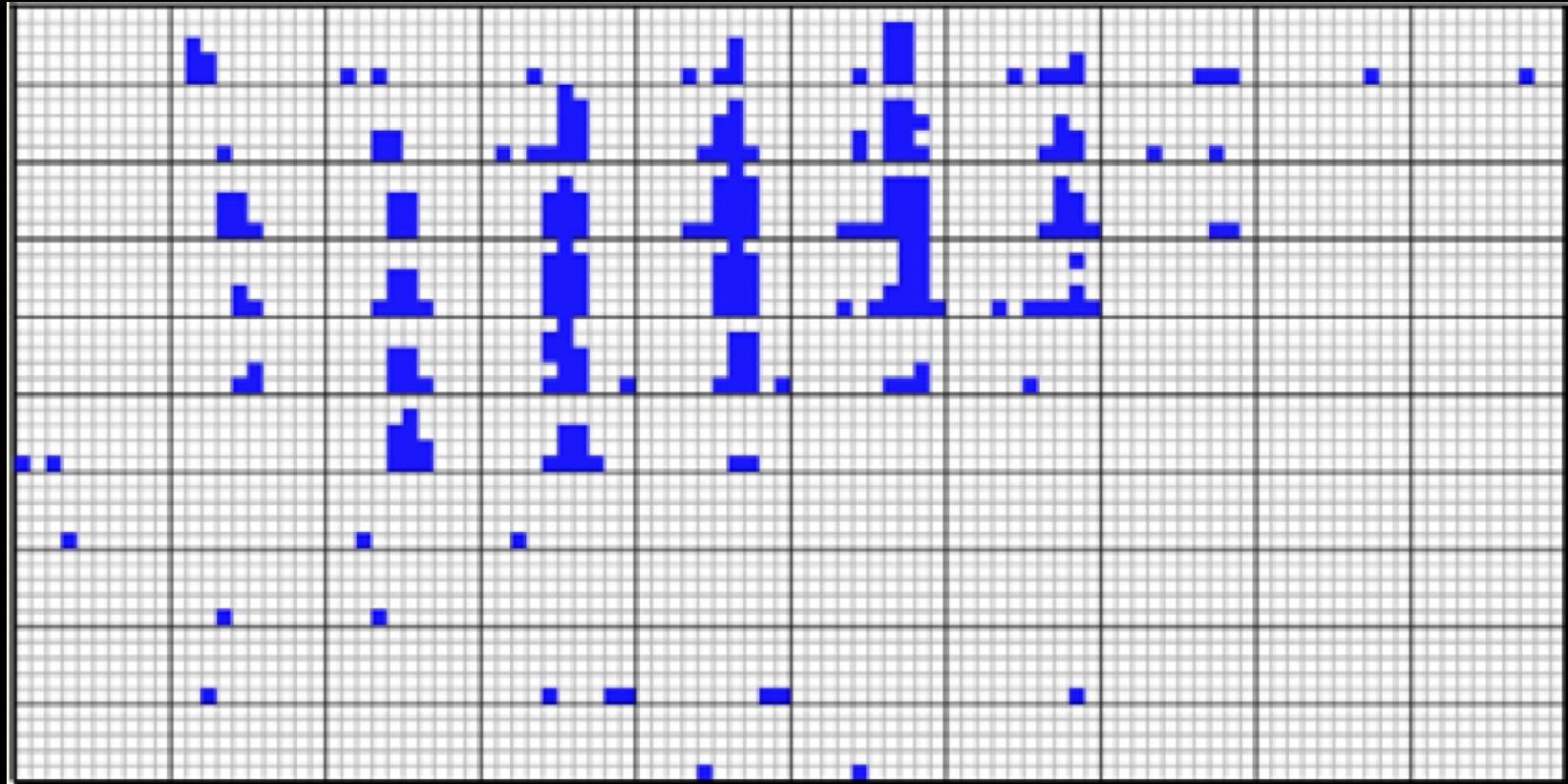


# Datos Multivariados

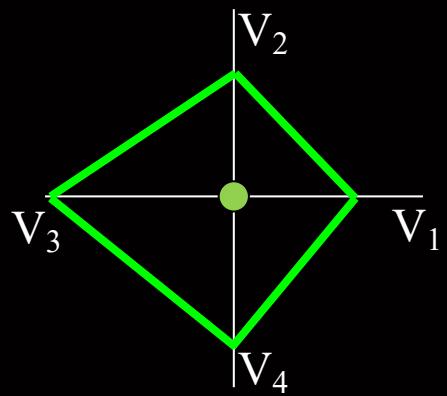
## Apilado Dimensional



Datos de extracción de petróleo con la longitud y latitud mapeadas sobre los ejes x e y externos, el grado y la profundidad sobre los ejes x e y internos.

# Datos Multivariados

## Gráficas con Glifos

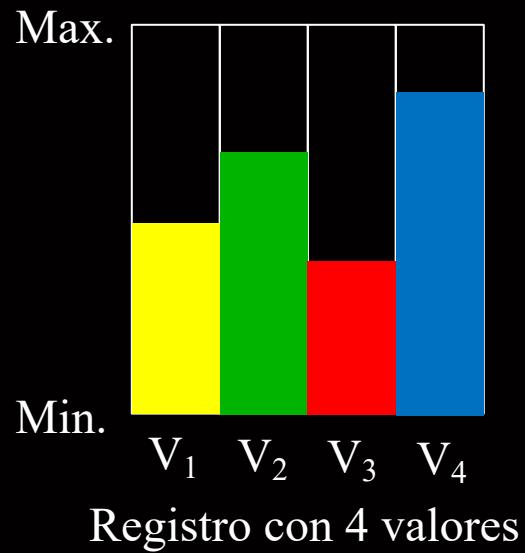


Registro con 4 valores



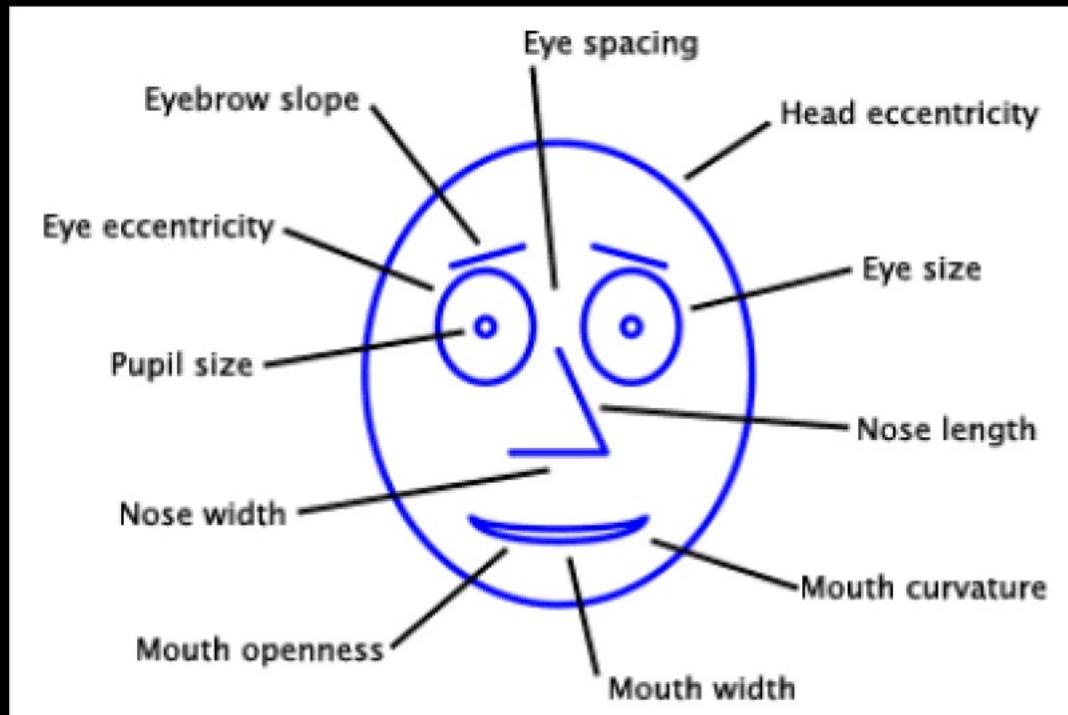
# Datos Multivariados

## Gráficas con Glifos



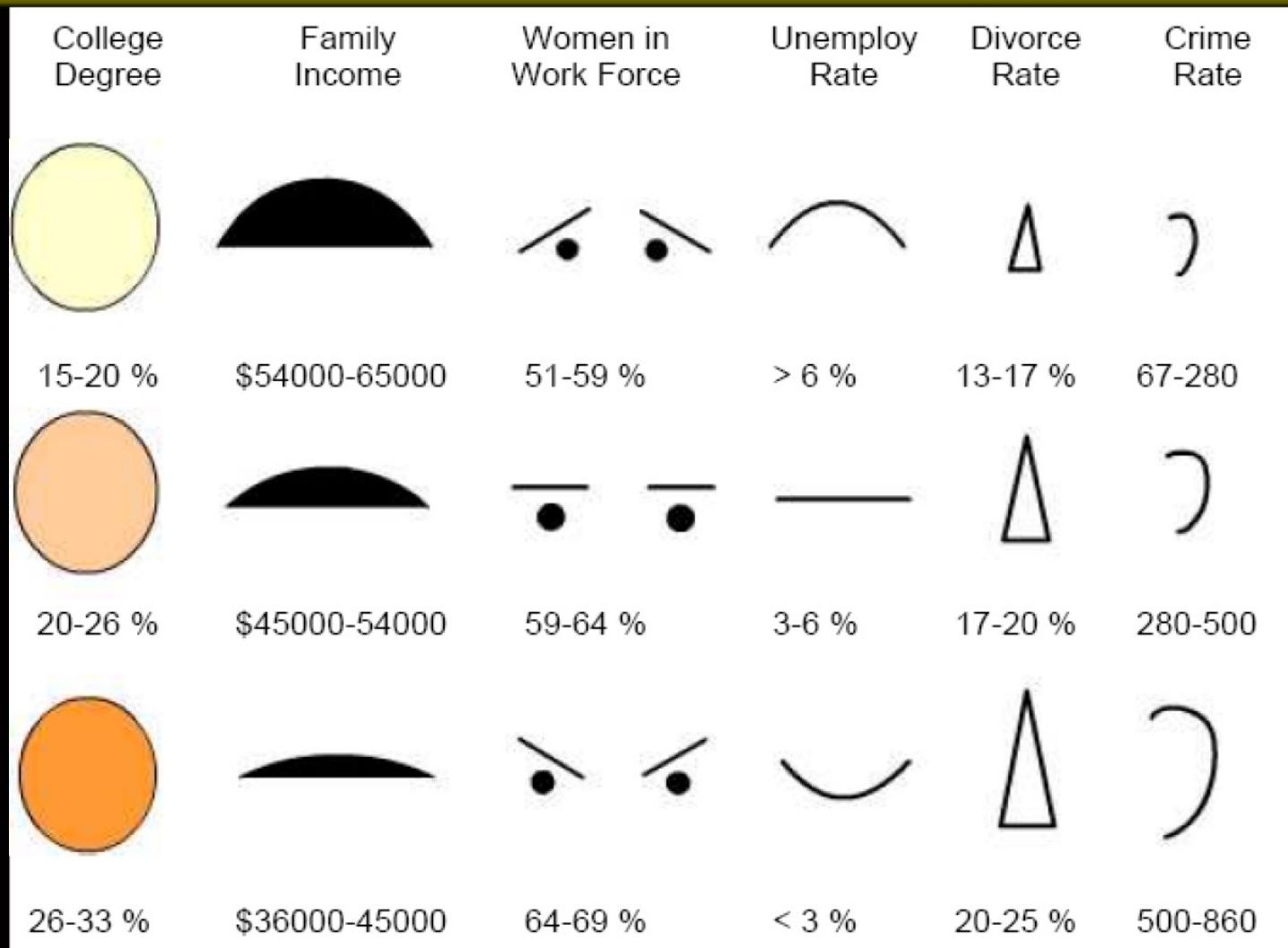
# Datos Multivariados

## Gráficas con Glifos (Rostros de Chernoff)



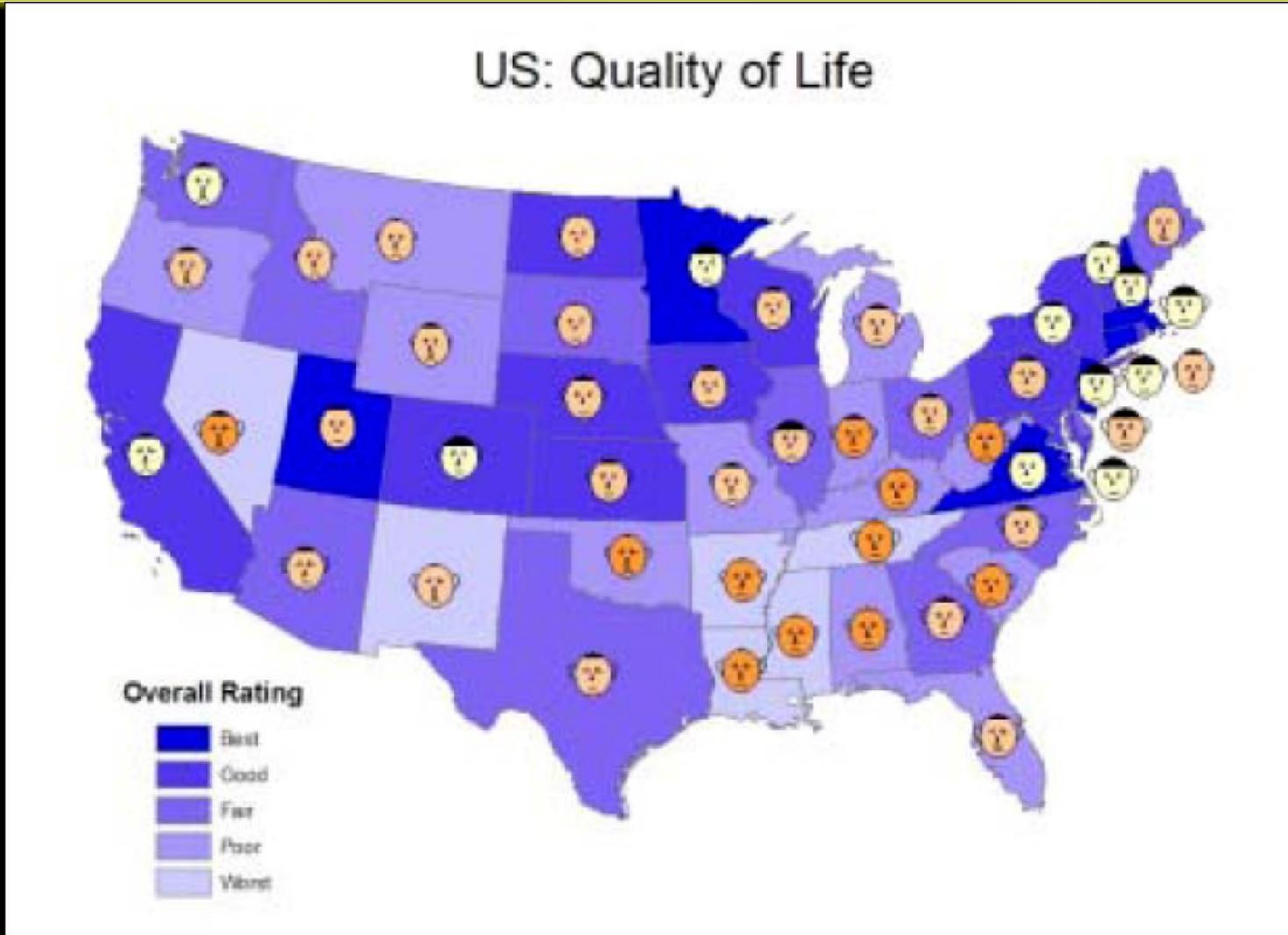
# Datos Multivariados

## Gráficas con Glifos (Rostros de Chernoff)



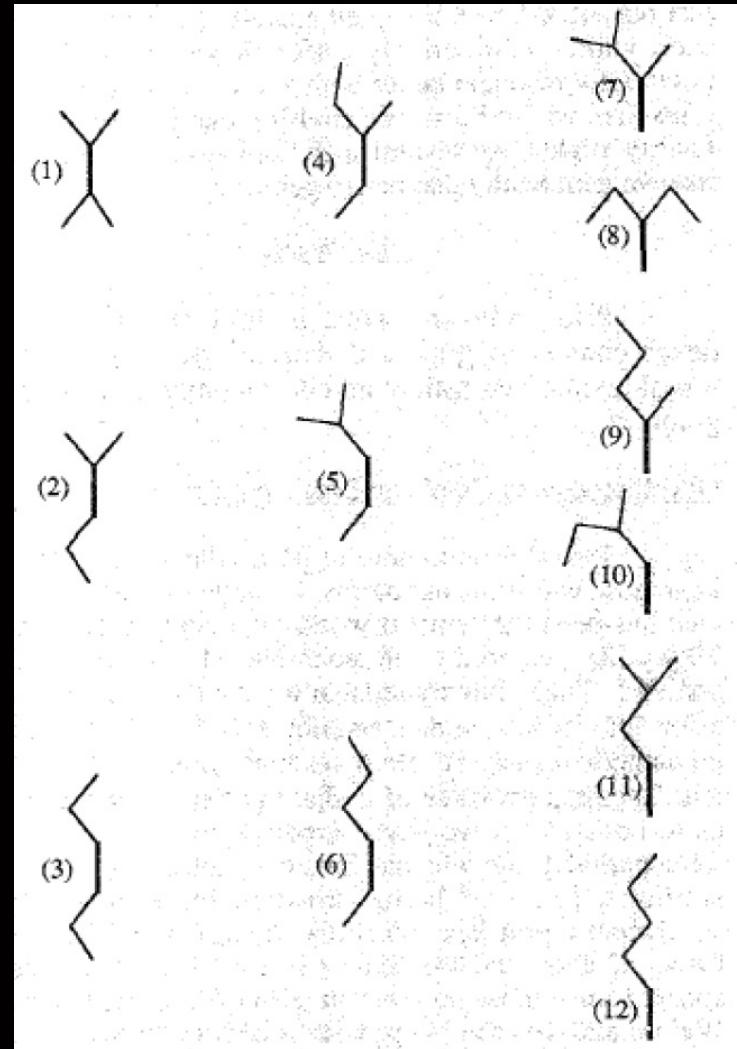
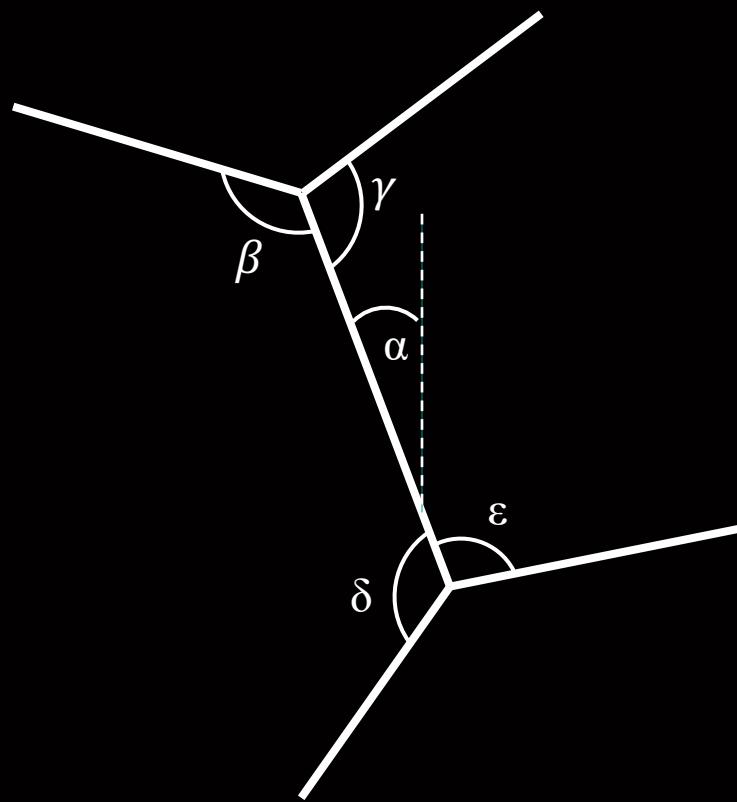
# Datos Multivariados

## Gráficas con Glifos (Rostros de Chernoff)



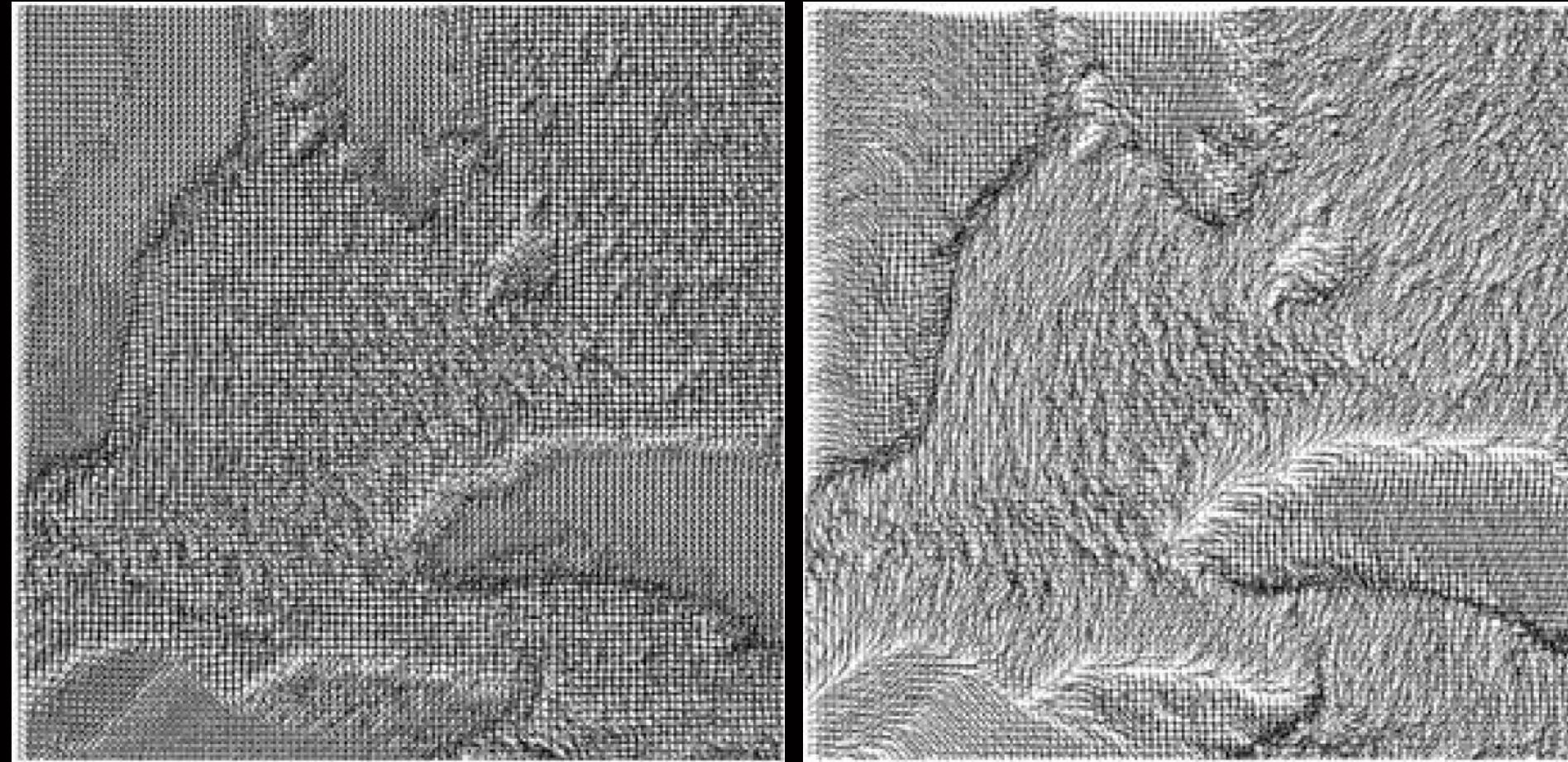
# Datos Multivariados

## Gráficas con Bastones/Palillos



# Datos Multivariados

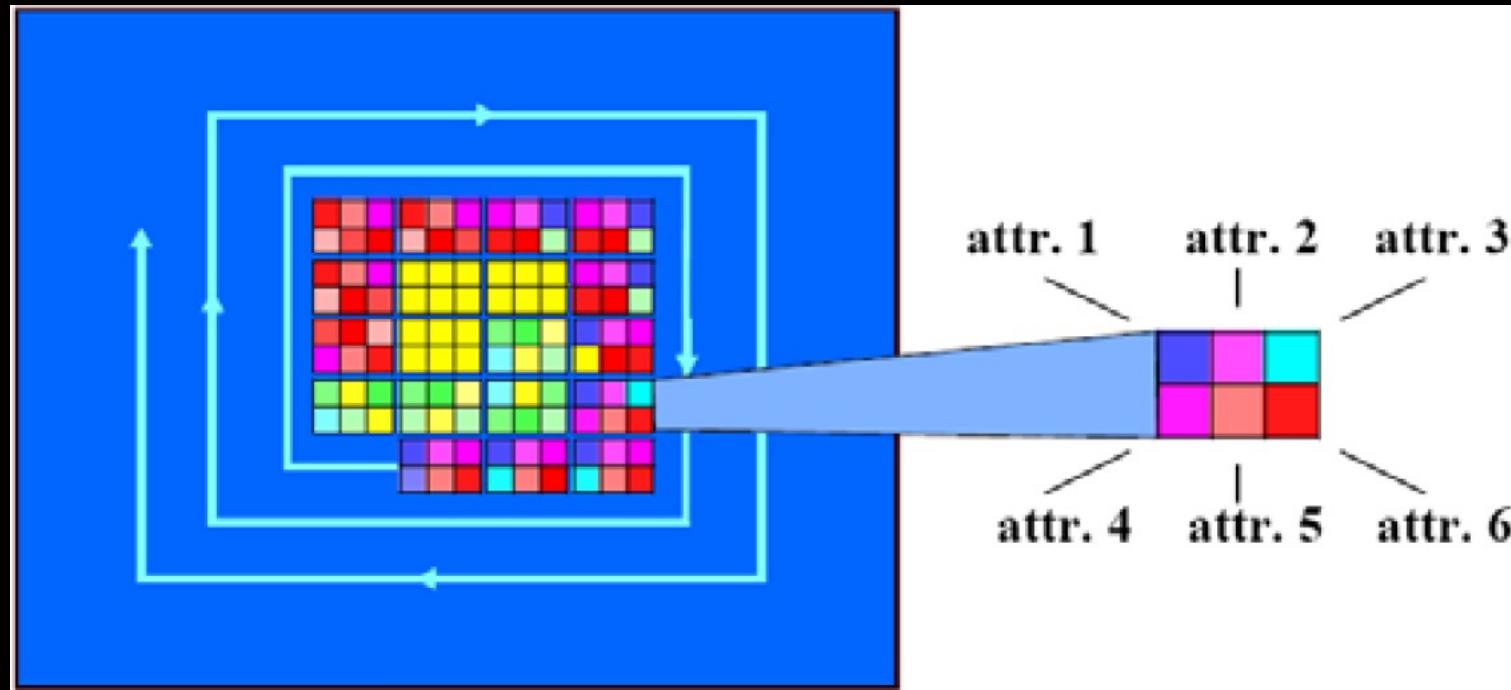
## Gráficas con Bastones/Palillos



Diferentes mapeos.

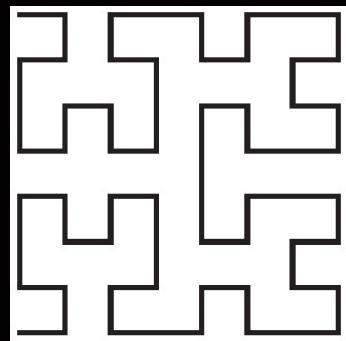
# Datos Multivariados

## Técnicas Basadas en Píxeles

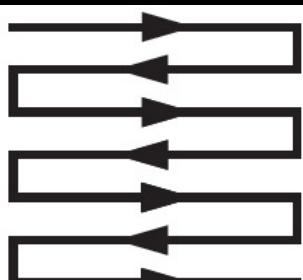


# Datos Multivariados

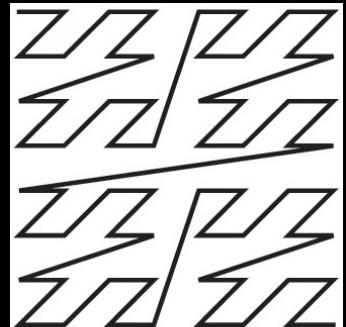
## Técnicas Basadas en Píxeles



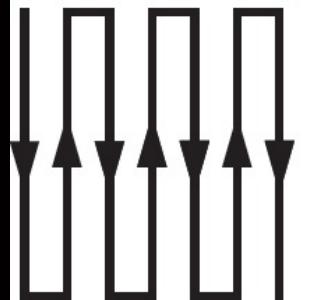
Peano-Hilbert



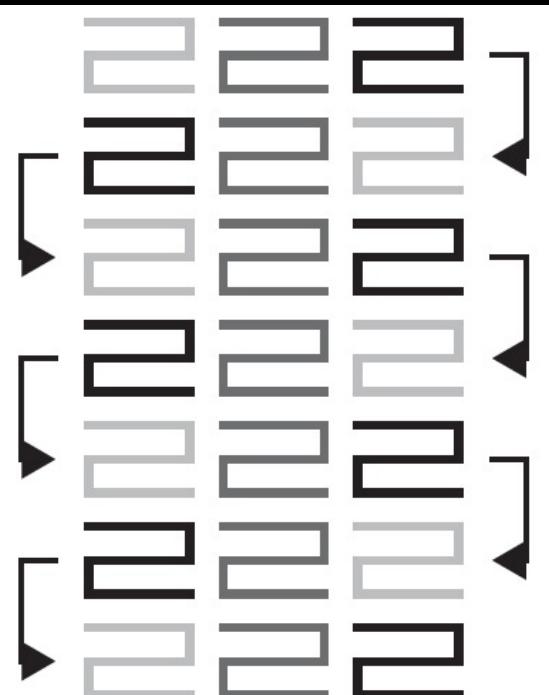
left-right



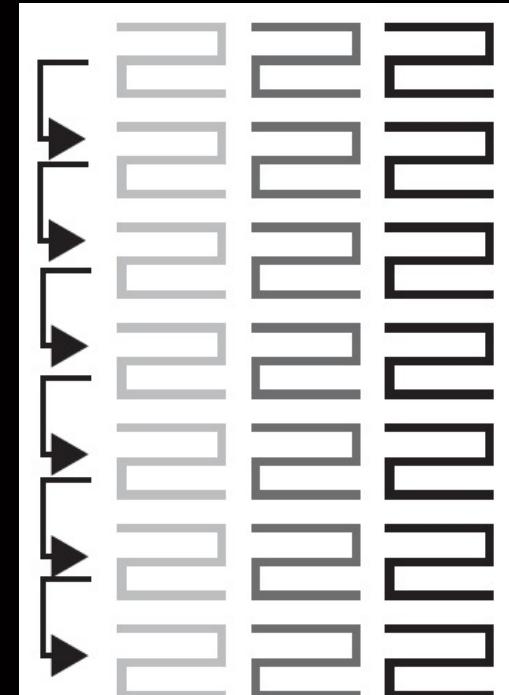
Morton



top-down



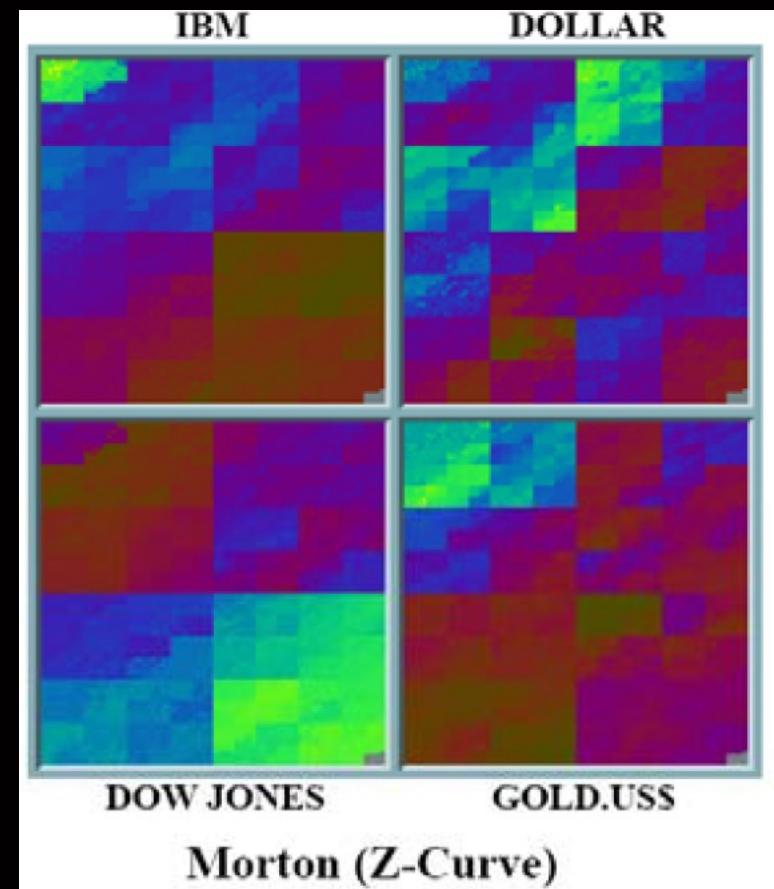
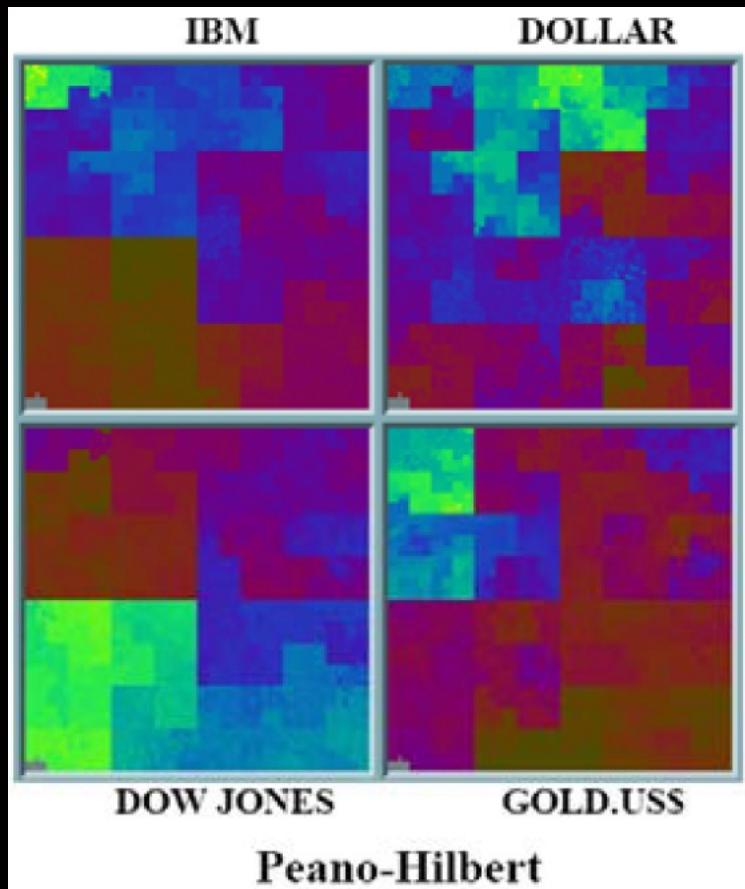
back-and-forth loop



line-by-line loop

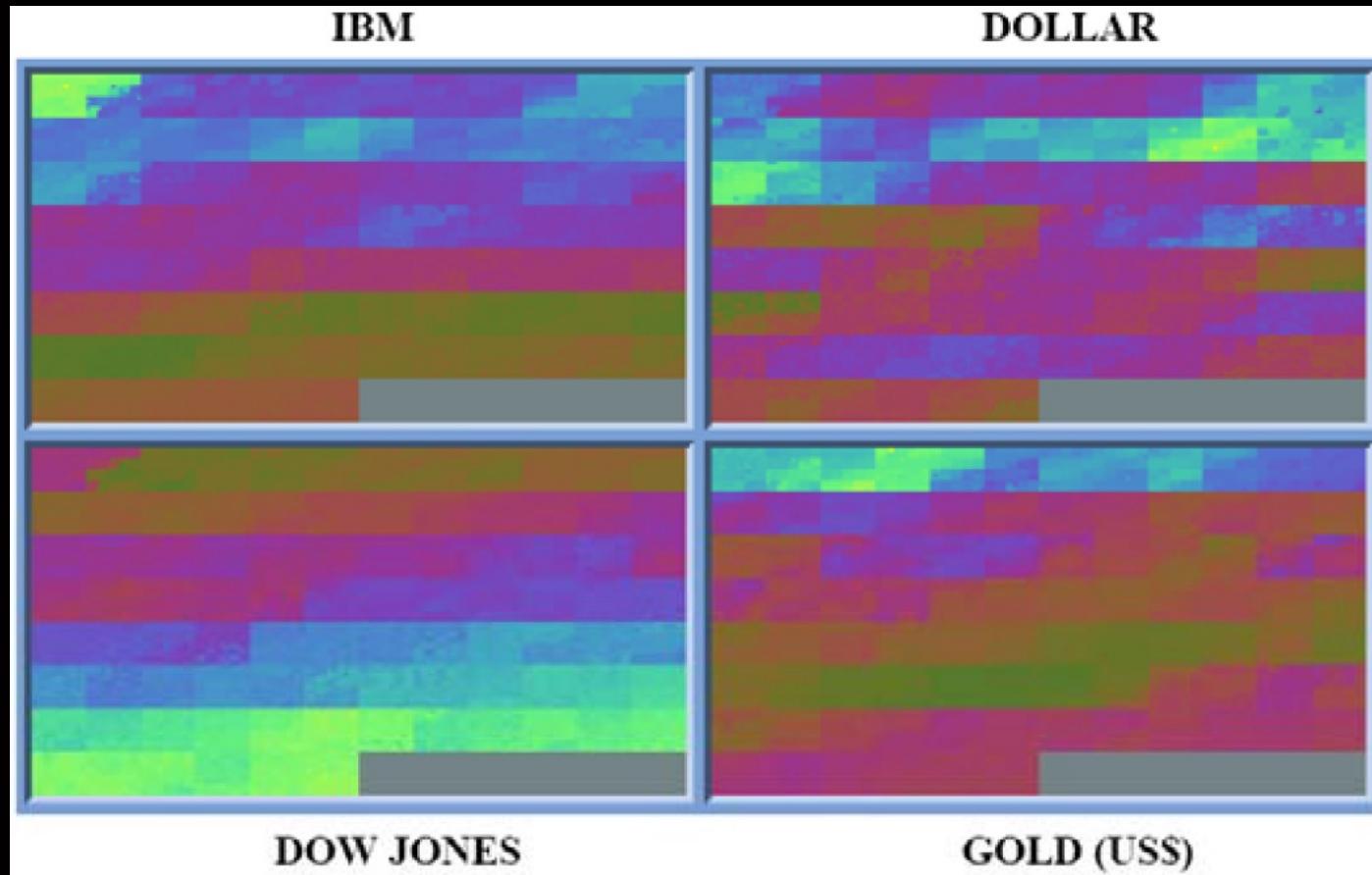
# Datos Multivariados

## Técnicas Basadas en Píxeles



# Datos Multivariados

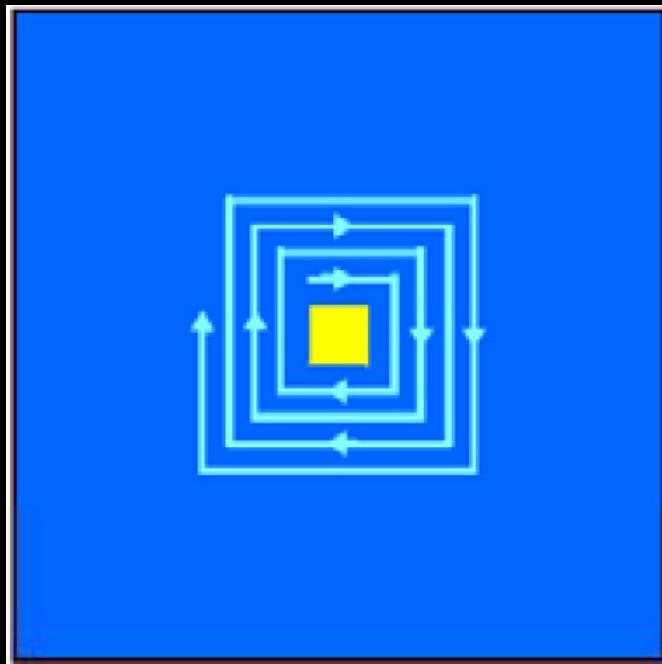
## Técnicas Basadas en Píxeles



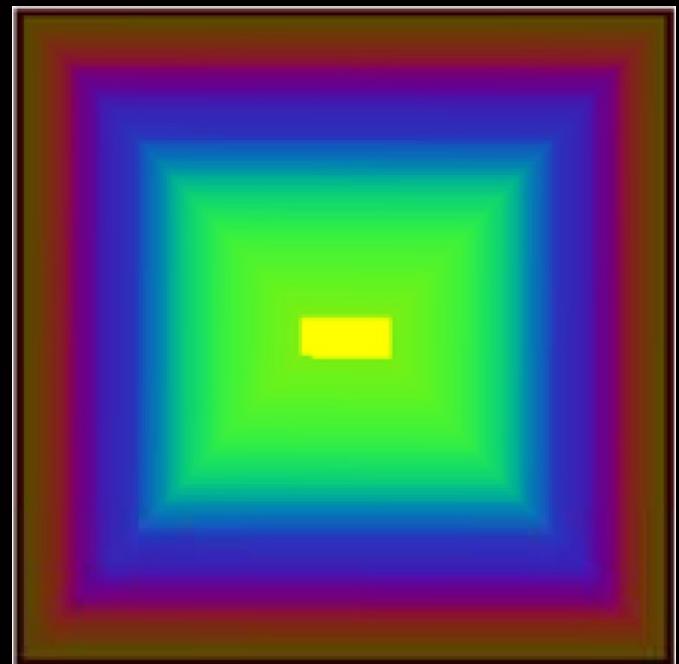
# Datos Multivariados

## Técnicas Basadas en Píxeles

### Técnica de Espiral



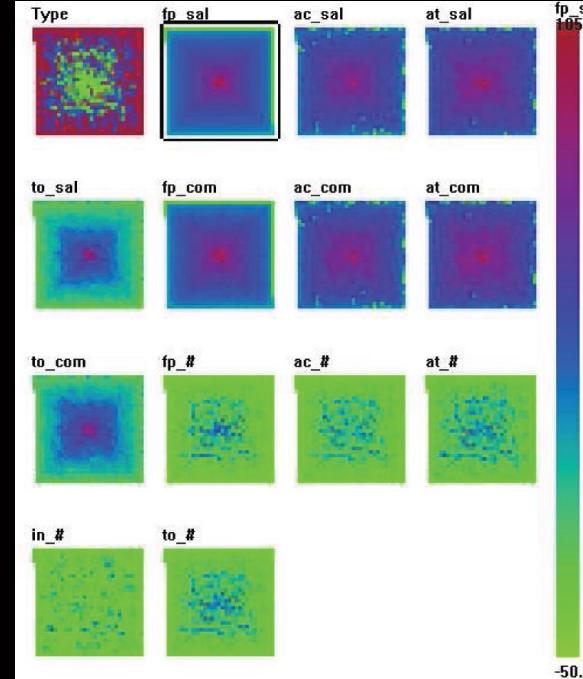
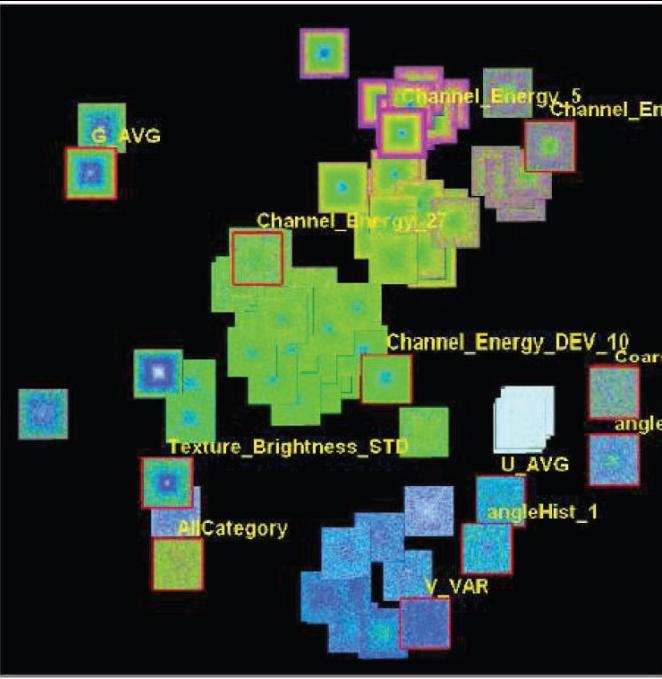
Arreglo



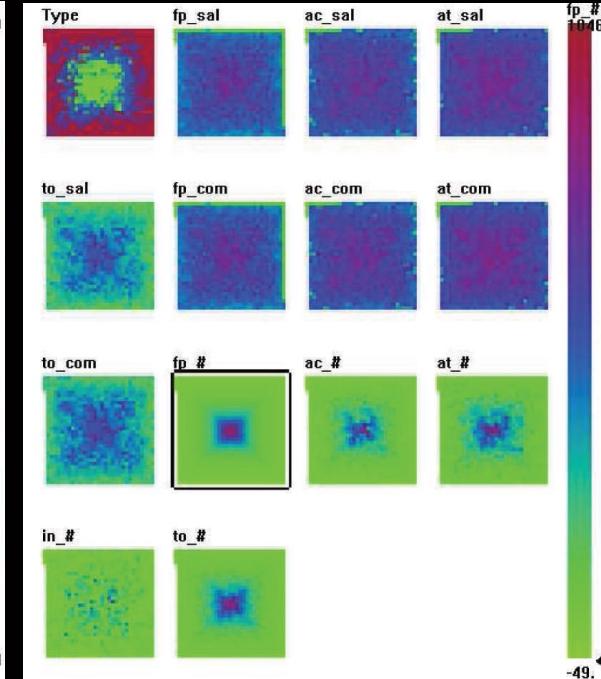
La  $m+1$  dimensión

# Datos Multivariados

## Técnicas Basadas en Píxeles



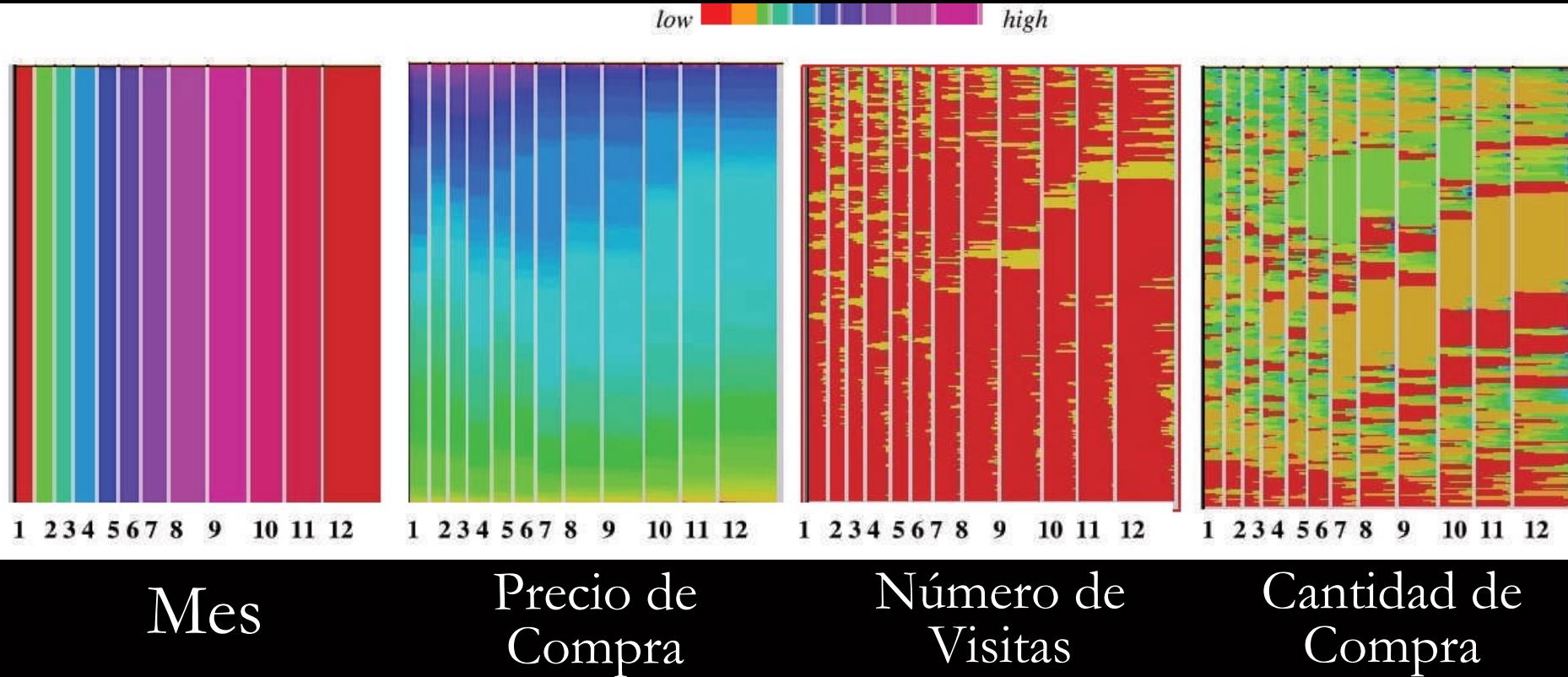
Orden por salario de profesores



Orden por número de profesores

# Datos Multivariados

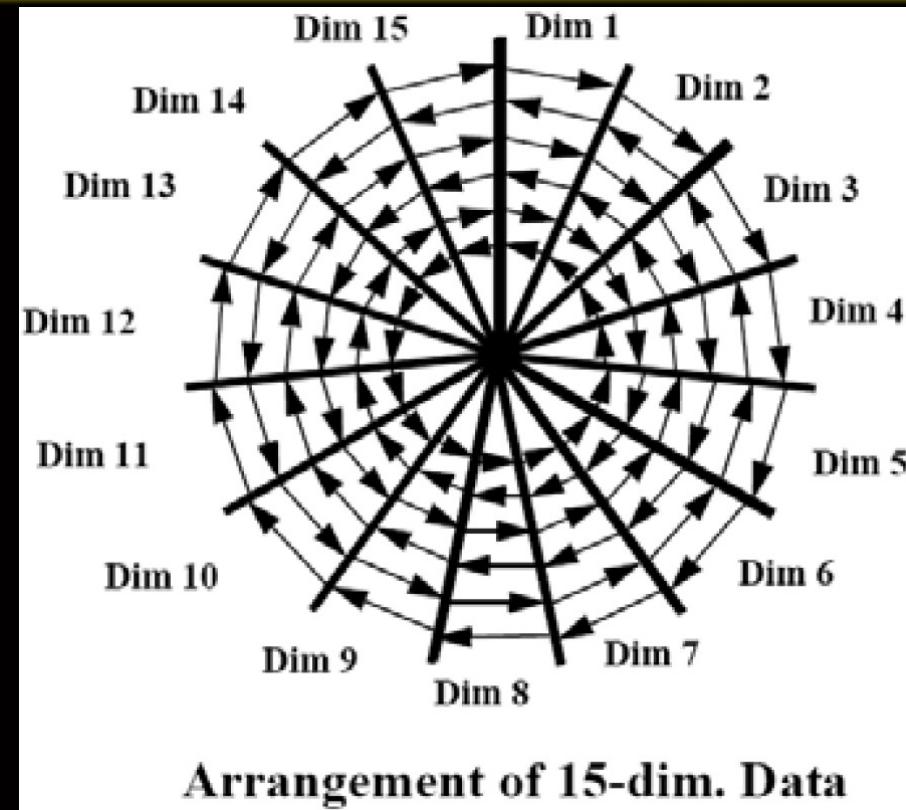
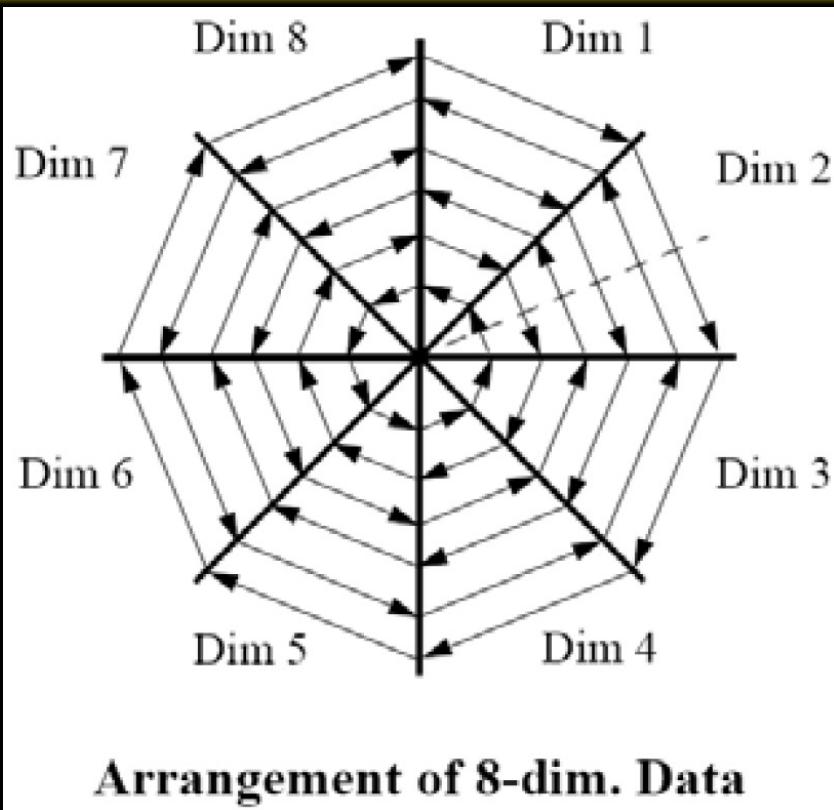
## Técnicas Basadas en Píxeles



Datos de Actividades de Compra de 150,000 e-Clientes

# Datos Multivariados

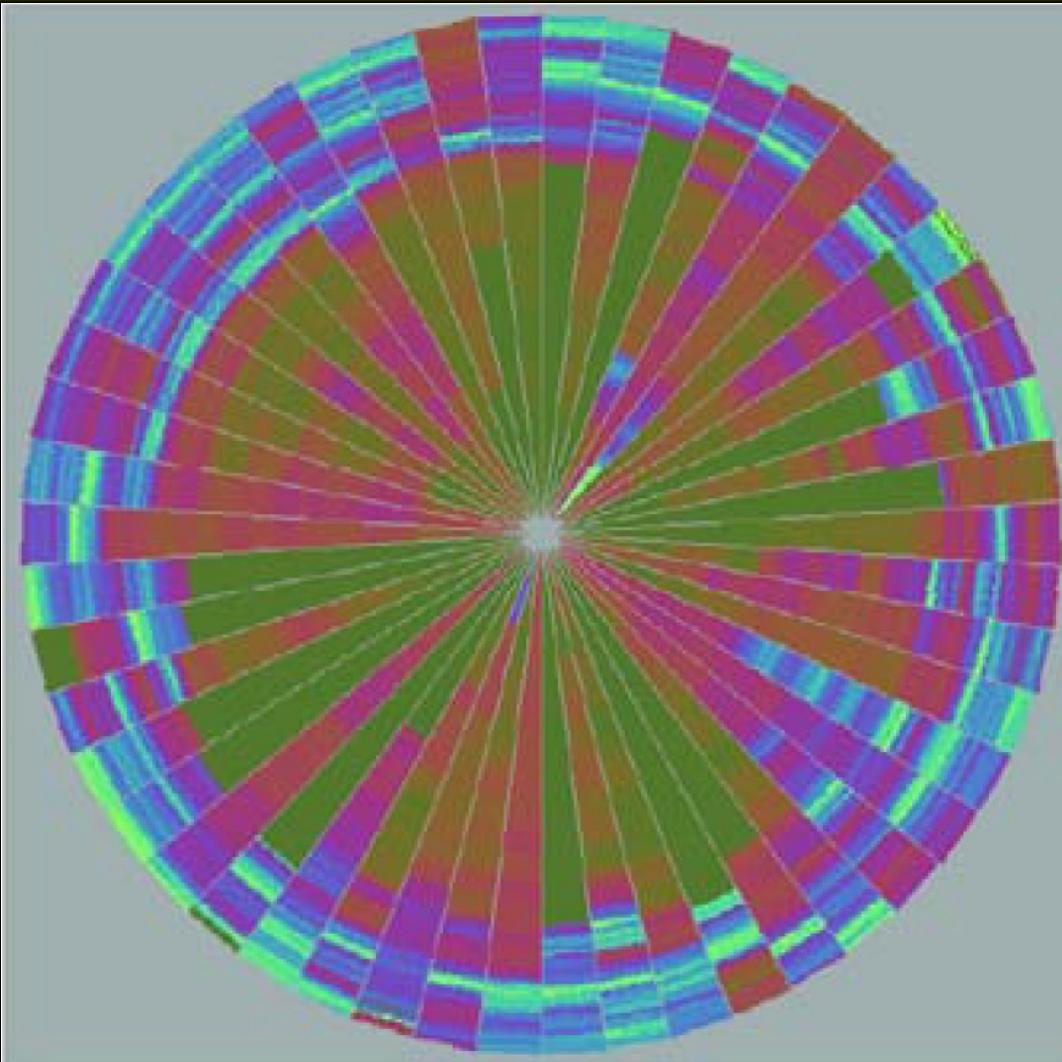
## Técnicas Basadas en Píxeles



Arreglo Circular

# Datos Multivariados

## Técnicas Basadas en Píxeles



Daily stock prices  
of 50 stocks over  
20 years

# **DATOS CON RELACIÓN AL TIEMPO**

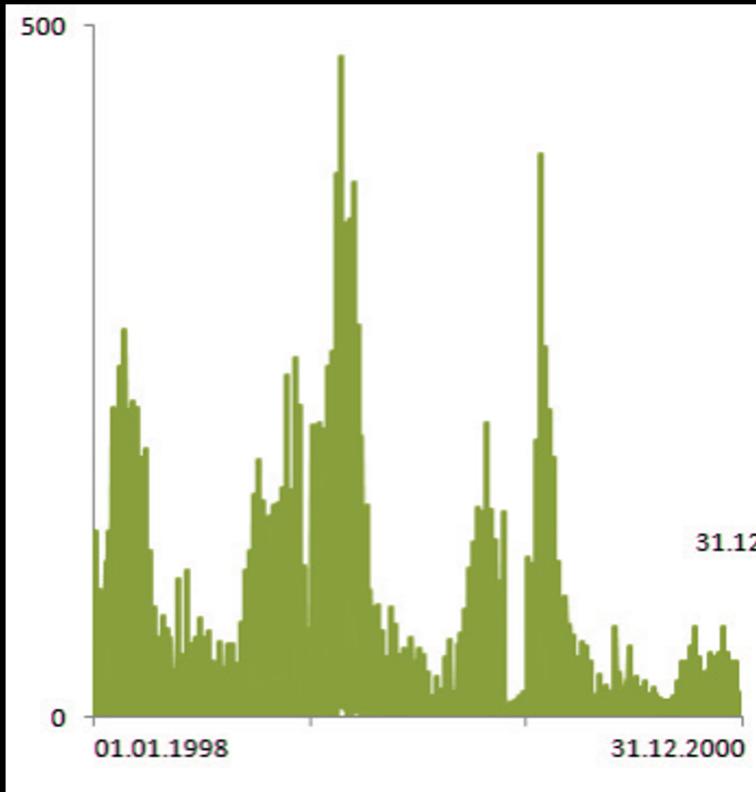
# Datos con Relación al Tiempo

- El tiempo es una dimensión inherente a los datos que es central a las tareas de descubrimiento de tendencias, patrones y relaciones de los datos. Debido a la importancia de los datos con relación al tiempo u orientación al tiempo, se han estudiado en varias publicaciones científicas.
- El siguiente ejemplo demuestra la importancia del tiempo.



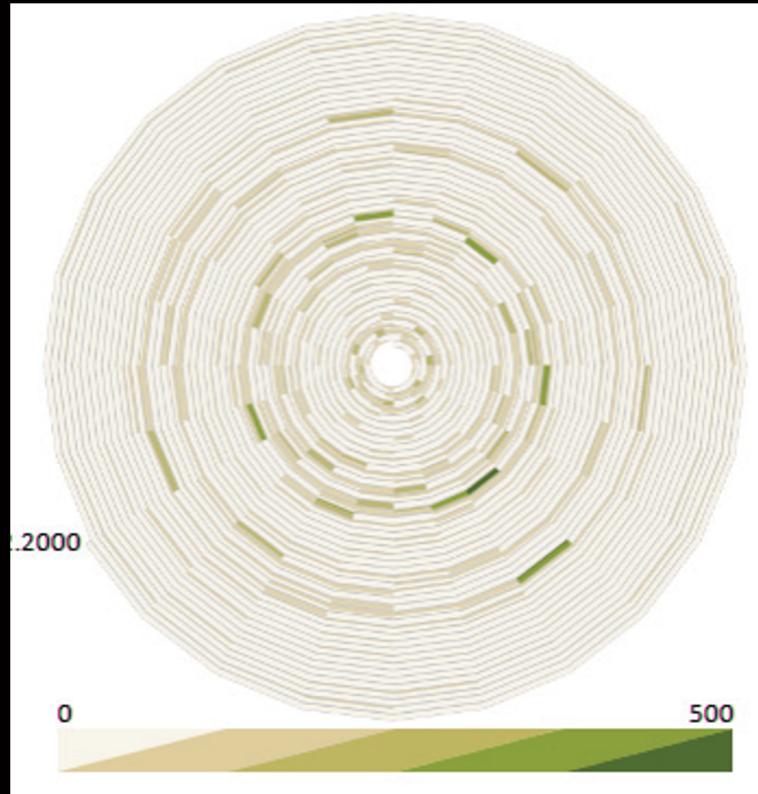
Tres diferentes representaciones del número de casos de influenza diarios en Alemania del norte durante un periodo de tres años.

# Datos con Relación al Tiempo



Un gráfica de línea simple para visualizar los datos. Se identifican bien los máximos en los datos, pero es muy difícil detectar el comportamiento cíclico.

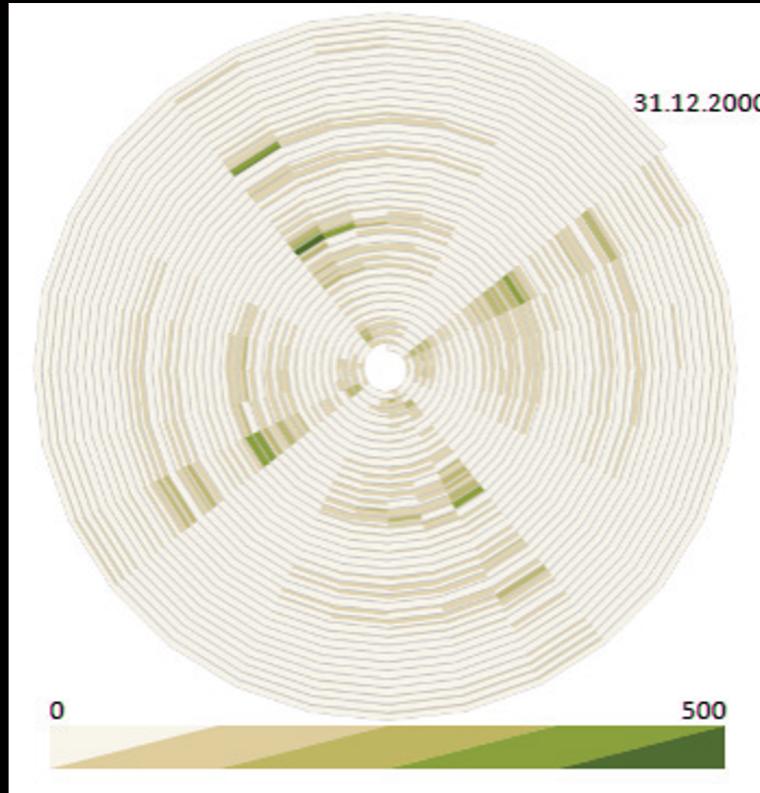
# Datos con Relación al Tiempo



Un gráfica circular para visualizar los datos que enfatiza las características de los datos con relación al tiempo al utilizar un eje de tiempo espiral con longitud de ciclo de 24 días.

# Datos con Relación al Tiempo

Ciclos con múltiplos de 7 también mostrarían patrones.



Un gráfica circular para visualizar los datos que enfatiza las características de los datos con relación al tiempo al utilizar un eje de tiempo espiral con longitud de ciclo de 28 días.

# Datos con Relación al Tiempo

Por lo tanto, es vital escoger una representación visual que sea apropiada para las características de los datos y parametrizar la representación visual de forma apropiada para ser capaz de detectar patrones ocultos en los datos.

Es importante distinguir entre el tiempo como la dimensión física y el tiempo como modelo en sistemas de la información. Cuando se modela el tiempo en sistemas de la información el objetivo es proveer un modelo que mejor se ajuste para reflejar los fenómenos bajo consideración y que apoye al análisis de la tarea, pero no en imitar la dimensión física. Más aún, NO existe un modelo único, o taxonomía única, del tiempo, por el contrario existen muchas formas para modelar el tiempo en los sistemas de la información y el tiempo se modela diferente para cada aplicación, dependiendo del problema en particular. Para formular la noción de tiempo, se ha realizado mucha investigación en varias áreas de ciencias de la computación (p.ej., IA, simulación en minería de datos, modelado, bases de datos).

# Datos con Relación al Tiempo

## Características del Tiempo

Las características del tiempo se pueden dividir en aspectos generales necesarios para modelar adecuadamente el dominio del tiempo y la organización jerárquica del tiempo y la definición de elementos de tiempo concretos.

Los aspectos generales son

1. **Escala:** En un dominio del tiempo ordinal, se presentan únicamente relaciones de orden relativas del tipo antes, después y durante. En dominios discretos, también se pueden considerar distancias temporales. Los valores de tiempo se pueden mapear a un conjunto de enteros, lo cual permite el modelado cuantitativo de los valores del tiempo (p.ej., distancias temporales cuantificables). Dominios del tiempo discretos se basan en la unidad más pequeña posible (p.ej., segundos o minutos) y son los modelos de tiempo más utilizados en sistemas de la información. Los modelos de tiempo continuos están caracterizados por un mapeo posible a los números reales (o sea, entre dos puntos en el tiempo cualquiera, existe otro punto en el tiempo, algo que se conoce como tiempo denso).

# Datos con Relación al Tiempo

## Características del Tiempo

2. **Alcance:** Los dominios del tiempo con base en puntos se pueden ver de forma similar a los puntos Euclidianos discretos en el espacio. Por lo tanto, no se da información acerca de la región entre dos puntos en el tiempo. Los dominios con base en intervalos se relacionan a subsecciones de tiempo que tienen una extensión mayor a cero (los puntos no tienen dimensión); este aspecto también está relacionado estrechamente a la noción de granulidad. Por ejemplo, el valor de tiempo 17 de mayo del 2021 puede relacionarse al instante 17 de mayo del 2021 en el instante 00:00:00 según el dominio basado en puntos, mientras que el mismo valor puede referirse al intervalo 17 de mayo del 2021 de 00:00:00 a 23:59:99 en el dominio basado en intervalos.

# Datos con Relación al Tiempo

## Características del Tiempo

3. **Disposición:** Lineal vs cíclico. Una forma natural de ver el paso del tiempo es lineal del pasado al futuro (cada valor de tiempo tiene un predecesor y sucesor únicos). Alternativamente, se pueden organizar los datos de manera cíclica para poder ver los valores de datos que son recurrentes (p.ej., temporadas del año); por lo tanto, un valor A está precedido y sucedido al mismo tiempo por otro valor de tiempo B (p.ej., el invierno es posterior al verano, pero también es anterior al verano).

# Datos con Relación al Tiempo

## Características del Tiempo

4. **Punto de vista:** Los dominios ordenados del tiempo consideran cosas que ocurren una después de otra. A un nivel más detallado, se podrían distinguir entre dominios totalmente ordenados y parcialmente ordenados. En un dominio totalmente ordenando, solo ocurre una cosa a la vez. En dominios parcialmente ordenados permiten la ocurrencia de eventos simultáneos o traslapados se permiten (o sea, primitivas de tiempo múltiples en un solo punto o traslape en el tiempo). Una forma más compleja de organización del dominio del tiempo es la derivación del tiempo; en esta organización, múltiples hilos de tiempo se derivan y permiten la descripción y comparación de escenarios alternativos (p.ej., planeación de proyectos). Otra opción es la de perspectivas múltiples que facilita vistas simultaneas del tiempo (aún las que son contrarias). Ejemplos de esta organización lo representan los reportes de testigos visuales que describen la misma situación, cada uno de ellos ligeramente diferente del otro, varias declaraciones de un desastre en diferentes países y zonas de tiempo, o varias ejecuciones de simulaciones estocásticas.

# Datos con Relación al Tiempo

## Características del Tiempo

La organización jerárquica del tiempo y los elementos de tiempo concretos son determinados con base en granularidad, primitivas de tiempo y determinación.

1. **Granularidad y calendarios:** Se pueden utilizar varias abstracciones para lidiar con la complejidad del tiempo y para proveer diferentes niveles de granularidad. Básicamente, la granularidad se puede considerar como una abstracción (de creación humana) del tiempo para hacer más fácil lidiar con el tiempo en la vida diaria (tales como minutos, horas, días). De manera más general, la granularidad describe mapeos desde valores de tiempo a unidades conceptuales más pequeñas o más grandes. En el caso de que el modelo de tiempo soporte una granularidad y un sistema de calendario, se puede caracterizar como granularidad múltiple. Además de esta variante compleja, puede existir únicamente la granularidad sencilla (p.ej., cada valor de tiempo está dado en términos de milisegundos) o ninguna de estas abstracciones son soportadas (p.ej., ticks/marcas abstractos).

# Datos con Relación al Tiempo

## Características del Tiempo

2. **Primitivas de tiempo:** Se pueden considerar como una capa intermediaria entre elementos de los datos y el dominio del tiempo. Básicamente, las primitivas de tiempo se pueden dividir en primitivas ancladas (absolutas) y sin anclar (relativas). Los intervalos e instantes son primitivas absolutas (en otras palabras, están localizadas en una posición fija a lo largo del dominio del tiempo). En contraste, un lapso/periodo es una primitiva relativa (no tiene una posición absoluta en el tiempo). Los instantes son un modelo para puntos únicos en el tiempo (a veces referidos como puntos en tiempo). Los intervalos cubren dos puntos en el tiempo y los lapsos son duraciones (de intervalos) sin una posición fija (p.ej., un mes).

# Datos con Relación al Tiempo

## Características del Tiempo

3. **Determinación:** La incertidumbre es otro aspecto importante cuando se consideran datos con relación al tiempo. Cuando no existe información completa o exacta acerca de las especificaciones de tiempo o cuando las primitivas de tiempo son convertidas de una granularidad a otra, se introduce incertidumbre y se tiene que lidiar con ella. Por ejemplo, el conocimiento inexacto, datos de planeación futura (“toma un par de meses”) o eventos temporales imprecisos (hace uno o dos días). Hay que hacer notar que la relatividad de referencias temporales y la indeterminación temporal son calificativos de sentencias u oraciones y no de los eventos que ellas denotan. La indeterminación se puede introducir al usar una especificación explícita o puede estar presente implícitamente en el caso de granularidades múltiples. Por ejemplo, la afirmación “la actividad A comenzó el 14 de mayo, 2014 y terminó el 16 de mayo, 2014” puede modelarse con granularidad de días por lo que se puede decir que comenzó en el instante 14 de mayo, 2014 y que terminó en el instante 16 de mayo, 2014. Si se modela con granularidad de horas, el intervalo puede comenzar y terminar en cualquier punto en el tiempo entre las 12:00 am y las 12:00 pm de los días. Por lo tanto, es necesario considerar la determinación de la especificación del tiempo dado. Una especificación determinada está presente cuando hay conocimiento completo de todos los aspectos temporales.