

Visualización de información y analítica visual

Clase 6: Datos espaciales, redes y árboles



**Hoy veremos cómo ubicar datos
espaciales y de redes/árboles en las
visualizaciones**



Datos espaciales

Lo más natural para acomodar datos con semántica espacial es usar dicha semántica para **guiar el layout**.

- Al hacer esto, el canal de posición ya no estará disponible para codificar directamente atributos

Datos espaciales

Lo más natural para acomodar datos con semántica espacial es usar dicha semántica para **guiar el layout**.

- Al hacer esto, el canal de posición ya no estará disponible para codificar directamente atributos



¿Por qué lo obvio tiene sentido?

Normalmente para los datos espaciales ocurre que la **información espacial** es el **atributo más relevante** del dataset.

Principio de efectividad

*La **importancia** del atributo codificado debe coincidir con la **prominencia** del canal escogido para codificarlo.*

¿Por qué lo obvio tiene sentido?

*La **importancia** del atributo codificado debe coincidir con la **prominencia** del canal escogido para codificarlo.*

Y cuál es el canal más prominente de las visualizaciones?




EL CANAL ESPACIAL



¿Por qué lo obvio tiene sentido?

Obviamente esto tiene que ver también con las tareas de los usuarios....

 Si la tarea principal no tiene que ver con el entendimiento de relaciones espaciales, entonces se puede considerar usar el canal espacial para otros atributos

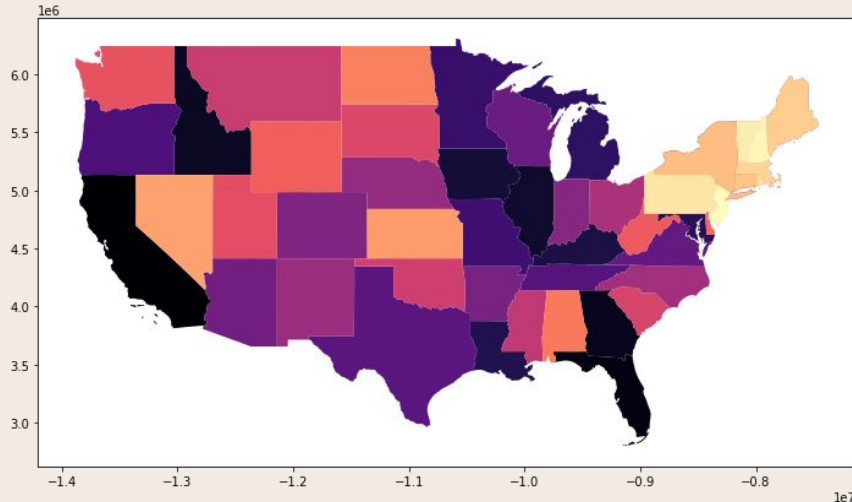
Tipos de datos espaciales

- **Datos Geométricos**
- **Campos espaciales**
 - Escalares
 - Vectoriales
 - Tensoriales

Geometría

Los datos geométricos se caracterizan por contener la información de **forma** de manera **directa** a través de la posición de sus elementos

Un ejemplo: datos geográficos de la tierra



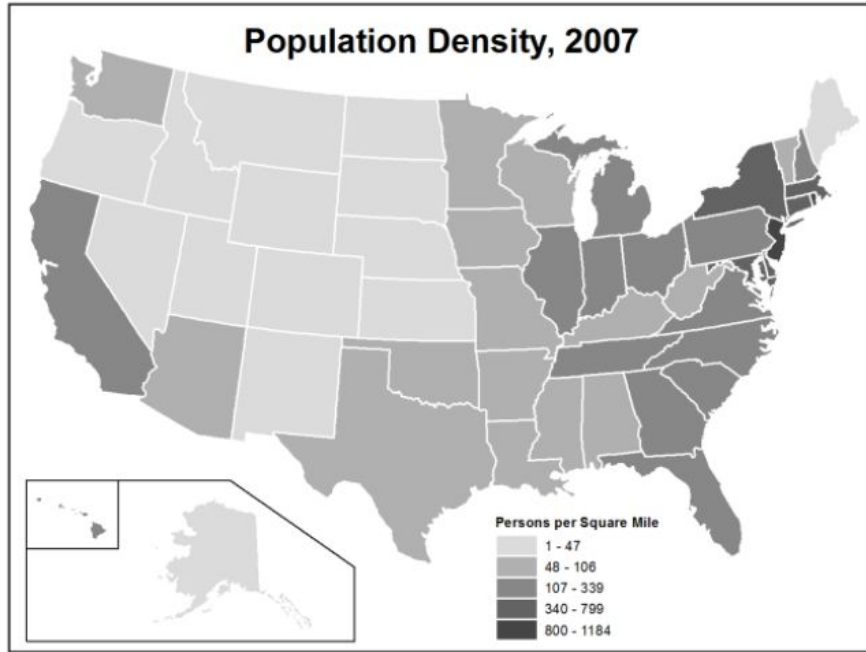
Para la visualización simplemente dibujamos las formas que aparecen en los datos

Geometría: datos geográficos

El campo de la cartografía ha creado *idioms* para representar datos geográficos. Se habla de generalización cartográfica en el mismo sentido que se habla de abstracción de datos/tareas.

Tiene que ver con el **CÓMO** derivar/codificar la geometría del dataset desde los datos en raw.

Choropleth Maps



Un **choropleth map** muestra un atributo cuantitativo con el **color** y los datos espaciales codificados en marcas de **área** cuya **forma** corresponde a la **información geométrica**.


Las decisiones más importantes del choropleth son las del **colormap usados**. Por lo general las formas de las áreas vienen dadas (o derivadas) del dataset

Choropleth Maps

Idiom	Choropleth Maps
Qué: datos	Data geométrica geográfica. Tabla con un dato cuantitativo por región.
Cómo: codificación	Espacio: uso de las formas geométricas para las marcas; Color: colormap secuencial
Por qué: tarea	Apreciar un valor en distintas zonas geográficas

Datos de campo: uno y múltiples valores

Otra manera en la que se presentan datos espaciales son a través de campos uni o multidimensionales (escalares, vectoriales o tensoriales)

-  Los campos son datos que se generan a partir de mediciones continuas de algo

Ejemplo: Isocontornos

Los isocontornos son *idioms* formados por marcas de tipo **líneas** que representan los contornos de un valor determinado.

Estas líneas se denominan **isolíneas** y representan los contornos de un valor escalar particular.



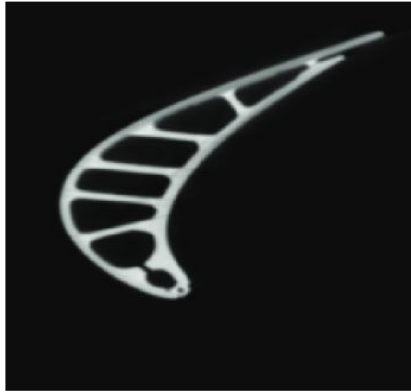
Topographic Terrain Maps



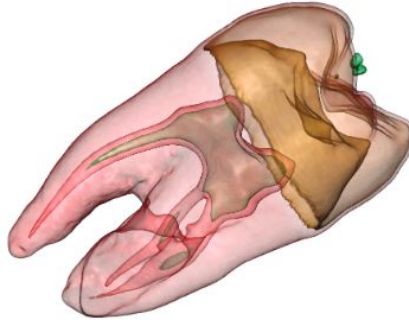
Topographic Terrain Maps

Idiom	Topographic Terrain Maps
Qué: datos	Campo espacial 2D, datos geográficos.
Qué: derivación de datos	Líneas que representan valor del campo
Cómo: codificación	Geometría y marcas tipo región. Líneas representando los campos de valor
Por qué: tarea	Consultar formas y valores
Escala	Docenas de niveles de contorno

Los datos de campo se dan en dominios muy específicos...



(a)



(b)



(c)

Figure 8.3. Spatial scalar fields shown with three different idioms. (a) A single 2D slice of a turbine blade dataset. (b) Multiple semitransparent isosurfaces of a 3D tooth dataset. (c) Direct volume rendering of the entire 3D turbine dataset. From [Kniss 02, Figures 1.2 and 2.1b].

Los datos de campo se dan en dominios muy específicos...

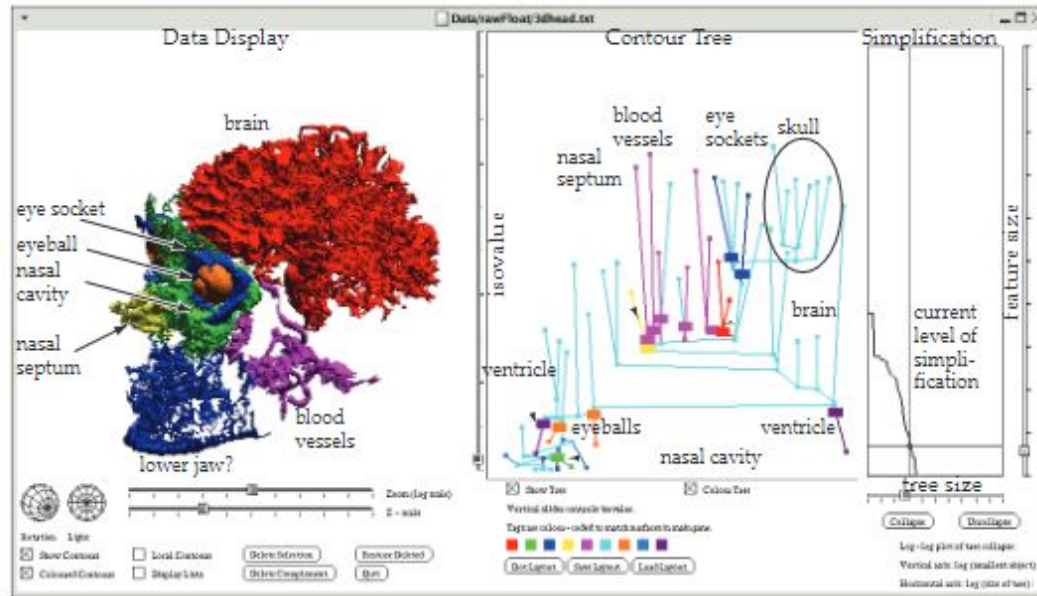


Figure 8.6. The flexible isosurfaces idiom uses the simplified contour tree of under 100 edges to help users identify meaningful structure. From [Carr et al. 04, Figure 1].

Las redes y árboles se originan a partir de datasets donde lo importante son aspectos topológicos y/o de relación/jerárquica en los datos

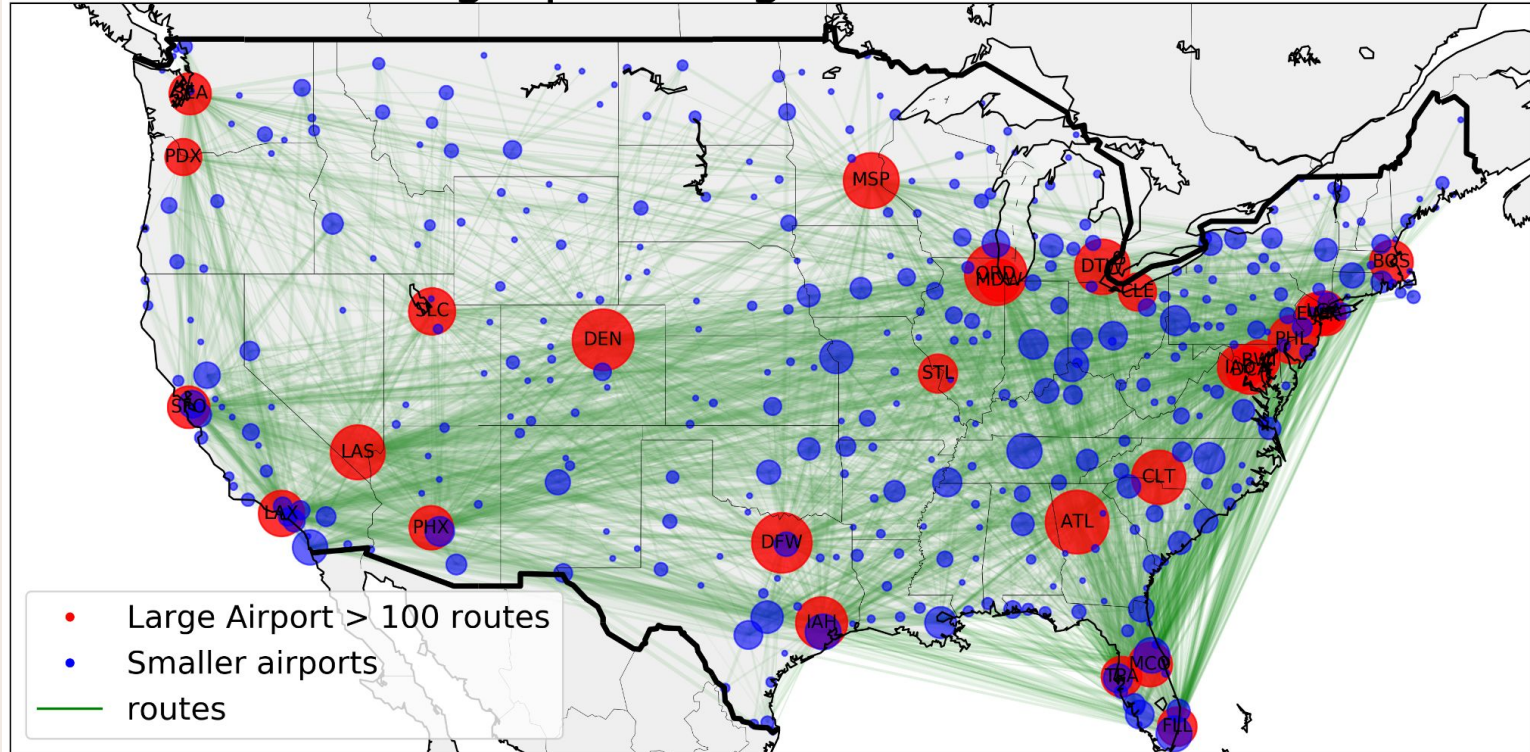
Redes y árboles

Dataset de redes



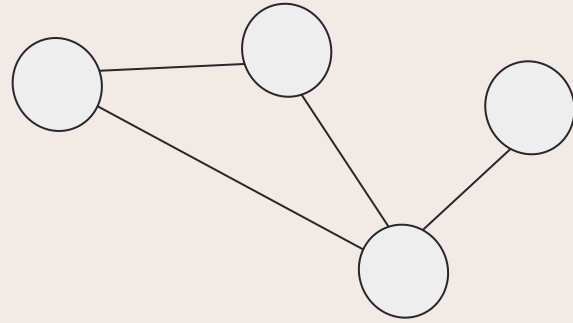
Dataset de redes

Network graph of flight routes in the USA

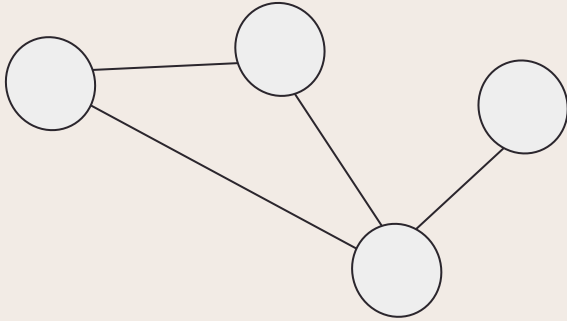


Dataset de redes: nomenclatura

- Items
- Enlaces
 - Enlaces dirigidos
 - Enlaces no-dirigidos
 - Enlaces con pesos
- Caminos



Grafos



Los grafos se definen matemáticamente como una tupla $G=(V,E)$ donde V es el conjunto de nodos (o vértices) y E es el conjunto de aristas.

Medidas en redes (grafos)

Existen medidas que permiten describir y comparar distintas redes

- Tamaño de la red (cantidad de nodos y aristas)
- Diámetro (camino más largo)
- Distancia promedio de nodos
- Grado de un nodo
- Cercanía de un nodo
- y muchas más ...

Objetivos comunes en dataset de redes

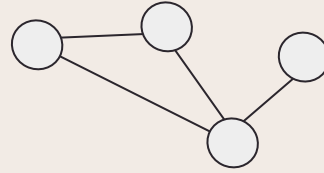
Tarea = acción (verbo) + objetivo (sustantivo) + ...

Por lo general, el objetivo en dataset de redes tendrá que ver con:

- Caminos
- Topología

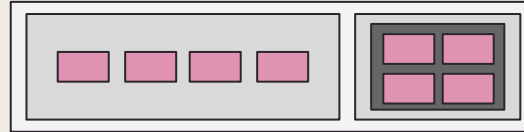
En esta clase exploraremos tres *idioms* para redes

1. Diagramas nodo-aristas



2. Matriz de adyacencia

3. Contención



Diagramas nodo-arista

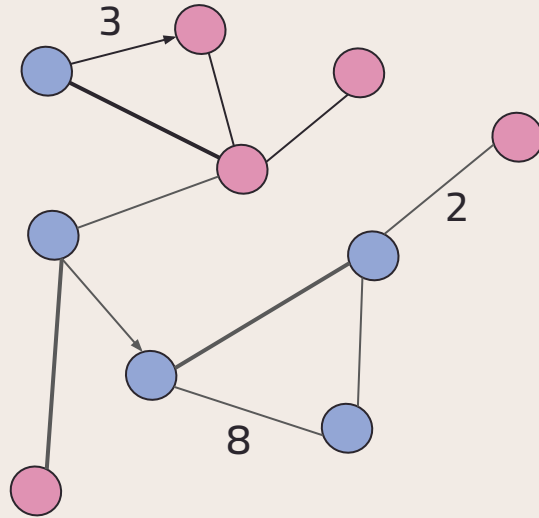
La manera tradicional de codificar árboles y redes son los grafos: diagramas con nodos y aristas.

- Los nodos se codifican en marcas tipo punto
- Las aristas se codifican en marcas tipo conexión (marca tipo línea).

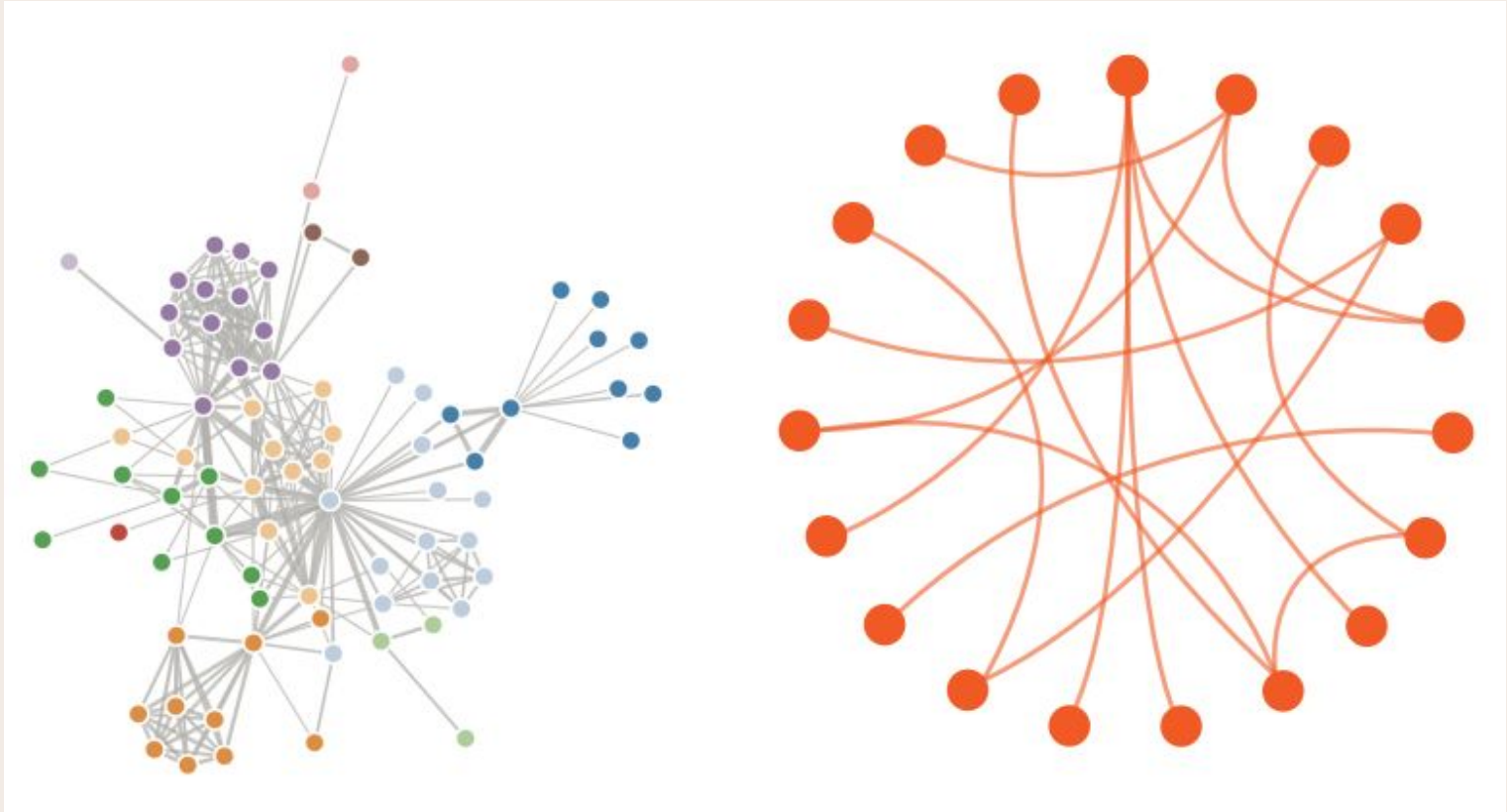
Con este idiom codificamos directamente las relaciones entre ítems



Diagramas nodo-arista



Conexiones rectas y curvas



Tareas en diagramas nodo-arista

- Entender topologías
- Encontrar caminos
- Diámetro
- Camino más corto
- Vecinos de un nodo
- Nodos con alto degree



Posicionamiento dirigido por fuerzas

Un aspecto relevante a considerar es donde ubicar los items para codificar el diagrama nodo-arista

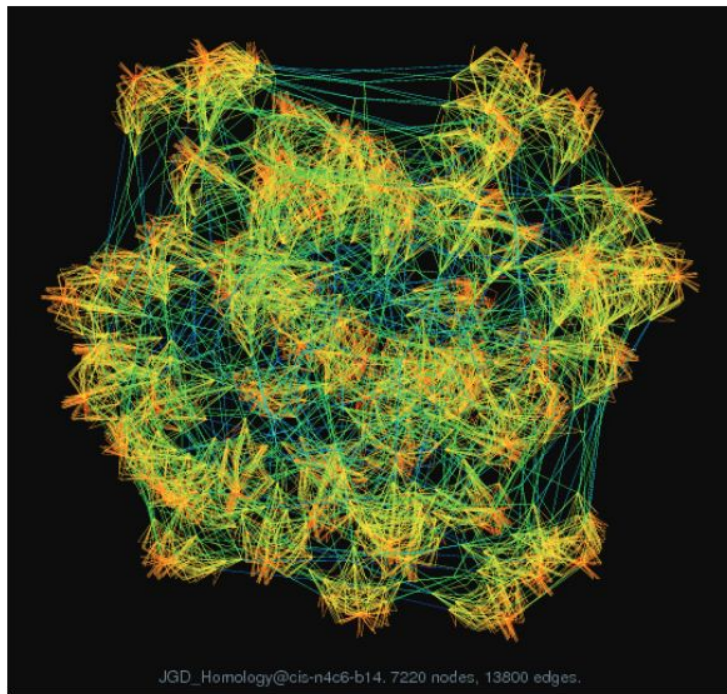
- Algunas veces hay algún aspecto espacial que nos indica dónde deben ir los ítems
- Pero otras veces no...

Hay algoritmos que se encargan de ubicar los nodos de manera que sea más fácil llevar a cabo algunas tareas: un algoritmo es el de posicionamiento dirigido por fuerzas

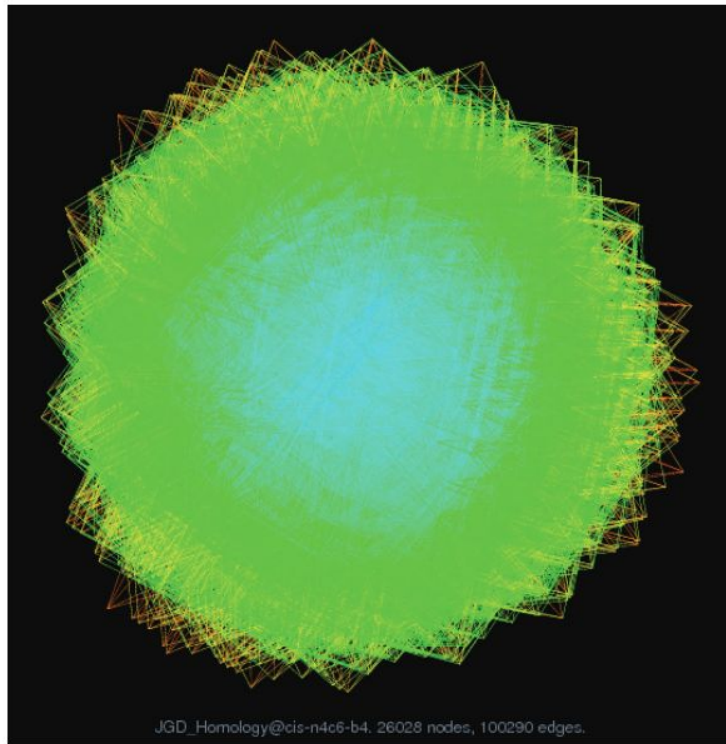
Posicionamiento dirigido por fuerzas



Escalabilidad: bolas de pelo



(a)















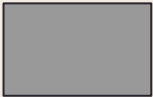











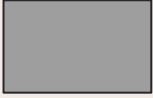

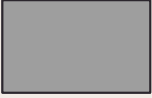



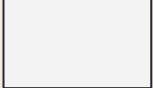
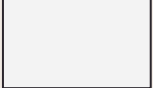
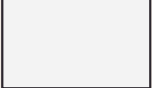

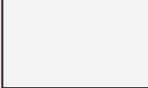
(b)

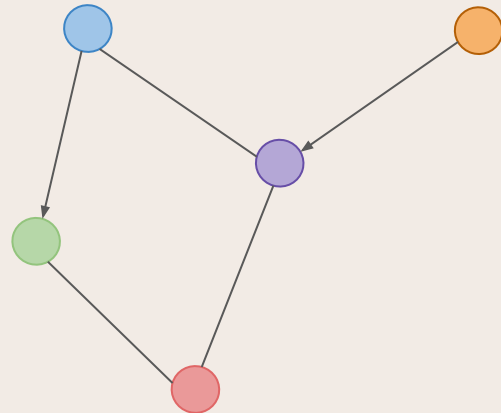
Matriz de adyacencia

Los datasets que codifican relaciones entre ítems (redes) pueden ser derivados a matrices de adyacencia:







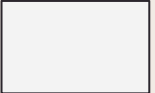

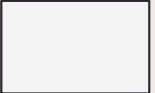

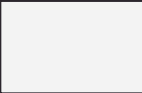














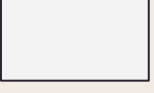

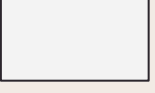
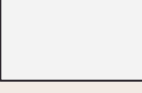

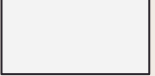
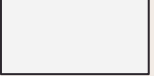
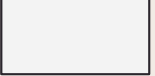

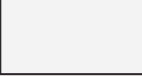
- Las filas y columnas representan ítems
- Los valores de las celdas representan relaciones entre ítems

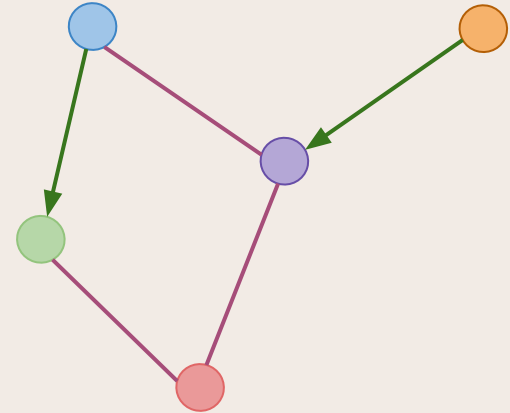
Matriz de adyacencia



Matriz de adyacencia: uso de colormaps



Matriz de adyacencia: ejemplos

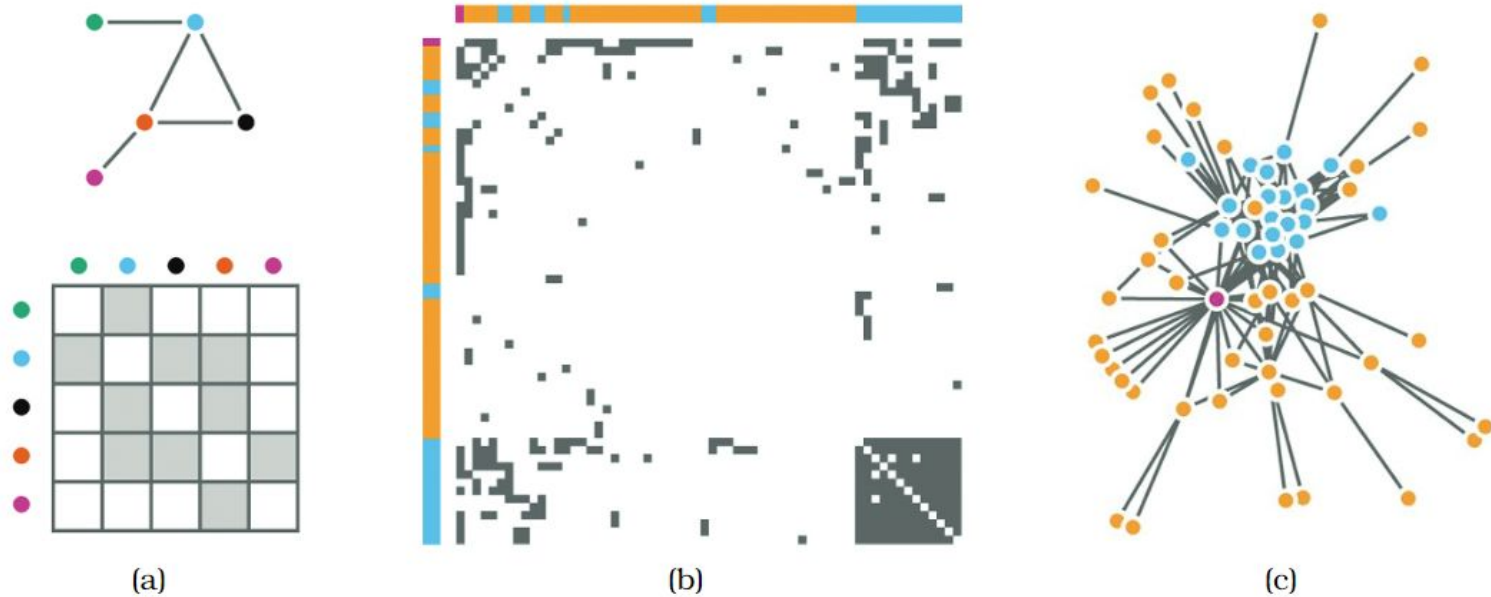
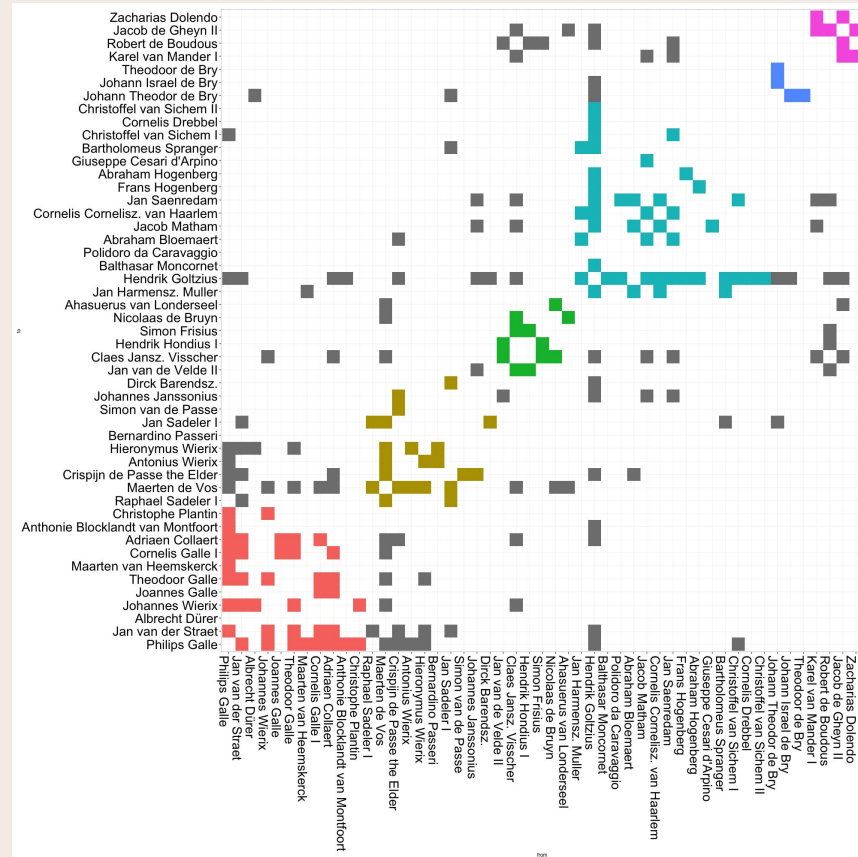
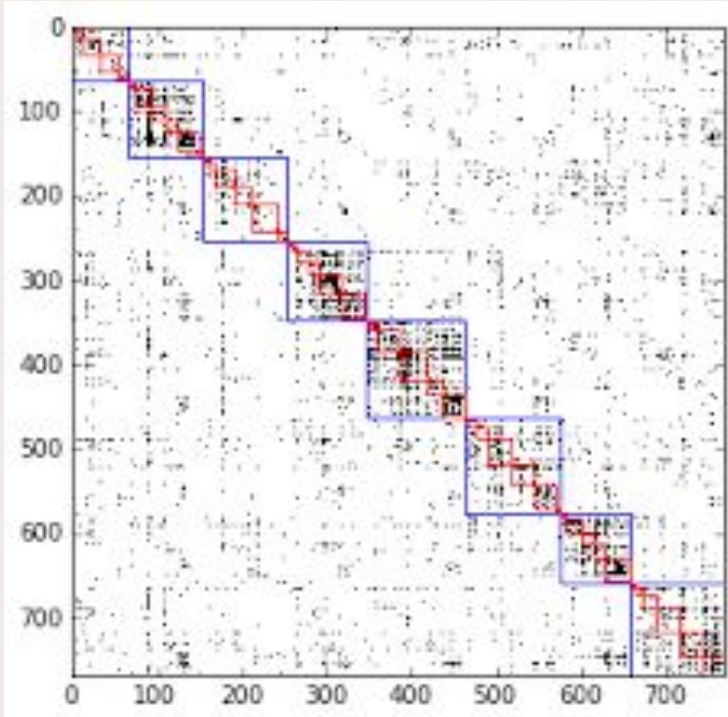


Figure 9.6. Comparing node-link matrix and matrix views of a network. (a) Node-link and matrix views of small network. (b) Matrix view of larger network. (c) Node-link view of larger network. From [Gehlenborg and Wong 12, Figures 1 and 2].

Matriz de adyacencia: ejemplos



Escalabilidad de matriz de adyacencia



A diferencia de los diagramas nodo-arista, la matriz de adyacencia es muy útil para redes particularmente grandes

Matriz de adyacencia

Idiom	Matriz de adyacencia
Qué: datos	Redes
Qué: datos derivados	Tabla: nodos como llaves, conexiones entre dos nodos como valores
Cómo: codificación	Marcas de area en alineación 2D tipo matriz
Por qué: tarea	Analizar saturación de la red, buscar clusters
Escala	Miles de nodos, millones de enlaces

Matrices versus diagrama nodo-enlace

Nodo-enlace

- Hasta cientos de items y 4x enlaces
- Se aprecian caminos
- Se aprecia la topología
- En general son buenos para todas las tareas, cuando el dataset es pequeño

Matriz de adyacencia

- Miles de items, millones de enlaces
- No se aprecian caminos
- Si se aprecian clusters
- Se aprecia directamente la densidad
- Muy buenas para datasets grandes y densos

Matrices versus diagrama nodo-enlace

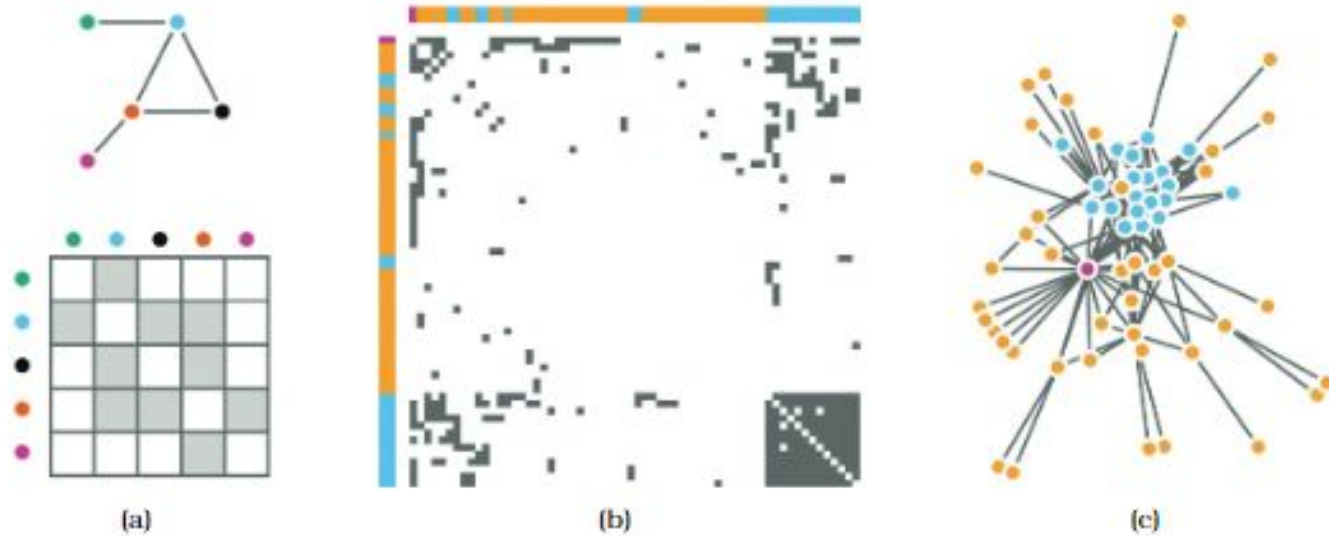
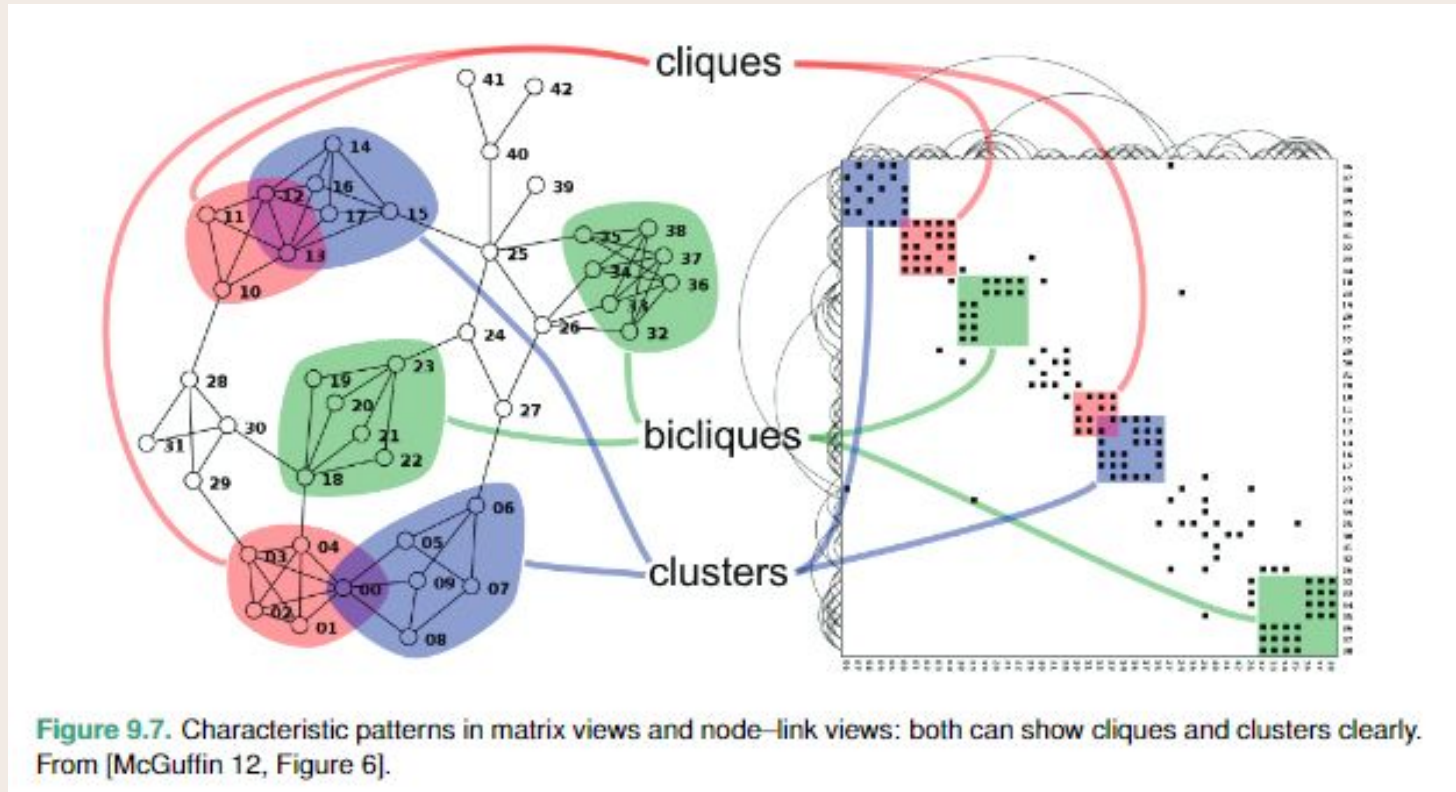
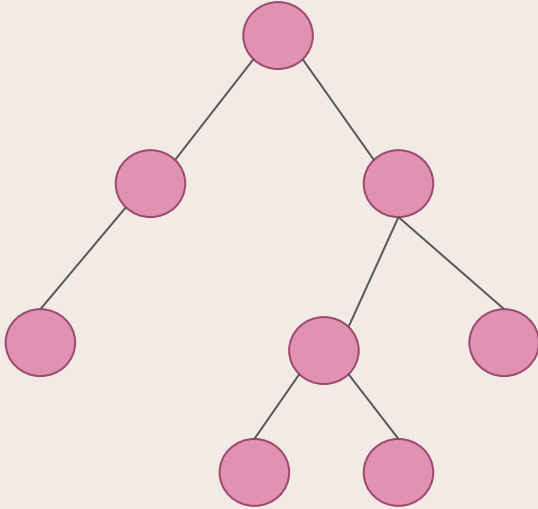


Figure 9.6. Comparing node-link matrix and matrix views of a network. (a) Node-link and matrix views of small network. (b) Matrix view of larger network. (c) Node-link view of larger network. From [Gehlenborg and Wong 12, Figures 1 and 2].

Matrices versus diagrama nodo-enlace



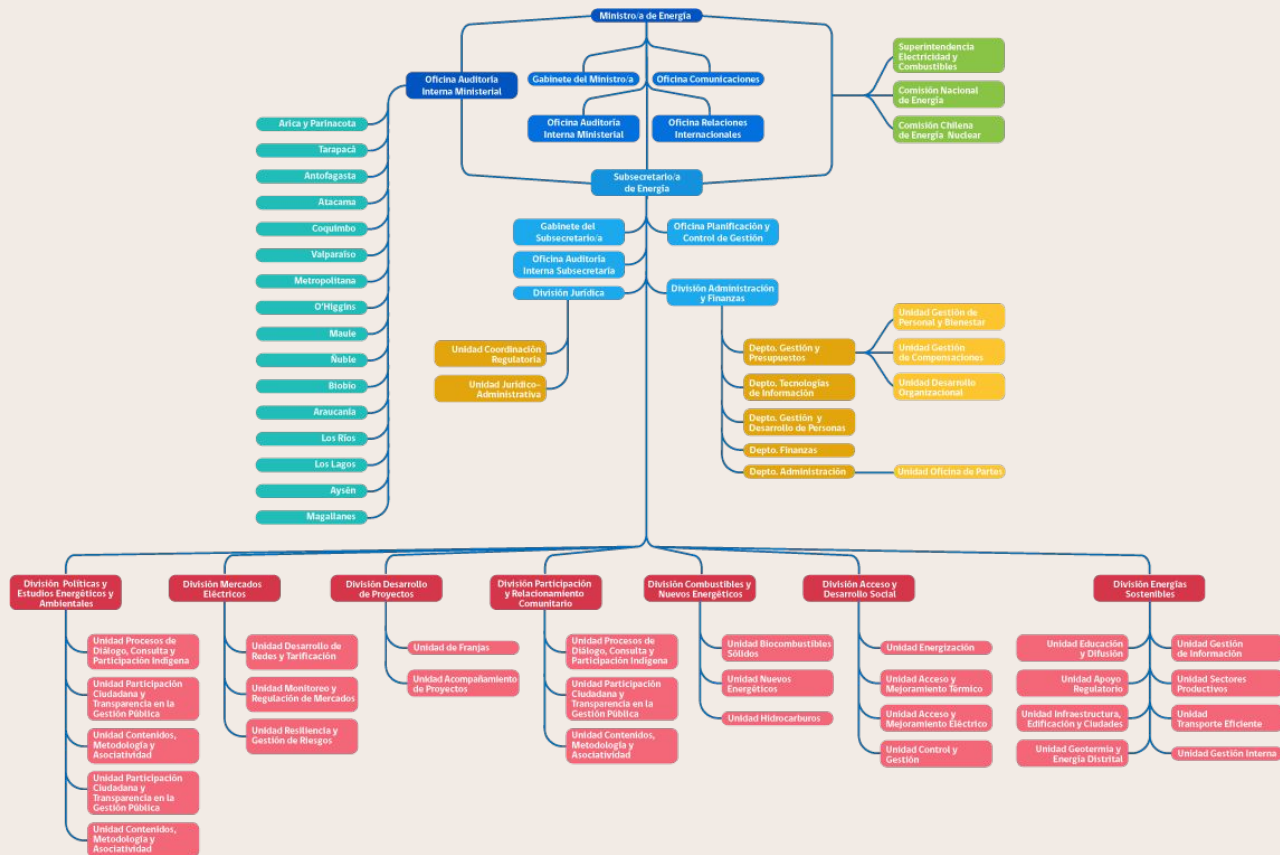
Jerarquías y árboles



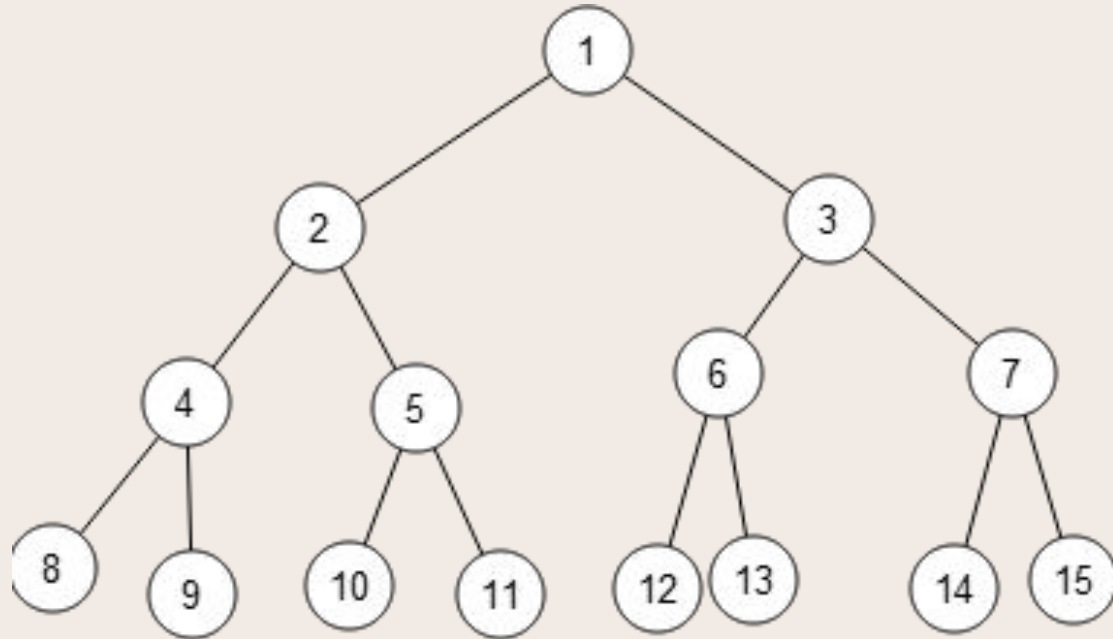
Los árboles son casos especiales de redes/grafos

- Hay un sentido de orden que se genera a partir de relaciones jerárquicas
- Las relaciones son de tipo padre-hijo
- Nodo especial: nodo raíz

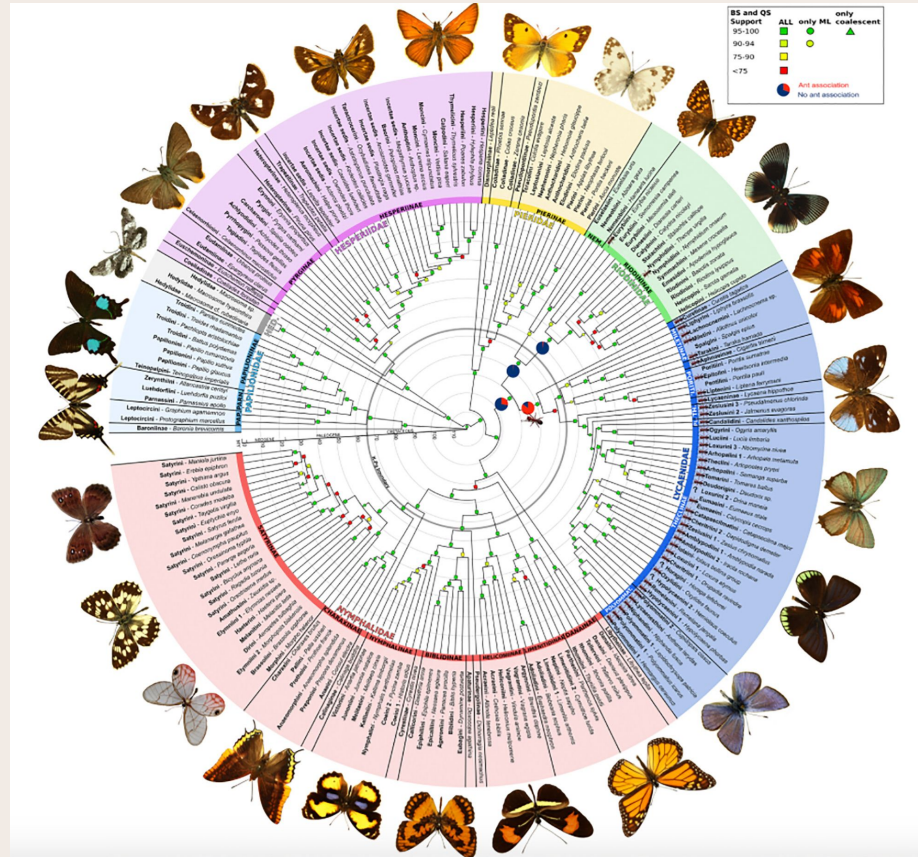
Jerarquías y árboles



Visualización en nodo-arista: árbol



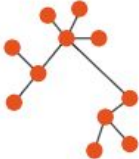
Visualización en nodo-arista: layout radial



Contención


➔ **Node-Link Diagrams**
Connection Marks

✓ NETWORKS ✓ TREES




➔ **Adjacency Matrix**
Derived Table

✓ NETWORKS ✓ TREES



➔ **Enclosure**
Containment Marks

✗ NETWORKS ✓ TREES



Como los árboles/jerarquías son redes, aplican todos los idioms que vimos para redes:

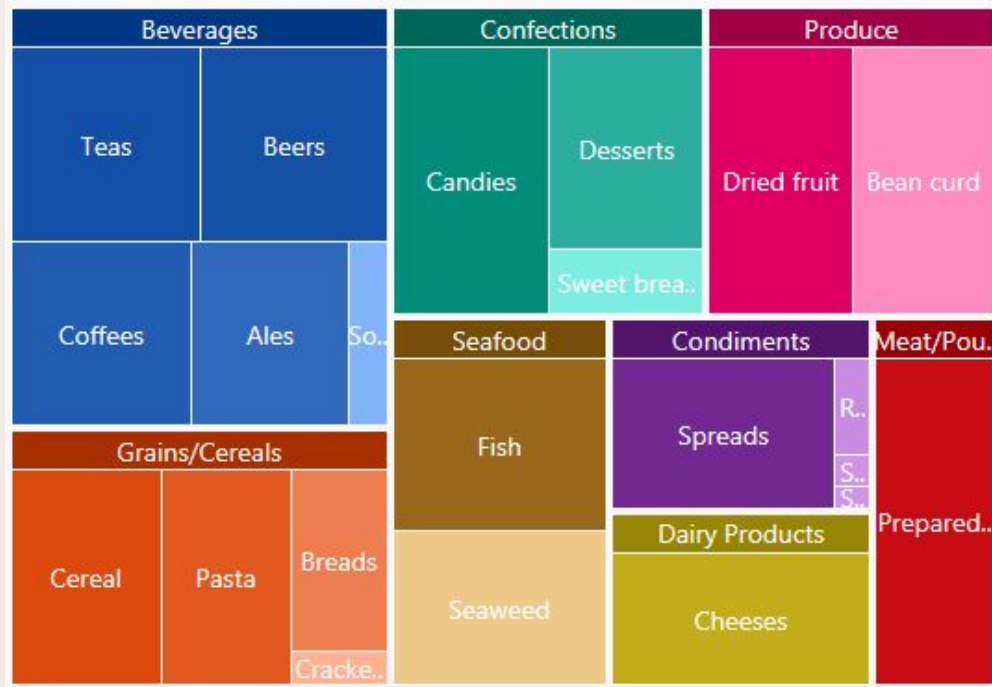
- diagrama nodo-enlace
- matriz de adyacencia

Pero además se suman nuevos idioms: como el de **contención**

Contención

El idiom de contención por lo general usa marcas de área y canales de proximidad, agrupamiento y contención espacial para codificar relaciones de jerarquía

Treemaps



Treemap: se usa la **contención** para codificar jerarquías.

Los tamaños de las áreas pueden mapear algún atributo cuantitativo de cada ítem.



Las marcas de contención no son tan buenas para re-armar caminos en el árbol, pero sí para aprender cosas de los valores de las hojas

Treemap

Idiom	Treemap
Qué: datos	Árbol
Cómo: codificación	Marcas de area y contención, con layout rectilíneo
Por qué: tarea	Analizar aspectos (atributos) de los ítems hoja
Escala	Millones

Hay varias alternativas para codificar jerarquías...

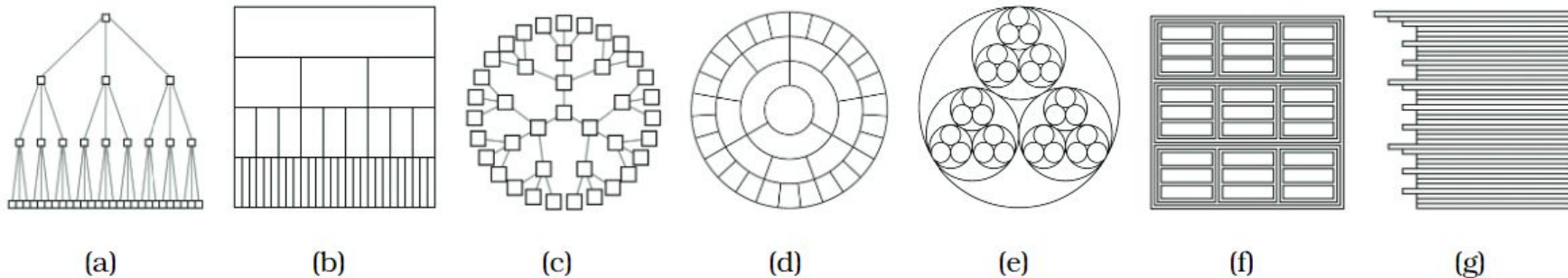


Figure 9.9. Seven visual encoding idioms showing the same tree dataset, using different combinations of visual channels. (a) Rectilinear vertical node-link, using connection to show link relationships, with vertical spatial position showing tree depth and horizontal spatial position showing sibling order. (b) Icicle, with vertical spatial position and size showing tree depth, and horizontal spatial position showing link relationships and sibling order. (c) Radial node-link, using connection to show link relationships, with radial depth spatial position showing tree depth and radial angular position showing sibling order. (d) Concentric circles, with radial depth spatial position and size showing tree depth and radial angular spatial position showing link relationships and sibling order. (e) Nested circles, using radial containment, with nesting level and size showing tree depth. (f) Treemap, using rectilinear containment, with nesting level and size showing tree depth. (g) Indented outline, with horizontal spatial position showing tree depth and link relationships and vertical spatial position showing sibling order. From [McGuffin and Robert 10, Figure 1].

Visualización de información y analítica visual

Clase 6: Datos espaciales, redes y árboles



Tutorial #3: visualizaciones de datos espaciales, redes y árboles



Control #1: Visualización de datos tabulares (continuación)

