COMP 150-04
Visualization

Lecture 12
Storyboarding and Evaluation
Please fill out the midterm course evaluation:

https://www.eecs.tufts.edu/cscourse/feedback.php
Final project

Identify interesting data sets
Apply visualization techniques to your own work

See “Project Ideas” link on the wiki

Beyond this semester
  Add to your portfolio
  Write a senior thesis
  Submit a conference paper or poster
  Continue research collaborations
Final project

Work individually or with a partner
Your choice of language, API, toolkit
Work expected (per person): on the order of two assignments

Proposal: In class, week of April 5th

Final presentations: Written report and demo session
Final project

Option 1: Design and implement a new visualization tool

Option 2: Extend an existing tool

Option 3: Conduct a user study to evaluate or compare tools
Review: Information seeking

Overview: see overall patterns in data

Zoom: see a subset of data

Filter: see a subset based on values

Detail on demand: see values of items

Relate: compare values

History: keep track of actions

Extract: mark and capture
Review:
Information seeking

Overview + Detail

Focus + Context
Review: Information seeking

Brushing and linking

Dynamic queries
Interacting with visualizations

Evaluation
- Controlled experiments
- Usability testing

Design
- Task analysis
- Storyboarding

http://gapminder.org
Evaluating visualization designs: Why?
Evaluating visualization designs: How?

Test with live participants
   Controlled experiment
   Usability test

Conduct an expert evaluation

Simulate with a model
   Goals, Operators, Methods, and Selection rules (GOMS)
<table>
<thead>
<tr>
<th>Usability test</th>
<th>Controlled experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guide a single design</td>
<td>Compare multiple designs</td>
</tr>
<tr>
<td>Few participants</td>
<td>Many participants</td>
</tr>
<tr>
<td>Qualitative feedback</td>
<td>Quantitative results</td>
</tr>
<tr>
<td>Identify usability problems</td>
<td>Identify statistical significance</td>
</tr>
<tr>
<td>Formative</td>
<td>Summative</td>
</tr>
</tbody>
</table>
Comparing graph visualizations

Node-Link

Matrix

Mohammad Ghoniem et al., A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations
In Proceedings of InfoVis'04
http://mohammad.ghoniem.info
Controlled experiment

The scientific method:

1. Form hypothesis
2. Collect data
3. Analyze data
4. Accept or reject hypothesis
Controlled experiment

Goal: Determine cause and effect
Procedure: Vary cause, measure effect

Cause: Independent variables
   Visualization
   Task (search, edit, ...)
   Data attributes

Effect: Dependent variables
   User performance (time, accuracy) on task
   Subjective impressions
## Experimental design: Variables

**Independent: Task**

<table>
<thead>
<tr>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
</tr>
</thead>
</table>

**Independent: Graph**

“A 2x4 design”
Experimental design: Variables

Independent: Task

Task 1  Task 2  Task 3  Task 4

Treatments (variable values)

“A 2x4 design”
Experimental design: Variables

Independent: Task

Task 1 | Task 2 | Task 3 | Task 4

Independent: Graph

Dependent: Performance time (Measured)

“A 2x4 design”
Experimental design: Groups

Between-subjects: One subject group per treatment

Within-subjects: All subjects perform all treatments

Counterbalance treatment order

Select uniform subject pool

- Spatial ability
- Colorblindness
- Handedness
Experimental design: Tasks

Should allow direct measurement

Should be “realistic”
  Goal: Ecological validity
  Do not cherry-pick
  May need to compress for time

Task analysis
Experimental design: Task analysis

Who? With others?

What tasks? How learned?

Under what conditions? How often?

Relationships between user and data?

Existing tools?

Error recovery?
Experimental design: Task selection

Real tasks from user interviews

Benchmark tasks
   Easy (common) to difficult (infrequent)
   Specific to open-ended

Representative distribution
Some end-to-end examples

During the test: State the task but not the detailed steps
Experimental procedure

Informed consent

Pre-survey: Demographics

Tutorial: Instructions and training runs

Actual runs: Measure performance on tasks

Post-survey: Subjective impressions

Max 1 hour
Controlled experiment: Performance analysis

Analyze at aggregate level

Determine statistical significance of effect
   Analysis of variance (ANOVA)

Often dealing with many variables
   Graph style: node-link, matrix
   Graph parameters: number of nodes, connectivity
   Tasks: complexity, frequency

Be mindful of what is being compared
Usability

ISO guidelines: “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.”
ISO guidelines: “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.”


ISO 9241: Ergonomics of Human System Interaction, Part 11
Usability testing

Goal: Identify usability problems
   Performance
   Learning
   Aesthetics

Procedure:
   Assign benchmark tasks
   Record qualitative feedback
   Record rough quantitative measures
Usability test: Procedure

Take notes
  Use videotaping or screen-recording if necessary

Prompt subjects to think aloud while performing tasks
  what they are thinking
  what they are trying to do
  where they are getting stuck

Avoid interfering
  Help resolve technical issues not related to the study topic
  Help if the user cannot continue the task
Usability test:
Tasks and scenarios

For each task, produce a scenario
  Step-by-step performance
  Storyboard steps: main screens, functions, transitions
  Estimate performance (e.g. completion time): best, worst, average cases

Unlike tasks, scenarios are design specific
Storyboarding interactions
Explaining a process

Depict:
Key steps of interaction to reach a goal
Context: Relate user, design, environment
Sketching interaction states
Sketching interaction states

Arie Stavchansky
http://stavchansky.net
Storyboard at multiple levels of detail

Build mock-ups early
  Low fidelity: Paper prototyping, copy-paste
  Medium fidelity: Interactive prototyping tools (Flash, VB, ...)
  High fidelity: GUI/Vis toolkits

“Play” for users and observe reactions
Iteratively evaluate and refine
Usability test: Performance analysis

Compare observed and estimated performance

Determine severity of a usability problem
   How frequently does it occur?
   How common is this task or scenario?
   What is the user’s reaction?
   What is the cost of changing the design?

Solve problems based on importance, cost

Favor simple solutions over major redesigns
Comparing tree visualizations

Indented list

Hyberbolic tree

Peter Pirolli et al., “The Effect of Information Scent on Searching Information Visualizations of Large Tree Structures”
Proceedings of AVI 2000
Storyboarding exercise

1. Food Journal: Daily calorie intake
2. March Madness: Basketball wins and losses
3. Jams: Real-time traffic map
4. Site Explorer: Website hierarchy
5. Personal Shopper: Clothing from multiple vendors
6. Taking Stock: Stock portfolio updates
7. Field Assistant: Database of building blueprints
8. Rolodex: Contact database synced to social networks
9. Build It: Furniture assembly instructions
10. Bargain Hunter: Price lookup by barcode