Week 4: Manipulate, Facet, Reduce

Tamara Munzner
Department of Computer Science
University of British Columbia

JRNL 520H, Special Topics in Contemporary Journalism: Data Visualization
Week 4: 4 October 2016

http://www.cs.ubc.ca/~tmm/courses/journ16
Whereabouts

• Caitlin on travel this week and next week
  – don’t expect email answers until she returns; email Tamara instead!

• Tamara on travel Thu Oct 6 - Mon Oct 10
  – in Portland Fri/Sat to give another keynote, will still be answering email
  – short office hours in Sing Tao next week: 12:30-1:30pm
News

• Assign 2 marks not out yet
  – stay tuned, just got back from Stanford late last night
• Today’s format
  – interleave foundations & demos
    • Tamara will walk through Tableau demos
    • you follow along step by step on your own laptop
    • Tamara will take breaks to rove the room to help out folks who get stuck
Last Time
Demo 1: Stone Color Workbook

• Credit: Maureen Stone, Tableau Research
  – designer of Tableau color defaults, author of *A Field Guide to Digital Color*
  – workbook from Tableau Customer Conference 2014 talk
    Seriously Colorful: Advanced Color Principles & Practices

• Tableau Lessons
  – more visual encoding practice
  – color palettes, univariate & bivariate
  – discrete (categorical) vs continuous (quantitative)

• Big Ideas
  – Tableau has many built-in features to get color right, but care still needed
Demo 2: Intro to Maps

• Tableau Lessons
  – handling spatial data
  – multiple data sources
  – paths on maps
  – more on handling missing data: filtering

• Big Ideas
  – integrating visual encoding design choices with given spatial data
How?

Encode

- Arrange
  - Express
  - Separate
- Order
  - Align
- Use

Map
- from categorical and ordered attributes
- Color
  - Hue
- Saturation
- Luminance
- Size, Angle, Curvature, ...
- Shape
- Motion
  - Direction, Rate, Frequency, ...

Manipulate

- Change
- Select
- Navigate

Facet

- Juxtapose
- Partition
- Superimpose

Reduce

- Filter
- Aggregate
- Embed

What?

Why?

How?
How to handle complexity: 1 previous strategy + 3 more

- Derive new data to show within view
- Change view over time
- Facet across multiple views
- Reduce items/attributes within single view

Derive

Manipulate

- Change

Facet

- Juxtapose

Reduce

- Filter

Select

- Partition

Navigate

- Superimpose

Embed

- Aggregate
Manipulate

- **Change over Time**
- **Navigate**
  - Item Reduction
    - Zoom
      - Geometric or Semantic
    - Pan/Translate
    - Constrained
  - Attribute Reduction
    - Slice
    - Cut
    - Project

- **Select**
Change over time

• change any of the other choices
  – encoding itself
  – parameters
  – arrange: rearrange, reorder
  – aggregation level, what is filtered...
  – interaction entails change
Idiom: **Re-encode**  
System: **Tableau**

made using Tableau, [http://tableausoftware.com](http://tableausoftware.com)
Idiom: **Reorder**

- **data**: tables with many attributes
- **task**: compare rankings

**System:** **LineUp**

Idiom: **Realign**

- stacked bars
  - easy to compare
    - first segment
    - total bar
- align to different segment
  - supports flexible comparison

**System: LineUp**

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

• example: animated transitions in statistical data graphics
  – https://vimeo.com/19278444

Select and highlight

- selection: basic operation for most interaction
- design choices
  - how many selection types?
    - click vs hover: heavyweight, lightweight
    - primary vs secondary: semantics (eg source/target)
- highlight: change visual encoding for selection targets
  - color
    - limitation: existing color coding hidden
  - other channels (eg motion)
  - add explicit connection marks between items
Navigate: Changing item visibility

- change viewpoint
  - changes which items are visible within view
  - camera metaphor
    - zoom
      - geometric zoom: familiar semantics
      - semantic zoom: adapt object representation based on available pixels
        » dramatic change, or more subtle one
    - pan/translate
    - rotate
      - especially in 3D
  - constrained navigation
    - often with animated transitions
    - often based on selection set
Idiom: **Semantic zooming**

- visual encoding change
  - colored box
  - sparkline
  - simple line chart
  - full chart: axes and tickmarks

**System: LiveRAC**

Navigate: Reducing attributes

• continuation of camera metaphor
  – slice
  • show only items matching specific value for given attribute: slicing plane
  • axis aligned, or arbitrary alignment
  – cut
  • show only items on far slide of plane from camera
  – project
  • change mathematics of image creation
    – orthographic
    – perspective
    – many others: Mercator, cabinet, ...

Previous Demos

• Tableau Lessons
  – changing visual encoding
  – changing ordering (sorting)
  – navigation
    • zoom/translate in maps
### How?

#### Encode

<table>
<thead>
<tr>
<th>Arrange</th>
<th>Order</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Express</td>
<td>Separate</td>
<td>Align</td>
</tr>
</tbody>
</table>

#### Manipulate

<table>
<thead>
<tr>
<th>Change</th>
<th>Juxtapose</th>
</tr>
</thead>
</table>

#### Facet

<table>
<thead>
<tr>
<th>Select</th>
<th>Partition</th>
<th>Superimpose</th>
</tr>
</thead>
</table>

#### Reduce

<table>
<thead>
<tr>
<th>Filter</th>
<th>Aggregate</th>
<th>Embed</th>
</tr>
</thead>
</table>

### What?

- Map from categorical and ordered attributes:
  - Color
    - Hue
    - Saturation
    - Luminance
  - Size, Angle, Curvature, ...
- Shape:
  - + • ■ ▲
- Motion:
  - Direction, Rate, Frequency, ...

### Why?

### How?

- Map
- Change
- Select
- Superimpose
Facet

- Juxtapose

- Partition

- Superimpose
Juxtapose and coordinate views

- Share Encoding: Same/Different
  - Linked Highlighting

- Share Data: All/Subset/None

- Share Navigation
**Idiom: Linked highlighting**

- see how regions contiguous in one view are distributed within another
  - powerful and pervasive interaction idiom

- encoding: different
  - *multiform*

- data: all shared

Demo 1: Seattle Construction

- Credit: Ben Jones

- Tableau Lessons
  - linking views with actions: highlight on hover
  - global filtering

- Big Ideas
  - linking views possible but somewhat clunky in Tableau
Idiom: bird’s-eye maps

- encoding: same
- data: subset shared
- navigation: shared
  - bidirectional linking

- differences
  - viewpoint
  - (size)

- overview-detail

Idiom: Small multiples

- encoding: same
- data: none shared
  - different attributes for node colors
  - (same network layout)
- navigation: shared

System: Cerebral

## Coordinate views: Design choice interaction

<table>
<thead>
<tr>
<th>Encoding</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Same</strong></td>
<td><strong>All</strong></td>
</tr>
<tr>
<td></td>
<td>Reapundant</td>
</tr>
<tr>
<td></td>
<td><strong>Subset</strong></td>
</tr>
<tr>
<td></td>
<td>Overview/Detail</td>
</tr>
<tr>
<td></td>
<td><strong>None</strong></td>
</tr>
<tr>
<td></td>
<td>Small Multiples</td>
</tr>
<tr>
<td><strong>Different</strong></td>
<td><strong>All</strong></td>
</tr>
<tr>
<td></td>
<td>Multiform</td>
</tr>
<tr>
<td></td>
<td><strong>Subset</strong></td>
</tr>
<tr>
<td></td>
<td>Multiform, Overview/Detail</td>
</tr>
<tr>
<td></td>
<td><strong>None</strong></td>
</tr>
<tr>
<td></td>
<td>No Linkage</td>
</tr>
</tbody>
</table>

- **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**
Why not animation?

• disparate frames and regions: comparison difficult
  – vs contiguous frames
  – vs small region
  – vs coherent motion of group

• safe special case
  – animated transitions
System: **Improvise**

- investigate power of multiple views
  - pushing limits on view count, interaction complexity
  - how many is ok?
    - open research question
- reorderable lists
  - easy lookup
  - useful when linked to other encodings

Partition into views

• how to divide data between views
  – split into regions by attributes
  – encodes association between items using spatial proximity
  – order of splits has major implications for what patterns are visible

• no strict dividing line
  – view: big/detailed
    • contiguous region in which visually encoded data is shown on the display
  – glyph: small/iconic
    • object with internal structure that arises from multiple marks
Partitioning: List alignment

- single bar chart with grouped bars
  - split by state into regions
    - complex glyph within each region showing all ages
  - compare: easy within state, hard across ages

- small-multiple bar charts
  - split by age into regions
    - one chart per region
  - compare: easy within age, harder across states
Partitioning: Recursive subdivision

- split by neighborhood
- then by type
- then time
  - years as rows
  - months as columns
- color by price

- neighborhood patterns
  - where it’s expensive
  - where you pay much more for detached type

Partitioning: Recursive subdivision

- switch order of splits
  - type then neighborhood
- switch color
  - by price variation
- type patterns
  - within specific type, which neighborhoods inconsistent

System: HIVE

Partitioning: Recursive subdivision

- different encoding for second-level regions
  - choropleth maps

System: HIVE

Partitioning: Recursive subdivision

- size regions by sale counts
  - not uniformly
- result: treemap

Previous Demos

• Tableau Lessons
  – partitioning: drag multiple pills into Row or Column
  – disaggregation: drag field into Detail/Color
    • aggregation is automatic and aggressive in Tableau
Superimpose layers

• **layer**: set of objects spread out over region
  – each set is visually distinguishable group
  – extent: whole view

• design choices
  – how many layers, how to distinguish?
    • encode with different, nonoverlapping channels
    • two layers achievable, three with careful design
  – small static set, or dynamic from many possible?
Static visual layering

• foreground layer: roads
  – hue, size distinguishing main from minor
  – high luminance contrast from background

• background layer: regions
  – desaturated colors for water, parks, land areas

• user can selectively focus attention
• “get it right in black and white”
  – check luminance contrast with greyscale view

Superimposing limits

• few layers, but many lines
  – up to a few dozen
  – but not hundreds

• superimpose vs juxtapose: empirical study
  – superimposed for local, multiple for global
  – tasks
    • local: maximum, global: slope, discrimination
  – same screen space for all multiples vs single superimposed

Dynamic visual layering

• interactive, from selection
  – lightweight: click
  – very lightweight: hover

• ex: 1-hop neighbors

Reduce items and attributes

• reduce/increase: inverses

• filter
  – pro: straightforward and intuitive
    • to understand and compute
  – con: out of sight, out of mind

• aggregation
  – pro: inform about whole set
  – con: difficult to avoid losing signal

• not mutually exclusive
  – combine filter, aggregate
  – combine reduce, change, facet
Idiom: **dynamic filtering**

- item filtering
- browse through tightly coupled interaction
  — alternative to queries that might return far too many or too few

*System: FilmFinder*

Idiom: **DOSFA**

- attribute filtering
- encoding: star glyphs

Idiom: **histogram**

- static item aggregation
- task: find distribution
- data: table
- derived data
  - new table: keys are bins, values are counts
- bin size crucial
  - pattern can change dramatically depending on discretization
  - opportunity for interaction: control bin size on the fly
Continuous scatterplot

- static item aggregation
- data: table
- derived data: table
  - key attrs x,y for pixels
  - quant attrib: overplot density
- dense space-filling 2D matrix
- color: sequential categorical hue + ordered luminance colormap

Idiom: **scented widgets**

- augment widgets for filtering to show *information scent*
  - cues to show whether value in drilling down further vs looking elsewhere
- concise, in part of screen normally considered control panel

Idiom: **boxplot**

- static item aggregation
- task: find distribution
- data: table
- derived data
  - 5 quant attribs
    - median: central line
    - lower and upper quartile: boxes
    - lower upper fences: whiskers
      - values beyond which items are outliers
  - outliers beyond fence cutoffs explicitly shown

---

[40 years of boxplots. Wickham and Stryjewski. 2012. had.co.nz]
Idiom: **Hierarchical parallel coordinates**

- dynamic item aggregation
- derived data: *hierarchical clustering*
- encoding:
  - cluster band with variable transparency, line at mean, width by min/max values
  - color by proximity in hierarchy

Spatial aggregation

- MAUP: Modifiable Areal Unit Problem
  - gerrymandering (manipulating voting district boundaries) is one example!

[http://www.e-education.psu.edu/geog486/l4_p7.html, Fig 4.cg.6]
Dimensionality reduction

- attribute aggregation
  - derive low-dimensional target space from high-dimensional measured space
  - use when you can’t directly measure what you care about
    - true dimensionality of dataset conjectured to be smaller than dimensionality of measurements
    - latent factors, hidden variables

Tumor Measurement Data

data: 9D measured space

derived data: 2D target space
Idiom: Dimensionality reduction for documents

Task 1

Input: HD data
Output: 2D data

What?
- In High-dimensional data
- Out 2D data

Why?
- Produce
- Derive

Task 2

Input: 2D data
Output: Scatterplot Clusters & points

What?
- In 2D data
- Out Scatterplot
- Out Clusters & points

Why?
- Discover
- Explore
- Identify

How?
- Encode
- Navigate
- Select

Task 3

Input: Scatterplot Clusters & points
Output: Labels for clusters

What?
- In Scatterplot Clusters & points
- Out Labels for clusters

Why?
- Produce
- Annotate
Demo 2: Internet Use

• Credit: Ben Jones

• Tableau Lessons
  – more maps, dual axes
  – linked views (apply filter to selected worksheets)
  – actions: highlight/hover

• Big Ideas
  – Tableau interactivity defaults not necessarily what you want
Demo 3: House Price Index

• Credit: Robert Kosara, from TCC 2014 talk Recreating News Visualizations in Tableau

• Tableau Lessons
  – more calculated field practice
  – create parameter
  – reference lines
  – interactive sliders

• Big Ideas
  – calculated fields plus interactivity gives you a lot of power and flexibility
Assignment 4

• finish/review House Price Index workbook

• add interactivity to last week’s story
  – update workbook
  – upload to Tableau Public
  – revise story to include embedded interactive

• final project proposal