IAT 814
Visualization

Getting to Design

Lyn Bartram
Administrivia: Assignment 1

- Visualization in practice
- Find one good and one bad example
- Put on Canvas discussion
- Present BRIEFLY in class
- We’ll need to run over
- We will revisit these choices later!
WHEN YOU WANT TO MEET WITH ME

https://connect.sfu.ca/home/lyn@sfu.ca?fmt=freebusy
The desired progression

http://www.systems-thinking.org/kmgmt/kmgmt.htm
Another way of looking at it ....
The value of visualization

• Capture information
  • Blueprints, photographs, sensors, seismographs, maps… metadata!

• Analyse data to support reasoning
  • Develop and test hypotheses
  • Discover errors
  • Find patterns
  • Expand memory

• Communicate
  • Share, persuade, educate
Excel

Information Visualization: Examples
SCHOOL OF INTERACTIVE ARTS & TECHNOLOGY PRE-REQUISITE MAP

MEDIA ARTS

4th
- IAT 443
- IAT 445
- IAT 344
- IAT 343

3rd
- IAT 203
- MATH 242
- IAT 267

INTERACTIVE SYSTEMS

4th
- IAT 265
- IAT 267

3rd
- IAT 201
- IAT 265

BACHELOR OF ARTS (BA)

REQUIRED:
- 309 - WRITING FOR IAT

BACHELOR OF SCIENCE (BSc)

1st
- IAT 100
- CMPT 166
- CMPT 166
- MATH 130

2nd
- IAT 202
- IAT 102

3rd
- IAT 235
- IAT 233

4th
- IAT 236
- IAT 237

Information Visualization: Examples
SARS

---

**FIGURE 1. Chain of transmission among guests at Hotel M — Hong Kong, 2003**

- **Hospital 2** Hong Kong
  - 2 family members

- **Hospital 3** Hong Kong
  - 156 close contacts of HCWs and patients
  - 3 HCWs
  - 99 HCWs (includes 17 medical students)

- **Hospital 4** Hong Kong
  - 28 HCWs
  - 4 other Hong Kong Hospitals

- **Hotel M** Hong Kong
  - **A**
    - 4 HCWs
    - 2 close contacts
    - 10 HCWs
  - **B**
    - 0 HCWs
  - **C**
    - 37 HCWs
  - **D**
    - 34 HCWs
  - **E**
    - 37 close contacts

- **Guangdong Province, China**
  - **A**
    - 4 family members

- **Canada**
  - **F**
    - **G**
      - **H**
        - **I**
          - **K**

- **Ireland**
  - **L**

- **United States**
  - **M**

- **Germany**

---

*Health-care workers*

- All guests except G and K stayed on the 9th floor of the hotel. Guest G stayed on the 14th floor, and Guest K stayed on the 11th floor.
- Guests L and M (spouses) were not at Hotel M during the same time as index Guest A but were at the hotel during the same times as Guests G, H, and I, who were ill during this period.

---

*Information Visualization: Examples*
## Train Schedule

**Train Information Visualization: Examples**

### Train Schedule Table

<table>
<thead>
<tr>
<th>Train Name</th>
<th>Silver Meteor</th>
<th>Crescent</th>
<th>Silver Star</th>
<th>Carolinian</th>
<th>Palmetto</th>
<th>Piedmont</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Train Number</strong></td>
<td><strong>98</strong></td>
<td><strong>20</strong></td>
<td><strong>92</strong></td>
<td><strong>80</strong></td>
<td><strong>90</strong></td>
<td><strong>74</strong></td>
</tr>
<tr>
<td><strong>Days of Operation</strong></td>
<td>Daily</td>
<td>Daily</td>
<td>Daily</td>
<td>Daily</td>
<td>Daily</td>
<td>Daily</td>
</tr>
</tbody>
</table>

**On Board Service**

<table>
<thead>
<tr>
<th>Connecting Train Number</th>
<th>R172/R174</th>
<th>R171/R176</th>
<th>R114/R116</th>
<th>E156</th>
<th>E158</th>
</tr>
</thead>
</table>

**Read Up**

<table>
<thead>
<tr>
<th>R17</th>
<th>R7</th>
<th>R10</th>
<th>R11</th>
<th>R15</th>
<th>R17</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Mile</th>
<th>Cp</th>
<th>Connecting Train Number</th>
<th>On Board Service</th>
</tr>
</thead>
</table>

### Route Information

- **Route Schedule**

<table>
<thead>
<tr>
<th>Section</th>
<th>Mile</th>
<th>Cp</th>
<th>Connecting Train Number</th>
<th>On Board Service</th>
</tr>
</thead>
</table>

**Continuation of service to/from Florida continues at right**
Nightingale’s Coxcomb

Causes of Mortality in the Army in the East
April, 1854 to March 1855

From: F. Nightingale, "Notes on Matters Affecting the Health, Efficiency and Hospital Administration of the British Army", 1858
Napoleon’s Invasion of Russia

Carte Figurative des ponts successifs et hommes de l'Armée Française dans la campagne de Russie 1812-1813.

TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.
Mediaeval Europe
Tabula Peutingeriana, road map of ancient Rome
• Portray data, usually abstract data

• Use visual features to represent properties, quantities, attributes
  • Explicitly
  • derived

• Give rise to *emergent features*
Review: why it helps

- Mean of the x values = 9.0
- Mean of the y values = 7.5
- Equation of the least-squared regression line: $y = 3 + 0.5x$
- Sums of squared errors (about the mean) = 110.0
- Regression sums of squared errors (variance accounted for by $x$) = 27.5
- Residual sums of squared errors (about the regression line) = 13.75
- Correlation coefficient = 0.82
- Coefficient of determination = 0.67
What the data look like …
The process of information visualization. Graphically encoded data is viewed in order to form a mental model of that data.
Recall: Why do we 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
Organize and capture information
Combine different data sources
Answer (and discover) questions

- Where do the most news stories originate?
- Inspire!

Newman & Gastner, PNAS
**Train schedule**

<table>
<thead>
<tr>
<th>PIEDMONT</th>
<th>PALMETTO</th>
<th>CAROLINA</th>
<th>SILVER STAR</th>
<th>SILVER SKY</th>
<th>CROISANT</th>
<th>TRAIN NUMBER</th>
<th>COMING/GOING</th>
<th>AIR</th>
<th>DAY</th>
<th>TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>73</td>
<td>89</td>
<td>79</td>
<td>91</td>
<td>97</td>
<td>19</td>
<td>98</td>
<td>Daily</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Information Visualization: Examples**

- Make decisions/expand memory

---

*Continuation of service to/from Florida continues at right*
How educated are world leaders?

In democracies, the driver of change is typically elections; in autocracies, it is typically biology. It is often said that autocratic leaders are therefore less likely to be educated. Leaders' education level has been found to be a positive predictor of, among other things, economic growth. Here's a look at the tenure of around 700 leaders sorted by political longevity in the world's autocratic and democratic countries, along with a selection from present-day democracies, since 1950 or thereabouts. For each country, the leader who has ruled the longest, continuously, is shown. Hover over a bar for more details.

Cuba
Held Cuba
1959–2006 (47 years)
Minimum term limit: none
Median time in office: 1 year (approx)

Time in office over time

Data give a sense of the number of leaders; zoom over for more info.

Here's a look at the number of leaders with an advanced degree, e.g., Master's or Doctorate, or equivalent - such as certain law degrees in some countries - for each of the above countries during the same time period. Lack of education is not necessarily indicative of ineptness, though it may be. It may also be the result of remaining life circumstances or lack of means, as in the case of former British Prime Minister James Callaghan, for example.

Cuba
6/6 have held a graduate degree or equivalent
0 out of 6 leaders

Education level over time
find patterns, compare

Information Visualization: Examples
Find patterns, expand calculations, find questions
Inspire and compel/tell a story

Gun ownership in Westchester County, USA
Image courtesy of S. Few, www.perceptualedge.com
Image courtesy of S. Few, [www.perceptualedge.com](http://www.perceptualedge.com)
Using this advanced method I was able to confirm my initial assessment that Idea C would be our best bet.
What’s really going on here?

• The purpose of visualization is INSIGHT not IMAGES (Stasko)
• Cognitive process of building a mental image and model and internalising understanding

“The use of computer-supported, interactive visual representations of data to amplify cognition.” [Card, Mackinlay Shneiderman ‘98]
Distributed Cognition

• Cognitive system is composed of people and the artifacts they use
• Cognition isn’t only internal
• Changes in external representation spur changes in internal representation and understanding
• It is interaction with the external representations that drives this process
Recap: how vis amplifies cognition

- Increasing memory and processing resources available
- Reducing search for information
- Enhancing the recognition of patterns
- Enabling perceptual inference operations
- Using perceptual attention mechanisms for monitoring
- Encoding info in a manipulable medium
Knowledge Crystallization
More simply

Acquire → Parse → Filter → Mine/Prune → Represent → Refine → Interact
Visualization Stages

Example: house real estate listing data

- Price
- Bedrooms
- Lot size
- Type
- etc
Visualization Stages

Data transformation – create a visual spatial model

- **Data transformation**
  - Map raw data into data tables – e.g. text to similarity matrix
Visualization Stages

Visual mapping—create a visual spatial model

- Data transformation
  - Map raw data into data tables – e.g. text to similarity matrix
- Visual Mappings:
  - Transform data tables into visual structures – e.g., house price, #bedrooms to 2 dims – x, y
Visualization Stages
Display the data that now have visual form

- Data transformation
  - Map raw data into data tables – e.g. text to similarity matrix
- Visual Mappings:
  - Transform data tables into visual structures – e.g. 2 dims – x, y
- View Transformations:
  - Create views of the Visual Structures by specifying graphical parameters such as position, scaling, and clipping
Visualization Stages

The user may change transformations and mappings

- Data transformation
  - Map raw data into data tables – e.g. text to similarity matrix
- Visual Mappings:
  - Transform data tables into visual structures – e.g. 2 dims – x, y
- View Transformations:
  - Create views of the Visual Structures by specifying graphical parameters such as position, scaling, and clipping
How do we find the right design?

- 4 levels of design [Munzner 2014]
- Validate against the right “threat”

**problem:** you misunderstood their needs

**abstraction:** you’re showing them the wrong thing

**encoding:** the way you show it doesn’t work

**algorithm:** your code is too slow
Solve the right problem

• Characterise the domain problem
  • identify a problem amenable to vis
  • provide novel capabilities
  • speed up existing workflow

• validation
  • immediate: interview and observe target users
  • downstream: notice adoption rates
• abstract from domain-specific to generic operations/tasks
  • sorting, filtering, browsing, comparing, finding trend/outlier, characterizing distributions, finding correlation
• data types
  • tables of numbers, relational networks, spatial data
  • transform into useful configuration: derived data model
  • more next time
• validation
  • deploy in the field and observe usage
• visual encoding: drawings they are shown
• interaction: how they manipulate drawings
• validation
  • immediate: careful justification wrt known principles
  • downstream: qualitative or quantitative analysis of results
  • downstream: lab study measuring time/error on given task
Validation

- Can’t necessarily answer “Is it worth it?”
- Can try to address “where is it failing”?
- Evaluation is a hard problem, more later
Human in the loop stages

- Threat: wrong problem
  - Validate: observe and interview target users
- Threat: bad data/operation abstraction
  - Threat: ineffective encoding/interaction technique
  - Validate: justify encoding/interaction design
  - Threat: slow algorithm
  - Validate: analyze computational complexity
  - Implement system
  - Validate: measure system time/memory
  - Validate: qualitative/quantitative result image analysis
  - [Informal usability study]
  - Validate: lab study, measure human time/errors for operation
  - Validate: field study, document human usage of deployed system
  - Validate: collect anecdotes about tool utility from target users
  - Validate: observe adoption rates
Matrix Explorer case study

- **domain**: social network analysis
  - early: participatory design to generate requirements
  - later: qualitative observations of tool use by target users

- **techniques**
  - interactively map attributes to visual variables
    - user can change visual encoding on the y axis
  - filtering
  - selection
  - sorting by attribute
requirements

- use multiple representations
- handle multiple connected components
- provide overviews
- display general dataset info
- use attributes to create multiple views
- display basic and derived attributes
- minimize parameter tuning
- allow manual finetuning of automatic layout
- provide visible reminders of filtered-out data
- support multiple clusterings, including manual
- support outlier discovery
- find where consensus between different clusterings
- aggregate, but provide full detail on demand
Techniques: dual views
• overviews: matrix, node-link, connected components
• details: matrix, node-link
• controls

Data Sketching exercise

- Data set 1: Titanic casualties
- Cabin class, Age, Gender, Survived, Survival

- What kinds of insights might you seek from these data?
- How would you represent it?
Data sketching exercise

Life Expectancy data

• #physicians/per capita, country, #Tvs/capita

Add in –

• Change in survival rates over time (20 years)
• Change in physicians
More speculation

- Highway deaths on the Labour Day weekend were much higher than anticipated in BC. You want to find an explanation.

- Decide what kinds of data you need
- Sketch how you would visualize it