

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HCM  
KHOA CÔNG NGHỆ THÔNG TIN

THIẾT KẾ  
MÔN HỌC: KHO DỮ LIỆU



BÁO CÁO ĐỒ ÁN  
ĐỀ TÀI  
**XÂY DỰNG KHO DỮ LIỆU CHO SÀN  
THƯƠNG MẠI ĐIỆN TỬ OLIST**

**GVHD:** Th.s Nguyễn Văn Thành

**Mã lớp:** DAWH430784\_23\_2\_02

**Nhóm sinh viên thực hiện:** Nhóm 5

1. Huỳnh Nhật Nam – 21110903
2. Nguyễn Thị Lan Anh – 21110368
3. Trịnh Thị Thanh Huyền – 21110477
4. Phan Minh Thuận – 21110667

TP. Hồ Chí Minh, tháng 05 năm 2024

## BẢNG PHÂN CÔNG

<b>STT</b>	<b>MSSV</b>	<b>Họ Và Tên</b>	<b>Nhiệm Vụ</b>	<b>Hoàn Thành</b>
1	21110903	Huỳnh Nhật Nam (Nhóm Trưởng)	Data Modeling Data integration with SSIS ( Fact Table)	100%
2	21110368	Nguyễn Thị Lan Anh	Data Modeling SQL Server Analyis Services	100%
3	21110477	Trịnh Thị Thanh Huyền	Data Modeling Data integration with SSIS ( Dimension Table)	100%
4	21110667	Phan Minh Thuận	Data Modeling Power BI	100%

# MỤC LỤC

<b>PHẦN 1: DIMENTION MODELING OLIST.....</b>	1
<b>1.1. Giới thiệu.....</b>	1
<b>1.2. Business requirement.....</b>	1
<b>1.3. Enterprise Bus Matrix .....</b>	2
<b>1.4. Detailed Bus Matrix .....</b>	2
<b>1.5. Detailed design.....</b>	2
<b>1.6. Detailed-Dimensional-Modeling-Workbook-KimballU .....</b>	3
<b>PHẦN 2: DATA INTERGRATION WITH SSIS.....</b>	9
<b>2.1. Tạo ETL project và solution .....</b>	9
<b>2.2. Thiết lập các Connection .....</b>	9
<b>2.3. Import Dimension and Fact .....</b>	14
<b>2.3.1. Date Dimension .....</b>	14
<b>2.3.2. Customers Dimension.....</b>	30
<b>2.3.3. Products Dimension.....</b>	32
<b>2.3.4. Sellers Dimension.....</b>	33
<b>2.3.5. Sales Fact .....</b>	34
<b>2.3.6. Order Fullfillment Fact.....</b>	38
<b>PHẦN 3: SQL SERVER ANALYIS SERVICES .....</b>	40
<b>3.1. Tạo một Analysis Services Project .....</b>	40
<b>3.2. Kết nối đến Analysis Services Server .....</b>	40
<b>3.3. Tạo một Data Cube .....</b>	43
<b>PHẦN 4: POWER BI .....</b>	68
<b>4.1. Tổng phí vận chuyển .....</b>	68
<b>4.2. Tổng doanh thu .....</b>	68
<b>4.3. Tổng đơn hàng .....</b>	68
<b>4.4. Tổng người bán .....</b>	69
<b>4.5. Tổng chi phí vận chuyển và tổng giá trị đơn hàng theo năm .....</b>	69
<b>4.6. Top 10 danh mục sản phẩm được mua nhiều nhất.....</b>	70

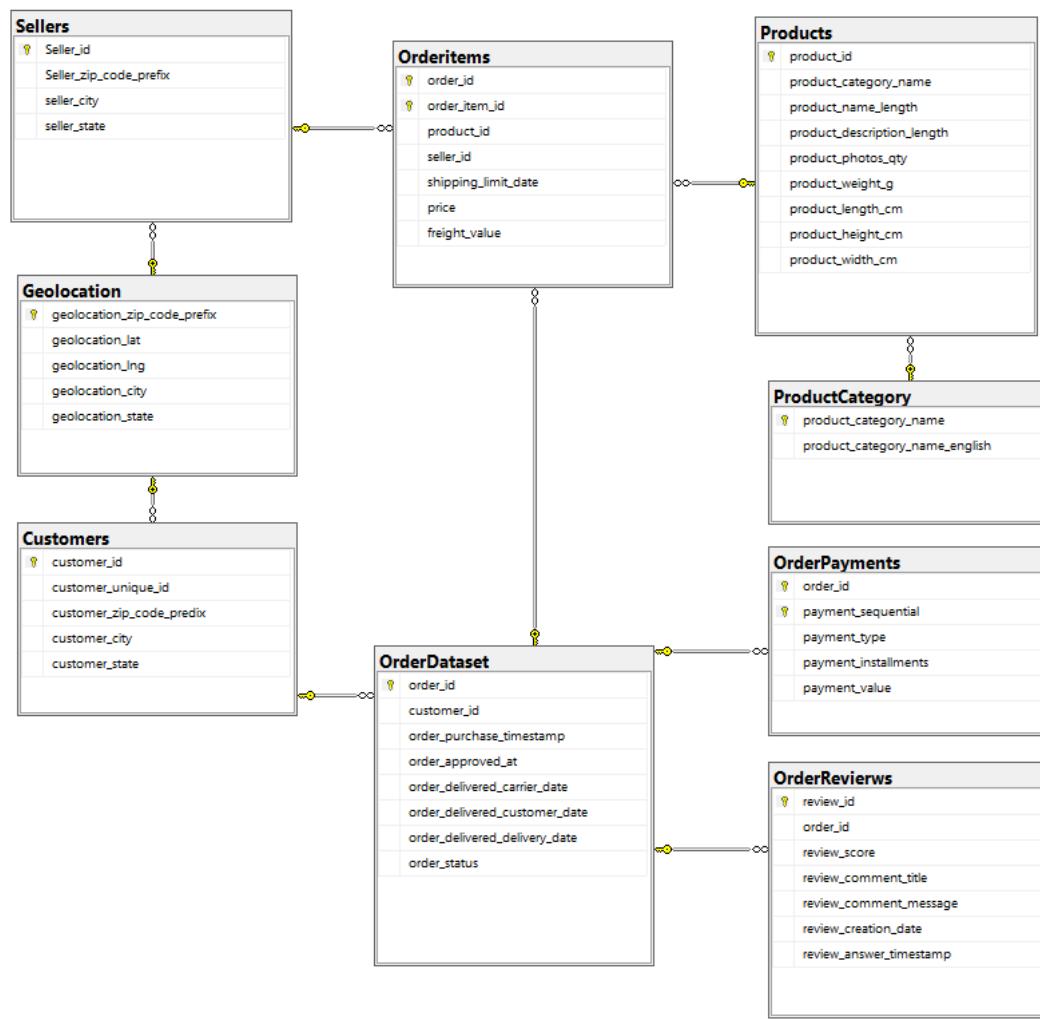
<b>4.7. Thời gian xác nhận đơn hàng trung bình theo từng tiểu bang .....</b>	<b>70</b>
<b>4.8. Số lượng đơn hàng theo loại thời gian giao hàng.....</b>	<b>71</b>
<b>4.9. Số lượng khách hàng theo năm.....</b>	<b>71</b>

## PHẦN 1: DIMENTION MODELING OLIST

### 1.1. Giới thiệu

Tập dữ liệu có thông tin về 100 nghìn đơn đặt hàng từ năm 2016 đến năm 2018 được thực hiện tại nhiều thị trường ở Brazil.

Đây là cơ sở dữ liệu sàn thương mại điện tử Olist:



### 1.2. Business requirement

Công ty Olist muốn tạo các data mart từ CSDL Olist với các mục đích sau:

- Báo cáo doanh số (Sales Reporting): Người quản lý theo dõi doanh số bán hàng theo danh mục sản phẩm (Category), khách hàng (Customers), người bán (Sellers) và thành phố đặt hàng của khách hàng (Geolocation\_city), để có thể biết được danh mục

sản phẩm nào bán chạy nhất, thành phố nào có doanh số cao nhất, người bán nào bán chạy nhất.

- Phân tích việc vận chuyển đơn hàng (Order Fulfillment and Delivery): Người quản lý muốn theo dõi, phân tích việc vận chuyển đơn hàng như thời gian giao hàng, chi phí giao hàng, tỉ lệ giao hàng đúng hạn hoặc trễ hạn, ... nhằm nâng cao chất lượng dịch vụ giao hàng.

### 1.3. Enterprise Bus Matrix

Dimension Bus. Process	Customers	Sellers	Products	Order_purchase_timestamp	order_approved_at	order_delivered_carrier_date	order_delivered_customer_date	order_delivered_estimate_date
Sales Reporting	X	X	X				X	
Order Fulfillment and Delivery	X	X		X	X	X	X	X

### 1.4. Detailed Bus Matrix

Instructions!				Customers	Sellers	Products	Order_purchase_timestamp	order_approved_at	order_delivered_carrier_date	order_delivered_customer_date	order_delivered_estimate_date
Business Process Name	Fact Table	Granularity	Facts								
Sales Reporting	FactSales	one row per order item	price, freight_value	X	X	X				X	
Order Fulfillment and Delivery	FactOrders	one row per order		X	X		X	X	X	X	X

### 1.5. Detailed design

Tham chiếu đến Attributes & Measures worksheet của High-Level-Dimensional-Modeling Excel Workbook.

Hình sau mô tả một phần của thiết kế chi tiết cho **Customers** dimension

Instructions!	Attribute / Fact Name	Description	Alternate Names	Sample Values
Dimension / Fact Table	customer_city	Customer's city		osasco
Customers	customer_state	State or province for the customer		SP

Hình sau mô tả một phần của thiết kế chi tiết cho **Products** dimension

Instructions!	Attribute / Fact Name	Description	Alternate Names	Sample Values
Products	product_category_name	Category or classification of the product.		perfumaria
Products	product_weight_g	Weight of the product in grams.		300
Products	product_length_cm	Length of the product in centimeters.		20
Products	product_height_cm	Height of the product in centimeters.		16
Products	product_width_cm	Width of the product in centimeters.		16

Hình sau mô tả một phần của thiết kế chi tiết cho Sellers dimension

Instructions!	Attribute / Fact Name	Description	Alternate Names	Sample Values
Seller	seller_city	City where the seller is located.		santo andre
Seller	seller_state	State where the seller is located.		SP

Hình sau mô tả một phần của thiết kế chi tiết cho Date dimension

Instructions!	Attribute / Fact Name	Description	Alternate Names	Sample Values
DateTime	Date	YYYY-MM-DD		
DateTime	FullDateUSA	MM/DD/YYYY		4/14/2024
DateTime	DayOfWeek	This property represents the day of the week, with values ranging from 1 (Sunday) to 7 (Saturday).	1	
DateTime	DayName	'Monday', 'Tuesday', etc.		Sunday
DateTime	DayOfMonth	This property represents the day of the month, with values ranging from 1 to 31.		14
DateTime	DayOfYear	This property represents the day of the year, with values ranging from 1 to 365.		104
DateTime	WeekOfYear	This property represents the week of the year, with values ranging from 1 to 53.		16
DateTime	MonthName	'January', 'February', etc.		April
DateTime	MonthOfYear	This property represents the month of the year, with values ranging from 1 to 12.		4
DateTime	Quarter	This property represents the quarter of the year, with values ranging from 1 to 4.		2
DateTime	QuarterName	This property represents the name of the quarter, with values such as 'Q1', 'Q2', etc.		Q2
DateTime	Year	This property represents the year, with values ranging from 1900 to 2155.		2024
DateTime	IsWeekday	This property represents whether the date is a weekday (TRUE) or a weekend (FALSE).		TRUE

## 1.6. Detailed-Dimensional-Modeling-Workbook-KimballU

### BlankDim

Table Name	BlankDimension	Home Page					
Table Type	Dimension						
Display Name	Blank						
Database Schema							
Table Description							
Comment							
Biz Filter Logic							
Size							
Generate Script?	N						
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Display Folder	ETL Rules
BlankDimKey	Blank Dim Key	Surrogate primary key	-1	1, 2, 3...	key		
BKBlankDimID	Blank Dim ID	Business key from source system (aka natural key)	-1		key		
Attribute1	Attribute1	Always describe your columns	Unk Attribute1		1		
Attribute2	Attribute2	Always describe your columns	Unk Attribute2		2		
Attribute3	Attribute3	Always describe your columns	Unk Attribute3		2		
RowIsCurrent	Row Is Current	Is this the current row for this member (Y/N)?	Y	Y, N	n/a	Exclude from cube	Standard SCD-2
RowStartDate	Row Start Date	When did this row become valid for this	1/1/1900	1/24/2011	n/a	Exclude from cube	Standard SCD-2
RowEndDate	Row End Date	When did this row become invalid? (12/31/9999 if current row)	12/31/9999	1/14/1998, 12/31/9999	n/a	Exclude from cube	Standard SCD-2
RowChangeReason	Row Change Reason	Why did the row change last?	N/A		n/a	Exclude from cube	Standard SCD-2
InsertAuditKey	InsertAuditKey	What process loaded this row?	-1		n/a	Exclude from cube	Standard Audit dim
UpdateAuditKey	UpdateAuditKey	What process most recently updated this row?	-1		n/a	Exclude from cube	Standard Audit dim

Table Name	BlankDimension	Target	Source										
Column Name	Display Name	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
BlankDimKey	Blank Dim Key	int				PK ID	N		Derived				
BKBlankDimID	Blank Dim ID	nvarchar	10				N						
Attribute1	Attribute1	nvarchar	50				N						
Attribute2	Attribute2	nvarchar	50				N						
Attribute3	Attribute3	nvarchar	50	1			N						
RowIsCurrent	Row Is Current	nchar	1				N			Derived			
RowStartDate	Row Start Date	datetime					N			Derived			
RowEndDate	Row End Date	datetime					N	12/31/9999		Derived			
RowChangeReason	Row Change Reason	nvarchar	200				N			Derived			
InsertAuditKey	InsertAuditKey	int				FK	DimAudit.AuditKey	N		Derived			
UpdateAuditKey	UpdateAuditKey	int				FK	DimAudit.AuditKey	N		Derived			

### BlankFact

A	B	C	E	G	H								
Table Name	BlankFact	Home Page											
Table Type	Fact	Always create a table description, which becomes a table extended property.											
Display Name	Blank Fact	You can put technical information here -- it doesn't get copied into the database in any way.											
Database Schema													
Table Description													
Comment													
Biz Filter Logic													
Size													
Generate Script?	N												
Column Name	Display Name	Description	Example Values	Display Folder	ETL Rules								
Dim1Key	Dim1 Key	Key to Dim1	1, 2, 3	key	Key lookup from Sales.CurrencyRate.FromCurrencyCode								
Dim2Key	Dim2 Key	Key to Dim2	1, 2, 3	key	Key lookup from Sales.CurrencyRate.CurrencyRateDate								
InsertAuditKey	Insert Audit Key	Key to Audit dimension for row insertion	1, 2, 3	key	Standard auditing								
UpdateAuditKey	Update Audit Key	Key to Audit dimension for row update	1, 2, 3	key	Standard auditing								
BKFactTable	BK Fact Table	The natural key for the fact table, if any (eg order number)		Exclude from cube									
Fact1	Fact 1	Always add a description!			Counts & rates								
Fact2	Fact 2	Always add a description!			Counts & rates								
Fact3	Fact 3	Always add a description!			Amounts								
Table Name	BlankFact												
Table Type	Fact												
Display Name	Blank Fact												
Database Schema													
Table Description													
Comment													
Biz Filter Logic													
Size													
Generate Script?	N												
Column Name	Display Name	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source	Source Schema	Source Table	Source Field Name	Source
Dim1Key	Dim1 Key	smalint				FK	Dim1.Dim1Key	N	ETL Process				
Dim2Key	Dim2 Key	int				FK	Dim2.Dim2Key	N	ETL Process				
InsertAuditKey	Insert Audit Key	int				FK	DimAudit.AuditKey	N	ETL Process				
UpdateAuditKey	Update Audit Key	int				FK	DimAudit.AuditKey	N	ETL Process				
BKFactTable	BK Fact Table	int						Y					
Fact1	Fact 1	float						Y					
Fact2	Fact 2	float						Y					
Fact3	Fact 3	money						Y					

## DimAudit

A	B	C	E	G	H
Table Name	DimAudit	Home Page			
Table Type	Audit	Audit dimension tags each data row with the the process that added or updated it.			
Display Name	Audit				
Database Schema					
Table Description					
Comment					
Biz Filter Logic					
Size		one row per execution of a package			
Generate Script?	N				
Column Name	Display Name	Description	Example Values	Unknown Member	
AuditKey	AuditKey	Surrogate primary key	-1	1, 2, 3...	
ParentAuditKey	ParentAuditKey	Foreign key to self, to identify calling package execution	-1	1, 2, 3...	
TableName	TableName	Name of the main table loaded by this package	Audit		
PkgName	PkgName	Name of the SSIS package	None: Dummy row		
PkgGUID	PkgGUID	Identifier for the package			
PkgVersionGUID	PkgVersionGUID	Identifier for the package version			
PkgVersionMajor	PkgVersionMajor	Major version number for the package			
PkgVersionMinor	PkgVersionMinor	Minor version number for the package			
ExecStartDT	ExecStartDT	Date-time the package started executing			
ExecStopDT	ExecStopDT	Date-time the package finished executing			
ExecutionInstanceGUID	ExecutionInstanceGUID	Identifier for the execution of the package			
ExtractRowCnt	ExtractRowCnt	Count of rows extracted from the source(s)			
InsertRowCnt	InsertRowCnt	Count of rows inserted in the destination table			
UpdateRowCnt	UpdateRowCnt	Count of rows updated in the destination table			
ErrorRowCnt	ErrorRowCnt	Count of error rows			
TableInitialRowCnt	TableInitialRowCnt	Count of rows in target table before we begin			
TableFinalRowCnt	TableFinalRowCnt	Count of rows in target table after package ends			
TableMaxSurrogateKey	TableMaxSurrogateKey	Maximum surrogate key value in table (if we're maintaining ourselves)			
SuccessfulProcessingInd	SuccessfulProcessingInd	Did the package finish executing successfully?	Y	Y, N	

Table Name		DimAudit								
Table Type		Audit								
Display Name		Audit dimension table								
Database Schema		Audit dimension table								
Table Description		Audit dimension table								
Comment		Audit dimension table								
Biz Filter Logic		Audit dimension table								
Size		Audit dimension table								
Generate Script?		Audit dimension table								
Column Name		Display Name	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source System
AuditKey	AuditKey	AuditKey	int			PK	ID	N		Derived
ParentAuditKey	ParentAuditKey	ParentAuditKey	int			FK	DimAudit.AuditKey	N	Unknown	Derived
TableName	TableName	TableName	nvarchar	50				N	Unknown	
PkgName	PkgName	PkgName	nvarchar	50				Y		
PkgGUID	PkgGUID	PkgGUID	uniqueidentifier					Y		
PkgVersionGUID	PkgVersionGUID	PkgVersionGUID	uniqueidentifier					Y		
PkgVersionMajor	PkgVersionMajor	PkgVersionMajor	smallint					Y		
PkgVersionMinor	PkgVersionMinor	PkgVersionMinor	smallint					Y		
ExecStartDT	ExecStartDT	ExecStartDT	datetime					N	getdate()	
ExecStopDT	ExecStopDT	ExecStopDT	datetime					Y		
ExecutionInstanceGUID	ExecutionInstanceGUID	ExecutionInstanceGUID	uniqueidentifier					Y		
ExtractRowCnt	ExtractRowCnt	ExtractRowCnt	bigint					Y		
InsertRowCnt	InsertRowCnt	InsertRowCnt	bigint					Y		
UpdateRowCnt	UpdateRowCnt	UpdateRowCnt	bigint					Y		
ErrorRowCnt	ErrorRowCnt	ErrorRowCnt	bigint					Y		
TableInitialRowCnt	TableInitialRowCnt	TableInitialRowCnt	bigint					Y		
TableFinalRowCnt	TableFinalRowCnt	TableFinalRowCnt	bigint					Y		
TableMaxSurrogateKey	TableMaxSurrogateKey	TableMaxSurrogateKey	bigint					Y		
SuccessfulProcessingInd	SuccessfulProcessingInd	SuccessfulProcessingInd	nchar	1				N	N	

## DimCustomers

Table Name		DimCustomers								
Table Type		Dimension								
Display Name		Customers								
Database Schema										
Table Description										
Comment										
Biz Filter Logic										
Size										
Generate Script?		Y								
Column Name		Display Name	Description			Unknown Member	Example Values		SCD Type	Display Folder
customerKey	customerKey	customerKey	Surrogate primary key			-1	1, 2, 3		key	
customer_id	customer_id	customer_id	Business key from source system (aka natural key)				00012a2ce68dcda20d059ce		key	
customer_city	customer_city	customer_city	Customer's city				98491703			
customer_state	customer_state	customer_state	State or province for the customer				osasco		2	
RowIsCurrent	Row Is Current	Row Is Current	Is this the current row for this member (Y/N)?				Y, N		n/a	Exclude from cube
RowStartDate	Row Start Date	Row Start Date	When did this row become valid for this member?				1/1/1900		1/24/2011	Standard SCD-2
RowEndDate	Row End Date	Row End Date	When did this row become invalid? (12/31/9999 if current row)				1/14/1998, 12/31/9999		n/a	Exclude from cube
RowChangeReason	Row Change Reason	Row Change Reason	Why did the row change last?				N/A		n/a	Standard SCD-2

Table Name		DimCustomers								
Table Type		Dimension								
Display Name		Customers								
Database Schema										
Table Description										
Comment										
Biz Filter Logic										
Size										
Generate Script?		Y								
Column Name		Display Name	Target			Datatype	Size	Precision	Key?	FK To
customerKey	customerKey	customerKey	Surrogate primary key			int			PK	ID
customer_id	customer_id	customer_id	Business key from source system (aka natural key)			varchar			PK	
customer_city	customer_city	customer_city							N	
customer_state	customer_state	customer_state							N	
RowIsCurrent	Row Is Current	Row Is Current							N	
RowStartDate	Row Start Date	Row Start Date							N	
RowEndDate	Row End Date	Row End Date							N	
RowChangeReason	Row Change Reason	Row Change Reason							N	

## DimProducts

Table Name		DimProducts								
Table Type		Dimension								
Display Name		Products								
Database Schema										
Table Description										
Comment										
Biz Filter Logic										
Size										
Generate Script?		Y								
Column Name		Display Name	Description			Unknown Member	Example Values		SCD Type	Display Folder
productKey	productKey	productKey	Surrogate primary key			-1	1, 2, 3		key	
product_id	product_id	product_id	Business key from source system (aka natural key)				00064a2aeab9300754bb9		key	
product_category_name	product_category_name	product_category_name	Category or classification of the product.				d3f3c38			
product_weight_g	product_weight_g	product_weight_g	Weight of the product in grams.				perfumaria		2	
product_length_mm	product_length_mm	product_length_mm	Length of the product in millimeters.				300		2	
product_height_mm	product_height_mm	product_height_mm	Height of the product in centimeters				25		2	
product_width_mm	product_width_mm	product_width_mm	Width of the product in centimeters				16		2	
RowIsCurrent	Row Is Current	Row Is Current	Is this the current row for this member (Y/N)?				Y, N		n/a	Exclude from cube
RowStartDate	Row Start Date	Row Start Date	When did this row become valid for this member?				1/1/1900		1/24/2011	Standard SCD-2
RowEndDate	Row End Date	Row End Date	When did this row become invalid? (12/31/9999 if current row)				12/31/9999		n/a	Exclude from cube
RowChangeReason	Row Change Reason	Row Change Reason	Why did the row change last?				N/A		n/a	Standard SCD-2

Table Name		DimProducts												
Table Type		Dimension												
Display Name		Products												
Database Schema														
Table Description														
Comment														
Biz Filter Logic														
Size														
Generate Script?		Y												
		<b>Target</b>												
Column Name	Display Name	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
productKey	productKey	int			PK ID	PK	N	-1	Derived	olist	dbo	Products	product_id	varchar
product_id	product_id	varchar					N							
product_category_name	product_category_name	varchar					N		olist	dbo	Products	product_category_name	varchar	
product_weight_g	product_weight_g	decimal							olist	dbo	Products	product_weight_g	decimal	
product_length_cm	product_length_cm	decimal							olist	dbo	Products	product_length_cm	decimal	
product_height_cm	product_height_cm	decimal							olist	dbo	Products	product_height_cm	decimal	
product_width_cm	product_width_cm	decimal							olist	dbo	Products	product_width_cm	decimal	
RowIsCurrent	Row Is Current	nchar				1			Derived					
RowStartDate	Row Start Date	datetime					N		Derived					
RowEndDate	Row End Date	datetime					N	12/31/9999	Derived					
RowChangeReason	Row Change Reason	nvarchar	200					N	Derived					

## DimSellers

Table Name		DimSellers												
Table Type		Dimension												
Display Name		Sellers												
Database Schema														
Table Description														
Comment														
Biz Filter Logic														
Size														
Generate Script?		Y												
		<b>Home Page</b>												
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Display Folder	ETL Rules	Comments						
sellerKey	sellerKey	Surrogate primary key	-1	1, 2, 3 0015a82c2db000af6aaaf3a 2ecb0532	key									
Seller_id	Seller_id	Business key from source system (aka natural key)												
seller_city	seller_city	City where the seller is located	santo andre	2										
seller_state	seller_state	State where the seller is located	SP	2										
RowIsCurrent	Row Is Current	Is this the current row for this member (Y/N)?	Y	Y, N	n/a	Exclude from cube	Standard SCD-2							
RowStartDate	Row Start Date	When did this row become valid for this member? (1/1/1900 if current row)	1/1/1900	1/24/2011	n/a	Exclude from cube	Standard SCD-2							
RowEndDate	Row End Date	When did this row become invalid? (12/31/9999 if current row)	12/31/9999	1/14/1998, 12/31/9999	n/a	Exclude from cube	Standard SCD-2							
RowChangeReason	Row Change Reason	Why did the row change last?	N/A		n/a	Exclude from cube	Standard SCD-2							

Table Name		DimSellers												
Table Type		Dimension												
Display Name		Sellers												
Database Schema														
Table Description														
Comment														
Biz Filter Logic														
Size														
Generate Script?		Y												
		<b>Target</b>												
Column Name	Display Name	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Source	Source System	Source Schema	Source Table	Source Field Name	Source Datatype
sellerKey	sellerKey	int			PK ID	PK	N	-1	Derived	olist	dbo	Sellers	Seller_id	varchar
Seller_id	Seller_id	varchar					N							
seller_city	seller_city	varchar					N		olist	dbo	Sellers	seller_city	varchar	
seller_state	seller_state	varchar					N		olist	dbo	Sellers	seller_state	varchar	
RowIsCurrent	Row Is Current	nchar			1				Derived					
RowStartDate	Row Start Date	datetime					N		Derived					
RowEndDate	Row End Date	datetime					N	12/31/9999	Derived					
RowChangeReason	Row Change Reason	nvarchar	200					N	Derived					

## DimDate

Table Name		DimDate												
Table Type		Dimension												
Display Name		Date												
Database Schema														
Table Description														
Comment														
Biz Filter Logic														
Size														
Generate Script?		Y												
		<b>Home Page</b>												
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Display Folder	ETL Rules	Comments						
DayOfYear	DayOfYear	This property represents the day of the year, with values ranging from 1 to 366.		104	2									
WeekOfYear	WeekOfYear	This property represents the week of the year, with values ranging from 1 to 53.		16	2									
MonthName	MonthName	January, February, March, April, May, June, July, August, September, October, November, December		April	2									
MonthOfYear	MonthOfYear	This property represents the month of the year, with values ranging from 1 to 12.		4	2									
Quarter	Quarter	This property represents the quarter of the year, with values ranging from 1 to 4.		2	2									
Year	Year	This property represents the year, with values ranging from 1900 to 2155.		2024	2									
IsWeekday	IsWeekday	This property represents whether the date is a weekday (TRUE) or a weekend (FALSE).		TRUE	2									
RowIsCurrent	Row Is Current	Is this the current row for this member (Y/N)?		Y, N	n/a	Exclude from cube	Standard SCD-2							
RowStartDate	Row Start Date	When did this row become valid for this member?		1/1/1900	1/24/2011	n/a	Exclude from cube	Standard SCD-2						
RowEndDate	Row End Date	When did this row become invalid? (12/31/9999 if current row)		12/31/9999	1/14/1998, 12/31/9999	n/a	Exclude from cube	Standard SCD-2						
RowChangeReason	Row Change Reason	Why did the row change last?		N/A	n/a	Exclude from cube	Standard SCD-2							

Table Name		DimDate						
Table Type	Dimension							
Display Name	Date							
Database Schema								
Table Description								
Comment								
Biz Filter Logic								
Size								
Generate Script?	Y							
Column Name		Display Name	Datatype	Size	Precision	Key?	FK To	NULL?
								Default Value
DayOfYear	DayOfYear	nvarchar	50				N	
WeekOfYear	WeekOfYear	nvarchar	50				N	
MonthName	MonthName	nvarchar	50				N	
MonthOfYear	MonthOfYear	nvarchar	50				N	
Quarter	Quarter	nvarchar	50				N	
Year	Year	nvarchar	50				N	
IsWeekday	IsWeekday	nvarchar	50				N	
RowIsCurrent	Row Is Current	nchar	1				N	
RowStartDate	Row Start Date	datetime					N	
RowEndDate	Row End Date	datetime					N	12/31/9999
RowChangeReason	Row Change Reason	nvarchar	200				N	

## Fact Sales Report

Table Name		FactSales						
Table Type	Fact							
Display Name	Sales							
Database Schema								
Table Description								
Comment								
Biz Filter Logic								
Size								
Generate Script?	Y							
Column Name		Display Name	Example Values		Display Folder		ETL Rules	
productKey	productKey		1, 2, 3		key			
customerKey	customerKey		1, 2, 3		key			
sellerKey	sellerKey		1, 2, 3		key			
price	price							
freight_value	freight_value							
OrderID	OrderID							
OrderItemID	OrderItemID							
orderDeliveredCustomerDateKey	orderDeliveredCustomerDateKey							

Table Name		FactSales						
Table Type	Fact							
Display Name	Sales							
Database Schema								
Table Description								
Comment								
Biz Filter Logic								
Size								
Generate Script?	Y							
Column Name		Display Name	Datatype	Size	Precision	Key?	FK To	NULL?
								Default Value
productKey	productKey	int					FK DimProducts.productKey	N
customerKey	customerKey	int					FK DimCustomers.customerKey	N
sellerKey	sellerKey	int					FK DimSellers.sellerKey	Y
price	price	(8, 3)						Y
freight_value	freight_value	(8, 3)						Y
OrderID	OrderID	50						Y
OrderItemID	OrderItemID	int						Y
orderDeliveredCustomerDateKey	orderDeliveredCustomerDateKey	int						

## Fact Order Fulfillment and Delivery

Table Name		FactOrderFullFulfillment						
Table Type	Fact							
Display Name	OrderFullFulfillment							
Database Schema								
Table Description								
Comment								
Biz Filter Logic								
Size								
Generate Script?	Y							
Column Name		Display Name	Description	Example Values		Display Folder		ETL Rules
								Comments
customerKey	customerKey			1, 2, 3		key		
sellerKey	sellerKey			1, 2, 3		key		
delivery_time	delivery_time							
delivery_delay	delivery_delay							
accept_time	accept_time							
order_purchase_timestamp_key	order_purchase_timestamp_key							
order_approved_at_key	order_approved_at_key							
order_delivered_carrier_date_key	order_delivered_carrier_date_key							
order_delivered_customer_date_key	order_delivered_customer_date_key							
order_delivered_estimate_date_key	order_delivered_estimate_date_key							

Table Name	FactOrderFullFilment														
Column Name	Display Name	Datatype	Size	Precision	Key?	Target	FK To	NULL?	Default Value	Source System	Source Schema	Source Table	Source Field Name	Source Datatype	
customerKey	customerKey	int			FK	DimCustomers.customerKey		N	ETL Process						
sellerKey	sellerKey	int			FK	DimSellers.sellerKey		N	ETL Process						
delivery_time	delivery_time					DimDate.DateKey		Y							
delivery_delay	delivery_delay					DimDate.DateKey		Y							
accept_time	accept_time					DimDate.DateKey		Y							
order_purchase_timestamp_key	order_purchase_timestamp_key	int				DimDate.DateKey		Y							
order_approved_at_key	order_approved_at_key	int				DimDate.DateKey		Y							
order_delivered_carrier_date_key	order_delivered_carrier_date_key	int				DimDate.DateKey		Y							
order_delivered_customer_date_key	order_delivered_customer_date_key	int				DimDate.DateKey		Y							
order_delivered_estimate_date_key	order_delivered_estimate_date_key	int				DimDate.DateKey		Y							

## Kết quả: Tạo thành công SQL scripts

Kimball Group, The Microsoft Data Warehouse Toolkit  
Generate a database from the datamodel worksheet, version: 4

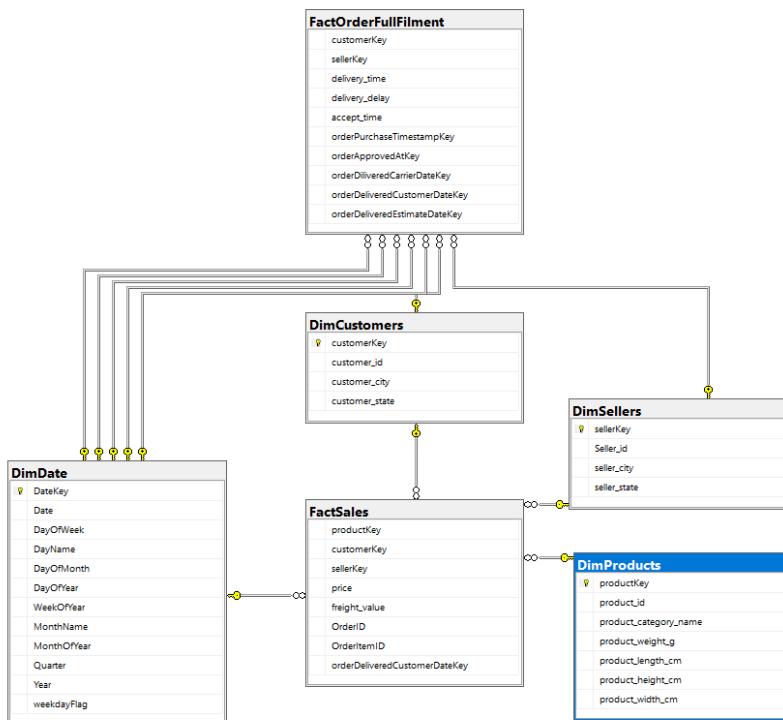
You can use this Excel workbook as a data modeling tool during the logical design phase of your project. As discussed in the book, it is in some ways preferable to a real data modeling tool during the initial design. We expect you to move away from this spreadsheet and into a real modeling tool during the physical design phase. The authors provide this macro so that the spreadsheet isn't a dead-end. You can 'import' into your data modeling tool by generating a database using this script, then reverse-engineering that database into your tool.

Uncomment the next lines if you want to drop and create the database

```
/*
DROP DATABASE OlistDW
GO
CREATE DATABASE OlistDW
GO
ALTER DATABASE OlistDW
SET RECOVERY SIMPLE
GO
*/
USE OlistDW
;

IF EXISTS (SELECT Name from sys.extended_properties where Name = 'Description')
    EXEC sys.sp_droptextendedproperty @name = 'Description'
EXEC sys.sp_addextendedproperty @name = 'Description', @value = 'Default description - you should change this.'
;
```

Sau đó đưa script qua SQL Server thì tạo thành công cơ sở dữ liệu OlistDW



## PHẦN 2: DATA INTERGRATION WITH SSIS

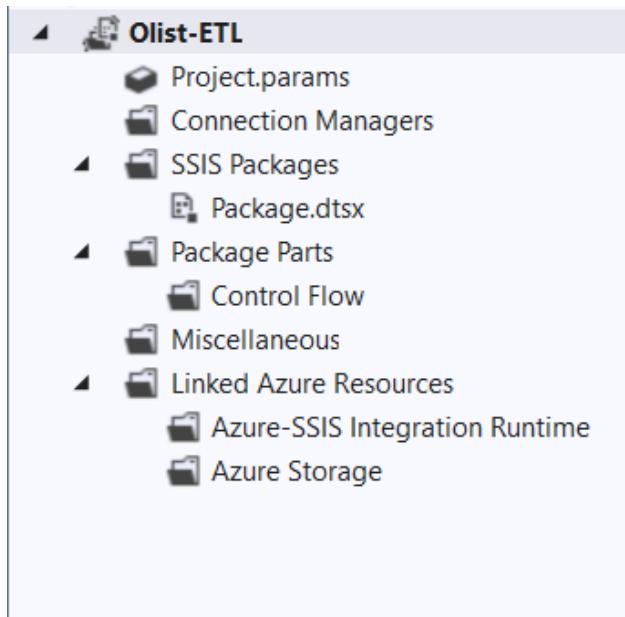
### 2.1. Tạo ETL project và solution

Từ Visual Studio menu, chọn **File -> New -> Project**.

Tùy dialog:

- Chọn **Integration Services Project**
- Đặt tên: **Northwind-ETL** (chọn đường dẫn phù hợp)

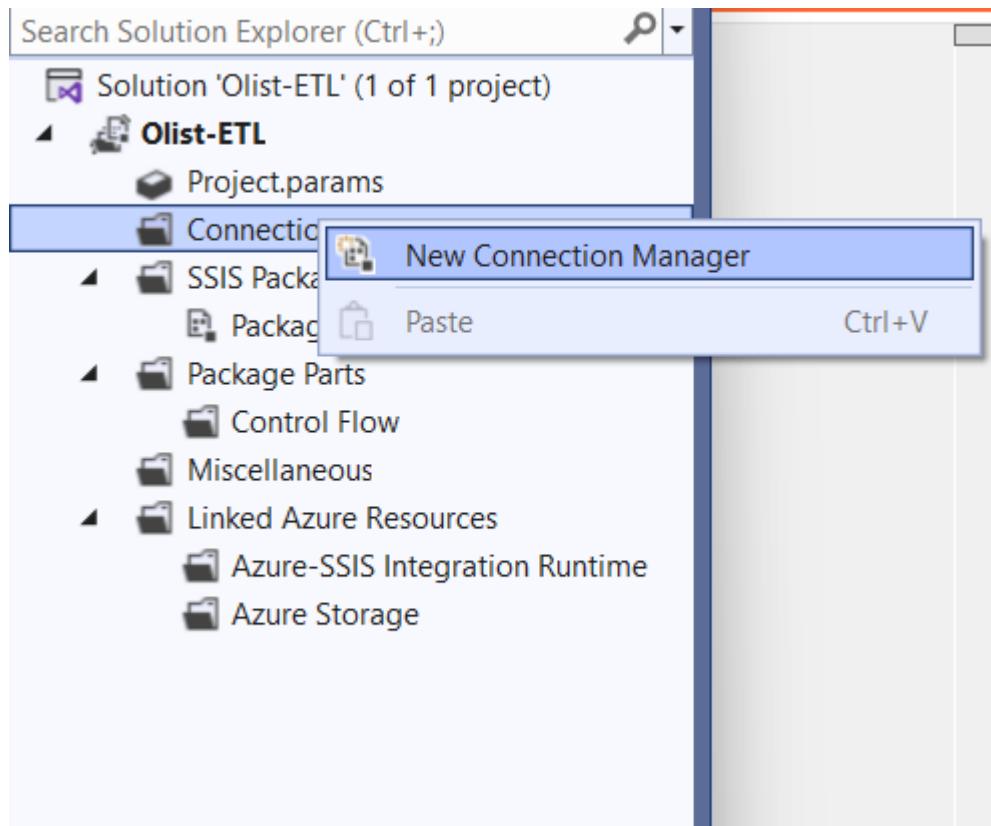
Click **OK**



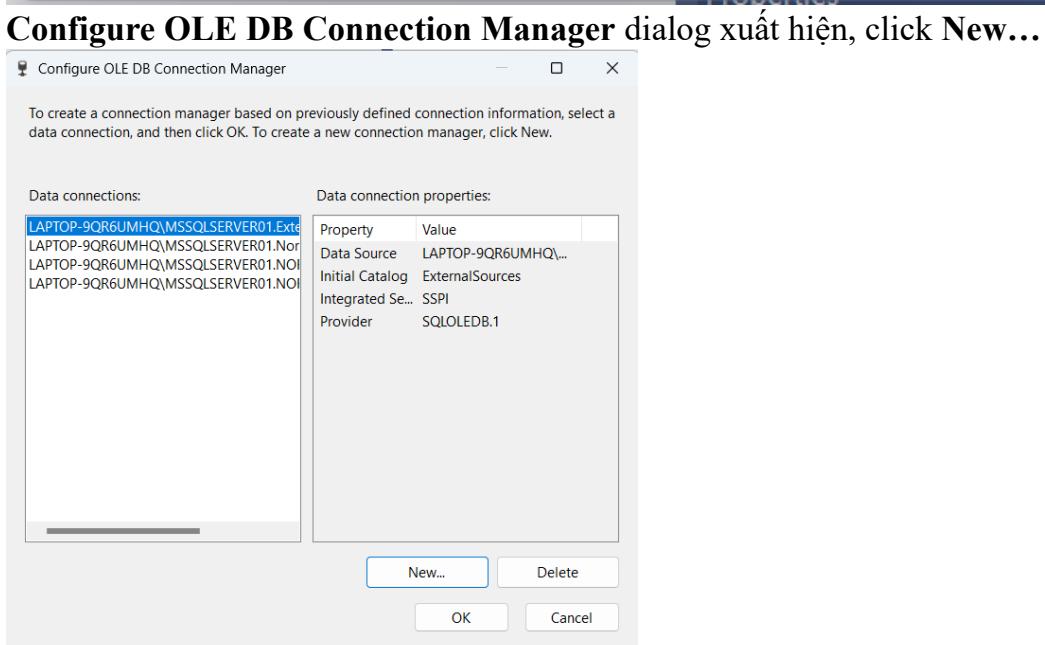
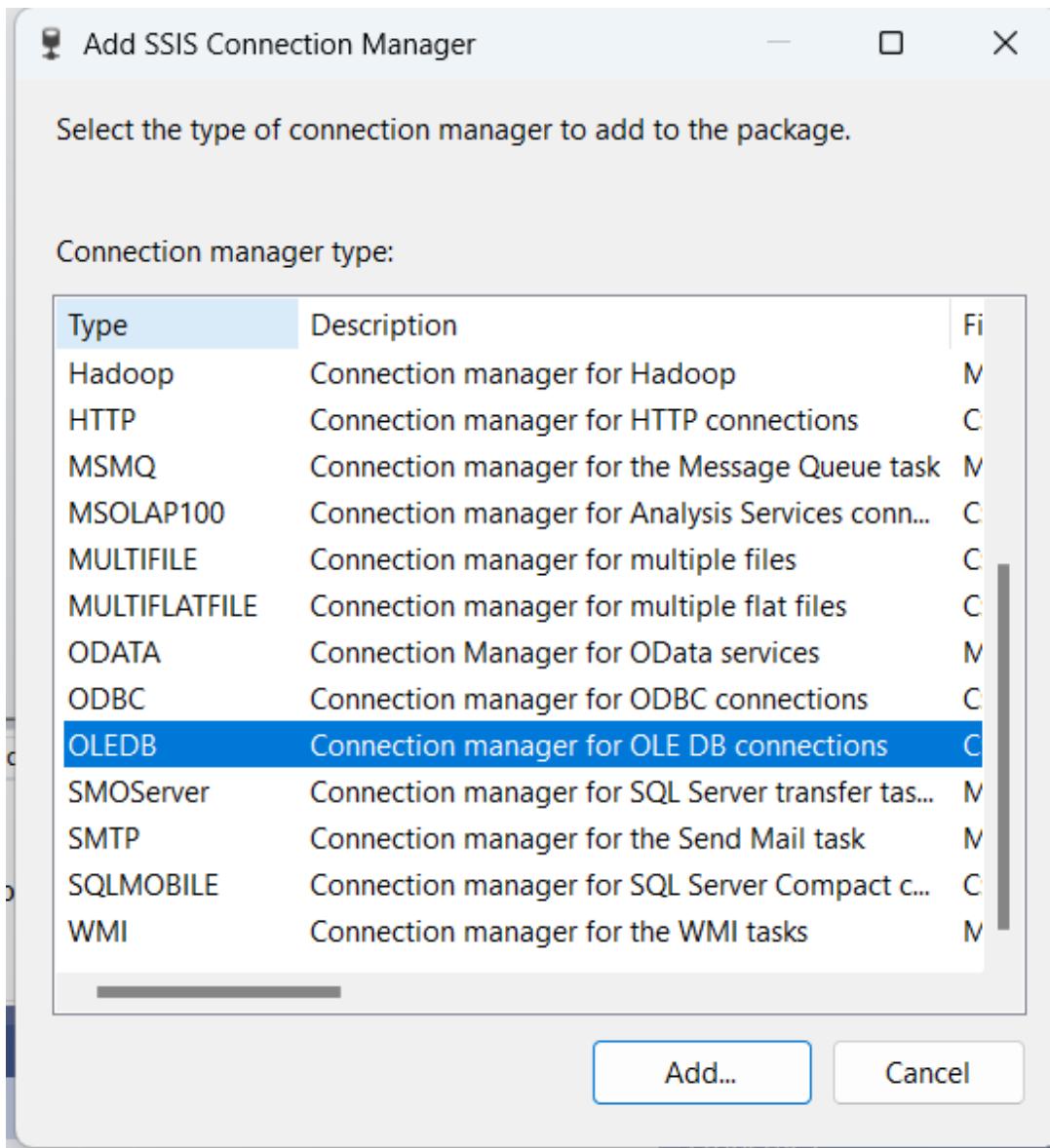
### 2.2. Thiết lập các Connection

Để có thể truy xuất dữ liệu cho quá trình ETL, ta cần thiết lập các **connection** tới nguồn (**source**) và đích (**target**). Ta có thể tạo các **connection** này trong mỗi **package**. Tuy nhiên, để có thể sử dụng lại các **connection** trong các **package** khác nhau, ta sẽ tạo các **connection** một cách **toàn cục (globally)**.

Tại **solution explorer**, right-click vào **Connection Managers**, chọn **New Connection Manager**.



Chọn **OLEDB** từ **Add SSIS Connection Manager** dialog và click **Add...**



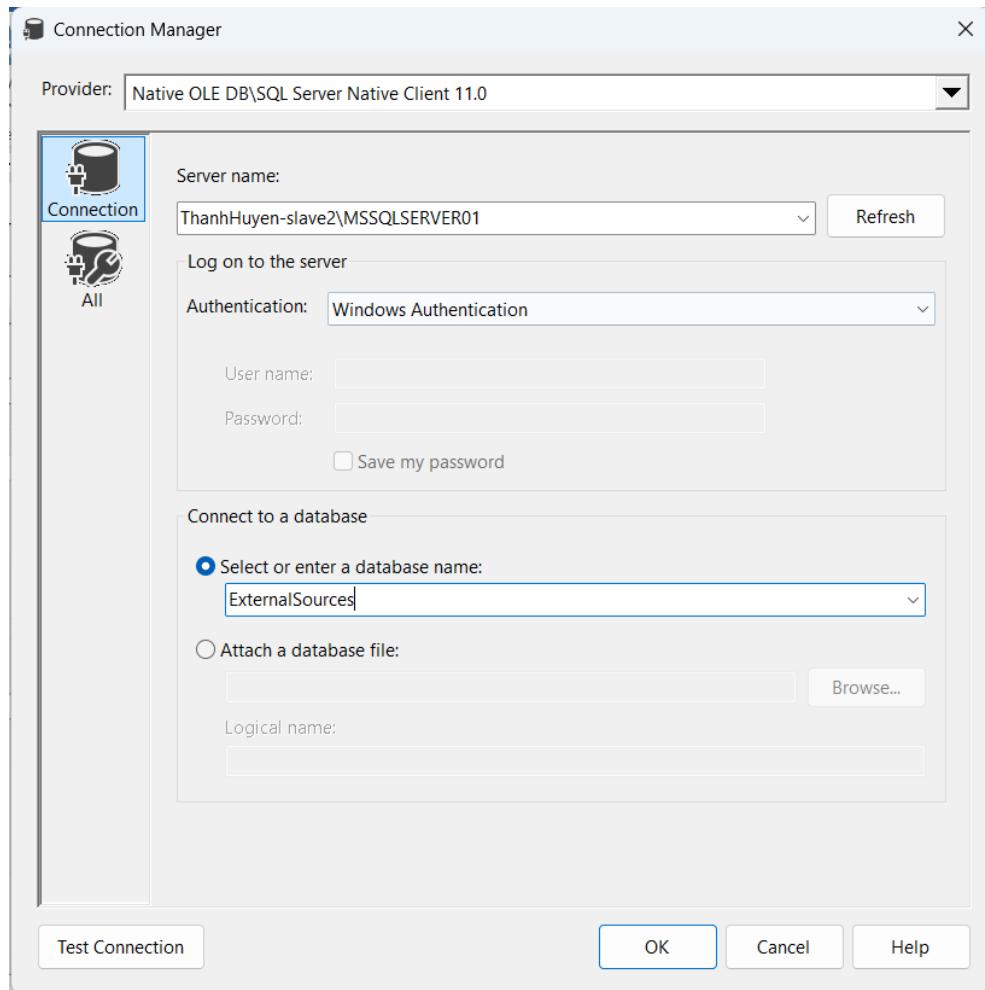
Bạn sẽ thấy **Connection Manager** dialog. Đặt tên cho các thông tin như bên dưới:

**Provider: SQL Server Native Client**

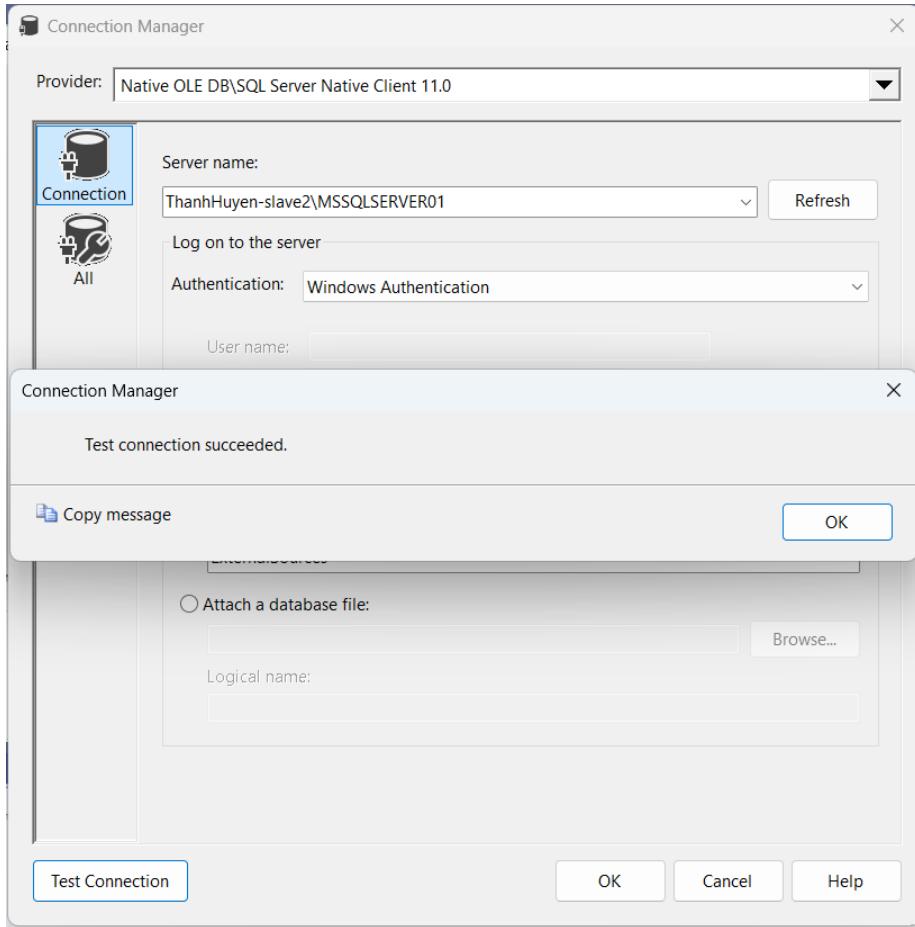
**Server Name: <tên database server>**

**Log on to the Server: Use Windows Authentication**

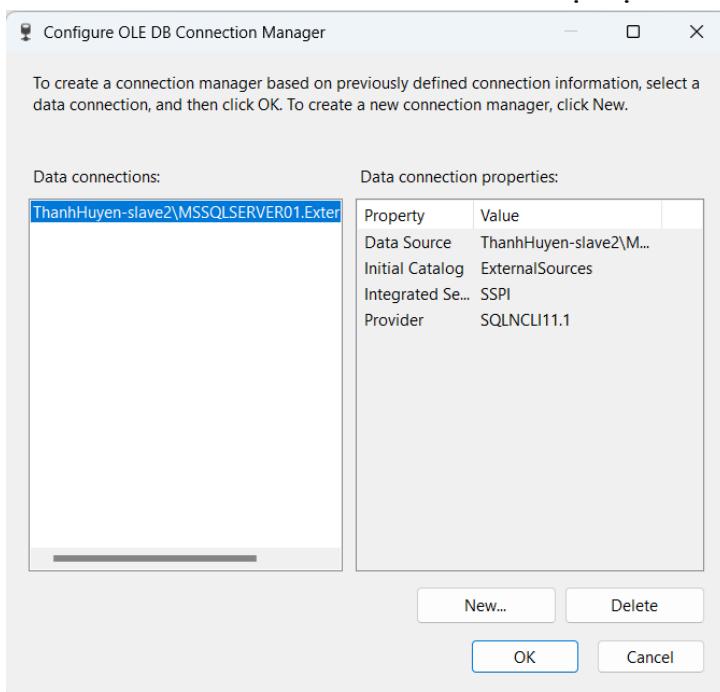
**Connect to a database: ExternalSources**



Click **Test Connection** để kiểm tra connection hoạt động, sau đó click **OK** để lưu connection.

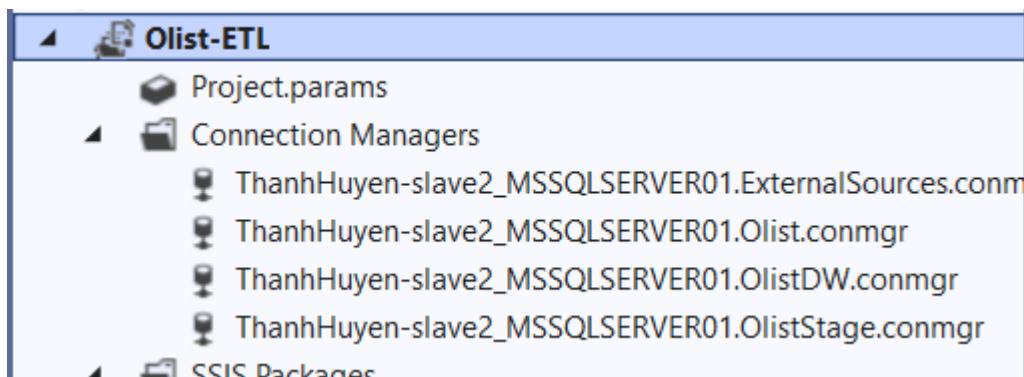


Trở lại màn hình **Configure OLE DB Connection Manager**, bạn sẽ thấy <địa chỉ database>.ExternalSources vừa được tạo:



Chọn **ExternalSources** connection và click **OK** để thêm connection này vào project.

Lặp lại các bước trên (bước 1->5) ba lần cho lần lượt **OlistDW**, **OlistStage** và **Olist** source database. Sau khi tạo xong **Connection Managers** folder sẽ như bên dưới.



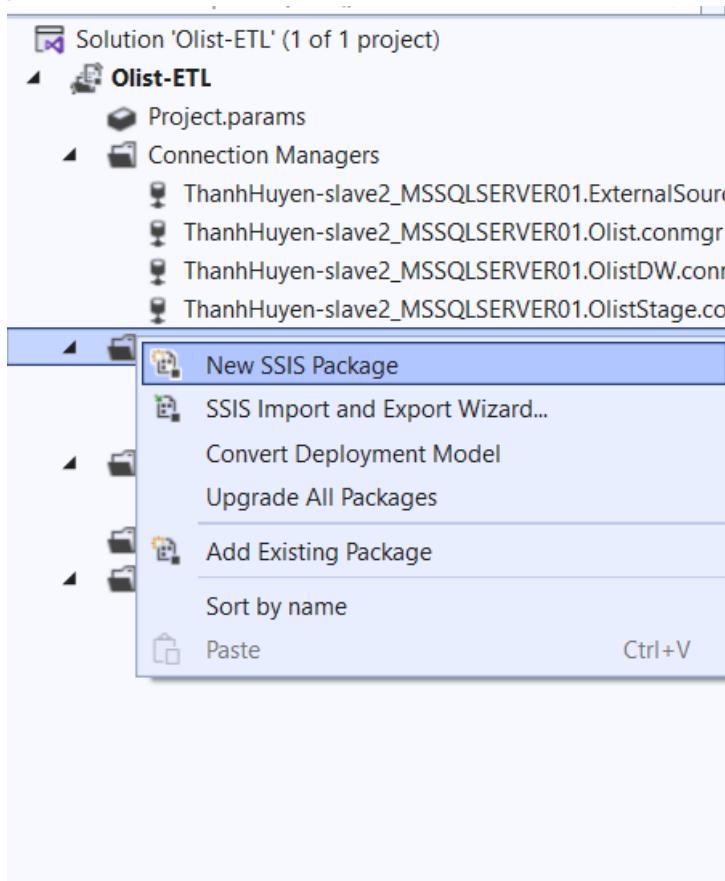
## 2.3. Import Dimension and Fact

### 2.3.1. Date Dimension

#### Tạo Package

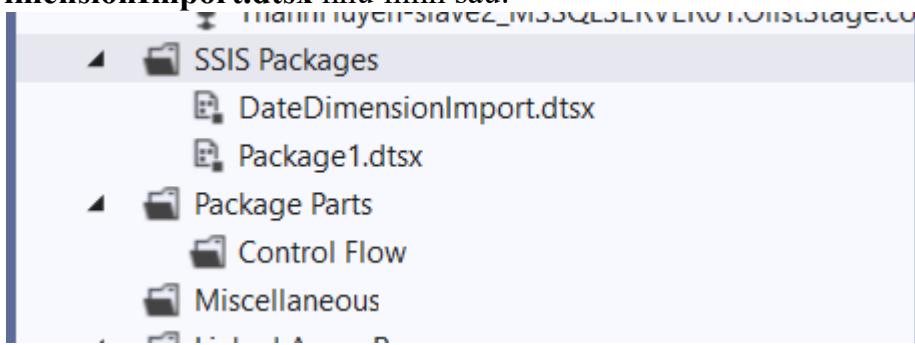
Tạo một package mới, **DateDimensionImport**.

Ở **Solution Explorer**, right-click vào **SSIS Packages** và chọn **New SSIS Package**



Right-Click vào package có tên **Package1.dtsx** và chọn **Rename**  
Đặt tên package là **DateDimensionImport**

Sau khi thực hiện xong bạn sẽ thấy **Package.dtsx** và **DateDimensionImport.dtsx** như hình sau:



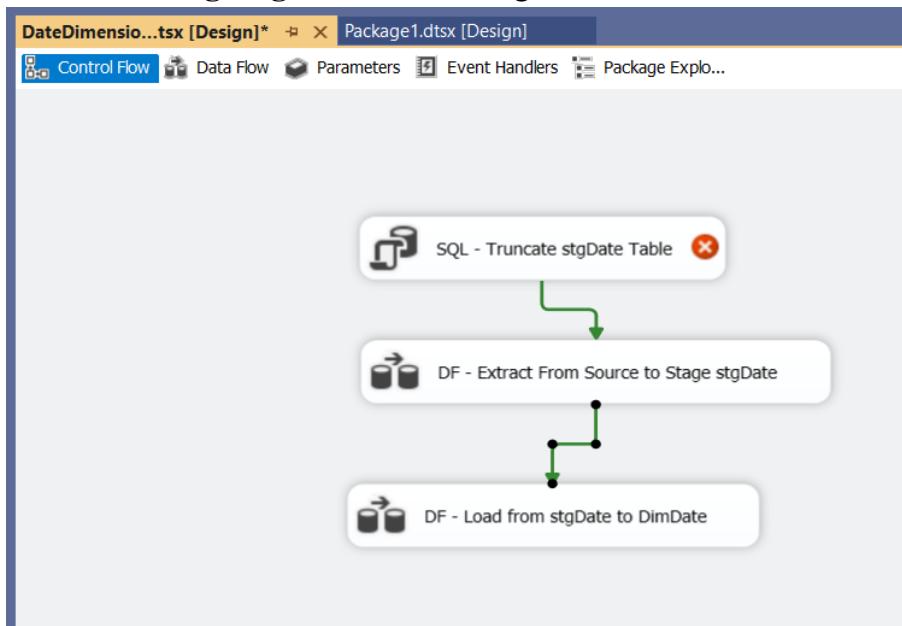
Double-click vào **DateDimensionImport.dtsx** package để mở nó.

### Thiết lập Control Flow

Từ SSIS Toolbox tạo 1 **Execute SQL task** và 2 **Data flow tasks** vào surface design như hình sau:

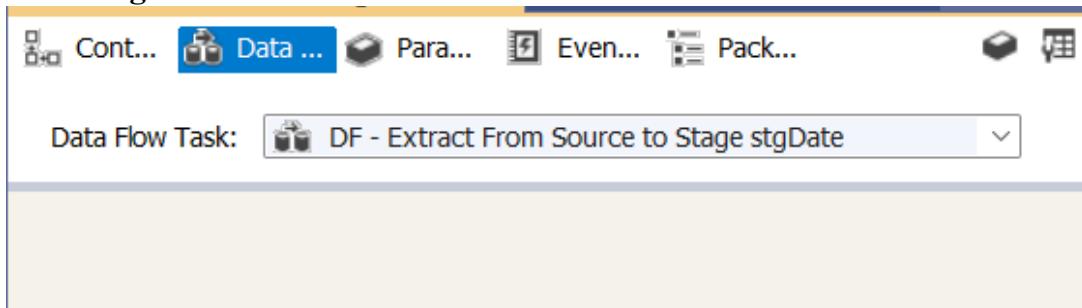
Đặt tên cho các **task**.

Connect các **tasks** để chúng có thể chạy theo thứ tự. Click vào **SQL - Truncate stgDate Table** task và một mũi tên xanh sẽ xuất hiện. Kéo rê nó vào **DF - Extract From Source To Stage stgDate** task. Tương tự, connect hai task còn lại

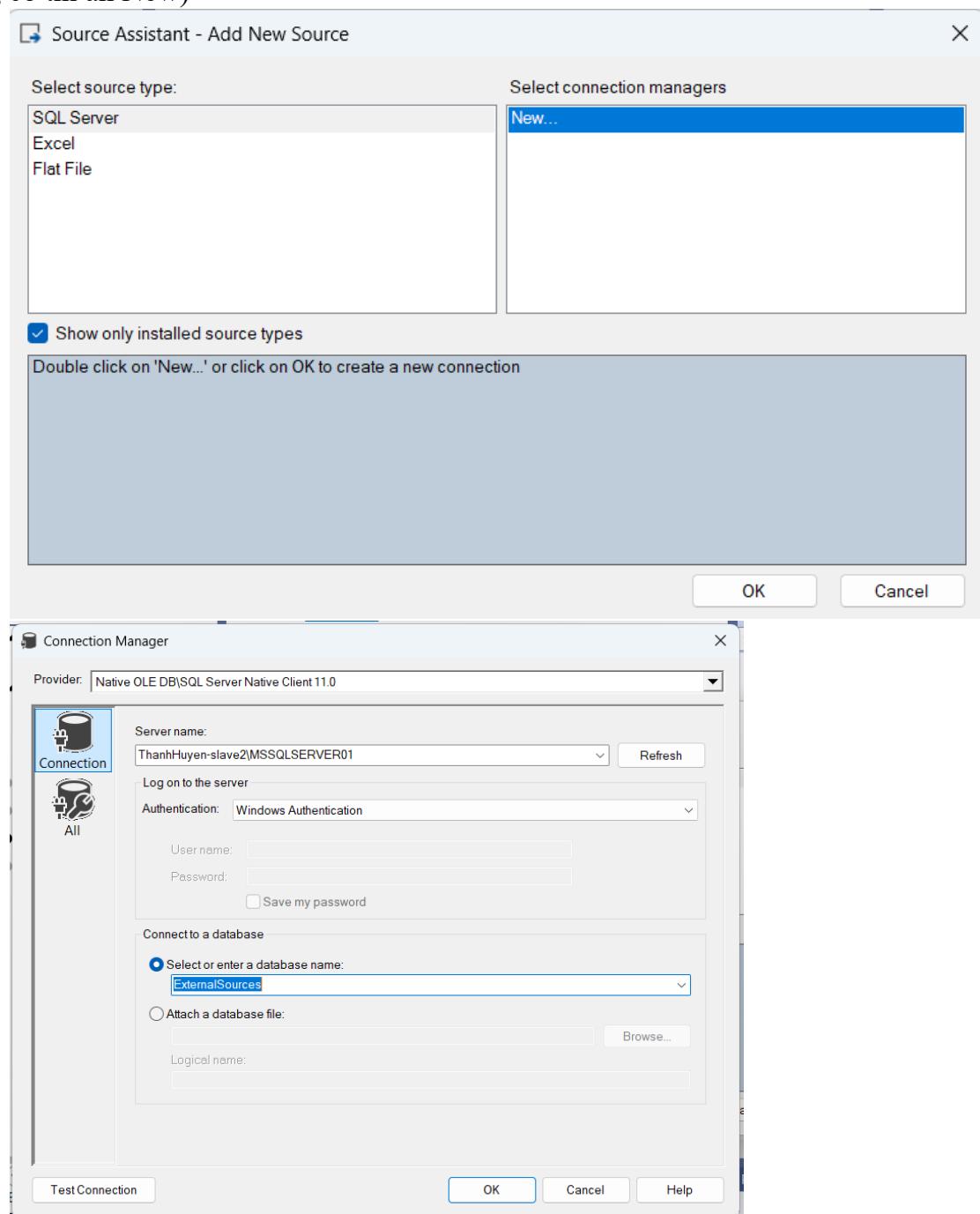


Cấu hình **Source -> Stage data flow**

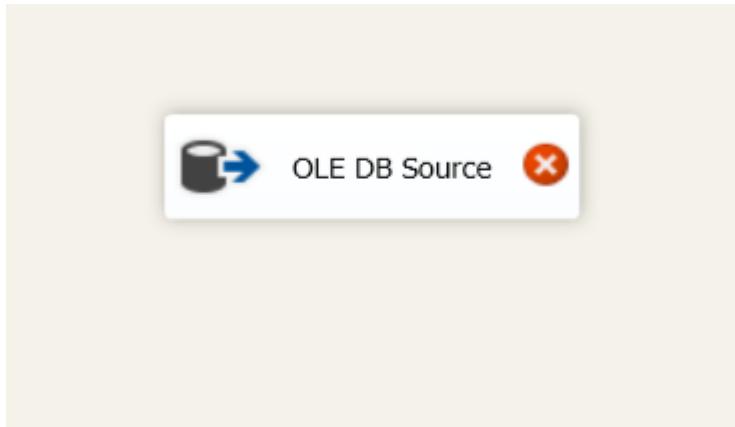
Double-click **DF - Extract From Source to Stage stgDate** task để mở nó trong **Data Flow design surface**:



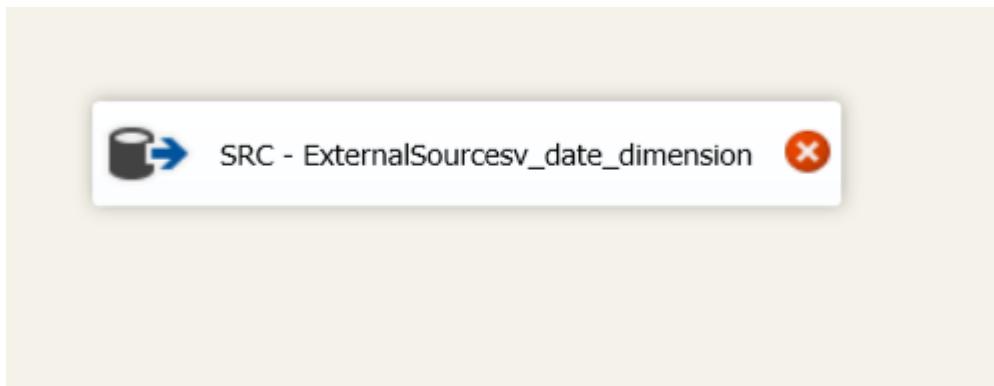
Từ SSIS Toolbox, kéo rê **Source Assistant** vào design surface, một dialog mở ra, Chọn source Type: **SQL Server** và **ExternalSources** connection manager.(nếu không có thì ấn New)



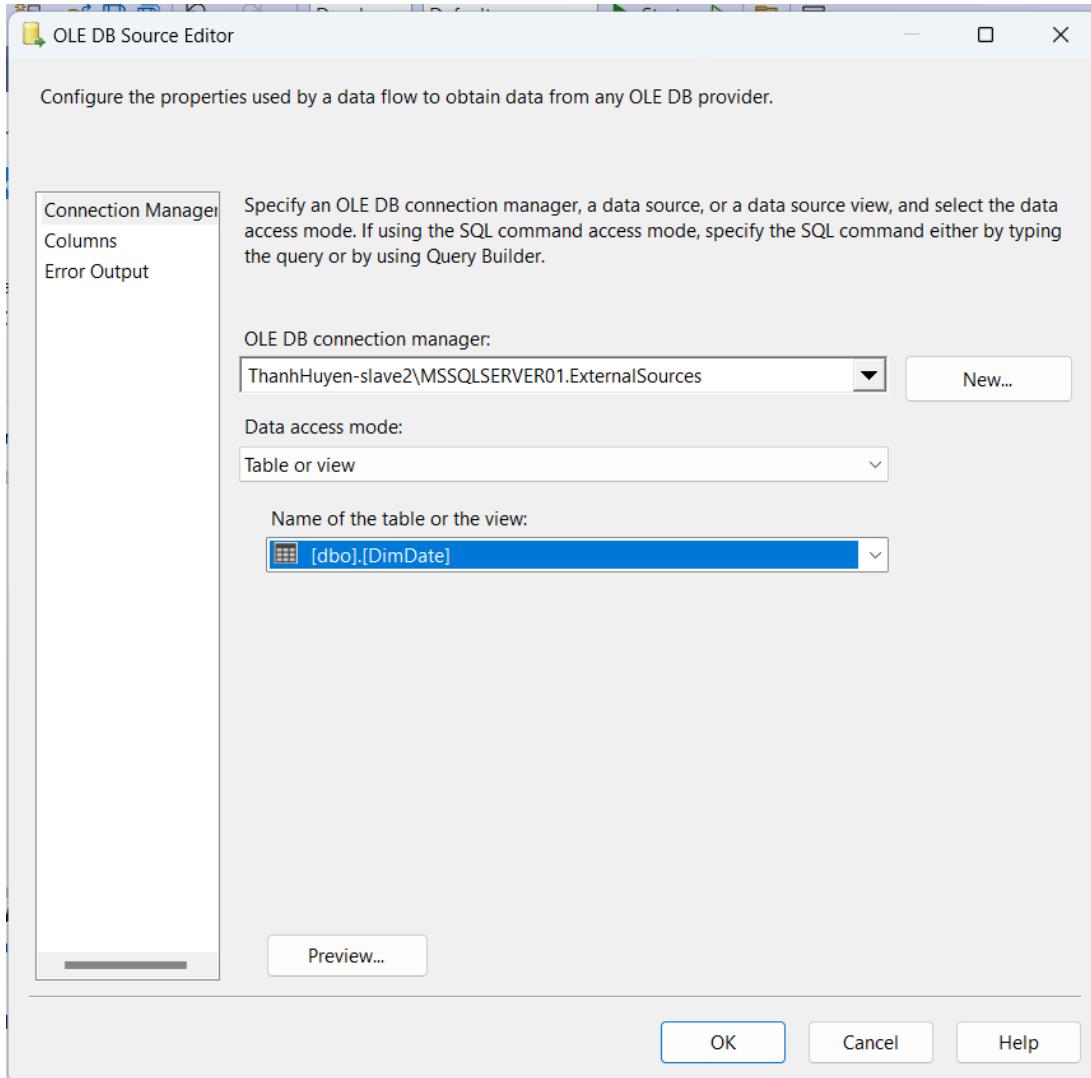
OLE DB Source sẽ xuất hiện như sau:



Đổi tên nó thành **SRC - ExternalSourcesv\_date\_dimension**



Double click vào source để cấu hình nó, **OLE DB Source Editor** sẽ mở ra. Từ dropdown **name of the table or view** chọn **[dbo].[DimDate]** là source. Click **Preview...** để xem source data



Click **OK** để lưu.

Sau khi cấu hình **source** xong, bạn sẽ không còn thấy dấu [x] màu đỏ (như bên dưới):

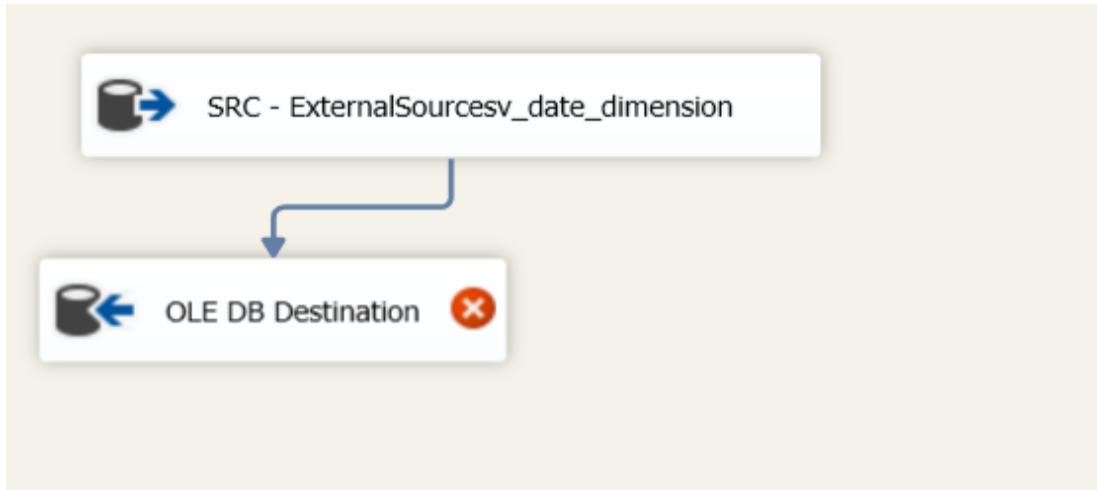


Tiếp tục, cấu hình destination. Từ **SSIS Toolbox**, kéo rê **Destination Assistant** vào design surface, một dialog xuất hiện. Chọn source Type: **SQL Server** và **NorthwindStage** là connection manager.

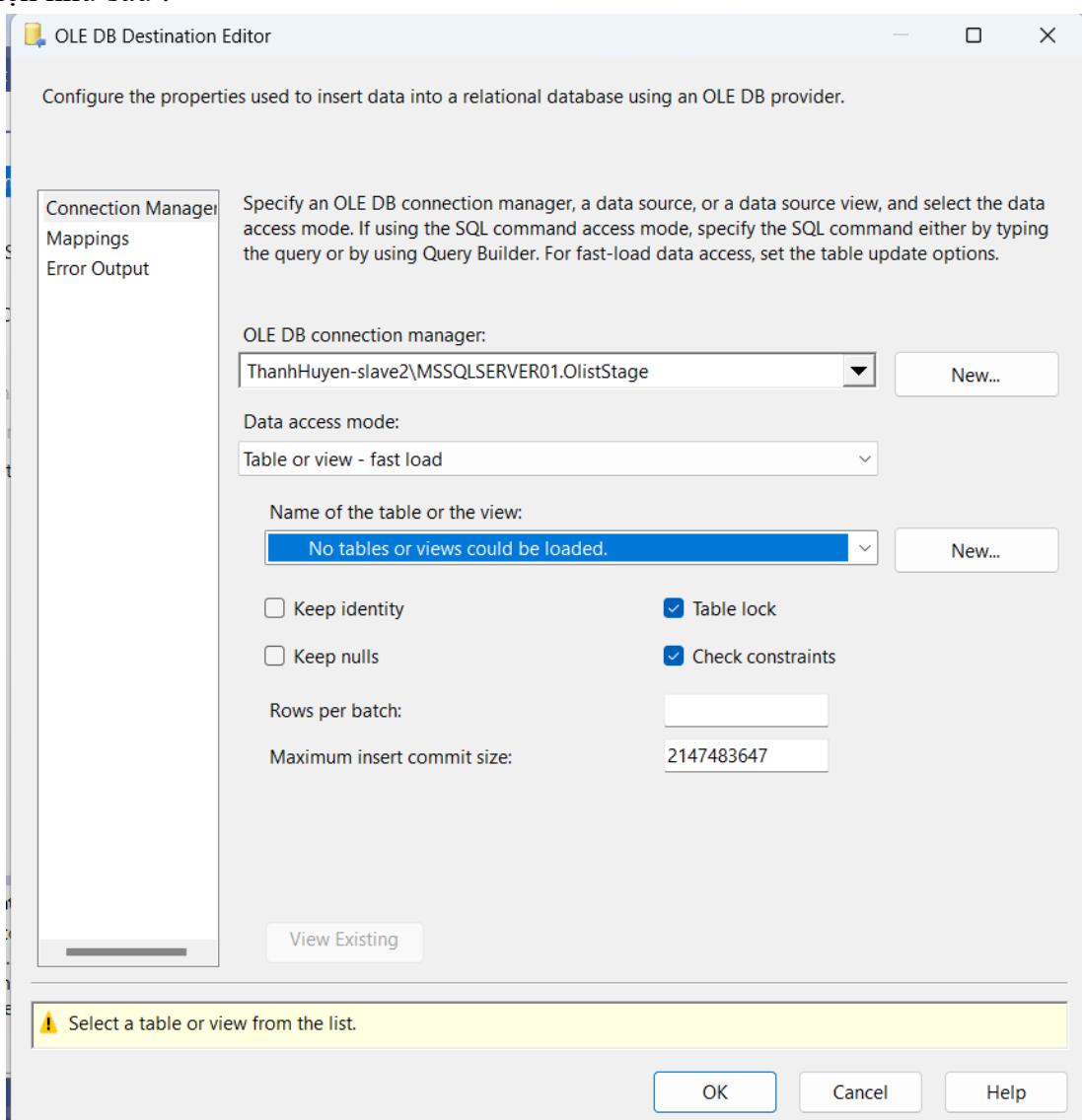
Ta có một **destination** như bên dưới

Đổi tên lại thành **DST – Stage stgDate**

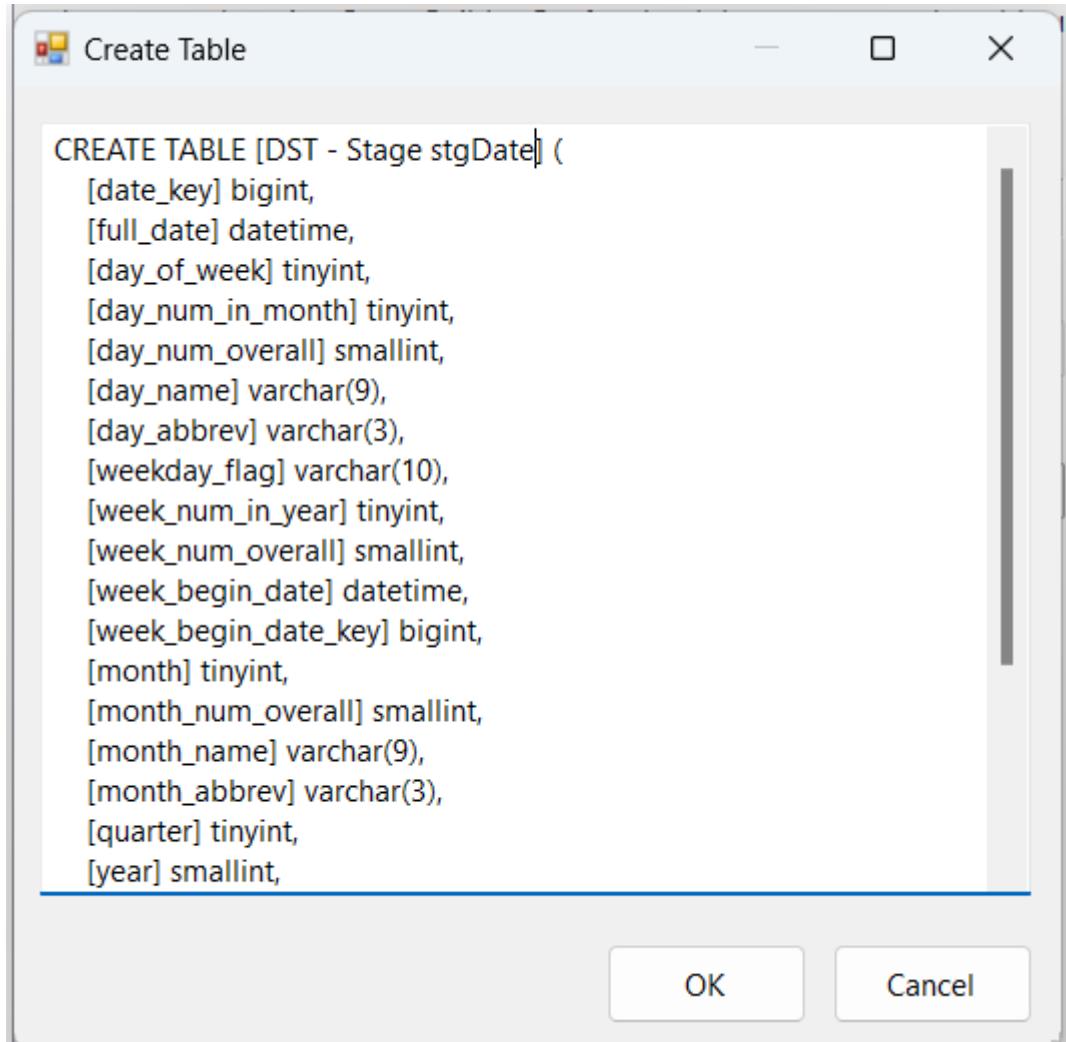
Click vào **source** và kéo rê nó vào **destination**.



Double click vào destination để cấu hình nó, **OLE DB Destination Editor** sẽ xuất hiện như sau :



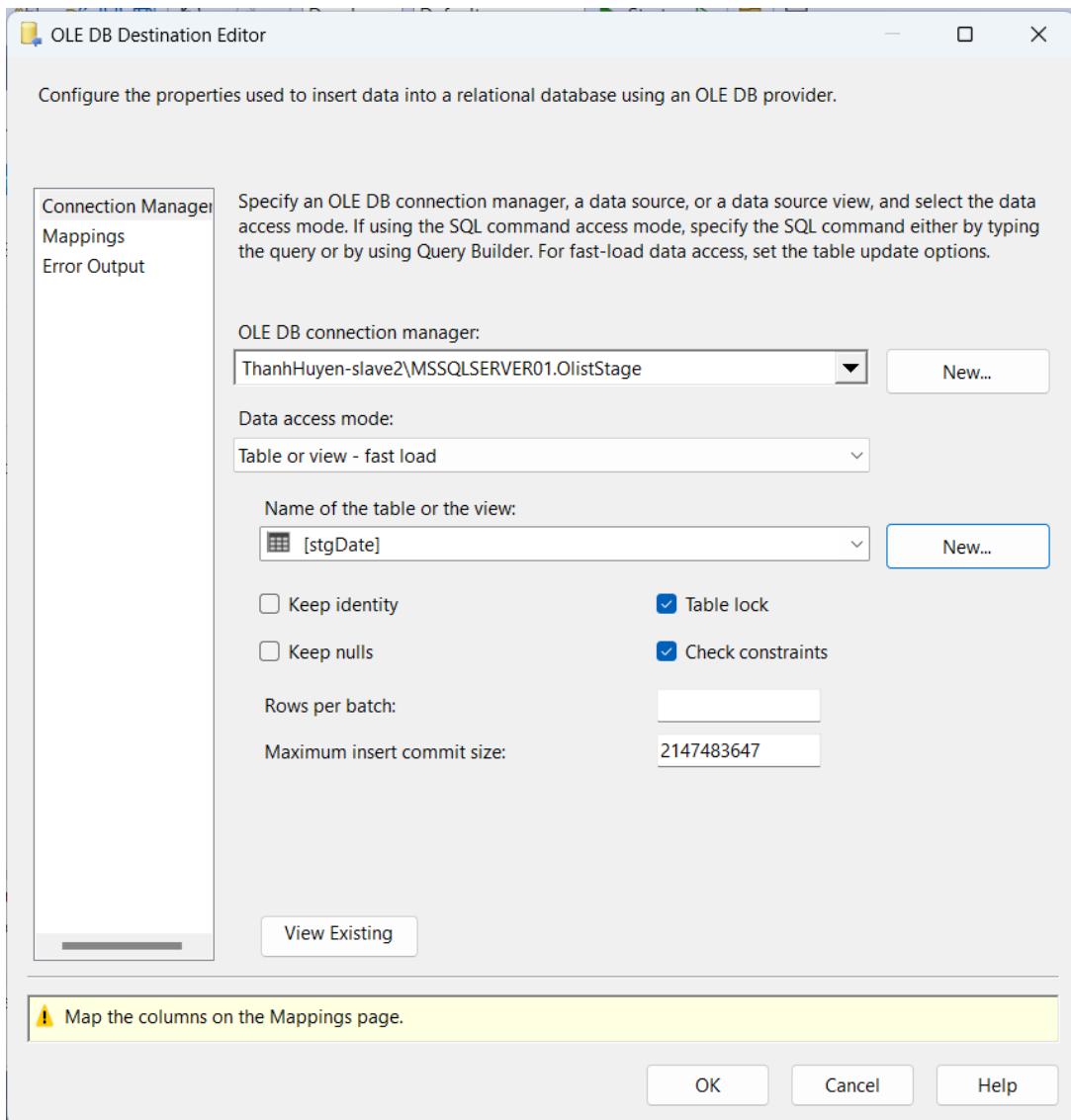
Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table** như sau:



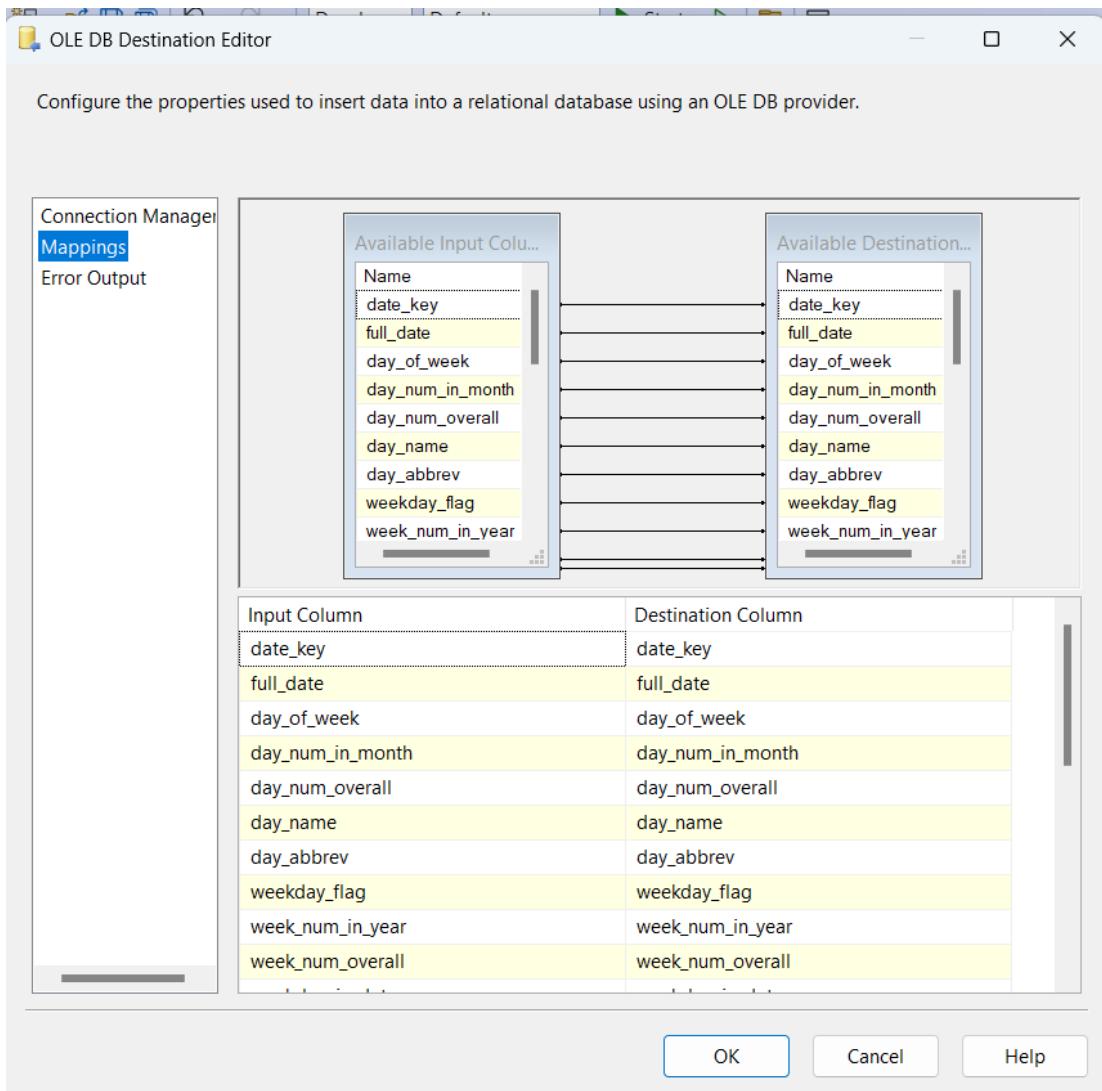
Đổi tên bảng lại thành **stgDate**. (**CREATE TABLE [stgDate]**).

Click OK.

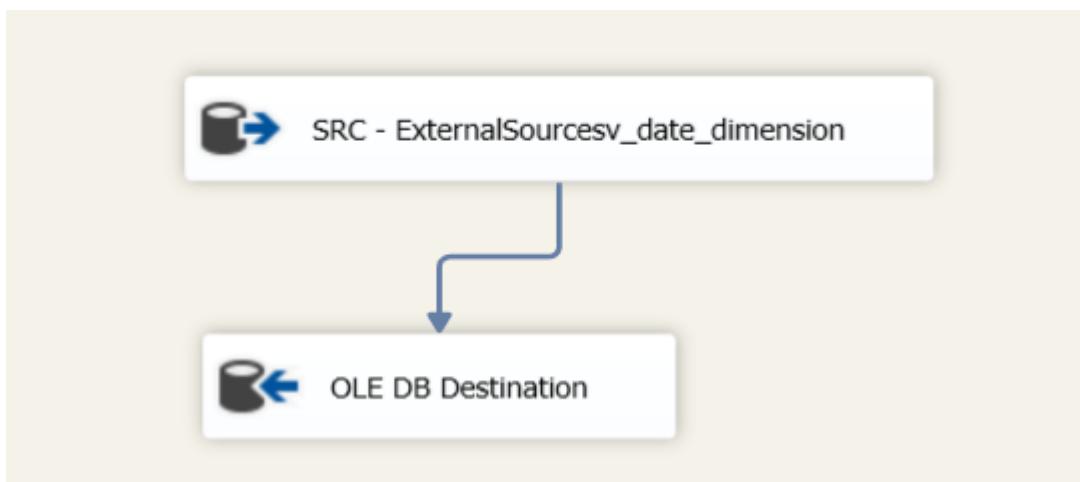
Trở lại màn hình **OLEDB Destination Editor**, tại mục **Name of the table or view** bạn sẽ thấy **[stgDate]**.



Click vào tab **Mapping** ta sẽ thấy như hình bên dưới:



Click **OK** để hoàn tất việc cấu hình. Màn hình data flow design surface sẽ như bên dưới.

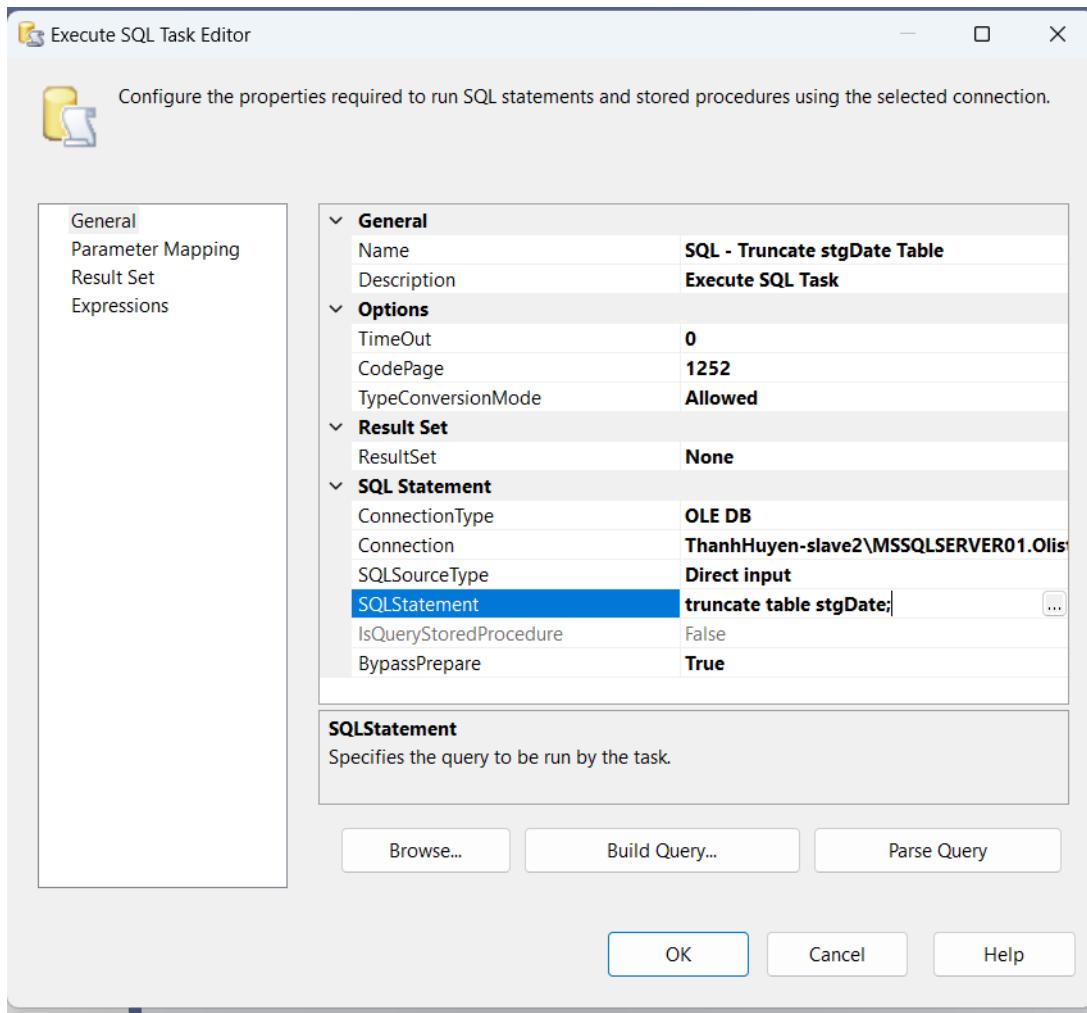


### Thiết lập truncate table SQL task

Click **Control Flow tab**.

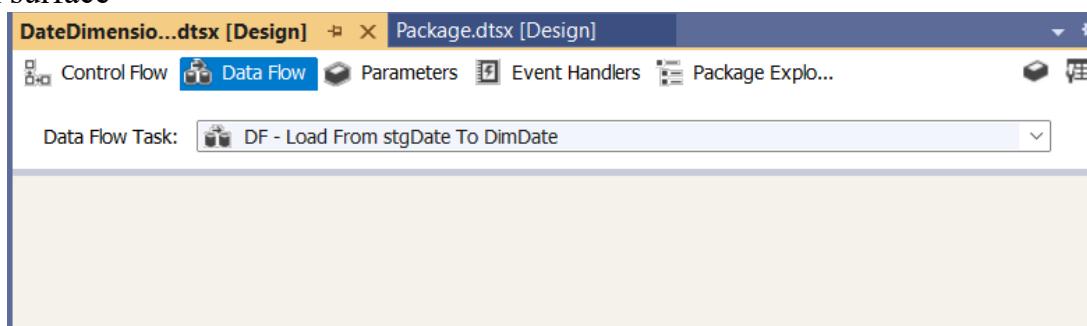
Double-click vào **SQL - Truncate stgDate Table** task để cấu hình nó, **Execute SQL Task Editor** sẽ mở ra.

Ở mục **SQL Statement**, chọn **Connection** là <tên database server>OlistStage và **SQLStatement** là truncate table stgDdate;



### Cấu hình Stage -> Target data flow

Double-click vào **DF - Load from stgDate to DimDate** task để mở data flow design surface



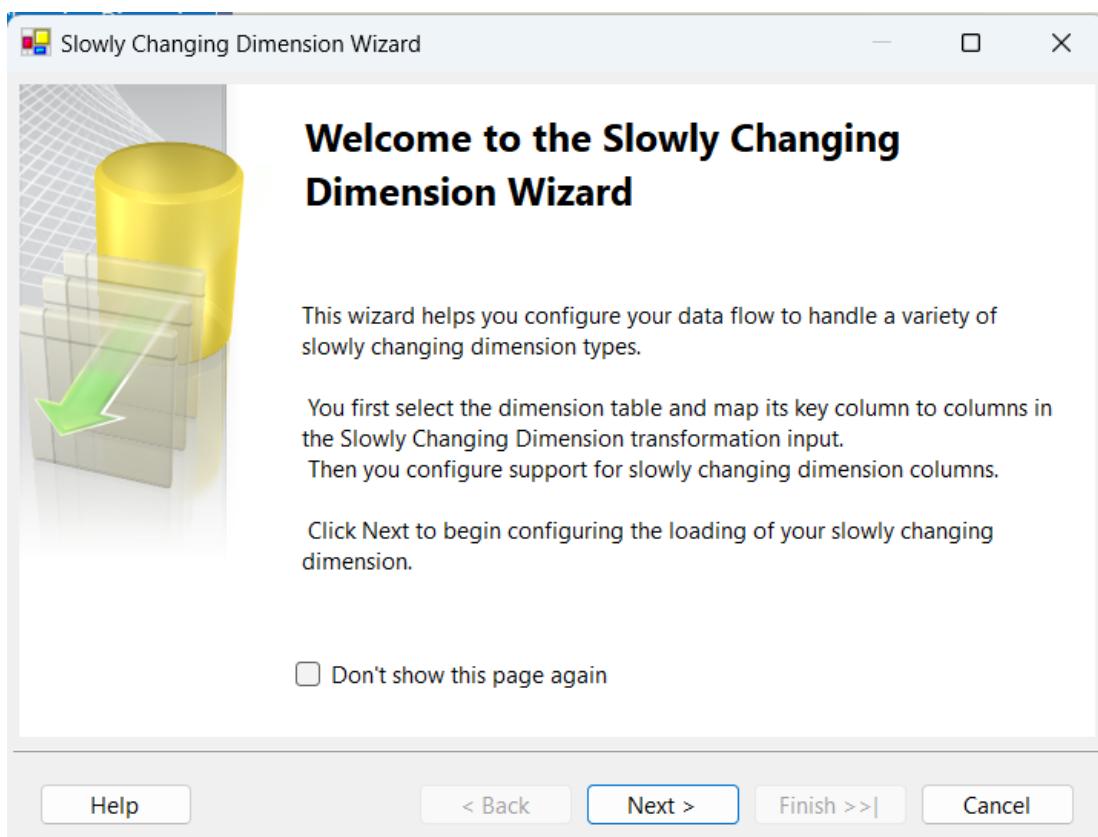
Kéo rê **Source Assistant** vào design surface, chọn **source type** là **SQL Server** và **connection manager** là **OlistStage**.

Đổi tên source thành **SRC - Stage stgDate**, và double-click vào nó để cấu hình.

Tù **OLE DB Source Editor** chọn [**dbo].[stgDate]**] table ở mục **name of the table or view**. Click OK để hoàn tất việc cấu hình.

Kéo rê **Slowly Changing Dimension** vào design surface. Đặt tên nó lại thành **DST - Slowly Changing Dimension DimDate** và **connect SRC -> DST**.

Double click vào **DST - Slowly Changing Dimension DimDate** để cấu hình **Slowly Changing Dimension Wizard** xuất hiện



Click **Next >**

Chọn bảng **DimDates** ở CSDL **OlistDW**. Ở mục **connection manager** chọn CSDL **OlistDW**.

Ở mục **Table or view** chọn bảng **[DimDates]**.

Chọn **DateKey** là **business key**. Tất cả các **cột khác** là “**Not a key column**”

Slowly Changing Dimension Wizard

### Select a Dimension Table and Keys

Select a dimension table to load and map columns in the transformation input to

Connection manager:

ThanhHuyen-slave2\MSSQLSERVER01.OlistDW ▾ New...

Table or view:

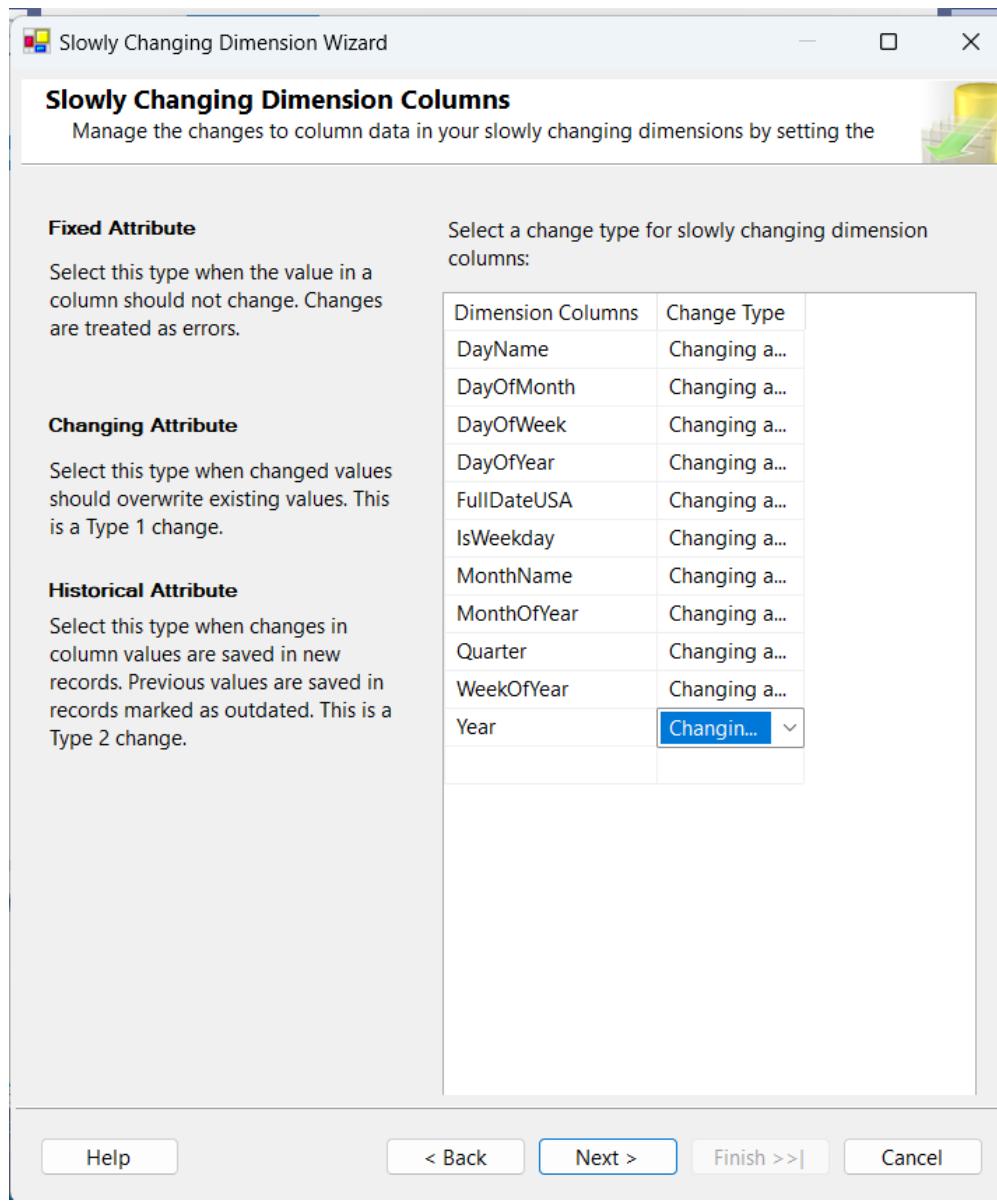
[dbo].[DimDates]

Input Columns	Dimension Columns	Key Type
date_key	Date	Business key
day_name	DayName	Not a key column
day_num_in_month	DayOfMonth	Not a key column
day_of_week	DayOfWeek	Not a key column
day_num_overall	DayOfYear	Not a key column
full_date	FullDateUSA	Not a key column
weekday_flag	IsWeekday	Not a key column

Help < Back Next > Finish >> Cancel

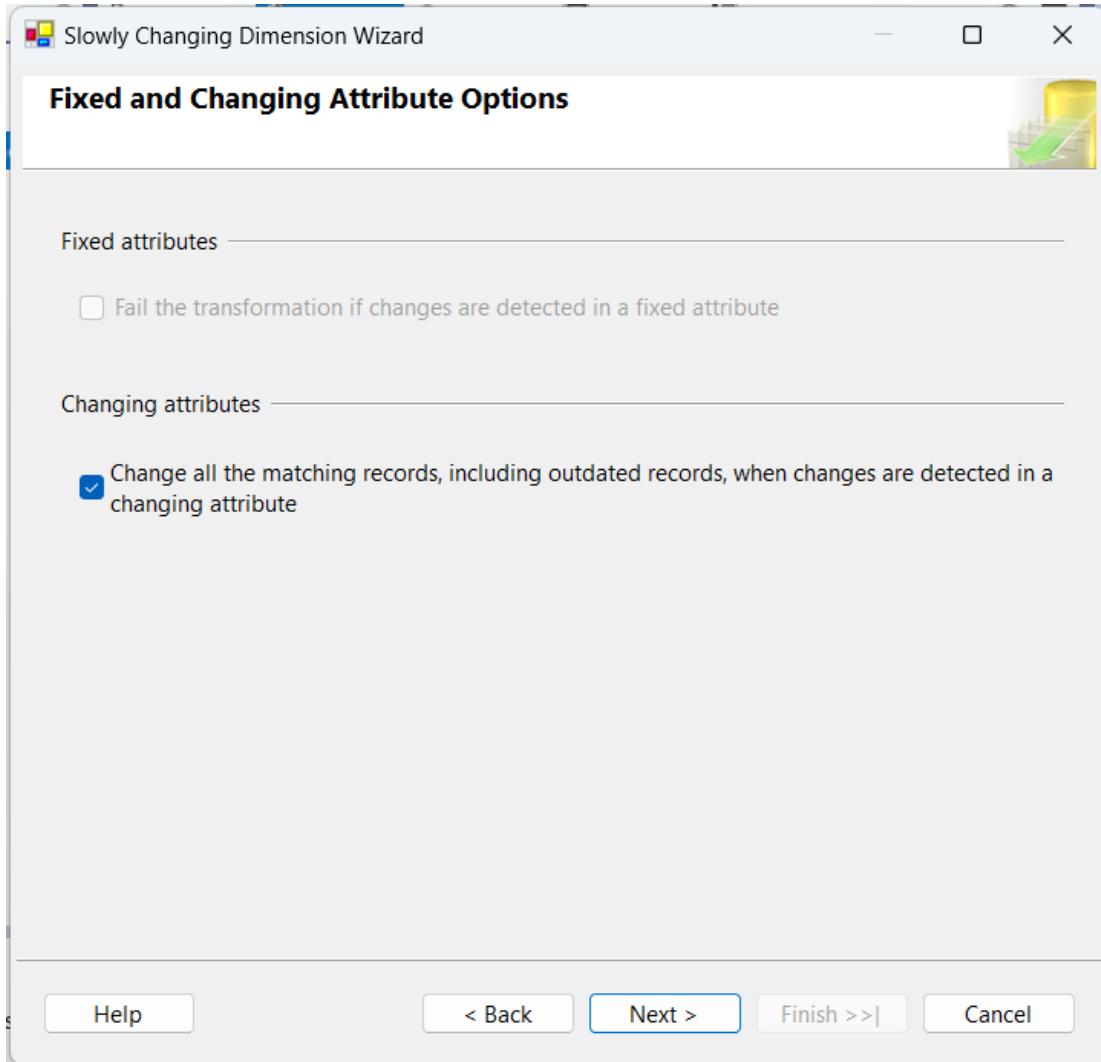
Cấu hình SCD type cho dimension. Các lựa chọn là:

- Fixed -> Không thay đổi
- Changing -> Type 1 (Update)
- Historical -> Type 2 (thêm dòng mới, dòng cũ inactive)

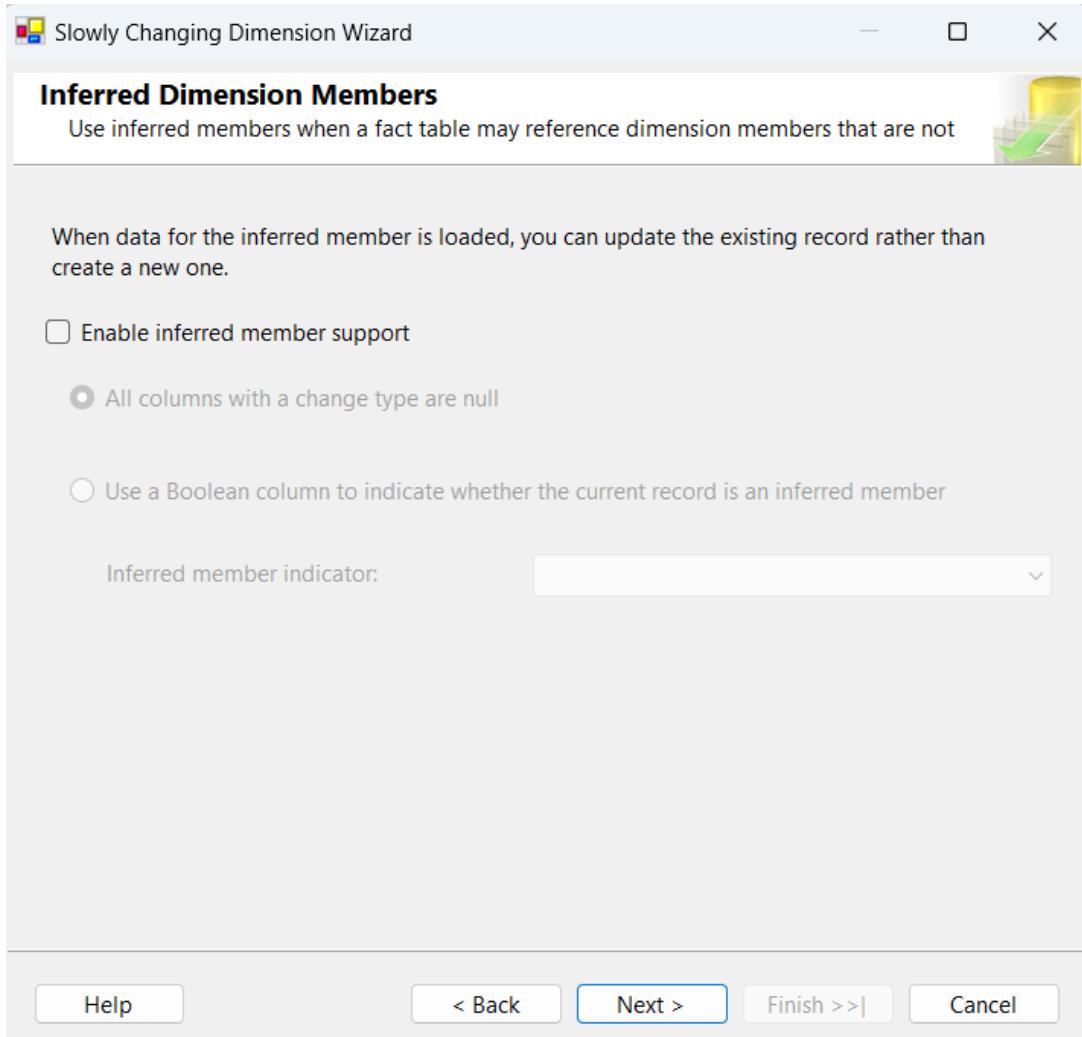


Click **Next**, xuất hiện màn hình **Fixed and Changing Attribute Options**.

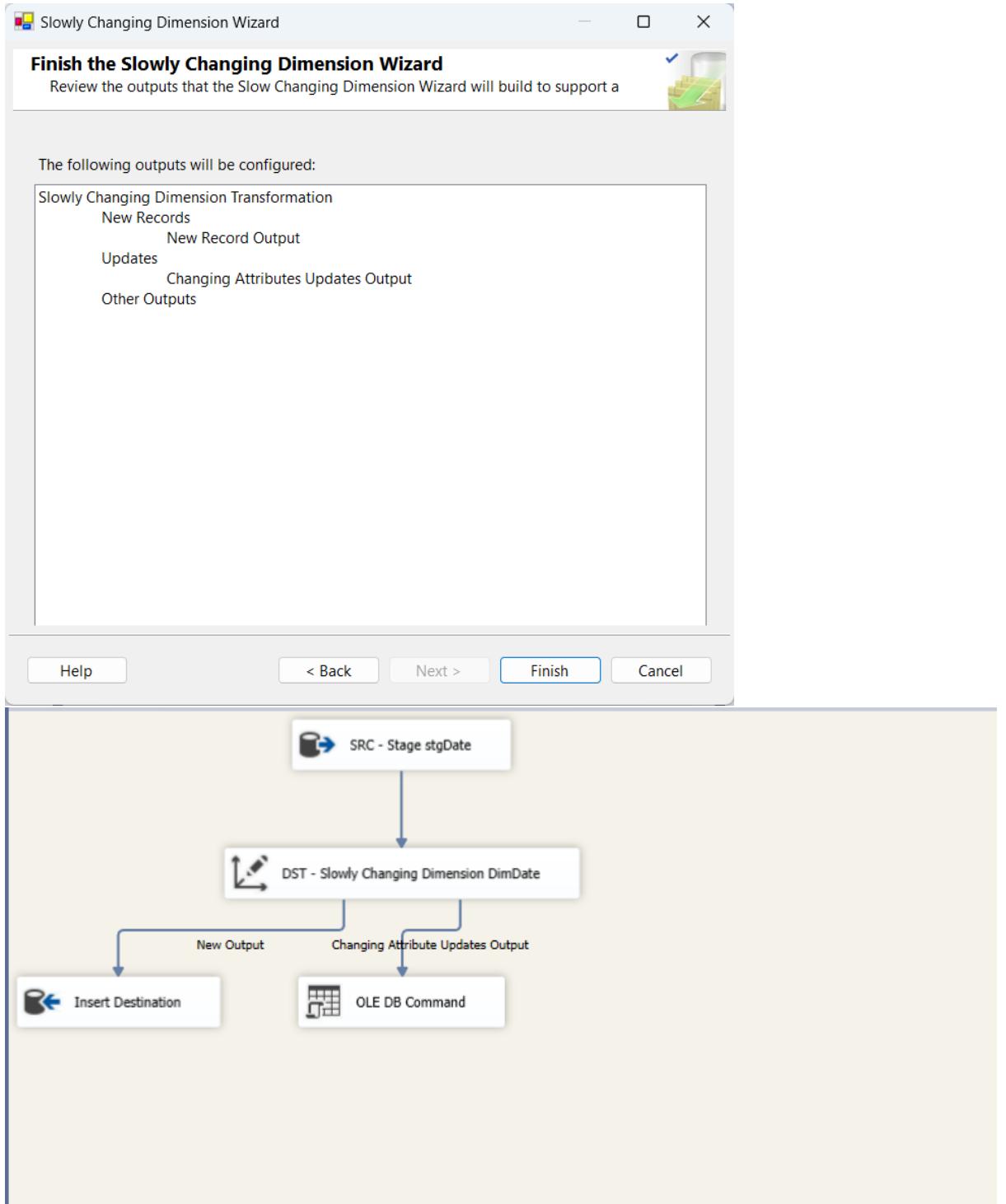
Click **Next**



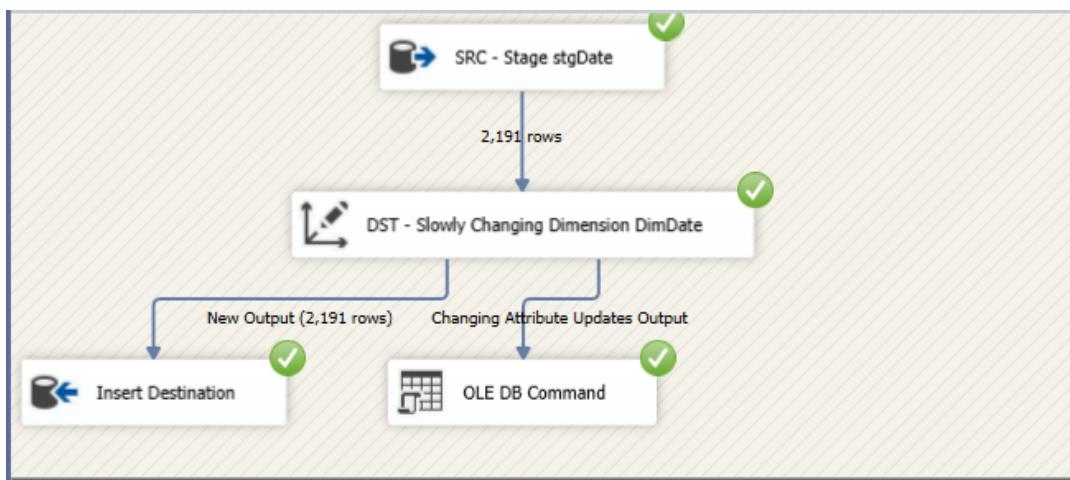
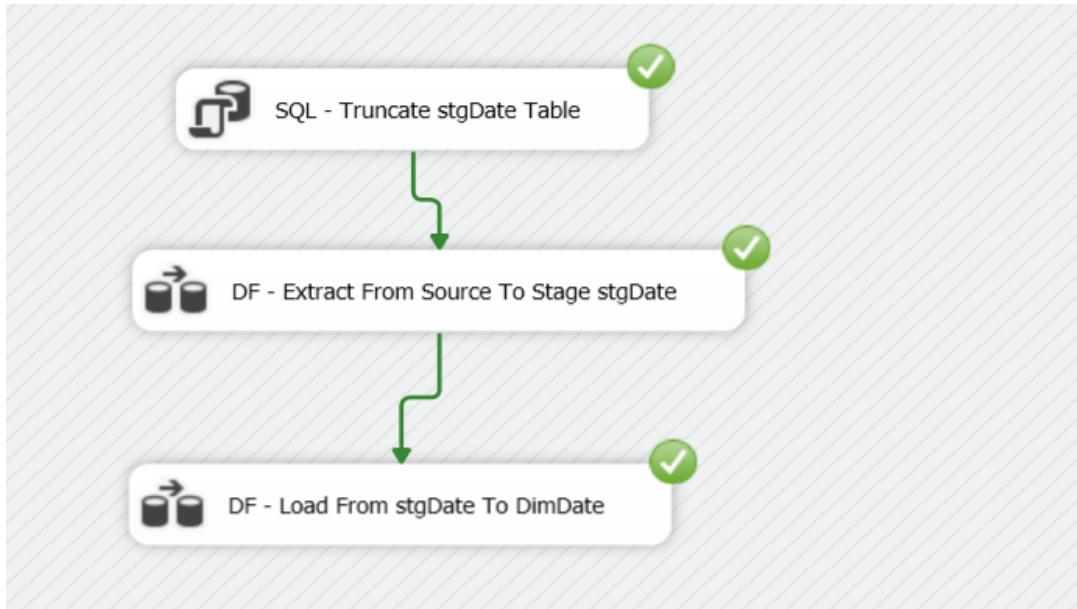
Xuất hiện màn hình **Inferred Dimension Members**. Bỏ chọn check box và click **Next**.



Xuất hiện màn hình **Finish**.

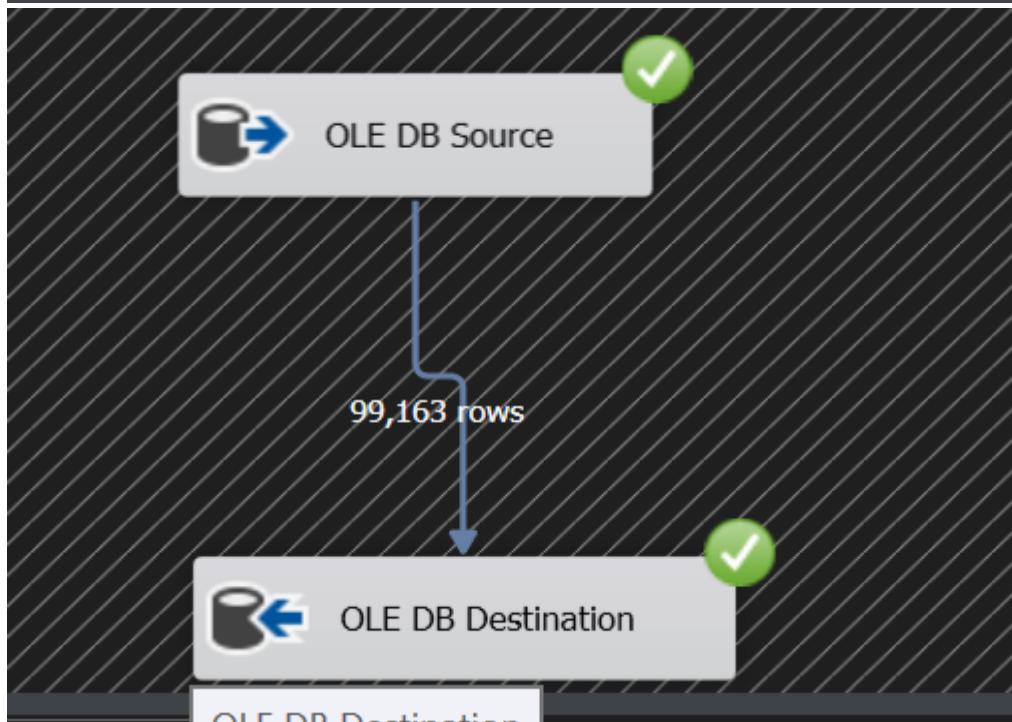
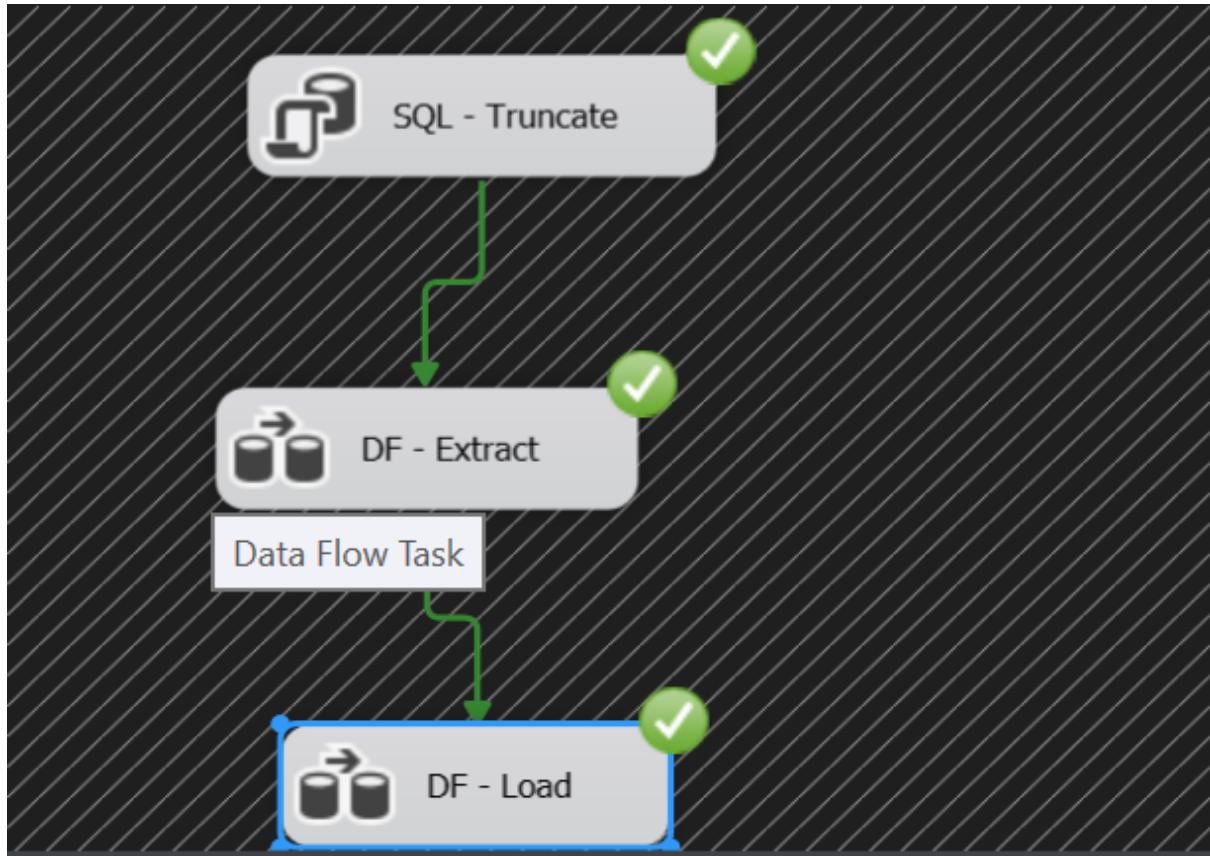


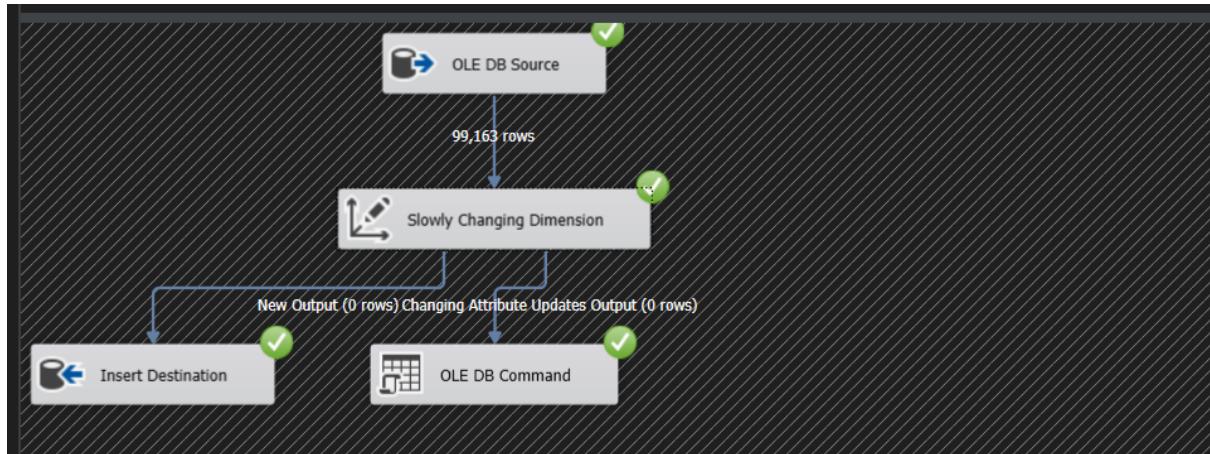
Bấm F5 để chạy. Nếu xảy ra lỗi, có găng đọc hiểu thông báo lỗi để tìm cách sửa.Nếu không có lỗi bạn sẽ thấy màn hình như sau:



### 2.3.2. Customers Dimension

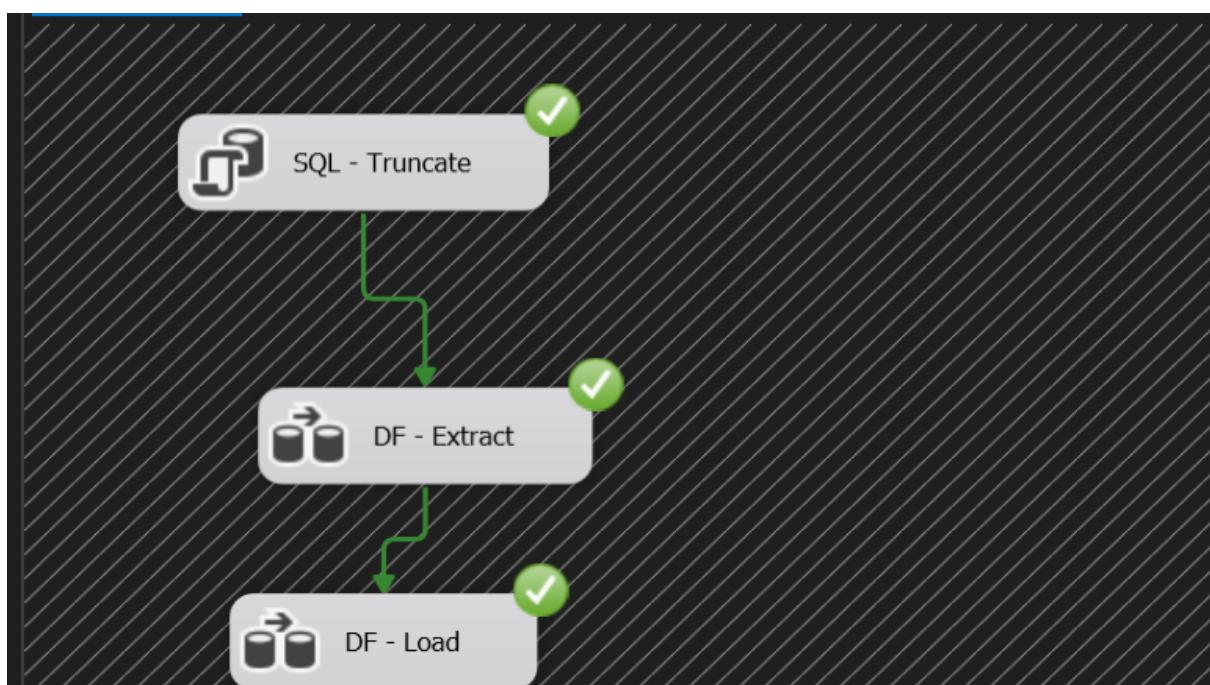
Làm các bước tương tự như Date dimension thì sẽ có kết quả như sau:

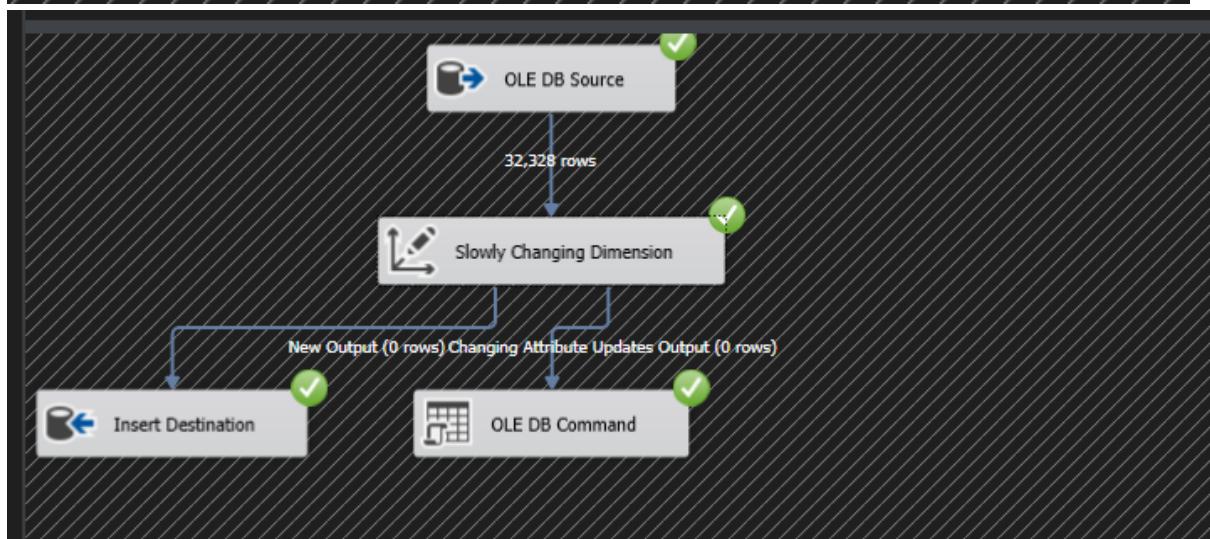
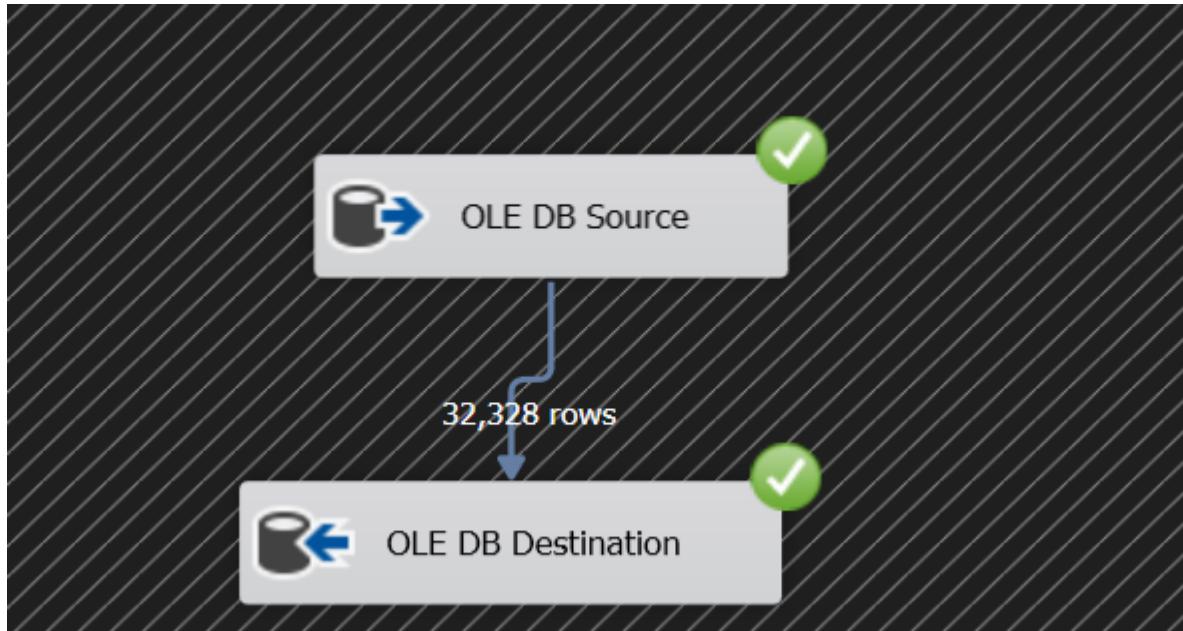




### 2.3.3. Products Dimension

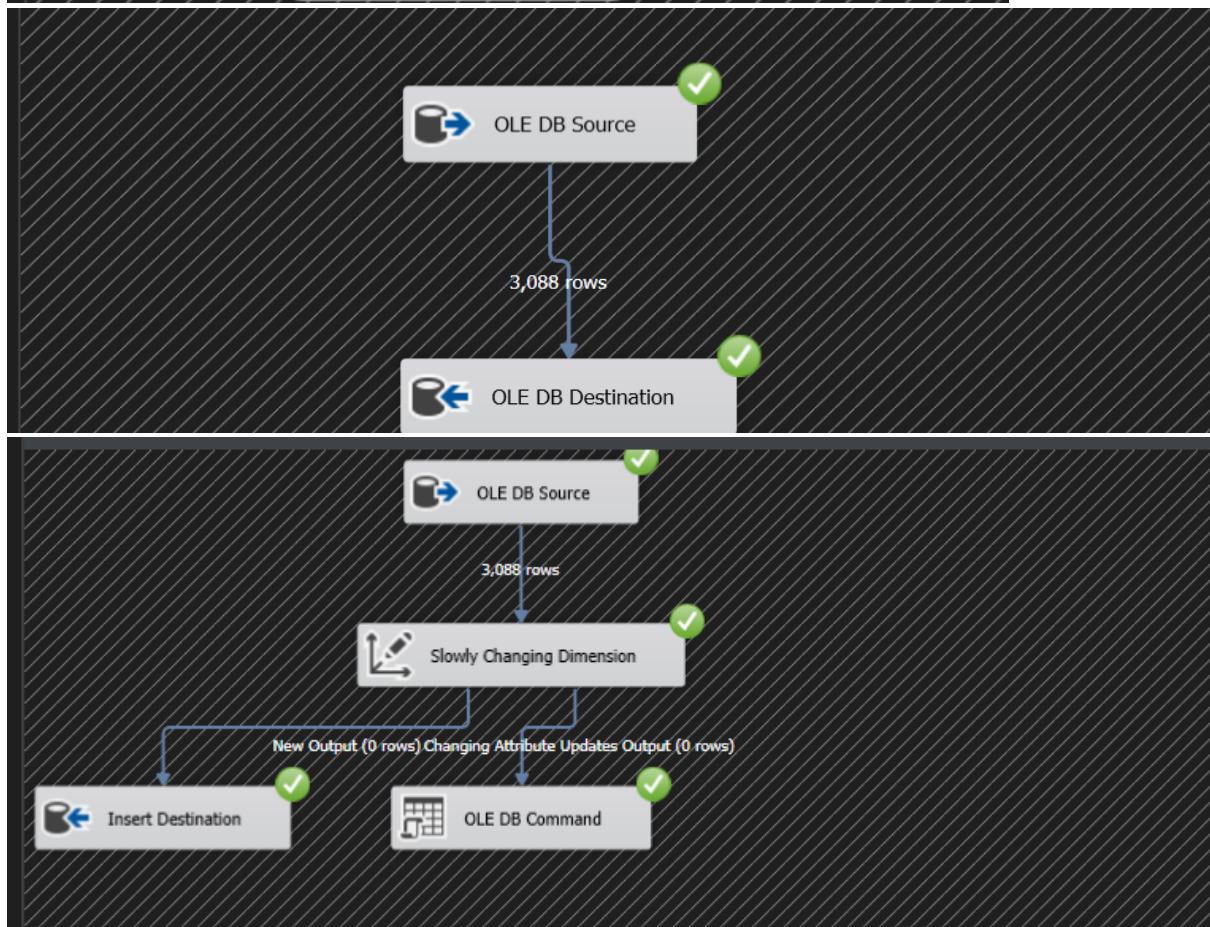
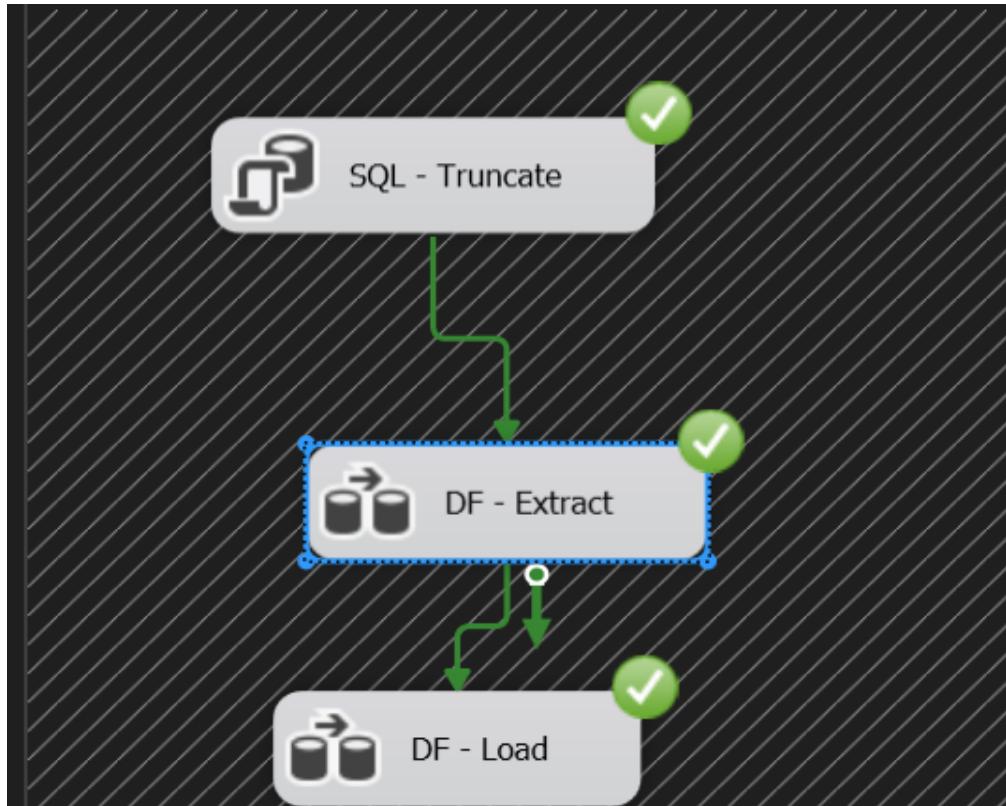
Làm các bước tương như như Date dimension thì sẽ có kết quả như sau:



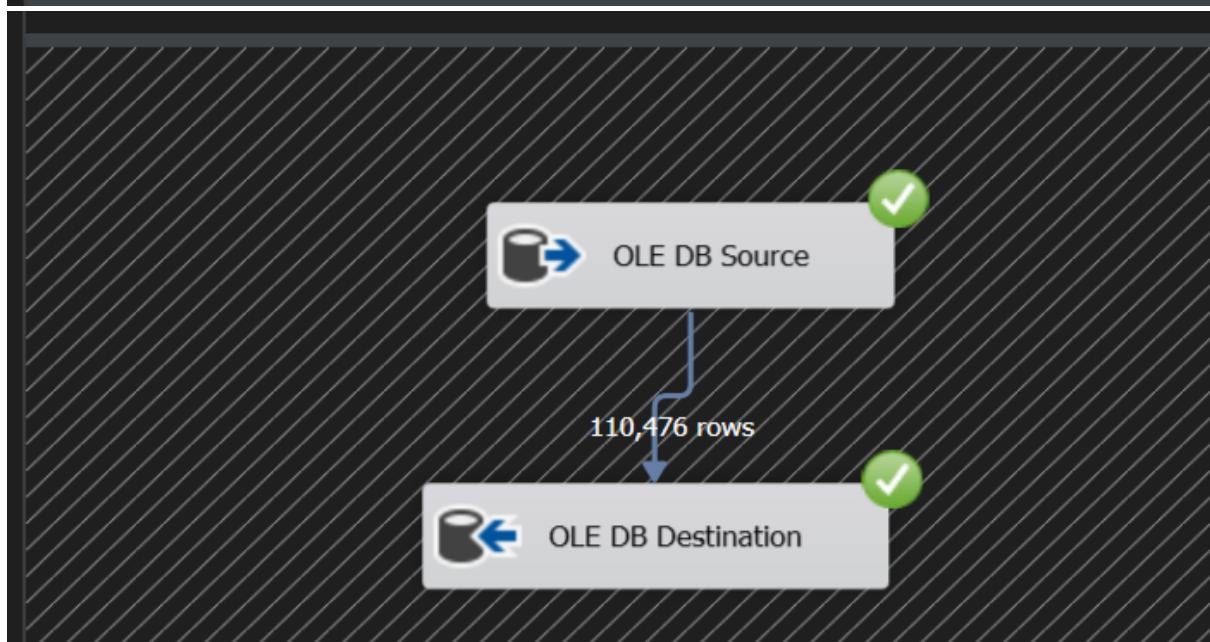
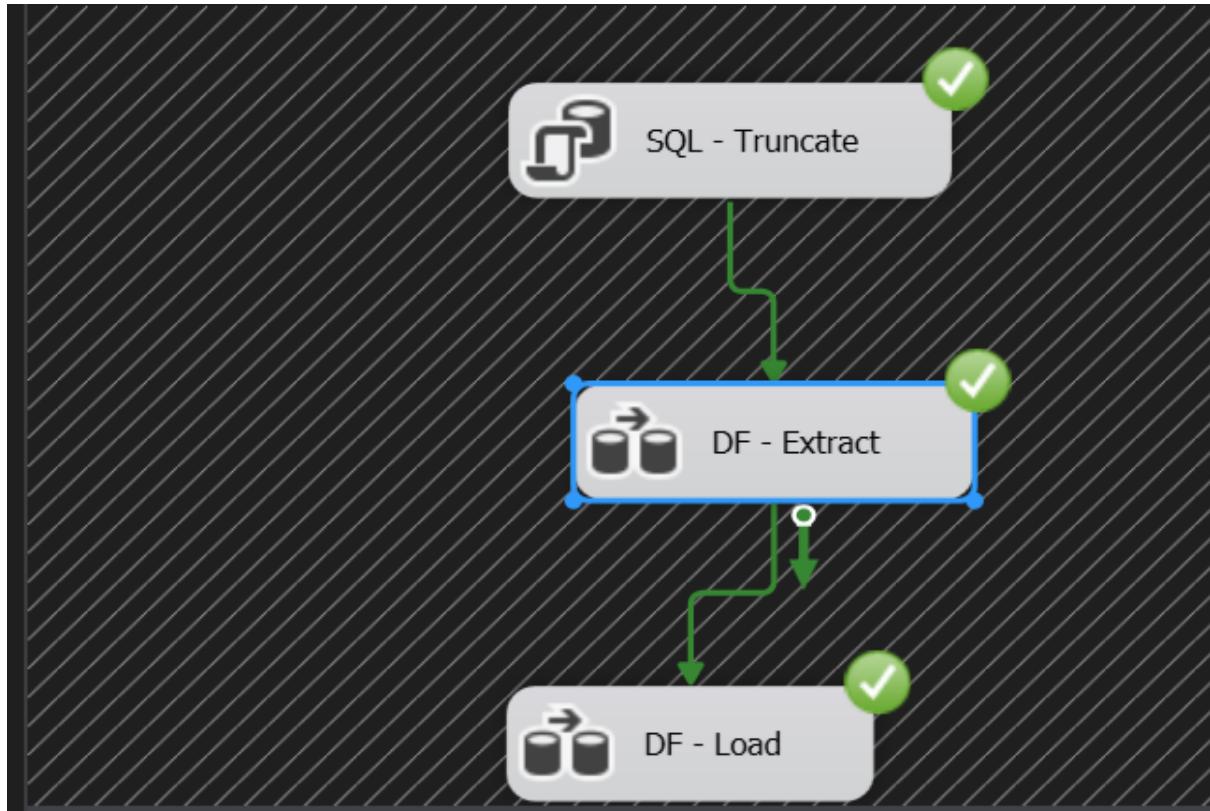


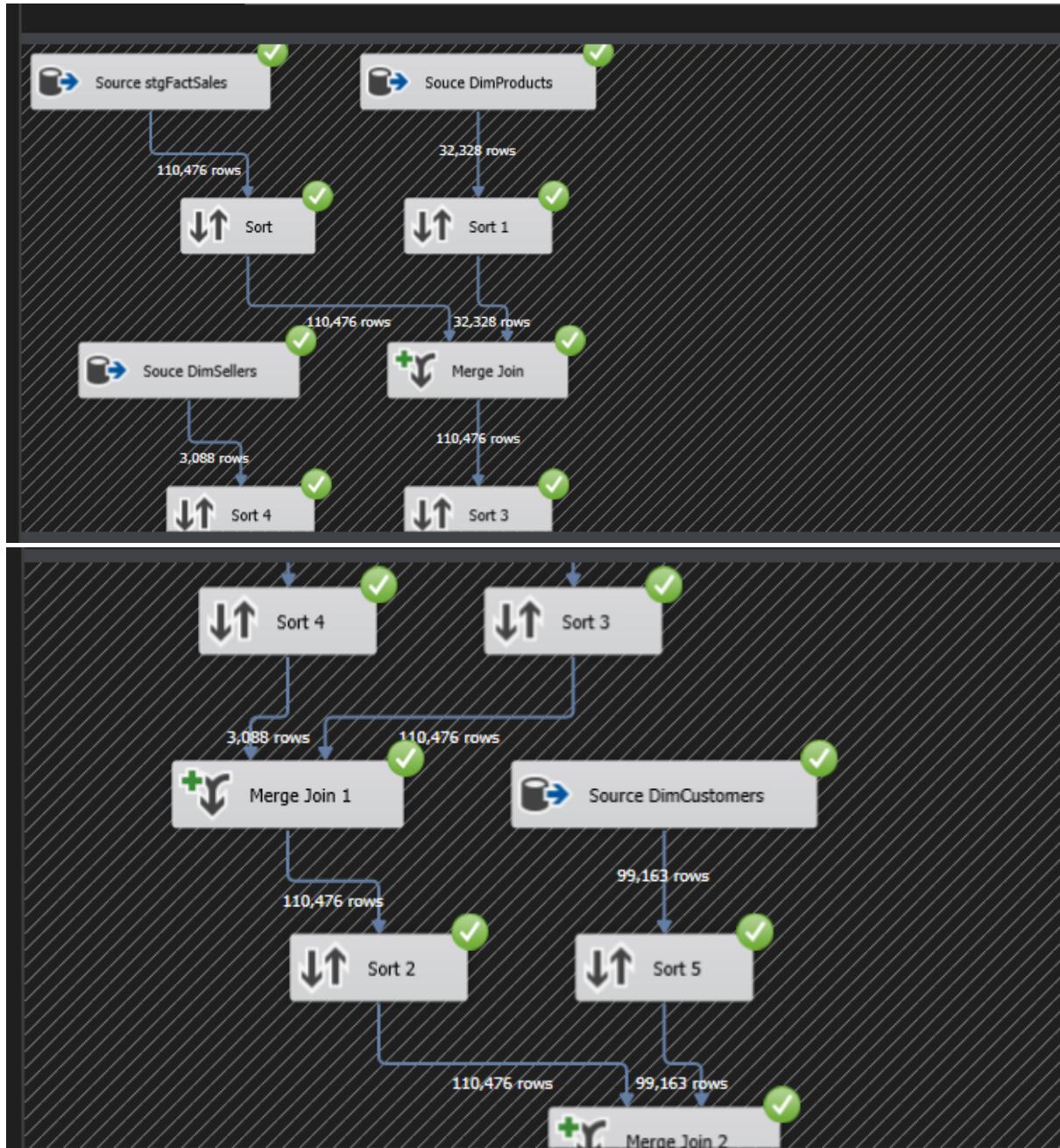
#### 2.3.4. Sellers Dimension

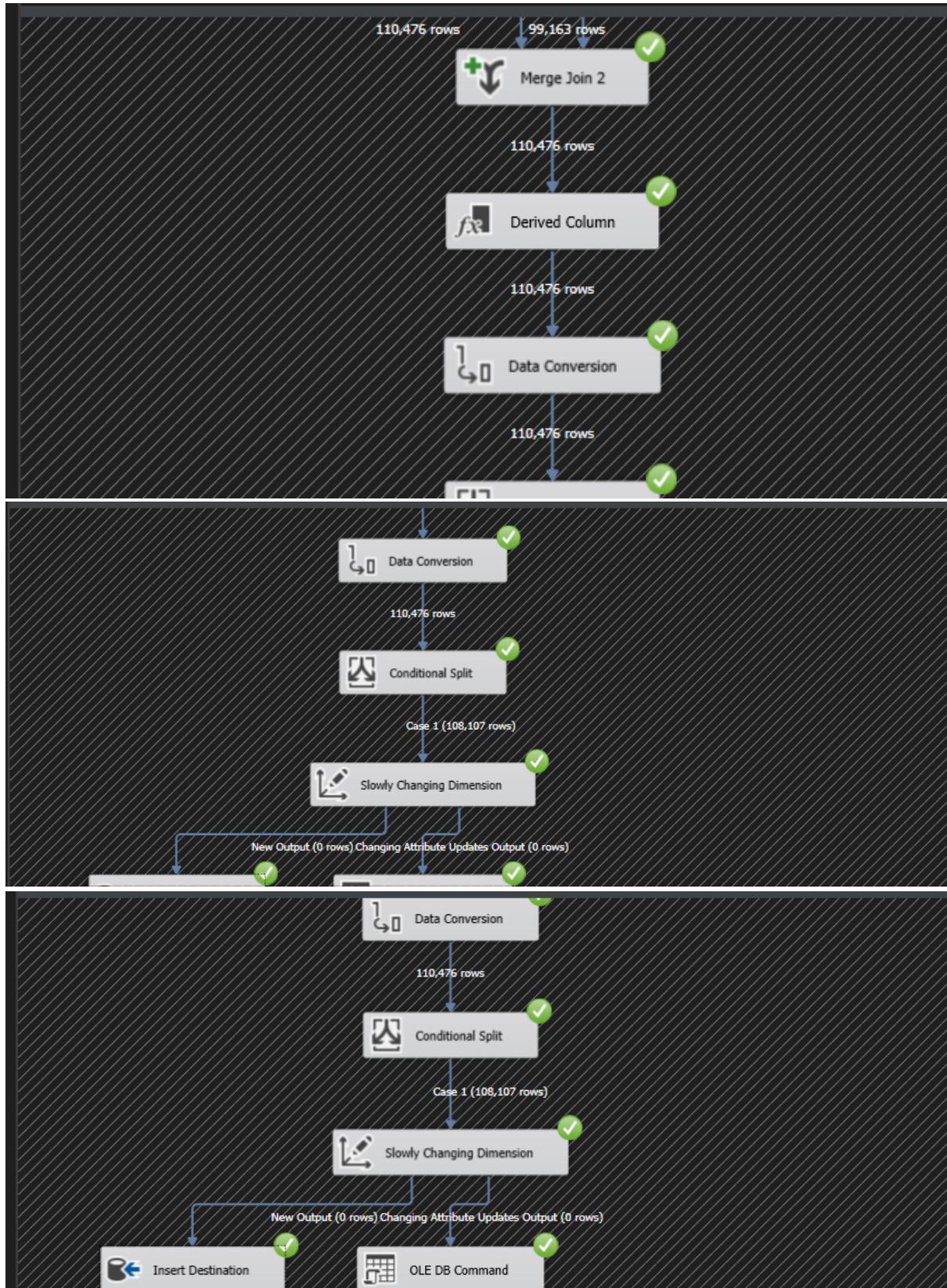
Làm các bước tương tự như Date dimension thì sẽ có kết quả như sau:



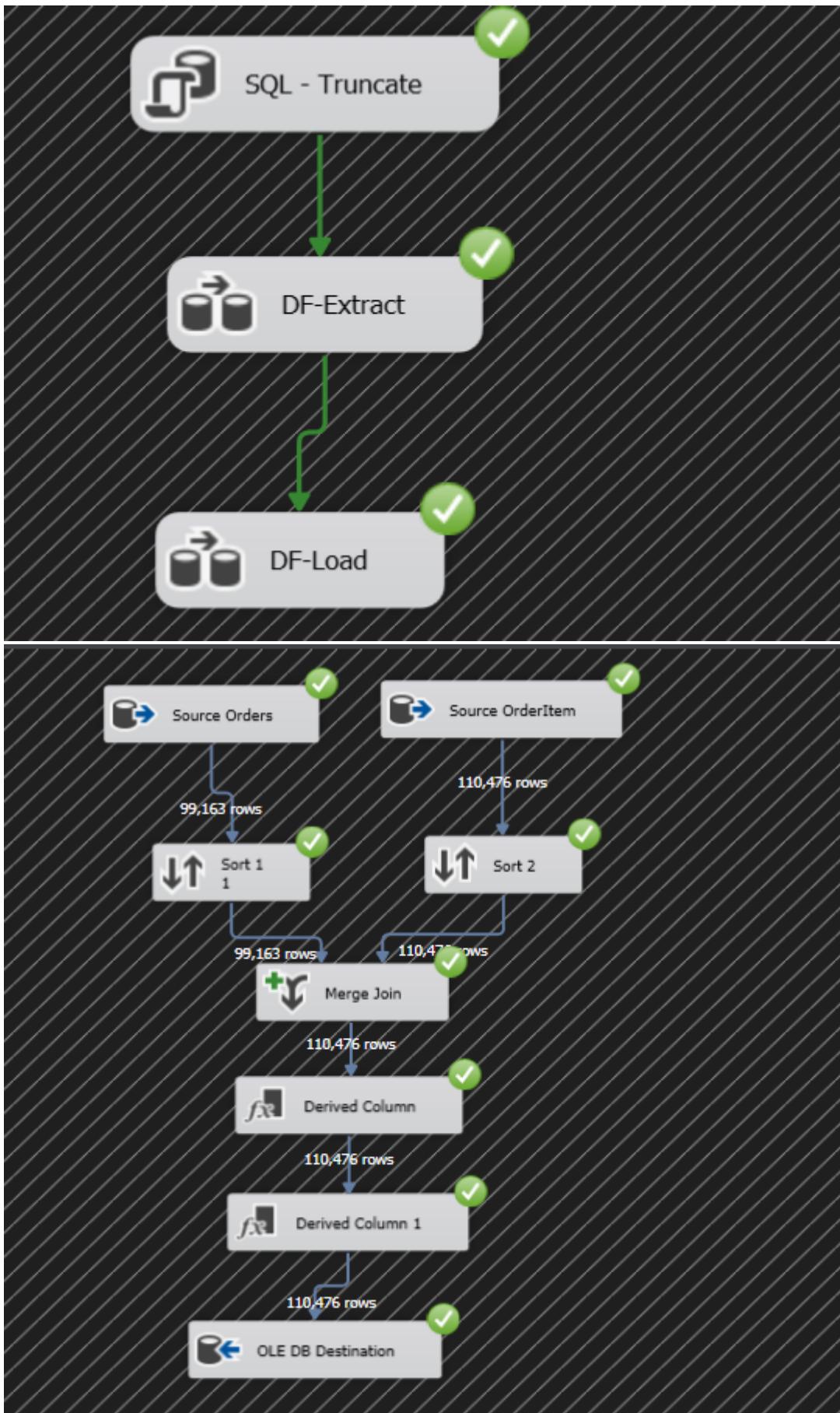
### 2.3.5. Sales Fact

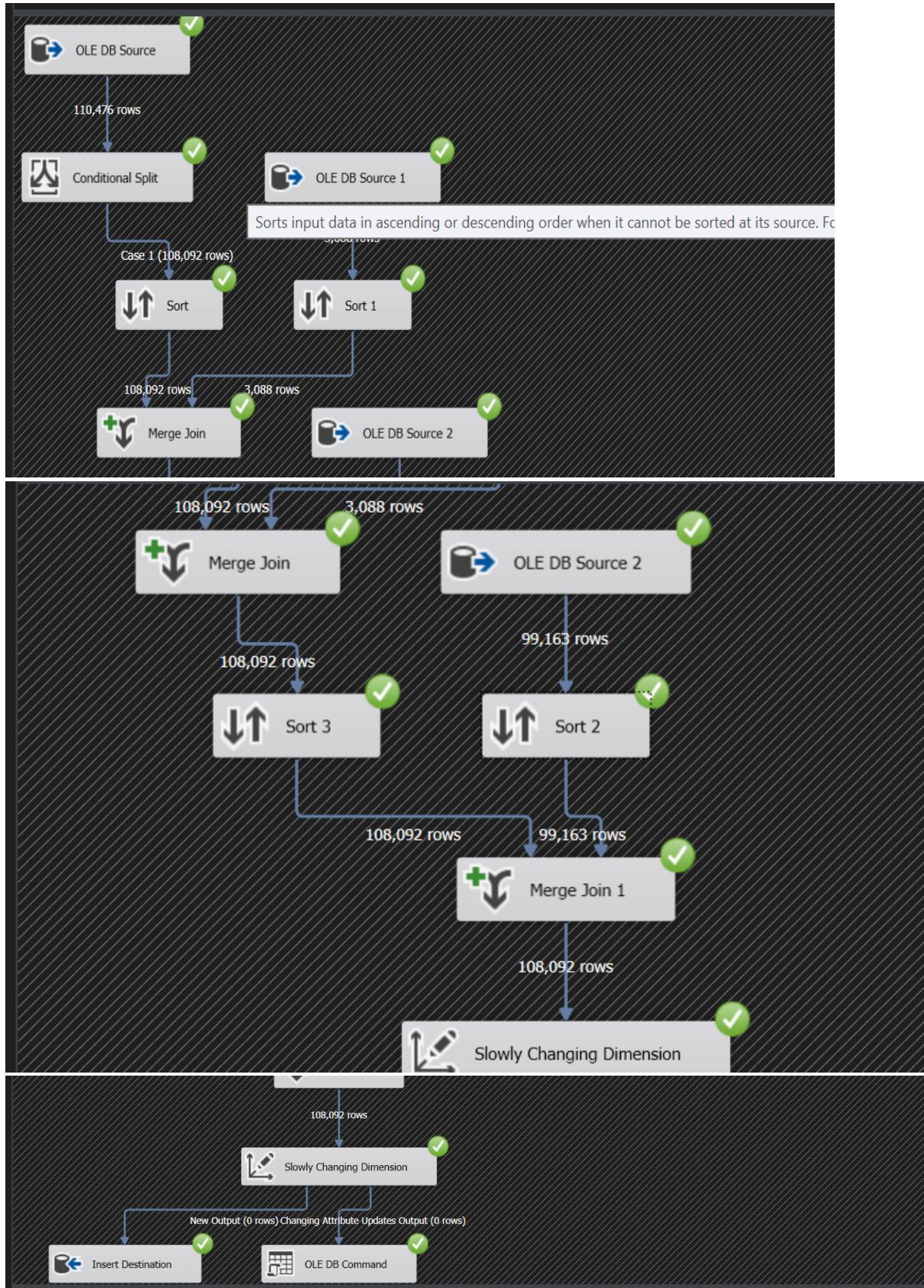






### 2.3.6. Order Fulfillment Fact

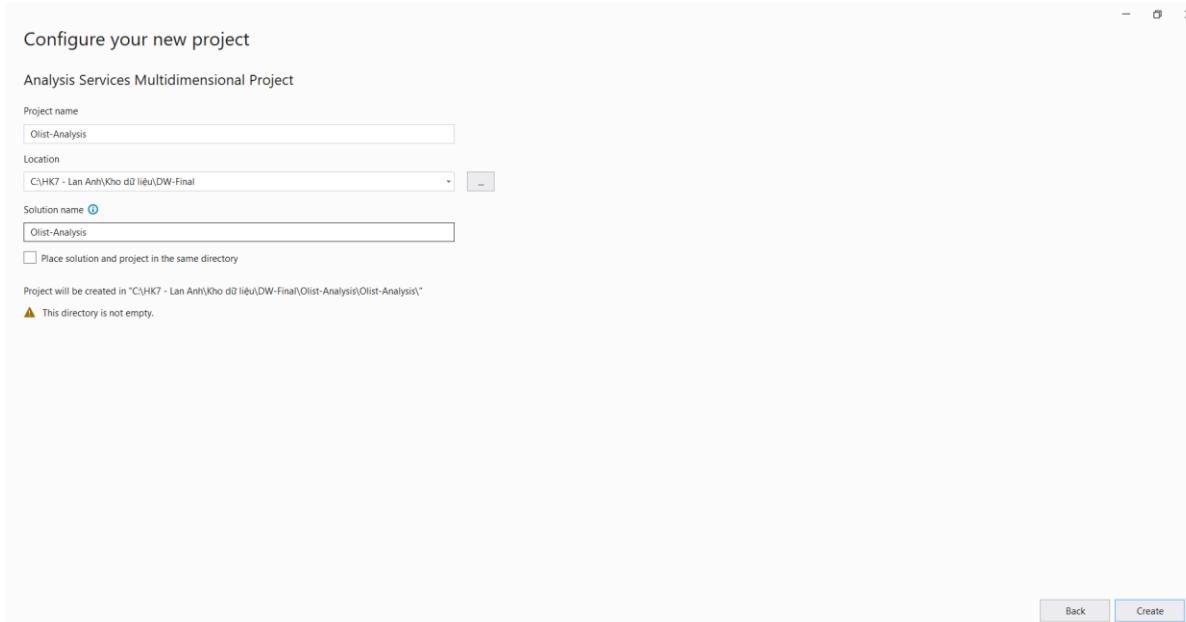




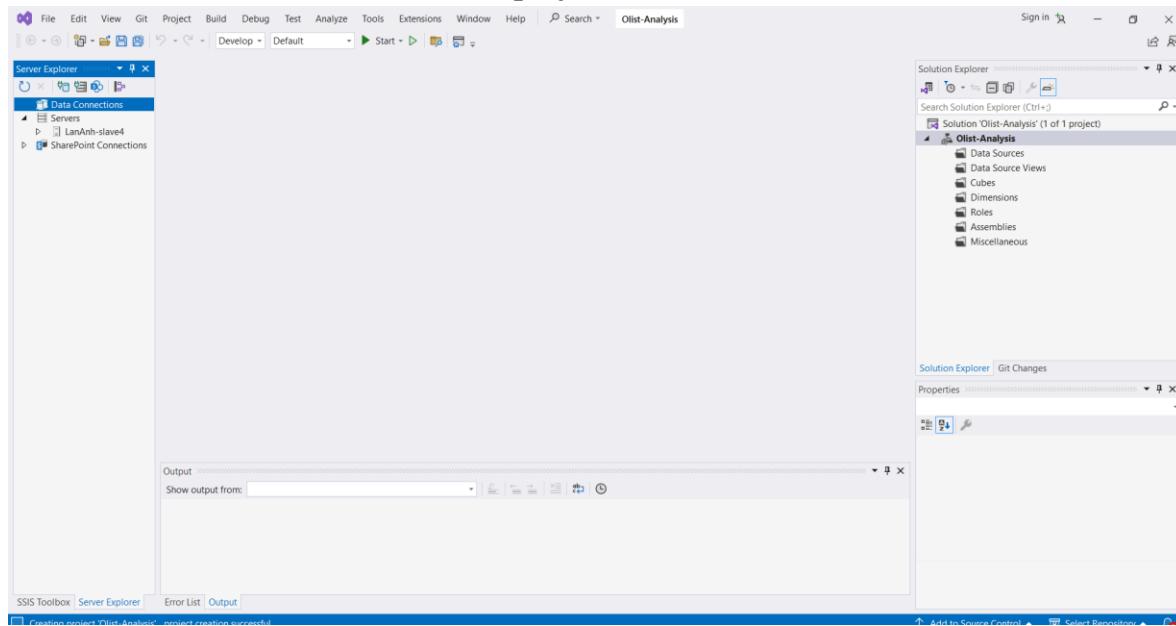
## PHẦN 3: SQL SERVER ANALYSIS SERVICES

### 3.1. Tạo một Analysis Services Project

Mở Visual Studio -> File -> New -> Project -> Analysis Services Multidimensional Project, đặt tên cho project và click OK.



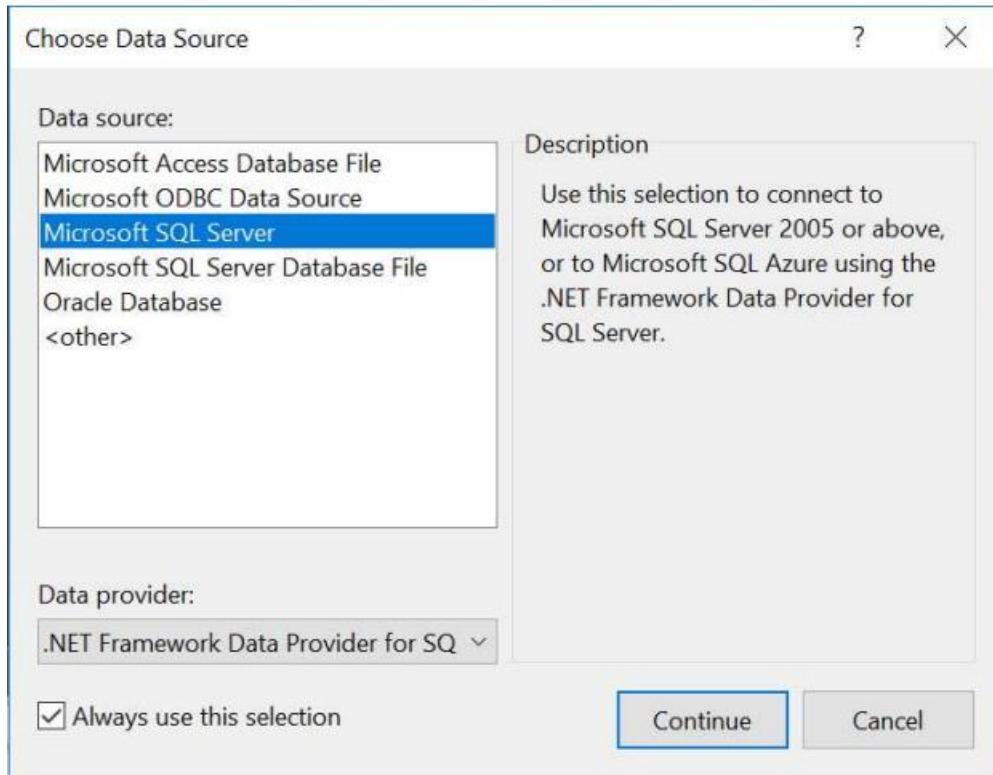
Giao diện của SSAS sau khi tạo một project mới như hình bên dưới:



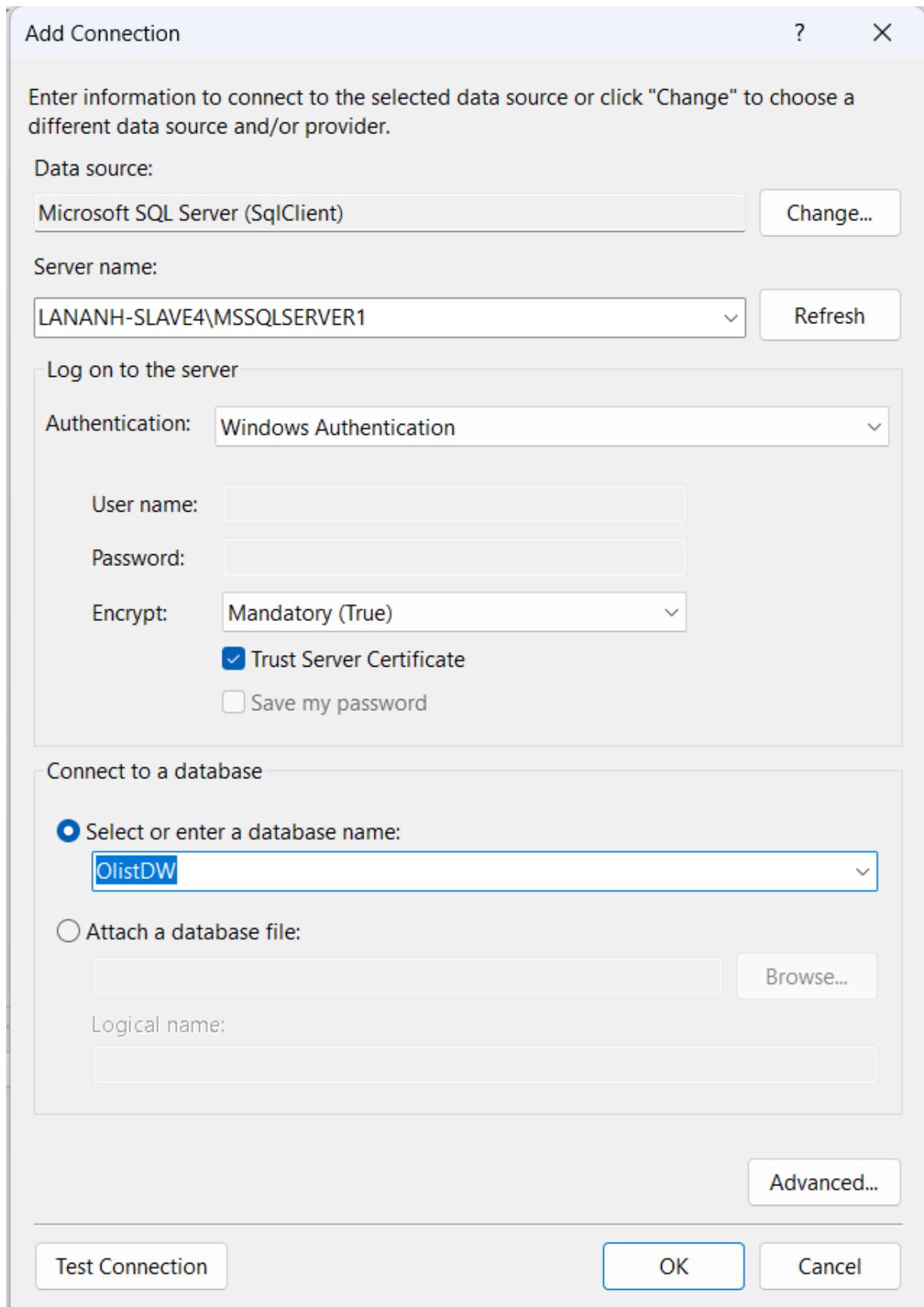
### 3.2. Kết nối đến Analysis Services Server

Right click vào “Data Connections” ở mục Server Explorer chọn “Add Connection ...” sẽ

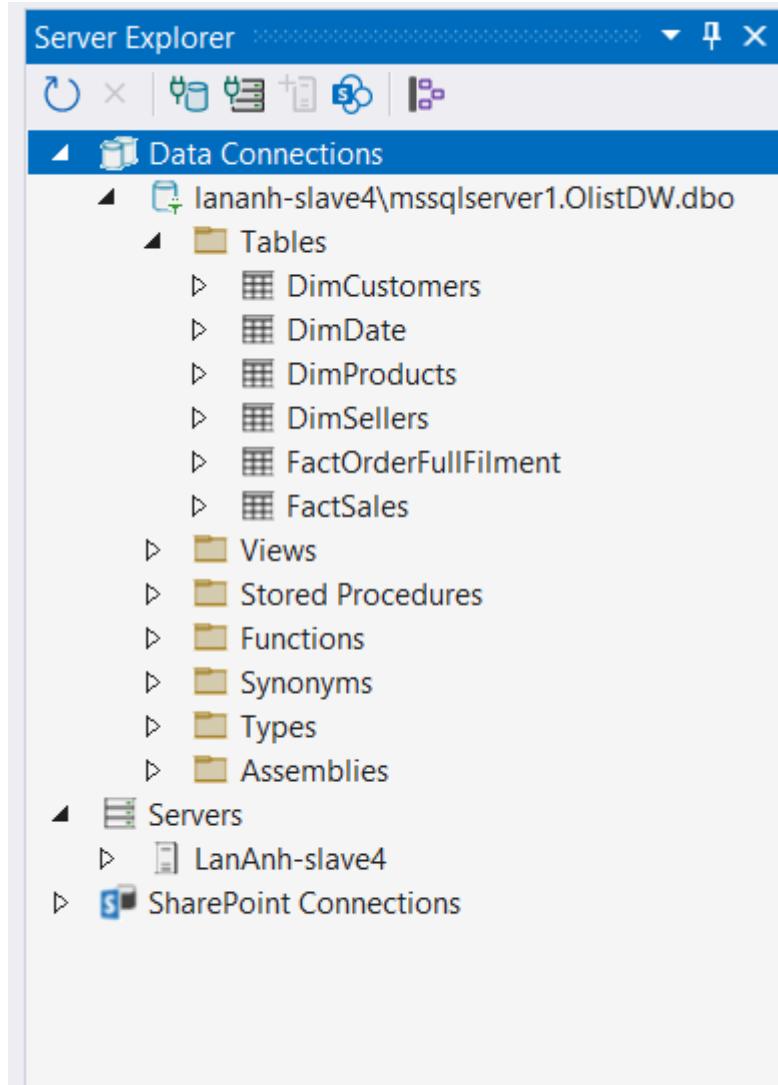
xuất hiện dialog “Choose Data Source”, chọn “Microsoft SQL Server” và click “Continue”.



Xuất hiện dialog “Add Connection”, chọn “Data Source”, “Server name”, và “database name” phù hợp. Click OK.



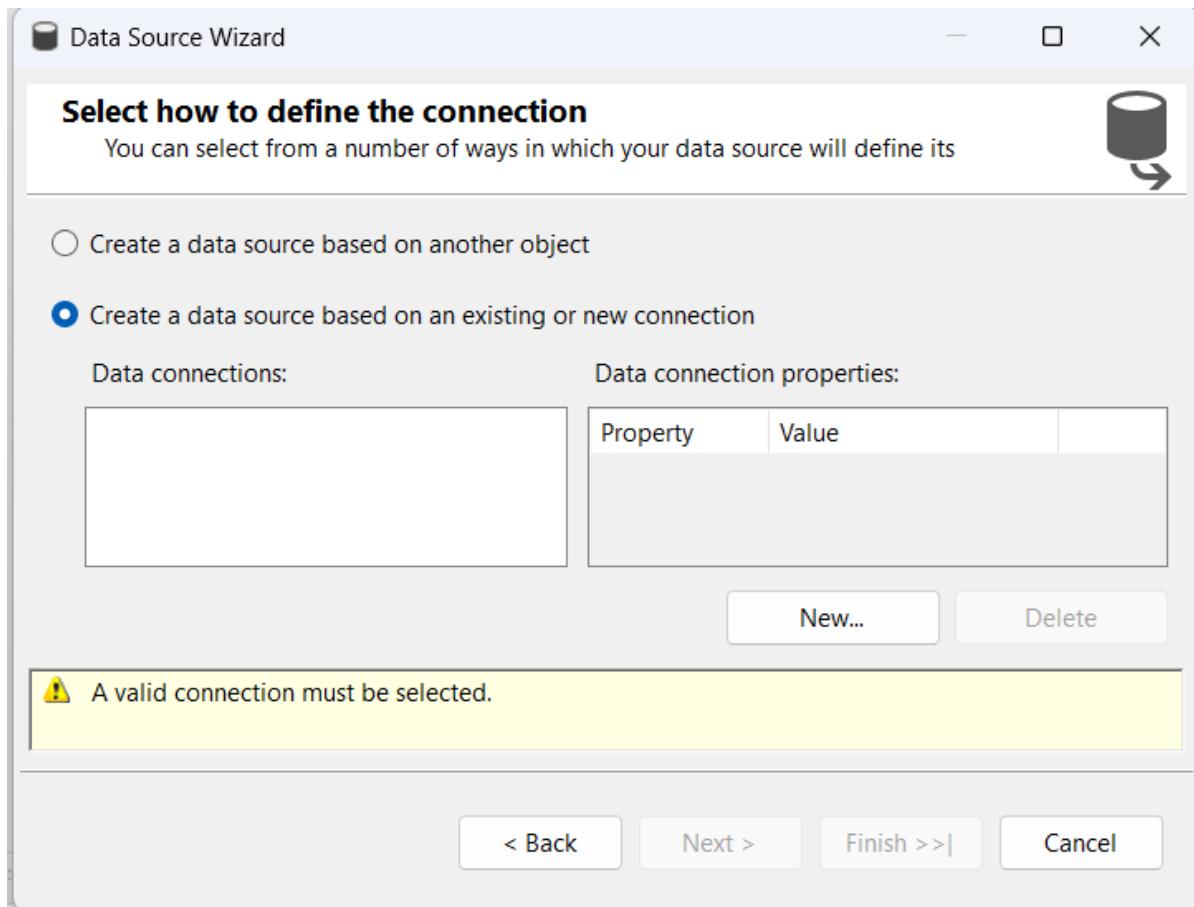
Trở lại màn hình chính, ta sẽ thấy thông tin về “Data Connections” được cập nhật và hiển thị như sau:



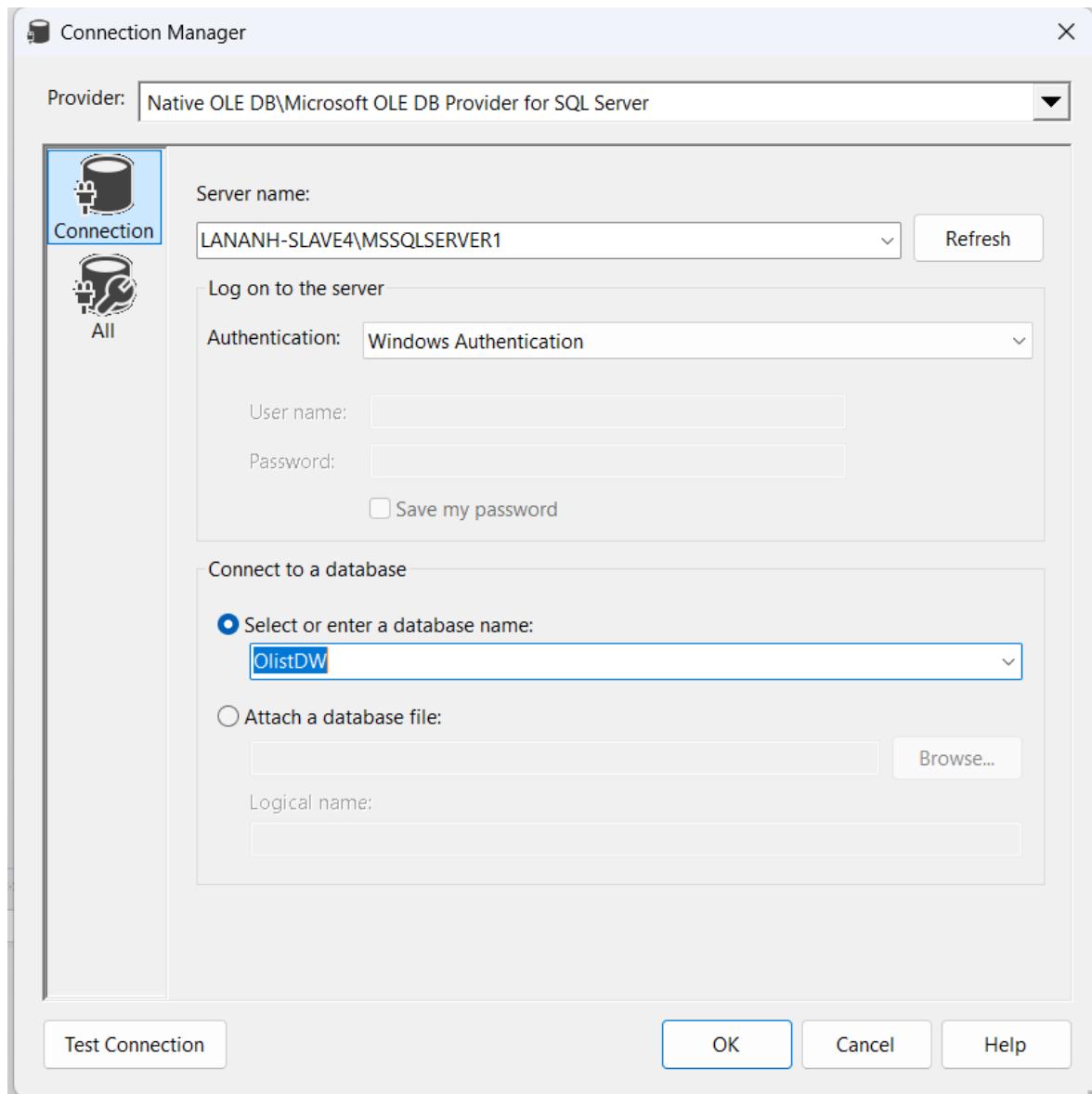
### 3.3. Tạo một Data Cube

#### Tạo một data source cho project

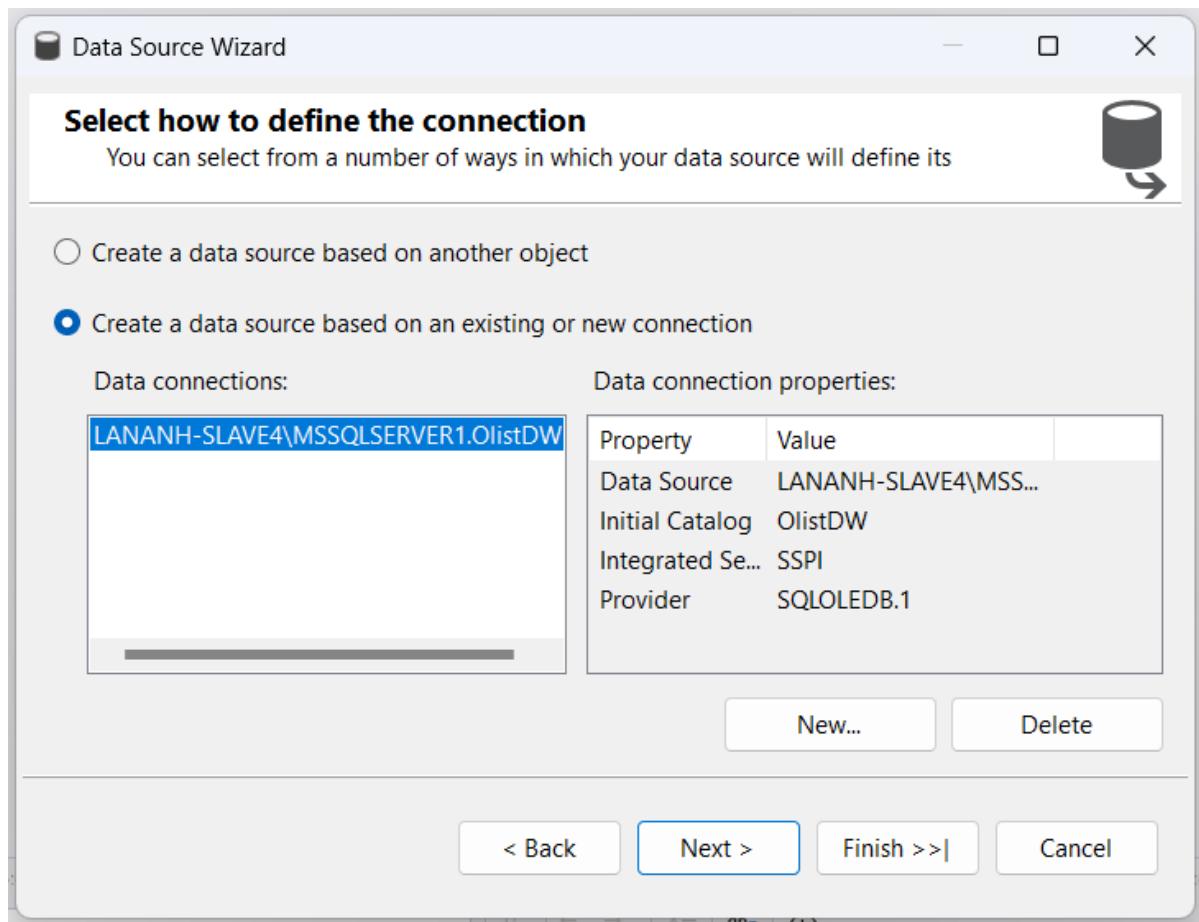
Right-click vào “Data Source” ở Solution Explorer và click New Data Source... sẽ xuất hiện Data Source Wizard dialog.



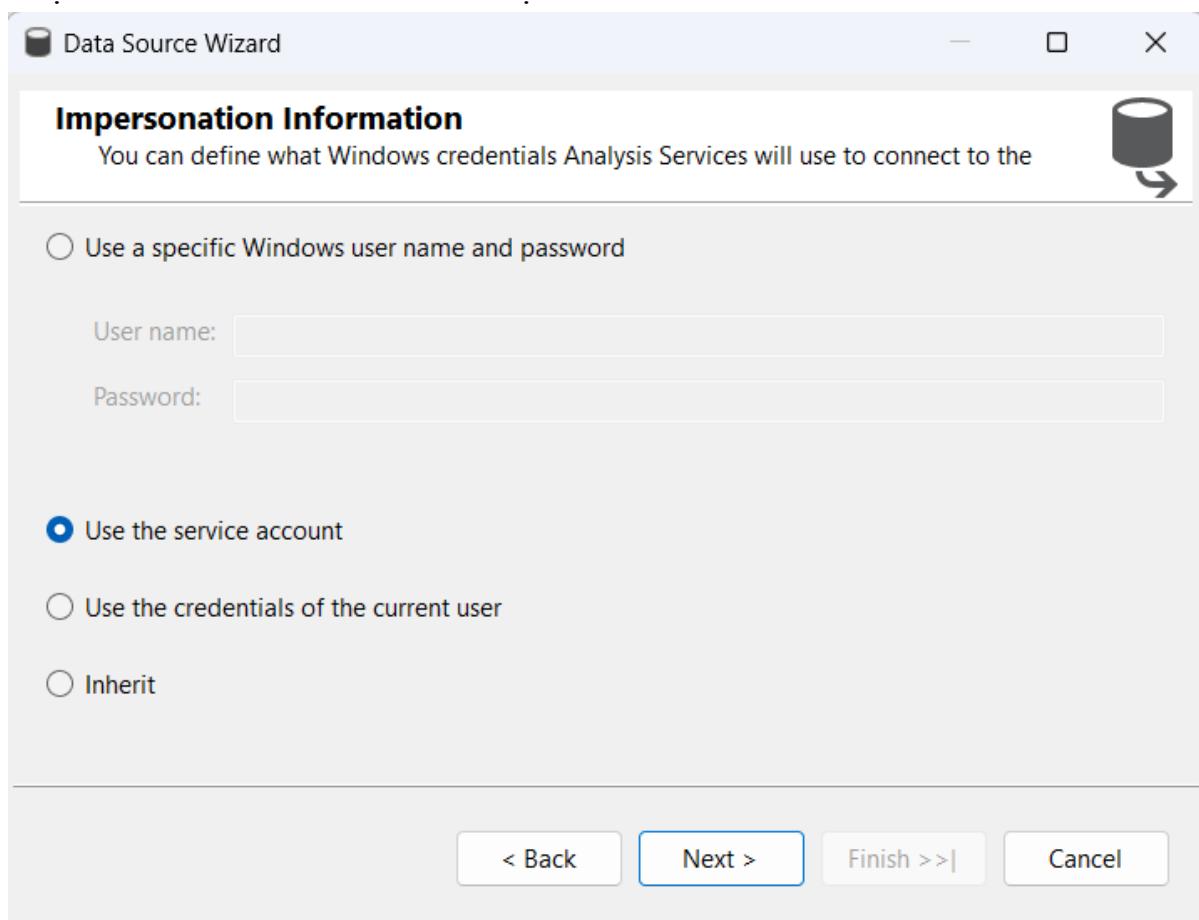
Click Next. Chọn option “Create a data source based on an existing or new connection” và click New...

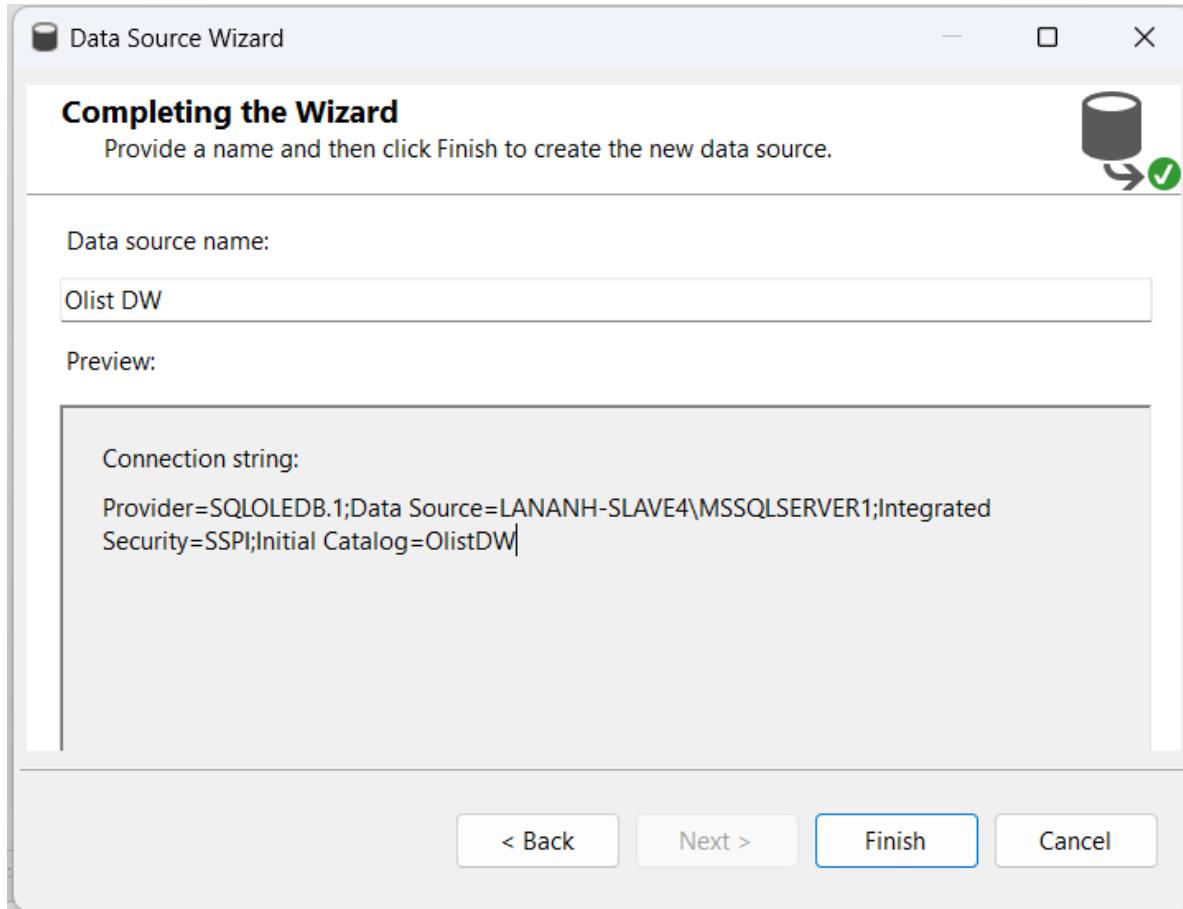


Điền thông tin Server name phù hợp và đặt tên database name và click OK.



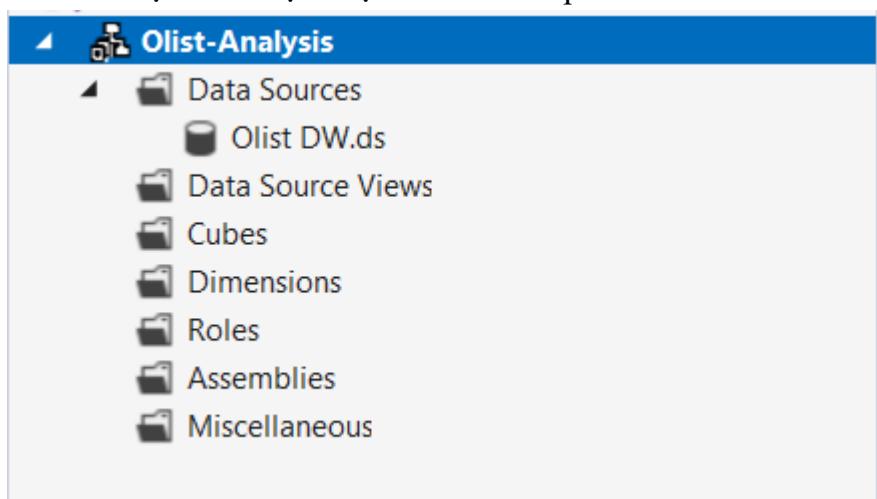
Chọn OlistDW connection vừa mới tạo và click Next





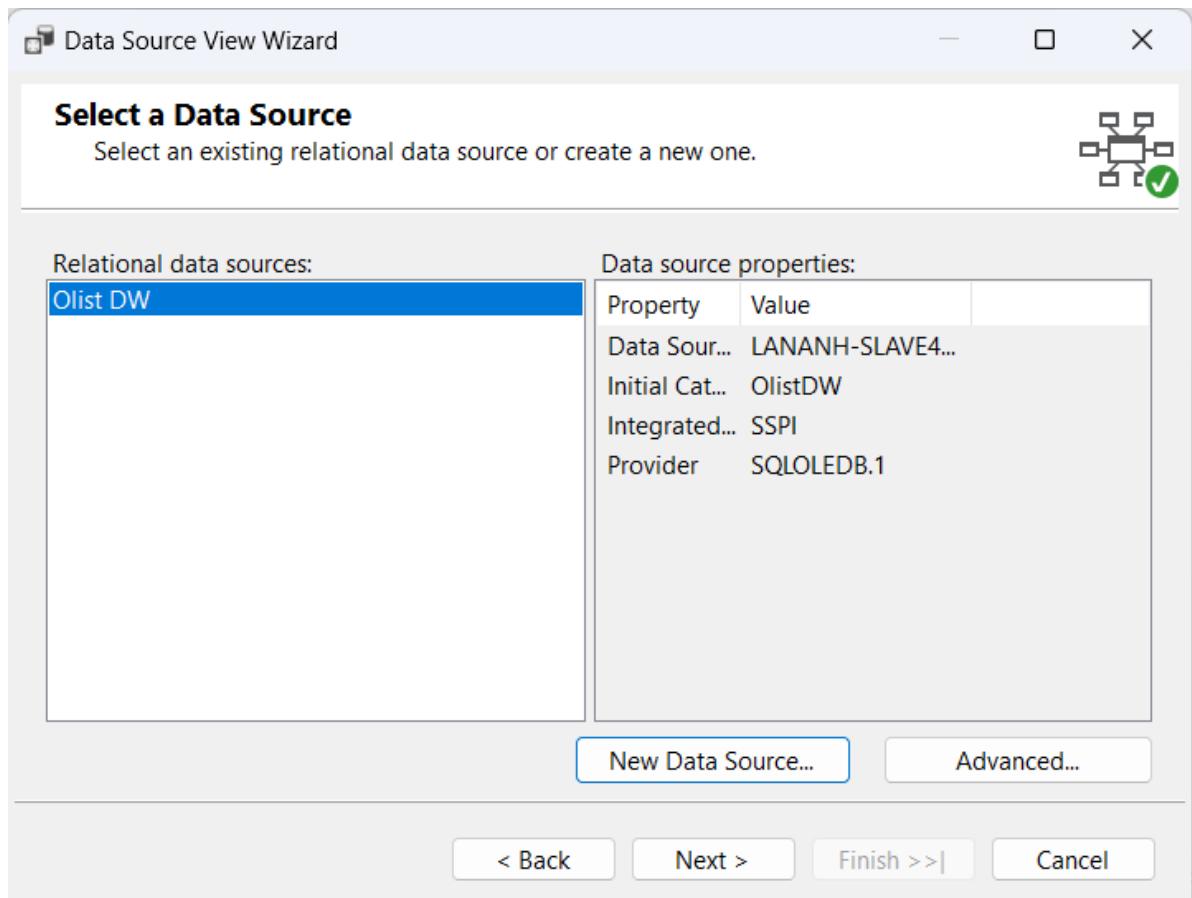
Click Finish để hoàn thành việc kết nối. Sau khi hoàn thành bạn sẽ thấy thông tin về Data

Source được hiển thị ở mục Solution Explorer như hình bên dưới.

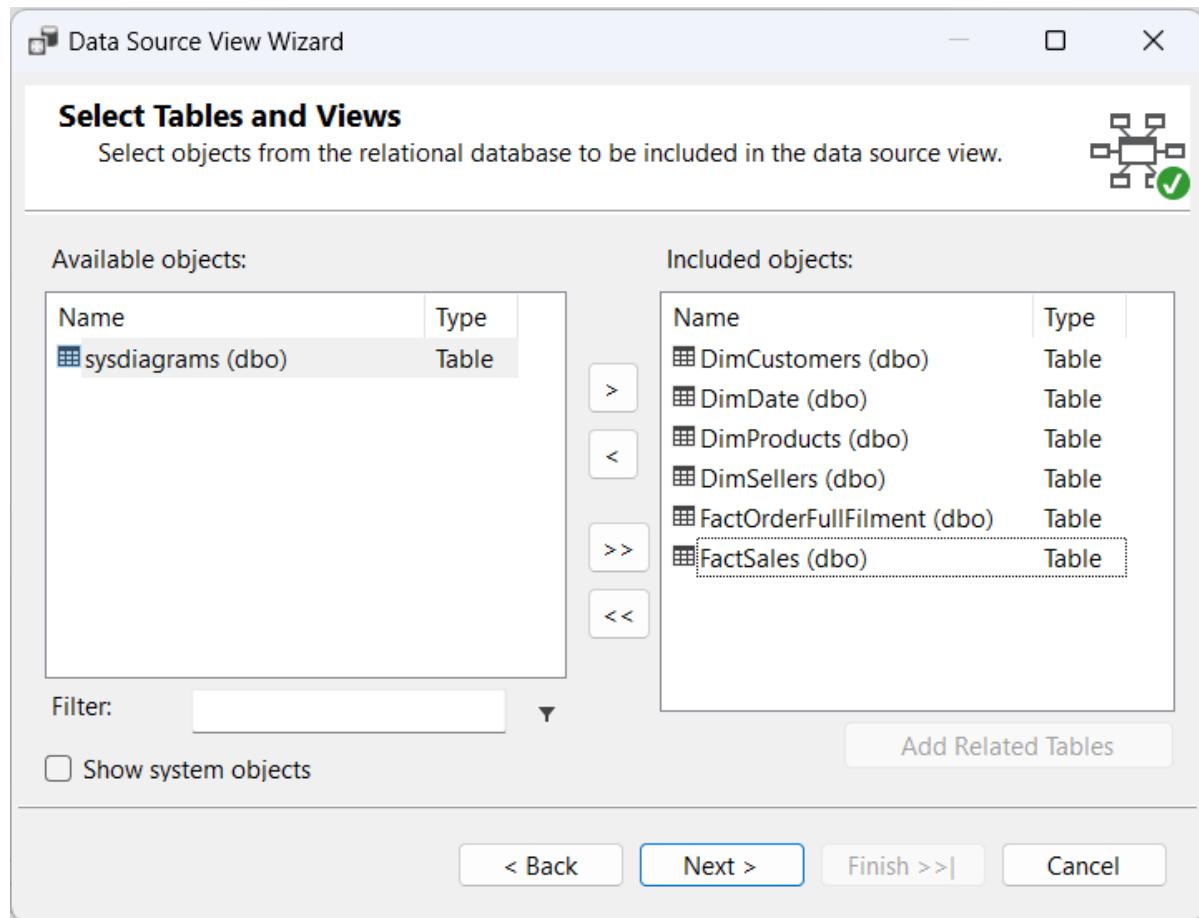


### Tạo data source view

Data source view chứa logical model của CSDL (tables, keys, columns, và các constraints) sẽ được sử dụng bởi OLAP database để tạo các data cube.

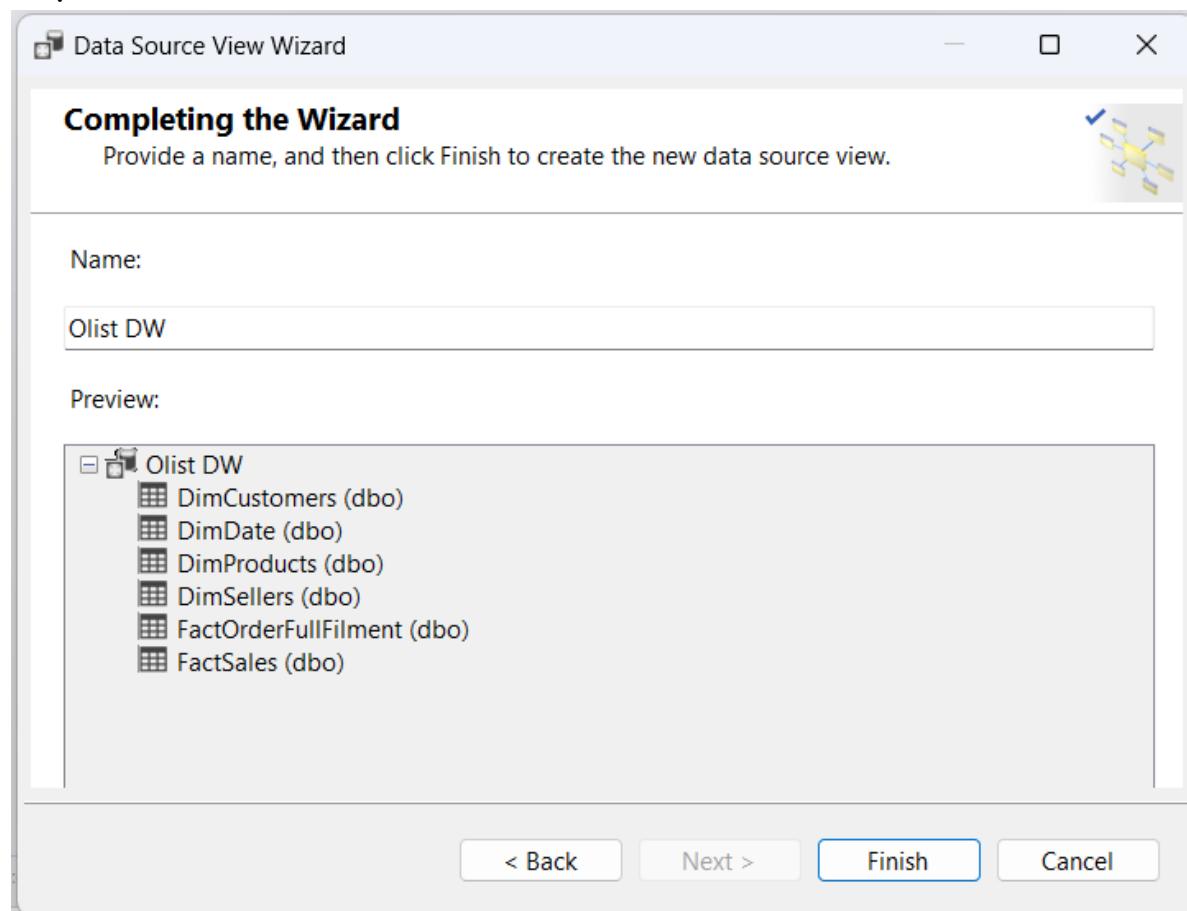


Right-click vào Data Source Views từ Solution Explorer và click New Data Source View... sẽ xuất hiện Data Source View Wizard, click Next,

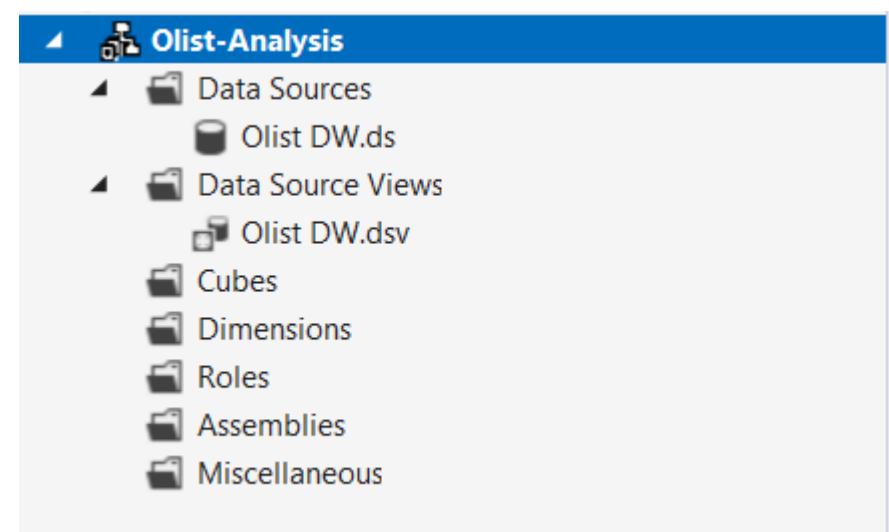


Click Next

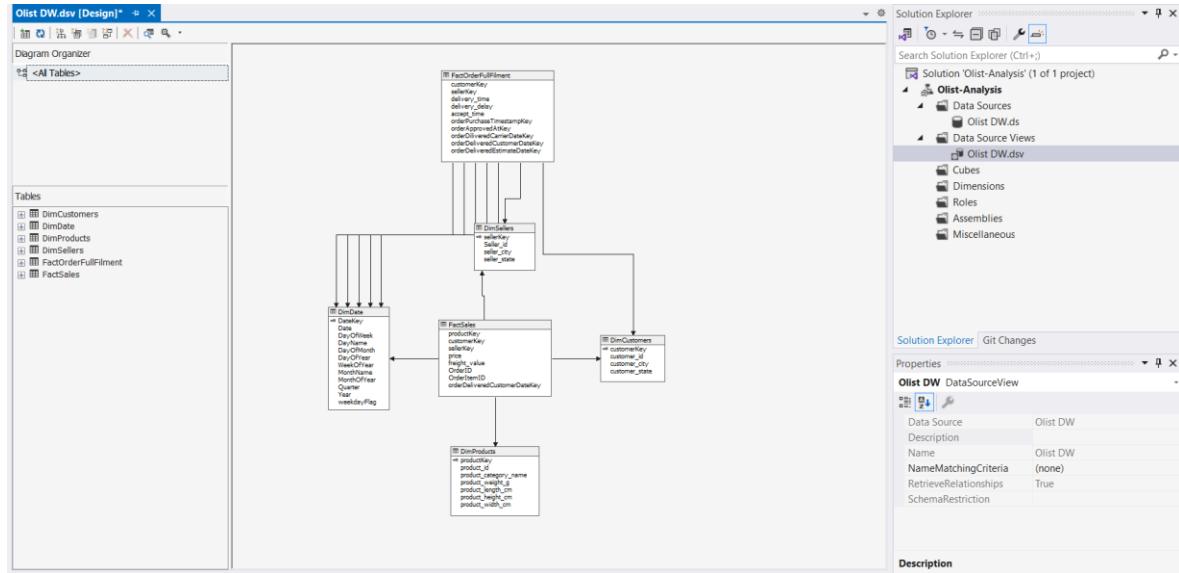
Chọn tất cả các table cho dimensional model và click Next.



Click Finish để kết thúc.



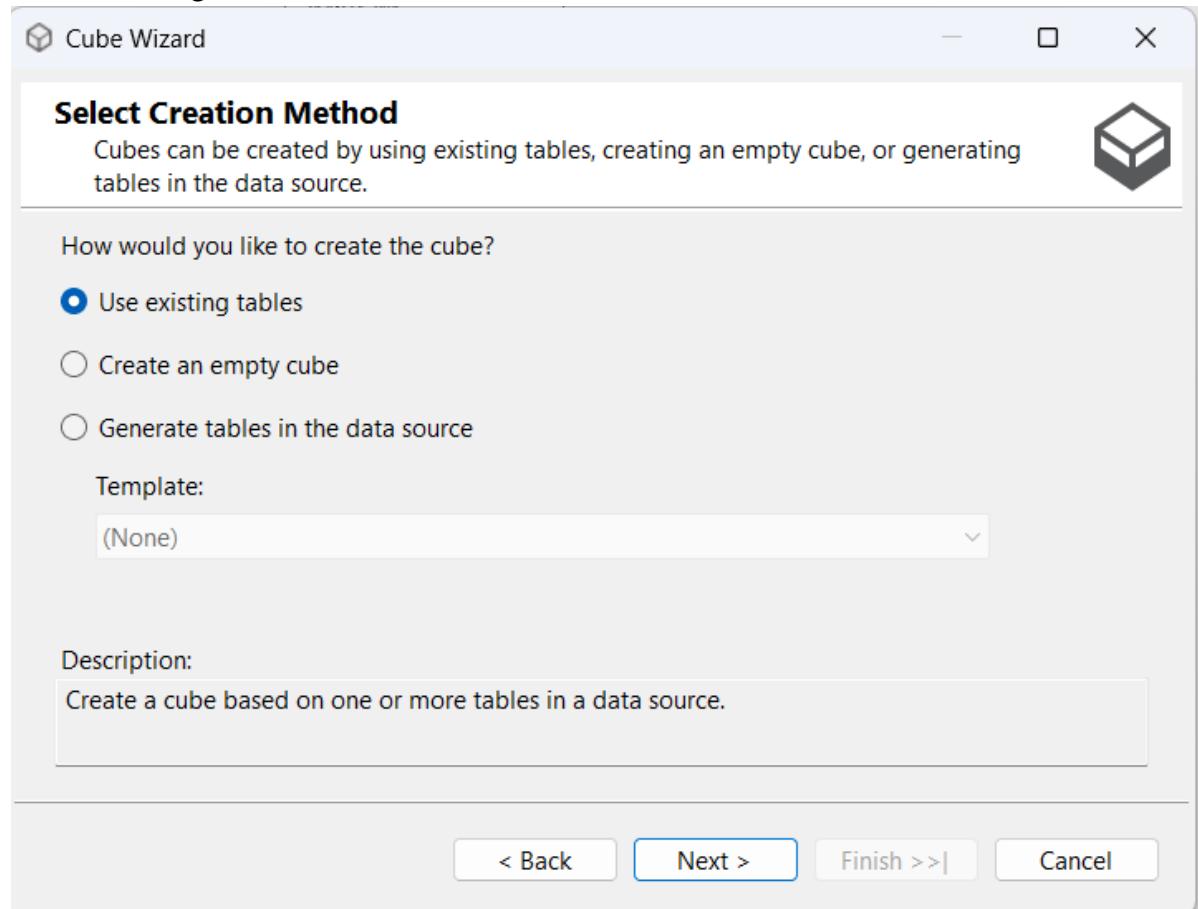
Double-click vào Store Sales Data Source View vừa mới tạo ta sẽ thấy star schema ở Design surface như sau:



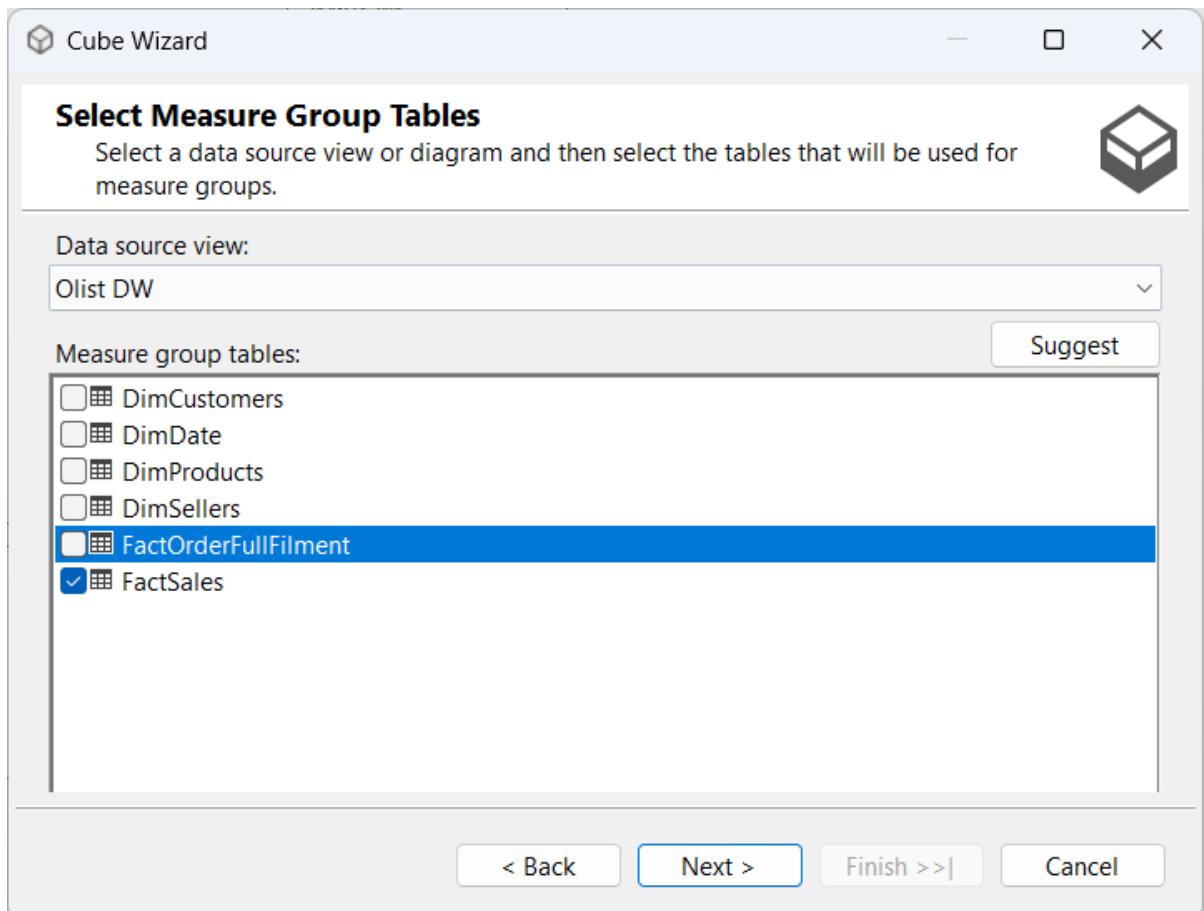
## Tạo data cube

Tiếp theo, ta sẽ tạo một data cube từ data source view.

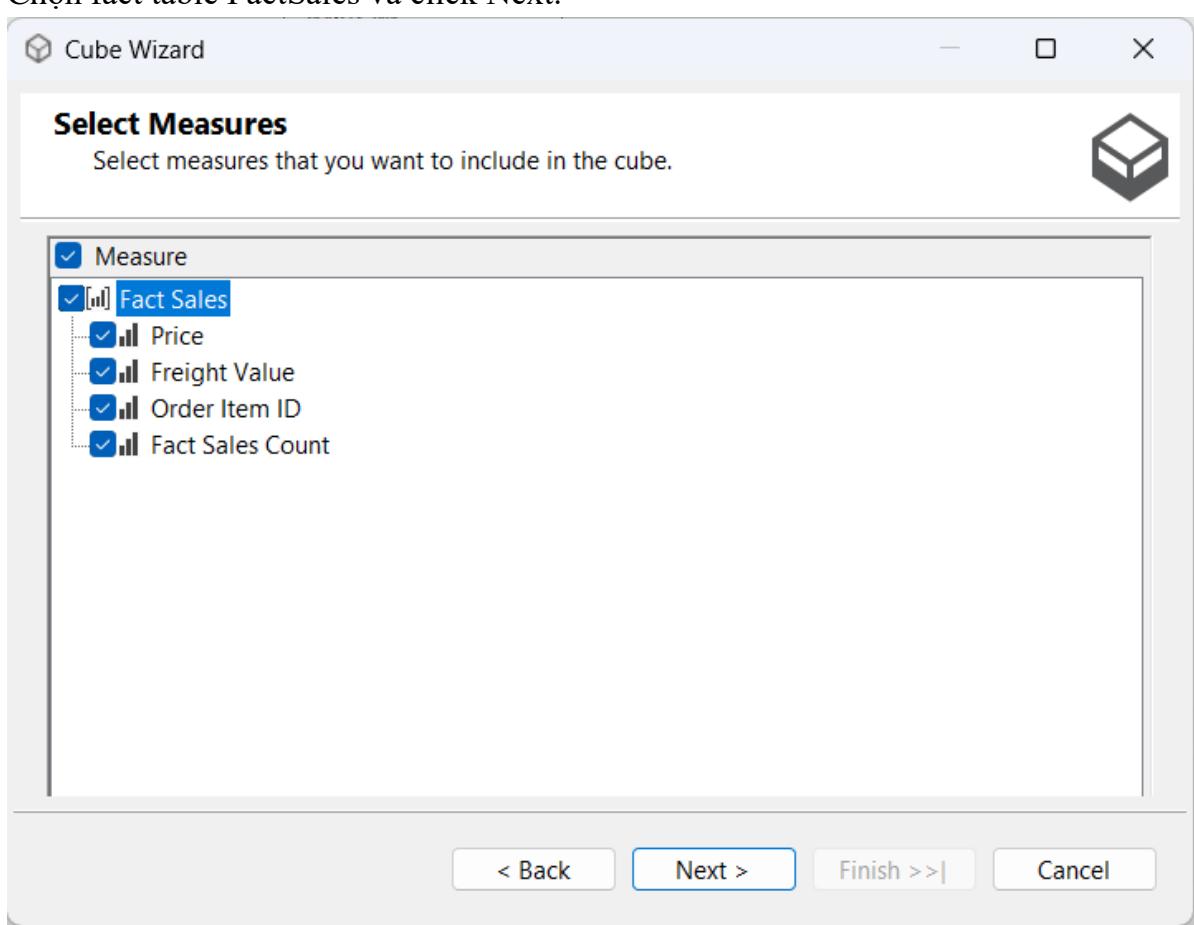
Right-click vào Cubes từ Solution Explorer và chọn New Cube... sẽ xuất hiện Cube Wizard dialog. Click Next.



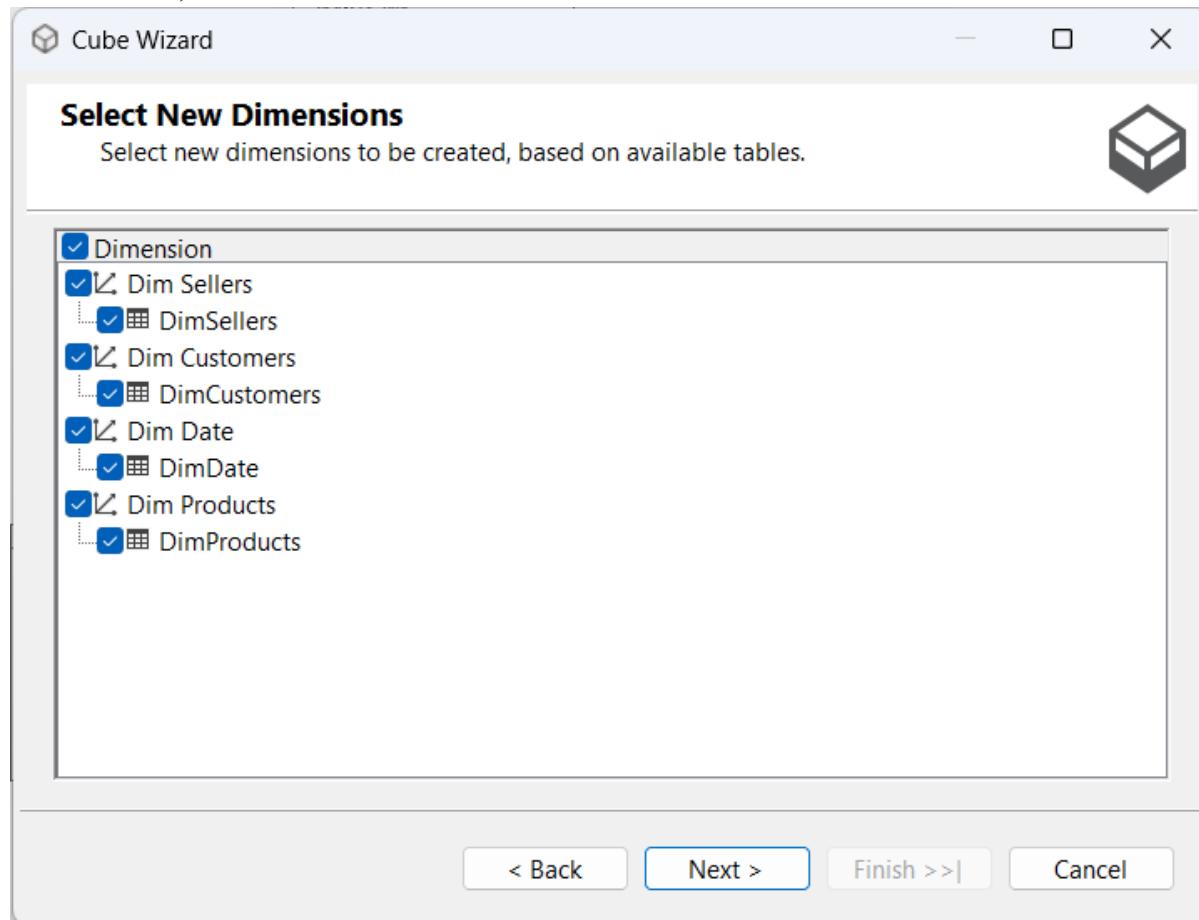
Click Next



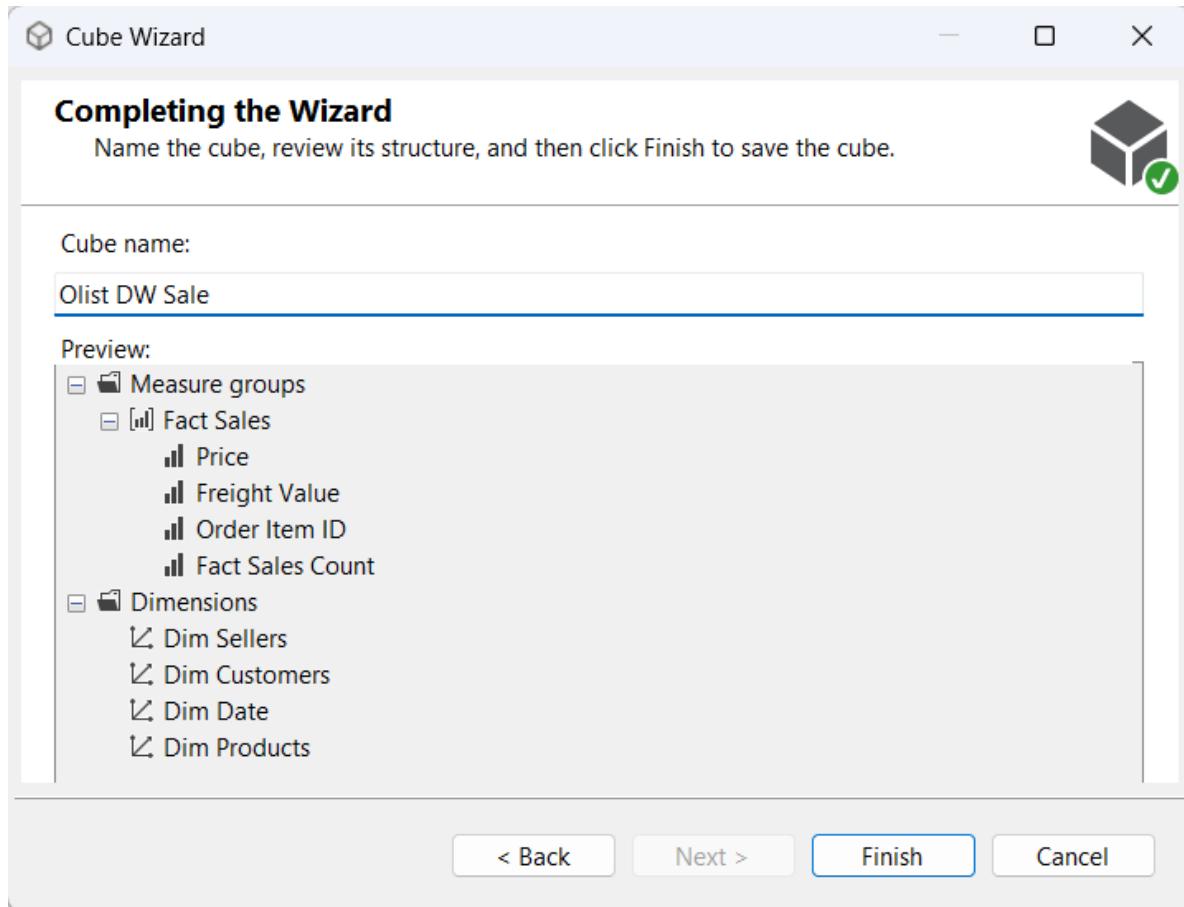
Chọn fact table FactSales và click Next.



Chọn (tất cả) các measures từ fact table để đưa vào cube và click Next.

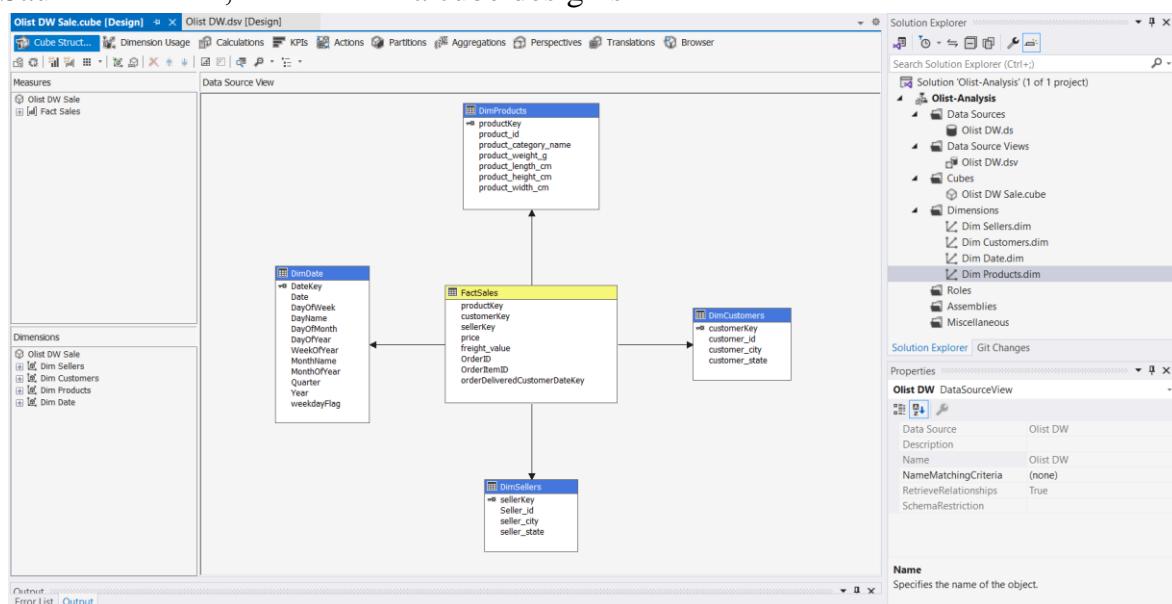


Chọn tất cả các dimension và click Next.



Đặt tên cho cube và click Finish để kết thúc.

Sau khi kết thúc, màn hình của cube design sẽ như bên dưới



## Cấu hình các dimension

Mặc dù Wizard ở bước trước đã tạo các dimension, ta cần chỉnh sửa lại chúng để có thể thay đổi các thuộc tính (attributes), các cây phân cấp (hierarchies), và các quan hệ (relationships).

Ta sẽ bắt đầu với DimSellers Dimension.

Từ Solution Explorer, double-click vào DimeSellers.dim, màn hình Design cho DimeSellers.dim sẽ hiển thị.

Attributes	Hierarchies	Data Source View		
<ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Dim Sellers           <ul style="list-style-type: none"> <li><input type="checkbox"/> Seller City</li> <li><input type="checkbox"/> Seller Key</li> <li><input type="checkbox"/> Seller State</li> </ul> </li> </ul>	<p>Hierarchies</p> <table border="1" style="margin-bottom: 10px;"> <tr> <td style="padding: 2px;">Hierarchy <input checked="" type="checkbox"/></td> <td style="padding: 2px;">&lt;new level&gt;</td> </tr> </table> <p>To create a new hierarchy, drag an attribute here.</p>	Hierarchy <input checked="" type="checkbox"/>	<new level>	<p>DimSellers</p> <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> DimSellers           <ul style="list-style-type: none"> <li><input type="checkbox"/> sellerKey</li> <li><input type="checkbox"/> Seller_id</li> <li><input type="checkbox"/> seller_city</li> <li><input type="checkbox"/> seller_state</li> </ul> </li> </ul>
Hierarchy <input checked="" type="checkbox"/>	<new level>			

Kéo rê (drag and drop) các cột seller\_state và seller\_city từ mục Data Source View vào mục Attributes.

Attributes	Hierarchies	Data Source View		
<ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Dim Sellers           <ul style="list-style-type: none"> <li><input type="checkbox"/> Seller City</li> <li><input type="checkbox"/> Seller Key</li> <li><input type="checkbox"/> Seller State</li> </ul> </li> </ul>	<p>Hierarchies</p> <table border="1" style="margin-bottom: 10px;"> <tr> <td style="padding: 2px;">Hierarchy <input checked="" type="checkbox"/></td> <td style="padding: 2px;">&lt;new level&gt;</td> </tr> </table> <p>To create a new hierarchy, drag an attribute here.</p>	Hierarchy <input checked="" type="checkbox"/>	<new level>	<p>DimSellers</p> <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> DimSellers           <ul style="list-style-type: none"> <li><input type="checkbox"/> sellerKey</li> <li><input type="checkbox"/> Seller_id</li> <li><input type="checkbox"/> seller_city</li> <li><input type="checkbox"/> seller_state</li> </ul> </li> </ul>
Hierarchy <input checked="" type="checkbox"/>	<new level>			

Tiếp theo, kéo rê Seller State từ mục Attributes vào mục the Hierarchies.

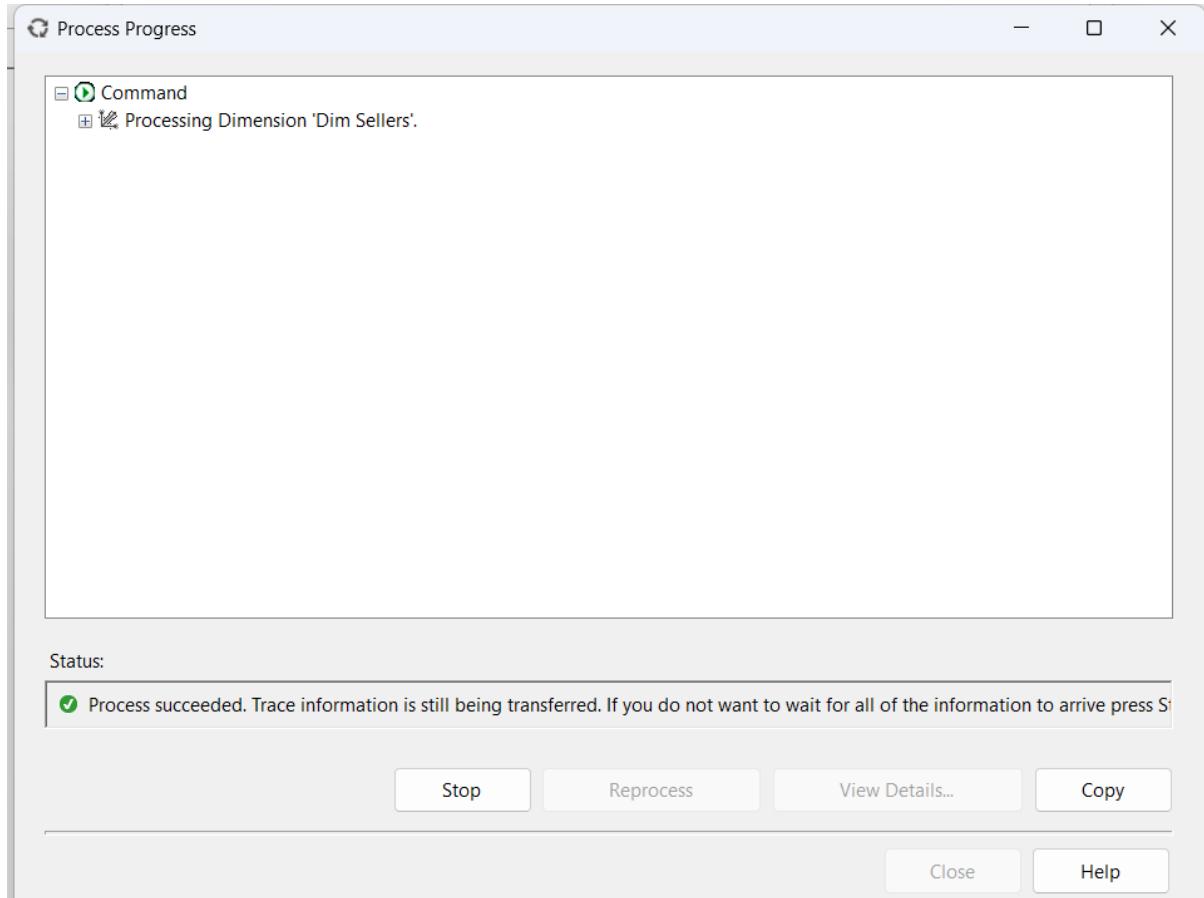
Đổi tên của Hierarchy thành SellHierarchy (right-click vào Hierarchy và click Rename).

Attributes	Hierarchies	Data Source View		
<ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> Dim Sellers           <ul style="list-style-type: none"> <li><input type="checkbox"/> Seller City</li> <li><input type="checkbox"/> Seller Key</li> <li><input type="checkbox"/> Seller State</li> </ul> </li> </ul>	<p>Hierarchies</p> <table border="1" style="margin-bottom: 10px;"> <tr> <td style="padding: 2px;">SellHierarchy <input checked="" type="checkbox"/></td> <td style="padding: 2px;">&lt;new level&gt;</td> </tr> </table> <p>To create a new hierarchy, drag an attribute here.</p>	SellHierarchy <input checked="" type="checkbox"/>	<new level>	<p>DimSellers</p> <ul style="list-style-type: none"> <li><input checked="" type="checkbox"/> DimSellers           <ul style="list-style-type: none"> <li><input type="checkbox"/> sellerKey</li> <li><input type="checkbox"/> Seller_id</li> <li><input type="checkbox"/> seller_city</li> <li><input type="checkbox"/> seller_state</li> </ul> </li> </ul>
SellHierarchy <input checked="" type="checkbox"/>	<new level>			

Cuối cùng, kéo rê the Seller City vào hierarchy.

Click Process icon ( ) ở toolbar click Yes khi được hỏi để build và deploy project. Nếu build thành công, dialog sau sẽ xuất hiện

Click Run để thực thi. Nếu thành công, Process Progress dialog sau sẽ xuất hiện.



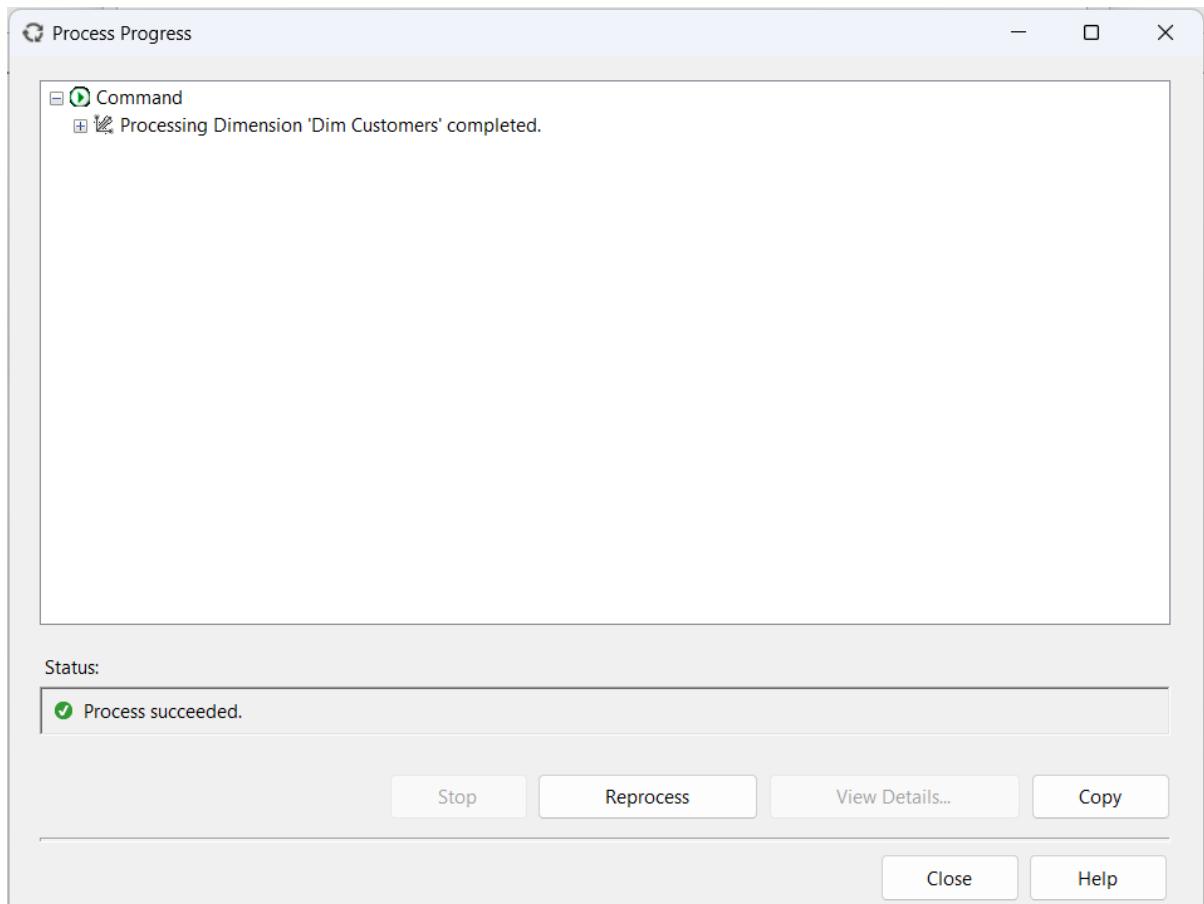
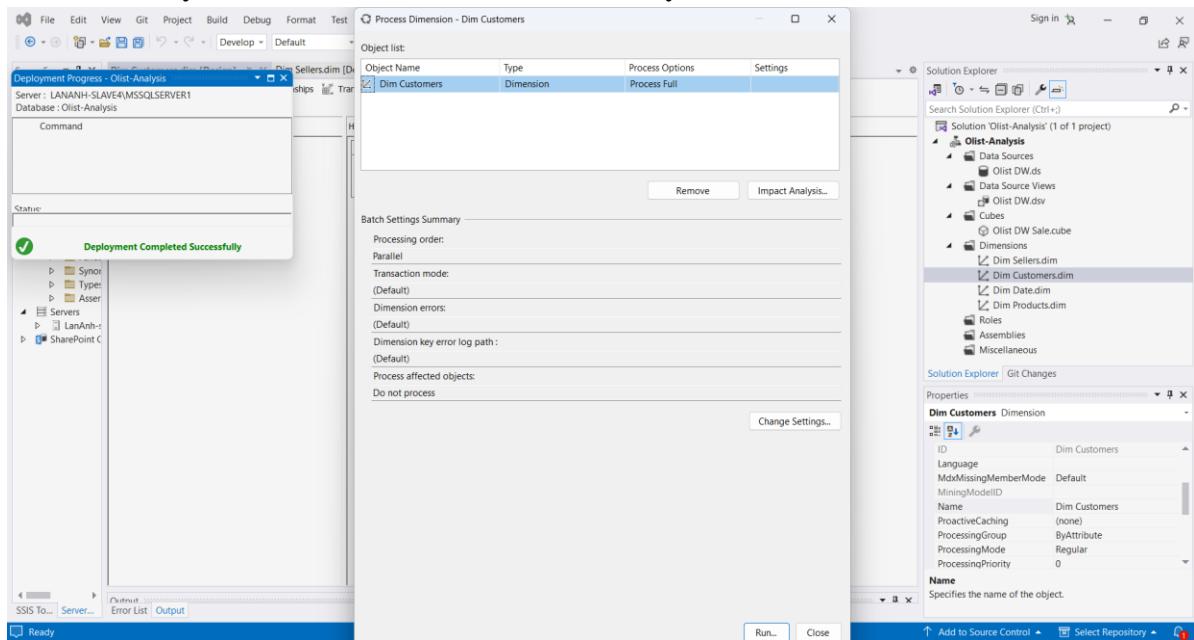
Làm tương tự cho các Dimension khác:

### DimCustomers

The screenshot shows the configuration screen for the 'DimCustomers' dimension, divided into three main sections:

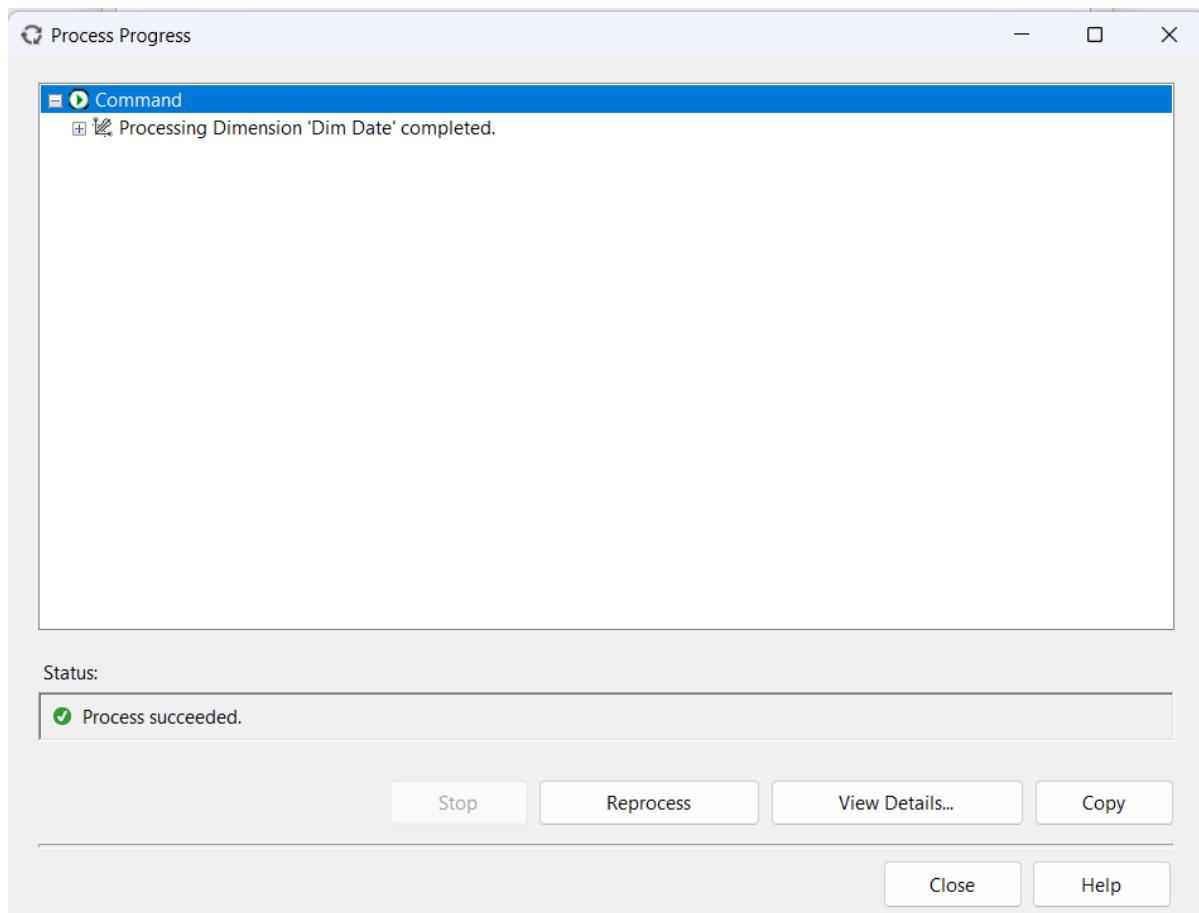
- Attributes:** Shows the 'DimCustomers' table with attributes: Customer City, Customer Key, and Customer State.
- Hierarchies:** Shows a hierarchy named 'CusHierarchy' with levels: Customer State, Customer City, and <new level>. A tooltip indicates: To create a new hierarchy, drag an attribute here.
- Data Source View:** Shows the 'DimCustomers' table with columns: customerKey, customer\_id, customer\_city, and customer\_state.

## CusHierarchy: Customer State -> Customer City



**DimDate**

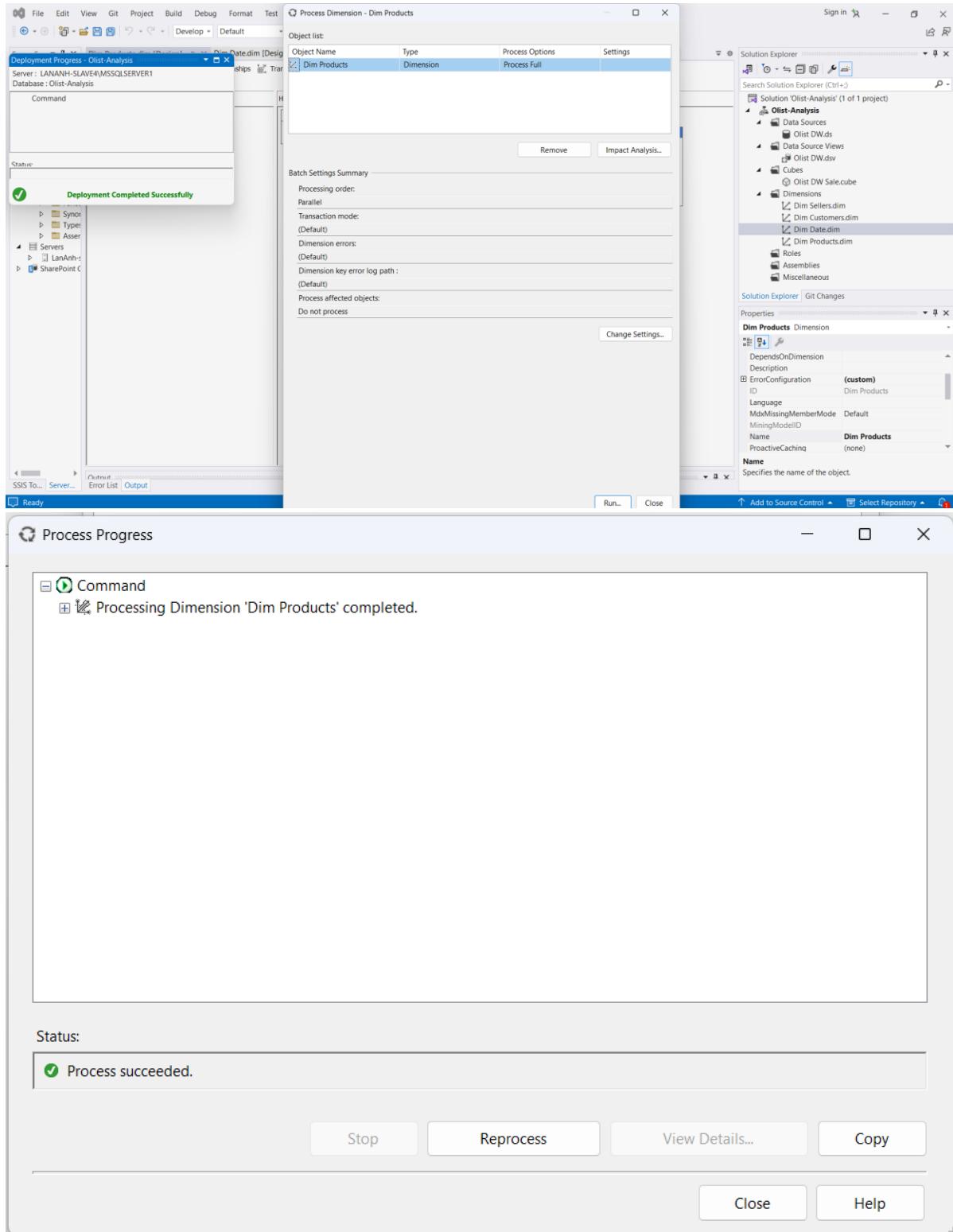
DateHierarchy: Year -> Quater -> Month Of Year -> Date



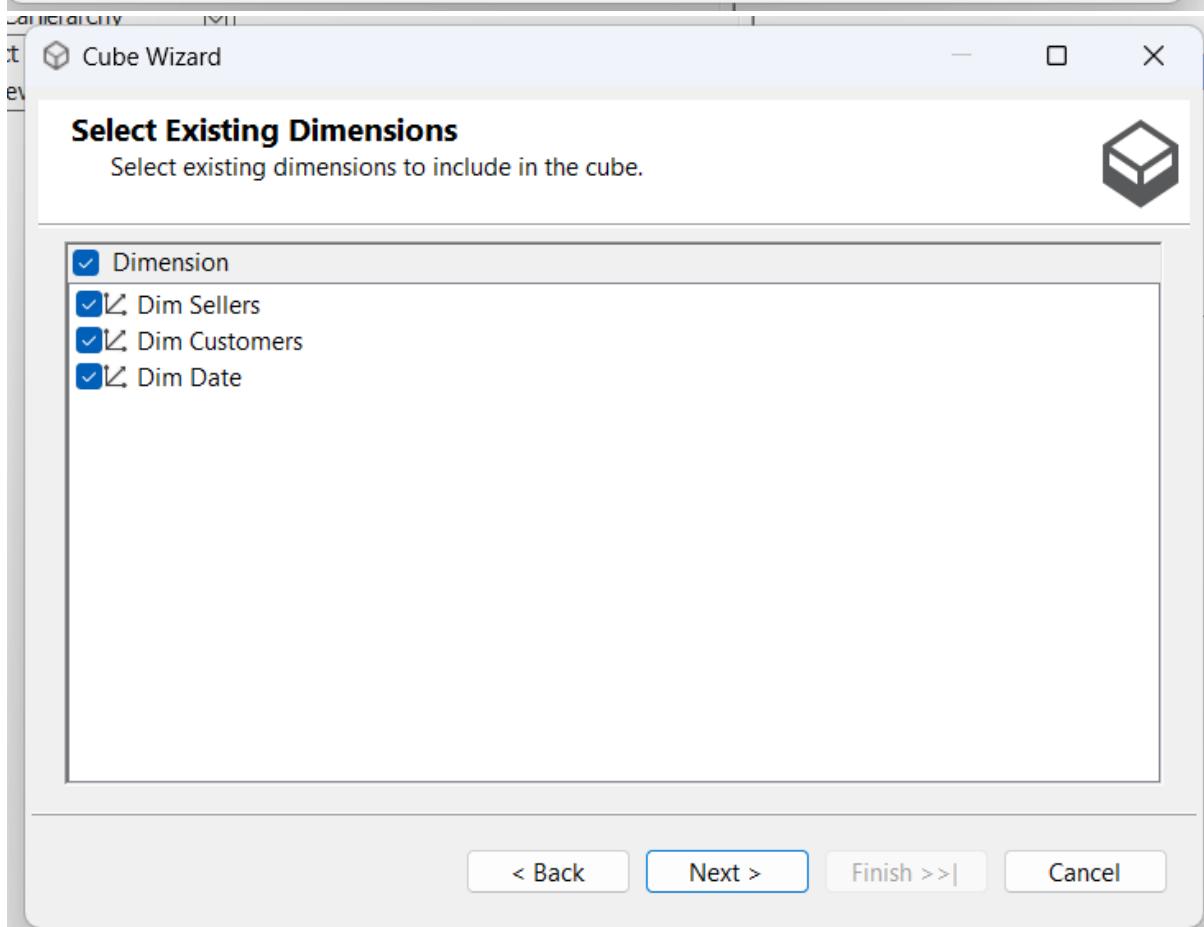
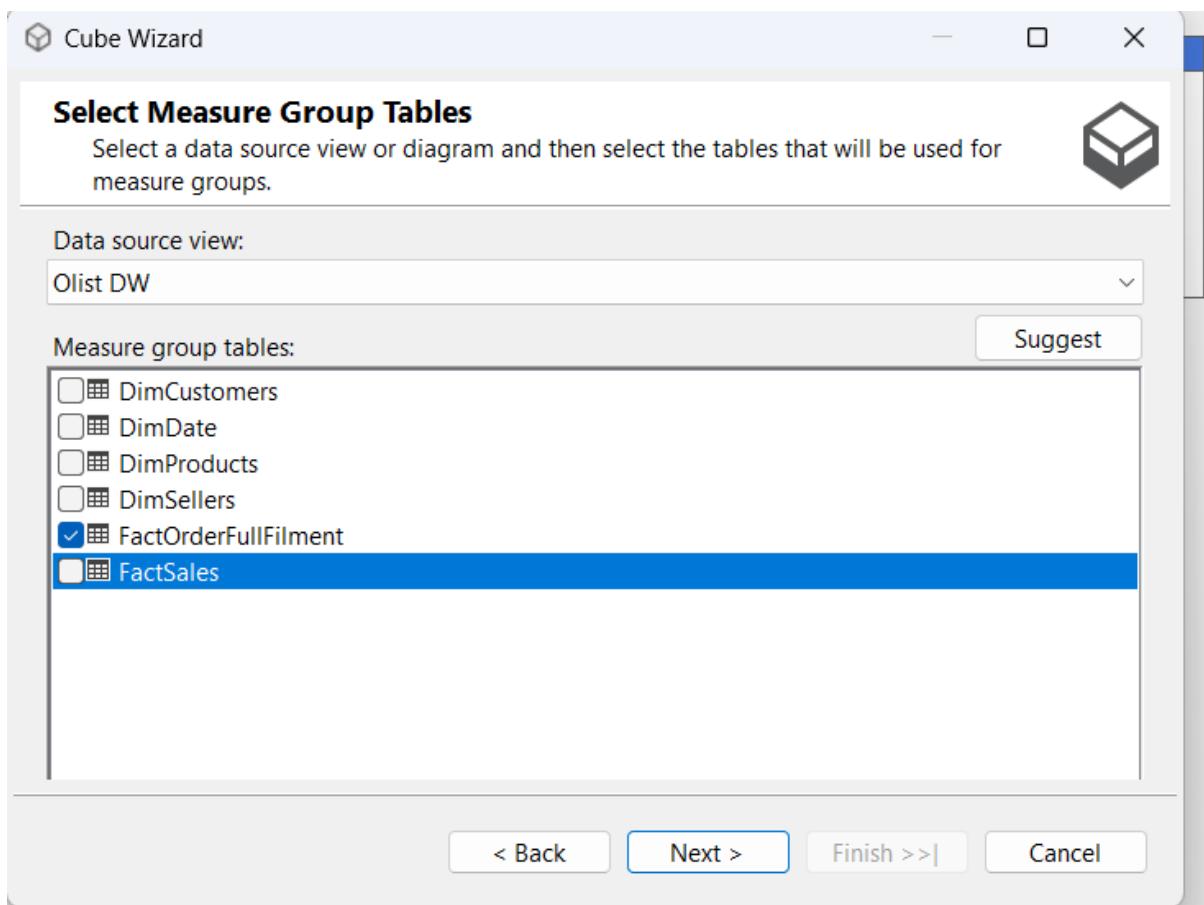
## DimProducts

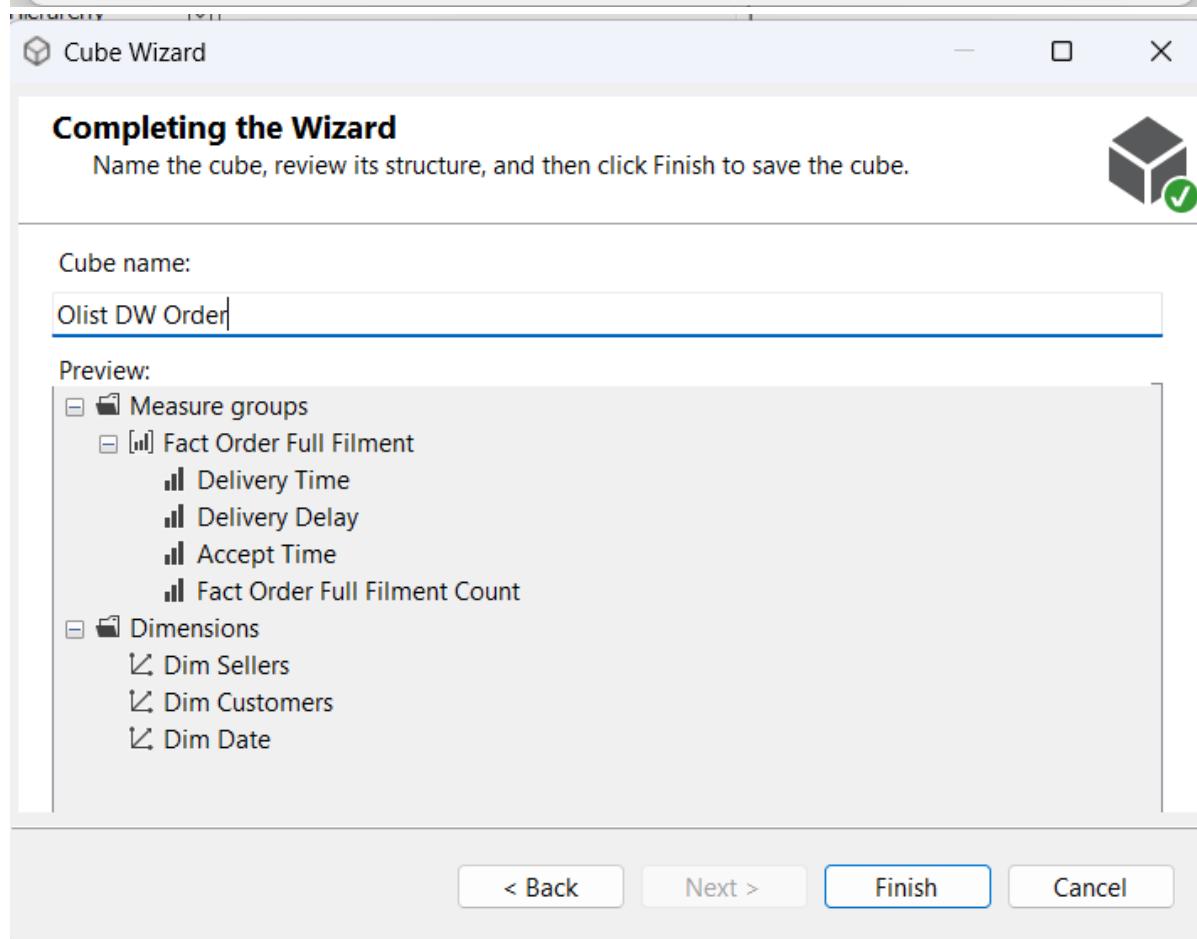
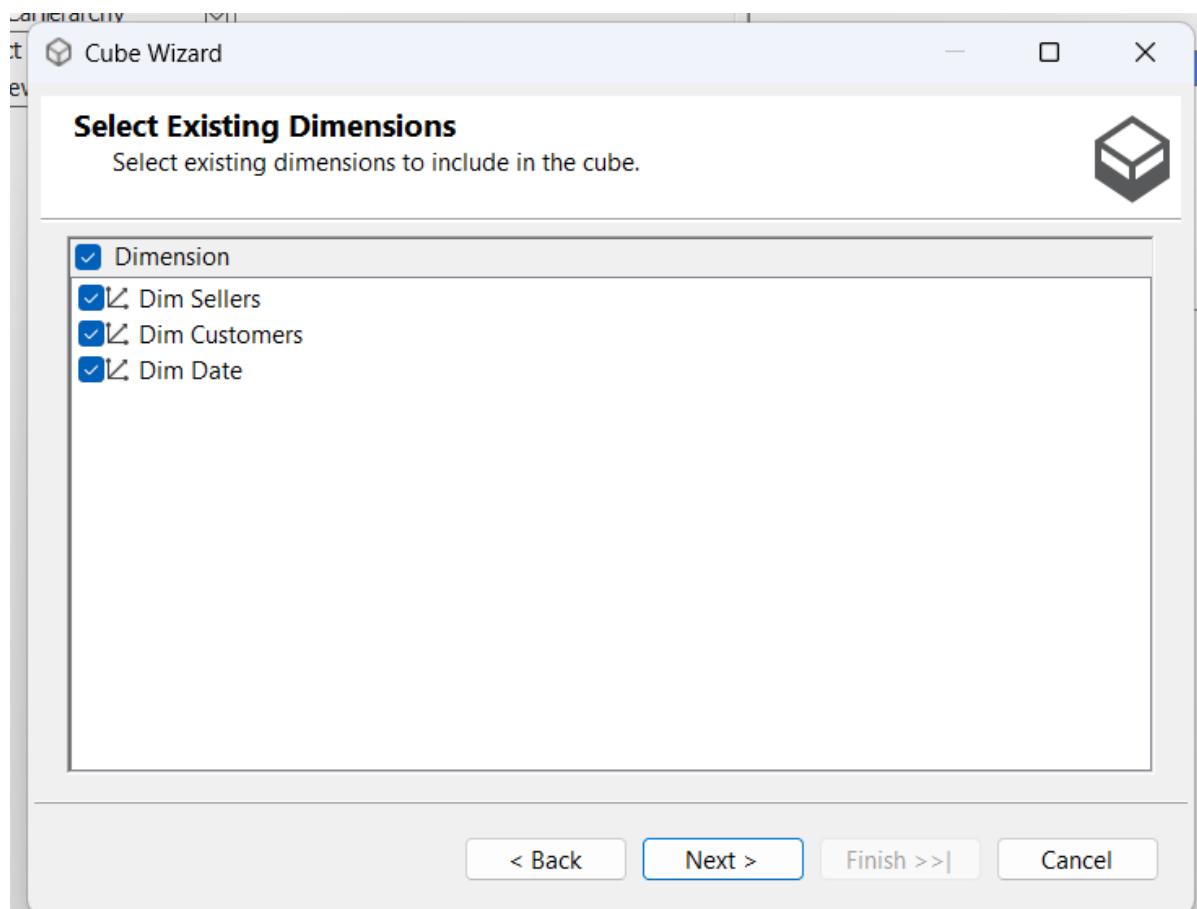
The screenshot shows the 'Dim Products.dim [Design]' tab in the Dimension Structure editor. The interface includes tabs for 'Attributes', 'Hierarchies', and 'Data Source View'. The 'Attributes' tab lists 'Dim.Products' with attributes 'Product Category Name' and 'Product Key'. The 'Hierarchies' tab shows a hierarchy named 'CaHierarchy' with a single item 'Product Category Name'. A note says 'To create a new hierarchy, drag an attribute here.' The 'Data Source View' tab displays the 'DimProducts' table with columns: productKey, product\_id, product\_category\_name, product\_weight\_g, product\_length\_cm, product\_height\_cm, and product\_width\_cm.

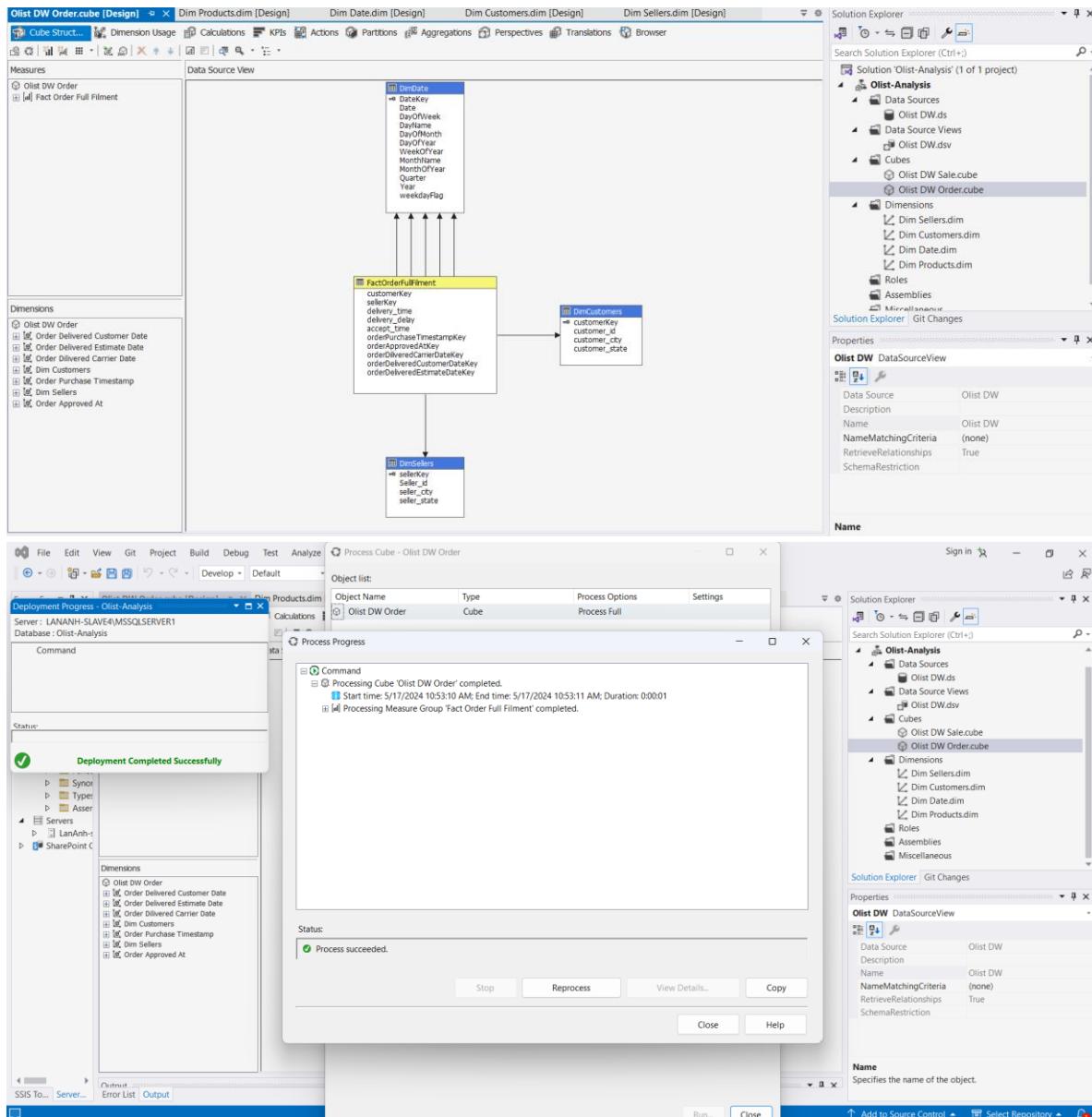
CaHierarchy: Product Category Name



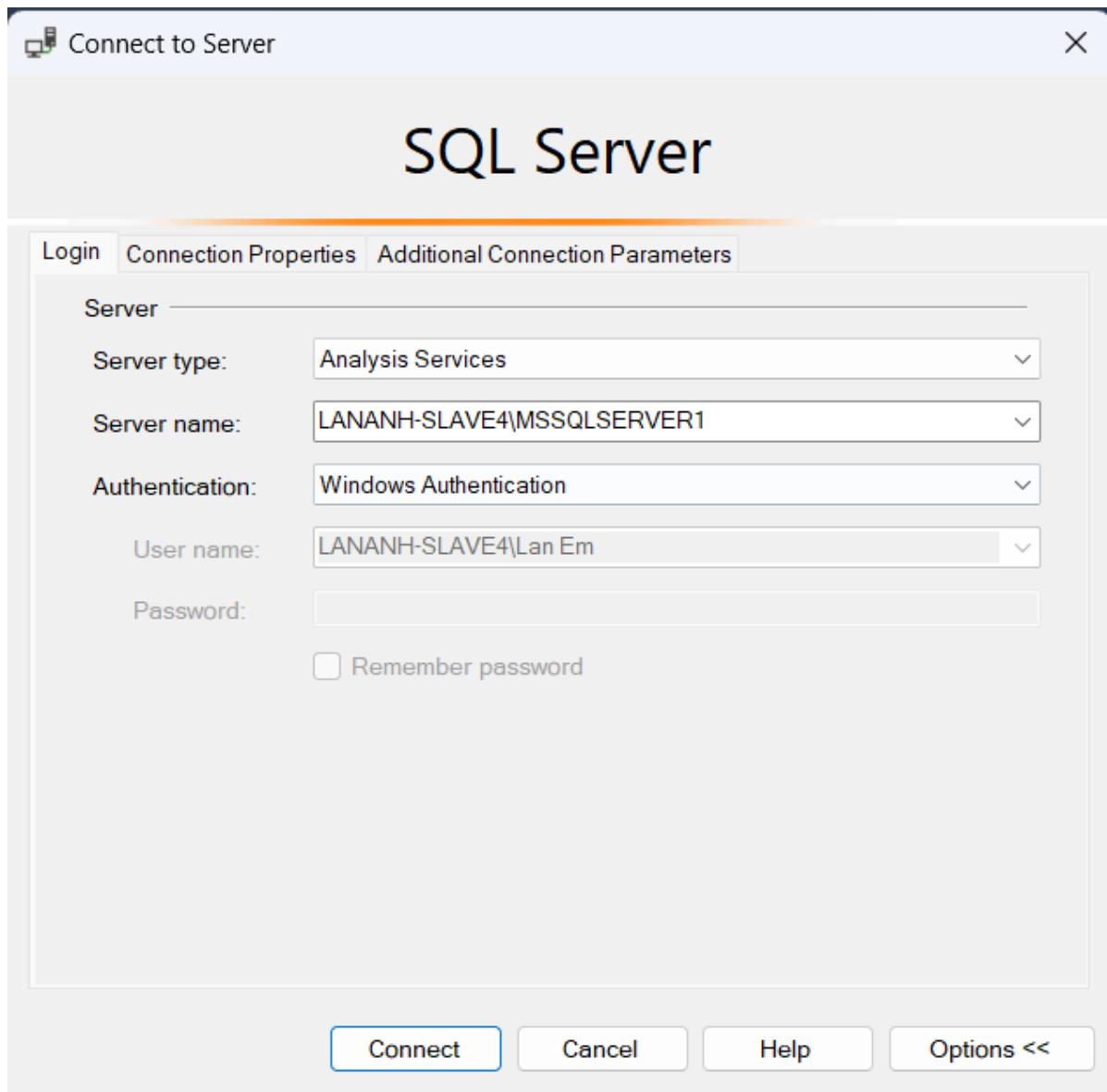
Tương tự với Data Cube Olist DW Order



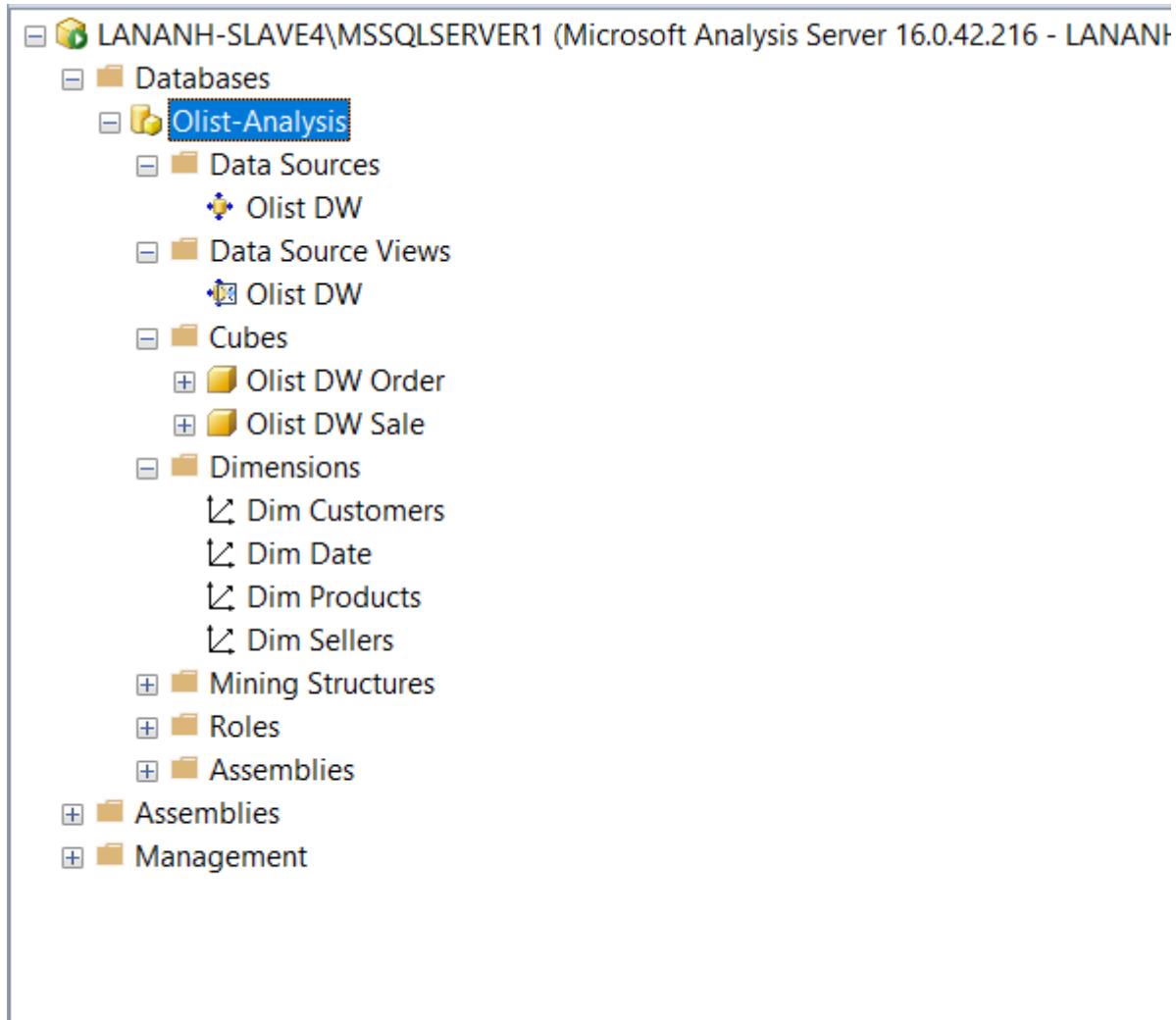




Sau khi tạo thành công 2 Data Cube, truy cập vào SSMS để kiểm tra



Hiển thị thành công 2 Data cube là Olist DW Sale và Olist DW Order trên SSMS với type là Analysis Services



Click chuột phải vào Olist DW Order và chọn Browse -> hiển thị chi tiết Bảng Fact và các bảng Dim liên quan

Phân tích dữ liệu: Phân tích việc vẫn chuyển đơn hàng theo khu vực và thành phố của khách hàng

Chọn kéo thả các thuộc tính bao gồm Customer State, Customer City, Accept Time, Delivery Delay, Delivery Time, Fact Order Full filment Count.

Nhấn Click to execute the query để hiển thị kết quả dữ liệu phân tích việc vận chuyển đơn hàng theo khu vực và thành phố của khách hàng

Customer State	Customer City	Accept Time	Delivery Delay	Delivery Time	Fact Order Full Filment Count
AC	brasileia	12	33	90	3
AC	cruzeiro do sul	10	100	71	3
AC	epitaciolandia	1	31	13	1
AC	manoel urbano	1	34	12	1
AC	porto acre	5	14	29	1
AC	rio branco	243	1472	1523	74
AC	senador guiomard	11	119	41	3
AC	xapuri	14	50	49	2
AL	agua branca	44	2	74	2
AL	anadia	3	42	28	2
AL	arapiraca	131	289	815	33
AL	atalaia	5	22	11	1
AL	barra de santo antonio	6	12	40	2
AL	barra de sao miguel	14	40	48	2
AL	batalha	6	55	48	3
AL	belem	5	27	38	2
AL	boca da mata	5	70	39	2
AL	cajuero	3	5	27	1
AL	campo alegre	12	-3	76	2
AL	canapi	8	7	39	1
AL	coite do noia	2	27	12	1
AL	colonia leopoldina	9	60	46	3
AL	coruripe	14	47	66	3
AL	delmiro gouveia	10	80	121	6

Tương tự cho Data Cube Fact Sale: Phân tích doanh thu theo loại sản phẩm.

Lựa chọn các thuộc tính bao gồm: Product Category Name, Price, Freight Value, Fact Sales Count và nhấn Click to execute the query để hiển thị kết quả dữ liệu Doanh thu theo loại sản phẩm.

Dimension Hierarchy Operator Filter Expression Parameters

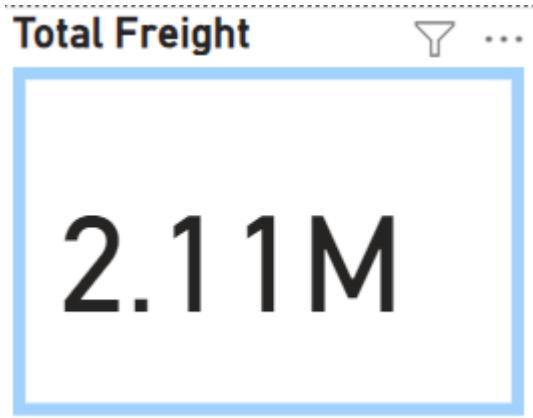
<Select dimension>

**Product Category Name**

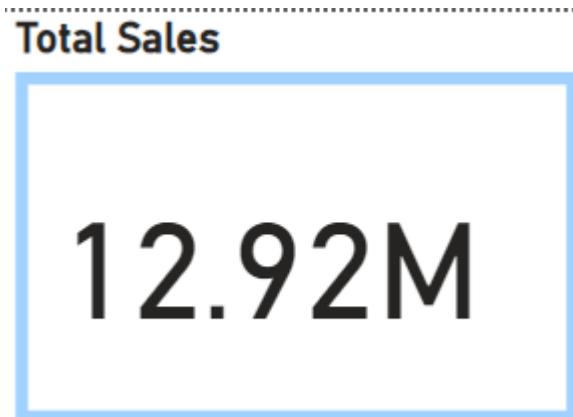
Product Category Name	Price	Freight Value	Fact Sales Count
agro_industria_e_comercio	70508	5545	206
alimentos	28191	6742	496
alimentos_bebedas	14801	4263	269
artes	23523	3744	197
artes_e_artesanato	1799	359	24
artigos_de_festas	4441	799	42
artigos_de_natal	8638	3115	150
audio	50359	5501	361
automotivo	575322	88344	4130
bebê	397071	64624	2973
bebidas	21380	5256	361
beleza_saude	1224960	173912	9440
brinquedos	463595	73192	3996
cama_mesa_banho	1013942	195752	10926
casa_conforto	57756	8220	429
casa_conforto_2	733	396	30
casa_construção	81090	13309	596
cds_dvds_musicais	730	218	14
cine_foto	6632	1196	70
climatização	53181	6387	289
consoles_games	147219	18503	1085
construção_ferramentas_construção	140882	19656	909
construção_ferramentas_ferramentas	15852	1982	103
construção_ferramentas_iluminação	39630	7291	300

## PHẦN 4: POWER BI

### 4.1. Tổng phí vận chuyển

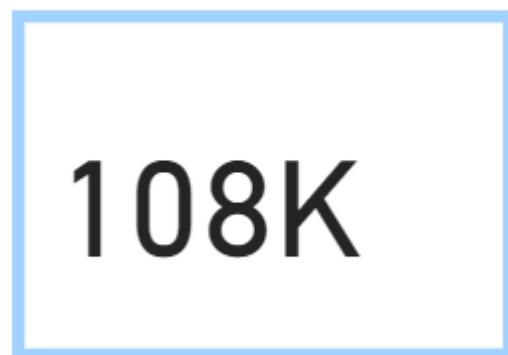


### 4.2. Tổng doanh thu



### 4.3. Tổng đơn hàng

Num of orders



#### 4.4. Tổng người bán

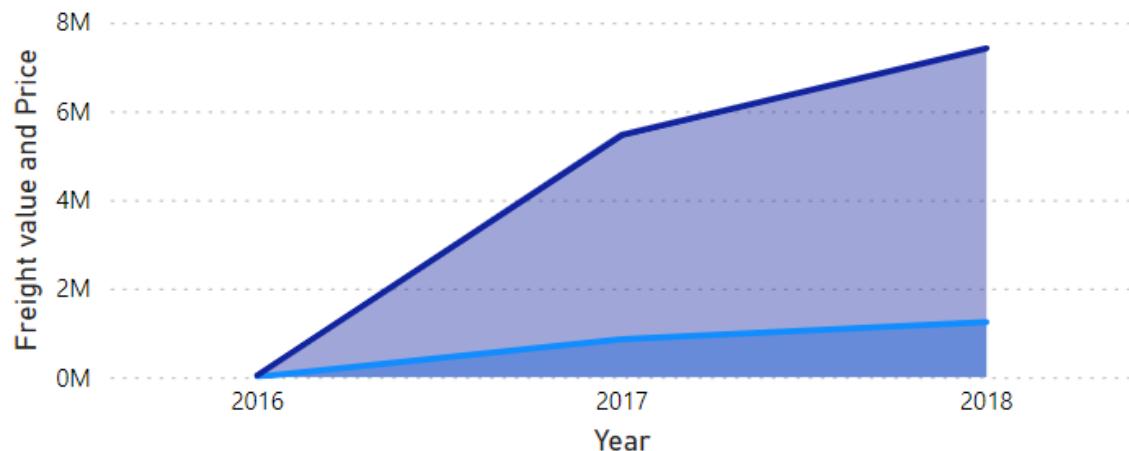
Num of sellers

98K

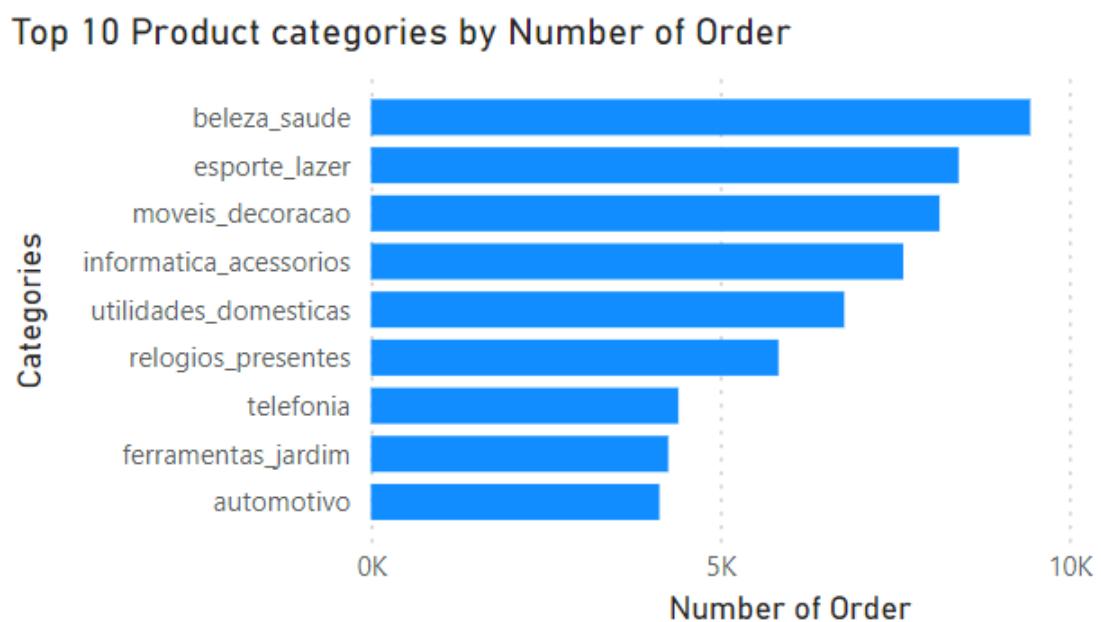
#### 4.5. Tổng chi phí vận chuyển và tổng giá trị đơn hàng theo năm

Sum of Freight Value and Sum of Price by Year

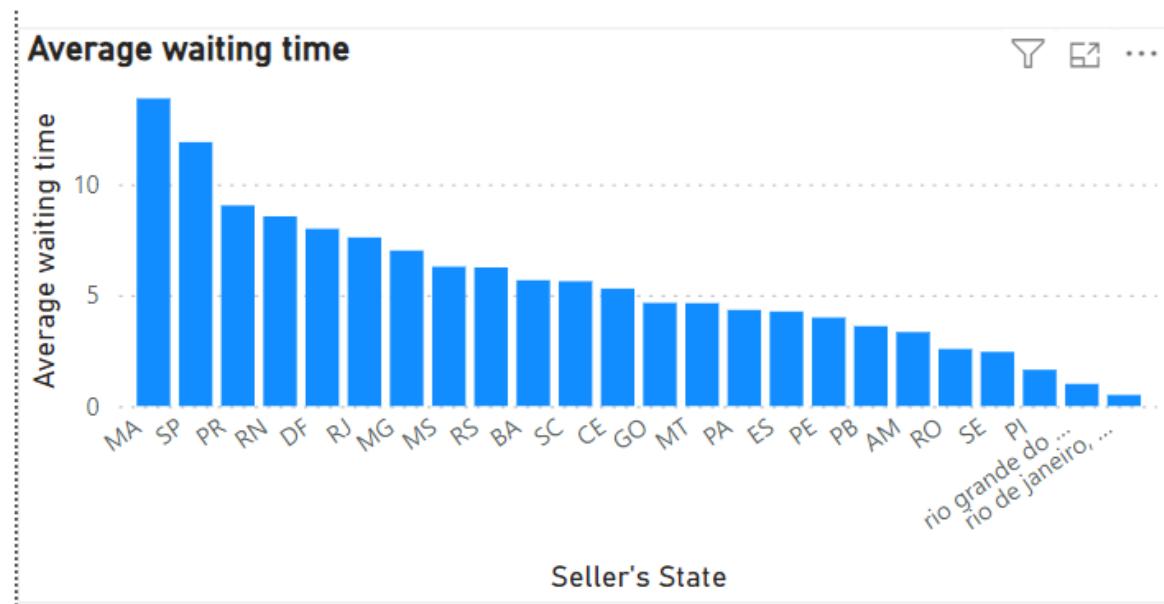
- Sum of Freight Value
- Sum of Price



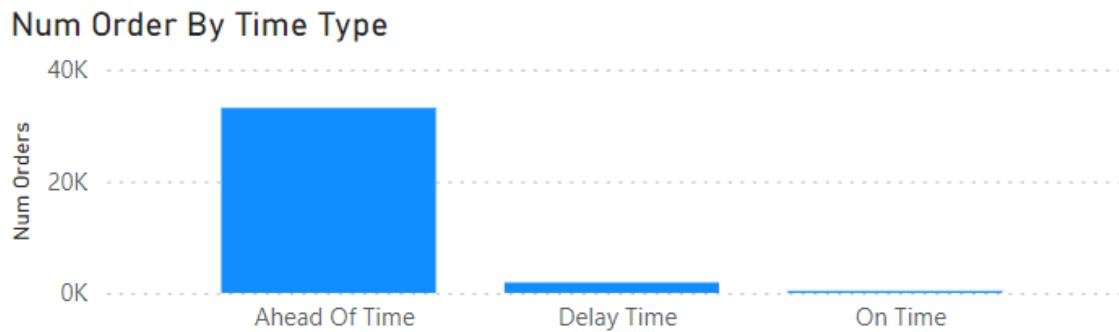
#### 4.6. Top 10 danh mục sản phẩm được mua nhiều nhất



#### 4.7. Thời gian xác nhận đơn hàng trung bình theo từng tiểu bang



#### 4.8. Số lượng đơn hàng theo loại thời gian giao hàng



#### 4.9. Số lượng khách hàng theo năm

