

# Análisis de datos (MDX)



**Alejandro Silva Rodríguez**

**Marta Cuevas Rodríguez**

*Almacenes De Datos*  
Universidad de Málaga

Diciembre 2024

# Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Objetivos</b>	<b>2</b>
<b>3. Creación del Cubo Multidimensional</b>	<b>2</b>
3.1. Pasos Iniciales . . . . .	2
3.2. Creación del Origen de Datos . . . . .	3
3.2.1. Configuración de la Conexión . . . . .	4
3.3. Creación de la Vista de Origen de Datos . . . . .	4
3.4. Creación del Cubo . . . . .	6
3.5. Relaciones y Dimensiones . . . . .	8
3.6. Cálculo de Medidas . . . . .	10
3.7. Jerarquías y Atributos . . . . .	10
3.8. Procesamiento del Cubo . . . . .	14
3.9. Conexión al Cubo Procesado . . . . .	15
3.10. Verificación de Jerarquías . . . . .	15
<b>4. Consultas MDX</b>	<b>16</b>
4.1. Consulta 1 . . . . .	16
4.2. Consulta 2 . . . . .	16
4.3. Consulta 3 . . . . .	17
4.4. Consulta 4 . . . . .	17
4.5. Consulta 5 . . . . .	17
4.6. Consulta 6 . . . . .	17
4.7. Consulta 7 . . . . .	17
4.8. Consulta 8 . . . . .	17
4.9. Consulta 9 . . . . .	17
4.10. Consulta 10 . . . . .	18
<b>5. Tutorial ejecutar consultas</b>	<b>18</b>
<b>6. Dificultades Encontradas</b>	<b>18</b>
<b>7. Conclusión</b>	<b>18</b>
<b>8. Acceso al Repositorio</b>	<b>19</b>

# 1. Introducción

En el contexto hospitalario actual, el análisis avanzado de datos se ha convertido en una herramienta indispensable para optimizar la toma de decisiones y mejorar la gestión de recursos en áreas críticas como las Unidades de Cuidados Intensivos (UCI). El análisis detallado del gasto en medicamentos, que representa una proporción significativa de los costos operativos, requiere técnicas especializadas que permitan explorar grandes volúmenes de datos desde múltiples perspectivas.

Tras la construcción de un almacén de datos orientado al análisis del gasto en medicamentos, el siguiente paso lógico es implementar estructuras que soporten consultas analíticas avanzadas. Los cubos multidimensionales permiten una visión integral de los datos, facilitando la identificación de patrones, tendencias y áreas críticas de gasto. Además, el uso de consultas MDX (Multidimensional Expressions) habilita a los usuarios para realizar análisis dinámicos y obtener insights clave de manera rápida y eficiente. Este trabajo se centra en la creación y explotación de un cubo multidimensional diseñado específicamente para analizar el gasto en medicamentos en pacientes ingresados en UCI en hospitales de EE.UU.

# 2. Objetivos

El objetivo principal de este trabajo es diseñar, implementar y explotar un cubo multidimensional para analizar el gasto en medicamentos en las UCI mediante el uso de consultas MDX. Este propósito se concreta en los siguientes objetivos específicos:

- Diseñar y construir un cubo multidimensional que permita explorar de manera eficiente el gasto en medicamentos desde múltiples dimensiones, como tiempo, tipo de medicamento y características del paciente.
- Implementar consultas MDX que permitan realizar análisis detallados.

# 3. Creación del Cubo Multidimensional

En esta sección se explicará el proceso de creación del cubo multidimensional a partir del almacén de datos cargado, de manera detallada.

## 3.1. Pasos Iniciales

Para comenzar con el proyecto, abre Visual Studio y crea un nuevo proyecto de tipo **Proyecto Multidimensional de Analysis Services**, como se muestra en la figura 1.

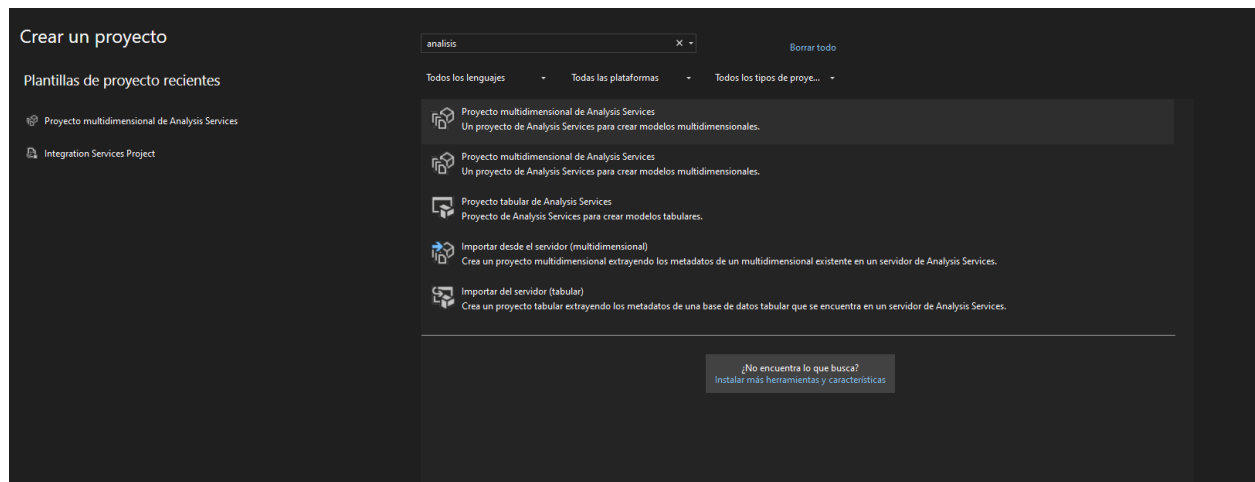


Figura 1: Creación de un nuevo proyecto multidimensional en Visual Studio.

### 3.2. Creación del Origen de Datos

El siguiente paso consiste en crear un nuevo **Origen de Datos**. Para ello, selecciona la opción de crear una nueva conexión, como se muestra en la figura 2. Posteriormente, configura la conexión con **Autenticación SQL Server**, utilizando una cuenta de servicio.

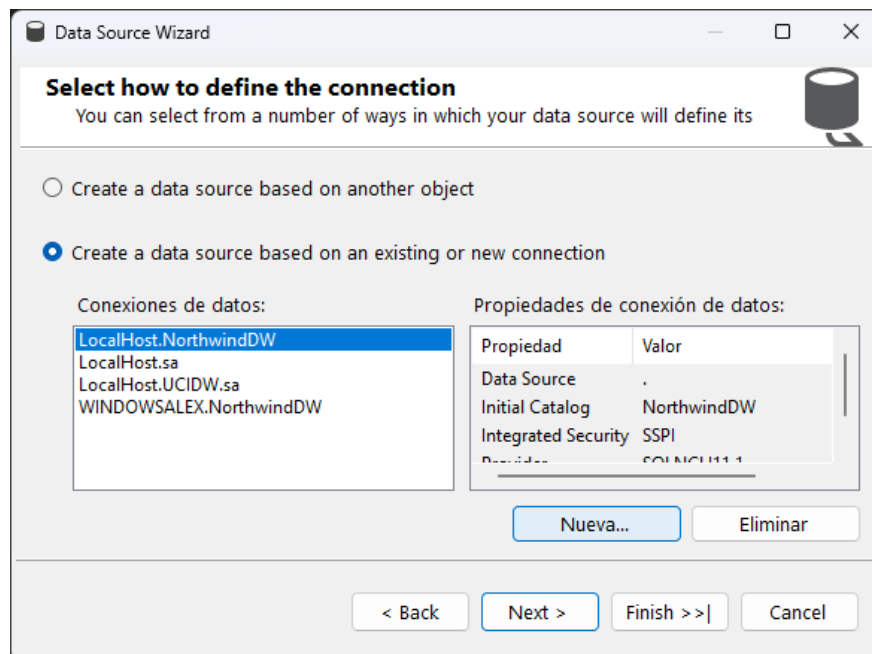


Figura 2: Nueva conexión en Visual Studio.

### 3.2.1. Configuración de la Conexión

A continuación, selecciona **Autenticación SQL Server** y utiliza la cuenta de servicio para configurar la conexión a la base de datos.

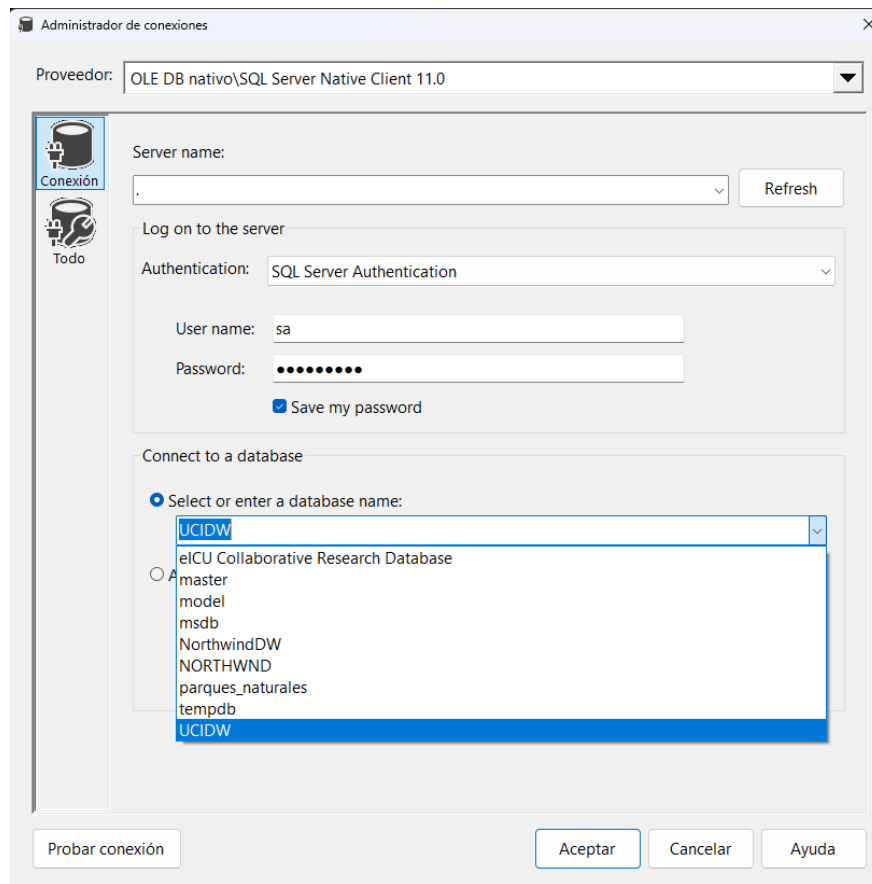


Figura 3: Configuración de la autenticación SQL Server.

### 3.3. Creación de la Vista de Origen de Datos

Una vez configurada la conexión, creamos una nueva vista de origen de datos. Para ello, seleccionamos el origen de datos previamente configurado, tal como se muestra en la figura 4.

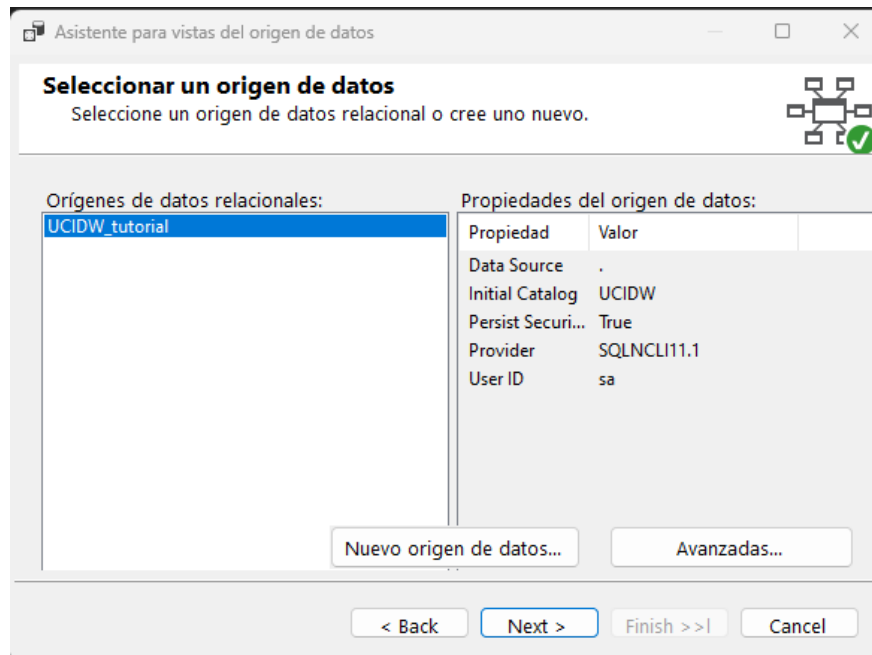


Figura 4: Selección del origen de datos configurado previamente.

A continuación, seleccionamos las tablas necesarias del origen de datos.

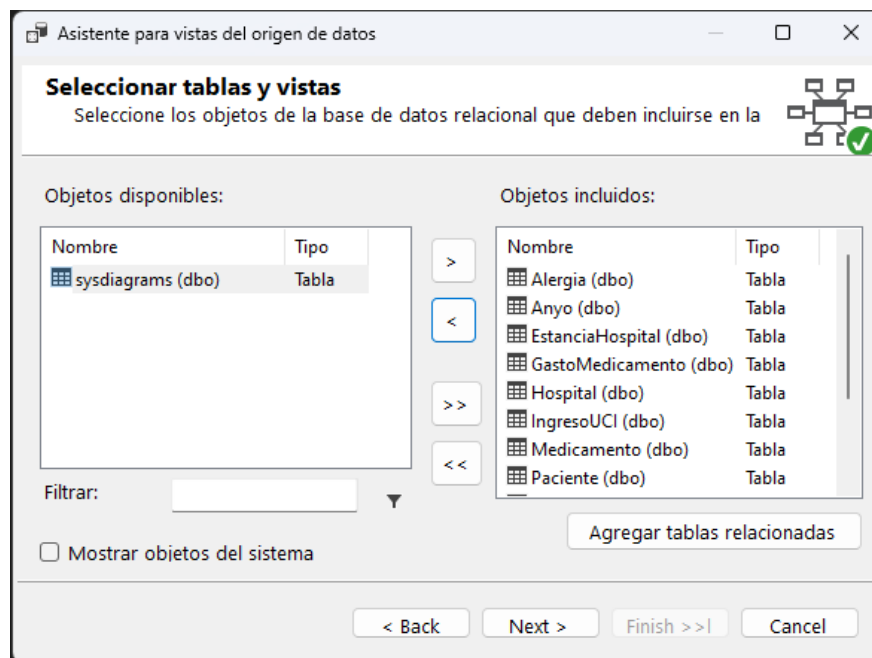


Figura 5: Selección de tablas del origen de datos.

Verifica la vista de la base de datos para asegurarte de que todas las tablas estén correctamente configu-

radas.

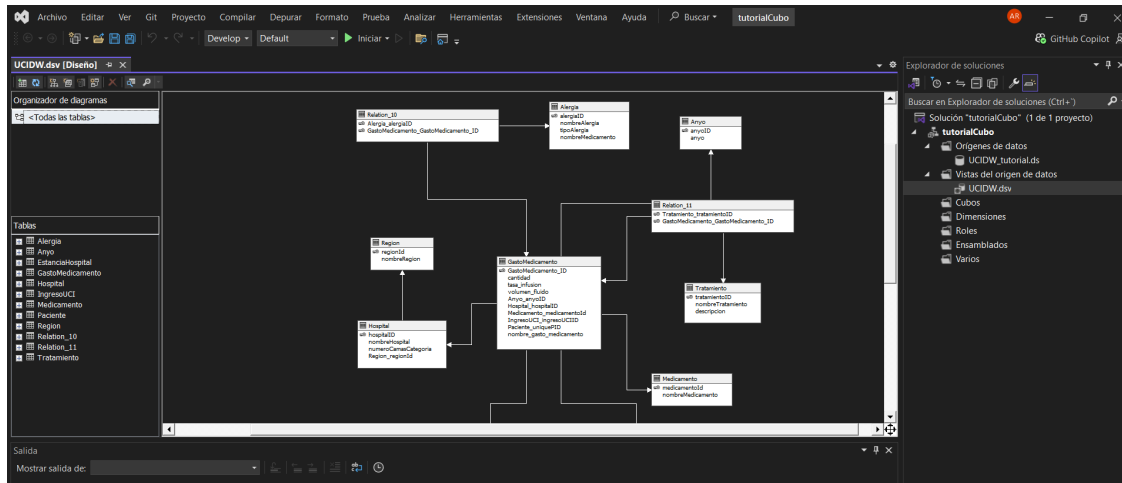


Figura 6: Verificación de la vista de la base de datos.

### 3.4. Creación del Cubo

Una vez configuradas las tablas, procede a crear el cubo multidimensional utilizando las tablas seleccionadas. En la figura 7 se muestra el proceso de creación del cubo.

The 'Asistente para cubos' dialog box is shown. It prompts the user to select a method for creating the cube. The 'Usar tablas existentes' option is selected. The description field contains the text: 'Cree un cubo basado en una o varias tablas de un origen de datos.'

Figura 7: Creación del cubo multidimensional con las tablas seleccionadas.

En este paso, seleccionamos las tablas del grupo de medida, como el **Gasto de Medicamento**, y los grupos de medida intermedios para las relaciones de tipo NM, como se muestra en la figura 8.

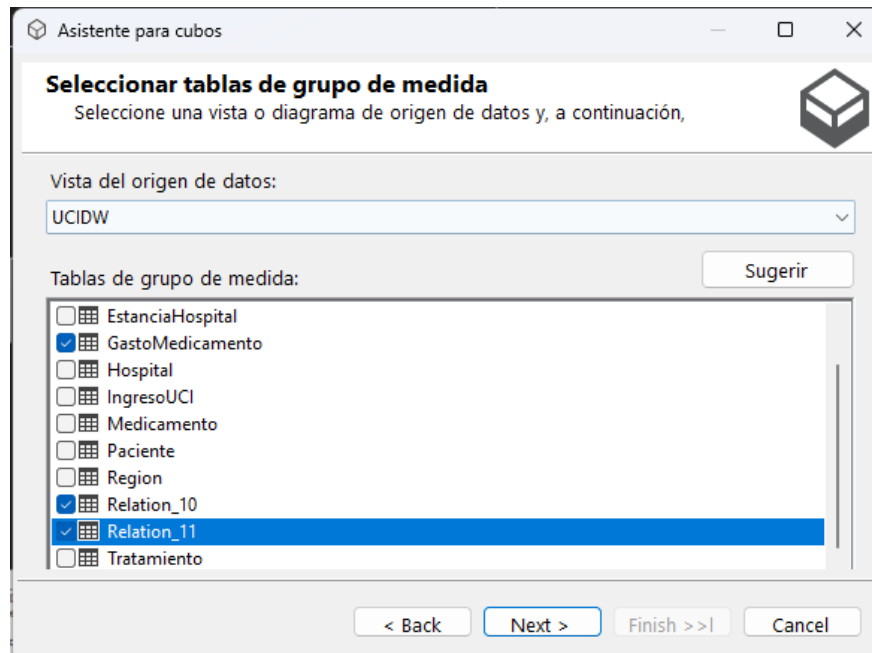


Figura 8: Selección de tablas del grupo de medida y relaciones intermedias.

Luego, seleccionamos las medidas y dimensiones del cubo, como se puede ver en la figura 10 y 9.

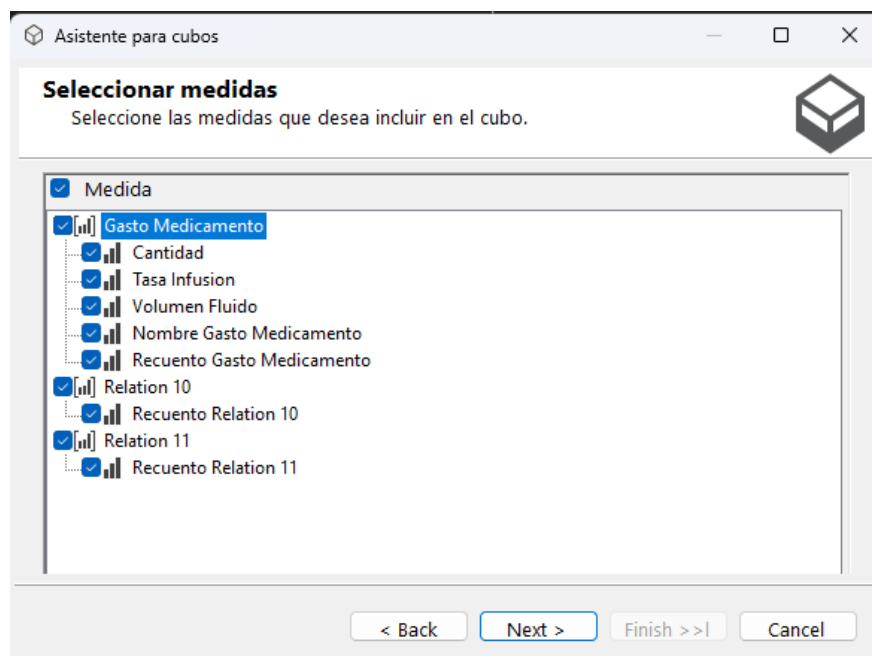


Figura 9: Configuración de medidas.



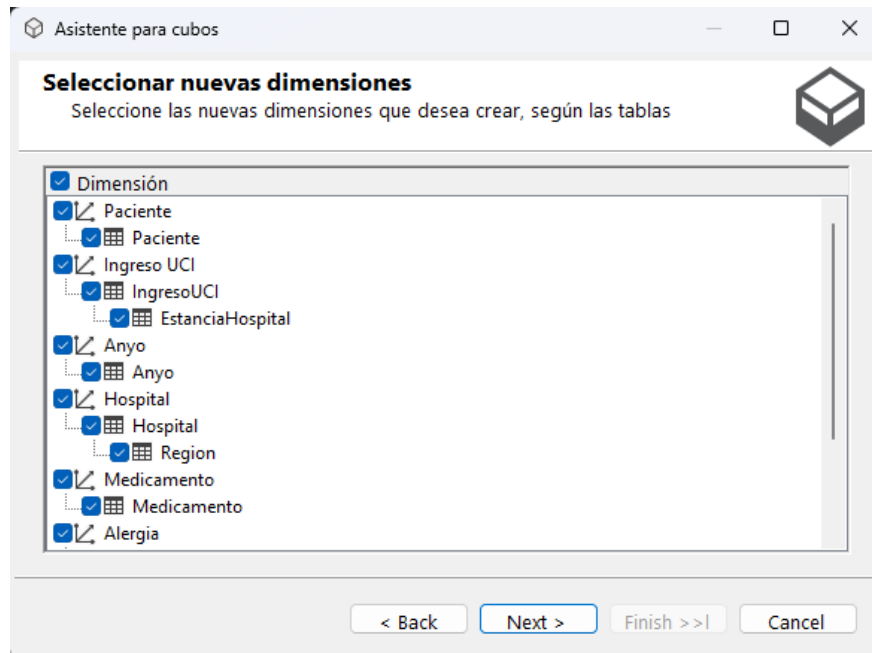


Figura 10: Configuración dimensiones en el cubo.

Finalmente, visualizamos el cubo creado en la figura 11.

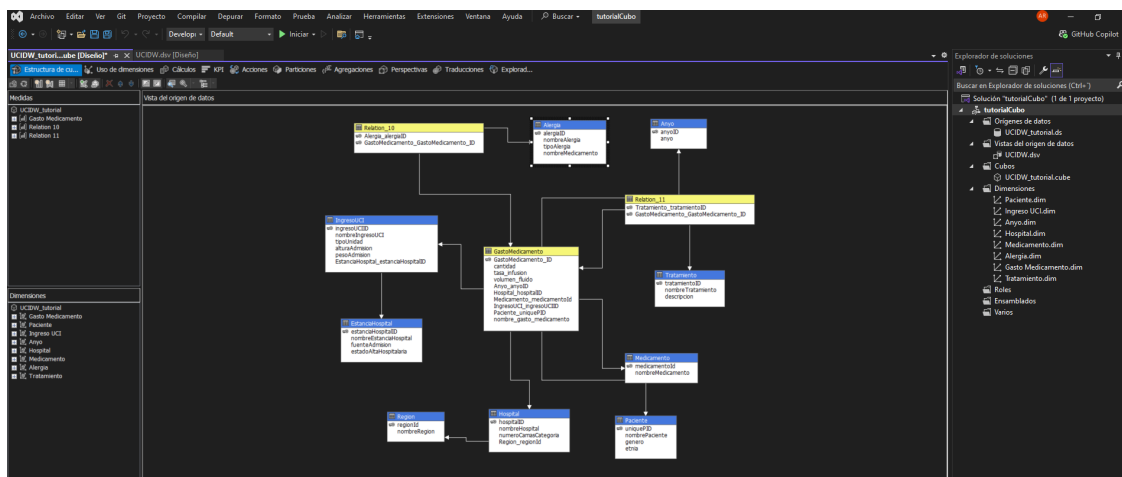


Figura 11: Visualización del cubo.

### 3.5. Relaciones y Dimensiones

Para configurar las relaciones, vamos a **Uso de Dimensiones**, donde definimos las relaciones entre el **Gasto** y el **Gasto** de tipo Hecho, así como las relaciones entre las dimensiones NM y los grupos de medidas intermedio. La figura 12 muestra este proceso.

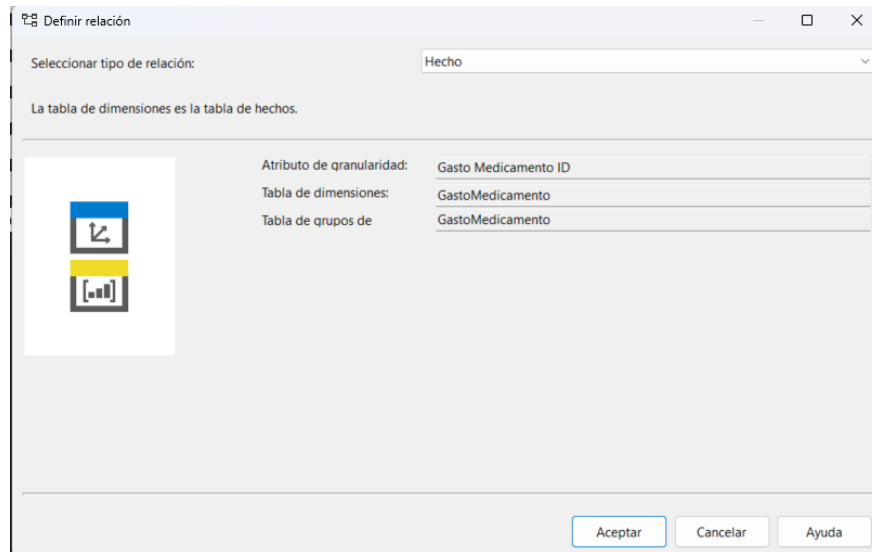


Figura 12: Relacion entre dimension Gasto y grupo de medidas Gasto de tipo Hecho

Además, configuramos relaciones de tipo **Varios a Varios** para dimensiones como **Alergia** y **Tratamiento**, con **Gasto de Medicamento**, como se muestra en las figuras 13 y 14.

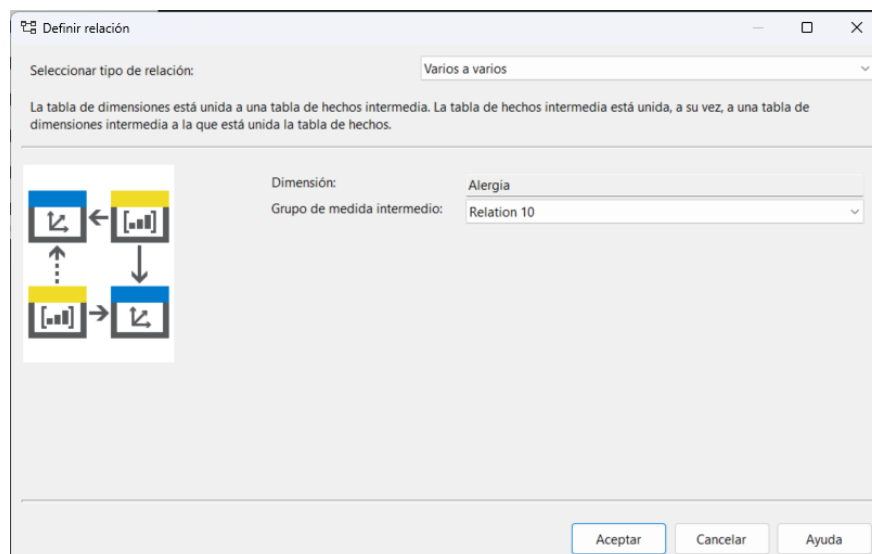


Figura 13: Configuración de relaciones de tipo Varios a Varios para Alergia.

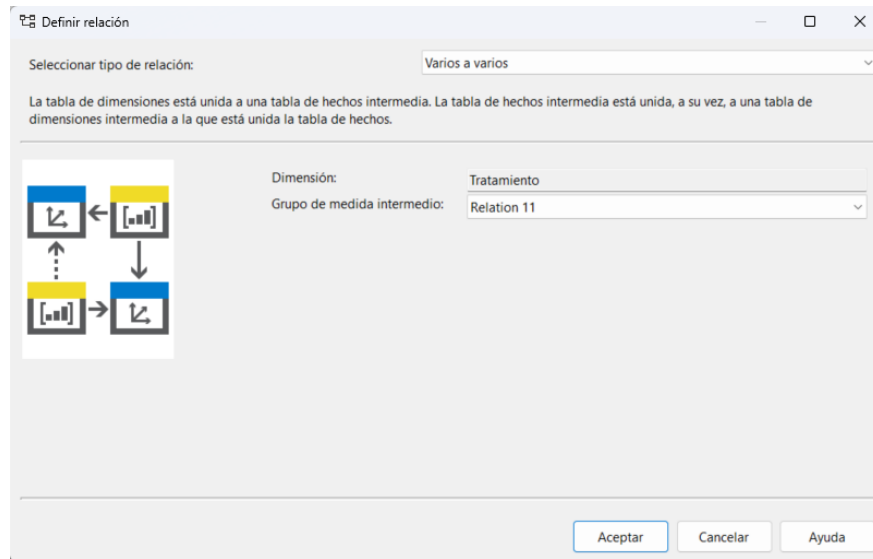


Figura 14: Configuración de relaciones de tipo Varios a Varios para Tratamiento.

Al final de este proceso el esquema de uso de dimensiones quedará como en la figura 15, donde las dimensiones alergia y tratamiento se relacionan con el gasto de medicamento de tipo varios a varios, pero para lograr esto las tablas intermedias(grupos de medidas intermedios) se tienen que relacionar con su dimension y el gasto.

Dimensiones	Grupos de medida		
	Gasto Medicamento	Relation 10	Relation 11
Gasto Medicamento	Gasto Medicamento ID	Gasto Medicamento ID	Gasto Medicamento ID
Paciente	Unique PID		
Ingreso UCI	Ingreso UCID		
Anyo	Anyo ID		
Hospital	Hospital ID		
Medicamento	Medicamento Id		
Alergia	Relation 10	Alergia ID	
Tratamiento	Relation 11		Tratamiento ID

Figura 15: Relaciones de dimensiones y grupos de medida del cubo.

### 3.6. Cálculo de Medidas

La medida **Tasa de Infusión**, se deben calcular utilizando la media, el resto de medidas (cantidad, recuento de gasto y volumen de fluido) son aditivas.

### 3.7. Jerarquías y Atributos

Para cada tabla, organizamos las jerarquías con los IDs correspondientes, configuramos los **nameColumn** con identificadores claros y añadimos atributos adicionales que no se incluyan en las jerarquías.

- **Paciente:** Configuramos la jerarquía con el ID de paciente, renombrado como *Paciente*, y añadimos atributos como género y etnia. La figura 16 muestra esta configuración.
- **Ingreso:** Configuramos la jerarquía con los IDs de ingreso en UCI y hospital, y añadimos atributos como fuente de admisión, estado de alta, tipo de unidad, altura y peso de admisión, como se muestra en la figura 17.
- **Año:** Configuramos el atributo **nameColumn** para el nombre del año, como se muestra en la figura 18.
- **Hospital:** Se configura la jerarquía con el ID de región y el ID de hospital, renombrando estos como *Región* y *Hospital*, respectivamente. La región también tendrá su nombre referenciado. Los atributos de esta jerarquía incluirán el número de camas y el nombre de la región. La configuración se muestra en la figura 19.
- **Medicamento:** Esta dimensión será simple, ya que solo tendrá el ID y el nombre del medicamento, el cual se incluirá en el atributo **nameColumn**. La figura 20 muestra esta configuración.
- **Alergia:** La jerarquía se basará en el ID de alergia, renombrado como *Alergia*, y el atributo **nameColumn** se usará para el nombre de la alergia. Además, se añadirán los atributos tipo de alergia y nombre de medicamento, en caso de que la alergia esté asociada a un medicamento. La configuración se muestra en la figura 21.
- **Tratamiento:** Se crea una jerarquía con el ID de tratamiento y la descripción del tratamiento, que se incluirá en el atributo **nameColumn** del ID. La figura 22 muestra esta configuración.

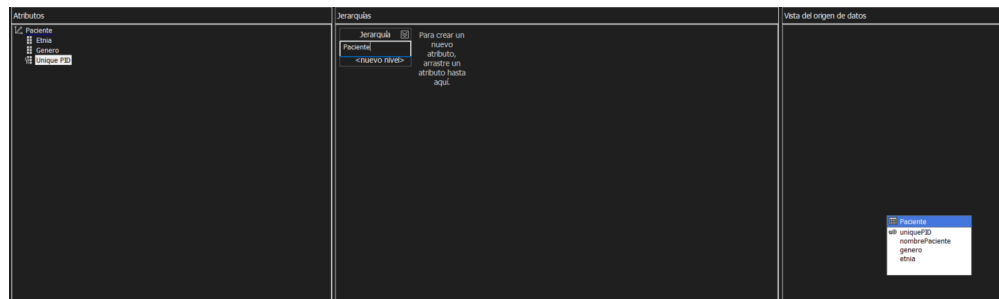


Figura 16: Jerarquías y atributos configurados para la tabla Paciente.

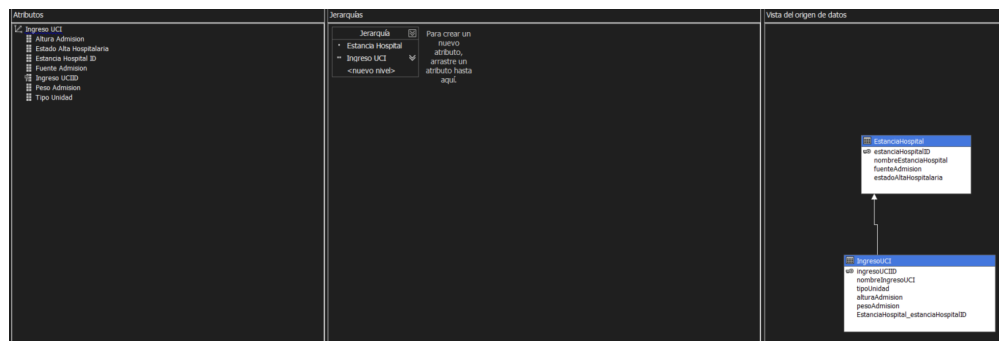


Figura 17: Jerarquías configuradas para la tabla Ingreso.

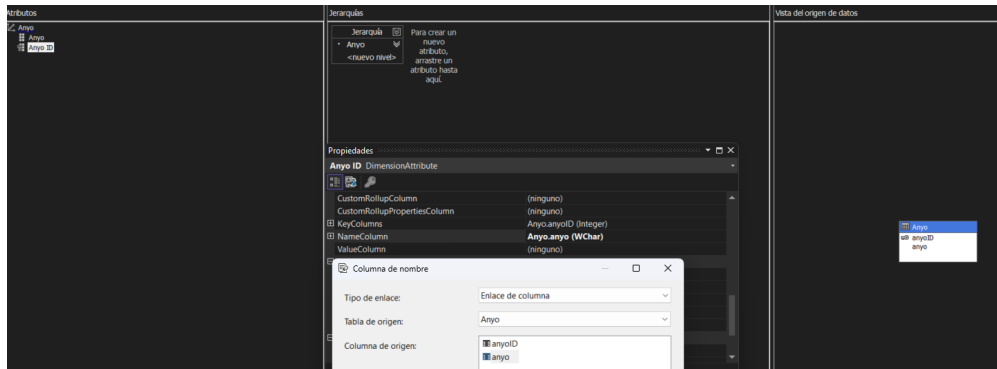


Figura 18: Configuración del atributo nameColumn para la tabla Año.

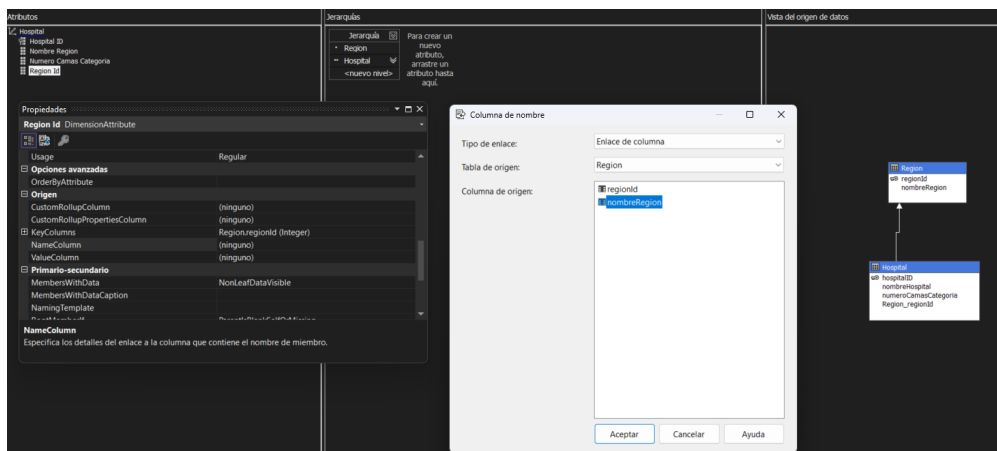


Figura 19: Jerarquías y atributos configurados para la tabla Hospital.

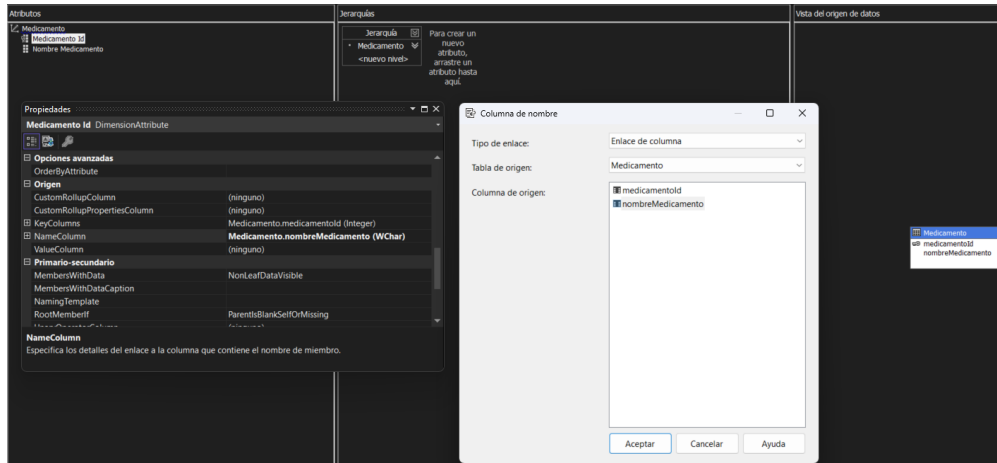


Figura 20: Configuración de la tabla Medicamento con el ID y el nombre del medicamento.

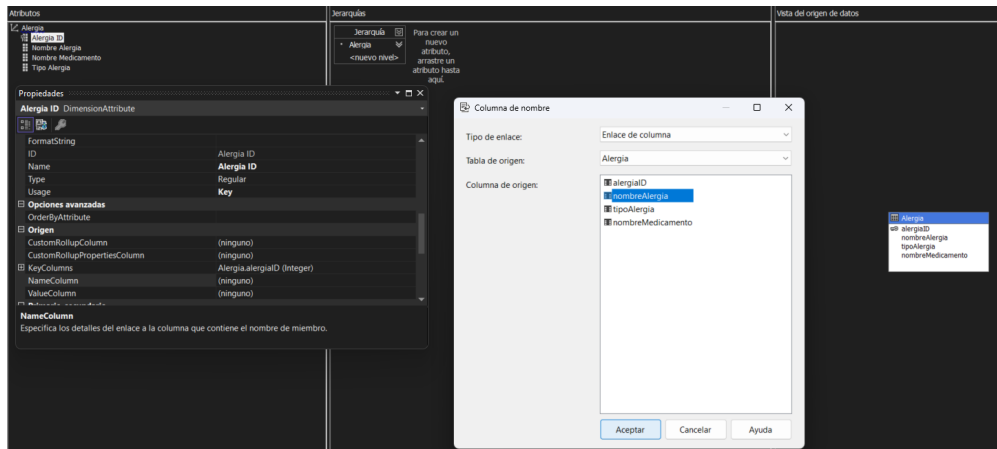


Figura 21: Jerarquías y atributos configurados para la tabla Alergia.

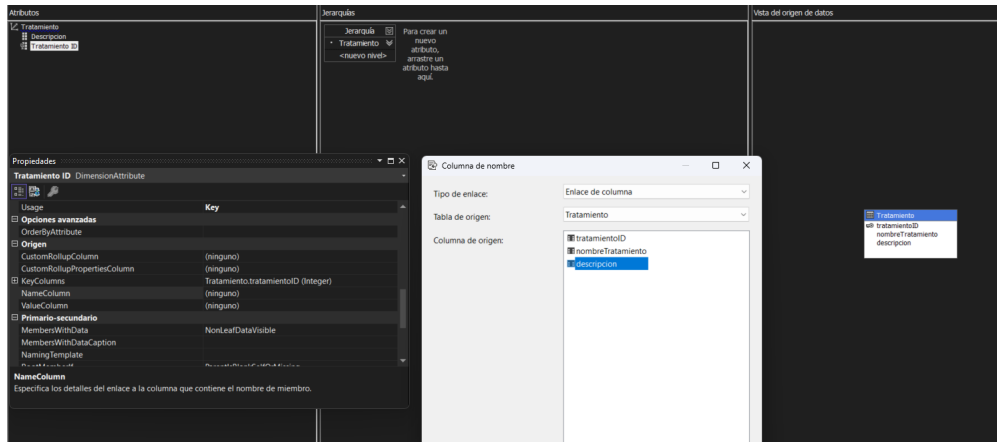


Figura 22: Jerarquías configuradas para la tabla Tratamiento.

### 3.8. Procesamiento del Cubo

Para procesar el cubo, haz clic derecho sobre él y selecciona la opción **Procesar**. En la figura 23 se muestra el procesamiento del cubo.

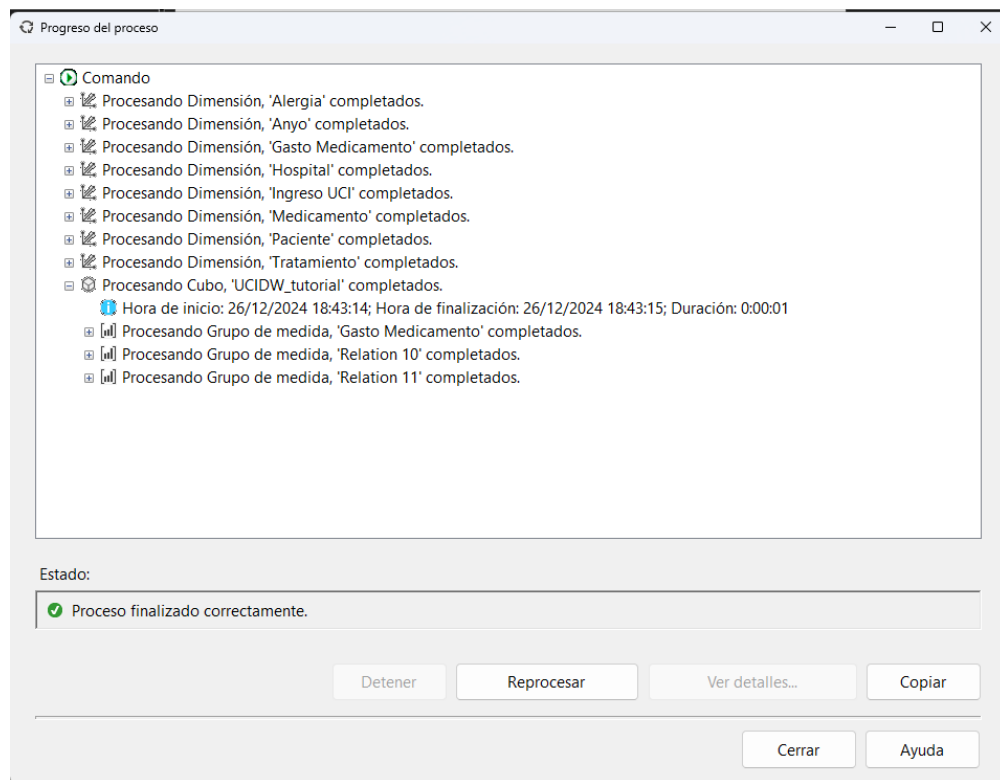


Figura 23: Procesamiento inicial del cubo multidimensional.

Para depurar errores, se recomiende el procesamiento de manera secuencial con transacciones indepen-

dientes.

### 3.9. Conexión al Cubo Procesado

Para acceder al cubo procesado, abre SQL Server, crea una conexión a Analysis Services en **localhost**, y navega hasta **Databases** para visualizar el cubo y realizar consultas MDX, como se muestra en la figura 24.

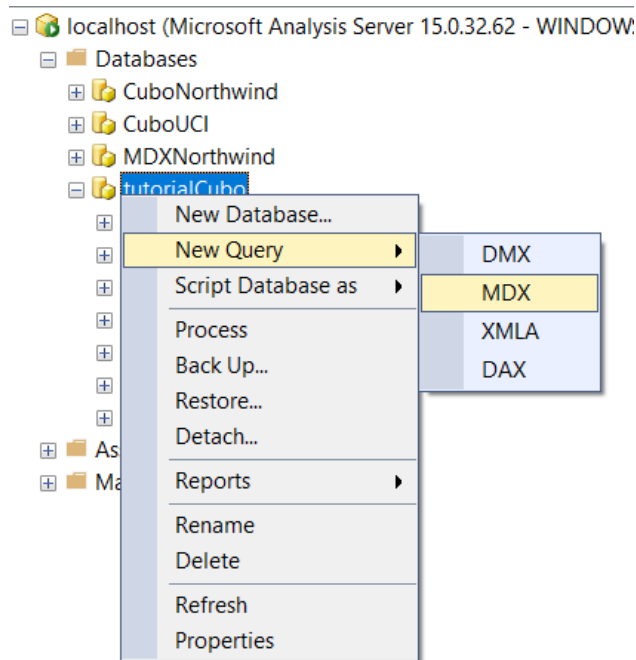


Figura 24: Visualización del cubo en SQL Server tras la conexión a Analysis Services.

### 3.10. Verificación de Jerarquías

Por último, verificamos las jerarquías creadas y navegamos por ellas para asegurarnos de que todo esté correcto, como se muestra en la figura 25.



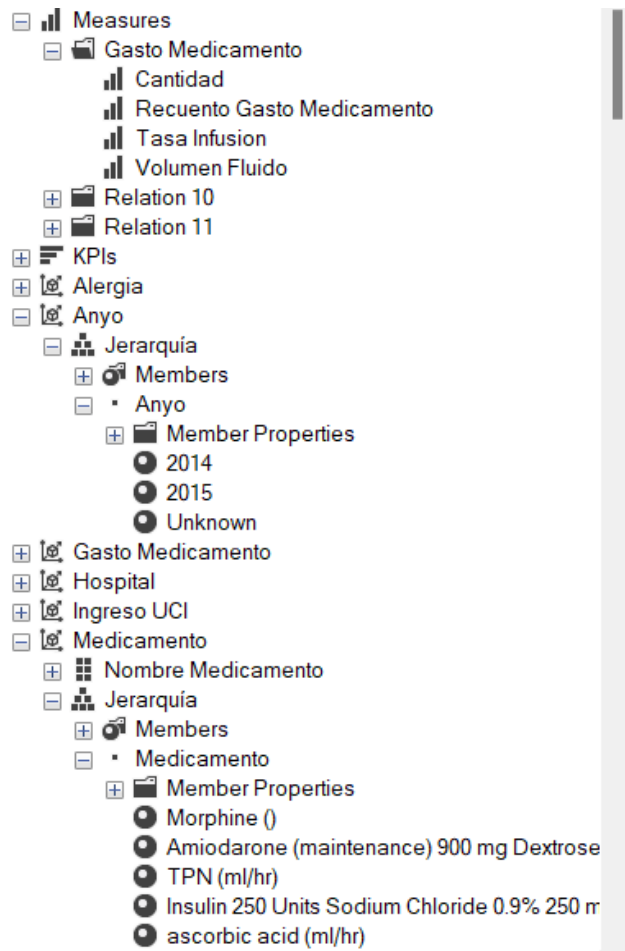


Figura 25: Verificación de jerarquías en el cubo multidimensional.

## 4. Consultas 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).

### 4.1. Consulta 1

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 4.2. Consulta 2

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 4.3. Consulta 3

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 4.4. Consulta 4

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 4.5. Consulta 5

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 4.6. Consulta 6

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 4.7. Consulta 7

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 4.8. Consulta 8

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 4.9. Consulta 9

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

#### 4.10. Consulta 10

**Explicación:** Aquí se describe el propósito y el enfoque de la consulta. Explica qué información se busca obtener y cómo se construye la consulta MDX.

**Resultado:** Incluye una breve descripción de los resultados obtenidos y, si es necesario, una tabla o gráfico para visualizarlos.

### 5. Tutorial ejecutar consultas

- Una sección con las instrucciones detalladas para que un evaluador pueda ejecutar las consultas en su máquina

### 6. Dificultades Encontradas

- 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.

### 7. Conclusión

Gracias a la correcta implementación del proceso ETL, se ha logrado organizar el almacén de datos de manera eficiente, optimizando la consulta de información relevante para la toma de decisiones. Este proceso ha permitido integrar y transformar datos provenientes de [1]. Además, la estructura obtenida no solo asegura la calidad y consistencia de los datos, sino que también sienta las bases para futuras ampliaciones o análisis más complejos, promoviendo la escalabilidad y adaptabilidad del sistema.

En conclusión, este proyecto ETL aporta un modelo sólido y adaptable para la gestión y análisis de datos en el ámbito hospitalario, contribuyendo a una administración más eficiente de los recursos y a una mejora potencial en la atención a los pacientes.

## 8. Acceso al Repositorio

Toda la información adicional, incluyendo el código fuente y la documentación completa de este proyecto, está disponible en el repositorio de GitHub [2].

## Referencias

- [1] MIT Laboratory for Computational Physiology. eICU Collaborative Research Database. <https://eicu-crd.mit.edu/>, 2020. Último acceso: 8 noviembre 2024.
- [2] Alex Silva. Healthcaredatawarehouse. <https://github.com/AlexSilvaa9/HealthcareDataWarehouse>, 2024. Último acceso: 1 octubre 2024.