



UNAH
UNIVERSIDAD NACIONAL
AUTÓNOMA DE HONDURAS

Facultad de Ingeniería
Bases de Datos II
DataWarehouse - AdventureWorks
Sección 1000

Presentado por:

Cesar Barahona	20141004268
Ana Iris Padilla Maradiaga	20141006787
Dilver Benavidez Alvarado	20131007823
Bessy Mariela Velásquez	20151004281
Allan Jafect Martínez Lagos	20151002529

Ciudad Universitaria, Tegucigalpa MDC, Francisco Morazán
29/04/2018

INDICE

INTRODUCCION 3

INTELIGENCIA DE NEGOCIOS 4

BASE DE DATOS OLTP 4

BASE DE DATOS OLAP 7

CUBO OLAP 10

REPORTES 13

REPLICACION..... 15

JOBS..... 15

INTRODUCCION

A continuación, se presenta el informe detallado del desarrollo del proyecto de inteligencia de negocios de la empresa Adventure Work, el cual consta de la construcción de la base de datos OLAP que será cargada desde la herramienta Visual Studio a través de un proceso denominado ETL, una vez cargada la base de datos con la información necesaria se procede a hacer análisis exhaustivos de los mismo haciendo uso de herramientas de soporte para la toma de decisiones (DSS. En este caso se utilizó Pentaho. A continuación, se detalla paso a paso cada proceso.

INTELIGENCIA DE NEGOCIOS

HERRAMIENTAS NECESARIAS:

- BASE DE DATOS OLTP
- BASE DE DATOS OLAP
- PROYECTO DE CONTRUCCION DE ETL
- CONFIGURACION Y OBTENCION DE REPORTES DE ANALISIS(PENTAHO)

BASE DE DATOS OLTP

Para llevar a cabo este proyecto se decidió utilizar la base de datos de la empresa Adventure Work Cycles, compañía dedicada al comercio de bicicletas, la versión a utilizar es la de las ventas y transacciones hechas en el año 2014, la definición y estructura es la siguiente:

Adventure Works Cycles

Adventure Works Cycles, es una gran empresa de fabricación multinacional. La empresa fabrica y vende bicicletas de metal y de metal compuesto en los mercados de Norteamérica, Europa y Asia. Si bien su sede central de operaciones se encuentra en Bothell, Washington, con 290 empleados, en toda su base de mercado tiene distribuidos varios equipos regionales de ventas.

Las actividades empresariales que realiza la empresa frecuentemente son como:

- ✚ Ventas
- ✚ Configuración de ventas y otros objetivos de la empresa
- ✚ Administración de inventarios
- ✚ Administración de comercio de muestra
- ✚ Administración de personal
- ✚ Presupuestos

Adventure Works Base de Datos

La base de datos es libre y se puede descargar desde la página oficial de Microsoft, es una base de datos transaccional desarrollada en sql server, cuenta con una estructura bien definida que permite el fácil mantenimiento las áreas que en ella se almacenan(ventas, recursos humanos, productos, etc.)

ESTRUCTURA DE ADVENTURE WORK BASE DE DATOS

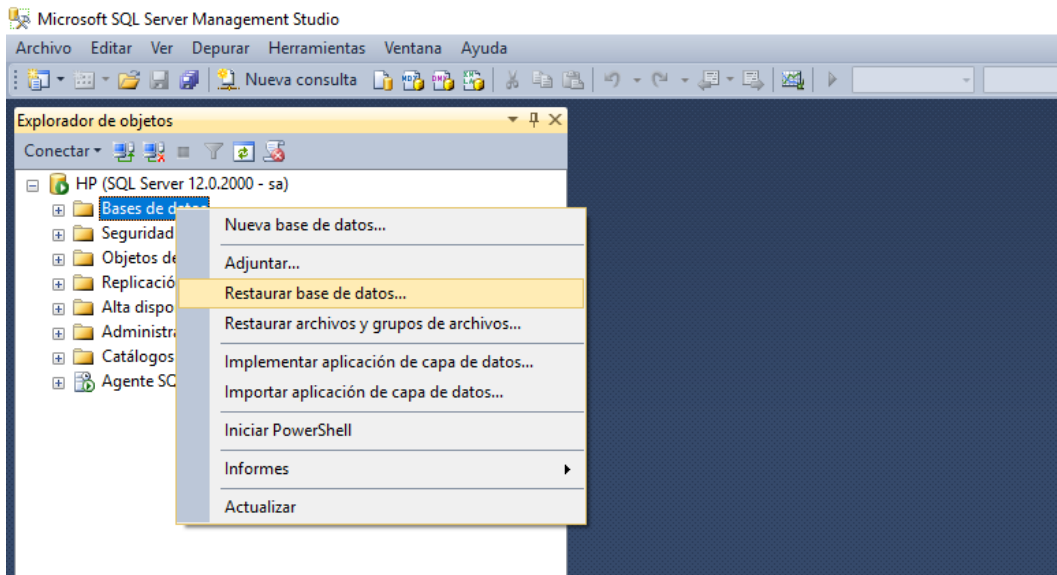
Actualmente esta base de datos maneja distintas áreas de trabajo, desde el manejo del personal hasta el área de ventas, a continuación se enlistaran de forma general Adventure Work base de datos:

- ✚ Recursos Humanos: aquí se lleva el control del, total de empleados que laboran para la empresa, así como el historial de sus trabajos, área asignada y direcciones.
- ✚ Personas (Empleados, Clientes): aquí se controlan datos personales de los empleados y clientes, es decir correos, números de teléfono, dirección de domicilio etc.
- ✚ Ventas: Se almacena toda la información relacionada a las ventas que a diario genera la empresa, fechas de ventas, vendedor, cliente, tienda, cantidad de productos, precios, totales etc.
- ✚ Compras: estructura utilizada para llevar el control ordenado de las compras que la empresa realiza, precios, totales etc.
- ✚ Producción: número de artículos, medidas, modelos, colores etc.

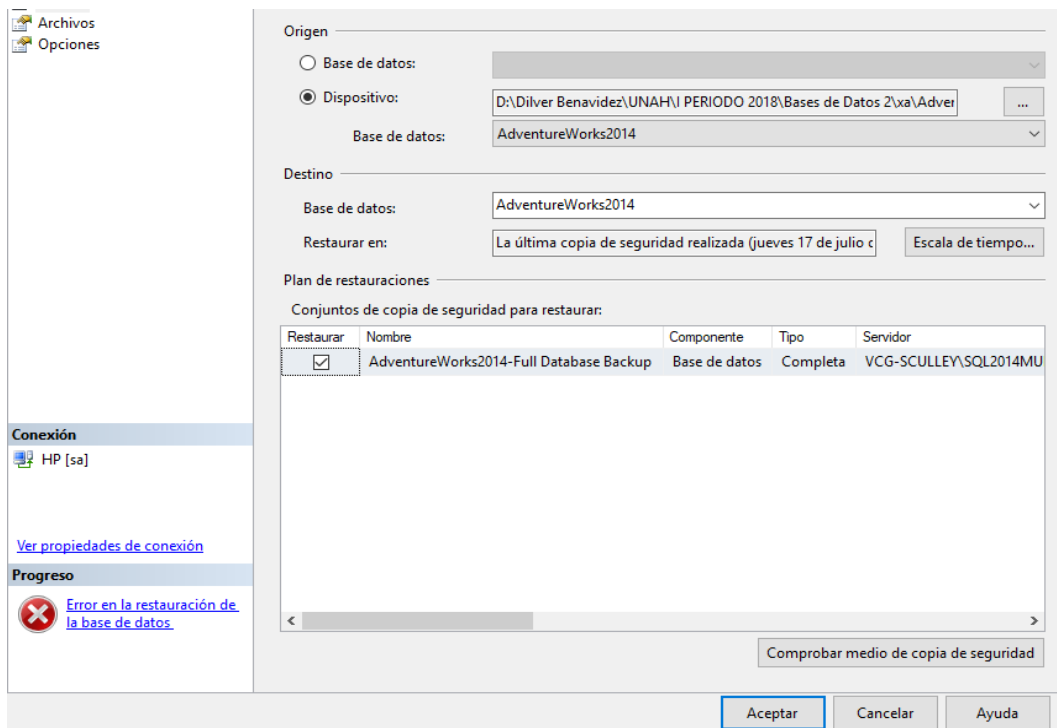
RESTAURACION DE LA BASE DE DATOS EN SQL SERVER

Una vez ya elegida la base de datos a utilizar no resta más que realizar la restauración del archivo .bak descargado desde el sitio web oficial de Microsoft. A continuación se detallan los pasos a seguir:

- Sobre el nombre de la instancia de sql server instalada se desplaza hasta la pestaña de Bases de datos, haciendo clic derecho en la opción restaurar base de datos.



- A continuación de apertura una ventana en la cual debe de seleccionar la ruta en la que tiene almacenado el archivo .bak que había descargado anteriormente. Presiona aceptar y la base de datos de debe restaurar sin mostrar ningún error.



- La base de datos ya debe de aparecer creada en la columna izquierda en la pestaña de base de datos.

En la carpeta del proyecto se adjunta la imagen del diagrama de la base de datos Adventure Works 2014.

BASE DE DATOS OLAP

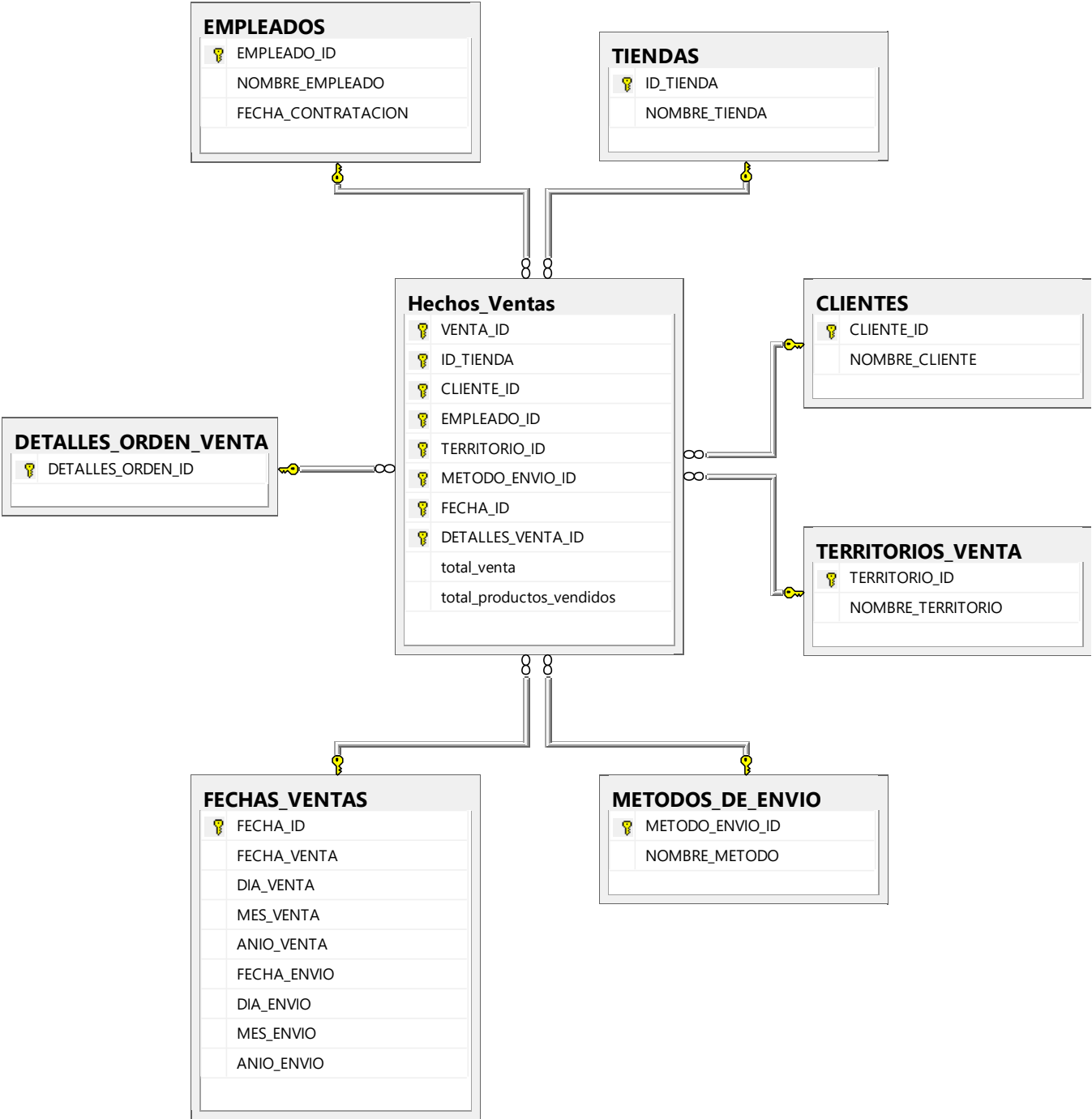
Los sistemas OLAP son bases de datos orientadas al procesamiento analítico. Este análisis suele implicar, generalmente la lectura de grandes cantidades de datos para llegar a extraer algún tipo de información útil: tendencia de ventas, patrones de los comportamientos de los consumidores, elaboración de informes complejos. Etc. Este sistema es típico de los datamarts.

Un Datamart es una base de datos departamental, es decir que se especializa en el almacenamiento de datos de un área del negocio.

Para este proyecto se utilizó una Datamarts del área de ventas, es decir, esta base de datos OLAP almacenara las ventas totales por fechas, además de los datos del empleado que vendió y que cliente fue el que realizo la compra. Además muestra la cantidad de productos vendidos en una venta en específico.

El modelo en que se diseñó el datamarts es en modelo estrella dado a que solo existe relación entre la tabla de Hechos y las tablas de dimensiones.

En la carpeta del proyecto se adjunta el script para la creación de la base de daros OLAP, la cual debe quedar de la siguiente forma:



PROYECTO DE CONTRUCCION DE ETL (EXTRACION TRANSFORMACION Y CARGA)

Una vez terminada e instalada la base de datos OLAP se debe llenar con los datos existentes en la base de datos OLTP, para realizar esto existen varias herramientas pero en este caso se hará uso de SQL Server Data Tools for Visual Studio 2012 para la creación de cada ETL. Esto significa que se hará un proceso de extracción de los datos de la base OLTP, seguida de una transformación de los datos para posteriormente ser cargados a la base de datos OLAP.

SQL Server Data Tools for Visual Studio 2012 permite realizar el proceso de ETL de manera rápida, con tan solo agregar 3 componentes y realizar algunas configuraciones de puede llenar la base de datos OLAP.

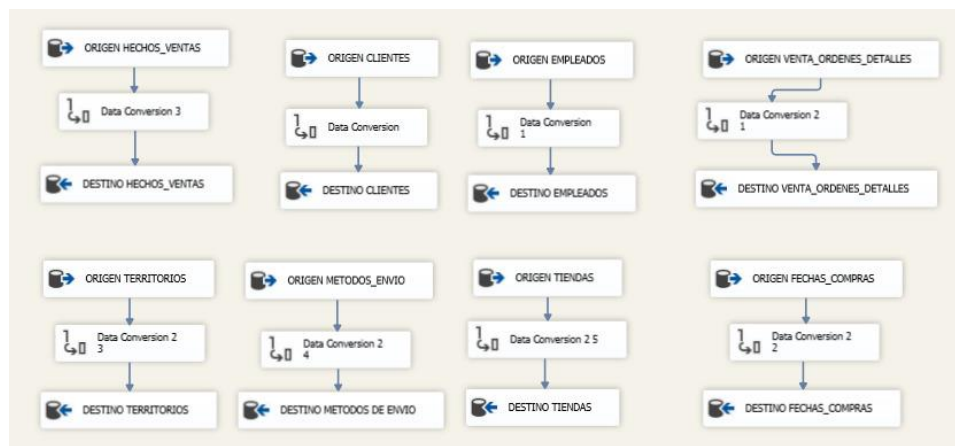
A continuación se describen los pasos para la creación de los ETL:

- ✚ El primer paso es obtener la consulta SQL que se encargan de obtener los datos que se desean mostrar en la base de datos OLAP.

*En la carpeta del proyecto se adjunta un archivo que contiene todas las consultas SQL necesarias para la creación de los ETL.

- ✚ Una vez estructurada la consulta se ejecuta SQL server data tolos y se inicia un nuevo proyecto. Ahora se debe agregar tres componentes:
 - ✓ OLE DB SOURCE: la función de este componente es obtener los datos de la base de datos OLTP en base a la consulta mencionada anteriormente.
 - ✓ DATA CONVERSION: Este componente es el encargado de hacer la conversión del tipo de dato de la base de datos OLTP a la base de datos OLAP.
 - ✓ OLE DB DESTINATION: aquí ocurre el proceso de carga, es decir, los datos ya con el tipo de dato correcto son cargados a la base de datos OLAP.



Los tres componentes anteriormente descritos deben ir unidos entre si y deben pertenecer a un mismo ETL.



CUBO OLAP

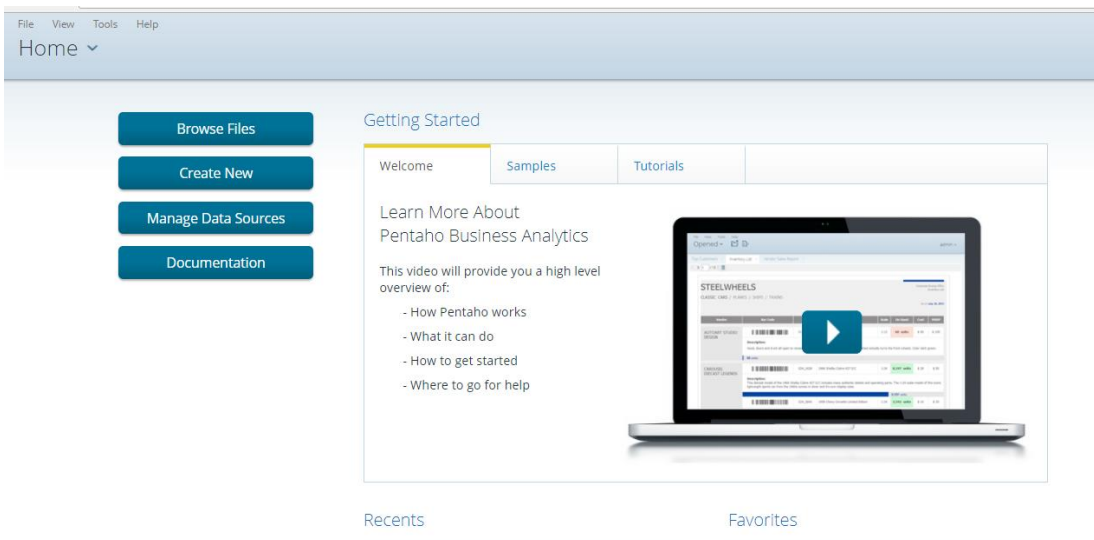
Un **cubo OLAP** es una estructura de datos que supera las limitaciones de las bases de datos relacionales y proporciona un análisis rápido de datos. Los **cubos** pueden mostrar y sumar grandes cantidades de datos, a la vez que proporcionan a los usuarios acceso mediante búsqueda a los puntos de datos.

Herramientas:

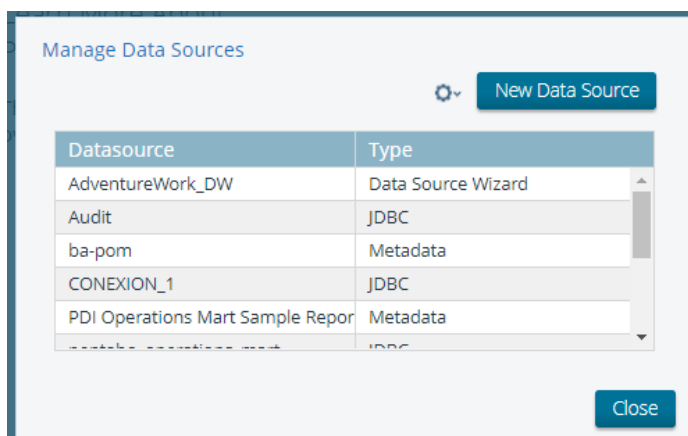
-  Sql Server Enterprise Edición 2014
-  Pentaho

Creación del cubo OLAP

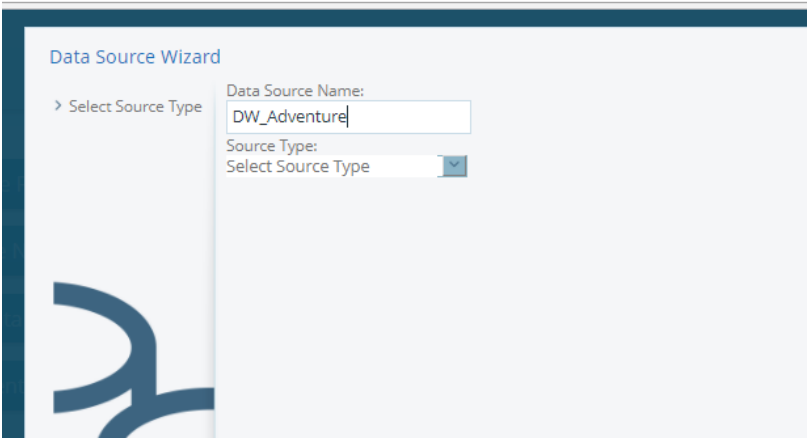
Inicio de pentaho después de haber iniciado sesión como “admin”



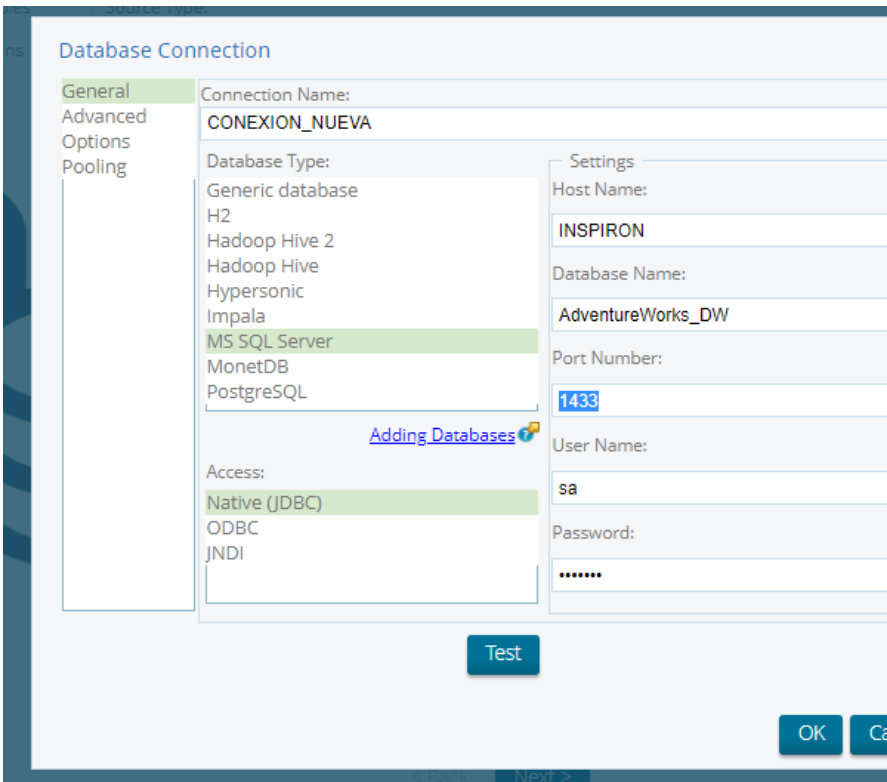
Conexión con la base de datos



Nombre del data source



Conexión con la base de AdventureWork



Tablas a utilizar y aquí se declara cual es la tabla de hechos

Data Source Wizard

Select Source Type: Select one table to finish or select multiple tables and click Next to define their joins.

Select Tables: Define Joins

Schema: **dbo**

Available Tables:

- 'dbo'. "MSpeer_conflictdetectionconfigrequest"
- 'dbo'. "MSpeer_conflictdetectionconfigresponse"
- 'dbo'. "MSpeer_jsns"
- 'dbo'. "MSpeer_originatorid_history"
- 'dbo'. "MSpeer_request"
- 'dbo'. "MSpeer_response"
- 'dbo'. "MSpeer_topologyrequest"
- 'dbo'. "MSpeer_topologyresponse"
- 'dbo'. "MSpub_identity_range"
- 'dbo'. "syncobj_0x3144414539413642"
- 'dbo'. "syncobj_0x3236414542433945"
- 'dbo'. "syncobj_0x3242454345374430"
- 'dbo'. "syncobj_0x3335463542383242"
- 'dbo'. "syncobj_0x3533343934354331"
- 'dbo'. "syncobj_0x3637334433373730"
- 'dbo'. "syncobj_0x3646354645373842"
- 'dbo'. "syncobj_0x3646413132433039"
- 'dbo'. "syncobj_0x3744423536363035"
- 'dbo'. "syncobj_0x4136333131424531"
- 'dbo'. "syncobj_0x4144423244464146"
- 'dbo'. "syncobj_0x4239353643343239"
- 'dbo'. "svncrhi_0x4337413445374332"

Selected Tables:

- 'dbo'. "CLIENTES"
- 'dbo'. "EMPLEADOS"
- 'dbo'. "DETALLES_ORDEN_VENTA"
- 'dbo'. "FECHAS_VENTAS"
- 'dbo'. "Hechos_Ventas"
- 'dbo'. "METODOS_DE_ENVIO"
- 'dbo'. "TERRITORIOS_VENTA"
- 'dbo'. "TIENDAS"

Fact Table: **"dbo". "Hechos_Ventas"**

< Back Next > Finish Cancel

Por último los JOIN

Data Source Wizard

Select Source Type: Define how the tables join to each other. All tables must have at least one join defined.

Select Tables: Define Joins

Left Table: **"dbo". "Hechos_Ventas"**

Right Table: **"dbo". "TIENDAS"**

Key Field: **VENTA_ID**

Key Field: **ID_TIENDA**

Left Table Fields:

- ID_TIENDA
- CLIENTE_ID
- EMPLEADO_ID
- TERRITORIO_ID
- METODO_ENVIO_ID
- FECHA_ID
- DETALLES_VENTA_ID

Right Table Fields:

- NOMBRE_TIENDA

Join(s):

- 'dbo'. "Hechos_Ventas". CLIENTE_ID - INNER JOIN - "dbo". "CLIENTES". CLIENTE_ID
- 'dbo'. "Hechos_Ventas". EMPLEADO_ID - INNER JOIN - "dbo". "EMPLEADOS". EMPLEADO_ID
- 'dbo'. "Hechos_Ventas". TERRITORIO_ID - INNER JOIN - "dbo". "TERRITORIOS_VENTA". TERRITORIO_ID
- 'dbo'. "Hechos_Ventas". METODO_ENVIO_ID - INNER JOIN - "dbo". "METODOS_DE_ENVIO". METODO_ENVIO_ID
- 'dbo'. "Hechos_Ventas". DETALLES_VENTA_ID - INNER JOIN - "dbo". "DETALLES_ORDEN_VENTA". DETALLES_VENTA_ID

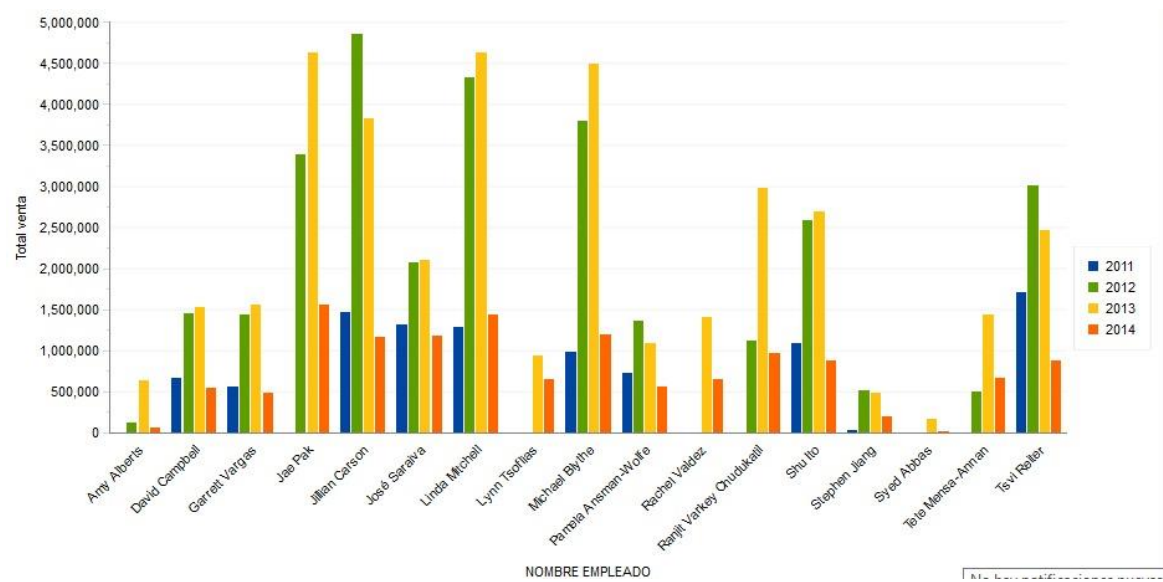
Create Join Delete Join

REPORTES

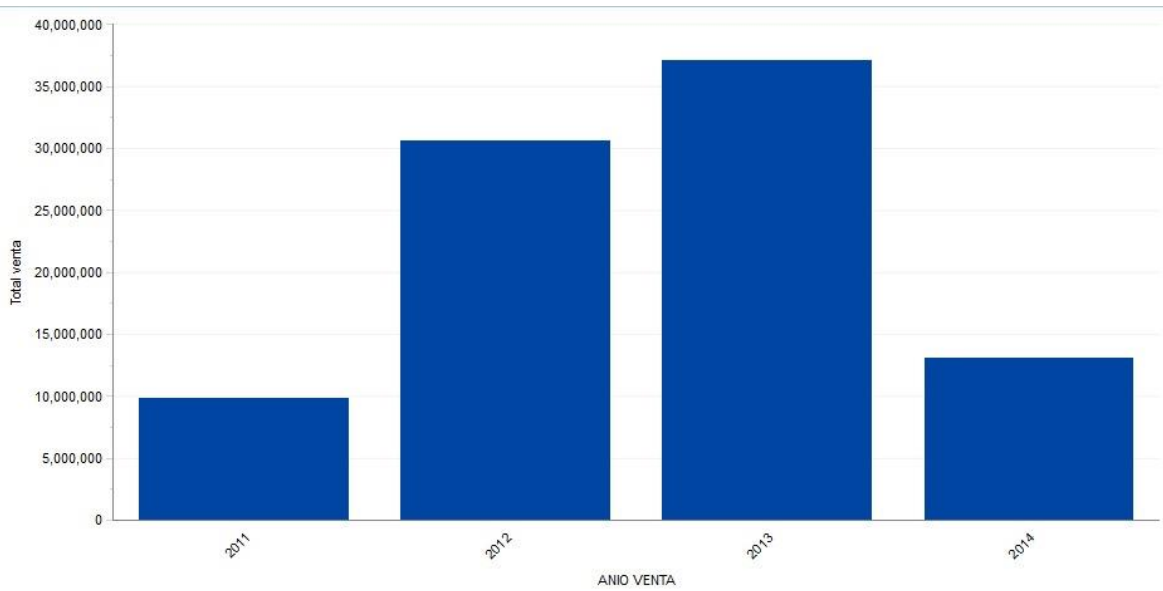
Los **reportes** son informes que organizan y exhiben la información contenida en una **base de datos**.

	AÑO VENTA			
	2011	2012	2013	2014
NOMBRE EMPLEADO	Total venta	Total venta	Total venta	Total venta
Amy Alberts	-	130667	632003	63748
David Campbell	675664	1453583	1525864	552784
Garrett Vargas	563327	1448036	1567162	490898
Jae Pak	-	3396243	4630339	1558542
Jillian Carson	1477158	4865651	3827980	1171596
José Saraiva	1323329	2070323	2112546	1177338
Linda Mitchell	1294820	4328860	4637218	1434121
Lynn Tsofilias	-	-	946893	659548
Michael Blythe	986298	3806298	4490942	1191828
Pamela Ansman-Wolfe	730273	1361053	1087554	569365
Rachel Valdez	-	-	1407494	654899
Ranjit Varkey Chudukatil	-	1122919	2987334	977725
Shu Ito	1089874	2598220	2695241	876233
Stephen Jiang	32568	516197	485881	201289
Syed Abbas	-	-	171607	23922
Tete Mensa-Annan	-	497285	1433786	677045
Tsvi Reiter	1713641	3018750	2467043	886639

REPORTE VENTAS_X_EMPLEADO_AÑO



REPORTE EMPLEADOS_VENTA_AÑO GRAFICO



REPORTE_VENTAS_AÑO

REPLICACION

Es una solución para aumentar la disponibilidad de la base de datos en la cual se tiene la base de datos principal en donde se guardan todos los datos de todas las transacciones de la empresa. Toda la información existente en la base de datos principal se copia fielmente a una o varias bases de datos secundarias, las cuales se tienen de respaldo en caso de que exista algún error en la base de datos principal.

*En el siguiente enlace se muestra un video para la configuración de la replicación
<https://drive.google.com/open?id=1EFy28SUcRpWMCY1AEDmdZkhL8si3IXig>

JOBS

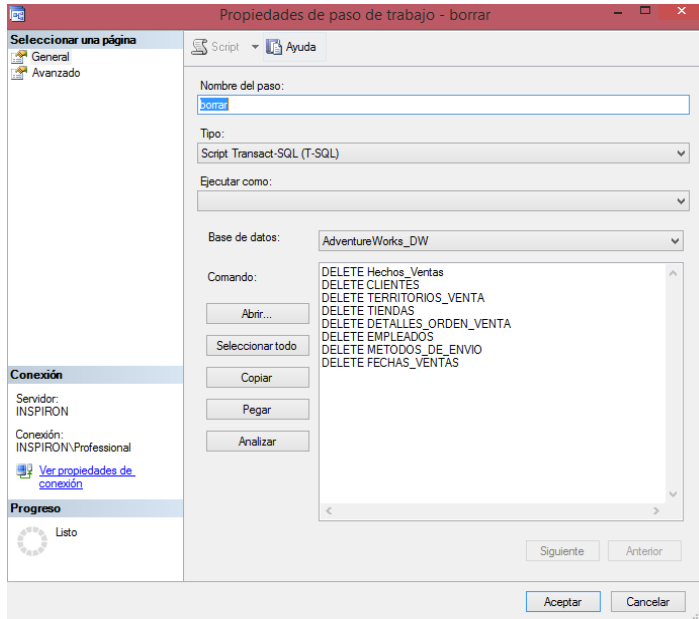
El Agente SQL Server es un producto de Microsoft que permite a los administradores de red automatizar procesos de bases de datos. El Agente SQL Server está incluido con el motor de la base de datos, por lo que los administradores pueden establecer procesos como copias de seguridad, importaciones, repeticiones y otra clase de mantenimiento de la base de datos. El Agente SQL Server también inicia el motor de la base de datos cada vez que el administrador da arranque al servidor

Para la Actualización del DataWarehouse se crearon 2 JOBS

1. vaciar la Base de datos
2. Cargar nuevamente los registros actualizados

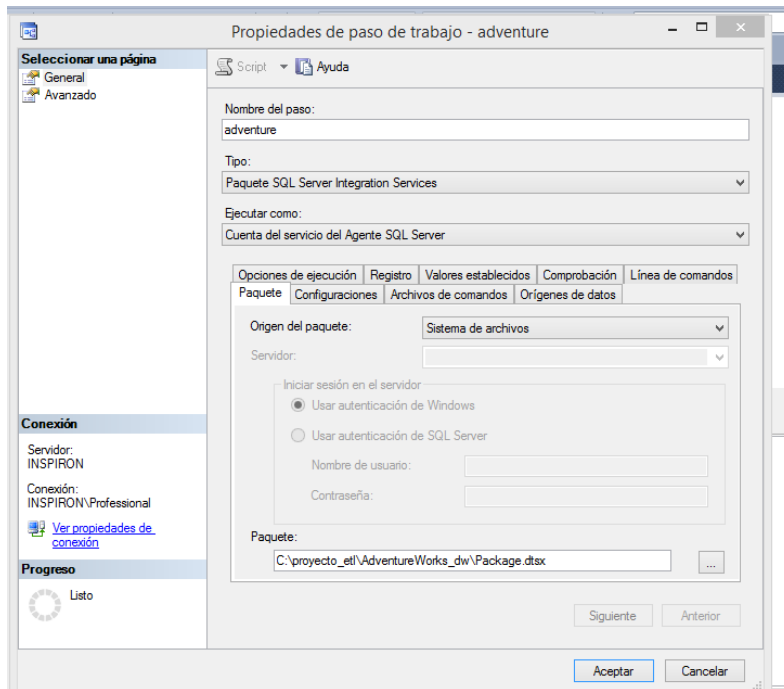
***Estos con el fin de actualizar todos los datos posibles**

JOBS BORRAR



JOBS ACTUALIZAR

Se muestra el Jobs que ejecuta el paquete de los ETL



Estos JOBS se ejecutan en una hora donde no perjudique las transacciones diarias