

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



**HCMUTE**

**BÁO CÁO ĐỒ ÁN CUỐI KỲ  
ĐỀ TÀI: XÂY DỰNG KHO DỮ LIỆU BÁN HÀNG  
CHO SIÊU THỊ**

**Môn học: KHO DỮ LIỆU  
Mã lớp học phần: DAWH430784\_23\_2\_01  
GVHD: ThS. Nguyễn Văn Thành**

**Nhóm sinh viên thực hiện: Nhóm 10**

Hoàng Mạnh Đức	21133027
Phạm Hữu Dũng	21133022
Trương Quốc Việt	21133092
Nguyễn Phương Khoa	21133048

**TP. Hồ Chí Minh, tháng 05 năm 2024**

## **DANH SÁCH THÀNH VIÊN THAM GIA THỰC HIỆN ĐỀ TÀI VÀ VIẾT BÁO CÁO**

Môn: Kho dữ liệu - HỌC KÌ II – NĂM HỌC 2023 – 2024

STT	HỌ VÀ TÊN	MSSV	TỶ LỆ ĐÓNG GÓP
1	Hoàng Mạnh Đức	21133027	100%
2	Phạm Hữu Dũng	21133022	100%
3	Trương Quốc Việt	21133092	100%
4	Nguyễn Phuong Khoa	21133048	100%

## Nhận xét của giảng viên:

Ngày ... tháng 05 năm 2024

## *Giảng viên chấm điểm*

*Ths. Nguyễn Văn Thành*

## LỜI CẢM ƠN

Chúng em xin gửi lời cảm ơn sâu sắc đến **thầy Nguyễn Văn Thành**, thầy đã tận tình giảng dạy và hỗ trợ trong suốt quá trình tìm hiểu và học tập môn **Khoa dữ liệu**. Trong quá trình học tập, thầy đã nhiệt huyết trong từng lời giải đáp, những kiến thức, kỹ năng, tài liệu quý giá mà thầy đã truyền đạt học không chỉ là nền tảng cho quá trình thực hiện báo cáo cuối kì mà còn là kiến thức nền tảng cho nhiều phần sau này. Rất nhiều công sức và nỗ lực đã được bỏ ra, nhưng do chưa có kinh nghiệm trong việc xây dựng đề tài và những hạn chế về kiến thức, bài báo cáo này không tránh khỏi sai sót. Chúng em luôn sẵn sàng tiếp thu những ý kiến từ thầy để cải thiện bài báo cáo cũng như đề tài lần này. Một lần nữa, nhóm em xin bày tỏ lòng biết ơn đến thầy. Xin kính chúc thầy luôn dồi dào sức khỏe, đạt được nhiều thành công trong công việc.

## MỤC LỤC

CHƯƠNG 1: TỔNG QUAN VỀ ĐỀ TÀI .....	1
1.1. Lý do chọn đề tài.....	1
1.2. Tổng quan về tập dữ liệu .....	1
1.2.1. Nguồn dữ liệu .....	1
1.2.2. Mô tả chi tiết tập dữ liệu .....	2
Tập Sample_SuperStore.csv .....	2
1.3. Giới thiệu các công cụ được sử dụng.....	2
CHƯƠNG 2: THIẾT KẾ XÂY DỰNG CƠ SỞ DỮ LIỆU TÁC NGHIỆP(OLAP) ....	3
2.1. Xác định các Business Process và bảng Fact.....	3
2.1.1. Business Process: Sales Analysis.....	3
2.1.2. Business Process: OrderFulfillment.....	3
2.2. Xây dựng các bảng Dimension .....	4
2.2.1. DimDate .....	4
2.2.2. DimProduct .....	4
2.2.3. DimCustomer .....	4
2.2.4. DimShipDate.....	4
2.2.5. DimOrder .....	4
2.3. Star Schema (Lược đồ hình sao).....	5
CHƯƠNG 3: TÍCH HỢP DỮ LIỆU VÀO KHO (SSIS).....	6
3.1. Import dữ liệu gốc vào database ExternalSources bằng SSIS .....	6
3.2. Import dữ liệu vào các bảng dimension: .....	13
3.2.1.OderDate Dimension .....	15
a) Load từ nguồn vào bảng stgDate .....	15
b) Load dữ liệu từ stgDate vào DimDate .....	20
3.2.2.Product Dimension.....	24
a) Load từ nguồn vào bảng stgProduct .....	24
b) Load dữ liệu từ stgProduct vào DimProduct .....	31
3.2.3.Customer Dimension .....	36
a) Load dữ liệu từ nguồn vào stgCustomer.....	36
b) Load dữ liệu từ stgCustomer vào bảng DimCustomer .....	43
3.2.4.ShipDate Dimension .....	48
a) Load từ nguồn vào bảng stgOrder .....	48
b) Load dữ liệu từ stgDate vào DimDate .....	54
3.2.5. Order Dimension.....	58
a) Load từ nguồn vào bảng stgOrder .....	58
b) Load dữ liệu từ stgDate vào DimDate .....	64
3.2.6. Package Stage_Dim .....	67
a) Setup Execute SQL Task .....	68
3.2.7. Package Load_Dim.....	69
a) Setup Execute SQL Task .....	70
3.3. Import dữ liệu vào bảng các bảng fact.....	72
3.3.1. Fact Sales .....	72
a) Load dữ liệu từ nguồn vào stgSales.....	72

b) Load dữ liệu vào Sales_fact.....	78
3.3.2. Fact OrderFulfillment .....	92
a) Load dữ liệu từ nguồn vào stgOrderFulfillment.....	92
b) Load dữ liệu vào Fact OrderFulfillment.....	99
3.3.3. Package Stage_Fact .....	112
a) Setup Execute SQL Task .....	113
3.3.4. Package Load_Fact .....	114
a) Setup Execute SQL Task .....	115
CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS).....	120
4.1. Quá trình xây dựng mô hình .....	120
4.1.1. Tạo Data Source.....	120
4.1.2. Tạo Data Source View .....	122
4.2. Quá trình xây dựng khối Cube.....	124
4.3. Cấu hình Hierarchy .....	126
4.3.1. Tạo Hierarchy cho Dim ShipDate.....	126
4.3.2. Tạo Hierarchy cho Dim Date .....	127
4.3.3. Tạo Hierarchy cho Dim Order .....	129
4.3.4. Tạo Hierarchy cho Dim Customer.....	131
4.3.5. Tạo Hierarchy cho Dim Product .....	133
4.4. Thực hiện phân tích dữ liệu .....	134
4.4.1. Câu hỏi: Số lượng các loại sản phẩm bán ra trong tháng/ quý/ năm. .....	134
4.4.1.1. Sử dụng công cụ SSAS .....	134
4.4.1.2. Sử dụng Pivot Table trong Excel .....	135
4.4.1.3. Sử dụng Power BI.....	136
4.4.2. Câu hỏi: Cho biết doanh số các sản phẩm trong tháng/ quý/ năm... .....	137
4.4.2.1. Sử dụng công cụ SSAS .....	137
4.4.2.2. Sử dụng Pivot Table trong Excel .....	137
4.4.2.3. Sử dụng Power BI.....	138
4.4.3. Câu hỏi: Những danh mục sản phẩm và danh mục phụ nào đóng góp nhiều nhất vào doanh thu bán hàng theo tháng/ quý/ năm? .....	139
4.4.3.1. Sử dụng công cụ SSAS .....	139
4.4.3.2. Sử dụng Pivot Table trong Excel .....	140
4.4.3.3. Sử dụng Power BI.....	141
4.4.4. Câu hỏi: Phương thức vận chuyển của từng phân khúc khách hàng khác nhau theo tháng quý năm? .....	142
4.4.4.1. Sử dụng công cụ SSAS .....	142
4.4.4.2. Sử dụng Pivot Table trong Excel .....	143
4.4.4.3. Sử dụng Power BI.....	144
4.4.5. Câu hỏi: Tháng có số lượng đơn hàng trong từng loại sản phẩm tăng cao theo từng phương thức vận chuyển. ....	145
4.4.5.1. Sử dụng công cụ SSAS .....	145
4.4.5.2. Sử dụng Pivot Table trong Excel .....	146
4.4.5.3. Sử dụng Power BI.....	146

CHƯƠNG 5: KẾT LUẬN .....	148
5.1. Kết quả đạt được .....	148
5.2. Những hạn chế .....	148
5.3. Tài liệu tham khảo .....	148

# PHÂN CÔNG NHIỆM VỤ

Nhiệm vụ	Quốc Việt	Hữu Dũng	Mạnh Đức	Phương Khoa
Tìm kiếm tập dữ liệu	X	X	X	X
<b>Mô hình hóa dữ liệu vào workbook</b>	X	X	X	X
+ DimDate				X
+ DimProduct				
+ DimCustomer		X		
+ DimShipDate				
+ DimOrder				
+ OrderFullfilmentFact			X	
+ SalesFact	X			
<b>Thực hiện ETL đưa dữ liệu vào các bảng dim và fact dùng SSIS</b>	X	X	X	X
+ Import dữ liệu gốc vào database bằng SSIS	X			
+ Tạo và đưa dữ liệu vào các StagesDim và các Dim			X	X
+ Tạo và đưa dữ liệu vào FactSales stage và FactSales	X			
+ Tạo và đưa dữ liệu vào FactOrderFullfilment stage và FactOrderFullfilment		X		
<b>Nhập dữ liệu vào SSAS, tạo data source, data source view, cube và hierarchy</b>			X	
Tạo các phân cấp hierarchy cho các bảng dim và các attribute relationships.	X	X	X	X
+ Hierarchy cho	X			

DimShipDate				
+ Hierarchy cho DimDate			x	
+ Hierarchy cho DimOrder		x		
+ Hierarchy cho DimCustomer, Product				x

Nhiệm vụ	Quốc Việt	Hữu Dũng	Mạnh Đức	Phương Khoa
<b>Thực hiện trả lời các câu hỏi đã đặt ra cho business process: Sales Analysis</b>	x	x	x	x
+ Số lượng các loại sản phẩm bán ra trong tháng/ quý/ năm.	x			
+ Cho biết doanh số các sản phẩm trong tháng/ quý/ năm.			x	
+ Câu hỏi: Những danh mục sản phẩm và danh mục phụ nào đóng góp nhiều nhất vào doanh thu bán hàng theo tháng/ quý/ năm?				x
+ Câu hỏi: Phương thức vận chuyển của từng phân khúc khách hàng khác nhau theo tháng quý năm?		x		
Câu hỏi: Tháng có số lượng đơn hàng trong từng loại sản phẩm tăng cao theo từng phương thức vận chuyển.				x
<b>Thực hiện trả lời các câu hỏi đã đặt ra cho business process: Inventory management: quản lý số lượng sản phẩm phân bổ trong các kho</b>	x	x	x	x
+ Trả lời câu hỏi	x			

dùng SSAS				
+ Trả lời câu hỏi dùng Power BI Desktop			x	x
+ Trả lời câu hỏi dùng Excel Pivot		x		

# **CHƯƠNG 1: TỔNG QUAN VỀ ĐỀ TÀI**

## **1.1. Lý do chọn đề tài**

Trong thời đại công nghệ 4.0 ngày nay, Hầu hết mọi nơi, vào bất kỳ thời điểm nào, đều bao gồm sự hiện diện và hoạt động của các công nghệ. Sự tăng cao về nhu cầu sử dụng các thiết bị công nghệ làm cho việc cung cấp, bảo trì và sửa chữa trở nên vô cùng quan trọng. Để đảm bảo rằng các quá trình từ sản xuất, vận chuyển, lưu trữ đến phân phối sản phẩm diễn ra một cách mượt mà và tự động, một hệ thống quản lý toàn diện là không thể thiếu. Thường thường, dữ liệu được thu thập và lưu trữ ở mỗi giai đoạn khác nhau, và việc kết hợp chúng một cách cẩn thận có thể giúp doanh nghiệp ra những quyết định chính xác và hiệu quả hơn. Vì lý do này, nhóm của chúng em đã quyết định nghiên cứu và thực hiện đề tài " Xây dựng kho dữ liệu cho siêu thị". Mục tiêu của chúng em là tạo ra một Kho dữ liệu linh hoạt và dễ quản lý, giúp công ty nắm bắt thông tin một cách hiệu quả và từ đó, đưa ra các quyết định chiến lược và có ý nghĩa .

## **1.2. Tổng quan về tập dữ liệu**

### **1.2.1. Nguồn dữ liệu**

- Nhóm sử dụng tập dữ liệu Sample\_SuperStore.csv lấy từ kaggle
- Đường dẫn tải tập dữ liệu <https://www.kaggle.com/datasets/vivek468/superstore-dataset-final>

## 1.2.2. Mô tả chi tiết tập dữ liệu

### Tập Sample\_SuperStore.csv

Gồm 21 cột và 9994 dòng

Tên cột	Mô tả
Order ID	ID đơn hàng
Order Date	Ngày đặt hàng
Ship Date	Ngày giao hàng
Ship Mode	Kiểu giao hàng
Customer ID	ID khách hàng
Customer Name	Tên khách hàng
Segment	Loại khách hàng
Country	Quốc gia
City	Thành phố
State	Bang
Postal Code	Mã bưu chính
Region	Địa phương
Product ID	ID sản phẩm
Category	Danh mục sản phẩm
Sub-Category	Danh mục sản phẩm phụ
Product Name	Tên sản phẩm
Sales	Doanh thu
Quantity	Số lượng
Discount	Giảm giá
Profit	Lợi nhuận

## 1.3. Giới thiệu các công cụ được sử dụng

- Visual Studio 2019/2022 tích hợp SSIS (SQL Server Integration Services) và SSAS (SQL Server Analysis Services).
- SQL Server Management Studio 2019
- Power BI.

# CHƯƠNG 2: THIẾT KẾ XÂY DỰNG CƠ SỞ DỮ LIỆU TÁC NGHIỆP(OLAP)

## 2.1. Xác định các Business Process và bảng Fact

- Xây dựng Detailed Bus Matrix xác định các Business Process, bảng Fact, bảng Dim cần thiết.

Instructions!		Fact Table	Fact Grain Type	Granularity	Facts					
Business Process Name						Customer	Product	Order	OrderDate	ShipDate
Sales reporting	Sale_fact	Transaction	One row per order detail	ShipDate, CustomerID, ProductID, OrderID, OrderDate, Quantity (số lượng sản phẩm bán được), Discount, Sales, Profit,	x x x x x					
Order Fulfillment	OrderFulfillment_fact	Accumulated	One row for each order	Order ID(mã đơn hàng), Order date/ngày đặt hàng, Shipdate/ngày giao hàng, Customer ID(mã khách hàng)	x x x x					

### 2.1.1. Business Process: Sales Analysis

Các câu hỏi cụ thể được đặt ra:

- Câu hỏi: Số lượng các loại sản phẩm bán ra trong tháng/ quý/ năm.
- Câu hỏi: Cho biết doanh số các sản phẩm trong tháng/ quý/ năm.
- Câu hỏi: Những danh mục sản phẩm và danh mục phụ nào đóng góp nhiều nhất vào doanh thu bán hàng theo tháng/ quý/ năm?

Bảng FactSales

FactSales	ShipDateKey	convert datetime to int or format YYYYMMDD	20041123
	OrderDateKey	Key of each orderdate	1,2,3
	ProductKey	Key of each product	1,2,3
	OrderKey	Key of each order	20041123
	CustomerKey	Surrogate key pipeline lookup for dimension key lookup using business key	1, 2, 5
	Quantity	Number of products sold in the order	1
	Profit	Price of each product unit	100
	Discount amount	Amount of discount applied (UnitPrice * Discount * Quantity)	0.1
	Sales	Total amount for the order after applying discounts (Quantity * UnitPrice * (1 - Discount))	200

### 2.1.2. Business Process: OrderFulfillment

Các câu hỏi cụ thể được đặt ra:

- Câu hỏi: Phương thức vận chuyển của từng phân khúc khách hàng khác nhau theo tháng quý năm?
- Câu hỏi: Tháng có số lượng đơn hàng trong từng loại sản phẩm tăng cao theo từng phương thức vận chuyển.

Bảng FactOrderFulfillment

FactOrderFulfillment	CustomerKey	Key of each customer	1,2,3...
	OrderDateKey	Key of each orderdate	20041123
	ShipDateKey	key of each shipdate	20041123
	OrderKey	Key of each order	1,2,3...

## 2.2. Xây dựng các bảng Dimension

### 2.2.1. DimDate

Date	OrderDate	each customer's order time.		11/8/2016
	OrderDayOfWeek	Each customer's order date by week.		6, 7,...
	OrderDayOfMonth	Each customer's order date by month.		3,4,5,...
	OrderDayOfYear	Each customer's order date by Year.		3,4,5,...
	OrderWeekOfYear	Each customer's order week by year.		1,2,3,4,...
	OrderMonthOfYear	Each customer's order Month by year.		1,2,3,4,...
	OrderQuarter	each customer's order quarter.		1,2,3,4
	OrderYear	each customer's order year.		2014,2015,...
	OrderIsWeekDay	each customer's order WeekDay.		N,Y

### 2.2.2. DimProduct

Products	Products ID	ID of each products.		OFF-PA-10001776
	Products Name	The name of the product.		Wirebound Message Books, Four 2 3/4" x 5"
	Category	The category of the product.		Forms per Page, 600 Sets per Book
	Sub-Category	The sub-category of the product.		Office Supplies, Technology

### 2.2.3. DimCustomer

Customers	Customer ID	ID of each customer.		CG-12520
	Customer Name	The name of customer's		Claire Gute
	Customer Segment	Segment of each customer.		Consumer, Corporate,...
	Contact State	Contact State of customer.		Kentucky, California,...
	Customer Country	Country of origin for the customer		Vietnam
	Customer Region	State or province for the customer (not aval sometimes)		South, West,...
	Customer City	Customer's city		Henderson, Los Angeles,...
	Customer Postal Code	Customer's postal code		42420, 90036,...

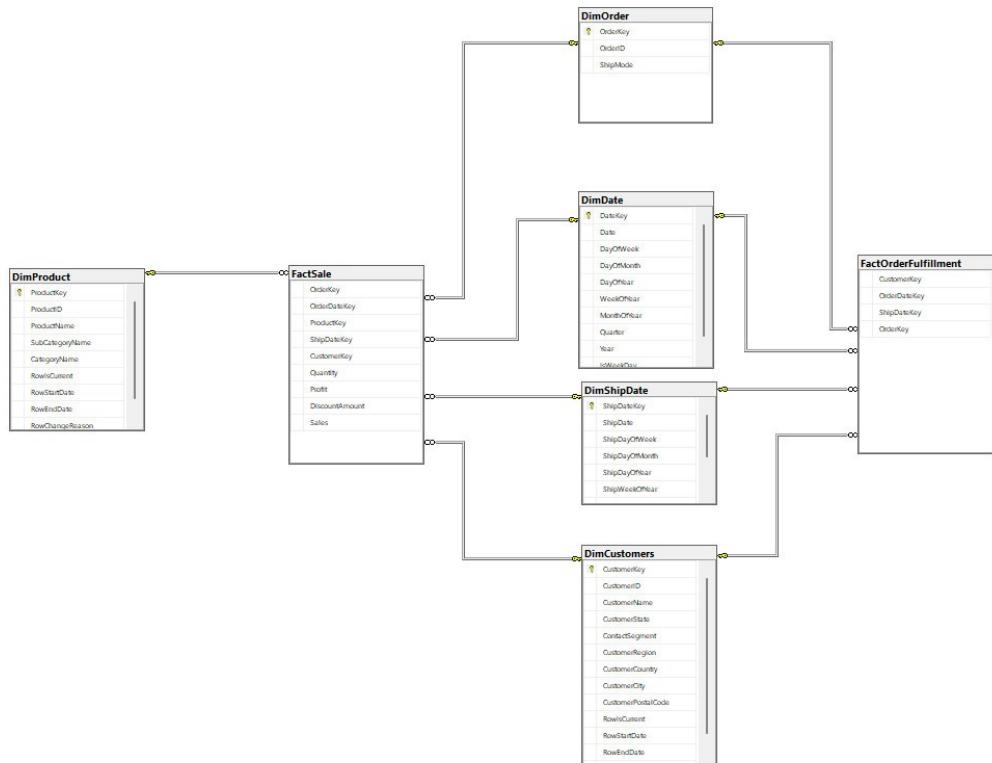
### 2.2.4. DimShipDate

ShipDate	ShipDate	The postal code or ZIP code of the employee's address		11/11/2016
	ShipDayOfWeek	The country where the employee is located		1,2,3,4,...
	ShipDayOfMonth	The personal telephone number of the employee		1,2,...30,31
	ShipDayOfYear	The internal extension or extension number within the organization		1,2,3,...
	ShipWeekOfYear	A photograph of the employee		2,3,4,...
	ShipMonthOfYear	Additional remarks or information about the employee		1,2,3,4,...
	ShipQuarter			1,2,3,4
	ShipYear	Information about the person to whom the employee reports or works under		2014, 2015,...
	ShipIsWeekDay	The file path or URL to the employee's photograph		N,Y

### 2.2.5. DimOrder

Orders	Order ID	ID of each order.		CA-2016-152156,...
	Ship Mode	How to handle each order.		Second Class, Standard Class,...

### 2.3. Star Schema (Lược đồ hình sao)

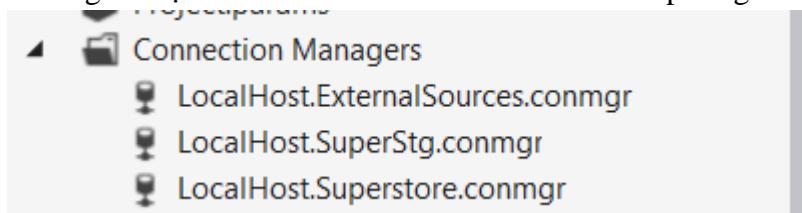


# CHƯƠNG 3: TÍCH HỢP DỮ LIỆU VÀO KHO (SSIS)

Tạo 3 kết nối tới 3 database

(SuperStg, ExternalSources là 2 database rỗng, Superstore là database được tạo từ script chưa có dữ liệu )

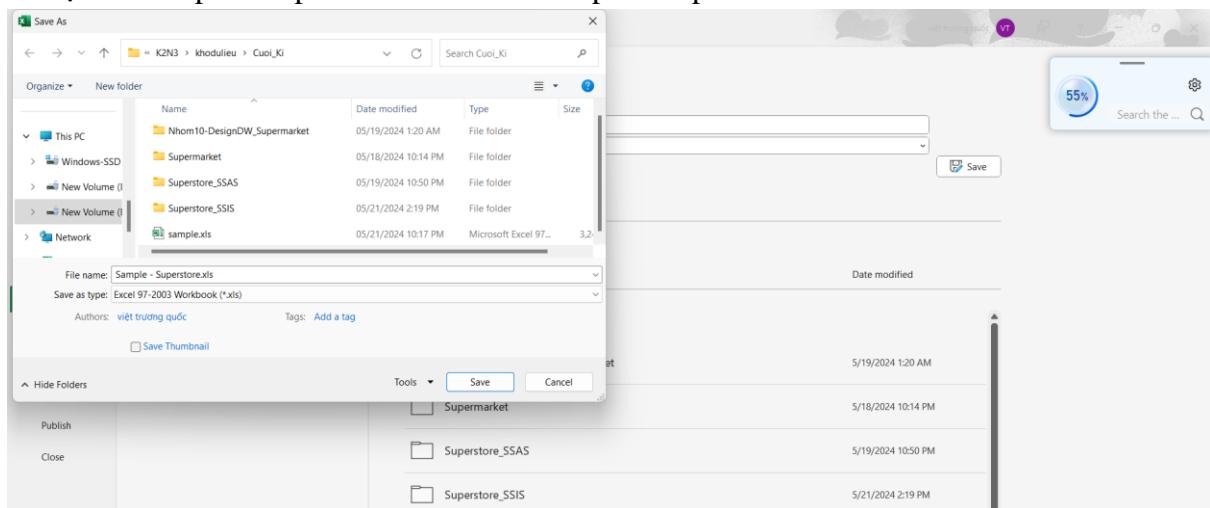
-Luồng dữ liệu đi từ excel ->ExternalSources -> SuperStg-> Superstore



## 3.1. Import dữ liệu gốc vào database ExternalSources bằng SSIS

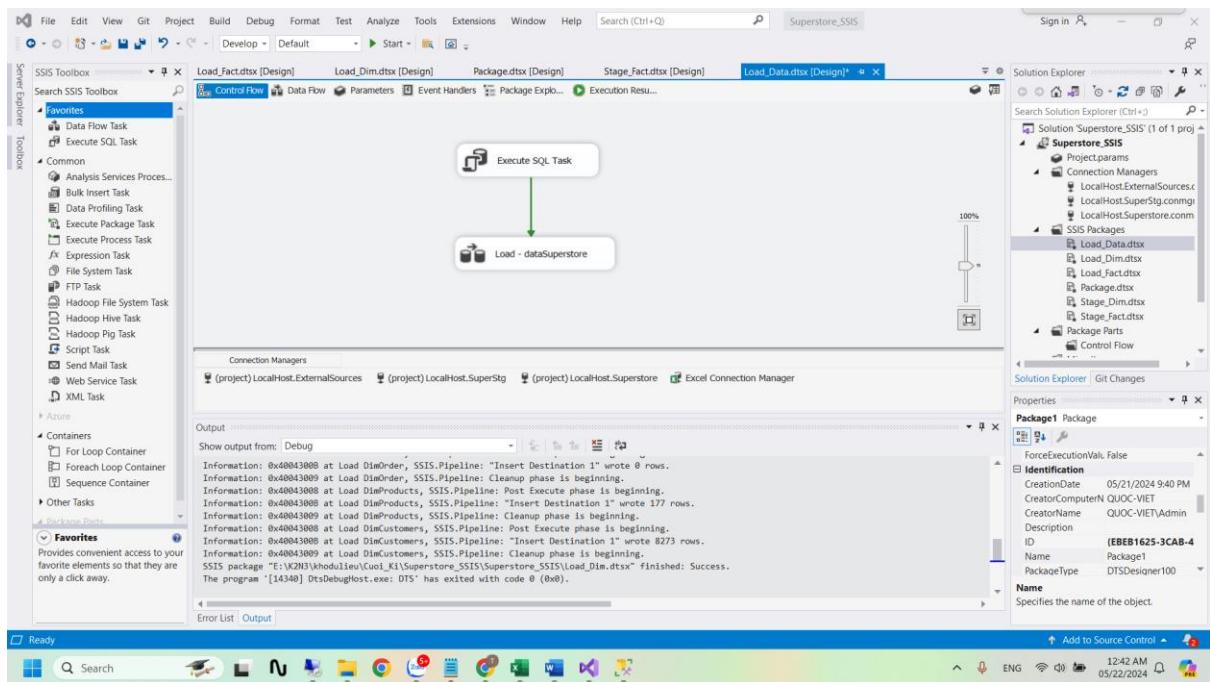
Chuyển dữ liệu gốc từ file có dạng .csv thành .xls or xlsx

Ví dụ : từ Sample - Superstore.csv thành Sample - Superstore.xls



Để tránh SSIS nhận dạng nhầm kiểu dữ liệu của cột.

**Package Load\_Data**



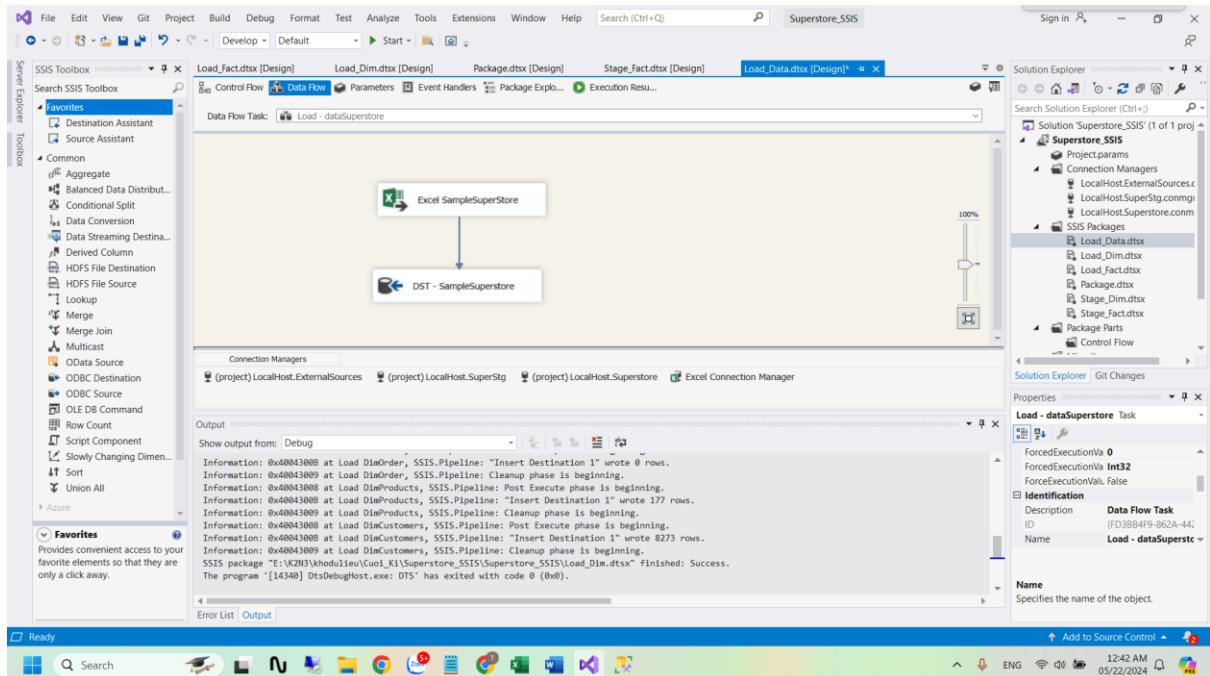
## Execute SQL Task

The screenshot shows the Execute SQL Task Editor dialog box. The General tab is selected, displaying the following configuration:

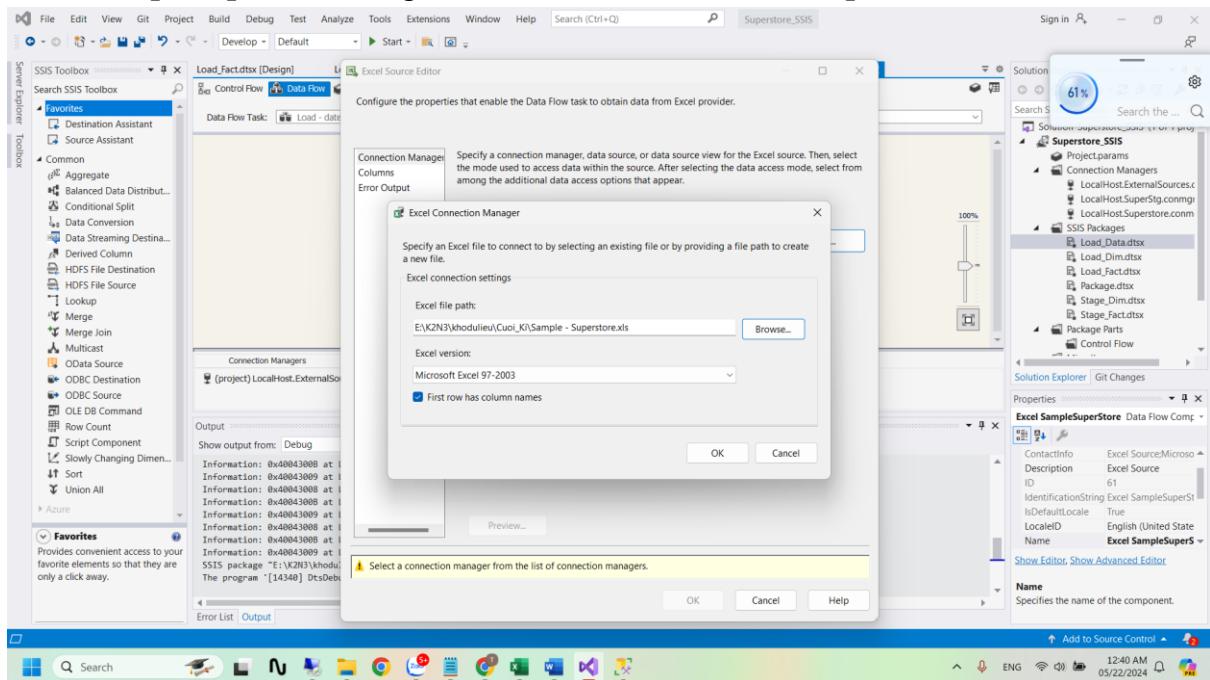
Name	<b>Execute SQL Task</b>
Description	<b>Execute SQL Task</b>
TimeOut	<b>0</b>
CodePage	<b>1252</b>
TypeConversionMode	<b>Allowed</b>
ResultSet	<b>None</b>
ConnectionType	<b>OLE DB</b>
Connection	<b>localhost.ExternalSources</b>
SQLSourceType	<b>Direct input</b>
SQLStatement	<b>Truncate table SampleSuperstore</b>
IsQueryStoredProcedure	<b>False</b>
BypassPrepare	<b>True</b>

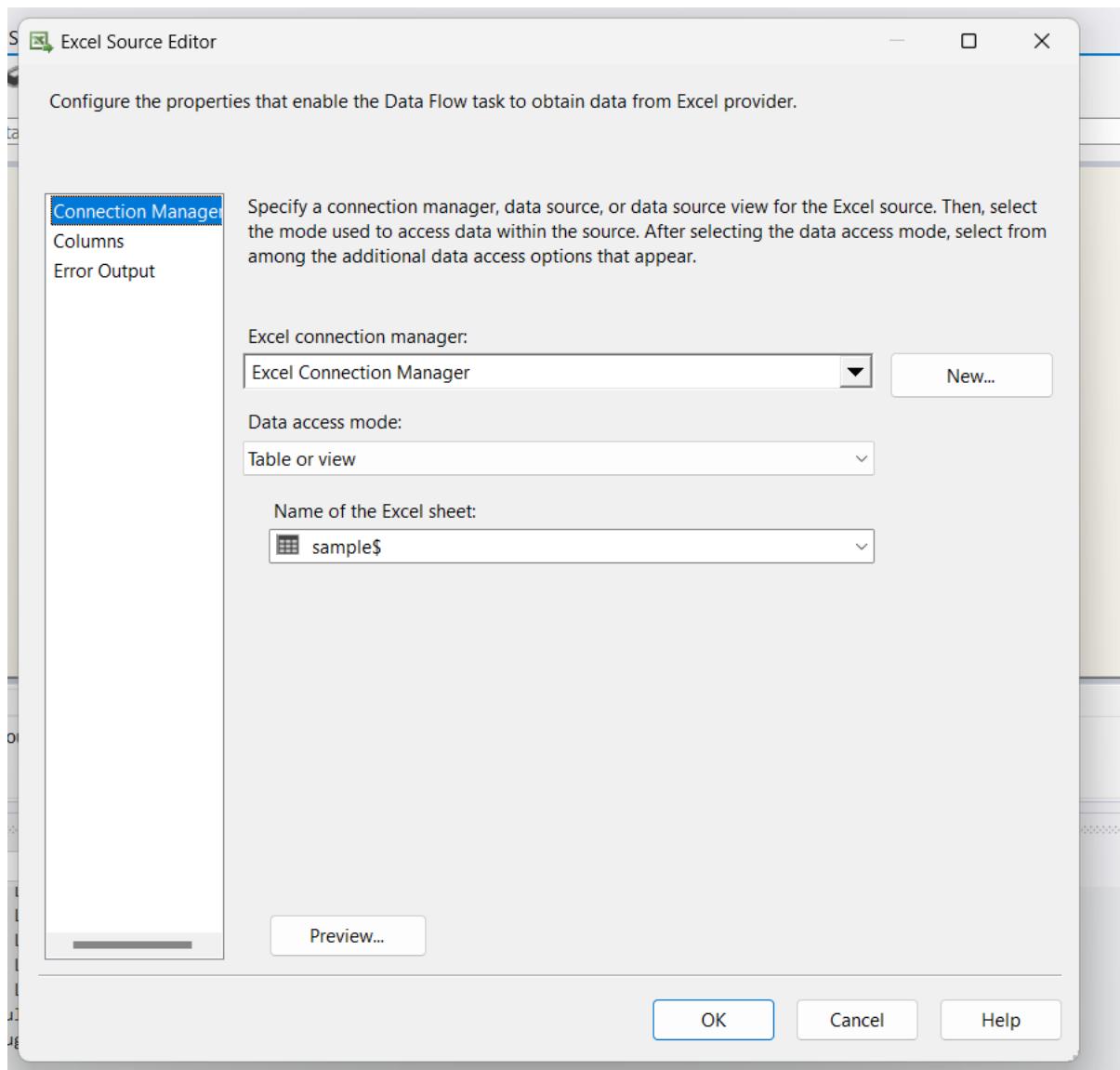
Below the table, there's a Name section with a description: "Specifies the name of the task." At the bottom of the dialog are three buttons: Browse..., Build Query..., and Parse Query. There are also OK, Cancel, and Help buttons.

**Load – dataSuperStore** dùng để truyền dữ liệu từ file excel **Sample - Superstore.xls** vào database ExternalSources



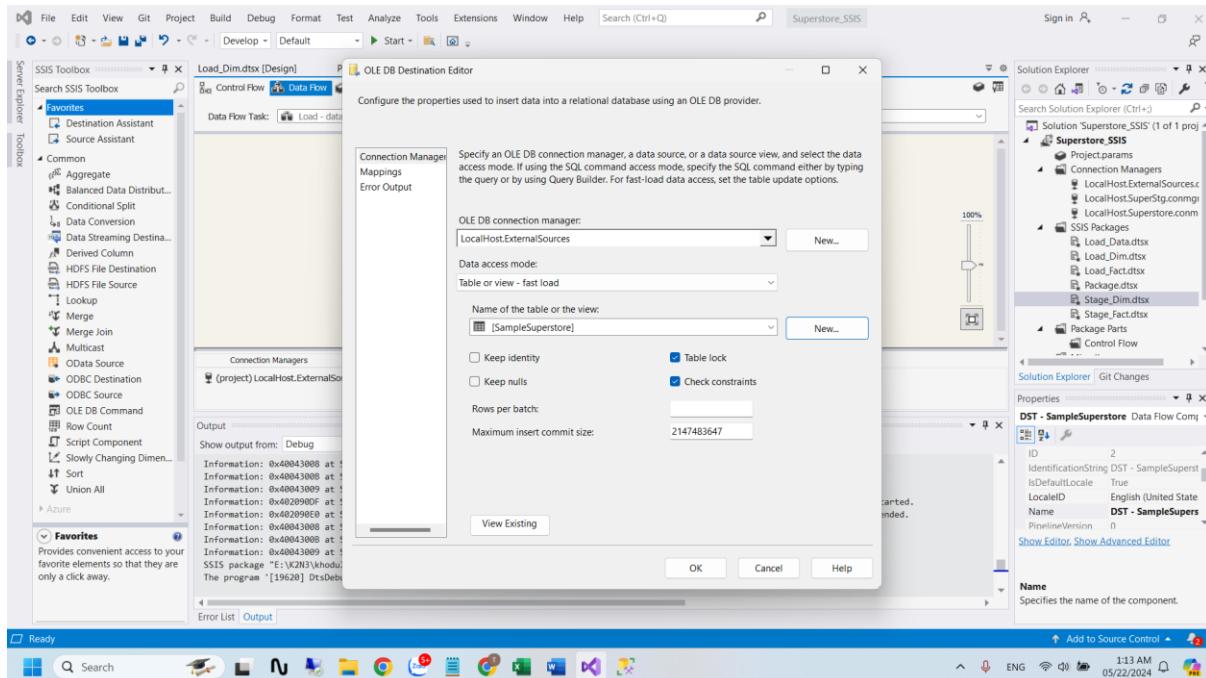
**Excel SampleSuperStore** dùng để chọn dữ liệu từ file sheet **sample\$**





Query result (up to the first 200 rows):																				
Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer	Segment	Country	City	State	Postal Code	Region	Product	Category	Sub-Cat	Product	Sales	Quantity	Discount	Profit	
1	CA-201...	11/08/...	11/11/...	Second...	CG-125...	Claire G...	Consu...	United ...	Hender...	Kentucky	42420	South	FUR-BO...	Furniture	Bookcas...	Bush S...	261.96	2	0	41,913.6
2	CA-201...	11/08/...	11/11/...	Second...	CG-125...	Claire G...	Consu...	United ...	Hender...	Kentucky	42420	South	FUR-CH...	Furniture	Chairs	Hon Del...	731.94	3	0	219,582
3	CA-201...	06/12/...	06/16/...	Second...	DV-130...	Darrin ...	Corpor...	United ...	Los An...	California	90036	West	OFF-IA...	Office S...	Labels	Self-Ad...	14,62	2	0	6,871.4
4	US-201...	10/11/...	10/18/...	Standard...	SO-203...	Sean O...	Consu...	United ...	Fort La...	Florida	33311	South	FUR-TA...	Furniture	Tables	Bretfor...	957,577.5	5	0.45	-383,031
5	US-201...	10/11/...	10/18/...	Standard...	SO-203...	Sean O...	Consu...	United ...	Fort La...	Florida	33311	South	OFF-ST...	Office S...	Storage	Eldon F...	22,368	2	0.2	2,516.4
6	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consu...	United ...	Los An...	California	90032	West	FUR-FU...	Furniture	Furnish...	Eldon E...	48.86	7	0	14,169.4
7	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consu...	United ...	Los An...	California	90032	West	OFF-AR...	Office S...	Art	Newell ...	7.28	4	0	1,965.6
8	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consu...	United ...	Los An...	California	90032	West	TEC-PH...	Technol...	Phones	Mitel 53...	907,152	6	0.2	90,715.2
9	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consu...	United ...	Los An...	California	90032	West	OFF-BE...	Office S...	Binders	DXL An...	18,504	3	0.2	5,782.5
10	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consu...	United ...	Los An...	California	90032	West	OFF-AP...	Office S...	Applian...	Bekin F...	114.9	5	0	34.47
11	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consu...	United ...	Los An...	California	90032	West	FUR-TA...	Furniture	Tables	Chrom...	1706,184	9	0.2	85,209.2
12	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consu...	United ...	Los An...	California	90032	West	TEC-PH...	Technol...	Phones	Konfidi ...	911,424	4	0.2	68,356.8
13	CA-201...	04/15/...	04/20/...	Standard...	AA-104...	Andrew ...	Consu...	United ...	Concord	North C...	28027	South	OFF-PA...	Office S...	Paper	Xerox 1...	15,552	3	0.2	5,443.2
14	CA-201...	12/05/...	12/10/...	Standard...	IM-15070	Irene M...	Consu...	United ...	Seattle	Washin...	98103	West	OFF-BE...	Office S...	Binders	Follow...	407,976	3	0.2	132,592.2
15	US-201...	11/22/...	11/26/...	Standard...	HP-148...	Harold ...	Home ...	United ...	Fort W...	Texas	76106	Central	OFF-AP...	Office S...	Applian...	Holmes ...	68.81	5	0.6	-123,858.6
16	US-201...	11/22/...	11/26/...	Standard...	HP-148...	Harold ...	Home ...	United ...	Fort W...	Texas	76106	Central	OFF-AP...	Office S...	Binders	Stor...ex	2,544	3	0.6	-3,816.0
17	CA-201...	11/11/...	11/18/...	Standard...	PK-190...	Pete Kro...	Consu...	United ...	Madison	Wiscon...	53711	Central	OFF-ST...	Office S...	Storage	Stur...D...	665.88	6	0	13,317.6
18	CA-201...	05/13/...	05/19/...	Second...	AG-102...	Alejand...	Consu...	United ...	West J...	Utah	84084	West	OFF-ST...	Office S...	Storage	Follow...	55.5	2	0	9.99
19	CA-201...	08/27/...	09/01/...	Second...	ZD-219...	Zuschi ...	Consu...	United ...	San Fra...	California	94109	West	OFF-AR...	Office S...	Art	Newell ...	8.56	2	0	2,482.4
20	CA-201...	08/27/...	09/01/...	Second...	ZD-219...	Zuschi ...	Consu...	United ...	San Fra...	California	94109	West	TEC-PH...	Technol...	Phones	Cisco S...	213,48	3	0.2	16,011.1
21	CA-201...	08/27/...	09/01/...	Second...	ZD-219...	Zuschi ...	Consu...	United ...	San Fra...	California	94109	West	OFF-BE...	Office S...	Binders	Wilson ...	22,72	4	0.2	7,384.0
22	CA-201...	12/09/...	12/13/...	Standard...	KB-165...	Ken Black	Corpor...	United ...	Fremont	Nebraska	68025	Central	OFF-AR...	Office S...	Art	Newell ...	19.46	7	0	5,059.6
23	CA-201...	12/09/...	12/13/...	Standard...	KB-165...	Ken Black	Corpor...	United ...	Fremont	Nebraska	68025	Central	OFF-AP...	Office S...	Applian...	Acco SU...	60.34	7	0	15,688.4
24	US-201...	07/16/...	07/18/...	Second...	SF-20065	Sandra ...	Consu...	United ...	Philadelphia	Pennsyl...	19140	East	FUR-CH...	Furniture	Chars	Global ...	71,372	2	0.3	-1,019.6
25	CA-201...	09/25/...	09/30/...	Standard...	EB-138...	Emily B...	Consu...	United ...	Orem	Utah	84057	West	FUR-TA...	Furniture	Tables	Bretfor...	104,63	3	0	240,264.9
26	CA-201...	01/16/...	01/20/...	Second...	EH-139...	Eric Hof...	Consu...	United ...	Los An...	California	90049	West	OFF-BE...	Office S...	Binders	Wilson ...	11,648	2	0.2	4,222.4
27	CA-201...	01/16/...	01/20/...	Second...	EH-139...	Eric Hof...	Consu...	United ...	Los An...	California	90049	West	TEC-AC...	Technol...	Access...	Imation...	90.57	3	0	11,774.1
28	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consu...	United ...	Philadelphia	Pennsyl...	19140	East	FUR-BO...	Furniture	Bookcas...	Riversid...	308,343	7	0.5	-1665.0
29	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consu...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-BE...	Office S...	Binders	Avery ...	9,618	2	0.7	-7,053.2
30	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consu...	United ...	Philadelphia	Pennsyl...	19140	East	FUR-FU...	Furniture	Furnish...	Howard...	124.2	3	0.2	15,525.0
31	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consu...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-EN...	Office S...	Enviro...	Poly Str...	3,264	2	0.2	1,101.6
32	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consu...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-AR...	Office S...	Art	BOSTO...	86,304	6	0.2	9,709.2
33	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consu...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-BE...	Office S...	Binders	Acco Pr...	6,858	6	0.7	-5,715.0
34	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consu...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-AD...	Office S...	Art	Lumbar...	1,746	7	0.5	5,546.0

## DST- SampleSuperStore lấy dữ liệu từ excel truyền vào table SampleSuperstore của database ExternalSources



**OLE DB Destination Editor**

Configure the properties used to insert data into a relational database using an OLE DB provider.

**Connection Manager**

**Mappings**

**Error Output**

**Create Table**

Specify an OLE DB connection manager, a data source, or a data source view, and select the data hand either by typing update options.

```
CREATE TABLE [SampleSuperstore] (
    [Row_ID] smallint,
    [Order_ID] nvarchar(30),
    [Order_Date] datetime,
    [Ship_Date] datetime,
    [Ship_Mode] nvarchar(20),
    [Customer_ID] nvarchar(30),
    [Customer_Name] nvarchar(40),
    [Segment] nvarchar(30),
    [Country] nvarchar(15),
    [City] nvarchar(40),
    [State] nvarchar(30),
    [Postal_Code] nvarchar(10),
    [Region] nvarchar(15),
    [Product_ID] nvarchar(40),
    [Category] nvarchar(15),
    [Sub_Category] nvarchar(40),
    [Product_Name] nvarchar(150),
    [Sales] float,
```

**New...**

**New...**

**OK**   **Cancel**

**View Existing**

**OK**   **Cancel**   **Help**

**Available Input...**

**Available D...**

Input Column	Destination Column
Segment	Segment
Country	Country
City	City
State	State
Region	Region
Category	Category
Sales	Sales
Quantity	Quantity
Discount	Discount
Profit	Profit
Row ID	Row_ID
Order ID	Order_ID
Order Date	Order_Date
Ship Date	Ship_Date
Ship Mode	Ship_Mode
Customer ID	Customer_ID
Customer Name	Customer_Name

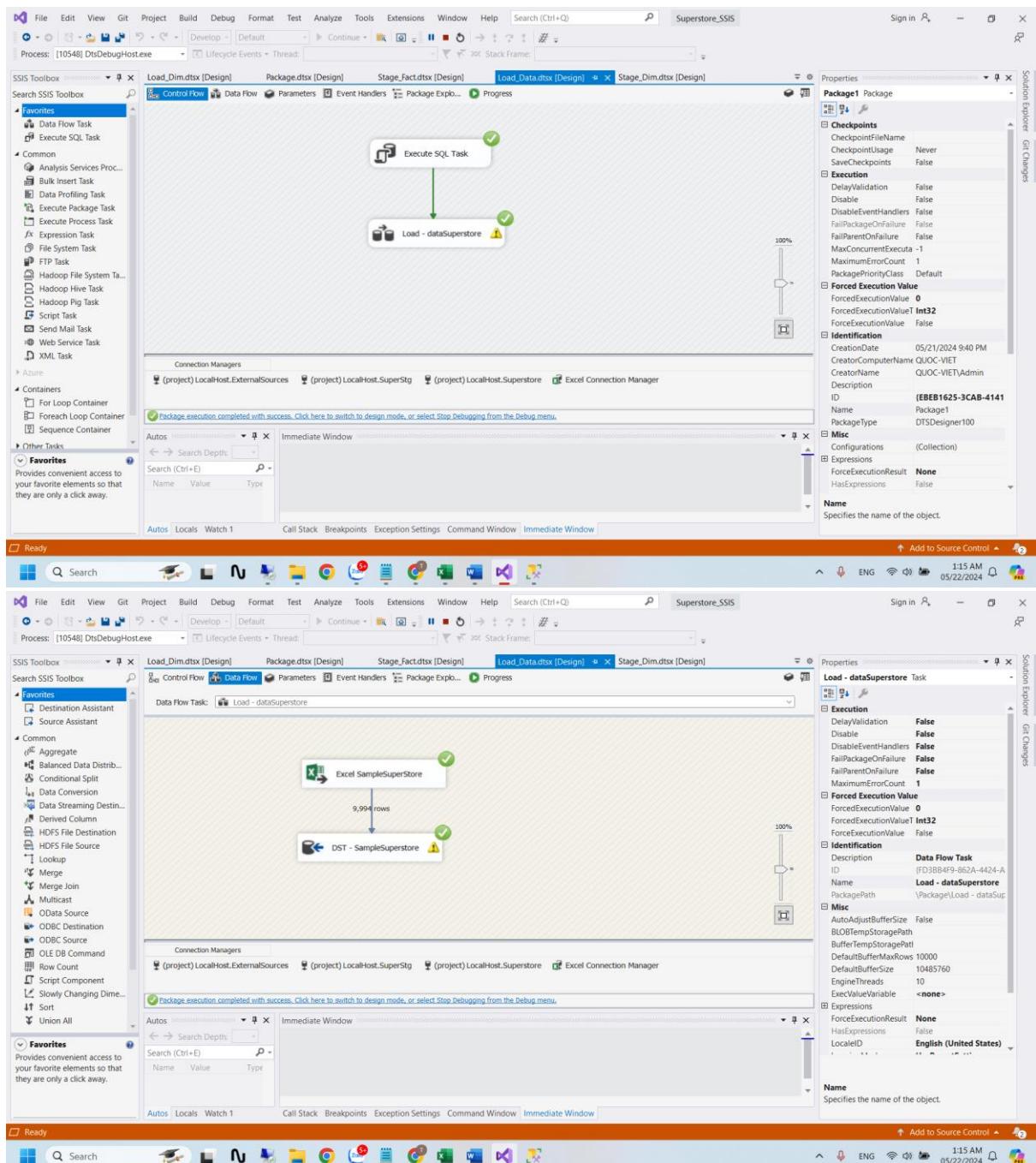
Configure the properties used to insert data into a relational database using an OLE DB provider.

**Connection Manager**

**Mappings**

**Error Output**

**OK**   **Cancel**   **Help**



Kiểm tra database ExternalSources trong SQL xem đã có dữ liệu chưa

The screenshot shows the Microsoft SQL Server Management Studio interface. In the center, there is a large grid displaying a dataset with 200 rows. The columns include Row\_ID, Order\_ID, Order\_Date, Ship\_Date, Ship\_Mode, Customer\_ID, Customer\_Lname, Segment, Country, City, and State. The data spans from 2013 to 2017. To the right of the grid is a Properties pane with tabs for [Qry] Query1.dtq, Identity, and Query Designer. The Query Designer tab shows a query with various clauses like GROUP BY, DISTINCT, and TOP. At the bottom of the screen is a taskbar with several icons.

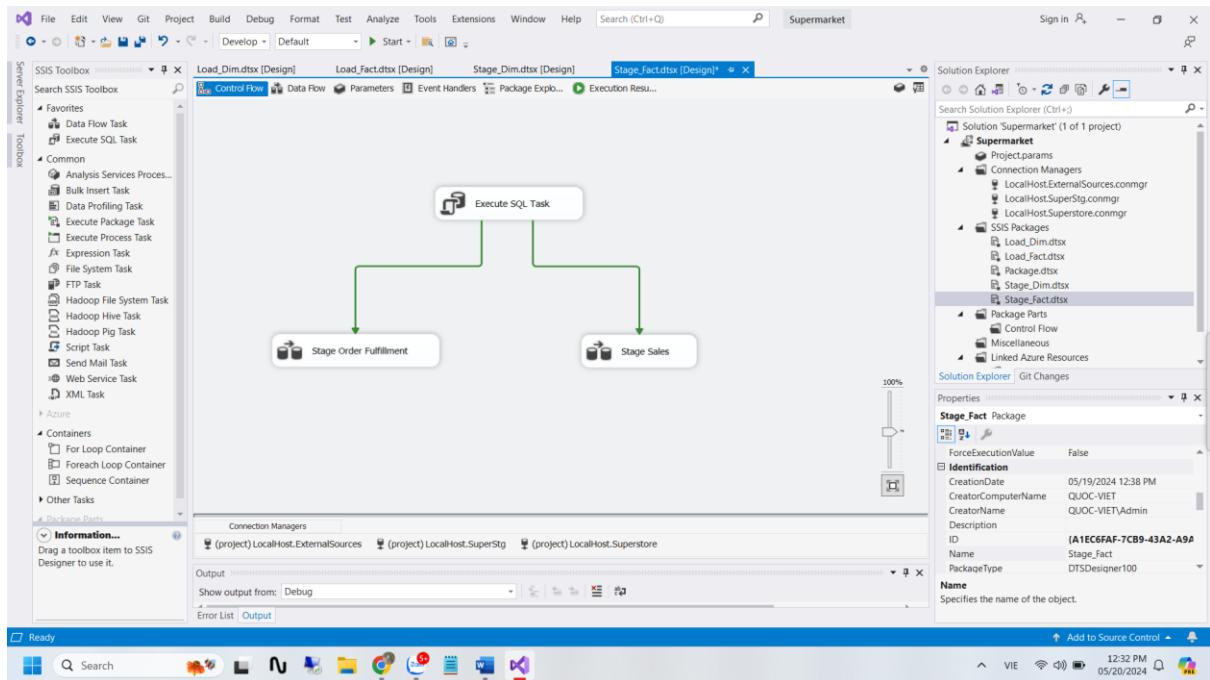
### 3.2. Import dữ liệu vào các bảng dimension:

Ta phải Load từ SRC vào Stage

Stage các bảng Dim

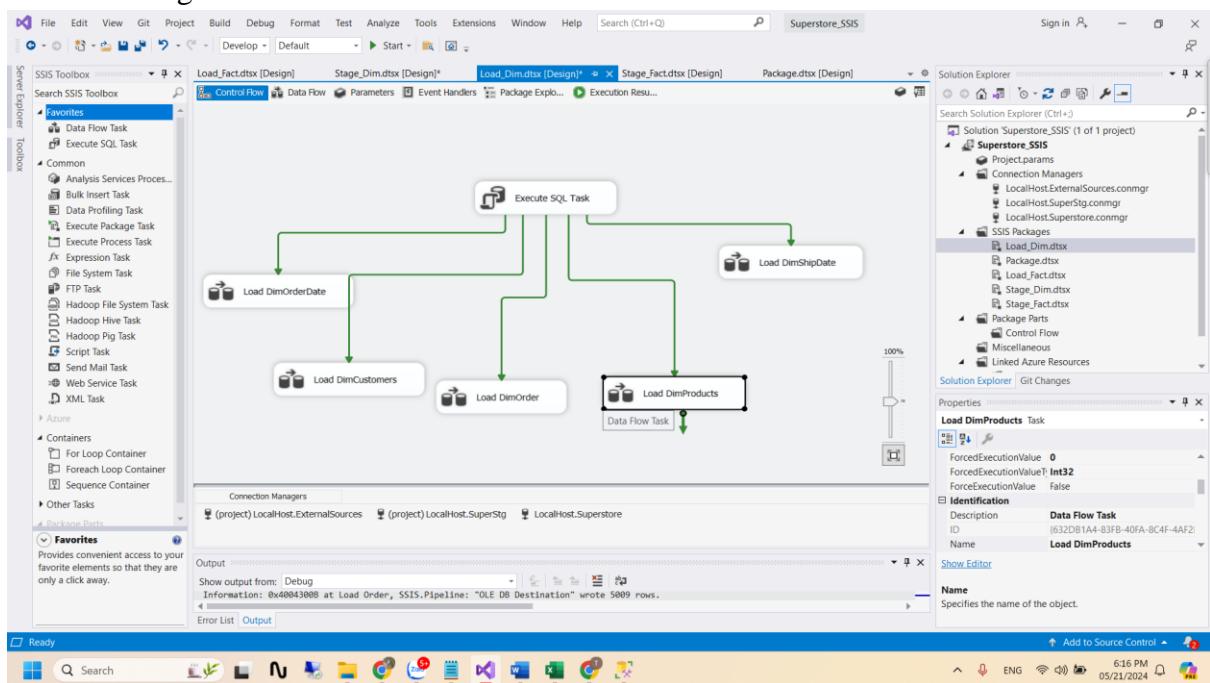
The screenshot shows the Microsoft Visual Studio interface with the SSIS Toolbox open. A control flow diagram is displayed, showing five data flow tasks: Stage OrderDate, Stage Product, Stage Order, Stage ShipDate, and Stage Customer. These tasks are connected to a central Execute SQL Task. The Execute SQL Task has three outgoing arrows pointing to the Stage OrderDate, Stage Product, and Stage Customer tasks. The Stage Order and Stage ShipDate tasks have outgoing arrows pointing to the Stage Customer task. On the left, the SSIS Toolbox shows various task categories like Data Flow Task, Execute SQL Task, and Bulk Insert Task. On the right, the Solution Explorer shows a project named Superstore\_SSIS with multiple packages: Stage\_Dim.dtsx, Stage\_Fact.dtsx, Load\_Dim.dtsx, Load\_Fact.dtsx, and Package.dtsx. The Stage\_Dim.dtsx package is selected. The Properties pane on the far right shows details for the Stage\_Dim package, including its ID, Name, and Description.

Stage các bảng Fact

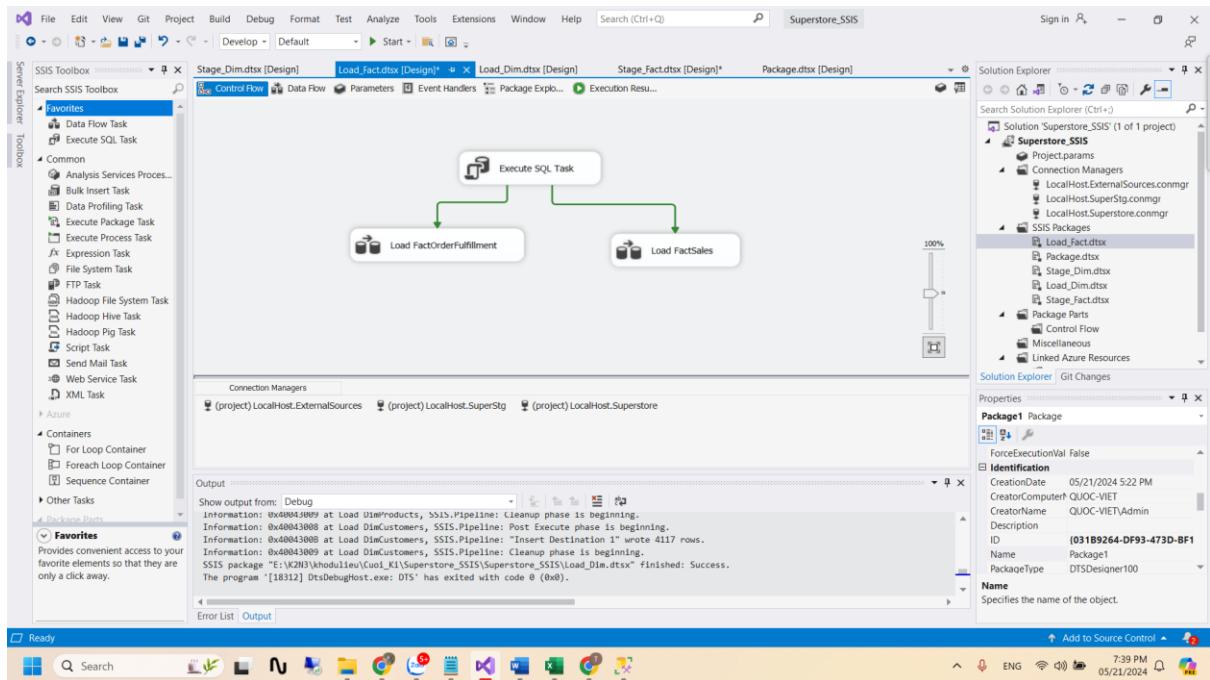


Sau đó Load từ Stage vào các bảng Dim và bảng Fact

Load các bảng Dim



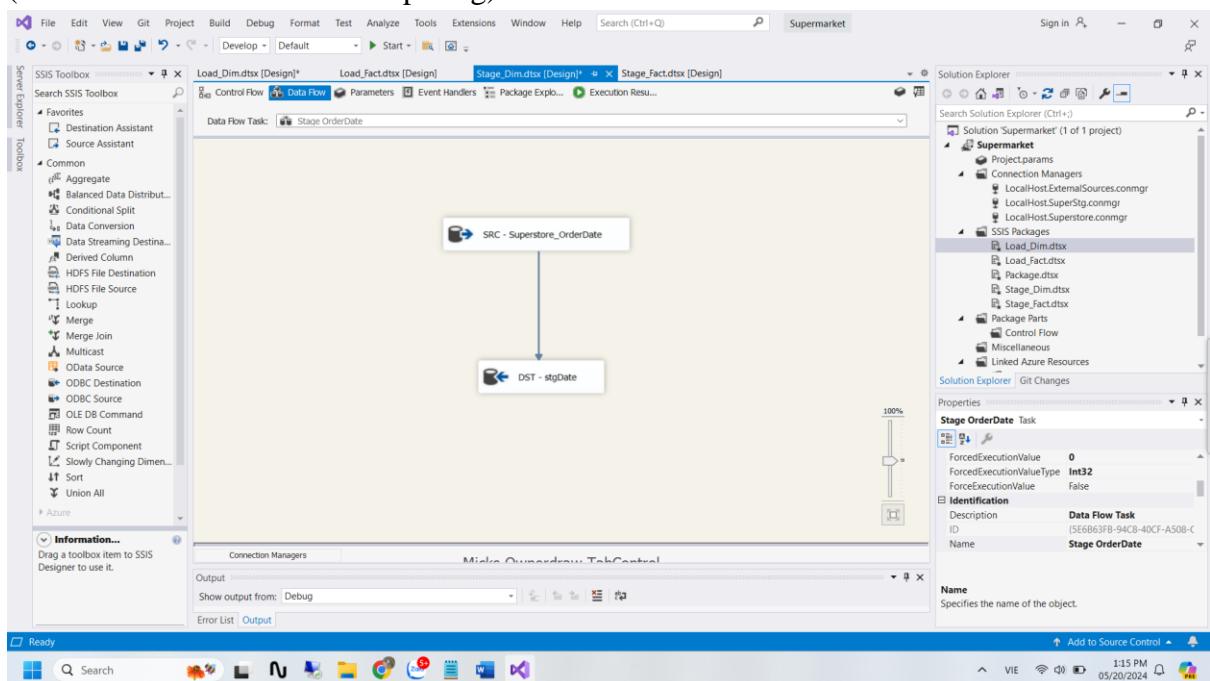
Load các bảng Fact



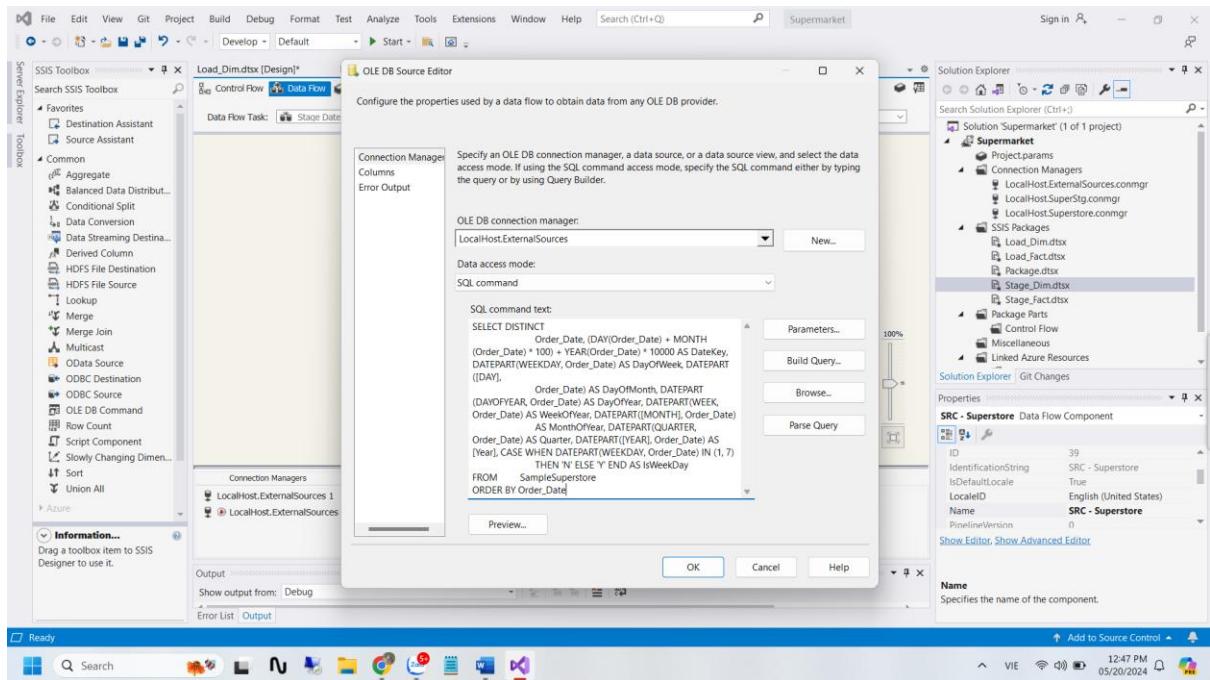
### 3.2.1.OderDate Dimension

#### a) Load từ nguồn vào bảng stgDate

(load từ ExternalSources -> SuperStg)



**Src – SuperStore\_OrderDate** là tập hợp các ngày có thực hiện đặt hàng trong bảng SampleSuperstore trong database ExternalSources



## Script :

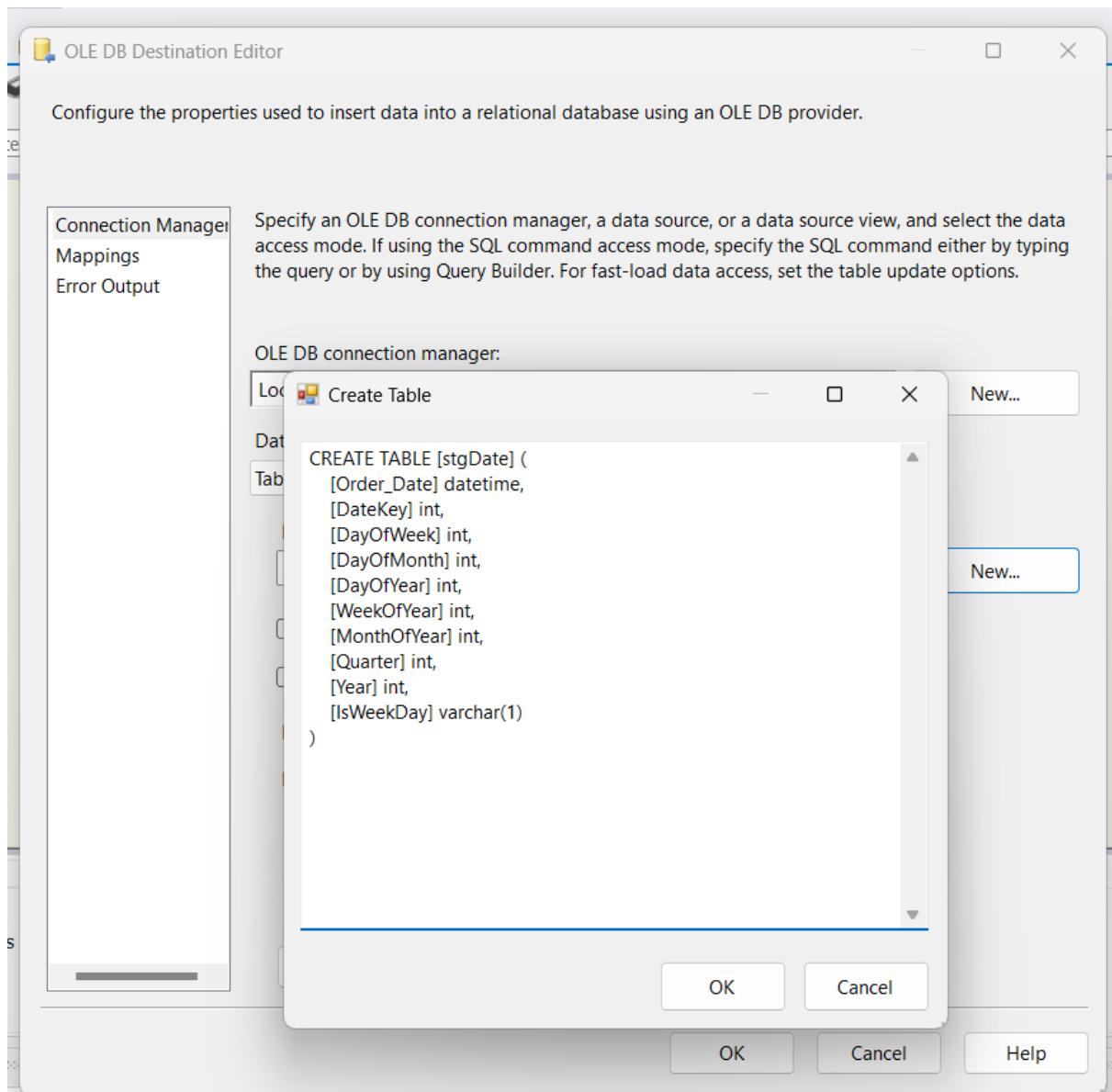
```

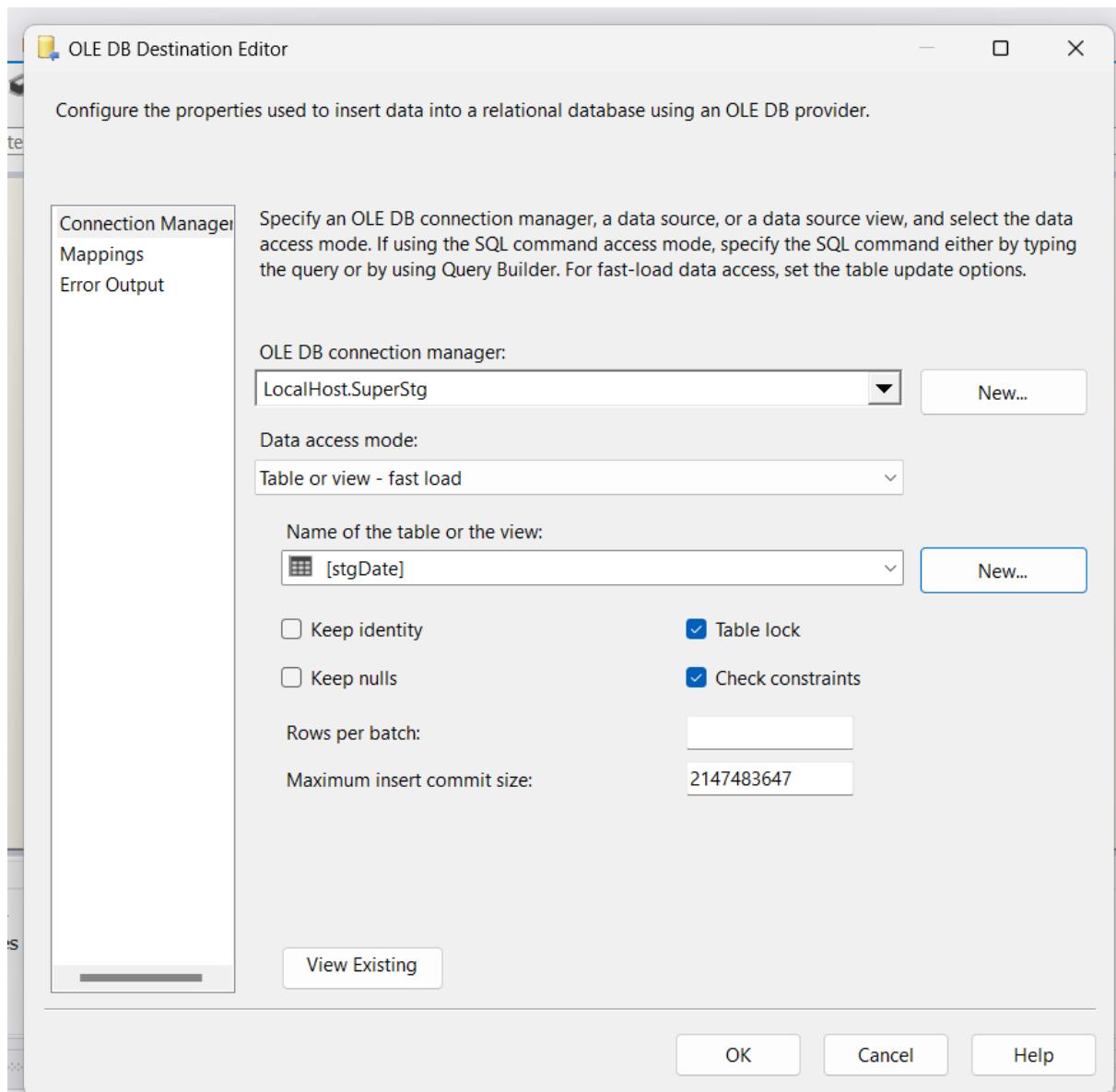
SELECT DISTINCT Order_Date,
(DAY(Order_Date) + MONTH(Order_Date) * 100) + YEAR(Order_Date) * 10000 AS DateKey,
DATEPART(WEEKDAY, Order_Date) AS DayOfWeek,
DATEPART([DAY],Order_Date) AS DayOfMonth,
DATEPART(DAYOFYEAR, Order_Date) AS DayOfYear,
DATEPART(WEEK, Order_Date) AS WeekOfYear,
DATEPART([MONTH], Order_Date) AS MonthOfYear,
DATEPART(QUARTER, Order_Date) AS Quarter,
DATEPART([YEAR], Order_Date) AS [Year],
CASE WHEN DATEPART(WEEKDAY, Order_Date) IN (1, 7) THEN 'N' ELSE 'Y' END AS IsWeekDay
FROM SampleSuperstore
ORDER BY Order_Date

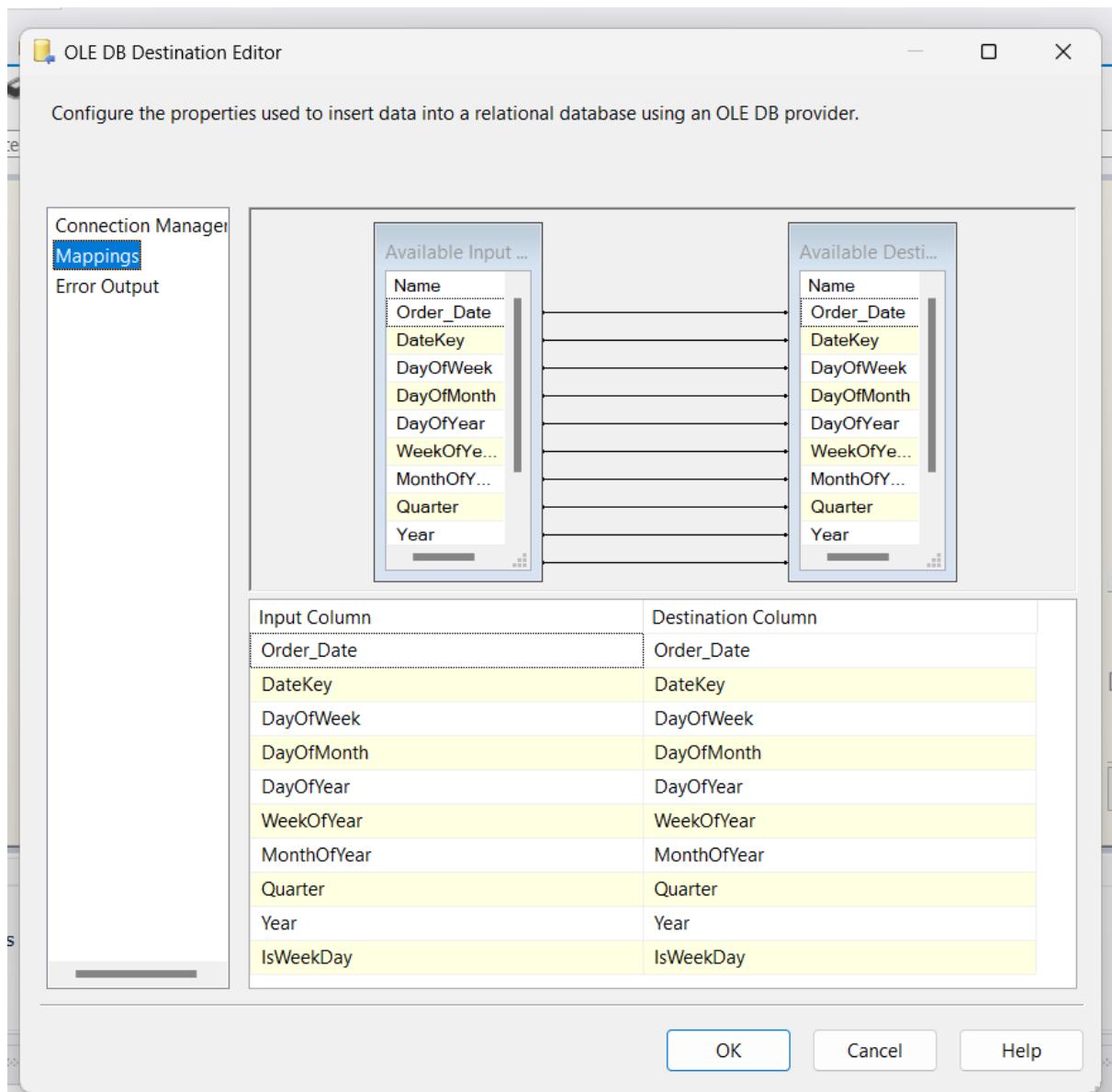
```

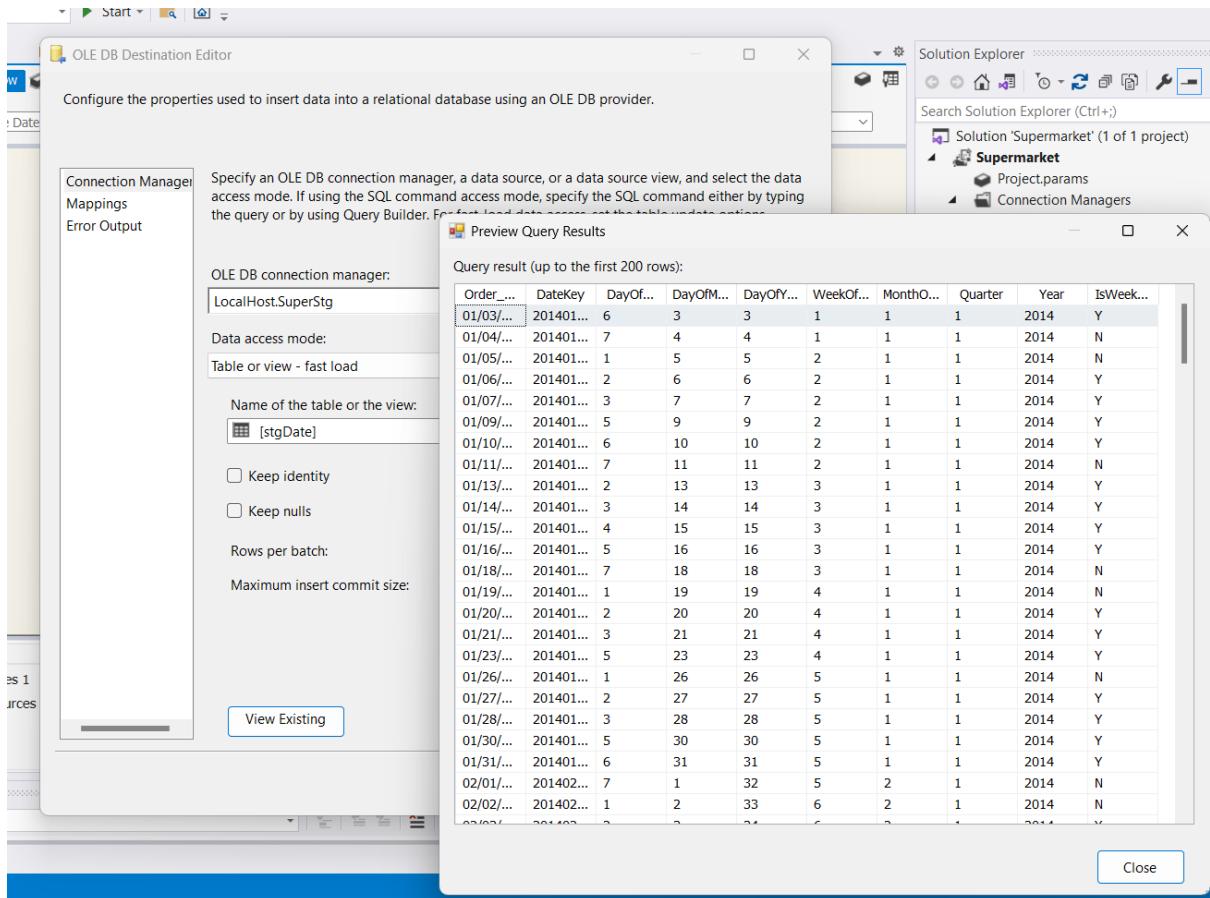
Ta sẽ lấy nguồn dữ liệu này để có thể tạo ra stgDate chứa thời gian thống kê kho và thời gian đặt hàng

## DST – stgDate:

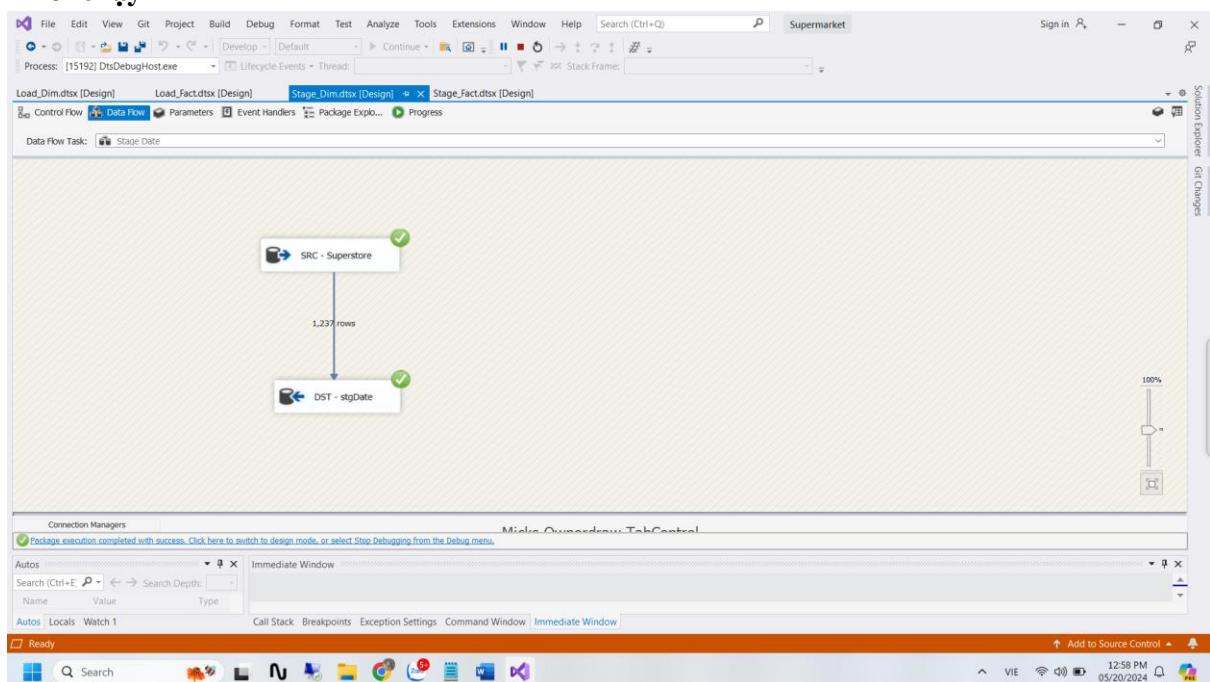






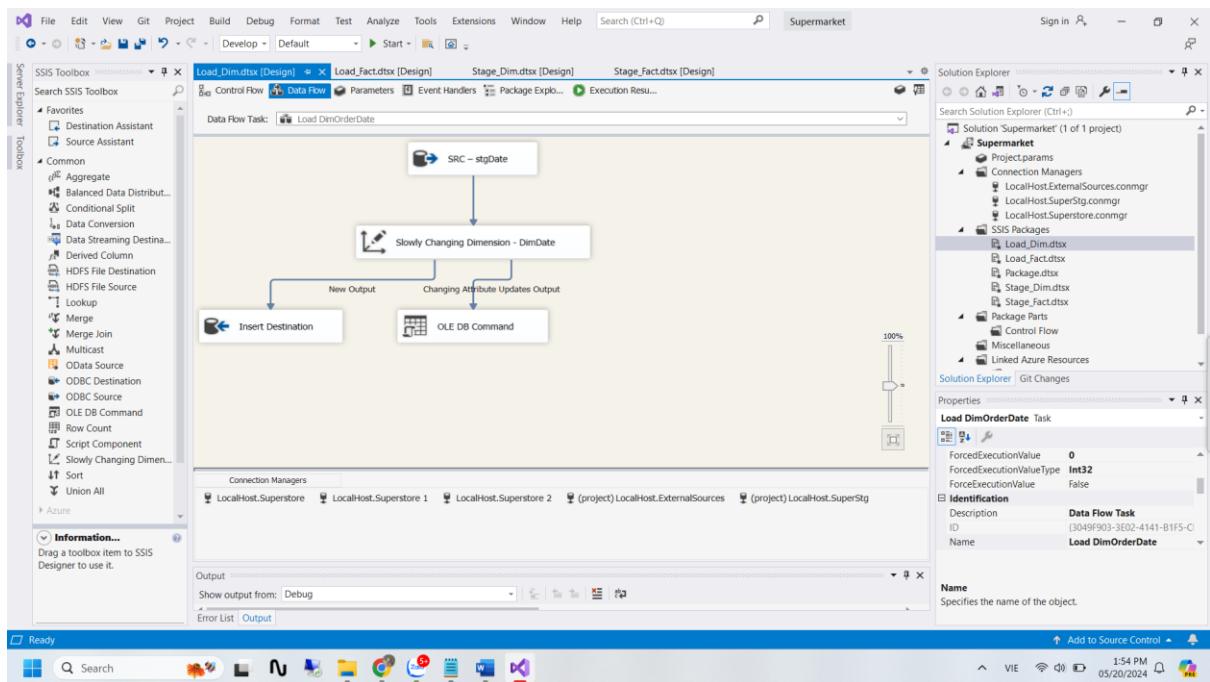


## Khởi chạy

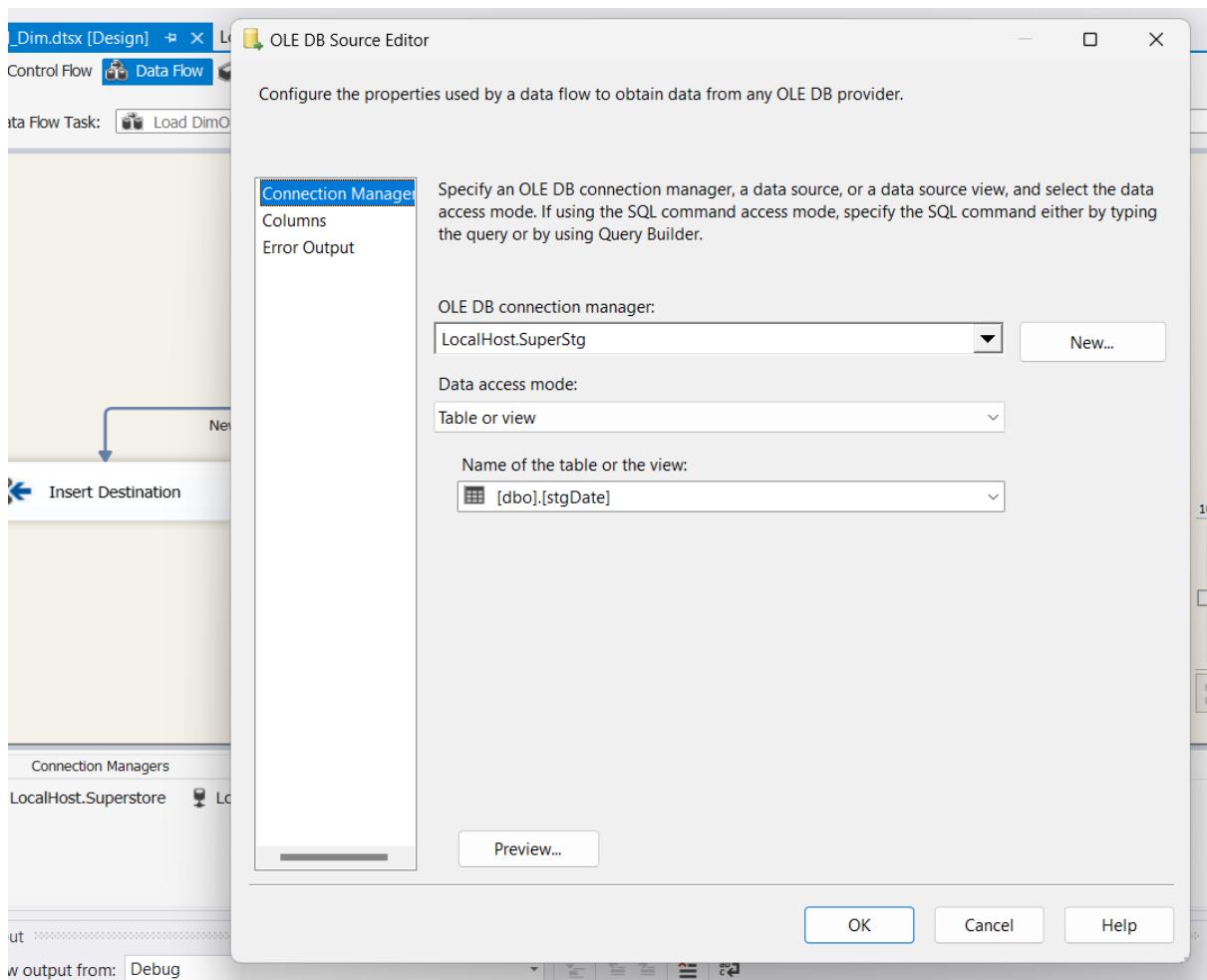


## b) Load dữ liệu từ stgDate vào DimDate

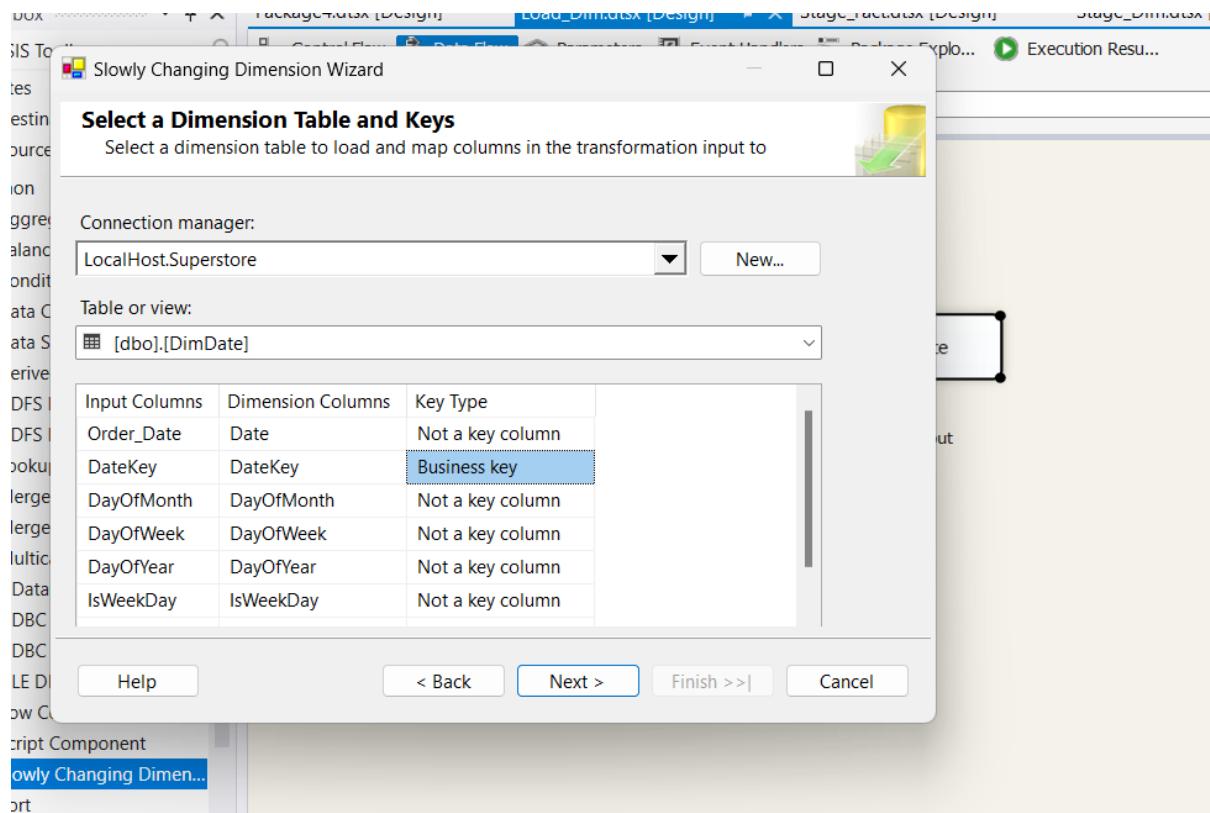
(load từ SuperStg-> Superstore)



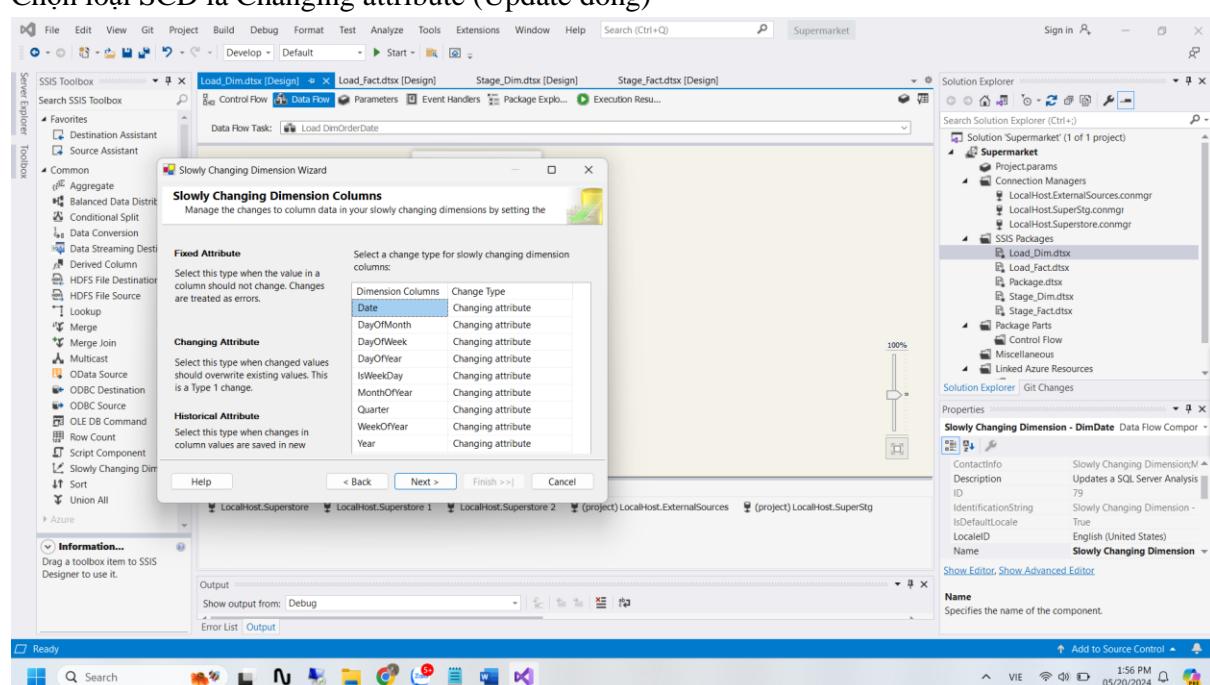
**SRC – stgDate** là dữ liệu từ bảng stgDate đã load ở trên



**Slowly Changing Dimension – DimDate** load dữ liệu từ stgDate vào DimDate

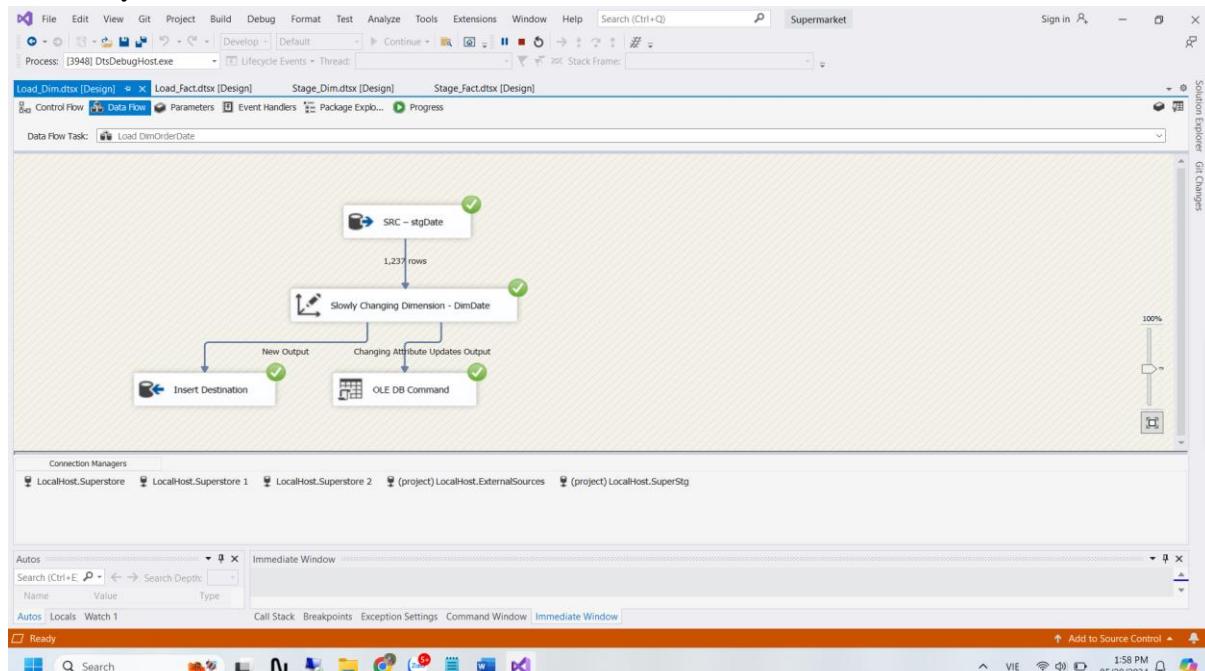


Chọn loại SCD là Changing attribute (Update dòng)



Khi thực hiện xong Slowly Changing Dimension – DimDate thì Insert Destination và OLE DB Command sẽ xuất hiện.

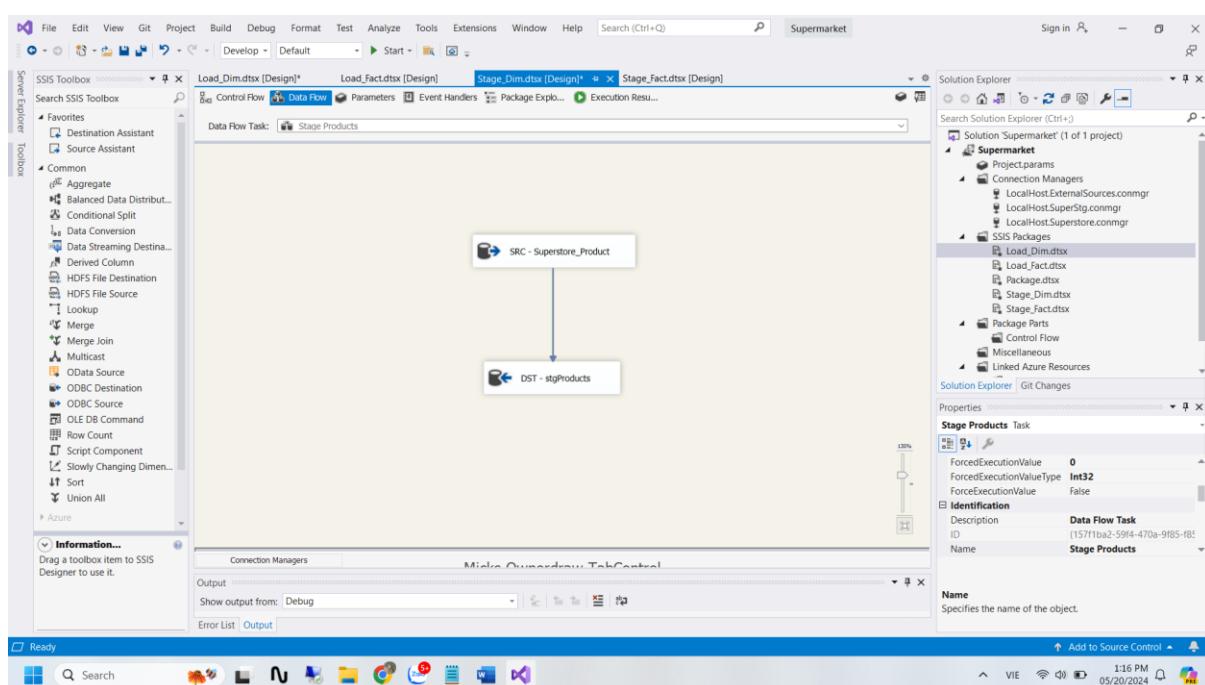
Khởi chạy



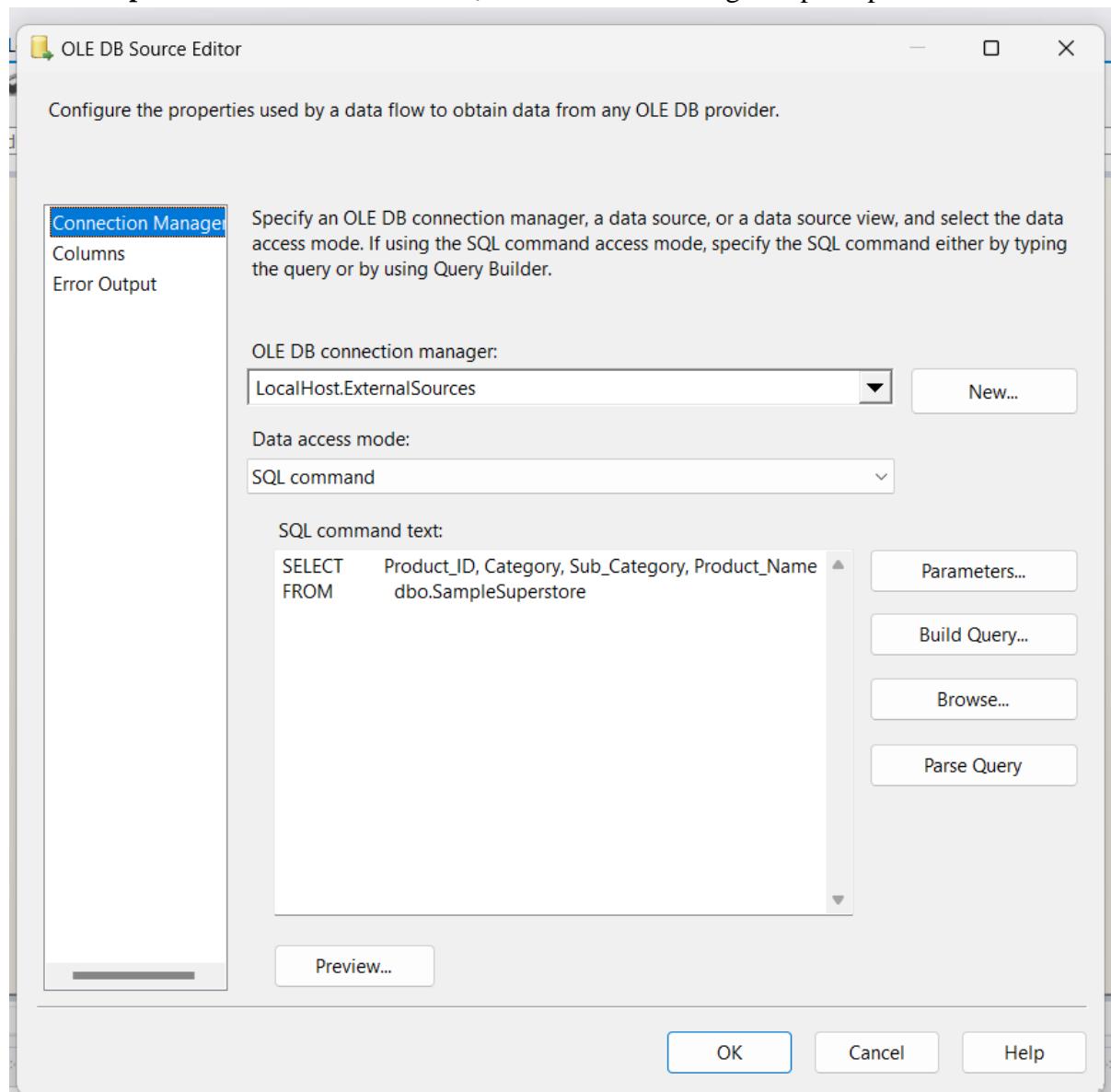
Khi chạy hoàn tất dữ liệu đã được Load vào DimDate

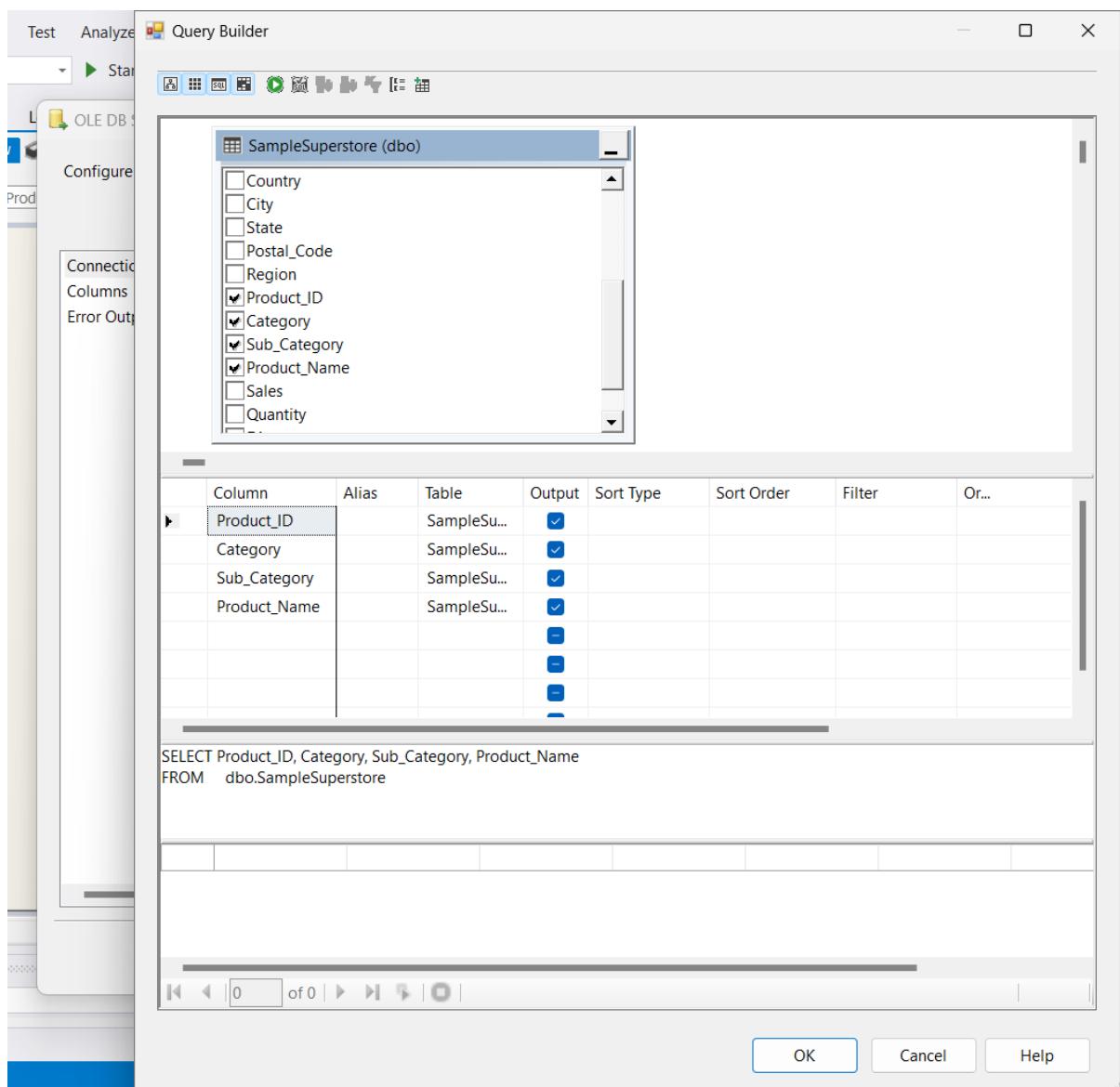
### 3.2.2. Product Dimension

#### a) Load từ nguồn vào bảng stgProduct

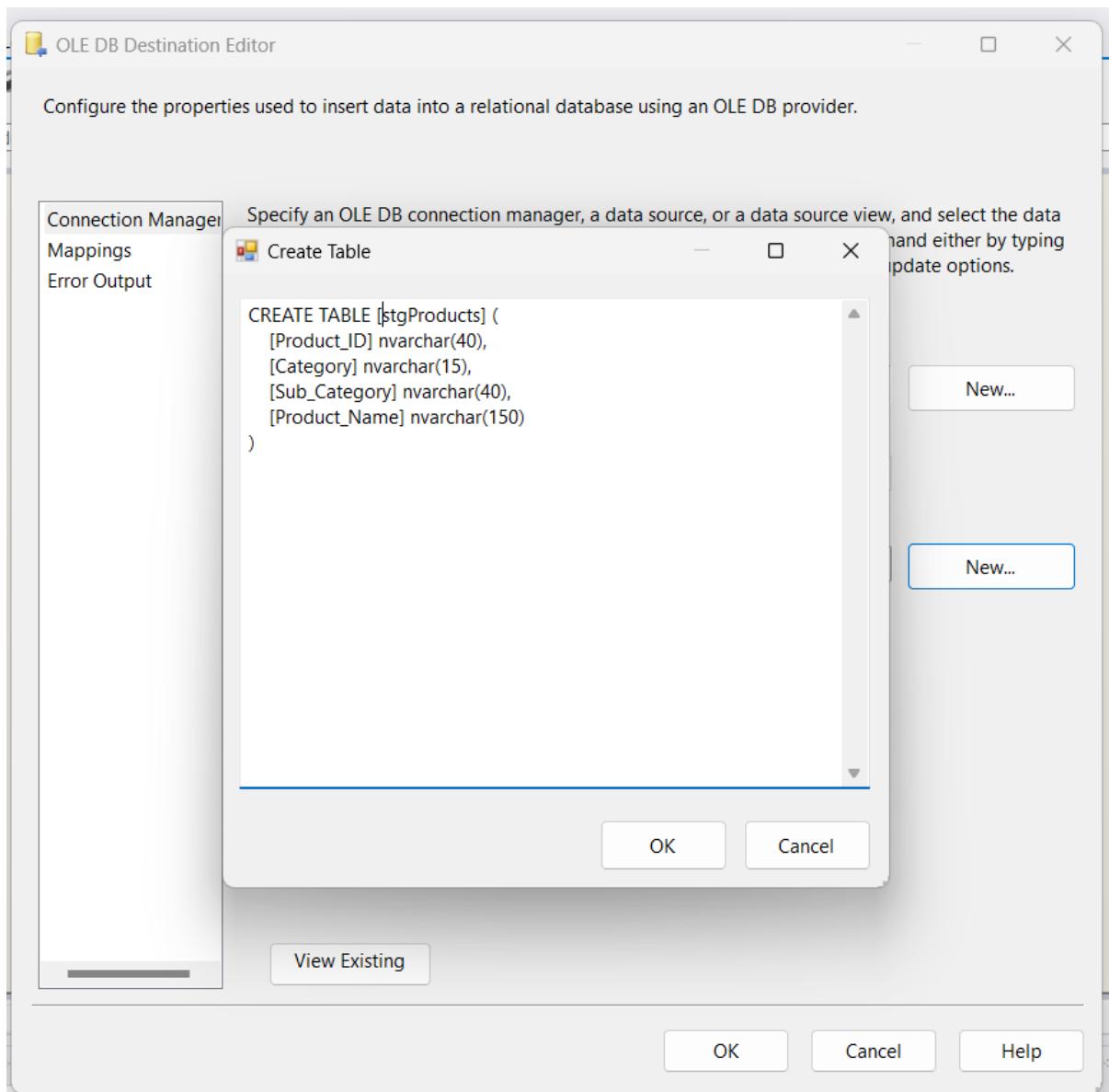


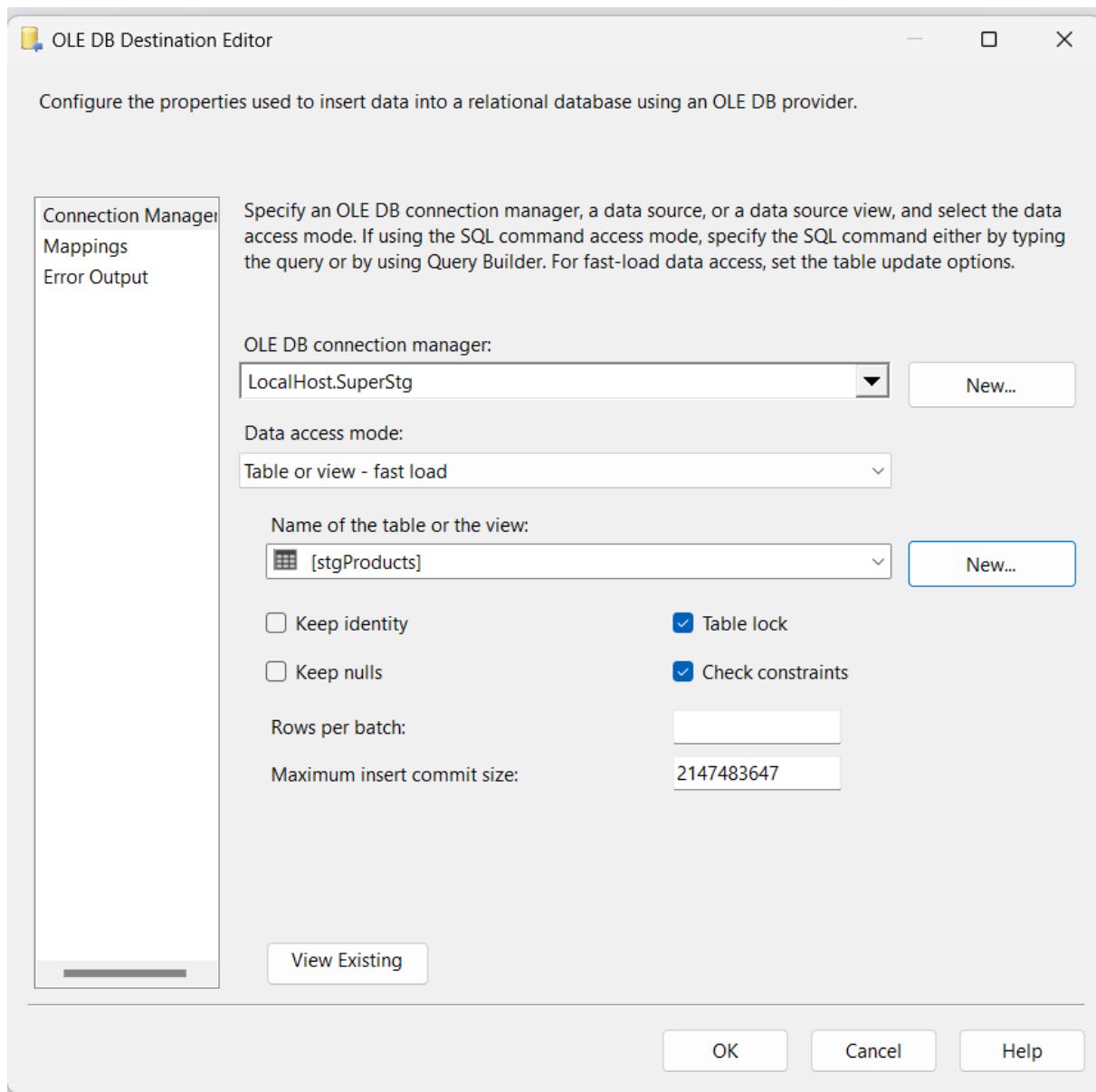
**SRC – SuperStore\_Product** chứa dữ liệu các Product từ bảng SampleSuperstore

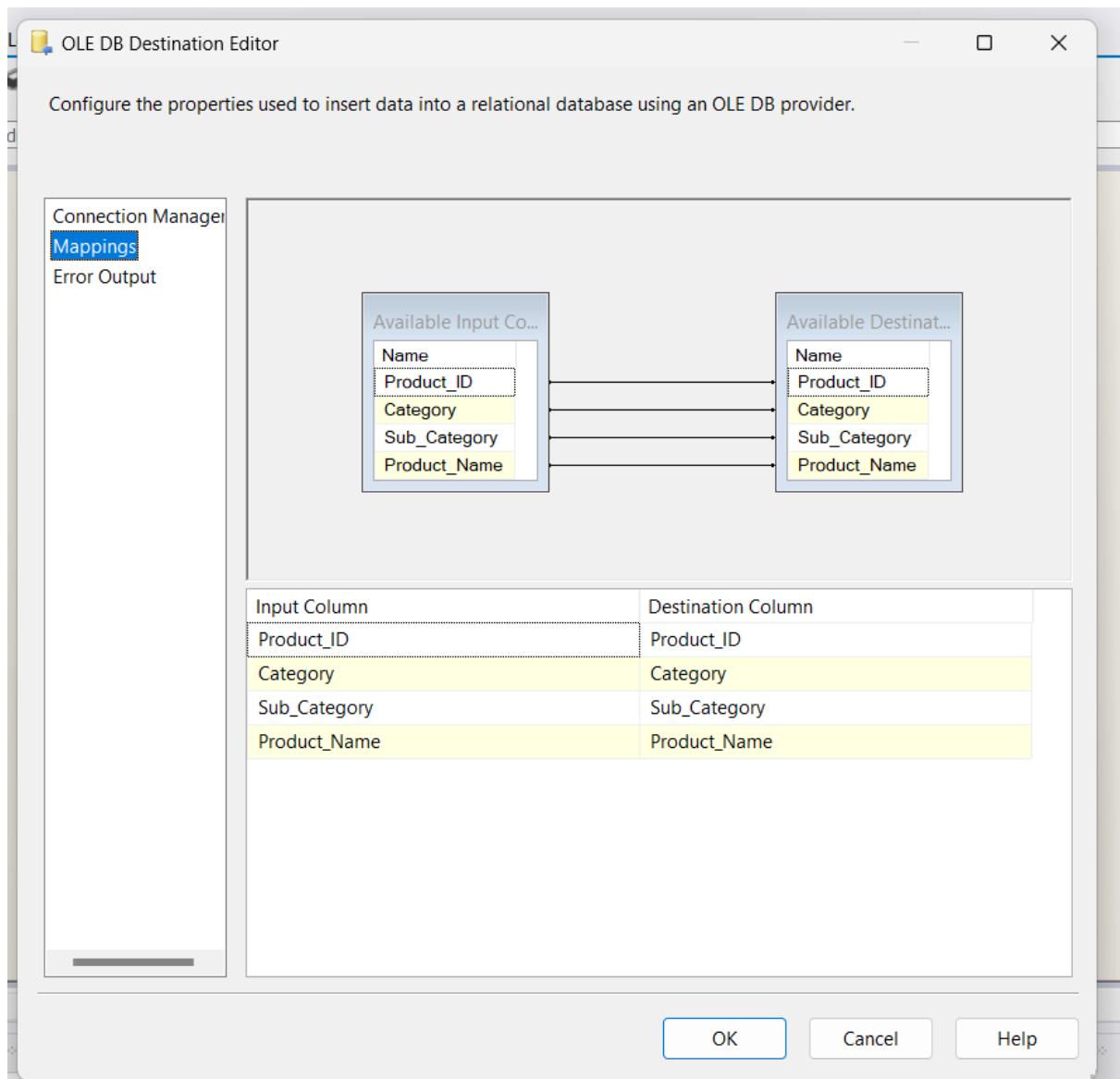


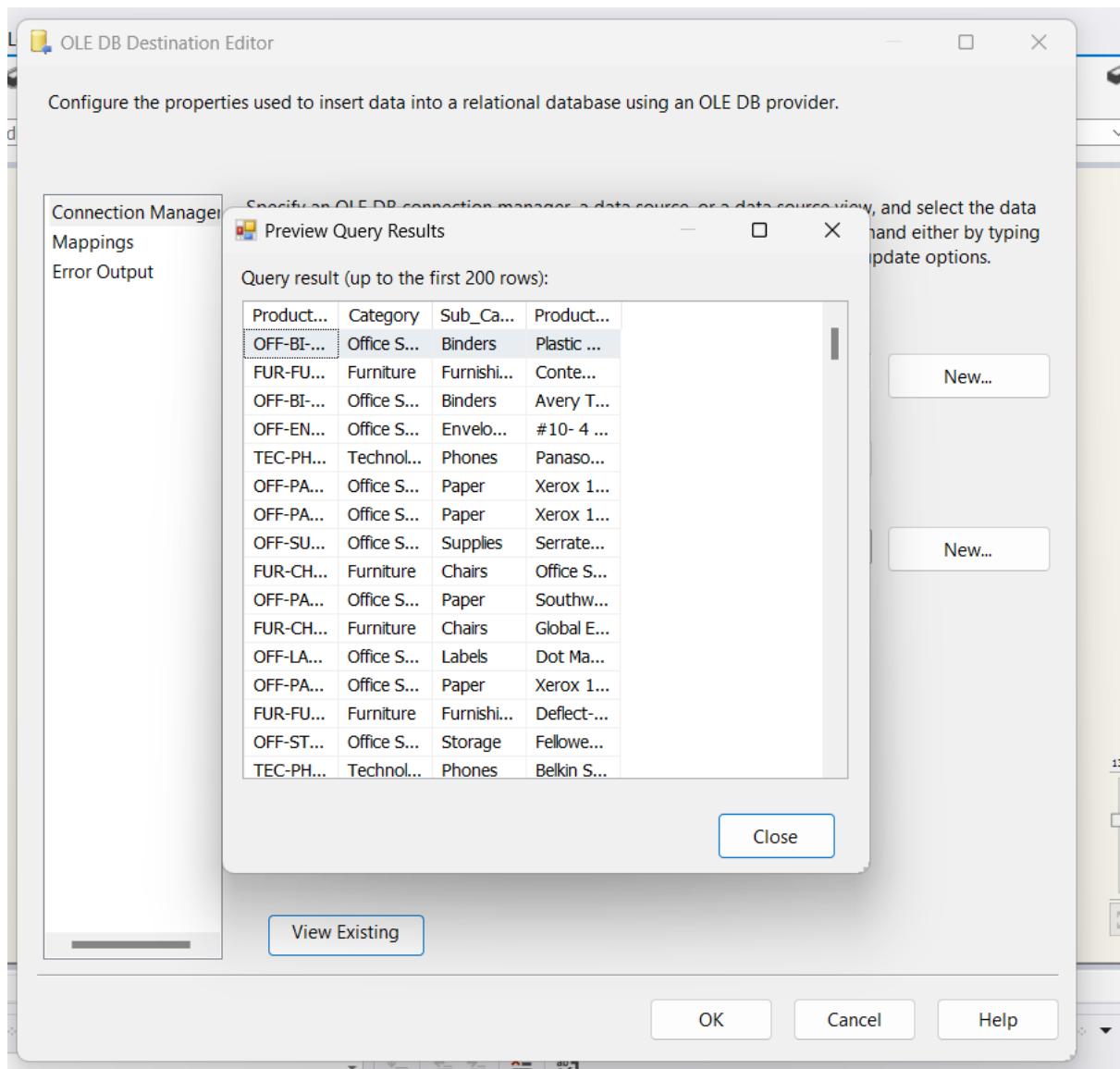


## DST – stgProduct

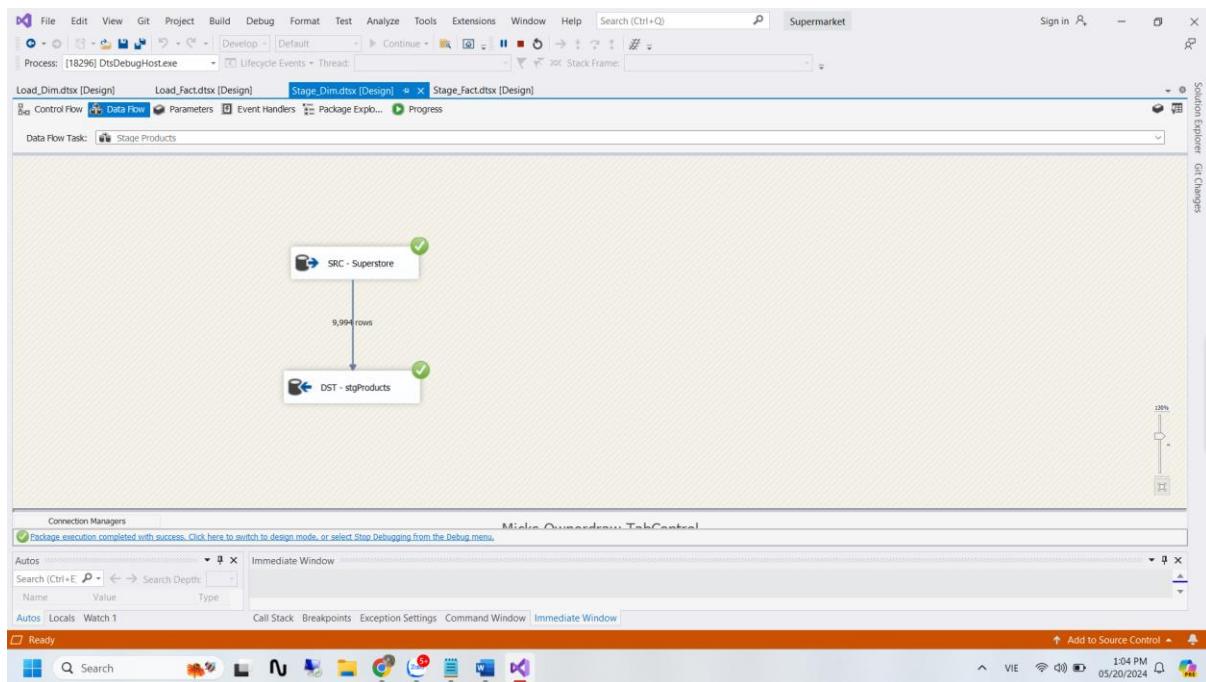




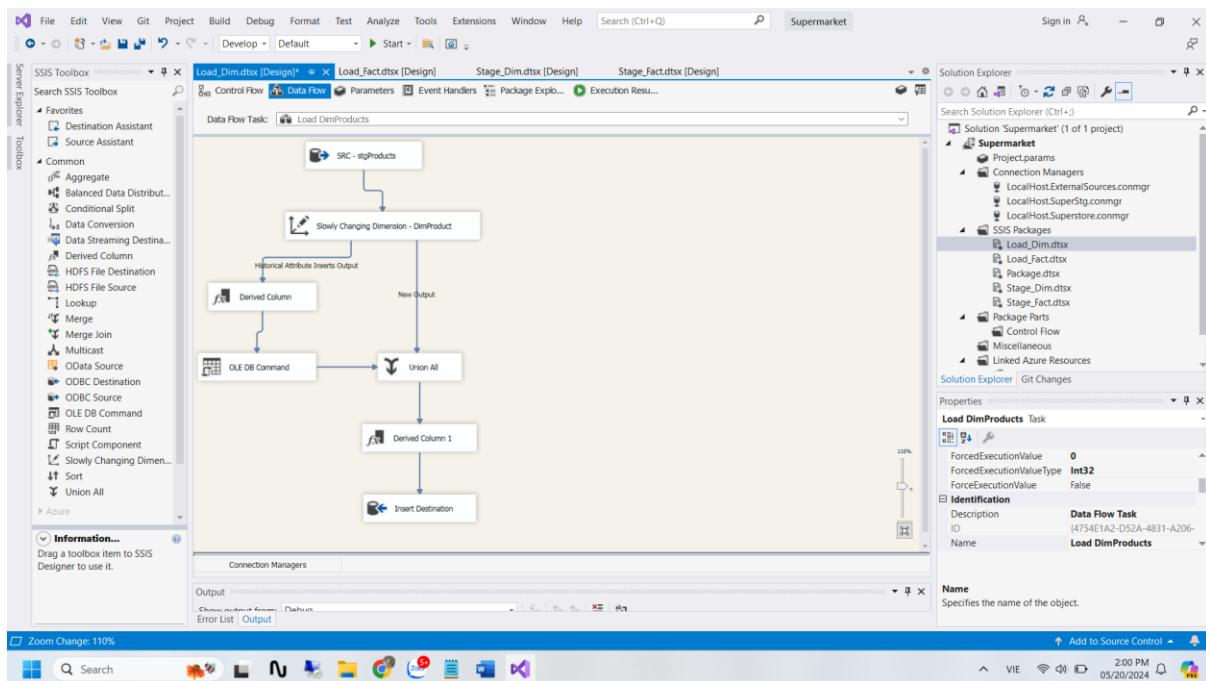




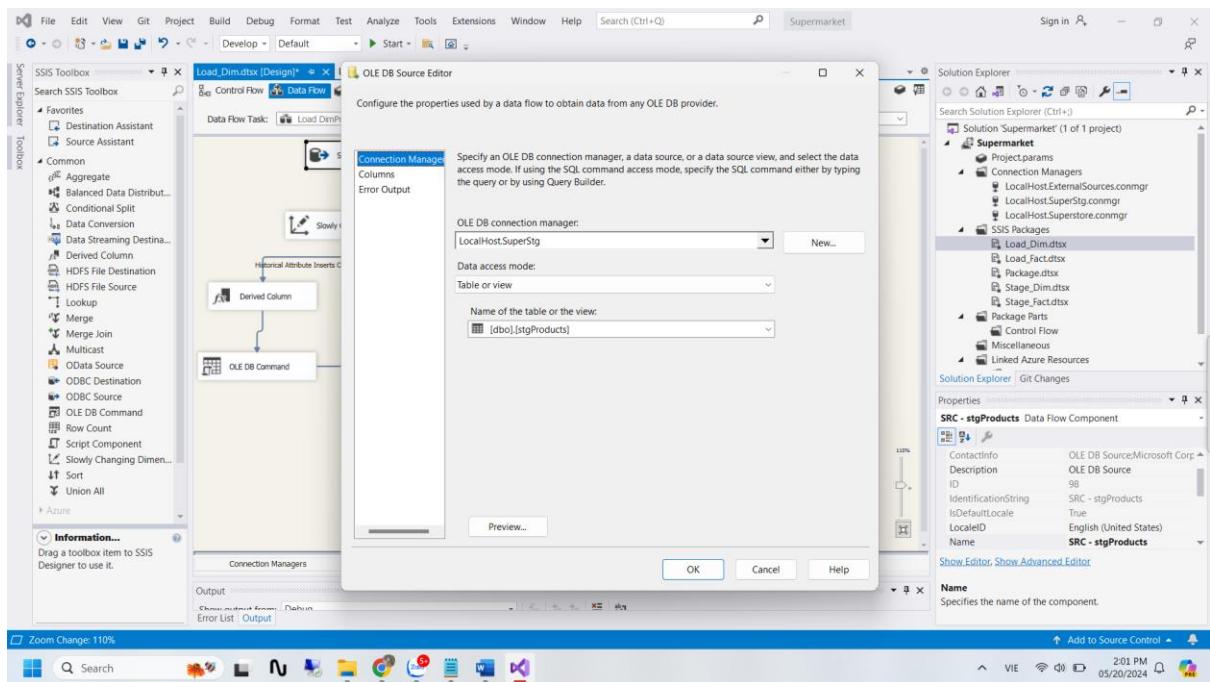
## Khởi chạy



## Load dữ liệu từ stgProduct vào DimProduct

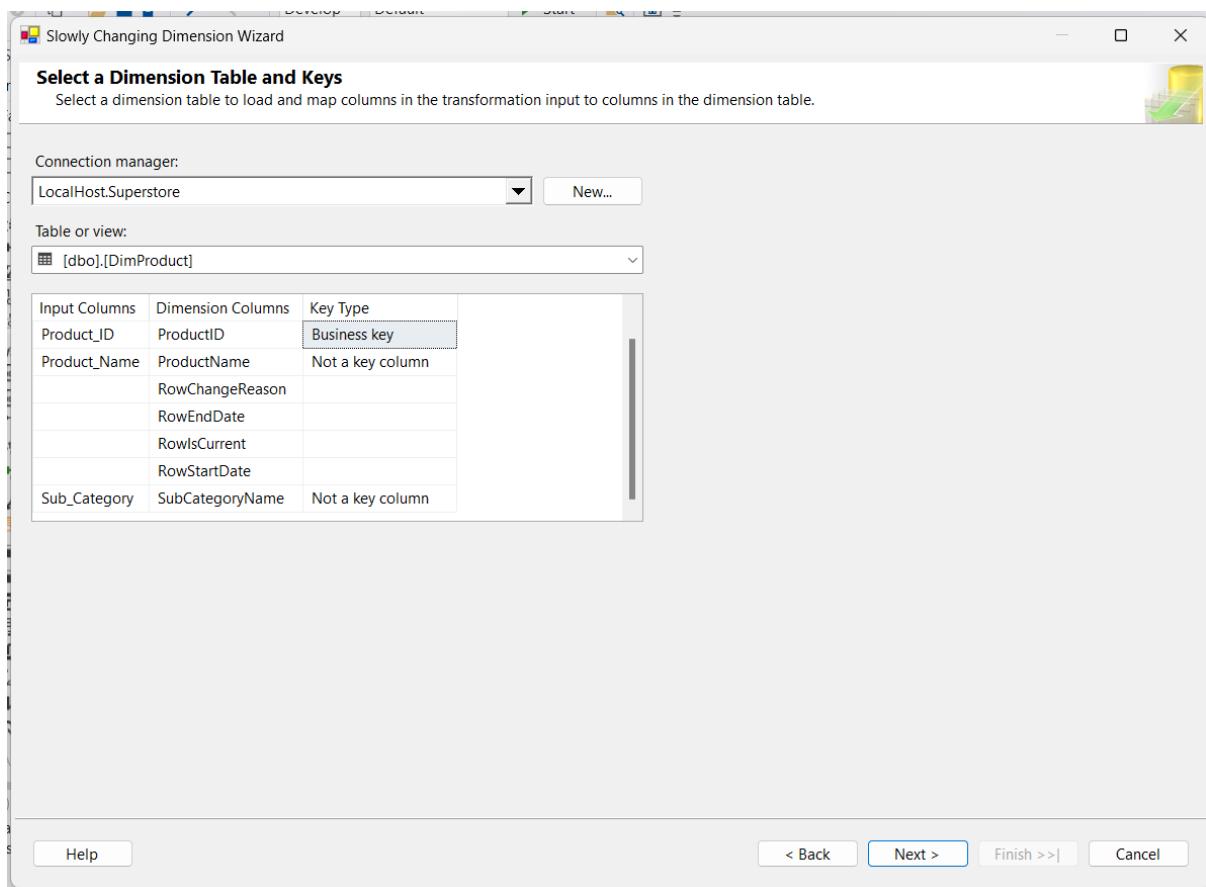


**\_SRC – stgProduct** là dữ liệu từ bảng stgProduct

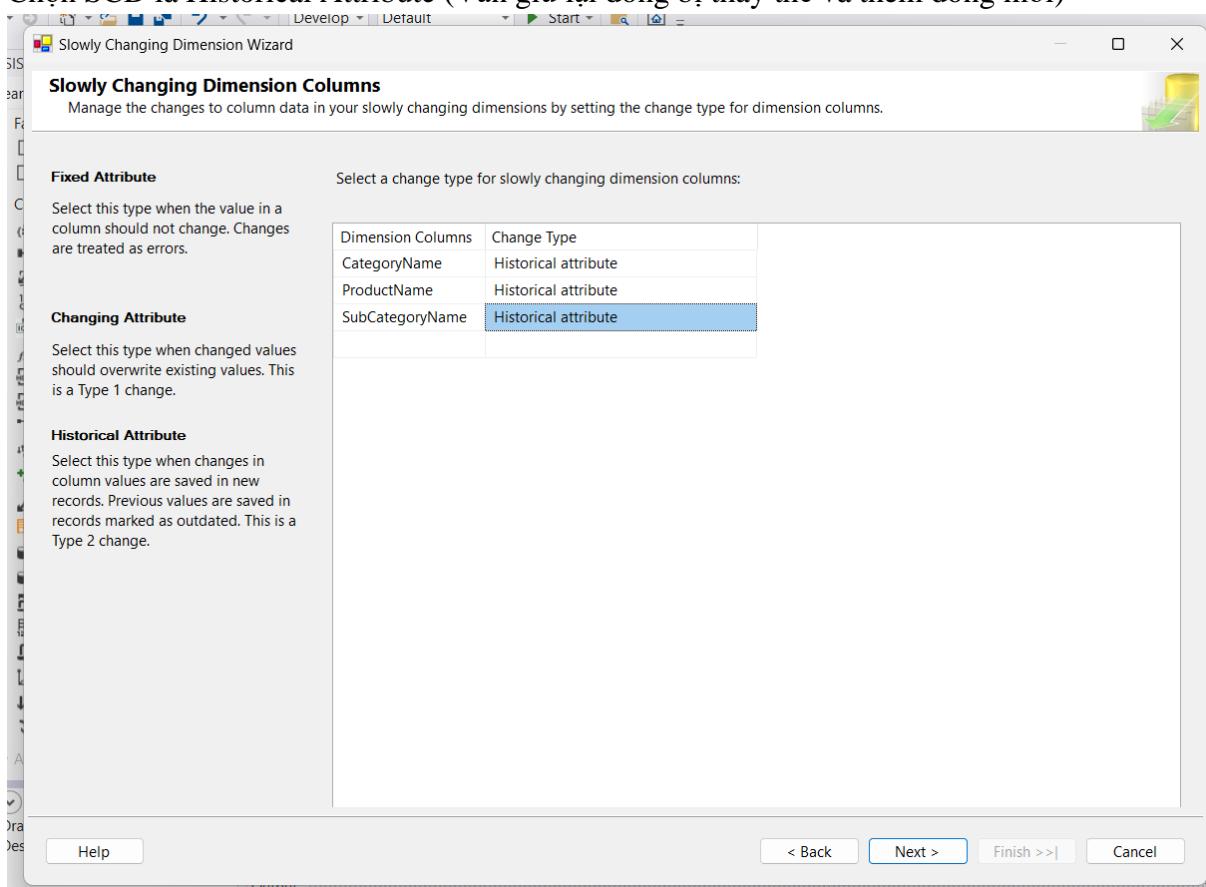


## Slowly Changing Dimension – DimProduct Load từ bảng stgProduct vào DimProduct

Input Columns	Dimension Columns	Key Type
Category	CategoryName	Not a key column
Product_ID	ProductID	Business key
Product_Name	ProductName	Not a key column
	RowChangeReason	
	RowEndDate	
	RowIsCurrent	
	RowStartDate	



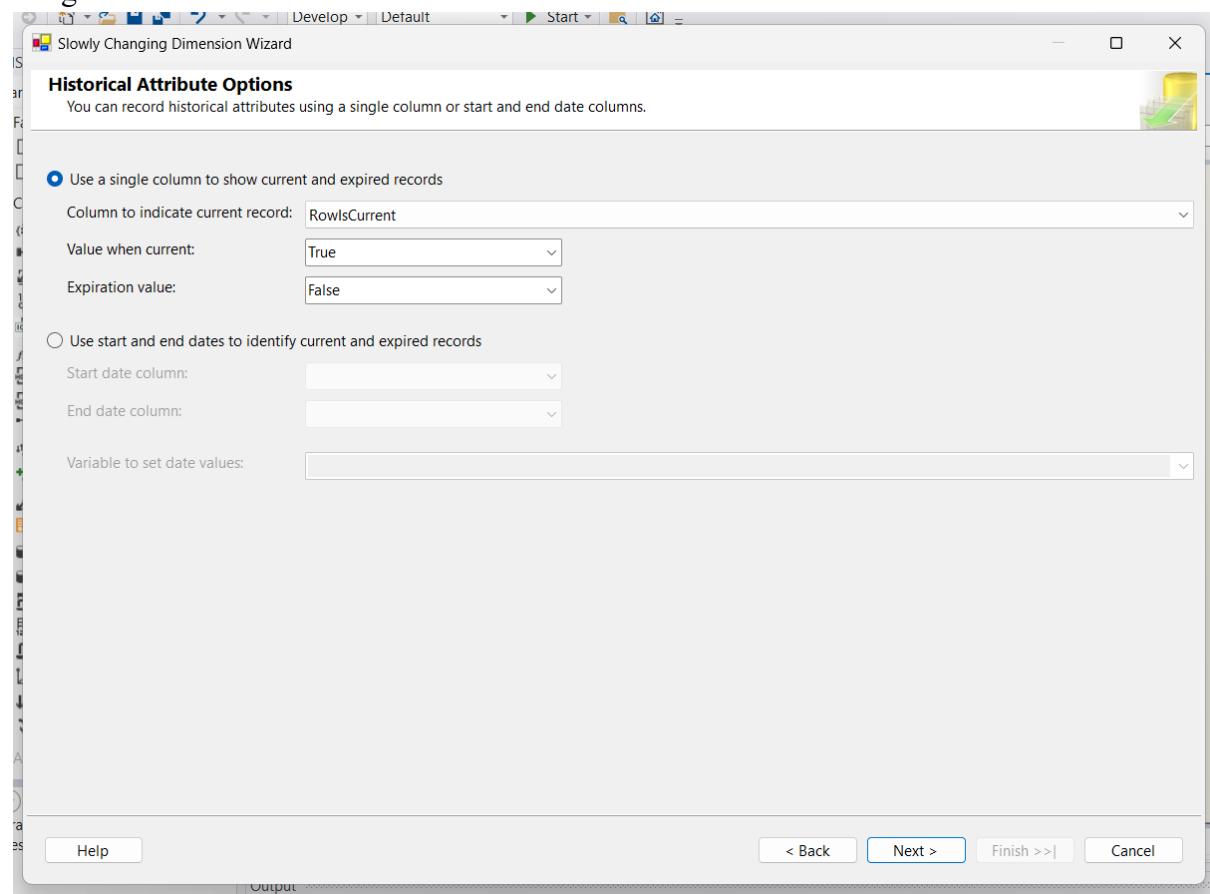
Chọn SCD là Historical Attribute (Vẫn giữ lại dòng bị thay thế và thêm dòng mới)



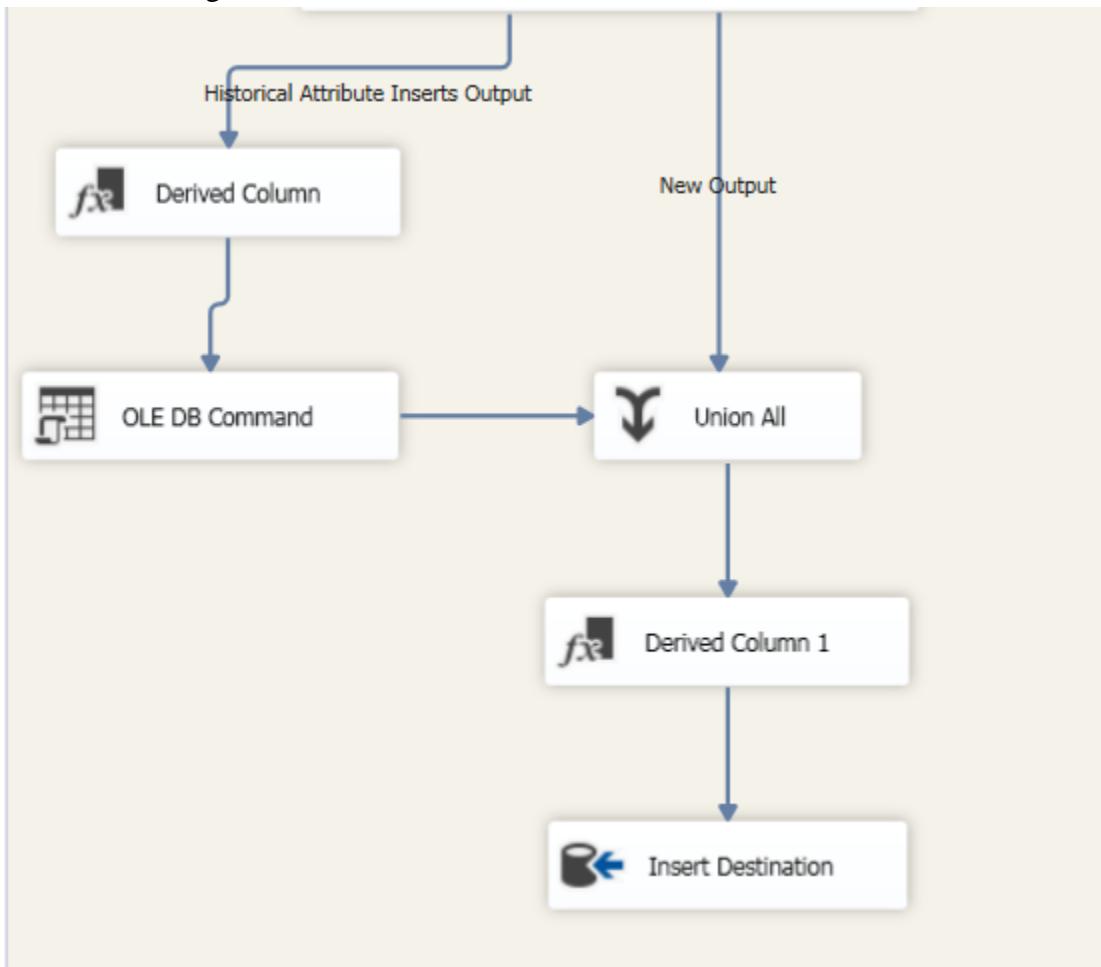
Khi một dòng trong bảng DimProduct cột RowIsCurrent sẽ có giá trị là 1

Ví dụ 1 khách hàng hiện tại đang có hạng tiêu dùng là bạc thì rowiscurrent là 1

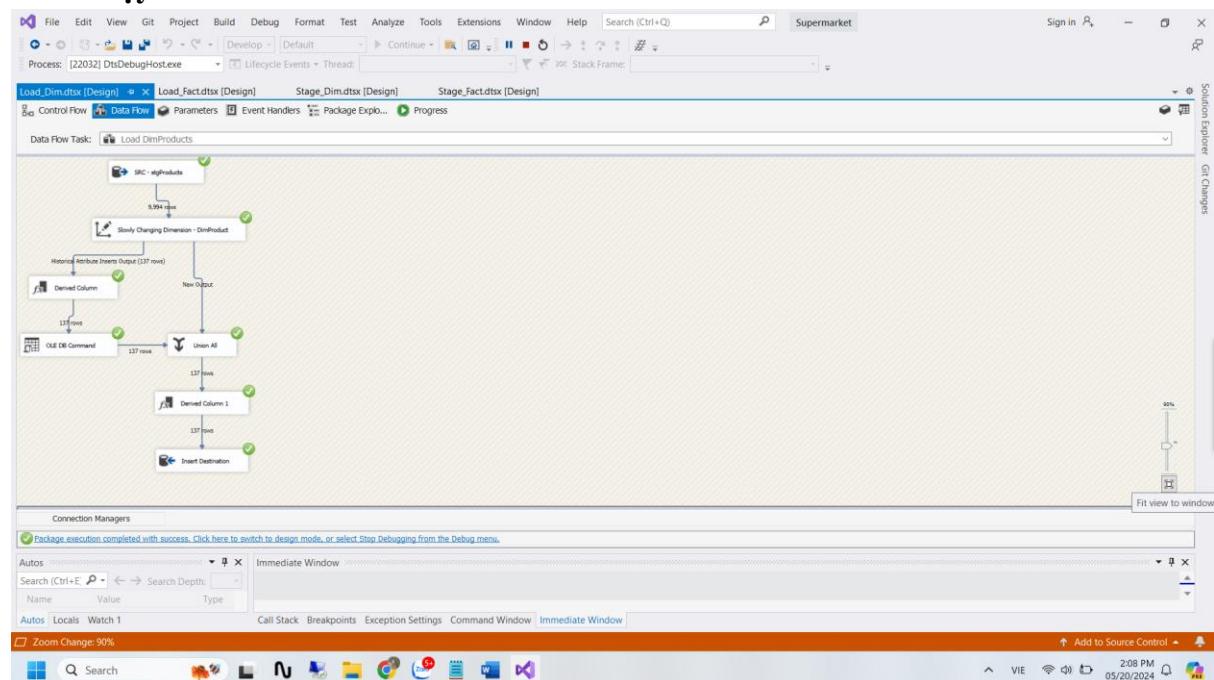
Sau đó được nâng hạng lên vàng thì sẽ thêm 1 dòng mới với mã khách hàng đó và hạng tiêu dùng là vàng với rowiscurrent là 1 và sẽ sửa rowiscurrent của dòng cũ thành 0 mà không xóa dòng cũ



Sau khi setup hoàn thành Slowly Changing Dimension – DimProduct thì các phần ở dưới sẽ được tạo tự động

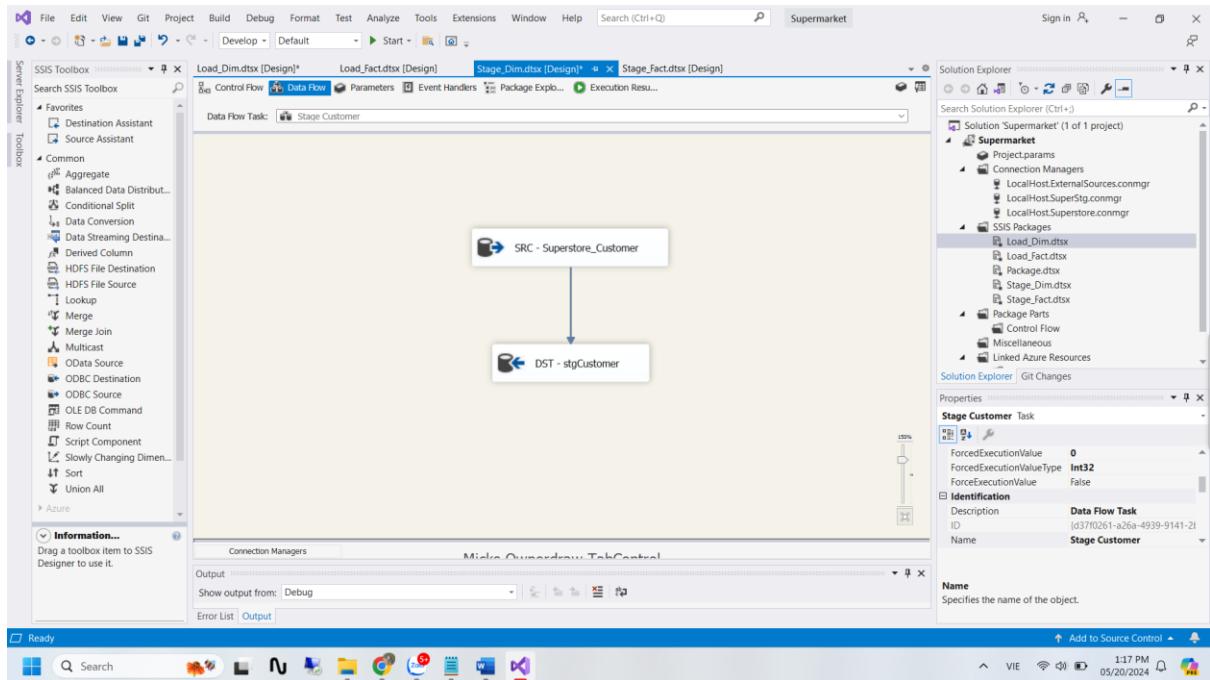


## Khởi chạy

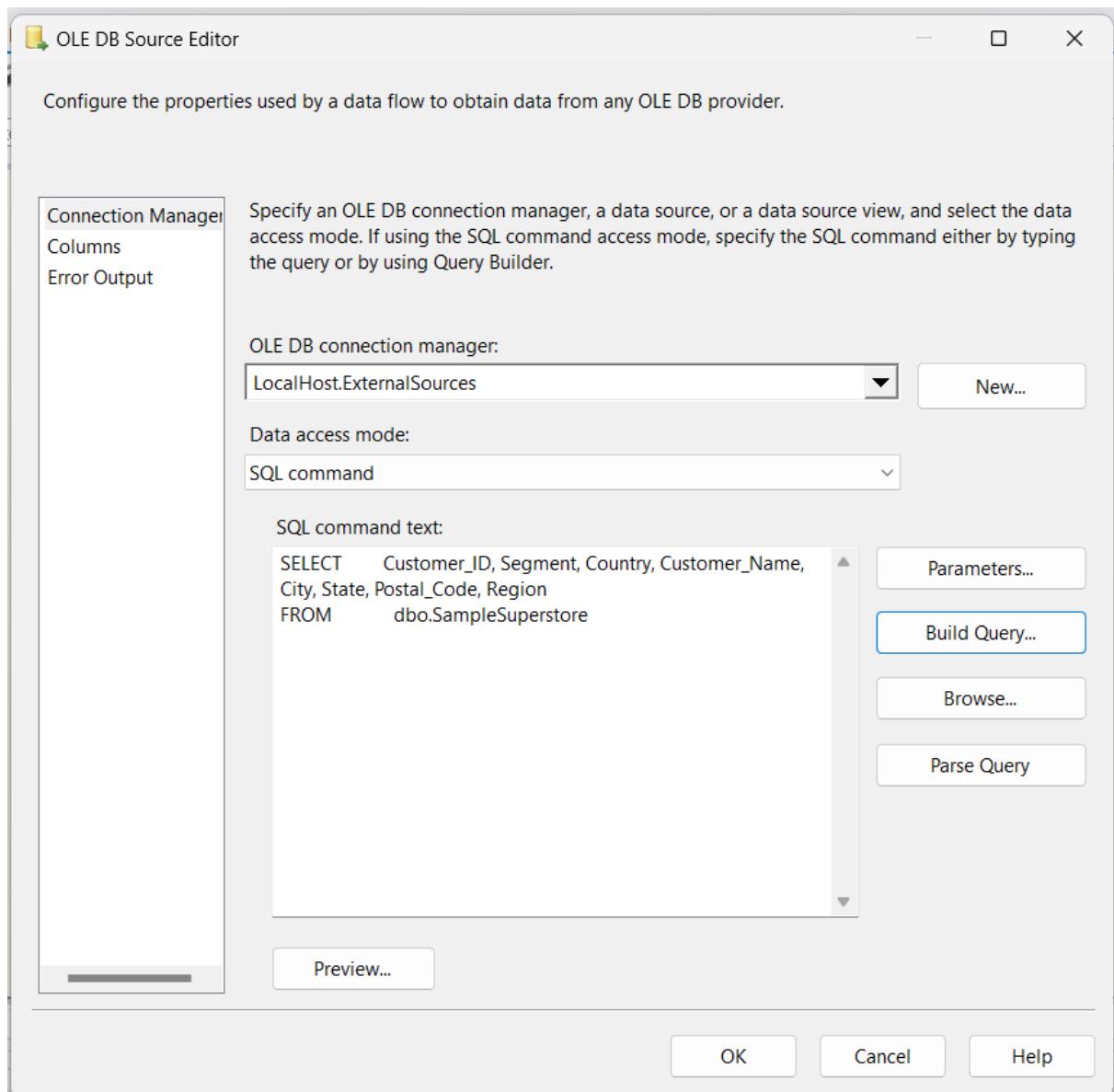


### 3.2.3.Customer Dimension

#### a) Load dữ liệu từ nguồn vào stgCustomer



**Src – SuperStore\_Customer** là dữ liệu Customer từ bảng SampleSuperstore trong database ExternalSources



Query Builder

SampleSuperstore (dbo)

Column	Alias	Table	Output	Sort Type	Sort Order	Filter	Or...	Or...
Customer_ID		SampleSu...	✓					
Segment		SampleSu...	✓					
Country		SampleSu...	✓					
Customer_Na...		SampleSu...	✓					
City		SampleSu...	✓					
State		SampleSu...	✓					
Postal_Code		SampleSu...	✓					
Region		SampleSu...	✓					

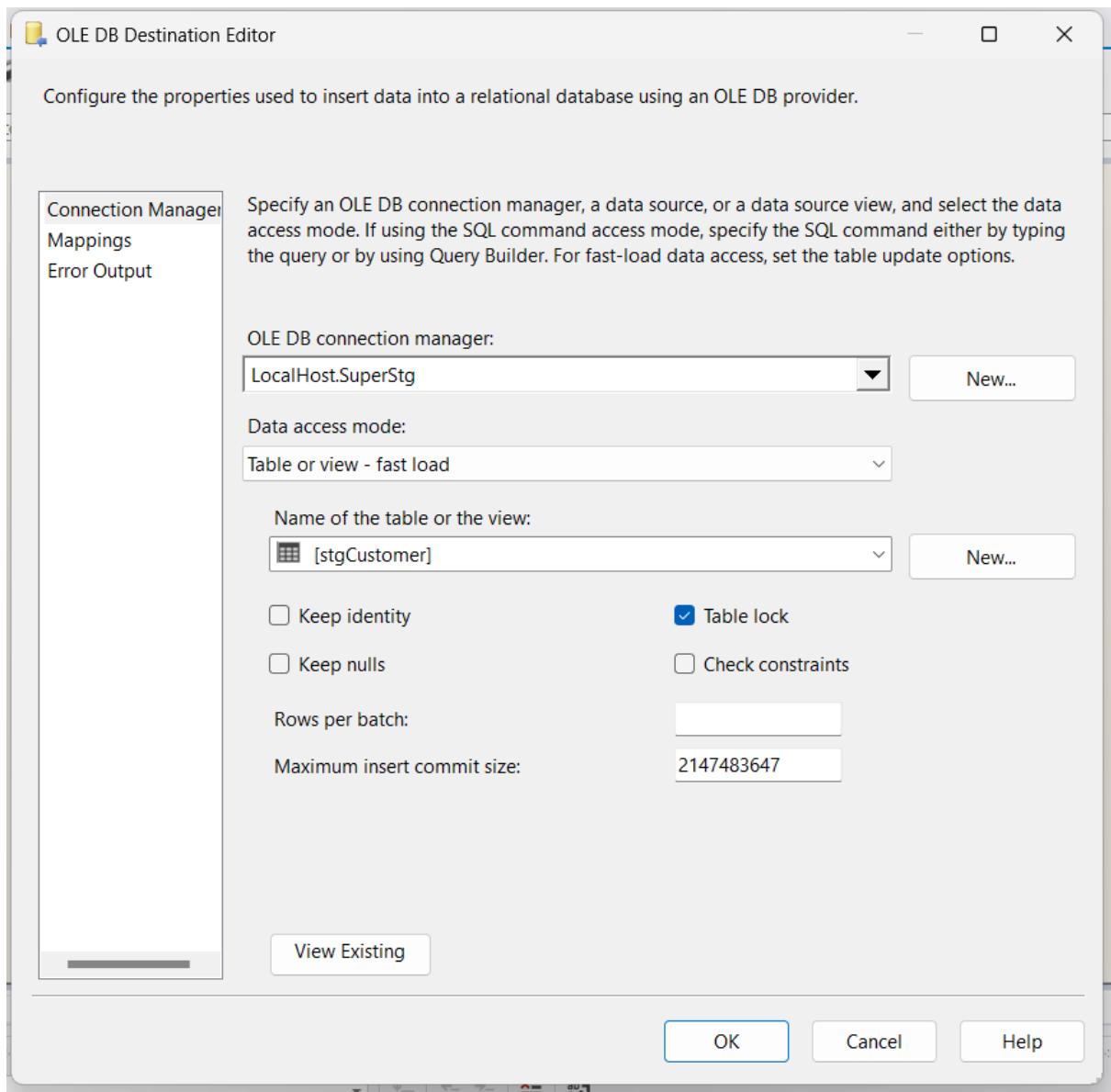
```
SELECT Customer_ID, Segment, Country, Customer_Name, City, State, Postal_Code, Region  
FROM   dbo.SampleSuperstore
```

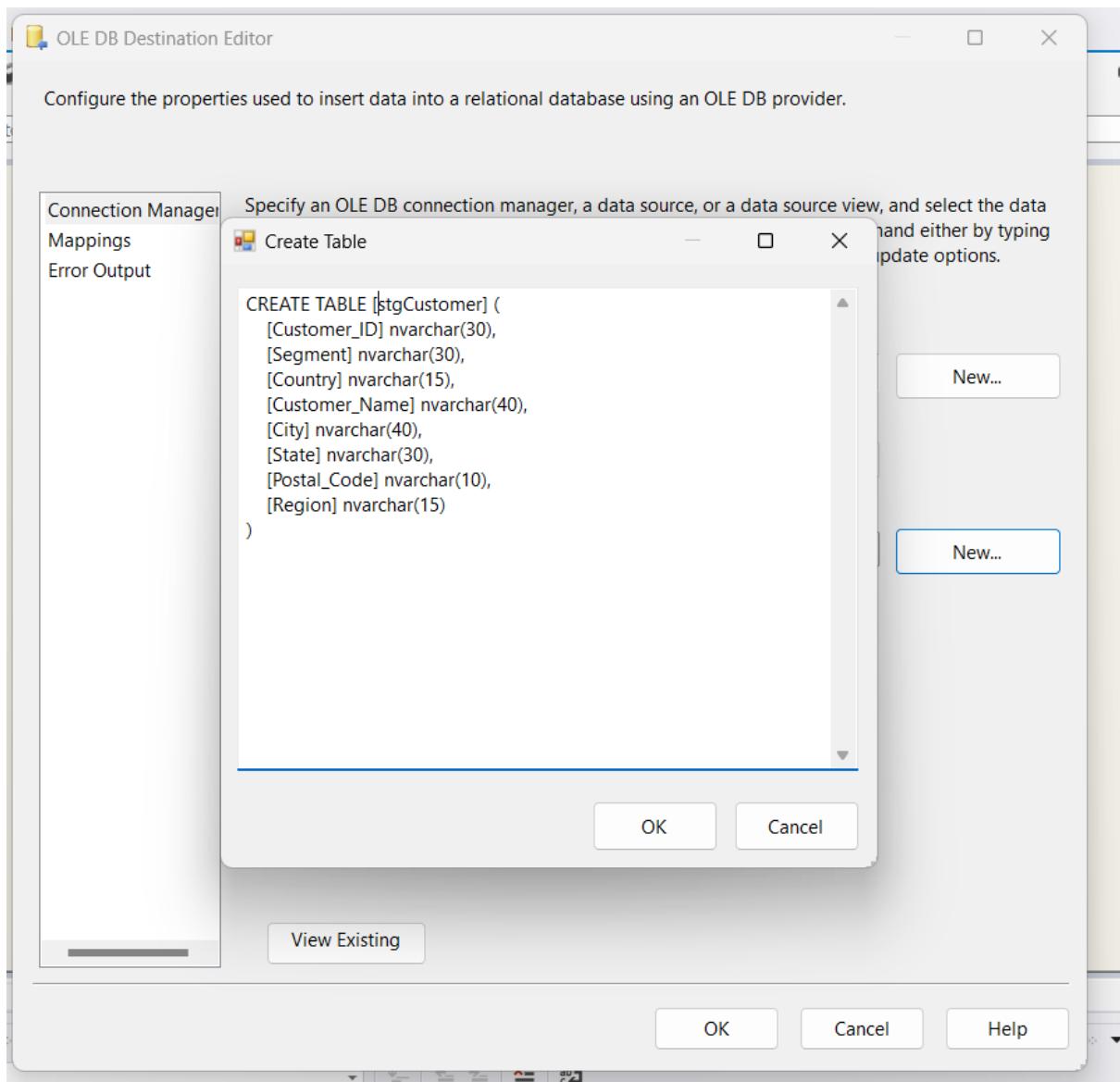
IG

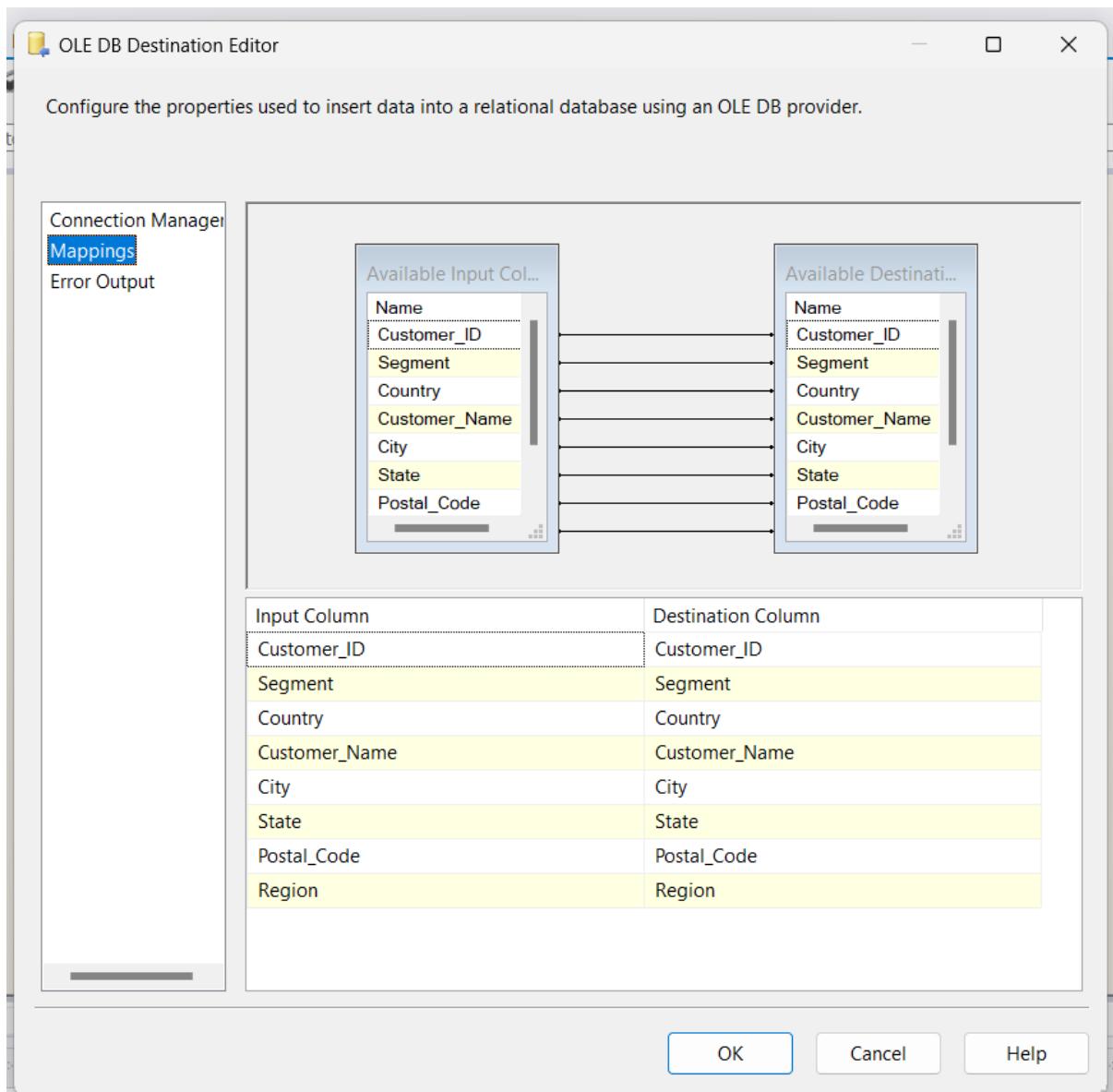
OK Cancel Help

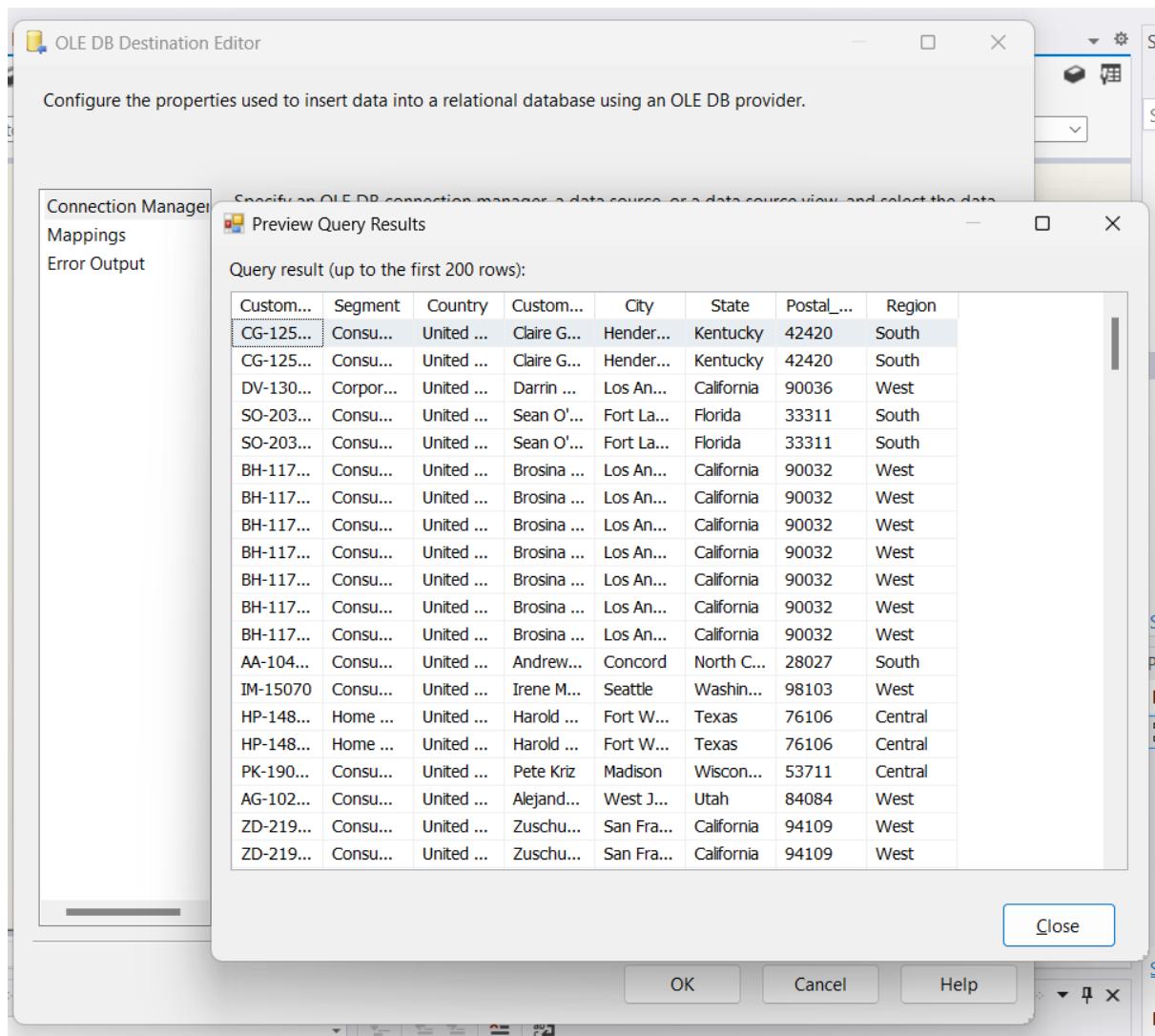
The screenshot shows the 'Query Builder' window with the title 'SampleSuperstore (dbo)'. On the left, a list of columns is displayed with checkboxes next to them. Most checkboxes are checked, except for 'Ship\_Mode' and 'Product ID'. Below this is a table view where each column is mapped to an alias and a table name ('SampleSu...'). The table has columns for Column, Alias, Table, Output, Sort Type, Sort Order, Filter, Or..., and Or... (partially visible). The 'Customer\_ID' row is currently selected. At the bottom, a SQL query is generated: 'SELECT Customer\_ID, Segment, Country, Customer\_Name, City, State, Postal\_Code, Region FROM dbo.SampleSuperstore'. A progress bar indicates the query is at 0 of 0. At the bottom right are 'OK', 'Cancel', and 'Help' buttons.

## DST – stgCustomer

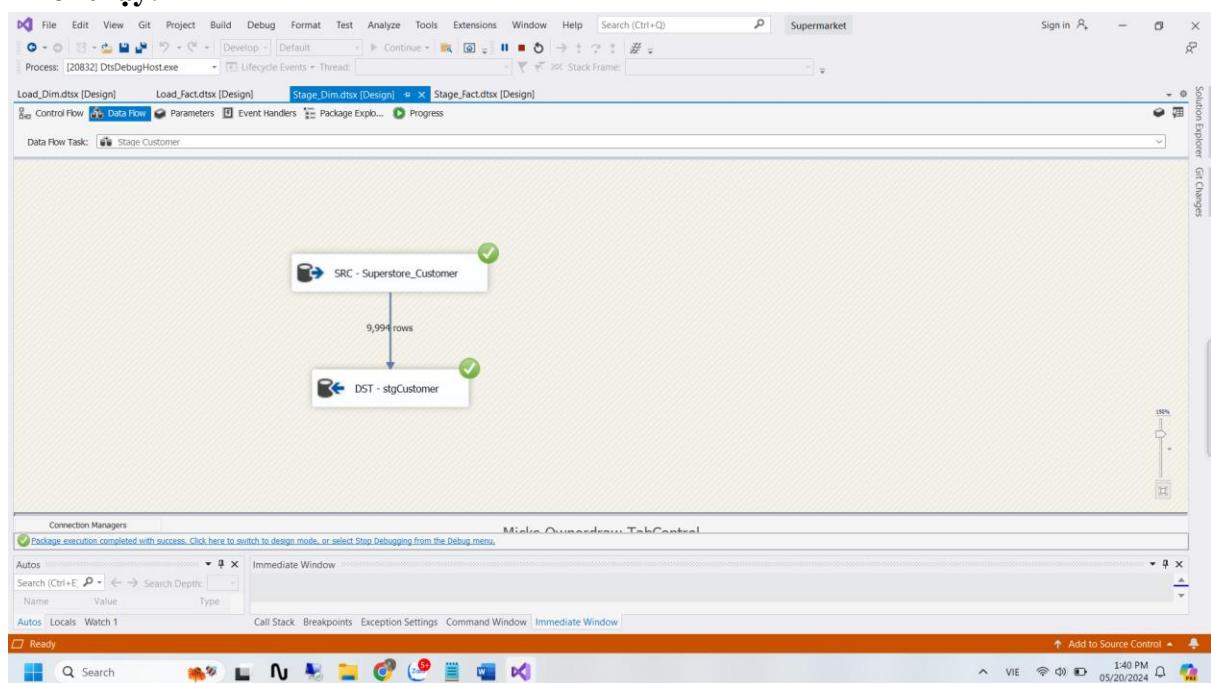




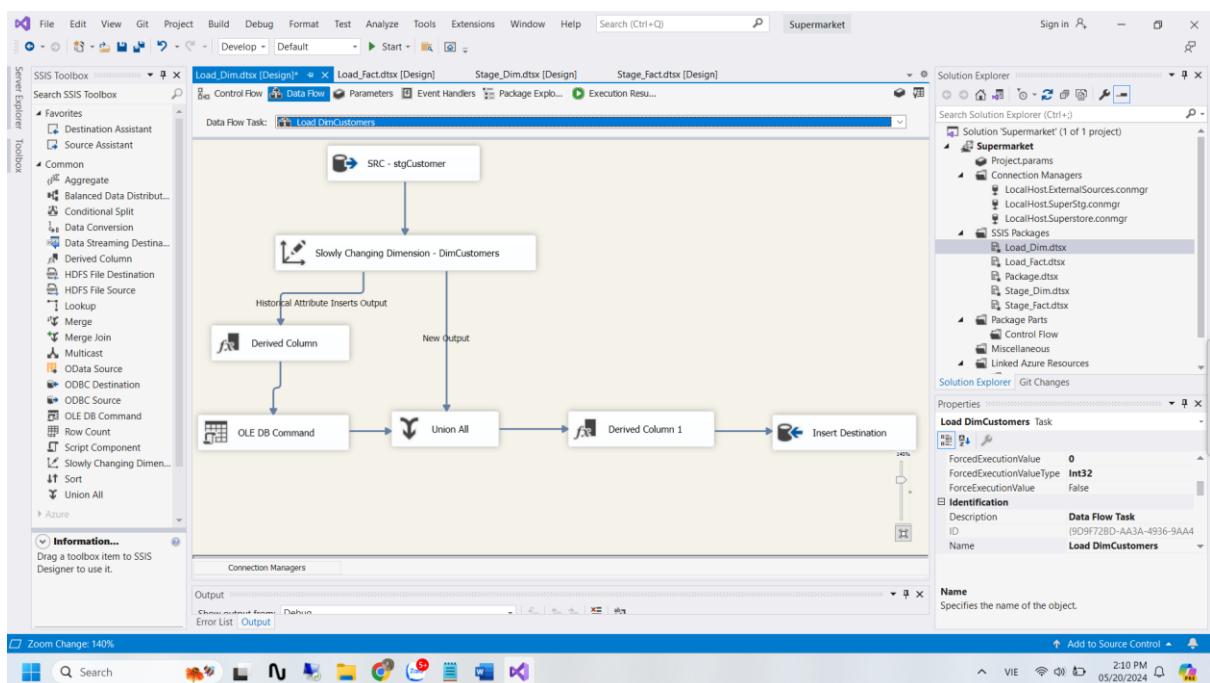




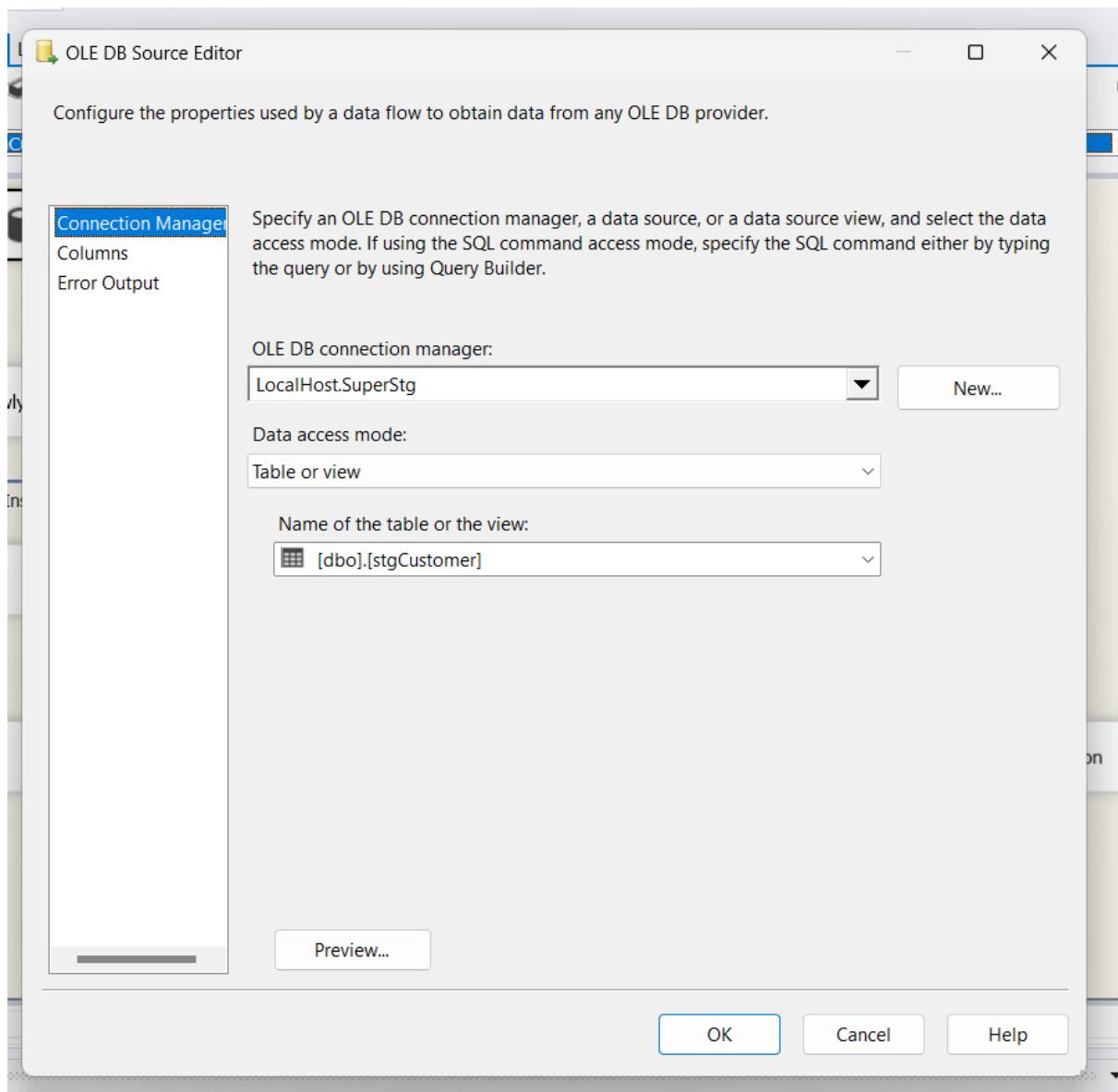
## Khởi chạy:



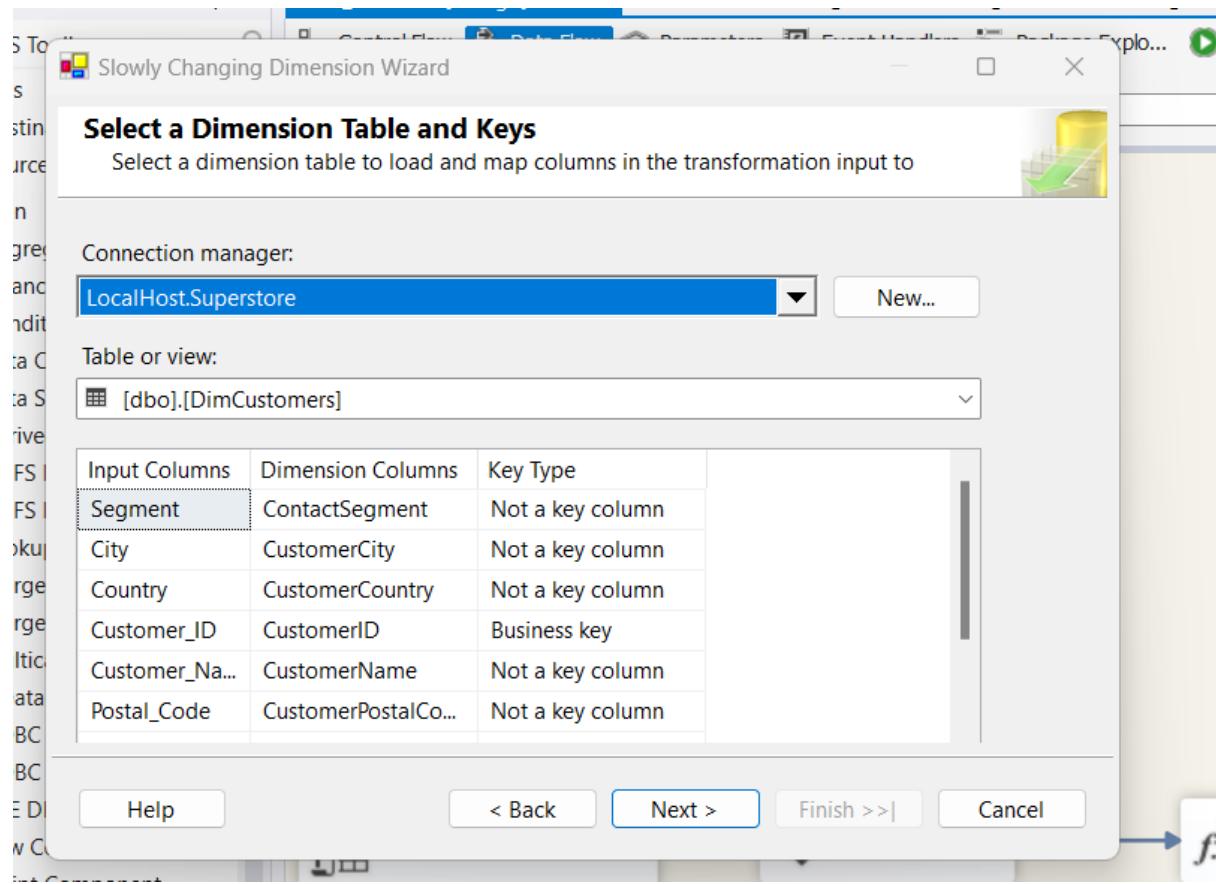
## b) Load dữ liệu từ stgCustomer vào bảng DimCustomer



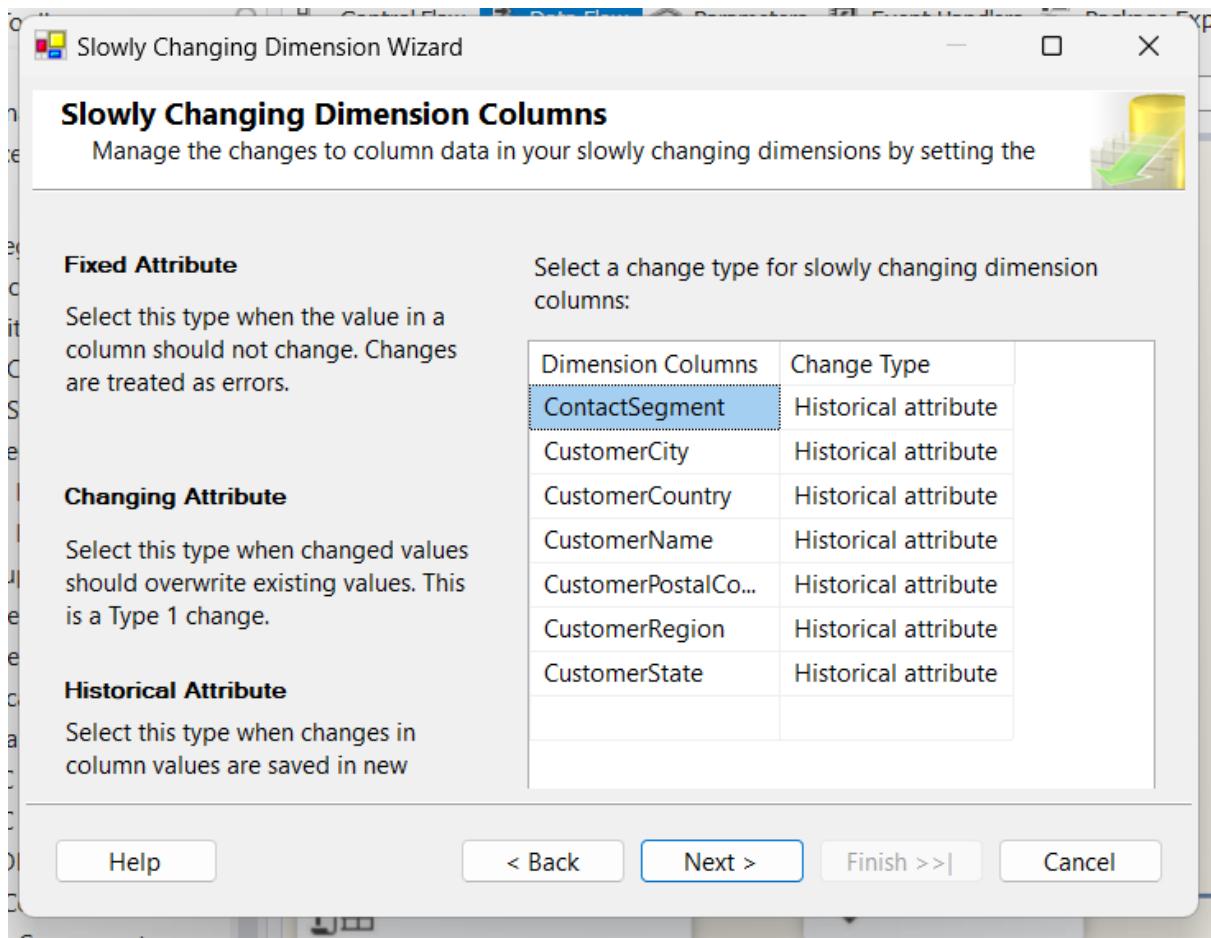
**\_SRC - stgCustomer** chứa dữ liệu bảng stgCustomer



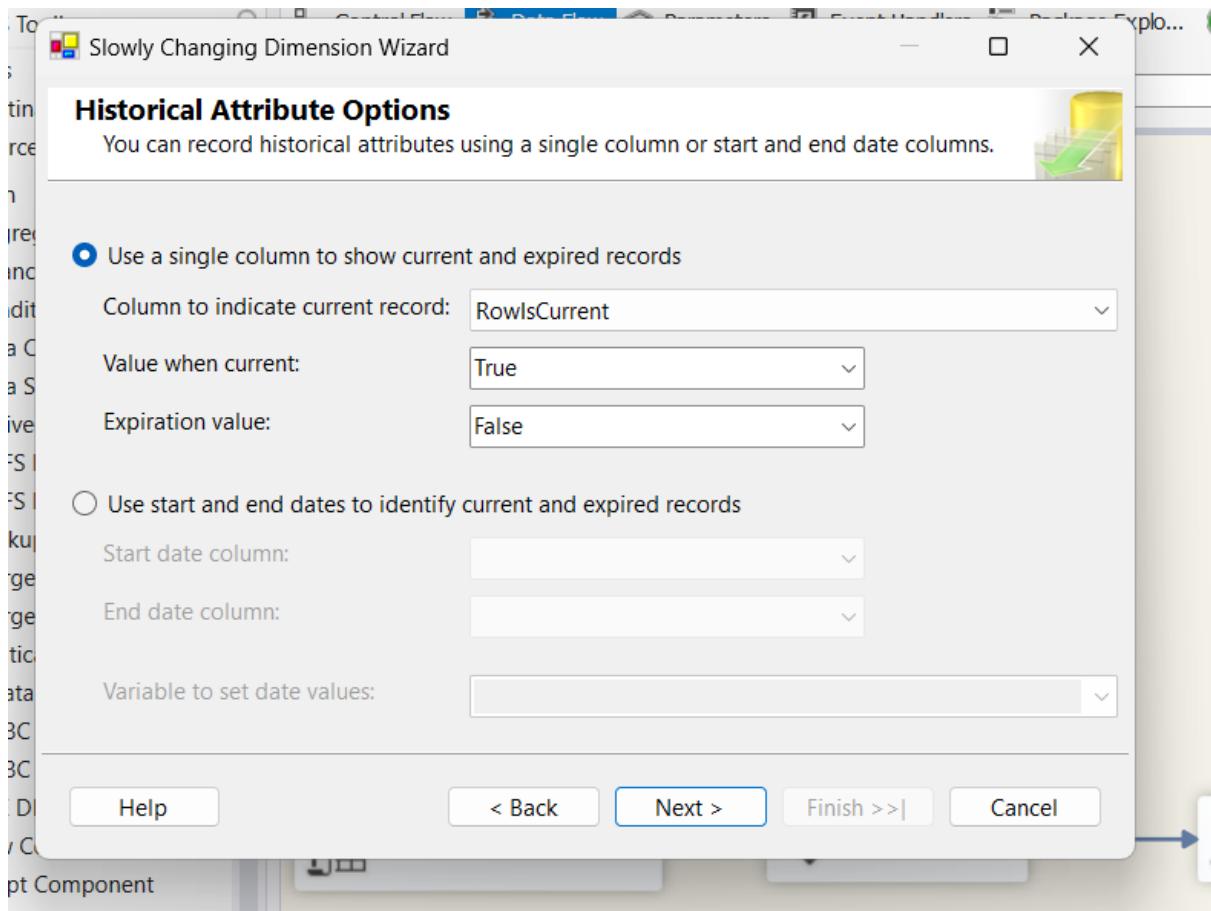
Slowly Changing Dimension – DimCustomers load dữ liệu vào bảng DimCustomer database SuperStore



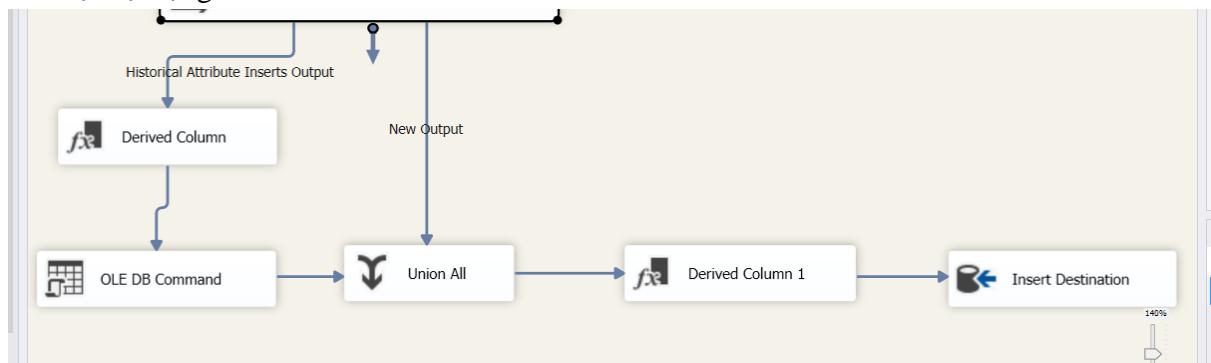
SCD chọn Historical attribute



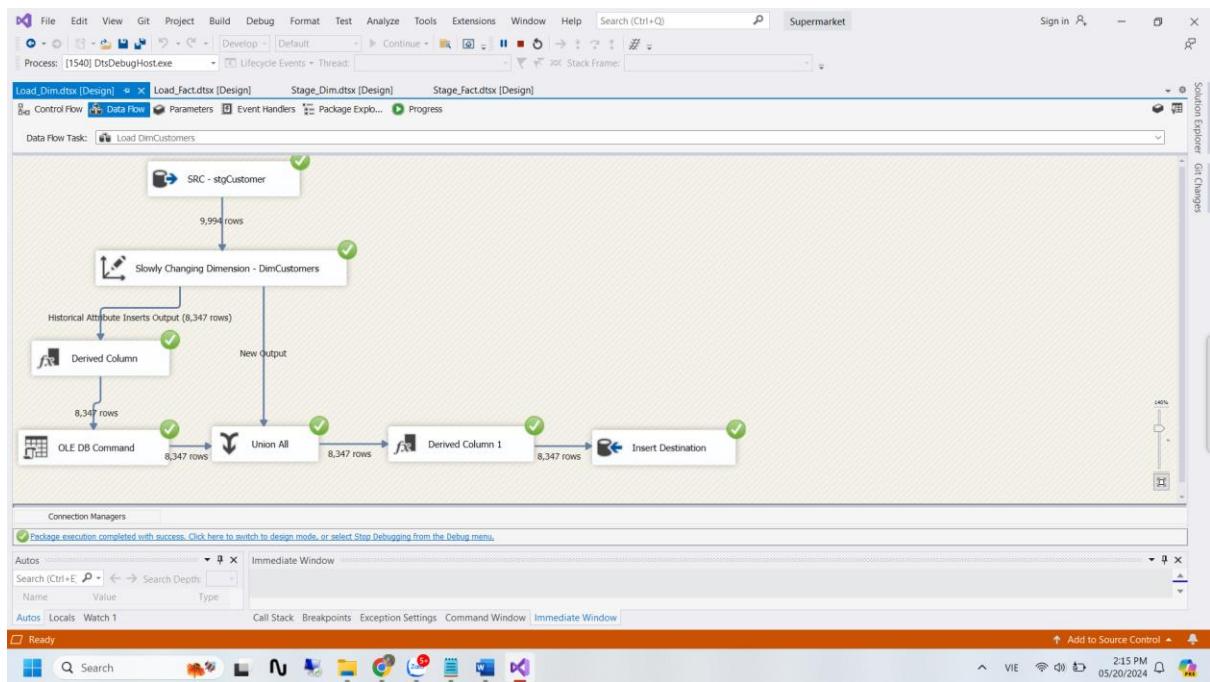
Historical attribute options chọn RowIsCurrent



Sau khi setup hoàn thành Slowly Changing Dimension – DimCustomers thì các phần ở dưới sẽ được tự động

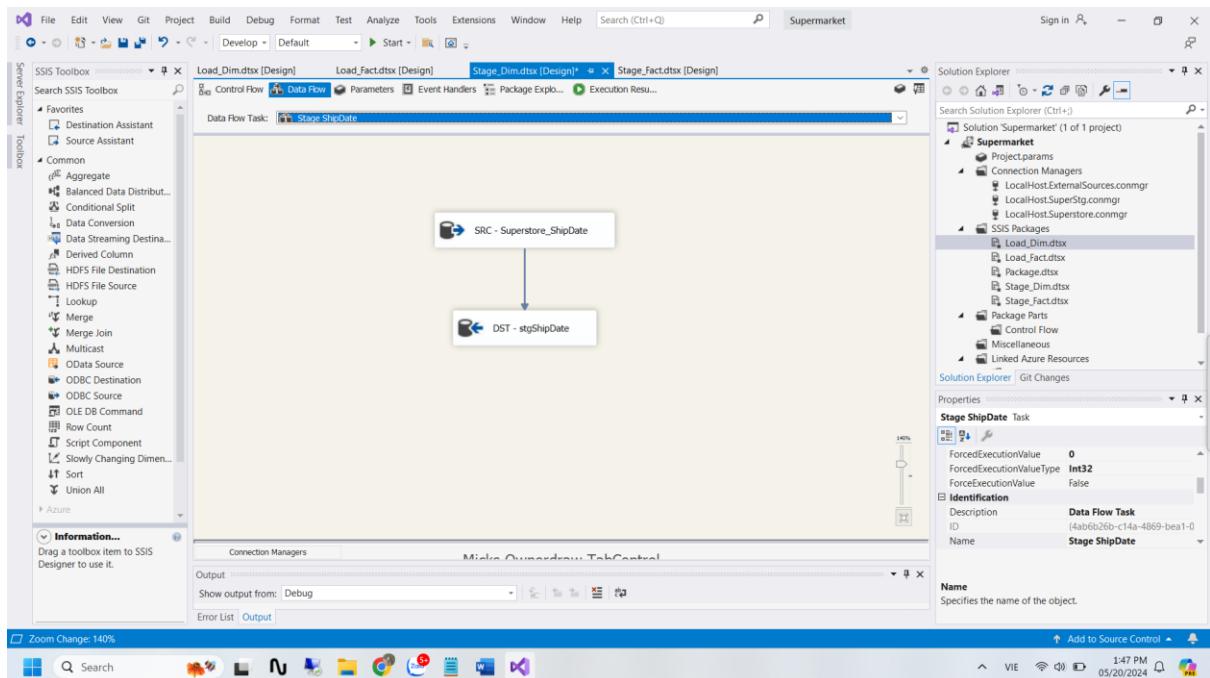


**Khởi chạy**

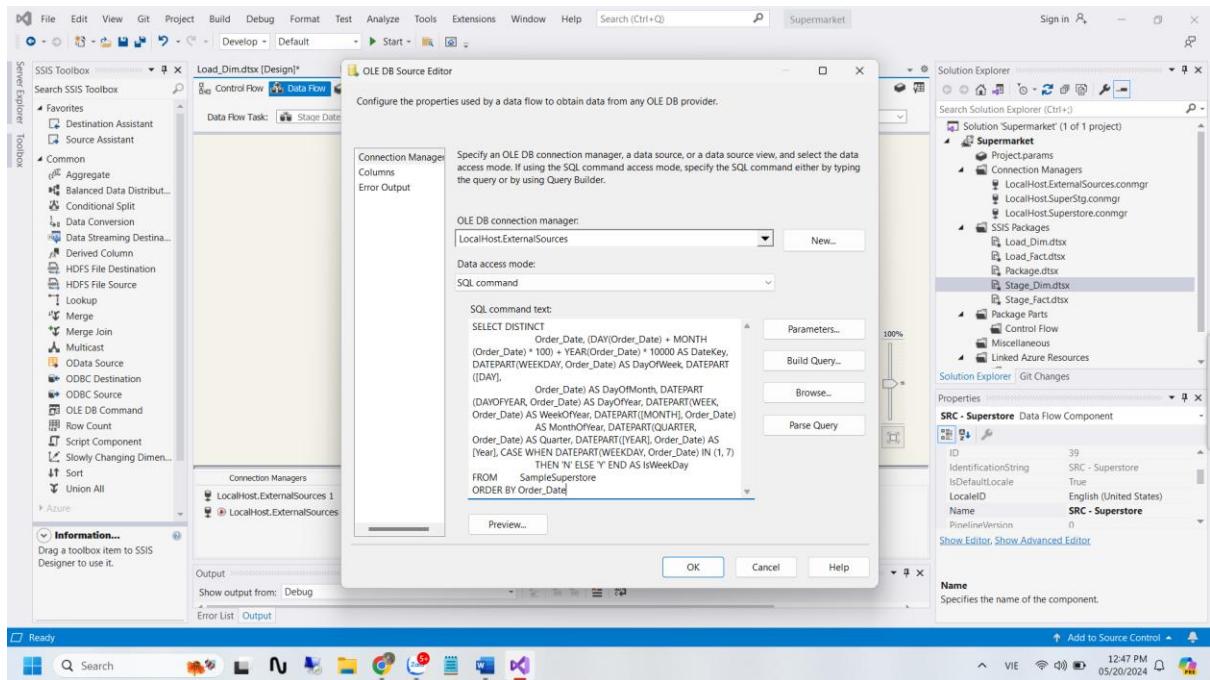


### 3.2.4. ShipDate Dimension

#### a) Load từ nguồn vào bảng stgOrder



**\_SRC – SuperStore\_ShipOrder** là tập hợp các ngày có thực hiện giao hàng



## Script :

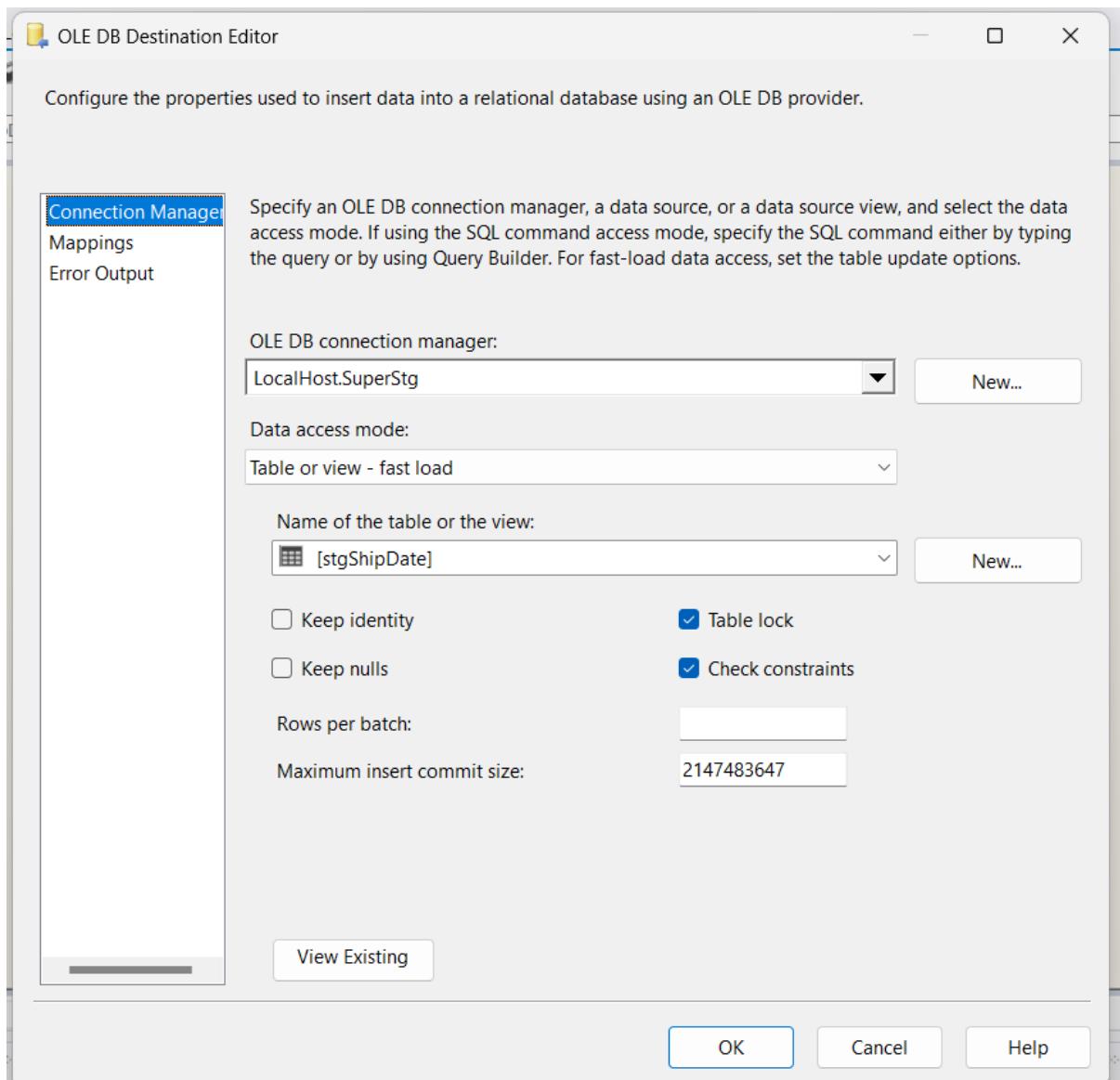
```

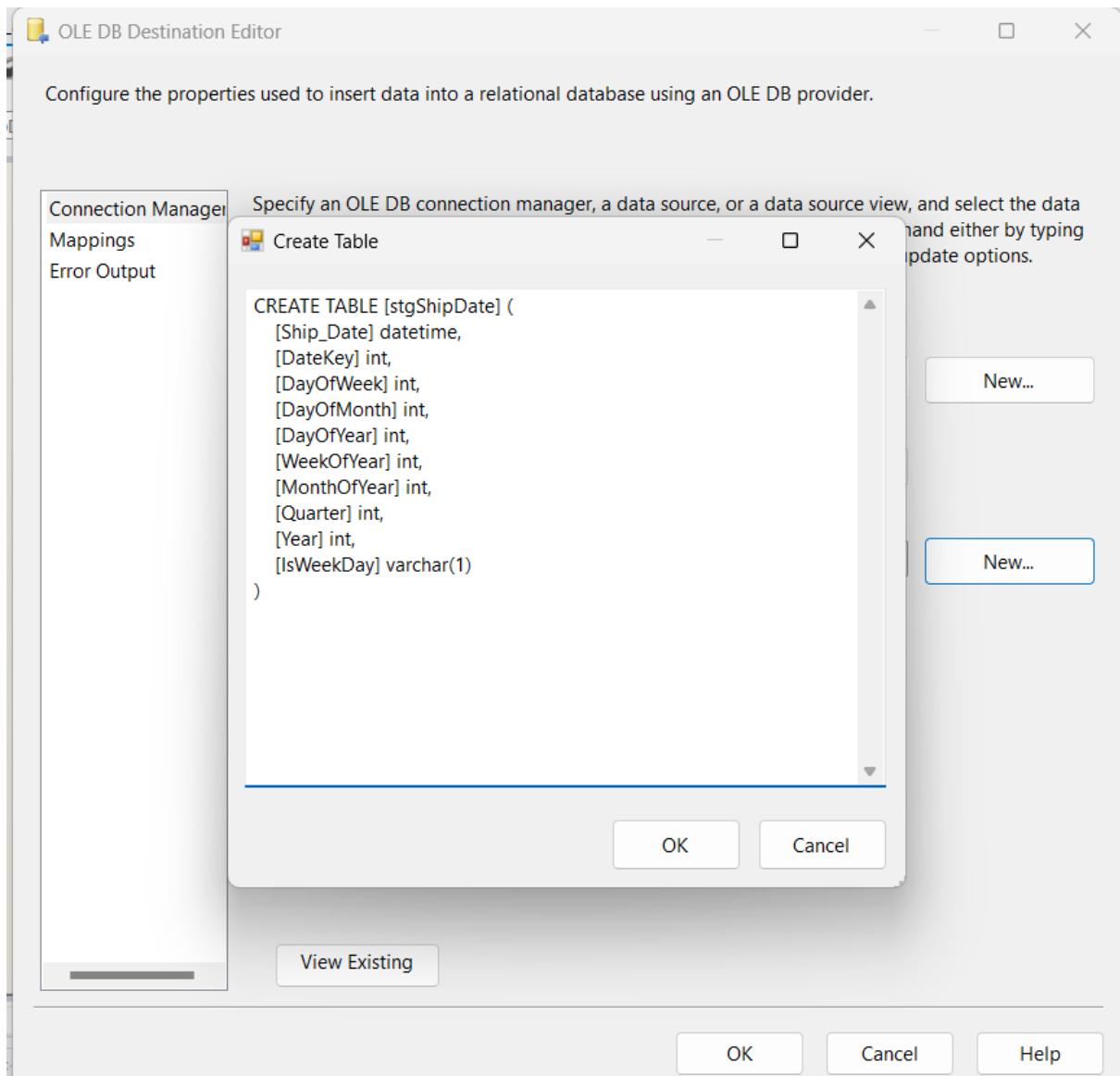
SELECT DISTINCT Ship_Date,
    (Day("Ship_Date") + MONTH("Ship_Date") * 100 + YEAR("Ship_Date") * 10000 ) as DateKey,
    DATEPART(WEEKDAY, Ship_Date) AS DayOfWeek,
    DATEPART(DAY, Ship_Date) AS DayOfMonth,
    DATEPART(DAYOFYEAR, Ship_Date) AS DayOfYear,
    DATEPART(WEEK, Ship_Date) AS WeekOfYear,
    DATEPART(MONTH, Ship_Date) AS MonthOfYear,
    DATEPART(QUARTER, Ship_Date) AS Quarter,
    DATEPART(YEAR, Ship_Date) AS "Year",
    CASE WHEN DATEPART(WEEKDAY, Ship_Date) IN (1, 7) THEN 'N' ELSE 'Y' END AS IsWeekDay
FROM dbo.SampleSuperstore
Order by Ship_Date

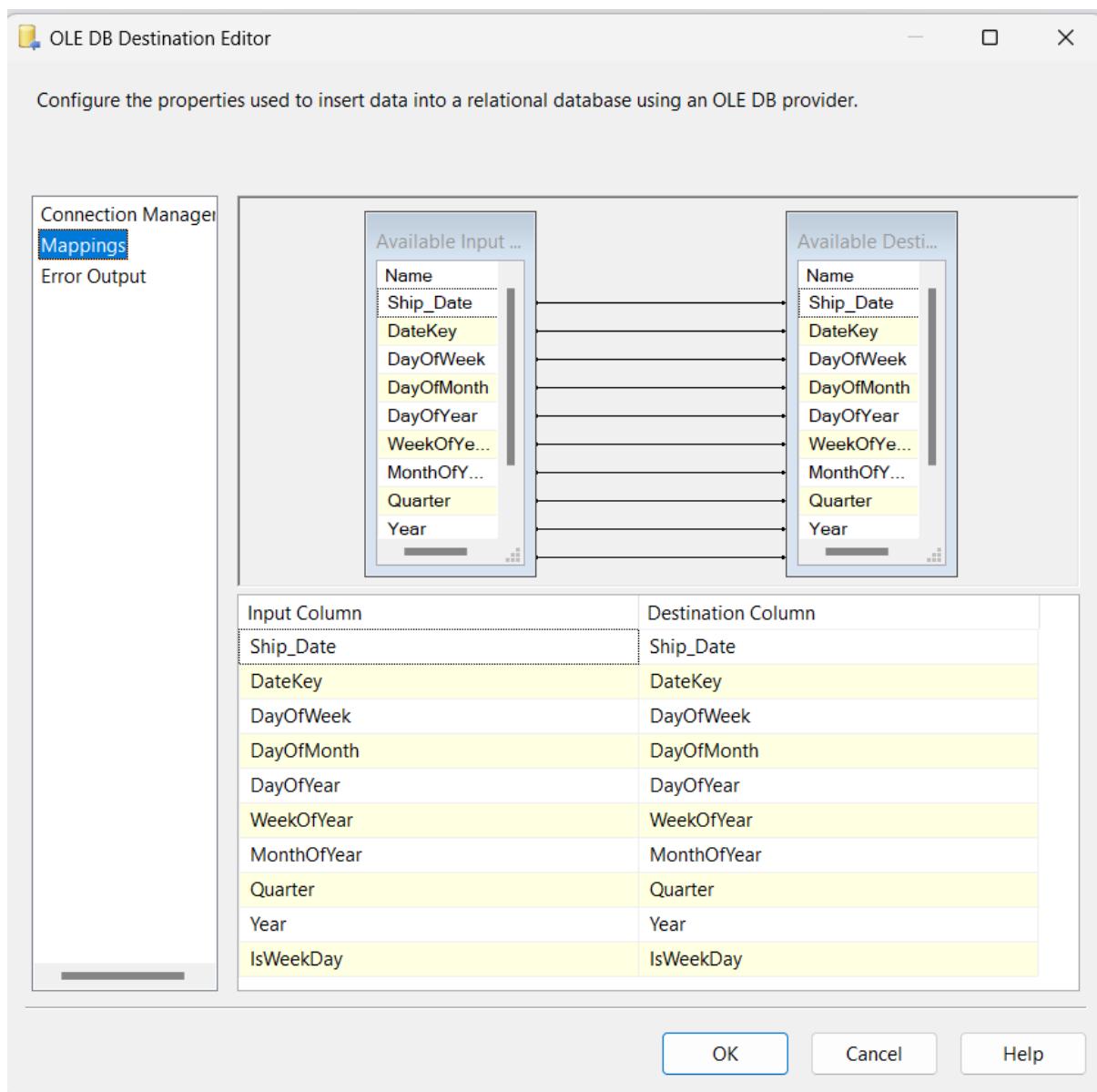
```

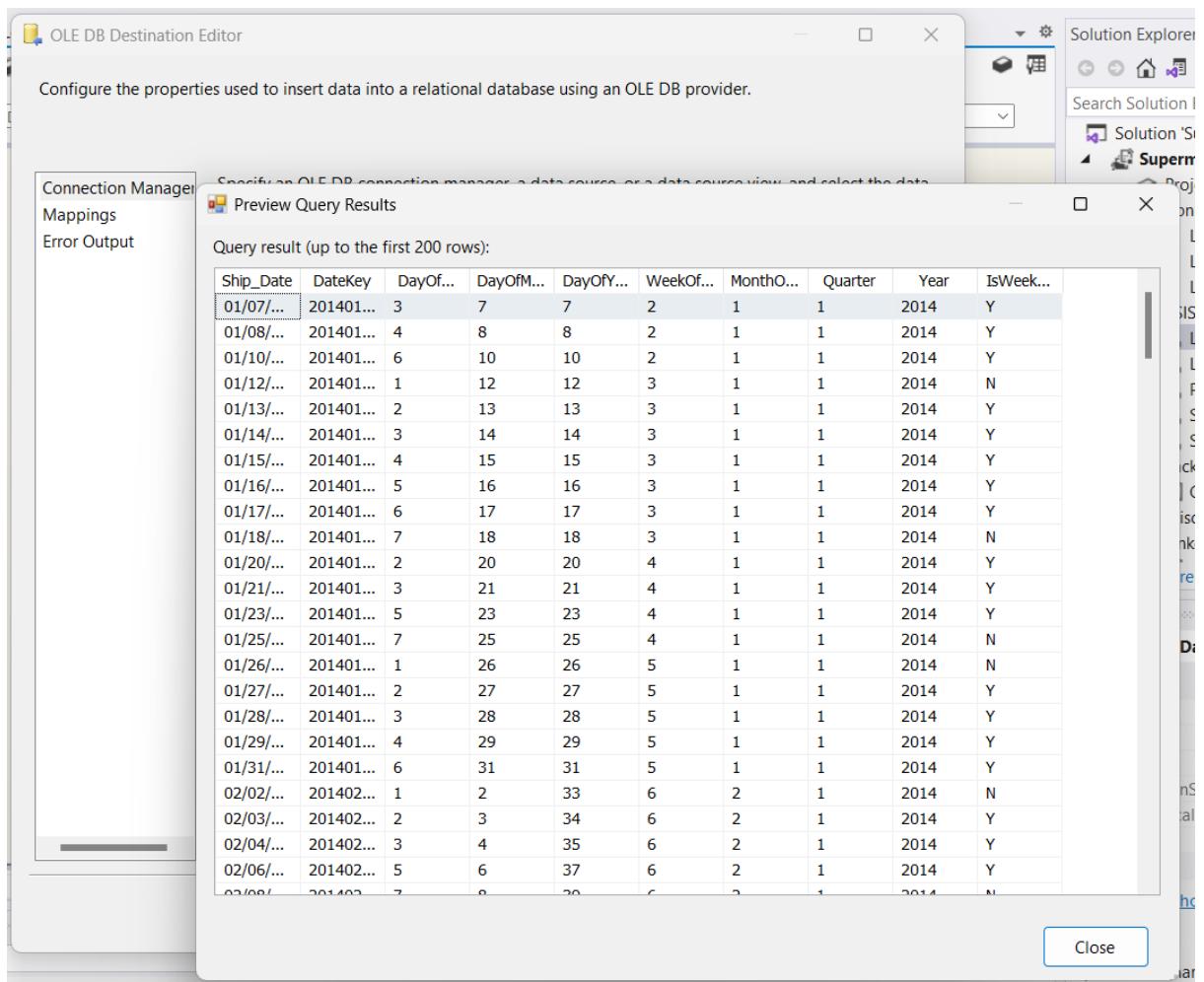
Ta sẽ lấy nguồn dữ liệu này để có thể tạo ra stgShipDate chứa thời gian thông kê kho và thời gian giao hàng

## DST – stgShipDate:

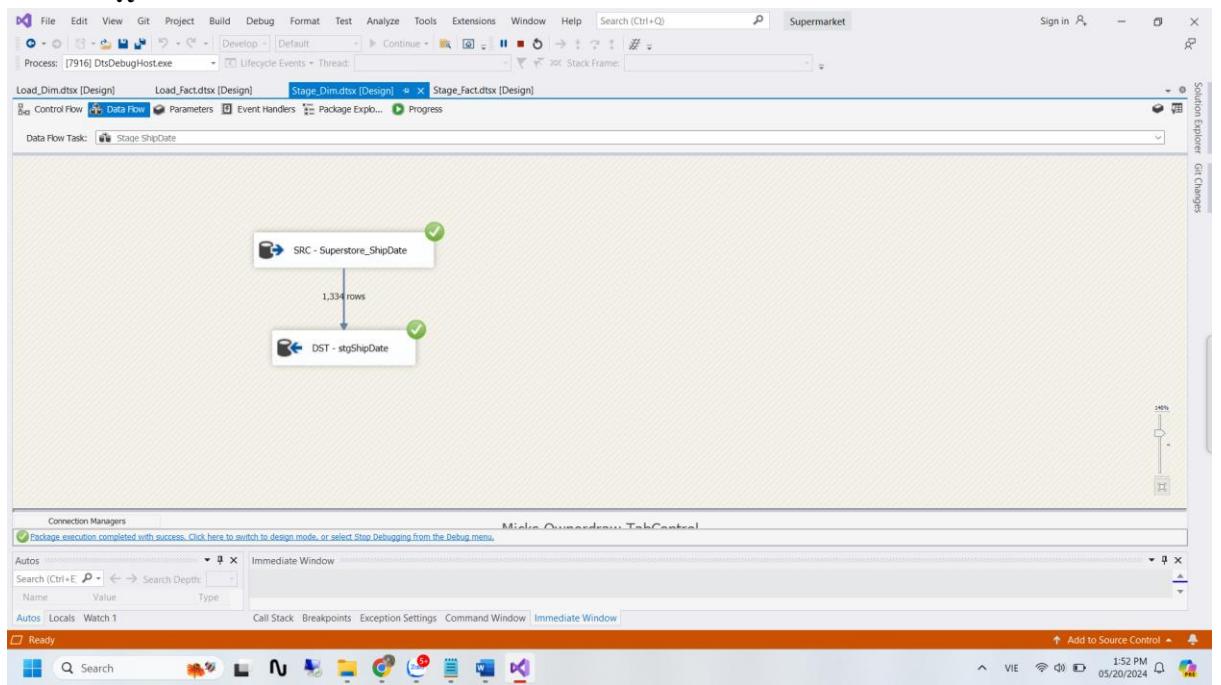




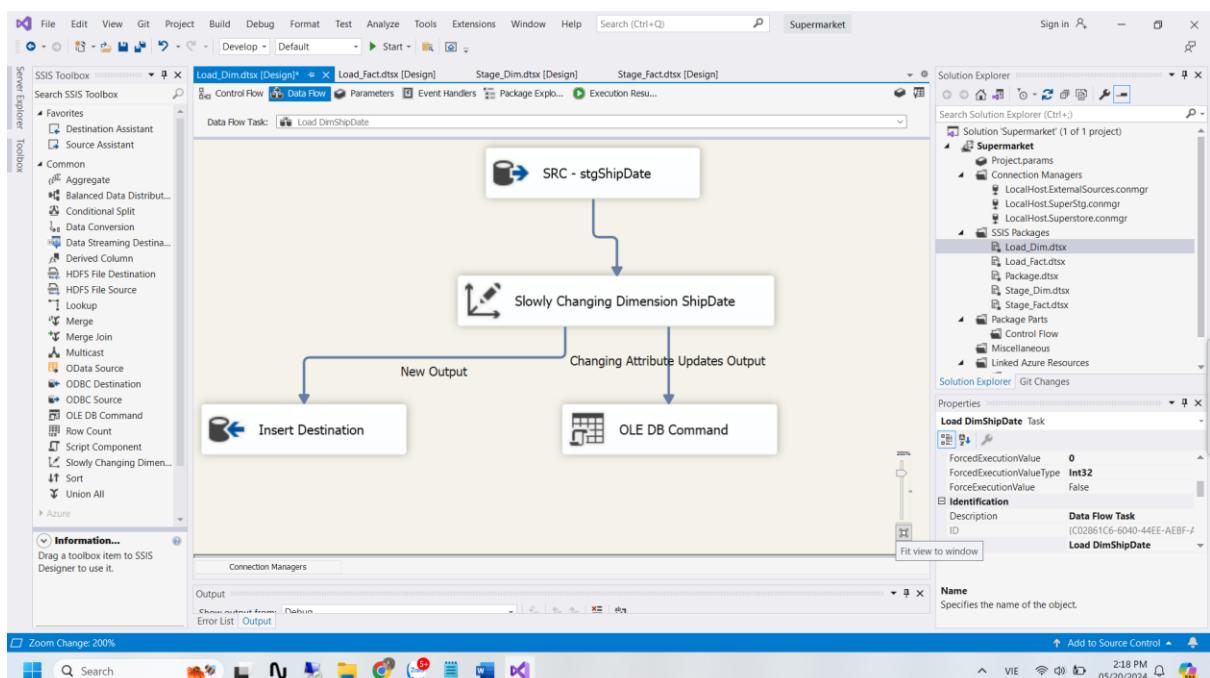




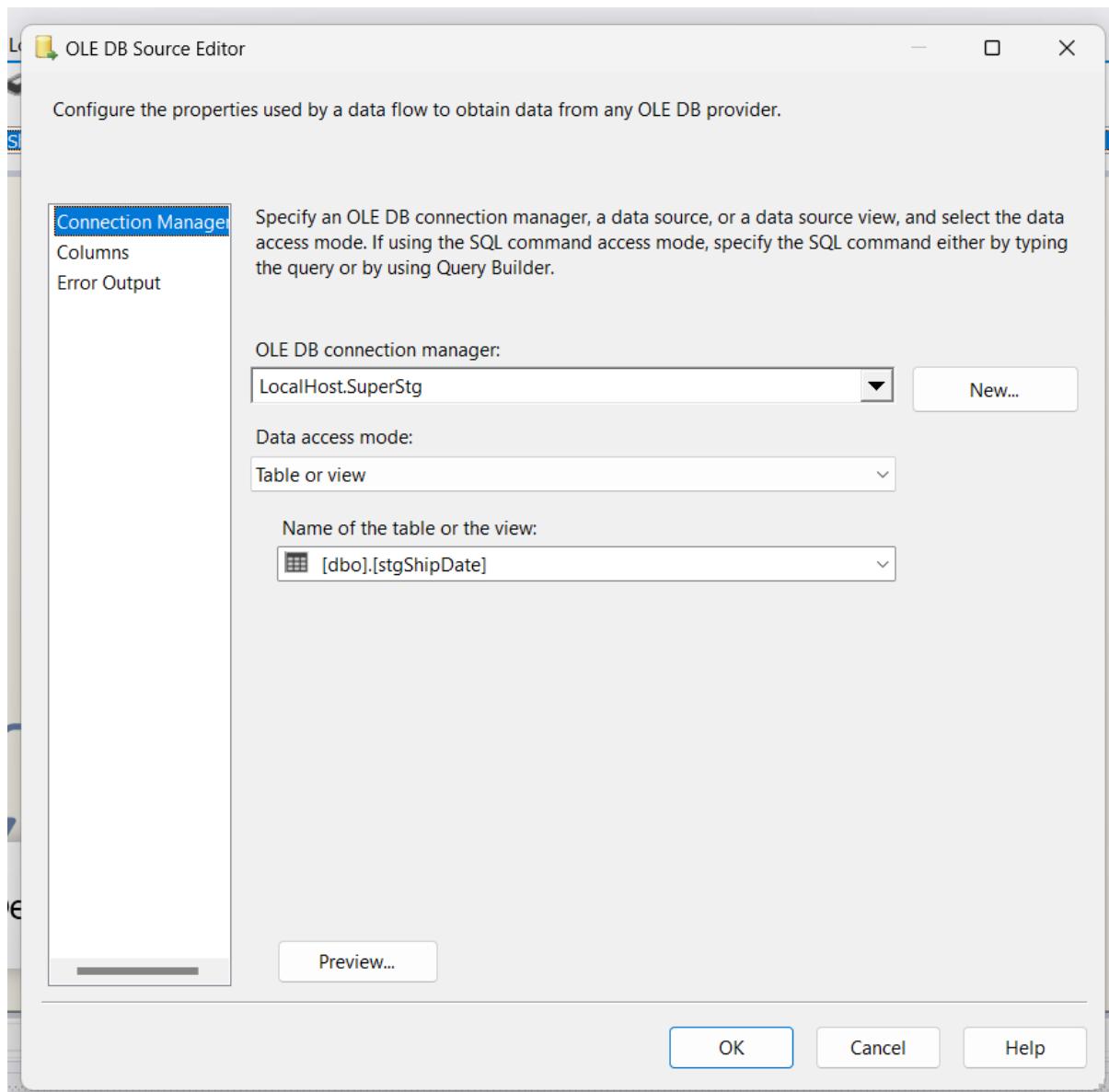
## Khởi chạy



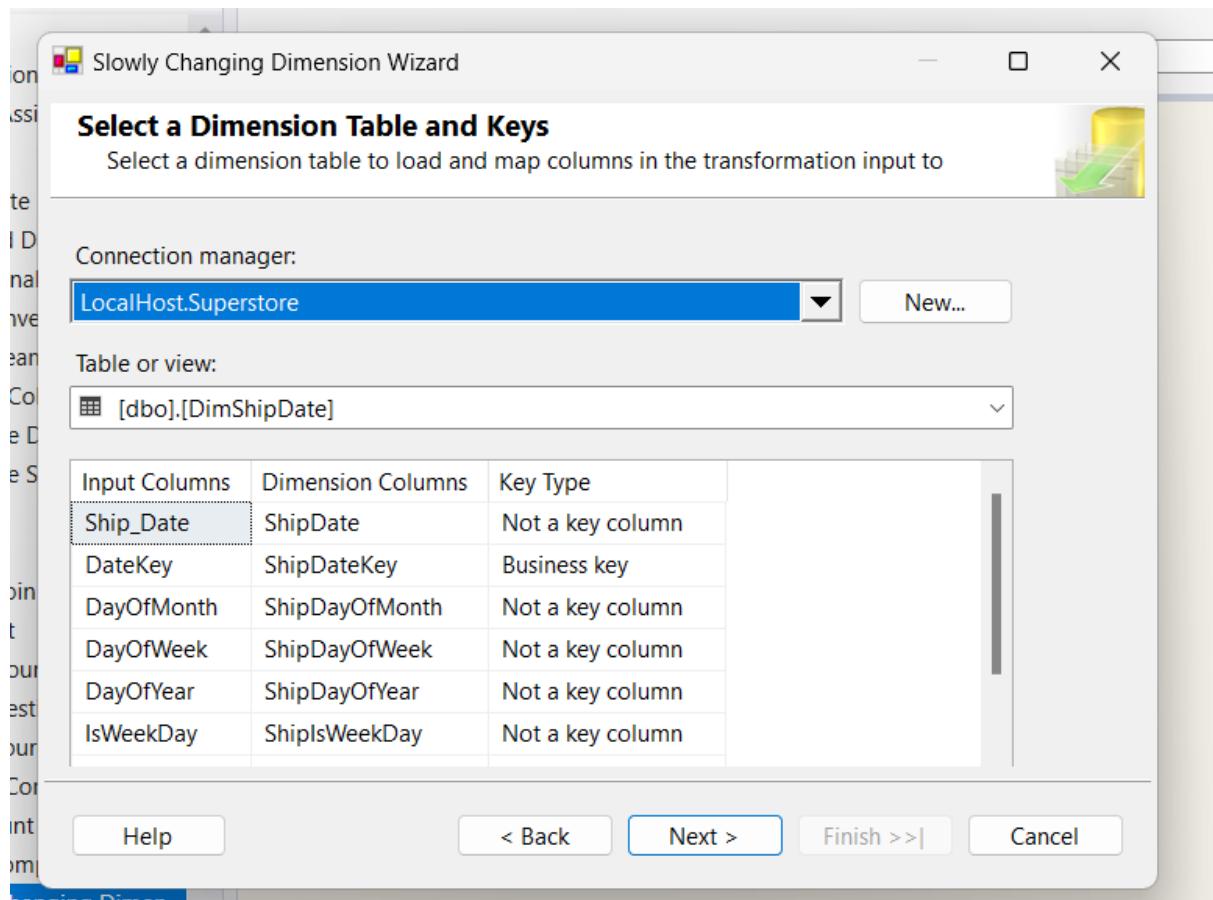
## b) Load dữ liệu từ stgShipDate vào DimDate



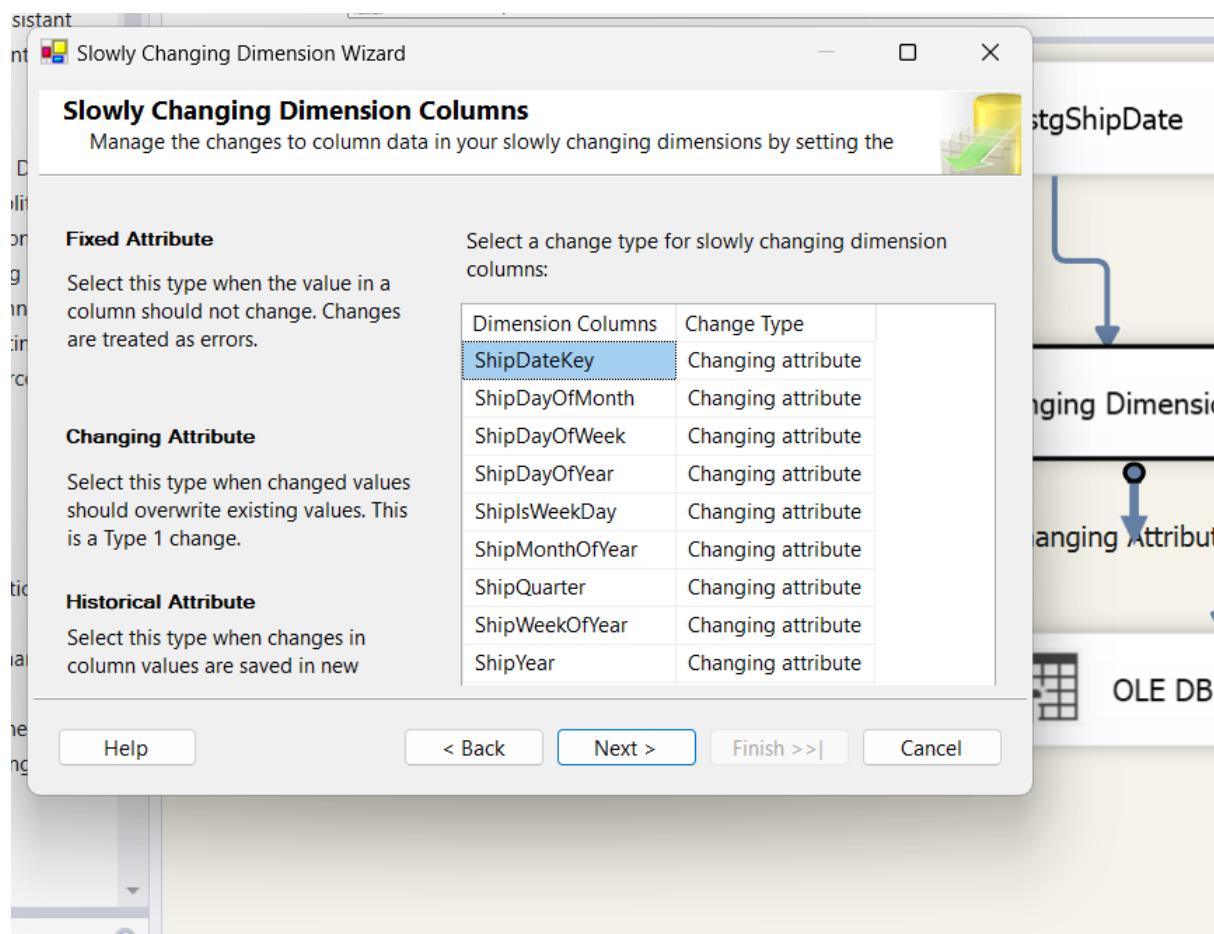
**SRC – stgShipDate** là dữ liệu từ bảng stgShipDate đã load ở trên



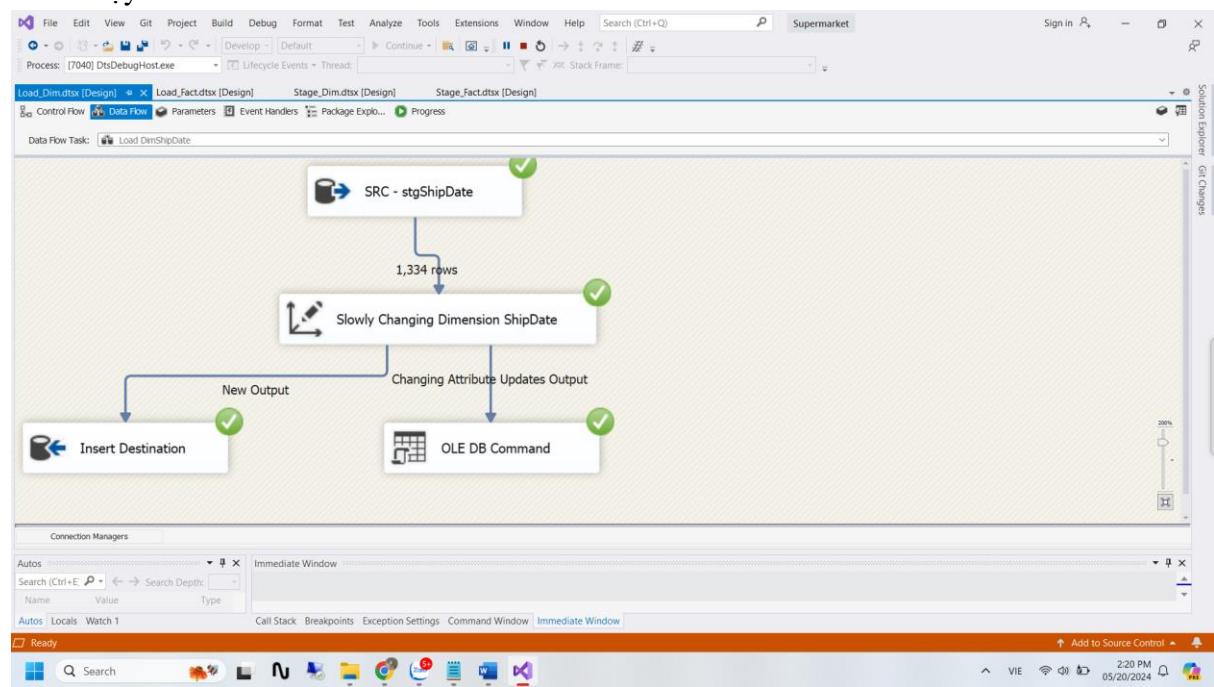
**Slowly Changing Dimension – DimShipDate** load dữ liệu từ stgShipDate vào DimShipDate



Chọn loại SCD là Changing attribute (Update dòng)

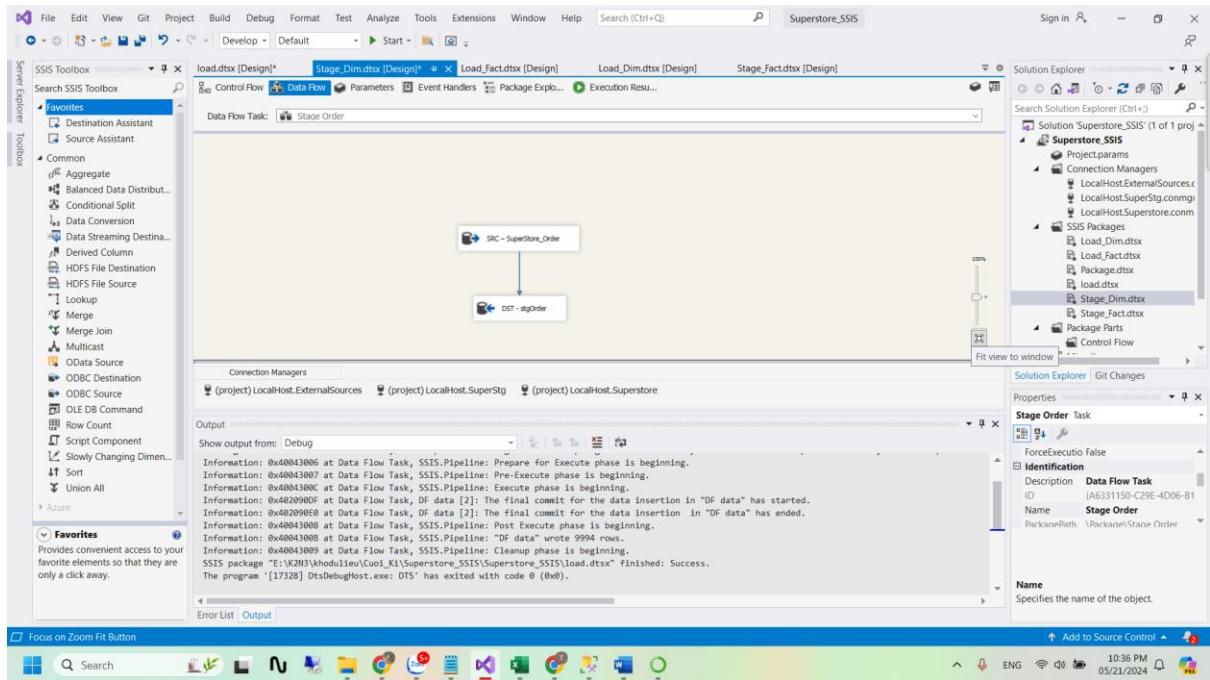


## Khởi chạy



### 3.2.5. Order Dimension

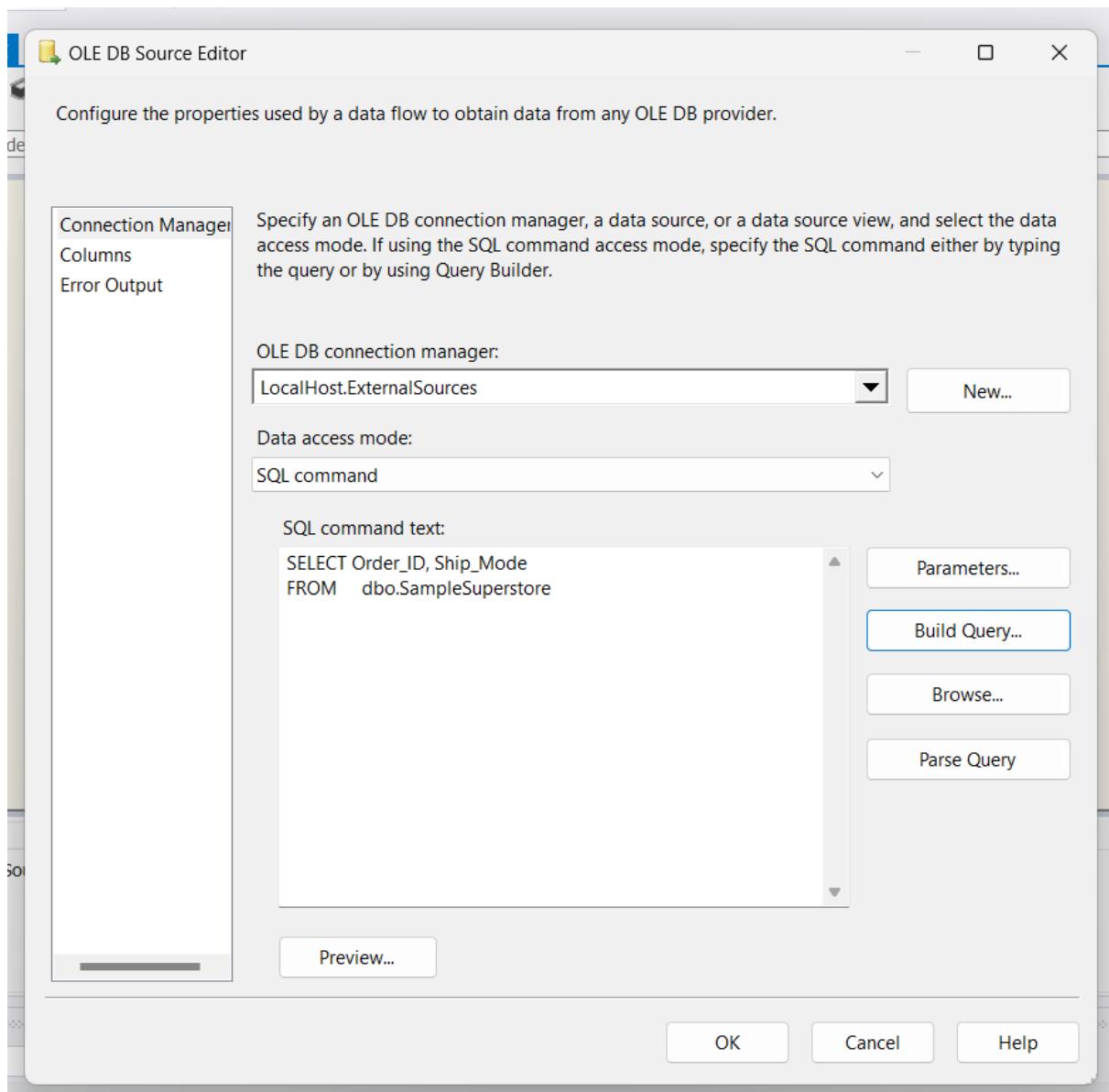
#### a) Load từ nguồn vào bảng stgOrder



SRC – SuperStore\_Order chứa dữ liệu các Order từ bảng SampleSuperstore

The screenshot shows the Microsoft Query Builder interface. A query is being constructed to select 'Order\_ID' and 'Ship\_Mode' from the 'SampleSuperstore' table. The results pane indicates 0 rows were returned.

```
SELECT Order_ID, Ship_Mode
FROM dbo.SampleSuperstore
```



## DST – stgProduct

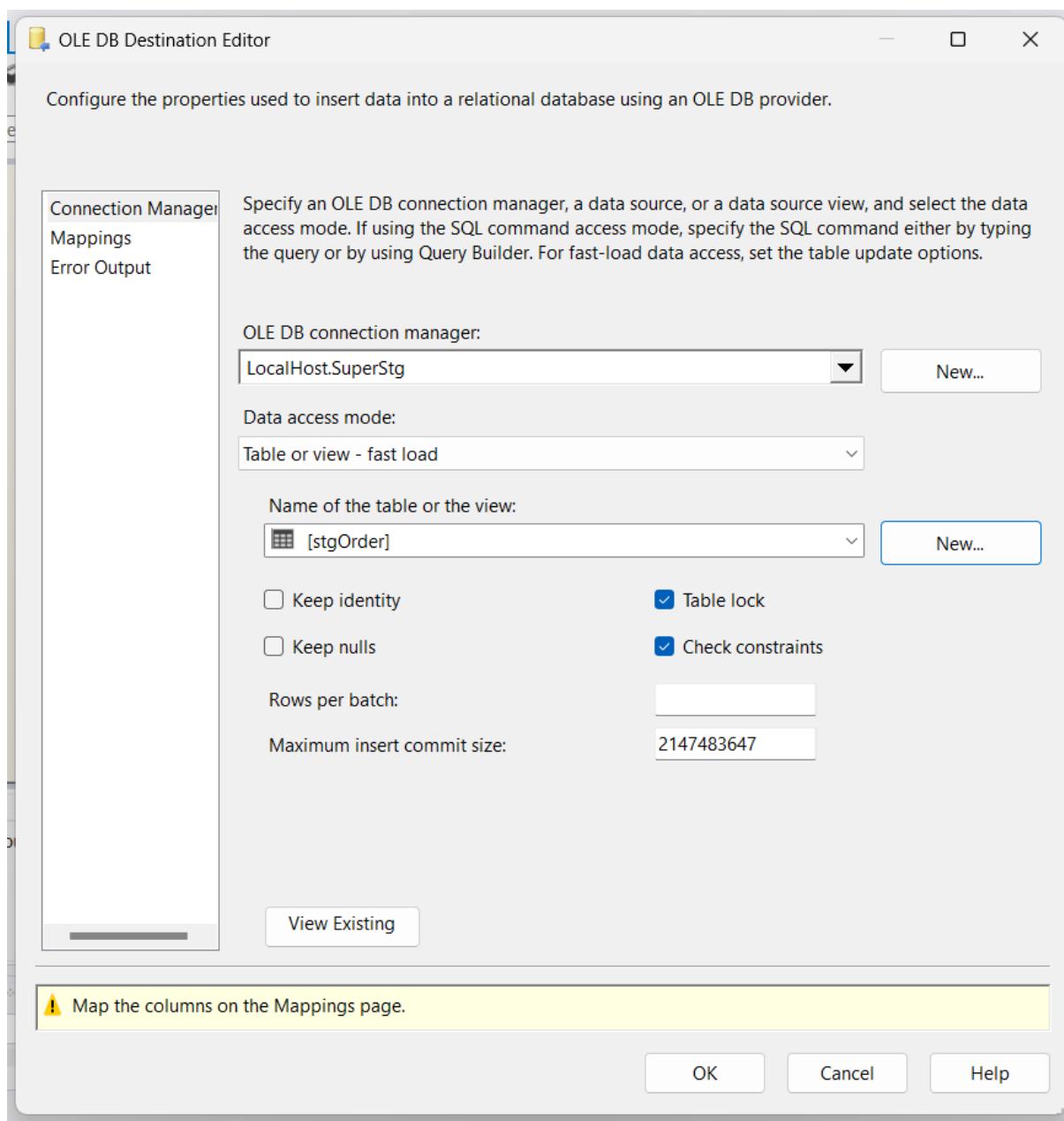
access mode. If using the SQL command access mode, specify the SQL command either

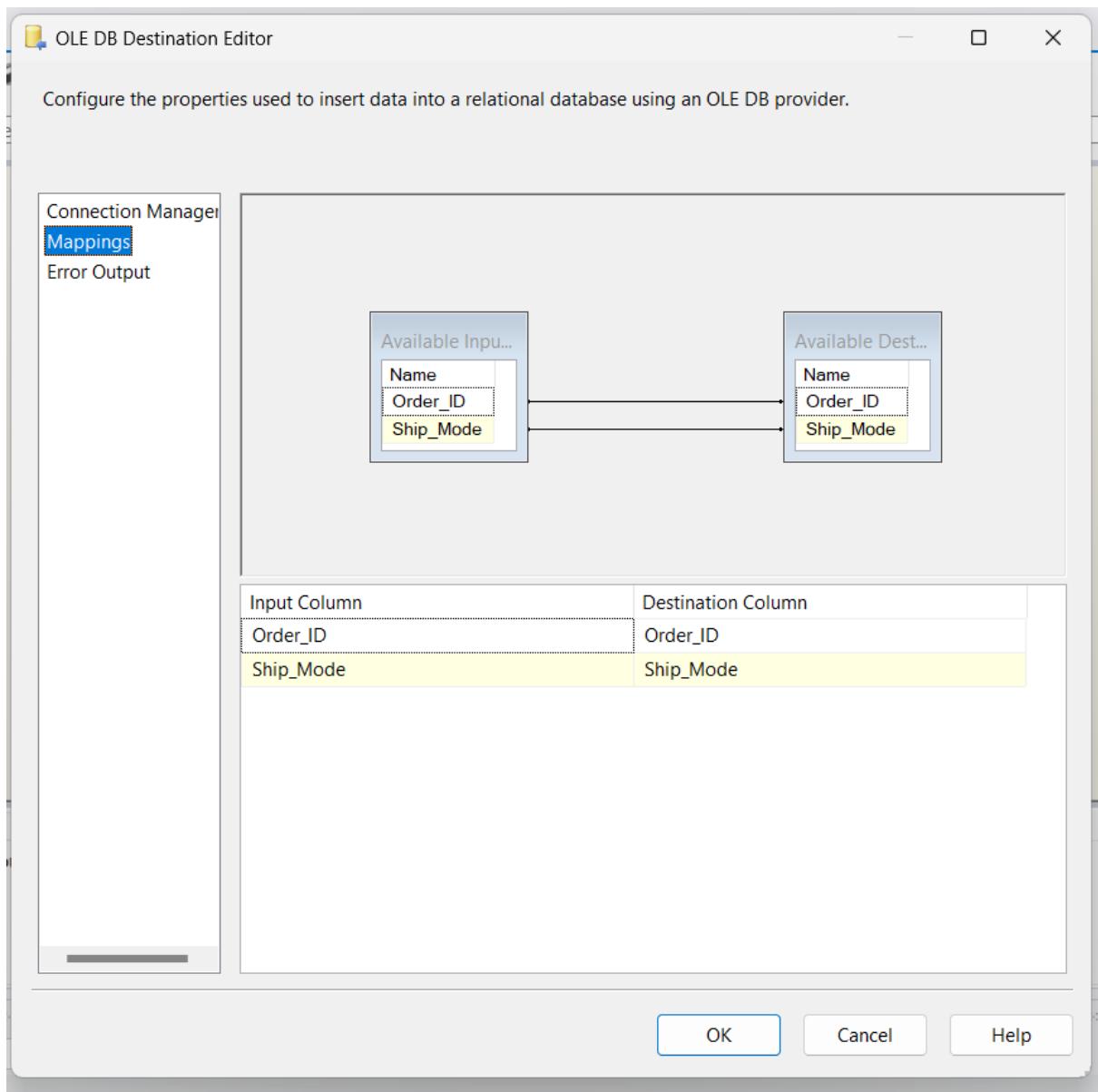
### Create Table

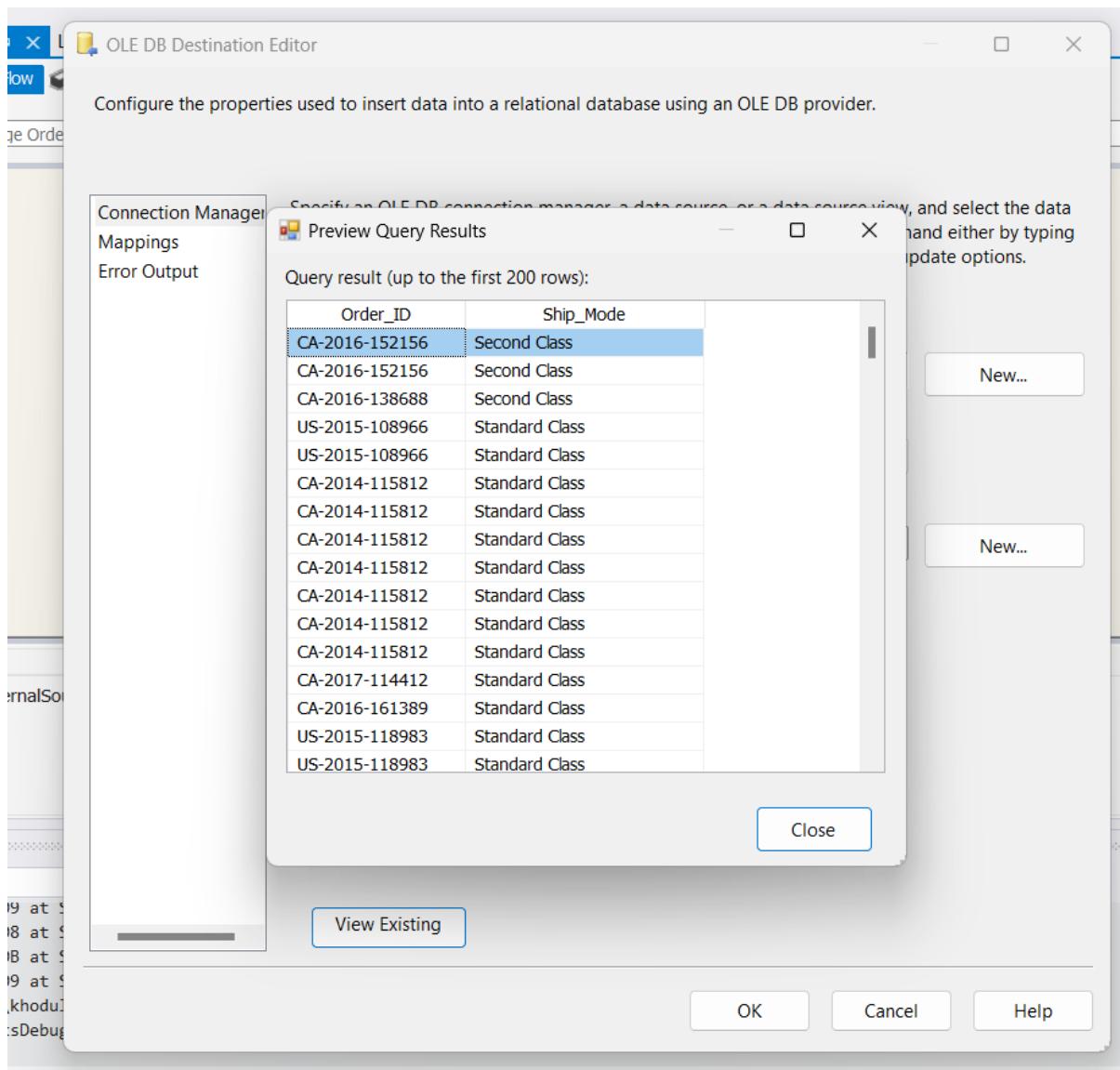
```
CREATE TABLE [stgOrder] (
    [Order_ID] nvarchar(30),
    [Ship_Mode] nvarchar(20)
)
```

OK

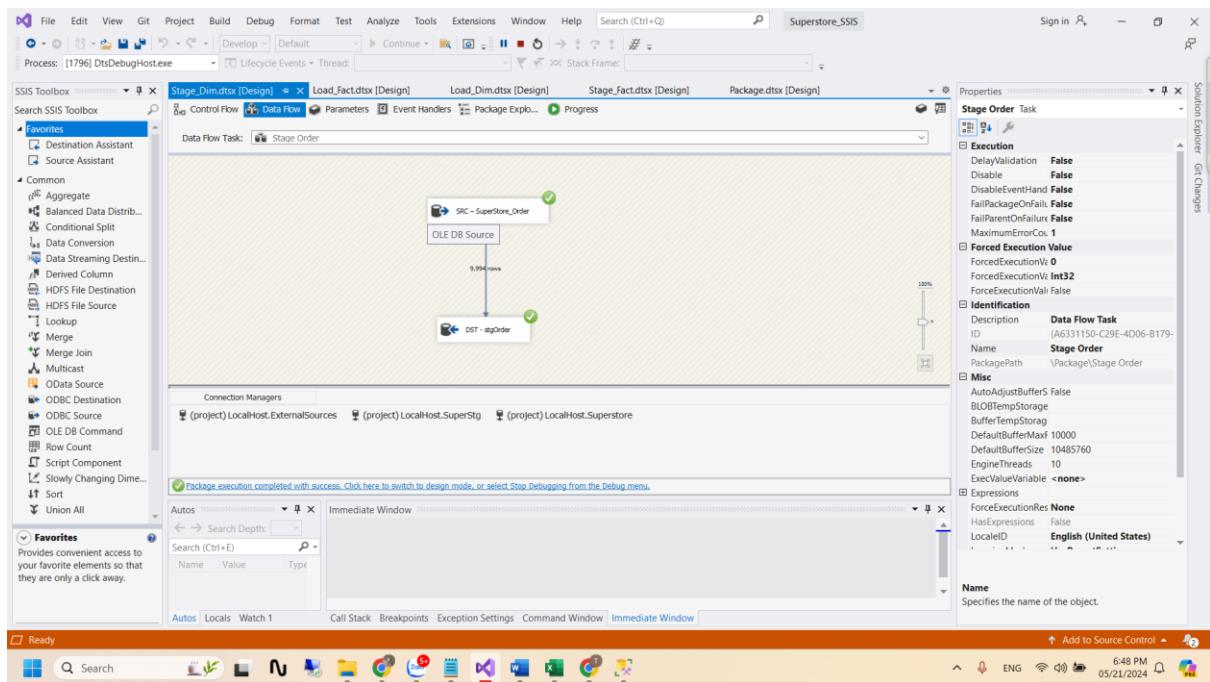
Cancel



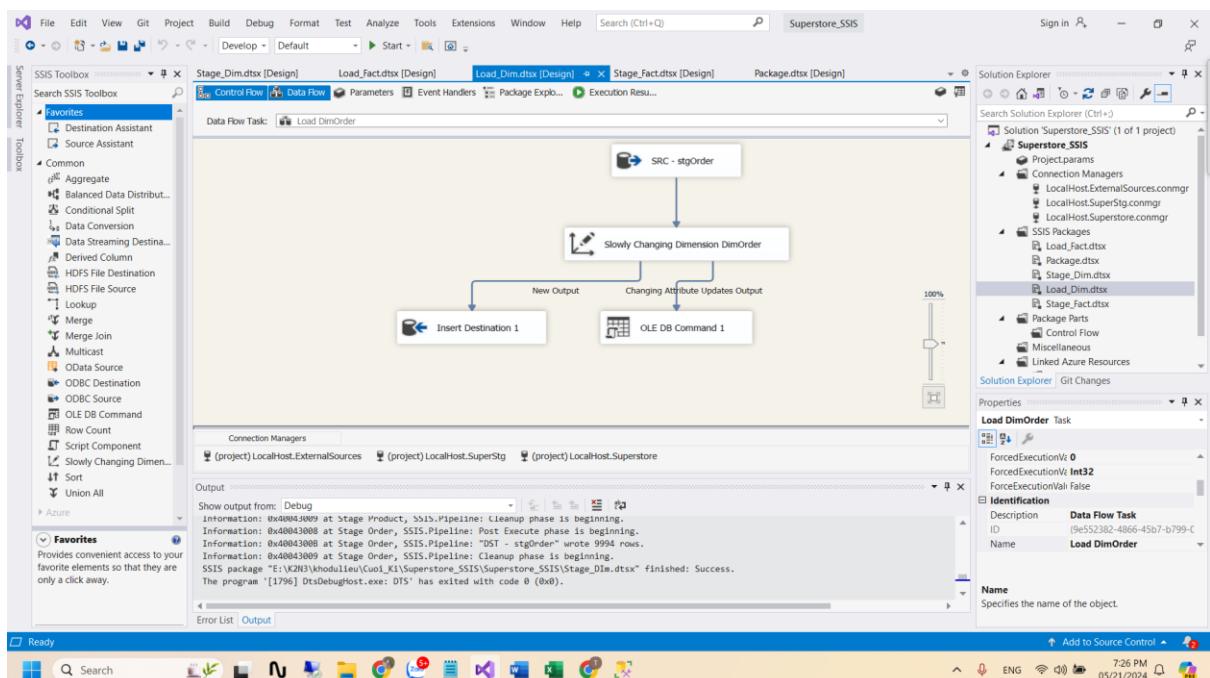




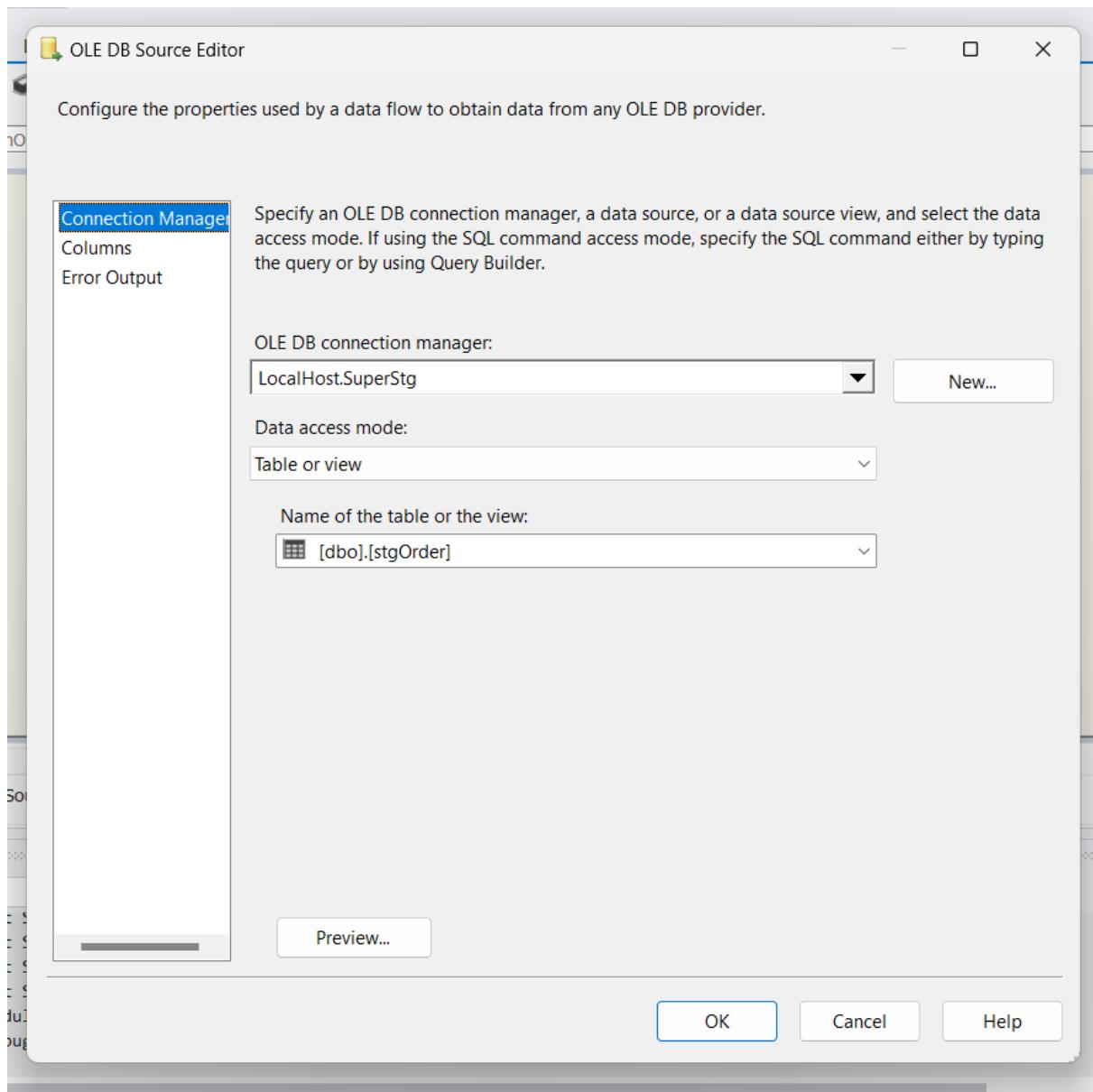
## **Khởi chạy**



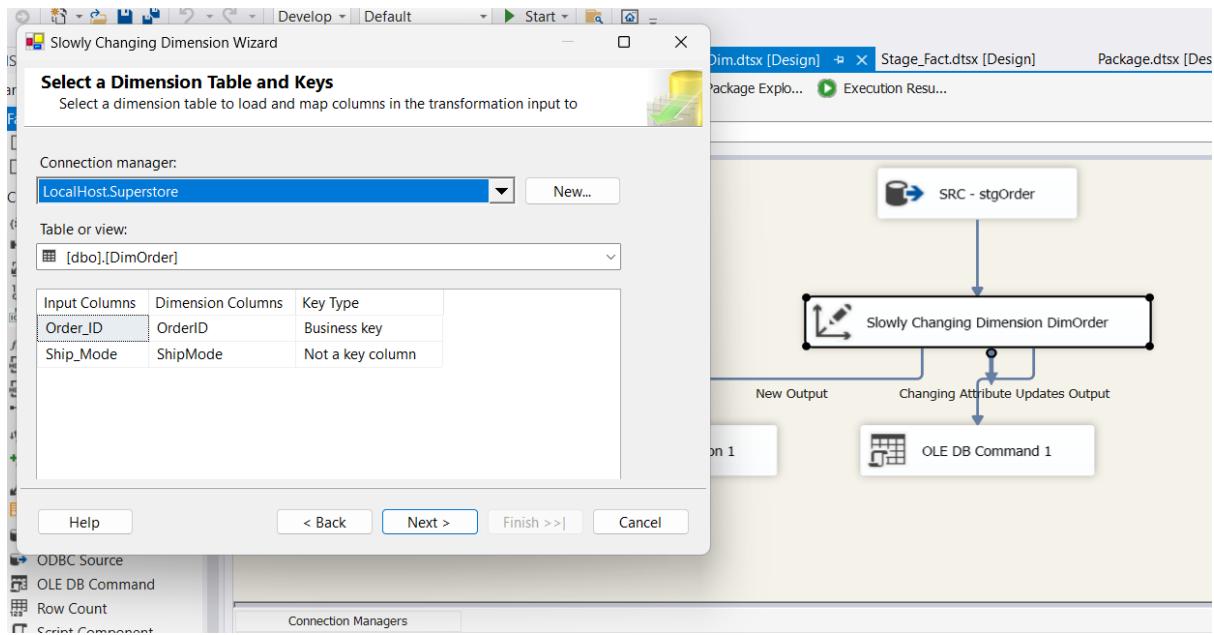
## b) Load dữ liệu từ stgOrder vào DimOrder



**SRC – stgOrder** là dữ liệu từ bảng stgOrder đã load ở trên



**Slowly Changing Dimension – DimOrder** load dữ liệu từ stgOrder vào DimOrder



Chọn loại SCD là Changing attribute (Update dòng)

Slowly Changing Dimension Wizard

### Slowly Changing Dimension Columns

Manage the changes to column data in your slowly changing dimensions by setting the

**Fixed Attribute**  
Select this type when the value in a column should not change. Changes are treated as errors.

**Changing Attribute**  
Select this type when changed values should overwrite existing values. This is a Type 1 change.

**Historical Attribute**  
Select this type when changes in column values are saved in new

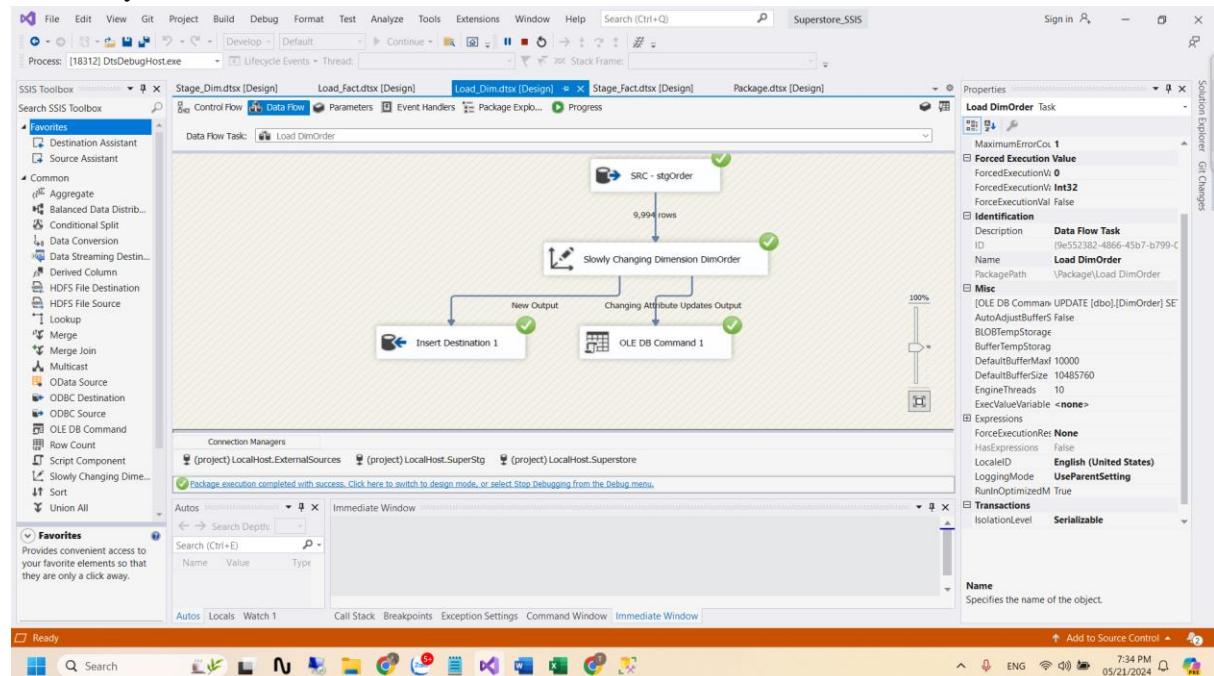
Dimension Columns	Change Type
ShipMode	Changing a...

Help < Back Next > Finish >> Cancel

ODBC Source

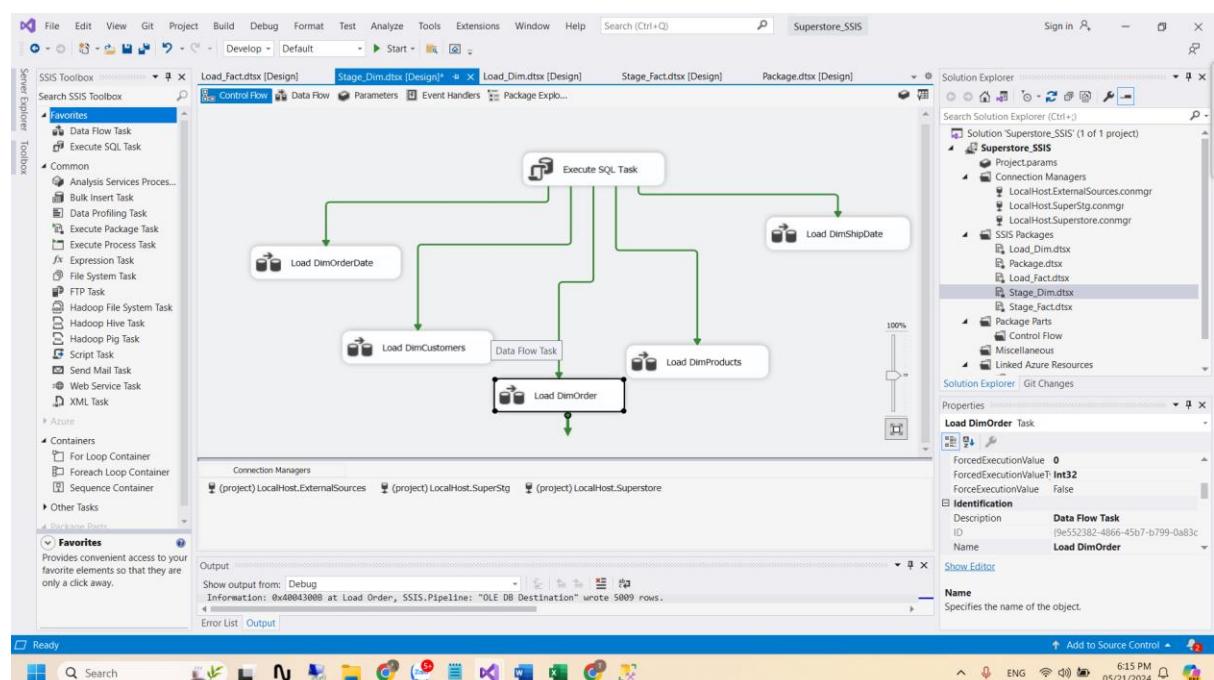
Khi thực hiện xong Slowly Changing Dimension – DimOrder thì Insert Destination và OLE DB Command sẽ xuất hiện.

Khởi chạy

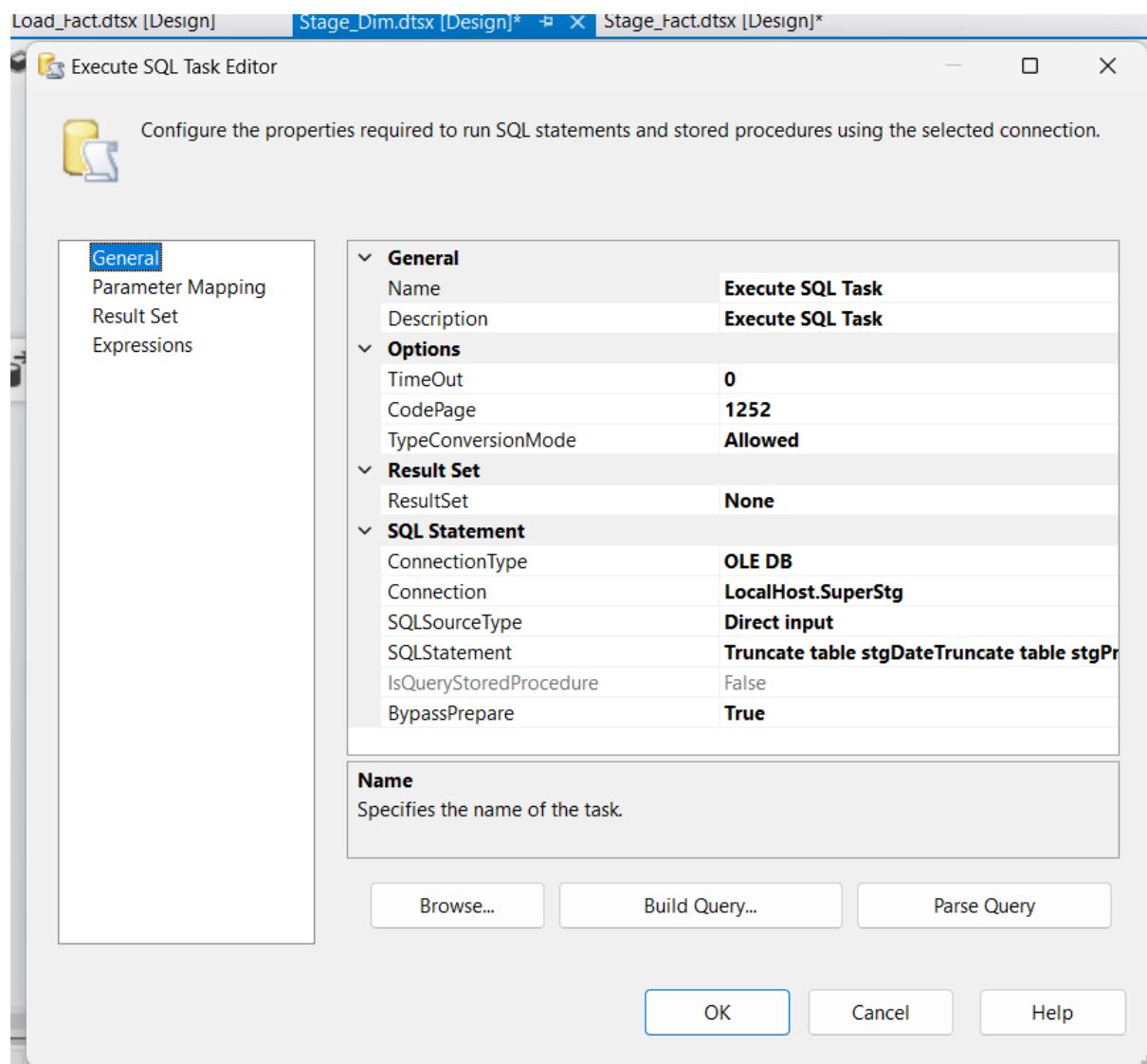


Khi chạy hoàn tất dữ liệu đã được Load vào DimOrder

### 3.2.6. Package Stage\_Dim



### a) Setup Execute SQL Task

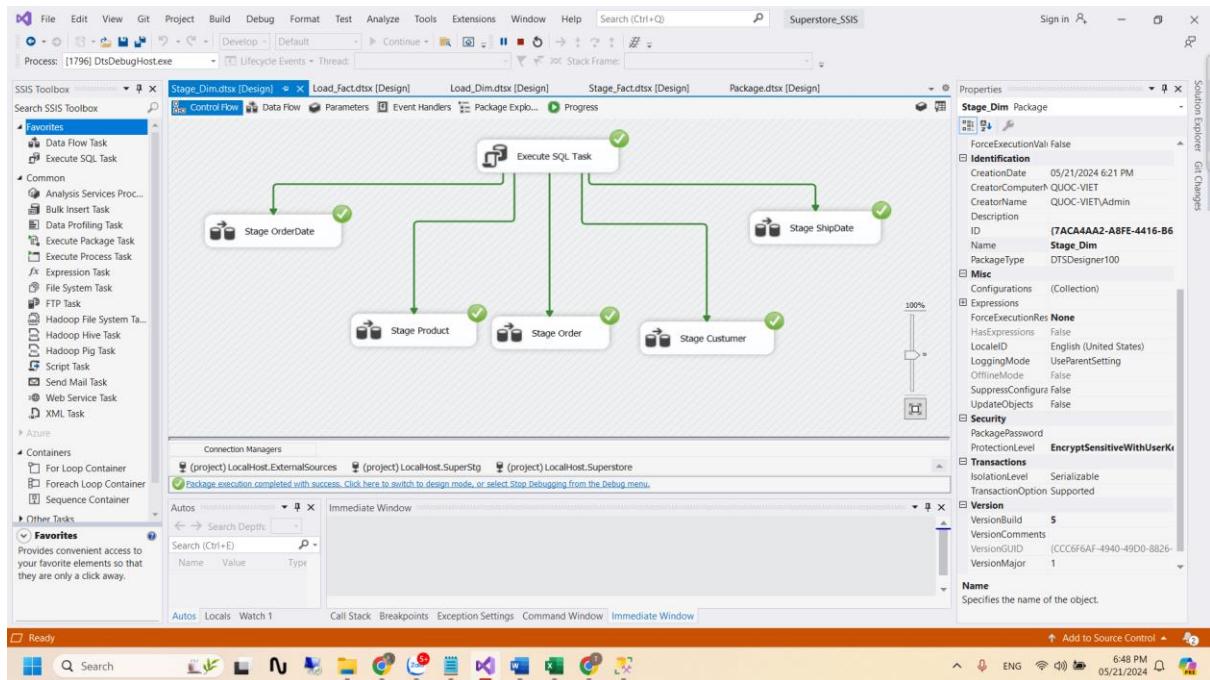


Connection đến database SuperStg

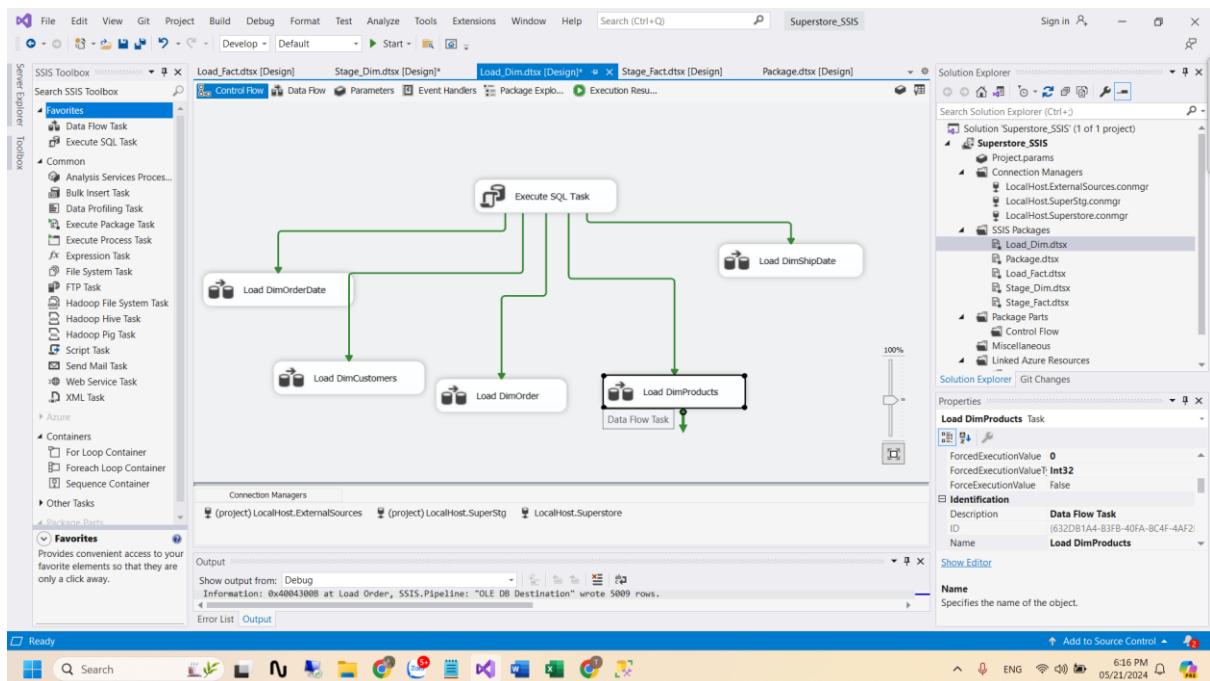
SQLStatement :

```
Truncate table stgOrderDate
Truncate table stgProducts
Truncate table stgShipDate
Truncate table stgCustomer
Truncate table stgOrder
```

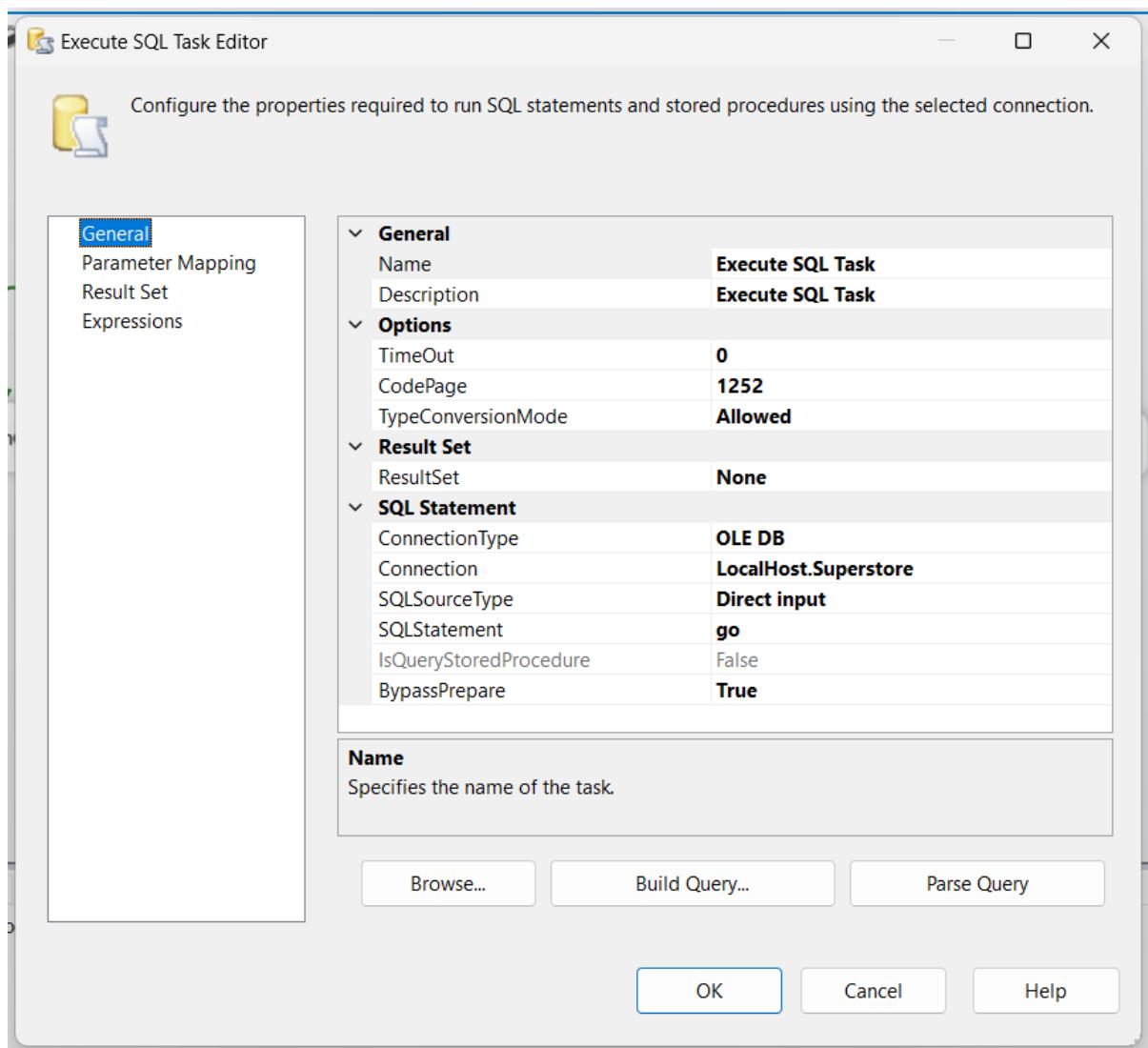
**Khởi chạy:**



### 3.2.7. Package Load\_Dim



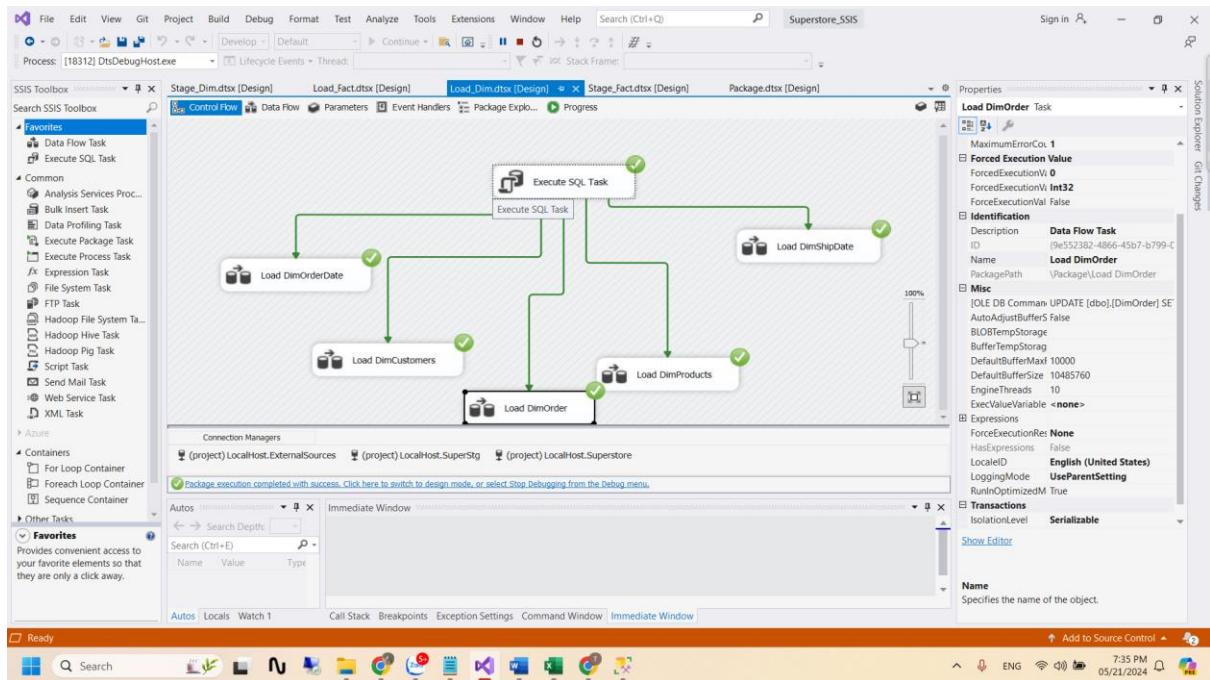
### a) Setup Execute SQL Task



Connection : Kết nối đến database Superstore

SQLStatement : go

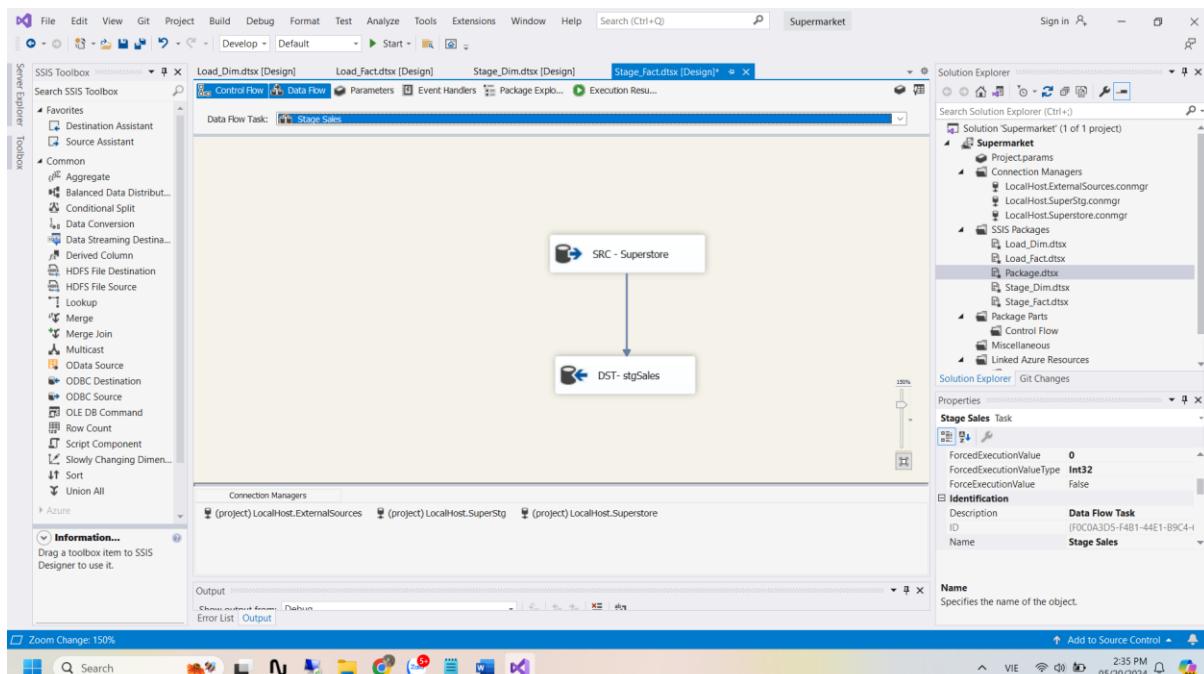
**Khởi chạy**



### 3.3. Import dữ liệu vào bảng các bảng fact

#### 3.3.1. Fact Sales

##### a) Load dữ liệu từ nguồn vào stgSales



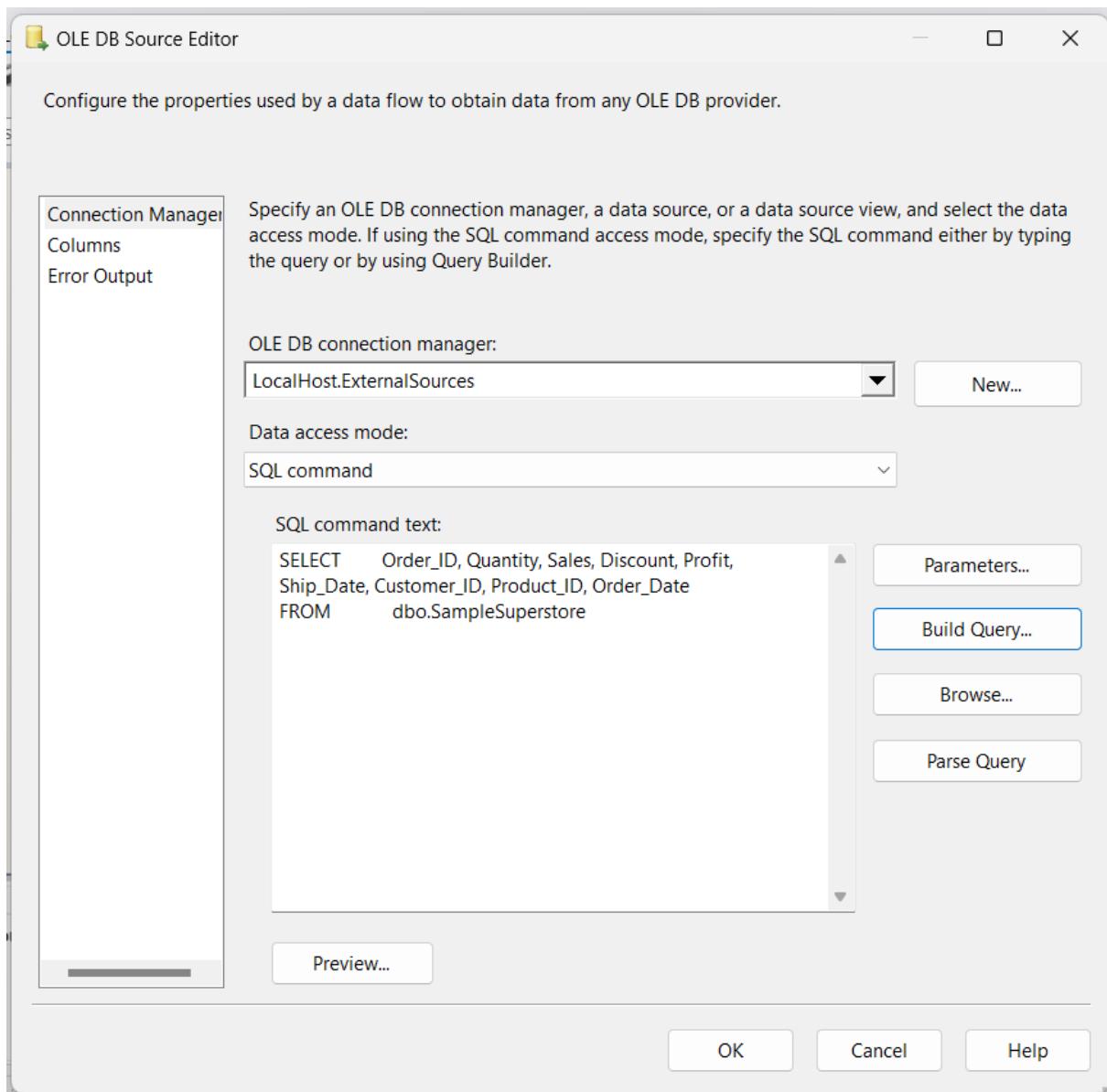
SRC – SuperStore chứa dữ liệu cần thiết để đưa vào bảng stgSales

The screenshot shows the Query Builder interface. At the top, there is a list of columns from the "SampleSuperstore (dbo)" table, including Row\_ID, Order\_ID, Order\_Date, Ship\_Date, Ship\_Mode, Customer\_ID, Customer\_Name, Segment, Country, City, State, Postal\_Code, Region, Product\_ID, Category, Sub\_Category, Product\_Name, and Sales. Below this, a mapping grid shows the columns being mapped to the "stgSales" table:

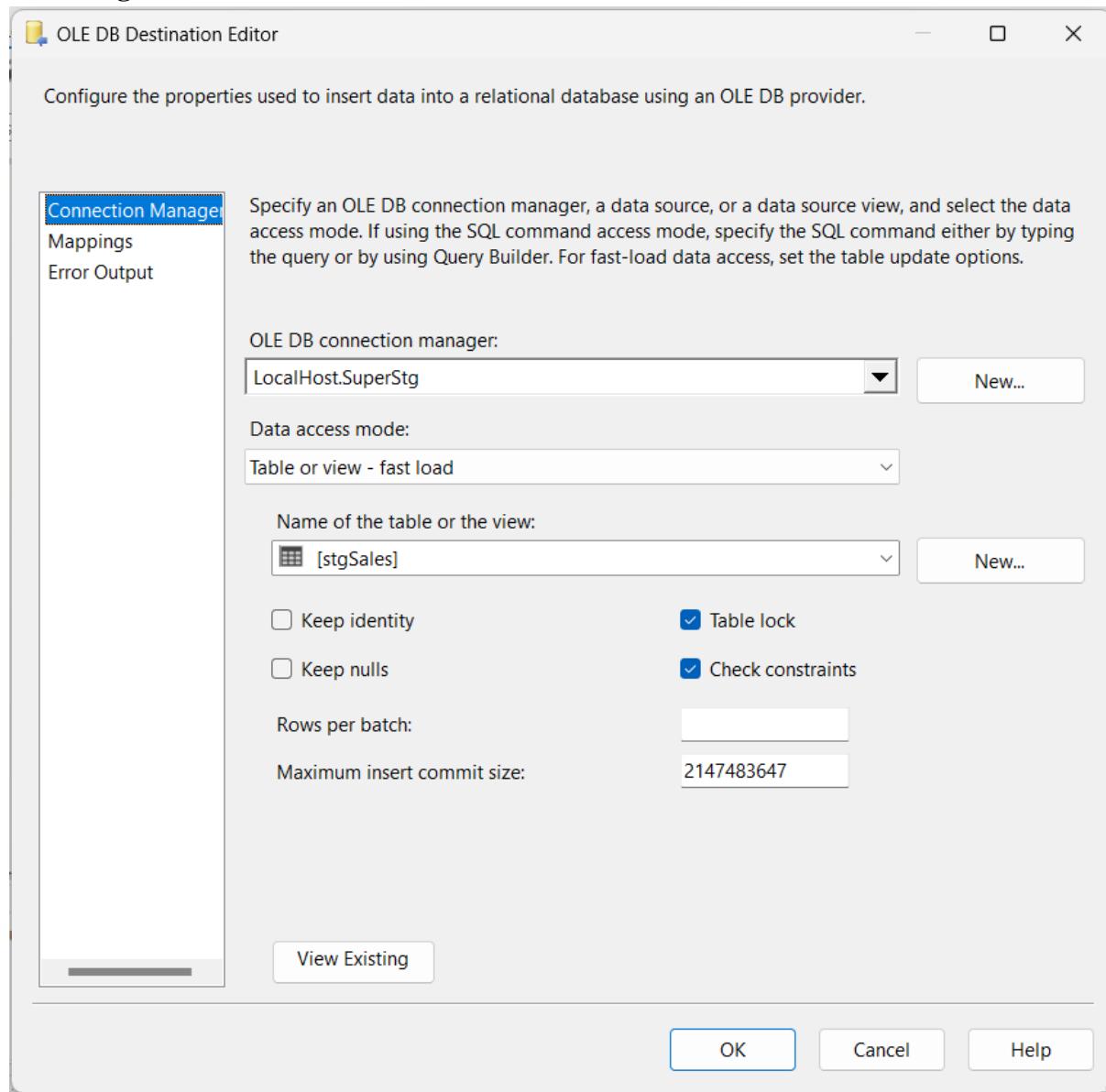
Column	Alias	Table	Output	Sort Type	Sort Order	Filter	Or...	Or...	Or...
Order_ID	Quantity	SampleSu...	SampleSu...	✓					
Quantity	Sales	SampleSu...	SampleSu...	✓					
Sales	Discount	SampleSu...	SampleSu...	✓					
Discount	Profit	SampleSu...	SampleSu...	✓					
Profit	Ship_Date	SampleSu...	SampleSu...	✓					
Ship_Date	Customer_ID	SampleSu...	SampleSu...	✓					
Customer_ID	Product_ID	SampleSu...	SampleSu...	✓					
Product_ID	Order_Date	SampleSu...	SampleSu...	✓					

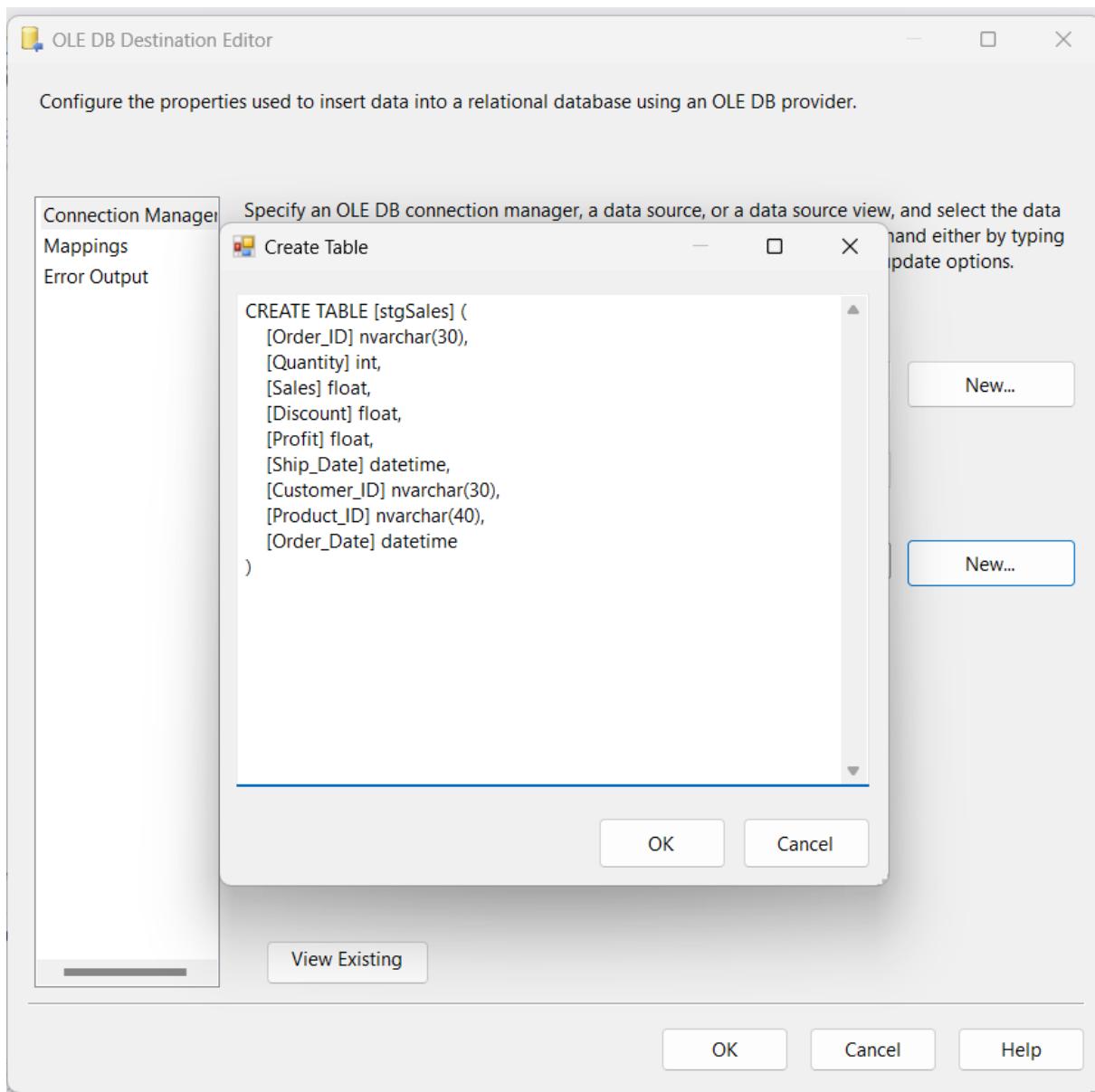
At the bottom of the Query Builder, the SQL query is displayed:

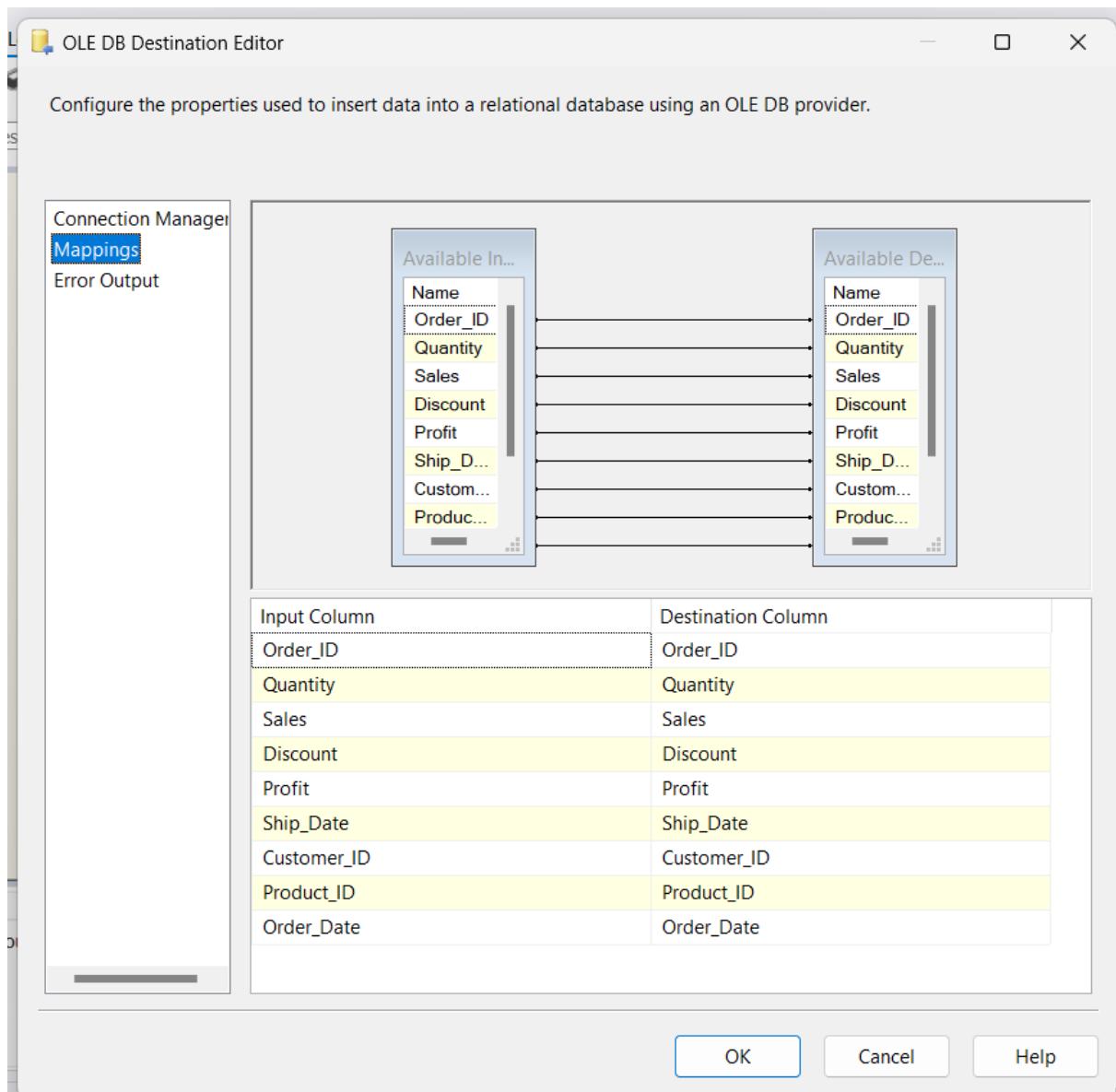
```
SELECT Order_ID, Quantity, Sales, Discount, Profit, Ship_Date, Customer_ID, Product_ID, Order_Date
FROM
    dbo.SampleSuperstore
```



## DST – stgSales







OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager  
Mappings  
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data to insert.

Preview Query Results

Query result (up to the first 200 rows):

Order_ID	Quantity	Sales	Discount	Profit	Ship_Date	Custom...	Product...	Order_...
CA-201...	2	261.95...	0	41.913...	11/11/...	CG-125...	FUR-BO...	11/08/...
CA-201...	3	731.94...	0	219.58...	11/11/...	CG-125...	FUR-CH...	11/08/...
CA-201...	2	14.619...	0	6.8713...	06/16/...	DV-130...	OFF-LA...	06/12/...
US-201...	5	957.57...	0.4499...	-383.0...	10/18/...	SO-203...	FUR-TA...	10/11/...
US-201...	2	22.368...	0.2000...	2.5164...	10/18/...	SO-203...	OFF-ST...	10/11/...
CA-201...	7	48.860...	0	14.169...	06/14/...	BH-117...	FUR-FU...	06/09/...
CA-201...	4	7.2800...	0	1.9656...	06/14/...	BH-117...	OFF-AR...	06/09/...
CA-201...	6	907.15...	0.2000...	90.715...	06/14/...	BH-117...	TEC-PH...	06/09/...
CA-201...	3	18.503...	0.2000...	5.7824...	06/14/...	BH-117...	OFF-BI...	06/09/...
CA-201...	5	114.90...	0	34.470...	06/14/...	BH-117...	OFF-AP...	06/09/...
CA-201...	9	1706.1...	0.2000...	85.309...	06/14/...	BH-117...	FUR-TA...	06/09/...
CA-201...	4	911.42...	0.2000...	68.356...	06/14/...	BH-117...	TEC-PH...	06/09/...
CA-201...	3	15.552...	0.2000...	5.4432...	04/20/...	AA-104...	OFF-PA...	04/15/...
CA-201...	3	407.97...	0.2000...	132.59...	12/10/...	IM-15070	OFF-BI...	12/05/...
US-201...	5	68.809...	0.8000...	-123.8...	11/26/...	HP-148...	OFF-AP...	11/22/...
US-201...	3	2.5439...	0.8000...	-3.815...	11/26/...	HP-148...	OFF-BI...	11/22/...
CA-201...	6	665.88...	0	13.317...	11/18/...	PK-190...	OFF-ST...	11/11/...
CA-201...	2	55.5	0	9.9899...	05/15/...	AG-102...	OFF-ST...	05/13/...
CA-201...	2	8.5600...	0	2.4823...	09/01/...	ZD-219...	OFF-AR...	08/27/...
CA-201...	3	213.47...	0.2000...	16.010...	09/01/...	ZD-219...	TEC-PH...	08/27/...
CA-201...	4	22.719...	0.2000...	7.3839...	09/01/...	ZD-219...	OFF-BI...	08/27/...
CA-201...	7	19.459...	0	5.0595...	12/13/...	KB-165...	OFF-AR...	12/09/...
CA-201...	7	60.340...	0	15.688...	12/13/...	KB-165...	OFF-AP...	12/09/...

Close

## Khởi chạy:

File Edit View Git Project Build Debug Format Test Analyze Tools Extensions Window Help Search (Ctrl+Q) Supermarket Sign in

Process: [22328] DtsDebugHost.exe Develop Continue Stack Frame: 11% 13%

Load\_Dim.dtsx [Design] Load\_Fact.dtsx [Design] Stage\_Dim.dtsx [Design] Stage\_Fact.dtsx [Design] Data Flow Control Flow Parameters Event Handlers Package Explorer Progress

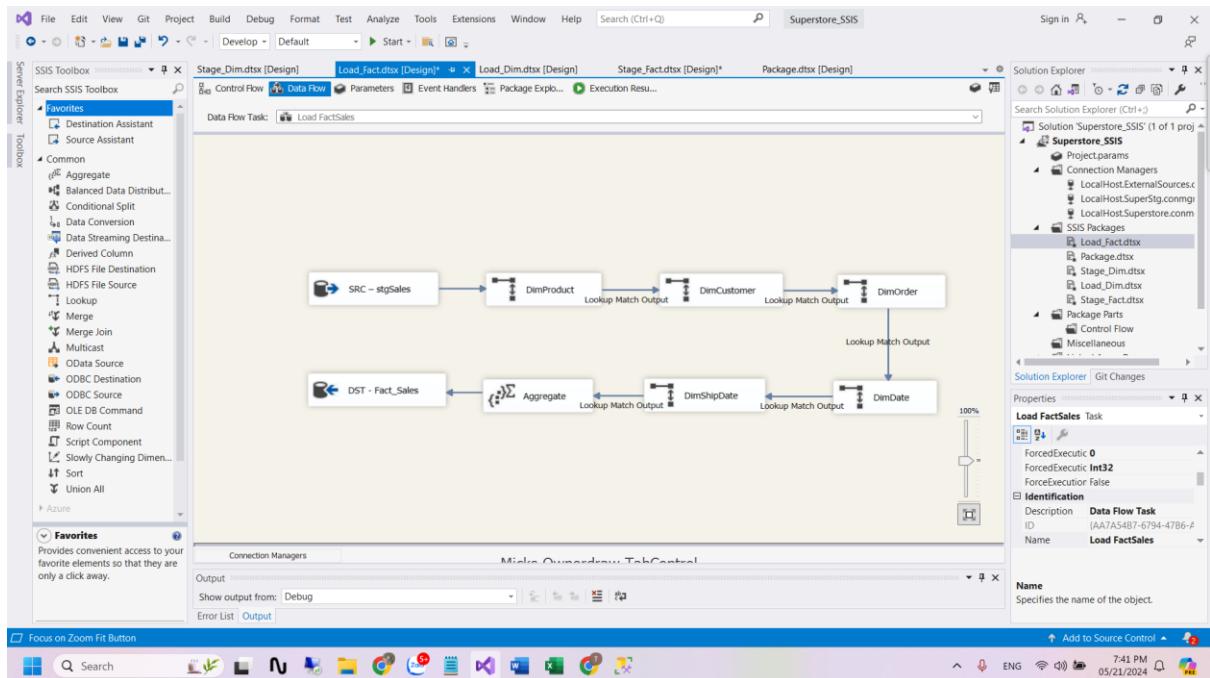
Data Flow Task: Stage Sales

Connection Managers: (project) localhost.ExternalSources, (project) localhost.SuperStore, (project) localhost.Superstore

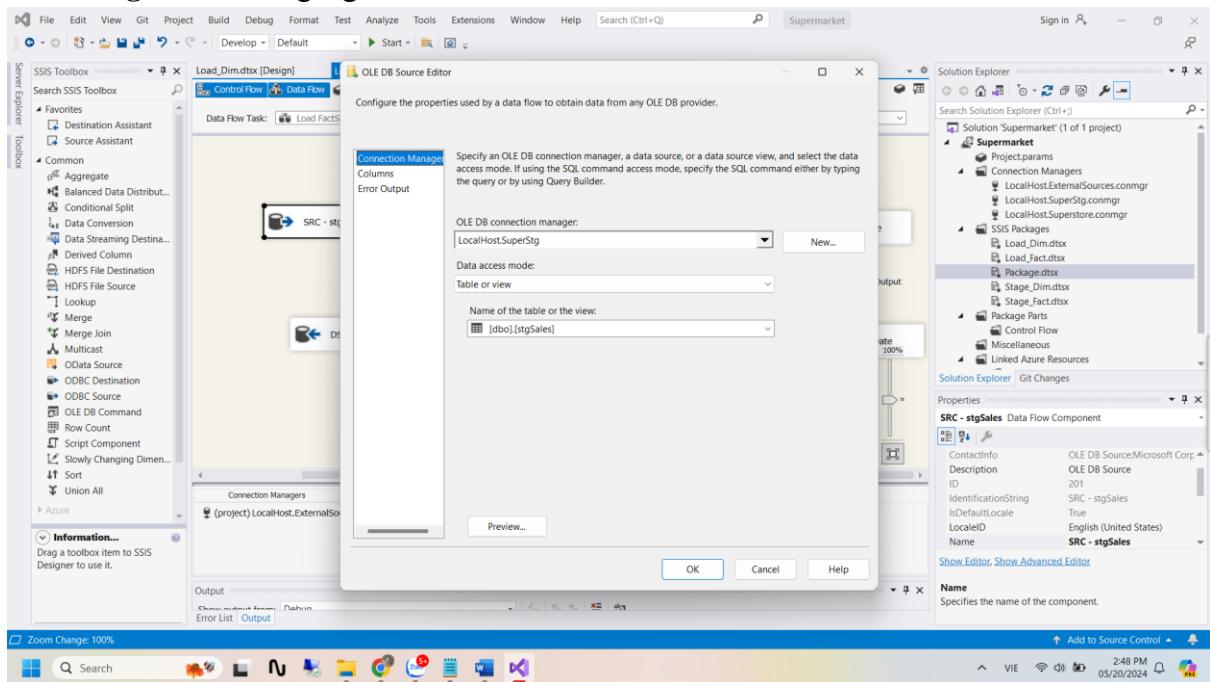
Package executor completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

Autos, Immediate Window, Call Stack, Breakpoints, Exception Settings, Command Window, Immediate Window, Add to Source Control, 2:38 PM, 05/20/2024

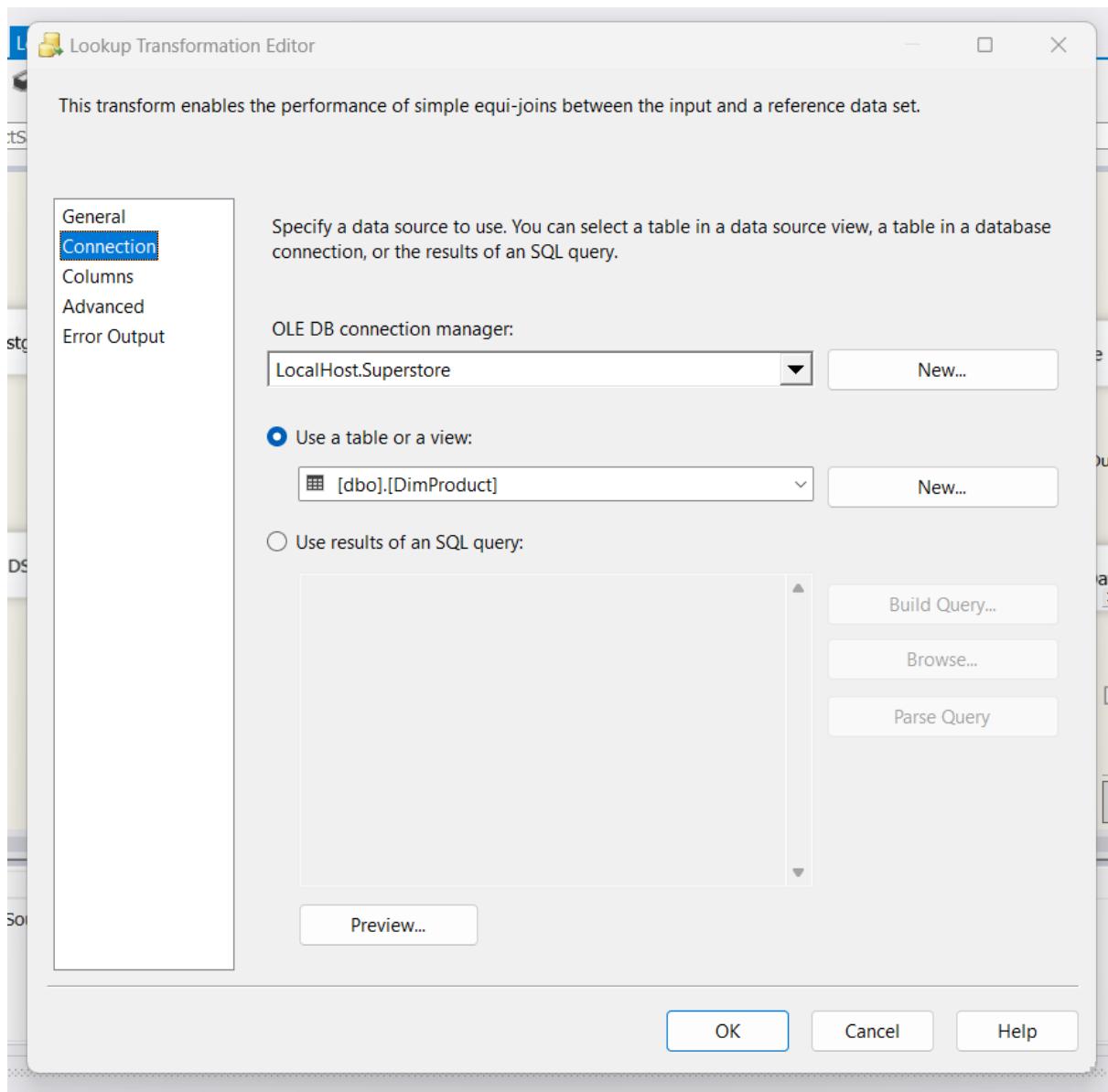
## b) Load dữ liệu vào Sales\_fact

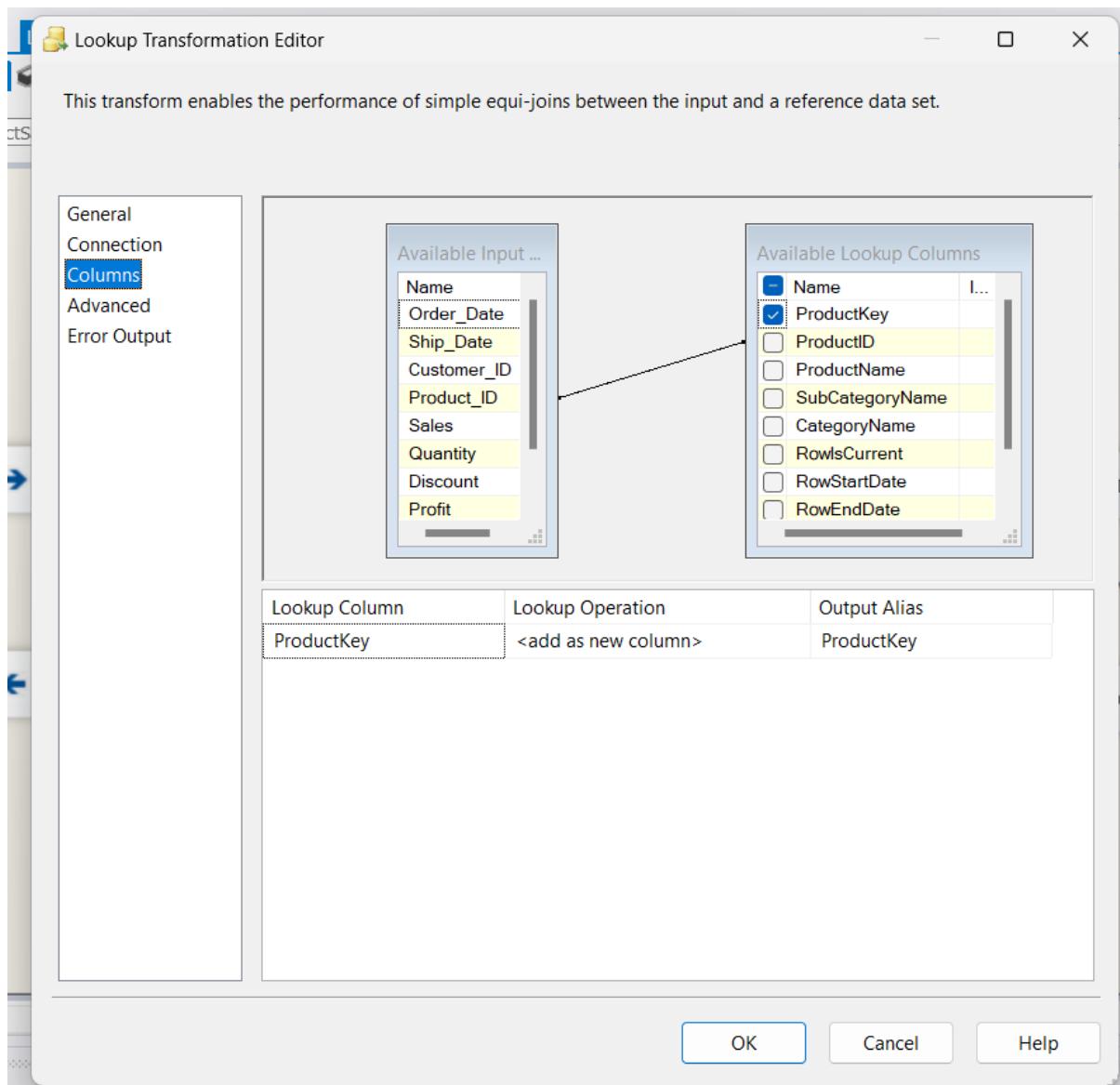


SRC – stgSales là bảng stgSales đã load từ trước

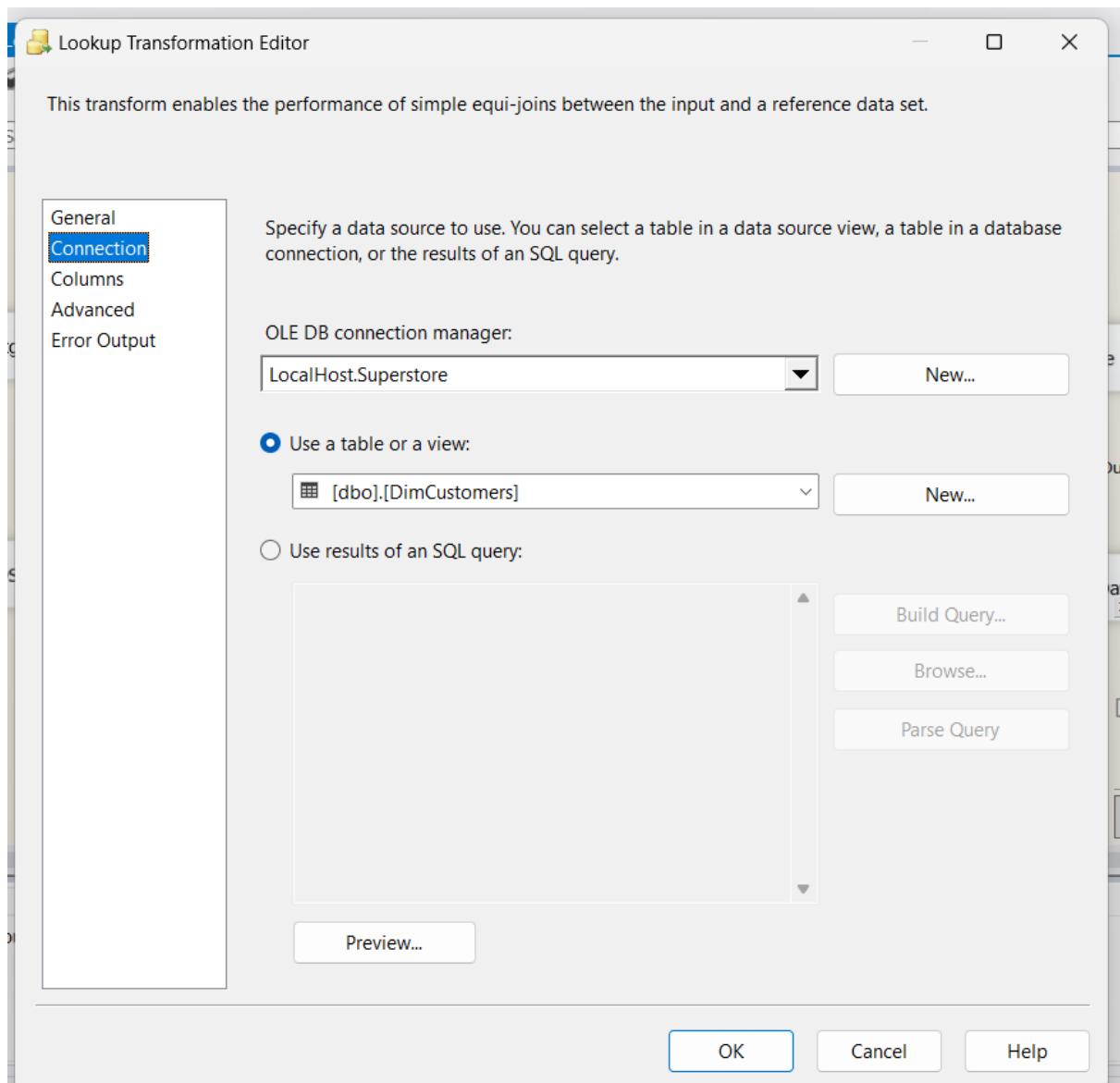


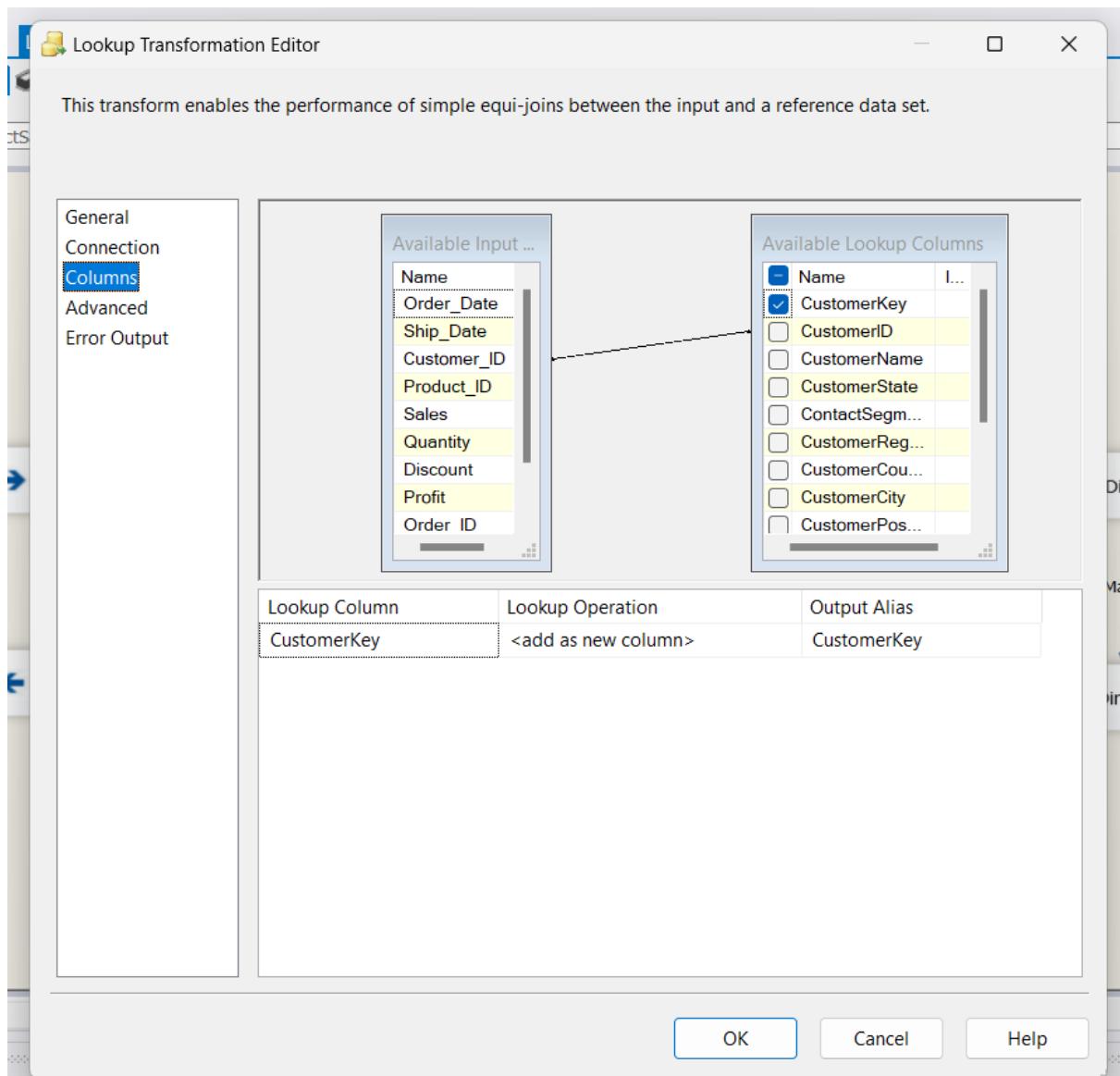
Lookup **DimProduct** ta sẽ map hai **ProductID** từ hai bảng lại với nhau và lấy ra thuộc tính **ProductKey** trong bảng DimProduct



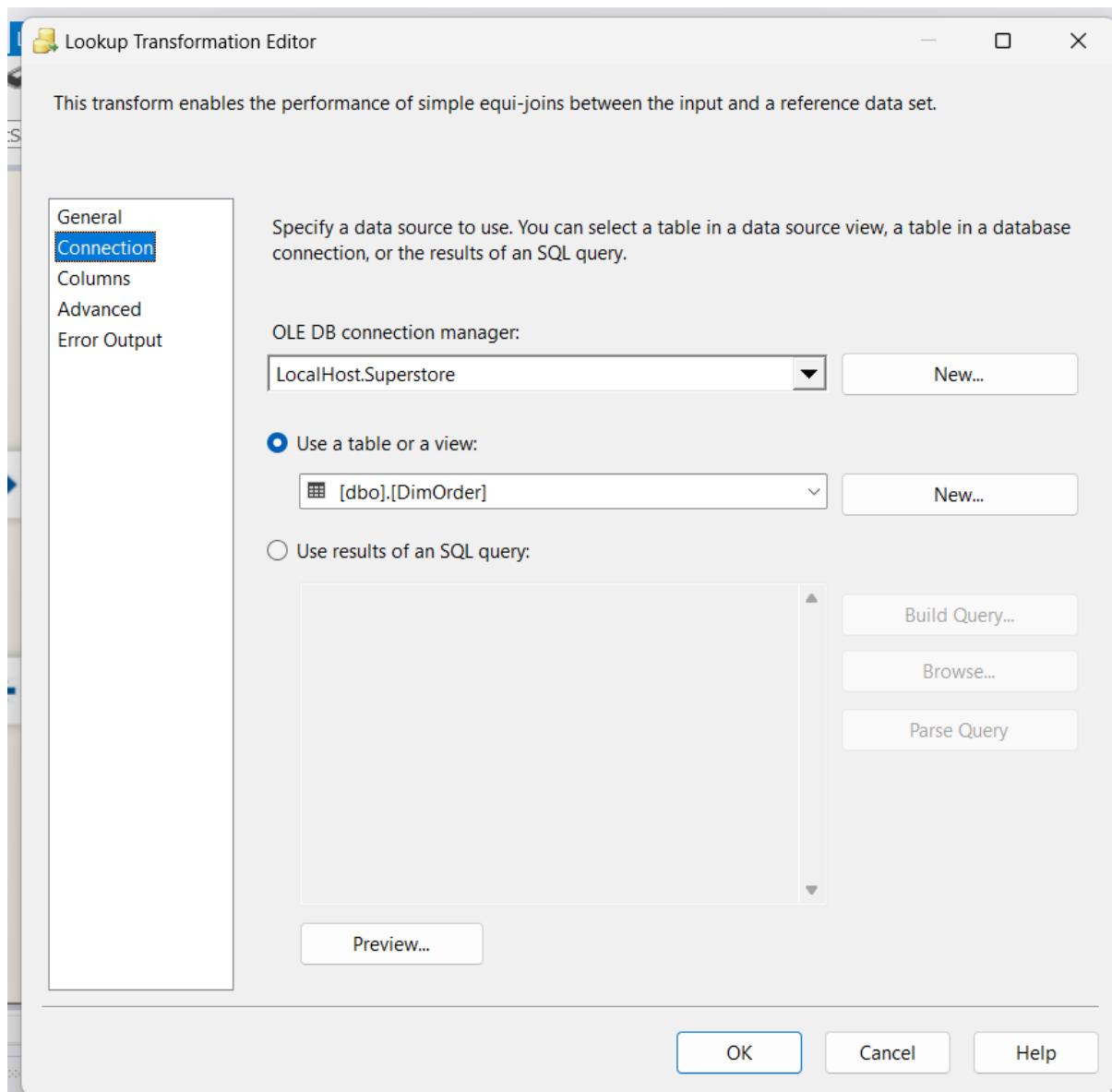


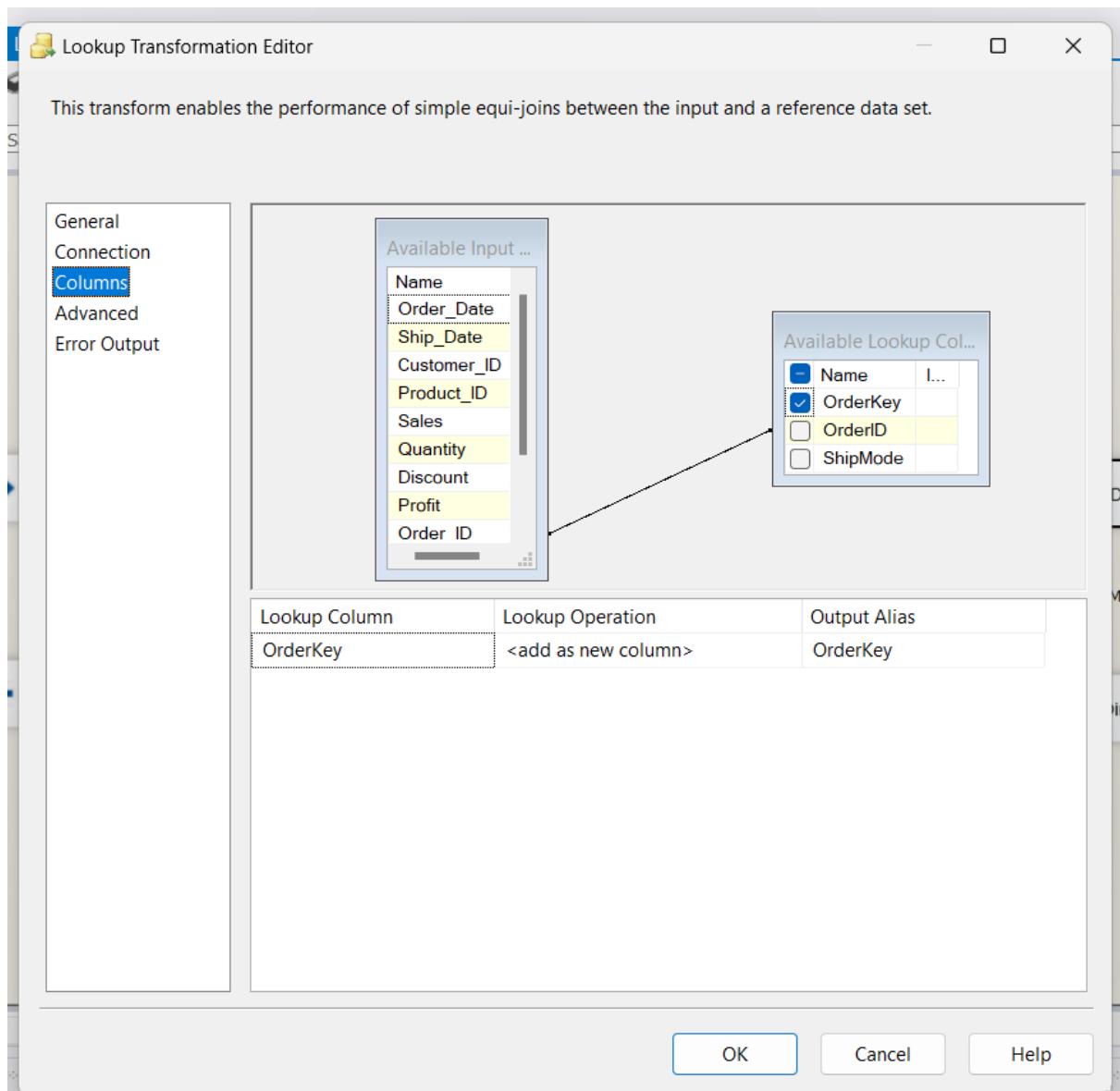
Lookup **DimCustomer** ta sẽ map cột customer\_id trong bảng stgSales với cột CustomerID trong bảng DimCustomer để có thể lấy ra cột CustomerKey



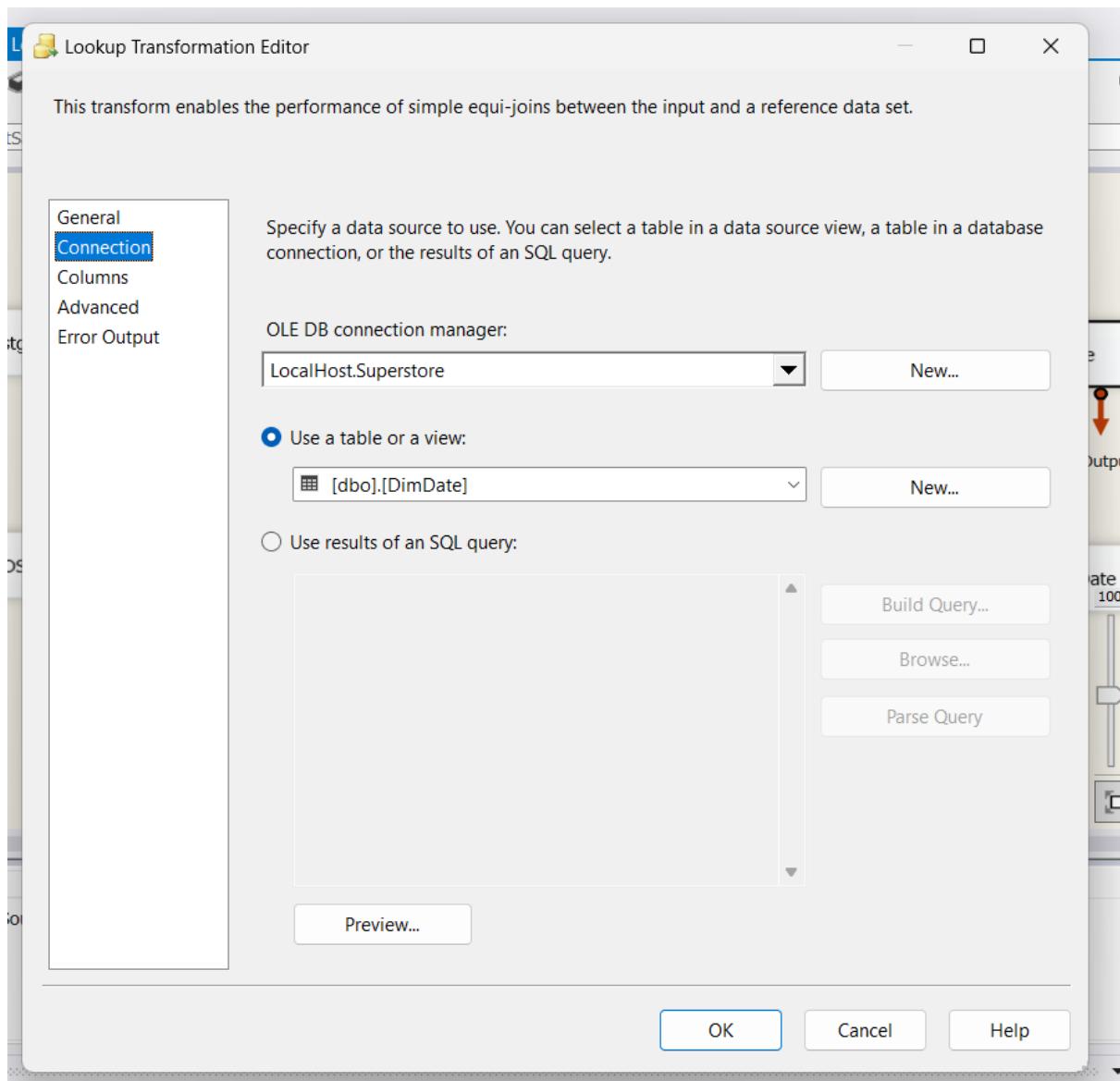


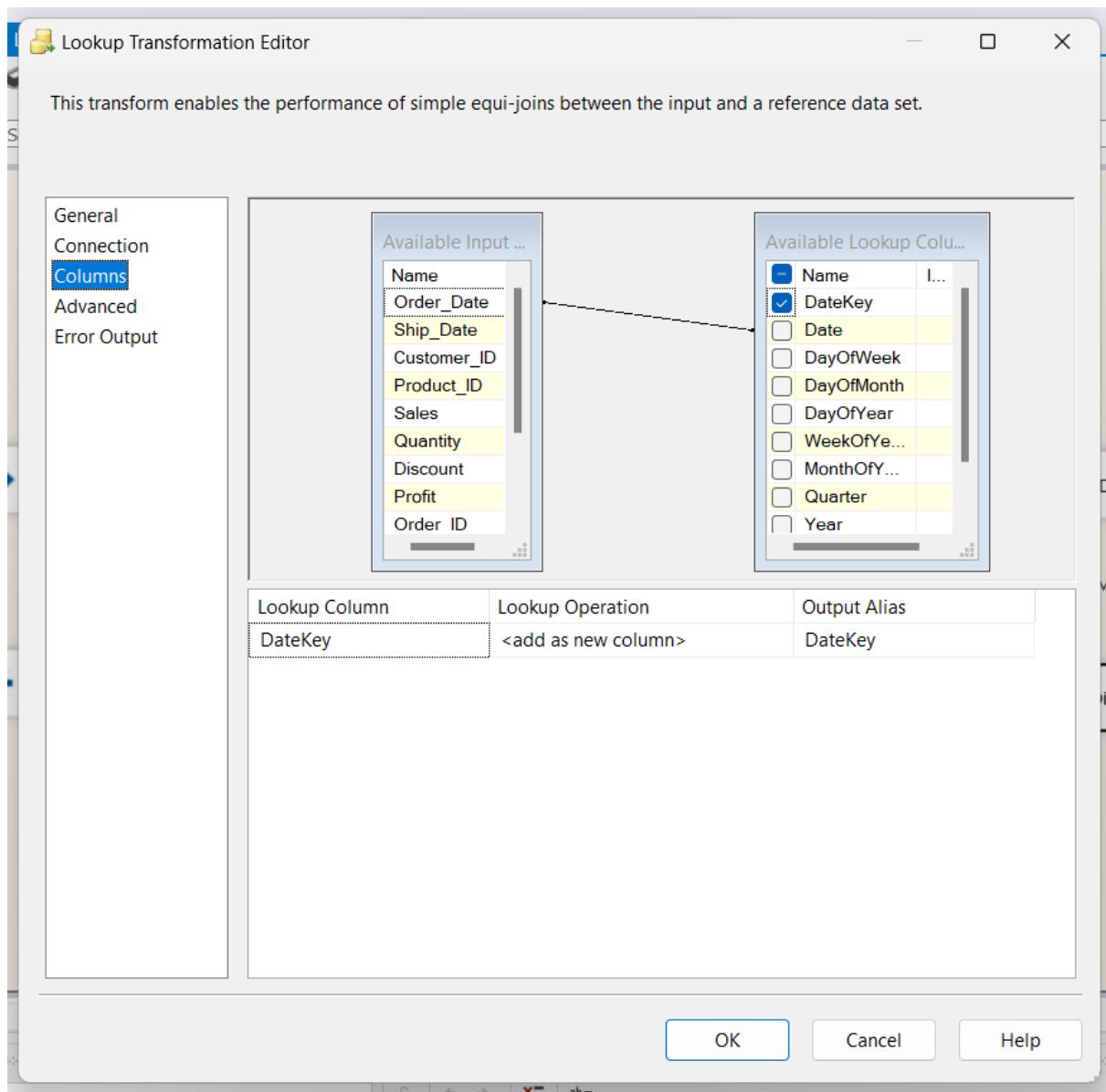
Lookup **DimOrder** ta sẽ map cột order\_id trong bảng stgOrder với cột OrderID trong bảng DimOrder để có thể lấy ra cột OrderKey



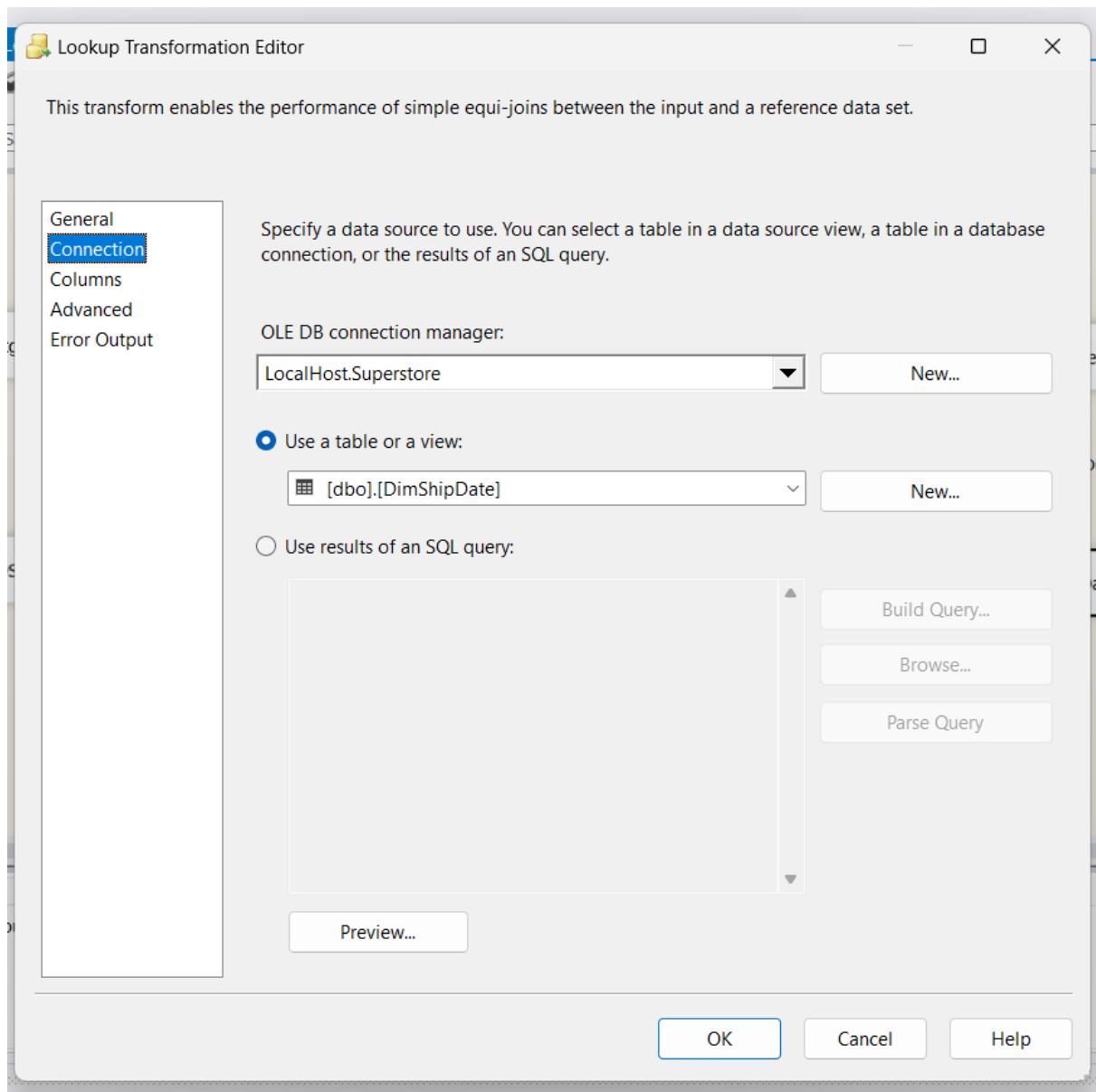


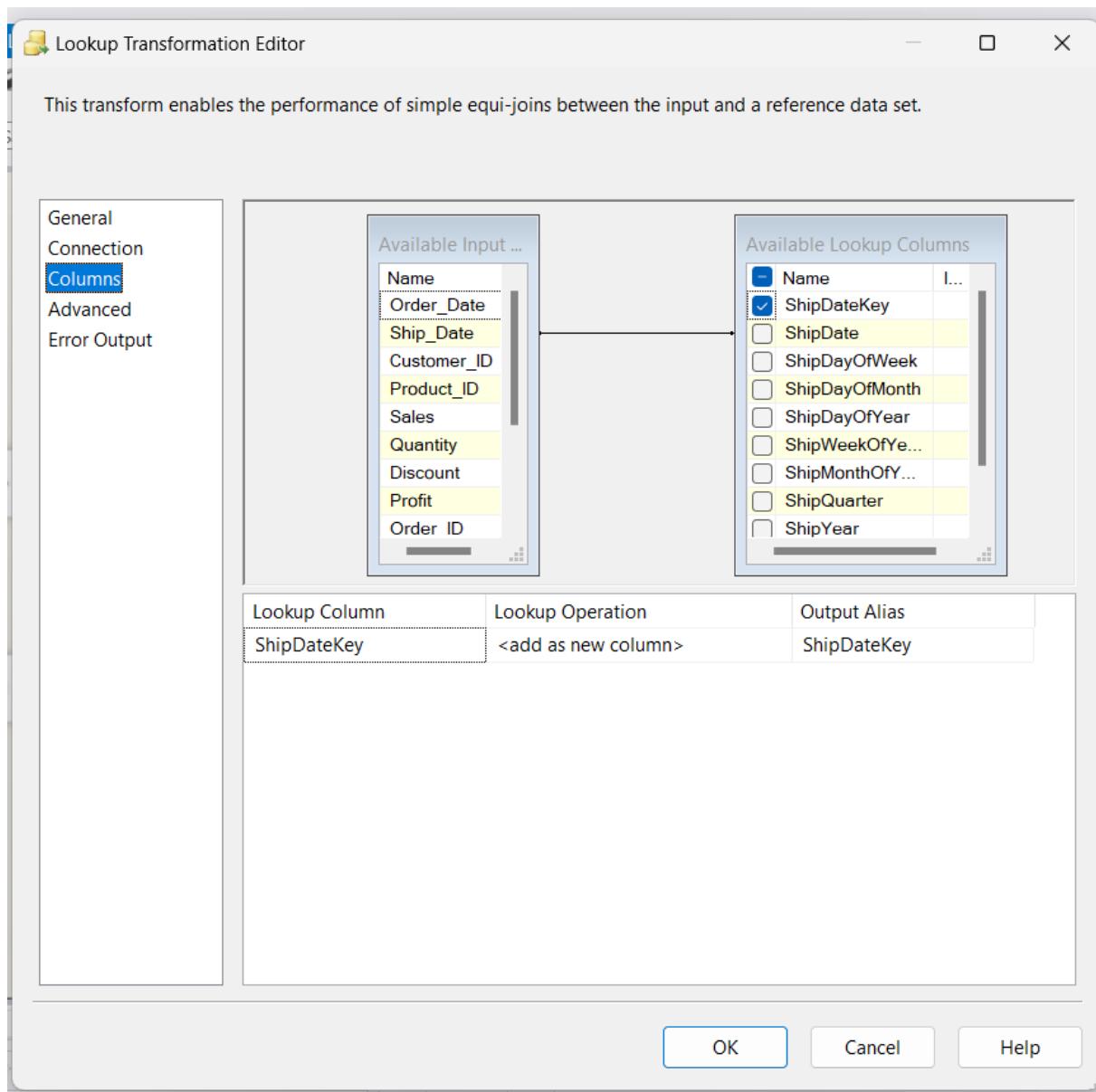
Lookup **DimDate** ta sẽ map cột order\_date ở bảng stgSales với cột Date trong bảng DimDate để lấy ra DateKey



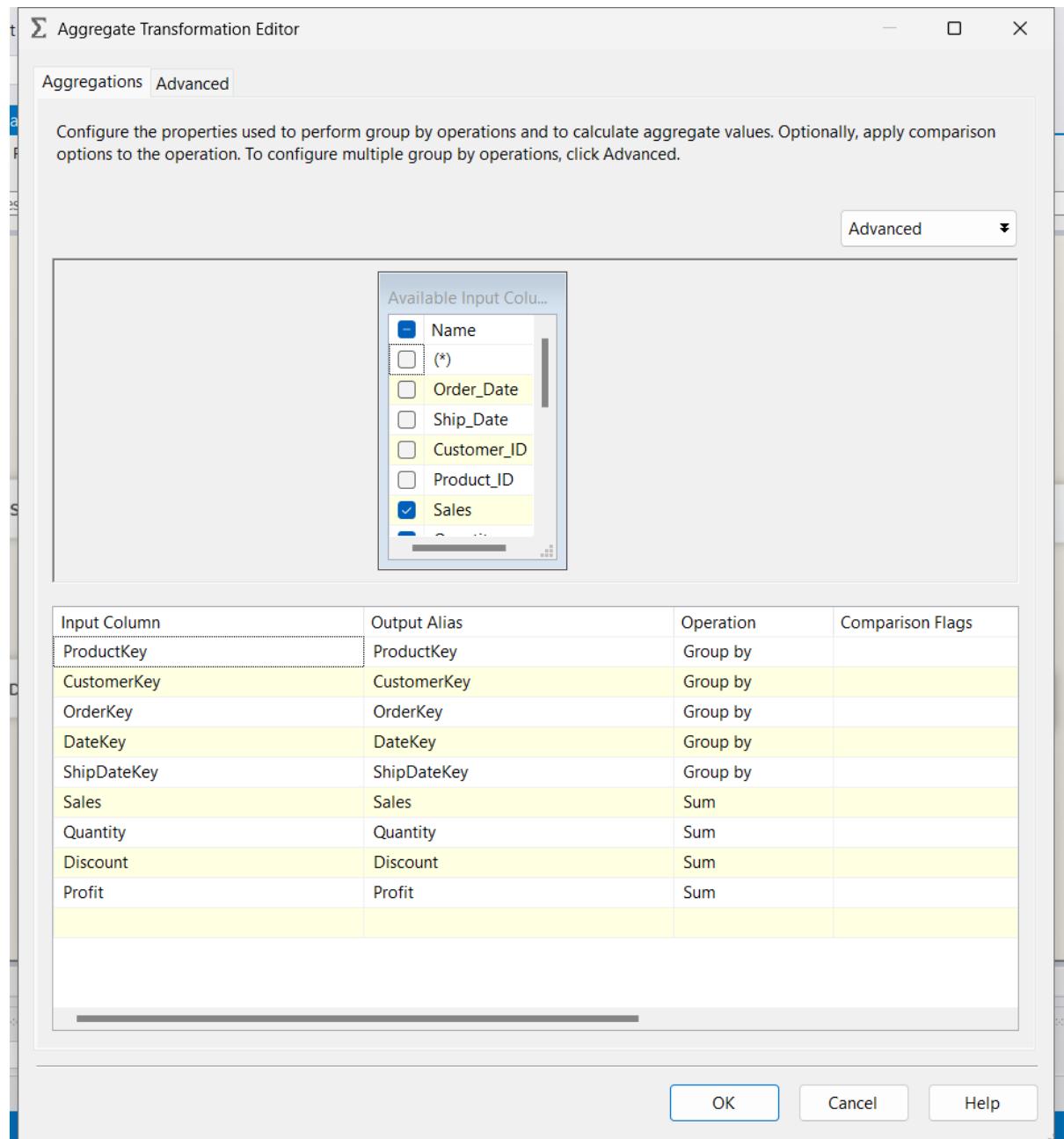


Lookup **DimShipDate** ta sẽ map cột ship\_date ở bảng stgSales với cột ShipDate trong bảng DimShipDate để lấy ra ShipDateKey

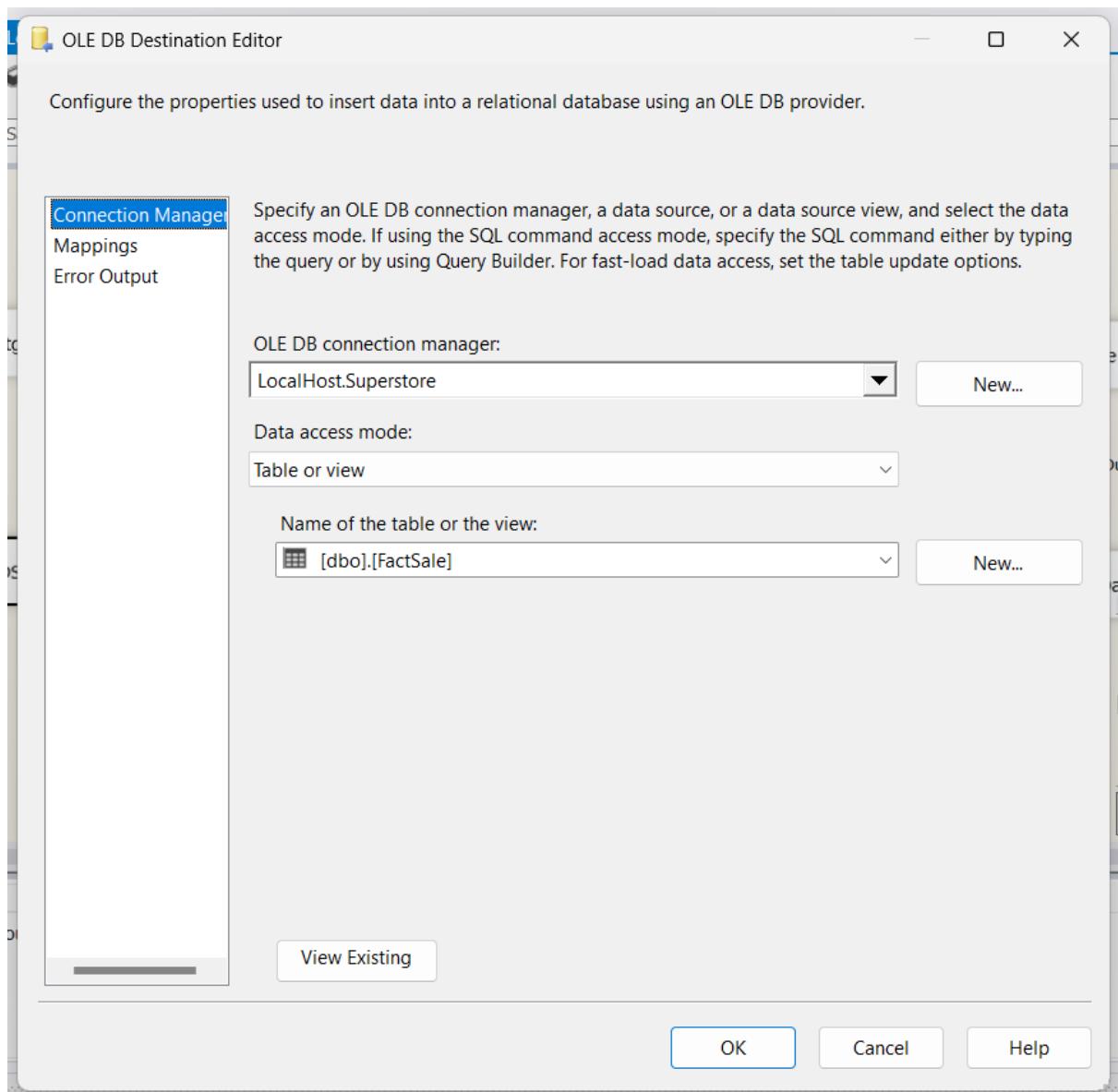


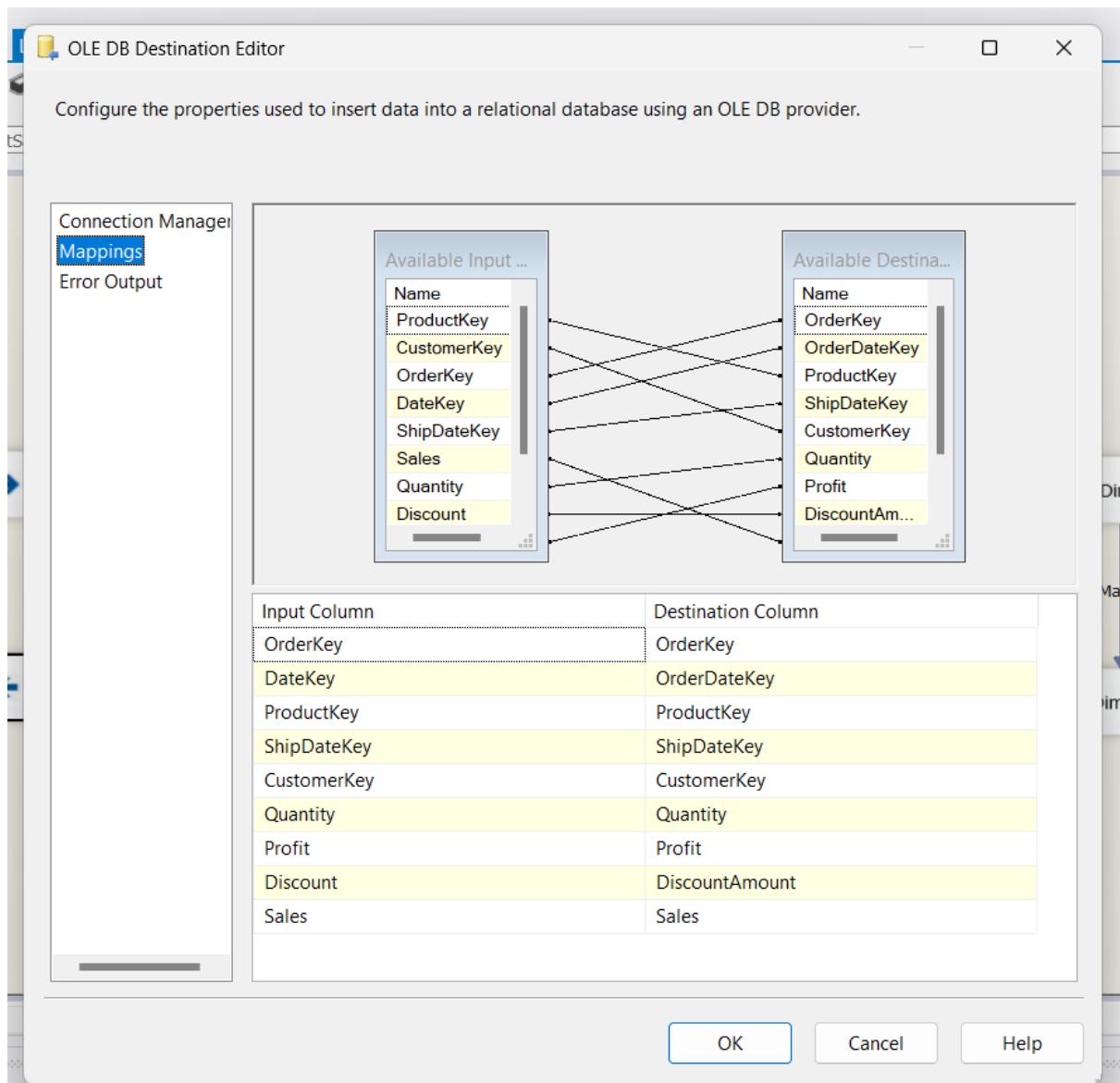


**Aggregate** Ta sẽ lấy những cột cần thiết để đưa vào bảng fact

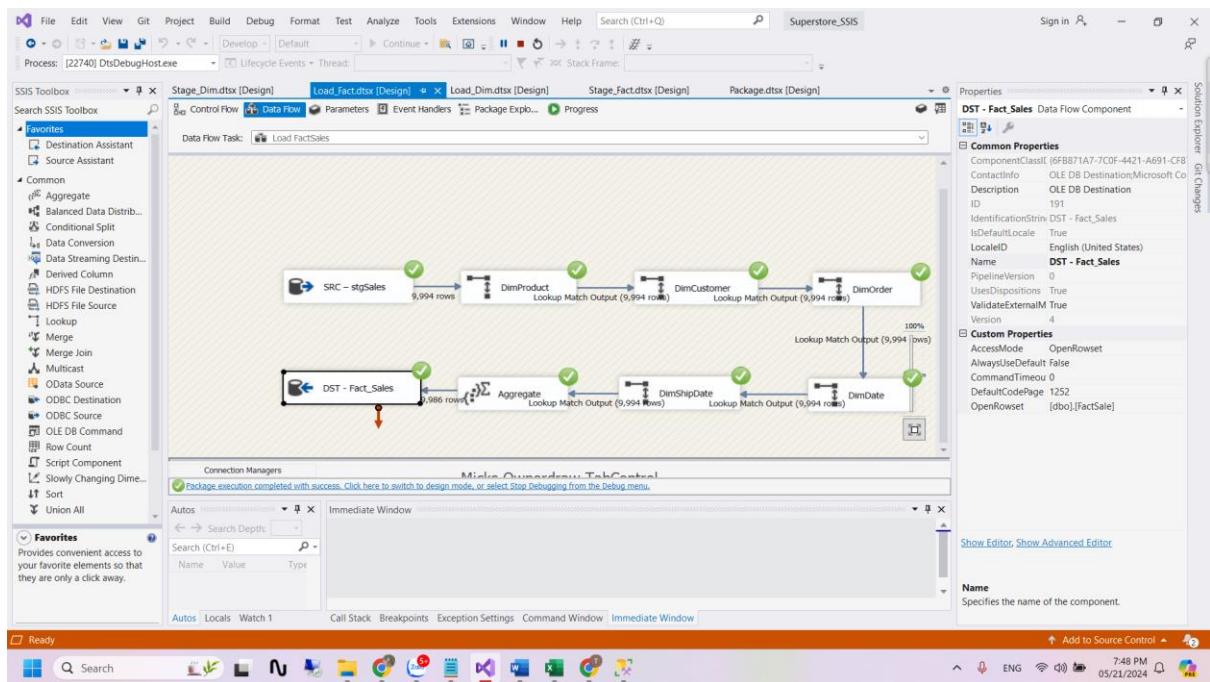


## DST - Fact\_Sales



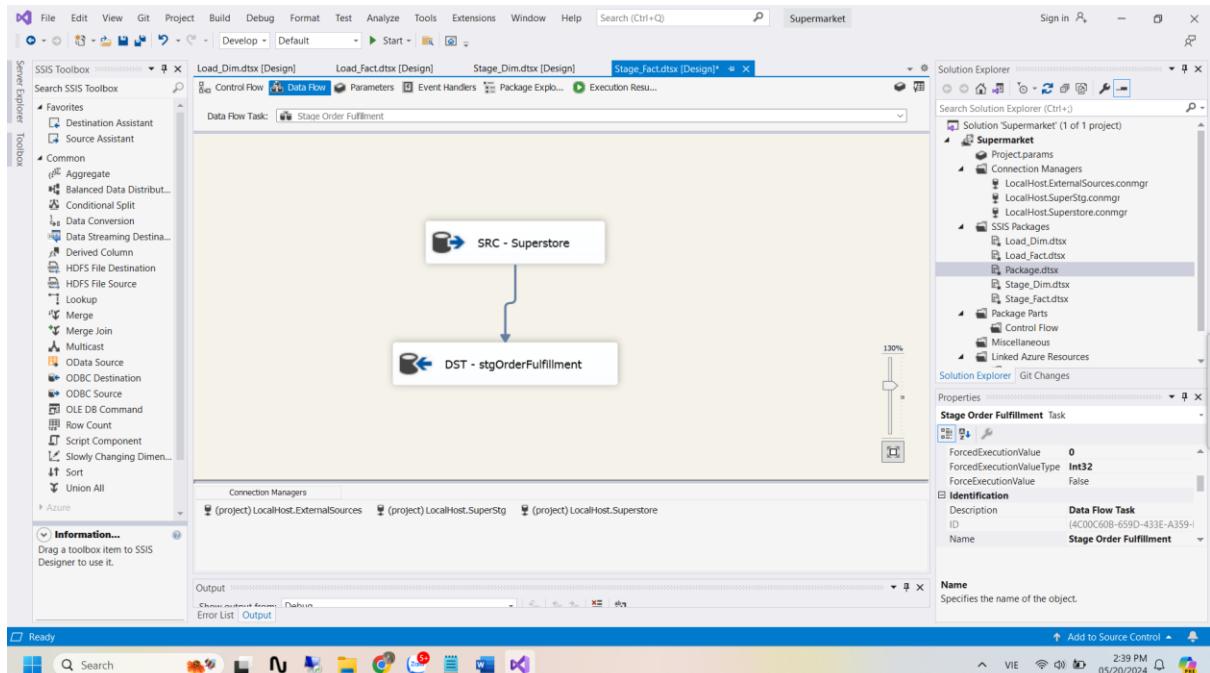


**Khởi chạy**

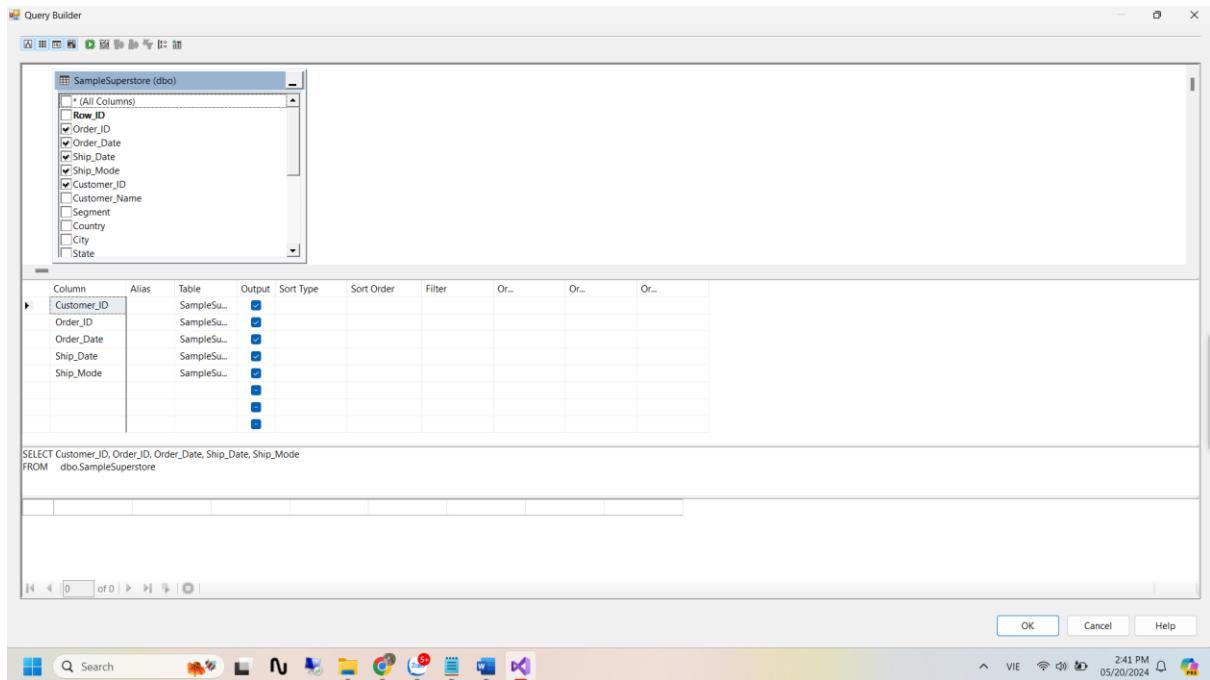


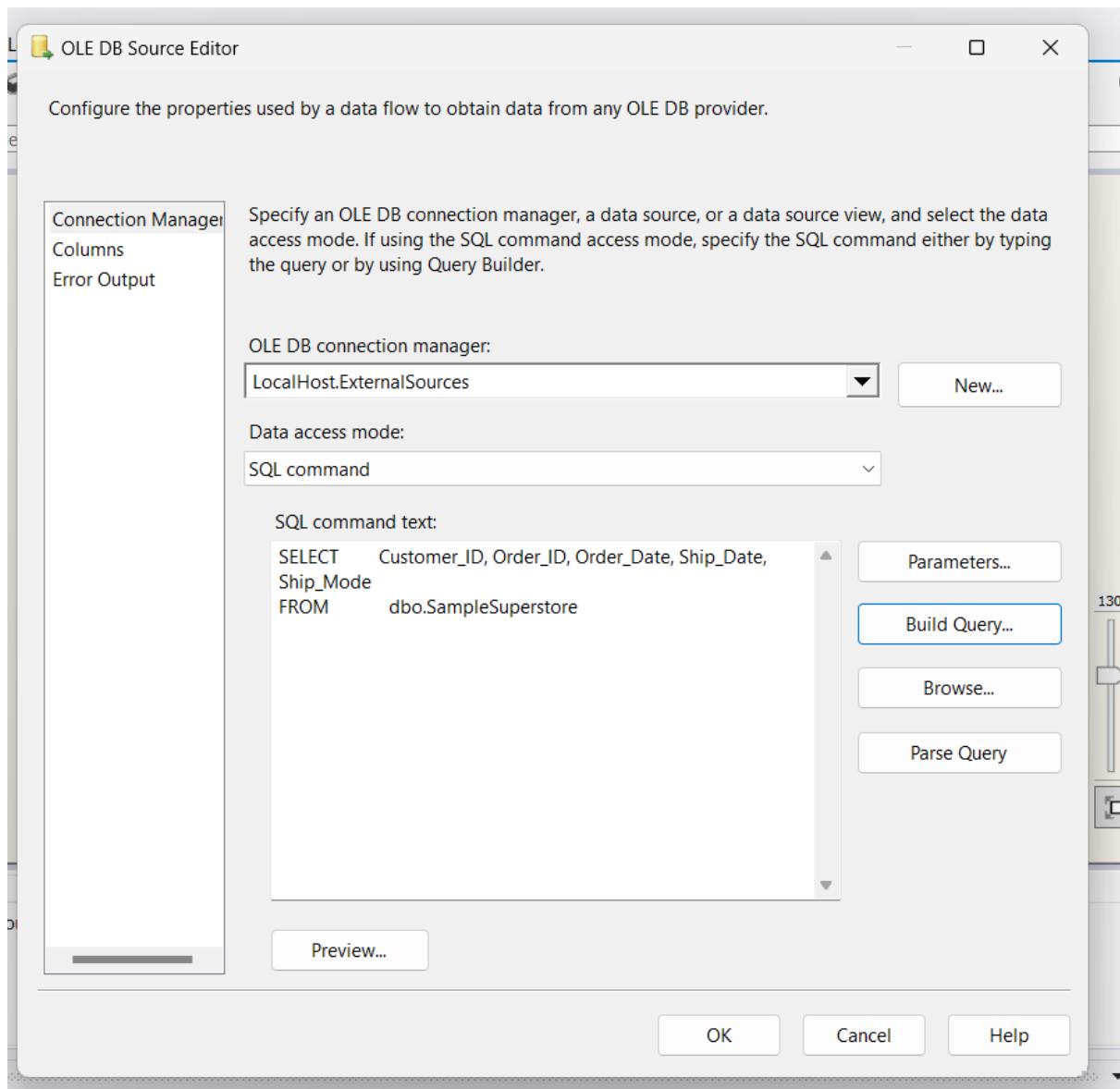
### 3.3.2. Fact OrderFulfillment

#### a) Load dữ liệu từ nguồn vào stgOrderFulfillment

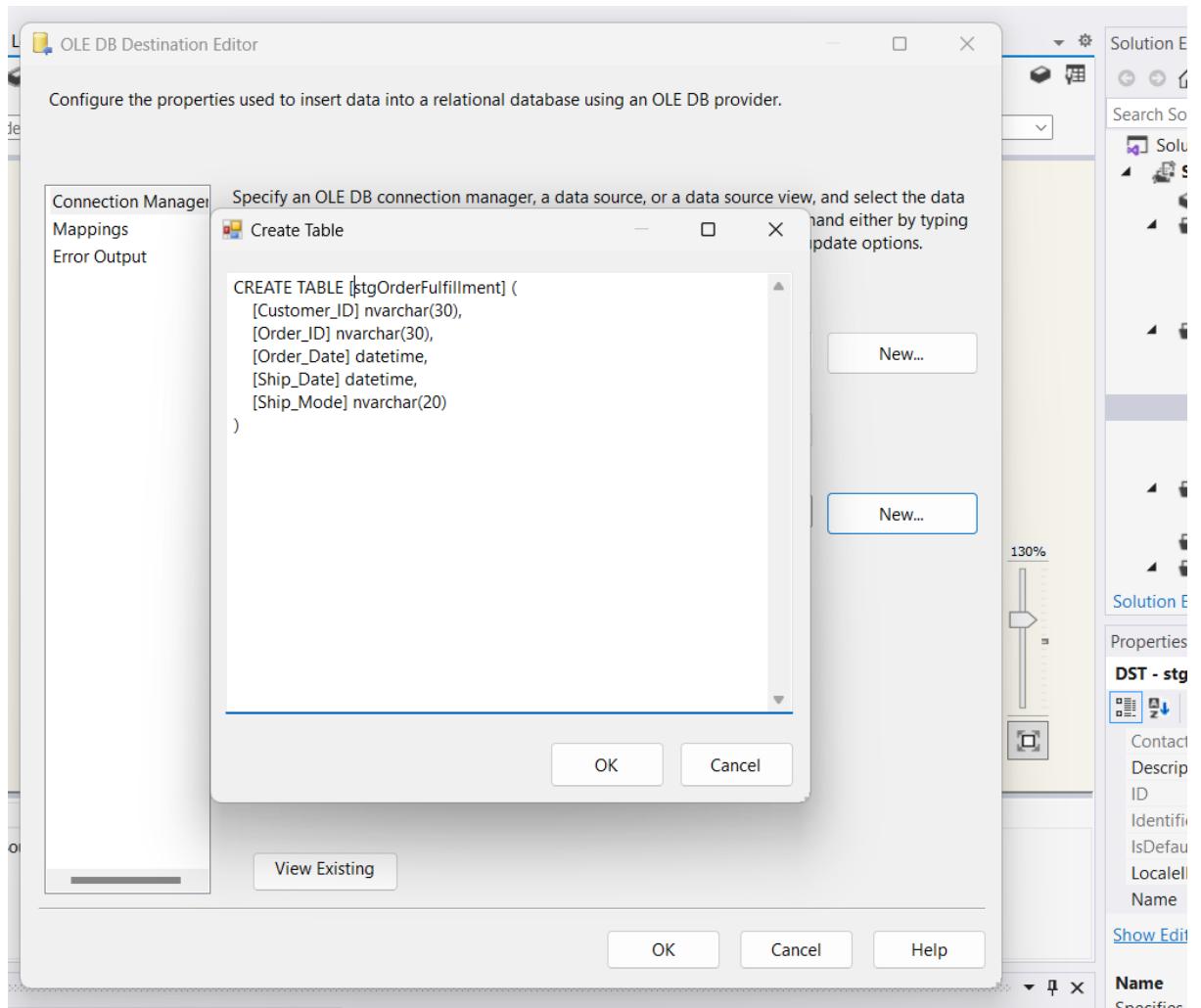


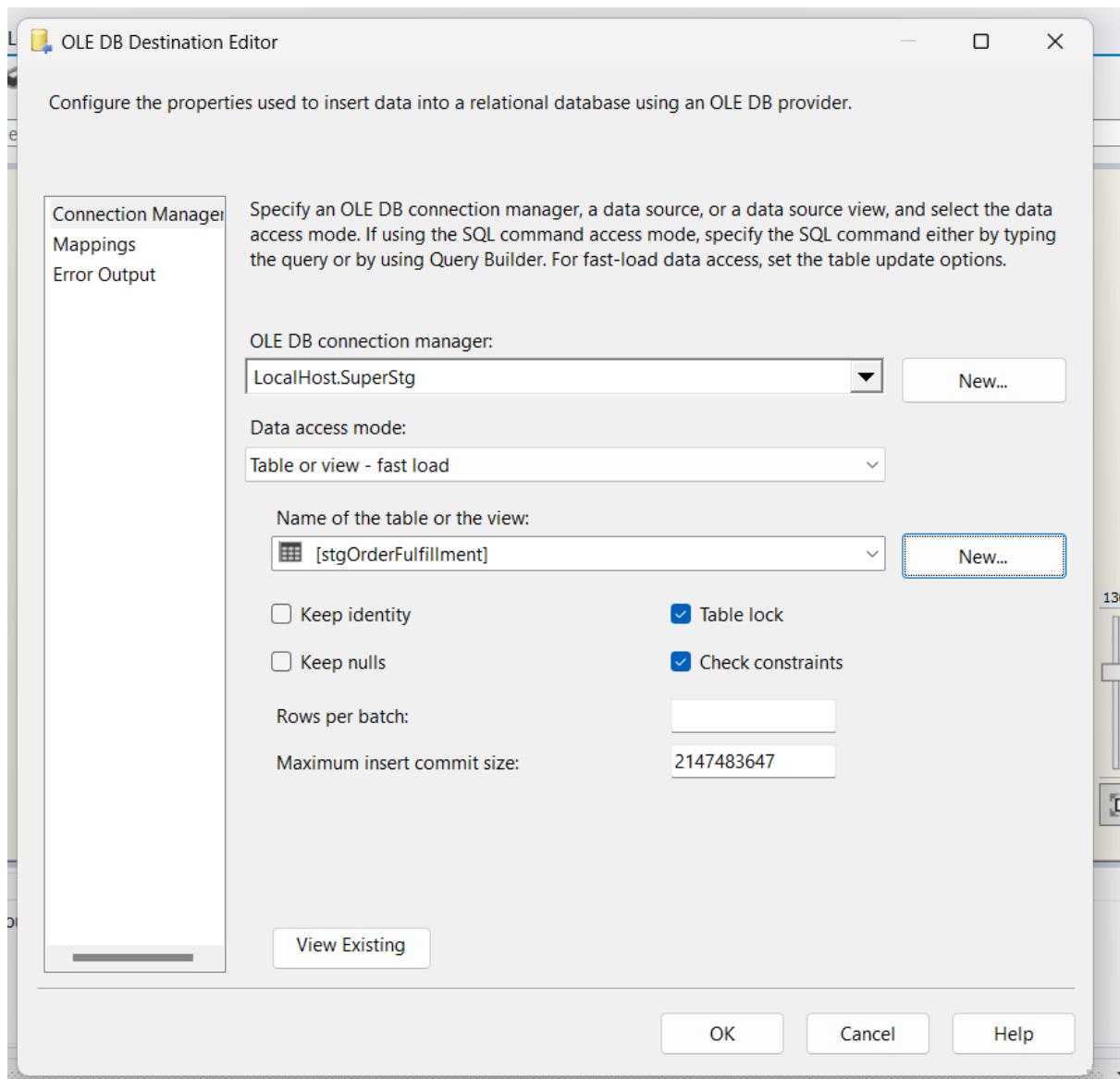
SRC – Superstore là dữ liệu nguồn ta lấy những thuộc tính cần thiết để đưa vào bảng stgOrderFulfillment

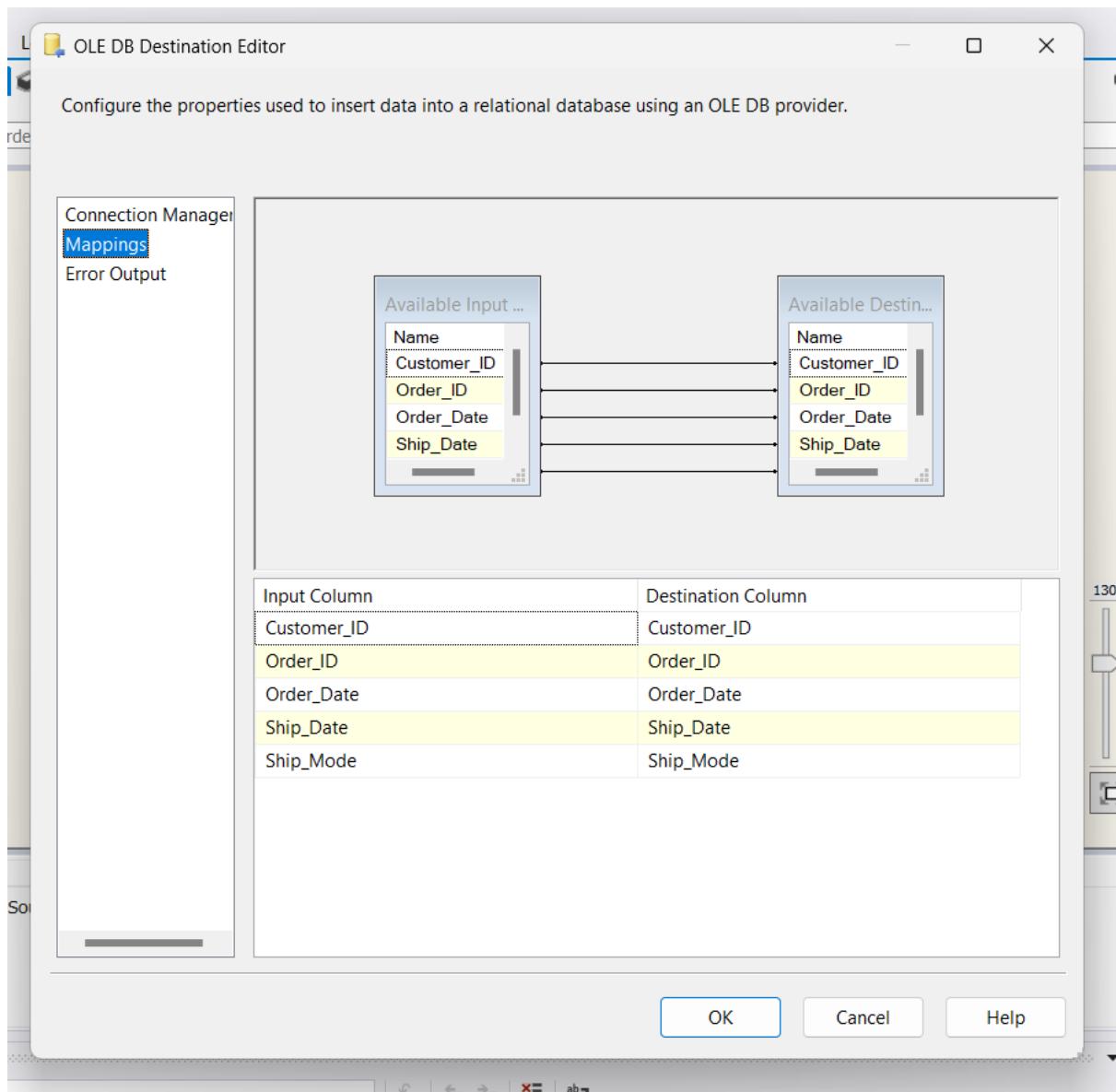


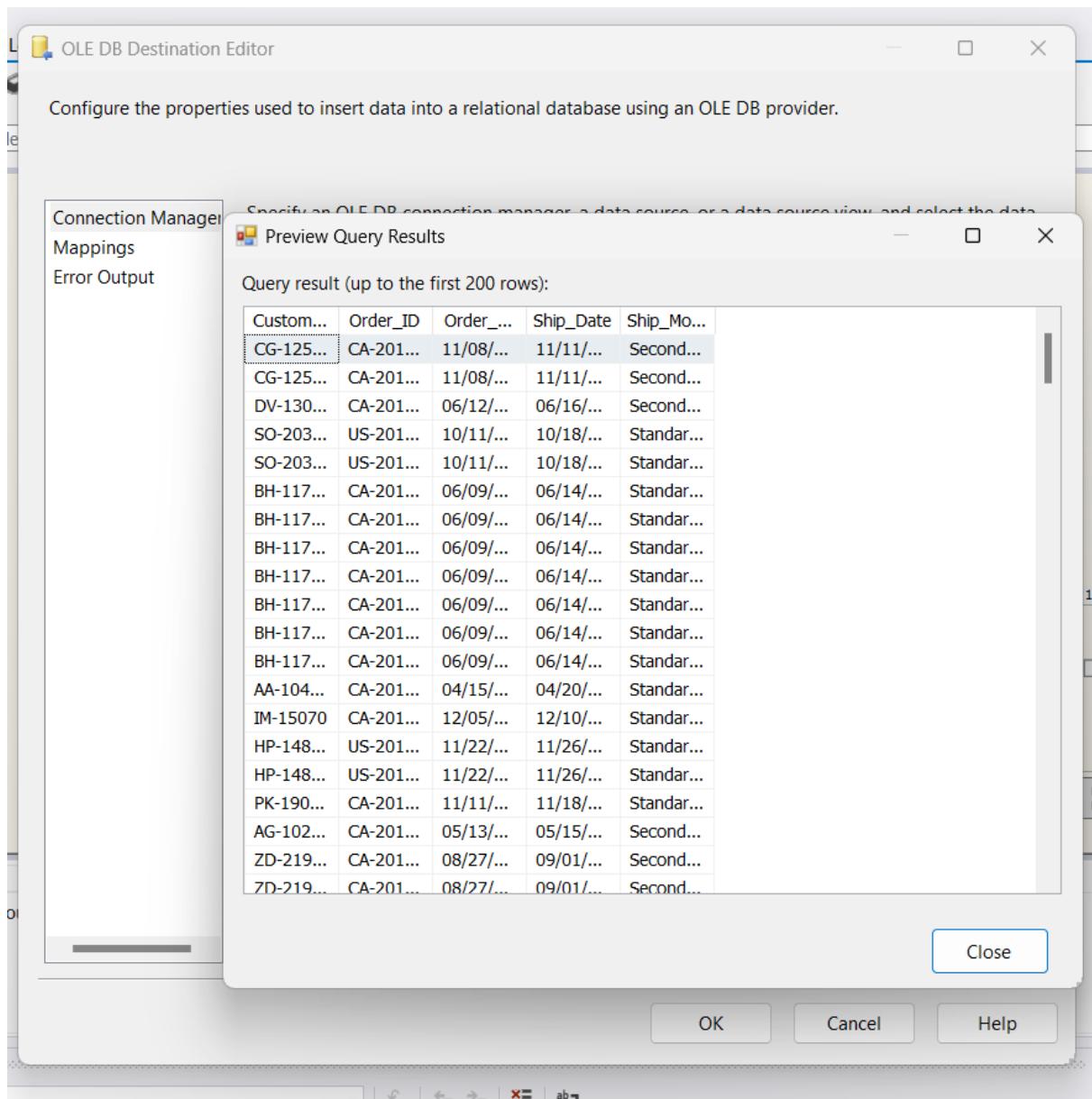


## DST – stgOrderFulfillment

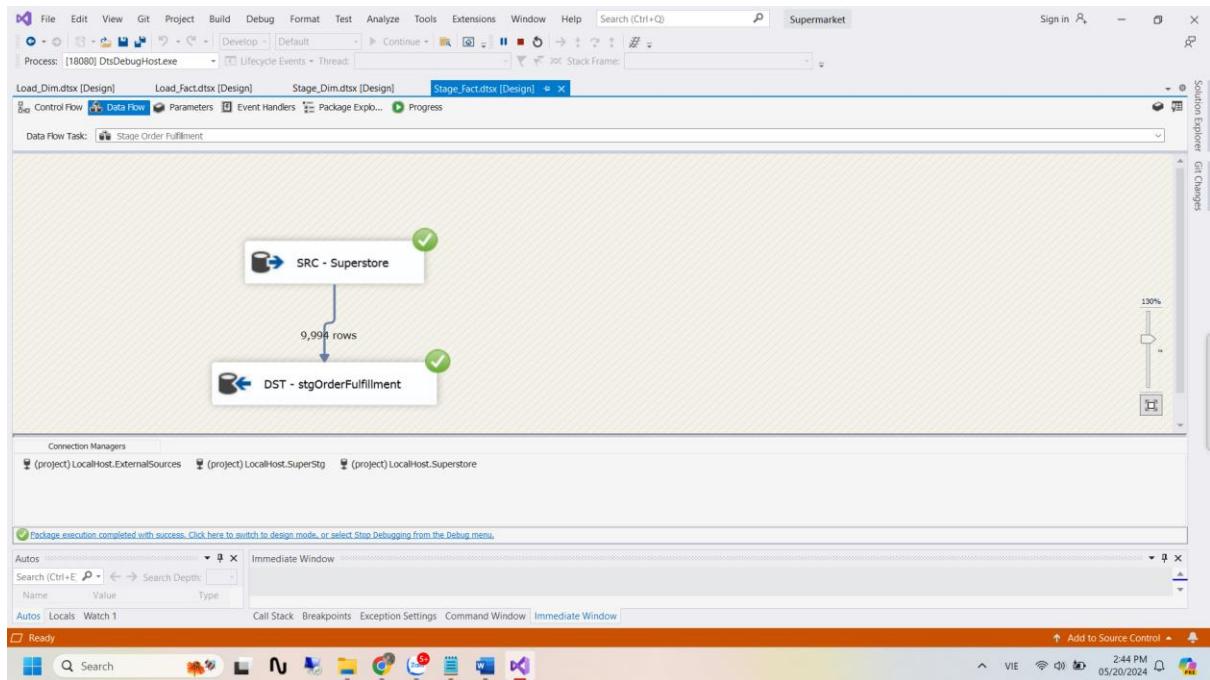




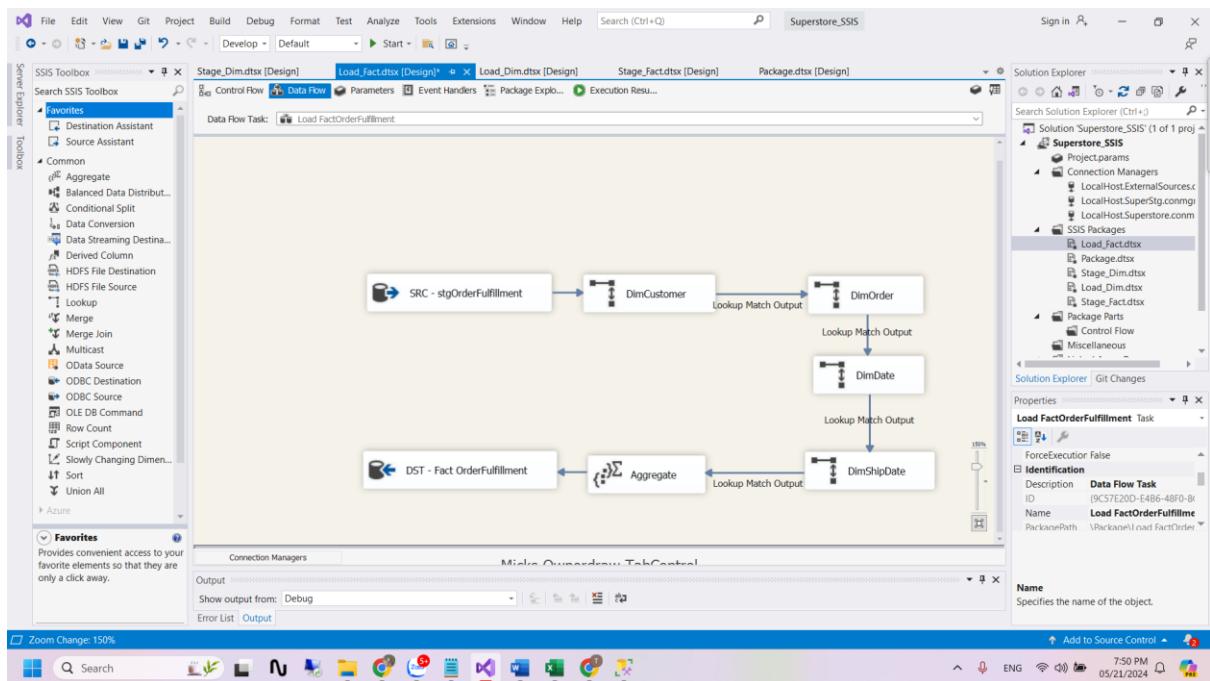




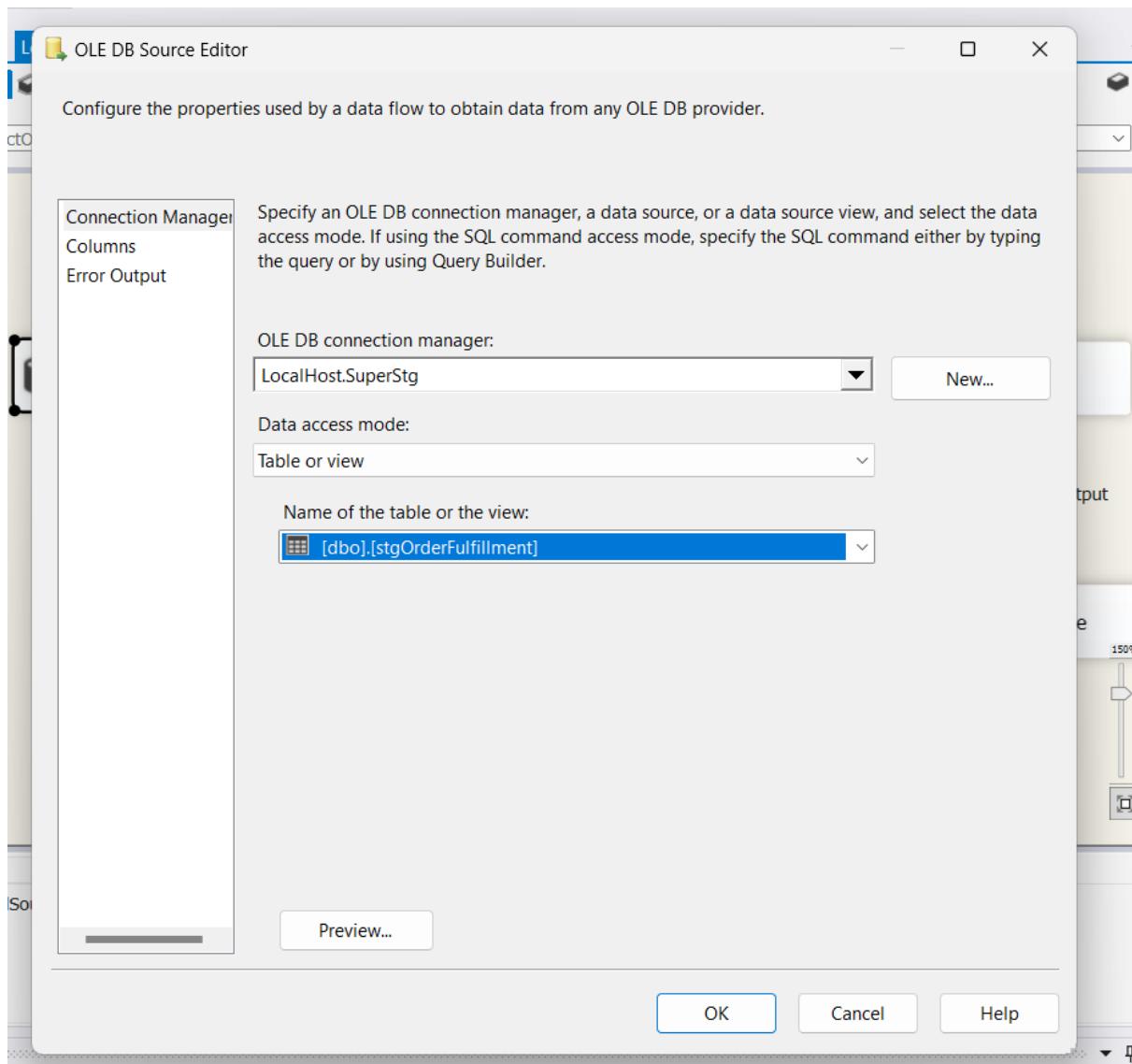
## Khởi chạy



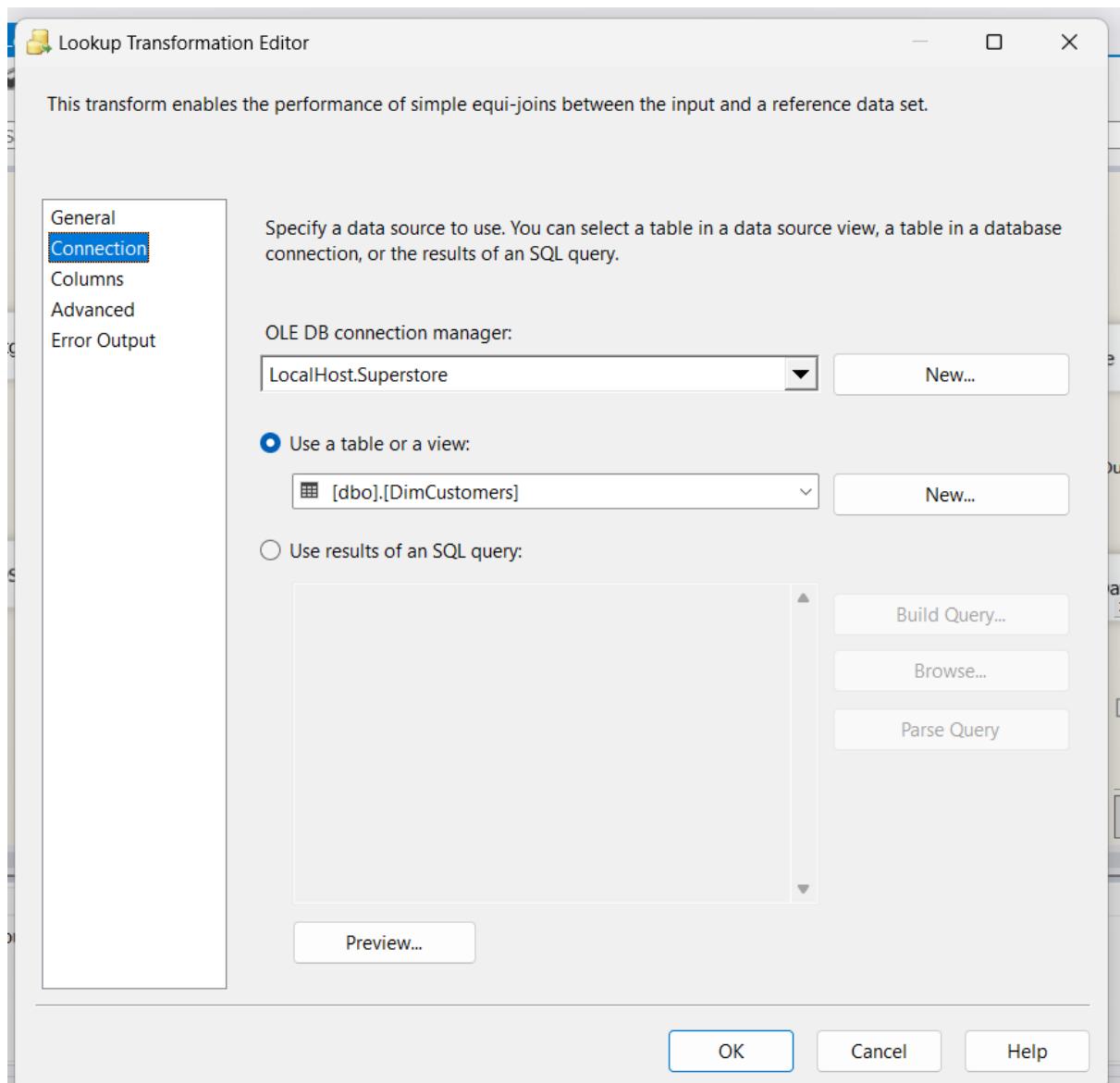
## b) Load dữ liệu vào Fact OrderFulfillment

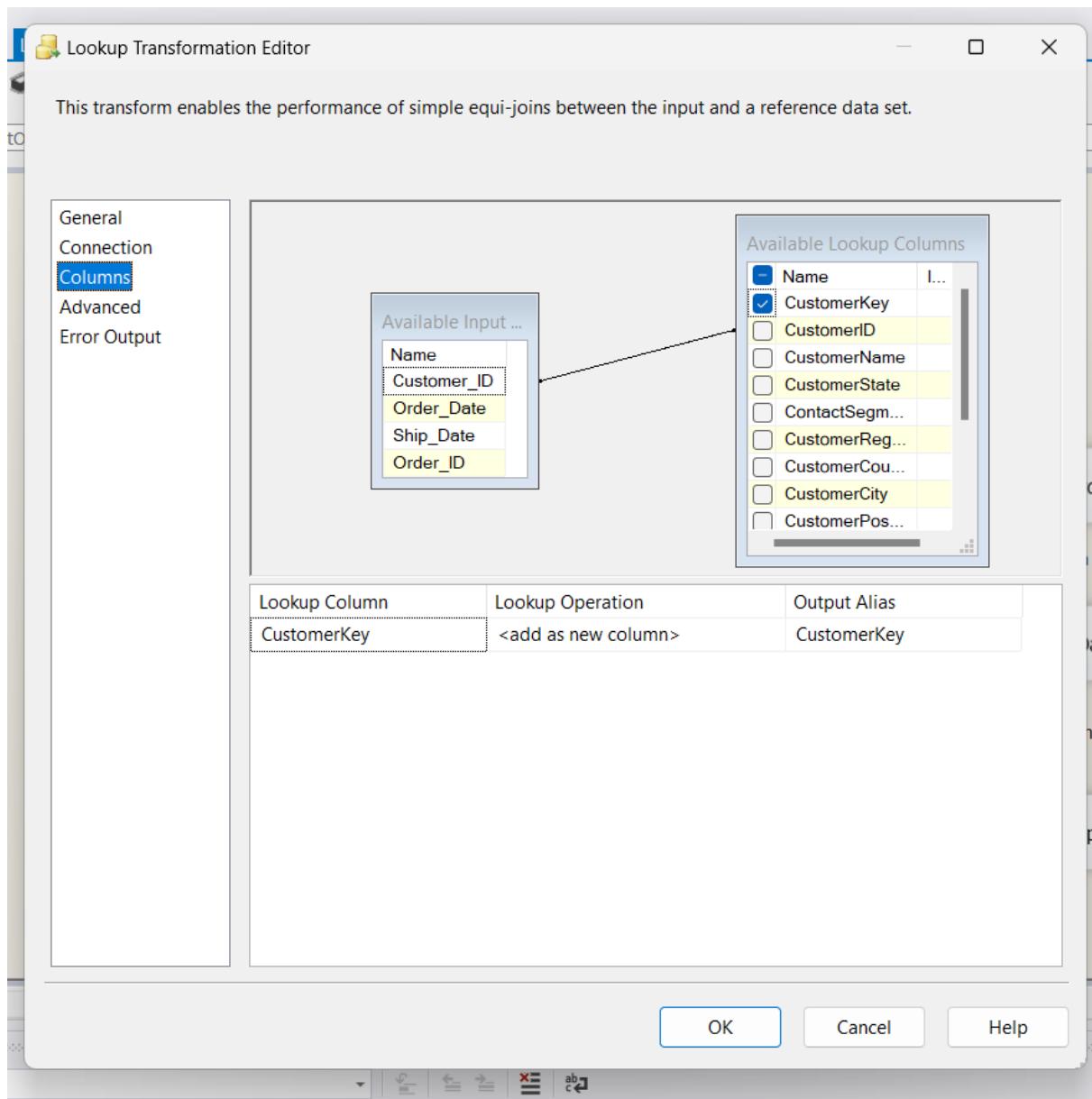


**SRC - stgOrderFulfillment** là bảng **stgOrderFulfillment**

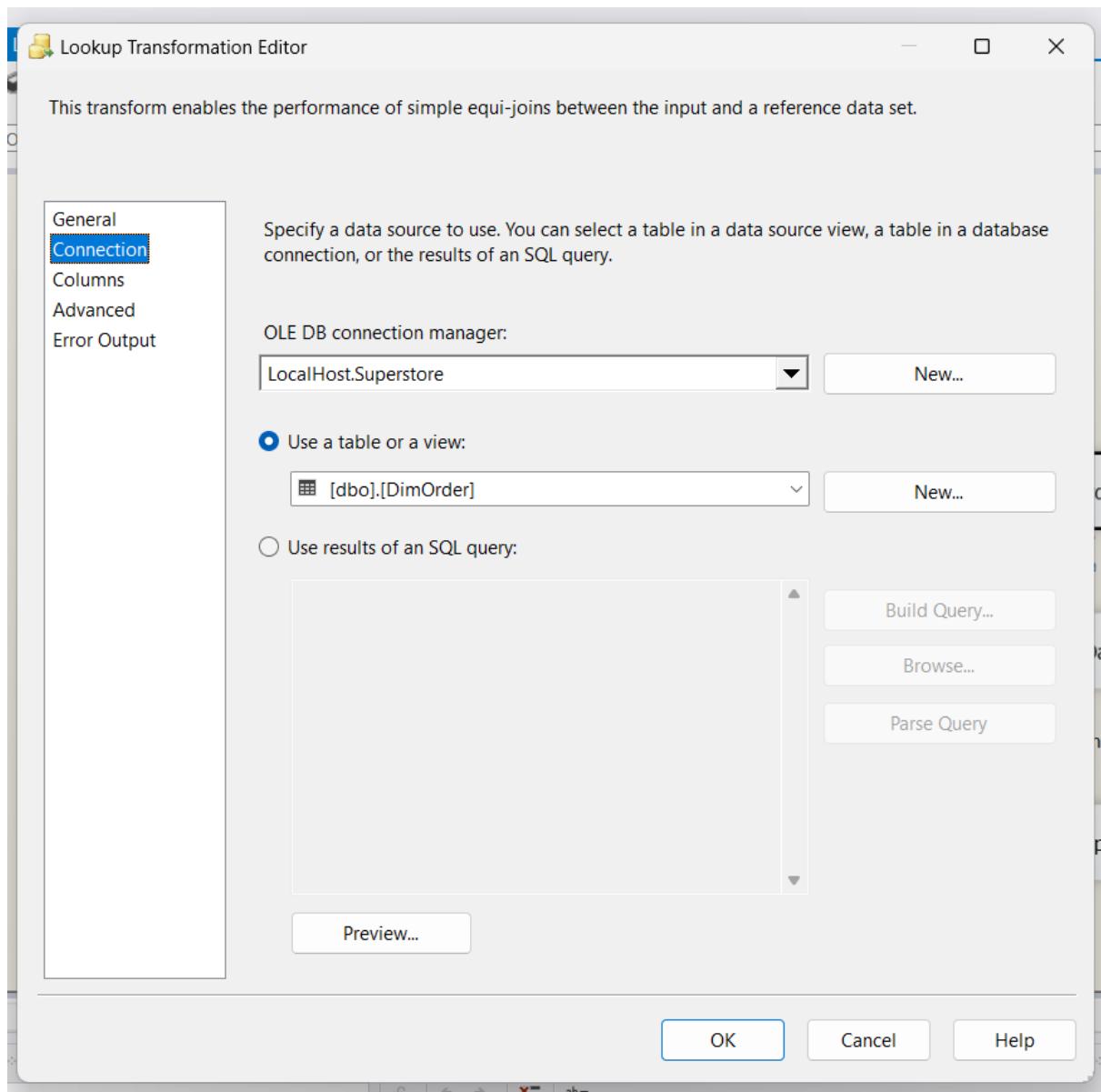


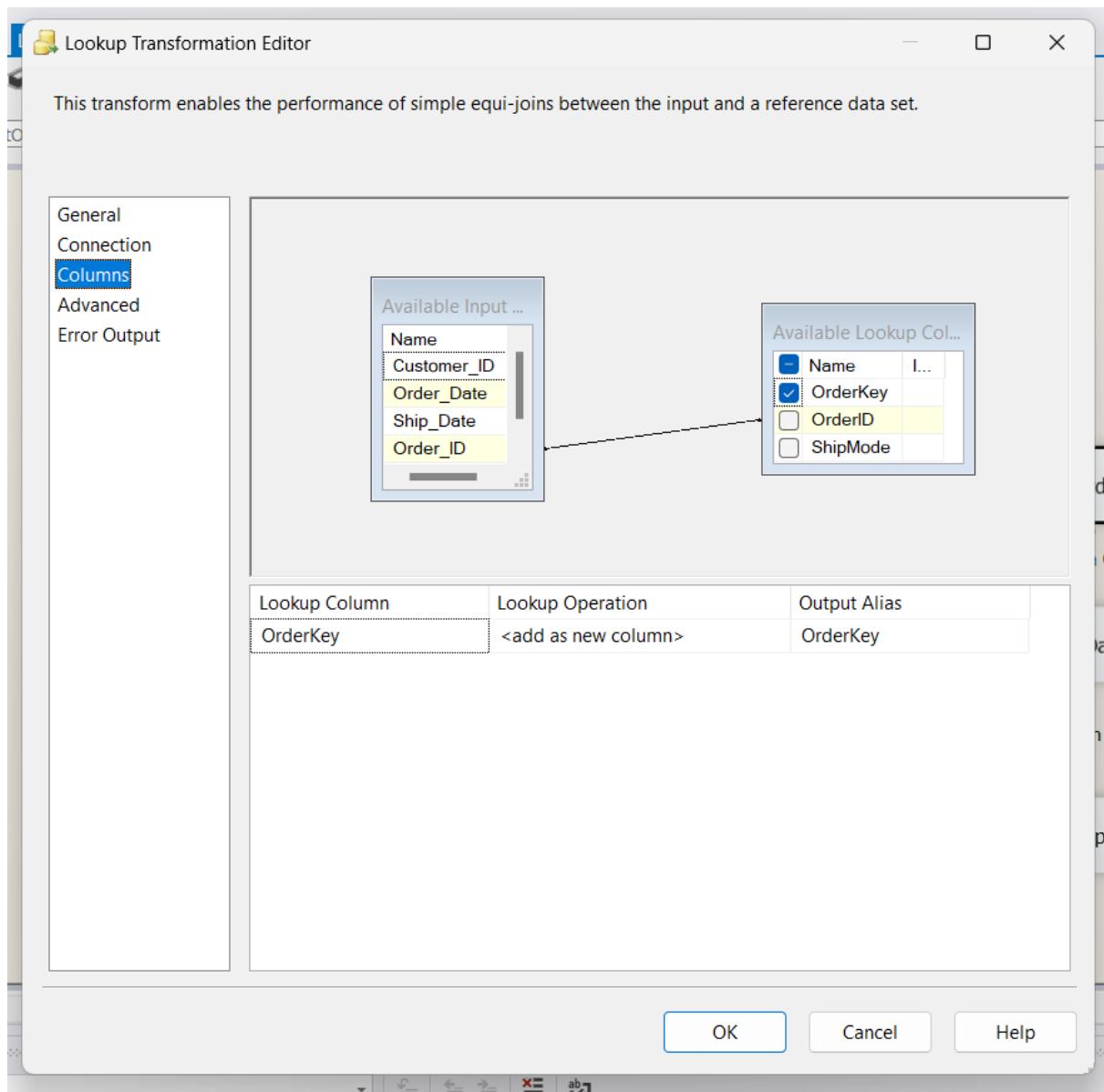
Lookup **DimCustomer** ta sẽ map cột customer\_id trong bảng stgSales với cột CustomerID trong bảng DimCustomer để có thể lấy ra cột CustomerKey



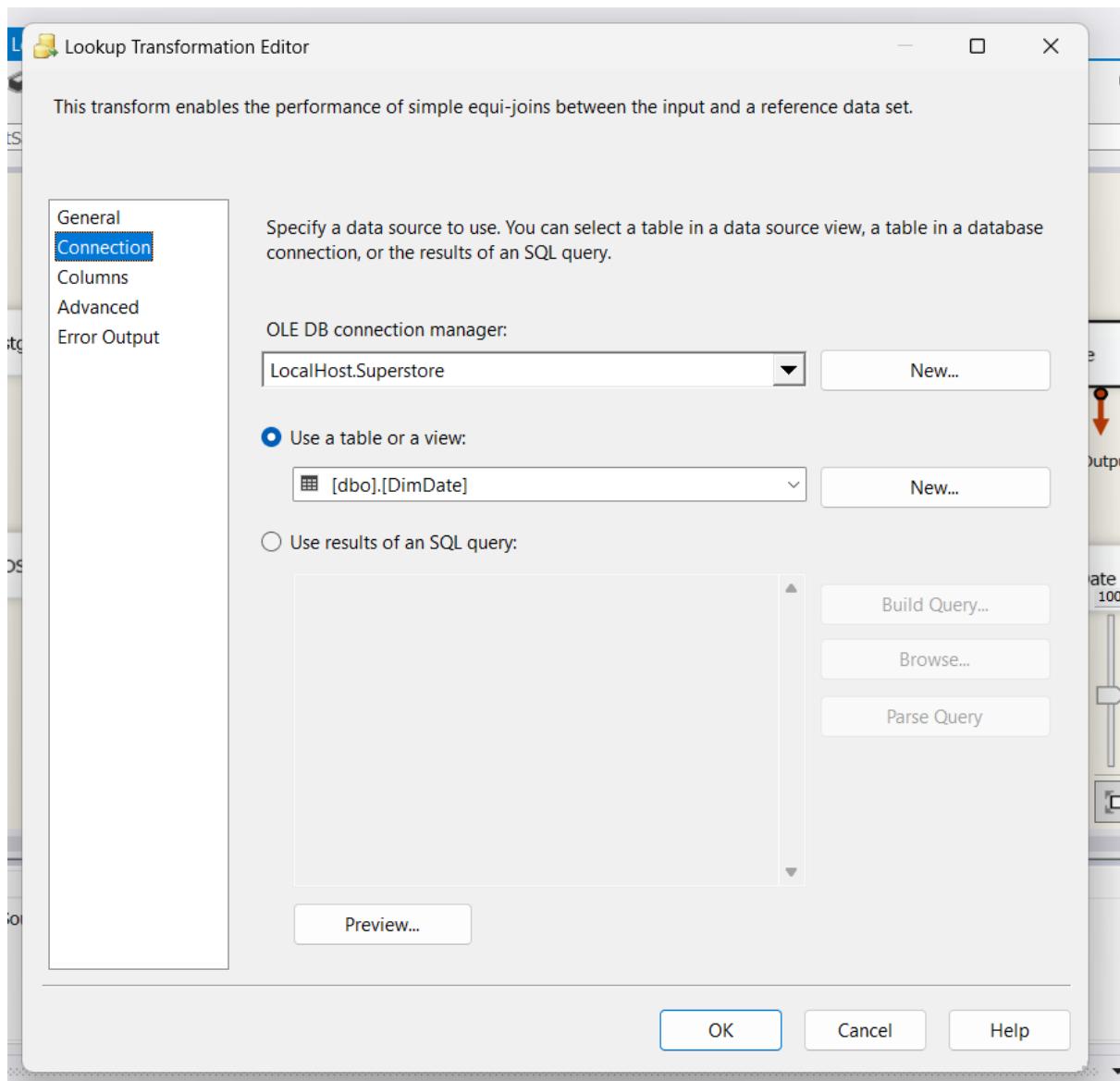


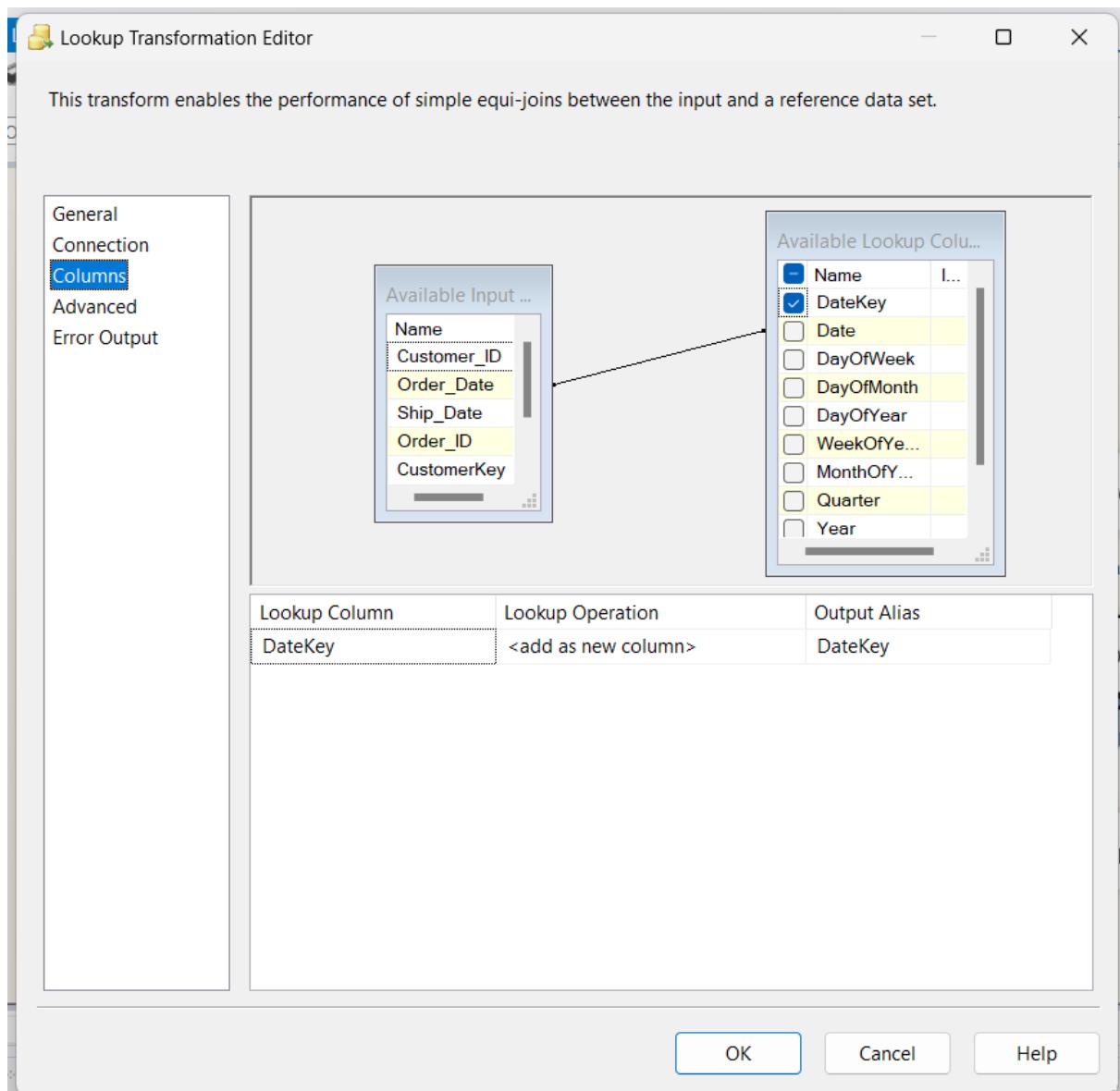
Lookup **DimOrder** ta sẽ map cột order\_id trong bảng stgOrder với cột OrderID trong bảng DimOrder để có thể lấy ra cột OrderKey



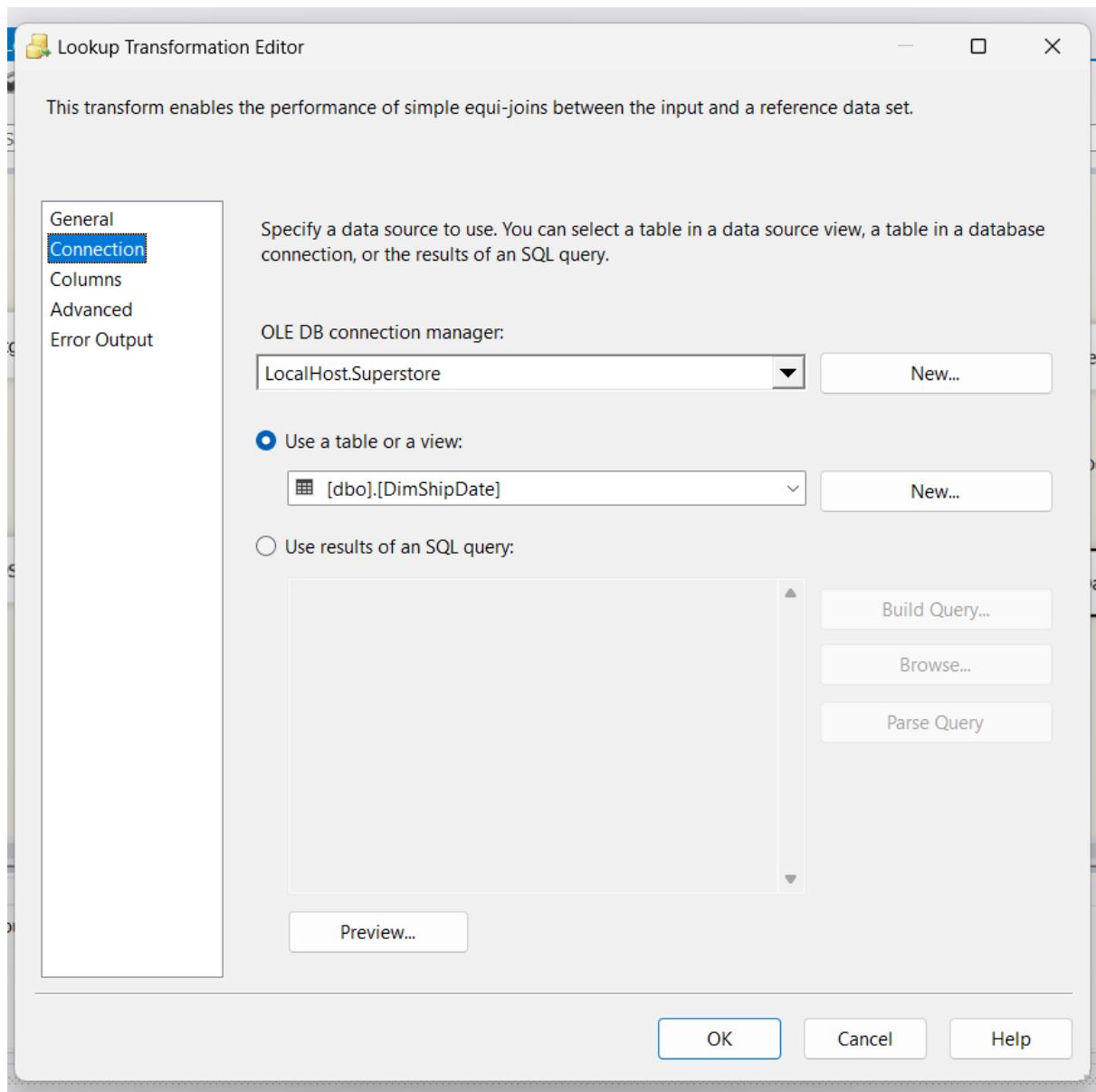


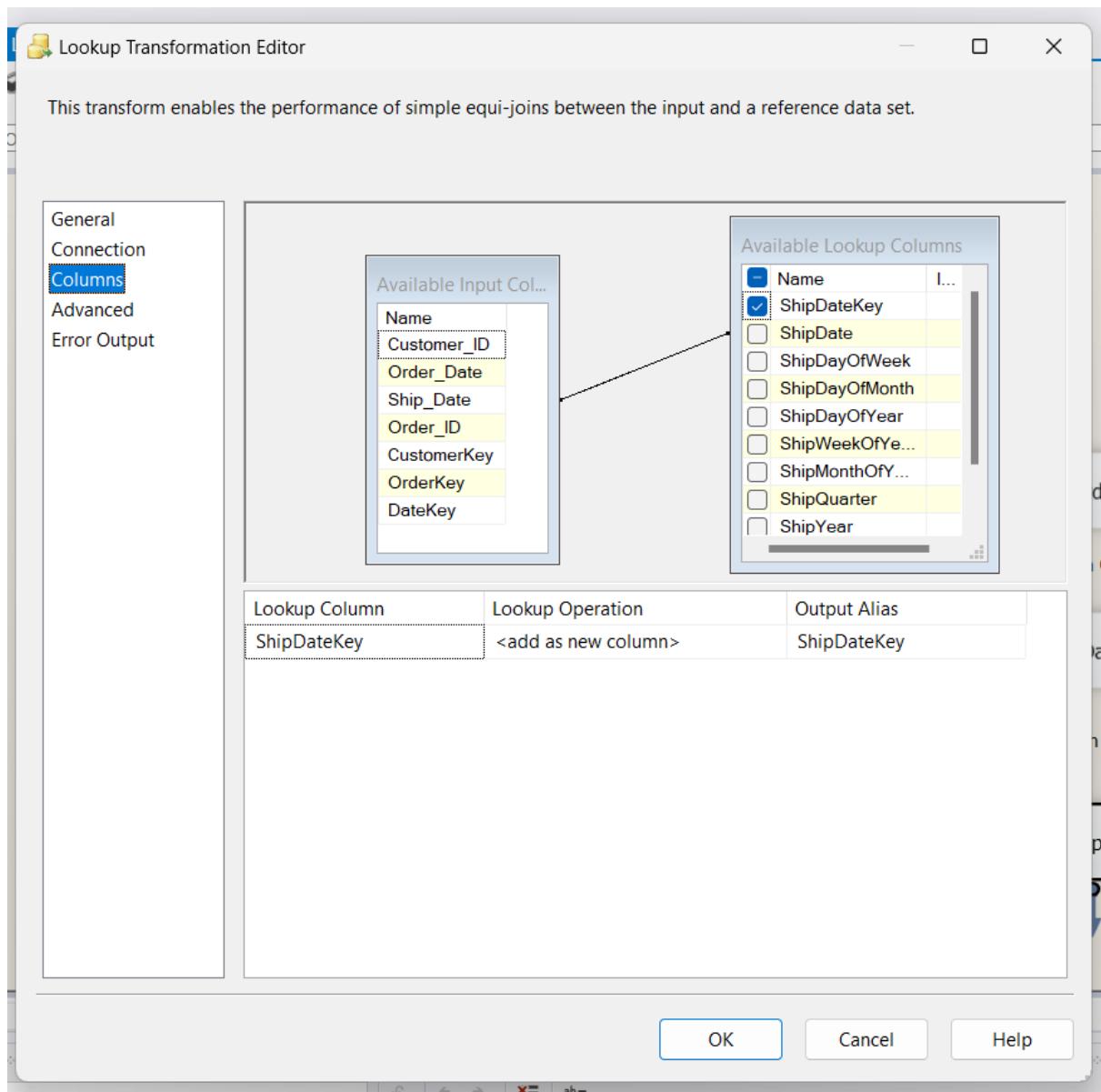
Lookup **DimDate** ta sẽ map cột order\_date ở bảng stgSales với cột Date trong bảng DimDate để lấy ra DateKey



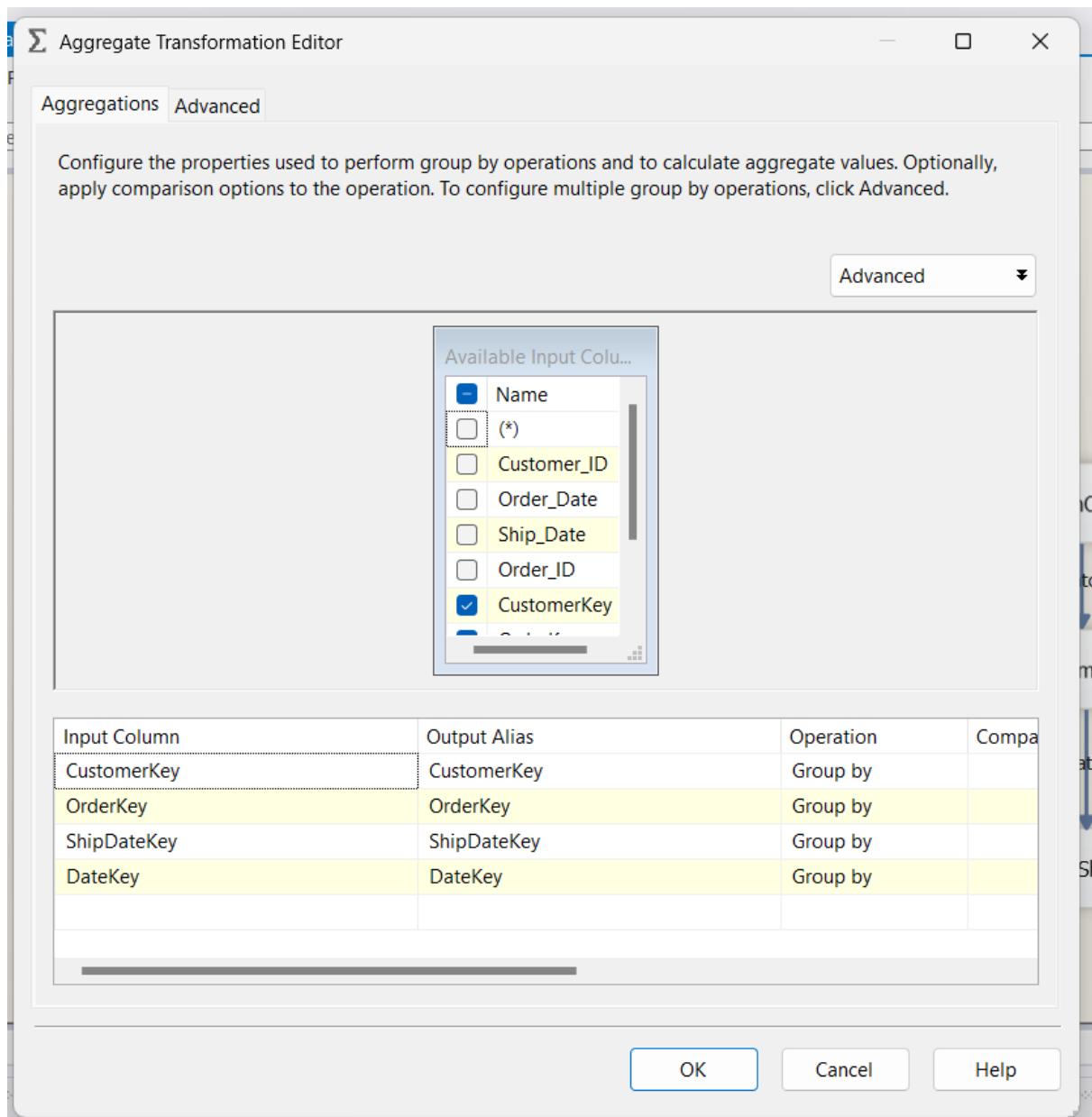


Lookup **DimShipDate** ta sẽ map cột ship\_date ở bảng stgSales với cột ShipDate trong bảng DimShipDate để lấy ra ShipDateKey

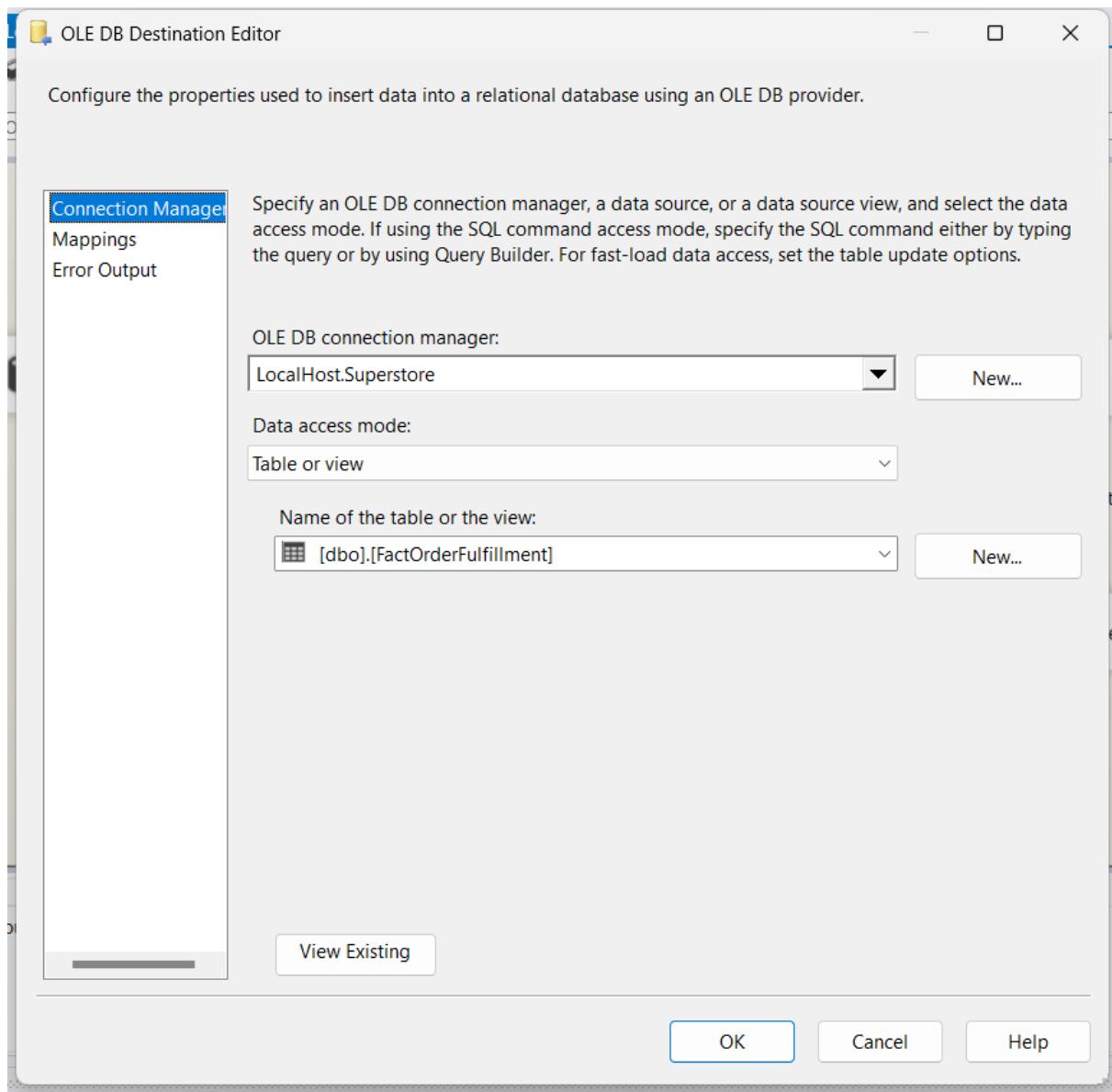


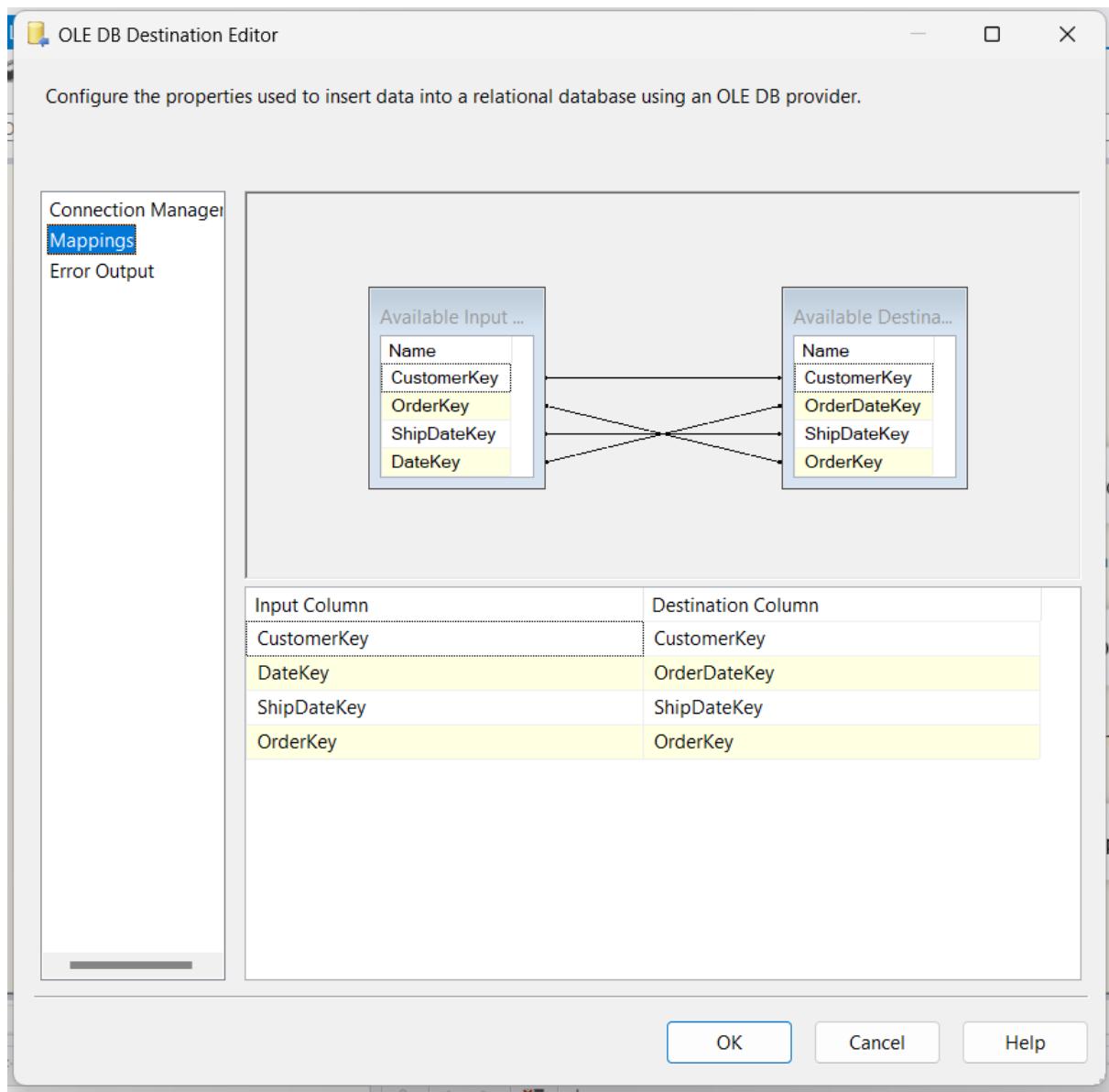


**Aggregate** ta sẽ chỉ lấy những cột cần thiết để đưa vào bảng Fact OrderFulfillment

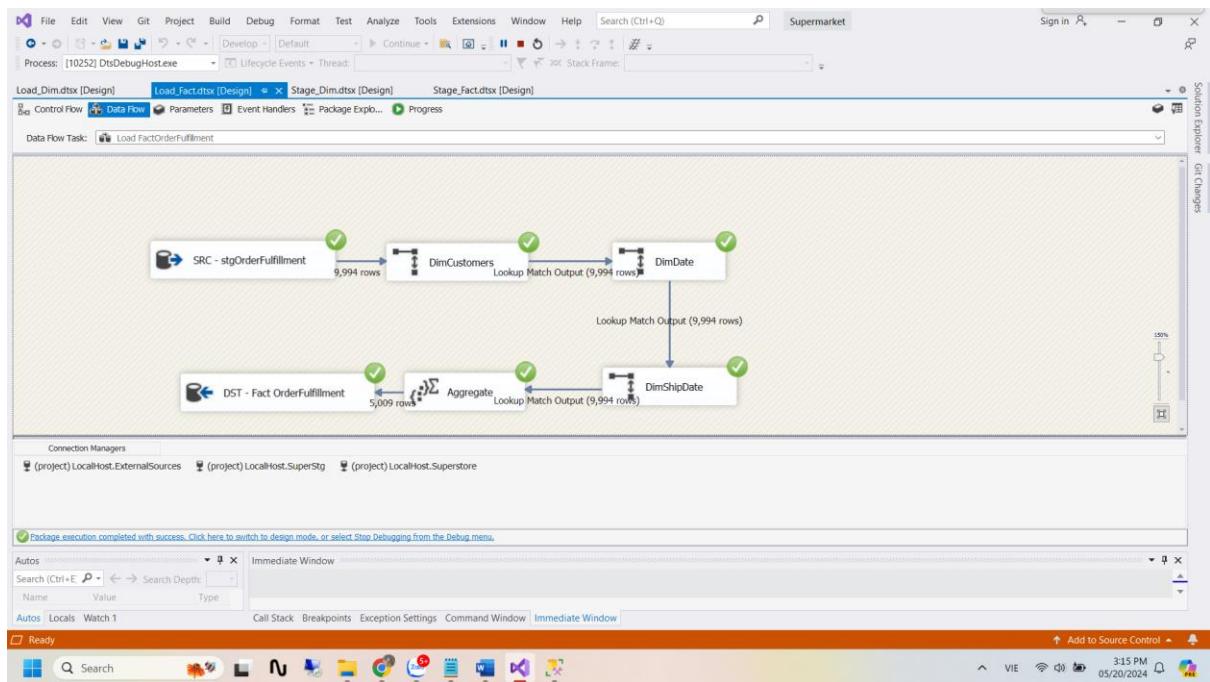


## DST - Fact OrderFulfillment

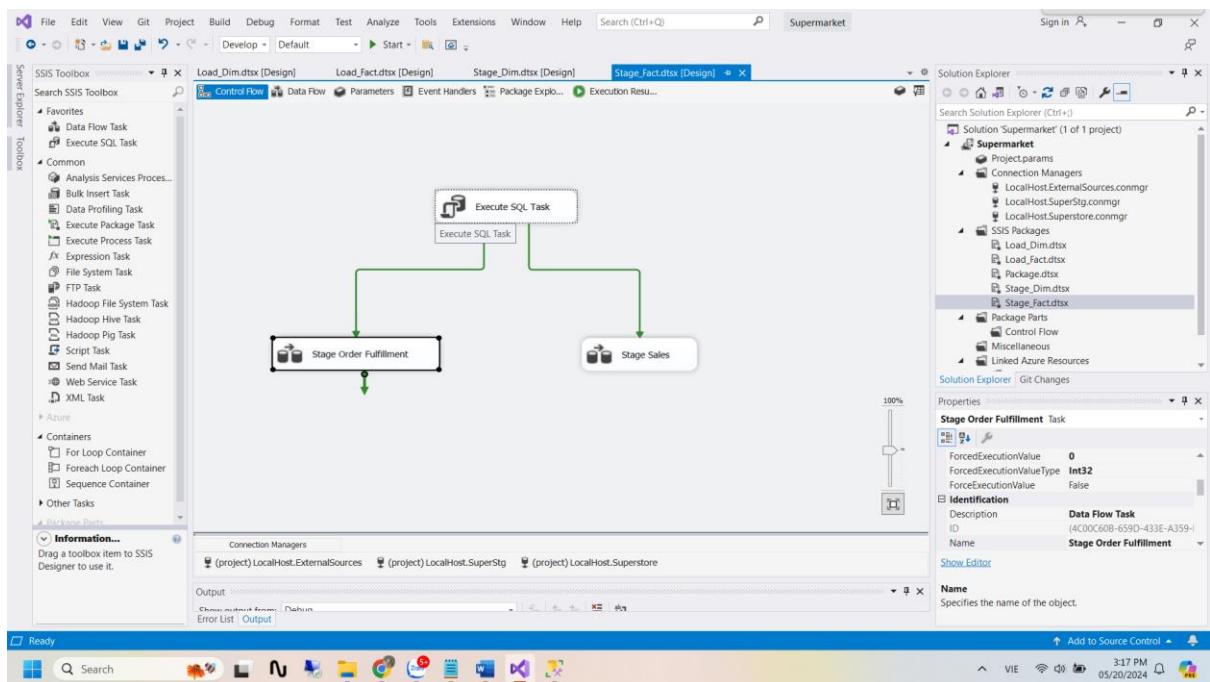




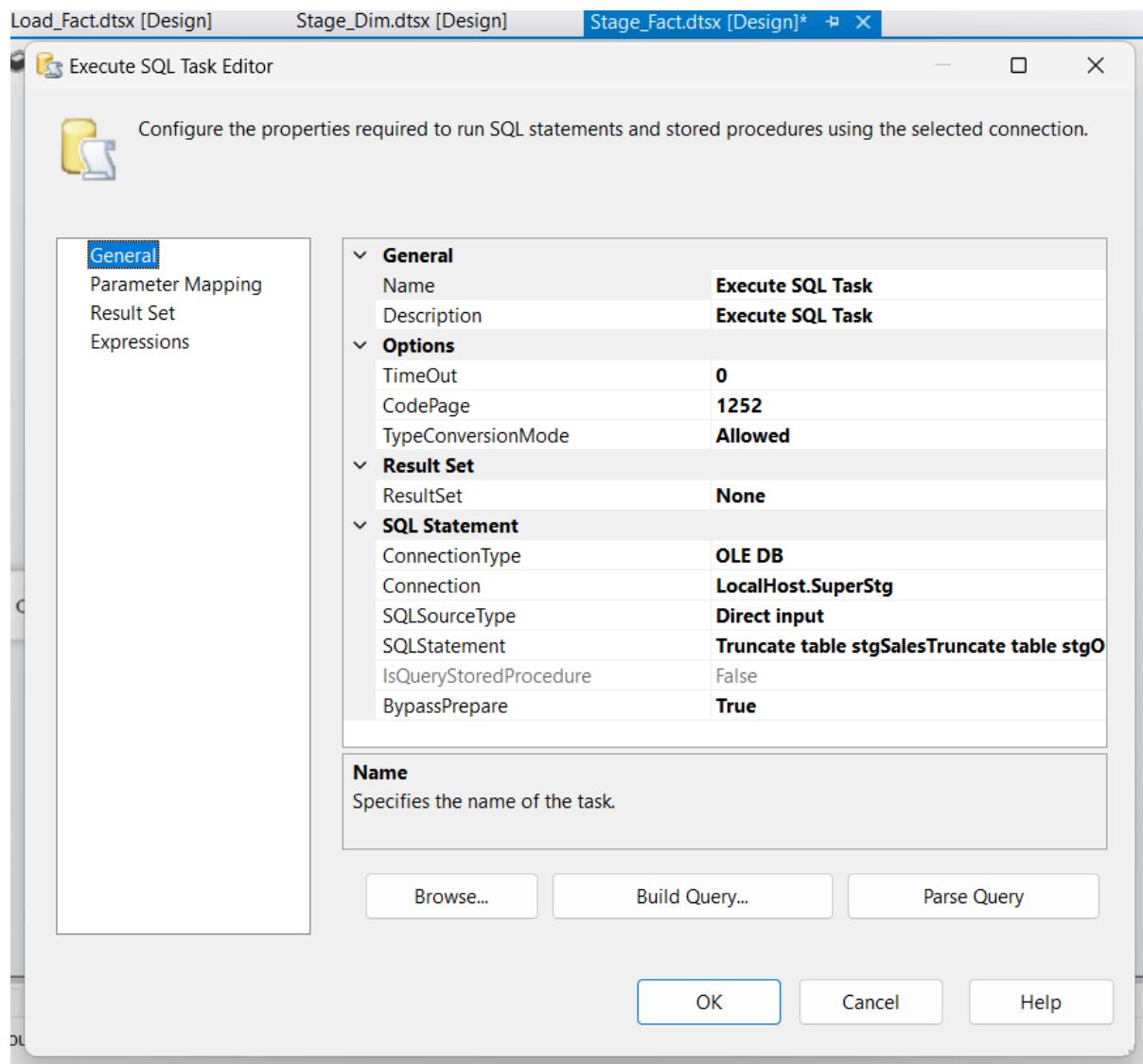
**Khởi chạy**



### 3.3.3. Package Stage\_Fact



## a) Setup Execute SQL Task



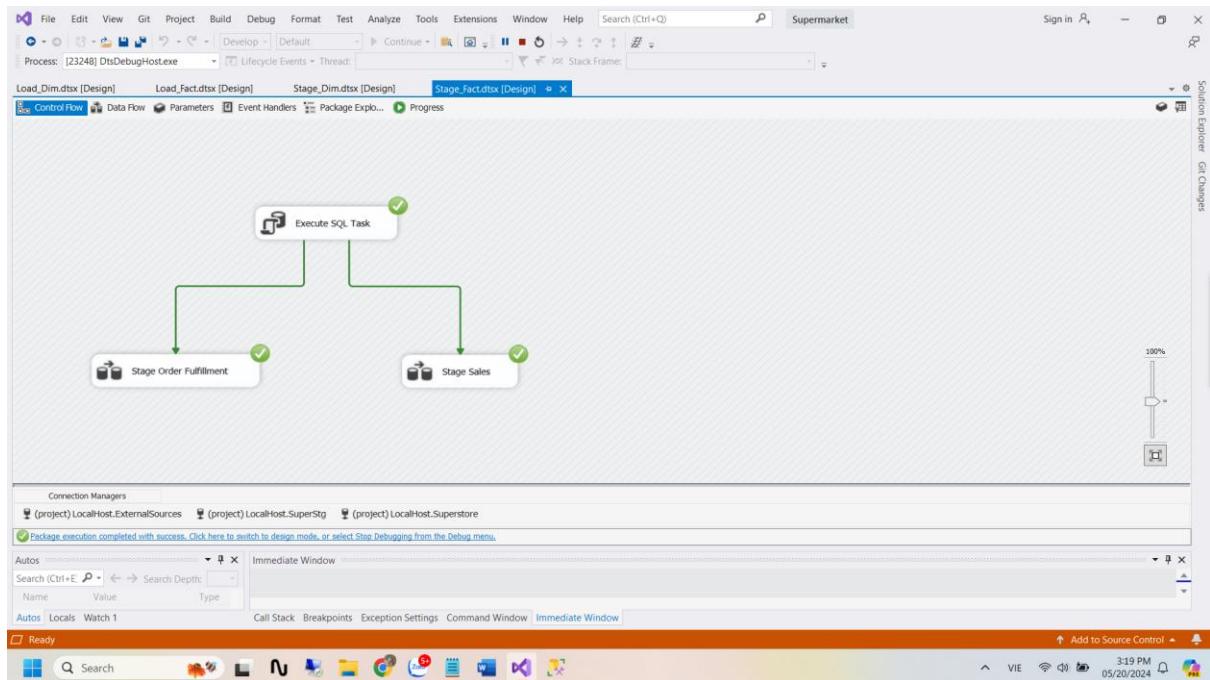
Connection : kết nối đến database SuperStg

SQLStatement :

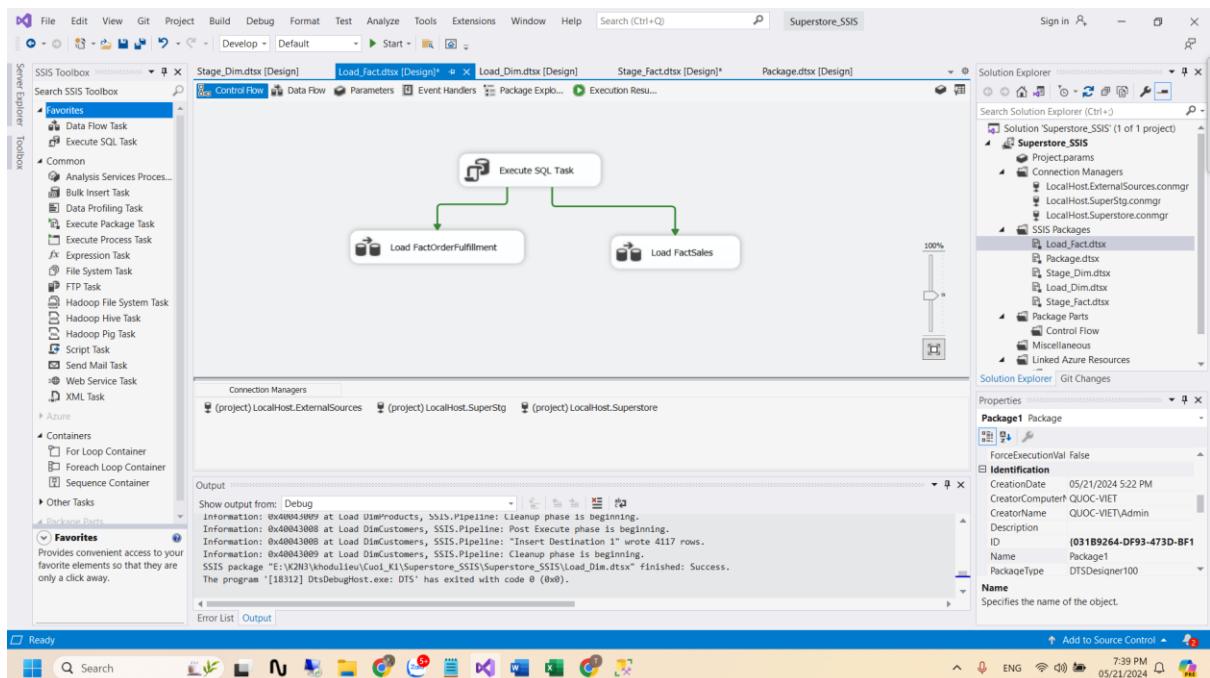
Truncate table stgSales

Truncate table stgOrderFulfillment

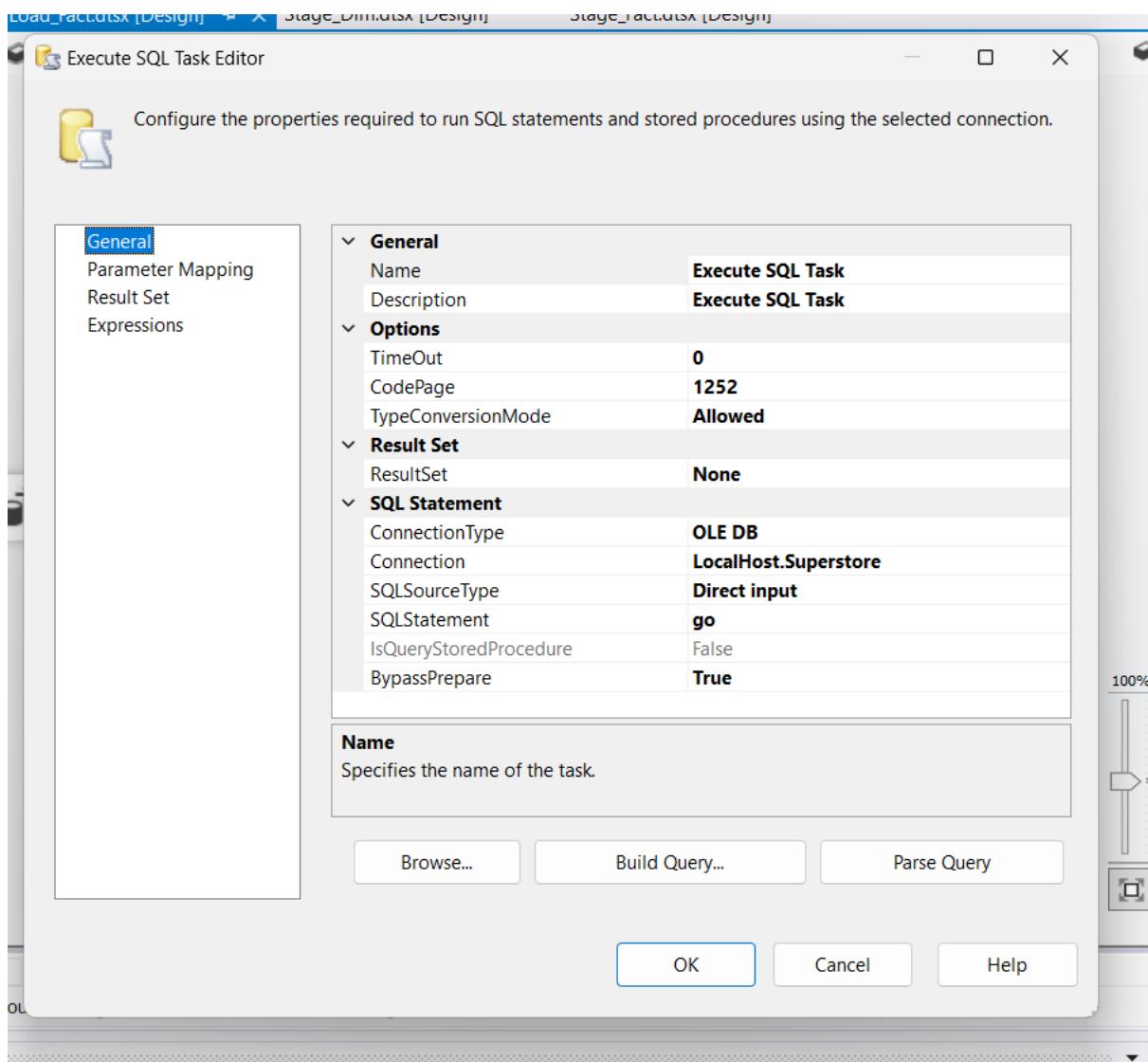
**Khởi chạy**



### 3.3.4. Package Load\_Fact



### a) Setup Execute SQL Task

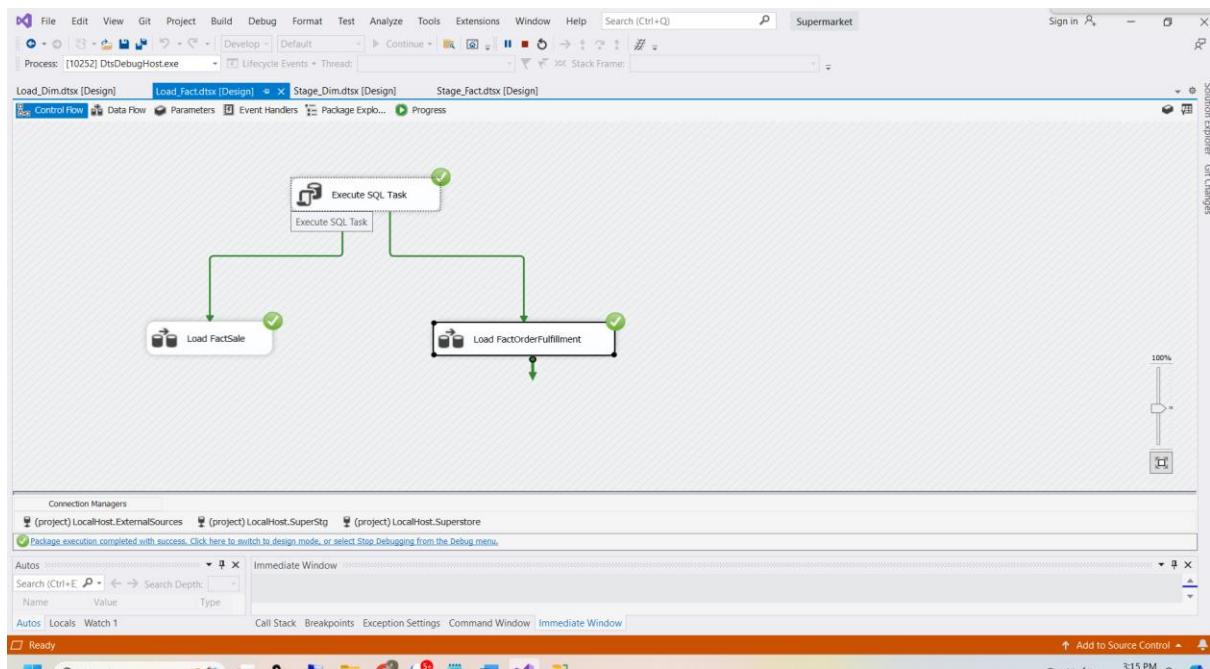


Connection : kết nối đến database Superstore

SQLStatement :

go

**Khởi chạy**



## SSIS hoàn tất

Kiểm tra dữ liệu data base Superstore ở sql :

### DimDate

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. The title bar says 'QUOC-VIET.Superstore - dbo.DimDate - Microsoft SQL Server Management Studio'. The left side has an 'Object Explorer' pane showing the database structure, including 'Databases', 'Tables', and 'System Tables'. The main area is a results grid for a query named '(Qty) Query1.dtrq'. The grid shows data for the 'DimDate' table, with columns: DateKey, Date, DayOfWeek, DayOfMonth, DayOfYear, WeekOfMonth, MonthOfYear, Quarter, Year, IsWeekDay. The data starts from 20140103 and continues through 20140208. The right side of the screen has a 'Properties' window for the query, showing details like 'Name: Query1.dtrq', 'Database Name: Superstore', 'Server Name: quoc-viet', and 'Top Specification: Yes'.

DateKey	Date	DayOfWeek	DayOfMonth	DayOfYear	WeekOfMonth	MonthOfYear	Quarter	Year	IsWeekDay
20140103	2014-01-03 .. 6	3	3	1	1	1	1	2014	Y
20140104	2014-01-04 .. 7	4	4	1	1	1	1	2014	N
20140105	2014-01-05 .. 1	5	5	2	1	1	1	2014	N
20140106	2014-01-06 .. 2	6	6	2	1	1	1	2014	Y
20140107	2014-01-07 .. 3	7	7	2	1	1	1	2014	Y
20140109	2014-01-09 .. 5	9	9	2	1	1	1	2014	Y
20140110	2014-01-10 .. 6	10	10	2	1	1	1	2014	Y
20140111	2014-01-11 .. 7	11	11	2	1	1	1	2014	N
20140113	2014-01-13 .. 2	13	13	3	1	1	1	2014	Y
20140114	2014-01-14 .. 3	14	14	3	1	1	1	2014	Y
20140115	2014-01-15 .. 4	15	15	3	1	1	1	2014	Y
20140116	2014-01-16 .. 5	16	16	3	1	1	1	2014	Y
20140118	2014-01-18 .. 7	18	18	3	1	1	1	2014	N
20140119	2014-01-19 .. 1	19	19	4	1	1	1	2014	N
20140120	2014-01-20 .. 2	20	20	4	1	1	1	2014	Y
20140121	2014-01-21 .. 3	21	21	4	1	1	1	2014	Y
20140123	2014-01-23 .. 5	23	23	4	1	1	1	2014	Y
20140126	2014-01-26 .. 1	26	26	5	1	1	1	2014	N
20140127	2014-01-27 .. 2	27	27	5	1	1	1	2014	Y
20140128	2014-01-28 .. 3	28	28	5	1	1	1	2014	Y
20140130	2014-01-30 .. 5	30	30	5	1	1	1	2014	Y
20140131	2014-01-31 .. 6	31	31	5	1	1	1	2014	Y
20140201	2014-02-01 .. 7	1	32	5	2	1	1	2014	N
20140202	2014-02-02 .. 1	2	33	6	2	1	1	2014	N
20140203	2014-02-03 .. 2	3	34	6	2	1	1	2014	Y
20140204	2014-02-04 .. 3	4	35	6	2	1	1	2014	Y
20140206	2014-02-06 .. 5	6	37	6	2	1	1	2014	Y
20140207	2014-02-07 .. 6	7	38	6	2	1	1	2014	Y
20140208	2014-02-08 .. 7	8	39	6	2	1	1	2014	N

## DimProduct

The screenshot shows the Microsoft SQL Server Management Studio interface. The title bar indicates the connection is to 'QUOC-VIET.Superstore - dbo.DimProduct'. The main window displays the 'DimProduct' table with columns: ProductKey, ProductID, ProductName, SubCategory, CategoryName, RowStartD..., RowEndD..., and RowChang... . The table contains 200 rows of product information. The 'Properties' pane on the right shows a query named 'Query1.dts' with various settings like 'Identity' and 'Query Designer'.

## Dim Order

The screenshot shows the Microsoft SQL Server Management Studio interface. The title bar indicates the connection is to 'QUOC-VIET.SuperStg - dbo.stgOrder'. The main window displays the 'stgOrder' table with columns: Order\_ID, Ship\_Mode, and a large text column 'order'. The table contains 200 rows of order data. The 'Properties' pane on the right shows a query named 'Query1.dts' with various settings like 'Identity' and 'Query Designer'.

## DimCustomer

QUOC-VIET.Superstore - dbo.DimCustomers - Microsoft SQL Server Management Studio

File Edit View Project Query Designer Tools Window Help

New Query Execute Change Type SQL Properties

OrderDate

Object Explorer

Connect Databases System Databases BankLoanDB qual-hang-QA ExternalSources QLNH QLVATTU QLVLT StoreSalesDW Super SupermarketDW SuperStg Superstore Database Diagrams Tables System Tables FileTables External Tables Graph Tables DimAudit DimCustomers DimDate DimProduct DimShipDate FactOrderFulfillment FactSale Dropped Ledger Tables Views External Resources Synonyms Programmability Oracle Store

QUOC-VIET.Superst...dbo.DimCustomers QUOC-VIET.Superst...dbo.DimProduct QUOC-VIET Superstore - dbo.DimDate

CustomerK...	CustomerID...	CustomerN...	CustomerS...	ContactS...	CustomerR...	CustomerC...	CustomerC...	CustomerP...	RowS...	RowStartD...
9995	HE-14800	Harold Engle	Delaware	Corporate	East	United States	Newark	19711	False	1899-12-31 ...
9996	RB-19435	Richard Bier...	New Jersey	Consumer	East	United States	Plainfield	7060	False	1899-12-31 ...
9997	RB-19435	Richard Bier...	New Jersey	Consumer	East	United States	Plainfield	7060	False	1899-12-31 ...
9998	RB-19435	Richard Bier...	New Jersey	Consumer	East	United States	Plainfield	7060	False	1899-12-31 ...
9999	MP-17470	Mark Packer	Georgia	Home Office	South	United States	Smyrna	30080	False	1899-12-31 ...
10000	MP-17470	Mark Packer	Georgia	Home Office	South	United States	Smyrna	30080	False	1899-12-31 ...
10001	RC-19960	Ryan Crowe	Texas	Consumer	Central	United States	Houston	77041	False	1899-12-31 ...
10002	AP-10720	Anne Pryor	California	Home Office	West	United States	Los Angeles	90032	False	1899-12-31 ...
10003	AP-10720	Anne Pryor	California	Home Office	West	United States	Los Angeles	90032	False	1899-12-31 ...
10004	AP-10720	Anne Pryor	California	Home Office	West	United States	Los Angeles	90032	False	1899-12-31 ...
10005	AP-10720	Anne Pryor	California	Home Office	West	United States	Los Angeles	90032	False	1899-12-31 ...
10006	AP-10720	Anne Pryor	California	Home Office	West	United States	Los Angeles	90032	False	1899-12-31 ...
10007	AP-10720	Anne Pryor	California	Home Office	West	United States	Los Angeles	90032	False	1899-12-31 ...
10008	AP-10720	Anne Pryor	California	Home Office	West	United States	Los Angeles	90032	False	1899-12-31 ...
10009	SW-20455	Shaun Weis	Louisiana	Consumer	South	United States	Lafayette	70506	False	1899-12-31 ...
10010	TB-21055	Ted Butterf...	Ohio	Consumer	East	United States	Fairfield	45014	False	1899-12-31 ...
10011	RC-19960	Ryan Crowe	Michigan	Consumer	Central	United States	Grand Rapids	49505	False	1899-12-31 ...
10012	RC-19960	Ryan Crowe	Michigan	Consumer	Central	United States	Grand Rapids	49505	False	1899-12-31 ...
10013	DV-13465	Dianna Vitt...	New York	Consumer	East	United States	Long Beach	11561	False	1899-12-31 ...
10014	DV-13465	Dianna Vitt...	New York	Consumer	East	United States	Long Beach	11561	False	1899-12-31 ...
10015	ML-17410	Maris LaWare	California	Consumer	West	United States	Los Angeles	90008	False	1899-12-31 ...
10016	RA-19885	Ruben Aus...	Georgia	Corporate	South	United States	Athens	30605	False	1899-12-31 ...
10017	RA-19885	Ruben Aus...	Georgia	Corporate	South	United States	Fairfield	30605	False	1899-12-31 ...
10018	TB-21400	Tom Beecke...	Florida	Consumer	South	United States	Miami	33180	False	1899-12-31 ...
10019	DB-13060	Dave Brooks	California	Consumer	West	United States	Costa Mesa	92627	False	1899-12-31 ...
10020	DB-13060	Dave Brooks	California	Consumer	West	United States	Costa Mesa	92627	False	1899-12-31 ...
10021	DB-13060	Dave Brooks	California	Consumer	West	United States	Costa Mesa	92627	False	1899-12-31 ...
10022	CC-12220	Chris Cortes	California	Consumer	West	United States	Westminster	92683	False	1899-12-31 ...
10023	CC-12520	Claire Gutter...	Kentucky	Consumer	South	United States	Henderson	42429	False	1899-12-31 ...

Properties [Qry] Query1.dtq (Identity) (Name) Query1.dtq Database Name Superstore Server Name quo-viet (Query Designer) Destination Tabl Distinct Values No GROUP BY Exten <None> Output All Colun No Query Parameter No parameters have SQL Comment \*\*\*\* Script for Select Top Specification Yes (Identity)

Ready

File Edit View Project Query Designer Tools Window Help

New Query Execute Change Type SQL Properties

OrderDate

Object Explorer

Connect Databases System Databases BankLoanDB qual-hang-QA ExternalSources QLNH QLVATTU QLVLT StoreSalesDW Super SupermarketDW SuperStg Superstore Database Diagrams Tables System Tables FileTables External Tables Graph Tables DimAudit DimCustomers DimDate DimProduct DimShipDate FactOrderFulfillment FactSale Dropped Ledger Tables Views External Resources Synonyms Programmability Oracle Store

QUOC-VIET.Superst...dbo.DimShipDate QUOC-VIET.Superst...dbo.DimCustomers QUOC-VIET Superstore - dbo.DimProduct

ShipDateKey...	ShipDate...	ShipDayOf...	ShipDayOf...	ShipWeekOf...	ShipMonth...	ShipQuarter...	ShipYear...	ShipS...	
20140107	2014-01-07 ...	3	7	2	1	1	2014	Y	
20140108	2014-01-08 ...	4	8	2	1	1	2014	Y	
20140110	2014-01-10 ...	6	10	2	1	1	2014	Y	
20140112	2014-01-12 ...	1	12	3	1	1	2014	N	
20140113	2014-01-13 ...	2	13	3	1	1	2014	Y	
20140114	2014-01-14 ...	3	14	3	1	1	2014	Y	
20140115	2014-01-15 ...	4	15	3	1	1	2014	Y	
20140116	2014-01-16 ...	5	16	3	1	1	2014	Y	
20140117	2014-01-17 ...	6	17	3	1	1	2014	Y	
20140118	2014-01-18 ...	7	18	3	1	1	2014	N	
20140120	2014-01-20 ...	2	20	4	1	1	2014	Y	
20140121	2014-01-21 ...	3	21	4	1	1	2014	Y	
20140123	2014-01-23 ...	5	23	4	1	1	2014	Y	
20140125	2014-01-25 ...	7	25	4	1	1	2014	N	
20140126	2014-01-26 ...	1	26	5	1	1	2014	N	
20140127	2014-01-27 ...	2	27	5	1	1	2014	Y	
20140128	2014-01-28 ...	3	28	5	1	1	2014	Y	
20140129	2014-01-29 ...	4	29	5	1	1	2014	Y	
20140131	2014-01-31 ...	6	31	5	1	1	2014	Y	
20140202	2014-02-02 ...	1	2	33	6	2	1	2014	N
20140203	2014-02-03 ...	2	3	34	6	2	1	2014	Y
20140204	2014-02-04 ...	3	4	35	6	2	1	2014	Y
20140206	2014-02-06 ...	5	6	37	6	2	1	2014	Y
20140208	2014-02-08 ...	7	8	39	6	2	1	2014	N
20140209	2014-02-09 ...	1	9	40	7	2	1	2014	N
20140210	2014-02-10 ...	2	10	41	7	2	1	2014	Y
20140212	2014-02-12 ...	4	12	43	7	2	1	2014	Y
20140215	2014-02-15 ...	7	15	46	7	2	1	2014	N
20140218	2014-02-18 ...	3	18	49	8	2	1	2014	Y

Properties [Qry] Query1.dtq (Identity) (Name) Query1.dtq Database Name Superstore Server Name quo-viet (Query Designer) Destination Tabl Distinct Values No GROUP BY Exten <None> Output All Colun No Query Parameter No parameters have SQL Comment \*\*\*\* Script for Select Top Specification Yes (Identity)

## FactSales

QUOC-VIET.Superstore - dbo.FactSale - Microsoft SQL Server Management Studio

File Edit View Project Query Designer Tools Window Help

New Query Execute

OrderID OrderDate ProductKey ShipDateKey CustomerK\_ Quantity Profit DiscountA.. Sales

OrderID	OrderDate	ProductKey	ShipDateKey	CustomerK_	Quantity	Profit	DiscountA..	Sales
2017-14032...	20170904	11356	20170909	10823	3	5.55000019...	0.20000000...	17.7600002...
CA-2014-11...	20140311	10844	20140313	11750	2	14.1728000...	0	32.9599990...
CA-2016-11...	20160610	10551	20160615	10579	2	9.9359998...	0	20.7000007...
CA-2017-11...	20171208	10402	20171212	10428	5	13.3035001...	0	27.1499996...
CA-2017-16...	20170514	16404	20170515	10576	3	16.1837997...	0.20000000...	47.9519996...
CA-2016-12...	20160409	10873	20160413	11582	9	12.8150997...	0.20000000...	35.3520011...
US-2017-10...	20170717	10024	20170723	10532	4	14.4596004...	0	39.0800018...
CA-2015-10...	20150901	12267	20150908	12838	2	27.3591995...	0	124.360000...
CA-2017-15...	20171116	12578	20171120	10156	3	4.47119998...	0	12.420000...
CA-2016-10...	20160422	10460	20160429	10485	2	-34.7519994...	0.20000000...	185.37000...
CA-2014-10...	20141124	10375	20141201	10786	3	22.3733997...	0	45.6599998...
CA-2017-13...	20170904	11901	20170906	10133	3	4.13280010...	0.20000000...	10.8709996...
CA-2016-10...	20161003	11906	20161007	12535	2	45.9753990...	0	97.8199996...
CA-2017-15...	20170518	10017	20170519	11310	2	4.48239994...	0	17.2399997...
US-2017-13...	20170514	10207	20170517	10888	5	37	0	74
CA-2014-13...	20140908	10681	20140909	16645	8	27.4855995...	0	58.4799995...
CA-2015-13...	20150404	11032	20150407	13168	3	-13.7340002...	0.80000001...	9.15600013...
CA-2017-16...	20170915	10100	20170917	15860	9	106.957801...	0.20000000...	295.055999...
CA-2015-12...	20150618	11179	20150623	11093	7	-3.32640004...	0.69999998...	4.53599977...
CA-2016-11...	20160925	11330	20160929	13342	3	-82.8840026...	0.20000000...	331.536010...
CA-2017-12...	20171222	10120	20171225	10621	2	70.0059967...	0	411.799987...
CA-2017-14...	20170603	11818	20170610	14744	3	5.4432001...	0.20000000...	15.532000...
CA-2014-15...	20141223	11491	20141227	10281	5	-1.04900002...	0.20000000...	83.9199981...
CA-2017-12...	20171214	13234	20171220	10111	3	19.7549991...	0.20000000...	52.6800003...
CA-2014-13...	20141220	12819	20141222	12057	2	2.64000010...	0.20000000...	8.44799995...
CA-2014-13...	20141101	10589	20141105	11073	2	22.0748004...	0	44.7200007...
CA-2015-10...	20151219	10262	20151221	11284	4	0.20160000...	0	5.03999996...
CA-2017-11...	20170917	12648	20170921	10327	3	80.7911987...	0	164.880004...
CA-2014-14...	20140529	13263	20140529	13290	1	12.5860004...	0	44.9500007...

Properties [Query1.dtq] (Identity) (Name) Query1.dtq Database Name Superstore Server Name quoc-viet

Query Designer Destination Table Distinct Values No GROUP BY Ext -None- Output All Column No Query Parameter No parameters have been specified SQL Comment \*\*\*\* Script for Select Top Specification Yes

Identity

## Fact OrderFulfillment

QUOC-VIET.SuperStg - dbo.stgOrderFulfillment - Microsoft SQL Server Management Studio

File Edit View Project Query Designer Tools Window Help

New Query Execute

Customer\_ID Order\_Date Ship\_Date Order\_ID

Customer_ID	Order_Date	Ship_Date	Order_ID
Cg-12526...	2016-11-08 00:00:00.000	2016-11-11 00:00:00.000	CA-2016-15...
CG-12520...	2016-11-08 00:00:00.000	2016-11-11 00:00:00.000	CA-2016-15...
DV-13045...	2016-06-12 00:00:00.000	2016-06-16 00:00:00.000	CA-2016-13...
SO-20335...	2015-10-11 00:00:00.000	2015-10-18 00:00:00.000	US-2015-10...
SO-20335...	2015-10-11 00:00:00.000	2015-10-18 00:00:00.000	US-2015-10...
BH-11710...	2014-06-09 00:00:00.000	2014-06-14 00:00:00.000	CA-2014-11...
BH-11710...	2014-06-09 00:00:00.000	2014-06-14 00:00:00.000	CA-2014-11...
BH-11710...	2014-06-09 00:00:00.000	2014-06-14 00:00:00.000	CA-2014-11...
BH-11710...	2014-06-09 00:00:00.000	2014-06-14 00:00:00.000	CA-2014-11...
BH-11710...	2014-06-09 00:00:00.000	2014-06-14 00:00:00.000	CA-2014-11...
AA-10480...	2017-04-15 00:00:00.000	2017-04-20 00:00:00.000	CA-2017-11...
IM-15070...	2016-12-05 00:00:00.000	2016-12-10 00:00:00.000	CA-2016-16...
HP-14815...	2015-11-22 00:00:00.000	2015-11-26 00:00:00.000	US-2015-11...
HP-14815...	2015-11-22 00:00:00.000	2015-11-26 00:00:00.000	US-2015-11...
PK-19075...	2014-11-11 00:00:00.000	2014-11-18 00:00:00.000	CA-2014-10...
AG-10270...	2014-05-13 00:00:00.000	2014-05-15 00:00:00.000	CA-2014-16...
ZD-21925...	2014-08-27 00:00:00.000	2014-09-01 00:00:00.000	CA-2014-14...
ZD-21925...	2014-08-27 00:00:00.000	2014-09-01 00:00:00.000	CA-2014-14...
ZD-21925...	2014-08-27 00:00:00.000	2014-09-01 00:00:00.000	CA-2014-14...
EH-13945...	2016-01-16 00:00:00.000	2016-01-20 00:00:00.000	CA-2016-12...
EH-13945...	2016-01-16 00:00:00.000	2016-01-20 00:00:00.000	CA-2016-12...
TB-21520...	2015-09-17 00:00:00.000	2015-09-21 00:00:00.000	US-2015-15...
TB-21520...	2015-09-17 00:00:00.000	2015-09-21 00:00:00.000	US-2015-15...

Properties [Query1.dtq] (Identity) (Name) Query1.dtq Database Name SuperStg Server Name quoc-viet

Query Designer Destination Table Distinct Value: No GROUP BY Ext -None- Output All Column No Query Parameter No parameters have been specified SQL Comment \*\*\*\* Script for SelectTop Specification Yes

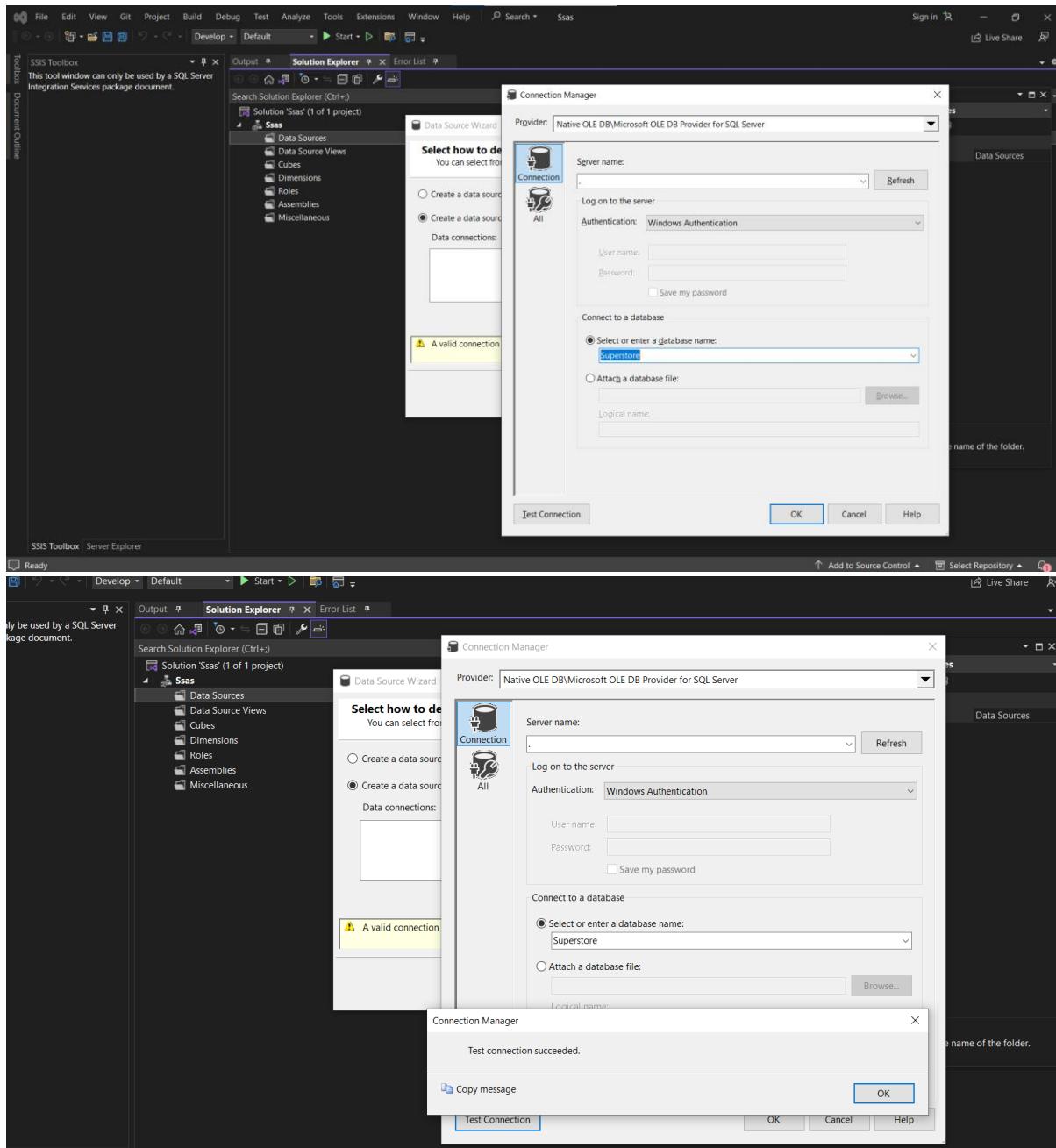
Identity

# CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS)

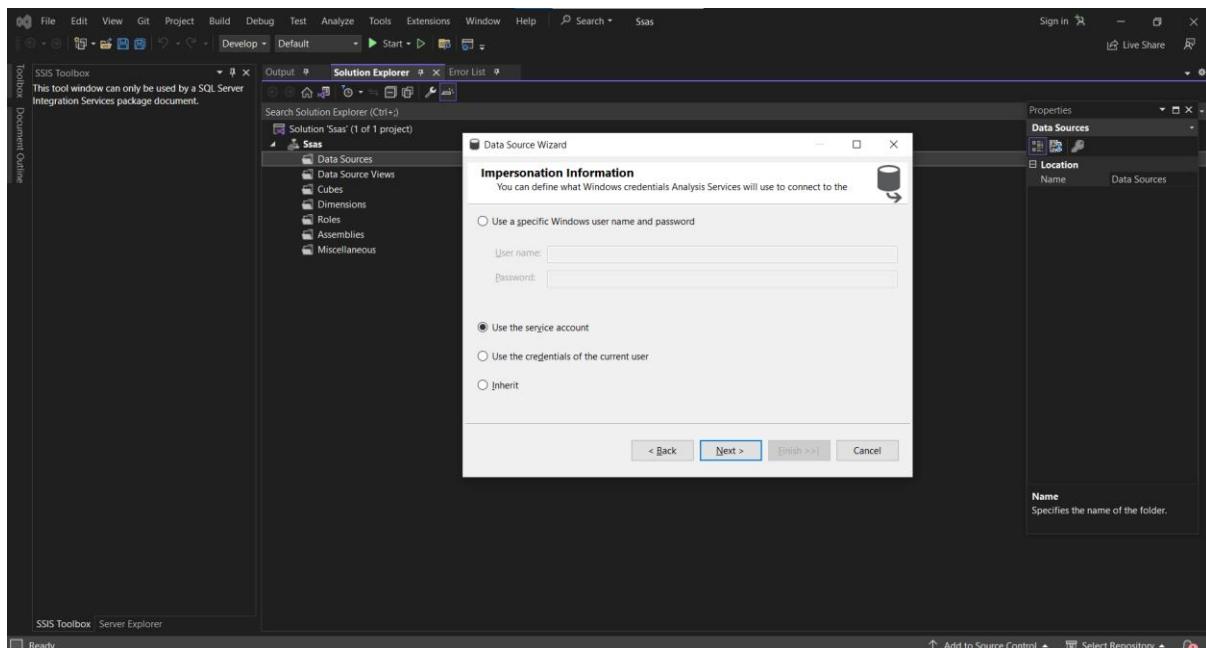
## 4.1. Quá trình xây dựng mô hình

### 4.1.1. Tạo Data Source

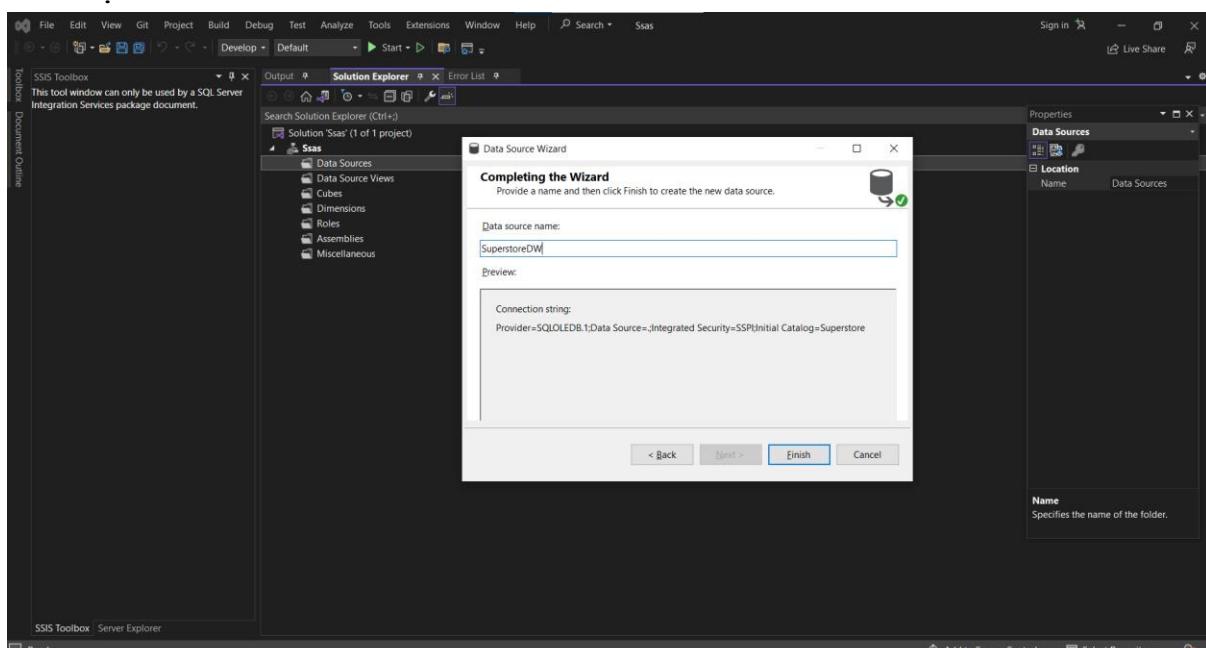
- Right-click vào Data Sources ở khung Solution Explorer. Chon New Data Source để tạo Data Source mới. Kết nối đến Database Superstore



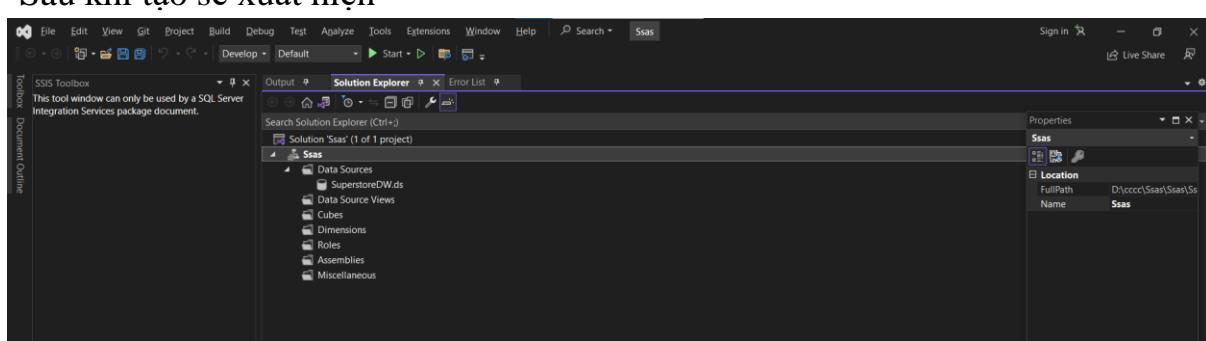
- Chọn Use the service account



- Chọn tên source và bấm Finish

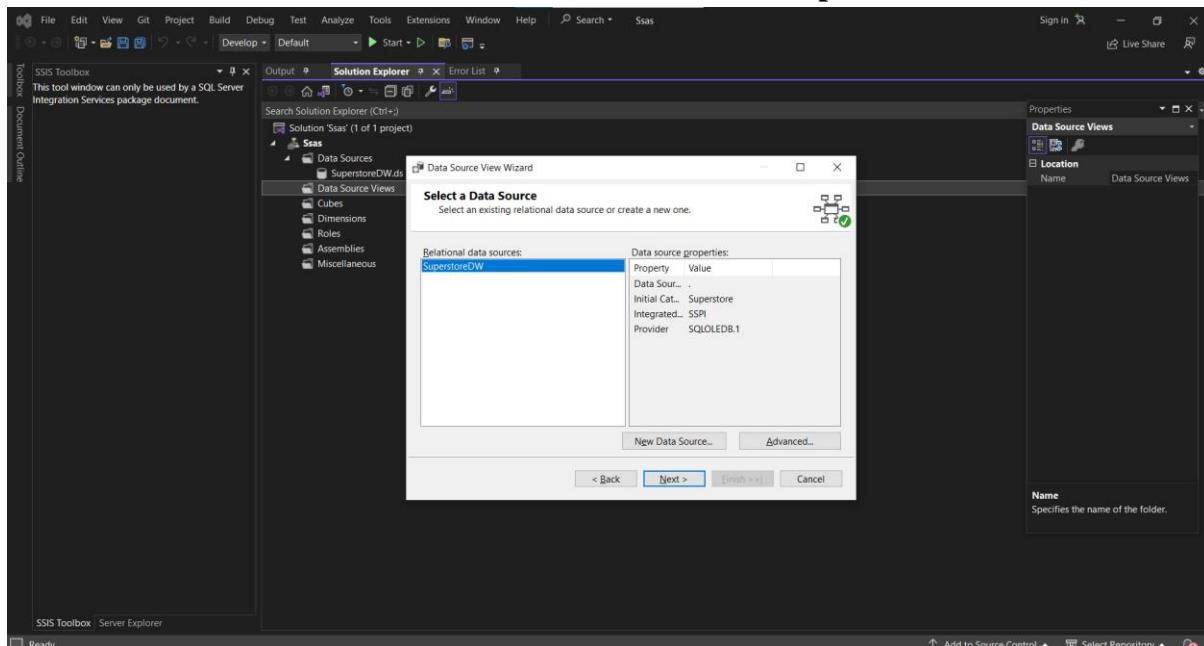


- Sau khi tạo sẽ xuất hiện

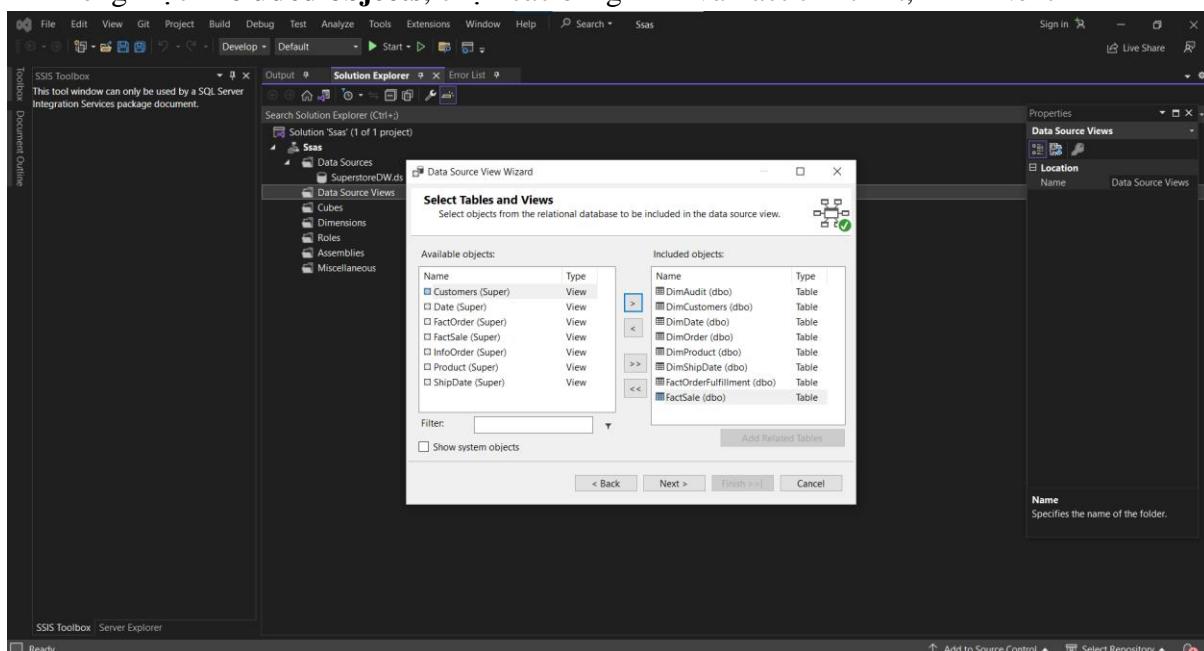


#### 4.1.2. Tạo Data Source View

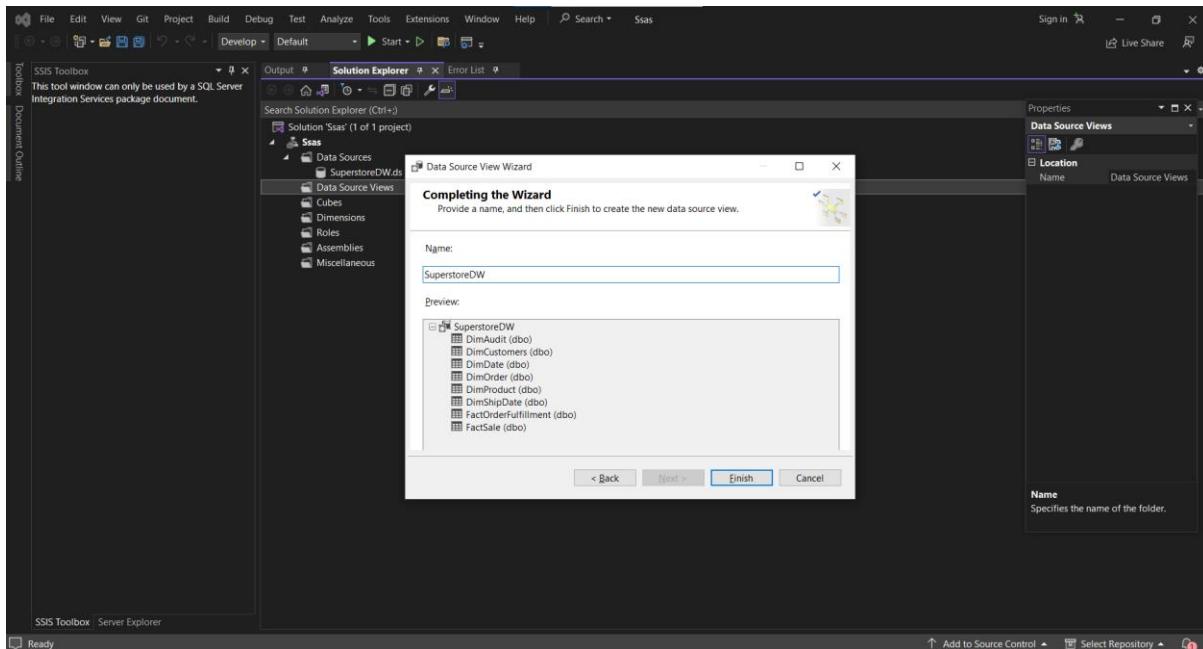
- Right-click vào Data Source Views ở khung Solution Explorer. Chon New Data Source View để tạo Data Source View mới. Chọn Data Soure là **SuperstoreDW** và chọn **Next**.



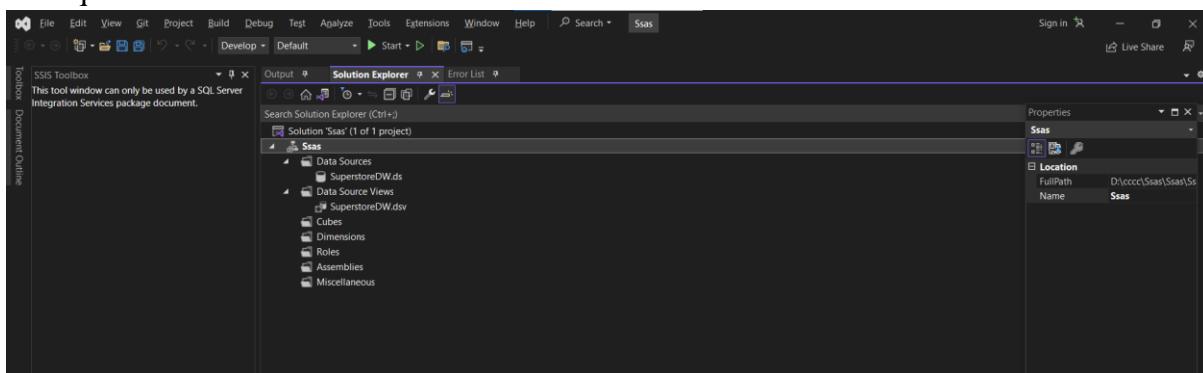
- Trong mục **Included objects**, chọn các bảng Dim và Fact cần thiết, nhấn Next



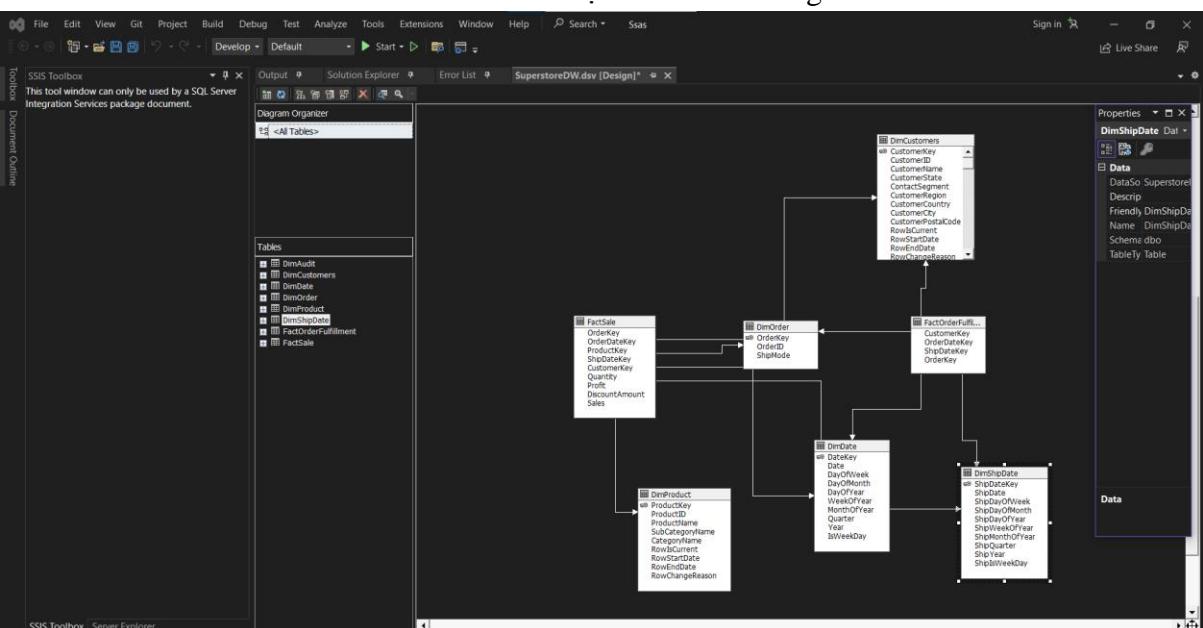
- Đặt tên cho Data Source View và nhấn Finish để hoàn tất việc tạo Data Source View



- Kết quả sau Finish

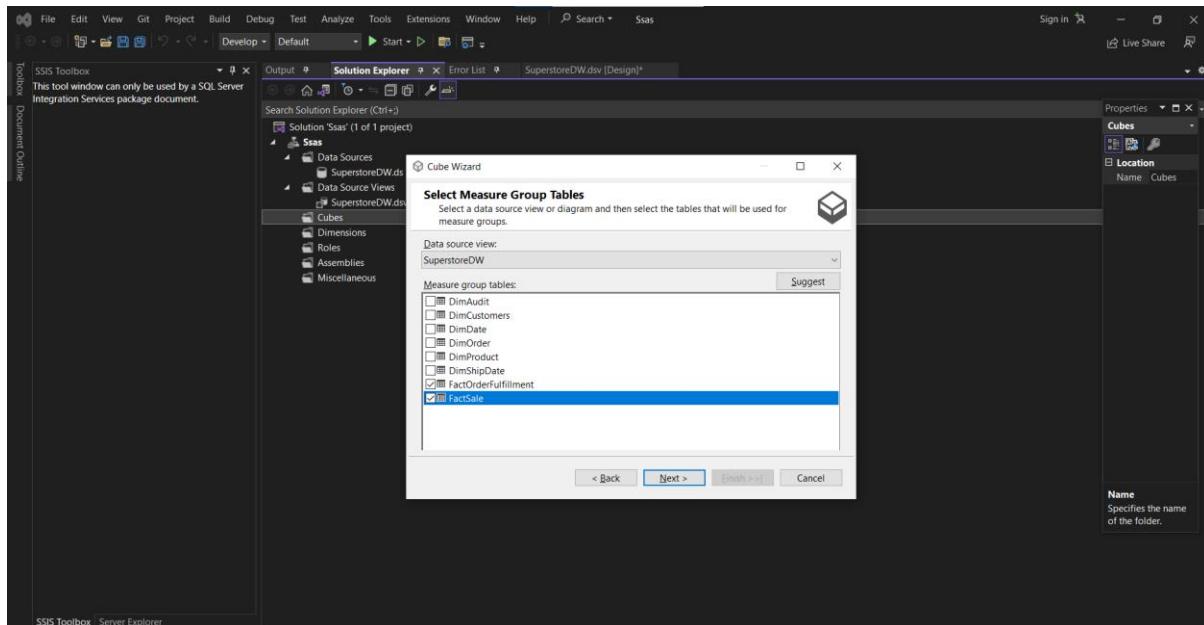


- Double-Click vào Data Source View vừa tạo để hiển thị diagram

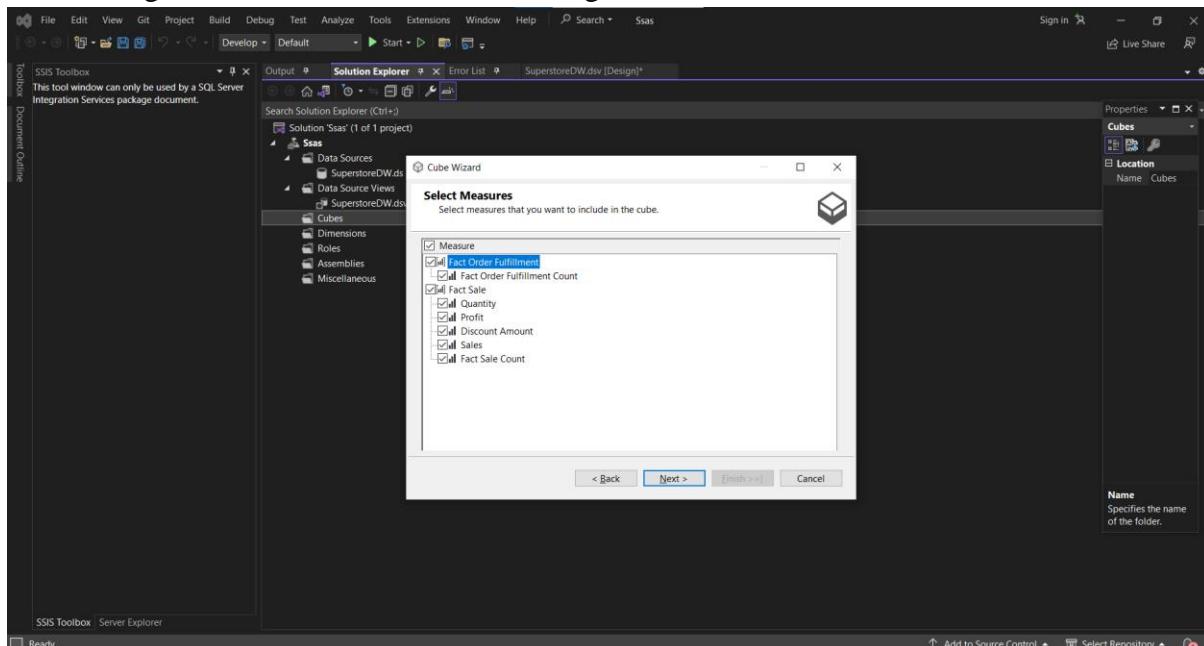


## 4.2. Quá trình xây dựng khối Cube

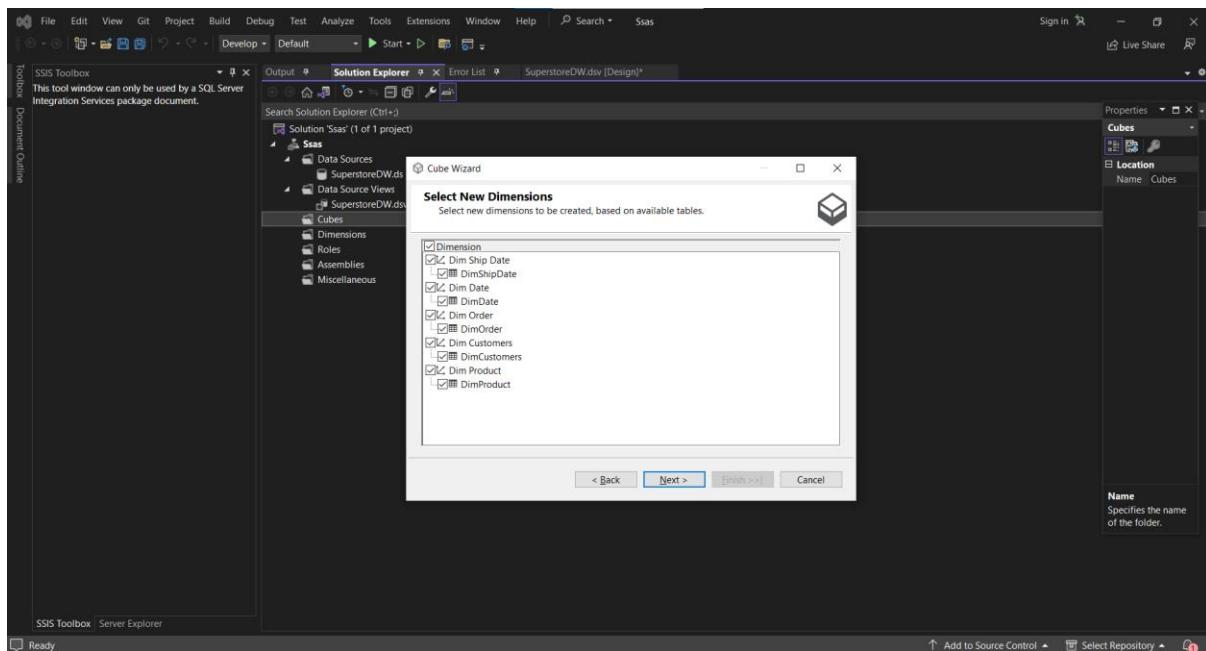
- Right-click vào **Cubes** ở khung Solution Explorer. Chon New Cube để tạo Cube mới.
- Chọn Data Source View là **SuperstoreDW**. Tích chọn 2 bảng **FactSale** và **FactOrderFulfilment**, nhấn Next



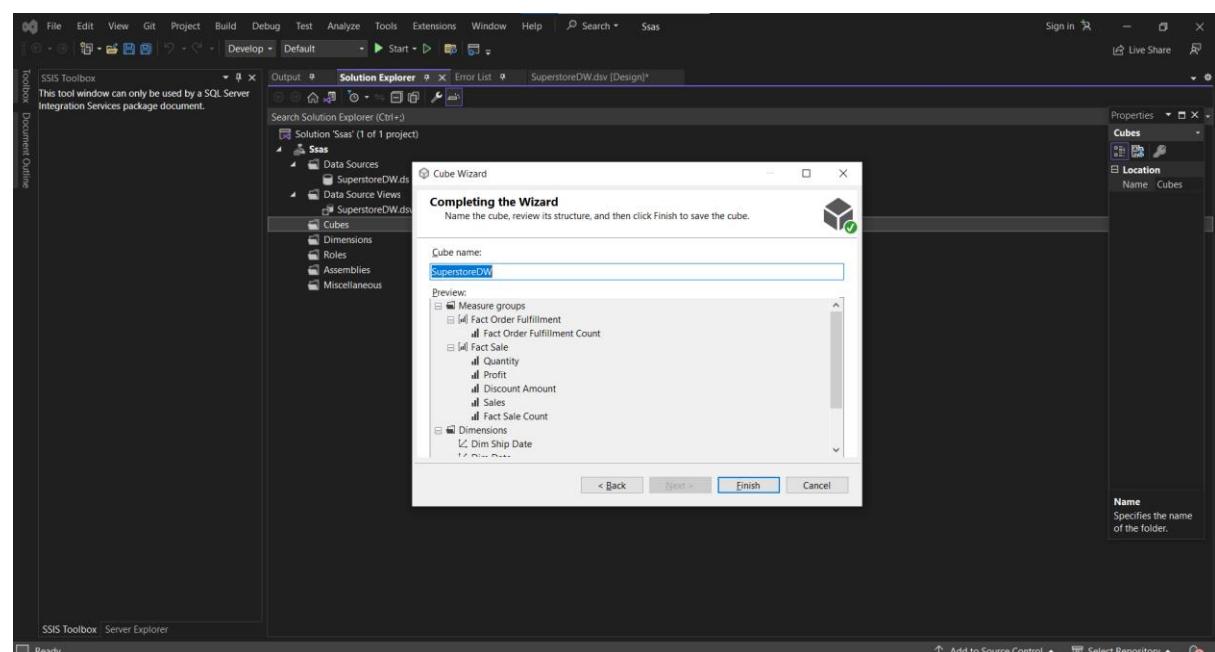
- Ở dialog Select Measures, tích chọn những measure cần thiết các mục sau



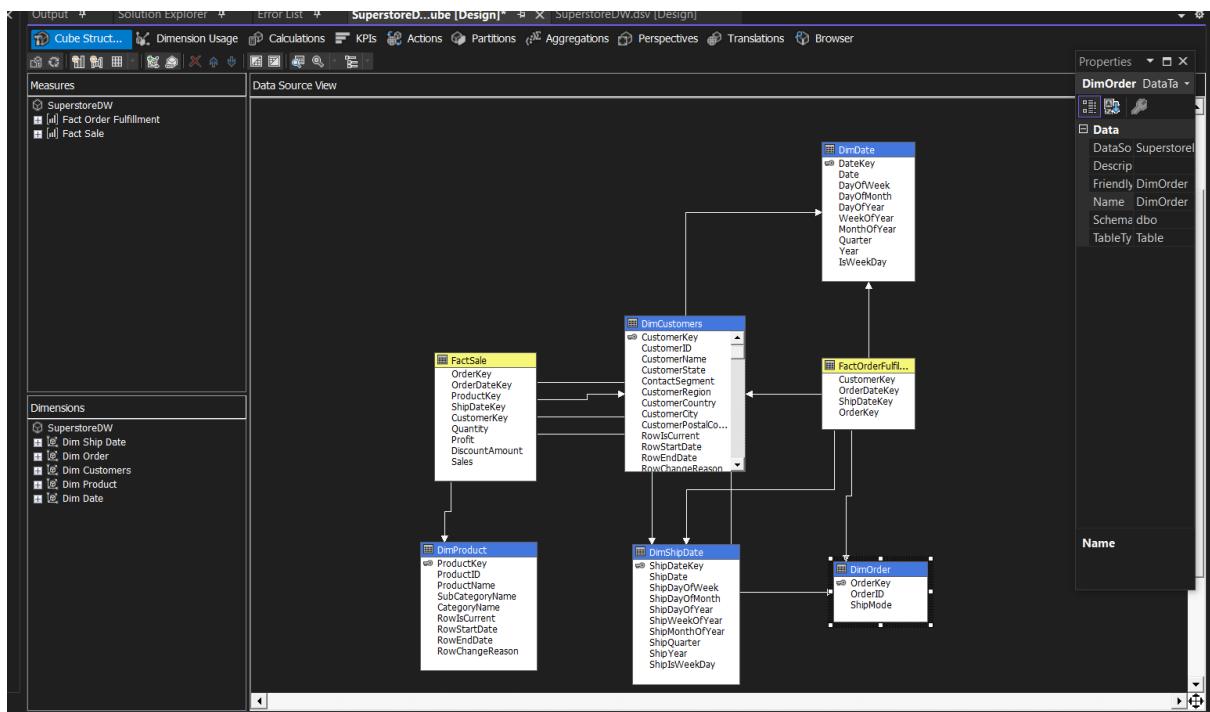
- Khởi tạo các bảng Dimension



- Đặt tên **Hardware DW Cube** để hoàn tất khởi tạo.



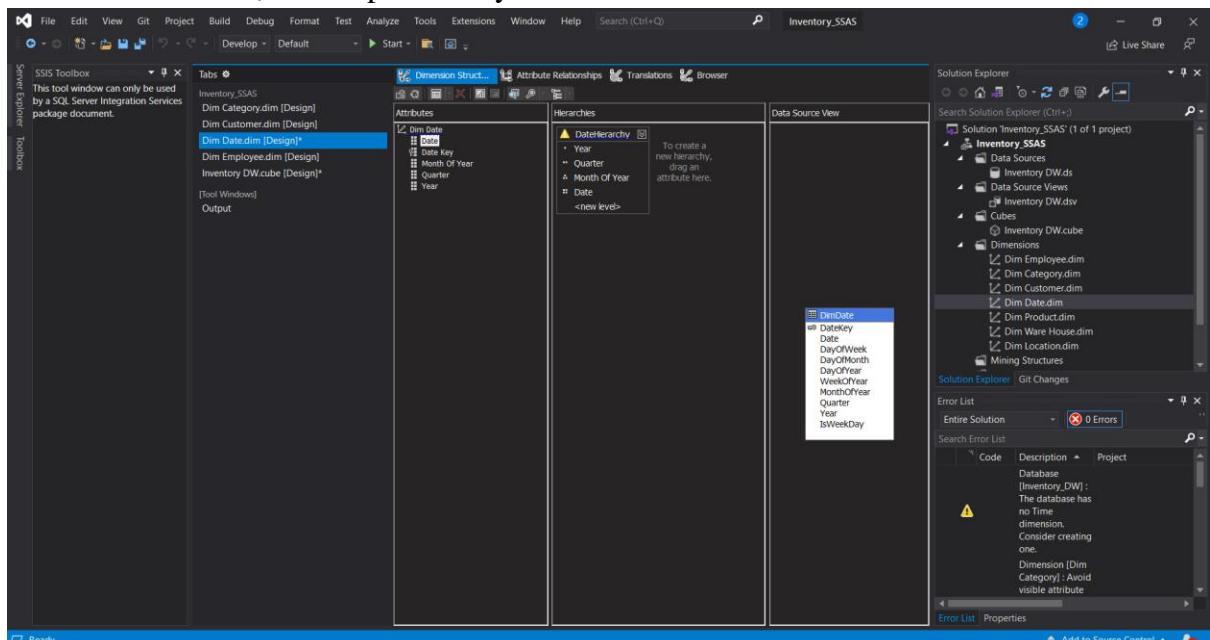
- Các bảng Fact, Dimension và quan hệ giữa chúng sẽ được thể hiện trong Diagram



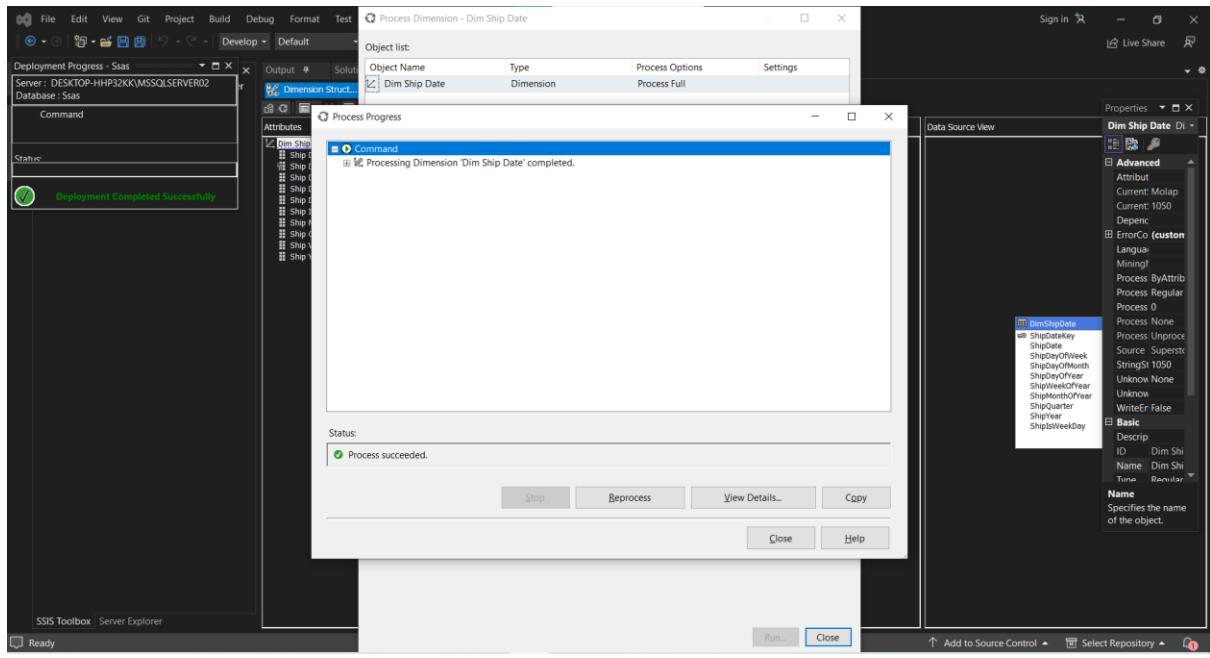
## 4.3. Cấu hình Hierarchy

### 4.3.1. Tạo Hierarchy cho Dim ShipDate

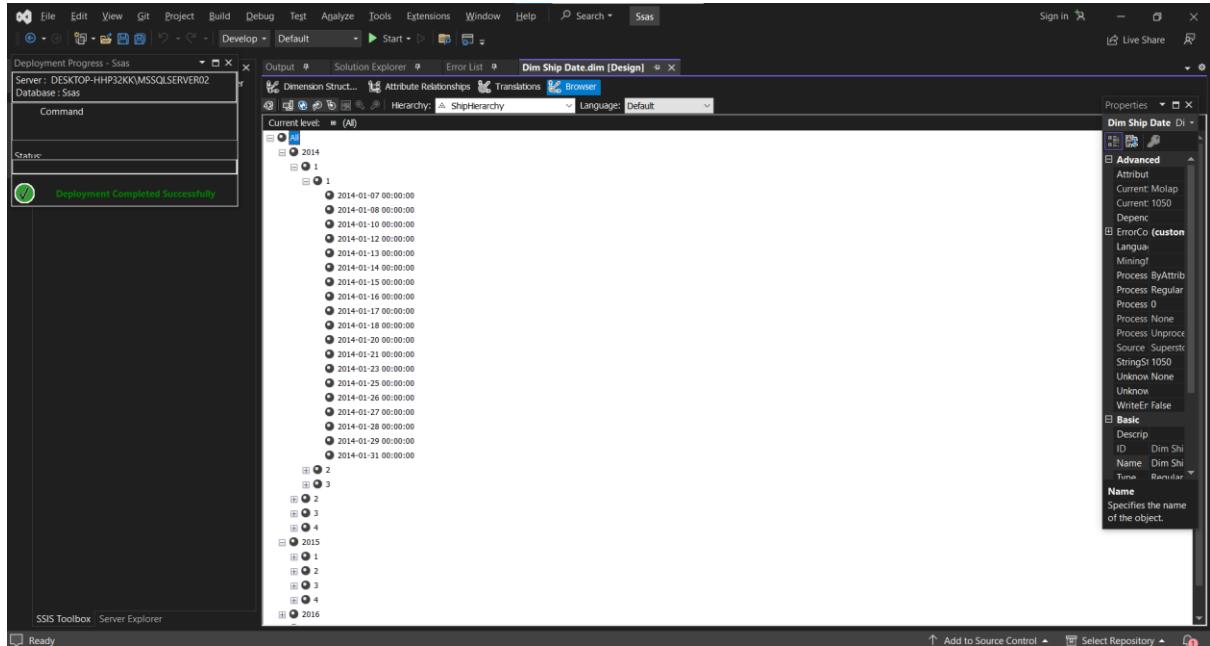
- Thêm thuộc tính vào bảng Dim: Tại table DimShipDate trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.
- Kéo thả 4 thuộc tính *Year*, *Quarter*, *Month Of Year*, *Date* từ Attributes vào panel Hierarchies để tạo ra ShipHierarchy



- Process để load dữ liệu cho bảng Dim ShipDate

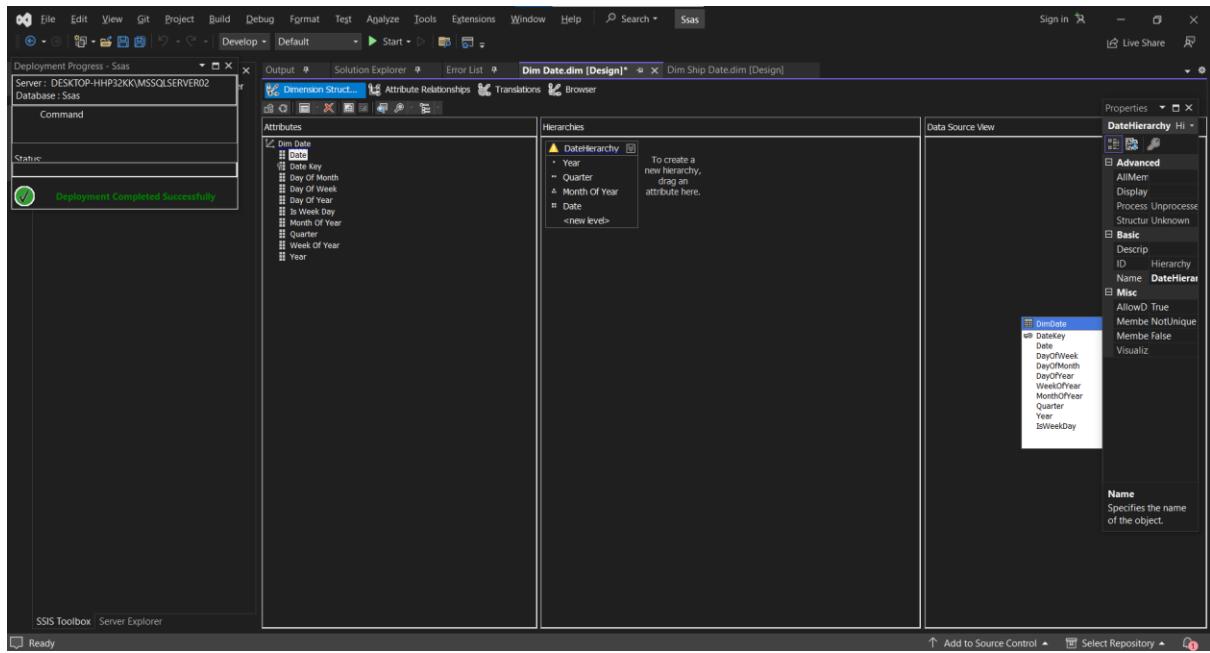


- Chuyển đến tab Browser. Ta thấy thời gian được phân cấp theo năm → quý → tháng.

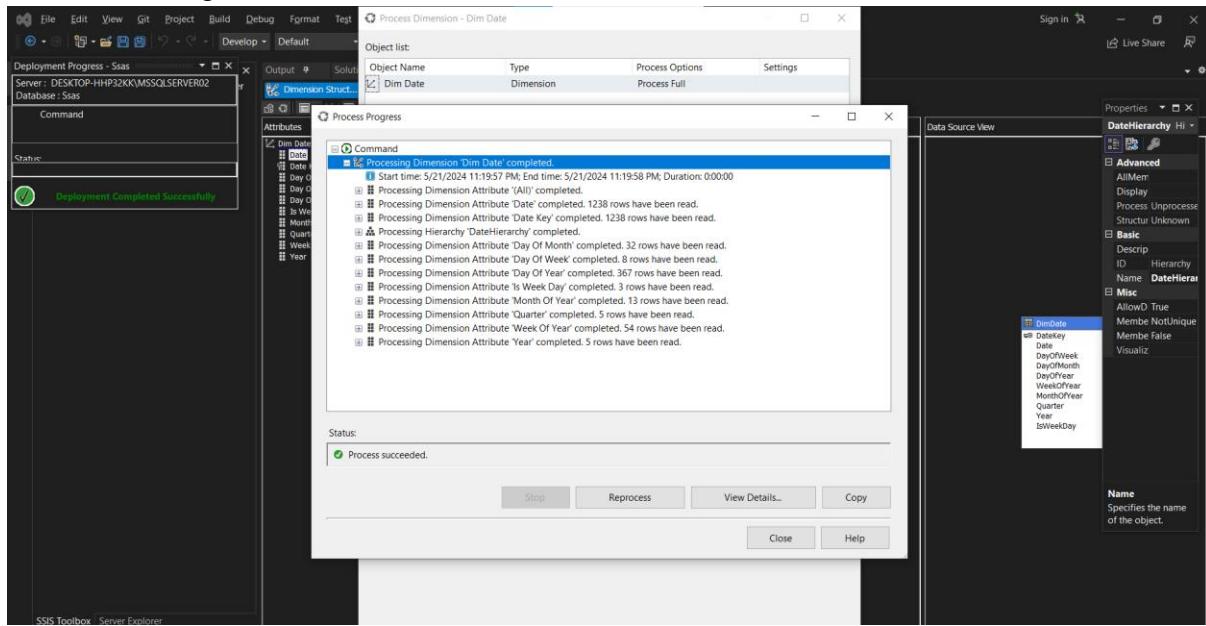


### 4.3.2. Tạo Hierarchy cho Dim Date

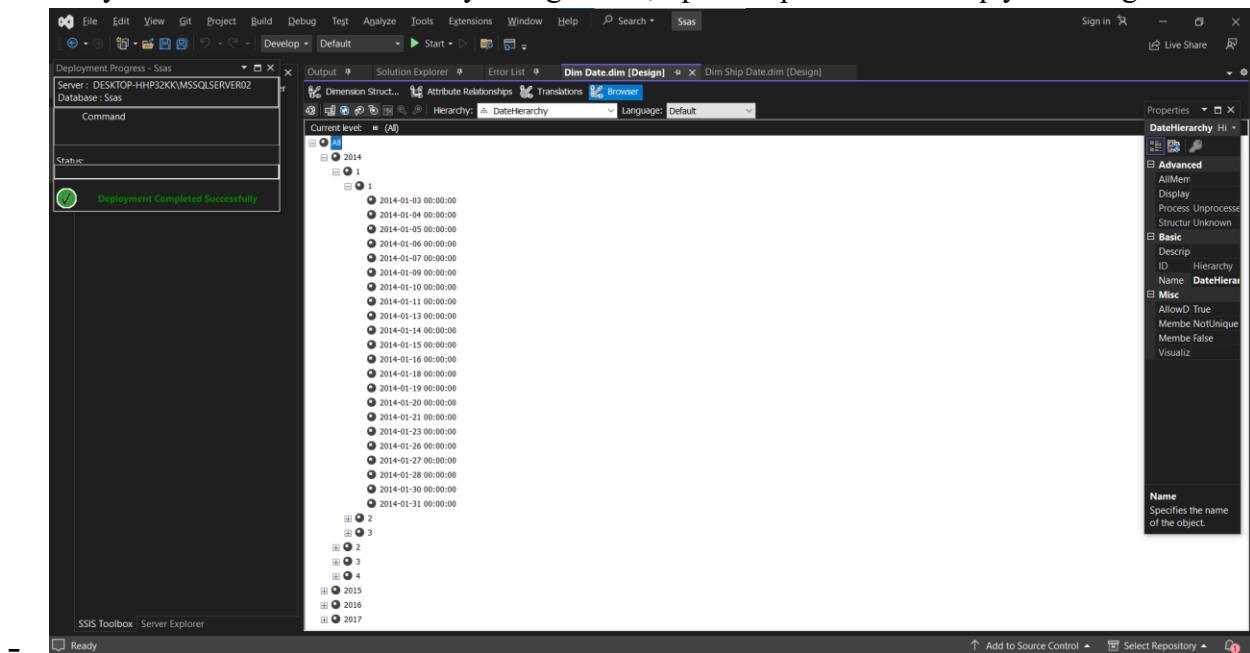
- Tại table Dim Date trong panel Data Source View, kéo 4 thuộc tính như hình vào panel Attributes. Sau đó kéo thả lần lượt Year, Quarter, Month of year, Date vào Hierarchies để tạo DateHierarchy



### - Process bảng Dim Date

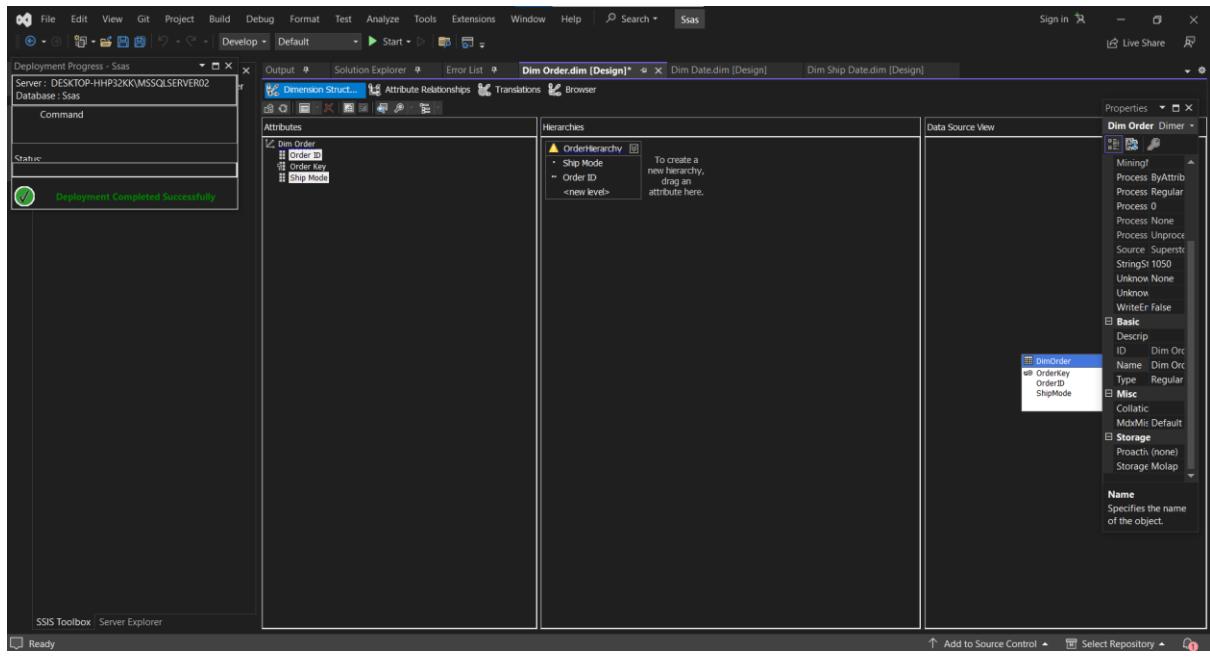


- Chuyển đến tab Browers. Ta thấy thời gian được phân cấp theo năm → quý → tháng.

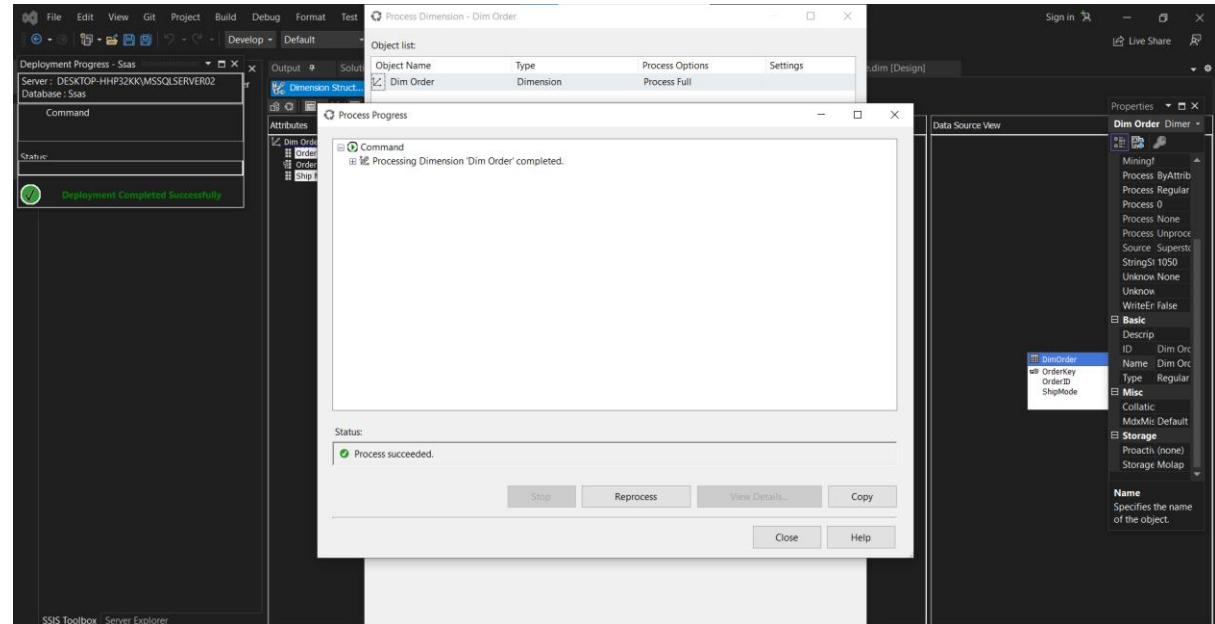


### 4.3.3. Tạo Hierarchy cho Dim Order

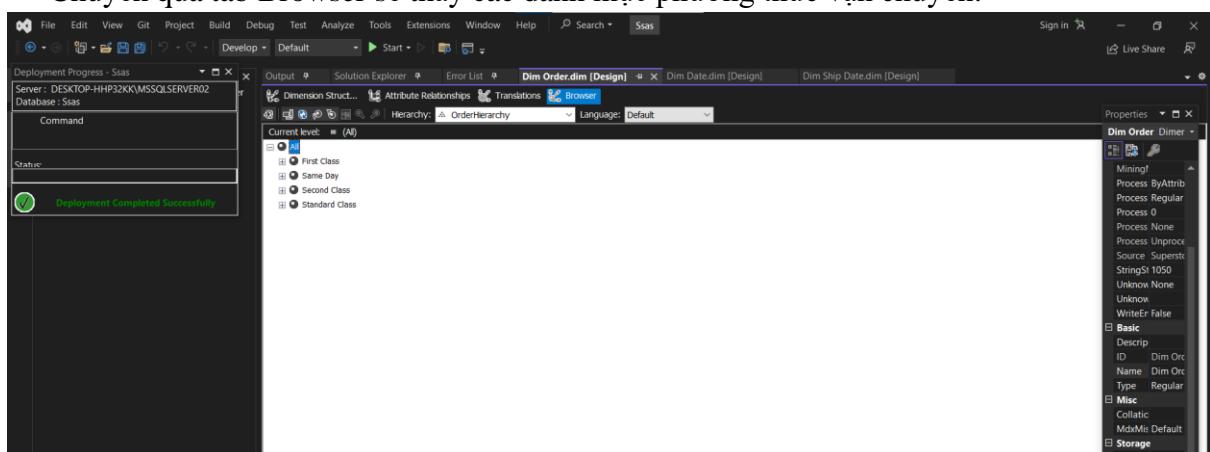
#### - Tạo OrderHierarchy



### - Process bảng Dim Order

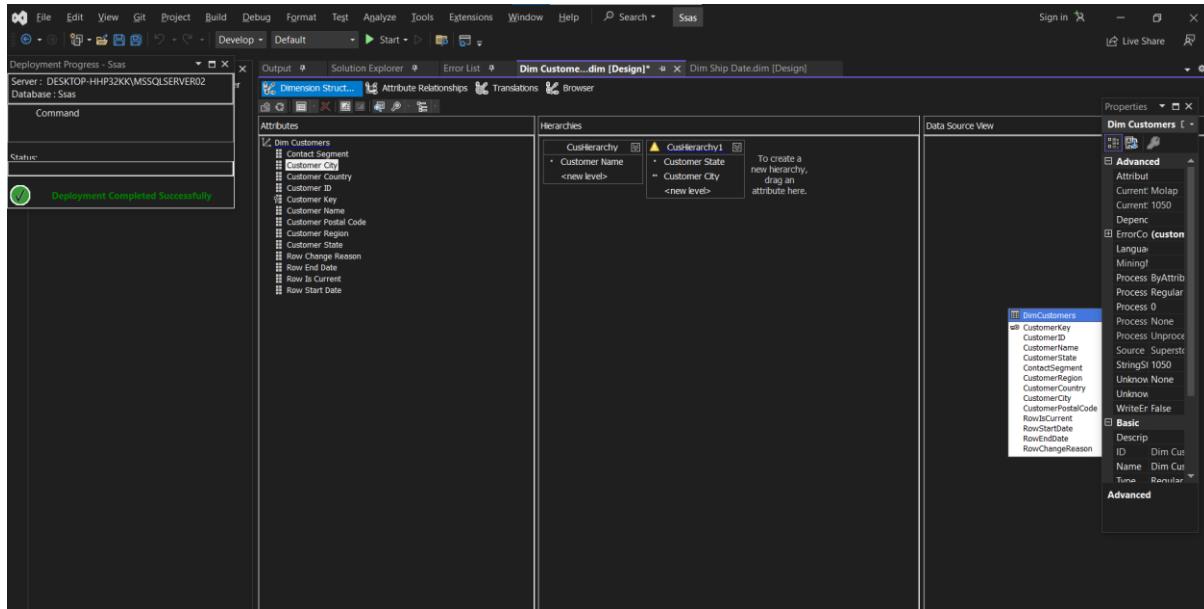


### - Chuyển qua tab Browser sẽ thấy các danh mục phương thức vận chuyển:

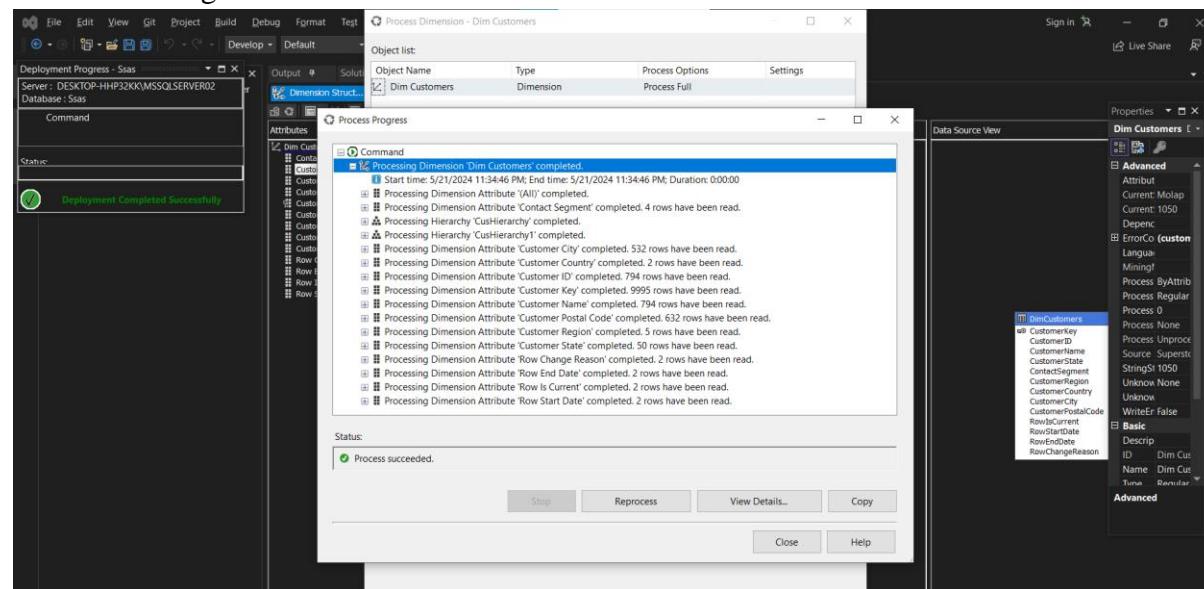


#### 4.3.4. Tạo Hierarchy cho Dim Customer

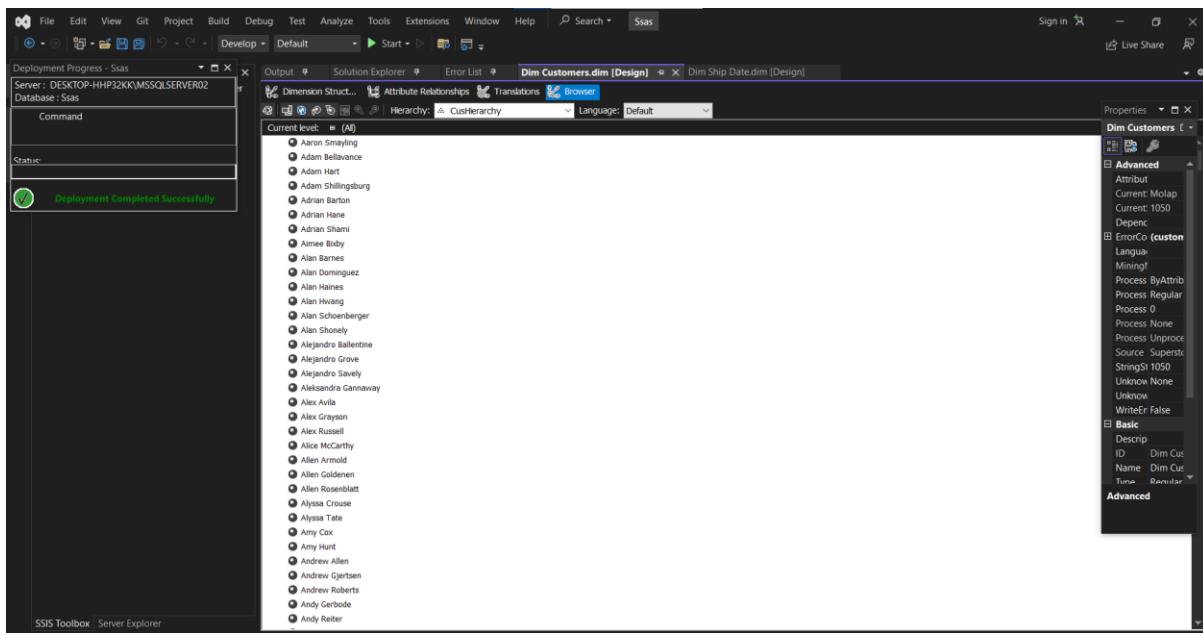
- Tạo CusHierarchy và CusHierarchy1



- Process bảng Dim Customer

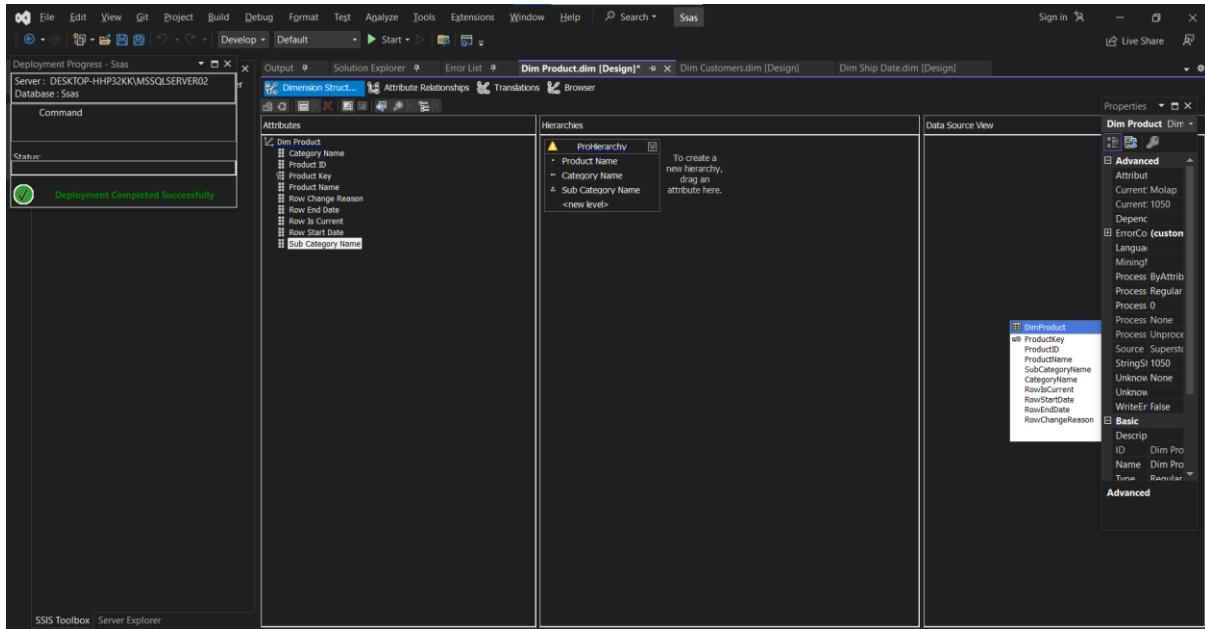


- Chuyển qua tab Browser Danh sách các khách hàng từng mua sản phẩm:

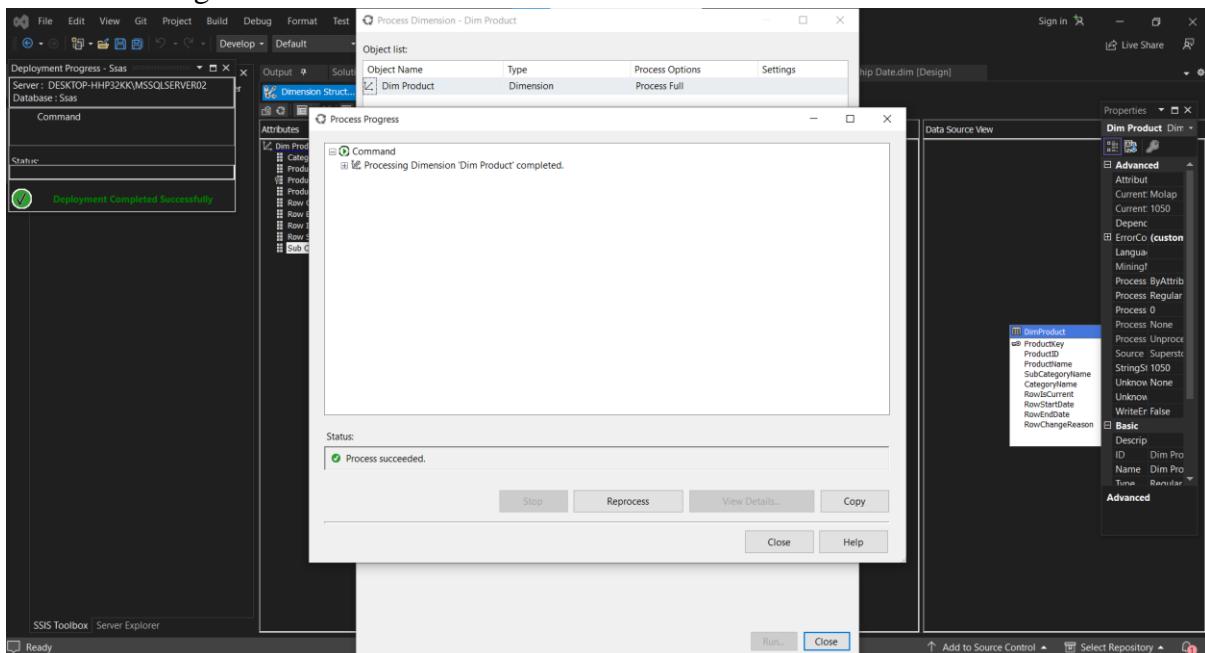


#### 4.3.5. Tạo Hierarchy cho Dim Product

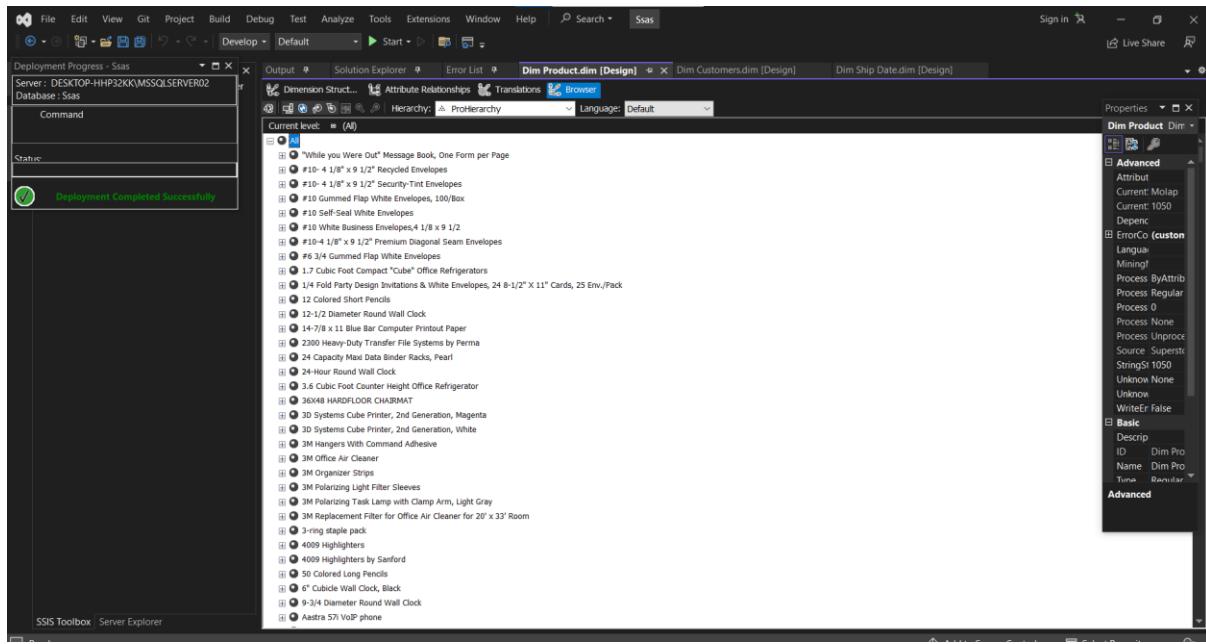
- Tạo ProHierarchy



- Process bảng Dim Product



- Chuyển đến tab browser để xem danh sách sản phẩm:



## 4.4. Thực hiện phân tích dữ liệu

### 4.4.1. Câu hỏi: Số lượng các loại sản phẩm bán ra trong tháng/ quý/ năm.

#### 4.4.1.1. Sử dụng công cụ SSAS

- Hiển thị số lượng các loại hàng và tên sản phẩm bán được từ trước đến nay

The screenshot shows a PivotTable in Power Pivot for Excel. The PivotTable has several dimensions across the top row: Dim DateYear, Dim DateQuarterQuarter, Dim DateMonth Of YearMonth Of Year, Dim ProductCategory NameCategory Name, Dim ProductProduct NameProduct Name, and MeasuresQuantity. The data rows show various products and their quantities, such as 'Adams Write n' Stick Phone Message Book' and 'Belkin Premiere Surge Master II 8-outlet su...'. The bottom status bar indicates 'Query Query1' and 'Records 1 of 9,235'.

Giải thích: Nhìn vào hình ta biết được tháng 12 năm 2014 chủ yếu bán các mặt hàng thuộc văn phòng phẩm với các sản phẩm có số lượng chủ yếu là 2 cái.

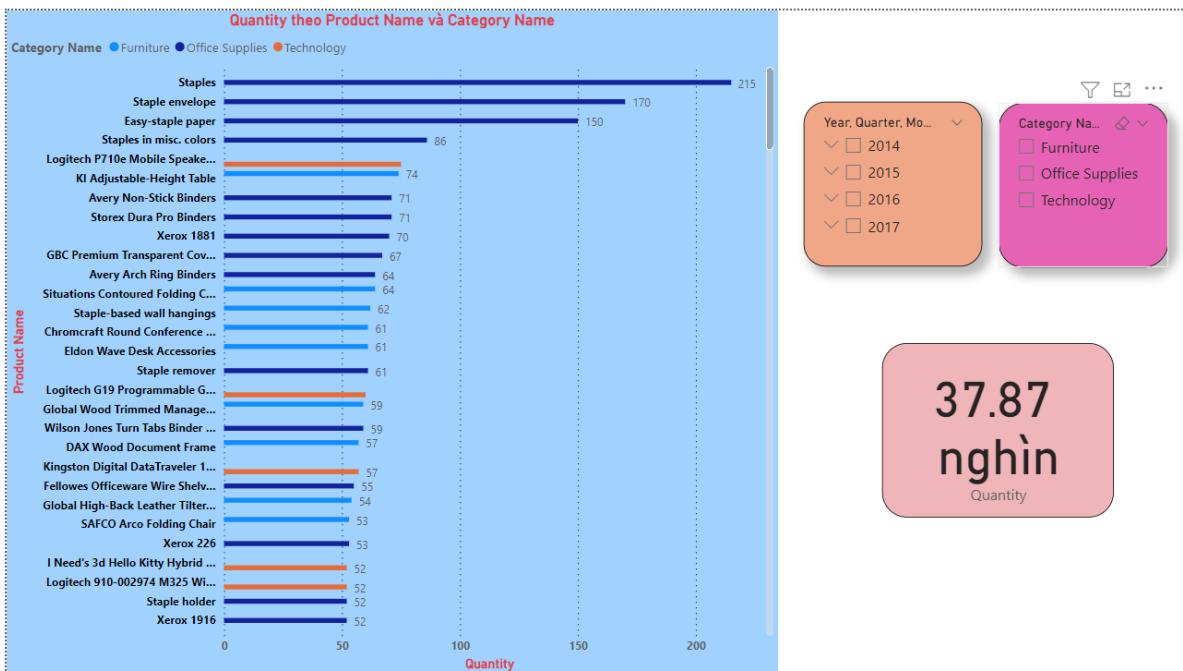
#### 4.4.1.2. Sử dụng Pivot Table trong Excel

- Hiển thị theo thời gian năm/ quý/ tháng

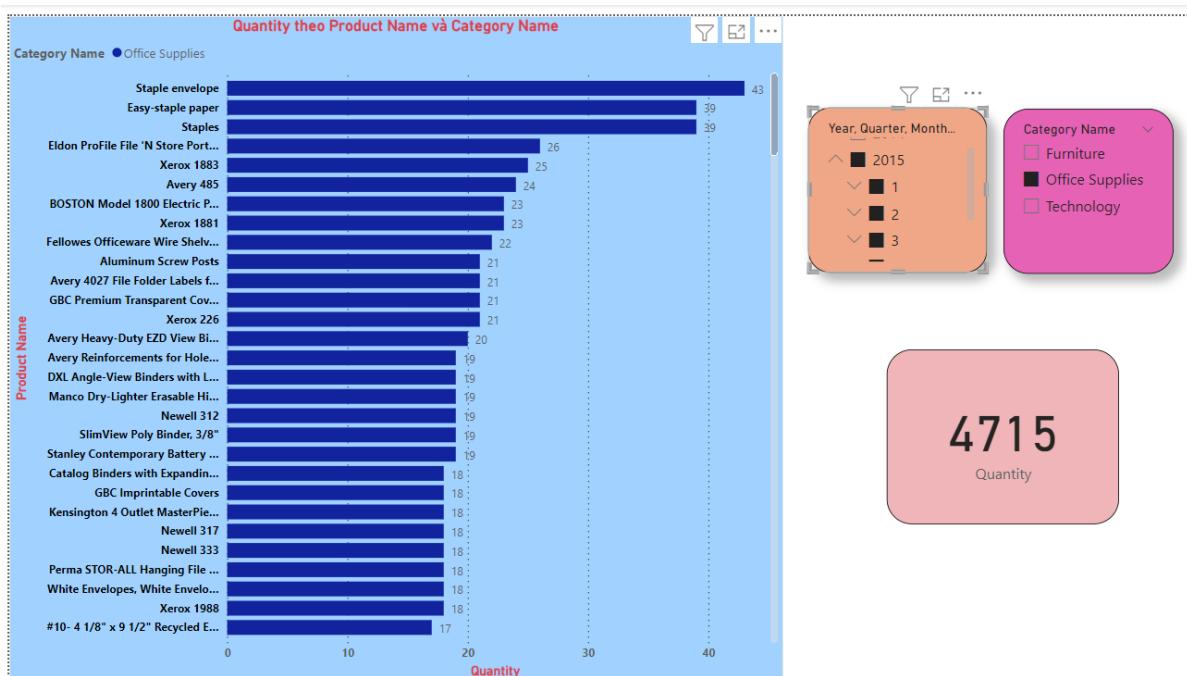
The screenshot shows an Excel spreadsheet with a PivotTable in the center. The PivotTable has 'Quantity' as the value field, 'Furniture' as the column label, and 'DateHierarchy' as the row label. The data shows sales for years 2014, 2015, 2016, and 2017. The PivotTable Tools ribbon tab is selected, and the 'PivotTable Fields' pane is open on the right side, showing fields like Product ID, Product Key, Product Name, etc., with 'Category Name' assigned to the columns and 'Quantity' assigned to the values.

- Giải thích: Với hình trên thì ta cũng có thể thấy được số lượng hàng bán ra hiển thị theo năm. Nhìn vào đây có thể thấy rằng số lượng mặt hàng bán ra vào năm 2014 còn khá ít và sau đó tăng dần theo thời gian. Trong đó, đơn hàng thuộc văn phòng phẩm chiếm số lượng nhiều nhất.

#### 4.4.1.3. Sử dụng Power BI



- Biểu đồ thể hiện các sản phẩm có số lượng được bán nhiều nhất theo danh mục sản phẩm và theo thời gian( Năm, quý, tháng, ngày). Nhìn vào biểu đồ, ta thấy được Staples thuộc văn phòng phẩm là sản phẩm bán được nhiều nhất trong 4 năm với số lượng 215 cái. Do đó, những năm tiếp theo nên đẩy mạnh sản xuất sản phẩm này và phát triển nó thành phiên bản mới.



- Ở biểu đồ trên, có thể thấy vào năm 2015, cửa hàng đã bán được tổng cộng 4715 sản phẩm trong danh mục văn phòng phẩm (Office Supplies). Ba sản phẩm bán chạy nhất trong danh mục này là phong bì kẹp ghim (Staple envelope), giấy dẽ ghim (Easy-staple

paper), và ghim kẹp (Staples). Những sản phẩm này đã đáp ứng tốt nhu cầu của khách hàng, góp phần vào thành công về doanh số của cửa hàng trong năm đó.

#### 4.4.2. Câu hỏi: Cho biết doanh số các sản phẩm trong tháng/ quý/ năm.

##### 4.4.2.1. Sử dụng công cụ SSAS

- Doanh số của các sản phẩm

The screenshot shows the Table Import Wizard dialog box in the Power Pivot ribbon. The 'Import from' section is set to 'SuperstoreDW'. The 'Dimension' dropdown is set to '<Select dimension>'. The 'Hierarchy' dropdown is set to 'Year, Quarter, Month'. The 'Operator' dropdown is set to 'Sum'. The 'Filter Expression' dropdown is set to 'Product Name'. The 'Param.' dropdown is set to 'Sales'. The main grid displays data from the 'Sales' table, including columns for Year, Quarter, Month, Date, Category Name, Product Name, and Sales. The data shows various office supplies and technology items sold in 2014.

The screenshot shows the Power Pivot ribbon with the 'Data View' tab selected. The main area displays a large table of sales data for 2017, organized by Date, Product Category, and Product Name. The table includes columns for Dim Date, Dim Date Hierarchy, Dim Date Hierarchy Year, Dim Date Hierarchy Quarter, Dim Date Hierarchy Month Of Year, Dim Date Hierarchy Date, Dim Product Category, Dim Product Product Name, and Measures Sales. The data shows a variety of products like Office Supplies, Xerox paper, and Fellowes binders.

- Nhìn vào hình, ta biết được sản phẩm GBC Therma-A-Bind 250T có doanh số cao nhất trong quý 4 năm 2017.

##### 4.4.2.2. Sử dụng Pivot Table trong Excel

- Doanh số các danh mục sản phẩm theo thời gian

**PivotTable Fields**

- Show fields: (All)
- Search:
- Fact Sale:
  - Sales
  - Profit
  - Quantity
  - Discount Amount
  - Fact Sale Count
- DateHierarchy

**Data Tools**

- Get & Transform Data
- Refresh All
- Workbook Links
- Sort & Filter
- Text to Columns
- Remove Duplicates
- Relationships
- What-If Forecast Analysis
- Manage Data Model
- Outline
- Analysis

**Book1 - Excel**

**Data**

**Review**

**View**

**Help**

**Power Pivot**

**PivotTable Analyze**

**Design**

**Tell me what you want to do**

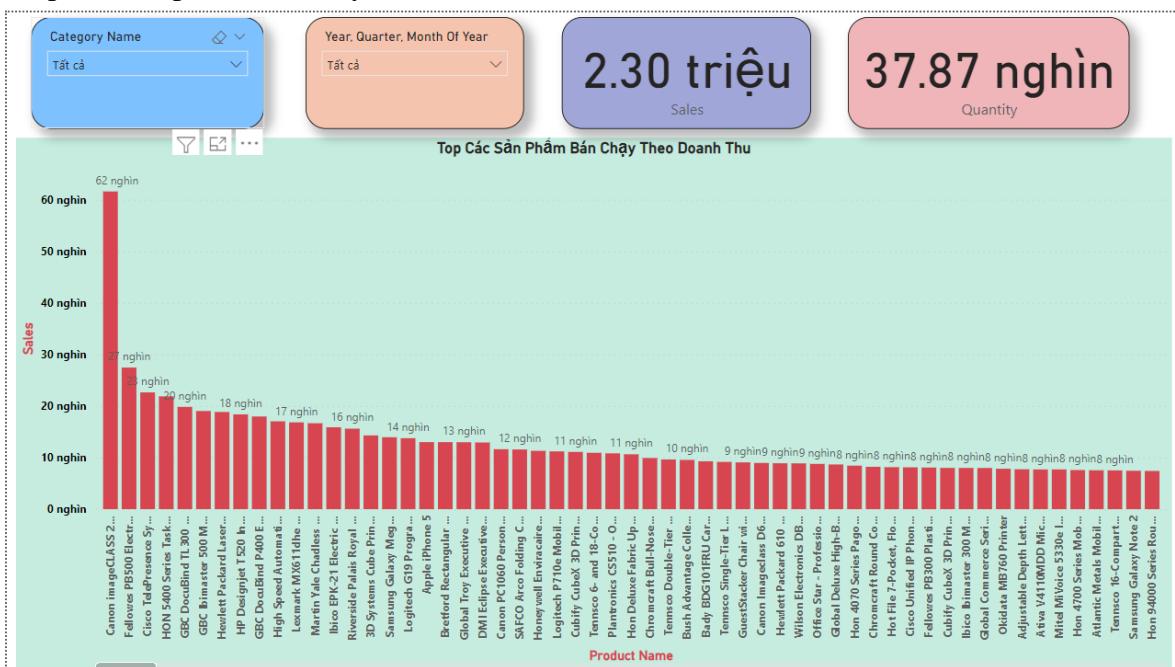
**File Home Insert Page Layout Formulas**

Sales	Column Labels	C	D	E	F	G	H
1	Row Labels	2014	2015	2016	2017	Grand Total	
3	Furniture	151792.8529	170518.237	198901.4352	215387.2685	741999.7937	
4	Office Supplies	151776.4115	137233.4632	183939.9824	246097.1742	719047.0313	
5	Technology	175278.2331	162780.8091	226364.1785	271730.8102	836154.0309	
6	Grand Total	484247.4975	470532.5093	609205.5962	73215.2528	2297200.856	
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							

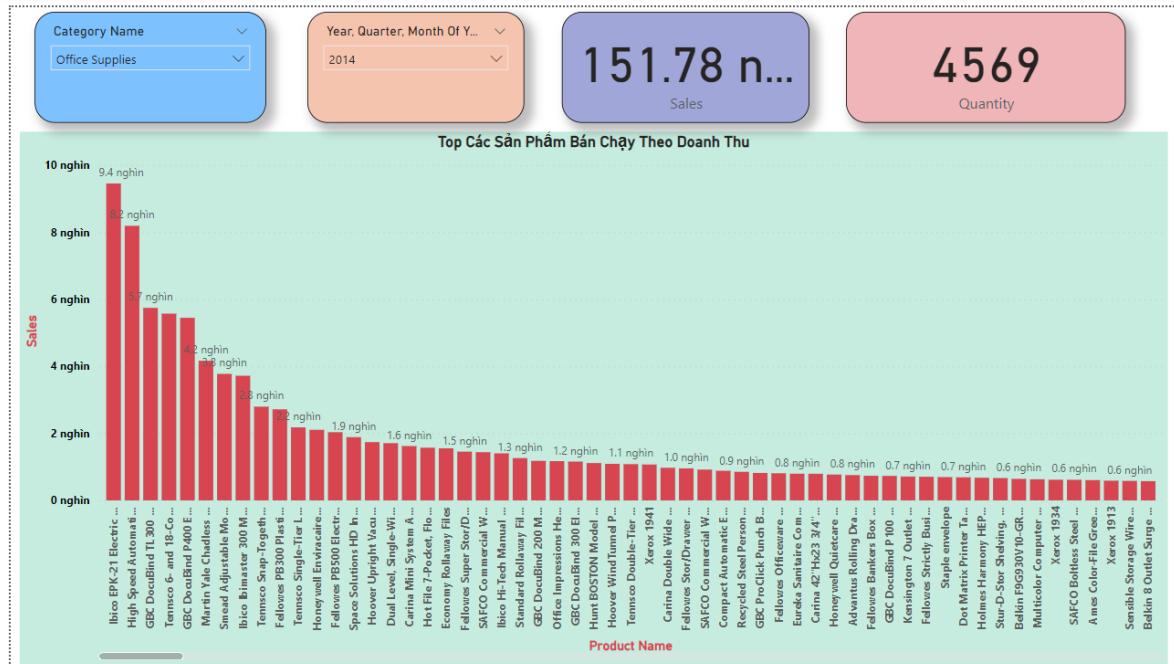
Giải thích: Hình trên thể hiện doanh thu các mặt hàng theo năm có thể thấy được Năm 2017 là năm có tổng doanh thu cao nhất với hơn 700 nghìn đô la và mặt hàng có doanh số cao nhất chính là Furniture(đồ nội thất) với hơn 740 nghìn đô la.

#### 4.4.2.3. Sử dụng Power BI

- Top các sản phẩm bán chạy dựa theo doanh thu



- Nhìn vào biểu đồ trên, ta thấy rằng từ năm 2014 đến 2017, sản phẩm Cannon imageCLASS đạt doanh số cao nhất, thu về 62 nghìn đô la. Điều này cho thấy sản phẩm rất được ưa chuộng, nhờ chất lượng in ấn vượt trội và dịch vụ hậu mãi tốt, giúp nó duy trì vị thế hàng đầu trên thị trường trong suốt giai đoạn này.
- Ngoài ra chúng ta có thể xem mặt hàng có sản phẩm bán chạy nhất theo thời gian bằng bộ lọc.



- Có thể thấy vào năm 2014, sản phẩm Ibico EPK-21 Electronic là sản phẩm có doanh thu cao nhất đạt 9.4 nghìn đô trong danh mục Văn phòng phẩm(Office Supplies)

#### 4.4.3. Câu hỏi: Những danh mục sản phẩm và danh mục phụ nào đóng góp nhiều nhất vào doanh thu bán hàng theo tháng/ quý/ năm?

##### 4.4.3.1. Sử dụng công cụ SSAS

- Hiển thị Sales và Profit của từng danh mục sản phẩm và danh mục phụ theo thời gian.

Table Import Wizard

Dimension Hierarchy Operator Filter Expression Parameters

Measure Group: <All>

SuperstoreDW

Metadata

Measure Group: <All>

Dimension Hierarchy Operator Filter Expression Parameters

Year Quarter Month Of Year Category Name Sub Category Name Profit Sales

Year	Quarter	Month Of Year	Category Name	Sub Category Name	Profit	Sales
2014	1	1	Furniture	Bookcases	-326...	1010...
2014	1	1	Furniture	Chairs	105...	4187...
2014	1	1	Furniture	Furnishings	90.8...	711...
2014	1	1	Furniture	Tables	-16....	333
2014	1	1	Office Supplies	Appliances	100...	312...
2014	1	1	Office Supplies	Art	52.5...	176...
2014	1	1	Office Supplies	Binders	190...	814...
2014	1	1	Office Supplies	Envelopes	78.4...	194...
2014	1	1	Office Supplies	Fasteners	0.31...	37.0...
2014	1	1	Office Supplies	Labels	15.7...	44.7...
2014	1	1	Office Supplies	Paper	101...	242...
2014	1	1	Office Supplies	Storage	249...	3028...
2014	1	1	Technology	Accessories	334...	827...
2014	1	1	Technology	Phones	521...	2315...
2014	1	2	Furniture	Chairs	-9.0...	190...
2014	1	2	Furniture	Furnishings	54.3...	392...
2014	1	2	Furniture	Tables	75.3...	1256...
2014	1	2	Office Supplies	Appliances	9.01...	89.9...
2014	1	2	Office Supplies	Art	12.2...	73.6...
2014	1	2	Office Supplies	Binders	61.5...	339...
2014	1	2	Office Supplies	Fasteners	5.80	12.3

OK Cancel

The screenshot shows a Microsoft Excel spreadsheet titled "Power Pivot for Excel - Book1". The ribbon tabs include File, Home, Design, Advanced, Paste, Copy, Paste Replace, From Database, From Service, From Other Sources, Existing Connections, Refresh, PivotTable, Data Type, Sort A to Z, Sort Z to A, Clear All, Sort by Column, Find, AutoSum, Create KPI, Data View, Diagram View, Show Hidden, Calculation Area, and View. The main area displays a PivotTable with the following structure:

	Dim ProductCategory NameCategory Name	Dim ProductSub Category NameSub Category Name	Sales	Profit	Total Profit
1 2014	1	1	Office Supplies	Appliances	100.067996501...
2 2014	1	1	Office Supplies	Art	52.548899769783
3 2014	1	1	Office Supplies	Binders	190.7386970520...
4 2014	1	1	Office Supplies	Envelopes	78.47200012207...
5 2014	1	1	Office Supplies	Fasteners	0.311199992895...
6 2014	1	1	Office Supplies	Labels	15.77929949760...
7 2014	1	1	Office Supplies	Paper	101.5603990554...
8 2014	1	1	Office Supplies	Storage	249.4720983505...
9 2014	1	2	Office Supplies	Appliances	9.01879882125
10 2014	1	2	Office Supplies	Art	12.21359968185...
11 2014	1	2	Office Supplies	Binders	61.55069839954...
12 2014	1	2	Office Supplies	Fasteners	5.804500102996...
13 2014	1	2	Office Supplies	Labels	16.41329956054...
14 2014	1	2	Office Supplies	Paper	56.50009965896...
15 2014	1	2	Office Supplies	Storage	9.817600488662...
16 2014	1	2	Office Supplies	Supplies	4.772399902343...
17 2014	1	3	Office Supplies	Appliances	-361.546808242...
18 2014	1	3	Office Supplies	Art	108.3423999845...
19 2014	1	3	Office Supplies	Binders	381.3023970127...
20 2014	1	3	Office Supplies	Envelopes	121.5210962295...
21 2014	1	3	Office Supplies	Fasteners	4.085399985313...

Nhìn vào hình ta thấy được sản phẩm Binders thuộc danh mục Văn phòng phẩm(Office Supplies) thu về lợi nhuận và doanh số cao nhất trong quý 1 năm 2014.

#### 4.4.3.2. Sử dụng Pivot Table trong Excel

- Hiển thị lợi nhuận và doanh thu của danh mục sản phẩm và danh mục phụ theo thời gian

The screenshot shows a Microsoft Excel spreadsheet with a PivotTable Fields pane open on the right side. The main area displays a PivotTable with the following structure:

	Column Labels	Sales	Profit	Total Sales	Total Profit
1	2014	2014	2014	2014	2014
2	2015	2015	2015	2015	2015
3	2016	2016	2016	2016	2016
4	Furniture				
5	Bookcases	20036.67757	38543.57256	26275.46651	30024.27937
6	Chairs	77241.57597	71734.52907	83198.64464	95554.35265
7	Furnishings	13826.23397	21089.71206	28784.12382	28915.09389
8	Tables	46088.36541	39150.42334	60833.20027	60893.54258
9	Office Supplies				
10	Appliances	15313.62506	23241.28909	26050.31515	42926.93201
11	Art	6057.98205	6236.83398	5960.90802	8863.06801
12	Binders	43488.2647	37453.09808	49683.32523	72788.04443
13	Envelopes	3855.75025	4512.187984	4729.889987	3378.573991
14	Fasteners	661.328005	545.223981	960.1339962	857.5939962
15	Labels	2841.393973	2956.461994	2827.240016	3861.21598
16	Paper	14838.95859	15287.63795	20661.89395	27694.71788
17	Storage	50329.04208	45048.24814	58788.70024	60977.61831
18	Supplies	14394.06969	1952.481981	14277.57584	16049.40957
19	Technology				
20	Accessories	25014.27192	40523.95995	41895.85402	59946.23199
21	Copiers	10849.7818	26179.45013	49599.40869	62899.38702
22	Machines	62023.37354	27763.69696	55906.88875	43544.67476
23	Phones	77390.80586	68313.70201	78962.03007	105340.5164
24	Grand Total	484247.4975	470532.5093	60925.5962	73215.2528

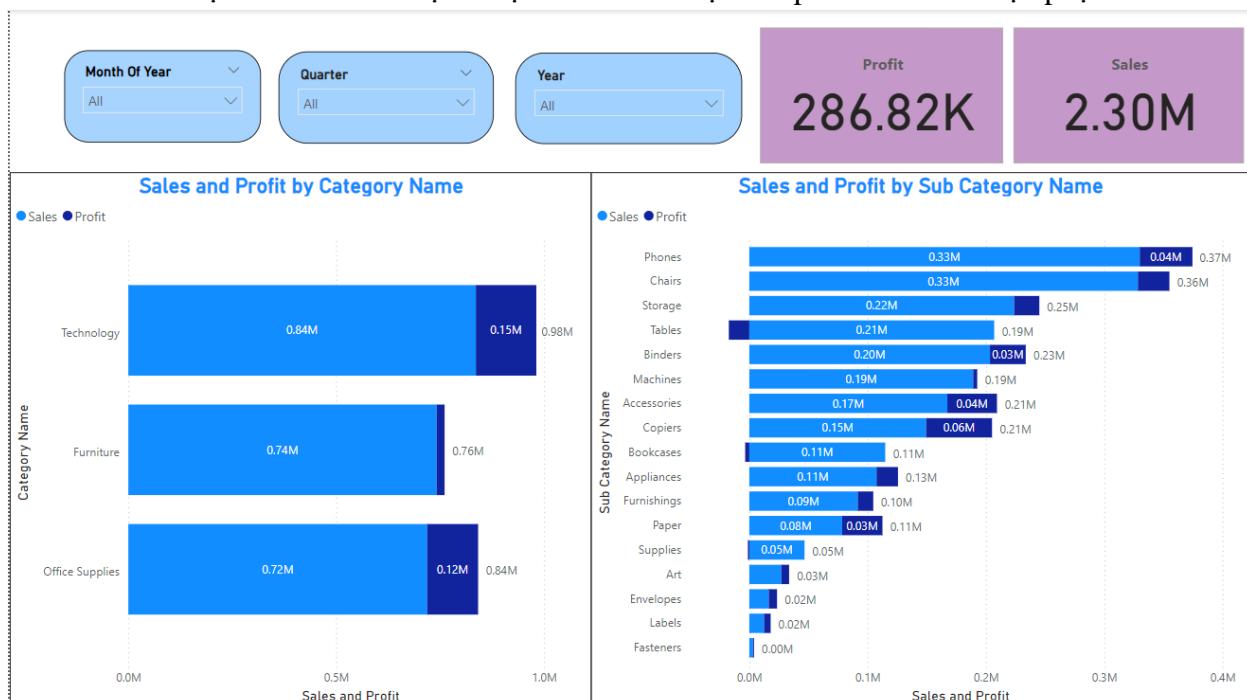
- Nhìn vào hình, ta có thể thấy được lợi nhuận và doanh thu tăng dần theo thời gian, cao nhất đều đạt ở năm 2017.
- Về doanh thu, danh mục có doanh thu cao nhất là Office Supplies vì danh mục này có số lượng mặt hàng nhiều nhất, danh mục phụ có doanh thu cao nhất là Phones(điện thoại) với doanh thu hơn 330 nghìn đô la.

- Về lợi nhuận, danh mục có lợi nhuận cao nhất là Office Supplies, danh mục phụ có lợi nhuận cao nhất là Copies với hơn 55 nghìn đô la.
- Ngoài ra, còn có thể xem doanh thu và lợi nhuận của các danh mục và danh mục phụ theo từng tháng/quý:

The screenshot shows an Excel spreadsheet titled "cau3 - Excel". The top ribbon includes tabs for File, Home, Insert, Page Layout, Formulas, Data, Review, View, Help, and Power Pivot. The Data tab is selected, showing options like Get & Transform Data, Queries & Connections, Sort & Filter, and Data Tools. The main content is a pivot table with columns for Sales, Profit, and various years (2014-2017) and categories (Furniture, Office Supplies, Technology). The pivot table displays detailed financial data for each category over time.

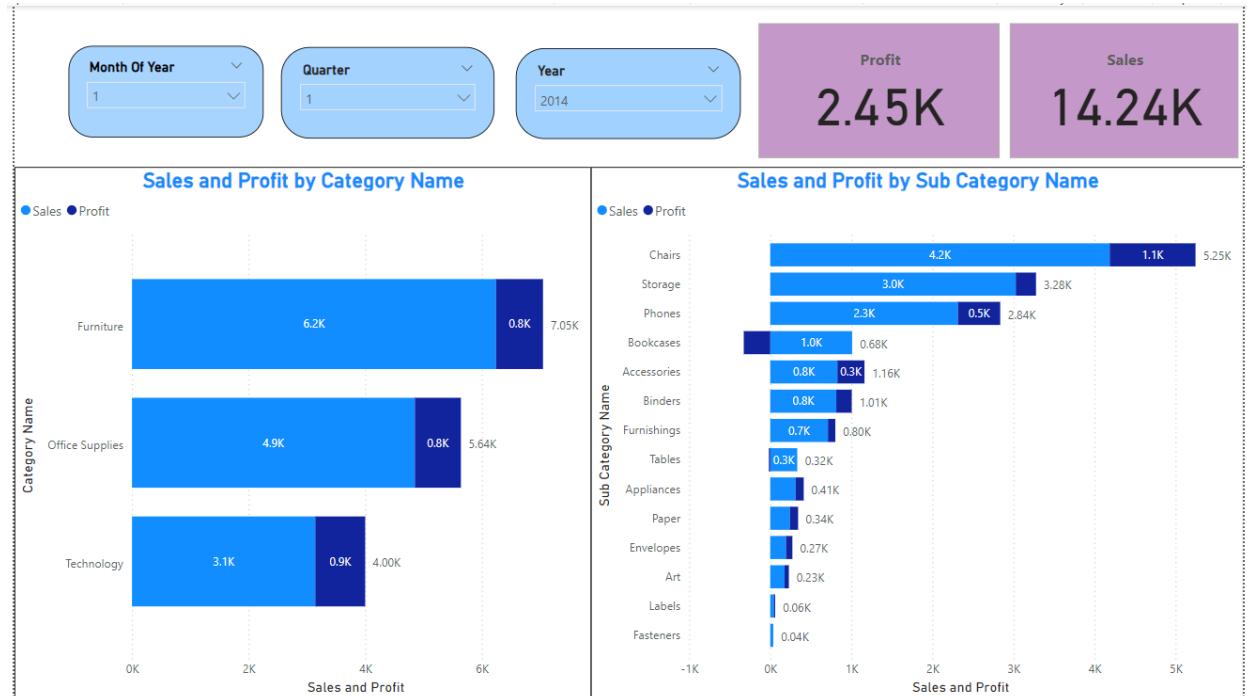
#### 4.4.3.3. Sử dụng Power BI

- Biểu đồ thể hiện doanh thu và lợi nhuận theo danh mục sản phẩm và danh mục phụ.



Giải thích: Số liệu được phân tích từ năm 2014 đến 2017. Ở biểu đồ hiển thị danh mục sản phẩm, ta thấy được Technology là danh mục thu về lợi nhuận và doanh thu cao nhất với tổng doanh thu và lợi nhuận là 0.98 triệu đô.

- Ở biểu đồ hiển thị danh mục phụ, ta thấy được Phones là danh mục phụ thu được lợi nhuận và doanh thu cao nhất với tổng là 0.37 triệu đô.
- Từ đó, kết luận được rằng ta nên đẩy mạnh phát triển danh mục Technology(công nghệ) và danh mục phụ Phones(điện thoại) cho siêu thị để nâng cao doanh số và lợi nhuận.
- Ngoài ra, có thể điều chỉnh biểu đồ để hiện thị doanh thu và lợi nhuận theo tháng /quý /năm.



#### 4.4.4. Câu hỏi: Phương thức vận chuyển của từng phân khúc khách hàng khác nhau theo tháng quý năm?

##### 4.4.4.1. Sử dụng công cụ SSAS

- Hiển thị phương thức vận chuyển của từng phân khúc khách hàng theo thời gian

Nhìn vào biểu đồ, ta thấy được đa số phân khúc Consumer đều sử dụng phương thức vận chuyển Standard Class phổ biến trong năm 2017.

#### 4.4.4.2. Sử dụng Pivot Table trong Excel

The screenshot shows a Microsoft Excel spreadsheet titled 'Book1 - Excel'. On the right, the 'PivotTable Fields' pane is open, showing fields for Contact Segment, Customer City, Customer Country, Customer ID, Customer Key, and Customer Name. The main area contains a PivotTable with data from 2014 to 2017. The columns are labeled 'Column Labels' (First Class, Same Day, Second Class, Standard Class, Grand Total) and rows are labeled 'Row Labels' (Year, Contact Segment). The data shows that Standard Class fulfillment counts are consistently high across all years and contact segments.

- Giải thích: chúng ta có thể thấy rằng trong suốt khoảng thời gian bốn năm, hầu hết các phân khúc khách hàng đều chủ yếu lựa chọn phương thức vận chuyển Standard Class. Sự phổ biến của phương thức vận chuyển này được thể hiện rõ ràng qua biểu đồ, cho thấy rằng Standard Class đã trở thành lựa chọn ưa thích của đa số khách hàng trong tất cả các nhóm khác nhau.

#### 4.4.4.3. Sử dụng Power BI

- Biểu đồ thể hiện phương thức vận chuyển của từng phân khúc khách hàng theo thời gian



Giải thích: Số liệu được phân tích từ năm 2014 đến 2017. Nhìn vào biểu đồ ta thấy rõ được rằng hầu hết mọi phân khúc khách hàng đều sử dụng phương thức vận chuyển Standard Class với số lượng vượt trội hơn so với 3 phương thức còn lại. Bất kể lý do là gì, sự ưu tiên cho

Standard Class trong mọi phân khúc khách hàng là một xu hướng rõ ràng và liên tục trong suốt bốn năm qua.

Do đó, chúng ta nên cân nhắc việc tuyển thêm nhân viên và phát triển thêm các dịch vụ cho phương thức vận chuyển Standard Class. Việc này nhằm mục đích cải thiện hiệu suất và đảm bảo rằng khách hàng có thể nhận hàng một cách nhanh chóng.

#### **4.4.5. Câu hỏi: Tháng có số lượng đơn hàng trong từng loại sản phẩm tăng cao theo từng phương thức vận chuyển.**

##### **4.4.5.1. Sử dụng công cụ SSAS**

- Hiển thị số lượng đơn hàng của từng danh mục sản phẩm và phương thức vận chuyển theo tháng.

The screenshot shows the 'Table Import Wizard' interface. On the left, there's a tree view of the 'SuperstoreDW' database's metadata, including 'Measure Group' (with 'All' selected) and 'Calculated Members'. In the center, a grid displays data from the 'Fact Order Fulfillment' table, filtered by 'Month Of Year' (1 or 10), 'Category Name' (Furniture, Office Supplies, Technology), 'Ship Mode' (First Class, Same Day, Standard ...), and 'Fact Order Fulfillment C...' (values ranging from 3 to 112). At the top, there are tabs for 'Dimension', 'Hierarchy', 'Operator', and 'Filter Expression', with 'Dimension' currently active. The bottom right of the window has 'OK' and 'Cancel' buttons.

Month Of Year	Category Name	Ship Mode	Fact Order Fulfillment C...
1	Furniture	First Class	25
1	Furniture	Same Day	3
1	Furniture	Second Cl...	38
1	Furniture	Standard ...	112
1	Office Supplies	First Class	25
1	Office Supplies	Same Day	3
1	Office Supplies	Second Cl...	38
1	Office Supplies	Standard ...	112
1	Technology	First Class	25
1	Technology	Same Day	3
1	Technology	Second Cl...	38
1	Technology	Standard ...	112
10	Furniture	First Class	73
10	Furniture	Same Day	33
10	Furniture	Second Cl...	75
10	Furniture	Standard ...	236
10	Office Supplies	First Class	73
10	Office Supplies	Same Day	33
10	Office Supplies	Second Cl...	75
10	Office Supplies	Standard ...	236
10	Technology	First Class	73

The screenshot shows a Microsoft Excel spreadsheet titled "Power Pivot for Excel - Book2". The ribbon tabs include File, Home, Design, Advanced, and Power Pivot. The Power Pivot tab is selected. A PivotTable is displayed in the main area, titled "Query1". The table has columns for "Dim DateMonth Of Year-Month Of Year", "Dim ProductCategory Name", "Dim OrderShip Mode", "Ship Mode", "MeasuresFact Order Fulfillment Count", and "Add Column". The data includes rows for Furniture, Office Supplies, Technology, and various shipping classes (First Class, Same Day, Second Class, Standard Class) across different months and years. The total count for Standard Class in Month 1 is 112.

- Nhìn vào biểu đồ, ta thấy được trong tháng 1, danh mục Furniture có số lượng đơn hàng nhiều nhất, cao nhất ở phương thức vận chuyển Standard Class.

#### 4.4.5.2. Sử dụng Pivot Table trong Excel

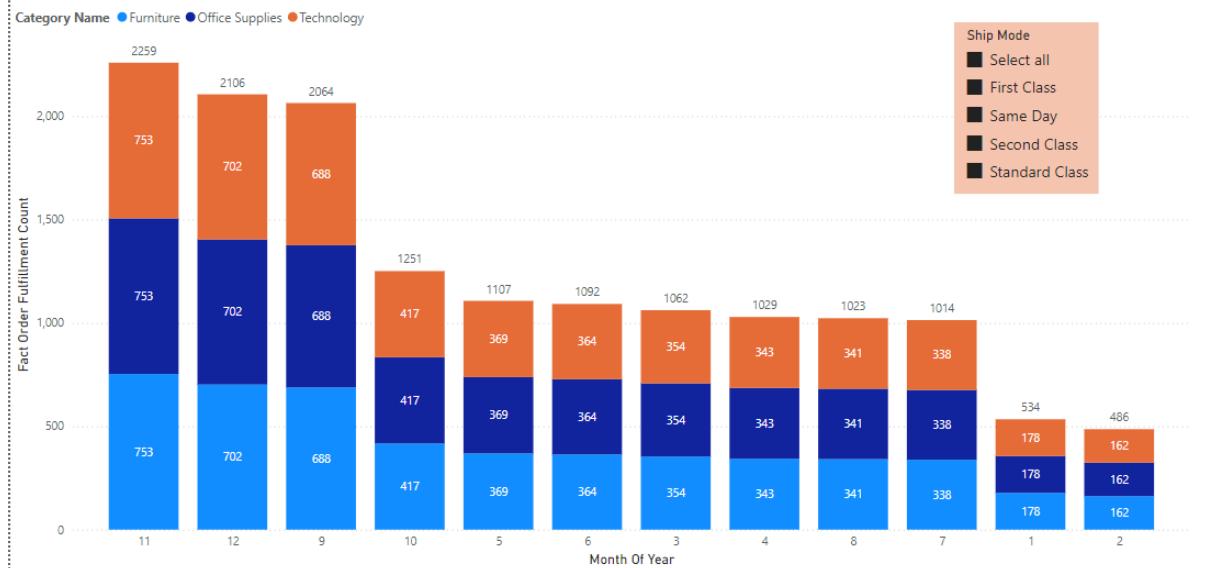
The screenshot shows a Microsoft Excel spreadsheet titled "Book2 - Excel". The ribbon tabs include File, Home, Insert, Page Layout, Formulas, Data, Review, View, Help, Power Pivot, PivotTable Analyze, and Design. The PivotTable Analyze tab is selected. A PivotTable is displayed in the main area, titled "Sheet1". The table has columns for "Fact Order Fulfillment Count", "Column Labels", and "Grand Total". The data is grouped by "Row Labels" such as "Furniture", "Office Supplies", and "Technology", and further broken down by "Ship Mode" like "First Class", "Same Day", "Second Class", and "Standard Class". The total count for Standard Class across all categories and months is 2994. The PivotTable Fields pane on the right shows fields for "Dim Product", "Dim Ship Date", and "Ship Mode".

- Nhìn vào biểu đồ, ta thấy được tháng 11 là tháng có số lượng đơn hàng nhiều nhất với tổng đơn hàng là 753 đơn trên tổng 5009 đơn. Các tháng còn lại có số lượng đơn hàng ở mức cao là tháng 9, 10, 12. Tháng có số lượng đơn hàng thấp nhất là tháng 1.

#### 4.4.5.3. Sử dụng Power BI

Biểu đồ thể hiện số lượng đơn hàng của từng danh mục sản phẩm và phương thức vận chuyển theo tháng.

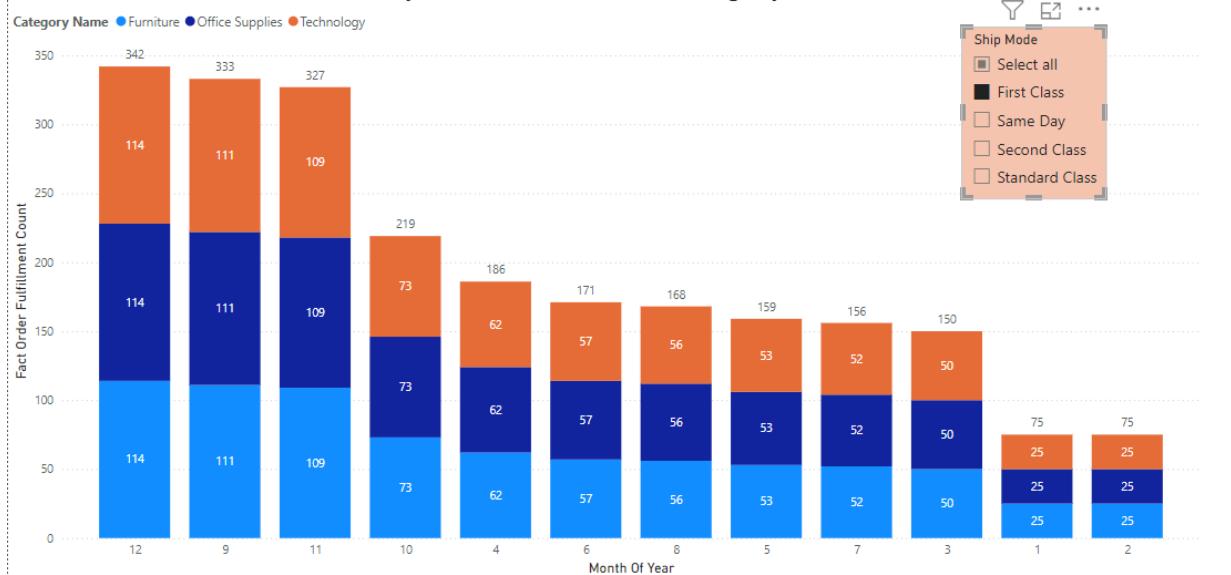
### Fact Order Fulfillment Count by Month Of Year and Category Name



Giải thích: Số liệu được phân tích trên cả bốn phương thức vận chuyển cho thấy rằng các tháng 11, 12, và 9 là những tháng có số lượng đơn hàng cao nhất, với tổng số đơn hàng vượt quá 2000 đơn mỗi tháng. Chính vì vậy, để đáp ứng nhu cầu tăng cao trong các tháng này, chúng ta cần bổ sung thêm nhân lực cho các công đoạn giao hàng và đóng gói. Điều này không chỉ nâng cao chất lượng dịch vụ mà còn giúp duy trì sự hài lòng của khách hàng trong những thời điểm bận rộn nhất trong năm.

- Ngoài ra, có thể xem số lượng các đơn hàng trong tháng theo từng phương thức vận chuyển ở biểu đồ sau:

### Fact Order Fulfillment Count by Month Of Year and Category Name



# **CHƯƠNG 5: KẾT LUẬN**

## **5.1. Kết quả đạt được**

Thực hiện xây dựng ‘XÂY DỰNG KHO DỮ LIỆU CHO SIÊU THỊ’ giúp chúng em áp dụng thực tế được các kiến thức đã học trong quá trình học môn Kho dữ liệu, nắm vững kiến thức lý thuyết cơ bản phục vụ tốt cho việc thiết kế chương trình. Thiết kế được một kho dữ liệu tương đối hoàn chỉnh, hiểu và sử dụng được thêm được một số Cung cấp lại các kiến thức đã được học, đặc biệt là kỹ năng phân tích, giải quyết vấn đề. Nắm rõ các khái niệm cơ bản về kho dữ liệu và OLAP, các tính chất của một kho dữ liệu cần có.

Thành thạo trong việc sử dụng các công cụ hỗ trợ quá trình ETL đưa dữ liệu vào kho dữ liệu như công cụ SSIS, SSAS, các công cụ để trực quan hóa dữ liệu như Excel Pivot, Power BI Desktop.

## **5.2. Những hạn chế**

Do thời gian hạn ngắn cộng với khối lượng công việc nhiều nên trong quá trình thực hiện đồ án nhóm còn gặp phải một số vấn đề :

- + Kiến thức thực tế và kiến thức tư duy phân tích của chúng em còn hạn chế, nên tính chuyên nghiệp của đề tài chưa cao
- + Khả năng xử lý nghiệp vụ chưa linh hoạt nhạy bén.

## **5.3. Tài liệu tham khảo**

[1] Tài liệu các file PDF hướng dẫn về Data Integration với SQL Server, SSIS, SSAS của GVHD Ths. Nguyễn Văn Thành.