



## Objetivos:

- ◆ Definir el concepto de Big Data
- ◆ Comprender la importancia del uso de Big Data en la vida diaria
- ◆ Describir las ventajas del uso de las Big Datas
- ◆ Explicar los diferentes tipos de Big Datas



## Introducción:

En la era actual, se están desarrollando numerosas aplicaciones sociales que resultan en un aumento masivo de datos todos los días y cuando hablamos de plataformas de redes sociales, millones de usuarios se conectan a diario, la información se comparte cada vez que los usuarios utilizan una plataforma de redes sociales o cualquier otro sitio web, por lo que surge la pregunta de cómo se maneja esta enorme cantidad de datos y a través de qué medio o herramientas se procesan y almacenan los datos. Aquí es donde el Big Data sale a la luz.

Big Data se refiere a cantidades masivas de datos producidos por diferentes fuentes como plataformas de redes sociales, registros web, sensores, dispositivos IoT y muchos más. Puede ser estructurado (como tablas en DBMS), semiestructurado (como archivos XML) o no estructurado (como audios, videos, imágenes).

Los sistemas tradicionales de gestión de bases de datos no son capaces de manejar esta gran cantidad de datos. Big Data ayuda a las empresas a generar información valiosa. Las empresas utilizan big data para refinar sus campañas y técnicas de marketing. Las empresas lo utilizan en proyectos de aprendizaje automático para entrenar máquinas, modelos predictivos y otras aplicaciones de análisis avanzado. No podemos equiparar big data a ningún volumen de datos específico. Las implementaciones de Big Data pueden involucrar terabytes, petabytes e incluso exabytes de datos capturados a lo largo del tiempo.



## ¿Qué es “Big Data”

Big data se refiere a los grandes y diversos conjuntos de información que crecen a tasas cada vez mayores. Abarca el volumen de información, la velocidad o velocidad a la que se crea y recopila, y la variedad o el alcance de los puntos de datos que se cubren (conocidos como las "tres v" de big data). Big Data a menudo proviene de la minería de datos y llega en múltiples formatos.

Como lo define Gartner: "Los Big Data son activos de información de alto volumen, alta velocidad o variedad que requieren nuevas formas de procesamiento para permitir una mejor toma de decisiones, descubrimiento de información y optimización de procesos". Profundicemos y entendamos esto en términos más simples.

## ¿Qué Significa “Big Data”

El término "big data" se explica por sí mismo, una colección de enormes conjuntos de datos que las técnicas informáticas normales no pueden procesar. El término no solo se refiere a los datos, sino también a los diversos marcos, herramientas y técnicas involucradas. El avance tecnológico y el advenimiento de nuevos canales de comunicación (como las redes sociales) y dispositivos nuevos y más fuertes han presentado un desafío para los actores de la industria en el sentido de que tienen que encontrar otras formas de manejar los datos.



## ***Importancia de las “Big Data”***

Lo que hace que Big Data sea tan útil para muchas empresas es el hecho de que proporciona respuestas a muchas preguntas que las empresas ni siquiera sabían que tenían. En otras palabras, proporciona un punto de referencia. Con una cantidad tan grande de información, los datos pueden ser moldeados o probados de cualquier manera que la empresa considere adecuada. Al hacerlo, las organizaciones son capaces de identificar los problemas de una forma más comprensible.

La recopilación de grandes cantidades de datos y la búsqueda de tendencias dentro de los datos permiten que las empresas se muevan mucho más rápidamente, sin problemas y de manera eficiente. También les permite eliminar las áreas problemáticas antes de que los problemas acaben con sus beneficios o su reputación.

El análisis de Big Data ayuda a las organizaciones a aprovechar sus datos y utilizarlos para identificar nuevas oportunidades. Eso, a su vez, conduce a movimientos de negocios más inteligentes, operaciones más eficientes, mayores ganancias y clientes más felices.



## Tipos de “Big Data”

Hay tres tipos de “Big Data” y son:

- ◆ **Estructuradas:** Cualquier dato que pueda almacenarse, accederse y procesarse en forma de formato fijo se denomina datos "estructurados". A lo largo del tiempo, el talento en informática ha logrado un mayor éxito en el desarrollo de técnicas para trabajar con este tipo de datos (donde el formato es bien conocido de antemano) y también en la obtención de valor de ellos. Sin embargo, hoy en día, estamos previendo problemas cuando un tamaño de tales datos crece en gran medida, los tamaños típicos están siendo furiosos por múltiples zettabytes.
- ◆ **Semi – estructuradas:** Los datos semiestructurados pueden contener ambas formas de datos. Podemos ver los datos semiestructurados como una forma estructurada, pero en realidad no se definen con, por ejemplo, una definición de tabla en DBMS relacional. Ejemplo de datos semiestructurados es un dato representado en un archivo XML
- ◆ **Desestructurados:** Cualquier dato con forma desconocida o la estructura se clasifica como datos no estructurados. Además de que el tamaño es enorme, los datos no estructurados plantean múltiples desafíos en términos de su procesamiento para obtener valor de ellos. Un ejemplo típico de datos no estructurados es una fuente de datos heterogénea que contiene una combinación de archivos de texto simples, imágenes, videos, etc. Hoy en día, las organizaciones tienen una gran cantidad de datos disponibles con ellos, pero desafortunadamente, no saben cómo obtener valor de ellos, ya que estos datos están en su forma cruda o en formato no estructurado.



## *Características de las “Big Datas”*

A continuación, las principales características:

- ◆ **Volumen:** El nombre Big Data en sí está relacionado con un tamaño que es enorme. El tamaño de los datos juega un papel muy crucial en la determinación del valor de los datos. Además, si un dato en particular puede considerarse realmente como un Big Data o no, depende del volumen de datos. Por lo tanto, el "volumen" es una característica que debe tenerse en cuenta al tratar con soluciones de Big Data.
- ◆ **Velocidad:** el término "velocidad" se refiere a la velocidad de generación de datos. La rapidez con la que se generan y procesan los datos para satisfacer las demandas determina el potencial real de los datos. Big Data Velocity se ocupa de la velocidad a la que los datos fluyen desde fuentes como procesos comerciales, registros de aplicaciones, redes y sitios de redes sociales, sensores, dispositivos móviles, etc. El flujo de datos es masivo y continuo.
- ◆ **Variedad:** La variedad se refiere a fuentes heterogéneas y a la naturaleza de los datos, tanto estructurados como no estructurados. Durante los primeros días, las hojas de cálculo y las bases de datos eran las únicas fuentes de datos consideradas por la mayoría de las aplicaciones. Hoy en día, los datos en forma de correos electrónicos, fotos, videos, dispositivos de monitoreo, PDF, audio, etc. también se están considerando en las aplicaciones de análisis. Esta variedad de datos no estructurados plantea ciertos problemas para el almacenamiento, la minería y el análisis de datos.



## *Beneficios de las “Big Datas”*

- ◆ **Mejorar la toma de decisiones:** cuando se estudia la información que proviene de los clientes y se tiene en cuenta la gestión de empleados, la toma de decisiones se puede llevar a cabo de forma mucho más acertada gracias al uso de la Big Data y la inteligencia de negocios.
- ◆ **Optimizar costes:** el análisis de la Big Data suele acelerar la forma en la que se desarrolla un producto. Por eso, cuando se acortan los procesos en una empresa, los costes que se tienen contemplados también disminuyen.
- ◆ **Orientar los productos hacia el cliente:** una vez se organizan todos los datos disponibles que ingresan a una empresa, es posible empezar a segmentar a los clientes y orientar los productos según las necesidades específicas de cada usuario
- ◆ **Garantizar mayor seguridad en los datos:** mantener un control de los datos de forma organizada permite detectar a tiempo posibles complicaciones o amenazas y además facilita encontrar información sensible que no se encuentre protegida adecuadamente.
- ◆ **Mejorar la accesibilidad a la información:** cuando en una empresa los datos se encuentran digitalizados y ordenados, acceder a la información relevante es mucho más sencillo. Por eso, la implementación del Big Data es una forma excelente de optimizar procesos orientados a la inteligencia de negocios.
- ◆ **Obtener ventajas competitivas:** debido a que el acceso a este tipo de datos permite actualizar, ajustar y aplicar herramientas de la analítica predictiva a un producto en tiempo real, es posible establecer una mejor posición en el mercado respecto a la competencia.



## *Ejemplos de Uso “Big Datas”*

- ◆ Experiencias de compra de comercio electrónico personalizadas
- ◆ Modelización de los mercados financieros
- ◆ Recopilación de billones de puntos de datos para acelerar la investigación del cáncer
- ◆ Recomendaciones de medios de servicios de transmisión como Spotify, Hulu y Netflix
- ◆ Predicción de los rendimientos de los cultivos para los agricultores
- ◆ Big data ayuda a los equipos deportivos a maximizar su eficiencia y valor
- ◆ Reconocer las tendencias en los hábitos educativos de los estudiantes individuales, las escuelas y los distritos
- ◆ Finanzas: las industrias de finanzas y seguros utilizan big data y análisis predictivo para la detección de fraudes, evaluaciones de riesgos, clasificaciones de crédito, servicios de corretaje y tecnología blockchain, entre otros usos.



## *Referencias:*

- ◆ <https://www.guru99.com/what-is-big-data.html>
- ◆ [https://www.sas.com/en\\_us/insights/big-data/what-is-big-data.html](https://www.sas.com/en_us/insights/big-data/what-is-big-data.html)
- ◆ <https://www.oracle.com/big-data/what-is-big-data/#use-cases>
- ◆ <https://builtin.com/big-data>
- ◆ <https://www.simplilearn.com/tutorials/big-data-tutorial/what-is-big-data>