UVA DS2003 Communicating with Data
Fall 2026

Networks and Trees

Sarah groves
based on materials by Hannah Bako and Zhicheng Liu

Finish up activity from last class: 10 min

- Use slides from last class to finish example 1
- If you have finished your example 1, look at another example and fill out the google doc for advice (things to copy and things to avoid)

Scrollytelling examples to explore

- **A Visual Introduction to Machine Learning** R2D3

  Scroller: one chart transforms as you read

- **Parable of the Polygons** Vi Hart & Nicky Case

  Explorable: you drag the shapes and run the model yourself

- **Film Dialogue by Gender** The Pudding

  Story first, then search and filter 2,000 films (martini glass)

- **Women's Pockets are Inferior** The Pudding

  Scroll essay built from measuring 80 pairs of jeans

- **Human Terrain** The Pudding

  Scroll-driven 3D map of world population, 1975 to 2015

NYT scrollytelling examples

- **Snow Fall: The Avalanche at Tunnel Creek** 2012

  The piece that made scrollytelling a genre: text, video and maps woven together

- **How the Recession Reshaped the Economy, in 255 Charts** 2014

  A guided scroll first, then all 255 charts to explore

- **You Draw It: How Family Income Predicts Children's College Chances** 2015

  You draw your guess before the data is revealed

- **How Much Hotter Is Your Hometown Than When You Were Born?** 2018

  Personalized: the story is built from your city and birth year

Dashboards

- **Gapminder Tools** Gapminder

  Income, health and population for every country since 1800

- **COVID-19 Data Explorer** Our World in Data

  Pick a metric, countries and a chart type

- **Unemployment Rate** FRED, St. Louis Fed

  75+ years of U.S. unemployment, plus thousands of other series

- **earth** nullschool

  Live global wind, weather and ocean conditions

- **Explore Census Data** U.S. Census Bureau

  Search any Census table, map or profile

Notebooks

- **Lego Price Explorer** marimo

  A Python notebook running in your browser: Lego set prices, 1970 to 2022

- **Movies Dashboard** marimo

  Pick a date range or a decade to explore films

- **Zoomable Sunburst** Observable

  Hierarchical data: you'll build sunbursts in Lab 9

- **Brushable Scatterplot Matrix** Observable

  Every pair of variables in one grid

Takeaways for your notebook

- **Hidden interaction** → say what to do: “drag to select”, “pick a country”
- **Scroll steps** → one markdown cell and one chart per idea
- **Martini glass** → author-driven About section, then the reader-driven Interactive section
- **Highlight the current step** → active titles, annotations, one highlighted series
- **Reader picks a path** → a dropdown that changes the chart's focus
- **Animated transition** → one chart and a slider (e.g., year), not ten charts

Agenda

- Hierarchical data, concepts & analysis tasks
- Node-link diagrams
  - Indentation, spatial layout, focus+context
- Space-filling techniques
  - Treemap and variants
- Layering techniques
  - Icicle Plots and Sunburst

Hierarchical data (trees) and networks

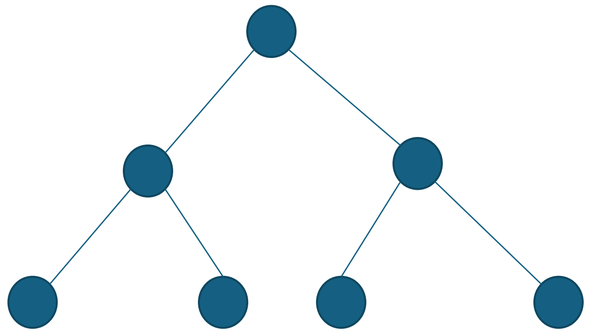
- A system of ranking data components, in which certain data components are parents or ancestors of the others
- Directed, acyclic
- Examples?
  - Family histories, genealogy
  - File systems on computers
  - Organization charts
  - Biology taxonomy: domain, kingdom, phylum, ….
  - Object-oriented programming design

Tasks

- Describe/understand structure
- Find items
  - Where are these data values located?
  - Where is this sub-hierarchy?
- Follow structure
  - What are the parent/children/siblings of x?
- Topological distance
  - How far are two things apart?
- Clustering/grouping
  - How similar are the nodes/sub-hierarchies?

Node-Link Diagrams

- Nodes assigned spatial locations
- Connected by straight or curved lines
- For 2D design, typically use one dimension (e.g. horizontal space) for breadth, another dimension (e.g. vertical space) for depth

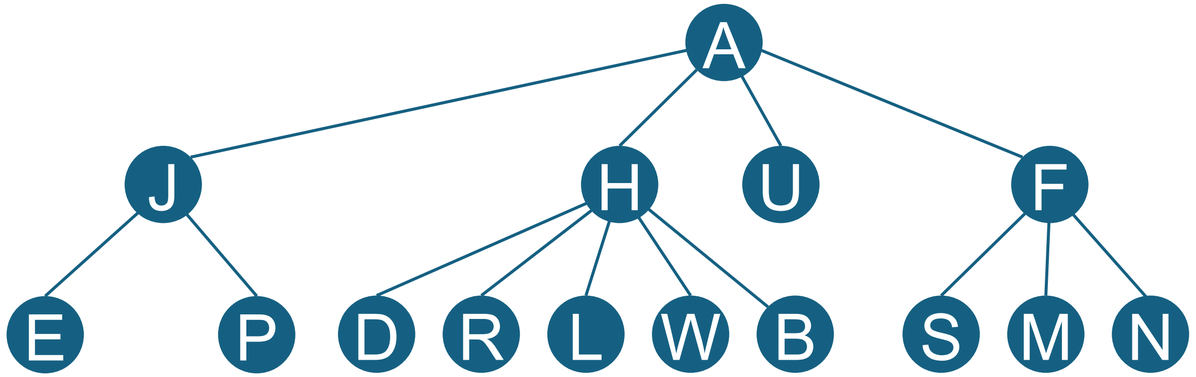


Breadth: how many children (each node has 2 children)
Depth: how many levels (3)

Exercise

Draw a representation for the following hierarchy:
Parent Node: Child1, Child2, … (order of children means nothing)

```
A: J, H, U, F
J: E, P
H: D, R, L, W, B
F: S, M, N
E: T, K
P: V, C, O, I
S: Q, G
```



A: J, H, U, F

J: E, P

H: D, R, L, W, B

F: S, M, N

E: T, K

P: V, C, O, I

S: Q, G

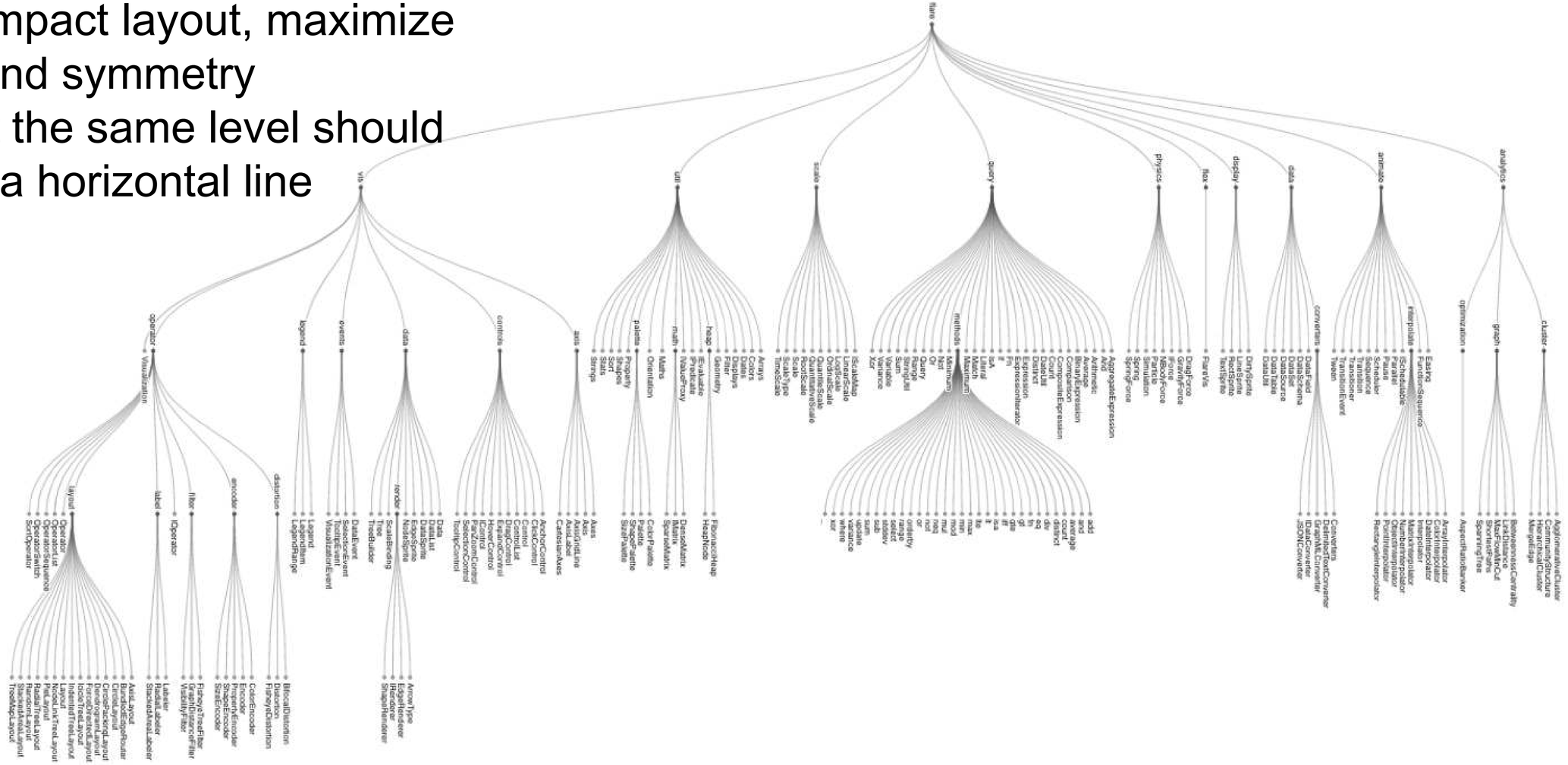
Potential Problems

- For top-down, width of fan-out uses up horizontal real estate very quickly
  - At level n, there are 2^n nodes
- Tree might grow a lot along one particular branch
  - Hard to draw it well in view without knowing how it will branch

Slide courtesy of John Stasko

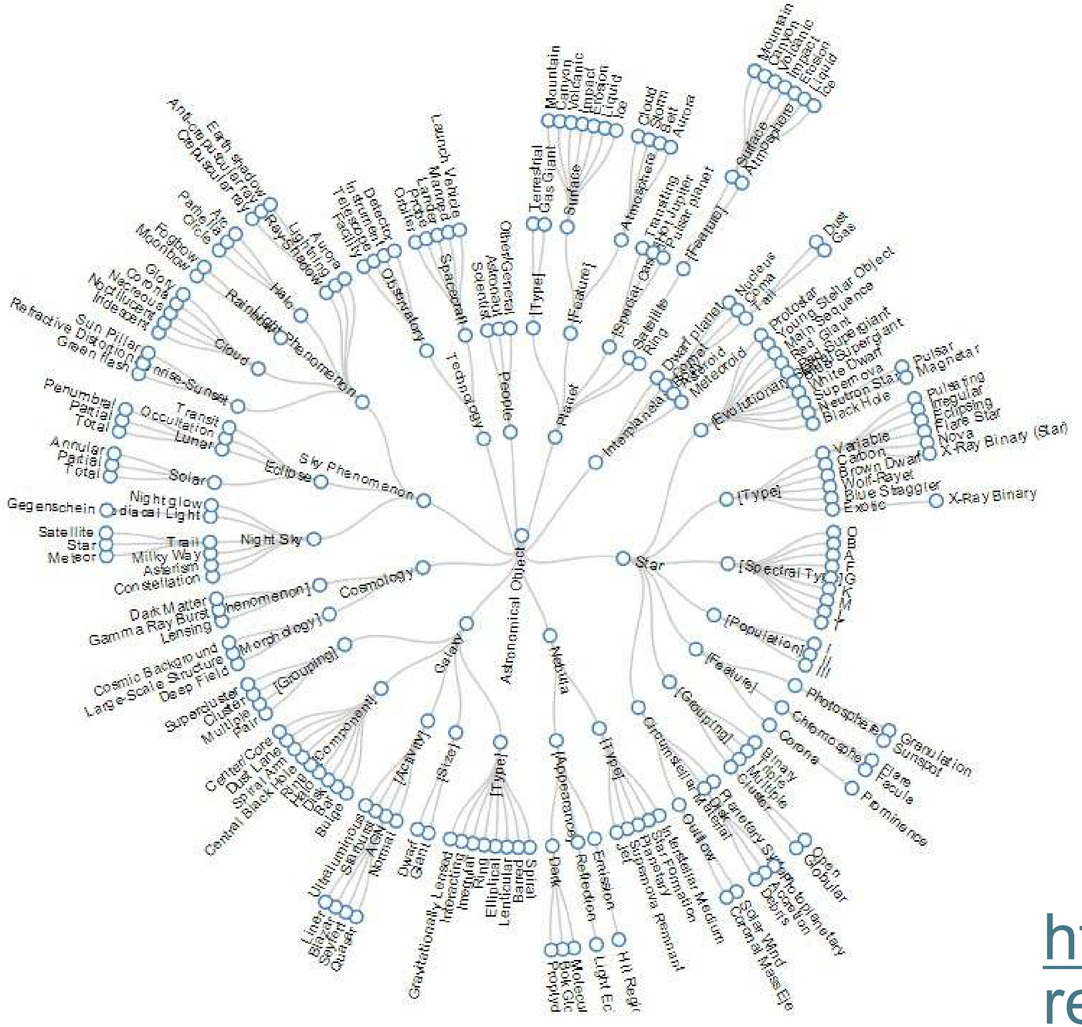
Reingold and Tilford Algorithm

Goal: compact layout, maximize density and symmetry
Nodes at the same level should lie along a horizontal line



https://observablehq.com/@d3/tidy-tree

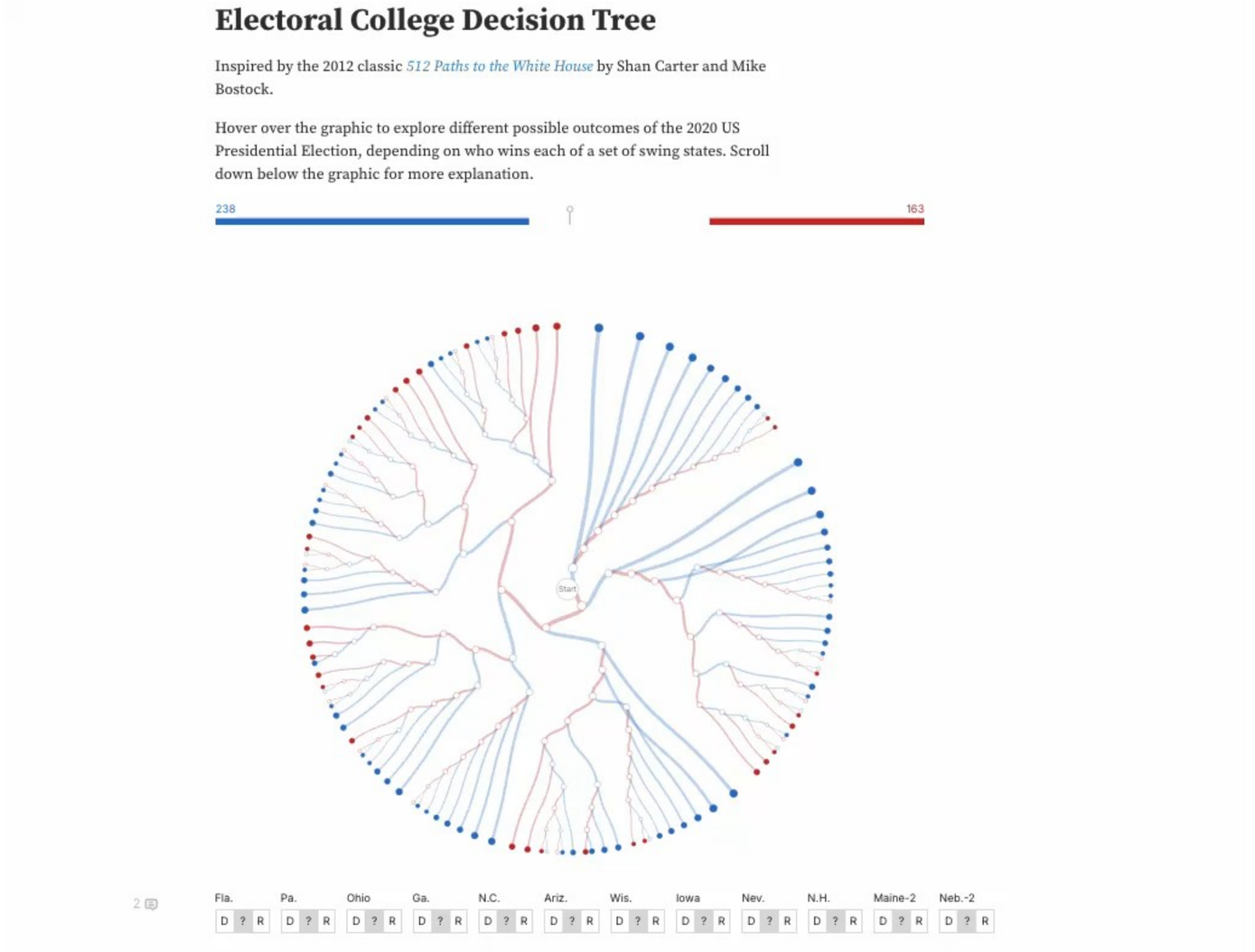
Radial Tree Layout



- Radius encodes depth, with root in the center
- Angular sectors assigned to subtrees (often with naïve recursive layout)
- Reingold-Tilford method can also be applied here

https://vega.github.io/vega/examples/radial-tree-layout/

Electoral College Decision Tree



https://observablehq.com/@observablehq/electoral-college-decision-tree

What if we want to show more attributes?

Node: Child1, Child2, … (order means nothing)

```
A: J, H, U, F
J: E, P
H: D, R, L, W, B
F: S, M, N
E: T, K
P: V, C, O, I
S: Q, G
```

```
Values:
B: 4
C: 6
D: 3
G: 2
I: 5
K: 8
L: 7
M: 1
N: 1
O: 6
Q: 2
R: 4
T: 3
U: 6
V: 4
W: 2
```

Difficult to encode more data attributes in a node-link diagram

- Color, shape, size: take up space and clash with basic node-link structure

Take 5 minutes to sketch out this data in a way that represents the values as well.

Space-filling Representations

- Treemaps
- Sunburst
- Icicle Plot
- Circle packing

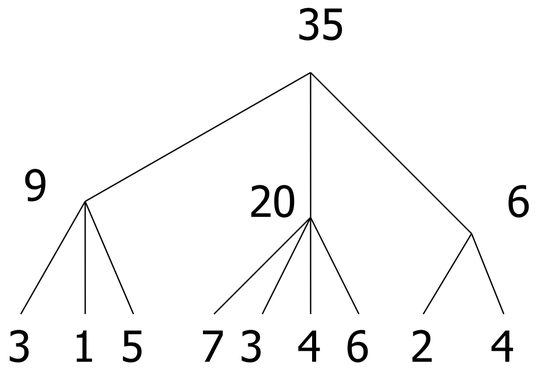
➢Each item occupies an area

➢Children are contained in the area

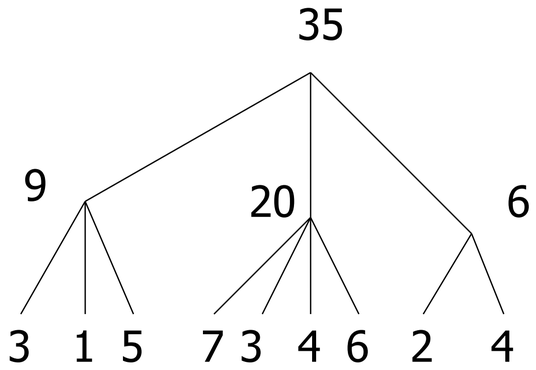
Treemap

- Space-filling representation developed by Shneiderman and Johnson, Vis ‘91
- Children are drawn inside their parent
- Alternate horizontal and vertical slicing at each successive level
- Use area to encode additional data attributes

Example

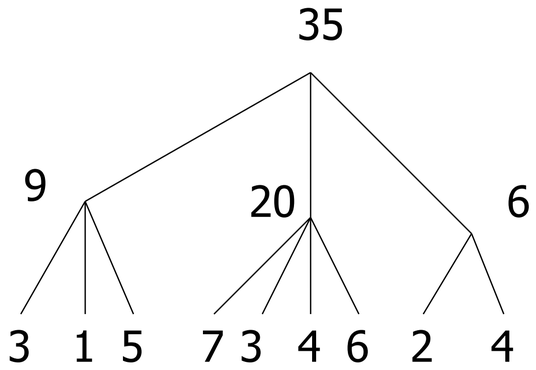


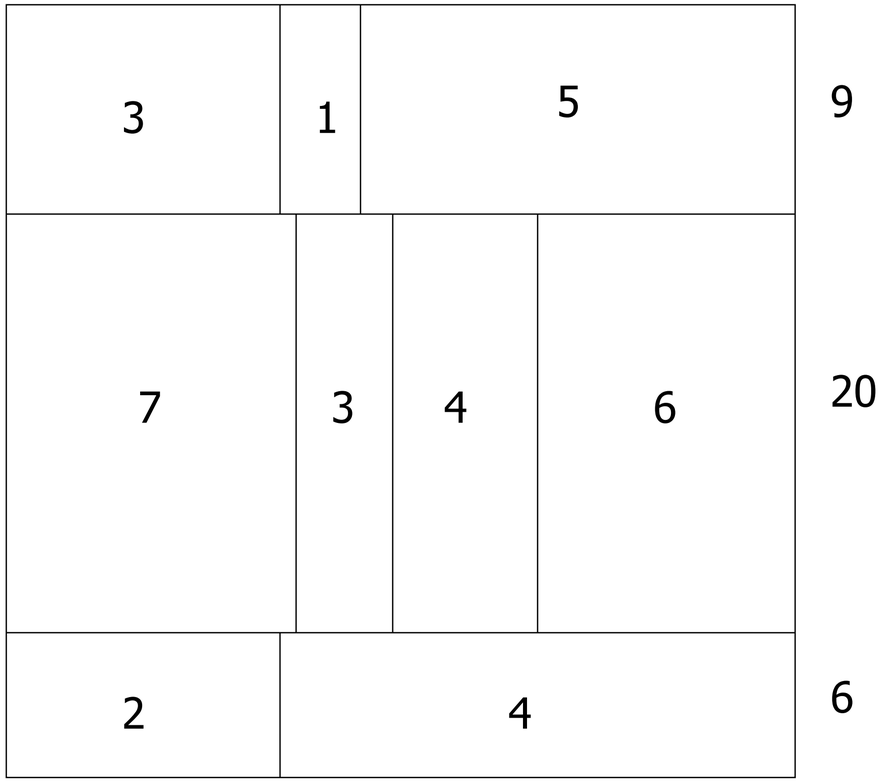
Example

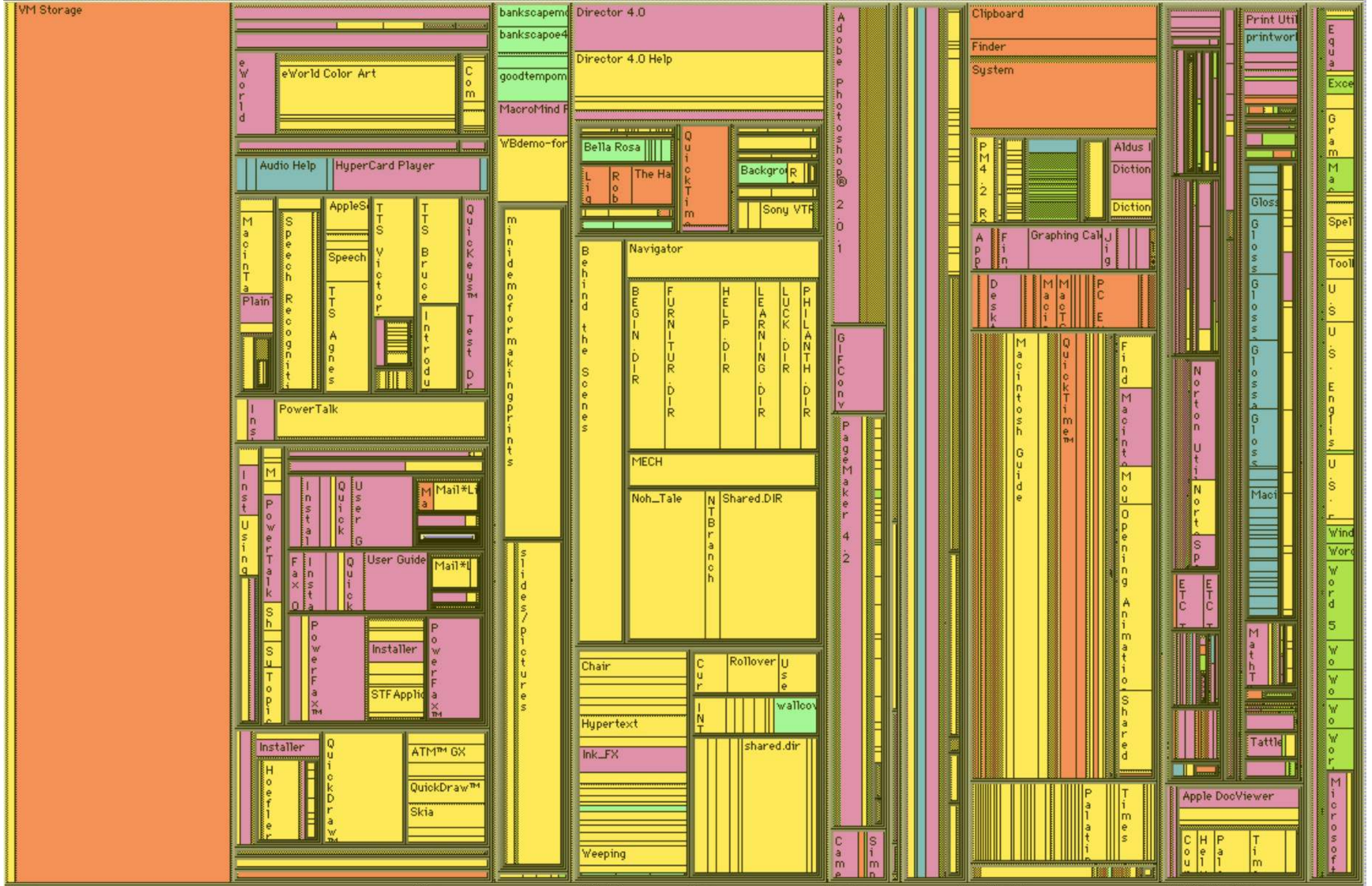


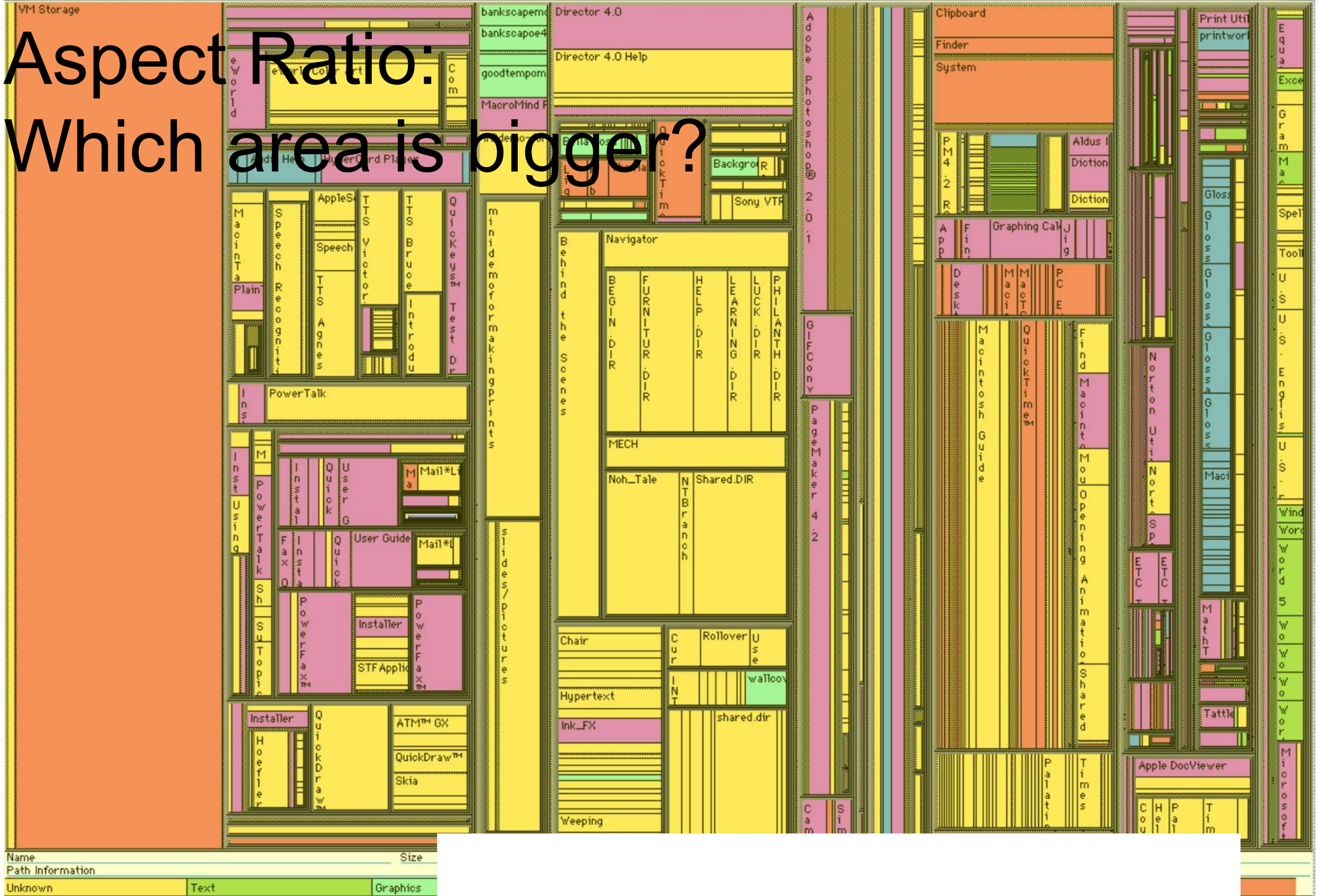


Example



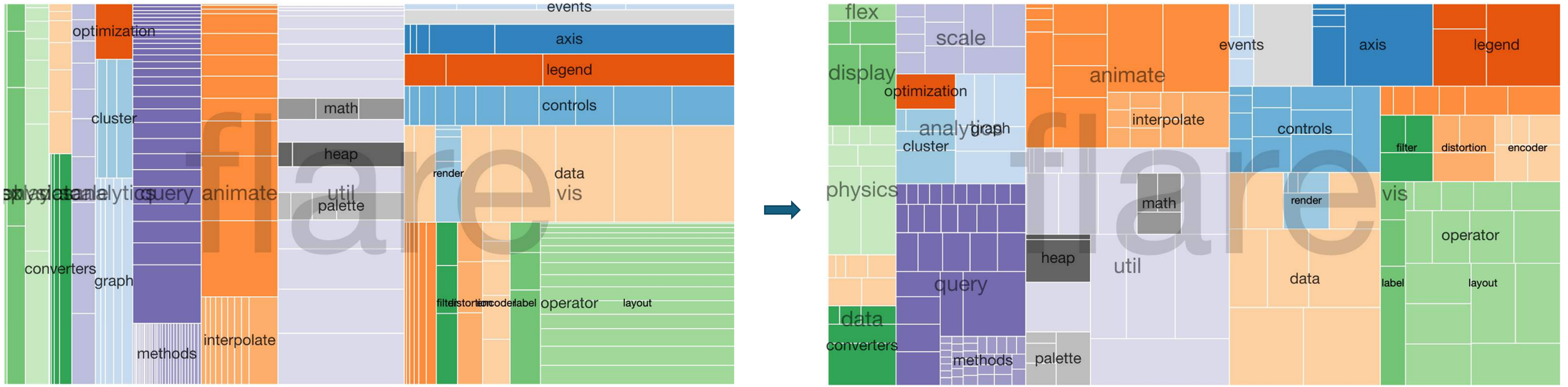



VM Storage
eWorld
eWorld Color Art
Com
Audio Help
HyperCard Player
MacintTa
Plain
Speech Recogniti
AppleSc
Speech
TTS Agnes
TTS Victor
TTS Bruce
Introdu
QuickKeys™ Test Dr
Inst
PowerTalk
Inst
Using
M
PowerTalk
Sh
Su
Topic
Instal
Quick
User G
Ma
Mail*Li
Fax Q
Insta
Quick
User Guide
Mail*L
PowerFax™
Installer
STF Applic
PowerFax™
Installer
Hoefler
QuickDraw™
ATM™ GX
QuickDraw™
Skia
bankscapemc
bankscapoe4
goodtempom
MacroMind P
WBdemo-for
minidemoformakingprints
slides/pictures
Director 4.0
Director 4.0 Help
Bella Rosa
Lig
Rob
The Ha
QuickTim
Backgrou
R
Sony VTR
Behind the Scenes
Navigator
BEGIN.DIR
FURNITUR.DIR
HELP.DIR
LEARNING.DIR
LUCK.DIR
PHILANTH.DIR
MECH
Noh_Tale
NTBranch
Shared.DIR
Chair
Hypertext
Ink_FX
Weeping
Cur
Rollover
Use
INT
wallcov
shared.dir
Adobe Photoshop® 2.0.1
GIFConv
PageMaker 4.2
Came
Simp
Clipboard
Finder
System
PM4.2R
Aldus I
Diction
Diction
App
Fin
Graphing Calc
Jig
Desk
Maci
MacTC
PCE
Macintosh Guide
QuickTime™
Find
Macinto
Mou
Opening Animatio
Shared
Palatin
Times
Norton Uti
Norto
SP
ETC
ETC
Print Uti
printwor
Gloss
Gloss
Gloss
Gloss
Gloss
Gloss
Maci
MathT
Tattle
Apple DocViewer
Cou
Hel
Pal
Tim
Equa
Exce
Gram
Mac
Spel
Tool
U.S.
U.S. Englis
U.S.
Wind
Word
Word 5
Wo
Wo
Wo
Wo
Wor
Microsoft
Name
Size
% Total
Type
Creator
Creation Date
Modification Date
Path Information
Unknown
Text
Graphics
Archives/Stacks
Programming
Applications
System


Aspect Ratio:
Which area is bigger?
VM Storage
Director 4.0
Director 4.0 Help
Clipboard
Finder
System
Navigator
BEGIN.DIR
FURNITUR.DIR
HELP.DIR
LEARNING.DIR
LUCK.DIR
PHILANTH.DIR
MECH
Noh_Tale
NTBranch
Shared.DIR
Behind the Scenes
Chair
Hypertext
Ink_FX
Weeping
shared.dir
Rollover
wallcov
Macintosh Guide
QuickTime™
Opening Animation
Shared
Find
Graphing Calc
Aldus
Diction
Adobe Photoshop® 2.0.1
PageMaker 4.2
Sony VTR
PowerTalk
Speech
TTS Victor
TTS Agnes
TTS Bruce
QuickKeys™ Test Dr
Installer
STF Applic
PowerFax™
User Guide
Mail*L
ATM™ GX
QuickDraw™
Skia
QuickDraw
Hoefler
Apple DocViewer
Norton Uti
Gloss
MathT
Tattle
Microsoft
Name
Size
Path Information
Unknown
Text
Graphics

Squarified Treemaps [Bruls et al. 2000]

- Greedy optimization: try to make the rectangles as square as possible
- Slice/dice within siblings; alternate whenever ratio worsens.

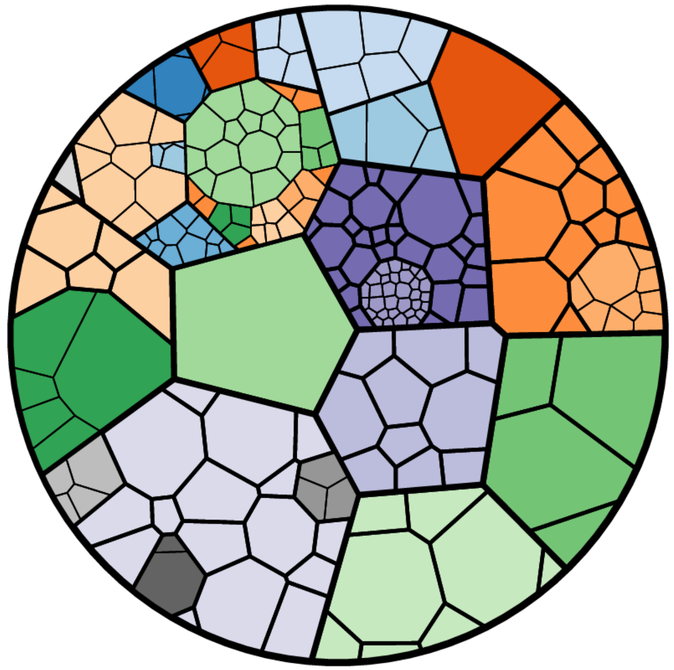


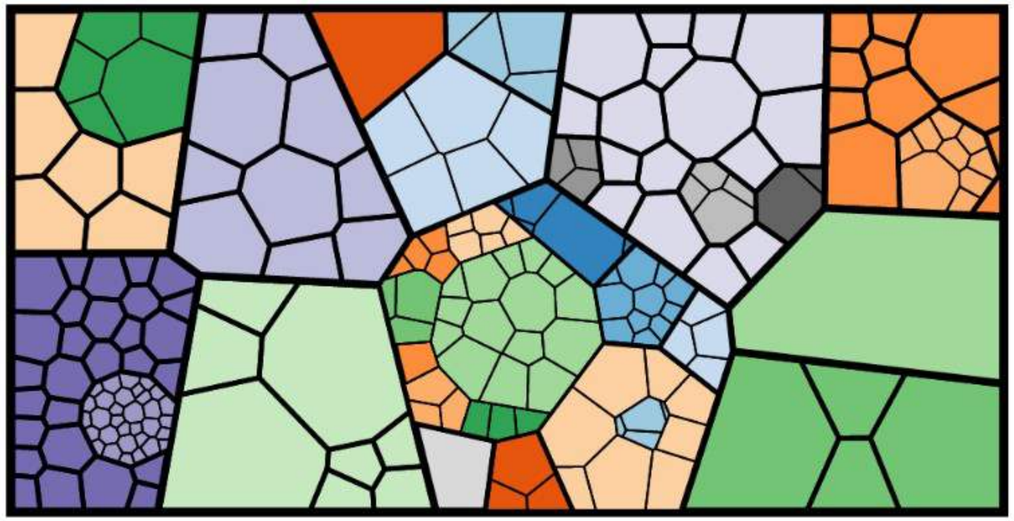
- a small change in data values makes all the rectangles drastically change position (unstable layout)

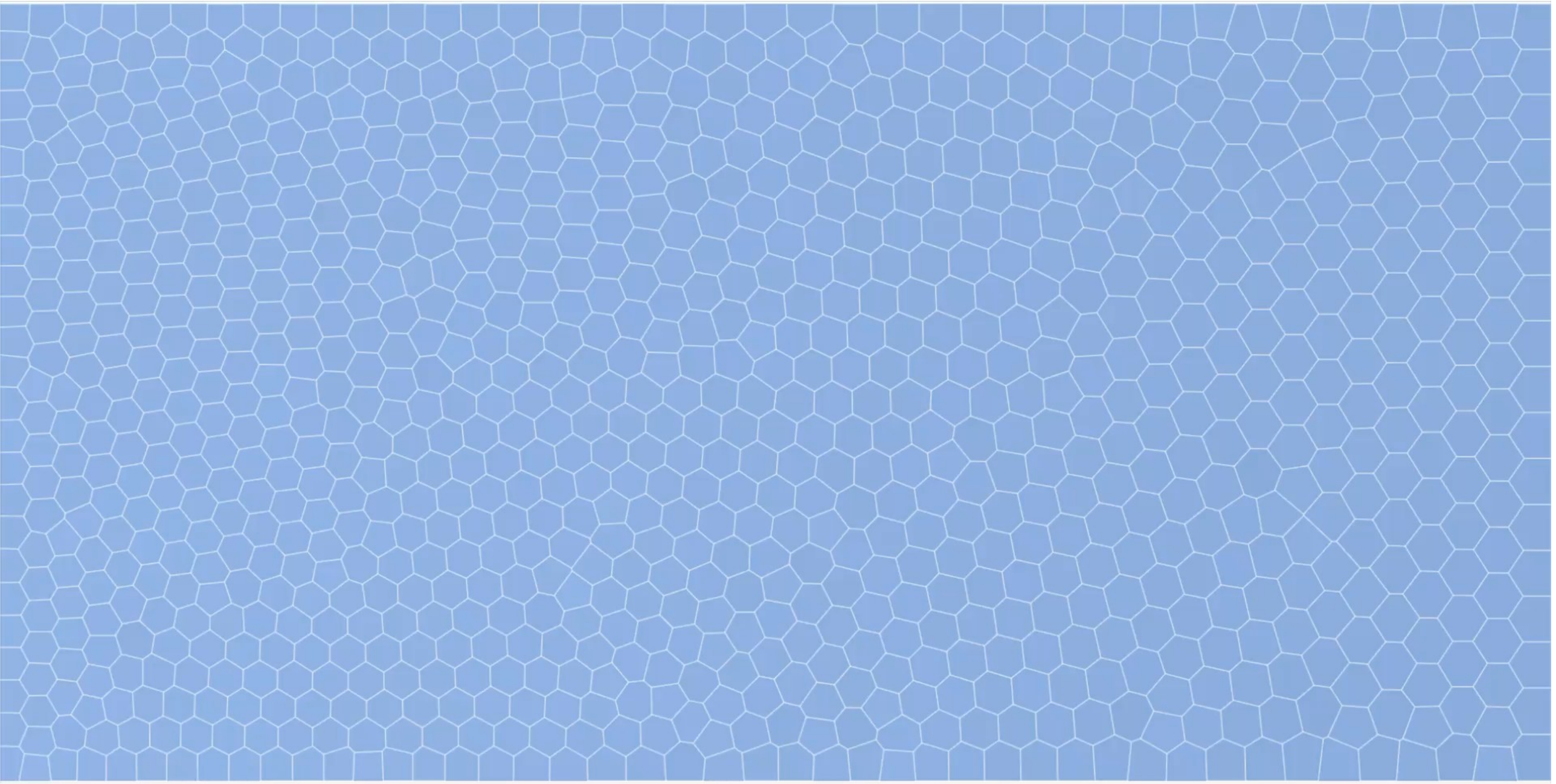
https://www.win.tue.nl/~vanwijk/stm.pdf

Voronoi Treemaps [Balzer et al. ’05]

- Instead of rectangles, divide area into arbitrary polygonal shapes
- Use iterative, weighted Voronoi tessellations to achieve cells with value-proportional areas.







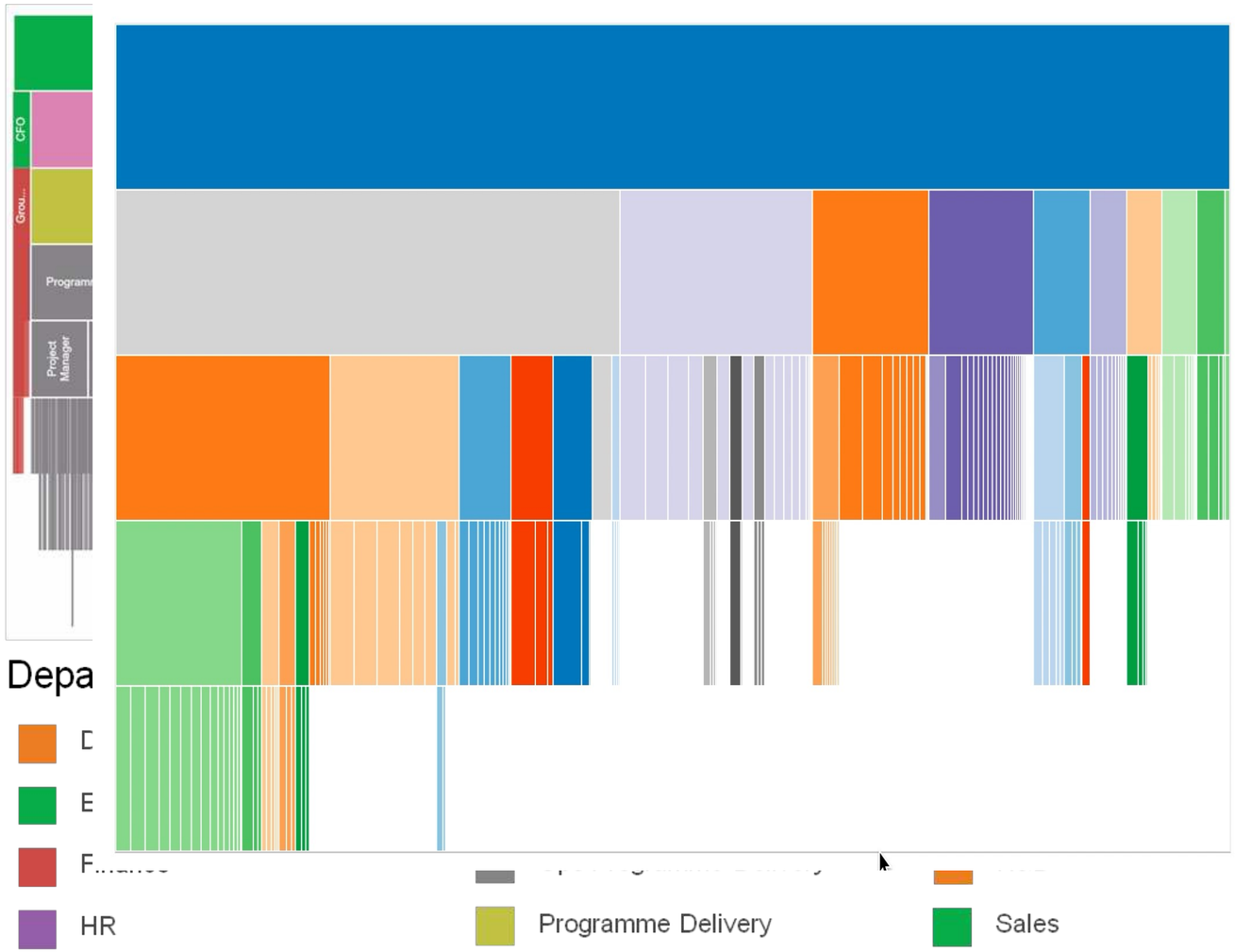
- https://www.jasondavies.com/lloyd/
- https://www.jasondavies.com/voronoi-treemap/

Layering Approaches

- Show tree structure using
  - Layering (parent-child relationships)
  - Adjacency (siblings & grouping)
  - Alignment (depth)

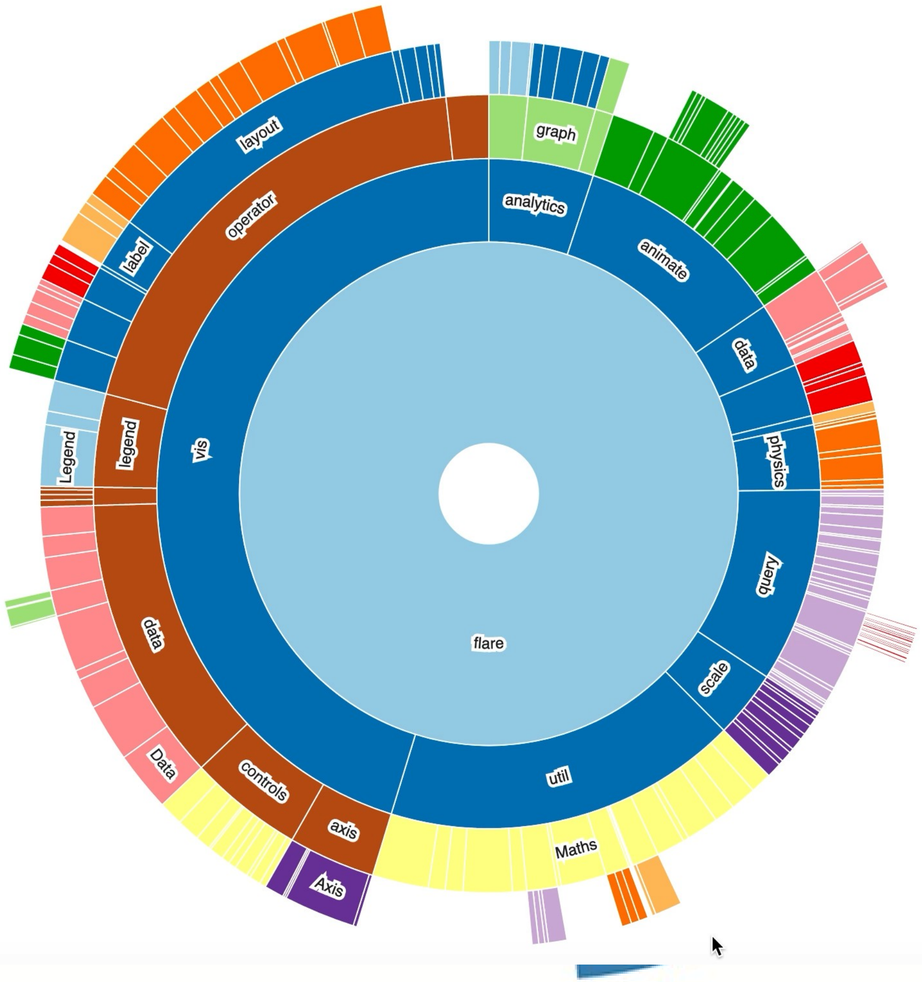
- Leaf nodes may be sized by value, parent size represents sum of descendant leaf values.

Icicle Plot



- Layering in Cartesian space

Sunburst



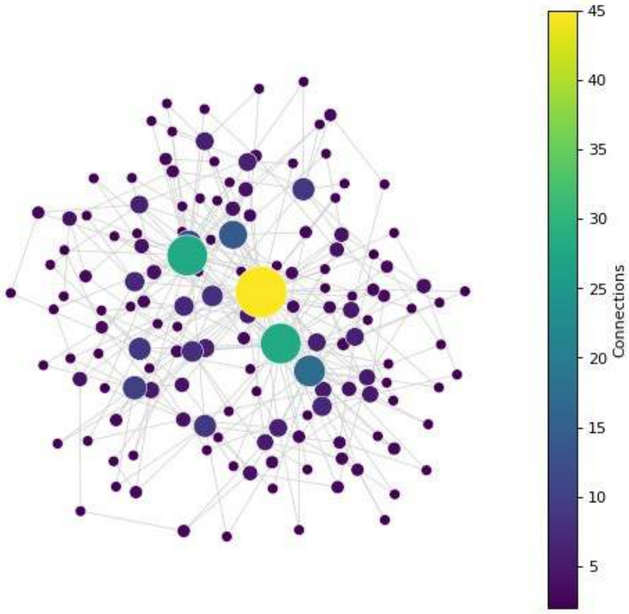
- Layering in Polar space
- Root at center, each successive level drawn farther out from center
- Sweep angle of item corresponds to size
- Color maps to categorical attribute

What if we have network connections that are not hierarchical?

- In some networks, each node has more than one “parent” so we can’t make a hierarchical tree
- You are each a “parent” node
- You have been given two connections to make with other students (“child nodes”)
- Can you arrange yourselves so that you are connected to your child nodes? (i.e. put a hand on their shoulder to “connect”)

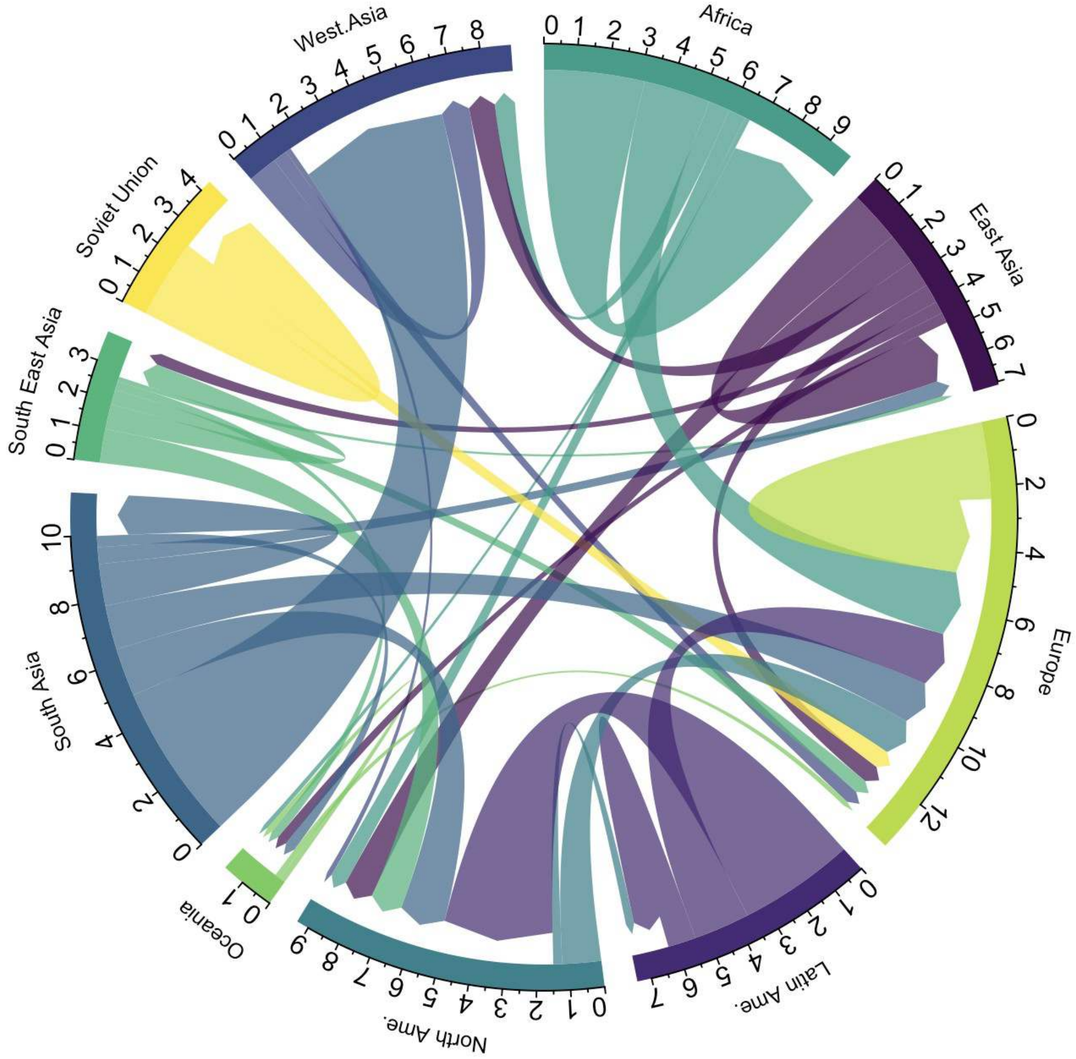
Takeaway: it is not always easy to figure out how to layout and visualize a network!

- We will explore different visualizations of networks on Friday
- Some good rules of thumb:
  - Using **size of node** to represent number of children means the most important (most connected) nodes will be most noticeable
  - Some layouts put these toward the center
- If there is a natural order to the nodes, a **circular layout** works well even though the links will overlap (like chord diagram)



Chord Diagram

- When the weight on the edge is meaningful (how “connected” are two nodes?)
- Vertices drawn as arcs
  - Arc size can encode attribute
- Edges drawn as chords
  - Good for weighted edges



https://www.data-to-viz.com/graph/chord.html

Network layouts

- Force directed/spring layout: more connected nodes get pulled to the center (think of edges as springs)
- Circular; hard to see edges, good for ordered nodes
- Random: almost never good!

