

Datos con Relación al Tiempo

Características de los Datos con Relación al Tiempo

Al igual que el tiempo, los datos tienen un impacto importante en el diseño de una visualización. Los criterios clave para los datos que están relacionados al tiempo son:

1. **Escala:** cuantitativa vs. cualitativa. Los datos cuantitativos se basan en una escala métrica (discreta o continua) mientras que los datos cualitativos describen conjuntos de elementos de datos ordenados (ordinales) o desordenados (nominales).
2. **Marco de referencia:** abstracto vs. espacial. Los datos abstractos (una cuenta de banco) han sido recolectados en un contexto no espacial y no están conectados *per se* a alguna conformación espacial inherente. Los datos espaciales (datos de un censo) contienen una conformación espacial inherente (posiciones geográficas).
3. **Tipo de datos:** eventos vs. estados. Los eventos se pueden considerar como marcadores de cambios de estado (p.ej., la partida de un autobús) mientras que los estados caracterizan las fases de continuidad entre eventos (p.ej., el tren está en el camino).
4. **Número de variables:** univariada vs. multivariados. Los datos univariados únicamente contienen un valor de datos por primitiva temporal mientras que los datos multivariados cada primitiva temporal contiene múltiples valores de datos. Estas categorías primarias forman una base para encontrar una técnica de visualización efectiva, expresiva y apropiada para los datos con relación al tiempo.

Datos con Relación al Tiempo

Relacionando Datos con el Tiempo

Cualquier conjunto de datos se relaciona a cualquiera de dos dominios temporales: tiempo interno o tiempo externo. El tiempo interno se considera como la dimensión temporal inherente en el modelo de datos; este tipo de tiempo describe cuando la información contenida en los datos es válida. Por otro lado, el tiempo externo se considera extrínseco al modelo de los datos; el tiempo externo es necesario para describir como evoluciona el conjunto de los datos en el tiempo (externo). Los conjuntos de datos relacionados al tiempo se pueden clasificar con base en el número de primitivas de tiempo en el tiempo externo y externo:

1. **Datos no-temporales estáticos.** Si el tiempo interno y externo están compuestos de únicamente un elemento temporal, los datos son completamente independientes al tiempo.
2. **Datos temporales estáticos.** Si el tiempo interno contiene más de una primitiva de tiempo, mientras que el tiempo externo contiene únicamente una primitiva de tiempo, entonces los datos se pueden considerar dependientes al tiempo. Debido a que los valores almacenados en los datos dependen del tiempo interno, los datos temporales estáticos se pueden interpretar como una visión histórica de cómo el mundo real o algún modelo observa a los varios elementos de tiempo interno. Las series de tiempo son un ejemplo prominente de los datos temporales estáticos. La mayoría de los métodos actuales de visualización que explícitamente consideran el tiempo como una dimensión de datos especial tratan datos temporales estáticos.

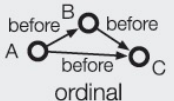






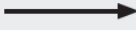


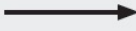







Datos con Relación al Tiempo




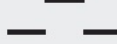


Relacionando Datos con el Tiempo

3. **Datos no-temporales dinámicos.** Si el tiempo interno contiene solo una primitiva de tiempo, pero el tiempo externo está compuesto de múltiples primitivas de tiempo, entonces los datos dependen del tiempo externo. En otras palabras, los datos cambian con el tiempo (son dinámicos). Ya que el tiempo interno no se considera, únicamente el estado actual de los datos se preserva (una vista histórica no se mantiene). Existen menos técnicas de visualización que explícitamente se enfocan en datos no-temporales dinámicos. Sin embargo, debido a que el tiempo interno y el externo pueden ser mapeados de uno al otro, algunas de las técnicas de visualización para datos temporales estáticos se pueden aplicar también a los datos no-temporales dinámicos.
4. **Datos temporales dinámicos.** Si ambos tiempos, interno y externo, están compuestos de múltiples primitivas de tiempo, entonces los datos son considerados dependientes de dos tiempos (bi-temporales). En otras palabras, los datos contienen variables que dependen del tiempo (interno) y el estado actual cambia con respecto al tiempo (externo). En este caso, usualmente, el tiempo interno y externo están acoplados fuertemente y pueden ser mapeados de uno de ellos al otro. No se hace una distinción explícita entre tiempo interno y externo por los métodos actuales de visualización porque considerar ambas dimensiones temporales es muy complicado.





Datos con Relación al Tiempo

Relacionando Datos con el Tiempo

scale			
scope			
arrangement			
viewpoint			
Abstractions			
granularity & calendars			
time primitives			
determinacy			

scale	3.14 3.27 4.88 quantitative	coconut banana apple qualitative
frame of reference	 abstract	 spatial
kind of data	 events	 states
number of variables	 univariate	 multivariate

Datos

internal time inherent in the data model	 non-temporal	 temporal
external time extrinsic to the data model	 static	 dynamic

Datos y Tiempo

C. Computación, I.I.M.A.S.

Tiempo

Datos con Relación al Tiempo

Visualización de Datos con Relación al Tiempo

Para comunicar/mostrar la dependencia de los datos con respecto al tiempo es necesario considerar bien como colocar el eje del tiempo. Existe una gran variedad de técnicas visuales para cubrir los diferentes métodos.

1. **Mapeo de Tiempo.** Existen dos opciones para mapear el tiempo: el mapeo de tiempo al espacio y mapeo del tiempo al tiempo. Cuando se habla de un mapeo del tiempo al espacio se quiere decir que los datos y el tiempo se representan coherentemente en una sola representación visual. Esta representación no cambia automáticamente con el paso del tiempo y es por ello que se les conoce como visualización estáticas de datos relacionados al tiempo. Por otra parte, las representaciones dinámicas utilizan la dimensión física del tiempo para transmitir la dependencia de los datos con respecto al tiempo; en otras palabras, el tiempo es mapeado a la dimensión física del tiempo. Esto resulta en visualizaciones que cambian automáticamente con el paso del tiempo (p.ej., animaciones o presentaciones de transparencias). Hay que hacer notar que la presencia, o ausencia, de interacción puede facilitar la navegación en el tiempo pero no tiene influencia en como se categoriza una visualización (en estática o dinámica).
2. **Dimensionalidad del espacio de presentación.** Se puede diferenciar entre presentaciones 2D o 3D de los datos relacionados al tiempo. Los métodos que usan el espacio de presentación 2D se tienen que asegurar que el eje del tiempo es enfatizado porque las dimensiones del tiempo y de los datos frecuentemente tienen que compartir las dos dimensiones del despliegue. En el caso de representaciones 3D, la dimensión extra se puede utilizar para el eje del tiempo para lograr una separación clara del tiempo de las otras dimensiones de los datos.

Datos con Relación al Tiempo

Categorización

Para facilitar una categorización clara y para mantener simple la categorización de las técnicas de visualización, uno se puede enfocar en los aspectos centrales de los datos, el tiempo y la visualización. Por ejemplo:

Datos:

- Marcos de referencia abstractos vs. espaciales.
- Variables univariados vs. multivariados.

Tiempo:

- Distribución lineal vs. cíclica.
- Primitivas del tiempo instantáneas vs. en intervalos.

Visualización:

- Mapeos estáticos vs. dinámicos.
- Dimensionalidad 2D vs. 3D.

Para estructurar las varias técnicas de visualización para los datos con relación al tiempo, se necesitan contestar:

¿Qué es lo que se presenta?, tiempo y datos.

¿Por qué se presenta?, tareas de usuario.

¿Cómo se presenta?, representación visual.

Hay tareas elementales que ayudan a determinar como se visualizan los datos, por ejemplo, para una consulta directa se puede interrogar “¿cuál es el valor de la glucosa el día 1º de marzo, 2021?”, o para una comparación directa se puede interrogar “compara los valores de glucosa y el nivel de actividad el 1º de marzo, 2021”, o para tareas sinópticas se puede preguntar “encuentra dos meses contiguos con tendencias opuestas en los valores de glucosa”.

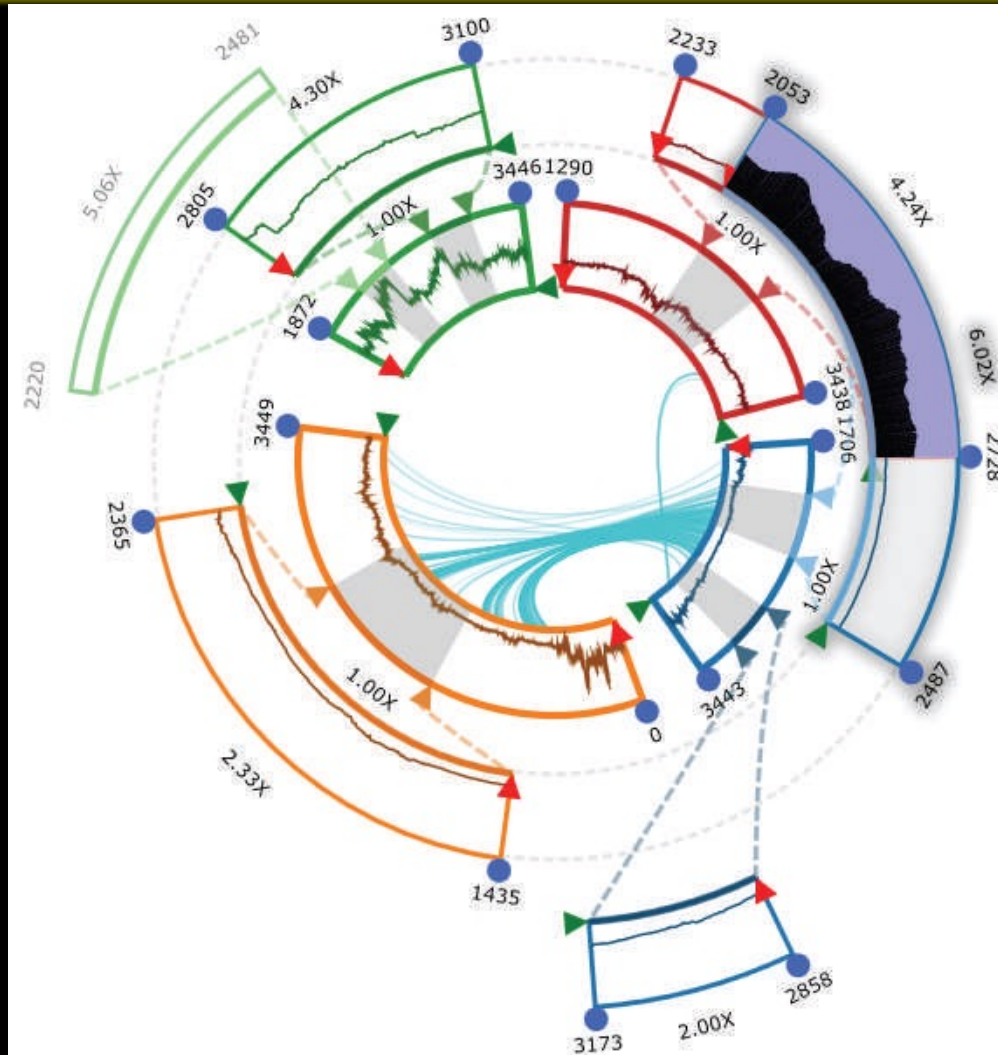
Datos con Relación al Tiempo

Datos: Marco de Referencia - Abstracto

KronoMiner es una herramienta multi-propósito para explorar las series de tiempo que provee altas capacidades de navegación y análisis. La forma de realizar visualización se basa en una distribución radial jerárquica lo cual permite a los usuarios adentrarse a los detalles de la información al enfocarse en piezas diferentes de la visualización. Porciones o piezas de los datos se pueden rotar, arrastrar, expandir o reducir de una manera fácil, con el apoyo de varias tareas de análisis y de exploración de series de tiempo. Este programa también utiliza dos técnicas analíticas: a) MagicAnalytics Lens que muestra las correlaciones entre dos partes de los datos cuando se traslapan en la visualización y b) Best Match mode que muestra un arco para indicar que dos partes de los datos coinciden bajo una medida de similitud específica.

Datos con Relación al Tiempo

Datos: Marco de Referencia - Abstracto



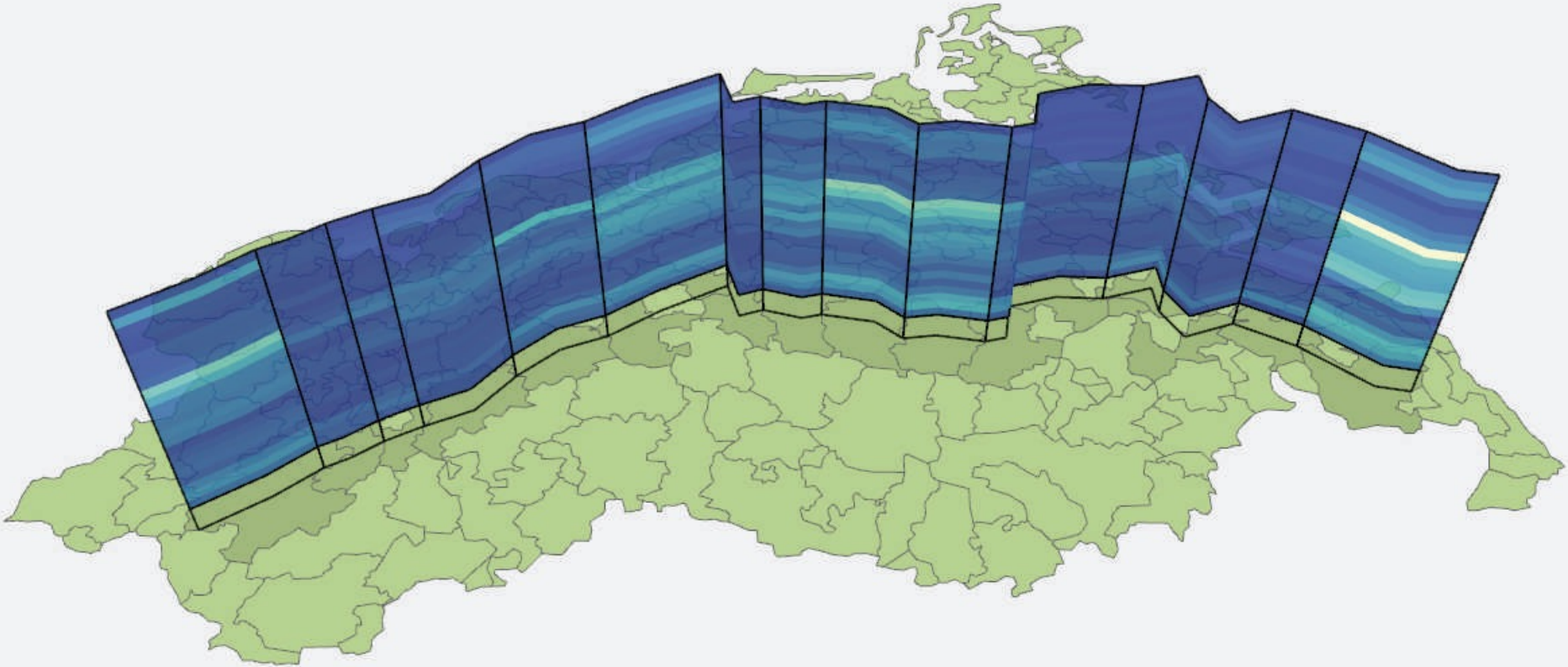
Datos con Relación al Tiempo

Datos: Marco de Referencia - Espacial

Tominski y Schulz han introducido una técnica de visualización para datos espacio-temporales que los refiere al espacio geográfico 2D y al tiempo lineal 1D. La idea detrás de esta técnica consiste en construir una rebanada no-plana (La Gran Muralla del Espacio-Tiempo) a través del espacio 3D, 2D + 1D, continuo. La construcción de la muralla se basa en los aspectos topológicos y geométricos del espacio geográfico. La trayectoria topológica se establece automáticamente o interactivamente con base en una gráfica de vecinos. El camino topológico se transforma a una trayectoria geométrica que respeta las propiedades geográficas del mapa. La trayectoria geométrica se extrude a una pared 3D, en donde la 3ª dimensión se puede utilizar para mapear el dominio del tiempo. Se pueden proyectar diferentes representaciones visuales sobre la muralla para poder mostrar los datos. La muralla tiene la ventaja de mostrar una trayectoria cerrada a través del espacio sin espacios vacíos entre los píxeles que contienen información.

Datos con Relación al Tiempo

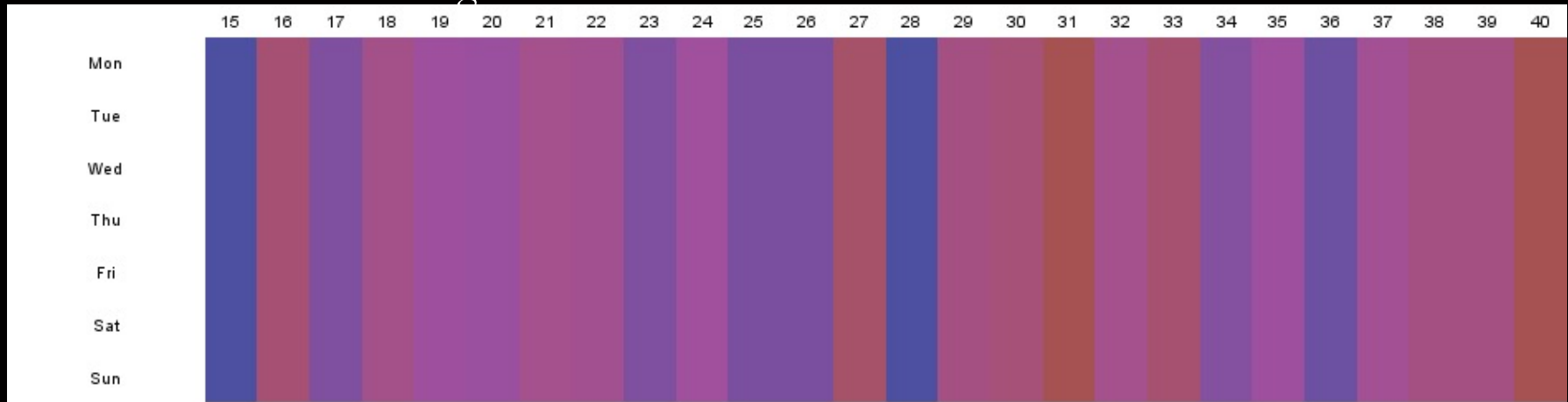
Datos: Marco de Referencia - Espacial



Datos con Relación al Tiempo

Datos: Número de Variables - Univariado

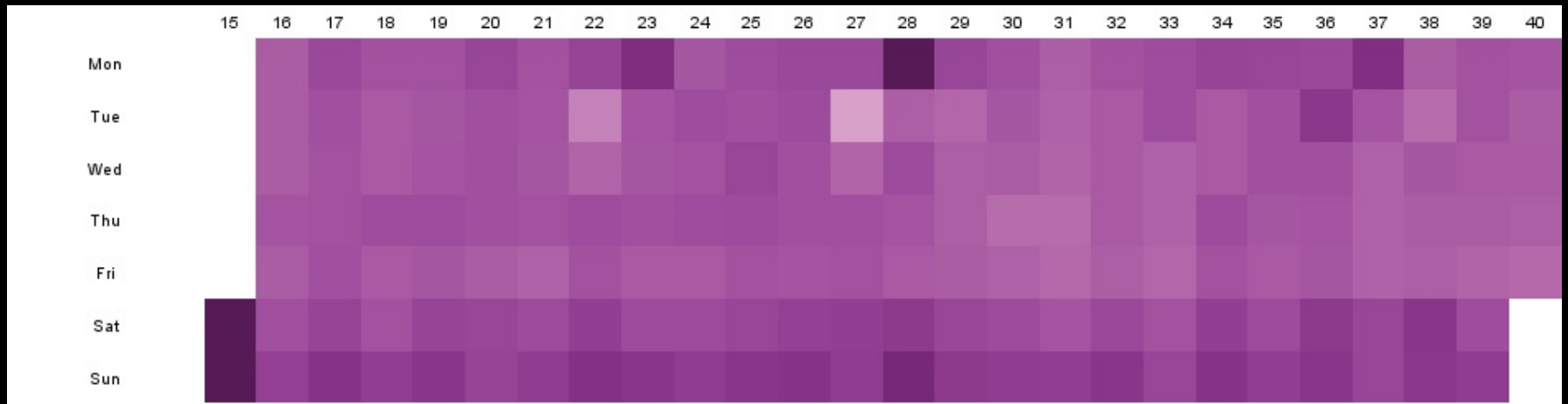
Las visualizaciones usando GROOVE (Granularity Overview OVERlay) extienden las visualizaciones basadas en píxeles al superponer varios niveles de agregación con uno o más métodos: a) traslape basado en color, b) traslape de opacidad o c) traslape espacial. Los niveles de agregación resultan de las diferentes granularidades del tiempo. Los traslapes permiten lecturas a niveles micro y macro y evita los movimientos oculares entre las representaciones detalladas y generales. Los posibles espacios vacíos aparecen debido a irregularidades impuestas por la necesidad de combinar las diferentes granularidades.



Traslape de colores. Datos del tráfico por varias semanas, cada bloque representa una semana. En esta gráfica se muestran sólo valores semanales.

Datos con Relación al Tiempo

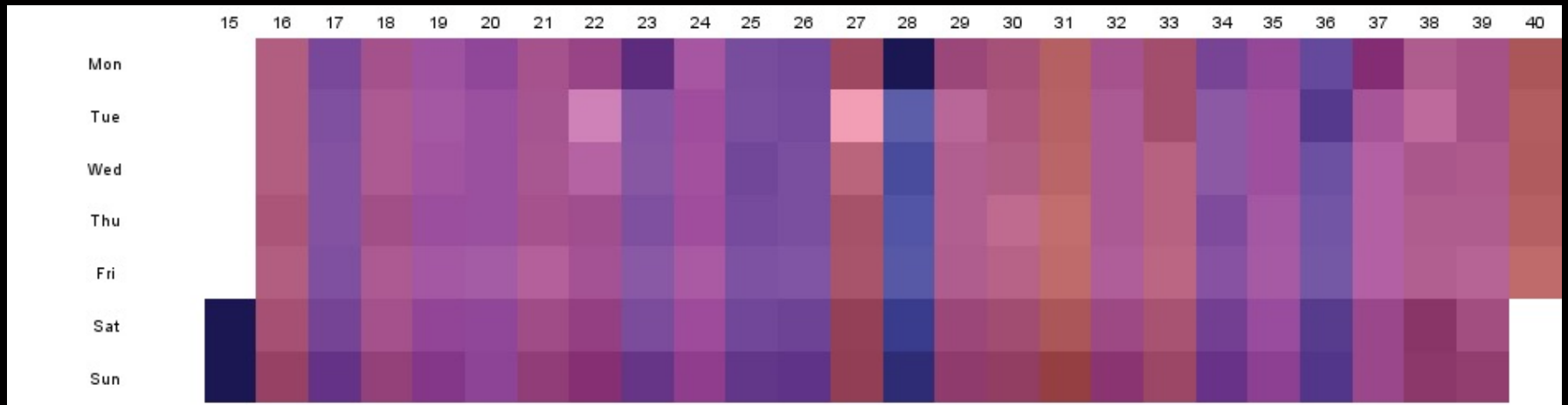
Datos: Número de Variables - Univariado



Traslape de colores. Datos del tráfico por varias semanas, cada bloque representa una semana. En esta gráfica se muestran sólo valores diarios.

Datos con Relación al Tiempo

Datos: Número de Variables - Univariado

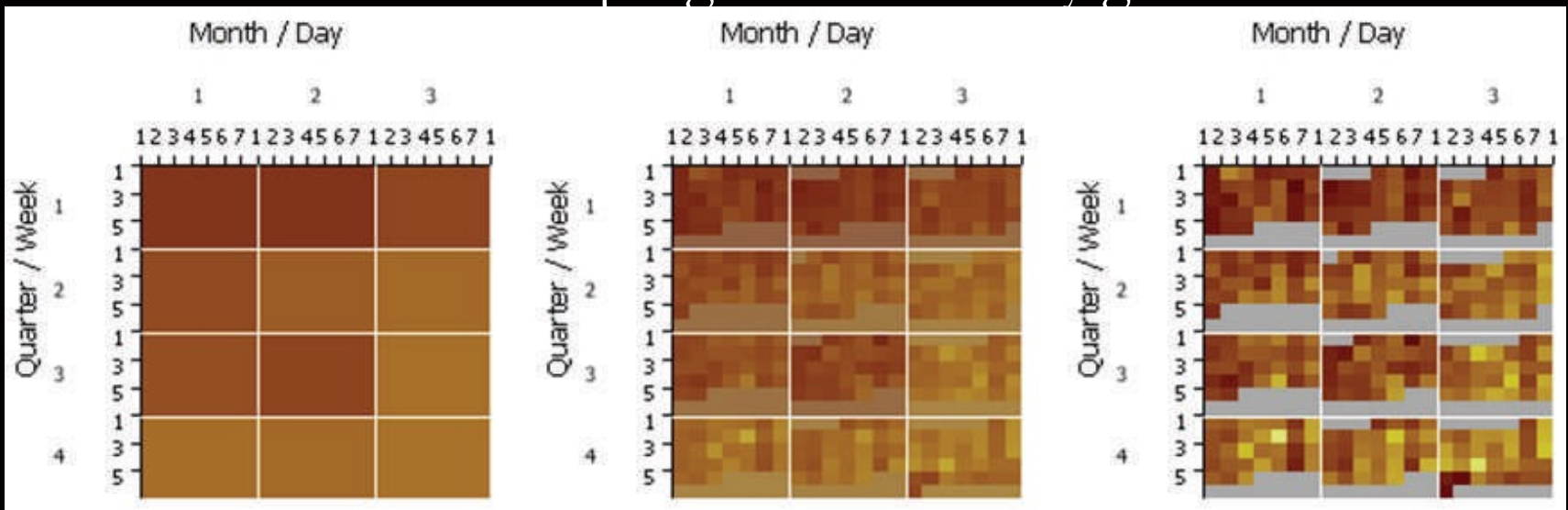


Traslape de colores. Datos del tráfico por varias semanas, cada bloque representa una semana. Combinación de granularidad diaria y semanal.

Datos con Relación al Tiempo

Datos: Número de Variables - Univariado

El traslape de opacidad aplica desvanecimiento cruzado interactivo entre los despliegues de detalle y general.



Asistencias por un año en una tienda utilizando distribución recursiva de patrones. Las columnas y las hileras de los bloques grandes se combinan con un arreglo de píxeles dentro de los bloques para la estructura de los detalles. Cada bloque muestra un mes y los bloques están acomodados con una hilera para cada cuarto. Dentro de los bloques, los píxeles están distribuidos con una hilera por cada semana y un píxel por cada día (parecido a una hoja de calendario).

Datos con Relación al Tiempo

Datos: Número de Variables - Univariado

El traslape espacial se puede utilizar al mostrar el promedio de valores como bordes alrededor de los valores de los detalles, lo cual es posible para más de dos niveles de agregación. El traslape espacial se puede combinar con los otros métodos y se puede aplicar interacción al expandir y colapsar áreas del gráfico.

Datos con Relación al Tiempo

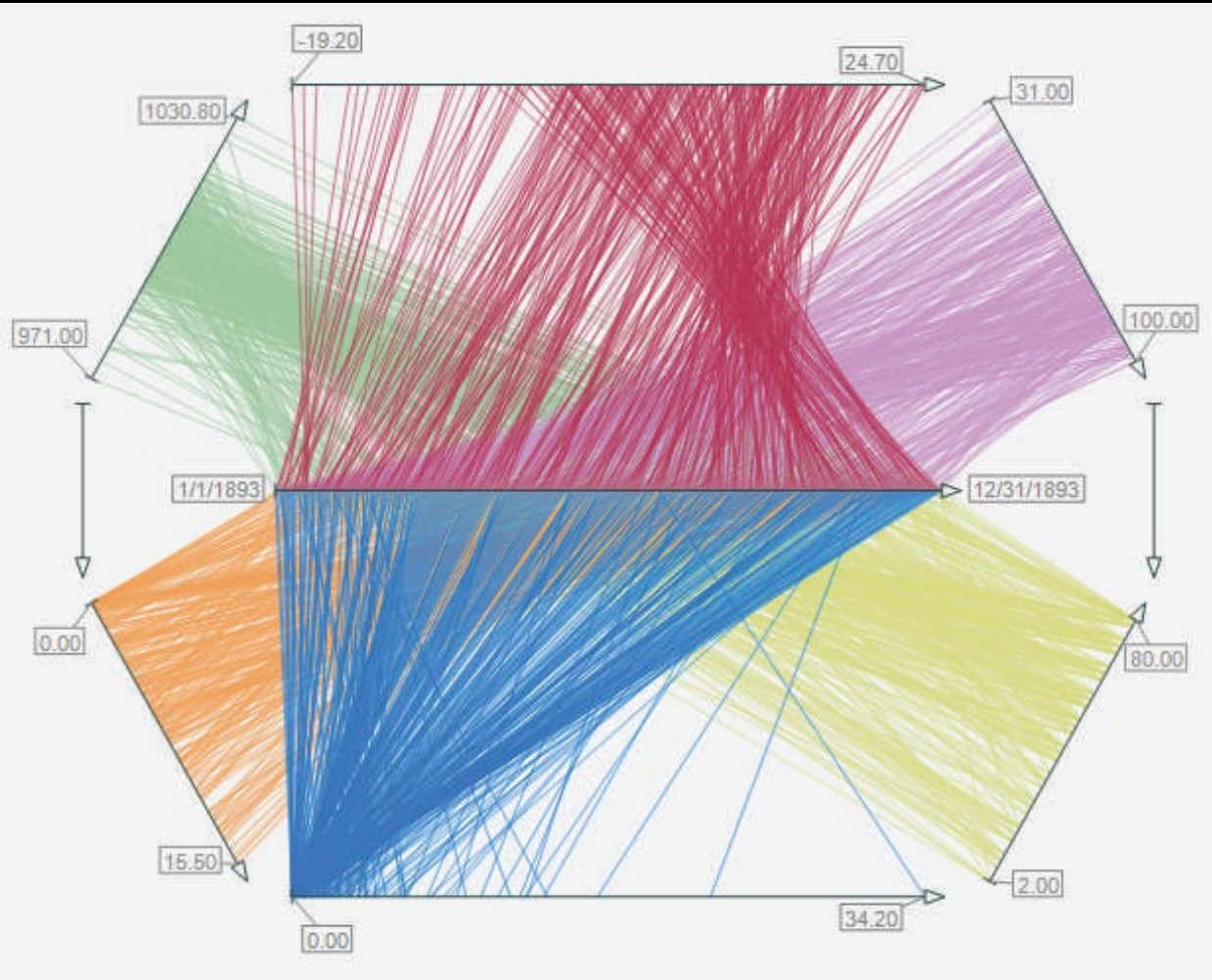
Datos: Número de Variables - Multivariados

La técnica TimeWheel permite visualizar múltiples variables dependientes del tiempo y consiste de un solo eje del tiempo y de múltiples ejes para las variables de datos. El eje del tiempo se coloca en el centro del despliegue para enfatizar el carácter temporal de los datos y los ejes de datos están asociados con colores individuales y están distribuidos circularmente alrededor del eje del tiempo. Para visualizar los datos, líneas emanan del eje del tiempo hacia cada uno de los ejes de datos para establecer una conexión visual entre puntos en el tiempo y los valores de datos asociados. Estas líneas forman patrones visuales que permiten al usuario identificar correlaciones positivas y negativas con el eje del tiempo, las tendencias y los valores atípicos. Tales patrones se pueden discernir mejor para aquellos ejes de datos que son paralelos al eje del tiempo.

Para enfocar los ejes de los datos, el usuario puede rotar la gráfica producida por Time Wheel. Los ejes de datos se enfatizan aún más al estirarlos, otorgándoles más espacio de despliegue. Los ejes de los datos que son perpendiculares al eje del tiempo son más difíciles de interpretar y son, por ello, atenuados usando color difuminado y encogiéndolos. Es claro que este tipo de modelo permite la exploración interactiva por medio de los diferentes tipos de ejes.

Datos con Relación al Tiempo

Datos: Número de Variables - Multivariados



Los ejes en la periferia representan ocho variables de datos meteorológicos. Las líneas rojas muestran la temperatura promedio (aumentando al principio y disminuyendo al final). Las líneas azules representan precipitación pluvial (valores atípicos de gran precipitación, pero precipitación moderada durante el año).

Datos con Relación al Tiempo

Tiempo: Distribución – Lineal

Una de las formas más directas para mostrar datos de series de tiempo es la utilización de un sistema coordenado Cartesiano con el tiempo en el eje horizontal y el valor correspondiente sobre el eje vertical. Cuando se usa un punto es utilizado por cada par de tiempo-valor observado/medido se refiere a una gráfica de puntos o gráfica dispersa (o de puntos dispersos). Existen muchas extensiones a este tipo de gráficas, tales como la utilización de diferentes símbolos en lugar de puntos o técnicas 3D usando gráficas de capas. Esta técnica es particularmente buena para enfatizar valores individuales. Además, mostrar los datos usando posición con respecto a una escala común puede ser percibida más precisamente por el sistema perceptual humano.

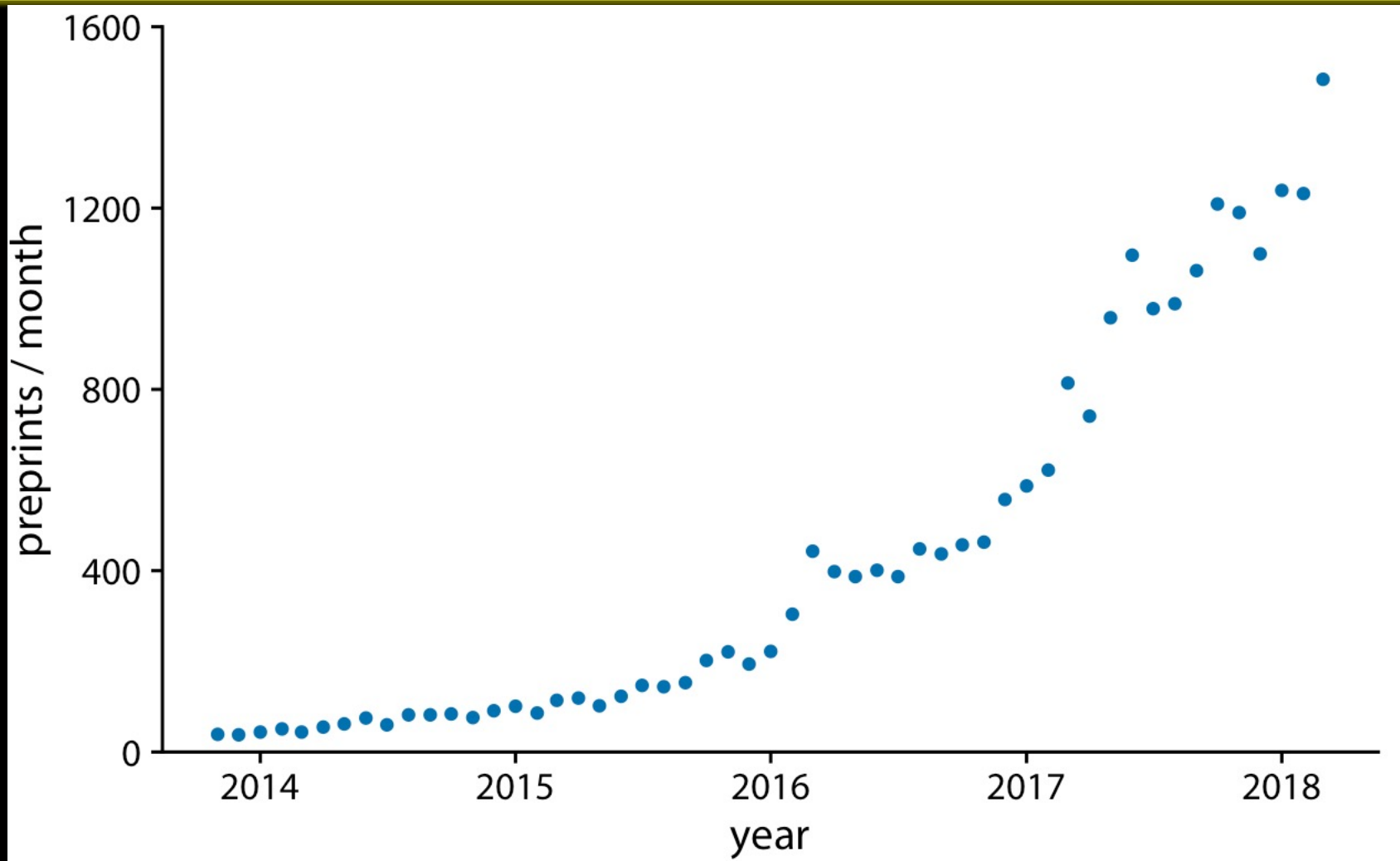
Datos con Relación al Tiempo

Tiempo: Distribución – Lineal

Las gráficas de líneas extienden las gráficas de puntos al enlazar los puntos de los datos con líneas, lo que enfatiza su relación temporal. Consecuentemente, las gráficas de líneas se enfocan en la forma general de los datos a lo largo del tiempo. Esto contrasta con las gráficas de puntos donde se enfatizan los puntos individuales. Existen diferentes tipos de conexiones para enlazar los puntos de los datos tales como líneas rectas, líneas de pasos (para cambios instantáneos de valores) o curvas de Bezier que se pueden usar dependiendo del fenómeno. Sin embargo, se tiene que tener en cuenta las conexiones son una aproximación buena o mala dependiendo del muestreo. Otro punto a tomar en cuenta es la falta de datos ya que conectar puntos de datos puede conllevar a conclusiones falsas sobre los datos. Por lo tanto, esto debe ser informado al observador, por ejemplo, usando líneas punteadas. Existen muchas extensiones o subtipos de estas gráficas, tales como gráficas de fiebre, gráficas de bandas, gráficas de superficie, gráficas de índices o gráficas de control.

Datos con Relación al Tiempo

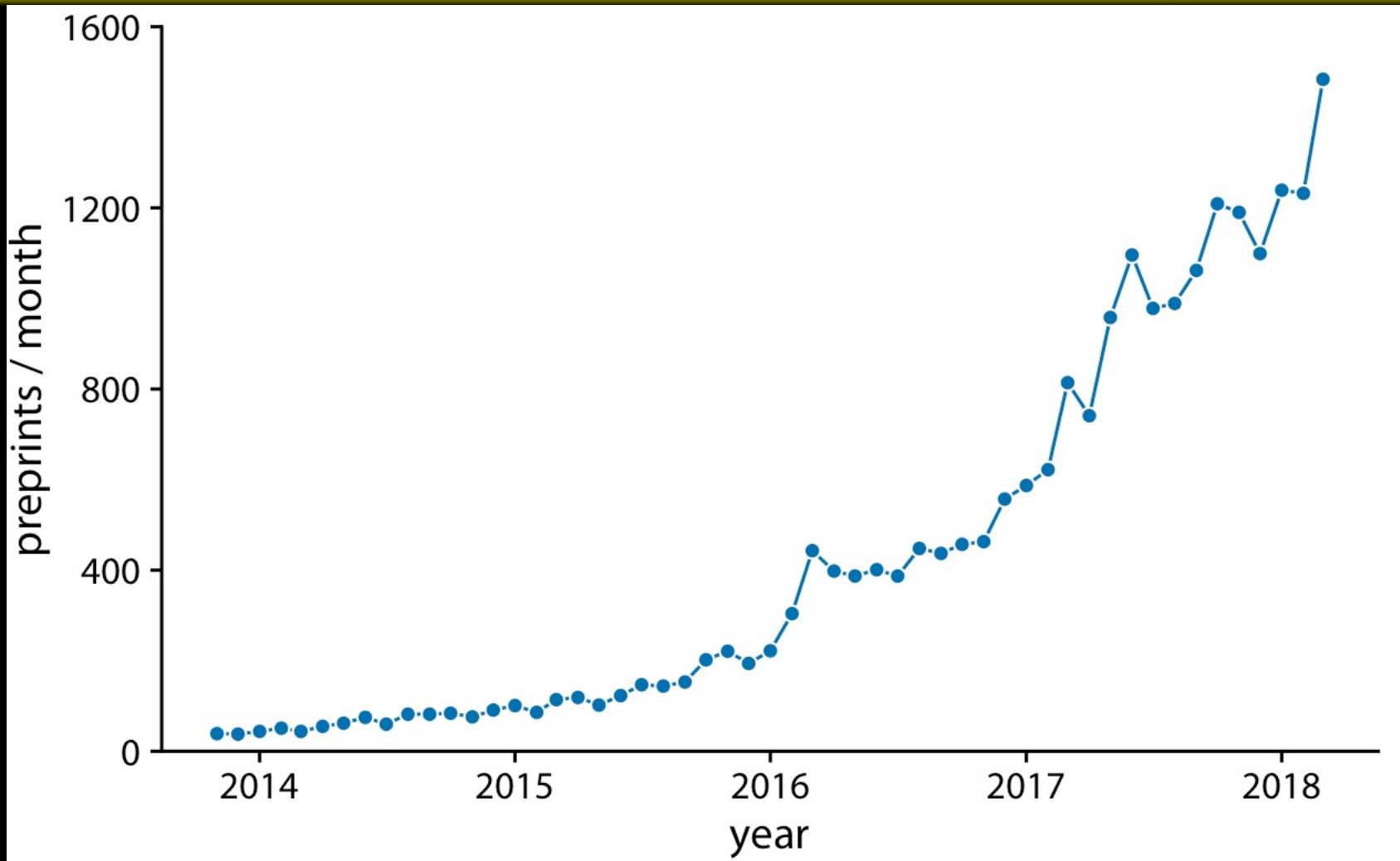
Series de Tiempo Individuales



Pre-publicaciones por mes entre nov. 2013 a abr. 2018

Datos con Relación al Tiempo

Series de Tiempo Individuales



Pre-publicaciones por mes entre nov. 2013 a abr. 2018

Datos con Relación al Tiempo

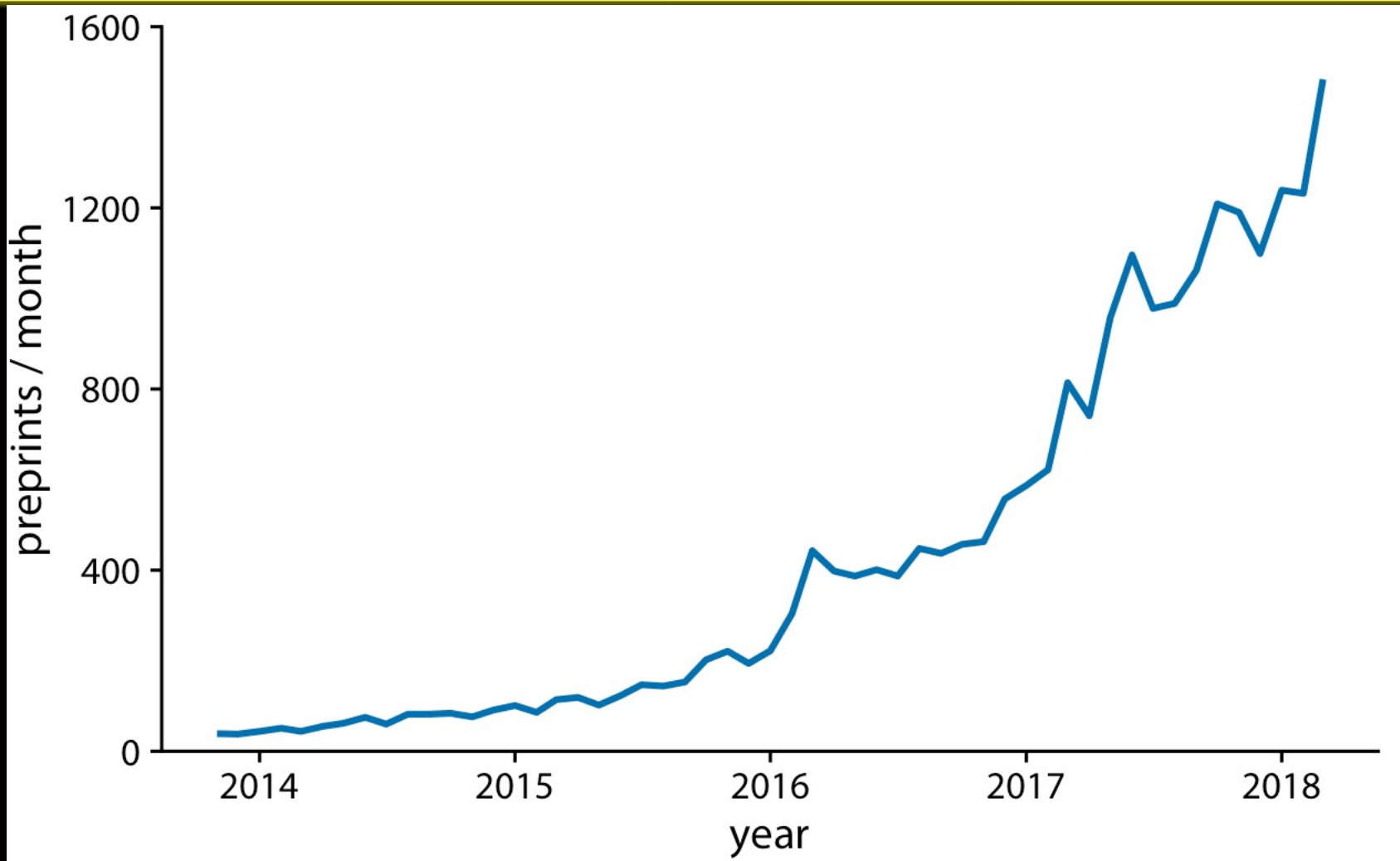
Series de Tiempo Individuales

Algunas personas se oponen a dibujar líneas entre los puntos de muestreo porque las líneas no representan mediciones (datos observados). Por ejemplo, en el caso de que solo existan pocas observaciones con mucha separación, trazar líneas puede implicar que ellas pasen por puntos que no pertenecen al fenómeno (*aliasing*). Sin embargo, la colocación de líneas puede ayudar a la percepción visual y se pueden utilizar líneas curvas si se conoce algo del fenómeno. En el caso de tener datos muy separados, se pueden colocar líneas haciendo énfasis en el pie de figura de que las líneas son sólo una guía para el observador.

El uso de líneas para representar series de tiempo es una práctica aceptada y en algunos casos se omiten los puntos que ilustran las muestras; en este caso se hace énfasis en la tendencia en los datos y menos en las observaciones individuales. También, desde el punto puramente visual, una figura sin puntos está menos poblada. En general, entre más densa sea una serie de tiempo, menos importante será mostrar las observaciones individuales con puntos.

Datos con Relación al Tiempo

Series de Tiempo Individuales



Pre-publicaciones por mes entre nov. 2013 a abr. 2018

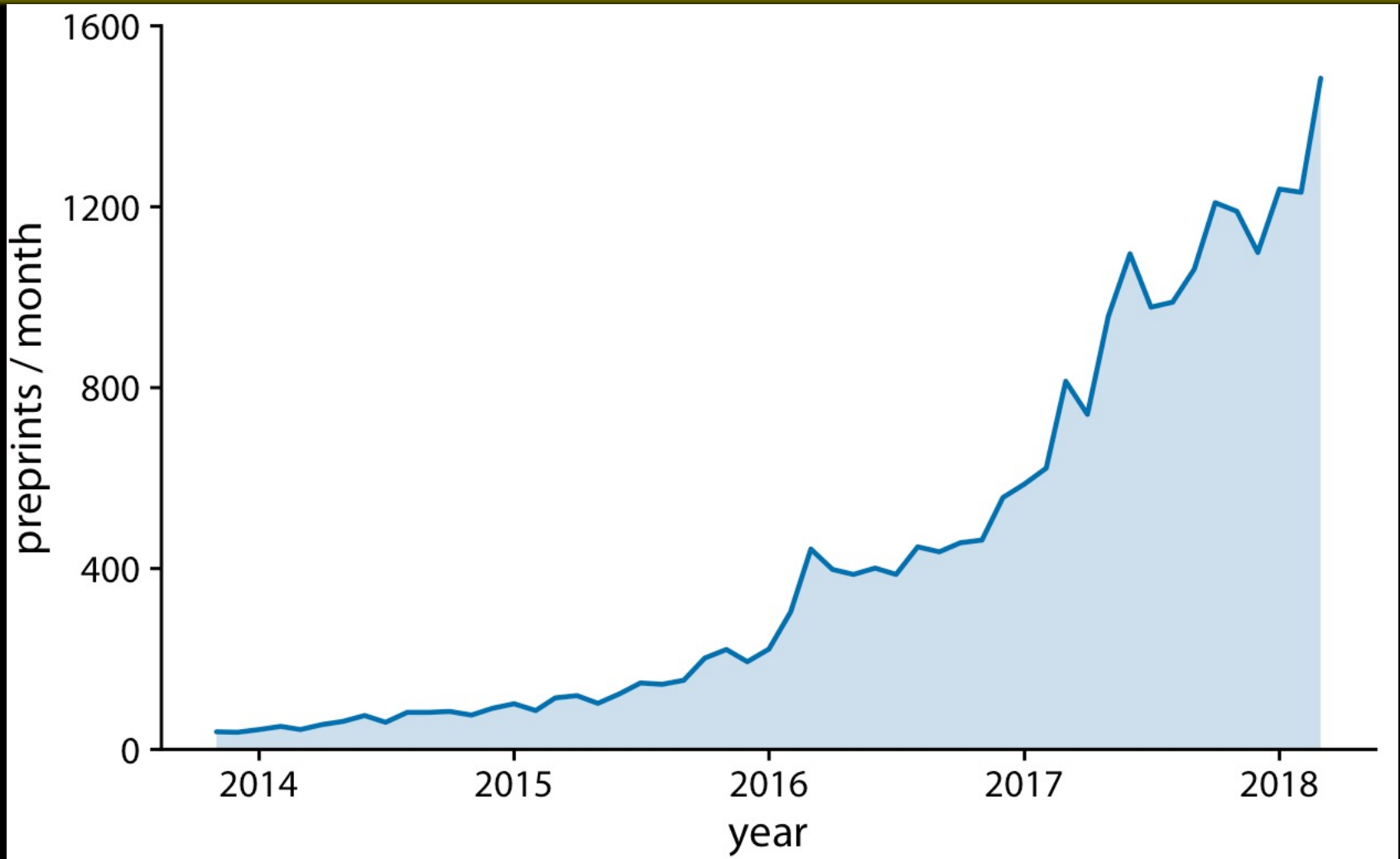
Datos con Relación al Tiempo

Series de Tiempo Individuales

También se puede rellenar el área debajo de la curva con un color solido de un color más claro que la línea de la serie. Esta opción enfatiza aún más la tendencia general en los datos porque separa visualmente las áreas superior e inferior a la curva. Pero, este tipo de visualización es valida sólo si el eje vertical comienza en cero de tal forma de que la altura del área sombreada en cada punto representa el valor de los datos en dicho punto.

Datos con Relación al Tiempo

Series de Tiempo Individuales



Pre-publicaciones por mes entre nov. 2013 a abr. 2018

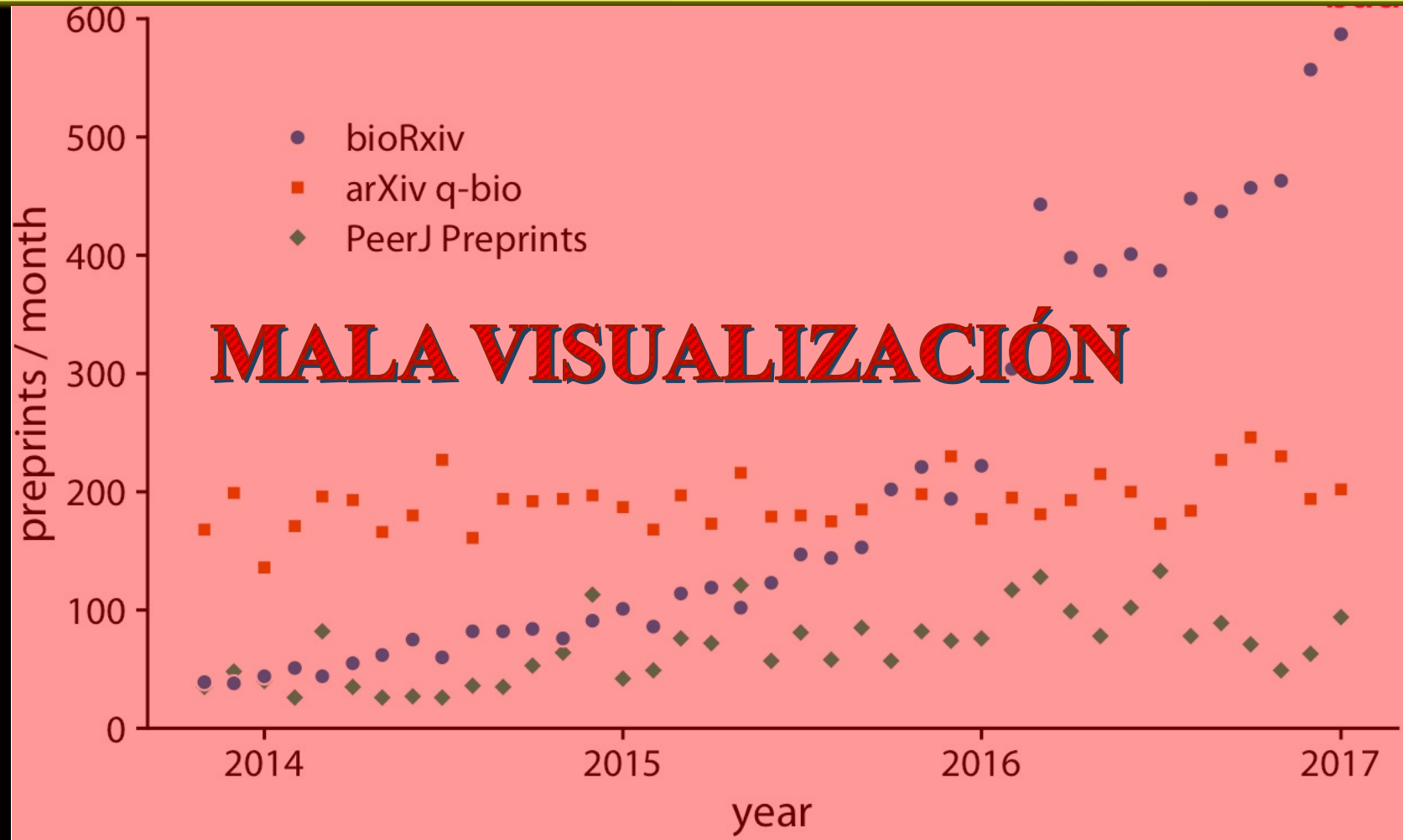
Datos con Relación al Tiempo

Series de Tiempo Múltiples

En varias ocasiones se desean visualizar varios procesos con respecto al tiempo. En este caso, se tiene que ser más cuidadoso en como se grafican los datos porque la visualización se puede volver confusa o difícil de leer.

Datos con Relación al Tiempo

Series de Tiempo Múltiples



Varías entregas de pre-publicaciones por mes entre nov. 2013 a abr. 2018

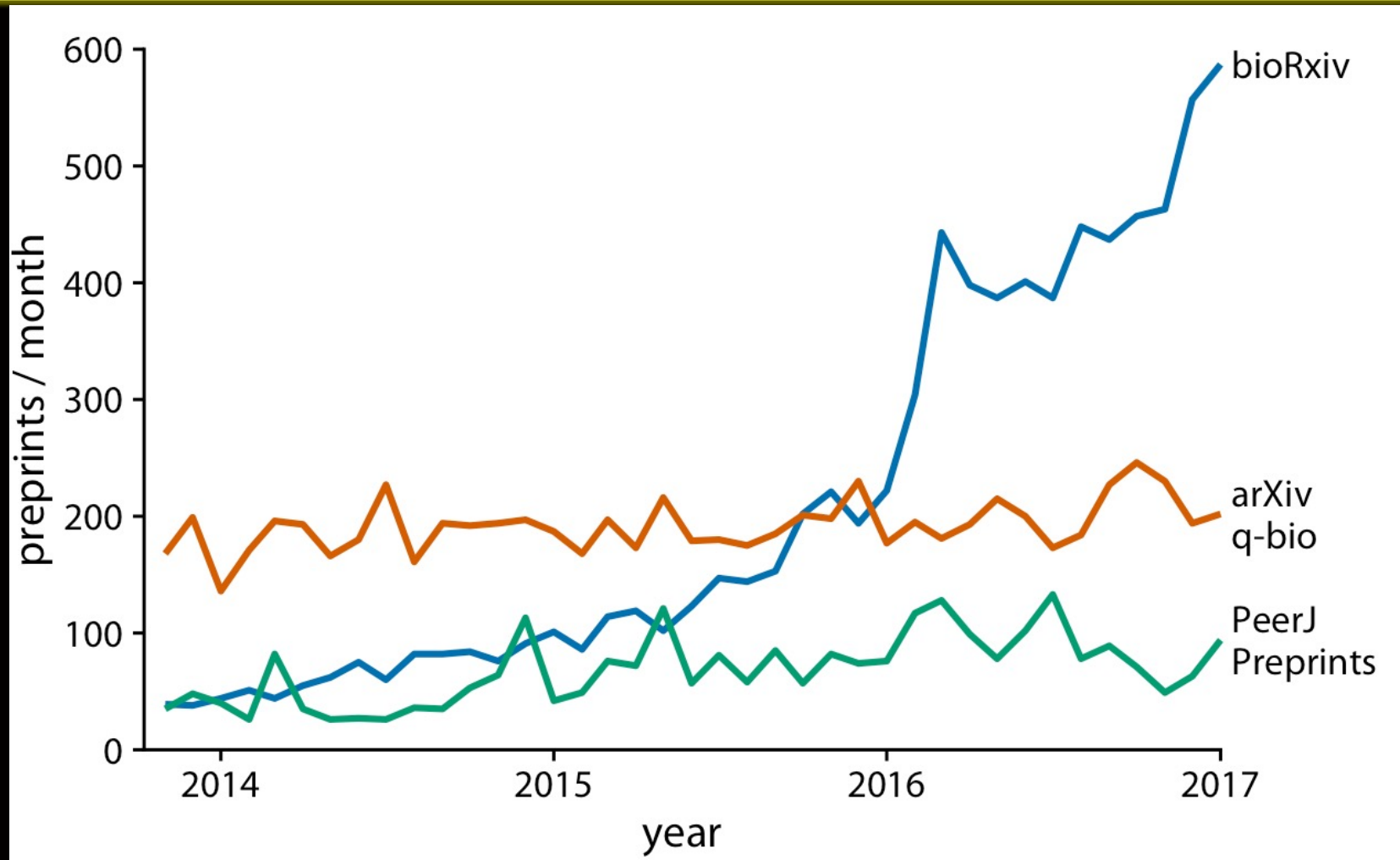
Datos con Relación al Tiempo

Series de Tiempo Múltiples

La utilización de gráficas de puntos no es una buena idea porque las trayectorias del tiempo individuales se pueden encimar. Esto se puede mejorar al conectar los puntos. Sin embargo, la utilización de leyendas separadas puede ocasionar un gráfica saturada y es mejor etiquetar las líneas directamente.

Datos con Relación al Tiempo

Series de Tiempo Múltiples



Varías entregas de pre-publicaciones por mes entre nov. 2013 a abr. 2018

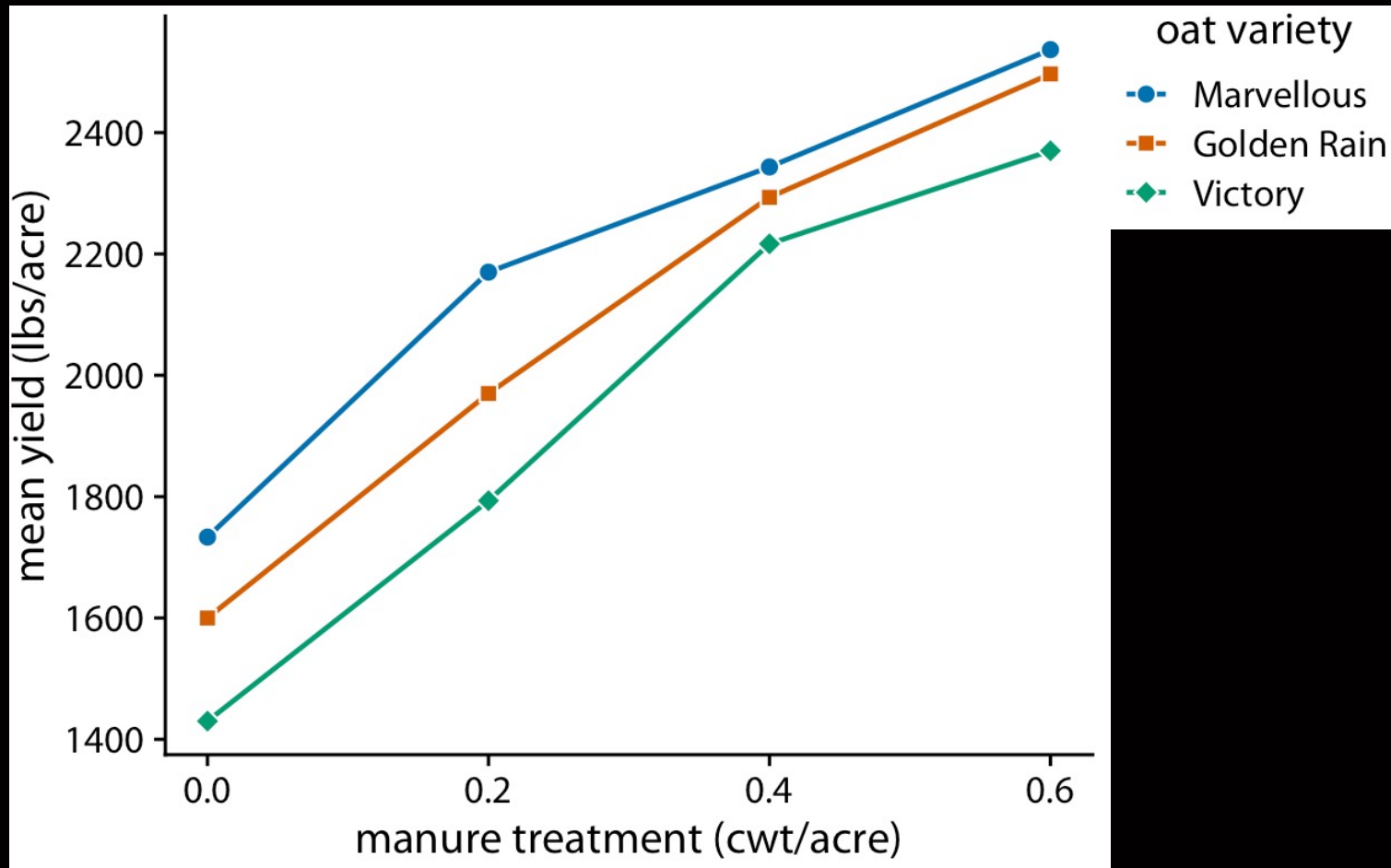
Datos con Relación al Tiempo

Series de Tiempo Múltiples

Las gráficas de líneas no están limitadas a series de tiempo y se pueden utilizar siempre que los puntos de los datos tengan un orden natural que se refleja en la variable independiente (abscisas) de tal manera que los puntos vecinos se puedan conectar por una línea. Un ejemplo de esta situación lo representan las curvas de respuesta a las dosis.. Este tipo de gráficas muestran como el cambio en un parámetro experimental (la dosis) afecta la respuesta buscada (la respuesta).

Datos con Relación al Tiempo

Series de Tiempo Múltiples



Producción de avena en respuesta a la cantidad de fertilizante

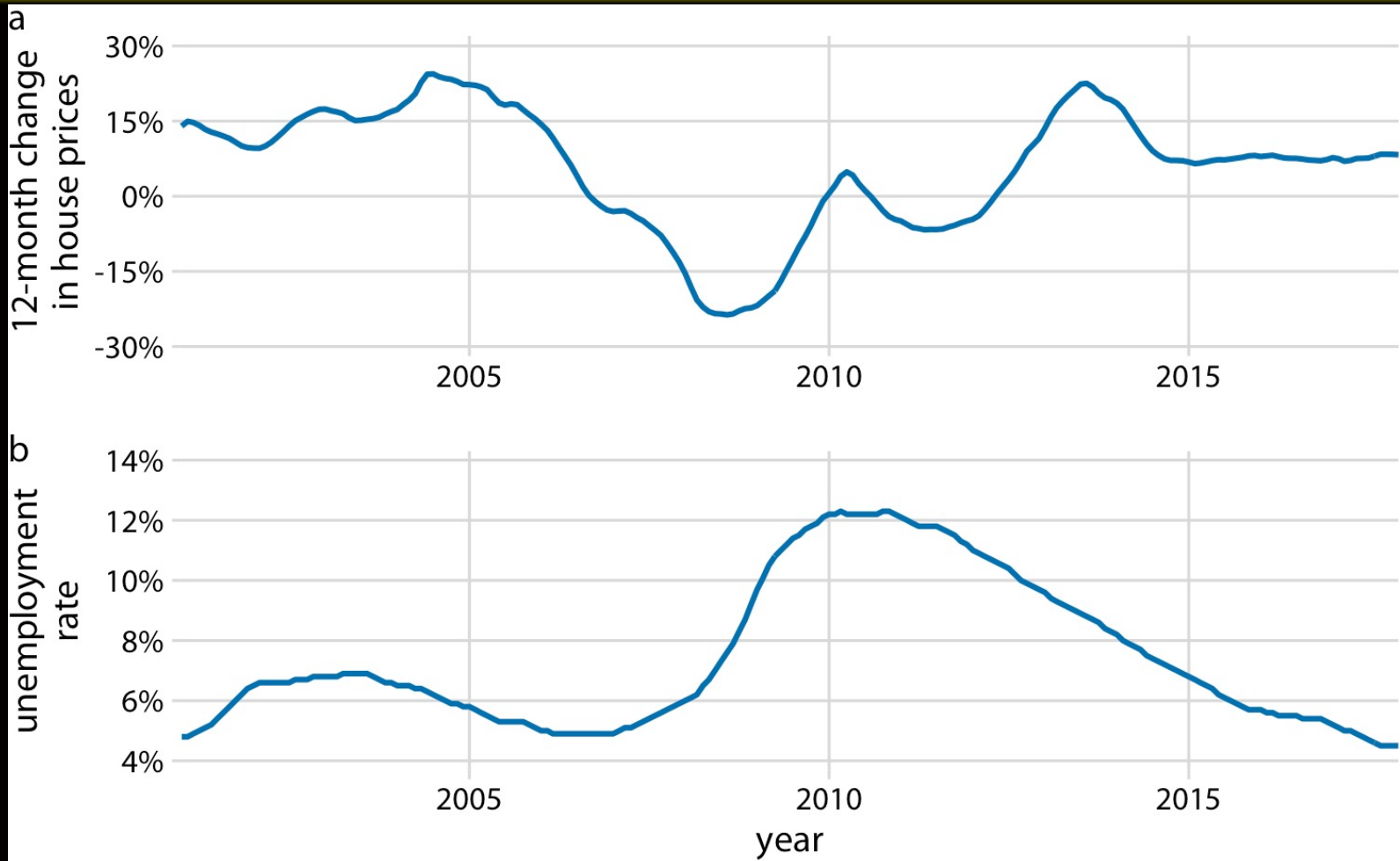
Datos con Relación al Tiempo

Series de Tiempo Multivariadas

En casos en que se tienen varias variables con relación al tiempo se pueden usar varias gráficas de líneas, cada una para representar cada variable, apiladas una sobre la otra. Este tipo de visualización permite una interpretación fácil, pero puede hacer difícil realizar comparaciones (cambiar de referencia) o se puede saturar la visualización rápidamente.

Datos con Relación al Tiempo

Series de Tiempo Múltiples



Gráficas de los precios de casas y de la tasa de desempleo.

Datos con Relación al Tiempo

Tiempo: Distribución – Lineal

Existen programas que permiten explorar este tipo de datos, entre ellos TiMoVA (Time series analysis, Model selection, applies Visual Analytics methods) que permiten usar gráficas de puntos, líneas y otro tipo de gráficas y pistas. Una ventaja de este tipo de programas es que proveen un ambiente de exploración interactiva para guiar al usuario en el proceso de selección del modelo de series de tiempo. El programa TiMoVA permite a un experto la selección interactiva del orden del modelo, obtener retroalimentación visual de los resultados del modelo mientras se selecciona el orden del modelo y ofrecer visualizaciones al transitar de un modelo a otro para poder ver si el modelo mejora o no.

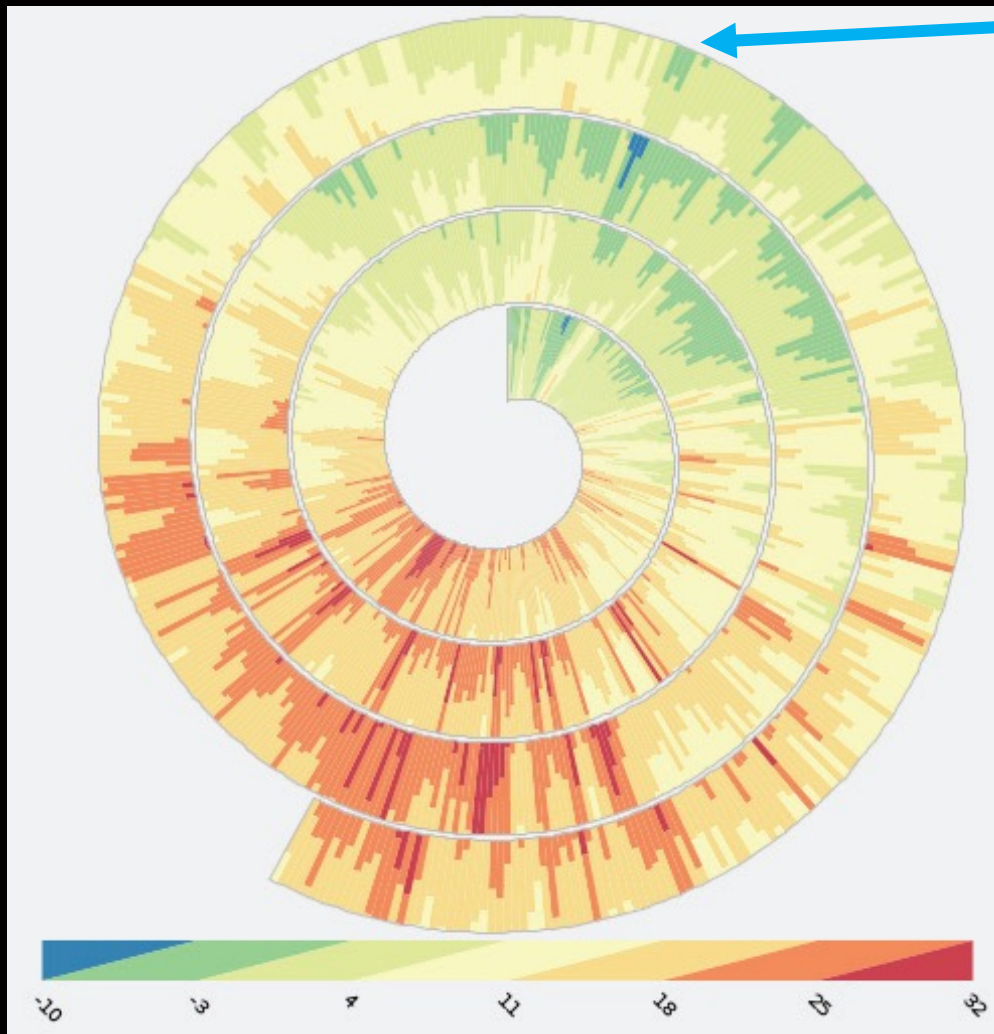
Datos con Relación al Tiempo

Tiempo: Distribución – Cíclica

Para los despliegues cíclicos de datos dependientes del tiempo se pueden aplicar una codificación de color de dos tonos. Cada primitiva de tiempo se mapea a un segmento único de la espiral y cada segmento es subdividido en dos partes que son coloreados de acuerdo al método de coloreado con dos tonos. La ventaja de utilizar el método de dos tonos es que logra el concepto del diseño general+detalle. Los dos colores usados en cada segmento de la espiral permite a un usuario reconocer rápidamente el rango de valores de dicho segmento (general). Si el rango de valores es de interés, la proporción de los dos colores indica más precisamente el valor de datos particular (detalle). La espiral interactiva, una mejora sobre la estática en varios sentidos. El número de primitivas de tiempo, el número de ciclos y otros parámetros geométricos influyen la forma de la espiral y, por lo tanto, el mapeo del dominio del tiempo. La representación de los datos está controlada por las escalas de color que se aplica y por los parámetros tales como el número de colores, la dirección del mapeo y la función de mapeo (lineal vs. logarítmica). La navegación a lo largo del tiempo es posible via manipulación directa de la espiral.

Datos con Relación al Tiempo

Tiempo: Distribución – Cíclica

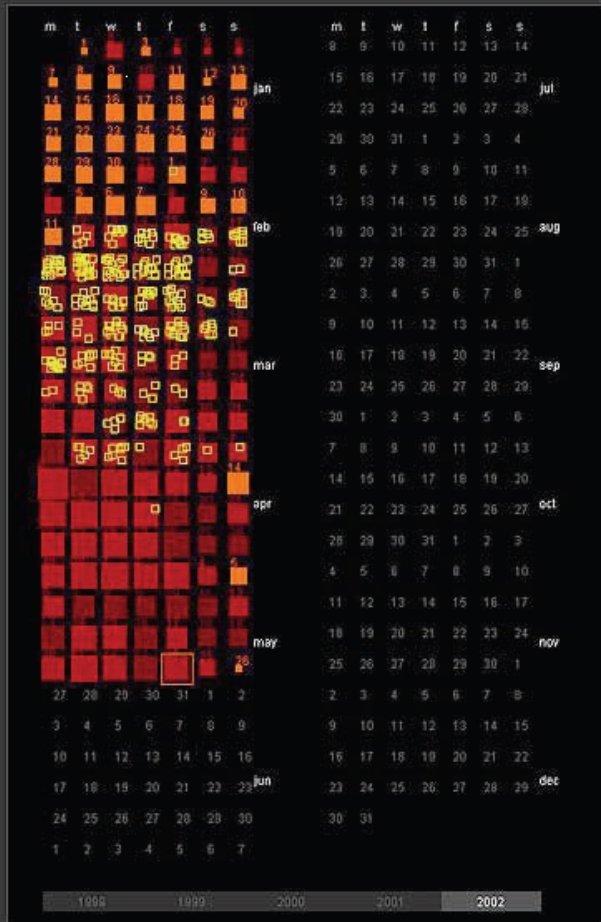


Invierno Templado

Gráfica que muestra 3 ½ años de datos del clima. Cada ciclo visualiza 365 valores del promedio diario de temperaturas en la Cd. De Rostock.

Datos con Relación al Tiempo

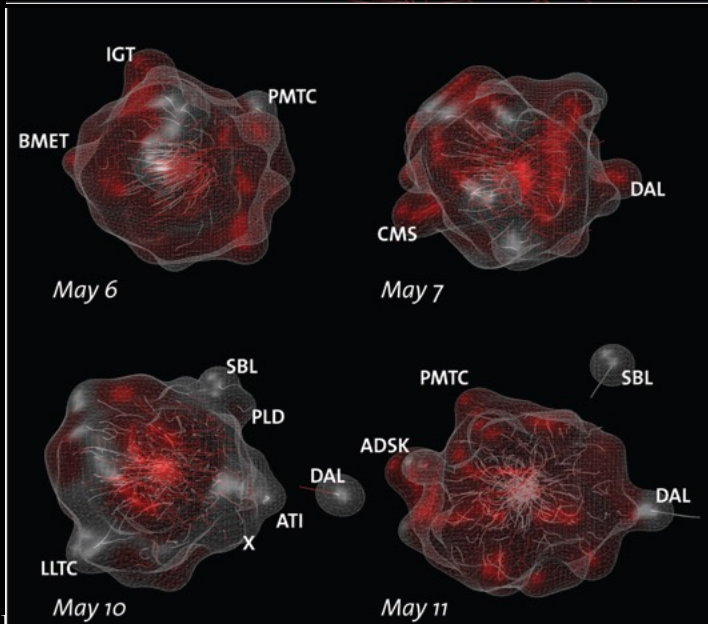
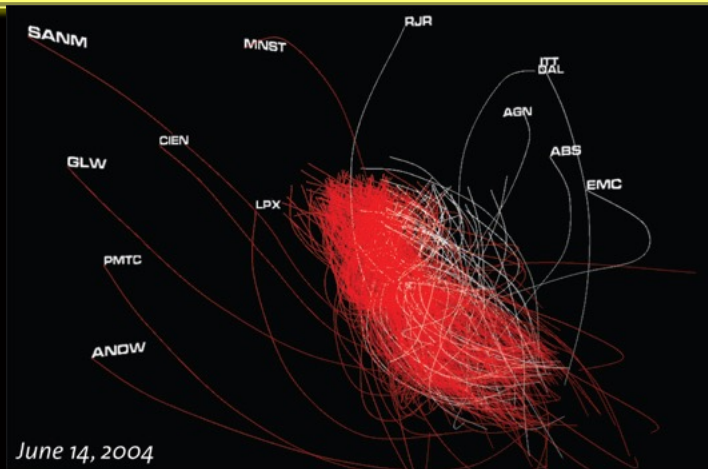
Visualización: Mapeo – Estático



PostHistory. El calendario a la izquierda muestra la actividad diaria de emails en donde el número de mensajes y el promedio de direcciones se mapean a tamaño de cajas y color, respectivamente. El panel de la derecha muestra los nombres de la gente que ha enviado los mensajes a cada usuario.

Datos con Relación al Tiempo

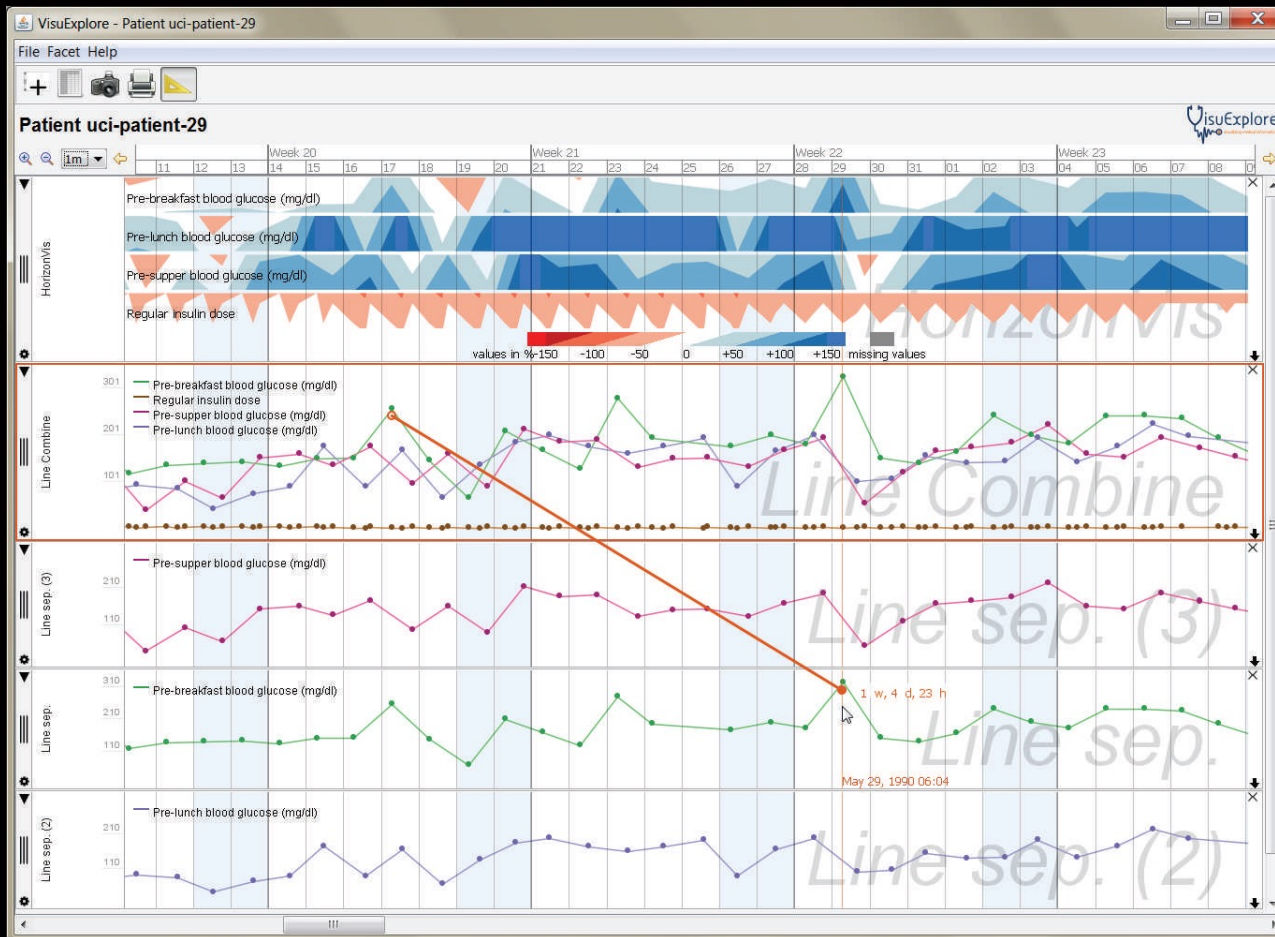
Visualización: Mapeo – Dinámico



Flocking Boids. Representación de datos de la bolsa como parvadas que se mueven en el espacio 3D. (arriba) Los pájaros dejan la parvada indican que los precios de las acciones se comportan diferente a la mayoría de los precios. (abajo) Superficies implícitas que rodean a las parvadas ayudan a los usuarios a reconocer la estructura espacial de ellas.

Datos con Relación al Tiempo

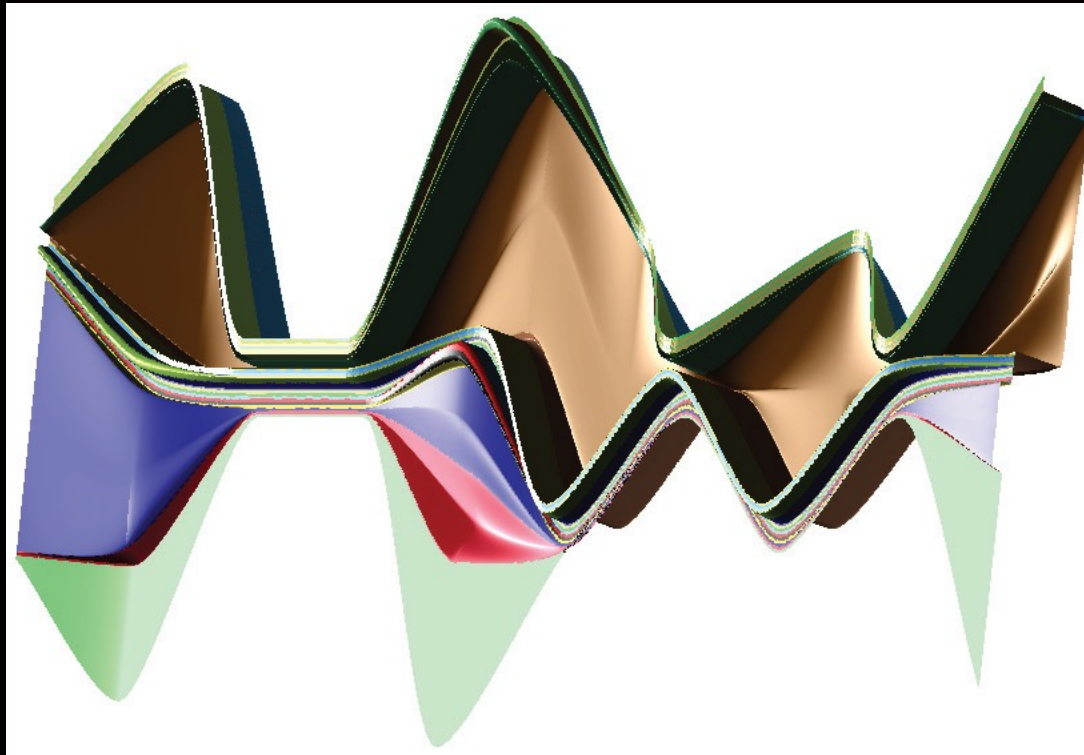
Visualización: Dimensionalidad – 2D



VisuExplore. Métodos para una representación visual fácil de entender usando gráficas de líneas, barras, eventos, líneas de tiempo, horizontales y de línea con zoom semántico. Se muestra una cinta métrica metafórica para el intervalo entre dos elementos que se encuentran en dos diagramas diferentes.

Datos con Relación al Tiempo

Visualización: Dimensionalidad – 2D



The 3D ThemeRiver. Las distintas corrientes coloreadas crean las formas generales. El ancho y la altura de las corrientes varían para visualizar datos relacionados con el tiempo. El ancho codifica la distribución general de 17 grupos de datos de aerosol y la altura indica la incidencia de zinc.