**¿Cuáles son las diferencias existentes entre una base de datos operacional y un Data Warehouse desde un punto de vista de la estructura de datos? Justifica tu respuesta.**

Las diferencias son:

* **Temporalidad:** en la base de datos operacionales el tiempo en que se guardan los datos son entre uno y dos años mientras que en un Data Warehouse se amplia de cinco a diez años.
* **Volumen:** en las bases de datos operacionales el volumen será relativamente pequeño respecto a los Data Warehouse.
* **Nivel de agregación:** mientras que en las bases de datos operacionales el nivel de agregación suele ser único y bastante bajo, en los Data Warehouse se suelen dar distintos niveles.
* **Actualización:** La actualización de los datos en una base de datos operacional se hace constantemente (de una manera atómica) puesto que la información es muy cambiante. En los Data Warehouse se hace de manera periódica (de un conjunto de registros) y con unos intervalos de tiempo definidos.
* **Estructura:** habitualmente las bases de datos operacionales tienen una estructura relacional ya que prima la estabilidad. En los Data Warehouse habrá una visión multidimensional y a la vez serán muy dinámicos.

**Enumera y describe brevemente las diferencias entre un Data Mart y un Enterprise Data Warehouse.**

Las diferencias de Data Warehouse y Data Marts son:

* **Alcance:** Mientras que un Data Warehouse es un sistema centralizado con datos globales de la empresa/organización y de todos sus procesos operacionales, un Data Mart es un subconjunto temático de datos, orientado a un proceso o un área de negocio específica.
* **Tema:** Un Data mart se centra en un solo tema, mientras que los datos de un Data Warehouse comprende datos de todos los departamentos de la organización donde se actualiza continuamente para eliminar datos redundantes.
* **Implementación:** El proceso de implementación un Data Warehouse puede extenderse muchos meses e incluso años. El proceso de implementación de Data Mart está restringido a unos pocos meses.
* **Nivel de detalle:** Los datos almacenados en un Data Warehouse siempre ofrecen más detalle en comparación con los de un Data Mart.

**¿Qué enfoque en la construcción de la FIC consideras más adecuado? Argumenta la respuesta.**

Realmente no existe un enfoque adecuado único, el enfoque dependerá de las circunstancias en las que se encuentre la organización que quiera construir la FIC.

Si la organización es reciente y no tiene ningún tratamiento de los datos en dicho momento, lo más idóneo es la **Construcción del almacén de datos corporativo a posteriori**, ya que la solución consiste en construir el almacén de datos corporativo antes o durante la construcción de los diferentes almacenes de datos departamentales, esto va a evitar la múltiple interpretación de los datos.

Si la organización ya existe y tiene ya almacenes de datos departamentales el mejor enfoque es la **Construcción de almacenes de datos departamentales**, puesto que es la manera más "natural" o intuitiva. Es un enfoque válido cuando se trata de construir almacenes de datos departamentales totalmente independientes, aunque tiene problemas ya que si los datos no son independientes los estaremos tratando más de una vez y además podemos tratarlos de manera distinta.

**¿Cuáles son las principales ventajas y desventajas de disponer de un cubo OLAP? Enumera y describe brevemente.**

**Ventajas**

* Tiene acceso a grandes cantidades de información.
* Analiza las relaciones entre muchos tipos de elementos empresariales.
* Involucra datos agregados.
* Compara datos agregados a través de periodos jerárquicos.
* Presentan los datos en diferentes perspectivas.
* Involucran cálculos complejos entre elementos de datos.
* Pueden responder con rapidez a consultas de usuarios.
* Robustez, consolidación y validación lo que aportan valor y calidad al sistema.

**Desventajas**

Su principal desventaja reside en la imposibilidad de realizar cambios en su estructura. Debido a su forma de funcionamiento y almacenamiento de la información, cuando los usuarios requieren realizar modificaciones en la estructura de este tipo de base de datos, deben rediseñar el Cubo OLAP, sin posibilidades de poder utilizar la estructura en la que se trabajó hasta el momento.

**Indica si las siguientes afirmaciones son verdaderas (V) o falsas (F). Justifica tu respuesta en aquellas afirmaciones que consideres falsas.**

* **El proceso de integración será el mismo independientemente de que el proceso de carga sea inicial o incremental. (V o F) FALSO**

El proceso de integración será diferente dependiendo de si hacemos la carga inicial del almacén de datos o una actualización de este. En el proceso de carga inicial se hará una integración de todos los datos previa a la carga en el almacén de datos. Cuando es incremental, es posible que no estén disponibles los datos de todas las fuentes al mismo tiempo e interese integrar los datos de las distintas fuentes por separado en el almacén de datos

* **El responsable de datos o Chief Data Officer (CDO) debe ser capaz de implementar algoritmos de procesamiento de datos y definir modelos predictivos. (V o F) VERDADERO**
* **Usando herramientas self-service BI varios analistas pueden obtener resultados diferentes a un mismo problema. (V o F) verdadero**
* **Los datos semiestructurados son datos con una estructura de datos conocida, y se almacenan principalmente en bases de datos relacionales. (V o F) FALSO**

Estos datos pueden tener una estructura interna, pero carecen de la estructura externa necesaria para los procesos estándares de administración de datos.

Los datos semiestructurados se refieren a cualquier información que utilice un esquema de autodescripción, como HTML, XML o JSON. Estos tipos de datos tienen un esquema abierto que permite la flexibilidad de datos de la aplicación.