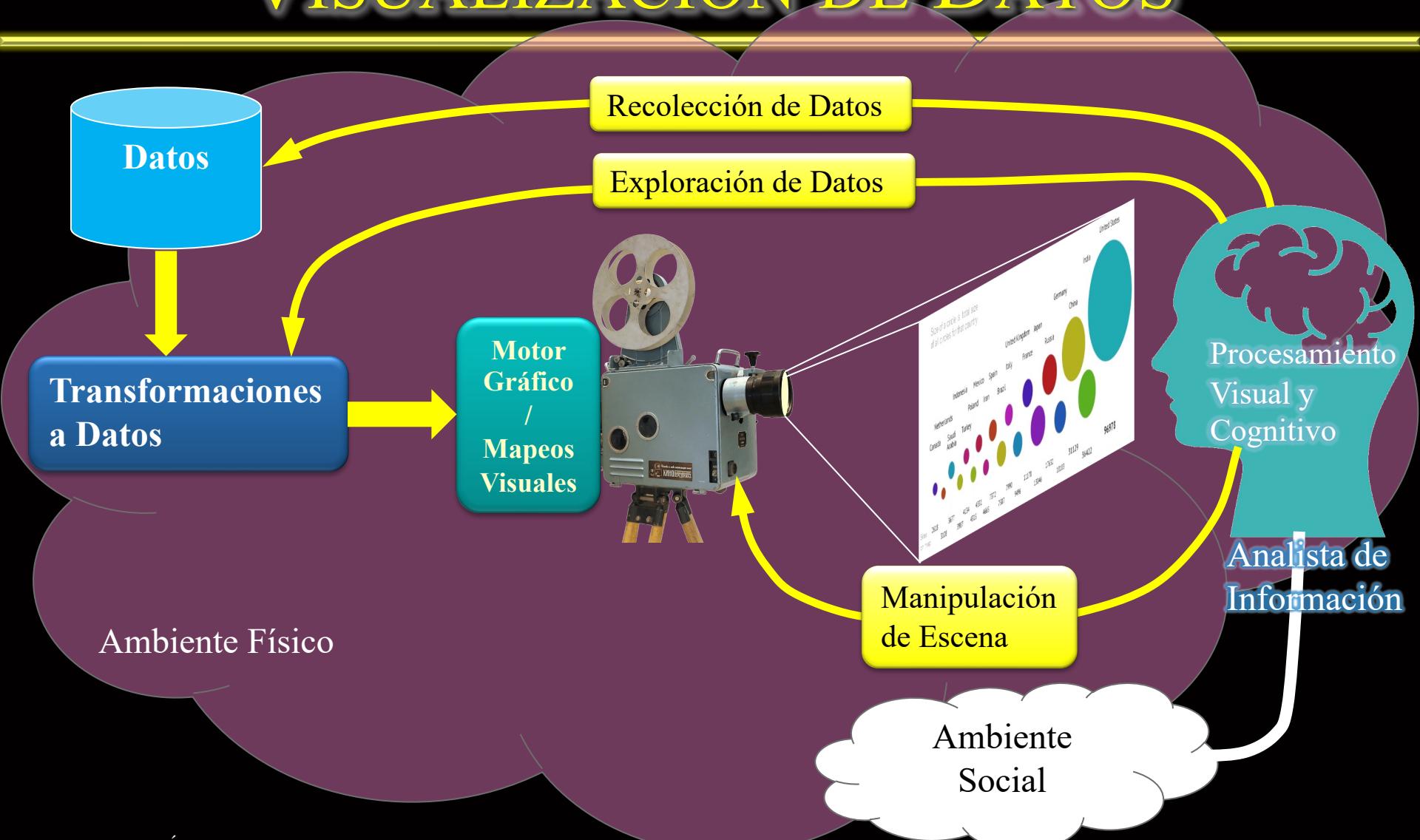
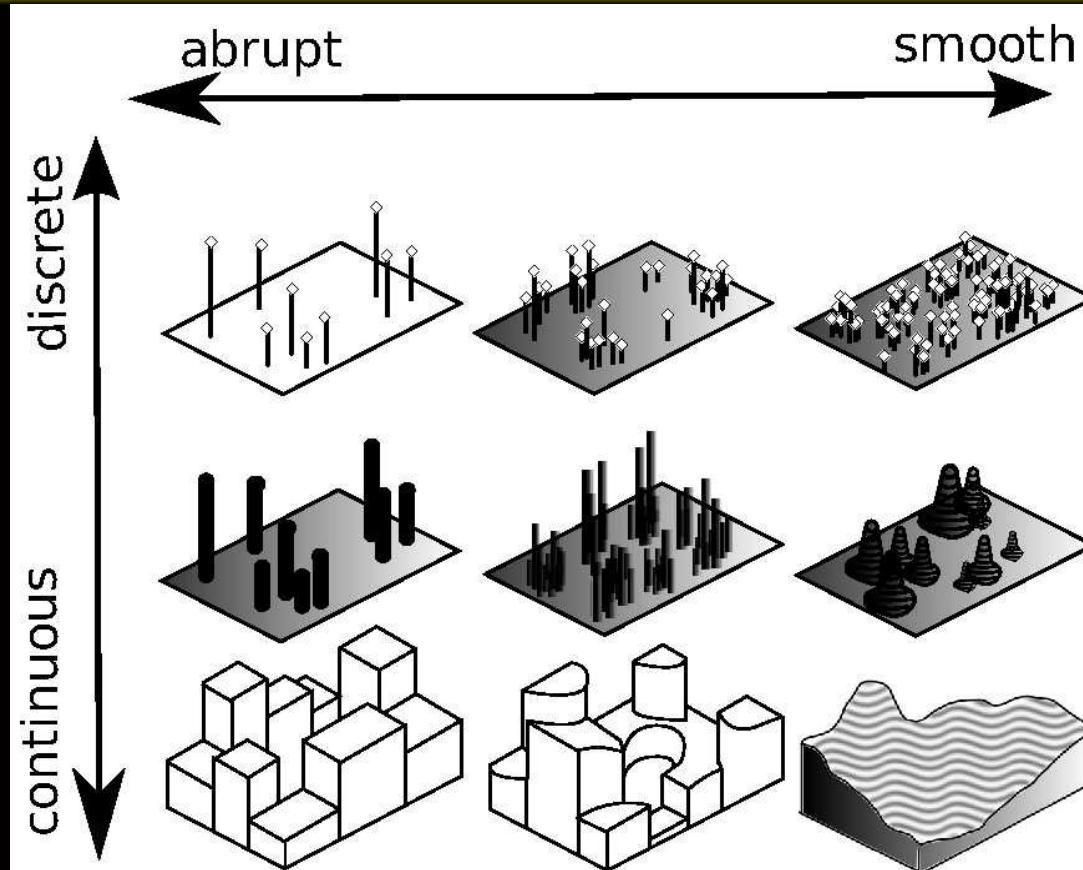


SÚPER RESUMEN DE VISUALIZACIÓN DE DATOS



Datos GeoEspaciales

Visualización de Datos Puntuales



Se presume que los datos discretos ocurren en posiciones diferentes mientras los continuos ocurren en “todas partes”. Los datos suaves se refieren a datos que cambian de forma gradual y los abruptos cambian repentinamente.

Datos GeoEspaciales

Visualización de Datos Puntuales

- **Mapas de puntos:** Los fenómenos puntuales se pueden visualizar colocando un símbolo o píxel sobre la posición en donde un fenómeno se ha registrado.
- Un parámetro cuantitativo puede ser mapeado al tamaño o al color del símbolo o píxel. Los círculos son los símbolos más usados en mapas de puntos, pero también se pueden usar cuadrados o barras.
- Si el tamaño del símbolo se usa para representar un parámetro cuantitativo, hay que saber como escalar los símbolos. Calcular el tamaño correcto de los símbolos **NO** significa que los símbolos serán percibidos correctamente. El tamaño percibido de los símbolos no necesariamente corresponde al tamaño actual debido a los problemas en la percepción del tamaño, por ejemplo, el tamaño percibido de los símbolos depende de su vecindario local y por ello **NO** existe una formula global para el escalamiento perceptual.
- Algo similar ocurre cuando se usa color para representar un parámetro cuantitativo (tomar en cuenta la percepción del color).

Datos GeoEspaciales

Visualización de Datos Puntuales



- Los mapas de puntos son una forma elegante de comunicar mucha información acerca de las relaciones de fenómenos de puntos espaciales en un formato conveniente, compacto y familiar. Sin embargo, cuando se dibujan conjuntos de datos grandes sobre un mapa, el problema de traslape o sobre graficación de los puntos de datos da lugar a áreas muy pobladas contra regiones poco pobladas (considerando que los datos espaciales reales no tienen una distribución uniforme).

Datos GeoEspaciales

Visualización de Datos Puntuales

- Es claro que es posible que sea necesario mostrar varios parámetros en varios mapas. Si todos los mapas muestran los datos de la misma manera, es posible relacionar los parámetros y detectar correlaciones, dependencias y otros patrones locales que puedan ser interesantes.
- Existen ya muchos métodos que lidian con datos espaciales densos, uno de ellos es una visualización 2.5D que muestra los datos de puntos agregados a regiones de un mapa o mapas; esta técnica se puede hallar en programas como In3D de VisualInsight y ESRI de ArcView.
- Otra forma de mostrar más detalles es la visualización de cada punto en los datos como una barra de acuerdo con su valor estadístico sobre el mapa. Esta técnica es parte de sistemas como MineSet y Swift 3D de AT&T.
- Un problema con este método es que un gran número de datos puntuales se gráfican sobre la misma posición y, por lo tanto, sólo una porción de los datos es visible. No sólo eso, debido a la oclusión en 3D, una fracción importante de los datos puede quedar oculta a menos que se cambie el punto de vista.

Datos GeoEspaciales

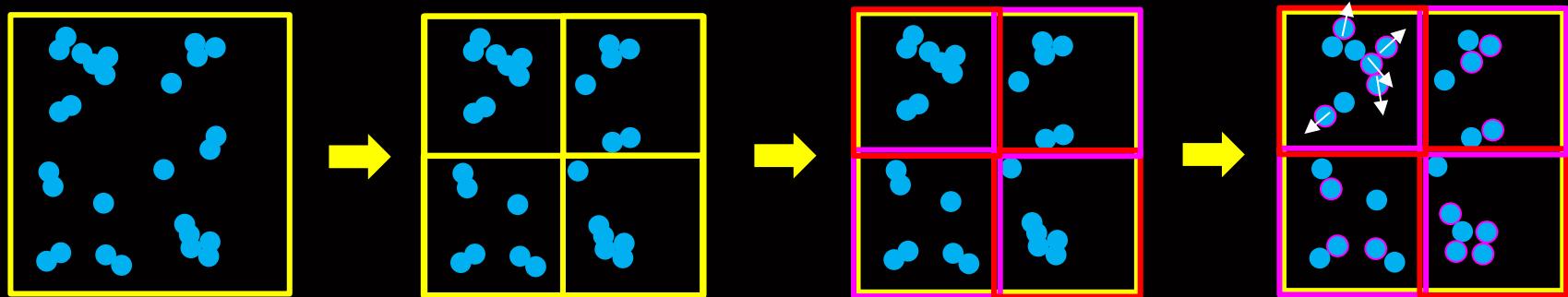
Mapas de Píxeles

- Un método que no superpone los datos en un despliegue 2D es el mapa de píxeles. La idea detrás del método es reposicionar los píxeles que potencialmente se superpondrían.
- La idea básica del algoritmo de reposicionamiento es dividir recursivamente el conjunto de datos en cuatro subconjuntos que contengan puntos de datos en subregiones del mismo tamaño. Debido a que los datos puntuales pueden salirse de las cuatro subregiones, será necesario determinar nuevos límites de las subregiones (sin cambiar los cuatro subconjuntos de datos), de tal forma que los datos en cada subconjunto se puedan visualizar en su subregión correspondiente.
- Para una implementación eficiente se utiliza una estructura similar a un *quadtree* con la que se administra la información necesaria y con la que se da apoyo al proceso de división recursivo. El proceso de división funciona como sigue.

Datos GeoEspaciales

Mapas de Píxeles

- La partición se lleva a cabo de la siguiente manera.



- Si después de algunas iteraciones quedan un número limitado de puntos en una subregión, los puntos se pueden colocar primero el punto inicial en su posición correcta y los subsecuentes que se superponen en posiciones en la pantalla que son cercanas y disponibles (posicionamiento quasi-aleatorio).
- Un problema con los mapas de píxeles es que en las áreas de gran superposición, la recolocación depende de la secuencia de los puntos en la base de datos.

Datos GeoEspaciales

Visualización de Datos de Lineales

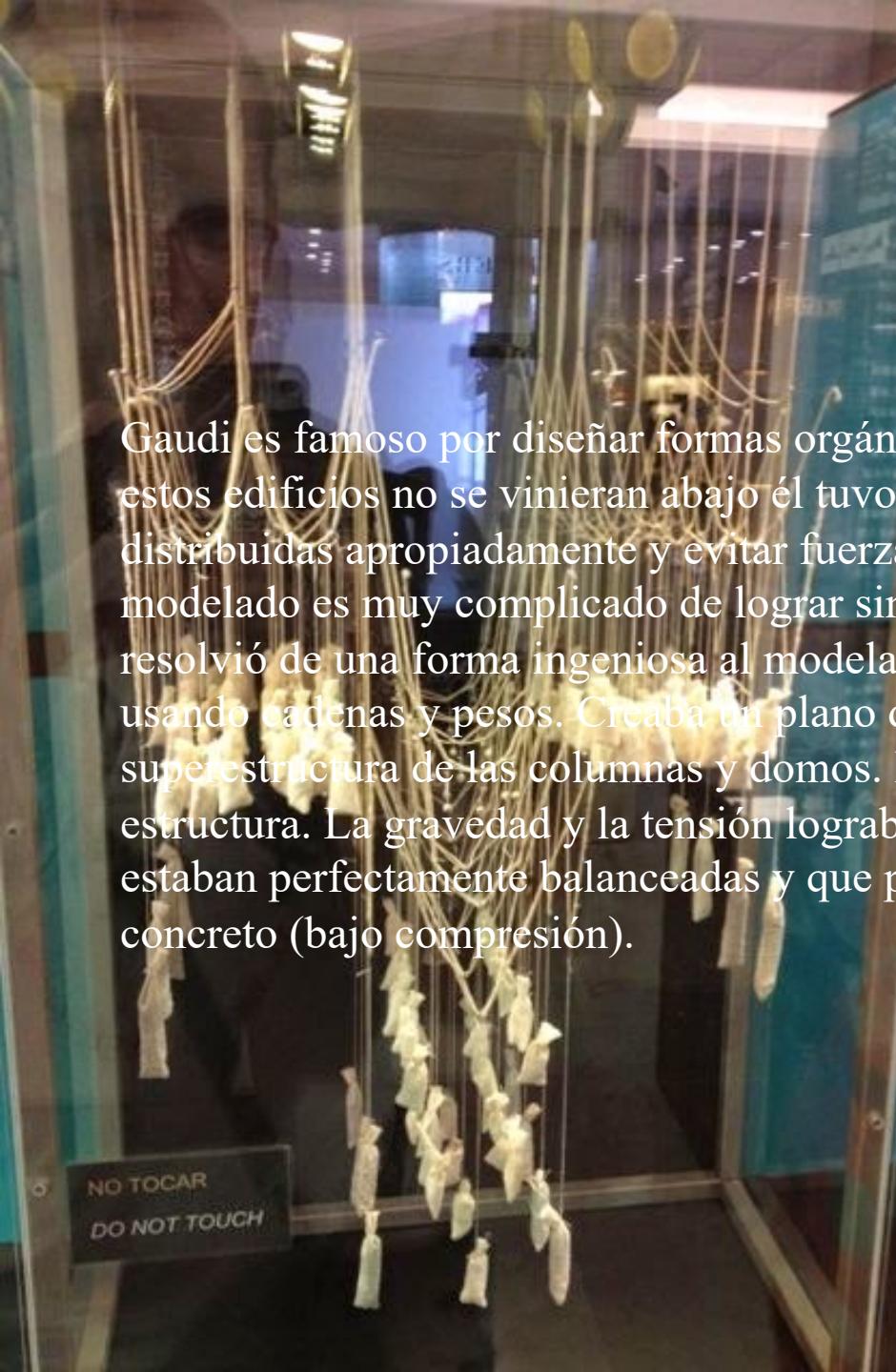
- La idea básica detrás de la visualización de datos espaciales que describen fenómenos lineales es representarlos como segmentos de líneas entre pares de puntos finales especificados por longitud y latitud.
- Un mapeo estándar de datos de líneas permite mapear los parámetros de los datos a ancho, patrón, color o etiqueta de línea. Además, las propiedades de los datos tanto para los puntos de inicio y terminación, como para la intersección de puntos pueden ser mapeados a los parámetros visuales de los nodos (tamaño, forma, color o etiqueta). Las líneas no tienen que ser rectas, pero pueden ser polilíneas (una secuencia de puntos con segmentos de líneas entre ellos) o *splines* para evitar desorden en el despliegue.
- Que despliegue es el mejor depende de la tarea que se debe cumplir.



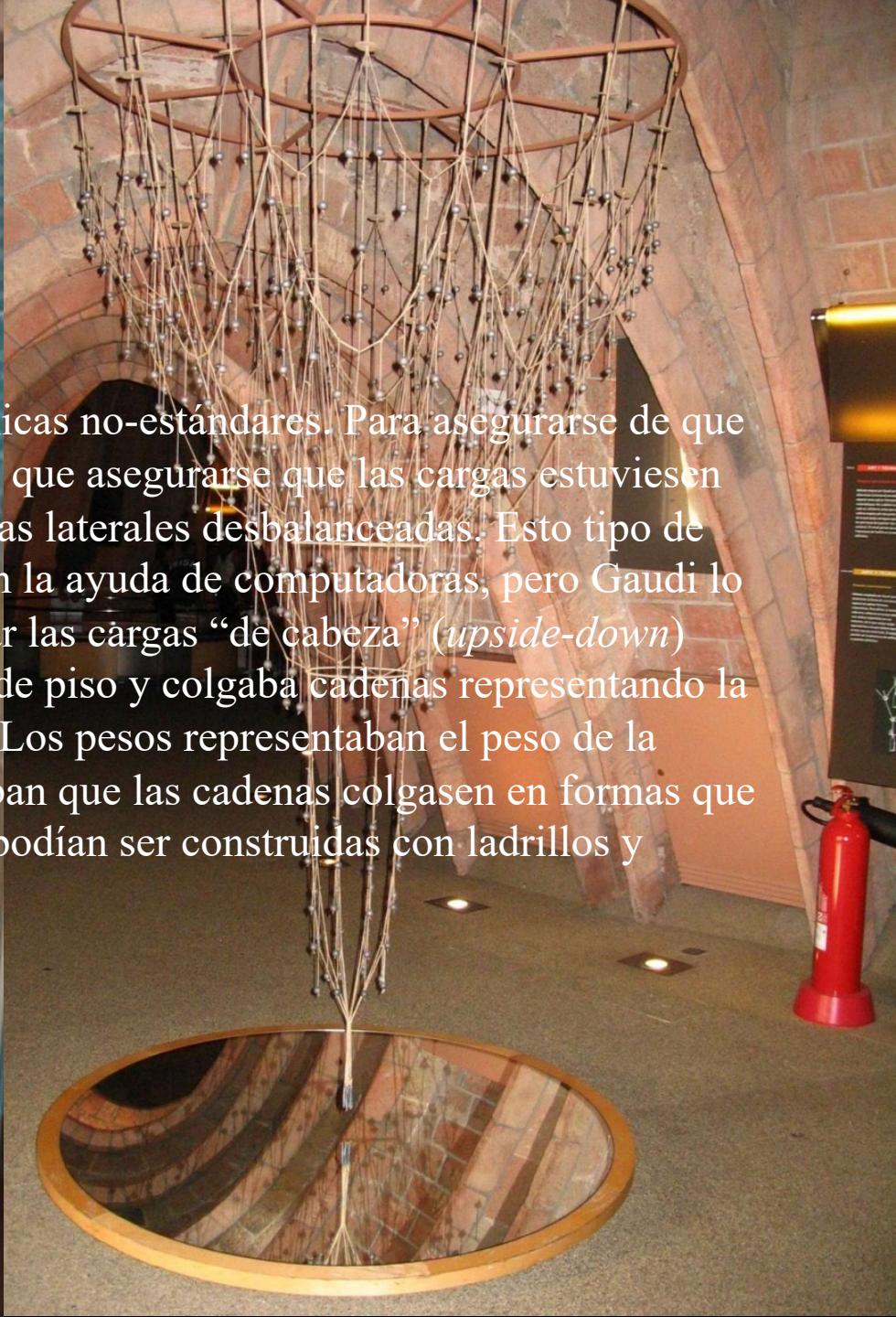








Gaudí es famoso por diseñar formas orgánicas no-estándares. Para asegurarse de que estos edificios no se vinieran abajo él tuvo que asegurarse que las cargas estuviesen distribuidas apropiadamente y evitar fuerzas laterales desbalanceadas. Esto tipo de modelado es muy complicado de lograr sin la ayuda de computadoras, pero Gaudí lo resolvió de una forma ingeniosa al modelar las cargas “de cabeza” (*upside-down*) usando cadenas y pesos. Creaba un plano de piso y colgaba cadenas representando la superestructura de las columnas y domos. Los pesos representaban el peso de la estructura. La gravedad y la tensión lograban que las cadenas colgasen en formas que estaban perfectamente balanceadas y que podían ser construidas con ladrillos y concreto (bajo compresión).







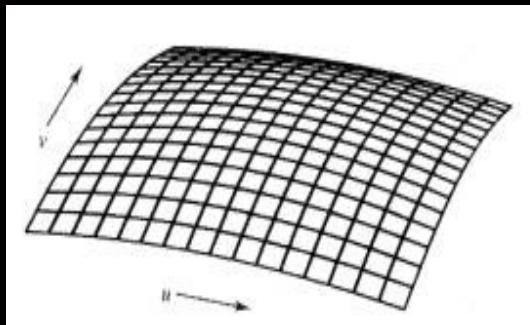




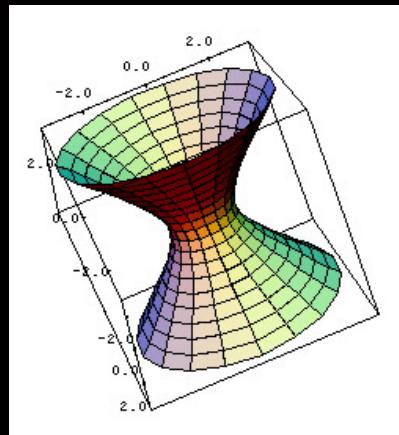
Splines

- Definición A: Una función que está definida en un intervalo, que a su vez puede ser usada para aproximar otra función, y está compuesta de funciones simples definidas en sub-intervalos y unidas en sus puntos finales con un grado de suavidad.
- Definición B: Una curva suave definida por un conjunto finito de puntos de control.

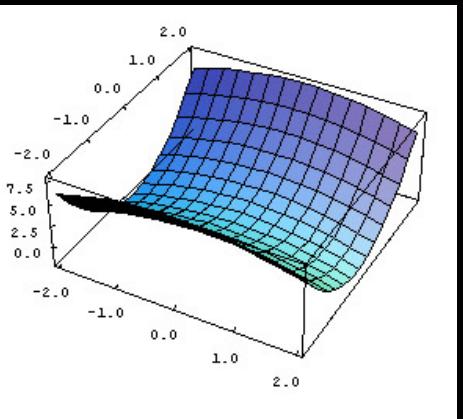
Primeras Alternativas



Parches/Superficies
Spline



Quadratics y otros
polinomios implícitos



Aproximación de Funciones

- Una de las formas más sencillas de aproximar una función desconocida es por medio de polinomios. Los polinomios son fáciles de almacenar en una computadora, son suaves y en un intervalo cerrado pueden aproximar una función tan bien como lo deseemos.
- Teorema de Aproximación de Weierstrass (Weierstrass Approximation Theorem). Si f es una función continua en el intervalo $[a,b]$ y dado $\varepsilon > 0$, entonces existe un polinomio P en $[a,b]$ de forma tal que
$$|f(x) - P(x)| < \varepsilon$$
para todo x in $[a,b]$.
- En otras palabras, cualquier función en un intervalo limitado y cerrado puede ser aproximado uniformemente en dicho intervalo por medio de polinomios con cualquier nivel de exactitud.

Aproximación de Funciones

- Sin embargo, los polinomios son bastante inflexibles. Entre más puntos se necesitan interpolar, se necesita usar un polinomio de mayor grado. Pero, los polinomios de un grado alto tienden a oscilar/menear.
- En particular, C. Runge probó qué el error máximo se puede aproximar a infinito conforme el grado del polinomio se incrementa.

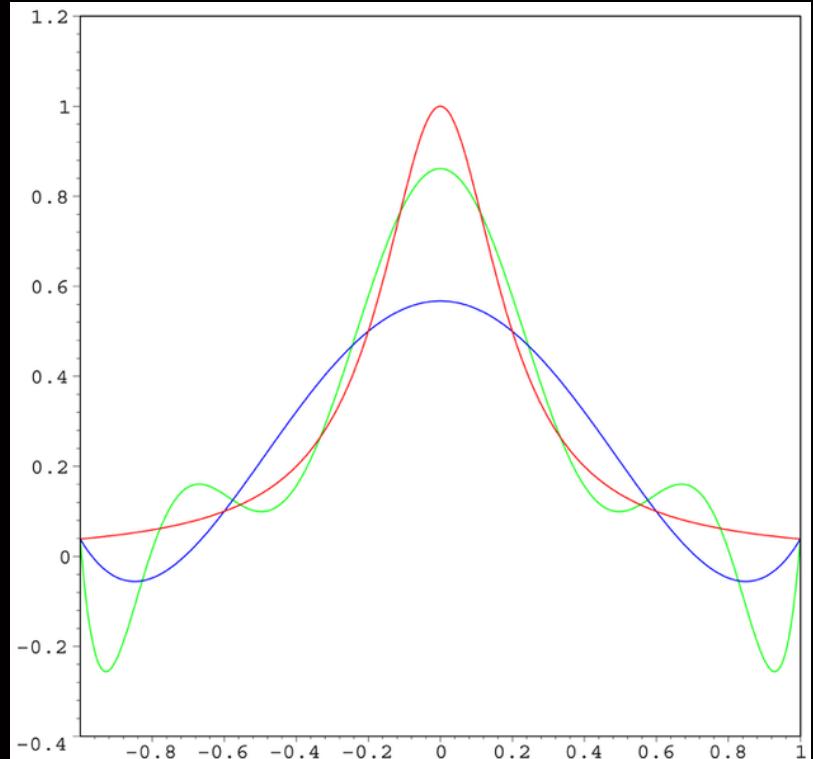
Fenómeno de C. Runge

C. Runge halló que si se interpola la función:

$$f(x) = \frac{1}{1 + 25x^2}$$

en puntos equidistantes en el intervalo $[-1, 1]$ con un polinomio que tiene un grado $\leq n$, la interpolación resultante puede oscilar cerca de los extremos del intervalo. De hecho, el error máximo tiende a infinito cuando $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} \left(\max_{-1 \leq x \leq 1} |f(x) - P_n(x)| \right) = \infty$$



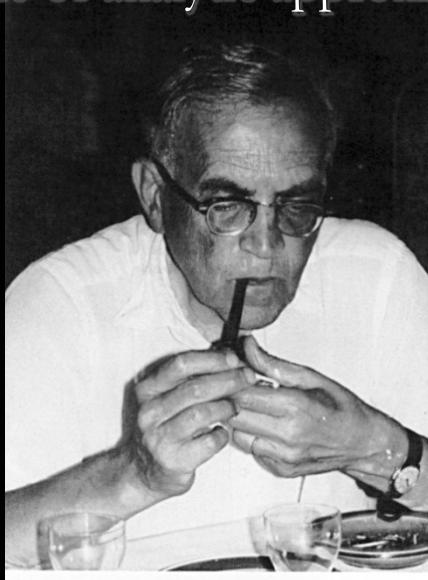
La línea roja representa $f(x)$, la línea azul es un polinomio de 5º orden y la línea verde representa un polinomio de 90º orden.

Splines

- El nombre se deriva de **El Spline de un Dibujante de Planos/Diagramas**, el cual es un cilindro flexible que se dobla para conformar a los puntos (y mantenidos en su lugar por medio de pesos).
- **Definición A:** Una función que está definida en un intervalo, que a su vez puede ser usada para aproximar otra función, y está compuesta de funciones simples definidas en sub-intervalos y unidas en sus puntos finales con un grado de suavidad.
- **Definición B:** Una curva suave definida por un conjunto finito de puntos de control.
- **Definición C:** Una función por parte polinomial que puede poseer una forma muy sencilla localmente y al mismo tiempo ser suave y flexible globalmente. Estas funciones son muy útiles para modelar funciones arbitrarias.

Splines Formales

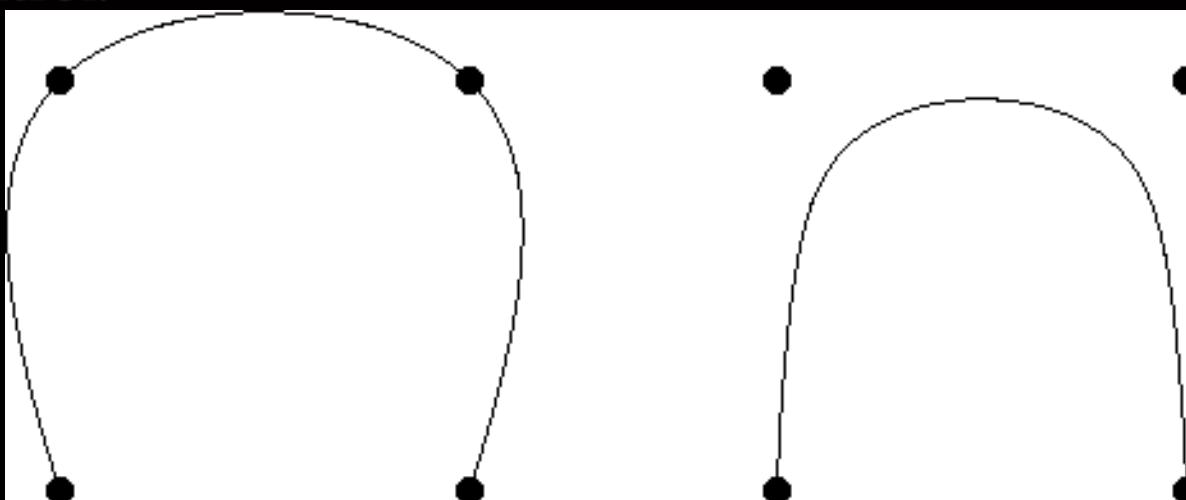
- Isaac J. Schoenberg introdujo en 1946 los *splines* en los siguientes artículos:
- I. J. Schoenberg (1946), “Contributions to the problem of approximation of equidistant data by analytic functions, Part A: On the problem of smoothing or graduation, a first class of analytic approximation formulas”, *Quart. Appl. Math.* 4, 45–99.
- I. J. Schoenberg (1946), “Contributions to the problem of approximation of equidistant data by analytic functions, Part B: On the problem of oscillatory interpolation, a second class of analytic approximation formulae”, *Quart. Appl. Math.* 4, 112–141.



I. J. Schoenberg

Categorías de *Splines*

- Básicamente existen dos tipos de *Splines*:
 - *Splines de Interpolación*: pasan a través de todos los puntos de control.
 - *Splines de Aproximación*: pasan cerca de todos los puntos de control.



Splines

- Dada una función $f(x)$ en el intervalo $[a,b]$ se subdivide el intervalo en nodos:

$$a = x_0 < x_1 < \dots < x_n = b$$

se requiere de una función $g(x)$ que interpole $f(x)$ en el intervalo $[a,b]$, y que sea igual a $f(x)$ en los puntos de control.

- En cada uno de los n subintervalos $g(x)$ está dada por un polinomio de orden bajo y diferenciable varias veces en los nodos.
- $g(x)$ es conocido como Spline.
- El spline más sencillo es cuando $g(x)$ es lineal:

$$Q(t) = (1-t)P_0 + tP_1 = (t-1)\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} P_0 \\ P_1 \end{pmatrix}$$

Power Base

$$Q = \mathbf{T}(t) \mathbf{G} \mathbf{P}$$

Spline Base

Geometría

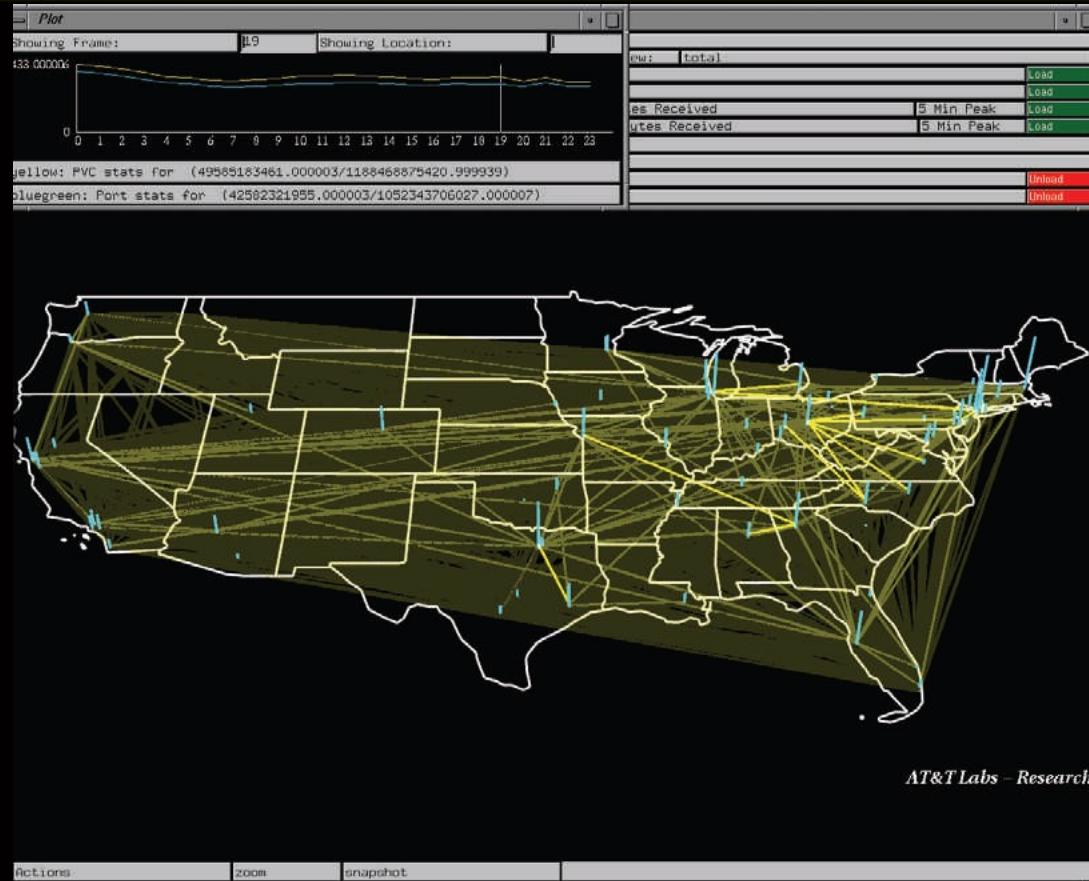
Datos GeoEspaciales

Visualización de Mapa de Redes

- Los mapas de redes son muy usados en varias aplicaciones. Algunos métodos únicamente despliegan la conectividad de las redes para transmitir información sobre su comportamiento y estructura.
- Hay autores que han usado agregación, información jerárquica, posición de nodos y enlazado de despliegues para investigar redes grandes con jerarquías y sin una estructura natural. Para ello usaron forma y color para codificar la información de los nodos y color y ancho de las líneas para codificar la información de los enlaces.
- Investigadores de la NCSA (National Center for Supercomputing Applications) agregaron gráficas 3D a sus mapas de redes para desplegar animaciones del tráfico de paquetes de Internet en su *backbone network*.
- El sistema SeeNet está motivado por investigación en gráficas estadísticas dinámicas. En este método se permite interacción con el usuario para que controle el despliegue para enfocarse en patrones interesantes. Se usan dos despliegues de redes estáticos para visualizar las relaciones geográficas y una matriz de enlace que otorga énfasis igualitario a todos los enlaces de la red.
- El sistema SWIFT-3D de AT&T integra varias técnicas de visualización que incluyen despliegues estadísticos a información general de píxeles y las combina con mapas 3D interactivos y herramientas de búsqueda del tipo *drag+drop*. El componente de visualización mapea los datos a un conjunto de vistas 2D y 3D creadas con diferentes técnicas de visualización: visualizaciones 2D estadísticas, visualizaciones 2D basadas en píxeles y visualizaciones 3D dinámicas.

Datos GeoEspaciales

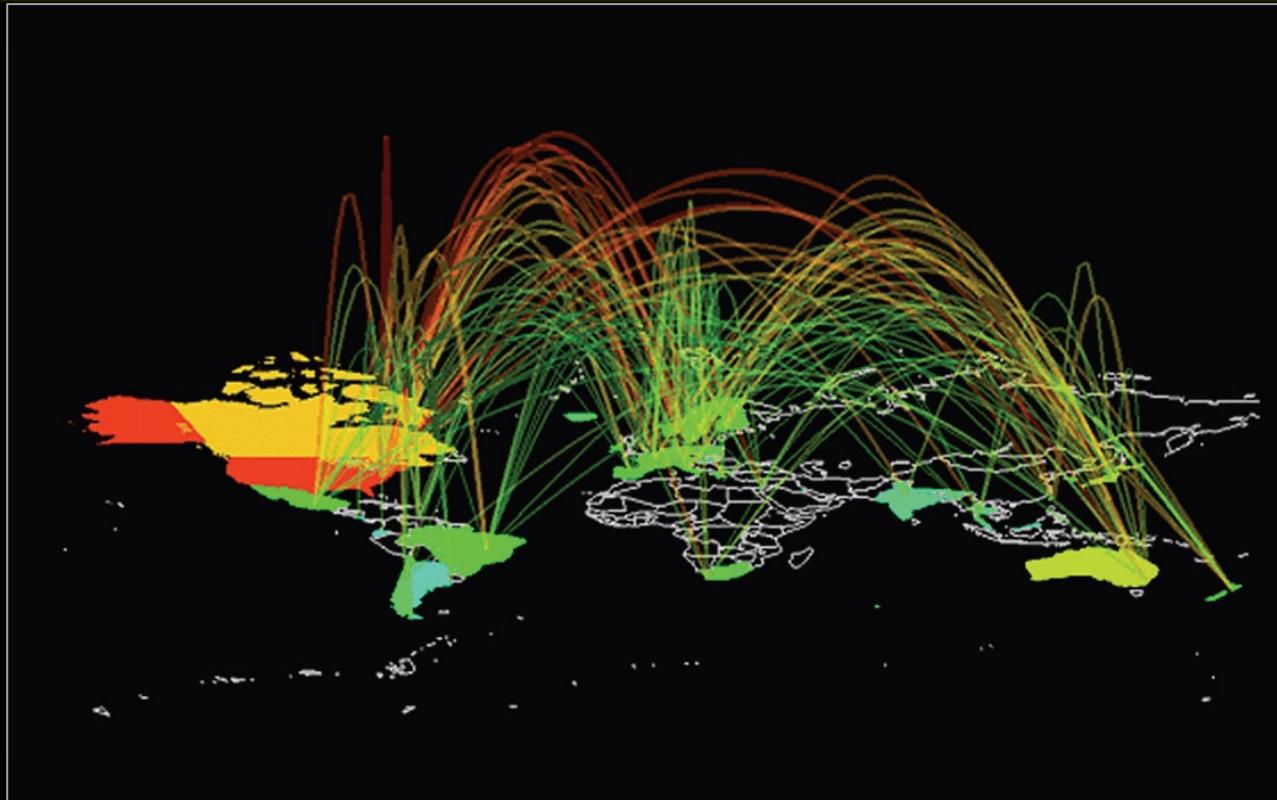
Visualización de Mapa de Redes



- Un problema importante en mapas de redes es la superposición de segmentos de líneas en áreas densas.

Datos GeoEspaciales

Mapas de Flujo y Agrupación de Aristas



Existen un número de métodos que intentan evitar el problema superposición de los mapas de redes tradicionales al utilizar líneas curvas en lugar de rectas. Cada vez más se utilizan algoritmos para lograr este propósito.

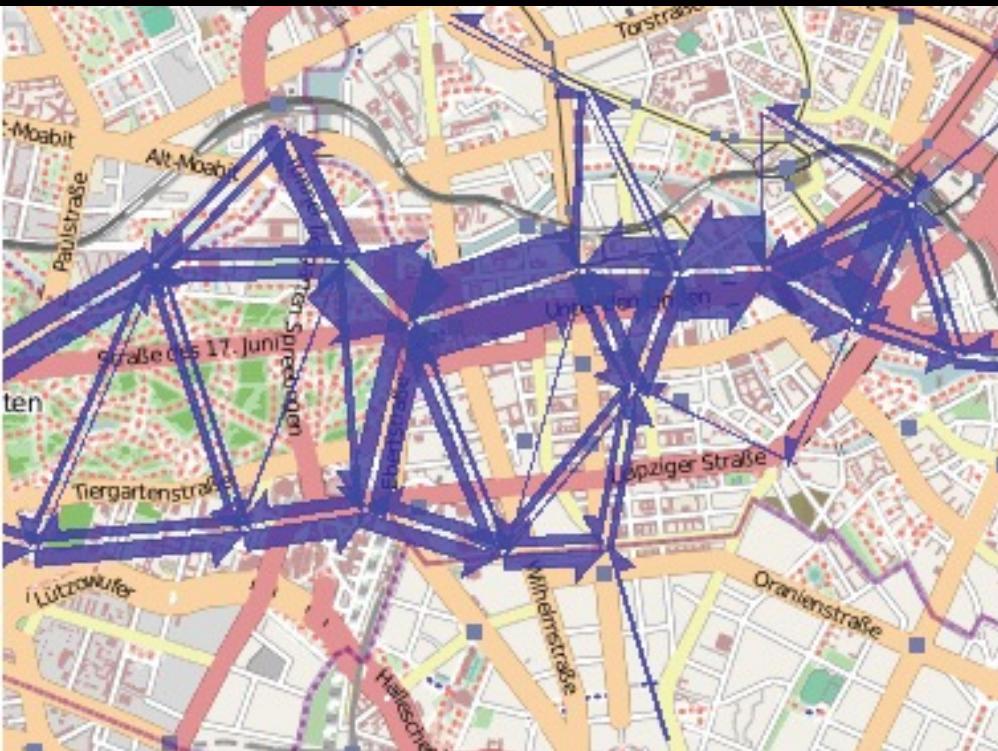
Datos GeoEspaciales

Mapas de Flujo y Agrupación de Aristas

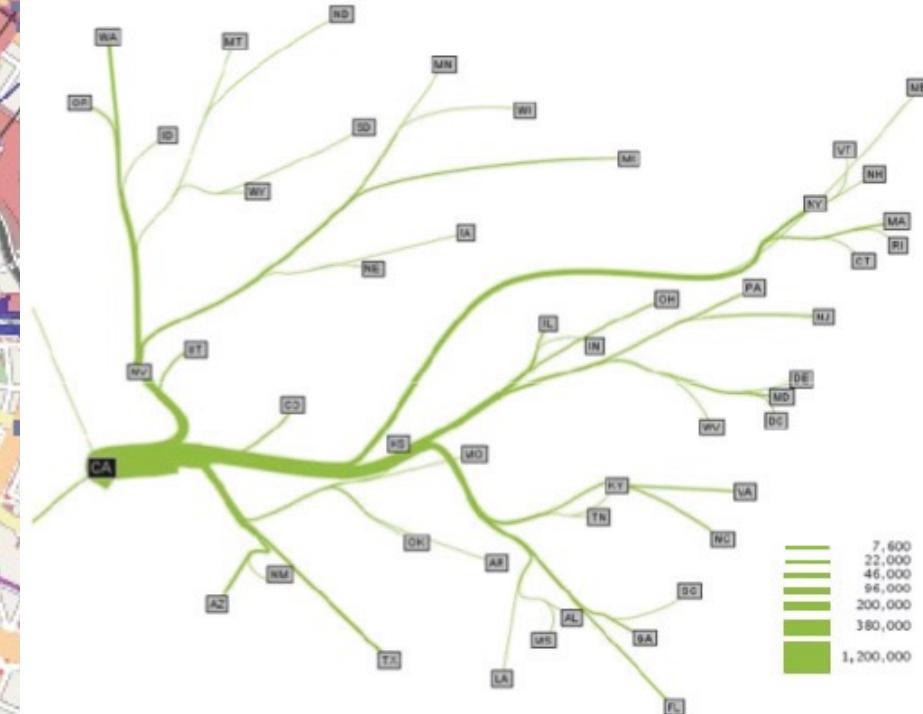
- Uno de los métodos es el propuesto en el sistema de mapas de flujo (Stanford) que se inspira en algoritmos de organización de gráficas que minimizan los cruces de aristas y distorsión de las posiciones de nodos (manteniendo sus posiciones relativas). Además, se utiliza un algoritmo de agrupamiento jerárquico basado en las posiciones de nodos y flujos entre ellos para calcular un re-ruteo y fusión de flujos.

Datos GeoEspaciales

Mapas de Flujo y Agrupación de Aristas



Flujo de Turistas en Berlín



Migración desde California

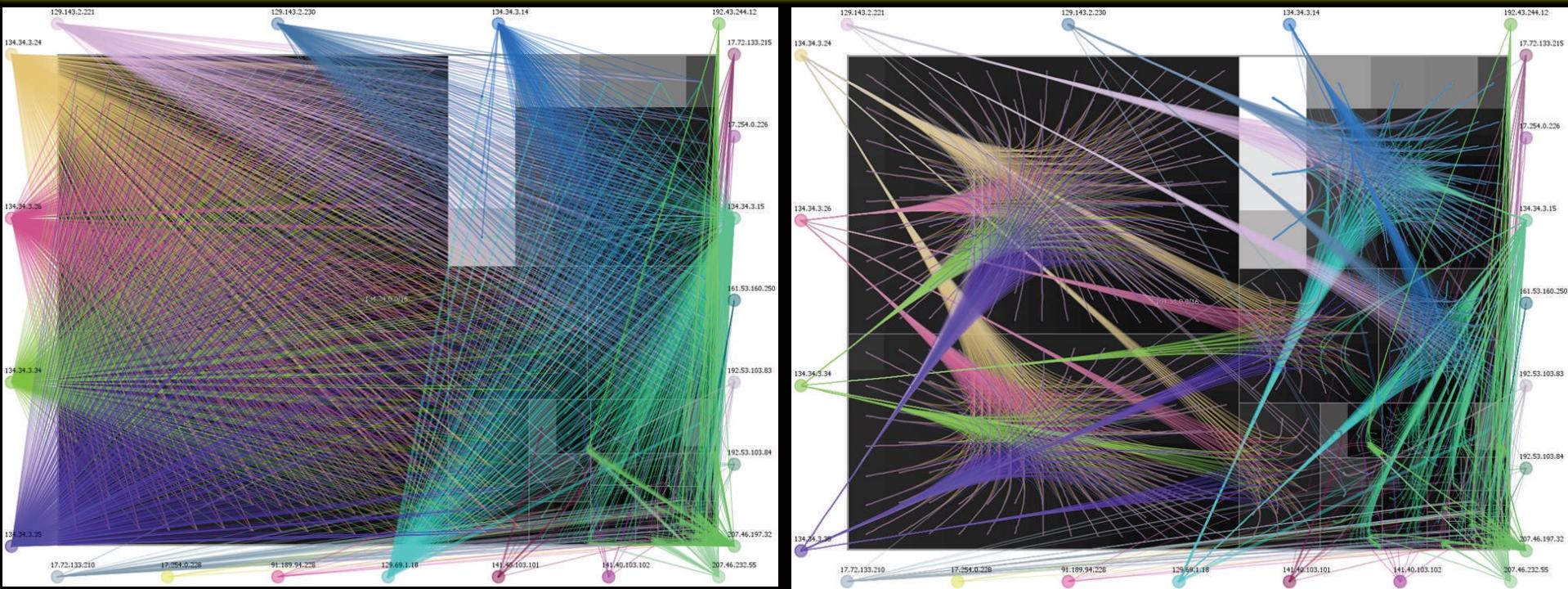
Datos GeoEspaciales

Mapas de Flujo y Agrupación de Aristas

- El **agrupamiento de aristas** también es un intento para reducir el desorden del trazado de líneas. Si se define una jerarquía en los nodos, las aristas pueden ser agrupadas de acuerdo a la jerarquía por medio de la utilización de la jerarquía como puntos de control para conectar puntos por medio de B-splines.
- Los nodos conectados pasando por la raíz de la jerarquía se doblan al máximo, mientras que aquellos dentro de la misma subjerarquía se doblan lo mínimo.
- La utilización de agrupamiento de aristas reduce el desorden visual tanto en mapas como visualización de árboles, grafos (gráficas) y otros métodos.

Datos GeoEspaciales

Mapas de Flujo y Agrupación de Aristas



Visualización del tráfico IP desde nodos externos a nodos internos, visualizados como mapas de árboles en el interior.

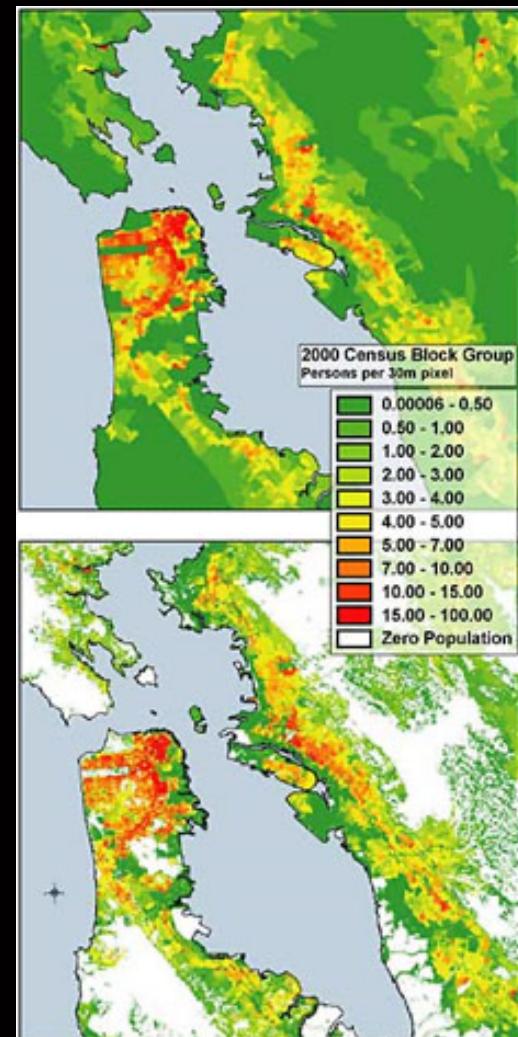
Datos GeoEspaciales

Mapas de Datos de Área

- Los mapas temáticos son el método principal para visualizar fenómenos de área y existen diferentes variantes de los mapas temáticos. El tipo de mapa temático más popular es el de coropletas (Griego choro=área y pleta=valor) en el cual los valores de un atributo o variable estadística se codifican como regiones coloreadas o sombreadas sobre el mapa.
- Se asume que para los mapas coropleticos existe un distribución uniforme en las regiones del atributo mapeado. Si el atributo tiene otras distribuciones diferentes a la partición en regiones, se pueden usar otras técnicas como los mapas dasimétricos.
- En un mapa dasimétrico la variable a ser visualizada forma áreas independientes de las regiones originales (p. ej., los bordes de las áreas se derivan de los atributos no tienen que coincidir con las de los mapas).
- Otro tipo de mapa es el mapa isarítmico o de contorno, el cual muestra los contornos de fenómenos continuos. Si los contornos se determinan de datos puntuales reales (tales como temperaturas medidas) los mapas se les llama mapas isométricos. Si los datos son medidos para cierta región (un país) y el centroide es el dato puntual, entonces se les conoce como mapas isopleticos. Una de las tareas principales para generar mapas isarítmicos es la interpolación de los datos puntuales para obtener contornos suaves.
- Los cartogramas son otra forma de mapear datos de área. En esta técnica las regiones son escaladas para reflejar una variable estadística, lo cual distorsiona la geometría de un mapa. Existen diferentes formas de cartogramas que van desde los continuos que retienen la topología de la malla poligonal a discontinuos que escalan cada polígono independientemente a áreas rectangulares o circulares.
- También se puede visualizar la información de área desplegando puntos discretos o símbolos sobre el mapa (usando símbolos que son dimensionados proporcionalmente al parámetro estadístico).

Datos GeoEspaciales

Mapas de Datos de Área

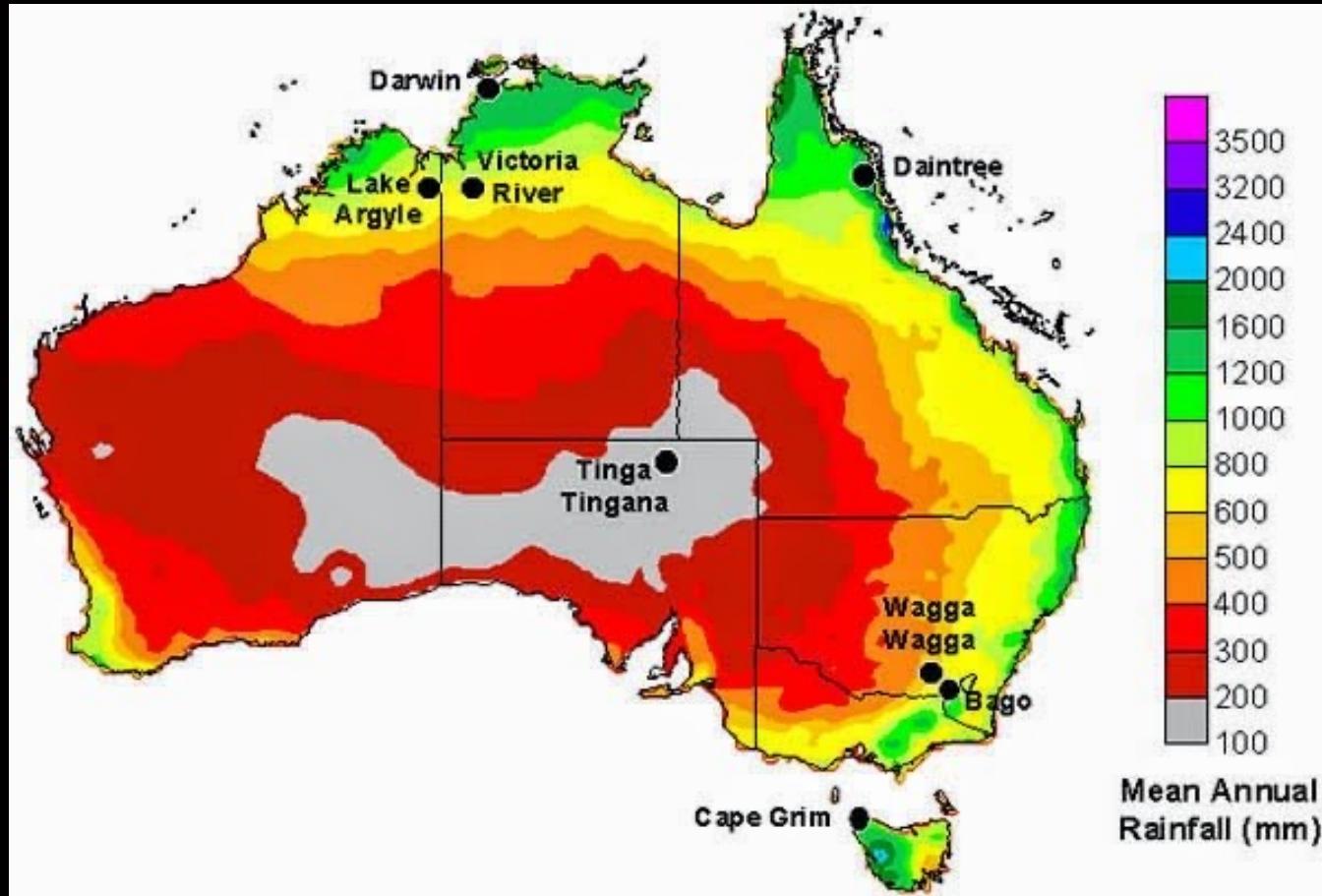


Mapa coropletico de la población de San Francisco, CA, EEUU

Mapa dasimétrico de la población de San Francisco, CA, EEUU

Datos GeoEspaciales

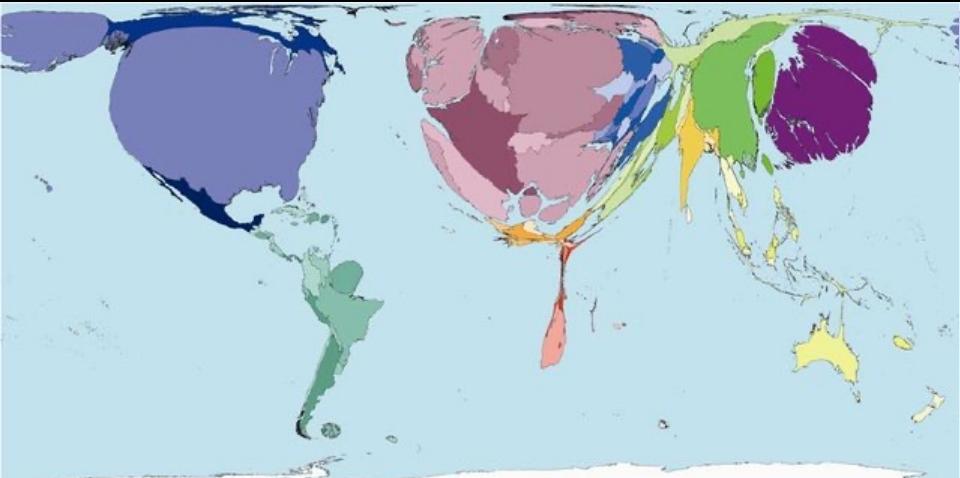
Mapas de Datos de Área



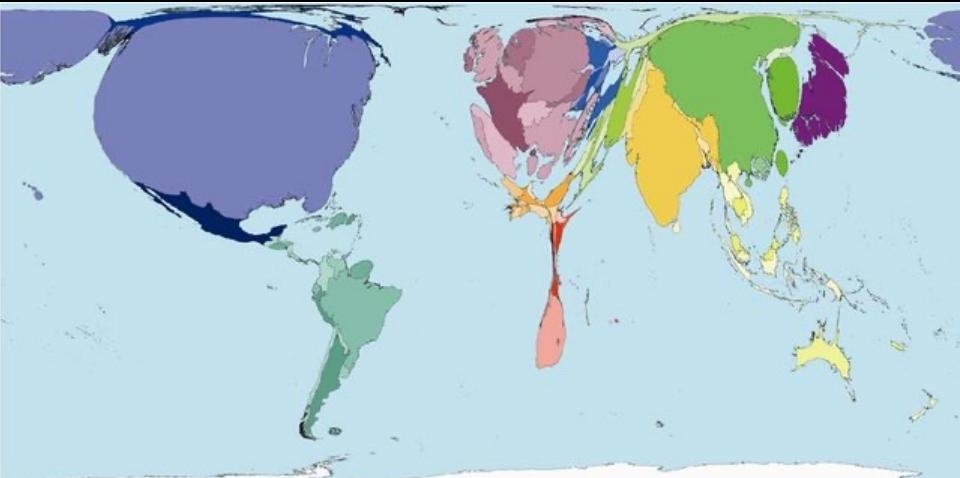
Mapa isarítmico de la precipitación media en Australia

Datos GeoEspaciales

Mapas de Datos de Área



Cartograma del gasto público en salud.



Cartograma del gasto privado en salud.