



Universidad
Tecmilenio®





Consulta en Microsoft SQL Server®

Implementación del flujo de control en un paquete de SQL Server Integration Services (SSIS)

Semana 10



Las empresas suelen tener una gran cantidad de información dispersa en distintos repositorios, por lo que contenerla en una sola base es de gran apoyo para su posterior análisis.

El **flujo de control** nos permite tener la jerarquía de acuerdo con los contenedores que usaremos para la extracción de información para al final solo tener un reporte con lo requerido de acuerdo con los expertos de la empresa y sus necesidades.



Introducción al flujo de control

Existen tres elementos del flujo de control dentro del SSIS o SQL Server Integration Services (Microsoft, 2022a):

Los repositorios de donde se obtendrán las estructuras del conjunto de tareas (los paquetes).

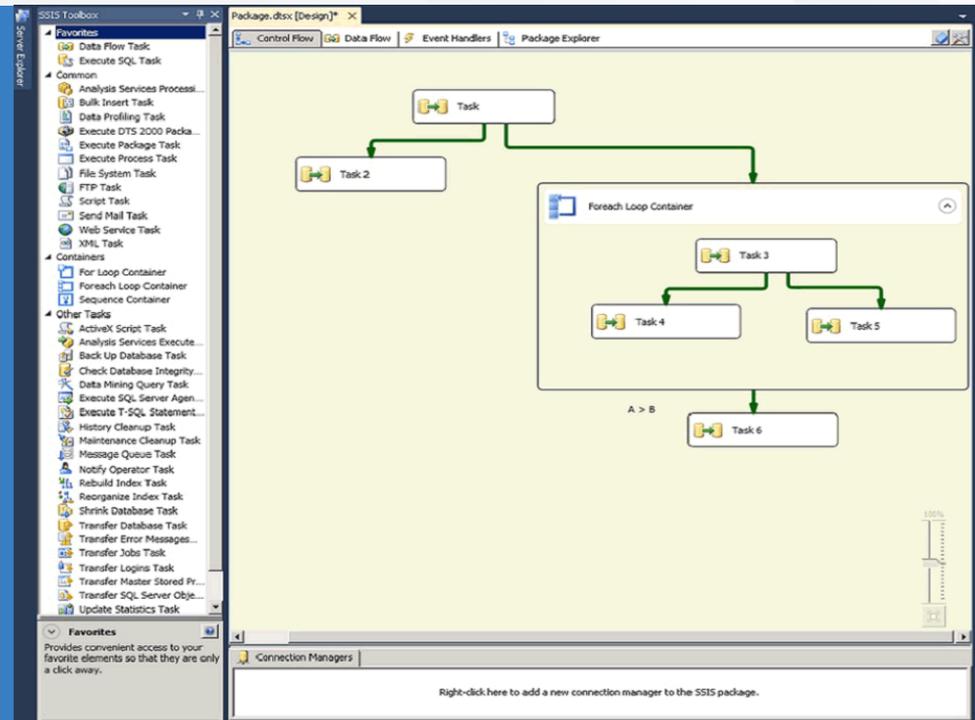
Las tareas que proporcionan la funcionalidad y las restricciones de precedencia que conectan los ejecutables y los contenedores.

Las tareas en un flujo de control ordenado.

Ejemplo:

Diagrama representativo de un conjunto simple de tareas creado desde la pestaña “Flujo de control” desde el diseñador SSIS.

El diseñador SQL Server Integration Services (SSIS) es capaz de agregar una conexión automática después de generar un elemento, ya sea una tarea o un contenedor.



Creación de paquetes dinámicos

Los paquetes dinámicos son pequeños, universales y de ajuste automático; esto gracias a la herramienta de implementación de Office (ODT), que habilita el ajuste automático, utilizando lo siguiente:

Version=MatchInstalled

Para mantener la armonía entre las versiones instaladas de las aplicaciones.

Language=MatchInstalled

Para tener acceso al mismo catálogo de idiomas de las aplicaciones ya instaladas.

Quitar los archivos de origen de los paquetes, haciéndolos esbeltos (*lean*), logra lo siguiente (Microsoft, 2022b):

- Reducir el peso del paquete de gigabytes a megabytes.
- Los clientes pueden seleccionar los elementos de su interés.
- Reducir la necesidad de un segundo escenario de instalación.
- Los archivos de origen no necesitan ser actualizados.

Uso de contenedores

Un archivo o una aplicación capaz de almacenar elementos incrustados o vinculados en sus propios documentos se conoce como **contenedor**.

Dentro de SQL Server Integration Services, los objetos **contenedores** brindan servicios a las tareas y ayudan a proporcionar estructura a los paquetes creados.

Los **contenedores** sirven para dar estructura y almacenar tareas, pero también pueden ser anidados y usados para las tareas repetitivas.



Tipos de contenedores

Contenedor Foreach Loop

Ejecuta un flujo de control repetidamente mediante un enumerador.

Contenedor de bucles For

Ejecuta un flujo de control repetidamente probando una condición.

Contenedor de secuencias

Agrupar tareas y contenedores en flujos de control que son subconjuntos del flujo de control del paquete.

Contenedor de tarea

Proporciona servicios a una tarea individual.

Microsoft. (2022c). *Contenedores de Integration Services*. Recuperado de <https://docs.microsoft.com/es-es/sql/integration-services/control-flow/integration-services-containers?view=sql-server-ver16>



Propiedades del contenedor

DelayValidation
Descripción
Deshabilitar
DisableEventHandlers
FailPackageOnFailure
FailParentOnFailure
ForcedExecutionValue
ForcedExecutionValueType
ForceExecutionResult
ForceExecutionValue
Id
IsolationLevel
LocaleID
LoggingMode
MaximumErrorCount
Nombre
TransactionOption



Con base en lo descrito en el tema, reflexiona sobre las siguientes preguntas:

01

→ ¿Cuál es la importancia de tener una jerarquía en el proceso de extracción dentro de un flujo de control?

02

→ Menciona un ejemplo donde se pueda implementar un contenedor de secuencias.





Como se puede observar, el **flujo de control** sirve para brindar la estructura a los paquetes, así como la jerarquía entre contenedores y tareas que se conectan por las restricciones de procedencia. Gracias a esto, es posible realizar la automatización de tareas repetitivas y que conllevan un tiempo largo de ejecución.

Finalmente, es importante reconocer que cada **contenedor** puede aportar un valor diferente al proyecto y, dependiendo de las necesidades de este último, se deberán seleccionar los contenedores más útiles.

Cierre





Bibliografía

- Microsoft. (2022a). *Flujo de control*. Recuperado de <https://docs.microsoft.com/es-es/sql/integration-services/control-flow/control-flow?view=sql-server-ver15>
- Microsoft. (2022b). *Procedimientos recomendados del campo: Crear paquetes dinámicos, eficientes y universales para Aplicaciones Microsoft 365*. Recuperado de <https://docs.microsoft.com/es-es/DeployOffice/fieldnotes/build-dynamic-lean-universal-packages>
- Microsoft. (2022c). *Contenedores de Integration Services*. Recuperado de <https://docs.microsoft.com/es-es/sql/integration-services/control-flow/integration-services-containers?view=sql-server-ver16>



Consulta en Microsoft SQL Server®

Carga de datos en un
almacén de datos



Semana 10



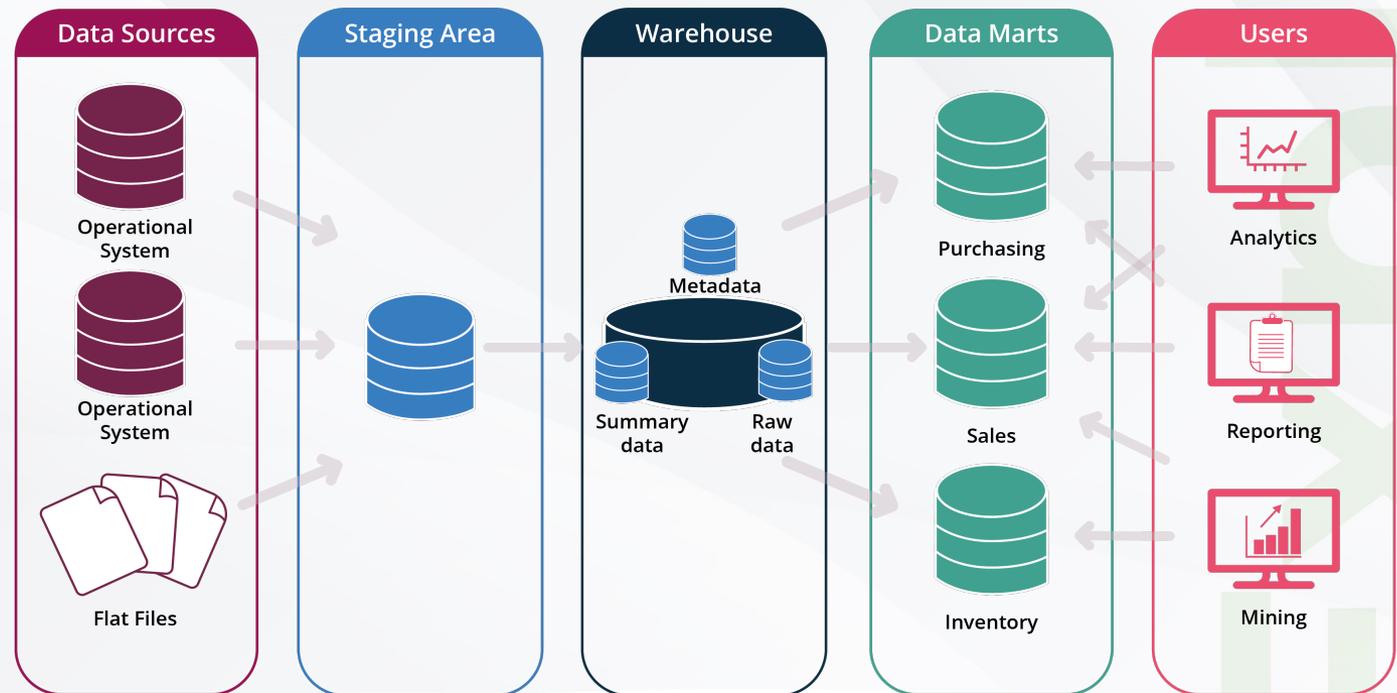
De acuerdo con Leland (2022), se le llama **carga de datos** al proceso de obtención de información o conjunto de datos desde una fuente a través de la selección, copia o recorte de información y su posterior pegado dentro de un software para su almacenamiento y resguardo.

Una práctica dentro de las industrias de **TI** es cargar los datos a una unidad de almacenamiento de destino con un formato diferente al utilizado por la fuente de donde se extrajo la información.



Planificación de cargas de datos

SSIS proporciona todos los elementos esenciales del procesamiento de datos, desde su origen hasta la “puesta en escena” y la creación de tablas en contenedores.



ETL

Es parte del proceso de *back end* de un almacén de datos y una forma sencilla para el resguardo de datos que se centra en un área específica de una organización.

Objetivos

- Reconocer las claves de origen que identifican de forma única un registro de origen.
- Realizar transformaciones de datos para alinear los datos de origen.
- Manejar los diferentes tipos de cambio para cada columna de origen.

Beneficios

- Business *Intelligence* mejorada.
- Rendimiento de consultas.
- Acceso oportuno a los datos.
- Calidad y consistencia mejoradas.
- Alto retorno de la inversión.

La solución **ETL** de SQL Server se basa en tener un resultado con velocidad de consulta y almacenamiento de datos, así como pensada en su escalabilidad.

Uso de SSIS para cargas incrementales

El proceso de carga es el movimiento físico de los datos desde los sistemas informáticos que almacenan las bases de datos fuentes al repositorio final que contendrá los datos resguardados.

A continuación, se presenta el proceso de transferencia de datos a un repositorio de almacenamiento (Acervo Lima, 2022):



Uso de técnicas de carga de T-SQL

Para ejecutar, gestionar y administrar paquetes SSIS es necesario utilizar técnicas de carga en **T-SQL**, que permiten parametrizar, implementar y administrar el servicio SSIS.

Transact-SQL (T-SQL) es un lenguaje de programación que permite almacenar temporalmente valores en variables, aplicar la ejecución condicional de comandos, pasar parámetros a procedimientos almacenados y controlar el flujo de los programas (Microsoft, 2022).

T-SQL resulta ser un lenguaje de uso masivo para realizar búsquedas estructuradas. Al ser estandarizado, su uso se vuelve más en la búsqueda y modificación de los datos de los repositorios a través de funciones y sentencias definidas.

Explicación



Diferentes sentencias y estructuras compatibles con la programación de cargas y búsquedas en lenguaje T-SQL:

IF...ELSE

WHILE

DECLARE

SET

BATCHES

Con base en lo descrito en el tema, reflexiona sobre las siguientes preguntas:

01

→ ¿Qué fuentes de origen de datos consideras que podría tener Facebook y necesite almacenar en un mismo repositorio?

02

→ ¿Cuál es el beneficio de tener la información estandarizada y centralizada en un mismo sitio?





Como hemos visto, se pueden aprovechar muchas oportunidades para usar **SQL** junto con **SSIS** en su solución **ETL** para desarrollar una solución de Business Intelligence (BI) que se ajuste a nuestras necesidades.

Todas las bases de datos pueden conformarse de información proveniente de múltiples fuentes, por lo que es importante comprender el proceso ETL.



Bibliografía

- Acervo Lima. (2022). *Carga de datos en el almacén de datos*. Recuperado de <https://es.acervolima.com/carga-de-datos-en-el-almacen-de-datos/>
- Leland. (2022). *Carga de datos*. Recuperado de <https://techinfo.wiki/carga-de-datos-2/>
- Microsoft. (2022). *Descripción de T-SQL para la programación*. Recuperado de <https://docs.microsoft.com/es-es/learn/modules/get-started-transact-sql-programming/2-describe-transact-sql-for-programming>



Consulta en Microsoft SQL Server®

Introducción a Business
Intelligence y modelado de datos



Semana 10



La empresa VasaBit está por expandirse en el mercado del software a la medida como parte de la estrategia de expansión.

Tú eres un recién graduado de la maestría de Ciencias en Datos y has tomado cursos relacionados con SQL, sin embargo, desconoces lo que sigue, que es precisamente la parte de modelado de datos e inteligencia de negocios, por lo que decides estudiar este tema.



Elementos de una solución BI empresarial

La **Business Intelligence** (BI) o inteligencia empresarial es una combinación de técnicas, como el minado, visualización, modelado y análisis de datos que sirven para tomar decisiones basadas en datos (Oracle, 2022).



Plataforma Microsoft Enterprise BI

Dentro del plan E5 de Microsoft Enterprise se incluye la herramienta **Power BI** que es la indicada para la inteligencia empresarial dentro del ecosistema de Microsoft.



The screenshot shows the Microsoft Store page for Power BI Desktop. It features the application icon (three yellow bars of increasing height), the title 'Power BI Desktop', and the developer 'Microsoft Corporation'. The page displays a 4.1 star average rating from 218 reviews and an ESRB 'EVERYONE' age rating. A blue progress bar at the top right indicates that 67.71 MB of 674.9 MB has been downloaded. The description states: 'Power BI Desktop puts visual analytics at your fingertips. With this powerful authoring tool, you can create interactive data visualizations and reports.'

Power BI también se conoce por:

- Ser una plataforma de autoservicio muy amigable en el momento de su utilización.
- Utiliza inteligencia artificial para brindar recomendaciones.
- Es segura y ofrece opciones en la nube para trabajar de manera remota y con otros usuarios.

Existen tres niveles en el modelado de datos:

- Conceptual: define la relación entre los datos y el negocio.
- Lógico: aterriza las ideas, conceptos y actores definidos. Después los aplica como objetos y/o como atributos de datos.
- Físico: para hacer real la representación lógica en términos de tablas, relaciones, conexiones y funcionalidad de la base de datos.

Araneda, P. (2020). *Capítulo 4 Modelamiento de Datos*. Recuperado de <https://bookdown.org/paranedagarcia/database/modelamiento-de-datos.html>

Planificación de un proyecto BI

Elementos de un proyecto de BI:



Con base en lo descrito en el tema, reflexiona sobre las siguientes preguntas:

01



Menciona al menos tres ventajas de la toma de decisiones basadas en datos dentro de una organización.

02



¿Por qué se debe invertir tiempo en la definición de un proyecto de tipo BI?





- A lo largo del contenido se tocaron diversos temas relacionados con el mundo de la inteligencia empresarial o de negocios.
- Se observó de manera superficial lo que ofrece Microsoft Enterprise y cuál es su relación con la herramienta por naturaleza de este ecosistema: **PowerBI**.
- Se definieron algunos de los puntos generales y específicos más importantes para emprender en un proyecto de **BI**.



Bibliografía

- Araneda, P. (2020). *Capítulo 4 Modelamiento de Datos*. Recuperado de <https://bookdown.org/paranedagarcia/database/modelamiento-de-datos.html>
- Oracle. (2022). *What is Business Intelligence (BI)?* Recuperado de <https://www.oracle.com/what-is-business-intelligence/>



Consulta en Microsoft SQL Server®

Creación de bases de datos
multidimensionales



Semana 10



El área de inteligencia de negocios o empresarial, la cual acaba de arrancar hace no más de seis meses, presenta un ROI negativo.

Después de esto te das cuenta de que el *data warehouse* y las bases de datos aledañas a este repositorio están fuera de fecha y, por ende, los sistemas de BI no están arrojando los resultados operacionales ni financieros esperados.



¿Qué harías en esta situación?

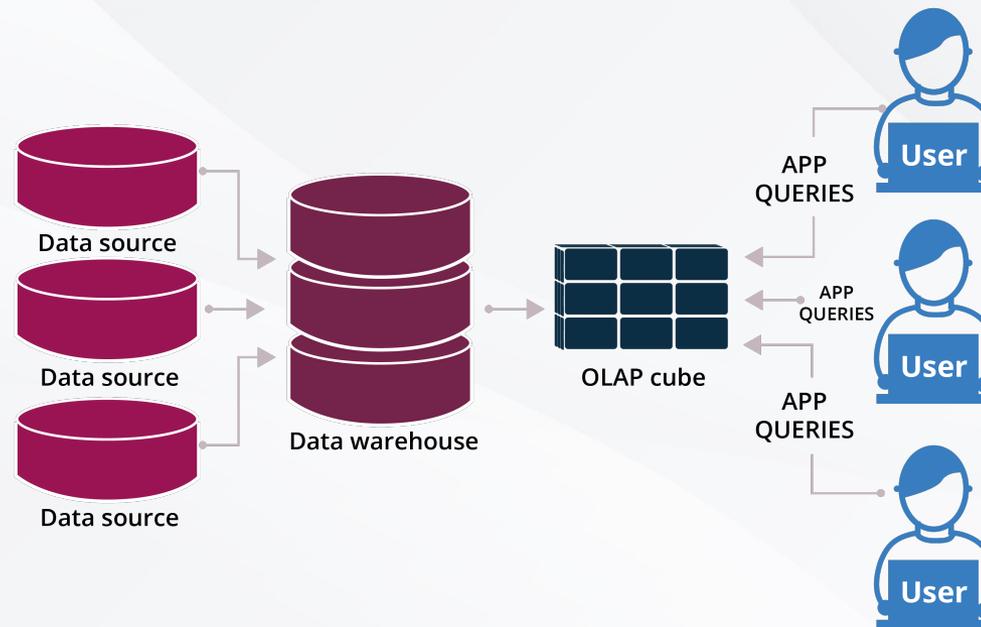
¿Será necesario implementar bases de datos multidimensionales?

¿Cómo se relacionarían estas con un *data warehouse*, considerando que este repositorio es rígido y costoso?

Introducción al análisis multidimensional

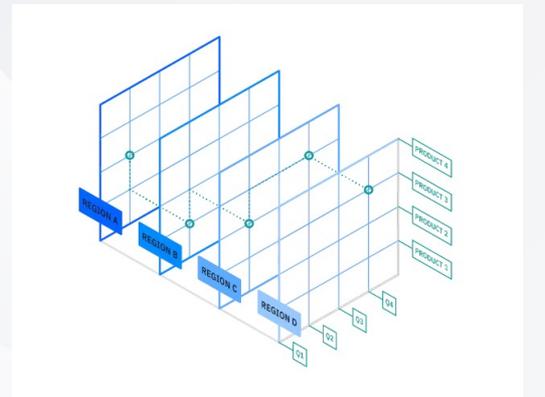
De acuerdo con **MyABCM** (2022), el análisis multidimensional se basa en categorías y se hace sobre grandes cantidades de datos.

Microsoft (2022) menciona un término conocido como **OLAP (online analytical processing)**, el cual es ampliamente utilizado en la práctica y se relaciona con el análisis multidimensional.



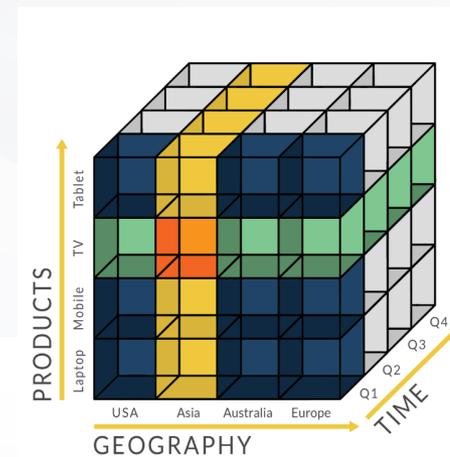
Los **cubos OLAP** son vistos como estructuras de datos que propician un análisis más rápido de datos.

De acuerdo con **IBM** (2020), un **cubo** es una base de datos multidimensional que se basa en matrices.



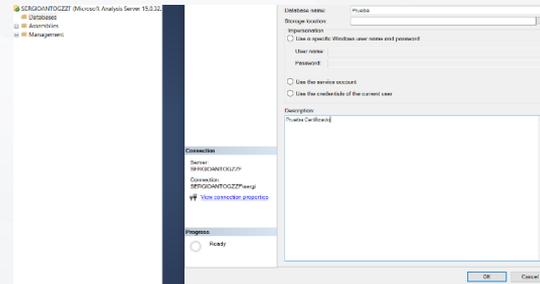
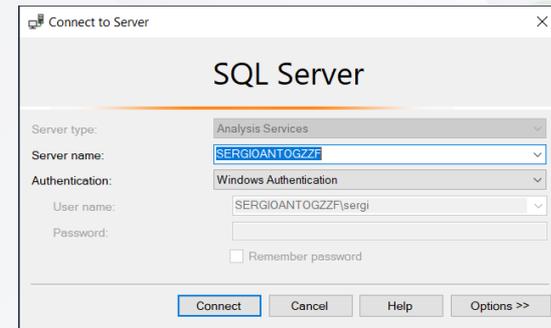
Ejes de un cubo.OLAP

Dimensiones de un cubo OLAP.

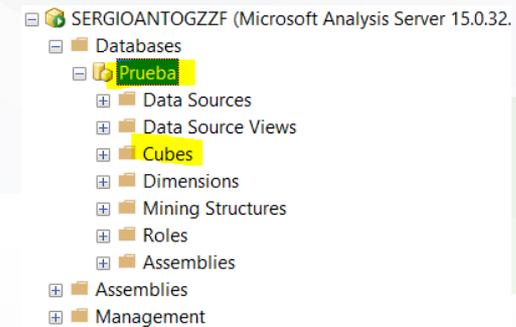


Para la **creación de un cubo** es necesario seguir los pasos para agregar funcionalidades al mismo sistema ya instalado de SQL Server.

1 Selecciona la opción de "Analysis Services".



2 Creación de la base de datos.

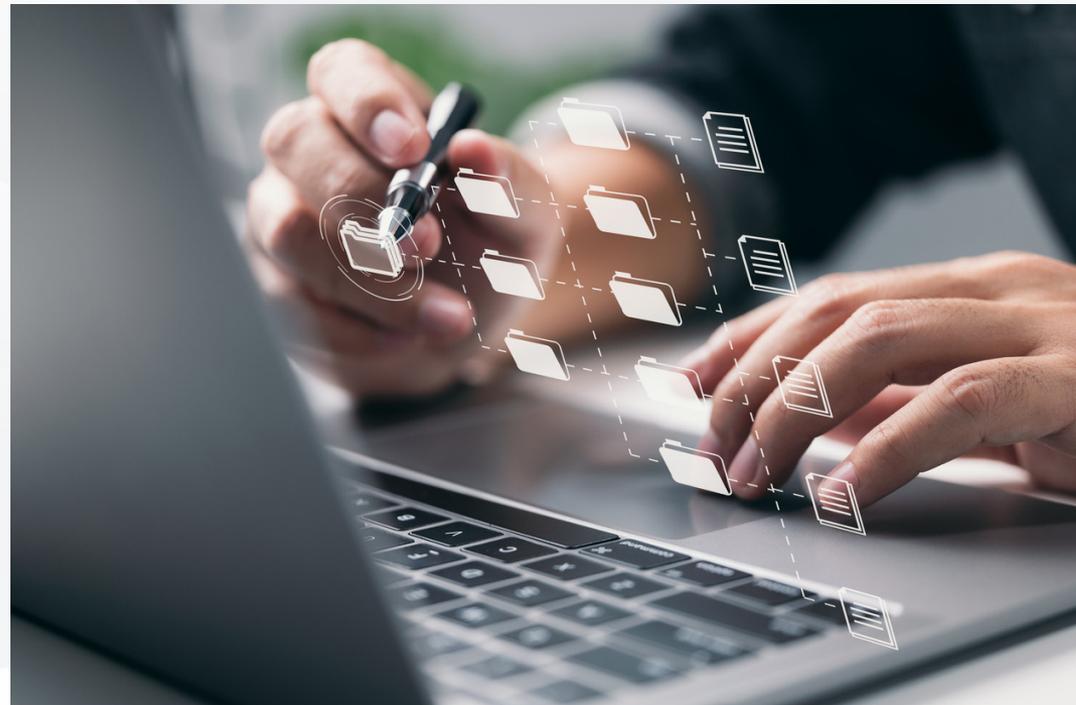


3 Elementos para la configuración de la base de datos.

Creación de fuentes de datos y vistas de fuentes de datos

Es necesario crear fuentes de datos y vistas para completar una configuración de un cubo.

Para esto, es necesario descargar la aplicación de SQL e instalar los servicios de análisis y reportes en Microsoft Visual Studio.

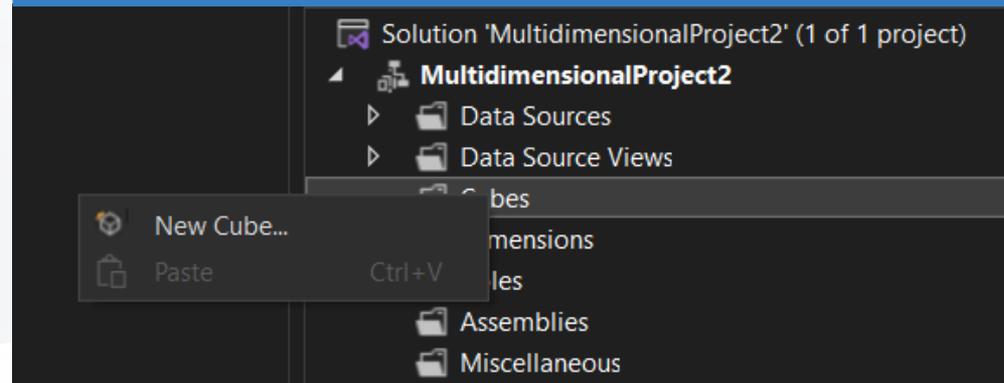


Creación de un cubo

Perspectivas de un cubo:

En su macroescala, el cubo es la base de datos multidimensional que alberga pequeños cubos en forma de matriz.

Este cubo mayor es la base de datos que ya está creada, sin embargo, para crear pequeños cubos dentro de este, es necesario crearlos por separado y con base en vistas y dimensiones.



Con base en lo descrito en el tema, reflexiona sobre las siguientes preguntas:

01

De acuerdo con tu interpretación, ¿cómo explicarías el significado de un cubo?

02

Menciona los beneficios de usar un cubo OLAP.



El análisis multidimensional se utiliza para temas de inteligencia de negocios, ya que actúa como base para los análisis que se realicen posteriormente.

Los **cubos OLAP** se utilizan ampliamente en la industria de la ciencia de datos.

Para crear un proyecto multidimensional es necesario descargar varias extensiones y sistemas de Microsoft, como los servicios de análisis e integración en **SQL Server** y el **Visual Studio**.

Es importante entender cuáles son las relaciones entre las tablas de la fuente de datos de una base de datos relacional con los grupos de medición de una base de datos multidimensional (**cubo**).



Bibliografía

- IBM. (2020). *What is OLAP?* Recuperado de <https://www.ibm.com/cloud/learn/olap>
- Microsoft. (2022). *Información general sobre el procesamiento analítico en línea (OLAP)*. Recuperado de <https://support.microsoft.com/es-es/office/informaci%C3%B3n-general-sobre-el-procesamiento-anal%C3%ADtico-en-l%C3%ADnea-olap-15d2cdde-f70b-4277-b009-ed732b75fdd6>
- MyABCM. (2022). *¿Cómo puede el análisis multidimensional ayudar a la empresa?* Recuperado de <https://myabcm.com/es/como-la-analisis-multidimensional-ayuda/>