The Purpose of Visualization

Maneesh Agrawala

CS 448B: Visualization
Winter 2020
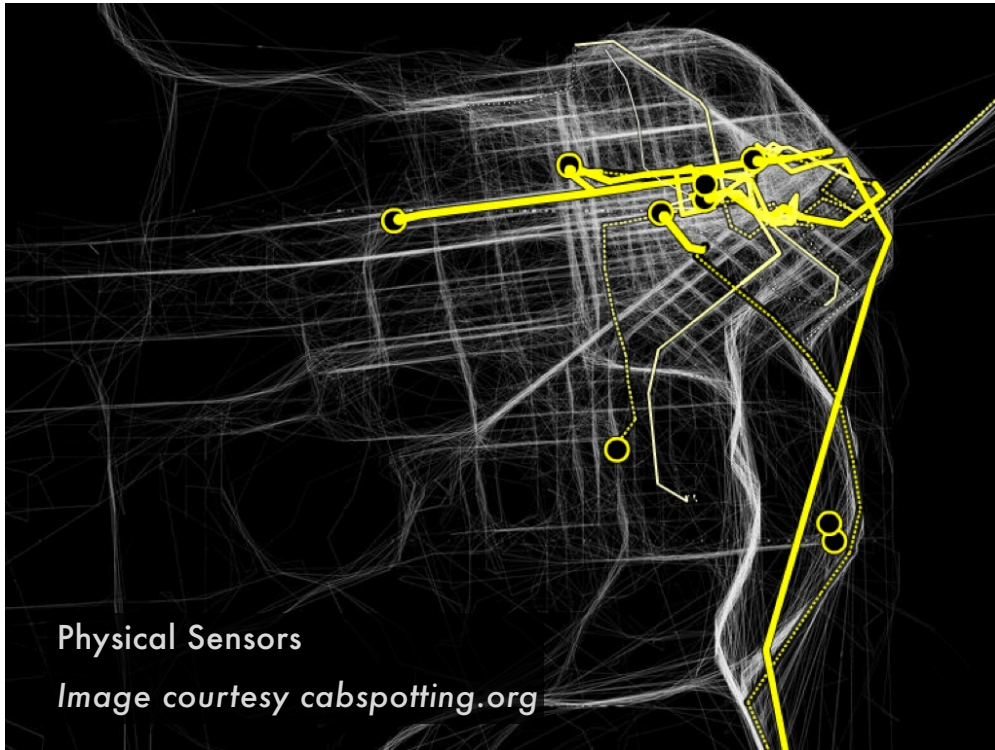

How much data (bytes) did we produce in 2016?
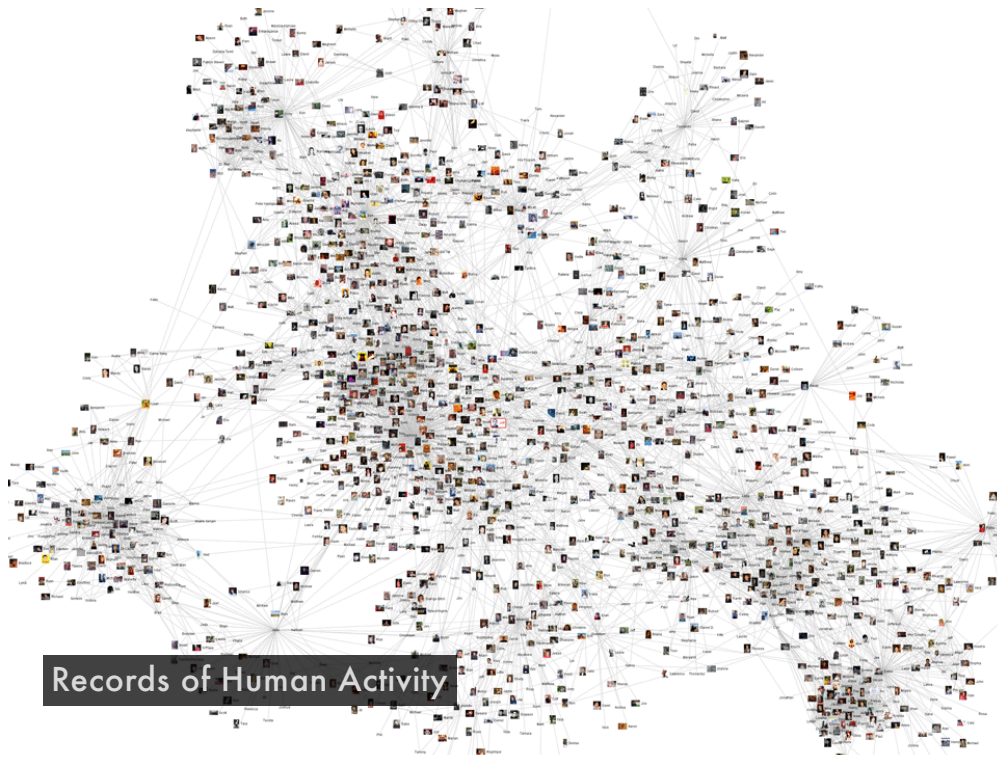
2016: 16.1 zetabytes

[Gantz 2017]

2016: 16.1 zetabytes
10x increase over 5 years

[Gantz 2017]

Physical Sensors
Image courtesy cabspotting.org

Health & Medicine

Records of Human Activity

Wikipedia: Collaborative Creation

Article   Talk   Read   View source   View history   Search Wikipedia   🔍

WIKIPEDIA
The Free Encyclopedia

Abortion   🔒

From Wikipedia, the free encyclopedia

For other uses, see Abortion (disambiguation).

Abortion is the ending of a pregnancy by removal or expulsion of an embryo or fetus before it can survive outside the uterus.[note 1] An abortion that occurs without intervention is known as a miscarriage or spontaneous abortion. When deliberate steps are taken to end a pregnancy, it is called an induced abortion, or less frequently "induced miscarriage". The unmodified word *abortion* generally refers to an induced abortion.[1][2] A similar procedure after the fetus has potential to survive outside the womb is known as a "late termination of pregnancy" or less accurately as a "late term abortion".[3]

When properly done, abortion is one of the safest procedures in medicine,[4][5] but unsafe abortion is a major cause of maternal death, especially in the developing world.[6] Making safe abortion legal and accessible reduces maternal deaths.[7][8] It is safer than childbirth, which has a 14 times higher risk of death in the United States.[9] Modern methods use medication or surgery for abortions.[10] The drug mifepristone in combination with prostaglandin appears to be as safe and effective as surgery during the first and second trimester of pregnancy.[10][11] The most common surgical technique involves dilating the cervix and using a suction device.[12] Birth control, such as the pill or intrauterine devices, can be used immediately following abortion.[11] When performed legally and safely on a woman who desires it, induced abortions do not increase the risk of long-term mental or physical problems.[13] In contrast, unsafe abortions (those performed by unskilled individuals, with hazardous equipment, or in unsanitary facilities) cause 47,000 deaths and 5 million hospital admissions each year.[13][14] The World Health Organization recommends safe and legal abortions be available to all women.[15]

Around 56 million abortions are performed each year in the world,[16] with about 45% done unsafely.[17] Abortion rates changed little between 2003 and 2008,[18] before which they decreased for at least two decades as access to family planning and birth control increased.[19] As of 2008, 40% of the world's women had access to legal abortions without limits as to reason.[20] Countries that permit abortions have different limits on how late in pregnancy abortion is allowed.[20]

Historically, abortions have been attempted using herbal medicines, sharp tools, forceful massage, or through other traditional methods.[21] Abortion laws and cultural or religious views of abortions are different around the world. In some areas abortion is legal only in specific cases such as rape, problems with the fetus, poverty, risk to a woman's health, or incest.[22] There is debate over the moral, ethical, and legal issues of abortion.[23][24] Those who oppose abortion often argue that an embryo or fetus is a human with a right to life, and they may compare abortion to murder.[25][26] Those who support the legality of abortion often hold that it is part of a woman's right to make decisions about her own body.[27] Others favor legal and accessible abortion as a public health measure.[28]

Contents
1 Types
  1.1 Induced
  1.2 Spontaneous

Abortion	
Other names	Induced miscarriage, termination of pregnancy
Specialty	Obstetrics and gynecology
ICD-10-PCS	O04 🔗
ICD-9-CM	779.6 🔗
MeSH	D000028
MedlinePlus	007382
	[edit on Wikidata]

Wikipedia: Collaborative Creation

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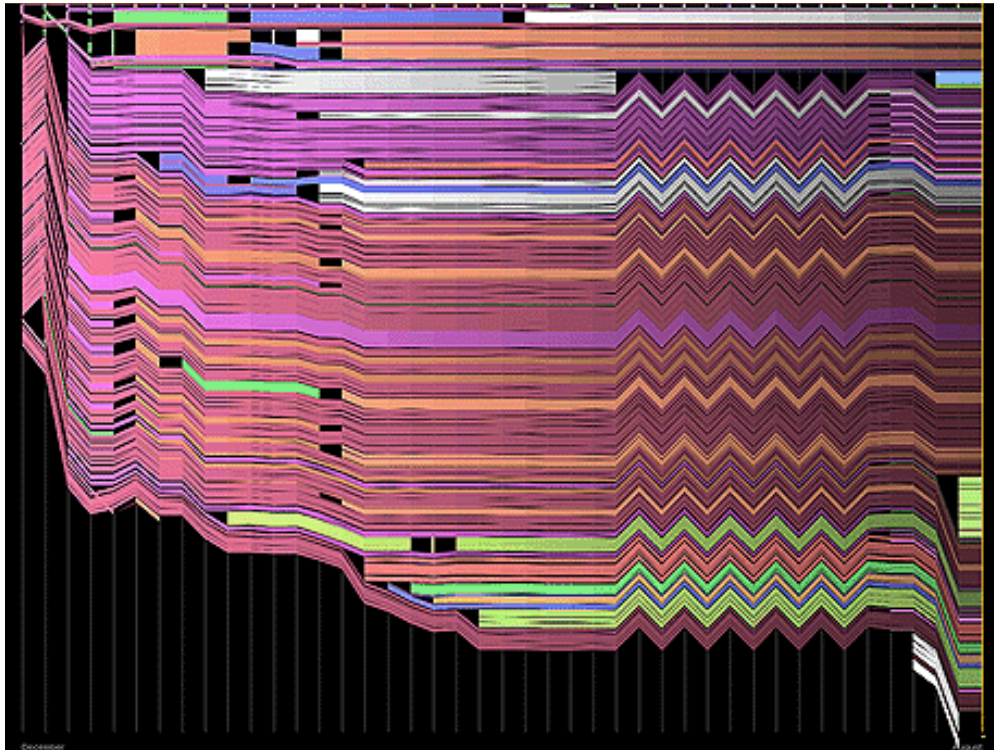
"The ability to take data—to be able to **understand** it, to **process** it, to **extract value** from it, to **visualize** it, to **communicate** it—that's going to be a hugely important skill in the next decades, … because now we really do have **essentially free and ubiquitous data**. So the complimentary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

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

"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."
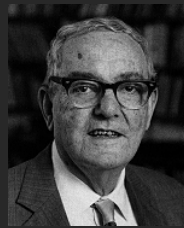
Herb Simon
as quoted by Hal Varian
Scientific American
September 1995

What is visualization?

Examples

Ages of first-time mothers in **1980**
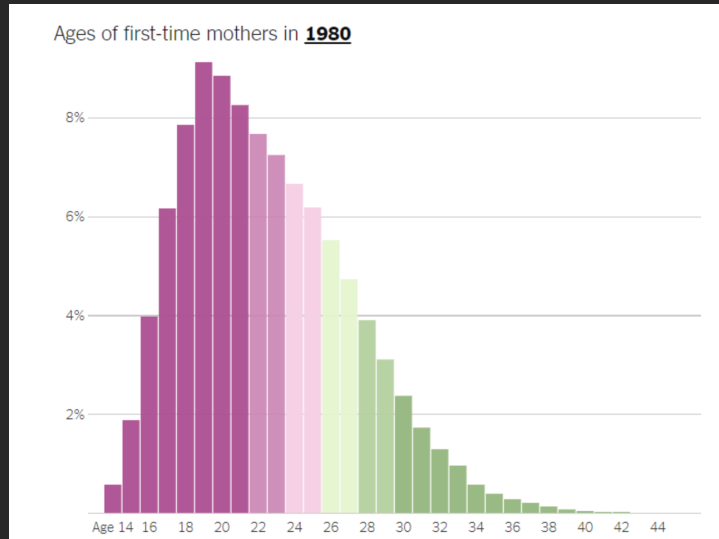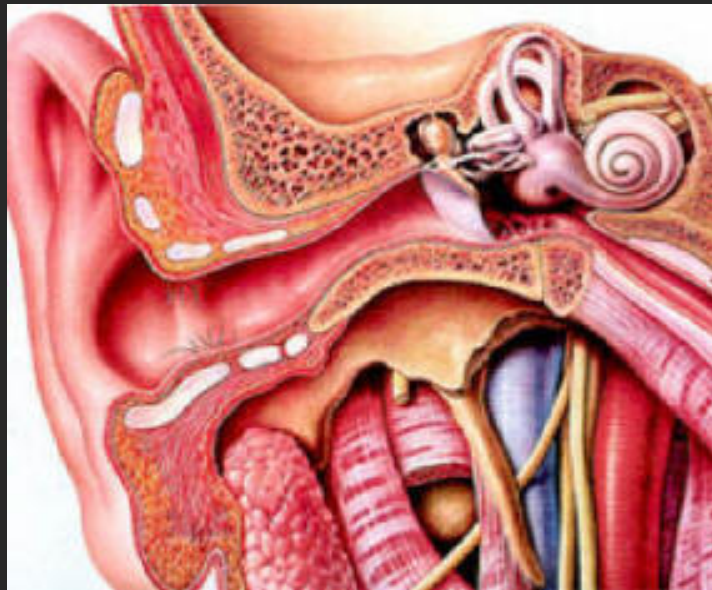
Examples

Examples

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

X	Y
10	8.04
8	6.95
13	7.58
9	8.81
11	8.33
14	9.96
6	7.24
4	4.26
12	10.84
7	4.82
5	5.68

Set B

X	Y
10	9.14
8	8.14
13	8.74
9	8.77
11	9.26
14	8.1
6	6.13
4	3.1
12	9.11
7	7.26
5	4.74

Set C

X	Y
10	7.46
8	6.77
13	12.74
9	7.11
11	7.81
14	8.84
6	6.08
4	5.39
12	8.15
7	6.42
5	5.73

Set D

X	Y
8	6.58
8	5.76
8	7.71
8	8.84
8	8.47
8	7.04
8	5.25
19	12.5
8	5.56
8	7.91
8	6.89

Summary Statistics

Linear Regression

$u_X = 9.0$ $\sigma_X = 3.317$
$u_Y = 7.5$ $\sigma_Y = 2.03$

$Y = 3 + 0.5\ X$
$R^2 = 0.67$

[Anscombe 73]

Set A, Set B, Set C, Set D scatter plots.
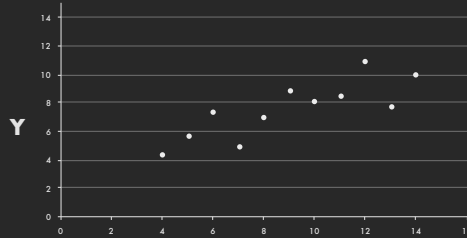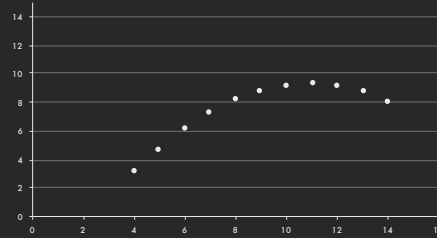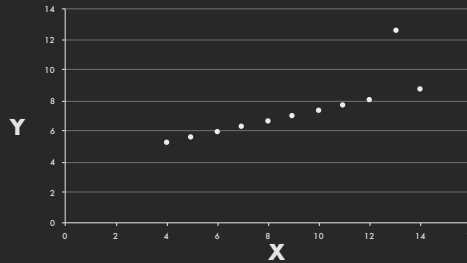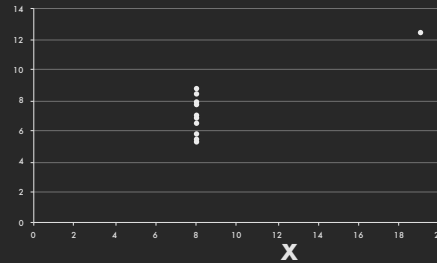
Why do we create visualizations?

Why do we create visualizations?

- Answer questions (or discover them)
- Make decisions
- See data in context
- Expand memory
- Support graphical calculation
- Find patterns
- Present argument
- Tell a story
- Inspire

Record Information

Answer question

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

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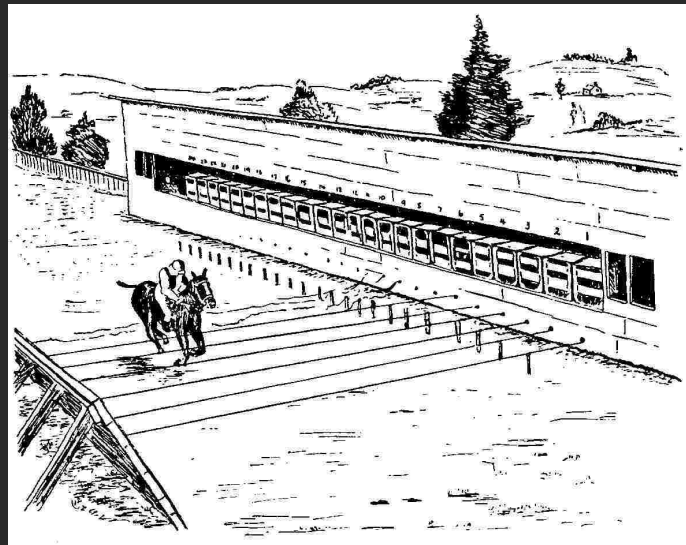
Gallop, Bay Horse "Daisy" [Muybridge 1884-86]

Photographs: Phases of the moon

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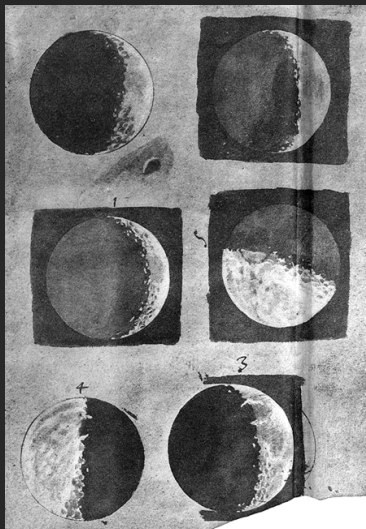
Galileo's drawings of the phases of the moon from 1616
http://galileo.rice.edu/sci/observations/moon.html

Other recording instruments
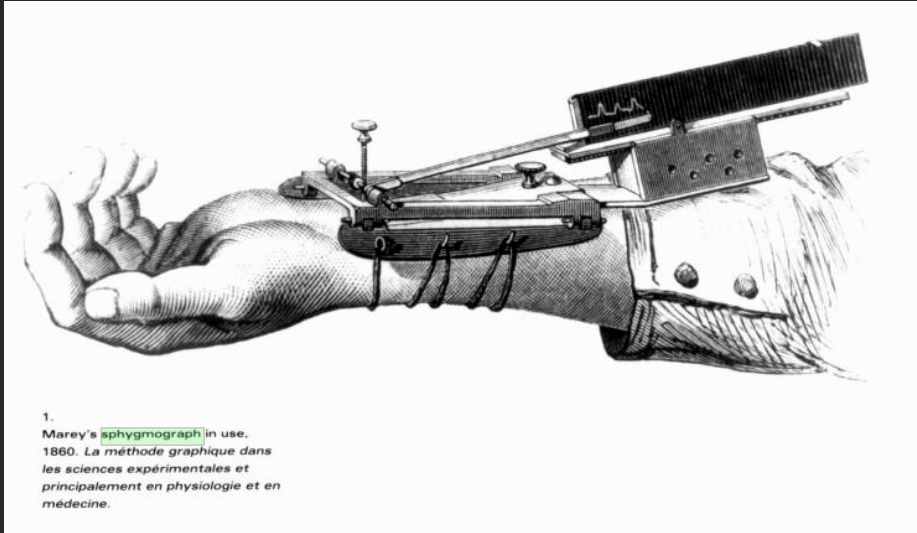
Marey's sphygmograph [from Braun 83]

Support Reasoning

Make a decision: Challenger

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

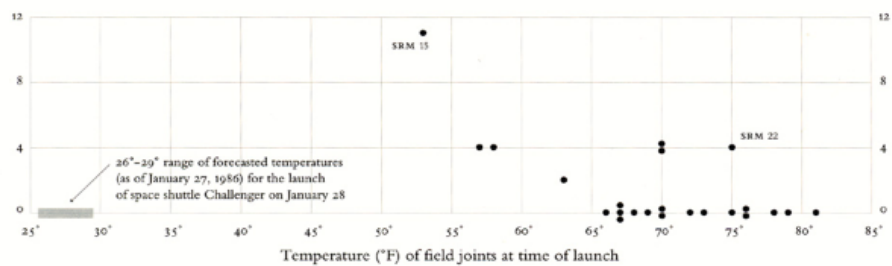# Make a decision: Challenger
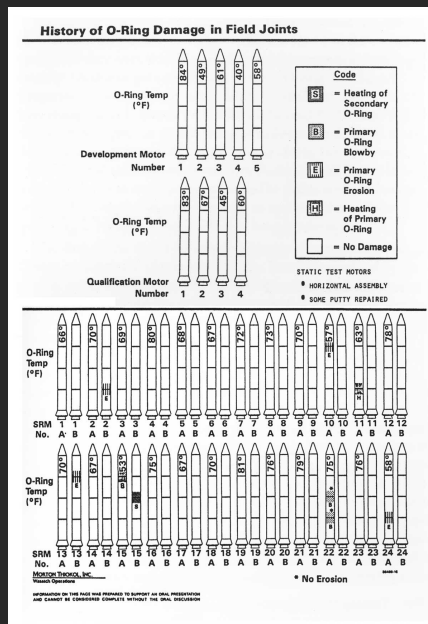
Make a decision: Challenger
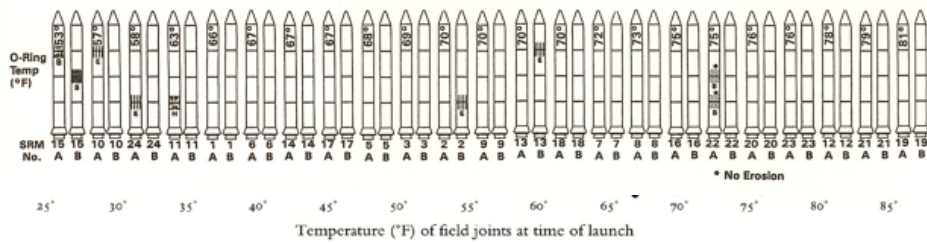
Visualizations drawn by Tufte show how low temperatures damage O-rings [Tufte 97]

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

```
    34
  x 87
  _____
```

Expand memory: Multiplication

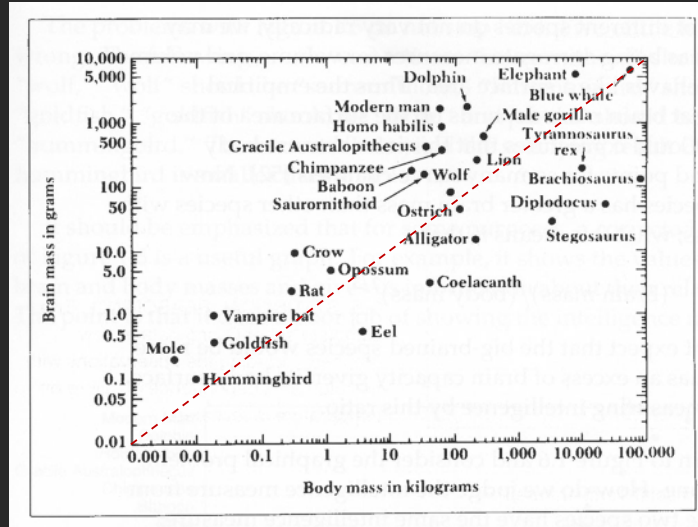$$\begin{array}{r} 34 \\ \times\ 87 \\ \hline 238 \\ 2720 \\ \hline 2958 \end{array}$$

Most powerful brain?

Most powerful brain?
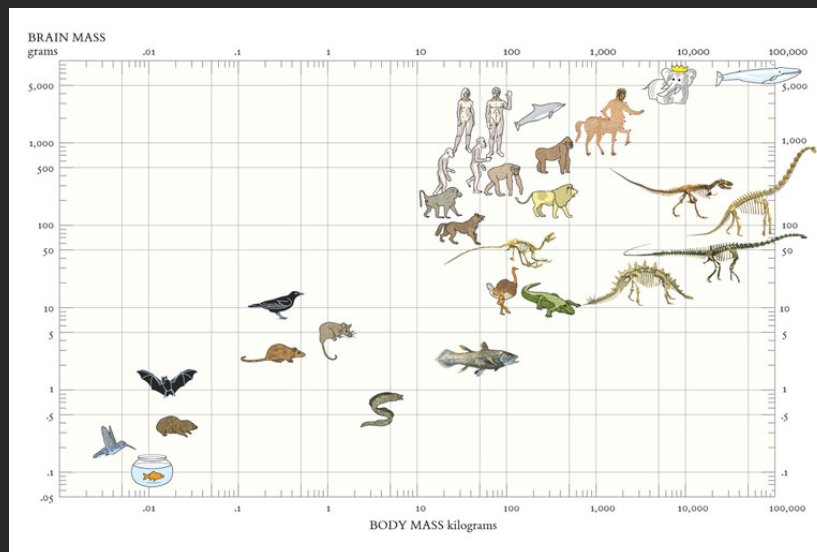
The Dragons of Eden [Carl Sagan]

Tell a story: Most powerful brain?
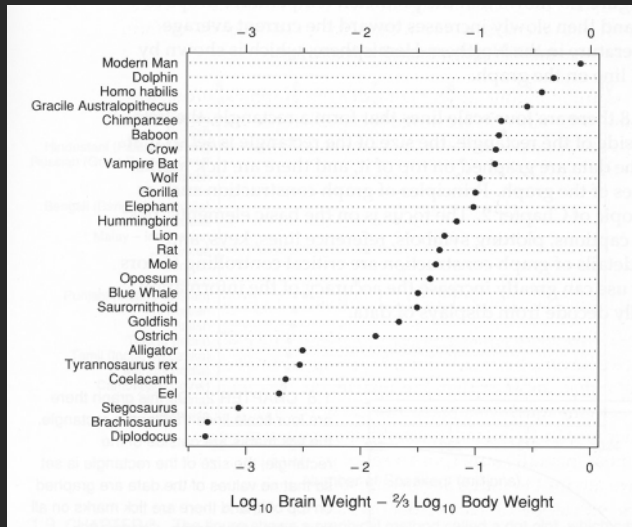
The Elements of Graping Data [Cleveland]

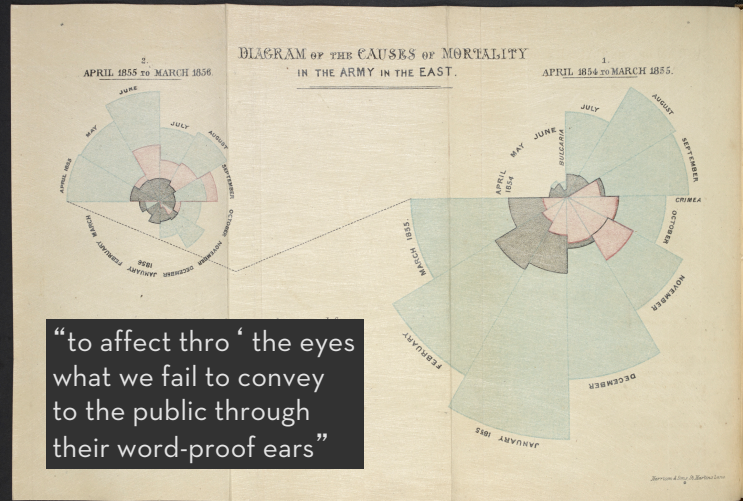# Convey Information to Others
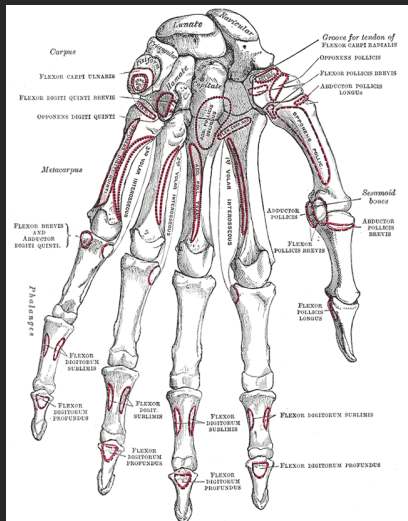
Most powerful brain?

Beautiful Evidence [Tufte]

Present argument

"to affect thro ' the eyes what we fail to convey to the public through their word-proof ears"
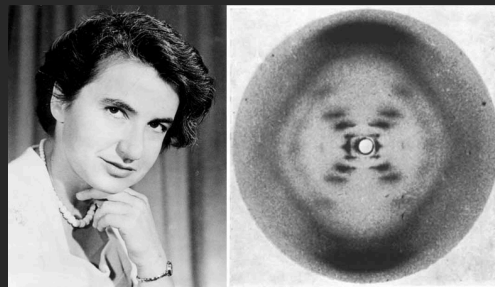
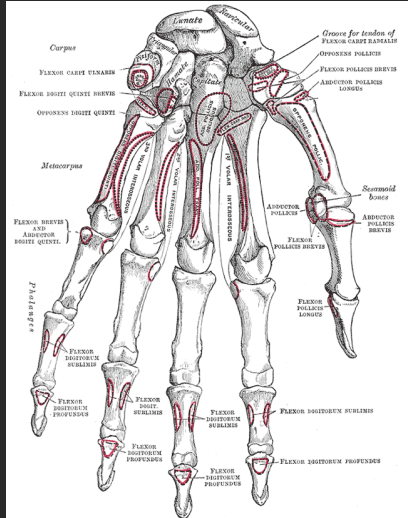Crimean War Deaths [Nightingale 1858]

Inspire

Bones in hand [from 1918 edition]

X-ray crystallography of DNA [Franklin 52]

Inspire

Bones in hand [from 1918 edition]

Double helix model [Watson and Crick 53]

The Purpose of Visualization

Record information
- Photographs, blueprints, …

Support reasoning about information (analyze)
- Process and calculate
- Reason about data
- Expand memory

Convey information to others (present)
- Share and persuade
- Emphasize important aspects of data

Goals of visualization research

1. Understand how visualizations convey information
- What do people perceive/comprehend ?
- How do visualizations correspond with mental models of data?

2. Develop principles and techniques for creating effective visualizations and supporting analysis
- Leverage perception and cognition
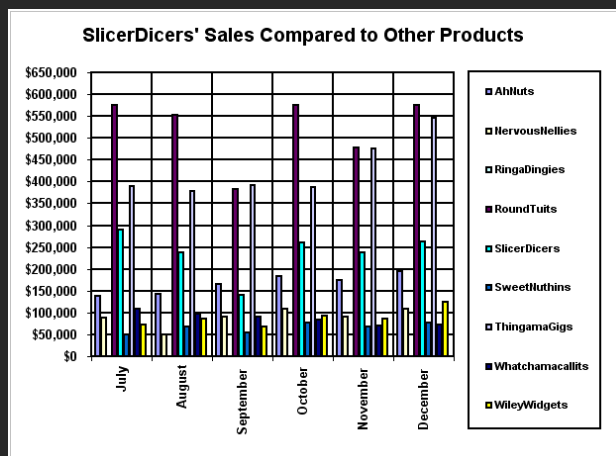- Strengthen connection between visualization and mental models

Course Topics

Data and image models
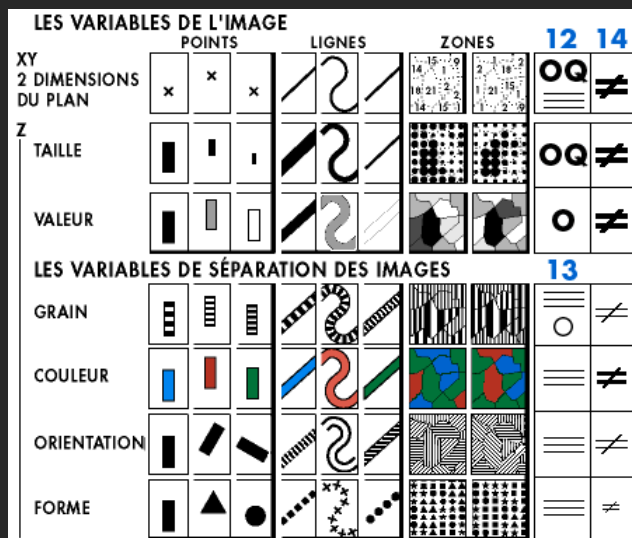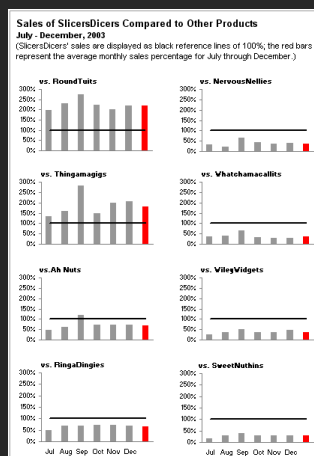

[Bertin, Graphics and Graphic Information Processing 1981]

Visualization Design & ReDesign

Problematic design   Redesign
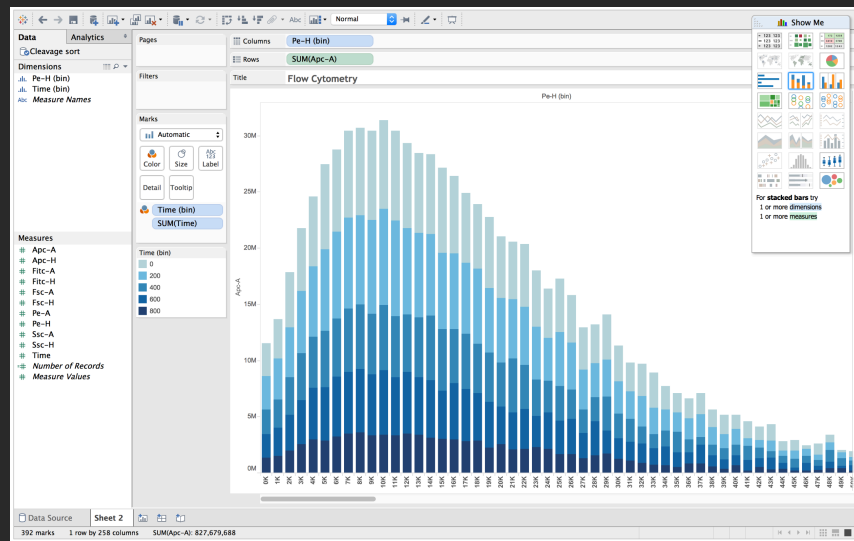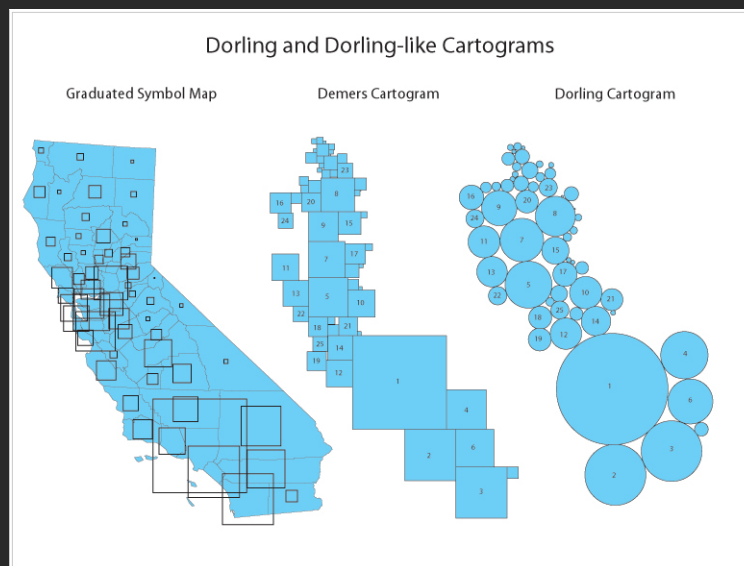
Exploratory Data Analysis

Tableau – based on Polaris [Stolte, Tang, Hanrahan]

Using Space Effectively

http://www.ncgia.ucsb.edu/projects/Cartogram_Central/types.html
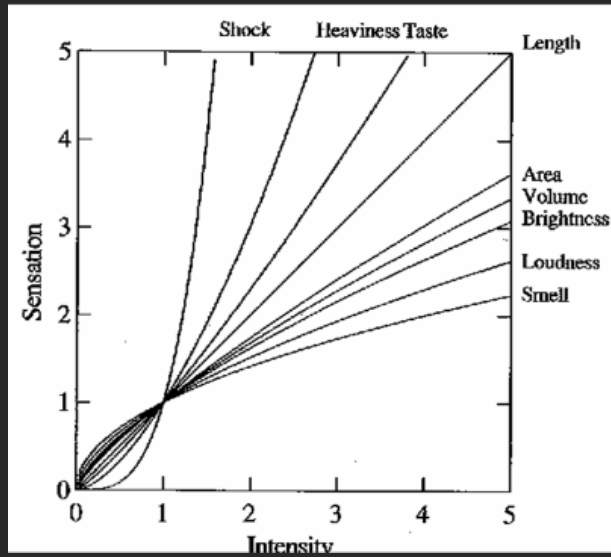
Introduction to D3

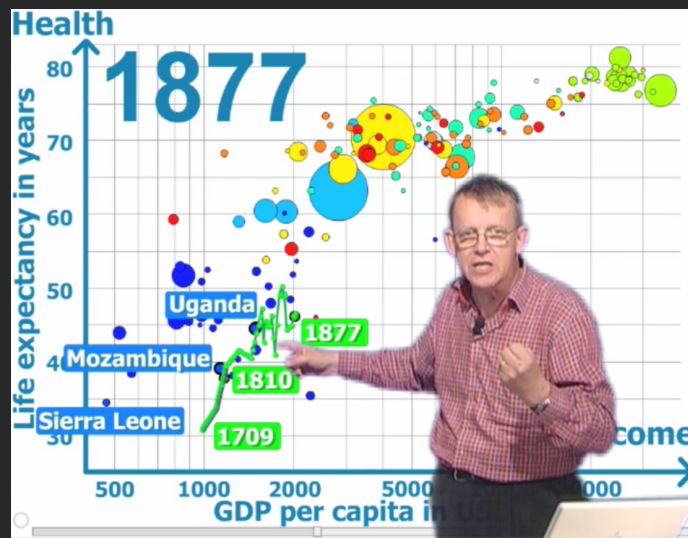D3: Data Driven Documents [Bostock 2011]

Interaction
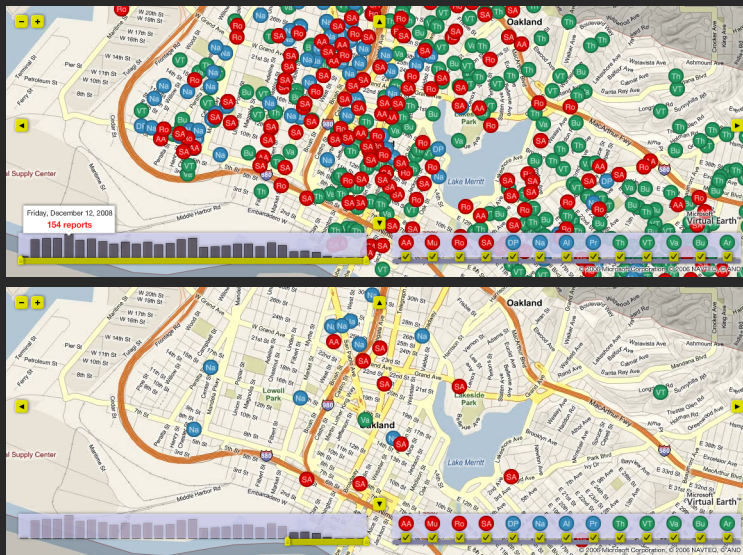
Oakland Crimespotting (crimespotting.org) [Stamen]

Perception

The psychophysics of sensory function [Stevens 61]

Visual Explainers

Gapminder [Rosling]

Color

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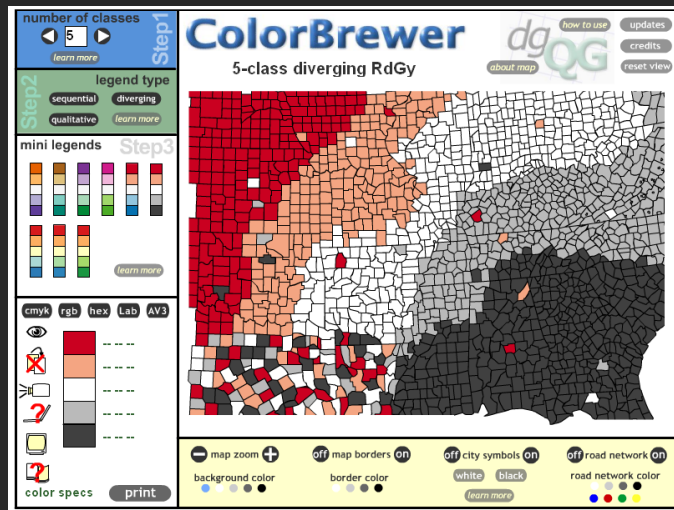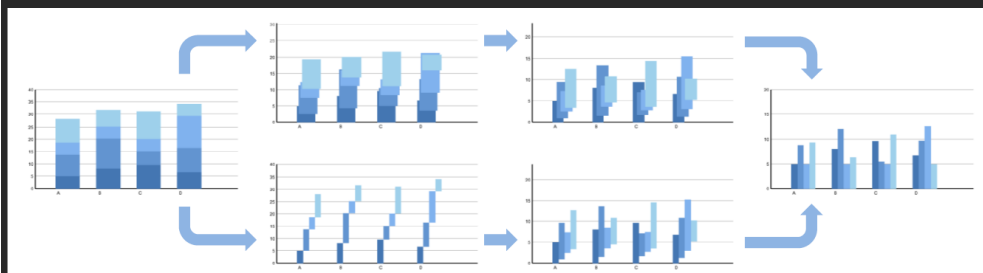
Animation

Animated Transitions [Heer 07]

Networks
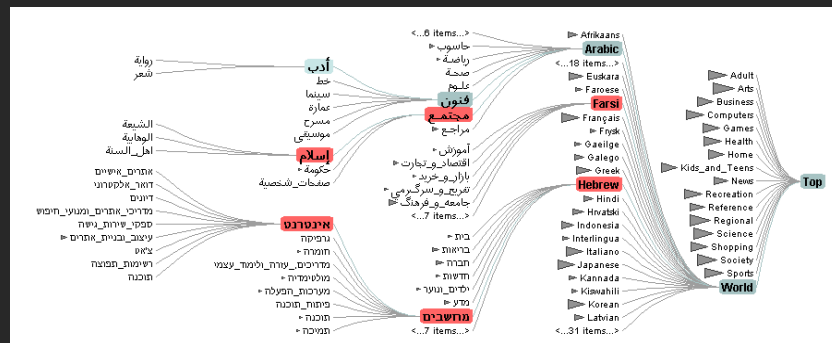
Degree-of-Interest Trees [Heer 2004]

You should expect to

1. *Design, evaluate and critique* **visualizations**

2. *Explore data* **using existing visualization tools**

3. *Implement* **interactive data visualizations**

4. *Develop* **a substantial visualization project**

Course Mechanics

Instructor: Maneesh Agrawala

Original

Restyled Result
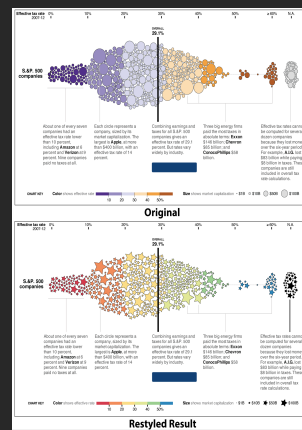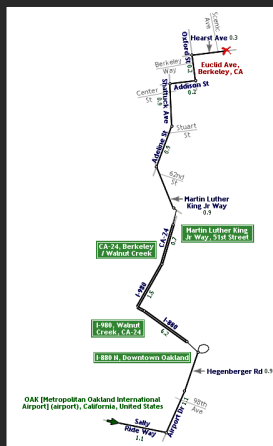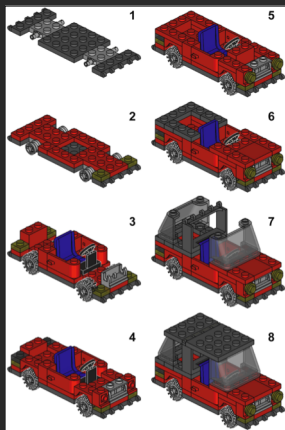
Course Assistant: Juliette Love
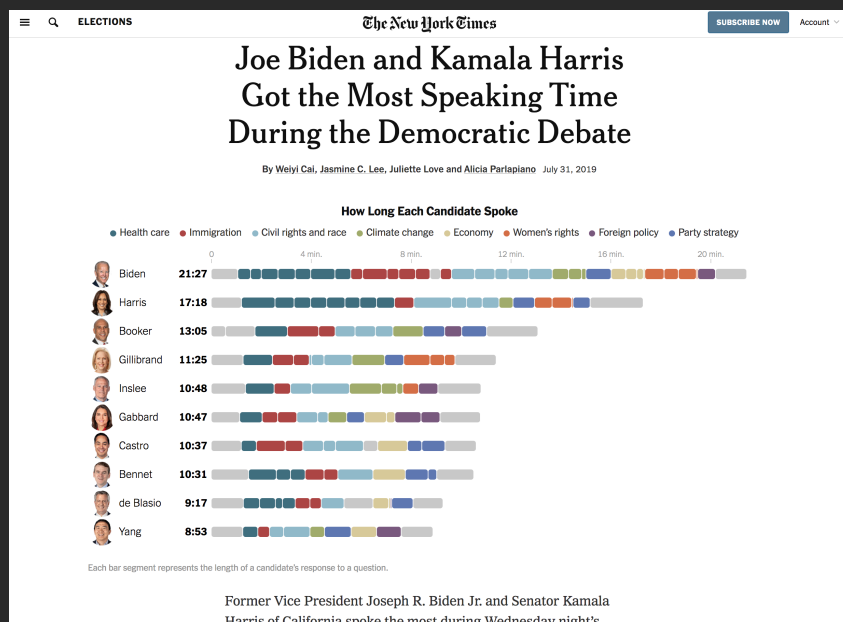
Office Hours

Maneesh: 1:30-2:30pm Thu, Gates 364 & by appt.
Juliette: 7-8:00pm Tue, Lathrop Tech Lounge & by appt.

Outside of OH use Piazza to conect with us
https://piazza.com/stanford/winter2020/cs448b/

Textbooks
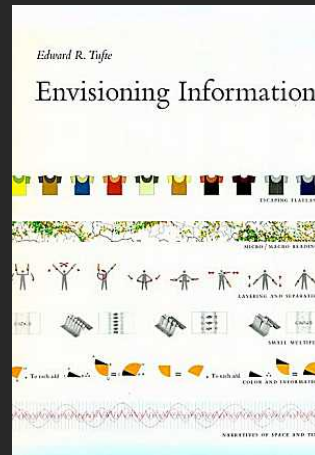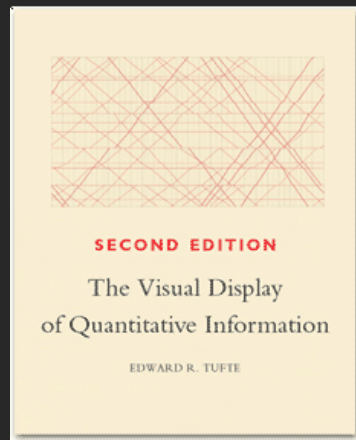
See also: www.edwardtufte.com

Interactive Notebooks
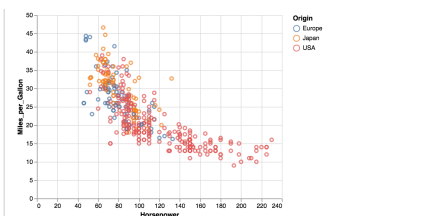
Hands-on engagement with course concepts and modern visualization tools (Vega-Lite / Altair), in both JavaScript (Observable) and Python (Jupyter)!

Optional Book
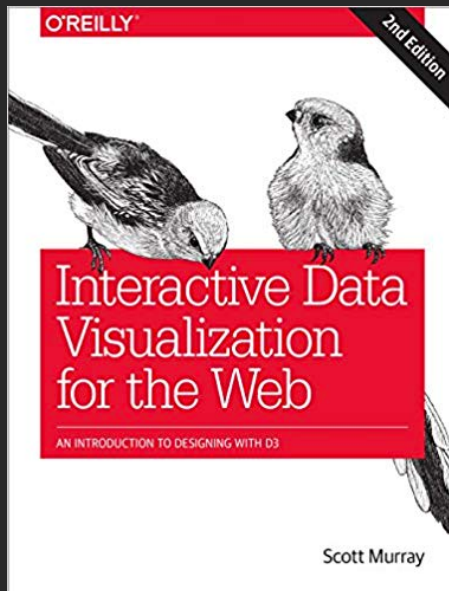
For learning D3!

Book available online
Code/examples on GitHub

We will be using D3 v5
https://d3js.org

Readings

- **From books, notebooks and linked articles**
  Many open to public, some may require SUNetID/Password
- **Material in class will be loosely based on readings**
- **Readings should be read by start of class**

- **Post discussion comments (about reading or lecture) using link on class webpage**
  One comment per week through week 9
  Must post by *noon the day after the lecture*
  You have 1 *pass* for the quarter

Class home page
https://magrawala.github.io/cs448b-wi20/

Lecture/Reading Responses

Good responses typically exhibit one or more

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

Responses should not be summaries

Discussion

Discussion is essential for effective design, evaluation and critique of visualizations

- ■ Attendance for non-SCPD students is mandatory
  (you have 2 passes before it will affect your grade)
- ■ Laptops not allowed
  (unless we specifically ask for them)

Assignments

Class participation (10%)

Assignment 1: Visualization Design **(10%)** due 1/13

Assignment 2: Exploratory Data Analysis **(15%)** due 1/27
> Learn to use Tableau

Assignment 3: Interactive Prototype **(25%)** due 2/10
> Should be familiar with Javascript (start now if you are not)
> Will cover basics of D3 in class

Final Project (40%) proposal due 2/19, milestone 3/9, final 3/16

Final project

Visualization research project on topic of choice

Initial prototype and peer evaluation

Design reviews and final presentation
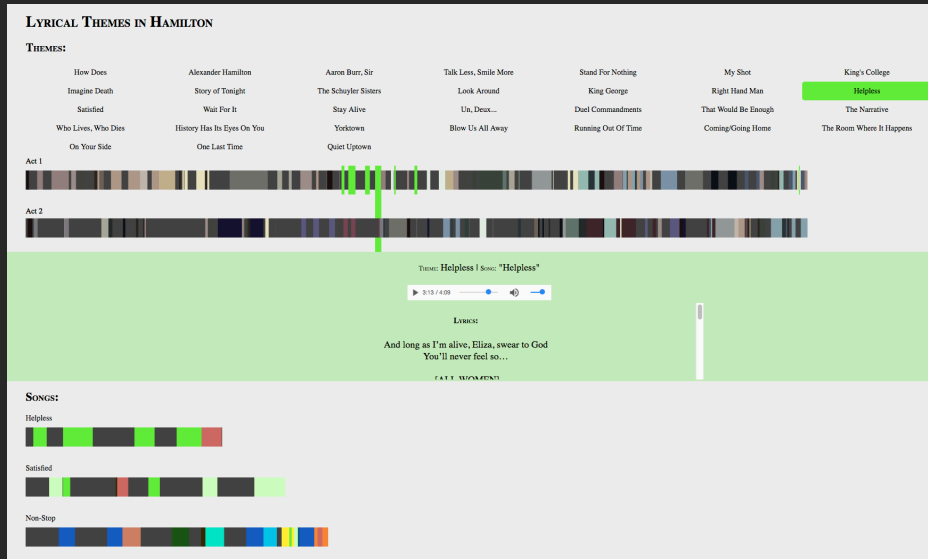
Submit and publish online (if feasible)

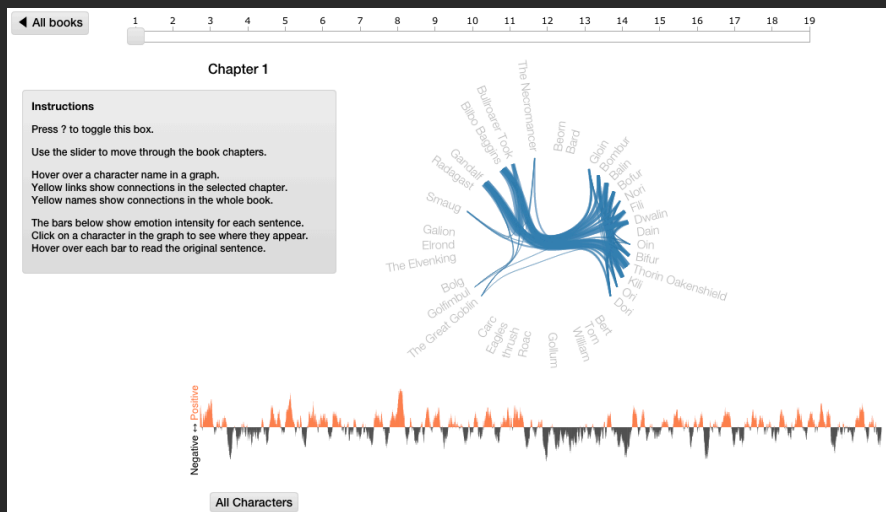Projects from previous classes have been:

- Published as research papers
- Gone viral on blogs
- Released as open source projects
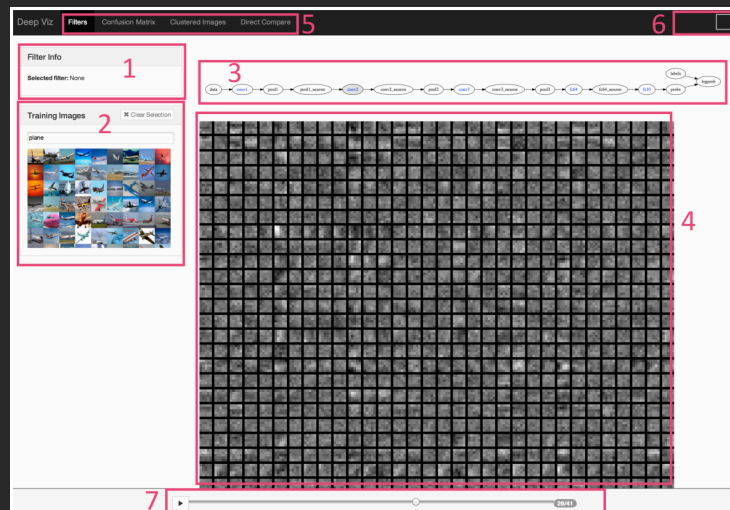
Structure of Musicals

Lyrical themes in Hamilton [Townley-Smith, Sterman, Cook 2016]

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

Design a static visualization for a data set.

You must choose the message you want to convey. What question(s) do you want to answer? What insight do you want to communicate?

Data: Coterminal Computer Science Master's Degrees at Stanford

Stanford Institutional Research and Decision Support collects a variety of data about the educational programs at Stanford. We have extracted and prepared a small data set about the the number of students who have pursued coterminal Master's degrees in Computer Science between 2014 and 2019. Our data set contains the following information:

Number of records: 112
Variable Names:
   Completion Year: Year in which the coterminal degrees were completed.
   Coterm Master's Plan: Name of Master's degree student completed.
   Coterm Undergraduate Plan: Name of Undergraduate degree student completed.
   Number of Completions: Num. of students that completed the corresponding (Master's, Undergraduate) coterminal degree plan.

The extracted dataset is available in csv format: StanfordCSCotermPlans_2014-1029.csv

Due by noon on Mon Jan 13

Assignment 1: Visualization Design

Pick a guiding question, use it to title your visualization

Design a static visualization for that question

You are free to use any tools (including pen & paper)

Deliverables (upload via Canvas; see A1 page)

PDF of your visualization with a short description including design rationale (≤ 4 paragraphs)

Due by noon on Mon Jan 13