Today

- **Before class**
  - Reading: Ch 1 - What's Vis, and Why Do It?
  - Reading: "What I Learned Recreating One Chart Using 24 Tools"

- **During class**
  - Highlight and discuss Ch 1
  - *We will not cover everything that you are responsible for during class time.*
What is visualization?

"The communication of information using graphical representations"
- Ward, Grinstein, Keim

"The use of computer-supported interactive visual representations of data to amplify cognition"
- Card, Mackinlay, Shneiderman, *Readings in Information Visualization: Using Vision to Think*

"The purpose of visualization is insight, not pictures."
- Ben Shneiderman
Where have you seen a visualization today?

http://hamptonroads.com/weather
http://hamptonroads.com/traffic

What's vis?

- Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- The design space of possible vis idioms is huge, and includes the considerations of both how to create and how to interact with visual representations.

- Vis design is full of tradeoffs, and most possibilities in the design space are ineffective for a particular task, so validating the effectiveness of a design is both necessary and difficult.

- Vis designers must take into account three very different kinds of resource limitations: those of computers, of humans, and of displays.

- Vis usage can be analyzed in terms of why the user needs it, what data is shown, and how the idiom is designed.
Why have a human in the loop?

- Vis allows people to analyze data when they don't know exactly what questions to ask in advance.

- Best path - put a human in the loop
  - exploit the pattern detection properties of human vision

Humans are great at pattern recognition

Create visualizations that let computers do what computers do well and lets humans do what humans do well.
Uses of vis tools

- Transitional
- Long-term
- Presentation

Why have a computer in the loop?
Why use an external representation?

- Vis allows people to offload cognition and memory usage to make space for other operations.

- Diagrams as external representations
  - information can be organized by spatial location
    - search - grouping items needed for problem-solving in one location
    - recognition - grouping relevant info for one item in the same location

Visualization can extend your memory

What is $57 \times 48$?

<table>
<thead>
<tr>
<th>paper</th>
<th>mental buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{57}{\underline{9}}$</td>
<td>$[8 \times 7 = 56]$</td>
</tr>
<tr>
<td>$\times 48$</td>
<td>$[8 \times 5 = 40 + 5 = 45]$</td>
</tr>
<tr>
<td>$1$</td>
<td>$[4 \times 7 = 28]$</td>
</tr>
<tr>
<td>456</td>
<td>$[4 \times 5 = 20 + 2 = 22]$</td>
</tr>
<tr>
<td>228</td>
<td>$[8 + 5 = 13]$</td>
</tr>
<tr>
<td>$\underline{2736}$</td>
<td>$[4 + 2 + 1 = 7]$</td>
</tr>
</tbody>
</table>

Example courtesy Tamara Munzner, Univ. of British Columbia
Why depend on vision?

- Visual system provides a high-bandwidth channel to our brains.
- Significant amount of visual information processing occurs in parallel at the pre-conscious level.

Can you find the red dot?

preattentive processing
Which state had the highest marriage rate?

U.S. Census Bureau, Statistical Abstract of the United States: 2012
http://www.census.gov/compendia/statatab2012/tables/12s0133.pdf,
http://www.census.gov/compendia/statatab2012/tables/12s0133.xls
Why show the data in detail?

- Vis tools can allow people to explore data to find patterns or to determine if a statistical model actually fits the data.

- Look out for questionable data
  - "just because it's numbers doesn't mean it's true"
  - is it a typo or something interesting?
    - "make sure you know which one it is"

---

### Anscombe's Quartet

<table>
<thead>
<tr>
<th>Anscombe’s Quartet: Raw Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>I</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Munzner, Figure 1.3
The four data sets are not the same

"Graphics reveal data"
- Edward Tufte, *The Visual Display of Quantitative Information*


Why use interactivity?

- The *design space of possible vis idioms is huge*, and includes the considerations of both *how to create* and *how to interact* with visual representations.

- Interaction allows for
  - handling complexity
  - displaying multiple aspects of a dataset
Why is the idiom design space huge?

- The design space of possible vis idioms is huge, and includes the considerations of both how to create and how to interact with visual representations.

- Vis idioms - approaches to creating and manipulating visual representations

- Simple examples: scatterplots, bar charts, line charts

Why focus on tasks?

- Vis usage can be analyzed in terms of why the user needs it, what data is shown, and how the idiom is designed.

- The intended task is just as important as the data to be visualized.

- Four categories of tasks
  - presentation
  - discovery
  - enjoyment of information
  - producing more information for later use
Why focus on effectiveness?

- Vis design is full of tradeoffs, and most possibilities in the design space are ineffective for a particular task, so validating the effectiveness of a design is both necessary and difficult.

- Effectiveness is an important measure for understanding if the user task was supported.
  - "The purpose of visualization is insight, not pictures." - Ben Shneiderman

- But, no picture can tell the truth, the whole truth, and nothing but the truth.

Why are most designs ineffective?

- Vis design is full of tradeoffs, and most possibilities in the design space are ineffective for a particular task, so validating the effectiveness of a design is both necessary and difficult.

- Design may not match with human perception
- Design may not match with intended task
### Why is validation difficult?

- Vis design is full of **tradeoffs**, and **most possibilities in the design space are ineffective** for a particular task, so validating the **effectiveness** of a design is both necessary and difficult.

- There are a ton of questions that could be asked.
  - what does "better" mean?
  - what does "effectively" mean?
  - how to measure insight?
  - who should perform the task to be measured?
Why are there resource limitations?

- Vis designers must take into account three very different kinds of **resource limitations**: those of computers, of humans, and of displays.

  - computational capacity
  - human perceptual and cognitive capacity
  - display capacity

Why analyze vis?

- Vis usage can be analyzed in terms of why the user needs it, what data is shown, and how the idiom is designed.

- Analyzing existing systems is a good stepping stone to designing new ones.

- High-level framework for analyzing vis use
  - *what* data the user sees
  - *why* the user intends to use a vis tool
  - *how* the visual encoding and interaction idioms are constructed in terms of design choices
Tools

Tools of the Trade

http://selection.datavisualization.ch/
Excel

http://chandoo.org/wp/2008/09/03/6-charts-to-never-use/

http://www.juiceanalytics.com/writing/recreating-ny-times-cancer-graph/

---

R

http://www.r-project.org

---

Edgar Anderson’s Iris Data

---

CS 725/825 - Spring 2017 - Weigle
Tableau

http://www.tableausoftware.com

Seattle Real Estate: Overview

Tableau

http://www.tableausoftware.com

D3

http://d3js.org

Data-Driven Documents
Many, many others

- "What I Learned Recreating One Chart Using 24 Tools" - *assigned reading*


---

History

---
~1000 - Al-Biruni's Moon Phases


1854 - John Snow's Cholera Map
1858 - Nightingale's Diagram

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

The blue wedges measured from the centre of the circle represent areas for due, the deaths from Preventible or Metabolic Syphilitic diseases, the red wedges measured from the centre the deaths from amputations & the black wedges measured from the centre the deaths from all other causes.

The black line across the red triangles in May 1854 marks the boundary of the deaths from all other causes during the months 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 excluding them.

1869 - Minard's Map

1933 - London Subway
1990s - Web Vis


2009 - Hans Rosling's Bubble Charts

http://www.gapminder.org/videos/200-years-that-changed-the-world-bbc/

Full TED talk (20 min):
2006, 2011 - Hierarchical Edge Bundling


---

2010 - New York Times Porcupine Chart

Budget Forecasts, Compared With Reality

Just two years ago, surpluses were predicted by 2012. How accurate have past White House budget forecasts been?

Topic Objectives

- Define visualization.
- Explain the importance of humans in the visualization process.
- Explain why human vision is particularly well-suited for information transfer.
- Give an example of a visualization idiom.
- Explain why it is best to consider multiple alternatives for vis before selecting a solution.
- Explain at a high-level the "what-why-how" framework for analyzing visualization use.
- Describe at least one historical visualization and explain its impact.
- Differentiate between R, D3, and Tableau and describe the type of tasks for which each tool might be most appropriate.