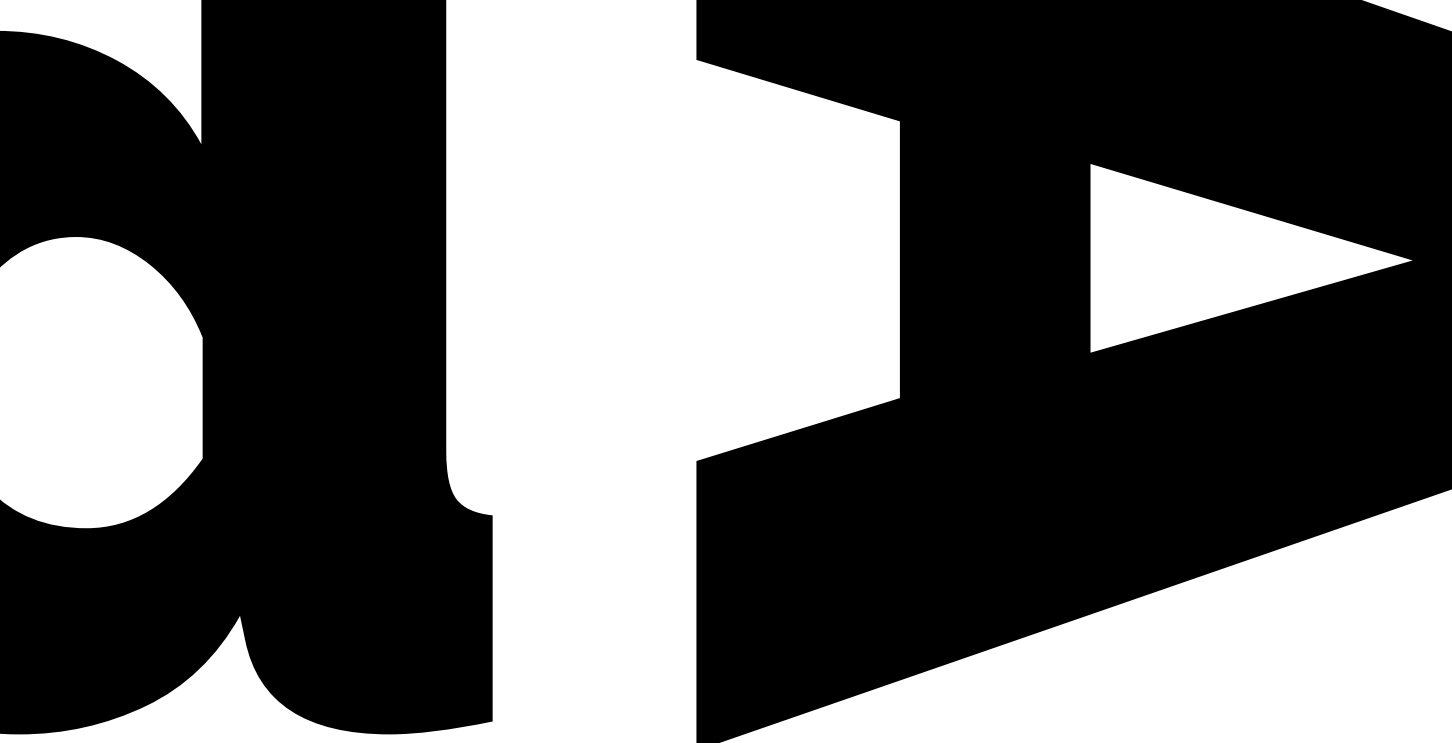




Cómo sacar partido del análisis de datos *big data*

Google, Amazon y Netflix fueron pioneras en el uso del big data. Hoy, son muchas las empresas que están aprovechando el potencial de los datos para mejorar sus procesos y crear nuevas prestaciones. El *big data* también puede ayudarte a tomar mejores decisiones en tu empresa y este artículo explica cómo hacerlo.

Por Inés Alegre, profesora Adjunta de Análisis de Decisiones, IESE; Miguel Ángel Ariño, profesor Ordinario de Análisis de Decisiones, IESE; Miguel Ángel Canela, profesor Agregado de Análisis de Decisiones, IESE.

Es difícil no haber oído hablar del *big data* a estas alturas, pero ¿qué es exactamente y para qué sirve? ¿Tengo *big data* en mi empresa? ¿Cómo puede el *big data* ayudarnos a tomar mejores decisiones? La expresión *big data* sugiere grandes volúmenes de

información, imposibles de procesar por los sistemas habituales de tratamiento de datos. Pero, en el *big data*, la cuestión no es únicamente el tamaño, sino otras características, que los expertos sintetizan en las tres V:

VOLUMEN. Estamos habituados a hablar de *megabytes* o de *gigabytes*. Un giga son 1000 megas. En *big data*, podemos hablar de órdenes de magnitud muy superiores, que pueden llegar a



La expresión *big data* sugiere grandes volúmenes de información, imposibles de procesar por los sistemas habituales de tratamiento de datos.

los *terabytes* (1000 *gigas*) o incluso a los *exabytes* (1000 millones de *gigabytes*). Se trata, pues, de grandes volúmenes de datos difíciles de manejar en un ordenador

personal y con el *software* convencional.

VELOCIDAD. En algunas aplicaciones de *big data*, los datos se generan de forma continua y se han de procesar en tiempo real, o dentro de una ventana de tiempo que se va reduciendo progresivamente. Es lo que llamamos *streaming data*.

VARIEDAD. A menudo, se trata de datos no estructurados. Al hablar de bases de datos, nos viene a la cabeza la imagen de una colección de tablas, con filas y columnas, en su mayor parte llenas de números. En *big data*, además, manejamos datos no estructurados, que pueden ser numéricos; textuales, como comentarios en Facebook, Twitter o en el blog de la empresa; o multimedia, como fotografías, canciones, etc.

Pueden provenir de fuentes muy diversas y hallarse en diferentes formatos (CSV, JPEG, PDF, etc.). La tecnología para trabajar con *big data* empezó a desarrollarse en empresas como Google, Amazon y Netflix, pero su expansión a organizaciones de todos los tipos y tamaños está siendo imparable, tanto para la mejora de los procesos existentes como para crear nuevas prestaciones. Veamos algunos ejemplos.

La tecnología para trabajar con *big data* empezó a desarrollarse en empresas como Google, Amazon y Netflix, pero su expansión a organizaciones de todos los tipos y tamaños está siendo imparable.

· Dos de los sectores en los que el uso del *big data* está más extendido son el bancario y el de las compañías telefónicas. Los bancos usan datos de sus clientes que captan a través de los pa-

gos con tarjeta y movimientos de sus cuentas para ofrecerles uno u otro producto. Las compañías telefónicas emplean los datos del uso del teléfono para predecir qué clientes tienen una mayor probabilidad de cambiar de compañía y así planificar una campaña de marketing a medida para lograr retenerlos.

· La página *web* de una empresa genera infinidad de datos continuamente. Cada vez que alguien la visita y hace un clic, se sabe a qué hora entra, cuánto tiempo está, a qué hora sale, a qué otra página de la misma *web* acude, de qué página procede, etc. La empresa puede diseñar experimentos para averiguar qué posiciones, dentro de la página, son las mejores para que los usuarios visiten lo que a la empresa más le interesa y lograr así una mayor retención.

· La cadena de establecimientos Macy's ha logrado, en los tres últimos años, un incremento de ventas del 10 % gracias al uso inteligente del *big data* en sus operaciones de logística y *pricing*. Macy's recoge datos en tiempo real de las ventas y el *stock* disponible en cada una de sus tiendas y adapta los precios y las ofertas de manera dinámica en más de 73 millones de artículos.

· Mediante el uso de datos en tiempo real, una conocida cadena de comida rápida controla la cantidad de personas que están esperando para ser atendidas. Si la cola es larga, anuncia en sus pantallas productos que son rápidos de cocinar y servir. Si no hay cola, anuncia productos que, aunque son más lentos de cocinar, tienen un margen de beneficios más alto.

· Una aerolínea de Estados Unidos está interesada en disponer de un *rating* de satisfacción del cliente que le permita compararse con la competencia. El estándar es el American Customer Satisfaction Index (ACSI), que aparece una vez al año. Una alternativa de muy bajo coste puede obtenerse mediante la captura de los mensajes de Twitter que mencionan a cada compañía. Cada semana se

La cadena de establecimientos Macy's ha logrado, en los tres últimos años, un incremento de ventas del 10 % gracias al uso inteligente del big data en sus operaciones de logística y pricing.

publican alrededor de 10 000 tuits mencionando cada una de las principales aerolíneas de Estados Unidos. Existen metodologías para diagnosticar de un modo agregado y suficientemente fiable el grado de satisfacción que transmiten los clientes a través de sus comentarios. Es la técnica llamada *sentiment analysis*. Esto permite medir con una frecuencia semanal la satisfacción de los clientes para cada aerolínea. De esta forma, se pueden captar las tendencias en la percepción de los clientes sobre los servicios prestados por las aerolíneas y tomar medidas de acuerdo con esta información. Esta técnica es muy útil para conocer el grado de aceptación y la imagen pública de una empresa, así como su evolución en el tiempo.

Estrategias de precio, de servicio al cliente, de *benchmarking*, de retención de clientes y *marketing*

son solo algunos ejemplos de los numerosos y variados usos con impacto real en los resultados de negocio que tiene el *big data*.

En la mayoría de los casos, no hemos de capturar nuevos datos, sino sacar partido de los que ya tenemos. Por ejemplo, un banco dispone de datos de los movimientos de las cuentas bancarias o de los pagos con tarjeta de cada uno de sus clientes; un supermercado sabe qué tipo de productos compran los clientes que tienen tarjeta de fidelización; una empresa dispone de datos sobre sus empleados, sus clientes, sus productos o su cadena de suministro y es posible extraer otros datos con relativa facilidad de redes sociales o de información pública.

Deténgase.

Lea atentamente.

Piense.

**“Aquellos que son
estáticos y repetitivos
son aburridos. Aquellos
que son dinámicos y
aleatorios son confusos.
En el medio yace el
arte”**

John Locke
(1632 – 1704), filósofo inglés

¿Por dónde hemos de empezar? El primer paso, que debe dar la propia dirección previo a la obtención de los datos, consiste en preguntarse qué provecho se va a sacar de ellos.

PLAN DE GESTIÓN DEL BIG DATA

Si mi empresa quiere usar el *big data*, ¿por dónde hemos de empezar? El primer paso, que debe dar la propia dirección (tú, lector), previo a la obtención de los datos,

consiste en preguntarse qué provecho se va a sacar de ellos. ¿Qué pregunta queremos responder? ¿Qué datos son necesarios para responder adecuadamente a esa pregunta? Y, una vez hecho el análisis, ¿de qué va a servir?

¿Qué decisión tomaremos y qué impacto tendrá en nuestra compañía? El segundo paso consiste en identificar los recursos que necesitamos para usar esos datos de la forma que queremos. Los recursos pueden ser de tres tipos:

A. *Hardware*. Con las inversiones en *hardware* hay que ser cauto. Los proveedores de *cloud computing* ofrecen un amplio abanico de prestaciones que nos permite ajustar los recursos contratados a nuestras necesidades de forma continua, de acuerdo con la fórmula “pago por uso”. Por consiguiente, no hay que pensar que debemos acometer grandes inversiones de entrada.

B. *Software*. También aquí hay que ser cauto. La mayor parte del *software* usado en la gestión y el análisis de *big data* es libre. En vista de la capacidad de innovación mostrada por la comunidad Open Source, podemos esperar que eso siga así en los años venideros.

C. *Personal*. Es, como en muchas otras cuestiones, el factor crítico, por varias razones. La primera es que la demanda supera a la oferta cualificada. La segunda, que hay que distinguir tres roles fundamentales. En primer lugar, la visión estratégica y la identificación de las oportunidades corresponden a la dirección. Sin esos inputs, el análisis de datos

no tiene sentido. En un plano más técnico, hay que distinguir entre *data engineering* y *data science*.

El data engineer, también llamado *data architect*, organiza los recursos para la captación, almacenamiento y circulación de los datos, mientras el *data scientist* se encarga del análisis. No obstante, en una organización pequeña, o con volúmenes de datos moderados, la separación entre estos dos últimos roles puede quedar difuminada.

Finalmente, no hay que olvidar la fase de evaluación de los resultados, reflexión y toma de decisiones. ●

Artículo publicado en la Revista de Antiguos Alumnos del IESE, abril-junio 2017.