Collaborative Visual Analysis

Maneesh Agrawala

CS 448B: Visualization
Fall 2016

Last Time: Deconstructing Visualizations
<table>
<thead>
<tr>
<th>Disease</th>
<th>Budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aids</td>
<td>70.0%</td>
</tr>
<tr>
<td>Alzheimer’s</td>
<td>5.0%</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>1.1%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>4.6%</td>
</tr>
<tr>
<td>Hepatitis B</td>
<td>4.1%</td>
</tr>
<tr>
<td>Hepatitis C</td>
<td>3.8%</td>
</tr>
<tr>
<td>Parkinson’</td>
<td>6.0%</td>
</tr>
<tr>
<td>Prostate</td>
<td>5.2%</td>
</tr>
</tbody>
</table>
Bar Charts

- Find Foreground Rectangles
- Identify Orientation and Baseline
- Recover Bar Values
- Associate Labels with Bars

Extract Marks

Extract Data

marks: lines

<table>
<thead>
<tr>
<th>y-value</th>
<th>x-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>A</td>
</tr>
<tr>
<td>35</td>
<td>B</td>
</tr>
<tr>
<td>4</td>
<td>C</td>
</tr>
<tr>
<td>75</td>
<td>D</td>
</tr>
</tbody>
</table>

Scale: 2 pixels/unit

Interactive Documents

How can we facilitate reading text and charts together?

Syrian refugees: how many are there and where are they?

The humanitarians listed in theoror in Syria receive new proportions on the number of individuals who are refugees who are in need of assistance. More data journalism and data visualizations from the Guardian.

Some contributions are made on a regional basis, but many donor countries prefer to contribute to efforts in a specific country. In line with the distribution of the refugees themselves, most support is funnelled towards Jordan (20%), followed by Lebanon (20%), Turkey (18%) and Iraq (11%).

Where the money goes

Where the international community has donated to help Syria's refugees

- Egypt
- Iraq
- Turkey
- Regional
- Lebanon
- Jordan

The Guardian
Announcements

Final project

Design new visualization method (e.g. software)
- Pose problem, Implement creative solution
- Design studies/evaluations less common but also possible (talk to us)

Deliverables
- Implementation of solution
- 6-8 page paper in format of conference paper submission
- Project progress presentations

Schedule
- Project proposal: 11/7
- Project progress presentation: 11/16 in class (3-4 min) slide presentation
- Final poster presentation: 12/9 Location: TBD
- Final paper: 12/11 11:59pm

Grading
- Groups of up to 3 people, graded individually
- Clearly report responsibilities of each member
In-Class Project Presentations

Dates: 11/16 In class presentation: **3 min 30 sec per group**
- Description of problem you are addressing
- Relevant prior work, how your work is different
- Description/storyboard and of approach
- A list of milestones for finishing the project by the deadline

Scheduling
- Slides pdf due 10am on 11/16 (make sure you can finish in time)
- When not presenting should post feedback (via google form)

Afterwards group should post on the wiki
- Review of relevant prior work with full references
- List of milestones (scope your project appropriately)

---

**A Tale of Two Visualizations**
**Observations**

Groups spent more time in front of the visualization than individuals

Friends encouraged each other to unearth relationships, probe community boundaries, and challenge reported information

Social play resulted in informal analysis, often driven by story-telling of group histories
Social Data Analysis

Visual sensemaking can be social as well as cognitive

Analysis of data coupled with social interpretation and deliberation

How can user interfaces catalyze and support collaborative visual analysis?
Excel + Email

[Excel spreadsheet and email interface images]

---

Wikimapia.org

[Map and search interface images]
Spotfire Decision Site Posters

Many-Eyes

[Viégas, Wattenberg, et al. CHI 07]
Tableau Public

Where can Collaborators Contribute?

Data Management
- Contribute Data
- Clean Data
- Categorize Data
- Moderate Data
- Create Metadata

Visualization
- Visual Structures
- Views

Visual Analytics
- Observations
- Hypotheses
- Evidence (+/-)
- Summarize
- Report / Presentation

Data Transformations
- Visual Mappings
- View Transformations
Design Considerations
[Heer & Agrawala VAST 07, IVS 08]

1. Division, allocation, and integration of work
2. Common ground and awareness
3. Reference and deixis (pointing)
4. Identity, trust, and reputation
5. Group formation and management
6. Incentives and engagement
8. Presentation and decision-making

Sense.us: Collaborative Visualization of Demographic Data
Exploratory Design Rationale

Sharing within visualization and across the web
Exploratory Design Rationale

Sharing within visualization and across the web
Pointing at interesting trends, outliers
Collecting and linking related views
Exploratory Design Rationale

Sharing within visualization and across the web
Pointing at interesting trends, outliers
Collecting and linking related views
Awareness of social activity

Building Off of Others
Data Jokes

Great depression "killed" a lot of brokers

Collaborative Sensemaking

Dentist  Technician - Medical  Dental

% of Work Force
Sharing in External Media

Designing for Social Data Analysis
Social Data Analysis

How can users’ activity traces be used to improve awareness in collaborative analysis?

Social Navigation
Scented Widgets [InfoVis 07]

Visual navigation cues embedded in interface widgets

Visitation counts
Comment counts

No scent (baseline)
Do activity cues affect usage?

Hypotheses: With activity cues, subjects will
1. Exhibit more revisitation of popular views
2. Make more unique observations

Controlled experiment with 28 subjects
Collect evidence for and against an assertion
Varied scent cues (3) and foraging task (3)

“Technology is costing jobs by making occupations obsolete.”
Results

Unique Discoveries
Visit scent had sig. higher rate of discoveries in first block
Less reliance on scent when subjects were familiar with data and visualization

Revisitation
Visit and comment scent conditions correlate more highly with seed usage than no scent

Social Data Analysis
How can users’ activity traces be used to improve collaborative analysis?

How should annotation techniques be designed to provide nuanced pointing behaviors?
Common Ground

The shared understanding enabling conversation and collaborative action
[Clark & Brennan ’91]

Do you see what I see? View sharing (URLs)

Principle of Least Collaborative Effort: participants exert just enough effort to successfully communicate
[Clark & Wilkes-Gibbs ’86]
“Look at that spike.”

“Look at the spike for Turkey.”
“Look at the spike in the middle.”

Use of Annotations

- Arrows: 25.1%
- Text: 24.6%
- Ovals: 17.9%
- Pencil: 16.2%
- Lines: 14.5%
- Rectangles: 1.7%

39.0% of comments included annotations

*Pointing* to specific points, trends, or regions (88.6%)

*Drawing* to socialize or tell jokes (11.4%)

Variety of subject responses

‘Not always necessary’, but ‘surprisingly satisfying’

Some concern about professional look
Many-Eyes: Social Data Analysis at an Internet Scale

[Viégas, Wattenberg, et al. CHI 07]
Templates for Visualizing Data

Create a visualization in three easy steps

1. CHOOSE A DATA SET
   - First thing first. Choose the data set that you would like to visualize.

2. CHOOSE A VISUALIZATION
   - Then choose the kind of visualization you would like to use.

3. CUSTOMIZE & PUBLISH
   - Customize what your visualization will look like when others visit, then publish it.

Start with a data set

You have two choices:
- Use one of the existing data sets on the site, or
- Upload your own data set

Commenting


- Veterans' benefits are going down, percentage-wise
- See more in the comment

- Health care costs escalating
- See more in the comment

- Spurious funding issues ongoing before the public
- See more in the comment

- Judicial funding issues ongoing before the public
- See more in the comment

- The national defense expenses in 1998 and 2014 are roughly equal; but the defense in 1998's entitlements do not compare with today’s
- See more in the comment
Outside Communities

Swivel.com, and others...
A graveyard of “YouTubes for Data”

Emergent analysis and discussion isn’t very good
Many Eyes – circa 4/2012

128,478 Visualizations

17,340 Comments

only ~11% of comments provided a plausible hypothesis or explanation for the data in the chart

CommentSpace: Structured Support for Social Data Analysis
Can we augment social data analysis to support deeper analysis and synthesis?
Tight coupling of comments & views

Tags and links for organization

- Hypothesis
- Question
- To-Do

- Evidence For
- Evidence Against
- (Related)
Hypothesis generation/evidence gathering

Contributors highlight important items with tags

Tags help late-joiners identify and build on important comments

Contributors use links to organize contributions and build narrative
Studies and Deployments

Controlled lab studies to test core analysis subtasks

Live deployments (www.commentspace.net)

Study: Use of Tags and Links

Hypothesis: Tags and links can provide common ground and encourage continued discussion.

A between subjects study (n=24) with 2 conditions.

“No-Tag” Condition

“Tag” Condition
Study: Prompt

Hypothesis: Stereotypically male jobs have remained almost entirely male even as women have joined the work force.

Study: Results

Participants who used tags and links classified comments more **consistently** and **accurately** than those who didn’t

- **(greater in-group agreement)**
- **(greater agreement with experts)**

Participants using tags and links generated significantly more replies to existing comments

- **Tag (Median=7)**
- **No-Tag (Median=2)**
But ....

In open-ended tasks, participants still engaged mostly in superficial, exploratory analysis.

We saw very little use of tags or links.

The Impact of Social Information on Visual Judgments
Can having access to social information be harmful to analysis?

Can evidence of others’ graphical judgments (accurate or not) impact subsequent users’ judgments?
Social Influence

Normative versus Informational

Using social signals as evidence when we’re choosing:
- Restaurants
- Music (Salganik et al. 2009)
- Tags (Golder et al. 2007)
Controlled Experiment

Interested in impact on accuracy of graphical judgments (given task where subjects are incentivized to be correct)

Two conditions:
Non-social: Visual judgment task with no social information
Social: Same task, presented with summary of judgments from non-social group

Visual judgment task

1. Which of the two marked sections is larger?
2. Make a quick visual judgment of what proportion the smaller section is of the larger.

Cleveland & McGill 1984

AMT study: Heer & Bostock 2010
Results

Can having access to others’ graphical judgments (accurate or not) impact subsequent users’ judgments?

Does the amount of social information matter?
Summary

Data analysis is usually a collaborative process

Supporting collaboration between strangers is difficult

Desire to fit in can influence analyst when working in a collaborative setting