How much data (bytes) did we produce in 2010?
2010: 1,200 exabytes and exponential growth...

Gantz et al., 2008, 2010
Data Created & Consumed
Source: IDC Digital Universe

~2x every 2 years

- 1.2 ZB
- 4.4 ZB
- 44.0 ZB
But what is in all this data?

Data Created & Consumed
Source: IDC Digital Universe

~2x every 2 years

1.2 ZB
4.4 ZB
44.0 ZB
Physical Sensors

Image courtesy cabspotting.org
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR/ECG</td>
<td>97</td>
</tr>
<tr>
<td>Art</td>
<td>82/60</td>
</tr>
<tr>
<td>SpO2</td>
<td>99</td>
</tr>
<tr>
<td>RR/CO2</td>
<td>------</td>
</tr>
<tr>
<td>HR/ECG</td>
<td>79</td>
</tr>
<tr>
<td>Art</td>
<td>152/79</td>
</tr>
<tr>
<td>SpO2</td>
<td>95</td>
</tr>
<tr>
<td>RR/CO2</td>
<td>------</td>
</tr>
<tr>
<td>HR/ECG</td>
<td>64</td>
</tr>
<tr>
<td>Art</td>
<td>93/55</td>
</tr>
<tr>
<td>SpO2</td>
<td>99</td>
</tr>
<tr>
<td>RR/Imp</td>
<td>------</td>
</tr>
</tbody>
</table>
Records of Human Activity
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.

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. “Free” to whom? “Ubiquitous” about whom? …to whose benefit?

Hal Varian, Google’s Chief Economist

*The McKinsey Quarterly*, Jan 2009
High potential for data abuse...
Inequality

Rise of the racist robots - how AI is learning all our worst impulses

There is a saying in computer science: garbage in, garbage out. When we feed machines data that reflects our prejudices, they mimic them - from antisemitic chatbots to racially biased software. Does a horrifying future await people forced to live at the mercy of algorithms?

...amplified by “big data” and ML systems.
We move from data to information to knowledge to wisdom, and separating one from the other, being able to distinguish among and between them that is, knowing the limitations and the danger of exercising one without the others while respecting each category of intelligence, is generally what serious education is about.

Toni Morrison, American Novelist

*The Source of Self Regard*
How might we use visualization to empower understanding of data and analysis processes?
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>X</th>
<th>Y</th>
<th>X</th>
<th>Y</th>
<th>X</th>
<th>Y</th>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>8.04</td>
<td>10</td>
<td>9.14</td>
<td>10</td>
<td>7.46</td>
<td>8</td>
<td>6.58</td>
</tr>
<tr>
<td>8</td>
<td>6.95</td>
<td>8</td>
<td>8.14</td>
<td>8</td>
<td>6.77</td>
<td>8</td>
<td>5.76</td>
</tr>
<tr>
<td>13</td>
<td>7.58</td>
<td>13</td>
<td>8.74</td>
<td>13</td>
<td>12.74</td>
<td>8</td>
<td>7.71</td>
</tr>
<tr>
<td>9</td>
<td>8.81</td>
<td>9</td>
<td>8.77</td>
<td>9</td>
<td>7.11</td>
<td>8</td>
<td>8.84</td>
</tr>
<tr>
<td>11</td>
<td>8.33</td>
<td>11</td>
<td>9.26</td>
<td>11</td>
<td>7.81</td>
<td>8</td>
<td>8.47</td>
</tr>
<tr>
<td>14</td>
<td>9.96</td>
<td>14</td>
<td>8.1</td>
<td>14</td>
<td>8.84</td>
<td>8</td>
<td>7.04</td>
</tr>
<tr>
<td>6</td>
<td>7.24</td>
<td>6</td>
<td>6.13</td>
<td>6</td>
<td>6.08</td>
<td>8</td>
<td>5.25</td>
</tr>
<tr>
<td>4</td>
<td>4.26</td>
<td>4</td>
<td>3.1</td>
<td>4</td>
<td>5.39</td>
<td>19</td>
<td>12.5</td>
</tr>
<tr>
<td>12</td>
<td>10.84</td>
<td>12</td>
<td>9.11</td>
<td>12</td>
<td>8.15</td>
<td>8</td>
<td>5.56</td>
</tr>
<tr>
<td>7</td>
<td>4.82</td>
<td>7</td>
<td>7.26</td>
<td>7</td>
<td>6.42</td>
<td>8</td>
<td>7.91</td>
</tr>
<tr>
<td>5</td>
<td>5.68</td>
<td>5</td>
<td>4.74</td>
<td>5</td>
<td>5.73</td>
<td>8</td>
<td>6.89</td>
</tr>
</tbody>
</table>

**Summary Statistics**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_X$</td>
<td>9.0</td>
<td>$\sigma_X$</td>
<td>3.32</td>
</tr>
<tr>
<td>$u_Y$</td>
<td>7.5</td>
<td>$\sigma_Y$</td>
<td>2.03</td>
</tr>
</tbody>
</table>

**Linear Regression**

$Y = 3 + 0.5X$

$R^2 = 0.67$

[Anscombe 1973]
Set A

Set B

Set C

Set D

[Anscombe 1973]
Abortion

In its most commonly used senses, abortion 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 causes, such as spontaneous abortion or miscarriage. In most cases, abortion is performed during the first 20 weeks of pregnancy. If performed after the twentieth week of pregnancy, it is known as a late abortion.

Methods

Depending on the stage of pregnancy and the location, abortion can be performed by a number of different methods. The most common method in the first trimester is a vacuum aspiration. In later pregnancy, a dilation and evacuation (D&E) procedure is used. In cases where the fetus is dead and expelled, it is called a miscarriage.

The controversy

The morality and legality of abortion is a subject of considerable debate and is also discussed by legal scholars and religious and ethical perspectives. The debate is also relevant to sociologists and historians.

Abortion has been common in most societies throughout history, although it has often been opposed by some religious and cultural traditions. In the 20th century, policies in the United States and Europe have generally favored legal abortion, whereas in China and other countries, it is more restricted.

The central question in the abortion debate is whether the right to life of the fetus is to be recognized and protected. On the one hand, a fetus is a human being with a right to life, and on the other, it is a product of human intervention (implantation).
Wikipedia History Flow [Viegas & Wattenberg]
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]
E.J. Marey’s sphygmograph [from Braun 83]
You Draw It: How Family Income Predicts Children’s College Chances

You Draw It: How Family Income Predicts Children’s College Chances

Support Reasoning
### HISTORY OF O-RING DAMAGE ON SRM FIELD JOINTS

<table>
<thead>
<tr>
<th>SRM No.</th>
<th>Erosion Depth (in.)</th>
<th>Perimeter Affected (deg.)</th>
<th>Nominal Dia. (in.)</th>
<th>Length Of Max Erosion (in.)</th>
<th>Total Heat Affected Length (in.)</th>
<th>Clamping Location (deg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>22A</td>
<td>None</td>
<td>None</td>
<td>0.280</td>
<td>None</td>
<td>None</td>
<td>36° - 66°</td>
</tr>
<tr>
<td>22A</td>
<td>None</td>
<td>None</td>
<td>0.280</td>
<td>None</td>
<td>None</td>
<td>330° - 18°</td>
</tr>
<tr>
<td>15A</td>
<td>0.010</td>
<td>154.0</td>
<td>0.280</td>
<td>4.25</td>
<td>5.25</td>
<td>163</td>
</tr>
<tr>
<td>15B</td>
<td>0.038</td>
<td>130.0</td>
<td>0.280</td>
<td>12.50</td>
<td>58.75</td>
<td>354</td>
</tr>
<tr>
<td>15B</td>
<td>None</td>
<td>45.0</td>
<td>0.280</td>
<td>None</td>
<td>29.50</td>
<td>354</td>
</tr>
<tr>
<td>13B</td>
<td>0.028</td>
<td>110.0</td>
<td>0.280</td>
<td>3.00</td>
<td>None</td>
<td>275</td>
</tr>
<tr>
<td>11A</td>
<td>None</td>
<td>None</td>
<td>0.280</td>
<td>None</td>
<td>None</td>
<td>--</td>
</tr>
<tr>
<td>10A</td>
<td>0.040</td>
<td>217.0</td>
<td>0.280</td>
<td>3.00</td>
<td>14.50</td>
<td>351</td>
</tr>
<tr>
<td>28</td>
<td>0.053</td>
<td>116.0</td>
<td>0.280</td>
<td>--</td>
<td>--</td>
<td>90</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-12 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>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>10 MPH</td>
</tr>
<tr>
<td>DM-2</td>
<td>76</td>
<td>45</td>
<td>52</td>
<td>10 MPH</td>
</tr>
<tr>
<td>GM-3</td>
<td>72.5</td>
<td>40</td>
<td>48</td>
<td>10 MPH</td>
</tr>
<tr>
<td>GM-4</td>
<td>76</td>
<td>48</td>
<td>51</td>
<td>10 MPH</td>
</tr>
<tr>
<td>SRM-15</td>
<td>52</td>
<td>64</td>
<td>53</td>
<td>10 MPH</td>
</tr>
<tr>
<td>SRM-22</td>
<td>77</td>
<td>78</td>
<td>75</td>
<td>10 MPH</td>
</tr>
<tr>
<td>SRM-25</td>
<td>55</td>
<td>26</td>
<td>29</td>
<td>25 MPH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27</td>
</tr>
</tbody>
</table>
Make Decisions: Challenger
But wait! What is an appropriate “damage index”? Which temperatures, O-ring or outside air?

Chart of temperatures vs. O-ring damage [Tufte 97]
Data in Context: Cholera Outbreak

In 1854 John Snow plotted the position of each cholera case on a map. [from Tufte 83]
Data in Context: Cholera Outbreak

Used map to hypothesize that pump on Broad St. was the cause. [from Tufte 83]
Find Patterns: NYC Weather

### Answer Questions: Brain Power?

<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 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
Communicate, Inform, Inspire

Visualizing Black America, Du Bois et al. 1900

Bones in hand, Gray’s Anatomy 1918 ed.
New deaths attributed to Covid-19 in European Union, United States, Brazil and United Kingdom

Seven-day rolling average of new deaths, by number of days since 3 average daily deaths first recorded

Source: Financial Times analysis of data from the European Centre for Disease Prevention and Control, the Covid Tracking Project, the UK Dept of Health & Social Care and the Spanish Ministry of Health.
Data updated September 25 2020 12.46pm BST. Interactive version: ft.com/covid19
The coronavirus crisis is different

Job growth (or loss) since each recession began, based on weekly earnings

**1990 recession**

**2001 recession**

**2008 recession**

**Coronavirus crisis**

Notes: Based on a three-month average to show the trend in volatile data.
Source: Labor Department via IPUMS, with methodology assistance from Ernie Tedeschi of Evercore ISI

THE WASHINGTON POST
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

**Convey** information
- Communicate, inform, inspire
- Collaborate and revise
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
   Leverage perception & augment cognition
   Improve ties between visualization & mental model
Course Topics
### Data and Image Models

#### Sémiologie Graphique [Bertin 67]

<table>
<thead>
<tr>
<th>XY</th>
<th>Points</th>
<th>Lignes</th>
<th>Zones</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Dimensions du Plan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAILLE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VALEUR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRAIN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COULEUR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORIENTATION</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Les variables de l'image</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 14</td>
</tr>
<tr>
<td>OQ</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Les variables de séparation des images</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
</tr>
<tr>
<td>OQ</td>
</tr>
</tbody>
</table>

---

*Note: The diagram illustrates various visual variables used in data visualization, including points, lines, zones, size, value, grain, color, and orientation.*
Visualization Design

Problematic design

Redesign
Exploratory Data Analysis
Dymaxion Maps [Fuller 46]
Recent elections have placed a heavy emphasis on “swing states” — Ohio, Florida and the other competitive states. You have to be careful to look beyond the numbers between the Democratic and Republican parties. A look at how the states shifted during the last elections and how they have shifted over past elections.

Obama Re-elected
The country voted about 5 percentage points more Republican in 2012 than in 2008. Obama lost North Carolina and Indiana, but won every tossup except Florida, which remains too close to call.

As Goes Ohio
Ohio, which has voted for the winner in every election since 1964, provided the decisive electoral votes in 2004, and it is the state likeliest to play that role again this year, according to the FiveThirtyEight model.
Visualization Software

D3: Data-Driven Documents
Vega-Lite / Altair
Animated transitions in statistical data graphics [Heer & Robertson 07]
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.
Graphical Perception

The psychophysics of sensory function [Stevens 61]
Uncertainty
Hierarchies

Degree-Of-Interest Trees [Heer & Card 04]
Networks
Scalability

Interactive querying of 180M flight records in Falcon [Moritz et al. 2019]
Course Mechanics
You should expect to:

1. Evaluate and critique visualization designs
2. Learn visualization techniques & theory
3. Implement interactive data visualizations
4. Develop a substantial visualization project
Instructors

Instructor
Leilani Battle  OH: Wed 4-5pm
Assistant Professor, CSE
https://homes.cs.washington.edu/~leibatt/

Teaching Assistants

Vishal Devireddy  OH: Mon 11am-12pm
Philip Garrison  OH: TBD
Brian Hou  OH: Online / Ed
Chandler Peterson  OH: TBD
Firn Tieanklin  OH: Wed 10:30-11:30am
Leilani Battle
Assistant Professor, UW CSE
Co-Director, CSE Interactive Data Lab
https://homes.cs.washington.edu/~leibatt/

Visualization / HCI / Data management / Data Science

I model relationships between analysts’ *intents*, i.e., analysis goals, and *behaviors*, i.e., patterns of interaction with data analysis systems.

I use these models to build *behavior-driven* optimizations, UI features, and performance benchmarks for interactive data analysis.

Hobbies: Travel, crafting, disc golf, board games, reading, etc.
Vishal Devireddy

<table>
<thead>
<tr>
<th>Email</th>
<th>vishald@cs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office hours</td>
<td>Mon. 11 am</td>
</tr>
</tbody>
</table>

I’m an MS student interested in web development, HCI, and perfectly aligning monospace text. My current research is with IDL on tools to support easily authoring responsive, interactive academic papers. Super excited to TA CSE 512!

<table>
<thead>
<tr>
<th>Some things I can help with:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>🌍 JavaScript</td>
<td>✨ C S S</td>
</tr>
<tr>
<td>🎨 Web design</td>
<td>📊 D3.js</td>
</tr>
</tbody>
</table>
Philip Garrison (he/him)

Hi! I’m a PhD student in CSE and I have been developing a data visualization platform for cold-chain equipment (refrigerators that store vaccines). My current research is about the social & political context of that platform. Outside of research and teaching, I like making music, and I stay involved in activism on and off campus.

philipmg@cs.uw.edu
Brian Hou
he/him

• PhD student in the Personal Robotics Lab

• Research interests: robot motion planning, reinforcement learning

• Non-research interests: baking, casual games, crosswords, basketball/baseball
Nussara ‘Firn’ Tieanklin

Office Hour: Wed 10:30-11:30 AM
nussara@cs

Research @ICTD Lab

- **Rideshare on different SES**: Understanding the effects of people’s SES on using rideshare/food delivery services in Southeast Asia.
- **Seattle Community Networks**: Providing internet access to resource-constrained communities in Washington

Technical Experience

- User research, Design process, Data Management, Web-programming

Things I do for fun

- Play Badminton 🏓
- Explore new bakeries and dessert cafes 🍪
- Play video games 🎮
- Travel 🚗
Chandler Petersen

- 2nd-year CSE PhD student
- Advisor: Georg Seelig
- Research Area: DNA computing and molecular programming
- Working to scale up the synthesis, computation, and readout of DNA logic circuits with Next Generation Sequencing
- Interested in the creating tools to better visualize and teach DNA strand displacement circuits
Readings

From books, notebooks, and linked articles. Material in class will loosely follow readings. Readings should be read by start of class.

Post comments & quizzes on class forum.

One comment per week.

Post comments by Friday 11:59pm.

You have 2 “passes” for the quarter.
“Textbooks”

See also: [www.edwardtufte.com](http://www.edwardtufte.com)
Interactive Notebooks

Hands-on engagement with course concepts and visualization tools (Vega-Lite / Altair), in both JavaScript (Observable) and Python (Jupyter).
Optional Book

Interactive Data Visualization for the Web, 2nd Edition

For learning D3!
Book available online.
Code / examples on GitHub.

We will be using D3 v6.
https://d3js.org

Scott Murray
Assignments

**CP** Class Participation (10%)

**A1** Visualization Design (10%) - Due 4/6

**A2** Deceptive Visualization (15%) - Due 4/22
Peer Evaluation - Due 4/29

**A3** Interactive Prototype (25%) - Due 5/10
Peer Evaluation - Due 5/17

**FP** Final Project (40%)
Proposal - Due 5/18
Milestone Prototype - Due 5/27
Final Projects Showcase - Posters Due 6/1
Final Prototype - Due 6/7
Final Project

Visualization research project on topic of choice
Initial prototype and design reviews
In-class demonstration video showcase
Submit and publish online (if feasible)

Projects from previous classes have been:
• Published as research papers
• Featured in the New York Times
• Released as successful open source projects
**Perfopticon** Distributed Query Performance

**Physical Query Plan:**

- Fragment 2
- Fragment 3
- Fragment 1
  - ShuffleConsumer
  - ShuffleConsumer
  - SymmetricHashJoin(((t0 = 3) and (t3 = 6)) and (t2 = 5)) and (t1 = 4));
- GroupBy(t0, COUNTALL)
- ShuffleProducer(h(t0))
- Fragment 0

**Overview / Operators inside fragment 1**

**Query time contribution** collapse/expand

- Shuffle Group: SymmetricHashJoin(((t0 = 3) and (t3 = 6)) and (t2 = 5)) and (t1 = 4));
  - 3% 3% 92%

**Detailed execution**

- 128 tuples returned, took 3 μs 456 ns

**Dominik Moritz et al.** [EuroVis '15]
Visualizing Galaxy Merger Trees

S. Loebman, J. Ortiz, L. Orr, M. Balazinska, T. Quinn et al. [SIGMOD '14]
A browser-based tool for visualization and analysis of diffusion MRI data

Jason D. Yeatman, Adam Richie-Halford, Josh K. Smith, Anisha Keshavan & Ariel Rokem

[Diagram of brain and diffusion MRI data plots]
Latent Space Cartography
Visual Analysis of Vector Space Embeddings
Yang Liu, Eunice Jun, Qisheng Li (CSE 512, Spring ’18)
locations of each train on the red, blue, and orange lines at 5:46 am. Hover over the diagram to the right to display trains at a different time.

trains are on the right side of the track relative to the direction they are moving.

see the morning rush-hour, midday lull, afternoon rush-hour, and the evening lull.
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+"°").textBaseline("center");

vis.add(pv.Line).data(napoleon.temp)
  .left(lon).top(tmp)
  .text(function(d) d.temp+d.date.substr(0,6)).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 Cornabas
Daniel de Pury
Daniel Defoe
Daniel Mathus
Daniel Marc Antoine Chardon
Daniel Muler

TOP CITIES AND AUTHORS

<table>
<thead>
<tr>
<th>Letters received</th>
<th>Letters sent</th>
</tr>
</thead>
<tbody>
<tr>
<td>London</td>
<td>346</td>
</tr>
<tr>
<td>Oates</td>
<td>304</td>
</tr>
<tr>
<td>Dublin</td>
<td>208</td>
</tr>
<tr>
<td>Paris</td>
<td>238</td>
</tr>
<tr>
<td>Twickenham</td>
<td>180</td>
</tr>
<tr>
<td>John Locke</td>
<td>350</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?
A1: Visualization Design

Design a static visualization for a data set.

The climate of a place can have a tremendous impact on people's lived experience. You will examine average monthly climate measurements for six major U.S. cities, roughly covering the edges of the continental United States.

You must choose the message you want to convey. What question(s) do you want to answer? What insight do you want to communicate?
A1: Visualization Design

Pick a **guiding question**, use it to title your vis. Design a **static visualization** for that question. 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 **11:59 pm, Wednesday April 6**.