**Visualization Analysis & Design**

*Tamara Munzner*

Department of Computer Science
University of British Columbia

November 21, 2005 San Francisco CA


---

**Why represent all the data?**

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- summaries lose information, details matter
- confirm expected and find unexpected patterns
- assess validity of data model

**Anscombe's Quartet**

**Why analyze?**

- imposes a structure on huge design space
- scaffold to help you think systematically about choices
- analyzing existing as stepping stone to designing new

**Types: Datasets and data**

- tables
- networks
- spatial
- geometry

**Attributes Types**

- categorical
- ordered
- quantitative

**Actions I: Analyze**

- consume
- discover vs present
- visual vs virtual
- explore vs explain
- inquiry
- produce
- annotate, record
- derive
- crucial design choices

**Derive**

- don't just draw what you're given!
- derive the right thing to show is
- create a series of transformations from the original dataset
- derive
- one of the four major strategies for handling complexity

**Analysis example: Derive one attribute**

- Skal of numbers
- centrality metric for trees
- derived quantitative attribute
- draw top 5% of OECD for good islands

**Actions II: Search**

- what does user know?
- targets, location

---

**Why is validation difficult?**

- different ways to get it wrong at each level
- don't need vis when fully automatic solution exists and is trusted
- design space
- cognitive science
- how to designing new...what
- deriving
- create it with a series of transformations from the original dataset
- classic split
- centroid metric for trees
- derived quantitative attribute
- draw top 5% of OECD for good islands

**Analysis framework: Four levels, three questions**

- domain situation
- who are the targets users?
- abstraction
- translate from specifics of domain to vocabulary of vis
- what is shown? data abstraction
- what do they want to see?
- task abstraction
- how is it shown?
- visual encoding idiom: how to draw
- interaction idiom: how to arrange
- algorithm
- efficient computation

---

**Why have a human in the loop?**

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- human-in-the-loop is crucial when there is a need to augment human capability rather than replace people with computational decision-making methods.
- don't need vis when fully automatic solution exists and is trusted
- many analysis problems ill-specified
- don't have exact questions to ask in advance
- possibilities
- long-term use for end users (e.g. exploratory analysis of scientific data)
- presentation of known results
- stepping stones to better understanding of requirements before developing models
- help developers of automatic solution rethink/redo; determine parameters
- help end users of automatic solutions verify build-trust

---

**Why use an external representation?**

Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

- external representation: replace cognition with perception

---

**Why is validation difficult?**

- solution: use methods from different fields at each level
How to handle complexity: 3 more strategies

- Change
- Select
- Navigate

- Reduce
- Facet
- Express

- Separate
- Align
- Filter

- Metrics
- Attributes
- Grouping

- Change
- Facet
- Reduce

- Distinct
- Same
- Different

- Link/Highlighting
- Share Data: All/Subset/None

Idiom: Animated transitions

- Smooth transition from one state to another
- Support for item tracking when amount of change is limited
- Example: Multilevel matrix views
  - Scope of what is shown narrows down
  - Reducing block in matrix view, additional attributes appear

Encoding visually

- Analyze idiom structure

Channels: Matching Types

- Perceptual channels
  - Magnitude
  - Identity
  - Spatial

Channels: Rankings

- Perceptual channels
  - Magnitude
  - Identity
  - Spatial

How to encode: Arrange space, map channels

- Arrange
- Express
- Separate
- Align
- Filter
- Aggregate

- Metrics
- Attributes
- Grouping

- Change
- Facet
- Reduce

- Distinct
- Same
- Different

- Link/Highlighting
- Share Data: All/Subset/None

How to encode: Arrange space, map channels

- Post-bin univariate scale
- Post-bin categorical scale

- Spatial region
  - Spatial extent
  - Color hue
  - Shape

- Magnitude channels
  - Ordered attributes
  - Identity channels
  - Categorical attributes

How to handle complexity: 3 more strategies

- Change
- Facet
- Reduce

- Distinct
- Same
- Different

- Link/Highlighting
- Share Data: All/Subset/None