Information Visualization
CPSC 683

Sheelagh Carpendale

Overview for today

• What is infovis?
• Why infovis?
• Structure of class
• A quick data sketch
• Semiotic approach
• A second data sketch
• Discussion
What is infovis?
Definitions from the OED

• Information
  – The action of informing; communication of the knowledge or ‘news’ of some fact or occurrence; the action of telling or fact of being told of something.

• Visualization
  – The action or fact of visualizing; the power or process of forming a mental picture or vision of something not actually present to the sight; a picture thus formed.
  – The action or process of rendering visible

The purpose of computing is insight, not numbers. (R. Hamming, 1962)

What is InfoVis?

Information visualization combines aspects of scientific visualization, human-computer interaction, graphics, data mining, imaging as well as drawing on cognitive science and graphic design. It focuses on information that is often abstract. Many interesting classes of information have no natural and obvious physical representation. A key research problem is to discover new visual metaphors for representing information and to understand which analytical tasks they support.

Information visualization is a new emerging field whose goal is the perceptualization of information. Information visualization differs from scientific and medical visualization in that the data to be visualized is inherently non-spatial.

Information visualization enables users to get information quickly, put it in a meaningful shape, and to make decisions in a relatively short time.
What is InfoVis?

Information visualization research is developing ways of visualizing information sources and data as a means toward helping people interpret and understand the information. In particular, it focuses on depicting large information spaces, which often contain more informational elements than there are pixels on the screen. The origins of this work are visualizations of large programs and software systems.

Information visualization research seeks to find methods for imposing structure on unstructured bodies of data, and using that structure to visually summarize the data.

Information visualization is a form of external cognition, using resources in the world outside the mind to amplify what the mind can do.

What is InfoVis?

Information Visualization is the use of computer-supported interactive visual representations of abstract data to amplify cognition. (Card et al.)
Why do Information Visualization?

- 2008
- 2010 – 1 zettabyte = 1 trillion gigabytes

It is estimated that 800 exabyte ($800 \times 10^{19}$) of digital information will be generated this year.
Example: Anscombe’s Quartet

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>y</td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>10.0</td>
<td>8.04</td>
<td>10.0</td>
<td>9.14</td>
<td>10.0</td>
</tr>
<tr>
<td>8.0</td>
<td>6.95</td>
<td>8.0</td>
<td>8.14</td>
<td>8.0</td>
</tr>
<tr>
<td>13.0</td>
<td>7.58</td>
<td>13.0</td>
<td>8.74</td>
<td>13.0</td>
</tr>
<tr>
<td>9.0</td>
<td>8.84</td>
<td>9.0</td>
<td>8.77</td>
<td>9.0</td>
</tr>
<tr>
<td>11.0</td>
<td>8.33</td>
<td>11.0</td>
<td>9.26</td>
<td>11.0</td>
</tr>
<tr>
<td>14.0</td>
<td>9.96</td>
<td>14.0</td>
<td>8.10</td>
<td>14.0</td>
</tr>
<tr>
<td>6.0</td>
<td>7.85</td>
<td>6.0</td>
<td>6.13</td>
<td>6.0</td>
</tr>
<tr>
<td>4.0</td>
<td>4.26</td>
<td>4.0</td>
<td>3.10</td>
<td>4.0</td>
</tr>
<tr>
<td>12.0</td>
<td>10.84</td>
<td>12.0</td>
<td>9.13</td>
<td>12.0</td>
</tr>
<tr>
<td>7.0</td>
<td>4.82</td>
<td>7.0</td>
<td>7.26</td>
<td>7.0</td>
</tr>
<tr>
<td>5.0</td>
<td>5.68</td>
<td>5.0</td>
<td>4.74</td>
<td>5.0</td>
</tr>
</tbody>
</table>
Statistical Analysis

For all 4 columns the stats are identical

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>y</td>
<td>x</td>
<td>y</td>
<td>x</td>
</tr>
<tr>
<td>10.0</td>
<td>8.04</td>
<td>10.0</td>
<td>9.14</td>
<td>10.0</td>
</tr>
<tr>
<td>8.0</td>
<td>6.95</td>
<td>8.0</td>
<td>8.14</td>
<td>8.0</td>
</tr>
<tr>
<td>13.0</td>
<td>7.58</td>
<td>13.0</td>
<td>8.74</td>
<td>13.0</td>
</tr>
<tr>
<td>9.0</td>
<td>8.81</td>
<td>9.0</td>
<td>8.77</td>
<td>9.0</td>
</tr>
<tr>
<td>11.0</td>
<td>8.33</td>
<td>11.0</td>
<td>9.26</td>
<td>11.0</td>
</tr>
<tr>
<td>14.0</td>
<td>9.96</td>
<td>14.0</td>
<td>8.10</td>
<td>14.0</td>
</tr>
<tr>
<td>6.0</td>
<td>7.24</td>
<td>6.0</td>
<td>6.13</td>
<td>6.0</td>
</tr>
<tr>
<td>4.0</td>
<td>6.26</td>
<td>4.0</td>
<td>5.10</td>
<td>4.0</td>
</tr>
<tr>
<td>12.0</td>
<td>10.84</td>
<td>12.0</td>
<td>9.13</td>
<td>12.0</td>
</tr>
<tr>
<td>7.0</td>
<td>4.82</td>
<td>7.0</td>
<td>7.26</td>
<td>7.0</td>
</tr>
<tr>
<td>5.0</td>
<td>5.68</td>
<td>5.0</td>
<td>4.74</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Mean of x: 9.0
Variance of x: 11.0
Mean of y: 7.5
Variance of y: 4.12
Correlation between x and y: 0.816
Linear regression line: $y = 3 + 0.5x$

Visual representation reveals a different story

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>y</td>
<td>x</td>
<td>y</td>
<td>x</td>
</tr>
<tr>
<td>10.0</td>
<td>8.04</td>
<td>10.0</td>
<td>9.14</td>
<td>10.0</td>
</tr>
<tr>
<td>8.0</td>
<td>6.95</td>
<td>8.0</td>
<td>8.14</td>
<td>8.0</td>
</tr>
<tr>
<td>13.0</td>
<td>7.58</td>
<td>13.0</td>
<td>8.74</td>
<td>13.0</td>
</tr>
<tr>
<td>9.0</td>
<td>8.81</td>
<td>9.0</td>
<td>8.77</td>
<td>9.0</td>
</tr>
<tr>
<td>11.0</td>
<td>8.33</td>
<td>11.0</td>
<td>9.26</td>
<td>11.0</td>
</tr>
<tr>
<td>14.0</td>
<td>9.96</td>
<td>14.0</td>
<td>8.10</td>
<td>14.0</td>
</tr>
<tr>
<td>6.0</td>
<td>7.24</td>
<td>6.0</td>
<td>6.13</td>
<td>6.0</td>
</tr>
<tr>
<td>4.0</td>
<td>6.26</td>
<td>4.0</td>
<td>5.10</td>
<td>4.0</td>
</tr>
<tr>
<td>12.0</td>
<td>10.84</td>
<td>12.0</td>
<td>9.13</td>
<td>12.0</td>
</tr>
<tr>
<td>7.0</td>
<td>4.82</td>
<td>7.0</td>
<td>7.26</td>
<td>7.0</td>
</tr>
<tr>
<td>5.0</td>
<td>5.68</td>
<td>5.0</td>
<td>4.74</td>
<td>5.0</td>
</tr>
</tbody>
</table>
Benefits of InfoVis

• expand human working memory
  • offload cognitive resources to visual system
• improve search
  • large amount of data in small space
• enhance patterns recognition
  • making patterns visual explicit
• aids monitoring of information
• manipulate-able medium supports exploration

Why visual data representations?

• Vision is our most dominant sense
• We are very good at recognizing visual patterns
• We need to see and understand in order to explain, reason, and make decisions

common examples:

graphs / hierarchies  charts  maps

all examples from: http://vis.stanford.edu/protovis/
CPSC 683: Information Visualization

- Wednesday 11:00-1:45
- MS 680J 220-6055,
- sheelagh@ucalgary.ca,
- Office Hours: by arrangement

Marks
- Class Participation & Assignments 20%
- Paper Presentations 15%
- Project 65%

Class Participation & Assignments
(worth 20%)

- You will be expected to take part in discussions
- Most weeks part will be conducted as seminars
- You will have assigned readings and part and/or questions in regards to the readings which you will discuss in class
- We will also hands on assignments such as
  - Data sketching,
  - Data construction,
  - Etc.
Paper Readings, Presentations, Duties

(15%) You will have assigned readings and questions about the readings which you will discuss in class

1. What did the author(s) think was the main point of the paper?
2. What do you think is the main point of the paper?
3. What practical use can you make of the contents of the paper (or your classmates)
4. Keep the notes – ask questions of the people answering 1, 2, and 3 to be clear about what you write.

P1, S1 does Q1, S2 does Q2, S3 does Q3; S4 takes notes
P2, S5 does Q1, S6 does Q2, S1 does Q3; S2 takes notes
P3, S3 does Q1, S4 does Q2, S5 does Q3; S6 takes notes

Keeping a Visual Journal

Sketch: ideas, data, concepts

• Collect
  • images that others have created for information visualization and/or visual communication in general
  • chosen because you liked/disliked them or because you can not figure out some else’s reaction

• React
  • why and which parts you like or dislike
  • annotate your images, draw on them, write on them, look at them in different scales
  • which parts might you use yourself, as is, with changes, or never!

• Generate
  • keeping track of your developing ideas
  • including, scribbles, sketches, math, and words
Course Project

*Projects must be approved*.

Two possible types of project
1) Visualize 3 aspects of your personal data. Data that affects your life or that you generate. Synthesize into a single application
2) Visualize data that you have or can readily obtain. There is lots of data on the web.

At least two possible manifestations
1) Visualize digitally
2) Visualize physically – physicalization
3) Or some approved alternative

---

Course Project: *Project Timeline*

Week1: Sept 21: data check – get your basic idea approved – tell me about your data is in 3 sentences: what it is; how you can get it; why you chose it.

Week4: Oct 5th: Project proposals due: Proposals (10%) consist of
- approximately 1 page description of what you propose to do
- a data ‘cheat sheet’ for your data
- 10 hand sketches – each conceptually distinct.
- 5 minute presentation – sketches on wall for discussion.

Week6: Oct 19th: project check in (5%):
- 5 minute presentation – 10 min discussion.
Course Project: *Project Timeline*

**Week 10: Nov 16th**: project check in (5%):
- 5 minute presentation – 10 min discussion

**Week 13: Dec 7th**: Project due (10+5+5+30+10+5 = 65) : projects consist of
- report-30% (lit-7%, representation-7%, presentation-7%, interaction-7%, - overall - 2%)
- illustrated, includes proposal, ~1 page for each check in
- video (10%)
- 10 minute presentation (5%)

---

**Visualization Pipeline**

```
start

Data Abstraction

Data Representation

Visual Representation

Visual Presentation

View

View Transformations

Presentation Transformation

Visual Abstraction

Data Abstraction

Interactions

Layout Spatial Organization
```

---

start
Infovis 683 topics

- Academic
- Algorithmic
- Alternate aesthetics
- Communication
- Constructive
- Data
- Design thinking
- Externalization
- Graph design
- Human in the loop
- Hybrid approaches

- Interaction
- Observation for design
- Cognitive / perception
- Personal visualization
- Physicalization
- Representation
- Semiotics
- Sketching
- Stats
- Story telling
- Task (infovis design study)

Infovis 683 topics: fall 2016

- Wk1: intro / Semiotics / Sketching
- Wk2: Externalization / Design thinking / project approval
- Wk3: task / design studies /proposals
- Wk4: Construction / physicalization
- Wk5: cognition / perception
- Wk6: representation / project check in

- Wk7: bio-mimicry
- Wk8: personal vis / story telling
- Wk9: statistics
- Wk10: communication / story telling / project check in
- Wk11: Interaction
- Wk12: Observation for design
- Wk13: final presentations