

CREACIÓN DE CUBOS

Eric Gustavo Coronel Castillo

youtube.com/DesarrollaSoftware

www.desarrollasoftware.com

gcoronel@uni.edu.com

Logro Esperado

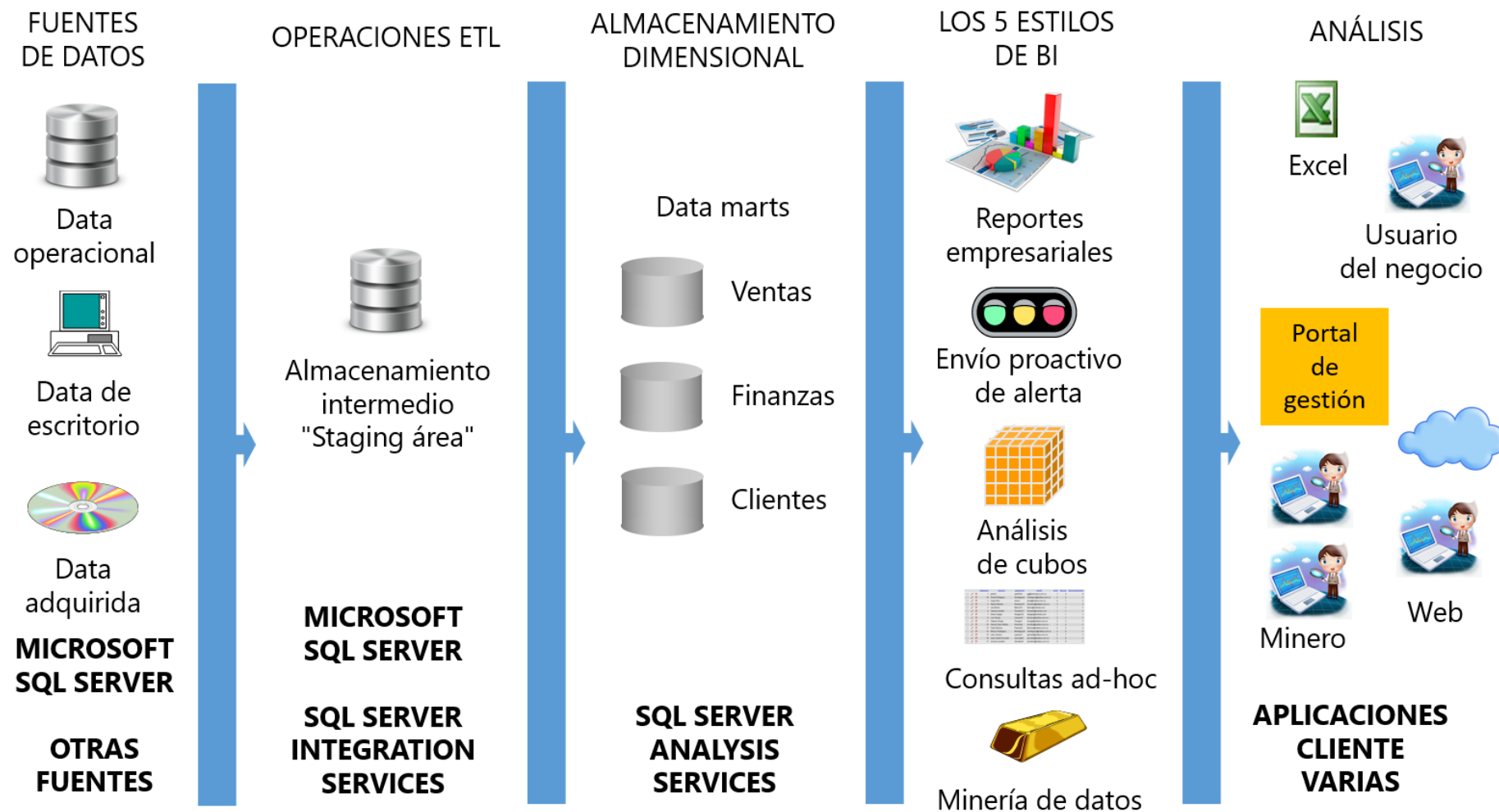
Al finalizar esta presentación, se espera que el participante aplique SSAS para crear CUBOS que le permitan analizar bases de datos dimensionales.



The background of the slide features a large, faint watermark of the University of Engineering and Technology (UNET) logo. The logo is circular and contains a profile of a person, a gear, and a book, with the text "UNIVERSIDAD NACIONAL DE INGENIERIA" and "LIMA PERU" around the perimeter.

ARQUITECTURA DE UNA SOLUCIÓN OLAP

ARQUITECTURA DE UNA SOLUCIÓN OLAP



ARQUITECTURA DE UNA SOLUCIÓN OLAP

- ❖ **Fuentes de datos:** la data operacional o transaccional puede provenir de fuentes variadas como base de datos de SQL Server, Access, Oracle, archivos planos u hojas de cálculo.
- ❖ **Operaciones ETL:** el almacenamiento intermedio ("staging área") lo podemos implementar en el motor de base de datos de SQL Server, y las operaciones ETL se programan en los paquetes de SQL Server Integration Services (SSIS).
- ❖ **Almacenamiento dimensional:** se implementa en SQL Server Analysis Services (SSAS), el motor OLAP de SQL Server mediante la creación de los cubos de datos.
- ❖ **Los estilos de BI:** hacen referencia a las diferentes formas de consolidación de la información para su análisis. El estilo depende del objetivo del análisis de información.
- ❖ **Análisis de la información:** se puede hacer uso de una herramienta de uso común como Excel, o herramientas más sofisticadas para generación de visualización de información como PowerBI, Tableau.

ALMACENAMIENTO MULTIDIMENSIONAL

| TRIMESTRE 1 | | CUZCO | LIMA | LORETO | TOTAL |
|-------------|----------|-------------|-------------|-------------|--------------|
| Ene-Mar | Peras | S/ 1,000.00 | S/ 575.00 | S/ 890.00 | S/ 2,465.00 |
| | Manzanas | S/ 3,000.00 | S/ 1,200.00 | S/ 700.00 | S/ 4,900.00 |
| | Melones | S/ - | S/ 125.00 | S/ 345.00 | S/ 470.00 |
| | Fresas | S/ 100.00 | S/ 1,250.00 | S/ 120.00 | S/ 1,470.00 |
| | Uvas | S/ 355.00 | S/ 2,345.00 | S/ 575.00 | S/ 3,275.00 |
| Total | Ene-Mar | S/ 4,455.00 | S/ 5,495.00 | S/ 2,630.00 | S/ 12,580.00 |
| ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... |
| TRIMESTRE 4 | | CUZCO | LIMA | LORETO | TOTAL |
| Oct-Dic | Peras | S/ 1,200.00 | S/ 465.00 | S/ 900.00 | S/ 2,565.00 |
| | Manzanas | S/ 2,500.00 | S/ 1,150.00 | S/ 500.00 | S/ 4,150.00 |
| | Melones | S/ 100.00 | S/ 200.00 | S/ 300.00 | S/ 600.00 |
| | Fresas | S/ 125.00 | S/ 1,300.00 | S/ 125.00 | S/ 1,550.00 |
| | Uvas | S/ 455.00 | S/ 2,100.00 | S/ 475.00 | S/ 3,030.00 |
| Total | Oct-Dic | S/ 4,380.00 | S/ 5,215.00 | S/ 2,300.00 | S/ 11,895.00 |

En la data dimensional que se muestra se identifican 3 dimensiones: Tiempo (trimestres), Ubicación (ciudades), y Producto (frutas).

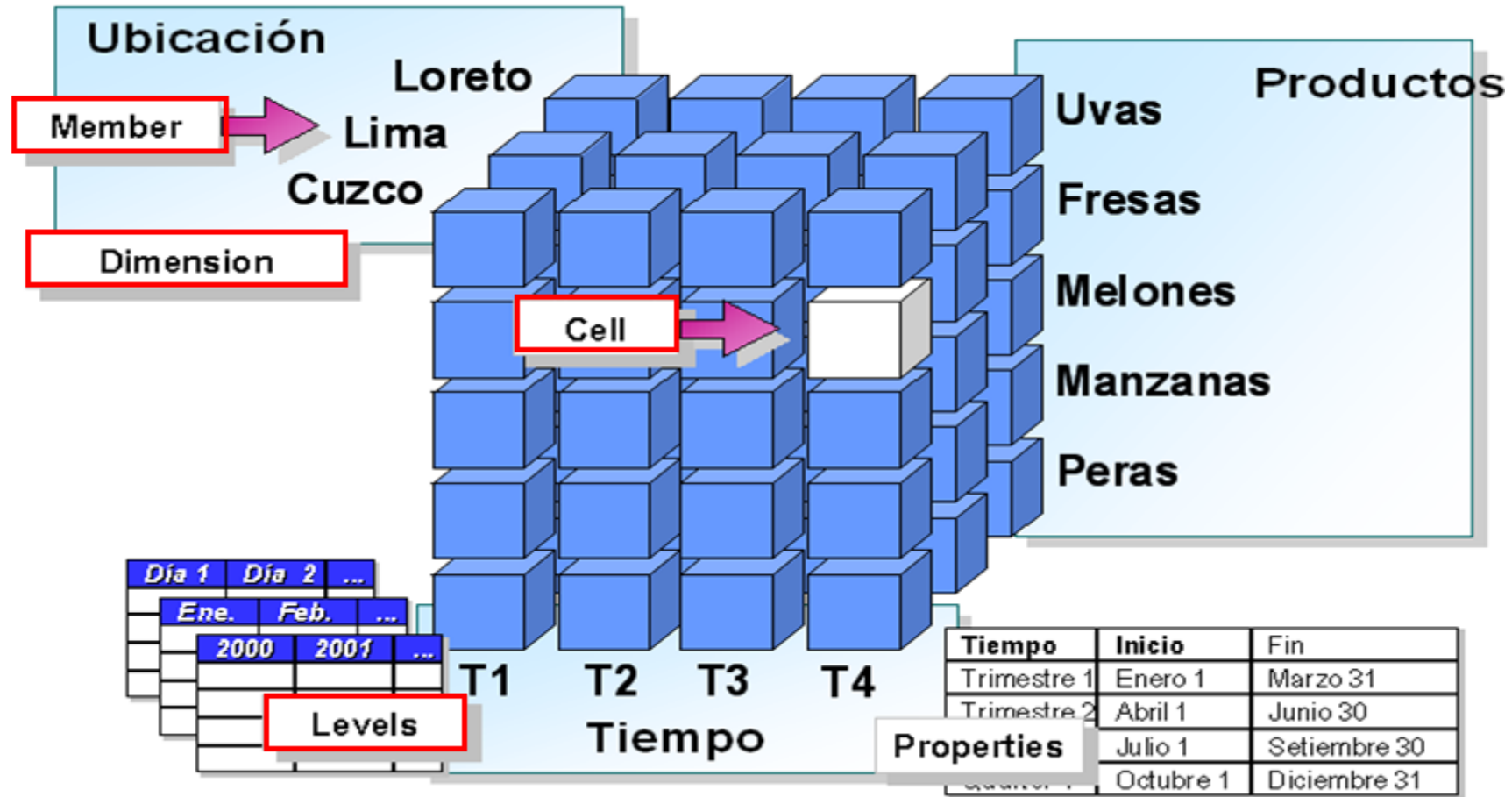
La data está representada en el cubo de datos que se muestra en la siguiente diapositiva.

En dicho cubo, la celda coloreada de blanco representa la métrica para el trimestre 4, ciudad Cuzco, producto fresas.



DEFINICIÓN DEL CUBO DE DATOS

DEFINICIÓN DEL CUBO DE DATOS



COMPONENTES DE UN CUBO DE DATOS

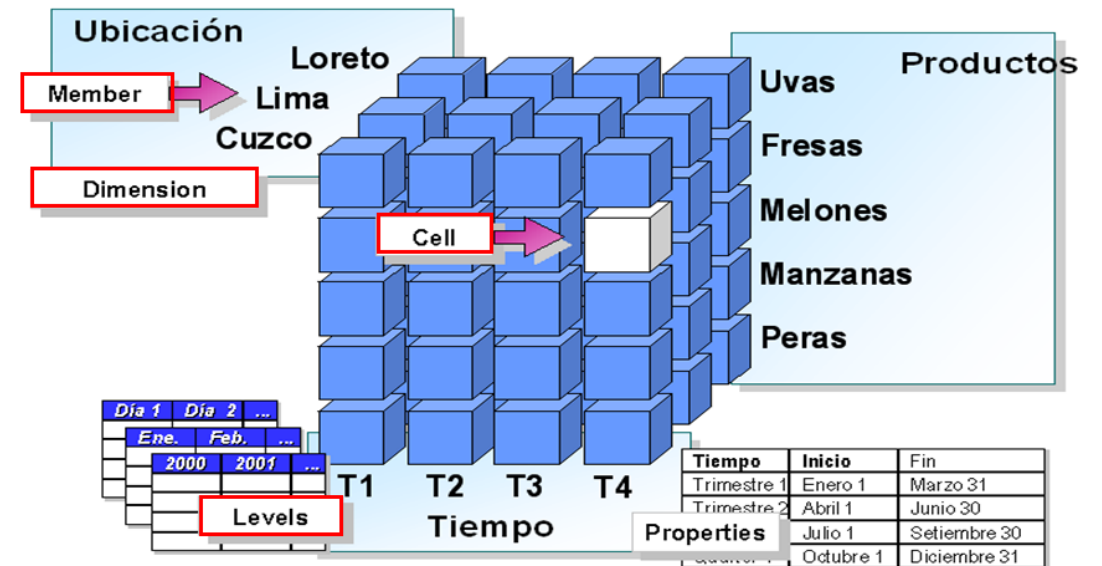
DIMENSIONES: cada uno de los criterios por los que se analiza el hecho. En el ejemplo se muestra las dimensiones **Ubicación**, **Productos** y **Tiempo**.

MIEMBRO de una dimensión: cada una de las ocurrencias en una dimensión. Por ejemplo, la dimensión **Ubicación** muestra 3 miembros: **Cuzco**, **Lima** y **Loreto**.

PROPIEDADES de una dimensión: cada uno de los atributos de la dimensión. Por ejemplo, para la dimensión **Tiempo** tenemos las propiedades **denominación**, **fecha de inicio**, **fecha de fin**.

NIVELES de una dimensión: jerarquías presentes en una dimensión. En la dimensión **Tiempo** tenemos los niveles: **año**, **trimestre**, **mes** y **día** que nos permiten analizar el hecho por periodos de tiempo.

CELDA: valor de una medida en el cubo. Si el cubo muestra la medida monto de ventas, la celda apuntada registra el monto de ventas del producto **Fresas**, en la ubicación **Cuzco**, para el periodo **T4**.



CREACIÓN DE UN CUBO DE DATOS

Para crear un cubo debemos:

- Definir la fuente de los datos para las dimensiones, las transacciones y la métrica (medidas) a procesar en el cubo.
- Definir la vista de la fuente de los datos que define la data relevante para cubo.
- Crear las dimensiones y personalizarlas según los requerimientos de información para análisis.
- Crear la tabla de hechos y su métrica.
- Procesar el cubo.

PROCESAMIENTO DE UN CUBO DE DATOS

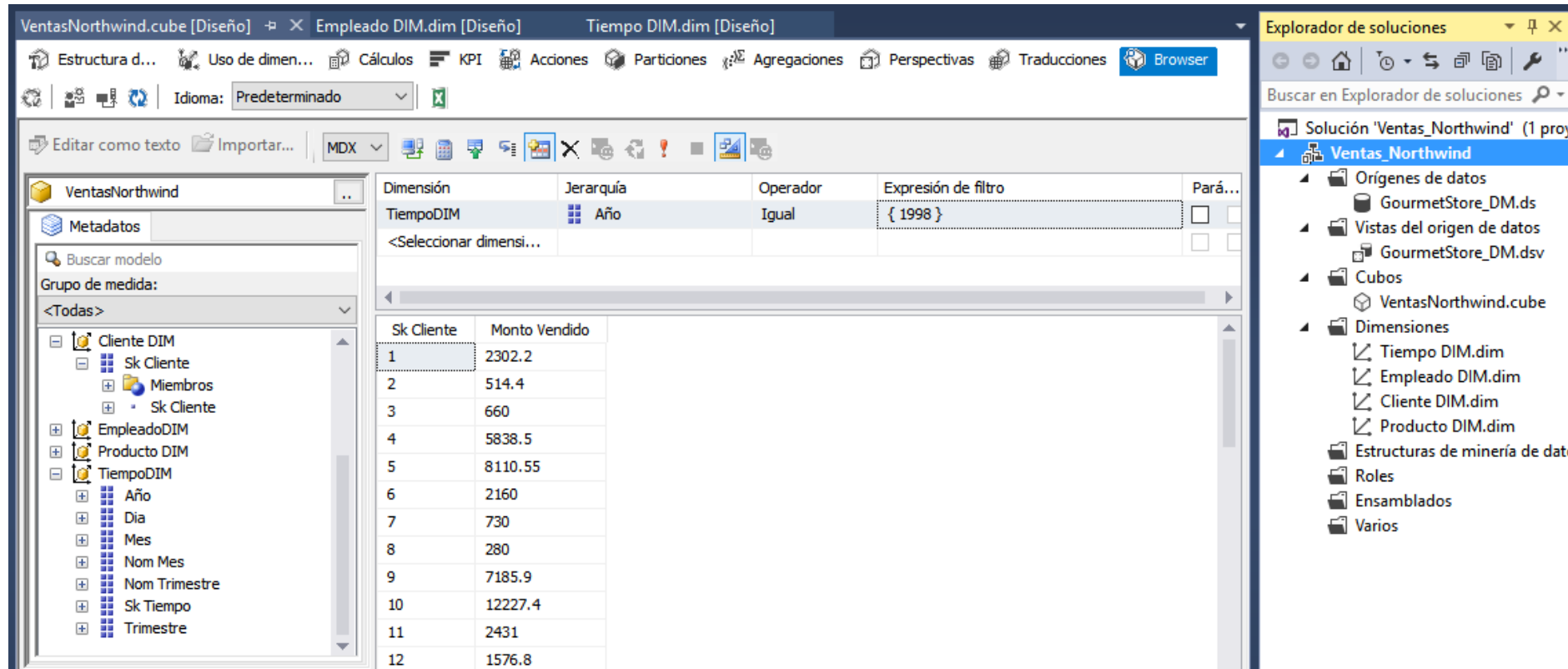
Para acceder a la información del cubo debemos:

- Conecta el cubo a la fuente de datos para extraer la información.
- Calcula la sumarización de los datos (agregaciones).
- Procesa el cubo y las dimensiones que no hayan sido procesadas.
- Mientras no se procese el cubo solo se tendrá creada su estructura, pero la información no estará disponible.

The logo of the University of Engineering and Technology is a circular seal. It features a central gear with a book and a torch inside. The text "UNIVERSIDAD NACIONAL DE INGENIERIA" is written around the top half, "LIMA PERU" around the bottom half, and "SCIENTIA ET LABOR" on the right side. The year "1826" is at the bottom.

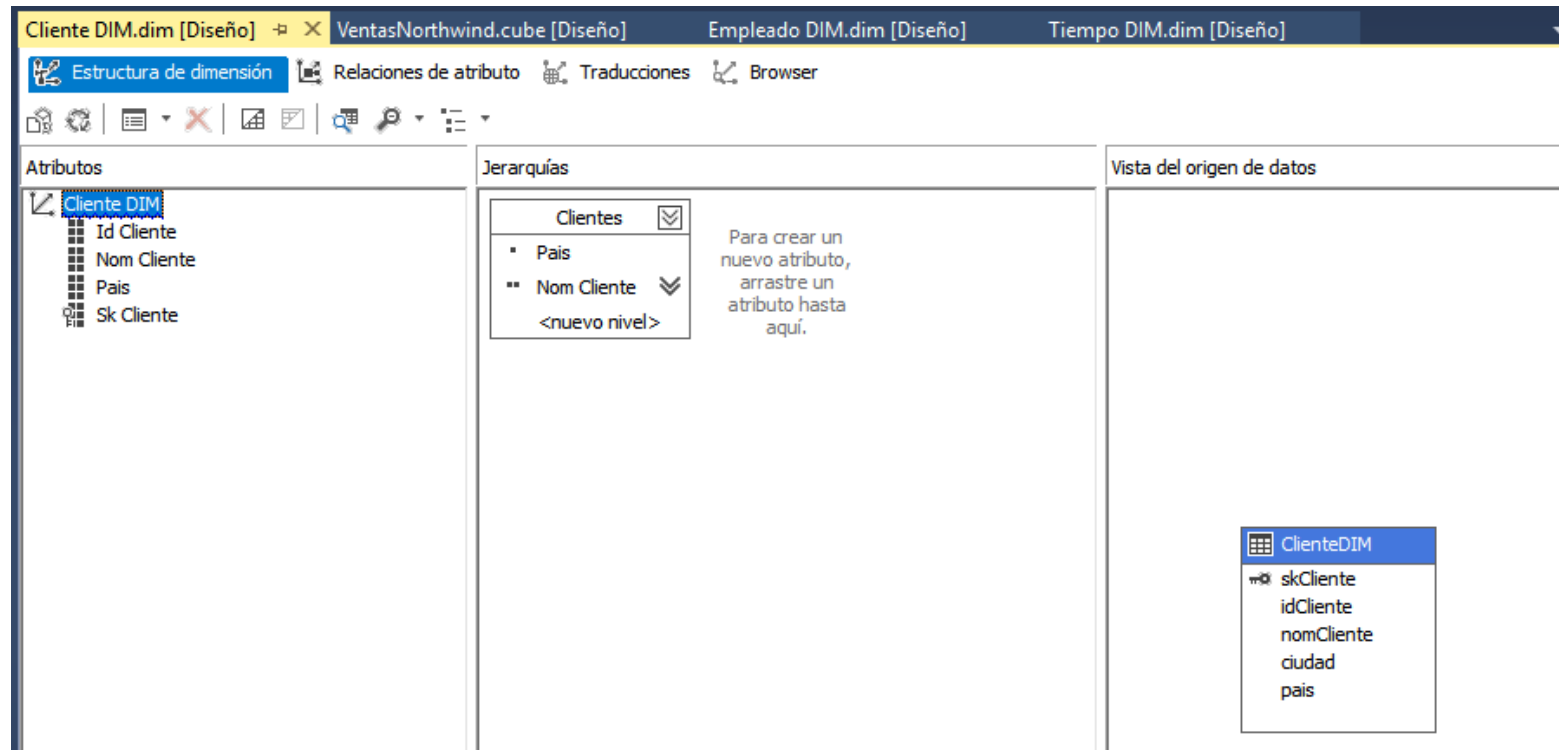
PERSONALIZACIÓN DEL CUBO Y DIMENSIONES

PERSONALIZACIÓN DEL CUBO (DISEÑADOR DE CUBOS)



La definición del cubo Analysis Services de la figura tiene 4 dimensiones: **Tiempo DIM**, **Empleado DIM**, **Cliente DIM** y **Producto DIM**. Observe que el cubo muestra la medida **Monto Vendido** por cliente, solo que para el cliente muestra su clave subrogada, lo que no es útil para hacer análisis del negocio. Esto debido a que para **Cliente DIM** se ha considerado como **atributo identificador** y **atributo descriptor** la clave subrogada. Se requiere configurar la dimensión **Cliente DIM** estableciendo como atributo descriptor el **nombre del cliente**.

PERSONALIZACIÓN DE DIMENSIONES (DISEÑADOR DE DIMENSIONES)



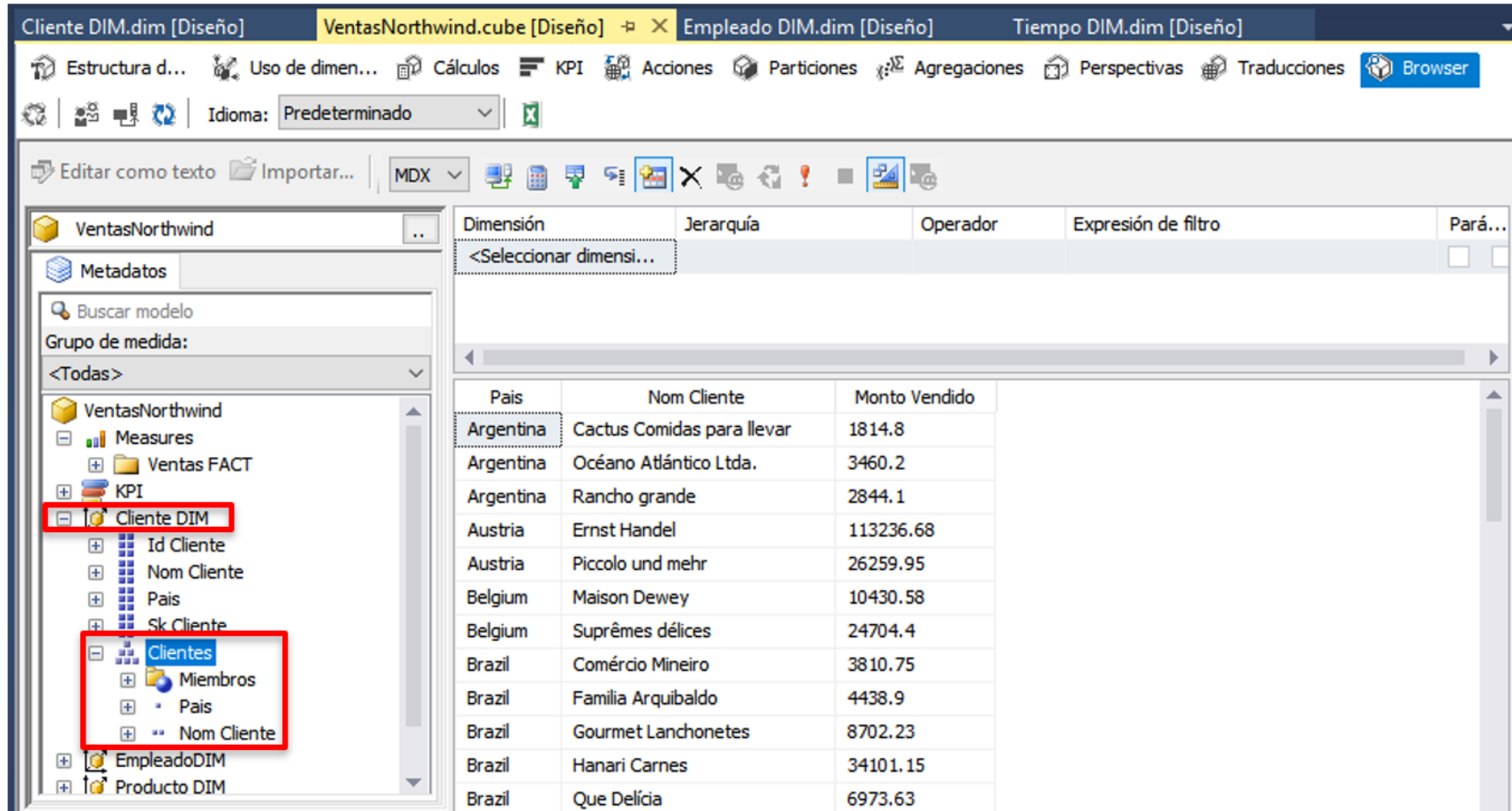
Panel Vista del origen de datos: atributos disponibles en el origen de datos.

Panel Atributos: atributos disponibles para la dimensión.

Panel Jerarquías: creadas en base a los atributos disponibles para la dimensión.

La jerarquía **Clientes** habilita al cubo para analizar las medidas por los niveles **país del cliente** y **nombre del cliente**. Ver panel de diseño del cubo en la siguiente diapositiva.

PERSONALIZACIÓN DEL CUBO Y SUS DIMENSIONES



The screenshot displays the SQL Server Data Tools (SSDT) interface for the 'VentasNorthwind.cube' design view. The left pane shows the 'VentasNorthwind' metadata tree, with 'Cliente DIM' and its hierarchy 'Clientes' highlighted. The right pane shows a table with columns 'Pais', 'Nom Cliente', and 'Monto Vendido'.

| Pais | Nom Cliente | Monto Vendido |
|-----------|----------------------------|---------------|
| Argentina | Cactus Comidas para llevar | 1814.8 |
| Argentina | Océano Atlántico Ltda. | 3460.2 |
| Argentina | Rancho grande | 2844.1 |
| Austria | Ernst Handel | 113236.68 |
| Austria | Piccolo und mehr | 26259.95 |
| Belgium | Maison Dewey | 10430.58 |
| Belgium | Suprêmes délices | 24704.4 |
| Brazil | Comércio Mineiro | 3810.75 |
| Brazil | Familia Arquibaldo | 4438.9 |
| Brazil | Gourmet Lanchonetes | 8702.23 |
| Brazil | Hanari Carnes | 34101.15 |
| Brazil | Que Delícia | 6973.63 |

La jerarquía **Clientes** de la dimensión Cliente DIM habilita al cubo para analizar las medidas por los niveles **país del cliente** y **nombre del cliente**.



EXPLORACIÓN DEL CUBO

EXPLORACIÓN DEL CUBO

The screenshot shows the 'VentasNorthwind.cube' design view in SSDT. The main workspace displays a table with the following data:

| Pais | Nom Cliente | Unidades Vendidas |
|------|-----------------|-------------------|
| Peru | MP Breña | 4000 |
| Peru | MP Callao | 4000 |
| Peru | MP Los Olivos | 4000 |
| Peru | MP Miraflores | 4000 |
| Peru | MP Pueblo Libre | 4000 |

The filter expression '{ 7 UP DESCARTABLE }' is applied to the 'Unidades Vendidas' column.

Para explorar el contenido del cubo utilice la pestaña **Browser**. Allí puede definir las filas y columnas que mostrará el informe, así como un filtro para seleccionar datos en base a determinada condición.



CONCLUSIONES

CONCLUSIONES

- ❖ El concepto de cubo permite explicar el porqué del modelado de datos multidimensional y cómo facilita el análisis de datos, ya que partiendo de un elemento en 3 dimensiones, podemos explicar la multidimensionalidad de los cubos de datos.
- ❖ El procesamiento de un cubo, además de extraer la data que requiere de la fuente de datos, ejecuta todas las agregaciones de datos (consultas GROUP BY) requeridas para obtener los reportes consolidados de forma más rápida.