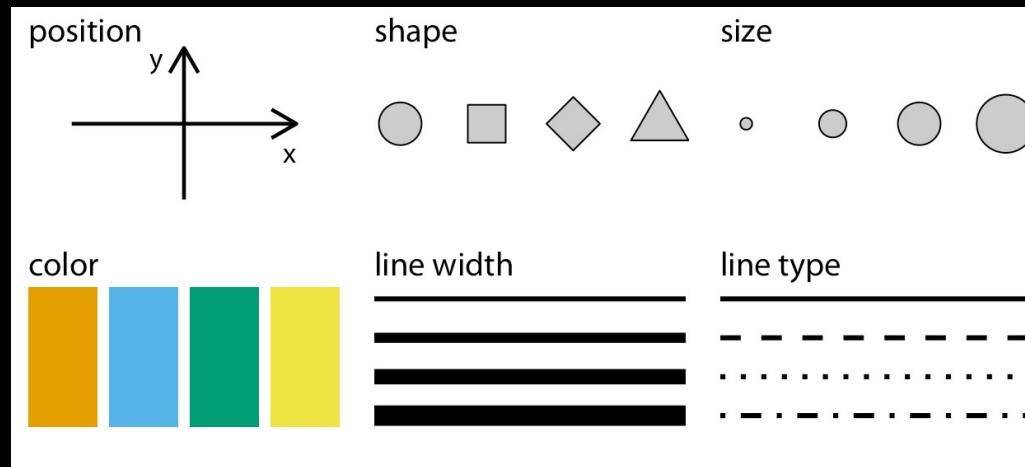


# VISUALIZACIÓN CIENCIA DE SÍMBOLOS GRÁFICOS Y MARCAS

# Semiología de los Símbolos Gráficos

- La ciencia de símbolos gráficos y marcas se conoce como semiología.
- Un símbolo gráfico se considera un objeto visual.
- La estética describe todos los aspectos de un elemento gráfico.
- Un componente critico de cualquier elemento gráfico es su posición, la cual describe el lugar en el que se localiza el elemento; el sistema de posicionamiento más común es el Cartesiano, pero hay otros que dependen del sistema de visualización.
- Todos los elementos gráficos poseen una forma, un tamaño y un color. Los elementos gráficos también pueden tener otras propiedades tales como ancho, patrón de relleno (textura), opacidad o tipo de letra.



# Semiología de los Símbolos Gráficos

- Todas las características estéticas caen en dos grupos: aquellos que pueden representar datos continuos y aquellos que no.
- Los valores de datos continuos son aquellos para los cuales existen intermediarios arbitrariamente pequeños. En contraste a los datos continuos existen los datos discretos.
- Además de datos numéricos continuos y discretos (datos cuantitativos) se pueden tener categorías discretas (fechas, tiempos) y texto (datos cualitativos).
- Las variables que tienen datos cualitativos se les llama **factores** y las diferentes categorías de esos datos se llaman **niveles**. Los niveles de un factor normalmente no tienen un orden, pero también es posible que los factores tengan un orden (p.ej., cuando hay un orden intrínseco entre los niveles del factor).

# Semiología de los Símbolos Gráficos

Tipo de Variable	Ejemplos	Escala Apropriada	Descripción
Numérica Continua (Cuantitativa)	1.3, 5.7, 83, $1.5 \times 10^5$	Continua	Valores numéricos arbitrarios (enteros, números irracionales, números reales)
Numérica Discreta (Cuantitativa)	1, 2, 3, ...	Discreta	Números en unidades discretas. Los números enteros son de los más comunes.
Categóricos sin orden (Cualitativos)	Perro, Gato, Pajaro	Discreta	Categorías sin orden. Categorías discretas y únicas que no tienen un orden intrínseco. A estas variables se les llama <i>factores</i> .
Categóricos ordenados (Cualitativos)	Bueno, Justo, Pobre	Discreta	Categorías con orden. Categorías discretas y únicas con un orden. A estas variables se les llama <i>factores</i> .
Fechas, horas o tiempos	Enero 5, 1980, 2:11 am	Continua o Discreta	
Texto	Un Texto Que importe algo	Ninguna o Discreta	Texto libre que puede ser tratado como categórico si es necesario.

Tipos de variables que se encuentran en escenarios de visualización típicos

# Símbolos y Visualizaciones

---

- Existen símbolos y gráficos que son reconocibles universalmente (adquieren dicha característica después de algún tiempo) y otros que requieren de mayor atención para poder entender su significado. El entendimiento de las gráficas o símbolos del segundo tipo de figuras se realiza en dos etapas. La primera etapa identifica los elementos principales de la imagen y la segunda identifica las diferentes relaciones entre ellos.
- En la primera etapa los símbolos son percibidos, con cierto esfuerzo, y son transferidos de la memoria de largo plazo. Los patrones son extraídos de la imagen. En la segunda etapa se identifican los aspectos más interesantes (los conjuntos de puntos, productos, o genes más interesantes), aquellos que tienen las características más interesantes o especiales.

# Símbolos y Visualizaciones

---

- El descubrimiento de las relaciones o patrones ocurre también a través de dos etapas. La primera es un mapeo entre cualquier relación de los símbolos gráficos y los datos que ellos representan (cualquier patrón en la imagen debe implicar un patrón en los datos). Si esto no ocurre, hay un artefacto en la representación gráfica seleccionada.
- Algo similar ocurre cuando cualquier variación en los patrones de los gráficos o símbolos implican cognitivamente dicha variación en los datos. Lo mismo ocurre para los ordenamientos, cualquier orden percibido en los símbolos gráficos está directamente correlacionado con un orden, percibido, en los datos y viceversa. Si un orden se percibe visualmente y dicho orden no está presente en los datos, hay un engaño.
- En resumen:
  - Similitud en la estructura de los datos  $\Leftrightarrow$  Similitud visual de los símbolos correspondientes.
  - Orden entre datos  $\Leftrightarrow$  Orden visual entre símbolos correspondientes.

# Símbolos y Visualizaciones

---

## IMPORTANTE

Es muy importante resaltar que sin identificación externa (cognitiva), una gráfica es inutilizable. La identificación externa debe ser legible y entendible directamente. Debido a que nuestra percepción se guía por interpretaciones físicas, las imágenes relevantes deben poseer dimensiones que sean de fácil interpretación y los elementos gráficos de la imagen deben ser claros.

# Reglas de Elementos Gráficos

---

- El objetivo de un elemento gráfico es el descubrimiento de grupos u ordenes horizontalmente (eje x) y grupos u ordenes verticalmente (eje y) que son formados con valores (sobre el eje z).
- La construcción sobre coordenadas tri-dimensionales (Cartesianas, esféricas, etcétera) permite el descubrimiento de estos grupos en todos los casos.
- Dentro de la construcción sobre coordenadas tri-dimensionales, las permutaciones y clasificaciones permiten resolver el problema de interpretación y meta-information.
- Cada elemento gráfico con más de tres factores que difieran de la construcción tri-dimensional destruye la unidad de la gráfica y la interpretación.
- Todas las imágenes deben ser legibles y entendibles por un ser humano.

# Análisis de un Elemento Gráfico

---

- Cuando se analiza un elemento gráfico primero se perciben grupos de objetos de manera preatentamente, después se intenta caracterizar dichos grupos (de manera cognitiva). Finalmente, se examinan casos especiales que no se encuentran dentro de los grupos o de las relaciones entre grupos (o alguna combinación de ambos).
- Este proceso se puede realizar a varios niveles y con muchas visualizaciones diferentes.

# Variables Visuales

- La aplicación de elementos gráficos para comunicar información requiere de un entendimiento de las primitivas gráficas (**marcas**) y sus propiedades.
- Una forma de codificar los datos para ser desplegados es mapear los diferentes valores de datos a diferentes **marcas** y sus atributos. Sin embargo, las **marcas** por sí mismas no definen despliegues informativos, ya que la simple colocación de marcas tiene el efecto de obscurecer marcas ya presentes; sólo a través de la colocación apropiada de las marcas que se logra un despliegue informativo.
- El posicionamiento de **marcas** dentro de un espacio de despliegue provee una forma de mapear o revelar propiedades adicionales de los datos que respaldan al despliegue, incluyendo similitud y distribución.
- Una vez que se ha especificado el arreglo y los tipos de las **marcas** es entonces que se pueden aplicar propiedades gráficas adicionales a las **marcas**.
- Las marcas pueden varias en tamaño, desplegarse con colores diferentes y pueden ser mapeados a orientaciones distintas (todas estas variaciones pueden estar determinadas por los datos para transmitir información).
- Son, básicamente, ocho las formas en que los objetos gráficos pueden codificar información (ocho variables visuales): **posición, forma, tamaño, brillantez, color, orientación, textura y movimiento**.
- Estas variables pueden ser ajustadas como sea necesario para maximizar la efectividad de una visualización para transmitir información.
- Sin embargo, cuando se define una visualización, es importante recordar que el resultado será una imagen que debe ser interpretada por el sistema de visión humano y que será sujeto a todos los problemas y todas las reglas de percepción de dicho sistema.

# Variables Visuales

## Posición

- Probablemente, la posición es la variable visual más importante: la colocación de elementos gráficos representativos dentro del espacio de despliegue.
- La posición tiene el mayor impacto en el despliegue de la información porque el arreglo espacial de elementos gráficos es el primer paso en la legibilidad de una visualización. Básicamente, la maximización del esparcimiento de los gráficos en todo el espacio de despliegue maximiza la cantidad de información comunicada, hasta cierto punto.
- El peor despliegue con el peor esquema de posicionamiento es aquel que mapea todos gráficos sobre la misma posición (saturación y poca transmisión de información). El mejor esquema de posicionamiento mapea cada gráfico a posiciones únicas, de tal forma que los gráficos pueden verse sin traslapos.
- Considerando un monitor de  $1024 \times 768$  (con 786,432 píxeles), no es posible desplegar muchos gráficos porque ellos típicamente cada uno de ellos ocupa más de un píxel.

# Variables Visuales

## Posición

- La selección de las variables usadas para organizar los datos dentro de un despliegue puede responder a una variedad de preguntas (en donde se localizan los datos, si los datos se ajustan a distribuciones conocidas, existen tendencias en los datos). La utilización de percepción y las relaciones de proximidad, simetría y otros principios Gestalt se pueden formar grupos y estructuras dentro de los datos.
- Además de la selección de variables apropiadas para organizar el espacio del despliegue y para presentar valores con gráficos representativos, se pueden aplicar cambios de escalas o escalamiento a variables para remapear valores para revelar estructuras. Aunque el cambio de escalas o escalamiento no es necesario en general, hay ocasiones en que se debe utilizar. El primer tipo de escala es la lineal (una alargamiento o encogimiento de un rango). El segundo tipo de escalamiento o cambio de escala es la logarítmica que se usa para mapear variables que cambian exponencialmente a rangos más compactos.
- Aunque las escalas lineales y logarítmicas se aplican a una variable de los datos, hay ocasiones en que es necesario hacer mapeos multidimensionales (p.e., bidimensionales), por ejemplo en cartografía. El mapeo de dos o más variables se conoce como **proyección** de un espacio de datos a otro (generalmente de menor dimensión).
- Adicionalmente al despliegue de gráficos representativos y al uso de varias escalas o proyecciones, cuando se trabaja con varias variables es práctica común agregar gráficos suplementarios para describir el espacio. Los ejes son elementos gráficos que proveen información adicional para entender como se define el espacio visual. Los ejes típicamente contienen marcas de graduación que indican los intervalos de los rangos de los datos y etiquetas que proveen valores a los datos. El título de un eje usualmente contiene el nombre de la variable de datos que se está mapeando.

# Conceptos Básicos

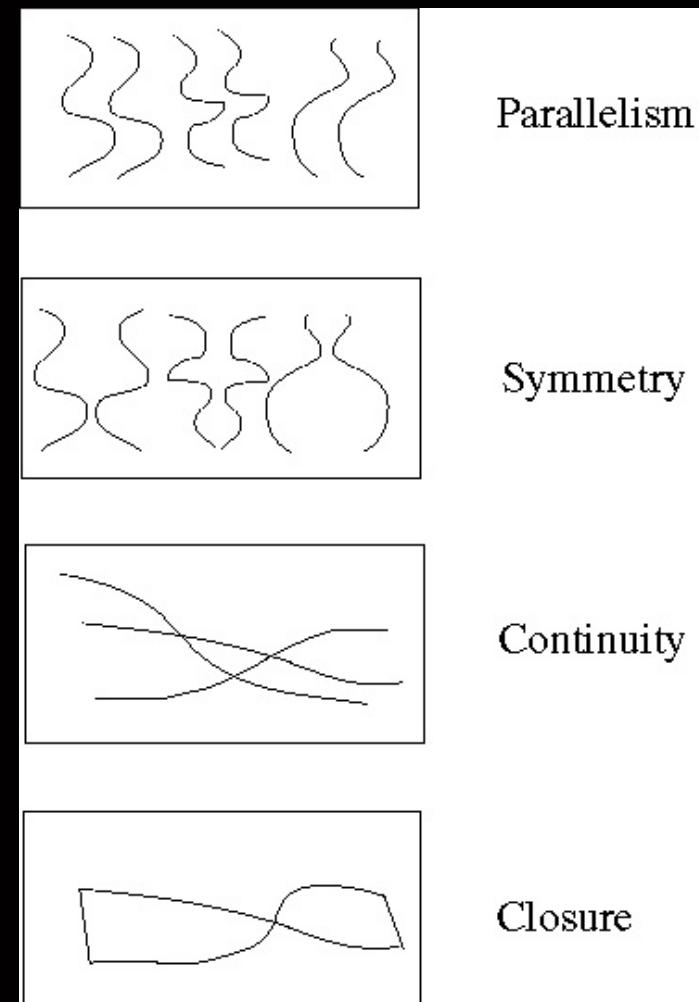
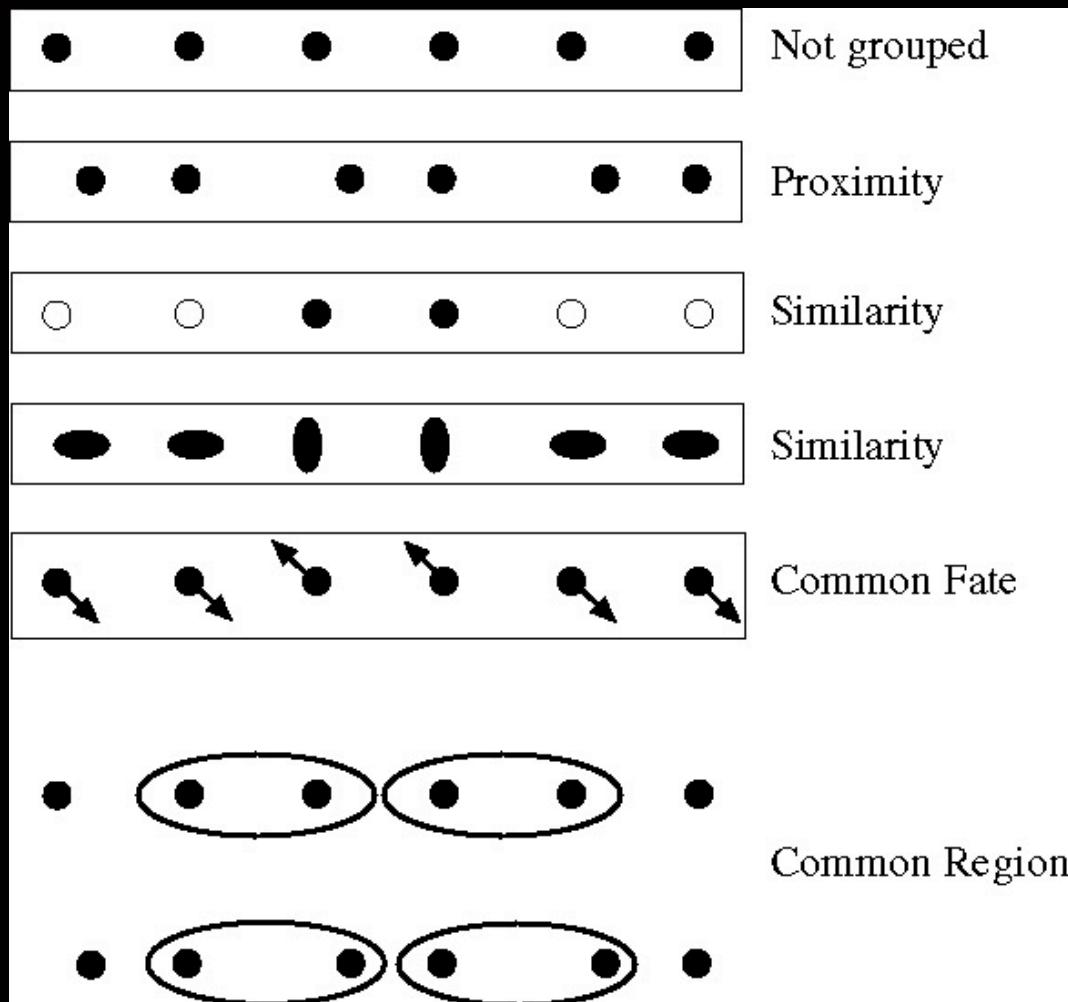
---

- **Tokens:** los elementos que se desean agrupar (i.e., píxeles, puntos, o elementos de superficie).
- **Segmentación Arriba-Abajo:** los tokens pertenecen a un grupo debido a que se encuentran en el mismo objeto.
- **Segmentación Abajo-Arriba:** los tokens pertenecen a un grupo porque poseen coherencia local.

# Agrupamiento Básico

---

- Propiedades Gestalt: los elementos en una colección pueden poseer propiedades que resultan de relaciones (efecto Muller-Lyer).
  - Cualidad Gestalt.
- Una serie de factores determinan que los elementos deben agruparse o no.
  - Factores Gestalt.



# Variables Visuales

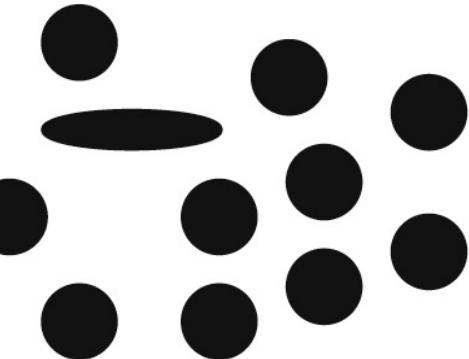
## Posición



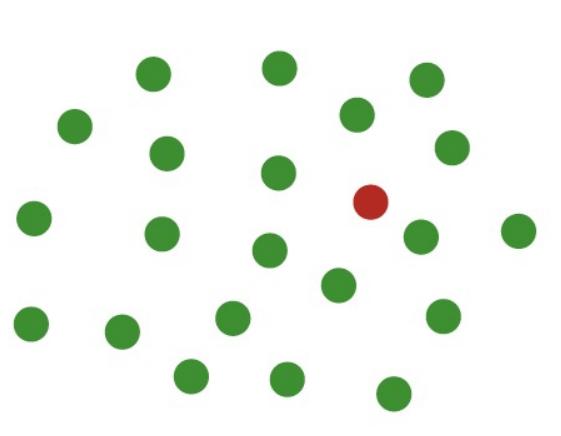
Orientación



Orientación



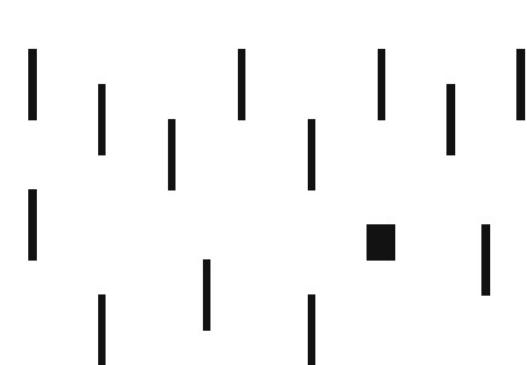
Forma



Color



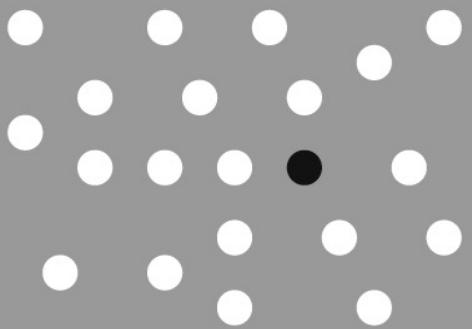
Tamaño



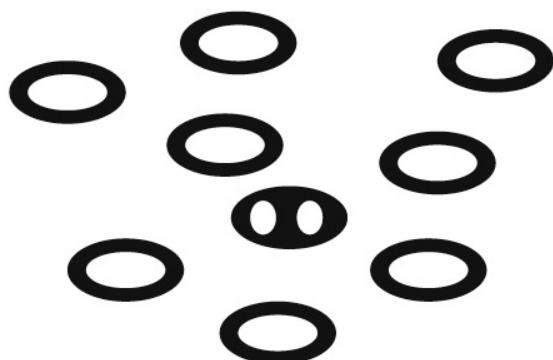
Forma

# Variables Visuales

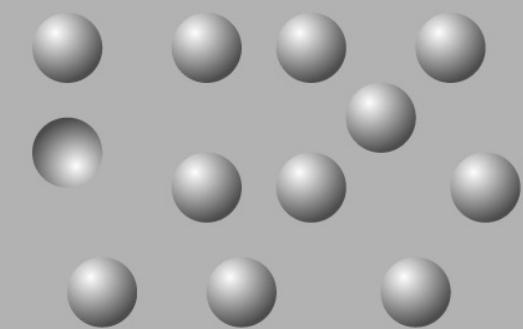
## Posición



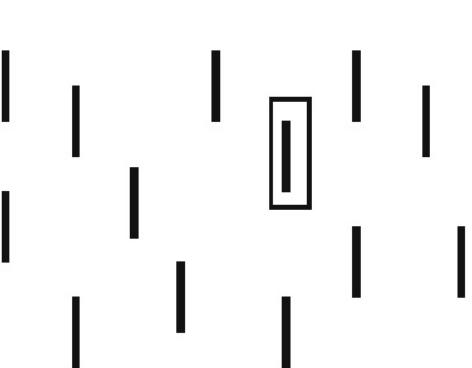
Claro/Oscuro



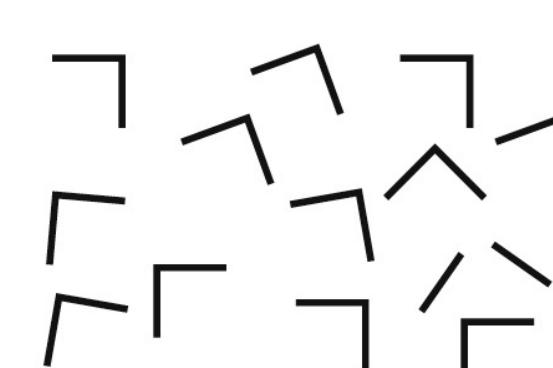
Topología/Conteo



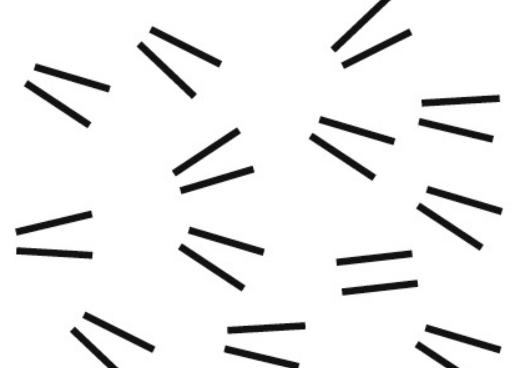
Convexidad/Concavidad



Adición



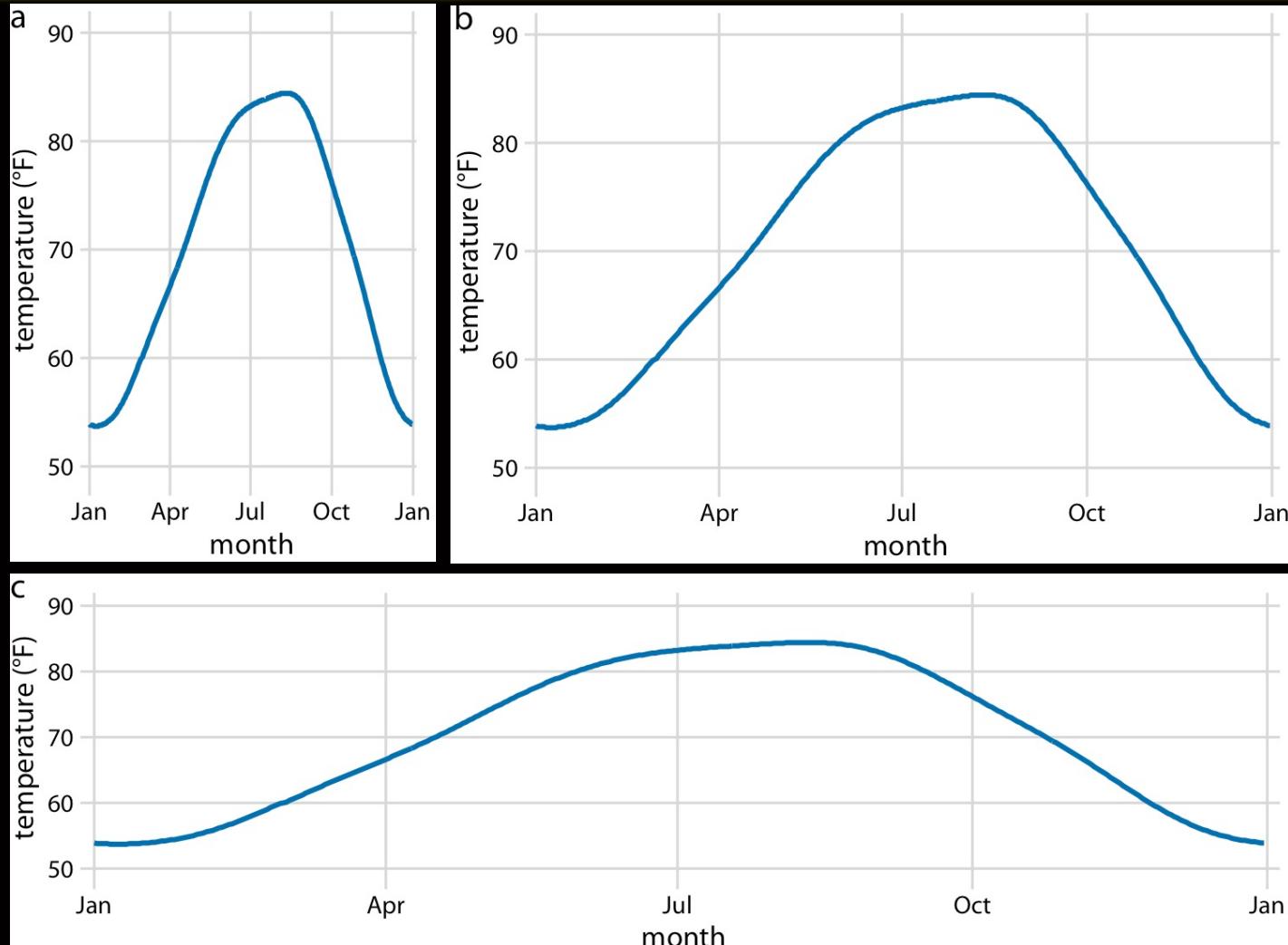
Coyuntura



Paralelismo

# Variables Visuales

## Posición



# Variables Visuales

## Posición

- Cuando los ejes usan diferentes unidades, se pueden comprimir o alargar los ejes y cambiar el aspecto de la figura final sin cambiar la validez de la visualización.
- La cuestión de que versión es preferible depende de las especificaciones del proyecto o de la historia que uno quiere transmitir.
- Por ejemplo, una figura alta y estrecha enfatiza el cambio a lo largo del eje vertical y una figura baja y ancha enfatiza el cambio a lo largo del eje horizontal. Idealmente, se desea un aspecto que asegure que las diferencias de posición importantes sean evidentes.
- Por otro lado, si todos los ejes usan las mismas unidades, entonces el espaciado de los ejes debe ser igual en todos para que la distancia en un eje corresponda en los otros.

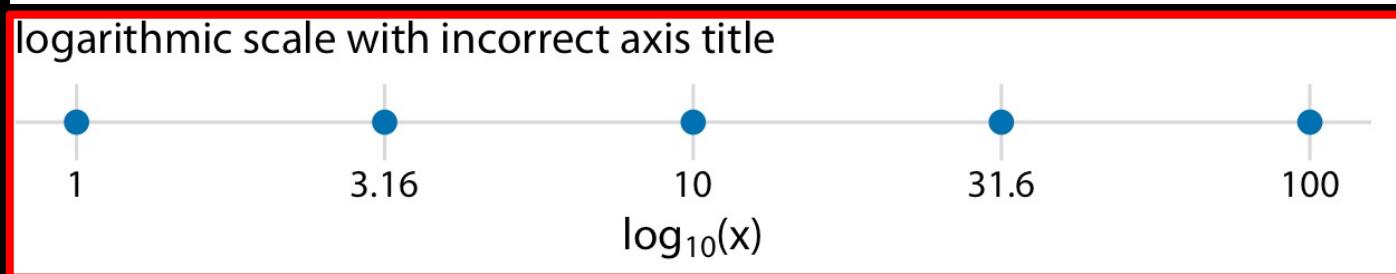
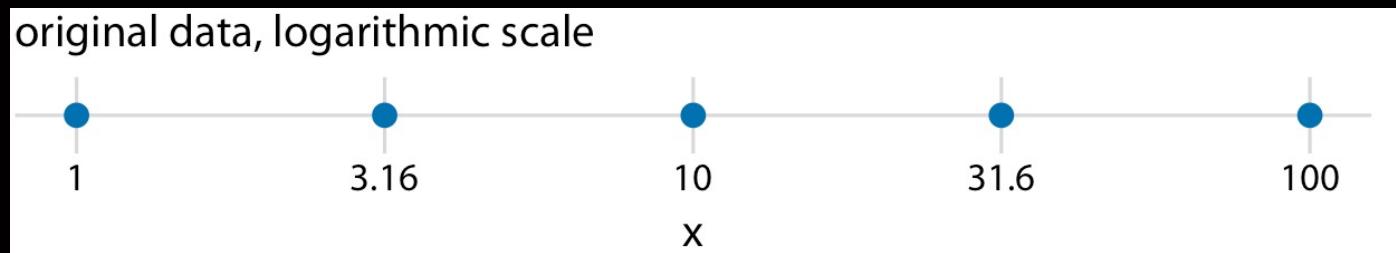
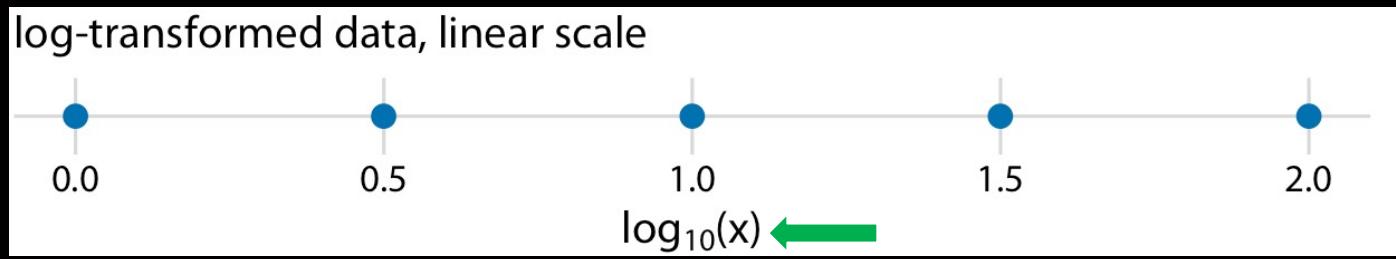
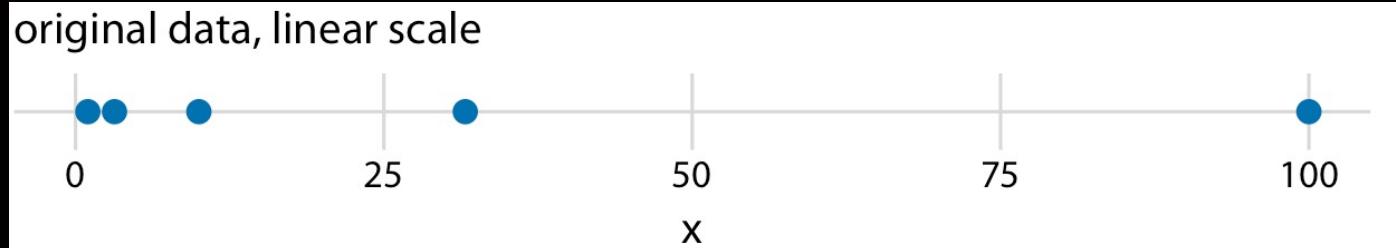
# Variables Visuales

## Posición

- Para el caso de ejes no lineales, un espaciado homogéneo en las unidades de los datos corresponde a un espaciado heterogéneo en la visualización y viceversa.
- La escala no-lineal más común es la escala logarítmica. Las escalas logarítmicas son lineales para la multiplicación, de tal forma que un paso unitario en la escala corresponde a la multiplicación por un valor fijo. Para crear un escala logarítmica, se necesita una transformación logarítmica de los datos mientras se elevan a un exponente los números que se muestran a lo largo de las líneas de los ejes.

# Variables Visuales

## Posición



# Variables Visuales

## Posición

- Debido a que la multiplicación en una escala logarítmica se parece a la suma en una escala lineal, las escalas logarítmicas son una opción natural para datos que han sido obtenidos a través de multiplicación o división.
- En particular, tasas (*ratios*), en general, deben ser mostrados usando una escala logarítmica.
- Las escalas logarítmicas se usan frecuentemente cuando los conjuntos de datos contienen números de magnitudes muy diferentes.
- Escalas que usan la raíz cuadrada comprimen números grandes en un rango más pequeño, pero a diferencia de escalas logarítmicas, estas permiten la presencia del número 0.

# Variables Visuales

## Posición

- Existen otros sistemas coordenados en donde los ejes son curvos. Por ejemplo, el sistema coordenado polar; en este sistema las posiciones se especifican a través de un ángulo y una distancia radial desde el origen coordenado; por lo tanto, el eje angular es circular.
- Las coordenadas polares son útiles para datos de naturaleza periódica, tales como sinusoidales o días del año.
- Otro contexto en el que se encuentran ejes curvos es con datos geo-espaciales. Las posiciones sobre el globo terráqueo se pueden especificar por su longitud y latitud. Sin embargo, debido a que la Tierra es una esfera, graficar las latitudes y longitudes como ejes Cartesianos es engañoso e inapropiado. Es por ello que se utilizan proyecciones no-lineales que intentan minimizar artefactos y balancear entre conservar áreas o ángulos y las líneas verdaderas de las formas sobre el globo terráqueo.