CSE 564
Visualization & Visual Analytics

INFOGRAPHICS DESIGN

Klaus Mueller

Computer Science Department
Stony Brook University
<table>
<thead>
<tr>
<th>Lecture</th>
<th>Topic</th>
<th>Projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intro, schedule, and logistics</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Applications of visual analytics, basic tasks, data types</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Introduction to D3, basic vis techniques for non-spatial data</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Data assimilation and preparation</td>
<td>Project #1 out</td>
</tr>
<tr>
<td>5</td>
<td>Data assimilation and preparation</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Bias in visualization</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Data reduction and dimension reduction</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Visual perception</td>
<td>Project #2(a) out</td>
</tr>
<tr>
<td>9</td>
<td>Visual cognition</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Visual design and aesthetics</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Cluster analysis: numerical data</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Cluster analysis: categorical data</td>
<td>Project #2(b) out</td>
</tr>
<tr>
<td>13</td>
<td>High-dimensional data visualization</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Dimensionality reduction and embedding methods</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Principles of interaction</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Midterm #1</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Visual analytics</td>
<td>Final project proposal call out</td>
</tr>
<tr>
<td>18</td>
<td>The visual sense making process</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Maps</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Visualization of hierarchies</td>
<td>Final project proposal due</td>
</tr>
<tr>
<td>21</td>
<td>Visualization of time-varying and time-series data</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Foundations of scientific and medical visualization</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Volume rendering</td>
<td>Project 3 out</td>
</tr>
<tr>
<td>24</td>
<td>Scientific and medical visualization</td>
<td>Final Project preliminary report due</td>
</tr>
<tr>
<td>25</td>
<td>Visual analytics system design and evaluation</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Memorable visualization and embellishments</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Infographics design</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Midterm #2</td>
<td></td>
</tr>
</tbody>
</table>
A clipped compound of "information" and "graphics"

A graphic visual representations of information, data or knowledge intended to present information
- quickly
- clearly

Can improve cognition by utilizing graphics to enhance the human visual system's ability to see patterns and trends
Graphical displays should

- show the data
- induce the viewer to think about the substance rather than about methodology, graphic design, the technology of graphic production, or something else
- avoid distorting what the data has to say
- present many numbers in a small space
- make large data sets coherent
- encourage the eye to compare different pieces of data
- reveal the data at several levels of detail, from a broad overview to the fine structure
- serve a reasonably clear purpose: description, exploration, tabulation, or decoration
- be closely integrated with the statistical and verbal descriptions of a data set.
Minard’s 1869 map of Napoleon’s campaign to Russia

Shows 6 variables in one 2D visualization:
- number of Napoleon's troops, temperature
- distance traveled, direction of travel
- latitude and longitude, location relative to specific dates
Some More Random Examples ...
The Time We Spend on the Internet

Almost 6 billion adults are connected to the Internet, with an average viewing time of 40 minutes.

We send over 200 million emails per minute.

...which includes time spent online at work.

Stress in the Workplace

- 70% of American workers experience stress-related illnesses.
- 34% think they'll burn out on the job in the next two years.
- 33% more heart attacks are occurring on Monday mornings.
- 45% of entrepreneurs said they were stressed.

The Average Work Week

- You get 168 hours a week.
- The average work week is closer to 50 hours than 40.
- The average commute time is 47 minutes round trip.
- Professionals and business leaders who use mobile devices like smartphones.
WHAT ARE WE EATING?
What the Average American Consumes in a Year

The Average American
Age: 36.6
Height: 5'9" (m) 5'4" (f)
Weight: 190 lbs (m) 164 lbs (f)

Fats & Oils 86.5 lbs
Red Meat 110 lbs
Poultry 73.6 lbs
Fish & Shellfish 16.1 lbs
Eggs 37.2 lbs
Cheese 31.4 lbs
Dairy Products (non-cheese) 600.5 lbs
Coffee, Cocoa & Nuts 24 lbs
Vegetables 415.4 lbs
Wheat flour 15.6 lbs
Beverage Milks 13.1 lbs
Caloric Sweeteners

*1,996.3 lbs*
EDUCATION AROUND THE WORLD

There are 1.4 Billion students on Earth.

Only 65.2 Million educators Globally.

THE CHALLENGE: Too many children remain out of school, and those who are in school aren’t learning the skills they need for life and work.

Children leaving school before completing their Primary Education

- In the Sub-Saharan, 11.07 million children leave school before completing their primary education.
- In South and West Asia, that number reaches 13.54 million.

Children out of primary school

- 61 Million children are still out of primary school.
- 32 Million of these children are Girls.

1 in 5

- 15 to 24 years old has not completed primary school and lacks skills for work.

An estimated 250 million children are not able to read or write.

The poorest and most marginalized are the most affected.

In some emerging economies, 3 in 10 youths cannot do basic arithmetic.

Fragile and conflict-affected countries account for more than 30% of all children not completing primary school.

OF THE 775 MILLION ILLITERATE ADULTS

- TWO-THIRDS ARE WOMEN

Connect. Inspire. Empower.

www.ubuntuemail.com
**STUDENT BULLYING**

**Effects of Bullying**
Bullying can have a significant impact on both child and teenage students. Students who are bullied often suffer from anxiety, fear, withdrawal, low self-esteem, and poor concentration. A bullied student will often avoid school, have lower grades, and become socially isolated. There have been numerous reports of suicide due to bullying.

**Effects on Bullies**
Bullying is often a warning sign that children and teens are heading for trouble and are at risk for serious violence. Teens (particularly boys) who bully are more likely to engage in other antisocial/inappropriate behavior (e.g., vandalism, shoplifting, truancy, and drug use) into adulthood. They are four times more likely than nonbullies to be convicted of crimes by age 24, with 60 percent of bullies having at least one criminal conviction.

**Most Common Types of Bullying**
- Hitting
- Threatening
- Intimidating
- Maliciously teasing and taunting
- Name-calling
- Making sexual remarks
- Stealing or damaging belongings
- More subtle, indirect attacks (such as spreading rumors or encouraging others to reject or exclude someone)

**Playground School Bullying Statistics**
- Every 7 minutes: 43% fear harassment in the bathroom at school
- Every 7 minutes: 280,000 students are physically attacked in secondary schools each month
- Every 7 minutes: 160,000 students miss school each day for fear of being bullied

**Sources**
- [www.coeps.usdoj.gov](http://www.coeps.usdoj.gov)
- [www.keepphysicalsafe.org](http://www.keepphysicalsafe.org)
- [www.billypolice.org](http://www.billypolice.org)

**ADULT INTERVENTION: 11% | PEER INTERVENTION: 4% | NO INTERVENTION: 85%**

**Worst States to Live In For Bullying K-12**
1 out of 4 students will be abused by another youth
1 out of 5 admit to being a bully, or doing some “bullying”
Five steps (from Venngage.com)

1. Outline the goals of your infographic
2. Collect data for your infographic
3. Visualize the data for your infographic
4. Layout your infographic using an infographic template
5. Add style to your infographic design
Step 1: Outline Your Goals

Use the question pyramid

1 Burning Problem
The main question the infographic will answer.

2-3 Supporting Questions
Questions to provide information (what, which).

1-2 Probing Questions
Questions to provide insight (why).

Venngage.com
EXAMPLE: FOOD – WINE PAIRING

Burning Problem
How to pair food and wine

Supporting Questions
Which food components matter?
What are the types/categories of wine?
Which category works with each component?

Probing Questions
Why do certain pairings work better?
We discussed this
Decide what are your primary goals
- what data aspect do you want to convey

The ICORE method
- Inform,
- Compare,
- Change,
- Organize,
- Reveal relationships, or
- Explore.

Venngage.com
Convey an important message or data point that doesn’t require much context to understand:

Make a numerical stat stand out with large, bold, colorful text:

The brain makes **700** neural connections per second before the age of 5.

Demand extra attention by pairing icons with text:

**4200**
VACCINES ADMINISTERED

Highlight a percentage or rate with a donut chart or a pictograph:

- **25%** were children
- **25%** were children

Venngage.com
Show similarities or differences among values or parts of a whole.

- **Pie Chart**: Use a pie chart, donut chart, pictograph, or tree map to compare parts of a whole.
- **Stacked Bar Chart**: Use a stacked bar chart or stacked column chart to compare categories and parts of a whole.
- **Stacked Area Chart**: Use a stacked area chart to compare trends over time.

Venngage.com
Show trends over time or space

Use a line chart or an area chart to show changes that are continuous over time.

**LINE CHART**

**AREA CHART**

Use a timeline to show discrete events in time.

**TIMELINE**

Use a choropleth map to show spatial data.

**CHOROPLETH MAP**

Use a map series to show data that changes over both space and time.

**MAP SERIES**
Organize

Show groups, patterns, rank or order

Use a list to show rank or order when you want to provide extra information about each element.

<table>
<thead>
<tr>
<th>LIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

Use a table to show rank or order when you want readers to be able to look up specific values.

<table>
<thead>
<tr>
<th>TABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Cars</td>
</tr>
<tr>
<td>BMW</td>
</tr>
<tr>
<td>Aston Martin</td>
</tr>
<tr>
<td>Bentley</td>
</tr>
</tbody>
</table>

Use a flowchart to show order in a process.

Otherwise, show rank or order with a bar chart, column chart, bubble chart, or pyramid chart.
Reveal more complex relationships among things

Use a scatter plot when you want to display two variables for a set of data.

Use a multi-series plot when you want to compare multiple sets of related data.
Add interaction

- filtering, sorting, and drilling down
- can’t do with a static chart
- but useful for online infographics displays
- will discuss later
Create a natural flow

- could use the question pyramid to guide the layout
- burning question into header
- follow with charts to address the supporting questions
- finish with the probing questions

Use a grid layout

- guides the reader’s eye
- symmetrical grid
- consider that people read
  - top to bottom
  - left to right
One Column for Linear Flow
Two Columns for Comparisons
Who Said What?

Democrats

- Character: 11
- Jobs: 14
- War: 9
- Hurricane: 8
- Change: 24
- McCain: 18
- God: 6
- Bush: 18
- Economy: 16
- Iraq: 12

Republican

- God: 25
- Character: 14
- Energy: 18
- War: 18
- Bush: 6
- Economy: 18
- Terror: 12
- Jobs: 8
- Iraq: 15
- Tax(es): 22

The Words They Used - Graphic - NYTimes.com
Three Column Example

CLean Water Systems

2.64 Billion people around the world do not have access to adequate sanitation

780 Million people do not have access to improved water sources

Every $1 spent on water and sanitation generates at least $4 in increased productivity
Multi-Section Layout For Random Visual Access
Overall goal
- make your infographics aesthetically pleasing
- make it easy to consume and understand

Minimize text
- should supplement the visuals
- short paragraphs (at most) at about a grade six reading level

Use font to point out importance
- readable font for the bulk of the text,
- amp up size and style of your main header, section headers, and data highlights
- make sure that the gist of your infographic is immediately apparent
Example

Reducing Poverty
We’re active in over 100 communities in Georgia alone. And we serve over 4 million Georgians each year. In 2017, the total investment amount was $152.7 million.

Find out how you can make an impact today
Repeat basic shapes to reinforce the underlying grid
Emphasize Using Shapes

Use basic shapes to emphasize headers and list elements.
Add icons to reinforce important concepts in the text

- keep icon color, style, and size consistent
- pull everything together with extra background shapes
BE Diligent

Make sure the repeated elements you add are aligned!
Negative (white) space is often used in advertising
- the space around and between the subject of an image
- helps to define a subject
- helps emphasizing a message
Just as important as any other element of a design

- simplest negative space are margins around the graphics

Venngage.com
Use gaps between unrelated elements

- makes it easier to immediately understand which elements are grouped together
Much easier to make colors play nicely with each other by adding a little whitespace.
Use color to highlight important information
COLOR FOR GROUPING

Use color to group related elements together
Neutral Colors

Use neutrals to offset bright colors

- helps put together a composition
- can group in a more subtle way

Venngage.com
There are Templates on the Web
Color Harmony

Non-harmonic colors

Harmonic colors

Hue wheel:
Harmonic Color Schemes

- i type
- V type
- L type
- I type
- T type
- Y type
- X type
- N type
Given arbitrary hue histogram $H(p)$ for image $X$, find the closest harmonic template $T_m$

- minimize the distance of the histogram to template coverage (delineated by template edges $E$)
- use an optimization procedure for this
- also find the orientation angle $\alpha$

$$F(X(m, \alpha)) = \sum_{p \in X} \| H(p) - E_{T_m(\alpha)}(p) \| \cdot S(p)$$

from Cohen ‘08
Given closest template and $\alpha$ has been found (user may specify other template)

- shift all hues $H(p)$ to the closest harmonic template position $H'(p)$ with width $w$
- a Gaussian $G$ controls the clustering of the hues around the sector mean $C$ of the template (greater $\sigma$ clusters more, we use $w/2$)

$$H'(p) = G(p) + \frac{w}{2} \left(1 - G_{\sigma}(||H(p) - C(p)||)\right)$$

This may break up coherent regions into disjointly colored regions

- to avoid this, may embed a graph-cut labeling into the shifting procedure

from Cohen ‘08
Color Harmonization: Example

Collage harmonization (from Cohen ’06):

non-harmonic  harmonized (T type)
COLOR HARMONIZATION: EXAMPLE

Collage harmonization (from Wang ‘08):