



UNIVERSIDAD
NACIONAL
AUTÓNOMA DE
NICARAGUA,
MANAGUA
UNAN - MANAGUA

Datawarehouse de NorthWind

Integrantes: Rodrigo Joao López Marengo.

Josefina del Carmen López Cruz.

Obeth Joshua Aburto López.

Asignatura: Principios de Inteligencia de Negocios

Docente: Msc. Ing. Manuel Mojica.

Índice

Introducción	3
Justificación	4
Descripción de caso de uso a desarrollar	5
Objetivos	7
Metodología y software	9
Análisis de Resultados	10
Anexos	11

Introducción

El presente trabajo tiene como objetivo principal utilizar la base de datos de NorthWind proporcionada por nuestro maestro para la creación de un datawarehouse. El datawarehouse permitirá consolidar y organizar de manera eficiente toda la información relacionada con NorthWind, con el fin de facilitar un análisis de datos y toma de decisiones estratégicas de manera eficiente. Mediante el uso de técnicas de extracción, transformación y carga de datos (ETL) esto con el fin de proporcionar una visión completa y consolidada de NorthWind al igual que la realización de consultas, análisis y generación de reportes.

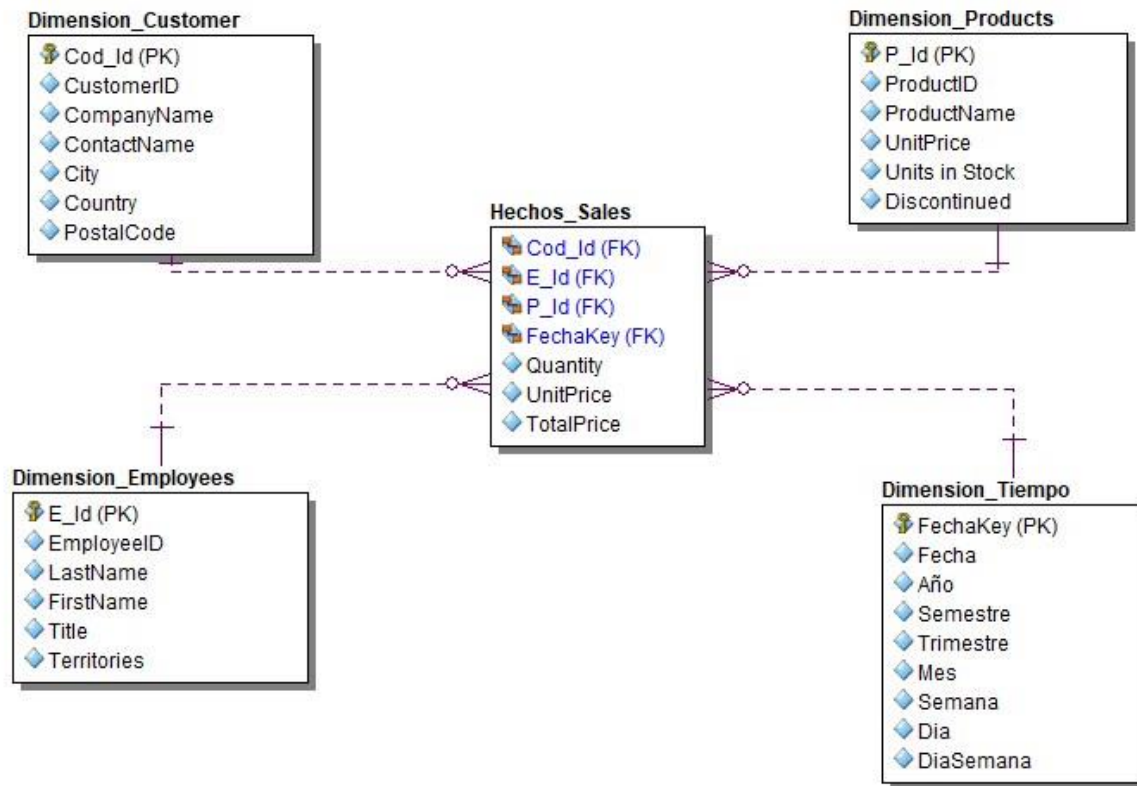
Justificación

La implementación de un almacén de datos y un cubo utilizando la base de datos NorthWind se justifica por varias razones importantes. En primer lugar, permite desarrollar competencias en tecnologías de ingeniería de datos, lo que es esencial en el ámbito actual de gestión y análisis de datos. Este proyecto proporciona una oportunidad para adquirir habilidades valiosas y altamente demandadas en el mercado laboral.

Además, trabajar con la base de datos NorthWind como fuente de datos real y estructurada brinda una experiencia práctica y concreta. Al utilizar datos existentes, se pueden aplicar conceptos y principios de ingeniería de datos en un entorno realista. Esto facilita el aprendizaje y la consolidación de conocimientos, y permite comprender cómo aplicar estas tecnologías en situaciones reales.

Por último, la implementación del almacén de datos y el cubo con la base de datos NorthWind mejora la toma de decisiones al proporcionar acceso a información más completa y precisa. Al tener un conjunto de datos centralizado y optimizado, se pueden realizar análisis y consultas complejas para identificar patrones, tendencias y oportunidades de negocio. Esto respalda la toma de decisiones informada y estratégica, lo que puede tener un impacto positivo en el rendimiento y la competitividad de una organización.

Descripción de caso de uso a desarrollar



El script de creación de tablas define las siguientes tablas en la base de datos DWNorthwind:

✚ **Tabla Dimension_Customers:** Esta tabla almacena información sobre los clientes.

Contiene columnas como Customers_Id (identificador único del cliente), Cod_Id (código de identificación del cliente), CompanyName (nombre de la empresa del cliente), ContactName (nombre del contacto del cliente), City (ciudad del cliente), Country (país del cliente) y PostalCode (código postal del cliente). La clave primaria es Customers_Id.

✚ **Tabla Dimension_Employees:** Esta tabla almacena información sobre los empleados.

Tiene columnas como Employee_Id (identificador único del empleado), LastName (apellido del empleado), FirstName (nombre del empleado), Title (cargo del empleado) y Territories (territorios asignados al empleado). La clave primaria es Employee_Id.

✚ **Tabla Dimension_Products:** Esta tabla almacena información sobre los productos. Incluye columnas como Product_Id (identificador único del producto), ProductName (nombre del producto), UnitPrice (precio unitario del producto), Units in Stock (unidades en stock del producto) y Discontinued (indicador de si el producto está descontinuado). La clave primaria es Product_Id.

✚ **Tabla Dimension_Tiempo:** Esta tabla es una dimensión de tiempo y almacena información relacionada con las fechas. Contiene columnas como FechaKey (clave única de la fecha), Fecha (fecha completa), Año (año), Semestre (semestre), Trimestre (trimestre), Mes (mes), Semana (semana), Dia (día) y DiaSemana (día de la semana). La clave primaria es FechaKey.

✚ **Tabla Hechos_Sales:** Esta tabla almacena información sobre las ventas. Contiene columnas como Employee_Id (identificador del empleado relacionado con la venta), Product_Id (identificador del producto vendido), FechaKey (clave de tiempo relacionada con la venta), Customers_Id (identificador del cliente relacionado con la venta), Quantity (cantidad vendida), UnitPrice (precio unitario del producto en la venta) y TotalPrice (precio total de la venta). Las claves foráneas se refieren a las tablas de dimensiones correspondientes.

En resumen, estas tablas forman parte del diseño de un almacén de datos para el negocio relacionado con la gestión de pedidos y ventas. La implementación del almacén de datos busca mejorar la capacidad de análisis y reporte de datos para tomar decisiones informadas en áreas como la gestión de clientes, empleados, productos y ventas.

Objetivos

Los objetivos que se desean desarrollar con la implementación del almacén de datos en el negocio relacionado con la gestión de pedidos y ventas son los siguientes:

1. **Mejorar la capacidad de análisis:** El almacén de datos permitirá consolidar y organizar los datos relacionados con clientes, empleados, productos y ventas en una estructura optimizada para análisis. El objetivo es facilitar el acceso y la consulta de datos, lo que permitirá realizar análisis más profundos y completos. Esto ayudará a identificar tendencias, patrones y oportunidades de negocio, lo que a su vez respaldará la toma de decisiones informadas y estratégicas.
2. **Optimizar la generación de informes:** Con el almacén de datos, se busca mejorar la generación de informes al contar con datos preprocesados y estructurados. En pocas palabras el objetivo es agilizar y simplificar el proceso de generación de informes, proporcionando información relevante y precisa en tiempo real. Esto permitirá a los usuarios obtener rápidamente los datos necesarios para evaluar el rendimiento del negocio, realizar análisis comparativos y tomar decisiones basadas en datos.
3. **Facilitar el análisis de la gestión de clientes:** Uno de los objetivos específicos es mejorar el análisis de la gestión de clientes. El almacén de datos permitirá analizar el comportamiento de los clientes, identificar segmentos de clientes más rentables, realizar análisis de lealtad y evaluar el impacto de las estrategias de marketing y ventas en la satisfacción del cliente. El objetivo final es mejorar la gestión de relaciones con los clientes y maximizar el valor de cada interacción con ellos.

4. **Optimizar la gestión de inventario y productos:** Otro objetivo es mejorar la gestión de inventario y productos mediante el análisis de datos. El almacén de datos proporcionará información detallada sobre las ventas, el rendimiento de los productos, las tendencias de demanda y otros factores clave. Esto permitirá optimizar los niveles de inventario, identificar productos más rentables, predecir la demanda futura y tomar decisiones más acertadas en cuanto a la introducción de nuevos productos o la eliminación de productos discontinuados.

Metodología y software

Tomamos la metodología Barquín 1 para nuestro proyecto con la base de datos NorthWind debido a sus beneficios y enfoque estructurado. Esta metodología nos permite realizar un análisis detallado de los requerimientos del negocio y definir claramente los objetivos del proyecto. Además, nos ayuda a diseñar un modelo de datos adecuado y a identificar las fuentes de datos relevantes en la base de datos NorthWind. A través de la metodología Barquín 1, podemos desarrollar un proceso de extracción, transformación y carga (ETL) eficiente para obtener datos precisos y coherentes que respalden nuestro sistema de BI en conjunto con herramientas como Visual Studio, SQL Server y ER Studio, es una combinación poderosa que garantiza el éxito de nuestro proyecto de BI con la base de datos NorthWind.

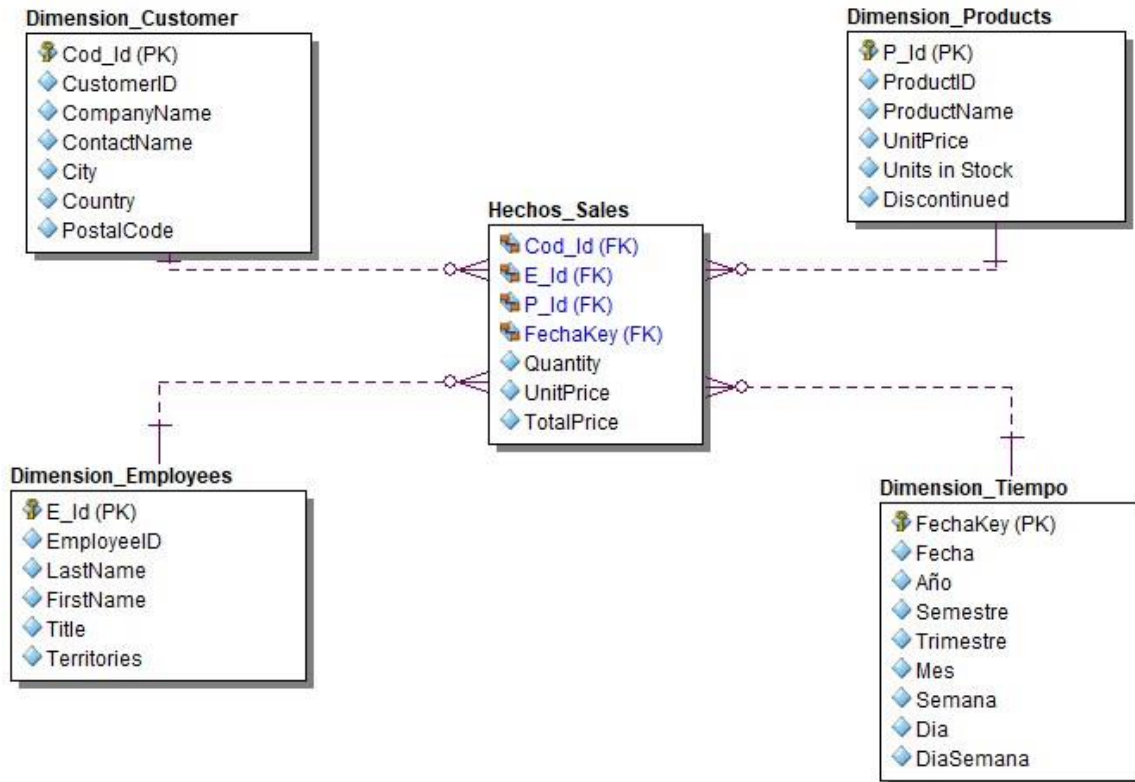
Análisis de Resultados

1. Inicialmente, se efectuó la correcta desnormalización de la base de datos original, Northwind, además del llenado de la misma.
2. Luego de tener como base la base de datos original, empezamos a diseñar, el diagrama Entidad-Relación del Datawarehouse, siguiendo como objetivo, el diseño en forma de estrella, para optimización de consultas al servidor, y correcta implementación sugerida por la documentación facilitada por el docente.
3. Ya teniendo el diagrama terminado, se usaron los metadatos de esto, para la correcta implementación a código SQL, y así usarlo avanzar al siguiente paso.
4. El siguiente paso es la creación de índices en las tablas del datawarehouse de Northwind, se crearon, siguiendo las indicaciones aprendidas anteriormente en clases, y luego de eso, se empezó a realizar el respectivo llenado de cada una de las tablas.
5. Teniendo el llenado y índices, empezamos con consultas y trabajos más extensos, como lo son la creación de las dimensiones lentamente cambiantes, en base al datawarehouse de Northwind.
6. Ya siguiendo con el proceso, ya teniendo instalados los programas y extensiones requeridas, empezamos a realizar el proceso de ETL en Visual Studio, con la carga y descarga de datos de origen y salida.
7. Ya teniendo correctamente implementados los pasos anteriormente mencionados, se empezó con la creación del cupo de datos OLAP, finalizado con el correcto proceso de transacciones de datos.

N/T: Las descripciones graficas de los resultados se encuentran en la parte de anexos.

Anexos

Diagrama ER Datawarehouse Northwind



Creación de Datawarehouse – SQL Server

```
DWNorthwind.sql - UIN3350,DELL (62) | SQLQuery1.sql - DE...UIN3350,DELL (61)
13 USE master
14 go
15 CREATE DATABASE DWNorthwind
16 go
17 USE DWNorthwind
18 go
19 /*
20 * TABLE: Dimension_Customer
21 */
22
23 CREATE TABLE Dimension_Customer(
24     Cod_Id int IDENTITY(1,1),
25     CustomerID nchar(10) NOT NULL,
26     CompanyName nvarchar(50) NOT NULL,
27     ...
28 )
29
30 CREATE TABLE Dimension_Employees(
31     E_Id int IDENTITY(1,1),
32     EmployeeID int NOT NULL,
33     LastName nchar(50) NOT NULL,
34     FirstName nchar(50) NOT NULL,
35     Title nchar(50) NOT NULL,
36     Territories nchar(50) NOT NULL
37 )
38
39 CREATE TABLE Dimension_Products(
40     P_Id int IDENTITY(1,1),
41     ProductID int NOT NULL,
42     ProductName nchar(50) NOT NULL,
43     UnitPrice float NOT NULL,
44     Units in Stock int NOT NULL,
45     Discontinued bit NOT NULL
46 )
47
48 CREATE TABLE Dimension_Tiempo(
49     FechaKey int IDENTITY(1,1),
50     Fecha datetime NOT NULL,
51     Año int NOT NULL,
52     Semestre int NOT NULL,
53     Trimestre int NOT NULL,
54     Mes int NOT NULL,
55     Semana int NOT NULL,
56     Día int NOT NULL,
57     DíaSemana int NOT NULL
58 )
59
60 CREATE TABLE Hechos_Sales(
61     Cod_Id int FOREIGN KEY REFERENCES Dimension_Customer(Cod_Id),
62     E_Id int FOREIGN KEY REFERENCES Dimension_Employees(E_Id),
63     P_Id int FOREIGN KEY REFERENCES Dimension_Products(P_Id),
64     FechaKey int FOREIGN KEY REFERENCES Dimension_Tiempo(FechaKey),
65     Quantity int NOT NULL,
66     UnitPrice float NOT NULL,
67     TotalPrice float NOT NULL
68 )
69
70 GO
71
72 Messages
73
74 <<< CREATED TABLE Dimension_Customer >>>
75 <<< CREATED TABLE Dimension_Employees >>>
76 <<< CREATED TABLE Dimension_Products >>>
77 <<< CREATED TABLE Dimension_Tiempo >>>
78 <<< CREATED TABLE Hechos_Sales >>>
79
80 Completion time: 2023-07-10T22:47:57.7059935-06:00
81
82 110 %
83
84 Query executed successfully. | DESKTOP-UIN3350 (15.0 RTM) | DESKTOP-UIN3350,DELL (62) | DWNorthwind | 00:00:02 | 0 rows
```

Indices Datawarehouse Northwind – SQL Server

```
Indices.sql - DESKTOP-UIN3350\DELL (65)  DWNNorthwind.sql - ...UIN3350\DELL (62)  SQLQuery1.sql - DE...UIN3350\DELL (61)
13
14 /*Indice Customers*/
15 Create Clustered Index Ix_Customer_Company ON dbo.Dimension_Customer (CompanyName)
16 GO
17 Create NonClustered Index Ix_Customer_City ON dbo.Dimension_Customer (City)
18 GO
19 Create NonClustered Index Ix_Customer_Country ON dbo.Dimension_Customer (Country)
20 GO
21
22 /*Indice Employees*/
23 Create Clustered Index Ix_Employees_FirstName ON dbo.Dimension_Employees (FirstName)
24 GO
25 Create NonClustered Index Ix_Employees_Title ON dbo.Dimension_Employees (Title)
26 GO
```

110 %

Messages

Commands completed successfully.

Completion time: 2023-07-10T22:48:51.8266061-06:00

Query executed successfully. DESKTOP-UIN3350 (15.0 RTM) DESKTOP-UIN3350\DELL (65) DWNNorthwind 00:00:00 0 rows

Llenado Datawarehouse Northwind – SQL Server

```
LlenadoDW.sql - DE...UIN3350\DELL (58)  SQLQuery1.sql - DE...UIN3350\DELL (61)
129 /*Llenado de hechos*/
130 INSERT INTO Hechos_Sales ( Cod_Id,E_Id,P_Id, FechaKey, Quantity, UnitPrice, TotalPrice)
131 SELECT
132     DC.Cod_Id,
133     DE.E_Id,
134     P.ProductID,
135     DT.FechaKey,
136     OD.Quantity,
137     P.UnitPrice,
138     OD.Quantity * P.UnitPrice AS TotalPrice
139 FROM Northwind.dbo.Orders o
140 INNER JOIN Northwind.dbo.[Order Details] OD ON o.OrderID = OD.OrderID
141 INNER JOIN Northwind.dbo.Products P ON OD.ProductID = P.ProductID
142 INNER JOIN Dimension_Employees DE ON o.EmployeeID = DE.EmployeeID
```

110 %

Results

	Cod_Id	CustomerID	CompanyName	ContactName	City	Country
1	1	ALFKI	Alfreds Futterkiste	Maria Anders	Berlin	Germany
2	2	ANATR	Ana Trujillo Emparedados y helados	Ana Trujillo	México D.F.	Mexico
3	3	ANTON	Antonio Moreno Taquería	Antonio Moreno	México D.F.	Mexico
4	4	AROUT	Around the Horn	Thomas Hardy	London	UK
5	5	BERGS	Berglunds snabbköp	Christina Berglund	Luleå	Sweden
6	6	BLAUS	Blauer See Delikatessen	Hanna Moos	Mannheim	Germany
7	7	BLONP	Blondee d'al père et fils	Frédérique Citeaux	Strasbourg	France
8	8	BOLID	Bólido Comidas preparadas	Martín Sommer	Madrid	Spain

	E_Id	EmployeeID	LastName	FirstName	Title	Territories
1	3	2	Fuller	Andrew	Vice President, Sales	Westboro
2	4	2	Fuller	Andrew	Vice President, Sales	Bedford

Query executed successfully. DESKTOP-UIN3350 (15.0 RTM) DESKTOP-UIN3350\DELL (58) DWNNorthwind 00:00:26 21,304 rows

Llenado solo mostrando datos afectados

```
LlenadoDW.sql - DE...UIN3350\DELL (58))  SQLQuery1.sql - DE...UIN3350\DELL (61))
129 /*Llenado de hechos*/
130 INSERT INTO Hechos_Sales ( Cod_Id,E_Id,P_Id, FechaKey, Quantity, UnitPrice, TotalPrice)
131 SELECT
132     DC.Cod_Id,
133     DE.E_Id,
134     P.ProductID,
135     DT.FechaKey,
136     OD.Quantity,
137     P.UnitPrice,
138     OD.Quantity * P.UnitPrice AS TotalPrice
139 FROM Northwind.dbo.Orders o
140 INNER JOIN Northwind.dbo.[Order Details] OD ON o.OrderID = OD.OrderID
141 INNER JOIN Northwind.dbo.Products P ON OD.ProductID = P.ProductID
142 INNER JOIN Dimension_Employees DE ON o.EmployeeID = DE.EmployeeID
```

Results Messages

(91 rows affected)
(91 rows affected)
(49 rows affected)
(49 rows affected)
Se cambió la configuración de idioma a Español.
(1 row affected)
(1 row affected)

Creación e Implementación del Merge – SQL Server

```
Merge.sql - DESKTO...UIN3350\DELL (59))  LlenadoDW.sql - DE...UIN3350\DELL (58))  SQLQuery1.sql - DE...UIN3350\DELL (61))
1 MERGE INTO Hechos_Sales AS destino
2 USING (
3     SELECT
4         DC.Cod_Id,
5         DE.E_Id,
6         P.ProductID,
7         DT.FechaKey,
8         OD.Quantity,
9         P.UnitPrice,
10        OD.Quantity * P.UnitPrice AS TotalPrice
11 FROM Northwind.dbo.Orders o
12 INNER JOIN Northwind.dbo.[Order Details] OD ON o.OrderID = OD.OrderID
13 INNER JOIN Northwind.dbo.Products P ON OD.ProductID = P.ProductID
14 INNER JOIN Dimension_Employees DE ON o.EmployeeID = DE.EmployeeID
```

Results Messages

	Cod_Id	E_Id	P_Id	FechaKey	Quantity	UnitPrice	TotalPrice
1	85	17	11	19960801	12	21	252
2	85	18	11	19960801	12	21	252
3	85	19	11	19960801	12	21	252
4	85	20	11	19960801	12	21	252
5	85	21	11	19960801	12	21	252
6	85	22	11	19960801	12	21	252
7	85	23	11	19960801	12	21	252
8	85	17	42	19960801	10	14	140
9	85	18	42	19960801	10	14	140
10	85	19	42	19960801	10	14	140
11	85	20	42	19960801	10	14	140

Query executed successfully. DESKTOP-UIN3350 (15.0 RTM) DESKTOP-UIN3350\DELL (59) DWNorthwind 00:00:27 10.129 rows

Dimensiones lentamente cambiantes –SQL Server

```
28
29
30
31 CREATE TABLE Dimension_Customer(
32     Cod_Id int IDENTITY(1,1),
33     CustomerID nchar(10) NOT NULL,
34     CompanyName nvarchar(50) NOT NULL,
35     ContactName nvarchar(50) NULL,
36     City nvarchar(50) NULL,
37     Country nvarchar(50) NULL,
38     Fecha_Inicial Datetime NOT NULL,
39     Fecha_Final Datetime NOT NULL
40     CONSTRAINT PK1 PRIMARY KEY NONCLUSTERED (Cod_Id)
41 )
```

110 %

Messages

Se están revirtiendo las transacciones no calificadas. Estimación de porcentaje de reversión: 0%.
Se están revirtiendo las transacciones no calificadas. Estimación de porcentaje de reversión: 100%.

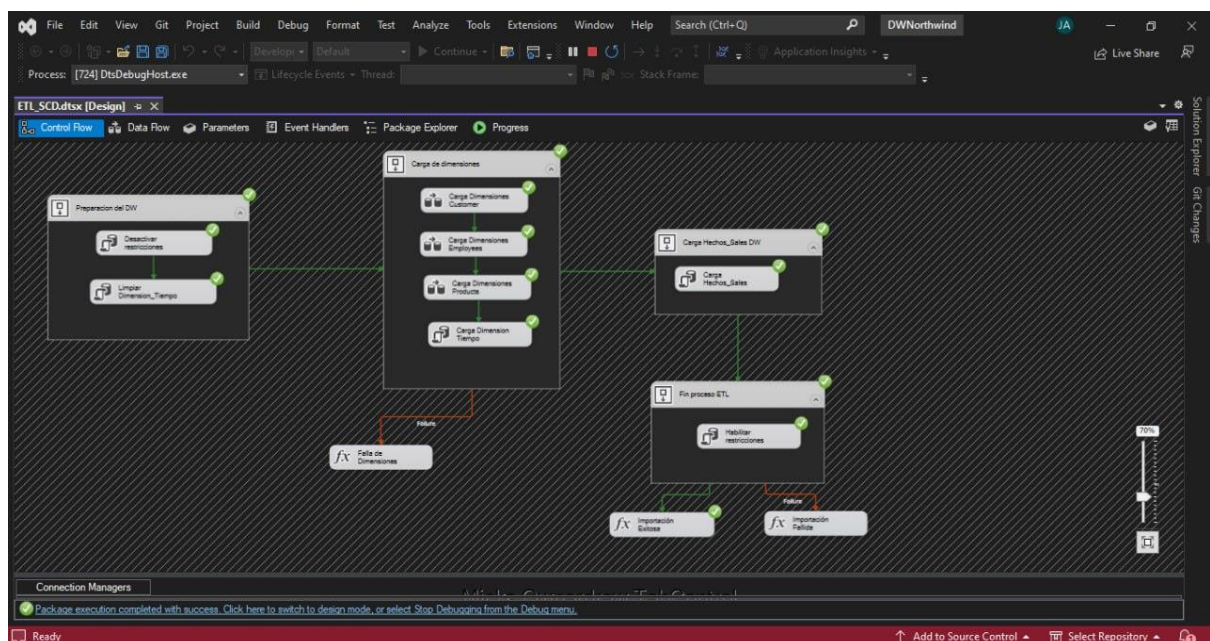
<<< CREATED TABLE Dimension_Customer >>>
<<< CREATED TABLE Dimension_Employees >>>
<<< CREATED TABLE Dimension_Products >>>
<<< CREATED TABLE Dimension_Tiempo >>>
<<< CREATED TABLE Hechos_Sales >>>

Completion time: 2023-07-10T22:54:15.7377078-06:00

110 %

Query executed successfully. | DESKTOP-UIN3350 (15.0 RTM) | DESKTOP-UIN3350\DELL (63) | DWNorthwind | 00:00:08 | 0 rows

Proceso de ETL Datawarehouse Northwind – VS



Cubo OLAP de Datawarehouse Northwind – VS



Métricas y Funciones Dashboard – Power BI

