CSE 512 - Data Visualization

The Value of Visualization

Jeffrey Heer  University of Washington
How much data (bytes) did we produce in 2010?
2010: 1,200 exabytes

Gantz et al, 2008, 2010
2010: 1,200 exabytes
10x increase over 5 years

Gantz et al, 2008, 2010
Physical Sensors

Image courtesy cabspotting.org
Records of Human Activity
Abortion

“Abortion,” in its most commonly used sense, refers to the deliberate early termination of pregnancy, resulting in the death of the fetus. Medically, the term also refers to the early termination of a pregnancy by natural (“spontaneous abortion” or “miscarriage”) or medical means. In most countries, abortion is legal in some form for a variety of reasons, including protecting the mother’s life or health, preventing serious health complications, or when the pregnancy is the result of rape or incest.

Methods

Depending on the stage of pregnancy and other factors, abortion can be performed by a number of different methods. In the earliest terminations, before nine weeks of pregnancy, a chemical abortion is the usual method, though there are others. Vacuum aspiration is usually the legal method in most countries, although research has uncovered similar efficacy and safety from methotrexate and misoprostol. Combination therapy with chemical abortion and vacuum aspiration is also common, with many women choosing to have the chemical abortion (Pills alone or pills followed by a vaginal or oral medication) at home and then the vacuum aspiration (a short and safe procedure) in a medical facility around the fifteenth week. Vacuum aspiration, vacuum aspiration dilatation and curettage (VAC), and from the fifteenth week up until around the eighteenth week, a surgical dilation and evacuation (D&E) is used.

As the fetus size increases other techniques are used to secure abortion in the third trimester. Premature expulsion of the fetus can be induced with prostaglandins, which can be coupled with injecting the amniotic fluid with saline or a solution. Very late abortions can be brought about by the controversial intact dilation and extraction (D&E) or a hysterotomy, similar to a caesarean section.

The controversy

The morality and legality of abortion is a controversial topic in many countries, and is also discussed by legal scholars and religious groups. Important facts about abortion are also researched by sociologists and historians.

Abortion has been common in most societies, although it has often been opposed by some institutionalized religions and governments. In the United States, abortion became commonly accepted by the end of the 20th century. Additionally, abortion is a legal right in China, India, and other countries. The Catholic Church remains opposed to the procedure, however, and in other countries, notably the United States and the (predominantly Catholic) Republic of Ireland, the controversy is extremely active, to the extent that even the laws of the respective positions are subject to经常 debate. While the debate has been generally peaceful, the issue of abortion is sometimes characterized by violence. Though true of both sides, this is more marked on the side of those opposed to abortion, because of what they see as the gravity and urgency of their views.

The central question

The central question in the abortion debate is whether the potential life of a fetus is similar to that of a newborn. On one hand, is a fetus a human being from the moment of conception? On the other hand, is abortion a choice of perceived or presumed rights. On the other hand, is a fetus (sometimes called the “pro-life/anti-abortion advocates”) a human being from the moment of conception, and if so, at what point does the fetus become human?
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*, Jan 2009
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]
<table>
<thead>
<tr>
<th>Set A</th>
<th>X</th>
<th>Y</th>
<th>Set B</th>
<th>X</th>
<th>Y</th>
<th>Set C</th>
<th>X</th>
<th>Y</th>
<th>Set D</th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>8.04</td>
<td></td>
<td>10</td>
<td>9.14</td>
<td></td>
<td>10</td>
<td>7.46</td>
<td></td>
<td>8</td>
<td>6.58</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>6.95</td>
<td></td>
<td>8</td>
<td>8.14</td>
<td></td>
<td>8</td>
<td>6.77</td>
<td></td>
<td>8</td>
<td>5.76</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>7.58</td>
<td></td>
<td>13</td>
<td>8.74</td>
<td></td>
<td>13</td>
<td>12.74</td>
<td></td>
<td>8</td>
<td>7.71</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>8.81</td>
<td></td>
<td>9</td>
<td>8.77</td>
<td></td>
<td>9</td>
<td>7.11</td>
<td></td>
<td>8</td>
<td>8.84</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>8.33</td>
<td></td>
<td>11</td>
<td>9.26</td>
<td></td>
<td>11</td>
<td>7.81</td>
<td></td>
<td>8</td>
<td>8.47</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>9.96</td>
<td></td>
<td>14</td>
<td>8.1</td>
<td></td>
<td>14</td>
<td>8.84</td>
<td></td>
<td>8</td>
<td>7.04</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>7.24</td>
<td></td>
<td>6</td>
<td>6.13</td>
<td></td>
<td>6</td>
<td>6.08</td>
<td></td>
<td>8</td>
<td>5.25</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.26</td>
<td></td>
<td>4</td>
<td>3.1</td>
<td></td>
<td>4</td>
<td>5.39</td>
<td></td>
<td>19</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>10.84</td>
<td></td>
<td>12</td>
<td>9.11</td>
<td></td>
<td>12</td>
<td>8.15</td>
<td></td>
<td>8</td>
<td>5.56</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>4.82</td>
<td></td>
<td>7</td>
<td>7.26</td>
<td></td>
<td>7</td>
<td>6.42</td>
<td></td>
<td>8</td>
<td>7.91</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5.68</td>
<td></td>
<td>5</td>
<td>4.74</td>
<td></td>
<td>5</td>
<td>5.73</td>
<td></td>
<td>8</td>
<td>6.89</td>
</tr>
</tbody>
</table>
Summary Statistics

**Set A**

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>8.04</td>
</tr>
<tr>
<td>8</td>
<td>6.95</td>
</tr>
<tr>
<td>13</td>
<td>7.58</td>
</tr>
<tr>
<td>9</td>
<td>8.81</td>
</tr>
<tr>
<td>11</td>
<td>8.33</td>
</tr>
<tr>
<td>14</td>
<td>9.96</td>
</tr>
<tr>
<td>6</td>
<td>7.24</td>
</tr>
<tr>
<td>4</td>
<td>4.26</td>
</tr>
<tr>
<td>12</td>
<td>10.84</td>
</tr>
<tr>
<td>7</td>
<td>4.82</td>
</tr>
<tr>
<td>5</td>
<td>5.68</td>
</tr>
</tbody>
</table>

**Set B**

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>9.14</td>
</tr>
<tr>
<td>8</td>
<td>8.14</td>
</tr>
<tr>
<td>13</td>
<td>8.74</td>
</tr>
<tr>
<td>9</td>
<td>8.77</td>
</tr>
<tr>
<td>11</td>
<td>9.26</td>
</tr>
<tr>
<td>14</td>
<td>8.1</td>
</tr>
<tr>
<td>6</td>
<td>6.13</td>
</tr>
<tr>
<td>4</td>
<td>3.1</td>
</tr>
<tr>
<td>12</td>
<td>9.11</td>
</tr>
<tr>
<td>7</td>
<td>7.26</td>
</tr>
<tr>
<td>5</td>
<td>4.74</td>
</tr>
</tbody>
</table>

**Set C**

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>7.46</td>
</tr>
<tr>
<td>8</td>
<td>6.77</td>
</tr>
<tr>
<td>13</td>
<td>12.74</td>
</tr>
<tr>
<td>9</td>
<td>7.11</td>
</tr>
<tr>
<td>11</td>
<td>7.81</td>
</tr>
<tr>
<td>14</td>
<td>8.84</td>
</tr>
<tr>
<td>6</td>
<td>6.08</td>
</tr>
<tr>
<td>4</td>
<td>5.39</td>
</tr>
<tr>
<td>12</td>
<td>8.15</td>
</tr>
<tr>
<td>7</td>
<td>6.42</td>
</tr>
<tr>
<td>5</td>
<td>5.73</td>
</tr>
</tbody>
</table>

**Set D**

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>6.58</td>
</tr>
<tr>
<td>8</td>
<td>5.76</td>
</tr>
<tr>
<td>8</td>
<td>7.71</td>
</tr>
<tr>
<td>8</td>
<td>8.84</td>
</tr>
<tr>
<td>8</td>
<td>8.47</td>
</tr>
<tr>
<td>8</td>
<td>7.04</td>
</tr>
<tr>
<td>8</td>
<td>5.25</td>
</tr>
<tr>
<td>19</td>
<td>12.5</td>
</tr>
<tr>
<td>8</td>
<td>5.56</td>
</tr>
<tr>
<td>8</td>
<td>7.91</td>
</tr>
<tr>
<td>8</td>
<td>6.89</td>
</tr>
</tbody>
</table>

**Linear Regression**

**Summary Statistics**

\[ u_X = 9.0 \quad \sigma_X = 3.317 \]

\[ u_Y = 7.5 \quad \sigma_Y = 2.03 \]

[Anscombe 1973]
Why Create Visualizations?
Why Create Visualizations?

Answer questions (or discover them)
Make decisions
See data in context
Expand memory
Support graphical calculation
Find patterns
Present argument or tell a story
Inspire
Record Information
Gallop, Bay Horse “Daisy” [Muybridge 1884-86]
E.J. Marey’s sphygmograph [from Braun 83]
Support Reasoning
### History of O-Ring Damage on SRM Field Joints

<table>
<thead>
<tr>
<th>Field Joints</th>
<th>Cross Sectional View</th>
<th>Top View</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SRM No.</td>
<td>Erosion Depth (in.)</td>
</tr>
<tr>
<td>61A LH Center Field**</td>
<td>22A</td>
<td>None</td>
</tr>
<tr>
<td>61A LH CENTER FIELD**</td>
<td>22B</td>
<td>None</td>
</tr>
<tr>
<td>51C LH Forward Field**</td>
<td>15A</td>
<td>0.010</td>
</tr>
<tr>
<td>51C LH Center Field (prim)**</td>
<td>15B</td>
<td>0.038</td>
</tr>
<tr>
<td>51C RH Center Field (sec)**</td>
<td>15B</td>
<td>0.040</td>
</tr>
<tr>
<td>41D RH Forward Field</td>
<td>13B</td>
<td>0.028</td>
</tr>
<tr>
<td>41C LH Aft Field*</td>
<td>11A</td>
<td>None</td>
</tr>
<tr>
<td>41B LH Forward Field</td>
<td>10A</td>
<td>0.040</td>
</tr>
<tr>
<td>STS-2 RH Aft Field</td>
<td>2B</td>
<td>0.053</td>
</tr>
</tbody>
</table>

---

*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 (50°, 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)

<table>
<thead>
<tr>
<th></th>
<th>MOTOR</th>
<th>MGT</th>
<th>AMB</th>
<th>O-RING</th>
<th>WIND</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM-1</td>
<td>68</td>
<td>36</td>
<td>47</td>
<td></td>
<td>10 mph</td>
</tr>
<tr>
<td>DM-2</td>
<td>76</td>
<td>45</td>
<td>52</td>
<td></td>
<td>10 mph</td>
</tr>
<tr>
<td>GM-3</td>
<td>72.5</td>
<td>40</td>
<td>48</td>
<td></td>
<td>10 mph</td>
</tr>
<tr>
<td>GM-4</td>
<td>76</td>
<td>48</td>
<td>51</td>
<td></td>
<td>10 mph</td>
</tr>
<tr>
<td>SRM-15</td>
<td>52</td>
<td>64</td>
<td>53</td>
<td></td>
<td>10 mph</td>
</tr>
<tr>
<td>SRM-22</td>
<td>77</td>
<td>78</td>
<td>75</td>
<td></td>
<td>10 mph</td>
</tr>
<tr>
<td>SRM-25</td>
<td>55</td>
<td>26</td>
<td>29</td>
<td></td>
<td>25 mph</td>
</tr>
<tr>
<td>SRM-29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Make a Decision: Challenger
Visualizations drawn by Tufte show how low temperatures damage O-rings [Tufte 97]
In 1854 John Snow plotted the position of each cholera case on a map. [from Tufte 83]
Used map to hypothesize that pump on Broad St. was the cause. [from Tufte 83]
Expand Memory: Multiplication

Class Exercise
Expand Memory: Multiplication

34
x 72
Expand Memory: Multiplication

\[ \begin{align*}
34 & \times 72 \\
& \quad 68 \\
& \underline{2380} \\
& \underline{2448}
\end{align*} \]

- Mental: 110 seconds
- Paper & Pencil: 28 seconds
Find Patterns: NYC Weather

The Most Powerful Brain?

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Body Weight</th>
<th>Brain Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lesser Short-tailed Shrew</td>
<td>5</td>
<td>0.14</td>
</tr>
<tr>
<td>2</td>
<td>Little Brown Bat</td>
<td>10</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>Mouse</td>
<td>23</td>
<td>0.3</td>
</tr>
<tr>
<td>4</td>
<td>Big Brown Bat</td>
<td>23</td>
<td>0.4</td>
</tr>
<tr>
<td>5</td>
<td>Musk Shrew</td>
<td>48</td>
<td>0.33</td>
</tr>
<tr>
<td>6</td>
<td>Star Nosed Mole</td>
<td>60</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Eastern American Mole</td>
<td>75</td>
<td>1.2</td>
</tr>
<tr>
<td>8</td>
<td>Ground Squirrel</td>
<td>101</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>Tree Shrew</td>
<td>104</td>
<td>2.5</td>
</tr>
<tr>
<td>10</td>
<td>Golden Hamster</td>
<td>120</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Mole Rate</td>
<td>122</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>Galago</td>
<td>200</td>
<td>5</td>
</tr>
<tr>
<td>13</td>
<td>Rat</td>
<td>280</td>
<td>1.9</td>
</tr>
<tr>
<td>14</td>
<td>Chinchilla</td>
<td>425</td>
<td>6.4</td>
</tr>
<tr>
<td>15</td>
<td>Desert Hedgehog</td>
<td>550</td>
<td>2.4</td>
</tr>
<tr>
<td>16</td>
<td>Rock Hyrax (a)</td>
<td>750</td>
<td>12.3</td>
</tr>
<tr>
<td>17</td>
<td>European Hedgehog</td>
<td>785</td>
<td>3.5</td>
</tr>
<tr>
<td>18</td>
<td>Tenrec</td>
<td>900</td>
<td>2.6</td>
</tr>
<tr>
<td>19</td>
<td>Arctic Ground Squirrel</td>
<td>920</td>
<td>5.7</td>
</tr>
<tr>
<td>20</td>
<td>African Giant Pouched Rat</td>
<td>1000</td>
<td>6.6</td>
</tr>
<tr>
<td>21</td>
<td>Guinea Pig</td>
<td>1040</td>
<td>5.5</td>
</tr>
<tr>
<td>22</td>
<td>Mountain Beaver</td>
<td>1350</td>
<td>8.1</td>
</tr>
<tr>
<td>23</td>
<td>Slow Loris</td>
<td>1400</td>
<td>12.5</td>
</tr>
<tr>
<td>24</td>
<td>Genet</td>
<td>1410</td>
<td>17.5</td>
</tr>
<tr>
<td>25</td>
<td>Phalanger</td>
<td>1620</td>
<td>11.4</td>
</tr>
</tbody>
</table>
The Elements of Graphing Data
[Cleveland]
Convey Information to Others
Inspire

Bones in hand [from 1918 edition]

Double helix model [Watson and Crick 53]
Diagram of the Causes of Mortality

in the Army in the East.

1. April 1854 to March 1855.

2. April 1855 to March 1856.

The areas of the blue, red, & black wedges are each measured from the centre as the common vertex.

The blue wedges measured from the centre of the circle represent areas: the deaths from Preventable or Mitigable Zymotic diseases; the red wedges measured from the centre, the deaths from wounds; & the black wedges measured from the centre, the deaths from all other causes.

The black line across the red triangle in Nov. 1854 marks the boundary of the deaths from all other causes during this month.

In October 1854 & April 1855, the black area coincides with the red; in January & February 1855, the blue coincides with the black.

The entire areas may be compared by following the blue, the red & the black lines enclosing them.

1856 "Coxcomb" of Crimean War Deaths, Florence Nightingale
“to affect thro’ the Eyes what we fail to convey to the public through their word-proof ears”

1856 “Coxcomb” of Crimean War Deaths, Florence Nightingale
The Value of Visualization

**Record** information
    - Blueprints, photographs, seismographs, ...

**Analyze** data to support reasoning
    - Develop and assess hypotheses
    - Find patterns / Discover errors in data
    - Expand memory

**Communicate** information to others
    - Share and persuade
    - Collaborate and revise
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 Authoring
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]
1,200 exabytes in 2010 [Gantz 10]

Necessitates better tools and algorithms for visually conveying information
“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
Goals of Visualization Research

1 **Understand** how visualizations convey information
   What do people perceive/comprehend?
   How do visualizations inform mental models?

2 **Develop principles and techniques** for creating effective visualizations and supporting analysis
   Amplify perception and cognition
   Improve ties between visualization & mental model
Course Topics
**Data and Image Models**

### Les Variables de l'Image

<table>
<thead>
<tr>
<th>Points</th>
<th>Lignes</th>
<th>Zones</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Y 2 Dimensions du Plan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z Taille</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valeur</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Les Variables de Séparation des Images

<table>
<thead>
<tr>
<th>Grain</th>
<th>Couleur</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sémiologie Graphique [Bertin 67]
Visualization (Re-)Design

SlicerDicers' Sales Compared to Other Products

Problematic design

Redesign

Sales of SlicerDicers Compared to Other Products
July - December, 2003
(SlicerDicers' sales are displayed as black reference lines of 100%, the red bars represent the average monthly sales percentage for July through December.)
Visualization Software

D3: Data-Driven Documents
The psychophysics of sensory function [Stevens 61]
CIE 1931 xy chromaticity diagram showing the gamut of the sRGB and Adobe RGB color spaces including the Planckian locus, with temperatures indicated. Wavelengths of monochromatic light are shown in blue.
Animation

Animated transitions in statistical data graphics [Heer & Robertson 07]
Graphs and Trees

Degree-Of-Interest Trees [Heer & Card 04]
Text Visualization

Creator: Martin Waltermont
Tags:

118 hits

i don't recall

want to

know

believe

think

have

Data source: CQ Transcript Wire via the Washington Post

Comments (4)
Collaboration and History

Dentist

1850 1870 1900 1920 1940 1960 1980 2000

0.14% 0.12% 0.1% 0.08% 0.06% 0.04% 0.02%
Collaboration and History

Where have all the dentists gone?
Collaboration and History
Course Mechanics
You should expect to:

1. Evaluate and critique visualization designs
2. Implement interactive data visualizations
3. Gain an overview of research & techniques
4. Develop a substantial visualization project
Instructors

Instructor

Jeffrey Heer  
Assoc Prof, CSE

OH: Th 10-11am, 642 CSE

http://jheer.org

Assistants

Dominik Moritz

OH: Fri 1:30-2:30, 218 CSE

Jeff Snyder

OH: Fri 1:30-2:30, 218 CSE
Graduated from University of Potsdam
1½ year PhD student with Bill Howe & Jeff Heer
Databases + Visualization
homes.cs.washington.edu/~domoritz
Hello
my name is...

Jeff Snyder

Office Hours: F 1:30-2:30 CSE 218

- graduated Princeton in 2013
- 2nd year PhD student with Jeff Heer
- part time at Google Seattle, working on coLaboratory
- loves javascript, warts and all
- more at http://jsnyd.es
Textbooks

See also: www.edwardtufte.com
Readings

Some from textbooks, also many papers
Material in class will loosely follow readings
Readings should be read by start of class
Post discussion comments on class Canvas forum
  Comments must be posted within 1 day of lecture
You have 2 “passes” for the quarter
Assignments

Class Participation (10%)
A1: Visualization Design (10%)
A2: Exploratory Data Analysis (15%)
A3: Interactive Visualization Software (25%)
FP: Final Project (40%)
Final Project

Visualization research project on topic of choice

Project write-up in form of a short research paper

Project check-ins and final poster/demo show

Projects from previous classes have been:

- Published (e.g., at the IEEE InfoVis conference)
- Featured in the New York Times
- Released as successful open source projects
Divided Edge Bundling – David Selassie
Visualizing Galaxy Merger Trees

S. Loebman, J. Ortiz, L. Orr, M. Balazinska, T. Quinn et al. [SIGMOD '14]
Perfopticon: Distributed Query Performance

Physical Query Plan:

Fragment 2

- ShuffleConsumer

Fragment 3

- ShuffleConsumer

Fragment 1

- SymmetricHashJoin(((($0 = $7) and ($3 = $6)) and ($2 = $5)) and ($1 = $4)); $1)

- GroupBy($0; COUNTALL)

- ShuffleProducer(h($0))

Fragment 0

Overview / Operators inside fragment 1

Query time contribution collapse/expand

- SymmetricHashJoin(((($0 = $7) and ($3 = $6)) and ($2 = $5)) and ($1 = $4)); $1)

3 % 3 % 92 %

Detailed execution

Dominik Moritz et al. [EuroVis '15]
Protovis: A Graphical Toolkit for Visualization

Mike Bostock
```javascript
var army = pd.nest(napoleon.army, "dir", "group");
var vis = new pv.Panel();

var lines = vis.add(pv.Panel).data(army);
lines.add(pv.Line)
  .data(function() army[this.idx])
  .left(lon).top(lat).size(function(d) d.size/8000)
  .strokeStyle(function() color[army[paneIndex][0].dir]);

vis.add(pv.Label).data(napoleon.cities)
  .left(lon).top(lat)
  .text(function(d) d.city).font("italic 10px Georgia")
  .textAlign("center").textBaseline("middle");

vis.add(pv.Rule).data([0, -10, -20, -30])
  .top(function(d) 300 - 2*d - 0.5).left(200).right(150)
  .lineWidth(1).strokeStyle("#ccc")
  .anchor("right").add(pv.Label)
  .font("italic 10px Georgia")
  .text(function(d) d.temp + "° " + d.date.substr(0, 6)).textBaseline("center");

vis.add(pv.Line).data(napoleon.temp)
  .left(lon).top(tmp)
  .strokeStyle("#0")
  .add(pv.Label)
  .top(function(d) 5 + tmp(d))
  .text(function(d) d.temp + "° " + d.date.substr(0, 6))
  .textBaseline("top").font("italic 10px Georgia");
```
Visualizing the Republic of Letters

Daniel Chang, Yuankai Ge, Shiwei Song

Republic of Letters
1700

FILTER BY AUTHOR
- Damien Desormes
- Daniel Carnabs
- Daniel de Pury
- Daniel Defoe
- Daniel Mathus
- Daniel Marc Antoine Chardon
- Daniel Muller

TOP CITIES AND AUTHORS

<table>
<thead>
<tr>
<th>Letters received</th>
<th>Letters sent</th>
</tr>
</thead>
<tbody>
<tr>
<td>London England</td>
<td>346</td>
</tr>
<tr>
<td>Oates England</td>
<td>304</td>
</tr>
<tr>
<td>Dublin Ireland</td>
<td>208</td>
</tr>
<tr>
<td>Paris France</td>
<td>238</td>
</tr>
<tr>
<td>Twickenham England</td>
<td>18</td>
</tr>
<tr>
<td>John Locke</td>
<td>359</td>
</tr>
<tr>
<td>Joseph Addison</td>
<td>30</td>
</tr>
<tr>
<td>Voltaire</td>
<td>26</td>
</tr>
<tr>
<td>Jonathan Swift</td>
<td>85</td>
</tr>
<tr>
<td>Alexander Pope</td>
<td>28</td>
</tr>
</tbody>
</table>
Questions?
Assignment 1: Visualization Design

Design a static visualization for a data set.

After the World War II, antibiotics were considered “wonder drugs.” To learn which drug is most effective for which bacterial infection, performance of the three most popular antibiotics were gathered.

You must choose the message you want to convey. What task do you want to support? What insight do you want to communicate?
Assignment 1: Visualization Design

Design a static visualization for the data set.
You are free to use any tools (inc. pen & paper).

Deliverables (upload via Canvas; see A1 page)
Image of your visualization (PNG or JPG format)
Short description + design rationale (≤ 4 paragraphs)

Due by 5:00 pm, Monday April 6.