



EJERCICIOS 4 TEMA 1

1. ¿Qué es lo que se conoce como Big Data?

Son datos cuyo volumen es demasiado grande para procesarlos con las tecnologías y técnicas tradicionales requiriendo nuevas arquitecturas de hardware, modelos de programación y algoritmos para su procesamiento.

2. ¿Cuáles son las tres Vs que describen al Big Data?

Volumen es demasiado grande.

Variedad de estructuras y formatos.

Velocidad para poder procesarlos rápido y que sean útiles.

3. ¿Cuáles son las 8 Vs del Big Data?

Volumen

Variedad

Velocidad

Volatilidad referida al tiempo durante el cual los datos recogidos son válidos

Valor se refiere a la utilidad de los datos obtenidos para extraer conocimientos y tomar decisiones a partir de ellos

Validez referida a lo precisos que son los datos para el uso que se pretende darles.

Veracidad relacionada con la confiabilidad del origen del cual provienen los datos con los que se trabajará así como la incertidumbre o el ruido que pudiera existir en ellos.

Variabilidad frente a la variedad de estructuras y formatos, hace referencia a la complejidad del conjunto de datos, es decir, al número de variables que contiene.

4. ¿Qué modelo de procesamiento utilizan las soluciones enfocadas al Big Data y qué diferencia tienen con el modelo anterior?

Utilizan el modelo OLAP y la diferencia es que el antiguo (OLTP) es un sistema de procesamiento de transacciones en línea, el segundo es un sistema de recuperación y análisis de datos en línea. Por tanto, OLAP complementa a SQL aportando la capacidad de analizar datos desde distintas variables y dimensiones, mejorando el proceso de toma de decisiones.



5. ¿Qué es un almacén de datos? ¿Qué le debe proveer a una empresa?

Es conocido como data warehouse es una solución de business intelligence combina tecnología y componentes con el objetivo de ayudar uso estratégico de los datos por parte de una organización. Esta solución debe proveer a la empresa, de forma integrada, de capacidad de almacenamiento de una gran cantidad de datos así como de herramientas de análisis del os mismo que frente al procesamiento de transacciones permita transformar los datos en información para ponerla a disposición de la organización y optimizar el proceso de toma de decisiones.

6. ¿Qué son las bases de datos NoSQL?

Es un tipo de base de datos que permite almacenar y gestionar tipos de datos que tradicionalmente han sido difíciles de gestionar por parte de las bases de datos relaciones. Así pues, las bases de datos NoSQL hace referencia a bases de datos relacionales a bases de datos documentales, bases de datos orientadas a grados, buscadores, etc.