The Purpose of Visualization

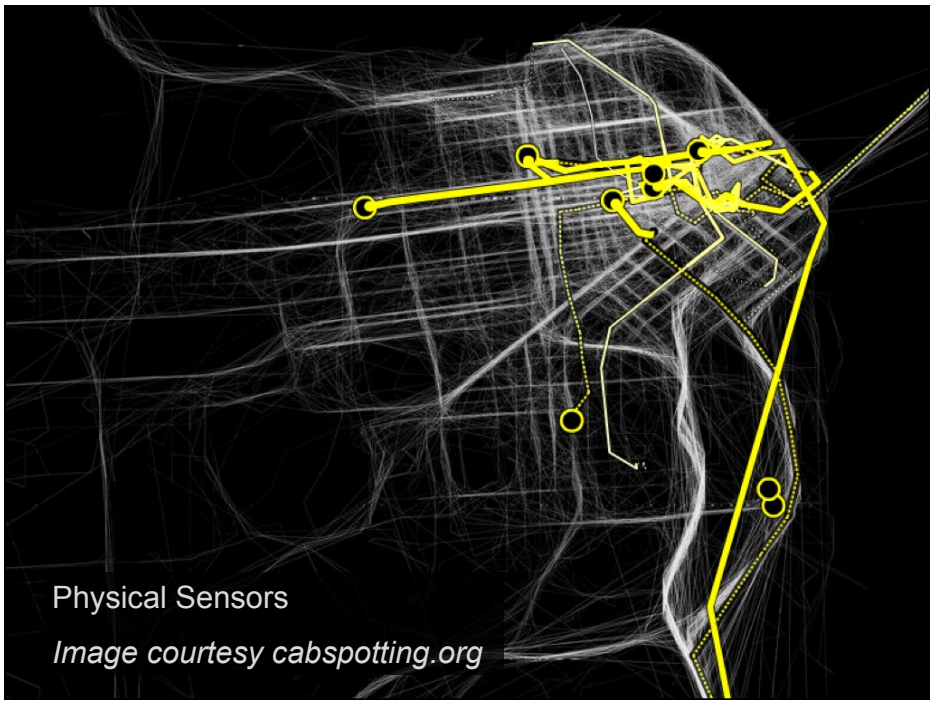
Maneesh Agrawala

CS 294-10: Visualization
Fall 2013

How much data (bytes) did we produce in 2011?

2011: 1800 exabytes
10x increase over 5 years

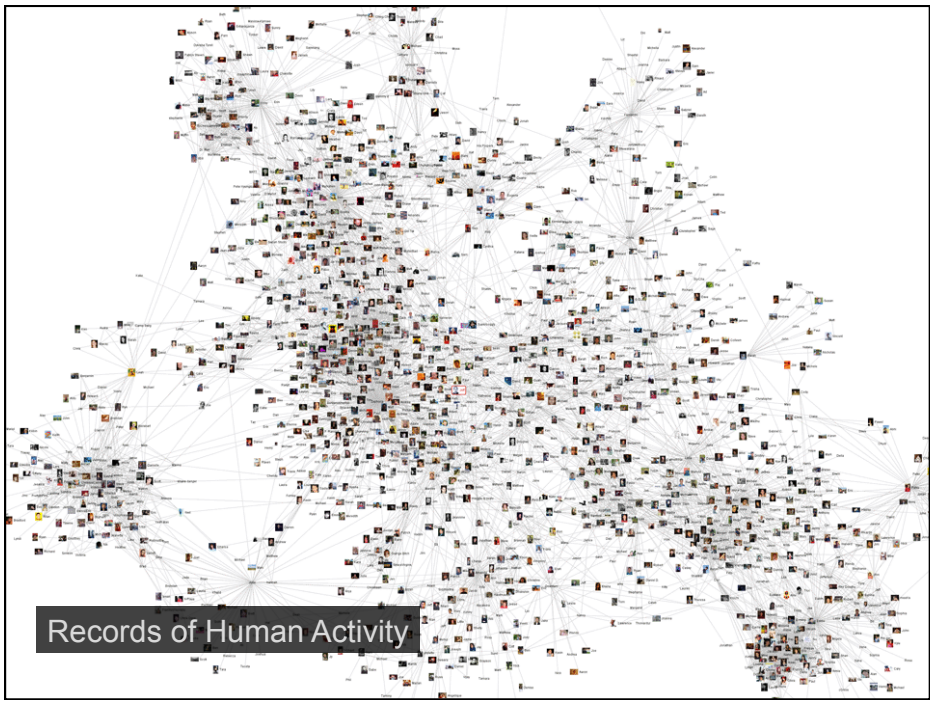
[Gantz 2007, 2011]

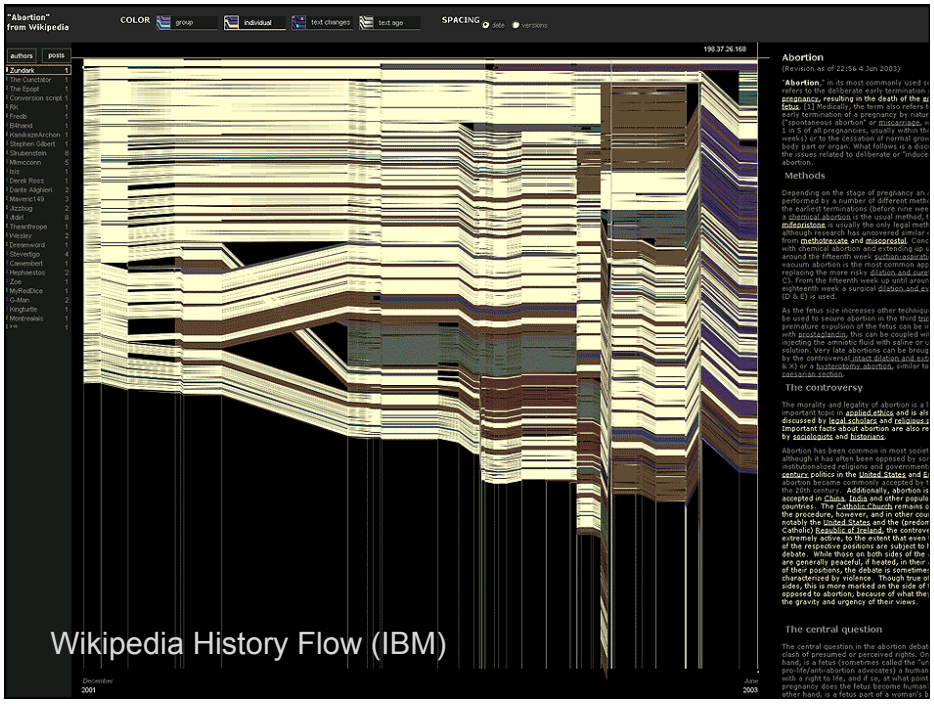


Physical Sensors

Image courtesy cabspotting.org

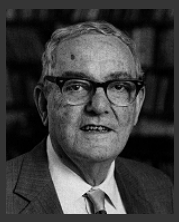

HR/ECG
1/min
97
Art
mmHg sys/dia
82/60
SpO2
%
99
RR/CO2
1/min
HR/ECG
1/min
79
Art
mmHg sys/dia
152/79
SpO2
%
95
RR/CO2
1/min
HR/ECG
1/min
64
Art
mmHg sys/dia
93/55
SpO2
%
99
RR/CO2
1/min
HR/ECG
1/min
99
SpO2
%
RR/Imp
1/min
Health & Medicine


Records of Human Activity



Wikipedia History Flow (IBM)

"What information consumes is rather obvious: **it consumes the attention of its recipients**. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it."



~Herb Simon
as quoted by Hal Varian
Scientific American
September 1995

What is visualization?

What is visualization?

Definition [www.oed.com]

1. **The action or fact of visualizing; the power or process of forming a mental picture or vision of something not actually present to the sight; a picture thus formed.**

2. **The action or process of rendering visible.**

What is visualization?

"Transformation of the symbolic into the geometric"
[McCormick et al. 1987]

"... finding the artificial memory that best supports our natural means of perception." [Bertin 1967]

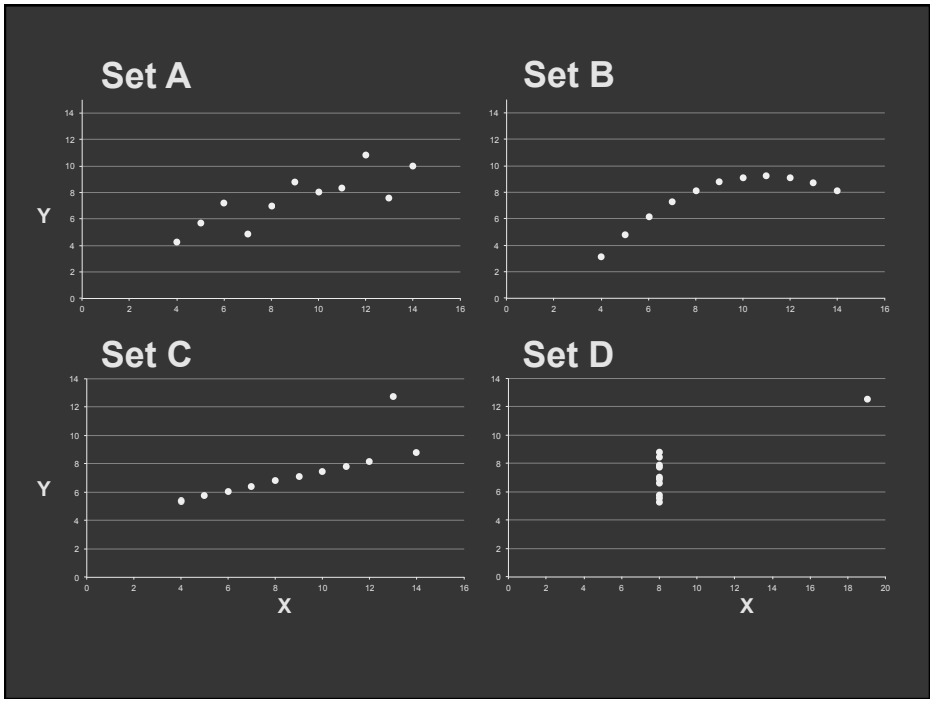
"The use of computer-generated, interactive, visual representations of data to amplify cognition."
[Card, Mackinlay, & Shneiderman 1999]

Set A		Set B		Set C		Set D	
X	Y	X	Y	X	Y	X	Y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.11	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Summary Statistics Linear Regression

$u_X = 9.0$ $\sigma_X = 3.317$ $Y = 3 + 0.5\ X$
$u_Y = 7.5$ $\sigma_Y = 2.03$ $R^2 = 0.67$

[Anscombe 73]



Why do we create visualizations?

Why do we create visualizations?

Help with pattern recognition
It would be stupid not to use it
Can show important aspects
Compresses the data (higher bandwidth)
Aesthetically pleasing/increzses engagement
Pushing a bias
Representatiojn similar to orgiinal form
Visual analogies
More direct perhaps than text

Three functions of visualizations

Record information

- Photographs, blueprints, …

Support reasoning about information (analyze)

- Process and calculate
- Reason about data
- Feedback and interaction

Convey information to others (present)

- Share and persuade
- Collaborate and revise
- Emphasize important aspects of data

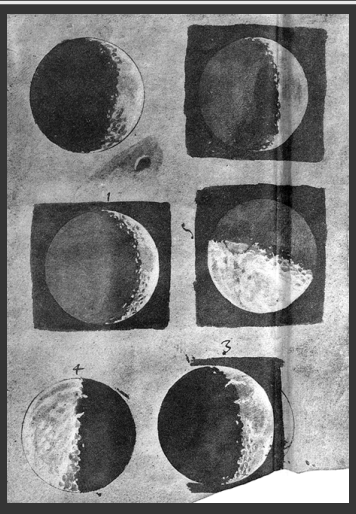
Record Information

Answer question

Gallop, Bay Horse “Daisy” [Muybridge 1884-86]

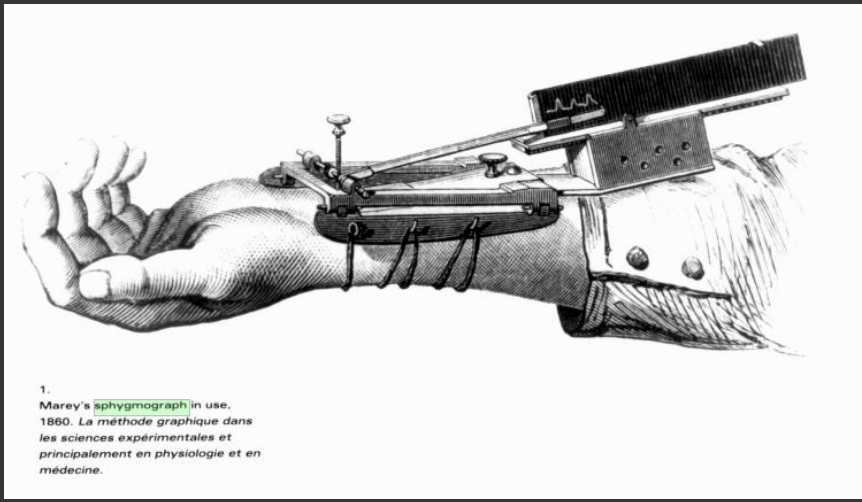
Photographs: Phases of the moon

Drawing: Phases of the moon



Galileo's drawings of the phases of the moon from 1616
http://galileo.rice.edu/sci/observations/moon.html

Other recording instruments

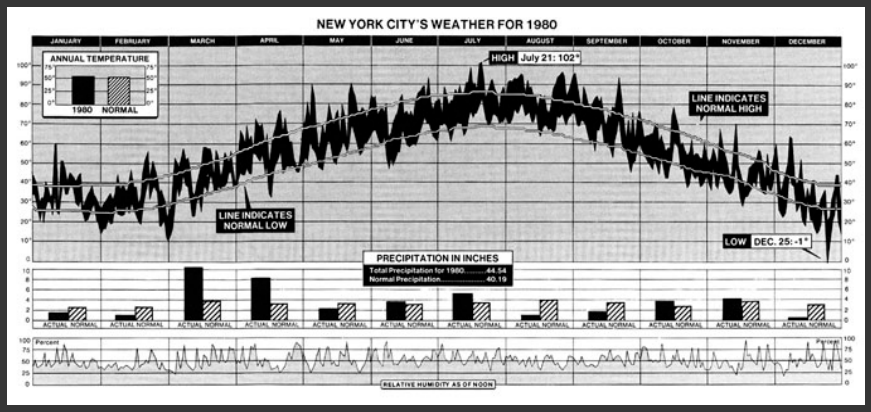


1.
Marey's sphygmograph in use.
1860. *La méthode graphique dans les sciences expérimentales et principalement en physiologie et en médecine.*

Marey's sphygmograph [from Braun 83]

Support Reasoning

Find patterns: New York weather



From the New York Times 1981

Make a decision: Challenger

HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

SRM No.	Cross Sectional View: Erosion Depth (in.)	Cross Sectional View: Perimeter Affected (deg)	Cross Sectional View: Nominal Dia. (in.)	Top View: Length Of Max Erosion (in.)	Top View: Total Heat Affected Length (in.)	Clocking Location (deg)
61A LH Center Field**	22A	None	None	0.280	None	None	36° -66°
61A LH CENTER FIELD**	22A	NONE	NONE	0.280	NONE	NONE	338°-18°
51C LH Forward Field**	15A	0.010	154.0	0.280	4.25	5.25	163
51C RH Center Field (prim)***	15B	0.038	130.0	0.280	12.50	58.75	354
51C RH Center Field (sec)***	15B	None	45.0	0.280	None	29.50	354
41D RH Forward Field	13B	0.028	110.0	0.280	3.00	None	275
41C LH Aft Field*	11A	None	None	0.280	None	None	--
41B LH Forward Field	10A	0.040	217.0	0.280	3.00	14.50	351
STS-2 RH Aft Field	2B	0.053	116.0	0.280	--	--	90

*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.

Clocking 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.

SRM-22 FORWARD FIELD JOINT HAD PUTTY PATH TO PRIMARY O-RING, BUT NO O-RING EROSION AND NO SOOT BLOWBY. OTHER SRM-22 FIELD JOINTS HAD NO BLOWHOLES IN PUTTY.

Blow By History
SRM-15 WORST BLOW-BY
- 2 CASE JOINTS (90°), (110°) ARC
- MUCH WORSE VISUALLY THAN SRM-22

SRM 22 BLOW-BY
- 2 CASE JOINTS (30-40°)

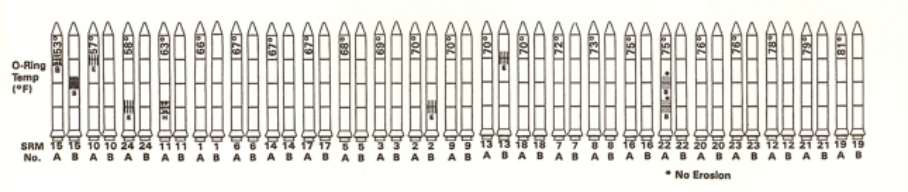
SRM-13A, 15, 16A, 18, 23A 24A
- NOZZLE BLOW-BY

HISTORY OF O-RING TEMPERATURES (DEGREES - F)

MOTOR	MAT	AMB	O-RING	WIND
DM-1	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

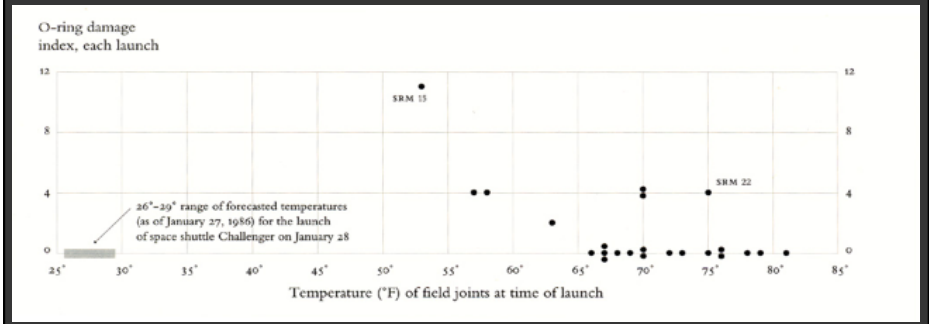
2 of 13 pages of material faxed to NASA by Morton Thiokol [from Tufte 1997]

Make a decision: Challenger



Visualizations drawn by Tufte show how low temperatures damage O-rings [Tufte 97]

Make a decision: Challenger



Visualizations drawn by Tufte show how low temperatures damage O-rings [Tufte 97]

See data in context: Cholera outbreak

In 1854 John Snow plotted the position of each cholera case on a map. [from Tufte 83]

See data in context: Cholera outbreak

Used map to support hypothesis Broad St. pump was the cause. [from Tufte 83]

Expand memory: Multiplication

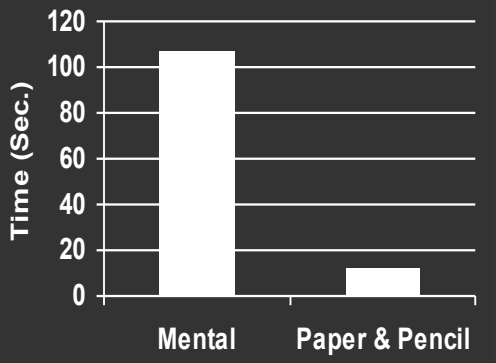
Class Exercise

Expand memory: Multiplication

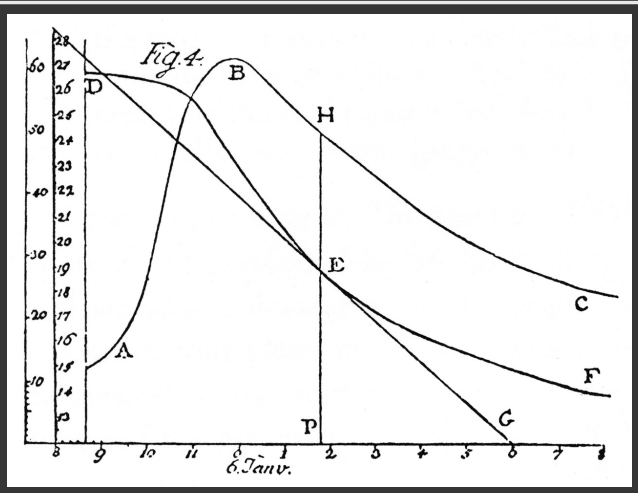
34
x 72

Expand memory: Multiplication

$$\begin{array}{r} 34 \\ \times\ 72 \\ \hline 68 \\ 2380 \\ \hline 2448 \end{array}$$

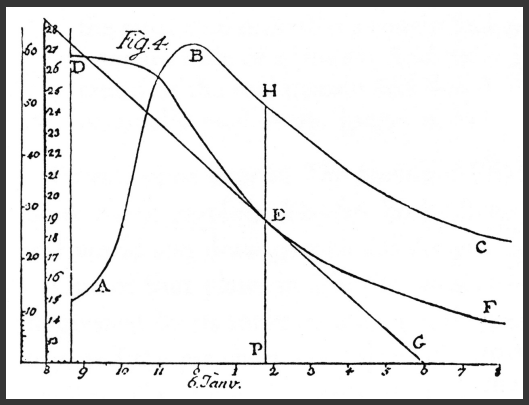


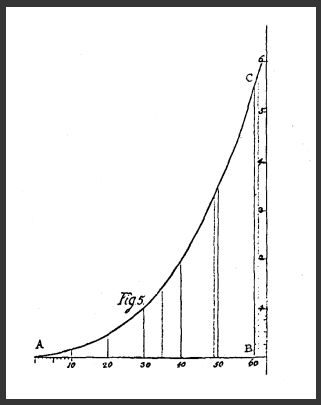
Graphical calculation: Evaporation



Johannes Lambert used graphs to study the rate of water evaporation as function of temperature [from Tufte 83]

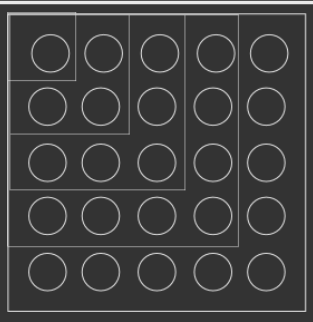
Graphical calculation: Evaporation



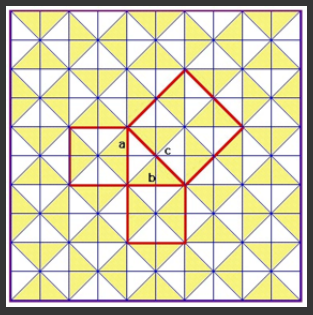


Johannes Lambert used graphs to study the rate of water evaporation as function of temperature [from Tufte 83]

Graphical calculation: Visual proofs



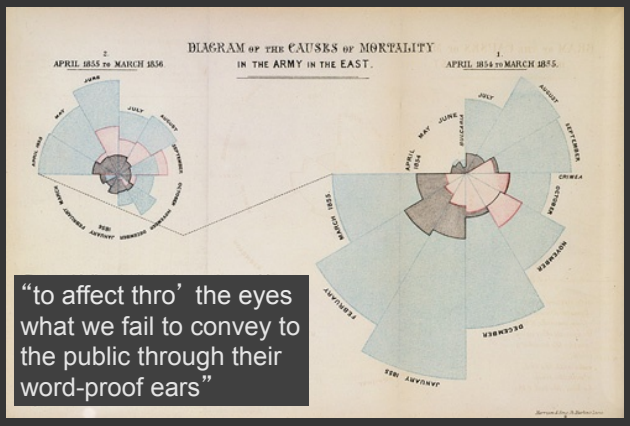
Sum of odd numbers:
$1 + 3 + 5 + 7 + 9 = 5^2$



Pythagorean theorem:
Chinese proof by dissection

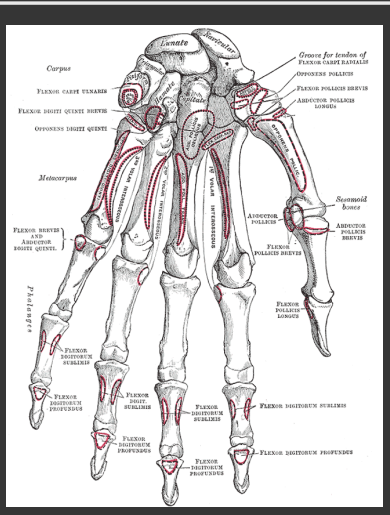
Convey Information to Others

Present argument



Crimean War Deaths [Nightingale 1858]

Inspire



Bones in hand [from 1918 edition]


Double helix model [Watson and Crick 53]

Visualization Research

Challenge

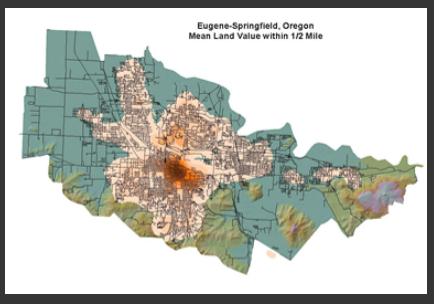
More and more unseen data

- Faster creation and collection

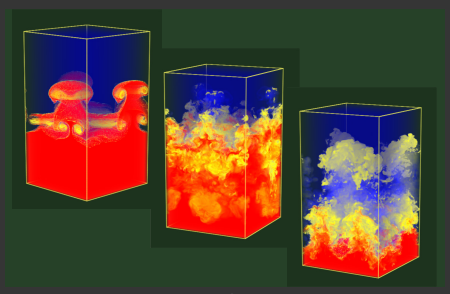
Challenge

More and more unseen data

- Faster creation and collection



Urban development planning
www.urbansim.org



Fluid flow
ctr.stanford.edu

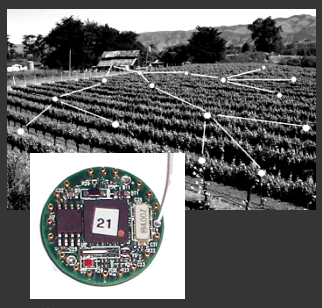
Simulation

Challenge

More and more unseen data

- Faster creation and collection

Sloan digital sky survey
www.sdss.org



Sensor networks [Hill 02]
www.xbow.com

Digital photography

Sensing

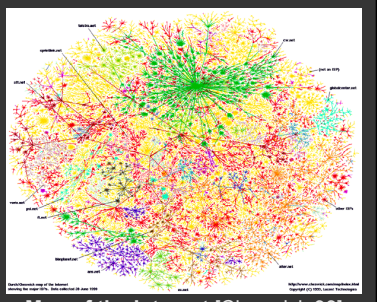
Challenge

More and more unseen data

- Faster creation and collection
- Faster dissemination

Photo sharing/annotation
flickr.com

Group Authored Encyclopedia
wikipedia.org



Map of the Internet [Cheswick 99]
research.lumeta.com

Internet

Challenge

More and more unseen data

- Faster creation and collection
- Faster dissemination

5 exabytes of new information in 2002 [Lyman 03]

161 exabytes in 2006 [Gantz 07]

1800 exabytes in 2011 [Gantz 11]

Need better tools and algorithms for visually conveying information

The ability to take data—to be able to **understand** it, to **process** it, to **extract value** from it, to **visualize** it, to **communicate** it—that's going to be a hugely important skill in the next decades, … because now we really do have **essentially free and ubiquitous data**. So the complimentary scarce factor is the ability to understand that data and extract value from it.

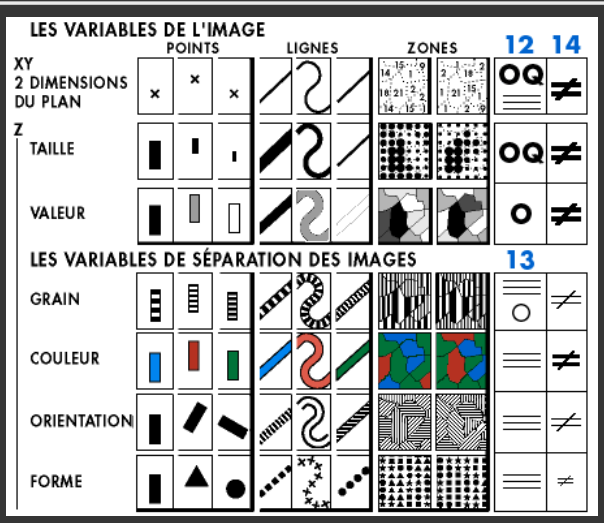
Hal Varian, Google's Chief Economist
The McKinsey Quarterly, Jan 2009

Goals of visualization research

1. **Understand how visualizations convey information to people**
   - What do people perceive/comprehend ?
   - How do visualizations correspond with mental models of data?
2. **Develop principles and techniques for creating effective visualizations**
   - Amplify perception and cognition
   - Strengthen connection between visualization and mental models of data

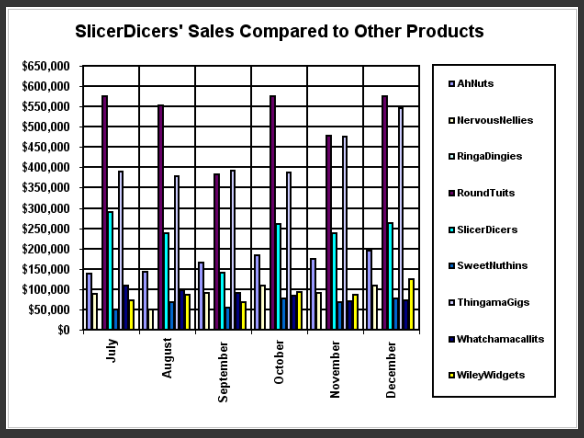
Topics

1. Data and image models

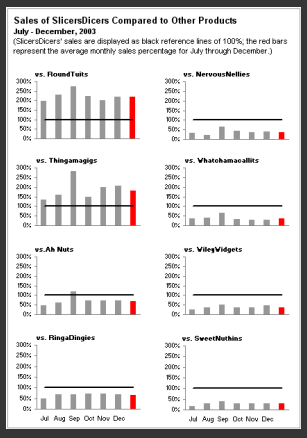


[Bertin, Graphics and Graphic Information Processing 1981]

2. Visualization Design

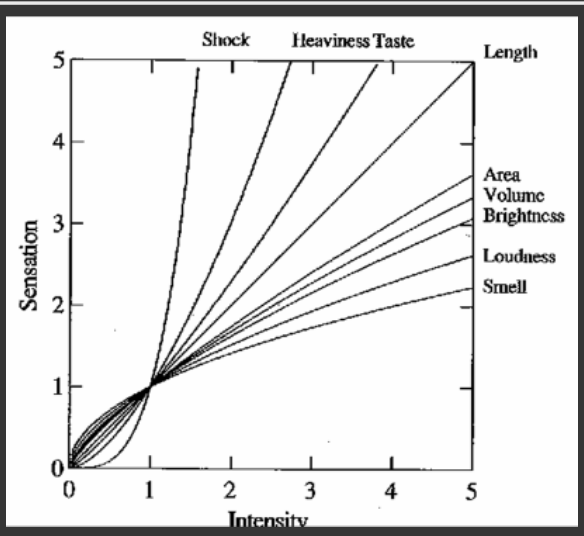


Problematic design



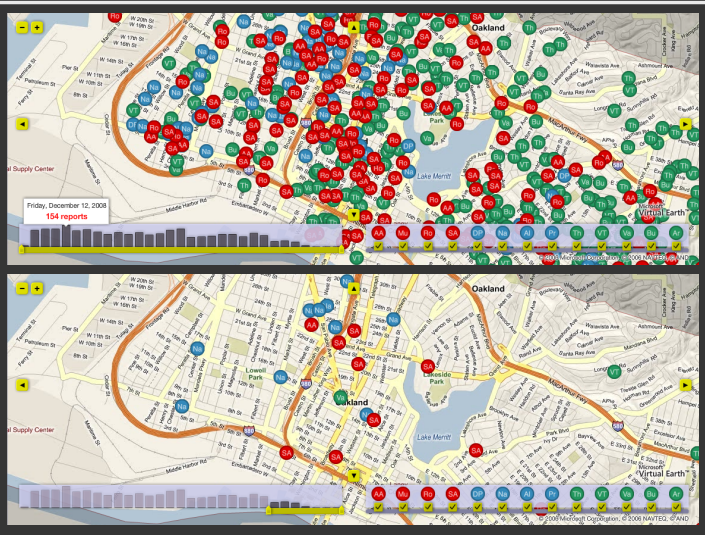
Redesign

3. Perception



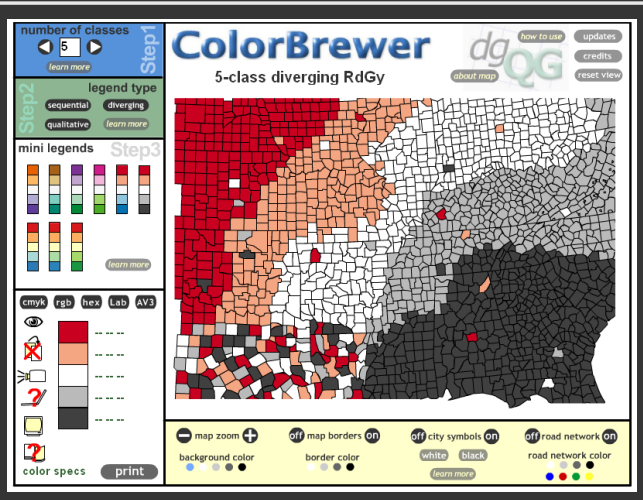
The psychophysics of sensory function [Stevens 61]

4. Interaction



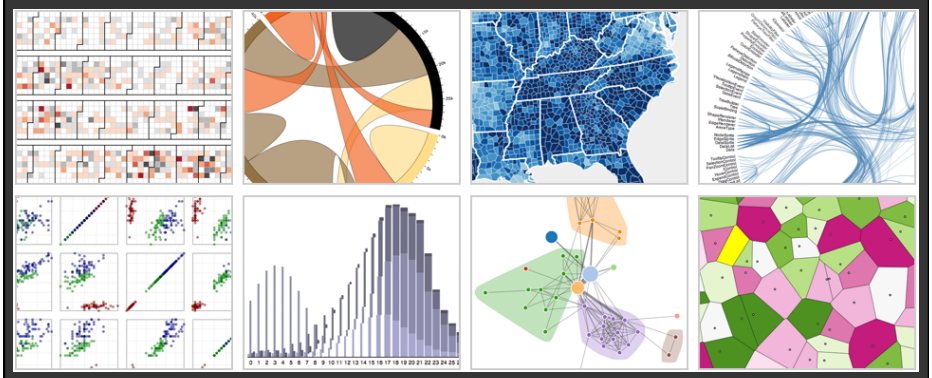
Oakland Crimespotting (crimespotting.org) [Stamen]

5. Color



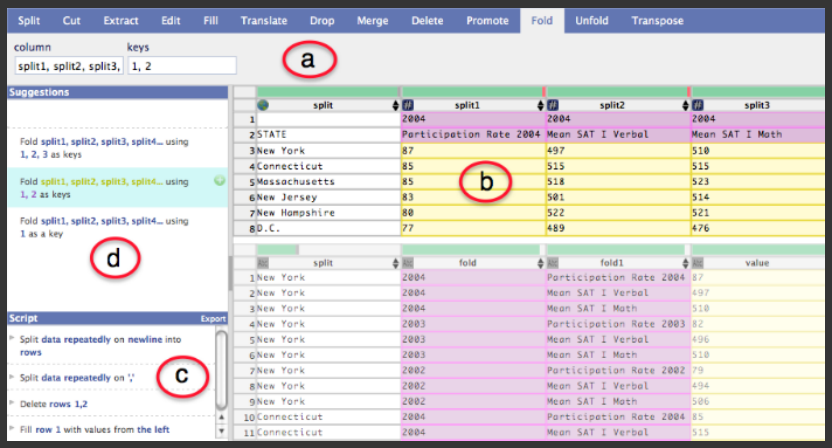
[from Cynthia Brewer http://www.personal.psu.edu/faculty/c/a/cab38/]

6. Building interactive visualizations with D3



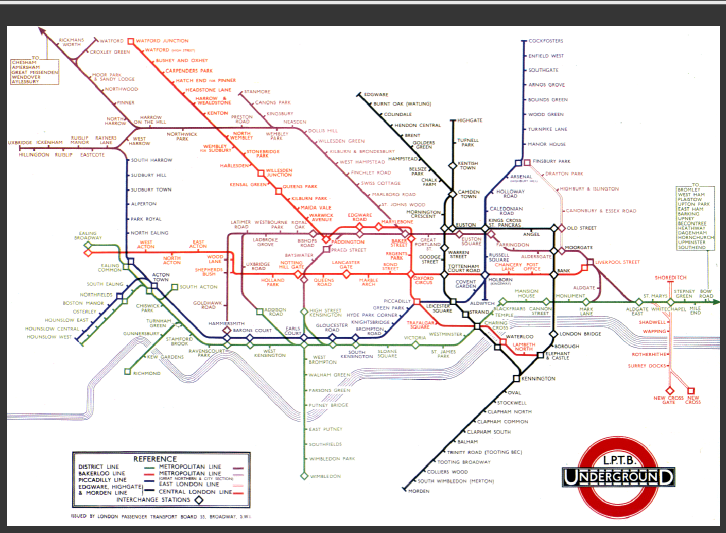
D3: Data Driven Documents [Bostock 2011]

7. Wrangling, Cleaning and Profiling



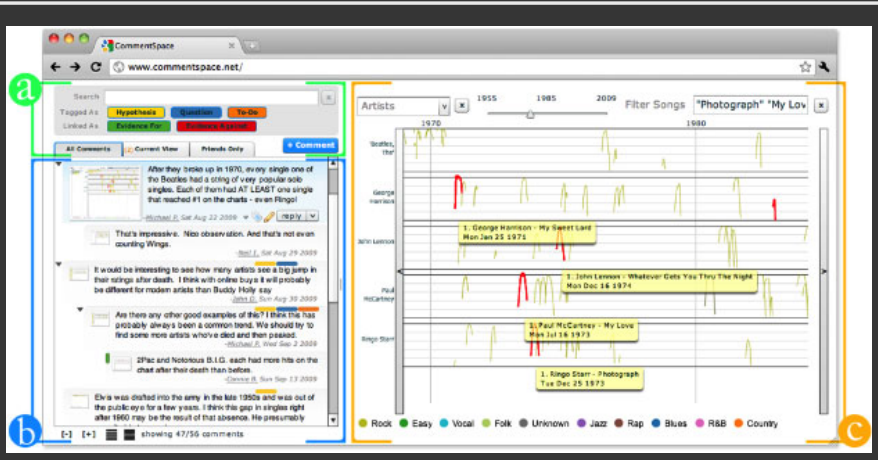
Proactive Wrangling [Guo 2011]

8. Spatial Layout



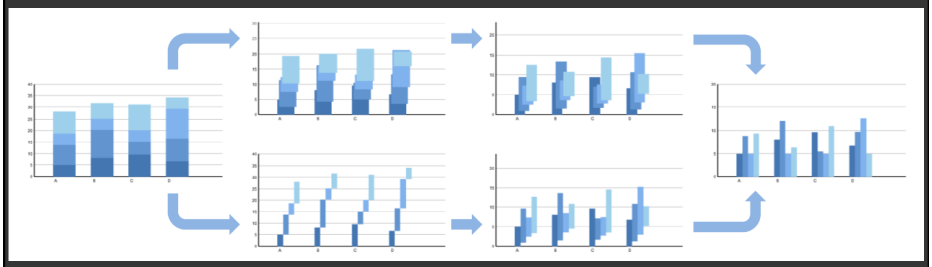
London underground [Beck 33]

9. Collaborative visualization



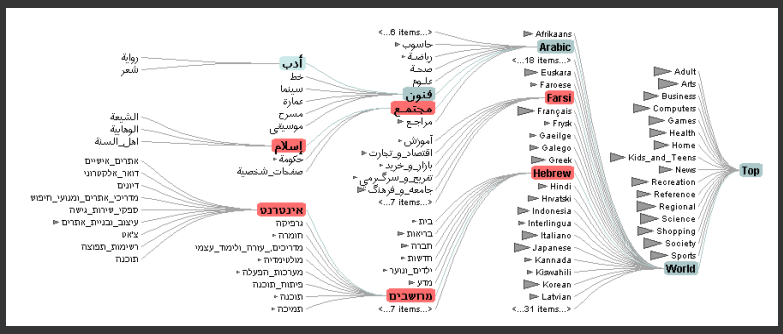
CommentSpace [Willett 2011]

10. Animation



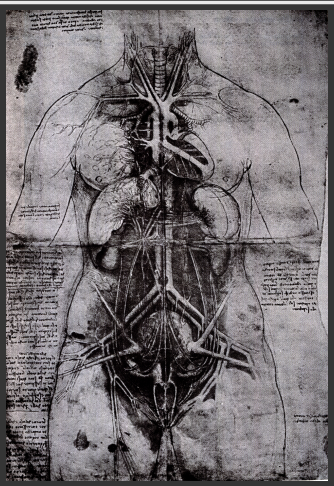
Animated Transitions [Heer 07]

10. Trees and graphs

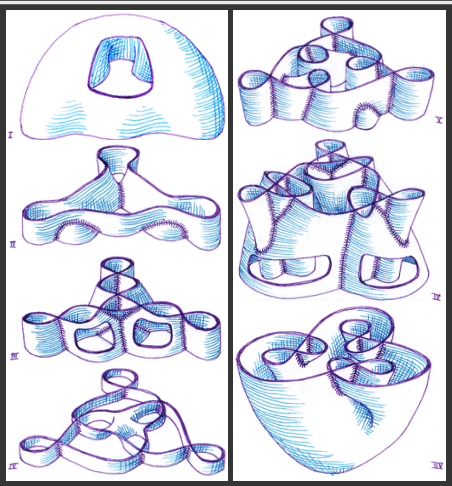


Degree-of-Interest Trees [Heer 2004]

11. Conveying structure

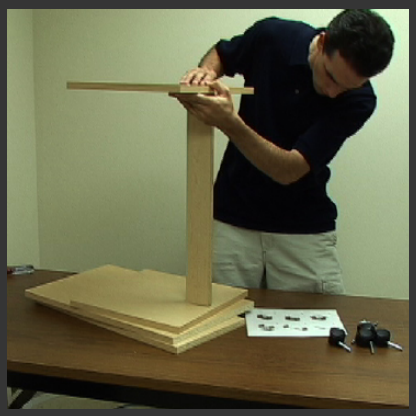


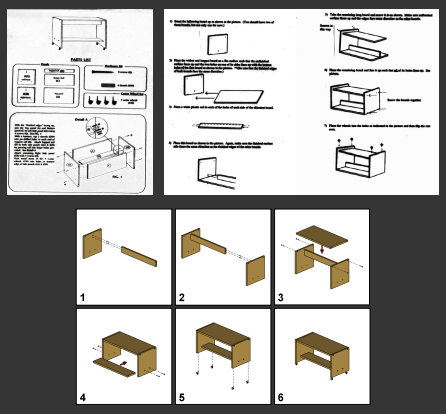
Principal Organs & Vascular System
[Leonardo da Vinci ca. 1490]



Strange Immersion of Torus in
3-Space [Curtis 92]

12. Identifying design principles

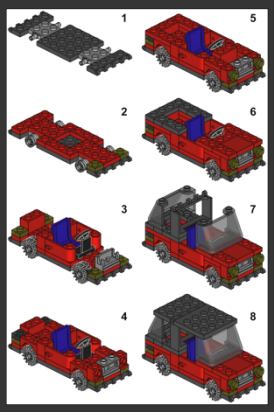


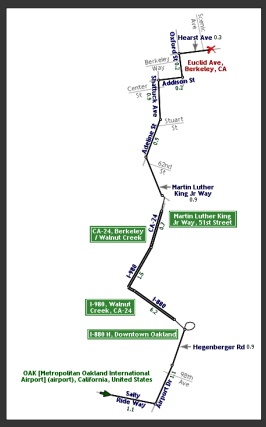


Testing effectiveness of 3 types of assembly instructions [Heiser 04]

Course Mechanics

Instructor: Maneesh Agrawala

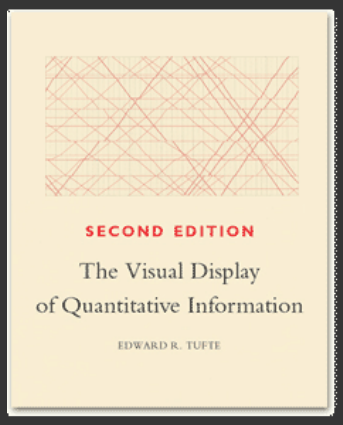


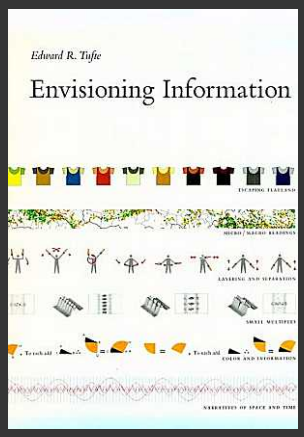


Course Goals

1. ***Design, evaluate and critique*** **visualizations**
2. ***Explore data*** **using existing visualization tools**
2. ***Implement*** **interactive data visualizations**
3. ***Gain*** **an overview of research and techniques**
4. ***Develop*** **a substantial visualization project**

Textbooks





See also: www.edwardtufte.com

Readings

- **Some from textbooks, also many papers**
  Username/Password: vis2013/vis2013Readings
- **Material in class will be loosely based on readings**
- **Readings should be read by start of class**
- **Post discussion comments on class wiki**
  Must post by 3pm on day of lecture
  You have 3 passes for the semester

Class home page
http://vis.berkeley.edu/courses/cs294-10-fa13/wiki

Requirements

Class participation (10%)

Assignment 1: Visualization Design **(10%)**

Assignment 2: Exploratory Data Analysis **(15%)**

Assignment 3: Creating Interactive Visualization Software **(25%)**

Final Project (40%)

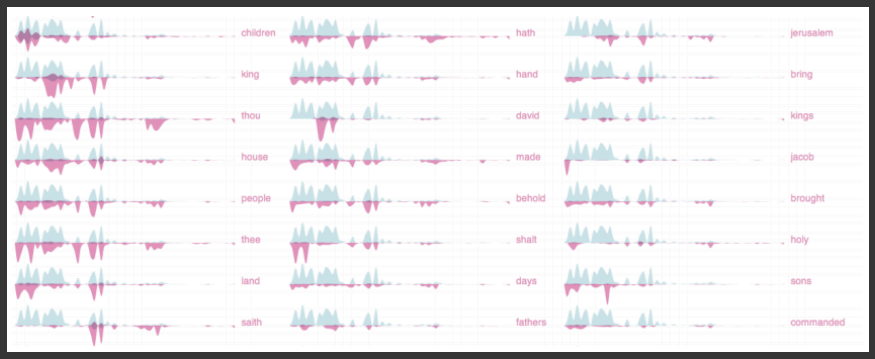
Final project

- **Visualization research project on topic of your choice**
- **2nd half of class**
- **Project write-up in form of a research paper**
- **Project presentations**
  1. Background research on project area
  2. Midway presentation on prototype solutions
  3. Final presentation – exact time to be determined

Projects from previous classes have been published

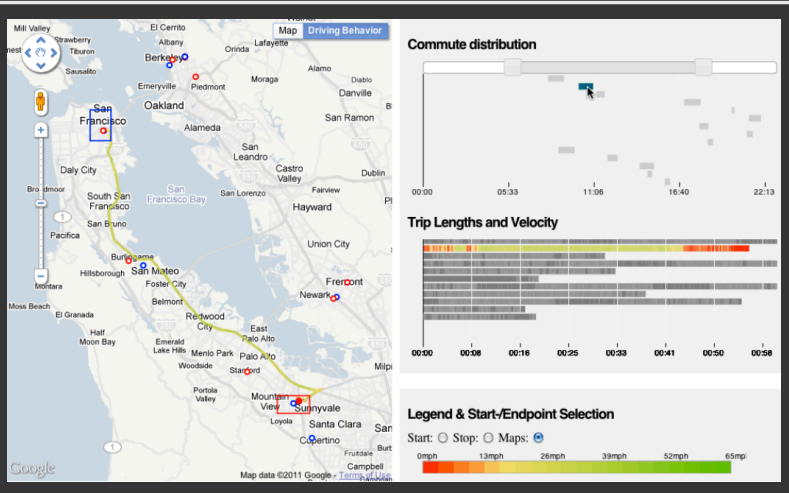
- IEEE Visualization
- IEEE Information Visualization
- SIGGRAPH

Text Co-Occurrences



Words that co-occur with Israel in the King James Bible [Brandon Liu 2011]

Visualizing Commuting Patterns



Interactive visualization of commuting patterns in San Francisco [Michael Porath 2011]

Assignment 1: Visualization Design

TEXAS DEPARTMENT of CRIMINAL JUSTICE

Home | Public Resources | Employment | About TDCJ | Online Services | Search

HOME | DEATH ROW | Death Row Information

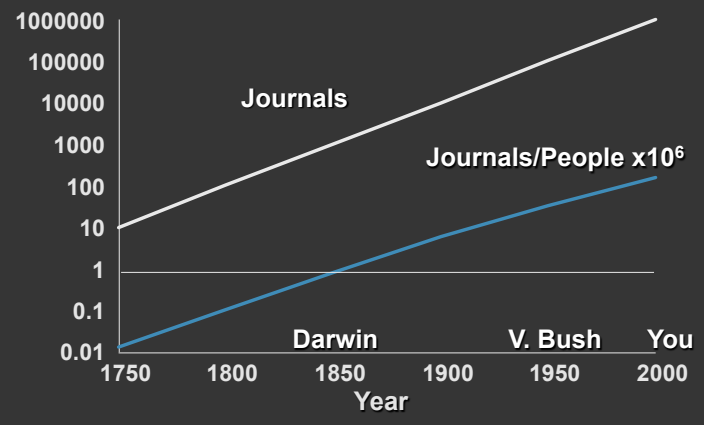
Executed Offenders

Execution	Link	Link	Last Name	First Name	TDCJ Number	Age	Date	Race	County
503	Offender Information	Last Statement	Feldman	Douglas	999326	55	07/31/2013	White	Dallas
502	Offender Information	Last Statement	Ross	Vaughn	999429	41	07/18/2013	Black	Lubbock
501	Offender Information	Last Statement	Quintanilla	John	999491	36	07/16/2013	Hispanic	Victoria
500	Offender Information	Last Statement	McCarthy	Kimberly	999287	52	06/26/2013	Black	Dallas
499	Offender Information	Last Statement	Chester	Elroy	999280	43	06/12/2013	Black	Jefferson
498	Offender Information	Last Statement	Williams	Jeffrey	999350	37	05/15/2013	Black	Harris
497	Offender Information	Last Statement	Parr	Carroll	999479	35	05/07/2013	Black	McLennan
496	Offender Information	Last Statement	Cobb	Richard	999467	29	04/26/2013	White	Cherokee

Due by 9am on Sep 11

Information

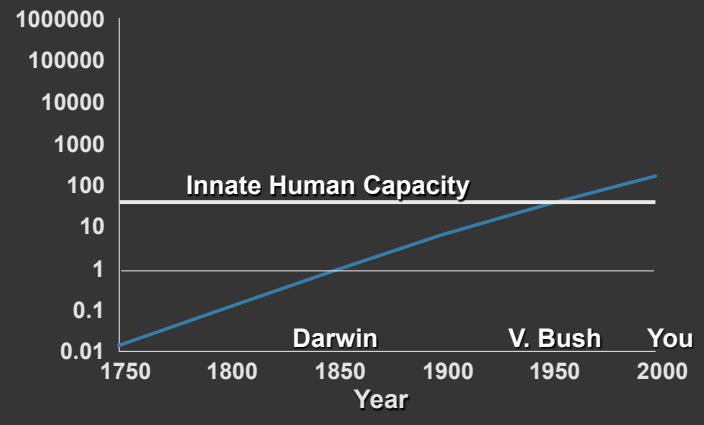
Journals/person increases 10X every 50 years



[slide from PARC UIR group]

Information Overload

Journals/person increases 10X every 50 years



[slide from PARC UIR group]

Assignment 1: Good and Bad Vis.

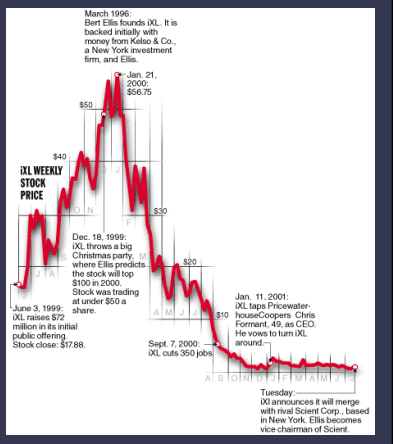
Find two visualizations one *good* and one *bad*

Use original sources

- Journals
- Science magazines
- Newspapers
- Textbooks

Make wiki page

- Clearly mark as good or bad
- Provide short explanation
- Be prepared to succinctly describe in class on Wed Jan 27



Due before class Mon Jan 25