**EJERCICIOS 4 TEMA 1**

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

La gran cantidad de datos disponibles junto con las herramientas, tanto hardware como software, que existen a disposición para analizarlos.

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

Volumen. Es demasiado grande para procesarlo con las tecnologías y técnicas convencionales, por lo que se necesitan nuevas arquitecturas, hardware, modelos de programación y algoritmos para su procesamiento.

Variedad. De estructuras y formatos para poder procesarlos: datos sintéticos, extraídos de sensores, numéricos, textuales, imágenes, audio, vídeo, etc.

Velocidad. Se requiere una enorme velocidad para poder extraer el valor y análisis de estos datos.

1. ¿Cuáles son las 8 Vs del Big Data?

En adición a las tres anteriores, se incluyen:

Volatilidad. El tiempo durante el cual los datos recogidos son válidos.

Valor. La utilidad de los datos obtenidos para tomar decisiones en base a ellos.

Validez. Lo precisos que son los datos para el uso propuesto.

Veracidad. La confiabilidad de su origen y si existe alguna incertidumbre en ellos.

Variabilidad. Con respecto a las estructuras y formatos, hace referencia a la complejidad de estos datos, es decir, el número de variables que contiene.

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

El modelo Online Analytical Processing, OLAP. Se diferencia del Online Transaction Processing, OLTP, en que mientras el anterior procesaba transacciones en línea, este nuevo modelo es un sistema de recuperación y análisis de datos en línea. Esto hace que OLAP complemente a SQL aportando la capacidad de analizar datos desde distintas variables y dimensiones, mejorando el proceso de toma de decisiones.

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

Un almacén de datos o data warehouse es una solución de business intelligence que combina tecnologías y componentes con el fin de ayudar al uso estratégico de los datos por parte de una organización. Provee a la empresa de una gran capacidad de almacenamiento de los datos y de herramientas de análisis, lo cual permite a la empresa transformar los datos en información para optimizar el proceso de toma de decisiones.

1. ¿Qué son las bases de datos NoSQL?

Son alternativas a las bases de datos relacionales, que se usan para gestionar datos que no son óptimos para estas últimas. Ejemplos de esto son bases de datos documentales, orientadas a grafos, buscadores, etc. Se caracterizan principalmente porque son independientes de los esquemas, no son relacionales y están distribuidas, es decir, no centralizadas.