Week 6:
Networks, Stories, Vis in the Newsroom

Tamara Munzner

Department of Computer Science
University of British Columbia

JRNL 520H, Special Topics in Contemporary Journalism: Data Visualization
Week 6: 18 October 2016

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

News

- Assignment 3, 4, 4-solo marks out
  - Assign 4: 90% pair mark, 10% solo mark for proposal
    - Assign 4 pair: min 74.8, avg 84.2, max 100
    - Assign 4 solo: min 50, avg 89.7, max 100 (5% of Assign 6 grade weight)
  - Assign 3
    - min 83.8, avg 93.6, max 100
- things to watch out for
  - diverging vs sequential colormaps
  - line charts vs bar charts (continuous vs discrete data)
  - absolute counts vs relative percentages
  - highlighting/emphasis vs filtering out completely
    - does your interaction make sense? does it help somebody?
  - please submit packaged workbooks (twbx) not plain workbooks (twb)

Schedule

- today: office hours 12:30-1:30pm, Tamara & Caitlin
- next week:
  - Tamara on travel Sat Oct 22 - Sat Oct 29
    - at VIS conference in Baltimore, likely extremely slow with email
  - Caitlin here
    - Tue Oct 25 9:30-12:30, 1:30-4:30 in Sing Tao bldg room 313, drop by for help/discussion
    - available by email throughout the week
- two weeks:
  - project 6 due Tue Nov 1 9am

Today

- stories
- networks
- (break)
- vis in the news
- beyond this class
- individual meetings on final project

Assignment 6

- find dataset, visualize it, and write story about it
  - you've now had practice in
    - effective visual encoding: space, color
    - finding the story within a dataset
    - wrangling
    - linking up and partitioning into multiple views
- you're encouraged to consult with us if you get stuck!
  - is your idea viable/newsworthy?
  - how can you do what you want inside Tableau?
  - is your visual encoding well justified?
- you're encouraged to post story publicly (but not required)
  - note you can embed vis within web page with Tableau Public

Last Time

Last Time: Rules of Thumb

- No unjustified 3D
- Resolution over immersion
- Overview first, zoom and filter, details on demand
- Responsiveness is required
- Function first, form next

Demos 1 & 2: Wrangling Tutorial, Simple Survey

- Credit: Caitlin Havlak

- Wrangling Lessons
  - first row for headers (right menu over source in
  - Tableau data interpeter
  - manual Excel/GoogleDoc cleaning

- Big Ideas
  - reshaping data: from wide to tall
  - joins: inner, left, right, outer
  - pivots: one observation per row, no cross-tabulation

Stories

Networks

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**
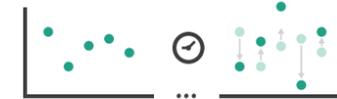
→ **Select**

→ **Navigate**

Facet

→ **Juxtapose**

→ **Partition**

→ **Superimpose**

Reduce

→ **Filter**
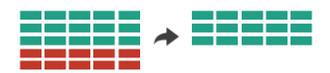
→ **Aggregate**

→ **Embed**

What?

Why?

How?

11
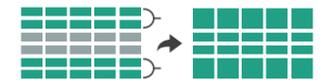
Arrange networks and trees

➔ **Node–Link Diagrams**
Connection Marks

✔ NETWORKS   ✔ TREES
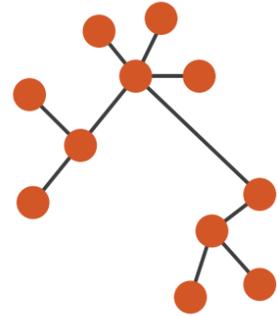
➔ **Adjacency Matrix**
Derived Table

✔ NETWORKS   ✔ TREES
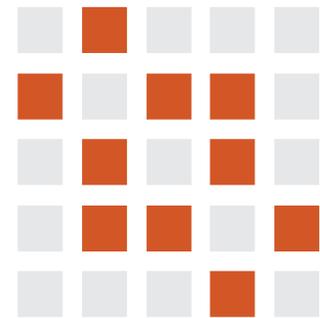
➔ **Enclosure**
Containment Marks

✘ NETWORKS   ✔ TREES
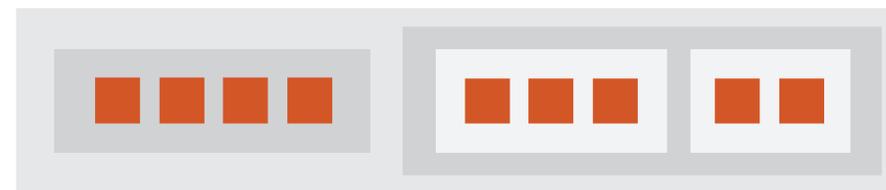
Idiom: **force-directed placement**

- 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

```
var width = 960,
    height = 500;

var color = d3.scale.category20();

var force = d3.layout.force()
```

http://mbostock.github.com/d3/ex/force.html

13
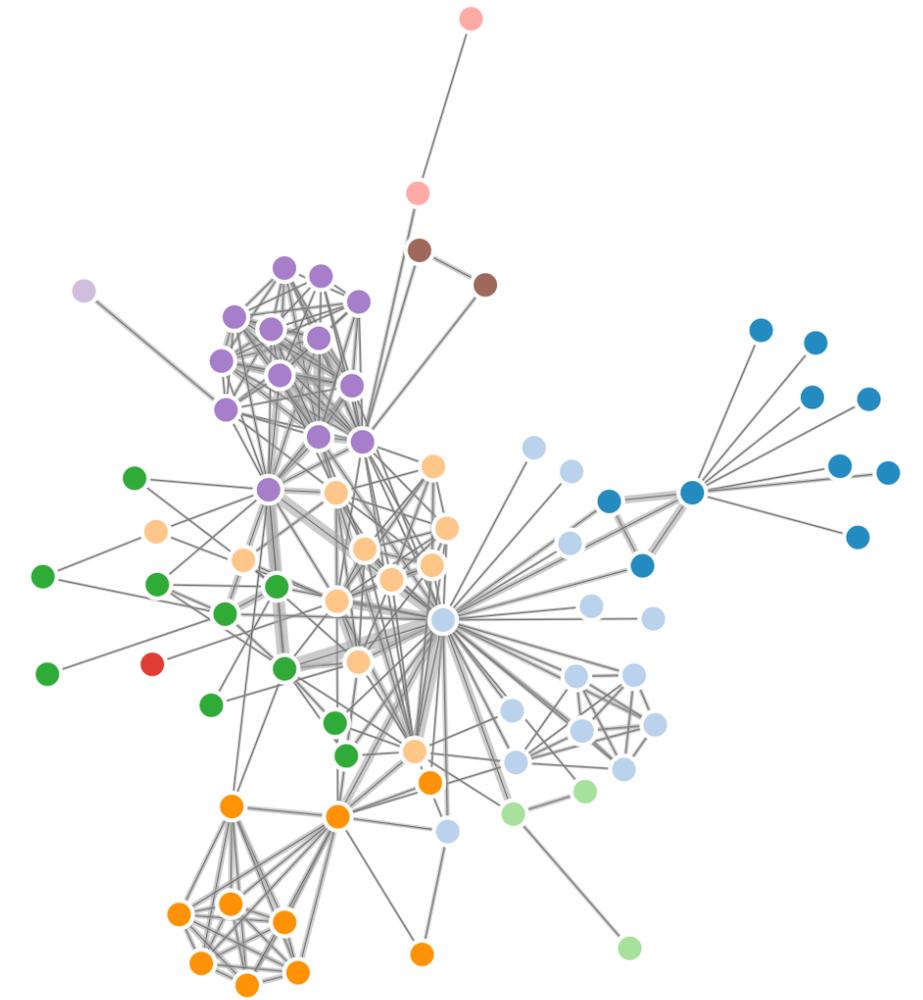
Idiom: **adjacency matrix view**

- 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
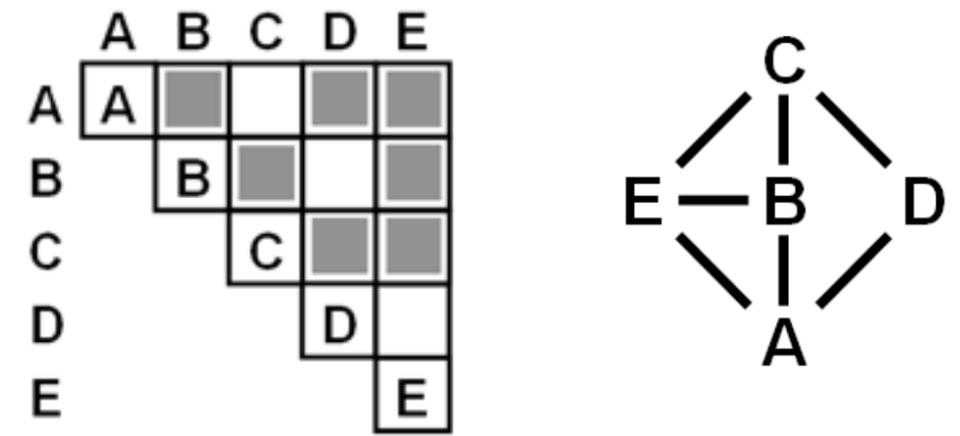
[NodeTrix: a Hybrid Visualization of Social Networks. Henry, Fekete, and McGuffin. IEEE TVCG (Proc. InfoVis) 13(6):1302-1309, 2007.]

[Points of view: Networks. Gehlenborg and Wong. Nature Methods 9:115.]

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!

http://www.michaelmcguffin.com/courses/vis/patternsInAdjacencyMatrix.png
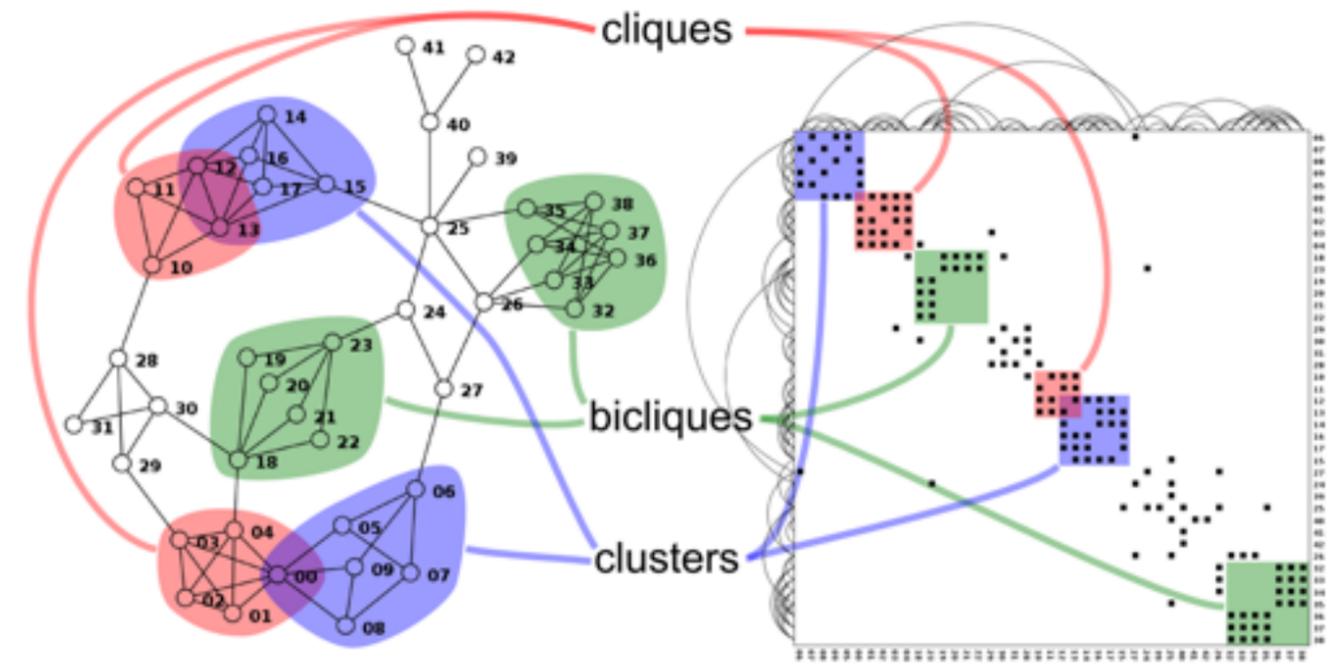
[On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. Ghoniem, Fekete, and Castagliola. Information Visualization 4:2 (2005), 114–135.]

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
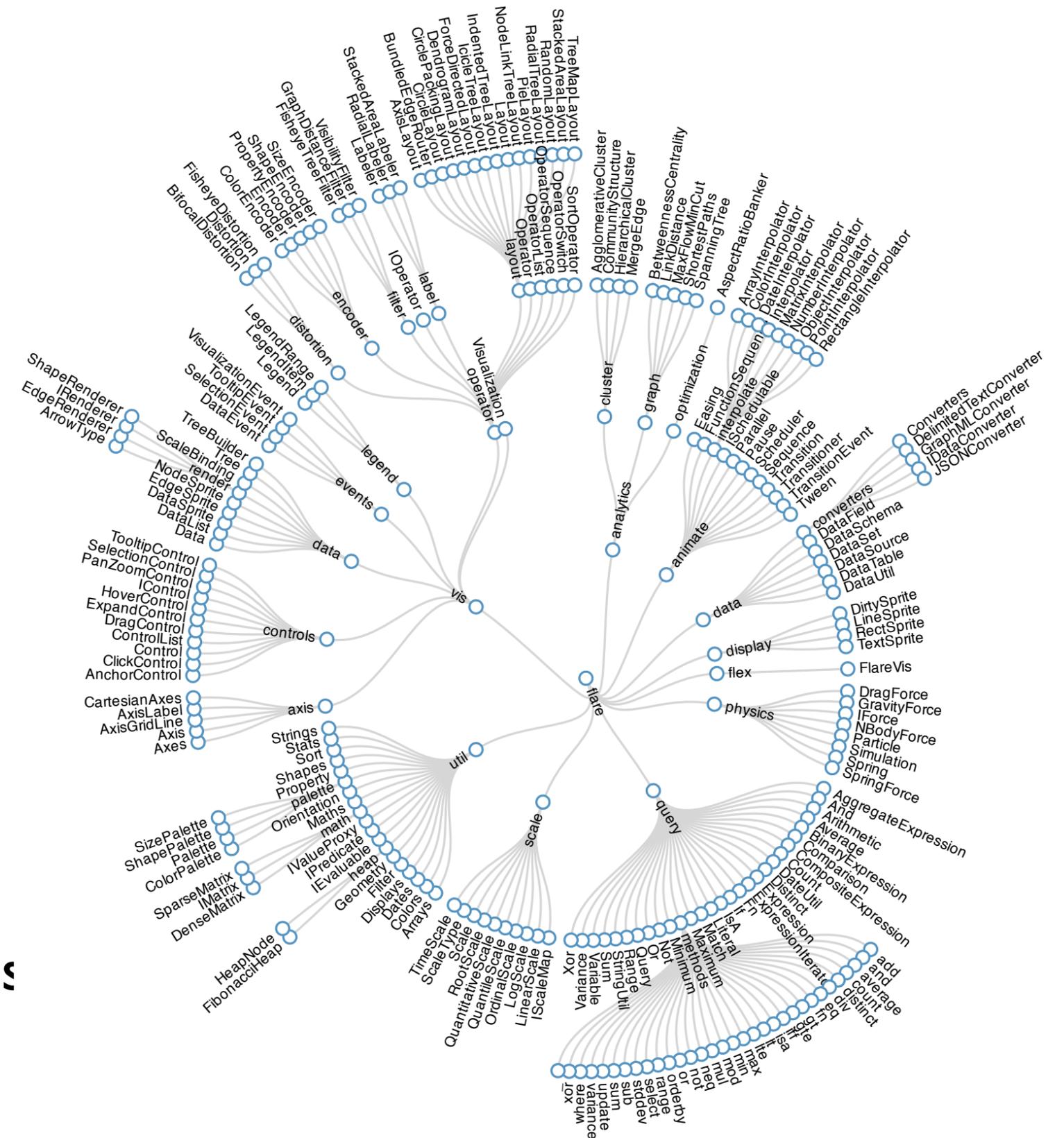
Idiom: **treemap**

- 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
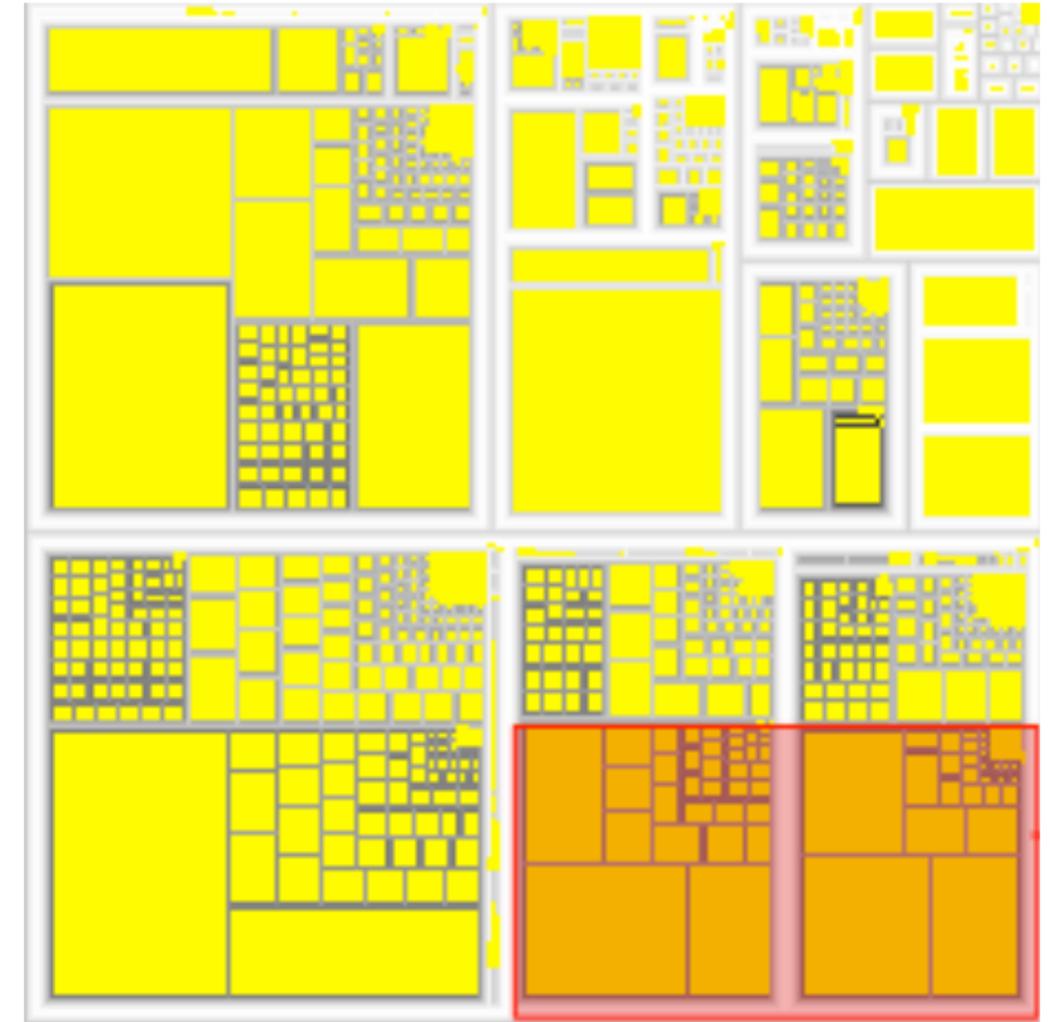
http://tulip.labri.fr/Documentation/3_7/userHandbook/html/ch06.html

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
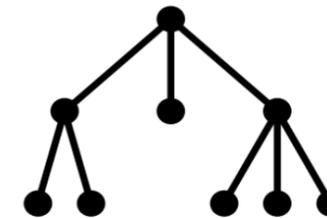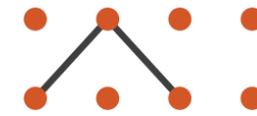
**Connection**   **Containment**
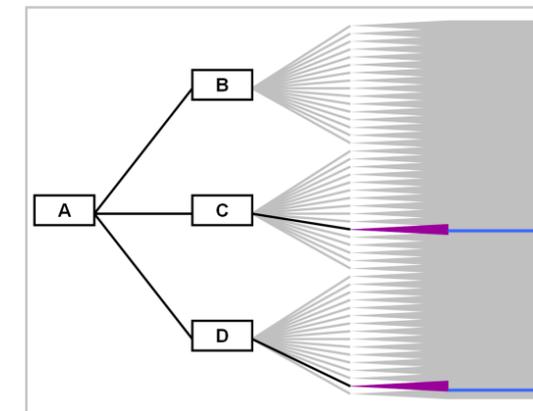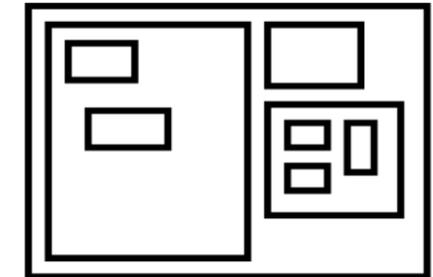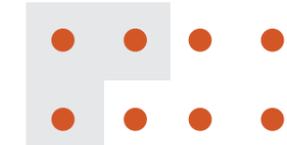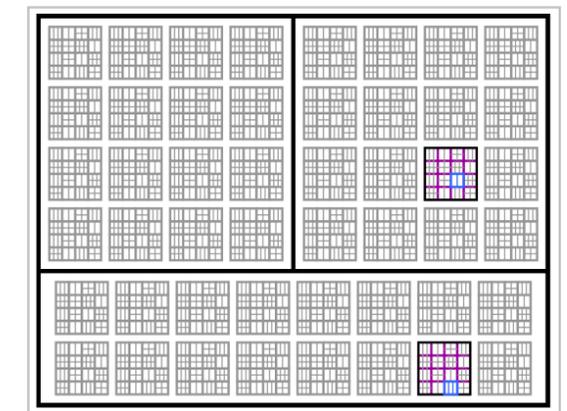
Node–Link Diagram   **Treemap**

[Elastic Hierarchies: Combining Treemaps and Node-Link Diagrams. Dong, McGuffin, and Chignell. Proc. InfoVis 2005, p. 57-64.]

Tree drawing idioms comparison

- data shown
  - link relationships
  - tree depth
  - sibling order

- design choices
  - connection vs containment link marks
  - rectilinear vs radial layout
  - spatial position channels

- considerations
  - redundant? arbitrary?
  - information density?
    - avoid wasting space
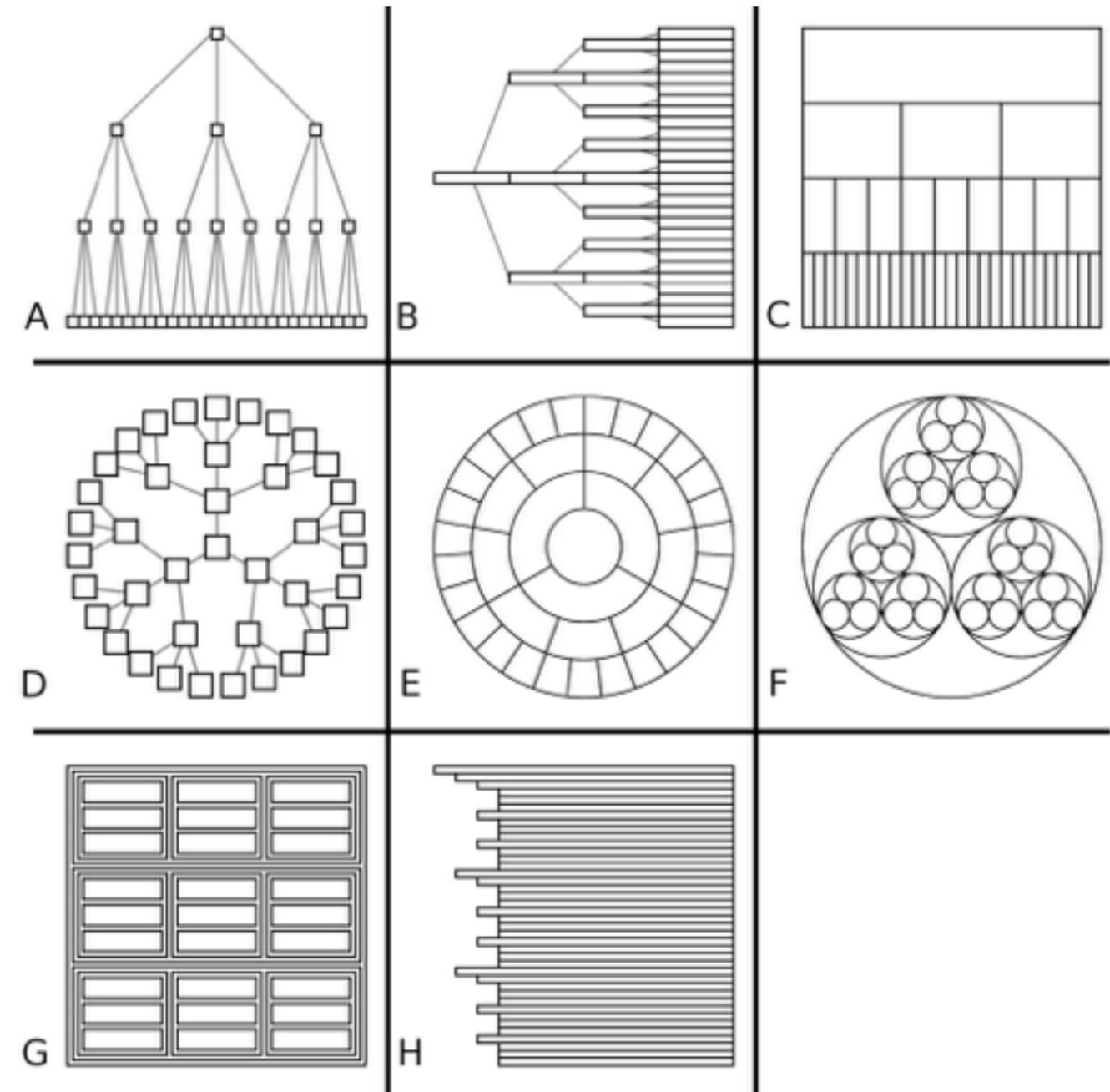
[Quantifying the Space-Efficiency of 2D Graphical Representations of Trees. McGuffin and Robert. Information Visualization 9:2 (2010), 115–140.]