
IIC2026

Visualización de Información

— Hernán F. Valdivieso López —
(2023 - 2 / Clase 22)

Temas de la clase - Visualización de Redes

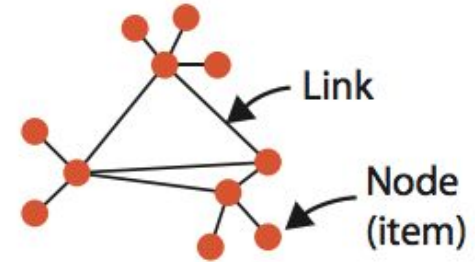
1. *Dataset* de red.
2. Nodo-enlace y Matriz de Adyacencia.
3. Jerarquías y Árboles.

Dataset de red

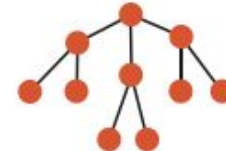
Dataset de red

- Representar situaciones donde existen relaciones entre dos o más ítems.
- Compuestos por datos de tipo ítem (nodos) y de tipo enlace.
- Algunos ejemplos típicos son redes sociales, red de tráfico, árbol jerárquico de una empresa, entre otros.

→ Networks

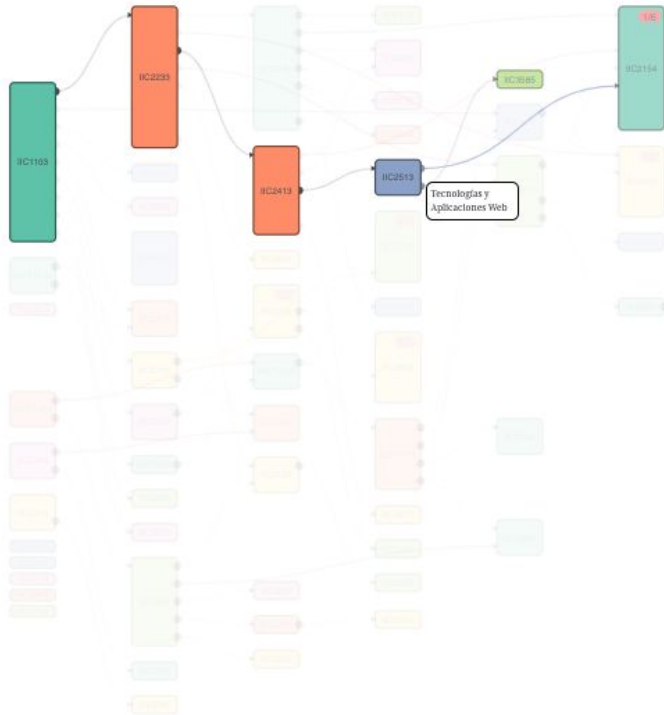


→ Trees



Dataset de red

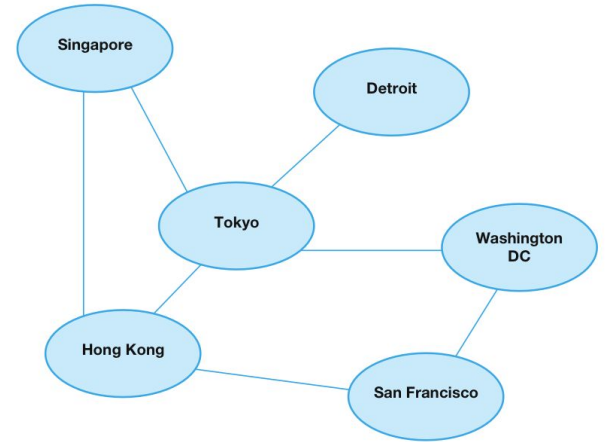
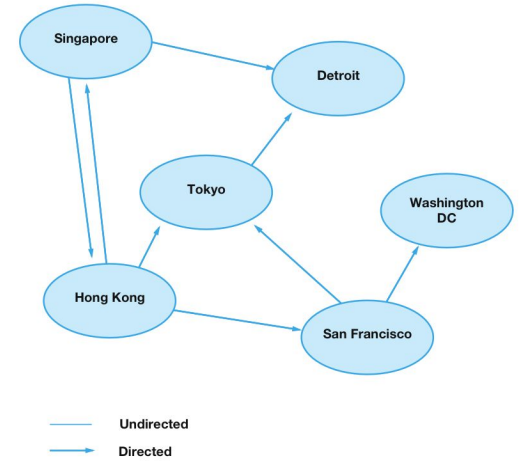
- Se pueden aplicar en diferentes contextos.



Dataset de red

Hay redes dirigidas y no dirigidas.

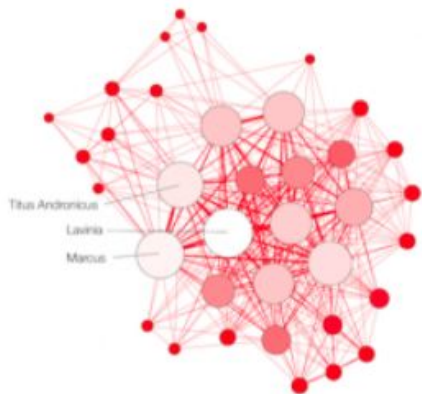
- **Dirigidas:** útil para representar relaciones asimétricas. Por ejemplo:
 - jefaturas: Sofía es la jefa de María en la empresa.
 - "seguir a un usuario" en redes sociales: Hernán siguió a Nicolas en Instagram y Twitter.
- **No dirigidas:** útil para representar relaciones simétricas. Por ejemplo:
 - amistades de facebook: Si Hernán es amigo de Nicolás, Nicolás también será amigo de Hernán.



Dataset de red

Los nodos y/o enlaces pueden contener atributos.

SHAKESPEAREAN TRAGEDY



TITUS ANDRONICUS
Number of characters 36 | 50% Network density



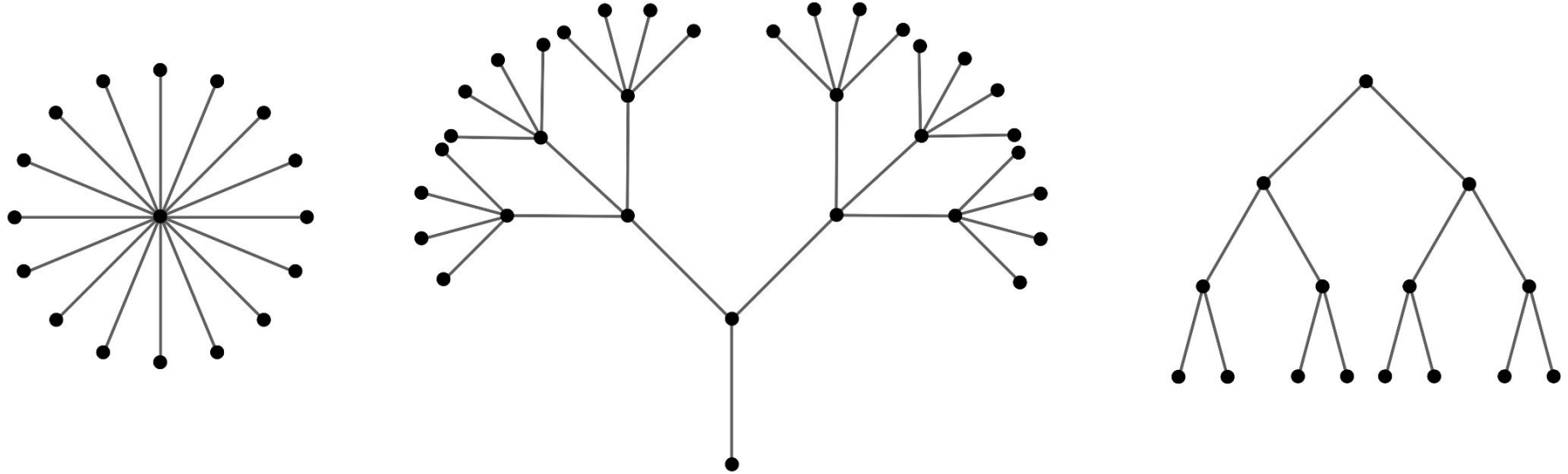
HAMLET
Number of characters 37 | 39% Network density



KING LEAR
Number of characters 33 | 45% Network density

Dataset de red

Pueden ser del tipo jerárquico

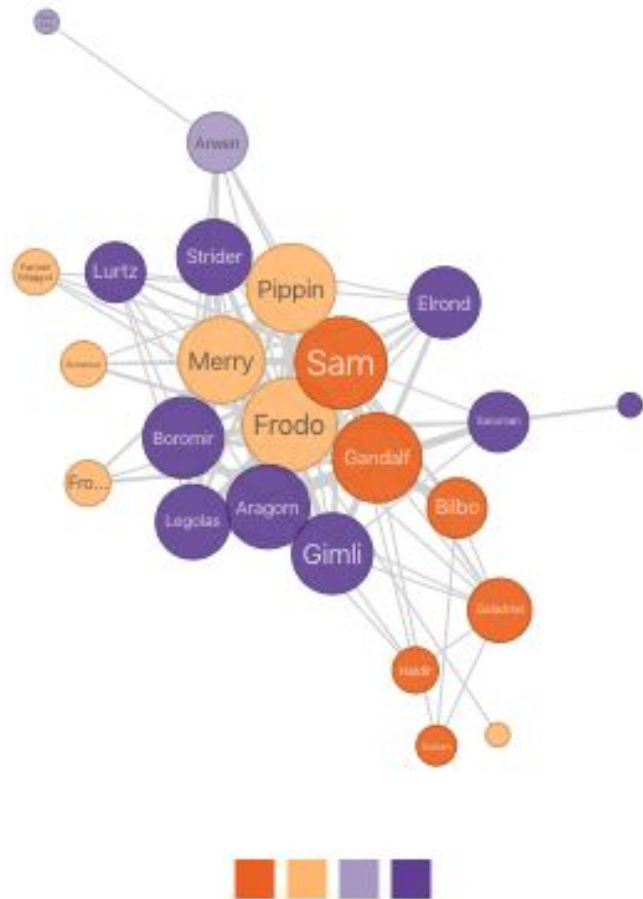


Dataset de red - Métricas

- En este tipo de dataset existen diferentes métricas, generalmente numéricas, que entregan una pequeña descripción de la red o nodos.
 - **De la red:** cantidad de nodos, cantidad de enlaces, diámetro (distancia máxima entre nodos de una red) y distancia promedio entre nodos
 - **De cada nodo:** grado (cantidad de conexiones) y cercanía (distancia promedio con otros nodos)
- Si están interesados en profundizar sobre medidas en redes, pueden revisar el siguiente enlace: [4. Measuring Networks Part 1: Centrality and Global Measures](#)

Dataset de red - Métricas

- Cantidad de nodos: 24
- Cantidad de enlaces: 106
- Diámetro (distancia máxima entre nodos de una red): 4
- Distancia promedio entre nodos: 2.23

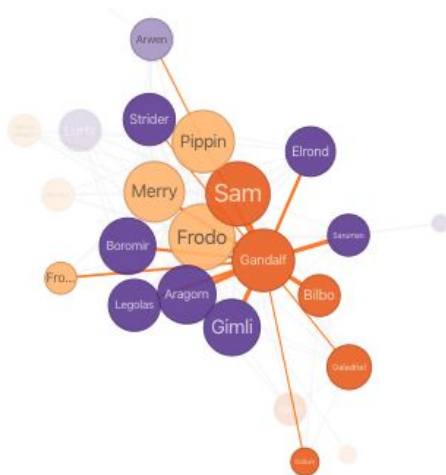


Dataset de red - Métricas

Nodo Gandalf

- Grado: 16
- Cercanía: 1.44

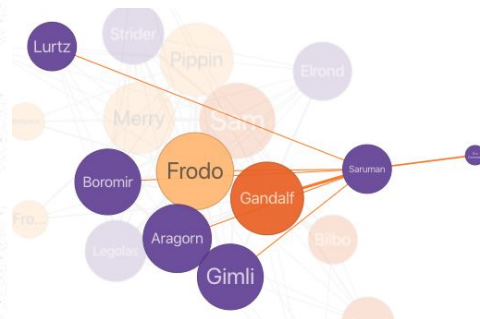
Name	Gandalf
Degree	16
Betweenness Centrality	27.9
Closeness Centrality	1.30
Eigenvector Centrality	0
Triangles	68
Eccentricity	2
Group	1
Pagerank	0.10



Nodo Saruman

- Grado: 7
- Cercanía: 1.74

Name	Saruman
Degree	7
Betweenness Centrality	22.7
Closeness Centrality	1.74
Eigenvector Centrality	0
Triangles	12
Eccentricity	3
Group	2
Pagerank	0.031



Dataset de red - Visualización

Algunos objetivos dentro de estos datos que motivan a realizar visualizaciones son:

- **Los caminos entre nodos.**
 - Conocer la distancia entre nodos específicos.
 - Conocer si están conectados directamente.
 - Conocer qué nodos hay en el camino.
- **La topología**, que es la forma general de la red, cómo se interconectan sus nodos: de forma tanto directa como indirecta.
 - Hay algún nodo que conecta con todos.
 - Hay *clusters* de nodos.
 - Hay nodos sin ninguna conexión.

Dataset de red - Visualización

Existen muchas formas de visualizar estos datos.

- [Graph Visualization Introduction / Brian Staats / Observable](#)

En esta clase abordaremos 3 principales *idioms*:

- Nodo-enlace.
- Matriz de adyacencia.
- *Idioms* de contención.

6 WAYS TO VISUALIZE GRAPHS

ROBERT GOVE • TWO SIX LABS

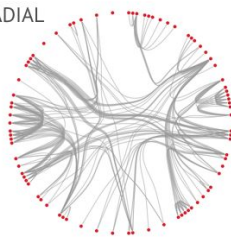
FORCE DIRECTED



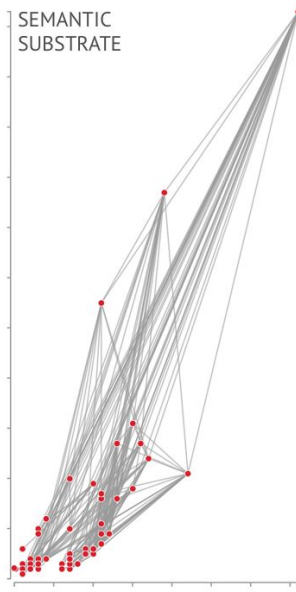
GROUP IN A BOX



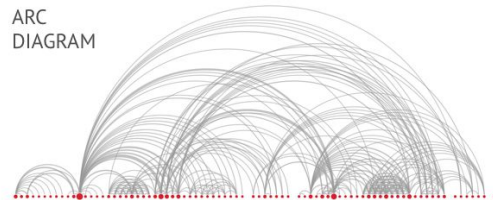
RADIAL



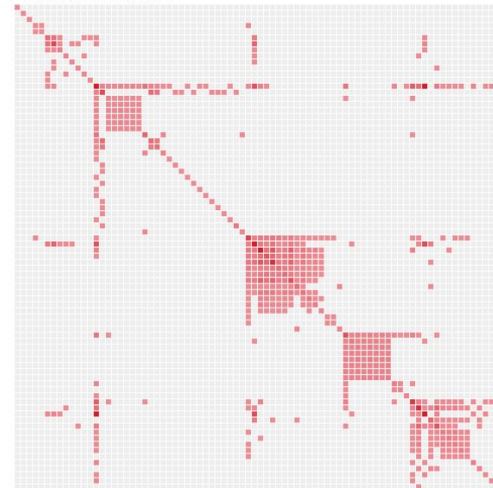
SEMANTIC SUBSTRATE



ARC DIAGRAM



MATRIX DIAGRAM

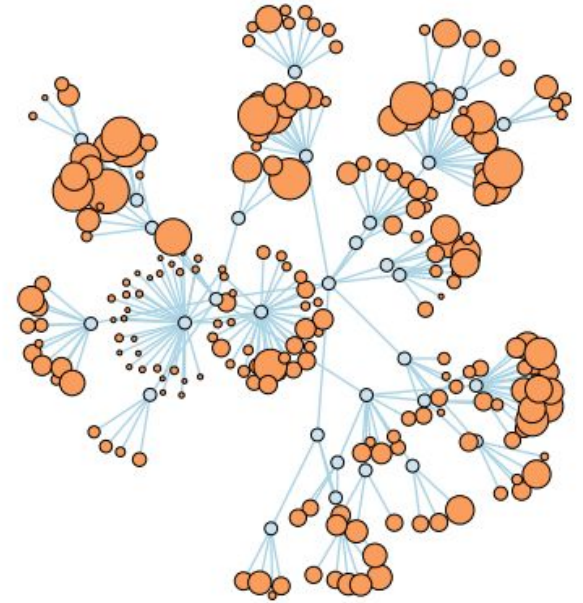


Nodo-enlace y Matriz de Adyacencia

Nodo-enlace y Matriz de Adyacencia

Nodo-enlace

- También conocido como **grafo**.
- Representa visualmente redes usando marcas de **puntos** para los nodos y marcas de **líneas** para mostrar la conexión entre pares de nodos.
- Visualmente, podemos codificar información de **diferente tipo en los nodos**: tamaño, color, y posición.
- En el caso de los enlaces, también podemos codificar información usando los canales de ancho, color y tipo de línea. También puede usarse una **flecha** (en lugar de línea) para identificar tipos de relaciones.

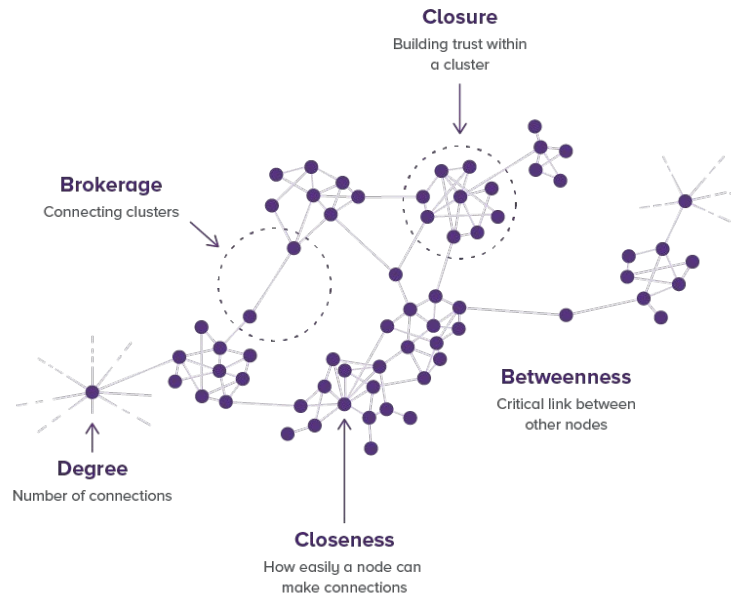


Nodo-enlace y Matriz de Adyacencia

Nodo-enlace

Idiom ideal para entender la topología de la red.

- Encontrar todos los caminos usando conexiones de la red entre dos nodos específicos.
- Encontrar el camino más corto entre dos nodos.
- Encontrar todos los nodos directamente adyacentes a un nodo específico.
- Encontrar nodos importantes dentro de la red en términos de conectividad.

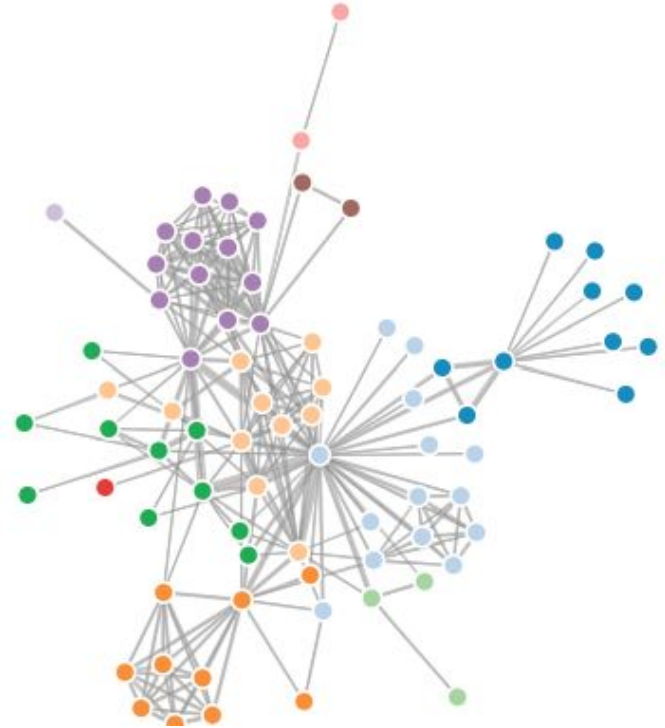
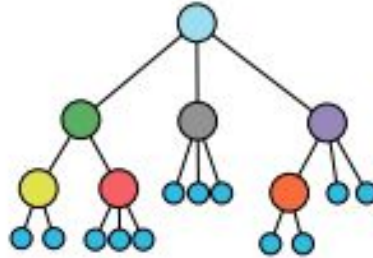
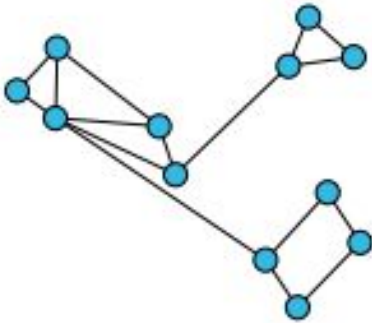


Nodo-enlace y Matriz de Adyacencia

Nodo-enlace

Una de las tareas más difíciles es intentar que el grafo se vea lo "más entendible posible".

- No se superpongan los nodos.
- Evitar la intercepción de líneas.

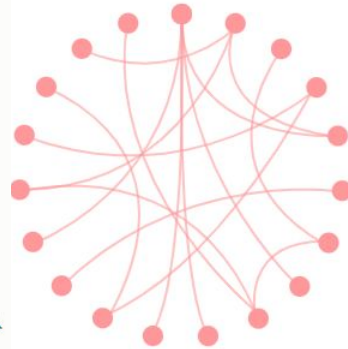
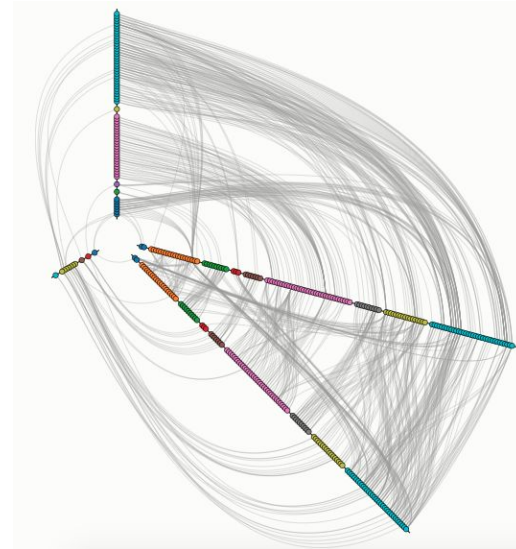
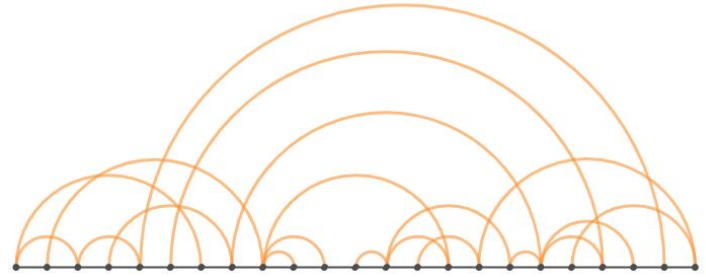


Nodo-enlace y Matriz de Adyacencia

Nodo-enlace

🤔 ¿Cómo determinar la posición de los nodos?

- Existen *idioms* del tipo nodo-enlace que fijan las posiciones de los nodos. Aunque su implementación es difícil de realizar.
 - **Diagrama de Arco** - Utiliza un eje (X o Y) para posicionar todos los nodos.
 - **Diagrama de Cuerdas Sin Cinta** - Utiliza una disposición radial para posicionar los nodos.
 - **Hive plot** - Genera múltiples ejes.

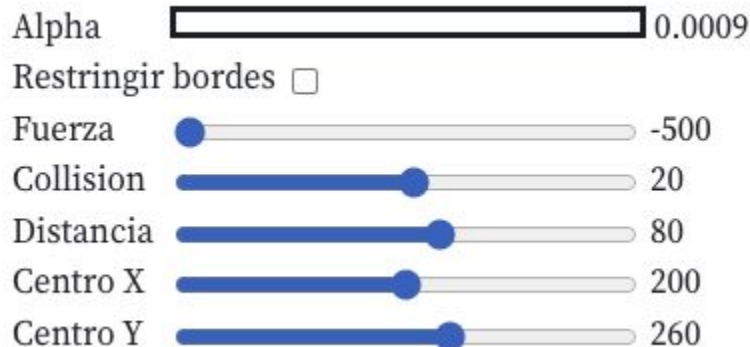
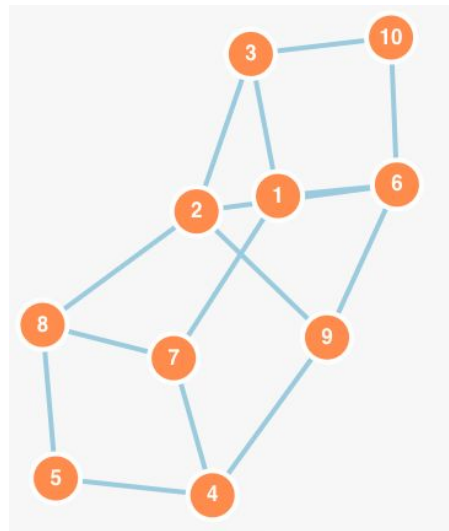


Nodo-enlace y Matriz de Adyacencia

Nodo-enlace

🤔 ¿Cómo determinar la posición de los nodos?

- Simular fuerzas físicas como si fueran partículas.
- Los enlaces provocarán, en la simulación, que 2 nodos se intenten juntar, mientras que la ausencia de estos permitirá que se alejen los nodos.
- Esta técnica hace que la posición ya no codifique ninguna información.
- Más fácil de implementar y barata computacionalmente que los otros enfoques.



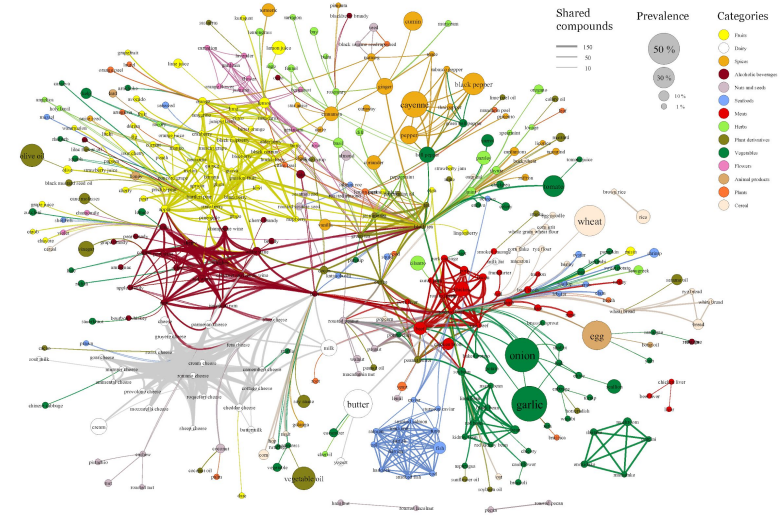
Nodo-enlace y Matriz de Adyacencia

Nodo-enlace - Ejemplo

- Cada nodo denota un ingrediente.
- El color del nodo indica la categoría de alimentos.
- El tamaño del nodo codifica la cantidad de recetas que usan dicho ingrediente.
- Dos ingredientes están conectados si comparten un número significativo de compuestos de sabor.
- El grosor del enlace representa el número de compuestos compartidos entre los dos ingredientes.

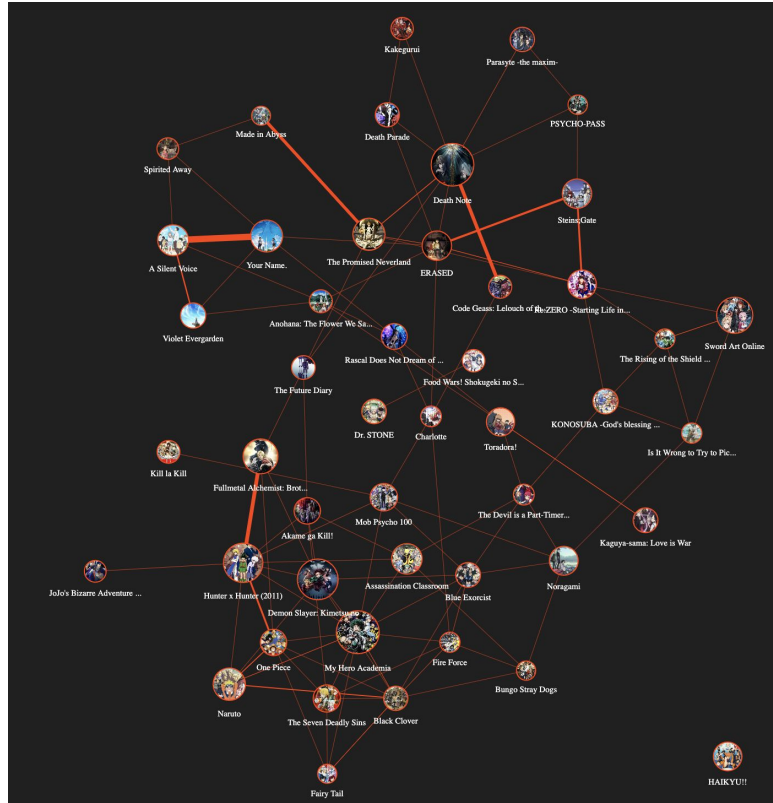
Flavor Network

Yong-Yeol Ahn, Sebastian Ahnert, James P. Bagrow, and A.-L. Barabási



Flavor network. Culinary ingredients (nodes) and their chemical relationship are illustrated. The color of each ingredient represents the food category that the ingredients belong to, and the size of an ingredient is proportional to the frequency we use (collected from online recipe databases) experiences on, also (spices, ingredients). Two culinary ingredients are connected if they share more flavor compounds. We extracted the list of flavor compounds in each ingredient from the food "chemically breakdown" of flavor ingredients (left) and then applied a backbone extraction method by Serres et al. (2015) and apply to pick statistically significant links between ingredients. The thickness of an edge represents the number of shared flavor compounds. To reduce clutter, edges are bundled based on the algorithm by Dorner (2010) (<http://www.cse.cmu.edu/~dorner/>).

Nodo-enlace y Matriz de Adyacencia



Nodo-enlace - Ejemplo 2

- Cada nodo es una serie.
- El tamaño del nodo codifica la cantidad de personas que ven esa serie.
- Dos series están conectadas si hay muchas personas que vieron ambas series.
- El grosor del enlace representa el número de personas que vieron ambas series.

Nodo-enlace y Matriz de Adyacencia

Nodo-enlace

Los ejemplos anteriores transmiten un mensaje:

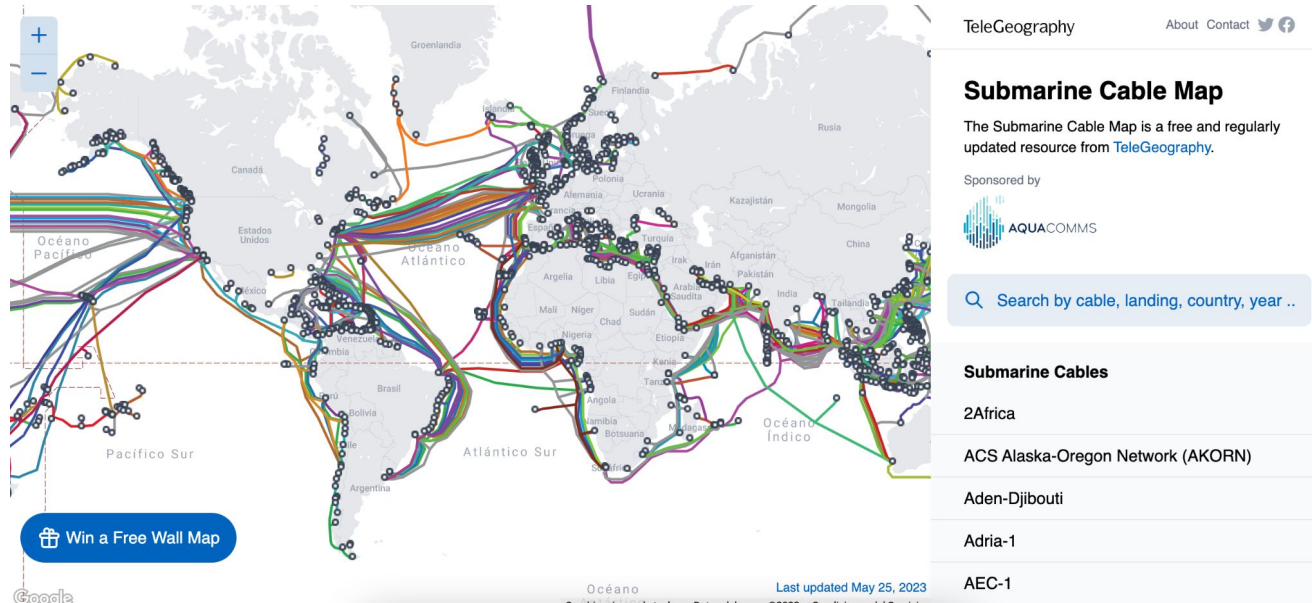
No necesariamente debes tener *datasets* intrínsecamente de relaciones (como red social, red de tráfico) para recurrir a visualizaciones de red.

Puedes definir alguna característica para "relacionar" 2 ítems.

Nodo-enlace y Matriz de Adyacencia

Nodo-enlace

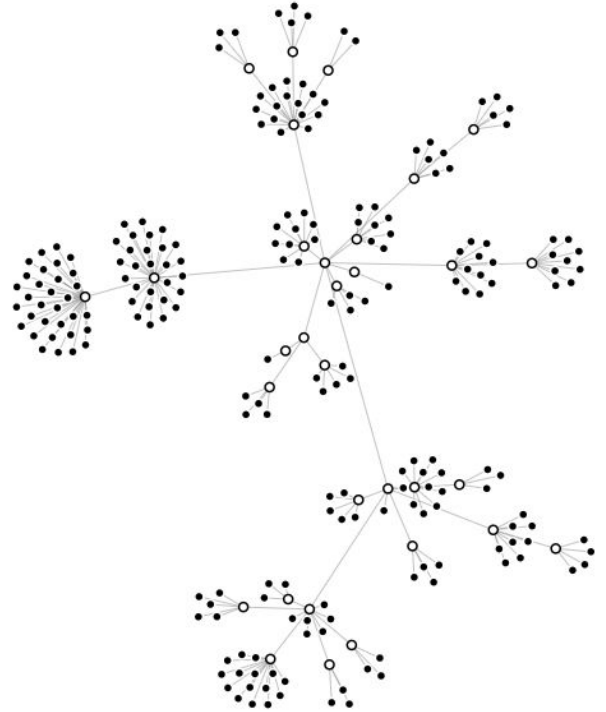
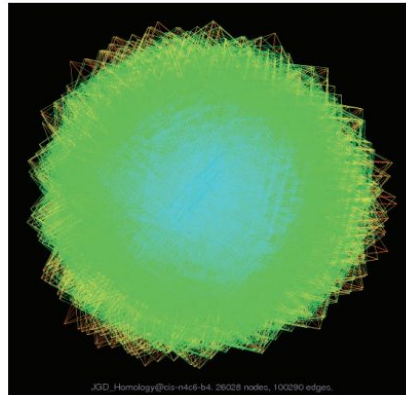
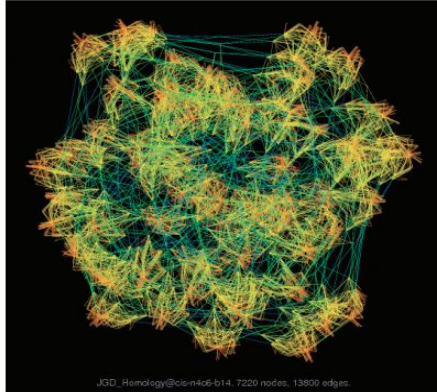
- También se puede ver como un mapa de flujo.
- Los enlaces unen el inicio y fin de un cable submarino.



Nodo-enlace y Matriz de Adyacencia

Nodo-enlace

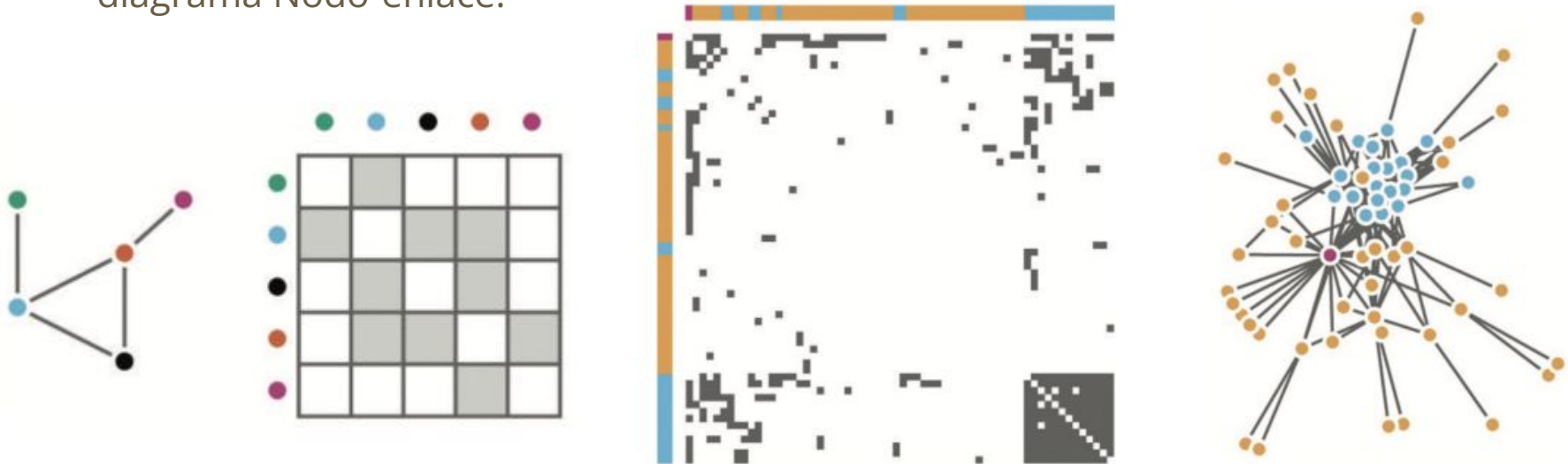
- **Ventaja:** Intuitivo para el humano de entender
- **Desventaja:** la escalabilidad. Si hay **más de 4 veces** enlaces que nodos, o son miles de nodos, pueden generarse bolas de pelos y se recomienda utilizar otro *idiom*.



Nodo-enlace y Matriz de Adyacencia

Matriz de Adyacencia

- Disposición de los nodos en los bordes de una matriz. Cada celda pintada indica la presencia de una conexión.
- Suele ir bien con redes grandes aunque es menos intuitiva de entender que el diagrama Nodo-enlace.



Fuente: Visualization Analysis & Design, Tamara Munzner y [Les Misérables Co-occurrence](#)

Nodo-enlace y Matriz de Adyacencia

Matriz de Adyacencia

Algunas tareas eficientes son:

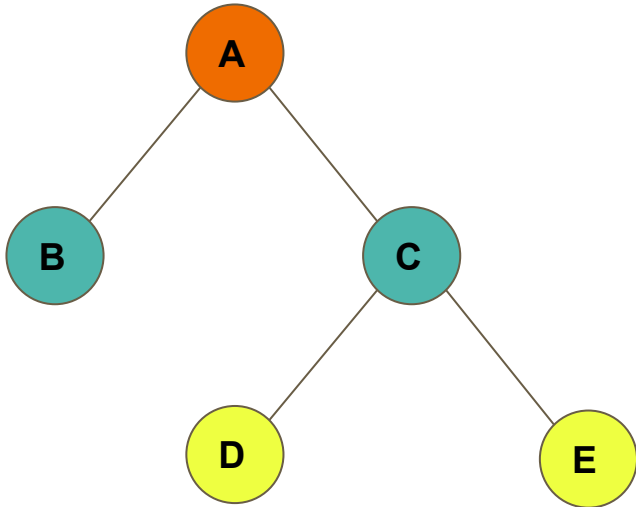
- Estimación del número de nodos. → Tamaño de la matriz
- Estimación del número de enlaces. → Cantidad de celdas pintadas
- Encontrar los nodos más conectados. → Filas/columnas muy pintada
- Encontrar un nodo según etiqueta. → Explorar nombres de fila/columna
- Encontrar un enlace entre un par de nodos. → Intercepción entre fila/columna

Jerarquías y Árboles

Jerarquías y Árboles

Caso particulares de red donde **no existen ciclos** en sus conexiones.

- Dado un nodo X, no existe un camino que parta en dicho nodo y termina en el mismo nodo pasando solo 1 vez por cada conexión del camino.

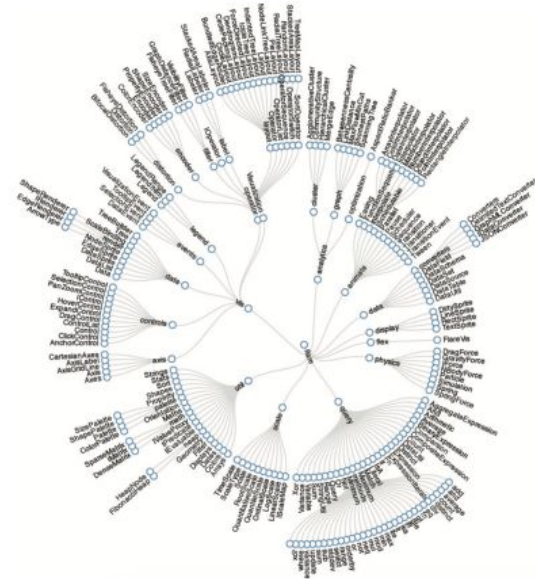
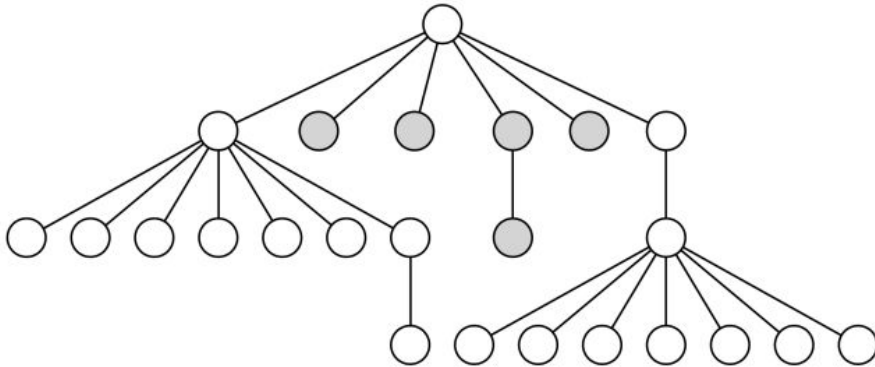


En este tipo de red, hay 4 conceptos importantes:

1. **Nodo hijo:** es el nodo receptor de una conexión. B es hijo de A, D es hijo de C.
2. **Nodo padre:** es el nodo de donde sale una conexión. A es padre de C, C es padre de E.
3. **Nodo hoja:** son los nodos que no tienen hijos. B, D y E son nodos hojas.
4. **Nodo raíz:** es el nodo que no tiene padre. A es nodo raíz.

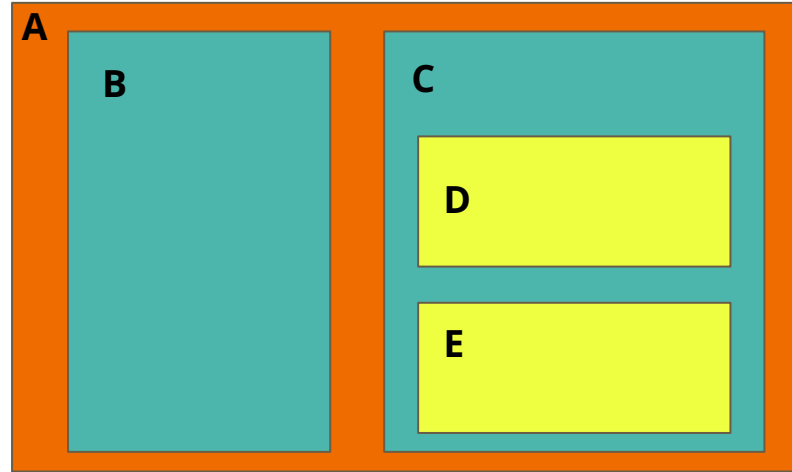
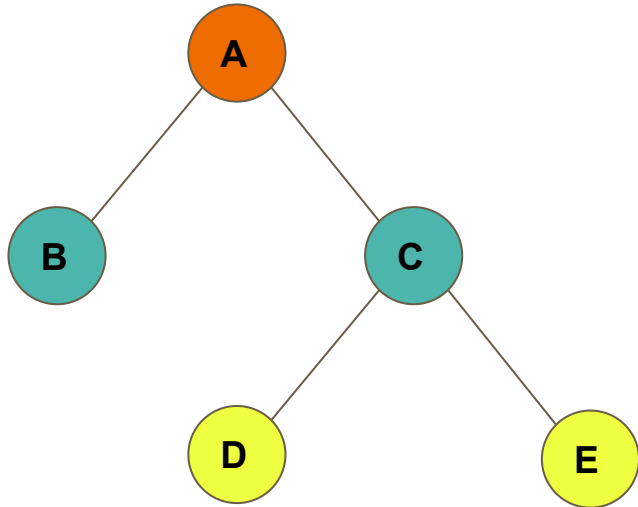
Jerarquías y Árboles

- Este tipo de dato se conoce como **árbol** o **de jerarquía**.
- Si bien podemos usar los *idioms* de nodo-enlace o matriz de adyacencia para codificar un árbol. El hecho de que exista un orden (la jerarquía) **las posibilidades de estos *idioms* cambian y aumentan.**



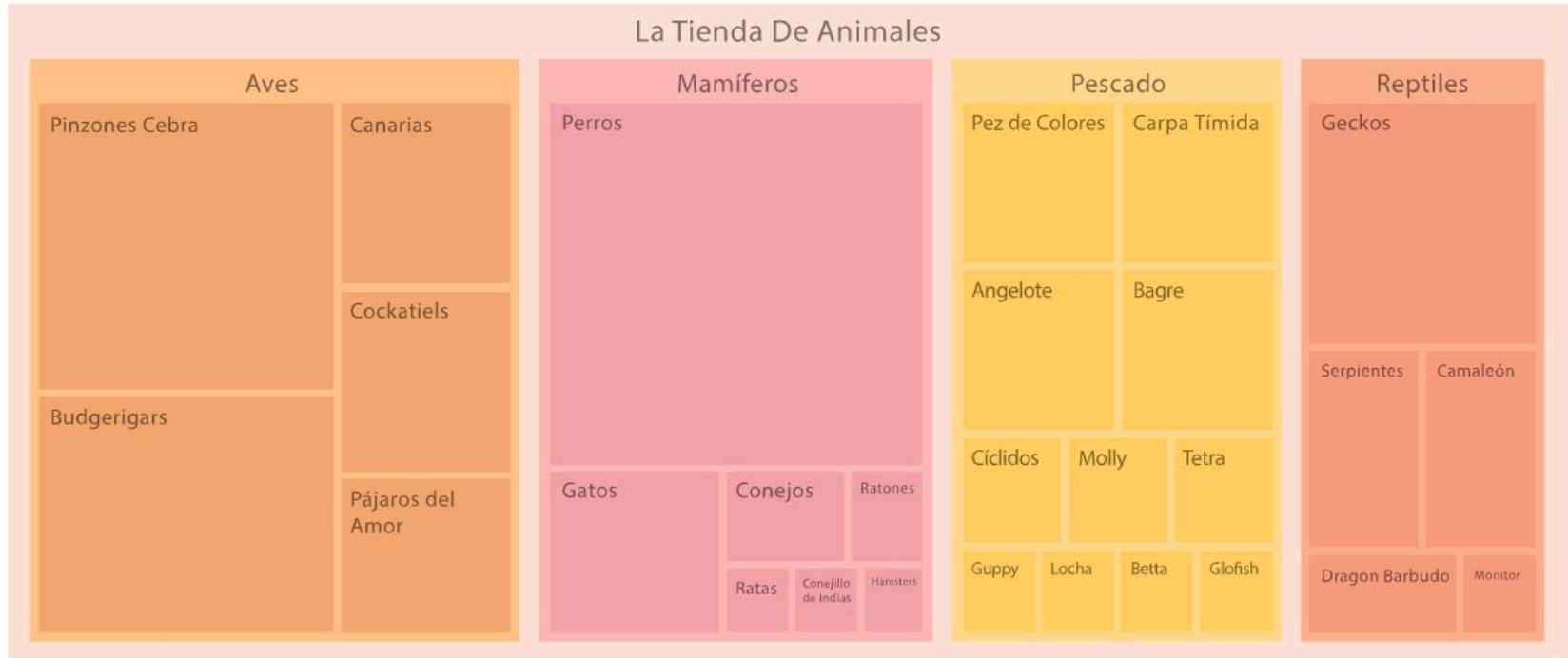
Jerarquías y Árboles

Es común recurrir al uso de **contención espacial** para comunicar jerarquía.



Jerarquías y Árboles

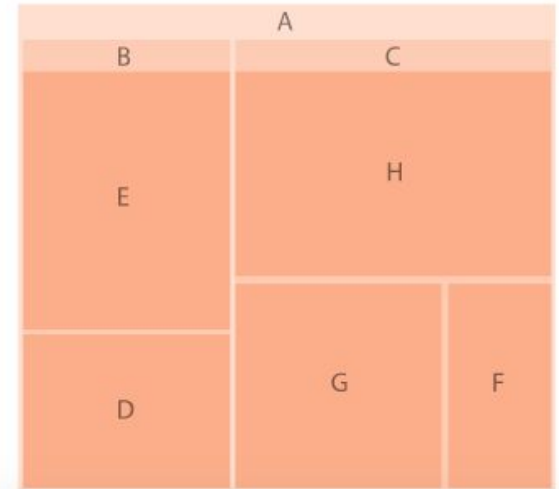
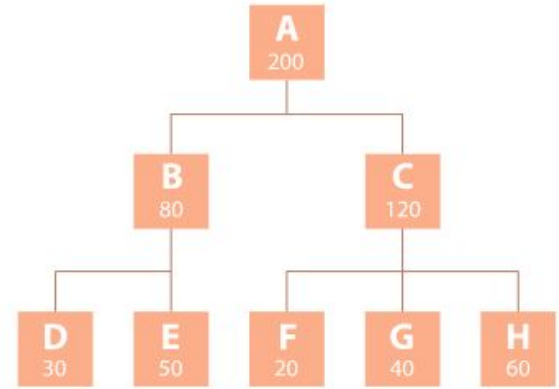
Treemap



Jerarquías y Árboles

Treemap

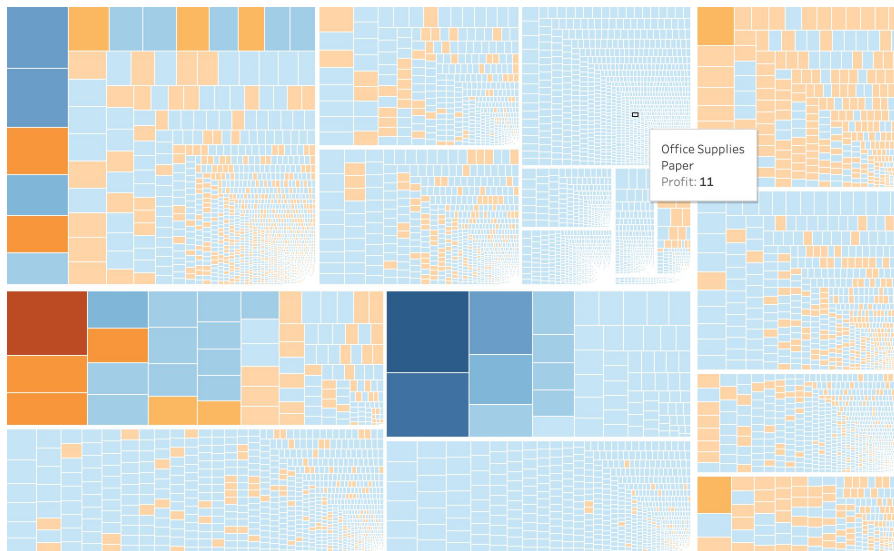
- Cada área representa un nodo.
- El tamaño **puede** codificar algún atributo numérico. Este canal se aplica a nodos hojas y el tamaño de los demás nodos se obtiene a partir de la suma de sus hijos.
 - Permite apreciar las contribuciones de distintas secciones de la jerarquía dentro de la totalidad de la red.
- El color **puede** codificar un atributo categórico.



Jerarquías y Árboles

Treemap

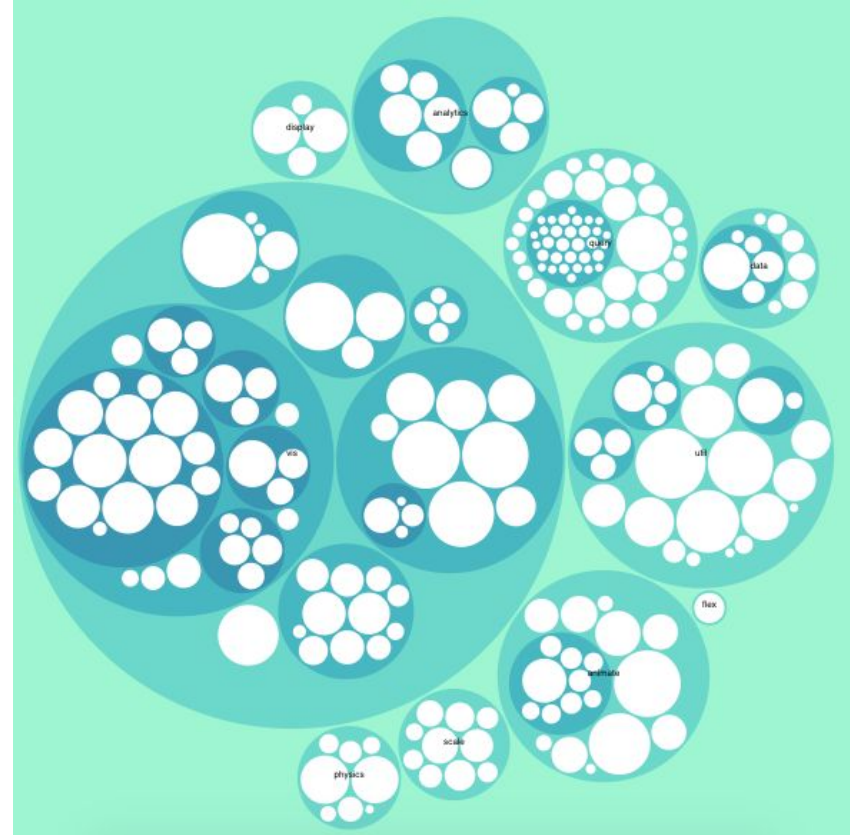
- Eficientes en espacio y buenos para hacer notar valores anormales para el atributo codificado.
- Ineficientes para tareas relacionadas a la topología de la red.



Jerarquías y Árboles

Circle packing

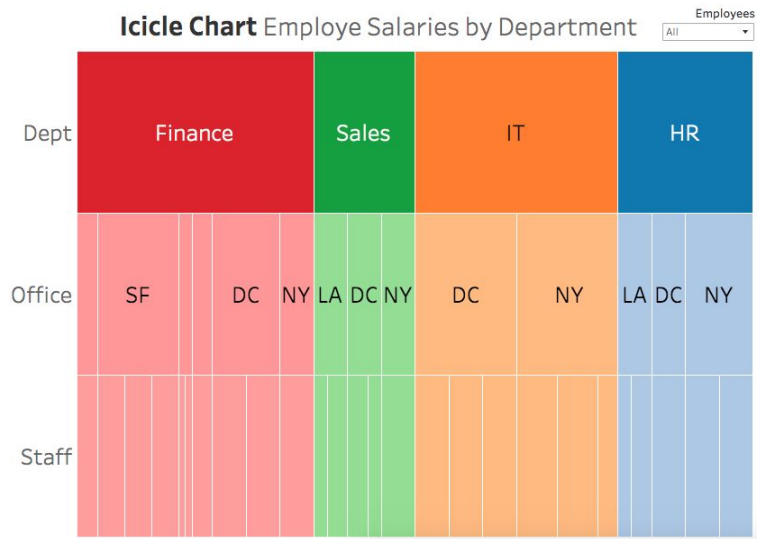
- Versión circular del *Treemap*.
- En muchos casos es más fácil de leer e interpretar que el *treemap*.
- Uso del espacio menos eficiente, y para jerarquías grandes el espacio necesario puede ser muy grande para apreciar detalles.



Jerarquías y Árboles

Icicle

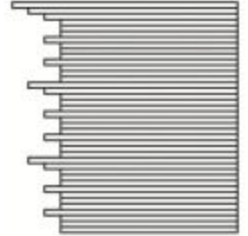
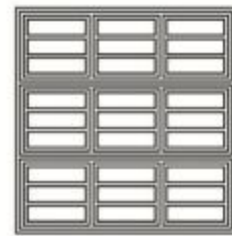
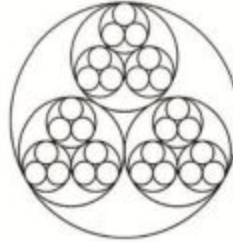
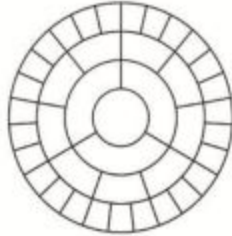
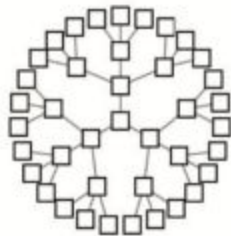
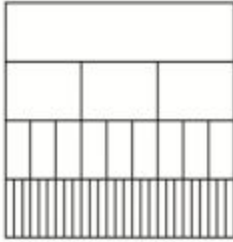
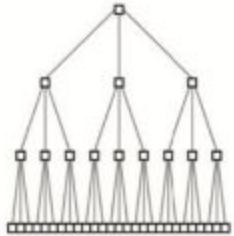
- Utilizar la posición espacial como método de codificación de jerarquía.
- Útil cuando la profundidad es baja.



Jerarquías y Árboles

¡Existen muchas opciones más!

Es nuestro trabajo evaluar cuál diseño es el adecuado para nuestros datos y tareas.



Resumen de conceptos importantes

- **Red (*network*):** tipo de *dataset* donde tenemos ítems (nodos en este caso) y conexiones/relaciones entre ítems (enlaces en este caso).
- **Grafo:** *idiom* donde cada nodo se representa con un área (generalmente círculo) y se usan líneas para representar la conexión.
- **Diagrama nodo-enlace:** otro nombre para un grafo.
- **Árbol:** tipo particular de red donde no hay ciclos en las conexiones.


Antes de salir... Revisión de contenidos (RC)

Se acaba de publicar un mini control de alternativas en Canvas sobre lo que **vimos en la clase de hoy**.

- **Duración:** 2 semanas para realizarlo a partir de hoy.
- **Intentos para responder:** ilimitados.
- **Extensión:** 8 preguntas de 1 punto c/u.
- **Condición para obtener el punto RC:** Al menos 7 puntos de 8.
- Cada vez que respondan, verán el puntaje total logrado, pero no cuáles preguntas están correctas e incorrectas.

Próximos eventos

Próxima clase

- Última clase donde veremos contenidos nuevos D3.
- Visualización de redes y árboles en D3 

Ayudantía de mañana

- Aplicar *brush* para crear una visualización de más alto nivel.

Tarea 4 (última tarea)

- Se entrega mañana.

IIC2026

Visualización de Información

— Hernán F. Valdivieso López —
(2023 - 2 / Clase 22)
