



Big Data  
Machine learning  
Business intelligence



## Big Data, Machine learning y Business intelligence

La revolución del dato ha venido a quedarse y las empresas tienen en su mano herramientas para poder ganarle el pulso a la demanda y ofrecer su mejor versión.

**Las herramientas de gestión y procesamiento de información están cambiando el mundo empresarial** y gracias a este documento podrás comprender las ventajas del conocimiento del dato. Hay oportunidades para procesar toda esa información que permitirán a cualquier negocio **mejorar la toma de decisiones, lograr la optimización de los activos, reducir gastos, personalizar los servicios o simplemente conocer la viabilidad de nuevas líneas de negocio.**

Cuando se tiene gran cantidad de datos surge la oportunidad de extraer conocimiento nuevo de ellos. No obstante, el volumen es tal que detectar patrones y sacar conclusiones escapa a las habilidades normales de un ser humano.



## Contexto

De siempre, cualquier empresa ha tenido necesidad de recabar datos, procesarlos y, con ellos, definir la estrategia para su negocio. Esto es lo que se llama análisis del dato o Data Analytics.

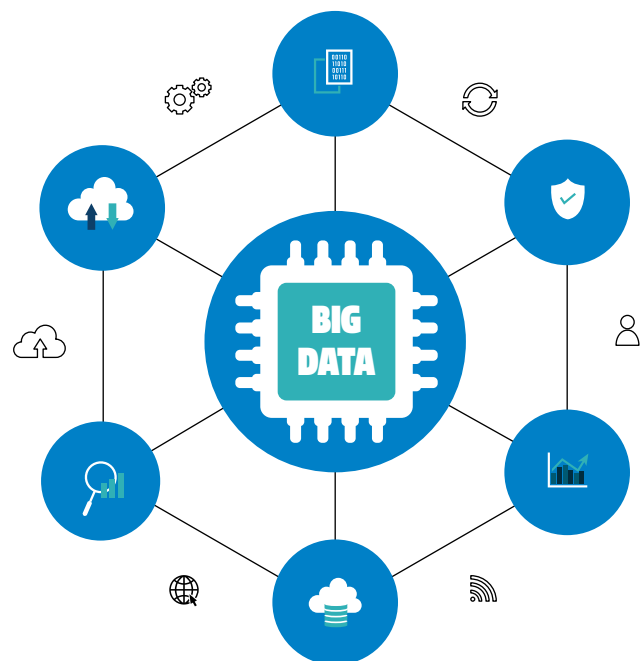
Hoy en día, **existe una revolución del dato gracias a la transformación digital. El volumen de datos es infinitamente más grande y su procesamiento más complejo.** Es en este panorama cuando surgen herramientas que permiten explotar esos volúmenes de información en poco tiempo, permitiendo a las organizaciones tomar decisiones basadas en información veraz sobre la realidad del mercado y/o de la empresa.

Todo esto hace que las empresas inviertan en estructuras y herramientas que propicien pasar de información sin procesar a datos y números procesables, herramientas que agilizan el día a día y optimizan el proceso de toma de decisiones.

Machine Learning, Big Data, Business Inteligente e Inteligencia Artificial son algunos de los conceptos que a día de hoy se emplean para llevar a cabo todo este análisis del dato, pero ¿sabemos cuál es la relación entre ellos? Por simplificar, poco se podría decir que todo se

resume en 2 pasos, siendo **el 1º o inicial la recopilación de datos (Data Analytics) para, a continuación, en 2º lugar, sacarle partido esos datos mediante el uso de procesos: Business Intelligence y Machine Learning.**

Son tres conceptos que buscan **exclusivamente mejorar el negocio, predecir patrones y encontrar datos que evidencien los argumentos para tomar determinadas decisiones.**



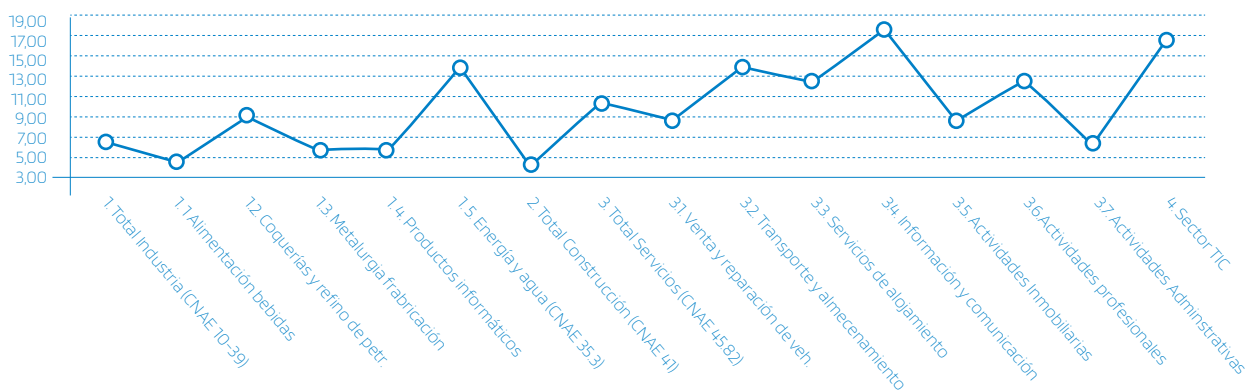
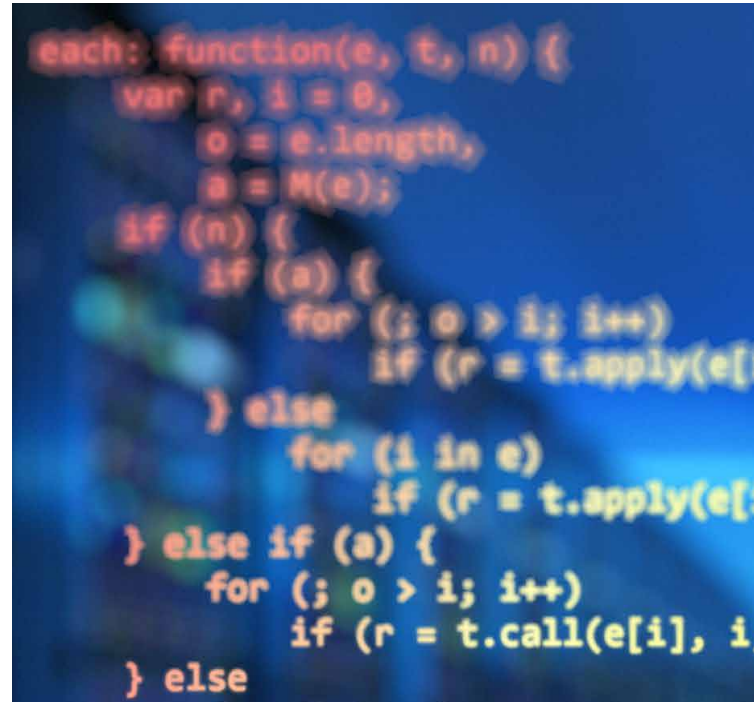


# Situación actual del uso de Big Data

Es muy distinto el uso que hacen las empresas en función del **sector** en el que se desarrolla su actividad económica.

El INE en su última encuesta sobre el Uso de Tecnologías de la Información y las Comunicaciones y del Comercio Electrónico en las Empresas **en lo referente al uso de Big Data**, destaca que los sectores que más uso hicieron son:

- Información y comunicaciones
- Transporte y almacenamiento
- Energía y agua
- Servicios de alojamiento



Encuesta sobre el Uso de Tecnologías de la Información y las Comunicaciones y del Comercio Electrónico en las Empresas, por agrupación de actividad (excepto CNAE 56 servicios de comidas y bebidas, 64-66 servicios financieros y auxiliares y servicios de seguros y 95.1, reparación ordenadores y equipos), y tamaño de la empresa

<https://www.ine.es/jaxi/Datos.htm?tpx=37753#!tabs-grafico>

# Situación actual del uso de Big Data

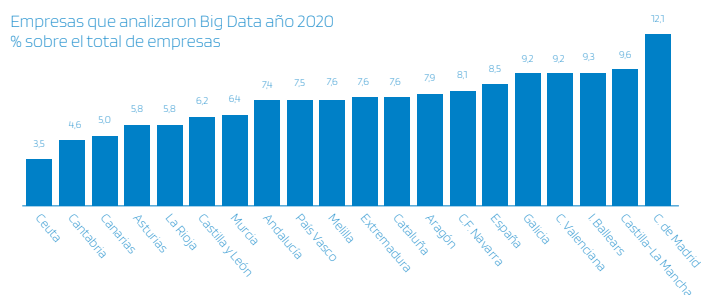
Por otro lado, por **comunidad autónoma, igualmente en lo referente al uso de Big Data**, solo Madrid supera los 2 dígitos en el porcentaje de uso de toda España

en el análisis del primer trimestre de 2020 llevado a cabo por el INE.

Primer trimestre de 2020	Internet y web <sup>(1)</sup>	Firma digital <sup>(1)</sup>	Medios Sociales <sup>(1)</sup>	Servicios nube <sup>(1)</sup>	Big Data	IoT	Impresión 3D	Robots
TOTAL	181	84.3	63	28.2	8.5	16.8	5	8.9
Andalucía	71.5	82.3	54.3	19.6	7.4	14.8	3.1	8.5
Aragón	85.4	83.3	64.7	22.5	7.9	15.9	5.2	11.3
Asturias, Pdo	81.6	85.4	61.3	24.4	5.8	14.4	6.3	7.8
Balears, Illes	81.4	76.1	73	23.3	9.3	1.4	1.9	5
Canarias	67.8	82	62.8	25.3	5	12.1	1.4	3.3
Cantabria	72.7	77	65.9	20	4.6	19.7	5.1	10.2
Castilla y León	73.2	85.2	61.3	23.1	6.2	18.3	4.7	8.8
Castilla-La Mancha	68.2	85.8	58.1	19.5	9.6	17.9	3.3	11.5
Cataluña	85.7	83.7	66.4	35.1	7.6	18.4	6.8	8.6
Comunitat Valenciana	77.4	86.5	65.2	28	9.2	17.3	4.9	10.1
Extremadura	68.8	83.3	59.5	15.4	7.6	21.6	4	10.1
Galicia	80.9	84.8	60.9	21	9.2	16.7	4.5	9.1
Madrid, Comunidad	79.1	86.9	67.8	38.5	12.1	17.4	6	8.4
Murcia, Región	70	86.1	56.3	16	6.4	16.2	3.7	10.5
Navarra, Comunidad Foral	82.2	85.2	62.0	28.7	8.1	23.3	6.6	15
País Vasco	81	81.9	58.3	31	7.5	15.7	5.6	10.4
Rioja, La	74.6	88.2	63.1	17.1	5.8	10	4	9.8
Ceuta	76.2	99.2	58.3	31	3.5	15	6.5	4.2
Melilla	56.2	86.2	33.7	7.1	7.6	12.9	0.7	2.2

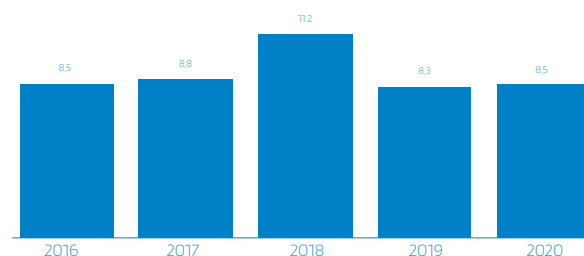
## (1) Porcentaje sobre el total de empresas con conexión a Internet

Empresas que analizaron Big Data año 2020 % sobre el total de empresas



Fuente: Instituto Nacional de Estadística (INE)  
[https://www.ine.es/prensa/tic\\_e\\_2019\\_2020.pdf](https://www.ine.es/prensa/tic_e_2019_2020.pdf)

INE en un informe, en 2018 disminuyendo en 2019 y 2020 a cifras similares a 2016 :



<https://www.ontsi.es/es/indicadores/Empresas/Equipamiento-TI-C/Empresas-que-utilizaron-Big-Data>

Y por último, destaca como dato relevante, el **uso del Big Data** en España tuvo su auge, tal y como indica el

# Conceptos

## Big Data

Es el tradicional Data Analytics pero con **volúmenes inmensos de datos en donde se depuran, transforman y modelan para poder extraer información útil** y lograr, con ello, agilizar las tomas de decisiones dentro de un negocio.

Lo más característico de BIG DATA son las 5 V's:

- **Volumen de datos**
- **Velocidad de procesamiento**
- **Variedad de fuente de datos a recabar**
- **Veracidad, ya que los datos deben permanecer limpios y ser correctos para dar frutos**
- **Valor de la información que se extrae, un beneficio para el negocio**

Para entender el Big Data es importante comprender **que todos estos datos se almacenan de manera regular en plataformas cloud, más concretamente en un data Lake** o lago de datos mediante cargas periódicas (también llamadas batch load) o en tiempo real (streaming).

En resumen, una amalgama de datos que permite tomar decisiones en base a información constantemente actualizada. **Esa depuración de datos y su forma de tratarlos se hace a través de procesos complejos como lo son Business Intelligence y Machine Learning.**



# Conceptos

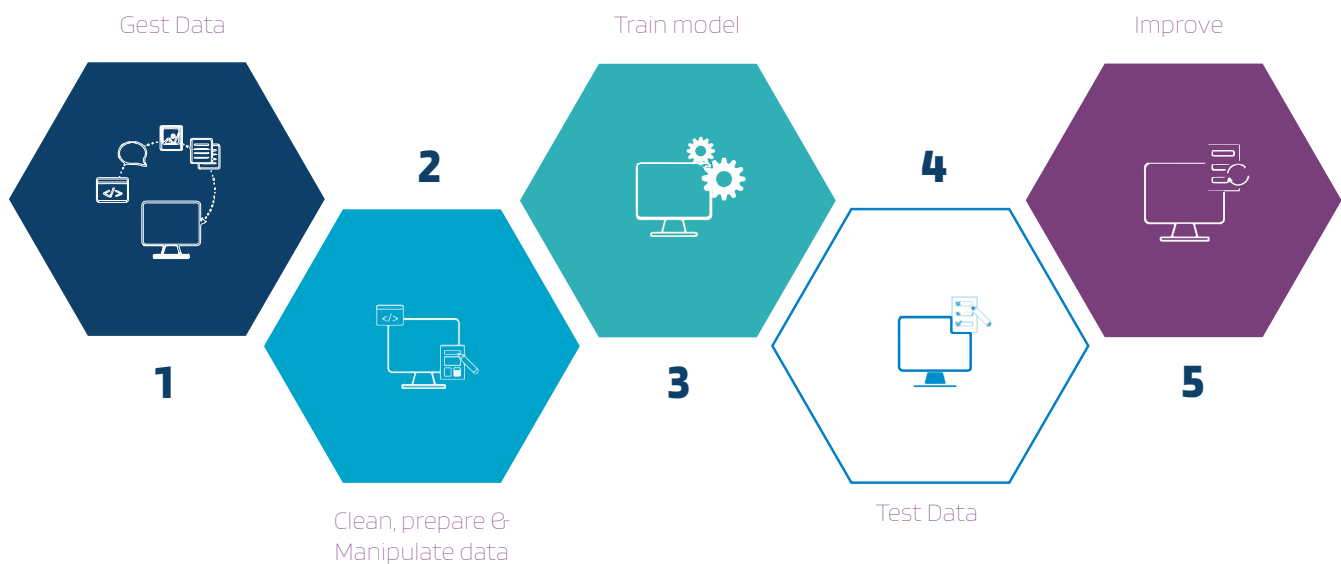
## Machine learning (ML)

Es una disciplina de la Inteligencia Artificial. Una vez que los datos han sido procesados y que el Big data ha cumplido su función, surge el proceso de **aprendizaje automático** para lograr una visión más profunda, rápida y sobre todo completa de los datos (y más amplia). Entre millones de datos **se identifican patrones complejos**. Se predicen comportamientos para que, a futuro, **la experiencia del análisis permita establecer un procedimiento interno automático**, sin necesidad de que intervenga persona alguna. De hecho, cuanto más dato haya, mayor capacidad de automatizar. Cabe

destacar que esto ocurre, de forma significativa, en negocios omnicanal, mejorando automáticamente con el tiempo, al disponer de gran variedad y volumen de datos.

ML se lleva a **cabo a través de algoritmos, diseñados para identificar patrones, predecir comportamientos y agilizar la toma de decisiones** en base a características útiles y predeterminadas. Esos algoritmos permiten adelantarse a la realidad y anticipar decisiones dentro de un negocio, definiendo así la estrategia futura de la organización.

## Machine learning



# Conceptos



## Business Intelligence (BI)

Otro proceso a tener en cuenta, una vez se ha llevado a cabo el almacenamiento de datos en el data Lake, es el llamado Business Intelligence. BI se encarga de **transformar el repositorio de data en información valiosa y práctica**, generando informes, gráficos, tablas, mapas y, por supuesto, dashboards para dar conocer el estado de un negocio de solo vistazo. Es decir, **transformar la información en conocimiento**. Los beneficios de cara a la toma de decisiones son inmensos ya que incluye todas las variables que influyen en el negocio, evitando la omisión de áreas importantes de la empresa que

pueden ponderar cualquier decisión, sin olvidar que **el análisis se hace en tiempo real y de forma automática**. De hecho, ya en 2020, en Orlando, durante el MicroStrategy World, se puso en evidencia la importancia de que todas las áreas y todos los empleados pudieran tener acceso a la información y, de esta manera, lograr un crecimiento estratégico en la empresa. **La realidad es que esto solo es posible si se estandariza la carga de datos, se organizan las fuentes y se tiene un lugar riguroso en que almacenarlos.**



# Conceptos

## Cuales son los beneficios del BI

El dato es actual y sobre todo no se duplica. Se controla en todo momento su practicidad y se asegura la [buena calidad del dato](#)

Autoconsumo de la información sin hacer uso de herramientas manuales sino de capacidades prescriptivas y predictivas [al alcance de todos](#).

Se logran [estandarizar los parámetros](#) de análisis para

que haya una coherencia en la predicción de tomas de decisiones y sirva en todas las áreas de la empresa.

Los datos son fiables [para todos](#).

Puesto que el primer paso es Data Analytics, veamos las diferencias entre Business Intelligence y Machine Learning

## Business Intelligence

Fórmulas matemáticas para análisis.

Permite definir la estrategia del negocio en base a datos y a resaltar oportunidades de negocio.

Información en bruto pasa a ser información ÚTIL.

Fundamental contar con un analista (su habilidad).

## Machine Learning

Software que va aprendiendo solo, a base de analizar datos.

Toma decisiones en base a programación, gracias a sistemas inteligentes.

Predice comportamientos o modelos a través del software: data mining o minería de datos.

El software y los algoritmos son la clave.

## Analítica de datos

**Descriptiva:** o ¿cómo ha funcionado nuestro negocio?

Consiste en almacenar datos históricos para poder visualizarlos y así entender mejor la situación actual del negocio.

**Predictiva:** o ¿cómo va a funcionar nuestro negocio?

Es analítica descriptiva unida a estadística con el fin de estimar datos que no conocemos del negocio o que son inciertos. Es muy útil en segmentación de clientes o en tema de valores a futuro.

**Prescriptiva:** o ¿cómo optimizamos el negocio (las estrategias) y gestión del mismo?

El resultado es, por ejemplo, recomendaciones automáticas sobre momentos idóneos para ejecutar acciones (temas de mantenimiento) ... etc.

**Diagnóstico:** analizar los resultados, pero teniendo en consideración su evolución ante determinadas acciones, para poder valorar posibles ajustes en las estrategias que se estén ejecutando.

# Conceptos



## Inteligencia artificial (AI)

### Simular la inteligencia humana en máquinas...

La inteligencia artificial se centra **en algoritmos**, buscando, con ello, explorar un territorio predictivo, extrayendo patrones de datos existentes y proyectándolos a futuro. Permite analizar tanto el pasado como el presente. **En base a la situación real, se puede predecir qué va**

**a pasar en el futuro**, algo más que atractivo para los negocios ya que supone adelantarse a las necesidades del mercado, ser proactivos en vez de reaccionar ante problemas y, sobre todo, evitar situaciones incómodas o de crisis

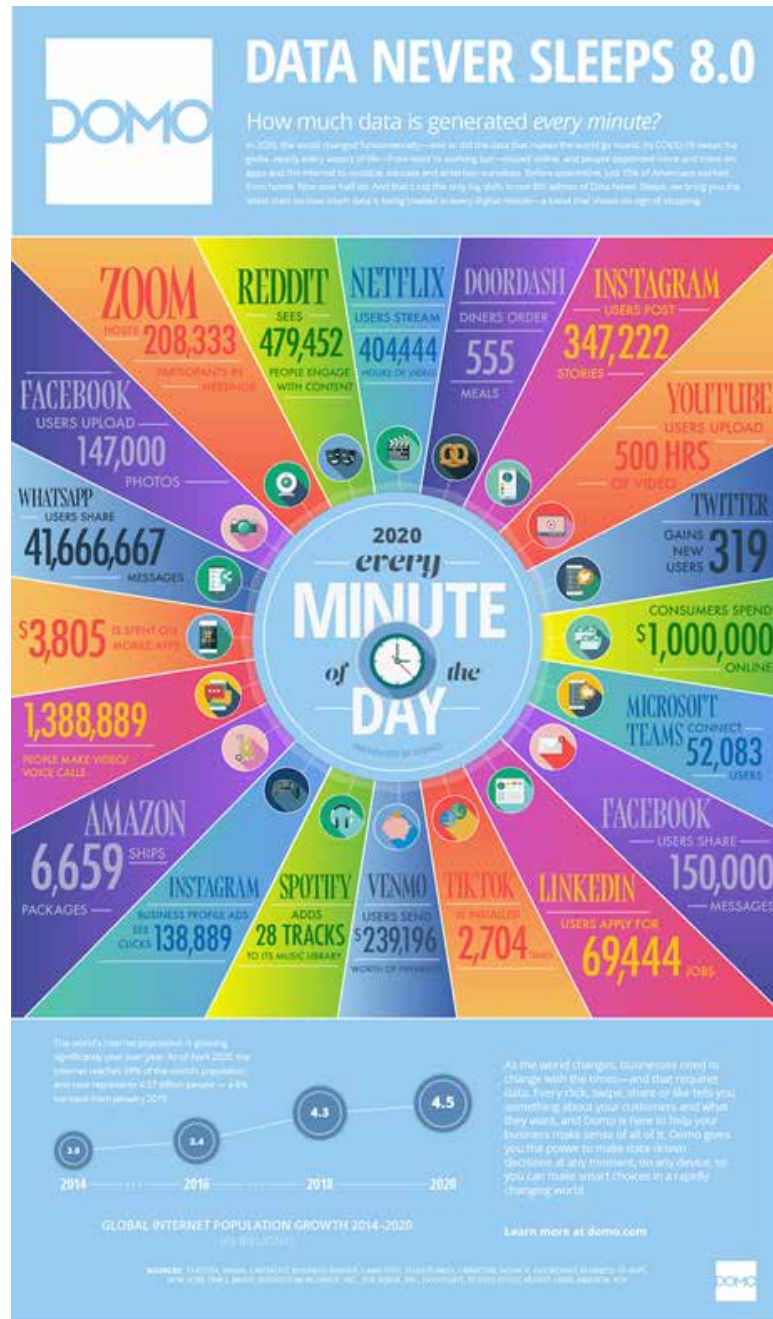
# Relevancia del dato

Partiendo de la premisa de que “todo puede medirse”, cualquier empresa de cualquier sector puede necesitar incorporar el **análisis de datos y de negocio a su estrategia de toma de decisiones con el fin de crecer (pese a la competencia feroz) y de adaptarse a un contexto cambiante (más acusado incluso tras la pandemia)**. Como estableció la revista WIRED: “Ya no basta con conocer los ingresos, los beneficios o los costes. Saber cuáles son los KPI que influyen en ellos y en el resto de objetivos empresariales es mucho más importante, porque se aprende a mejorar el negocio”.

Data never sleep 7.0 (los datos nunca duermen) pone de manifiesto la importancia de los datos, conocer el origen de la información que se usa en estas técnicas de análisis. Debemos tener en cuenta que cada día utilizamos el móvil desde donde se manda una cantidad ingente de información; pero además generamos información cada vez que hacemos clic en una web, pagamos con tarjeta, publicamos imágenes en rrs, hacemos uso del GPS, etc. Todas estas (y muchas otras) acciones generan datos masivos, susceptibles de ser tratados. Los tipos de datos pueden ser internos o externos y, a su vez, se diferencian o agrupan de la siguiente forma:

**Estructurados:** normalmente salen de bases de datos y son de fácil acceso, como puede ser el caso de CRM ("Customer Relationship Management" o Gestión de la Relación con el Cliente) o ERP ("Enterprise Resource Planning" o Planificación de Recursos Empresariales).

**Semi estructurados:** también de fácil acceso, pero requieren preparación ya que pueden provenir de plataformas o fuentes tales como redes sociales, dispositivos móviles, etc.



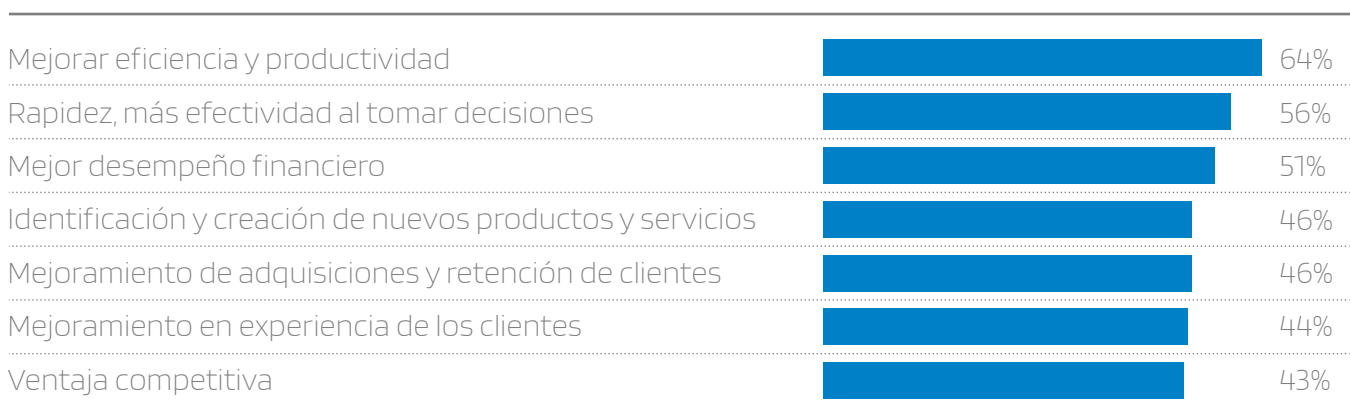
**No estructurados:** sin formato específico, por lo que son más complicados de analizar al provenir de conversaciones telefónicas, de imágenes, de e-mails o incluso de videos.

## Relevancia de los datos

En lo que se refiere a los negocios, las empresas cuentan con diversos sistemas informáticos que generan importantes cantidades de información, día a día, con distintas estructuras y formatos.

Para poder tratarlos, es fundamental establecer un proceso ETL (Extract, Transform, Load) que permite la extracción de los datos de las diferentes y múltiples fuentes, formatearlos, limpiarlos y cargarlos en un datawarehouse (repositorio de información).

## Qué beneficios aporta el estudio del dato



<https://www.larepublica.co/internet-economy/un-buen-uso-de-la-big-data-mejora-la-productividad-en-las-companias-de-todo-tipo-2962180>

# Sectores de uso del análisis y procesamiento de datos (Ejemplos)

**Marketing:** segmentación de clientes y poder conocer mejor, permitiendo, así, ponerlos en el centro de la estrategia. También la gestión de datos masivos de cara a adaptar sus productos y servicios a las necesidades de sus clientes, poder personalizar productos, optimizar operaciones e infraestructuras, y la posibilidad de explorar nuevas oportunidades de negocio.

**Deportes:** optimización del rendimiento. Dispositivos como los smart watches ayudan a tener información cualificada que incrementan la actividad física, así como su mantenimiento.

**Salud pública:** codificación del mapa genético. Muy útil a la hora de comprender enfermedades, prevenir posibles distorsiones y sobre todo encontrar tratamientos más eficientes. Los datos y su tratamiento también sirven para prescripción y prevención temprana de enfermedades, diagnosticar enfermedades e incrementar el éxito y la salud de los pacientes. Se logra con la analítica y procesamiento la medicina personalizada (diagnósticos adaptados a la casuística de cada persona).

**Automoción:** todo lo relacionado con Big Data sirve para la detección, mediante imágenes, de piezas defectuosas en la cadena de montaje y evitar su uso. Información para incrementar la rentabilidad de los vehículos, prolongar o afinar el mantenimiento, incrementando en todo momento su seguridad. Otros de los usos son adecuar la oferta a la demanda, para hacer más rentable la industria.

**Agricultura:** el análisis del dato aumenta el ROI de cualquier negocio del sector ya que puede ser útil en situaciones muy amplias, como, por ejemplo, sistemas de riego, análisis de rendimientos de cultivos y predicción de resultados, tractores inteligentes, datos sobre

animales y su salud, control de temperatura y climatología en invernaderos, control de plagas. Gestión ganadera, de cultivos, forestal, etc.

En cualquier negocio centrado en el consumo: chatbots para incrementar la satisfacción y la adecuación de los procesos de compra digitales. Robots en fábricas para reposición, que adecuen la cadena de valor a la realidad del mercado; incluso para predecir la demanda.

**Financiero:** el conocimiento de datos es útil para una infinidad de acciones del sector como evitar fraude, analizar el comportamiento de los clientes, agilizar procesos internos (calificación crediticia), suscripción de seguros optimizada y más ajustada a la realidad. Simulación de gestión de carteras para asegurar resultados en base a variables de datos (Robo Advisors). Todo tipo de automatizaciones que evitan el trabajo manual costoso y largo.

**Industria manufacturera:** análisis históricos para evitar fallos futuros. Optimizar la producción en base al acumulado histórico, estrategias de marketing para optimizar precios, lograr mayor rentabilidad al inventario, almacenamiento, stock... Posibilidad de automatizar pedidos a la carta para tener la customización como una opción real y rápida. También para la predicción vida útil de fábrica/maquinaria. En definitiva, todo lo necesario para mejorar la experiencia del usuario y fidelizar está en los datos. fábrica/maquinaria. En definitiva, todo lo necesario para mejorar la experiencia del usuario y fidelizar está en los datos.

**E-Commerce:** evitar y prevenir actividades fraudulentas, recopilar datos para mejorar atención, pero también para adaptar los productos y servicios a la necesidad real de la audiencia. Llevar a cabo estrategias de marketing omnicanal (un único usuario, aunque cambie de canal o de dispositivo).



**Energético:** mantener el equilibrio entre oferta y demanda, prevención de roturas o de mantenimiento, tras analizar los fallos funcionales de las redes eléctricas. Casas inteligentes que hacen uso de la información y datos para adaptar los consumos a las variaciones de precio.

o fracaso escolar, evaluación del alumno con amplia información para orientación de carreras.

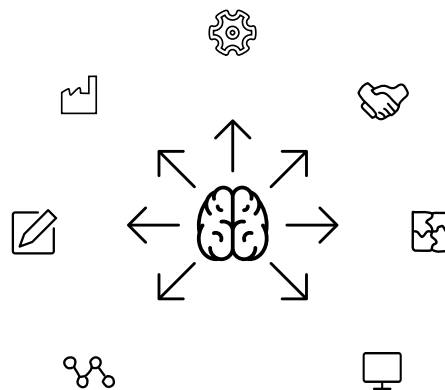
**Turismo:** el dato es el rey ya que la atención mejora y la oferta se adapta a las necesidades. Además, se pueden crear chatbots para

### Supply Chain

- Stock level predictions
- Demand forecasting
- Returns forecasting

### Predictive Maintenance

- Machine failure
- Preemptive maintenance
- Fault detection



### IT Optimization

- Security Cyberlake
- Master Data Management

### Freight and Shipping

- Dispatch failures
- Arrival delays
- Damaged packages
- Returns

**Administración pública:** chat y chatbot para agilizar procesos con información reglados, fraude fiscal identificable más rápidamente, seguridad ciudadana, movilidad y tráfico. Incluso para inspecciones de urbanismo (incendios).

**Educación:** sector donde la información es clave para priorizar al usuario o alumno y lograr así un aprendizaje personalizado, aumentando la capacidad de cada uno y fomentando las aptitudes. Gestión administrativa eficiente y ágil, formación orientada a vencer debilidades, análisis de rendimiento en tiempo real. Datos para poder reaccionar ante posible absentismo

atención al cliente. Ofrecer lo mejor al viajero en base al histórico de conocimiento y a sus preferencias: lograr una experiencia superior. Análisis de datos para ofrecer el mejor momento para un viaje o el mejor viaje para un momento/presupuesto. Sistemas de precio inteligente que casen mejor oferta - demanda - oportunidad. Proactividad ante índices de cancelación. Predicción en ocupación, en públicos objetivos.

**Retail:** predicción de gustos y de posibles compras, en base a comportamiento en red; flexibilidad de precios en base a ubicación, la actitud, la competencia...

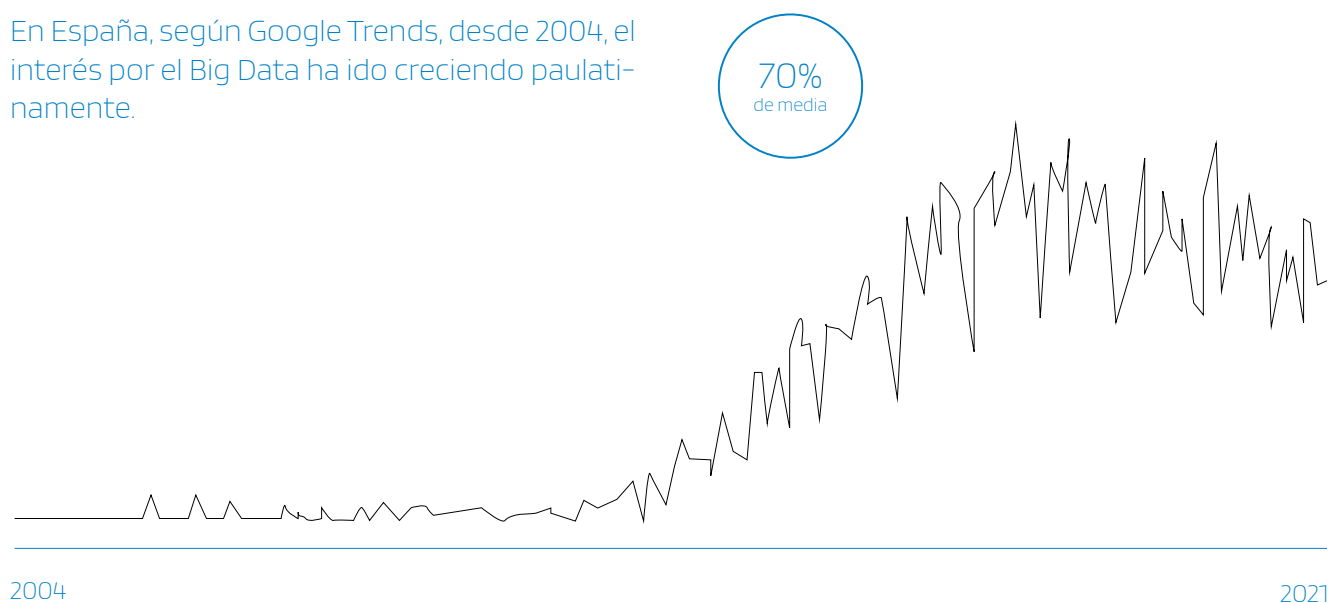
# El conocimiento del dato tras la pandemia

## Interés por el Big Data

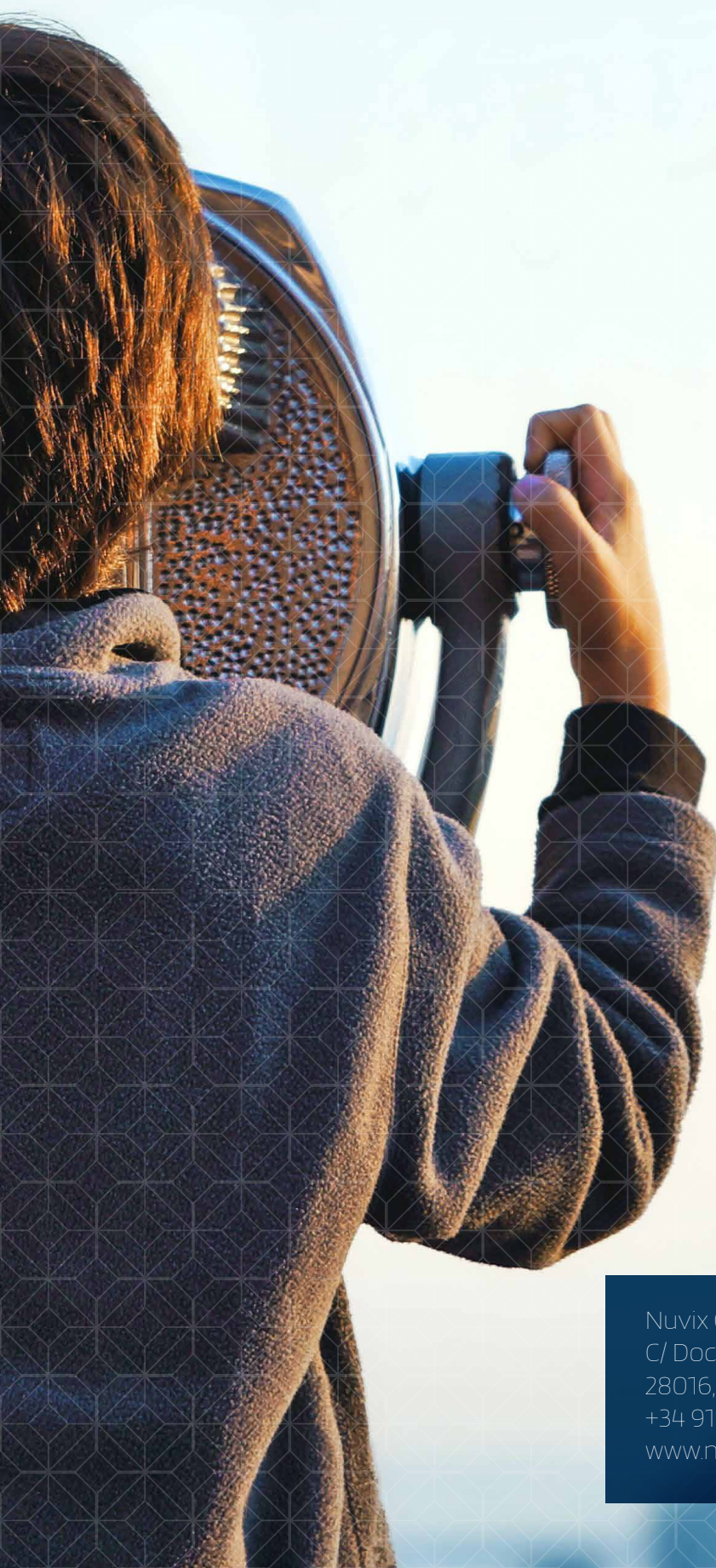
**Sin duda, el 2021 va a ser el año del conocimiento del dato y de la gestión del mismo, un año en que, tras la pandemia, las empresas tendrán entre sus pilares básicos los datos, la información.** La pandemia ha generado oportunidades, o incluso obligaciones, de digitalización para las empresas o microempresas y, ese engranaje digital que tanto dato arroja, ha generado posibilidades para las empresas y trabajadores que antes eran impensables. Las tendencias de BIG DATA son el futuro y prueba de ello ha sido el uso de la

misma en la lucha contra el coronavirus, permitiendo mejorar los resultados y la eficacia de los modelos epidemiológicos (app de rastreo, investigación, comportamiento del virus...). Si en algo tan crucial como el sistema sanitario ha sido útil, su aplicación al resto de sectores y su necesidad es infinita. Tal y como muestra este gráfico, el interés por el Big Data y el uso de los datos, como conocimiento clave en una empresa, es infinitamente mayor a partir de la pandemia.

En España, según Google Trends, desde 2004, el interés por el Big Data ha ido creciendo paulatinamente.



<https://bigdatamagazine.es/2021-el-ano-del-big-data>



Nuvix Consulting SL  
C/ Doctor Arce 14,  
28016, Madrid  
+34 91 769 86 48  
[www.nuvix.es](http://www.nuvix.es)

**Big Data**  
**Machine learning**  
**Business intelligence**