

Análisis de datos (MDX)



Alejandro Silva Rodríguez

Marta Cuevas Rodríguez

Almacenes De Datos
Universidad de Málaga

Diciembre 2024

Índice

1. Introducción	2
2. Objetivos	2
3. Creación del Cubo Multidimensional	2
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
4. Consultas MDX	16
4.1. Consulta 1	16
4.2. Consulta 2	18
4.3. Consulta 3	19
4.4. Consulta 4	20
4.5. Consulta 5	21
4.6. Consulta 6	22
4.7. Consulta 7	23
4.8. Consulta 8	24
4.9. Consulta 9	26
4.10. Consulta 10	26
5. Tutorial ejecutar consultas	27
6. Dificultades Encontradas	27
7. Conclusión	28
8. Acceso al Repositorio	29

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.

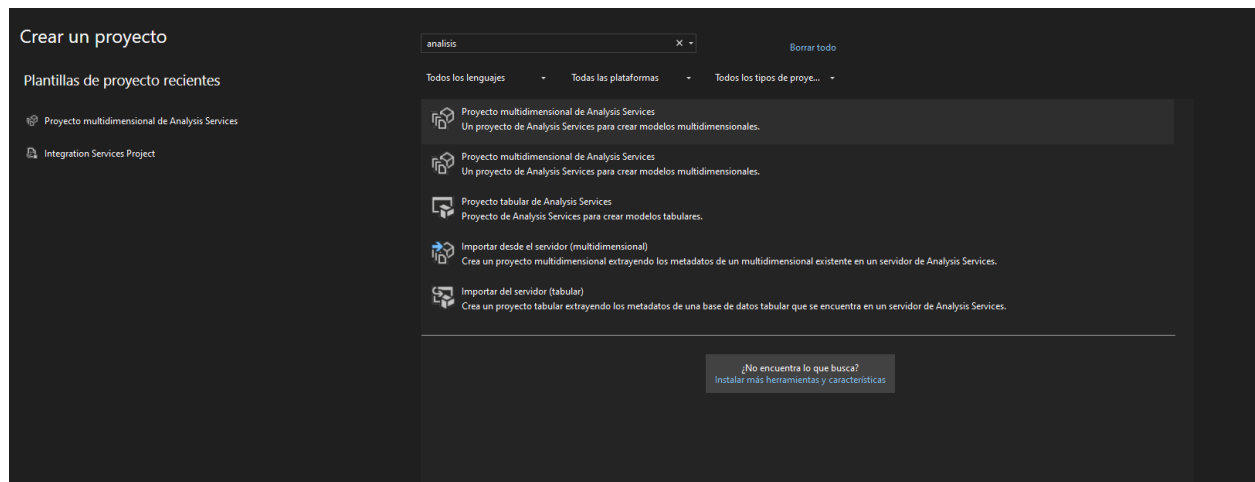


Figura 1: Creación de un nuevo proyecto multidimensional en Visual Studio.

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.

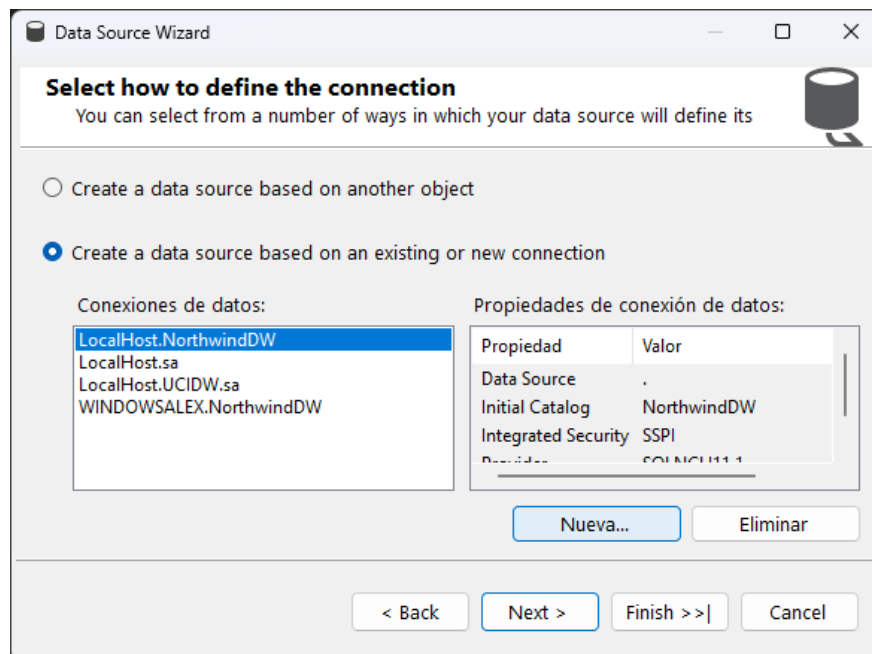


Figura 2: Nueva conexión en Visual Studio.

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.

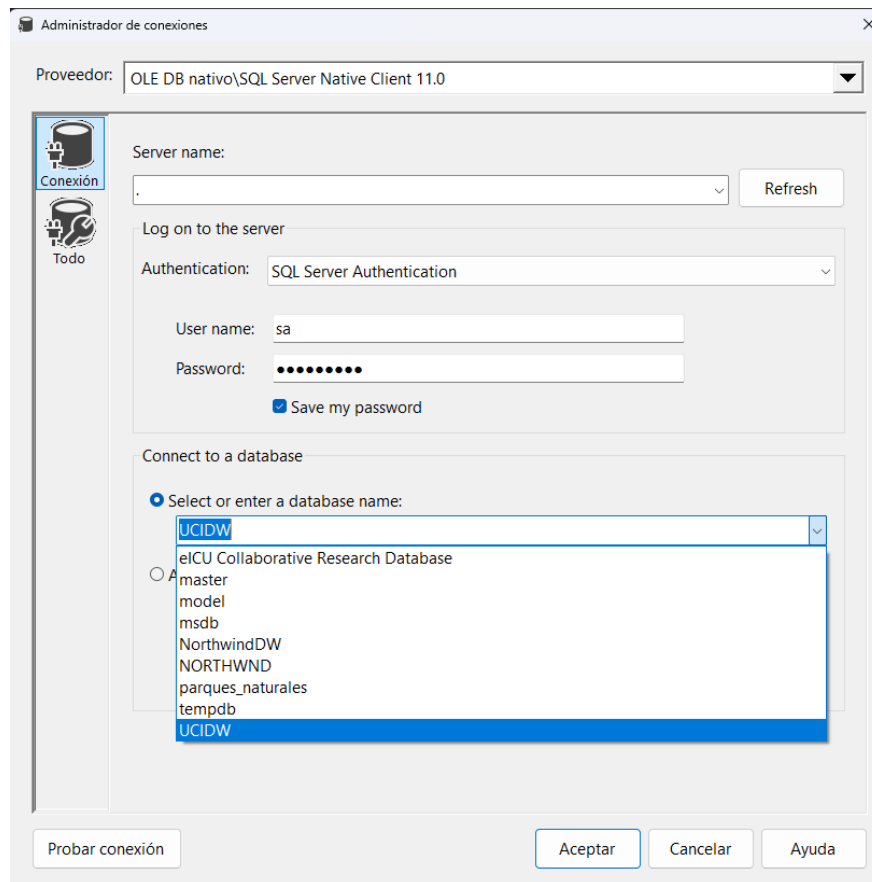


Figura 3: Configuración de la autenticación SQL Server.

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.

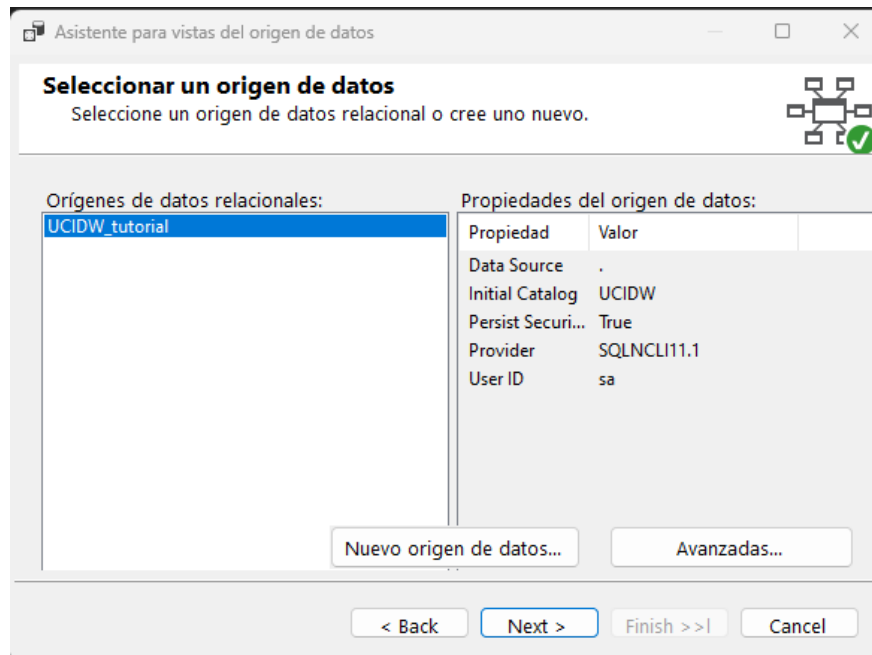


Figura 4: Selección del origen de datos configurado previamente.

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

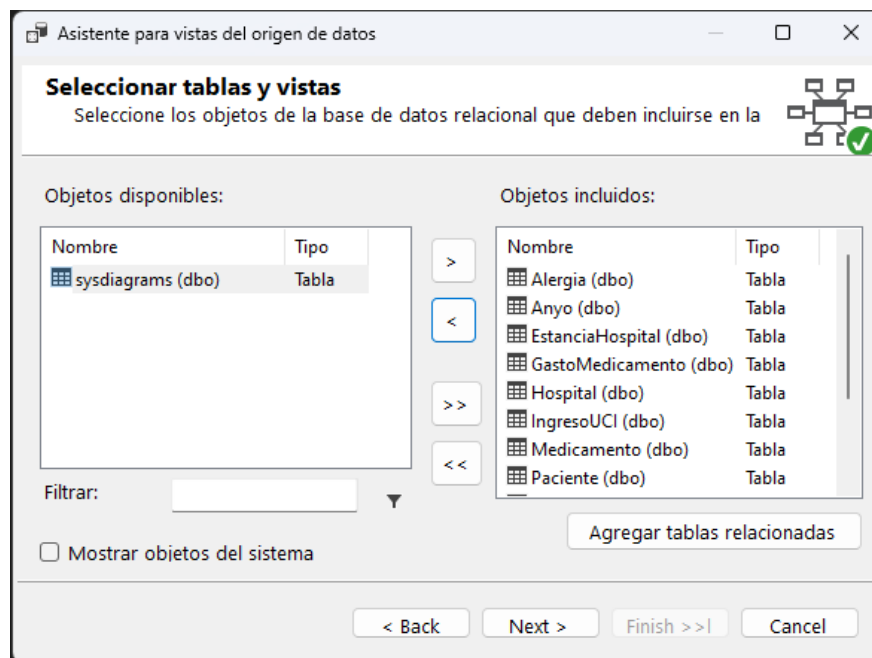


Figura 5: Selección de tablas del origen de datos.

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

radas.

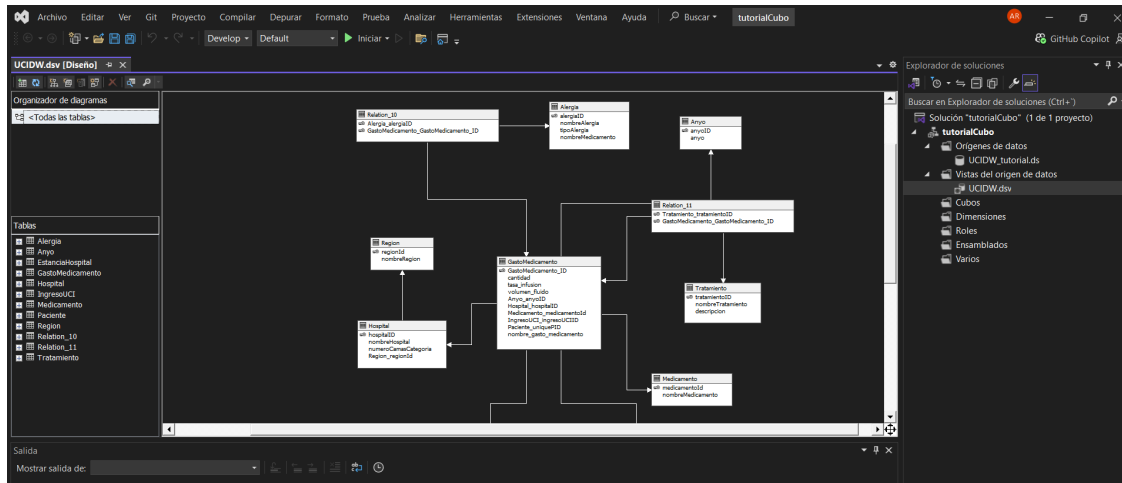


Figura 6: Verificación de la vista de la base de datos.

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.

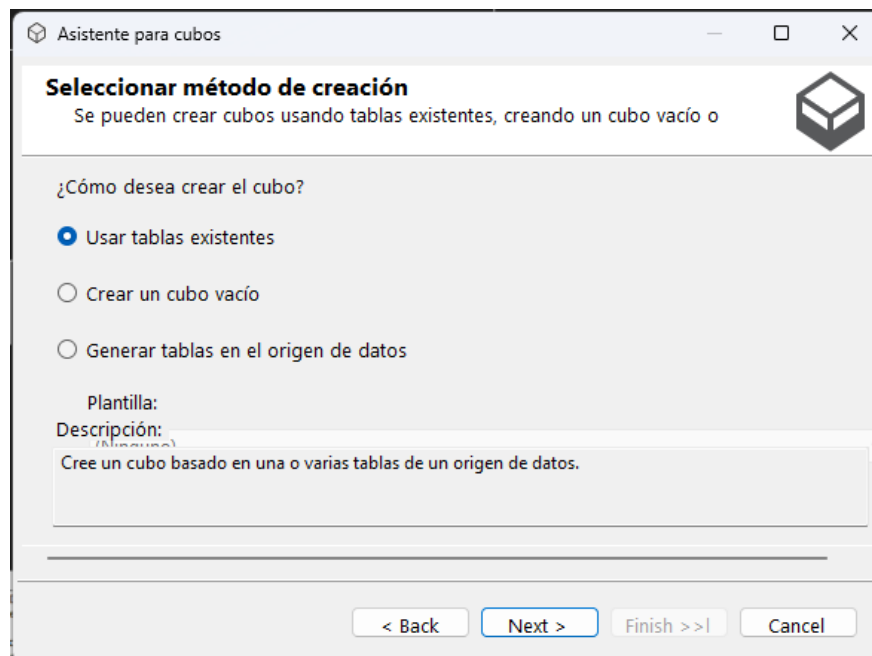


Figura 7: Creación del cubo multidimensional con las tablas seleccionadas.

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.

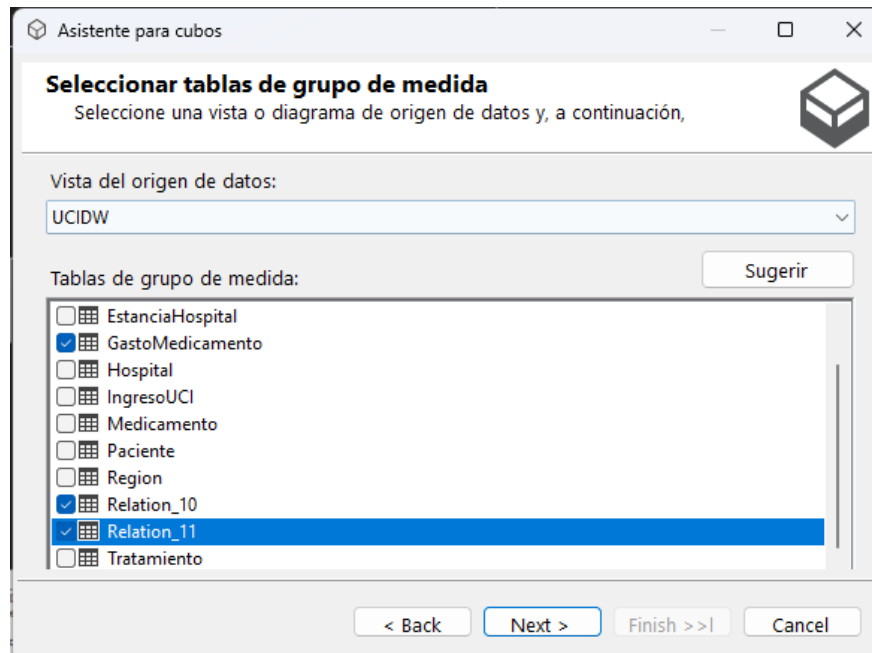


Figura 8: Selección de tablas del grupo de medida y relaciones intermedias.

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

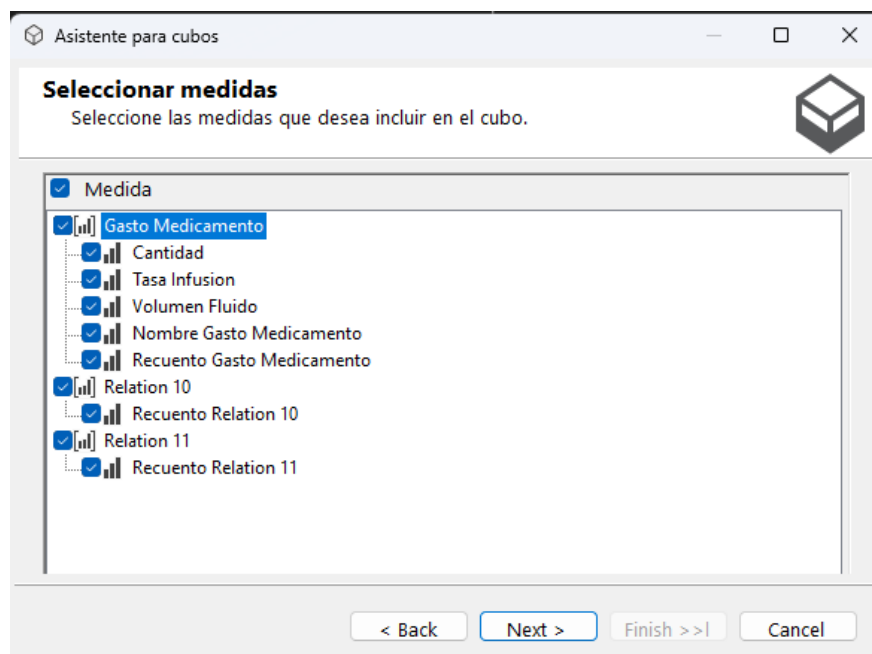


Figura 9: Configuración de medidas.

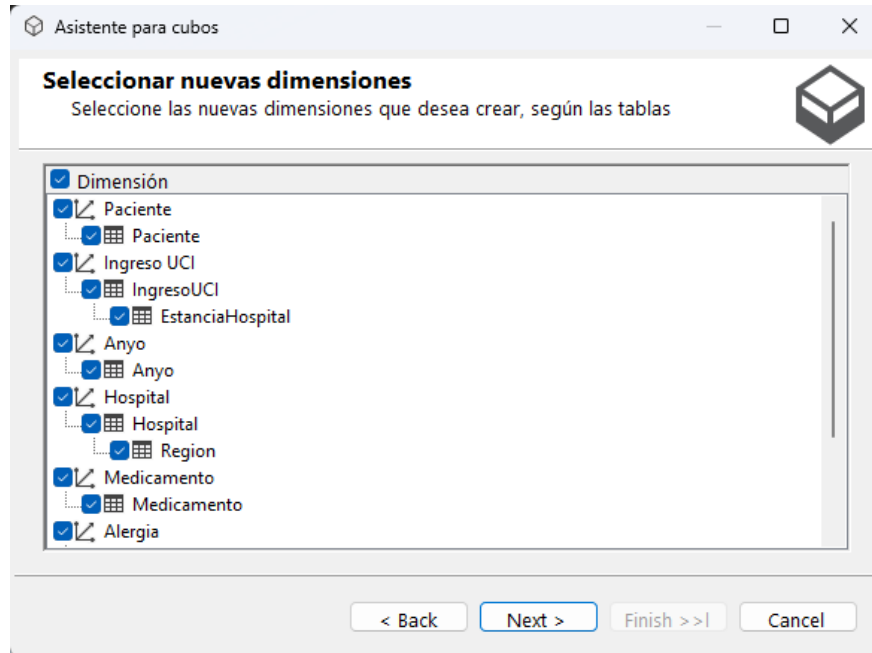


Figura 10: Configuración dimensiones en el cubo.

Finalmente, visualizamos el cubo creado en la figura 11.

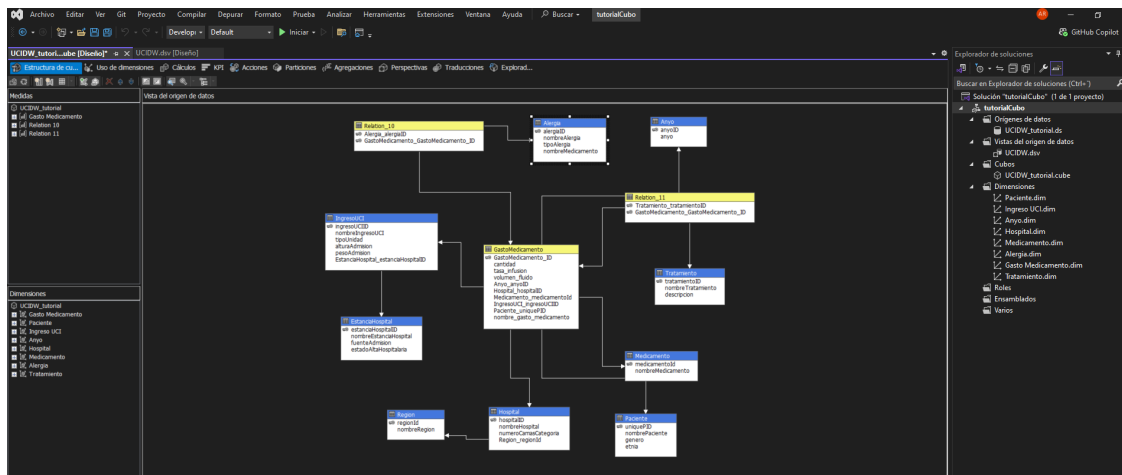


Figura 11: Visualización del cubo.

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.

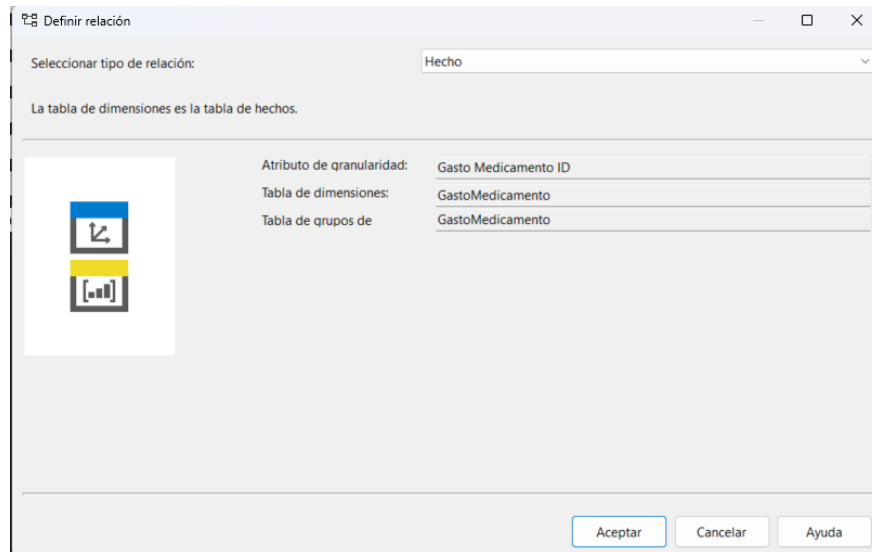


Figura 12: Relacion entre dimension Gasto y grupo de medidas Gasto de tipo Hecho

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.

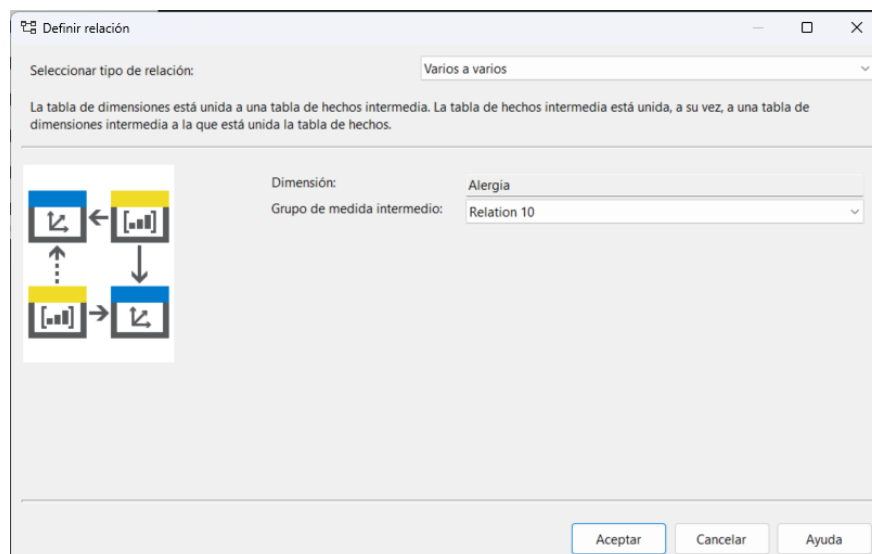


Figura 13: Configuración de relaciones de tipo Varios a Varios para Alergia.

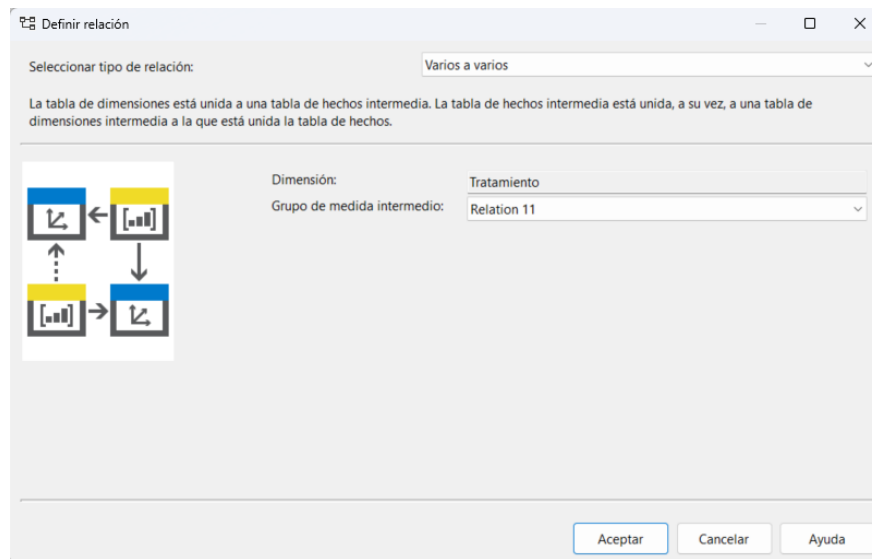


Figura 14: Configuración de relaciones de tipo Varios a Varios para Tratamiento.

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.

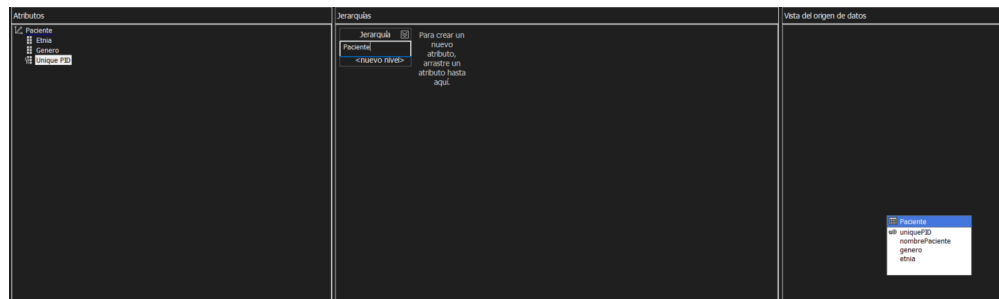


Figura 16: Jerarquías y atributos configurados para la tabla Paciente.

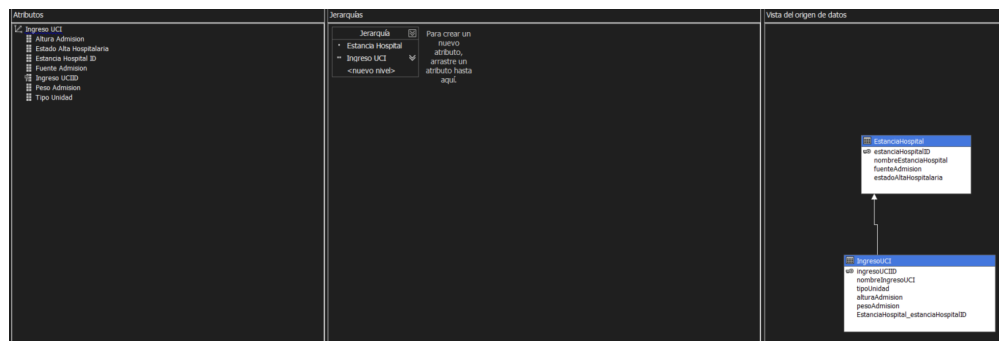


Figura 17: Jerarquías configuradas para la tabla Ingreso.

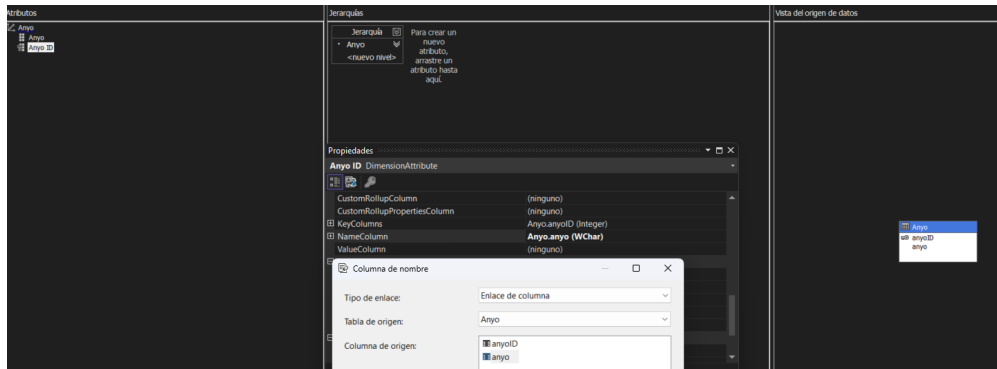


Figura 18: Configuración del atributo nameColumn para la tabla Año.

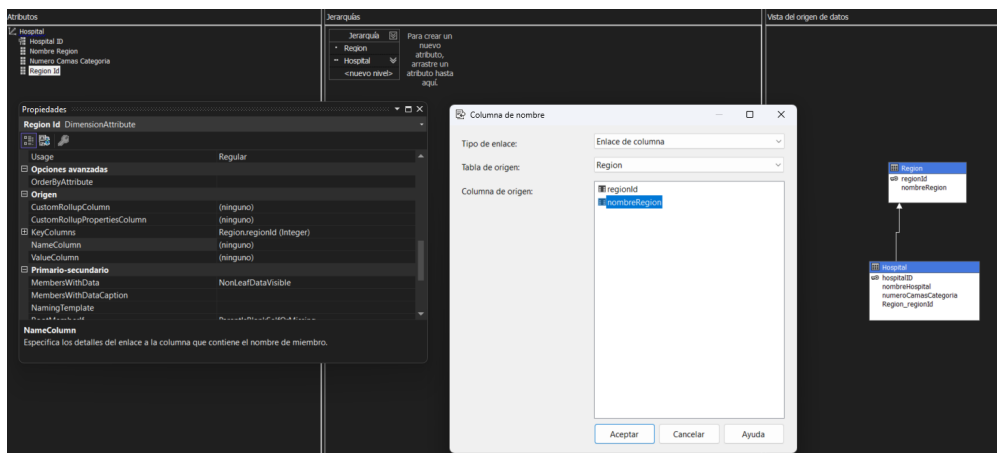


Figura 19: Jerarquías y atributos configurados para la tabla Hospital.

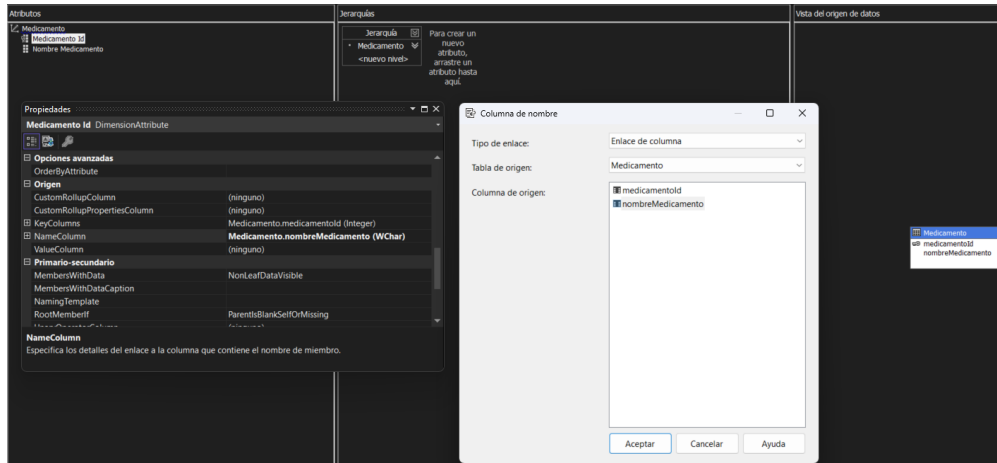


Figura 20: Configuración de la tabla Medicamento con el ID y el nombre del medicamento.

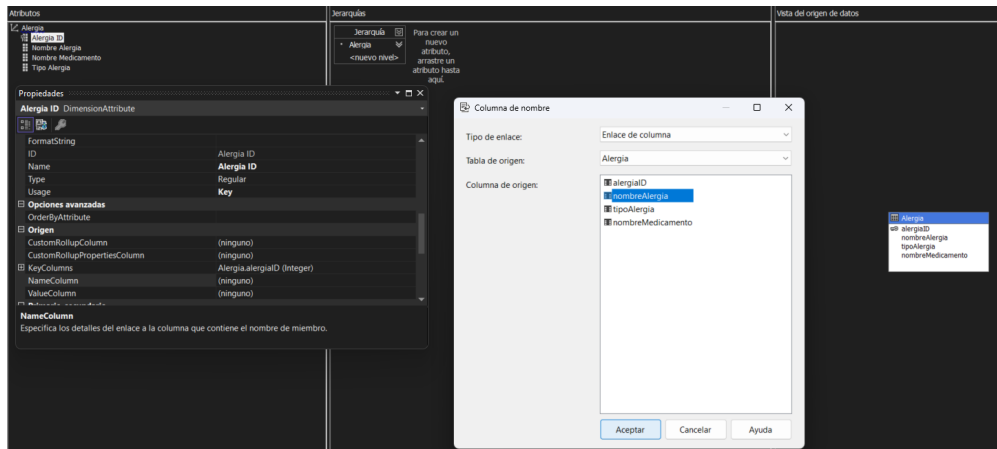


Figura 21: Jerarquías y atributos configurados para la tabla Alergia.

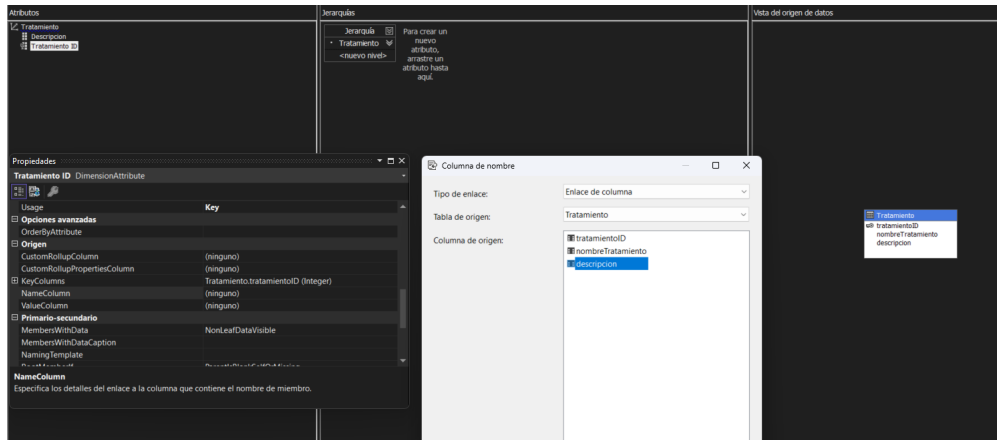


Figura 22: Jerarquías configuradas para la tabla Tratamiento.

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.

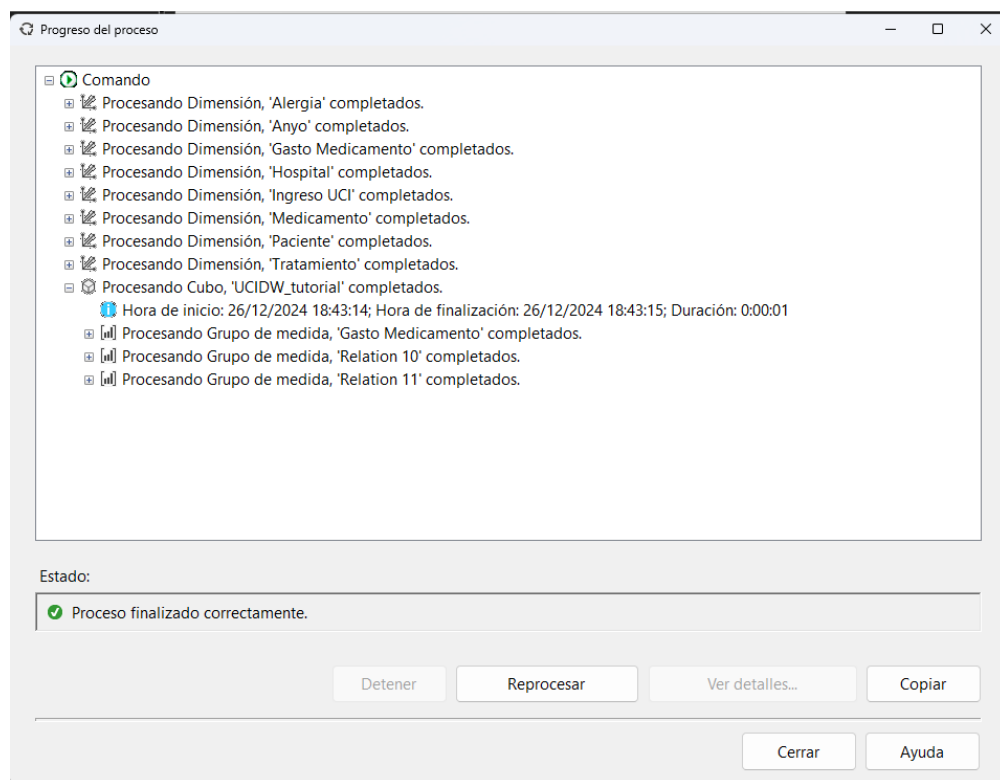


Figura 23: Procesamiento inicial del cubo multidimensional.

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.

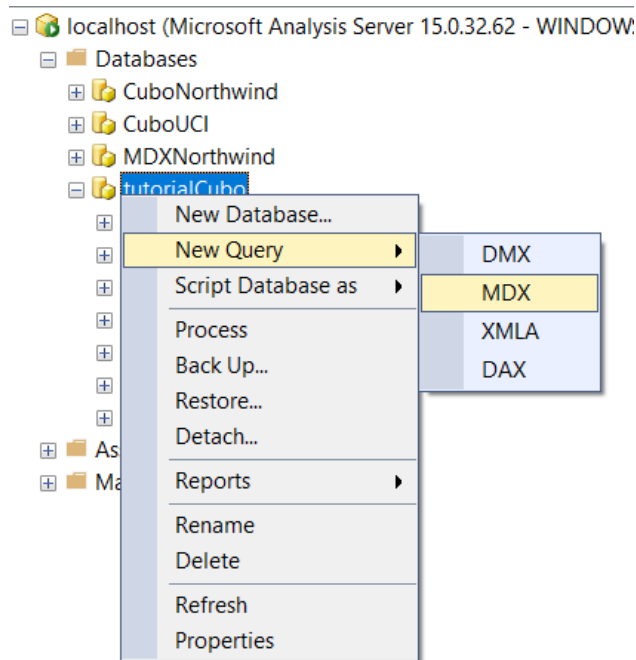


Figura 24: Visualización del cubo en SQL Server tras la conexión a Analysis Services.

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.

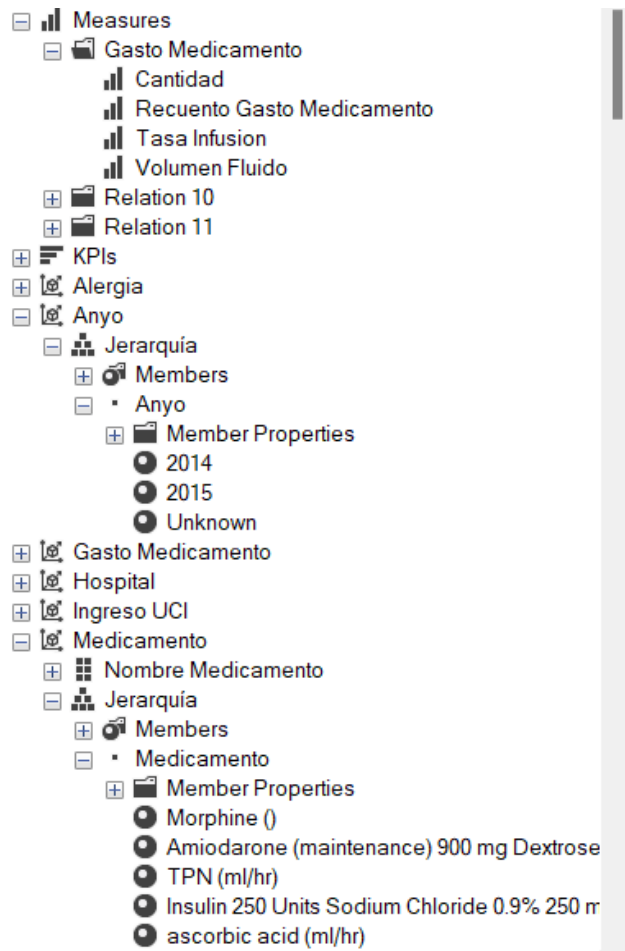


Figura 25: Verificación de jerarquías en el cubo multidimensional.

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

Ranking de Hospitales según Volumen de Fluido

Esta consulta en MDX genera un ranking de hospitales basado en el volumen de fluido, utilizando la base de datos multidimensional [UCIDW]. Primero, define una métrica calculada, [Rank Volumen Fluido], que asigna un rango a cada hospital según el volumen de fluido (medida [Measures].[Volumen Fluido]) en orden descendente (BDESC). Luego, en la selección (SELECT), construye dos ejes: en las filas (ON ROWS), se muestran los hospitales que tienen un valor no vacío de volumen de fluido, ordenados ascendentemente por el ranking recién calculado (BASC). En las columnas (ON COLUMNS), se incluyen las medidas [Volumen Fluido] y [Rank Volumen Fluido].

```

1
2 WITH
3     MEMBER [Measures].[Rank Volumen Fluido] AS
4         RANK([Hospital].[Hospital].CURRENTMEMBER,
5             ORDER([Hospital].[Hospital].[Hospital].MEMBERS,
6                 [Measures].[Volumen Fluido], BDESC))
7
8 SELECT
9     NON EMPTY
10        ORDER(
11            FILTER(
12                [Hospital].[Hospital].[Hospital].MEMBERS,
13                NOT ISEMPTY([Measures].[Volumen Fluido])
14            ),
15            [Measures].[Rank Volumen Fluido],
16            BASC
17        ) ON ROWS,
18        {[Measures].[Volumen Fluido], [Measures].[Rank Volumen Fluido]} ON
19        COLUMNS
20 FROM
    [UCIDW]

```

Listing 1: Consulta 1: Ranking de Hospitales según Volumen de Fluido

Messages		Results	
	Volumen Fluido	Rank Volumen Fluido	
142	537450	1	
425	256300	2	
420	213150	3	
141	190450	4	
144	107020	5	
394	85914	6	
194	75075	7	
428	72050	8	
444	70265	9	
188	61310	10	
419	56290	11	
301	54540	12	
436	50360	13	
422	45475	14	
273	29600	15	
281	28900	16	

Figura 26: Resultados de la consulta 1: Ranking de Hospitales según Volumen de Fluido

Como se muestra en la Figura 26, los resultados de la consulta presentan dos columnas principales: la primera indica el volumen de fluido correspondiente a cada hospital, mientras que la segunda muestra su posición en el ranking. Este ranking ordena los hospitales de manera descendente, desde el que registra el mayor gasto de fluido hasta el que registra el menor.

4.2. Consulta 2

Tres medicamentos más gastados por cada unidad de UCI

Esta consulta en MDX identifica los tres medicamentos más utilizados en cada tipo de unidad de UCI. Primero, se define la métrica calculada [Rank Medicamento], que asigna un ranking a cada medicamento dentro de su categoría según la cantidad utilizada ([Measures].[Cantidad]) en orden descendente (BDESC). En la selección (SELECT), se configuran dos ejes: en las columnas (ON COLUMNS), se incluye la medida [Cantidad]; y en las filas (ON ROWS), se genera una lista utilizando TOPCOUNT para seleccionar los tres medicamentos con mayor cantidad en cada tipo de unidad de UCI. Para lograr esto, se emplea NONEMPTYCROSSJOIN para cruzar los miembros actuales de las unidades de UCI con los medicamentos disponibles, y NONEMPTY para excluir combinaciones sin datos.

Además para los medicamentos, se utilizó .children en lugar de .MEMBERS para evitar que apareciera el conjunto completo (ALL) como primera posición del ranking. Sin embargo, si se dejó en el tipo de UCI para observar cuales son los medicamentos más utilizados en general.

```
1 WITH
2 MEMBER [Measures].[Rank Medicamento] AS
3 RANK (
4     [Medicamento].[Nombre Medicamento].CURRENTMEMBER,
5     ORDER (
6         [Medicamento].[Nombre Medicamento].children,
7         [Measures].[Cantidad],
8         BDESC
9     )
10 )
11
12 SELECT
13     {[Measures].[Cantidad]} ON COLUMNS,
14     NONEMPTY (
15         GENERATE (
16             [Ingreso UCI].[Tipo Unidad].members,
17             TOPCOUNT (
18                 NONEMPTYCROSSJOIN (
19                     {[Ingreso UCI].[Tipo Unidad].
20                     CURRENTMEMBER},
21                     [Medicamento].[Nombre Medicamento]
22                     ].children
23                 ),
24                 3,
25                 [Measures].[Cantidad]
26             )
27         ) ON ROWS
28 FROM
29     [UCIDW]
```

Listing 2: Consulta 2: Tres medicamentos más gastados por cada unidad de UCI

En la Figura 27 se presenta el resultado esperado de la consulta. Este muestra dos columnas principales: la primera enumera los tipos de unidades de UCI, mientras que la segunda identifica los medicamentos más utilizados en cada tipo específico de unidad. Además, junto a cada medicamento, se muestra la cantidad correspondiente, lo que permite un análisis detallado de su consumo en cada tipo de UCI.

Messages		Results
		Cantidad
All	Heparin (units/hr)	8200000
All	Propofol (mg/kg/min)	8021500
All	Propofol (mcg/kg/min)	1151854
Cardiac ICU	Heparin (units/hr)	1050000
Cardiac ICU	Amiodarone (mg/min)	45000
Cardiac ICU	Propofol (mcg/kg/min)	24000
CCU-CTICU	Heparin (units/hr)	500000
CCU-CTICU	Dobutamine (mcg/kg/min)	41000
CCU-CTICU	Midazolam (mg/hr)	9300
CTICU	Propofol (mcg/kg/min)	110000
CTICU	Insulin (units/hr)	21700
CTICU	Norepinephrine (mcg/min)	3184
Med-Surg ICU	Propofol (mg/kg/min)	8017500
Med-Surg ICU	Heparin (units/hr)	4000000
Med-Surg ICU	Propofol (mcg/kg/min)	778514
MICU	Heparin (units/hr)	2625000

Figura 27: Resultados de la consulta 2: Ranking de Hospitales según Volumen de Fluido

4.3. Consulta 3

Cantidad de cada medicamento usado para hombres y para mujeres

Esta consulta en MDX calcula y muestra la cantidad de cada medicamento utilizado, desglosado por género (hombres y mujeres), utilizando nuevamente la base de datos multidimensional [UCIDW]. Define dos medidas calculadas: [Cantidad Hombres] y [Cantidad Mujeres], que calculan la suma total de la cantidad de medicamentos consumidos por pacientes masculinos ([Paciente].[Genero].[Male]) y femeninos ([Paciente].[Genero].[Female]), respectivamente, utilizando la función SUM. En la cláusula SELECT, se configuran dos ejes: en las filas (ON ROWS), se listan los medicamentos disponibles ([Medicamento].[Nombre Medicamento].MEMBERS); y en las columnas (ON COLUMNS), se muestran las dos medidas calculadas ([Cantidad Hombres] y [Cantidad Mujeres]).

```

1 WITH
2
3     MEMBER [Measures].[Cantidad Hombres] AS
4         SUM(
5             {[Paciente].[Genero].[Male]},
6             [Measures].[Cantidad]
7         )
8     MEMBER [Measures].[Cantidad Mujeres] AS
9         SUM(
10            {[Paciente].[Genero].[Female]},
11            [Measures].[Cantidad]
12        )
13 SELECT
14     NON EMPTY
15         {[Medicamento].[Nombre Medicamento].MEMBERS} ON ROWS,
16         {[Measures].[Cantidad Hombres], [Measures].[Cantidad Mujeres]} ON
17         COLUMNS
18 FROM
19     [UCIDW]

```

Listing 3: Consulta 3: Cantidad de cada medicamento usado para hombres y para mujeres

	Cantidad Hombres	Cantidad Mujeres
All	15705035	3322061
Abciximab (mcg/min)	54	(null)
alteplase (mg/hr)	132	(null)
Amiodarone (mg/min)	126584	29700
Angiomax (mg/kg/hr)	750	(null)
Angiomax (mg/kg/min)	750	(null)
Bumetanide (mg/hr)	519	(null)
clevipine (mg/hr)	200	(null)
Dexmedetomidine (mcg/kg/hr)	13410	538
Dilaudid (mg/hr)	1860	(null)
Dilaudid PCA (mg/hr)	300	(null)
Diltiazem (mg/hr)	38056	16120
Dobutamine (mcg/kg/min)	64000	(null)
Dopamine (mcg/kg/min)	54000	11600
Epinephrine (mcg/min)	535	148
Epinephrine (mg/kg/min)	8	(null)

Figura 28: Resultados de la consulta 3: Cantidad de cada medicamento usado para hombres y para mujeres

Como resultado, obtenemos dos columnas, una correspondiente a los hombres y otra a las mujeres (Figura 28), que muestran la cantidad de cada medicamento utilizado por cada género.

4.4. Consulta 4

Pacientes con Mayor Uso de Medicamentos por Hospital (con mas de 500)

Esta consulta en MDX filtra los registros para mostrar únicamente aquellos en los que la cantidad de medicamento utilizada es superior a 500. En la cláusula **SELECT**, en las filas (**ON ROWS**), se combinan las jerarquías **[Hospital].[Hospital].[Hospital]** y los hijos de **[Paciente].[Unique PID].CHILDREN**, para evitar que se mostrara el total de pacientes como paciente con mayor uso de medicamentos. La función **FILTER** se utiliza para filtrar estas combinaciones, mostrando solo aquellas en las que la medida **[Measures].[Cantidad]** es mayor a 500. En las columnas (**ON COLUMNS**), se incluye la medida **[Measures].[Cantidad]**, que muestra la cantidad utilizada para cada combinación.

```

1  SELECT
2
3      NON EMPTY
4          FILTER (
5              [Hospital].[Hospital].[Hospital] * [Paciente].[Unique PID
6                  ].CHILDREN,
7                  [Measures].[Cantidad] > 500
8              ) ON ROWS,
9          {[Measures].[Cantidad]} ON COLUMNS
10 FROM
    [UCIDW]
```

Listing 4: Consulta 4: Pacientes con Mayor Uso de Medicamentos por Hospital (con mas de 500)

Messages		Results
		Cantidad
79	003-10024	50000
92	003-10438	101200
92	003-10535	8007226
93	003-14605	804
93	003-16531	4125
95	003-11756	25000
95	003-12237	125000
95	003-11948	25025
108	003-10606	25000
108	003-10799	750
108	003-10003	10000
108	003-10799	750
108	003-10799	750
108	003-10003	10000
256	017-101737	564
256	017-101737	564

Figura 29: Resultados de la consulta 4: Pacientes con Mayor Uso de Medicamentos por Hospital (con mas de 500)

El resultado es una tabla (Figura 29) que presenta solo aquellos hospitales y pacientes donde la cantidad de medicamento excede los 500. Se muestran dos columnas de filas, la primera representa los hospitales y la segunda los pacientes.

4.5. Consulta 5

Tres tratamientos con mayor tasa por fuente de admisión

Esta consulta en MDX identifica los tres tratamientos con mayor tasa de infusión por cada fuente de admisión, es decir, por cada motivo de ingreso en la UCI. En la cláusula **SELECT**, en las columnas (**ON COLUMNS**), se incluye la medida **[Measures].[Tasa Infusion]**, que representa la tasa de infusión para los tratamientos. En las filas (**ON ROWS**), se utiliza la función **GENERATE** para iterar sobre los hijos de **[Ingreso UCI].[Fuente Admision].children**, (una vez más eligiendo **.children** para evitar que en el ranking influya el numero total de fuentes de admisión al igual que haremos con la descripción de los tratamientos) representando las diferentes fuentes de admisión. Para cada fuente de admisión, se aplica **TOPCOUNT** para seleccionar los tres tratamientos con las mayores tasas de infusión. Esto se logra mediante la función **NONEMPTYCROSSJOIN**, que combina las fuentes de admisión con los tratamientos disponibles y excluye combinaciones sin datos.

```

1 SELECT
2     {[Measures].[Tasa Infusion]} ON COLUMNS,
3     NONEMPTY(
4         GENERATE(
5             [Ingreso UCI].[Fuente Admision].children,
6             TOPCOUNT(
7                 NONEMPTYCROSSJOIN(
8
```

```

9      { [Ingreso UCI].[Fuente Admision].
10          CURRENTMEMBER},
11      [Tratamiento].[Descripcion].children
12      ),
13      3,
14      [Measures].[Tasa Infusion]
15      )
16      ) ON ROWS
17 FROM
18      [UCIDW]

```

Listing 5: Consulta 5: Tres tratamientos con mayor tasa por fuente de admisión

Messages Results		
		Tasa Infusion
	pulmonary consultations Pulmonary medici...	274133400.00%
	cardiovascular cardiac surgery CABG	271189400.00%
	cardiovascular consultations Cardiac surge...	271189400.00%
Acute Care/Floor	cardiovascular shock vasopressors phenyl...	27976800.00%
Acute Care/Floor	cardiovascular consultations Cardiology co...	27944800.00%
Acute Care/Floor	pulmonary radiologic procedures / broncho...	27944800.00%
Chest Pain Center	cardiovascular consultations Cardiology co...	80000.00%
Chest Pain Center	cardiovascular myocardial ischemia / infarct...	80000.00%
Chest Pain Center	cardiovascular myocardial ischemia / infarct...	80000.00%
Direct Admit	cardiovascular consultations Cardiology co...	2600000.00%
Direct Admit	cardiovascular arrhythmias antiarrhythmics ...	2525000.00%
Direct Admit	cardiovascular arrhythmias antiarrhythmics ...	2500000.00%
Emergency Department	pulmonary ventilation and oxygenation mec...	38997200.00%
Emergency Department	infectious diseases medications therapeuti...	31751000.00%
Emergency Department	cardiovascular shock vasopressors vasopr...	25819900.00%
Floor	pulmonary radiologic procedures / broncho...	801623900.00%

Figura 30: Resultados de la consulta 5: Tres tratamientos con mayor tasa por fuente de admisión

En la Figura 30 se presentan los resultados de la quinta consulta, organizados en dos columnas de datos principales. La primera columna muestra las diferentes fuentes de ingreso a la UCI junto con la descripción de los tratamientos aplicados, mientras que la segunda columna indica la tasa de infusión administrada para cada tratamiento con las tasas más altas. Finalmente, la columna de Tasa Infusión muestra los datos obtenidos de cada tratamiento por tipo de ingreso en UCI.

4.6. Consulta 6

Gasto medio de cantidad y volumen de fluido por numero de camas de hospitales

Esta consulta en MDX calcula el gasto medio de cantidad y volumen de fluido por categorías de número de camas en hospitales. Define dos medidas calculadas: [Media Cantidad] y [Media Volumen Fluido], que utilizan la función AVG para calcular el promedio de las medidas [Cantidad] y [Volumen Fluido]. El uso de EXISTING es crucial, ya que asegura que los cálculos se limiten a la búsqueda de los datos únicamente asociados a ese número de camas en concreto, evitando incluir datos del resto de categorías. En la cláusula SELECT, las medidas calculadas se muestran en las columnas (ON COLUMNS), mientras que las filas (ON ROWS) listan las categorías de número de camas mediante [Hospital].[Numero Camas Categoria].CHILDREN.

```

1 WITH
2     MEMBER [Measures].[Media Cantidad] AS
3     AVG(
4         EXISTING [Hospital].[Hospital].[Hospital].MEMBERS,
5         [Measures].[Cantidad]
6     )
7     MEMBER [Measures].[Media Volumen Fluido] AS
8     AVG(
9         EXISTING [Hospital].[Hospital].[Hospital].MEMBERS,
10        [Measures].[Volumen Fluido]
11    )
12
13 SELECT
14     {[Measures].[Media Cantidad],
15      [Measures].[Media Volumen
16      Fluido]} ON COLUMNS,
17     [Hospital].[Numero Camas Categoria].CHILDREN ON ROWS
18
19 FROM
20     [UCIDW]

```

Listing 6: Consulta 6: Gasto medio de cantidad y volumen de fluido por numero de camas de hospitales

	Media Cantidad	Media Volumen Fluido
	183721	12718.125
<100	15010	7034
>= 500	154592.5	137805
100 - 249	254074.166666667	45053.7777777778
250 - 499	1089290.18181818	40625.8181818182
Unknown	(null)	(null)

Figura 31: Resultados de la consulta 6: Gasto medio de cantidad y volumen de fluido por numero de camas de hospitales

El resultado (Figura 31) muestra la media tanto de cantidad de gasto de medicamento como de volumen de fluido para cada cantidad de número de camas de los hospitales.

4.7. Consulta 7

Gasto en medicamentos superior a 1000 por hospital, etnia, y año para pacientes masculinos en la UCI

En este caso, la consulta identifica los gastos en medicamentos superiores a 1000 para pacientes masculinos en la UCI, desglosados por hospital, etnia y año. Se define una medida calculada, [Gasto Filtrado], que utiliza la función IIF para verificar si el gasto en medicamentos ([Measures].[Cantidad]) supera 1000; si es así, devuelve el valor de la medida, y en caso contrario, devuelve NULL. En la cláusula SELECT, en las filas (ON ROWS), se ordena la combinación de los años ([Anyo].[Jerarquía].[Anyo]), hospitales ([Ingreso UCI].[Jerarquía].[Ingreso UCIDW]) y las etnias de los pacientes ([Paciente].[Etnia].children) en orden ascendente. En las columnas (ON COLUMNS), se incluye la medida [Gasto Filtrado]. Además, se aplica un filtro (WHERE) para limitar los resultados a pacientes masculinos ([Paciente].[Genero].[Male]).


```

1 WITH
2     MEMBER [Measures].[Gasto Filtrado] AS
3     IIF([Measures].[Cantidad] > 1000, [Measures].[Cantidad], NULL)
4
5
6 SELECT
7     NON EMPTY
8         ORDER (
9             {[Anyo].[Jerarquia].[Anyo].MEMBERS *
10              [Ingreso UCI].[Jerarquia].[Ingreso UCID].[MEMBERS]*
11              [Paciente].[Etnia].children,
12              [Anyo].[Anyo].CURRENTMEMBER,
13              ASC
14            ) ON ROWS,
15         {[Measures].[Gasto Filtrado]} ON COLUMNS
16 FROM
17     [UCIDW]
18 WHERE
19     ([Paciente].[Genero].[Male])

```

Listing 7: Consulta 7: Gasto en medicamentos superior a 1000 por hospital etnia y año para pacientes masculinos en la UCI

El resultado es una tabla que detalla, para cada combinación de hospital, etnia y año, los gastos en medicamentos superiores a 1000 para pacientes masculinos en la UCI como se muestra a continuación.

Messages		Results	
			Gasto Filtrado
2014	243097	Caucasian	25000
2014	261520	Caucasian	101200
2014	296927	Caucasian	2250
2014	426975	Hispanic	49854
2014	426976	Hispanic	198372
2014	441732	Hispanic	2024
2014	472811	Hispanic	2460
2014	496831	Hispanic	29220
2014	498848	Hispanic	1656812
2014	498849	Hispanic	1055082
2014	531662	Hispanic	7632
2014	533168	Hispanic	19500
2014	1059638	African American	22000
2014	1718412	Caucasian	1258
2014	1725601	Caucasian	29440
2014	1728151	Caucasian	50100

Figura 32: Resultados de la consulta 7: Gasto en medicamentos superior a 1000 por hospital, etnia, y año para pacientes masculinos en la UCI

4.8. Consulta 8

Volumen de fluido utilizado en el ingreso, clasificado por la altura y peso de los pacientes

La octava consulta calcula el volumen total de fluido utilizado para cada combinación de altura y peso de admisión en las unidades de UCI. Se define la medida calculada [Volumen Fluido Utilizado], que utiliza

la función SUM para sumar el volumen de fluido ([Measures].[Volumen Fluido]) a través de todos los tipos de unidad de UCI ([Ingreso UCI].[Tipo Unidad].members). En la cláusula SELECT, la medida [Volumen Fluido Utilizado] se coloca en las columnas (ON COLUMNS), mientras que las filas (ON ROWS) presentan las combinaciones de valores de altura y peso de admisión ([Ingreso UCI].[Altura Admision].children y [Ingreso UCI].[Peso Admision].children), generadas mediante NONEMPTYCROSSJOIN para excluir las combinaciones sin datos.

```

1 WITH
2     MEMBER [Measures].[Volumen Fluido Utilizado] AS
3     SUM(
4         [Ingreso UCI].[Tipo Unidad].members,
5         [Measures].[Volumen Fluido]
6     )
7
8 SELECT
9     {[Measures].[Volumen Fluido Utilizado]} ON COLUMNS,
10    NONEMPTYCROSSJOIN(
11        [Ingreso UCI].[Altura Admision].children,
12        [Ingreso UCI].[Peso Admision].children
13    ) ON ROWS
14 FROM
15    [UCIDW]

```

Listing 8: Consulta 8: Volumen de fluido utilizado en el ingreso clasificado por la altura y peso de los pacientes

Messages		Results
		Volumen Fluido Utilizado
142	100	200
152	55	2200
152	48	2400
152	68	21000
152	72	9000
152	58	6500
152	38	44000
154	64	53500
154	109	7400
154	64	8700
155	72	500
157	59	500
157	40	8400
157	55	11200
157	41	78000
157	58	4400

Figura 33: Resultados de la consulta 8: Volumen de fluido utilizado en el ingreso, clasificado por la altura y peso de los pacientes

En la Figura 33 podemos observar el resultado obtenido de la consulta con una columna que muestra el volumen de fluido clasificado en función de la altura (izquierda) y peso (derecha) de cada uno de los pacientes durante su ingreso.

4.9. Consulta 9

Medicamentos que causan alergias y su gasto total

Esta consulta en MDX identifica los medicamentos que causan alergias y muestra la cantidad utilizada de estos. En la cláusula **SELECT**, en las columnas (**ON COLUMNS**), se incluye la medida `[Measures].[Cantidad]`, que representa el gasto total asociado a cada medicamento. En las filas (**ON ROWS**), se aplica la función **FILTER** para seleccionar únicamente aquellos medicamentos (`[Medicamento].[Nombre Medicamento].children`) que coinciden con los registrados en la dimensión de alergias (`[Alergia].[Nombre Medicamento].CURRENTMEMBER`). Esto asegura que solo se incluyan los medicamentos asociados con reacciones alérgicas.

```
1 SELECT
2     {[Measures].[Cantidad]} ON COLUMNS,
3     FILTER(
4         [Medicamento].[Nombre Medicamento].children,
5         [Alergia].[Nombre Medicamento].CURRENTMEMBER
6     ) ON ROWS
7 FROM
8     [UCIDW]
```

Listing 9: Consulta 9: Medicamentos que causan alergias y su gasto total

Messages Results	
	Cantidad
Abciximab (mcg/min)	54
alteplase (mg/hr)	132
Amiodarone (mg/min)	156284
Angiomax (mg/kg/hr)	750
Angiomax (mg/kg/min)	750
Bumetanide (mg/hr)	519
clevipine (mg/hr)	200
Dexmedetomidine (mcg/kg/hr)	13948
Dilaudid (mg/hr)	1860
Dilaudid PCA (mg/hr)	300
Diltiazem (mg/hr)	54176
Dobutamine (mcg/kg/min)	64000
Dopamine (mcg/kg/min)	66400
Epinephrine (mcg/min)	683
Epinephrine (mg/kg/min)	8
Eptifibatide (mcg/kg/min)	750

Figura 34: Resultados de la consulta 9: Medicamentos que causan alergias y su gasto total

El resultado observado en la Figura 34 es una tabla que lista únicamente los medicamentos que causan alergias junto la cantidad utilizada de los mismos.

4.10. Consulta 10

Volumen de fluido de cada medicamento por región en 2014

La décima muestra el volumen de fluido utilizado para cada medicamento en cada región concretamente el año 2014. En la cláusula **SELECT**, en las columnas (**ON COLUMNS**), se incluye la medida `[Measures].[Volumen`

Fluido], que representa el volumen total de fluido. En las filas (ON ROWS), se genera una combinación de regiones ([Hospital].[Region].MEMBERS) y medicamentos ([Medicamento].[Nombre Medicamento].MEMBERS) mediante la función NONEMPTYCROSSJOIN, que asegura que solo se incluyan combinaciones con datos disponibles. La cláusula WHERE restringe los resultados al año 2014 ([Anyo].[Jerarquía].[Anyo].[2014]).

```

1 SELECT
2     {[Measures].[Volumen Fluido]} ON COLUMNS,
3     NONEMPTYCROSSJOIN(
4         [Hospital].[Region].MEMBERS,
5         [Medicamento].[Nombre Medicamento].MEMBERS
6     ) ON ROWS
7 FROM
8     [UCIDW]
9 WHERE
10     ([Anyo].[Jerarquía].[Anyo].[2014])
11

```

Listing 10: Consulta 10: Volumen de fluido de cada medicamento por region en 2014

Messages		Results
		Volumen Fluido
Midwest	All	211901
Midwest	alteplase (mg/hr)	6000
Midwest	Amiodarone (mg/min)	4750
Midwest	Dexmedetomidine (mcg/kg/hr)	7550
Midwest	Diltiazem (mg/hr)	4800
Midwest	Dobutamine (mcg/kg/min)	2750
Midwest	Dopamine (mcg/kg/min)	2250
Midwest	Eptifibatide (mcg/kg/min)	1000
Midwest	Esomeprazole (mg/hr)	12000
Midwest	Esomeprazole (mg/min)	1800
Midwest	Fentanyl (mcg/hr)	4020
Midwest	Fentanyl (mcg/kg/hr)	430
Midwest	Furosemide (mg/hr)	1400
Midwest	Heparin (units/hr)	18750
Midwest	Heparin (units/kg/hr)	3500
Midwest	Insulin (units/hr)	6850

Figura 35: Resultados de la consulta 10: Volumen de fluido de cada medicamento por región en 2014

Como resultado obtenemos la tabla de la figura 35 donde se muestran los resultados de volumen de fluido clasificado por región de cada medicamento

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.