



# INICIO GRABACIÓN

MINERIA DE DATOS // DATA WAREHOUSE.  
IMPLEMENTACION BIG DATA

<https://github.com/rb-one/curso-practico-machine-learning/blob/master/Notes/Notes.md>



**SANJOSÉ**  
FUNDACIÓN DE EDUCACIÓN SUPERIOR



## INDICE

1

**IMPLEMENTACION BIG DATA**

2

**CONCEPTOS**

3

**EJERCICIOS**

4

**CONCLUSIONES**

Bibliografía: LaValle, Steve,  
Michael Hopkins.

“The new path to value: How  
the smartest organizations



**SANJOSÉ**  
FUNDACIÓN DE EDUCACIÓN SUPERIOR

PRESENTACIÓN

# Implementar Big Data en su Organización



- Cuando el objetivo es la implementación de Big Data en la organización, lo primero que se debe hacer es conocer y cuantificar cuánto valor puede ofrecer al negocio. Es necesario llevar a cabo una autoevaluación que permita conocer cuáles son los límites que impiden avanzar, o acelerar el progreso, qué expectativas se tienen a medio y largo plazo o qué restricciones afectan a la recogida, procesamiento y análisis de la información.
- Una vez respondidas todas estas cuestiones, se está en posición de decidir cuál es el proyecto de Big Data más adecuado para la empresa



## Elegir el Data Warehouse más Adecuado para Big Data

- El primer paso es la elección de tecnología. Hay que seleccionar la que mejor se adapta a las necesidades corporativas, en otras palabras, la que facilitará la consecución de los objetivos de negocio. Son las metas las que determinan la tecnología, y no al revés.
- Por ejemplo, si se quiere realizar un análisis de eventos en tiempo real, hará falta contar con una tecnología del CEP para poder determinarlo; mientras que, si el objetivo es realizar un análisis predictivo, entonces se necesitan bases de datos con capacidades de procesamiento MPP que puedan hacer un análisis de lo actual contra todo el histórico.







## Conocer y Anticiparse al Negocio es Imperativo

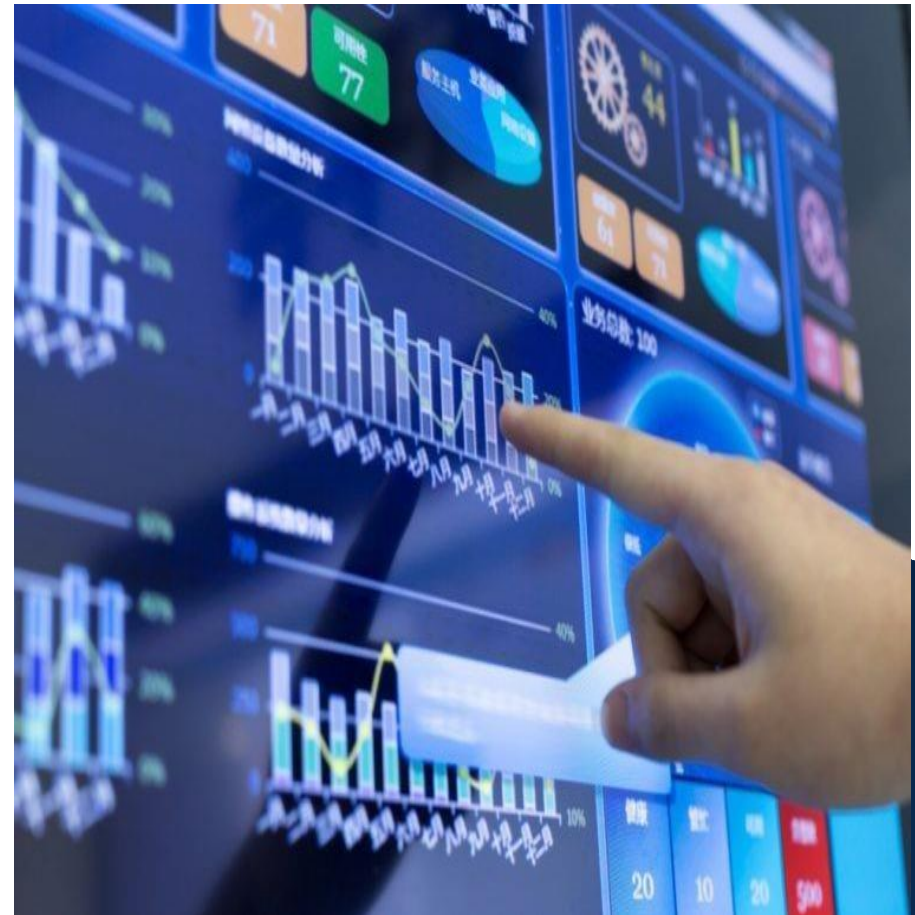
- Toda iniciativa de Big Data necesita perfiles de negocio que conozcan la empresa y el sector en detalle, ya que van a aparecer nuevos roles, como el Data Scientist o Data Analyst, que harán posible, en base a esa información, acceder a un nivel de conocimiento superior, inalcanzable hasta el momento.
- Por ejemplo, para realizar un análisis completo de sentimiento de la percepción que se tiene de una tu marca y la empresa que hay detrás, es necesario conocer a la perfección el producto que se ofrece, los canales que se emplean, los perfiles a que se dirige o el tipo de campañas de marketing que se lanzan. Sólo así es posible encontrar toda la valiosa información que hay detrás de los datos.

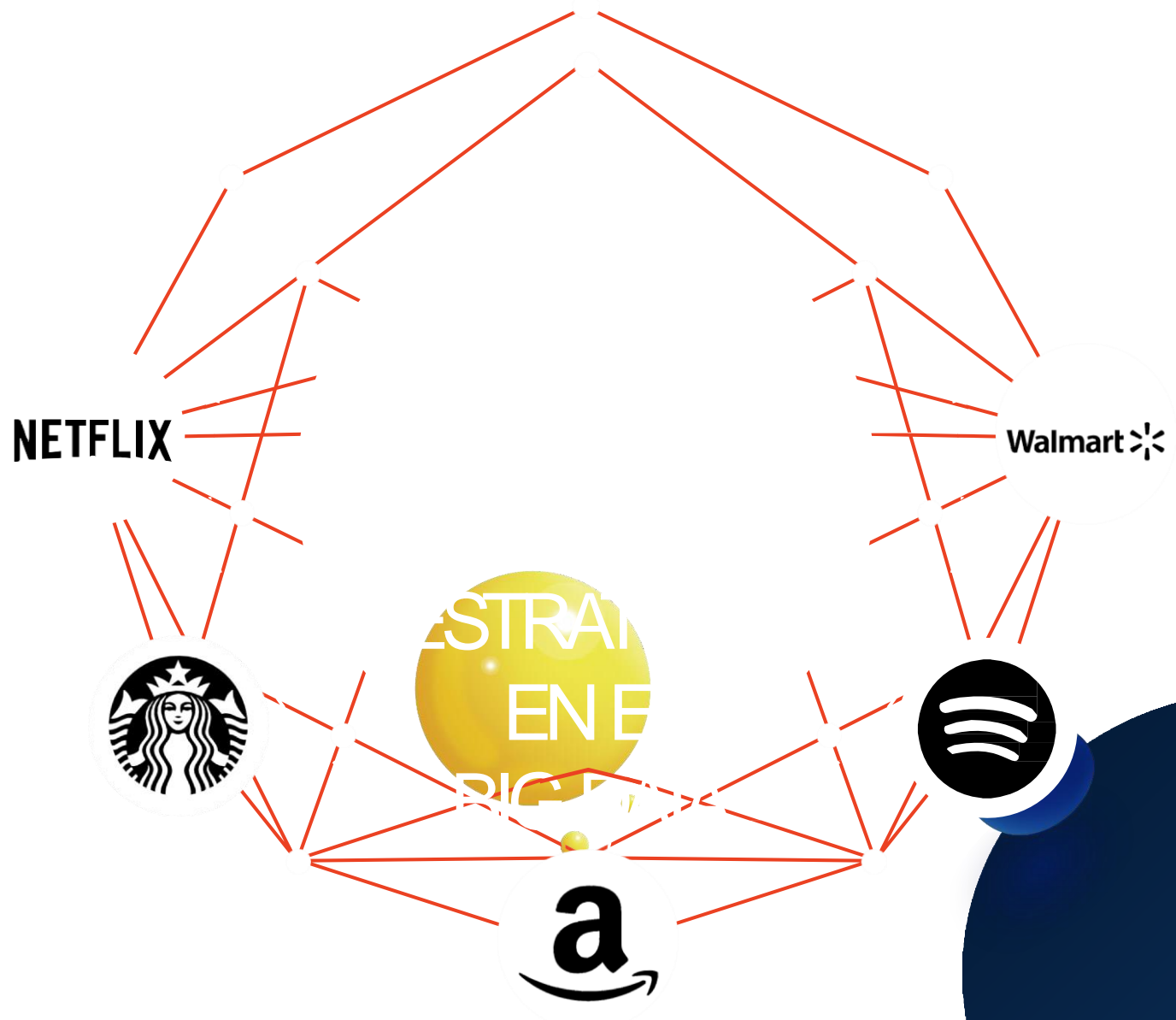
## Encontrar las herramientas de análisis correctas



- En función del tipo de análisis que se quiera realizar se elegirán unas herramientas u otras. No

son iguales las herramientas que se emplean al realizar un análisis predictivo que las usadas para llevar a cabo un análisis en tiempo real.

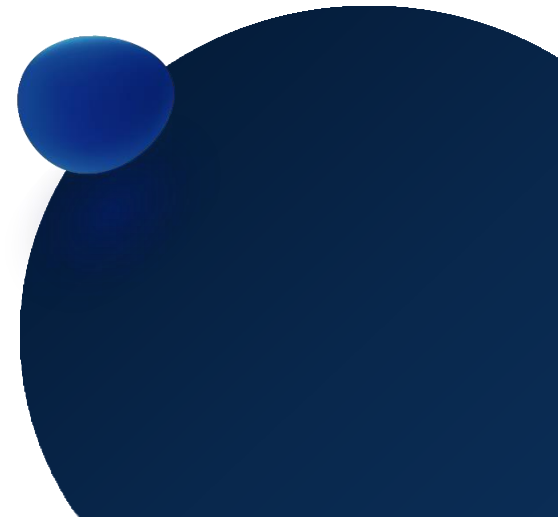








**El Big Data supone una gran oportunidad para tu negocio.  
¿Cómo? Aplicando las estrategias y técnicas del Big Data podrás  
alcanzar tus objetivos, diferenciarte de la competencia o conocer  
con mayor precisión el comportamiento y las necesidades de tus  
consumidores.**



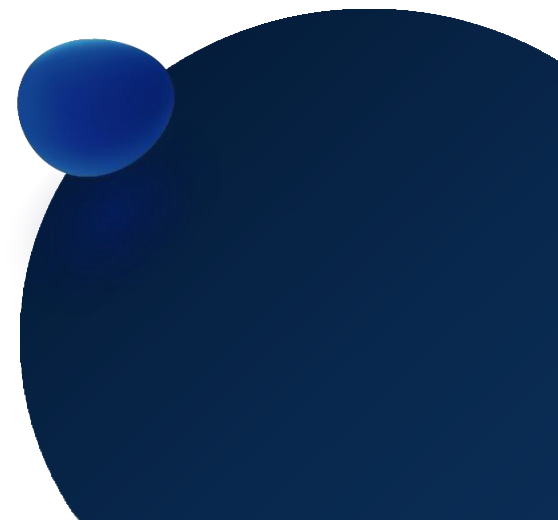


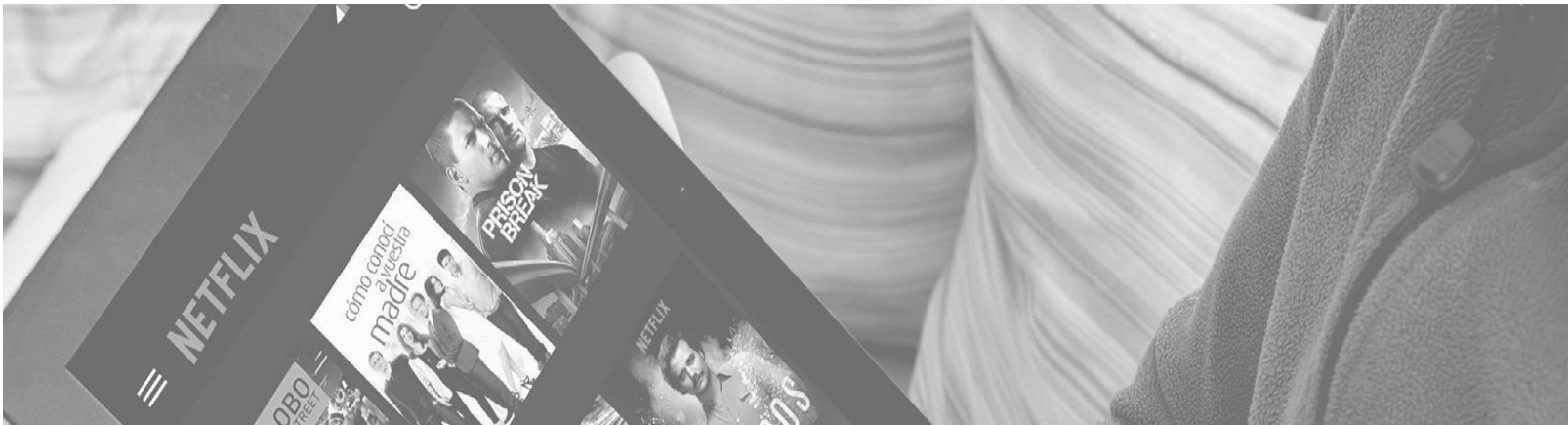
# 1. Netflix

Netflix empezó siendo una empresa que alquilaba películas en formato de DVD por correo. En 2009 comenzó a usar un **algoritmo para predecir los gustos de sus suscriptores**.

Hoy día tiene más de 80 millones de usuarios en todo el mundo.

**80 MILLONES DE USUARIOS**





### ¿Qué datos de clientes recoge?

- Horarios de visualización
- Dispositivos que usan
- Tipos de búsquedas que realizan
- Su comportamiento durante la serie (pausa, rebobinado, créditos, etc.)
- Etiquetado de las películas (taggers)
- Valoraciones de los usuarios



### obtiene?

186,6 millones de dólares de beneficios en 2016, un 52% más de beneficios que en 2015.

¿Qué resultados

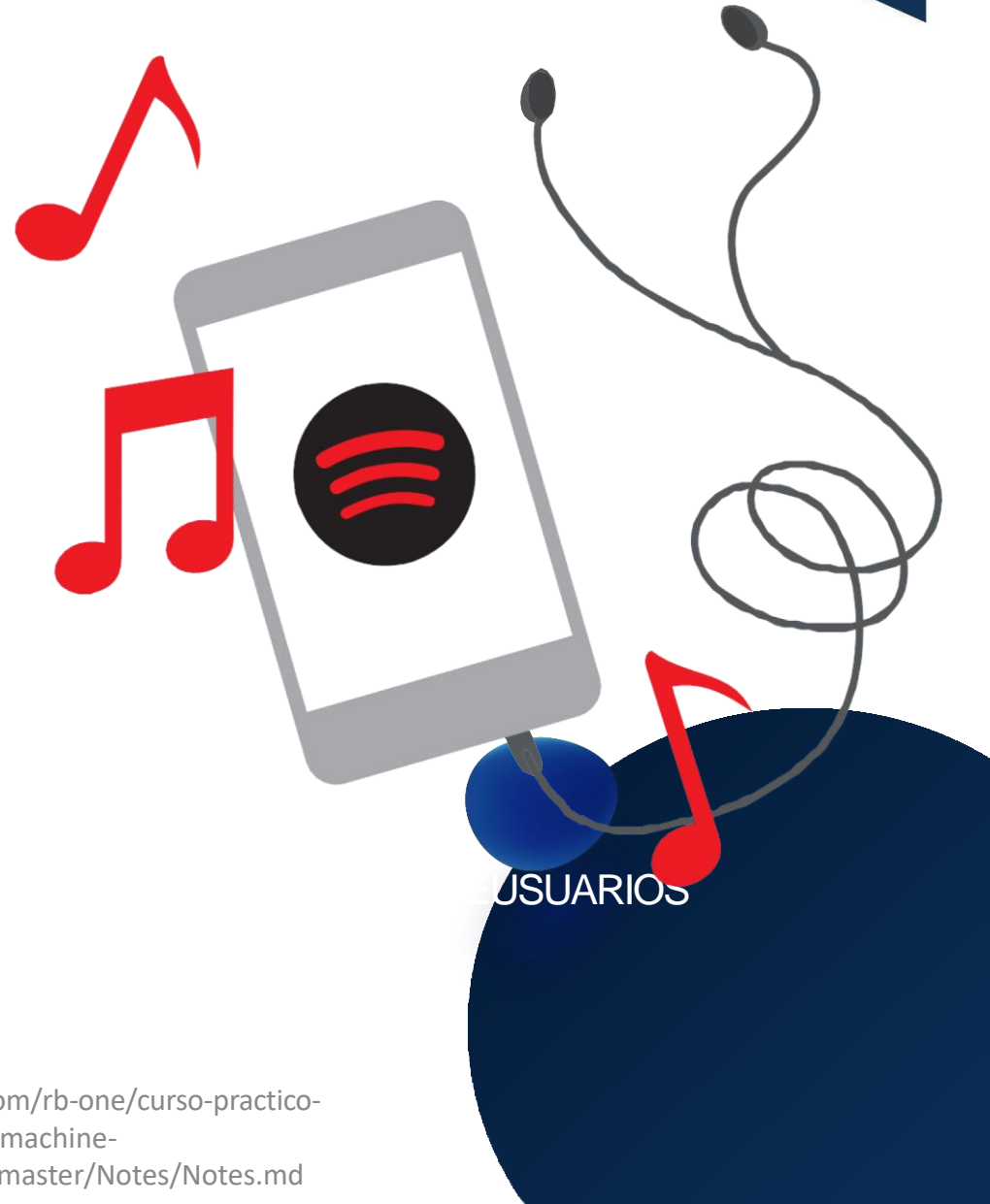
Gracias a:

- Hipersegmentar el cliente y adaptar las series y películas a su demanda
- Ofrecer a sus clientes recomendaciones de contenido afín a sus gustos
- Generar contenido en función de la s.p.r.e.f.e.r.e.n.c.i.a.s de sus

## 2.Spotify

La plataforma de música en streaming se lanzó en 2008 y desde entonces ha registrado más de 24 millones de usuarios activos, de los cuales 6 millones son premium.

En 2006 Spotify compró The Echo Nest (compañía especializada en Big Data musical) con la que extrae datos para reconocer las preferencias de los usuarios y las tendencias musicales.



¿Qué datos recoge?



- Géneros más escuchados
- A qué hora escuchan música
- Dónde la escuchan
- Suscripciones a una playlists concreta
- Popularidad de los artistas
- Tiempo de reproducción (si pasas la canción antes de 30seg es un “no me gusta”)



¿Qué resultados obtiene?

50 millones de suscriptores de pago  
Aumento de 5 millones en altas de usuarios en 2016 respecto a 2015

Gracias a:

- Establecer patrones de conducta para poder hacer recomendaciones y listas de descubrimiento en función de los gustos musicales de cada usuario



### 3Amazon

Es el mayor marketplace del mundo. Utiliza una **estrategia basada en el análisis predictivo** capaz de construir un sistema de recomendación de productos para las personas que visitan la web.

En 2016 cerró con 152 millones de cuentas de clientes.

152

MILLONES DE CUENTAS DE  
CLIENTES





¿Qué datos recoge?

- Secuencias de clics (valoraciones y likes)
- Datos históricos de compra de los usuarios
- Tiempo en el sitio web
- Duración de la vista
- Comparación de compras similares con otros clientes
- Ítems en los carritos de compra virtuales



¿Qué resultados obtiene?

Amazon cuadruplicó su beneficio en 2016, hasta los 2.371 millones de dólares. Sus ventas crecieron un 27% en 2016, 135.987 millones de dólares.

Gracias a:

- Ofrecer a los clientes una experiencia de compra personalizada
- Los motores de recomendación con los que logran el 29% de sus ventas



## 4. Starbucks

¿Has visto alguna vez cuatro Starbucks a pocos metros de distancia? No están ahí por casualidad. La famosa cadena de cafeterías utiliza el **BigData** para **determinar cuál es el preciso lugar donde abrir una tienda.**

Adiá de hoy, más de 6 millones de personas están registradas en su loyalty program.



**6 MILLONES DE  
PERSONAS**

**REGISTRADAS EN SU LOYALTY  
PROGRAM**





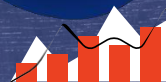
¿Qué datos recoge?

• Localización de la tienda:

- Datos demográficos de los clientes
- Población
- Patrones de tráfico
- Paradas de transporte público
- Tipos de negocio en los alrededores

• Ofertas personalizadas

- Productos que gustan a los clientes
- Cómo toman el producto (con leche, sin azúcar...)
- Horario
- Lugar



¿Qué resultados obtiene?

Un 10% más de ingresos en 2016

Gracias a:

- Determinar la rentabilidad de abrir una tienda en un lugar u en otro
- Ofrecer a sus clientes ofertas y promociones que les resulten interesantes

<https://github.com/ab-sme/cursopractico-strategy/#4d39d6903d71>

learning/blob/master/Notes/Notes.md



## 5. Walmart

Con 245 millones de clientes, Walmart es el retailer más grande del mundo. **Gracias al BigData identifica patrones de comportamiento.**

Su objetivo actual es optimizar la experiencia de compra de los clientes tanto en la web como en el punto de venta.



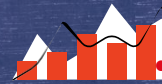
**245 MILLONES DE  
CLIENTES**





### ¿Qué datos recoge?

- Análisis semántico del texto en el buscador de la web - Ticket medio
- Duración de las visitas
- Frecuencia de las visitas



### ¿Qué resultados obtiene?

- Ingresos de 482.1 billones de dólares en 2016. Un 10% más de conversión de sus visitas en la web.
- Gracias a:
- El lanzamiento de nuevos productos según las preferencias de sus clientes
- Las recomendaciones a sus clientes en función de sus necesidades y gustos
- Los modelos de análisis predictivos que le permitan identificar patrones de conducta





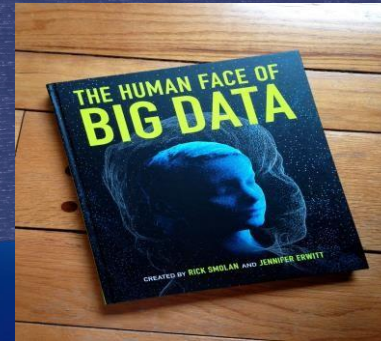
Como ves, las empresas más grandes del mundo ya basan sus estrategias de negocio en el Big Data con grandes resultados.

Establece tú también una estrategia de Big Data en tu empresa

y  
**aumenta la certeza y éxito** de tus decisiones de negocio.



## Humanizando el Big Data



- Existe un documental llamado “The Human face of Big Data” que muestra que Big Data es algo más habitual en el día a día de lo que se cree.
- Muchas de las decisiones que se toman, derivan en acciones que se hacen realidad gracias a la utilización de Big Data. Ejemplos de ello son las compras a través de internet, tras cada una de las cuales se esconden multitud de datos a tener en cuenta; la conducción de un vehículo mediante sensores; o el simple hecho de acudir al médico, donde el historial de cada paciente está digitalizado, centralizado y permite ser comparado con el de otros pacientes.

Cómo Big Data Impacta en tu Día a Día





## Big Data: Su efectividad está en las personas

- Digo que tenemos un problema porque, cuando no se entiende ni se sabe trabajar con él, el Big data se convierte en “Junk data”, es decir, en “datos basura”. Ves datos, pero realmente no sabes lo que ves. Y para poder entender y aprovechar el potencial del «Big data», un profesional de Recursos Humanos necesita saber lo que ve, algo prácticamente imposible si no se domina Excel, o los conceptos de bases de datos en general, y no se han adquirido unas nociones mínimas de estadística, psicología social y antropología social, por citar unas pocas, a lo que habría que añadir alguna de las competencias recomendadas por Davos, como por ejemplo el “**Pensamiento Crítico**”. Si estos requisitos faltan, los datos solo sirven de adorno.
- Un ejemplo. Desde hace años, los departamentos de RRHH manejan volúmenes ingentes de información. ¿Quién la está usando y para qué? He visto entrevistas de salida de varios años olvidadas sin que nadie dedicara un segundo a analizarlas. ¿Para qué se hacen entonces? Con frecuencia he escuchado afirmaciones carentes de fundamento y, lo que es peor, he visto tomar e implementar decisiones a partir de esas afirmaciones carentes de





## Big Data: Su efectividad está en las personas

- El hecho es que RRHH no progresa adecuadamente en su nivel de influencia organizativa y una de las razones es que no sabe usar los datos. Y guste o no guste, “sin datos solo eres una opinión más”, como solía repetir uno de mis managers.
- El Big data es fantástico y sus posibilidades son enormes. Sin duda. Pero, para aprovecharlas, hacen falta unos perfiles profesionales con determinadas competencias que a día de hoy distan mucho de ser el estándar. Y hasta que esto no cambie, el Big data será simplemente otro palabra de moda más, con escaso valor. Porque, en realidad, y como ocurre con otras muchas cosas, el Big data por sí mismo sirve para poco. Su efectividad está en las personas.







## Big Data Permite Innovar

- En una empresa, no es lo mismo copiar e ir detrás de los competidores que ser el primero en ofrecer un servicio nuevo. De ahí proviene la importancia de la innovación que Big Data propicia, que permite hacer volar la imaginación sin límites al eliminar por completo cualquier tipo de restricción en la gestión de datos.





# Seguridad

- La cantidad de datos que hoy día se puede analizar debe ser tratada teniendo en cuenta el nivel de privacidad que los clientes y usuarios de negocio están dispuestos a aceptar.
- Por ejemplo, en Google no todo es gratis. El precio que se paga es el de la privacidad. Al interactuar con el motor de búsqueda, éste recopila información personal que ayuda a completar el perfil individual de cada usuario.
- Las empresas deben encontrar el equilibrio entre la cantidad y tipos de datos que son capaces de analizar de sus clientes y los límites de tolerancia de éstos con respecto al acceso a su privacidad. Descubrir tendencias puede no interesar tanto si para hacerlo hay que sobrepasar determinadas barreras.





# EVERYBODY LIES

BIG DATA, NEW DATA,  
AND WHAT THE INTERNET  
CAN TELL US ABOUT WHO  
WE REALLY ARE



SETH STEPHENS-DAVIDOWITZ

FOREWORD BY STEVEN PINKER

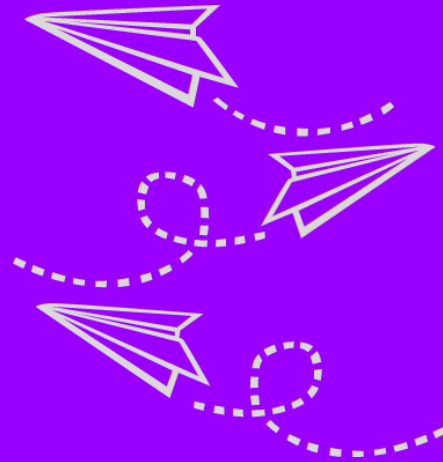
<https://github.com/rb-one/curso-practico-machine-learning/blob/master/Notes/Notes.md>



# CONCLUSIONES

- El bigdata debe perseguir un fin, tenemos que trazar una ruta clara y lógica para conseguir los objetivos propuestos mediante estos análisis de datos.
- El bigdata es una estrategia a seguir para el tratamiento de la información, una hoja de ruta para elaborar esos objetivos y las herramientas de software necesarias que necesitamos para ponerlo en práctica.

Bibliografía: LaValle, Steve, Michael Hopkins.  
“The new path to value: How the  
smartest organizations embedding analytics to  
transform insights.”



## FRASE

---

“Los datos son el activo estratégico de las compañías por excelencia”

**Christian Gardiner**

“Los datos son el petróleo del siglo XXI. El despliegue de sensores y el incremento de la capacidad del procesamiento, son claves en la transformación de muchos sectores y en la creación de un mundo más medible y programable”

**Cesar Alierta**







**INICIO RECESO**

<https://github.com/rb-one/curso-practico-machine-learning/blob/master/Notes/Notes.md>



**FIN DE RECESO**

<https://github.com/rb-one/curso-practico-machine-learning/blob/master/Notes/Notes.md>



FUNDACIÓN DE EDUCACIÓN SUPERIOR  
**SAN JOSÉ**  
INSTITUCIÓN TECNOLÓGICA

<https://github.com/rb-one/curso-practico-machine-learning/blob/master/Notes/Notes.md>

FIN DE  
GRABACIÓN