

Big Data

1. ¿Qué es el big data?
 - A. El procesamiento y gestión de un volumen de datos tan alto que no puede llegar a ser procesado por herramientas tecnológicas convencionales.
 - B. Un software avanzado que permite realizar estadísticas.
 - C. Análisis de grandes cantidades de datos que pueden reutilizarse con las herramientas de almacenamiento y procesamiento convencionales.
 - D. Proceso de desarrollar nuevos modelos de análisis para conseguir que los datos ofrezcan más información.

2. ¿Cómo se traduce al español el término Big Data?
 - A. Datos masivos.
 - B. Procesado de grandes volúmenes.
 - C. Grandes datos.
 - D. Datos grandes.

3. ¿Cuál de las siguientes características **NO** define a los conjuntos de Big Data?
 - A. Volumen.
 - B. Velocidad.
 - C. Vistosidad.
 - D. Variedad.

4. ¿Cuáles son las tres Vs principales del Big Data?
 - A. Valor, velocidad y variedad.
 - B. Volumen, velocidad y variedad.
 - C. Vistosidad, volumen y variedad.
 - D. Vistosidad, volumen y velocidad.

5. ¿Qué entendemos por velocidad en el Big Data?
 - A. La rapidez para almacenar los datos.
 - B. La cantidad de datos que pueden analizarse por segundo.
 - C. El tiempo que tardan las redes informáticas en actualizarse.
 - D. La velocidad no es una de las cinco Vs que definen al Big Data.

6. Para representar volúmenes de datos se utiliza el exabyte, que es igual a...
 - A. 10^9 bytes.
 - B. 10^{12} bytes.
 - C. 10^{15} bytes.
 - D. 10^{18} bytes.
7. Para medir la información que generan los miles de millones de usuarios de internet se utilizan unidades de medida de la información del orden de...
 - A. Kilobytes.
 - B. Megabytes.
 - C. Petabytes.
 - D. Gigabytes.
8. Los datos obtenidos de las bases de datos SQL serían datos...
 - A. Estructurados.
 - B. Semiestructurados.
 - C. No estructurados.
 - D. No son un tipo de datos.
9. Los datos obtenidos de las bases de datos No SQL y el código HTML serían datos...
 - A. Estructurados.
 - B. Semiestructurados.
 - C. No estructurados.
 - D. No son un tipo de datos.
10. Dentro del entorno del Big data, los datos provenientes de las redes sociales, suelen ser datos...
 - A. Estructurados o semiestructurados.
 - B. Sólo semiestructurados.
 - C. No estructurados o semiestructurados.
 - D. Dentro del entorno del Big Data no se utilizan datos provenientes de las redes sociales.
11. Los datos semiestructurados...
 - A. Son datos que tienen un formato definido claro.
 - B. Son datos en crudo que no tienen un formato específico.
 - C. Son datos que no se limitan a campos determinados, pues contienen marcadores para separar las diferentes secciones. Son datos que tienen su propio esquema (metadato).
 - D. Los datos semiestructurados no existen.

12. ¿Cuál dirías que es uno de los principales beneficios de utilizar un modelo Big Data?
- A. Permite un mayor control sobre las personas.
 - B. Únicamente la automatización de los procesos.
 - C. La capacidad para que los datos existentes queden almacenados y accesibles.
 - D. Pueden aprovecharse mejor las nuevas oportunidades.
13. ¿Cuál sería una desventaja del Big Data?
- A. Dificultad en el análisis y la gestión de datos.
 - B. Nuevas oportunidades de negocio.
 - C. Permite un mundo de oportunidades.
 - D. Ninguna de las anteriores es una desventaja del Big Data.
14. ¿Cuál de las siguientes opciones dirías que es uno de los principales usos del Big data?
- A. Encontrar correlaciones entre múltiples patrones vistos en los datos.
 - B. Analizar el perfil de un usuario para comprender sus hábitos al hacer uso de una aplicación o servicio.
 - C. Hacer más rápido el análisis de datos y toma de decisiones convencional.
 - D. Ninguna de las opciones anteriores se corresponde con usos del Big Data.
15. Los dispositivos que llevamos encima y que se encargan de recoger datos de nuestra actividad, reciben el nombre de...
- A. disponibles.
 - B. scrapers.
 - C. wescrapers.
 - D. weareables.
16. ¿Cuál de las siguientes afirmaciones es **FALSA**?
- A. La arquitectura Big Data incluye muchos tipos de fuentes de datos.
 - B. Las aplicaciones de Big data son relevantes sobre todo para industrias intensivas en conocimiento.
 - C. Las interacciones en las redes sociales son una forma de Big Data.
 - A. Big Data incluye cantidades aproximadamente iguales de datos estructurados y no estructurados.

17. ¿Cuál de las siguientes afirmaciones relacionadas con la huella digital es **CIERTA**?

- A. Nuestra huella digital está formada por datos que proporcionamos voluntariamente.
- B. Nuestra huella digital está formada por datos que se recopilan automáticamente.
- C. Nuestra huella digital está formada por datos que proporcionamos voluntariamente y datos que se recopilan automáticamente.
- D. Ninguna de las afirmaciones anteriores es cierta.

18. ¿Cuál de las siguientes afirmaciones relacionada con el Data Scraping es **FALSA**?

- A. El web scraping es la forma más común de scraping de datos, y consiste en la extracción de contenido de páginas web, que luego puede ser utilizado para distintos fines como análisis, seguimiento de precios, investigaciones de mercado, entre otros.
- B. El scraping de datos, data scraping o extracción de datos, es una técnica automatizada mediante la cual un programa informático extrae datos de un sitio web o una aplicación.
- C. La función básica del data scraping es la de convertir datos que son estructurados en semiestructurados para hacerlos fáciles de leer y entender.
- D. El web scraping o raspado web es una técnica que utiliza programas de software para extraer información de sitios web.

19. ¿De dónde se obtiene el Big Data?

- A. De las redes sociales.
- B. De los dispositivos móviles.
- C. De fuentes de datos abiertos.
- D. Todas las respuestas anteriores son correctas.

20. ¿Cuál de las siguientes **NO** es una buena práctica para la protección de la información personal?

- A. Desconocer las configuraciones de seguridad de las aplicaciones.
- B. Ser cauteloso con la información que se comparte en línea.
- C. Utilizar contraseñas seguras.
- D. Todas las respuestas anteriores son buenas prácticas para proteger nuestra información personal.