

qué datos son estos

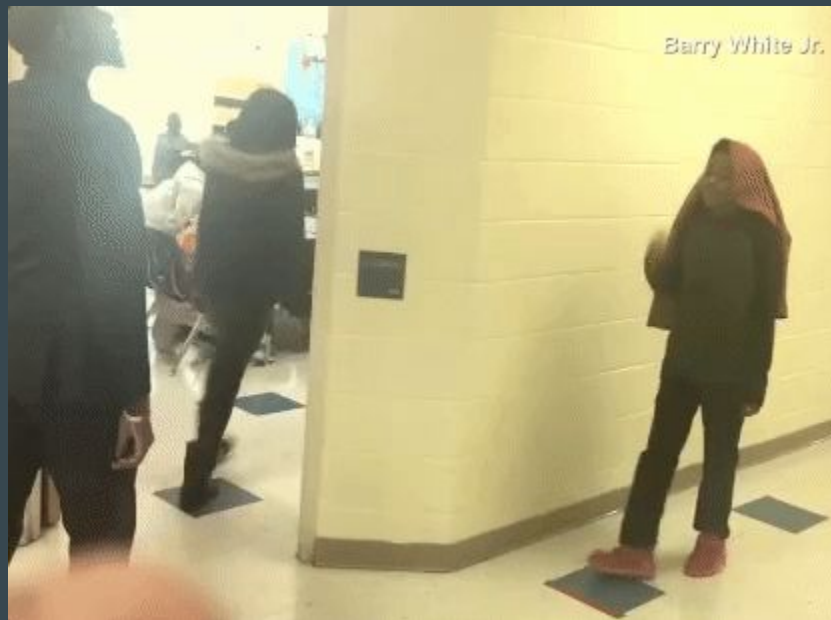




@voxdotcom mission: explain the
news, not explore the news

@JZarracina #malofiej25

conocer



qué datos son estos



qué datos son estos



-- ¿ qué representan? --

-- ¿de dónde vienen? --

-- ¿son suficientes? --

- - qué representan - -

Variables: ¿qué información recogen? (género, edad, población, evolución temporal....) **significado:** el paro: SEPE vs. EPA

Ámbito geográfico: los datos sólo son válidos para el ámbito geográfico en el fueron creados. ¿coincide con el mío?

Fecha: a qué momento se refieren los datos, ¿qué ‘validez temporal’ tienen?

- - de dónde vienen - -

A decorative graphic consisting of several vertical columns of horizontal lines. The lines are mostly grey, but some columns have a single line highlighted in red, green, or yellow.

Metodología: si cambia la metodología, los resultados no tienen nada que ver. La fiabilidad tampoco.

- es un censo: recuento de todos los casos.
- un muestreo: pregunta a unos pocos ‘representantes’ elegidos.
- autoadministrada: se presentan voluntarios.

-- son suficientes --

Cantidad, compatibilidad y comparabilidad

-- ¿ qué representan? --

Variables

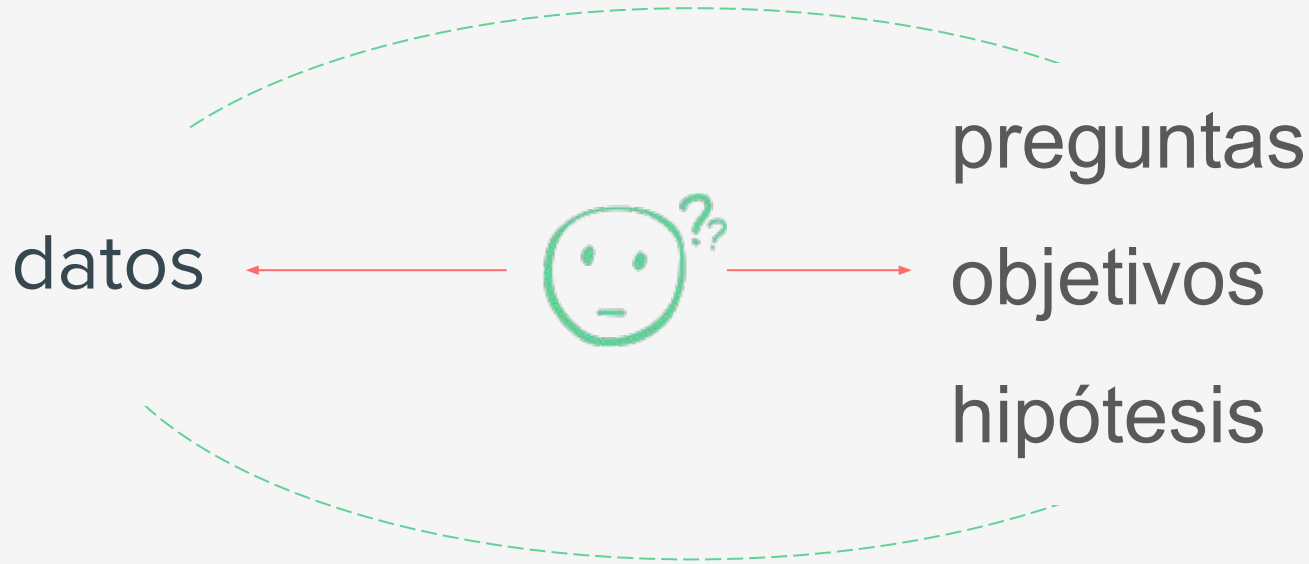
Ámbito geográfico

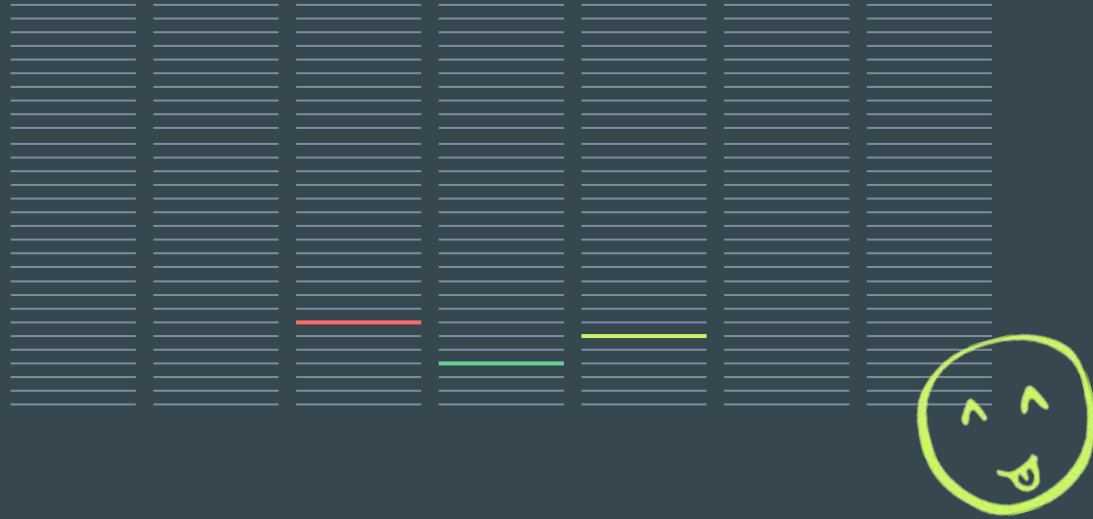
Fecha

-- ¿de dónde vienen? --

Metodología

- - son suficientes - -







@voxdotcom mission: explain the
news, not explore the news

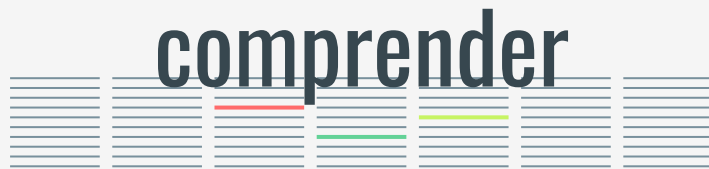
@JZarracina #malofiej25

comprender





comprender


A decorative graphic consisting of several vertical columns of horizontal lines. The central columns have one line highlighted in red, green, and yellow respectively.

----- objetivos del proyecto -----

se trata de responder preguntas...
y de hacerse las preguntas adecuadas

- Buscar los datos más altos / los más bajos.
- Cuáles son los más habituales.
- Ver las tendencias en el tiempo: cuáles suben, bajan, son planos, tienen picos.... ¿por qué? Cuales cambian más rápido.
- Cómo se comportan por grupos (hombre - mujeres; rural - urbano; jóvenes - adultos - mayores....)

comprender

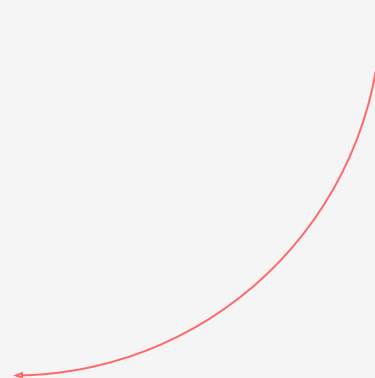
A decorative graphic consisting of several vertical columns of horizontal lines. The lines are mostly grey, but some are highlighted in red, green, and yellow.

Esta es la parte de los conceptos básicos de estadística !!

-- contar --

-- relativizar --

-- comparar / contextualizar --





Frecuencias y tablas de contingencia.

¿cuántos?

En Arcos de la Sierra (Cuenca) hay 45 mujeres.

En Arcos de la Sierra (Cuenca) hay 30 afiliados a la Seguridad Social.

En Arcos de la Sierra (Cuenca) hay 27 extranjeros.

por grupos, por cruces ¿cuántos?

Hay 12 mujeres afiliadas a la Seguridad Social.

Hay 8 mujeres extranjeras.



Datos absolutos y relativos: porcentajes y tasas

Absolutos

En Arcos de la Sierra (Cuenca) hay 45 mujeres.

Hay 12 mujeres afiliadas a la Seguridad Social.

Hay 8 mujeres extranjeras.

Relativos (porcentajes)

Hay 45 mujeres de 80 habitantes - 56%

Hay 45 mujeres de 300 habitantes - 15%

Relativo a qué (porcentajes)

8 extranjeras de 45 mujeres - 17% de las mujeres son extranjeras.

8 mujeres de 27 extranjeros - 30% de los extranjeros son mujeres.



Datos absolutos y relativos: porcentajes y tasas

Evita hacer de todo un mapa de población

Coches por provincia: Madrid, Barcelona, Valencia, Bilbao....

Árboles en capitales de provincia: Madrid, Barcelona, Valencia, Bilbao....

km de fibra óptica: Madrid, Barcelona, Valencia, Bilbao....

Gatos: Madrid, Barcelona, Valencia, Bilbao....

Drones: Madrid, Barcelona, Valencia, Bilbao....

Prácticamente todo lo que podáis imaginar: Madrid, Barcelona, Valencia, Bilbao....

- - comparar / contextualizar - -

Medidas de tendencia central: media, moda y mediana

Comparar

En Arcos de la Sierra el 15 % son mujeres, en Villalpando de Arriba, pueblo de Cuenca de similar población, el 90% de la población es mujer ¿por qué? ¿qué pasa en otros sitios?

Contextualizar (medidas de tendencia central)

En Arcos de la Sierra el 15% son mujeres, la media nacional en el ámbito rural es del 47% ¿qué pasa en Arcos de la Sierra?

Medidas de tendencia central:

media, mediana y moda

media

El centro de la distribución, el punto de equilibrio.

mediana

La mitad de la distribución

moda

El valor que más se repite.

Medidas de tendencia central:
media, *mediana* y *moda*

moda

El valor que más se repite.

• mode - most frequent number
3

1, 1, 1, 1, 2, 3, 20 = 1

Medidas de tendencia central:
media, *mediana* y *moda*

media

El centro de la distribución, el punto de equilibrio.

2, 3, 3, 3, 4, 4, 10

Average

• arithmetic mean:
$$\frac{2+3+3+3+4+4+10}{7} = \frac{29}{7} = \left[4\frac{1}{7}\right]$$

$$1, 1, 1, 1, 2, 3, 20 = 29 / 7$$

Medidas de tendencia central:
media, mediana y moda

mediana

La mitad de la distribución.

• median - "middle"
2, 3, 3 | (3) | 4, 4, 10

12, 61, 64, 68, 72, 73, 74 = 68

Medidas de tendencia central: *media*, *mediana* y *moda*

2, 3, 3, 3, 4, 4, 10, 1,000,000
outlier

Average

* arithmetic mean: $\frac{2+3+3+3+4+4+10}{7} = \frac{29}{7} = 4\frac{1}{7}$

* median - "middle" - 3
2, 3, 3 | 3 | 4, 4, 10

* mode - most frequent number
3

2, 3, 4, 5
 $\frac{3+4}{2} = 3.5$

2+3+4+5 = 14
4

<https://www.youtube.com/watch?v=GrynkZB3E7M>

www.khanacademy.org

visualizados.com



@voxdotcom mission: explain the
news, not explore the news

@JZarracina #malofiej25

conocer



comprender





visualizados.com

Beatriz Martínez

Data visualizer

Design + Code