

The Purpose of Visualization

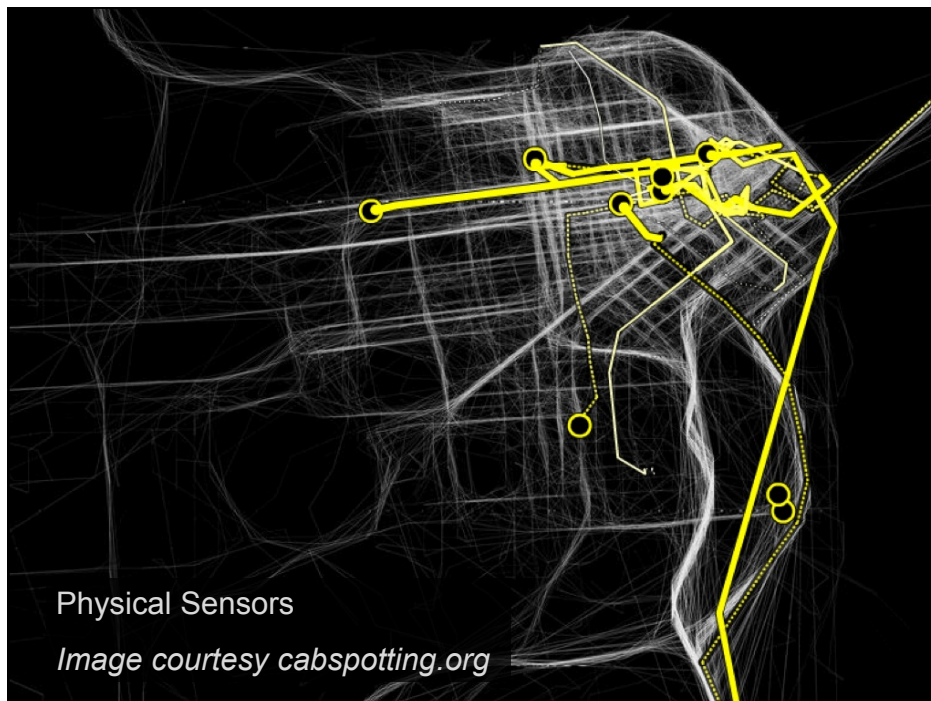
Maneesh Agrawala

CS 448B: Visualization
Spring 2016

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

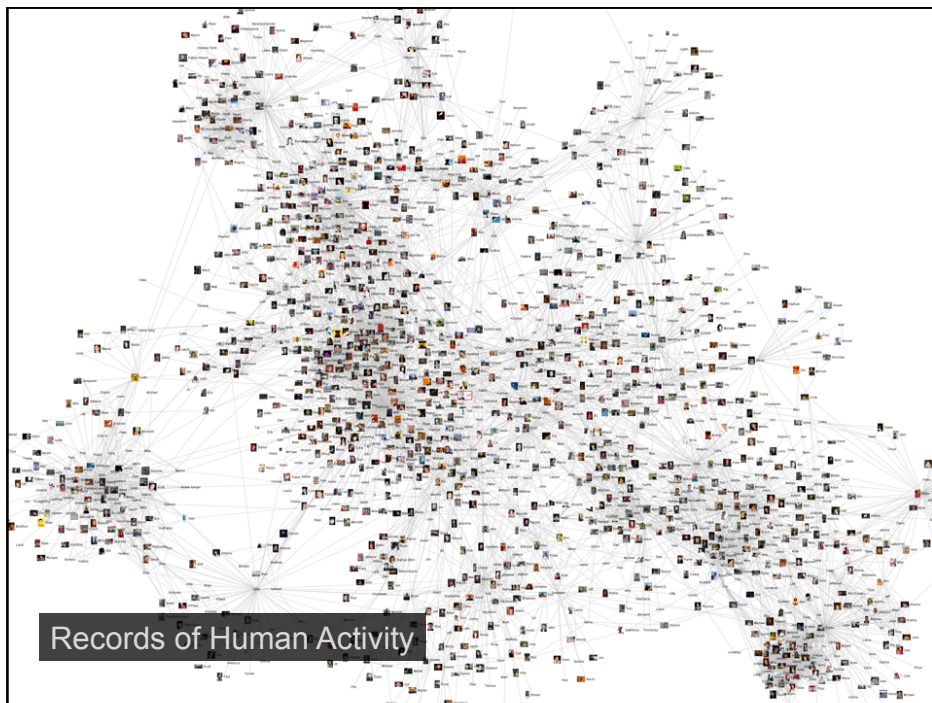
2011: 1800 exabytes
10x increase over 5 years

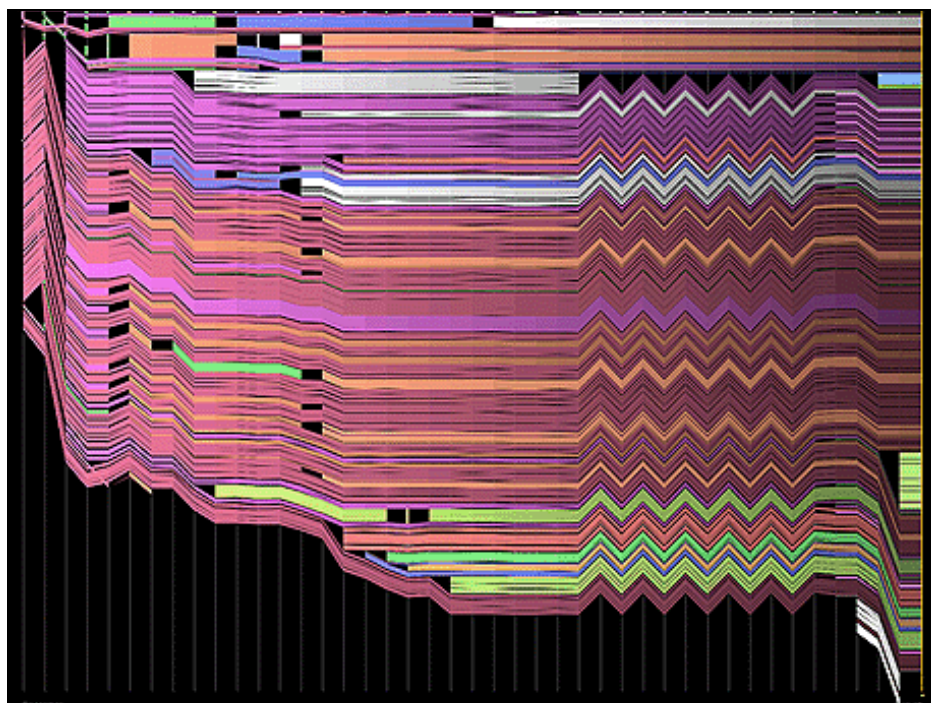
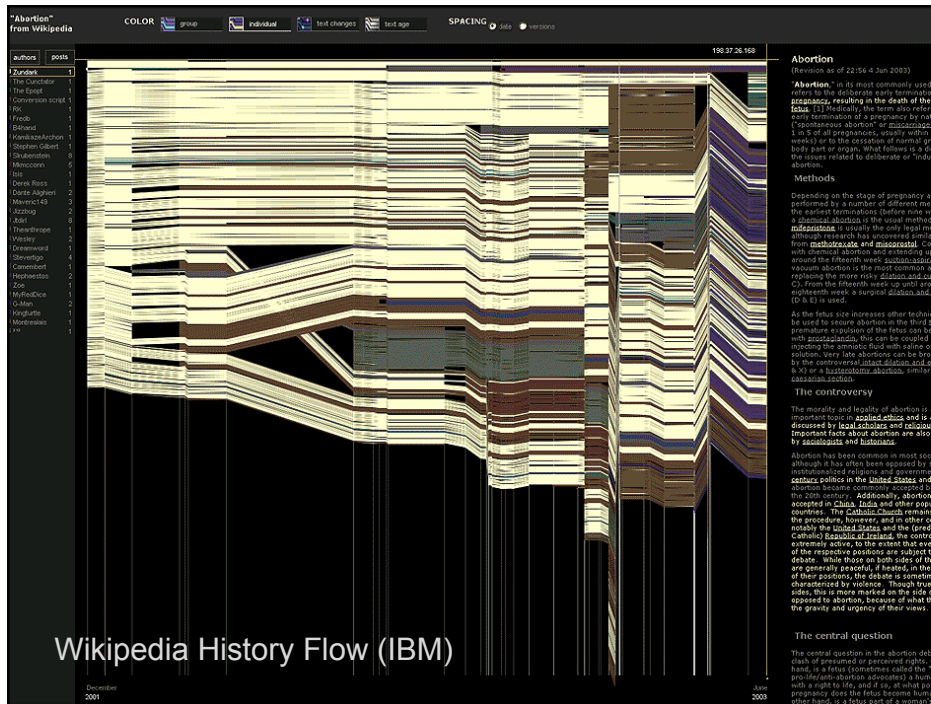
[Gantz 2007, 2011]



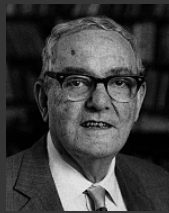
Physical Sensors

Image courtesy cabspotting.org





“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]

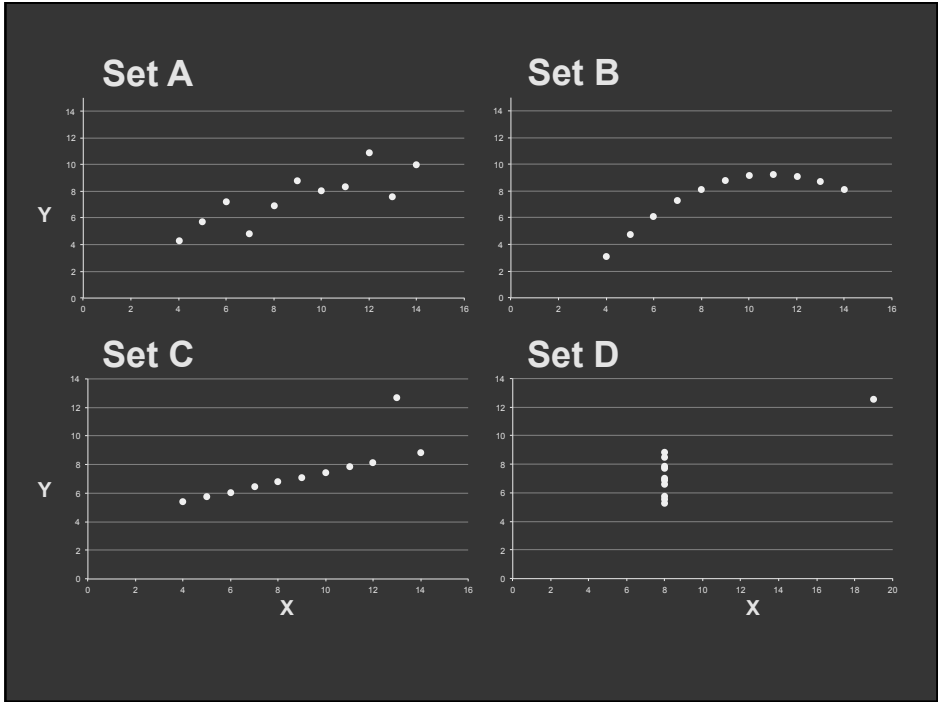
“The use of computer-generated, interactive, visual representations of data to amplify cognition.”
[Card, Mackinlay, & Shneiderman 1999]

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?

- Make info easier/faster to grasp (communicate)**
- To show patterns – more readily visible**
- Detect outliers**
- To condense the information**
- Add emotion to the data (point of view)**
- May allow interaction**
- People less intimidated**
- Help make predictions**

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

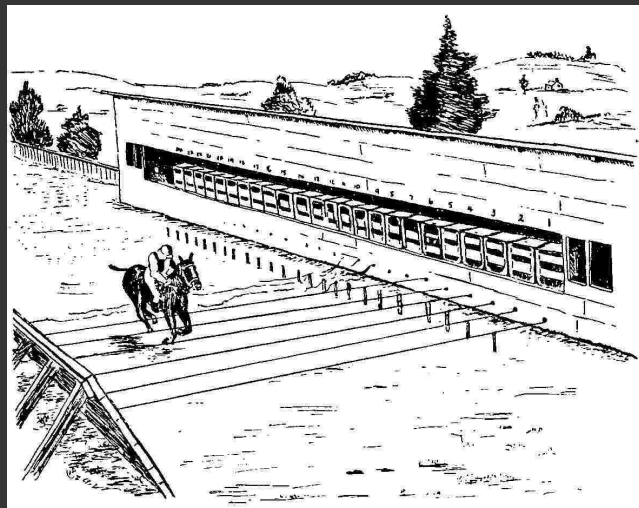
Record Information

Answer question



Gallop, Bay Horse "Daisy" [Muybridge 1884-86]

Answer question

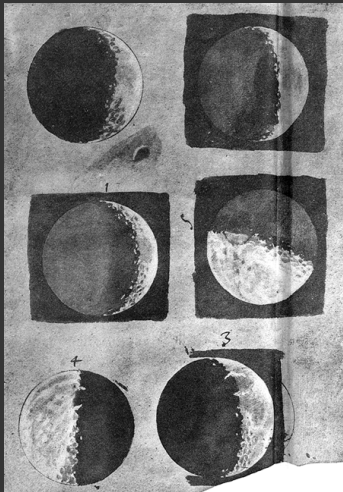


Gallop, Bay Horse "Daisy" [Muybridge 1884-86]

Photographs: Phases of the moon

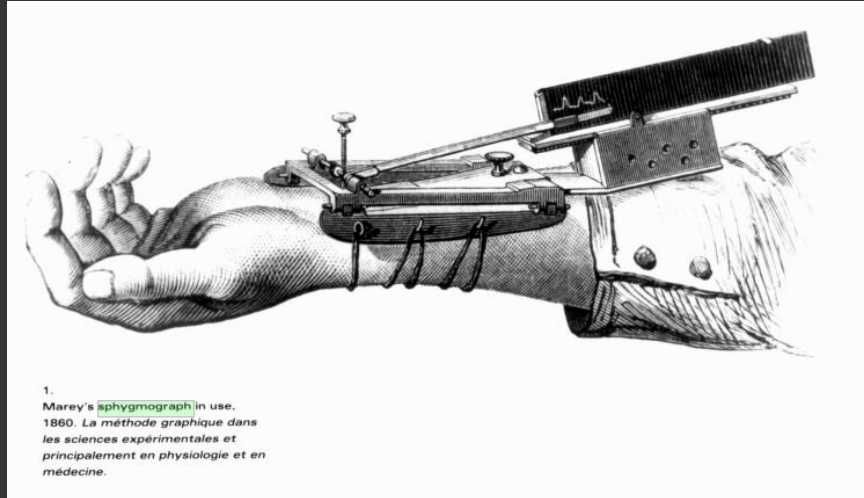


Drawing: Phases of the moon



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

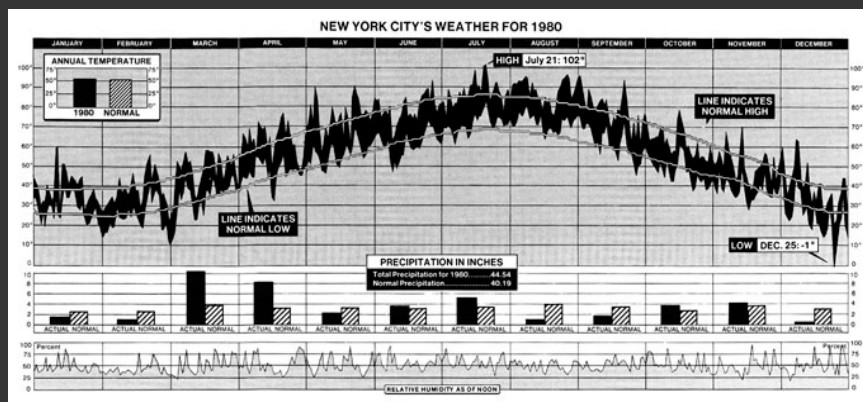
Other recording instruments



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	Cross Sectional View				Top View		Clocking Location (deg)
	Erosion Depth (in.)	Putty Gap (deg)	Nominal Dia. (in.)	Length of Max Erosion (in.)	Total Heat Affected Length (in.)		
SRM-15	None	None	0.280	None	None	367-368	181
SRM-15 LH Center Field**	0.010	154.0	0.280	4.25	5.25	354	275
SRM-15 LH Forward Field**	0.038	130.0	0.280	12.50	16.75	354	351
SRM-15 RH Center Field (prim)***	None	45.0	0.280	None	29.50	354	275
SRM-15 RH Center Field (sec)***	None	None	None	None	None	354	275
SRM-15 RH Forward Field	0.028	110.0	0.280	3.00	None	351	90
SRM-15 LH Aft Field*	None	None	0.280	None	None	351	90
SRM-15 LH Forward Field	0.040	217.0	0.280	3.00	14.50	351	90
SRM-15 RH Aft Field	0.053	116.0	0.280	--	--	351	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
 o 2 CASE JOINTS (30'), (110") SSC
 o MUCH WORSE VISUALLY THAN SRM-22

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

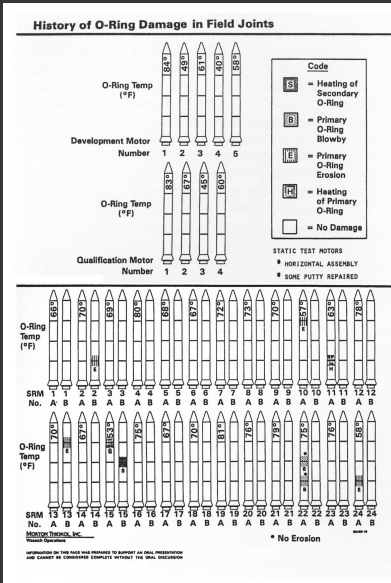
SRM-18 A, 15, 16A, 18, 23A 24A
 o NOZZLE BLOW-BY

HISTORY OF O-RING TEMPERATURES (DEGREES - F)

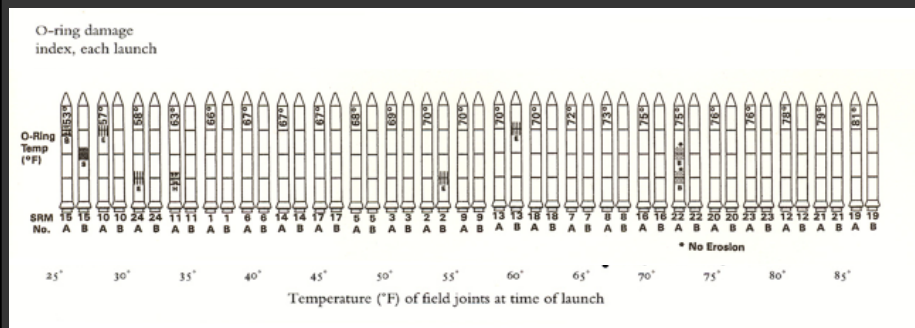
MOTOR	MBT	AMB	O-RING	WIND
DM-1	69	36	47	10 MPH
DM-2	76	45	52	10 MPH
DM-3	72.5	40	48	10 MPH
DM-4	76	48	51	10 MPH
SRM-15	52	64	53	10 MPH
SRM-22	77	78	75	10 MPH
SRM-25	55	26	29	10 MPH
			27	25 MPH

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

Make a decision: Challenger

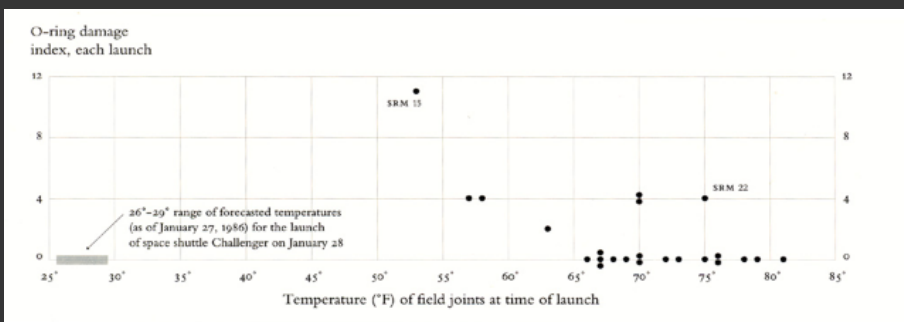


Make a decision: Challenger



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

Make a decision: Challenger



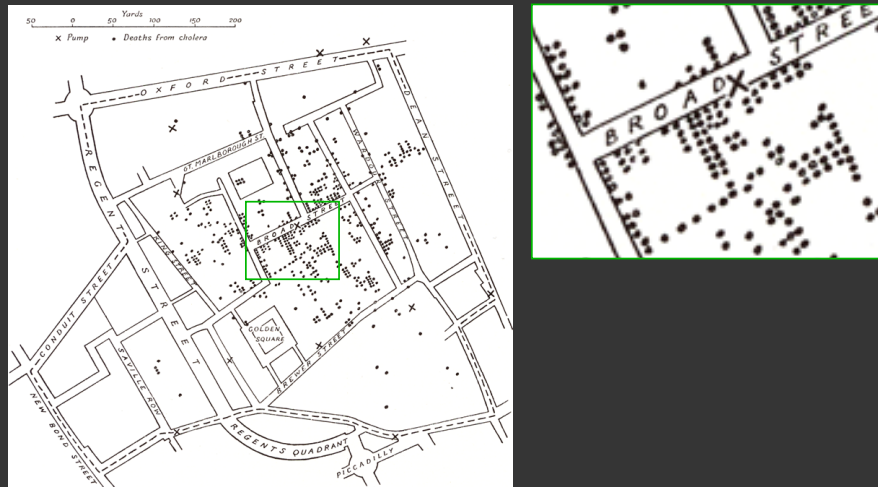
Visualizations drawn by Tufte show how low temperatures damage O-rings [Tufte 97]

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

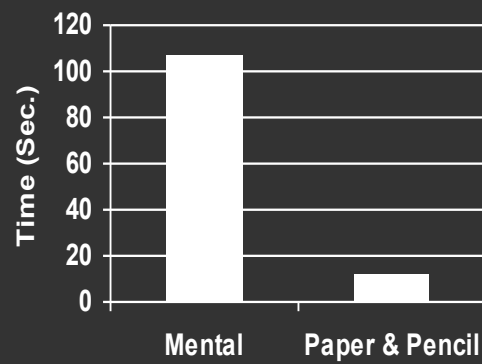
Class Exercise

Expand memory: Multiplication

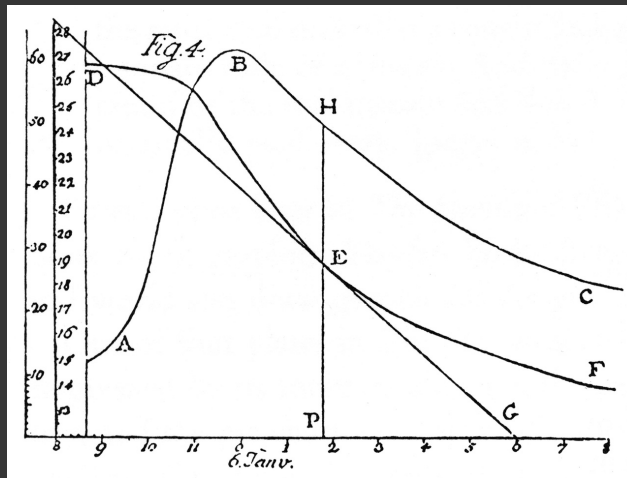
$$\begin{array}{r} 34 \\ \times 78 \\ \hline \end{array}$$

Expand memory: Multiplication

$$\begin{array}{r} 34 \\ \times 78 \\ \hline 272 \\ 2380 \\ \hline 2652 \end{array}$$

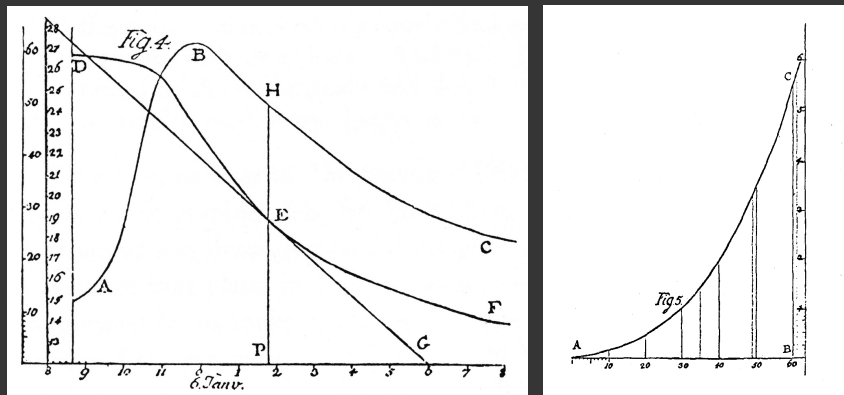


Graphical calculation: Evaporation



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

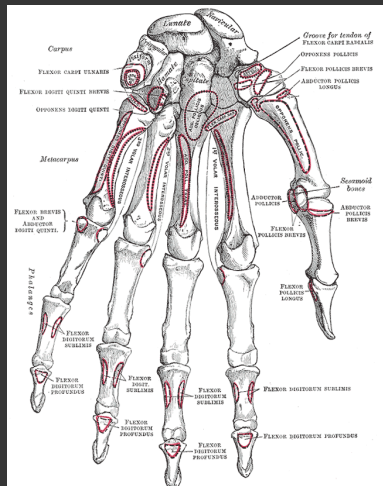
Graphical calculation: Evaporation



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

Convey Information to Others

Inspire

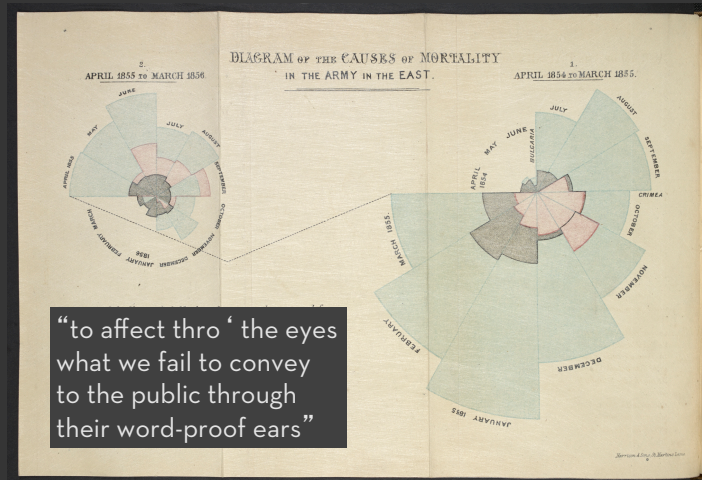


Bones in hand [from 1918 edition]



Double helix model [Watson and Crick 53]

Present argument



Crimean War Deaths [Nightingale 1858]

Visualization Research

Challenge

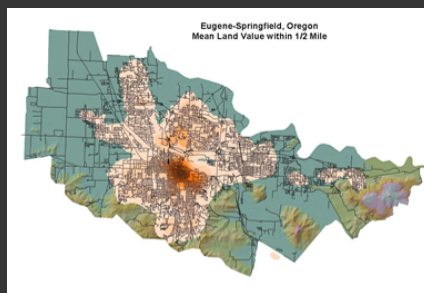
More and more unseen data

- Faster creation and collection

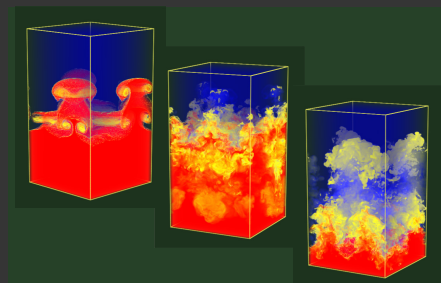
Challenge

More and more unseen data

- Faster creation and collection



Urban development planning
www.urbansim.org



Fluid flow
ctr.stanford.edu

Simulation

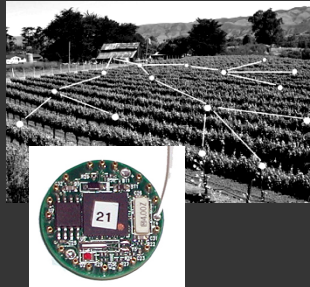
Challenge

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

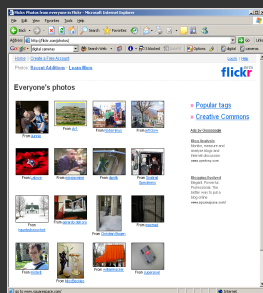
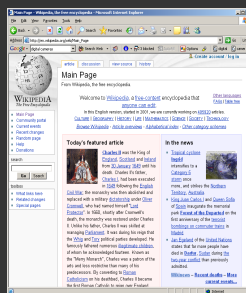
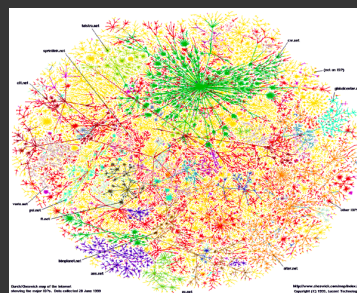


Photo sharing/annotation
flickr.com



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

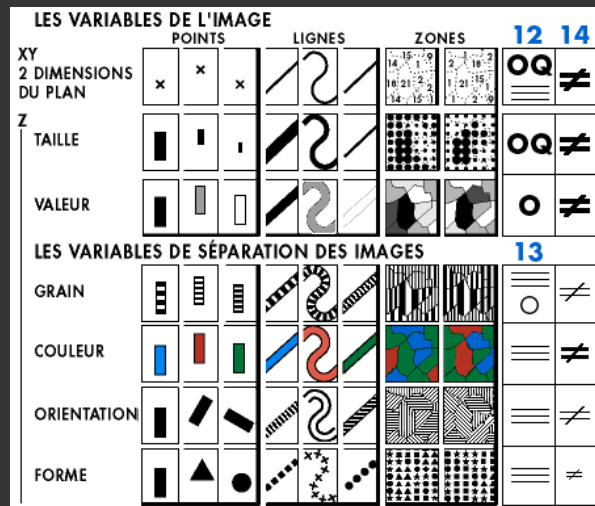
- What do people perceive/comprehend ?
- How do visualizations correspond with mental models of data?

2. Develop principles and techniques for creating effective visualizations

- Amplify perception and cognition
- Strengthen connection between visualization and mental models of data

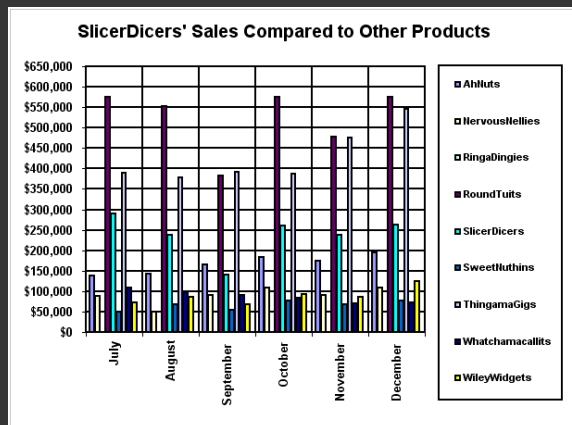
Topics

1. Data and image models

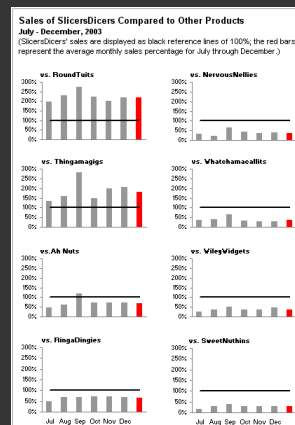


[Bertin, Graphics and Graphic Information Processing 1981]

2. Visualization Design

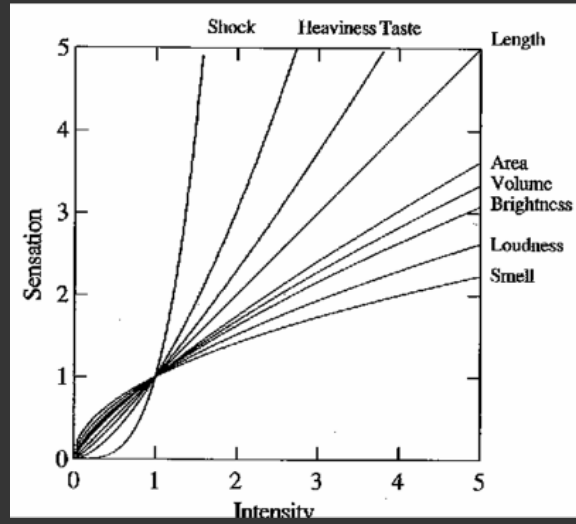


Problematic design



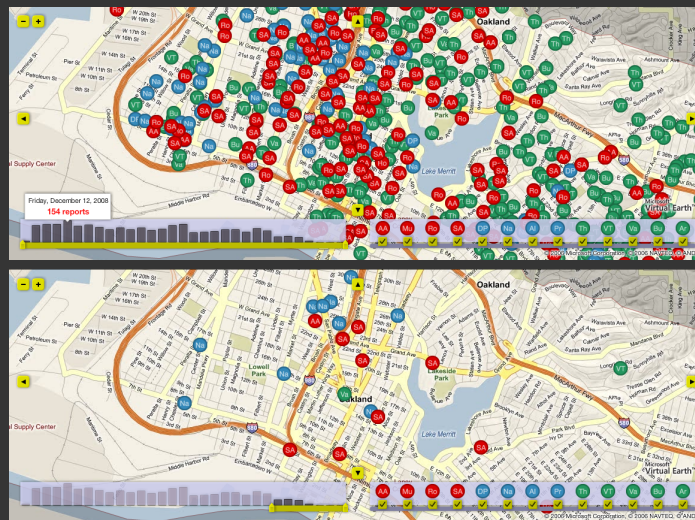
Redesign

3. Perception



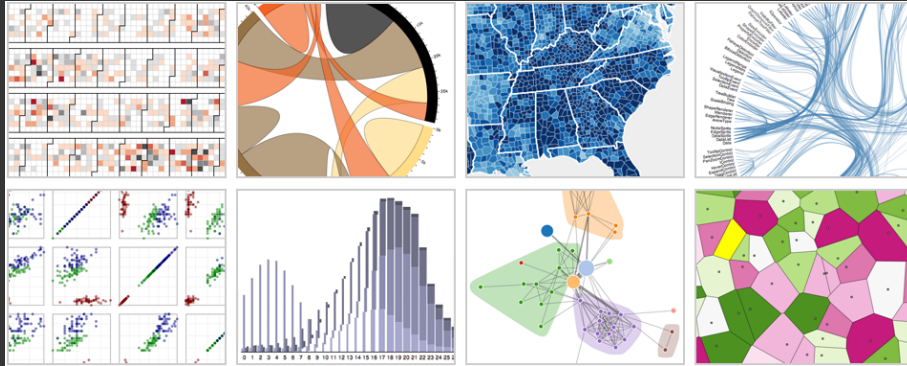
The psychophysics of sensory function [Stevens 61]

4. Interaction



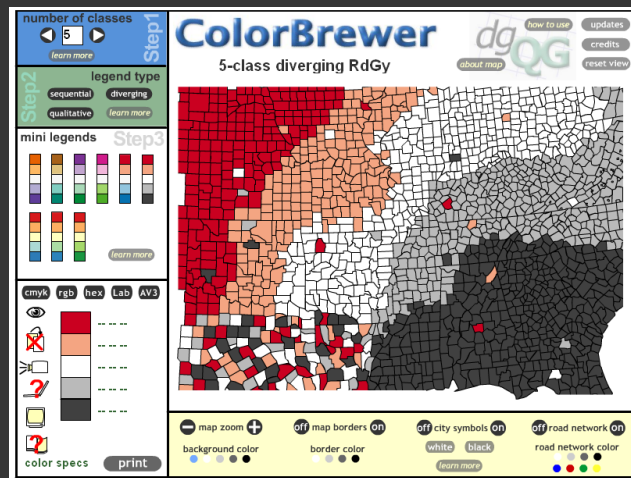
Oakland Crimespotting (crimespotting.org) [Stamen]

5. Building interactive visualizations with D3



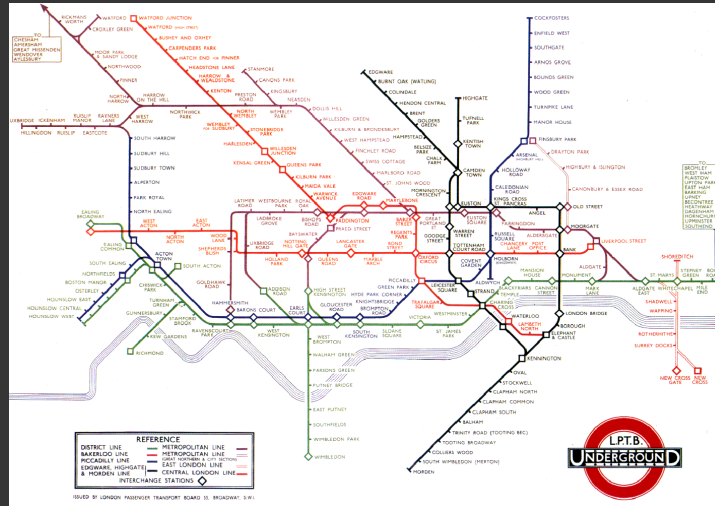
D3: Data Driven Documents [Bostock 2011]

6. Color



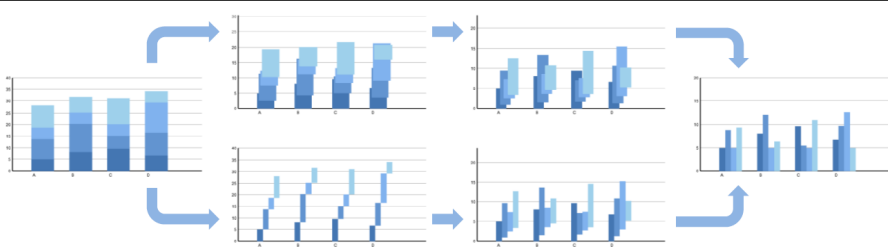
[from Cynthia Brewer <http://www.personal.psu.edu/faculty/c/a/cab38/>]

7. Spatial Layout



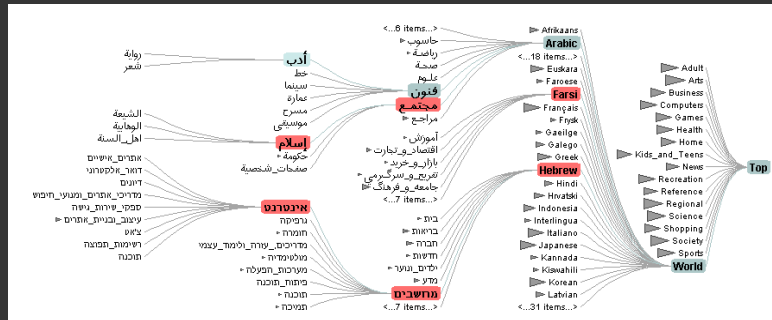
London underground [Beck 33]

8. Animation



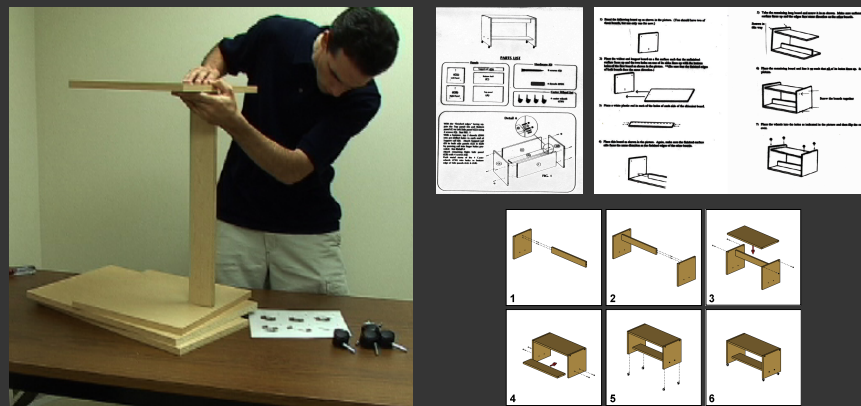
Animated Transitions [Heer 07]

9. Trees and graphs



Degree-of-Interest Trees [Heer 2004]

10. Identifying design principles



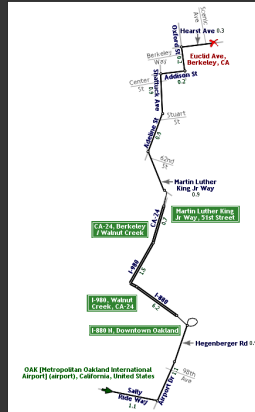
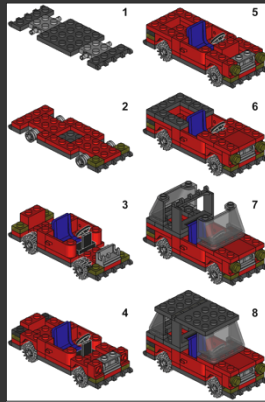
Testing effectiveness of 3 types of assembly instructions [Heiser 04]

Course Goals

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

Course Mechanics

Instructor: Maneesh Agrawala



Course Assistants

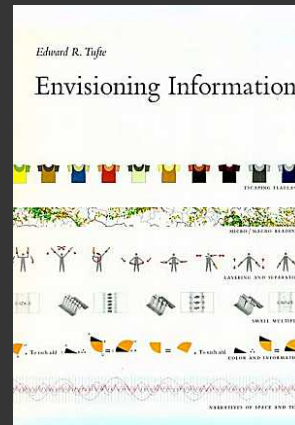
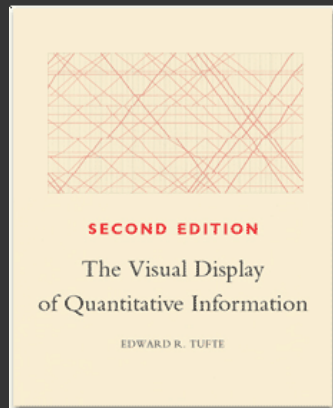
Scott Cheng
Ludwig Schubert
Peter Washington

Piazza is the best way to interact with us

<http://piazza.com/stanford/spring2016/cs448b>

Laptops?

Textbooks



See also: www.edwardtufte.com

Readings

- **Some from textbooks, also many papers**
Many open to public, some require SUNetID/Password
- **Material in class will be loosely based on readings**
- **Readings should be read by start of class**
- **Post discussion comment using link on wiki**
Must post by 1:30pm on day of lecture
You have 2 passes for the quarter

Class home page

<http://web.stanford.edu/class/cs448b/>

Reading Responses

Good responses typically exhibit one or more

- Critiques of arguments made in the papers
- Analysis of implications or future directions for ideas in readings
- Insightful questions about the readings

Responses should not summaries

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%)

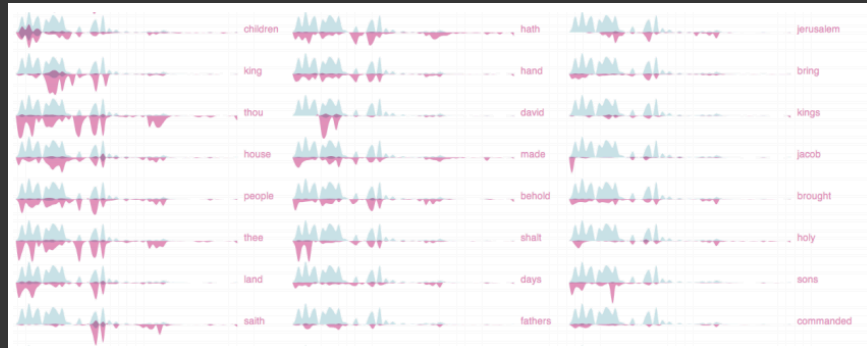
Final project

- Visualization research project on topic of your choice
- Last 4 weeks of course
- Project write-up in form of a research paper (6-8 pages)
- Two in-class project presentations
 1. Initial in-class status report (dates TBD)
 2. Final poster presentation (tentatively 6/1)

Projects from previous classes have been published

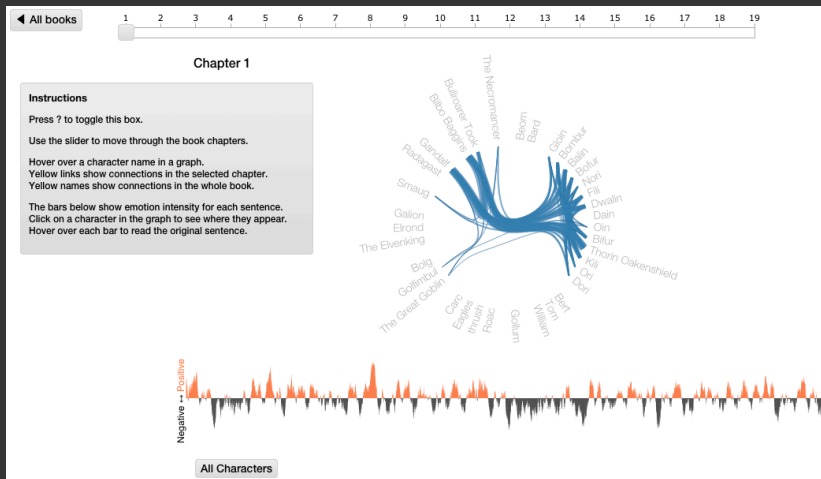
- IEEE Visualization
- IEEE Information Visualization
- SIGGRAPH

Text Co-Occurrences



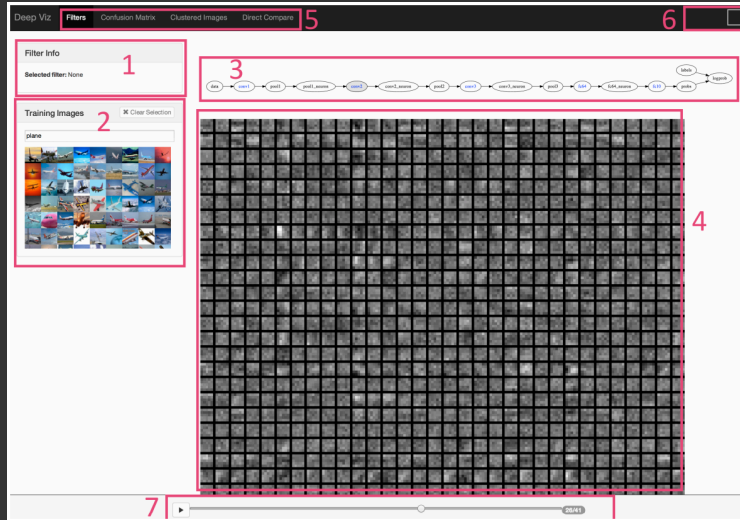
Words that co-occur with Israel in the King James Bible [Brandon Liu 2011]

Visualization of Narrative Structure



Character interactions and sentiment in *The Hobbit* [Bilenko, Miyakawa 2013]

deepviz: Visualizing Convolutional NNs



1) Filter details 2) Image selector 3) Network overview 4) Filter visualization
 5) Visualization selector 6) Selection helper 7) Animation slider [Bruckner,Rosen,Sparks 2013]

Assignment 1: Visualization Design

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	name	age	gender	raceethnicity	month	day	year	streetaddress	city	state	latitude	longitude	state_fp	county_fp
2	A'donte Was	16	Male	Black	February	23	2015	Clearview Ln Millbrook	AL	32.529577	-86.362829		1	51
3	Aaron Rutlec	27	Male	White	April	2	2015	300 block Irlt Pineville	LA	31.3217392	-92.43486		22	79
4	Aaron Siler	26	Male	White	March	14	2015	22nd Ave ani Kenosha	WI	42.5835597	-87.83571		55	59
5	Aaron Valdes	25	Male	Hispanic/Lati	March	11	2015	3000 Seminc South Gate	CA	33.9392976	-118.21946		6	37
6	Adam Jovicic	29	Male	White	March	19	2015	364 Hiwood Munroe Falls	OH	41.1485748	-81.429878		39	153
7	Adam Reinht	29	Male	White	March	7	2015	18th St and F Phoenix	AZ	33.4693799	-112.04332		4	13
8	Adrian Hernt	22	Male	Hispanic/Lati	March	27	2015	4000 Union I Bakersfield	CA	35.3956975	-119.00274		6	29
9	Adrian Solis	35	Male	Hispanic/Lati	March	26	2015	1500 Baylen Wilmington	CA	33.7930495	-118.27093		6	37
10	Alan Alverso	44	Male	White	January	28	2015	Pickett Runn Sunset	TX	30.6653042	-96.401482		48	41
11	Alan James	31	Male	White	February	7	2015	200 Abbie St Wyoming	MI	42.8932381	-85.660584		26	81
12	Albert Hansc	76	Male	White	April	26	2015	7th Ave and Hanford	CA	36.2109603	-119.58288		6	31
13	Alec Ouzoun	40	Male	White	May	12	2015	28 Paseo Vie Rancho Santi	CA	33.6533852	-117.61337		6	59
14	Alejandro Sa	Unknown	Male	Hispanic/Lati	February	20	2015	1200 E Airtes Houston	TX	29.9832049	-95.403857		48	201
15	Alexander Lc	31	Male	White	February	25	2015	25th St and F Terre Haute	IN	39.4629302	-87.37886		18	167
16	Alexander M	23	Male	White	April	6	2015	5700 block A indianapolis	IN	39.7669106	-86.149963		18	97
17	Alexander Ri	39	Male	Hispanic/Lati	May	30	2015	1128 Murfne Nashville	TN	36.1259117	-86.709015		47	37
18	Alexia Christl	25	Female	Black	April	30	2015	141 Pryor St Atlanta	GA	33.7512627	-84.393028		13	121
19	Alfredo Rials	54	Male	Hispanic/Lati	May	19	2015	4219 2nd Ro Arlington	VA	38.8731527	-77.10501		51	13
20	Alice Brown	24	Female	White	March	17	2015	Van Ness Av San Francisco	CA	37.7894309	-122.4221		6	75
21	Alvin Haynes	57	Male	Black	January	26	2015	1 Moreland I San Francisco	CA	37.6279793	-122.45393		6	81

Police Killings 2015

Due by 11:59pm on Sun Apr 3