, decide, judge, evaluate, assess, determine, ...
- Ultimately, it's about solving **problems**, answering **questions**, performing **tasks**, accomplishing **goals**
 - we will cover categories of goals/tasks/questions later in the semester

an aside - **When (not) to use InfoVis**

- Remember, there are many other techniques for data analysis
 - Statistics, data mining, machine learning, DB querying/searching
 - These give you explicit answers/results
 - E.g., the average temperature in March in Atlanta is 74.82 degrees
- Good when you know what questions to ask
 - and when they have a known, well-defined stop rule

When to use InfoVis

- Visualization is most useful in **exploratory data analysis**
 - Don't know exactly what you're looking for
 - Don't have a priori questions
 - Want to know what questions to ask
 - Want to "get a feel for the data" (aka "overview")
- Should be open to **let yourself discover**, but also explain and answer specific questions

*"A graphic display has many purposes but it achieves its highest value when it **forces us to see what we were not expecting.**"*

H. Wainer

EDA Example 1

- Business
 - Did a recent marketing campaign work as intended?
 - How do sales reflect this? By region? By demographic? Compared to norms?
 - Did their product improve in quality, or just perceived quality? What defines quality?
 - How does their product compare to competitors?

EDA Example 2

- Airlines
 - What are the key factors causing flight delays in the US?
 - Are delays worse in the summer or winter?
 - Is the seasonal effect influenced by geographic location?
 - How does competition at an airport affect flight delays?

EDA Example 3

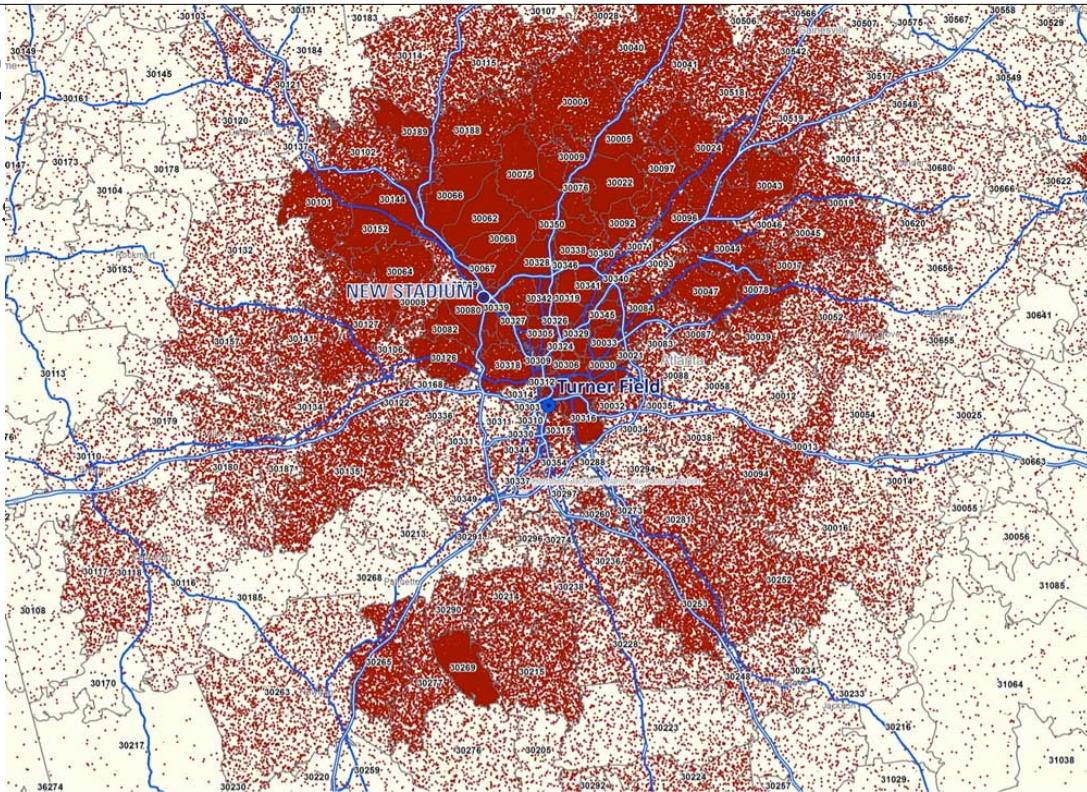
- House Hunting
 - Why do you “like” some homes, and not others? # of bedrooms, baths, price // smell? feel of area?
 - What will happen to the market? Is the home a “good deal”?
- Interesting thought:
Could an AI explicitly tell me what house to buy? What is the optimization function? Would you trust it?
 - Which car to purchase? Which TV show to watch? What courses to take?
 - In these cases, some vis is helpful. The decision is very important.

EDA Example 4

- Where should the Atlanta Braves move their stadium to increase attendance/profit?

EDA E

- Where is it?



Analytic Tasks for InfoVis

- **Search (OK)**

- Finding a specific piece of information
- *How many games did the Braves win in 1995?*
- *What novels did Ian Fleming author?*

- **Browsing (Better)**

- Look over or inspect something in a more casual manner, seek interesting information
- *Understand how baseball works*
- *What has Jane been up to lately?*
- *How can I explain why a decision was made?*

Analytic Tasks for InfoVis

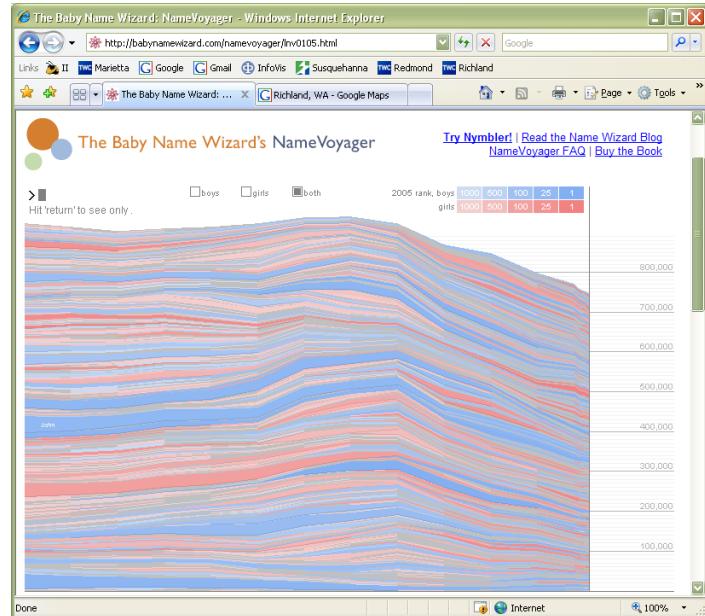
- **Analysis**

- Comparison-Difference
- Outliers, Extremes
- Patterns

- Monitoring - what has changed? what changes matter?
- Awareness (of something happening // the impact if something happens)
- lots more on tasks later in the semester

<http://babynamewizard.com/namevoyager/>

let's try it! Baby Name Wizard



let's discuss

- Chat with 2 or 3 people near you
- How does this vis support analysis?
- Take 3 minutes.
 - Be ready. I might call on you!

2. Presentation

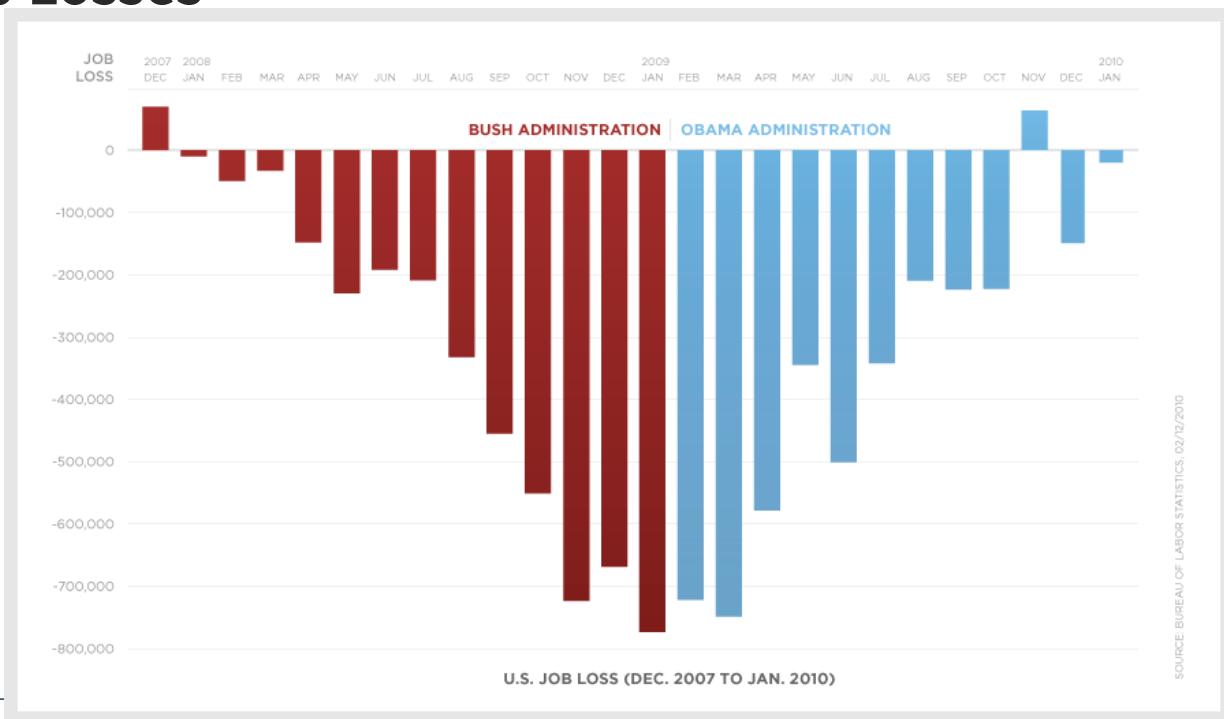
- Use visualization to **communicate** ideas, influence, explain, persuade
- Visuals can serve as evidence or support
- Tell a story
 - we will look at the famous Gapminder video later in the semester as an example of doing this really well

When to Apply?

- Visuals can frequently take the place of many words
- Visuals can summarize, aggregate, unite, explain, ...
- Sometimes words are needed, however

Job Losses

Controversial, see <http://soquelbythecreek.blogspot.com/2010/02/what-does-obama-job-chart-really-mean.html>

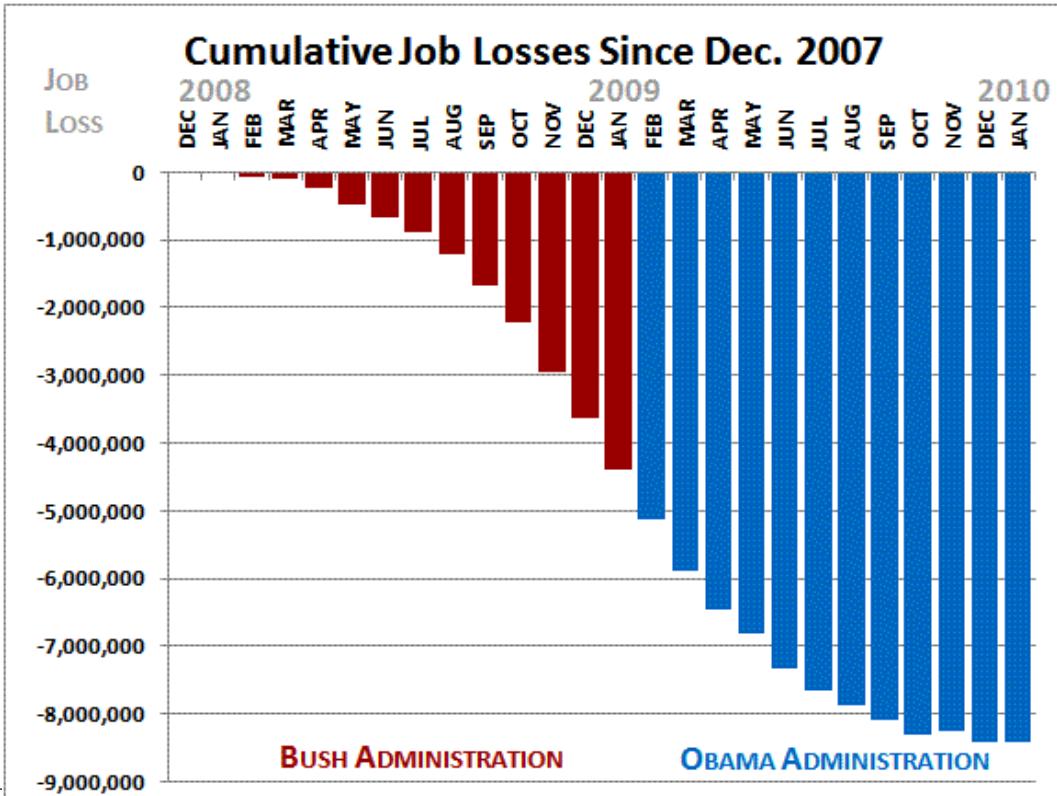


SOURCE: BUREAU OF LABOR STATISTICS, 02/12/2010

dert@gatech.edu

Georgia Tech

same data, different story

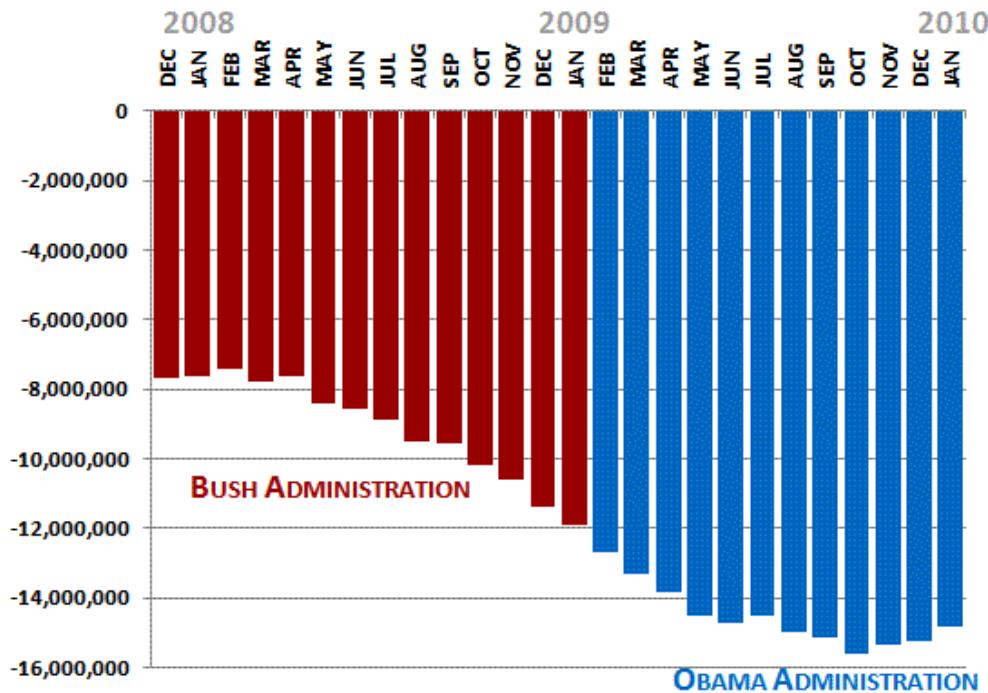


gatech.edu

Georgia Tech

when looking at totals
(baseline)

Total Unemployment Since Dec. 2007



endert@gatech.edu

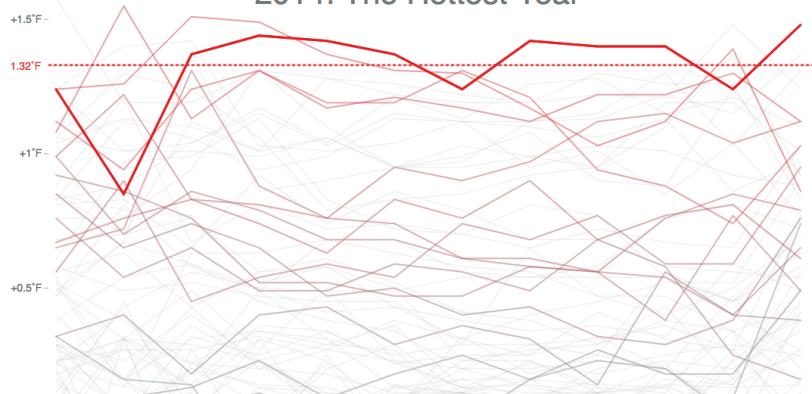
Storytelling, Narration, Animation

2014 Was the Hottest Year on Record

By Tom Randall and Blacki Migliozzi | January 16, 2015
Deny this. The **animation below** shows the Earth's warming climate, recorded in monthly measurements from land and sea over 135 years. Temperatures are displayed in degrees above or below the 20th-century average. Thirteen of the 14 hottest years are in the 21st century.
Update: see new records in 2015 and 2016.

[Record Years](#) [All Years](#) [See It Again](#)

2014: The Hottest Year



<https://www.bloomberg.com/graphics/2014-hottest-year-on-record/>

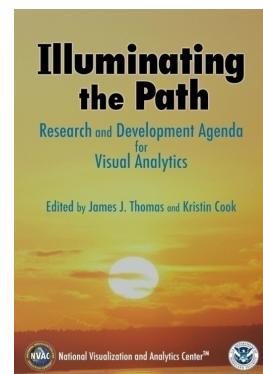
Georgia Tech

Key Benefits of Visualization — stepping back

- Facilitating awareness and understanding
- Helping to raise new questions and supply answers
- Generating insights
- Telling a story faithfully, and making a point

New Area Emerging: Visual Analytics

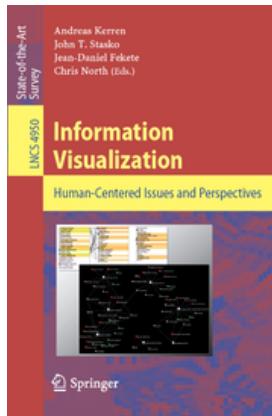
Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces



Available at <http://nvac.pnl.gov/>
in PDF form

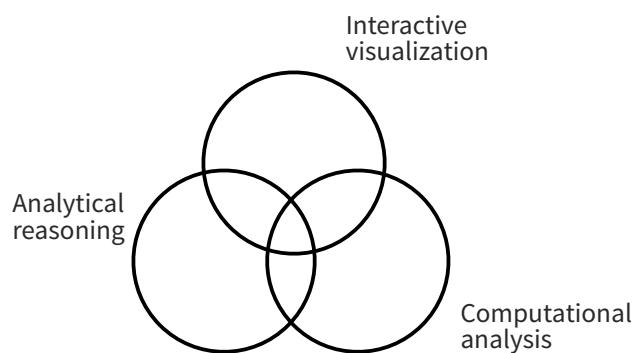
Alternate Definition

- Visual analytics **combines automated analysis techniques with interactive visualizations** for an effective understanding, reasoning and decision making on the basis of **very large and complex data sets**



Keim et al
chapter in Information Visualization: Human-Centered
Issues and Perspectives, 2008

Main Components of VA



We will cover VA in either 1 or 2 lectures during the semester

So,

- What are the 2 main **purposes** of InfoVis?

So,

- What are the 2 main purposes of InfoVis?
 - 1. Analysis
 - 2. Presentation

Components of InfoVis

- Two key components of InfoVis

- **Representation**

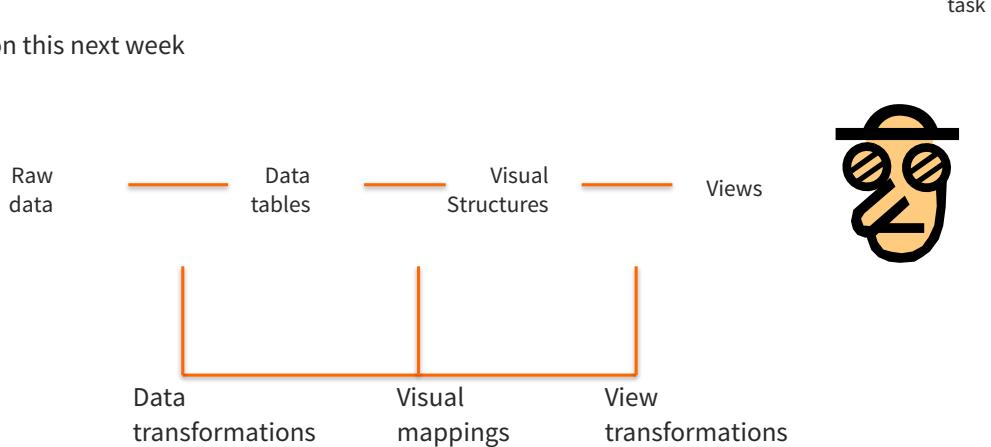
- **Interaction** (too often overlooked)

“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

InfoVis Process Model

more on this next week



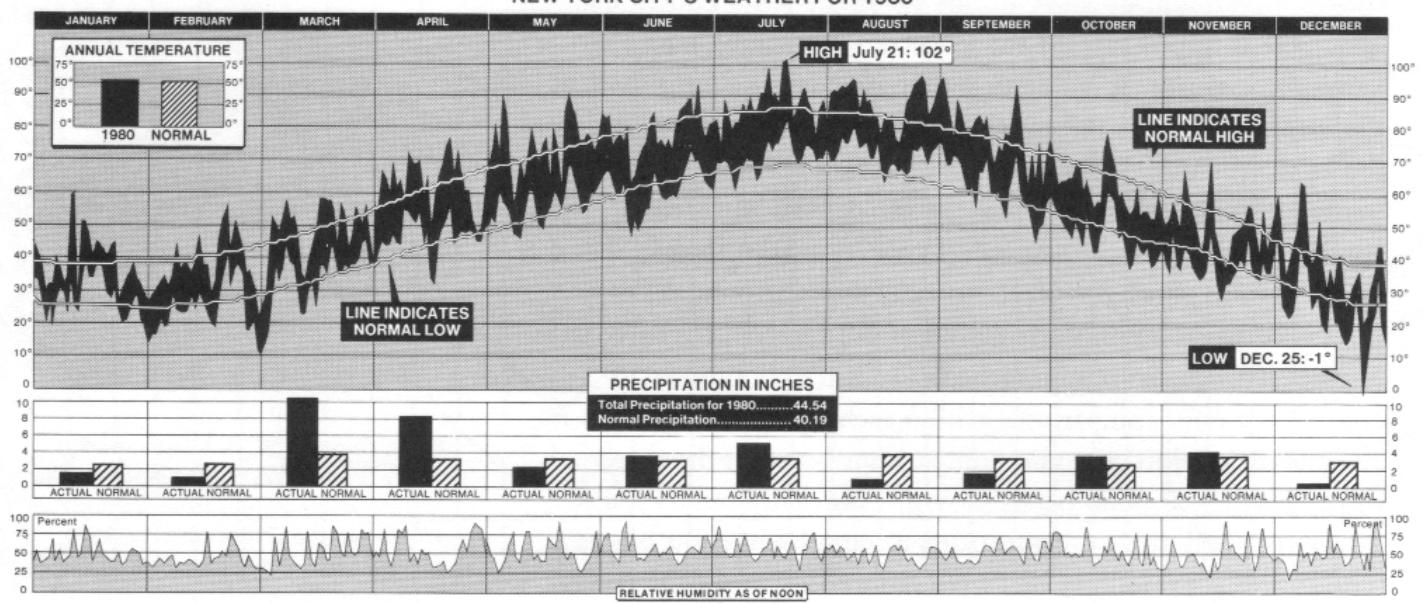
From: Card, Mackinlay, Shneiderman '99

InfoVis (a few Examples)

NYC Weather

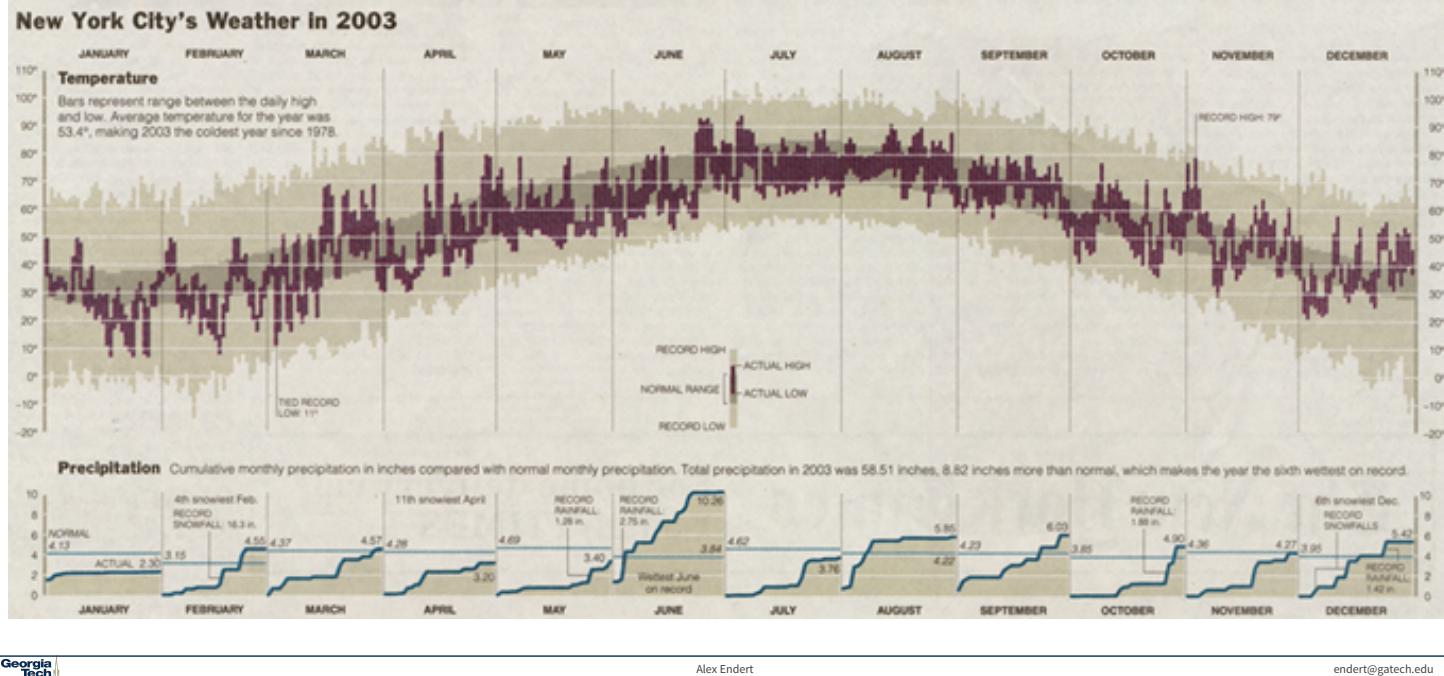
2220 measurements

NEW YORK CITY'S WEATHER FOR 1980



Updated Version

http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=00014g



Soccer

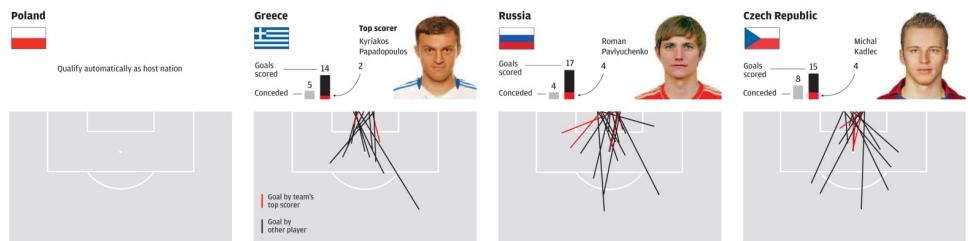
Simon Scarr
South China
Morning Post

All of the goals

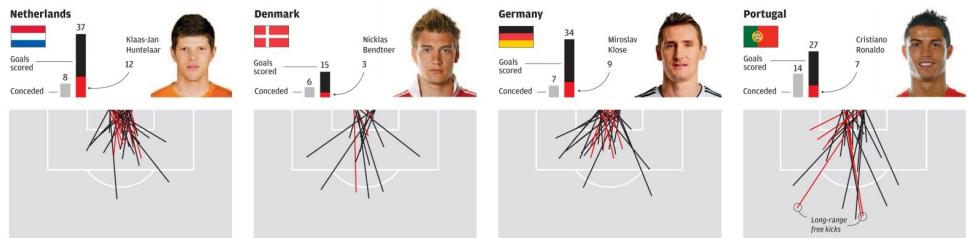
which took the teams to Euro 2012

Euro 2012 kicks off tomorrow night and promises to keep soccer fans glued to their television sets for the rest of the month. This 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



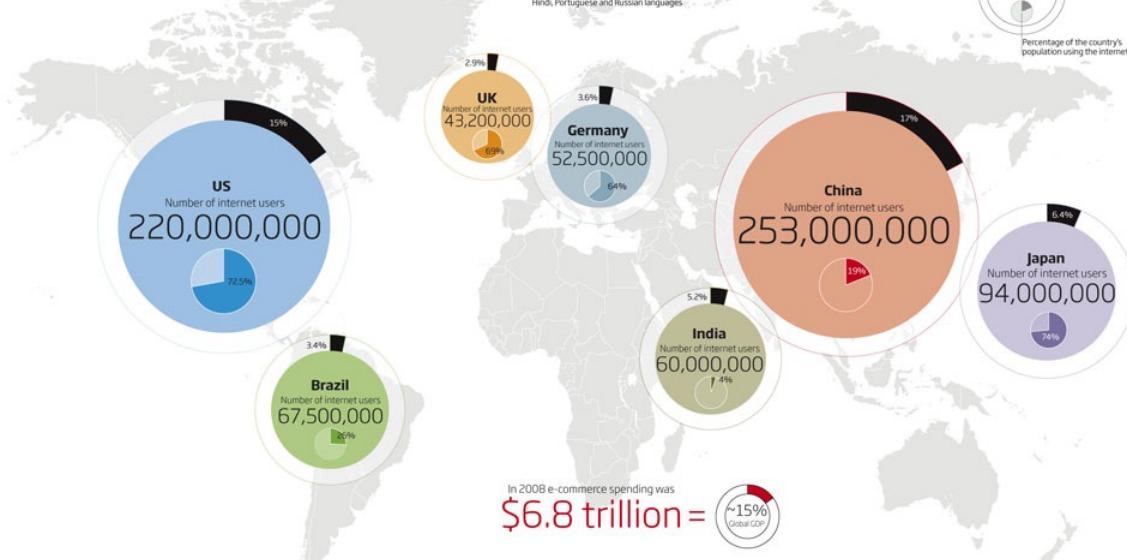
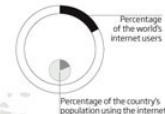
More Internet — getting more InfoGraphic-y

The exploding internet 2008

The exploding Internet
2008

Internet traffic is growing by about 50 per cent annually, with video and music streaming rising fastest. At the same time, the web is becoming divided along language lines

By 2012, Asian web surfers, including about 490 million Chinese, will outnumber North Americans by 3 to 1 and Indians will become the third-largest group online. Tomorrow's web will probably be dominated by a mixture of the English, Mandarin, Hindi, Portuguese and Russian languages.



http://www.newscientist.com/data/images/ns/cms/mg20227062.200/mg20227062.200-6_1000.jpg

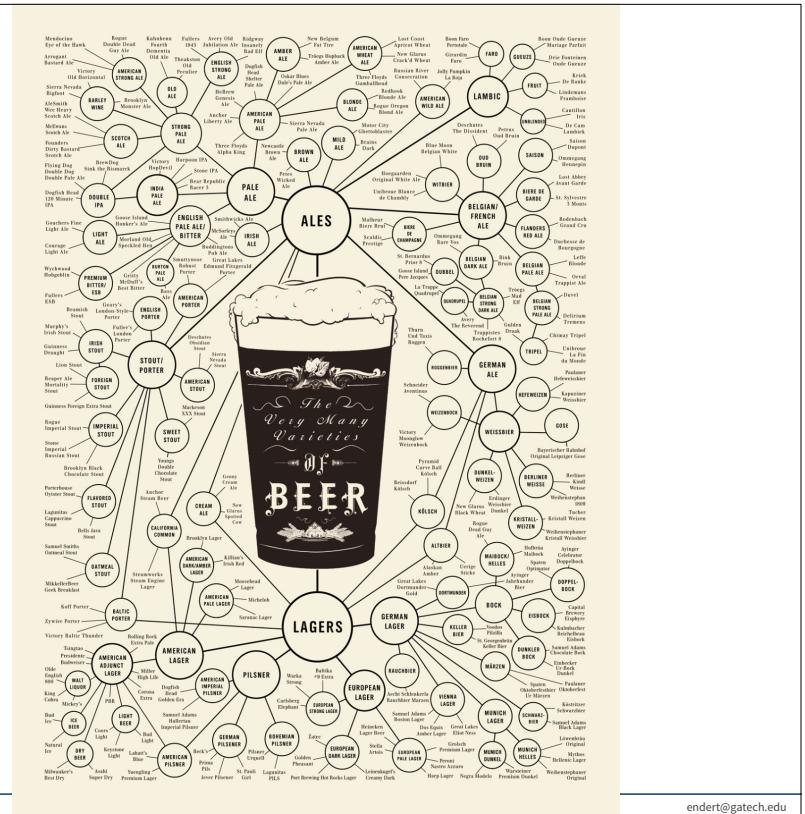
endert@gatech.edu

Beer



note how there are visual artifacts shown that do not directly map to data (InfoGraphic)

Beer!



http://images.fastcompany.com/upload/poster_beer_1300.jpg

endert@gatech.edu

More Beer! <http://thebeermongers.com/beers/>

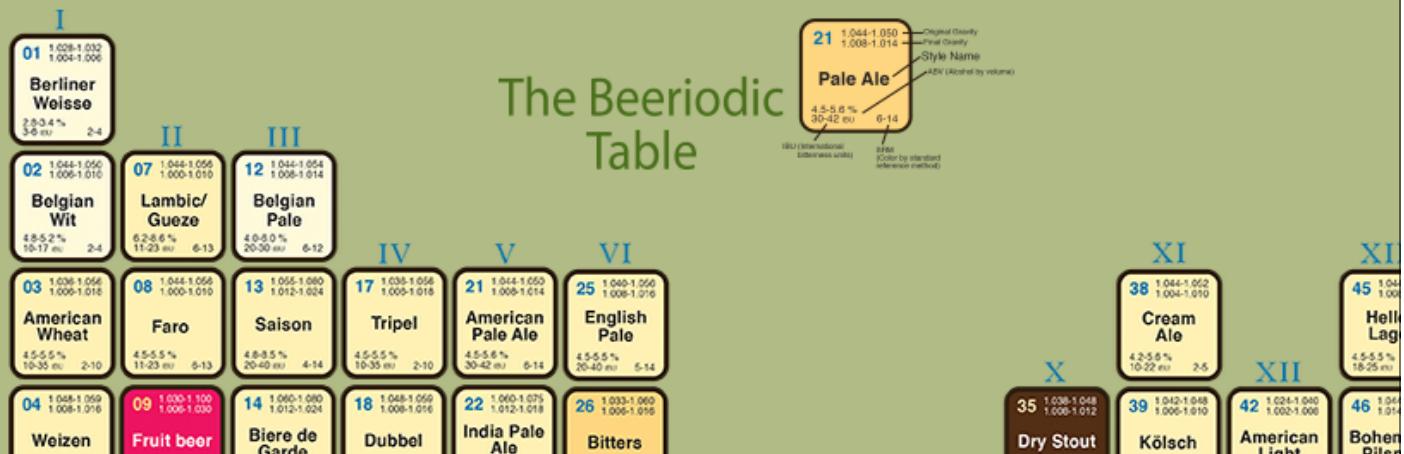
Home

Events

Beers

MONGERS

Click on any image below to see full description

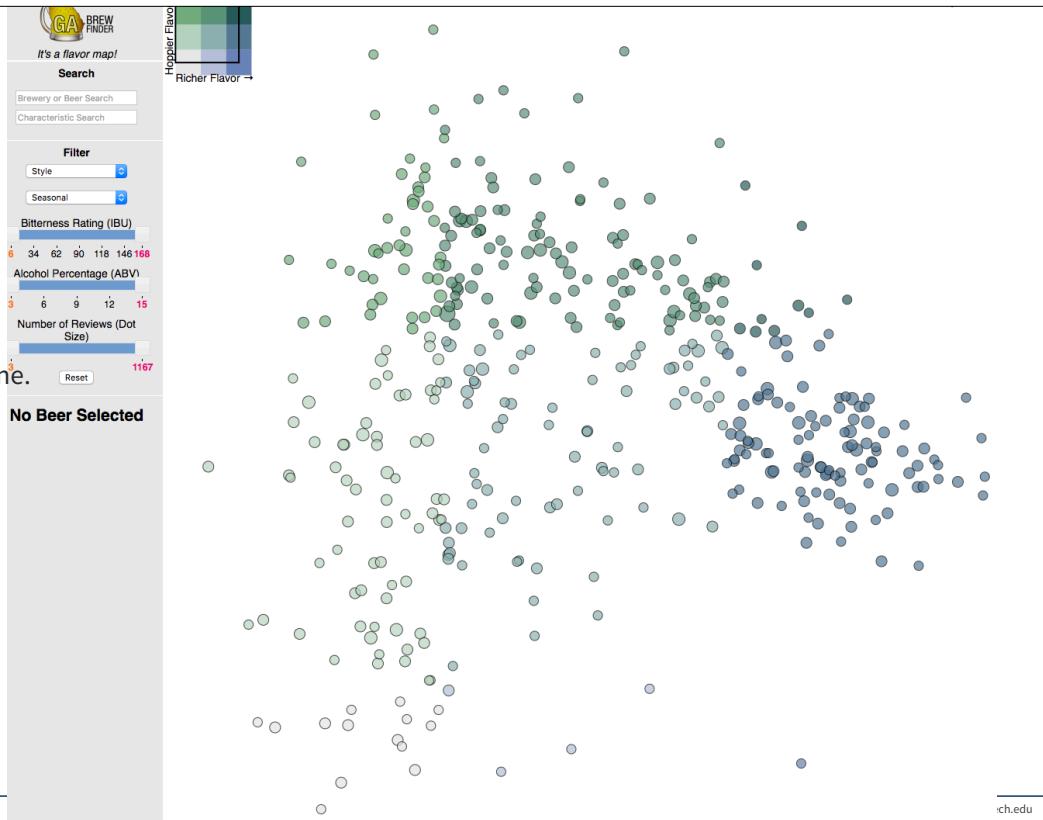


more beer

gabrewfinder.com

one of my students did this one.³

each dot is a beer
distance = similarity in taste



Wordle

<http://www.wordle.net>

Women's Rights

Women's Rights —macdoodle11 11 minutes ago

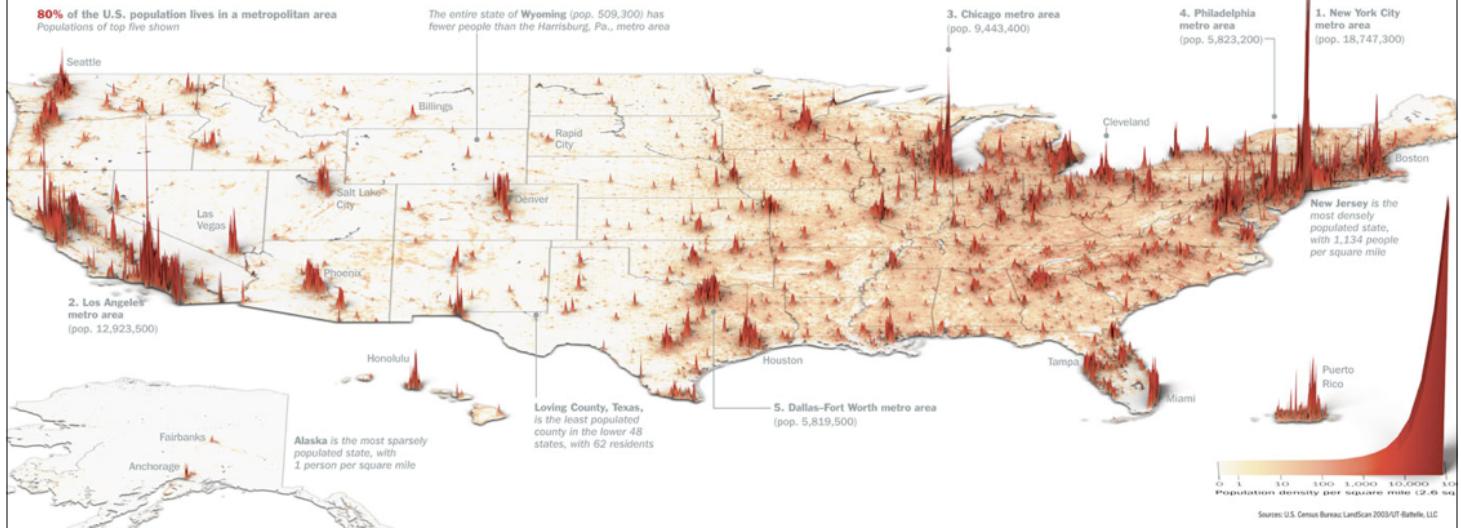


Population

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

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



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

Country Music

Geospatial

size is mapped to how often a state's name is mentioned in country song lyrics



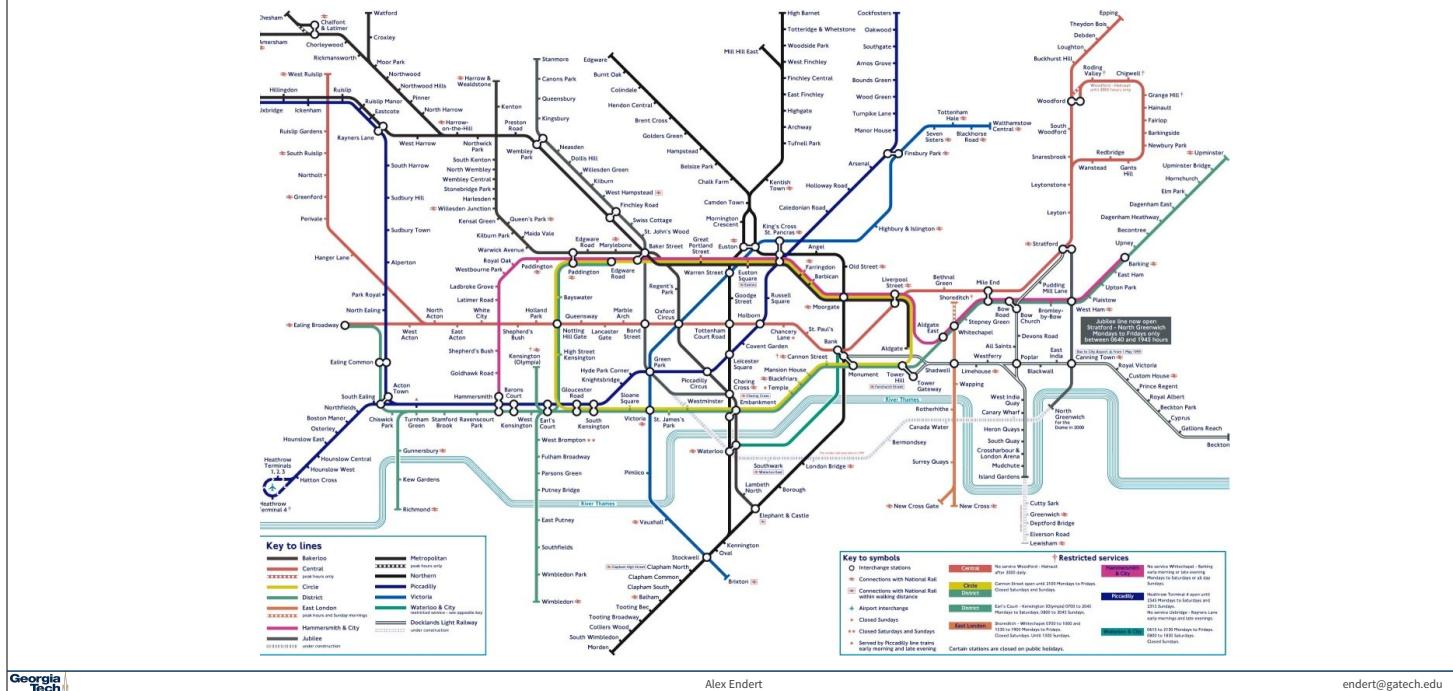
Figure 14. States Mentioned in Country-Music Lyrics

Source: Ben Marsh, "A Rose-Colored Map," Harper's, July 1977, 80. Used by permission.

Note: The size of each state is proportional to the number of times it is mentioned.

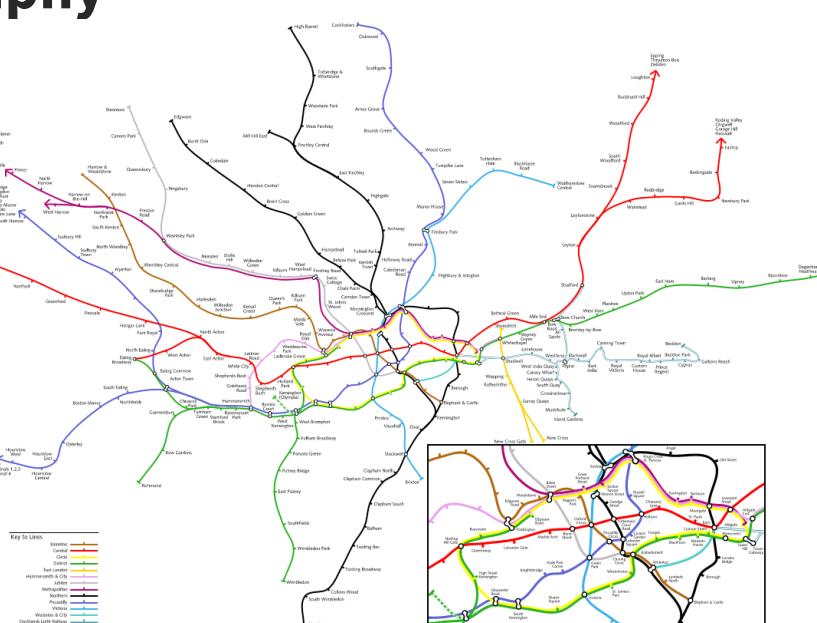
London Subway

www.thetube.com



True Geography

www.kottke.org/plus/misc/images/tubegeo.gif



Easy Walking Lines Added



Georgia Tech

endert@gatech.edu

Atlanta MARTA

They have rivers,
we have highways
:(

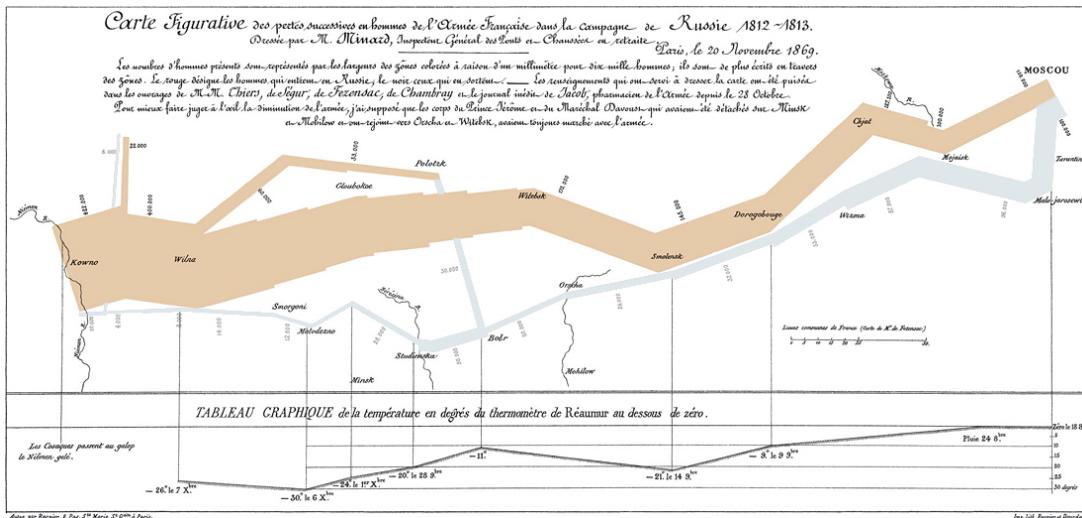


Georgia Tech

endert@gatech.edu

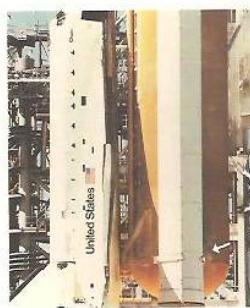
Napoleonic's March

From E. Tufte
The Visual Display of
Quantitative Information

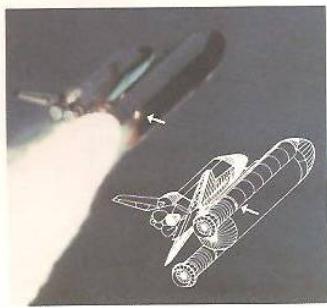


Example





Less than 1 second after ignition, a puff of smoke appeared at the aft joint of the right booster, indicating that the O-rings burned through and failed to seal. At this point, all was lost.



On the launch pad, the leak lasted only about 2 seconds and then apparently was plugged by insulation as the shuttle rose, flying through rather strong cross-winds. Then 58.768 seconds after ignition, when the Challenger was 6 miles up, a flicker of flame emerged from the leaky joint; 2 seconds later, the flame grew and engulfed the fuel tank (containing liquid hydrogen and liquid oxygen). That tank ruptured and exploded, destroying the shuttle.



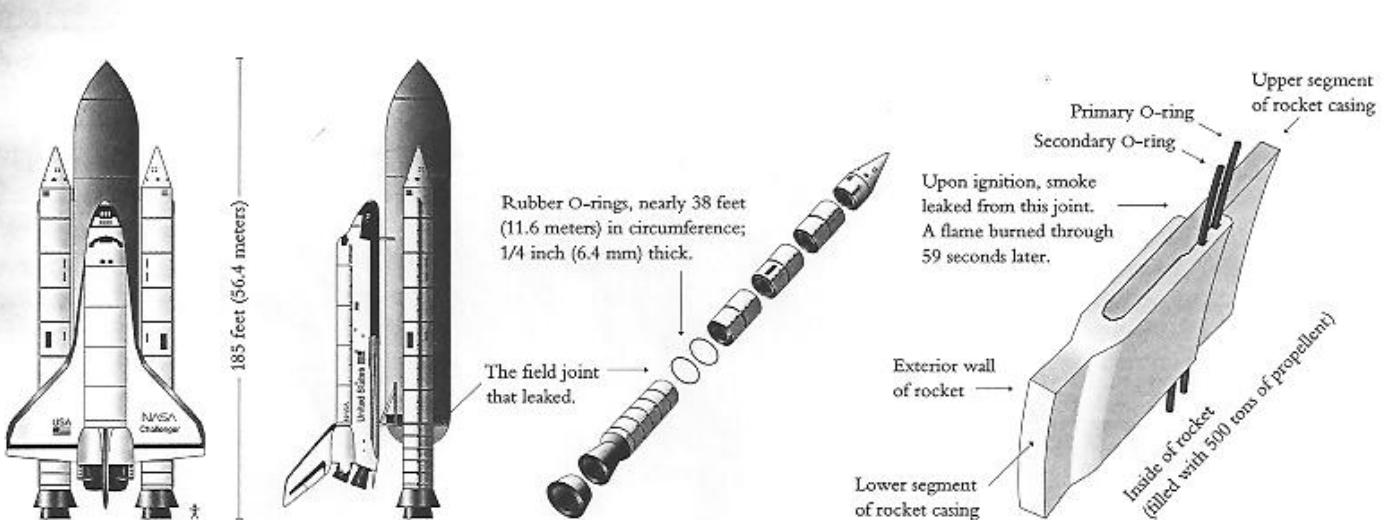
As the shuttle exploded and broke up at approximately 73 seconds after launch, the two booster rockets crisscrossed and continued flying wildly. The right booster, identifiable by its failure plume, is now to the left of its non-defective counterpart.



The flight crew of Challenger 51-L. Front row, left to right: Michael J. Smith, pilot; Francis R. (Dick) Scobee, commander; Ronald E. McNair, payload specialist. Back row: Ellison S. Onizuka, scientist-pilot; Sharon Christa McAuliffe, teacher; Gregory B. Jarvis, payload specialist; Judith A. Resnik.

endert@gatech.edu

Georgia Tech



The shuttle consists of an *orbiter* (which carries the crew and has powerful engines in the back), a large liquid-fuel *tank* for the orbiter engines, and 2 solid-fuel *booster rockets* mounted on the sides of the central tank. Segments of the booster rockets are shipped to the launch site, where

they are assembled to make the solid-fuel rockets. Where these segments mate, each joint is sealed by two rubber O-rings as shown above. In the case of the Challenger accident, one of these joints leaked, and a torch-like flame burned through the side of the booster rocket.

Georgia Tech

Alex Endert

endert@gatech.edu

TEMPERATURE CONCERN ON SRM JOINTS

27 JAN 1986

HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

SRM No.	Cross Sectional View			Top View	
	Erosion Depth (in.)	Perimeter Affected (deg)	Nominal Dia. (in.)	Length Of Max Erosion (in.)	Total Heat Affected Length (in.)
22A	None	None	0.280	None	None
22A	NONE	NONE	0.280	NONE	NONE
15A	0.010	154.0	0.280	4.25	5.25
15B	0.038	130.0	0.280	12.50	58.75
15B	None	45.0	0.280	None	29.50
13B	0.028	110.0	0.280	3.00	None
11A	None	None	0.280	None	None
10A	0.040	217.0	0.280	3.00	14.50
2B	0.053	116.0	0.280	--	--

*Hot gas path detected in putty. Indication of heat on O-ring, but no damage.

**Soot behind primary O-ring.

***Soot behind primary O-ring, heat affected secondary O-ring.

Clockwise location of leak check port - 0 deg.

OTHER SRM-15 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY AND NO SOOT NEAR OR BEYOND THE PRIMARY O-RING.

PRIMARY CONCERN -

ELD JOINT - HIGHEST CONCERN

EROSION PENETRATION OF PRIMARY SEAL REQUIRES RELIABLE SECONDARY SEAL
FOR PRESSURE INTEGRITY

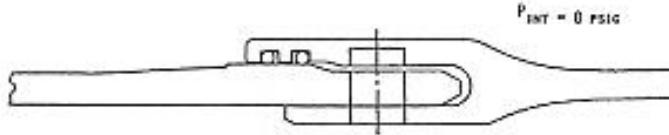
- o IGNITION TRANSIENT - (0-600 MS)
 - o (0-170 MS) HIGH PROBABILITY OF RELIABLE SECONDARY SEAL
 - o (170-330 MS) REDUCED PROBABILITY OF RELIABLE SECONDARY SEAL
 - o (330-600 MS) HIGH PROBABILITY OF NO SECONDARY SEAL CAPABILITY

STEADY STATE - (600 MS - 2 MINUTES)

- o IF EROSION PENETRATES PRIMARY O-RING SEAL - HIGH PROBABILITY OF NO SECONDARY SEAL CAPABILITY
- o BENCH TESTING SHOWED O-RING NOT CAPABLE OF MAINTAINING CONTACT WITH METAL PARTS GAP OPENING RATE TO MEOP
- o BENCH TESTING SHOWED CAPABILITY TO MAINTAIN O-RING CONTACT DURING INITIAL PHASE (0-170 MS) OF TRANSIENT

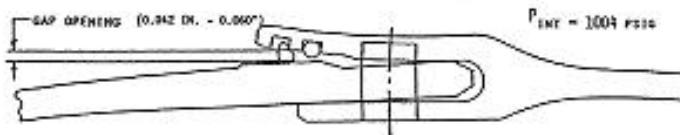
PRIMARY CONCERN - CONT

SEGMENT CENTERLINE



UNPRESSURIZED JOINT - NO ROTATION

SEGMENT CENTERLINE



PRESSURIZED JOINT - ROTATION EFFECT (EXAGGERATED)

BLOW BY HISTORY

SRM-15 WORST BLOW-BY

- o 2 CASE JOINTS (80° , (110°)) ARC
- o MUCH WORSE VISUALLY THAN SRM-22

SRM 22 BLOW-BY

- o 2 CASE JOINTS ($30-40^\circ$)

SRM-13A, 15, 16A, 18, 23A 24A

- o NOZZLE BLOW-BY

HISTORY OF O-RING TEMPERATURES (DEGREES - F)

MOTOR	MBT	AMB	O-RING	WIND
DM-4	68	36	47	10 MPH
DM-2	76	45	52	10 MPH
QM-3	72.5	40	48	10 MPH
QM-4	76	48	51	10 MPH
SRM-15	52	64	53	10 MPH
SRM-22	77	78	75	10 MPH
SRM-25	55	26	29	10 MPH
			27	25 MPH

Recommendation

CONCLUSIONS :

- o TEMPERATURE OF O-RING IS NOT ONLY PARAMETER CONTROLLING BLOW-BY
SRM 15 WITH BLOW-BY HAD AN O-RING TEMP AT 53°F
SEM 22 WITH BLOW-BY HAD AN O-RING TEMP AT 15°F
FOUR DEVELOPMENT MOTORS WITH NO BLOW-BY
WERE TESTED AT O-RING TEMP OF 47° TO 52°F

DEVELOPMENT MOTORS HAD PUTTY PACKING WHICH
RESULTED IN BETTER PERFORMANCE
- o AT ABOUT 50°F BLOW-BY COULD BE
EXPERIENCED IN CASE JOINTS
- o TEMP FOR SRM 25 ON 1-28-86 LAUNCH WILL
BE 29°F 9 AM
38°F 2 PM
- o HAVE NO DATA THAT WOULD INDICATE SRM 25 IS
DIFFERENT THAN SRM 15 OTHER THAN TEMP

RECOMMENDATIONS :

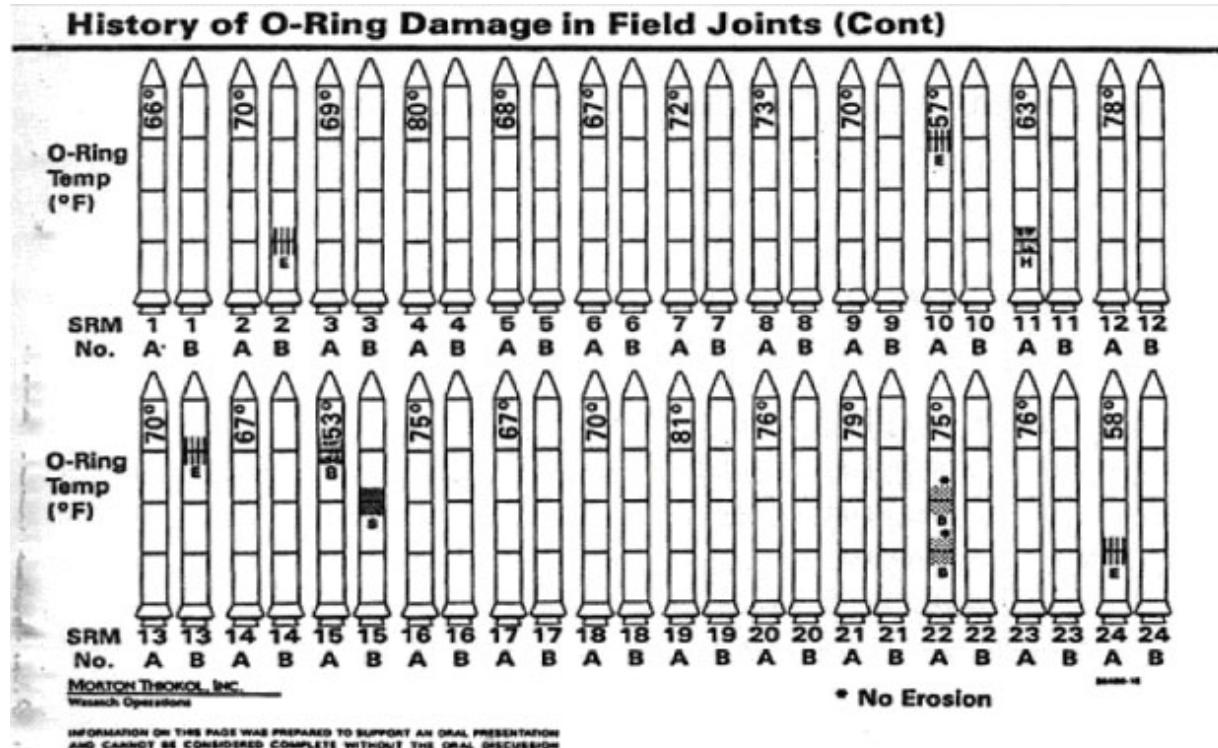
- o O-RING TEMP MUST BE \geq 53°F AT LAUNCH
DEVELOPMENT MOTORS AT 47° TO 52°F WITH
PUTTY PACKING HAD NO BLOW-BY
SRM 15 (THE BEST SIMULATION) WORKED AT 53°F
- o PROJECT AMBIENT CONDITIONS (TEMP & WIND)
TO DETERMINE LAUNCH TIME

Predict Temperature

The data

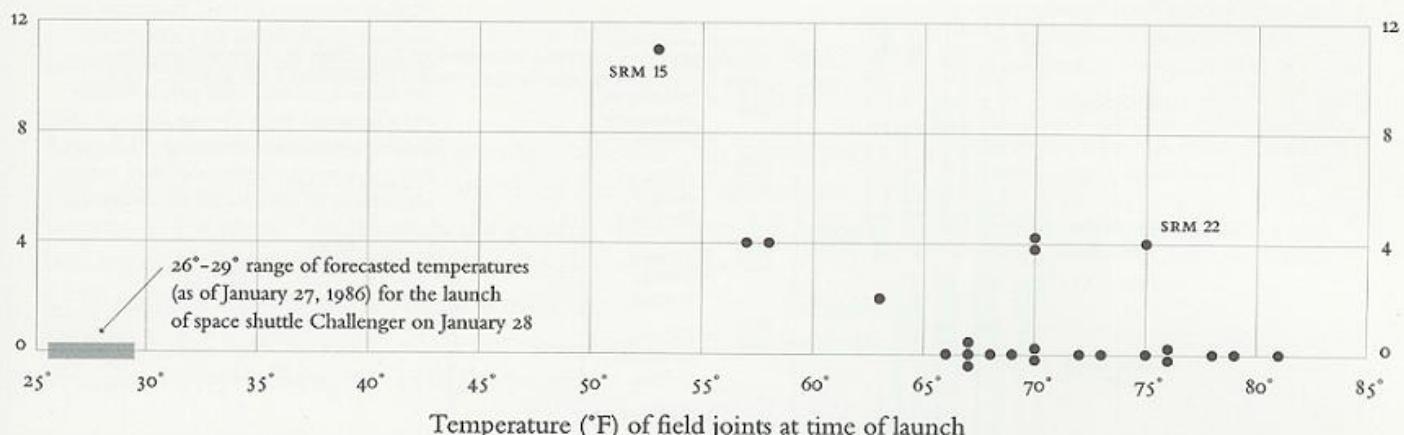
Flight	Date	Temperature °F	Erosion incidents	Blow-by incidents	Damage index	Comments
51-C	01.24.85	53°	3	2	11	Most erosion any flight; blow-by; back-up rings heated.
41-B	02.03.84	57°	1		4	Deep, extensive erosion.
61-C	01.12.86	58°	1		4	O-ring erosion on launch two weeks before Challenger.
41-C	04.06.84	63°	1		2	O-rings showed signs of heating, but no damage.
1	04.12.81	66°			0	Coolest (66°) launch without O-ring problems.
6	04.04.83	67°			0	
51-A	11.08.84	67°			0	
51-D	04.12.85	67°			0	
5	11.11.82	68°			0	
3	03.22.82	69°			0	
2	11.12.81	70°	1		4	Extent of erosion not fully known.
9	11.28.83	70°			0	
41-D	08.30.84	70°	1		4	
51-G	06.17.85	70°			0	
7	06.18.83	72°			0	
8	08.30.83	73°			0	
51-B	04.29.85	75°			0	
61-A	10.30.85	75°		2	4	No erosion. Soot found behind two primary O-rings.
51-I	08.27.85	76°			0	
61-B	11.26.85	76°			0	
41-G	10.05.84	78°			0	
51-J	10.03.85	79°			0	
4	06.27.82	80°			?	O-ring condition unknown; rocket casing lost at sea.
51-F	07.29.85	81°			0	

The Vis: Can you see a pattern?

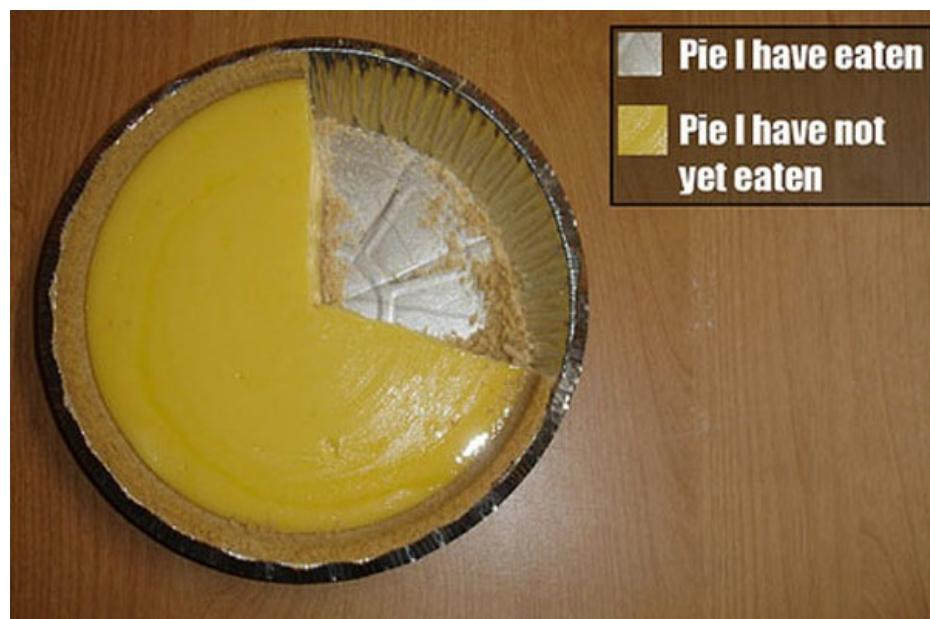


Another look at the data

O-ring damage index, each launch



Or, for fun...

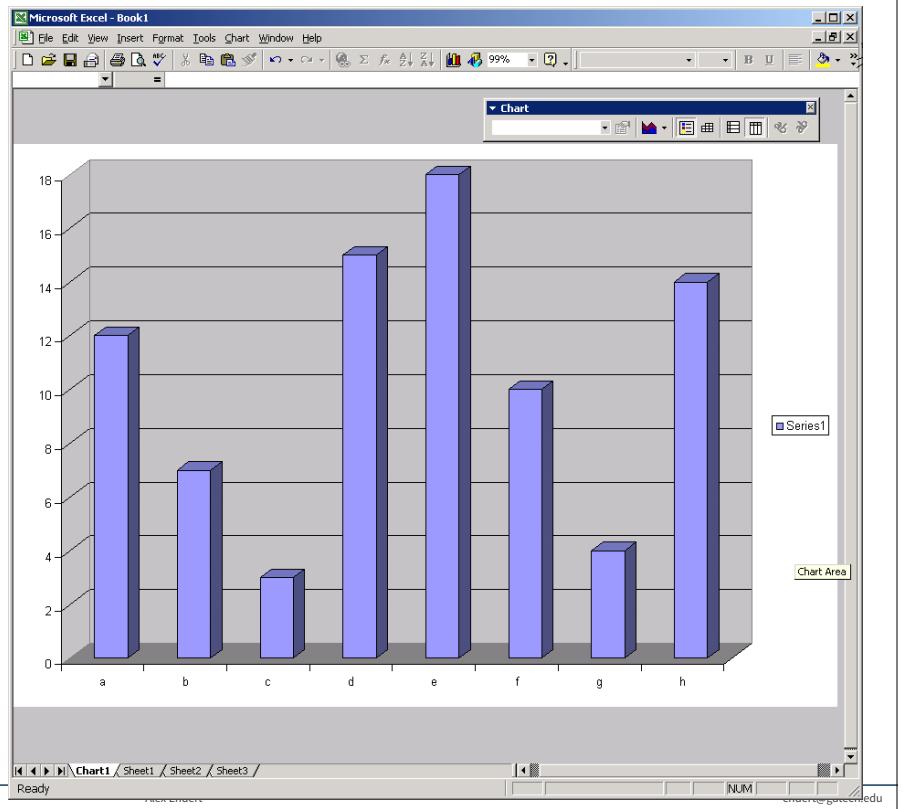


http://infosthetics.com/archives/2008/09/funniest_pie_chart_ever.html

But Don't Do This

Excel

What information does the third dimension on the bars show?



USA Today Graphics



Can anyone figure out what information is being shown here?

Alex Endert

endert@gatech.edu

USA Today Graphics

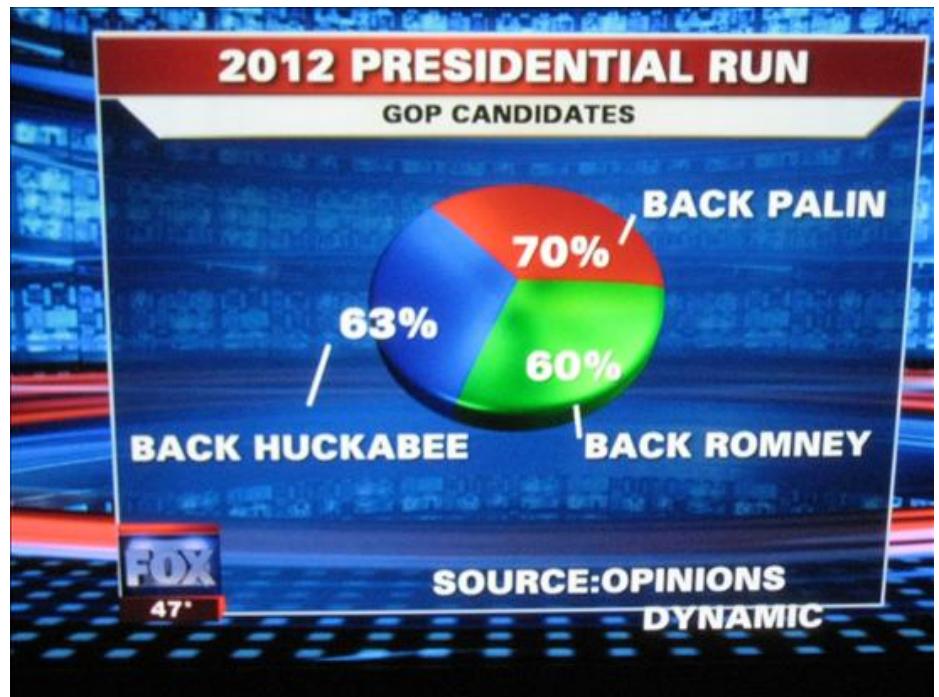


Can anyone figure out what information is being shown here?

Alex Endert

endert@gatech.edu

see anything wrong here?



<http://wonkette.com/412361/all-193-of-republicans-support-palin-romney-and-huckabee>

Recap from today

- You should have learned about:
 - Why do we visualize data? (and when should we not)
 - What are the 2 high-level purposes of InfoVis
 - Get an overview of different types of visualizations (more detail on domain-specific one's later)
 - Learned about the specific example of why data presentation is important (Challenger use case)

Reminders

- The course website has all the information about assignments, schedule, etc.
- When using GitHub, use **private repos**
- HW1 due soon (no submission on Canvas, just do the survey)