

Visualización de la Información

Maestría en Explotación de Datos y Descubrimiento del Conocimiento

Clase teórica 3: Consideraciones para el diseño de visualización de datos (parte I)

Claudio Delrieux – cad@uns.edu.ar

Laboratorio de Ciencias de las Imágenes – www.imaglabs.org

Universidad Nacional del Sur – www.uns.edu.ar

Emmanuel Iarussi - earussi@conicet.gov.ar

FRBA UTN - CONICET

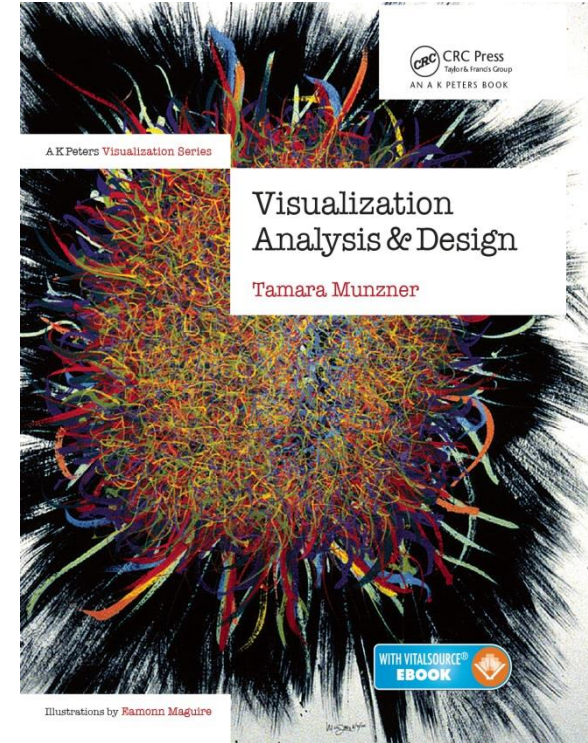
- **Disposición de la Información**
- Diseño de vistas (views)
- Validación
- Reglas y heurísticas

Consideraciones para el diseño de visualización de datos

“Vis(ualization) design is full of trade-offs, and most possibilities in the design space are ineffective for a particular task, so validating the effectiveness of a design is both necessary and difficult.

Vis designers must take into account three very different kinds of resource limitations: those of computers, of humans, and of displays.

Vis usage can be analyzed in terms of why the user needs it, what data is shown, and how the idiom is designed.”



Disposición de la Información en una Vista

La disposición de la información en una vista es un aspecto clave y esencial, dado que implícita o explícitamente utiliza el espacio 2D de la vista para organizar la representación (se utilice el espacio como metáfora visual o no).

El espacio es el máspreciado, preciso, y escaso de los atributos visuales, y también domina el modelo mental generado por la vista, y puede ser proclive a generar percepciones espúrias o indeseadas.

Armado de Vistas

Datos Tabulares	Datos Espaciales	Árboles – Grafos - Redes
Valores Cuantitativos Datos Matriciales Arreglos Radiales	Geometría Original Datos Geoespaciales Datos Anatómicos Geometrías Derivadas	Nodo-Arco Matrices de Adyacencia Estructuras Contenidas

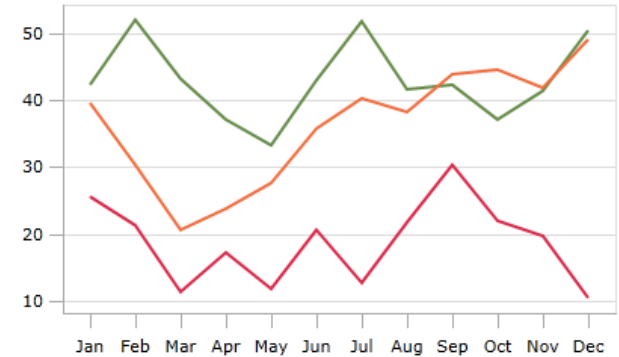
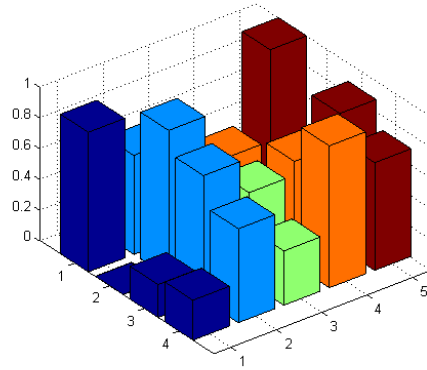
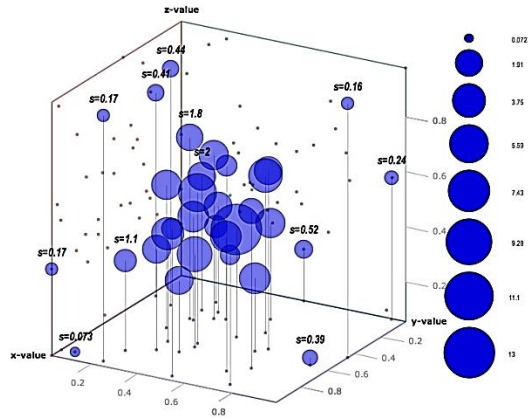
Armado de Vistas: Datos Tabulares

Los datos tabulares son especialmente comunes y si bien la ontología subyacente es sencilla, no siempre la generación de vistas es trivial.

Aún en los casos “rectangulares” (matrices de ítems vs. atributos), la existencia de dos o más dimensiones puede presentar dificultades a la hora de diseñar una visualización adecuada.

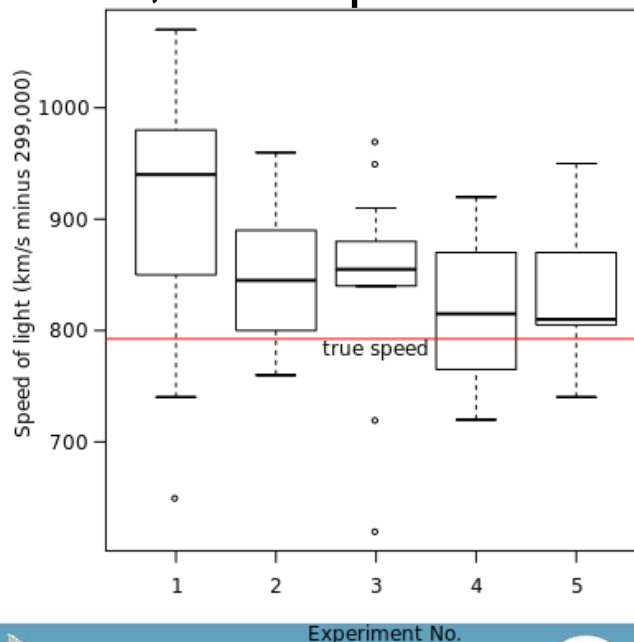
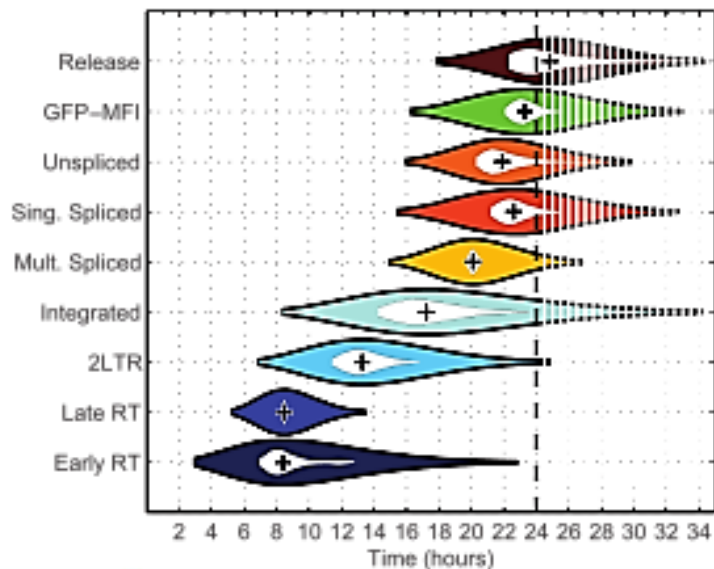
Datos Tabulares: Representación de Valores Cuantitativos

Los casos más sencillos fueron mencionados en clases pasadas (scatterplots, histogramas, histogramas agrupados, coordenadas paralelas, line-charts, etc.).



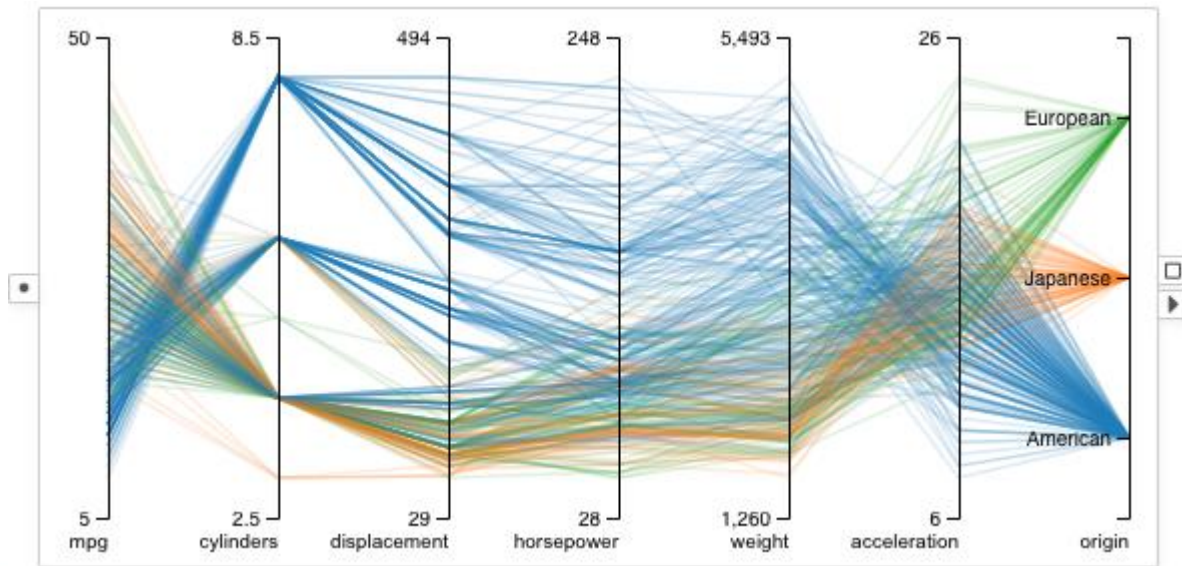
Datos Tabulares: Representación de Distribuciones

La representación de distribuciones permite interpretar los parámetros estadísticos de las mismas, o comparar varias.



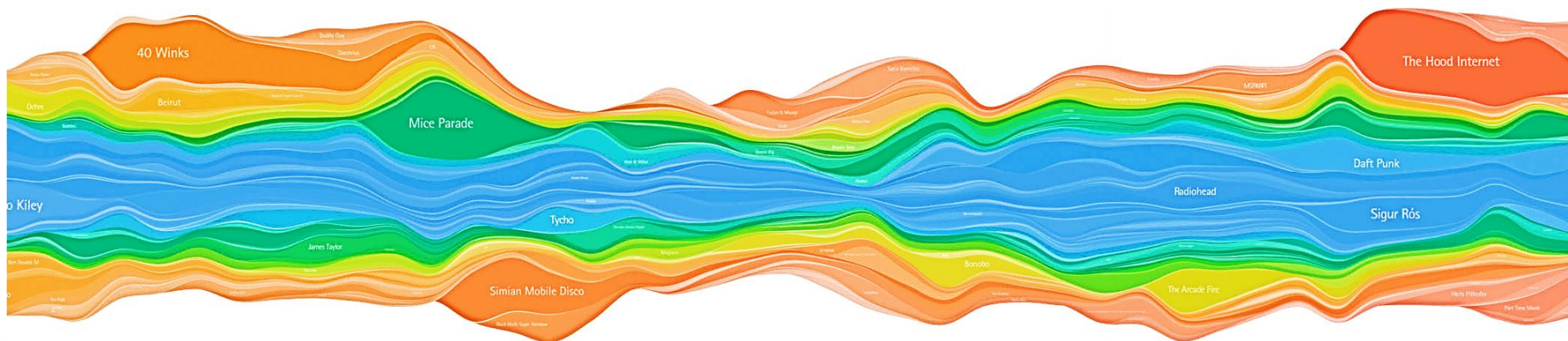
Datos Tabulares: Representación de Valores Cuantitativos

Los datos tabulares $1N \rightarrow nQ$ son muy y claramente los scatterplot no escalan. Para eso conviene utilizar coordenadas paralelas.



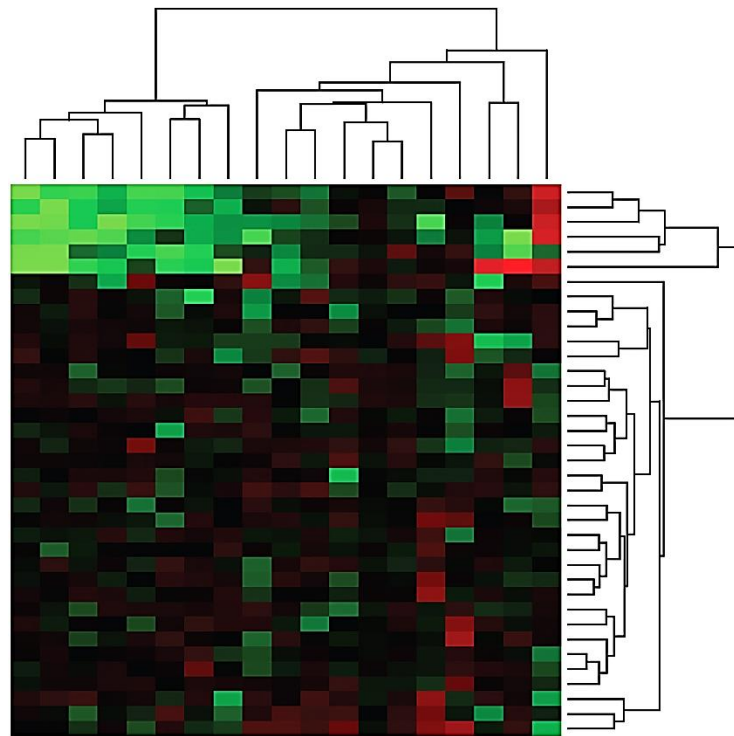
Datos Tabulares: Representación de Valores Cuantitativos

Una manera de representar datos tabulares 1Q ->nQ es a través de *streamgraphs*. Si bien representa la misma información que N line-charts coordinados, algunas percepciones son más precisas.



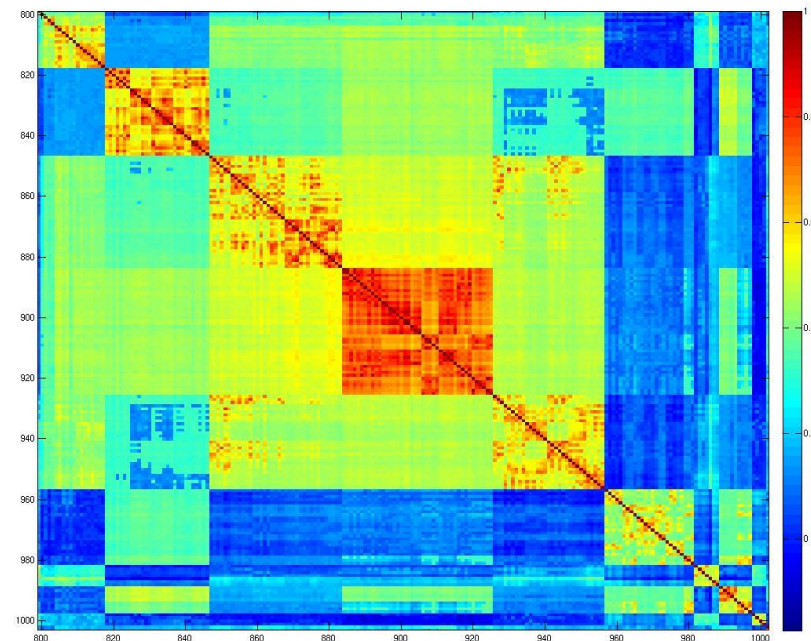
Datos Tabulares: Representación de Datos Matriciales

La representación de matrices de datos cuantitativos con alturas es muy difícil de interpretar, por lo que es común utilizar color al estilo de los mapas de calor.



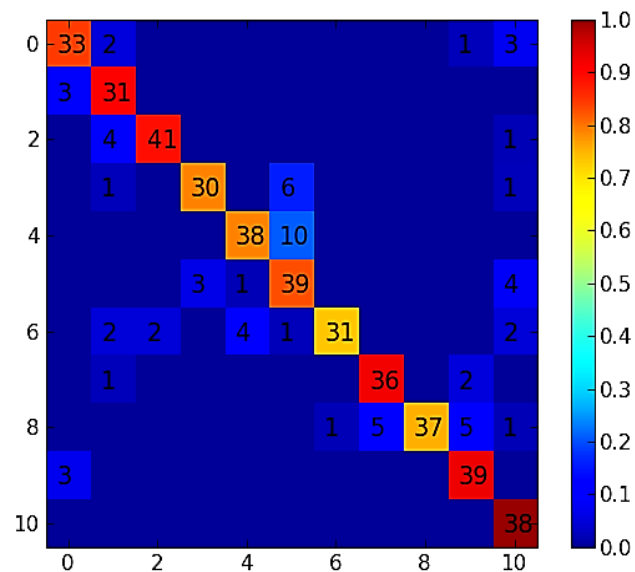
Datos Tabulares: Representación de Datos Matriciales

Un caso particular en nuestra carrera son las matrices de covariancia, dado que la visualización puede mostrar con más claridad la correlación entre variables.



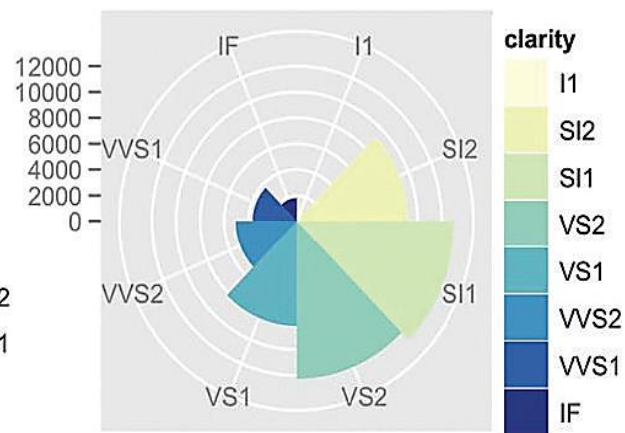
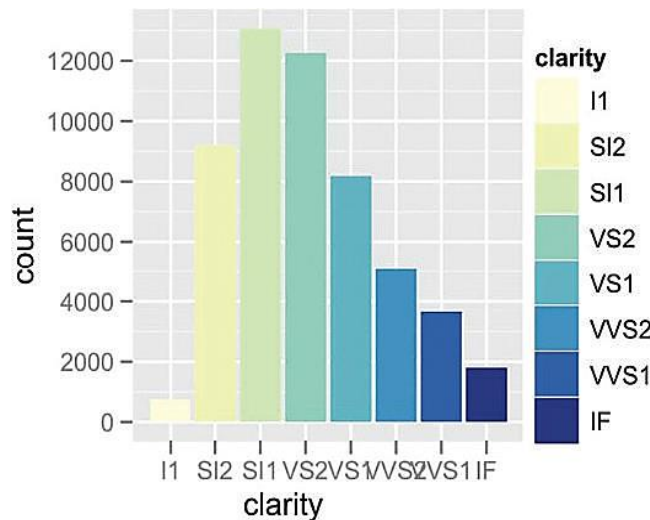
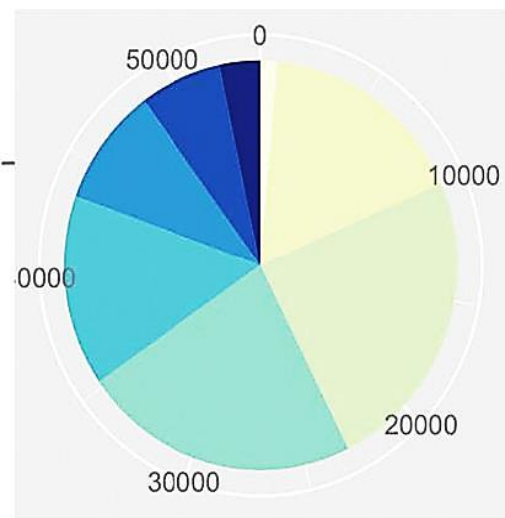
Datos Tabulares: Representación de Datos Matriciales

Otro caso son las matrices de confusión (heat map, mosaic plot).



Datos Tabulares: Arreglos radiales

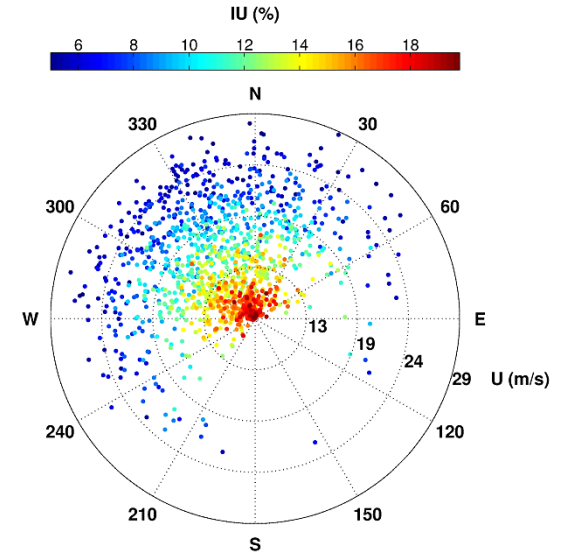
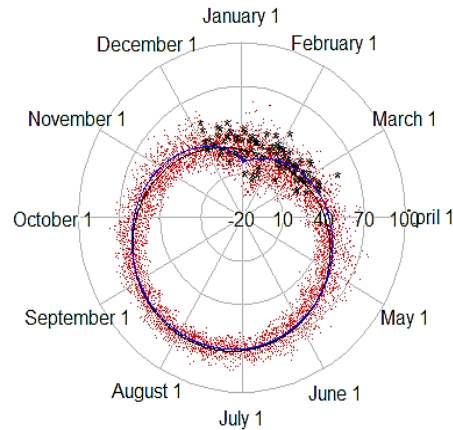
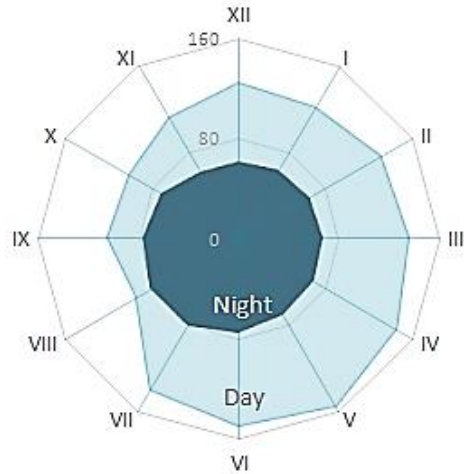
La exactitud depende de la percepción del área, de la longitud, y también de la orientación y el color:



Datos Tabulares: Arreglos radiales

Muchas veces son necesarios por el tipo de información a representar (p. ej. eventos periódicos, puntos cardinales).

Pageviews per Hour by Hour of the Day



Vistas y más vistas

<https://datavizcatalogue.com>



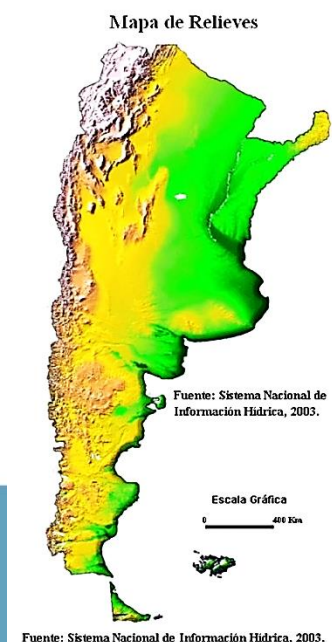
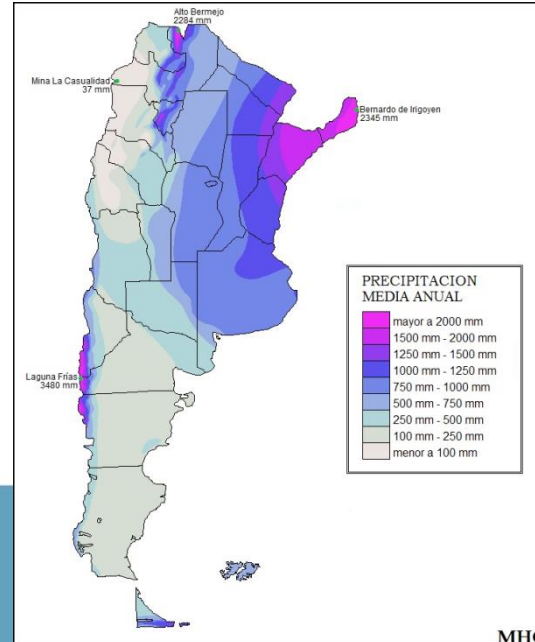
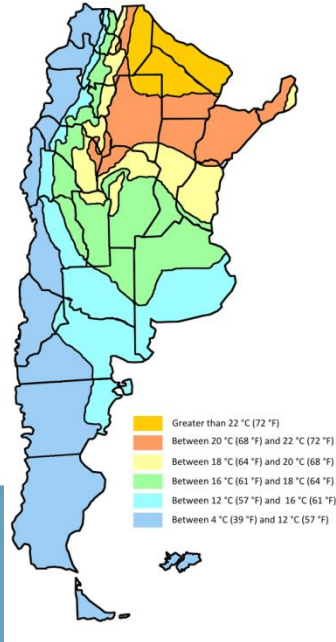
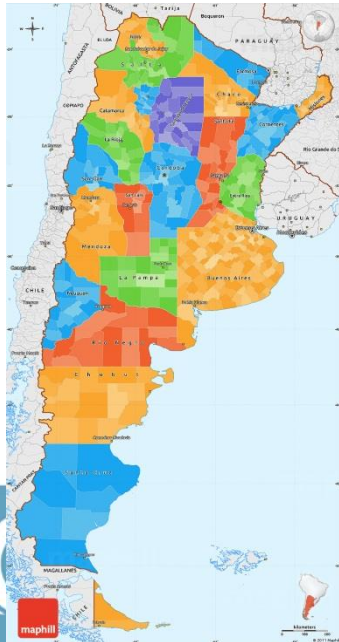
Armado de Vistas: Datos Espaciales

Los datos espaciales requieren que el espacio en la vista represente la geometría (posición) de su ubicación original. Los casos más destacados son los datos geográficos y las imágenes médicas.

En algunos casos el mapeo entre el espacio original y el de la vista no es lineal (por ejemplo, distorsiones), pero pese a ello debe existir siempre algún tipo de correspondencia.

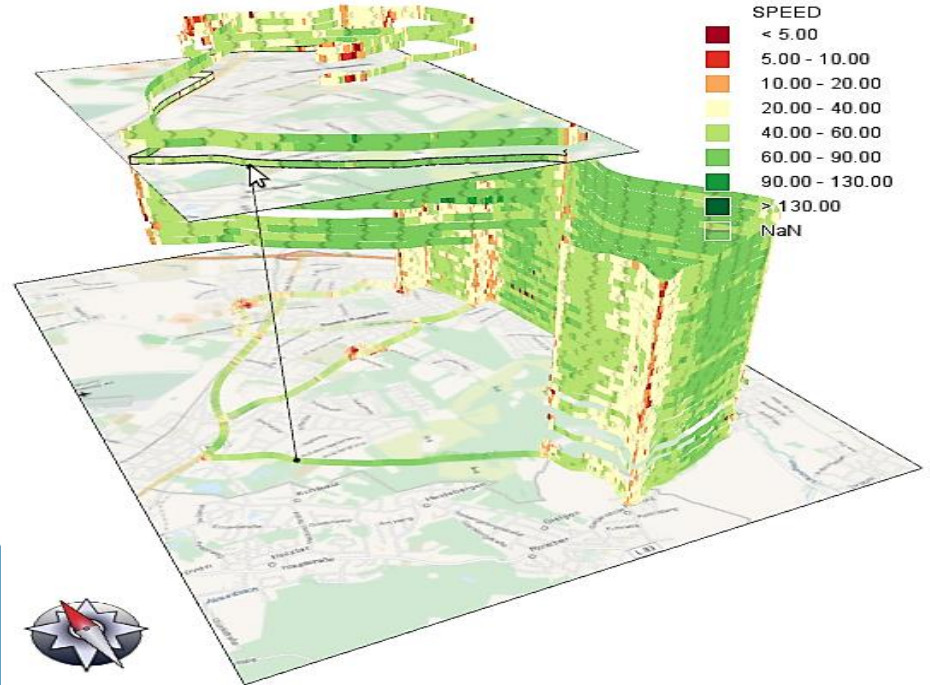
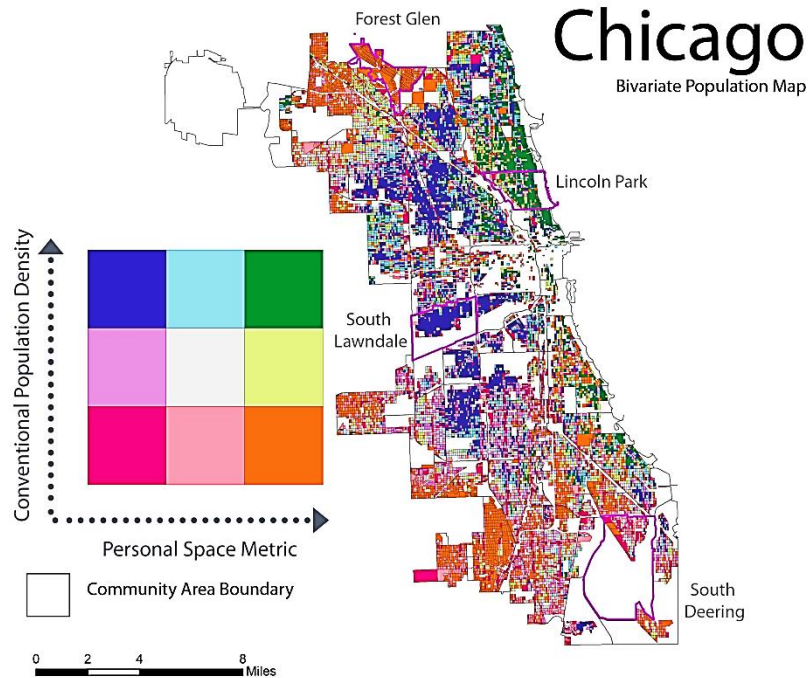
Datos Espaciales: Geometría original

El caso más sencillo de datos geoespaciales es la representación de valores 1DN, O ó Q sobre áreas geográficas (físicas o políticas) utilizando colores.



Datos Espaciales: Geometría original

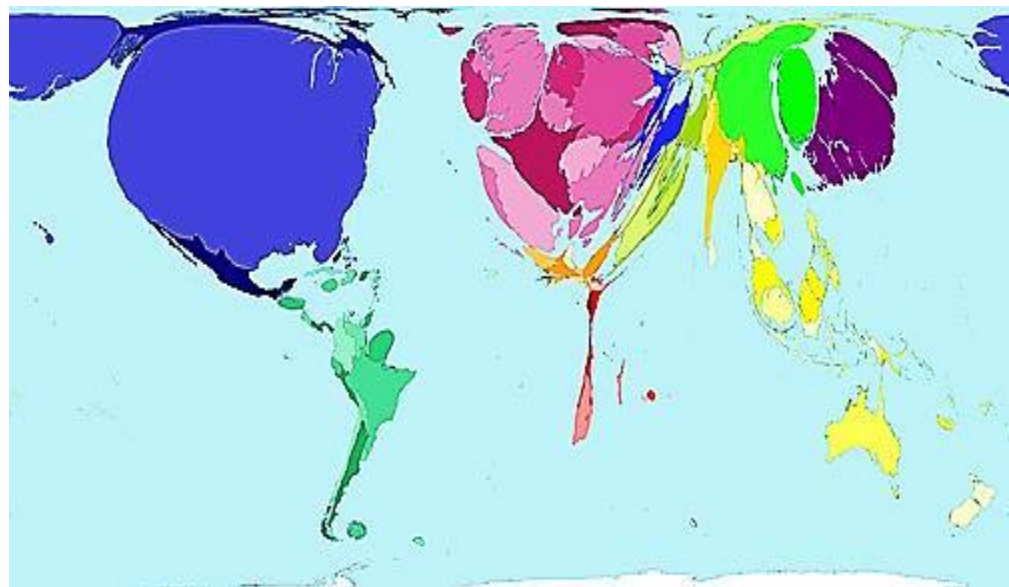
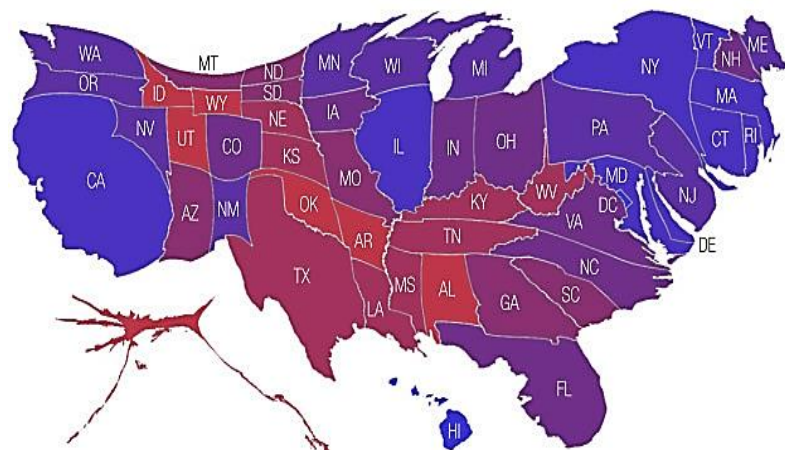
Para representación de valores geoespaciales 2D+ se requieren metáforas más elaboradas (paletas bivariadas, glifos, etc.).



Datos Espaciales: Geometría original

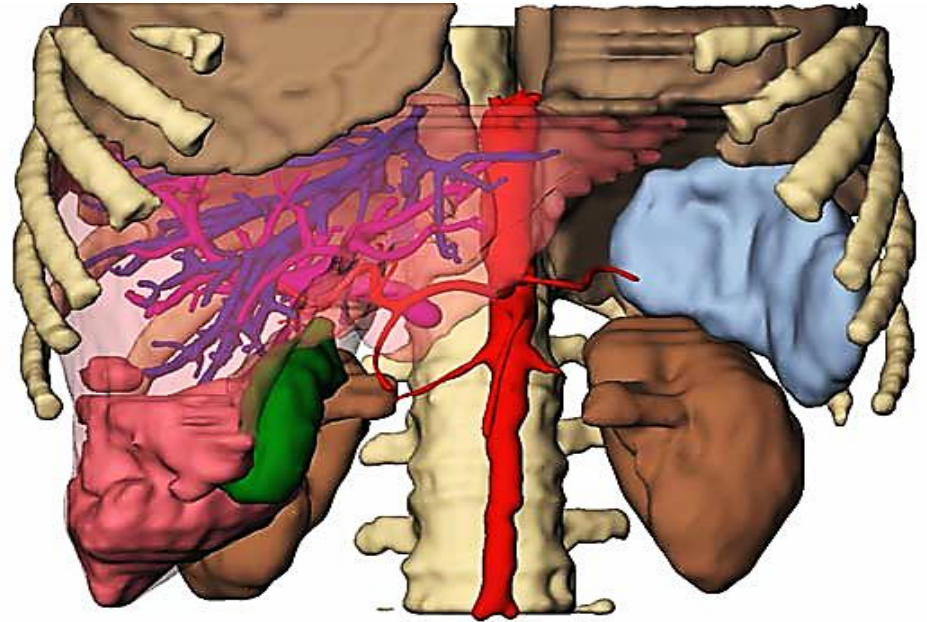
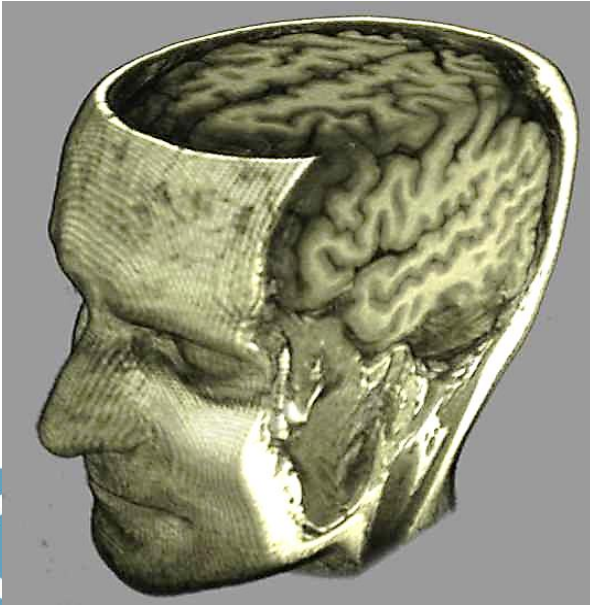
En algunos casos se utilizan escalas geográficas no isotrópicas.

Electoral Votes



Datos Espaciales: Geometría original

Para representación de datos anatómicos basados en valores escalares en un dominio 3D (CT, MR) se requiere el uso de transparencias o isosuperficies, o metáforas ad-hoc.



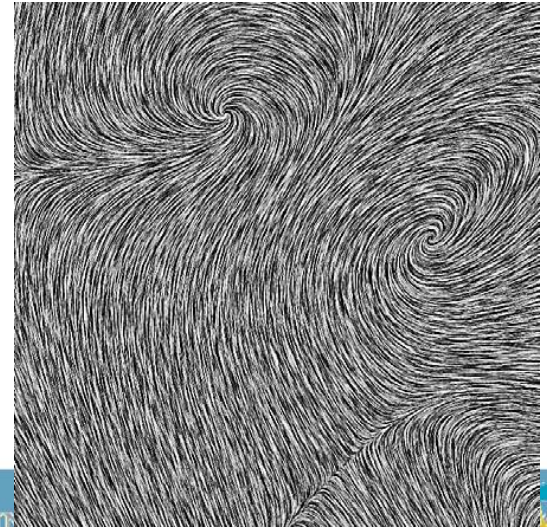
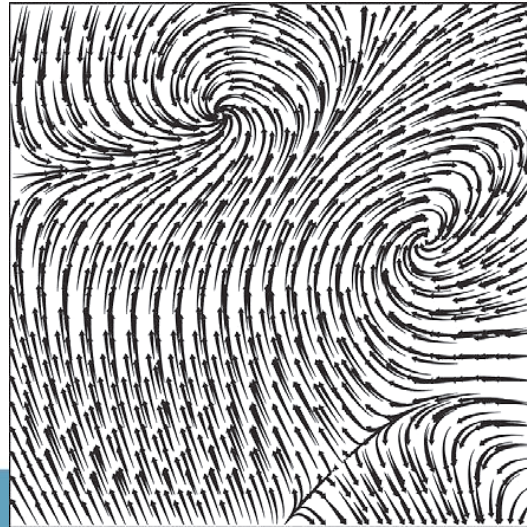
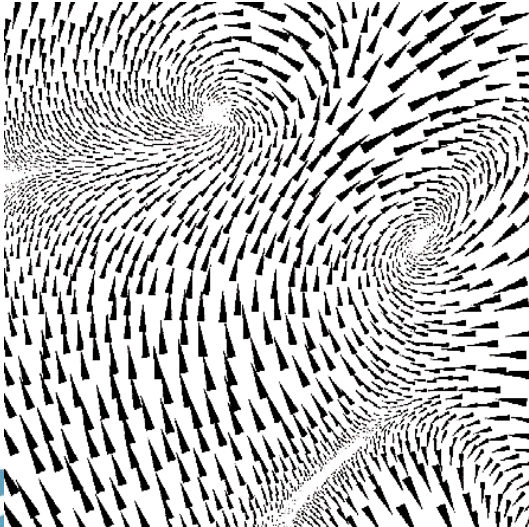
Armado de Vistas: Geometrías Derivadas

Una geometría derivada es la que surge al utilizar espacio como metáfora de una magnitud no espacial. Un ejemplo sencillo es la variación de una señal en el tiempo.

Las geometrías derivadas en dominios 2D o 3D ocurren típicamente en los *espacios de fases* de sistemas de ecuaciones utilizados para modelar un sinnúmero de fenómenos.

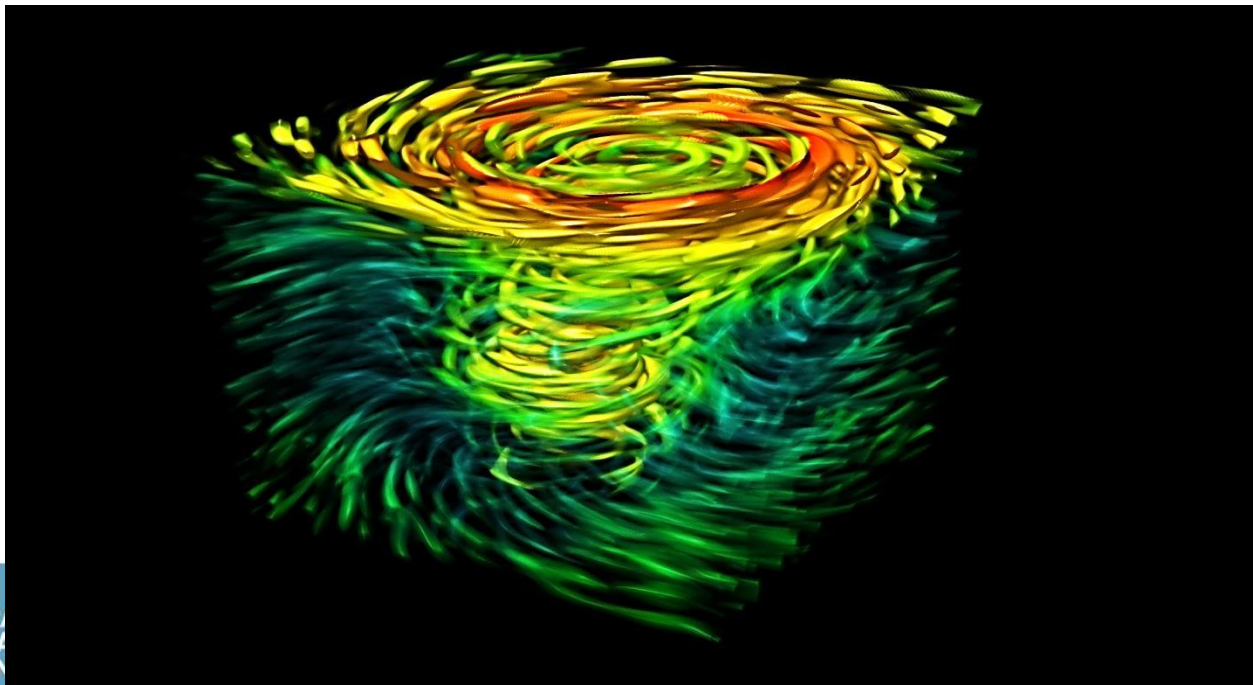
Geometrías Derivadas: Espacios de Fases

Un uso típico de espacios de fases representar la evolución de sistemas dinámicos, que modelan fenómenos muchas veces muy complejos en ecología, física, etc.



Geometrías Derivadas: Espacios de Fases

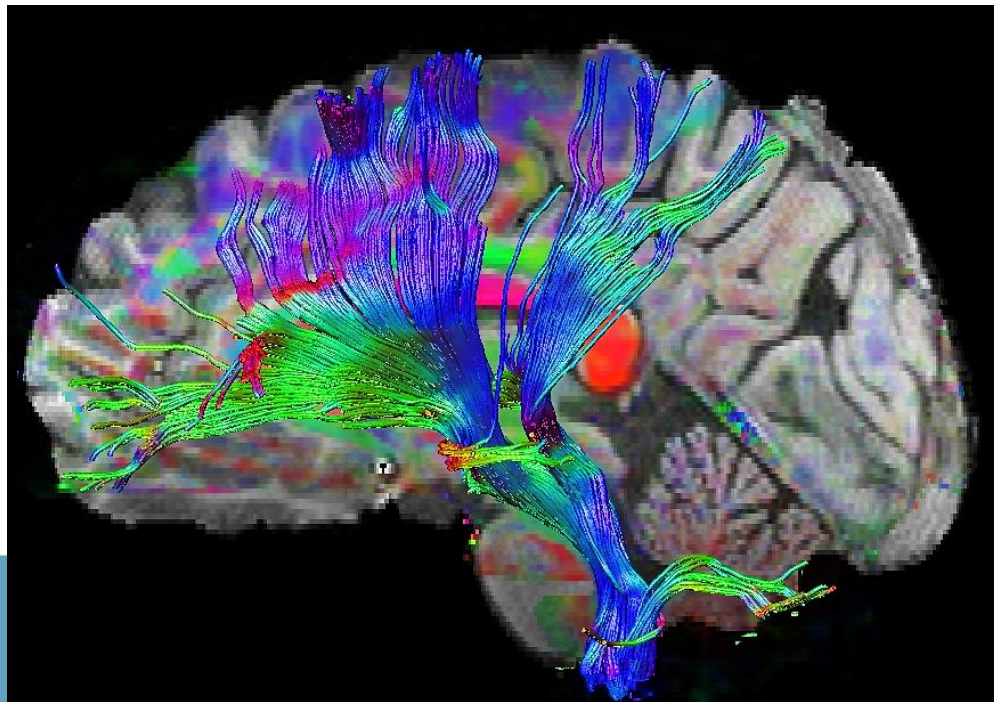
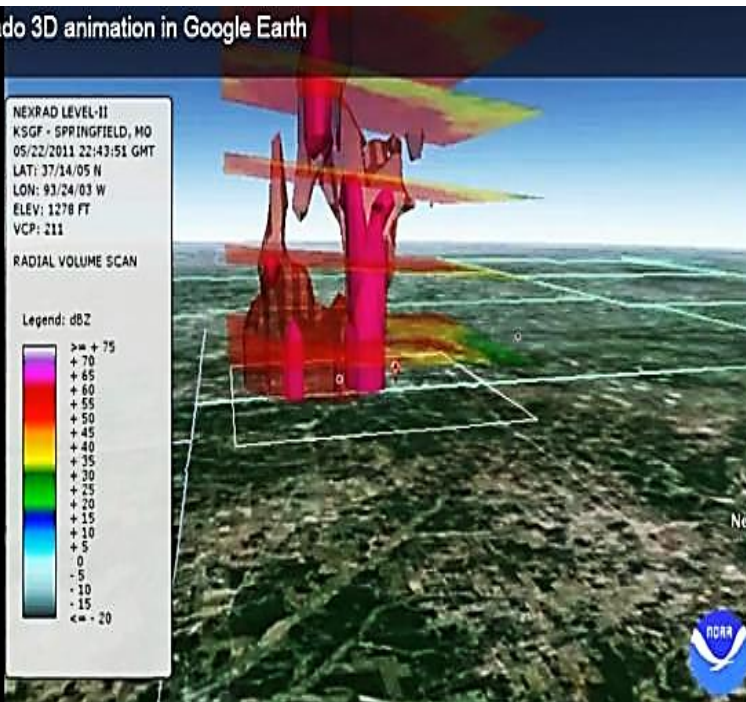
En muchas situaciones los datos en los espacios de fases son además multivaluados (por ejemplo, tensores).



Geometrías Derivadas: Espacios de Fases + Geometría

Finalmente, podemos tener datos tensoriales sobre geometrías (por ejemplo, tornados, TDI).

Joplin tornado 3D animation in Google Earth
by steve16624



Armado de Vistas: Grafos, Árboles, Redes

Visualizar estructuras de información donde la conectividad es el elemento esencial representa un particular desafío.

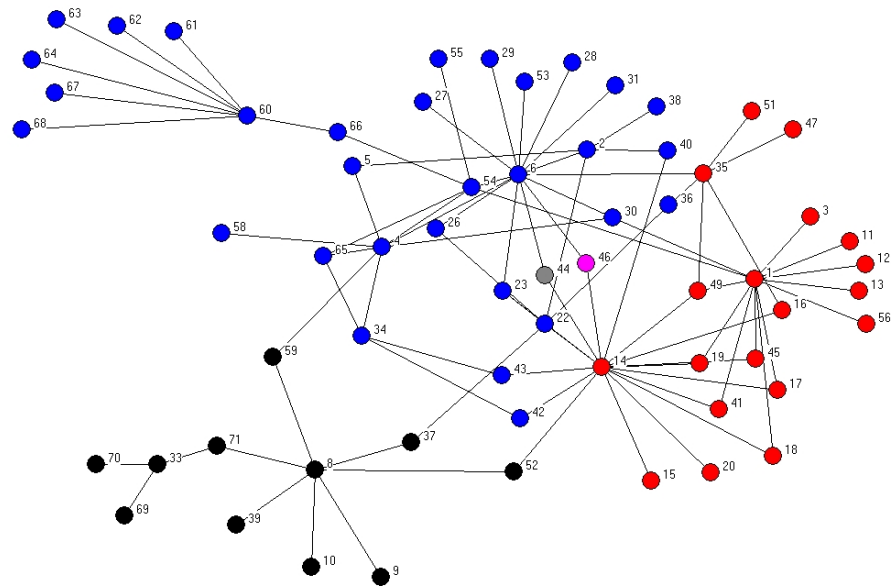
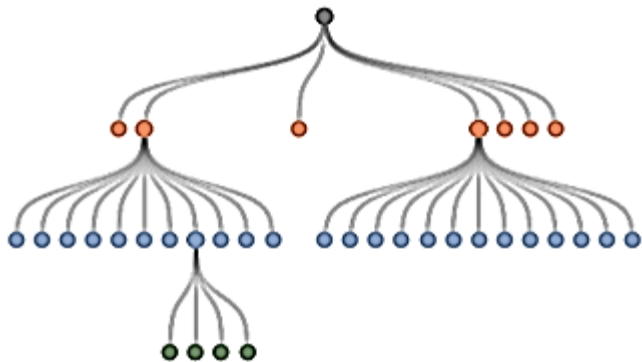
Las metáforas más utilizadas y naturalizadas se basan en representar gráficamente la relación **nodo-arco**, o en representar la **adyacencia**, o bien alguna forma de **inclusión**.

Armado de vistas: Resumen

Datos Tabulares	Datos Espaciales	Árboles – Grafos - Redes
Valores Cuantitativos Datos Matriciales Arreglos Radiales	Geometría Original Datos Geoespaciales Datos Anatómicos Geometrías Derivadas	Nodo-Arco Matrices de Adyacencia Estructuras Contenidas

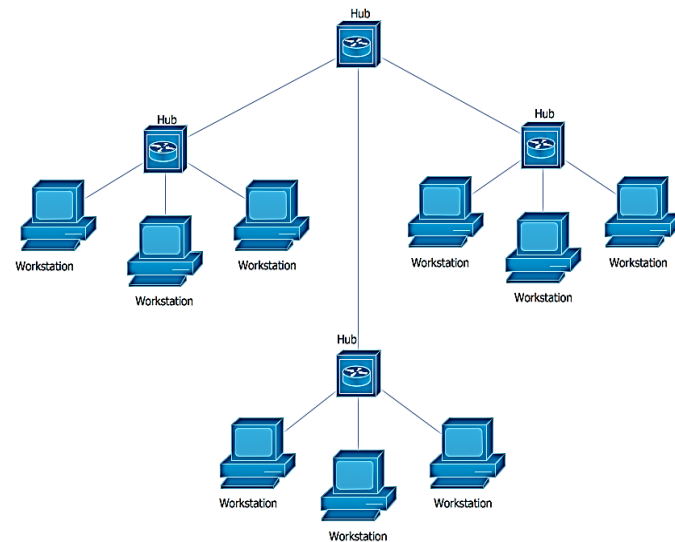
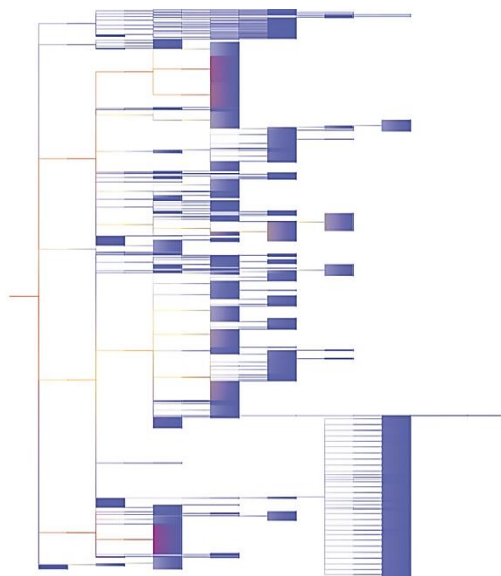
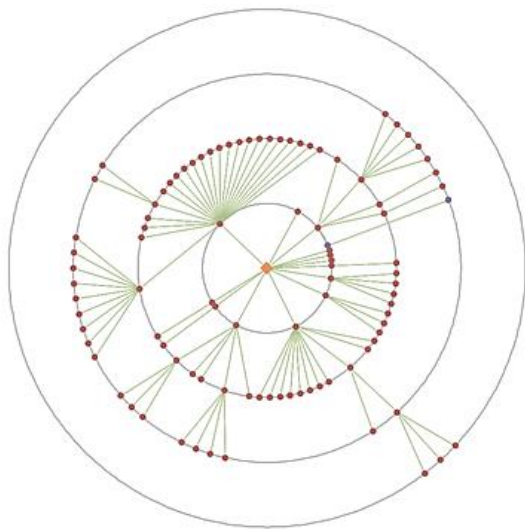
Grafos, Árboles, Redes: Nodo-Arco

Es la vista más usual y natural para grafos, árboles y redes.



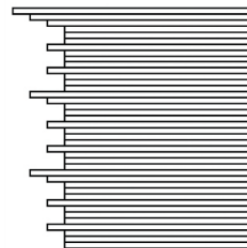
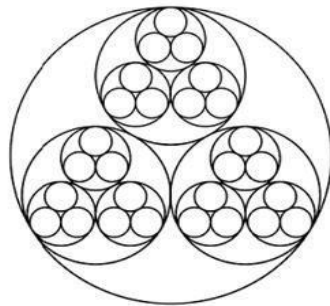
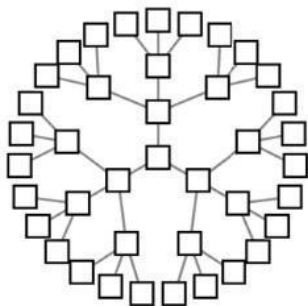
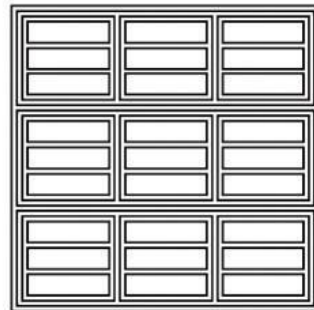
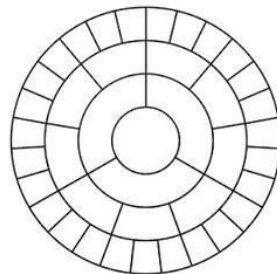
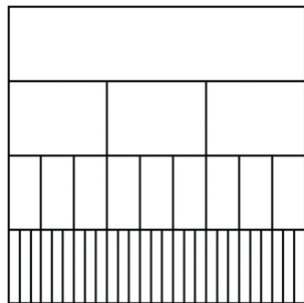
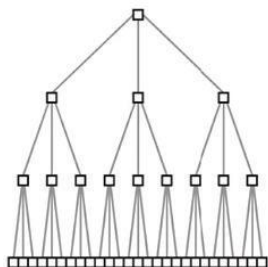
Grafos, Árboles, Redes: Nodo-Arco

Existe una gran cantidad de variantes (ver en viz-cat)



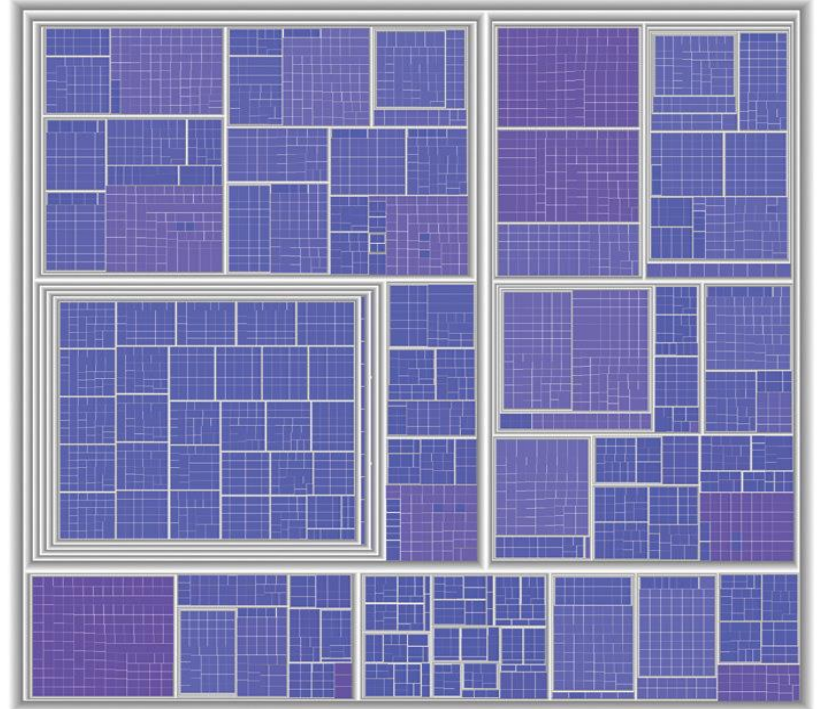
Grafos, Árboles, Redes: Nodo-Arco

En el caso de árboles, existen muchas vistas posibles.



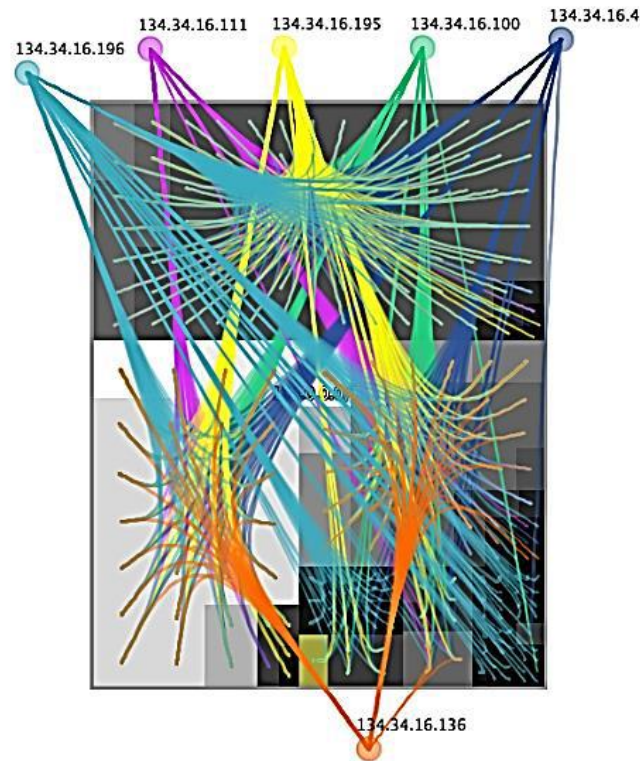
Grafos, Árboles, Redes: Estructuras contenidas

Los árboles son estructuras jerárquicas, por lo que pensar en estructuras que contienen a otras estructuras puede ser una buena representación.



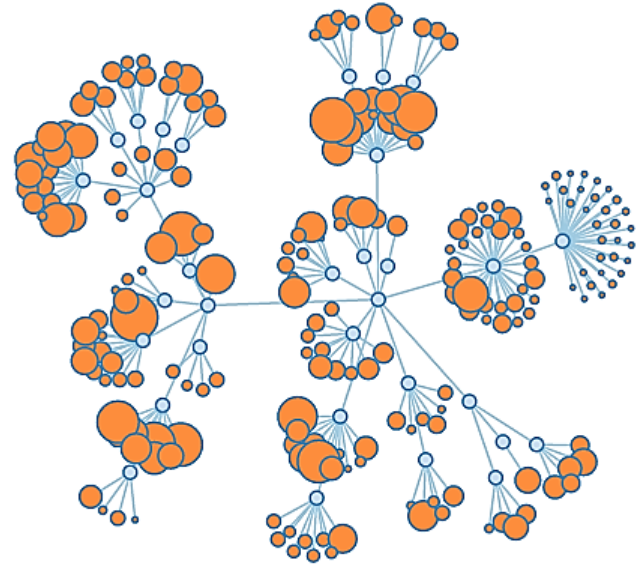
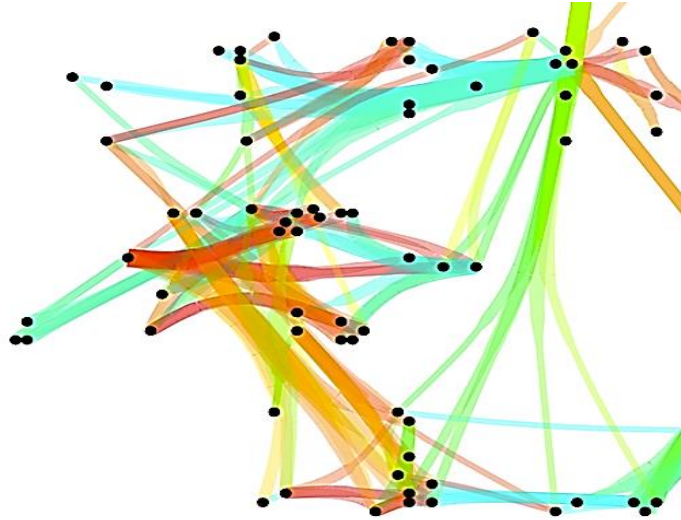
Grafos, Árboles, Redes: Estructuras contenidas

Representación de un cyber ataque (nodos atacantes por IP y dominios atacados como mapa de calor).



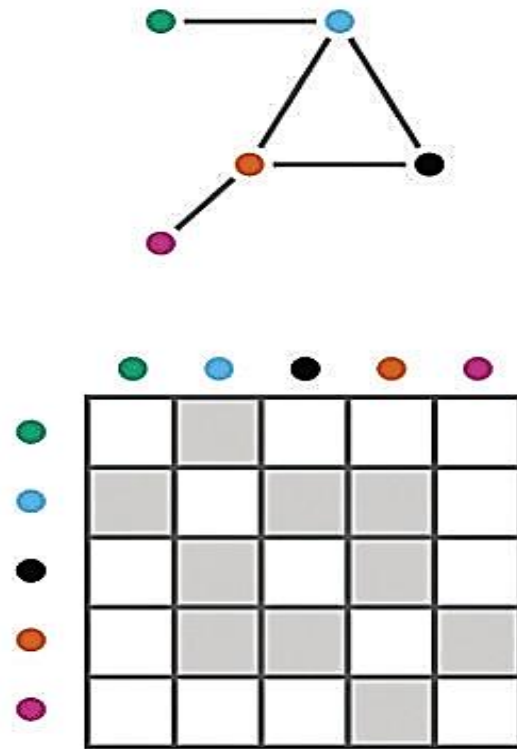
Grafos, Árboles, Redes: Nodo-Arco

Adicionalmente, otras metáforas pueden representar atributos de los arcos o de los nodos.



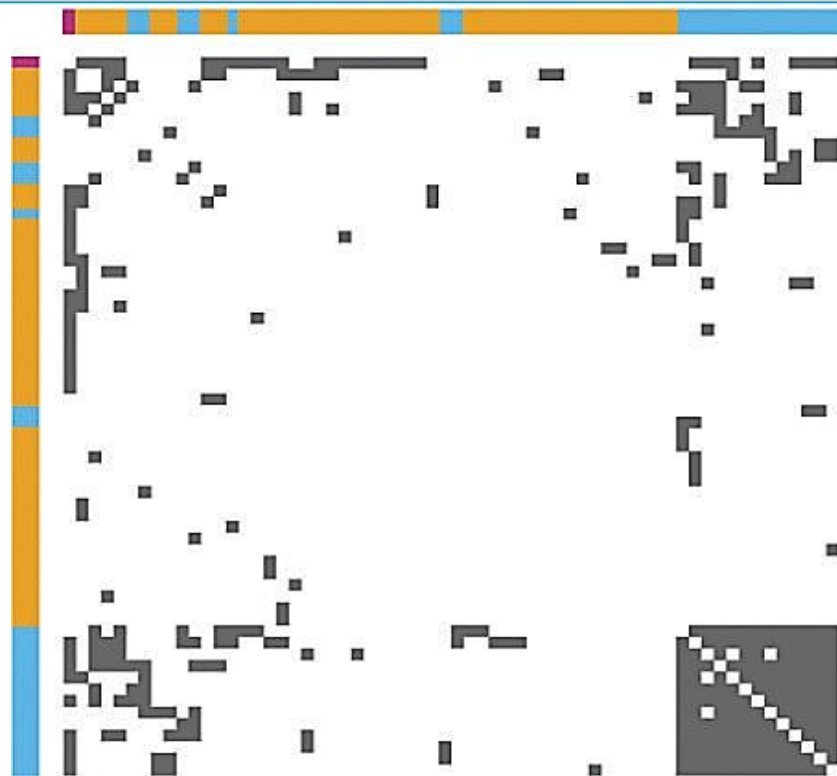
Grafos, Árboles, Redes: Matrices de Adyacencia

La información presente en un árbol o grafo puede representarse en su **matriz de adyacencia** asociada.



Grafos, Árboles, Redes: Matrices de Adyacencia

En algunos casos visualizar la matriz de adyacencia puede no ser especialmente útil.



Grafos, Árboles, Redes: Matrices de Adyacencia

En otros casos, la matriz de adyacencia hace visibles propiedades del grafo que no se detectan fácilmente en la representación nodo-arco.

