Information Visualization Design Process

Ronald Metoyer
Associate Professor
Oregon State University
Infovis Reference Model

Design pattern for visualization systems
Focus on separating data, representation, presentation
So...how do we get there!
Information Visualizations are ...

• Interactive interfaces to data
• Software artifacts

Subject to a development cycle/process like any other piece of software

Design the right Vis: Matches user’s needs
Design the Vis right: It’s appropriate for the data/task at hand
Software Development Cycles/Processes
Tamara Munzner’s Nested Model

domain problem characterization

data/operation abstraction design

encoding/interaction technique design

algorithm design

requirements

design

implementation
detailed questions the target users ask
identification of the data of interest

<table>
<thead>
<tr>
<th>species</th>
<th>food_plant</th>
<th>capture_loc</th>
<th>temp</th>
<th>capture date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gymno</td>
<td>conifer</td>
<td>trap_1</td>
<td>75</td>
<td>6/20/1997</td>
</tr>
<tr>
<td>Gymno</td>
<td>herb</td>
<td>trap_53</td>
<td>80</td>
<td>6/21/1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Question: What are the effects of clear-cutting on moth species diversity?

Which moths are most common? rare?
What do common moths eat?
Where do rare moths live?
Minard’s Data

What happened?
Where were the losses?
What role did the cold weather play?
Where was the army at a particular date?
When/Where were the largest losses?
Carte Figurative des pertes suscitées en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.


Les nombres d'hommes portés sur l'axe ordonnance (à raison d'un millier pour dix milliers d'hommes) le sont en même temps que ceux désignés par des flèches. Le tracé, dans lequel sont représentés les pertes, suit la base des pertes dans les campagnes de M. O. Bourgeois, de Vieu, de Ferruch, de Chambry, et le journal militaire de Jacob, président de l'Armée depuis le 13 Octobre.

Pour mieux faire jouer à l'œil la diminution de l'Armée, j'ai supposé que le corps de la Seine formé en la Marne au Bourbonnais, qui arrivait se détaché sur Moscou et se réunissait avec la division de M. O. Bourgeois.

TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Courbes passent au golpe le Milieu, où les températures sont les suivantes:

- 30° de degré
- 35° de degré
- 40° de degré
- 45° de degré
- 50° de degré
- 55° de degré
- 60° de degré
- 65° de degré
- 70° de degré
- 75° de degré
- 80° de degré
- 85° de degré
- 90° de degré
- 95° de degré
- 100° degré

Le 24 Janvier.

Avec les bons vœux de M. XX.

Adieu.

Paris, le 24 Janvier 1869.
How do we get this information?
Exercise: Student Data Questions

• Make a list of the questions that might be asked of this data by (pick one!)
  – a student,
  – a parent
  – a subject teacher
  – the principal/head of the school

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Art</td>
<td>10</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Science</td>
<td>1</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>History</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sport</td>
<td>2</td>
<td>9</td>
<td>5</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Physics</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>English</td>
<td>2</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Chemistry</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Mathematics</td>
<td>10</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>
• Move from domain specific vocabulary to abstract operations

<table>
<thead>
<tr>
<th>species</th>
<th>food plant</th>
<th>capture loc</th>
<th>temp</th>
<th>capture year</th>
<th>capture period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gymno</td>
<td>conifer</td>
<td>trap_1</td>
<td>75</td>
<td>1991</td>
<td>8.2</td>
</tr>
<tr>
<td>Gymno</td>
<td>herb</td>
<td>trap_53</td>
<td>80</td>
<td>1991</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Which are the most common moths? → count, sort
Do common moths emerge earlier in warmer years? → filter, correlate
What’s the distribution of emergence dates? → count, characterize
Where did the losses occur? → derive
Was the weather a factor? → sort, correlate
Where was the army on a particular date? → retrieve
When and where were the largest losses? → derive, sort
Exercise: Student Data Operations

• For your student data questions, take 5 minutes to map each question to at least one operation
  – e.g. How am I doing in math compared to my classmates
    • sort, retrieve
    • compare
    • derive (average)
### Identify data abstractions

<table>
<thead>
<tr>
<th>Logical/Math Ops</th>
<th>Categorical (Qualitative)</th>
<th>Quantitative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nominal</td>
<td>Interval</td>
</tr>
<tr>
<td></td>
<td>Ordinal</td>
<td>Ratio</td>
</tr>
<tr>
<td>* /</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ -</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; &gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>= !</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Example:
- Gender
- Grade Level (Fresh, Soph, Junior, Senior)
- Date (from 1457 BD to AD 2013)
- Lat, Long
- Age Length

Carte Figurative des pertes successives en hommes de l’Armée Française dans la Campagne de Russie 1812-1813.

Direction

Lat/Lon

Size

Location at Dates

Temp

Dates

Q
Exercise: Student Data Abstract Types

<table>
<thead>
<tr>
<th>Subjects:</th>
<th>Students:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Art</td>
<td>10 1 5 3 2</td>
</tr>
<tr>
<td>Science</td>
<td>1 10 5 4 8</td>
</tr>
<tr>
<td>History</td>
<td>8 5 7 1 1</td>
</tr>
<tr>
<td>Sport</td>
<td>2 9 10 4</td>
</tr>
<tr>
<td>Physics</td>
<td>1 3 1</td>
</tr>
<tr>
<td>English</td>
<td>2 8 6 8 5</td>
</tr>
<tr>
<td>Chemistry</td>
<td>4 1 1 1 4</td>
</tr>
<tr>
<td>Mathematics</td>
<td>10 1 5 4 2</td>
</tr>
<tr>
<td>Scores:</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Q</td>
<td></td>
</tr>
</tbody>
</table>
- Map data to visual marks
- Define appropriate interactions to support the operations from previous level
Encoding

- **Mark**: Visual abstractions that represent your objects
- **Map data to visual properties of marks**
  - Position, x, y, z
  - Size, length, area, volume
  - Orientation, angle, slope
  - Color, gray scale, texture
  - Shape
  - Animation, blink, motion


Cleveland and McGill 1984

- Accuracy of judgement of *quantity*

![Diagram showing accuracy of judgement for different quantities: Position, Length, Angle, Slope, Area, Volume, Colour, Density. The most accurate is at the top, and the least accurate is at the bottom.](diagram.png)
Mackinlay’s Ranking 1986

- **Quantitative**
  - Position
  - Length
  - Angle
  - Slope
  - Area
  - Volume
  - Density
  - Shape

- **Ordinal**
  - Position
  - Density
  - Colour saturation
  - Colour hue
  - Texture
  - Connection
  - Containment
  - Length
  - Angle
  - Slope
  - Area
  - Volume

- **Categorical**
  - Position
  - Colour hue
  - Texture
  - Connection
  - Containment
  - Density
  - Colour saturation
  - Shape
  - Length
  - Angle
  - Slope
  - Area
  - Volume
Using Rankings – Rule of Thumb

• Rank your questions/operations
• Identify the data variables that can answer those questions
• Apply best possible (highest ranking) encoding to most important question/operation to encode those variables
Deconstruction: Minard 1861
Napoleonic’s March on Moscow

Based on slides from Jeffrey Heer
Composition Along X axis
Mark Composition

Y: Temperature (Q)

X: Time (Q) and Latitude (Q)

Temperature over space and time
Mark Composition

Y: Longitude (Q)

X: Latitude (Q)

Width: Army Size (Q)

Color: Advance/Retreat (O)
Putting it all back together...

- **Y:** Longitude (Q)
- **X:** Latitude (Q)
- **Width:** Army Size (Q)
- **Color:** Advance/Retreat (N)
- **Y:** Temperature (Q)
- **X:** Time (Q) and Latitude (Q)
Exercise: Student Data Encoding

• Assume you’re designing for students only.
• Pick encodings for the grade data
  – Consider rankings
  – Consider questions
  – variables: students, subjects, scores
• Sketch a static representation of this data
• Which questions does it answer/not answer?
Exercise: Encode Student Data

<table>
<thead>
<tr>
<th>Subjects</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Art</td>
<td>10</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Science</td>
<td>1</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>History</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sport</td>
<td>2</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Physics</td>
<td>-</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>2</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Chemistry</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Mathematics</td>
<td>10</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>
What can go wrong?

threat: wrong problem
validate: observe and interview target users

threat: bad data/operation abstractions

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: lab study, measure human/time errors for operation

validate: test on target users, collect anecdotal evidence of utility
validate: field study, document human usage of deployed system

validate: observe adoption rates
Summary

• Infovis development *is* software development

• Benefits greatly from *user-centered* design
  – Get into target user’s process
  – Understand their needs & use cases
    (e.g. questions and operations)

• Rapid prototyping/iteration is extremely valuable

• These are guidelines, not hard and fast rules
Be an informed shopper!

Google Charts

http://guns.periscopic.com/?year=2013