

---

---

## Visualización, algunas pautas y ejemplos

---

---

# Motivación

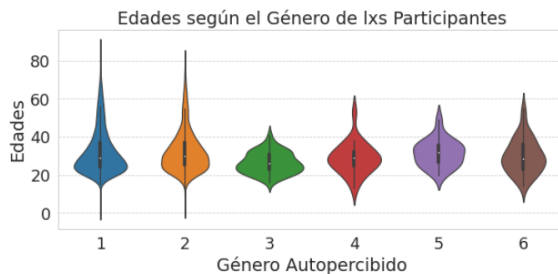
A la hora de explorar los datos con los que nos enfrentamos o cuando queremos comunicar nuestros análisis a determinada audiencia, la **visualización** aparece como una de las herramientas fundamentales

# Motivación

A la hora de explorar los datos con los que nos enfrentamos o cuando queremos comunicar nuestros análisis a determinada audiencia, la **visualización** aparece como una de las herramientas fundamentales



Género Autopercebido	Media de la Edad	Desviación Estándar de la Edad
1	32.18	11.68
2	32.49	11.00
3	26.45	5.52
4	28.83	8.15
5	31.58	7.01
6	30.00	9.12



# Motivación

A la hora de explorar los datos con los que nos enfrentamos o cuando queremos comunicar nuestros análisis a determinada audiencia, la **visualización** aparece como una de las herramientas fundamentales

Entonces, es necesario entender el **contexto** para el cual necesitamos la visualización, si la misma va ir acompañada de información extra. Además, decidir la correcta herramienta de visualización y la **estética** necesaria para resaltar lo que queremos indicar.

# Temario

Generalidades

Tipos de Datos

Sistemas de Coordenadas

Colores

Algunos Ejemplos

Algunos Ejemplos en Práctica

# Generalidades

Si bien visualizar puede llegar a sernos algo intuitivo, está bueno tener en cuenta, por lo menos, las siguientes cuestiones:

- Contexto: exploratorio, académico, divulgativo

# Generalidades

Si bien visualizar puede llegar a sernos algo intuitivo, está bueno tener en cuenta, por lo menos, las siguientes cuestiones:

- Contexto: exploratorio, académico, divulgativo
- Tipo de datos: cuantitativos, categóricos, discretos,

# Generalidades

Si bien visualizar puede llegar a sernos algo intuitivo, está bueno tener en cuenta, por lo menos, las siguientes cuestiones:

- Contexto: exploratorio, académico, divulgativo
- Tipo de datos: cuantitativos, categóricos, discretos,
- Tipo de gráfico: scatter, barras, torta



# Generalidades

Si bien visualizar puede llegar a ser algo intuitivo, está bueno tener en cuenta, por lo menos, las siguientes cuestiones:

- Contexto: exploratorio, académico, divulgativo
- Tipo de datos: cuantitativos, categóricos, discretos,
- Tipo de gráfico: scatter, barras, torta
- Aspectos estéticos: colores, formas, ejes

# Tipos de Datos

Dependiendo del problema que estemos estudiando, vamos a tener datos de distinto tipo, y resulta importante poder identificarlos para saber cómo trabajar con los mismos

# Tipos de Datos

## Tipo de Dato

## Ejemplo

Numérico continuo

Cantidad de contagios

Frecuencia de uso de palabras

Numérico discreto

Nombres de vacunas

Cantidades de algún bien

Categorico

Tiempos de adquisición

Nombre de medios

Categorico ordenable

Fechas de ingresos de vacunas

Fechas

Textos de noticias

Bueno, regular, malo

Texto

Distancias

Rangos etarios

# Tipos de Datos

En general, nos enfrentaremos a estos tipos de datos:

Tipo de Dato	Ejemplo	Posibles Visualizaciones
Numérico continuo	Distancias, cantidades de algún bien, etc	Scatter, línea,
Numérico discreto	Cantidad de contagios, frecuencia de uso de palabras	Barras, scatter
Categorico	Vacunas, nombre de medios	Torta, barras agrupadas (?)
Categorico ordenable	Bueno, regular, malo / rangos etarios	“
Fechas	Tiempos de adquisición (continuo), fechas de ingreso de vacunas (discreto)	
Texto	Observaciones, textos de noticias	Nubes de palabras

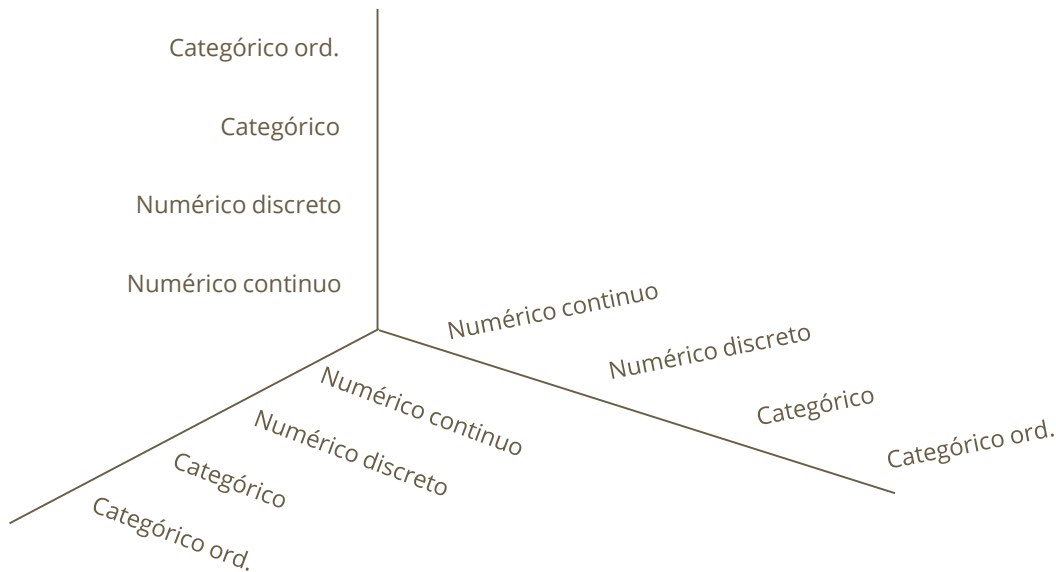
# Tipos de Datos

Pero, en realidad, va a ser más complejo:

Tipo de Dato	Numérico continuo	Numérico discreto	Categorógico	Categorógico ordenable	Fechas	Texto
Numérico continuo	scatter		barras		lineal	nube de palabras
Numérico discreto	-					
Categorógico	-	-				
Categorógico ordenable	-	-	-			
Fechas	-	-	-	-		
Texto	-	-	-	-	-	

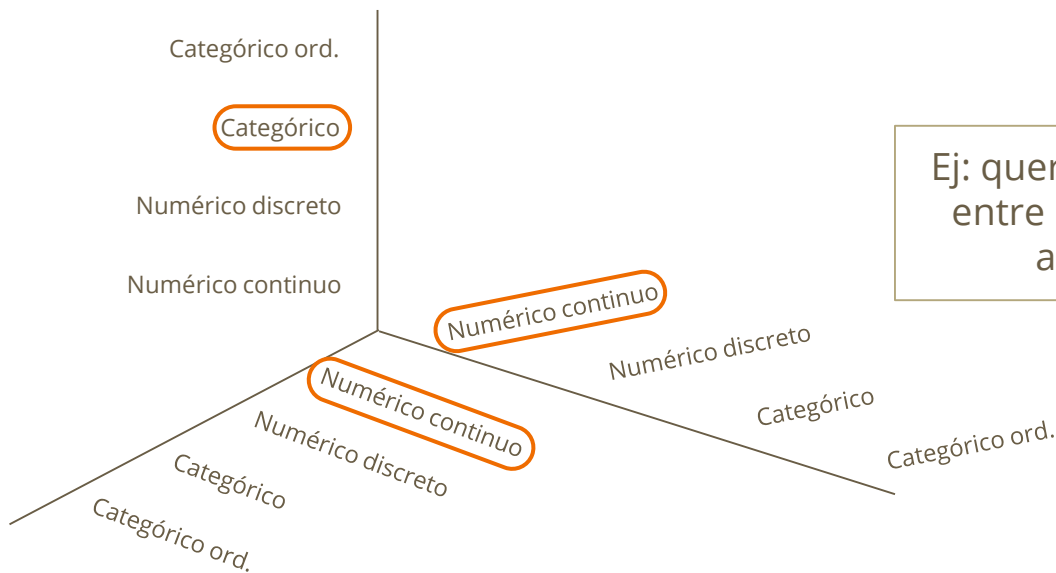
# Tipos de Datos

Pero, en realidad, va a ser más y más complejo:



# Tipos de Datos

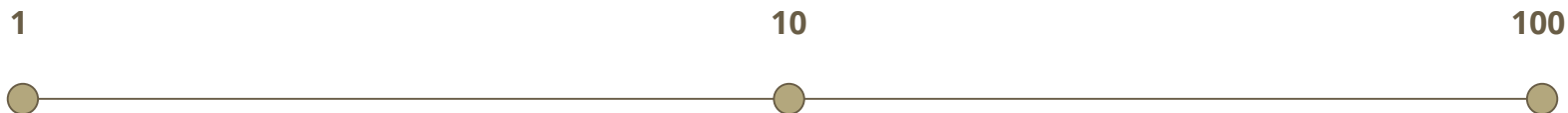
Pero, en realidad, va a ser más y más complejo:



Ej: queremos ver correlación  
entre dos variables, según  
alguna categoría

# Sistemas de Coordenadas

Dependiendo del tipo de datos, podemos elegir un sistema de coordenadas adecuado para mostrar los mismos





# Sistemas de Coordenadas - Lineal

El sistema que más frecuentemente usamos, los ejes son ortogonales entre sí. Al poner una grilla sobre los ejes, las marcas son equiespaciadas.

A tener en cuenta:

- Podemos manejar los tamaños relativos entre ejes, cuando son magnitudes distintas
- La distancia entre dos puntos en cualquier lugar del plano implica la misma distancia en los datos.
- Invarianza frente a transformaciones lineales de escala

# Sistemas de Coordenadas - No lineal

Muchas veces, veremos que la linealidad no traduce satisfactoriamente. Datos ampliamente distribuidos a lo largo de varios órdenes de magnitud suelen agolparse en los ejes cartesianos.

# Sistemas de Coordenadas - No lineal

Muchas veces, veremos que la linealidad no traduce satisfactoriamente. Datos ampliamente distribuidos a lo largo de varios órdenes de magnitud suelen agolparse en los ejes cartesianos.

¿Qué hacemos? Transformar los datos (o los ejes) no linealmente:

- escalas logarítmicas (ojo con el cero)
- escalas raíz cuadrada (ojo, poco intuitiva)

# Sistemas de Coordenadas - Radial

Por fuera de ejes ortogonales, aparecen los ejes radiales.

Puede servirnos para:

- Visualizar distribuciones: si tenemos pocas categorías, como alternativa a stacked bars
- Visualizar datos intrínsecamente radiales: polarización en función de un ángulo

# Colores

Los colores nos ayudan en dos cuestiones claves:

- Separar categorías
- Trazar una equivalencia visual entre valores numéricos y un mapa de colores

# Colores para Categorías

Usamos colores discretos. Deberíamos respetar que:

- Los colores no sean parecidos
- Ningún color sobresalga sobre el resto
- Que no haya una secuencia entre los mismos que implícitamente induzcan un ordenamiento



# Colores para Valores Continuos

Usamos una escala secuencial cuando querramos dar cuenta de valores no necesariamente centrados

ColorBrewer Blues



Heat



Viridis



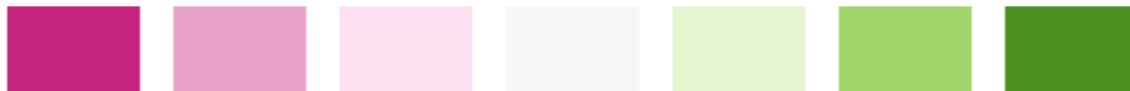
# Colores para Valores Continuos

Usamos una escala divergente cuando querramos dar cuenta de valores centrados en un determinado punto (por ejemplo, graficamos z-scores)

CARTO Earth



ColorBrewer PiYG



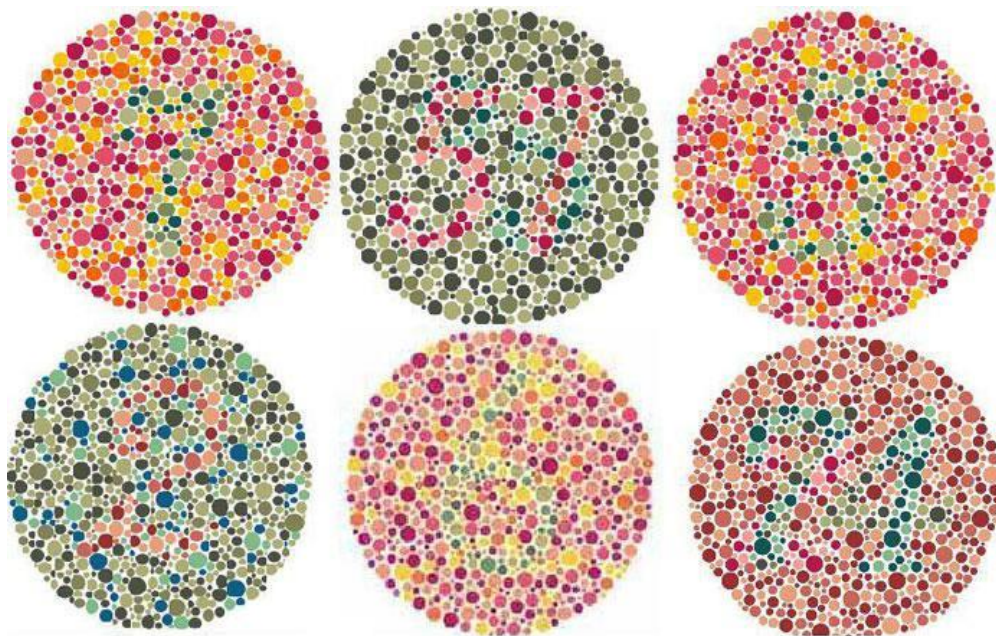
Blue-Red





# Colores para Todxs

Ojo, no todxs podemos ver todos los colores por igual



# Colores para Todxs

Ojo, no todxs podemos ver todos los colores por igual

Es importante conocer nuestra audiencia para poder tener en cuenta esta (y muchas otras consideraciones)

Existen repositorios que generan las escalas de colores aptas para todo público:

- <https://colorbrewer2.org/>
- <http://www.gisandbeers.com/mapas-de-diseno-para-daltonicos/>

# Colores para Todxs

Ojo, no todxs podemos ver todos los colores por igual

Es importante conocer nuestra audiencia para poder tener en cuenta esta (y muchas otras consideraciones)

Existen repositorios que generan las escalas de colores aptas para todo público:

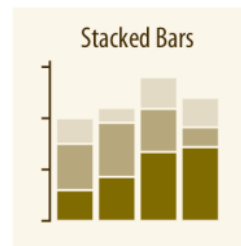
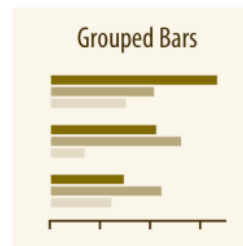
- <https://colorbrewer2.org/>
- <http://www.gisandbeers.com/mapas-de-diseno-para-daltonicos>

AH! Usar también las paletas de colores, por ejemplo, la de esta presentación

# Algunos Ejemplos - Cantidades y Categorías

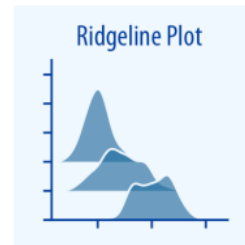
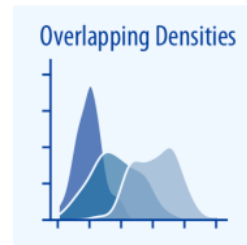
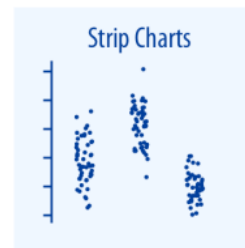
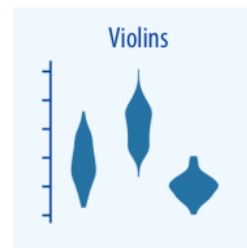
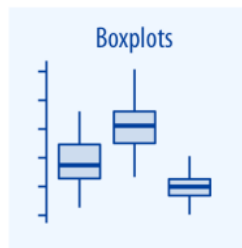
Tenemos una cantidad, cuyo valor depende de distintas categorías

En general, alguna de los datos categóricos es ordenable (ej. fechas)



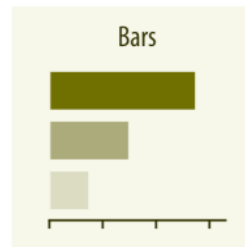
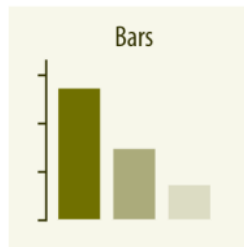
# Algunos Ejemplos - Distribuciones

Tenemos distribuciones de ciertos campos, y podemos llegar a tenerlos diferenciados según categorías



# Algunos Ejemplos - Proporciones

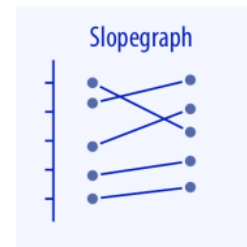
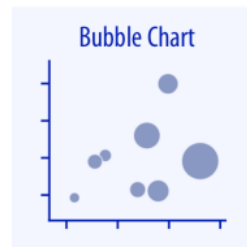
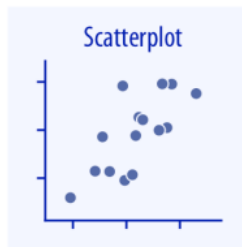
Similar al caso de cantidades/categorías, pero en los valores no dependen de la categoría



# Algunos Ejemplos - X vs Y

Tenemos datos numéricos y queremos graficar cómo se relacionan

Si los pares  $(x,y)$  pueden diferenciarse por categorías, podemos marcarlo con colores o formas.



# Algunos Ejemplos en Práctica

Ahora sí, vayamos al colab...

Para más de estos ejemplos:

- <https://clauswilke.com/dataviz/>
- <https://www.python-graph-gallery.com/>