

Práctica 2: Base de Datos Multidimensional

Michel Dennis Quitaquis
NIUB: 14991045

Tian Lan
NIUB: 16129094

ÍNDICE

[INTRODUCCIÓN](#)

[Proceso de la práctica](#)

[OBJETIVOS](#)

[DIAGRAMA DEL DATAMART](#)

[DESARROLLO](#)

[Tablas y inserciones](#)

[Schema Workbench](#)

[Explotación con Mondrian](#)

[DIFICULTADES SURGIDAS](#)

[Diseño del Datamart](#)

[Diseño del cubo](#)

[Explotación con Mondrian](#)

[CONCLUSIÓN](#)

[ENTREGA](#)

INTRODUCCIÓN

En esta práctica vamos a crear un Datamart de Ventas (**Privalia moda UB**) mediante el esquema estrella.

Empezaremos por el diseño del Datamart, cumpliendo los requisitos del enunciado y cumpliendo el esquema elegido.

Después daremos paso a la implementación del paso a tablas generando un script SQL (*main.sql*) que se cargará en el servidor Postgres de la Universidad de Barcelona.

Entonces procederemos a crear el cubo mediante la interfaz *Schema Workbench*, que nos permite crear y testear esquemas de cubos HOLAP de Mondrian, dicho programa nos generará un modelo XML con una específica estructura usada por el engine de MONDRIAN.

Mondrian, hace uso de motor multidimensional, open source. Es una herramienta cliente-servidor que sobre la base de datos nos proporcionará las herramientas de análisis multidimensional.

Éste modelo XML se generará en base a los datos residente en el servidor POSTGRES linkado con las dimensiones del cubo. Entonces podremos explotar el cubo mediante el cliente JPivot, que nos conecta con el cubo pudiendo mostrar los resultados de las consultas filtrando por los campos de la tabla de manera que se pueden poner diferentes criterios de búsqueda.

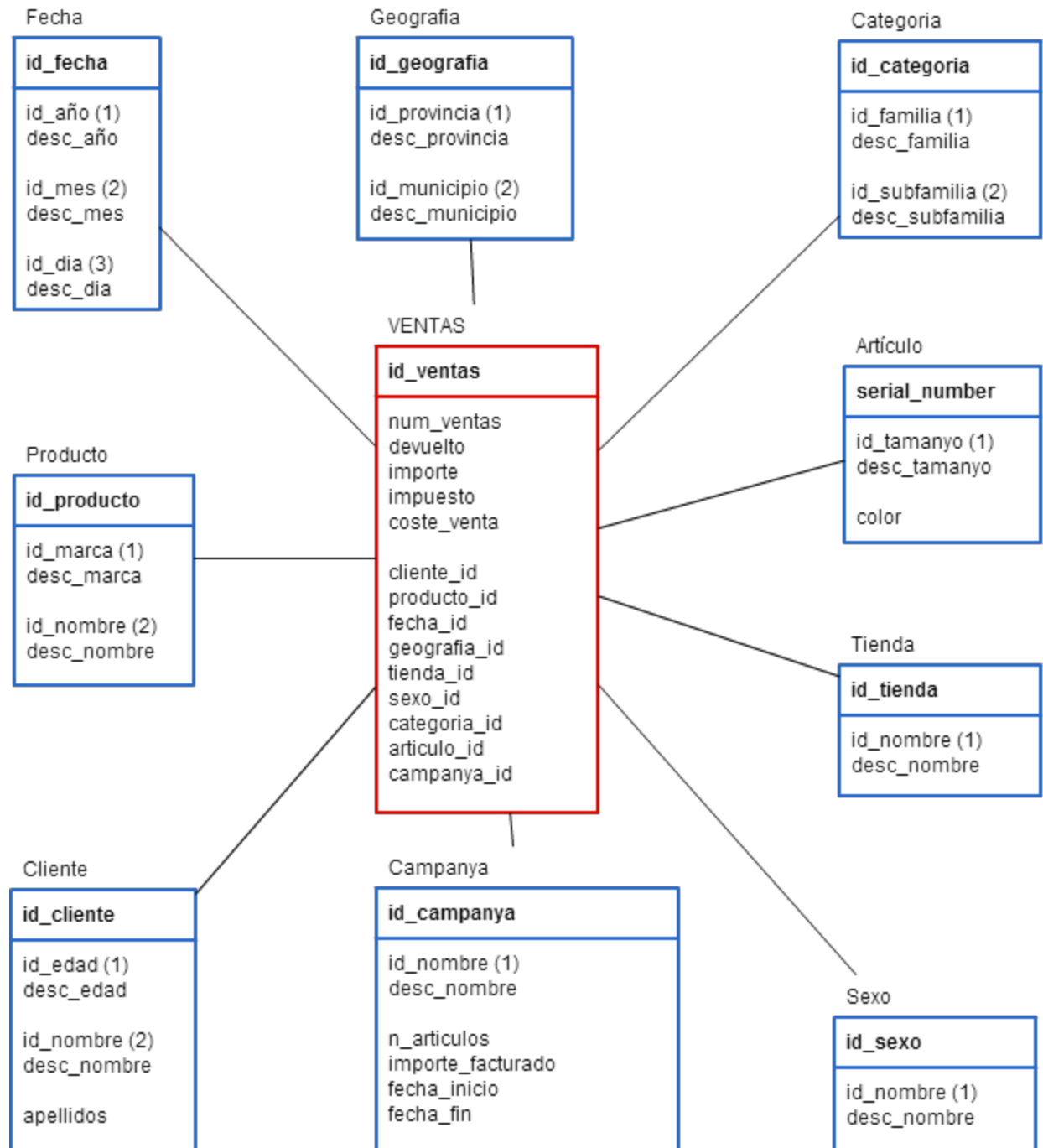
Proceso de la práctica

- Diseño del cubo
- Paso a tablas (tabla de hechos)
- Diseño del cubo (mapeo de tablas con el cubo)
- Implementación del cubo (con schema-workbench)
- Publicación al servidor (MONDRIAN)
- Pruebas de navegación (explotación del cubo)

OBJETIVOS

- Aprender a utilizar las herramientas para esta práctica (Schema Workbench, pgAdmin)
- Aprender a diseñar el almacén de datos (Datamart).
- Implementar la tabla de hechos y sus dimensión.
- Crear los datos “dummy” para nuestra base de datos.
- Diseñar el cubo para poder analizarla y explotarla posteriormente de forma multidimensional.
- Navegar por el cubo.

DIAGRAMA DEL DATAMART



DESARROLLO

Tablas y inserciones

Tal como se ha enseñado en el diagrama, hemos implementado para cada dimensión una tabla con sus respectivos atributos y también para la tabla de hecho otra tabla.

Se ha creado un script SQL con las definiciones de cada tabla que al ejecutarlo en Postgres las carga en las base de datos.

Luego además los datos “*dummy*” para cada tabla estan todos definidos en un archivo excel llamado “*dades.xlsx*” y luego independientemente cada tabla tiene sus datos en un archivo csv el cual se puede cargar en la base de datos.

Schema Workbench

Cubo: (1)

- Ventas

Tablas de hecho: (1)

- ventas

Dimensiones: (9)

- producto
- cliente
- tienda
- fecha (compartida)
- geografia (compartida)
- ...

Métricas: (5)

- num_ventas
- devolucion
- importe
- impuestos
- coste_venta

Explotación con Mondrian

El Schema-workbench genera un fichero xml, que es el que representa a nuestro cubo. Éste fichero se ha de mover a la carpeta de “*queries*” dentro del WEB-INF de Mondrian.

A partir de entonces para realizar las consultas de prueba, hemos estado modificando el fichero “*arrows.jsp*”.

DIFICULTADES SURGIDAS

1. Diseño del Datamart

- Confusión de nivel y atributo: a veces no sabes distinguir si un atributo tiene que ser un nivel o una dimensión
- Id de los niveles: utilizábamos un id para cada tabla de dimensión, al final entendimos que cada nivel tiene que tener asociado un id, ya que el filtro por índices será más eficiente. Y éste además ha de tener una descripción para que es dicho id.
- La diferencia entre la perspectiva sobre OLTP y OLAP

2. Diseño del cubo

- Identificación de los atributos de los elementos del cubo para poder enlazarlos bien con las tablas en postgresSQL.

3. Explotación con Mondrian

- Errores en la consulta: por la falta de conocimiento sobre el lenguaje de consulta del cubo, nos ha sido fácil tener errores en el código de consultas
- Muchas veces hay que reiniciar el servidor Mondrian sin darnos cuenta previamente

CONCLUSIÓN

Para esta práctica, hemos implementado una mini-versión de tienda Online, es decir, hemos acotado tanto el número de tablas (de hecho, de dimensión) como los datos “dummy” (muy pocos datos para cada tabla). Sin embargo creemos que el objetivo principal de entender el proceso total del diseño y la práctica nos ha sido bastante provechosa para dar un primer paso al mundo de BD multidimensional, la cuál es altamente utilizada por las empresas y instituciones.

ENTREGA

- El main.sql donde están implementadas las tablas.
- Los ficheros .csv que contienen los datos “dummy” para la tabla de hecho (ventas) y las de dimensiones.
- Cuenta postgres:
 - dquitata12_adm
 - curs1415
- Un fichero “Privalia.xml” generado que traduce de multidimensional a relacional (tablas) que representa el cubo.
- Un fichero “indice.html” donde se definen los enlaces a las diferentes consultas.
- Los ficheros “.jsp” donde está configurado el acceso a la nuestra base de datos en postgres; también está una consulta al cubo de los elementos que se mostrarán en la web.
 - query_campanyas.jsp
 - query_categorias.jsp
 - query_clientes.jsp
 - query_geografia.jsp
 - query_all.jsp