

Práctica 3: Informes y Consultas sobre un Data Mart

Almacenes y Minería de Datos
Curso 2020/2021

Grado en Ingeniería en Informática
Dpto. de Informática e Ingeniería de Sistemas
Universidad de Zaragoza
Escuela de Ingeniería y Arquitectura

Sergio Ilarri y Raquel Trillo

9 de noviembre de 2020

1. Objetivos / tareas a realizar

En esta práctica se diseñarán e implementarán los mecanismos necesarios para formular preguntas y elaborar informes a partir de la información almacenada en el data mart implementado en la práctica anterior. En mayor detalle, los objetivos de esta práctica son:

- Definir y probar diferentes consultas SQL en la base de datos analítica (OLAP) implementada en la práctica anterior.
- Realizar diferentes informes o cuadros de mando a partir de la información almacenada en dicha base de datos analítica.

2. Contexto del problema

Se pretende analizar los datos almacenados en el data mart que contiene información sobre vuelos comerciales en Estados Unidos implementado en las prácticas anteriores.

2.1. Parte 1: consultas

Consultas básicas

En primer lugar se requiere que se formulen las siguientes consultas empleando el lenguaje SQL:

- Obtener el retraso medio de los vuelos que salen de cada aeropuerto en función de la ciudad destino, junto con el retraso medio total por aeropuerto origen (con independencia del destino).
- Obtener el retraso medio de los vuelos que salen de cada aeropuerto en función de la ciudad destino, junto con el retraso medio total por aeropuerto origen (con independencia del destino) y el retraso medio total por destino (con independencia del aeropuerto origen).

Consultas propias (operadores de GROUP BY)

Además, se requiere que se definan al menos 5 consultas adicionales no triviales donde se utilicen operadores como *CUBE*, *ROLLUP*, y *GROUPING SETS* (operadores que permiten extender la cláusula *GROUP BY*), así como otros operadores que puedan determinarse como especialmente interesantes para las consultas analíticas.

Además de los operadores indicados, se recomienda considerar, en la medida de lo posible, funcionalidades que vayan más allá de las vistas en clase (lo que permitirá potenciar más el aprendizaje autónomo, la creatividad y la experimentación personal), indicando si forman o no parte del estándar de SQL.

Sobre la sentencia *CREATE DIMENSION* y similares

Por último, considerando los tipos de consultas que podéis formular, se requiere determinar la relevancia de la sentencia *CREATE DIMENSION* de Oracle: ¿es necesaria la utilización de esta sentencia? ¿aporta alguna utilidad? En este punto resulta también de interés revisar la documentación de Oracle e identificar y probar, si resulta relevante, otras sentencias que puedan ser de interés para la creación de esquemas en estrella, reflejando en la memoria de la práctica el trabajo realizado y las pruebas que se han llevado a cabo para evaluar estas funcionalidades.

2.2. Parte 2: informes

A continuación, se generarán al menos 2 informes (*reports*) o cuadros de mando (*dashboards*), de contenido libre siempre y cuando sean relevantes. Para ello, se considerará preferentemente el posible uso de una de las siguientes herramientas: *QlikView*¹, *Pentaho Report Designer*², o *Tableau*³ (por ejemplo, *Tableau Desktop*⁴).

También podría considerarse el uso de más de una herramienta y/o de otras adicionales a las indicadas. Por ejemplo, se puede considerar la utilización de *Excel* para la creación de informes de tablas dinámicas / *pivot tables* o la realización de posibles pruebas con *Microsoft Power BI*⁵ (para sistemas Windows) o con otras herramientas que se consideren de interés.

Aunque se pueden utilizar los recursos disponibles en las máquinas del laboratorio, se recomienda utilizar una máquina virtual, donde se dispondrá de una mayor libertad para probar las herramientas deseadas, así como una mayor flexibilidad para resolver posibles problemas.

¹<https://www.qlik.com/us/products/qlikview>

²[https://sourceforge.net/projects/pentaho/files/Report Designer/](https://sourceforge.net/projects/pentaho/files/Report%20Designer/)

³<https://www.tableau.com>

⁴<https://www.tableau.com/products/desktop>

⁵<https://powerbi.microsoft.com/>

3. Entrega de la práctica

La práctica se realizará en equipos de tres personas (salvo que existan problemas logísticos, que deberán comentarse y resolverse previamente con el profesor). Al finalizar la práctica se debe entregar un fichero .zip denominado *pN_NIP1_NIP2_NIP3.zip* (donde NIP1, NIP2 y NIP3 son los NIP de los autores de la práctica, con $NIP1 < NIP2 < NIP3$, y en *pN* la *N* representa el número de práctica) con el siguiente contenido:

1. Un fichero PDF denominado **informe.pdf**, que contendrá la memoria de prácticas (respuestas a las cuestiones planteadas, esquemas y código desarrollados, etc., según el caso). En la portada de dicho documento se indicarán los componentes del grupo de prácticas (nombre y apellidos y NIP de cada uno) y el número de práctica. El documento contendrá todo lo necesario para la evaluación de las prácticas. El informe de prácticas debe contener al final un apartado de conclusiones personales que incluirá, entre otras cosas, información sobre el tiempo invertido por cada miembro del grupo de prácticas en la realización de la misma.
2. Todos los fuentes, proyectos, ficheros de configuración y programas de prueba desarrollados, si es el caso, en una carpeta **fuentes**.

Al descomprimir el fichero .zip se deben extraer los ficheros y directorios necesarios en el directorio *pN_NIP1_NIP2_NIP3*. Es importante seguir las convenciones de nombrado y la estructura de ficheros y directorios descrita.

Para la entrega del fichero .zip, se utilizará Moodle del Anillo Digital Docente de la Universidad de Zaragoza. La fecha límite de entrega es el día anterior al de la siguiente sesión de prácticas en el laboratorio a las 23:59.

La práctica entregada debe contener, además de la memoria, todos los ficheros fuentes y proyectos realizados para su resolución. Se recomienda estructurar la memoria de acuerdo con la división en las 2 partes y apartados indicados en este guión. Para la parte de informes, además de los fuentes es aconsejable incluir capturas u otros elementos que permitan apreciar adecuadamente su contenido en la memoria.

4. Recomendaciones

Una vez realizadas las prácticas y entregadas, cada grupo debe presentárselas al profesorado de prácticas en la siguiente sesión de prácticas. Al realizar la presentación el profesorado le formulará cuestiones sobre las decisiones que ha tomado. La práctica debe entregarse en los términos indicados anteriormente y debe ser original (no debe haber sido copiada). Se considerarán los siguientes aspectos:

- La consecución de los objetivos planteados y la adecuada realización de las tareas correspondientes.
- La justificación de las afirmaciones incluidas en el informe.
- La adecuada inclusión de referencias a fuentes de información, si procede.
- La metodología para llevar a cabo la búsqueda de información necesaria para la realización de la práctica.

- La creatividad para la búsqueda de soluciones apropiadas y el correcto análisis de alternativas.
- La estructura y presentación del informe elaborado (completitud del trabajo, precisión, explicaciones adecuadas y completas, justificación de las decisiones tomadas, referencias adecuadas, coherencia de discurso, ausencia de errores tipográficos y ortográficos, etc.).

Hay que asegurarse de que, en caso de que la práctica incluya algún tipo de desarrollo, funciona correctamente en los ordenadores del laboratorio o puede probarse fácilmente sobre una máquina virtual base proporcionada por el profesorado. También es importante someter código limpio (donde se ha evitado introducir mensajes de depuración y comentarios que no proporcionan información relevante).

Es importante extraer conclusiones: resultados obtenidos, indagaciones realizadas, diferencias observadas entre las distintas alternativas y/o herramientas, dificultades encontradas, valoración de los aprendizajes o mejoras de habilidades conseguidas, opinión personal, etc. Se recomienda destacar especialmente aquellos aspectos que muestren la capacidad de indagación, búsqueda de información, autonomía, y curiosidad personal que se hayan desarrollado.