Soluciones de Inteligencia de Negocios

Tema Nº2:Base de Datos OLTP – OLAP

Indicador de logro Nº2:Identifica fundamentos de las bases de datos OLTP y OLAP, diferencia entre un data Mining y Big Data haciendo uso de los conocimientos impartidos.

**TEMA 01 Teoría de los**

Imagen que contiene Icono

Descripción generada automáticamente

**TEMA Nº2:**

Base de Datos OLTP – OLAP

**Subtema 1.1:**

Conceptos Base de Datos OLTP – OLAP

**MARCO TEÓRICO**

**OLTP**

Se caracteriza por bases de datos transaccionales con una gran cantidad de transacciones cortas (INSERT, DELETE, UPDATE) on-line. Suelen ser BD de sistemas críticos como ERP, CRM, etc.

**OLAP**

Se caracteriza por ser un modelo con volumen de transacciones bajo. Hay **dos modelos**: **multidimensional** o **tabular** y están optimizadas para leer datos que normalmente implican agregaciones y donde el tiempo de respuesta de consulta de datos debe ser rápido (es una medida de efectividad). Normalmente hay información histórica agregada almacenada en esquemas multidimensionales (generalmente esquema en estrella).

* Almacén operacional de datos, también llamado ODS (del inglés Operational Data Store), es un contenedor de datos activos, es decir, está diseñado para integrar datos de múltiples fuentes con los que realizar operaciones adicionales en los propios datos. A diferencia de un almacén de datos maestro, los datos no vuelven de nuevo a los sistemas operacionales de origen, sino que se preparan para la realización de otras operaciones o para ser incorporados a un almacén de datos (o data warehouse) para ser explotados en la realización de informes. Por tanto, se puede decir que un sistema ODS estará ubicado entre uno o varios sistemas OLTP (origen) y un sistema OLAP (destino).
* Un sistema OLTP diría qué facturas no han sido liquidadas, el monto que falta por liquidar de cada una, cuando se espera que se paguen etc.
* Un sistema OLAP diría cuanto debe un cliente, cuantas facturas son, el promedio de días que tarda en pagar.
* En un ODS un gerente podría enterarse a grandes rasgos si el cliente está listo para un cambio en los límites de crédito, si se puede o debe reestructurar la deuda y algunas otras notas que le haya escrito su predecesor o algún otro empleado.
* Hoy día se considera que la estructura informática de una empresa madura contiene OLTPs + ODS + Data warehouse + OLAPs + Minería de datos.

**Subtema 1.2:**

Data Mining / BigData

**MARCO TEÓRICO**

**Data Mining**

* Es un proceso para detectar información de conjuntos grandes de datos, de la manera más automáticamente posible. Su objetivo es encontrar patrones, tendencias o reglas que expliquen el comportamiento de los datos en un contexto específico.

**Big Data**

* Es un término que describe el gran volumen de datos, tanto estructurados como no estructurados, que inundan los negocios cada día. Pero no es la cantidad de datos lo que es importante. Lo que importa con el Big Data es lo que las organizaciones hacen con los datos.

1. **MATERIALES**

Según la experiencia a realizar, de acuerdo al catálogo de equipamiento del taller.

1. EQUIPOS

Computadora personal o Laptop.

1. HERRAMIENTAS

Bloc de Notas, Excel y SQL Server.

1. CONSUMIBLES
   1. Archivo Excel, proporcionado por Almacén, conteniendo los datos de los productos, categorías de productos y proveedor del producto.
   2. Archivo Texto, proporcionado por Personal, conteniendo los datos de todos los empleados que intervienen en el proceso de Ventas.
   3. Script SQL con estructuras de las tablas de Ventas, proporcionado por el área de Sistemas.
   4. Script SQL con datos de las tablas de Ventas, proporcionado por el área de Sistemas.
2. **Abrir el SQL Server Management Studio para crear una Base de Datos de tipo Stage o Landing**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. **Crear la estructura de la tabla Personal.**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. **Crear la estructura de la tabla Categorías.**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. **Crear la estructura de la tabla Proveedores**

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

1. **Crear la estructura de la tabla Productos.**

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

1. **Luego, abrir el Visual Studio 2019 Data Tools, para crear un nuevo proyecto de Integration Services.**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. **Le damos un nombre a nuestro proyecto: ETL\_STG\_NW.**
2. **Seguidamente creamos un objeto de tipo Execute SQL Task, donde colocaremos las instrucciones necesarias para tener las tablas de la BD STAGE sin datos.**

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. **En propiedades del objeto le agregamos a la propiedad Connection una conexión a la Base de Datos STG\_NW.:**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. **En la propiedad SQLStatement, le colocamos el siguiente código, para limpiar el contenido de las tablas existentes:**

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. **La propiedad BypassPrepare debe de encontrarse en False, luego Aceptar.**
2. **Creamos un objeto de tipo Data Flow Task, que servirá para recuperar los datos de las Categorías desde el archivo en Excel.**

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

1. **Ingresamos al objeto Data Flow Task con nombre Categorías y procedemos a crear los siguientes objetos:**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. **El objeto de tipo Excel Source, nos permitirá acceder al archivo en Excel, específicamente a la hoja que contiene los datos de las categorías, para ello debemos configurar la conexión que utilizará el archivo en Excel.**

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. **Luego especificamos la hoja de la cual tomaremos los datos.**

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. **Seguidamente seleccionaremos los campos que serán utilizados.**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Con el objeto **Data Convertion**, procederemos a realizar cambios de ajustes en cuanto a tipos de datos y longitudes de los campos utilizados.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. En el objeto de tipo **OLE DB Destination** indicaremos como serán recibidos los datos, en este caso apuntaremos a la BD STG\_NW, tabla Categorías.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Con la opción **Mappings** procederemos a indicar que campos serán utilizados para la recepción de datos.

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza baja

1. Agregamos un control de tipo **Data Flow Task** para cargar los datos de los proveedores que se encuentran en el archivo en Excel.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. El objeto **Excel Source** nos permitirá conectarnos al archivo de Excel para recuperar los datos de proveedor, para este caso no es necesario volver a crear la conexión puesto que se utiliza el ya existente.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Los campos para utilizar son los siguientes:

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente

1. Con el control **Data Convertion** procederemos a realizar los ajustes con los datos, como son el tipo y la longitud.

Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente

1. Finalmente, usamos un objeto de tipo **OLE DB Destination** para recepcionar los datos en la tabla Proveedor de la BD STG\_NW.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Emparejamos los campos de origen con los de destino, considerando aquellos que fueron convertidos.

Tabla

Descripción generada automáticamente

1. Agregamos un nuevo control de tipo **Data Flow Task** para recuperar los datos de los productos.

Diagrama

Descripción generada automáticamente

1. Este control nos permitirá recibir los datos desde el archivo de Excel hacia la BD STG\_NW.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. El control de tipo **Excel Source** apuntará al archivo de Excel para recuperar los datos de los productos.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Los campos para usar son los siguientes:

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente

1. Se realizará un ajuste en los tipos de datos y las longitudes para que puedan ser recibidos en la BD.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Ahora usamos un control de tipo **OLE DB Destination** para que pueda recibir los datos la tabla Producto.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Seguidamente se enlazan las parejas de datos.

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente

1. Ahora procederemos a recuperar los datos de los empleados desde el archivo de texto.

Diagrama

Descripción generada automáticamente con confianza baja

1. Se utilizarán los siguientes controles para recuperar los datos desde el archivo de texto hacia la BD STG\_NW.

Interfaz de usuario gráfica

Descripción generada automáticamente

1. Desde el control de tipo **Flat File Source** apuntamos al archivo de texto de empleados.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Se indican las columnas a utilizar, especificando el tipo de salto de registro y delimitador de columna.

Interfaz de usuario gráfica

Descripción generada automáticamente

1. Validamos los campos que serán utilizados.

Interfaz de usuario gráfica, Tabla

Descripción generada automáticamente

1. Con el objeto **Data Conversion** procedemos a validar los tipos de datos y sus longitudes.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Luego apuntamos a la tabla de destino Employees.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Seguidamente emparejamos los campos de origen con los de destino.

Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente

1. Finalmente, ejecutamos el proyecto, debiéndose mostrar de la siguiente manera:

Diagrama

Descripción generada automáticamente

**ACTIVIDAD VIRTUAL:**

Ingresa a la plataforma virtual, luego desarrolla la siguiente actividad propuesta:

1. ¿Cuál es el objetivo de tener los datos en una Base de Datos Operacional?
2. ¿Es necesario que todas las tablas se encuentren relacionas? Explique
3. ¿Es necesario que se apliquen restricciones en las tablas ODS? Comente
4. ¿Las Bases de Datos Stage y ODS se pueden encontrar en el mismo servidor? Explique
5. Con la experiencia obtenida ¿Qué problemas tuvo para implementar el modelo de BD ODS? Comente

**CONCLUSIONES DE LA EXPERIENCIA**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_