

EL ARTE DE LA ESTADÍSTICA

CÓMO LOS DATOS REVELAN LA VERDAD

Desde asesinos en serie hasta la felicidad global.

Los números no pueden hablar por sí mismos. Nosotros
hablamos por ellos. Les conferimos significado.

— Nate Silver

UN RECORRIDO NARRATIVO BASADO EN LA OBRA DE DAVID SPIEGELHALTER



EL MISTERIO DEL «DOCTOR BONDADOSO»

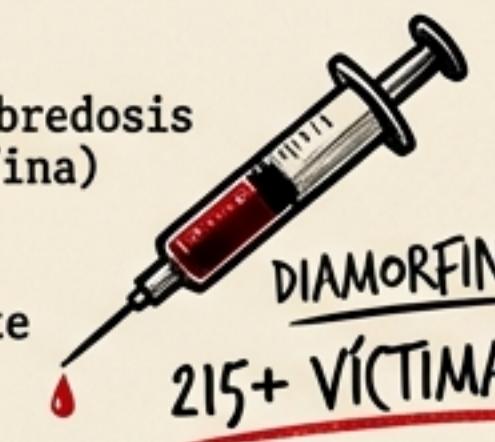
BIOGRAFÍA DEL CASO

EL PERSONAJE:

Harold Shipman era un médico de familia de maneras suaves en Gran Bretaña. No encajaba en el perfil arquetípico de un asesino en serie. Era un pilar de la comunidad.

EL CRIMEN:

Entre 1975 y 1998, inyectó sobredosis masivas de opiáceos (diamorfina) a un mínimo de 215 pacientes, mayoritariamente ancianos.



Last of Will

FALSO



LA CAPTURA:

Un error de codicia. Falsificó el testamento de una víctima. Su hija, abogada, sospechó de la repentina herencia.



EVIDENCIA DIGITAL

[TDESTMP 15:42:16]

BASE DE DATOS MÉDICA - REGISTROS CLÍNICOS

PACIENTE: [ELFACR800 P08 REHOCRMN]	FECH: 18/05/1998	DIAGNÓSTICO: INMOCITOSCLS DIRECHA CODEL	TESTIMONIO: DR. R. SHIPMAN	HORA: 15:46:00	SST00: TALLEEDO
PACIENTE: [R008 TE98RA8]	FECH: 19/05/1998	DIAGNÓSTICO: INMOCITOSCLS DIRECHA CODEL	TESTIMONIO: DR. R. SHIPMAN	HORA: 16:01:00	SST00: TALLEEDO [ALT00640]
PACIENTE: [R008 TE98RA8]	FECH: 22/05/1998	DIAGNÓSTICO: INMOCITOSCLS DIRECHA CODEL	TESTIMONIO: DR. R. SHIPMAN	HORA: 16:46:16	SST00:

PROBLEMA: [ELFACR800 P08 REHOCRMN] | FECH: 22/05/1998 | DIAGNÓSTICO: INMOCITOSCLS DIRECHA CODEL | TESTIMONIO: DR. R. SHIPMAN | HORA: 16:46:16 | MODIFICACIÓN: 15:42:16 [TDESTMP BIE1ML AGJL] | ESTADO: FALLECIDO [ALTERADO]

PACIENTE: [R008 TE98RA8]	FECH: 22/05/1998	DIAGNÓSTICO: INMOCITOSCLS DIRECHA CODEL	TESTIMONIO: DR. R. SHIPMAN	HORA: 16:46:16	SST00: FUE EJECUTADO [ALT00640]
PACIENTE: [R008 TE98RA8]	FECH: 22/05/1998	DIAGNÓSTICO: INMOCITOSCLS DIRECHA CODEL	TESTIMONIO: DR. R. SHIPMAN	HORA: 16:46:16	SST00: FUE EJECUTADO [ALT00640]

ALERTA: MODIFICACIÓN POST-NOREM DETECTADA.
SEGUINIENTE DE ACTIVIDAD REGISTRADO.

[TDESTMP 16:57:35]

[TDESTMP 17:01:45]

PRUEBAS
EGRESOS
0102TRLS - 2598

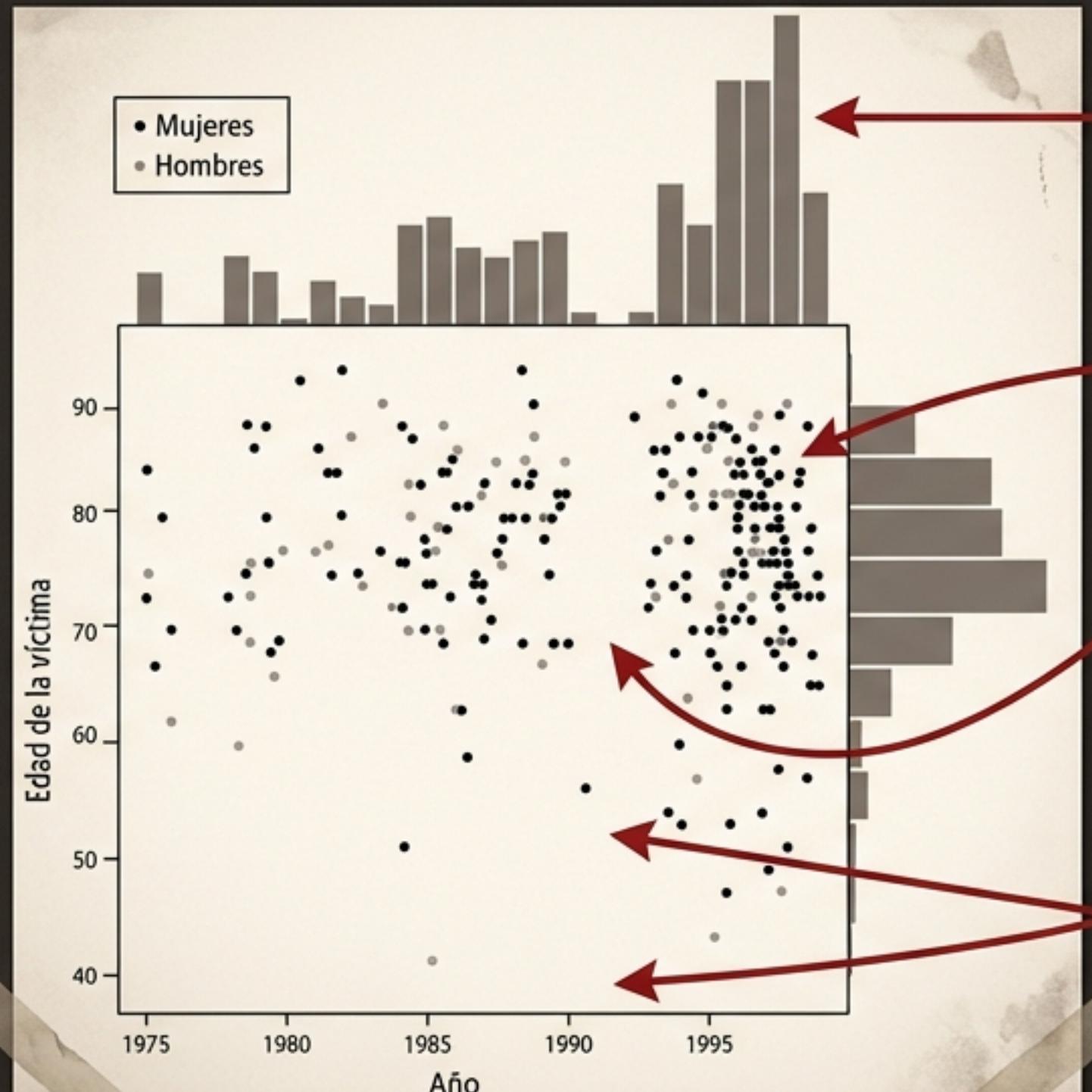


EVIDENCIA DIGITAL: Shipman modificó los historiales médicos para que sus víctimas parecieran más enfermas. La tecnología registró la fecha y hora exacta de cada falsificación.

CASO RESUELTO:
EVIDENCIA IRREFUTABLE

Visualizando el Horror

Análisis de dispersión: Edad de la víctima vs. Año de muerte



Aceleración final: La actividad criminal se intensificó drásticamente antes de su captura en 1998.

Demografía: Note la densidad de puntos negros. Las víctimas fueron mayoritariamente mujeres ancianas.

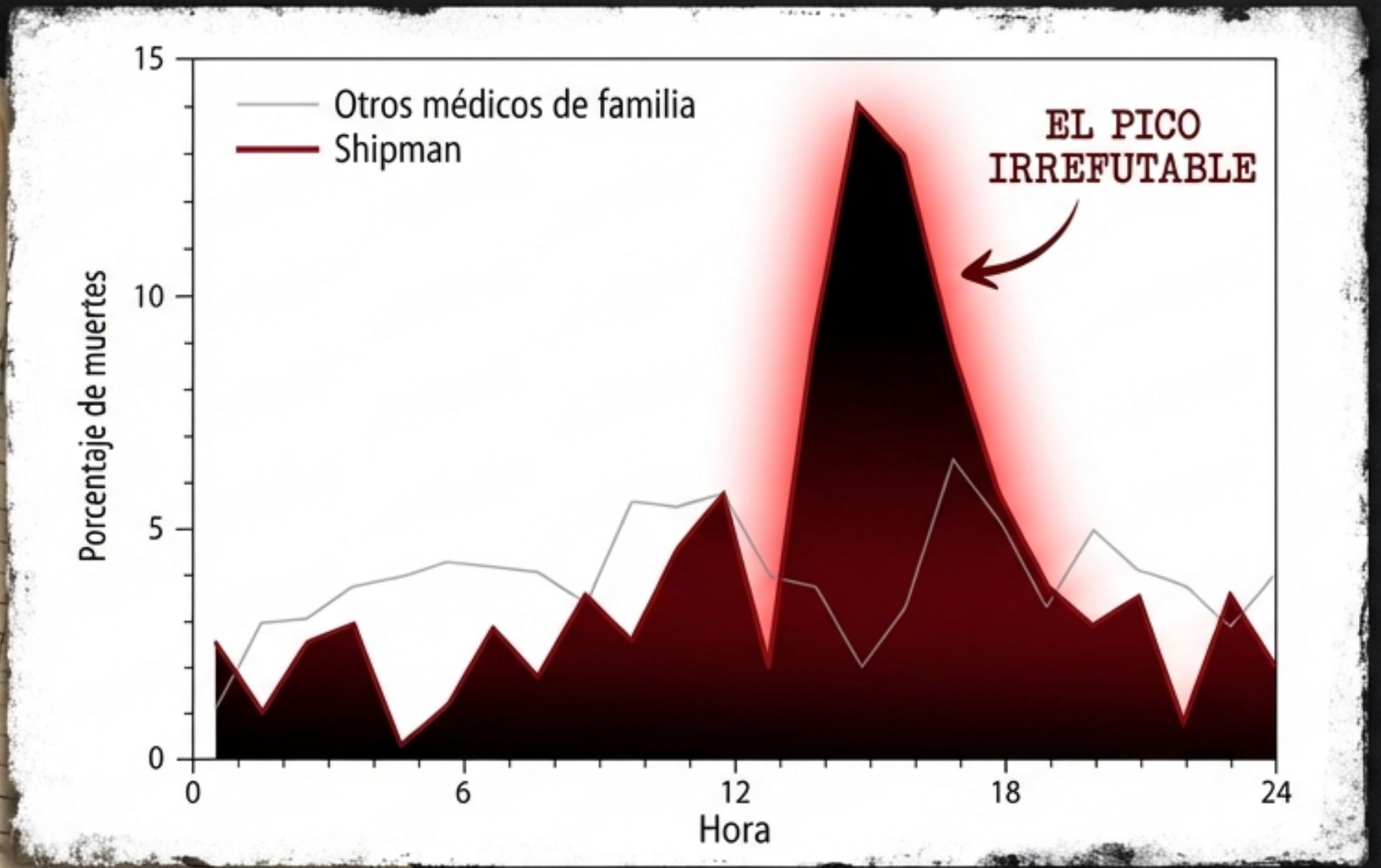
El Hiato (1992): No hubo asesinatos mientras trabajaba en una consulta compartida. Al abrir su propia consulta, los crímenes se reanudaron.

Aceleración final: La actividad criminal se intensificó drásticamente antes de su captura en 1998.

Una sola imagen convierte 215 tragedias individuales en una narrativa criminal clara.

La Prueba «Interocular»

Cuando la verdad te entra directamente por los ojos.



EL PATRÓN:

Shipman realizaba visitas a domicilio después de comer. A solas con los ancianos, les ofrecía una inyección para «sentirse cómodos».

EL PICO:

La mayoría de sus pacientes morían a primera hora de la tarde, una anomalía estadística imposible de ignorar.

“Indescriptiblemente horrible... que día tras día fingiese ser ese doctor maravillosamente bondadoso mientras llevaba en su maletín su arma letal.”

— Magistrada Janet Smith

Estadística Forense: La Búsqueda de Pautas

Lo que acabamos de ver es un **trabajo exploratorio**.
No hubo matemáticas complejas ni teoría abstracta.
Fue una búsqueda de patrones que generó preguntas críticas:

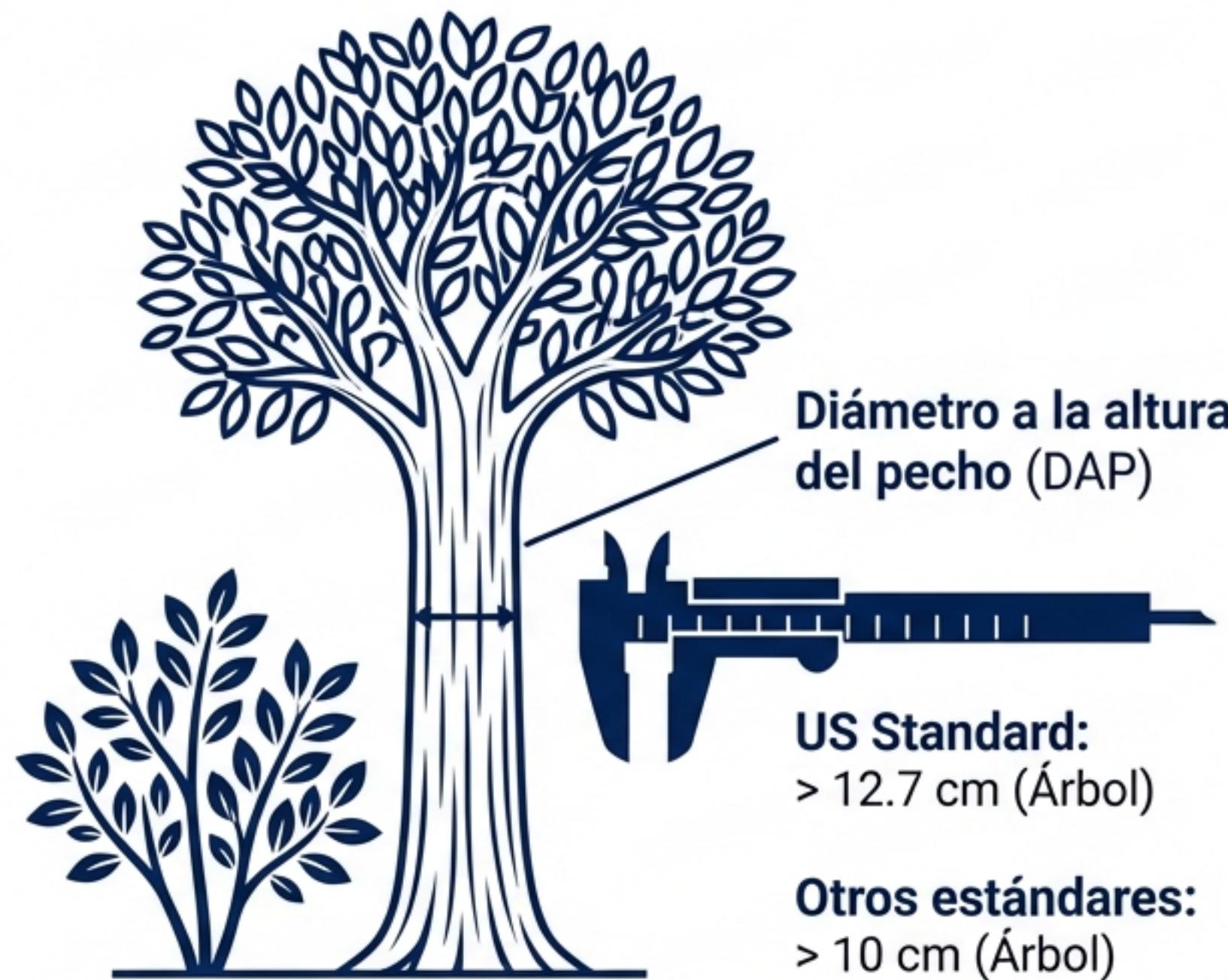
¿Por qué mueren a esta hora?
¿Por qué hubo una pausa en 1992?

La estadística no solo responde preguntas;
nos dice qué preguntas debemos hacer.



El Desafío de Clasificar el Mundo

¿Cuántos árboles hay en el planeta?



EL PROBLEMA DE LA DEFINICIÓN:

Antes de contar, hay que definir. ¿Qué es un «árbol»? Su juicio puede diferir del de un experto. ¿Es un arbusto grande o un árbol pequeño?

EL RESULTADO:

Usando imágenes satelitales y modelización, se estiman 3.04 billones de árboles. Sin embargo, los investigadores creen que la cifra real podría ser el doble, dependiendo simplemente de cómo definamos la palabra «árbol».

Definiendo lo Indefinible

Las estadísticas se construyen sobre juicios humanos,
no verdades absolutas.



DESEMPLEO

En el Reino Unido, la definición oficial cambió al menos **31 veces** entre 1979 y 1996.

Cada cambio alteró la cifra final sin que la realidad cambiara.



PIB (Producto Interior Bruto)

Se revisa continuamente. En 2014, el Reino Unido añadió el **comercio de drogas ilegales** y la **prostitución** al cálculo (usando webs de reseñas para estimar precios).



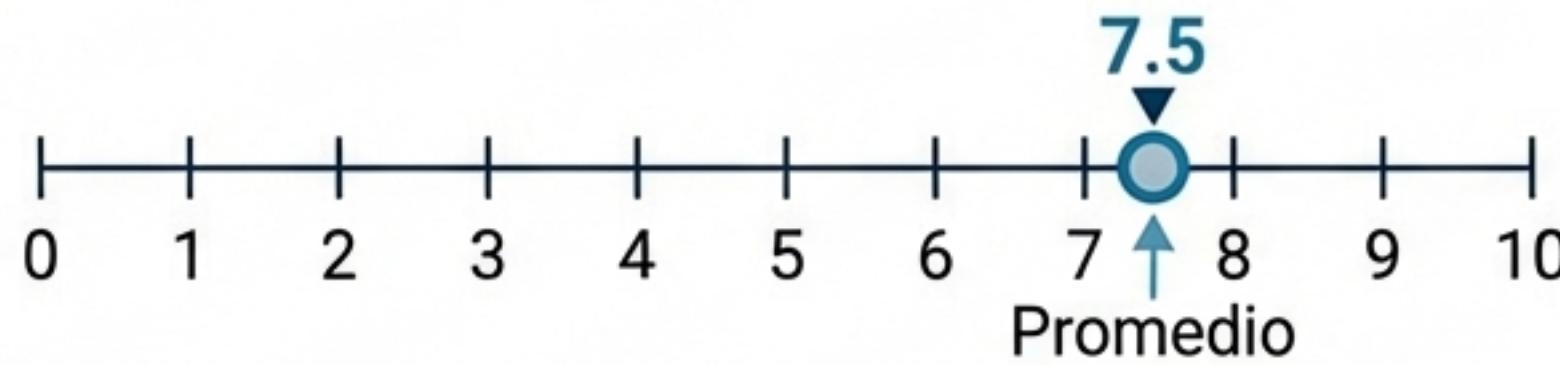
MUERTE

Varía legalmente. Alguien declarado muerto en Alabama podría no estarlo en Florida, donde se requieren dos médicos para certificarlo.

Midiendo Sentimientos en una Hoja de Cálculo

Sociological Study

Pregunta: En líneas generales, ¿cómo de feliz se sintió usted ayer?



EL PATRÓN DE LA EDAD (Reino Unido):

- **Puntuaciones más bajas:** Personas entre 50 y 54 años.
- **Puntuaciones más altas:** Personas entre 70 y 74 años.

LAS DOS LIMITACIONES DE LOS DATOS

1. **Medida Imperfecta:** Un número del 0 al 10 difícilmente captura el bienestar emocional real.
2. **Variabilidad (Ruido):** Todo difiere de una persona a otra. El reto es encontrar la señal entre el ruido aleatorio.

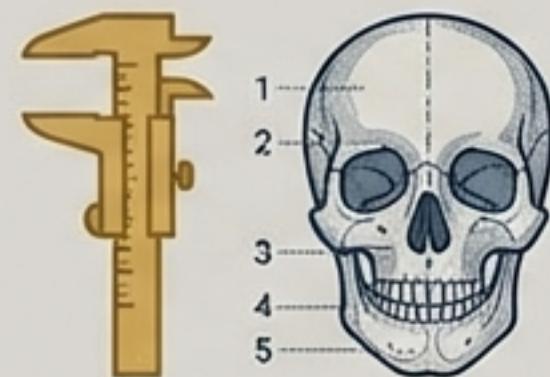
La Evolución de la Disciplina

1650s



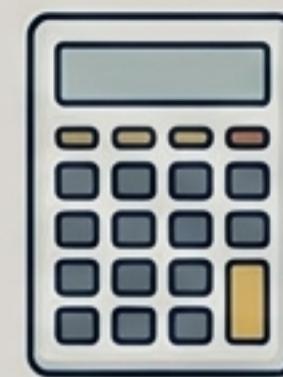
Pascal y Fermat dominan la probabilidad (matemáticas de la variabilidad).

Era Victoriana



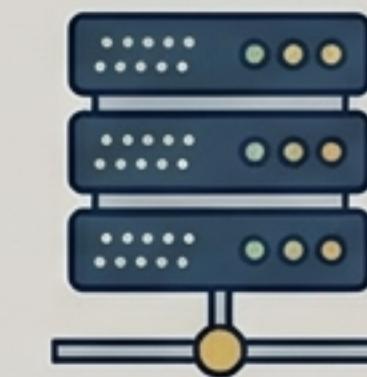
Obsesión por reunir datos del cuerpo humano y genética.

Siglo XX



La estadística se vuelve una «caja de herramientas» matemática mecánica.

Hoy: Big Data



Ciencia de Datos y Crisis de Reproducibilidad.

LA CRISIS MODERNA: El acceso fácil al software permite el «dragado de datos»—buscar patrones aleatorios hasta encontrar algo interesante pero falso (ej. poses de poder).

LA NECESIDAD DE LA «ALFABETIZACIÓN DE DATOS»

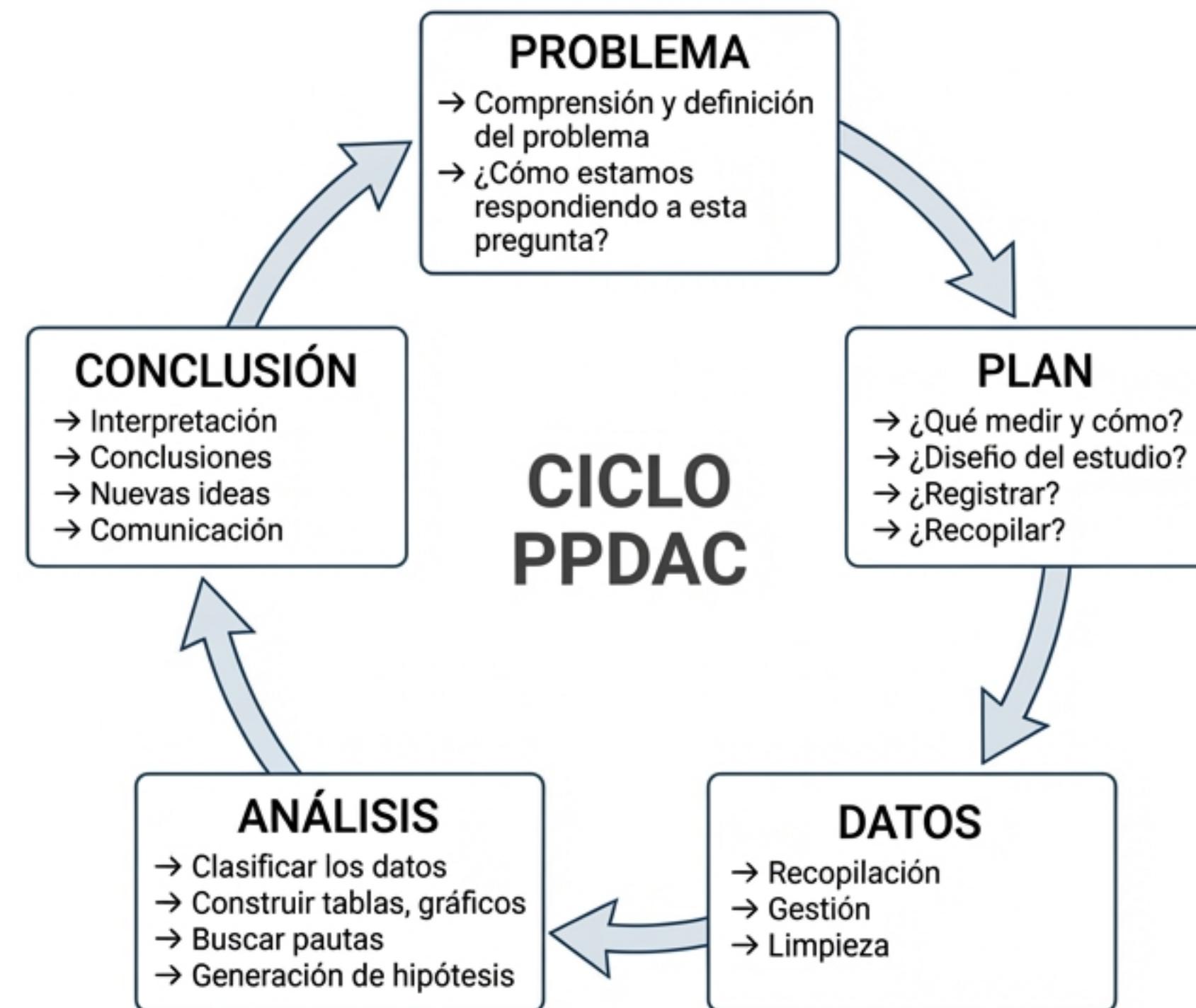
Más datos no significan automáticamente más verdad. A menudo significan más ruido.

- Criticar conclusiones de otros.
- Entender el contexto de la evidencia.
- Reconocer los peligros de la presentación selectiva.

OBJETIVO: Pasar de aplicar fórmulas mecánicas a un ciclo de resolución de problemas.

El Ciclo PPDAC

La estadística no es lineal. Es un ciclo iterativo.



PPDAC Pasos 1 y 2: Definir y Diseñar



1. PROBLEMA

Todo empieza con una pregunta clara.

Ejemplo: «¿A qué tipo de personas asesinó Shipman?» o «¿El tratamiento X funciona?»



2. PLAN

La etapa más tentadora de saltar, pero la más crítica. ¿Qué medir y cómo? Aquí es donde definimos qué es un «árbol» o qué cuenta como «desempleo». Sin un diseño correcto, no hay conclusiones creíbles.

PPDAC Paso 3: La Realidad de los Datos



3. DATOS

En la ciencia de datos moderna, los datos son «desordenados» (found data). A menudo provienen de fuentes cotidianas y requieren habilidades de codificación para detectar errores antes de cualquier análisis.



PPDAC Pasos 4 y 5: Encontrando el Significado

4. ANÁLISIS



Tradicionalmente el foco de la estadística.
Puede ser matemática compleja, PERO a
veces basta una visualización bien
presentada (como el gráfico de Shipman).

5. CONCLUSIÓN



Interpretación y comunicación.
Debe reconocer las limitaciones
de la evidencia.

EL CICLO CONTINÚA: Las conclusiones plantean nuevas
ideas, reiniciando el ciclo hacia un nuevo Problema.

Encontrando la Señal en el Ruido

El mundo está lleno de variabilidad y definiciones subjetivas. Desde los registros de un asesino hasta el conteo de árboles, los datos son nuestra herramienta para navegar esa incertidumbre.

ADOPTE LA MENTALIDAD PPDAC. Sea escéptico. Sea riguroso. Permita que los datos cuenten su historia, pero nunca olvide quién define los términos.

«La estadística es la ciencia de aprender de los datos.»