

EVIDENCIA DE APRENDIZAJE N° 3

ALVARO JAVIER RADA ROMERO

ESTUDIANTE DE INGENIERIA DE SOFTWARE Y DATOS

DOCENTE

VICTOR HUGO MERCADO

CURSO: BASES DE DATOS II

INSTITUCIÓN UNIVERSITARIA DIGITAL DE ANTIOQUIA

CÓRDOBA 12/10/2024

A continuación, realizaremos el modelo en estrella para la creación de un Data Mart partiendo des de una base de datos relacional llamada “jardineria”, esta base datos contiene varias entidades con sus respectivos atributos, llaves primarias y llaves foráneas, tomando el caso de estudio anterior vamos a desnormalizar las distintas entidades, desnormalizar nos va ayudar a dar mejores tiempos de respuesta, algo que nos interesa ya que estamos creando un sistema orientado a consultas.

Estudiando el caso de estudio, Definiremos la tabla de hechos y sus respectivas dimensiones para así reagrupar los atributos de las entidades, desechar entidades innecesarias para nuestro caso de estudio y enlazar las nuevas relaciones entre las tablas de carácter dimensional y la tabla de hechos.

La realización asertiva de los “pasos” anteriormente mencionados nos dará como resultado el modelo en estrella que necesitamos, al igual tendremos un script totalmente funcional que derivará en una nueva base de datos.

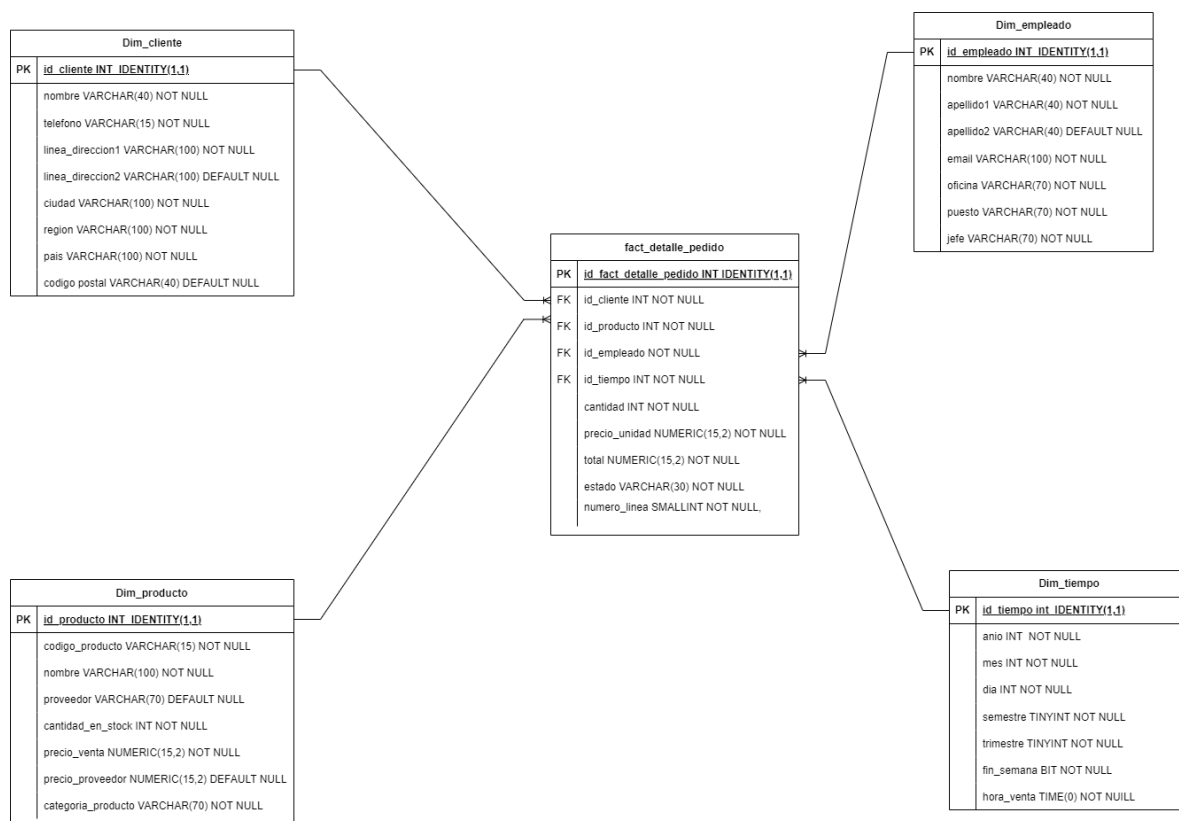
Con el modelo de estrella para nuestro Data Mart ya construido, comenzaremos la extracción de los datos necesarios para llenar el Data Mart, esto lo vamos hacer mediante la creación de un proyecto SSIS donde en primer lugar generaremos la extracción o bien el primer paso del ETL(Extract, Transform and Load) llevando los datos a una base de datos temporal llamada Staging.

Con la extracción finalizada, se inicia la Transformación de esos datos, con la transformación de estos datos se busca generar las dimensiones estipuladas en nuestro modelo de estrellas, con esta transformación buscamos normalizar los datos.

Como Objetivos tenemos.

- Desnormalizar la base de datos “jardineria” de nuestro caso de estudios.
- Ajustar los atributos y relaciones.
- Crear nuestra modelo en estrella.
- Mostrar nuestro diagrama de modelo en estrella
- Crear el script del modelo en estrella
- Crear un proyecto SSIS
- Crear una base de datos temporal o Staging
- Extraer los datos necesarios de la base de datos jardinería
- Cargar los datos en Staging
- Transformación de los datos extraídos
- Creación temporal de tablas en

Como solución para el modelo de estrella en base a las reglas del negocio, propuse eliminar las entidades o tablas “pagos”, “categoría_producto” y “oficina”; designar la tabla de “detalle_pedido” como la tabla de hechos y como tablas dimensionales propongo la dimensión tiempo, dimensión cliente, dimensión empleado, la cual se le agrega el atributo de “oficina”; dimensión producto a la cual se le suma el atributo “categoría_producto” y por último la dimensión pedido. Esto lo vamos a ver de la siguiente manera.



[VISUALIZAR MEJOR]

https://drive.google.com/file/d/162jSz78UQ7AhH_GYIs5bIhiLktobsHzP/view?usp=s_haring

Dimensiones propuestas

- Tiempo
- Cliente
- Empleado
- Producto

Script de las dimensiones:

- Dimensión Tiempo

```
CREATE TABLE dim_tiempo {  
    id_tiempo INT IDENTITY(1,1),  
    anio INT NOT NULL,  
    mes INT NOT NULL,  
    dias INT NOT NULL,  
    semestre TINYINT NOT NULL,  
    trimestre TINYINT NOT NULL,  
    fin_semana BIT NOT NULL,  
    hora_venta TIME(0) NOT NULL,  
    PRIMARY KEY (id_tiempo)  
}
```

- Dimension cliente

```
CREATE TABLE dim_cliente {  
    id_cliente INT IDENTITY(1,1),  
    nombre VARCHAR(40) NOT NULL,  
    apellido1 VARCHAR(40) NOT NULL,  
    apellido2 VARCHAR(40) DEFAULT NULL,  
    telefono VARCHAR(15) NOT NULL,  
    linea_direccion1 VARCHAR(100) NOT NULL,  
    linea_direccion2 VARCHAR(100) DEFAULT NULL,  
    ciudad VARCHAR(100) NOT NULL,  
    region VARCHAR(100) NOT NULL,  
    pais VARCHAR(100) NOT NULL,  
    codigo_postal VARCHAR(40) DEFAULT NULL,  
    PRIMARY KEY (id_cliente)  
}
```

- Dimensión empleado

```
CREATE TABLE dim_empleado {  
    id_empleado INT IDENTITY(1,1),  
    nombre VARCHAR(40) NOT NULL,  
    apellido1 VARCHAR(40) NOT NULL,  
    apellido2 VARCHAR(40) DEFAULT NULL,  
    email VARCHAR(100) NOT NULL,  
    oficina VARCHAR(70) NOT NULL,  
    puesto VARCHAR(70) NOT NULL,  
    jefe VARCHAR(70) NOT NULL,  
    PRIMARY KEY (id_empleado)  
}
```

- Dimensión producto

```
CREATE TABLE dim_producto {  
    id_producto INT IDENTITY(1,1),  
    codigo_producto VARCHAR(15) NOT NULL,  
    nombre VARCHAR(100) NOT NULL,  
    proveedor VARCHAR(70) DEFAULT NULL,  
    cantidad_en_stock INT NOT NULL,  
    precio_venta NUMERIC(15,2) NOT NULL,  
    precio_proveedor NUMERIC(15,2) DEFAULT NULL,  
    categoria_producto VARCHAR(70) NOT NULL,  
    PRIMARY KEY (id_producto)  
}
```

Tabla de hechos

- Detalle pedido

```
CREATE TABLE fact_detalle_pedido {  
    id_fact_detalle_pedido INT IDENTITY(1,1),  
    id_cliente INT NOT NULL,  
    id_producto INT NOT NULL,  
    id_empleado NOT NULL,  
    id_tiempo INT NOT NULL,  
    cantidad INT NOT NULL,  
    precio_unidad NUMERIC(15,2) NOT NULL,  
    total NUMERIC(15,2) NOT NULL,  
    estado VARCHAR(30) NOT NULL,  
    numero_linea SMALLINT NOT NULL,  
    PRIMARY KEY (id_fact_detalle_pedido),  
    FOREIGN KEY (id_cliente) REFERENCES dim_cliente (id_cliente),  
    FOREIGN KEY (id_producto) REFERENCES dim_producto (id_producto),  
    FOREIGN KEY (id_empleado) REFERENCES dim_empleado (id_empleado),  
    FOREIGN KEY (id_tiempo) REFERENCES dim_tiempo (id_tiempo)  
}
```


SCRIP PARA OBTENER LOS CAMPOS DE LA DIMENCION TIEMPO

SELECT

[ID_DimTiempo],

[fecha_pedido],

[anio],

[mes],

[dia],

CASE

WHEN DATEPART(month, [fecha_pedido]) BETWEEN 1 AND 6 THEN 1

ELSE 2

END AS [Semestre],

CASE

WHEN DATEPART(month, [fecha_pedido]) BETWEEN 1 AND 3 THEN 1

WHEN DATEPART(month, [fecha_pedido]) BETWEEN 4 AND 6 THEN 2

WHEN DATEPART(month, [fecha_pedido]) BETWEEN 7 AND 9 THEN 3

ELSE 4

END AS [Trimestre],

CASE

WHEN DATEPART(WEEKDAY, [fecha_pedido]) IN (1, 7) THEN 'Si'

ELSE 'No'

END AS [EsFinDeSemana]

FROM

[StagingJD].[dbo].[STDimTiempo]

Con todos los objetivos que nos habíamos planteado ya logrados creo decir que el ejercicio de desnormalización de las tablas o entidades fue lo más complicado, ya que naturalmente se nos enseña a normalizar nuestras bases de datos, con ese desafío culminado presento mi modelo de estrella para un Data Mart con sus respectivas dimensiones y tabla de hecho.

Los scripts propuestos para el gestor de base de datos MySQL server nos ayudan a asimilar las relaciones y atributos encontrados en mi diagrama del modelo de estrella y propuesta para el Data Mart.

Con la creación de un proyecto SSIS la extracción de los datos se vuelve algo repetitiva, por lo que es fácil trabajar la extracción de los datos con SSIS, a pesar de a ver tenidos problemas al final se lograron los objetivos propuestos para la extracción y carga de datos en Staging.

Al igual que la extracción, la transformación de datos se vuelve algo repetitiva dentro del proyecto SSIS, aquí dentro de la misma Staging creamos nuestras dimensiones aun sin cargar en el Data Mart, crear estas dimensiones con los datos extraídos es algo fácil de hacer, pero, el problema o reto en particular para un iniciado como yo, estuvo en hacer la transformación de los datos con las herramientas del SSIS.