



<https://hit.activehalton.co.uk/rethink-your-drink/>

# Visualización con gráficas

# Índice

---

Componentes de la visualización de datos

Tipos de datos

Mapeado

Gráficas

Glifos

Datos multidimensionales

Gráficas apiladas

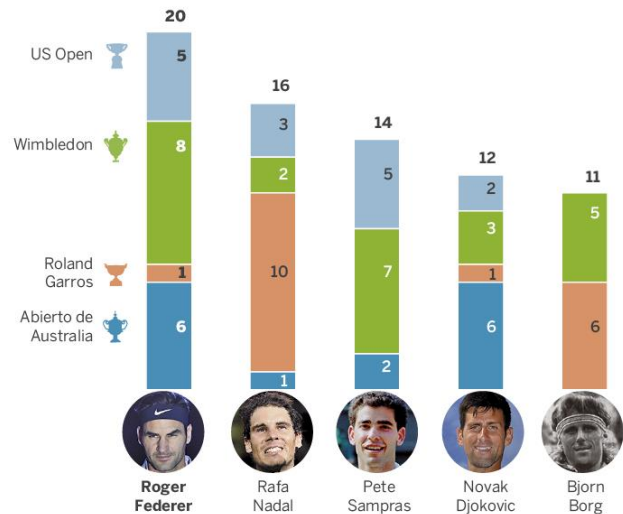
# Introducción

Una gráfica (o gráfico o diagrama) es una representación gráfica de datos

- Los datos se representan mediante símbolos (puntos, líneas, barras, sectores, etc.)

Existen muchos tipos de gráficas, con diferentes características y diseñadas para visualizar distintos tipos de datos

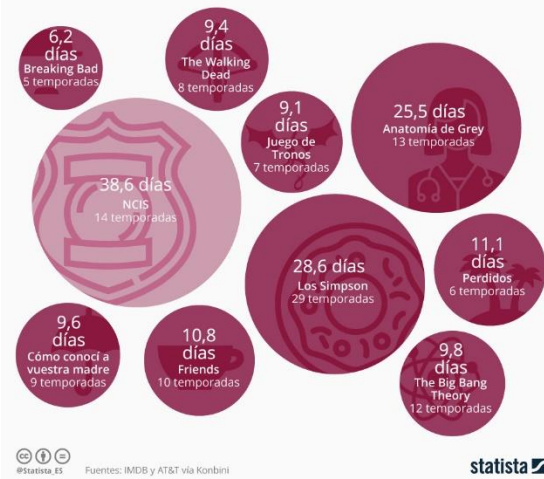
TENISTAS CON MÁS TÍTULOS DE GRAND SLAM



[https://elpais.com/elpais/2018/01/28/media/1517165943\\_254213.html](https://elpais.com/elpais/2018/01/28/media/1517165943_254213.html)

Si ver series fuera (de verdad) un trabajo...

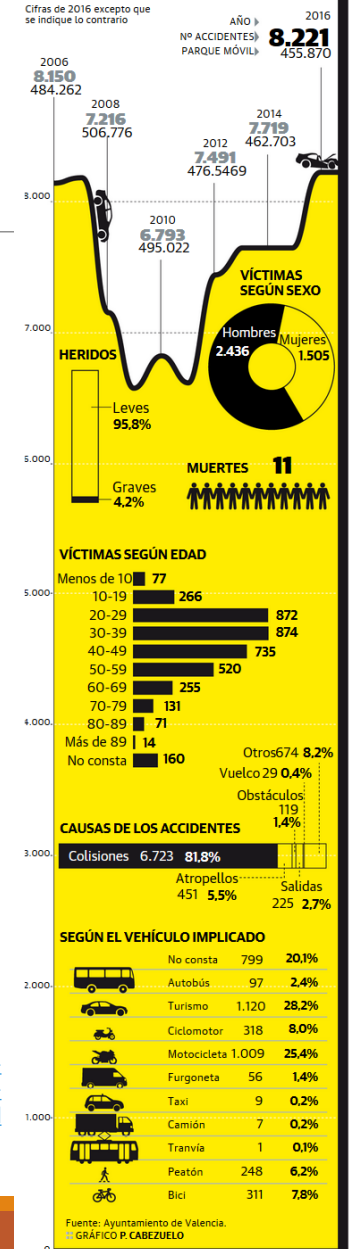
Número de jornadas de trabajo de 8 horas necesarias para ver las siguientes series



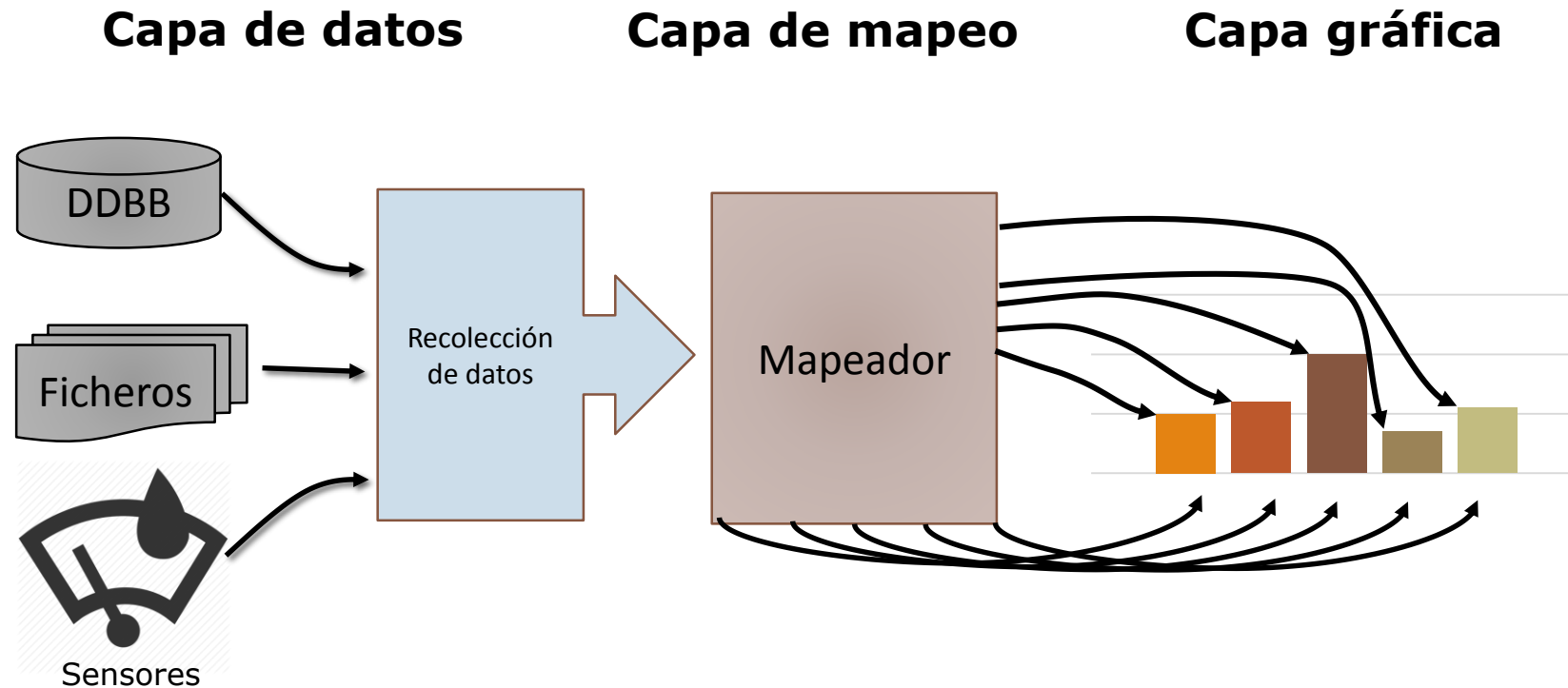
<https://es.statista.com/grafico/12552/las-series-de-mayor-duracion-de-la-tv/>

## Igual que hace diez años

Malas noticias para el tráfico en Valencia que registró en 2016 incluso más accidentes que en 2006



# Componentes de la visualización de datos



# Componentes de la visualización de datos

---

## Capa de datos

- Localizar y obtener los datos
- Importar los datos al formato apropiado
- Vincular los datos que tengan relación
- Análisis de datos y agregación

## Capa de mapeo

- Asociar una geometría adecuada a cada canal de datos
- Análisis de datos y algoritmos (p.e., extracción de contornos)

## Capa gráfica

- Conversión de los objetos geométricos en imágenes visualizables
- Decoraciones
- Gestión de la interacción

# Tipos de datos

---

	Discretos	Continuos
Ordenados (comparables)	<b>Ordinales</b> (tallas: S, M, L, XL...) <b>Cardinales</b> (cantidades)	<b>Campos</b> (altitud, temperatura, presión)
No ordenados (no comparables)	<b>Nominales</b> (formas: □ ○ ◇) <b>Categorías</b> (nacionalidad)	<b>Valores cíclicos</b> (direcciones, colores)

# Datos como variables

---

## Variable independiente

- En un experimento, es la variable que controlamos (tipo de señal de entrada, tipo de comida suministrada, duración de la exposición, temperatura...)
- En una ecuación, se suele representar por la variable  $x$
- En bases de datos, se representa por la clave

$$y = f(x)$$

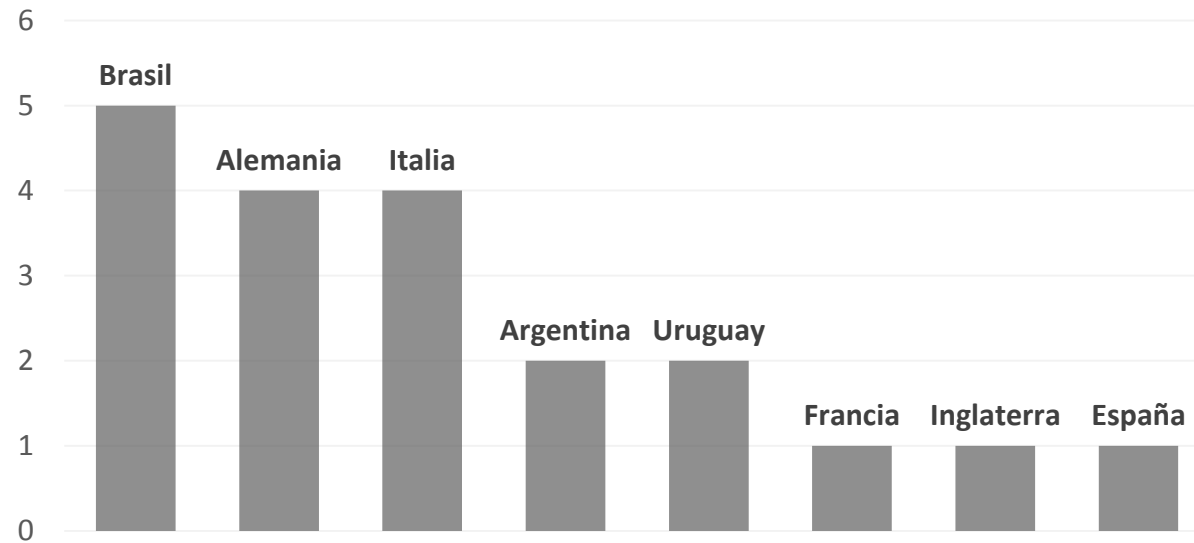
## Variable dependiente

- En un experimento, es la variable que medimos, y cuyo valor cambia dependiendo de la variable independiente (por ejemplo, señal de salida, peso de los ratones después de una semana, velocidad de reacción, etc)
- En una ecuación, se suele representar por la variable  $y$
- En bases de datos, representa el valor obtenido por la consulta

# Mapeado

Consiste en convertir los datos en elementos geométricos que se puedan visualizar

Brasil	5
Alemania	4
Italia	4
Argentina	2
Uruguay	2
Francia	1
Inglaterra	1
España	1

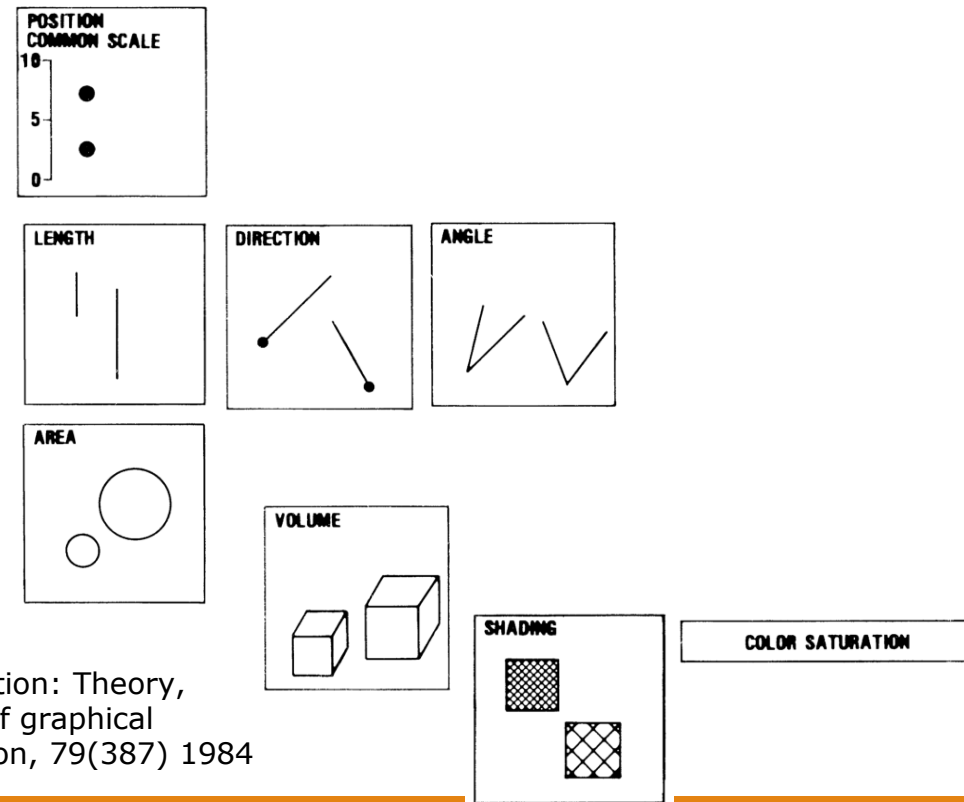




# Mapeado de valores cuantitativos

En los 80 se estudió cómo se percibe mejor valores numéricos dependiendo de su representación gráfica:

- Posición
- Longitud, dirección, ángulo
- Área
- Volumen
- Color, densidad

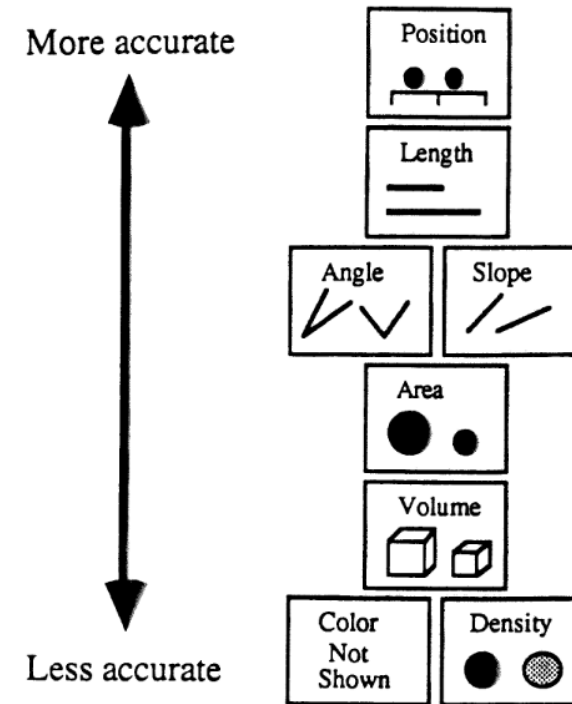


CLEVELAND, W. S., AND MCGILL, R. Graphical perception: Theory, experimentation and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387) 1984

# Mapeado de valores cuantitativos

En los 80 se estudió cómo se percibe mejor valores numéricos dependiendo de su representación gráfica:

- Posición
- Longitud
- Ángulo, Pendiente
- Área
- Volumen
- Densidad, Saturación, Matiz

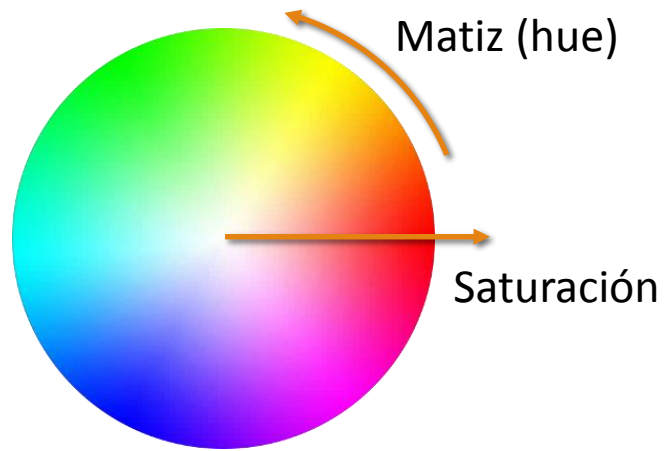


J. Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM Transactions on Graphics 5(2), 1986

# Mapeado de valores cuantitativos

Avance del color

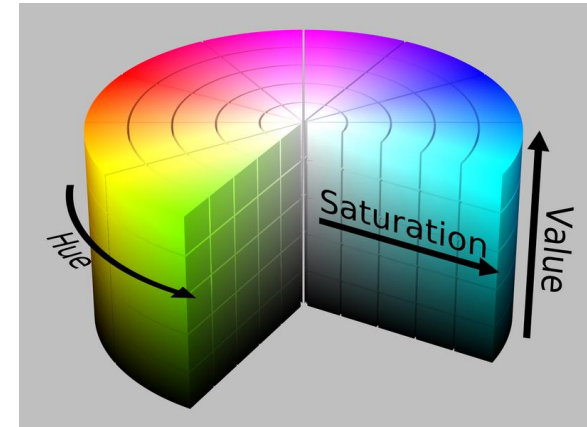
Espacio de color HSV:



Saturación:



Valor (o brillo):



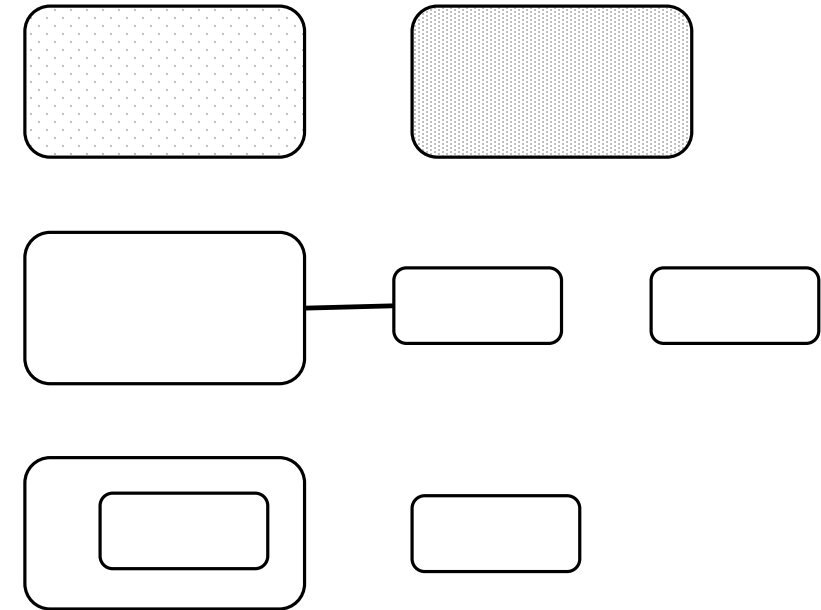
# Mapeado de valores ordinales

## Cuantitativo

Posición  
Longitud  
Ángulo  
Pendiente  
Área  
Volumen  
Densidad  
Saturación de color  
Matiz de color

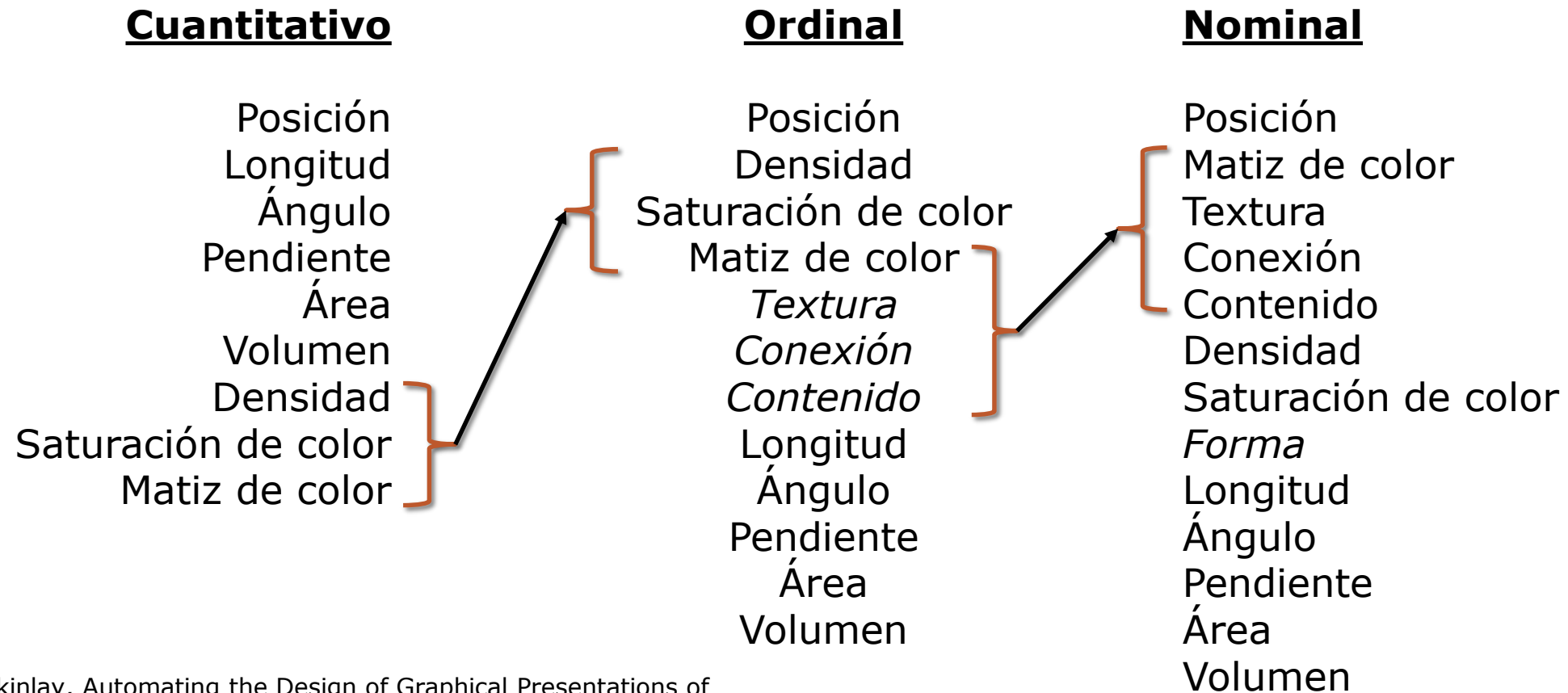
## Ordinal

Posición  
Densidad  
Saturación de color  
Matiz de color  
*Textura*  
*Conexión*  
*Contenido*  
Longitud  
Ángulo  
Pendiente  
Área  
Volumen



J. Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM Transactions on Graphics 5(2), 1986

# Mapeado de valores ordinales



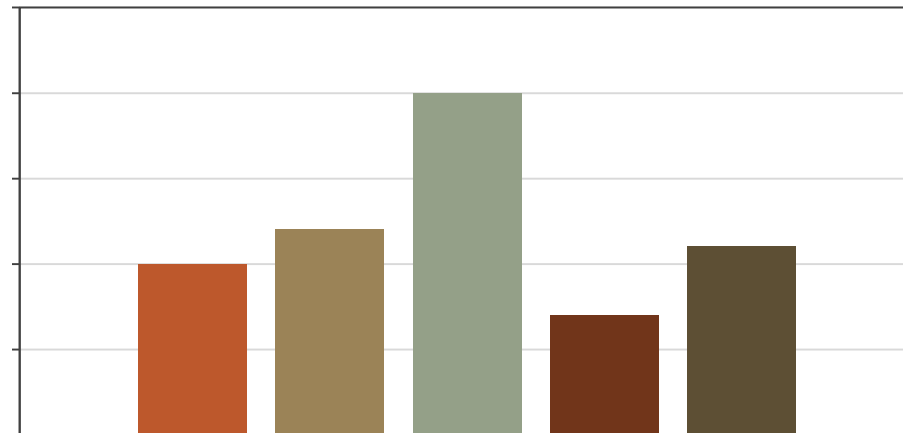
J. Mackinlay, Automating the Design of Graphical Presentations of Relational Information, ACM Transactions on Graphics 5(2), 1986

# Gráficas

## Gráfica de barras

### *Bar Chart*

↑  
Variable  
dependiente  
cuantitativa  
↓



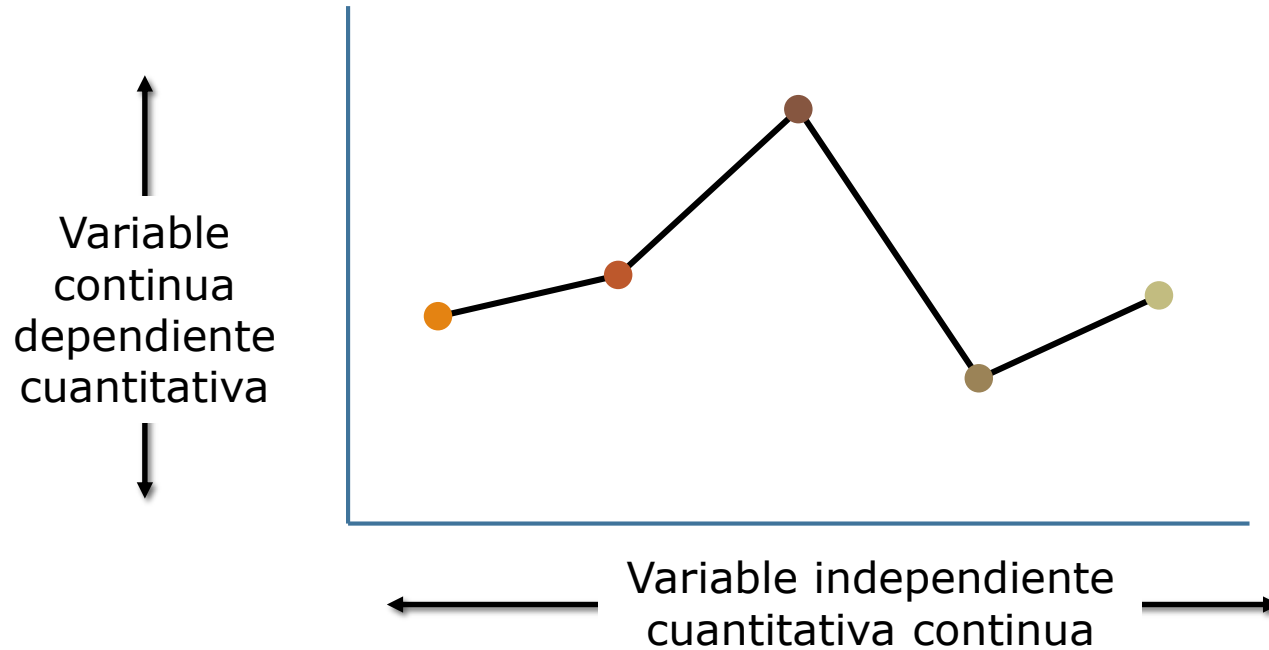
← Variable independiente  
discreta/nominal →

Uso de tanto la  
posición (borde  
superior de la barra)  
como del tamaño  
(longitud)

# Gráficas

## Gráfica de líneas

### *Line Chart*

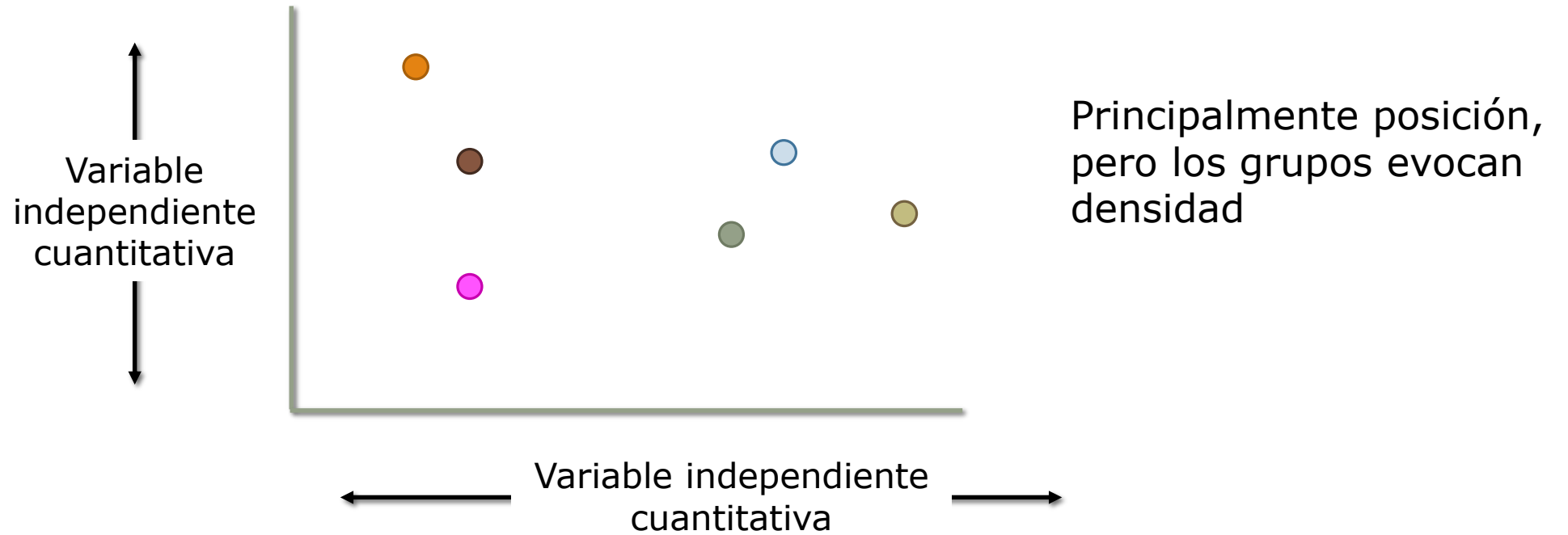


Uso de la posición,  
pero no de la longitud

# Gráficas

## Diagrama de dispersión

*Scatter Plot*

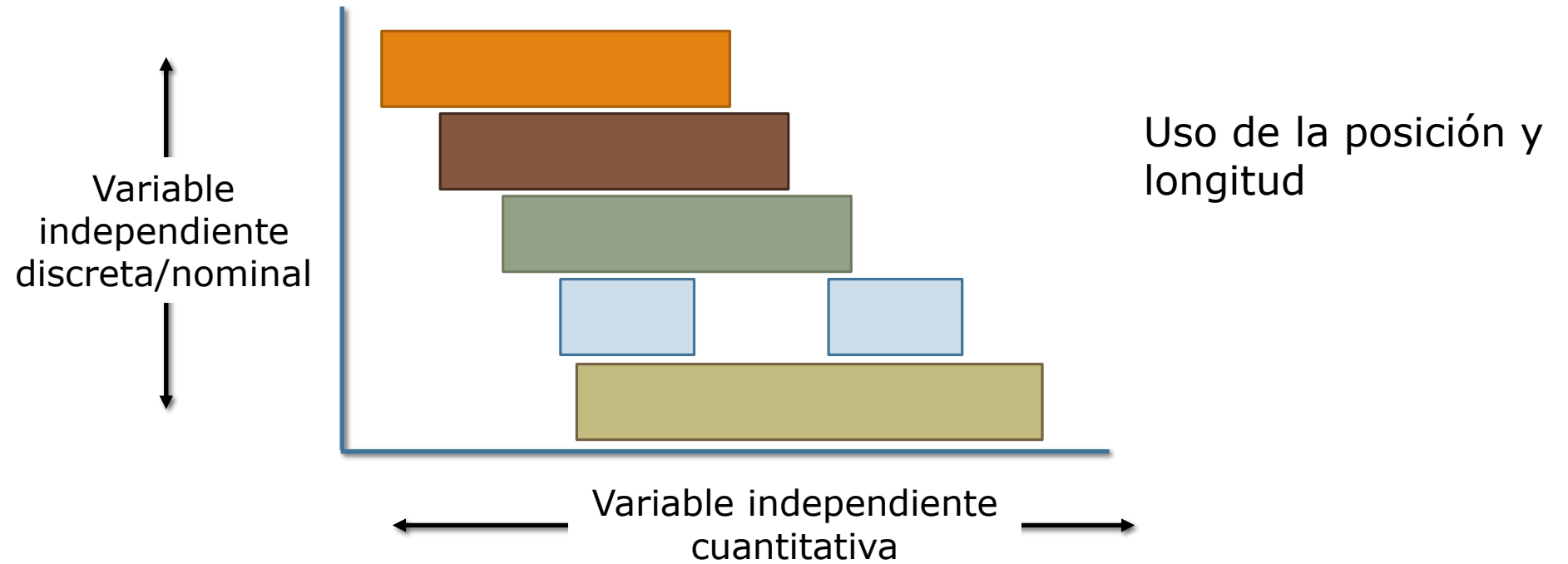




# Gráficas

## Diagrama de Gantt

*Gant Chart*



# Gráficas

## Tablas

---

↑  
Variable  
independiente  
discreta/nominal  
↓


← Variable independiente  
discreta/nominal →

Uso tan sólo de la  
posición

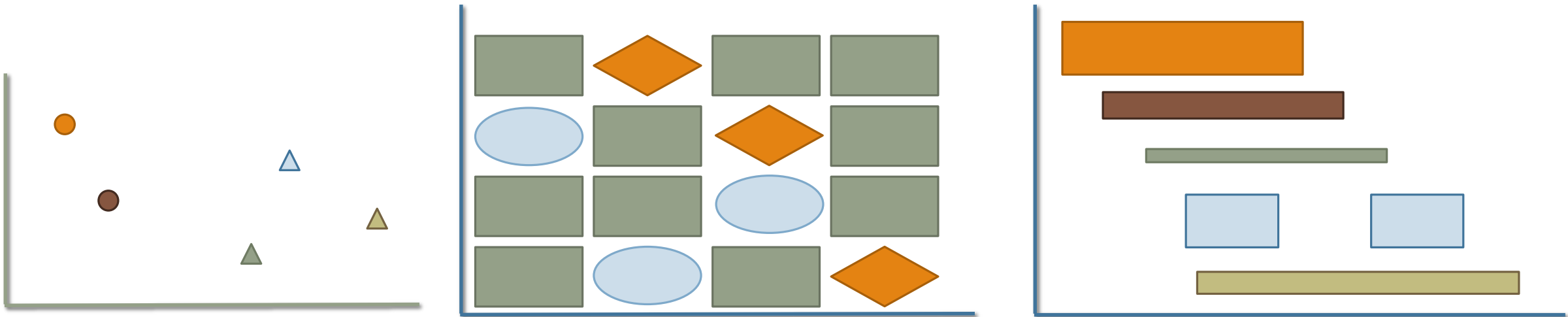
# ¿Qué gráfica usar?

Dependiente	Cuantitativa continua	Barras	Líneas
	Cuantitativa discreta	Barras	Barras
Independiente	Cuantitativa continua	Gantt	Dispersión
	Nominal o cuantitativa discreta	Tabla	Gantt
		Nominal o cuantitativa discreta	Cuantitativa continua
		Independiente	

# Glifos

Representa un dato en una gráfica

- Formas en las gráficas que permiten codificar información extra
- Pueden codificar información con su forma, color, tamaño, orientación, etc.



# Glifos

---

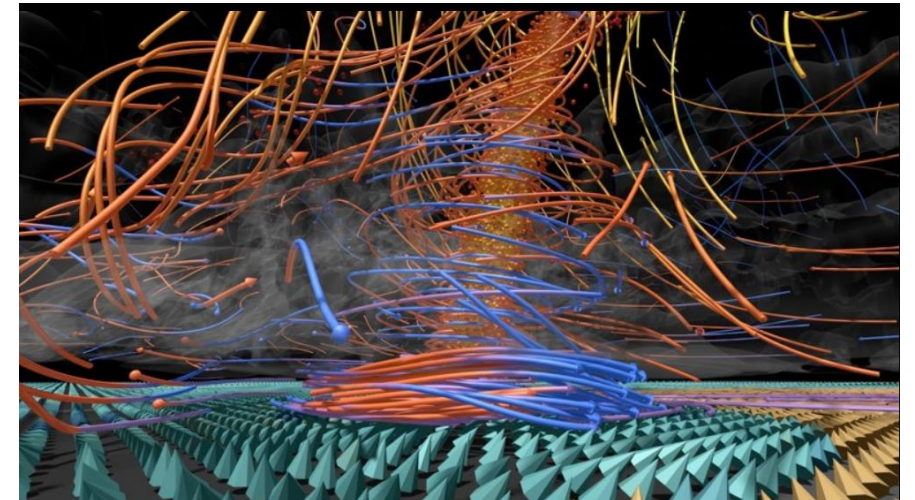
Los conos indican:

- dirección (orientación),
- temperatura (color) y
- velocidad (inclinación)

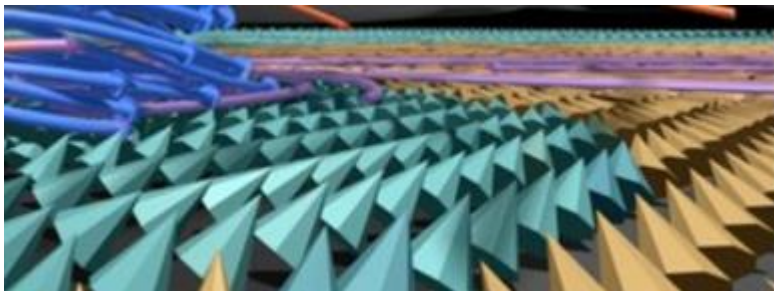
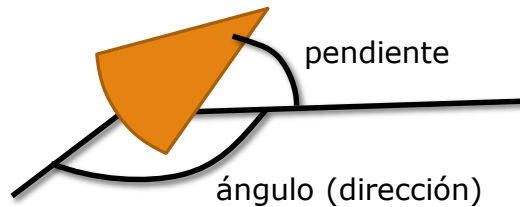
El color de los tubos codifica la dirección:

- Naranja: ascendiente
- Azul: descendiente

Añaden dimensiones adicionales a la visualización de los datos



# Glifos



## Cuantitativo

Posición  
Longitud  
Ángulo  
Pendiente  
Área  
Volumen  
Densidad  
Saturación de color  
Matiz de color

## Ordinal

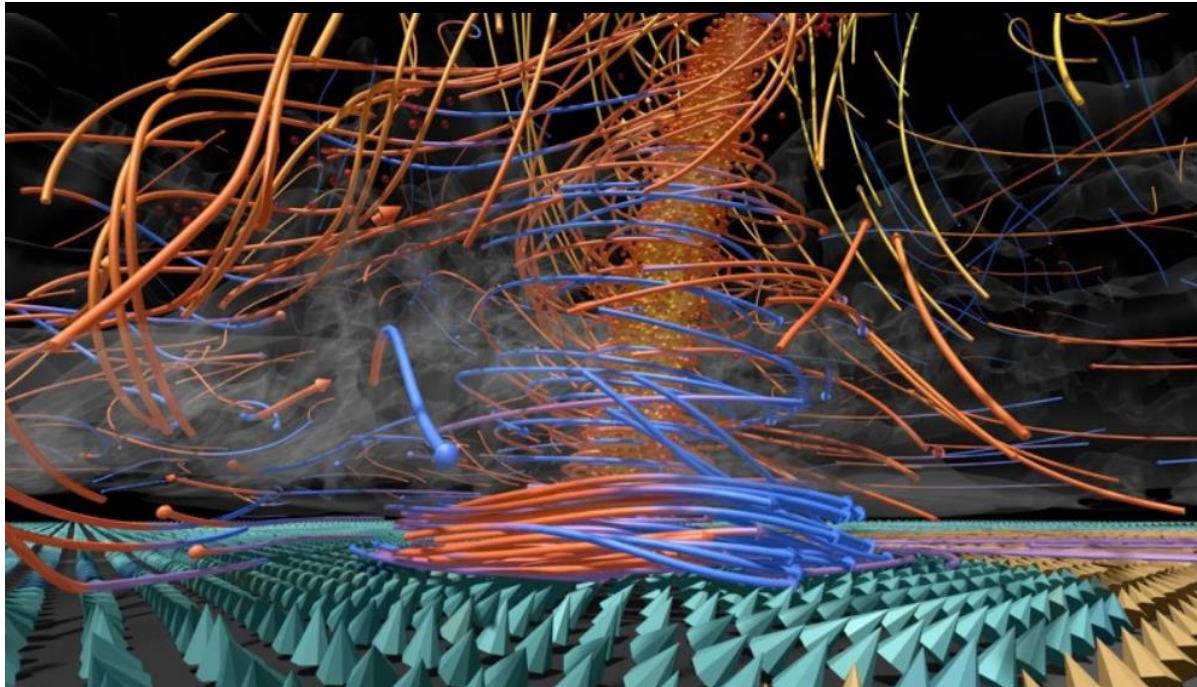
Posición  
Densidad  
Saturación de color  
Matiz de color  
*Textura*  
*Conexión*  
*Contenido*  
Longitud  
Ángulo  
Pendiente  
Área  
Volumen

## Nominal

Posición  
Matiz de color  
Textura  
Conexión  
Contenido  
Densidad  
Saturación de color  
*Forma*  
Longitud  
Ángulo  
Pendiente  
Área  
Volumen

# Glifos

---



## **Nominal**

Posición

Matiz de color

Textura

Conexión

Contenido

Densidad

Saturación de color

*Forma*

Longitud

Ángulo

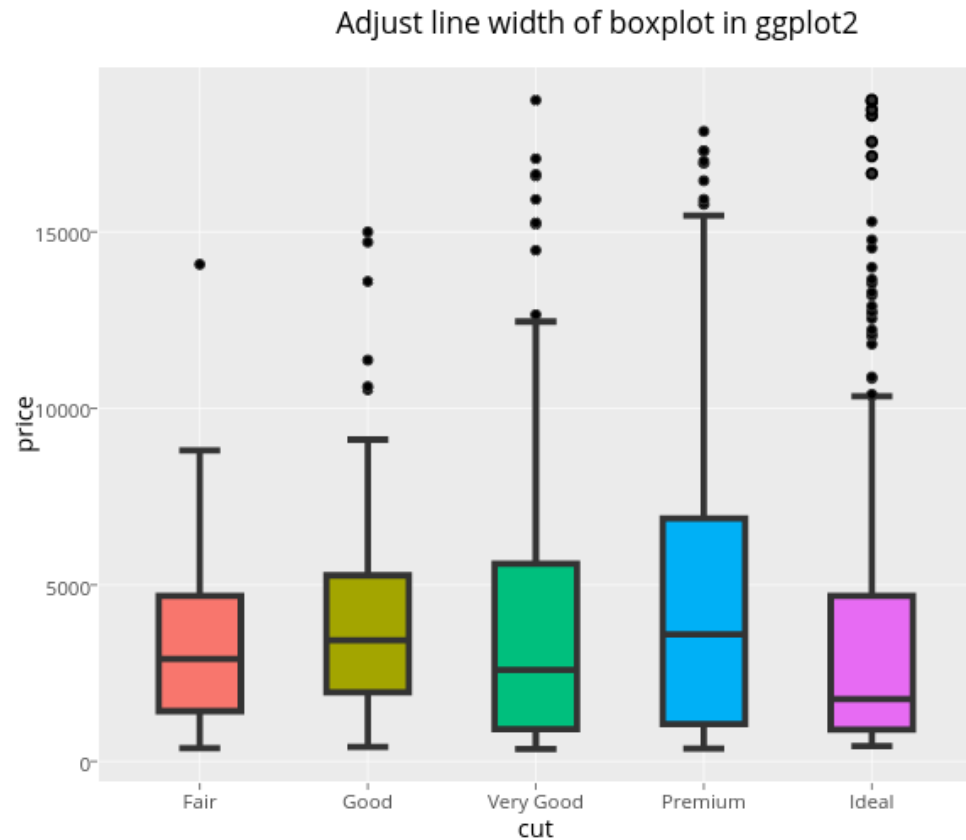
Pendiente

Área

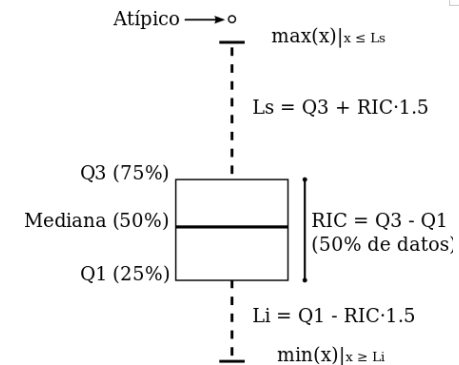
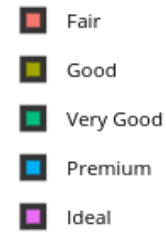
Volumen

# Glifos

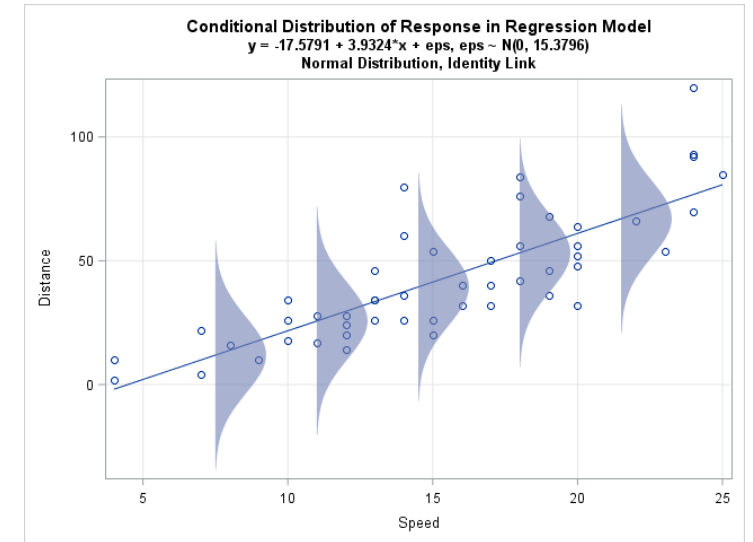
## Representación del error



cut



[https://es.wikipedia.org/wiki/Diagrama\\_de\\_caja](https://es.wikipedia.org/wiki/Diagrama_de_caja)



<https://blogs.sas.com/content/iml/2015/09/16/plot-distrib-exp.html>



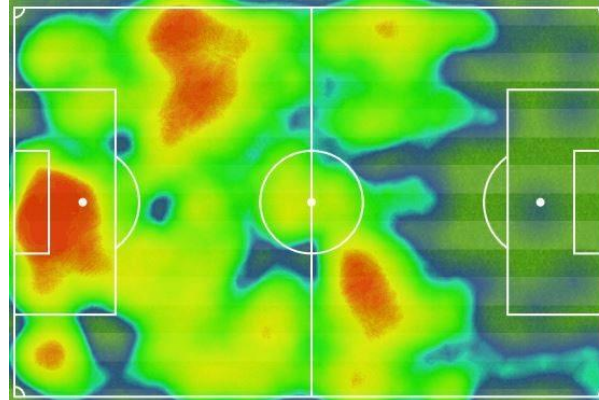
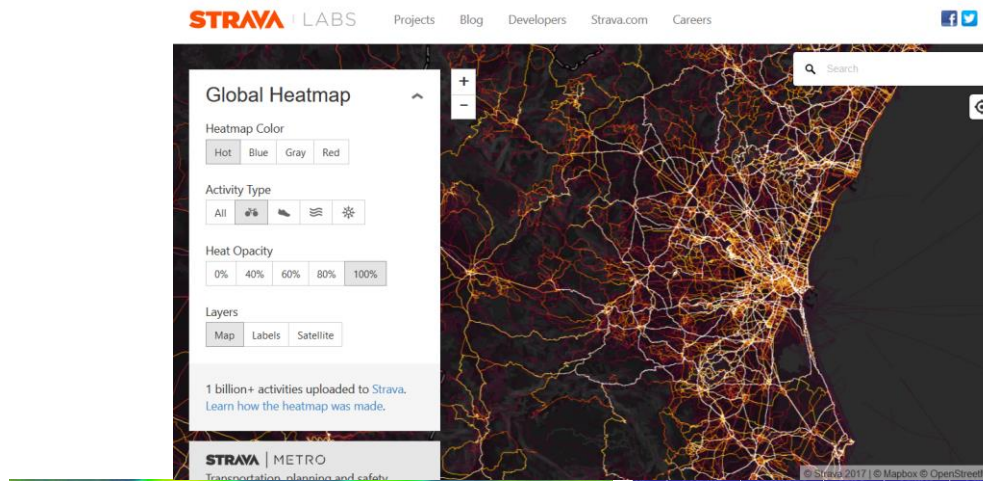
# Glifos

## Mapas de calor

Muestran los valores de una tabla codificados en colores

	Domingo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Total
00:00	6	15	19	15	21	14	12	102
01:00	2	9	9	12	19	10	70	47
02:00	7	4	7	8	5	10	6	47
03:00	3	4	7	5	6	6	37	36
04:00	3	7	2	3	5	8	8	36
05:00	1	3	5	5	3	5	3	25
06:00	3	3	5	2	5	3	2	23
07:00	3	2	2	14	7	10	3	41
08:00	4	9	19	25	28	26	5	116
09:00	4	26	40	46	60	40	4	220
10:00	13	64	48	55	62	49	9	300
11:00	13	48	61	38	54	16	298	372
12:00	7	60	83	68	83	57	14	307
13:00	9	59	55	58	67	50	9	272
14:00	8	42	50	47	65	57	3	285
15:00	8	55	48	63	58	46	7	339
16:00	15	66	64	66	66	50	12	305
17:00	7	65	56	53	74	40	10	268
18:00	12	56	49	52	43	38	18	20
19:00	20	4						11
20:00	11	3						13
21:00	14	2						8
22:00	8	2						2
23:00								
Total general	194	744	768	811	842	680	212	4251

<http://www.ombilbao.com/mapa-de-calor-de-actividad-en-google-analytics-manual/>



## Cuantitativo

Posición  
Longitud  
Ángulo  
Pendiente  
Área  
Volumen

Densidad  
Saturación de color  
Matiz de color

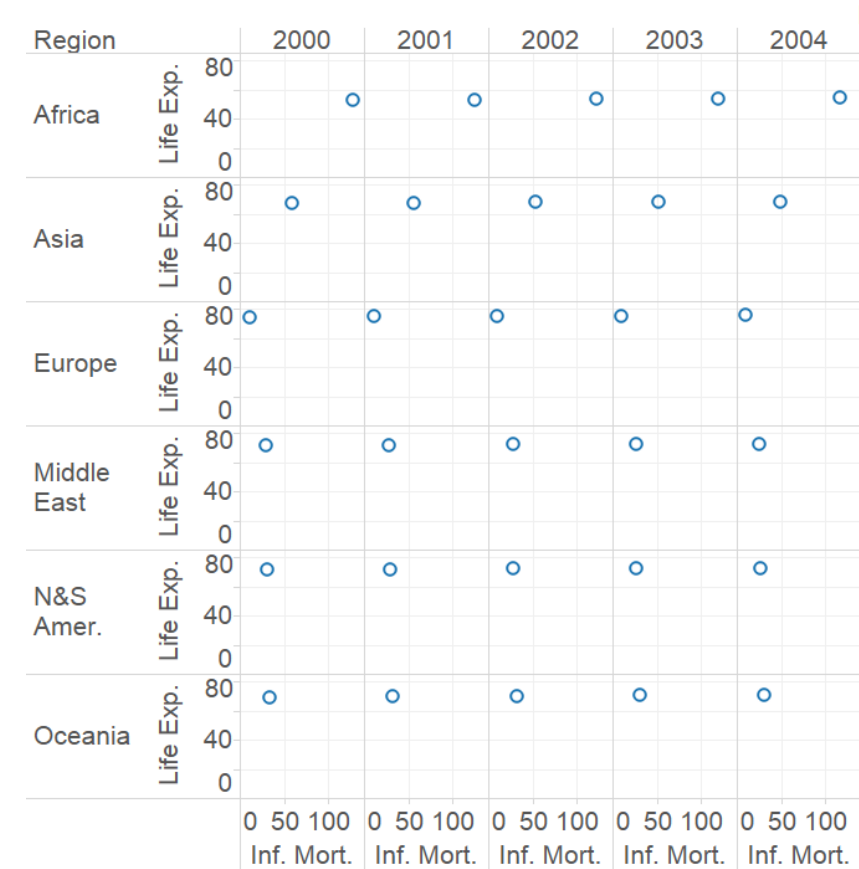
# Glifos

## Gráficas dentro de gráficas

Cada glifo puede ser otra gráfica

P.e., una tabla de tablas

Hay que usar distintas escalas para los ejes mayores y menores tanto para los horizontales como los verticales



Gráfica de Tableau con indicadores del Banco Mundial

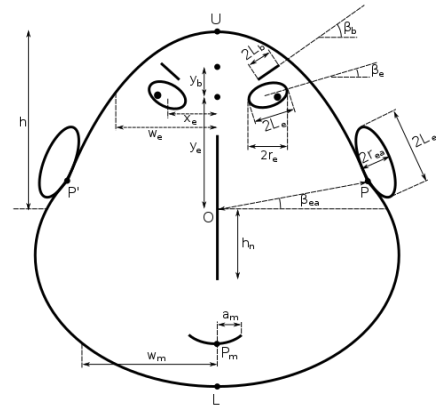
# Glifos

## Caras de Chernoff

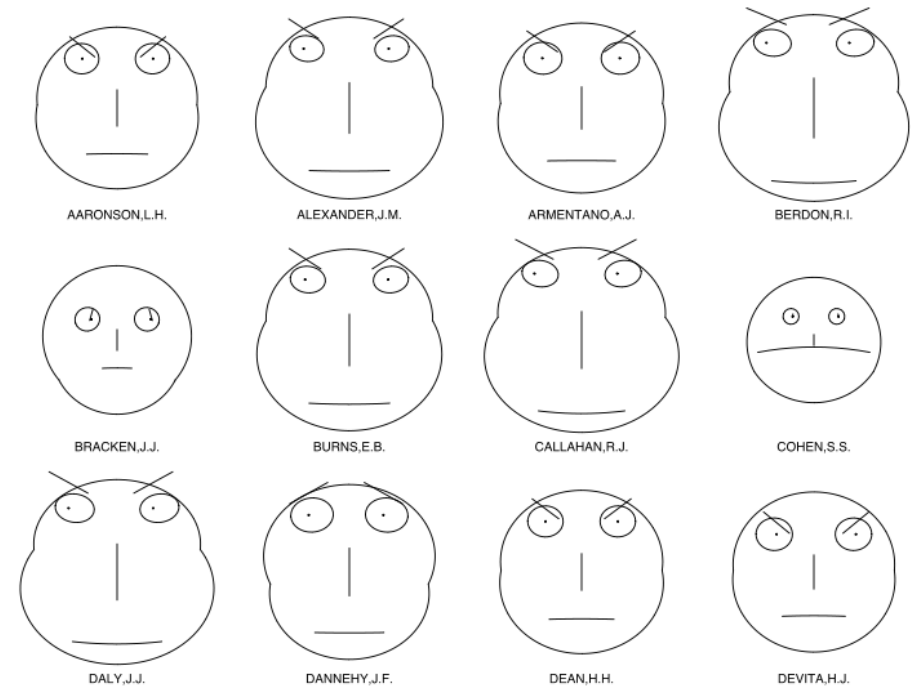
Los glifos son bocetos de una cara humana

Mapean los datos a características (forma, posición, orientación y tamaño) de los elementos faciales (cara, cejas, ojos, nariz, boca...)

Se basa en la capacidad que tenemos las personas de analizar y recordar caras



<https://de.wikipedia.org/wiki/Chernoff-Gesichter>



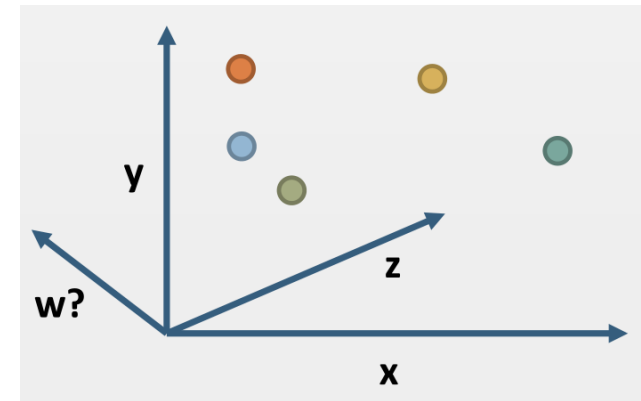
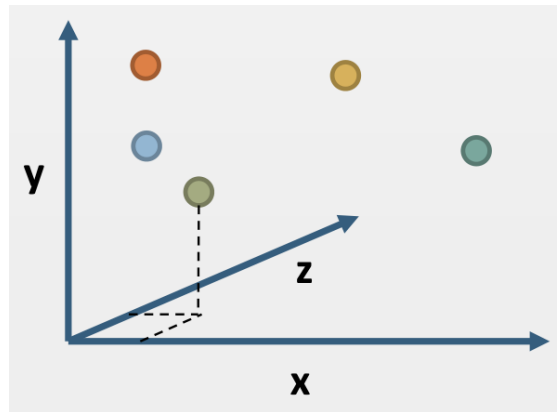
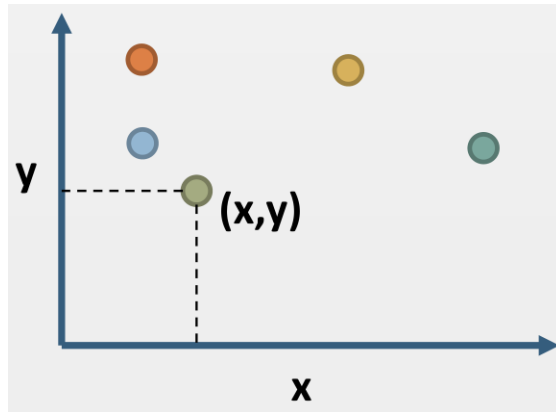
Opinión de abogados sobre 12 jueces de EE.UU., generado en R por el usuario "Avenue" de Wikipedia

# Datos multidimensionales

Estamos acostumbrados a percibir 3 (o 4) dimensiones

Percibir más dimensiones se hace más difícil

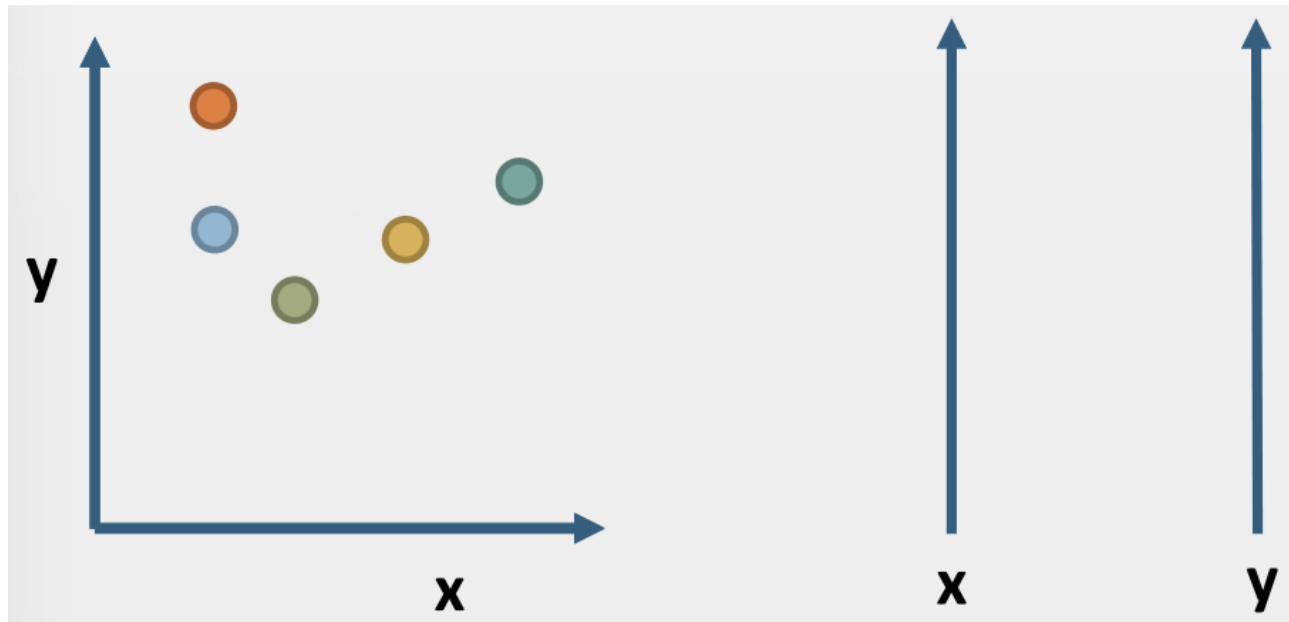
Hay herramientas que nos permiten representar espacios de más de 3 dimensiones



# Datos multidimensionales

## Coordenadas paralelas

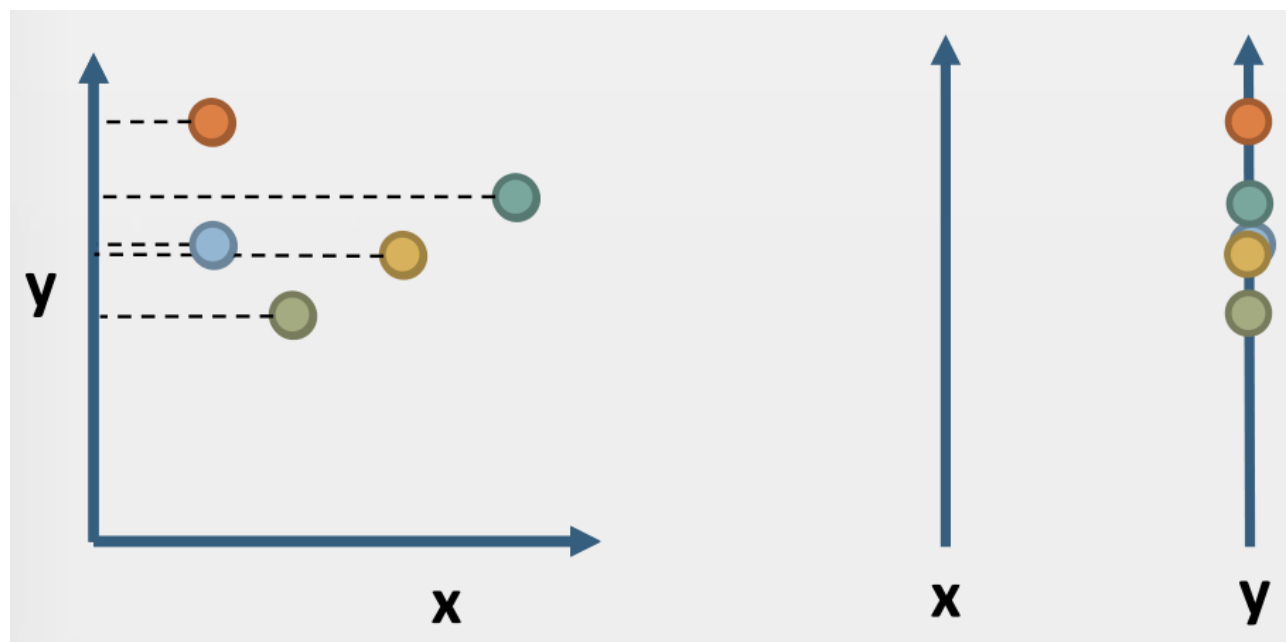
Es una técnica que permite visualizar información sobre datos multidimensionales, y visualizar algunos patrones de los datos



# Datos multidimensionales

Coordenadas paralelas

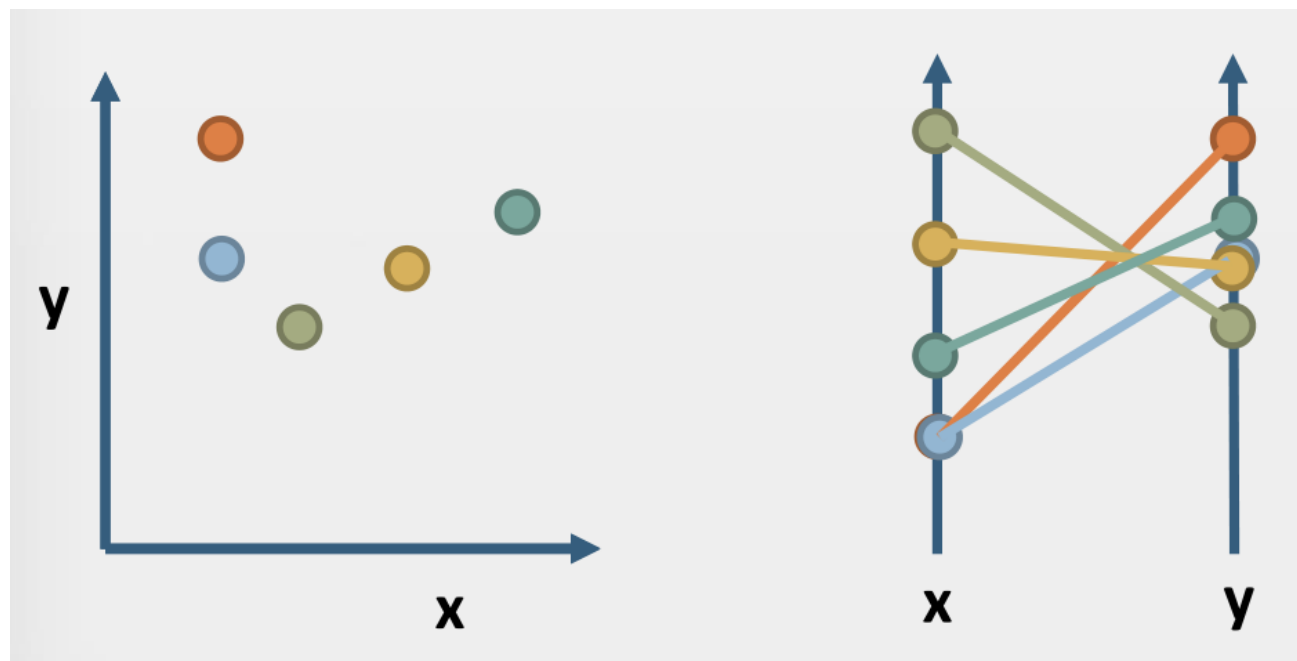
---



# Datos multidimensionales

## Coordenadas paralelas

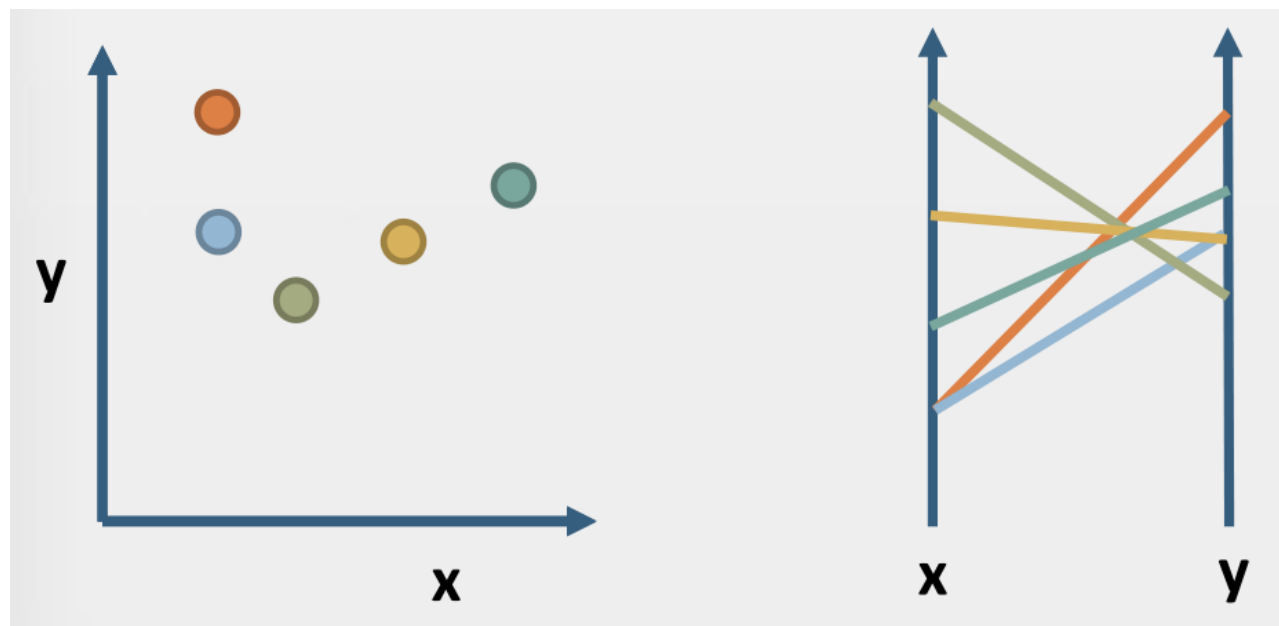
---



# Datos multidimensionales

Coordenadas paralelas

---

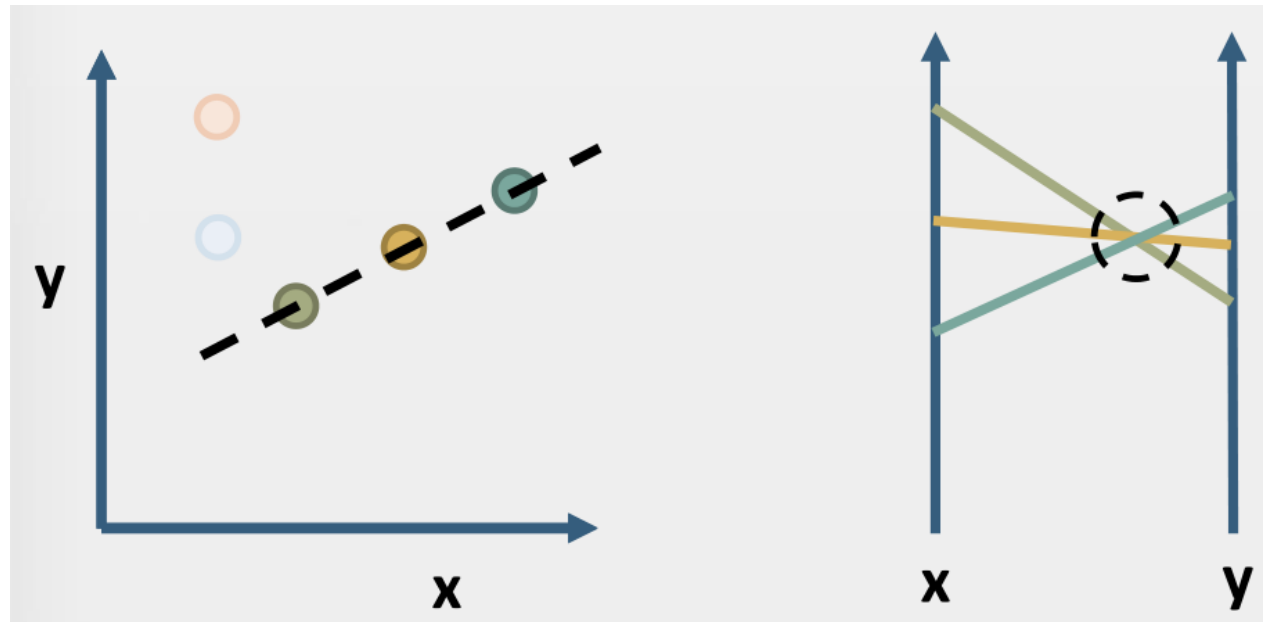




# Datos multidimensionales

## Coordenadas paralelas

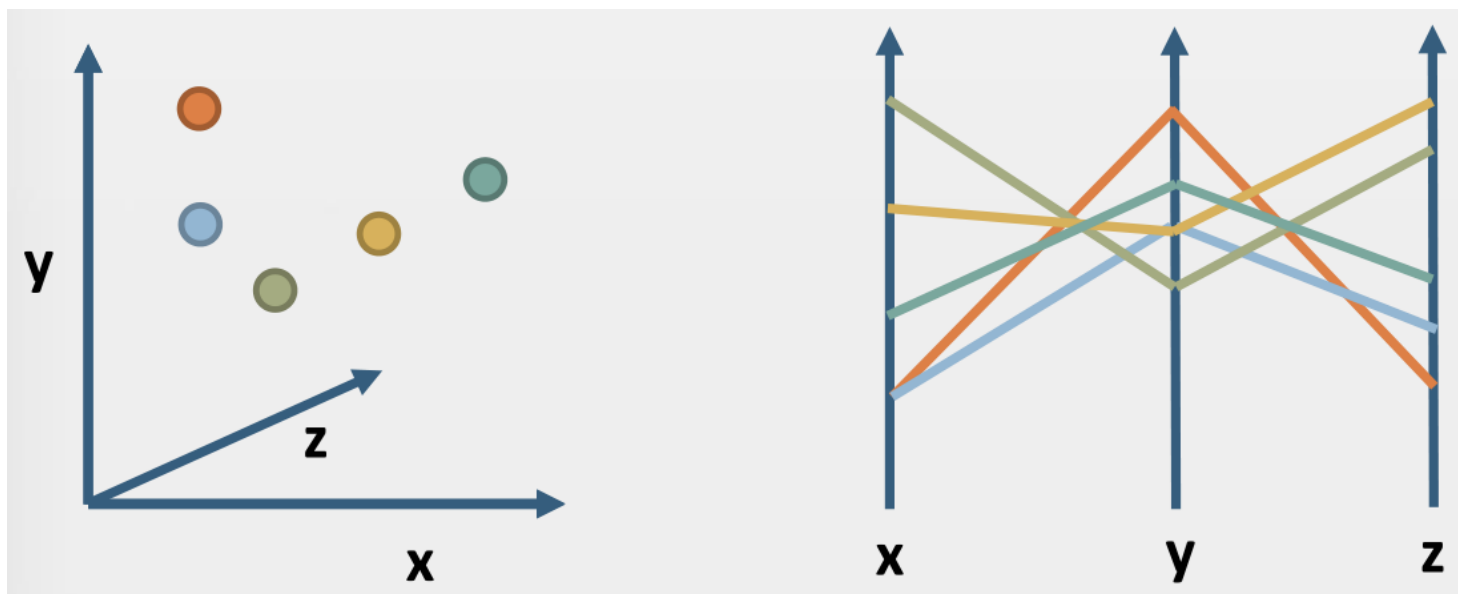
Las gráficas de coordenadas paralelas muestran claramente puntos colineales



# Datos multidimensionales

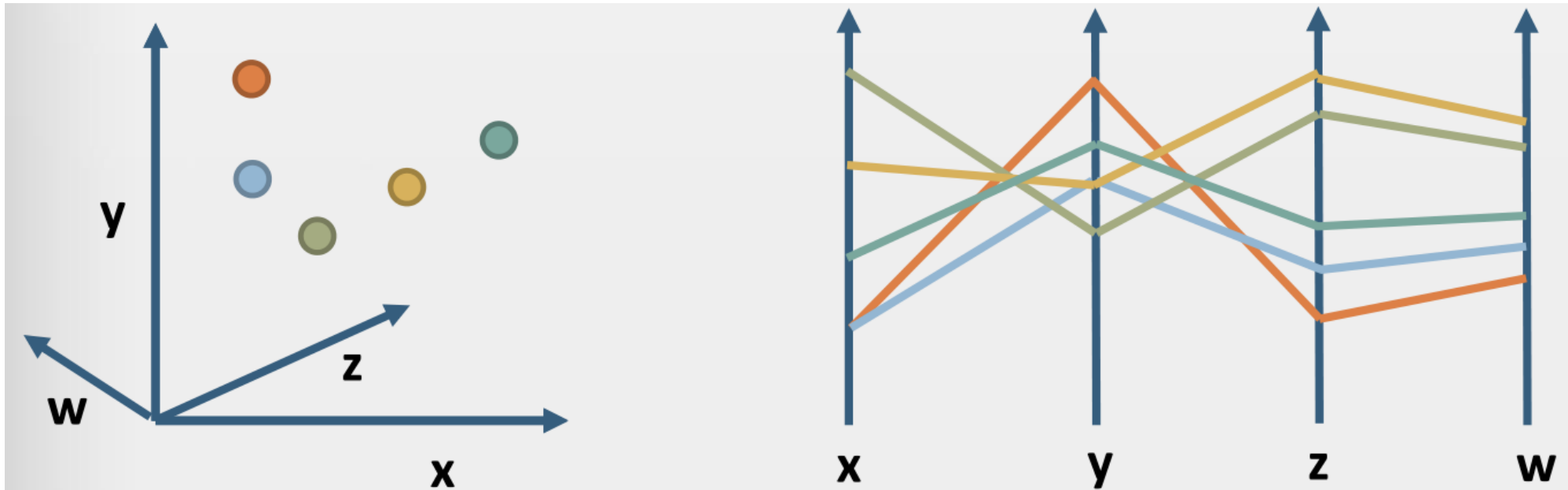
Coordenadas paralelas

---



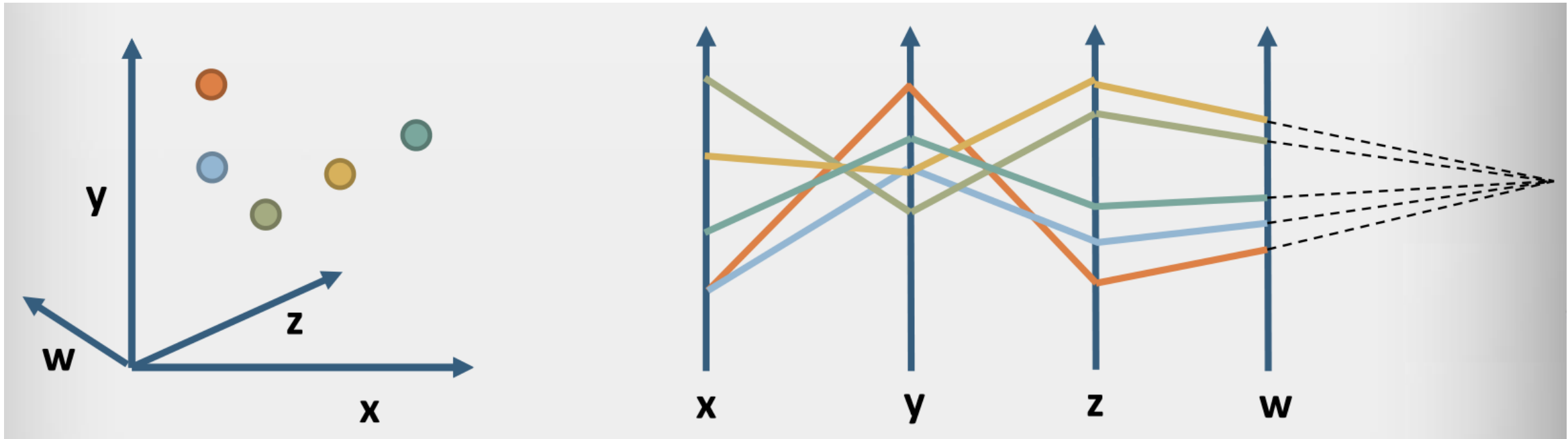
# Datos multidimensionales

Coordenadas paralelas



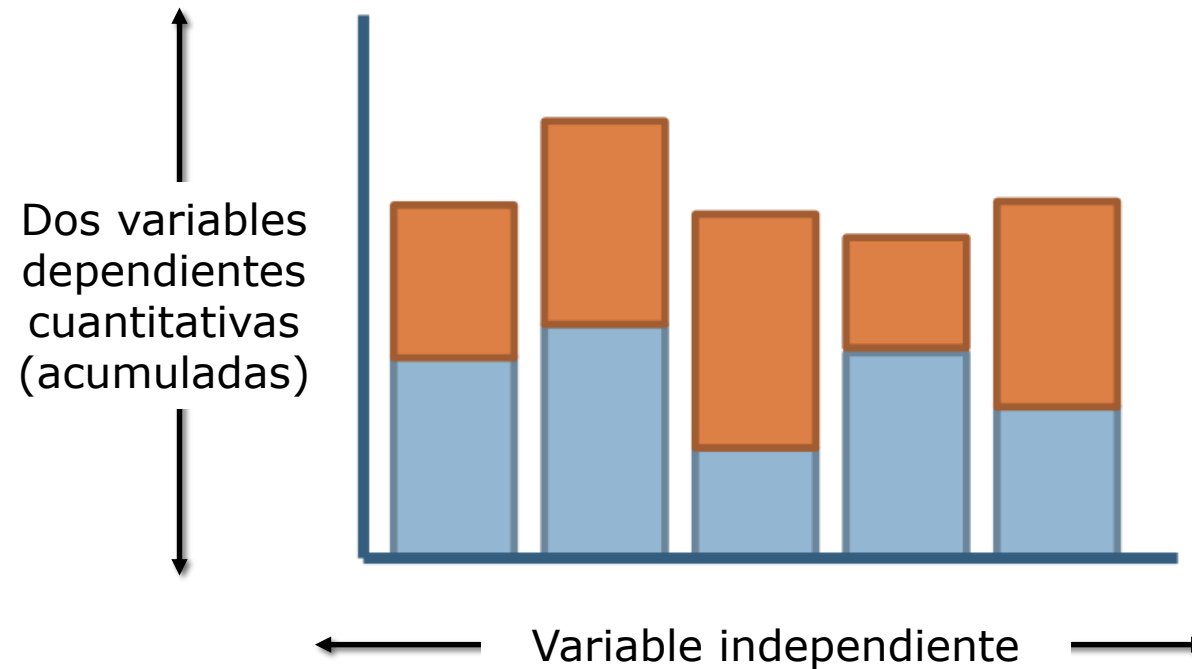
# Datos multidimensionales

Coordenadas paralelas



# Gráficas de barras apiladas

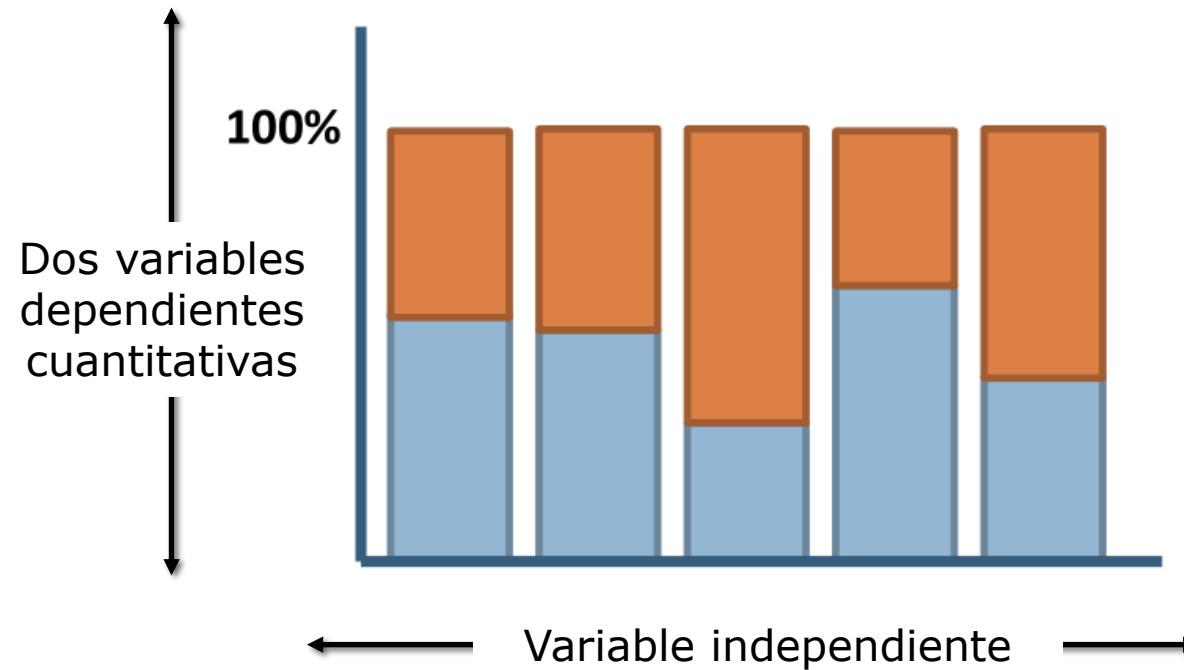
Permiten mostrar el valor de varias variables que varían a lo largo de una misma dimensión



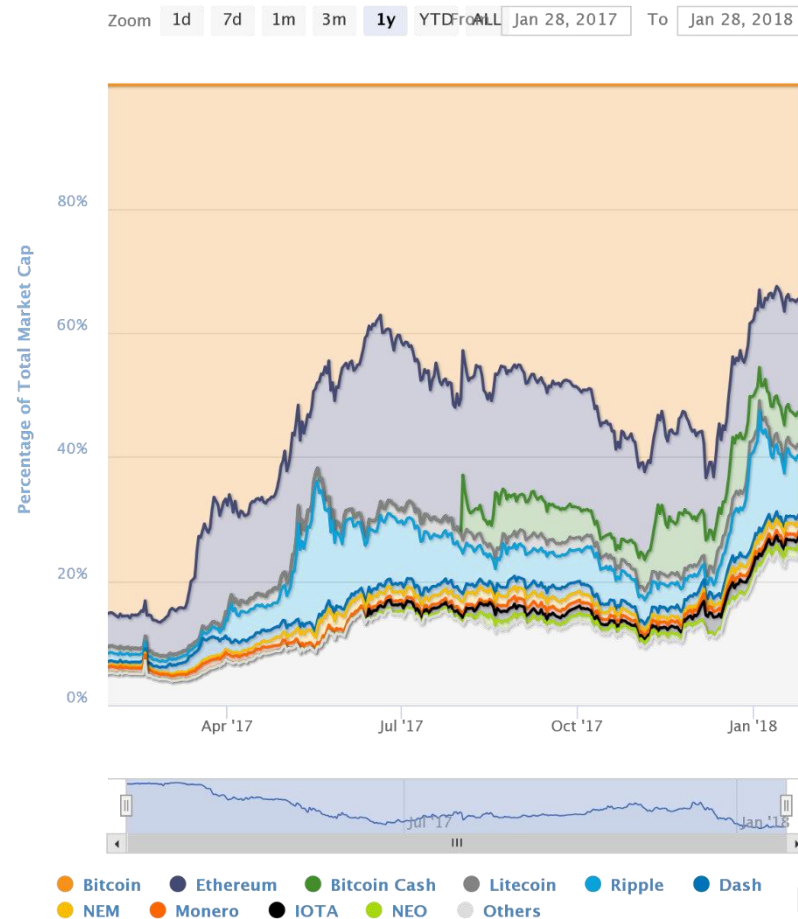
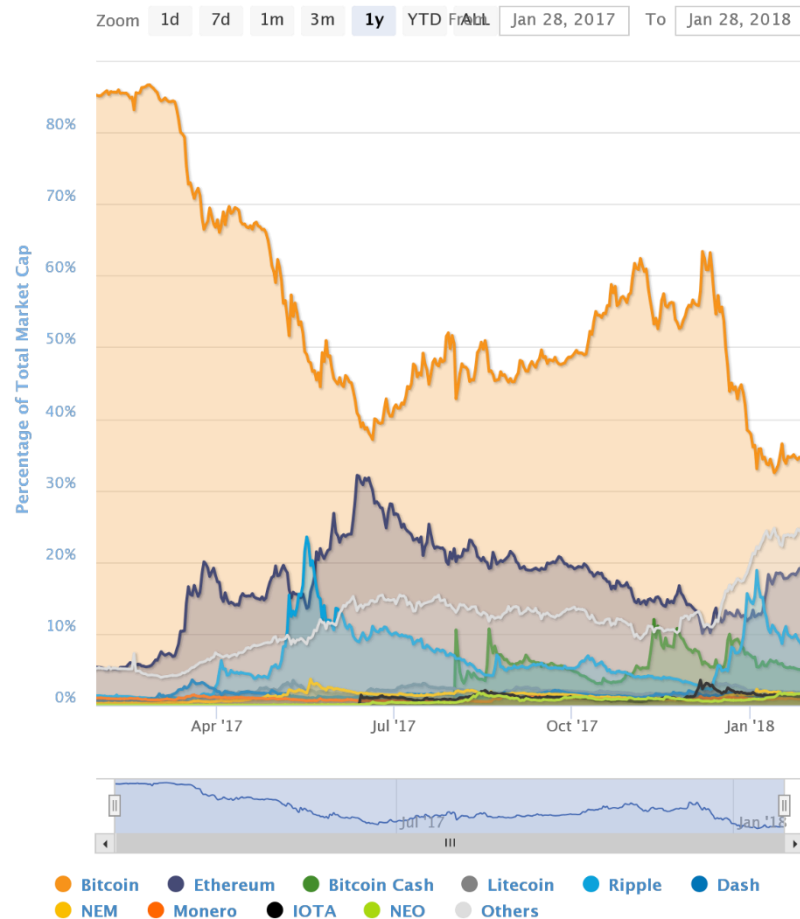
# Gráficas de barras apiladas

## Relativas

Muestran de manera relativa la contribución de cada variable al total



# Gráficas de áreas



Apiladas

<https://coinmarketcap.com/charts/>

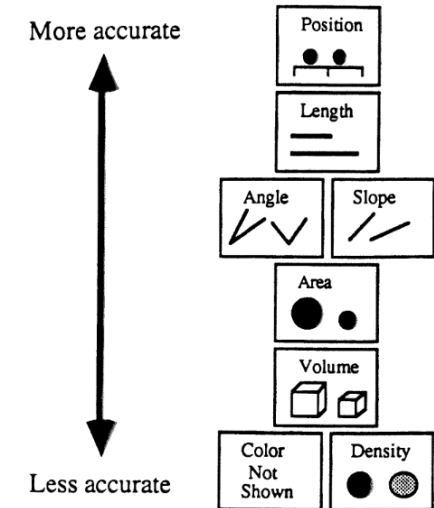
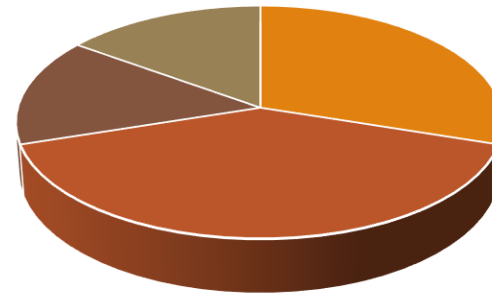
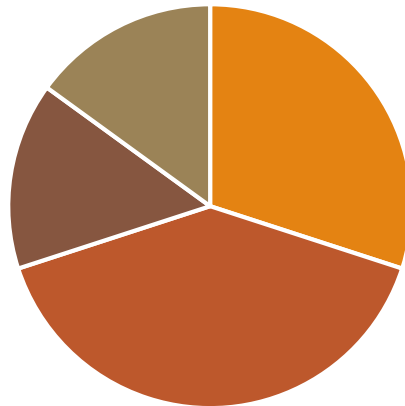
# Gráficas apiladas

## Relativas

Las gráficas de tartas (*Pie Chart*) también indican la proporción de las variables dependientes cuantitativas en una dimensión

Mapean porcentajes sobre el total al ángulo de un sector

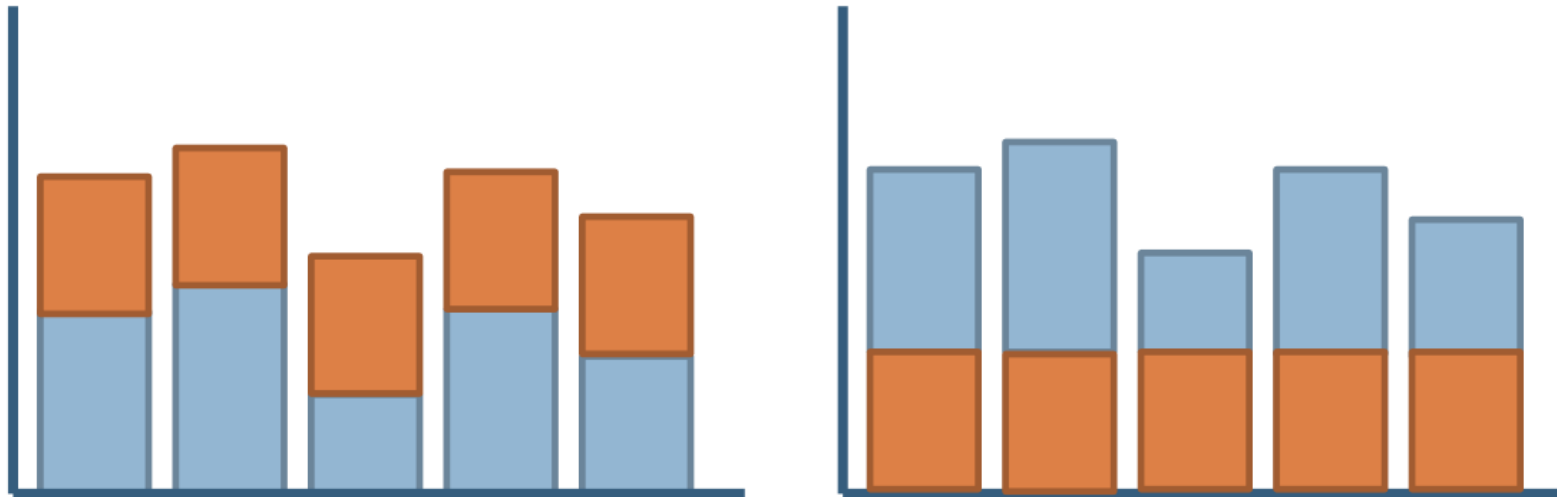
La perspectiva introduce confusión en la percepción del ángulo





# Gráficas de barras apiladas

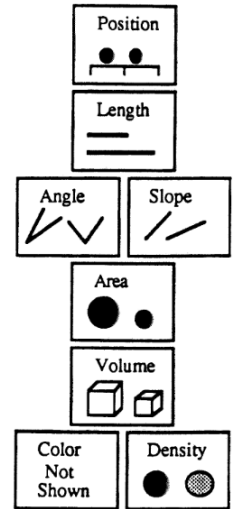
El orden de las variables afecta al resultado



La varianza de los elementos inferiores de cada barra afectan a la percepción de los superiores

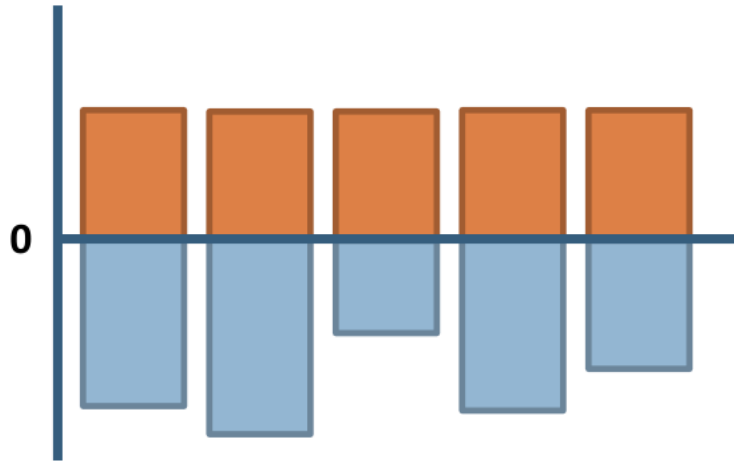
More accurate

Less accurate



# Gráficas de barras apiladas

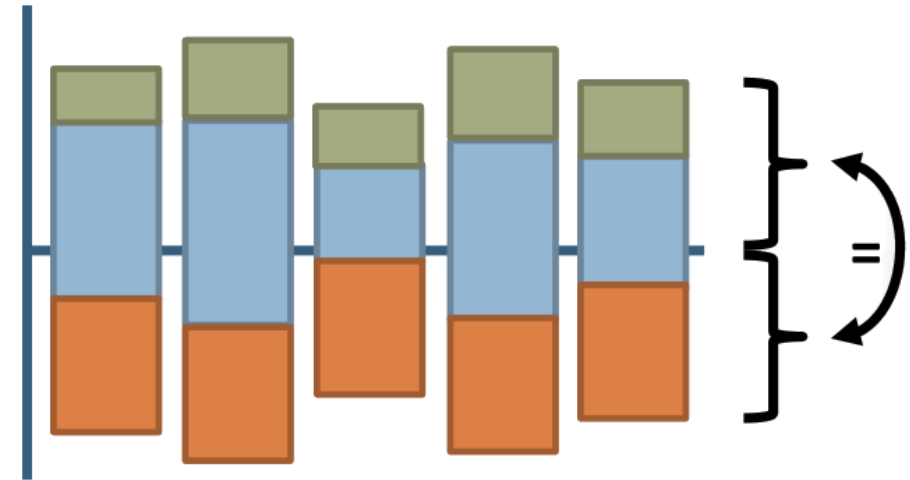
## Divergentes



Usamos tanto posición como tamaño

Sólo funciona para dos variables

Connotaciones negativas para la variable inferior



Sólo se usa el tamaño

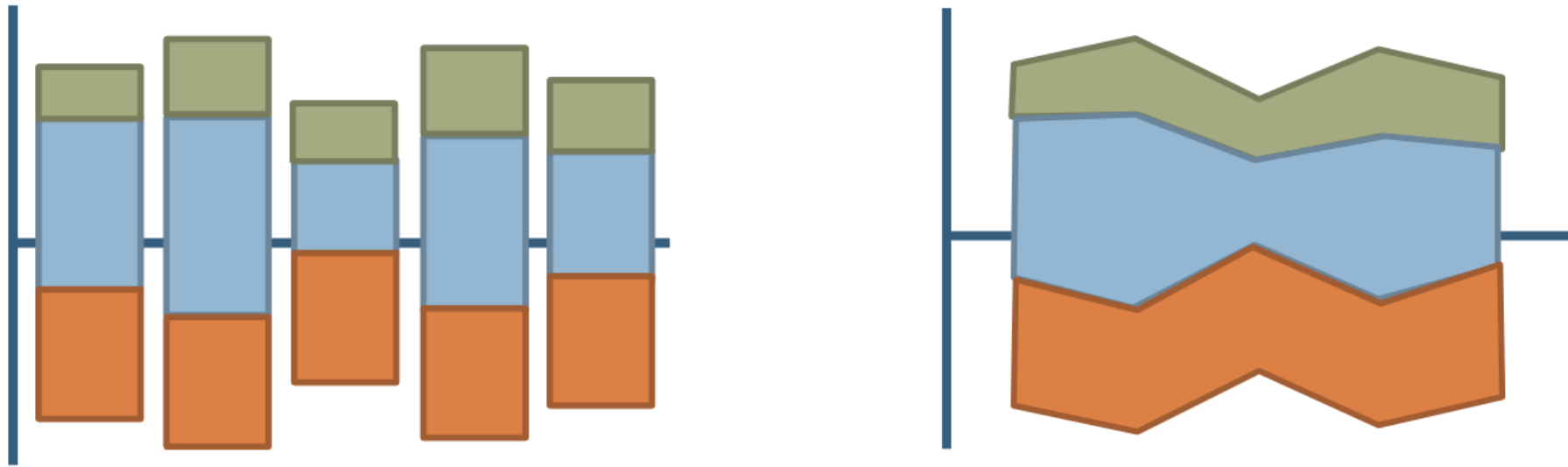
Sirva para muchas variables

Las tendencias de una variable pueden quedar ocultas por la varianza de las barras vecinas

# Gráficas de líneas apiladas

## Divergentes

---



Para datos continuos sobre una variable independiente continua

También se puede suavizar las regiones usando curvas en vez de segmentos

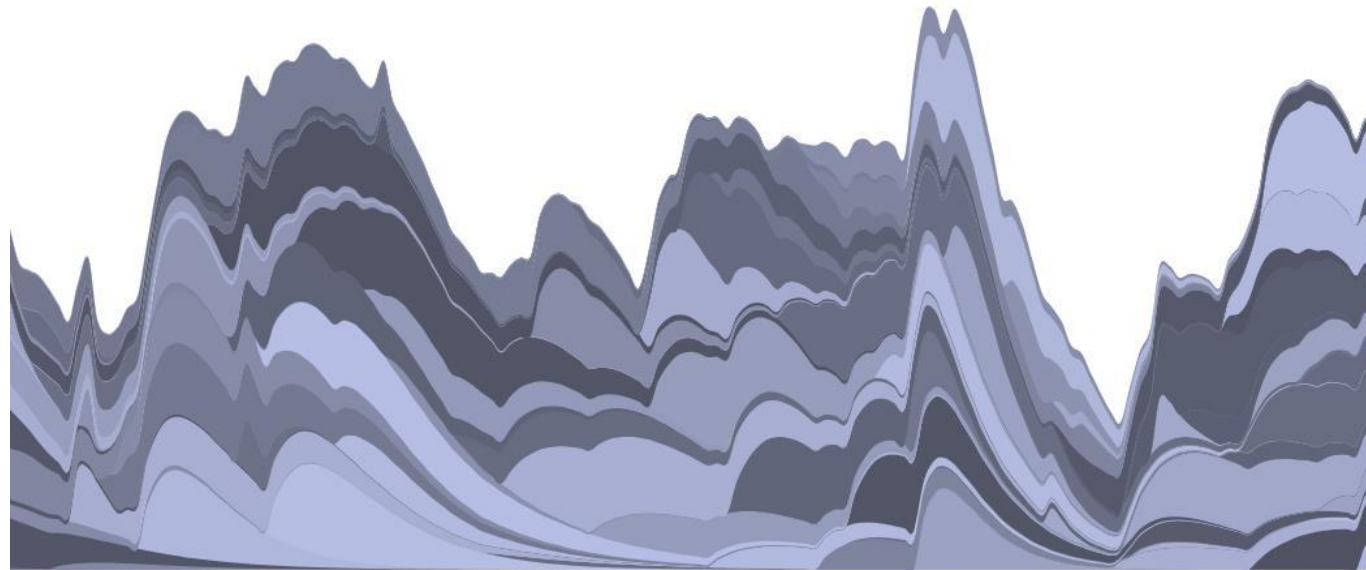
# Diseño de gráficas apiladas

---

Sea  $g_i$  la posición de la cima de la  $i$ -ésima barra:

$$g_i = g_0 + f_1 + f_2 + \cdots + f_i$$

donde  $g_0$  es la posición de la base. Para  $g_0 = 0$ , obtenemos la gráfica de áreas apilada vista



<https://github.com/leebyron/streamgraph>

# Diseño de gráficas apiladas

ThemeRiver

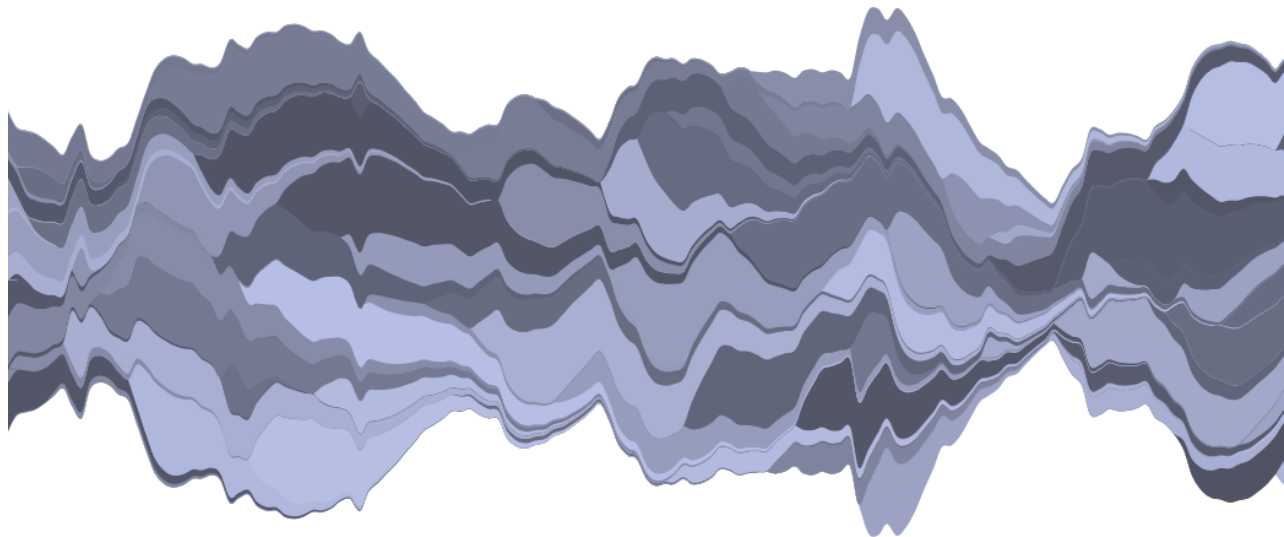
---

Sea  $g_i$  la posición de la cima de la  $i$ -ésima barra:

$$g_i = g_0 + f_1 + f_2 + \cdots + f_i$$

Este diseño define  $g_0 = -1/2 (f_1 + f_2 + \cdots + f_n)$

Minimiza la distancia de la gráfica al eje horizontal y las pendientes de los límites superior e inferior



# Diseño de gráficas apiladas

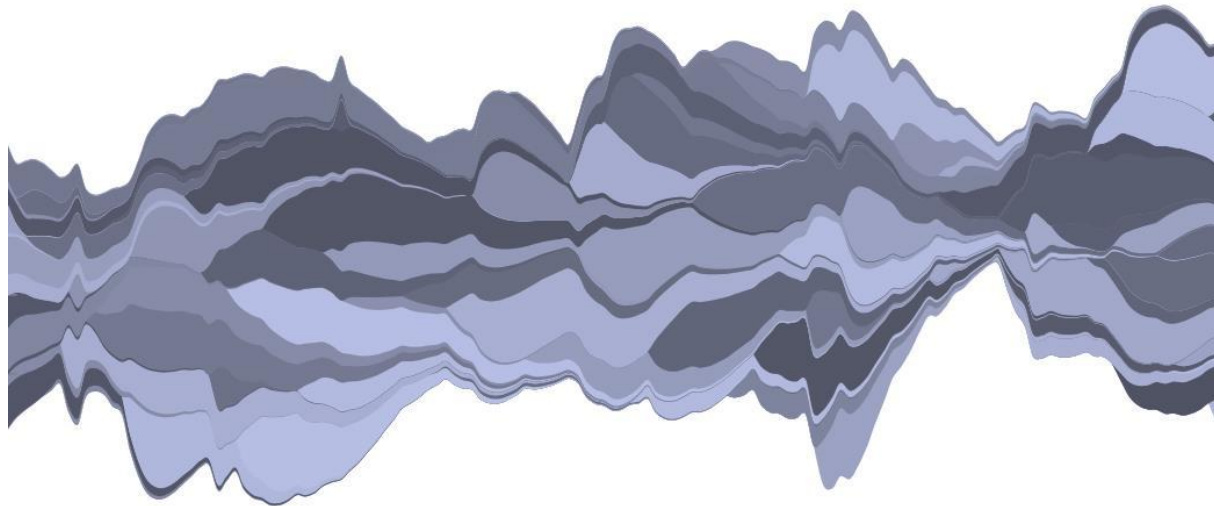
## Streamgraph

---

Este diseño establece la base en:

$$g_0 = -\frac{1}{n+1} \sum_{i=1}^n (n-i+1) f_i$$

Hace más fácil seguir la evolución de cada variable, minimizando los cambios en su posición a lo largo de la gráfica, y su pendiente

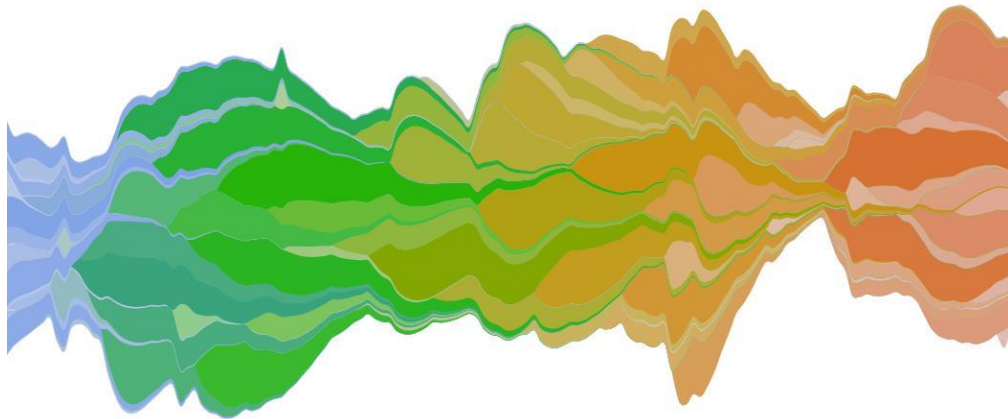


# Diseño de gráficas apiladas

## Streamgraph

---

Suponiendo que las variables aparecen (toman valores distintos de cero) y desaparecen (se hace cero), se pueden añadir las nuevas variables:



En un orden arbitrario

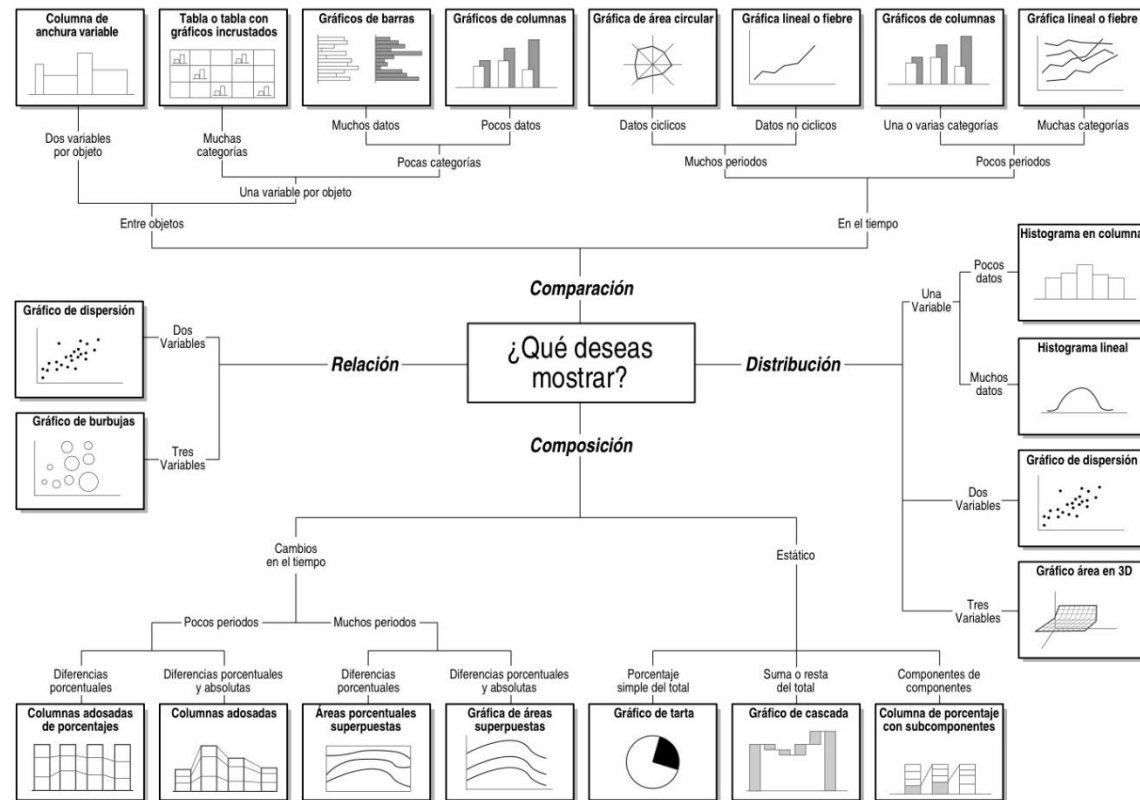


Introduciendo las nuevas variables en el exterior

- Así, el pico inicial no afecta al resto de variables

# Ayuda

## ¿Qué gráfico elegir?



Traducción a cargo de Victor Caballero con autorización del autor  
www.Vectart.com / contact@vectart.com

www.ExtremePresentation.com  
© 2009 A. Abela - a.v.abela@gmail.com

[http://extremepresentation.typepad.com/blog/2006/09/choosing\\_a\\_good.html](http://extremepresentation.typepad.com/blog/2006/09/choosing_a_good.html)



# Bibliografía

---

Curso de John C. Hart. Data Visualization. Universidad de Illinois en Urbana-Champaign

- Disponible online en Coursera