

Por el autor de *El economista camuflado*

Tim Harford

10 reglas
para comprender
el mundo



Cómo los números pueden explicar
(y mejorar) lo que sucede

conecta

*Para todos los profesores, y, en particular,
para mis profesores; en memoria de Peter Sinclair*

Introducción

Cómo mentir con estadísticas

El problema real [...] no está en demostrar que algo es falso, sino en demostrar que el objeto auténtico es auténtico.

UMBERTO ECO [1]

¿Te suena aquello de que a los bebés los traen las cigüeñas? Es verdad.

Puedo demostrarlo con estadísticas.

Echemos un vistazo a la población estimada de cigüeñas en cada país, y luego al número de niños que nacen cada año. En toda Europa hay una correlación notablemente fundamentada. Más cigüeñas, más bebés; menos cigüeñas, menos bebés.

El patrón es lo bastante sólido para cumplir con los requisitos de publicación de una revista académica. De hecho, se ha publicado un ensayo científico con el título «Las cigüeñas traen a los bebés ($p = 0,008$)». Para no ponernos demasiado técnicos, todos esos ceros significan que no es una coincidencia. [2]

Quizá ya te hayas dado cuenta de dónde está el truco. Países europeos de gran tamaño, como Alemania, Polonia y Turquía, son el hogar de muchos bebés y muchas cigüeñas. Países pequeños, como Albania y Dinamarca, tienen menos bebés y menos cigüeñas. Aunque en los datos hay un patrón claro, ese patrón no significa que las cigüeñas sean la causa de los bebés.

Al parecer, con las estadísticas se puede «demostrar»

cualquier cosa..., incluso que a los bebés los traen las cigüeñas.

Sin duda eso es lo que pensarías después de leer *Cómo mentir con estadísticas*. Publicado en 1954 por un periodista freelance llamado Darrell Huff, este librito ocurrente y cínico recibió de inmediato una crítica elogiosa de *The New York Times*, llegó a convertirse en el que quizá sea el libro sobre estadística más popular jamás publicado y vendió más de un millón de ejemplares.

El libro se merece la popularidad y los elogios. Es una maravilla de la comunicación estadística. También convirtió a Darrell Huff en una leyenda entre los entendidos. Ben Goldacre, epidemiólogo y autor superventas de *Mala ciencia*, afirmó con admiración que «The Huff» había escrito un libro «impactante». El escritor estadounidense Charles Wheeland describe su libro *Naked Statistics* como un «homenaje» al «clásico» de Huff. La respetada revista *Statistical Science* organizó una retrospectiva de Huff cuando se cumplieron cincuenta años de su publicación.

Yo antes también pensaba eso. De adolescente me encantaba leer *Cómo mentir con estadísticas*. Brillante, perspicaz e ilustrado con diagramas atractivos, este libro me permitió mirar bajo la alfombra de la manipulación estadística y me enseñó cómo se falseaban los datos, con lo que ya no volverían a engañarme.

Huff da un montón de ejemplos. Comienza sopesando cuánto dinero ganan los graduados en Yale. Según una encuesta de 1950, la promoción de 1924 tenía unos ingresos medios cercanos a los 500.000 dólares anuales, al valor actual. Parece posible, creíble —al fin y al cabo, se trata de Yale—, pero medio millón de dólares al año es mucho dinero. ¿De verdad es esa la media?

No. Huff explica que esta cifra tan «improbablemente generosa» proviene de los datos que aportan los exalumnos, y cabe sospechar que exageren sus ingresos por una simple cuestión de vanidad. Además, solo se tuvieron en cuen-

ta las encuestas que los exalumnos se molestaron en responder, y la encuesta solo les llegó a los exalumnos que Yale consiguió encontrar. ¿Y quiénes son fáciles de encontrar? Los ricos y famosos. «¿Quiénes son los corderitos perdidos que Yale etiqueta como “dirección desconocida”?», se pregunta Huff. Yale controlaba a los exalumnos millonarios, pero es muy posible que pasara por alto a algunos de sus compañeros con menos éxito. Todo esto significa que la encuesta presentará una conclusión extraordinariamente sobredimensionada.

Huff se mueve como pez en el agua en un amplio abanico de crímenes estadísticos, desde anuncios de dentífricos basados en investigaciones elegidas por conveniencia hasta mapas que cambian de significado dependiendo de cómo los coloreemos. Como escribió Huff: «Los timadores conocen todos estos trucos; las personas honestas deben aprenderlos para defenderse».

Si lees *Cómo mentir con estadísticas*, te volverás más eséptico respecto a cómo pueden engañarte los números. Es un libro inteligente e instructivo.

Pero yo me he pasado más de una década tratando de comunicar ideas estadísticas y hechos contrastados numéricamente, y a medida que pasan los años cada vez me incomoda más *Cómo mentir con estadísticas* y lo que representa. ¿Qué dice de las estadísticas —y de nosotros— que el libro más exitoso sobre esta cuestión sea, de la primera a la última página, una advertencia sobre la desinformación?

Darrell Huff publicó *Cómo mentir con estadísticas* en 1954. Pero ese mismo año ocurrió otra cosa: dos investigadores británicos, Richard Doll y Austin Bradford Hill, elaboraron uno de los primeros estudios convincentes para demostrar que fumar causa cáncer de pulmón. [3]

Doll y Hill no habrían podido llegar a esta conclusión sin estadísticas. Las tasas de cáncer de pulmón se habían mul-

tiplicado por seis en el Reino Unido en solo quince años; en 1950, las del Reino Unido eran las más altas del mundo, y las muertes por cáncer de pulmón superaron por primera vez a las de tuberculosis. Incluso para tomar conciencia de esta situación era necesaria una perspectiva estadística. Un médico, en soledad, solo habría podido hacerse una idea vaga.

Para demostrar que los cigarrillos eran los culpables, se necesitaban, de nuevo, estadísticas. Mucha gente pensaba que los coches con motor eran la causa del aumento de cáncer de pulmón. Tenía sentido. En la primera mitad del siglo xx, los coches con motor se volvieron omnipresentes, con los gases de los tubos de escape y las emanaciones extrañamente atractivas del alquitrán de las carreteras nuevas. El cáncer de pulmón se incrementó en la misma época. Descubrir la verdad —que la causa eran los cigarrillos y no los coches— requería algo más que observar la situación. Era necesario que los investigadores empezaran a contar, y a comparar, con mucho detalle. Se necesitaba, en suma, la estadística.

Muchos valoraban con escepticismo la hipótesis de los cigarrillos, aunque no era del todo nueva. Por ejemplo, en la Alemania nazi se invirtieron considerables recursos para conseguir pruebas de que el tabaco era peligroso; Adolf Hitler lo detestaba. Sin duda lo complació que los médicos descubrieran que causaba cáncer. No obstante, por razones obvias, el hecho de que los nazis lo detestaran no impidió su popularidad.

Así que Doll y Hill decidieron llevar a cabo sus propias investigaciones estadísticas. Richard Doll era un joven apuesto, sereno e indefectiblemente educado. Había vuelto de la Segunda Guerra Mundial con la cabeza rebosante de ideas sobre cómo la estadística podía revolucionar la medicina. Su mentor, Austin Bradford Hill, había sido piloto en la Primera Guerra Mundial antes de estar al borde de la muerte por tuberculosis. (1) Hill era un hombre carismático, con un

humor mordaz, y se lo consideraba el mejor médico estadístico del siglo xx . [4] Su labor como detectives de datos logró salvar vidas.

Su primer estudio sobre la relación entre el hecho de fumar y el cáncer empezó el día de Año Nuevo de 1948. Se centró en veinte hospitales del noroeste de Londres, y Richard Doll estuvo al cargo.

Cada vez que un paciente con cáncer llegaba al hospital, las enfermeras —de forma aleatoria— buscaban a otro paciente del mismo centro con una edad parecida y del mismo sexo. Ambos pacientes eran interrogados en profundidad sobre dónde vivían y trabajaban, su estilo de vida y su dieta, y su historial con el tabaco. Semana tras semana, mes tras mes, los resultados fueron llegando.

En octubre de 1949, menos de dos años después de que empezaran las pruebas, Doll dejó de fumar. Tenía treinta y siete años y había sido fumador durante toda su vida adulta. Él y Hill descubrieron que los fumadores empedernidos no solo doblaban, triplicaban o cuadruplicaban el riesgo de padecer cáncer de pulmón, sino que multiplicaban por dieciséis las posibilidades de tener cáncer de pulmón. [5]

Hill y Doll publicaron los resultados en septiembre de 1950, y enseguida se embarcaron en un estudio mayor, a largo plazo y más ambicioso. Hill escribió a cada uno de los médicos del Reino Unido —un total de 59.600— para que completaran un «cuestionario» sobre su salud y sus hábitos con los cigarrillos. Doll y Hill supusieron que los médicos serían capaces de hacer un seguimiento de lo que fumaban. Estarían en el registro médico, así que serían fáciles de encontrar. Y cuando muere un médico cabe esperar un buen diagnóstico de la causa de su muerte. Todo lo que tenían que hacer Hill y Doll era esperar.

Más de 40.000 médicos accedieron a la petición de Hill, pero no todos lo hicieron de forma complaciente. Debe-

mos tener en cuenta que fumar era de lo más común en aquella época, no fue ninguna sorpresa descubrir que el 85 por ciento de los médicos varones de la primera muestra de Doll y Hill eran fumadores. A nadie le gusta que le digan que se están matando poco a poco, sobre todo si el método de suicidio es altamente adictivo.

Un médico acorraló a Hill en una fiesta en Londres. «Tú eres el tipo que quiere que dejemos de fumar», le dijo de manera acusadora.

«En absoluto —contestó Hill, que seguía siendo fumador de pipa—. Me interesa saber de qué morirás si sigues fumando y me interesa saber de qué morirás si dejas de fumar. Tú eliges, o lo dejas o sigues. A mí me es indiferente. Registraré tu muerte de todas formas.» [6]

¿He mencionado que Hill, al principio, se formó como economista? Fue entonces cuando adquirió este encanto conversacional.

El estudio de los médicos se alargó durante décadas, pero no pasó mucho tiempo antes de que Doll y Hill recabaran datos suficientes para publicar una conclusión clara: fumar provoca cáncer de pulmón, y cuanto más se fuma, más alto es el riesgo. Más aún —y esto era algo nuevo—: fumar también provoca ataques al corazón.

Los médicos no son idiotas. En 1954, cuando se publicó el estudio en su revista profesional, la *British Medical Journal*, sacaron sus propias conclusiones. Hill dejó de fumar aquel año, y muchos de sus compañeros también lo hicieron. Los médicos se convirtieron en el primer grupo social identificable en el Reino Unido que dejó de fumar de manera generalizada.

En 1954, por lo tanto, aparecieron al mismo tiempo dos visiones de la estadística. Para muchos lectores de *Cómo mentir con estadísticas*, de Darrell Huff, las estadísticas eran un juego lleno de timos y de trampas, y podía ser divertido desenmascarar a los estafadores y sus trucos. Pero para Austin Bradford Hill y Richard Doll, las estadísticas no

eran algo de lo que reírse. Era una cuestión en la que había mucho en juego y, si se abordaba con sinceridad y diligencia, podía salvar vidas.

En la primavera de 2020 —cuando estaba dando los últimos retoques a este libro—, lo que estaba en juego con las estadísticas rigurosas, oportunas y transparentes resultó, de repente, muy claro. Un nuevo coronavirus arrollaba al mundo. Los gobernantes debían tomar las decisiones más relevantes que se habían tomado en décadas, y tenían que hacerlo rápido. Muchas de estas decisiones dependían de los datos que recababan los epidemiólogos, los estadísticos médicos y los economistas. Millones de vidas estaban potencialmente en riesgo, así como el sustento de miles de millones de personas.

Escribo estas palabras a principios de abril de 2020: muchos países de todo el mundo llevan un par de semanas de confinamiento, las muertes globales han superado las 60.000, y no está nada claro cómo acabará todo esto. Quizá, cuando este libro llegue a tus manos, estemos enfangados en la peor depresión económica desde la década de 1930 y el número de muertos se habrá disparado. Quizá, gracias al ingenio humano o a la buena fortuna, estos miedos apocalípticos se hayan desvanecido y solo formen parte del recuerdo. Hay muchas situaciones posibles. Y ese es el problema.

El epidemiólogo John Ioannidis escribió a mediados de marzo que el Covid-19 «podría ser el fracaso del siglo». [Z] Los detectives de datos hacen cuanto pueden, pero deben trabajar con datos incompletos, inconsistentes y terriblemente inadecuados para tomar decisiones de vida o muerte con la seguridad con que nos gustaría.

Los detalles de este fracaso se estudiarán, sin duda, en los años por venir. Pero algunas cosas ya están bastante claras. Al principio de la crisis, por ejemplo, parece que los

políticos han entorpecido la transmisión libre de estadísticas claras, un problema del que volveremos a ocuparnos en el capítulo octavo. Taiwán se quejó de que a finales de diciembre de 2019 ya informó de indicios importantes sobre la transmisión entre humanos a la Organización Mundial de la Salud (OMS), pero, a mediados de enero la OMS tuiteaba sin asomo de duda que China no había detectado pruebas de transmisión entre humanos. (Taiwán no es miembro de la OMS porque China reclama la soberanía sobre este territorio y exige que no se lo trate como un Estado independiente. Es posible que este obstáculo geopolítico haya provocado este supuesto retraso.) [8]

¿Influyó en algo? Casi seguro que sí. Con el número de casos doblándose cada dos o tres días, nunca sabremos qué habría cambiado de haber estado advertidos un par de semanas antes. Está claro que muchos líderes se tomaron su tiempo para valorar la gravedad potencial de la amenaza. El presidente Trump, por ejemplo, declaró a finales de febrero: «Va a desaparecer. Un día, como un milagro, desaparecerá». Cuatro semanas después, con 1.300 estadounidenses muertos y con más casos confirmados que ningún otro país, el señor Trump todavía anunciaba con esperanza que en Pascua todos podrían ir a la iglesia. [9]

Mientras escribo estas líneas los debates son encarnizados. ¿Podrán las pruebas rápidas, el aislamiento y el rastreo de contactos contener los brotes indefinidamente, o solo retrasarán los contagios? ¿Deberíamos preocuparnos más por las reuniones de pocas personas en espacios interiores o por las reuniones de mucha gente en espacios exteriores? ¿Cerrar las escuelas impide el contagio del virus, o es más dañino que los niños se queden en casa con unos abuelos vulnerables? ¿Hasta qué punto ayuda llevar mascarilla? Estas y muchas otras preguntas solo podrán responderlas buenos datos sobre quién se ha infectado y cuándo.

Pero un número considerable de infecciones no se registraron en las estadísticas oficiales debido a la falta de prue-

bas. Y las pruebas que sí que se llevaban a cabo proporcionaban un panorama sesgado, pues se centraba en el personal médico, los pacientes críticos y —afrontémoslo— personas ricas y famosas. Cuando escribo estas palabras, los datos todavía no pueden decirnos cuántos casos leves o asintomáticos hay y, por lo tanto, cuán mortífero es el virus. A medida que el número de muertes aumentaba exponencialmente en marzo, doblándose cada dos días, no había tiempo que perder. Los líderes pusieron a las economías en un coma inducido: más de tres millones de estadounidenses se quedaron en el paro en solo una semana de finales de marzo, lo cual multiplicó por cinco el récord anterior. La siguiente semana fue incluso peor: seis millones y medio de personas se quedaron sin trabajo. ¿Las posibles consecuencias para la salud eran lo bastante catastróficas como para justificar esta pérdida de empleos? Eso parecía, pero los epidemiólogos tuvieron que hacer sus mejores conjeturas con una información muy limitada.

Es difícil imaginar un ejemplo más extraordinario de hasta qué punto solemos dar por garantizada una recopilación de cifras sistemática y precisa. Las estadísticas de un amplio abanico de cuestiones importantes que precedieron al coronavirus se recabaron a lo largo de los años gracias al esfuerzo de estadísticos diligentes, y con frecuencia se pueden descargar desde cualquier lugar del mundo sin cargo alguno. Pero este lujo nos ha malcriado, y despreciamos con despreocupación las «mentiras, las malditas mentiras y las estadísticas». El caso del Covid-19 nos recuerda lo desesperada que puede ser una situación cuando carecemos de estadísticas.

Darrell Huff hizo que las estadísticas parecieran el truco de un mago: todo muy divertido, nunca deben tomarse demasiado en serio. Mucho antes del coronavirus, empecé a preocuparme porque esta actitud no nos ayudaría. Hemos

dejado de lado la idea de que las estadísticas pueden ayudarnos a comprender el mundo. No es que creamos que todas las estadísticas son una mentira, sino que no tenemos forma de desgranar las verdades. Así que creemos cualquier cosa que queramos creer (más sobre esta cuestión en el próximo capítulo), y, por lo demás, adoptamos la reacción de Huff: risa ácida, encogimiento de hombros, o las dos cosas a la vez.

Este cinismo estadístico no solo es una vergüenza, es una tragedia. Si cedemos a la idea de que ya no tenemos el poder de saber lo que es verdad, habremos perdido una herramienta vital. Es una herramienta que nos enseñó que los cigarrillos son mortíferos. Es nuestra única oportunidad real de abrirnos paso en la crisis del coronavirus o, desde una perspectiva más amplia, de comprender el complejo mundo en el que vivimos. Pero esta herramienta es inútil si nos limitamos al desprecio automático de cualquier estadística que no nos guste. No debemos ser crédulos, por supuesto, pero el antídoto de la credulidad no es no creer en nada, sino tener confianza para sopesar la información con curiosidad y un escepticismo saludable.

Las buenas estadísticas no son un truco, aunque son una especie de magia. Las buenas estadísticas no son humo ni espejos; de hecho, nos ayudan a ver con más claridad. Las buenas estadísticas son como el telescopio para el astrónomo, el microscopio para el bacteriólogo o los rayos X para el radiólogo. Si estamos dispuestos a permitirselo, las buenas estadísticas nos ayudarán a ver cosas del mundo que nos rodea y de nosotros mismos —tanto grandes como pequeñas— que no seríamos capaces de ver de otra manera.

El principal objetivo de este libro es persuadirte para que adoptes la visión de Doll y Hill, no el cinismo de Huff. Quiero convencerte de que las estadísticas pueden iluminar la realidad con claridad y honestidad. Para ello, necesito mostrarte que puedes beneficiarte del razonamiento estadístico evaluando las afirmaciones de los medios, de las redes so-

ciales y de las conversaciones cotidianas. Quiero ayudarte a valorar las afirmaciones tal como son y, no menos importante, a saber dónde encontrar ayuda en la que confiar.

La buena noticia es que va a ser divertido. Hay una satisfacción real en llegar al fondo de la estadística: ganas confianza, alimentas tu curiosidad y acabas sintiendo que dominas algo. Entiendes en lugar de mirar con desconfianza desde el banquillo. La perspectiva de Darrell Huff es comida basura: superficialmente atractiva, pero tediosa al cabo de un rato. Y es mala para ti. Pero lo contrario de la comida basura estadística no son copos de avena y nabos, sino un menú variado y satisfactorio.

En este libro voy a relatar lo que he aprendido desde 2007, cuando la BBC me propuso que presentara un programa de radio llamado *More or Less*, un espacio sobre los números en las noticias y la vida. Los creadores del programa, el periodista Michael Blastland y el economista sir Andrew Dilnot, querían seguir otros caminos. Yo estaba menos cualificado para el cargo de lo que quizá creía la BBC: me formé en teoría económica, no en estadística. Sí, esta formación me daba cierta seguridad respecto a los números, pero era en esencia defensiva: había aprendido a detectar defectos y trucos, pero no mucho más.

Fue entonces cuando comenzó mi divergencia de perspectiva con Darrell Huff. Semana tras semana, mis colegas y yo evaluábamos afirmaciones estadísticas que articulaban las bocas de los políticos o que aparecían publicadas con letras mayúsculas en los periódicos. Con frecuencia, estas afirmaciones eran exageradas, pero el mero cotejo de los hechos no parecía una respuesta satisfactoria. Descubríamos que tras cada afirmación —verdadera, falsa o en el límite— había un mundo fascinante por descubrir y explicar. Tanto si evaluábamos la prevalencia de las apoplejías, las pruebas de que la deuda merma el crecimiento económico,

o incluso cuántas veces se utiliza la palabra «ella» en *El Hobbit*, los números podían iluminar el mundo o encriptarlo.

Como muestra la epidemia del coronavirus, dependemos de cifras en las que podamos confiar para tomar decisiones como individuos, organizaciones y sociedad. Y, como ocurre con el coronavirus, solo recabamos estadísticas cuando nos enfrentamos a una crisis. Piensa en la tasa de desempleo, una variable que nos sirve para saber cuántas personas quieren un trabajo pero no lo encuentran. En la actualidad es un dato básico para cualquier gobierno que quiera comprender el estado de la economía, pero en 1920 nadie tenía ni la más remota idea de cuántas personas buscaban trabajo. [10] Solo cuando varias recesiones severas lo convirtieron en una cuestión políticamente pertinente los gobiernos empezaron a recopilar información para obtener ese dato.

Nuestro mundo, enorme y convulso, está lleno de preguntas que solo una atención adecuada a los números puede responder. ¿Facebook tiende a hacernos felices o infelices? ¿Podemos predecir por qué personas diferentes reaccionan de formas diferentes? ¿Cuántas especies hay en peligro de extinción? ¿Es una gran proporción del total? ¿Y cuál es la causa? ¿El cambio climático, la propagación de la agricultura, o algo distinto por completo? ¿La innovación humana se está acelerando o ralentizando? ¿Hasta qué punto es grave el impacto de la crisis de opioides en la salud de los estadounidenses? ¿Beber en la adolescencia es ahora menos común? Y, si es así, ¿por qué?

Cada vez me incomodaba más cuando los oyentes de *More or Less* nos felicitaban por «desbaratar estadísticas falsas». Es verdad, lo hacíamos, y era divertido. Pero, poco a poco, a medida que fui aprendiendo, me di cuenta de que la auténtica felicidad no estaba en cargarse falsedades sino en intentar comprender qué era verdad.

Trabajando en *More or Less* aprendí que los principios del sentido común pueden ayudarnos mucho si somos detecti-

ves de datos. Esos son los principios que voy a resumir en este libro. La mayoría del equipo de investigadores y productores, como yo, carecían de una formación seria sobre cómo analizar las cifras. Pero, incluso en ámbitos muy técnicos, algunas preguntas simples —y quizá un par de búsquedas en internet— con frecuencia generaban respuestas muy satisfactorias. Sí, en ocasiones un grado avanzado en estadística habría sido útil, pero nunca lo necesitamos para hacer las preguntas adecuadas. Y tú tampoco lo necesitarás.

Justo antes de las Navidades de 1953, unos veteranos ejecutivos de la industria tabaquera se reunieron en el hotel Plaza de Nueva York. El exhaustivo estudio de Doll y Hill todavía iba a tardar un año en publicarse, pero las grandes compañías tabaqueras ya eran conscientes de que la ciencia empezaba a tenerlas en el punto de mira. Se reunieron para decidir cómo reaccionar a esta crisis inminente.

La reacción fue —mal que nos pese— bastante brillante, y preparó el terreno de la propaganda a partir de entonces.

Revolvieron las aguas. Cuestionaron las investigaciones existentes; pidieron más investigaciones; financiaron investigaciones en otros ámbitos que interesaran a los medios, como el síndrome del edificio enfermo y la enfermedad de las vacas locas. Fabricaron dudas. [11] Un memorándum secreto de la industria recordó más tarde a los trabajadores que «la duda es nuestro producto». [12]

Como es lógico, cuando pensamos en la persuasión, pensamos en personas a las que se embauca para que crean algo que no deberían (hablaremos de esto en el próximo capítulo). Pero a veces el problema no es que creamos en algo con entusiasmo, sino que encontremos razones para no creer en nada. A los fumadores les gustaba fumar, dependían físicamente de la nicotina, y querían seguir fumando si podían. Una situación en la que los fumadores se encogie-