Outline

• **Session 1 10-11:30am**
  Data Visualization Pitfalls to Avoid
  – Introduction
  – Color
  – Space: 2D vs 3D

• **Session 2 12:30-3pm**
  Visualization Analysis & Design, In More Depth
  – Marks and Channels, Perception
  – Arrange Tables
  – Arrange Spatial Data
  – Arrange Networks
  – Manipulate: Change, Select, Navigate
  – Facet: Juxtapose, Partition, Superimpose
  – Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft
@tamaramunzner
What?

Why?

How?

### Datasets

<table>
<thead>
<tr>
<th>Types</th>
<th>Data Types</th>
<th>Data and Dataset Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tables</td>
<td>Items</td>
<td>Networks &amp; Trees</td>
</tr>
<tr>
<td></td>
<td>Attributes</td>
<td>Items (nodes)</td>
</tr>
<tr>
<td></td>
<td>Links</td>
<td>Positions</td>
</tr>
<tr>
<td></td>
<td>Positions</td>
<td>Grids</td>
</tr>
<tr>
<td></td>
<td>Grids</td>
<td>Items</td>
</tr>
<tr>
<td></td>
<td>Clusters, Sets, Lists</td>
<td>Items</td>
</tr>
</tbody>
</table>

#### Attributes

- **Attribute Types**
  - Categorical
  - Ordered
    - Ordinal
    - Quantitative

#### Ordering Direction

- Sequential
- Diverging
- Cyclic

#### Dataset Availability

- Static
- Dynamic
• \{action, target\} pairs
  – discover distribution
  – compare trends
  – locate outliers
  – browse topology

**Analyzing**

- **Consume**
  - Discover
  - Present
  - Enjoy

- **Produce**
  - Annotate
  - Record
  - Derive

**Searching**

<table>
<thead>
<tr>
<th>Target known</th>
<th>Target unknown</th>
</tr>
</thead>
</table>

- **Location known**
  - Lookup
  - Browse

- **Location unknown**
  - Locate
  - Explore

**Querying**

- **Identify**
- **Compare**
- **Summarize**

**Actions**

**Targets**

- **All Data**
  - Trends
  - Outliers
  - Features

- **Attributes**
  - One
    - Distribution
  - Many
    - Dependency
    - Correlation
    - Similarity

- **Network Data**
  - Topology
    - Paths

- **Spatial Data**
  - Shape
## How?

### Encode

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

### Manipulate

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

### Facet

- **Change**
  - ![Change Diagram]
- **Select**
  - ![Select Diagram]
- **Navigate**
  - ![Navigate Diagram]
- **Superimpose**
  - ![Superimpose Diagram]

### Reduce

- **Filter**
  - ![Filter Diagram]
- **Aggregate**
  - ![Aggregate Diagram]
- **Embed**
  - ![Embed Diagram]

### Why?

- **What?**
- **Why?**
- **How?**
Channels: Rankings

**Magnitude Channels: Ordered Attributes**

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

**Identity Channels: Categorical Attributes**

- Spatial region
- Color hue
- Motion
- Shape

- **expressiveness principle**
  - match channel and data characteristics

- **effectiveness principle**
  - encode most important attributes with highest ranked channels
Channels: Expressiveness types and effectiveness rankings

**Magnitude Channels: Ordered Attributes**

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

**Identity Channels: Categorical Attributes**

- Spatial region
- Color hue
- Motion
- Shape

- expressiveness principle
  - match channel and data characteristics
- effectiveness principle
  - encode most important attributes with highest ranked channels
  - spatial position ranks high for both
Accuracy: Fundamental Theory

Steven’s Psychophysical Power Law: $S = I^N$
Accuracy: Vis experiments

Discriminability: How many usable steps?

• must be sufficient for number of attribute levels to show
  – linewidth: few bins

[mappa.mundi.net/maps/maps_014/telegeography.html]
Separability vs. Integrality

<table>
<thead>
<tr>
<th>Position</th>
<th>Size</th>
<th>Width</th>
<th>Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Hue (Color)</td>
<td>+ Hue (Color)</td>
<td>+ Height</td>
<td>+ Green</td>
</tr>
</tbody>
</table>

- **Fully separable**: 2 groups each
- **Some interference**: 2 groups each
- **Some/significant interference**: 3 groups total: integral area
- **Major interference**: 4 groups total: integral hue
Popout

• find the red dot
  – how long does it take?

• parallel processing on many individual channels
  – speed independent of distractor count
  – speed depends on channel and amount of difference from distractors

• serial search for (almost all) combinations
  – speed depends on number of distractors
Popout

• many channels: tilt, size, shape, proximity, shadow direction, ...
• but not all! parallel line pairs do not pop out from tilted pairs
Grouping

- containment
- connection

Marks as Links

- Containment
- Connection

Identity Channels: Categorical Attributes

- Spatial region
- Color hue
- Motion
- Shape

- proximity
  - same spatial region
- similarity
  - same values as other categorical channels
Relative vs. absolute judgements

- perceptual system mostly operates with relative judgements, not absolute
  - that’s why accuracy increases with common frame/scale and alignment
  - Weber’s Law: ratio of increment to background is constant
    - filled rectangles differ in length by 1:9, difficult judgement
    - white rectangles differ in length by 1:2, easy judgement

Relative luminance judgements

- perception of luminance is contextual based on contrast with surroundings

http://persci.mit.edu/gallery/checkershadow
Relative color judgements

• color constancy across broad range of illumination conditions

http://www.purveslab.net/seeforyourself/
Further reading

  –Chap 5: Marks and Channels


Outline

• **Session 1 10-11:30am**
  Data Visualization Pitfalls to Avoid
  – Introduction
  – Color
  – Space: 2D vs 3D

• **Session 2 12:30-3pm**
  Visualization Analysis & Design, In More Depth
  – Marks and Channels, Perception
  – Arrange Tables
  – Arrange Spatial Data
  – Arrange Networks
  – Manipulate: Change, Select, Navigate
  – Facet: Juxtapose, Partition, Superimpose
  – Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft
@tamaramunzner
# How?

## Encode

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

- **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?**
Encode tables: Arrange space

Encode

- **Arrange**
  - Express
  - Separate
  - Order
  - Align

- Encode tables: Arrange space
Keys and values

- **key**
  - independent attribute
  - used as unique index to look up items
  - simple tables: 1 key
  - multidimensional tables: multiple keys

- **value**
  - dependent attribute, value of cell

- classify arrangements by key count
  - 0, 1, 2, many...

Express Values

- 1 Key
  - List
- 2 Keys
  - Matrix
- 3 Keys
  - Volume
- Many Keys
  - Recursive Subdivision
0 Keys

Express Values

1 Key
List

2 Keys
Matrix

3 Keys
Volume

Many Keys
Recursive Subdivision
Idiom: **scatterplot**

- **express values**
  - quantitative attributes
- **no keys, only values**
  - data
    - 2 quant attrs
  - mark: points
  - channels
    - horiz + vert position
- **tasks**
  - find trends, outliers, distribution, correlation, clusters
- **scalability**
  - hundreds of items

Some keys

Express Values

1 Key
List

2 Keys
Matrix

3 Keys
Volume

Many Keys
Recursive Subdivision
Some keys: Categorical regions

- **regions**: contiguous bounded areas distinct from each other
  - using space to *separate* (proximity)
  - following expressiveness principle for categorical attributes

- use ordered attribute to *order* and *align* regions

<table>
<thead>
<tr>
<th>1 Key</th>
<th>2 Keys</th>
<th>3 Keys</th>
<th>Many Keys</th>
</tr>
</thead>
<tbody>
<tr>
<td>List</td>
<td>Matrix</td>
<td>Volume</td>
<td>Recursive Subdivision</td>
</tr>
</tbody>
</table>

Separate  

Order  

Align
Idiom: **bar chart**

- one key, one value
  - data
    - 1 categ attrib, 1 quant attrib
  - mark: lines
  - channels
    - length to express quant value
    - spatial regions: one per mark
      - separated horizontally, aligned vertically
      - ordered by quant attrib
        » by label (alphabetical), by length attrib (data-driven)
  - task
    - compare, lookup values
  - scalability
    - dozens to hundreds of levels for key attrib
Separated and Aligned but not Ordered

LIMITATION: Hard to know rank. What’s the 4th most? The 7th?

[Slide courtesy of Ben Jones]
Separate, Aligned and Ordered

[Slide courtesy of Ben Jones]
Separated but not Ordered or Aligned

LIMITATION: Hard to make comparisons

[Slide courtesy of Ben Jones]
Idiom: **line chart**

- one key, one value
  - data
    - 2 quant attribs
  - mark: points
    - line connection marks between them
  - channels
    - aligned lengths to express quant value
    - separated and ordered by key attrib into horizontal regions
- task
  - find trend
    - connection marks emphasize ordering of items along key axis by explicitly showing relationship between one item and the next
Idiom: **line chart / dot plot**

- one key, one value
  - data
    - 2 quant attribs
  - mark: points
    - line connection marks between them
  - channels
    - aligned lengths to express quant value
    - separated and ordered by key attrib into horizontal regions
- task
  - find trend
    - connection marks emphasize ordering of items along key axis by explicitly showing relationship between one item and the next
- scalability
  - hundreds of key levels, hundreds of value levels
Choosing bar vs line charts

• depends on type of key attrib
  – bar charts if categorical
  – line charts if ordered
• do not use line charts for categorical key attribs
  – violates expressiveness principle
  • implication of trend so strong that it overrides semantics!
    – “The more male a person is, the taller he/she is”

Chart axes

• labelled axis is critical
• avoid cropping y-axis
  – include 0 at bottom left
  – or slope misleads

• dual axes controversial
  – acceptable if commensurate
  – beware, very easy to mislead!

http://www.thefunctionalart.com/2015/10/if-you-see-bullshit-say-bullshit.html
Idiom: connected scatterplots

• scatterplot with line connection marks
  – popular in journalism
  – horiz + vert axes: value attribs
  – line connection marks: temporal order
  – alternative to dual-axis charts
    • horiz: time
    • vert: two value attribs

• empirical study
  – engaging, but correlation unclear


http://steveharoz.com/research/connected_scatterplot/
Idiom: Indexed line charts

• data: 2 quant attires
  – 1 key + 1 value

• derived data: new quant value attrib
  – index
  – plot instead of original value

• task: show change over time
  – principle: normalized, not absolute

• scalability
  – same as standard line chart

https://public.tableau.com/profile/ben.jones#!/vizhome/CAStateRevenues/Revenues
Idiom: **Gantt charts**

- one key, two (related) values
  - data
    - 1 categ attrib, 2 quant attribs
  - mark: line
    - length: duration
- channels
  - horiz position: start time (+end from duration)
- task
  - emphasize temporal overlaps, start/end dependencies between items
- scalability
  - dozens of key levels
  - hundreds of value levels

Idiom: **Slopegraphs**

- two values
  - data
    - 2 quant value attribs
  - mark: point + line
    - line connecting mark between pts
- channels
  - 2 vertical pos: express attrib value
- task
  - emphasize changes in rank/value
- scalability
  - hundreds of value levels

https://public.tableau.com/profile/ben.jones#!/vizhome/Slopegraphs/Slopegraphs
Breaking conventions

- presentation vs exploration
  - engaging/evocative
  - inverted y axis
- blood drips down on Poe

https://public.tableau.com/profile/ben.jones#!/vizhome/EdgarAllanPoeViz/EdgarAllanPoeViz
2 Keys

Express Values ➔ 1 Key List ➔ 2 Keys Matrix ➔ 3 Keys Volume ➔ Many Keys Recursive Subdivision
Idiom: heatmap

- two keys, one value
  - data
    - 2 categ attribs (gene, experimental condition)
    - 1 quant attrib (expression levels)
  - marks: area
    - separate and align in 2D matrix
      - indexed by 2 categorical attributes
  - channels
    - color by quant attrib
      - (ordered diverging colormap)
  - task
    - find clusters, outliers
- scalability
  - 1M items, 100s of categ levels, ~10 quant attrib levels
Axis Orientation

- Rectilinear
- Parallel
- Radial
Idioms: **scatterplot matrix, parallel coordinates**

- **scatterplot matrix (SPLOM)**
  - rectilinear axes, point mark
  - all possible pairs of axes
  - scalability
    - one dozen attribs
    - dozens to hundreds of items

- **parallel coordinates**
  - parallel axes, jagged line representing item
  - rectilinear axes, item as point
    - axis ordering is major challenge
  - scalability
    - dozens of attribs
    - hundreds of items

---

Idioms: **radial bar chart, star plot**

- **radial bar chart**
  - radial axes meet at central ring, line mark

- **star plot**
  - radial axes, meet at central point, line mark

- **bar chart**
  - rectilinear axes, aligned vertically

- **accuracy**
  - length unaligned with radial
    - less accurate than aligned with rectilinear

Radial Orientation: Radar Plots

LIMITATION: Not good when categories aren’t cyclic

[Slide courtesy of Ben Jones]
“Radar graphs: Avoid them (99.9% of the time)”


[Slide courtesy of Ben Jones]
Idioms: *pie chart, polar area chart*

- **pie chart**
  - area marks with angle channel
  - accuracy: angle/area less accurate than line length
    - arclength also less accurate than line length

- **polar area chart**
  - area marks with length channel
  - more direct analog to bar charts

- **data**
  - 1 categ key attrib, 1 quant value attrib

- **task**
  - part-to-whole judgements
Idioms: **normalized stacked bar chart**

- task
  - part-to-whole judgements

- **normalized stacked bar chart**
  - stacked bar chart, normalized to full vert height
  - single stacked bar equivalent to full pie
    - high information density: requires narrow rectangle

- **pie chart**
  - information density: requires large circle

http://bl.ocks.org/mbostock/3887235
http://bl.ocks.org/mbostock/3886208
http://bl.ocks.org/mbostock/3886394.
Idiom: *glyphmaps*

- rectilinear good for linear vs nonlinear trends

- radial good for cyclic patterns

---

Orientation limitations

- rectilinear: scalability wrt #axes
  - 2 axes best
  - 3 problematic
    - more in afternoon
  - 4+ impossible
- parallel: unfamiliarity, training time
- radial: perceptual limits
  - angles lower precision than lengths
  - asymmetry between angle and length
  - can be exploited!

Layout Density

Dense

Arrange tables

Express Values

Separate, Order, Align Regions

Separate

Order

Align

1 Key
List

2 Keys
Matrix

3 Keys
Volume

Many Keys
Recursive Subdivision

Axis Orientation

Rectilinear

Parallel

Radial

Layout Density

Dense

Space-Filling

Rectilinear

Parallel

Radial

Space-Filling

Layout Density

1 Key
2 Keys
3 Keys
Many Keys

List
Matrix
Volume
Recursive Subdivision
Further reading

  —Chap 7: Arrange Tables


• A Brief History of Data Visualization. Friendly. 2008.
  http://www.datavis.ca/milestones
Outline

• **Session 1 10-11:30am**
  Data Visualization Pitfalls to Avoid
  – Introduction
  – Color
  – Space: 2D vs 3D

• **Session 2 12:30-3pm**
  Visualization Analysis & Design, In More Depth
  – Marks and Channels, Perception
  – Arrange Tables
  – Arrange Spatial Data
  – Arrange Networks
  – Manipulate: Change, Select, Navigate
  – Facet: Juxtapose, Partition, Superimpose
  – Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft @tamaramunzner
Idiom: choropleth map

- **use** given spatial data
  - when central task is understanding spatial relationships

- **data**
  - geographic geometry
  - table with 1 quant attribute per region

- **encoding**
  - use given geometry for area mark boundaries
  - sequential segmented colormap [more later]

http://bl.ocks.org/mbostock/4060606
Beware: Population maps trickiness!

[ https://xkcd.com/1138 ]
Population maps trickiness

• beware!
• absolute vs relative again
  • population density vs per capita
• investigate with Ben Jones Tableau Public demo
  • http://public.tableau.com/profile/ben.jones#!/vizhome/PopVsFin/PopVsFin

Are Maps of Financial Variables just Population Maps?
  • yes, unless you look at per capita (relative) numbers

[ https://xkcd.com/1138 ]
Idiom: Bayesian surprise maps

• use models of expectations to highlight surprising values
• confounds (population) and variance (sparsity)

https://medium.com/@uwdata/surprise-maps-showing-the-unexpected-e92b67398865
Outline

• **Session 1 10-11:30am**
  Data Visualization Pitfalls to Avoid
  – Introduction
  – Color
  – Space: 2D vs 3D

• **Session 2 12:30-3pm**
  Visualization Analysis & Design, In More Depth
  – Marks and Channels, Perception
  – Arrange Tables
  – Arrange Spatial Data
  – Arrange Networks
  – Manipulate: Change, Select, Navigate
  – Facet: Juxtapose, Partition, Superimpose
  – Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft  @tamaramunzner
Arrange networks and trees

Node–Link Diagrams
Connection Marks
✔️ NETWORKS✔️ TREES

Adjacency Matrix
Derived Table
✔️ NETWORKS✔️ TREES

Enclosure
Containment Marks
❌ NETWORKS✔️ TREES
Idiom: **force-directed placement**

- **visual encoding**
  - link connection marks, node point marks

- **considerations**
  - spatial position: no meaning directly encoded
    - left free to minimize crossings
  - proximity semantics?
    - sometimes meaningful
    - sometimes arbitrary, artifact of layout algorithm
    - tension with length
      - long edges more visually salient than short

- **tasks**
  - explore topology; locate paths, clusters

- **scalability**
  - node/edge density \( E < 4N \)

Idiom: **adjacency matrix view**

- **data:** network
  - transform into same data/encoding as heatmap
- **derived data:** table from network
  - 1 quant attrib
    - weighted edge between nodes
  - 2 categ attribs: node list x 2
- **visual encoding**
  - cell shows presence/absence of edge
- **scalability**
  - 1K nodes, 1M edges

---

**Figure 7.5:** Comparing matrix and node-link views of a five-node network.

- **a)** Matrix view.
- **b)** Node-link view. From [Henry et al. 07], Figure 3b and 3a.

Network matrix views can achieve very high information density, up to a limit of one thousand nodes and one million edges, just like cluster heatmaps and all other matrix views that use small area marks.

**Technique**

**network matrix view**

**Data Types**

**network**

**Derived Data**

**table:** network nodes as keys, link status between two nodes as values

**View Comp.**

**space:** area marks in 2D matrix alignment

**Scalability**

**nodes:** 1K

**edges:** 1M

---

7.1.3.3 Multiple Keys: Partition and Subdivide

When a dataset has only one key, then it is straightforward to use that key to separate into one region.


Connection vs. adjacency comparison

- **adjacency matrix strengths**
  - predictability, scalability, supports reordering
  - some topology tasks trainable
- **node-link diagram strengths**
  - topology understanding, path tracing
  - intuitive, no training needed
- **empirical study**
  - node-link best for small networks
  - matrix best for large networks
    - if tasks don’t involve topological structure!

Idiom: radial node-link tree

• data
  – tree

• encoding
  – link connection marks
  – point node marks
  – radial axis orientation
    • angular proximity: siblings
    • distance from center: depth in tree

• tasks
  – understanding topology, following paths

• scalability
  – 1K - 10K nodes

Idiom: treemap

• data
  – tree
  – 1 quant attrib at leaf nodes

• encoding
  – area containment marks for hierarchical structure
    – rectilinear orientation
  – size encodes quant attrib

• tasks
  – query attribute at leaf nodes

• scalability
  – 1M leaf nodes

Link marks: Connection and containment

- marks as links (vs. nodes)
  - common case in network drawing
  - 1D case: connection
    - ex: all node-link diagrams
    - emphasizes topology, path tracing
    - networks and trees
  - 2D case: containment
    - ex: all treemap variants
    - emphasizes attribute values at leaves (size coding)
    - only trees

Further reading

  – Chap 9: Arrange Networks and Trees


Outline

• Session 1 10-11:30am  
  Data Visualization Pitfalls to Avoid  
  – Introduction  
  – Color  
  – Space: 2D vs 3D

• Session 2 12:30-3pm  
  Visualization Analysis & Design, In More Depth  
  – Marks and Channels, Perception  
  – Arrange Tables  
  – Arrange Spatial Data  
  – Arrange Networks  
  – Manipulate: Change, Select, Navigate  
  – Facet: Juxtapose, Partition, Superimpose  
  – Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft  
@tamaramunzner
 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
How to handle complexity: 1 previous strategy + 3 more

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

- Manipulate
  - Change
    - Select
      - Navigate

- Facet
  - Juxtapose
    - Partition
      - Superimpose

- Reduce
  - Filter
    - Aggregate
      - Embed
Manipulate

Change over Time

Navigate

Item Reduction

Attribute Reduction

Select

Zoom

Geometric or Semantic

Slice

Pan/Translate

Cut

Constrained

Project
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 interpolation from one state to another
  - alternative to jump cuts, supports item tracking
  - best case for animation
  - staging to reduce cognitive load
- example: animated transitions in statistical data graphics


video: vimeo.com/19278444
Idiom: *Animated transitions - visual encoding change*

- smooth transition from one state to another
  - alternative to jump cuts, supports item tracking
  - best case for animation
  - staging to reduce cognitive load
Idiom: **Animated transition - tree detail**

- animated transition
  - network drilldown/rollup

[Collapsible Tree](https://bl.ocks.org/mbostock/4339083)
Idiom: **Animated transition - bar detail**

- example: hierarchical bar chart
  - add detail during transition to new level of detail

[Hierarchical Bar Chart](https://bl.ocks.org/mbostock/1283663)
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
Further reading

  –*Chap 11: Manipulate View*


Outline

• **Session 1 10-11:30am**
  Data Visualization Pitfalls to Avoid
  – Introduction
  – Color
  – Space: 2D vs 3D

• **Session 2 12:30-3pm**
  Visualization Analysis & Design, In More Depth
  – Marks and Channels, Perception
  – Arrange Tables
  – Arrange Spatial Data
  – Arrange Networks
  – Manipulate: Change, Select, Navigate
  – Facet: Juxtapose, Partition, Superimpose
  – Reduce: Filter, Aggregate

http://www.cs.ubc.ca/~tmnr/talks.html#vad17can-aft  @tamaramunzner
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

Idiom: **bird’s-eye maps**

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

- differences
  - viewpoint
  - (size)

- **overview-detail**

**System: Google Maps**

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>Same</td>
<td>All: Redundant</td>
</tr>
<tr>
<td>Different</td>
<td>Multiform</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

System: HIVE

Partitioning: Recursive subdivision

• switch order of splits
  – type then neighborhood

• switch color
  – by price variation

• type patterns
  – within specific type, which neighborhoods inconsistent

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

System: HIVE

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

Idiom: **Trellis plots**

- superimpose within same frame
- color code by year
Dynamic visual layering

- interactive, from selection
  - lightweight: click
  - very lightweight: hover

- ex: 1-hop neighbors

Dynamic visual layering

• one-hop neighbour highlighting demos: click vs hover

Further reading

  – Chap 12: Facet Into Multiple Views


Outline

• **Session 1 10-11:30am**
  Data Visualization Pitfalls to Avoid
  – Introduction
  – Color
  – Space: 2D vs 3D

• **Session 2 12:30-3pm**
  Visualization Analysis & Design, In More Depth
  – Marks and Channels, Perception
  – Arrange Tables
  – Arrange Spatial Data
  – Arrange Networks
  – Manipulate: Change, Select, Navigate
  – Facet: Juxtapose, Partition, Superimpose
  – Reduce: Filter, Aggregate

[http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft](http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft)
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

Reducing Items and Attributes

Filter

Items

Attributes

Aggregate

Items

Attributes
Idiom: **dynamic filtering**

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

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

• augmented widgets show information scent
  – cues to show whether value in drilling down further vs looking elsewhere

• concise use of space: histogram on slider
Scented histogram bisliders: detailed

Idiom: **cross filtering**

- item filtering
- coordinated views/controls combined
  - all scented histogram bisliders update when any ranges change

System: **Crossfilter**

[http://square.github.io/crossfilter/]
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]
Spatial aggregation

• MAUP: Modifiable Areal Unit Problem
  – gerrymandering (manipulating voting district boundaries) is only one example!
  – zone effects

  ![Image](http://www.e-education.psu.edu/geog486/l4_p7.html, Fig 4.cg.6)

  – scale effects

  ![Image](https://blog.cartographica.com/blog/2011/5/19/the-modifiable-areal-unit-problem-in-gis.html)
Further reading

  – Chap 13: Reduce Items and Attributes


More Information

- this talk
  http://www.cs.ubc.ca/~tmm/talks.html#vad17can-aft

- book page (including tutorial lecture slides)
  http://www.cs.ubc.ca/~tmm/vadbook
  – 20% promo code for book+ebook combo: HVN17

  – illustrations: Eamonn Maguire

- papers, videos, software, talks, courses
  http://www.cs.ubc.ca/group/infovis
  http://www.cs.ubc.ca/~tmm

Visualization Analysis and Design.