CSE 557A | Jan 17, 2017

INFORMATION VISUALIZATION

Alvitta Ottley

Washington University in St. Louis

Slide Acknowledgements:
Mariah Meyer, University of Utah
Remco Chang, Tufts University

GETTING TO KNOW YOU…

Show of hand:

- Know Java?
- Taken Graphics?
- Taken HCI?
- Proficient with GUI frameworks or toolkit
  - Eg. Processing, D3, Javascript/ HTML 5, openGL, Java Swing, etc.
- Design Background?
  - Have taken courses in art or design

YOUR EXPECTIONS

What does **Visualization** mean to **you**?

What do you hope to get from the course?

MY EXPECTATIONS

- Try
- Be creative
- Participate
- Integrity

A visualization is a visual representation of abstract **data** to aid **human cognition**

- Must be based on data
- The results must be readable, recognizable and useful

How much data are there?

2010: 1.2 zettabytes
2013: 4.4 zettabytes
2020: ~40 zettabytes

Gantz et al 2011

2010: 1.2 zettabytes
2013: 4.4 zettabytes
2020: ~40 zettabytes

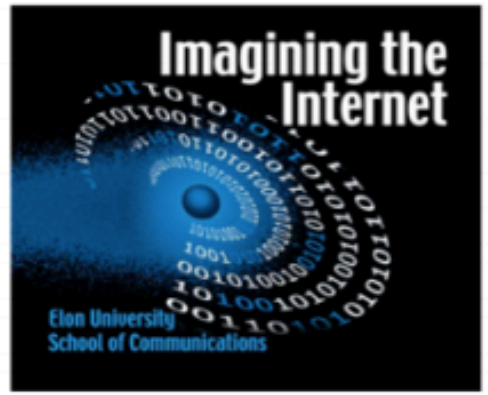
Zettabyte ~= 1,000,000,000,000,000,000,000 or 10^{21}
200x **all** words ever spoken by humans

Gantz et al 2011

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... [Hal Varian, Google's Chief Economist]

PewResearchCenter

Pew Internet
Pew Internet & American Life Project



Big Data: Experts say new forms of information analysis will help people be more nimble and adaptive, but worry over humans' capacity to understand and use these new tools well

Tech experts believe the vast quantities of data that humans and machines will be creating by the year 2020 could enhance productivity, improve organizational transparency, and expand the frontier of the "knowable future." But they worry about "humanity's dashboard" being in government and corporate hands and they are anxious about people's ability to analyze it wisely

Janna Quitney Anderson, Elon University
Lee Rainie, Pew Research Center's Internet & American Life Project
July 20, 2012

Pew Research Center's Internet & American Life Project
An initiative of the Pew Research Center
1615 L St., NW – Suite 700

WHY does Visualization work?

- Cognition is limited

The "Door" Study

Daniel Simons

Subscribe 10,299

1,662,215 views

Add to Share More

7,501 80

WHY does Visualization work?

- Cognition is limited
- Memory is limited

calculation exercise…

calculation exercise…

$$\begin{array}{r} 34 \\ \times\ 28 \\ \hline \end{array}$$

calculation exercise…

calculation exercise…

$$\begin{array}{r} 79 \\ \times\ 16 \\ \hline \end{array}$$

HOW does Visualization work?

- Uses perception to point out interesting things.

How many R's are there?

GLNSAGGKLDSANGNASDGN
KLANSDGLKNASDGNDFVMD
GJERKJVERJVJKENJLVNEKVJEN
VJEAJVJNDJVNAAJBVRKLVLJKD

How many R's are there?

GLNSAGGKLDSANGNASDGN
KLANSDGLKNASDGNDFVMD
GJERKJVERJVJKENJLVNEKVJEN
VJEAJVJNDJVNAAJBVRKLVLJKD

What is the complexity (in BigO Notation)
to find the single B?

GLNSAGGKLDSANGNASDGN
KLANSDGLKNASDGNDFVMD
GJERKJVERJVJKENJLVNEKVJEN
VJEAJVJNDJVNAAJBVRKLVLJKD

What is the complexity (in BigO Notation)
for a human to find the single B?

GLNSAGGKLDSANGNASDGN
KLANSDGLKNASDGNDFVMD
GJERKJVERJVJKENJLVNEKVJEN
VJEAJVJNDJVNAAJBVRKLVLJKD

Consider the following dataset:

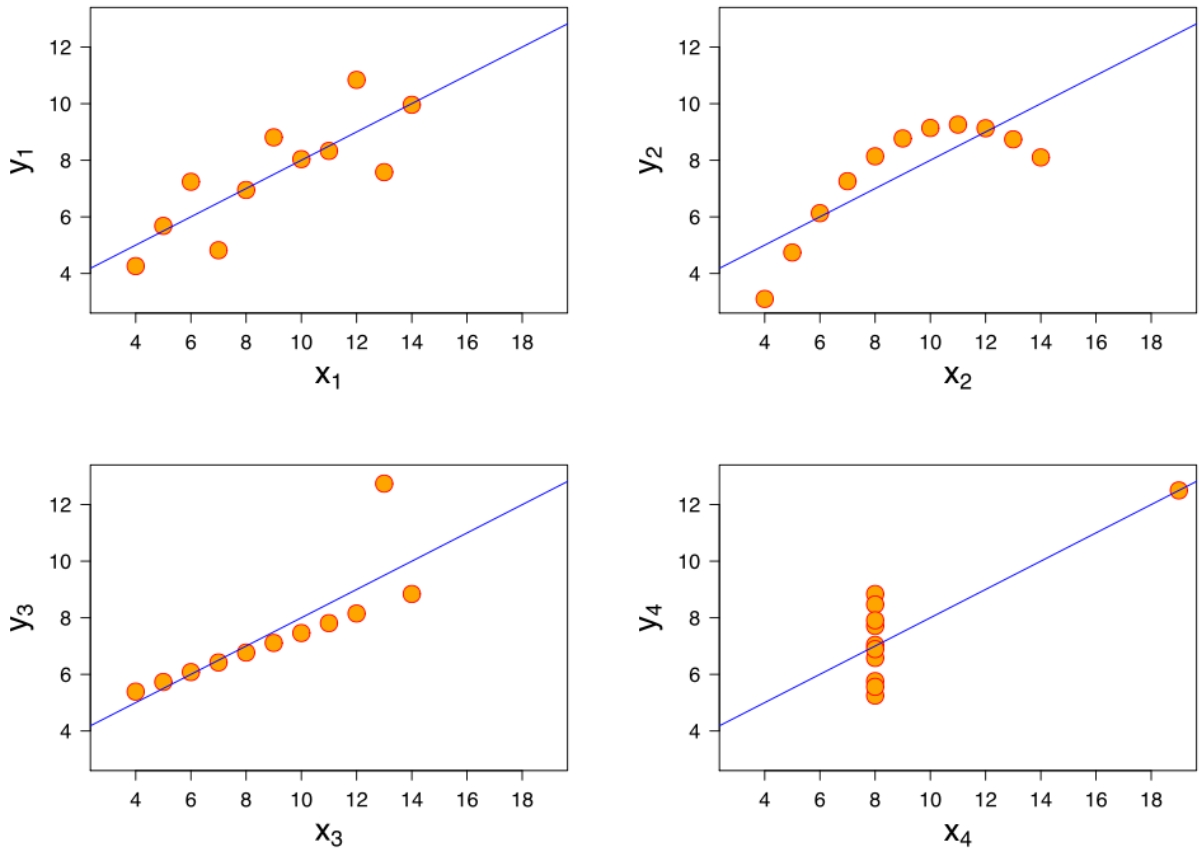
I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	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

Consider the following dataset:

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

Property	Value
Mean of *x* in each case	9 (exact)
Sample variance of *x* in each case	11 (exact)
Mean of *y* in each case	7.50 (to 2 decimal places)
Sample variance of *y* in each case	4.122 or 4.127 (to 3 decimal places)
Correlation between *x* and *y* in each case	0.816 (to 3 decimal places)
Linear regression line in each case	$y = 3.00 + 0.500x$ (to 2 and 3 decimal places, respectively)

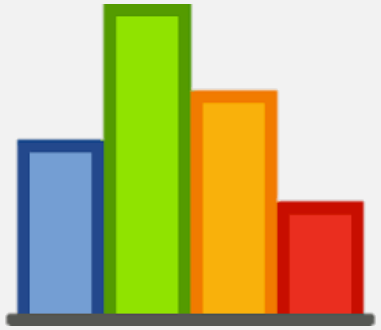
Anscombe's Quartet



WHY do we create visualizations?

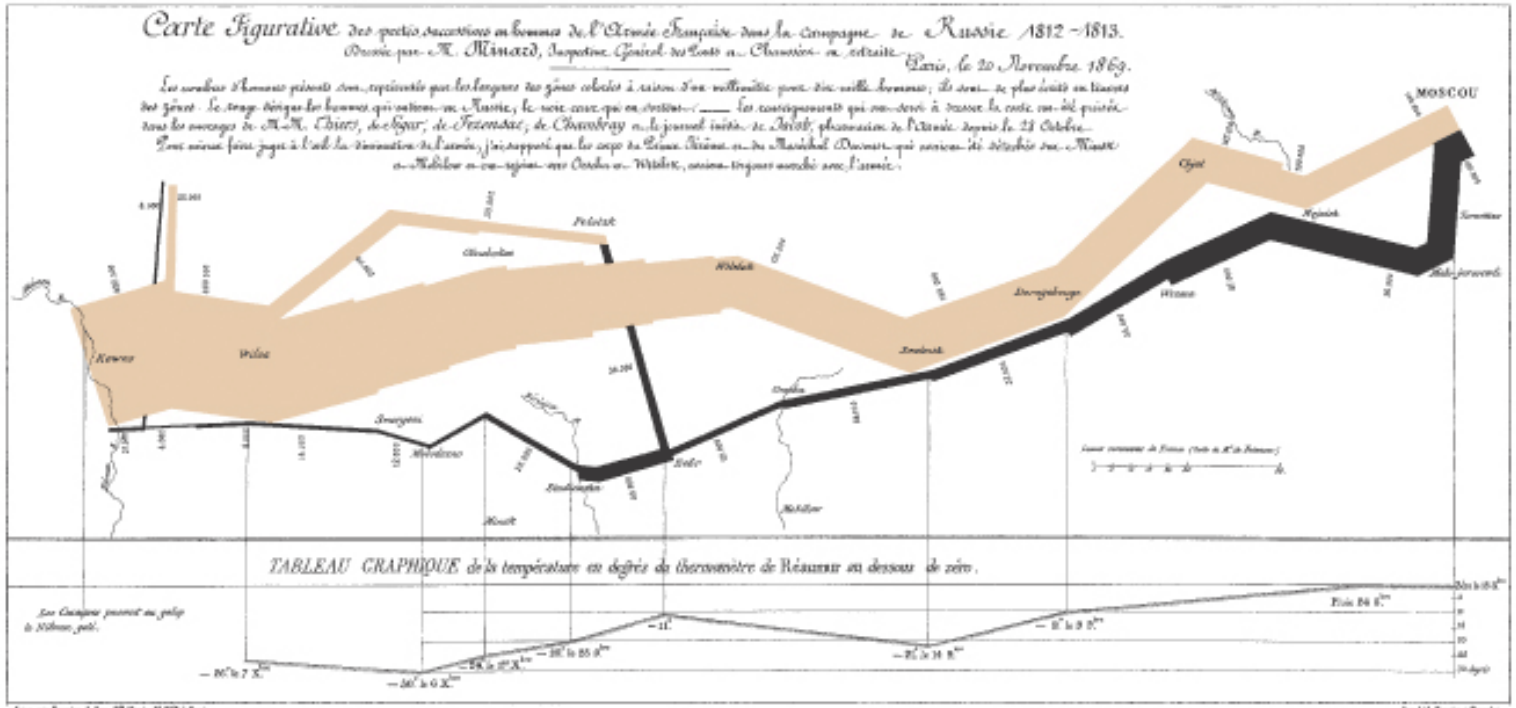
- answer questions
- generate hypotheses
- make decisions
- see data in context
- expand memory
- support computational analysis
- find patterns
- tell a story
- inspire

Examples of visualizations

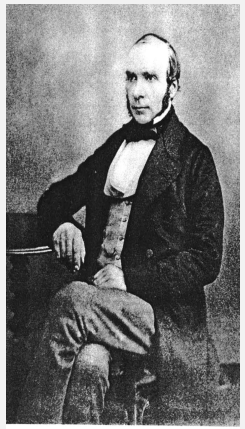


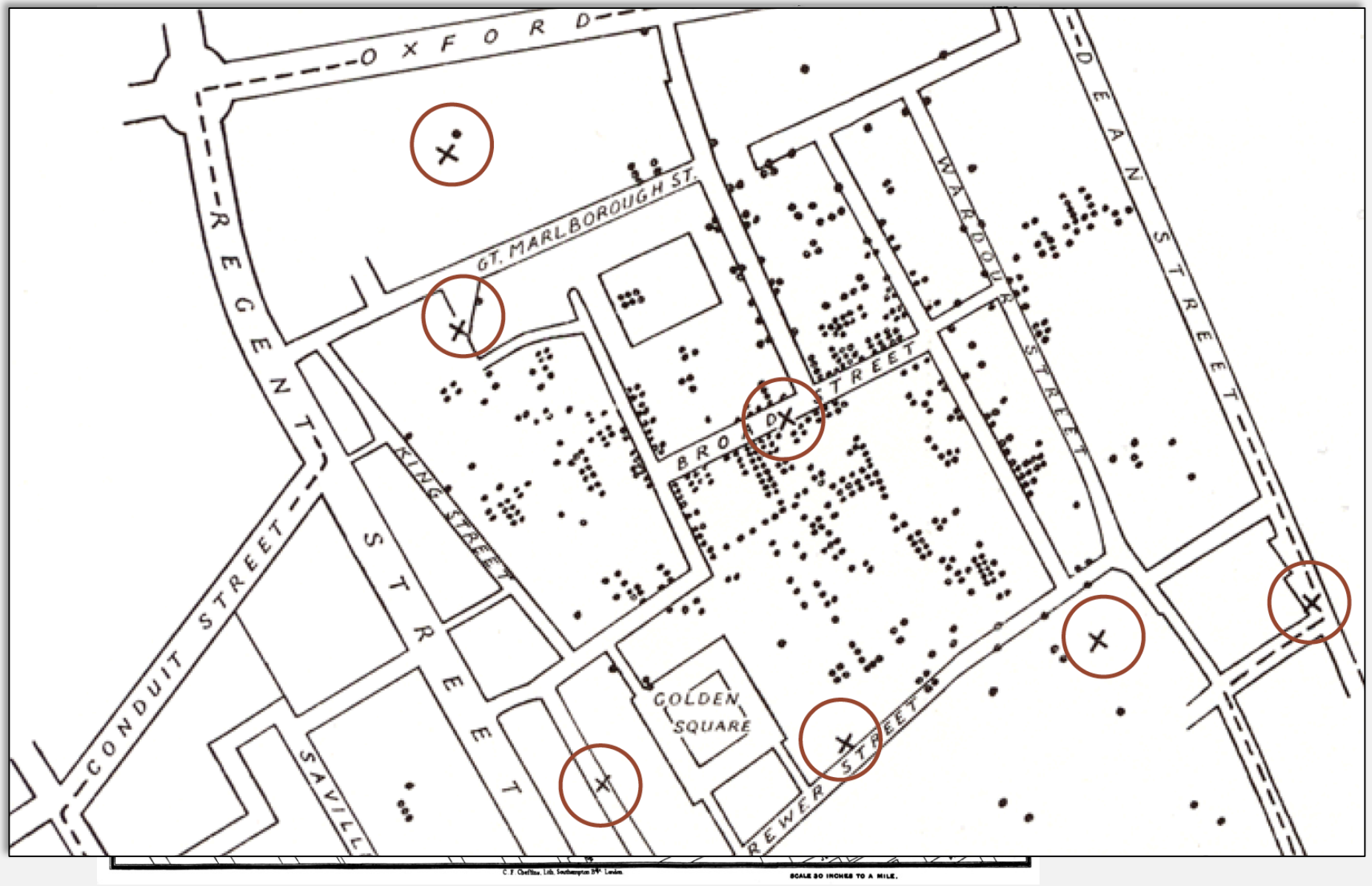
Visualization for Storytelling

Napoleon's March to Moscow (War of 1812)

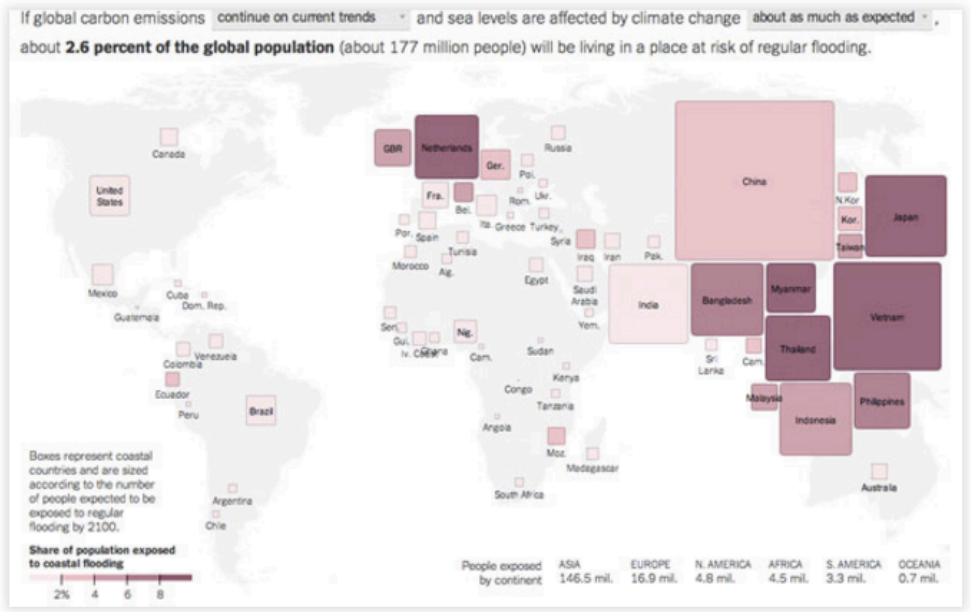


Visualization for spotting trends





Visualization for information dissemination



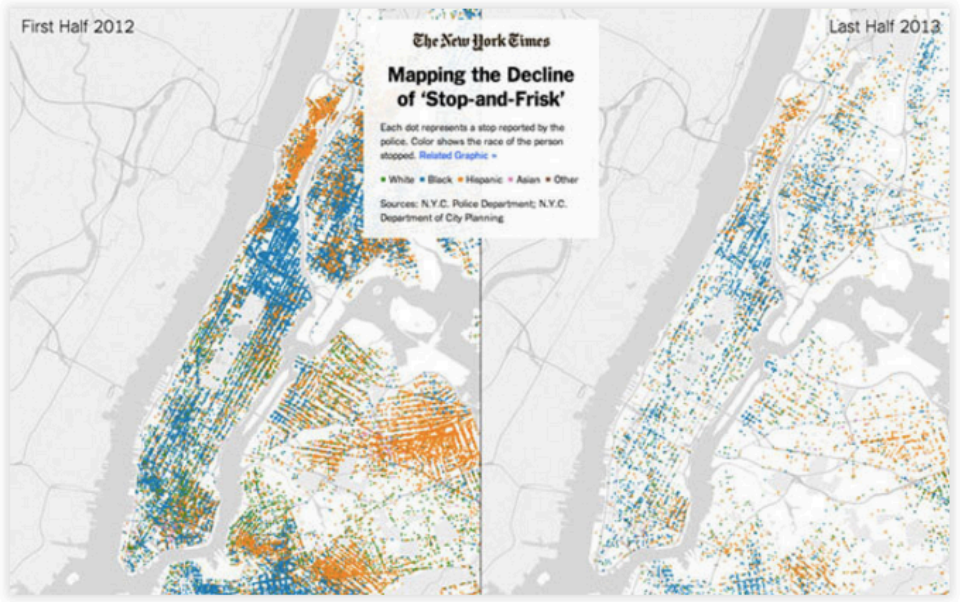
Flooding Risk From Climate Change, Country by Country

SEPT. 23

A new analysis of sea levels and flood risk around the world offers more evidence that the brunt of climate change will not be borne equally.

FACEBOOK TWITTER

Visualization for information dissemination



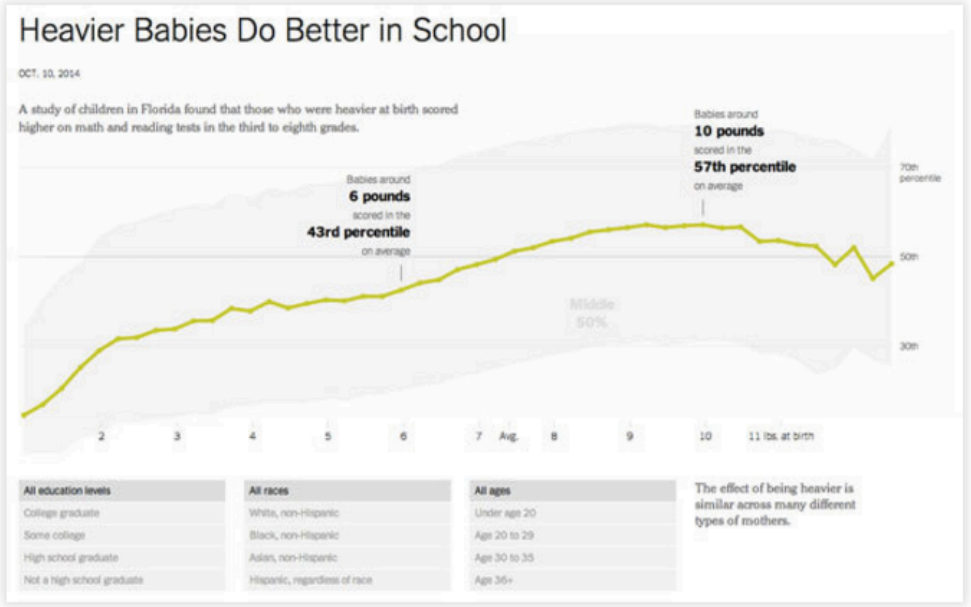
Mapping the Decline of 'Stop-and-Frisk'

SEPT. 19

Mapping and charting the decline of a controversial police practice.

FACEBOOK TWITTER

Visualization for information dissemination



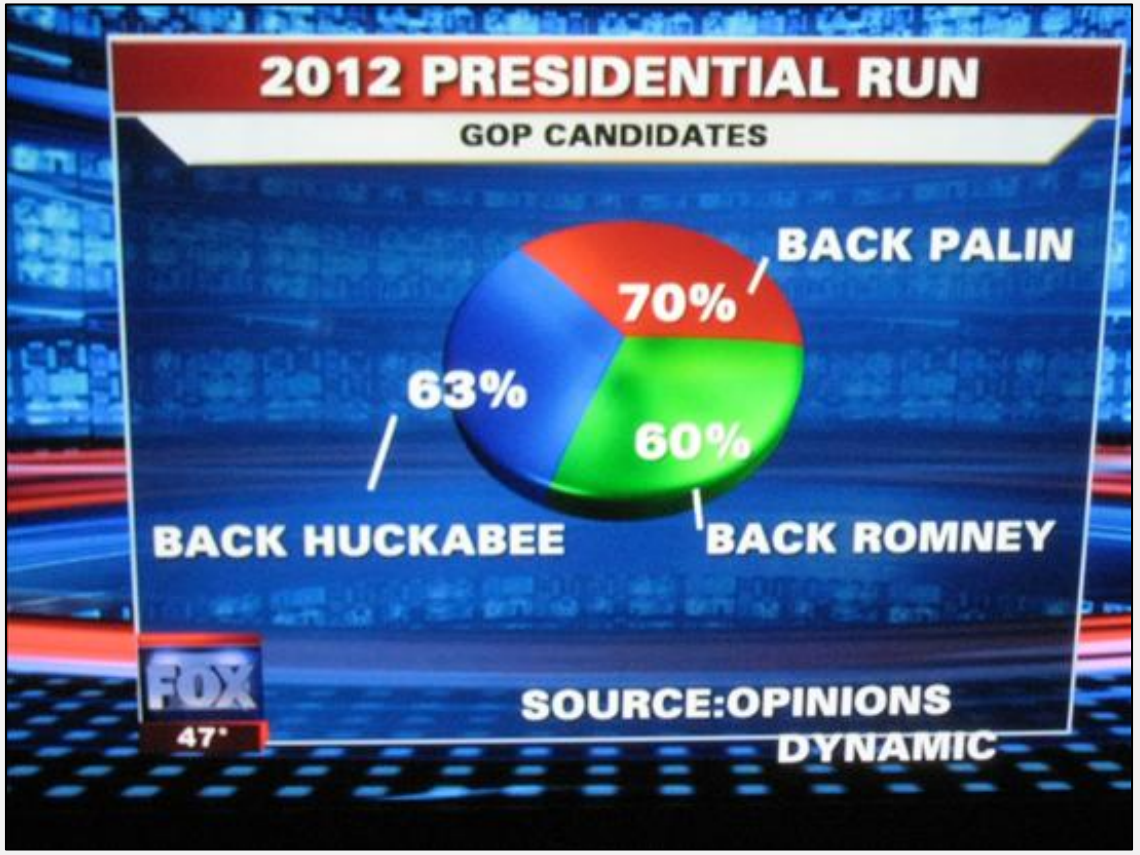
Heavier Babies Do Better in School

OCT. 10

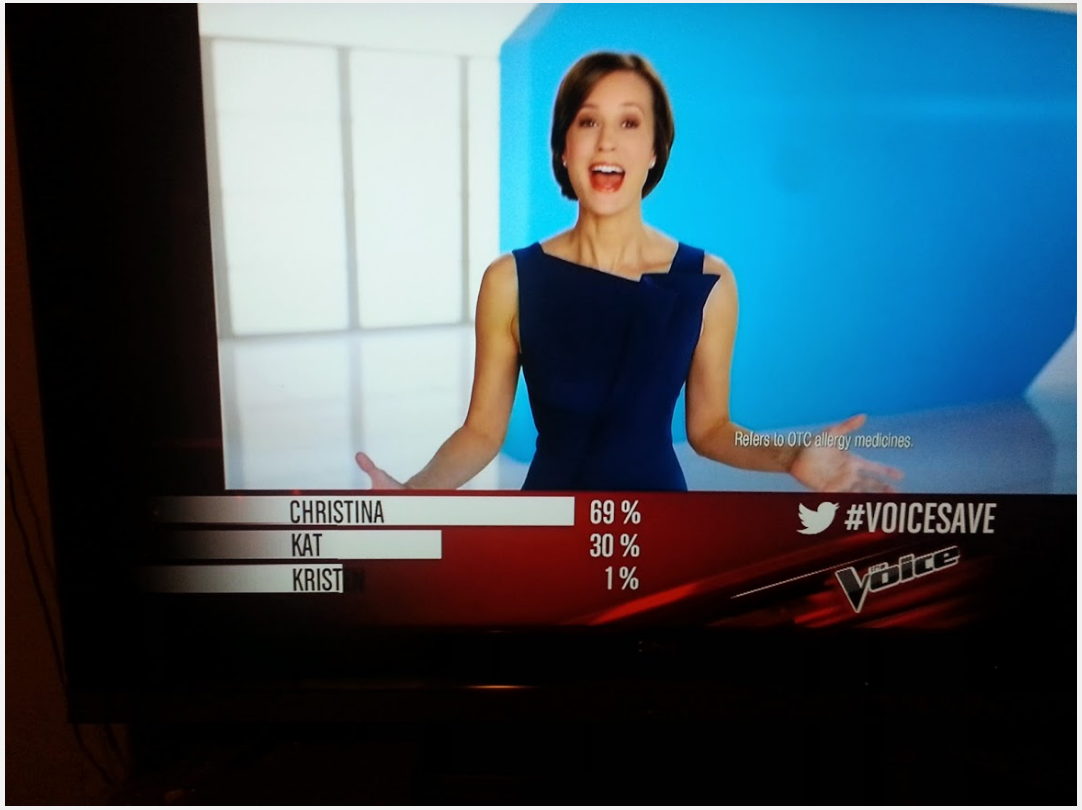
Babies who are allowed to linger in the womb are often healthier and do better in school.

FACEBOOK TWITTER

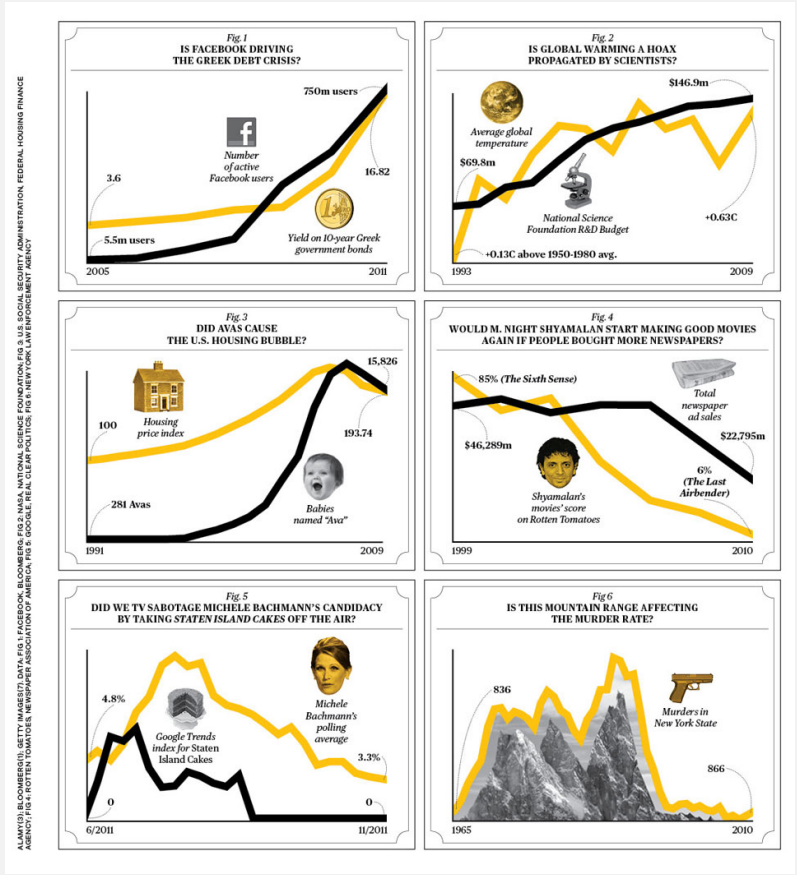
Examples of “bad” visualizations



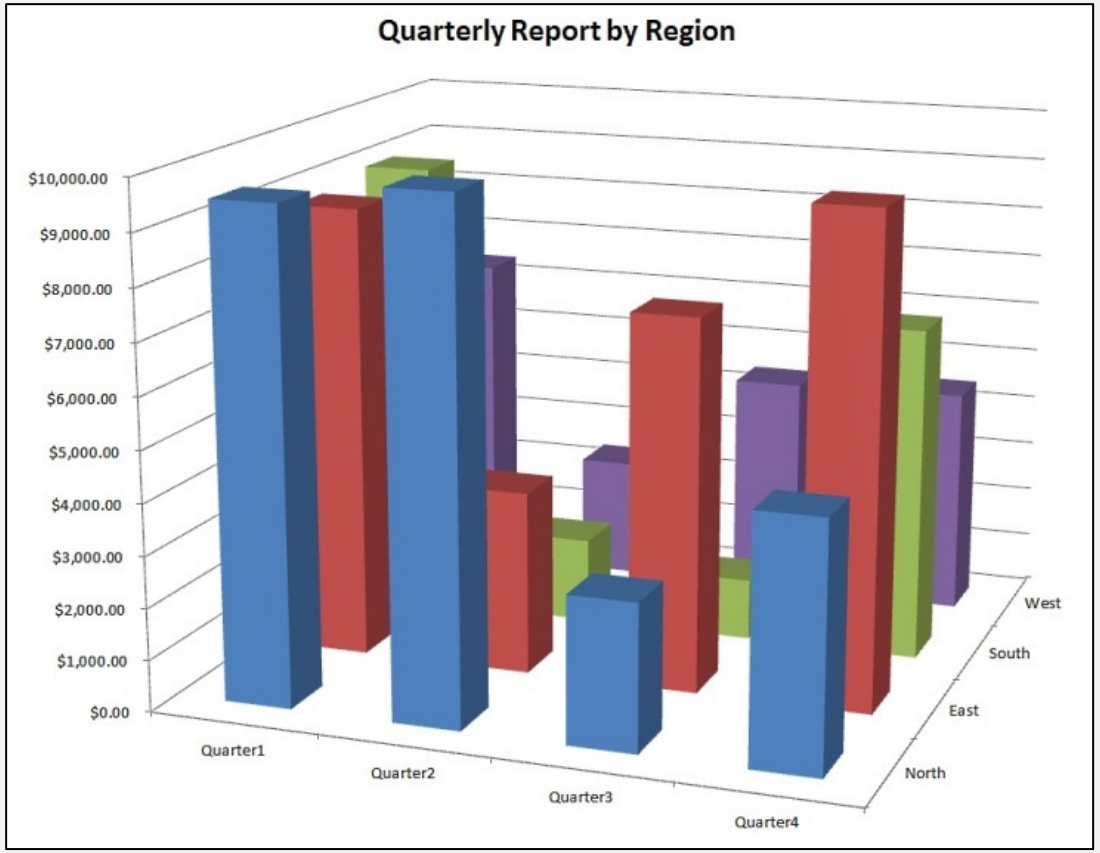
Examples of “bad” visualizations



Examples of “bad” visualizations



Visualizations that probably should not exist



Questions?

CSE 557A: INFORMATION VISUALIZATION

Course Number	CSE 557A
Semester	Spring 2017
Hours	TR 11:30AM-1:00PM
Location	Rudolph Hall 203
Instructor	Alvitta Ottley
Email	alvitta at wustl.edu
Office	Jolley 410
Office Hours	TR 1:00PM-2:00PM
TA	Josh Landman
Email	landman at wustl.edu
Office	Psychology 402F
Office Hours	W 1:00PM-2:30PM, S 5PM-6PM

Schedule

Syllabus

Assignments

Final Project

Resources

Course Description

In this course, we study the principles for transforming abstract data into effective information visualizations. We learn about the state-of-the-art in visualization research and development, and we gain hands-on experience with designing and developing information visualizations. We also learn how to critique existing visualizations and how to evaluate the systems we build. Weekly readings include current research papers from the Information Visualization community.

Schedule

WEEK	DATE	TOPIC	DATE	TOPIC	PAPER	ASSIGNMENT
1	1/17	Introduction	1/19	Discussion: Chart Junk Lab 1: Intro to Processing	Bateman et al. 2010	Design Critique | due 1 in week
2	1/24	Design	1/26	Discussion: Visualization Purpose Lab 2: Click Me!	Borkin et al. 2013	Bar and Line | due in 2 weeks
3	1/31	Data Types	2/02	Discussion: Bars vs. Pies Lab 3: Bars	Cleveland and McGill	
4	2/07	Visual Encodings	2/09	Discussion: Data Mapping Lab 4: Animated Transition	Polaris	Parallel Coordinates | due in 2 weeks
5	2/14	Tablular Data	2/16	Discussion: TDB Lab 5: Intersection Detection		
6	2/21	Perception	2/23	Discussion: Perception Lab 6: Web Basics	Visual Thinking for Design (Chapters 1 & 2) Perception in Visualization	ThemeRiver | due in 2 weeks
7	2/28	Trees	3/02	Discussion: Tree Visualizations Lab 7: d3	ThemeRiver Web Tutorial	
8	3/07	Graphs	3/09	Discussion: Graph Simplification Lab 8: Electoral Map	Schniederman (Eyes) Scatter Plot Example Basic Template	TreeMap in D3 | due in 2 weeks
9	3/15	**SPRING BREAK**	3/17	**SPRING BREAK**		
10	3/21	Storytelling	3/23	Discussion: Storytelling Lab 8: Final Project	Yi et al. An Empire Built on Sand	Final Project Proposal | due in 1 week

General Information

Five Assignments
One Final Project
Weekly Paper Discussions & Labs

Assignment 1: Design Critique
Due: 01-24-2017, 11:59pm (midnight)

In this assignment, you will be looking for visualizations "in the wild" (in books, newspapers, magazines, on the internet, etc.). Specifically, you will be looking for two visualizations – one that you like, and one that you dislike. For the visualization that you like, you need to express why you like the visualization (what is it that makes the visualization good). For the visualization that you dislike, you need to provide a critique, as well as design a better alternate visualization.

Due next Tuesday

Basic Requirements for this Assignment:

1. Find two visualizations in the wild
   a. One that you like
   b. One that you dislike
2. For the visualization that you like, provide a description of what makes the visualization good.
3. For the visualization that you dislike, explain why you dislike it.
4. In addition, for the visualization that you dislike, design an alternate visualization that is better than the original.
5. Explain your design and what problem(s) your design addresses.

To do for next class:

- Download and install Processing
- Bring laptops
- Read paper

Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts

Scott Bateman, Regan L. Mandryk, Carl Gutwin,
Aaron Genest, David McDine, Christopher Brooks
Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
scott.bateman@usask.ca, regan@cs.usask.ca, gutwin@cs.usask.ca,
aaron.genest@usask.ca, dam085@mail.usask.ca, cab938@mail.usask.ca

ABSTRACT
Guidelines for designing information charts often state that the presentation should reduce 'chart junk' – visual embellishments that are not essential to understanding the data. In contrast, some popular chart designers wrap the presented data in detailed and elaborate imagery, raising the questions of whether this imagery is really as detrimental to understanding as has been proposed, and whether the visual embellishment may have other benefits. To investigate these issues, we conducted an experiment that compared embellished charts with plain ones, and measured both interpretation accuracy and long-term recall. We found that people's accuracy in describing the embellished charts was no worse than for plain charts, and that their recall after a two-to-three-week gap was significantly better. Although we are cautious about recommending that all charts be produced in this style, our results question some of the premises of the minimalist approach to chart design.

Author Keywords
Charts, information visualization, imagery, memorability.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

data-ink – or the ink in the chart used to represent data.

Despite these minimalist guidelines, many designers include a wide variety of visual embellishments in their charts, from small decorations to large images and visual backgrounds. One well-known proponent of visual embellishment in charts is the graphic artist Nigel Holmes, whose work regularly incorporates strong visual imagery into the fabric of the chart [7] (e.g., Figure 1).

MONSTROUS COSTS
Total House and Senate campaign expenditures, in millions