



UNIVERSIDAD DE MÁLAGA



Trabajo Almacenes de Datos

Análisis de datos (MDX)

Crear un cubo multidimensional del almacén para poder realizar consultas en MDX.

Realizado por
De Pablo Diego y Soriano Juan

Profesor encargado:
Luque Baena, Rafael Marcos

Departamento
Lenguajes y Ciencias de la Computación

MÁLAGA, DICIEMBRE de 2024



UNIVERSIDAD
DE MÁLAGA



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA
ESTUDIANTES DE INGENIERÍA BIOINFORMÁTICA

Análisis de datos (MDX)

Almacenes de Datos

Realizado por
De Pablo Diego y Soriano Juan

Profesor encargado:
Luque Baena, Rafael Marcos

Departamento
Lenguajes y Ciencias de la Computación

UNIVERSIDAD DE MÁLAGA
MÁLAGA, DICIEMBRE DE 2024

Contents

1	Introducción	3
2	Objetivos	3
3	Modificación del almacén de datos	3
4	Creación del cubo	4
5	Dificultades encontradas	5
6	Conclusión	5
7	Github y conjunto de instrucciones para su correcto despliegue en SQL Server.	5

1 Introducción

Este documento es una continuación de las fases anteriores del proyecto, donde se desarrollaron el **diseño conceptual y lógico** de un almacén de datos basado en la información proporcionada por la *Base de Datos de Investigación Colaborativa eICU* [1], y la **integración de datos** mediante un proceso de Extracción, Transformación y Carga (ETL). Para la creación de un almacén enfocado a **pacientes con patologías respiratorias**.

En esta fase, se busca construir un **cubo multidimensional** utilizando los datos del almacén previamente desarrollado, En el ámbito hospitalario, la creación de un cubo multidimensional ofrece ventajas al permitir analizar datos clínicos complejos de manera ágil. Posteriormente se realizarán consultas en *MDX* (Multidimensional Expressions) sobre dicho cubo. Permitiendo extraer información de manera eficiente al navegar por las dimensiones y medidas del cubo.

Además, se describirán los problemas encontrados durante el desarrollo y las soluciones aplicadas, asegurando que las consultas generen resultados útiles para el análisis de datos.

2 Objetivos

Los objetivos principales de este informe son los siguientes:

- Realizar las correcciones pertinentes en el proceso ETL que permitan el desarrollo del cubo multidimensional.
- Implementar un cubo multidimensional adaptado para el análisis de **pacientes con patologías respiratorias**.
- Desarrollar al menos 8 consultas en *MDX* que exploren diferentes dimensiones y hechos del cubo, aplicando funciones avanzadas cuando sea necesario.
- Documentar de manera clara y replicable el proceso de creación del cubo multidimensional y las consultas *MDX*, proporcionando instrucciones detalladas para su ejecución.
- Describir los problemas encontrados durante el desarrollo del cubo y las consultas, y detallar las soluciones empleadas.

3 Modificación del almacén de datos

Eliminamos Age y PatientUnitStayID de paciente Cogemos los Pacientes únicos quedándonos con su primera aparición en la base de datos Antes nos daba 1849 y ahora después del cambio mencionado en la segunda línea nos salen 1841 como debe ser El tiempo seleccionarlo con distinct y diagnosis coger el primer string del DiagnosisString (la primera dimension) Corregimos la creación de las tablas intermedias para las relaciones NxM

- En caso de que se haya tenido que modificar el almacén y el ETL, se ha incluido una sección que indica brevemente las modificaciones realizadas y por qué ha sido necesario acometerlas.

4 Creación del cubo

- Una sección que describa, a modo de tutorial, el proceso para la creación del cubo de vuestro proyecto.
 - La memoria es descriptiva y completa, describiendo correctamente, a modo de tutorial, los pasos para la creación del cubo multidimensional.
 - Existe un conjunto de instrucciones para desplegar el proyecto en el equipo y es posible su implementación siguiéndolas.

5 Consultas en MDX

- Una sección con las consultas en MDX y una captura con el resultado de cada una de ellas (la imagen capturada no tiene por qué mostrar todas las tuplas resultantes).

6 Instrucciones para ejecutar las consultas.

- Una sección con las instrucciones detalladas para que un evaluador pueda ejecutar las consultas en su máquina.
 - Para cada consulta MDX, se muestra, además del enunciado de la consulta en sí, la consulta realizada en MDX y una captura del resultado generado.
 - Indica el número de consultas MDX que se han intentando, es decir, que se ha escrito código MDX y se ha mostrado el resultado de la misma, independientemente de si son correctas o no:
 - (***) 8 o más.
 - Según tu experiencia, valora cómo están realizadas las consultas en MDX, seleccionando la opción más adecuada:
 - (****) Todas las consultas parecen ser correctas, teniendo sentido la salida de cada una de ellas.
 - En algunas consultas se han usado funciones más avanzada de MDX, que impliquen el uso de métodos para recorrer una jerarquía (PREVMEMBER, CURRENTMEMBER, PARENT, etc.).

7 Problemas encontrados

Maldita maquina virtual

- Una sección "problemas encontrados" que explique con cierto detalle los problemas que se han encontrado durante la realización de la práctica. Se permite que la información incluida en esta sección se encuentre dividida o dispersada a lo largo del documento.
- La sección de "problemas encontrados" (o similar) es coherente y explica con cierto detalle los problemas que se han encontrado durante la realización de la práctica (ojo, no tienen por qué haber sido resueltos). Se permite que la información incluida en esta sección se encuentre dividida o dispersada a lo largo del documento.
- Calificación final.

8 Conclusión

XD

9 Github y conjunto de instrucciones para su correcto despliegue en SQL Server.

Todo el proyecto está accesible en github [2] donde se detalla más específicamente como desplegar en SQL.

References

- [1] eICU Collaborative Research Database. *eICU Collaborative Research Database*. <https://eicu-crd.mit.edu/about/eicu/>. Accessed: 2024-11-14. 2024. URL: <https://eicu-crd.mit.edu/about/eicu/>.
- [2] Diegodepab. *Almacén UCI Sanitaria*. https://github.com/Diegodepab/almacen_UCI_Sanitaria. Accessed: 2024-11-14. 2024. URL: https://github.com/Diegodepab/almacen_UCI_Sanitaria.



UNIVERSIDAD
DE MÁLAGA

| **uma.es**

E.T.S. DE INGENIERÍA INFORMÁTICA

E.T.S de Ingeniería Informática
Bulevar Louis Pasteur, 35
Campus de Teatinos
29071 Málaga