Visualization Analysis & Design

Tamara Munzner
Department of Computer Science
University of British Columbia
University of Washington, Data Science Seminar
September 30 2015, Seattle WA
http://www.cs.ubc.ca/~tmm/talks.html#vad15uw

Visualization Analysis & Design
Tamara Munzner
Department of Computer Science
University of British Columbia
University of Washington, Data Science Seminar
September 30 2015, Seattle WA

Analysis framework: Four levels, three questions
- domain situation
  - who are the target users?
- abstraction
  - translation from specific to vocabulary of vis
  - what is shown in data abstraction
  - what don’t you draw when you give transforms to new form
- why is the user looking at it? task abstraction
  - idiom
    - how is it shown?
    - visual encoding idiom: how to draw
    - interaction idiom: how to manipulate
  - algorithm
    - efficient computation

Types: Datasets and data
- Dataset Types
  - Tables
  - Networks
  - Fields (Continuous)
  - Geometry (Spatial)
- Attribute Types
  - Categorical
  - Ordered
  - Ordinal
  - Quantitative

Actions II: Search
- what does user know?
  - target, location
- how much of the data matters?
  - one, some, all
- analyze, search, query
  - independent choices for each

Actions III: Query
- what does user know?
  - target, location
- how much of the data matters?
  - one, some, all
- analyze, search, query
  - independent choices for each

Factors: Why analyze?
- imposes a structure on huge design space
  - scaffold to help you think systematically about choices
- reduces cognitive load
  - saves time/memos
- consumes time/memos
- produces
  - annotate, record
  - derive
  - crucial design choice

Why use an external representation?
- external representation: replace cognition with perception
- computer-based visualization provides visual representations of datasets designed to help people carry out tasks more effectively.

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

Why?...

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 statistical model

Anscombe’s Quartet

- x mean
- y mean
- x variance
- y variance
- x mean
- y variance
- x correlation
- y correlation

Validation methods from different fields for each level
anthropology/ethnography
design
computer science
cognitive psychology

Dataset Availability
Static Dynamic

How?
Encode Manipulate Facet Reduce

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.
- don’t need visual if full automatic solution exists and is trusted
- many analysis problems ill-specified
- don’t know exactly what questions to ask in advance
- possibilities
  - long-term use for end users (e.g. exploratory analysis of scientific data)
  - preservation of known results
  - stepping stones for better understanding of requirements before developing models
  - help developers of automatic solution rethinking; determine parameters
  - help end-users of automatic solution verify build-truth
How to encode: Arrange space, map channels

Encoding visually
- analyze idiom structure

Definitions: Marks and channels
- marks
  - geometric primitives
- channels
  - visual appearance of marks

Channels: Expressiveness types and effectiveness rankings

Idiom: Linked highlighting
- see how regions contiguously in one view are distributed within another
  - powerful and pervasive interaction idiom
- encoding different
  - multiform
- data all-shared

Idiom: Animated transitions
- smooth transition from one state to another
  - alternative to jump cuts
  - support for item tracking when amount of change is limited
- example: multilevel matrix views
  - scope of what is shown narrows down
- either blocks expand down to increasingly aggregated representations

Idiom: bird’s-eye maps
- encoding: same
- data: subsets shared
- navigation: shared
  - bidirectional linking
- differences: viewpoints
  - overview–detail

Facet
- Partition
  - Superimpose
    - Share Navigation
      - Share Encoding: Same/Different
      - Linked Highlighting
    - Data: All/Subset/None

Encoding visually with marks and channels
- analyze idiom structure
  - as combination of marks and channels

Channels: Matching Types

Channels: Rankings

How to handle complexity: 3 more strategies

System: EDV

System: Google Maps

System: Cerebral

Definition: Marks and channels
- marks
- channels
- shared
- encoding: same
- data: subsets shared
- navigation: shared
- bidirectional linking
- differences: viewpoints
- overview–detail

Mark: point
- position
- value
- attributes

Shape
- size
- area
- length
- volume

Color
- spectrum
- hue
- luminance
- saturation

Resources:
Partitioning: Design choice interaction

Coordinate views: Data and Design choice interaction

Partitioning into views

• how to divide data between views
  - encodes associations between what is seen using spatial proximity
  - major implications for how patterns are visible
  - split according to attributes
  - design choices
  - how many splits
  - all the way down: one mark per region?
  - easy enough for more complex structure within region?
  - order in which attributes used to split
  - how many views

• why juxtapose views?
  - benefits: eyes vs memory
  - lower cognitive load to move eyes between 2 views than remembering previous state with single changing view
  - costs: display area, 2 views side by side each have only half the area of one view


FIGURES

Idiom: Dimensionality reduction for documents

• attribute aggregation
  - derive low-dimensional target space from high-dimensional measured space

Idiom: Partitioning: List alignment

• single bar chart with grouped bars
  - split by state into regions
  - compare gdp within each region showing all ages

• small-multiple bar charts
  - split by age into regions
  - compare easy within age, harder across ages

• small-multiple pie charts
  - one pie chart per region
  - compare easy within age, harder across ages

How to handle complexity: 3 more strategies

• reduce what is shown within single view
  - reduce/increase: inverses
  - filter
  - aggregation

Reduce and attributes

• reduce/increase: inverses
  - filter
  - aggregation

• to/whence
  - from/whence

• how many splits
  - all the way down: one mark per region?
  - easy enough for more complex structure within region?
  - order in which attributes used to split
  - how many views

• why juxtapose views?
  - benefits: eyes vs memory
  - lower cognitive load to move eyes between 2 views than remembering previous state with single changing view
  - costs: display area, 2 views side by side each have only half the area of one view


FIGURES

Idiom: More Information

• reduce what is shown
• reduce/increase: inverses
• filter
  - pro: straightforward and intuitive
  - to/whence
  - from/whence

• aggregation
  - pro: infomin about whole set
  - con: difficult to avoid losing signal

• not mutually exclusive
  - combine filter, aggregate

Reduce and attributes

• reduce/increase: inverses
• filter
  - pro: straightforward and intuitive
  - to/whence
  - from/whence

• aggregation
  - pro: infomin about whole set
  - con: difficult to avoid losing signal

• not mutually exclusive
  - combine filter, aggregate

Reduce and attributes

• reduce/increase: inverses
• filter
  - pro: straightforward and intuitive
  - to/whence
  - from/whence

• aggregation
  - pro: infomin about whole set
  - con: difficult to avoid losing signal

• not mutually exclusive
  - combine filter, aggregate

Reduce and attributes

• reduce/increase: inverses
• filter
  - pro: straightforward and intuitive
  - to/whence
  - from/whence

• aggregation
  - pro: infomin about whole set
  - con: difficult to avoid losing signal

• not mutually exclusive
  - combine filter, aggregate