

# Introduction to Information Visualization



CS 4460 - Information Visualization

Fall 2016

Prof Jim Foley

GtAs Sreshta Vijayaraghavan, Lei Xu,  
Qixuan LeiHou

# Agenda for Today



- Introductions of your teacher & GTAs
- An example InfoViz
- Course information
- More about InfoViz
- Project information

# Introductions - Jim Foley



- Founded GVI Center @ GT in 1991
- Industry and consulting and education
- Recently directed MS-HCI Degree
- Research Interests
  - HCI
    - Technology in education
    - Information Visualization
  - Computer Graphics (four books, no more!)
- Office hours
  - I'm available a lot – before and sometimes after class
  - No one ever comes to my office in TSRB 355
  - Arrange other times via email; [foley@cc.gatech.edu](mailto:foley@cc.gatech.edu)
- Something about me .....









# Sreshta Vijayaraghavan



From Singapore

Pursuing MSCS in HCI

BS in CS from NUS '13

Worked at Goldman Sachs  
and Interned at Pinterest

My last semester in school ☹️

Love cooking, running, and  
reading books

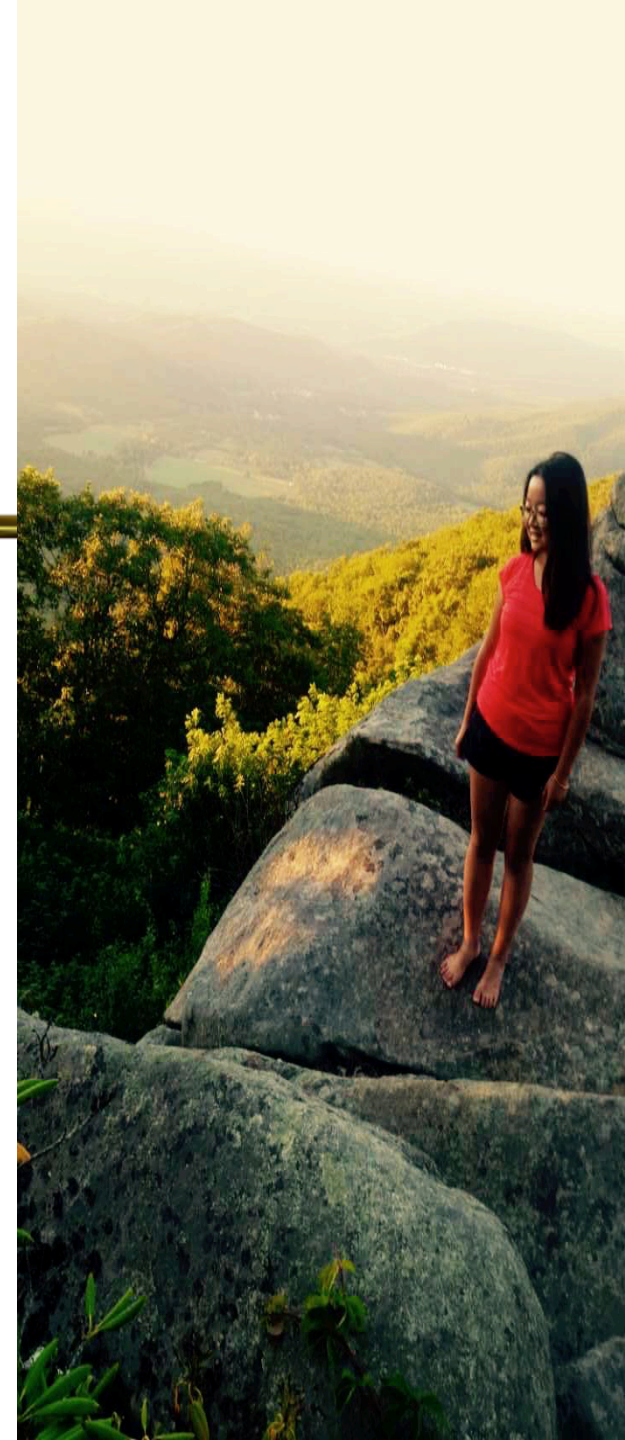
Favorite country - Japan

# Qixuan Hou

## Queency

---

- From Gansu, China
- Pursuing BS in CS & Math
- Interned at Cox Automotive
- Interested in hiking, kayaking, caving...





# Lei Xu

From the other side of the Facebook “like” sign

Pronounces as: Lay Shoo

From China

MS-HCI: Interactive Computing

B.S. in CS, Tsinghua Univ.

Interned in Google (UX) and  
Facebook (PE)

Struggled with D3 force  
layout in CS 7450 last year

Finished 3 full Marathons





# What is Information Visualization?

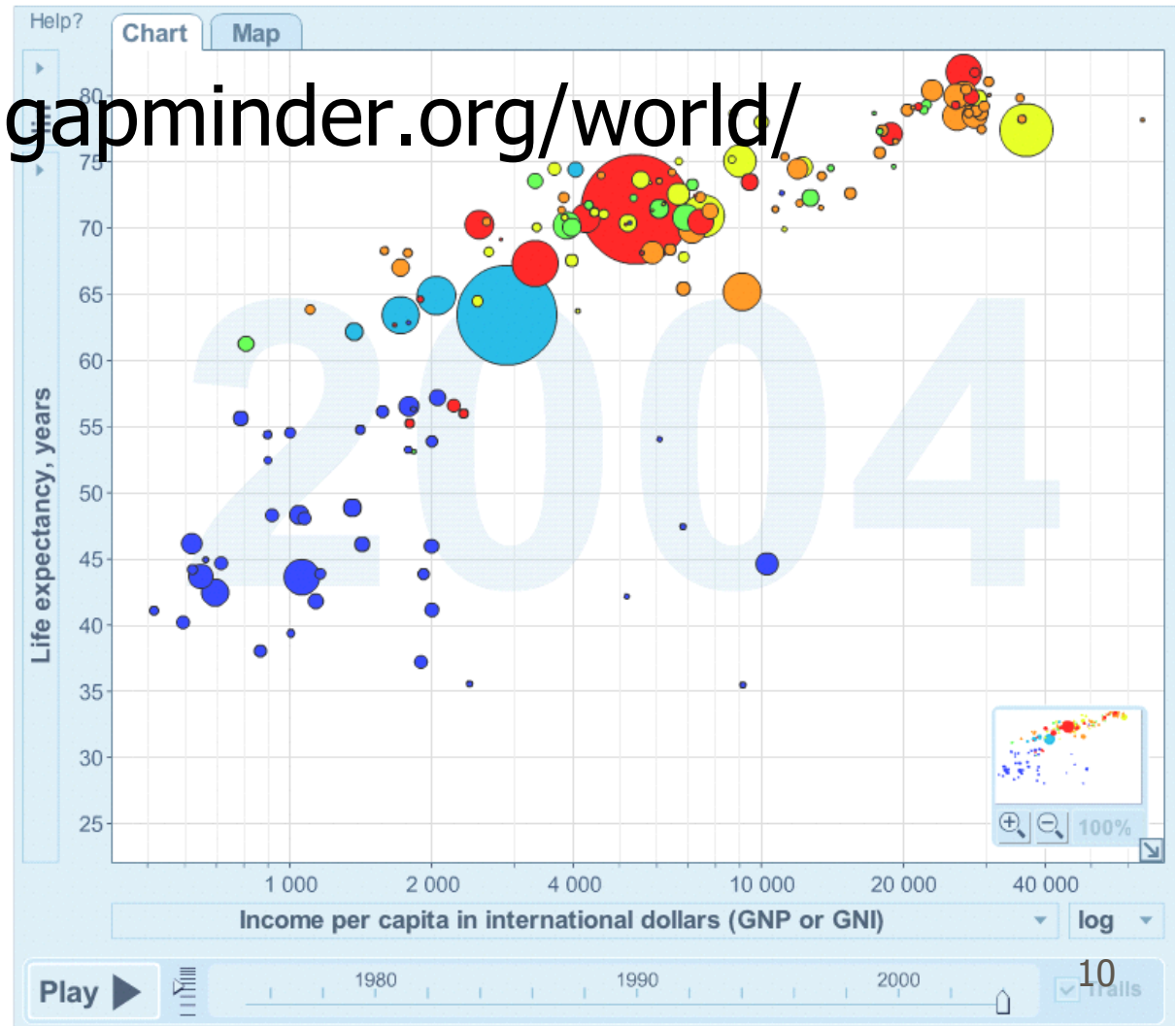


- My elevator speech
  - Presenting data via INTERACTIVE charts, graphs, maps so that users can understand the data, answer questions about the data and gain insights from the data
  - NOT just printed (static) information presentation – interaction is key element!
- Longer version
  - The whole course

# Information Visualization Example



- Gapminder
- <http://www.gapminder.org/world/>
- Interaction!
- Animation!





# Interaction Methods in Gapminder



- Change Representation (Map, chart)
- Linking/brushing/wiping
- Details on Demand (DoD)
- Animation
- Change binding of data to a visual representation

# Course Goals: Be Able to



- Know how to design and implement information visualizations.
- Use design principles to create effective information visualizations.
- Choose appropriate type of visualization for various types of data and for different goals.
- Choose appropriate dynamic interaction methods.
- Apply understanding of human perceptual and cognitive capabilities to the design of information visualizations.
- Critiquing different visualization techniques in the context of user goals and objectives.
- Implement compelling Information Visualizations



# Policies - <https://cs4460infovis.wordpress.com/>



- No late homework accepted without documented personal issues (serious illness, family emergency, etc.)
- Review the Georgia Tech Academic Honor Code
  - <http://www.deanofstudents.gatech.edu/Honor/>
  - Individual HWs done individually
  - Group projects are the work of your group alone
    - Talk to others for feedback; look at other systems for ideas; group synthesizes an **original** design
- Attendance expected

# Device Usage in Class



- During lectures please put your smart phones away, and use your computers and tablets for course-relevant work. Games, browsing & emailing is distracting to those around you: messengers and notifications are designed to grab your attention and are de-facto irresistible. BTW, even note-taking by hand versus on your computer has been shown to be more efficient for learning (also see this news story). But that's your call to make.



# Texts and Readings



- Listed on web site
- Extensive reading list on CS 7460 web site
  - Most can be accessed on-line via GT library
- You may want to buy one of:
  - For those interested in design: Any of Edward Tufte's three books: *The Visual Display of Quantitative Information*; *Envisioning Information*; and *Visual Explanations*.
  - For those interested in business intelligence and business dashboards: Wayne Eckerson, *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*, Wiley, 2005, ISBN 978-0471724179
  - For those interested in Network Visualization, particularly Social Networks: Hansen, Shneiderman and Smith, *Analyzing Social Media Networks with NodeXL*, Morgan Kaufman, 2011, ISBN 978-0-12-382229-1.
  - For those interested in the psychological/perceptual factors affecting information visualization: Colin Ware, *Information Visualization: Perception for Design*, 2nd Edition, Morgan Kaufman Elsevier 2004, ISBN 978-1558608191.
  - For a deeper treatment of many aspects of InfoViz: *Visualization Analysis and Design*, Tamara Munzer, CRC Press, ISBN 9781466508910.

# Grading



- Point distribution
  - Test 1                    18 points
  - Test 2                    18 points
  - Pop quizzes (4)            4 points
  - Homeworks                25 points
  - Project                    35 points
- Cutoffs generally around 80 and 90, I do give Cs and Ds and the occasional F
- YOU MUST be a responsible project team member; slough off at your own peril!
- Participation and attendance matter if you are close to a boundary

# Web Sites



- Class schedule, assignments, PPTs:
  - <http://cs4460infovis.wordpress.com/>
  - Note HW1 and HW2 due Thursday!
- Turn-ins via T-Square
  - [CS-4460-A](#)
- Piazza – you will receive invite to join
  - Use it 😊!

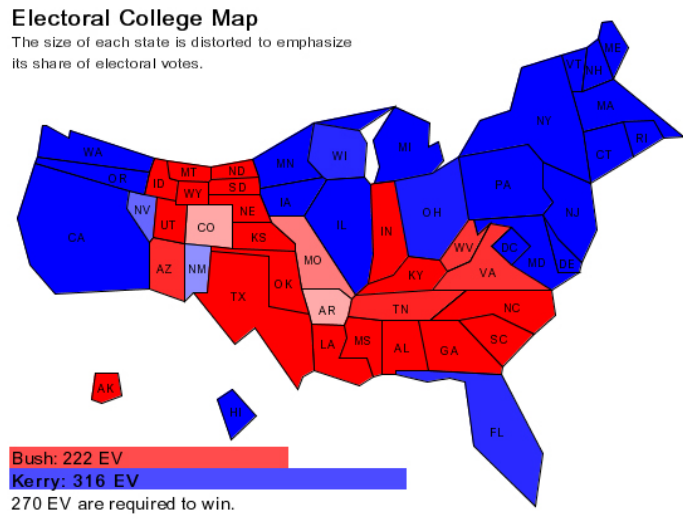
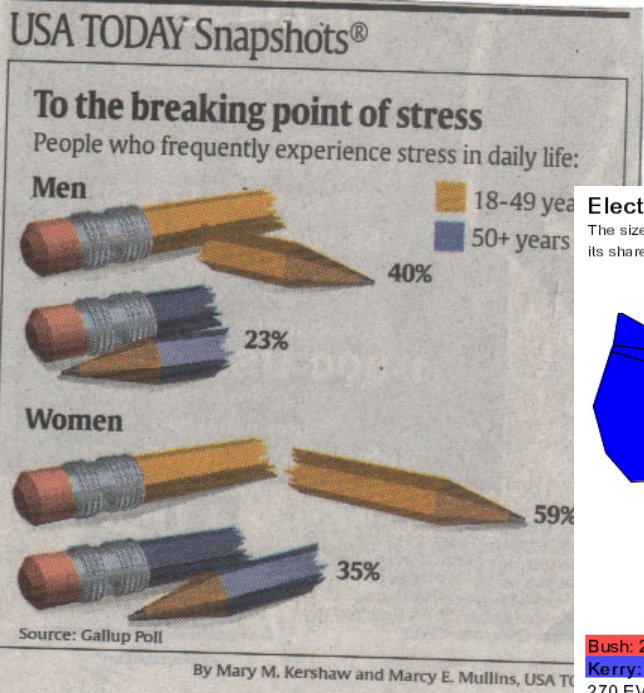
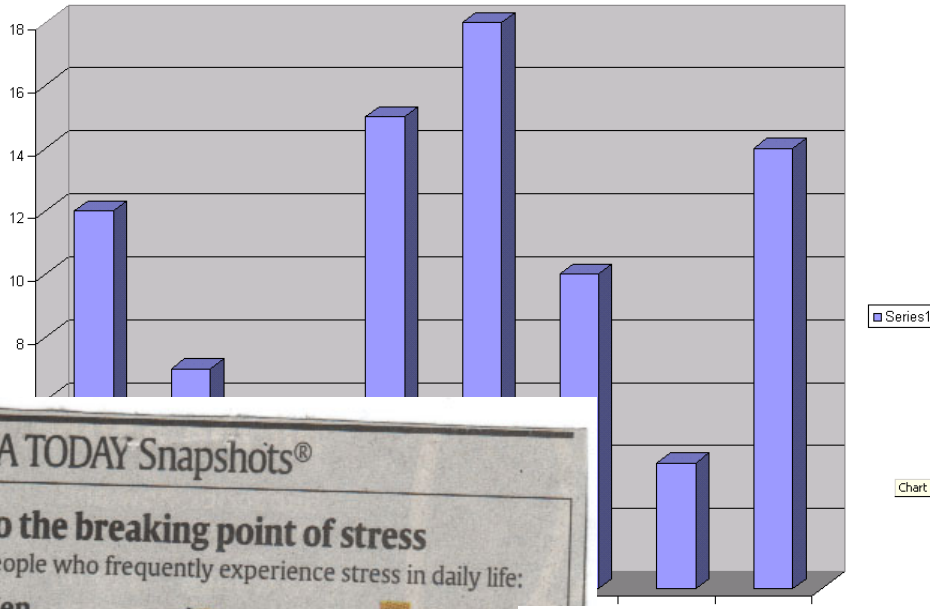
# Active Learning



- Talk with your neighbor about course information.
- I'll ask a few groups if have questions
- BTW a common pedagogical technique throughout the course



# Information Presentation Examples



## WIDE WORLD OF WAYANS

Whatever happened to...the cut-ups of *In Living Color*?

**Keenen Ivory Wayans** | **Jim Carrey** | **Rosie Perez** | **Marlon Wayans** | **Jennifer Lopez**

**1991** Stars in *Night on Earth* after a brilliant turn in *Do the Right Thing* as "Toreats" that get rubbed with ice.

**1993** Gets Oscar nomination for playing a plane-crash survivor in *Fearless*. Unfortunately, she doesn't grieve by rubbing ice on her breasts.

**1994** Makes a butt-load for *Ace Ventura: Pet Detective* by talking out of his ass. Hey, we can do that! (read in ass accent) Where's our moohah?

**1996** Ditches depressingly unfamous wife Melissa Womer like flaming pair of underwear; later acquires suitable trophy girlfriend Lauren Holly.

**1996** Earns \$20 million for *The Cable Guy*, which, after a huge opening weekend, earns 47¢ at the box office.

**1997** Stars in *Selena* as the Latin songstress' stalker. Her voice is mercifully dubbed for the musical numbers.

**1997** Stars in and produces HBO's *Subway Stories: Tales From the Underground*. Hey, who's the "man who reeks of urine and screams to himself" episode?

**1997** Stars in *Men in Black II* as the wildly successful *Scary Movie*, which lampoons homosexuals, ugly people, and the disabled. "Mr. Wayans, I've got Fox on line two..."

**1999** The cover of her debut album, *On the 6*, sells millions of copies...and each one comes with a free audio CD!

**1999** Wears a sheet of toilet paper to the Grammys. Ratings skyrocket as 500 million men are unable to stand for hours.

**2000** The *Grinch*. Carrey subjects himself to hours of makeup each day to make his face rubbery and funny.

**2001** It's reported that Lopez will make personal appearances—for \$750,000 an hour. In unrelated news, a \$3 million office party plunges *Maxim* into debt.

**1995** "Somehow" gets a gig on the WB series *The Wayans Brothers* as the fifth choice for the part of "Wayans Brother."

**1996** Gets role in brother Keenen's *Damn Be a Menace to South Central While Drinking Your Juice in the Hood* by spelling his name correctly on the application.

**1998** Senseless: A guy takes a portion that heightens some senses but makes him lose others. Many scripts are rejected each year so classics like *Senseless* can be made.

**2000** Arrested for protesting the U.S. Navy's bombing range on the Puerto Rican island of Vieques. Perez is immediately sentenced to rub ice on her breasts.

**2000** Appears in those wonderful, string-C-A-L-L-A-T-I commercials proving there truly can be something more annoying than David Arquette.

**1997** Hosts the short-lived late-night *Keenen Ivory Wayans Show*. This horrible unfunny streak prompts rumors Keenen has contracted a crippling case of Dan Aykroyd.

**1996** Joins perennial Oscar favorite Steven Seagal in *The Glimmer Man*. Superviolent...and yet, somehow, it sucks!

**1997** Writes, *Low Down* intended the *Shaft* tended joke: receipts.



# What is Information Visualization?



- Others' definitions
  - The use of computer-supported, *interactive* visual representations of data to amplify cognition (Card, Mackinlay Shneiderman '98)
- These two definitions purport to define InfoVis, but what are they missing?
  - Transformation of the symbolic into the geometric (McCormick et al., 1987)
  - The depiction of information using spatial or graphical representations, to facilitate comparison, pattern recognition, change detection, and other cognitive skills by making use of the visual system (Hearst 03).

# BTW



- Information Visualization aka
  - InfoVis
  - InfoViz
- Other related terms
  - Infographics
  - Dashboards
  - Data Visualization



# Some InfoVis Application Areas



- Financial/business data – “business intelligence”
- Internet information – traffic, topology
- Software development
- Intelligence analysis
- Social networking analysis
  
- Others you have seen??

# Information Visualization



- What is “information”?
  - Items, entities, things which do not have a direct physical correspondence
  - Notion of abstractness of the entities is important too
  - Examples: baseball statistics, stock trends, connections between criminals, car attributes...

# The Challenge of InfoVis



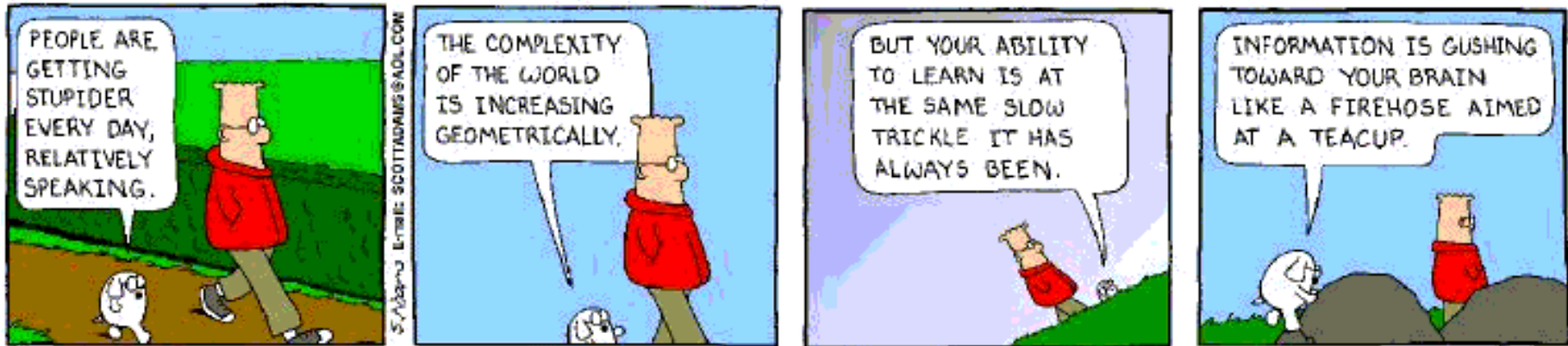
- Transform *data* into *information* (understanding, insight) thus making it useful to people
  - To understand the data
  - To make decisions based on that understanding
- Visual Analytics
  - New term - focus on decisions/action



# Data Overload



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





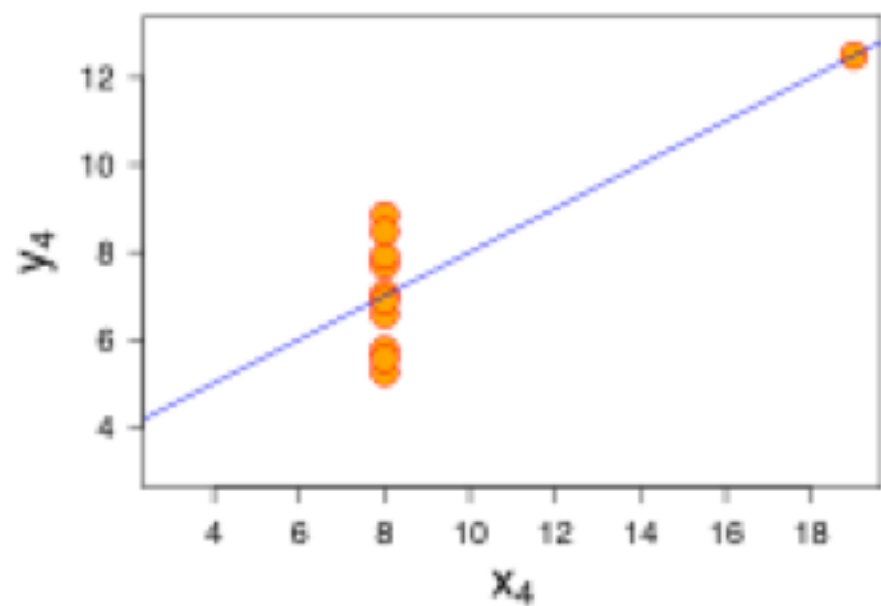
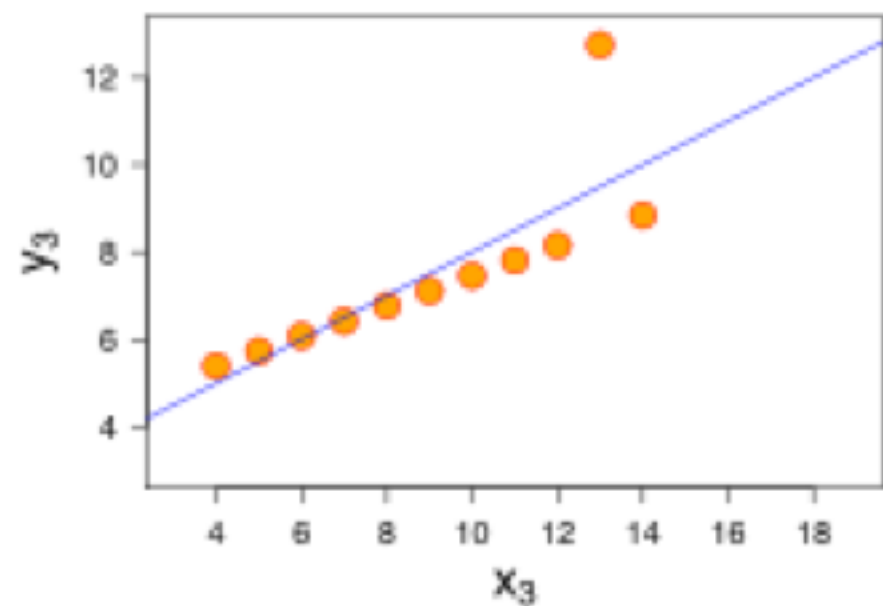
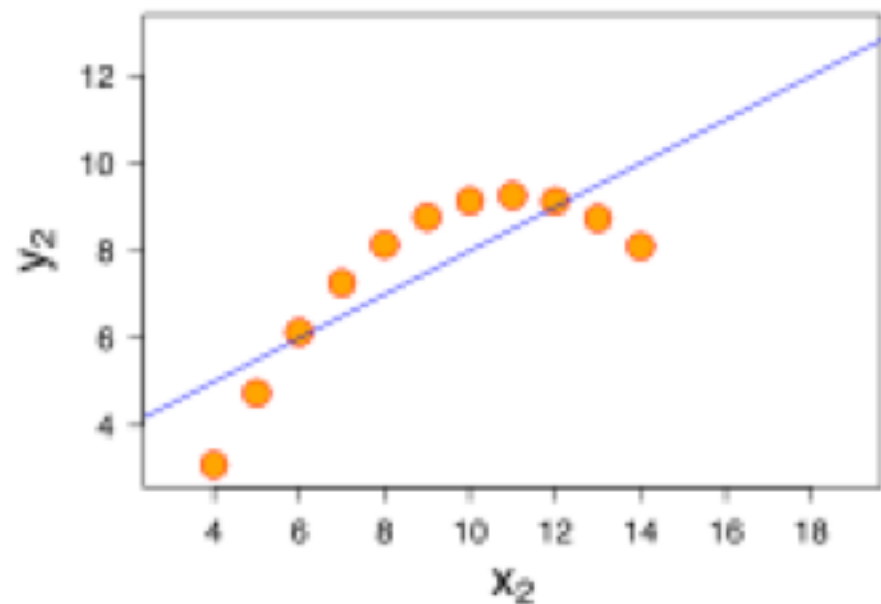
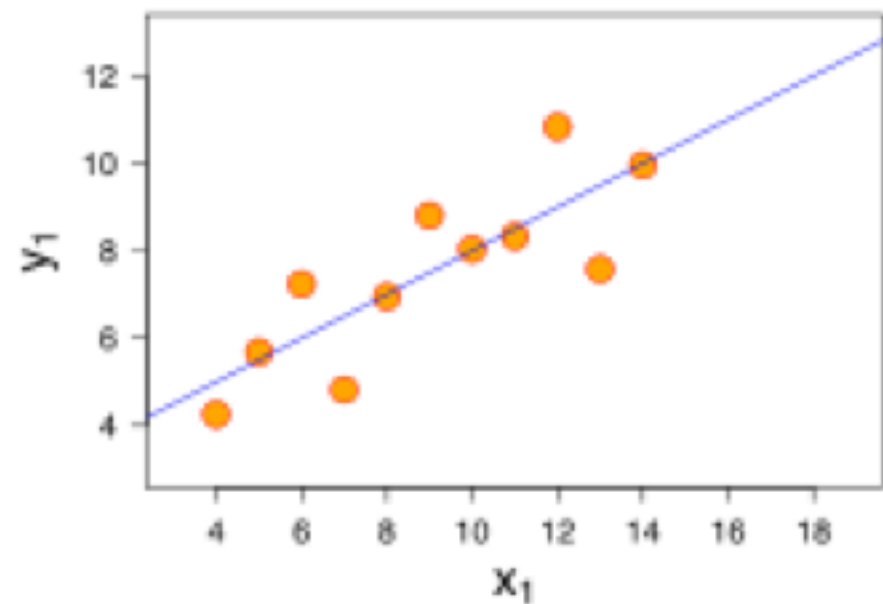
# Anscombe's Quartet – Tables or Graph?



- Statistics the same for each of the four  $x$ - $y$  tables (the quartet)
  - Mean = 9
  - Variance = 9
  - Correlation = 0.816
  - Linear regression
  - $y = 3 + 0.5x$
- So what's different about the data?
- You could study the tables very closely
  - And make a little progress
- Or you could graph the data 😊

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

From F.J. Anscombe, "Graphs in Statistical Analysis", *American Statistician*, February 1973, 17-21.



# Data – Tables or Graphs?



Questions:

Which state has the highest income?

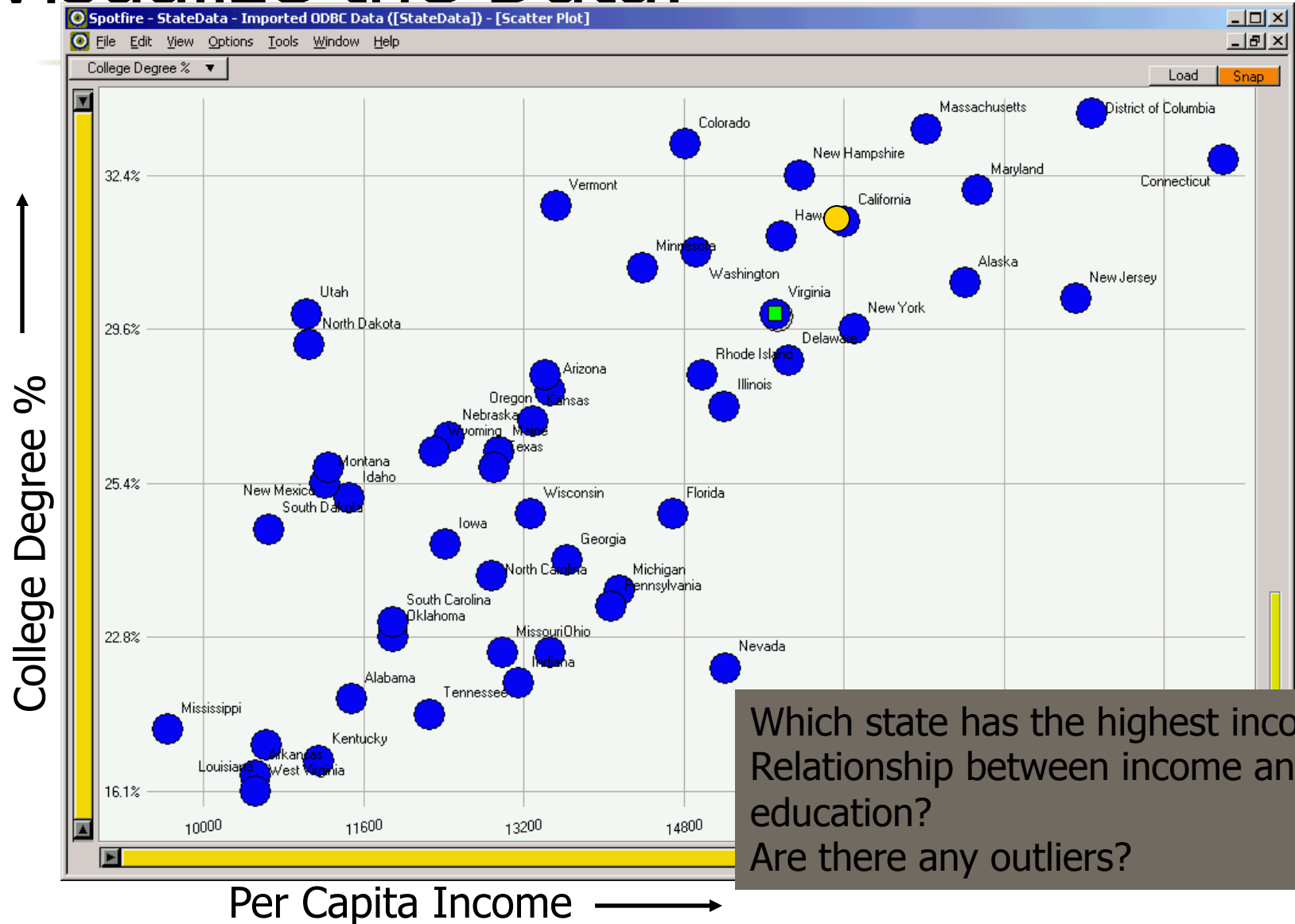
Is there a relationship between income and education?

Are there any outliers?

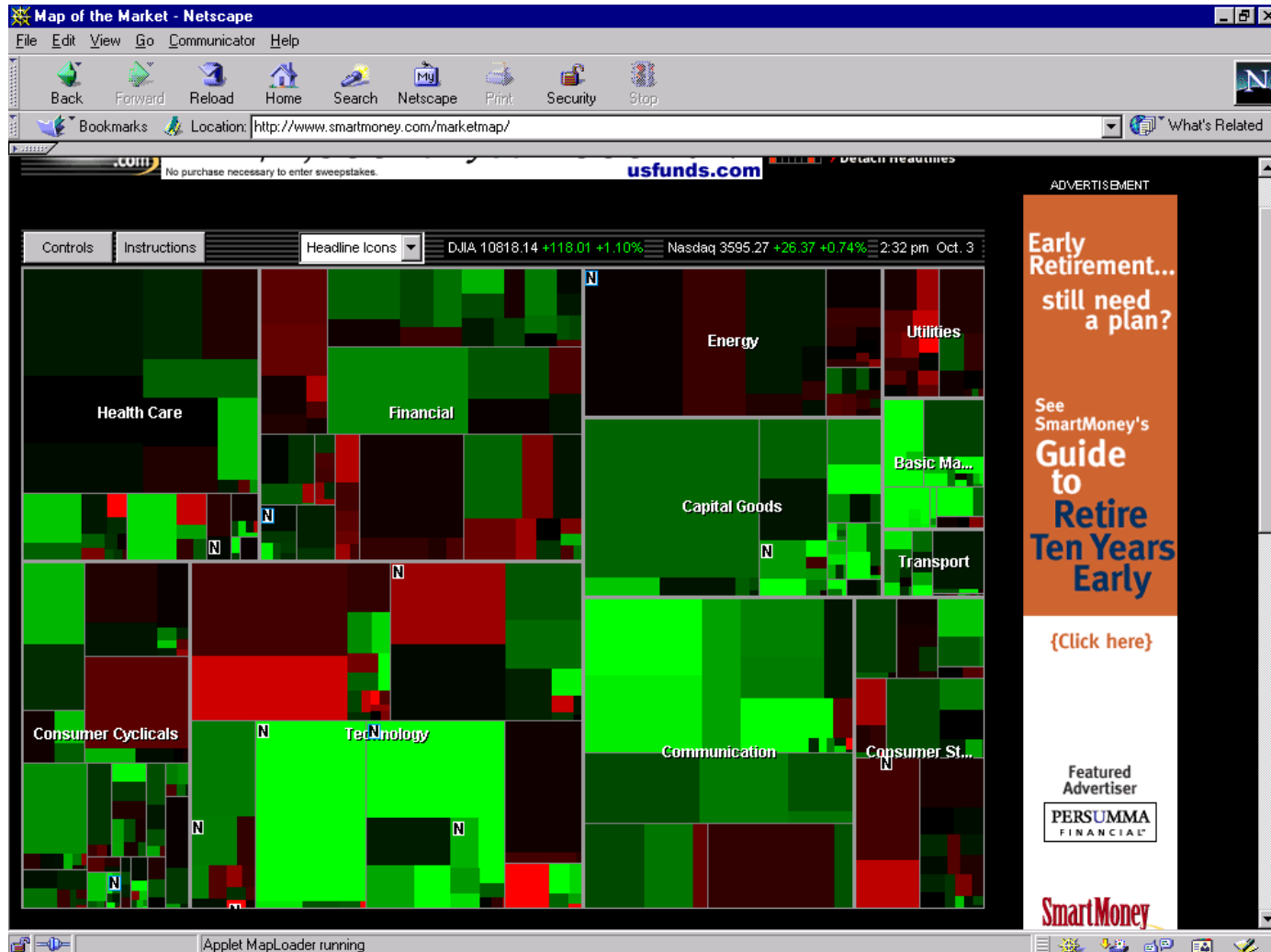
State	College Degree %	Per Capita Income
Alabama	20.6%	11486
Alaska	30.3%	17610
Arizona	27.1%	13461
Arkansas	17.0%	10520
California	31.3%	16409
Colorado	33.9%	14821
Connecticut	33.8%	20189
Delaware	27.9%	15854
District of Columbia	36.4%	18881
Florida	24.9%	14698
Georgia	24.3%	13631
Hawaii	31.2%	15770
Idaho	25.2%	11457
Illinois	26.8%	15201
Indiana	20.9%	13149
Iowa	24.5%	12422
Kansas	26.5%	13300
Kentucky	17.7%	11153
Louisiana	19.4%	10635
Maine	25.7%	12957
Maryland	31.7%	17730
Massachusetts	34.5%	17224
Michigan	24.1%	14154
Minnesota	30.4%	14389

State	College Degree %	Per Capita Income
Minnesota	30.4%	14389
Mississippi	19.9%	9648
Missouri	22.3%	12989
Montana	25.4%	11213
Nebraska	26.0%	12452
Nevada	21.5%	15214
New Hampshire	32.4%	15959
New Jersey	30.1%	18714
New Mexico	25.5%	11246
New York	29.6%	16501
North Carolina	24.2%	12885
North Dakota	28.1%	11051
Ohio	22.3%	13461
Oklahoma	22.8%	11893
Oregon	27.5%	13418
Pennsylvania	23.2%	14068
Rhode Island	27.5%	14981
South Carolina	23.0%	11897
South Dakota	24.6%	10661
Tennessee	20.1%	12255
Texas	25.5%	12904
Utah	30.0%	11029
Vermont	31.5%	13527
Virginia	30.0%	15713
Washington	30.9%	14923
West Virginia	16.1%	10520
Wisconsin	24.9%	13276
Wyoming	25.7%	12311

# Visualize the Data!



# Map of the Market



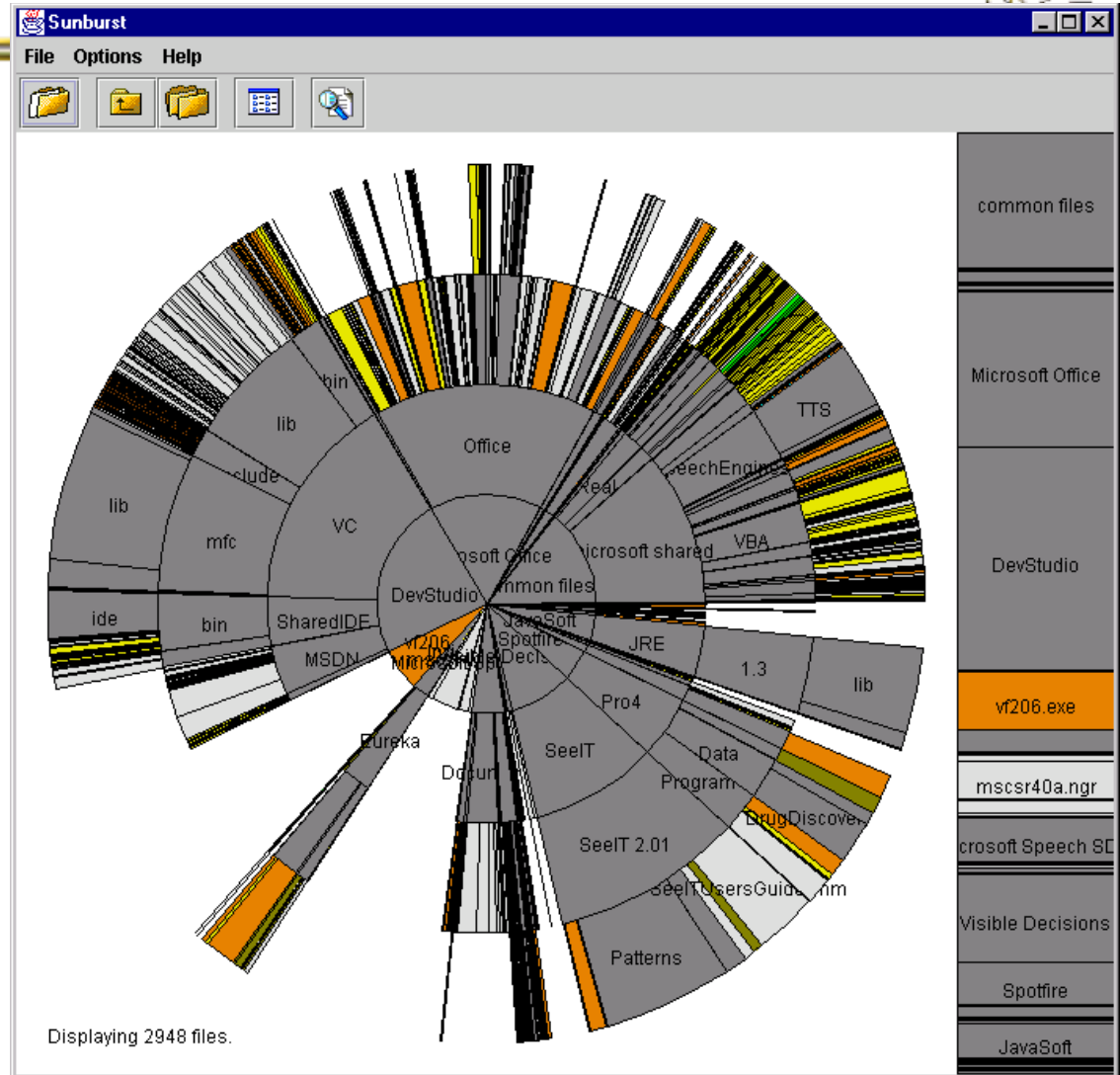
Demo



# SunBurst

[www.cc.gatech.edu/gvu/ii/sunburst](http://www.cc.gatech.edu/gvu/ii/sunburst)

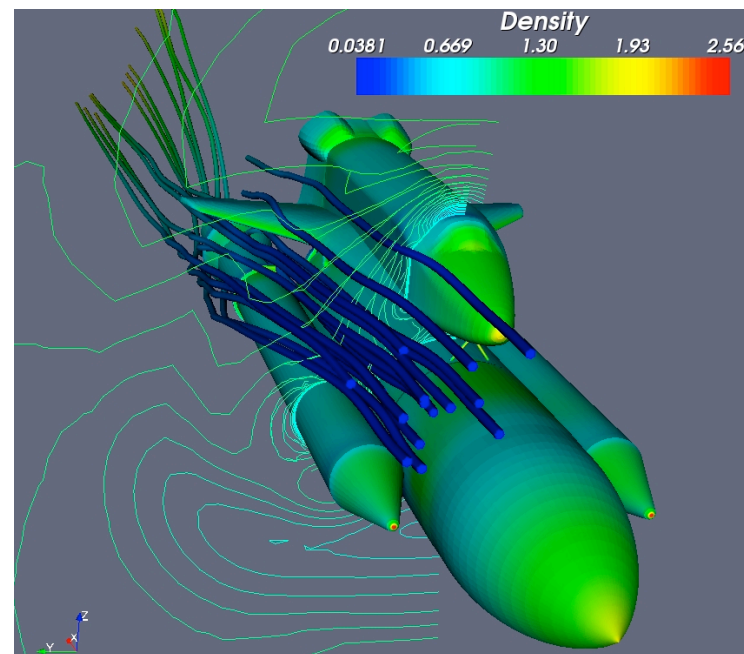
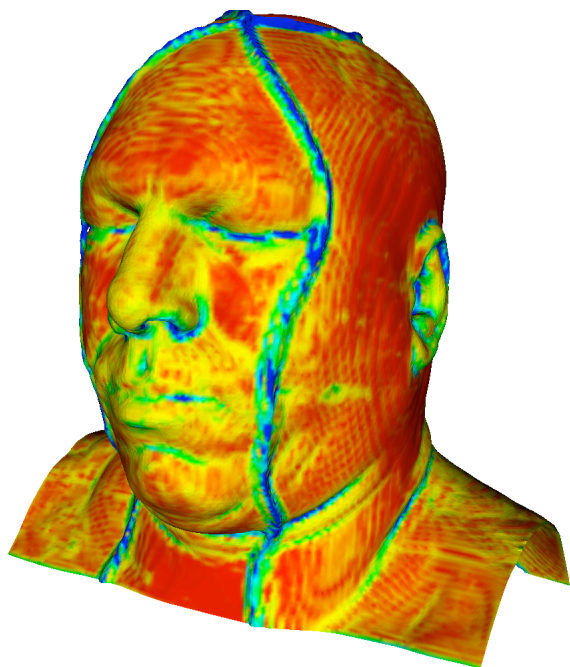
File browser



# InfoViz $\rightarrow$ = SciViz or MedViz



- SciViz: Scientific Data Visualization
- MedViz: Medical Data Visualization (body)



# SciViz & MedViz



- Data generally associated with physical positions in a 2D or 3D space – has a geometry
- InfoViz – data generally abstract
  - We have to create a geometry with which the data is encoded

# Information Visualization

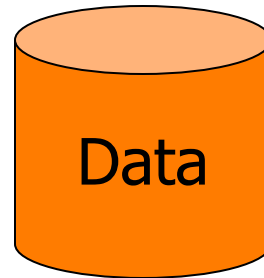


- Consists of:
  - Taking elements of information without a direct physical correspondence and mapping them to a 2-D or 3-D physical space.
  - Giving the information elements a *visual representation* that is useful for analysis and decision-making
  - Adding *interactions* to facilitate exploration

# The Problem



Web,  
Books,  
Papers,  
Game scores,  
Scientific data,  
Biotech,  
Shopping  
People  
Stock/finance  
News



Data Transfer →

## How?



Vision: 100 MB/s

Ears: <100 b/s

Telepathy

Haptic/tactile

Smell

Taste

Two slides courtesy  
of Chris North



# More about the Course



- Wordpress web site – syllabus, schedule, homework, project info
- T-Square for turn-ins, grades, Wiki
- Piazza
- Project
- Due on Thursday
  - HW1 – survey
  - HW2 – bring to class an Info Presentation showing data of interest to you