



## BIG DATA PARA CEOS Y DIRECTORES DE MARKETING

Isaac González Díaz  
Publicación independiente  
(2017)

El Big Data es uno de los fenómenos con mayor impacto en el mercado desde hace varios años. La transformación digital, impulsada por el COVID-19, todavía ha acelerado más su adopción a lo largo de los sectores de la economía afectando a todo tipo de empresas, incluida la administración pública. Lo digital son datos y los datos son el activo más valioso que poseen actualmente las empresas. Todos los directivos son conscientes de su importancia y saben que no integrar en su estrategia los beneficios de Big Data significa perder competitividad y posiblemente en unos años quedar fuera de un mercado dominado cada vez más por los algoritmos. Están obligados a dar este paso. A partir de ese momento y una vez tomada la decisión, surgen las siguientes cuestiones lógicas:

- ¿Para qué me va a servir todo esto?
- ¿Qué puedo esperar de una forma realista?
- ¿Realmente me puede servir para mis objetivos o es solo una moda de la que todo el mundo habla, pero en realidad pocos entienden?
- ¿Todas las tecnologías son iguales? ¿Por qué hay tanta variedad y para qué sirve cada una?

Lo que parece una gran oportunidad se convierte instantáneamente en incertidumbre relacionada con el riesgo de no "conocer el terreno", de tomar una mala decisión o de desperdiciar tiempo, dinero y esfuerzo. En este contexto y con el objetivo de dar respuesta a estas incógnitas surge el presente libro que pretende ser de gran ayuda.

En la actualidad existen muchos textos escritos orientados a personal técnico, pero muy pocos dirigidos a los responsables del negocio o decisores finales, que expliquen el ámbito de Big Data de una forma sencilla, práctica y orientada a la aportación de la cuenta de resultados. Esto es posible porque el autor presenta un perfil y una amplia experiencia que combinan la parte técnica y de negocio de una manera docente y comprensible, que se transmite a lo largo de todo el texto. Sitúa la explicación en un nivel intermedio y funcional para aportar valor (sin tecnicismos, ni líneas de código) y sin recurrir a clichés vacíos de contenido que no aportan valor ni

ejecución práctica. De esa forma, el libro constituye un "tour guiado" y progresivo que recorre los componentes que un directivo debe conocer sobre Big Data para poder aplicarlo en su empresa.

El manual se estructura en 5 pilares. En el primer pilar se abordan los conceptos generales que son necesarios comprender. Se comienza determinando qué es Big Data desde un punto de vista práctico. Esto es, no son sólo las "3 Uves", si no una definición que rápidamente permite entender el momento en el que es necesario utilizar Big Data, en contraposición a otros en los que se puede obtener grandes beneficios del análisis de datos sin la utilización preceptiva de Big Data. Acto seguido, se explican los usos más directos por los que se puede comenzar. Durante el resto del libro se explican muchos otros, pero los que se exponen en este pilar son los imprescindibles para que una empresa puede empezar a obtener los primeros resultados, especialmente en el área de la reducción de costes. Adicionalmente, se describen las diferencias entre términos que muchas veces se confunden como Big Data, Data Science, Machine Learning o Inteligencia Artificial, entre otros. Esto es muy importante en actividades empresariales tan necesarias como la contratación del talento o de consultoría para poder seleccionar en el mercado de forma adecuada, en función de las necesidades y recursos de que se dispongan.

El segundo pilar corresponde a la tecnología, donde se explica con un lenguaje comprensible y sin tecnicismos las principales tecnologías que se encuentran en el ámbito de Big Data. Se comienza por Hadoop. Cómo funciona, cuáles son sus fortalezas y sus debilidades, en qué tipo de casos tiene sentido y cuáles no, etc. Se explican a continuación las tecnologías accesorias que surgen precisamente para cubrir las debilidades de Hadoop. En este capítulo se presentan otras tecnologías de amplia utilización en el mercado como Pig, Hive, Storm o MongoDB, donde se especifica su concepto y funcionalidad. También se describen desde una perspectiva práctica conceptos como procesamiento batch frente a tiempo real o las ventajas y desventajas de la implementación on-premise frente al cloud. Tras Hadoop el resto del capítulo se centra en exponer Spark, que en la actualidad es la gran tecnología dominante para el desarrollo de Big Data. Por último, se finaliza con un resumen ilustrativo y clarificador: un gráfico que muestra en una sola página la relación entre todas las tecnologías Big Data. Constituye un cierre de gran valor en este capítulo ya que permite al lector obtener un "mapa mental" que le ayudará a tener una visión general de las tecnologías existentes, que es posiblemente la parte más difícil de entender en Big Data.

El tercer pilar es la parte analítica, que se centra en las metodologías y los algoritmos. Comienza exponiendo los cuatro grandes tipos de analítica en el entorno empresarial. El primero es la analítica descriptiva. Simple, pero de gran utilización en cualquier

análisis. Posteriormente se aborda el segundo tipo, el Business Intelligence, y conceptos imprescindibles como Datamart, Datawarehouse, Datalake, OLAP, Cuadro de Mando, KPIs, ETL, etc., que aportan al directivo el conocimiento técnico mínimo necesario para no estar desorientado y poder entender los desarrollos de sus equipos técnicos. El Business Intelligence permite conocer y entender lo que ha pasado. Pero cada vez más las empresas están interesadas en conocer lo que va a suceder en el futuro próximo con mayor probabilidad. Esto se presenta en el tercer y cuarto tipo de analítica: la predictiva y la prescriptiva. Este ámbito es donde se genera más valor en Big Data y, por tanto, corresponde a la parte más extensa del libro.

Comienza explicando la metodología para construir un modelo predictivo. Se explican las actividades que se desarrollan en las fases de calidad de datos, transformación y creación de variables, modelización, y evaluación. También se repasan los principales algoritmos usados en aplicaciones empresariales y las métricas que permiten evaluar la calidad de los modelos predictivos. Esto es de gran utilidad para el directivo para entender el desarrollo de proyectos llevados a cabo por sus equipos o consultores, así como para poder medir los resultados a través de la calidad de los modelos generados. También se exponen de forma práctica las ventajas e inconvenientes de algoritmos como los árboles de decisión, la regresión logística y redes neuronales, así como cuándo se recomienda usar cada uno de ellos. Adicionalmente, se dedica un apartado sobre técnicas más avanzadas. Esto responde al enfoque conocido como Social Network Analytics, que permite identificar las relaciones entre los propios clientes de la empresa y, cómo estas relaciones, ocultan información de gran valor que suele pasar desapercibida.

El autor lo explica a través de casos ilustrativos. Por ejemplo, a la hora de evaluar el riesgo de una empresa quizá ésta presente unas cuentas preliminares saneadas. Sin embargo, si se encuentra en un ecosistema de clientes-proveedores donde varios de sus clientes muestran problemas financieros, seguramente tendrá dificultades para cobrar sus facturas a 30/60/90 días. Por tanto, donde el análisis individual pudiera arrojar preliminarmente un riesgo bajo, el análisis social lo completaría aportando indicios de alto riesgo y activando las alarmas correspondientes. También se detallan las bases del text mining, que permite el análisis de información en texto totalmente desestructurada con la misma facilidad que se analiza una tabla de datos estructurados.

El cuarto pilar corresponde a los casos de uso. Una vez que el lector ha entendido de manera progresiva la tecnología y los algoritmos este capítulo responde a la gran pregunta: ¿y todo esto para qué sirve en el contexto empresarial? Se presenta de manera resumida e ilustrativa a través de un gráfico que reúne gran parte de los casos de uso más frecuentes en la empresa. A continuación, el resto del capítulo se cen-

tra en explicar cada caso de uso en el que se aporta una visión general que permite situar al lector en las aplicaciones prácticas de Big Data en casi todos los ámbitos. Se describen desde aplicaciones en estrategia y multicanalidad, hasta todo lo relacionado con el ámbito comercial, de marketing y de gestión de clientes. También las aplicaciones en los ámbitos de riesgos, operaciones, fraude, y por supuesto del mundo digital. En este capítulo el libro realiza una aportación diferencial en el mercado, ya que hay pocos textos que expliquen esta capa de negocio.

El quinto y último pilar del libro hace referencia al management. Para aquellos lectores que tengan o vayan a tener responsabilidad en algún área relacionada con Big Data, este es posiblemente el capítulo que le puede aportar un mayor valor desde el punto de vista de aplicación práctica. La gestión es un ámbito lleno de incertidumbre ante la dificultad de encontrar soluciones objetivamente óptimas. Hay que tomar decisiones de alto impacto con poca información, donde el componente cualitativo es en muchos casos mayor que el cuantitativo. Por eso, los directivos requieren la participación de asesores o mentores que aporten conocimiento y experiencia ante los retos y desafíos que se les presentan. Este es el papel que desarrolla el autor en este capítulo. Ofrece su experiencia y comparte consejos claves para quienes tengan que crear productos basados en datos, crear y liderar una unidad analítica y gestionar un equipo de alto talento, vender o gestionar proyectos de Big Data o saber cómo gestionar las fricciones que puedan aparecer con otras áreas de

la empresa. Es un capítulo breve pero muy valioso para quien se encuentre en esta posición y responsabilidad.

Este libro destaca sobre todo en dos cosas: claridad y estructura. En cuanto a la claridad, la mayor aportación radica en que permitirá al lector entender por fin qué es el Big Data y cómo puede aplicarlo en la empresa de una manera sencilla, comprensible y práctica. Simplemente qué es, para qué sirve, cómo se puede usar y los conceptos mínimos que un directivo o de perfil de negocio debe adquirir para poder abordarlo de forma exitosa. Y en cuanto a la estructura es la responsable en gran parte de la claridad. La organización del contenido como una hoja de ruta gradual y acumulativa a través de los distintos pilares facilita su adopción, avance y comprensión final de una manera progresiva y sencilla. En definitiva, se trata de un libro muy indicado para quien tenga o vaya a tener algún tipo de responsabilidad en el ámbito de Big Data. Además, su extensión permite que se puede leer en dos fines de semana.

Una recomendación final: el libro no debe guardarse muy lejos porque en la medida que se gana experiencia en este ámbito, resulta imprescindible consultarlo de nuevo, ya que en cada revisión se aprecian cosas nuevas que seguramente habían pasado desapercibidas anteriormente.

■ **Javier Cabezas Arias**