

Visualización de Datos y Storytelling

Magíster en Data Science - Facultad de Ingeniería UDD



Carlos Elías Pérez Pizarro

Agradecimientos: Eduardo Graells-Garrido

Método Científico

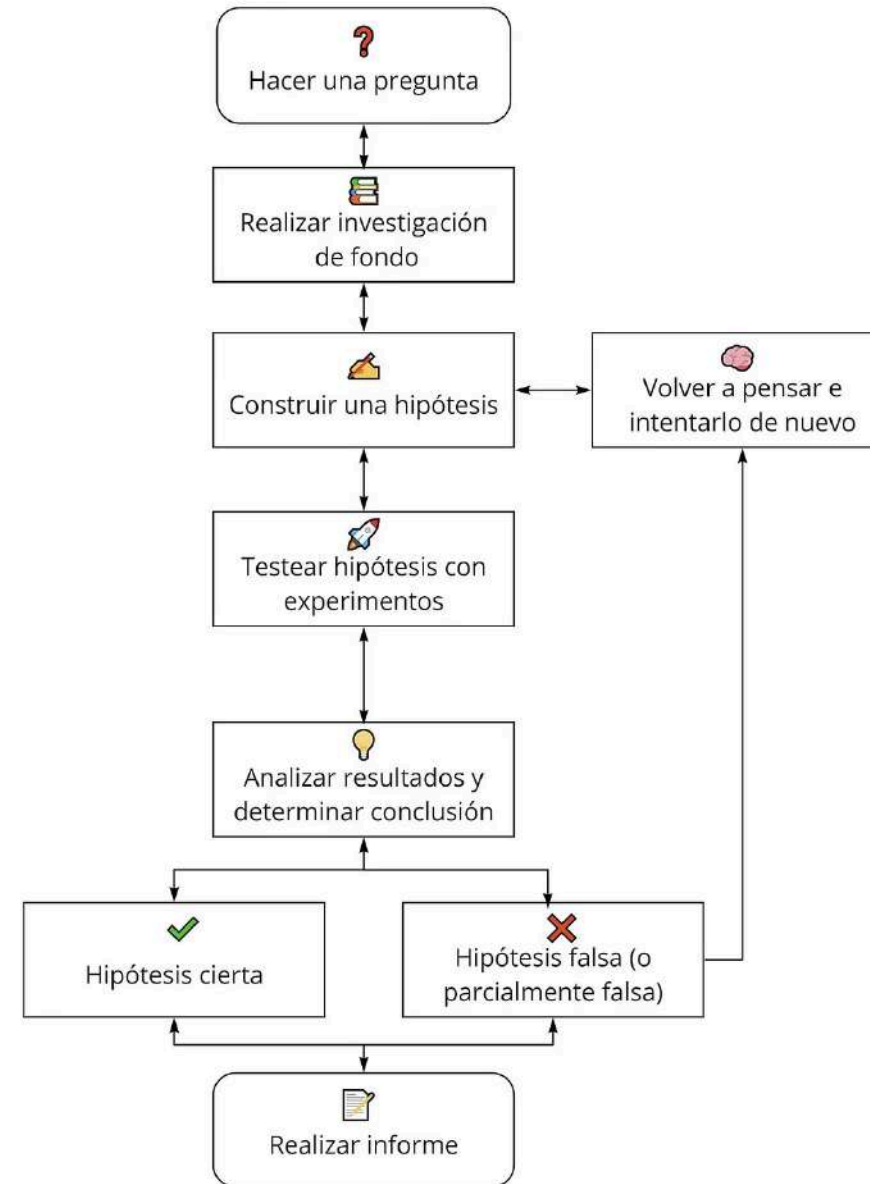
Más que el conocimiento en sí mismo,

Asociamos a la ciencia con un método

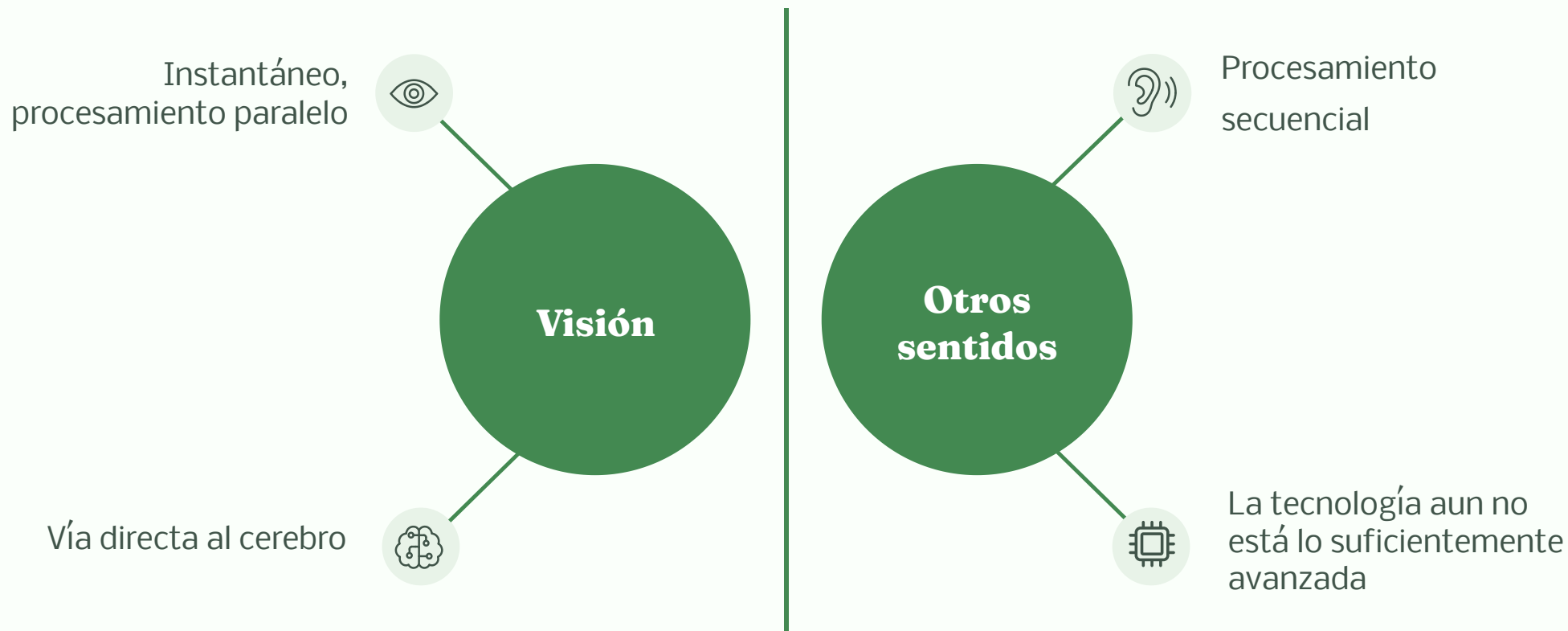
Para **conseguir dicho conocimiento**

Noten que **el método científico ya incluye datos**: son parte de los experimentos y del análisis de resultados

Modelo simplificado de las etapas del método científico

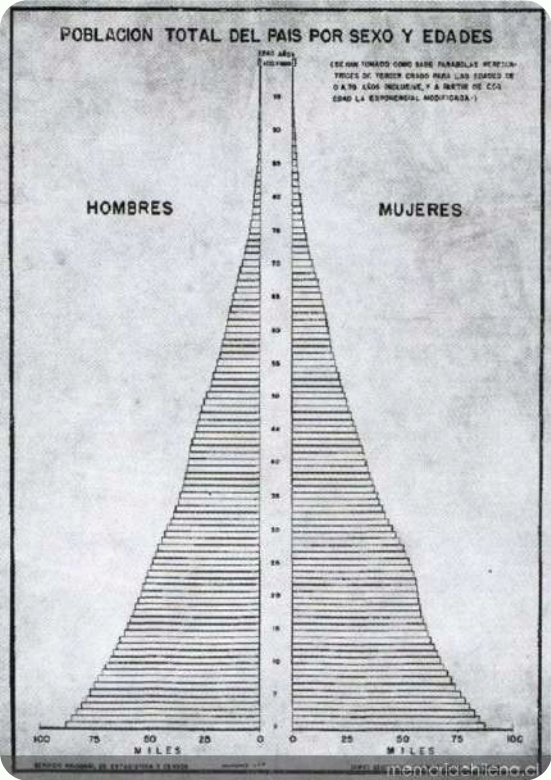


¿Por qué depender de la visión?



Datos no son algo "nuevo"

Censos de 1865 y 1940



REPÚBLICA DE CHILE.

REPÚBLICA DE CHILE.

Resúmen de las profesiones en toda la República, con especificacion del sexo.

PROFESIONES.	HOMBRES.	MUJERES.
.....	2
.....	75
.....	97
.....	516	65
.....	4
.....	106	10
.....	25
.....	206
.....	18
.....	5
.....	16
.....	4
.....	2
.....	2
.....	968
.....	55
.....	4	21
.....	9
.....	63
.....	97	83
.....	375
.....	178766	489
.....	11
.....	32
.....	3561
.....	150
.....	115	37490
.....	418
.....	148	15
.....	39	1
.....	64
.....	191	9
.....	189	3
.....	189	28
.....	3533
.....	82
.....	1	28904
.....	360	6
.....	2
.....	659	13
.....	9
.....	230	2726
.....	3812
.....	463
.....	8	488
.....	81
.....	230
.....	20625	118
.....	7081
.....	1	467
.....	806	6
.....	1	107
.....	520	310
.....	3

Siempre visualiza tus datos

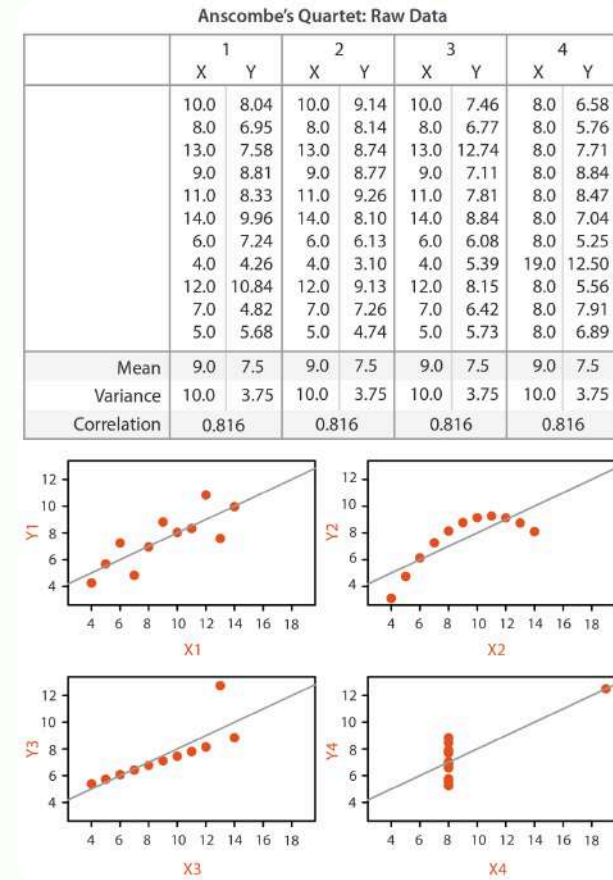
“Sin datos, eres simplemente otra persona opinando”

– W. Edwards Deming

Dataset

Un dataset es un conjunto de observaciones o registros que contiene variables y atributos, usualmente numéricos.

Derecha: **El Cuarteto de Anscombe.**
Presenta 4 subdatasets que tienen las mismas propiedades estadísticas básicas.



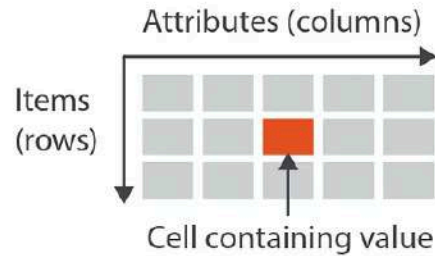
Datasaurus dozen



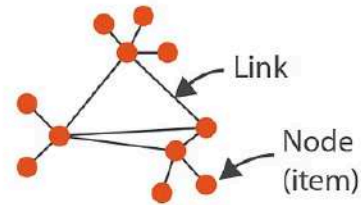
Common statistical values for each group in the dataset							
Dataset	Summary statistics					Regression results	
	Mean x	Mean y	Std Dev x	Std Dev y	Corr x y	Intercept	Coefficients
Away	54.27	47.83	16.77	26.94	-0.06	53.43	-0.10
Bulkytype	54.27	47.83	16.77	26.94	-0.07	53.81	-0.11
Circle	54.27	47.84	16.76	26.93	-0.07	53.80	-0.11
Dino	54.26	47.83	16.77	26.94	-0.06	53.45	-0.10
Dots	54.26	47.84	16.77	26.93	-0.06	53.10	-0.10
H_lines	54.26	47.83	16.77	26.94	-0.06	53.21	-0.10
High_lines	54.27	47.84	16.77	26.94	-0.07	53.81	-0.11
Slant_down	54.27	47.84	16.77	26.94	-0.07	53.85	-0.11
Slant_up	54.27	47.83	16.77	26.94	-0.07	53.81	-0.11
Star	54.27	47.84	16.77	26.93	-0.06	53.33	-0.10
V_lines	54.27	47.84	16.77	26.94	-0.07	53.89	-0.11
Wide_lines	54.27	47.83	16.77	26.94	-0.07	53.63	-0.11
X_shape	54.26	47.84	16.77	26.93	-0.07	53.55	-0.11

Tipos de Datos

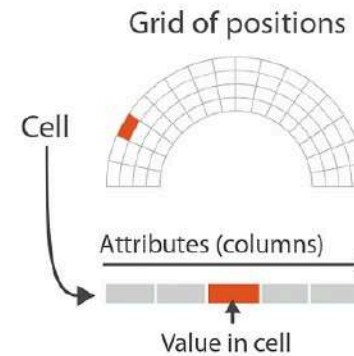
→ Tables



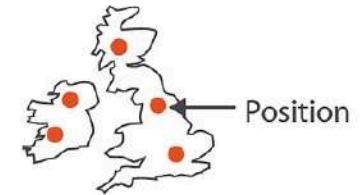
→ Networks



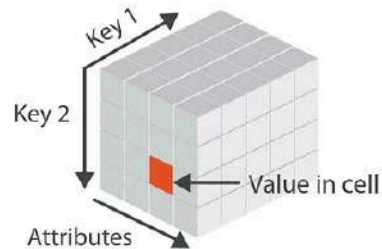
→ Fields (Continuous)



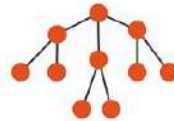
→ Geometry (Spatial)



→ Multidimensional Table

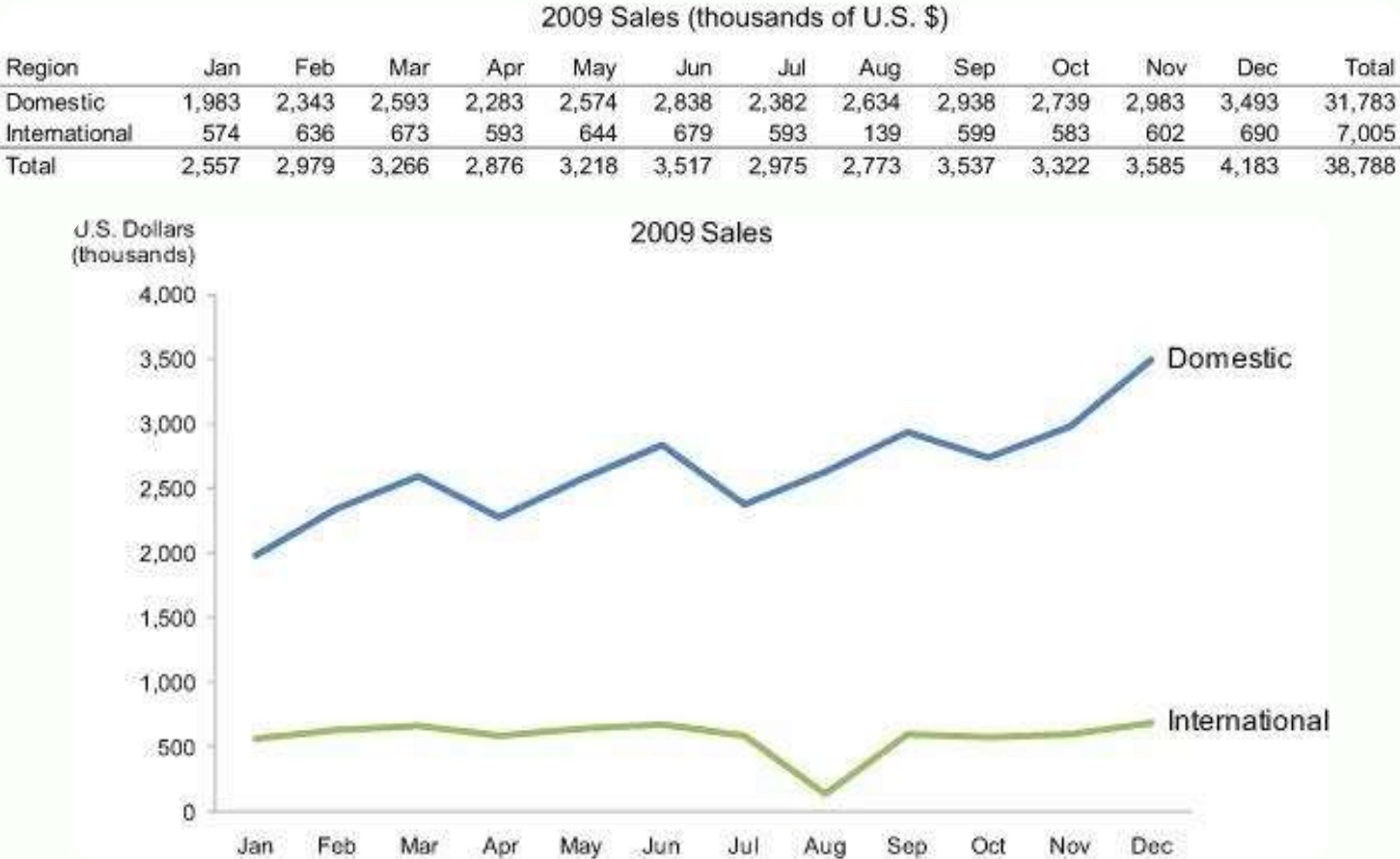


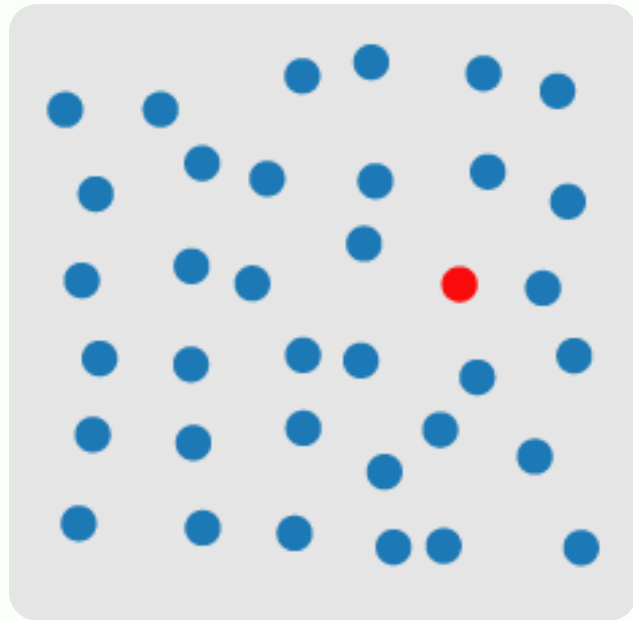
→ Trees



Nos da superpoderes

Una visualización nos puede organizar la información de manera que podamos realizar consultas de manera eficiente.





¿Dónde está el punto rojo?

Fuente:

Perception in Visualization

Christopher G. Healy

<https://www.csc2.ncsu.edu/faculty/healey/PP/>

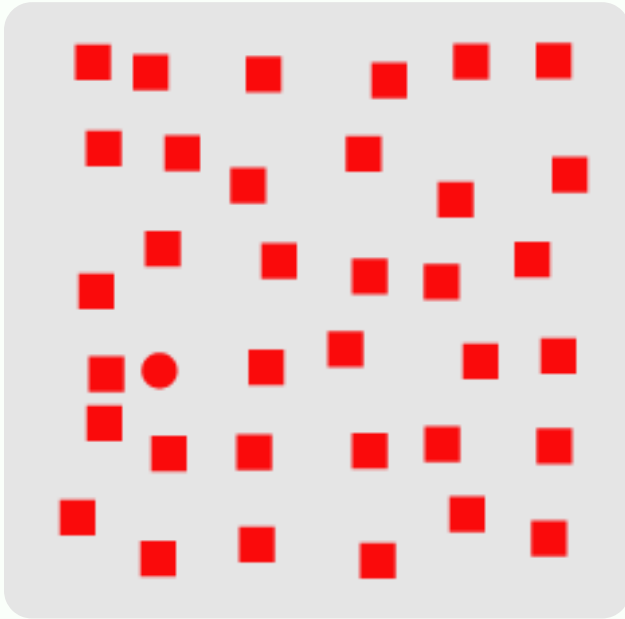
¿Dónde está el punto rojo?

Fuente:

Perception in Visualization

Christopher G. Healy

<https://www.csc2.ncsu.edu/faculty/healey/PP/>

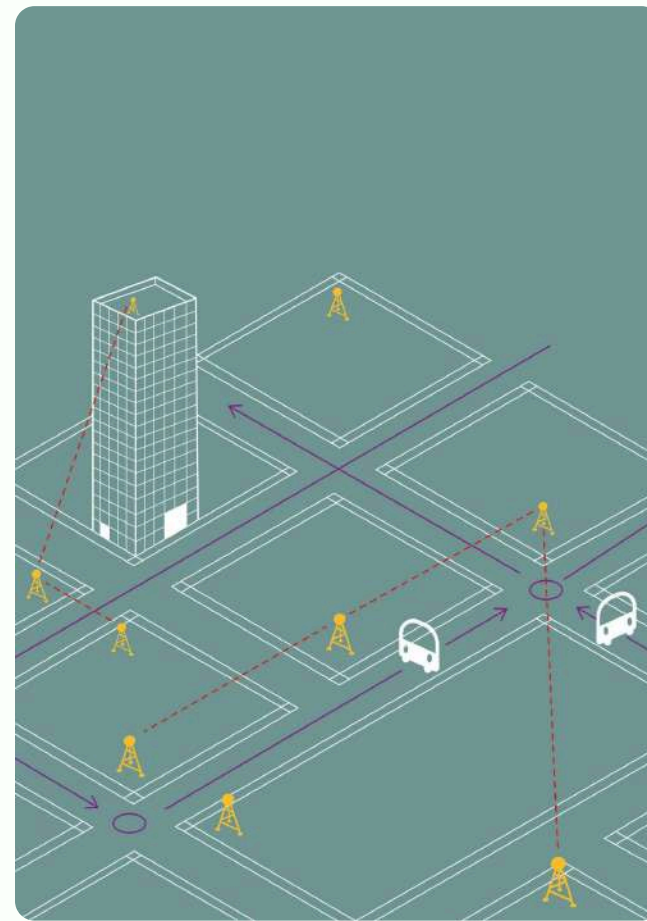


El mundo es una traza digital

Trazas Digitales



Mejora nuestro entendimiento del mundo



Muchos datos?

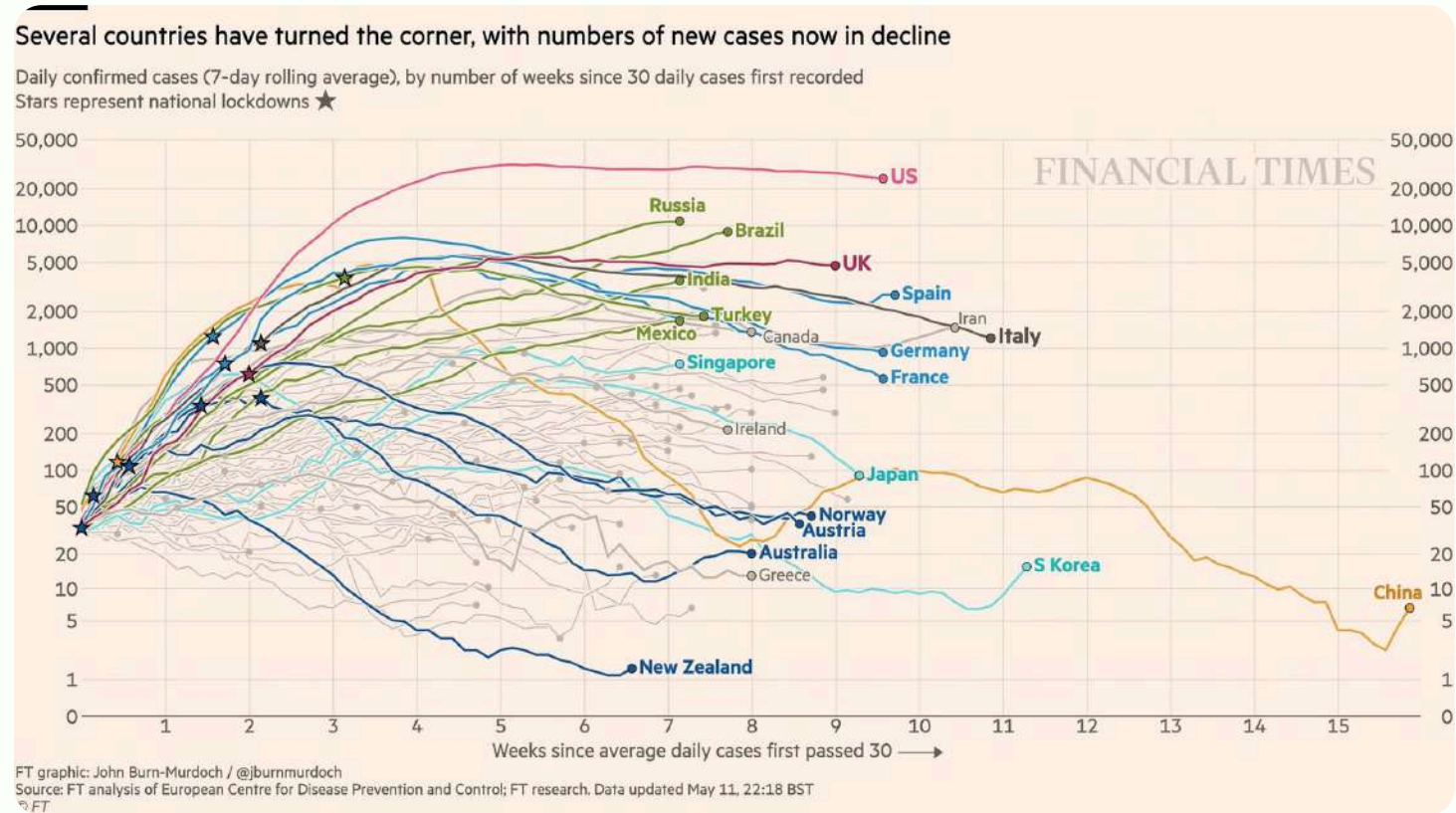


Gráfico por John Burn-Murdoch sobre la evolución del COVID-19.