

Práctica – Data Warehouse (4 punto) – EAD Q2

Enunciado

Diseñar, implementar y explotar un DW para una aplicación de citas online.

Tarea 1 (1+1 puntos): Diseño del modelo multidimensional utilizando las fases del modelo de Kimball.

1. Determinar el proceso de negocio a analizar.
2. Establecer el nivel de detalle o granularidad (de la tabla de hechos).
3. Elegir las dimensiones.
4. Determinar los hechos medibles.

Más concretamente, se debe incluir:

- Una explicación de las suposiciones realizadas sobre los requisitos de explotación y los atributos a considerar. Debe considerarse aquí las necesidades de consulta. Para la justificación de las dimensiones y sus atributos considerados se debe incluir ejemplos de consultas analíticas en donde se usarían.
- Un diagrama, siguiendo la notación DFM vista en clase, que refleje los hechos, dimensiones y atributos seleccionados.
 - Se recomienda seleccionar nombres descriptivos para los atributos, evitando el uso de abreviaturas. No es necesario incluir explicaciones detalladas de los atributos, salvo cuando sean necesarias para entender su significado o funcionamiento.
 - Acompañando al diagrama, debe incluirse una justificación de las decisiones más relevantes tomadas, y en particular las relacionadas con:
 - Atributos que se hayan considerado inicialmente y posteriormente descartado como no relevantes para la explotación.
 - Dimensiones que tengan particularidades relevantes: dimensiones degeneradas, dimensiones que cambian en el tiempo y la estrategia seleccionada para representarlas (y la justificación de la elección de dicha estrategia), etc.
 - Medidas/métricas seleccionadas con su significado. Se debe incluir información sobre el carácter aditivo, semiaditivo o no aditivo de cada una de las métricas.

Tarea 2 (1 punto): Diseño del modelo lógico e implementación del DW en Oracle

- Proponer el diagrama del modelo lógico para el diseño realizado (normalmente será un modelo sencillo, en estrella o en copo de nieve), incluyendo una breve explicación de las decisiones tomadas.
- Crear las tablas resultantes en Oracle e insertar algunos datos de ejemplo.

Tarea 3 (1 punto): Explotación del DW

- Se deben plantear y resolver tres consultas analíticas relevantes sobre el DW implementado. Para cada consulta, debe incluirse una breve descripción de la consulta (qué hace, y por qué es relevante o para qué podría aplicarse), y la sentencia SQL que resuelve dicha consulta sobre el esquema seleccionado. Las consultas deben de ser realistas y diferentes (utilizando diferentes tipos de consultas analíticas).

Comentarios:

- El enunciado es intencionadamente abierto, a fin de permitir múltiples interpretaciones. Se valorará fundamentalmente la consistencia entre las suposiciones realizadas y el modelo generado, al no tener conocimiento experto del dominio.
- Se espera que se incluyan al menos unas 4-5 dimensiones, de forma que el alcance del ejercicio sea suficiente para valorar los conocimientos del tema, y se valorará la complejidad del modelo presentado (teniendo en cuenta los conceptos avanzados de diseño de DW).

Grupos

La práctica se realizará en grupos de cuatro estudiantes. Serán, salvo causa de fuerza mayor, los mismos que se han establecido el campus virtual.

Fechas de entregas y defensa

Las entregas se realizarán a través del campus virtual.

Primera entrega (sólo tarea 1): 28 de marzo de 2025, hasta las 23:59 (1 punto)

Segunda entrega (tarea 1 revisada + tareas 2 y 3): 13 de mayo de 2025, hasta las 23:59 (3 puntos, que se otorgarán después de la defensa).

Defensa: durante último día de clase de prácticas de la asignatura, es decir, 14/16 de mayo de 2025. La defensa será presencial, y permitirá comprobar que los estudiantes han adquirido correctamente los fundamentos del diseño de DW utilizados.

Entregables

Primera entrega: un documento pdf que refleje el trabajo realizado en la primera tarea (descripción del entorno analítico, diagrama y justificaciones).

Segunda entrega: un fichero comprimido que incluya

- Un documento pdf con el detalle del trabajo realizado, e incluya lo solicitado en cada tarea (descripción del entorno analítico, diagramas, enunciados de consultas, justificaciones, etc.).
- Scripts de creación de tablas e inserción de datos.
- Scripts para resolución de consultas.