

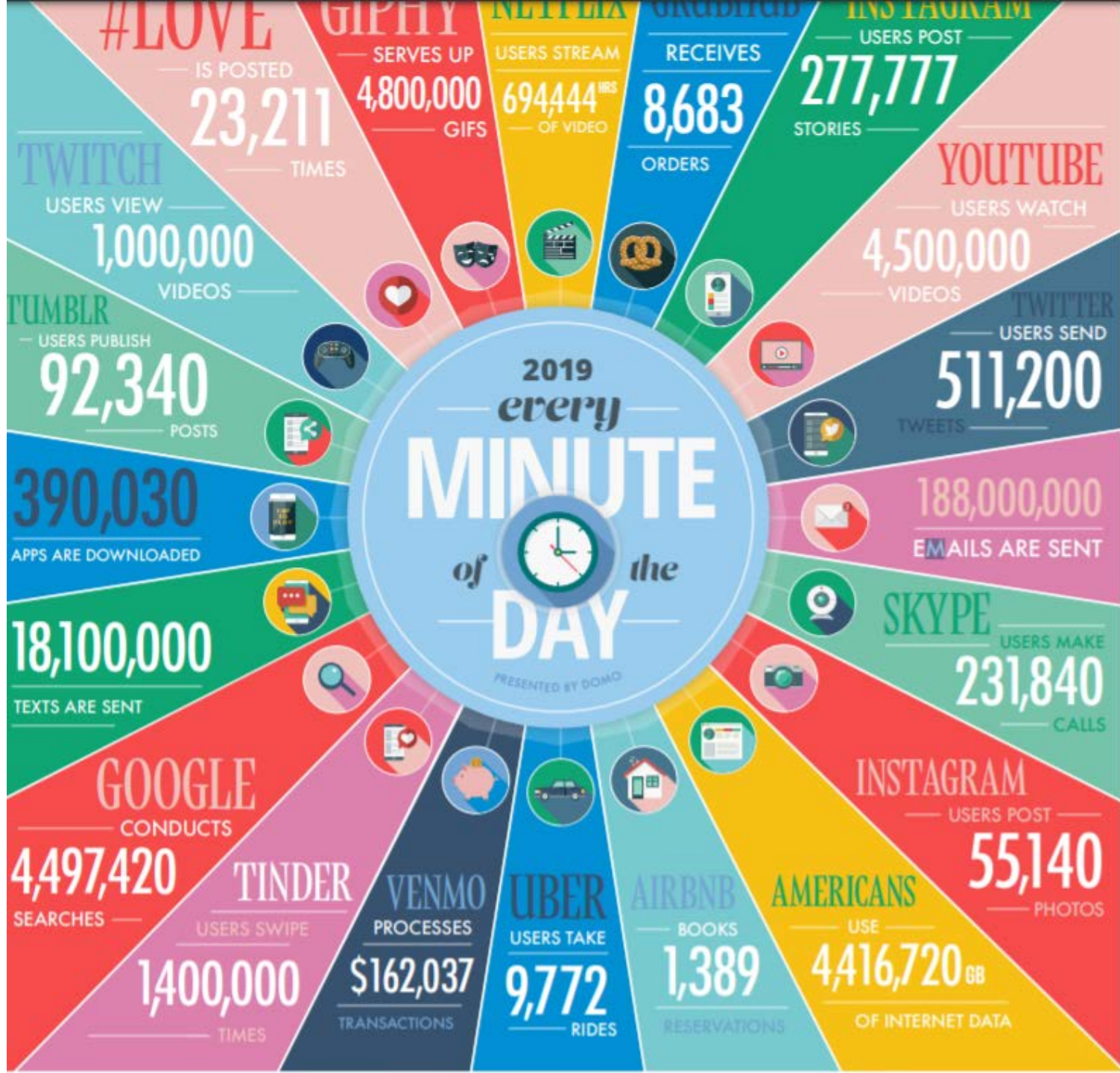
# **La ciencia y el arte de visualización de datos (parte 1)**

**Johan Van Horebeek**

**CIMAT**

**Julio 2019**

# D A T A T O S



**D**

**Cantidades enormes**

**A**

**T**

**O**

**S**

**D**

**Cantidades enormes**

**Crecimiento rápido**

**A**

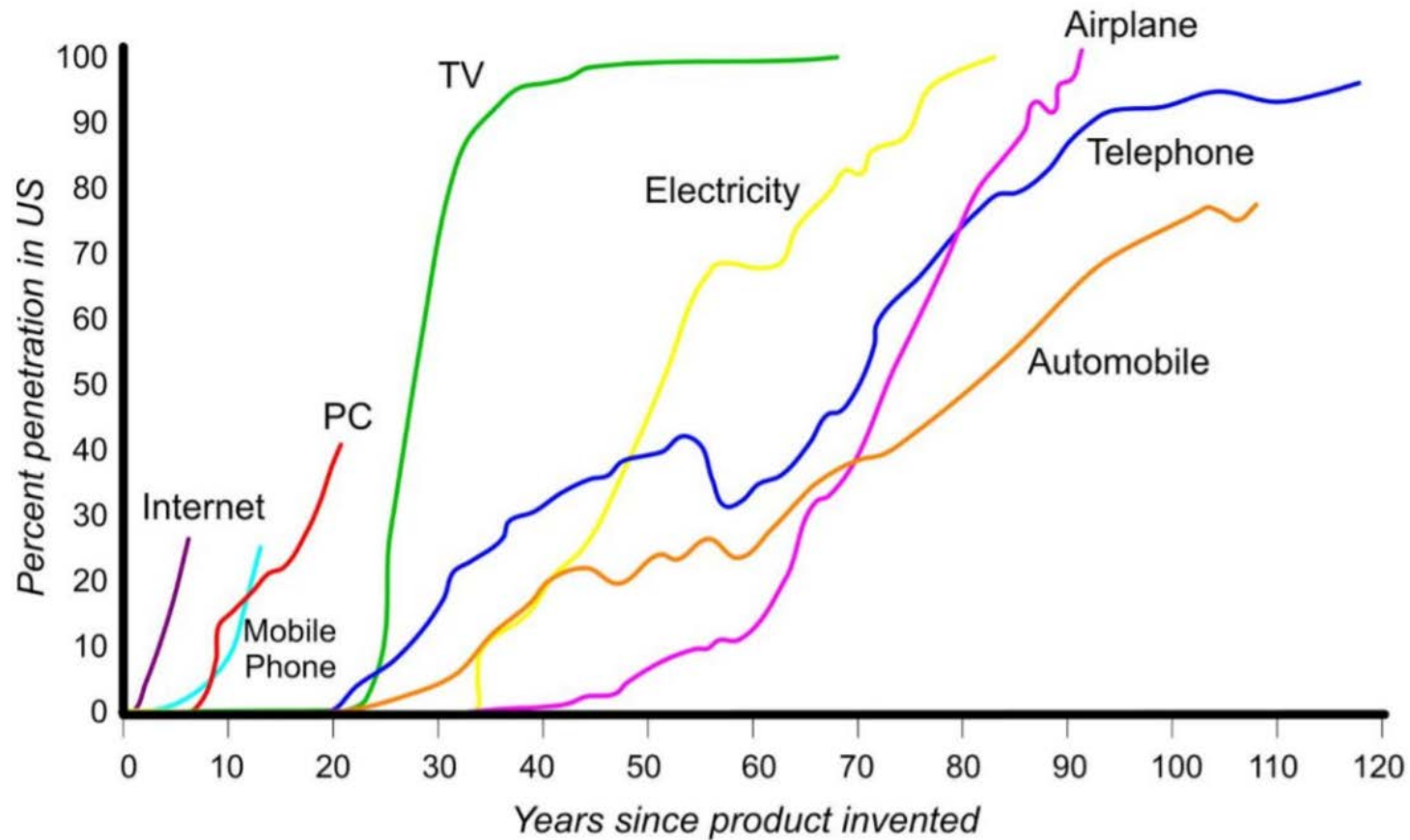
**T**

**O**

**S**

# D A T O S

**Cantidades enormes**  
**Crecimiento rápido**



# ¿Cómo aprovechar datos?



# ¿Cómo aprovechar datos?

- Para tomar una decisión
- Para predecir
- Para *demostrar* algo
- Para entender mejor el mundo

# ¿Cómo aprovechar datos?

- Para tomar una decisión
- Para predecir
- Para *demostrar* algo
- Para entender mejor el mundo

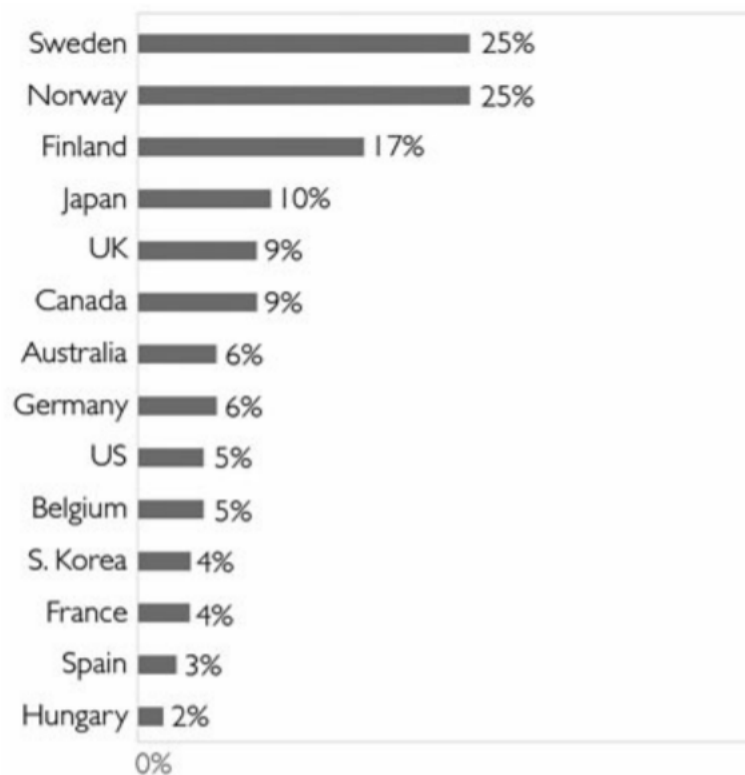
**OJO:**

**1. Seguido pensamos que conocemos bien los datos pero no es así**



En los últimos 20 años la proporción de personas que viven en pobreza extrema:

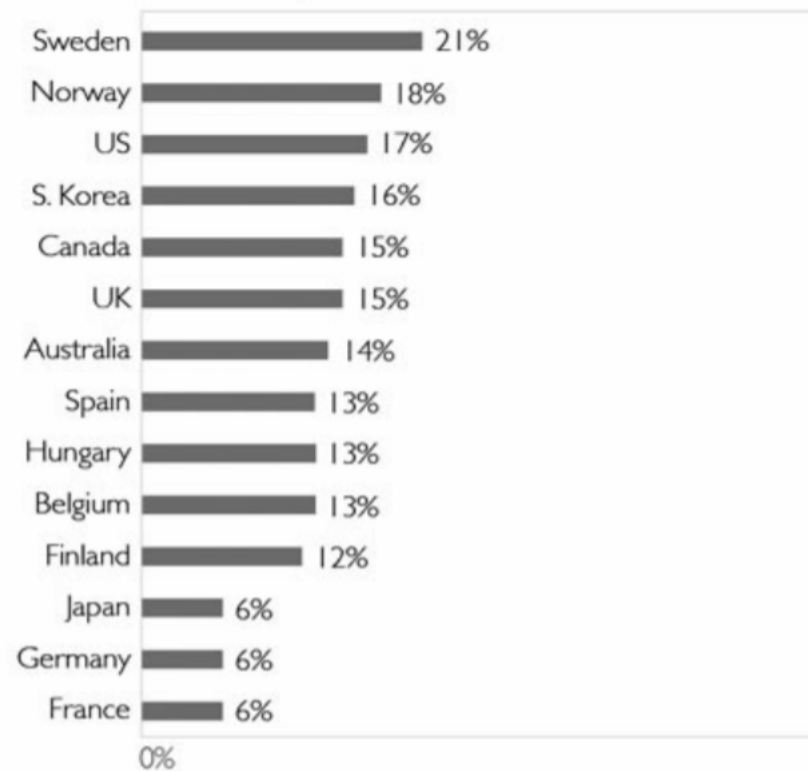
1. Se duplicó
2. Ha quedado más o menos igual
3. Se redujo a la mitad



**% respuestas correctas (libro factfulness)**

¿Cuál es el porcentaje de niños de 1 años que han sido vacunados en el mundo?

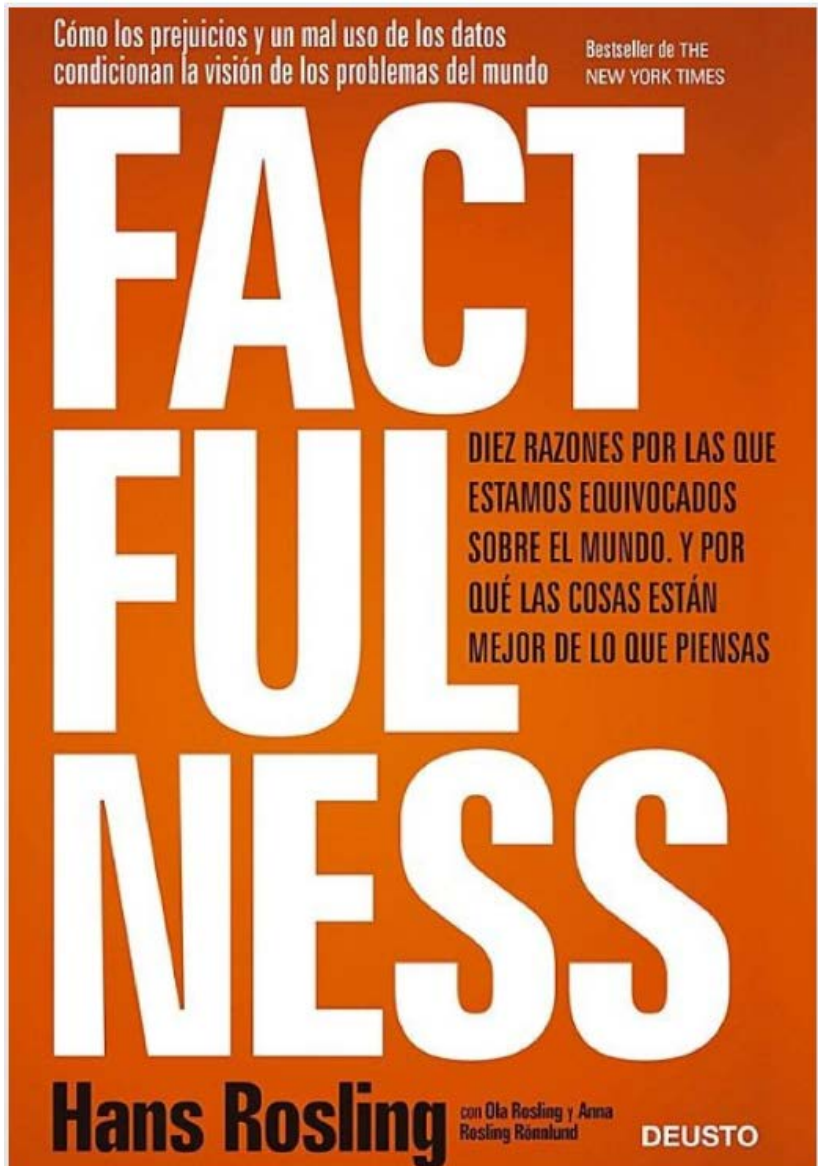
1. 20%
2. 50%
3. 80%



**% respuestas correctas (libro factfulness)**

**OJO:**

## 1. Seguido pensamos que conocemos bien los datos pero no es así



Algunas razones:

- Pesimismo
- Polarización
- Extrapolación equivocada
- Miedo
- Etc etc

## **OJO:**

- 1. Seguido pensamos que conocemos bien los datos pero no es así**
- 2. Seguido pensamos que tenemos los datos buenos pero no es así**

## **OJO:**

- 1. Seguido pensamos que conocemos bien los datos pero no es así**
- 2. Seguido pensamos que tenemos los datos buenos pero no es así**

**¿Cuántos niños en promedio hay en una familia Guanajuatense?**

**¿Para quién votarás en las siguientes elecciones? (encuesta por teléfono)**

**¿Cuánto tiempo quedan los turistas en Guanajuato?**

## **OJO:**

- 1. Seguido pensamos que conocemos bien los datos pero no es así**
- 2. Seguido pensamos que tenemos los datos buenos pero no es así**
- 3. Seguido visualizamos los datos para entenderlos pero con errores**

## **OJO:**

- 1. Seguido pensamos que conocemos bien los datos pero no es así**
- 2. Seguido pensamos que tenemos los datos buenos pero no es así**
- 3. Seguido visualizamos los datos para entenderlos pero con errores**

**Visualizar:** para **entender** mejor los datos  
y para motivar a estudiarlos

**¿Lo que observamos en la visualización es  
característica intrínseca,  
coincidencia o artefacto del diseño ?**

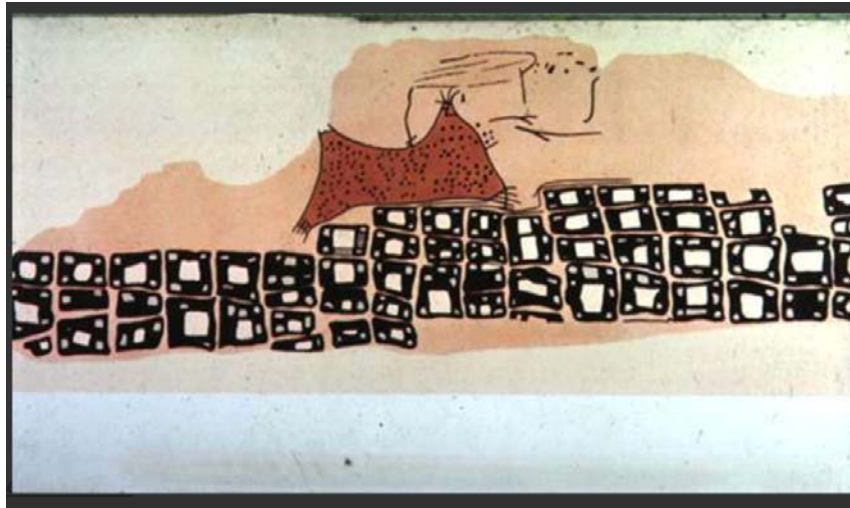
**Visualizar:** para **entender** mejor los datos  
y para motivar a estudiarlos

Algo de todos los tiempos:



**Visualizar:** para **entender** mejor los datos  
y para motivar a estudiarlos

Algo de todos los tiempos:



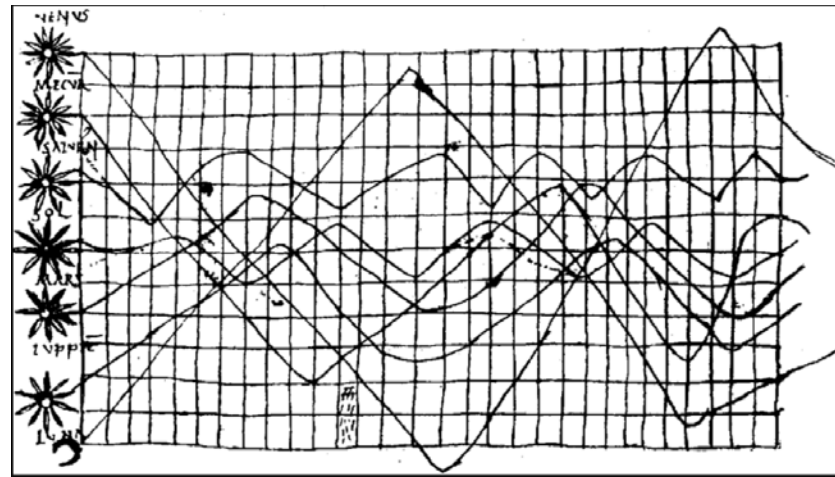
6500 AC

Mapa de ciudad

(<http://datavis.ca/>)

**Visualizar:** para **entender** mejor los datos  
y para motivar a estudiarlos

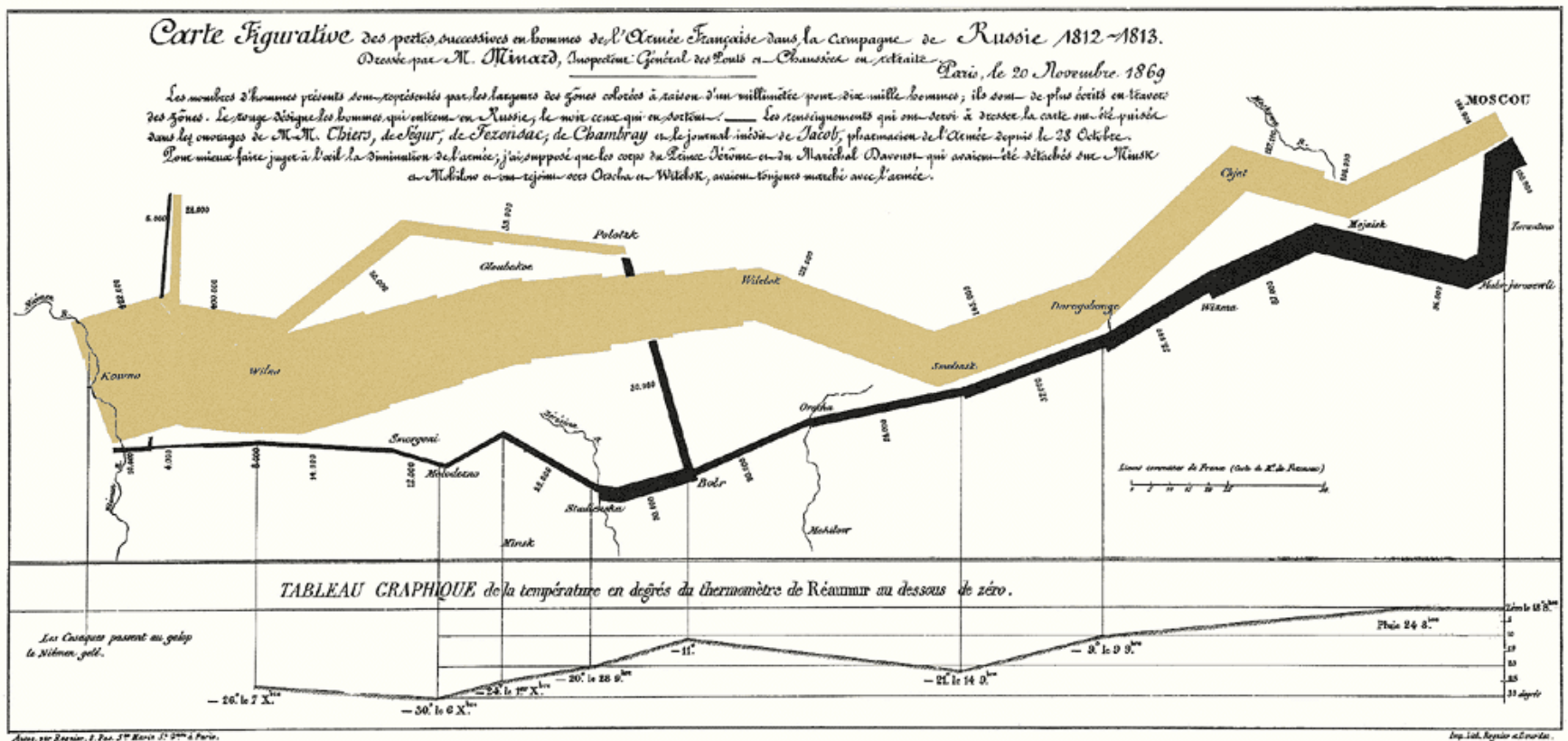
Algo de todos los tiempos:



950 DC  
Movimientos planetas  
(<http://datavis.ca/>)

# Visualizar: para entender mejor los datos y para motivar a estudiarlos

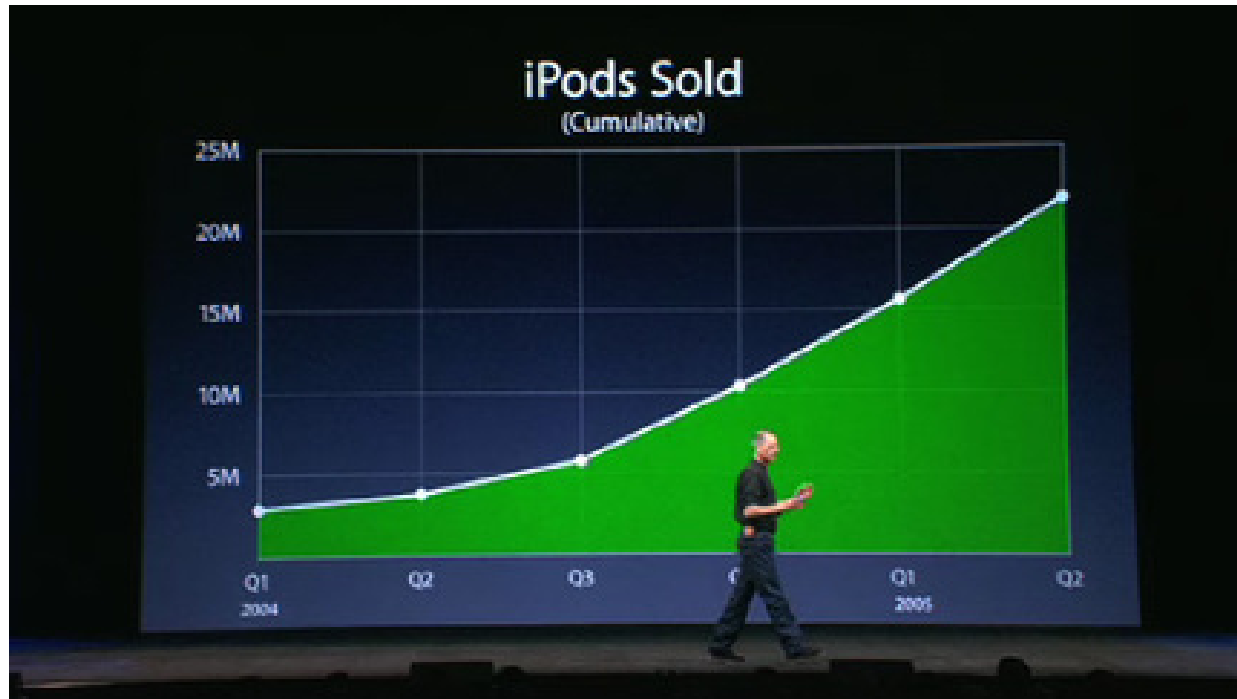
Algo de todos los tiempos:



Siglo XVII  
Invasión de Napoleon en Rusia

**Visualizar:** para **entender** mejor los datos  
y para motivar a estudiarlos

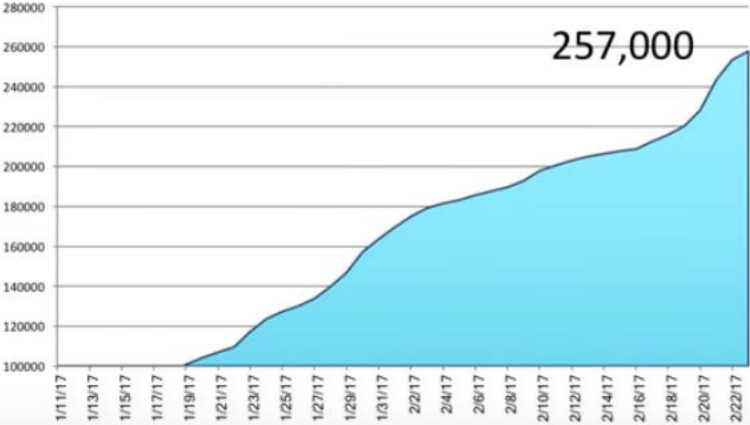
No tan evidente:



# Explosive growth at [callingbullshit.org](http://callingbullshit.org)

1-D

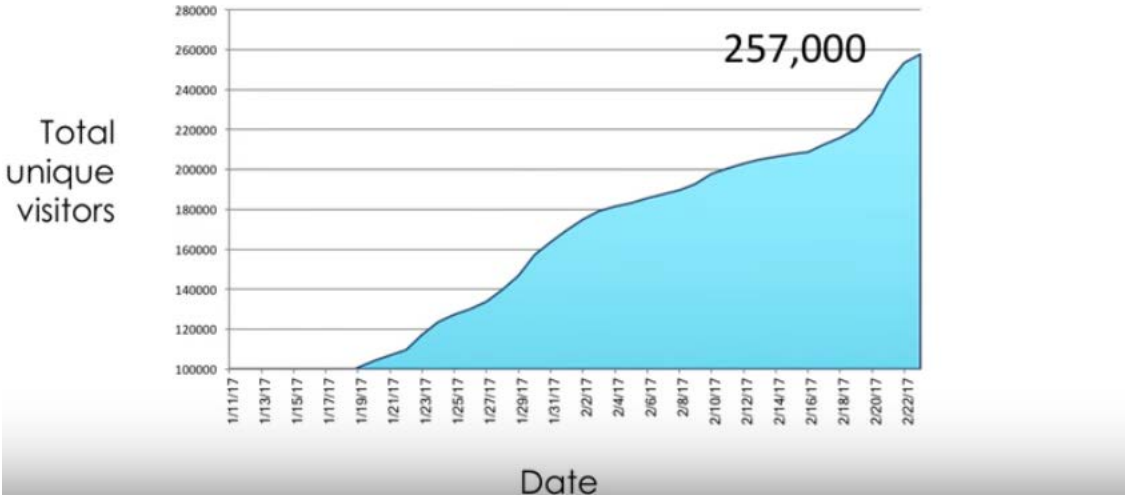
Total unique visitors



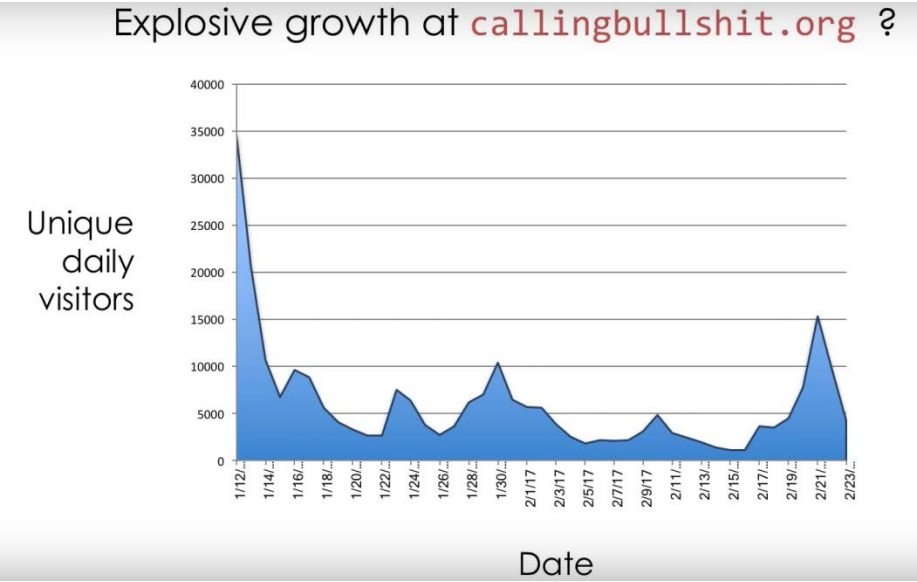
Date

# Explosive growth at callingbullshit.org

1-D

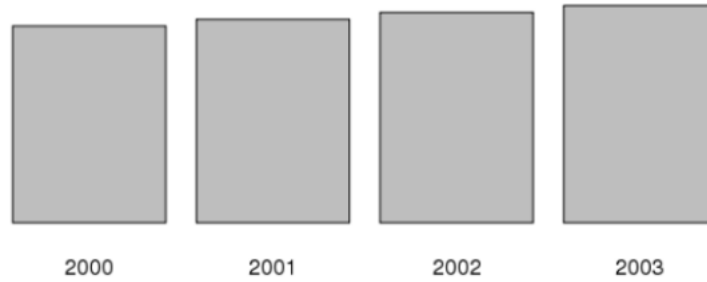


# Explosive growth at callingbullshit.org ?

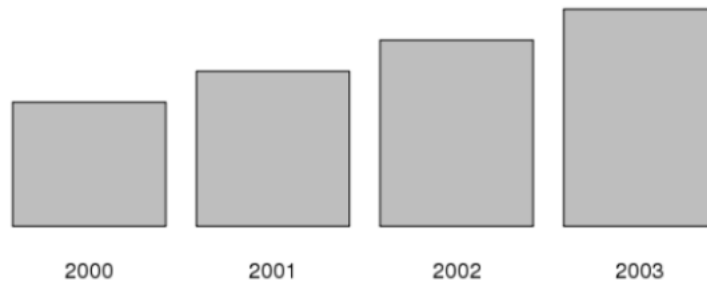


1-D

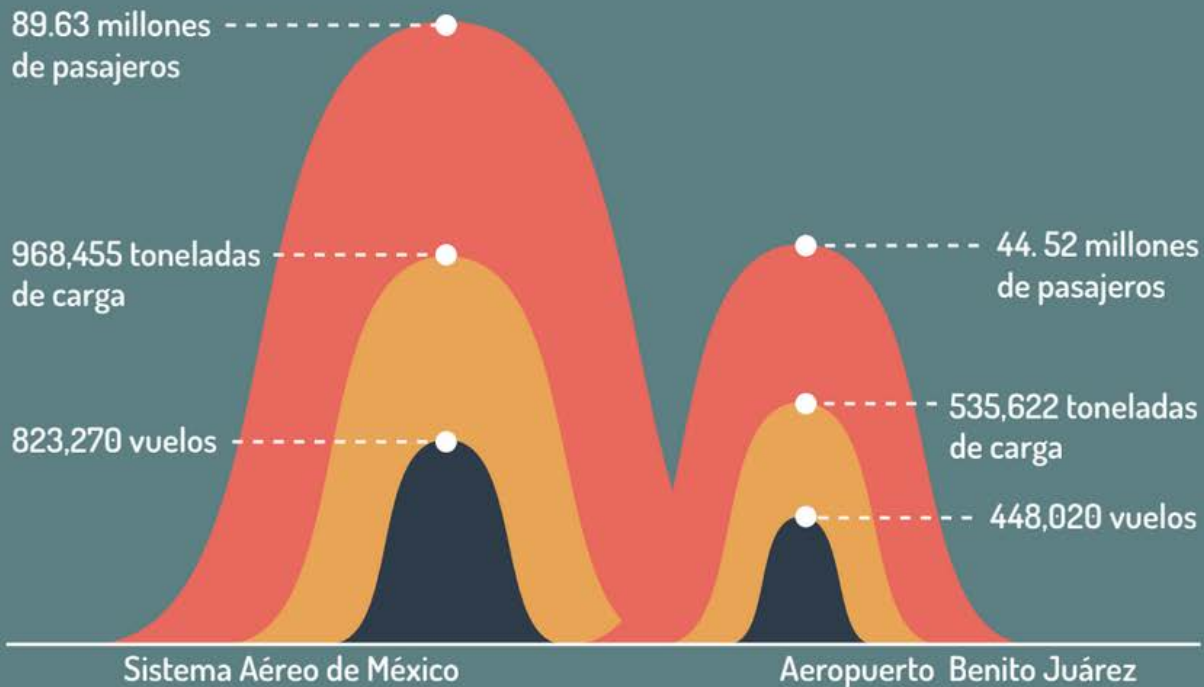
## Ventas anuales empresa A



## Ventas anuales empresa B



### Comparativo del Aeropuerto Benito Juárez y el Sistema Aéreo de México en 2017



Fuente: Balance 2017, CANAERO

**RUTAS**  
nacionales  
ternacionales



1-D

# THE SHRINKING FAMILY DOCTOR In California

Percentage of Doctors Devoted Solely to Family Practice

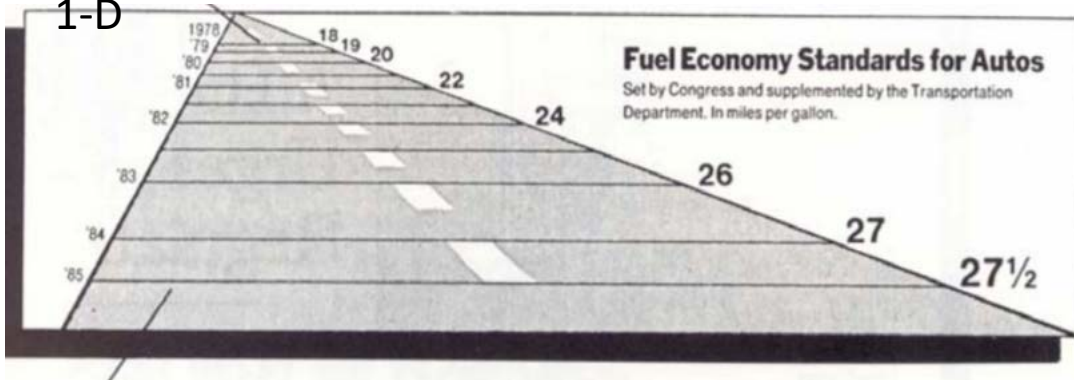
1964	1975	1990
27%	16.0%	12.0%

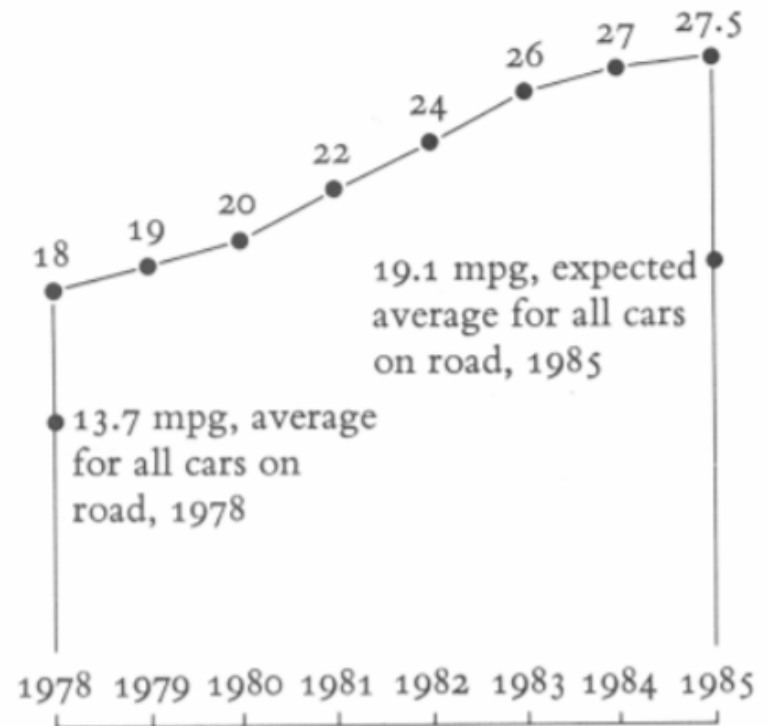
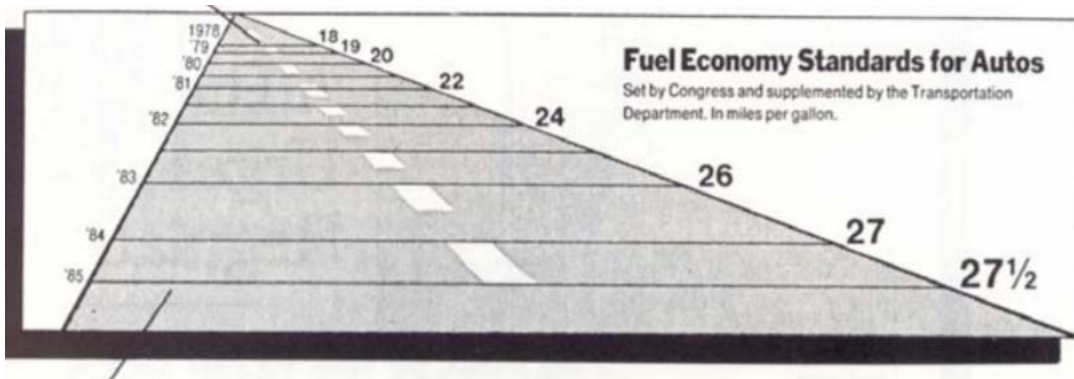


(<http://datavis.ca/>)

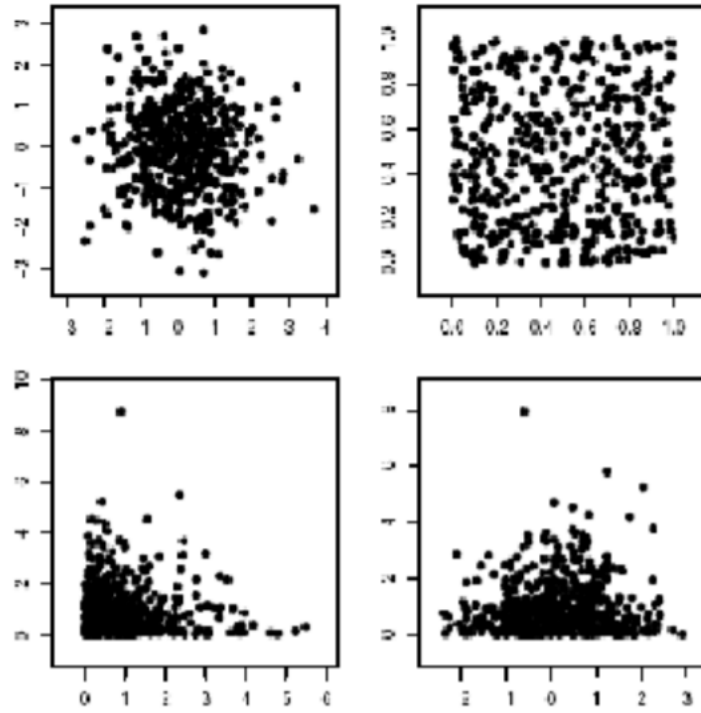
[\(http://datavis.ca/\)](http://datavis.ca/)

1-D

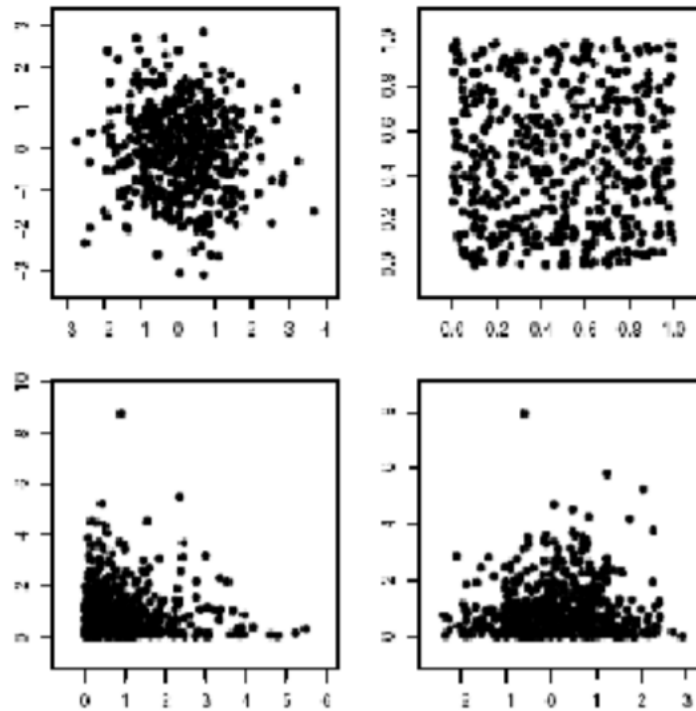




2-D



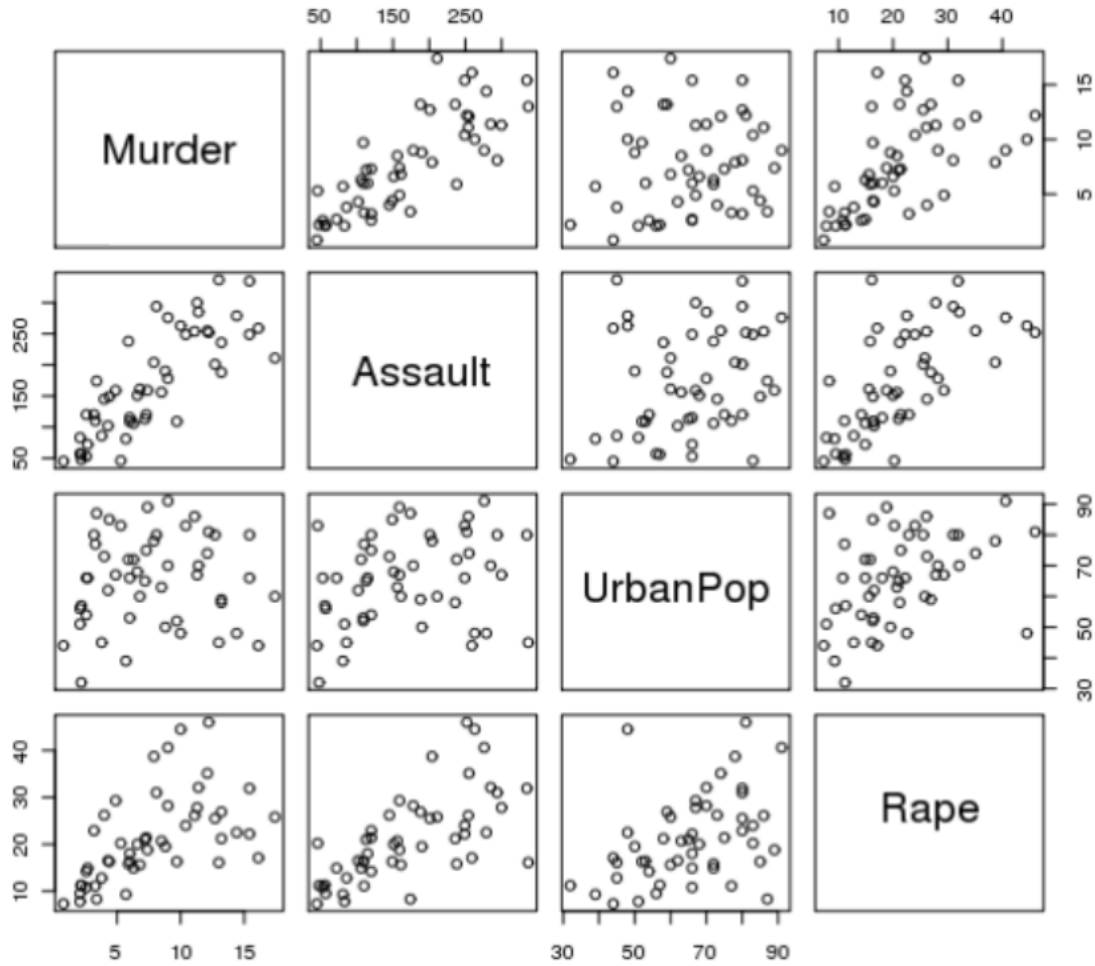
2-D



**4 veces X es independiente de Y**

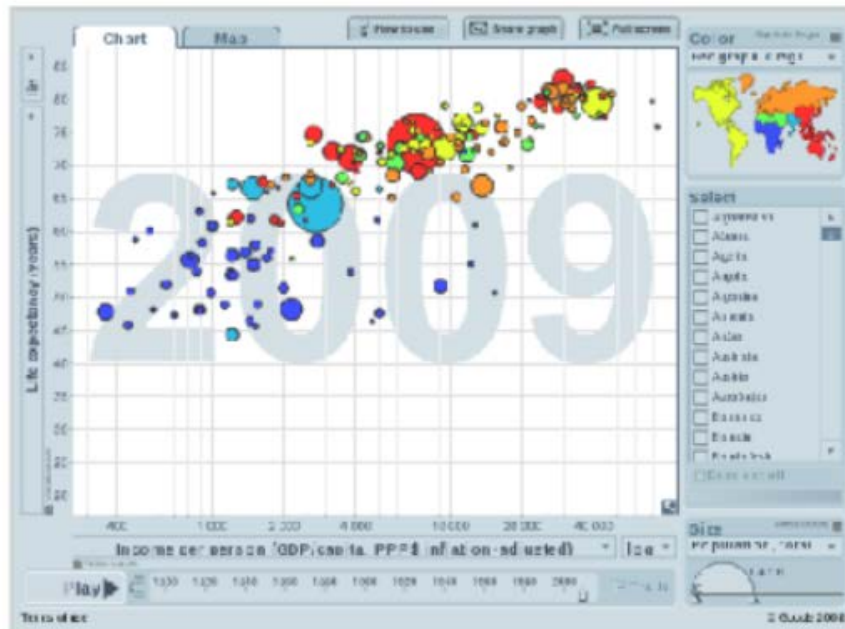
¿En más dimensiones?

¿En más dimensiones?



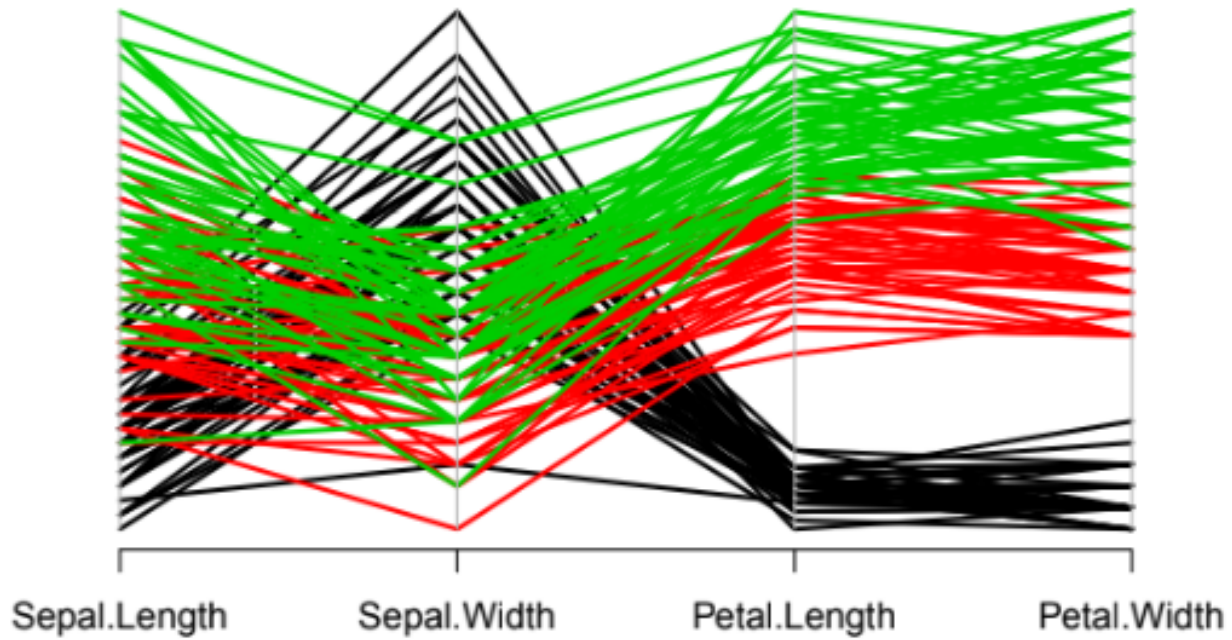
¿En más dimensiones?

## bubbleplots (Gapminder)

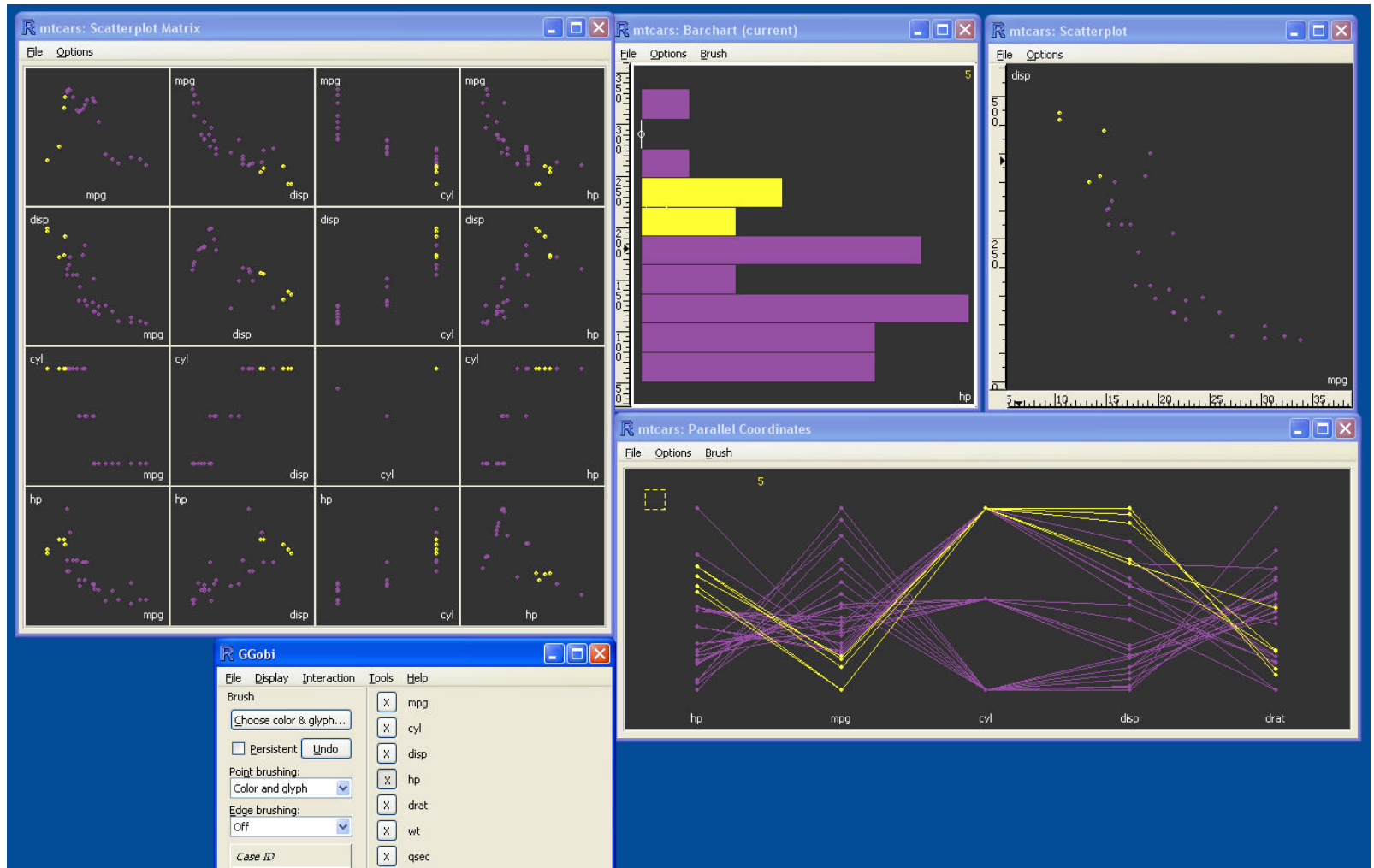




¿En más dimensiones?

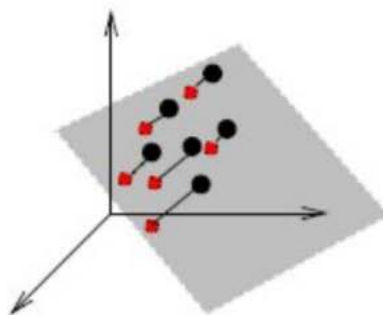
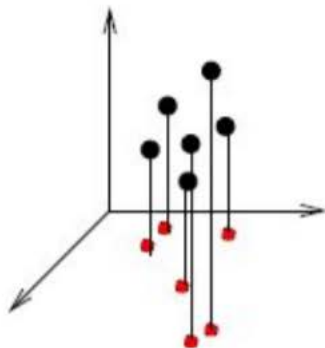


¿En más dimensiones?



ggobi

¿En más dimensiones?



## **OJO:**

- 1. Seguido pensamos que conocemos bien los datos pero no es así**
- 2. Seguido pensamos que tenemos los datos buenos pero no es así**
- 3. Seguido visualizamos los datos para entenderlos pero con errores**
- 4. Seguido confundimos causalidad con correlación**