

CHRISTIAN RUDDER

dataclismo

Amor, sexo, raza e identidad;
lo que nuestra vida *online*
cuenta de nosotros



Cuando creemos que nadie
nos está mirando



Lectulandia

¿Has pensado alguna vez en las consecuencias de colgar tu intimidad en las redes sociales?

Dataclismo es una investigación audaz e irreverente sobre el comportamiento humano, un nuevo enfoque a una revolución en marcha.

Durante décadas las únicas herramientas para estudiar el comportamiento humano partían de encuestas o pequeños experimentos. Hoy es posible un nuevo enfoque. Debido a Internet cada vez mostramos más aspectos de nuestra vida y los investigadores pueden observar nuestro comportamiento directamente, extrapolando a grandes cifras y sin filtros. Los analistas son a día de hoy los nuevos sociólogos.

En Dataclismo Christian Rudder nos explica de forma original cómo a través de los likes de Facebook se puede predecir con precisión de cirujano la orientación sexual de una persona e incluso su inteligencia, cómo las mujeres atractivas reciben más solicitudes de entrevistas, y por qué hay que tener haters para estar de moda. Muestra cómo se expresa la gente en público y en privado y otras muchas cosas más.

Lectulandia

Christian Rudder

Dataclismo

ePub r1.0

XcUiDi 29.04.2019

Título original: *Dataclysm*
Christian Rudder, 2014
Traducción: Darío Giménez Imirizaldu

Editor digital: XcUiDi
ePub base r2.1



más libros en lectulandia.com

Dataclismo

Quiénes
somos*

Christian
Rudder

* Cuando creemos que nadie nos mira

Introducción

A estas alturas ya habrás oído hablar mucho sobre el fenómeno *Big Data*: su ingente potencial, sus funestas consecuencias y el nuevo paradigma, demolidor de paradigmas, que presagia para la humanidad y sus benditas páginas web. Te da vueltas la cabeza solo de pensarlo, como si te hubiesen atizado con un objeto contundente. Pero no vengo aquí a darle más bombo y platillo a este fenómeno de los datos. Vengo con la cosa en sí misma: los datos, despojados de todo el fenómeno. Vengo con un montón enorme de información real que se está recopilando y de la cual la suerte, el trabajo, la persuasión y algo más de suerte me han colocado en la posición privilegiada de poseedor y analista.

Fui uno de los fundadores de OkCupid, una web de contactos que, tras una trayectoria nada efervescente de diez largos años, ha llegado a ser una de las mayores del mundo. La fundamos entre cuatro amigos. Todos éramos de mentalidad muy matemática y la web triunfó en parte porque aplicamos esa mentalidad a las citas de pareja: llevamos algo de análisis y de rigor a lo que desde siempre habían sido los dominios de «expertos» en amor y embaucadores sonrientes como el doctor Phil. La manera de funcionar de la web no es que sea muy sofisticada —resulta que las únicas matemáticas que hacen falta para dar forma al proceso que facilita que dos personas queden para una cita consisten en un poco de aritmética básica— pero, por algún motivo, nuestro sistema ha triunfado y 10 millones de personas van a usar este año la web para buscar pareja.

Como sé por propia experiencia, a los fundadores de sitios web nos encanta alardear de grandes cifras, pero con toda certeza la mayoría de la gente que sabe usar la cabeza ha aprendido a ignorarlas: uno escucha hablar de millones por aquí y miles de millones por allá y sabe que básicamente se trata de un «Mira cómo molo» pero expresado con una ristra de ceros. A diferencia de Google, Facebook, Twitter y otras fuentes cuyos datos aparecen de manera habitual a lo largo de este libro, OkCupid es un nombre mucho menos familiar: si tus amigos y tú lleváis años felizmente casados, seguramente nunca habréis oído hablar de nosotros. Así que he pensado mucho acerca de cómo describir el alcance de nuestra web a alguien que

nunca la ha usado y a quien, por consiguiente, poco le interesan las estadísticas de interacción de usuarios de una *startup* fundada por determinado individuo. He optado por hacerlo en términos personales. Esta noche, unas 30 000 parejas acudirán a su primera cita gracias a OkCupid. Aproximadamente 3000 de ellas acabarán manteniendo una relación duradera. Unas 200 acabarán casándose y muchas de estas, naturalmente, tendrán hijos. Ahora mismo, ya hay por ahí niños vivitos y coleando, humanos gruñones en miniatura que se niegan a ponerse los zapatos y que nunca habrían existido de no ser por los tejemanejes de nuestro html.

No voy a venirte con la pretenciosa idea de que hayamos perfeccionado nada, y cabe señalar aquí que aunque estoy orgulloso de la web que fundamos mis amigos y yo, la verdad es que no me importa en absoluto que ya seas miembro o que te vayas a crear una cuenta en ella ni nada de eso. Nunca en toda mi vida he ido a una cita *online* ni tampoco lo han hecho los otros dos fundadores, así que si eso tampoco va contigo, créeme, lo entiendo. El evangelismo tecnológico es una de mis cosas menos favoritas y no estoy aquí para intercambiar mis relucientes abalorios digitales por la preciosa isla de nadie. Sigo estando suscrito a revistas en papel. Todos los fines de semana me llega el *Times*. Me da vergüenza publicar en Twitter. No pretendo convencerte de que uses, respetes o «creas en» Internet o las redes sociales más de lo que ya lo haces, si es que lo haces. Te ruego que sigas pensando lo que has pensado siempre sobre el universo *online*. Pero si hay una cosa que con toda sinceridad espero que este libro te haga replantearte es lo que piensas acerca de ti mismo. Porque de eso va precisamente este libro. OkCupid no es más que el medio a través del cual yo he llegado a esta historia.

He dirigido el equipo de análisis de OkCupid desde 2009, y mi tarea consiste en dar sentido a los datos que crean nuestros usuarios. Mientras que mis tres socios cofundadores han hecho casi todo el trabajo pesado de crear el sitio web, yo me he pasado años jugando con los números. Parte de mi trabajo nos ayuda a dirigir la empresa: por ejemplo, comprender por qué y cómo hombres y mujeres ven de manera distinta el sexo y la belleza es algo que resulta esencial en una web de contactos. Pero muchos de mis hallazgos no son útiles de una manera directa, solo interesantes. No es que pueda hacerse gran cosa con el hecho de que, estadísticamente, el grupo de música menos negro del mundo sea Belle & Sebastian, o que usar el *flash* en una foto haga que el modelo parezca siete años más viejo, excepto para decir «¡Qué cosas!» y, como mucho, contarlo en una cena con amigos. Eso es prácticamente lo único que hemos hecho con todos esos datos durante un tiempo; la

información que hemos cosechado no ha ido más allá de alguna que otra miserable nota de prensa. Pero con el tiempo llegamos a analizar tal cantidad de información que empezaron a hacerse visibles grandes tendencias, patrones de mayor alcance dentro de los más pequeños y, lo que es aún mejor, me di cuenta de que podía emplear esos datos para examinar de manera directa temas tabú como la raza. Es decir, que en lugar de preguntar cosas a la gente por medio de encuestas o llevar a cabo experimentos a pequeña escala, que es como habitualmente se ponía en práctica la sociología en el pasado, ahora puedo ir a ver directamente lo que ocurre cuando, pongamos por caso, 100 000 hombres blancos y 100 000 mujeres negras interactúan en privado. Los datos estaban ahí, esperando en nuestros servidores. Se trataba de una oportunidad sociológica irresistible.

Seguí ahondando y al ir incrementándose los descubrimientos, como hace todo aquel que tiene más ideas que público, creé un blog para írselos contando al mundo. Tras una mejora considerable, aquel blog se acabó convirtiendo en este libro. Para hacer *Dataclismo* he ido mucho más allá de OkCupid. De hecho, probablemente he reunido el conjunto de datos de interacción persona a persona más denso y variado que el que haya podido recopilar ningún otro individuo por su cuenta, que abarca la mayoría de las principales fuentes de información *online* de nuestro tiempo, si no todas. En estas páginas me serviré de esos datos no solo para hablar de los hábitos de los usuarios de una web, sino también de una serie de patrones universales.

El debate público sobre el tema de los datos se ha centrado principalmente en dos aspectos: por un lado, el espionaje gubernamental y, por otro, las posibilidades comerciales. Respecto al primero, no creo que sepa más de lo que sabéis vosotros, solo lo que he leído. Por lo que sé, el aparato de seguridad nacional nunca se ha dirigido a ninguna web de contactos para pedir acceso y, a menos que tengan pensado criminalizar la exhibición de abdominales marcadísimos sin rostro o las incesantes declaraciones de mujeres de Brooklyn sobre cuánto les gusta el *whisky*, cuando todo el mundo sabe que no les gusta, no se me ocurre qué interés podrían tener en ellas. Sobre el segundo aspecto, los datos como ingresos económicos, tengo algo más de conocimiento. Cuando empezaba a escribir este libro, a la prensa especializada en tecnología se le caía la baba con la OPV de Facebook: habían recopilado los datos personales de todo el mundo y estaban empezando a convertir esos datos en un montón de dinero, y estaban a punto de convertir todo aquel dinero en más dinero todavía en el mercado de valores. Un titular del *Times*, tres días antes de la oferta de acciones, lo dice

todo: «Facebook debe convertir los datos en oro». Casi podía esperarse que apareciera Rumpelstiltskin en la página de opinión para decir: «Sí, América, esta es una compra segura».

Como fundador de una web financiada con publicidad, puedo confirmar que los datos son útiles para vender. Cada una de las páginas de un sitio web es capaz de registrar la interacción íntegra de determinado usuario —todo aquello donde hace clic, cualquier texto que teclea e incluso el tiempo que permanece en la página— y a partir de eso no cuesta nada formarse una imagen clara de cuáles son sus gustos y de cómo ponérselos en bandeja. Pero por muy increíble que pueda ser esa capacidad, no es mi intención en absoluto centrarme en cosas como que la misión oculta de nuestro país es vender loción corporal a las personas que comparten con sus amigos publicaciones sobre loción corporal. Merced a mi acceso a esos mismos datos, me propongo darle un uso distinto a esas experiencias de los usuarios —los clics de ratón, las teclas que han pulsado y los milisegundos de permanencia—. Si las dos historias recurrentes sobre el *Big Data* han sido la vigilancia y el dinero, yo me he pasado los últimos tres años trabajando en una tercera: la historia humana.

Puede que Facebook sepa que eres uno de los muchos fans de los chocolates M&M y, por lo tanto, te muestre ofertas relacionadas con ese producto. También sabe cuándo has roto con tu novio, te has mudado a Texas o has empezado a aparecer en multitud de fotos con tu exnovio y, por tanto, que vuelves a salir con él. Google sabe que estás buscando un coche nuevo y puede mostrarte la marca y el modelo preseleccionados para ti basándose solo en tus características psicodemográficas. ¿Eres un individuo del tipo B, amante de las emociones fuertes y socialmente consciente, mujer y de 25 a 34 años? Aquí tienes tu coche Subaru. Al mismo tiempo, Google también sabe si eres gay, si estás indignado, si eres racista o si te preocupa el cáncer que sufre tu madre. Twitter, Reddit, Tumblr, Instagram: todas estas empresas son en primer lugar un negocio, pero en un ajustado segundo lugar son también muestras demográficas de un alcance, una profundidad y una relevancia sin precedentes. Los datos digitales, de un modo casi accidental, tienen ahora la facultad de mostrarnos cómo luchamos, cómo amamos, cómo envejecemos y cómo cambiamos. No tenemos más que mirarlos: con solo alejarnos un poco, los datos nos descubren cómo se comporta la gente cuando cree que nadie la está mirando. En estas páginas os enseñaré lo que he visto. Y, por cierto, a la mierda la loción corporal.



Si lees mucha literatura de divulgación, encontrarás en *Dataclismo* un par de cosas poco habituales. La primera es el color gris. La segunda es que en el libro aparecen totales y grandes cifras y eso genera una curiosa ausencia en una historia que supuestamente versa sobre la gente: aquí hacen acto de presencia muy pocas personas concretas. Abundan los gráficos, las tablas y los diagramas, pero casi ningún nombre. Ha empezado a ser todo un cliché que en la divulgación científica se parta de algo pequeño y peculiar para explicar grandes acontecimientos: contar la historia del mundo partiendo de un nabo, rastrear el devenir de una guerra hasta llegar a un pez o iluminar un prisma con una linterna de bolsillo y proyectar todo un precioso arcoíris en la pared del dormitorio. Yo voy a ir en dirección contraria. Voy a partir de algo grande —un ingente paquete de terabytes de datos sobre lo que la gente hace, piensa y dice—, y pretendo filtrar a partir de eso muchas cosas pequeñas: lo que tu red de amistades opina sobre la estabilidad de tu matrimonio, por qué los asiáticos (y también los blancos, negros y latinos) son más reacios a describirse a sí mismos, dónde y por qué permanecen sin salir del armario los gais, cómo ha cambiado la manera de escribir en los últimos diez años y cómo no ha cambiado la indignación. La idea es apartar de toda narrativa la comprensión que tenemos de nosotros mismos y dirigirla hacia los números o, mejor dicho, pensar de tal modo que esos propios números sean la narrativa.

Esta manera de enfocarlo ha ido evolucionando después de mucho hurgar en las cloacas de la estadística. *Dataclismo* es una extensión de lo que mis colegas de trabajo y yo hemos estado haciendo durante años. Una web de contactos sirve para unir a personas, y para hacerlo de un modo creíble debe llegar hasta sus deseos, sus hábitos y sus aversiones. A partir de ahí, te dedicas a recopilar un montón de datos y detalles y te esfuerzas al máximo para traducirlos en forma de teorías generales de la conducta humana. Lo que acaba desarrollando cualquiera que se dedique a esto de trabajar entre tantísima información, a diferencia, digamos, de trabajar para la sección de enlaces matrimoniales del periódico dominical, es una afinidad especial con el confuso conjunto de la humanidad en lugar de con una pareja de individuos concretos. Acabas por comprender a la gente de forma parecida a como lo haría un químico con las partículas que se arremolinan en su solución, y a través de esa comprensión llegas a amarla.

Dicho esto, todas las webs y todos los científicos que trabajan con datos acaban por deshumanizarlos. Los algoritmos no funcionan bien con nada que no sean números, así que si quieres que un ordenador comprenda una idea tienes que convertir todo lo que puedas de ella en dígitos. Por lo tanto, lo

difícil a la hora de enfrentarse a webs y apps es desmenuzar y embutir en pequeños paquetitos —uno, dos, tres...— una experiencia humana que está en continuo desarrollo sin que nadie se dé cuenta: dividir algún proceso enorme e indescriptible —en el caso de Facebook, la amistad; en el de Reddit, la comunidad; en las webs de contactos, el amor— en pedacitos que sea capaz de manejar un servidor informático. Al mismo tiempo, tienes que conservar ese no sé qué que tiene el asunto, de tal modo que el usuario tenga la certeza de que lo que le estás ofreciendo representa la vida real. Internet es una ilusión delicada: imagina que es una zanahoria cortada tan limpiamente en rodajas que mantiene su forma en la tabla de cortar y aparenta estar intacta. Y esta misma tensión —entre la condición humana en eterno progreso y la fragmentación de la base de datos— que puede hacer que sea complicado mantener un sitio web es también lo que hace que funcione la historia que pretendo contar. Las aproximaciones a cosas como la lujuria y la amistad tal como las aborda la tecnología nos brindan una oportunidad verdaderamente novedosa: poner en números puros y duros unos cuantos misterios atemporales, obtener algo de comprensión sobre experiencias que hasta ahora nos habíamos contentado con calificar de «incuantificables». A medida que esas aproximaciones han ido mejorando cada vez más, y a medida que la gente ha ido dejando que penetren cada vez más en su vida, esa comprensión se ha incrementado con sorprendente rapidez. Os daré un ejemplo rápido, pero antes quiero decir que el eslogan de OkCupid debería haber sido «Hacer descriptible lo indescriptible». Lástima.

Todo en Internet está plagado de sistemas de valoración o puntuación. Ya sea mediante los votos arriba/abajo de Reddit, las reseñas de los usuarios de Amazon o incluso el botón «Me gusta» de Facebook, las webs te piden que votes porque ese voto tuyo convierte algo que es fluido e idiosincrásico —tu opinión— en algo que pueden comprender y usar. Las webs de contactos piden a sus usuarios que se valoren entre ellos porque eso les permite transformar primeras impresiones como:

Tiene ojos bonitos

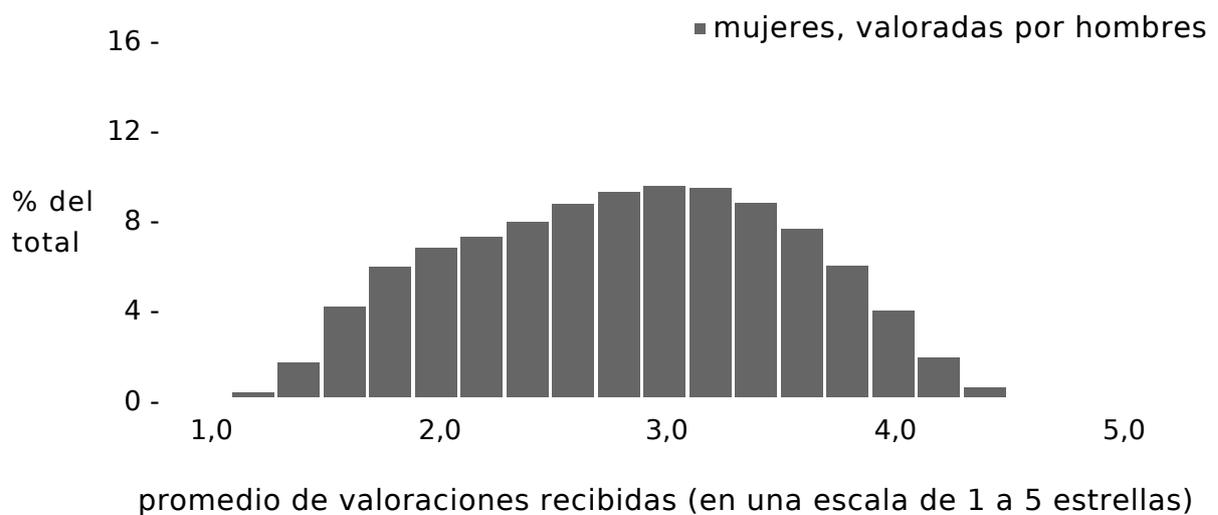
Hummm, es mono, pero no me gustan los pelirrojos

Aj, qué asco

... en simples números como, digamos, 5, 3 y 1 en una escala de una a cinco estrellas. En los sitios web se han recopilado miles de millones de estos microjuicios, de estos destellos de opinión de una persona acerca de otra. Todo ese conjunto de minúsculas reflexiones compone una inmensa fuente de

conocimientos acerca de cómo llega la gente a determinadas opiniones sobre los demás.

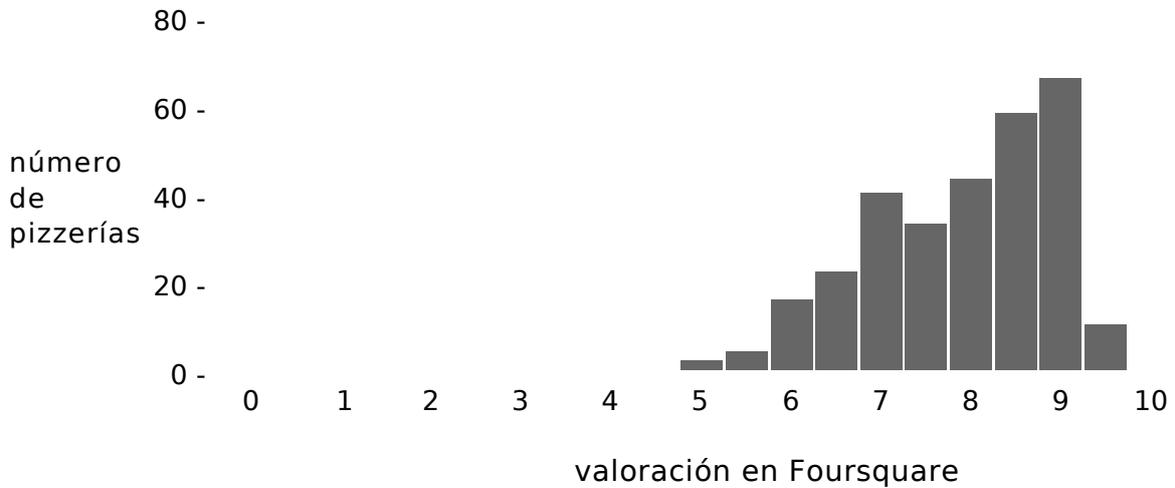
Lo más básico que uno puede hacer con estas valoraciones de persona a persona es contarlas. Contar el total de cuánta gente tiene una media de una estrella, de dos estrellas, etcétera, y después comparar los totales. Abajo he hecho exactamente eso con los promedios de votos otorgados a mujeres heterosexuales por parte de hombres heterosexuales. La curva tiene esta forma:



Unos 51 millones de preferencias personales quedan reducidas a esta sencilla presentación de barras verticales. Se trata, en esencia, de la opinión de los hombres sobre la belleza femenina recopilada por OkCupid. Reúne todas las pequeñas historias (lo que un hombre opina de una mujer, multiplicado por millones de veces) y todas las anécdotas (si nos extendiésemos en cualquiera de ellas, este sería otro tipo de libro) en forma de un todo inteligible. Mirar de este modo a la gente es como observar el planeta Tierra desde el espacio: te pierdes los detalles, pero logras ver de un modo completamente nuevo algo que te es familiar.

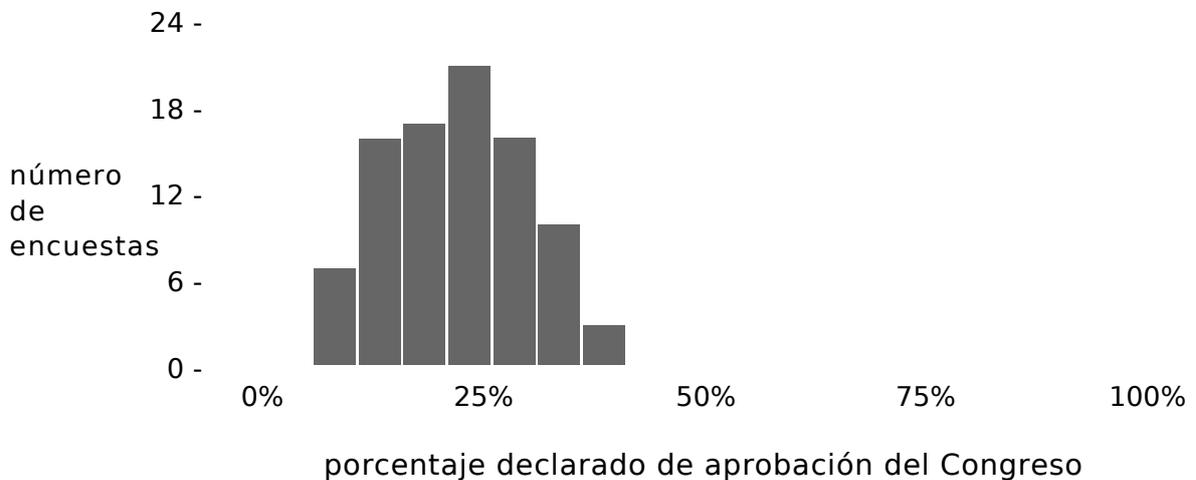
Así pues, ¿qué nos dice esta curva? Es fácil que nos parezca corriente esta forma básica —una curva normal— porque los libros de texto probablemente nos lleven a esperar un gráfico de este tipo, pero los porcentajes podrían haber sido fácilmente más acusados en uno de los extremos o en el otro. Eso es algo que suele ocurrir cuando tratamos con preferencias personales. Tomemos las valoraciones de pizzerías en Foursquare, por ejemplo, que tienden a ser muy positivas:

valoraciones de los usuarios de Foursquare sobre pizzerías de Nueva York, en una escala de 0 a 10



O podemos mirar los porcentajes de aprobación del Congreso que, puesto que los políticos son moralmente lo opuesto a la *pizza*, tienden a ir al otro extremo:

popularidad del Congreso en las principales encuestas de los medios desde noviembre de 2008

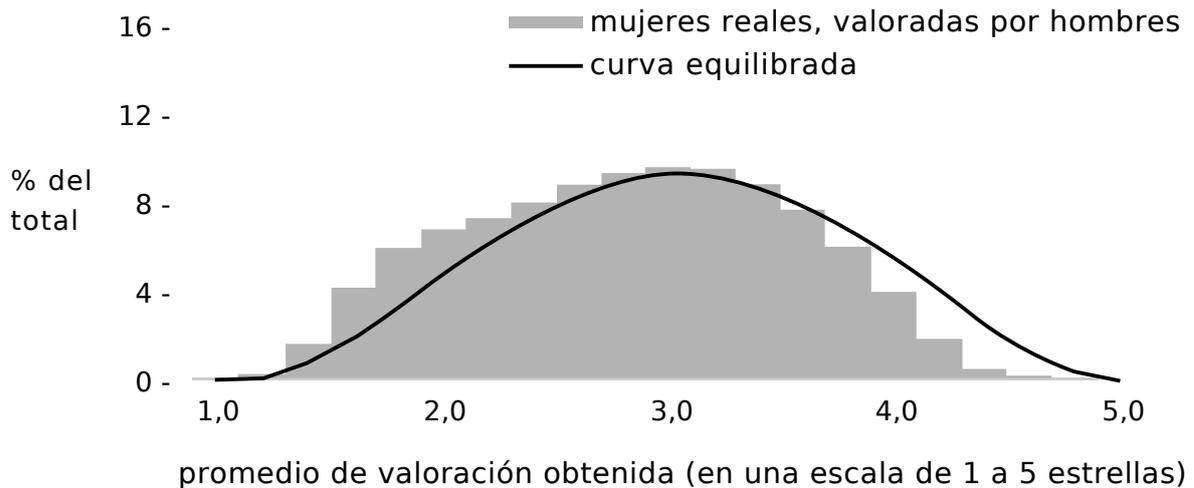


Además, nuestra curva de valoraciones de hombre a mujer es unimodal, lo que quiere decir que las notas asignadas a las mujeres tienden a agruparse alrededor de un solo valor. También esto puede no parecer nada fuera de lo normal, pero muchas situaciones presentan varios modos o valores «típicos». Si separas a los jugadores de la NBA por la frecuencia con la que han jugado en el quinteto inicial en la temporada 2012-2013, verás que te aparecen un montón de atletas agrupados en cada extremo y prácticamente ninguno en el centro:

*jugadores de la NBA por porcentaje de partidos jugados desde el inicio,
temporada 2012-2013*

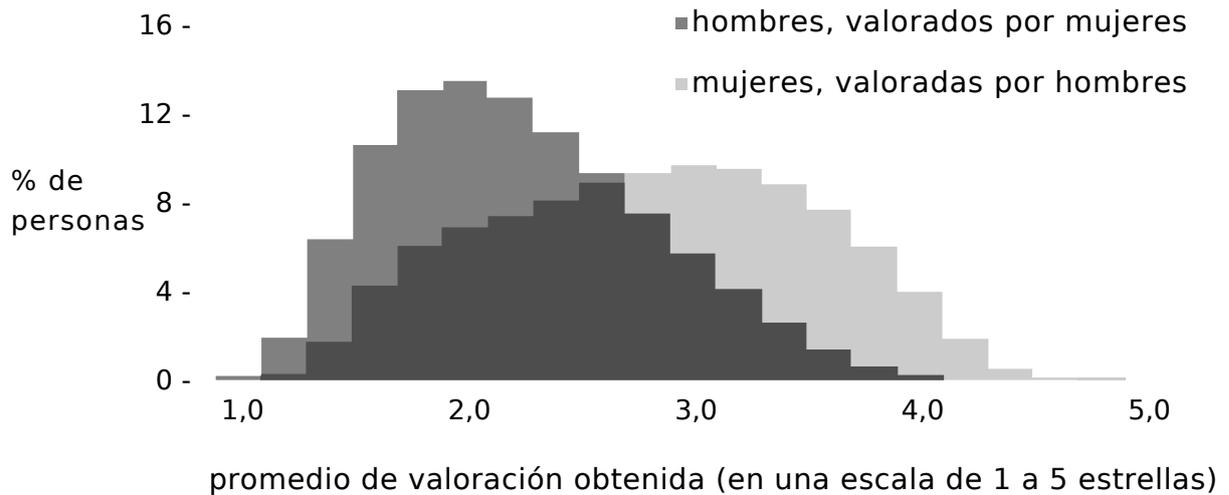
Ahí los datos nos dicen que los entrenadores creen que determinado jugador vale o no vale para jugar de inicio, y ese hombre, por tanto, juega o no en el quinteto inicial del equipo. Ahí tenemos claramente un sistema binario. De forma parecida, en nuestros datos de valoración, los hombres en tanto que grupo podrían haber visto a las mujeres como «preciosas» o «feas» y haberlas valorado en consonancia: igual que en el *ranking* de jugadores de inicio de los equipos de la NBA, la belleza podría considerarse un tema que «se tiene o no se tiene». Pero la curva que hemos mostrado al principio no nos dice eso. Pretender extraer luz de los datos muchas veces es cuestión de comparar nuestros resultados con este tipo de datos contrafácticos. Algunas veces, cuando un resultado sencillo y directo se topa con un sinfín de alternativas, se hace más notable precisamente por ser sencillo y directo. De hecho, nuestro gráfico se acerca bastante a lo que se denomina «distribución beta simétrica» —una curva que suele emplearse como modelo de decisiones equilibradas básicas—, que muestro a continuación:

percepción del atractivo femenino



Nuestros datos del mundo real solo difieren ligeramente (un 6 por ciento) de esa formulación ideal, lo que indica que este gráfico del deseo masculino es más o menos lo que podríamos haber supuesto si careciésemos de los datos: de hecho, es uno de esos ejemplos típicos de libro de texto que he mencionado antes. Así que la curva es predecible, centrada, tal vez incluso aburrida. ¿Y qué? Bueno, este es un contexto en el que lo aburrido resulta especial: implica que los hombres que asignaron sus valoraciones son asimismo predecibles, centrados y, sobre todo, imparciales. Y si nos paramos a pensar en las supermodelos, el porno, las chicas de portada, las superheroínas del tipo de Lara Croft, los anuncios de cerveza Bud Light y, ya si nos ponemos muy retorcidos, en los retoques hechos con Photoshop que todos esos hombres ven cada día por doquier, el hecho de que la opinión masculina sobre el atractivo femenino se mantenga donde se supone que debería estar constituye para mí, a todas luces, un pequeño milagro. Es prácticamente de sentido común que los hombres deberían de tener unas expectativas poco realistas acerca de la apariencia de las mujeres, pero aquí vemos claramente que no es así. En todo caso, son incluso más generosos que las mujeres, cuyos votos responden a esta curva:

percepción del atractivo masculino y del femenino



El centro del gráfico gris oscuro se ubica más o menos en el primer cuarto de la escala. Solo uno de cada seis hombres está «por encima de la media» en términos absolutos. El atractivo sexual no es algo que se acostumbre a cuantificar de este modo, por lo que me voy a permitir explicarlo en un contexto más familiar: si viésemos este gráfico en términos de cociente intelectual tendríamos un mundo en el que las mujeres opinarían que el 58 por ciento de los hombres sufre algún daño cerebral.

Vamos a ver, los hombres de OkCupid no son para nada feos —lo he comprobado haciendo un experimento: comparé un grupo extraído al azar de nuestros usuarios con otro parecido extraído de una red social y obtuve las mismas valoraciones para ambos—, y resulta que se obtienen unos patrones muy parecidos a este en cualquier otra web de contactos de las que he visto: Tinder, Match.com y DateHookup, que juntas abarcan a cerca de la mitad de la población soltera de Estados Unidos. Lo que ocurre es que hombres y mujeres hacen unos cálculos diferentes respecto al sexo. En la revista *Harper's* lo explicaron a la perfección: «Las mujeres son proclives a quejarse de las experiencias sexuales que han tenido y hombres, de las que no han tenido». En el gráfico puede apreciarse perfectamente cómo funciona esto. Y yo añadiría: los hombres del gráfico deben de estar cargadísimos de quejas.

En una curva beta se plasma lo que podríamos considerar el resultado de haber tirado una moneda al aire multitud de veces: representa las probabilidades superpuestas de numerosos eventos binarios. En este caso la moneda masculina está bien compensada y salen más o menos las mismas veces cara (que equipararé a positivo) que cruz. Pero en nuestros datos vemos que la moneda femenina está descompensada: solo sale cara una vez de cada cuatro. Hay muchos procesos naturales, como la predicción del tiempo, a los

que se les puede aplicar el modelo de curva beta, y gracias a los registros obsesivos de unos cuantos fanáticos de la meteorología he podido comparar nuestras valoraciones de persona a persona con algunos patrones climáticos históricos. En el caso que nos ocupa, la actitud masculina recuerda bastante a la función de predicción de cielos nublados en la ciudad de Nueva York. La psique femenina, siguiendo la misma técnica de medición, se acercaría a una ubicación con los cielos más encapotados que los de Seattle.

Seguiremos tirando de este mismo hilo a lo largo de los tres primeros grandes apartados que abarca *Dataclismo*: los datos de las personas que contactan. El atractivo sexual —cómo cambia y qué lo causa— será nuestro punto de partida. Veremos por qué, técnicamente, una mujer está en su momento cumbre a los 21 años, así como la importancia de tener un tatuaje destacado, pero enseguida iremos más allá de los contactos puramente carnales. Veremos qué nos pueden decir los tuits sobre la comunicación moderna y lo que pueden decirnos las amistades de Facebook sobre la estabilidad de un matrimonio. Las fotos de perfil que ponemos en la red son una bendición y a la vez una cruz: convierten casi cualquier sitio (Facebook, las webs de empleo y, por descontado, las de contactos) en un concurso de belleza. Echaremos un vistazo a lo que ocurrió cuando en OkCupid las eliminamos durante un día para ver qué pasaba. El amor no es ciego, aunque hemos encontrado pruebas de que debería serlo.

La segunda parte se centra en los datos que evidencian discordia. Empezaremos por fijarnos en esa primerísima división que existe entre los humanos: la raza, un asunto que por primera vez podemos abordar al nivel de persona a persona. Nuestros datos privilegiados ponen de manifiesto actitudes que la mayoría de la gente nunca admitiría en público, y veremos que el sesgo racial no solo es fuerte y habitual, sino que se repite casi palabra por palabra (o, mejor dicho, número por número) en cualquier sitio web. El racismo puede ser también una cosa interna: un solo hombre, su teclado y sus prejuicios. Veremos qué nos tienen que decir las búsquedas de Google sobre la palabra más odiada del país, y qué nos dicen también sobre el propio país. Pasaremos luego a explorar la disensión sobre la belleza física apoyándonos en un paquete de datos miles de veces mayor que cualquier otro que se haya analizado nunca antes. La fealdad tiene un sorprendente coste social que por fin somos capaces de cuantificar. A partir de ahí, veremos lo que nos descubre Twitter sobre la indignación. Esta red social permite que la gente esté conectada minuto a minuto, pero también puede separarnos con idéntica rapidez. La indignación colectiva que posibilita aporta una nueva nota de

violencia a la más antigua modalidad de reunión humana: la muchedumbre. Veremos si logra también que lleguemos a un nuevo entendimiento.

Cuando lleguemos al tercer apartado del libro ya habremos visto cómo son los datos cuando dos personas interactúan, para bien o para mal. Aquí vamos a centrarnos en el individuo en singular. Veremos cómo se expresa la identidad étnica, sexual y política centrándonos en las palabras, las imágenes y los referentes culturales que escoge la gente para expresarse. Estas son cinco de las frases más típicas de una mujer blanca:

mis ojos azules
pelirrojo y
conducir un todoterreno
chica de campo
me encanta el aire libre

¿Es un haiku de Carrie Underwood o son datos? ¡Tú eliges! Exploraremos las palabras que escribe la gente en público. También veremos cómo actúa y habla la gente en privado, con un ojo puesto en aquellos asuntos en los que las etiquetas y los actos divergen: los hombres bisexuales, por ejemplo, hacen que nos cuestionemos nuestra noción de una identidad clara. Después recurriremos a un amplio surtido de fuentes —Twitter, Facebook, Reddit e incluso Craigslist— para vernos a nosotros mismos cuando nos encontramos en casa, tanto físicamente como figuradamente. Y concluiremos con la pregunta que surge de manera natural en un libro sobre este tema: ¿cómo consigue uno preservar su intimidad en un mundo donde es posible llevar a cabo una exploración de este nivel?

A lo largo de todo el libro veremos que Internet puede ser un sitio emocionante, brutal, amoroso, indulgente, decepcionante, sensual e indignante. Y claro que lo es: lo componen seres humanos. Sin embargo, al recopilar toda esta información, me percaté de que los datos no recogen las vidas de todos. Si no tienes un ordenador o un *smartphone*, no sales aquí. Lo único que puedo hacer es reconocer el problema, rodearlo y esperar a que desaparezca.

Entretanto, me gustaría decir que el alcance de sitios como Twitter o Facebook, e incluso el de los datos de mi web de contactos, es sorprendentemente exhaustivo. Si no usas muchos de estos servicios *online*, puede ser que no lo aprecies. Cerca de un 87 por ciento de los estadounidenses están conectados a Internet, y esa cifra se mantiene prácticamente en todos los estamentos demográficos. Urbanos y rurales, ricos

y pobres, negros, asiáticos, blancos y latinos, todos están conectados. El uso de Internet es menor (alrededor del 60 por ciento) entre los muy ancianos y quienes tienen menor nivel de estudios, motivo por el que he trazado en estas páginas una «línea de edad» bastante previa a la tercera edad —los 50— y por el que he optado por no abordar el tema de la educación. Más de un tercio de los estadounidenses entran en Facebook cada día. La red social tiene 1300 millones de cuentas en todo el mundo. Dado que cerca de una cuarta parte de la población mundial tiene menos de 14 años, eso significa que algo así como el 25 por ciento de los adultos del planeta tienen cuenta en Facebook. En las webs de contactos que he estudiado para *Dataclismo* se han registrado unos 55 millones de usuarios estadounidenses durante los últimos tres años (como he dicho antes, eso es una cuenta por cada dos personas solteras del país). Twitter es un caso de especial interés demográfico. Se trata de una brillante historia de éxito tecnológico y la empresa está consiguiendo casi por sí sola que se enriquezca una amplia franja de la población de San Francisco. Pero se trata de una red social fundamentalmente populista, tanto por la «apertura» de la plataforma como por quienes optan por utilizarla. Por ejemplo, no presenta diferencias significativas en usuarios por género; personas con estudios de bachillerato tuitean igual que usuarios universitarios; los latinos emplean la red social igual que los blancos, y los negros, el doble. Y luego, por supuesto, está Google. Si el 87 por ciento de los estadounidenses usa Internet, el 87 por ciento de ellos ha usado Google.

Estas grandes cifras no demuestran que me haya hecho una idea completa de nada, pero por lo menos sugieren que esa idea está por llegar. Y, en todo caso, lo perfecto no debería estar reñido con lo mejor que ha existido hasta ahora. Los datos con los que trabajaremos abarcan a miles de veces más personas que cualquier encuesta Gallup o Pew, eso por descontado. Lo que ya no es tan obvio es que en realidad son mucho más inclusivos que la mayoría de los estudios del comportamiento llevados a cabo en el ámbito académico.

Es un problema conocido —aunque apenas se ha debatido públicamente— que casi todas las ideas fundacionales de las actuales ciencias del comportamiento se basaron en el estudio de pequeños grupos de alumnos universitarios. Cuando estaba en la universidad, en el hospital Mass General me pagaron unos 25 dólares por inhalar durante una hora un gas ligeramente radiactivo que actuaba de marcador y después me pidieron que llevara a cabo diversas tareas mentales mientras me tomografiaban el cerebro. «No te va a causar ningún daño», me dijeron. «Será como si hubieras pasado un año en un avión», añadieron. «No es nada», continuaron. Lo que no me dijeron —y de

lo que entonces tampoco me percaté— fue que mientras estaba ahí tumbado con una leve resaca en una especie de escáner o TAC, leyendo palabras y pulsando botones con el pie, lo que hacía era ejercer del típico varón humano. Un amigo mío también se presentó a aquel estudio. Era un universitario blanco, igual que yo. Me apuesto lo que sea a que la mayoría de los sujetos del estudio lo eran. Nada más lejos del típico varón humano.

Comprendo cómo funciona: en persona, conseguir una serie de datos que sean verdaderamente representativos de algo es muchas veces más difícil que llevar a cabo el propio experimento que queremos ejecutar. Si eres un profesor o un estudiante de posdoctorado que quiere progresar, te basas en lo que se llama un «muestreo por conveniencia», es decir, los alumnos de tu universidad. Pero eso supone un gran problema, sobre todo si investigas sobre creencias y comportamientos. Hasta tiene un nombre en inglés. Se llama investigación WEIRD (raro, en inglés): acrónimo de blanco (*white*), educado (*educated*), industrializado (*industrialized*), rico (*rich*) y democrático (*democratic*). Y la mayoría de los estudios sociales que se publican son WEIRD^[1].

Varios de estos problemas proliferan también en mis datos. Todavía tiene que pasar cierto tiempo hasta que los datos digitales logren tachar definitivamente de la lista eso de «industrializado». Pero dado que la tecnología se suele ver como un «ámbito elitista» —una imagen que muchos del propio sector se empeñan en incentivar—, me veo obligado a diferenciar entre los emprendedores e inversores de capital riesgo que uno encuentra en el panorama público de la tecnología, que se dedican a gesticular mientras sueltan su cháchara por el micrófono inalámbrico y que, de hecho, suelen ser gente muy WEIRD, y los propios usuarios de esos servicios, que son mucho más normales. No pueden evitar serlo, porque el uso de esos servicios — Twitter, Facebook, Google y demás— es la norma.

En cuanto a la autenticidad de los datos, gran parte de ella viene avalada por los propios hechos, puesto que Internet ha pasado ya a formar parte de la vida cotidiana. Pensemos en los datos de OkCupid. Tú le indicas a la web cuál es tu ciudad, tu sexo, tu edad y a quién buscas y esta te ayuda a encontrar a alguien con quien quedar para tomar un café o una cerveza. Se supone que tu perfil eres tú, la versión verdadera de ti. Si subes una foto de alguien más guapo para ponerla en tu perfil, o si finges ser mucho más joven de lo que eres, seguramente conseguirás más citas. Pero piensa en cuando quedes en persona con alguien: esperará encontrar a la persona que ha visto en la web. Si tu verdadero yo no se le parece, la cita prácticamente habrá terminado en

cuanto aparezcas. Este es un ejemplo de la tendencia general: conforme se mezclan los mundos *online* y *offline*, la inevitable presión social mantiene a raya muchos de los impulsos más embusteros de Internet.

La gente que usa estos servicios, tanto webs de contactos y citas como redes sociales y agregadores de noticias, anda a tientas por la vida, como se ha hecho siempre. Pero ahora lo hacen con sus móviles y sus portátiles. Casi sin darse cuenta, han generado un archivo extraordinario: ahora hay bases de datos por todo el mundo que contienen años y años de anhelos, opiniones y caos. Y como esos datos están almacenados con precisión escrupulosa, no solo pueden acabar analizándose, sino que se pueden analizar con un alcance y una flexibilidad que eran inimaginables hace solo una década.

Me he pasado varios años recopilando y descifrando estos datos. No solo los de OkCupid, sino los de casi todos los demás sitios web importantes. Y aun así nunca he podido librarme del todo de una duda inquietante y que, dadas mis simpatías por el ludismo, me irrita todavía más: me da la impresión de que escribir un libro sobre Internet es un poco como hacer un dibujo muy bonito sobre las películas de cine. ¿Para qué molestarse? Esa es la pregunta que me hago en mis horas bajas.



Hay un documental genial sobre Bob Dylan titulado *Don't Look Back* que vi varias veces cuando iba a la facultad. Mi mejor amigo, Justin, estudiaba cine. En un determinado momento de la película, después de una fiesta, Bob discute con un tipo cualquiera sobre quién ha tirado o quién no ha tirado alguna cosa de vidrio en la calle. Están los dos claramente borrachos. El clímax de la discusión es este fragmento del diálogo, que no me he quitado de la cabeza durante quince años:

Dylan: Conozco a mucha gente que es como tú y habla como tú.

Tipo de la fiesta: Que te den por culo. Eres un pez gordo, ¿sabes?

Dylan: Lo sé, tío. Sé que soy un pez gordo.

Tipo de la fiesta: Sé que lo sabes.

Dylan: Soy un pez más gordo que tú, tío.

Tipo de la fiesta: Yo soy un don nadie.

Dylan: Exacto.

Y entonces aparece alguien y se ponen todos a hablar de poesía. Es una de esas noches. Pero aquí está el asunto: estrellas del *rock* o no, los peces gordos han sido siempre la banda sonora de la humanidad. Conquistadores, magnates, mártires, salvadores, incluso granujas (sobre todo granujas)... a través de sus vidas hemos narrado nuestra historia, hemos ido marcando nuestra progresión desde las riberas fangosas de un par de ríos a lo que sea que seamos ahora. Desde el faraón Narmer, en el año 3100 a. C., el primer ser humano cuyo nombre conocemos, hasta Steve Jobs y Nelson Mandela, la gente ha organizado el mundo basándose en un armazón de héroes. Narmer fue el primero de una antigua lista de reyes. Los escribas han cambiado, pero la lista ha seguido creciendo. Los años sesenta, con todo aquello de «la gente al poder» y tal, son un ejemplo perfecto: fue la época de Lennon y McCartney, de Dylan y Hendrix, no de un «tipo de la fiesta». Vamos, que la existencia del hombre de a pie nunca ha merecido quedar registrada, salvo cuando se cruzaba en la vida de alguien que se convirtió en una leyenda.

Pero esta asimetría está llegando a su fin: el pez chico, la masa crepitante que somos todos los demás, está logrando por fin plasmar su vida en los registros. Dado que Internet ha democratizado el periodismo, la fotografía, la pornografía, la beneficencia, la comedia y tantas otras materias de la trayectoria de las personas, acabará, espero, por democratizar nuestras historias más esenciales. Esa banda sonora no es más que incipiente todavía y está sin refinar. Pero escribo este libro para sacar a la luz aquellos tenues patrones que yo, entre otros, he detectado. Es el eco del tren que se acerca cuando apoyas la oreja en las vías. La de los datos no es una ciencia exacta: todavía tenemos que entender, reconocer y refinar cosas como los sesgos a la hora de seleccionarlos. Pero la distancia entre lo que podría ser y lo que es se reduce un poco cada día que pasa, y yo escribo para que llegue un día en el que ambas converjan.

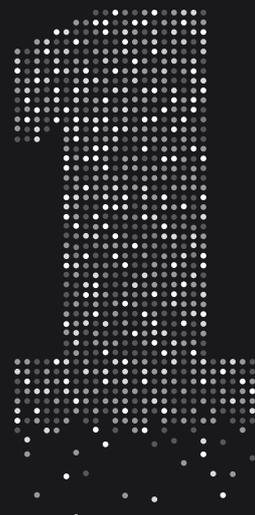
Sé que a mucha gente se le llena la boca cuando habla de los datos, y no es mi intención aquí decir que van a cambiar el curso de la historia — ciertamente no en la manera en la que lo hicieron el motor de combustión o el acero—, pero sí que creo que cambiarán lo que es la historia. Con datos, la historia puede llegar más al fondo. Puede hacerse más grande. A diferencia de lo que ocurre con la escritura en tablillas de arcilla, en papiro y pergamino, con el papel impreso, con el celuloide o con el papel fotográfico, el espacio de disco es barato y casi inagotable. En un disco duro hay sitio para que quepan más cosas además de los héroes. Por tanto, como yo no soy ningún héroe,

sino alguien a quien lo que más le gusta es pasar ratos con su familia y sus amigos y vivir la vida dignamente, eso resulta importante para mí.

Ahora bien, por mucho que me gustaría verme a mí mismo, a ti y a FulanitoDeTal81 en la misma página que al presidente en futuras obras dedicadas a la historia de esta década, imagino que la gente corriente siempre será más o menos anónima, tal como lo es incluso en estas mismas páginas. Ni los mejores datos pueden cambiar eso. Pero nos contarán a todos. Cuando en diez, veinte o cien años alguien quiera tomarle la temperatura a estos tiempos y pretenda entender los cambios producidos en esa historia —que quiera ver cómo la legalización del matrimonio homosexual reflejó y a la vez propició una mayor aceptación de la homosexualidad o cómo la población rural asiática se desarraigó y se volvió a consolidar en grandes núcleos urbanos— y llegar hasta su mismo tuétano, lo hará con los datos de Facebook, Twitter, Reddit y demás. Y si no, nuestro hipotético escritor habrá fracasado.

He tratado de abarcar todo esto en ese revoltijo que es mi título. *Kataklysmos* es como se llamaba en griego al diluvio del Antiguo Testamento y de ahí llegó a su actual acepción. La alusión tiene doble sentido: está, por supuesto, el hecho del aluvión de datos sin precedentes. Lo que se recopila hoy es tan profundo que roza lo insondable; es claramente un diluvio de cuarenta días y cuarenta noches en comparación con la pobre llovizna de las épocas precedentes. Pero también está la esperanza de un mundo transformado, de que el diluvio arrastre consigo tanto la raquítica comprensión que teníamos ayer como la todavía limitada visión que tenemos hoy.

Este libro lo componen una serie de viñetas, de ventanitas que se abren a nuestra vida: lo que nos une, lo que nos separa, lo que nos hace ser quienes somos. Conforme sigan llegando más y más datos, las ventanas se harán cada vez más grandes, pero hoy tenemos ya muchas cosas que ver, y el primer vistazo es siempre el más emocionante. Te animo a que te encarames al alféizar.



PARTE 1

Lo que nos une

1.

La ley de Wooderson

En los lugares más escarpados del mundo, como en los Andes, la gente se sirve de teleféricos para llegar a los sitios: un par de vagonetas que circulan por cables conectados a una polea montada en lo alto de la montaña. El peso de una vagoneta al descender hace que suba la otra: los dos vehículos se desplazan por contrapeso. He aprendido que en eso consiste ser padre. Los mismos años que me harán ir para abajo, harán que mi hija vaya para arriba. Y, por favor, que sea así. Yo me pliego gustosamente a ese tránsito, por supuesto, sobre todo porque cada momento transcurrido es un momento que he vivido con ella, pero eso no implica que no eche de menos los tiempos en los que tenía todo el pelo castaño y ninguna mancha rara en la piel. Mi niña tiene dos años y puedo asegurar que no hay nada que te haga ver más claro el paso del tiempo que las arrugas que descubres que tienes en el dorso de la mano mientras le estás enseñando a otra manita regordeta a contar: uno, dos tres...

Pero un tipo que ha tenido un bebé y tiene arrugas no es noticia. Puedes empezar por lo que sea que estén pregonando esta semana los del departamento de *marketing* de Oil of Olay —mientras escribo esto están con la idea de «corrección de color» del rostro con una pasta cremosa de tono beis que bien podría ser barro de las laderas de Alsacia o la misma esencia de la caca de vaca— e ir retrocediendo hasta el mito de Hera y los celos furiosos. La gente siempre se ha obsesionado con lo de envejecer y volverse fea, lo ha hecho desde que existen la obsesión y la fealdad. ¿Acaso no son nuestros temas eternos e ineludibles la muerte y los impuestos? Y, dependiendo de cómo cierre el ejercicio el gobierno de turno, estos últimos son cada vez menos de fiar. Pues ahí lo tienes.

De adolescente —y me sorprende darme cuenta de que estaba entonces más cerca de la edad que tiene ahora mi hija que de mis actuales 38 años— me encantaba la música punk, sobre todo el pop punk. Los grupos eran básicamente como Green Day pero en versión más creída y menos capaz. Cuando los recuerdo y vuelvo a escucharlos hoy, todo aquel fenómeno me parece irreal: hombres de pelo en pecho que se juntan en tríos y cuartetos por obra de alguna fuerza invisible para gimotear sobre cosas como las novias y

lo que come la gente. Pero en su día aquellos grupos me parecían la hostia. Y como eran demasiado molones para imprimirse carteles, me tuve que conformar con pegar las carátulas de sus discos y los flyers de sus conciertos en las paredes de mi cuarto. Mis padres se cambiaron de casa hace mucho — dos veces, de hecho—, así que estoy bastante seguro de que mi antiguo dormitorio es ahora la buhardilla de alguien y no tengo ni idea de adónde ha ido a parar toda aquella parafernalia que coleccioné. Ni siquiera me acuerdo de cómo era la mayor parte de todo aquello. Solo logro recordarlo y sonreír; y avergonzarme.

Hoy en día un chaval de 18 años cuelga una foto en su muro y ese muro nunca será derribado. Y no solo el chaval podrá retroceder en el tiempo, rebuscar entre aquella basura y preguntarse: «¿En qué estaría yo pensando?», sino que podemos hacerlo todos los demás, incluidos los investigadores. Y, lo que es más, pueden hacerlo con todo el mundo, no con una sola persona. Y más aún, pueden relacionar a ese chaval con lo que ocurrió antes y con lo que está por ocurrir, porque ese muro cubierto de tótems le sigue desde aquel dormitorio de la casa de sus padres hasta su primer piso, hasta el piso de su novia, hasta su luna de miel y, sí, también, hasta la sala de maternidad cuando nace su hija, momento a partir del cual procederá a empapelarlo con millones de actualizaciones de la niña comiendo papilla.

Puede que un padre primerizo sea más sensible a las señales de las distintas etapas del crecimiento de su hijo. Es prácticamente lo único de lo que hablas con otra gente y cada pocos meses sales del pediatra con una nueva serie de mediciones. Pero esas señales siguen llegando mucho después de dejar de consultar babycenter.com y de que el pediatra deje de darte citas. Lo que pasa es que dejamos de prestar atención a esos datos. Los ordenadores, sin embargo, no tienen nada mejor que hacer; prestar atención a los datos es su única tarea. No pierden el álbum de recortes, ni viajan, ni se emborrachan, ni se vuelven seniles, ni siquiera pestañean. Lo único que hacen es estar ahí y recordar. La mirada de fases de nuestra vida, que antes quedaban relegadas a la memoria o, como mucho, a una caja de zapatos llena de fotos, se están volviendo permanentes. Y por mucho que les pese a todos aquellos que hayan puesto una autofoto borrachos en Instagram, la oportunidad que ello nos brinda para comprender las cosas, si la manejamos con cuidado, es evidente.

Lo que acabo de describir, el muro y lo mucho que acumulamos a lo largo de la vida, es lo que los sociólogos denominan *datos longitudinales* —datos que se obtienen de seguir a las mismas personas a lo largo del tiempo— y el

resto son especulaciones más sobre cómo será la investigación en el futuro. Todavía no disponemos plenamente de esa capacidad porque Internet, como registro humano generalizado, es aún demasiado reciente. Aunque parezca difícil creerlo, incluso Facebook, por muy piedra de toque y caballo de batalla que sea, es algo grande desde hace solo seis años. ¡Ni siquiera tiene edad de entrar en la escuela primaria! Una información de esa profundidad es todavía algo en lo que estamos trabajando, literalmente, día a día. Dentro de diez o veinte años seremos capaces de responder a preguntas como por ejemplo: ¿cuánto le altera a una persona tener publicados a la vista de todo el mundo todos los momentos de su vida, desde la infancia? Pero también sabremos mucho más sobre por qué se separan los amigos o de qué manera se van filtrando en la corriente general las nuevas ideas. Puedo imaginarme el potencial a largo plazo de las hileras y columnas de mis bases de datos, y todos podemos verlo, por ejemplo, en lo que nos promete la Cronología de Facebook: los datos generan un nuevo modo de llenar ese transcurso del tiempo, si no exactamente una nueva ciencia.

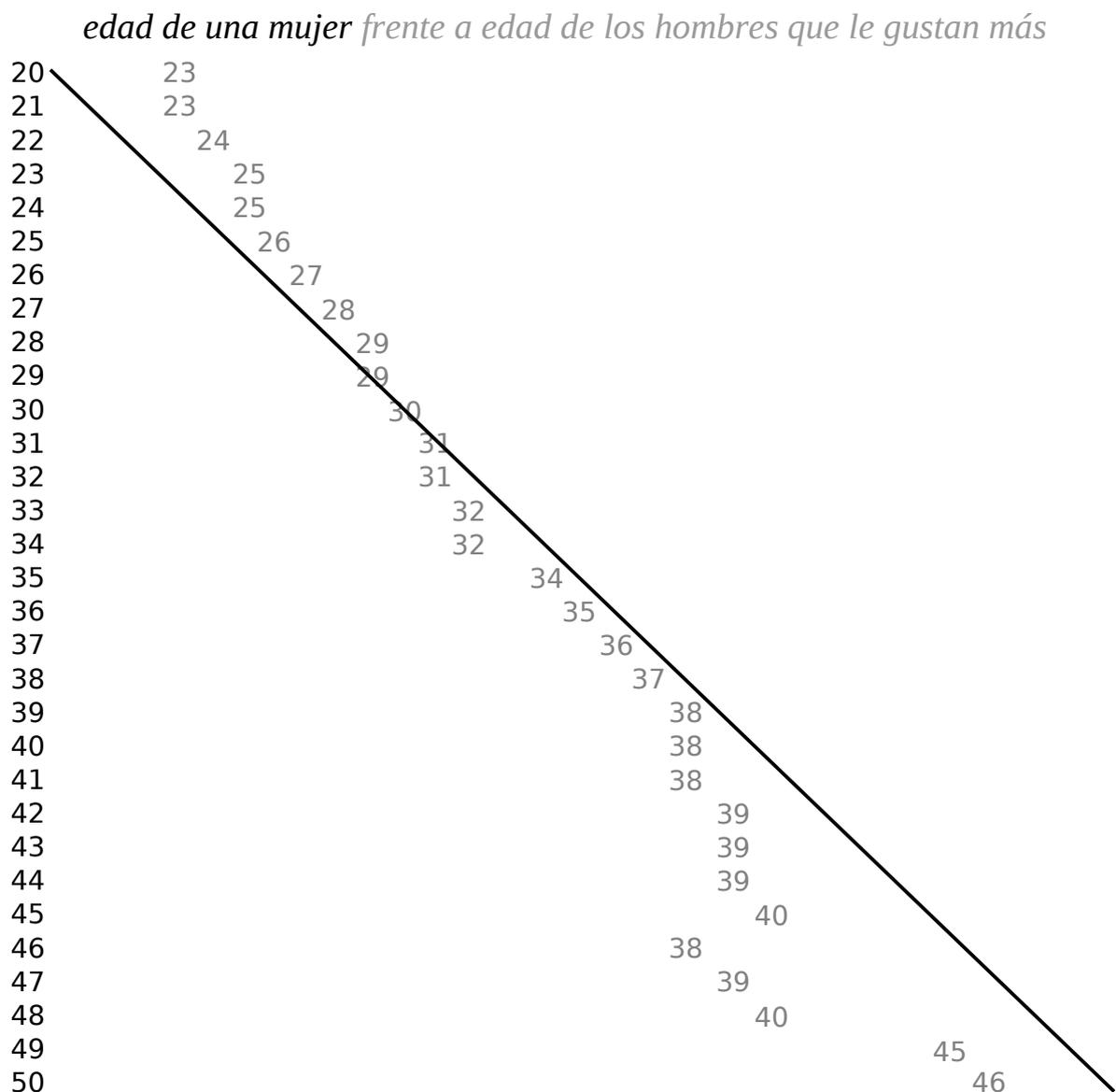
Incluso ahora, en determinadas situaciones, logramos encontrar una excelente muestra de ello, una especie de anticipo de esas posibilidades. Podemos tomar grupos de personas en distintos momentos de su vida, compararlos y obtener un esbozo básico del marco de la vida. Este método no funcionará con los gustos musicales, por ejemplo, porque la propia música evoluciona también a lo largo del tiempo, de modo que el análisis no se puede controlar. Pero hay ciertos universales fijos que pueden ayudar y, según los datos de que dispongo, el nexo entre belleza, sexo y edad es uno de ellos. En esto también existe la posibilidad de ir marcando puntos cruciales, así como de ir dejando al descubierto vanidades y vulnerabilidades que quizá hasta ahora no habían sido sino meras sombras de la verdad. Y al hacerlo nos acercaremos a un tema que ha obsesionado a escritores, pintores, filósofos y poetas desde que existen tales vocaciones, tal vez con menos arte (aunque algo de arte tiene), pero con una nueva y centelleante precisión. Como siempre, lo interesante reside en el espacio que hay entre el pensamiento y el acto, y yo os voy a enseñar cómo encontrarlo.

Empezaré por las opiniones de las mujeres. Todas las tendencias que se muestran a continuación las confirman los diversos datos que tengo sobre sexos, pero en pro de la concreción emplearé solo cifras extraídas de OkCupid. Esta tabla muestra la edad de los hombres que una mujer encuentra más atractivos. La he distribuido de manera poco habitual; enseguida veréis por qué.

edad de una mujer frente a edad de los hombres que le gustan más

20	23
21	23
22	24
23	25
24	25
25	26
26	27
27	28
28	29
29	29
30	30
31	31
32	31
33	32
34	32
35	34
36	35
37	36
38	37
39	38
40	38
41	38
42	39
43	39
44	39
45	40
46	38
47	39
48	40
49	45
50	46

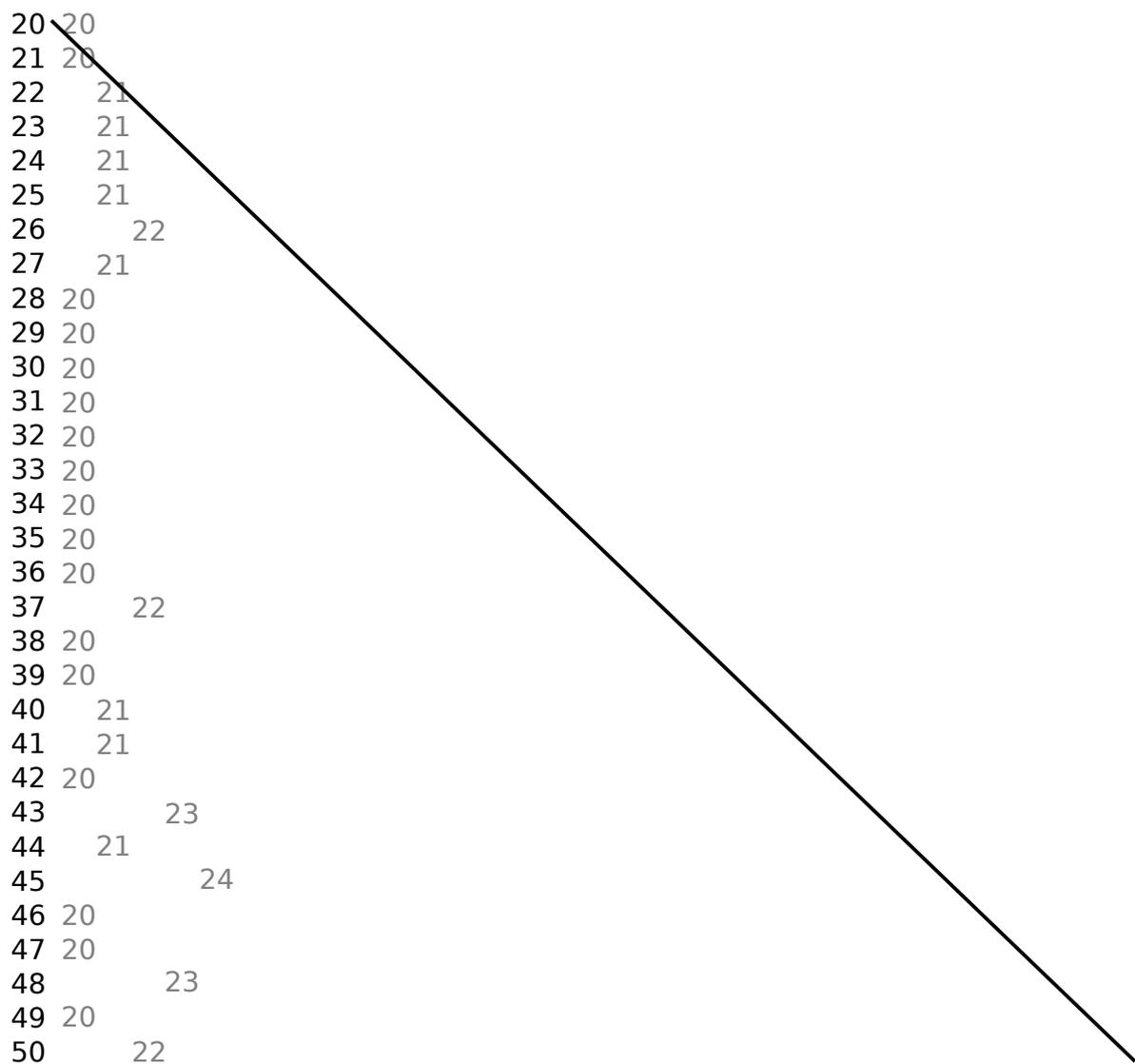
Si empezamos a leer por arriba, vemos que las mujeres de 20 y 21 años prefieren a hombres de 23, que las mujeres de 22 prefieren a hombres de 24, y así sucesivamente hasta que la mujer alcanza la cincuentena, donde vemos que valora como mejores a los hombres de 46. Estos no son datos de una encuesta, son datos basados en millones de preferencias expresadas en el acto de buscar una cita de pareja. Y con solo fijarnos en las primeras entradas nos queda claro cuál es el quid de la tabla: una mujer quiere que el hombre sea más o menos de la misma edad que ella. Si tomamos una edad en negro de menos de 40, veremos que el número en gris siempre se le acerca mucho. La tendencia general se aprecia mejor si distribuimos en horizontal la progresión de los valores en gris:



Esa diagonal punteada es la línea de «paridad de edades», en la que la edad de hombres y mujeres sería idéntica. No es que sea un canon matemático, sino que la he superpuesto únicamente a modo de guía para que la sigas con los ojos. Muchas veces, una situación lleva aparejada de manera inherente determinada geometría —por algo esta fue la primera ciencia— y eso lo aprovecharemos cuando sea posible^[2]. En este caso concreto, la línea punteada nos descubre dos transiciones, que coinciden con edades redondas. El primer punto de cambio es a los 30, cuando la tendencia de los números — las edades de los hombres— atraviesan la línea para ponerse debajo y nunca volver a ponerse por encima. Esa es la manera que tienen los datos de decirnos que, hasta los 30, una mujer prefiere a tipos algo mayores que ella; después, los prefiere un poco menores. Luego, a los 40, la progresión se aparta de la diagonal y desciende prácticamente en línea recta durante nueve años. Es decir, que el gusto de la mujer parece chocar contra un muro. Eso, o el atractivo físico del hombre cae en picado, según como uno quiera entenderlo. Si pretendemos señalar el punto en el que el atractivo sexual de un hombre alcanza su punto álgido, ahí lo tenemos: a los 40.

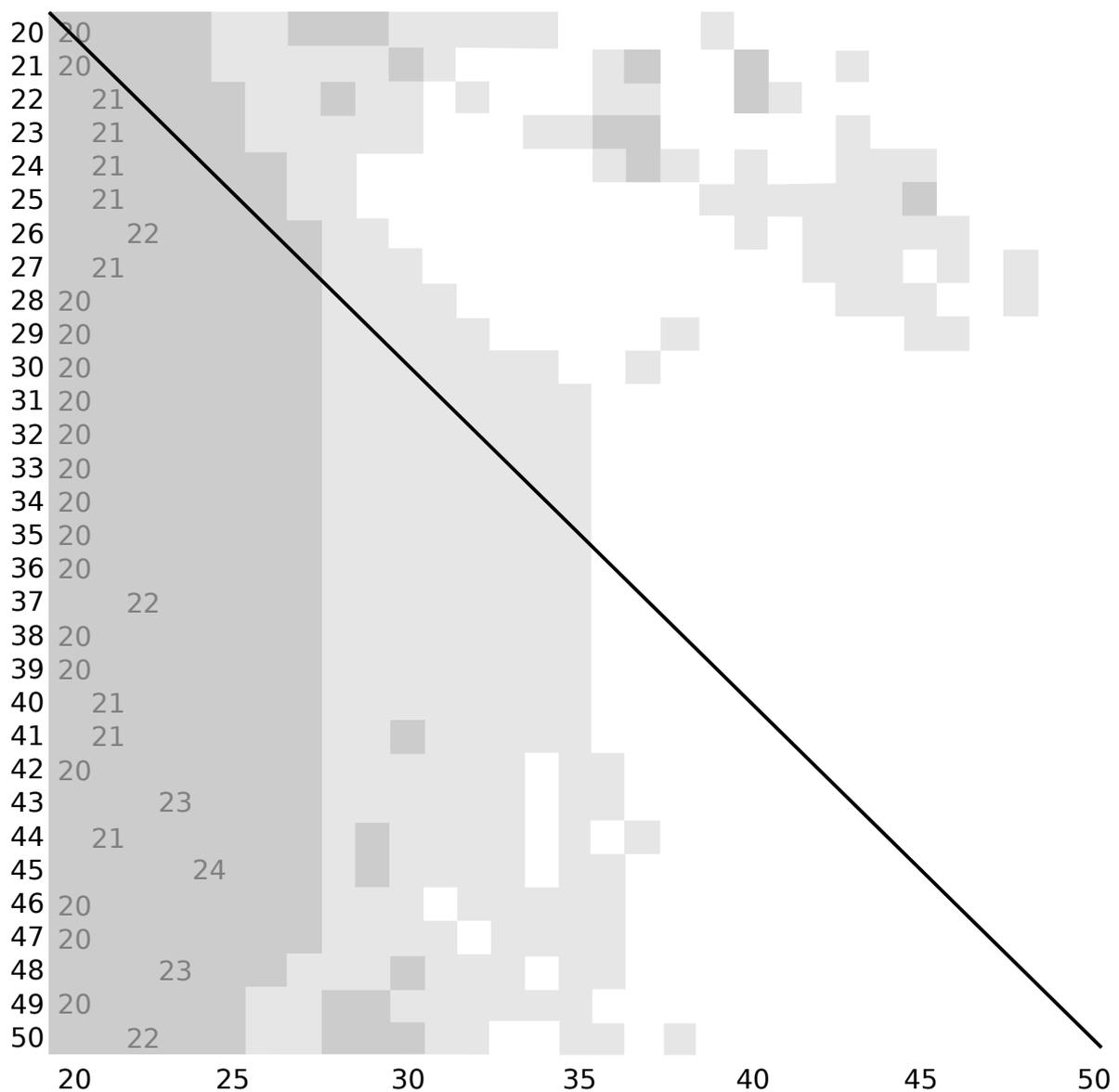
Ambas perspectivas (la de la mujer que hace la valoración y la del hombre que es valorado) son las dos caras de una misma moneda. A medida que una mujer se hace mayor, sus estándares evolucionan; y desde el punto de vista del hombre, la relación 1:1 entre los números grises y los negros indica que a medida que madura, las expectativas de sus contrapartes femeninas también maduran, prácticamente año a año. Él envejece y el punto de vista de ellas se adapta a él. Las arrugas, los pelos en la nariz, el renovado gusto por llevar bermudas militares... de algún modo todo eso pasa a ser satisfactorio, o al menos se ve compensado por otras virtudes. Compara esto con la caída en picado de las valoraciones que se dan en sentido contrario, de los hombres hacia las mujeres:

edad de un hombre frente a edad de las mujeres que le gustan más



Este gráfico —que apenas puede calificarse de tal, sino más bien de una simple tabla con un par de columnas— expresa un mensaje tan simple y directo como lo es el propio espacio negativo que vemos en él. Una mujer está de mejor ver cuando tiene veintipocos años. Punto. Y en realidad mi argumento no logra demostrarlo con la suficiente contundencia. Las cuatro edades femeninas con mayor valoración son los 20, los 21, los 22 y los 23 según todos los grupos de hombres excepto uno. El patrón general puede verse en el gráfico siguiente, donde he superpuesto unas tramas en los dos cuartiles superiores (es decir, en la mitad superior) de las valoraciones. También he añadido una escala de edades de las mujeres, en números negros, para que te orientes mejor:

edad de un hombre frente a edad de las mujeres que le gustan más

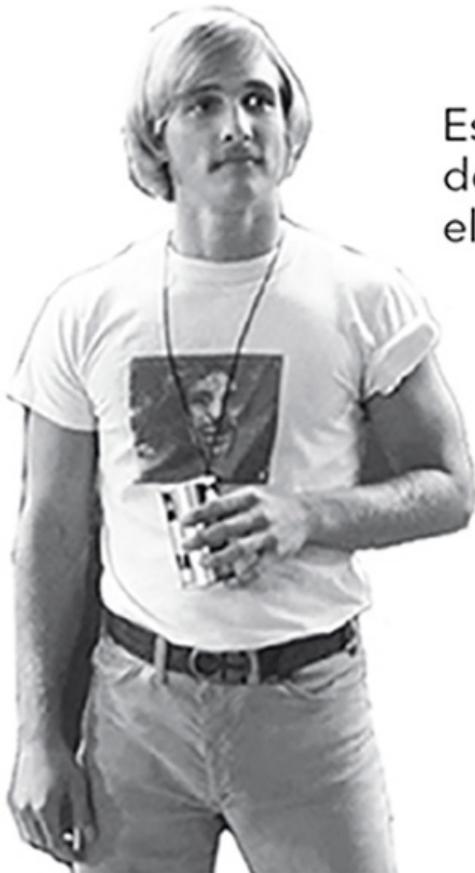


De nuevo vemos que la geometría habla por sí sola: el patrón masculino va mucho más allá de una simple predilección por las veinteañeras. Y después de que el hombre cumpla 30, la mitad de nuestra franja de edad (es decir, las mujeres de más de 35) podría perfectamente no existir. Cuanto más jóvenes, mejor; y según esto, si eso de que «se le pasa el arroz» es el momento en el que empieza su declive como persona, a una mujer heterosexual se le pasa el arroz en cuanto tiene la edad suficiente para beber alcohol.

Naturalmente, otra manera de ver esta fijación por la juventud es el hecho de que las expectativas del hombre nunca maduran. La idea que tiene un cincuentón de lo que es *sexy* es básicamente la misma que la que tiene un chico universitario, al menos en lo que atañe a la variable edad; como mucho, los hombres veinteañeros están más dispuestos a citarse con mujeres mayores. Fíjate en ese grupo aislado que vemos en la parte superior derecha del gráfico;

bien, pues ahí tienes a las presas de las mujeres asaltacunas, unos tipos que salen a dar una vuelta y, de repente, pam.

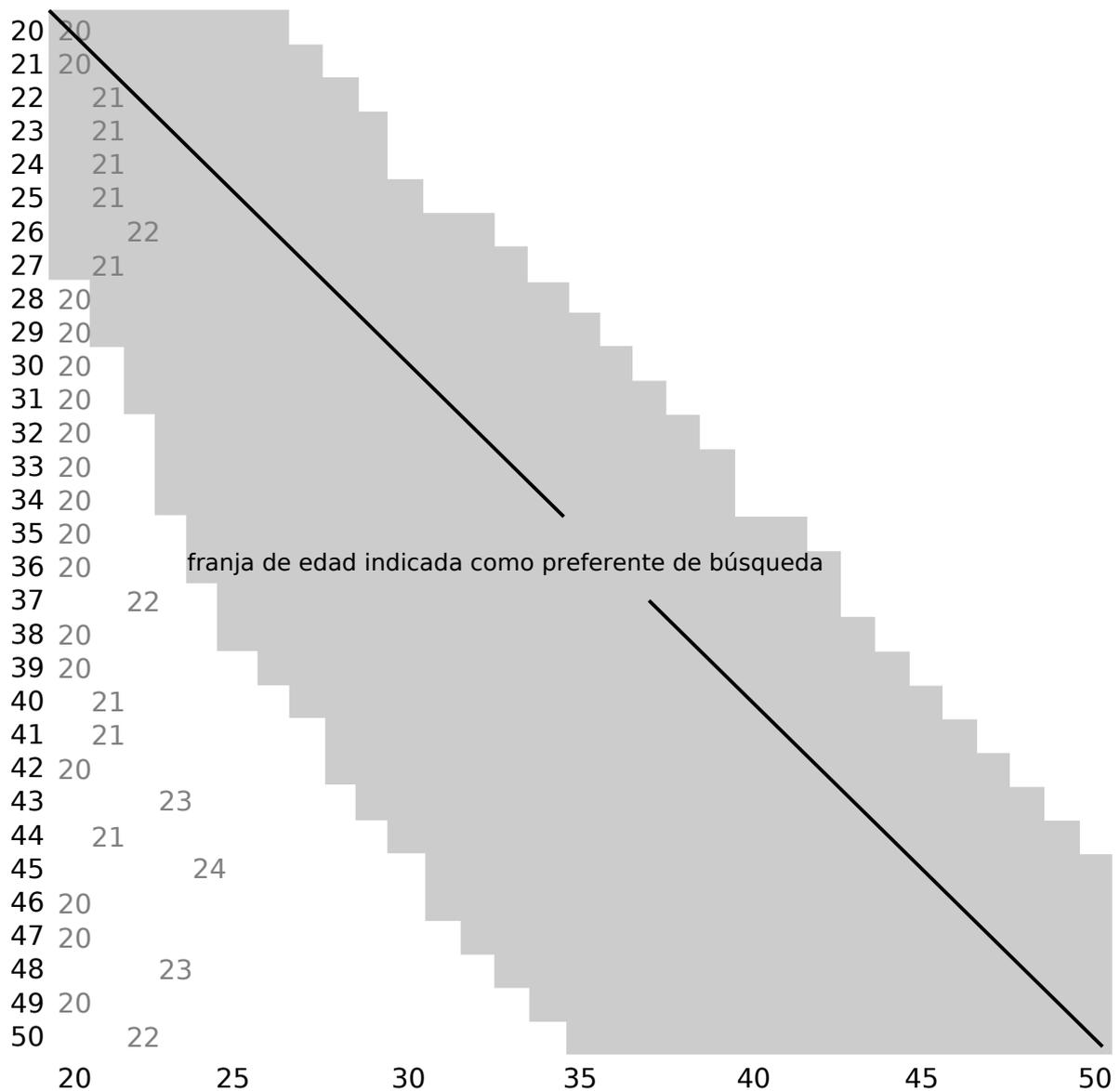
En términos matemáticos, la edad de un hombre y sus objetivos sexuales son variables independientes: la primera va cambiando mientras que la segunda nunca varía. A esto lo llamo la ley de Wooderson, en honor de su valedor más famoso, el personaje que interpreta Matthew McConaughey en *Movida del 76*.



Eso es lo que me gusta de las chicas de instituto: yo me hago mayor pero ellas siguen teniendo la misma edad.

A diferencia de Wooderson, lo que los hombres declaran querer es bastante distinto de los votos privados que acabamos de ver. Las valoraciones del gráfico anterior se asignaron sin ninguna otra exigencia específica que no fuese «valora a esta persona». Pero cuando les pides directamente a los hombres que escojan la edad de las mujeres que andan buscando obtienes resultados muy distintos. El espacio gris del gráfico siguiente es lo que los hombres nos dicen que quieren cuando se lo preguntamos:

edad de un hombre frente a edad de las mujeres que le gustan más

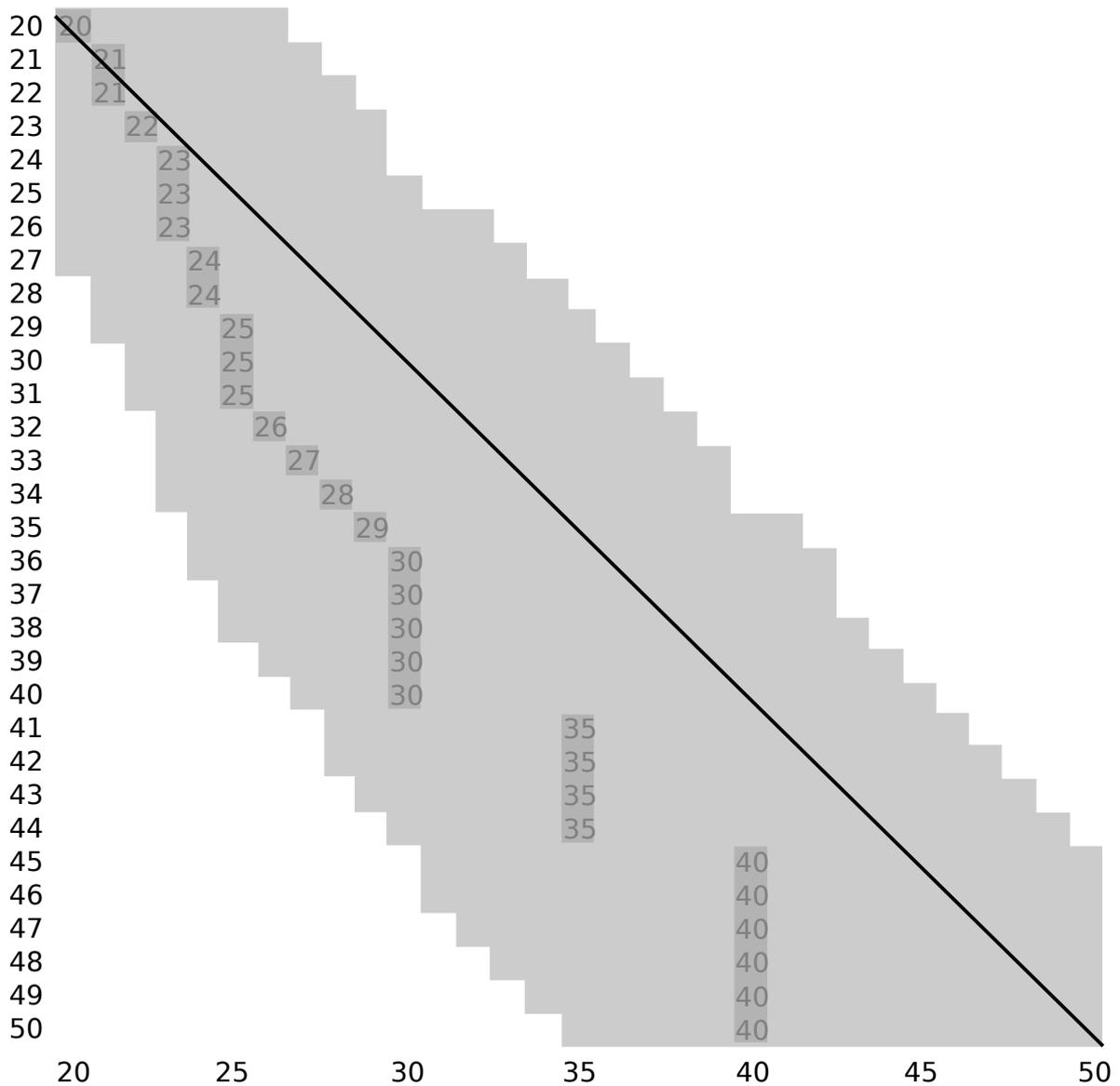


Como creo que nadie pretende engañarnos de manera intencionada cuando indica sus preferencias en OkCupid —hacerlo tiene pocos incentivos, ya que entonces estarías en una web que te da aquello que sabes que no quieres—, lo veo como una declaración de lo que los hombres piensan que supuestamente deben desear, en contraposición con lo que realmente desean. La brecha entre ambas ideas sigue ampliándose con la edad, aunque esa tensión parece quedarse en una especie de entrañable concesión a la hora de dejar de votar y pasar a la acción, como veremos.

El siguiente gráfico (el último de este tipo que vamos a ver) relaciona la edad con la mayor densidad de tentativas de contacto. Aquellas edades en las que más mensajes se reciben se expresan en los cuadrados gris oscuro que van discurriendo por la parte izquierda de la franja gris. Esos tres bloques verticales de cuadrados grises que se ven en la parte inferior del gráfico nos

muestran los saltos en el concepto que tienen los hombres de sí mismos al llegar a la edad madura. Casi puede oírse el ruido que hace el motor al reducir las marchas. A los 44, el hombre se encuentra cómodo acercándose a una mujer de 35. Luego, al año siguiente... se lo piensa mejor. Mientras que una diferencia de edad de nueve años le parece bien, por lo visto una diferencia de diez ya es demasiado.

edad de un hombre frente a edad de las mujeres a quienes más mensajes envía



Es en esta especie de tierra de nadie calculada —el equilibrio entre lo que quieres, lo que dices y lo que haces— donde debe surgir el romance: independientemente de lo que la gente vote en privado o de lo que prefiera en abstracto, no hay muchos hombres de 50 años que triunfen buscando mujeres

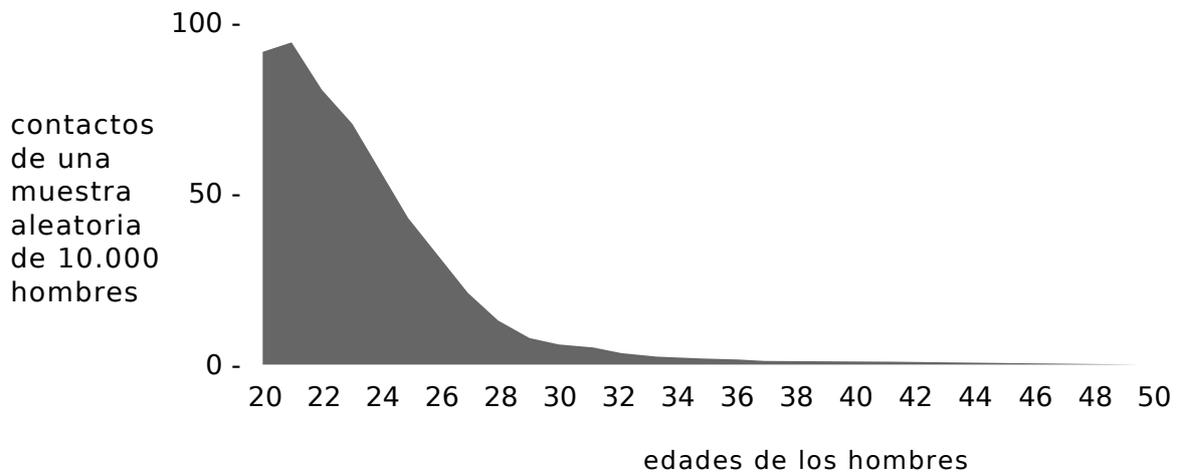
de 20. Para empezar, tienen en contra convenciones sociales. Y para seguir, lo de conseguir una cita es una cosa recíproca. Lo que quiere una persona es solo la mitad de la ecuación.

Cuando se trata de mujeres que toman la iniciativa y buscan hombres, dada la relación de atracción mujer-hombre que hemos visto al principio del capítulo (1 año:1 año), más las motivaciones ajenas a la apariencia física que empujan a las mujeres hacia los hombres mayores —las económicas, por ejemplo—, las mujeres envían más mensajes, y no menos, a un hombre a medida que este se hace mayor, hasta que supera en poco la treintena. A partir de ahí, la cantidad de contactos desciende, pero a un ritmo que no se aleja mucho del descenso general del propio número de mujeres disponibles. Piénsalo así: imagina al típico tío de 20 años que empieza a quedar con chicas como adulto (definición: no cuentan fiestas con colegas durante al menos uno de los actos de cortejo/consumación/ruptura) y, de un modo u otro, podrás tener una idea de todas las mujeres que estarán interesadas en él. Si entonces pudieras seguir el rastro de todo ese lote de mujeres a lo largo del tiempo, verías que la principal razón de que el chico vaya perdiendo opciones en ese grupo se debe a que algunas de ellas han dejado de estar solteras por haberse emparejado con otro. De hecho, ese segmento total de «interesadas» se incrementaría, porque a medida que él madura y supuestamente se enriquece y triunfa, esas cualidades atraen a mujeres más jóvenes al lote. Sea como sea, la edad por sí sola no le perjudica. Durante las dos primeras décadas de su vida de citas adultas, conforme él y las mujeres de su segmento maduran, a las que todavía están disponibles les parecerá una opción tan deseable como cuando tanto ellas como él tenían 20 años.

Si pudieras hacer lo mismo con la típica mujer de 20 años verías una historia muy distinta. Con los años, también ella iría perdiendo hombres de su reserva a causa del matrimonio, pero además perdería opciones debido al propio paso del tiempo: a medida que pasasen los años, irían quedando cada vez menos hombres solteros que la encontrasen atractiva. Su grupo de posibles citas es como una lata con dos agujeros: se vacía el doble de rápido.

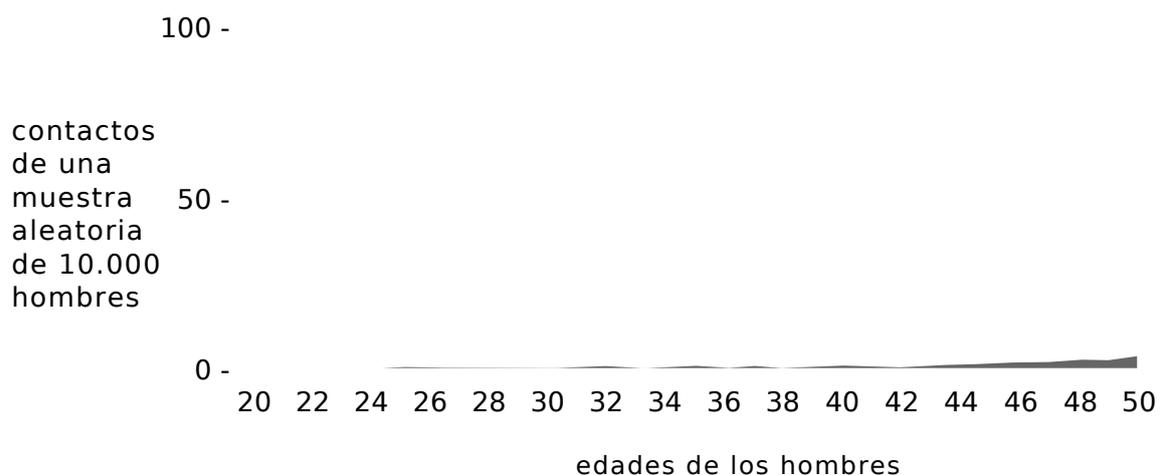
El número de hombres solteros mengua velozmente con la edad: según el censo de Estados Unidos, hay 10 millones de hombres cuya edad está entre los 20 y los 24 años, pero solo 5 millones de entre 30 y 34 años y solo 3,5 millones de 40 a 44. Si superponemos los patrones de preferencias que hemos visto antes a estos datos demográficos descendentes, nos haremos una idea de cómo cambian las opciones reales de una mujer con el tiempo. Este es el gráfico real de las posibles citas de una mujer de 20 años:

para una mujer de 20 años: número de hombres interesados, por edades (de 20 a 50 años)



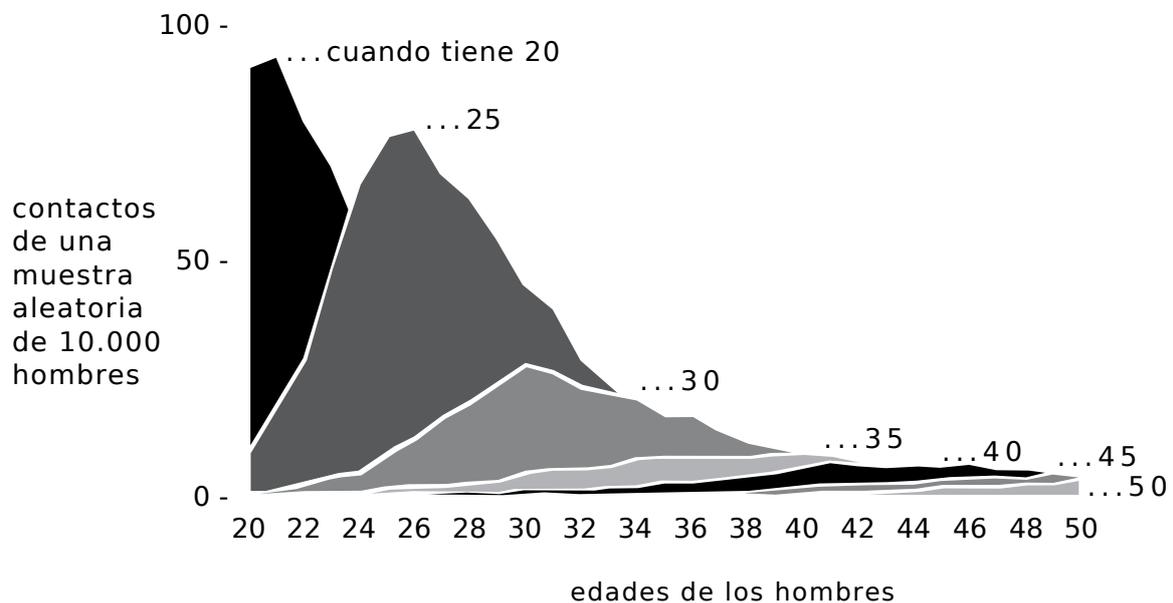
Sus coetáneos (hombres de veintipocos años) componen el mayor grupo, y a partir de ahí los números decrecen con rapidez. Los hombres de 30, por ejemplo, componen una parte ya muy pequeña del total. Son menos proclives a contactar con alguien tan joven, pese al interés que puedan haber expresado en privado, y además muchos hombres ya se han emparejado al llegar a esa edad. Cuando la mujer alcanza la cincuentena, estos son los que quedan (y siguen interesados), presentados en una escala idéntica. Es Bridget Jones hecha gráficos.

para una mujer de 50 años: número de hombres interesados, por edades (de 20 a 50 años)



Si comparamos las dos áreas de color, por cada cien hombres interesados en esa veinteañera, solo hay nueve que buscan a alguien treinta años mayor.

Este es el gráfico con las progresiones completas, según las diversas edades de la mujer, desde los 20 a los 50 años:



Es muy frecuente que en mi trabajo vea a dos individuos que están solos pero que, por algún motivo, no conectan. En este caso que estamos viendo ahora, se trata de dos grupos de personas que se buscan mutuamente con objetivos contradictorios. Las mujeres quieren que los hombres envejecen con ellas. Y los hombres siempre tienden a buscar la juventud. Una mujer de 32 años se registra, indica su preferencia de edades filtrándola entre 28 y 35 y empieza a explorar. Un hombre de 35 años entra, ajusta su preferencia de edades entre los 24 y los 40 y rara vez contacta con ninguna mujer que supere los 29. Ninguno de los dos encuentra lo que está buscando. Podría decirse que son como dos barcos que se cruzan durante la noche, pero tampoco sería exactamente eso. Los hombres sí que parecen haberse hecho a la mar, atraídos por un horizonte que se ve en la distancia. Pero, según yo lo veo, las mujeres siguen todavía en tierra firme, en la orilla, viéndoles perderse en alta mar.

2.

Sobredosis de ni fu ni fa

Para los Oscar de 2002 contrataron al director Errol Morris con el fin de que filmase un corto sobre por qué nos encanta el cine. La Academia quería iniciar la transmisión televisiva con un montaje trepidante de gente, famosos y no famosos, que hablase de sus películas favoritas. Mi amigo Justin era el director de *casting* de Morris, así que me puso en la lista. Nada garantizaba que acabase saliendo en el montaje final del corto, pero podía ponerme ante la cámara para la entrevista y ver qué tal salía.

Con ese enchufe, me citaron el mismo día que a gente de renombre: Donald Trump, Walter Cronkite, Iggy Pop, Al Sharpton y Mijaíl Gorbachov. Trump y Gorbachov se daban la espalda y por ahí, en alguna parte, hay una foto en la que salgo yo entre ambos, chupando cámara antes de que existiese el fenómeno del *photobombing*. Digo «en alguna parte» porque justo cuando saltó el *flash*, Trump chasqueó los dedos y su guardaespaldas le quitó la cámara a Justin. Trump escogió *King Kong* como su película favorita, porque naturalmente le gustan los simios que tratan de «conquistar Nueva York». Gorbachov, a través de un intérprete cuyo bigote debía de pesar cinco kilos, optó por *Gladiator*. En el minuto 2:01 del corto de Morris, los ojos superabiertos y la voz que dice «*La profecía*» son míos.

Es cierto que a mí me gusta una buena peli sobre el Anticristo más que a la mayoría de la gente, pero escogí *La profecía* más o menos al azar. Hay tantas películas buenas que no estoy muy seguro de cuál es mi favorita. Pero sí sé que la película que menos me gusta es *Pecker*, de John Waters. Me salí del cine. Dos veces. Fui a verla una vez con amigos, no pude soportar ese rollo del submundo *mondo-trasho*, por no hablar de aquellos acentos exagerados, y tuve que irme. El fin de semana siguiente, otros amigos fueron a verla y, puesto que John Waters es un autor respetado y yo soy un tío guay que está al día, me dije que podía existir alguna mínima posibilidad de que me hubiese equivocado la primera vez. Tampoco tenía nada mejor que hacer. Así que fui a verla otra vez.

La locura temporal de los 22 años es así. No digo que John Waters haga películas objetivamente malas, pero no son para mí. Ni para mucha gente. Y él eso lo adora, ese rechazo es prácticamente su carta de presentación como

director. Pongámoslo así: nadie sale de ver *Pecker* pensando que es una película que «ni fu ni fa», o te encanta o te largas de allí a los veinte minutos como hice yo, dos veces. Está hecho aposta^[3].

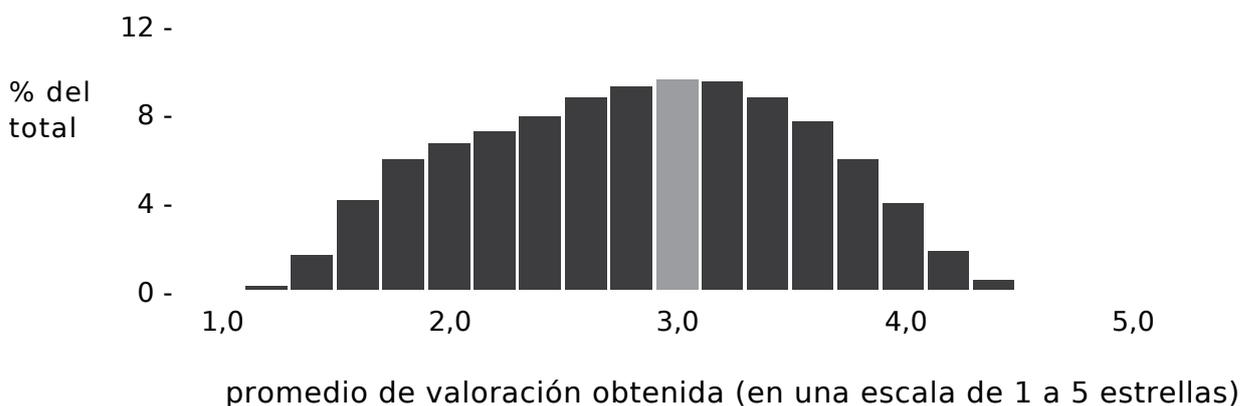
Los fans de Waters parecen adorarlo más todavía por ser pocos. Si haces una búsqueda en los perfiles de usuarios de OkCupid obtienes más resultados con su nombre que los que dan George Lucas y Steven Spielberg juntos. En Reddit tiene su página dedicada: [/r/JohnWaters](#)^[4] y, aunque no es que sea la URL de mayor tráfico del mundo, la gente pone cosas allí: noticias, vídeos antiguos, preguntas acerca de él y demás. También existe una página [/r/GeorgeLucas](#): contiene una sola publicación, la única que ha contenido siempre. Si escribes [/r/StevenSpielberg](#) en el campo de la URL, te saldrá el mensaje del servidor Reddit que reza «parece que no hay nada por aquí» porque, por muy buenas que sean sus películas, nadie se ha entusiasmado tanto como para hacerle una página. Incluso hay directores con mucho tirón en Internet como J.J. Abrams que no tienen página allí. Hace falta cierta motivación especial para hacer un *fan site*, y esa motivación suele intensificarse si tienes la impresión de formar parte de un grupo selecto y acosado. La devoción es como el vapor contenido en un pistón: la presión ayuda a ponerlo en marcha.

Como muchos otros artistas que le han precedido y seguido, Waters sabe exactamente cómo funciona el asunto: repeler a alguna gente atrae más a otra, y no solo he sacado este caso a colación por mi contienda de siempre con *Pecker*, sino porque Waters también es consciente de la universalidad del principio: no es solo cierto para el arte. Hay muchas citas tuyas buenas, pero esta en particular a mí me dice mucho: «La belleza es una apariencia externa que nunca vas a olvidar. Un rostro debería sobresaltar, no confortar». Tiene toda la razón, ya que en la música, en el cine y en una amplia variedad de fenómenos humanos un defecto es una cosa muy poderosa. Incluso al nivel de persona a persona, agradar de manera universal es ser relativamente ignorado. Desagradar a unos cuantos lleva aparejado gustar mucho más a otros. Y, concretamente, el atractivo sexual de una mujer se ve incrementado cuando algunos hombres la encuentran fea.

Esto se puede ver en las valoraciones de los perfiles de OkCupid. Dado que el sistema de valoración es de cinco estrellas, los votos tienen algo más de sustancia que un simple sí o no. La gente opina en diversos grados, lo que nos da margen para explorar. Para exponer este descubrimiento tendremos que emprender un breve viaje matemático. Este tipo de ejercicios son los que hacen que funcione la ciencia de los datos. Para armar un puzle tienes que

extender todas las piezas y después empezar a hacer pruebas. En el caso de terabytes de datos en bruto, dada la imposibilidad de recurrir a la búsqueda minuciosa, a la reducción y a la parsimonia, hay muy pocas cosas que simplemente te «salten a la vista».

Pensemos en un grupo de mujeres cuyo atractivo sea aproximadamente el mismo, pongamos que las que se han valorado en el centro de este gráfico:



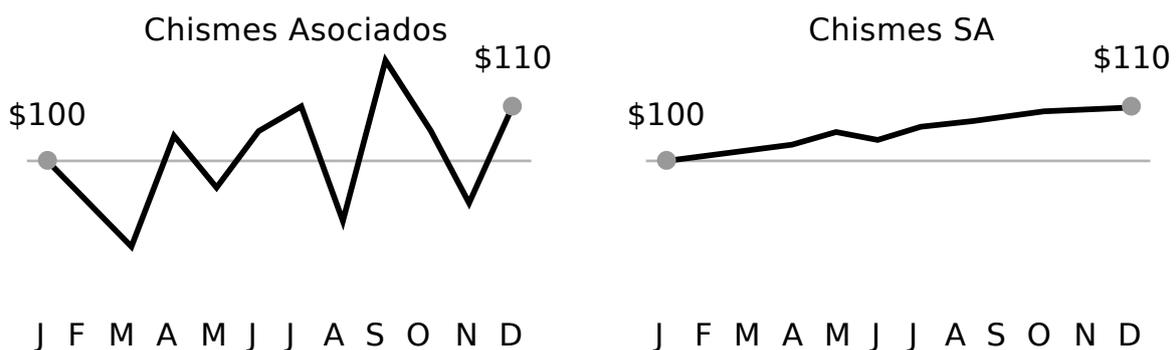
Ahora imaginemos a una mujer de ese grupo y pensemos en los muchos votos distintos que pueden asignarle los hombres: pensemos básicamente en cómo ha acabado estando ahí en el medio. Hay miles de posibilidades. Estas son unas pocas que me he inventado: combinaciones de una, dos, tres, cuatro y cinco estrellas, que todas suman un promedio de 3.

	número de hombres que han votado...					promedio del patrón
	"1"	"2"	"3"	"4"	"5"	
patrón A			100			3,0
patrón B		10	80	10		3,0
patrón C	10	20	40	20	10	3,0
patrón D	25	25		25	25	3,0
patrón E	50				50	3,0

Como habrás advertido, los patrones de voto que he escogido se van polarizando cada vez más desde el patrón A hasta el patrón E. El promedio de todas las hileras da ese mismo 3 central, pero se llega a ese promedio de

formas distintas. El patrón A encarna el consenso. Aquí, los hombres que otorgan el voto han hablado con absoluta unanimidad: esta mujer está exactamente en el medio. Pero cuando llegamos al final de la tabla, el promedio sigue centrado aunque ninguno de los individuos particulares exprese esa opción centrada. El patrón E muestra el método más extremo posible para llegar a un promedio equidistante: por cada hombre que ha asignado un voto de 1, otro ha asignado un voto de 5. Y el total obtenido es de 3 casi a pesar de sí mismo. Ese el método de John Waters.

Estos patrones ejemplifican un concepto llamado varianza. Se trata de una medida de cuánto se desvían los datos respecto de un valor central. La varianza se incrementa cuanto más se alejan los datos del promedio. En la tabla de arriba, la mayor varianza corresponde al patrón E. Una de las aplicaciones más habituales de la varianza es medir la volatilidad (y, por tanto, el riesgo) en los mercados financieros. Pensemos en estas dos empresas:



Ambas tuvieron beneficios de un 10 por ciento al final del año, pero como inversión difieren mucho. El valor de Chismes Asociados experimentó grandes oscilaciones a lo largo del año, mientras que Chismes SA creció poco a poco, con ganancias constantes cada mes. Calcular la varianza permite a los analistas condensar esas diferencias en un simple número y, aunque no difieran en nada más, los inversores prefieren con mucho el ritmo sosegado del gráfico de la derecha. Iguales beneficios, menos taquicardias. Naturalmente, cuando hablamos de relaciones amorosas, esas taquicardias son los beneficios, y ahí está el quid de la cuestión. Resulta que la varianza afecta tanto a la atención sexual que recibe una mujer como a su atractivo general.

La cantidad de mensajes que reciben las mujeres de cualquier grupo cuyo grado de belleza sea idéntico está sumamente relacionada con la varianza. Desde los ganadores de concursos de belleza hasta los menos agraciados, pasando por cualquier término medio, los individuos que obtienen más

atención serán siempre los que más polarizan la opinión. Y los efectos no son nimios: generar opiniones muy opuestas hará que recibas cerca de un 70 por ciento más de mensajes. Eso significa que la varianza te ayuda a subir varias «ligas» en la jerarquía de las citas amorosas. Por ejemplo, una mujer con una valoración muy baja (con percentil 20) y una varianza elevada entre sus puntuaciones recibe más o menos los mismos mensajes que una típica mujer de percentil 70.

Esto se debe en parte a que la varianza implica, por definición, que hay más personas a quienes les gustas mucho (y también a quienes les desagradas mucho). Y esos entusiastas —llamémosles tus *fanboys*— son los que más mensajes envían. Por lo tanto, si logras que más gente te ponga en la parte superior de las calificaciones (las cinco estrellas), te garantizas más acción.

Pero los votos negativos también tienen su papel en el juego. También atraen cierta atención. Por ejemplo, los patrones reales que en el siguiente gráfico se han ejemplificado como mujeres C y D obtienen cerca de un 10 por ciento más de mensajes que las mujeres A y B, por mucho que a esas dos mujeres de lo alto de la tabla se las haya valorado mejor en general:

	número de hombres que han votado...					promedio del patrón
	"1"	"2"	"3"	"4"	"5"	
mujer A	2	22	27	29	20	3,4
mujer B	10	13	31	28	18	3,3
mujer C	32	22	12	16	18	2,7
mujer D	47	13	6	19	15	2,4

He estado hablando de los mensajes como si fueran un fin en sí mismos, pero en una web de contactos los mensajes son precursores de resultados más tangibles como conversaciones en profundidad, intercambio de datos de contacto y, finalmente, citas cara a cara. La gente cuya varianza es más alta también obtiene más de estas cosas. Así, por ejemplo, la mujer D del gráfico tendría cerca de un 10 por ciento más de conversaciones, de encuentros y, probablemente, de relaciones sexuales que la mujer A, pese a que en términos de valoración absoluta sea mucho menos atractiva.

Por añadidura, los hombres que otorgan esas valoraciones de 1 y 2 no mensajean a esas mujeres: la gente casi nunca contacta con nadie a quien ha

valorado bajo^[5]. La cuestión es que tener detractores induce de algún modo a todos los demás a quererte más. La gente a la que no le gustas atrae por sí sola más atención hacia ti. Y sí, en su búnker subterráneo, Karl Rove esboza una sonrisa autocomplaciente mientras acaricia a un enorme sapo.

Y lo que contribuye al misterio del fenómeno es que OkCupid no hace públicas las valoraciones concretas sobre el atractivo (ni la varianza, por descontado) de ninguno de sus usuarios. Nadie toma decisiones conscientes basándose en esos datos. Pero la gente tiene una manera de captar las matemáticas que hay en las cosas, sean o no conscientes de ello, y aquí va lo que yo pienso al respecto. Supongamos que a un tipo le atrae una mujer cuyo aspecto sabe que no es convencional. Esa ausencia de convencionalidad implica que otros hombres la habrán descartado, o sea, que habrá menos competencia. Tener menos rivales incrementa sus posibilidades de triunfar. Puedo imaginarme a ese hombre mirando el perfil de la mujer, dando vueltas con el cursor y pensando para sí: «Me juego lo que sea a que no conoce a muchos tíos que piensen que es preciosa. De hecho, a mí me atrae por sus peculiaridades, no a pesar de ellas. Este es mi diamante en bruto», etcétera. Hasta cierto punto, es su impopularidad la que la hace atractiva para él. Y si nuestro usuario interesado estuviera dudando sobre si presentarse o no, eso podría darle ese empujoncito que le hacía falta.

Si miramos este fenómeno desde el ángulo opuesto —el lado de una varianza baja—, una mujer relativamente atractiva con valoraciones uniformes es alguien a quien cualquier hombre consideraría convencionalmente guapa. Y por tanto podría parecer más popular de lo que realmente es. Una mujer de atractivo estándar genera la impresión de que habrá otros tipos que irán detrás de ella, y eso le restará cada vez más interés. Nuestro amigo interesado pero indeciso pasará de largo.

Esta es mi teoría, al menos. Pero la idea de que la varianza es algo positivo se ha demostrado bastante bien en otros ámbitos. Los psicólogos lo llaman «efecto Pratfall»: mientras que en general eres una persona competente, cometer un pequeño error de vez en cuando hace que la gente piense que eres más competente. Los defectos hacen que lo bueno destaque más. Esta necesidad de imperfección puede ser sencillamente la manera de que todo encaje en nuestro cerebro. Nuestro sentido del olfato, que es el más conectado con el centro emocional del cerebro, prefiere las notas discordantes a las homogéneas. Los científicos lo han demostrado en sus laboratorios mezclando olores fétidos con otros agradables, pero la naturaleza, con su sabiduría evolutiva, se percató de ello mucho antes. El aroma

agradable que emiten muchas flores, como el azahar o el jazmín, contiene una porción significativa (alrededor de un 3 por ciento) de una proteína llamada indol, que se encuentra habitualmente en el intestino grueso y que huele en consonancia con ese hecho. Pero las flores no olerían igual de bien sin ella. Un poquito de mierda atrae a las abejas. El indol está presente también como ingrediente en los perfumes sintetizados por los humanos.

Podemos ver una aplicación de los datos de OkCupid en versión pública, por así decirlo, en el exclusivo mundo de las supermodelos. Las mujeres son todas profesionalmente hermosas: 5 estrellas de 5, por supuesto. Pero incluso a ese nivel tan elevado la cuestión es diferenciarse del resto por medio de la imperfección. La carrera de Cindy Crawford despegó cuando dejó de taparse el lunar. Linda Evangelista lucía aquel peinado tan duro: no se puede decir que la hiciese más hermosa, pero sí que la hacía más interesante. Kate Upton, al menos según los estándares del sector, tiene unos cuantos kilos de más. Si extraemos unas cuantas muestras del paquete de datos, mejor unas que nos sean más familiares y accesibles que las modelos de ropa de baño, nos ayudará a ver cómo funciona esto con las personas normales. Aquí tienes a seis mujeres, todas ellas con valoraciones generales moderadas pero que tienden a suscitar reacciones extremas en un sentido u otro: muchos síes y muchos noes, pero pocos ni fu ni fa.



Gracias a todas ellas por la confianza que han mostrado al permitirnos poner aquí su foto y hablar sobre ellas. Lo que ves en esta selección es lo que puedes encontrar en todo el corpus. Estas son personas que han optado deliberadamente por salirse del montón: por medio del arte corporal, adoptando una mueca sarcástica o porque les da por comer queso gratinado como si no hubiese un mañana. Así que te encuentras a mujeres relativamente normales con una peculiaridad inusual: como la mujer de la foto central de la hilera inferior, cuyo pelo de color azul no puedes apreciar en esa foto en blanco y negro. Y te fijas sobre todo en mujeres que han optado por lucir sus activos/pasivos particulares. Si eres capaz de sacarte, pongamos por caso, una valoración de 3,3 a pesar de los kilos de más o de la gente que odia los tatuajes o lo que sea, entonces te habrás dotado literalmente de más poder.

Así pues, al final, dado que todo el mundo en este planeta tiene algún tipo de defecto, la moraleja es esta: sé tú mismo y hazlo con coraje. Lo cierto es que tratar de encajar por el mero hecho de encajar resulta contraproducente. Sé que esto recuerda peligrosamente a esas frases que se bordan en punto de cruz, y que los cuadros en punto de cruz, que son las presentaciones en PowerPoint de eras pasadas, son lo opuesto a la ciencia. También suena mucho a los consejos que le da una madre, junto con una palmadita en la cabeza, a su hijo de nariz desproporcionada y aparato en los dientes cuando este tiene 14 años y no entiende por qué no es más popular. Pero sea como sea, ahí está, en los números. Como ya he dicho, la gente es capaz de captar las matemáticas que hay en las cosas, sobre todo las madres. Solo quisiera que la mía me hubiese dicho que cuando tienes 14 años las camisetas de ositos no molan.



3.

Escribir en el muro

En otros tiempos a la nostalgia se la conocía como *mal du Suisse*, la enfermedad suiza. Los mercenarios suizos luchaban por toda Europa y al parecer destacaban por su deseo de volver a casa. Se ponían sentimentales y cantaban baladas pastoriles en lugar de luchar. Y cuando eres el rey de Francia y tienes que quemar hugonotes, las canciones no te sirven. Se prohibieron las baladas. En la Guerra Civil americana la nostalgia supuso tal problema que dejó fuera de combate a unos 5000 soldados y 74 de ellos murieron por su causa, al menos según los archivos médicos. Dadas las circunstancias, ponerse mortalmente enfermo es bastante comprensible, pero también es cierto que aquellos eran tiempos de sanguijuelas y amputaciones con serrucho, así que quién sabe lo que de verdad ocurría. Es curioso pensar que mucha de la gente que abandonaba su hogar lo hacía para irse a la guerra y que en gran parte de los primeros escritos sobre la nostalgia, que por aquel entonces se consideraba una verdadera enfermedad, se menciona a los soldados. No puedo sino pensar en ese pasado de tonos sepia: me gusta imaginarme que en 1863 había científicos que trabajaban afanosamente y a contrarreloj a ambas orillas del Potomac para desarrollar las superarmas que acabarían definitivamente con las guerras: los anuarios escolares.

La verdad es que ni siquiera sé si en los institutos se siguen haciendo esos anuarios. Cuesta entender para qué harían falta ahora que tenemos Facebook, aunque según el último informe cuatrimestral de la empresa, los menores de 18 años ya no usan tanto Facebook como hace un tiempo. Así que puede ser que los chavales necesiten de nuevo tener el anuario impreso, no sé^[6]. Pero sea cual sea la manera de entablar contacto que emplean hoy los adolescentes —a través de Snapchat, WhatsApp o Twitter—, estoy seguro de que lo hacen mediante palabras. Las imágenes forman parte del atractivo de todos estos servicios, pero son pocas las cosas que puedes decir sin un teclado. Hasta en Instagram los comentarios y los pies de foto son esenciales: al fin y al cabo, la foto no ocupa más que unos cuantos centímetros cuadrados. Pero las palabras son las palabras. Siguen siendo la manera de transmitir sentimientos y de entablar contacto.

De hecho, por mucho que se denueste el efecto que ejerce en nuestra cultura la tecnología, estoy convencido de que hasta el adolescente más reticente ha escrito en 2014 mucho más en su vida que lo que yo o mis compañeros de colegio habíamos escrito a principios de los noventa. Por entonces, cuando tenías que hablar con alguien usabas el teléfono. Yo me limitaba a escribir alguna que otra nota de agradecimiento y tal vez una carta al año. El típico alumno de instituto de hoy debe de superar eso en una sola mañana. Internet tiene muchas cosas deplorables, pero hay una cosa que a mí siempre me ha convencido: es un mundo de escritores. En la Red, tu vida pasa por tus palabras. Trabajas, socializas, ligas... y todo ello lo haces tecleando. Creo sinceramente que esta iniciativa de Internet tiene algo de epistolar, de austeniano. Da igual qué palabras usemos o cómo mecanografiemos nuestras cartas: nos escribimos los unos a los otros más que nunca. Aunque, a veces, cosas como:

k pasa tia

sean lo único que tenemos que decirnos.



El comandante Sullivan Ballou era uno de los soldados del ejército de la Unión que estaban en el Potomac, sufriendo y con morriña. Al principio de la serie *The Civil War*, de Ken Burns, un narrador nos lee la carta de despedida que le escribió a su mujer, a su «muy querida Sarah», y es un momento importante y conmovedor del documental. El comandante la escribió desde el campamento justo antes de la primera gran batalla de la contienda y fue herido de muerte pocos días después. Aquellas palabras tuyas fueron las últimas que escuchó su familia y llevaron a su hogar la enorme aflicción que el país tendría que soportar en los años venideros. Debido a su difusión, la carta de Ballou se ha convertido en una de las más famosas que se han escrito; si busco «carta famosa» en Google, me aparece en segundo lugar. Es un hermoso escrito, pero piensa en todas las demás cartas que nunca serán leídas en voz alta, que fueron quemadas, se traspapelaron, se las llevó el viento o, sencillamente, acabaron descomponiéndose en algún lugar.

Hoy no tenemos que depender del feliz accidente de la conservación para saber lo que alguien pensaba o cómo hablaba y tampoco necesitamos que ese alguien sea representativo de muchos. Todo se conserva, no solo lo que un hombre dirige a una esposa antes de una batalla, sino lo de todos dirigido a

todos, antes y después e incluso en medio de todas y cada una de nuestras batallas personales. En YouTube podemos encontrar lecturas de la carta de Ballou, y muchos de los comentarios vienen a ser de este estilo: «Ya no las hacen como antes». Eso es cierto. Pero lo que sí se hace, o sí hacemos, aporta una belleza y una riqueza de índole distinta: una poesía no de frases líricas, sino de conocimiento. Estamos en la cúspide de una época de cambios trascendentales en el estudio de la comunicación humana y de lo que esta trata de fomentar: la conexión personal y comunitaria.

Cuando uno quiere conocer cómo escribía la gente, sus palabras desnudas y expuestas son el mejor sitio para empezar a aprender, y tenemos paletadas de ellas. Se habrán escrito más palabras en Twitter durante los dos próximos años que las que contienen todos los libros impresos hasta hoy. Es el paradigma de la nueva comunicación: breve y en tiempo real. De hecho, Twitter fue el primer servicio que no solo fomentaba la brevedad y la inmediatez, sino que la exigía. La pregunta que hace a sus usuarios es: «¿Qué está pasando?», y les ofrece 140 caracteres para que se lo cuenten al mundo. Y la repentina popularidad de Twitter, al igual que su repentina redefinición de la escritura, pareció confirmar los temores de que Internet estaba «matando la cultura». ¿Cómo iba la gente a seguir escribiendo bien (y pensando bien) en ese nuevo espacio condensado? ¿En qué se iba a convertir una mente tan limitada? El actor Ralph Fiennes habló por boca de muchos cuando afirmó: «No tienes más que mirar Twitter para demostrar el hecho de que muchas de las palabras inglesas que se emplean, digamos, en las obras de Shakespeare o en las novelas de P. G. Wodehouse... se usan ahora tan poco que la gente ni siquiera sabe ya qué significan».

Hasta el análisis más básico demuestra que el lenguaje que se usa en Twitter poco tiene de degradado. En la siguiente tabla he comparado las palabras inglesas más usadas en Twitter con el Oxford English Corpus, una recopilación de cerca de 2500 millones de palabras de fuentes modernas de toda índole: textos periodísticos, novelas, blogs, estudios académicos, de todo. El OEC es el censo canónico del vocabulario inglés actual. Solo he tabulado las cien palabras más usadas de las decenas de miles que emplea la gente, lo que puede parecer una muestra un tanto insignificante, pero prácticamente la mitad de todo lo que se escribe se compone a base de esas palabras (tanto en Twitter como en el OEC). Lo más reseñable de la lista de palabras de Twitter es esto: pese a todas las quejas de los adalides encaramados a las almenas de la fortaleza del inglés, en la lista de las cien más usadas solo hay dos palabras de la «ciberlengua»: «rt», de *retuit*, y «u»,

de *you* (tú). Podría pensarse que las contracciones, gramaticalmente correctas o no, serían un recurso habitual en una forma de comunicación que nos limita a 140 caracteres, pero parece que en lugar de eso la gente escribe bien a pesar de esa limitación y no se empeña en chocar tercamente con ella. Por otra parte, cuando calculas la extensión media de las palabras de la lista de Twitter, ves que son más largas que las del OEC: 4,3 caracteres en lugar de 3,4. Y, más allá de la longitud, fijémonos en el contenido del vocabulario de Twitter. He resaltado las palabras que aparecen allí exclusivamente para facilitar la comparación:

	OEC	Twitter		OEC	Twitter
1	the	to	51	when	back
	be	a		make	an
	to	i		can	see
	of	the		like	more
	and	and		time	by
	a	in		no	today
	in	you		just	twitter
	that	my		him	or
	have	for		know	as
10	I	on	60	take	make
	it	of		people	who
	for	it		into	got
	not	me		year	here
	on	this		your	want
	with	with		good	need
	he	at		some	happy
	as	just		could	too
	you	so		them	u
	do	be		see	best
20	at	rt	70	other	people
	this	out		than	some
	but	that		then	they
	his	have		now	life
	by	your		look	there
	from	all		only	think

	they	up		come	going
	we	love		its	why
	say	do		over	he
	her	what		think	really
30	she	like	80	also	way
	or	not		back	come
	an	get		after	much
	will	no		use	only
	my	goof		two	off
	one	but		how	still
	all	new		our	right
	would	can		work	night
	there	if		first	home
	their	day		well	say
40	what	new	90	way	great
	so	time		even	never
	up	from		new	work
	out	go		want	would
	if	how		because	last
	about	we		any	first
	who	will		these	over
	get	one		give	take
	which	about		day	its
	go	know		most	better
50	me	when	100	us	them

Mientras que la lista del OEC es bastante deslucida —muchas partículas y modificadores, lenguaje diligente que te conduce hasta la recompensa de algún verbo o sustantivo—, en la de Twitter no hay cabida para ese tipo de funcionariado: cada una de las palabras es un jefazo. Y te encuentras que cosas tan brillantes como estas:

love (amor)
happy (feliz)
life (vida)
today (hoy)

best (mejor)
never (nunca)
home (hogar)

... están entre esas cien palabras. Es incluso posible que Twitter esté mejorando la escritura de sus usuarios, ya que los obliga a expresar unas pocas letras para transmitir significado. Encarna la famosa máxima de William Strunk, «Prescinde de las palabras innecesarias», llevándola a la mínima expresión de las pulsaciones en un teclado. A quien tuitea no le queda más opción que ser conciso y, si nos paramos a pensarlo, la limitación de caracteres en realidad explica el porqué de esa longitud de palabras ligeramente mayor. Cuando se dispone de un espacio restringido para trabajar, usar palabras más largas implica dejar menos espacios entre ellas, es decir, desperdiciar menos caracteres. Aunque los pensamientos que se expresan en Twitter puedan ser más o menos sesgados, no existen pruebas de que estén mermados.

Mark Liberman, profesor de Lingüística en la Universidad de Pennsylvania, concluyó más o menos lo mismo: en una respuesta directa al señor Fiennes, calculó la longitud media de las palabras de *Hamlet* (3,99 caracteres) y de una recopilación de relatos de Wodehouse (4,05) y descubrió que ambas eran más cortas que las de su muestra de Twitter (4,80)^[7]. Se trata de uno de los muchos lingüistas comparativistas que han empezado a hurgar en los datos de Twitter. Un equipo de estudiosos del estado de Arizona fue más allá del recuento y la longitud de las palabras y ahondó en las emociones y el estilo de lo escrito, donde descubrieron varias cosas sorprendentes: para empezar, Twitter no altera la forma de escribir de una persona. Entre los muchos ejemplos que fueron siguiendo, vieron que el usuario que emplea «u» en lugar de «you» para la segunda persona en correos electrónicos o mensajes de texto, lo seguirá usando en Twitter. Pero, del mismo modo, si en general lo escribe «you», lo hará en todas partes: en Twitter, en los textos, en el correo, etcétera. La decisión de referirse a uno mismo en primera persona como «I», en mayúscula, o «i» sigue el mismo patrón. Es decir, que el estilo de determinada persona no cambia al pasar de un medio a otro; no existe ninguna «bajada de nivel». Uno escribe como escribe, sea donde sea. Los lingüistas también calcularon la densidad léxica de Twitter, la proporción de palabras portadoras de significado como verbos y nombres, y descubrieron que no solo era más elevada que en los correos electrónicos, sino que era equiparable a la que muestran los textos del sistema de control del nivel sintáctico que se usa en la revista *Slate*. Todo apunta a la misma conclusión: que Twitter no ha

alterado nuestra manera de escribir sino que más bien la ha obligado a encajar en un espacio más pequeño. Al estudiar los datos, en lugar de encontrarnos un erial de árboles talados, dimos con un bosque de bonsáis.

Este tipo de análisis en profundidad (densidad léxica, frecuencia de términos) apunta a la verdadera naturaleza de la transformación que está hoy en curso. El cambio al que ha dado pie Twitter en el propio lenguaje no es nada comparado con el cambio que está ocasionando en el estudio del lenguaje. Twitter nos proporciona una concepción de las palabras en la que no solo se las considera los ladrillos que componen el pensamiento como conector social, que ha sido ciertamente el objetivo del lenguaje desde que la humanidad empezó a deambular erguida por el Serengeti. Y, a diferencia de los medios de comunicación anteriores, Twitter nos brinda un modo de rastrear esos vínculos a nivel individual. No solo puedes ver lo que dice otra persona, sino a quién se lo dice, cuándo y con qué frecuencia. La lingüística comparada lleva mucho tiempo siguiendo la pista a rasgos comunes de determinados grupos a través del lenguaje. Las palabras básicas suelen compartir sonidos parecidos (como *tres*, *trois*, *drei*, *three* y *thran*, respectivamente del español, el francés, el alemán, el inglés y el gujarati de la India) y esas raíces nos han proporcionado una idea de los movimientos de los genes y de la cultura a lo largo del tiempo. Los investigadores ya empiezan a agrupar a la gente por la lengua que emplean en Twitter. Aquí presento un extracto de uno de los primeros intentos de identificar las tribus y los dialectos emergentes, procedente de un corpus de 189 000 tuiteros que se han enviado 75 millones de tuits en inglés entre ellos.

subgrupos de Twitter por patrones de escritura

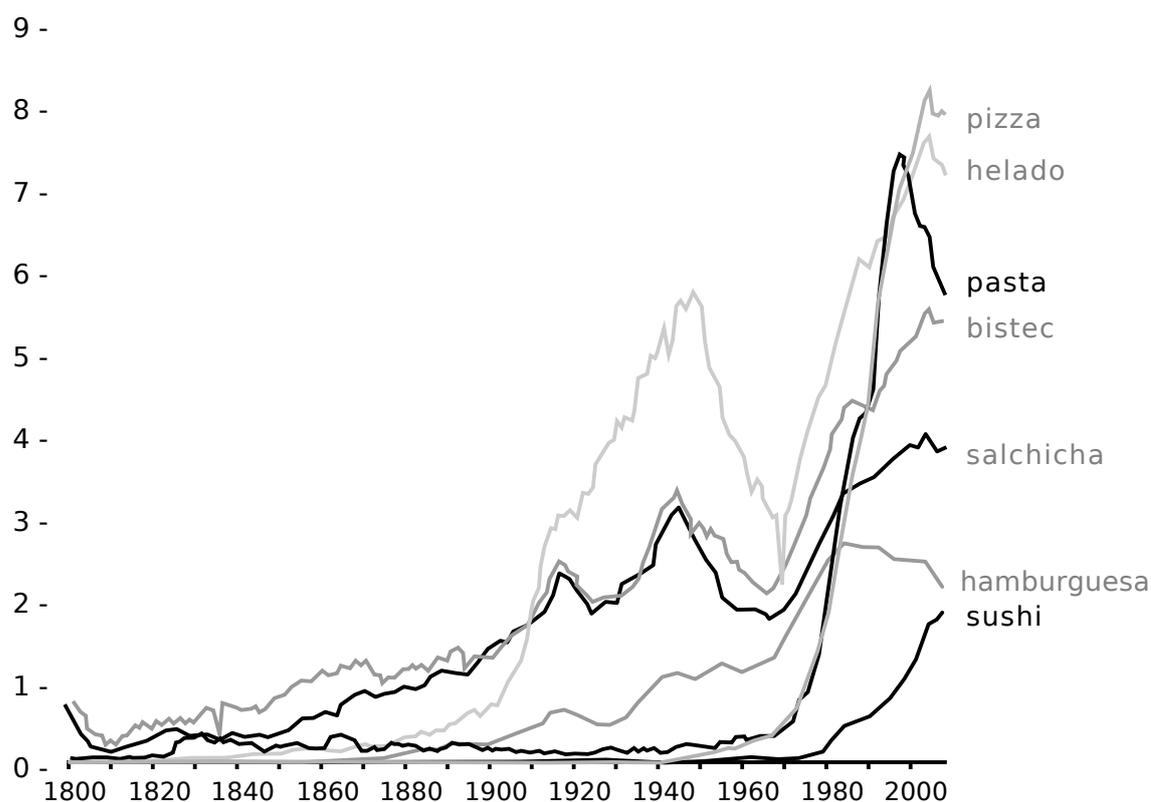
ejemplos habla	características	porcentaje de la muestra
nigga, popping, chillin	finales acortados (p.ej. -er => -a o -ing => -in)	14
tweetup, metrics, innovation	tech palabrería corporativa	12
inspiring, webinar, affiliate	tips autoayuda de marketing	11
etsy, adorable, hubby	jerga de artesanías varias	5
pelosi, obamacare, beck, libs	cháchara política	4
bibier, pleasee, youu, <33	tips autoayuda de marketing	2
anipals, pawesome, furever	juego de palabras basados en animales	1

Es importante destacar que el estudio agrupó a los usuarios solamente por sus palabras, a quién se las enviaron y lo que escribieron; estos grupos de lenguaje no estaban determinados *a priori*. El grupo que encabeza la lista es, de hecho, el más numeroso que encontraron los lingüistas, y resulta ser también el más locuaz (el que más tuits por cabeza envía), así como el más aislado. Cerca del 90 por ciento de los tuits enviados por ese grupo tiene su destino en ese mismo grupo y el lenguaje que emplean sus usuarios es el más acentuadamente característico: la mitad de sus cien palabras más representativas encajan en el patrón «finales acortados». En el gráfico se distinguen grupos tipificados por el argot, las referencias a la cultura popular, las jergas o los juegos simplones de palabras: gente que se junta por medio de su especial manera de hablar, un tipo de lenguaje (y de información) que es exactamente el que hasta ahora ha quedado excluido de la historia. Igual que saber cuáles fueron las últimas palabras dirigidas por un hombre a su esposa, saber cómo habla la gente cuando está entre amigos nos brinda un conocimiento más hondo de quiénes son. Tecnócratas, pirados de la política, gurús del *marketing*, fans de Robert Pattinson..., será interesante ver cómo todos estos grupos se mezclan y se recombinan en los próximos años, y seremos capaces de seguir todo el proceso a través de sus textos.

Una vez tenemos reunidos el lenguaje y los datos, la otra dimensión, el tiempo, es la que añade interés. Si miramos al futuro, vemos que servicios como Twitter serán indispensables. Si miramos al pasado, vemos que Google Libros se esmera en corregir nuestro punto ciego histórico: en colaboración con bibliotecas de todo el mundo, ha digitalizado 30 millones de libros distintos, grandes y pequeños, y además, fieles a su experiencia, los han hecho consultables. Ese corpus de datos ha dado pie a la creación de una nueva disciplina de los estudios culturales cuantitativos llamada culturonomía; su método principal consiste en controlar los cambios que experimenta el uso de las palabras con el transcurrir del tiempo. El largo alcance de esos datos (se remontan a 1800) permite una manera insólita de ver a las personas y aquello que les importa. Este es un pequeño gráfico que me gusta titular «Pizza ahora, pizza siempre»:

frecuencia de menciones escritas de determinados alimentos, 1800-2008

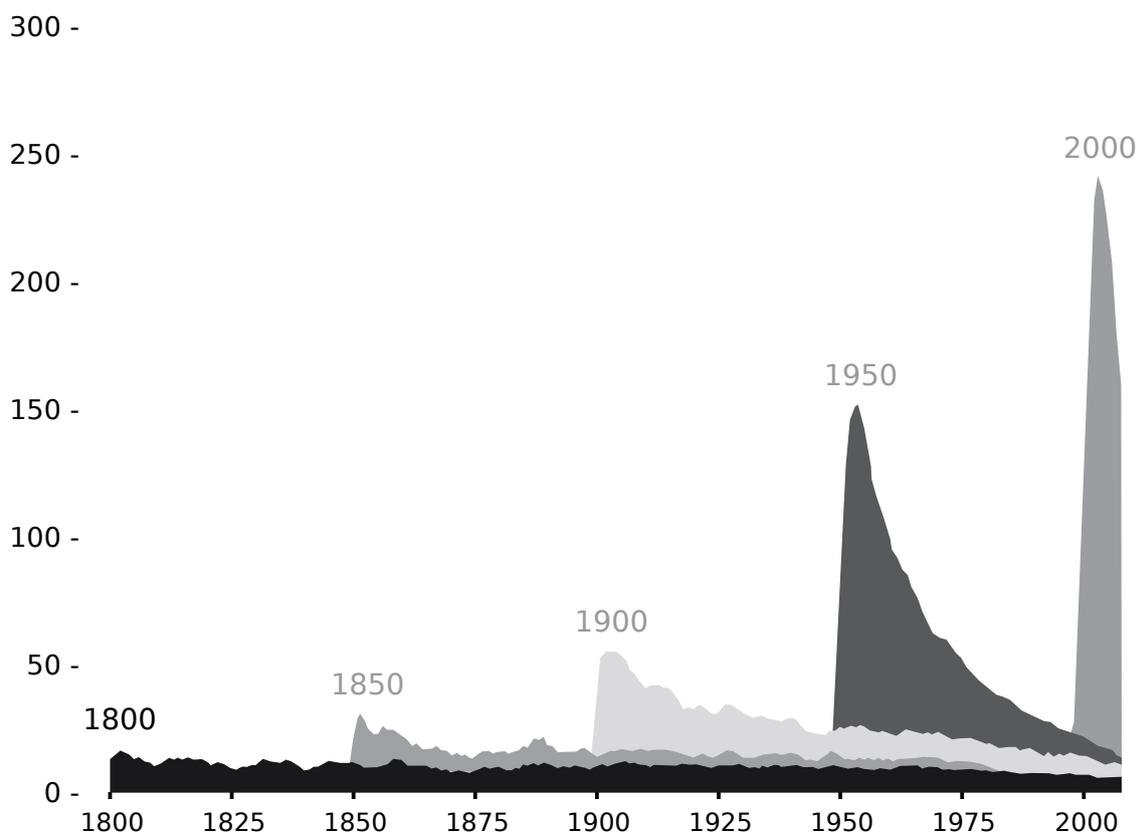
menciones por cada millón de palabras



De estos datos se pueden extraer también unos cuantos bocaditos de historia no culinaria. «Helado» despegó en la década de 1910, justo cuando General Electric presentó el congelador eléctrico doméstico. ¿Nos fijamos en el desplome que sufrió «pasta» a finales de la década de 1990? Fue cuando se popularizó la dieta Atkins. En las dos guerras mundiales nos gustaba la carne roja. Estas son aplicaciones triviales de una técnica con la que se puede llegar muy adentro en nuestra psique colectiva^[8]. Las frecuencias de aparición de palabras pueden incluso mostrarnos cómo percibimos abstracciones, como el paso del tiempo, algo que cuesta mucho investigar directamente. Preguntarle a alguien qué significan «diez años» es como pedirle que describa un color: obtienes impresiones cuando lo que buscas son hechos. Pero observar cómo cambia la escritura con el tiempo logra que nos hagamos una idea.

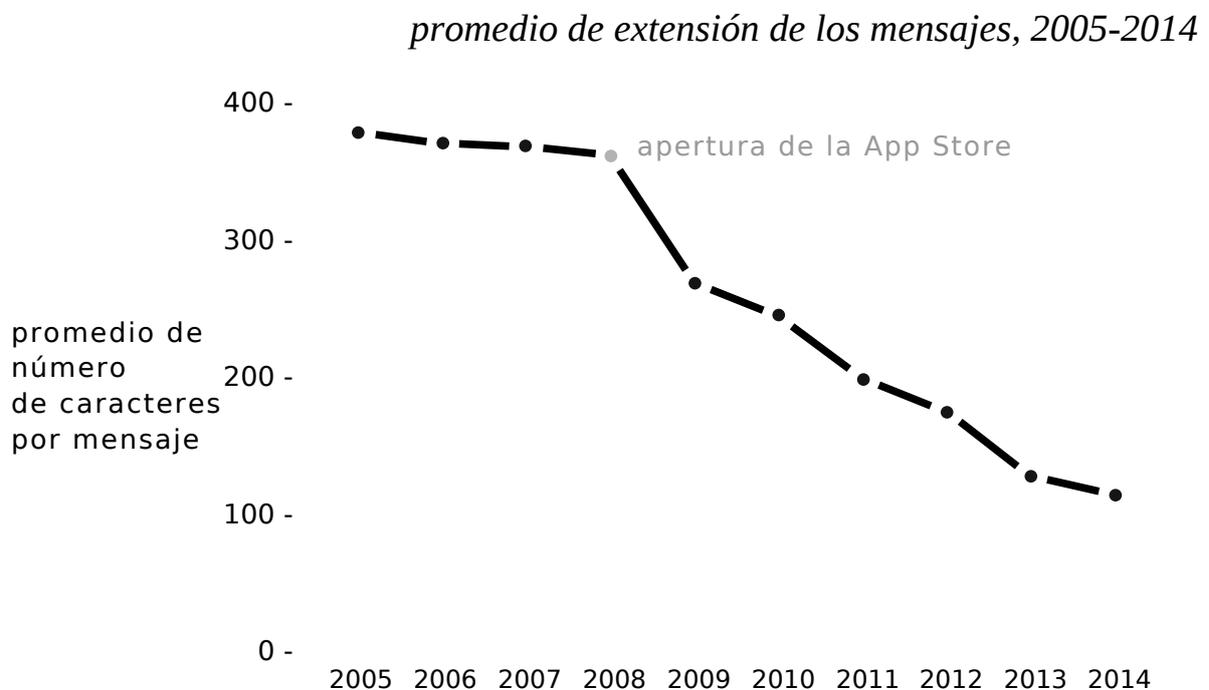
Los datos muestran que cada año que pasa nos centramos más en el presente. Por ejemplo, las menciones escritas del año 1850 alcanzaron (en 1851) un total de aproximadamente 35 por cada millón de palabras escritas. Las menciones del año 1900 llegaron hasta un total de 58 por millón. Las menciones de años recientes alcanzan picos de tres veces esa cantidad. Aquí están las trayectorias de las menciones a los años de cambio de siglo y de medio siglo según esos datos:

menciones por millón de palabras



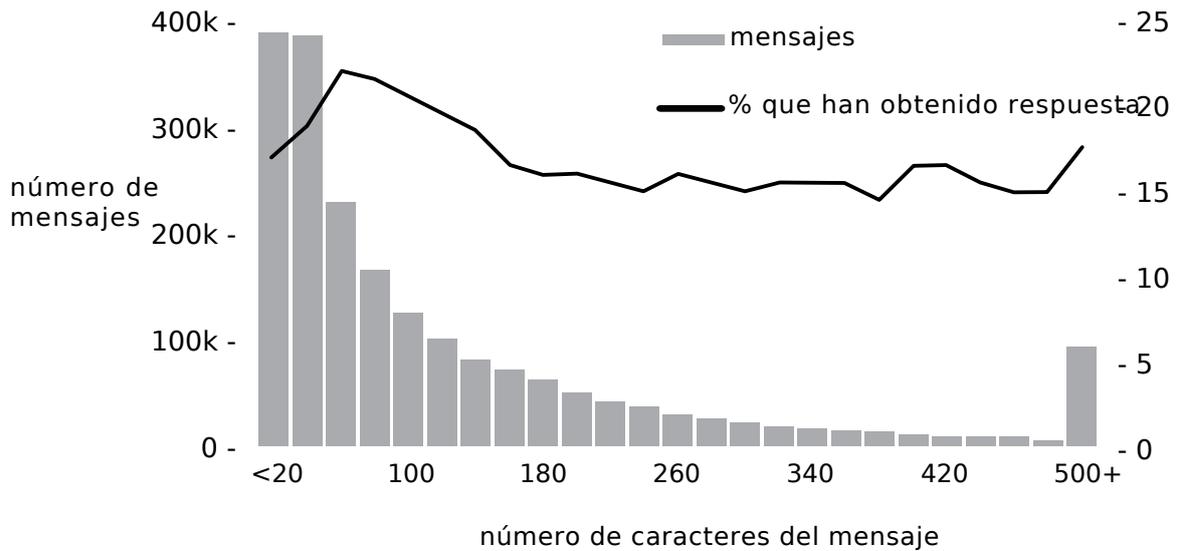
Este tipo de trabajos basados en la palabra impresa nos ayudan a entender nuestra cultura en sentido amplio. Twitter nos permite ver los grupos que se forman en su seno. Pero tanto libros como tuits son formas de comunicación de un solo emisor para muchos receptores, y muchas veces, como en el caso del comandante Ballou, las palabras más importantes que emitimos son de un solo emisor para un único receptor. Los usuarios de OkCupid intercambian alrededor de 4 millones de mensajes cada día. Naturalmente, lo hacen con un objetivo concreto —quedar—, pero la interfaz no les fuerza a nada específico ni les impone ningún límite sobre qué o cuánto deben teclear. Podemos verlo como una especie de Gmail para desconocidos: la comunicación en la web consiste en que dos personas se vayan conociendo; el amor viene mucho después, fuera de la Red. Los investigadores externos casi nunca trabajan con mensajes de este tipo: es el contenido más íntimo que generan los usuarios y los datos de mensajes rara vez salen de lo más profundo y sagrado de la base de datos, ni siquiera para tratarlos como conglomerado anónimo. Pero mi puesto privilegiado en OkCupid nos abre las puertas de manera especial.

Para empezar, la década de historia de la empresa nos permite ver cómo ha alterado la tecnología la manera de comunicarse de la gente. OkCupid conserva registros de la era previa a los *smartphones*, a Twitter y a Instagram; hombre, si ya estábamos *online* cuando Myspace era todavía un servicio de almacenamiento de archivos..., a juzgar por los mensajes de todos esos años, se puede afirmar que la cultura de la escritura, en el sentido más amplio, está cambiando, y ese cambio lo propician los teléfonos. Apple abrió su App Store a mediados de 2008, y OkCupid, como casi todos los grandes servicios de Internet, no tardó en lanzar una app. El efecto que tuvo en la manera de escribir fue inmediato. Los usuarios empezaron a escribir en teclados menores que la palma de la mano y, desde entonces, la extensión de los mensajes ha caído en unos dos tercios:



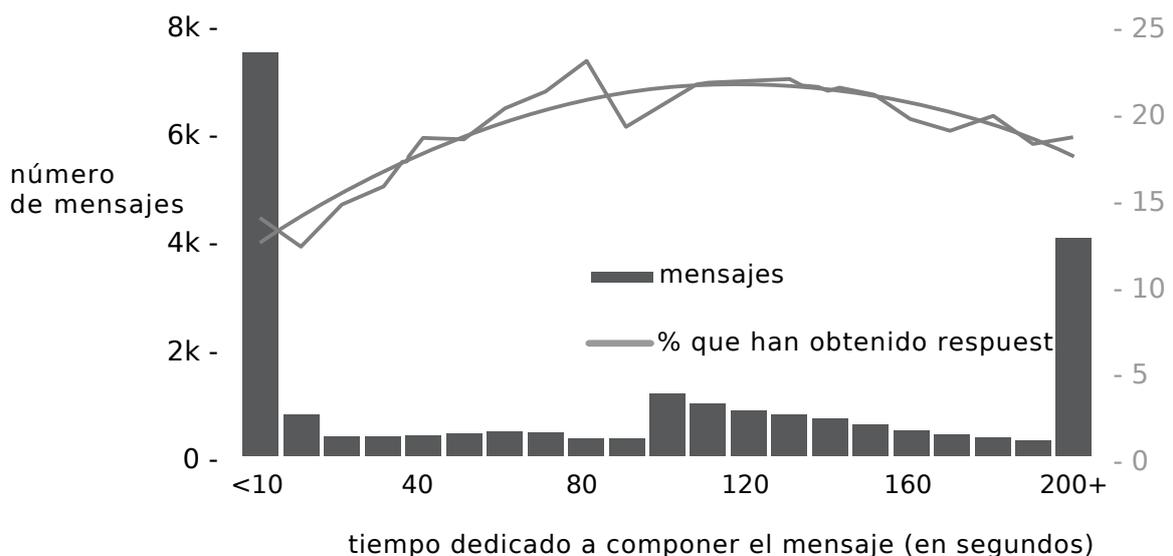
El mensaje típico de hoy supera por poco los 100 caracteres; de hecho, es la extensión de un tuit. Y en lo que se refiere a su efecto, parece ser que los lectores se han adaptado. Los mejores mensajes, los que consiguen un mayor índice de respuestas, tienen hoy una extensión de 40 a 60 caracteres.

extensión del mensaje frente a índice de respuesta obtenido



Si tenemos en cuenta solo los mensajes de determinada longitud y preguntamos cuántos segundos costó componer el mensaje, podemos hacernos una idea de cuánta revisión y cuánto esfuerzo se traducen en mejores resultados. A continuación vemos mensajes de entre 150 y 300 caracteres, enfrentados al tiempo que se tardó en escribirlos. Como se puede apreciar, tomarte tu tiempo ayuda, hasta cierto punto. Pero la curva descendente de las líneas de tendencia es como un fiel escudero numérico que te dice: «¡No te lo pienses tanto!».

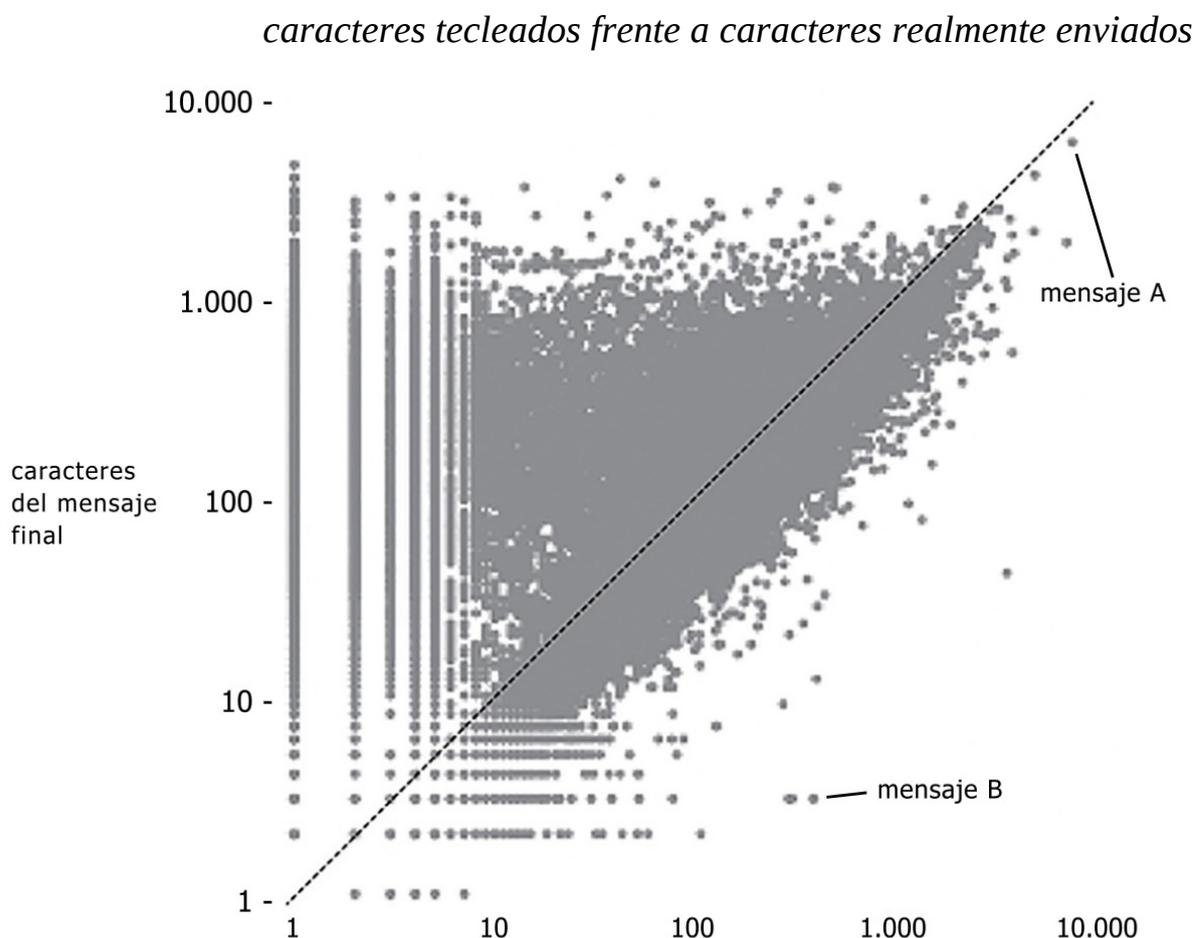
tiempo de composición frente a índice de respuesta obtenido, para mensajes de entre 150 y 300 caracteres



Bien, pues la primera barra vertical de la izquierda muestra los mensajes que se tardaron menos de diez segundos en escribir, supone una porción

desorbitada respecto al total y debería provocar alguna reacción de sorpresa. A mí ciertamente me la provocó, y eso que a estas alturas estoy más que insensibilizado a estas cosas después de pasarme diez años trabajando en una web de contactos. ¿Cómo puede ser que haya tanta gente que escribe mensajes tan largos a tal velocidad? La respuesta breve es: no la hay, y ahora explico cómo lo sé.

Abajo vemos un diagrama de dispersión de 100 000 mensajes, donde se muestran los caracteres tecleados frente a los caracteres realmente enviados^[9]. Puesto que contiene una amplia gama de recuentos, que van de uno hasta casi 10 000 caracteres, este gráfico es logarítmico:

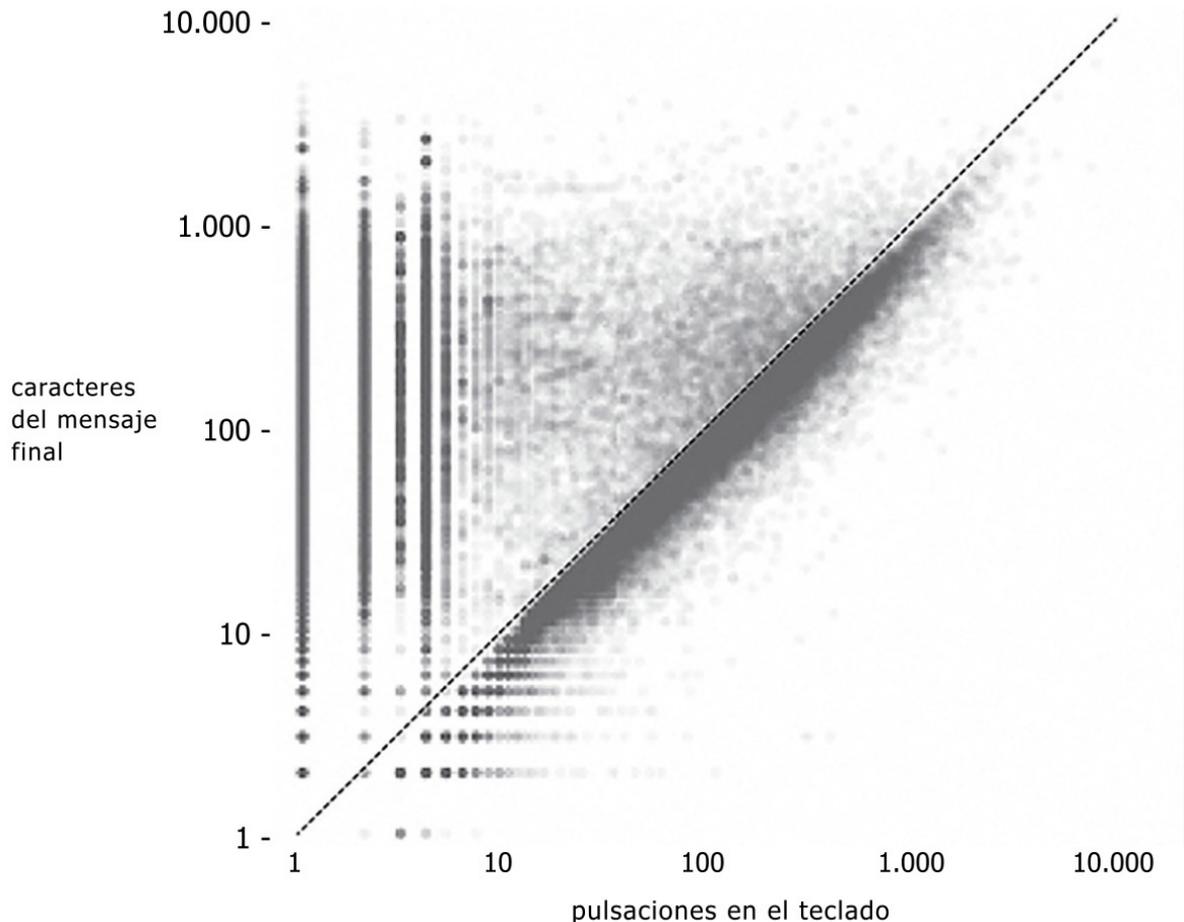


He añadido otra diagonal punteada que, como antes, señala el lugar en el que ambos ejes se igualan, es decir, que en esos puntos grises que la recorren el texto final coincide con los caracteres tecleados. Básicamente, el emisor tecleó lo que tenía en la cabeza y le dio a «Enviar», sin borrar ni editar nada. Por tanto, sabemos que el mensaje A que vemos en la esquina superior derecha se tecleó de manera más o menos apresurada, sin apenas revisarlo. Al

ir a mirar los registros, comprobé que el emisor tardó 73 minutos y 41 segundos en aporrear esos 5979 caracteres del mensaje de saludo; el mensaje final ocupa como cuatro páginas de este libro. No obtuvo respuesta. Ni tampoco el caballero que envió el mensaje B, que merece el premio Raymond Carver al esfuerzo más laborioso por ser breve. Necesitó pulsar 387 teclas para acabar escribiendo «Hey».

Pero estos son ejemplos muy extremos. La clave de este gráfico es que a medida que te acercas a la diagonal los mensajes presentan menos revisión. Si te vas hacia la esquina inferior derecha encuentras una edición más exhaustiva; si te mueves hacia la esquina superior derecha encuentras... imposibilidad física. La geometría de nuestro gráfico implica que en cuanto cruzas la diagonal para pasar a la parte superior, estás entre las personas que deben de haber tecleado menos caracteres que los que contienen sus mensajes. ¿Y quiénes son estos misteriosos hechiceros capaces de escribir palabras a partir de sus pensamientos? Son los que cortan y pegan, y son legión.

Nos aclararemos mejor con el gráfico si adjudicamos a los puntos una transparencia del 90 por ciento. Así podremos ver la verdadera densidad subyacente. Es como hacer una radiografía de los datos y, al hacerla, podemos ver hasta sus huesos:



La densa franja de puntos que recorre la diagonal por debajo corresponde a los que escriben sus mensajes desde cero. Es sorprendentemente compacta. Tenemos, claro, el límite definido que marca la diagonal por la parte de arriba, que separa a los que escriben mensajes de cosecha propia de los que pegan, como la línea de un frente de batalla entre dos facciones contendientes. Pero el borde inferior de la franja es casi igual de nítido. Parece existir un límite natural para el esfuerzo máximo que una persona es capaz de hacer a la hora de escribir un mensaje. Si echas las cuentas, te salen tres caracteres tecleados por cada carácter del mensaje final.

Por encima de la diagonal está la gente que decidió que ese esfuerzo era demasiado. Esa dispersión de puntos que ocupa el centro de la parte superior izquierda la componen todos aquellos que copiaron una plantilla de mensaje y se limitaron a aplicarle unos cuantos cambios. Aquí puede despistarte la naturaleza logarítmica del gráfico: hasta el más mínimo desvío por encima de esa diagonal significa que la mayor parte del contenido del mensaje es copiado. Si vas al extremo izquierdo del gráfico, verás unas líneas verticales densas, como surcos. Esos son los mensajes que se «escribieron» pulsando

unas pocas teclas. Son muchos: en total, un 20 por ciento de la muestra registró cinco pulsaciones del teclado o menos. Estos escritores han optado por poner algo que les gusta o que les funciona y listo. No se trata de *spam* en el sentido que le damos habitualmente al término; OkCupid detecta enseguida cualquier cuenta falsa o cualquier bot externo a la web. Son intentos de contacto de personas reales; básicamente, frases de ligoteo digital memorizadas. Muchas son tan facilonas y prosaicas como cabría esperar: «Hey, eres muy mona» o «¿Quieres hablar?», meros equivalentes digitales del «¿Vienes mucho por aquí?». Pero algunos de esos mensajes repetidos son tan idiosincrásicos que cuesta creer que se pudieran aplicar a más de una persona. Aquí tienes uno, reproducido tal como se escribió:

Yo también soy fumador. Me enganché cuando estuve de viaje en mayo. Antes lo hacía cuando bebía, pero ahora me levanto y, coño, me apetece un cigarro. A veces me gustaría trabajar en una oficina de *Mad Men*. ¿Has visto la exposición de Le Corbusier en el MoMA? Parece bastante interesante. La semana pasada fui a ver la de Frank Gehry (¿se escribe así?) en Montreal sobre cómo usó el modelado por ordenador para diseñar una casa rarísima en Ohio.

Ese es el texto íntegro del mensaje; su emisor intentaba contactar con mujeres fumadoras e interesadas por el arte. Lo de «(¿se escribe así?)» es mi floritura preferida. Un total de 42 mujeres recibieron este mismo mensaje.

En todo el sitio, la estrategia de copiar-pegar se ve superada en un 25 por ciento por los mensajes escritos desde cero, pero en términos de esfuerzo aplicado y resultados obtenidos siempre sale ganando: si lo medimos por las respuestas recibidas por unidad de esfuerzo, es varias veces más eficaz enviar más o menos lo mismo a todo el mundo que componer un mensaje nuevo cada vez. Se lo he dicho a la gente, lo de copiar y pegar, y la respuesta suele ser alguna variación de «Qué penoso...». Cuando les digo que estos textos cliché tienen una eficacia del 75 por ciento si los comparamos con la de un escrito original, se muestran escépticos —seguro que casi todos son capaces de apreciar lo que hay detrás de la fórmula—. Pero este último mensaje es un ejemplo de texto replicado a través del cual es imposible ver nada y el emisor, en una fracción del tiempo que le hubiese costado haber escrito algo original, obtuvo cinco respuestas del tipo exacto de mujer que estaba buscando. Y, permíteme que te diga, prácticamente todo lo que tengo sobre la mesa, lo que llevo puesto encima y, probablemente, todo lo que hay en mi casa, se hizo en

una fábrica junto con quién sabe cuántas copias más. Acabo de pelearme con un montón de gente para conseguir mi almuerzo, que consistía en un bocadillo escogido entre un montón enorme de otros bocadillos iguales. Las plantillas funcionan. Nuestro viajero fumador social y amante de la arquitectura no hace sino lo que la gente ha hecho siempre: aprovechar la tecnología. En este caso su innovación es haber pulsado unas cuantas teclas para ahorrarse un poco de tiempo.

Como ya hemos visto, los teléfonos y las redes sociales como Twitter exigen determinadas adaptaciones. Lo que no deja de ser eterno es que la escritura, como la vida misma, sigue adelante. Cambia de forma, se replica de maneras extrañas, encuentra nichos insospechados... a veces incluso, como cualquier otra cosa que esté viva, es una mierda. Pero ten por cierto esto: estamos viviendo una explosión cámbrica de la escritura, no su extinción masiva. El lenguaje es más variado que nunca, pese a que parte de él lo copiamos directamente desde el portapapeles, y la variedad es la preservación de un arte, no una amenaza. Desde el lenguaje de altos vuelos de la ficción literaria a la más simple y mal escrita actualización de un estado, todo eso que escribimos persigue un mismo objetivo. Ya se trate de un amigo a otro amigo, de un desconocido a otro, de un amante a otro o de un autor a su lector, empleamos las palabras para conectar. Siempre que haya alguien aburrido, emocionado, indignado, transportado, enamorado, curioso o que eche de menos su hogar y tema por su futuro, escribiré acerca de ello.

4.

Tú tienes que ser lo que nos
una

Uno de los principales inconvenientes de las webs de contactos es que apenas nos dicen nada de la gente que ha quedado para verse. Una vez que se han citado en persona ya no necesitan enviarse mensajes ni darse valoraciones ni nada de eso. Es una ironía que está tanto en los datos como en el propio trabajo: lo haces bien y los clientes se van. ¡Y se van por parejas, encima!

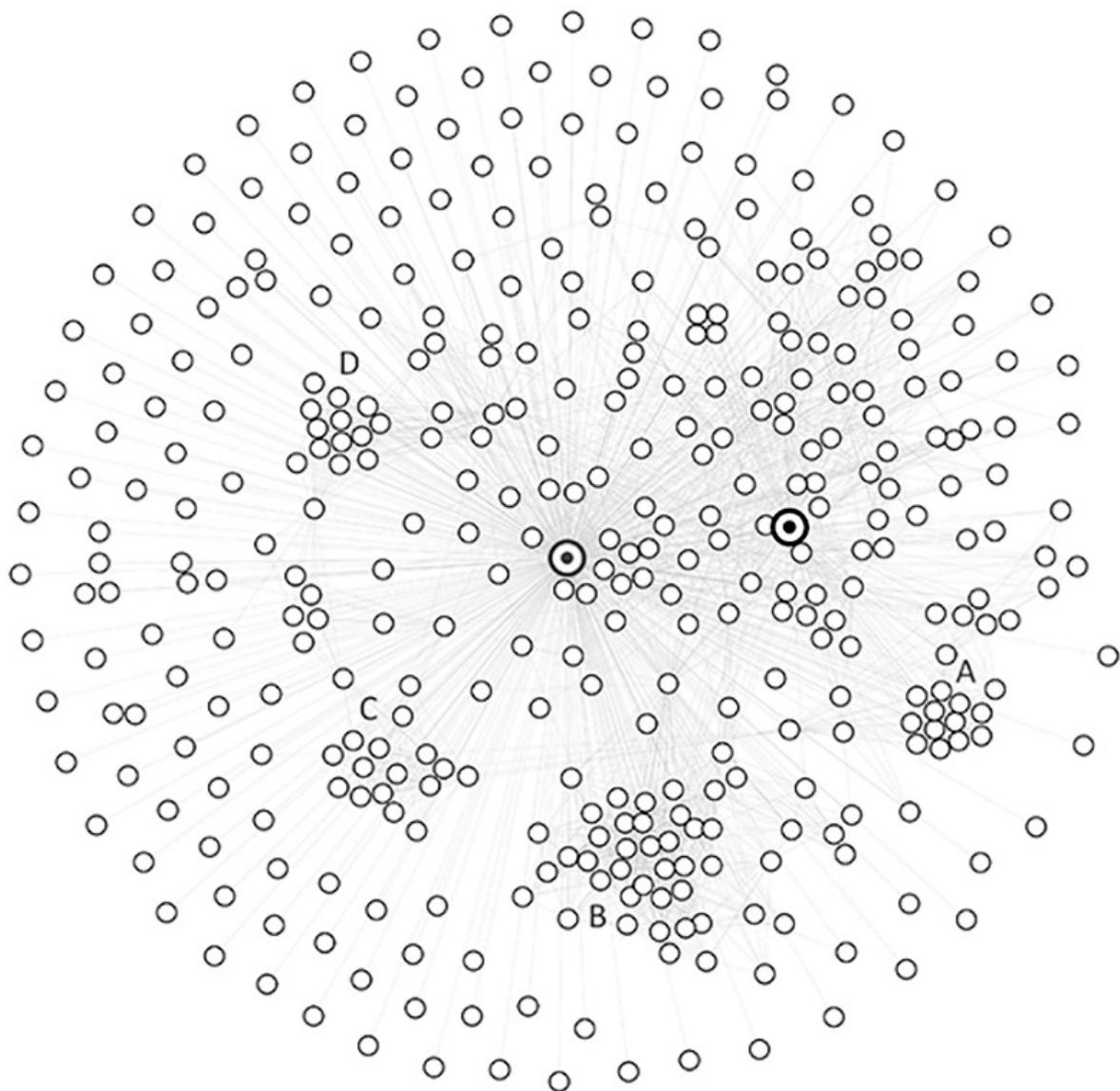
Ese sitio al que van, obviamente, es el mundo real: a un bar, a plena luz, a conocerse en carne y hueso. Abandonan el mundo fácilmente cuantificable de los bits y los píxeles y, en definitiva, se adentran mutuamente en sus vidas. Piensa en la progresión de una relación reciente: dos individuos se encuentran por primera vez en persona. Charlan, beben, empiezan a conocerse. Lo siguiente, si es que lo hay, es ir a casa de uno de ellos. El número poco familiar que hay sobre la puerta, un pomo de latón cuando el tuyo es de acero. El extraño pero agradable olor de las sábanas ajenas. Los champús que hay en la ducha, usados, pero nuevos para ti. Zumo de moras... Vale, ¿por qué no? La vez siguiente toca en tu casa: ella abre la nevera y solo hay... mostaza. Ay, lo siento. Todos hemos pasado alguna vez por el dormitorio de alguien, por su guarida, rodeados de *souvenirs* de acontecimientos y personas que no recordamos, sorprendidos al principio por todos esos chismes pero preguntándonos enseguida si podremos llegar a hacer nuestras cosas como el trofeo del tercer puesto en la Reunión de Natación Ponderosa de 1985, pese al hecho —o quizá debido a ello— de que solo las hayamos conocido a través de esa otra persona.

Conoces a sus amigos. A su mejor amiga. A su otra mejor amiga. A su segunda otra mejor amiga, que, en serio, se conocen de toda la vida. Con la cantidad de bebida suficiente y la gente adecuada, ya te has hecho nuevos amigos. Conocidos y compañeros de trabajo pasan también el filtro, algunos de pasada, otros a propósito. Por último, tal vez, si la cosa se pone realmente seria, llegan los padres. Cuentas una versión adornada de la historia de tu vida, partes de la cual ya incluyen a tu pareja porque ya estáis familiarizados el uno con el otro... Te alejas un momento de la mesa y cuando vuelves los padres saben más de ti que al marcharte. Te vuelves a acomodar en la silla:

«M me comenta que...», y se da la circunstancia perfecta para que cuentes una de tus anécdotas preferidas. Dos vidas se están fundiendo. Y entonces muchas veces, y muchas veces de repente, todo vuelve a empezar con alguien distinto.

Hasta ahora hemos echado un vistazo a las maneras que tienen dos personas de juntarse por un flechazo. No estoy muy seguro de que un ordenador pueda llegar a ser capaz de registrar su trayecto hasta la unión plena, pero sí que nos da una imagen de su vida una vez que lo han logrado. Ese patrón de una pareja, esa urdimbre de lo que se ha dado en llamar sus «gráficos sociales» está ahora bien documentada.

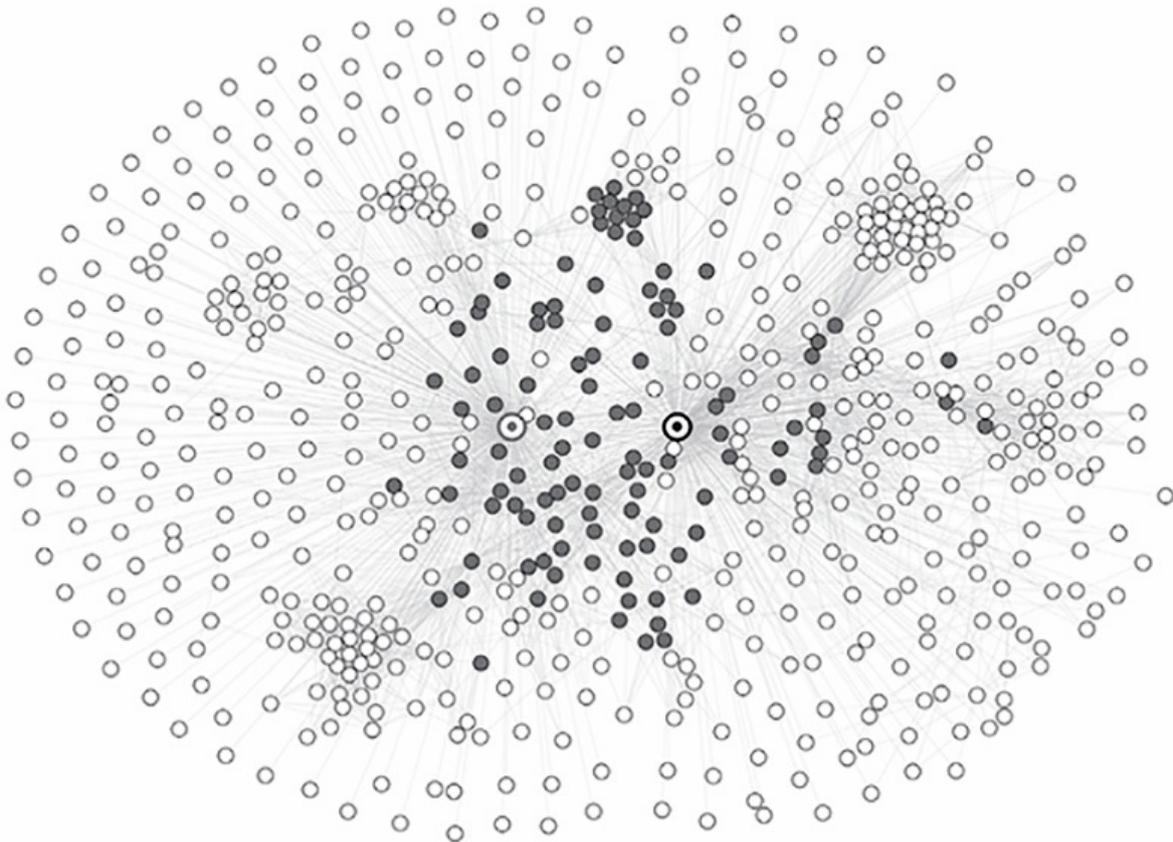
Tengo 384 amigos en Facebook; aquí están. Yo soy el punto gris del centro; mi mujer, Reshma, es el punto negro que está hacia las tres en punto. Todas las demás conexiones entre todas las personas se han representado mediante líneas grises:



Aunque mis amigos aparecen bien agrupados, este gráfico no se distribuyó a mano, sino que mi ayudante de investigación, James Dowdell, escribió un *software* especial para generarlo. Los puntos se agrupan en función de la cantidad de conexiones que comparten. Es como si fuesen ralladuras de hierro imantadas por el poder de la amistad que se han esparcido sobre una mesa para que se asienten. Aunque apenas uso Facebook salvo para la tarea sumamente circular y cerrada de aceptar peticiones de amistad, aquí pueden verse todas las facetas de mi vida. Mi denso conjunto de familiares políticos, que se solapan todo lo que permite el *software*, es el grupo A; la gente con la que fui al instituto son el B; mis compañeros de trabajo son el C; mis amigos jugones, el D. Hasta puede verse mi pasada y prometedor carrera como músico en el gráfico. Me pasé años yendo de gira con un grupo y esos puntos solitarios que recorren el perímetro izquierdo son sobre todo personas

que conocí entonces. Sus vínculos mutuos son nuestra música, que resulta invisible para los algoritmos.

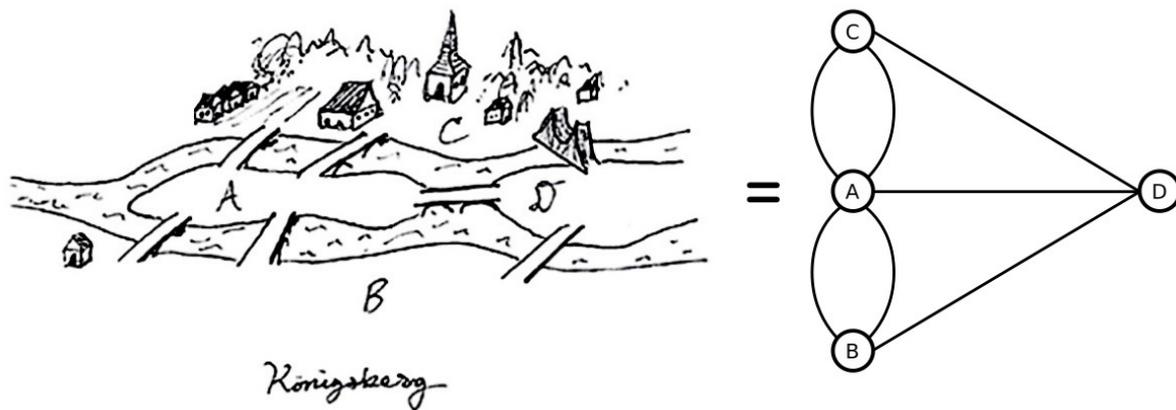
A continuación amplió el gráfico para incorporar también los contactos de Reshma y poder apreciar el alcance de nuestra red como pareja. Las conexiones que compartimos, nuestros amigos comunes, se han marcado en gris oscuro.



Puede que esto parezca una abstracción insulsa de la vida en común de una pareja, pero este tipo de diagrama conjunto nos dice muchísimo sobre el vínculo existente entre las dos personas en las que se basa. Fijándonos solamente en el gráfico, en la imagen, podemos concluir que Reshma y yo tenemos menos probabilidades de romper que otras parejas.

El análisis de redes, el estudio de puntos y líneas como los que vemos arriba, es toda una ciencia desde hace al menos cien años, y algo del actual auge de los datos (el paso de un mero goteo a un verdadero cataclismo) lo podemos ver en su evolución. El primer problema de redes fue una suerte de acertijo rudimentario, en realidad una leyenda urbana de la era de la Ilustración, consistente en que era imposible recorrer la ciudad prusiana de Königsberg atravesando todos sus puentes una sola vez cada uno. En 1735

apareció Leonhard Euler y, como es propio de los genios, redujo a una abstracción de puntos y líneas (formalmente, nodos y extremos) lo que había sido un problema coloquial sobre barrios y rutas a pie, y de ese modo demostró con rigor que la leyenda era cierta. Representó la ciudad en forma de red y así nació una nueva disciplina científica.



La reflexión que se hizo Euler fue que si solo debías cruzar cada puente una vez, para pasar por cada barrio necesitarías una pareja de puentes: uno para entrar y otro para salir. De modo que la solución era tan sencilla como mirar el diagrama y preguntarte si a todos los puntos de tu recorrido, aparte de los de partida y destino, les llegaban un número par de líneas (un par de puentes). En el caso de Königsberg, ninguno las tenía, así que el problema quedaba resuelto. Que de unos orígenes tan prosaicos pueda surgir una ciencia tan duradera y floreciente, que solo ahora empieza a llegar a su máxima expresión, me parece el mejor ejemplo posible sobre cómo es el espíritu humano^[10]. El concepto euleriano de nodos y extremos, que en origen sirvió solo para desenmarañar una caminata, nos ha ayudado desde entonces a comprender las enfermedades y su propagación, los camiones y sus rutas, los genes y sus vínculos y, por supuesto, a las personas y sus relaciones. Y en las décadas más recientes, esta última aplicación de la teoría de redes ha experimentado una explosión, porque las propias redes han explotado.

Hace cuarenta años, Stanley Milgram se dedicó a enviar paquetes por correo (con una serie de instrucciones y un sobre prefranqueado) a un centenar de personas de Omaha, en su experimento de los «seis grados de separación», con la esperanza de que tal vez una docena de almas aventureras se animasen a participar. Sus peculiares métodos —por muy ingeniosos que fuesen— le granjearon su teoría, pero sin llegar a demostrarla del todo. En

2011, el alcance abrumador y sin precedentes de Facebook nos permitió ver que estaba en lo cierto: el 99,6 por ciento de los 721 millones de cuentas que existían entonces estaban conectadas por seis pasos o menos.

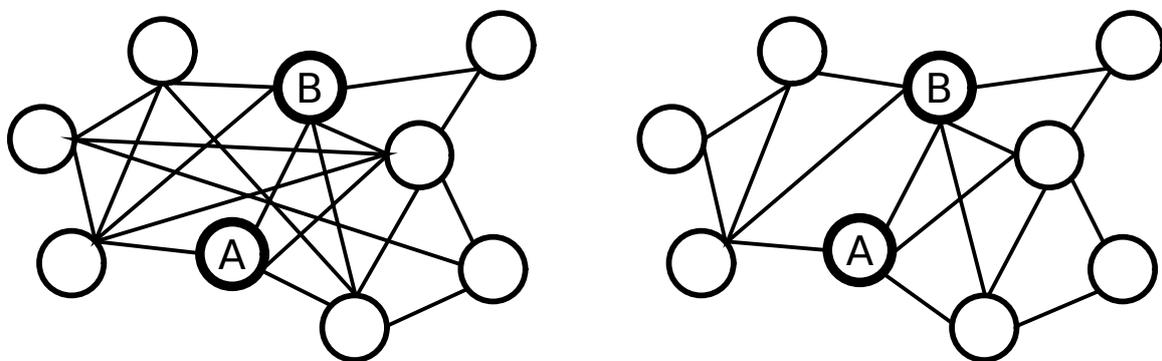
Hoy en día, la teoría de redes, que trabaja con grandes paquetes de datos gracias a la tecnología, nos muestra que la gente puede encontrar trabajo, filtrar información a partir de tonterías e incluso hacer mejores películas. Es sabido que los de Pixar, cuando edificaron su sede, pusieron los únicos cuartos de baño del edificio dentro del atrio central, para obligar a los trabajadores de los distintos departamentos a que charlasen entre ellos, sabedores de que la innovación muchas veces surge de la colisión fortuita de ideas. Aquella fue una puesta en práctica de «la fuerza de los vínculos débiles», un concepto que se acuñó en la década de 1970 mediante decenas de muestras y que desde entonces se ha amplificado a base de nuevos y sólidos datos procedentes de las redes: nos descubre que la gente de nuestra vida que no conocemos demasiado bien es la que contribuye a difundir las ideas, sobre todo las novedosas^[11].

Otra idea que se asocia desde hace tiempo con la teoría de redes es la de «integración». Una de sus expresiones es la cantidad de solapamiento que se da en dos gráficos sociales: la integración entre Reshma y yo consiste sencillamente en comparar el tamaño de la parte gris de nuestro gráfico con todo el conjunto. La investigación a través de fuentes diversas (correo electrónico, mensajería instantánea, teléfono) ha demostrado que cuantos más amigos comunes comparten dos personas, más intensa es su relación. Más contactos implican más tiempo compartido, más intereses comunes y más estabilidad. Pero a diferencia, digamos, de los registros telefónicos o incluso del correo electrónico, las redes sociales de Internet aportan muchos datos succulentos a los extremos y nodos de un gráfico (de forma parecida, las webs de contactos han tomado el sempiterno ritual del cortejo y le han añadido la edad y la belleza como variables de estudio) y, naturalmente, de esas redes Facebook es la más rica que jamás se ha creado. Los efectos de esa riqueza de datos ahora empiezan a notarse.

El análisis de gráficos sociales empezó como un asunto de «quién conoce a quién», y en gran parte así sigue siendo. El alcance de los datos de Facebook —puedes profundizar varios grados sin apenas esfuerzo adicional— está empezando a poner esa idea patas arriba. En lo que respecta a las relaciones, y concretamente a las relaciones amorosas, estos datos nos han brindado recientemente una nueva y potente manera de medir la fortaleza de un vínculo entre dos personas. Resulta que sus vidas no solo deberían

entrelazarse, sino que deberían hacerlo de determinada manera. Y, lo que resulta raro de ver como medida del análisis de redes, la cantidad relevante es «quién no conoce a quién».

Dos científicos, Lars Backstrom y Jon Kleinberg, trabajaron con 1,3 millones de parejas de Facebook y plasmaron esta idea en un artículo científico en 2013. Su medición se basó en contar la cantidad de veces que una persona y su cónyuge actuaban de puente entre partes inconexas de su red de pareja. Me refiero a esto: el gráfico de la izquierda representa un escenario guay del Paraguay, en el que más o menos todo el mundo se conoce; presenta una integración muy elevada. Pero el matrimonio más sólido es el del diagrama de la derecha, donde los miembros de la pareja, A y B, son los únicos conectores de los dos colectivos que los rodean:



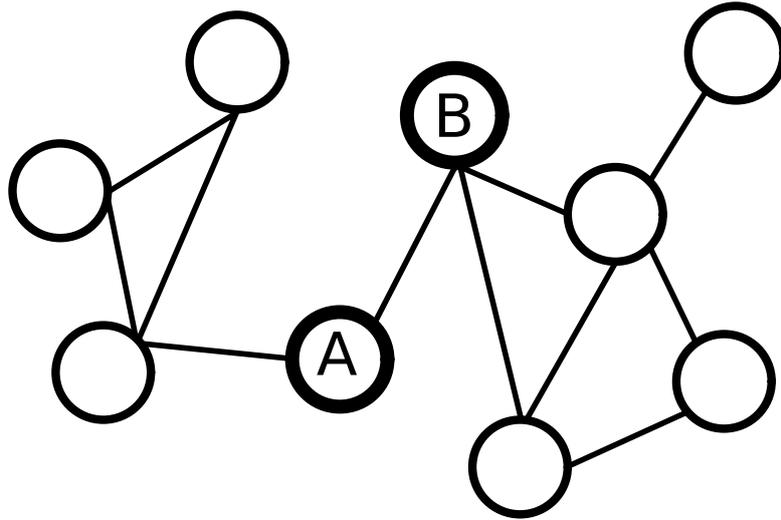
Esto puede parecer un poco raro: ¿por qué querría alguien que su red estuviese más desconectada salvo para uno mismo y su pareja? Pero, como ocurre con las mejores ideas, en la vida real esto se produce de manera intuitiva. Sheel, el primo de Reshma, está muy integrado en su vida. Se criaron juntos y tanto él como ella tienen contacto prácticamente con todos los miembros de su extensa familia, incluida mucha gente que yo ni siquiera conozco. Se conocen de toda la vida, mientras que Reshma y yo llevamos casados solo siete años. Si colocásemos la relación de Sheel y Reshma como eje central, el diagrama sería muy parecido al que aparece a la izquierda. Sin embargo, Sheel no conoce a los colegas de trabajo de Reshma. Tampoco conoce a los compañeros de danza de Reshma. No conoce a los amigos de la facultad de Reshma. Yo sí los conozco a todos y, lo que es más, soy la única persona en la vida de Reshma que esos tres grupos independientes tienen en común, al menos de manera directa. Para esos colectivos, nosotros encarnamos el ideal del diagrama de la derecha. Cabe señalar que, por ejemplo, si Reshma y yo trabajásemos juntos, o si ella no fuese a danza, o si

hubiésemos asistido a la misma facultad, no podríamos desempeñar el papel que desempeñamos en nuestras mutuas redes.

Backstrom y Kleinberg denominan el nivel en el que una relación cumple ese ideal su «dispersión», porque muestra lo desconectado que estaría tu diagrama sin ti; es decir, que tu círculo social sería completamente arrastrado por el viento si tu pareja y tú fueseis apartados de algún modo del centro (por el nacimiento de un segundo hijo, por ejemplo). Prefiero usar el término «asimilación», porque creo que capta mejor el resultado: la gente asimilada cumple un papel decisivo como pareja en sus redes mutuas. Las parejas muy asimiladas funcionan —los dos juntos— como el vínculo que une a muchos colectivos que en su ausencia estarían desconectados. Son como un aglutinante especial de determinada distribución de puntos y, más aún, son como una especie de resina epoxi: se necesita mezclar los dos ingredientes para que peguen bien.

El poder de asimilación surge del hecho de que tu pareja es una de las pocas personas (si no la única) a quien presentas a los individuos más alejados del centro de tu vida. Está contigo en las fiestas del trabajo, en las reuniones y en esos días en los que tus amigos jugones se te instalan en casa para daros el atracón de videojuegos que lleváis todo el año esperando (bueno, puede que ese día ella no esté, si puede evitarlo, pero seguro que captas la idea). Sin embargo, esos colegas de trabajo, esos compañeros de clase y esos amigos jugones pese a ser grupos muy densamente interconectados, no guardan ninguna relación entre sí salvo por ti y tu pareja.

Y he aquí la importancia del asunto: para los casados que están en Facebook, su cónyuge es el miembro más asimilado de su red en un alucinante 75 por ciento del tiempo. Y, lo que es más importante aún para considerar la asimilación como unidad de medida de la fortaleza de una relación, las parejas jóvenes para las que no es ese el caso tienen un 50 por ciento más de probabilidades de romper. En las relaciones más estables, las dos personas desempeñan un papel específico en la vida el otro. Si observamos los gráficos de una pareja no asimilada sabremos con certeza por qué: en una pareja excesivamente integrada, como el ejemplo anterior de la izquierda, tú y tu pareja acabáis compitiendo por el tiempo y la atención con todos los demás. Queda todo demasiado nivelado, hay muy poco que sea especial. Demasiadas noches «solo para chicas». Y en una red de colectivos carentes de asimilación, «llevar vidas separadas» puede convertirse en poco tiempo en «llevar vidas secretas», lo que podría tener más o menos este aspecto:



Además de mediante la asimilación, Backstrom y Kleinberg probaron muchos otros modos de evaluar una relación, y encontré un detalle en su artículo, presentado casi como un aparte, que me pareció especialmente irónico. En los inicios de una relación, el mejor indicador no depende en absoluto del gráfico social de la pareja; durante el primer año o así de la relación, el mejor método consiste en ver con qué frecuencia consultan mutuamente sus perfiles. Solo cuando ha pasado cierto tiempo, cuando empiezan a disminuir las visitas a sus páginas y su red mutua se va engrosando, empieza la asimilación a adquirir relevancia en los cálculos. Dicho de otro modo, la curiosidad, el descubrimiento y la estimulación (visual) del enamoramiento acaban siendo reemplazados por lo que en la teoría de redes sería el equivalente a la anidación.



En informática existe la idea de que uno debe ser su propio cliente: que debes tener por lo menos la suficiente confianza en el sitio web o en el *software* que le estás vendiendo al resto del mundo para usarlo tú mismo. Igual que cuando Jonas Salk se inyectó su recién descubierta vacuna contra la polio, tienes que demostrar que lo que estás haciendo es bueno. Los programadores lo llaman *dogfooding*, «comerse la comida del perro», en un guiño a su incapacidad como colectivo de tomar buenas decisiones en lo que respecta a la alimentación. En algunas empresas es imperativo comerse la comida del perro. Vete a una reunión con los de Microsoft y verás que todos sacan sus teléfonos con Windows y sus tabletas con Surface, como perros diligentes que mastican sus trocitos de huesos y tendón.

Tú y yo no tenemos que acatar ese tipo de órdenes, por supuesto. Pero he basado deliberadamente este capítulo en datos de mi vida porque, para empezar, necesitaba trabajar sobre conceptos abstractos basándome en ejemplos claros, pero también porque he querido mostrar que, en un libro que recoge los datos sumamente personales de mucha gente, no tengo ningún inconveniente en aplicarme ese mismo análisis a mí mismo.

Te doy la misma oportunidad. Para que puedas poner a prueba tu matrimonio, tu círculo social o tu amistad malsana y codependiente mediante los principios tratados en este capítulo, pongo a tu disposición el algoritmo Backstrom/Kleinberg aquí:

dataclysm.org/relationshipstest

Si le proporcionas un par de credenciales de Facebook, no solo te mostrará el gráfico mutuo y su correspondiente integración, sino que también establecerá una jerarquía de tus relaciones más asimiladas. El mundo ha llegado por fin a un punto en el que podemos hacer algo con nuestros datos propios: no tenemos que esperar a que un Milgram, por no decir un Euler, vengan a enseñarnos nada sobre nosotros mismos. Del mismo modo que redes como Facebook o Twitter dejan nuestros datos expuestos al escrutinio académico, también nos los devuelven para que podamos analizarlos nosotros. Ya tenemos disponibles algunas herramientas ligeras para capturar y analizar nuestra actividad física, y, en breve, dispondremos de otras más potentes. Cuando ves a cargos intermedios de alguna empresa jugueteando con sus dispositivos Fitbit en el ascensor, te das cuenta de que el Movimiento Autocuantificador ha llegado para quedarse. El ejemplo de arriba es mi pequeña aportación a las posibilidades que todo esto nos ofrece.

Si usas mi app para Facebook con alguna otra persona, espero que ambos estéis en lo más alto de vuestras correspondientes listas. Y, recuerda: un poco de eliminación creativa de amigos puede darle el impulso necesario a tu puntuación de asimilación. Lo digo porque esto de medirse a uno mismo está muy bien hasta que en la lista aparece alguna exnovia por encima de tu mujer.

5.

No hay mejor éxito que el
fracaso

Hay un Tumblr genial que se llama *Clients from Hell* (clientes del infierno), donde cualquiera puede colgar sus propias historias de terror relacionadas con su sector de trabajo. Allí se recogen despistes y descuidos de toda índole y cada pocas horas se publican nuevas entradas. Esta es una de las publicaciones típicas, de alguien que estaba maquetando unas fotos:

Cliente: ¿Podemos ponerle también un titular a la foto?

Diseñador: Hombre, ya tiene su pie de foto.

Cliente: Si el lector no ve el pie, siempre le quedará el titular.

Diseñador: Sería un poco extraño ponerle a una foto un titular y también un pie de foto...

Cliente: Lo que dices tiene sentido. Pues ponle un titular al pie de foto.

Mi cita favorita de cliente de las que hay en la web es: «No me gusta este dinosaurio del dibujo. Se ve muy falso. Mejor pon una foto real de un dinosaurio». La mayoría de las contribuciones al blog son de diseñadores gráficos, pero la popularidad de *Clients from Hell* evidencia que se trata de verdades universales. La gente odia a sus clientes.

No me refiero a que los odien a nivel personal, pero los clientes, en masa, como cualquier muchedumbre, son gente de temer. Quien te diga lo contrario, ya sea el dueño de la tienda de *cupcakes* de la esquina o el CEO que está en un consejo de administración, te miente. En parte se debe a eso de «... siempre tiene la razón»: a nadie le gusta una persona dotada de todo ese poder. Pero, con mucho, la mayor causa de frustración radica en que la gente no entiende ni es capaz de expresar lo que de verdad necesita. Como dijo Steve Jobs: «La gente no sabe lo que quiere hasta que se lo enseñas». Lo que no dijo es que enseñárselo, sobre todo cuando se trata de tecnología, significa jugar a ponerle la cola al burro mientras varios millones de personas te dan consejos a gritos.

Pongamos que eres un fabricante de coches y que a la gente no le gusta cierta parte de tu producto, la mayoría te lo dirá de manera indirecta dejando de comprártelo. Históricamente, nunca ha habido un canal abierto entre Ford y la gente que quiere que los portabebidas del coche sean de color verde o la que piensa que sería mejor que el volante fuese cuadrado porque, oye, la mayoría de los giros son de noventa grados. Por eso las empresas tradicionales gastan tanto en estudios de mercado: tienen que estar siempre muy por delante de esta clase de cosas, porque para cuando una empresa como Ford se percata de determinado problema, normalmente a través del departamento de contabilidad, suele ser demasiado tarde.

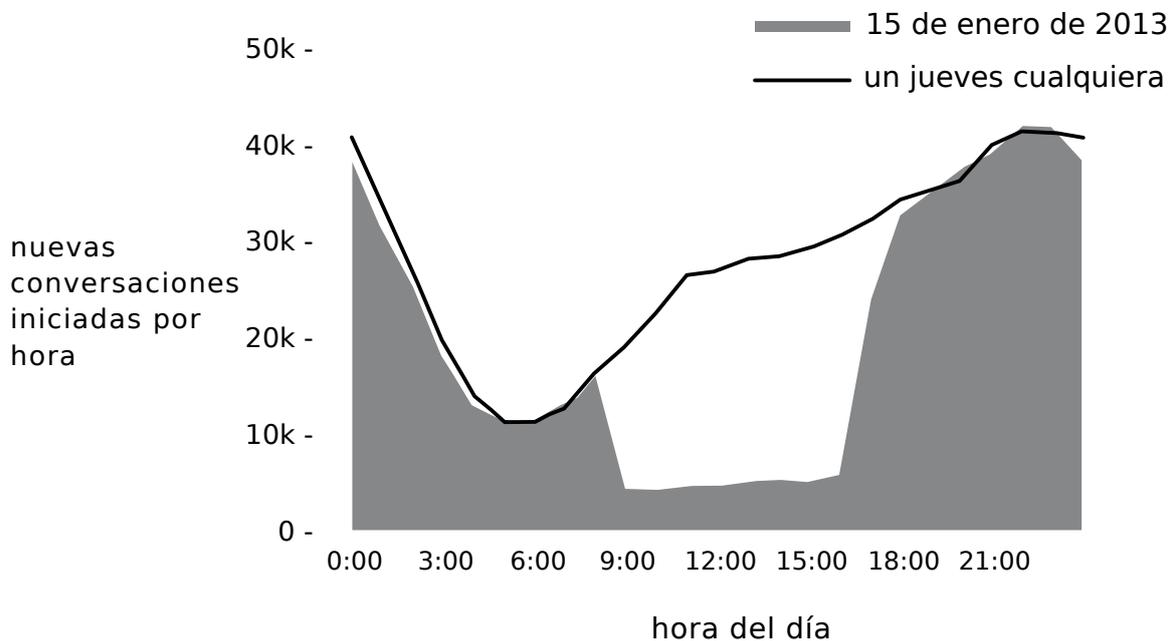
En el mundo de la web es diferente: si a uno se le ocurre una idea peregrina, tiene a alguien de la empresa a tiro de *e-mail*. Y si la gente no usa algo, la web se percata de ello de inmediato. Las mediciones tienen un seguimiento en tiempo real, hasta el mínimo detalle y el último rincón. Que sepas que cada vez que ves algo en tu web favorita —Google, Facebook, LinkedIn, YouTube o la que sea— y haces clic allí, alguien, que probablemente lleve cascos y coma Doritos, ve cómo en un pequeño contador se añade un número. Eso es lo que pasa cuando la riqueza de los datos puede hacer que alguien enloquezca: uno de los mejores diseñadores de Google, Douglas Bowman, que de hecho fue el encargado de montar su equipo de diseño, acabó por dimitir porque el proceso se había vuelto demasiado microscópico. La empresa no lograba decidirse entre dos tonos de azul para un botón, así que publicaron 41 tonos intermedios para ver cuál funcionaba mejor. «Conócete a ti mismo» es lo que estaba grabado en un sillar del templo de Apolo en Delfos. Pero como toda la mejor sabiduría que la historia nos puede ofrecer, se queda en nada en cuanto alguien enciende un ordenador.

No saber qué necesitan los consumidores de un coche, o incluso de una interfaz web concreta, eso es materia para una escuela de negocios o un taller de diseño. A mí lo que me interesa es qué pasa cuando las personas no entienden su propio corazón. Encontrar gente que dice una cosa y hace otra es de lo más común en las redes sociales, pero yo he tenido la excepcional oportunidad de ver a gente que actúa de dos maneras contradictorias. Y eso ocurrió porque tampoco yo sabía qué era lo que querían.



El 15 de enero de 2013 OkCupid proclamó que era el «Día del amor ciego» y suprimió todas las fotos de los perfiles del sitio durante unas horas. La idea era hacer algo distinto y, a la vez, llamar un poco la atención sobre un nuevo

servicio que estábamos lanzando. Los programadores «apagaron el interruptor» a las nueve de la mañana:



Aquello fue un sincero ataque de desesperación. ¡Lo nunca visto! El nuevo servicio que OkCupid trataba de promocionar era una aplicación para móviles llamada Crazy Blind Date (cita a ciegas loca). Con un par de toques en la pantalla te emparejaba con una persona y escogía un punto cercano y una hora próxima para encontrarse. La interfaz de la app permitía que los dos confirmasen la cita, pero no daba opción alguna a que nadie se pudiese comunicar directamente antes de la cita. La única información que te proporcionaba sobre la otra persona era el nombre de pila y un rompecabezas en miniatura de su foto, como la que se ve abajo. Se trataba de que aparecieses allí y cruzases los dedos a ver qué te encontrabas.

Seguramente te habrás dado cuenta de que estoy hablando de Crazy Blind Date en pasado. Pese a llegar al cuarto de millón de descargas, fracasó porque al final la gente insiste en ver en qué se está metiendo. La app fue una de esas ideas que parecen geniales en la pizarra y penosas cuando las ves a todo color en la realidad. Fue como un «Día del amor ciego» muy largo y sin posibilidad de volver a encender al interruptor. A los pocos meses del lanzamiento cerramos el servicio, pero antes de que Crazy Blind Date se fuese definitivamente al cielo de las apps (un dato poco conocido: no hay *bugs* en el cielo, solo *software* molón), unas 10 000 personas habían compartido una cerveza o un café con alguien a quien nunca habían visto y con quien nunca habían hablado.



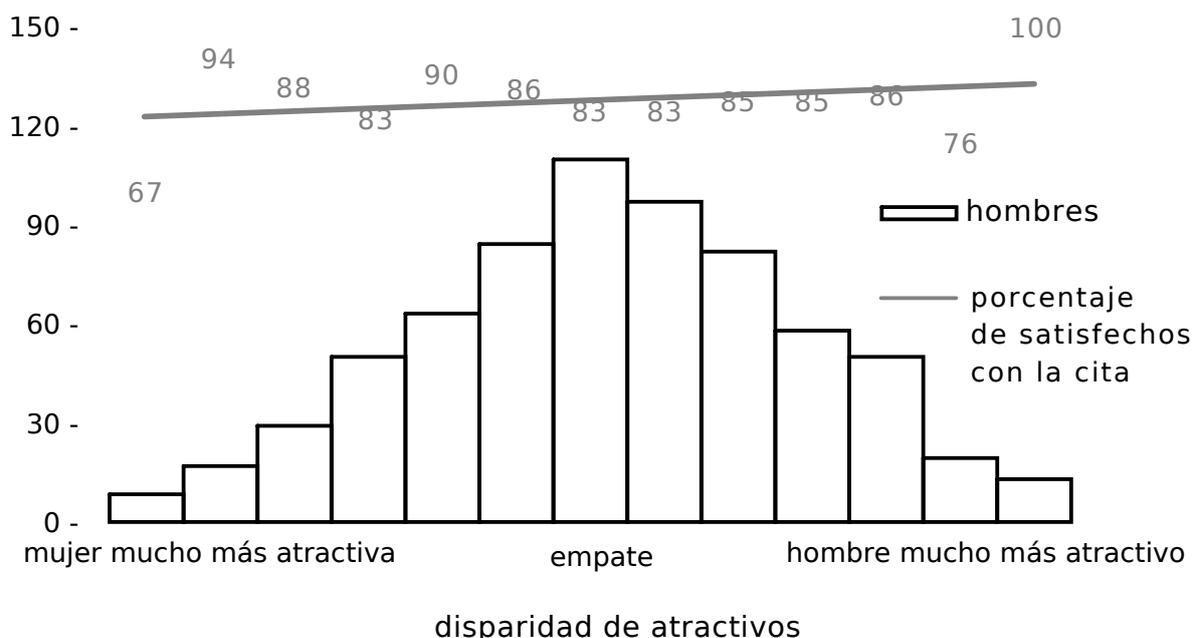
Una foto de banco de imágenes compuesta al modo de los rompecabezas de CBD.

A partir de esos pocos intrépidos, la aplicación nos dejó una curiosa serie de datos para la posteridad. Crazy Blind Date no solo registró que el ligón A y la ligona B se habían reunido en persona, sino también sus opiniones mutuas. Después de cada una de las citas, como un compañero de habitación cotilla, la app te preguntaba qué tal había ido. Dado que la mayoría de sus usuarios también tenían cuenta en OkCupid, pudimos cruzar estos datos con todo tipo de información demográfica. De repente podíamos combinar registros de encuentros presenciales con nuestra ingente colección de interacciones digitales. Cuando combinamos esas dos fuentes, dimos con algo reseñable: el aspecto físico de las dos personas apenas afectó al hecho de que lo hubiesen

pasado bien o no. Independientemente de cuál de las dos personas fuese más guapa, o cuánto más guapa que la otra —incluso en aquellos casos en los que un participante en la cita a ciegas ganase al otro por KO absoluto—, el porcentaje de personas que dieron valoraciones positivas fue bastante unánime. El atractivo no importaba. Estos datos, obtenidos de citas reales, dieron al traste con todo lo que había visto en los diez años que llevaba dirigiendo la web de contactos.

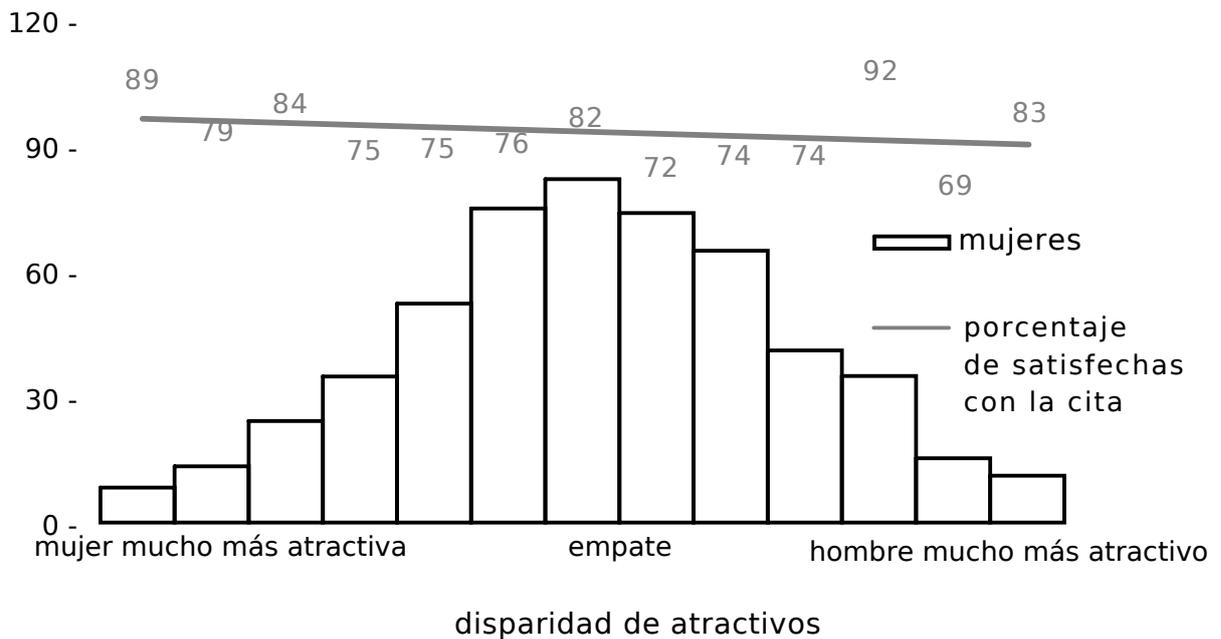
Estos son los números correspondientes a los hombres. En el gráfico he expresado el atractivo como la diferencia relativa entre las valoraciones individuales de cada pareja, en lugar de usar valores absolutos. Lo he hecho para registrar el hecho de que la felicidad de una persona que al otro lado de la mesa tiene a alguien con, digamos, un 6 depende en gran medida de su propia apariencia física. Si esa persona tiene un 1, podría encantarle la cita: eso significaría que está dando un braguetazo. Si tiene un 10, la cosa cambia. He dado forma de barras al recuento de las citas para demostrar que el equilibrio en lo que respecta al atractivo entre hombres y mujeres que acudieron a esas citas era el que cabría esperar al ser emparejados al azar. No hay pruebas de que nadie engañase al sistema, por ejemplo, por medio de recomponer la fotografía antes de acudir a la cita y después escabullirse sin ser visto al llegar su cita a ciegas y no ser lo que esperaba. Las cifras de satisfacción (de los hombres) son los porcentajes en gris:

cómo afecta el atractivo a la satisfacción masculina en una cita



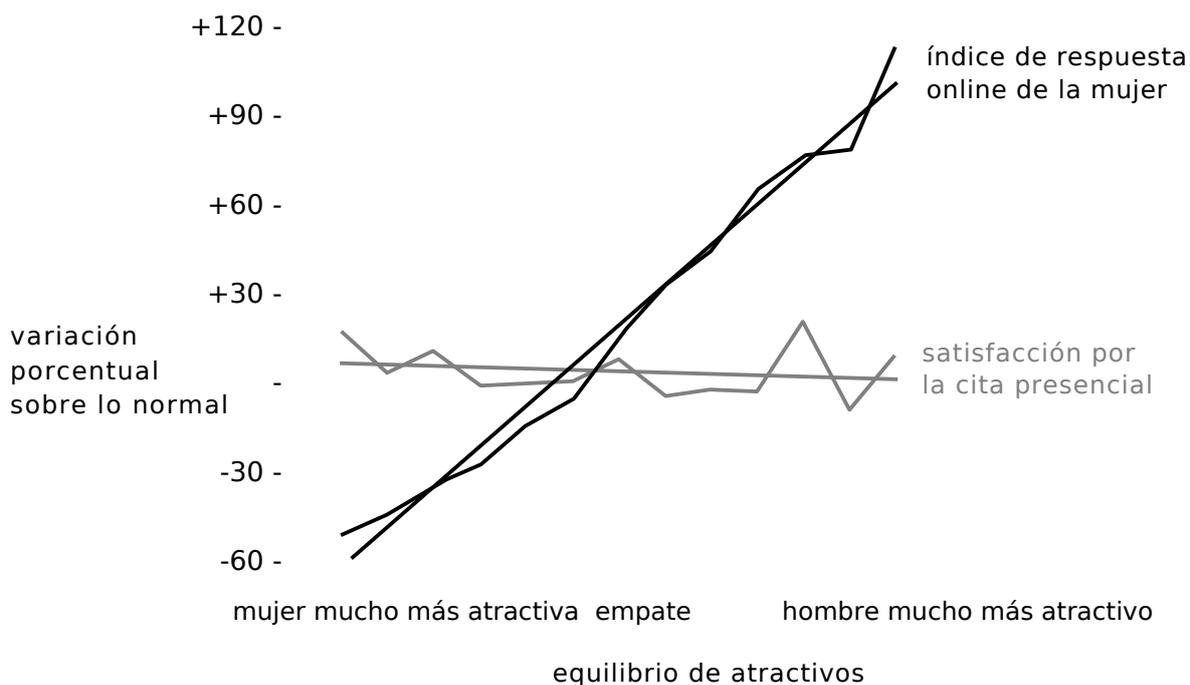
Y aquí están los mismos datos, pero ahora en el caso de las mujeres:

cómo afecta el atractivo a la satisfacción femenina en una cita



De las dos series de datos de Crazy Blind Date se desprende que la gente no pareció preocuparse mucho de la apariencia física de sus acompañantes. Las mujeres lo pasaron bien en un 75 por ciento de las veces; los hombres, en un 85 por ciento. El resto de las variaciones son básicamente ruido. Esa indiferencia por el aspecto es justo lo contrario de lo que se aprecia en los datos de OkCupid. Por ejemplo, he hecho un gráfico de los datos de satisfacción presencial de arriba (los números en gris) junto con el índice de respuesta de esas mismas mujeres a los mensajes en la web. Para facilitar su comparación, las líneas muestran las desviaciones respecto al promedio de sus respectivas cantidades:

respuesta femenina al atractivo masculino



En el caso masculino, el gráfico comparativo es muy parecido a este y, para ser claros, los datos que dan pie a las dos líneas del diagrama proceden del mismo grupo de personas. La línea negra es la experiencia en OkCupid y la gris, la de Crazy Blind Date. Dicho en dos palabras: la gente parece ser sumamente selectiva en la web respecto a algo que, una vez están sentados cara a cara, no parece importarles.

Ese tipo de preselección superficial la encontramos por doquier. De hecho, ahí hay mucho dinero en juego. ¿Sabes cuál es la diferencia entre el paracetamol de las marcas Tylenol y Kroger? La caja. A menos que te mediques más que un yonqui y pretendas engullirte el paquete entero, no hay motivo alguno para pagar el doble por unas moléculas que tienen otro nombre pero cuyas propiedades las determinan las inmutables leyes de la química. Aun así, tengo un gran frasco rojo de Tylenol encima de la cómoda.

Naturalmente, prestamos mucha más atención a las etiquetas cuando son las personas las que las llevan adheridas. En términos de compatibilidad superficial, los que se autoproclaman demócratas y republicanos son los que peor se llevan de los principales grupos de OkCupid, peor incluso que los protestantes y los ateos. Esto lo sé por las muchas preguntas que hace la web, que lo abarcan prácticamente todo y de las que el usuario medio viene a responder a unas 300. El sitio web te permite decidir la importancia de cada una de las preguntas que contestas, y puedes precisar qué preguntas acertarías (y cuáles no) de una posible pareja. Pese a todo este control, en el caso de la política el sistema se viene abajo. Cuando miras más allá de las etiquetas para

ver de verdad quién manda mensajes a quién y quién contesta (y, por tanto, quién acaba quedando con otra persona), lo que realmente importa para la compatibilidad mutua es la preocupación por la política, en un sentido u otro, más que los detalles de cualquier creencia concreta. Esto lo confirmamos mediante un experimento que duró todo el verano de 2011.

La gente tiende a volverse loca con todas esas preguntas para buscar coincidencias; marcan todo tipo de cosas como obligatorias y, básicamente, ponen a la vista de todos una lista de cosas imprescindibles: busco a una persona liberal, no fumadora, agnóstica y amante de los perros que no tenga hijos —y que sea buena en la cama, claro—. Pero hay preguntas muy humildes, como «¿Te gustan las películas de terror?» o «¿Has viajado alguna vez a otro país?», cuyo poder de predicción es increíble. Si alguna vez te bloqueas y no sabes qué preguntarle a alguien en la primera cita, prueba con esas. Cerca de tres cuartas partes de las parejas duraderas que se han conocido a través de OkCupid, ambas personas contestaron de igual forma: o los dos dijeron «sí» o los dos «no». La gente tiende a poner demasiado énfasis en las cosas grandes y llamativas: la fe, la política y, por supuesto, el aspecto físico, pero son cosas que no importan tanto como todo el mundo cree. A veces no importan nada.

Pese a que acabase siendo un fiasco, el «Día del amor ciego» nos proporcionó un ejemplo más bien visceral de lo que hace la gente cuando carece de información. Al ocultar las imágenes sin cambiar nada más creamos un experimento en tiempo real que pudimos cotejar con la actividad habitual de la web. Durante siete horas, nuestros usuarios obraron sin tener acceso a la mismísima cosa que nuestros datos previos indicaban que era la información más relevante que podía ofrecer OkCupid: el aspecto que tenían los demás.

Parte de lo que ocurrió era predecible. La gente se enviaba mensajes sin los típicos prejuicios, o sin sesgos raciales o estéticos. Lo que el usuario no podía ver, tampoco podía juzgarlo. Pero de los 30 333 mensajes que se enviaron a ciegas, 8912 acabaron obteniendo respuesta, un 40 por ciento más de lo habitual. Y a los que estuvieron allí, en la oscuridad, les pasó algo asombroso. El 24 por ciento de las parejas que estaban hablando cuando se ocultaron las fotos habían intercambiado ya datos de contacto para cuando volvimos a colocar las imágenes. Eso ocurrió solo en la ventana de siete horas del «Día del amor ciego». Las cifras que habíamos previsto para ese tiempo apenas llegaban a la mitad. No es solo que la gente escribiese mensajes con más probabilidades de obtener respuesta, sino que estaban dando muchos más

números de teléfono y direcciones de correo... y a personas a quienes nunca habían visto.

Sin embargo, para las parejas que habían empezado a hablar y a conocerse cuando volvimos a activar las fotos a las cuatro de la tarde, el día tuvo un efecto contrario. Esas habían estado a oscuras y de repente se encendieron las luces y, solo con mirar los datos, podemos hacernos una idea de la cara de espanto que pusieron. Conversaciones que hasta entonces habían sido extensas, en cuanto le dimos al interruptor duraron una media de 4,4 mensajes más. Si se comparan con los datos de control que tenemos, deberían haberse alargado por lo menos 5,6 mensajes más. Los intercambios de datos de contacto que se producen al término de esas conversaciones también cayeron en picado de forma parecida cuando dimos la luz.

Las webs de contactos están diseñadas para brindar a la gente los instrumentos y la información que les permita conseguir lo que quieran para dejar de estar solos: sexo ocasional, unas cuantas citas para pasarlo bien, una pareja, un matrimonio... lo que sea. Hay cosas que están ahí y son muy fáciles de buscar y de encontrar, como la estatura, las opiniones políticas, las fotos, los escritos y demás. Están ahí para ayudar a la gente a que juzgue y satisfaga sus deseos, y por muy fascinante que sea desmenuzar y analizar esos juicios y esos deseos, creo que tiene algo de despropósito. Las personas eligen a partir de la información que les proporcionamos porque pueden, no necesariamente porque deban.

No puedo evitar pensar en la cantidad de gente que es rechazada por culpa de algún impedimento percibido que en realidad no le preocupa a nadie y me pregunto si Internet ha cambiado las relaciones amorosas del mismo modo que ha cambiado tantas otras cosas... y por la misma razón. Si se me permite dar voz a mi «anti Mick Jagger» interior, diré que en la Red uno sí que puede siempre obtener lo que quiere. Pero si hablamos de lo que uno necesita, eso ya es mucho más difícil de encontrar.



PARTE 1

Lo que nos separa

6.

El factor de interferencia

Si te paras en la esquina suroeste del cruce entre las calles Cincuenta y ocho y Cincuenta con un cuaderno para apuntar y te dedicas un rato a observar a la gente, concluirás enseguida que los neoyorquinos son guapos, delgados y, sobre todo, ricos. No hay tela, broche o pliegue que no reluzca como el oro. Es cierto que muchos neoyorquinos son ricos, pero no es oro todo lo que aquí reluce. Estás parado delante de Bergdorf Goodman, y eso es un factor de interferencia.

Ese es un término técnico que define algo que no has tenido en cuenta en tu análisis pero que, sin embargo, afecta a su resultado. Asegurarte de que no estás centrándote con obcecación en alguna versión en bits digitales del Upper East Side es una de las tareas más arduas y tediosas de trabajar con datos digitales. Cuando te parece que tienes todas las variables y todas las posibilidades disponibles para el análisis y la especulación, entonces eres libre de dirigir tu investigación allí donde te lleve la curiosidad. Pero, para ser fieles al cliché, esa libertad exige una eterna vigilancia.

Y aquí es donde tengo que admitir una cosa. En lo que hemos visto hasta ahora en estas páginas, allí donde presento datos de opiniones de unas personas sobre otras, en los votos, en los datos de Crazy Blind Date, en las tablas y los gráficos —en todos los porcentajes y en todos los totales—, allí donde un usuario ha valorado a otro, ambas personas eran blancas. He tenido que hacerlo así, porque cuando te fijas en cómo se comportan dos estadounidenses desconocidos en un contexto amoroso, la raza es el colmo del factor de interferencia. Y para asegurarme de que lo que quiero decir sobre la atracción o el sexo estuviese referido únicamente a esas ideas, tenía que despejar el panorama.

Como estadounidense que soy, el instinto reflejo de barrer la raza debajo de la alfombra es una cosa innata en mí, así que, en cierta medida y aunque los números me obligasen, lo que he hecho me ha salido de manera natural. Y aun dejando de lado la peculiar relación que tiene nuestro país con el tema, la larga tradición que tenemos de falsa integración y lamentable pseudociencia hacen que cualquier análisis cuantitativo de la raza sea especialmente peliagudo. Eso no quiere decir que no dispongamos de nuestros buenos

números. Los tenemos a montones, de cierto tipo; mis datos preferidos son los de persona a persona, pero estos otros los veo más de persona a cosa: un grupo u otro frente a las tasas de desempleo, los exámenes de admisión a la universidad, el sistema de justicia, el cáncer... Si bien este tipo de estudios nos han ayudado a señalar y (en ocasiones) a abordar la desigualdad, hay algo que les falta. Falta el ser humano que se encarga de llevar a cabo (o no llevar a cabo) la contratación, la enseñanza, el trabajo policial, el cuidado preventivo; falta la gente que generó los resultados que todos estos estudios afirman cuantificar. Así que acabas sacando conclusiones como esta: «Los imputados de raza negra tienen por lo menos un 30 por ciento más de probabilidades de ser encarcelados que los blancos por el mismo delito». La voz pasiva del titular lo dice todo. ¿Quién está cometiendo aquí el error judicial? Sintácticamente hablando, nadie. En la práctica, lo tengo bastante claro. Pero se trata de un singular estudio que mira más allá de las instituciones, a ese valor binario de «nosotros contra ellos» tan fundamental en las relaciones raciales.

Detrás de cada uno de los bits de mis datos hay dos personas, el que actúa y el que recibe la acción, y el hecho de que ahora los veamos a ambos como iguales en el proceso es algo nuevo. Si en todo esto de los datos hay verdaderamente algo de «-clismo», si el título de este libro no es algo más que un mero juego de palabras semigracioso o un accidente del alfabeto, es esto: nos permite ver de una vez la experiencia humana en su totalidad, no solo aquella parte a la que, por lo que sea, estamos prestando atención en un momento determinado.

Antes del advenimiento de esta nueva modalidad de datos, uno de los panoramas más cuantificados de la vida pública era el de los deportes. Disponemos de cifras en tiempo real sobre cualquier interacción imaginable, que podemos diseccionar y recombinar a placer. Tal vez sorprenda, entonces, que sea en los deportes donde el debate sobre la raza resulta menos analítico. La polémica sobre el «quarterback negro» que se prolongó durante la primera década de este milenio es el ejemplo perfecto. Durante años se produjo el mismo ciclo de noticias: un quarterback afroamericano salía escogido al principio del *draft* o lo habían puesto de titular en un partido importante y siempre salía alguno a decir que los negros no pueden triunfar en ese puesto en la NFL, la liga de fútbol americano. El motivo que solían aducir era que les faltaba inteligencia para ello. Se producían reacciones airadas, debates y muchas discusiones sobre que aquello no eran más que estereotipos miserables. Pero entre tanto comentario y tanta protesta, y tantas protestas

sobre las protestas, en los 97 000 resultados que arroja Google sobre «quarterback negro» solo he encontrado un artículo en el que de verdad se han calculado las puntuaciones de quarterbacks blancos y negros, que resultan ser idénticas hasta en el segundo decimal: 81,55. En un ámbito tan obsesionado con las estadísticas en el que hordas de chalados de los números calculan la tasa de éxito del 54 por ciento que tiene Fulanito Pateabolas de marcar un gol de campo desde las 50 yardas en un partido como visitante que se decide por siete puntos o menos contra un rival de la AFC (federación americana de fútbol), uno podría pensar que lo de comparar estadísticamente a quarterbacks blancos y negros sería casi un instinto primario. Sin embargo, siempre hubo —y sigue habiendo en general— un sepulcral silencio numérico al respecto. En lugar de ello encontramos que se recurre a la retórica y a la anécdota. Pero lo único que se consigue con un debate de ese estilo es que todo el mundo siga creyendo tener la razón cuando, en realidad, pese a toda esa palabrería, con un simple número —81,55— se puede demostrar claramente que unos están equivocados. El artículo donde se daba ese cálculo tenía cero tuits y cero «Me gusta» en Facebook, por cierto, y no es que estuviese colgado en algún blog ignoto: aparecía publicado en *The Big Lead*, que es propiedad de *USA Today*. A veces tiene uno la sensación de que la gente no quiere saber.

Mientras que en situaciones como esta parecemos carecer de la voluntad de examinar la raza a través de una lente estadística, en muchos otros ámbitos lo que nos falta es sencillamente tener datos. Hay pocos aspectos de la vida que hayan sido tan obsesivamente cuantificados como el fútbol americano. Pero eso está cambiando a toda velocidad.

En OkCupid, una de las maneras más sencillas de comparar a una persona negra y a una blanca (o a dos personas de cualquier raza) es mirar su «porcentaje de coincidencia». Ese es el término que empleamos en la web para referirnos a la compatibilidad. A los usuarios se les hacen un puñado de preguntas, ellos responden y un algoritmo predice si se van a llevar bien, digamos, cuando se tomen una cerveza o durante una cena. A diferencia de otras funciones de OkCupid, el porcentaje de coincidencia no tiene componente visual alguno. El número correspondiente a dos personas solo refleja lo que podríamos llamar su yo interior, todo aquello sobre lo que opinan, lo que necesitan y desean, incluso lo que les parece gracioso, pero nada acerca de su apariencia física. Si los valoramos solo por esta medida de compatibilidad, los cuatro mayores grupos raciales que hay en OkCupid — asiáticos, negros, latinos y blancos— se llevan más o menos igual de bien^[12].

De hecho, la raza afecta menos al porcentaje de coincidencia que la religión, la política o la educación. De los detalles que los usuarios consideran importantes, lo más comparable a la raza es el signo zodiacal, que tampoco es que afecte demasiado. Para un ordenador no entrenado a distinguir estas categorías, «asiático», «blanco» y «negro» podrían perfectamente ser «aries», «virgo» y «capricornio».

Pero esta neutralidad racial solo es teórica; la cosa cambia cuando entran en juego las opiniones de los usuarios y no solo los mecanismos de un algoritmo, que son insensibles al color. A partir del perfil completo, en el que la foto domina la página, así es como los usuarios de OkCupid se valoran unos a otros por su raza:

valoración media asignada por hombres a mujeres en OkCupid

		raza de la mujer			
		asiática	negra	latina	blanca
raza del hombre	asiático	3,16	1,97	2,74	2,85
	negro	3,40	3,31	3,43	3,23
	latino	3,13	2,24	3,37	3,19
	blanco	2,91	2,04	2,82	2,98

Aquí he puesto los datos en bruto, sin adornos, porque a estas alturas ya estarás familiarizado con el sistema de valoraciones de una a cinco estrellas de OkCupid. Pero para facilitar que se vean mejor las tendencias, voy a usar esa misma tabla y a «normalizar» cada hilera. En la tabla de abajo, las entradas son la diferencia de porcentaje (+/-) respecto a la media (lo «normal») de la hilera. Se trata de la misma información pero expresada de manera un poco distinta. Piensa en el número normalizado como la preferencia relativa de los hombres por las mujeres. Por ejemplo, como se puede ver, los hombres asiáticos piensan que las mujeres asiáticas son un 18 por ciento más guapas que la media, mientras que a los hombres negros les parece que están solo un 2 por ciento por encima. Y lo mismo con el resto:

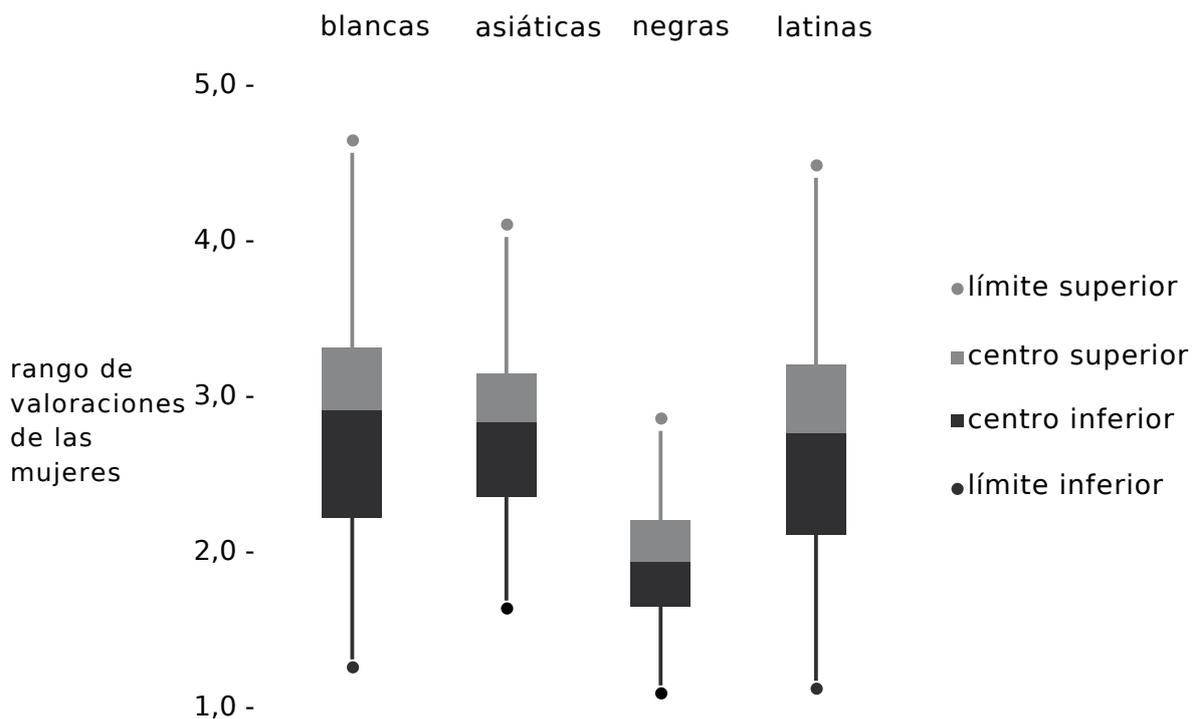
valoración normalizada asignada por hombres a mujeres en OkCupid

		raza de la mujer			
		asiática	negra	latina	blanca
raza del hombre	asiático	+18 %	-27 %	+2 %	+7 %
	negro	+2 %	-1 %	+3 %	-4 %
	latino	+5 %	-25 %	+13 %	+7 %

| blanco +8 % -24 % +5 % +11 %

Enseguida veremos otras cosas más allá de OkCupid, y cuando presente tablas parecidas lo haré exponiendo directamente cifras normalizadas. Pero, de momento, los dos patrones básicos de la atracción entre hombres y mujeres están claros: a los hombres tienden a gustarles mujeres de su misma raza. Y, lo que es más, no les gustan las mujeres negras. Los datos obtenidos de los mensajes están muy correlacionados con estas valoraciones, de modo que también siguen este patrón^[13].

Solo para que se vea que estas tendencias de votación no las está arrojando algún oscuro artilugio estadístico, he colocado las cifras en bruto de los votos por cabeza en lo que se llama un diagrama de cajas: te dice dónde está el grueso de un paquete de datos. Abajo puedes ver que el bloque central de mujeres negras se ha valorado casi íntegramente por debajo de los otros tres grupos étnicos y que el extremo superior de las mujeres negras está hacia la mitad de la altura que ocupan los otros tres colectivos:



Matemáticamente hablando, esto es un producto con precio de descuento: ser negra básicamente te cuesta unas tres cuartas partes de estrella de tu valoración, aunque estés en lo más alto de tu grupo. Además, si haces el análisis a la inversa y te fijas en la gente que emite los votos, verás un patrón de parecido componente mayorista. La mayor parte de los hombres que no son negros aplican esa reducción de tres cuartos a las mujeres negras. No es

que haya un escuadrón de racistas que se dedique a bajar los promedios por su cuenta.

Por muy sorprendente que pueda parecernos esto, solo refleja una serie de datos, lo que piensa un grupo de personas. Así que este es un buen momento para hacer una pausa y responder a una pregunta que tal vez te hayas planteado antes, dado lo mucho que he hablado hasta ahora de los datos de OkCupid en estas páginas: ¿Quién es esta gente?

En términos sumamente superficiales, los miembros de OkCupid representan a la composición general de los usuarios de Internet, claro que con la particularidad de que (casi) todos los que están en la web son solteros. Los usuarios del sitio son más jóvenes que el promedio nacional (la edad media de OkCupid es de 21) y tienden a ser menos religiosos. La composición racial es la que cabría esperar. A continuación vemos nuestras cifras comparadas con las del «usuario estadounidense de Internet» genérico de Quantcast, la principal empresa de medición de audiencias de Internet, que es como la Nielsen de la Red.

	Usuarios de OkCupid	Usuarios estadounidenses de Internet
asiáticos	6 %	4 %
negros	7 %	9 %
latinos	8 %	9 %
blancos	80 %	78 %

Si bajamos un nivel demográfico más, los usuarios de OkCupid son, en todo caso, bastante más urbanos, formados y progresistas que el resto del país. Los mayores mercados de la web son, con mucho, sitios como Nueva York, San Francisco, Los Ángeles, Boston y Seattle. El 85 por ciento de los usuarios ha asistido a la universidad. Los que se proclaman liberales superan a los que se describen como conservadores en una proporción de más de dos a uno. En toda la web existe una mentalidad abierta de forma generalizada. Y un total del 85 por ciento de los usuarios, que por hilarante que parezca no es intencionado, responde a esta pregunta:

¿Te plantearías salir con alguien que haya verbalizado un fuerte sesgo negativo hacia determinado grupo étnico?

con una negativa rotunda (escogen «No» por encima de «Sí» y de «Depende»). A la vista de los datos anteriores, eso significa que el 84 por

ciento de los que están en OkCupid no se plantearía quedar con alguien de OkCupid.

Básicamente, cualquier cosa que en teoría hace que un grupo de personas sea «menos racista», eso es lo que son los usuarios de OkCupid. Esto es lo que le digo a la gente que, como yo, vive tranquilamente en ciudades grandes y plurales, que considera que sus opiniones y gustos son, si acaso, tolerantes, que se relaja por la noche con una copa de vino y una dosis o dos de rectitud progresista en Facebook: cuando muestro aquí que las mujeres negras y, luego, los hombres negros, reciben un trato de segunda y que añadir el color de piel blanco a la identidad de un usuario lo hace más atractivo, sea hombre o mujer, no estoy describiendo ninguna pesadilla digna de los catetos de lo más profundo de la América montañosa. Estoy describiendo nuestro mundo, el mío y el tuyo. Si estás leyendo este libro de divulgación sobre el fenómeno *Big Data* y todos sus portentos, ten por seguro que los datos que en él encuentras se corresponden contigo.

Pero fíjate otra vez en la pregunta de arriba, que fue escrita por uno de los usuarios de OkCupid y se ha contestado cerca de un millón de veces: «verbalizar» es una palabra muy rara. Si te deshaces de ella, lo que queda sigue siendo, más o menos, «¿Saldrías con una persona racista?», que siempre pensé que era la verdadera intención de la pregunta. El que la escribió, no obstante, comprendió las sutilezas de los datos antes que yo. En una web de contactos uno puede actuar por instintos que en otros sitios mantendría ocultos. En cierta medida, los usuarios se dedican a juzgar y a ser juzgados por otros y todos ellos acceden a la web libres del contexto de su vida cotidiana. No hay nada que se publique en los muros de los amigos. El juego consiste en esto: la web te muestra a personas y a ti te gustan o no, hablas con ellas o no. Eso es todo lo que hay. En un mundo digital que, por otra parte, está obsesivamente conectado por redes, lo de los contactos y las citas tiene algo de solitario y de vieja escuela. Tu experiencia consiste en ti mismo y la gente con la que eliges estar; y lo que haces es secreto. Muchas veces, el propio hecho de que uno tenga una cuenta —por no hablar de lo que hace con ella— no lo saben ni sus amigos. De modo que la gente puede obrar con una actitud y unos deseos relativamente exentos de cualquier presión social.

En la mente del lego, Facebook, «la red social», es el *sine qua non* de las fuentes de datos *online*. Y no cuesta entender por qué: Facebook es enorme y omnipresente, y una muestra de sus usuarios es prácticamente una muestra de la población mundial que tiene acceso a Internet; dicho de otro modo, puedes extraer de allí un corpus representativo de lo que quieras. Y sus datos son

muy sólidos y variados: saben con quién fuiste al instituto, qué canción acabas de escuchar en Spotify, dónde viven tus padres, etcétera.

Pero así como muchas veces esa riqueza es un valor, también puede ser una carga. Rara vez encuentras a un desconocido en Facebook. Es, por su propio diseño, un sitio de personas que ya conoces y acerca de quienes ya te has formado una opinión; son tus amigos, al fin y al cabo. Los datos de Facebook sobre la raza son la encarnación de ese solipsismo tan frecuente de «Pero si yo tengo amigos negros». El modo en que tratas a tus amigos es, por definición, la excepción a cómo tratas al resto de la humanidad. Y, para empezar, tanto tú como la relación con tus amigos os formasteis fuera de las redes.

Por añadidura, la gente se inhibe cuando sus amigos los están mirando. Este aspecto de escaparate nos explica por qué el primer paso de la mayoría de las aplicaciones de contactos y citas que hay en Facebook consiste en sacarte de Facebook: tu existencia allí está siempre bajo custodia. Hace mucho tiempo probamos a meter funciones «sociales» en OkCupid, pero fallaron estrepitosamente, igual que ocurrió con opciones parecidas cuando las probaron en Match.com. Por el motivo que sea, la gente no quiere tener a su red de contactos alrededor cuando liga por Internet. Ese deseo de soledad tiene el mismo origen, supongo, que la claustrofobia que nos atenazaría a la mayoría si, en una prometedora primera cita en un restaurante, se apostasen en la mesa de al lado dos amigos de toda la vida. Esto no es para desmerecer en absoluto el negocio ni la comunidad que ha creado Facebook, pero las relaciones de la «vida real» sobre las que se afianza Facebook y que predominan allí les confieren a sus datos un poder especial. Cuando quieres detenerte en algo como la raza, en lo que, al menos entre gente decente, existe cierta presión para comportarse de determinada manera en público, las webs de contactos te brindan una serie de datos excepcionales: todo el mundo es un desconocido, están todos solos y lo que hacen es decirte quién les gusta y quién no^[14].

Así que vamos a confrontar los datos de OkCupid con los de otras webs de contactos y a ver qué sale del cóctel. Si miramos las cifras de otros usuarios que se mueven en otras interfaces nos haremos una idea general mucho mejor del verdadero patrón. Y eso es lo que vemos abajo: aquí tenemos datos de OkCupid, DateHookup y Match.com —sitios que, combinados, sumaron unos 20 millones de usuarios estadounidenses solo el año pasado— presentados unos junto a los otros. Las matrices varían un poco en los detalles —recordemos que estos valores reflejan acciones de personas

distintas que emplean distinto *software*—, pero si obviamos esas diferencias el patrón general es el mismo. En términos de la «dirección» de las opiniones, de lo que gusta o no gusta, las tablas son prácticamente idénticas:

		raza de la mujer			
OkC		asiática	negra	latina	blanca
raza del hombre	asiático	+18 %	-27 %	+2 %	+7 %
	negro	+2 %	-1 %	+3 %	-4 %
	latino	+5 %	-25 %	+13 %	+7 %
	blanco	+8 %	-24 %	+5 %	+11 %
Match		asiática	negra	latina	blanca
asiático		+50 %	-68 %	-14 %	+31 %
negro		+9 %	-13 %	+8 %	-3 %
latino		+4 %	-67 %	+33 %	+29 %
blanco		+13 %	-68 %	+8 %	+47 %
DH		asiática	negra	latina	blanca
asiático		+11 %	-24 %	+9 %	+4 %
negro		+7 %	-9 %	+9 %	-7 %
latino		+12 %	-27 %	+10 %	+6 %
blanco		+18 %	-30 %	+6 %	+5 %

Probablemente conozcas Match.com. Lleva casi dos décadas siendo la web de contactos más popular de Estados Unidos. Emiten toneladas de publicidad en la televisión nacional y, a resultas de ello, cuentan exactamente con los datos demográficos de los «americanos típicos» que cabría esperar. DateHookup es una web gratuita con varios millones de miembros y muy popular entre los que buscan citas esporádicas; su base de usuarios consiste en algo menos del 20 por ciento de negros y el 13 por ciento de latinos. Es el más diversificado de los tres sitios que estudiamos aquí. Yo lo veo como un Atlanta o un Houston en comparación con el Portland que sería OkCupid y el Dallas que sería Match. Pero como puedes ver, en lo que se refiere a hombres que valoran a mujeres, en los tres sitios web encontramos el mismo patrón.

Los votos en la otra dirección, mujeres que valoran a hombres, no son tan uniformes en los tres sitios, pero son bastante parecidos:

		raza de la hombre			
		asiático	negro	latino	blanco
raza de la mujer	OkC				
	asiática	+19 %	-38 %	-15 %	+35 %
	negra	-34 %	+52 %	-17 %	-1 %
	latina	-35 %	-20 %	+19 %	+37 %
	blanca	-26 %	-19 %	-1 %	+46 %
	Match				
	asiática	+3 %	-7 %	-5 %	+9 %
	negra	-9 %	+10 %	-1 %	+0 %
	latina	-8 %	-6 %	+6 %	+8 %
	blanca	-7 %	-5 %	-0 %	+12 %
	DH				
	asiática	-	-34 %	+14 %	+20 %
	negra	+9 %	+25 %	-12 %	-22 %
	latina	-18 %	-14 %	+21 %	+10 %
	blanca	-12 %	-25 %	+7 %	+31 %

Estas matrices muestran dos tendencias negativas y dos positivas. Los negros vuelven a estar infravalorados por los usuarios no negros, pero los hombres asiáticos se les han unido en los números grises. Por el lado positivo, las mujeres prefieren claramente a hombres de su misma raza —son más «leales a su raza» que los hombres—, pero también manifiestan una clara preferencia secundaria por los hombres blancos.

En OkCupid se nos abre otro camino que nos permite adentrarnos en las jerarquías raciales y que refuerza esta «preferencia blanca». Dado que los usuarios pueden escoger más de un grupo étnico en su perfil, podemos estudiar las mezclas raciales casi como si lo hiciésemos en un laboratorio. Por ejemplo, tenemos a hombres que seleccionan «asiático» como identidad étnica. También tenemos a hombres que escogen «asiático» y «blanco». Si comparamos los dos grupos nos haremos una idea de lo que una persona consigue al añadir esa «blancura». Y resulta que consigue bastante. Cuando añades el color blanco las valoraciones suben, y eso ocurre de manera generalizada. A continuación he volcado todos los datos. Es una tabla grande y liosa, pero que vale la pena explorar.

En la columna de la derecha se ven las mejoras en las valoraciones generadas al añadir el blanco al componente racial de la persona. El mensaje principal es que el descuento racial que hemos visto aplicado a hombres y mujeres negros y a hombres asiáticos en las tablas anteriores queda anulado. Es lo opuesto a la vieja «regla de una gota».

Por desgracia, no hay suficiente gente que marque «negro» y «latino» o «asiático» y «negro» para engrosar plenamente esta alquimia, pero constituye un intrigante vistazo a cómo está el espectro étnico:

		raza de la mujer			
hombres valorando a las mujeres		latina	latina + blanca	% de cambio	
raza del hombre	asiático	2,7	2,8	+4	
	negro	3,4	3,4	-2	
	latino	3,4	3,4	+1	
	blanco	2,8	3,0	+7	
			negra		
			negra + blanca		
	asiático	2,0	2,3	+19	
	negro	3,3	3,5	+5	
	latino	2,2	2,9	+28	
	blanco	2,0	2,5	+24	
			asiática		
			asiática + blanca		
asiático	3,2	3,0	-5		
negro	3,4	3,6	+5		
latino	3,1	3,3	+5		
blanco	2,9	3,0	+2		
		raza del hombre			
mujeres valorando a los hombres		latino	latino + blanco	% de cambio	
	asiática	1,7	1,8	+7	
	negra	2,0	2,4	+18	
	latina	2,1	2,2	+8	
	blanca	1,8	2,1	+15	

		negro	negro + blanco	
raza de la mujer	asiática	1,5	1,6	+6
	negra	2,7	2,6	-4
	latina	1,7	1,9	+17
	blanca	1,6	2,0	+26
		<hr/>		
		asiático	asiático + blanco	
	asiática	2,0	2,1	+4
	negra	1,8	2,7	+48
	latina	1,5	2,2	+44
	blanca	1,5	2,0	+32

Ahora bien, todo esto está extraído de las valoraciones de una web de contactos. Pero los datos de ese tipo son básicamente datos de la primera impresión, del primer flechazo —los usuarios tienen que llegar a conocerse mutuamente, al menos un poco, antes de querer besarse—, y ese es el mismo espíritu básico que se da cuando cualquier par de personas se encuentran: «Vale, ¿qué es lo que estoy mirando? ¿A quién veo?». Los datos miden el escalofrío de conocer a alguien nuevo: ese estallido de juicio, instinto y química que determina si te gusta o no una persona y que se produce antes incluso de que sepas prácticamente nada de ella. He aquí a unos cuantos usuarios de OkCupid que lo explican con sus propias palabras:

Y entonces, un día, creo que estaba mirando mis coincidencias diarias, y allí estaba él. Hice clic en su perfil al instante... Había algo en él que me hizo sonreír.

Bella, acerca de Patrick

Bueno, todo empezó cuando un día estaba consultando mis coincidencias y vi a esa chica que me pareció atractiva a primera vista.

Dan, acerca de Jenn

Pero si existe el amor a primera vista, existe también el desagrado a primera vista, ¿verdad? ¿Y no es ese mismo escalofrío de la atracción, pero al contrario, cuando alguien se arruga, aunque sea de manera inconsciente, al ver a un desconocido? También, en palabras de alguien:

Hay muy pocos hombres afroamericanos que no hayan pasado por la experiencia de ir paseando por la calle y oír cómo se van cerrando los seguros de los coches a su paso. Eso me pasa a mí... Hay muy pocos afroamericanos que no hayan pasado por la experiencia de meterse en un ascensor y que una mujer se aferre con nerviosismo a su bolso y contenga la respiración hasta que por fin llega a su piso. Eso pasa mucho.

Barack Obama, 19 de julio de 2013

Estos destellos de intuición que hay en el corazón de los datos — extrapolaciones de las cantidades más diminutas de información— no atañen solo a lo sentimental, sino a quién escoges para alquilarle tu piso, a la hora de decidir si dar el visto bueno a un préstamo o no y, con certeza, al trabajo policial, en el que muchas veces no hay tiempo más que para un destello. Hasta en las situaciones más pausadas, la primera impresión es la que cuenta. En un artículo se preguntaba: «¿Son Emily y Greg más contratables que Lakisha y Jamal?» y obtuvo un rotundo «Sí» como respuesta por parte de los profesionales de recursos humanos del país. Los científicos enviaron currículos idénticos, algunos de ellos encabezados por nombres que «sonaban a negro» y otros con nombres que «sonaban a blanco», y se percataron de que los segundos recibían un 50 por ciento más de respuestas, independientemente del cargo o del sector. Y las empresas que dicen contratar según «igualdad de oportunidades» discriminan lo mismo que todos los demás.

Este tipo de ironías nos llevan a ver que los grandes estudios son importantes pero las pequeñas mediciones a nivel de persona a persona resultan esenciales: cuando lees cosas como el ejemplo antes mencionado y ves que Jamal no consigue el trabajo, es fácil que menees la cabeza con desaprobación ante esos pocos encargados de contratación racistas que han inclinado la balanza en su contra. Pero los datos que vemos en este capítulo demuestran que el racismo no es un problema de extremos atípicos. Está por todas partes. Hemos visto los mismos patrones repetidos en tres sitios web con distintos usuarios y experiencias diferentes: hombres, mujeres, gratuitos, solo por suscripción, informales, serios, de demografía más «urbana» y más «convencionales». En definitiva, los sujetos de la investigación representan una buena porción de los adultos jóvenes de este país y los datos demuestran que los no negros infravaloran los perfiles de los afroamericanos. No se trata de un problema ocasionado por un pequeño grupo de usuarios negros «feos» ni por un pequeño colectivo de racistas impenitentes que desbaratan un patrón por otra parte regular.

Ser abiertamente racista ya no es algo socialmente aceptado. En respuesta a esa presión, parte de la población se ha salido por la tangente: si ya no puedo meterme a gritos con unos cuantos chavales en edad escolar, pues vale, le gritaré a la tele. No es ese el típico estadounidense. La mayoría de nosotros —casi todos, en realidad— reconocemos que el racismo es malo. Pero todavía está implícito en muchas de las decisiones que tomamos^[15]. Los psicólogos le dan un nombre a la estructura interior de nuestro pensamiento que nos ayuda a organizar la información al toparnos con ella: esquema. Y nuestro esquema está todavía fuera de sincronía con cómo creemos la mayoría que debería ser el mundo. Por medio de centenares de pequeños actos diarios, ninguno de ellos llevados a cabo con intenciones o sentimientos racistas, reflejamos una cultura que es, de hecho, racista. Como hemos visto, ese patrón está tan arraigado en nuestro tejido que adiciones recientes a nuestra sociedad, como los asiáticos y los latinos, también lo han adoptado.

En lo que se refiere a estos patrones, los individuos están en cierto modo libres de culpa. Que la gente negra obtenga tres cuartas partes del afecto que reciben los blancos en webs de contactos es prácticamente un accidente. No se puede culpar a nadie por no querer salir con otra persona. En esa decisión rara vez interviene la malevolencia. Juicios como votar a alguien se hacen en un instante, y son actos pequeños y aparentemente carentes de significado. Navegas por la web y tal vez una de cada doce caras es negra. Y al ver a esa persona, tu acción en ese momento podría ir en cualquier dirección, igual que si estuvieses mirando a una persona blanca; te dejas llevar. ¿Y qué pasa si no te gusta una persona concreta en un momento determinado? Todo el mundo tiene derecho a pensar lo que quiera de cualquier individuo; de hecho, considerar en primer lugar a cada persona como un individuo, y no como una categoría, es dar un enorme paso en la dirección adecuada. Lo único que pasa es que si nos fijamos en el conjunto de patrones, estos nos indican que, en general, los dados siguen estando trucados. En realidad, podemos usar una metáfora de la misma categoría general algo más adecuada: nos indican que la casa se sigue llevando lo suyo; no es el crupier, no es la mano, no es ni siquiera la partida, son las reglas del juego lo que hace que determinados grupos de personas ganen y que otros pierdan.

El profesor de sociología Osagie K. Obasogie emprendió hace poco una curiosa investigación: entrevistó a personas ciegas de nacimiento y descubrió que tenían las mismas actitudes acerca de la raza que las que se dan en el mundo de los videntes. Su muestra era relativamente pequeña —solo 106 individuos—, pero lo que descubrió fueron mis datos de OkCupid hechos

carne y hueso. Cita numerosos ejemplos de jóvenes ciegos que están contentos con su acompañante en una cita hasta que alguna «pista» — normalmente el tacto de su cabello, pero también a veces el susurro de un desconocido— les revela que la otra persona es negra. Se acabó la cita.

Obasogie afirma que la actitud de la gente ciega sobre la raza es fruto de toda una vida de absorción cultural, a falta de cualquier realidad visual. A partir de estos datos parece imposible argumentar otra cosa. Además, observó que el sexo es el punto clave de la discordia más acentuada entre lo que estamos mirando y lo que nuestra cultura nos dice que estamos viendo. Como dijo en el *Boston Globe*, le dejó atónito la vigilancia con la que, incluso entre sus sujetos de estudio invidentes, «se patrullan las fronteras raciales, fundamentalmente en el ámbito de las citas románticas». Para llevar esta metáfora un paso más allá, una patrulla protege el interior, y allí las citas son solo la frontera de una ingente masa cultural que nos costará décadas reorganizar.

Sea como sea, soy muy consciente de la larga y vergonzosa historia de la «ciencia» llevada a cabo por investigadores blancos para «demostrar» la creencia del científico de que las personas blancas son mejores. Y soy igualmente consciente de que nos encontraremos con datos que muestren, por ejemplo, que «las mujeres encuentran atractivos a los hombres blancos». No pretendo decir con eso que los hombres blancos sean inusualmente guapos. Ni tampoco que los datos «demuestren» que las personas negras no son atractivas. De hecho, esos patrones de OkCupid cambian en países que no son Estados Unidos. En el Reino Unido, los usuarios negros del sitio obtienen un 98,9 por ciento de los mensajes que reciben los blancos. En Japón, un 97,8 por ciento. En Canadá, un 90 por ciento. Muchos de los usuarios negros de estos últimos dos países, sobre todo los de Japón, son estadounidenses expatriados.

El sexo a veces no tiene nada que ver con la estructura ósea, los músculos y la carne, cuyos defectos y virtudes comparten todas las razas en igual cantidad. También está la cultura, y las expectativas, y los condicionantes. Eso es lo que muestran mis datos, y puesto que son de persona a persona y se han recopilado hasta en el más mínimo detalle, pueden mostrarlo de un modo imposible de conseguir mediante ninguna otra investigación.

Cuando iba al instituto estuve un verano como estudiante de intercambio en Japón y los funcionarios de la agencia de mi ciudad de acogida, Utsunomiya, nos recogían de vez en cuando a mí y a otros estadounidenses para ir de visita a un colegio o una fábrica cercanos. El objetivo era tanto que

viésemos el país como que el país nos viese a nosotros. Era a principios de los noventa, en la era preinternet, y Japón, y no China, era nuestro gran rival económico. Había tensión: los japoneses habían comprado el Rockefeller Center unos años antes y el yen amenazaba al dólar. El nombre de mi programa de intercambio resumía el tono de la visita en cuatro palabras: «Jóvenes para el entendimiento».

A pesar de aquel nombre, su cultura me desconcertó. Hasta recuerdo que los nombres del *Street Fighter II* estaban todos mal: Vega se llamaba Balrog, y Balrog era M. Bison... Aquello era de locos. Pero tenían televisión americana; *Los vigilantes de la playa* se convirtió enseguida en la serie número uno del país. En un colegio al que nos mandaron tuvimos que ponernos en pie y decir unas palabras ante toda la asamblea de estudiantes. Yo me levanté del suelo y subí al estrado, dije cualquier tontería y volví a bajar. La siguiente a la que le tocaba era la única rubia que había en nuestro grupo y, cuando se levantó, nunca lo olvidaré, se oyó a la sala contener la respiración. La persona que estaba allí de pie era una chica normal —teníamos 16 años, así que todos éramos feos y desgarbados— pero un estremecimiento recorrió la multitud como si tuviesen allí delante a Pamela Anderson en carne y hueso.

Mucha gente ha experimentado ese estremecimiento. Y, durante décadas, frenólogos, racistas y charlatanes han hecho malabarismos para darle a esa reacción esencialmente cultural (y por tanto inmutable) un fundamento biológico. El libro *The History of White People*, de Nell Irvin Painter, hace un excelente repaso de la «ciencia racial» y en sus páginas la autora nos proporciona una cita de un texto de la Ilustración sobre las maravillas de la raza «caucásica», escrito, por supuesto, por un hombre blanco:

La sangre de Georgia es la mejor del Este, y tal vez del mundo entero. No he observado ni un solo rostro feo en ese país, de ningún sexo, pero los he visto angelicales. La naturaleza se ha prodigado allí en las bellezas femeninas, que no se encuentran en ninguna otra parte... Sería imposible hallar rostros más cautivadores ni figuras más bellas que las de los georgianos.

Johann Blumenbach escribió esto. Desarrolló sus teorías raciales mediante la recolección y la comparación de cráneos humanos. Parece que la erudición ha avanzado algo. En cuanto al subconsciente, eso ya es otro cantar.

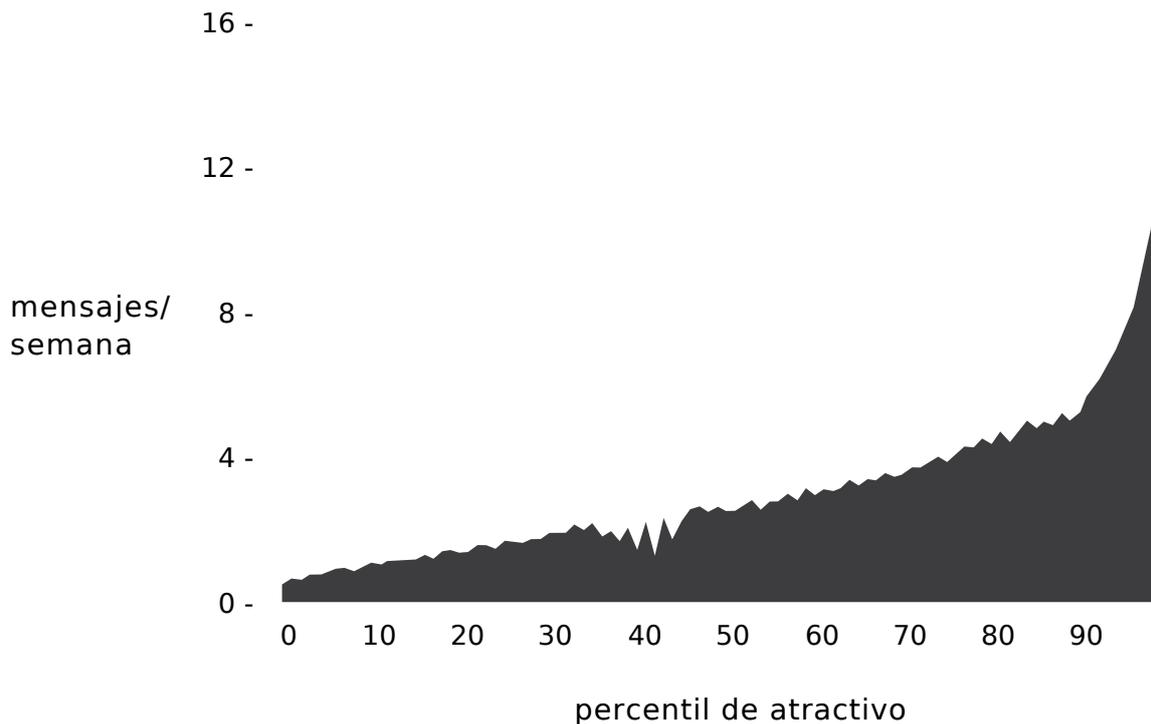
7.

El mito de la belleza en la
apoteosis

Trabajo en un universo en el que la gente se identifica de todas las maneras posibles: como fumadores y no fumadores, como cristianos y ateos, como nerds, geeks y frikis, por no hablar de blancos y asiáticos o gais y heteros, o ninguno, o ambos. La especie humana son tribus dentro de tribus. O dicho con más gracia en un proverbio coreano: «Detrás de las montañas, montañas». Eso responde tanto a lo agreste de su península como a la interminable orografía de nuestro resquebrajado terreno humano.

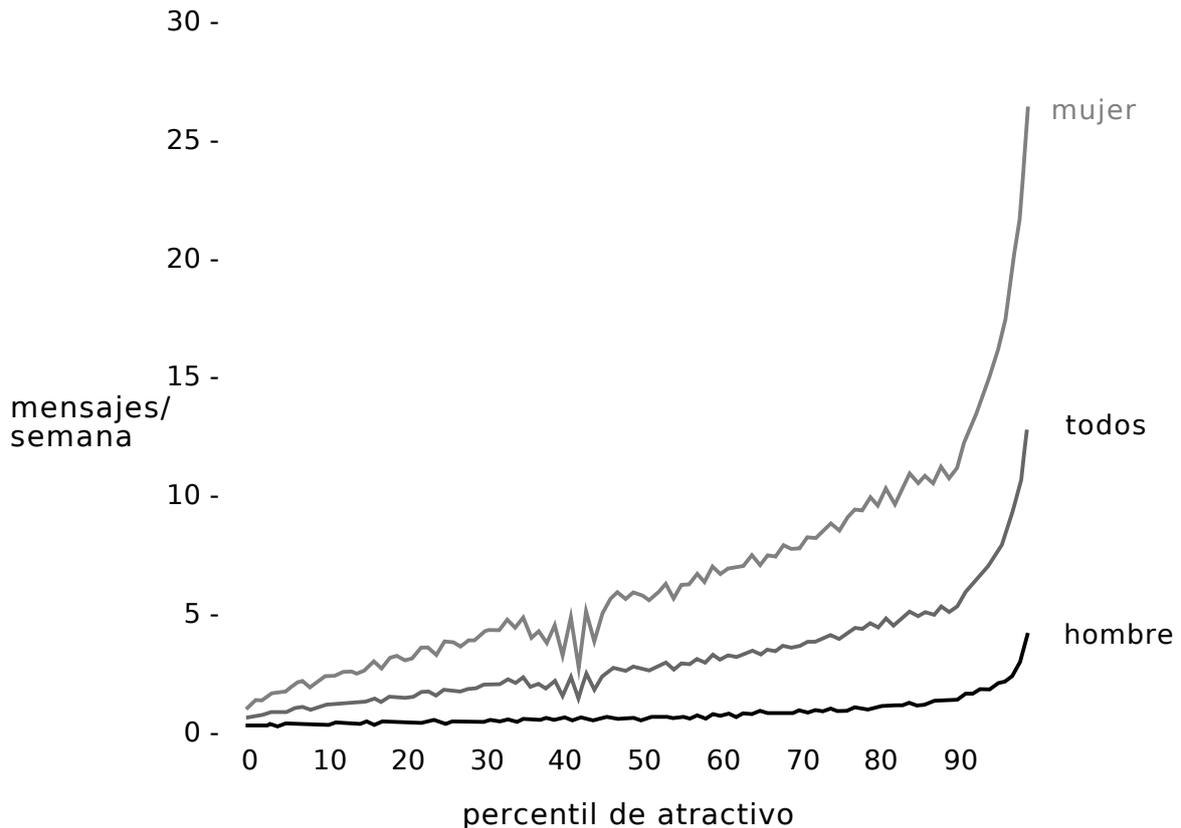
Cuando diriges una web de contactos te das cuenta de que hay una subdivisión que por una parte parece frívola, pero por otra es tan innata como la raza o la sexualidad de las personas. Y, como estas últimas, suele resistirse al análisis directo. En OkCupid —como en Match y en Tinder—, se da una primera división, acaso la más profunda, entre los guapos y los demás. Estas son nuestras clases altas y bajas, nuestros ricos y nuestros pobres, y en lo que se refiere a la atracción sexual, los ricos cosechan los frutos de su herencia igual que cualquier heredero, mientras que los pobres pasan prácticamente sin nada. De forma parecida a la raza, la belleza es una carta que te toca en el reparto, y tiene repercusiones enormes.

Abajo he representado gráficamente los nuevos mensajes recibidos por semana según el atractivo físico del receptor:



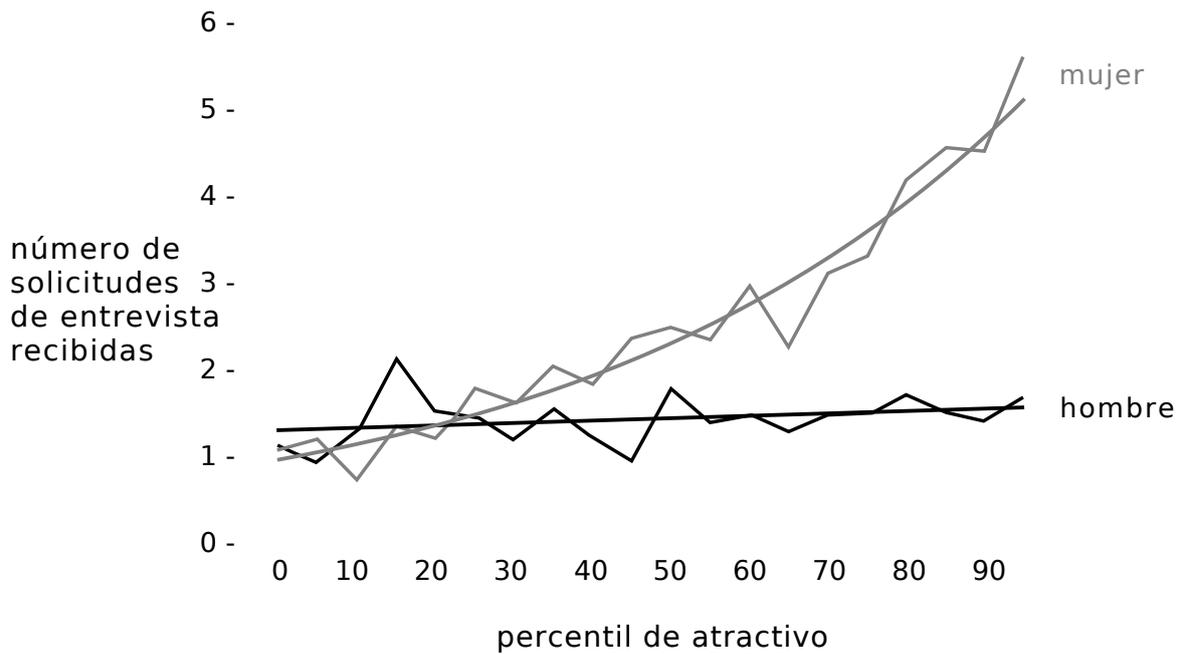
El acusado incremento que se ve a la derecha desmerece un poco al resto de la curva, y su verdadera naturaleza es algo confusa; pero desde el percentil más bajo en adelante, vemos una función que es prácticamente exponencial. Es decir, que obedece a los mismos cálculos que emplean los sismólogos para medir la energía liberada por los terremotos: la belleza opera según una escala de Richter. En lo que respecta a sus efectos, no hay mucho que destacar entre, digamos, un 1.0 y un 2.0: ambos causan temblores que solo varían en su grado de imperceptibilidad. Pero en el extremo superior de la escala, una mínima variación genera el impacto de un cataclismo. Un 9.0 es intenso, pero un 10.0 puede desgarrar el mundo. O lanzar mil cohetes.

Lo que seguro que no se aprecia en el gráfico anterior, porque he juntado todos los datos para ocultarlo, es que hombres y mujeres experimentan la belleza de manera distinta. Este es el mismo gráfico de densidad de mensajes de OkCupid, pero desglosado por géneros, con la media en forma de línea punteada en el centro:

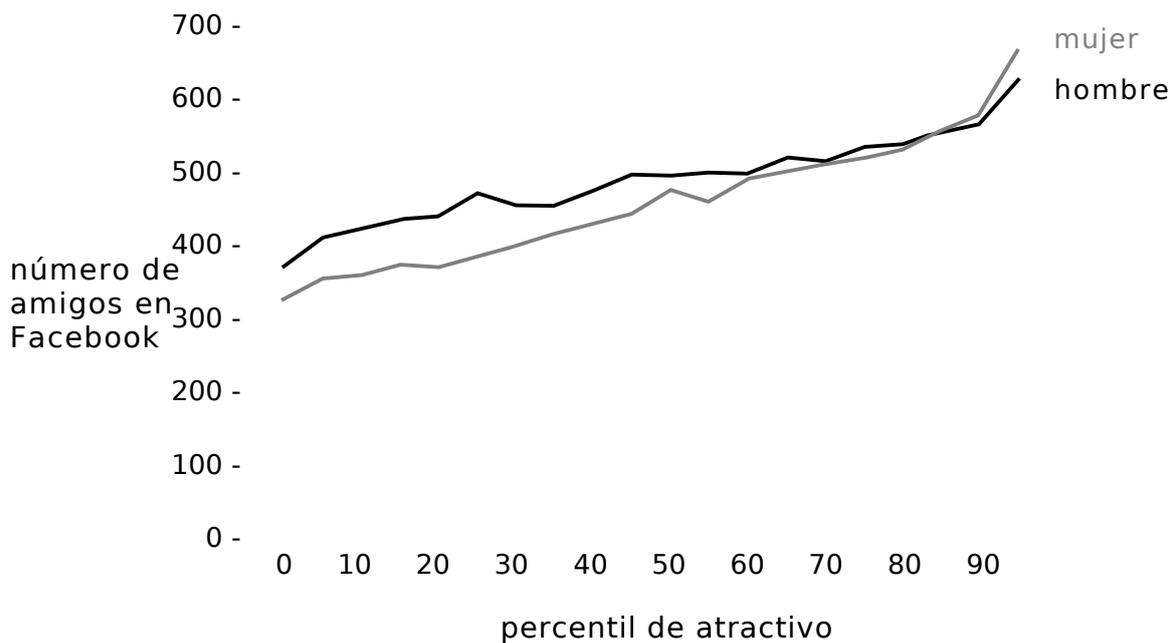


Me cuesta transmitir cuánta atención me merece el extremo superior derecho de esta curva, a menos que me dedicase a perseguirte y a darte la paliza a grito pelado sobre mis *hobbies*. Sobre todo en las grandes ciudades, donde el flujo de mensajes es incluso un 50 por ciento más elevado que lo que ves aquí, una mujer que está en lo alto de la escala se encuentra toneladas de «Hola, ¿qué tal? ¿Te gustan las motos? A mí me encantan las motos» cada vez que entra en la web. Un «plastaclismo», podríamos llamarlo. No obstante, ni los efectos de la belleza ni la diferencia entre hombres y mujeres están confinados al ámbito sexual.

Estos datos son de solicitudes de entrevistas extraídos de Shiftgig, un portal de búsqueda de empleo para el sector servicios^[16]:



Y aquí, los datos del número de amigos en Facebook:



El éxito y la belleza están correlacionados en ambos sexos, pero se aprecia claramente que la pendiente de la línea gris es más acusada. En Facebook, cada percentil de atractivo de un hombre le granjea dos nuevos amigos. A la mujer, tres. En Shiftgig, las curvas no se pueden comparar del mismo modo. La curva femenina es exponencial y la masculina, lineal. Y, lo que es más, varían según si el encargado de contratación, la persona que hace la

entrevista, es hombre o mujer. En ambos casos, la curva de los candidatos masculinos es una línea plana —el aspecto físico de un hombre no afecta a sus posibilidades— y la femenina es exponencial. Es decir, que a estas mujeres se las trata como si estuviesen en OkCupid, aunque se estén presentando a un puesto de trabajo. Los encargados de selección de personal sopesan la belleza de las candidatas como lo harían en un contexto romántico, lo cual es o deprimente o muy, muy emocionante, dependiendo de si eres o no un abogado especializado en ese tipo de demandas por discriminación. Y si la contratadora es mujer lo ve a través de la misma lente (aparentemente sexualizada), pese a que (habitualmente) no albergue propósito sexual alguno.

No es que sea ninguna novedad que la belleza importa, y que a las mujeres les importa más. Por ejemplo, uno de los artículos fundacionales de la psicología social se titula «What Is Beautiful Is Good» («Lo bello es bueno»). Fue el primero de una larga lista de estudios para establecer que a la gente apuesta se la considera más inteligente, más competente y más fiable que a los demás. La gente más atractiva consigue mejores trabajos. También se los absuelve más en los juicios y, en caso de que no sea así, se les imponen condenas más suaves. Como señala Robert Sapolsky en el *Wall Street Journal*, dos neuropsicólogos de Duke están trabajando en el porqué de eso: «La corteza orbitofrontal medial interviene a la hora de valorar tanto la belleza de un rostro como la bondad de un comportamiento, y el nivel de actividad en esa región mientras se produce una de esas tareas predice el nivel que se producirá durante la otra. Dicho de otro modo, el cerebro [...] asume que los pómulos te están diciendo algo sobre el intelecto y la emoción». A nivel neurológico, el cerebro registra ese toque de atracción sexual —«¡Oh, qué buena está!»— y todo lo demás pasan a ser daños colaterales.

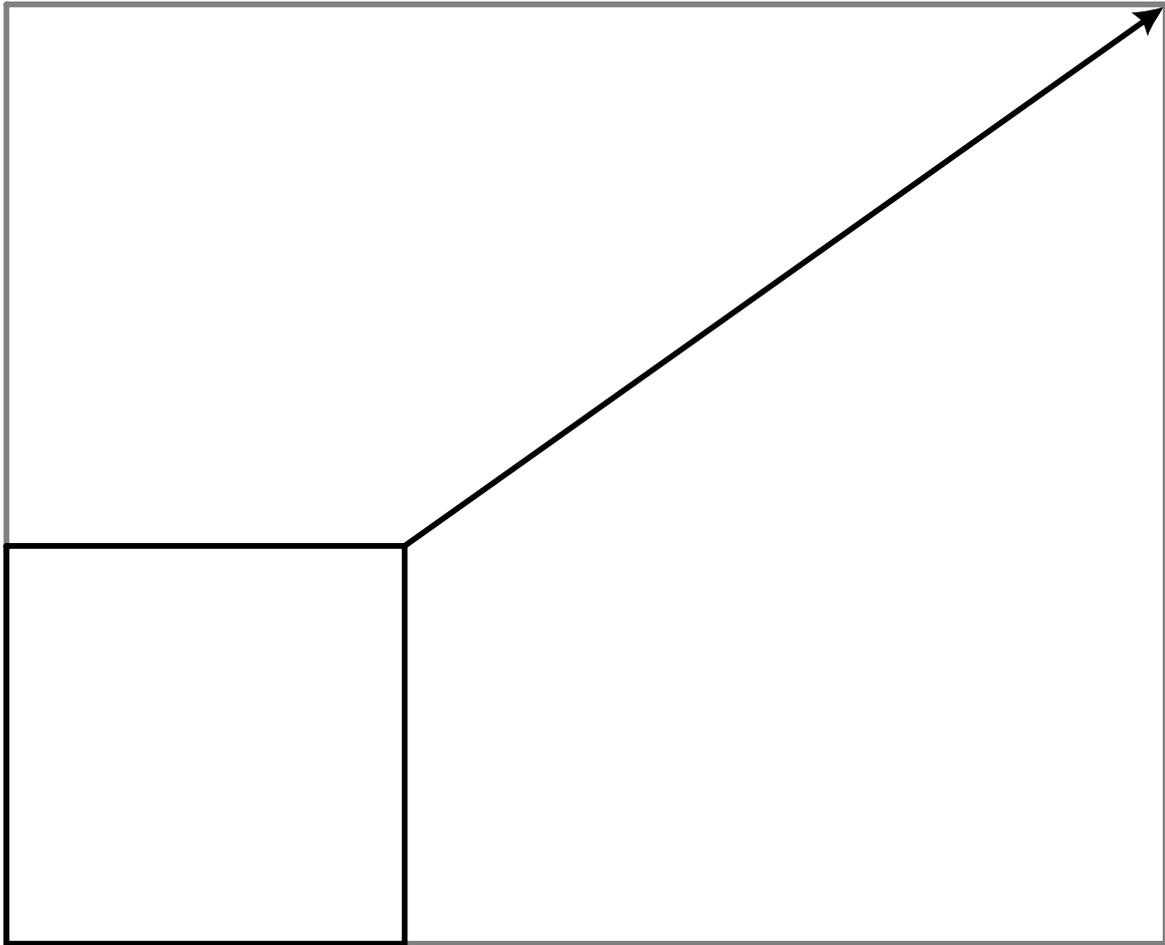
En cuanto a mi segunda afirmación, que la belleza afecta especialmente a las mujeres, el *best seller* de Naomi Wolf *El mito de la belleza* lo demostró mejor de lo que yo podría hacerlo. Estos descubrimientos míos no son nuevos. Lo que es nuevo es nuestra capacidad para poner a prueba ideas establecidas, incluso ideas famosas, a través de las acciones atomizadas de millones de personas. Ese grado de detalle aporta fuerza y matices a los trabajos previos e incluso sugiere nuevas maneras de trabajar a partir de ellos.

El artículo «Lo bello es bueno» se basó en una muestra de solo 60 sujetos de estudio, algo que a duras penas parece suficiente para demostrar el efecto, por no hablar de sus muchas facetas^[17]. Pero ahora podemos pasar del «Lo bello es bueno» a preguntarnos «¿Cuán bueno es?» y en qué contextos. En el sexo, la belleza es buena. En la amistad, solo es un poco buena, y cuando

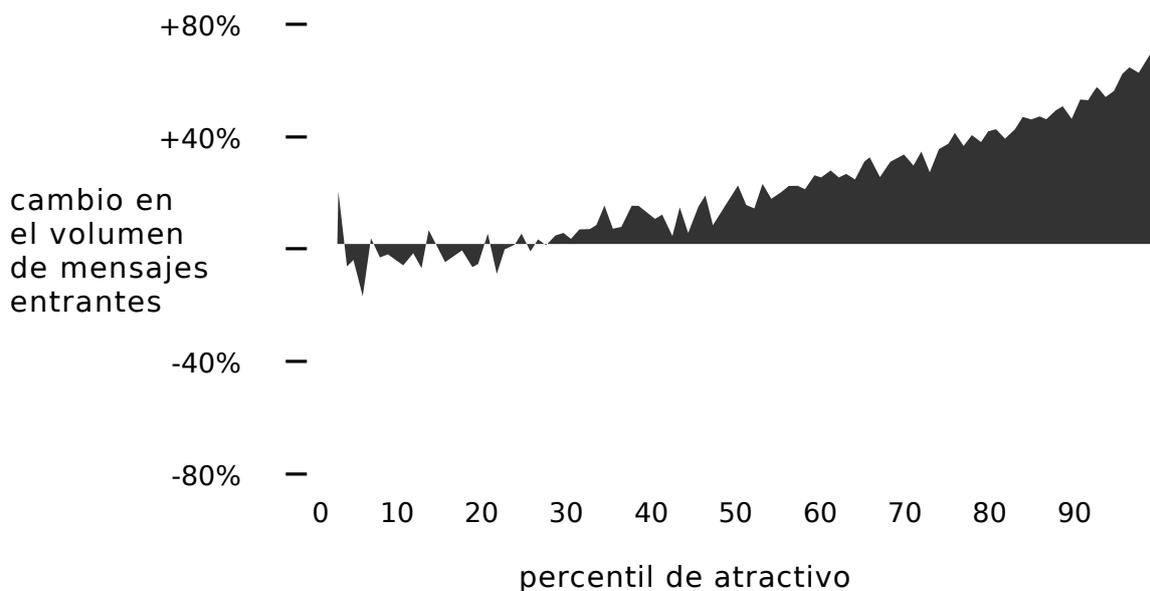
buscas trabajo, el efecto depende en realidad de cuál sea tu género. En cuanto a la obra seminal de Wolf, podemos confirmar la verdad que yace tras su vaga observación de que «la mujer de hoy se ha convertido en su “belleza”»: tres sólidos grupos de sujetos de estudio demuestran que la correlación es muy intensa. Y, mejor aún, podemos ampliar algunos de sus argumentos más convincentes sobre que la belleza constituye una forma de control social. Piensa en cómo cambian los datos de Shiftgig nuestra manera de entender la percepción del rendimiento de la mujer en el puesto de trabajo. Es obvio que las buscan (y de manera exponencial) por un atributo que no tiene nada que ver con su capacidad de desempeñar bien un trabajo. Entretanto, a los hombres no se les impone esa exigencia. Así, es sencillo prever que las probabilidades de fracaso de las mujeres sean más elevadas. Y lo crucial es que la culpa es de los criterios, no de las personas. Imagina que a los hombres, independientemente del trabajo, se los contratase por su fortaleza física. Acabaríamos viendo, por definición, a hombres fuertes enfrentarse a desafíos en los que la fuerza no tiene nada que ver. Del mismo modo, contratar a mujeres basándose en su aspecto es garantizar (estadísticamente) un mal rendimiento. O eso, o les limitas sus oportunidades. Por eso, según Wolf: «El mito de la belleza siempre dicta el comportamiento y no la apariencia». Se refería básicamente al ámbito sexual, pero aquí vemos cómo esto se produce, con equivalencia matemática, en el trabajo.

Como ya he mencionado antes, tengo una hija pequeña, y en nuestros escasos momentos de respiro, Reshma y yo especulamos sobre su vida y adónde podría llevarla. Todos los padres lo hacen: dales un momento e inevitablemente lo harán, del mismo modo que dos borrachos en un bar siempre acabarán discutiendo. Cada familia tendrá sus momentos de sueños e ilusiones particulares, pero los nuestros son más o menos los normales, me imagino. Mi mujer o yo, en realidad no importa, empezamos diciendo: «Nuestra niñita va a ser muy lista. Oh, sí, le vamos a enseñar todo lo que podamos. Será amable, de buen corazón». Estas cosas son muy importantes en la vida, en eso estamos de acuerdo. «Y luego, claro, mira qué piel tiene, y qué ojos; será preciosa. Es que mira, uau. Sí, tendremos que poner candados en las puertas cuando llegue a la adolescencia». Y ahí la conversación da un pequeño giro. «Pero que no sea demasiado guapa, eh... Sí, tampoco queremos eso». Volvemos los dos a apoyar la espalda en la silla y la conversación pasa a otros temas. Y todo se reduce a esto: no soy capaz de imaginar a nadie que le desee ninguna limitación a su hijo.

Por desgracia, este es un problema que Internet, con toda certeza, está haciendo que empeore: según *El mito de la belleza*, las redes sociales señalan el Día del Juicio Final. Tu foto aparece prácticamente en todas partes, en tu currículum, en cada aplicación, en cada cosa que firmas. Si alguien se interesa por lo que haces, acabará averiguando qué aspecto tienes. No porque deban hacerlo, sino porque pueden hacerlo: Facebook y LinkedIn básicamente han ampliado el problema del «amor es ciego» de OkCupid al máximo. No hace ni diez años era casi imposible vincular el nombre de una persona con su fotografía; ahora solo tienes que meter el nombre en Google —todo el mundo lo hace— y te aparece una foto en miniatura de alguna red social. Todos hemos tenido que escoger entre varias fotos para decidir cuál es la «mejor». Pues escoge con cuidado, amigo, porque eso te va a definir como nunca nada te ha definido. Esta tendencia está experimentando un subidón que puede no parecer tan obvio a quienes trabajan fuera del sector. El nuevo estándar de diseño de los últimos dos o tres años, más abierto y más fotocéntrico —lo que yo denomino «pinterístico»— está haciendo que no solo las fotos, sino concretamente la belleza, sea lo más importante. En OkCupid hemos cambiado algunos formatos de las fotos, pasando del tamaño del recuadro negro al del gris, como vemos:



Los diseñadores solo querían que la página se viese más moderna. Lo que no previeron (y después tuvieron que paliar) fue esto: todos esos píxeles adicionales hacían que las caras bonitas eclipsasen todavía más al resto. Los ricos, aún más ricos. Fue el equivalente en el diseño web a la política nacional estadounidense.



Vista esta presión, no sorprende que los blogs sobre imagen corporal estén tan en boga. Ni que etiquetas que promueven la delgadez extrema, como #thinspiration #thinspo #loseweight #keeplosing #proana #thighgap, se hayan hecho tan habituales que tanto Tumblr como Pinterest (que son independientes uno del otro) tuvieran que alterar sus Condiciones de servicio para prohibir este tipo de contenidos. Si te estás preguntando qué quieren decir esos dos últimos hashtags: #proana es la abreviación de «pro anorexia», gente que está a favor de morirse de hambre como método de adelgazamiento; #thighgap hace referencia a tener unos muslos tan delgados que no se toquen cuando estás de pie con las rodillas juntas. Este último atributo es una especie de fijación entre las adolescentes y, aparte de lo cuestionable que sea querer eso, para la mayoría de ellas resulta biológicamente imposible. No eres consciente de la total depravación de este fenómeno hasta que haces una búsqueda de estas etiquetas y te enfrentas a una interminable página de fotos de cuerpos quebrados ladeándose ante la cámara: no solo las mujeres «inspiradoras» están mortalmente delgadas, sino que con frecuencia posan en ropa interior, bikini o lencería. Los blogs, creados por mujeres, son verdaderamente el paradigma de la mirada masculina —y esto lo digo como persona que se toma con reflexivo escepticismo el lenguaje de la izquierda intelectual.

Que tanto Tumblr como Pinterest prohibieran ese contenido no resolvió nada, por supuesto, y menos aún los problemas de imagen corporal de sus usuarios, de modo que ambas plataformas han optado por otra estrategia. Puesto que en estos blogs se emplean etiquetas, han logrado intervenir en los

algoritmos: si buscas «thighgap» en Tumblr, la pantalla se pone en blanco y aparece el siguiente mensaje:

«... Si padeces algún trastorno de la alimentación o conoces a alguien que se encuentre en esta situación...».

Y le siguen varios enlaces a páginas de ayuda. Es una medida pequeña, pero antes de que se digitalizase el comportamiento no había prácticamente manera alguna de atacar directamente este problema, al menos no hasta que se había producido un perjuicio visible. Solo existían rumores, alguien que había pegado la oreja a la puerta del baño, las tristes sospechas de un padre o una madre... Los datos nos dicen lo que de verdad estamos sintiendo; lo que sentimos unos por otros, sí, pero también lo que sentimos por nosotros mismos. Si son capaces de encontrar divisiones en nuestra cultura, nuestras opiniones políticas, nuestros hábitos o nuestras tribus, también las encuentran en nosotros mismos. Y ese es un pensamiento esperanzador, porque para lograr que algo sea pleno, el primer paso es saber qué es lo que falta.

8.

Lo que cuenta es el interior

Antes había dos maneras de averiguar lo que de verdad pensaba una persona. La primera, pillarlo en un momento en el que tuviese la guardia baja. Cotilleabas, provocabas, elaborabas cualquier pretexto en un laboratorio, hacías lo que pudieras para que tu presa olvidase que estaba siendo observada. Ese tipo de investigación debía de ser muy divertida —una bata de laboratorio, una cámara oculta... quién sabe, hasta un bigote postizo—, pero era imposible llevarla a cabo a gran escala. Así que para obtener datos en masa te quedaba la segunda opción: hacer una pregunta y confiar en que te dieran una respuesta sincera. Esa ha sido la norma estándar desde que Gallup fundó el American Institute of Public Opinion en 1935.

Por desgracia, las encuestas han sido históricamente incapaces de sacar a la luz actitudes sinceras sobre temas como la raza, la conducta sexual, el consumo de drogas e incluso las funciones corporales, porque los encuestados editan sus respuestas. Los datos obtenidos de la observación del comportamiento son muy útiles, como ya hemos visto. Pero hay determinadas cosas —pensamientos, creencias— que no conllevan una acción explícita. Y muchas veces las actitudes más feas y que causan más desencuentros se ocultan tras un telón de ego y de normas culturales que es prácticamente imposible de descorrer, al menos mediante preguntas directas. Es una maldición que sufre el científico social: aquello a donde quieres llegar es exactamente lo que tus encuestados esconden con más ahínco. Esta tendencia recibe el nombre de inclinación a la deseabilidad social, y está muy bien documentada: por todo el mundo, los encuestados responden a las preguntas para quedar bien. El caso más famoso fue el del llamado efecto Bradley: en 1982, los votantes de California dijeron en las encuestas a pie de urna que habían votado a un gobernador negro, Tom Bradley, por un margen significativo, pero en la intimidad de la urna le habían brindado a su oponente blanco una ajustada victoria. A lo largo de las décadas de 1980 y 1990, los candidatos negros muchas veces recibían más apoyos en las encuestas que en las elecciones. Hay otros problemas, aparte del racismo, como la depresión o las adicciones, que son igualmente difíciles de diagnosticar a nivel social porque a la gente le cuesta ser sincera. Incluso en las preguntas que se hacen

en OkCupid para encontrar coincidencias con otros usuarios —que en líneas generales no las ve nadie salvo el usuario en cuestión— los usuarios son reacios a reconocer determinadas actitudes, incluidas aquellas que han puesto en práctica en otras partes de la web. El propio acto de preguntar hace que aparezca la autocensura. Prácticamente todos los sitios web en los que se registran opiniones o se recopilan datos descriptivos sufren el mismo problema. Pero hay un lugar en el que no hace falta preguntar nada y donde, por tanto, los datos se liberan: en las búsquedas no se te pregunta nada, lo dices tú todo.

La única invitación a interactuar de Google consiste en su famosa página casi en blanco con su solitario campo de texto, ese pequeño rectángulo vacío, con su cursor parpadeando y preparado, a la espera de tus pensamientos. La actividad de la empresa consiste en ayudar a la gente a encontrar cosas en la vasta espesura de Internet, y eso lo ha hecho de manera espectacular. Pero ese éxito mundial de Google lleva aparejado, casi como un añadido no previsto, el hecho de que con cada usuario que introduce un nuevo deseo en su base de datos se haya ido convirtiendo en un almacén de la identidad colectiva humana. Escucha nuestras confesiones, nuestras preocupaciones, nuestros secretos. Es nuestro médico, párroco, psiquiatra, confidente y, sobre todo, Google no tiene que preguntarnos nada, porque la pregunta está siempre implícita en el campo en blanco de su interfaz: «¿Qué estás pensando?». Ahab y su ballena, Arturo y su grial. Lo que una persona está buscando muchas veces te entrega a la propia persona. Pero hasta ahora el truco era: ¿cómo logramos ver esa búsqueda?

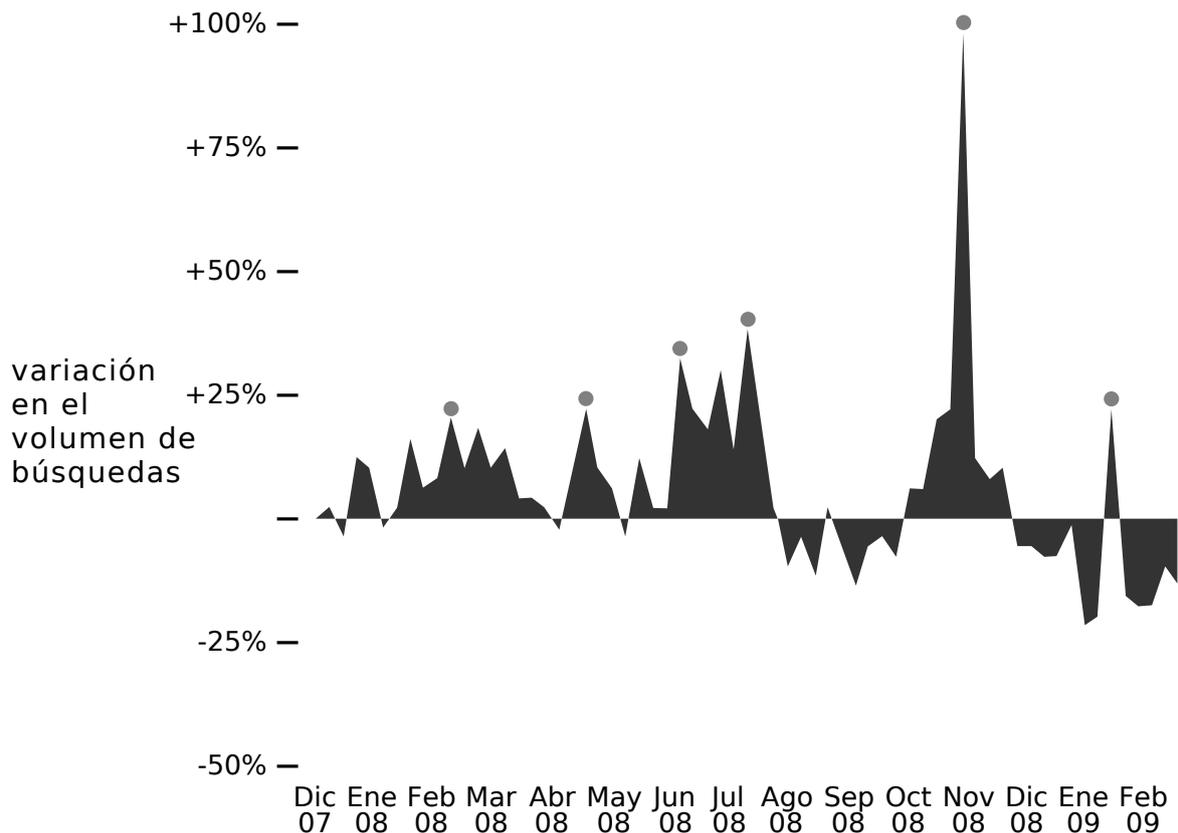
Desde 2008, Google nos ha proporcionado esa información mediante su herramienta de Tendencias de búsqueda. Permite que cualquiera interroge a su base de datos de búsquedas agregadas y, con la sintaxis adecuada y un poco de tabulación cruzada, se puede emplear para extraer una excelente muestra de cualquier mente individual, de los mecanismos internos que hasta ahora habían quedado fuera del alcance de la investigación, desde que esta existe. Desde que se hizo público este servicio, los científicos han usado Google Tendencias para hacer predicciones en bolsa, para sacar a la luz impulsores de actividad económica (a los países más ricos les preocupa más el futuro que el pasado) y, por lo que es más famoso, para llevar un seguimiento de las epidemias de gripe y dengue en tiempo real —y, por tanto, atajarlas con la misma rapidez—. Cuando la gente enferma, busca síntomas y remedios. La página de Evolución de la gripe de Google detecta cómo está la

enfermedad y alerta al CDC (centro de control y prevención de enfermedades de Estados Unidos).

El sitio registra también otro tipo de virulencias. Puesto que no se pregunta nada y que, a diferencia de las redes sociales, no hay nadie al otro lado de la línea, la gente da rienda suelta a sus impulsos más viles en Google. La palabra despectiva «nigger», por ejemplo, es un término que aparece habitualmente en 7 millones de búsquedas al año. En Estados Unidos, el volumen de estas búsquedas es más alto donde cabría esperarlo —Virginia Occidental—, pero se mantiene uniforme en todo el país. Brooklyn tiene pocas cosas en común con la ciudad en la que me crié, Little Rock, pero esta es una de esas cosas: «nigger» es tan común en Nueva York como en el centro de Arkansas, y tan frecuente en Chicago como en Fresno^[18]. A juzgar por el volumen de búsquedas, el término es literalmente más norteamericano que «apple pie», por una ventaja del 30 por ciento. Y resulta esclarecedor que aparezca con mucha más frecuencia en Google que lo que lo hace en un sitio que para la psique es mucho más público, Twitter. Si empleamos «nigga» como término de control, dado que en significado es similar pero carece del mismo bagaje, «nigger» aparece unas 30 veces más asiduamente en las búsquedas que en las redes sociales.

A diferencia de lo que ocurre con los acusados ciclos de las enfermedades, el racismo sigue un curso lento y laborioso —actúa a nivel generacional, no metabólico— y es uno de los pocos ámbitos en los que podemos empezar a vislumbrar las amplias posibilidades longitudinales de los datos. Por otra parte, si vinculamos los flujos y reflujos de las búsquedas a acontecimientos reales podemos disipar parte de las sombras emocionales que se ocultan tras los datos. Por ejemplo, si representamos en un gráfico las búsquedas de la palabra «nigger» a lo largo del ciclo de la campaña electoral estadounidense de 2008, podemos observar cómo se enfrentó el país a la perspectiva de tener un presidente negro.

variación en las búsquedas en Google de «nigger», diciembre 2007-febrero 2008



Si nos fijamos en los seis picos marcados con gris, de izquierda a derecha vemos: el «Supermartes» del 5 de febrero, seguido de las primarias de Pennsylvania del 22 de abril, que tantas protestas suscitaron; el 6 de junio las búsquedas experimentan un nuevo pico, Hillary suspendió su campaña y Obama obtuvo la candidatura; el 15 de julio, para complicar algo los datos (y también el debate moral), el grupo Nas sacó un disco cuyo título ofensivo fue *Nigger* y que llegó al número uno. Pero incluso pese a la interferencia causada por ese acontecimiento, el volumen general de búsquedas se desplomó al consolidarse el predominio de Obama. La tensión racial, y hasta la política, se disipó cuando los candidatos, ninguno oficial todavía, se posicionaron para el otoño que les esperaba. De hecho, el volumen de búsquedas con carga racial alcanzó su índice más bajo de toda la campaña la semana de la Convención Nacional Republicana de principios de septiembre^[19].

No obstante, tras haber alcanzado ese mínimo, los ánimos volvieron rápidamente a su cauce normal, para acabar explotando en la misma noche electoral, cuando las búsquedas de «nigger» llegaron a un nivel nunca antes igualado. Al día siguiente, cuando el país se levantó con la confirmación que tenía un presidente negro, cerca de una de cada cien búsquedas del nombre

«Obama» incluía también el término «KKK» en el campo de búsqueda. Pero casi inmediatamente después el volumen de estas búsquedas con carga racial cayó en picado y, salvo por un último estertor de indignación en la ceremonia de inauguración, ese nivel se ha mantenido (un 25 por ciento por debajo de los niveles pre-Obama). Se oye mucho hablar de nuestro «debate nacional» sobre la raza, pero si observas los datos ves que se trata más de una serie de convulsiones nacionales. También ves, sin embargo, que pese a todas las promesas fallidas de su famoso lema, Obama sí cambió el curso del epíteto favorito de nuestro país:

índice de «nigger» en las búsquedas de Google, por fecha



De hecho, se han producido tres verdaderos saltos en las búsquedas de «nigger» durante la presidencia de Obama. El primero lo originó una de esas marcianadas en las que parecen especializarse los políticos del Tea Party: el volumen de búsquedas alcanzó un pico en octubre de 2011, la semana en la que el mundo descubrió que el gobernador de Texas, Rick Perry, tenía una finca llamada «Niggerhead Lake». Los otros dos picos, ambos comparables en altura y brusquedad al de la noche de la victoria electoral de Obama, fueron dos pilares de la misma historia. El primero llegó a los servidores de datos a finales de marzo de 2012, y el segundo la última semana de junio del año siguiente. Coinciden, el primero, con el momento en el que los padres de Trayvon Martin pusieron la muerte de su hijo en el foco de atención nacional, y el segundo, con la presentación de la acusación contra George Zimmerman; puede que sean las dos veces en las que los blancos se han sentido más atacados desde la primera campaña de Obama. No se produjo ningún pico comparable durante la fase del juicio en la que se planteó la defensa, ni tampoco cuando se leyó el veredicto. Y, como ocurrió una vez celebradas las elecciones de 2008, las búsquedas alcanzaron un nuevo mínimo justo después de la absolución, lo que mostró una vez más el ciclo de tensiones y catarsis que caracteriza a las relaciones raciales en Estados Unidos.

Cuando uno busca palabras con carga racial, «nigger» es el punto de partida más obvio, pero enseguida se da uno cuenta de que no hay mucho más significado que buscar ahí fuera: es verdaderamente el alfa y el omega del discurso del odio. Otros términos peyorativos como «spik» (hispano) y

«chink» (oriental) se usan tan rara vez que, comparativamente, existen pocos datos que analizar. No son los epítetos mismos los que más significado contienen, no obstante, sino la mentalidad que conllevan, una certeza que puede observarse si nos fijamos en cómo cambia la carga del término «nigger» según la identidad del hablante. Si, en lugar de Nas, hubiese sido Toby Keith el que lanzó aquel álbum en 2008, tendríamos entre manos una historia muy distinta. La opción de autorrelleno de Google es muy útil para esto: proporciona pensamientos completos más que una palabra fuera de contexto.

Si no estás muy familiarizado con esa función de autorrelleno, cuando empiezas a escribir una frase, por ejemplo: «Quién es el...», Google se ofrece a terminar tu pensamiento con el texto de otras búsquedas habituales. Si escribes «Quién es el...» te sugerirá «... hombre más rico del mundo». Si trasteas un poco con esta función obtendrás una fugaz mirada a lo que media humanidad se pregunta acerca de cómo vive la otra media.

¿Por qué las mujeres...
... viven más que los hombres?
... gimen?
... van juntas al baño?
¿Por qué los hombres...
... no escuchan?
... entienden los mapas?
... mienten?

Y si te dedicas a buscar estereotipos, es como jugar al juego Tabú, pero sin ningún tabú. ¿Por qué a los negros... les gusta el pollo frito? ¿Por qué los musulmanes... odian a América? ¿Por qué los asiáticos... se parecen? La opción de autorrelleno te regala todo tipo de cosas: esos son todos ejemplos verídicos. De hecho, uno de los resultados, el de «¿Por qué los blancos tienen los labios finos?», es el título de un reciente artículo de investigación que ahonda en el doble propósito al que sirve esa función: desvela tendencias, por supuesto, pero dada la ubicuidad de Google también tiene el poder de consolidarlas. El artículo sugiere que la función de autorrelleno acabará perpetuando los estereotipos que debería limitarse a reflejar, y no cuesta entender por qué: un usuario teclea una pregunta sin relación con el asunto y lo que ocurre es que le salen al paso los prejuicios de otras personas. Por ejemplo: «¿Por qué las parejas... gays se parecen?» no era para mí un

estereotipo hasta ahora mismo. Es Google, que actúa no ya de Gran Hermano, sino de Hermano Mayor, y te pervierte dándote cigarrillos mentales.

Si te fijas en los entresijos de esta función obtienes otra visión de la humanidad. Es como ponerse al lado de alguien delante del espejo de su baño. Prueba a escribir en el campo de búsquedas:

«¿Por qué a...» y luego
«¿Por qué b...», y así sucesivamente,

y Google completará tus preguntas con un alfabeto entero de problemas, incluidos algunos tan brillantes como estos:

por qué salen las verrugas
por qué tengo tantos gases
por qué un hombre no eyacula
por qué vomitamos

Dolencias que, debo señalar, son probablemente fruto de estar demasiado rato sentado delante de un ordenador.

Así que, gracias a toda esta confusión, nuestros pensamientos están pasando a formar parte del mundo. Con un poco de tecleo creativo, dándole unas vueltas y aplicando un poco de matemáticas, le estamos dando al monólogo interno de la humanidad un público más amplio. Sacamos a relucir las partes hirientes de nosotros mismos y también las ridículas, y los datos de búsqueda proporcionan a esos impulsos dañinos una muy necesaria exposición pública. Ya no se acepta que digamos cosas racistas en público, pero ahora podemos saber que se siguen diciendo aunque la tendencia a la deseabilidad social nos indique lo contrario. Y, lo que es más, aunque nuestra capacidad de detectar actitudes latentes u ocultas sea nueva, nuestro poder para explotarlas no lo es, lo que hace que todos estos datos sean más importantes si cabe. Dejaré que lo explique el estratega republicano Lee Atwater, que habla sobre la llamada «estrategia sureña» de su partido en una entrevista con el politólogo Alexander P. Lamis. En 1981, cuando era miembro de la administración Reagan, dijo:

Empiezas en 1954 diciendo: «Nigger, nigger, nigger». En 1968 ya no puedes decir «nigger»: te da mala imagen. Te sale el tiro por la culata. Así que acabas hablando de cosas como «integración forzosa», «derechos de los estados» y demás. Y ya

te pones tan abstracto [que] acabas hablando sobre rebajar impuestos, y todas esas cosas de las que hablas son temas de economía, y uno de los subproductos de ellas es [que] a los negros se les perjudica más que a los blancos.

Atwater creía estar hablando de manera extraoficial («A ver, esto no me lo vas a publicar, ¿verdad?»). Los datos de búsqueda implican que ya no tenemos que esperar a que ocurran ese tipo de accidentes para poder examinar la desconexión existente entre el debate público y la conversación privada sobre un tema como el racial. Nos muestran que vamos hacia un mundo mejor. También nos enseñan que todavía nos queda mucho camino por recorrer.

Retomemos el asunto de Obama donde lo dejamos, en el día de la investidura de 2009. Se habló mucho entonces, y en un tono esperanzador, de que Estados Unidos se habían convertido en una sociedad «posracial», lo que tampoco era necesariamente una idea disparatada. En el fondo, esa historia de lo «posracial» fue un intento de extrapolar el triunfo de la campaña de Obama a otros rincones de la vida estadounidense, así como para afirmar que su victoria demostraba que «la raza ya no era un factor» en nuestras vidas.

Pese a esa prometedora posibilidad, Seth Stephens-Davidowitz, de Google, concluyó que la raza de Obama probablemente le había costado de tres a cinco puntos porcentuales del voto popular en 2008, y esos votos perdidos no correspondían a los republicanos, sino a gente que hubiese votado a un demócrata blanco como John Kerry. Esta oscilación del 5 por ciento, a ese nivel, habría alterado bastante más de la mitad de las elecciones celebradas desde la Segunda Guerra Mundial y es un resultado que jamás hubiésemos detectado sin la ayuda de los datos de búsqueda. La idea que se le ocurrió al investigador fue remontarse a antes de que Obama apareciese en la escena política nacional, entre 2004 y 2007, y escarbar en Google Tendencias en busca de actitudes preexistentes de índole racial (de ese modo se evitaba que el propio Obama nublase el panorama). Empleando esos datos para obtener un «índice de la animosidad racial» estado por estado, logró comparar ese índice con los votos totales que obtuvo finalmente Obama y con los resultados previstos de un candidato genérico (es decir, blanco) demócrata (de lo que, por supuesto, se tienen abundantes datos previos). Como cabía esperar, cuanto mayor era ese índice de animosidad, peores resultados obtuvo Obama. He aquí un ejemplo del método, en palabras del hombre que lo puso en práctica:

Pensemos en dos mercados de difusión mediática como Denver y Wheeling (que es un mercado dividido entre Ohio y Virginia Occidental). Kerry obtuvo más o menos el 50 por ciento de los votos en ambos mercados. Basándonos en el gran avance de los demócratas en 2008, Obama debería haber obtenido cerca del 57 por ciento de los votos tanto en Denver como en Wheeling. Pero Denver y Wheeling presentan actitudes raciales distintas. Denver tenía el cuarto índice más bajo de búsquedas con carga racial del país. Obama obtuvo allí el 57 por ciento de los votos, justo como se preveía. Wheeling estaba en el séptimo puesto en cuanto a índice de búsquedas con carga racial del país. Obama obtuvo menos del 48 por ciento de los votos de Wheeling.

Históricamente, un candidato a presidente puede esperar una modesta subida, de unos dos puntos porcentuales, en el voto popular de su estado natal. Debido a la animosidad racial, John McCain consiguió en 2008 en el país entero un incremento mayor que el típico del estado natal de un candidato. Si buscamos pruebas de que ser blanco es una ventaja en la vida estadounidense, aquí las tenemos. McCain era el hijo predilecto del país por la simple razón de que se enfrentaba a un negro.

En mi opinión, Muhammad Ali es uno de los estadounidenses más valientes que ha habido. En 1967, cuando era campeón de los pesos pesados, se negó a servir en Vietnam y no solo lo despojaron de su título, sino que lo inhabilitaron para competir durante tres años y medio. Perdió lo mejor de su carrera y lo sentenciaron a cinco años de prisión (que tuvo que revocar el Tribunal Supremo) debido a sus convicciones. Esa es una posición inimaginable en los líderes políticos de hoy, y no digamos en nuestros deportistas y otros famosos. De Kanye, Glenn Beck, Rachel Maddow o Sarah Palin se obtiene un montón de ira, pero poco sacrificio. Todos podemos tener nuestra opinión sobre la actitud de Ali contra lo de Vietnam —y como hijo de veterano que soy (Huê, 1969), conozco al menos a una persona que no está de acuerdo conmigo—, pero datos como estos nos pueden ayudar a entender por qué la adoptó. Como dijo entonces Ali: «Ningún viet-cong me ha llamado “nigger”», y probablemente tenía razón. Pero imagínate lo que habría ocurrido en el campo de búsqueda de los navegadores si Google hubiese existido entonces. E imagínate cómo debía de ser la desventaja en su estado natal de un hombre negro en aquellos tiempos.

Queda por ver qué derroteros tomarán estas actitudes en lo sucesivo. A pesar de todo lo anterior, Obama venció, y por muy deprimentes que sean

algunas de estas cosas, hay muchas otras que invitan al optimismo; para empezar, no hay pruebas de que la tendencia perjudicase al presidente en 2012, aunque para entonces ya era un valor conocido, tal vez menos «un negro» y más «Barack Obama». Una de las cosas que se pierde en todas las cifras totales de este libro es que a nivel individual, los efectos de todas estas variadas fuerzas sociales en la persona suelen ser muy sutiles. Si nos remitimos a los datos que hemos visto en un capítulo anterior, la experiencia de los numerosos usuarios negros de OkCupid es buena: todos consiguen citas con personas y son rechazados por otras, como cualquier otro. Colectivamente, obtienen más de lo segundo, eso sí. Si vas persona por persona, la experiencia de cada individuo es demasiado pequeña y variada para poder concluir con fundamento que haya sucedido algo «racial». Puede que sea el color de tu piel o puede que seas sencillamente tú. Por otra parte, da risa pensar que un tipo de cara colorada se ponga a buscar «chistes de negros» porque Barack Obama resultase elegido. Pero es mucho menos gracioso cuando eres capaz de ver que son miles y miles los que hacen esa búsqueda. Y es todavía menos gracioso cuando ves el enorme efecto que pueden seguir teniendo esas actitudes privadas, incluso en la vida pública. De ahí lo de la historia de uno solo de nosotros frente a la historia de todos nosotros. Por eso son necesarios este tipo de datos: sirven para zanjar discusiones que nunca podrían zanjarse mediante anécdotas. Proporcionan hechos que es necesario afrontar.

Conozco a gente que solo lee buenos libros, y me refiero a cosas que les recomiendan amigos, profesores, críticos, Amazon. Tiene sentido: leer es lento y el tiempo es muy valioso, así que ¿por qué arriesgarse? Pero ese no es mi estilo. Me gusta la historia y cuando voy a una librería arramblo con un montón de cosas variadas de la estantería de esa sección, a ver qué me encuentro. He leído bastante porquería. Y demasiados libros sobre Napoleón. Pero de los muchos hallazgos felices, mi preferido es *La otra historia de Estados Unidos*. Sí, ya sé que ahora es un clásico, pero eso no altera el hecho de que nunca había oído hablar del libro hasta que lo cogí del estante. Google Libros lo describe muy bien: es una crónica de «la historia estadounidense desde abajo para arriba»; y mientras que la mayoría de los libros versan sobre líderes y grandes acontecimientos, *La otra historia de Estados Unidos* nos muestra los hogares, las tiendas, las granjas, las fábricas y las pequeñas preocupaciones de antaño. La cuestión es que, por mucho que me guste el libro y por mucho que le dé la vuelta por completo a la versión escolar de la historia de Norteamérica, Howard Zinn solo logró contarnos lo que veía, las

acciones observables, las palabras pronunciadas en voz alta. El corazón de hombres y mujeres quedaba fuera de su alcance. La tensión de la crisis de los misiles de Cuba, el aburrimiento de las trincheras, la liberación que representó la píldora... ¿Qué pasaría si hubiésemos tenido los datos que hoy tenemos de todos esos momentos de callada felicidad y de angustia interior? ¿Cuánto habríamos ganado en riqueza de conocimiento?

9.

Días de furia

En Nochevieja, aburrida en su sofá y esperando a que diesen las doce, Safiyyah Nawaz tuiteó una tontería en broma.

\$afiyyah @safiyyahn 

este hermoso planeta tierra tiene ahora 2014 años, alucinante

Consiguió 16 000 retuits, la mayoría en las siguientes veinticuatro horas. Para que nos hagamos una idea, la felicitación de Año Nuevo de Katy Perry a sus 49 millones de seguidores se retuiteó poco más de 19 000 veces. La de Lady Gaga, que aprovechó para anunciar un esperado vídeo, obtuvo 20 000 retuits. Safiyyah Nawaz no es ninguna estrella emergente del pop ni esta es la historia de cómo Twitter permite a cualquier advenedizo desafiar al orden cultural. Si nunca has oído hablar de Safiyyah es porque es una alumna de un instituto de Carolina del Norte cuyo chiste hizo que Twitter explotase.

Probablemente, al principio se trataba solo de gente que no daba crédito, que se preguntaba si aquello iba en serio, pero si te fijas en cómo se fueron sucediendo los tuits durante aquella noche en la que cada persona que retuiteaba se iba separando un paso más de Safiyyah, el ser humano, y cada una de aquellas personas se iba dando cuenta de su ridícula aportación a un fenómeno —al ver que el número de retuits iba creciendo y creciendo—, serás capaz de ver cómo el público digital se convierte en una turba furiosa. En un periquete, los risueños LOL se convirtieron en OMG y luego en WTF, y después el tono empezó a ponerse así:

Cocaine Burger @Cocaine_Burger 
@safiyyahn Suicídase

Rick Huijbers @HARDEBAKSTEEN 
@safiyyahn Suicídase puta imbécil

La cosa, como explicaron en el blog *Gawker*, se salió de madre en cuestión de minutos. Dada la violencia de la reacción, la señorita Nawaz manejó la

situación bastante bien para ser una chica de 17 años, y más tarde resumió aquel clamor a la perfección:

\$afiyyah @safiyyahn 

los jóvenes de hoy se toman muy a pecho lo de la verdadera edad de la Tierra

Nawaz no era consciente de ello, pero en el punto de mira la acompañaba gente famosa. Solo quince minutos después de que hubiese tuiteado su broma, la cómica Natasha Leggero estaba en Times Square, en un programa de la televisión con Carson Daly, cachondeándose de la campaña de relaciones públicas de SpaghettiOs para el día de Pearl Harbor. La marca estaba en el candelero por animar a los ciudadanos a recordar a los caídos comprando espaguetis de lata —sí, hasta esos extremos ha llegado este mundo— y dijo: «Es penoso que la única comida que todavía pueden masticar los supervivientes de Pearl Harbor se burle ellos».

El presentador y la invitada se rieron y pasaron a otros asuntos, sin darse cuenta de que también Natasha había rozado sin notarlo el interruptor de encendido de la máquina de la indignación de Internet. Y esta se puso en marcha con toda su rectitud. Leggero publicó más tarde en Tumblr varios ejemplos selectos de los tuits que recibió:

Mike Oswald @SDPStudio 

@natashaleggero Pero que puta vil eres.

Mark Tichenor @hotrod607 

@natashaleggero Zorra, zorra irrespetuosa

Y el que personalmente más me gusta y que si algún día Internet muere, debería ser su epitafio:

Chris McAllister @macdawg22 

@natashaleggero Eres una puta estúpida e ignorante.

Me fijé especialmente en estos dos episodios porque algo parecido le acababa de pasar a una colega mía del trabajo. El 20 de diciembre, Justine Sacco, directora de comunicación de una empresa colaboradora de OkCupid, IAC, estaba en Heathrow haciendo transbordo para coger un vuelo a Johannesburgo. Subió al avión, se acomodó en su asiento y tecleó:

Justine Sacco @justinesacco 

Voy para África. Espero no coger el sida. Es broma: ¡soy blanca!

Y acto seguido apagó el móvil. Su tuit era un chiste menos obvio que los otros dos ejemplos y, como mucho —como mucho—, se trataba de una pulla torpe sobre los privilegios de los blancos. Pero lo que empezó en forma de unos cuantos gestos de incredulidad ante su torpeza acabó convirtiéndose en un carnaval de atroz odio personal. Recibió los habituales insultos y amenazas, pero el blanco de los ataques no se limitó a su Twitter personal. Fotos de su familia circularon por la Red, junto con los datos de su domicilio. Unos tipos llamaron a sus sobrinos y amenazaron con secuestrarlos. Se reunió gente en el aeropuerto de Johannesburgo para esperarla al aterrizar. Su incapacidad para responder mientras estaba en el vuelo añadió un aliciente más de animosidad a su aterrizaje. Hacia la mitad del trayecto de su vuelo, se había acuñado ya la etiqueta #HasJustineLandedYet (ha aterrizado ya Justine), que no tardó en convertirse en *trending topic* en Twitter. Las búsquedas de su nombre en Google empezaron a dar como resultado su número de vuelo y su hora de llegada, porque aquello era lo que la gente buscaba: los algoritmos de búsqueda habían vuelto a ejercer de espejo. Durante las once horas que Justine estuvo en el aire, Internet esperó sedienta de sangre y con la boca seca el momento en el que volviese a conectarse para descubrir que su vida era un montón de ruinas.

Ron Geraci @RonGeraci 

Hay como 2 millones de personas esperándola con las luces apagadas para ver la cara que pone cuando la Tierra explote.

I'm Gary @noyokono 

#HasJustineLandedYet La gente no ha esperado con tanta ansiedad el aterrizaje de un avión desde Amelia Earhart.

V. Hussein Savage @Kennymack1971 

Hostia... A ver si acabo este trabajo y pillo unas cervezas y unas alitas de pollo a la barbacoa. Está a punto de empezar... #HasJustineLandedYet

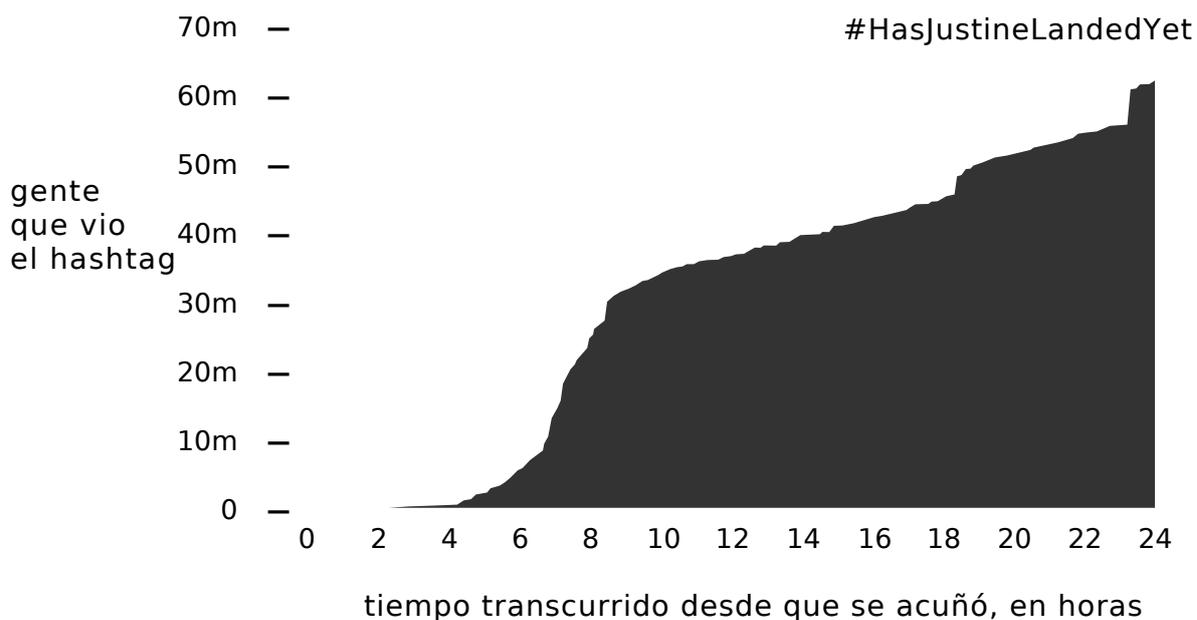
Entonces la presa era una persona que tenía unos centenares de seguidores y carecía de perfil público. Yo no conocía demasiado bien a Justine, pero me había gustado trabajar con ella y me puso enfermo observar la evidente emoción de la gente ante el dolor y el miedo que estaban a punto de causarle.

Me fui como un loco a Facebook a desahogarme. No hacía ni diez minutos que había colgado mi publicación cuando un conocido (y futuro examigo de Facebook), con quien en aquel momento hacía quince años que no había hablado, comentó «su padre es multimillonario» e insinuó que eso venía de algún modo a justificar aquella destrucción personal^[20]. Por supuesto, su padre no es multimillonario: ese es otro de los bulos que se fueron acumulando en aquella historia. Fue como si te topas con una turba enfurecida que está lapidando a alguien y, mientras tratas de apartar a la gente de aquello, te encuentras con alguien a quien conoces —uf, por fin alguien con quien se puede razonar— y ese conocido está ahí mirando ojiplático y te suelta: «¡Tío, mira todas esas piedras!».

La metáfora de la lapidación vuelve a salir una y otra vez si lees los comentarios sobre episodios de esta índole. No es casual que esa sea la pena de muerte preferida de las religiones antiguas: no hay un único verdugo, es la comunidad la que impone el castigo. Nadie puede saber quién asestó la pedrada fatal porque todos lo hicieron a la vez^[21]. ¿Qué mejor receta podría haber para una tribu floreciente que se esfuerza por sobrevivir y por conservar a su dios en un mundo hostil? En la culpa colectiva hay fuerza, y la culpa se difumina al ser compartida. Extirpas a «los otros» y vuelves a convertirte en un grupo compacto.

En el caso de Justine se juntaron gentes de tres continentes para destruirla. Si miras las descripciones que se hacen de ellos mismos en sus bios de Twitter verás que los hay de todo tipo: lobista, comunista, *hater*, asperger, líder, amante de la naturaleza, bloguero, caimán, padre, cristiano imperfecto, detector profesional de sombras, virtuoso de la cultura pop, hija del mar, hermana del viento... Lo único que tenía en común toda esa gente era un blanco de sus iras y un hashtag a medida, y consiguieron la sangre que habían venido a buscar. Justine perdió el trabajo. En la página de inicio de BuzzFeed pusieron su cara con un enorme «LOL» sobreimpreso.

El alcance que tienen las redes sociales hace que estas reuniones de gente sean inmensas. A las veinticuatro horas de haber puesto su tuit, a Safiyyah la habían regañado alrededor de 7,4 millones de personas. Y 62 millones vieron la etiqueta #HasJustineLandedY et ese primer día.



No todos los que componen esa curva leyeron los tuits o les importaron, pero muchos sí, y todos fueron en cierta medida testigos de lo ocurrido.

Sir Qwap Qwap @BeardedHistoria 

Todos y cada uno de los primeros 20 tuits que me aparecen en casa llevan #HasJustineLandedYet. Me debo de haber perdido algo, amigos del Twitter.

Es obvio que una reacción de semejante volumen debería ser una vergüenza en las redes sociales, una prueba no solo de su poder, sino de cuán vacío puede ser ese poder. En el caso de Justine, el sida, el racismo y la impenitente y vergonzante pobreza del África poscolonial constituyen problemas que no se resuelven en absoluto tuiteando.

Podemos considerar el sacrificio humano como una cosa del pasado salvaje, y puede que el acto físico solo exista hoy en películas sobre templos y fatalidades, pero el instinto sigue presente en nosotros, al parecer grabado a fuego por el paso del tiempo en las profundidades de nuestra mente animal. Cuando escasea el alimento, los leones matan a sus cachorros. Los peces se comen sus propios huevos. En los embarazos múltiples humanos, el útero a veces absorbe uno de los fetos para preservar al resto. Destruir a uno por el bien de muchos es probablemente una práctica tan antigua como la vida misma. Ahora que ese ritual se lleva a cabo mediante bits (y, afortunadamente, sin que nadie se manche las manos de sangre, aunque al leer algunos de estos tuits te llevas la impresión de que hay gente que eso lo

considera más un fallo informático que una ventaja), se ha convertido en un asunto que por primera vez podemos estudiar con rigor. Los sociólogos han dedicado considerables esfuerzos a la cuestión de por qué y cómo se difunden las ideas negativas, e Internet les ha brindado tanto una cantidad ilimitada de material de estudio como un potente mecanismo de seguimiento. Los biólogos marinos etiquetan a los tiburones en libertad para conocer sus movimientos y para limitar la amenaza que suponen para los humanos^[22]. Aquí las que tienen dientes son las palabras. Los tres casos que he mencionado no son precisamente rumores ni cotilleos, pero la ira de las multitudes sigue muchos de los mismos caminos, tanto neurológicos como de persona a persona, y el estudio de los rumores puede ayudarnos a comprender qué les ha pasado a gente como Natasha, Safiyyah y Justine, y por qué.

Los rumores se mencionan ya en nuestros primeros textos. En todos los panteones de la antigüedad —el noruego, el egipcio, el griego— hay un dios dedicado a las oscuras artes del chismorreó. El Libro de los Proverbios trata el tema exhaustivamente; uno de los muchos versículos al respecto advierte de que «el hombre falto de juicio menosprecia a su vecino, pero el hombre sensato guarda silencio». La de «no juzgues a menos que quieras ser juzgado» es una de las frases más famosas de toda la Biblia. Varias fuentes sostienen que los romanos adoraban a una diosa llamada «Rumor», un demonio alado con cien ojos y cien bocas que solo decía el lado más hiriente de la verdad. La verdad sea dicha, esto no puedo confirmarlo.

Los biólogos evolutivos creen que el chismorreó y los rumores surgieron de la necesidad que tenían nuestros ancestros de comprender su entorno a través del habla. Según la teoría, cuando el hombre de la antigüedad tuvo que averiguar si x era verdad, el lenguaje le proporcionó una manera de investigar. Así pues, habló de ello. Y, fuese cierto o falso, se corrió la voz. Los rumores —que son básicamente una especulación colectiva sobre la verdad de una idea— se convirtieron en un modo de establecer vínculos y afianzar el capital social. Las historias generan estatus para quienes las cuentan, sobre todo cuando atañen a individuos de importancia, porque la información sobre los poderosos es una forma de poder en sí misma.

Pero el auge de las redes sociales ha propiciado un par de cambios en los cálculos. Primero, nos brinda maneras de medir —recuento de seguidores, de retuits, de favoritos— para evaluar nuestro estatus. Si eres el primero en difundir la noticia, más retuits te hacen. Si dices algo especialmente mordaz, tus seguidores aplaudirán tu ingenio. El capital social que acumulas al difundir información es ahora explícito; de hecho, reside en esos numeritos

que se van incrementando ante tus mismos ojos. Jesse Singal lo dijo al escribir sobre las motivaciones del cotilleo tradicional de persona a persona en el *Boston Globe*, pero podría haber estado hablando perfectamente sobre Twitter: «En gran medida, cuando la gente mantiene una intención de difundir rumores, esta intención se dirige más a la gente a la que se los están comunicando que al propio sujeto del rumor». Internet proporciona un público mucho más nutrido que nunca.

El segundo cambio consiste en que Internet también ha hecho que todo el mundo sea una figura pública. Los individuos que tenían un estatus alto solían ser caciques tribales, y después famosos y presidentes, pero ahora la guadaña arrasadora de la tecnología no deja títere con cabeza. Si cualquiera puede convertirse en famoso de la noche a la mañana, cualquiera puede convertirse en apestado de la noche a la mañana. Uno de los argumentos del discurso evangelizador de Internet que menos me gusta es el de que la tecnología «empodera» a la gente: los más empoderados de todos son siempre el que lo dice y sus inversores. Pero algo de verdad encontramos en ese cliché: las redes sociales te empoderan hasta el punto de que te hacen merecedor de cualquier crítica demoledora. Al mismo tiempo, dan a todo el mundo los instrumentos para hacerlo. El demonio Rumor ahora tiene un millón de bocas.

Así pues, mucho de lo que hace que Internet sea útil para la comunicación —asincronía, anonimato, escapismo, ausencia de cualquier autoridad central— también la hace terrorífica. La gente puede actuar como le plazca (y decir lo que quiera) sin consecuencias, un fenómeno que estudió por primera vez John Suler, profesor de psicología en la Universidad de Rider. El nombre que le puso es el de «efecto de desinhibición *online*». En el webcómic *Penny Arcade* se explica algo mejor:

Teoría del Gran Tarado de Internet

persona normal + anonimato + público = tarado total

Pero aquí lo que es peculiar no son los sapos y culebras que se expulsan por la boca, ni siquiera el anonimato. Internet no ha sido del todo la revolución de los troles que podíamos creer. Es hartos sabido que los viejos canales de las emisoras de radio que usaban los camioneros se llenaban de diatribas racistas y fantasías masturbatorias^[23]. Antes de que la identificación de llamadas telefónicas suprimiese ese necesario aditivo, el anonimato, los Jerky Boys estuvieron produciendo «tocacojonismo» en masa durante décadas. Los radioaficionados siguen atacándose a través de las ondas, como si ser radioaficionado en 2014 no fuese ya bastante penoso. No, lo único que aporta

Internet a nuestra larga historia de negatividad es que ahora por fin podemos responder de manera constructiva. En cierto modo, la intervención de Tumblr en el caso de los *thighgap* que hemos visto en el capítulo 7 es solo un caso especial de lo que ahora es posible hacer en general. Podemos localizar al emisor, las palabras, el momento e incluso la latitud y la longitud de la comunicación humana. Como he señalado antes, los usuarios de Twitter habrán intercambiado en 2015 más palabras de las que jamás se hayan impreso. La cuestión es qué uso le damos a toda esa cháchara.

El gobierno es quien más interés tiene en hacer un seguimiento de la negatividad. Existen ya modelos matemáticos para predecir el resultado de un conflicto armado —cuánto durará, quién vencerá y cuánta gente morirá— y los modelos recientes han aprendido a incluir la guerra de guerrillas, puesto que esa es la forma actual de la guerra. Pero la insurgencia armada suele venir precedida por una agitación no armada, que muchas veces se propaga, e incluso se coordina, a través de las redes sociales^[24]. Esos movimientos incipientes, al ser digitalizados, han atraído la atención de los investigadores.

Peter Gloor, del MIT, sirviéndose de los movimientos de Occidente como sujetos de estudio, ha desarrollado un *software* para seguir las fluctuaciones de opinión de las redes de disidentes. Lo ha llamado Condor, porque así es como parecen llamarse siempre los proyectos de este tipo: cóndor, el ave-espíritu de las subvenciones del gobierno. Sea como sea, el *software* establece primero la personalidad central de los que componen determinado grupo analizando su gráfico social; el *software*, de manera parecida a como antes hemos retratado un matrimonio en forma de nodos y extremos, despliega gráficamente la red y después determina por medio de algoritmos cuáles son sus puntos más importantes. A continuación, se centra en lo que están diciendo esos puntos. Condor ha descubierto que mientras los focos de un movimiento son positivos en cuanto a las palabras que emplean, el movimiento prospera. Pero palabras negativas como «odio», «no», «inútil» o «nunca» indican su declive, y cuando, como publicó *The Economist*, empiezan a aparecer «quejas sobre compañeros idiotas en el propio movimiento o adversidades como el robo de una cerveza por parte de un compañero de manifestación», ese movimiento está ya acabado. ¡Oh, Occupy!

En cuanto a lograr desvelar los objetivos de esas agitaciones, que es donde la tecnología puede ir más allá del mero espionaje y aportar algo bueno, se han empleado análisis textuales de parecida índole para determinar, por ejemplo, qué ciudades egipcias se verán más alteradas por los incidentes

fronterizos con Israel, así como para localizar puntos de escasez de agua en un territorio asolado por la sequía.

Cualquier *software* que siga el hilo de un pensamiento a través de una red no solo debe seguir la idea, sino la «susceptibilidad» de la gente que se ve expuesta a ella. Debe ver qué es lo que arraiga, qué es lo que se repite y quién lo hace circular. Transmitir la opinión de otro no es más exclusivo de Internet que lo que lo es la negatividad: la televisión y la radio acuñaron la frase «tema de debate» mucho antes de que llegase AOL, por no hablar de Twitter. Los fans más acérrimos de Rush Limbaugh se autodenominan «Dittoheads» (algo así como «loritos repetitivos» de lo que dice él). Pero nada hace que la idea de repetir como un loro sea más simple, o más cuantificable, que los botones de «Me gusta», «Ping», «Reblog» o «Retweet». Recordemos: el 27,5 por ciento de los 500 millones de tuits diarios son retuits, gente que hace circular la reflexión de otro.

El equipo encargado de datos de Facebook se puso a investigar su versión del fenómeno y lo hicieron siguiendo la evolución en la Red de una sola actualización de estado en el muro relativa al debate sobre el sistema sanitario estadounidense de 2009:

Nadie debería morir porque no pueda permitirse atención sanitaria, y nadie debería arruinarse porque esté enfermo. Si estás de acuerdo, publica esto en tu estado durante todo el día.

Esto se republicó, textualmente, más de 470 000 veces y también generó 121 605 variaciones distintas, que a su vez se republicaron alrededor de 800 000 veces más. Alguno que creyó que esa actualización no le representaba del todo le aplicó alguna modificación y así las versiones se fueron propagando en distintos círculos sociales. Cuando se coteja cada una de las versiones con la inclinación política de las personas que las publicaron (-2,0 es liberal en extremo y +2,0 es conservador en extremo), no solo se obtiene una visión interesante del espectro político estadounidense — izquierda y derecha extremas, más un centro que opta por no entrar en el debate—, sino que se aprecia cómo se traducen en palabras las creencias políticas. Los que aparecen en lo más alto y lo más bajo de esta lista se basan en el mismo marco para entablar una suerte de diálogo de besugos:

Nadie debería

Tendencia política
del que publica

...morir porque no pueda permitirse atención sanitaria...

-0.87 más liberal

...ser congelado en carbonita porque no pueda pagar a Jabba el Hutt...	-0.37
...morir a manos de los zombis si no puede comprarse una escopeta...	-0.30
...preocuparse por morir mañana, pero los enfermos de cáncer lo hacen...	-0.02
...quedarse sin cerveza porque no pueda permitirse comprar una...	+0.22
...morir porque el gobierno se encargue de la atención sanitaria...	+0.88
...morir porque el plan de Obamacare le racione la asistencia sanitaria...	+0.96
...arruinarse debido a los impuestos y los gastos del gobierno...	+0.97 más conservador

En 1950, en los albores de la era de la televisión, la American Political Science Association reclamó una mayor polarización de la política nacional: los partidos se habían acercado demasiado y el electorado ya no tenía claras las opciones. La APSA vio cumplidos sus deseos, pero, como ocurrió con el célebre genio y su lámpara maravillosa, hubo mucho de lo que lamentarse por los deseos concedidos. Ahora, sesenta años después, estamos más divididos que nunca, y eso también lo podemos ver si nos fijamos en las palabras. La reiteración del discurso sesgado tanto en el Congreso como en el texto impreso (como podemos verificar en Google Libros) está en correlación con el atasco político, que está en su punto más álgido de la historia. Que estamos divididos puede que sea la única cosa, de hecho, en la que estemos de acuerdo.

Tomé conciencia de esta paradoja cuando me metí en Facebook después de lo del tuit de Justine. En mi publicación puse un enlace a un artículo de breitbart.com, la web homónima del instigador del Tea Party, Andrew Breitbart. Gran parte del artículo era lamentable, pero el autor era uno de los pocos que señalaban lo desproporcionado de aquella reacción. Siempre había imaginado que la indignación ciega era un vicio de la derecha política: había oído hablar de la ridícula «Guerra contra la Navidad» o de la creencia de que Obama les estaba «quitando las armas a la gente» y había pensado: ¿Cómo puede ser toda esa gente tan estúpida para creerse esas cosas? ¿Por qué hablar de las cosas en términos tan radicales? ¿Por qué ver las cosas a través del peor prisma posible? Pero hizo falta aquel incidente en Twitter para que me diese cuenta de que la gente de la «izquierda» puede estar tan intolerantemente desinformada como cualquiera. Fue revelador y me avergüenzo por no haberlo visto antes.

Así pues, dejando a un lado las teorías —y la ciencia es tan nueva que sin duda un *software* como Condor nos parecerá Zork dentro de pocos años—, este es, en mi opinión, el motivo de que los datos generados a partir de la indignación puedan llegar a ser tan relevantes. Encarnan (y, por tanto, nos

permiten estudiarlas) las contradicciones que todos albergamos. Nos muestran que luchamos con más ímpetu contra quienes menos pueden defenderse. Y, ante todo, nos hacen ver nuestro deseo inmemorial de ascender a costa de los demás. Los científicos han determinado que ese impulso es tan viejo como el tiempo, pero eso no significa que hayan conseguido comprenderlo. Como dijo Gandhi: «Siempre me ha parecido un misterio el que los hombres puedan sentirse honrados por humillar a su prójimo».

Te invito a imaginarte el momento en el que eso deje de ser un misterio. Esa será la verdadera transformación, saber no solo que la gente es cruel y en qué medida y cuándo, sino por qué. Por qué buscamos «chistes de negros» cuando gana un negro, por qué la inspiración tiene ojeras, posa desnuda y, sobre todo, está #delgada, por qué la gente discute a gritos acerca de la verdadera edad del planeta Tierra. Y por qué, al parecer, nos definimos tanto por lo que odiamos como por lo que amamos.



PARTE 3

Lo que nos hace ser
quienes somos

10.

Muy alto para ser asiático

Cuando hice la solicitud para entrar en la universidad tuve que escribir acerca de mí. Seguro que tú también lo hiciste. No logro recordar la pregunta que figuraba en la solicitud porque fuera lo que fuese lo que preguntaban no venía al caso. La redacción tenía el propósito de hacer que hablase de Christian Rudder, de modo que el personal de admisiones pudiera decidir si les gustaba lo que leían. Tal como figura ahora en el formulario de solicitud habitual: «La redacción personal nos ayuda a conocerte como persona».

Siendo ya entonces un bisoño amante del melodrama, escribí acerca de la pena que me iba a dar dejar a mi perro en casa cuando fuese a la universidad. *Frosty* llegó a casa cuando yo tenía seis años y habíamos crecido juntos. Pero tal como funciona lo de los años de los perros, se había hecho viejo demasiado deprisa. Mi familia se había mudado bastantes veces y el perro era el último vínculo que me quedaba con la infancia profunda: clubes deportivos, piscinas del barrio, amigos; todo eso lo había dejado atrás en Houston, en Cleveland o en Louisville, pero *Frosty* siempre permaneció a mi lado. El siguiente traslado, sin embargo, sabía que lo iba a tener que hacer solo.

En cualquier caso, perdido y a la deriva en aquellos tránsitos y en aquellas camisetas enormes con dibujos de M. C. Escher que llevaba, rellené mi solicitud para la universidad. No he escrito muchos textos de autoafirmación desde entonces, pero implicado como estoy en el negocio de entender a la gente, no puedo evitar rememorar mi yo de 17 años y la redacción que opté por escribir. ¿Por qué hablar de *Frosty* y de hacerse viejo? ¿Por qué no hablar de béisbol? ¿O de baloncesto? ¿O de tenis? ¿O de estadísticas de béisbol? ¿O de cualquier otra de mis aficiones? ¿Qué fue lo que, al preguntarme «Quién eres», me incitó a responder como lo hice? Y, lo que es más importante, ¿cómo contestaban a esa pregunta los demás chavales?

Ahora, pasados veinte años, me veo leyendo millones de redacciones — miles de millones de palabras— escritas más o menos como respuesta a esa pregunta: «¿Quién eres?». Y este corpus de textos me permite abordar ese mismo proceso de admisión en la universidad pero en sentido inverso. En

lugar de cotejar escritos uno a uno con un ideal preconcebido (es decir, el «material universitario»), puedo juntar todos los textos, triturarlos y ver qué ideales me descubren. Algunas veces, el conjunto de datos es tan sólido que si encaras bien tu análisis no te hace falta preguntar nada: el propio análisis te lo dice todo. ¿Cómo se describe la gente a sí misma? ¿Qué es importante, qué es típico, qué es atípico? Cuando a todo el mundo le toca el turno de poner en palabras quiénes son, ¿qué identidades esbozan?

Aquí vamos a tratar con categorías amplias: negros, blancos, asiáticos, mujeres, hombres, etcétera. Uno de los problemas que surgen a la hora de estudiar cualquier grupo concreto es que uno siempre lleva consigo sus prejuicios y sus preconcepciones. Lo que eliges advertir, recordar y transcribir tiene tanto que ver con tu modo de mirar como con aquello que en realidad tienes delante. En ciencias sociales, el conocimiento, como el agua, muchas veces adquiere la forma de su contenedor. Así que si queremos coger todas las declaraciones escritas que he recopilado y a partir de ellas formarnos una idea de quiénes son sus escritores —qué hace que grupos étnicos, sexos y orientaciones sean únicos— tendremos que desarrollar un algoritmo que extraiga el «nosotros» de esos textos y deje solamente el «ellos».

Los textos de presentación que ponen los usuarios en OkCupid son lo más parecido que se puede encontrar a un autorresumen. Son de final abierto:

«Mi autorresumen es...»

«Se me da especialmente bien...»

«Normalmente, lo primero en lo que la gente se fija de mí es...»

«Dedico mucho tiempo a pensar en...»

Y en tanto que la gente se esfuerza por empezar con buen pie no difieren tanto de las redacciones para la universidad. Imagino que mucha gente se enfrenta a ellos con parecido pavor. No hay limitaciones de extensión ni directrices en esas manifestaciones escritas. En total, la gente ha aportado al sitio web 3200 millones de palabras de autodescripción. Además, a diferencia de otros grandes bloques de texto —por ejemplo, los que ha recopilado Google Libros—, detrás de cada palabra hay datos demográficos: la edad del autor, dónde vive, su raza y demás. Pero inferir de un texto la identidad grupal de, pongamos, las mujeres asiáticas no es tan sencillo como contar quién teclea más caracteres, que en gran medida es la manera en la que hemos tratado los textos hasta ahora en este libro. Si contamos palabras solo obtenemos esto:

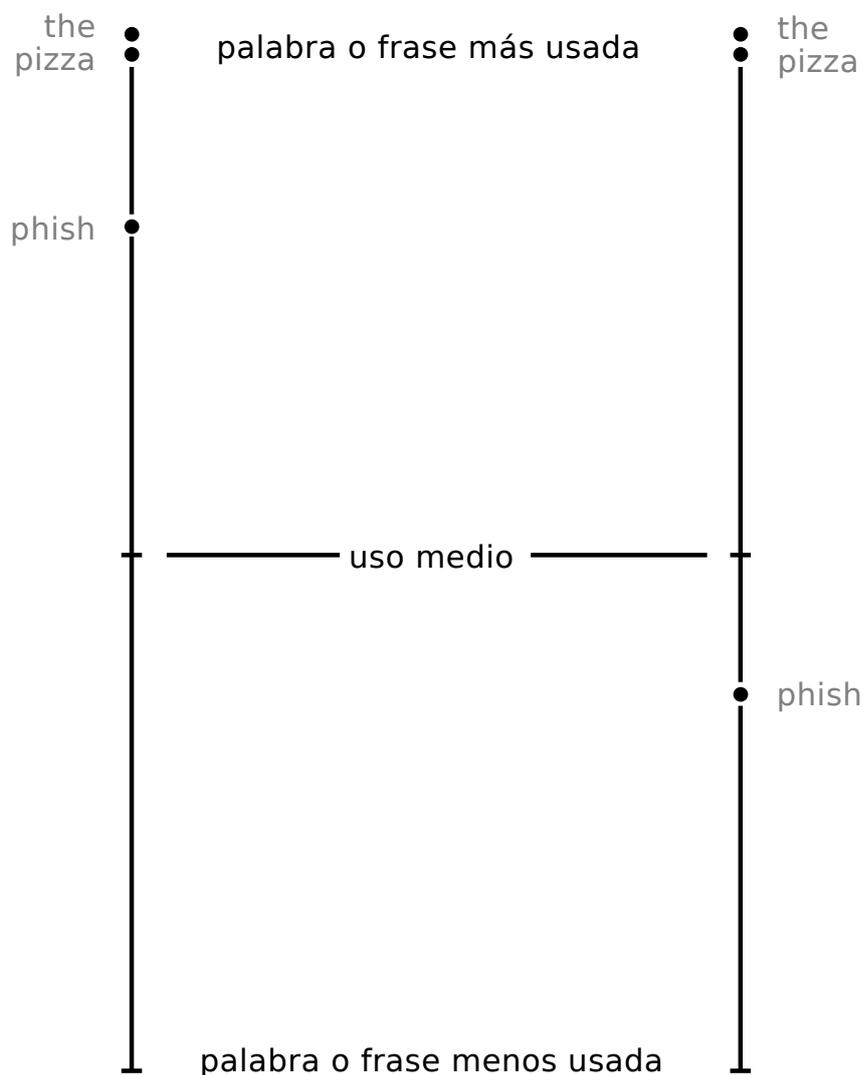
1. the
2. of
3. and
4. ...

y así sucesivamente; básicamente, esas 100 palabras más usadas del Oxford English Corpus que hemos visto antes. Las mujeres asiáticas, los hombres blancos y todos los hablantes del inglés emplean los mismos pronombres, artículos y preposiciones para hablar de sí mismos. Para averiguar qué es de verdad lo que hace que determinado grupo, y solo ese grupo, sea especial hemos de clasificar el texto de manera algo distinta.

Me serviré de los hombres blancos como ejemplo ilustrativo, porque son los que mejor conozco. El primer paso es separar los textos de esos hombres blancos de los del resto. Después, en los dos conjuntos de autodescripciones —las de los hombres blancos y las de los que no lo son— ordenamos todas las frases y palabras de los textos por la frecuencia con la que aparecen. Las ponemos en dos listas, de la más usada a la menos usada, y obtenemos así una cosa parecida al diagrama de abajo. He extraído tres ejemplos y los he puesto en los sitios que les corresponden de la línea; las listas completas constan de alrededor de 360 000 frases cada una:

autodescripción de hombre blanco

autodescripción del resto



Ya vamos llegando a alguna parte, pero antes de seguir avanzando y mientras la lista sigue siendo sencilla, debo decir que en estos gráficos hay algo que resulta confuso. No, no tiene nada que ver con la banda Phish, aunque Dios es testigo de que han confundido a muchos. Se trata de que las palabras «pizza» y «the» parecen mencionarse prácticamente el mismo número de veces. De acuerdo con que la *pizza* es la reina de la comida, pero ese «the» es la palabra absolutamente más popular de la lengua inglesa. Y en nuestros datos, aunque el «the» está en su pertinente primer puesto en lo más alto, al parecer «pizza» está justo a su lado, con un percentil de 98. Esto puede causar la impresión de que hay algo que está mal en mis datos o en mi método, pero la clasificación de las palabras es la correcta. Es solo que los humanos empleamos el lenguaje de un modo peculiar: siempre nos estamos

repitiendo. Así que un puñado de palabras de lo más alto de la clasificación acaparan la mayor parte de lo que escribimos. Y, a la inversa, la frecuencia de una palabra cae en picado en cuanto te alejas aunque sea mínimamente de las «más populares».

Esta relación contradictoria entre la popularidad de una palabra (su «rango», en el vocabulario especializado) y el número de veces que aparece la describe lo que se denomina ley de Zipf, una propiedad estadística que se ha observado en el lenguaje y que, como ocurre con otras joyas de las matemáticas, reside en alguna parte entre el milagro y la coincidencia^[25]. Su enunciado es que en cualquier cuerpo extenso de texto, la popularidad de una palabra (su puesto en el vocabulario, en el que el 1 es el rango más alto) multiplicada por el número de veces que figura es idéntica para cualquier palabra de ese texto. O, expresado con más elegancia:

$$\text{rango} \times \text{número} = \text{constante}$$

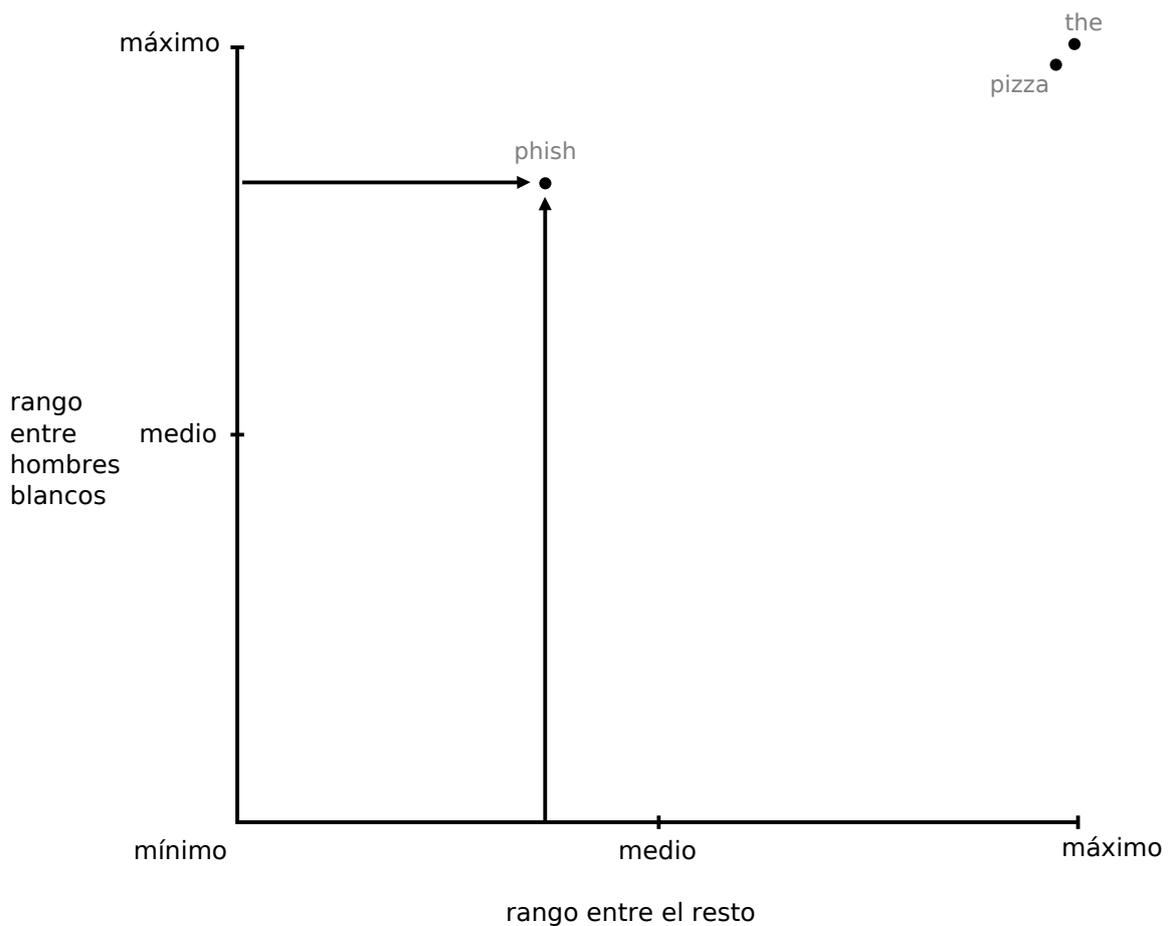
Esta ley vale para la Biblia, para las letras de canciones de los años sesenta, para el corpus canónico de la literatura inglesa (el Oxford English Corpus) y también, claro, para los textos de los perfiles. Para ver lo bien que funciona en la práctica, incluso en un texto sumamente idiosincrásico, he aquí la ley aplicada al *Ulises* de James Joyce^[26]:

palabra	rango	número de apariciones	rango × número
's	10	2826	28 260
is	20	1435	28 700
what	30	975	29 250
has	100	289	28 900
wife	200	140	28 260
Ireland	300	90	28 260
college	1000	26	28 260
morn	5000	5	28 260
builder	10 000	2	28 260
Zurich	29 055	1	28 260

La relación invariable entre rango y número parece ser una propiedad de la mente tanto como lo es del lenguaje. Como puede verse, la tabla abarca nombres propios arbitrarios, como «Ireland» y «Zurich», e incluso palabras transcritas de un dialecto, como esa «'s».

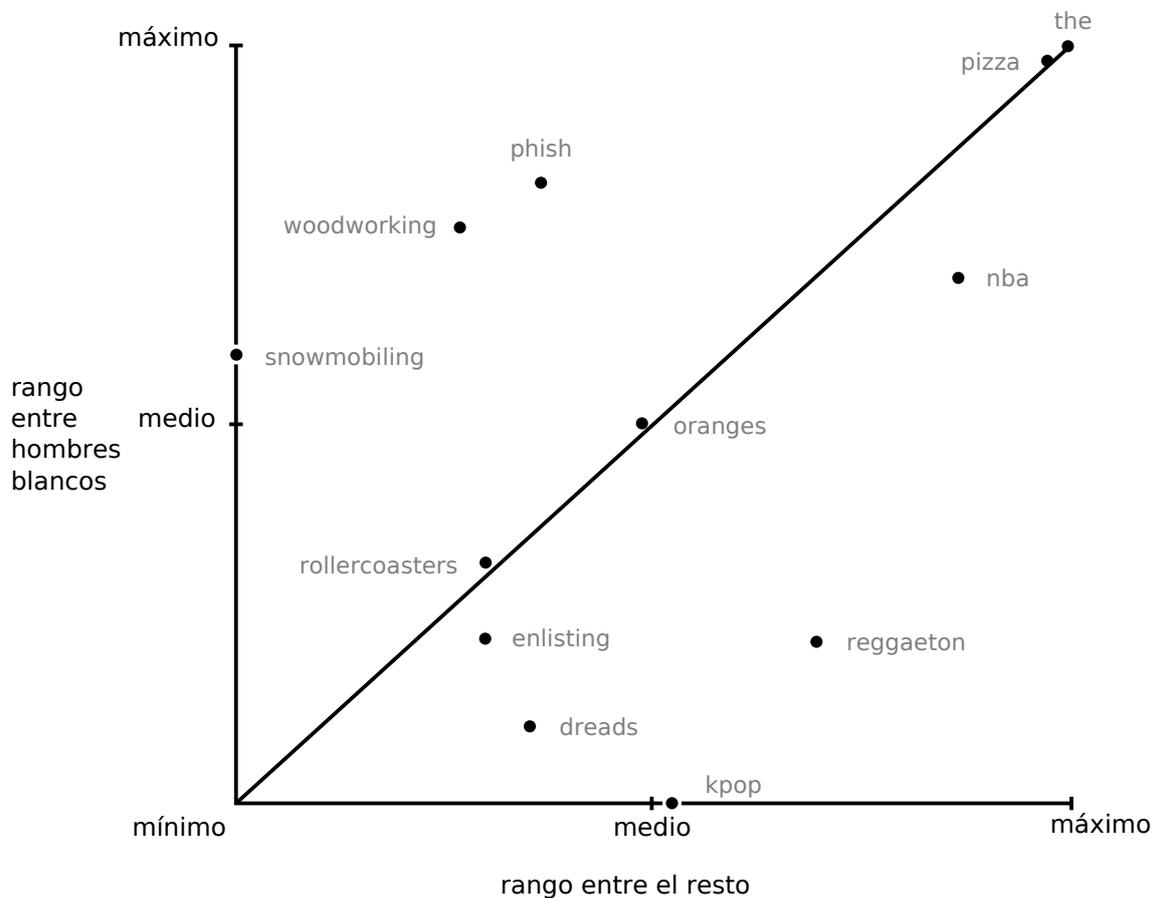
Y para mayor prueba de su estrecho vínculo con la experiencia humana, la ley de Zipf también sirve para describir varias de nuestras creaciones sociales: el tamaño de las ciudades, por ejemplo, o la distribución de ingresos entre determinada población. Lo que eso implica para lo que aquí nos ocupa es que dado que el lenguaje no es más que un pequeño cuerpo de patrones repetidos, el uso de una palabra decae enseguida. El artículo «the» aparece en casi todos los perfiles. «Pizza» figura en uno de cada 14. «Phish», incluso en el caso de los hombres blancos, para quienes alcanza un rango tan elevado como el percentil 80, aparece en menos de uno de cada 200 perfiles. Ahora que ya entendemos cómo comparar los rangos y la frecuencia de uso, el siguiente paso es emplear esos rangos en nuestro provecho.

A continuación he colocado las dos listas en ángulo recto, formando un cuadrado, y he ubicado las palabras en su interior empleando como coordenadas sus rangos de popularidad en ambas listas. He puesto unas flechas en la palabra «phish» para que quede claro:



Aquí, la posición de una palabra tiene doble significado. Cuanto más cerca esté del margen máximo, más popular es entre los hombres blancos.

Cuanto más a la derecha aparece, más popular es entre todos los demás. Si añadimos unas cuantas palabras más al gráfico nos haremos una idea de cómo funciona esta geometría antes de que pasemos a ver el corpus completo:

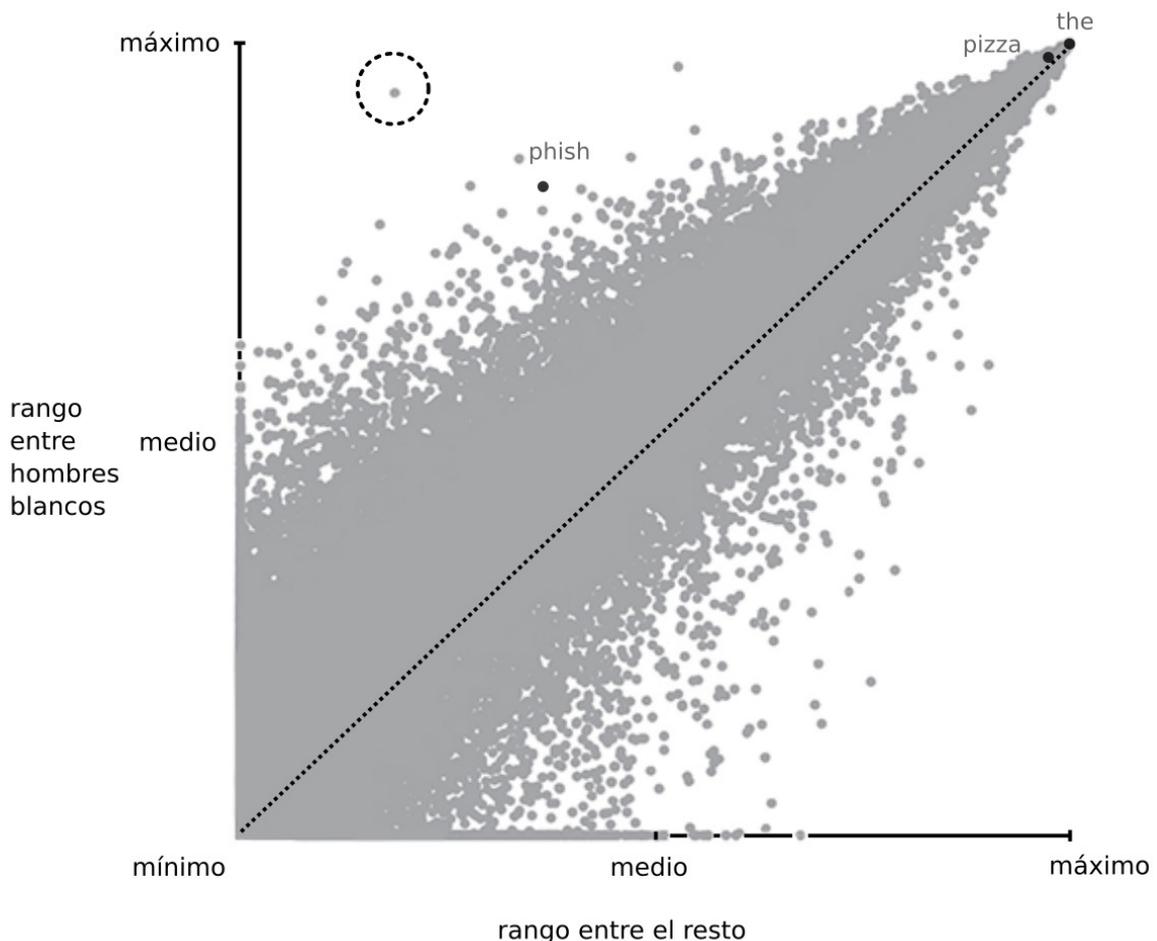


También aquí he añadido una línea en diagonal para mostrar la paridad de los datos. Las palabras que más cerca están de la diagonal son igual de importantes para todos. Y cuanto más hacia arriba y a la derecha están, más universalmente importantes son. Pero recordemos que no estamos buscando universalidades. Estamos buscando particularidades. Queremos saber qué es especial para la gente que estamos analizando: en este caso, hombres blancos. Para lo cual deberemos fijarnos en la parte superior izquierda: cuanto más se acerque una palabra a ese punto, con más frecuencia la usan los hombres blancos y con menos frecuencia la usan todos los demás. De hecho, cuanto más cerca se halla una palabra de ese remoto ámbito de la masculinidad blanca, el vértice superior izquierdo del cuadrado, más los tipifica a ellos y solo a ellos. Imaginemos que hay un punto justo en esa esquina; para estar allí, la palabra debería aparecer en todos y cada uno de los perfiles de hombres blancos y, a la vez, no aparecer en ningún otro perfil. Al menos en lo

que concierne a las palabras que aparecen en un escrito de presentación, ese sería el ideal platónico de la identidad. Este sistema, y esa medida —la distancia del vértice superior izquierdo— les facilita a los datos una manera de hablarnos, de ayudarnos a entender cómo hablan las personas de sí mismas.

Puesto que cada conjunto de datos tiene sus peculiaridades, muchas veces los investigadores crean herramientas desde cero, como hemos hecho aquí. Siempre que hagas esto conviene que verifiques tu método mediante resultados ya conocidos. Imagina que un armador tiene un buque nuevo: quién sabe lo que ocurrirá cuando esté en alta mar, así que mejor revisar bien el casco cuando esté cerca de la orilla por si tiene agujeros. En este caso, si hubiésemos encontrado que «Kpop» (pop coreano) o «dreads» aparecían en la esquina superior izquierda, en mi supuesto rincón de la hombría blanca, tendríamos un indicio muy claro de que mis datos o mi método son una porquería. Pero, como se puede ver, funciona a la perfección.

Así pues, por fin, he aquí el aspecto que presenta el corpus completo de palabras y frases:



He rodeado con un círculo el punto que más se acerca al vértice superior izquierdo; es la cosa más masculina blanca que una persona puede escribir sobre sí misma: «mis ojos azules». Y para obtener una lista más extensa de las cosas que definen singularmente a los hombres blancos solo es cuestión de ir alejándose de ese vértice: por ejemplo, los 30 puntos más cercanos son las 30 cosas más típicas. La geometría encuentra los clichés por nosotros.

He hecho gráficos como este para todos los que componen mi conjunto de datos, no solo los tipos blancos, y con este mismo método matemático he obtenido también listas de sus palabras y frases exclusivas. Pero antes de pasar a enumerar todo eso, quiero señalar una cuestión importante. Si repasamos cualquiera de las combinaciones de sexo × grupo étnico × orientación obtendremos $2 \times 4 \times 3 = 24$ gráficos como este que hay arriba, y en todos ellos la masa de puntos tiene esa misma forma ahusada, ancha por abajo a la izquierda y que se estrecha en la esquina superior derecha. Es decir, cuanto más cerca está una frase de esa esquina superior derecha, más se aproxima a la línea diagonal. Lo que eso significa es que tendemos a estar de acuerdo en las cosas más importantes. En cuanto a las cosas en las que no coincidimos, las he enumerado con todo detalle a continuación. Empezaré por los hombres^[27]:

palabras más típicas de hombres...

blancos	negros	latinos	asiáticos
my blue eyes	dreads	colombian	tall for an asian
blonde air	jill scott	salsa merengue	asian
ween	haitian	cumbia	taiwanese
brown air	soca	una	taiwan
hunting and fishing	neo soul	merengue bachata	cantonese
allman brothers	jamie foxx	mana	infernal affairs
woodworking	zane	banda	seoul
campfire	paid in full	puertorican	infernal
redneck	nigga	colombia	shangai
dropkick murphys	luther vandross	gusta	boba
they might be giants	coldest winter	puerto rican	kbbq
brewing beer	tyler perry	tejano	kpop
robert heinlein	swagg	corridos	badminton
tom robbins	jerome	bachata merengue	kimchi

townes	dreadlocks	hector	chungking express
old crow medicine show	spike lee	espa	chou
mystery science theater	holla at me	por	viet
skis	menace to society	salsa bachata	jiro
sailboat	brotha	aventura	dash berlin
around a fire	shottas	english and spanish	ucsd
caddyshack	boomerang	musica	beijing
blond hair	nigerian	espa ol	hk
bill bryson	hearthbeats	como	norwegian wood
wheelers	anthony hamilton	fiu	jiro dreams of sushi
pogues	gud	pero	lin
barenaked ladies	wayans	soledad	philippines
mst3k	dickey	espanol	noodle soup
truckers	isley	amor	malaysian
jethro tull	interracial	muy	for my next meal
canoe	nigeria	reggaeton	gangnam style

Puede que Phish ya no figure en esa lista, pero en el interior del hombre blanco sigue bullendo un festival de música para leñadores.

En cuanto a las otras tres listas, nunca había oído hablar de Zane, ni de Anthony Hamilton, ni de *The Coldest Winter Ever*, ni de *Chunking Express*, ni de Dash Berlin ni de muchas de las cosas que ahí aparecen antes de que mis scripts las escupiesen, y no voy a pretender que unos cuantos minutos consultando Wikipedia sustituyan a toda la comprensión de una cultura. Se trata de usuarios que hablan con voz propia, y mi intención es dejarles que hagan eso mismo, pero sí quiero señalar unas cuantas tendencias generales: los blancos se diferencian a sí mismos sobre todo por su pelo y sus ojos, los asiáticos por su país de origen y los latinos por su música. Pero debido a la configuración matemática, las tres listas de no blancos evidencian unas culturas que yo, como hombre blanco, no se supone que tenga que conocer. Naturalmente, todos estamos familiarizados con Spike Lee, Pekín y Shanghái, pero estas listas nos proporcionan la visión que tiene un «autóctono» de determinada cultura. Es una visión que un foráneo no puede lograr mediante la función de autorrelleno, ni de cualquier otra manera que implique ir de arriba abajo, porque no podemos preguntar sobre algo de cuya existencia no somos conscientes. «¿Por qué a los asiáticos les gusta *Norwegian Wood (Tokio Blues)*?» no es un estereotipo porque no son muchos los asiáticos que

conozcan la novela (de Haruki Murakami) o la película. Yo creí que se refería a la canción de los Beatles, y si antes de escribir este capítulo alguien me hubiese preguntado si había visto *Norwegian Wood*, le habría contestado: «No creo que por aquel entonces hiciesen videoclips». Estas listas que hemos visto son nuestro santo y seña. Como tales, no son algo que uno pueda generar *a priori*, ya sea escribiendo cosas en Google Tendencias o buscando entre millones de hashtags. A veces hace falta un algoritmo ciego para ver realmente los datos.

Aquí van las listas de las mujeres. Como podrás observar, son muy parecidas a las de los hombres en cuanto a su esencia. Si acaso, contienen unas cuantas baladas más.

palabras más típicas de mujeres...

blancas	negras	asiáticas	latinas
my blue eyes	soca	taiwan	latina
red hair and	eric jerome dickey	tall for an asian	colombian
blonde hair and	haitian	philippines	una
love to be outside	imitation of life	taiwanese	cumbia
mudding	zane	beijing	banda
campfire	coldest winter ever	coz	tejano
four wheeling	nigerian	boba	merengue bachata
phish	interracial	filipina	gusta
hiunting fishing	rb and gospel	cantonese	puertorican
campfires	five heartbeats	asians	colombia
green eyes and	anita baker	won kar wai	mana
redneck	crooklyn	shangai	vida
auburn	neosoul	seoul	bachata merengue
ride horses	octavia butler	macarons	amor
old crow medicine show	housewives of atlanta	viet	musica
grateful dead	luther vandross	kimchi	english and spanish
mountain goats	zora	for my next meal	espanol
love country music but	waiting to exhale	singapore	salsa merengue
gillian welch	anthony hamilton	malaysian	todo
country girl	chrisette	hk	por
christmas vacation	locs	malaysia	mariachi

bill bryson	outside my race	noodle soup	marc anthony
riding horses	kem	cambodian	espa ol
eric church	octavia	norwegian wood	novelas
barn	real housewives of atlanta	hong kong	como
allman	calypso	chungking express	pero
willie nelson	know why the caged	rachmaninoff	venezuela
harley	did i get married	southeast asia	soledad
brunette	spike lee	vienna	mas
flogging molly	braxton	mandarin	tacuba

Mientras trabajaba con esto he descubierto que el algoritmo que empleamos para confeccionar estas listas es flexible. Puedes hacer las cuentas a la inversa con la misma facilidad. De ese modo obtienes la antítesis de un grupo: las cosas de las que específicamente no hablan, que pueden arrojar tanta luz como aquellas de las que hablan especialmente. Estas son las listas de los hombres; están impresas sobre un fondo más oscuro para acentuar gráficamente que estas listas son lo opuesto a las anteriores. Son las palabras menos usadas por estos grupos pero que más usan todos los demás, el espacio negativo de nuestro test de Rorschach verbal. Vale la pena leerlas de cabo a rabo:

palabras más antitéticas de hombres...

blancos	negros	asiáticos	latinos
slow jams	borges	sence	southern accent
trey songz	social distortion	layed	from the midwest
robin thicke	tallest man on earth	layed back	ann arbor
smh	gaslight anthem	sence of humor	midwestern
musiq	snorkeling	truck driver	gumbo
merengue	belle and sebastian	6'4	freakanomics
laker	xkcd	really	equity
ig	diet coke	anything ese you wanna	discworld
kevin hart	surfboard	like what u see	shanghai
raised in nyc	totoro	and my son	scallops
hip hop rap rb	magnetic fields	u like what u	slopes
kpop	gogol bordello	care of my kids	university of michigan
george lopez	dropkick murphys	makeing	assessment

neo soul	rebelution	welder	parentheses
rb and hip hop	peru	hunting fishing	snowboarder
neyo	horrible's sing along blog	care of my son	nyt
knw	wakeboarding	wanna know anything else	dominion
gud	herzog	else you wanna know	msu
follow me	my blue eyes	raising my son	ellipses
jordans	guitar and sing	ask and ill	maple
handball	dr horrible's sing along	comedys	nigerian
soulchild	coachella	dnt	kenya
ne yo	dr horrible's sing	woman who wants	john irving
bachata	yo la tengo	i'm a single father	over a decade
basketball	airborne toxic event	somthing	cheesesteaks
paid in full	yosemite	careing	wall street journal
mos def talib	feynman	writting	alternatively
mangas	coppola	and my daughter	mistborn
abt	wind up bird	haveing	weber
utada	kar	brown hair	gravitate toward

La lista de los «opuestos a los latinos» es la que más me sorprende. Los demógrafos muchas veces fusionan las identidades hispana y blanca; por ejemplo, el censo de Estados Unidos lleva años tratando de separarlas. Pero el único recurso que tienen para hacerlo son las casillas de verificación de sus formularios. Las listas de palabras «más típicas» y sus «opuestas» de los latinos que vemos aquí definen los extremos. La primera nos brinda lo más extendido de la cultura latina (la música y la lengua) y la segunda nos da el estereotipo del blanco «criado con maíz» del Medio Oeste, que es una de las pocas subculturas blancas exenta de influencias de los latinos. Nótese también que las cosas «menos asiáticas» son todas palabras mal escritas, oficios de la clase trabajadora y otras cosas de índole pedestre, como los padres solteros. Y, naturalmente, también están esos «6,4» (pies de estatura).

Las listas correspondientes a las mujeres son igualmente interesantes, y vuelvo a sugerir leerlas enteras. En las antítesis asiáticas está ese sorprendente «me llamo Ashley». Y debo decir, como toque de orgullo profesional, que cuando le preguntas a un algoritmo «¿De qué no hablan las mujeres negras?» y te contesta «bronceado», sabes que lo has hecho bien.

palabras más antitéticas de mujeres...

blancas	negras	asiáticas	latinoas
filipino	belle and sebastian	bbw	midwestern
neo soul	tanning	god my children	cincinnati
musiq	bruins	single mother of two	classically
slow jams	tahoe	grandson	kenya
rich dad poor dad	simon and garfunkel	god my daughter	neal
corinne bailey rae	magnetic fields	mother of three	shanghai
bailey rae	sf giants	human services	financial services
salsa bachata	floggin molly	degree in criminal justice	classically trained
aaliyah	head and the heart	single mom of two	southern belle
jpop	dodgers	notice my eyes and	cutting for stone
smh	wavy	wanna know just ask	in new england
salsa merengue	naked and famous	mexican and chinese	antarctica
nujabes	social distortion	they are my world	kavalier
48 laws of power	mountain biking	being the best mom	full disclosure
musiq soulchild	portugal, the man	raising my children	gravitate toward
neyo	camera obscura	a better life for	brussels
2ne1	rancid	associates degree in	toronto
esperanza	yo la tengo	curly hair and	nyt
mangas	paddle boarding	madea	cambridge
zane	armin	im a single mom	adventure of kavalier
n.e.r.d	santa cruz	mexican and italian food	creole
coldest winter ever	ecuador	i'm a country girl	meetup
mines	ccr	ellen hopkins	parentheses
ratchet	the dog park	people notice my eyes	arbor
aventura	bbqing	my name is ashley	curl up with a
malcolm x	origami	brittany	for my next meal
asians	handshake	at a daycare	singer songwriters
carne	gabriela	my family my cell	ann arbor
hw	line is it anyway	want a man that	raleigh
earphones	sunblock	me and my son	interpreter of maladies

Hasta ahora he hablado mucho de la raza, y lo he hecho, como ya he señalado, porque es algo que no se suele tratar mucho de manera analítica. Y los datos de que dispongo son ideales para derribar tabúes. Pero el sexo es lo

más importante, con diferencia, de lo que agrupa a los humanos. Siempre ha existido, aunque nos remontemos a cuando no éramos más que un solo pueblo, y tal vez a causa de esas raíces tan profundamente hincadas en el tiempo; los roles de género son más universales y más tercamente inmutables que cualquier otro. Es fácil olvidar, dado lo imposible de erradicar que parece la segregación racial, que las ideas raciales son producto del tiempo y la ubicación geográfica. Los irlandeses y los europeos orientales no se consideraban «blancos» hasta la década de 1900; en México, los indígenas mayas y los mestizos de sangre española llevan siglos siendo grupos étnicos diferentes (y adversarios políticos). Aun así, para mucha gente de Estados Unidos, ambos son simplemente «hispanos». Pero la división sexual es un hecho propio de la cultura humana, de toda cultura y de cualquier época.

Paradójicamente, OkCupid no es el mejor sitio para estudiar las diferencias entre hombres y mujeres, al menos no mediante el método que hemos desarrollado aquí. Tu sexo es algo innato en la manera en que usas una web de contactos, de modo que, por ejemplo, lo más destacado que observamos en el texto de los perfiles de las mujeres (heteros) es que van en busca de un hombre, y así con el resto. El sexo y el texto del perfil son indisolubles, y cualquier análisis no arroja más que redundancias. La fuente ideal para analizar la diferencia de géneros es aquella en la que el género del usuario es nominalmente irrelevante, en la que no importa que la persona sea hombre o mujer. Ese terreno neutral es para mí Twitter. Las listas que vienen a continuación se elaboraron mediante el mismo proceso matemático que las que hemos visto de OkCupid, pero empleando el texto de los tuits de los usuarios.

palabras más típicas de...

hombres	mujeres
good bro	my nails done
ps4	my sissy
james harden	mani pedi
mark sanchez	my makeup
my beard	my purse
cp3	girls night
in 2k	my hair for
bynum	prom dress
the squad	girls day

bro we	retail therapy
manziel	thanks girl
in nba	my future husband
year deal	to dye
iverson	dress shopping
yeah bro	too girl
kyrie	happy girl
hoopin	bobby pins
free agent	wanelo
tim duncan	my boyfriend and
scorer	my belly button
offseason	my roomie
hof	girlies
xbox one	dying my
david stern	cute texts
yds	girl crush
fantasy team	my boyfriends
gameplay	eyebrows done
gasol	curl my
lbj	my hubby
bro u	us girl

Con esto obtenemos la esencia destilada de hombres y mujeres: lee la lista y vuélvete más estúpido. Recuerda, antes de deprimirte, que el método está diseñado para hallar lo que es exclusivo de cada grupo, para encontrar las cosas que no tienen en común y ponerlas en primer plano. Es la versión matemática del dibujante callejero: la caricatura a través del algoritmo, en lugar del aerógrafo.

Estas palabras son los extremos, pero tanto en hombres y mujeres, al igual que en los grupos étnicos que hemos visto, el vocabulario básico («the», «pizza» y demás) es común. De hecho, existe un consenso cada vez mayor entre los psicólogos respecto a que hombres y mujeres son fundamentalmente muy parecidos, pese a la cosmología popular que los coloca en planetas distintos. Los investigadores de la Universidad de Rochester proclamaron recientemente que «Los hombres son de ~~Marte~~ la Tierra y las mujeres son de ~~Venus~~ la Tierra», y concluyeron:

Desde la empatía y la sexualidad hasta la inclinación por la ciencia y la extroversión, el análisis estadístico de 122 características referidas a 13 301 individuos demuestra que hombres y mujeres, en líneas generales, no pertenecen a grupos distintos.

Y aun así, aunque mi método está hecho para extraer las diferencias, cuesta imaginar dos series de intereses más contrapuestos que los que aparecen en la lista de arriba. No sé por cuál de los dos decantarme: por una parte, no es un mundo mejor aquel en el que las mujeres tienen fijación por su aspecto y los hombres viven a base de chutes de proteínas; por otra parte, si hombres y mujeres fuesen exactamente iguales, la vida no tendría mucha gracia. Lo mismo ocurre con las listas por grupos étnicos. Las diferencias culturales, aunque a veces den risa, hacen del mundo un lugar más rico.

Lo de Marte y Venus, por muy metafórico que sea, me recuerda que el firmamento ha sido un punto de referencia para la ciencia desde antiguo. Aristóteles miraba al vacío que tenía sobre su cabeza en busca de verificación de su éter. Newton confirmó su ley de la inversa del cuadrado a través del movimiento de Marte. Incluso Einstein no fue del todo Einstein hasta que el sol y la luna lo dijeron, en un eclipse producido en 1919 que confirmó su teoría de la relatividad. Pese a que aquí no estamos trabajando en algo tan grandioso como todo eso, debo decir que confío en que el texto irónicamente tachado del título de ese artículo sea precipitado, al menos en cuanto a las cosas de las que hablamos y a nuestra manera de emplear el tiempo. Míralo de este modo: si no hubiese ningún otro planeta más que la Tierra, el universo sería un sitio muy aburrido.

11.

¿Te has enamorado alguna vez?

Hace unos años, un par de alumnos del MIT emplearon datos de Facebook en un proyecto de clase para crear un «radar gay» que funcionase. Se trataba de una sencilla aplicación de *software* que se comportaba de manera muy parecida a como lo hace una persona que trata de hacer una conjetura al respecto: se fijaba en quiénes eran los amigos de cierto individuo. El programa aprendía rápidamente a reconocer que determinado equilibrio de gais y heteros en el círculo social de un hombre indicaba con fiabilidad cuál era su sexualidad; no necesitaba saber nada más de ese individuo. Como publicó entonces el *Boston Globe*: «Puede que la gente esté efectivamente “saliendo del armario” por el simple hecho de relacionarse con determinadas compañías virtuales». Después de que los estudiantes entrenasen el *software* aplicándolo a perfiles conocidos, este era capaz de predecir correctamente si un hombre era gay el 78 por ciento de las veces, simplemente a partir de su gráfico social. Ese es un resultado muy sólido cuando se tiene en cuenta que el índice de acierto que cabe esperar de un programa que se limite a especular a ciegas habría sido de solo..., ejem..., ¿un 10 por ciento?, ¿un 2 por ciento?, ¿un 8?, ¿ $\pi/2$?

Esa es justo la cuestión —y el motivo, en parte, de que aquellos chavales crearan un programa para averiguarlo—, que nadie sabe en realidad cuántos gais hay. Los cálculos previos varían enormemente, como suele ser habitual en los cálculos del pasado^[28]. Uno de los primeros intentos científicos de dar con un cómputo real fue el *Informe Kinsey*, de 1948; hizo que muchos ceños se fruncieran tras sus respectivas gafas de concha al sugerir que el 10 por ciento de los hombres y el 6 por ciento de las mujeres eran homosexuales. Estudios posteriores, muchos de ellos con motivaciones políticas y todos realizados a partir de datos extraídos de encuestas o pergeñados en laboratorios, han rebajado la cifra hasta el 1 por ciento o la han elevado hasta el 15 por ciento^[29]. Ahora somos capaces de hacernos una idea siguiendo una ruta distinta, y aquí es importante mejorar la precisión, puesto que, como afirmaba un estudio sin gracia alguna: «Este trabajo puede ser de utilidad para determinar las políticas públicas». Todas las elecciones presidenciales desde 1952, salvo cuatro, habrían dado resultados muy distintos de haber cambiado

de opinión el 5 por ciento del electorado, de modo que la cuestión de si un grupo constituye el 1, el 5 o el 10 por ciento del país es de interés primordial para los cálculos políticos. Pese a que la cantidad de población gay no conlleva peso moral alguno —aunque solo hubiese una persona homosexual en todo Estados Unidos, él o ella merecería tener los mismos derechos que cualquier otro—, no es sino una realidad práctica que las decisiones políticas dependen del tamaño real de la población.

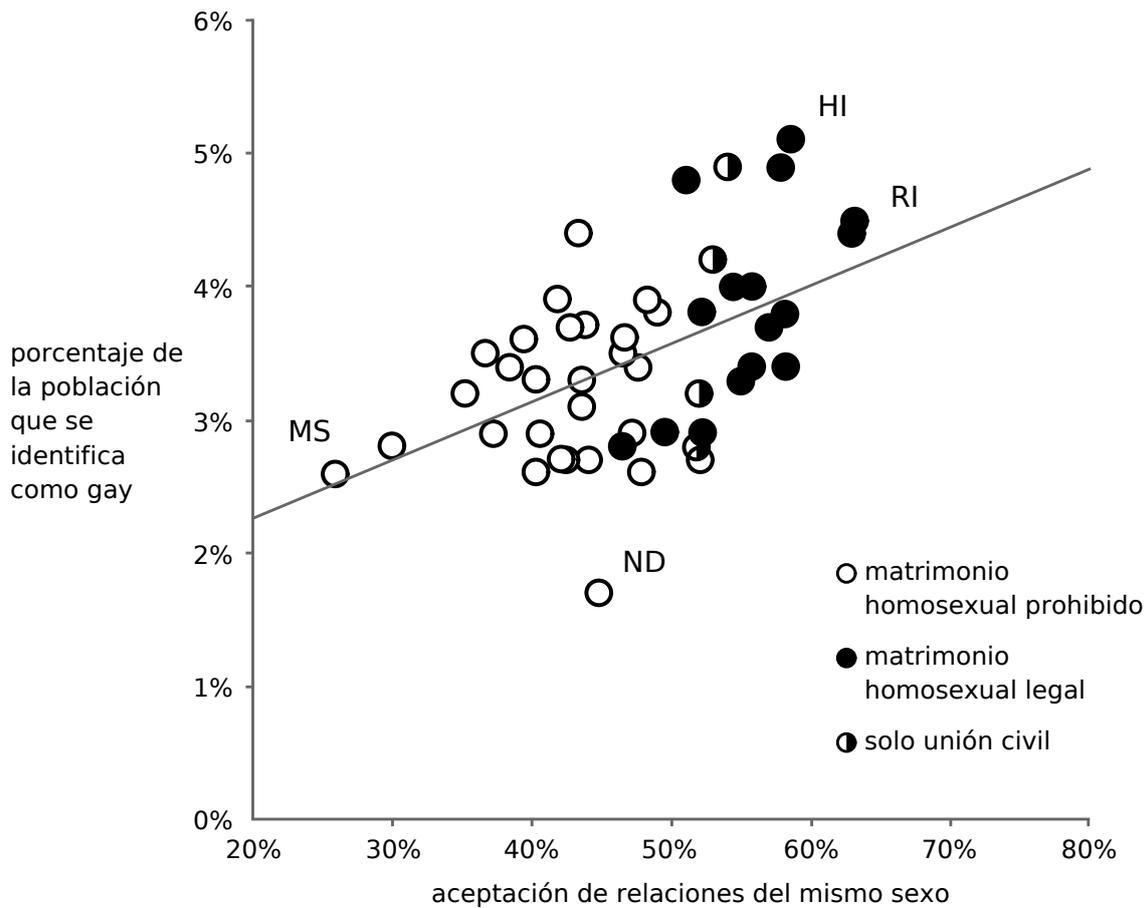
Además, en el caso de un grupo tan estigmatizado históricamente, un número contundente se hace oír allí donde el individuo no lo logra. Dice: estoy aquí. Los homosexuales son una minoría un tanto inusual porque pueden parecer heteros, al menos de manera superficial, si deciden que deben hacerlo. Esto implica ciertamente un doloroso dilema entre la autoconservación y la autoexpresión al que muy poca gente tiene que enfrentarse. Pero dejando de lado el obvio coste que supone para el individuo, «el armario» también le sale caro a nuestra sociedad, ya que esa discreción permite que continúen inalterables viejas actitudes, y un prejuicio no cuestionado es un prejuicio perpetuado. Al obligar a la gente a esconderse, la intolerancia genera su propia y cínica lógica: cuando una gran porción de un grupo no se ve reconocida, lo único que se consigue es marginalizar con más facilidad al conjunto. Por otra parte, la visibilidad genera aceptación. Aun tirando de los cálculos más bajos, la homosexualidad no es más infrecuente que el cabello rubio natural, con el que nace algo así como el 2 por ciento de la humanidad. De hecho, ser gay parece ser mucho más habitual que eso. Sencillamente se acepta menos y por consiguiente se aparta de la vista. Piensa en ello la próxima vez que hojees una revista de famosos.

Volviendo a los datos, Google Tendencias saca de nuevo a relucir su poder para descubrirnos que la gente no dice lo que piensa. Según Stephens-Davidowitz, el investigador de Google, el 5 por ciento de las búsquedas de porno en Estados Unidos era de lo que él denomina «representaciones de hombres homosexuales» —eso es un comodín que abarca solicitudes directas como «porno gay» y búsquedas relacionadas como «Rocket Tube», un portal gay muy popular—. Por añadidura, esa proporción de 1 a 20 es la misma en todos los estados, lo que significa que el deseo por personas del mismo sexo no se ve afectado por el ámbito político y religioso de un hombre. Esa uniformidad implica unas cuantas cosas de importancia. En primer lugar, desarma todo argumento de que la homosexualidad es cualquier cosa menos genética. Si la misma cantidad de hombres de entornos tan distintos como Misisipi y Massachusetts buscan porno gay, esa es una prueba fehaciente de

que cualquier supuesta influencia externa afecta poco a la atracción por gente del mismo sexo.

Para llegar a lo segundo que indica esa uniformidad entre estados en los datos —es decir, lo que nos revela no tanto sobre los gais, sino sobre la intolerancia— necesitaremos un poco más de tiempo. A principios de 2013, cuando Nate Silver todavía cubría la sección de política para el *Times*, aplicó su famosa técnica de predicción electoral a las iniciativas en pro de votar sobre el matrimonio homosexual que se presentaron por todo el país. Como había hecho en las elecciones presidenciales, recopiló datos para obtener una instantánea de la opinión pública de cada estado y luego elaboró ciertos análisis de predicción para averiguar cómo podrían evolucionar esas actitudes. Silver calculó que para 2020 el matrimonio homosexual sería legal en 44 estados.

Una cosa interesante del trabajo de Silver sobre esta cuestión, que se basó en las encuestas políticas, es su relación con otra fuente de datos: lo que la gente de los distintos estados dijo a Gallup sobre su sexualidad. Aquí están esas cifras que declararon en un gráfico, cotejadas con las proyecciones de Silver sobre la aceptación del matrimonio homosexual, estado por estado. He codificado a cada estado según su tratamiento legal del matrimonio homosexual y, además, he añadido las iniciales de unos cuantos de los que presentan valores más atípicos:



Según el eje horizontal vemos que, para Silver, Misisipi es el estado menos tolerante y que Rhode Island es el más tolerante. En cuanto al eje vertical, los números de Gallup oscilan entre el 1,7 por ciento de Dakota del Norte y el 5,1 por ciento de Hawái. Y, como puede apreciarse en la inclinación de la línea de tendencia, cuanto más acepta un estado la homosexualidad, más elevado es el índice de su población que se declara gay. Resulta notable que si seguimos esa línea punteada hasta llegar al cien por cien de apoyo al matrimonio gay (imaginando estadísticamente un futuro de total tolerancia), veremos que implica que alrededor del 5 por ciento de la población diría que es gay, al no existir presión social alguna para que no lo dijese. Esa es la misma cifra que extrajimos de las búsquedas de Google, donde la ausencia de presión social no es solo teórica.

Es más, esa línea de tendencia no representa una función de tios que se limitan a vivir donde mejor se los acepta. La regularidad que se aprecia entre estados en las búsquedas de porno gay nos da pruebas de ello, así como los datos de movilidad de Facebook. Si comparamos las ciudades natales de usuarios homosexuales con sus actuales lugares de residencia veremos que la reubicación solo explica una pequeña fracción de la varianza en los índices de

homosexualidad de Gallup que hemos visto antes. Los gais no se trasladan de manera masiva a sitios más tolerantes. Por una parte, esto es un testimonio de la fortaleza de los vínculos con el hogar, de la crianza y de la simple inercia. Por otra, significa que por cada persona que coge sus bártulos y se larga a San Francisco o Nueva York para vivir plenamente, lo más probable es que haya decenas que siguen viviendo en la autonegación.

Si aceptamos estos dos cálculos independientes del 5 por ciento, a los que se ha llegado a través de tres de las mayores potencias modernas en lo que a datos se refiere —Nate Silver, Google y Facebook, con la ayuda de ese superviviente de la vieja escuela de las encuestas que es Gallup—, empezaremos a ver con otros ojos esos números declarados por los propios sujetos. Cuando Gallup nos dice, por ejemplo, que el 1,7 por ciento de los habitantes de Dakota del Norte son gais, entonces es que posiblemente algo así como el 3,3 por ciento de ese estado es gay y no quiere reconocerlo. En Nueva York, cerca del 4 por ciento de la población es abiertamente gay, lo que deja tal vez un 1 por ciento de gais que no lo cuentan. Y lo mismo en los demás estados. Ante esa uniformidad de los datos, los altibajos que vemos en las declaraciones de población gay adquieren un nuevo significado: nos muestran un país de estadounidenses que viven vidas secretas. Esto aporta conocimiento específico a la cita poética que se suele atribuir a Thoreau: «La mayoría de los hombres llevan una vida de callada desesperación y se van a la tumba con la canción todavía sonando en su interior». Estos son refugiados del alma, y lo vemos claramente en los datos.

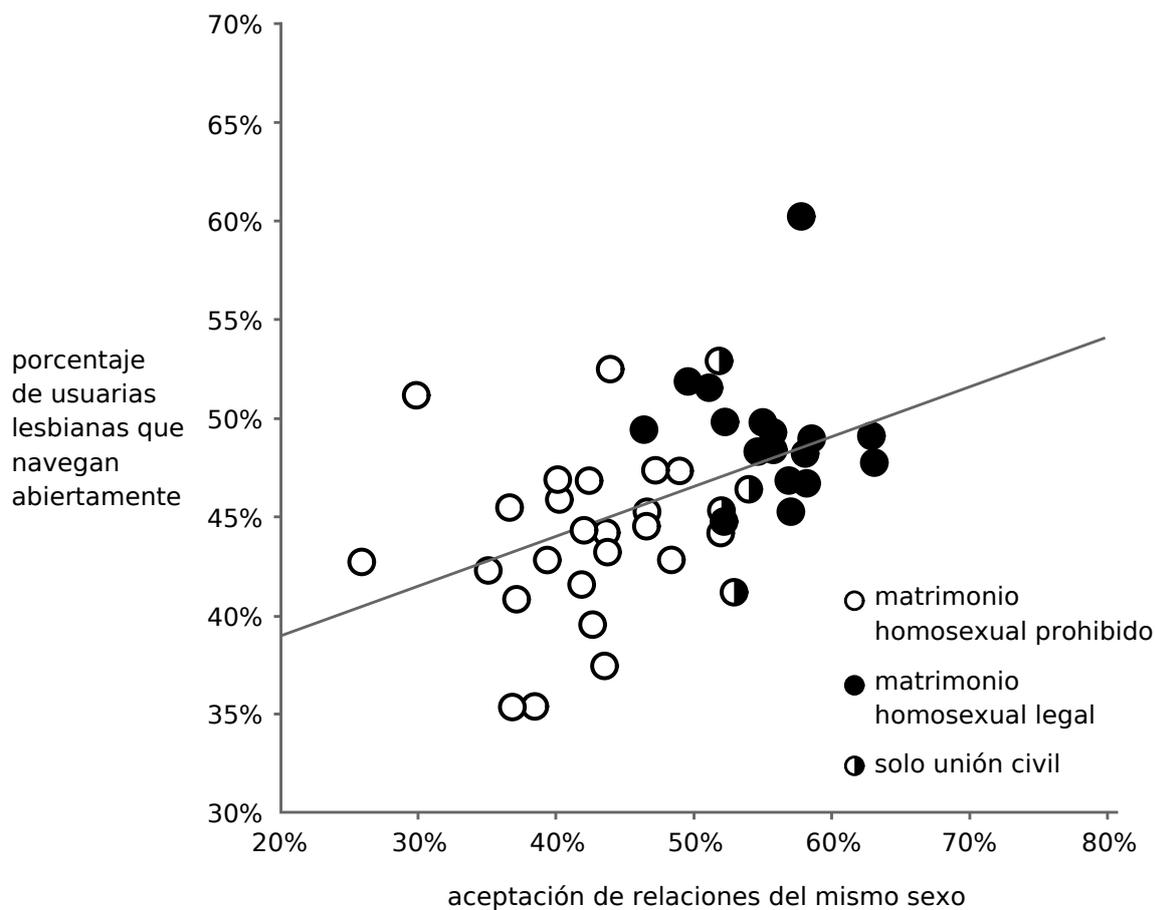
Los datos nos brindan un panorama de los daños colaterales. He aquí de nuevo a Stephens-Davidowitz:

En Estados Unidos, a todas las búsquedas en Google que empiezan con «¿Es mi marido...» la palabra que más comúnmente les sigue es «gay». El término «gay» es un 10 por ciento más habitual en esas búsquedas que la palabra que ocupa el segundo puesto, «infidel». Es ocho veces más frecuente que «alcohólico» y diez veces más que «depresivo».

Y esas búsquedas interrogativas son más habituales allí donde la represión es más acentuada: Carolina del Sur y Luisiana, por ejemplo, muestran los índices más elevados, y la aceptación del matrimonio homosexual está por debajo de la media nacional en 21 de los 25 estados en los que esta búsqueda es más frecuente. Uno se pregunta qué pensarán de estos datos aquellos que tanto se esfuerzan en mantener oculta la homosexualidad (o «curarla»), o de los

matrimonios sin sexo y los niños con padres infelices que esos esfuerzos suyos tan claramente propician. Insisto en que esto no es retórica: son números. El antiguo «índice de miseria» económico consistía en inflación + desempleo. En mi opinión, su versión social la compone la parte de la población que vive en lugares donde no pueden ser ellos mismos. Se trata de una situación que lo único que aporta es sufrimiento^[30].

Por desgracia, las búsquedas de Google no son eficaces para calcular el número de lesbianas que hay en el país. Los muchos hombres heteros que buscan porno de mujeres con mujeres embrollan los datos. No obstante, podemos apreciar sombras de los cálculos de aceptación de Silver en los datos de OkCupid, con algunos giros interesantes. Calculo que más de una cuarta parte de la población gay del país que busca pareja ha usado OkCupid en 2013^[31]. Los que ligan por Internet deberían ser por lo general más abiertos que la media sobre su sexualidad; al fin y al cabo, ponen su perfil en una web a la vista de todos. No obstante, sabedores de que mucha gente prefiere no difundir su identidad sexual en Internet, OkCupid brinda a sus usuarios gais la posibilidad de «ocultar» su perfil a todos los demás salvo a otros usuarios gais. El 59 por ciento de los hombres gais y el 53 por ciento de las mujeres lesbianas han aprovechado esa opción. También en estos datos es visible la correlación entre la tolerancia de determinado estado y el grado de apertura, aunque más en el caso de las mujeres, a las que he representado en el siguiente gráfico:



Si dejamos de lado la cuestión de la «salida del armario», los usuarios homosexuales de OkCupid son muy parecidos a los demás. En las preguntas para buscar coincidencias, los usuarios gais del sitio muestran los mismos índices de consumo de drogas, prejuicios raciales y excitación que los heteros, y los gais buscan el mismo tipo de relaciones. De hecho, en cuanto a la actitud sexual, si hay un grupo que se sale del montón es el de las mujeres heterosexuales. En comparación, son más mojigatas: el 6,1 por ciento de hombres heteros, el 6,9 por ciento de hombres gais y el 7 por ciento de lesbianas están en OkCupid buscando explícitamente sexo ocasional. Solo el 0,8 por ciento de mujeres heterosexuales lo busca, lo que probablemente dice más acerca del tabú que existe en contra de la iniciativa sexual de las mujeres (heterosexuales) que cualquier otra cosa^[32].

La cifra de parejas sexuales declaradas a lo largo de la vida es básicamente la misma en los cuatro grupos. La media de los hombres gais y de las mujeres heteros es de cuatro; la de las lesbianas y los hombres heteros es de cinco, pero a duras penas^[33]. Si existe una diferencia significativa en la conducta sexual, está en lo más extremo, donde encontramos que se cumple en parte un estereotipo. Los hombres gais sumamente promiscuos (el montón

que ha declarado haber tenido 25 parejas o más) superan en número a sus contrapartes femeninas heterosexuales en relación de dos a uno. Curiosamente, en el sexo, como ocurre con la riqueza o el lenguaje, tenemos un problema de desigualdad. Según estos datos, el 2 por ciento de hombres gays del extremo superior practica alrededor del 28 por ciento del sexo gay total.

Para ver cómo se configuran identidades alrededor de las etiquetas «gay» y «hetero», podemos aplicar el método del «rango de palabras» que hemos empleado en el último capítulo para investigar las descripciones personales que hacen los propios usuarios. Como antes, los textos de los perfiles nos dan una idea de lo que hace único a cada grupo frente a los demás: lo que tienen de especial las lesbianas, lo que diferencia a los hombres gays de los heteros, etcétera, y el método se basa únicamente en las propias palabras de los usuarios. Los datos de comportamiento que hemos visto antes nos muestran que no es tan diferente «cómo» amamos, pero a continuación veremos que sí lo es, naturalmente, «a quién» amamos. Los números hacen que salga a la luz el vocabulario más típico de cada grupo:

palabras más típicas de...

hombres gays	mujeres lesbianas	hombres heteros	mujeres heteros
first wives	i am gay	knows what she wants	honest man
velvet rage	old lesbian	i have no kids	man to share
tales of the city	i'm a lesbian	treat a woman	to meet a man
you're nice guy	i am a lesbian	care of herself	a man who knows
anything on bravo	femme side	never been married	care of himself
music madinna	attracted to women who	daughter family	meet a man who
music britney	lesbian friends	for a good woman	find a man who
ltr oriented	are femme	treat a lady	who knows what he
romy and michelle's	butch femme	good women	meet a man
new guys	lesbian movies	my kids my family	man who knows how
barefoot contessa	single lesbian	hello ladies	a nice guy who
kathy griffin	u haul	type of girl	honest guy
single gay	butch but	woman that can	a man who has
the comeback	are feminine	real woman	are a nice guy
hiv positive	femme who	my son family	christian man
density of souls	elena undone	woman to share	like a man who

modern family glee	the butch	my daughter family	a guy who has
ab fab	not butch	intelligent woman	man that knows
most gay	movies imagine	god my kids	love jesus
muriel's	music brandi	girl that i can	a man who will
christopher rice	walls could	meet a woman who	man that has
muriel's wedding	lesbian romance	have no children	true gentleman
other gay	femme women	son family	you are a gentleman
flipping out	debs	with the right woman	guy to share
find mr	feminine women	treat her	nice guy who
sordid lives	soft butch	great woman	a guy that can
stereotypical gay	my future wife	a woman who can	christian woman
flight attendant	hunter valentine	nice woman	for a good guy
are you there vodka	lesbian looking	i like a woman	you're a gentlemen

Como antes, te dejo que escudriñes al detalle las palabras de los usuarios y me voy a limitar a señalar unas cuantas tendencias generales. Las dos listas de heterosexuales se centran decididamente en la (potencial) pareja de la persona. Todas y cada una de las entradas de las mujeres heterosexuales se centran en el hombre que están buscando (aquí cuento a Jesús, ya que es soltero), y la única en la que los hombres se apartan un poco de hablar de las mujeres es cuando apuntan a la presencia o ausencia de hijos. Estas listas, en conjunto, son como una versión más larga del «Yo Tarzán, tú Jane». O como si las hubiese adaptado el escritor Nicholas Sparks.

La lista de las lesbianas es más introspectiva, con más autodescripciones, pero aun así sigue siendo bastante parecida a las de los heteros. Como las mujeres heterosexuales, las lesbianas aparecen sumamente tipificadas por la relación que están buscando («si eres femenina», «mi futura esposa»); sencillamente emplean palabras distintas.

La lista de hombres gays es muy distinta de las otras tres. Está plagada de referencias a la cultura popular y presenta bastantes menos referencias a la pareja o a la familia del usuario. Lo de «Anything on Bravo» debe de ser la generalización más acertada de todos los tiempos. Dicho esto, resulta interesante que los hombres gays sean los menos centrados en el sexo y la identidad sexual de los tres grupos. O, mejor dicho, que obtienen su identidad de otras cosas que no son el sexo.

Repito que este método tiene la finalidad de acentuar las diferencias entre los grupos, pero otros datos nos muestran que los límites son permeables. Una

de las cosas más intrigantes que he encontrado en OkCupid es la respuesta a esta pregunta para hallar concordancias, que responden únicamente los usuarios del sitio que se identifican como heterosexuales:

P: ¿Has mantenido alguna vez un encuentro sexual con alguien de tu mismo sexo?

	mujeres		hombres	
Sí, y me gustó.	22 308	26 %	12 070	7 %
Sí, y no me gustó.	6153	7 %	10 100	6 %
No, pero me gustaría.	14 896	17 %	7632	5 %
No, y nunca lo haría.	42 286	49 %	137 455	82 %
	85 643		167 257	

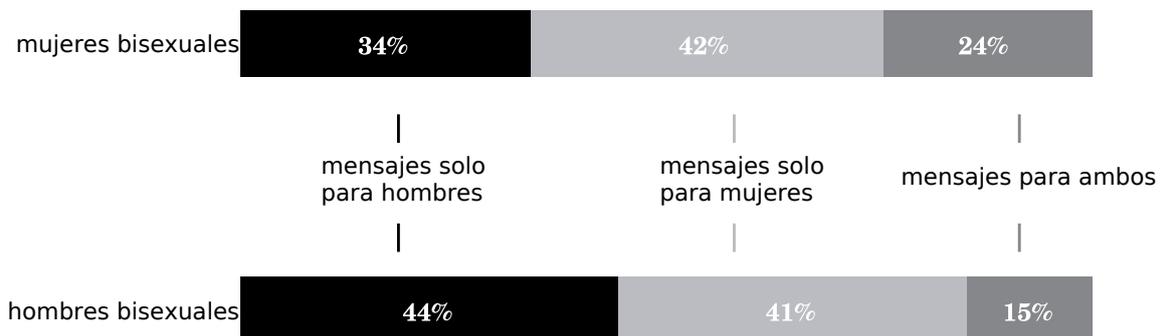
O sea, que el 51 por ciento de las mujeres y el 18 por ciento de los hombres han tenido o les gustaría tener una experiencia con alguien del mismo sexo. Esas cifras son mucho más altas que cualquier cálculo plausible de la población homosexual real, de modo que no solo descubrimos que la sexualidad es más fluida de lo que logran clasificar las categorías de un sitio web, sino que además observamos que el sexo con una persona del mismo género es relativamente habitual, independientemente de que la gente lo considere parte de su identidad o no.

Los datos de la tabla anterior son de usuarios que seleccionaron «heterosexual» al registrarse, pero en ese mismo menú desplegable de OkCupid se les sugiere la opción de «bisexual». Cerca del 8 por ciento de las mujeres y el 5 por ciento de los hombres la seleccionan. He observado, tanto en OkCupid como en otros sitios, que a los bisexuales les causa mucha frustración el hecho de que no se considere la bisexualidad como una orientación; que, por ejemplo, a los hombres bisexuales se los considere gays que todavía no lo han asumido. Mucha gente homosexual ve la bisexualidad como una actitud evasiva. Un reciente estudio de la Facultad de Salud Pública de la Universidad de Pittsburgh lo explica muy claro, aunque algo fríamente: «Los encuestados que se identificaron como gays o lesbianas respondieron de manera significativamente menos positiva hacia la bisexualidad [...] lo que indica que incluso en el seno de la comunidad sexual minoritaria los bisexuales se enfrentan a un profundo estigma».

Gerulf Rieger, de la Universidad de Essex, en colaboración con psicólogos de Northwestern y Cornell, concluyó en un artículo científico de 2005 que en términos de reacción genital al estímulo, casi todos los hombres

que se declararon bisexuales eran gais, algunos eran heterosexuales y muy pocos se excitaban físicamente por ambos sexos. Por lo tanto, definió la bisexualidad masculina como un «estilo» de interpretar la excitación más que como una excitación por sí misma. Comprensiblemente, esto indignó a la comunidad bisexual; Rieger revisó el asunto al cabo de un tiempo y concluyó que la bisexualidad masculina debe de ser «una cuestión de curiosidad», que «el interés por ver a otros desnudos, por ver a alguien mientras practica el sexo, por ver películas pornográficas o por participar en orgías» explicaba la aparente desconexión entre la atracción por ambos sexos que declaraban los bisexuales y la atracción física observada en ellos por uno solo de los sexos. Su mente disfrutaba de todo tipo de sexo, pero su cuerpo ya discriminaba más.

En OkCupid encontramos cierta corroboración del espíritu de las conclusiones de Rieger, si bien no de su vaga terminología. La gran mayoría de los hombres y mujeres bisexuales de la web buscan exclusivamente a personas de un sexo u otro. A continuación muestro a quién manda mensajes en realidad la gente que se identifica como bisexual:



Para estar en cualquiera de las franjas de «mensajes solo para», un usuario debe haber contactado al menos el 95 por ciento de las veces con personas de ese sexo, así que el umbral que tenemos aquí es bastante alto; no se trata de ningún cambalache de cifras. Solo una pequeña fracción de la base de usuarios bisexuales establece contacto de manera significativa con ambos sexos. Sea cual sea el mecanismo empleado, la afirmación de Rieger de que la bisexualidad declarada no concuerda con la conducta observada parece ser correcta en este caso. Curiosamente, en el caso de los hombres, los mensajes van cambiando con la edad. En ese cambio observamos pruebas plausibles del argumento de que es una actitud evasiva: más de la mitad de los hombres bisexuales más jóvenes envían mensajes únicamente a otros hombres, y ese porcentaje va disminuyendo paulatinamente hasta que llegan a la treintena,

punto en el que la mayor parte de la base de usuarios bisexuales mandan mensajes únicamente a mujeres. Eso es lo que cabría esperar si los hombres interesados en otros hombres dejaran de identificarse a sí mismos como bisexuales al ir haciéndose mayores y empezasen a sentirse más cómodos si se les llama «gais». Pero para responder a esta cuestión sería necesario disponer de datos longitudinales, que no tenemos todavía.

Dicho lo cual, quiénes decimos que somos y cómo nos comportamos son dos cosas distintas, y la segunda no debería descalificar automáticamente a la primera. La gente es absolutamente libre de definirse a sí misma como le plazca, y pedirles que sus etiquetas cumplan a rajatabla la definición de un investigador (o de una web) no tiene ningún sentido. La culpa de cualquier discrepancia, de todos modos, la tiene la etiqueta: los individuos aman de la manera que les parece adecuada, y a veces las palabras que la definen no están al día. El día de San Valentín de 2014, por ejemplo, Facebook propuso en su web más de cincuenta opciones de género distintas (que permitían a los usuarios definirse como «transgénero» o «andrógino» en lugar de masculino y femenino). Ellyn Ruthstrom, presidenta del Bisexual Resource Center de Boston, hablaba de la orientación y el trabajo de Rieger, pero podía haber estado hablando perfectamente de mis datos, cuando declaró al *Times*: «Esto, por desgracia, reduce la sexualidad y las relaciones a una mera simulación sexual. Los investigadores quieren encajar la atracción bisexual en una cajita pequeña: tienes que sentirte atraído exactamente igual por hombres y mujeres para ser bisexual. Eso es una estupidez. Lo que me gusta es que la gente expresa su bisexualidad de maneras muy distintas».

Lo cierto es que esta variedad de expresiones la encontramos cuando nos fijamos en las palabras «típicas» que aparecen en los perfiles de los hombres bisexuales de OkCupid. Entre las primeras treinta están «bisexual», «pansexual», «travestido» y «heteroflexible». En la lista de sus antítesis encontramos «muy unido a mi familia» y «me gusta de verdad mi trabajo», indicadores tal vez de la soledad y la insatisfacción que conlleva sentirse diferente, incluso entre otros que también son diferentes.

La bisexualidad en las mujeres es un caso algo distinto. Es más convencional, o al menos lo es la versión que venden Miley Cyrus y demás. Tal vez porque las gentes de *marketing* saben que «el sexo vende» y que las estrellas tienen que traspasar los límites, el caso es que en la actual cultura popular es bastante común ver una especie de moda del «gay de pago». En el caso de Miley —aunque, naturalmente, no puedo saberlo con certeza— parece ser un disfraz para vender discos, apenas distinto de la pintura facial de

Gene Simmons. También ocultos por un disfraz, hay estafadores que van en pos de hombres en Internet que muchas veces eligen la bisexualidad como identidad en sus cuentas falsas. En Facebook, el 58 por ciento de los perfiles falsos son de «mujeres bisexuales», mientras que las cuentas que no son falsas suman en total solo el 6 por ciento. En OkCupid el problema no es tan acentuado, pero seleccionar la bisexualidad junto con unos cuantos indicadores claves más te garantiza que los administradores de la web te dediquen una revisión especial.

Pero incluso en nuestros perfiles legítimos, que son casi todos, la bisexualidad femenina y la fantasía heterosexual masculina están vinculadas. Esto se aprecia con claridad en los datos cuando nos fijamos en el texto de los perfiles: sobre todo se trata de mujeres que invitan a todo el mundo a participar en tríos con sus novios o maridos.

palabras más típicas de...

mujeres bisexuales

bi female

bisexual female

me and my husband

me and my man

my boyfriend is

hubby and

we are a couple

i am bisexual and

me and my boyfriend

fun couple

couple we

married couple

we are not looking

fun with me and

do have a boyfriend

my bf and

female to join

girl to join

another couple

bi woman

my boyfriend my
i am bi sexual
my hubby and
join me and my
female for
my boyfriend and i
we are looking to
a triad
no single
send us

Si a esta lista pudiera ponerle un ritmo y consiguiera que Pitbull me la cantase al compás, llegaba seguro al número uno. Dicho esto, pese a toda la zafiedad de vincular identidad sexual con planes de negocio, es buena señal que la mayoría social considere que vale más apropiarse de una identidad minoritaria que suprimirla. De hecho, observamos que las cosas están cambiando en lo que respecta a la sexualidad, y con rapidez. Mientras concebía las predicciones que hemos visto antes, Nate Silver cronometró un acentuado cambio en las actitudes de los estadounidenses durante la pasada década. La aceptación del matrimonio homosexual se aceleró ostensiblemente en 2004, y Silver declaró: «Ya no hace falta hacer suposiciones optimistas para determinar que quienes apoyan el matrimonio con personas del mismo sexo pronto constituirán una mayoría nacional».

Así pues, todo vuelve a reducirse al recuento, y el porcentaje cada vez nos es más favorable. Aunque siempre ha habido gais, a finales del siglo XIX la gente empezó a proclamar voluntariamente su homosexualidad a modo de acto político. La frase «salir del armario» se acuñó pocos años después. Ahora, el objetivo de vivir y amar abiertamente, que los hombres y las mujeres homosexuales han perseguido durante tanto tiempo, prácticamente se ha cumplido. Ese cambio se ve encarnado en los famosos que han salido del armario, pero más aún en los millones de personas cuyos nombres nunca conoceré pero que han contribuido a que las tasas de aceptación hayan ido subiendo puntos poco a poco. Llegará el día en el que los encuestadores puedan dejar sus bolígrafos, los científicos puedan apartar sus lentes y los estudiantes emprendedores puedan emplear sus algoritmos para calcular otras cosas. Llegará el día en el que el mundo será tan abierto que nadie tendrá que especular sobre los demás.

12.

Saber cuál es tu sitio

Cuando iba a secundaria, la pausa para comer era muy larga, y como a esa edad todos éramos demasiado mayores para ponernos a jugar o para pasarlo bien, después de comer nos limitábamos a plantarnos fuera de la escuela a esperar que sonase el timbre que anunciaba la reanudación de las clases. Los primeros días de clase nos fuimos organizando en la acera asfaltada y, una vez establecido, ese orden apenas varió en tres años. Tal como lo recuerdo, el orden, desde los que estaban más cerca de la puerta de la cafetería a los que estaban más lejos, era el siguiente:

- los ultramolones (sobre todo los de los Heights, que era el barrio más rico de la ciudad)
- los pijos en general
- los que escuchaban a REM/The Cure en las radios universitarias (era la época anterior al indie)
- los skaters
- los heavys (los que llamábamos «colgaos del metal», y todo aquel para quien el pegamento no era solo un adhesivo)
- mis amigos y yo
- UN CONTENEDOR MARRÓN ENORME
- los estudiantes de intercambio y los chavales con dificultades de aprendizaje

Obviamente, esta organización no tenía nada de casual. El contenedor, bendito sea, constituía un punto de reunión natural para los intocables, y a partir de ahí ya todo dependía de la polaridad positiva/negativa de las moléculas de cada alumno. Puesto que en un extremo de la fila los míos jugaban a ver quién le partía antes un lápiz a otro y discutían sobre las bondades de las Tortugas Ninja —el juego de rol, no la serie de la tele, porque la serie de la tele es para niños—, todos los demás fueron encontrando su sitio de manera más o menos natural.

Una de las cosas bonitas de los datos digitales, además de su ingente volumen, es que, al igual que en el patio del instituto Pulaski Heights, tienen una dimensión física pero también social. Una hoja de papel tiene dos dimensiones, el espacio-tiempo tiene cuatro. La teoría de cuerdas predice que

nuestra existencia física exige que haya entre 10 y 26 dimensiones. Nuestro universo emocional tiene esas y más, seguro. Y combinar estos espacios — nuestro paisaje interior y el mundo exterior— nos permite retratar la existencia con una profundidad nueva.

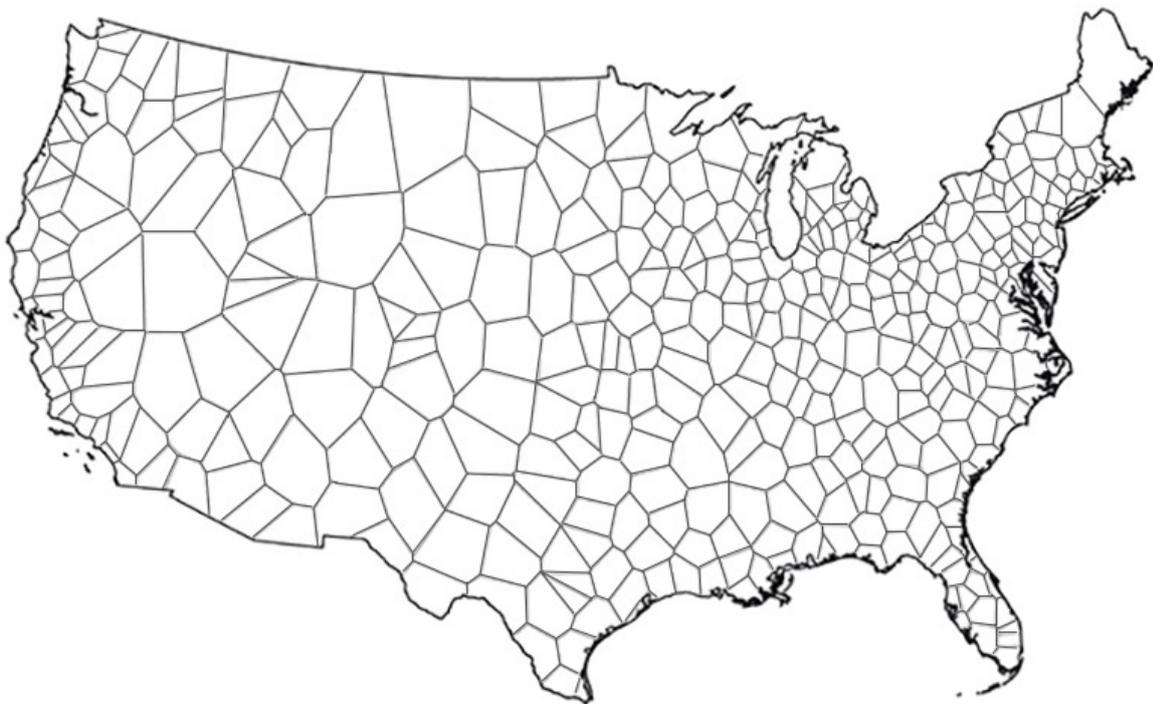
Hasta ahora hemos mirado a la gente y sus interacciones —contactos, texto del perfil, valoraciones, etcétera— ignorando prácticamente el emplazamiento físico, pero los sitios web y los *smartphones* recopilan muchos datos de localización, por supuesto. Los tuits se geolocalizan mediante latitud y longitud, Facebook te pregunta cuál es tu ciudad natal, la ciudad donde fuiste a la universidad y tu ciudad de residencia actual, muchas apps registran hasta el propio edificio en el que te encuentras en cada momento. Aquí vamos a superponer identidad, emoción, comportamiento y opinión a los espacios físicos que ocupamos y a ver qué nuevos conocimientos surgen de ahí. Nos fijaremos en cómo afecta la ubicación a la persona y en cómo ha trazado la gente nuevas fronteras en nuestro viejo planeta.

Las fronteras de muchas comunidades se fijaron por decreto o por accidente, o por ambas cosas. Estados Unidos y la URSS dividieron Corea por el paralelo 38 porque esa línea destacaba en un mapa de un ejemplar de *National Geographic* que tenía un funcionario. Ese mismo mes, un poco antes, Alemania fue dividida en zonas de ocupación que reflejaban, más que otra cosa, qué tropas de qué países estaban destacadas en cada sitio en aquel momento. Muchos de los propios estados de Estados Unidos se crearon por decreto real o por una ley del Congreso, y sus límites los trazaron personas que nunca llegarían a pisar esas tierras. Esa cartografía no presencial fue y es todavía un problema mucho más pernicioso en África, en el subcontinente indio, en Oriente Medio y en todas aquellas tierras donde el imperio estampó su huella. Solo en muy contadas ocasiones se han trazado mapas que reflejan la «voluntad del pueblo», e incluso en dichos casos, como hemos visto en Israel, cuya historia moderna comenzó oficialmente como Mandato británico de Palestina, la pregunta obvia es: qué pueblo y la voluntad de quién.

En el caso de los sitios web, las fronteras políticas y naturales son otra serie de datos que hay que tener en cuenta. Cuando la información —fluida, sin restricciones y abstracta— es tu moneda de cambio, el mundo físico y sus muchos límites arbitrarios suelen ser más que nada un incordio. En OkCupid, los ríos son un engorro perpetuo para los algoritmos que buscan coincidencias por distancias. Queens está a la vez a un kilómetro y a un mundo de distancia de Manhattan. Intenta explicarle eso a un ordenador. El problema es que

cuando una persona está conectada, él o ella está a la vez en el mundo y fuera de este. Pero esa dualidad también implica que podemos reinventar nuestros espacios físicos siguiendo nuevas líneas, tal vez dotadas de mayor significado que aquellas trazadas por la tectónica de placas o por los dictados de determinado pedazo de pergamino.

Este es un mapa de cómo divide el país la web Craigslist: cada una de las regiones corresponde al territorio que abarca determinada lista de anuncios clasificados de esa web. Un cartógrafo lo llamó «Estados Unidos de Craigslist», pero me da la sensación de que «unidos» no es la palabra idónea; esto es una división y, dentro del conjunto, cada una de las pequeñas zonas es un reino en miniatura. Es una especie de Sacro Imperio Romano de los muebles de segunda mano.



A medida que empezamos a insertar contenido en los espacios, el mapa se va volviendo más interesante. A continuación está el imperio de Craigslist otra vez, sobre el que se han superpuesto las localizaciones más populares que aparecen en los numerosos tablones de «Contactos perdidos» de la web, donde un corazón solitario puede publicar cosas como esta:

Ambos subimos al tren Q en la calle 34. Tú llevabas un chaquetón marinero y tenías en los ojos un centelleo como de Audrey Hepburn. Nuestras miradas se cruzaron unas cuantas veces. Si lees esto, envíame un correo.

Esa es la versión de uno de Manhattan, por lo menos. Los de Portland suelen cruzar las miradas en el autobús. Los californianos flirtean junto a las máquinas elípticas. Pero para casi todo el resto del país, el lugar de los deseos incumplidos es Walmart.

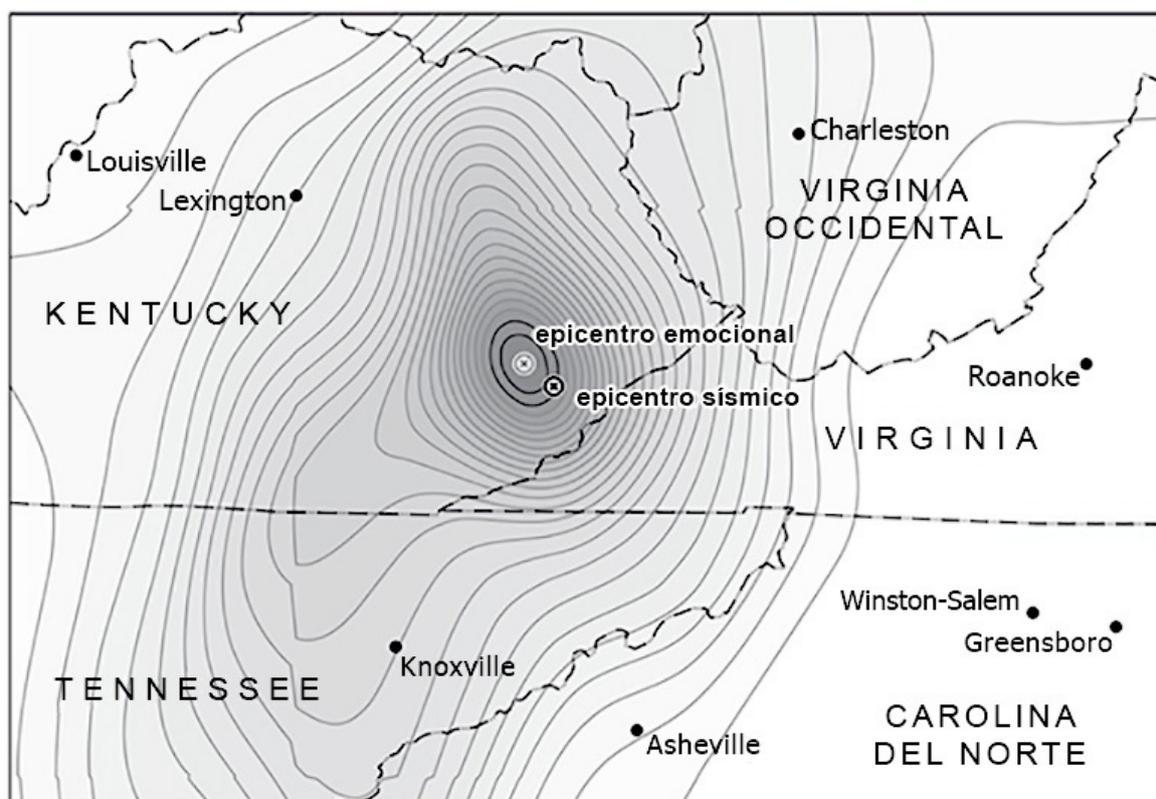


Ahora ya vamos llegando a un punto al que ningún cartógrafo tradicional puede llevarnos y que ningún satélite puede fisgonear. Lo de arriba no es más que una página sencilla y hasta torpona de una nueva clase de atlas: la unión de los terrenos geográficos y del comportamiento.

En los ejemplos anteriores, Craigslist delimitó sus fronteras *a priori* y lo hizo escogiendo aquellos mercados a los que quería dirigirse. La mayoría de las webs recopilan datos en lugar de proyectarlos, y a partir de estos podemos generar un mapamundi verdaderamente alternativo, mover las fronteras y los contornos para adaptarlos al paisaje humano. Años atrás, un *hacker* estadounidense se dedicó a reunir datos de Facebook y elaboró un gráfico con los contactos compartidos de los 210 millones de perfiles que había recopilado. A partir de los datos que tenía, dividió el país en unos estados ficticios definidos más por la amistad que por la política. Eran siete: Pacífica (la zona noroeste/Pacífico), Socialistán (California), Mormonia, el Oeste Nómada, el Gran Texas (que incluía Arkansas, Oklahoma y Luisiana), Dixie (en el sureste) y por último, formando una franja de un verde intenso que iba

desde Minnesota hasta el Atlántico, pasando por Ohio y abarcando toda Nueva Inglaterra, Caserolandia. Ese es mi tipo de país.

Desde entonces, los *smartphones*, equipados todos ellos con su minúscula señal de GPS, han revolucionado la cartografía. Matthew Zook, geógrafo de la Universidad de Kentucky, ha creado conjuntamente con especialistas en datos lo que denominan el proyecto DOLLY (acrónimo de *Digital OnLine Life and You*, es decir: vida digital *online* y tú). Se trata de un repositorio en el que se pueden buscar todos los tuits geolocalizados desde diciembre de 2011, lo que supone que Zook y su equipo han recopilado miles de millones de sentimientos escritos, todos con su correspondiente longitud y latitud. DOLLY es un recurso increíblemente versátil, cuyas posibilidades justo ahora están empezando a vislumbrarse. El mismo Zook ya le ha dado unas cuantas aplicaciones sumamente personales. En febrero de 2012 su oficina de Lexington sufrió la sacudida de un terremoto y él se asomó a la base de datos para observar cuáles fueron las réplicas emocionales. El siguiente mapa muestra la densidad de las reacciones que se produjeron en Twitter, superpuestas al epicentro geográfico del seísmo. Aquí podemos apreciar los contornos de la sorpresa generada por el movimiento de la tierra:



Zook descubrió que el epicentro emocional del terremoto quedaba un poco al noroeste del epicentro sísmico, en Hazard, Kentucky. Y, por muy simple que pueda parecer, este tipo de descubrimiento es verdaderamente novedoso. Los mapas de Craiglist, por ejemplo, podrían haberse elaborado en la década de 1970; al fin y al cabo, la idea de la sección de «contactos perdidos» de esa web se copió de los periódicos. De modo que, antes de que existiese Internet, si uno hubiese querido, podría haber recopilado los anuncios clasificados de todo un mes del principal periódico, pongamos, de las cien ciudades más importantes del país, procesar los datos y obtener algo muy parecido a lo que hemos visto unas páginas atrás. Incluso esa redefinición geográfica Facebook/Caserolandia era teóricamente posible hace décadas, siempre que se contase con un equipo de encuestadores para entrevistar a millones de personas en sus casas y hacer un seguimiento de los contactos que estos declarasen tener.

Pero el mapa de Zook nos muestra la reacción instantánea de la gente ante un acontecimiento que duró un instante. Si, aun con infinito esfuerzo, hubiese logrado encuestar después a los habitantes de Kentucky, no podría haber generado un informe realista; no es solo que las emociones cambien cuando las recordamos, sino que la cobertura de los medios y todo lo que se hubiera hablado sobre el terremoto habrían contaminado sin remedio los datos. Tampoco es que la gente con *smartphone* haga que los sismógrafos queden obsoletos, pero el gráfico de Zook refleja el «impacto» del terremoto de un modo mucho más directo que la vieja escala de Richter. En caso de no saber nada más acerca de un terremoto, si hubiese que distribuir ayuda a las víctimas, los contornos de la reacción de Twitter constituirían una guía mucho mejor que las tradicionales ondas expansivas que se ven alrededor del epicentro en un mapa sísmico^[34].

Pese a que todos ellos son transitorios, los tuits que se recopilan pueden capturar más que acontecimientos efímeros. Un vídeo en YouTube que demuestra el poder de DOLLY nos enseña el seguimiento que se hizo de la festividad holandesa de Sint Maarten, una especie de Halloween germánico en el que los niños van cantando puerta por puerta para pedir golosinas. En los datos se aprecia que la gente no solo lo celebra en los principales núcleos de población del norte de los Países Bajos, como cabría esperar, sino también en la parte occidental de Bélgica, donde los tuits reconectan a la vieja Holanda con Flandes, su prima hermana cultural. De ese modo logramos una visualización animada de puntos geolocalizados mediante GPS y vislumbramos la sombra de los Habsburgo.

Vistas las posibilidades de lo que ya podemos lograr gracias a un *software* como DOLLY, la falta de datos longitudinales se hace especialmente dolorosa. En el actual corpus de búsquedas, el tiempo se echa en falta como si se tratase de un miembro amputado fantasma. Twitter ya nos proporciona ahora bastante de esa promesa multidimensional: tenemos todas las emociones y todos los puntos del planeta, pero solo tenemos unos cuantos años recientes con los que trabajar. En Europa, donde la combinación de geografía, cultura y lengua lleva siglos siendo una cosa muy volátil, imaginemos que pudiésemos hacer el seguimiento de Alsacia-Lorena y todos sus cambios de manos —ahora alemana, ahora francesa, ahora alemana, ahora francesa—, con cada nuevo gobierno que impuso su cultura a la gente, como si la región fuese una casa a la que se le dan sucesivas manos de pintura. O pensemos en el Caribe de finales del siglo xv y que fuésemos capaces de ver primero a los soldados que llegaron, luego su religión, luego su lengua que dominó a todas las de la zona, desde las lenguas arawak a las aztecas. DOLLY fue concebido para ver los flujos y las fracturas de una cultura a lo largo de décadas. Ahora lo único que nos hace falta son esas décadas^[35].

También podemos encontrar información geocultural en otras fuentes, y aunque en la mayoría de ellas perdemos la inmediatez que tiene Twitter, en su lugar obtenemos otro tipo de detalles. Cuando los sitios web plantean preguntas directamente a sus usuarios, no solo tenemos la posibilidad de afinar el trazado de las fronteras, sino de demostrar que en realidad no existen tal y como habitualmente las concebimos.

A continuación muestro el millón de respuestas a la pregunta «¿Debería ser ilegal quemar la bandera?» que hemos recopilado en OkCupid. En este caso, mi *software* cartográfico no ha dibujado límites políticos o naturales, sino que se ha limitado a distribuir las opiniones según su latitud y longitud. Este es ciertamente un país definido por sus principios o, mejor dicho, como se puede apreciar, dos países: el urbano y el rural. Se puede distinguir incluso la parte donde uno se inserta en el otro: las comunidades rurales que están río arriba siguiendo el Hudson y la zona vinícola del norte de California; construidas a base de dinero procedente de grandes ciudades, presentan también opiniones propias de una gran ciudad.

mapa de las respuestas a «¿Debería ser ilegal quemar la bandera?»

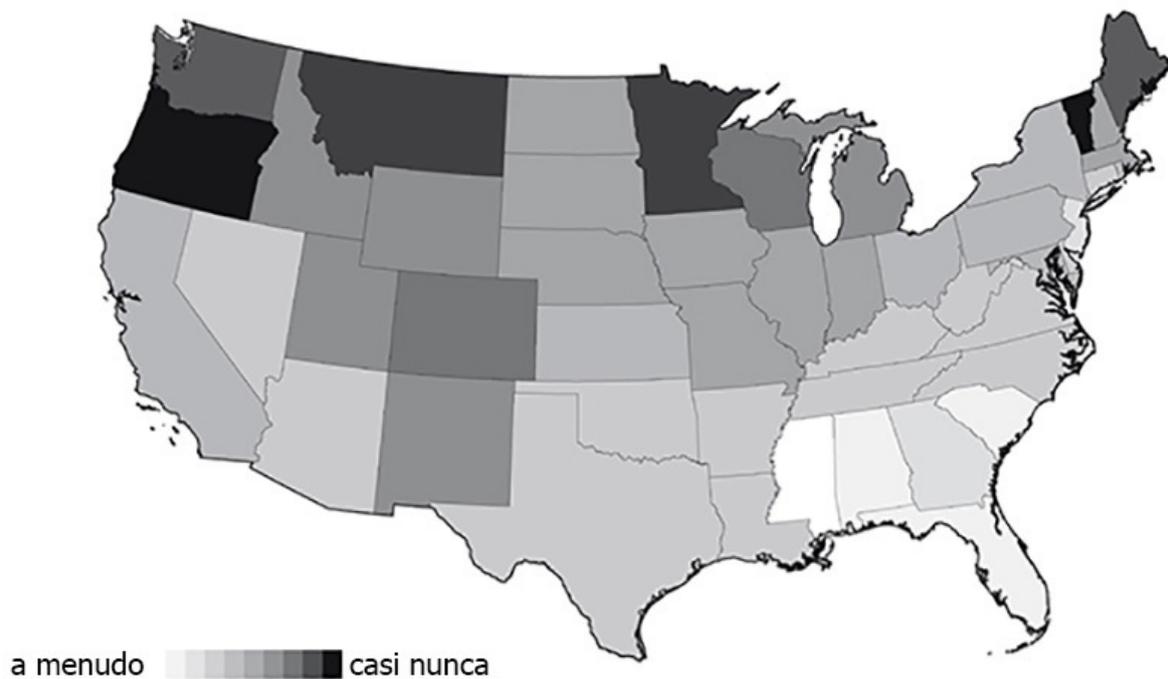


De forma parecida, y para confirmar la afirmación anterior de que según Google Tendencias la homosexualidad es universal, vemos que las búsquedas relacionadas con la homosexualidad no tienen fronteras ni estados ni países. Abajo vemos un mapa de las descargas de porno gay, según la dirección IP, extraídas de la mayor fuente de descargas de archivos torrent, The Pirate Bay. Este mapa también carece de cualquier línea divisoria predibujada, y a diferencia del anterior que hemos visto de OkCupid, su característica es la solidaridad: aquí vemos dónde vive la gente, desde Edmonton y Calgary a Monterrey y Chihuahua.



Hay tantas maneras de trazar mapas como fuentes de datos existen. Hasta ahora hemos ido abriéndonos camino lentamente, elaborando una dimensión psicológica —cómo nos sentimos acerca de la bandera, del porno...— a partir de nuestros mapas. Pero también se puede hacer a la inversa: los datos permiten partir de abstracciones e ir bajando hasta poner los pies bien firmes en el suelo. Tomemos como ejemplo la higiene, de nuevo a través de OkCupid. Esta es la frecuencia con la que dice la gente que se ducha:

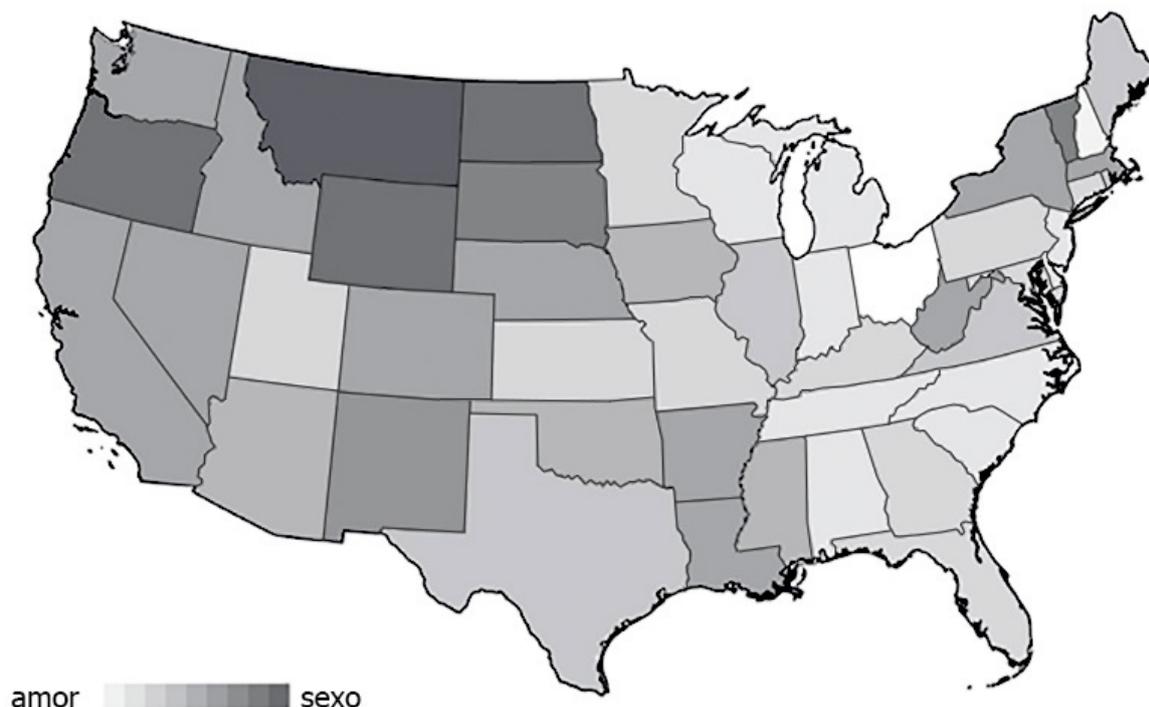
mapa de las respuestas a «¿Con qué frecuencia te duchas?»



Por una parte, la tendencia general refleja el clima: donde hace calor la gente se ducha más. Pero si nos remitimos a los detalles encontramos un par de buenas historias. El color claro del estado de Jersey lo podemos achacar a la obsesión de las estrellas televisivas Pauly D y The Situation por el gimnasio, el bronceado y la lavandería (Jersey es mucho más escrupuloso que los estados que lo rodean). Y en Vermont encontramos la filosofía opuesta: lo ecologista alternativo es más que un simple estereotipo. Vermont es el estado donde la gente se ducha menos en general, y verdaderamente destaca entre sus vecinos inmediatos. Según Google, el animal que tiene ese estado como símbolo es el caballo Morgan. Debería ser un tío blanco con rastas.

La política, el clima, Walmart y, ciertamente, los terremotos son cosas muy relacionadas con el mundo físico, pero en algunos de nuestros datos empezamos a vislumbrar una geografía exclusivamente interna. Pensemos en la lujuria, que en teoría no debería ser de ningún estado en particular. Pues aquí vemos que sí, y es uno que nos sorprende:

mapa de las respuestas a «¿Qué es más importante para ti ahora mismo, el sexo o el amor?»



Este patrón aparece una y otra vez en los datos de OkCupid: el centro septentrional y el oeste del país son sexualmente más abiertos, sexualmente más atrevidos y sexualmente más agresivos. Tal vez en la costa del Pacífico cabría esperar estas actitudes tan poco convencionales, pero no parecen pegar en el caso de muchos de estos estados consumidores de carne roja del centro. Los usuarios de OkCupid de las dos Dakotas, por poner un ejemplo, son políticamente tan conservadores como indica su reputación. El texto de sus perfiles no difiere mucho del de cualquier otro. Por lo que respecta a los demás indicadores, esos dos estados no deberían ser oscuros, pero en los datos apreciamos una misteriosa intensificación sexual. Este patrón inesperado nos desvela otro poder de los datos de Internet: ahora podemos descubrir comunidades que, más que reflejar la geografía, la trascienden.

Estos datos no demuestran que la región de la franja horaria central del país sea una enorme orgía multitudinaria en las praderas. De hecho, la explicación es bastante banal: si buscas a alguien con quien acostarte en un sitio como Pierre, Dakota del Sur, tus opciones locales son escasas. Así que pruebas en una web de contactos para encontrar lo que andas buscando. Es una simple tendencia selectiva de nuestros datos, pero no carece de significado: allí donde la gente no encuentra satisfacción en persona, crea comunidades digitales alternativas. En una web de contactos eso se traduce en comunidades con parecidos intereses sexuales. En otros sitios web de fines

más variados, donde los usuarios no se conectan solo para ligar en grupos de dos (y de vez en cuando, de tres), los datos son aún más ricos.

Reddit es la culminación de esa primitiva ambición de Internet: juntar a gente de sitios distantes para que hablen, compartan, discutan, difundan noticias y se rían, donde se suprime el espacio y se genera cercanía personal. Es uno de los sitios más concurridos de la Red^[36] y se autoproclama, con razón, «la portada de Internet»: muchas de las chorradas virales que circulan por los grandes sitios agregadores de contenidos proceden de allí. Mientras escribo esto, y no es broma, veo que hay un vídeo como más visitado en el *Huffington Post* que se titula: «Este ciervo pensaba que nadie estaba observando mientras se tiraba un pedo. Ahora todo el mundo lo ha visto». Te puedo asegurar que el primero en verle tirarse un pedo fue Reddit.

Lo que resulta raro es que, aun con toda la influencia que tiene, Reddit en realidad no hace nada. No tiene apps, ni juegos, ni perfiles como tales. Sus oficinas de Nueva York consisten en un espacio en un despacho compartido más pequeño que mi dormitorio. El propio sitio web no consiste más que en una serie de enlaces publicados por los usuarios, quienes votan, comentan y recomentan sobre los comentarios, modifican y redifunden, de tal modo que parecen el grupo de amigos más grande del mundo sentados en el sofá más grande del mundo. Son pocos los usuarios de Reddit que se conocen unos a otros por su nombre, por no hablar de que se conozcan en persona, pero sus vínculos no son menos estrechos a causa de ese anonimato: una mujer de 40 años de la bahía de San Francisco publicó que estaba sola para el Día de Acción de Gracias. El hilo de su publicación acumuló más de 500 comentarios en unas cuantas horas (incluidas, por supuesto, numerosas invitaciones para ir a comer al día siguiente) y se fue extendiendo a toda velocidad, tal cual, hasta conectar a usuarios de Reddit de muchas otras ciudades.

El sitio está organizado en forma de miles de subreddits temáticos. Cada uno de ellos lo crea y lo modifica el usuario, y todos cuentan con sus seguidores, comentaristas y publicadores devotos. Son lugares donde la gente ha creado verdaderas comunidades virtuales a partir de la nada, del espacio vacío. Encuentras cosas como juegos, tecnología, música o fútbol americano junto a multitud de temas de cosecha propia que solo encontrarás en Reddit, como en los ejemplos siguientes:

explícamelocomosituviesecinco años – «En el hinduismo y en el budismo, en los que los muertos se reencarnan, ¿cómo miden el crecimiento poblacional?»

soyun – «Soy un periodista que cubre al gobernador de New Jersey Chris Christie. ¡PLQQ! [pregúntame lo que quieras]»

hoyheaprendido – «HHA que la ciudad de Boring (aburrido), en Oregón, se ha hermanado con la de Dull (tedioso), en Escocia, para promocionar el turismo conjuntamente»

preguntaenreddit – «Exfumadores de Reddit, ¿qué os funcionó de verdad para conseguir dejar de fumar?»

quiénganaría – «Superman Prime contra Superman con Guantelete del Infinito»

En la página siguiente presento un gráfico con los 200 temas más populares, y ese mapa bien podría calificarse de «Estados Unidos de Reddit». Es un tipo de división geográfica parecida a la de Craiglist que hemos visto antes —de hecho, se ha elaborado mediante el mismo algoritmo—, pero en lugar de una geografía física, constituye una geografía de los intereses de la psique colectiva de Reddit. Y presenta comunidades distintas pero relacionadas. El tamaño de cada uno de los estados corresponde con la popularidad del tema, y el *software* yuxtapone los temas semejantes en función de los comentarios cruzados entre subreddits.

Igual que hicimos antes al encontrarnos con una manera poco usual de presentar los datos verbales, conviene buscar algún término conocido para hacernos una idea de cómo encaja todo. A mí me resultó fácil. Mi juego favorito, *Magic: The Gathering (magicTCG)*, se halla atinadamente rodeado por sus desafortunados compañeros naturales como *Mens-Rights* (derechos de los hombres), *whowouldwin* (quiénganaría) y *mylittlepony*. De modo parecido, muchos deportes (*nfl*, *nba*, *fórmula 1* y demás) se agrupan en la parte inferior. Todo lo relacionado con *pokemon* se agolpa hacia la parte izquierda. Alineados con el margen derecho están temas como *Britishproblems*, *Australia* y *soccer*. También tiene lógica que los subreddits más populares se hallen en el centro, es decir, no demasiado lejos de nada. El color gris corresponde a lo densamente tejido que está cada subreddit. Muestra hasta qué punto la gente que publica allí lo hace solo allí. Cuanto más oscuro es el gris, más aislado está el hilo. Todo esto es una abstracción, pero nos enseña cómo puede irse ubicando la gente según las cosas que le interesan o le parecen divertidas o importantes en lugar de por el sitio en el que duermen por las noches. Es el mapa de una conciencia colectiva particular.

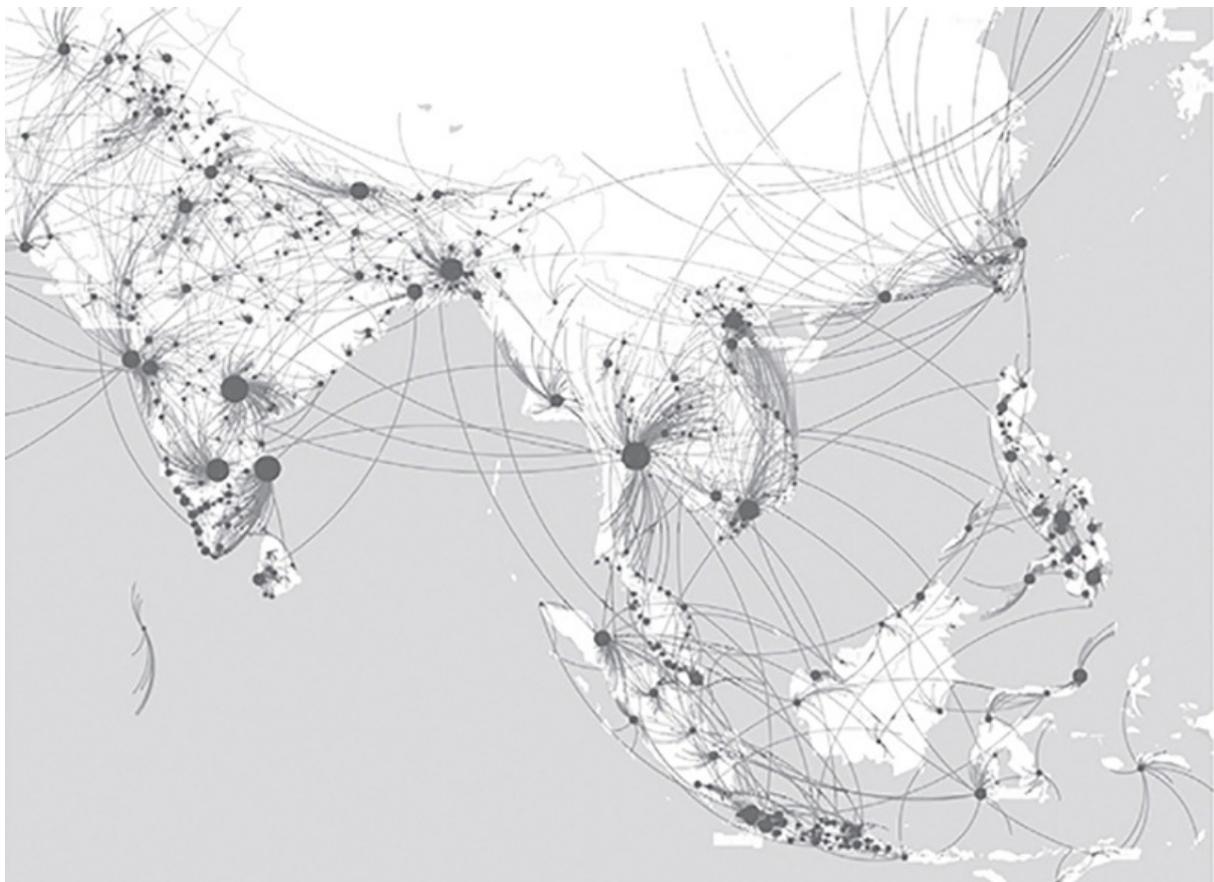
Benedict Anderson es profesor en la Universidad de Cornell y escribió un libro que tuve mucho tiempo en mi estantería sin abrirlo. Tenía que haberlo leído para una asignatura de la facultad y no lo hice, aunque siempre lo he llevado conmigo en mis muchas mudanzas a lo largo de los años; ha viajado de polizón en todas las furgonetas de mudanzas. El libro se titula *Comunidades imaginadas*, y lo abrí hace poco porque me pareció que por fin el título podía aplicarse a algo. Los temas principales que aborda Anderson son el nacionalismo y la construcción de naciones, y sugiere que una nación «es imaginaria porque hasta los habitantes de la menor de las naciones nunca llegarán a conocer a la mayoría de sus compatriotas, ni a verlos, ni siquiera a oír hablar de ellos, pero en la mente de cada uno de ellos habita la imagen de su comunión». Lo escribió en 1981, pero bien podría estar refiriéndose a Internet. No sé si Reddit es una nación, pero desborda comunión. Y resulta interesante ver a otra comunidad puramente digital definir su floreciente identidad. Antes hemos visto ese sempiterno afán por la violencia colectiva, dirigida en este caso a Safiyyah, Natasha y Justine en Twitter. Aquí, en Reddit, apreciamos algunas de las mayores bondades de una nación: sentido de pertenencia, compasión y participación.

Llevo ya doce años viviendo en Brooklyn —*Comunidades imaginadas* se había cubierto de una buena capa de la típica mugre neoyorquina para cuando lo saqué de la estantería—, pero el primer sitio al que me acompañó el libro fue a Texas. Justo después de la universidad estuve compartiendo piso con varios chicos, y uno de ellos, Andrew Bujalski, que hoy es director de cine, decidió mudarse a Austin porque le encantaban *Movida del 76* y *Slacker*. Se iba en peregrinación en busca de Richard Linklater. Los demás no teníamos ningún plan, así que nos apuntamos al suyo.

Naturalmente, lo de recoger los bártulos y mudarse de ese modo es un privilegio solo al alcance de unos veinteañeros sin nada mejor que hacer que perseguir el sueño de otro. Habíamos oído decir que Austin molaba, así que allí nos fuimos. Es un ejemplo un poco descafeinado, pero este tipo de movimientos en grupo, basados en poco más que el boca a boca y la esperanza de encontrar algo mejor, han creado el mundo tal como lo conocemos. La llamada Gran Migración —en la que millones de afroamericanos abandonaron el Sur de las leyes segregadoras de Jim Crow para irse a ciudades como Detroit, Chicago o Nueva York a principios del siglo xx— supuso un salto cultural transformador para el país y lo constituyeron miles de pequeñas decisiones de recoger las cosas y marcharse a otro sitio. Lo mismo ocurrió con la fiebre del oro que se produjo en

California. Y lo mismo con muchos de los primeros asentamientos europeos que llevaron el Viejo Mundo al continente americano. Y lo mismo, supongo, que los primeros grupos de la cultura Clovis que cruzaron el estrecho de Bering sobre el hielo hace trece mil años para constituirse en la primera nación de estas tierras. Las comunidades se trasladan para buscar un entorno que las sustente y en el que encuentren seguridad, pero también para hallar un espacio físico que esté en consonancia con lo que sienten en su fuero interno.

Recientemente, el equipo de análisis de datos de Facebook se dedicó a observar los actuales movimientos a gran escala del mundo entero, migraciones coordinadas en las que una proporción significativa de la población de un sitio se ha trasladado, como grupo, a otra parte. La gente ya no se mueve en masa de este modo en Estados Unidos, pero en muchos otros sitios sí que está empezando a hacerlo. Los investigadores cartografiaron los movimientos coordinados por todo el globo. Aquí he extraído una pequeña porción de su mapa del sureste asiático: las líneas muestran cómo ciudades pequeñas y pueblos se trasladan y se reubican enteros en grandes núcleos urbanos. Es una foto fija de una región que está cambiando rápidamente. Pero bien podría tratarse de Inglaterra alrededor de 1850 o de Estados Unidos cincuenta años más tarde.



En un sentido amplio, estos movimientos suelen deberse a cuestiones económicas: ciudades como Chicago o Bangkok prometen puestos de trabajo. Pero aunque las líneas y los puntos del mapa sean acumulativos, todas las migraciones que representan son pequeñas, personales y, sin duda, excepcionales para las personas que las emprenden. ¿Fue un padre el que tomó la decisión de hacer las maletas y marcharse? ¿Fue un amigo el que abrió el camino? ¿Con quién se juntó esa gente al llegar a su nueva ciudad? ¿A quiénes dejaron atrás en su antigua ciudad? ¿Se llevaron todas sus cosas? ¿Las abandonaron todas? Y no dejo de preguntarme también si todo el mundo tiene un libro que les sigue allá donde van hasta que lo leen. Y si es así, ¿cuál es su libro?

13.

Nuestra marca podría ser tu
vida

El logo triangular de la cerveza Bass fue la primera marca registrada del mundo angloparlante, y hoy esa robusta antigualla gráfica sigue formando una parte importante del atractivo de la marca. Lo ponen ahí en su etiqueta: «Primera marca comercial registrada de Inglaterra». Pero lo que no te dicen es que Bass fue la primera solo porque un empleado de la cervecera resultó ser el primero de la cola en la oficina del registro la mañana en la que entró en vigor la ley de registro británica. Se han valido de un accidente burocrático para labrarse una reputación que, al menos a juzgar por lo que hoy contienen esas botellas marrones, supera con mucho a la verdadera calidad del producto. Bass es una marca construida a partir de poco más que el propio acto de creación de la marca.

Antes de Bass había ya muchas marcas y firmas, las suficientes como para que el Reino Unido empezase a regularlas, al fin y al cabo, y las etiquetas y la elaboración de una imagen preceden incluso a la Revolución Industrial. Vamos, que las primeras marcas se grababan a fuego sobre la piel. Es difícil ser más primitivo. Los arqueólogos han desenterrado de tumbas del desierto selladas hace cinco mil años aceite y vino con sus marcas. En una de las etiquetas encontradas en Egipto se lee «El mejor aceite de Tjehenu» debajo del emblema real y de un pictograma que representa una prensa de aceite dorada. Compara eso con el eslogan «Los lúpulos, arroces y cebadas más selectos» que se lee debajo del título de «Reina de las cervezas» en una lata de Budweiser. Por mucho que haya avanzado el *branding*, en gran medida será siempre una ciencia de la Edad del Bronce, porque las emociones a las que responde son eternas.

Pero aunque las ambiciones y el prestigio de asociarse puedan ser conceptos intemporales, la marca se ha adentrado recientemente en nuevos territorios: las personas. En 1997, Tom Peters, especialista en motivación y asesor de empresas, publicó un artículo titulado «The Brand Called You» (la marca llamada tú) en la revista *Fast Company*. Ahí nació la era de la marca personal.

En su artículo, que en realidad es más un discurso promocional, pide a los lectores que determinen en primer lugar cuál es su «modelo de características-

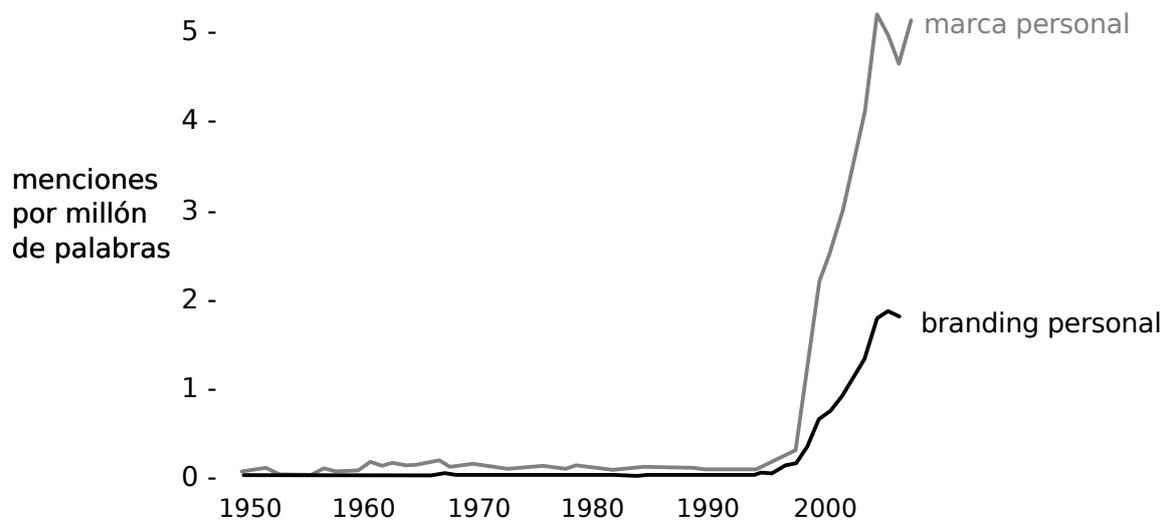
ventajas» y que después se lo vendan de manera implacable a sus posibles contratadores, a sus compañeros de trabajo, al mundo en general... ¡o la habrán pifiado! Así es como termina literalmente el artículo, y sirve de punto y final para la sarta de típicas sandeces («Siéntate y pregúntate: ¿por qué cosas quiero ser famoso? Eso es: ¡ser famoso!» y «Eres un líder. ¡Lidérate a ti mismo!») que nos ofrece la peor literatura empresarial. Al leerlo te imaginas al señor Peters con su micro y andando por el escenario sin cesar como un león enjaulado, enjaulado por el dichoso paradigma que está a punto de hacer explotar ante tus mismos ojos, con sus verdades devastadoras, su gran saber hacer y muchos signos de exclamación. Exhibe esa clase de convicción que otro tipo de persona canalizaría rompiendo listines telefónicos por la mitad. En la firma que aparece al final del artículo se lee: «Tom Peters es la marca líder del mundo en lo que respecta a escribir, hablar y pensar sobre la nueva economía». En aquel momento no solo era también la marca líder, sino la única persona que se llamaba a sí misma marca. Por tanto, tenemos a un adalid de la «nueva economía» que aprovecha una página del manual de estrategia victoriano de Bass. ¿Y por qué no? Ten a todo el mundo engañado hasta que consigas lo que quieres. El artículo dio el pistoletazo de salida a la idea de venderse a uno mismo como marca para ir directo al éxito y todavía hoy se lee en los cursos de *marketing*.

Pocos años después, un hombre llamado Peter Montoya amplió la idea de Peters en un segundo e influyente manifiesto titulado *La marca llamada tú*. Sí, llevaba el mismo título que el primer manifiesto, y no, él y Peters no trabajaban conjuntamente; de hecho, si acaso, los dos tipos son rivales en el negocio de los gurús del *branding*. Fundir el frío acero de la ignorancia absoluta para venderte una moto es el bien remunerado talento de los conferenciantes que hay por doquier, y el señor Montoya bien podría ser el maestro herrero. *La marca llamada tú* (su versión) es básicamente una lista muy larga y este es el primer punto de la lista, que aparece en la página 2:

1. Eres diferente. La diferenciación —la capacidad de ser considerado nuevo y original— es el factor más importante de la marca personal.

Por supuesto, *La marca llamada tú*, el remake, fue un superventas, y Montoya, igual que Peters, ha continuado hasta hoy con su boyante trayectoria como conferenciante. Pero si ese discurso de «tu marca personal» no hubiese ido más allá de las salas de congresos y los salones de conferencias de los hoteles del país y se hubiese limitado a ser absorbido

como los restos de café y las migajas de bollería por la raída moqueta de nuestros tiempos, yo ahora no estaría escribiendo al respecto. La idea tenía su miga, una miga succulenta, y ahora cada vez que a algún personaje de talla nacional se le escapa una metedura de pata o la pifia, la pregunta que se hace todo el mundo es: ¿cómo va a afectar eso a su marca personal? Peters y Montoya fueron innovadores, y lo digo en serio. Algunas de las personas más inteligentes y de más merecido éxito que conozco dicen las palabras «mi marca» sin ninguna ironía. En Google Libros se puede ver la génesis de la idea y su subsiguiente auge a través de las menciones en letra impresa:



Claro que los principios del *branding* personal no son nuevos. Ni Montoya ni Peters^[37] son tan distintos de Dale Carnegie, que se recicló como marca al pasar del simple «Dale Carnegie» al nombre que tomó prestado del magnate del acero Andrew Carnegie y que, como estos otros más recientes, reducía el carácter a puntos de una lista y consideraba que la influencia era, por encima de todo lo demás, la clave del éxito. Los objetivos del *branding* personal son los mismos que podemos encontrar en cualquier seminario de motivación o en cualquier sermón evangélico sobre la prosperidad de cualquier década. La meta ha sido siempre lograr riqueza y poder.

Lo que tiene de nuevo es que el «branding personal» te exige que alcances esos fines tratándote a ti mismo como si fueras un producto y no como un ser humano. De nuevo, Peters:

A partir de hoy eres una marca. Eres tanto una marca como lo son Nike, Coke, Pepsi o The Body Shop. Para empezar a pensar como tu propio responsable de *branding*, hazte la misma pregunta que se hacen los responsables de *branding* de Nike,

Coke, Pepsi o The Body Shop: ¿Qué hace que mi producto o servicio sea diferente?

Ese es el concepto nuclear de la marca personal, y al igual que el cristianismo con la imprenta o el fútbol con la televisión, la idea ha encontrado en las redes sociales la tecnología ideal para llegar al mundo entero. No voy a insistir sobre las distintas maneras que tienen sitios como Facebook, Twitter e Instagram de posibilitar que nos proyectemos al mundo. Pero sí señalaré que no hace mucho solo las grandes empresas, con grandes presupuestos, podían hacerse oír y querer por desconocidos que estaban al otro lado del planeta. Ahora también yo puedo hacerlo, y también tú, y todo el mundo. Que alguien te escuche ya es más difícil

La mejor manera de lograrlo es ser ameno, interesante y divertido. Pero por algo son tan pocos los cómicos capaces de hacer reír a la gente. Porque es difícil. Es más probable que un aficionado que trate de hacerse con una buena base de seguidores en Twitter siendo irónico o provocador acabe convirtiéndose en el siguiente Justine Sacco que en Justin Halpern (@ShitMyDadSays), que tiene 3 millones de seguidores y un contrato para publicar un libro. Por cada chaval que logra entrar en una universidad o conseguir un buen trabajo en la revista *New Yorker* gracias a Twitter —como han hecho algunos— debe de haber decenas cuyos tuits acaban llevándolos al despacho del director o, más probablemente, a estrellarse contra un muro de embarazoso silencio.

Podremos hacernos una idea de lo que se necesita para conseguir seguidores si volvemos a usar nuestro algoritmo de análisis de textos. Estas son las palabras típicas de lo que yo llamaría niveles de seguidores de un «amateur raso» y un «profesional del colegio»:

palabras más típicas de...

gente con <100 seguidores	gente con >1000 seguidores
#thehungergames	partnering
#upset	#heyboo
#worthit	vamping
#mhyme	optimizing
roethlisberger	sourcing
#workaholics	marketer
#wordsofwisdom	tweetup

#hurryup	visibility
#depressed	monetize
#wishmeluck	industry's
#getonmylevel	optimize
#studying	brownskin
#idiots	merchants
cincy	influencers
#collegeproblems	robust
#sunny	yeen
#notokay	guwop
#finalsweek	talmbout
#tebow	innovators
#silly	partnered
#impatient	bezos
#leavemealone	infographics
#holysht	livest
#suckstosuck	strategist
pujols	entrepreneurial
#saveme	slideshare
#yeahbuddy	yass
pattys	amplify
#girlproblems	goodmorning
#killme	creatives

En la columna de la izquierda vemos el tipo de intereses sencillos y fugaces que uno esperaría encontrar en Twitter. En la de la derecha vemos que prácticamente todo es jerga empresarial: si tienes muchos seguidores, lo más probable es que acabes hablando como una gran empresa. Pero algunas de las palabras de la derecha no son típicamente profesionales: #heyboo (saludo adolescente), talmbout (contracción de «*talking about*», hablando sobre), yeen («*you ain't*», tú no), yass («*your ass*», tu culo) y unas cuantas más. Esa es gente que usa Twitter igual que los de la columna de la izquierda —para hablar, quejarse, fanfarronear— pero la diferencia es que lo hacen para círculos más amplios, de miles de seguidores. Los usuarios que escriben esas palabras son negros y la presencia de esos términos en la columna de la derecha es otra prueba de que los afroamericanos tienden a usar esa red social

de manera distinta (quiero poner énfasis en lo de que «tienden» porque no hay ningún grupo que sea monolítico). Los observadores llaman a este fenómeno el Twitter Negro, que Farhad Manjoo describe así en *Slate*:

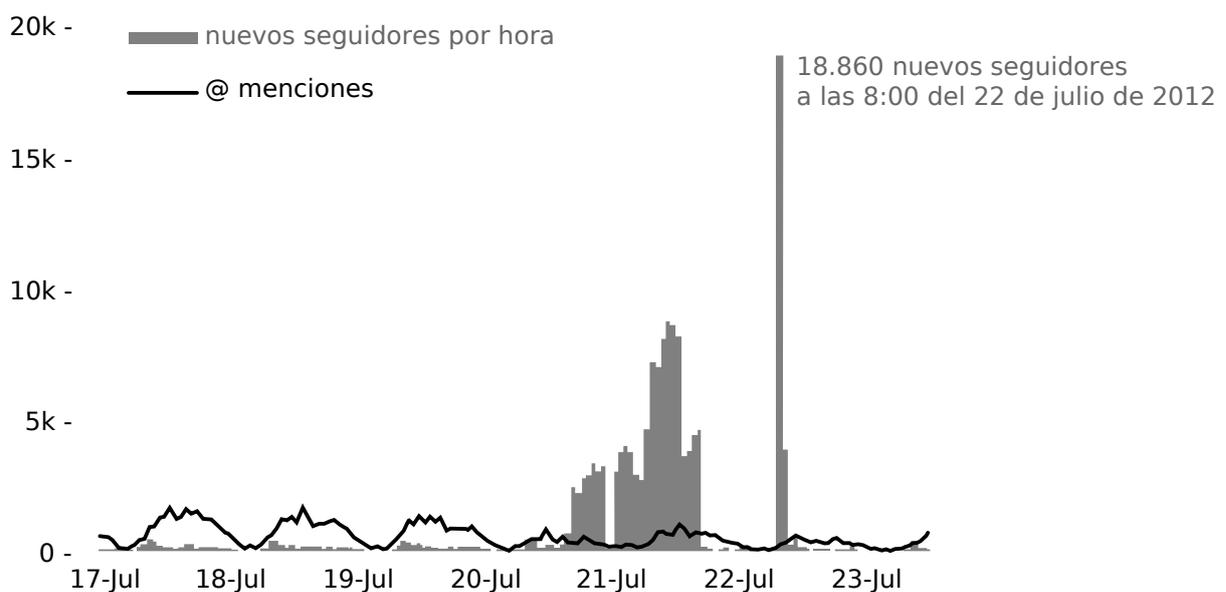
Los negros —más concretamente, los negros jóvenes— parecen usar Twitter de un modo distinto de como lo hacen todos los demás. Forman grupos más cerrados en la Red, son más propensos a seguirse unos a otros y a retuitearse, y una mayor proporción de sus tuits son respuestas a los de otros. Este comportamiento, intencionado o no, es lo que propicia que los negros —concretamente, los adolescentes negros— dominen la conversación en Twitter.

Cuando habla de «dominar» se refiere al hecho de que en los primeros años de Twitter había mucha confusión entre los usuarios blancos cuando llegaban a la lista de *trending topics* etiquetas como #uainthittinitright (no lo estás haciendo bien) o #ifsantawasblack (si Santa Claus fuese negro), junto con la última reflexión profunda de Ryan Seacrest o la última ocurrencia publicitaria de Old Spice (igual que lo de #heyboo puede parecer extraño cuando aparece junto a la palabra «monetize» en la lista). La mayoría de los usuarios de Twitter sigue a instituciones de un tipo u otro (famosos, periodistas, productos) y esas instituciones no los siguen a ellos. La cultura de esa red social se organiza convencionalmente alrededor de una comunicación de un emisor y muchos receptores, o sea, alrededor de la marca. Pero los usuarios negros tienden a centrarse en un uso personal y son sumamente recíprocos; de ahí sus elevados recuentos de seguidores y la capacidad extrema que tienen de poner sus memes en lo más alto de las listas.

Cualquiera que espere crearse una marca en esa red social de la manera convencional —de uno para muchos— debería darse cuenta de que Twitter es el mundo del Uno Por Ciento. Su recurso más preciado, los seguidores, se distribuyen de manera mucho más desigual que la riqueza. En mi muestra, el 1 por ciento de las cuentas tiene el 72 por ciento de los seguidores. Y el 0,1 por ciento tiene algo más de la mitad. Resulta mucho más difícil conseguir un millón de seguidores que ganar un millón de dólares. Según el IRS, en 2011 hubo 300 890 personas que declararon ingresos superiores a un millón de dólares. Ahora mismo hay en todo el mundo 2643 cuentas de Twitter con un millón de seguidores. Cerca de la mitad de ellos están en Estados Unidos. Ser un estadounidense con un millón de seguidores en Twitter es más o menos equivalente a ser multimillonario^[38].

Naturalmente, eso suponiendo que los seguidores sean de verdad. Compré unos cuantos desde una de mis cuentas para ver cómo funcionaba. En sitios como TwitterWind puedes elegir una cantidad de un menú (yo seleccioné 1000), pagar (17 dólares) y un par de días después, y casi todos a la vez, tienes ahí a todos esos amigos nuevos e inútiles. Lo único que hacen esos seguidores de alquiler es existir, y prácticamente todos aquellos que tienen una gran base de seguidores en Twitter seguramente han comprado unos cuantos, sobre todo esa gente para quien parecer popular lo es prácticamente todo, como los famosos y los políticos. Cuando todavía no se había decidido la candidatura republicana, Newt Gingrich fanfarroneó: «Tengo seis veces más seguidores en Twitter que todos los demás candidatos juntos». La única pega es que había pagado para obtener cerca del 90 por ciento de ellos^[39]. Mitt Romney (casi con toda seguridad) también compró seguidores; por poner un ejemplo, obtuvo 20 000 seguidores en cuestión de minutos un día de julio, algo así como doscientas veces el incremento de seguidores que había tenido antes y que tuvo después. Ahora bien, hay que destacar dos cosas: una, que una persona puede comprar seguidores para otra, así que este podría haber sido una especie de caso Nixon de juego sucio político del siglo XXI; y la verdad es que fue una buena manera de hacer quedar a Mitt como un tontaina. Y dos, estoy seguro de que Obama y muchos, muchos demócratas han comprado también seguidores. Este tipo de cobardes intentos de hacer trampas con el sistema son moneda de cambio corriente en ambos partidos. Lo que pasa es que normalmente no se los pilla tan fácil como en este caso:

actividad de @MittRomney durante julio de 2012



Se puede llegar a entender por qué lo hacen. Cuanto más popular parece alguien, más popularidad adquiere. Es lo más parecido que se puede llegar a la compra de votos, al menos hasta que el Tribunal Supremo lo legalice en 2018.

Los titulares de cuentas más de andar por casa tampoco escapan a la tentación de los amigos fáciles, aunque no manejen presupuestos como los de Barack o Mitt. Dos de los cinco hashtags más comunes de mi paquete de datos extraídos aleatoriamente de Twitter (que aparecen en el número uno y cinco, respectivamente) son #ff y #teamfollowback. El primero significa «Follow Friday» y es una tradición que se remonta a los primeros tiempos de Twitter: los viernes destacabas a la gente que te gustaba para que tus seguidores los siguiesen a su vez. Ahora se ha convertido en un código general (que se usa en cualquier momento) para decir: «Hey, seguid a estas cuentas» y con el que normalmente bombardean a los demás quienes pretenden aumentar su número de seguidores. El segundo, #teamfollowback, es la etiqueta que emplean determinadas cuentas de Twitter que básicamente hacen de manera gratuita lo que los políticos pueden permitirse pagar. La idea consiste en que tú sigues a TeamFollowBack y los demás seguidores de esa cuenta te siguen a ti. Tú, a tu vez, los sigues a ellos, y así se incrementan los números de todos. Es como el viejo concepto del «web ring», que en los tiempos anteriores a Google era una manera de que las webs se enlazasen unas con otras para garantizarse más tráfico. También se parece a la vieja idea de la paja colectiva. Así es como se autodefinen los de esa cuenta TeamFollowBack:

¡Te ayudamos a conseguir seguidores! LA CUENTA ORIGINAL [sic] Y LA MEJOR. Promociona nuestros hashtags #WILLFOLLOWBACK y #TEAMFOLLOWBACK

Así que a esto puede llevarte lo de la marca personal: a buscar números vacíos. Yo sé lo que tuiteo, tengo el mismo interés en saber quién lo difunde y con qué rapidez lo hace que el interés que puse en lo que inicialmente trataba de comunicar. Las pocas veces que he publicado en Facebook me he quedado sentado y dándole a recargar la página para ir viendo los nuevos comentarios, como si nunca antes hubiese entrado en Internet. Jena Wortham, del *Times*, describe muy bien esta mentalidad: «Nosotros, los usuarios, los productores, los consumidores —toda nuestra frenética energía, nuestro afán por hacernos notar, por que nos reconozcan las decisivas contribuciones a la conversación — somos el problema. Y lo alimenta nuestra creciente necesidad de atención,

de valoración, a través de los “Me gusta”, de favoritos, respuestas, interacciones. Es un círculo que se retroalimenta y que no hay manera de cerrar, al menos no por ahora». Yo puedo confirmarlo desde dentro: las empresas diseñan sus productos para que ese círculo quede abierto. OkCupid te muestra unos cuantos recuentos de tus mensajes, visitas y posibilidades. Sabemos que esas cifras mantienen interesados a los usuarios, sobre todo cuando prosperan. Sin esas pizcas de emoción, una web o una app le parece muerta al usuario y la abandona. El término general que se usa para esto es «implicación del usuario», la cantidad de gente que se conecta cada semana, cada día, cada hora. Básicamente, se trata de a qué velocidad corren en la rueda de hámster que se les ha preparado junto a su nidito de serrín, y es una de las medidas que el sector observa más obsesivamente. Los sitios web te enseñan recuentos, totales e insignias porque saben que volverás para ver cómo van aumentando. Después podrán poner tu creciente implicación en una diapositiva para impresionar a sus inversores.

Por ahí va la cosa: se trata de que tú mismo te reduzcas a un número. Cuando es otro quien te reduce a un número no te sienta bien. Klout es una de las principales firmas de análisis personal en la Red. Analizan todas tus cuentas en las redes sociales y, gracias a una magia negra que solo ellos conocen, te ofrecen una medición general de tu influencia *online*, puntuándote de cero a 100. Como recordarás gracias a Montoya (y a Carnegie): la influencia es lo importante de una marca personal, y Klout te ayuda a averiguar qué tal lo estás haciendo. Ahora mismo, mi puntuación en Klout es un bastante penoso 34. TeamFollowBack tiene 60, lo que no sé si me hace reír o llorar. Por una parte, esta gente ha obtenido el equivalente a un suspenso en lo que es su único motivo para existir. Por otra, tienen una puntuación más alta que cualquiera que conozca.

En 2012, Salesforce.com, el mastodonte de la informática en la nube, publicó una oferta de trabajo que incluía entre las «aptitudes deseables» una puntuación en Klout de al menos 35. No figuraba como requisito, pero lo pusieron allí junto a otras cosas de esas que te comentan aunque puedan parecer obvias, como «capacidad de trabajo en equipo», de modo que supuestamente era una parte importante del trabajo. El negocio de Salesforce se centra en la cuantificación: ayudan a las empresas a vender con la ayuda de datos^[40]. Así que no sorprende que afronten la contratación de personal de la misma manera cuantificadora. Pero aunque cifras como los índices de capacidad crediticia llevan ya algún tiempo formando parte de los odiosos

procesos de recursos humanos, ver una puntuación de Klout en una oferta de trabajo molestó a bastante gente.

El artículo «¿Quieres trabajar para Salesforce? Mejor que tengas una puntuación de 35 o más en Klout», publicado en *BetaBeat*, suscitó una reacción general acorde con la única palabra que componía el subtítulo que le pusieron al artículo: «Puf». No obstante, lo que de verdad preocupa, que vamos a vernos todos reducidos a un número, y pronto, merece discusión aparte. Salesforce era, y sigue siendo, una empresa que marca tendencias, especialmente en el mundo del *marketing online*. Fue la «Empresa más innovadora de Estados Unidos» según *Forbes* el mismo año que publicaron esa oferta. Contratan a centenares de personas cada año y, lo que es más importante aún, cuando una empresa innovadora que ha recibido premios hace algo nuevo, las demás la copian. Si Salesforce pide puntuaciones de Klout, todos los demás empezarán pronto a pedir puntuaciones de Klout. La gente no quiere verse reducida a un número de dos dígitos elaborado por una empresa que incluso en el vaporoso mundillo de las redes sociales parece un tanto fantasma.

Pero puesto que Klout emplea muchas de las mismas herramientas de reducción que yo he usado para recopilar datos, ¿cómo nos deja eso a ti y a mí y a este libro en el que llevamos tanto tiempo invertido? Bueno, pues la respuesta corta es: nos deja en el mismo lugar que a Klout y Salesforce. La reducción es inevitable. Los algoritmos son descarnados. Los ordenadores son máquinas. La ciencia de los datos está intentando dar sentido digital a un mundo analógico. Es un derivado de la naturaleza física más básica del microchip: un chip no es más que una sucesión de puertas diminutas. No en el sentido figurado en que decimos que Internet es una «serie de tuberías», sino en realidad. Las puertas se abren y se cierran al paso de los electrones, y cuando una de estas puertas quiere saber en qué estado debe estar, las dos opciones son todo o nada: como cualquier puerta, un circuito está abierto o no está abierto, no hay matices ni quizás que valgan. A partir de esa realidad microscópica se propaga un absolutismo por todo el sistema, hasta que en los niveles superiores encontramos las definiciones, los tipos de datos y las clases que forman la esencia de la programación de lenguajes como C o JavaScript.

Por tanto, la información se ve reducida por pura necesidad. Pero las objeciones que se le pusieron a aquel requisito de la puntuación de Klout se referían fundamentalmente al hecho de que se redujera a dígitos a las personas, no solo a su información. Y aquí es donde *Dataclismo* difiere de la

oferta de empleo de Salesforce y también del modelo de negocio entero de Klout.

Por muchos números que aparezcan aquí, no tienen el propósito de representar a ninguna persona concreta. Un número nunca podría hacerlo. Ese es un hecho que resume muy bien la anécdota apócrifa que dice que Einstein suspendió las matemáticas en el instituto. No es verdad. Pero podría haberlas suspendido. Y si lo hubiese hecho, ¿a quién le importa? Si hubiese sacado un 35 en Álgebra II, ¿qué? ¿De repente deja de ser inteligente? Ningún número, ningún test, ninguna medida singular —ni el CI, ni la estatura ni tampoco, por supuesto, la puntuación en Klout, la cantidad de seguidores o el porcentaje de respuestas en OkCupid— es una persona como tal, motivo por el cual no aparecen personas concretas en este libro, salvo con fines ilustrativos. Pero cuando sumamos un puñado de estas partes pequeñas e incompletas de nosotros, obtenemos algo grande. La regla de los números grandes es un concepto que he tocado de pasada unas cuantas veces ya, pero quiero dejarlo claro de manera explícita: la verdad completa de los datos solo se nos descubre cuando examinamos una muestra grande. Imagina un dado misterioso: no sabemos cuántos lados tiene pero podemos tirarlo unas cuantas veces. Si lo tiramos una sola vez nos dará cualquier número y seguiremos sin saber nada. Pero si lo tiramos unas cuantas veces obtendremos su distribución, el promedio, y eso nos dirá al momento cómo es el dado. Solo sabremos qué forma tiene por acumulación.

Y, lo que es más, la reducción y la repetición son fundamentales en la larga historia de la ciencia, no solo de la ciencia que estudia los datos o la informática, sino de la Ciencia con mayúscula, la eterna iniciativa humana. Los experimentos se elaboran a base de reducir un proceso a su aspecto más sencillo y manejable. El método científico exige control, y no podemos tenerlo sin reducir la complejidad a lo más simple y llano y decir que eso, *eso*, es lo importante. Solo cuando hemos simplificado la cuestión podemos ponerla a prueba una y otra vez. La mayor parte del conocimiento que poseemos se adquirió de ese modo, por reducción, ya fuera en una mesa de laboratorio o en un ordenador portátil.

De modo que aquí hemos destilado la humanidad hasta reducirla a unos números en lugar de, pongamos, anécdotas. Tal como yo lo veo —y esto no resta valor alguno a Malcolm Gladwell—, este libro es lo contrario de los valores atípicos o las cosas fuera de serie. En lugar de detenerme en los datos errabundos que aparecen en los extremos más alejados —los casos únicos, las excepciones, las peculiaridades, los Einstein de quienes tienes que conocer

toda la historia para entenderlos—, me baso en el conjunto indiferenciado. Nos centramos en las acumulaciones densas, en los centros masivos, en los datos duplicados una y otra vez por la repetición y el carácter común de nuestra experiencia humana. Es ciencia a modo de puntillismo. Puede que esos puntos sean una mínima fracción de ti, pero el conjunto somos nosotros.

La acumulación y la reducción también nos permiten manejar tendencias generales, cuyo suave flujo puede que no presente los picos y valles de las habituales narraciones heroicas, pero que por eso mismo son más aplicables. El hecho de que Paul McCartney y John Lennon ensayasen música *rock* durante 10 000 horas y después se convirtiesen en los Beatles nos dice algo acerca del valor que tienen el ensayo y la persistencia, pero ese número por sí solo no significa nada. Yo mismo me he pasado un tiempo parecido tocando la guitarra, como tantos otros cuya música jamás llegaremos a escuchar. Sea lo que fuere lo que permitió que Lennon y McCartney convirtiesen la práctica en genialidad, es algo exclusivo de ellos. Por otra parte, cada uno de los números que aparecen en este libro tiene detrás muchos cientos, y a veces muchos miles, de personas, ninguna de ellas famosa. Ahí está el meollo de la cuestión: la frase «uno en un millón» está en la esencia de muchísimas obras de arte maravillosas. Implica que hay una persona que es tan especial, tan talentosa, tan «algo» que es prácticamente única, y esa misma rareza es lo que la hace significativa. Pero en matemáticas —y lo mismo pasa con los datos y por tanto con este libro— la frase significa justo lo contrario: $1/1\ 000\ 000$ es un error de redondeo.

Pero si lo que hace falta para comprender los macrodatos es simplificar, a mí me preocupa otro tipo de reduccionismo: que la gente no se reduzca exactamente a un número, sino a una identificación de usuario deshumanizada y atrapada en el molino de un algoritmo de *marketing*, en la harina de la marca de otro. Los datos les restan mucha especulación a las ventas. Es una de esas pocas leyendas urbanas que resultan ser verídicas, pero Target, al analizar las compras de una consumidora, supo que estaba embarazada antes de que se lo hubiese dicho a nadie. La pega fue que era una adolescente y que empezaron a enviar publicidad maternal a casa de su padre.

En cierta medida, ese tipo de intromisión corporativa es mejor que algunas marcas que intentan hacerse las «colegas». El verano pasado, una campaña de *marketing* de la empresa de gelatinas Jell-O se hizo eco (¿se apropió?) de la etiqueta #fml, que es la abreviatura que se usa en Internet para «fuck my life» (puta vida). Sus responsables de redes sociales empezaron a responder a los tuits que contenían ese hashtag con una oferta no solicitada en

la que cambiaron el «fuck» por «fun» (diversión) y repartieron vales de regalo. Así, gente que lo estaba pasando mal empezó a recibir alegres ofertas de una gelatina, como en esta conversación:

Pyrrhus Nelson @suhrryp 
Estoy viendo desaparecer mi cuenta corriente en la consulta del médico #fml

JELL-O @JELLO 
@suhrryp Fun my life? Claro que lo haremos. De hecho, nos encantará. prmtns.co/dkTq Entrega 48h.

Es muy fácil colar este tipo de intercesión no solicitada en las redes sociales porque está todo muy cuantificado. Los hashtags aparecen uno tras otro en la pantalla del encargado de gestionar la marca y este no tiene más que ofrecer descuentos. Por lo menos la misma tecnología que les permite meterse en nuestra vida nos permite a nosotros defendernos. Hace unos cuantos años, McDonald's publicó un par de tuits con unas anécdotas entrañables sobre sus proveedores que llevaban la etiqueta #McDStories, y a cambio obtuvieron respuestas de tipo #fml. Esta es solo una de las muchas respuestas:

MUZZFUZZA @Muzzafuzza 
Hace muchos años que no voy a McDonald's, porque antes preferiría comerme mi propia diarrea. #McDStories

McDonald's había pagado para promocionar ese hashtag y retiró la campaña al cabo de un par de horas, cuando vieron que se les había ido de las manos. Una semana después la reconvertida etiqueta #McDStories seguía coleando con fuerza. Sus estrategias para las redes sociales deberían haber sabido lo que iba a ocurrir: unos meses antes, Wendy's había intentado promocionar #HeresTheBeef (aquí está la carne), y su eslogan fue despojado completamente de su contexto previsto. La gente se dedicó a usarla para quejarse sobre cualquier cosa que no les gustaba («beef» es también «queja» en inglés), haciendo caso omiso de la marca:

Remi Mitchison @RemiBee 
#HeresTheBeef cuando una tía ve que a otra tía le va mejor y tiene más que ella... lo que quiere es darle una buena zurra.

Jeremy Baumhower @jeremytheproduc 
#HeresTheBeef Las farmacéuticas ya han curado el VIH y el cáncer, pero les es mucho más rentable mantener a la gente con un hilo de vida a base de medicamentos.

Más recientemente, Mountain Dew lanzó un concurso llamado «Dub the Dew» (ponle un nombre a Dew), con el que trataron de subirse a la ola del «crowdsourcing» para ponerle un nombre molón a un nuevo refresco y tal vez pensando que, si todo iba bien y las mediciones mostraban la atracción suficiente para conseguir la confianza de determinada gente influyente, lograrían hacerse con unos cuantos embajadores de la marca en la blogosfera. Reddit y 4chan se hicieron eco de la historia y el nombre «Hitler no hizo nada malo» encabezó las votaciones durante un tiempo, hasta que en el último momento irrumpió la opción «Diabeetus» y se oyó la voz del pueblo: «Ponte nombre a ti mismo, capullo».

Internet puede ser un sitio desquiciante, pero ese potencial que tiene para lo inesperado, incluso para lo demente, es lo que muchas veces la redime. ¡No me puedo imaginar nada peor para ti, La Marca, que votarle a Hitler! Además, menuda pérdida de tiempo, porque está claro que Mountain Dew no va a imprimir ni una sola palabra que desfavorezca a ninguna de sus líneas de productos tan valiosas y distinguidas. No me molestan la tontería ni la frivolidad, ni siquiera la estupidez. Troleear a un refresco es algo que ninguna fórmula recomendaría jamás. No es una buena práctica industrial. Y demuestra que por mucho que el corporativismo pueda invadir nuestros feeds de noticias, nuestros álbumes de fotos *online*, nuestros muros e incluso, como algunos desearían, nuestras mismas almas, una pequeña parte de nosotros sigue quedando fuera de su alcance. Eso es lo que siempre quiero tener presente: no son los números los que nos niegan nuestra humanidad, es la decisión premeditada de dejar de ser humanos.

14.

Miguitas de pan

Facebook lanzó el botón de «Me gusta» en 2009, y eso cambió el modo que tenía la gente de compartir contenidos. La idea no era nueva: en otros sitios que fueron populares y ahora eran marginales, como digg.com y del.icio.us, hacía ya años que la gente podía dar «Me gusta» a los artículos que publicaban. Pero en esas empresas la estrella era el contenido. Facebook aportó cierto grado de edición a una red social que ya era sólida y, para bien de los creadores de contenido, facilitó que cualquiera pudiese añadir ese iconito del pulgar levantado para aprobar su trabajo. Crearon una nueva microdivisa universal: a lo mejor no te pago por tu texto, tu música o lo que sea, pero te doy un estímulo en forma de aprobación y difundo lo que has creado entre mis amigos. En mayo de 2013, Facebook sumaba 4500 millones de «Me gusta» al día, y en septiembre de ese mismo año declaró que el botón se había pulsado 1,13 billones de veces en total.

Aquellos estudiantes del MIT desarrollaron su radar gay el mismo año en el que se lanzó el botón de «Me gusta». Su algoritmo era bastante bueno a la hora de adivinar la sexualidad de un hombre, pero también funcionaba de una manera bastante obvia: no es ciertamente ningún gran secreto que los hombres gais suelen tener amigos gais. La innovación de su radar gay fue emplear datos a nivel macro para hacer algo que la gente había estado haciendo a pequeña escala toda la vida. Desde entonces, la potencia del *software* predictivo se ha incrementado con rapidez; esta clase de programas se van haciendo más veloces e inteligentes conforme se dispone de más datos. En 2012, un equipo del Reino Unido descubrió que solo a partir de los «Me gusta» de una persona eran capaces de averiguar lo siguiente, con diversos grados de precisión:

si alguien es...

gay o hetero	88 %
lesbiana o hetero	75 %
caucasiano o afroamericano	95 %
hombre o mujer	93 %
demócrata o republicano	85 %

consumidor de drogas	65 %
hijo de padres que se divorciaron antes de que cumpliera los 21	60 %

Insisto en que esto no se basa en observar las actualizaciones de estado, los comentarios, el contenido compartido ni nada que haya tecleado el usuario. Solo sus «Me gusta». Te das cuenta de que la ciencia se adentra en terrenos inexplorados cuando oyes a tus padres luchando con el clic-clic-clic del ratón. El patrón de los «Me gusta» de una persona puede darnos incluso una buena idea de su inteligencia: este modelo podía predecir de manera fiable la puntuación de alguien en un test estándar de CI (que se le hiciese por separado), sin que dicha persona hubiese contestado ni una sola pregunta directa.

Esto se analizó informáticamente a partir de tres años de datos recopilados de gente que se registró en Facebook después de haber vivido décadas sin esa red social. ¿Qué podrá ocurrir cuando alguien lleve usando estos servicios desde la infancia? Ese es el lado oscuro de los datos longitudinales que tanto me entusiasman. Hace mucho que empresarios, universidades y ejércitos emplean *tests* como los de Myers-Briggs o Stanford-Binet. Te sientas, lo haces lo mejor que puedes y te califican. Por lo general, tú eliges hacerlo. Pero cada vez se da más el caso de que estás haciendo estos *tests* simplemente por vivir tu vida. Y los resultados quedan a la vista de todos para que los lean y los juzguen. Una cosa es ver que alguien tiene una puntuación de 51, o lo que sea, en Klout antes de hacerle una entrevista de trabajo y otra muy distinta es saber cuál es su CI.

Si los contratadores de personal empiezan a utilizar algoritmos para averiguar cuán inteligente eres o si consumes drogas, tu única opción será simular seguirle el juego al sistema o, empleando una expresión prestada del capítulo anterior, «gestionar tu marca». Para vencer a la máquina deberás actuar como una máquina, lo que significa que la máquina te habrá vencido. Y eso suponiendo en primer lugar que seas capaz de deducir qué es lo que tienes que hacer. Por lo visto, una de las cosas que más correlacionadas aparecían con la inteligencia en la investigación era que te gustasen las «patatas fritas onduladas». ¿Quién va a ser capaz de sacar algo de ese dato mediante ingeniería inversa?

Pero aunque es verdad que Facebook sabe muchas cosas de ti, es más como un «amigo del trabajo»: pese a todo el tiempo que pasáis juntos, vuestra relación tiene unos límites definidos. Facebook solo sabe lo que haces en Facebook. Hay muchos sitios cuyo alcance es mucho más profundo. Si tienes un iPhone, Apple puede disponer de tu agenda de contactos, tu calendario, tus

fotos, tus textos, toda la música que escuchas, los lugares a los que vas y hasta cuántos pasos das para llegar hasta ellos, puesto que los *smartphones* llevan dentro un pequeño giroscopio. ¿Que no tienes un iPhone? Entonces cambia Apple por Google, Samsung o Verizon. ¿Llevas una FuelBand? Nike sabe si duermes bien o no. ¿Tienes una Xbox One? Microsoft conoce tu ritmo cardiaco^[41]. ¿Usas tarjeta de crédito? Cómprate algo en una tienda y tu PII (información personal de identificación) adjuntará el código UPC a tu identificador en el *software* CRM (gestión de relaciones con el cliente), que se pondrá a trabajar acto seguido en lo siguiente que vas a querer comprar.

Esto es solo una migaja del alcance de los datos a nivel corporativo, cuya descripción completa podría llenar varias páginas. En cuanto a lo que tiene el gobierno, lo único que sé es otra migaja, porque eso es lo único que a todos nos han dejado ver al respecto. Sabemos que en el Reino Unido hay 5,9 millones de cámaras de seguridad, una por cada 11 ciudadanos. En Manhattan, justo por debajo de la calle Catorce, hay 4176. Los satélites y los drones acaban de completar el panorama desde más arriba del asfalto. Aunque no hay manera de saber qué es lo que ven, podemos asegurar que si el gobierno está interesado en tu paradero y tus actividades, seguro que habrá uno vigilándote. Y además, como desveló Edward Snowden, gran parte de lo que no pueden observar a través de una lente lo pueden monitorizar a placer desde la pantalla de un terminal de NSANet, cuya ubicación es secreta.

Dado que ocurren tantas cosas con tan poco conocimiento del público, la comprensión que puede tener de los datos una persona de a pie va siempre varios pasos por detrás de la realidad. Debo decir que por el simple hecho de haberme parado a escribir este libro estoy seguro de que ya me he quedado atrás. El análisis de datos ha superado ya en gran medida a la propia información como la verdadera palanca que hay que empujar. Las *cookies* de tu navegador y los tíos que piratean números de tarjetas de crédito acaparan la mayor parte de los titulares y son con toda certeza los recolectores de datos más acusadamente molestos. Pero también se han hecho con una pequeña porción de tu vida y para ese pedacito minúsculo han tenido que meter muchas horas de trabajo. No importa lo intrincado que sea el JavaScript, que ellos son villanos como los de las películas mudas, con mostachos atusados y chistera. O, siguiendo una referencia más contemporánea, son como otros tantos Doctores Malignos, reliquias de los bajos fondos que chantajejan al mundo por *un... millón... de dólares...* mientras que otros miles de millones pasan a las manos de los verdaderos cerebros, como Acxiom. Estos traficantes corporativos de datos, que llegan hasta ámbitos como los registros bancarios y

de las tarjetas de crédito, los historiales de ventas de los minoristas y archivos gubernamentales como los registros tributarios, saben cosas sobre la conducta humana que jamás podría saber un investigador académico que se dedica a pescar en los sitios web en busca de patrones^[42]. Entretanto, los recursos y la experiencia que pone a trabajar el aparato de seguridad nacional hace que el *software* de minería de datos que se emplea a nivel corporativo parezca el Buscaminas.

Estos datos, pese a la metáfora «minera», no son un recurso que se genere de forma natural; vienen de alguna parte, y esa parte eres tú. Las empresas y el gobierno van recopilando piezas dispersas de tu vida privada y tratan de reorganizarlas para hacerse con una imagen que puedan manejar. Cuanta más intimidad pierdas, más eficaces son. El asunto fundamental en cualquier discusión sobre privacidad es ese intercambio: lo que tú obtienes a cambio de perderla. Nos pasamos todo el tiempo calculando ese intercambio. Los personajes públicos venden su vida personal para progresar en su carrera. Cualquiera que haya reservado un hostel en Europa o que haya comprado un billete de tren en la India ha tenido que decidir si vale la pena gastar un poco de dinero adicional para tener un espacio privado. Y no es por desviarme del asunto que nos ocupa, pero mucha gente, hombres y mujeres, comercia con su intimidad cada noche cuando salen por la puerta de casa y la entregan, ya sea mediante una falda corta o un pantalón ceñido, a cambio de atención. Así que ese intercambio no es nuevo. Pero sí lo son nuestros socios comerciales y sus condiciones. Por el lado corporativo, lo que rinden nuestros datos (nuestros beneficios) no resulta tan interesante a menos que seamos economistas. En teoría, tus datos significan que los anuncios están mejor dirigidos, lo que supone menos gasto en *marketing* y, por lo tanto, precios más bajos. Por lo menos, los datos que venden implican que podamos utilizar servicios verdaderamente útiles como Facebook y Google sin tener que pagar por usarlos. Lo que obtenemos a cambio por la intrusión gubernamental ya está menos claro.

¿Nos hace estar más seguros esa vigilancia? ¿Proporciona cobertura total el aparato de seguridad? Bueno, no se han producido atentados terroristas contra civiles estadounidenses desde 2001, al menos no por parte de los sindicatos. Eso no carece de significado, sobre todo para un neoyorquino. Pero un argumento basado en la ausencia de algo no es muy sólido, y al menos hasta que se nos informe de las amenazas que se han desbaratado, y no de aquellas que nunca llegaron a planearse, nos cuesta creer lo que nos han dicho. El «Nivel de amenaza» codificado por colores del que tanto se habló

durante los años posteriores al 11-S, cuyo recuerdo casi se ha disipado ya, como ocurre en Texas con la arena de las tormentas, siempre me pareció un anuncio publicitario muy elaborado de Halliburton. Cuesta creer en una información que te llega con el argumento de la «necesidad de saber» y procedente de una entidad que cree que no tienes por qué saber nada. Ya cada vez hemos de preocuparnos menos por lo que nos dicen y más por el porqué de lo que nos dicen. En cualquier caso, no tengo ni idea de cuántos crímenes, si es que hubo alguno, ha evitado esa gran cosecha de datos de la NSA. Nos han dicho que funciona, pero no cuándo, dónde ni cómo.

Por quijotesco que parezca, en el caso de aquellos crímenes que la vigilancia total no logró evitar, sí que resultó útil a la hora de resolverlos. Todas esas cámaras de seguridad resolvieron el caso de la bomba en la maratón de Boston, al igual que lo hicieron en los atentados del metro de Londres en 2005^[43]. Sobre todo en el caso de los crímenes de índole asíncrona, necesitas poder recurrir a todos los datos posibles, porque los delincuentes cometen sus actos mucho antes de que se produzcan las víctimas. En esas investigaciones, el poder de los servicios de inteligencia forma parte de la historia que interesa a los medios, es el momento en el que el «estado de la vigilancia» brilla con máximo esplendor. Entonces los datos tienen un propósito definido y nadie discute sobre el equilibrio entre privacidad y protección cuando las calles están manchadas de sangre. Pero entre uno y otro de esos momentos de «mantengámonos unidos», gran parte de lo que nos enteramos que sabe el gobierno procede de soplones como Snowden.

La NSA es el brazo operativo de los servicios de inteligencia del gobierno encargados de la interceptación de señales, y aquí las señales que andan buscando están en nuestros datos. Estoy personalmente un tanto familiarizado con esa organización. Como ya he dicho, estudié matemáticas. Y lo hice en Harvard. Mi licenciatura parece igual que la de mis compañeros de curso, pero había extraoficialmente dos itinerarios en el departamento. Uno, el mío, era para los chavales a quienes les gustaban las mates y se les daban bastante bien. El otro era para los prodigios extraordinarios. En primero había una asignatura difícil llamada Matemáticas 25, para la que yo no era suficientemente bueno y a partir de la cual se escogía a la ultraélite para pasarlos a una clase llamada Matemáticas 55 por invitación especial del departamento. Estos matemáticos de verdad se saltaron completamente las asignaturas más duras que yo cursé. Muchos de los profesores auxiliares de mis cursos más avanzados, la gente que se encargaba de la mayor parte de la

formación y de poner todas las notas, no solo eran más jóvenes que yo (uno tenía 16 años) sino que ya de adolescentes habían cursado el currículo de licenciatura completo. Recuerdo que me entusiasmaba (y me parecía un desafío) la clase de Análisis real, que era una asignatura que la mayoría de mis compañeros —como si esa palabra pudiese definirlos— habría encontrado aburrida cuando iba a secundaria. Cada vez que veo las siglas NSA me acuerdo de aquellos tiempos, porque era en ese segundo itinerario donde reclutaban a la gente.

Saco esto a colación porque para mucha gente los empleados del gobierno tienen una reputación anodina: burócratas, funcionarios, lo que sea. Y también es verdad que los que trabajan en análisis de datos en el sector privado tienen las mismas posibilidades de ser competentes como de no serlo. Pero la gente que nos espía es extremadamente lista. Podemos esperar que, como ya hicieron Feynman y Einstein antes que ellos, sean capaces de aplicar a su trabajo cierta sagacidad y humanidad, pero lo que sí sabemos con certeza es que, como ya hicieron Feynman y Einstein antes que ellos, aquello en lo que trabajan es inhumanamente poderoso.

Sabiendo que los algoritmos se alimentan a base de datos, el señor Snowden nos ha hecho ver que los de la NSA están cebados con superalimentos. O, mejor dicho, con todos los alimentos que existen. Recogen llamadas de teléfono, correos electrónicos, mensajes de texto, fotos... básicamente todo aquello que viaja a través de impulsos eléctricos. Está claro que no se trata de una operación pasiva: según un documento filtrado, el objetivo declarado al más alto nivel es el de «dominar Internet». El descaro de ese proyecto es una de las cosas que más impresionan de él. Entre los primeros documentos que se publicaron (conjuntamente por el *Guardian* y el *Washington Post*) se halla una presentación en PowerPoint sobre un programa denominado PRISM. Las diapositivas no se andan con chiquitas:



Current Providers

- Microsoft (Hotmail, etc.)
- Google
- Yahoo!
- Facebook
- PalTalk
- YouTube
- Skype
- AOL
- Apple

What Will You Receive in Collection
(Surveillance and Stored Comms)?

It varies by provider. In general:

- E-mail
- Chat - video, voice
- Videos
- Photos
- Stored data
- VoIP
- File transfers
- Video Conferencing
- Notifications of target activity - logins, etc.
- Online Social Networking details
- **Special Requests**

Complete list and details on PRISM web page:
Go PRISMFAA

TOP SECRET//SI//ORCON//NOFORN

La deberían haber llamado Operación Te-Lo-Birlo. Por una parte, la vida en la Tierra solo puede empeorar cuando alguien que lleva pistola empieza a pensar en nuestras cuentas de Facebook. Por otra, cuesta tener miedo de una gente que emplea las herramientas de dibujo de un producto de Microsoft.

Nadie puede ver sin una orden judicial los datos que tiene PRISM sobre determinado individuo, al menos en teoría, porque el programa es sumamente invasivo. Otra de las intromisiones se centra en los metadatos, esos imponderables de la comunicación. He aquí la descripción de una parte de otro proyecto por parte de la propia Junta de Supervisión de la Intimidad y las Libertades Civiles del Gobierno estadounidense:

De cada uno de los millones de números de teléfono que cubre el programa Section 215 de la NSA, la agencia obtiene un registro de todas las llamadas entrantes y salientes, la duración de estas y la hora exacta del día en la que se realizaron. Cuando la agencia apunta a un número de teléfono para su análisis, [se obtiene] esa misma información de cada uno de los números con los que ha entrado en contacto el número de teléfono

original, así como de cualquier otro número con el que hayan entrado en contacto cualquiera de esos números.

Hay que decir que nada de esto atañe al contenido en sí de las comunicaciones de nadie. A ese respecto, no hay mucha diferencia con los datos que hemos analizado en este libro. Aquí hacemos que los patrones sustituyan a la vida de cualquier persona concreta, igual que hace esta gente. En la NSA, de nuevo según ellos mismos, si tu red de llamadas cumple con el perfil de una «amenaza», solo entonces empiezan a prestarte atención. Pero los metadatos no son necesariamente menos invasivos por muy indirectos que sean.

La gente deja unos asombrosos rastros de miguitas de pan para quienquiera que tenga interés en seguirlos. Ya has visto bastante al respecto, unas doscientas páginas. Aun así, hay otros muchos rastros que no hemos seguido. Por ejemplo, un pequeño archivo de texto llamado Exif se adjunta a todas las imágenes que se capturan con una cámara digital, desde las SLR de gama alta a tu iPhone. En el archivo queda registrado no solo cuándo se hizo la foto, sino otras cosas variadas como la relación focal o la velocidad del obturador y, muchas veces, la latitud y longitud del lugar donde se disparó. El Exif es lo que permite que programas como iPhoto ordenen tus fotos con facilidad en «momentos» y que coloquen alfileres por todo el mapamundi para mostrarte dónde has estado. Pero el Exif también puede decirnos otras cosas. Tomemos, por ejemplo, las fotos de los perfiles de OkCupid. Cuanto mejor es una foto, más posibilidades hay de que sea antigua. Es decir, que a la gente le parece que esa es una «gran foto» y se la dejan puesta ahí para siempre. Esto lo sabemos gracias al Exif, que nos dice cuándo se hizo la fotografía. Este tipo de datos rémora son más que habituales. Las coordenadas del GPS te acompañan por la Red en todo momento cuando abres tu app favorita. Casi todas las páginas web que has abierto en tu vida tienen decenas de imágenes de un píxel (un simple puntito transparente) enterradas en los márgenes que, al cargarse junto con la página «real», registran tu visita; los píxeles no pueden saber lo que estás haciendo, pero sí cuándo y dónde has estado. Estas cosas tan sencillas, esos cuándo y dónde, pueden darle a una empresa una buena conjetura sobre tu perfil demográfico completo.

¿Y qué pasa con la gente que no quiere compartir nada de ese modo, los que prefieren comprar a su aire y acicalarse a solas? Yo mismo conozco el valor de la intimidad. Ese es uno de los motivos por los que no soy muy fan de las redes sociales, la verdad. Nunca he colgado una foto de mi hija en Internet. Empecé a usar Instagram a principios de 2011, cuando la plataforma

todavía no era muy grande, y la usaba como aplicación de galería de fotos porque me gustaron los filtros. Me parecía que era un poco como Hipstamatic, que tenía poco de red social —ya sé que esto me hace parecer un abuelo—. Cuando mi mujer se dio cuenta de lo que estaba haciendo el carroza de su marido, me informó de que podía conectar mi cuenta con las cuentas de otros, cosa que hice, porque, mira, solo hay que darle a un botón. Pero en cuanto aquello ya dejó de ser solo yo y mis fotos, perdió todo interés para mí.

Este tipo de reticencia no es lo habitual. Con la cantidad de quejas que se oyen, cuesta argumentar que la mayoría de los usuarios muestran indiferencia en este asunto de la privacidad. Cada vez que Facebook actualiza sus Condiciones de servicio para incrementar su ya exhaustivo dominio de nuestros datos, nos pasamos un día rabiando y al siguiente ya estamos otra vez allí, como cuando se provoca a las abejas y, al ver que no tienen a nadie a quien picarle, no les queda otra que volverse todas a su colmena. Como a la tecnología le gusta forzar los límites y los límites siguen cediendo, el *software* se ha vuelto casi agresivamente invasivo. Hay apps para perder peso. Apps para medir el ritmo cardíaco. Apps para que los demás valoren tu ropa: envías tu combinación de prendas a la multitud y esta te aconseja sobre moda. Las mujeres emplean apps para predecir y controlar sus ciclos menstruales: «El mercado está inundado de ellas», escribe Jenna Wortham, para añadir: «Casi todas las mujeres que conozco las usan». Le indicas a la aplicación el día en el que te empieza la regla y te avisa de tu momento de máxima fertilidad, para que lo aproveches o lo evites, según prefieras. Claro que, como los datos que declara la propia gente no son suficientemente invasivos, ha aparecido una startup que asegura poder deducir cuándo tiene la regla una mujer a partir de su historial de enlaces. Cualquiera de estas apps menstruales —por lo menos si tienen detrás a un científico de datos competente— sabrá, por supuesto, cuándo una mujer está embarazada, hace demasiado ejercicio, se hace vieja o ha tenido relaciones sexuales sin protección, ya que cuando tienes una falta consultas la aplicación con excesiva frecuencia.

Pero pese a esa actitud despreocupada sobre la privacidad que adopta alguna gente, por no decir mucha, no he querido poner en peligro la identidad de nadie a la hora de escribir este libro. Como ya he dicho, todos los análisis se hicieron de forma anónima y en formato agregado, y he manejado el material original en bruto con sumo cuidado. Ninguno de mis datos contenía información personal de identificación (PII). Al tratar con las palabras escritas por los usuarios —los textos de sus perfiles, tuits, actualizaciones de estado y demás— me aseguré siempre de que fuesen públicas. Cuando utilicé registros

de interacciones entre dos usuarios concretos, las identidades de los usuarios estaban encriptadas. Y en todos los análisis el alcance de los datos se limitó a las variables más esenciales, de forma que nada pudiese vincularse con ningún individuo concreto.

En ningún momento he pretendido vincular los datos a los individuos, por supuesto. Mi objetivo era vincularlos con todo el mundo en general. Ese es el valor que veo yo en los datos y, por tanto, también en la privacidad que se pierde con su misma existencia: lo que podemos aprender. Jaron Lanier, autor de *¿Quién controla el futuro?* e informático que ahora trabaja en Microsoft Research, escribió en *Scientific American* que «se está almacenando y analizando una ingente cantidad de información acerca de nuestra vida privada, de la que se hace uso antes de haber demostrado la posibilidad de darle un uso válido». Tiene razón de manera incuestionable en lo de la «ingente cantidad», pero discrepo en la última parte de la frase. ¿Cómo puede algo llegar a ser útil alguna vez si no se le puede «dar uso antes de haber demostrado la posibilidad de darle un uso válido»? El concepto mismo de investigación científica se afianza sobre la exploración. El mineral de hierro no era sino una roca más hasta que alguien empezó a experimentar con él. El moho del pan se pasó milenios haciendo que la gente se pusiese enferma hasta que Alexander Fleming descubrió que también servía para hacer penicilina.

La ciencia de los datos está generando ya grandes descubrimientos que no solo sirven para describir cómo vive la gente, sino para cambiar su manera de vivir. Ya he mencionado antes el servicio de Evolución de la gripe de Google; está disponible desde 2008 y ahora rastrea epidemias incipientes en más de 25 países. No es una herramienta perfecta, pero es un buen comienzo. Los datos combinados ya se emplean para prevenir enfermedades, no solo para minimizarlas. Como publicó el año pasado el *New York Times*: «Gracias a datos introducidos en los motores de búsqueda de Google, Microsoft y Yahoo, científicos de Microsoft y de las universidades de Stanford y Columbia han podido por primera vez detectar pruebas de efectos secundarios de medicamentos recetados antes de que las detectase el sistema de alertas de la Food and Drug Administration». Los investigadores determinaron que la paroxetina y la pravastatina provocaban hipoglucemia en los pacientes. En este caso, la recompensa por vivir con un poco menos de intimidad es vivir con una salud un poco mejor.

Cada día, por lo que parece, nos trae noticias de algún nuevo avance. Hoy he descubierto que una web llamada geni.com va camino de crear un árbol genealógico de toda la humanidad por medio del *crowdsourcing*. Si funciona,

la empresa básicamente habrá generado una red social de nuestro material genético. La semana pasada, dos politólogos echaron por tierra la idea preconcebida de que los republicanos deben su mayoría parlamentaria a la manipulación de circunscripciones electorales. Los autores habían confeccionado modelos de todas las posibles elecciones en todas las configuraciones posibles de Estados Unidos y, mientras en el ordenador sonaba la opereta *Candide*, concluyeron que nuestro mundo dividido en dos es lo mejor que podemos esperar. Es la geografía política del país, y no los mapas, lo que propicia esta situación estancada de bipartidismo.

Esto es solo el comienzo. Los datos llevan una buena ventaja —Facebook, ya en 2002, recopilaba 500 terabytes de información cada día—, pero los análisis ya les empiezan a ir a la zaga. El periodismo basado en datos lo puso de actualidad Nate Silver, pero se está convirtiendo en algo indispensable para los medios: cuantificamos para comprender. El *Times*, el *Washington Post* y el *Guardian* cuentan con impresionantes equipos de análisis y visualización de datos y no dejan de dedicar recursos a publicar datos sobre nuestra vida, incluso en el actual clima de recortes económicos que sufren los periodistas y su trabajo.

Por lo que respecta al bando corporativo, Google, que tantas veces hemos mencionado en estas páginas, va en cabeza en cuanto a aplicar los datos al bien común. Hemos visto lo de la gripe y el trabajo de Stephens-Davidowitz, pero hay también un montón de proyectos todavía más ambiciosos, si bien menos publicitados, como Constitute, un sistema de diseño de constituciones basado en datos. A los ciudadanos de la mayoría de los países normalmente solo les preocupa una constitución, la de su país, pero Google ha reunido los 900 documentos legislativos de ese tipo que se han redactado desde 1787. Combinados y cuantificados, proporcionan a los países nacientes —cada año se escriben cinco nuevas constituciones— mayores posibilidades de disfrutar de un gobierno duradero porque pueden ver qué es lo que ha funcionado y lo que no en el pasado. En este caso, los datos abren la puerta a un futuro mejor porque, como señala la web de Constitute, en una constitución «hasta una simple coma puede marcar una enorme diferencia».

Como ya hemos visto, el equipo de Facebook encargado de datos ha empezado a publicar investigaciones de gran valor a partir de su inmenso almacén de acciones y reacciones humanas. Alex Pentland, del MIT, se hace heredero de esa interacción newtoniana y denomina a esta nueva ciencia la «física social». Su equipo y él han empezado a trasladar los datos sociales al mundo físico. Trabajando conjuntamente con el gobierno local, las operadoras

de comunicaciones y los ciudadanos han «datificado» una ciudad entera. Los residentes de Trento, en Italia, ahora pueden enfrentarse, mediante números puros y duros, a lo que para el resto de nosotros son preguntas sin respuesta que nos hacemos a diario: ¿en qué gastan el dinero las demás familias?, ¿cuánto salen y socializan?, ¿con qué guarderías o médicos permanece la gente más tiempo?

Tal vez sea ese el futuro que debemos esperar. He tratado de explicar lo que hemos aprendido mediante la combinación de los trabajos que están ya a la vista de todos y de mis propias investigaciones. Al hacerlo, más que levantar los brazos para exclamar: «¡Hemos llegado a la cima!», pretendo transmitir las posibilidades de lo que está por llegar. Watson y Crick resolvieron el acertijo del ADN en 1953, y seis décadas después los científicos siguen desentrañando el genoma humano. A la ciencia de nuestra humanidad compartida —la búsqueda de la expresión completa de esos genes que pronto tendremos totalmente cartografiados— le faltan todavía años para llegar a esa cima.

En lo que respecta a equilibrar las posibles ventajas e inconvenientes, me gustaría poder proponer un camino a seguir. Pero, para ser sincero, no veo una solución sencilla. Puede ser que lo vea desde demasiado cerca. Conuerdo con la opinión de Lanier de que la regulación no funcionará, aunque seguro que habrá quien pruebe ese enfoque. Las nuevas leyes se redactarán con toda la buena intención, estoy convencido, pero sus letras se habrán quedado obsoletas antes de que se seque la tinta. Y estando como estoy en el bando de los coleccionistas de datos, he visto con mis propios ojos que puedes darle a la gente todos los controles de privacidad del mundo, que la gente no los usará. OkCupid les pregunta a las mujeres: «¿Has abortado alguna vez?» (es la pregunta número 3686 de las que se hacen; ya he dicho que lo abarcan todo). Justo debajo de la pregunta hay una casilla para marcar si quieres que tu respuesta sea privada. De las mujeres que responden afirmativamente, menos de la mitad marcan esa casilla.

De modo que la mayoría de la gente no va a usar las herramientas que les proporcionas, pero tal vez nos estemos equivocando de objetivo con eso de «la mayoría de la gente». Por una parte, lo que hay que hacer es facilitar maneras de borrar los datos, o al menos de recuperarlos, independientemente de cuán pocos sean los usuarios que las empleen. Por otra, es posible que la privacidad haya cambiado y que haya dejado atrás a quienes escriben sobre ella. Lanier y yo somos unos viejos según los criterios de Internet, y «los generales siempre libran la última guerra» no solo con sus ejércitos. Mis

expectativas sobre lo que es correcto y permisible pueden ser erróneas. Las diferentes culturas y generaciones definen de manera distinta la privacidad.

A la gente ni siquiera le molesta demasiado la NSA, por muy repugnantes que sean sus extralimitaciones. Se han celebrado muchas «protestas del millón» en Washington: Million Man, Million Mom, etcétera. Hace poco, el colectivo de *hackers* Anonymous convocó a un millón de enmascarados para manifestarse, entre otras cosas, contra el programa PRISM y la vigilancia masiva del gobierno. El *Washington Post* captó la escasez de interés del público en la mismísima primera palabra de su crónica: «Cientos de manifestantes...».

Lanier, en su artículo de *Scientific American*, propone que se nos compense por nuestros datos personales y que sean las fuerzas del mercado las que equilibren la ecuación privacidad/valor. Propone que los recopiladores de datos emitan micropagos a los usuarios cada vez que se venden sus datos. Pero ese gasto pasaría directamente a devolverse al consumidor, como un impuesto, o provocaría una carrera en espiral descendente en la que los sitios web tendrían que sacar su margen de donde fuese, como hacen ahora las aerolíneas comerciales.

La opción de Pentland es mucho más factible: lo llama su «nuevo acuerdo sobre los datos». Por irónico que parezca, sus principios se remiten al antiguo derecho anglosajón. Opina que, como ocurre con cualquier otra de tus pertenencias, deberías tener derechos fundamentales de posesión, uso y destrucción de tus datos. Eso significa que deberías poder eliminar tus datos de una web (o de cualquier otro sitio donde estén almacenados) siempre que creas que se están utilizando de manera incorrecta. También deberías poder «llevártelos contigo», en teoría para revenderlos, en caso de que se crease un mercado para ello. Ese sencillo mecanismo —el botón de borrar, con la opción de copiar/pegar— no solo es más viable, sino más justo que cualquier compensación forzosa.

De hecho, por la parte corporativa, podría argumentar que a la gente ya se la está compensando por sus datos: pueden usar plataformas como Facebook y Google —contactar con viejos amigos, encontrar lo que buscan— de manera gratuita. Como ya he dicho, apporto poco a estos servicios, pero también los aprovecho poco. La gente tiene que decidir cuál es su compensación en esto. Aunque pronto puede que solo haya una decisión que tomar: ¿voy a usar esos servicios o no? El análisis de datos se está volviendo tan potente que es posible que lo que intentas conservar deje de tener importancia. Los algoritmos ya logran extrapolar o inferir muchas cosas sobre

una persona partiendo únicamente de la información más básica; y eso que solo llevan unos cuantos años trabajando con los datos. Pronto las medias opciones que proporcionan los menús para que «gestiones tus ajustes de privacidad» dejarán de dar protección alguna, porque el resto del mundo no será tan celoso de su información. Las empresas y el gobierno te encontrarán enseguida en los gráficos. Todo este debate podría convertirse pronto en un anacronismo.

Sea como sea, cuando hace mucho, mucho rato me referí a los datos como una inundación, tal vez no hice suficiente hincapié: las aguas todavía siguen agitándose. Solo cuando empiecen a calmarse podrá la gente saber con certeza qué nivel han alcanzado y aprovechar lo que pueda de esa riada. Yo estoy ansioso por hacerlo. Entretanto, la gente que almacena, analiza y utiliza los datos tiene la responsabilidad de seguir demostrando el valor de su trabajo y de revelar con exactitud qué es lo que están haciendo. Si no, pese a todas mis objeciones, Lanier tiene razón: no deberíamos estar haciéndolo.

La tecnología es el nuevo mito. Indudablemente, tiene algo de mágico. Pero más grandiosa aún que su sustancia es la imagen que da. Dioses de la tecnología. Titanes. Colosos puestos a horcajadas sobre la Tierra entera, porque, la verdad, Rodas está pasado de moda. Así es como muchas veces se proyecta este sector ante el público y, tristemente, también es como muchas veces se ve a sí mismo. Pero aunque es cierto que hay monstruos, no hay dioses. Todos haríamos bien en recordarlo. Todos somos falibles, humanos y mortales, y todos caminamos bajo el mismo oscuro firmamento. Nosotros trajimos el diluvio: ¿nos mandará a pique o nos levantará sobre las olas? Lo que espero de mí y de otros como yo es que saquemos algo bueno, real y humano de los datos. Y mientras lo hacemos, cada vez que la tecnología, los aparatos y los algoritmos nos parezcan demasiado épicos, deberemos recordar al Ulises envejecido de Tennyson y optar por buscar nuestra verdad de un modo ligeramente distinto. Esforzarse, buscar, encontrar, pero luego, siempre, devolver algo a cambio.

Conclusión

Para diseñar las tablas y los gráficos de este libro me basé en el trabajo del estadístico y diseñador gráfico Edward R. Tufte. Más que basarme en su trabajo, traté de copiárselo. Sus libros ocupan ese estrecho nicho que se halla entre la hermosura de los libros ilustrados de sobremesa y la claridad de los libros de texto, y en su interior plasma principios del diseño de información que beben de los ejemplos más famosos de la historia en lo que a representar datos de manera narrativa se refiere: el gráfico de Charles Minard sobre la caída de Napoleón en Rusia; la *Descripción de un buque negrero*, obra de un abolicionista anónimo, que muestra el cargamento humano estibado en las peores condiciones, una imagen que sigue siendo un icono gráfico de los horrores de la ruta de los esclavos; el mapa del doctor John Snow del brote de cólera de 1854 que ubicó por primera vez el origen de la enfermedad. Tufte obtiene enseñanzas de todo ello y lo convierte en gráficos útiles para el contexto actual. Le exige al diseñador de información que saque el máximo provecho a la relación entre datos y consumo de tinta, que dé a cada esquema una historia clara que contar, que use el color para destacar la esencia de los datos, que emplee el blanco como una dimensión más, no como un espacio muerto. Yo lo he hecho lo mejor que he podido.

Entre los muchos mapas, gráficos y tablas que figuran en los libros de Tufte, hay uno a doble página que representa el monumento a la guerra de Vietnam, no en forma de grabado en piedra ni a modo histórico, sino como un puro diseño de datos. Me gustaría poder reproducirlo íntegro aquí, pero el núcleo del gráfico es este:

Desde cierta distancia, esa recopilación de los nombres de 58 000 soldados fallecidos dispuestos sobre el granito negro proporciona una medida visual de lo que significa el número 58 000, al difuminarse en una forma gris las letras de cada uno de los nombres para, acumulándose, dar lugar al recuento total de bajas.

Encontrar significado en esa forma gris difuminada es lo que ansía todo científico de los datos, y yo he procurado repetidamente guardar esa distancia

y hallar ese gris en estas páginas, mirando a los datos en forma de grandes paquetes, contemplando las historias más genéricas, todo para incrementar mis opciones de dar con la verdad.

El monumento se digitalizó en 2008. Se fotografió cada centímetro cuadrado y se cotejó con los archivos militares, y la versión *online* permite que los usuarios añadan fotos y comentarios a cada uno de los nombres. La versión web confronta al visitante con un campo de texto vacío y le insta a «Buscar en el muro». Tras una pausa, empecé a escribir el nombre de mi padre, porque es casi un acto reflejo pensar en él cada vez que me viene a la mente Vietnam. Pero luego recordé que, por fortuna, David Patton Rudder no está en esa lista. Así que introduje un nombre cualquiera, una conjetura: «John», por supuesto, y luego, puesto que Smith me parecía demasiado insulso y Doe demasiado falso, escribí «Wilson». La página se recargó durante medio segundo y luego en la parte superior apareció esto:

Lorne John Wilson

Fecha de inicio de servicio 17/3/1967

Fecha de fin de servicio 28/3/1967

Fecha de deceso 28/3/1967

Edad 20

En esa entrada habían añadido dos fotos, una era un retrato en uniforme de gala y la otra, una instantánea, tal vez tomada uno de esos once días que el soldado de primera clase Wilson pasó vivo en aquel país. En ella se ve a cuatro jóvenes alrededor de un *jeep*, uno de ellos subido en la parte trasera; están charlando y es por la tarde. La foto tiene grano y está descolorida, pero si no fuese porque llevan uniforme de faena, podría haber salido de Instagram. Quienquiera que la hubiese colgado había conservado la foto, y a sus amigos, durante décadas.

Una web no puede sustituir al granito. Tampoco puede sustituir a la amistad, el amor o la familia. Pero lo que sí puede hacer —como conductor de nuestra experiencia compartida— es ayudar a que nos comprendamos y a que comprendamos nuestra vida. La era de los datos ya está aquí; ahora estamos todos registrados. Eso, como todo cambio, asusta, pero entre el gris metálico-armamentístico del gobierno y el rosa chillón de las ofertas comerciales que no podemos rechazar hay una vía abierta y sensata: usar los datos para saber pero sin manipular, explorar pero sin fisgonear, proteger pero sin asfixiar, observar pero sin destapar y, sobre todo, devolver ese regalo impagable que legamos al mundo cuando compartimos nuestra vida para que

otras vidas puedan mejorar. Y que todos satisfagan la más antigua de las esperanzas humanas, desde Gilgamesh y Ramsés hasta nuestros días: que nuestro nombre sea recordado, no solo en piedra sino como parte de la propia memoria.

Acerca de los datos

Los números son engañosos. Aun sin contexto, dan la apariencia de ser hechos, y su especificidad anula todo argumento: «20 679 médicos aseguran que los Lucky Strike son menos irritantes». ¿Qué más necesitamos saber sobre fumar, verdad? La ilusión se refuerza todavía más cuando los números se visten de estadísticas. No pretendo rescatar aquí ningún saber ancestral, pero detrás de cada número hay una persona que toma decisiones: qué analizar, qué excluir, cómo enmarcar las imágenes que nos pintan los números. Enunciar algo, aunque sea haciendo un gráfico sencillo, es tomar decisiones, y en esas decisiones surge inevitablemente la imperfección humana. Por lo que a mí respecta, no he tomado ninguna decisión consciente que haya dado determinado sesgo al resultado de mi trabajo, los datos de personas que viven su vida ya son lo suficientemente interesantes sin necesidad de que intervenga yo para inclinarlos hacia un lado u otro. Pero sí he tomado decisiones, y esas decisiones han afectado al libro. Me gustaría repasar unas cuantas.

Mi primera decisión fue con toda probabilidad la más difícil: la decisión de centrarme en las relaciones entre hombre y mujer al hablar de la atracción y el sexo. El espacio, claro, fue un factor importante: incluir relaciones entre personas del mismo sexo habría supuesto reproducir por triplicado cada uno de los gráficos y tablas. Pero más que eso, lo determinante fue descubrir que las relaciones entre personas del mismo sexo no son excepcionales, sino que siguen las mismas tendencias. Los hombres gays, por ejemplo, prefieren parejas más jóvenes, igual que los heteros. En cuestiones que solo tienen que ver con el sexo de manera indirecta, como las valoraciones entre personas de distintas razas, los gays y los heteros también muestran patrones parecidos. Las relaciones entre hombre y mujer me posibilitaban repetir menos y lograr una mayor repercusión por unidad de espacio, así que opté por centrarme en ellas.

Mi segunda decisión, la de dejar fuera los tecnicismos estadísticos, me costó mucho menos tomarla. No menciono en *Dataclismo* los intervalos de confianza, los tamaños de la muestra, los valores p y demás lindezas porque el libro tiene que ver, ante todo, con la popularización de los datos y de su

estudio. La idea no era hacer entender la parafernalia matemática. Pero, como ocurre con los pilares y las vigas de una casa, el rigor no deja de estar presente aunque no se vea. Muchas de las investigaciones que aparecen en el libro proceden de fuentes académicas corroboradas por expertos. Yo he aplicado los mismos criterios en mis investigaciones: gran parte del análisis de OkCupid lo realicé primero personalmente y después lo verificó de manera independiente un empleado de la empresa. También separé el análisis de la selección y organización de los datos para asegurarme de que lo primero no afectase a lo segundo. Una persona extraía la información y otra trataba de averiguar cuál era su significado.

A veces presento una tendencia y le atribuyo una causa determinada. Muchas veces esa causa es la mejor que se me ocurre, dado mi conocimiento de todas las fuerzas que intervienen. Para interpretar resultados —algo necesario en un libro que no se limite a presentar listados de cifras— he tenido que optar por una explicación de entre varias posibles. ¿Intervienen otras fuerzas aparte de la edad en la que yo llamo ley de Wooderson (el hecho de que a los hombres heterosexuales de todas las edades les interesen más las mujeres de 20 años)? Puede ser. Pero me parece muy improbable. La frase «La correlación no implica causalidad» es algo que conviene tener en mente, además de constituir un excelente control de las extralimitaciones narrativas. Pero que esa sea una frase concisa no significa que la cuestión de la causalidad no sea interesante, y he intentado atribuir causas solo cuando estaba más que justificado hacerlo.

En el caso de todas aquellas partes de *Dataclismo* que se solapan con entradas publicadas en el blog de OkCupid, he optado por rehacer el trabajo desde cero a partir de los datos más recientes en lugar de limitarme a citar mis descubrimientos previos. Lo he hecho así porque, la verdad, quería volver a comprobar lo que había analizado antes. La investigación publicada allí entre 2009 y 2011 se fue realizando de manera fragmentada. Varias personas —por lo menos cinco, que recuerde— estuvieron extrayendo índices de mensajes entre hombres y mujeres para mí durante esos tres años, por mencionar uno de los asuntos que trato con más frecuencia en este libro, y aunque quisiera volver a repasar todos mis datos, no habría manera de saber con certeza cuál de aquellas series de datos fue la que generó los resultados. Volviendo a repetir todo el proceso me aseguraba y también conseguía aplicar unos criterios uniformes a toda mi investigación (por ejemplo, restringir el análisis solo a personas de edades comprendidas entre los 20 y los 50, una decisión

que tomé porque esas son las edades de las que sabía que disponía datos representativos).

Puesto que este estudio es nuevo, los números impresos en *Dataclismo* difieren de los publicados en el blog. Las curvas adoptan formas algo distintas. Los gráficos son un poco más gruesos o más finos en determinados puntos. Sin embargo, los resultados de la investigación del libro y del blog son coherentes. Por irónico que parezca, en este tipo de investigación la precisión es muchas veces menos apropiada que la generalización. Por eso a menudo redondeo las cifras hasta el 5 o el 10 más cercanos y por eso también aparecen mucho en estas páginas términos como «alrededor de», «aproximadamente» y «cerca de». Cuando lees en un artículo que el «89,6 por ciento» de la gente hace tal o cual cosa, el verdadero interés es que «muchos» o «casi todos» o «cerca del 90 por ciento» de la gente hace tal o cual cosa, pero lo más probable es que al redactor le pareciera que los decimales molaban más y le daban un toque de autoridad al texto. Si un científico volviese a calcular esa cifra, tal vez el resultado le diese un 85,2 por ciento. Y la siguiente vez igual le daba un 93,4 por ciento. Observa el mar cuando hay marejadilla y pregúntate qué ola se encuentra exactamente «al nivel del mar». Como mucho, se trata de un ejercicio inútil; o peor aún, engañoso.

Si seguimos la pista a los resultados que aparecen en *Dataclismo* hasta sus fuentes originales, no encontraremos discrepancias solo en los datos de OkCupid. Estos datos sobre nuestra vida, siendo como son prácticamente una cosa viva, están siempre cambiando. Por ejemplo, mi puntuación en Klout, que se mantiene en 34 mientras escribo estas líneas, sin duda habrá subido cuando las estés leyendo, puesto que parte de mi compromiso con la editorial consistirá en tuitear acerca de este libro. ¡Eso es implicación del usuario!

A veces los números varían sin razón aparente. Mi editora y yo las pasamos canutas para dar con los textos de autorrelleno de Google en preguntas como «¿Por qué las mujeres...». Google nos había dado resultados ligeramente distintos a los dos [«... llevan tanga?» fue el tercer resultado que me salía, seguramente porque esa es una pregunta típicamente masculina (?)]. El que le salía a ella era («... llevan sujetador?»). Cuando volví a comprobarlo unas semanas después, me salió otra cosa distinta: «... van juntas al baño?». Como me pareció el resultado más decente, es el que acabé incluyendo en el libro.

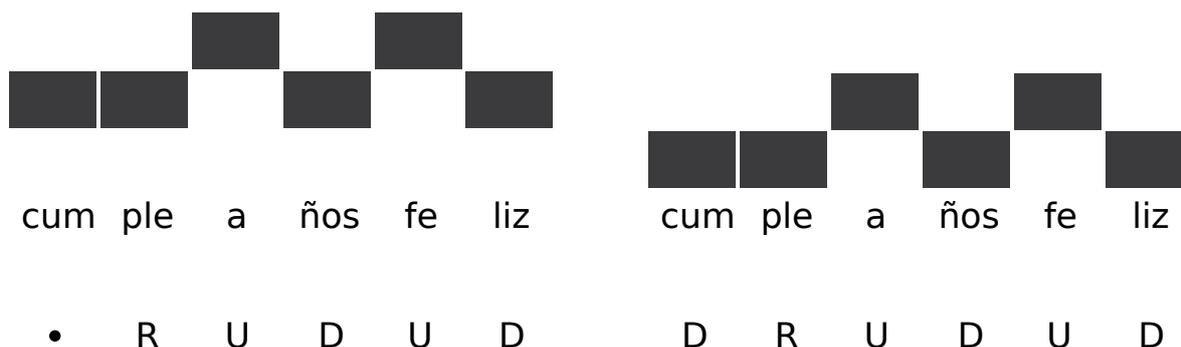
Por muy interesante que sea esa caja negra que es el autorrelleno de Google (y de Google Tendencias, para el caso), ejemplifica una de las peores

cosas de la actual ciencia de los datos: su opacidad. La corroboración, tan importante en el método científico, es difícil, porque muchísima de la información es de propiedad exclusiva (y en esto OkCupid es tan culpable como el que más). Pese a lo mucho que pregonan las empresas de redes sociales la inmensidad y el potencial de sus datos, el grueso de estos se ha mantenido fuera del alcance del mundo en general. Los paquetes de datos se mueven en la actualidad por la comunidad de investigadores como el yeti: «Tengo un puñado de cosas interesantes pero no puedo decir de dónde... He oído que alguien de Temple tiene toneladas de reseñas de Amazon... Creo que L tiene un pellizco de Facebook...». Esto último me lo han dicho tres académicos sin relación entre ellos. Se referían a otro científico por su nombre, que aquí he preferido ocultar. Es cierto que L tiene ese pellizco birlado a Facebook —me lo confirmó él mismo—, pero no se lo puede enseñar a nadie. Para empezar se supone que no debería tenerlo. Los datos son dinero, por lo que las empresas los tratan como tal, y aunque determinados datos digitales están a la vista de todos, se hallan blindados tras muros legales tan gruesos como los de una cámara acorazada. Si miras la página de Facebook de tu amiga Lisa, observas que se llama Lisa y publicas esa observación tuya (¡en cualquier sitio!), técnicamente habrás robado datos de Facebook. Si alguna vez te has registrado en un sitio web y has dado un código postal o una fecha de nacimiento falsos, habrás infringido alguna ley, como la Computer Fraud and Abuse Act estadounidense. Cualquier chaval de menos de 13 años que visite newyorktimes.com está infringiendo sus Condiciones de servicio y es un delincuente, no solo en teoría, sino según la doctrina de trabajo del Departamento de Justicia^[44]. Estos ejemplos son muy extremos, sí, pero las leyes que atañen a estas cosas son tan vagas en su redacción que prácticamente garantizan que todos los estadounidenses que usan Internet sean verdaderos malhechores con una larguísima estela de perversas navegaciones por la Red. Que alguien te penalice por tus «crímenes» ya es otro cantar, pero legalmente estás postrado y con una bota pisándote el cuello. El abogado de una empresa o un fiscal de distrito deseoso de complacer a algún importante donante corporativo puede arruinarte la vida con solo apretar un poco. Y cuando se tercia lo hacen. Por eso los que estudian los datos son muy reservados con sus paquetes de datos; en realidad, más que como al yeti, los tratan como si fuesen grandes bolsas de marihuana, mostrándose posesivos, algo paranoicos y siempre curiosos por saber quién más tiene hierba y si es mejor que la suya.

Cada vez es más habitual la práctica de contratar investigadores internos que la de dar información a procesar fuera^[45]. Y esa manera de funcionar nos ha dado, entre muchos otros frutos, la novedosa investigación del equipo de datos de Facebook y el excelente trabajo de Stephens-Davidowitz en Google, cuyos resultados he aprovechado aquí. Confío en que cada vez más empresas sigan ese modelo y que al final los propietarios de sitios web acabemos encontrando una manera de liberar nuestros datos para provecho público sin poner en peligro la intimidad de nuestros usuarios al hacerlo.



Ahora ya está más vista que el tebeo, pero la aplicación Shazam me pareció una de las grandes maravillas del iPhone. Es un programita que sirve para identificar música: si está sonando una canción y quieres saber cuál es, pones en marcha la aplicación y sostienes el teléfono en el aire. Shazam escucha a través del micrófono y algo así como dos segundos después te dice qué es lo que estás escuchando. La primera vez que alguien lo hizo delante de mí me quedé atónito, no solo por lo poco que le costaba al *software* acertar la canción (muchas veces funciona incluso a través de la pared o sobre el barullo de un bar), sino por lo rápido que funcionaba. Era lo más parecido a la magia que había visto jamás, al menos hasta que un día conocí a un nigromante de gran talento que, en un periquete, era capaz de recopilar todas mis facturas y desgravarme la renovación de mi cocina. Pero, en fin, luego descubrí que Shazam se basa en un principio increíble: que casi todos los temas musicales pueden identificarse por el patrón de altos y bajos de la melodía; puedes ignorar todo lo demás: la clave, el ritmo, la letra, los arreglos... Para saber cuál es la canción solo te hace falta un mapa de las subidas y bajadas de las notas. Ese contorno melódico se llama el código de Parsons, por el musicólogo que lo desarrolló en la década de 1970. El código de las dos primeras líneas de «Cumpleaños feliz» es •RUDUDDRUDUD, donde la U significa «melodía arriba» (de *up* en inglés), la D «melodía abajo» (de *down*) y la R «nota repetida». El topo • solo señala el principio de la canción, que obviamente no está ni arriba ni abajo. Tararéala para comprobarlo:



Por absurdo que parezca, el código de «Cumpleaños feliz» es prácticamente único en el catálogo entero de música grabada que existe, como lo es el de casi todas las canciones. Y a la descripción tan concisa que proporcionan esas letras se debe el que Shazam funcione tan rápido: en lugar de con una guitarra, la voz de Paul McCartney y la cantidad correcta de reverberación, la canción «Yesterday» empieza con este código: •DRUUUUUDR. Eso es mucho más fácil de entender.

Igual que una aplicación que se esfuerza por acertar una canción, la ciencia de los datos se esmera por encontrar patrones. Yo —y otros muchos que hacen mi mismo trabajo— he tenido que concebir métodos, estructuras e incluso atajos, una y otra vez, para hallar la señal en medio del ruido. Todos buscamos nuestro código de Parsons. Algo tan sencillo y a la vez tan potente es un descubrimiento irreplicable en la vida, pero por suerte hay muchas otras vidas ahí fuera. Y por muchos que puedan ser los problemas que deba afrontar la ciencia de los datos, este libro ha sido mi manera de decir: me gusta lo que nos espera.

Notas

Vivimos en un mundo en el que ya ningún lector depende de las notas finales para «obtener más información» ni para buscar pruebas de cualquier hecho o declaración. Por ejemplo, me imagino que cualquier lector interesado en la figura de Sullivan Ballou lo habrá buscado en Google antes de consultar estas notas y de escribir en la barra superior de su navegador los enlaces que ofrezco. Así que he aprovechado este apartado para centrarme en las muchas fuentes que han aportado a este libro no solo referencias, sino también ideas. También, para corroborar o explicar cuestiones referidas a los datos que he aportado yo en exclusiva.

Puesto que el tema sobre el que versa *Dataclismo* cambia casi a diario, he decidido complementar este apartado en Internet, en dataclysm.org/endnotes, donde encontrarás material original adicional y resultados de ulteriores investigaciones.

(A continuación se reproducen fragmentos de texto referenciados en el apartado de notas de la edición original. Cada fragmento destacado se completa con bibliografía).

Introducción

10 millones de personas van a usar. En esta cifra he contado a todo aquel que accedió a su cuenta de OkCupid durante los doce meses anteriores a abril de 2014: 10 922 722.

Esta noche, unas 30 000 parejas. Es la gran incógnita de dirigir una web de contactos y citas: ¿cuántos usuarios acaban por quedar en persona? ¿Y qué ocurre luego? Estos párrafos representan mis mejores conjeturas sobre algunos de los cálculos de esos encuentros. He empleado dos métodos distintos:

1. He supuesto que alguien que usa OkCupid de manera activa consigue una cita cada dos meses. Me parece una cifra conservadora. Con unos 4000 millones de usuarios activos cada mes, eso significa que cerca de 65 000 quedan cada día, o sea, unas 30 000 parejas.

2. Cada día 300 parejas se afanan en rellenar nuestra interfaz de «anulación de cuenta» para hacernos saber que ya no necesitan OkCupid precisamente porque han encontrado una relación estable en OkCupid. Estas son parejas que (a) mantienen una relación lo suficientemente seria para cerrar sus cuentas en OkCupid y que (b) están dispuestas a pasar el trago de rellenar un montón de formularios para hacernos saber cuál es el estado actual de su relación. Calculo que el grupo (b) representa solo a una de cada diez parejas formadas en la web. Y calculo que el grupo (a) representa el resultado de una de cada diez primeras citas. Por lo tanto, deben de salir unas 3000 parejas con relación estable de las 30 000 primeras citas de cada día. De cada 3000 parejas estables, pienso que menos de una de cada diez llega a casarse. Una manera de verlo es esta: ¿cuántas relaciones serias has mantenido antes de establecerte con la persona definitiva? Me imagino que el promedio se acerca a las 10.

Estas valoraciones se complementan, al menos en la cifra de «primeras citas», y aunque sean aproximadas, creo que el resto de cálculos que las siguen son plausibles.

valoraciones de pizzerías en Foursquare. Valoraciones de una muestra aleatoria de 305 pizzerías de la ciudad de Nueva York, a las que se ha accedido a través de la API pública de Foursquare.

porcentajes de aprobación del Congreso. Se recopilaron a partir de las 529 encuestas que medían la «aprobación del trabajo parlamentario», publicadas en el sitio web clearpolitics.com desde el 26 de enero de 2009 hasta el 14 de septiembre de 2013. Véase realclearpolitics.com/epolls/other/congressional_job_approval-903.html#polls.

jugadores de la NBA por la frecuencia. La tabla muestra el porcentaje de partidos iniciados de cada uno de los jugadores que aparecen en la alineación de su equipo durante la temporada 2012-2013 en espn.com. Sí, cuento a los 76ers como equipo de la NBA.

6 por ciento. Esta cifra se obtiene de tomar la media geométrica de las distancias entre cada uno de los 21 puntos que marcan las curvas. Así, para las curvas a y b, he calculado:

Lo que da un total de 0,056.

$$\sqrt{\sum_{k=1}^{21} (a_k - b_k)^2}$$

58 por ciento de los hombres. El centro de la curva de atractivo masculino está bastante más desviado y bajo que el de la femenina. Si trasladamos esa misma disparidad al CI, veremos que el CI medio masculino es algo más bajo que 85, que es el umbral de lo que se denomina «capacidad intelectual límite». El ejército, por ejemplo, no admite a candidatos cuyo CI sea menor de 85. Cuando digo «daño cerebral» estoy usando una especie de hipérbole para reflejar esta diferencia. Hablando con propiedad, quiero decir que el 58 por ciento de los hombres tendrían un CI menor de 85.

la mitad de la población soltera de Estados Unidos. Aquí me supuso todo un reto especificar el alcance de los datos sobre citas de que dispongo. Me he esmerado en hacerlo con términos amplios y fáciles de entender porque sé que, a diferencia de Facebook o Twitter, muchos de mis lectores nunca han usado una web de contactos. Si llevas casado o mantienes una relación desde finales de los noventa o antes, nunca habrás necesitado una web de contactos. Según las cifras del censo de 2011, en Estados Unidos hay 103 millones de personas solteras con edades comprendidas ente los 15 y los 64 años; ahí se contabiliza a todo aquel que no esté legalmente casado, incluida mucha gente que mantiene relaciones estables, así como prácticamente todas las personas homosexuales. Tinder, OkCupid, DateHookup y Match.com registraron en conjunto 57 millones de cuentas en Estados Unidos entre 2011 y 2013, y 23 millones solo en los últimos tres años. «La mitad» es mi cálculo aproximado de 57/103, menos el 10-15 por ciento de pérdidas por solapamiento o cuentas duplicadas.

«Las mujeres son proclives a quejarse». Esta cita es de la sección «Findings», de Rafil Kroll-Zaidi, del número de febrero de 2014 de la revista *Harper's*.

En una curva beta se plasma. Mi investigador de datos, Tom Quisel, me ayudó a explicar en términos sencillos la naturaleza binómica de las curvas beta. También señaló que se suelen emplear en modelos de predicción del tiempo y se encargó de establecer las comparaciones con los patrones climáticos de ciudades en weatherbug.com.

Cerca de un 87 por ciento de los estadounidenses. Véase Susannah Fox y Lee Rainie, «Summary of Findings», Pew Research Internet Project, Pew Research Center, 27 de febrero de 2014, pewinternet.org/2014/02/27/summary-of-findings3/.

esa cifra se mantiene. Por ejemplo, el uso de Internet entre blancos, afroamericanos e hispanos es del 85, 81 y 83 por ciento, respectivamente. Uno no puede sino conjeturar que el de los asiáticos será parecido. La adopción de Internet está por encima del 80 por ciento en todos los grupos de edad, excepto los de 65 años o más. Susannah Fox y Lee Rainie, «Internet Users in 2014», Pew Research Internet Project, Pew Research Center, 27 de febrero de 2014, pewinternet.org/files/2014/02/12-internet-users-in-2014.jpg.

Más de un tercio de los estadounidenses entran en Facebook. Facebook declaró tener 128 millones de usuarios en Estados Unidos en agosto de 2013. Facebook tenía por lo menos 1260 millones de usuarios en todo el mundo en septiembre de 2013. Las estadísticas mundiales y de Estados Unidos proceden de Wikipedia. Véase expandedramblings.com/index.php/by-the-numbers-17-amazing-facebook-stats/.

fundamentalmente populista. Esto es algo así como de dominio público entre quienes estudian la adopción de las redes sociales más allá del mero uso de Google Glasshole/Technocrat. Véase «Demographics of Key Social Networking Platforms» de Pew Research Center (2013). El informe no muestra diferencias estadísticas significativas en los índices de uso de Twitter entre los grupos con formación «secundaria o más baja» y «universitaria o más alta» (que son del 17 y el 18 por ciento, respectivamente). Pew ha encuestado a una muestra representativa aleatoria de estadounidenses de 18 años o más, así que muy pocos del grupo «secundaria o más baja» están ahí solo porque siguen estando en el instituto. Por grupos étnicos, Pew informa de índices de adopción del 29 por ciento entre negros y del 16 por ciento tanto entre blancos como entre hispanos. El informe íntegro, de Maeve Duggan y Aaron Smith, se puede ver aquí: pewinternet.org/2013/12/30/demographics-of-key-social-networking-platforms/.

Se llama investigación WEIRD. Este hecho y mi enfoque general del fenómeno se han adaptado de «Psychology is WEIRD», de Bethany

Brookshire, en *Slate*. Véase también «The Roar of the Crowd», *The Economist*, 24 de mayo de 2012, economist.com/node/21555876.

faraón Narmer. Como cabe imaginar, esto da para mucho debate, aunque Narmer, también llamado Serket, es una opción defendible. En borradores anteriores había puesto aquí a Gilgamesh, el héroe acadio, porque J. M. Roberts, en su *Historia del mundo* (Debate, 2010), elige a Gilgamesh. Acabé optando por Narmer porque su vida está datada varios siglos antes y me pareció igualmente probable que hubiese vivido en realidad. Yahoo! Answers menciona también a Elvis Presley.

1. La ley de Wooderson

Estos no son datos de una encuesta. Este es un buen momento para señalar que para que tuviese en cuenta el atractivo de una persona en el análisis realizado en este libro, dicha persona tenía que haber recibido votos de 25 personas, por lo menos. Para algo tan idiosincrásico como la atracción física, me parecía que un promedio de menos de 25 votos era poco fiable.

según el censo de Estados Unidos. Estas cifras son del «Marital Status of People 15 Years and Over, by Age, Sex, Personal Earnings, Race, and Hispanic Origin, 2011», de la Oficina del Censo de Estados Unidos.

2. Sobredosis de ni fu ni fa

«*La belleza es una apariencia externa que nunca vas a olvidar*». John Waters, *Shock Value: A Tasteful Book About Bad Taste* (Running Press, Philadelphia, 2005), pág. 128.

un concepto llamado varianza. He empleado la desviación estándar para medir la varianza en todo este capítulo.

«*efecto Pratfall*». Una búsqueda en Google de «efecto Pratfall» arrojará muchos ejemplos. Me basé especialmente en el resumen de Bill Snyder, «The Positive Effect of Negative Information», y en el artículo original que este resume, «When Blemishing Leads to Blossoming: The Positive Effect of Negative Information», de Danit Ein Gar, Zakary Tormala y Shiv Tormala, *Journal of Consumer Research* 38, n.º 5 (2012): 846-859.

Nuestro sentido del olfato. Para estos párrafos me basé en Fabian Grabenhorst *et al.*, «How Pleasant and Unpleasant Stimuli Combine in Different Brain Regions: Odor Mixtures», *Journal of Neuroscience* 27, n.º 49 (2007): 13532-13540, doi: 10.1523/JNEUROSCI.3337-07.2007. En la entrada de

Wikipedia sobre «Indol» se señala su «intenso olor fecal». Para más información sobre el uso del indol en perfumería y su presencia en el aroma natural de las flores, véase, como yo hice, **perfumeshrine.blogspot.com/2010/05/jasmine-indolic-vs-non-indolic.html**.

Aquí tienes a seis mujeres. Para proteger la intimidad de las usuarias, estos permisos los obtuvimos mediante un sistema de doble ciego. Envié los criterios (mujeres, puntuaciones con varianza elevada, atractivo general medio) al equipo de datos de OkCupid, que generó una lista de nombres que luego pasó a nuestra administradora. Ella recibió entonces una lista de nombres sin ninguna otra información adjunta y se le pidió que contactase con ellas para lograr la autorización total de uso de sus fotos (acostumbramos a recibir peticiones de fotos de usuarios para usarlas en prensa, así que este tipo de procedimientos no son inusuales). Solo se vinculó una foto con sus atributos particulares bajo expresa autorización de la usuaria.

3. Escribir en el muro

a la nostalgia se la conocía. Dado que el fenómeno es tan interesante (e insólito) y que un enlace lleva a otro, mis fuentes para esta parte fueron varias. Me informé directamente en estas:

«Dying to Go Home», de Jackie Rosenhek, *Doctor's Review*, diciembre de 2008, **doctorsreview.com/history/dying-to-go-home/**.

«Beware Social Nostalgia», de Stephanie Coontz, *New York Times*, 19 de mayo de 2013, **nytimes.com/2013/05/19/opinion/sunday/coontz-beware-social-nostalgia.html**.

«When Nostalgia Was a Disease», de Julie Beck, *The Atlantic*, agosto de 2013, **theatlantic.com/health/archive/2013/08/when-nostalgia-was-a-disease/278648/**.

La entrada sobre «Nostalgia» en qi.com: qi.com/infocloud/nostalgia.

los menores de 18 años ya no usan tanto Facebook. En la conferencia de presentación de los resultados en cuestión se revisó el rendimiento de Facebook en el cuarto trimestre de 2013. Véase Joanna Stern, «Teens Are Leaving Facebook and This Is Where They Are Going», ABC-News, 31 de octubre de 2013, **abcnews.go.com/story?id=20739310**.

El comandante Sullivan Ballou. La historia básica sobre esta carta puede leerse aquí: pbs.org/civilwar/war/ballou_letter.html. Pese a que la carta

nunca llegó a enviarse, se incluyó entre los objetos personales de Ballou que se devolvieron a su familia a su muerte.

Se habrán escrito más palabras en Twitter. Esto lo calculo así: se han escrito 129.864.880 libros, al menos según Google. Esa cifra es de una precisión que da risa, sí, pero dado que ya han subido a la Red 30 millones de libros y que indexar cosas es su ocupación, deberíamos considerar que ese cálculo es plausible. Véase Ben Parr, «Google: There Are 129,864,880 Books in the Entire World», *Mashable*, 5 de agosto de 2010, mashable.com/2010/08/05/number-of-books-in-the-world/.

Según Amazon, la extensión media de una novela es de 64.000 palabras. Puesto que es muy probable que la media y la mediana se acerquen bastante en este caso, me parece bien usarla como promedio. No creo que las novelas sean necesariamente más breves o más extensas que otros libros. Véase Gabe Habash, «The Average Book Has 64,500 Words», *PWxyz*, 6 de marzo de 2012, blogs.publishersweekly.com/blogs/PWxyz/2012/03/06/the-average-book-has-64500-words.

La combinación de estos dos números da un total de 8.311.352.320.000 palabras impresas.

Twitter declaró un total de 500 millones de tuits diarios en agosto de 2013. Véase blog.twitter.com/2013/new-tweets-per-second-record-and-how.

Calculo que cada tuit contiene 20 palabras. Así que, con 10.000 millones de palabras diarias, a Twitter le harían falta ochocientos treinta y un días (2,3 años) para superar en volumen de palabras a toda la literatura impresa hasta hoy. Esto, obviamente, pretende ser una aproximación, y bastante conservadora, por cierto. Con toda probabilidad, Twitter llegará a esa cifra mucho antes, ya que la cantidad de tuits diarios no deja de crecer.

«No tienes más que mirar Twitter». La cita del señor Fiennes recibió mucha atención en los medios. Véase Lucy Jones, «Ralph Fiennes Blames Twitter for “Eroding” Language», *Telegraph*, 27 de octubre de 2012, telegraph.co.uk/technology/twitter/8853427/Ralph-Fiennes-blames-Twitter-for-eroding-language.html.

Hasta el análisis más básico demuestra. Tanto aquí como en todos los análisis que hago de Twitter empleo los tuits y los seguidores generados

por un corpus representativo de 1,2 millones de cuentas, recopiladas al azar por mi equipo de investigadores.

El OEC es el censo canónico. Más información sobre el OEC y sus palabras más comunes en: **en.wikipedia.org/wiki/Most_common_words_in_English**.

El OEC incluye solo lemas, es decir, la raíz básica de un patrón léxico relacionado. Por ejemplo, cuenta solo la forma verbal *have* para derivados como *had*, *having*, *has*, etcétera. Yo he optado por no hacer esto en mi estudio de Twitter. Aunque esta decisión hace que sea más difícil comparar las listas directamente, he preferido presentar los datos de la manera menos procesada posible.

Mark Liberman. El blog *Language Log* del profesor Liberman (**languagelog.ldc.upenn.edu/nll/**) contiene un tesoro de interesantes análisis textuales. Véase «Up in UR Internets, Shortening All the Words», 28 de octubre de 2011, languagelog.ldc.upenn.edu/nll/?p=3532, sobre el caso concreto de la cita de Fiennes.

Un equipo de estudiosos del estado de Arizona. El análisis textual de Twitter que ocupa el resto de este capítulo está inspirado en «*Dude, srsly?: The Surprisingly Formal Nature of Twitter's Language*», de Yuheng Hu, Kartik Talamadupula y Subbarao Kambhampati, artículo presentado en la séptima conferencia anual International AAAI Conference on Weblogs and Social Media, Cambridge, Massachusetts, 8-11 de julio de 2013, **aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6139**.

Aquí presento un extracto. La tabla y la discusión subsiguiente acerca de la palabra «tribus» en Twitter proceden de «Word Usage Mirrors Community Structure in the Online Social Network Twitter», de John Bryden, Sebastian Funk y Vincent AA Jansen, *EPJ Data Science* 2, n.º 3 (2013). También me he servido de su «Additional Material», que contiene listas en bruto de palabras por comunidades que no se han empleado en el artículo. El artículo íntegro, junto con enlaces al material adicional, se puede leer aquí: **epjdatascience.com/content/2/1/3**.

Ese corpus de datos ha dado pie. Este método de rebuscar tendencias culturales en Google Libros lo propuso por primera vez la revista *Science* en el artículo «Quantitative Analysis of Culture Using Millions of Digitized Books», de Jean Baptiste Michel *et al.*, *Science* 331, n.º 6014 (2011): 176-182, doi:10.1126/science.1199644.

Mi gráfico de palabras relacionadas con la comida a lo largo del tiempo es una reproducción de su exploración de esos mismos términos en dicho artículo. Mi gráfico de años mencionados a lo largo del tiempo es una adaptación de su método, más que una reproducción. En el artículo se hace referencia a una «vida media» de la memoria que no he sido capaz de reproducir. No obstante, la afirmación por parte de los autores de que «cada año que pasa nos olvidamos con más rapidez de nuestro pasado» va claramente en la dirección correcta. El artículo contiene mucho más de interés que los dos simples gráficos que he incluido aquí y vale la pena leerlo entero.

Abajo vemos un diagrama de dispersión de 100.000 mensajes. Ningún mensaje privado fue leído por nadie al llevar a cabo este análisis. Nuestro actual *software* de detección de *spam* registra de manera automática el número de teclas pulsadas y el tiempo de tecleado de una muestra de los usuarios de OkCupid. Dado que no he leído ningún mensaje de los usuarios, el texto citado del mensaje de tres letras «Hey» es una probabilidad más que una certeza. Alrededor del 80 por ciento de los mensajes de tres letras que se escriben en la web son «hey». El siguiente más frecuente es «sup», seguido de «wow». Dada la aplastante popularidad de «hey», dado que estaba haciendo una broma y dado que las otras opciones habrían funcionado igual de bien, me pareció adecuado optar por «hey» en este contexto.

«*Yo también soy fumador*». Este mensaje privado, que presento aquí literal e íntegro, llamó mi atención en un contexto ajeno a este libro, y además recibí autorización del emisor para imprimirlo y para hablar sobre él.

4. Tú tienes que ser lo que nos una

«*gráficos sociales*». Los diagramas de redes de las páginas 80 y 81 son obra de James Dowdell, que ha usado el mismo sistema gráfico general que emplearon Lars Backstrom y Jon Kleinberg en su artículo científico «*Romantic Partnerships and the Dispersion of Social Ties: A Network Analysis of Relationship Status on Facebook*», presentado en la 18th ACM Conference on Computer Supported Cooperative Work and Social Computing, Baltimore, Maryland, 15-19 de febrero de 2014, delivery.acm.org/10.1145/2540000/2531642/p831-backstrom.pdf.

Me pasé años yendo de gira con un grupo. Mi grupo se llama Bishop Allen; Justin Rice es la otra mitad. Puedes escuchar nuestras canciones en

Spotify, o en el torrent más cercano, o en iTunes. Por si alguien está interesado, mis recomendaciones personales son los temas «Like Castanets», «Click Click Click Click», «Chinatown Bus», «Start Again» y «Little Black Ache».

En 175 apareció Leonhard Euler. Pese a que ya estaba familiarizado desde mis tiempos de la facultad de Matemáticas con Euler, el problema de los puentes y su papel en la génesis de la teoría de grafos, me he basado en la entrada de Wikipedia «Problema de los siete puentes de Königsberg» para informarme de los pormenores del problema y su solución.

nos ha ayudado desde entonces a comprender. Esta es una buena fuente sobre los usos tanto clásicos como modernos de la teoría de grafos: **world.mathigon.org/Graph_Theory**.

Stanley Milgram. Al igual que a Euler, conocía desde hace años a Milgram y su trabajo. Sin embargo, me basé en su entrada de Wikipedia para informarme de los detalles de su experimentos de los «seis grados de separación».

Facebook nos permitió ver. Véase «The Anatomy of the Facebook Social Graph», de Johan Ugander *et al.*, (prepublicaciones de arXiv.org, 2011, arXiv: 1111.4503).

Es sabido que los de Pixar. La idea fue de Steve Jobs. La primera noticia que tuve de esta anécdota fue en *Imagine*, de Jonah Lehrer (Canongate, Edimburgo, RU, 2012). Véase en BuzzFeed «Inside Steve Jobs' Mind-Blowing Pixar Campus», de Adam B. Vary, para más información. Vary en una entrevista sorprendente a Craig Payne, directivo de Pixar: **buzzfeed.com/adambvary/inside-steve-jobs-mindblowing-pixar-campus**.

«*la fuerza de los vínculos débiles*». Véase «The Strength of Weak Ties», de Mark S. Granovetter, *American Journal of Sociology* 78, n.º 6 (1973): 1360-1380.

Otra idea que se asocia desde hace tiempo. Aunque la integración ya la propuso Granovetter en 1985, el resto de mi exposición sobre la integración y las redes interpersonales se basa en la principal fuente que he empleado en todo el capítulo, «Romantic Partnerships», de Backstrom y Kleinberg. He aplicado su heurística a mis redes y he simplificado en cierta medida su trabajo original para adaptarlo a un público no académico.

un alucinante 75 por ciento del tiempo. Backstrom y Kleinberg definen numerosos tipos muy sutiles de dispersión matemática. Mi cifra hace referencia a la precisión que lograron con el método que ellos denominan «dispersión recursiva».

un 50 por ciento más de probabilidades. Esto procede del siguiente párrafo del artículo de Backstrom y Kleinberg: «Vemos que aquellas relaciones en las que la dispersión recursiva no logra identificar a la pareja tienen significativamente más probabilidades de pasar a un estado de “soltería” [es decir, de romperse] en un periodo de sesenta días. Este efecto se produce en relaciones de todas las edades y es especialmente acentuado en relaciones de hasta doce meses de duración; en este caso la probabilidad de transición es de cerca del 50 por ciento mayor cuando la dispersión recursiva no logra reconocer a la pareja».

Vete a una reunión con los de Microsoft. Puede que esto no sea algo extrapolable en general a todos los empleados de Microsoft; sin embargo, los equipos responsables de los productos para móviles y tablets de Microsoft son, según mi experiencia, perros diligentes de primer orden. Windows Mobile es tan raro de ver que destaca, así que lo recordarás cuando lo veas. Este es un buen momento para señalar que llevo toda la vida usando Microsoft Office y que todas las tablas y gran parte de los análisis que figuran en este libro se hicieron con Excel.

5. No hay mejor éxito que el fracaso

uno de los mejores diseñadores de Google. La marcha de Google de Douglas Bowman es un acontecimiento famoso en los círculos tecnológicos. Véase su propia publicación «Goodbye, Google» en stopdesign.com/archive/2009/03/20/goodbye-google.html.

No hay pruebas de que nadie engañase al sistema. Era bastante fácil recomponer una foto de Crazy Blind Date; sabíamos que iba a ser así. Es verdad que una semana después del lanzamiento, unos cuantos *hackers* habían creado aplicaciones para recomponer las fotos. No obstante, esas aplicaciones no acabaron de cuajar, sobre todo porque eran difíciles de usar y solo funcionaban algunas veces. Estos «recomponedores» no afectaron a la trayectoria de producto de Crazy Blind Date ni a los datos que generó. La foto descompuesta de ejemplo que se ha impreso en este

libro es una fotografía de banco de imágenes, con licencia de Getty Images.

6. El factor de interferencia

de cierto tipo. Véase, por ejemplo, «Blacks Still Dying More from Cancer Than Whites», de Jordan Lite, *Scientific American*, febrero de 2009. Véase también «Criminal Justice Primer for the 111th Congress», en The Sentencing Project, donde se detallan muchas disparidades deprimentes en las condenas impuestas a blancos en comparación con las impuestas a acusados de etnias minoritarias: sentencingproject.org/doc/publications/cjprimer2009.pdf.

conclusiones como esta. El titular citado procede de ThinkProgress.org. «Study: Black Defendants Are at Least 30 % More Likely to Be Imprisoned Than White Defendants for the Same Crime», de Inimai Chettiar, 30 de agosto de 2012, thinkprogress.org/justice/2012/08/30/770501/study-black-defendants-are-at-least-30-more-likely-to-be-imprisoned-than-white-defendants-for-the-same-crime.

en los 97.000 resultados. Hay que hacer un poco de trampa para conseguir que Google te dé una cifra en este caso. Mi búsqueda exacta fue «“black quarterback” -adsffsdada». Si pones el signo menos seguido de una palabra tonta evitas que la página te proponga imágenes en lugar del texto de «Aproximadamente 97.000 resultados». Estoy seguro de que si no tienes ahora mismo el navegador delante todo esto te va a sonar a chino. Pruébalo si te interesa y verás de inmediato a qué me refiero. Este es otro ejemplo de un número que ha cambiado en el transcurso de la elaboración de este libro. Lo he probado otra vez y me ha dado «89.800 resultados».

solo he encontrado un artículo. Véase Jason Lisk, «Quarterbacks and Whether Race Matters», *The Big Lead*, 2 de diciembre de 2010, thebiglead.com/2010/12/02/quarterbacks-and-whether-race-matters/. Por supuesto, el hecho de que solo haya encontrado a un autor que ha calculado las puntuaciones de los quarterbacks por razas no demuestra que ningún otro no las haya calculado. Sin embargo, me pasé varias horas peinando los resultados y solo encontré a Lisk.

los cuatro mayores grupos raciales. El 15 por ciento de los usuarios de OkCupid que seleccionan un grupo étnico escogen más de una raza; el 3

por ciento opta por una raza distinta de las cuatro mayoritarias. Esta gente está excluida de los análisis, así como aquellos que optaron por no seleccionar ninguna raza.

a «normalizar» cada hilera. Esa normalización la he aplicado frente a la media simple de cada hilera, no a la media ponderada. Dado el predominio de personas blancas, la segunda técnica habría dado un sesgo a la matriz, ya que habría usado eso que todo el mundo piensa de que los blancos son lo «normal». Con una media simple se obtiene lo siguiente: «Cuando una persona de raza A conoce a una persona al azar de raza B, ¿cómo evalúa A a B en relación con las evaluaciones de A sobre otras razas?». Esta es la pregunta interesante y lo que queremos investigar.

No es que haya un escuadrón de racistas. Un análisis de la predilección individual que tienen los hombres no negros por los perfiles de mujeres negras muestra una deducción media de 0,6 estrellas, mientras que la mayoría de la muestra aplica una deducción que oscila entre 0,2 y 1,0 estrellas. El 82 por ciento de la muestra presenta al menos cierta tendencia antinegros.

A continuación vemos nuestras cifras. Pese a que las cifras que presento aquí de OkCupid se generaron a partir de datos internos, pueden verse esas mismas cifras corroboradas y comparadas con los promedios nacionales de Quantcast en: <https://www.quantcast.com/okcupid.com?country=US>. Selecciona «Ethnicity» en el menú «Demographics» y amplía la opción «US average».

usuarios de OkCupid que lo explican con sus propias palabras. Estos extractos proceden de las «Success Stories» escritas por los propios usuarios y publicadas en la web. La de Bella y Patrick está aquí: <https://www.okcupid.com/success/story?id=2855>. La de Dan y Jenn: <https://www.okcupid.com/success/story?id=2587>.

«Hay muy pocos». La cita de Barack Obama está extraída de sus comentarios sobre el veredicto de George Zimmerman: whitehouse.gov/the-press-office/2013/07/19/remarks-president-trayvon-martin.

En un artículo se preguntaba. Véase «Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination», de Marianne Bertrand y Sendhil Mullainathan, *American Economic Review* 94, n.º 4 (2004): 991-1013, doi: 10.1257/0002828042002561.

Osagie K. Obasogie. Mi exposición sobre el trabajo de Obasogie se basa en el artículo de Francie Latour, publicado en el *Boston Globe*, «How Blind People See Race», 19 de enero de 2014. Latour esboza un resumen del libro de Obasogie *Blinded by Sight: Seeing Race Through the Eyes of the Blind* (Stanford University Press, Redwood City, CA, 2014) y entrevista al autor.

Los vigilantes de la playa. Estuve en Japón en 1992. *Los vigilantes de la playa* era ya famosa en el mundo entero, pero no se hizo popular en Japón hasta un año más tarde. No obstante, la cultura del surf, California y los cabellos rubios dorados por el sol estaban ya por todas partes. Cuando entrabas en una tienda de ropa *cool*, sonaban los Beach Boys. Temas como «Surfin' Safari», no «Kokomo».

7. El mito de la belleza en la apoteosis

proverbio coreano. Esto lo saqué de la biografía de Douglas MacArthur *American Caesar* (Little, Brown & Co, Nueva York, 1978), obra de William Manchester, que estuve leyendo para apartar mi mente de tanto dato cuando este libro estaba ya en su última fase.

la belleza opera según una escala de Richter. Ya conocía la naturaleza logarítmica de la escala de Richter, pero me he basado en la entrada de Wikipedia «Escala sismológica de Richter» para comprender mejor las magnitudes de referencia. Al comparar la belleza con esa escala me estoy concediendo, naturalmente, una licencia poética; las funciones no son exactamente las mismas.

Estos datos son de solicitudes de entrevistas. Los datos de Shiftgig los proporcionó su equipo de datos, con la gentil colaboración de Eddie Lou.

Y aquí, los datos del número de amigos. Estas son las cuentas en formato agregado y anónimas de usuarios de OkCupid que han optado por vincular sus cuentas de OkCupid con sus cuentas de Facebook.

uno de los artículos fundacionales de la psicología social. Véase «What Is Beautiful Is Good», de Karen Dion, Ellen Berscheid y Elaine Walster en *Journal of Personality and Social Psychology* 24 (1972): 285-290.

Fue el primero de una larga lista de estudios. En esta parte he adaptado conclusiones y notas directas de «Pretty Smart? Why We Equate Beauty with Truth», de Robert M. Sapolsky, en el *Wall Street Journal*, 17 de

enero de 2014. Los neuropsicólogos de Duke a los que aludo son Takashi Tsukiura y Roberto Cabeza. Véase también «Jurors Biased in Sentencing Decisions by the Attractiveness of the Defendant» en *Psychology and Crime News* para leer un resumen de los efectos de la atracción física en los procesos judiciales: crimepsychblog.com/?p=1437, publicado por la usuaria EmmaB, 3 de abril de 2007.

Que tanto Tumblr como Pinterest. Véase «Nueva política dirigida a los blogs que promueven conductas autodestructivas», blog del equipo de Tumblr, 27 de febrero de 2012, equipo.tumblr.com/post/18381730589/nueva-politica-dirigida-a-los-blogs-que-promueven.

Véase también «Pinterest “Thinspiration” Content Banned According to New Acceptable Use Policy», de Ellie Krupnick, *Huffington Post*, 26 de marzo de 2012, huffingtonpost.com/2012/03/26/pinterest-thinspiration-content-banned_n_1380484.html.

El *Huffington Post* ha tratado de manera activa el fenómeno de la «thinspiration». Véase «The Hunger Blogs: A Secret World of Teenage “Thinspiration”», de Carolyn Gregoire, 8 de febrero de 2012, huffingtonpost.com/2012/02/08/thinspiration-blogs_n_1264459.html.

Para más información sobre el «thighgap» (y para verificar que alterar los Términos del servicio no resolvió el problema), véase «The Sexualization of the Thigh Gap», de Allie Jones, en *The Wire*, 22 de noviembre de 2013, thewire.com/culture/2013/11/sexualization-thigh-gap/355434/.

8. Lo que cuenta es el interior

Esa ha sido la norma estándar desde. Estos datos básicos sobre los orígenes de Gallup se extrajeron de la entrada de la Wikipedia «Gallup (company)».

las encuestas han sido históricamente. Como menciono en el texto y en las notas al pie de este capítulo, la idea de usar Google Tendencias para buscar tabúes fue ocurrencia de Seth Stephens-Davidowitz. Su artículo del 9 de junio de 2012 en el *New York Times*, «How Racist Are We? Ask Google», y su defensa de la tesis doctoral en Harvard en 2013, «Essays Using Google Data», <http://nrs.harvard.edu/urn-3:HUL.InstRepos:10984881>, fueron la inspiración de este capítulo. Respecto al asunto de exactamente cuánto le costó su raza a Obama en las elecciones de 2008, que se recoge más adelante en este mismo capítulo,

me he basado directamente en el trabajo de Stephens-Davidowitz. En cuanto al uso del término «nigger» a lo largo del tiempo y las demás citas directas de Google Tendencias de este capítulo, el trabajo es mío, aunque he adaptado un método sugerido por él.

Aunque Stephens-Davidowitz trabaja ahora en Google, tengo que recalcar que su investigación sobre las búsquedas siempre se ha basado en fuentes públicas anónimas y no en el acceso privilegiado al historial privado de búsquedas de nadie. Mis investigaciones también se basan en una fuente pública y anónima, es decir, en Google Tendencias: google.com/trends.

Esta tendencia recibe el nombre. He usado la entrada «Deseabilidad social» de Wikipedia como fuente para los datos más básicos de este tema.

El caso más famoso. La primera vez que el efecto Bradley me llamó la atención fue durante la campaña electoral de 2008, cuando los críticos se preguntaban cómo afectarían a Obama los sondeos el día de las elecciones. Para los detalles sobre la derrota de Bradley me he basado en la entrada «Efecto Bradley» de la Wikipedia.

Desde que se hizo público este servicio. Véase Nick Bilton, «Google Search Terms Can Predict Stock Market, Study Finds», blog Bits del *New York Times*, 26 de abril de 2013. Véase también Casey Johnston, «Google Trends Reveals Clues About the Mentality of Richer Nations», *Arstechnica*, 5 de abril de 2012, arstechnica.com/gadgets/2012/04/google-trends-reveals-clues-about-the-mentality-of-richer-nations/; y Tobias Preis *et al.*, «Quantifying the Advantage of Looking Forward», *Scientific Reports* 2, n.º 350 (2012), doi: 10.1038/srep00350.

seguimiento de las epidemias de gripe. El concepto del servicio de Evolución de la gripe de Google se desarrolló inicialmente en el artículo científico «Detecting Influenza Epidemics Using Search Engine Query Data», de Jeremy Ginsberg *et al.*, en *Nature* 457 (2009): 1012-1014, doi:10.1038/nature07634. Últimamente, la eficacia del servicio está dejando que desear: véase Kaiser Fung, «Google Flu Trends' Failure Shows Good Data > Big Data», *Harvard Business Review Blog Network*, 25 de marzo de 2014.

en 7 millones de búsquedas al año. Stephens-Davidowitz, «How Racist Are We?».

más norteamericano que «apple pie». Índice de Google Tendencias para búsquedas en Estados Unidos, desde enero de 2004 a septiembre de 2013, de «apple pie»: 25. De «nigger»: 32.

Y resulta esclarecedor. La relación «nigga»:«nigger» es treinta veces mayor en los tuits enviados desde mi corpus de Twitter que lo que se aprecia en Google Tendencias. Es decir, en Twitter, «nigger» aparece con treinta veces menos frecuencia.

cerca de una de cada cien búsquedas. Stephens-Davidowitz me comentó este hecho por correo electrónico.

un 25 por ciento por debajo de los niveles pre-Obama. Stephens-Davidowitz, «How Racist Are We?». Esto también se puede verificar de primera mano a través de Google Tendencias.

Otros términos peyorativos. Estos epítetos raciales son mucho menos frecuentes en Twitter, en los mensajes privados de OkCupid y en las búsquedas de Google, tal como ha confirmado Stephens-Davidowitz a través del correo electrónico.

Si no estás muy familiarizado con. El algoritmo mediante el cual se genera la función de autorrelleno de Google es la más negra de las cajas negras. Tenemos muy poca información fidedigna sobre cómo funciona. Danny Sullivan, de searchengineland.com, ofrece información exhaustiva — aunque es bastante *ad hoc*— en searchengineland.com/how-google-instant-autocomplete-suggestions-work-62592. Dado que el autorrelleno parece tener en cuenta el historial personal de búsquedas, los resultados que arroja son muy variables. Si intentas reproducir mis resultados por tu cuenta, asegúrate de abrir Chrome en modo incógnito, como hice yo, para que Google no disponga de datos personales previos en los que basarse. Si usas Safari, selecciona «Navegación privada».

uno de los resultados. Véase Paul Baker y Amanda Potts, «“Why Do White People Have Thin Lips?” Google and the Perpetuation of Stereotypes Via Auto-Complete Search Forms», *Critical Discourse Studies* 10, n.º 2 (2013): 187-204.

Prueba a escribir en el campo de búsquedas. Esta larga cadena de búsquedas me la sugirió Sean Mathey en la furgoneta que nos traía de vuelta a casa después de una excursión de acampada en la que jugamos mucho a *Magic: the Gathering*.

Dejaré que lo explique el estratega republicano Lee Atwater. Véase Rick Perlstein, «Exclusive: Lee Atwater's Infamous 1981 Interview on the Southern Strategy», *The Nation*, 13 de noviembre de 2012, **thenation.com/article/170841/exclusive-lee-atwaters-infamous-1981-interview-southern-strategy**. Cita original del libro de Alexander P. Lamis *The Two Party South* (Oxford University Press, Nueva York, 1984), a través de la entrada «Lee Atwater» de Wikiquote.

«*Pensemos en dos mercados de difusión mediática*». Stephens-Davidowitz, «How Racist Are We?».

En mi opinión, Muhammad Ali. Leí el libro de David Remnick *Rey del mundo* (Debate, 2001) en 1999 y desde entonces admiro a Ali. He verificado determinados hechos relativos a la protesta de Ali contra Vietnam en su entrada de Wikipedia. Para la famosa cita de Ali sobre el Viet-Cong, opté por la popular y mucho más concisa cita errónea que por sus palabras exactas, que fueron: «Mi conciencia no me permite ir a dispararle a mi hermano, ni a una gente de tez más oscura, ni a una pobre gente que vive en medio del barro, por la gran y poderosa América. ¿Por qué iba a dispararles? Ellos nunca me han llamado “nigger”, nunca me han linchado, no me han soltado los perros, nunca me han robado mi nacionalidad, ni han violado y robado a mi madre y a mi padre... ¿Por qué dispararles? ¿Cómo voy a disparar a esa pobre gente? Métanme en la cárcel y ya está». La cita errónea es idéntica en espíritu, pero al ser mucho más breve y mucho más conocida decidí que resultaba aceptable en lugar de la cita auténtica.

Se le puede oír pronunciando esas palabras (la cita larga) en el vídeo de YouTube «Muhammad Ali on the Vietnam War-Draft», en **<https://www.youtube.com/watch?v=HeFMyrWIZ68>**. En ese vídeo parece estar hablando justo después de un combate y se expresa con lentitud y deliberación. Se le puede ver hablando con más fluidez sobre el mismo tema dos años más tarde en el vídeo «Muhammad Ali Interview with Ian Wooldridge (1969)», en **https://www.youtube.com/watch?v=dLam_GiQ2Ww**.

9. Días de furia

Safiyyah Nawaz tuiteó una tontería en broma. Mis fuentes de información sobre Safiyyah y sobre los tuits del desastroso asunto han sido:

Neetzan Zimmerman, «Teen Posts Joke on Twitter, Internet Orders Her to Kill Herself», *Gawker*, 2 de enero de 2013, gawker.com/1493156583.

Ryan Broderick, «Meet the 17-Year-Old Girl Who Stood Up to Death Threats After Her Tweet Went Viral on New Year's Eve», BuzzFeed, 2 de enero de 2014, buzzfeed.com/ryanhatesthis/meet-the-17-year-old-girl-who-stood-up-to-death-threats-aft.

Ryan Broderick, «After Twitter Started Viciously Attacking Her over a Silly Joke, This Girl Handled It Like a Champ», BuzzFeed, 2 de enero de 2014, buzzfeed.com/ryanhatesthis/after-twitter-started-attacking-her-over-a-silly-joke-this-g.

En estos artículos se informa de un total de 14.000 retuits de su tuit original, pero se publicaron solo un día después. Mi cifra de 16.000 corresponde a mediados de enero de 2014.

Katy Perry/Lady Gaga. El recuento de los retuits de sus tuits de felicitación del Año Nuevo corresponde a mediados de enero de 2014 y lo más probable es que se haya incrementado algo desde entonces.

la cómica Natasha Leggero. Mis fuentes sobre la broma de Leggero y el alboroto subsiguiente han sido:

«“I'm Not Sorry”: Comedian Natasha Leggero Refuses to Apologize Mocking Pearl Harbor Survivors on NBC», del legendario sabueso «DAILY MAIL REPORTER», *Mail Online*, 4 de enero de 2014, dailymail.co.uk/news/article-2533809/.

Ross Luippold, «Natasha Leggero's Stunning “Not Sorry” Response over Controversial Pearl Harbor Joke», *Huffington Post*, 4 de enero de 2014, huffingtonpost.com/2014/01/04/natasha-leggero-not-sorry-for-pearl-harbor-joke_n_4541354.html.

Los tuits despectivos enviados a Leggero se han extraído de una carta que publicó en su Tumblr: natashaleggero.com/letter/.

Fotos de su familia. El tuit de Justine y el clamor que provocó recibieron una exhaustiva atención mediática. Un repaso bastante decente del alboroto se puede encontrar aquí: «Justine Sacco: 5 Fast Facts You Need to Know», de Matthew Guariglia, en *Heavy*, 21 de diciembre de 2013, heavy.com/news/2013/12/justine-sacco-iac-racist-pr-tweet-africa/.

«This Is How a Woman's Offensive Tweet Became the World's Top Story», de Alison Vingiano, en BuzzFeed, es un estudio más concienzudo, pero que convenientemente omite el papel que tuvo

BuzzFeed en el enardecimiento de las masas: buzzfeed.com/alisonvingiano/this-is-how-a-womans-offensive-tweet-became-the-worlds-top-s.

«The Case of Justine Sacco and the Twitter Lynch Mob», de Sharon Waxman, en *The Wrap*, es un artículo de alguien que, como yo, había trabajado con Justine: thewrap.com/case-justine-sacco-twitter-lynch-mob/.

«Justine Sacco: How to Kill a Career with One Tweet», de Juana Poareo, es uno de los muchos artículos despiadados al respecto, repleto de capturas de pantalla posteriores de los tuits de Justine: *The Guardian*, «Liberty Voice», 22 de diciembre de 2013, guardianlv.com/2013/12/justine-sacco-how-to-kill-a-career-with-one-tweet/.

Se puede ver una captura de pantalla de la implicación de Google en la etiqueta #HasJustineLandedYet en «Justine Sacco Saga Sparks Criticism of Twitter Lynch Mob», de Lauren O'Neil, en [CBCnews.com: cbc.ca/newsblogs/yourcommunity/2013/12/justine-sacco-saga-sparks-criticism-of-twitter-lynch-mob.html](http://CBCnews.com/cbc.ca/newsblogs/yourcommunity/2013/12/justine-sacco-saga-sparks-criticism-of-twitter-lynch-mob.html).

Internet esperó sedienta de sangre. Aquí, aunque había muchos miles de tuits malvados de entre los que escoger en mis datos, he optado por imprimir solo tuits que ya habían sido publicados por otros medios:

El tuit de @RonGeraci aparece en su blog, The Minty Plum, en una entrada muy reflexiva, «View from the Pitchfork Mob», 12 de enero de 2014, themintyplum.com/?p=486.

El tuit de @noyokono aparece en Frazier Tharpe, «PR Woman Tweets Racist Joke Before Flight, Twitter Waits for Her to Land and Get Fired», *Complex.com*, 21 de diciembre de 2013, complex.com/pop-culture/2013/12/justine-sacco-racist-tweet/.

El tuit de @Kennymack1971 aparece en el artículo de Sharon Waxman antes citado, «The Case of Justine Sacco and the Twitter Lynch Mob».

su padre no es multimillonario. Alec Hogg, «Rubbish Rumours. Tweeting Idiot Justine Sacco No Relation to Desmond Sacco, SA Mining Billionaire», *Biz News.com*, 27 de diciembre de 2013, biznews.com/tweeting-idiot-justine-sacco-no-relation-to-desmond-sacco-sa-mining-billionaire/.

El alcance que tienen las redes sociales. En esta investigación no usamos nuestro habitual corpus aleatorio de Twitter. En su lugar, optamos por adoptar un enfoque más completista. Para estas cifras y el gráfico relacionado, mi equipo y yo extrajimos todos los retuits de la broma de Safiyyah y de #HasJustineLandedYet. Estos números reflejan nuestros mejores cálculos de quiénes vieron las dos cosas.

Los biólogos marinos. Alan Yu, «More Than 300 Sharks in Australia Are Now on Twitter», *All Tech Considered*, 31 de diciembre de 2013, NPR, npr.org/blogs/alltechconsidered/2013/12/31/258670211/.

Los rumores se mencionan. Mi fuente sobre la historia y la ciencia de los rumores ha sido el artículo de Jesse Singal «How to Fight a Rumor», *Boston Globe*, 12 de octubre de 2008, boston.com/bostonglobe/ideas/articles/2008/10/12/how_to_fight_a_rumor/. La reflexión sobre el vínculo entre rumores y violencia en las redes sociales es suya. También cita el pasaje de la Biblia «el hombre falto de juicio...». La cita «no juzgues a menos que...» es un añadido propio, así como lo del demonio Rumor.

También he usado «Rumor, Gossip and Urban Legends», de Nicholas DiFonzo y Prashant Bordia, en *Diogenes* 54, n.º 1 (2007): 19-35, y el artículo de DiFonzo «Rumour Research Can Douse Digital Wildfires», en *Nature* 493, n.º 7431 (2013): 135.

un fenómeno que estudió por primera vez. Llegué al trabajo de Suler desde *Penny Arcade*. He extraído información básica acerca de Suler y del efecto de desinhibición *online* de la entrada «Efecto de desinhibición en línea» de Wikipedia, que enlaza con el webcómic. El cómic se puede ver aquí: penny-arcade.com/comic/2004/03/19.

los viejos canales de las emisoras de radio. Me enteré de esto a través de la entrada de Wikipedia sobre el «Efecto de desinhibición en línea», donde se cita el artículo de Kenneth Tynan «Fifteen Years of the Salto Mortale», *The New Yorker*, 20 de febrero de 1978, como la fuente original.

los Jerky Boys. Para quien pueda estar interesado en el mundo del humor telefónico, Longmont Potion Castle es una especie de versión a lo Mitch Hedberg del Dane Cook de los Jerky Boys. No podría recomendar más encarecidamente el disco *Longmont Potion Castle II*.

Los radioaficionados siguen atacándose. Véase Todd Dugdale, «Sandbaggers and Trolls», kd0tls Ham Radio Experience, 6 de enero de 2014,

kd0tls.blogspot.com/2014/01/sandbaggers-and-trolls.html/.

El gobierno es quien más interés. Mis textos sobre la vigilancia gubernamental de las agitaciones y el trabajo de Peter Gloor en el MIT se basan en «What Makes Heroic Strife», *Economist*, 21 de abril de 2012, economist.com/node/21553006/.

el 27,5 por ciento de los 500 millones de tuits. Esta cifra procede del análisis de la muestra aleatoria de mi estudio.

El equipo encargado de datos de Facebook. El análisis de datos de Facebook se hace siempre mediante datos anónimos y en modo agregado. Este asunto de las repeticiones de «Nadie debería...» y la tabla adyacente proceden de Lada Adamic *et al.*, «The Evolution of Memes on Facebook», 18 de enero de 2014, facebook.com/notes/facebook-data-science/the-evolution-of-memes-on-facebook/10151988334203859. La publicación no deja claro cómo se determinó la tendencia política. Mi conjetura es que se hizo por medio de los patrones de «Me gusta» de los usuarios.

En 1950. Este párrafo que trata sobre la polarización de la política estadounidense se basa en Jill Lepore, «Long Division», *The New Yorker*, 2 de diciembre de 2013.

«*Siempre me ha parecido un misterio*». Leí *Life of Mahatma Gandhi*, de Louis Fisher (Harper & Brothers, Nueva York, 1950) en 2007, y desde entonces esta cita me acompaña siempre.

10. Muy alto para ser asiático

Para averiguar qué es de verdad lo que hace que determinado grupo. El método de reducción del texto de los perfiles recopilados de un grupo a sus fundamentos idiosincrásicos que presento en este capítulo es de cosecha propia. No obstante, en la entrada en el blog de OkCupid que inspiró ese trabajo —«The Real Stuff White People Like»— se empleó un método distinto, desarrollado con la ayuda de Max Shron y Aditya Mukerjee. No habría podido desarrollar el método que he aplicado en este libro sin el ejemplo previo de esa publicación. Lo desarrollé porque el que se empleó en dicha publicación me obligaba a ir separando las tonterías de los «datos reales» como fase final. Para este libro quería algo completamente logarítmico, que no implicase intervención humana alguna en la selección. El método es tal como se describe: colocas las

palabras y las frases en la retícula por sus percentiles y luego las clasificas por su distancia euclídea respecto al vértice deseado del cuadrado.

El elemento humano solo tuvo que intervenir en los pocos casos en los que aparecían juntas en la lista frases redundantes, como «my blue eyes and», «blue eyes and» y «my blue eyes». En esos casos usé la palabra o frase más representativa y eliminé las demás. Eso no alteró las listas de manera significativa. El método tenía en cuenta todas las frases de cuatro palabras o menos que aparecían en más de 30 perfiles.

Por razones de espacio, tres entradas largas se recortaron para evitar que partiesen la línea. En la tabla de antítesis masculinas empleé «follow me» en lugar de «follow me on instagram». En la de antítesis femeninas, usé «malcolm x» en lugar de «biography of malcolm x». Y en la tabla de palabras según orientación sexual del capítulo siguiente empleé «feminine women» en lugar de «attracted to feminine women».

lo que se denomina ley de Zipf. Ya conocía de antes las leyes de potencias y las distribuciones. Sin embargo, he aprovechado la entrada de Wikipedia sobre la «Ley de Zipf» para informarme más al respecto. Otra fuente fue «Zipf's Law and Vocabulary», de C. Joseph Sorell, *The Encyclopedia of Applied Linguistics*, 5 de noviembre de 2012. La tabla que presento aquí está extraída de una tabla más extensa que figura en ese artículo.

Los irlandeses y los europeos orientales. De *The History of White People*, de Nell Irvin Painter (W. W. Norton, Nueva York, 2010).

en México. De niño viví en México varios años y conservo cierto interés en su política. Véase Ronald Loewe, *Maya or Mestizo?: Nationalism, Modernity, and Its Discontents* (University of Toronto Press, Toronto, 2010).

«*Desde la empatía y la sexualidad*». Véase Bobbi J. Carothers y Harry T. Reis, «Men and Women Are from Earth: Examining the Latent Structure of Gender», *Journal of Personality and Social Psychology* 104, n.º 2 (2013): 385-407. «Men Are from Mars Earth, Women Are from Venus Earth» es el título del resumen de este artículo: sciencedaily.com/releases/2013/02/130204094518.htm.

Aristóteles miraba al vacío. Ya conocía el papel del firmamento en los trabajos de Einstein y Newton. Para el tercer ejemplo, y el más antiguo, rebusqué en Wikipedia hasta dar con uno que me gustó. Véase la entrada de «Éter (física)».

11. ¿Te has enamorado alguna vez?

Hace unos años, un par de alumnos del MIT. Aquí he usado el artículo «Project “Gaydar”», de Carolyn Y. Johnson, *Boston Globe*, 20 de septiembre de 2009, así como el trabajo original de los estudiantes, «Gaydar: Facebook Friendships Expose Sexual Orientation», de Carter Jernigan y Behram F. T. Mistree, *First Monday* 14, n.º 10 (2009), firstmonday.org/article/view/2611/2302.

el Informe Kinsey, de 1948. Véase la entrada «Informe Kinsey» de Wikipedia, donde se resumen las ediciones masculina y femenina del trabajo de Kinsey. La cifra del 10 por ciento de los hombres es clara y literal. Existe menos certeza en el informe en cuanto a la sexualidad femenina. El informe dice que del 2 al 6 por ciento de las mujeres de entre 20 y 35 años son «exclusivamente» homosexuales.

Estudios posteriores. Véase en Wikipedia «Demografía de la orientación sexual» para consultar todo tipo de cifras. Véase también la entrada «LGBT demographics of the United States».

«*Este trabajo puede ser de utilidad*». Dan Black *et al.*, «Demographics of the Gay and Lesbian Population in the United States: Evidence from Available Systematic Data Sources», *Demography* 37, n.º 2 (2000): 139-154.

Esto implica ciertamente un doloroso dilema. Véase Assi Azar, «Op-ed: To You There, in the Closet», *The Advocate*, 16 de abril de 2013, advocate.com/commentary/2013/04/16/op-ed-you-there-closet/.

no es más infrecuente que el cabello rubio natural. Mi fuente es el profesor C. George Boeree, de la Universidad de Shippensburg. Véase su publicación «Race» en la web space.ship.edu/cgboer/race.html. Su argumento se demuestra hasta haciendo la cuenta de la vieja: hay cerca de 1000 millones de europeos, canadienses, estadounidenses y australianos en el mundo. Si uno de cada seis de ellos es rubio natural, que en mi círculo personal sería un cálculo muy generoso, eso daría un 2 por ciento de la población mundial.

Según Stephens-Davidowitz. Mis cuatro páginas de disertación sobre las búsquedas de porno gay y sus diversas implicaciones están adaptadas del artículo de Stephens-Davidowitz «How Many American Men Are Gay?», *New York Times*, 7 de diciembre de 2013. Tanto los datos de Google

Tendencias que cito como la ampliación a los descubrimientos de Nate Silver y las cifras estado por estado de Gallup se basan en ese artículo. El artículo original de Silver es «How Opinion on Same Sex Marriage Is Changing, and What It Means», de su blog *fivethirtyeight* del *New York Times*, **fivethirtyeight.blogs.nytimes.com/2013/03/26/how-opinion-on-same-sex-marriage-is-changing-and-what-it-means/**.

Las cifras de Gallup proceden de Gary J. Gates y Frank Newport, «LGBT Percentage Highest in D. C., Lowest in North Dakota», **gallup.com/poll/160517/lgbt-percentage-highest-lowest-north-dakota.aspx**.

así como los datos de movilidad de Facebook. En su artículo, Stephens-Davidowitz también amplió su investigación a los datos de perfiles de Facebook que están a disposición del público.

que se suele atribuir a Thoreau. La cita en sí es una combinación de un fragmento del *Walden* de Thoreau con dos versos del poema de Oliver Wendell «The Voiceless». Véase The Walden Woods Project: **walden.org/Library/Quotations/The_Henry_D._Thoreau_Mis-Quotation_Page**.

El antiguo «índice de miseria» económico. Véase la entrada de Wikipedia «Misery index (economics)». Arthur Okun sugirió la formulación original.

«*Los encuestados que se identificaron*». Véase Mackey Friedman, «Considerable Gender, Racial and Sexuality Differences Exist in Attitudes Toward Bisexuality», *ScienceDaily*, 5 de noviembre de 2013, **sciencedaily.com/releases/2013/11/131105081521.htm**.

Gerulf Rieger, de la Universidad de Essex. Hago referencia a un par de artículos académicos del profesor Rieger y su equipo: Gerulf Rieger, Meredith L. Chivers y J. Michael Bailey, «Sexual Arousal Patterns of Bisexual Men», *Psychological Science* 16, n.º 8 (2005): 579-584, y su sucesor, Gerulf Rieger *et al.*, «Male Bisexual Arousal: A Matter of Curiosity?», *Biological Psychology* 94, n.º 3 (2013): 479-489.

Ellyn Ruthstrom. Véase David Tuller, «No Surprise for Bisexual Men: Report Indicates They Exist», *New York Times*, 22 de agosto de 2011, y Meredith Melnick, «Scientific Study Finds That Bisexuality Really Exists», *Time*, 23 de agosto de 2011, **healthland.time.com/2011/08/23/scientific-study-finds-that-bisexuality-really-exists/**.

En Facebook, el 58 por ciento. Véase Chris Taylor, «Fake Facebook Users Likely to Be Popular Bisexual College Women», *Mashable*, 3 de febrero de 2012, mashable.com/2012/02/03/fake-facebook-users-bisexual-college-women/.

Aunque siempre ha habido gais. Véanse en Wikipedia las entradas «Cronología de la historia LGBT» y «Salir del armario». La idea de sincerarse (es decir, salir del armario) como acto de autoafirmación procede de Karl Heinrich Ulrichs.

12. Saber cuál es tu sitio

Estados Unidos y la URSS dividieron Corea. Algo conocía de este proceso, sobre todo por el libro *American Caesar*, pero esta increíble anécdota se menciona en la entrada de Wikipedia «Division of Korea», donde se cita como fuente original el libro de Don Oberdorfer *The Two Koreas* (Basic Books, Nueva York, 2001). Confirmé la anécdota mediante una búsqueda en el texto en Google Libros: books.google.com/books/about/The_Two_Koreas.html?id=yJZKpYXh2SAC.

Este es un mapa. Este mapa, como todos los de Estados Unidos que figuran en este capítulo, y el de Reddit son obra de James Dowdell. Este en concreto se confeccionó mediante una partición Voronoi estándar de Estados Unidos, en la que cada uno de los mercados de Craigslist opera como «capital» de un «estado» (llamados «semillas» y «células»). Aunque el gráfico parece complejo, en realidad es de una sencillez muy elegante: los segmentos son todos los puntos equidistantes a las dos semillas más cercanas. He visto varias versiones más de este mapa. La mía se inspiró en la elaborada por IDV Solutions que publicó «john.nelson» en su blog de UX: uxblog.idvsolutions.com/2011/07/chalkboard-maps-united-states-of.html.

el lugar de los deseos incumplidos es Walmart. Esta es la misma distribución Voronoi, pero combinada con los datos de los estados del mapa de Dorothy Gambrell «Missed Connections», publicado en *Psychology Today*. Las células están codificadas según el mejor resultado de «contacto perdido» del estado en el que se ubica la semilla. Aquí se puede

ver el mapa original:
psychologytoday.com/blog/brainstorm/201302/misled-connections-0.

He trasladado los datos a la anterior distribución Voronoi con el fin de mantener la coherencia con el mapa previo de Craiglist.

Años atrás, un hacker estadounidense. El *hacker* es Pete Warden y su publicación es «How to Split Up the US», que se puede ver aquí: **petewarden.com/2010/02/06/how-to-split-up-the-us/**. Como señala Warden en una publicación posterior, «Why You Should Never Trust a Data Scientist», su agrupamiento de Estados Unidos en esas siete zonas es arbitraria, la versión en modo ciencia de datos del clásico «con fines de entretenimiento». Y con ese talante lo cito aquí.

Matthew Zook, geógrafo. El profesor Zook y su equipo llevan un blog fantástico titulado *Floating Sheep*, que ha sido mi principal fuente de información sobre su trabajo: floatingsheep.org.

La información y el mapa sobre el terremoto están extraídos de «Mapping the Eastern Kentucky Earthquake», publicado en el blog *Floating Sheep* por Taylor Shelton. Mi imagen es una reproducción de la original, simplificada para impresión: **floatingsheep.org/2012/11/mapping-eastern-kentucky-earthquake.html**.

El equipo de DOLLY lo componen Matthew Zook, Mark Graham, Taylor Shelton, Monica Stephens y Ate Poorthuis. Poorthuis es quien narra la historia de Sint Maarten que puede verse aquí: **www.youtube.com/watch?v=pD9HWAaQGUA**.

Mi texto sobre la revuelta de estudiantes está extraído del artículo «Beyond the Geotag: Situating “Big Data” and Leveraging the Potential of the Geoweb», de Jeremy W. Crampton *et al.*, *Cartography and Geographic Information Science* 40, n.º 2 (2013): págs. 130-139.

Abajo vemos un mapa de las descargas de porno gay. La dirección IP no señala la ubicación exacta de ninguna persona concreta (o, para ser exactos, la de ninguna dirección de un ordenador), sino solo un radio de unos 15 a 80 kilómetros. Es básicamente la misma tecnología que usa, por ejemplo, weather.com para decidir de manera predeterminada el tiempo que hace en una ciudad concreta hasta que introduces un código postal. Solo indica la zona general desde la cual accede a Internet un ordenador. De este estudio, lo único que sabemos de los propios ordenadores es el porno que se descargaban. No sabemos absolutamente nada sobre quién

estaba utilizando dichos ordenadores ni, en algunos casos, si en realidad había una persona implicada.

una mujer de 40 años de la bahía de San Francisco. Véase «I'm Just Gonna Throw This Out There. Any Redditors in the SF Bay Area Have a Empty Spot at Their Table for a Lonely Thanksgiving Orphan?», publicado por el usuario «MeMyselfOhMy» en [Reddit: reddit.com/r/AskReddit/comments/ebhh1/](https://www.reddit.com/r/AskReddit/comments/ebhh1/).

temas de cosecha propia que solo encontrarás en Reddit. Todas las publicaciones de ejemplo que se mencionan aparecieron en la primera página de sus respectivos subreddits el 30 de enero de 2013.

Los temas principales que aborda Anderson. Como muestra de la flexibilidad de su teoría, muchas de las ideas de Anderson sobre el nacionalismo son aplicables a las comunidades de Internet. Describe las naciones como «a la vez inherentemente limitadas y soberanas» y «concebidas como una intensa camaradería horizontal». Y este pasaje es especialmente aplicable a Internet: «Esta novedosa sincronía solo podía surgir históricamente cuando determinados grupos sustanciales de personas se hallaban en una posición que les permitía considerar que estaban viviendo una vida paralela a la de otros grupos sustanciales de personas, pues aunque nunca llegaban a conocerse sí que avanzaban siguiendo una misma trayectoria». Benedict Anderson, *Imagined Communities* (Verso, Londres, 1983), 6, págs. 191-192.

observar los actuales movimientos a gran escala del mundo entero. Obtuve permiso de los investigadores de Facebook Aude Hofleitner, Ta Viroit Chiraphadhanakul y Bogdan State para reproducir su mapa y hablar sobre sus resultados. Me pidieron que incluyese una explicación más consistente sobre la «migración coordinada» y sobre su estudio. Estas son sus palabras:

En una migración coordinada, una proporción significativa de la población de una ciudad ha emigrado, en grupo, a otra ciudad. Más concretamente, el flujo de población de la ciudad A (ciudad natal) a otra ciudad B (residencia actual) se considera migración coordinada si, entre aquellas ciudades en las que habita gente de la ciudad natal A, la ciudad B es aquella que alberga a la mayor cantidad de individuos cuya ciudad actual es B y cuya ciudad natal es A.

Se producen numerosas migraciones a y desde Estados Unidos, pero no exhiben esta cualidad de coordinación porque ninguna ciudad destaca sobre las demás por su atracción y la gente se traslada a distintas zonas. Este mapa muestra pedazos enteros de ciudades pequeñas y pueblos del sureste asiático que se han reubicado en masa, de forma coordinada, en núcleos urbanos.

Se puede obtener más información y ver el estudio entero en la nota «Coordinated Migration» de Facebook Data Science: www.facebook.com/notes/facebook-data-science/coordinated-migration/10151930946453859.

Como advertirás si consultas esa página, a la hora de reproducir su trabajo he modificado su mapa original eliminando las etiquetas y centrándome en una zona más pequeña de la región con el fin de que el mapa fuese más legible al imprimirlo. Gracias a Mike Develin, también de Facebook, por ayudarme a conseguir el permiso para reproducirlo. Todo el trabajo de Facebook Data Science se lleva a cabo con datos anónimos y en formato agregado.

13. Nuestra marca podría ser tu vida

Pero lo que no te dicen. Véase Clare Baker, «Behind the Red Triangle: The Bass Pale Ale Brand and Logo», Logoworks.com, 8 de noviembre de 2013, logoworks.com/blog/bass-pale-ale-brand-and-logo/.

Los arqueólogos han desenterrado. Mi texto sobre las marcas en la antigüedad se basa en David Wengrow, «Prehistories of Commodity Branding», *Current Anthropology* 49, n.º 1 (2008): 7-34, y Gary Richardson, «Brand Names Before the Industrial Revolution», NBER Working Paper n.º 13930, National Bureau of Economic Research, Cambridge, MA, 2008, <http://papers.nber.org/paper/w10411>.

En 1997, Tom Peters. Véase «The Brand Called You», de Tom Peters, publicado en *Fast Company*, agosto-septiembre de 1997, fastcompany.com/28905/brand-called-you.

todavía hoy se lee en los cursos de marketing. Véase «What a great article. I was given this to read for a class of mine, and it is written brilliantly. Great insight and information on branding. Thanks!!» (Genial artículo. Me lo dieron para que lo leyese en clase y está brillantemente escrito.

Geniales reflexiones e información sobre el branding. ¡Gracias!), comentario del usuario «Morgan» en el artículo de Peter publicado en **Fastcompany.com**.

un hombre llamado Peter Montoya. El primer trabajo de Montoya sobre el tema se titulaba *The Brand Called You: The Ultimate Brand-Building and Business Development Handbook to Transform Anyone into an Indispensable Personal Brand*, de Peter Montoya y Tim Vandehey (autopublicado, 2003). Luego volvió a publicarse como *The Brand Called You: Make Your Business Stand Out in a Crowded Marketplace* (McGraw-Hill, Nueva York, 2008), y según Amazon fue un «superventas internacional». Aquí puede leerse un PDF del primer capítulo: **petermontoya.com/pdfs/tbcy-chapter1.pdf**. La web personal de Montoya redirige a **marketinglibrary.net**, donde se le puede contratar para dar conferencias.

En Google Libros se puede ver la génesis de la idea. Para hacer este gráfico, me abstuve de contar la frase hecha «marca personal de» (como en «marca personal de su liderazgo») en los resultados de la búsqueda de «marca personal» para aislar de ese modo el fenómeno del *marketing* personal.

Dale Carnegie. Me he basado en la entrada de Wikipedia «Dale Carnegie» para informarme sobre los detalles básicos de su vida.

Por cada chaval que logra entrar en una universidad. Los dos incidentes a los que aludo aquí son la campaña de Bernie Zak para entrar en la UCLA, como se detalla en «Brookline Student Lobbies UCLA on Twitter», de Brock Parker, *Boston Globe*, 7 de mayo de 2013, y el fichaje de Rob Meyer por el *Atlantic Monthly*, como se explica en «How to Actually Get a Job on Twitter», de Alexis C. Madrigal, *Atlantic Monthly*, 31 de julio de 2013.

Véase también «The Construction of a Twitter Aesthetic», de Jason Fagone, *The New Yorker*, 12 de febrero de 2014, **newyorker.com/online/blogs/culture/2014/02/the-construction-of-a-twitter-aesthetic.html**.

los afroamericanos tienden a usar esa red social. Mi exposición sobre el fenómeno del Twitter Negro se basa en las siguientes fuentes:

Choire Sicha, «What Were Black People Talking About on Twitter Last Night?», *The Awl*, 11 de noviembre de 2009,

theawl.com/2009/11/what-were-black-people-talking-about-on-twitter-last-night.

Farhad Manjoo, «How Black People Use Twitter», *Slate*, 1 de agosto de 2010,

slate.com/articles/technology/technology/2010/08/how_black_people_u

Al artículo de Manjoo le pone el contrapunto «Why “They” Don’t Understand What Black People Do on Twitter», del doctor Goddess, en blogspot. Goddess objeta sobre todo la manera en que se retrata a los negros que están Twitter como un «monolito»; la palabra aparece dos veces en la entrada del blog y yo me he hecho eco de ello en mi texto. Véase **<http://drgoddess.blogspot.com.es/2010/08/why-they-dont-understand-what-black.html>**.

«How to Be Black Online», presentación en diapositivas de Baratunde Thurston, es un inteligente repaso del tema del Twitter Negro y da fe mejor que la mayoría de fuentes de que, como muchos tropos raciales, «Twitter Negro» es tanto «divertido porque es cierto» como erróneo al mismo tiempo. Véase **slideshare.net/baratunde/how-to-be-black-online-by-baratunde**.

Se pueden encontrar datos rigurosos sobre el uso de Twitter por grupos étnicos en el informe de Pew Research «Demographics of Key Social Networking Platforms» (2013), de Maeve Duggan y Aaron Smith: **pewinternet.org/2013/12/30/demographics-of-key-social-networking-platforms/**.

Para más pruebas de la confusión de los blancos sobre el Twitter Negro, véase Nick Douglas, «Micah’s “Black People on Twitter” Theory», *Too Much Nick*, 21 de agosto de 2009, **toomuchnick.com/post/168222309/**.

Ahora mismo hay en todo el mundo 2643 cuentas. En la web de Social Bakers ordenan cuentas de Twitter por número de seguidores. Sin duda, esa cifra habrá cambiado. Visita socialbakers.com/twitter/ y recorre los *rankings* para verlo con tus propios ojos. Para obtener información sobre los contribuyentes de Estados Unidos por ingresos, visita la página del IRS «SOI Tax Stats-Individual Statistical Tables by Filing Status», en **irs.gov/uac/SOI-Tax-Stats---Individual-Statistical-Tables-by-Filing-Status**. La información sobre la lista de multimillonarios de Forbes es de Elizabeth Barber, «Forbes’ Richest People: Number of Billionaires up Significantly», *Christian Science Monitor*, 3 de marzo de 2014,

csmonitor.com/USA/USA-Update/2014/0303/Forbes-richest-people-number-of-billionaires-up-significantly-video.

Newt Gingrich fanfarroneó. Véase Jeff Neumann, «Newt Gingrich Brags About His Twitter Followers», *Gawker*, 1 de agosto de 2011, **gawker.com/5826477/**. Véase también John Cook, «Update: Only 92 % of Newt Gingrich's Twitter Followers Are Fake», *Gawker*, 2 de agosto de 2011, **gawker.com/5826960/**.

Mitt Romney. Véase «Is Mitt Romney Buying Twitter Followers?», de Zach Green, en 140elect: **140elect.com/twitter-politics/is-mitt-romney-buying-twitter-followers/**. Mis datos y el gráfico se han adaptado a partir de los que aparecen en ese artículo.

«*Nosotros, los usuarios*». Véase Jenna Wortham, «Valley of the Blahs: How Justin Bieber's Troubles Exposed Twitter's Achilles' Heel», blog *Bits* del *New York Times*, 25 de enero de 2014, **bits.blogs.nytimes.com/2014/01/25/valley-of-the-blahs-how-justin-biebers-downfall-exposed-twitters-achilles-heel/**.

En 2012, Salesforce.com. Mi exposición sobre la oferta de trabajo de Salesforce procede de las siguientes fuentes:

Drew Olanoff, «Klout Would Like Potential Employers to Consider Your Score Before Hiring You. And That's Stupid», *TechCrunch*, 29 de septiembre 2012, **techcrunch.com/2012/09/29/klout-would-like-potential-employers-to-consider-your-score-before-hiring-you-and-thats-stupid/**.

Jessica Roy, «Want to Work at Salesforce? Better Have a Klout Score of 35 or Higher», *BetaBeat*, 27 de septiembre de 2012, **betabeat.com/2012/09/you-may-not-work-at-salesforce-unless-you-have-a-klout-score-of-35-or-higher/**.

La oferta de trabajo original seguía estando activa mientras escribía este libro, pero ya la han eliminado.

Las puertas se abren y se cierran. Véase Larry Wissel, «How Does a Logic Gate in a Microchip Work? A Gate Seems Like a Device That Must Swing Open and Closed, Yet Microchips Are Etched onto Silicon Wafers That Have No Moving Parts. So How Can the Gate Open and Close?», *Scientific American*, «Ask the Experts», 21 de octubre de 1999, **scientificamerican.com/article/how-does-a-logic-gate-in/**.

Las puertas de un microchip no son puertas en el sentido tradicional ni giran sobre sus bisagras. Emplean el voltaje para controlar el movimiento, mientras que una vieja puerta está hecha de listones de madera. Pero, como las puertas de toda la vida, controlan el paso de un espacio a otro y pueden estar abiertas o cerradas.

Target, al analizar las compras de una consumidora. Véase Kashmir Hill, «How Target Figured Out a Teen Girl Was Pregnant Before Her Father Did», *Forbes*, 16 de febrero de 2012, forbes.com/sites/kashmirhill/2012/02/16/how-target-figure-out-a-teen-girl-was-pregnant-before-her-father-did/.

una campaña de marketing de la empresa de gelatinas Jell-O. El texto sobre Jell-O y los tuits que lo ilustran proceden de Harry Bradford, «Jell-O's Fun My Life Twitter Campaign: Social Media Genius or Just "Funning" Annoying?», *Huffington Post*, 24 de mayo de 2013, huffingtonpost.com/2013/05/24/jello-fun-my-life-twitter_n_3332230.html.

McDonald's publicó un par de tuits. Extraído de Hannah Roberts, «#McFail! McDonalds' Twitter Promotion Backfires as Users Hijack #Mcdstories Hashtag to Share Fast Food Horror Stories», *Daily Mail*, 24 de enero de 2012, dailymail.co.uk/news/article-2090862/.

Wendy's había intentado. Extraído de «When Twitter Hashtag Promotion Marketing Goes Bad #HeresTheBeef», de la bloguera «stacie», en el blog *Divine Miss Mommy*: thedivinemissmommy.com/when-twitter-hashtag-promotion-marketing-goes-bad-heresthebeef/.

Más recientemente, Mountain Dew. Véase Everett Rosenfeld, «Mountain Dew's "Dub the Dew" Online Poll Goes Horribly Wrong», *Time*, 14 de agosto de 2012, newsfeed.time.com/2012/08/14/mountain-dews-dub-the-dew-online-poll-goes-horribly-wrong/.

14. Miguitas de pan

En mayo de 2013, Facebook sumaba. Véase Craig Smith, «By the Numbers: 98 Amazing Facebook Stats», *Digital Marketing Ramblings*, 13 de marzo de 2014, expandedramblings.com/index.php/by-the-numbers-17-amazing-facebook-stats/#.U1AArPldXko.

un equipo del Reino Unido. Este texto y la tabla están basados en «Private Traits and Attributes Are Predictable from Digital Records of Human

Behavior», de Michal Kosinska, David Stillwell y Thore Graepel, *Proceedings of the National Academy of Sciences* 110, n.º 15 (2013): 5802-5805.

Xbox One. Véase Stephen Fairclough, «Physiological Data Must Remain Confidential», *Nature* 505, n.º 7483 (2014): 263.

en el Reino Unido hay 5,9 millones. Véase David Barrett, «One Surveillance Camera for Every 11 People in Britain, Says CCTV Survey», *Telegraph*, 10 de julio de 2013, telegraph.co.uk/technology/10172298/.

En Manhattan. Véase Brian Palmer, «Big Apple Is Watching You», *Slate*, 3 de mayo de 2010, slate.com/articles/news_and_politics/explainer/201%5/big_apple_is_watching_you.html.

Todas esas cámaras de seguridad. Véase Jon Healey, «Surveillance Cameras and the Boston Marathon Bombing», *Los Angeles Times*, 17 de abril de 2013, articles.latimes.com/2013/apr/17/news/la-ol-boston-bombing-surveillance-suspects-20130417.

Véase también «The Need for Closed Circuit Television in Mass Transit», de Michael Greenberger, University of Maryland Legal Studies Research Paper n.º 2006-15, Law Enforcement Executive Forum (2006): 151, digitalcommons.law.umaryland.edu/cgi/viewcontent.cgi?article=1065&context=fac_pubs.

«*dominar Internet*». Esta frase en concreto se refiere a la cooperación de la NSA con los aparatos de vigilancia de otros países, como parte de la alianza «Five Eyes». Véase la entrada de Wikipedia «Mastering the Internet». La diapositiva que reproduzco circuló ampliamente después de que la publicase el *Guardian*. Véase theguardian.com/world/interactive/2013/nov/01/prism-slides-nsa-document.

«*De cada uno de los millones*». Véase David Medine *et al.*, «Report on the Telephone Records Program Conducted under Section 215 of the USA PATRIOT Act and on the Operations of the Foreign Intelligence Surveillance Court», Privacy and Civil Liberties Oversight Board (2014), <http://www.fas.org/irp/offdocs/pcllob-215.pdf>.

Las mujeres emplean apps. Mi texto sobre las apps de control de la menstruación se basa en Jenna Wortham, «Our Bodies, Our Apps: For the Love of Period-Trackers», *New York Times*, 23 de enero de 2014.

ha aparecido una startup que asegura. Este dato procede de Jaron Lanier, «How Should We Think About Privacy?», *Scientific American*, noviembre de 2013, 65-71.

todos los análisis se hicieron de forma anónima. Cabe insistir en que en ningún momento se ha vinculado ningún dato con ningún individuo. En cuanto a las fotos y el texto citados en el libro, véanse las notas anteriores al respecto.

Jaron Lanier. Mi exposición sobre el trabajo de Lanier se centra en su artículo «How Should We Think About Privacy?».

«*Gracias a datos introducidos en los motores*». Véase John Markoff, «Unreported Side Effects of Drugs Are Found Using Internet Search Data, Study Finds», *New York Times*, 7 de marzo de 2013, **nytimes.com/2013/03/07/science/unreported-side-effects-of-drugs-found-using-internet-data-study-finds.html**.

un árbol genealógico de toda la humanidad. Geni.com cuenta con más de 75 millones de entradas en su árbol. Es propiedad de MyHeritage, que declara contar con 1500 millones de entradas.

dos politólogos echaron por tierra. Véase Jowei Chen y Jonathan Rodden, «Don't Blame the Maps», *New York Times*, 26 de enero de 2014, **nytimes.com/2014/01/26/opinion/sunday/its-the-geography-stupid.html**.

Facebook, ya en 2002, recopilaba 500 terabytes. Véase Eliza Kern, «Facebook Is Collecting Your Data-500 Terabytes a Day», Gigaom, 22 de agosto de 2012, **gigaom.com/2012/08/22/facebook-is-collecting-your-data-500-terabytes-a-day/**.

Alex Pentland, del MIT. Mi exposición sobre Pentland se basa en su artículo «Reality Mining of Mobile Communications: Toward a New Deal on Data», en *Global Information Technology Report 2008-2009*, edición de Soumitra Dutta e Irene Mia (World Economic Forum, Ginebra, 2009), págs. 75-80, y en una entrevista con él, «An Interview with Alex “Sandy” Pentland About “Social Physics”», en IDcubed: **idcubed.org/?post_type=home_page_feature&p=880**.

El Washington Post captó la escasez. Véase «Million Mask March descends on Washington», en el blog PostTV del *Washington Post*: **<http://wapo.st/1b5Kt5J>**.

Conclusión

los libros de Tufte. Estos párrafos sobre el monumento a los caídos en Vietnam y la cita que reproduzco proceden de *Beautiful Evidence* (Graphics Press, Cheshire, CT, 2006), pero los textos de Tufte *Envisioning Information* (Graphics Press, Cheshire, CT, 1990) y *The Visual Display of Quantitative Information* (Graphics Press, Cheshire, CT, 2001) fueron también indispensables.

El monumento se digitalizó en 2008. Véase fold3.com/thewall y Mallory Simon, «Vets Pay Tribute to Fallen Comrades at Virtual Vietnam Wall», CNN.com, 1 de abril de 2008, cnn.com/2008/TECH/04/01/vietnam.wall/.

En esa entrada habían añadido dos fotos. El perfil del soldado Wilson en [fold3](http://fold3.com) se puede ver en fold3.com/page/631972608_lorne_john_wilson/stories/. No está claro si él aparece en la foto de grupo. Se trata claramente de una instantánea auténtica tomada en Vietnam, pero está borrosa.

Agradecimientos

Este proyecto y mi vida misma habrían volado con el viento hace mucho tiempo, como páginas sin encuadernar, sin mi mujer, Reshma. Gracias por tu apoyo incondicional, tu abnegación y tu amor.

Gracias a Max Krohn, Sam Yagan y Chris Coyne por crear OkCupid y por tenerme a su lado. Ha sido un privilegio trabajar con y para vosotros, chicos, estos últimos quince años.

Gracias a mi agente, Chris Parris-Lamb, que convirtió *Dataclismo* en una propuesta sólida a partir de una conversación inconexa en un bar, y a Amanda Cook, mi editora en Crown, que se hizo cargo del asunto a partir de entonces. Cualquier éxito que pudiera derivarse de este libro se debe a su paciencia y su pericia, que convirtieron meras ideas en algo que vale la pena leer. Gracias también a Emma Berry, asistente de edición, y al equipo de diseño, en especial a Chris Brand, por dar vida a *Dataclismo*, y a Annsley Rosner, Sarah Breivogel, Sarah Pekdemir y Jay Sones por ayudarlo a salir al mundo. El apoyo y la visión de Molly Stern, Jacob Lewis y David Drake hicieron posible todo lo anterior. Gracias también a Allison Lorentzen, de Penguin, por guiarme en mis primeros pasos por el mundo editorial.

Gracias a James Dowdell, mi versátil investigador de datos y programador. James se ha encargado del trabajo fundamental con bases de datos que hay detrás de *Dataclismo* y además ha elaborado muchos de los mapas y los gráficos del libro. Gracias a Tom Quisel y Mike Maxim por extraer (¡y reextraer!) datos de OkCupid y por haber ejercido de excelentes bancos de pruebas para mis diversas ideas estadísticas.

Gracias a mis padres y a mi hermana por sus ánimos y por ser los cimientos de mi vida. Gracias a la familia Patel por su apoyo, y en especial a Reshma, mientras dedicábamos nuestros días, semanas y meses a acabar este libro.

Gracias a Eddie Lou, de Shiftgig; a Tim Abraham, de StumbleUpon (y ahora de Twitter); a Ryan Ogle y Sean Rad, de Tinder; a Jim Talbot, de Match; a Tom Jacques, de Datehookup, y a Erik Martin, de Reddit, por sus datos y por brindarme acceso a ellos. Gracias a Michael Tapper y Ben Murray por leer borradores, y a Sean Mathey, de Mathey & Tree; a Eric Brown, de

Franklin, Weinrib, Rudell & Vassallo, y a John Therien, de Smith Anderson, por las labores legales. Gracias a Doug Demay por unos consejos que no por informales fueron menos sabios. Por último, gracias a Jed McCaleb y Justin Rice, quienes tanto me han enseñado, desde los dados de veinte caras al bitcoin, pasando por Dylan y *Ulises*; mi vida y este libro se han enriquecido mucho con vuestra amistad.

***Dataclismo* es una nueva forma de vernos a nosotros mismos, un libro de alquimia en el que las matemáticas se humanizan y los números se convierten en los trovadores de nuestro tiempo.**

Bestseller en The New York Times

La crítica ha dicho...

«Rudder nos demuestra que la información que proporcionamos a título individual dice mucho sobre quiénes somos como colectivo. Una lectura que es visualmente cautivadora y un tema fascinante hacen de este libro una elección ideal.»

Library Journal

«Demógrafos, emprendedores, estudiantes de historia y sociología y ciudadanos de a pie por igual encontrarán multitud de provocaciones y, sí, muchos datos en las páginas de este libro.»

Kirkus

«Estudiar la conducta humana es como explorar una jungla: es trabajoso y lioso, y uno se pierde con facilidad. Pero Christian Rudder es un guía experto que nos desvela verdades fundamentales sobre quiénes somos.»

Dan Ariely,
autor de *Predictably Irrational*

«*Dataclismo* es un libro repleto de secretos succulentos: secretos sobre a quién amamos, lo que anhelamos, por qué nos gusta lo que nos gusta y cómo nos hacemos cambiar de opinión y de vida unos a otros muchas veces sin siquiera saberlo.»

Jane McGonigal,
autora de *Reality is Broken*

«Una mirada fascinante, casi voyeurística, sobre quiénes somos en realidad.»

Steven Strogatz,
profesor en la universidad de Cornell y autor de *The Joy of X*

«Un libro divertido y profundo sobre temas importantes. La raza, el amor, el sexo... ¿Somos la suma de los datos que generamos? Lee este libro de inmediato y verás si eres capaz de responder a esa pregunta.»

Errol Morris



Christian Rudder (1 de septiembre de 1975, Little Rock, Arkansas, Estados Unidos) es empresario, escritor y músico.

Rudder se graduó de Little Rock Central High School en 1993. Asistió a la Universidad de Harvard y se graduó en matemáticas en 1998.

Rudder se unió a SparkNotes en octubre de 1999, pocos meses después de su fundación. Rudder fue la voz creativa de TheSpark.com, que fue el brazo de contenido viral de SparkNotes durante el ascenso temprano del sitio a la popularidad. Se convirtió en el director creativo de TheSpark en marzo de 2001. Poco después de la venta del sitio a Barnes y Noble, Rudder y los fundadores de SparkNotes (Chris Coyne, Sam Yagan y Max Krohn) se fueron y comenzaron a trabajar en OkCupid, un sitio de citas.

Notas

[1] En un artículo publicado en la revista *Slate* se señala: «Los sujetos de estudio WEIRD, procedentes de países que representan solo un 12 por ciento de la población mundial, difieren de otros grupos de población en su toma de decisiones morales, en su manera de razonar, en su imparcialidad e incluso en cosas como su percepción visual. Esto se debe a que gran parte de estas conductas y percepciones se basan en el entorno y el contexto en el que nos hemos criado». <<

[2] Esto es, en mi opinión, lo que diferencia una auténtica representación de los datos de, digamos, un simple gráfico u obra de arte impresionista que, por lo que sea, incluye números. En una representación veraz el propio espacio físico nos indica relaciones. <<

[3] Waters, sobre el cine: «Para mí, el entretenimiento no se entiende sin el mal gusto. Si alguien vomita mientras ve una de mis películas, es como si me dedicasen una ovación». <<

[4] Estas páginas, en Reddit, se llaman *subreddits*. Más adelante entraré en detalles sobre esa plataforma y sus peculiaridades. <<

[5] Solo un 0,2 por ciento de los mensajes de la web los ha enviado un usuario a una persona a quien ha asignado menos de tres estrellas. <<

[6] Definición de ignorancia: enterarse de «en qué andan hoy los jóvenes» por la Comisión del Mercado de Valores. <<

[7] Tanto Liberman como yo excluimos del análisis las URL y los símbolos espaciales como @ y #, para no hinchar artificialmente las cifras con material ajeno a las palabras. <<

[8] Los datos de Google Libros recogen el hecho de que ahora se publican más libros que los que se publicaron, pongamos, en el siglo XIX. Recopilan determinado número de libros de cada año. Así pues, aunque los dos gráficos muestran incrementos en las menciones de los términos representados a lo largo del tiempo, esa es verdaderamente una función de interés creciente. No todos los términos siguen ese patrón: «Dios», por ejemplo, decae de manera constante desde hace décadas y actualmente se emplea cerca de una tercera parte de lo que se empleaba a principios de 1800 en los textos estadounidenses. Los investigadores Jean-Baptiste Michel y Erez Lieberman Aiden acuñaron el término «culturomía» en su artículo científico «Quantitative Analysis of Culture Using Millions of Digitized Books» («Análisis cuantitativo de la cultura a partir de millones de libros digitalizados»). Los gráficos y descubrimientos que muestro aquí se han basado en su trabajo. <<

[9] He capturado los caracteres tecleados mediante un script creado ex profeso para elaborar este capítulo. <<

[10] Pruebas en contra: cuatro de los siete famosos puentes de la época de Euler han sido derribados desde entonces, dos por bombas y otros dos por una superautopista. <<

[11] El artículo original se ha citado más de 20 000 veces. <<

[12] Por supuesto, no todas las personas que están en OkCupid se incluyen en estas nítidas categorías. No obstante, para simplificar y centrarnos en el asunto, limitaremos nuestro análisis a los usuarios que han seleccionado una de las cuatro. <<

[13] Las mujeres negras obtienen cerca del 75 por ciento de la cantidad de primeros mensajes que reciben las demás mujeres. Sus mensajes son respondidos también con una frecuencia del 75 por ciento en comparación con las demás. <<

[14] Ahora bien, está claro que las webs de contactos no son ni de lejos una fuente general perfecta. Como todos sabemos, casi todos los usuarios son solteros, y eso tiene sus consecuencias. Por ejemplo, si me sentase a investigar hábitos de consumo y concluyese a partir de nuestros datos que el hombre medio estadounidense se gasta todos sus ingresos disponibles en restaurantes y entradas de cine, estaría haciendo el tonto. Una cosa así, que obviase la peculiar naturaleza de mis datos, sería absurda. <<

[15] Para ser claros, ese «nosotros» no es retórico: me incluye también a mí.
<<

[16] Aquí he superpuesto las líneas de tendencia porque los datos son ligeramente más escasos y, por tanto, presentan más interferencias de lo que es habitual. La muestra es de ≈ 5000 personas. <<

[17] El estudio de la belleza por métodos tradicionales resulta especialmente susceptible al problema de la insuficiencia. Si el tema de tu estudio es, digamos, la salud, puedes hacer con facilidad una medición de los ingresos netos de una persona y acto seguido pasar al atributo subordinado que pretendes analizar. Pero para estudiar la belleza primero tienes que determinar cuán guapos son tus sujetos, un proceso que consume muchos recursos. Al ser la belleza tan extremadamente subjetiva (en comparación con, digamos, el color del cabello, que si preguntas por él obtienes ligeras variaciones — castaño, marrón, cobrizo— que son básicamente sinónimos), se obtienen opiniones que oscilan enormemente y que solo pueden absorberse mediante un muestreo de estudio muy amplio y variado. Como hemos visto en los estudios WEIRD, ese no es que haya sido uno de los puntos fuertes de la investigación académica previa. <<

[18] Google Tendencias expresa la popularidad de una búsqueda mediante un simple número o índice proporcional a la cantidad de búsquedas de determinada palabra o frase. Los índices de este epíteto no varían más de un 10 por ciento entre las citadas áreas metropolitanas. La variante «nigga» no está incluida en la búsqueda, puesto que la mayoría de las búsquedas que la incluyen corresponden a letras de canciones rap (los términos de búsqueda exactos que generaron los datos que presento en este capítulo fueron «nigger -nigga -song»). Las búsquedas relacionadas más frecuentes de «nigger» son, con creces, «jokes» (chistes) y «nigger jokes» (chistes de negros). Para mi análisis de búsquedas raciales me baso en un método creado por Stephens-Davidowitz, estadístico y economista de Google. Al respecto de su visión interna de los datos, escribe: «Una gran proporción de las búsquedas [de “nigger”] fueron acerca de chistes sobre afroamericanos». Para su estudio emplea datos públicos y también anónimos. <<

[19] Esto no solo tiene que ver con que la gente se fuese de vacaciones: términos neutros como «pasta», «pizza», «familia» y «camión» se mantuvieron estables durante todo el año. <<

[20] Si alguna vez se cansan en Facebook de esa «f» tan minimalista y buscan un nuevo icono gráfico, sugiero este: sobre fondo azul, dos personas blancas que discuten sobre lo que otra persona blanca ha dicho sobre África. <<

[21] Sería interesante ver si los residentes de los países en los que todavía se usa la lapidación como castigo se divierten tanto con la versión digital. <<

[22] En Australia, esas etiquetas van equipadas con transmisores que advierten a los bañistas de determinada playa sobre la presencia de un tiburón. Y esas etiquetas se comunican con nosotros... a través de Twitter. <<

[23] Y, al igual que ahora en Internet, los usuarios tenían hasta «apodos». <<

[24] La Primavera Árabe, por ejemplo, supuso el debut de Twitter como herramienta de relevancia global, y la red social ha facilitado asimismo protestas en Guatemala, Moldavia, Rusia y Ucrania. <<

[25] Otro ejemplo, mucho menos famoso, es: $e^{i\pi} + 1 = 0$. Aquí, sorprendentemente, los cinco valores más importantes de las matemáticas componen una sola ecuación. Se llama la identidad de Euler, por cierto. No se rompió los cuernos para bautizarla. <<

[26] Este ejemplo se ha adaptado de «Zipf's Law and Vocabulary», de C. Joseph Sorell, Victoria University de Wellington. Como cualquier ley empírica, la de Zipf es un buen marco descriptivo (y de eficacia comprobada), pero como se puede apreciar aquí se produce cierta varianza en los resultados observados. Es como saber que al tirar al aire una moneda saldrá cara la mitad de las veces. Sin embargo, después de tirarla mil veces es muy poco probable que exactamente la mitad de ellas haya dado cara. <<

[27] El algoritmo ha puesto todas las palabras en minúsculas, de modo que así las presento aquí. <<

[28] Véase un mapamundi de alrededor de 1491 para mayor información. <<

[29] Los datos de encuestas suelen estar contaminados por factores externos; por ejemplo, cómo formula el investigador las preguntas o cómo decide comparar la experiencia sexual con la identificación sexual. <<

[30] Salvo a la carrera política, religiosa y del espectáculo de las personas que la perpetúan. <<

[31] Esto se basa en dos suposiciones: 1) que alrededor del 5 por ciento del país es gay, y 2) que, de los 93 millones de solteros que hay en Estados Unidos según el censo, la mitad buscan pareja. El gobierno considera «soltero» a todo aquel que no está casado, lo que obviamente es problemático a la hora de calcular la verdadera población soltera, sobre todo entre la gente homosexual. En 2013, OkCupid registró actividad en 650 000 perfiles declaradamente gais, lo que, según esta aritmética, es un 26,8 por ciento de la población gay estadounidense que busca pareja activamente. Una pequeña fracción de las cuentas son duplicadas o «fantasmas» (que apenas se usan), pero pese a ello la porción que aloja el sitio web del mercado de contactos gais del país es significativa. En esta nota, así como en el resto del capítulo, se cuenta por separado a los usuarios «gais» y «bisexuales», y estos cálculos no incluyen a los segundos. <<

[32] Existen apps de ligoteo homosexual específicas para sexo ocasional: Grindr y Scruff son las más conocidas para hombres. El equivalente hetero de estas apps es Tinder. En proporción es igual de popular, tal vez más incluso. Por lo tanto, no creo que esa tendencia selectiva (por buscar relaciones duraderas) de la población homosexual de OkCupid sea peor que la de la población heterosexual, aunque debo admitir que esa es una cosa imposible de saber con certeza. <<

[33] El 49 por ciento de los hombres gays y de las mujeres lesbianas han declarado haber tenido menos parejas. <<

[34] Dos meses más tarde, Zook midió una convulsión de distinta índole: los Kentucky Wildcats ganaron el campeonato de la NCAA y los estudiantes se emborracharon y empezaron a quemar de todo como los futuros líderes que sin duda alguna son. El hashtag #LexingtonPoliceScanner empezó a ser *trending topic*, basado sobre todo en este tuit de @TKoppe22: «Em... Tenemos a un varón parcialmente desnudo con un tanque de propano #LexingtonPoliceScanner». Zook rastreó esa etiqueta para mostrar cómo lo que antes habría sido una tontería local puede ahora tener repercusiones a nivel mundial. La esquizofrenia entre intelectualidad y zafiedad que tiene Twitter nunca deja de sorprenderme. Es el Chris Farley de las tecnologías. <<

[35] Soy consciente de que una condición adicional es que la gente afectada use Twitter, y de que en el contexto de la Mesoamérica precolombina esa es una expectativa absurda. No obstante, como ya he señalado, la red social tiene mucha más penetración y es mucho más democrática de lo que la gente cree, y si hoy fuese a ocurrir algo parecido a la conquista española, con toda certeza veríamos sus reverberaciones en Twitter. <<

[36] En diciembre de 2013 tuvo 101 millones de visitantes únicos y sirvió 5 millones de páginas. <<

[37] Su mantra, por cierto, es: «Distinto... o extinto». <<

[38] La lista Forbes de multimillonarios de 2014 tiene 1645 miembros. <<

[39] Uno de los antiguos miembros del equipo de Newt contó en Gawker: «Alrededor de un 80 por ciento de esas cuentas están inactivas o son cuentas fantasma creadas por diversas “agencias de seguidores”, otro 10 por ciento son personas reales que forman parte de una red de gente que se sigue mutuamente y que también paga por tener seguidores (el perfil de Newt forma parte de estas redes porque él las utiliza, pero no les sigue a cambio), y el 10 por ciento restante puede que sean de verdad personas reales con sentimientos a quienes les gusta Newt Gingrich». <<

[40] Como muestra de su capacidad analítica, cabe señalar que son los dueños de data.com. <<

[41] De la reseña en la revista *Nature* sobre la consola: «Está equipada con una cámara que es capaz de monitorizar el ritmo cardiaco de las personas que están en la misma habitación. El sensor está diseñado principalmente para los juegos en los que hay que hacer ejercicio y permite que los jugadores controlen los cambios de su ritmo cardiaco durante la actividad física, pero en principio ese mismo tipo de sistema podría controlar y transmitir datos sobre respuestas fisiológicas ante anuncios televisivos, películas de terror o incluso... retransmisiones sobre política». <<

[42] De la web de Acxiom: «[Damos] a nuestros clientes el poder de gestionar audiencias de manera satisfactoria, personalizar la experiencia de los usuarios y crear relaciones provechosas con el cliente». Una paradoja interesante: cada vez que ves escrita la palabra «personalizar», sabes que las cosas se han puesto sumamente impersonales. <<

[43] Después de lo de Boston, Reddit y 4chan se esforzaron tenazmente (es decir, que se escribió mucho al respecto) por seguir la pista de los terroristas y acabaron por «cazar» a un hombre inocente. Pese a toda la palabrería vertida en la nube por las masas, el *hardware* acabó resolviendo el crimen. <<

[44] Para más información sobre las implicaciones kafkianas de la ley CFAA, véanse los artículos «Until Today, If You Were 17, It Could Have Been Illegal to Read Seventeen.com Under the CFAA» y «Are You a Teenager Who Reads News Online? According to the Justice Department, You May Be a Criminal», ambos publicados por la Electronic Frontier Foundation. <<

[45] Me encantaría que eso se llamase «fumadero hermético de maría». Pero por desgracia no es así. <<