How much data (bytes) did we produce in 2018?
2018: 33 zettabytes
Up from 1.2 zettabyte (2010)

Statista 2018, Gantz et al 2010
Physical Sensors

Image courtesy cabspotting.org
<table>
<thead>
<tr>
<th>HR/EKG</th>
<th>Art mmHg sys/dia</th>
<th>SpO2 %</th>
<th>RR/CO2 1/min</th>
</tr>
</thead>
<tbody>
<tr>
<td>97</td>
<td></td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td></td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td></td>
<td>99</td>
<td></td>
</tr>
</tbody>
</table>

Heart rate and blood pressure data displayed on a medical monitor.
Records of Human Activity
Wikipedia History Flow (IBM)
The ability to take data—\textit{to understand} it, \textit{to process} it, \textit{to extract value} from it, \textit{to visualize} it, \textit{to communicate} it—\textit{that’s} going to be a hugely important skill in the next decades, \textit{...} because now we really do have \textit{essentially free} and \textit{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

\textit{The McKinsey Quarterly, Jan 2009}
A Poverty of Attention

“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?

“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></th>
<th>Set A</th>
<th>Set B</th>
<th>Set C</th>
<th>Set D</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>10  8.04</td>
<td>10  9.14</td>
<td>10  7.46</td>
<td>8   6.58</td>
</tr>
<tr>
<td>Y</td>
<td>8   6.95</td>
<td>8   8.14</td>
<td>8   6.77</td>
<td>8   5.76</td>
</tr>
<tr>
<td>X</td>
<td>13  7.58</td>
<td>13  8.74</td>
<td>13  12.74</td>
<td>8   7.71</td>
</tr>
<tr>
<td>Y</td>
<td>9   8.81</td>
<td>9   8.77</td>
<td>9   7.11</td>
<td>8   8.84</td>
</tr>
<tr>
<td></td>
<td>11  8.33</td>
<td>11  9.26</td>
<td>11  7.81</td>
<td>8   8.47</td>
</tr>
<tr>
<td></td>
<td>14  9.96</td>
<td>14  8.1</td>
<td>14  8.84</td>
<td>8   7.04</td>
</tr>
<tr>
<td></td>
<td>6   7.24</td>
<td>6   6.13</td>
<td>6   6.08</td>
<td>8   5.25</td>
</tr>
<tr>
<td>Y</td>
<td>4   4.26</td>
<td>4   3.1</td>
<td>4   5.39</td>
<td>19  12.5</td>
</tr>
<tr>
<td></td>
<td>12  10.84</td>
<td>12  9.11</td>
<td>12  8.15</td>
<td>8   5.56</td>
</tr>
<tr>
<td></td>
<td>7   4.82</td>
<td>7   7.26</td>
<td>7   6.42</td>
<td>8   7.91</td>
</tr>
<tr>
<td></td>
<td>5   5.68</td>
<td>5   4.74</td>
<td>5   5.73</td>
<td>8   6.89</td>
</tr>
</tbody>
</table>

**Summary Statistics**

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

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

**Linear Regression**

\[
Y = 3 + 0.5 \, X
\]

\[
R^2 = 0.67
\]

[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]
<table>
<thead>
<tr>
<th>Edge of Frame</th>
<th>DN</th>
<th>50-45</th>
<th>DARK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>45-40</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40-35</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30-25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25-20</td>
<td>LIGHT</td>
<td></td>
</tr>
</tbody>
</table>
Support Reasoning
## History of O-Ring Damage on SRM Field Joints

<table>
<thead>
<tr>
<th>Cross Sectional View</th>
<th>Top View</th>
<th>Clamping Location (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRM No.</td>
<td>Erosion Depth (in.)</td>
<td>Perimeter Affected (deg)</td>
</tr>
<tr>
<td>22A</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>22A</td>
<td>NONE</td>
<td>NONE</td>
</tr>
<tr>
<td>15A</td>
<td>0.010</td>
<td>154.0</td>
</tr>
<tr>
<td>15B</td>
<td>0.038</td>
<td>130.0</td>
</tr>
<tr>
<td>15B</td>
<td>0.038</td>
<td>130.0</td>
</tr>
<tr>
<td>15B</td>
<td>None</td>
<td>45.0</td>
</tr>
<tr>
<td>13B</td>
<td>0.028</td>
<td>110.0</td>
</tr>
<tr>
<td>11A</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>10A</td>
<td>0.040</td>
<td>217.0</td>
</tr>
<tr>
<td>2B</td>
<td>0.053</td>
<td>116.0</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 (30°, 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>Motor</th>
<th>MBT</th>
<th>AMB</th>
<th>O-Ring</th>
<th>Wind</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM-4</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></td>
</tr>
</tbody>
</table>
Make a Decision: Challenger
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]
Data in Context: Cholera Outbreak

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

\[ \begin{array}{c}
34 \\
\times 72 \\
\hline \\
68 \\
\hline \\
2380 \\
\hline \\
2448 \\
\end{array} \]

Time (Sec.)

- Mental: 110 seconds
- Paper & Pencil: 0 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 Dragons of Eden [Carl Sagan]
The Elements of Graphing Data
[Cleveland]
Convey Information to Others
Bones in hand [from 1918 edition]

Double helix model [Watson and Crick 53]
“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
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
Sémiologie Graphique [Bertin 67]
Visualization Design

SlicerDicers' Sales Compared to Other Products

Problematic design

Redesign
Exploratory Data Analysis
Recent elections have placed a heavy emphasis on “swing states” — Ohio, Florida and the other competitive states. You can see how many votes shifted between the Democratic and Republican parties. A look at how the states 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
Interaction

Crimespotting.org

Friday, December 12, 2008
154 reports
Graphical Perception

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.
Uncertainty
Dymaxion Maps [Fuller 46]
Animation

Animated transitions in statistical data graphics [Heer & Robertson 07]
Hierarchies

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

Zephoria

User ID: 21721
Friends: 756
Age: 27
Gender: Female
Status: Single
Location: San Francisco, CA
Hometown: Lancaster, PA
Occupation: researcher, social networks, identity, context
Interests: schizotypy, observing people, culture, questioning power, reading, Buddhism, depth psychology, computer-mediated communication, social networks, technology, anthropology, stamping
Music: post-punk/indie, Mushroom, Sun Kilu, blog/Digital Structures, Ani DiFranco, downtempo, Thievery Corporation, Beth Orton, Machete, Vase, White Stripes
Books: Various authors including: Gottman, Stanley Mignam, Jeanette Winterson, Eric Schlosser, Leslie Feinberg, Dorothy Allison, Italo Calvino, Hermann Hesse
TV Shows: ??
Movies: X-Men, X-Files, The Matrix, Clockwork Orange, American Beauty, Fight Club, Boys Don't Cry
Member Since: ??
Last Login: 2003-12-21
Last Updated: 2003-12-21
About: Some know me as danah...

I'm a geek, an activist, and an academic, fascinated by people and society. I see life as a very large playground and enjoy exploring its intricacies. I revel in life's chaos, while simultaneously providing my own insane element.

My musings: http://www.zephoria.org/thoughts

Want to Meet: Someone who makes life's comically avoidable simply elegant... A partner in crime with no agenda.
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
Matthew Conlen
mconlen@cs.washington.edu

Research on interactive documents and data-driven storytelling.

I maintain an interactive markup language called Idyll, publish a data-driven digital magazine called the Parametric Press, and work with NASA’s climate communications team.
Instructors
Matthew Conlen
PhD Student, CSE
https://mathisonian.com

cse442@cs

Assistants
Eunice Jun
Chanwut (Mick) Kittivorawong
Andrew Wang
Zhu (Ruby) Li
JL (Jialiang) Liu
Instructors
Matthew Conlen
PhD Student, CSE
https://mathisonian.com

cse442@cs

Assistants
Eunice Jun
Chanwut (Mick) Kittivorawong
Andrew Wang
Zhu (Ruby) Li
JL (Jialiang) Liu

Office hours posted tomorrow.
Eunice Jun
emjun@cs.washington.edu
OH: Friday 1-2pm GATES 152 or by appointment

Research
Languages and tools for statistical data analysis
designed for non-experts*  🌿 tea-lang.org

Ask me about...
Much experience with Python, R, backend web programming

Some experience with Javascript, D3, general web programming

Visualization + functional programming

*Interested in research? Contact me!
Chanwut (Mick) Kittivorawong
chanwutk@cs.washington.edu

Research Contributions:

Vega: Labeling automation in visualizations

Vega-Lite: Enabling syntax for creating composite marks

Technical Experience:

TypeScript, Vega-Lite/Vega, and web development
Andrew Wang

aywang@cs.washington.edu
OH: TBD

Work experience in data vis @ Microsoft Gaming
Other: data infrastructure @ Citadel, machine learning infrastructure @ Stripe

Research experience in computer vision

Experienced in JavaScript, D3, and general web programming
Zhu (Ruby) Li
liz67@cs.washington.edu

Research on interactive mapping tool for accessibility

Experience with JavaScript, D3, prototyping, and general web programming

Interest in perceptual and cognitive psychology
JL (Jialiang) Liu

jl262@cs.washington.edu

Love making interactive and meaningful websites.

Taken INFO and HCDE courses related to design.

Experience with JavaScript, D3, and general web programming.
Interactive Data Visualization for the Web, 2nd Edition

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

We will be using D3 v5.
https://d3js.org
Some from D3 book, others from papers & web. Material in class will loosely follow readings. Readings should be read by start of class. Post discussion comments on class Canvas forum. One comment per week (ending week 8). Comments posted by Friday 11:59pm. You have 1 “pass” for the quarter.
Assignments

Class Participation (10%)
A1: Visualization Design (10%)
A2: Exploratory Data Analysis (15%)
A3: Interactive Prototype (25%)
  Peer Evaluation
FP: Final Project (40%)
  Initial Prototypes
  Project Deliverables
Final Project

Produce interactive web-based visualizations
Initial prototype and design review
Final deliverables and video presentation
Submit and publish on GitHub
Projects from previous classes have been:
• Published as research papers
• Featured in the New York Times
• Released as successful open source projects
Final Project Theme

Interactive dashboard for a public audience.

Goal: find data of public interest, design visualizations to explore and communicate it effectively.

- Politics
- Sports
- Climate
- ...

You must identify a topic, dataset, and target audience. We’ll provide some examples and potential datasets.
Inspiration...
Change In Times (CSE 442, Spring 2017)
Gunnar Olson, Halden Lin, Lilian Liang, and Shobhbit Hathi
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**.

Service starts at 5AM on Monday morning. Each line represents the path of one train. Time continues downward, so steeper lines indicate slower trains.

Since the red line splits, we show the Ashmont branch first then the Braintree branch. "Jump over" the Ashmont branch.

Train frequency increases around 6:30AM as morning rush hour begins.
Visualizing Galaxy Merger Trees

S. Loebman, J. Ortiz, L. Orr, M. Balazinska, T. Quinn et al. [SIGMOD '14]
Visualizing the Republic of Letters

Daniel Chang, Yuankai Ge, Shiwei Song

Republic of Letters
1700

FILTER BY AUTHOR
Clear All
Damien Desormes
Daniel Cornabs
Daniel de Pury
Daniel Defoe
Daniel Mathus
Daniel Marc Antoine Chardon
Daniel Muller

TOP CITIES AND AUTHORS

<table>
<thead>
<tr>
<th>City</th>
<th>Letters received</th>
<th>Letters sent</th>
</tr>
</thead>
<tbody>
<tr>
<td>London</td>
<td>346</td>
<td>357</td>
</tr>
<tr>
<td>Oates</td>
<td>304</td>
<td>250</td>
</tr>
<tr>
<td>Dublin</td>
<td>208</td>
<td>154</td>
</tr>
<tr>
<td>Paris</td>
<td>238</td>
<td>112</td>
</tr>
<tr>
<td>Twickenham</td>
<td>18</td>
<td>101</td>
</tr>
<tr>
<td>John Locke</td>
<td>350</td>
<td>253</td>
</tr>
<tr>
<td>Joseph Addison</td>
<td>30</td>
<td>244</td>
</tr>
<tr>
<td>Voltaire</td>
<td>26</td>
<td>231</td>
</tr>
<tr>
<td>Jonathan Swift</td>
<td>85</td>
<td>159</td>
</tr>
<tr>
<td>Alexander Pope</td>
<td>28</td>
<td>150</td>
</tr>
</tbody>
</table>
Questions?
A1: Visualization Design

Design a static visualization for a data set.
Every 10 years, the census documents the demographic make-up of the U.S., influencing congressional districting and social services. This dataset contains a summary of census data for two years a century apart: 1900 and 2000.

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 ($\leq$ 4 paragraphs)

Due by **11:59 pm, Tues Jan 14**.