Data Visualization (CIS/DSC 468)

Networks

Dr. David Koop
Arrange Tables

Express Values

Separate, Order, Align Regions

Separate
Order
Align

1 Key
2 Keys
3 Keys
Many Keys

List
Recursive Subdivision
Volume
Matrix

Rectilinear
Parallel
Radial

Dense
Space-Filling

[Munzner (ill. Maguire), 2014]
Scatterplot & Axes

Wickham, 2014
Bar Charts, Dot Charts, and Line Charts

![Bar Chart: Height (inches) vs. Gender](image1)

- Female: 60 inches
- Male: 50 inches

![Bar Chart: Height (inches) vs. Age](image2)

- 10-year-olds: 50 inches
- 12-year-olds: 40 inches

![Line Chart: Height (inches) vs. Gender](image3)

- Female: 40 inches
- Male: 30 inches

![Line Chart: Height (inches) vs. Age](image4)

- 10-year-olds: 30 inches
- 12-year-olds: 20 inches

[Zacks and Tversky, 1999, Munzner (ill. Maguire), 2014]
Stream graphs


Sources: Baseline StudioSystems; Box Office Mojo

Mathew Bloch, Lee Byron, Shan Carter and Amanda Cox
Banking

Sunspot Cycles
Aspect Ratio = 3.96

![Graph showing sunspot cycles from 1700 to 1950 with two plots, one showing low-frequency oscillations and the other bringing individual cycles into greater relief.]

Aspect Ratio = 22.35

[Heer and Agrawala, 2006]
Heatmaps

• Data: Two keys, one quantitative attribute
• Task: Find clusters, outliers, summarize
• How: area marks in grid, color encoding of quantitative attribute
• Scalability: number of pixels for area marks (millions), <12 colors
• Red-green color scales often used
  - Be aware of colorblindness!

[fastpitchanalytics.com]
Cluster Heatmap

[File System Similarity, R. Musăloiu-E., 2009]
Bertin’s Encodings

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>text</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Text</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Grayscale</strong></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Circle</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dual bar chart</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bar chart</strong></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Line</strong></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Black and white bar chart</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average bar chart</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[C.Perrin et al., 2014]
Assignment 2


- Notes:
  - Be careful about what `selectAll` is doing
  - In the updated refugees array, every object has a new property named "Total"
Exam 1

• Wednesday, March 1 in class (12-12:50pm)
• Format:
  - Multiple Choice
  - Short Answer
• Sample questions on web site:
Radial Axes
Radial Axes

- Polar Coordinates (angle + position along the line at that angle)
- What types of encodings are possible for tabular data in polar coordinates?
Radial Axes

• Polar Coordinates (angle + position along the line at that angle)

• What types of encodings are possible for tabular data in polar coordinates?
  - Radial bar charts
  - Pie charts
  - Donut charts
Pie Charts

• Part-Of-Whole Relationship
• Comparing Wedges
  - By angle, area, or arc length?
  - vs. bar charts
• Space-usage?

[Pie Chart, Bostock, 2017]
Part-of-whole: Relative % comparison?

(Stacked Bar Chart, Bostock, 2017)
Normalized Stacked Bar Chart

[Normalized Stacked Bar Chart, Bostock, 2017]
Scatterplot Matrix (SPLOM)

• Data: Many quantitative attributes
• Derived Data: names of attributes
• Task: Find correlations, trends, outliers
• How: Scatterplots in matrix alignment
• Scale: attributes: ~12, items: hundreds?

• Visualizations in a visualization: at high level, marks are themselves visualizations...

[R command from Wikipedia]
Spatial Axis Orientation

• So far, we have seen the vertical and horizontal axes (a *rectilinear* layout) used to encode almost everything

• What other possibilities are there for axes?

[Munzner (ill. Maguire), 2014]
Spatial Axis Orientation

• So far, we have seen the vertical and horizontal axes (a **rectilinear** layout) used to encode almost everything

• What other possibilities are there for axes?
  - Parallel axes

[Munzner (ill. Maguire), 2014]
Spatial Axis Orientation

• So far, we have seen the vertical and horizontal axes (a **rectilinear** layout) used to encode almost everything

• What other possibilities are there for axes?
  - Parallel axes
  - Radial axes

[Munzner (ill. Maguire), 2014]
Parallel Coordinates

• Data: many quantitative attributes
• Task: Find trends, extremes, correlation
• How: vertical spatial position for each attribute, connection marks for identity, axes horizontally spaced
• Scalability: <40 attributes, hundreds of values

• Connection marks help visualize trends between particular values
• Ordering the horizontal axes is important
• Not as well-known, often requires learning
Comparing SPLOMs and Parallel Coordinates

Scatterplot Matrix

Parallel Coordinates

100
90
80
70
60
50
40
30
20
10
0

Math Physics Dance Drama

Math

Physics

Dance

Drama

Math

Physics

Dance

Drama

[100, 90, 80, 70, 60, 50, 40, 30, 20, 10, 0]

19

D. Koop, CIS 468, Spring 2017

[Munzner (ill. Maguire), 2014]
Correlation in Parallel Coordinates

[Wegman, 1990]
Overdraw in Parallel Coordinates

[Fua et al., 1999]
Hierarchical Parallel Coordinates

[Figure 4: This image sequence shows a Fatal Accident data set of 230,000 data elements at different levels of details. The first image shows a cut across the root node. The last image shows the cut chaining all the leaf nodes. The rest of the images show intermediate cuts at varying levels of detail. (See Color Plates).

Figure 6: Left image shows Iris data set without proximity-based coloring. Right image shows Iris data set with proximity-based coloring revealing three distinct clusters depicted by concentrations of blue, green and pink lines. (See Color Plates).

[Fua et al., 1999]
Networks

• Why not graphs?
  - Bar graph
  - Graphing functions in mathematics

• Network: nodes and edges connecting the nodes

• Formally, $G = (V,E)$ is a set of nodes $V$ and a set of edges $E$ where each edge connects two nodes.

• Nodes == items, edges connect items

• **Both** nodes and edges may have attributes
Arrange Networks and Trees

- **Node–Link Diagrams**
  - Connection Marks
  - ![Networks and Trees]

- **Adjacency Matrix**
  - Derived Table
  - ![Networks and Trees]

- **Enclosure**
  - Containment Marks
  - ![Networks and Trees]

[Munzner (ill. Maguire), 2014]
Molecule Network
Molecule Network

- Nodes may have attributes (e.g. element)
Molecule Network

- Nodes may have attributes (e.g. element)
- Edges may have attributes (e.g. number of bonds)
Web Sites as Networks (amazon.com)

[M. Salathe, 2006]
Social Networks

[P. Butler, 2010]
Networks as Data

**Nodes**

<table>
<thead>
<tr>
<th>ID</th>
<th>Atom</th>
<th>Electrons</th>
<th>Protons</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>N</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>C</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>S</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>N</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

**Edges**

<table>
<thead>
<tr>
<th>ID1</th>
<th>ID2</th>
<th>Bonds</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>
Node-Link Diagrams

- Data: nodes and edges
- Task: understand connectivity, paths, structure (topology)
- Encoding: nodes as point marks, connections as line marks
- Scalability: hundreds

- ...but high **density** of links can be problematic! Why?

- How do we generate the visualization?
Arc Diagram

[D. Eppstein, 2013]
Network Layout

• Need to use spatial position when designing network visualizations
• Otherwise, nodes can **occlude** each other, links hard to distinguish
• How?
  - With bar charts, we could order using an attribute…
  - With networks, we want to be able to see connectivity and topology (not in the data usually)
• Possible metrics:
  - Edge crossings
  - Node overlaps
  - Total area
Force-Directed Layout

- Nodes push away from each other but edges are springs that pull them together
- Weakness: nondeterminism, algorithm may produce different results each time it runs

[Force-Directed Layout, M. Bostock, 2012]