

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 Phương 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

# 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	Customer	Product	Order	OrderDate	ShipDate
Business Process Name	Fact Table					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 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 of 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,...

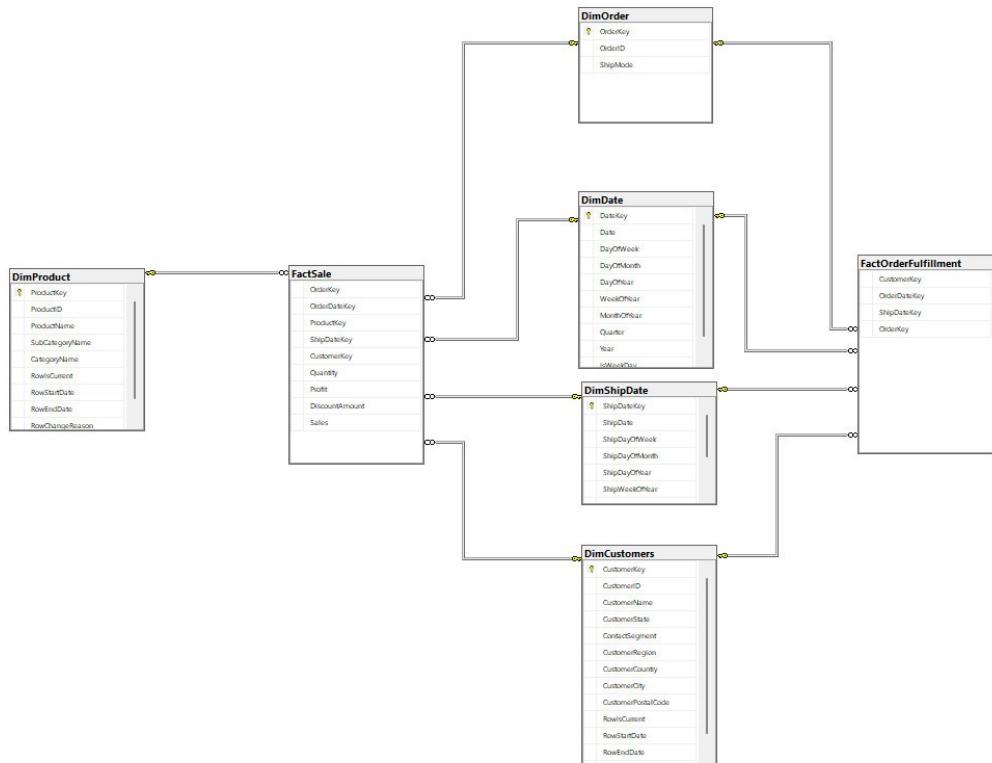
### 2.2.4. DimShipDate

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)

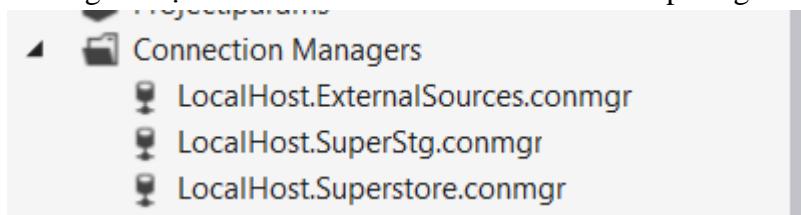


# CHƯƠNG 3: TÍCH HỢP DỮ LIỆU VÀO KHO (SSIS)

Tạo 3 kết nối tới 3 database

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

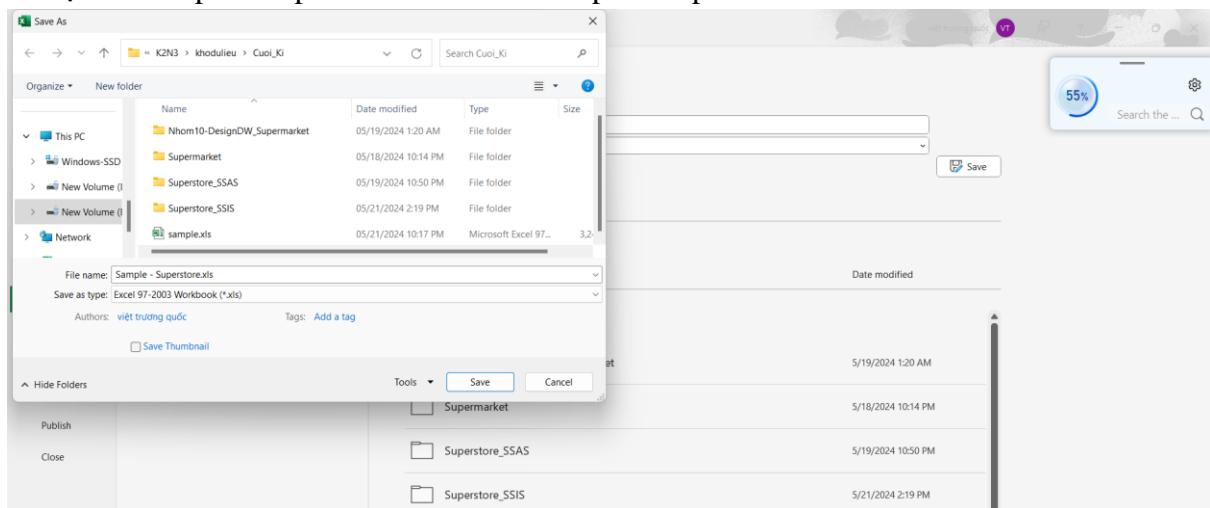
-Luồng dữ liệu đi từ excel ->ExternalSources -> SuperStg-> Superstore



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

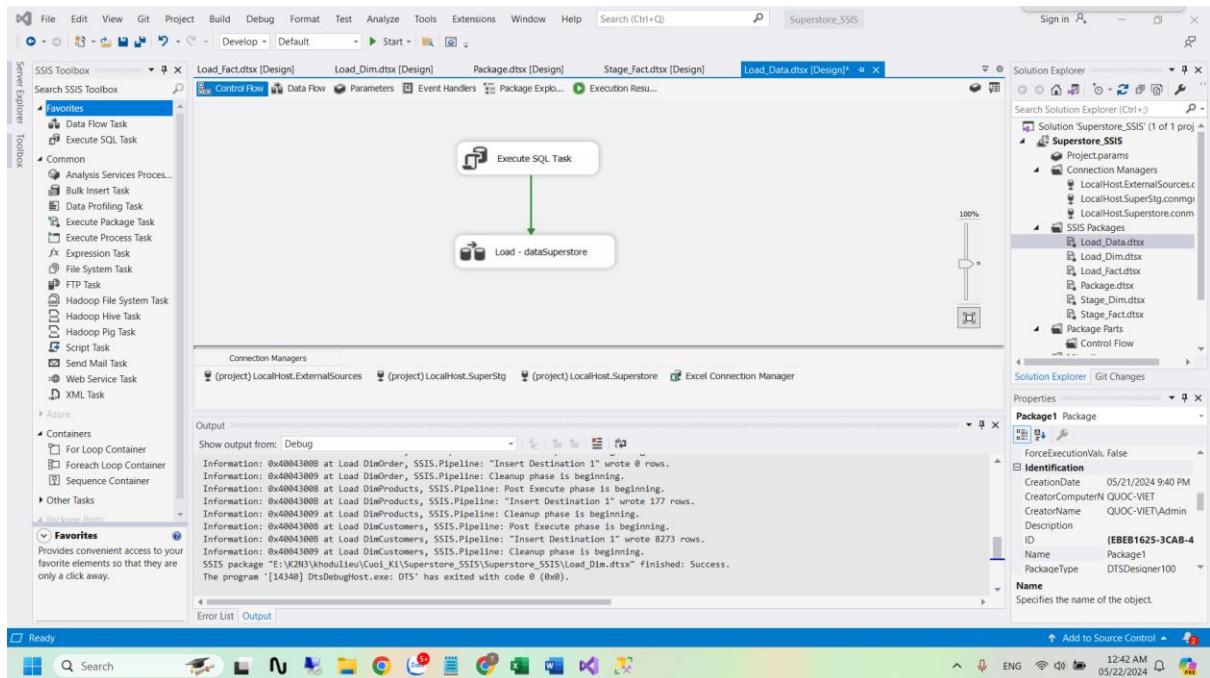
Chuyển dữ liệu gốc từ file có dạng .csv thành .xls or xlsx

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

Execute SQL Task Editor

Configure the properties required to run SQL statements and stored procedures using the selected connection.

General

Parameter Mapping  
Result Set  
Expressions

Name	Execute SQL Task
Description	Execute SQL Task
Options	
TimeOut	0
CodePage	1252
TypeConversionMode	Allowed
Result Set	None
SQL Statement	
ConnectionType	OLE DB
Connection	localhost.ExternalSources
SQLSourceType	Direct input
SQLStatement	Truncate table SampleSuperstore
IsQueryStoredProcedure	False
BypassPrepare	True

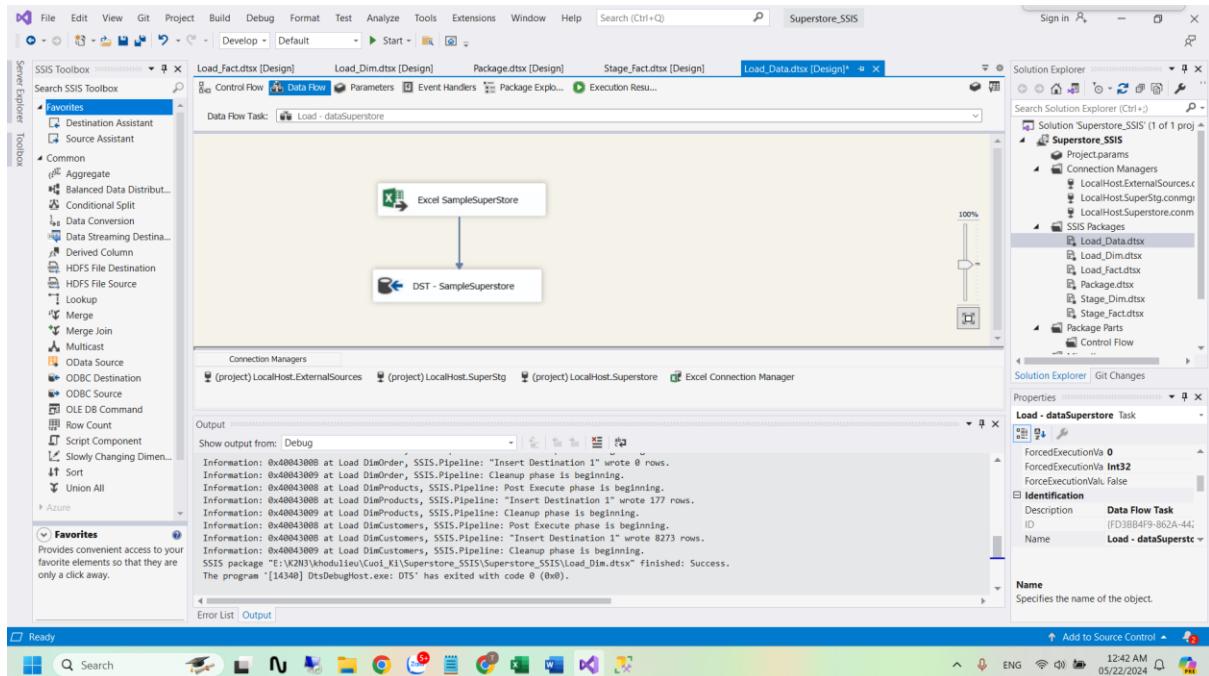
Name

Specifies the name of the task.

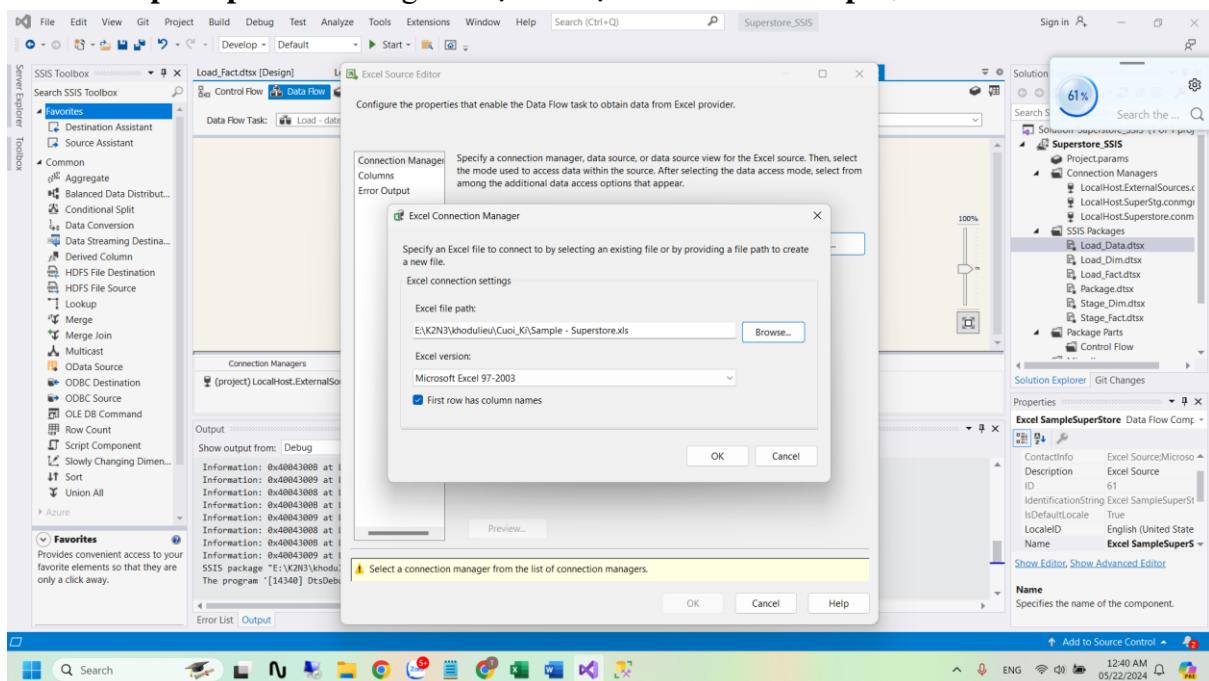
Browse... Build Query... Parse Query

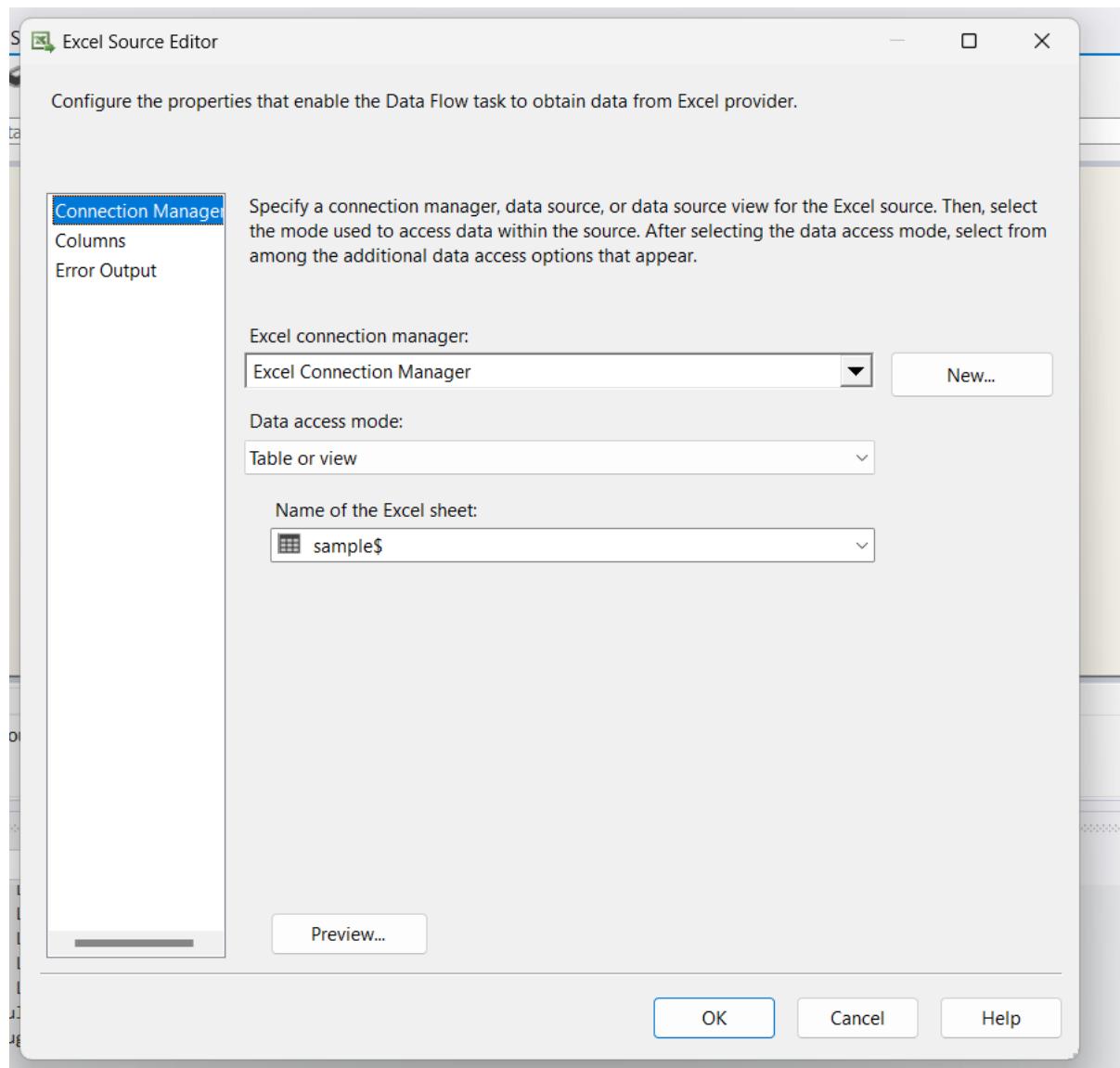
OK Cancel Help

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\$

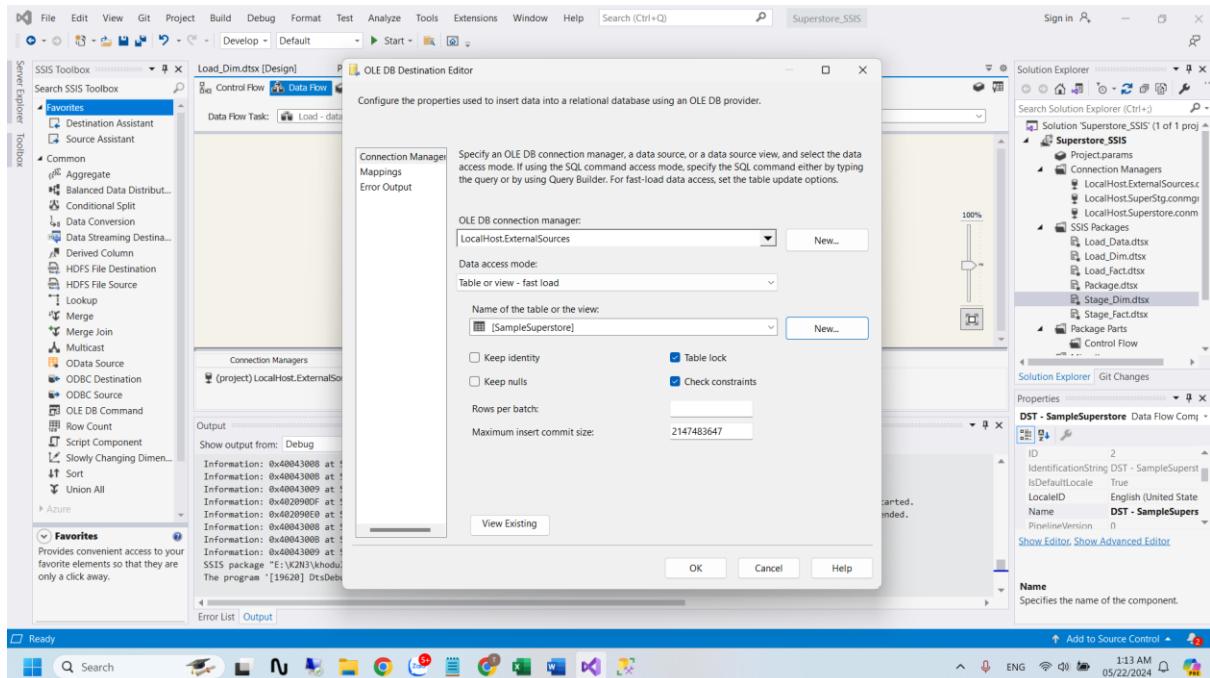


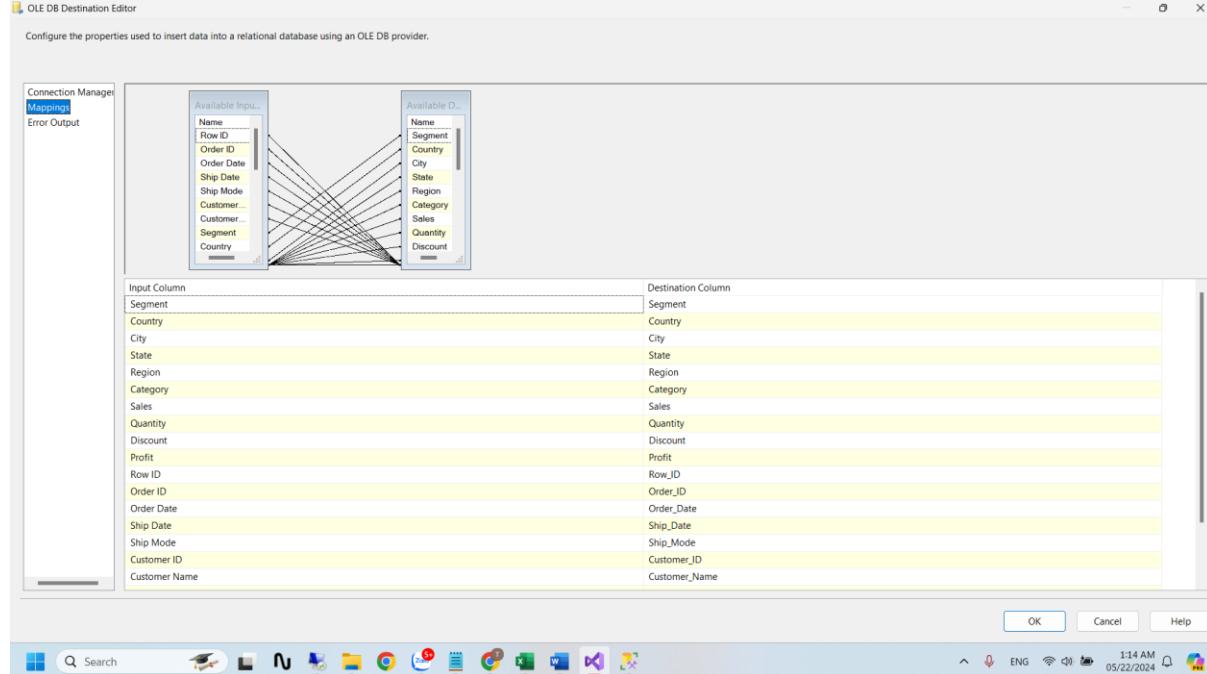
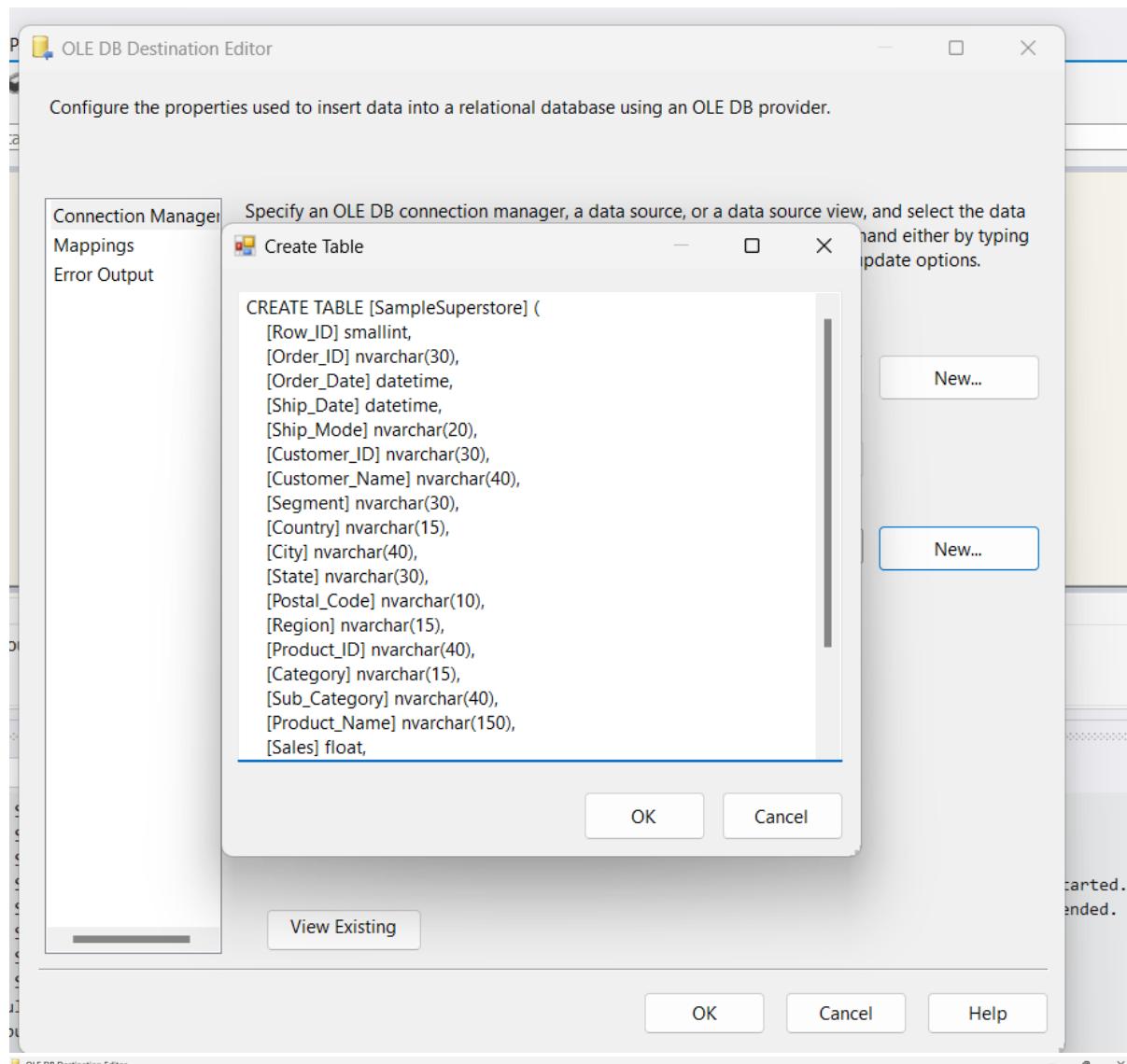


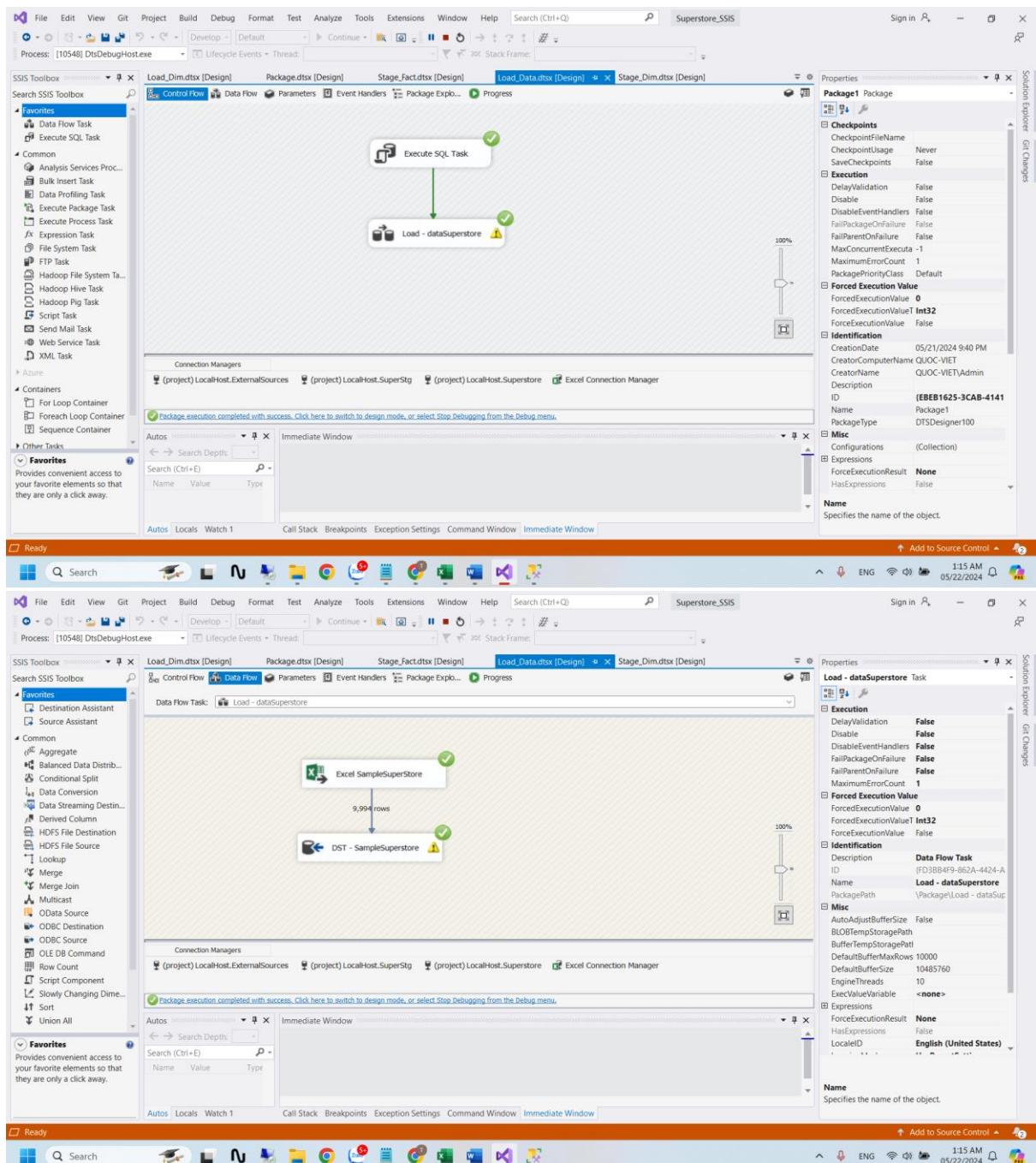
Query result (up to the first 200 rows):

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer...	Customer...	Segment	Country	City	State	Postal C...	Region	Product...	Category	Sub-Cat...	Product...	Sales	Quantity	Discount	Profit
1	CA-201...	11/08/...	11/11/...	Second...	CG-125...	Claire G...	Consume...	United ...	Hender...	Kentucky	42420	South	FUR-BO...	Furniture	Bookcas...	Bush S...	261.96	2	0	41,9136
2	CA-201...	11/08/...	11/11/...	Second...	CG-125...	Claire G...	Consume...	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...	Darin R...	Corpor...	United ...	Los An...	California	90036	West	OFF-IA...	Office S...	Labels	Self-Ad...	14,62	2	0	6,8714
4	US-201...	10/11/...	10/18/...	Standard...	SO-203...	Sean O...	Consume...	United ...	Fort La...	Florida	33311	South	FUR-TA...	Furniture	Tables	Bretfor...	957,5775	5	0,45	-383,031
5	US-201...	10/11/...	10/18/...	Standard...	SO-203...	Sean O...	Consume...	United ...	Fort La...	Florida	33311	South	OFF-ST...	Office S...	Storage	Elton F...	22,368	2	0,2	2,5164
6	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consume...	United ...	Los An...	California	90032	West	FUR-FU...	Furniture	Furnish...	Elton E...	48,86	7	0	14,1694
7	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consume...	United ...	Los An...	California	90032	West	OFF-AR...	Office S...	Art	Newell ...	7,28	4	0	1,9656
8	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consume...	United ...	Los An...	California	90032	West	TEC-PH...	Technol...	Phones	Mtel 53...	907,152	6	0,2	90,7152
9	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consume...	United ...	Los An...	California	90032	West	OFF-BE...	Office S...	Binders	DXL An...	18,504	3	0,2	5,7825
10	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consume...	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 ...	Consume...	United ...	Los An...	California	90032	West	FUR-TA...	Furniture	Tables	Chrom...	1706,184	9	0,2	85,2092
12	CA-201...	06/09/...	06/14/...	Standard...	BH-117...	Brosna ...	Consume...	United ...	Los An...	California	90032	West	TEC-PH...	Technol...	Phones	Konfidi ...	911,424	4	0,2	68,3568
13	CA-201...	04/15/...	04/20/...	Standard...	AA-104...	Andrew...	Consume...	United ...	Concord	North C...	28027	South	OFF-PA...	Office S...	Paper	Xerox 1...	15,552	3	0,2	5,4432
14	CA-201...	12/05/...	12/10/...	Standard...	IM-15070...	Irene M...	Consume...	United ...	Seattle	Washin...	98103	West	OFF-BE...	Office S...	Binders	Follow...	407,976	3	0,2	132,5922
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
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	0,8	-3,816
17	CA-201...	11/11/...	11/18/...	Standard...	PK-190...	Pete Kro...	Consume...	United ...	Madison	Wiscon...	53711	Central	OFF-ST...	Office S...	Storage	Stor...D...	665,88	6	0	13,3176
18	CA-201...	05/13/...	05/19/...	Second...	AG-102...	Alejand...	Consume...	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...	Zuchu ...	Consume...	United ...	San Fra...	California	94109	West	OFF-AR...	Office S...	Art	Newell ...	8,56	2	0	2,4824
20	CA-201...	08/27/...	09/01/...	Second...	ZD-219...	Zuchu ...	Consume...	United ...	San Fra...	California	94109	West	TEC-PH...	Technol...	Phones	Cisco S...	213,48	3	0,2	16,0111
21	CA-201...	08/27/...	09/01/...	Second...	ZD-219...	Zuchu ...	Consume...	United ...	San Fra...	California	94109	West	OFF-BE...	Office S...	Binders	Wilson ...	22,72	4	0,2	7,384
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	5,0596
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	15,6884
24	US-201...	07/16/...	07/18/...	Second...	SF-20065...	Sandra ...	Consume...	United ...	Philadelphia	Pennsyl...	19140	East	FUR-CH...	Furniture	Chairs	Global ...	71,372	2	0,3	-1,0196
25	CA-201...	09/25/...	09/30/...	Standard...	EB-138...	Emily B...	Consume...	United ...	Orem	Utah	84057	West	FUR-TA...	Furniture	Tables	Bretfor...	104,63	3	0	24,2649
26	CA-201...	01/16/...	01/20/...	Second...	EH-139...	Eric Hof...	Consume...	United ...	Los An...	California	90049	West	OFF-BE...	Office S...	Binders	Wilson ...	11,648	2	0,2	4,2224
27	CA-201...	01/16/...	01/20/...	Second...	EH-139...	Eric Hof...	Consume...	United ...	Los An...	California	90049	West	TEC-AC...	Technol...	Access...	Imation...	90,57	3	0	11,7741
28	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consume...	United ...	Philadelphia	Pennsyl...	19140	East	FUR-BO...	Furniture	Bookcas...	Riversid...	308,343	7	0,5	-1665...
29	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consume...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-BE...	Office S...	Binders	Avery ...	9,618	2	0,7	-7,0532
30	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consume...	United ...	Philadelphia	Pennsyl...	19140	East	FUR-FU...	Furniture	Furnish...	Howard...	124,2	3	0,2	15,525
31	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consume...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-EN...	Office S...	Envelope...	Poly Str...	3,264	2	0,2	1,1016
32	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consume...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-AR...	Office S...	Art	BOSTO...	86,304	6	0,2	9,7092
33	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consume...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-BE...	Office S...	Binders	Acco Pr...	6,858	6	0,7	-5,715
34	US-201...	09/17/...	09/21/...	Standard...	TB-215...	Tracy B...	Consume...	United ...	Philadelphia	Pennsyl...	19140	East	OFF-AD...	Office S...	Art	Lumbar...	1,574	7	0,5	5,544

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







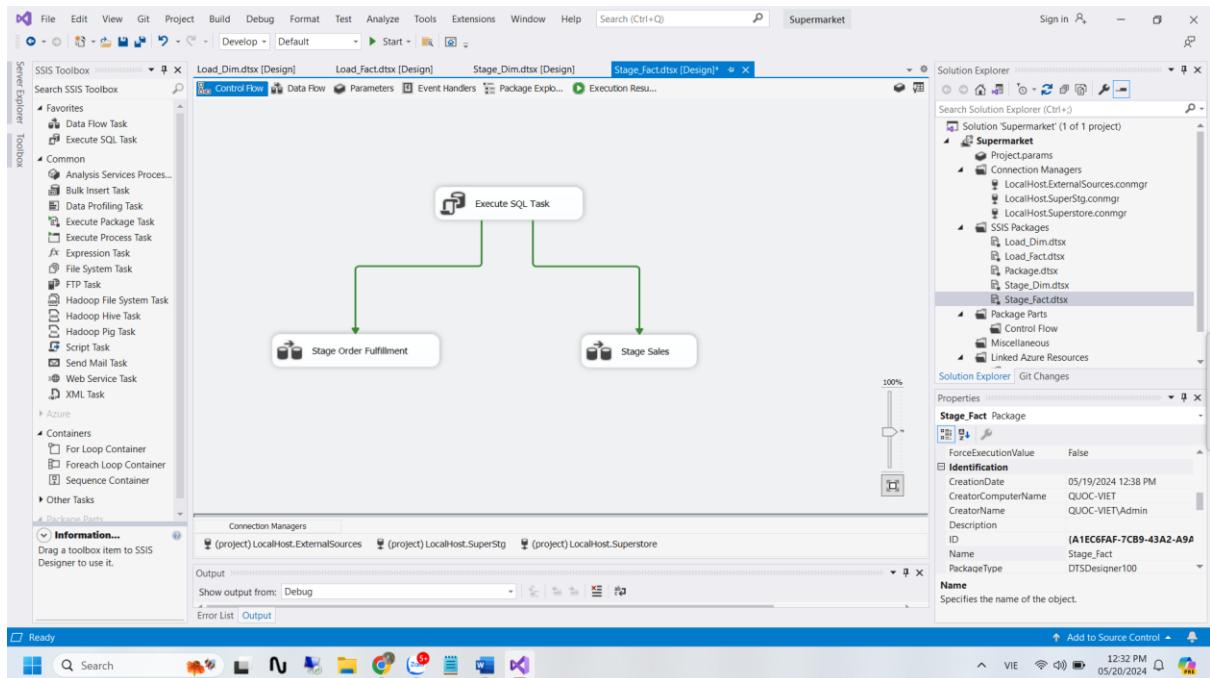
Kiểm tra database ExternalSources trong SQL xem đã có dữ liệu chưa

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

Ta phải Load từ SRC vào Stage

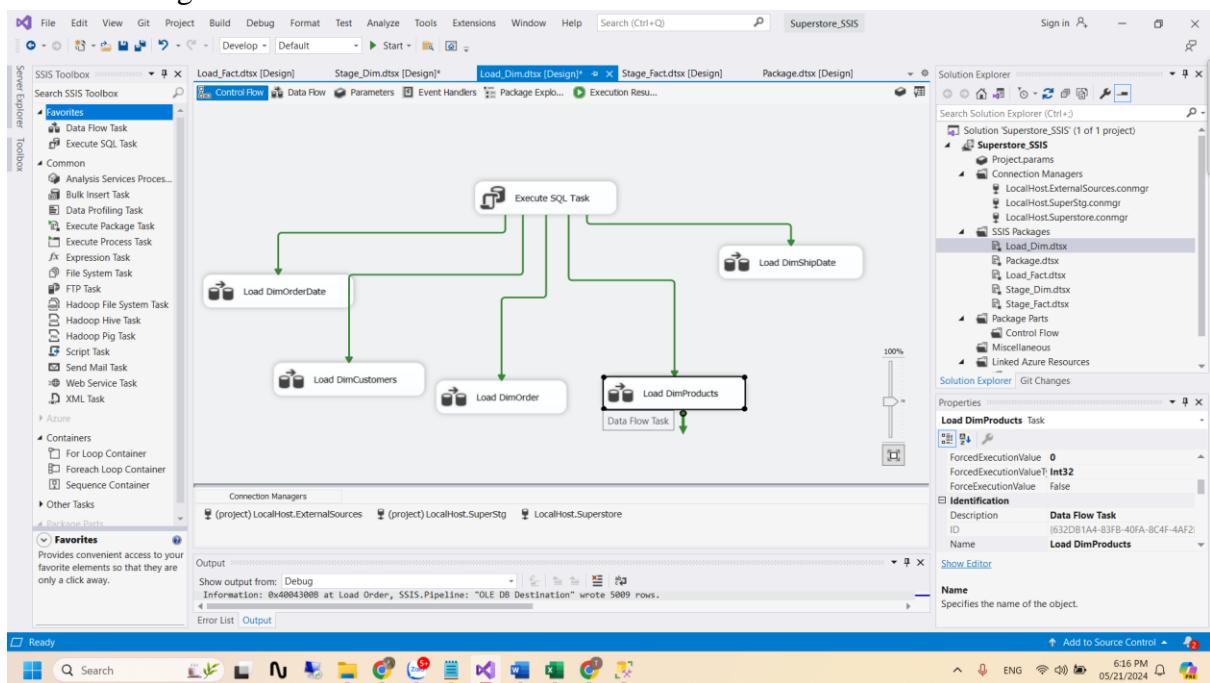
Stage các bảng Dim

Stage các bảng Fact

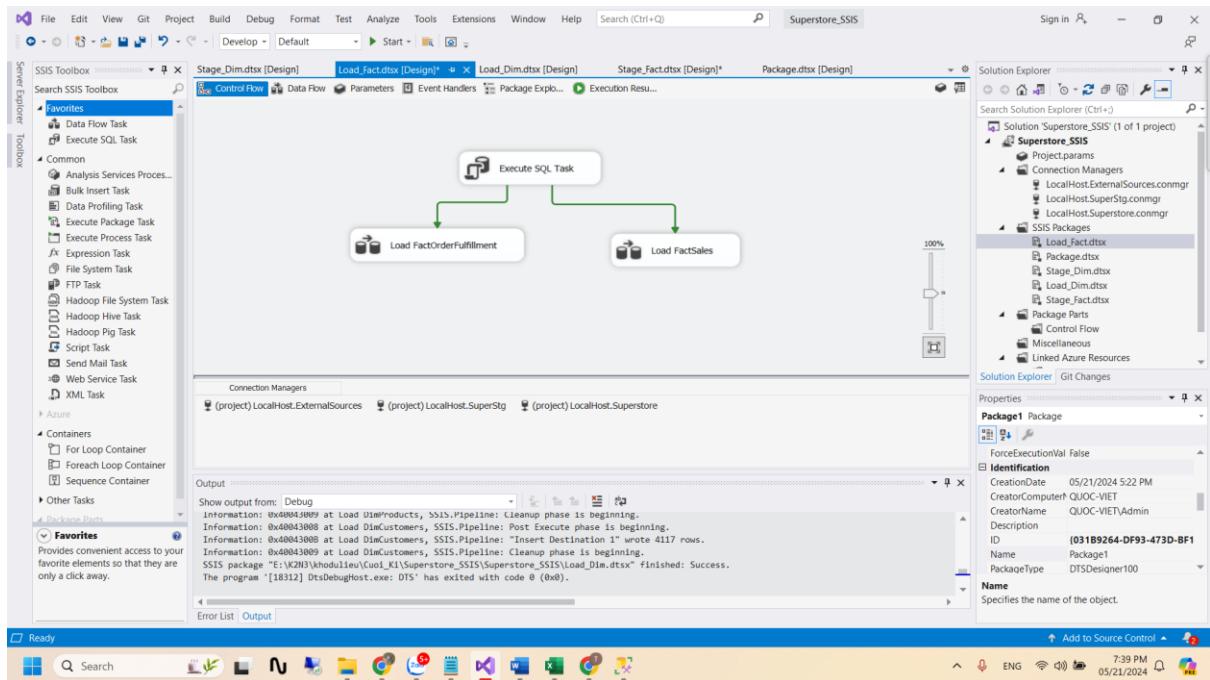


Sau đó Load từ Stage vào các bản Dim và bảng Fact

## Load các bảng Dim



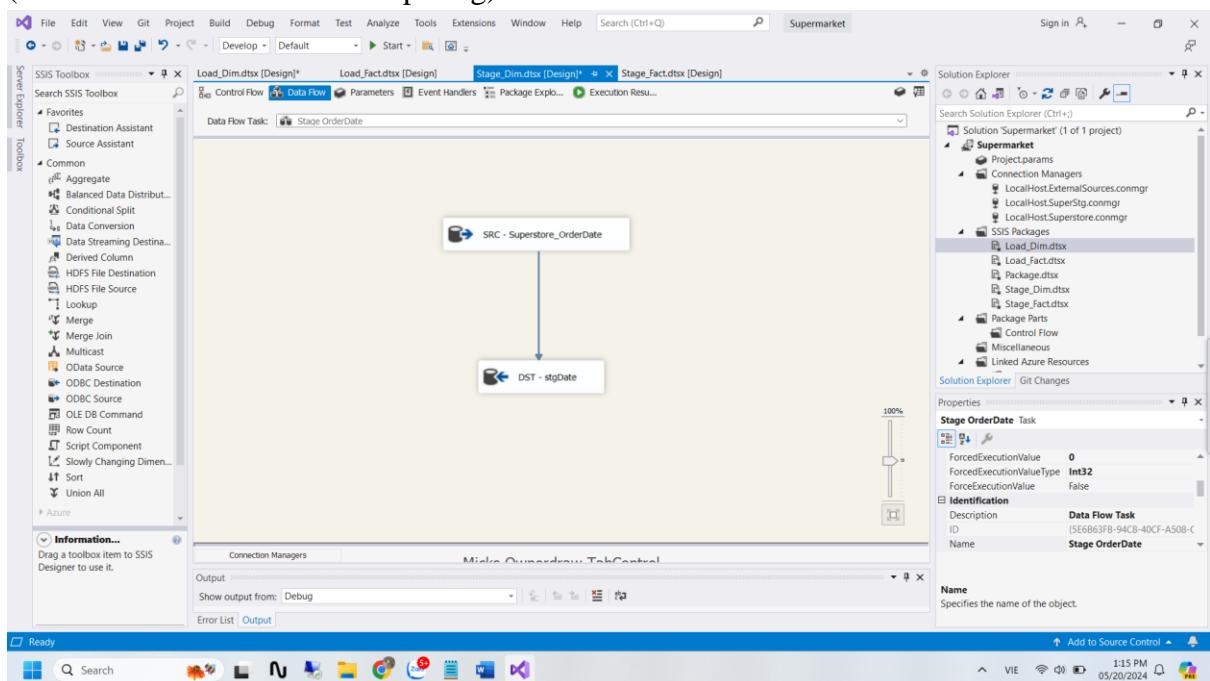
## Load các bảng Fact



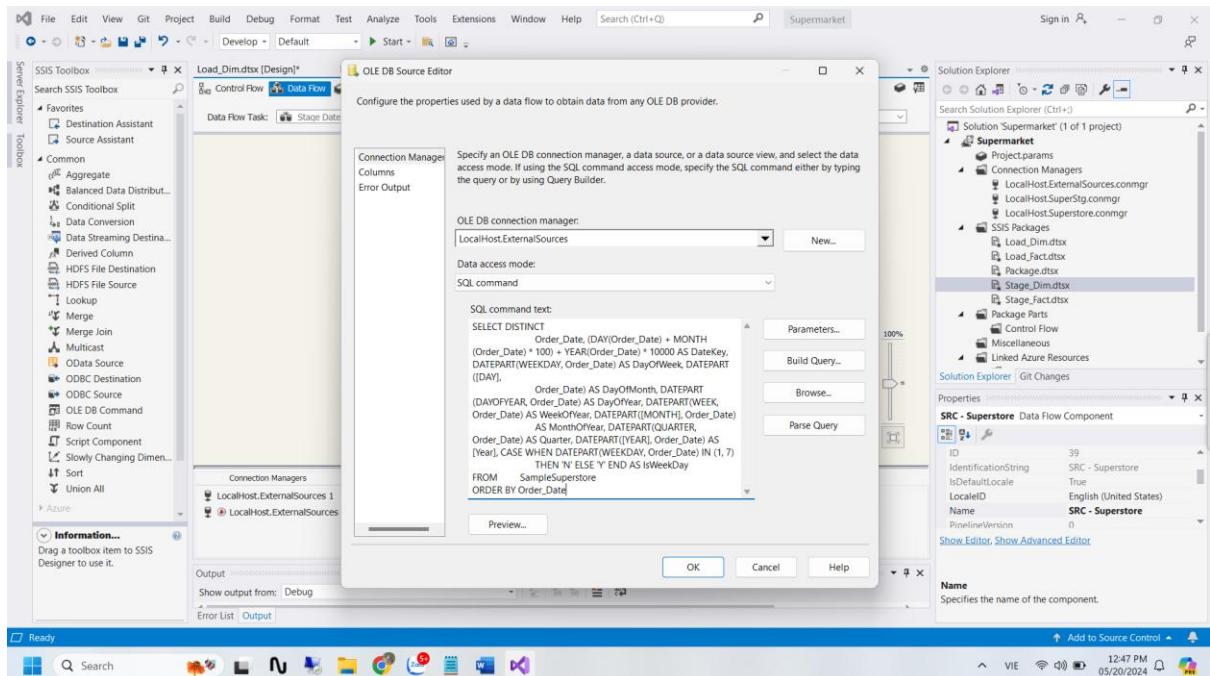
### 3.2.1. OrderDate Dimension

#### 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 :

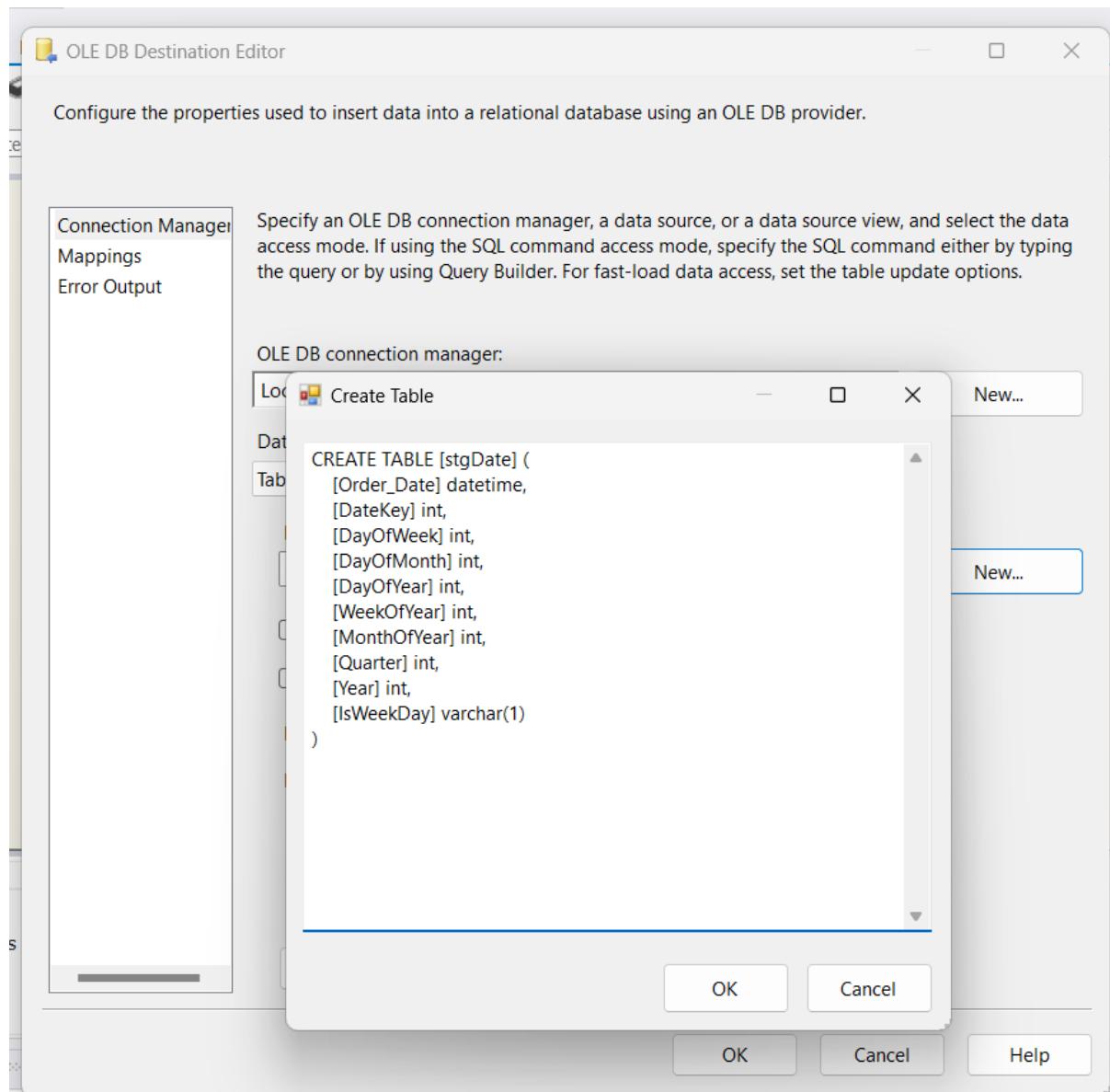
```

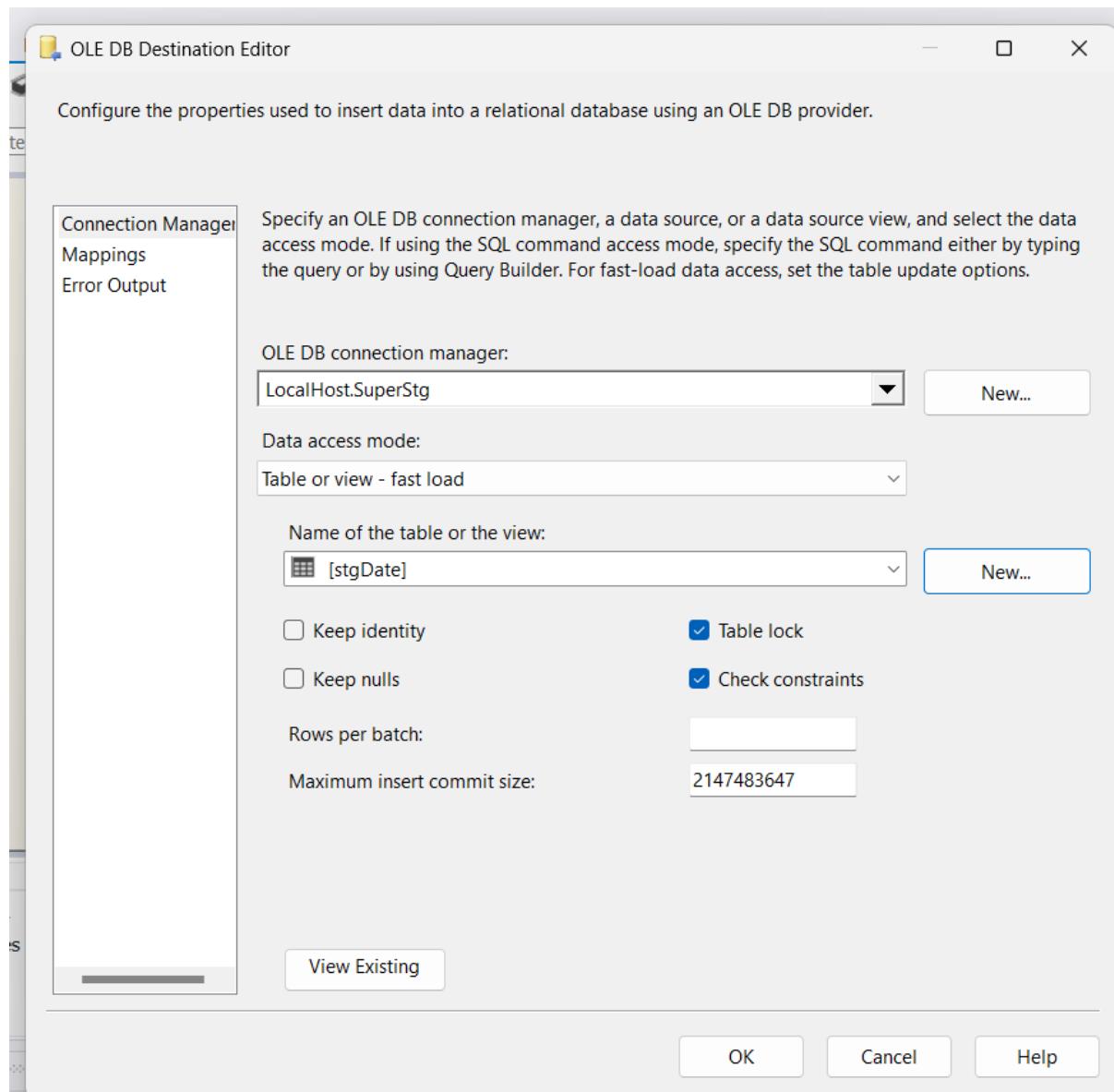
SELECT DISTINCT Order_Date,
(DAY(Order_Date) + MONTH(Order_Date) * 100) + YEAR(Order_Date) * 10000 AS DateKey,
DATEPART(WEEKDAY, Order_Date) AS DayOfWeek,
DATEPART([DAY], Order_Date) AS DayOfMonth,
DATEPART(DAYOFYEAR, Order_Date) AS DayOfYear,
DATEPART(WEEK, Order_Date) AS WeekOfYear,
DATEPART([MONTH], 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

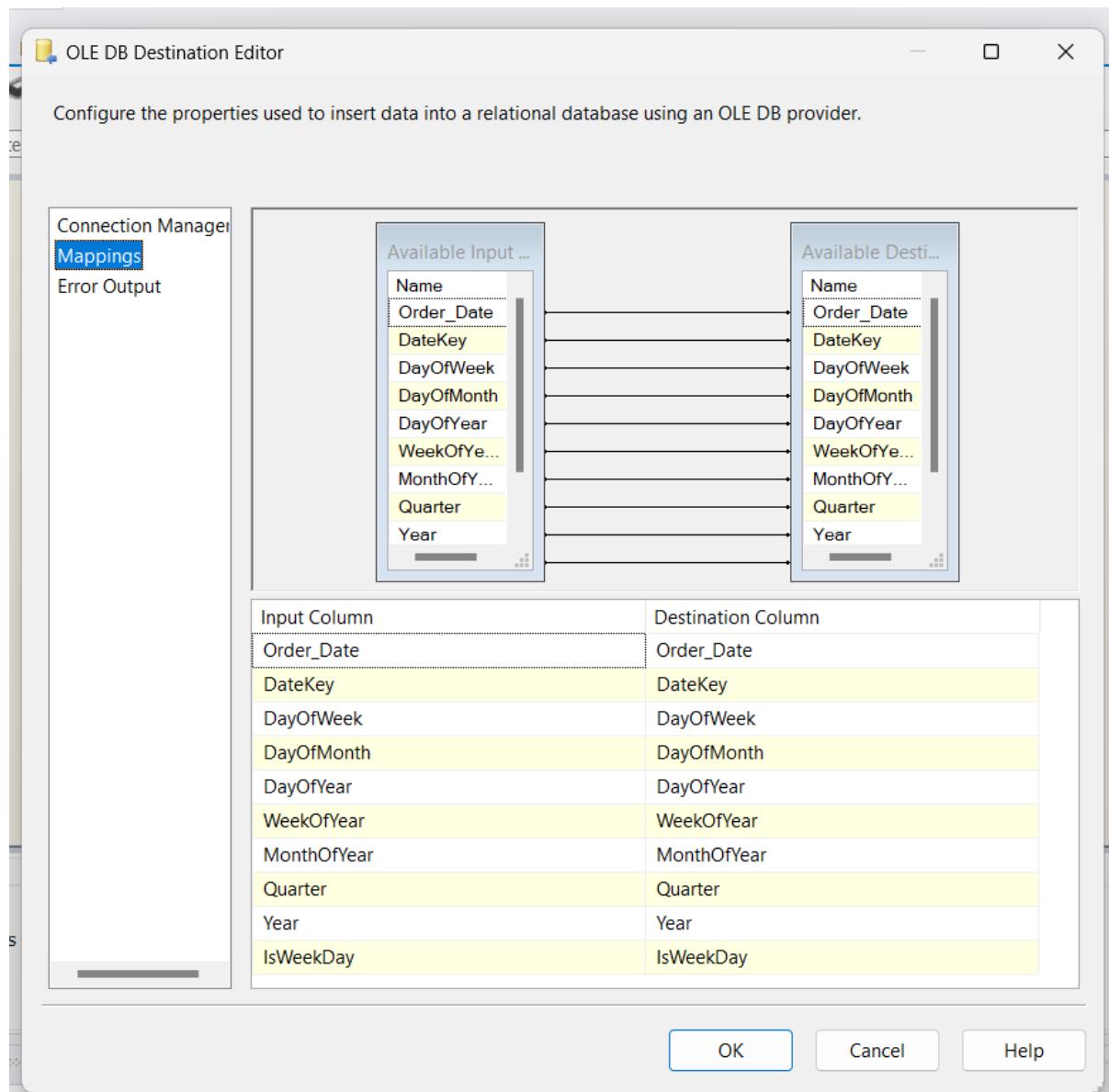
```

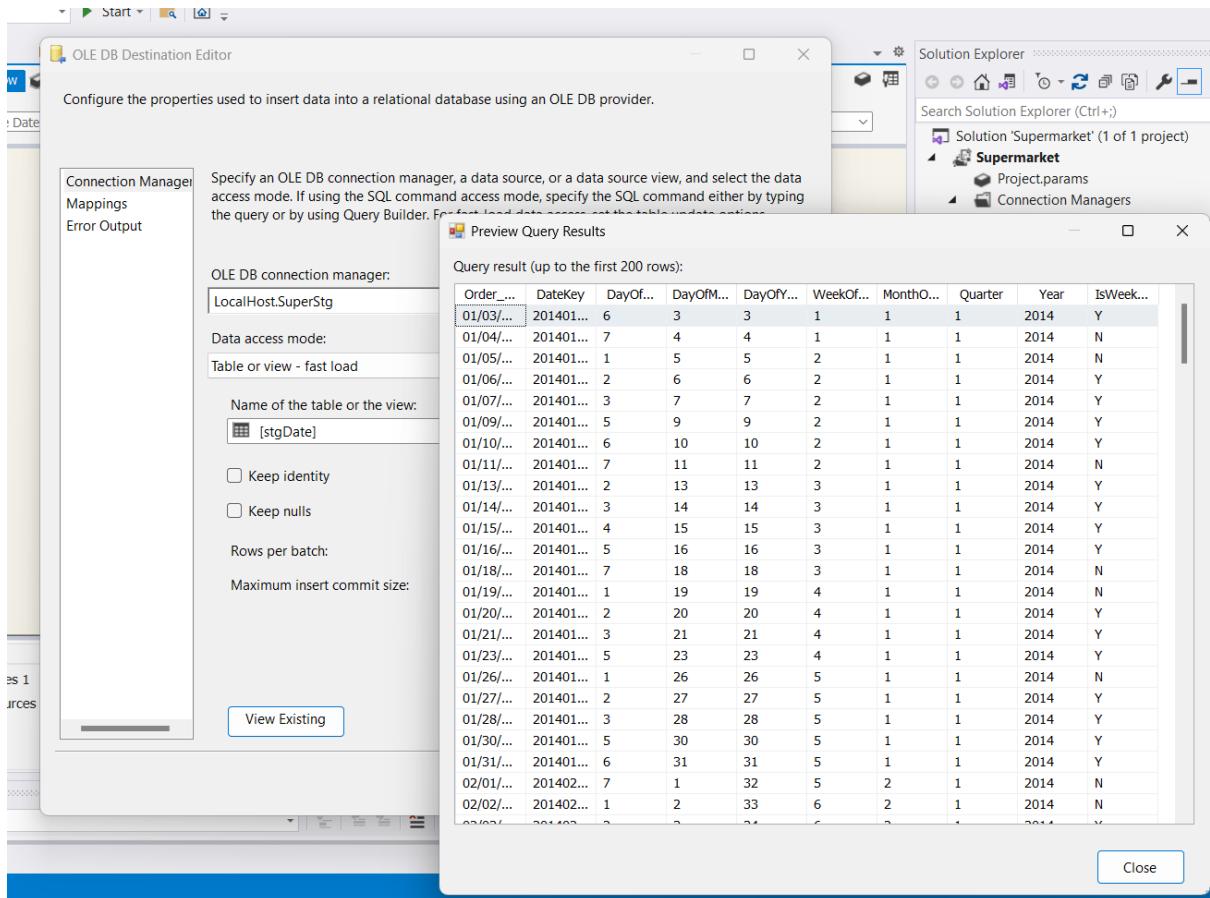
Ta sẽ lấy nguồn dữ liệu này để có thể tạo ra stgDate chứa thời gian thống kê kho và thời gian đặt hàng

## DST – stgDate:

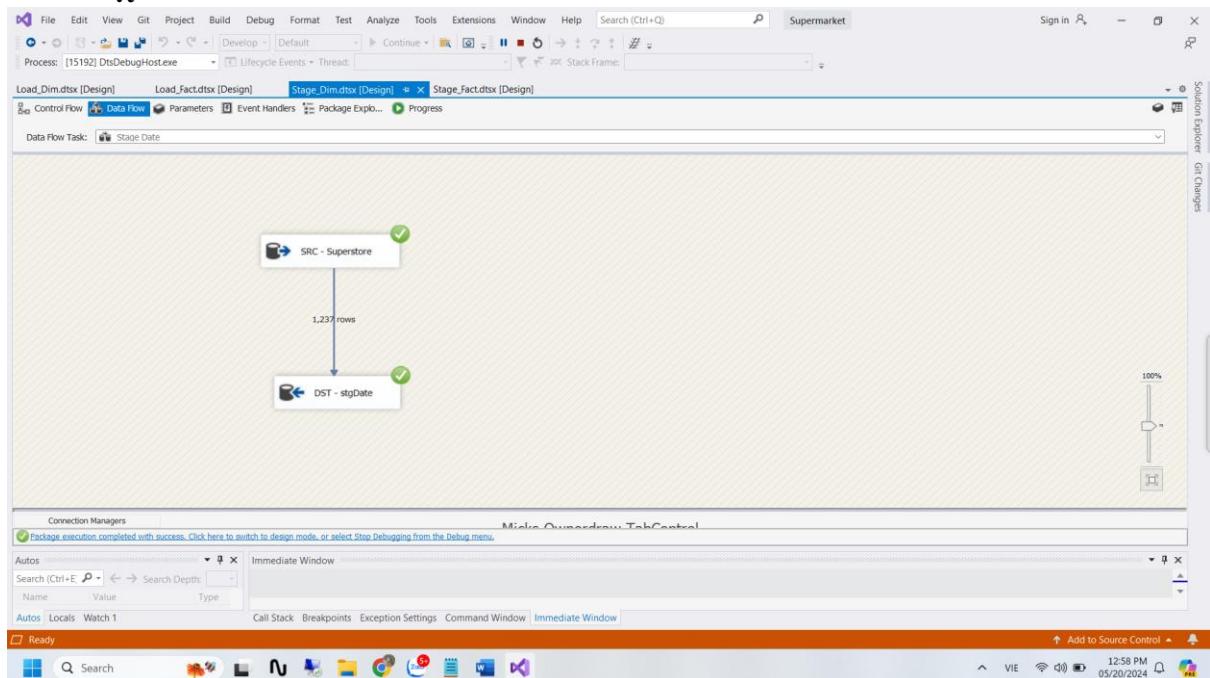






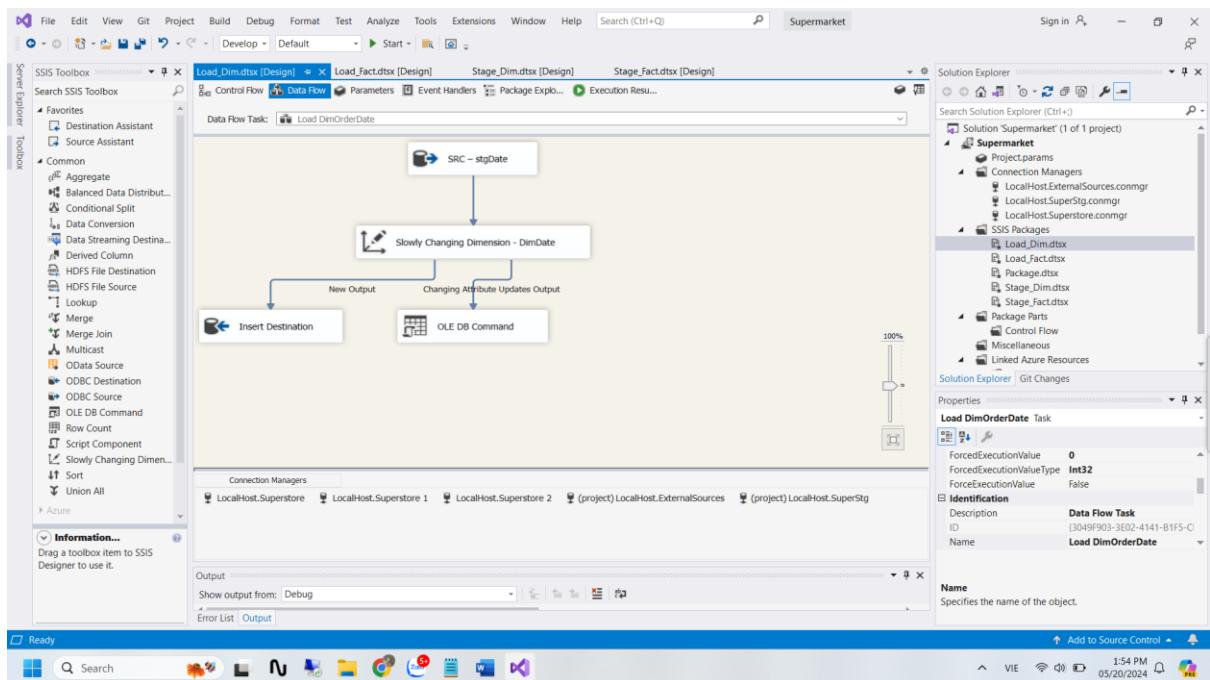


## Khởi chạy

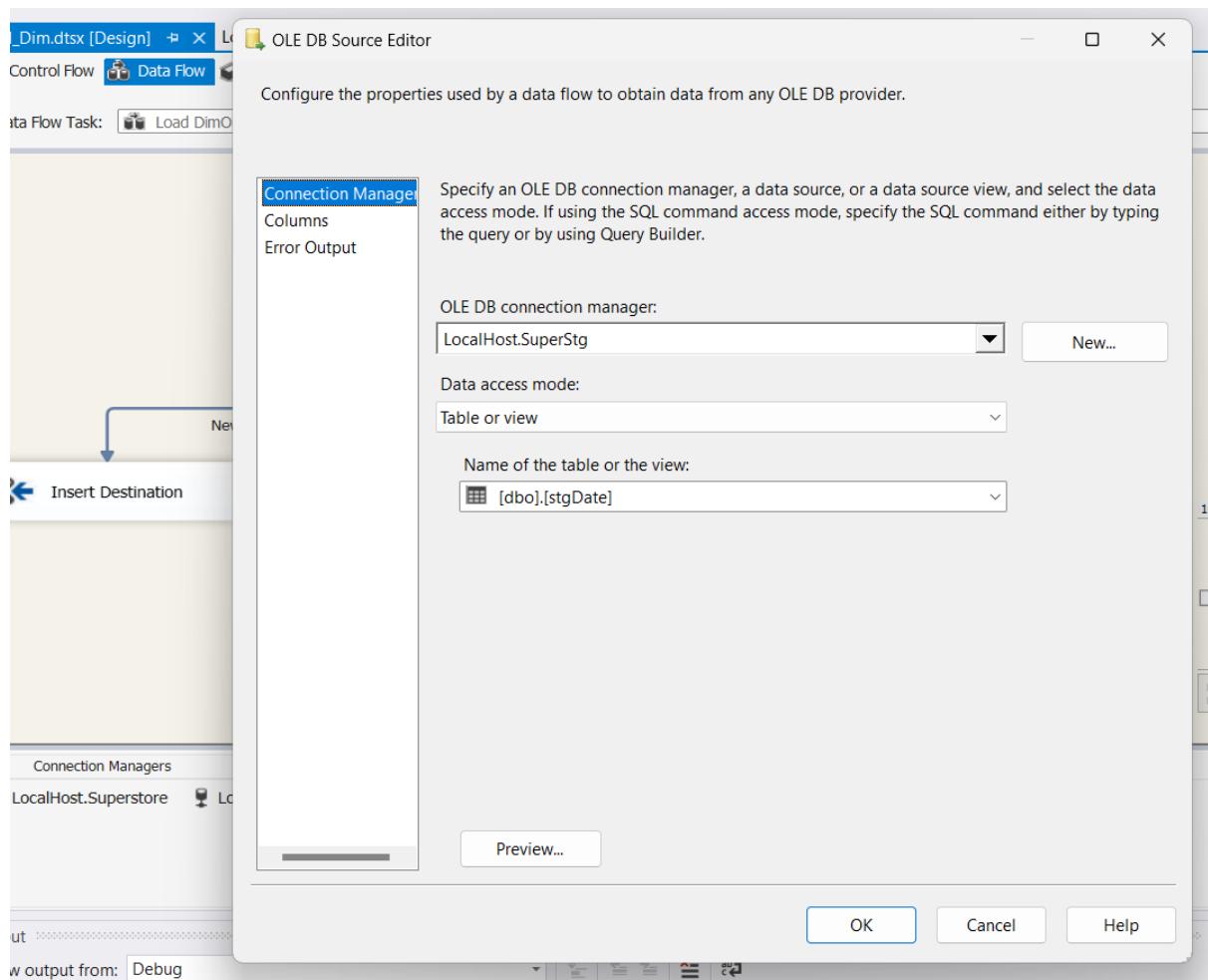


## b) Load dữ liệu từ stgDate vào DimDate

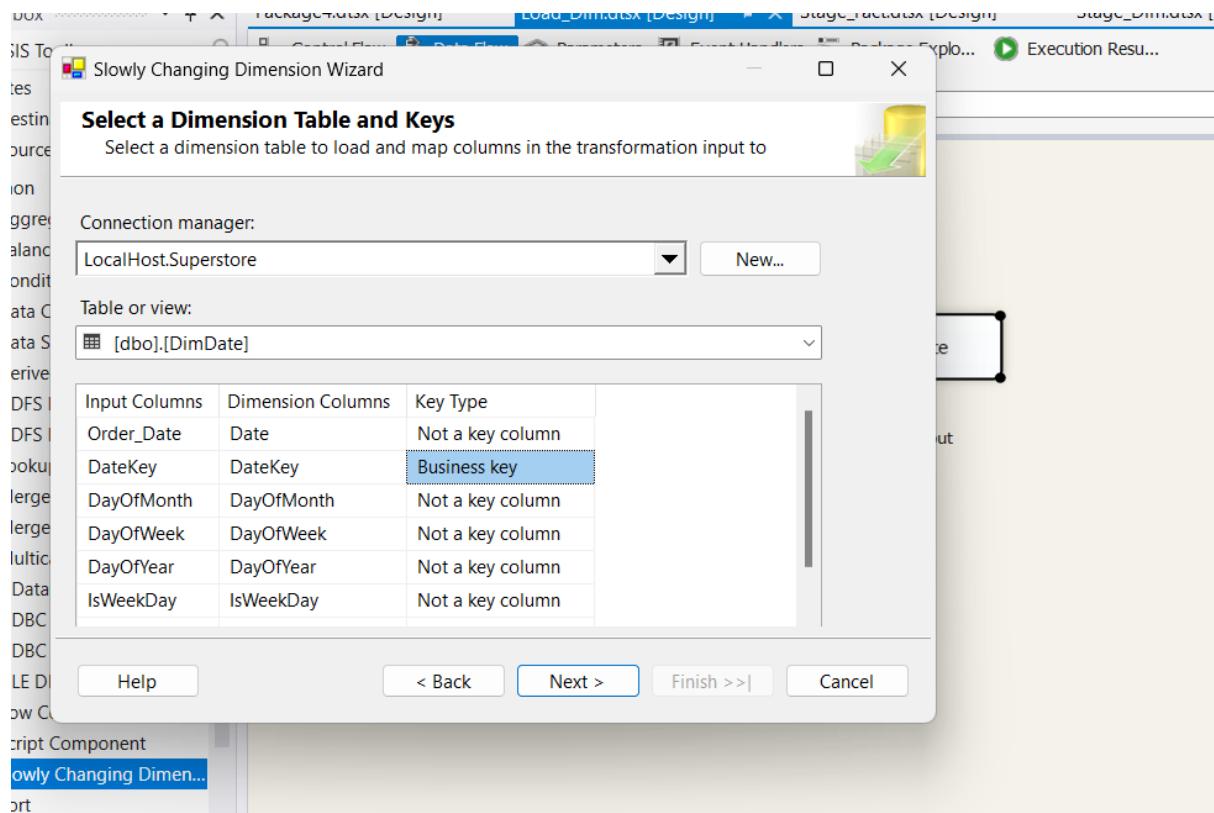
(load từ SuperStg-> Superstore)



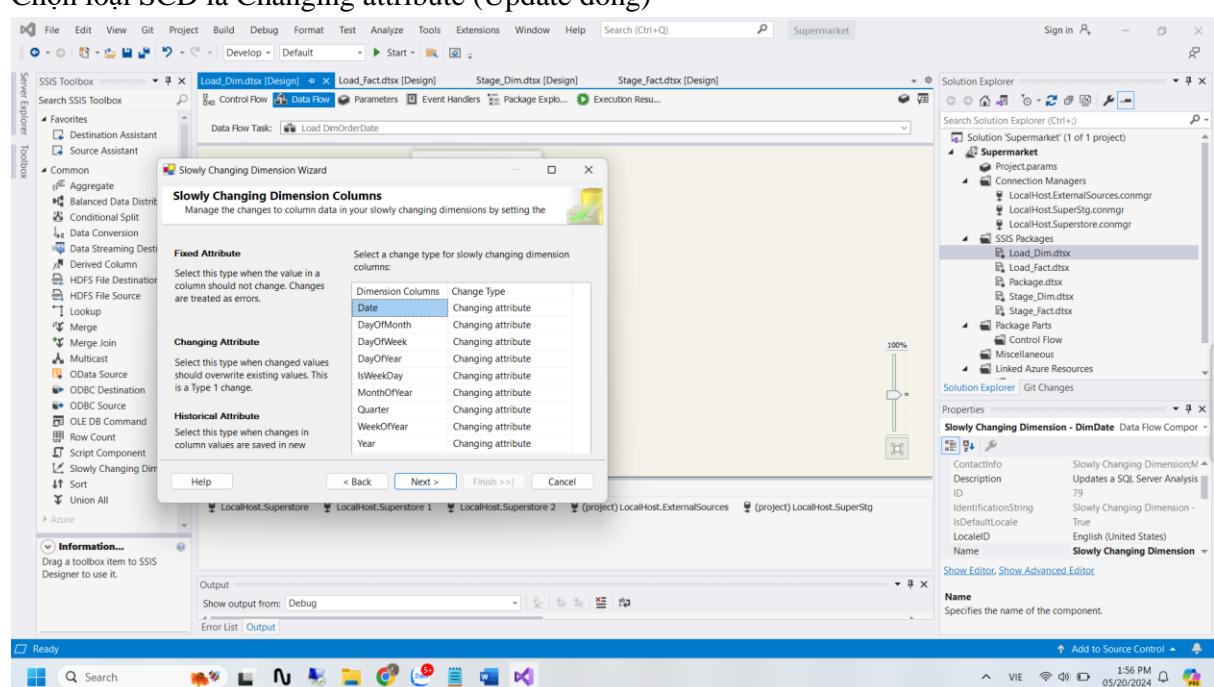
**SRC – stgDate** là dữ liệu từ bảng stgDate đã load ở trên



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

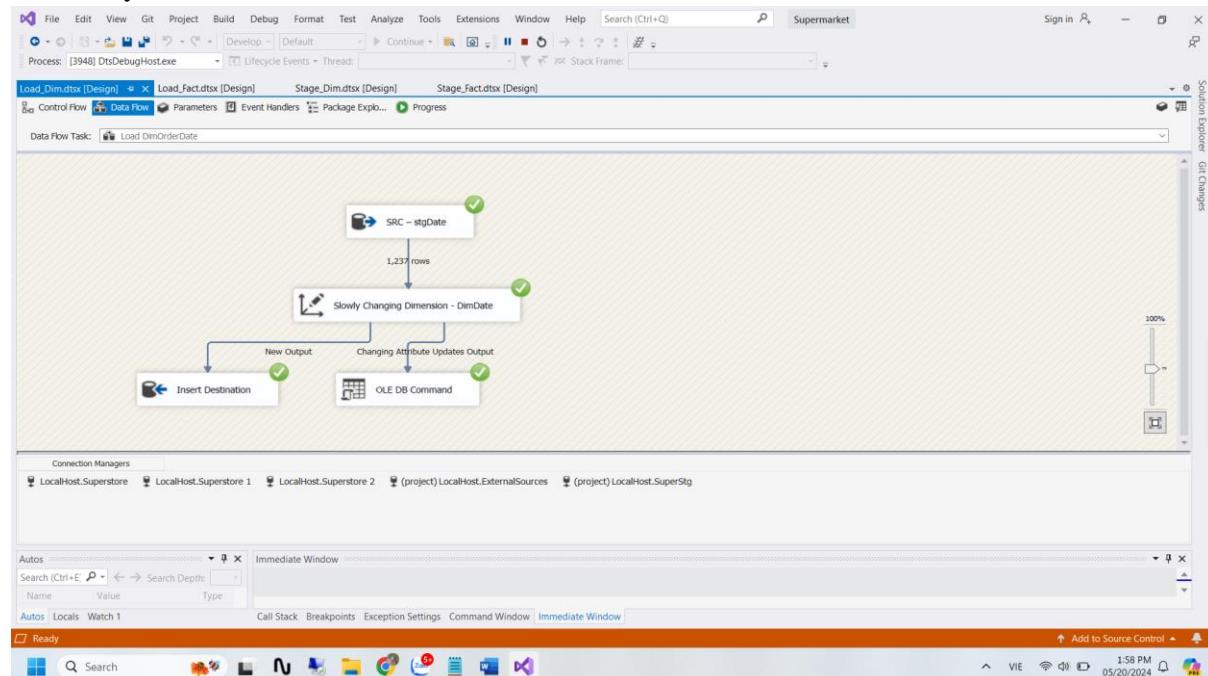


## Chọn loại SCD là Changing attribute (Update dòng)



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

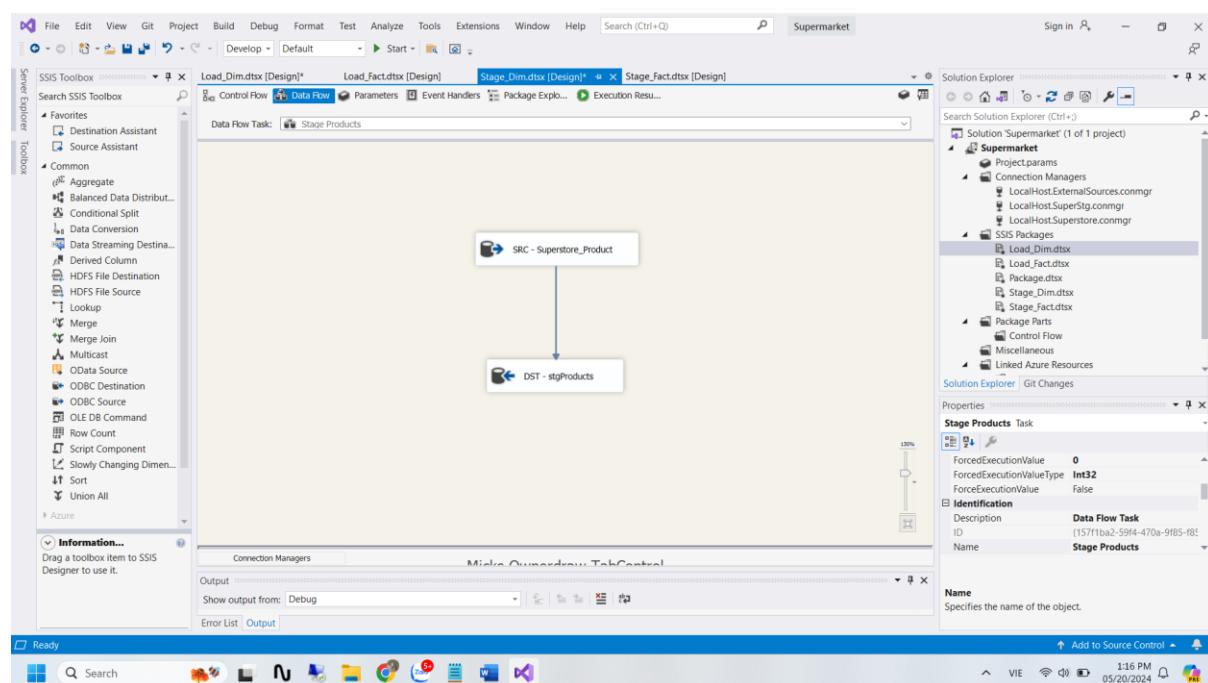
Khởi chạy



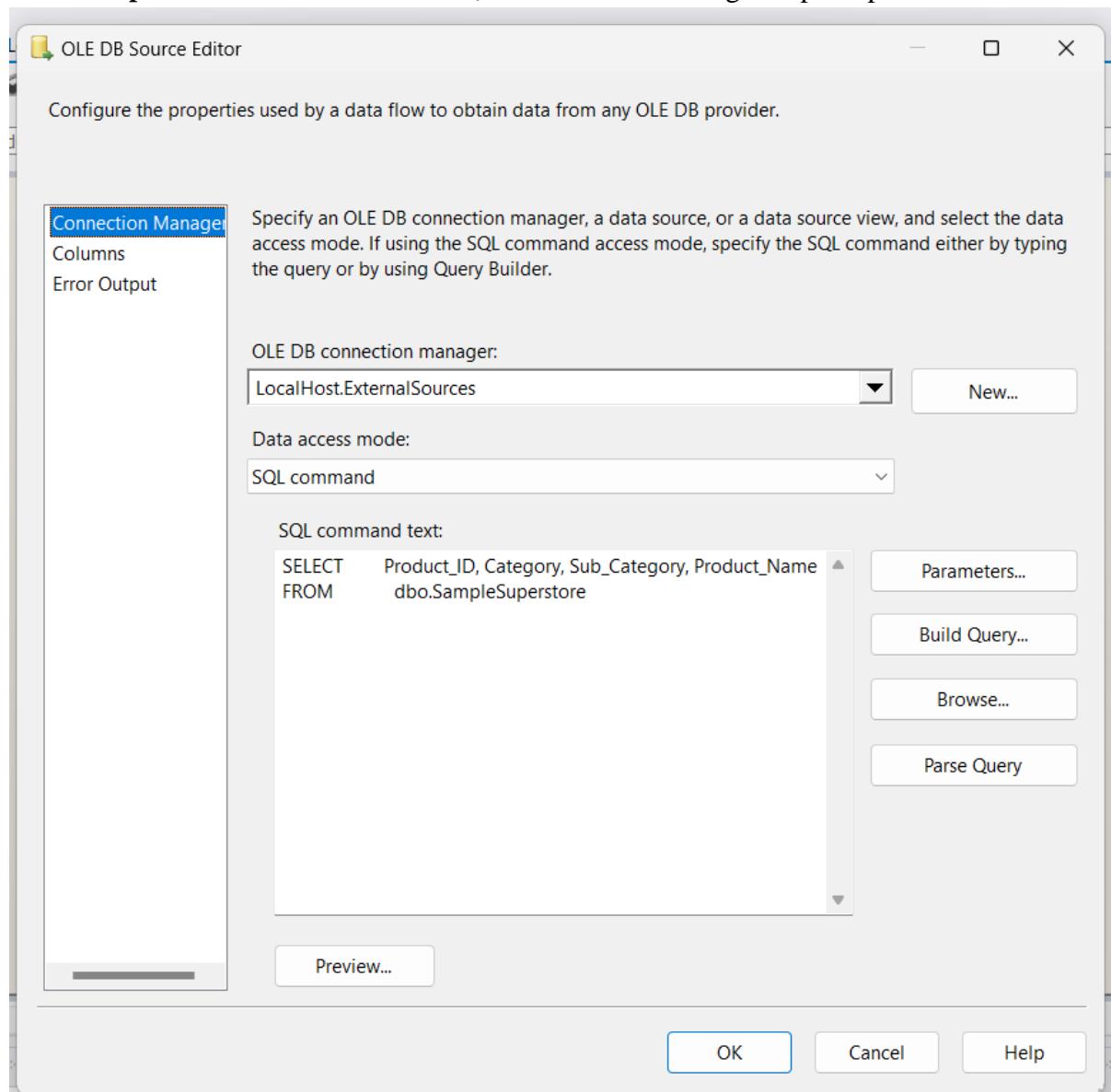
Khi chạy hoàn tất dữ liệu đã được Load vào DimDate

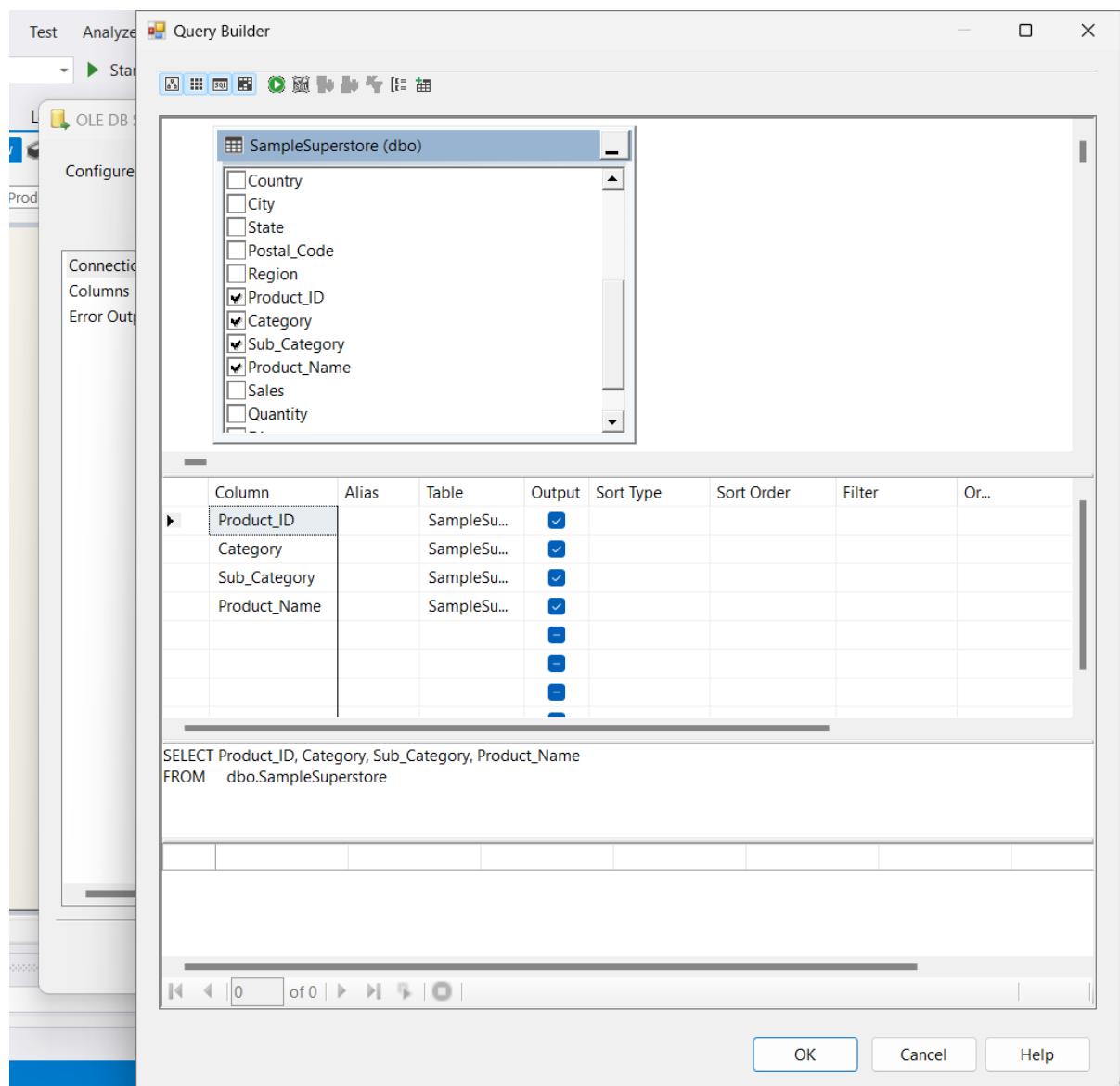
### 3.2.2. Product Dimension

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

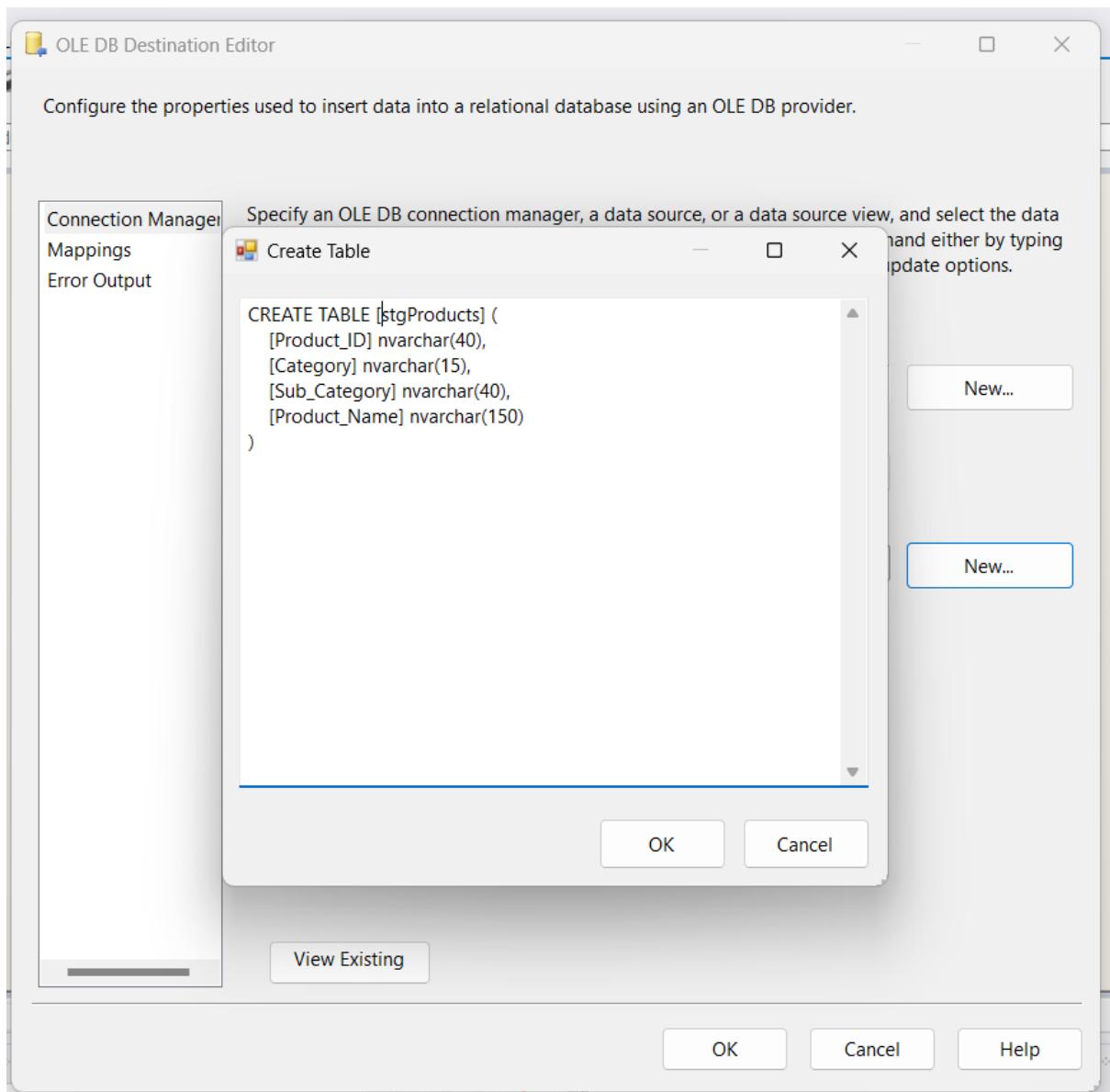


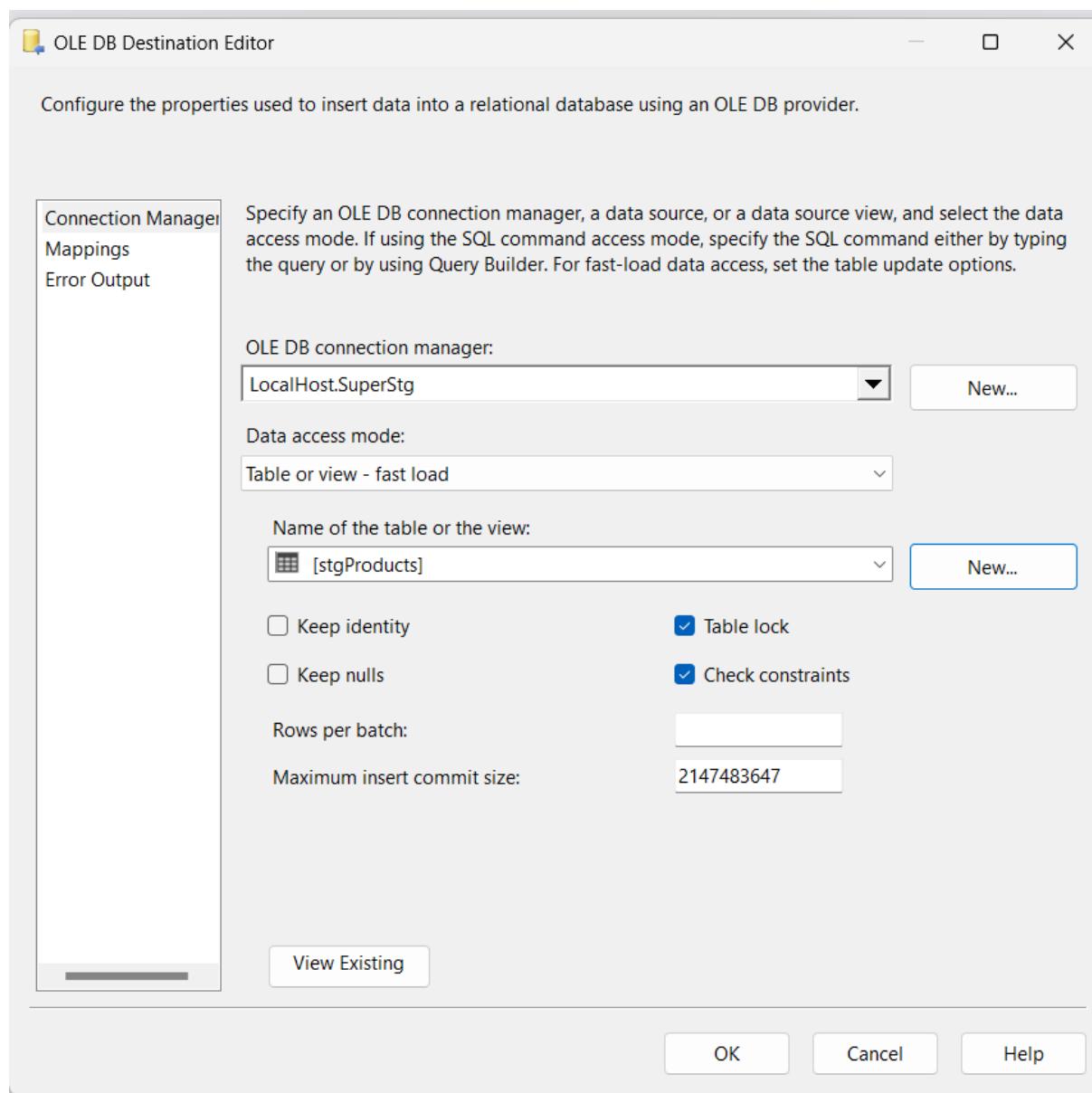
**SRC – SuperStore\_Product** chứa dữ liệu các Product từ bảng SampleSuperstore

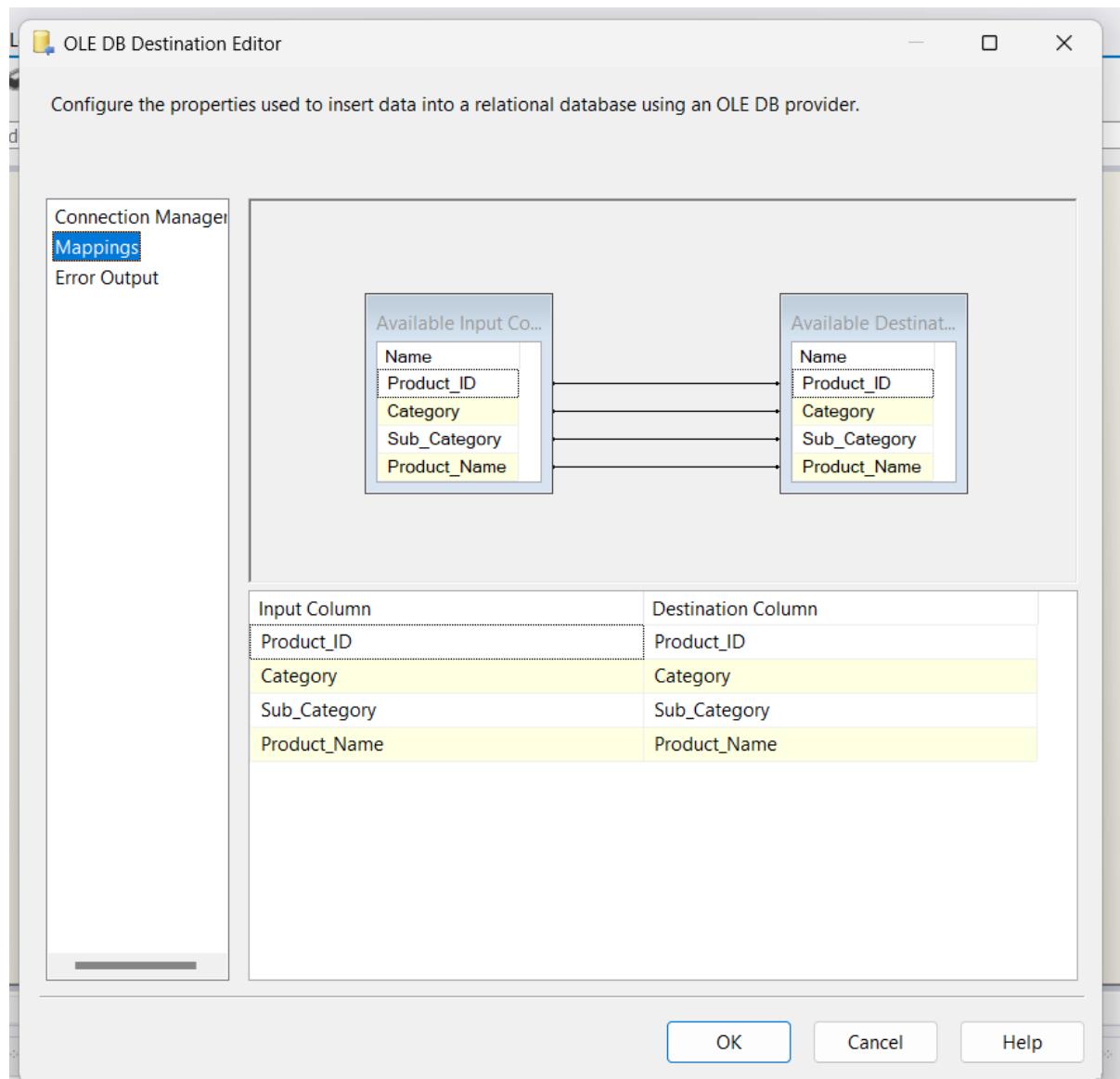


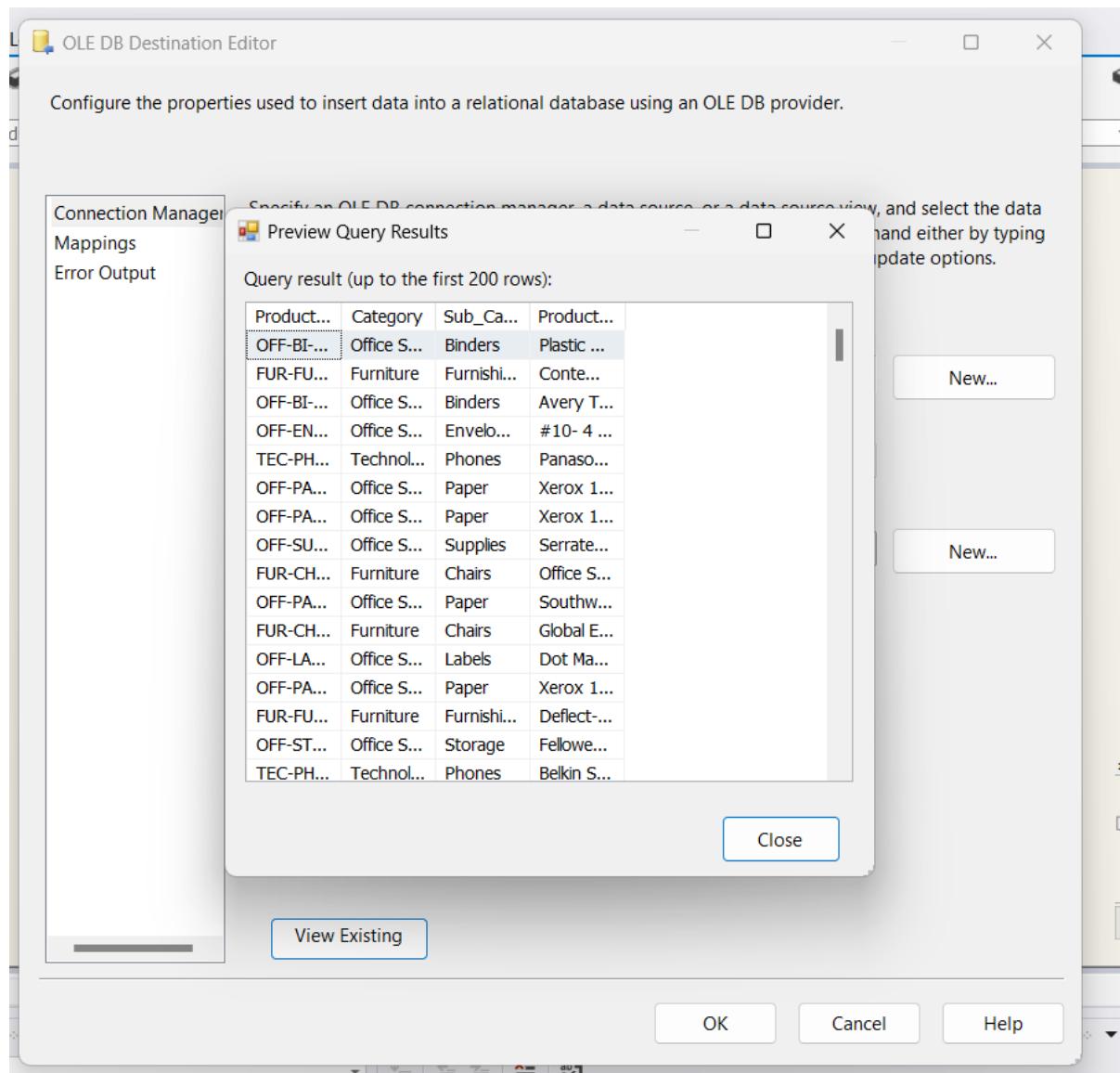


## DST – stgProduct

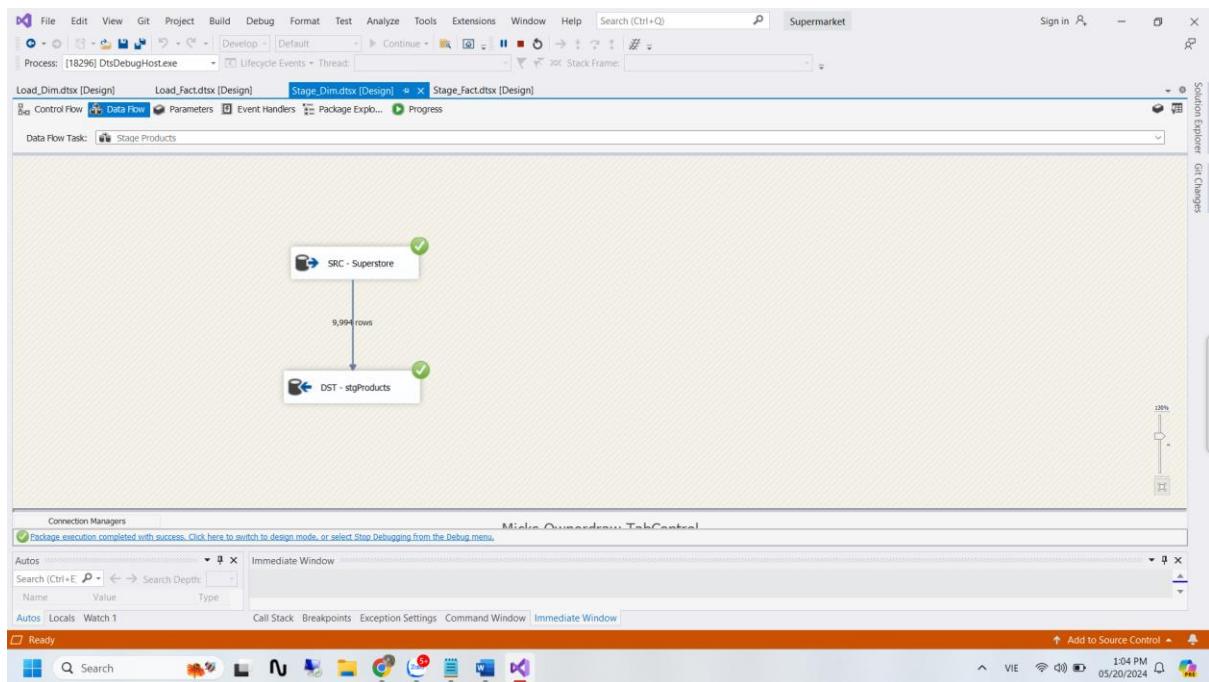




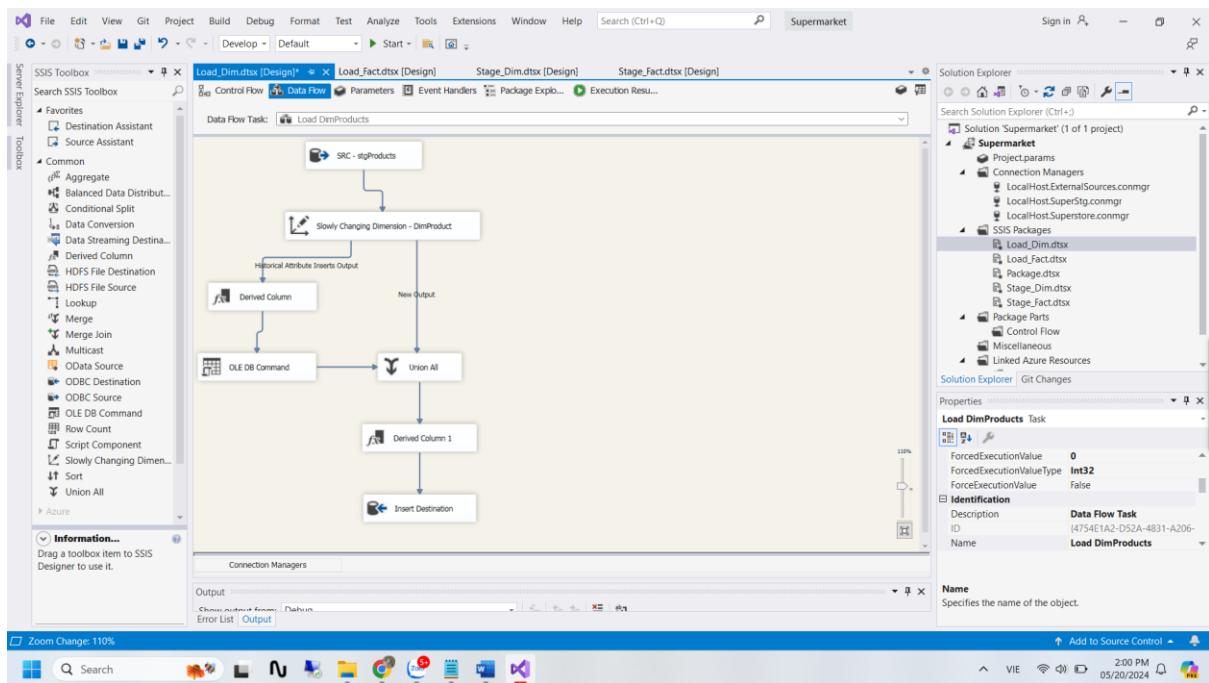




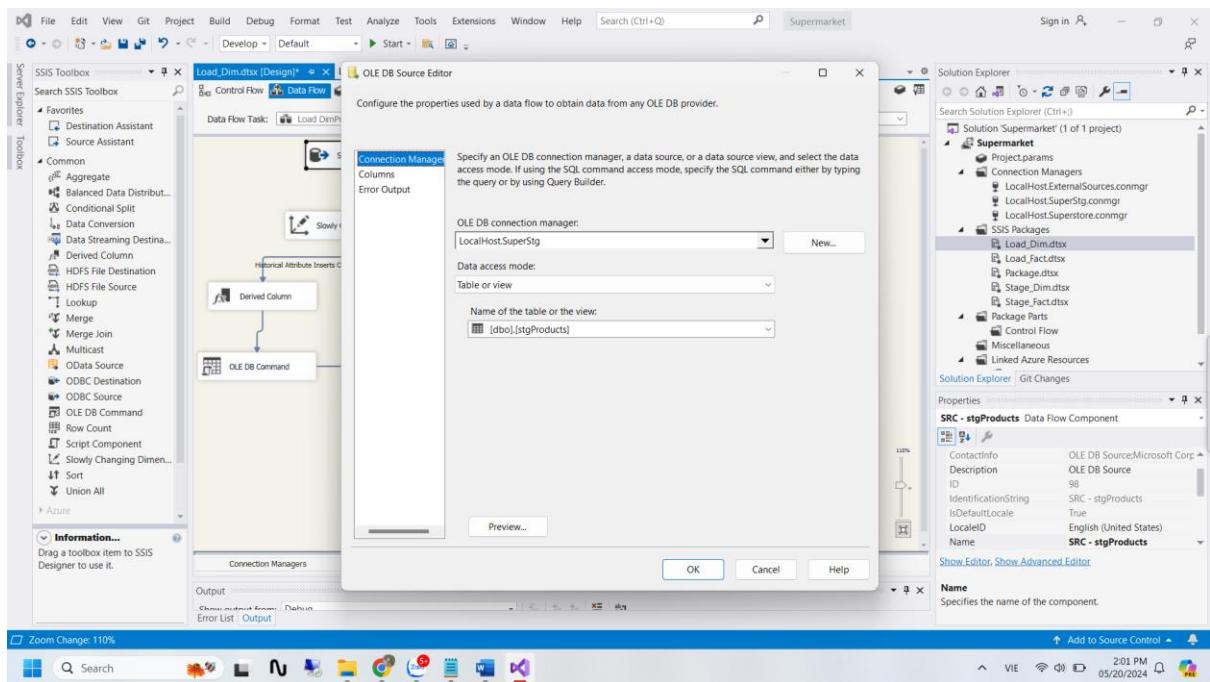
## Khởi chạy



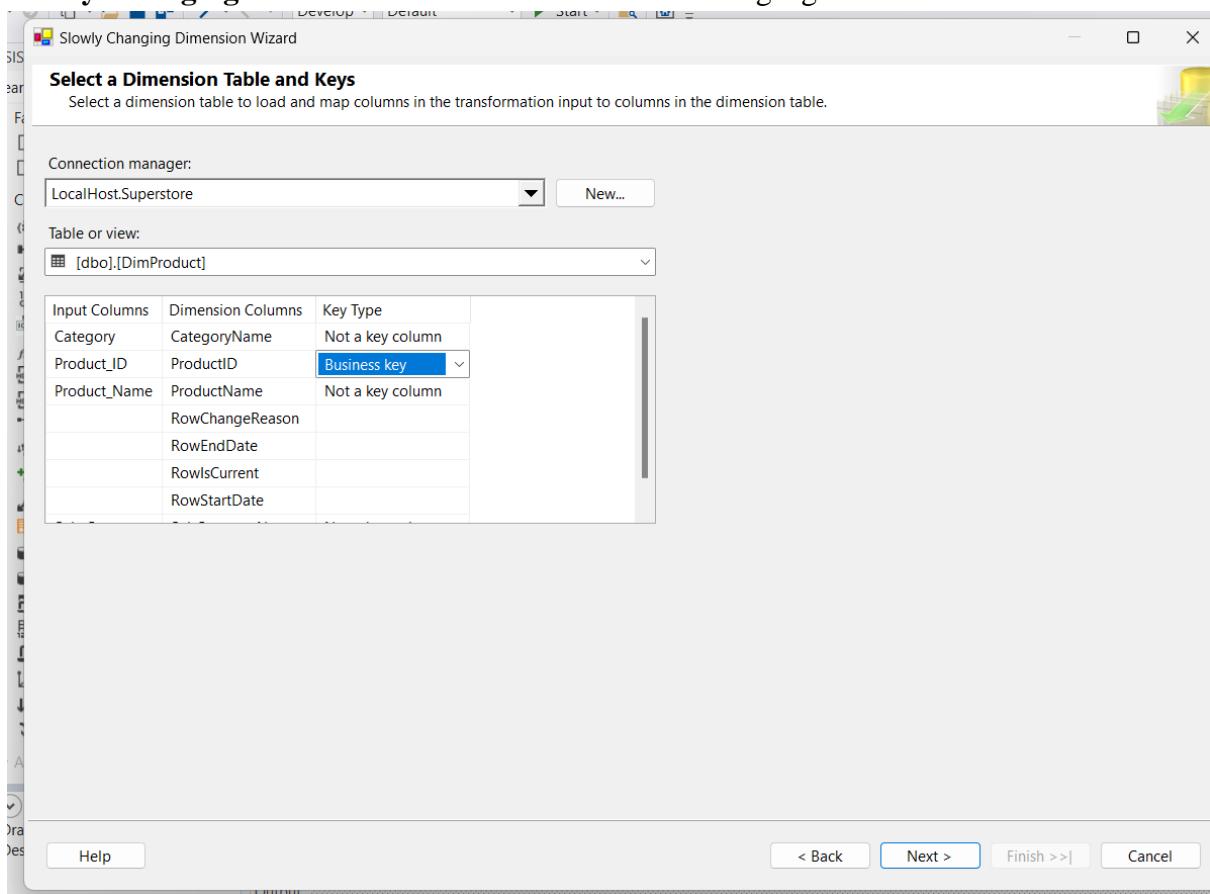
## Load dữ liệu từ stgProduct vào DimProduct

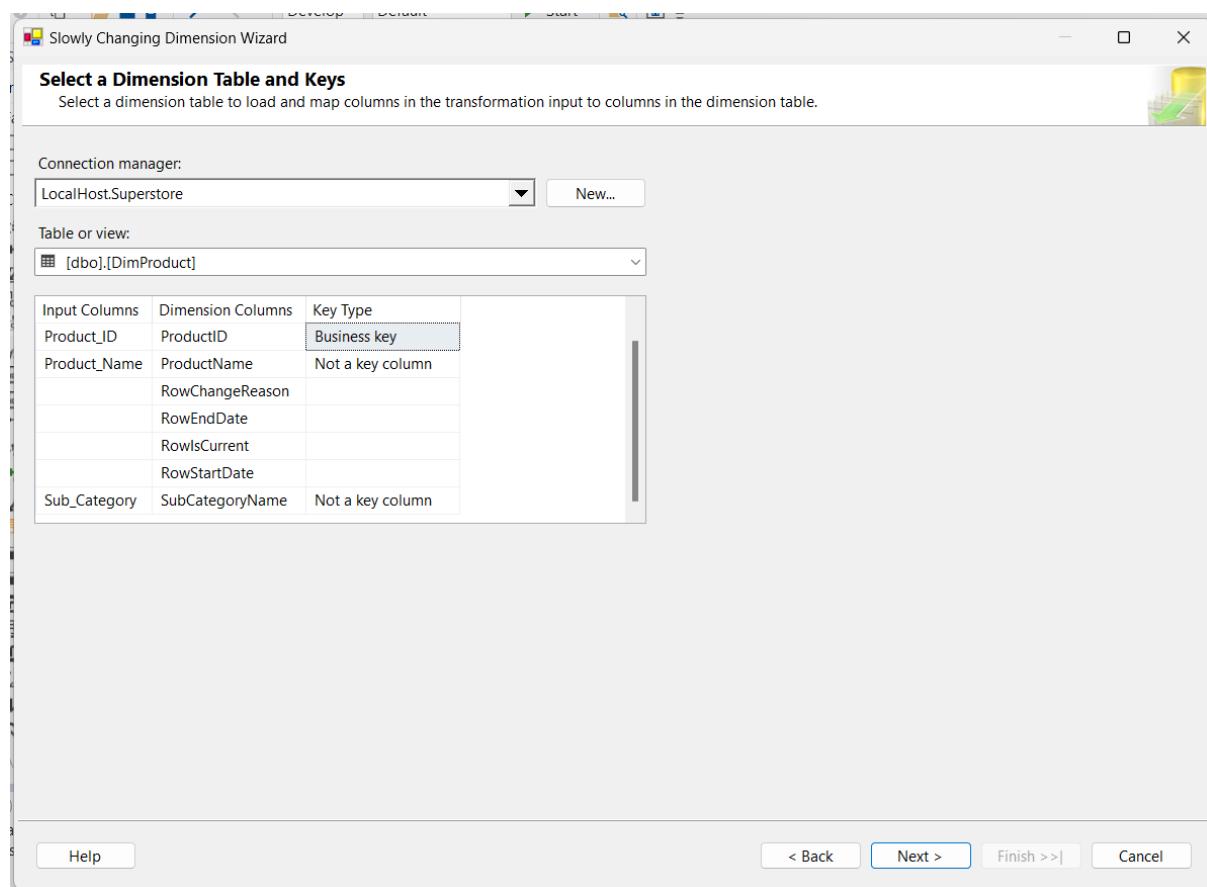


**SRC – stgProduct** là dữ liệu từ bảng stgProduct

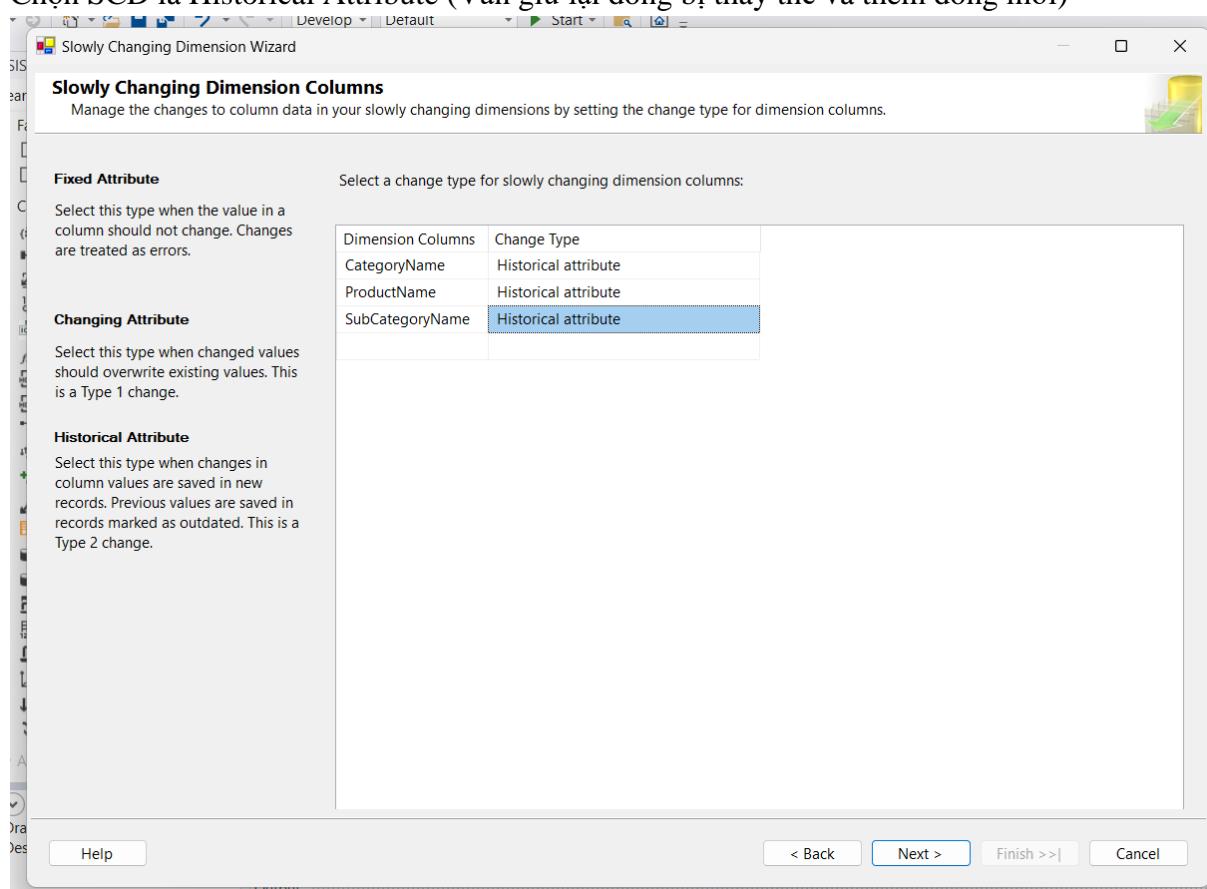


## Slowly Changing Dimension – DimProduct Load từ bảng stgProduct vào DimProduct





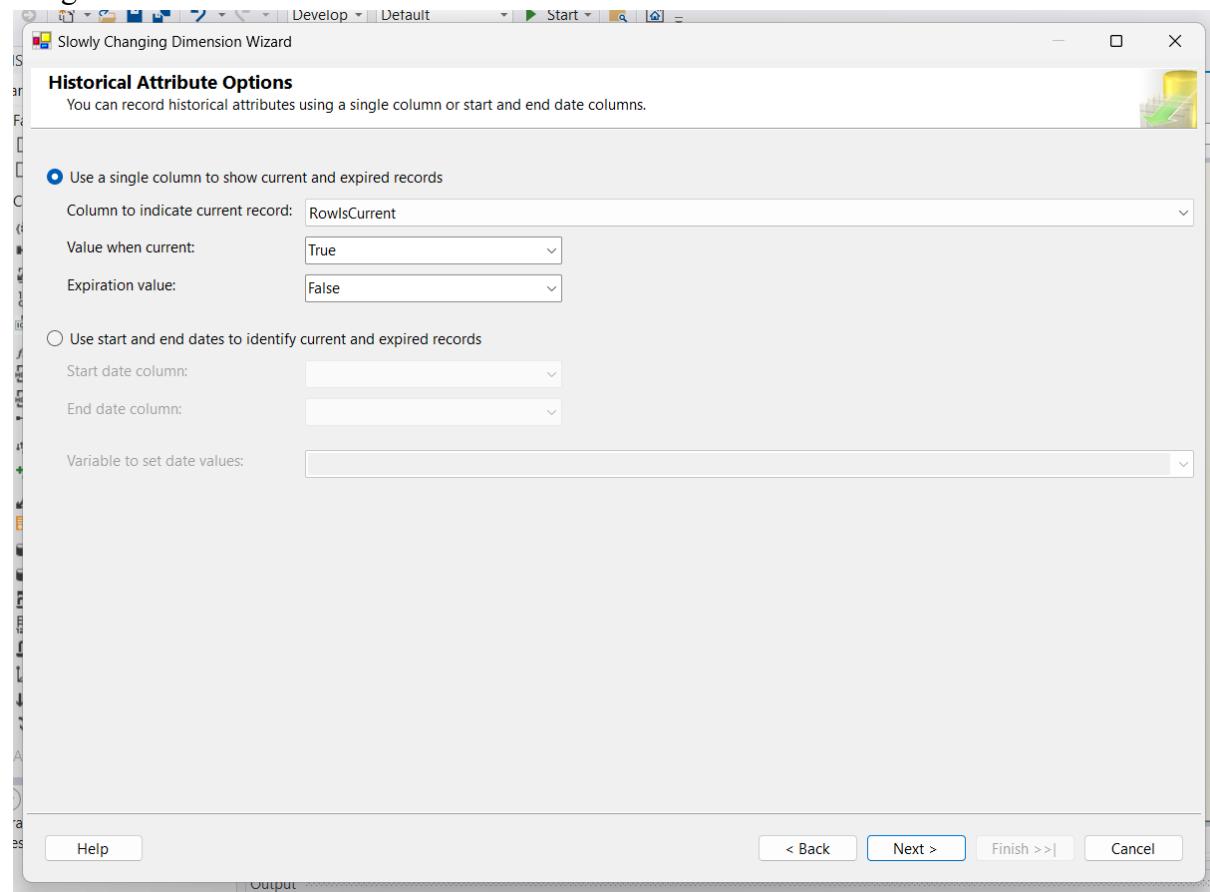
Chọn SCD là Historical Attribute (Vẫn giữ lại dòng bị thay thế và thêm dòng mới)



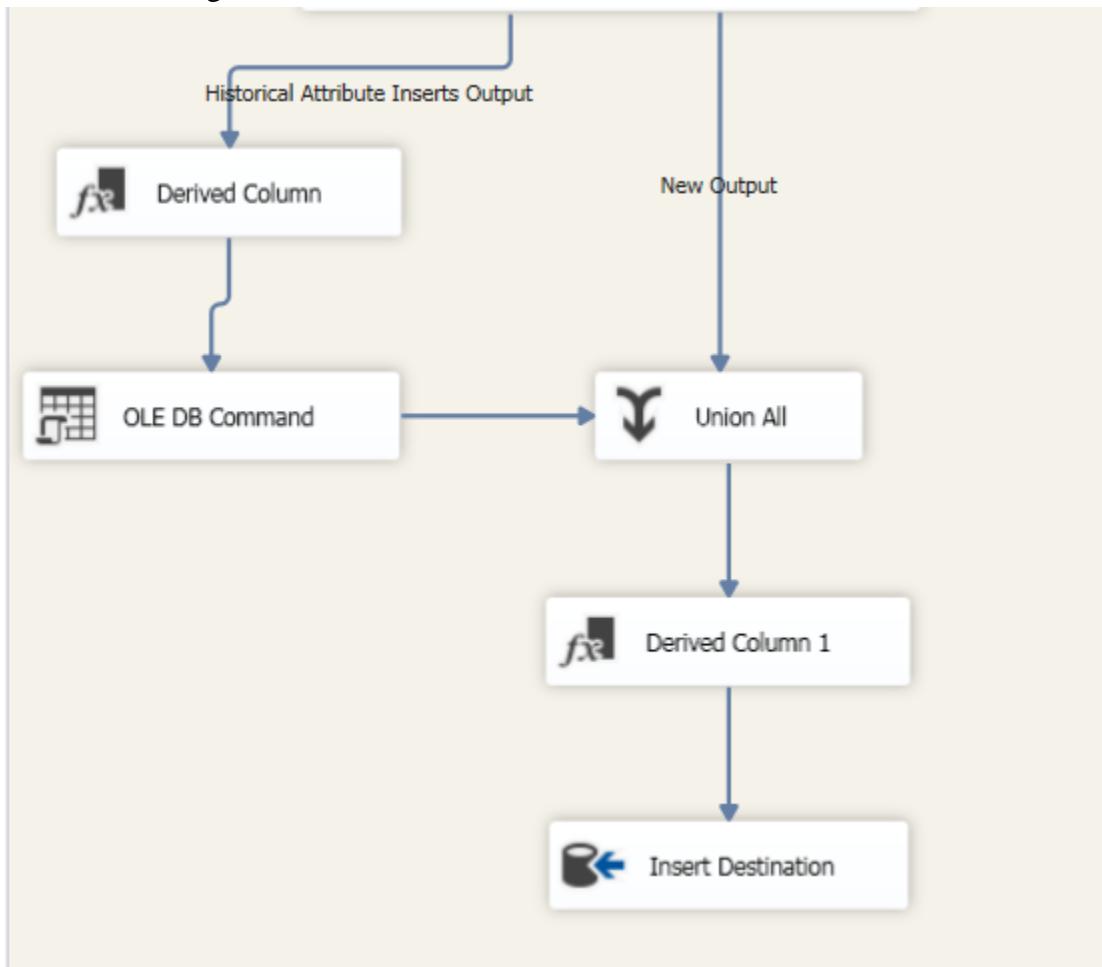
Khi một dòng trong bảng DimProduct cột RowIsCurrent sẽ có giá trị là 1

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

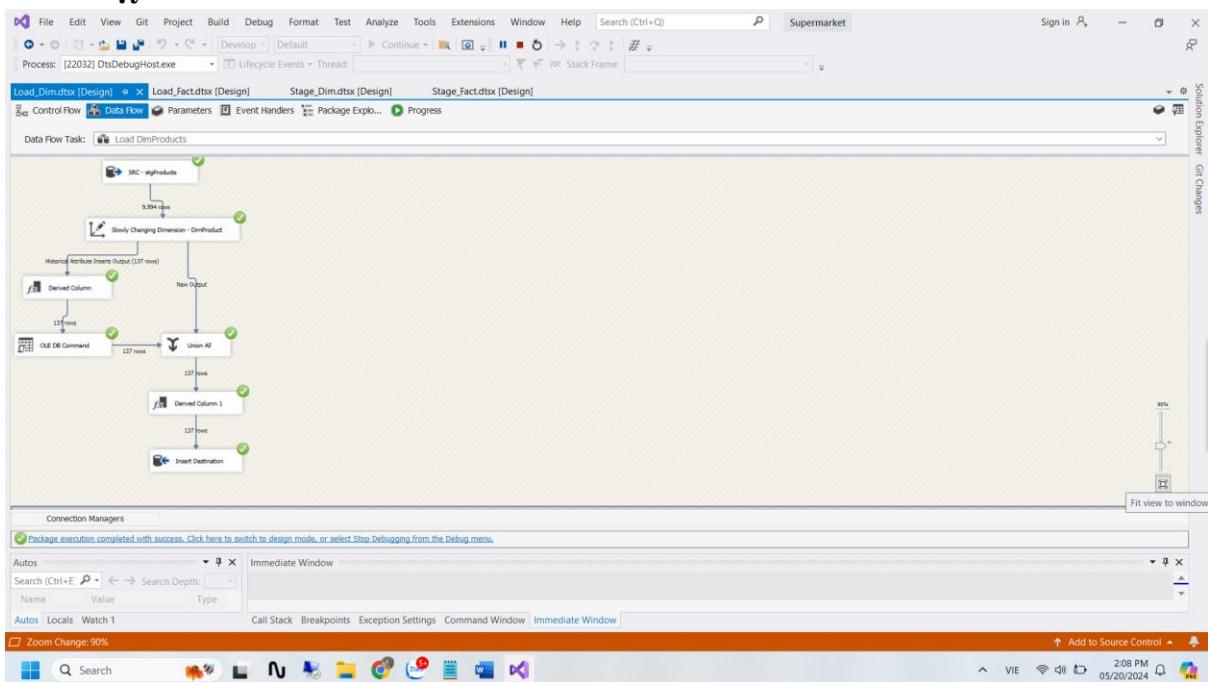
Sau đó được nâng hạng lên vàng thì sẽ thêm 1 dòng mới với mã khách hàng đó và hạng tiêu dùng là vàng với rowiscurrent là 1 và sẽ sửa rowiscurrent của dòng cũ thành 0 mà không xóa dòng cũ



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

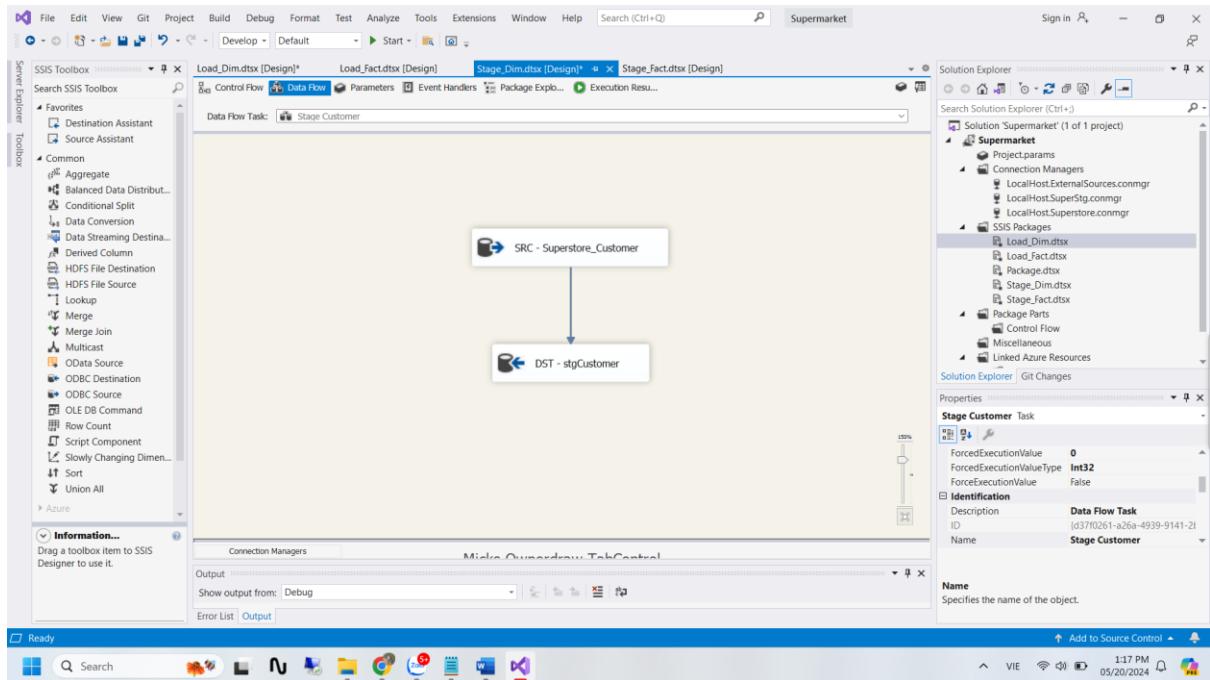


## Khởi chạy

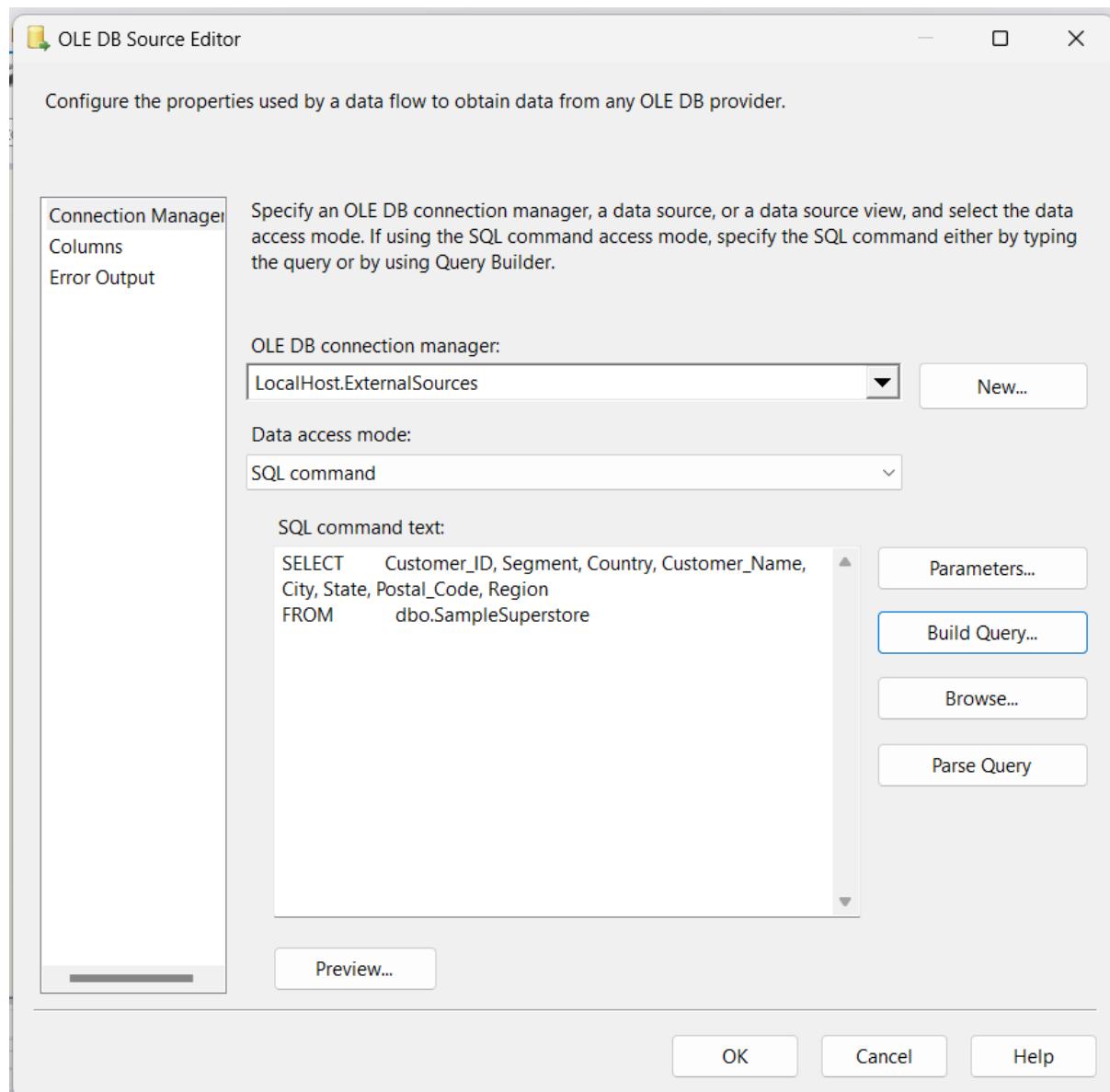


### 3.2.3.Customer Dimension

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



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



Query Builder

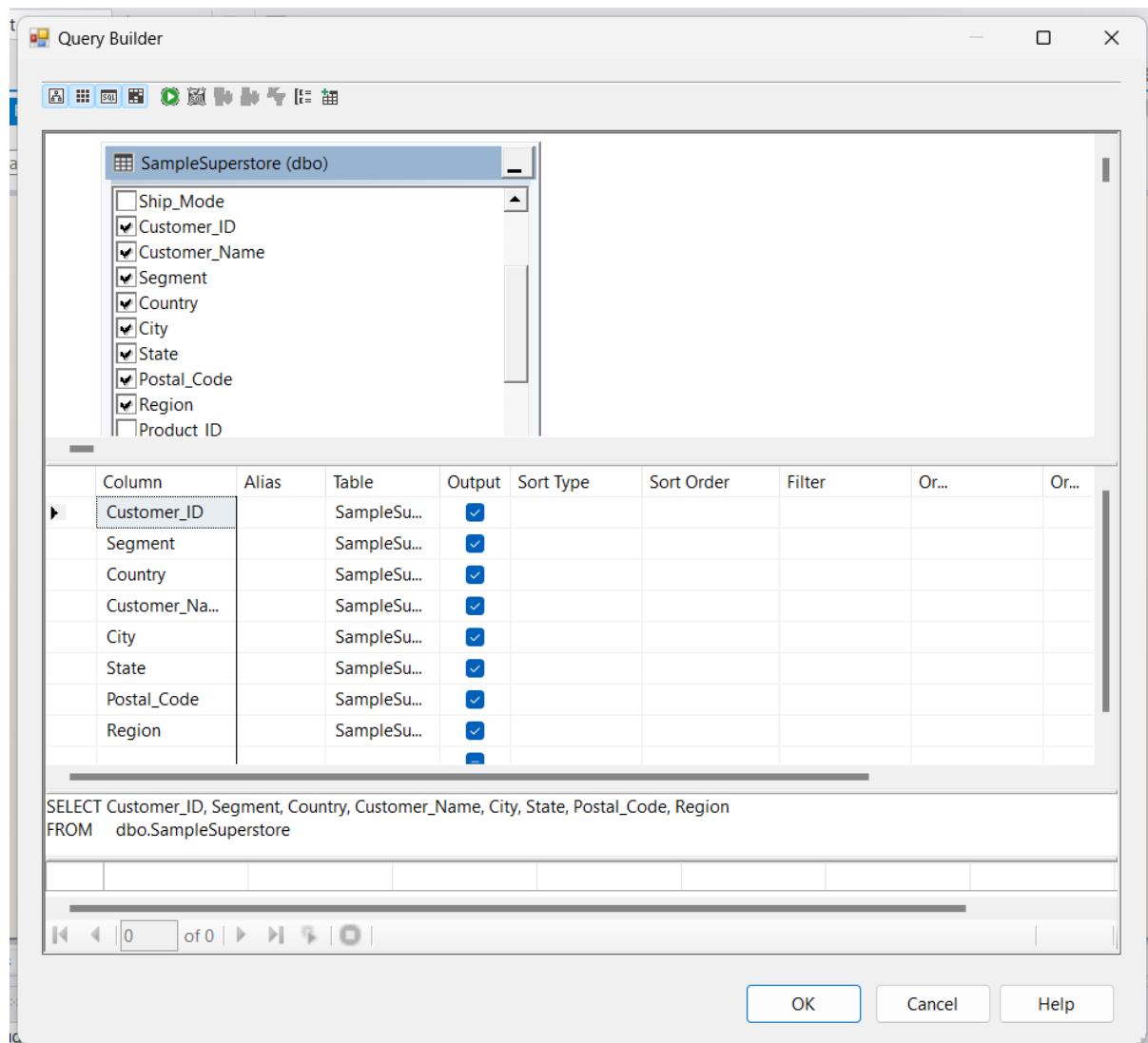
SampleSuperstore (dbo)

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...	✓					

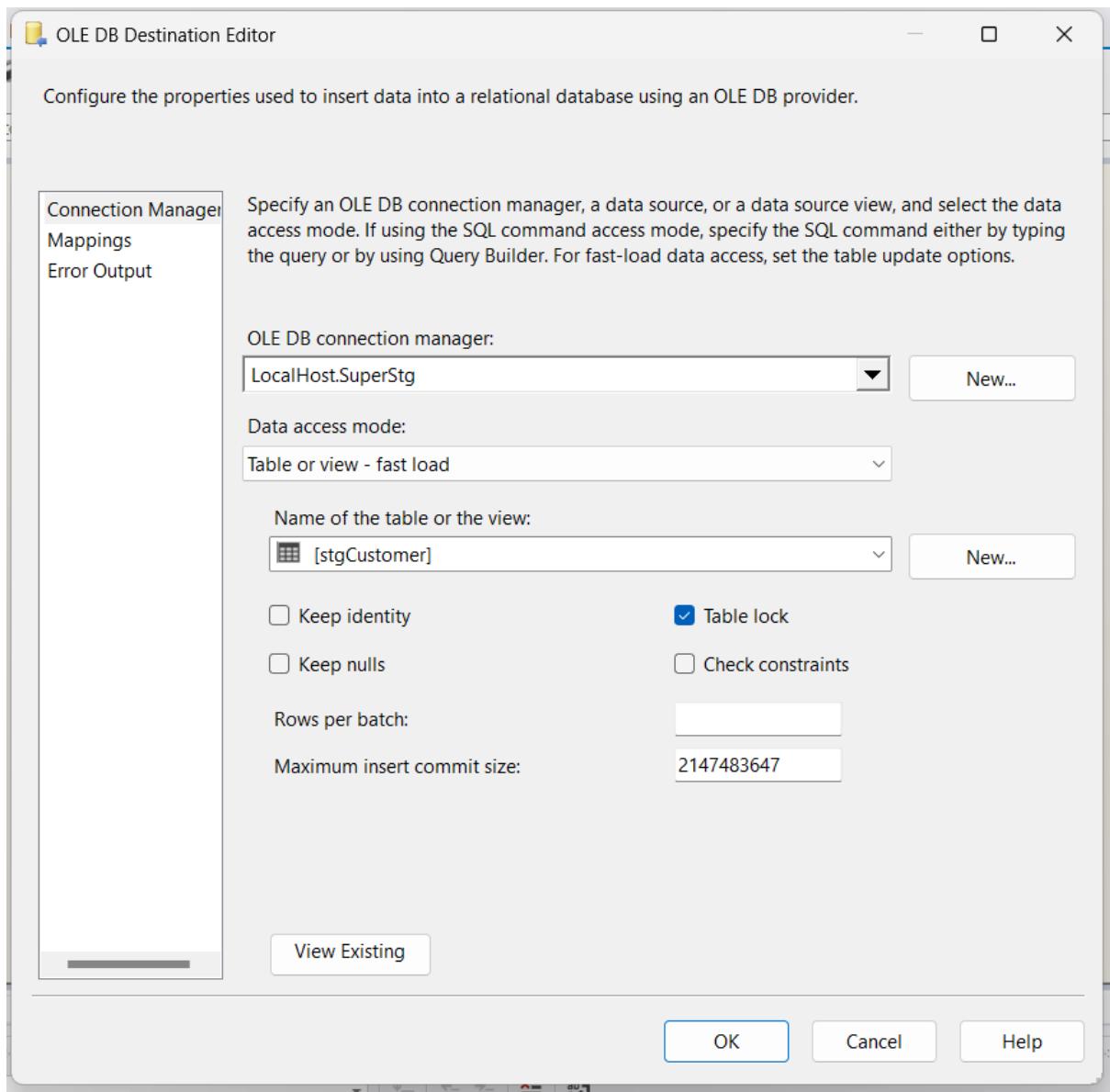
```
SELECT Customer_ID, Segment, Country, Customer_Name, City, State, Postal_Code, Region
FROM   dbo.SampleSuperstore
```

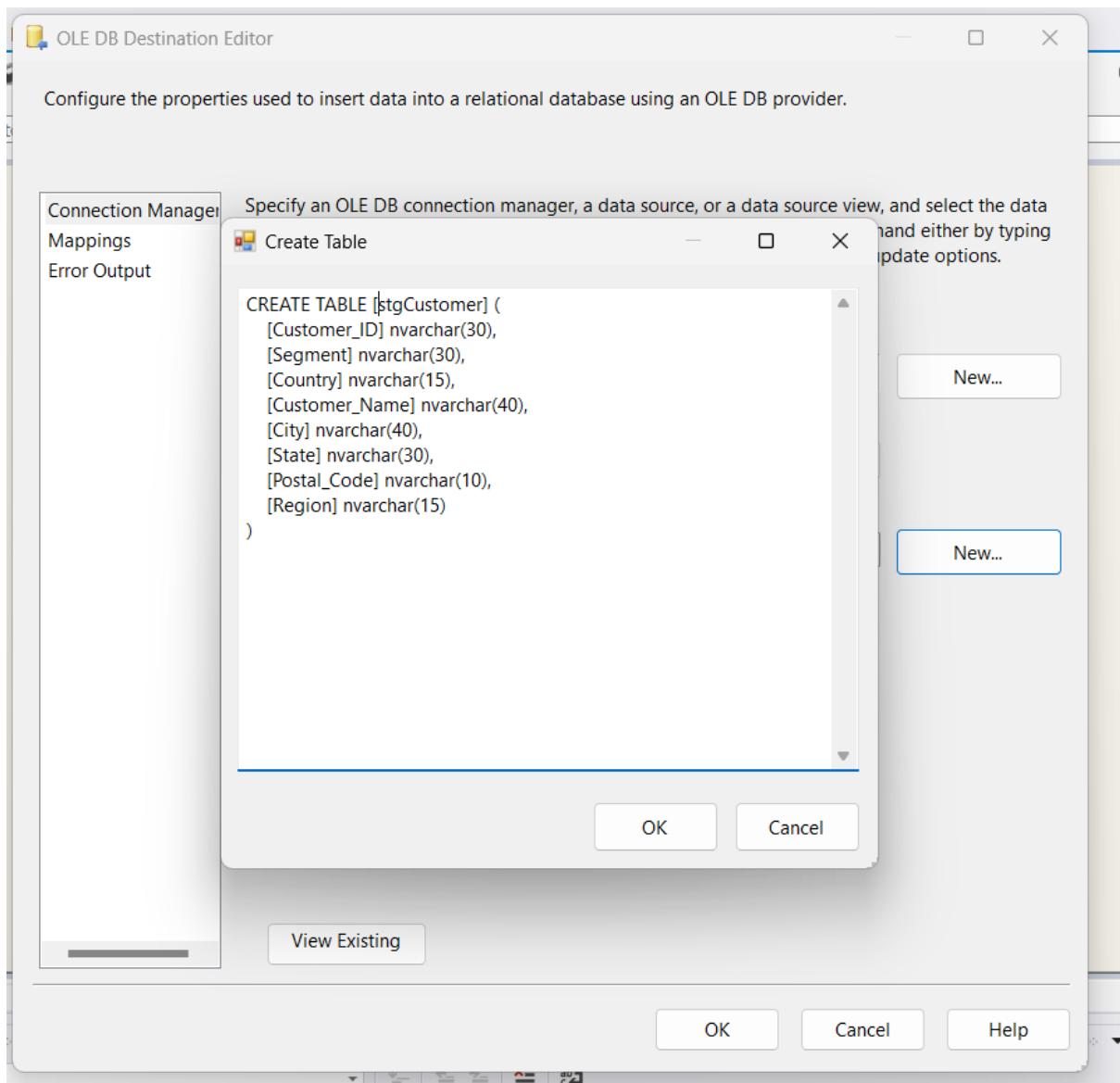
IG

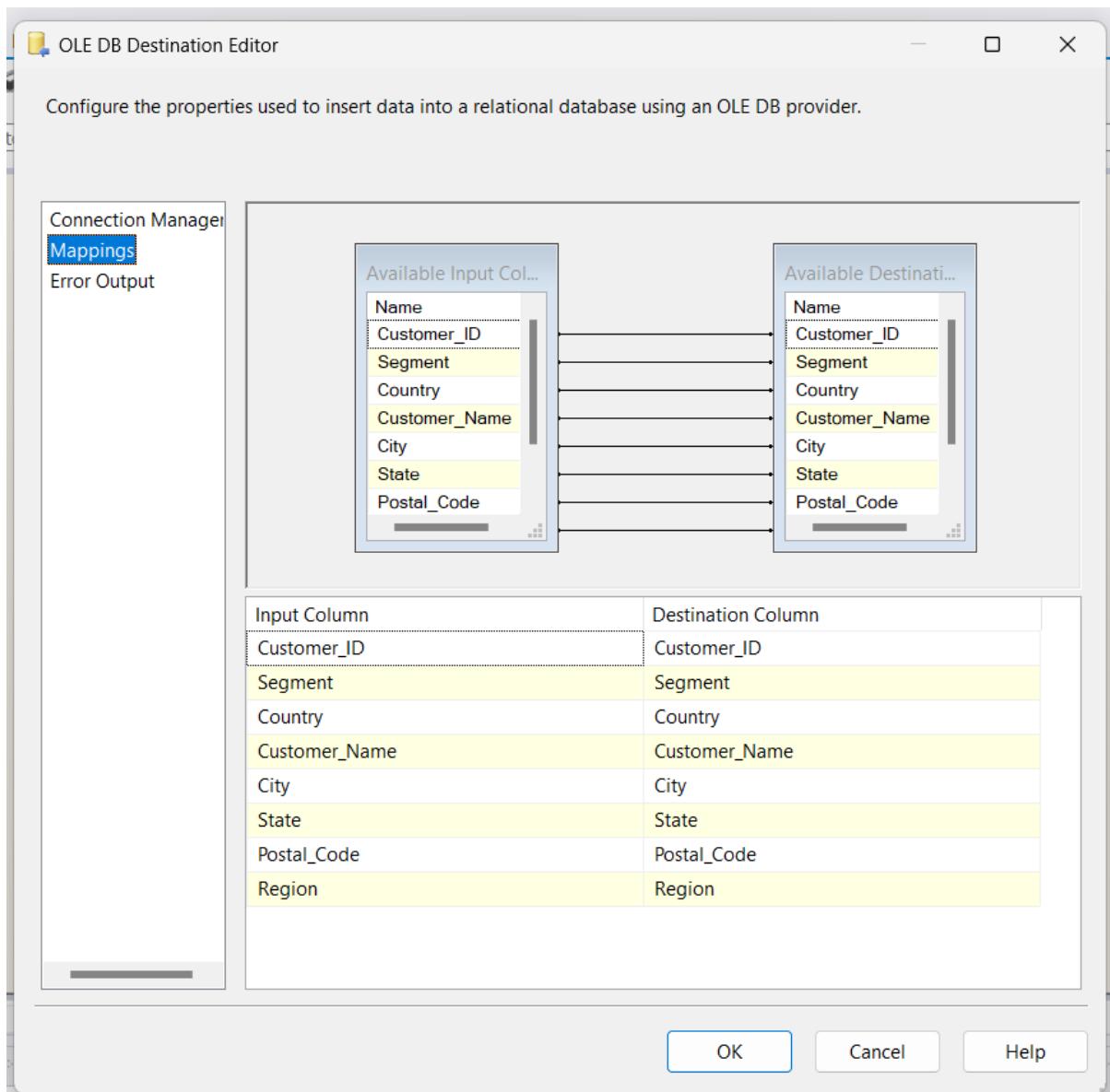
OK Cancel Help

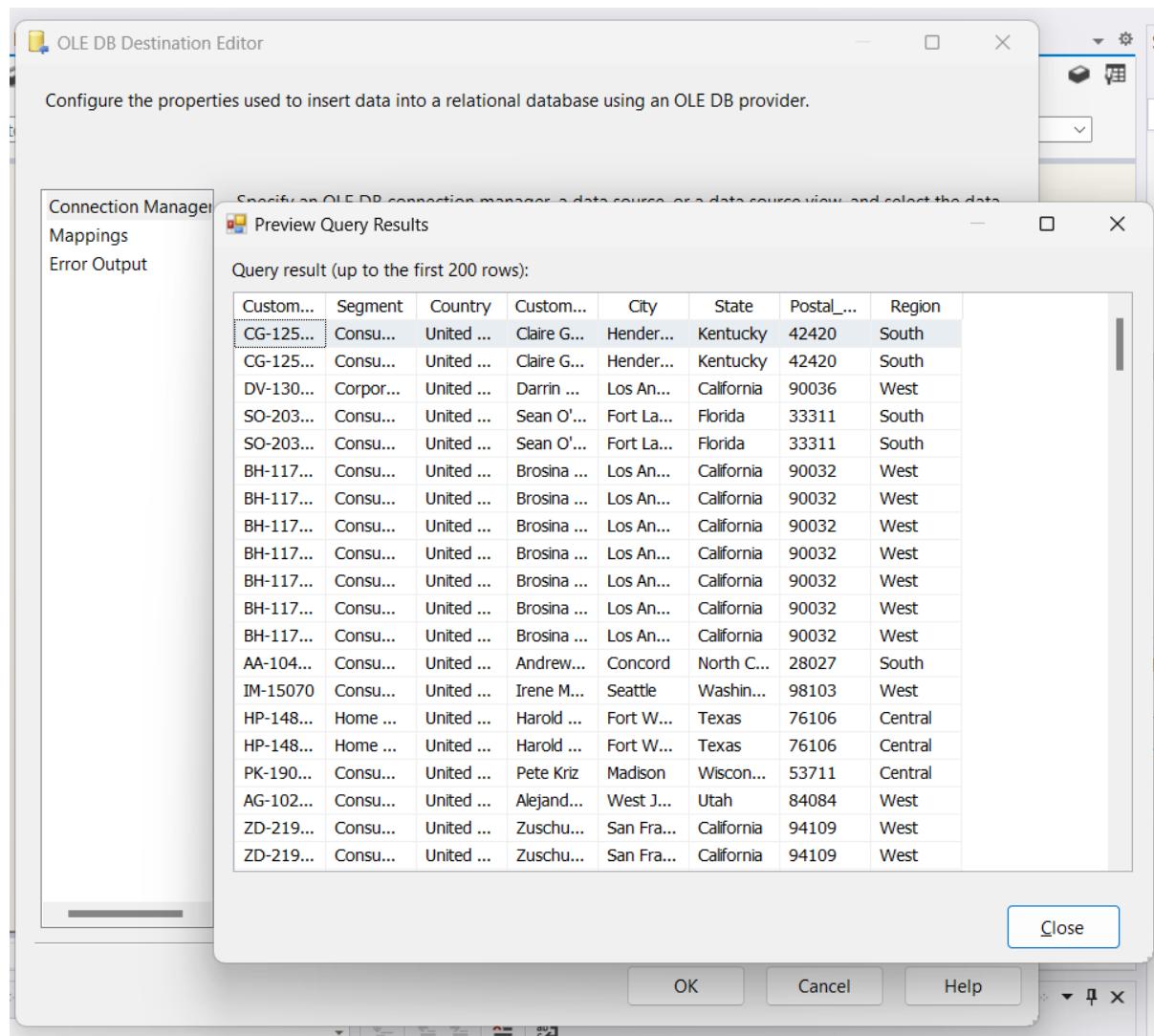


## DST – stgCustomer

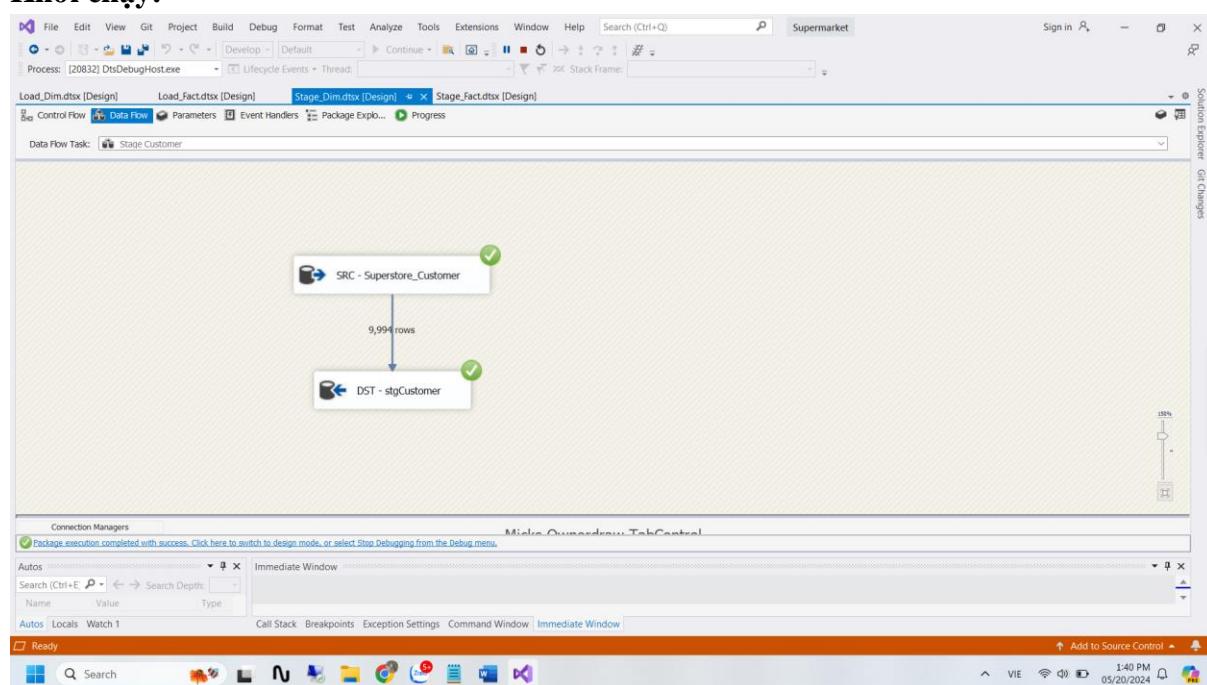




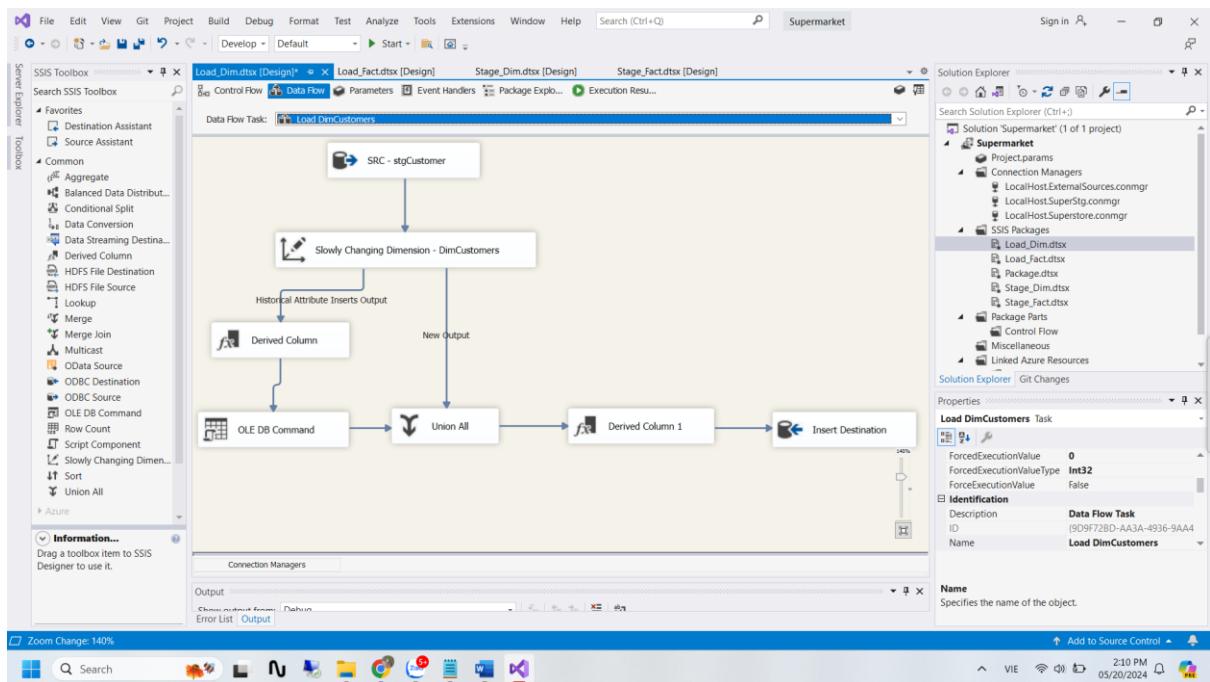




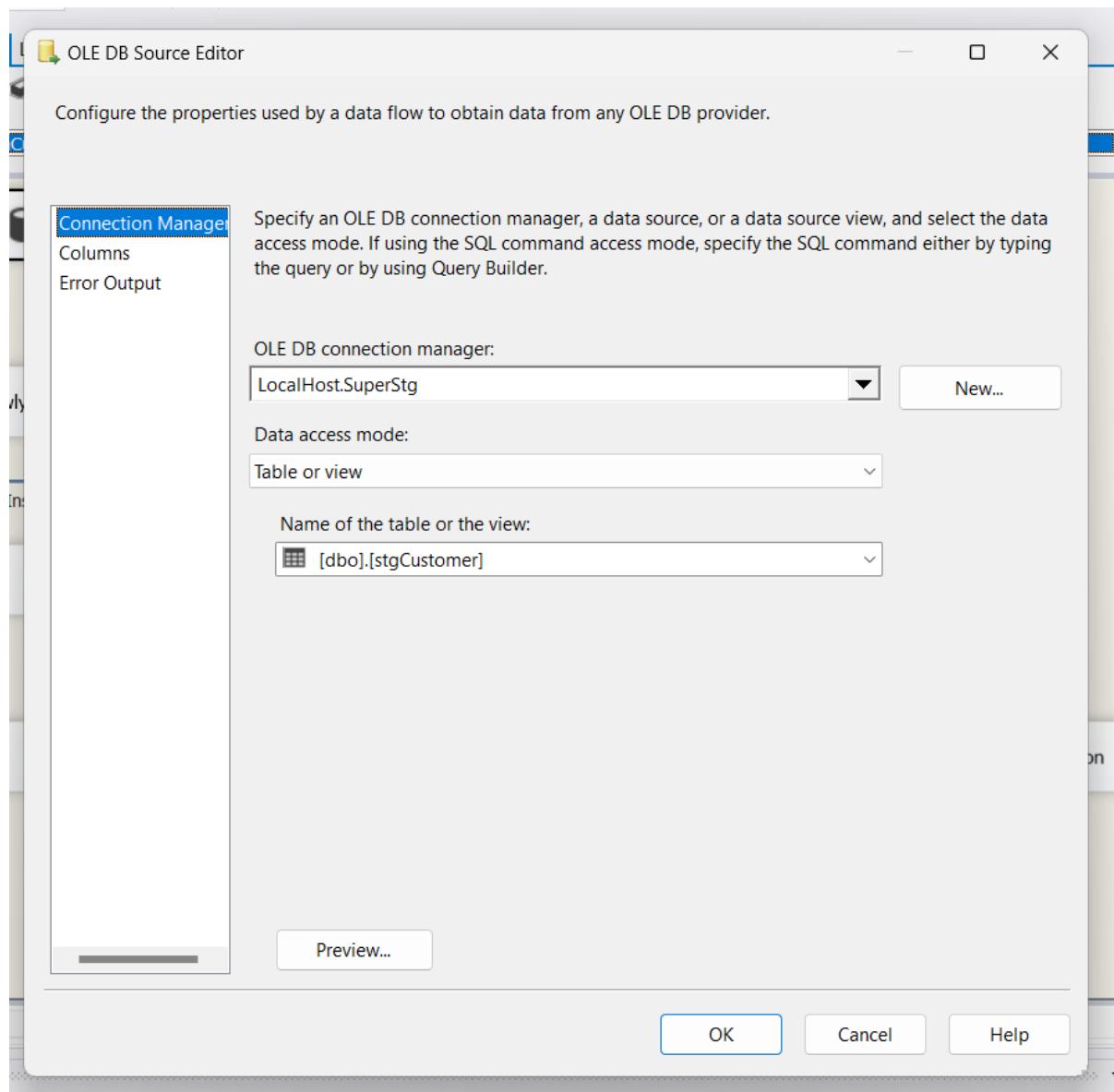
## Khởi chạy:



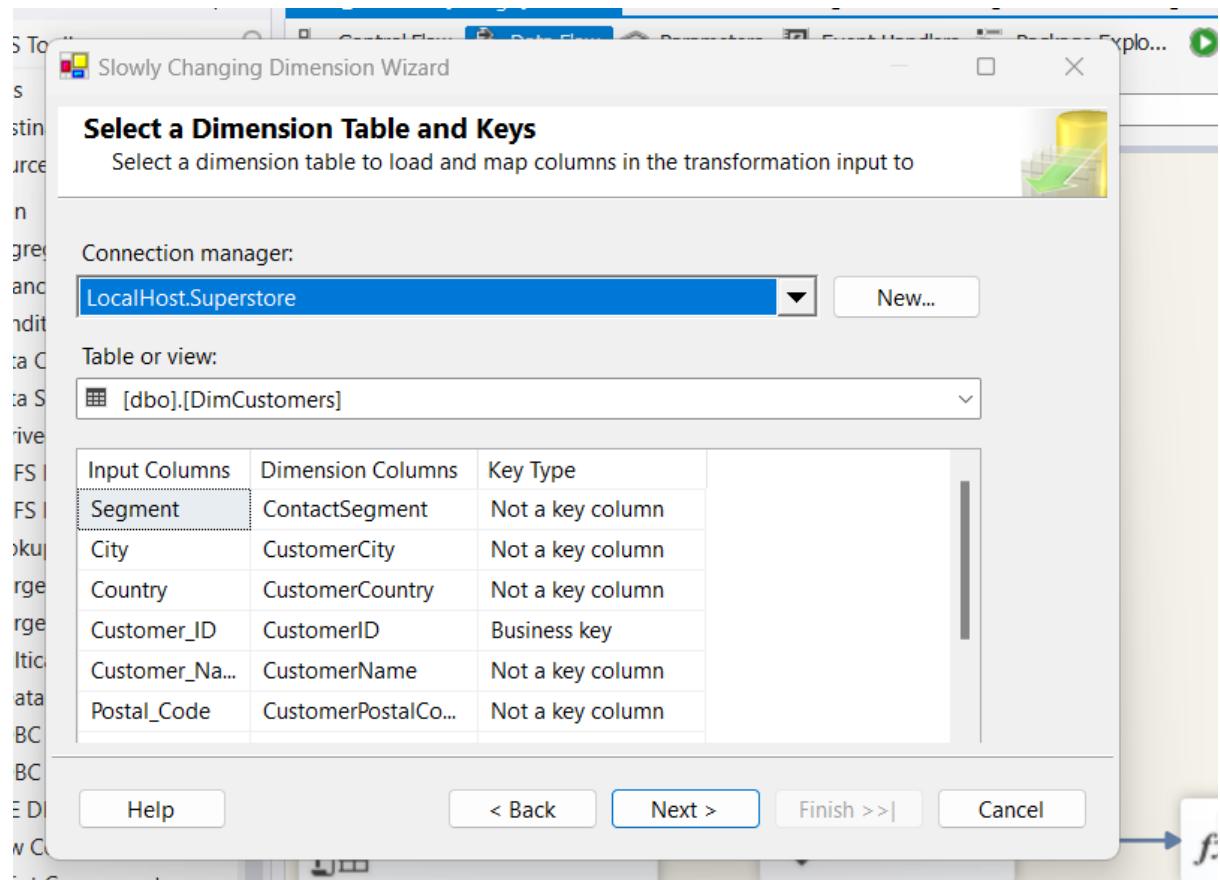
b) Load dữ liệu từ stgCustomer vào bảng DimCustomer



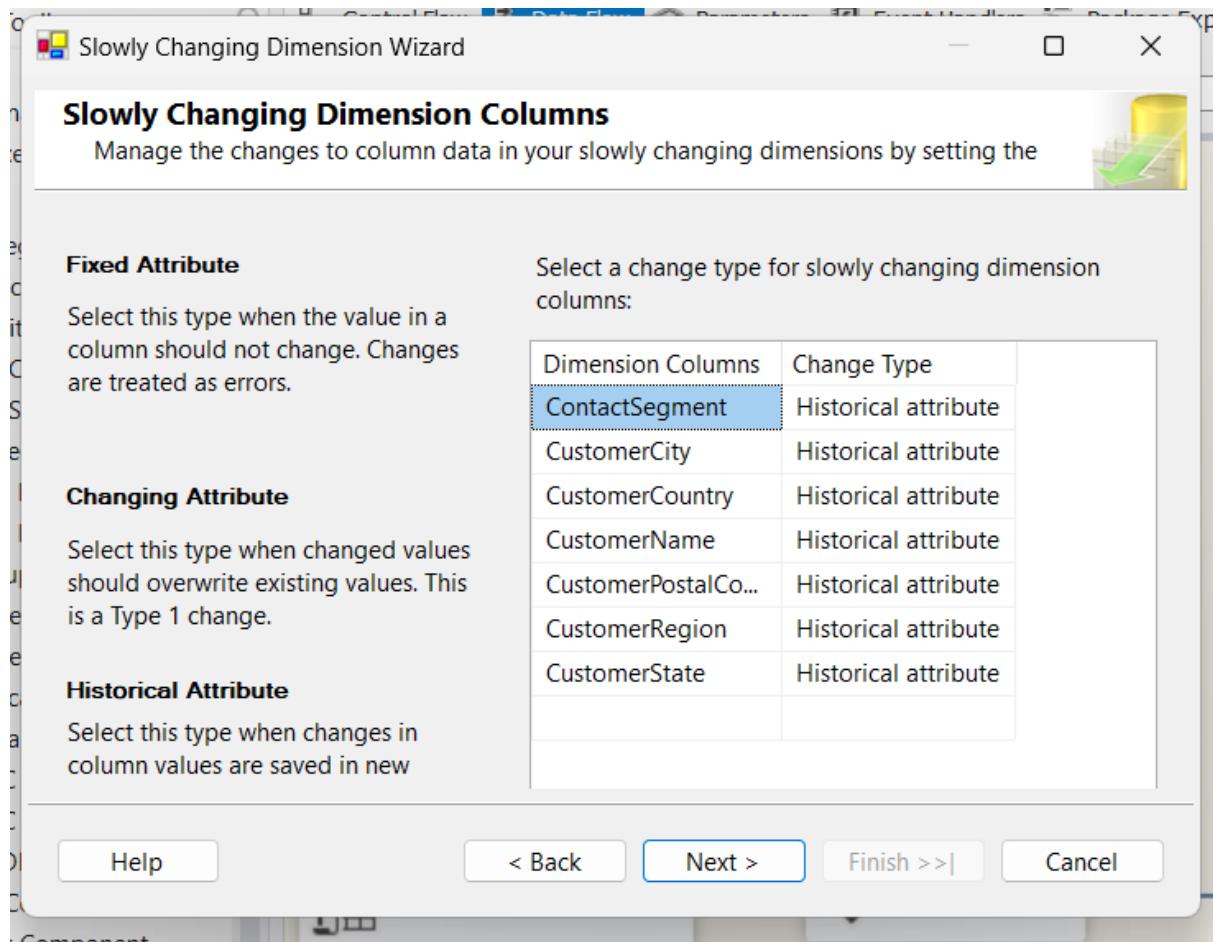
**SRC - stgCustomer** chứa dữ liệu bảng stgCustomer



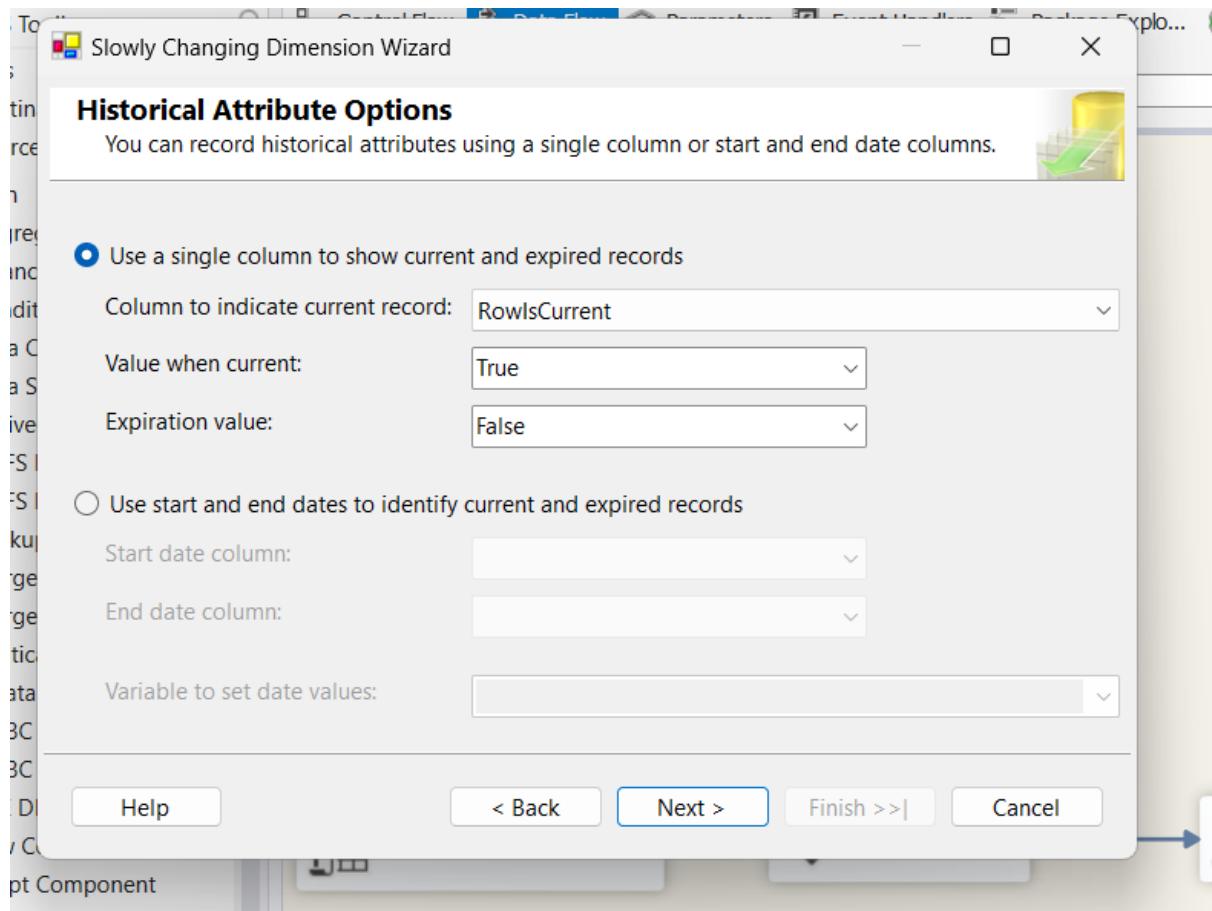
Slowly Changing Dimension – DimCustomers load dữ liệu vào bảng DimCustomer database SuperStore



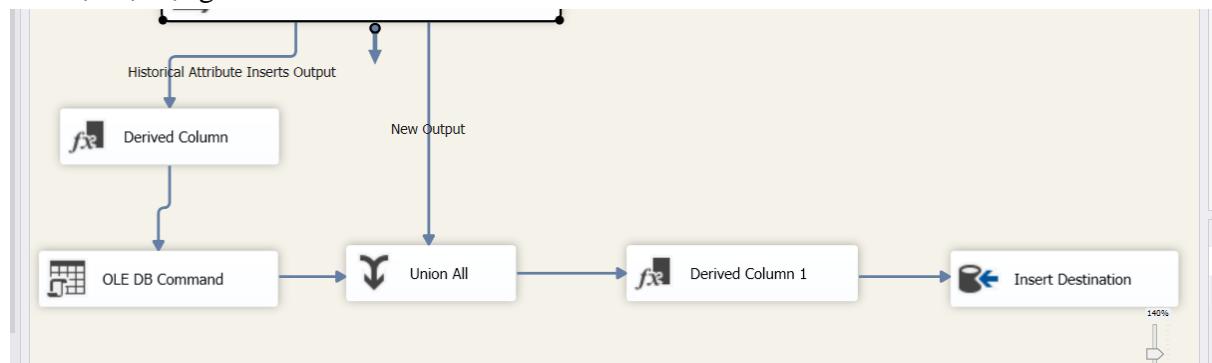
SCD chọn Historical attribute



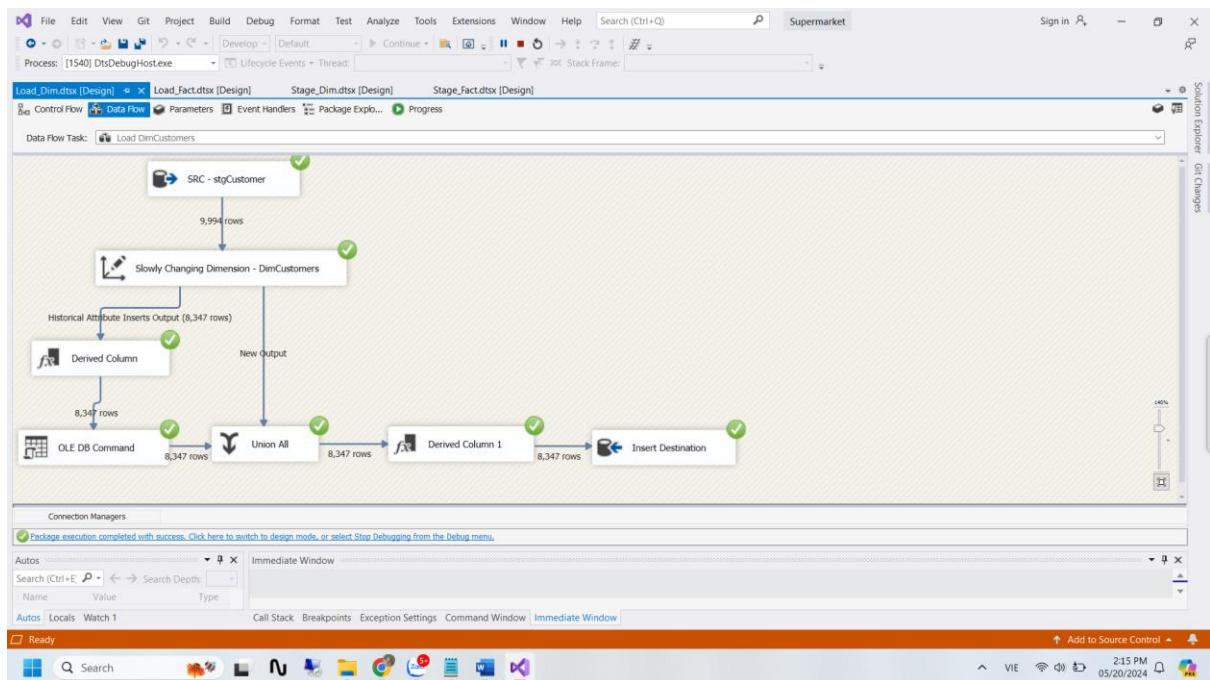
Historical attribute options chọn RowIsCurrent



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

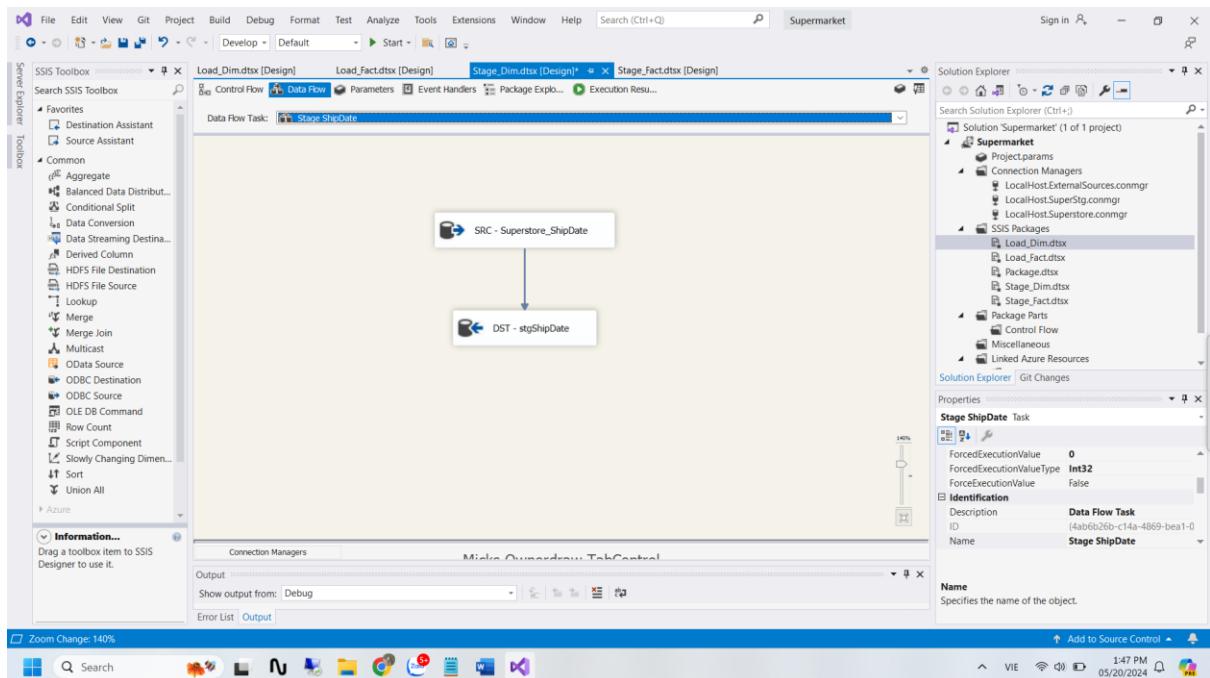


**Khởi chạy**

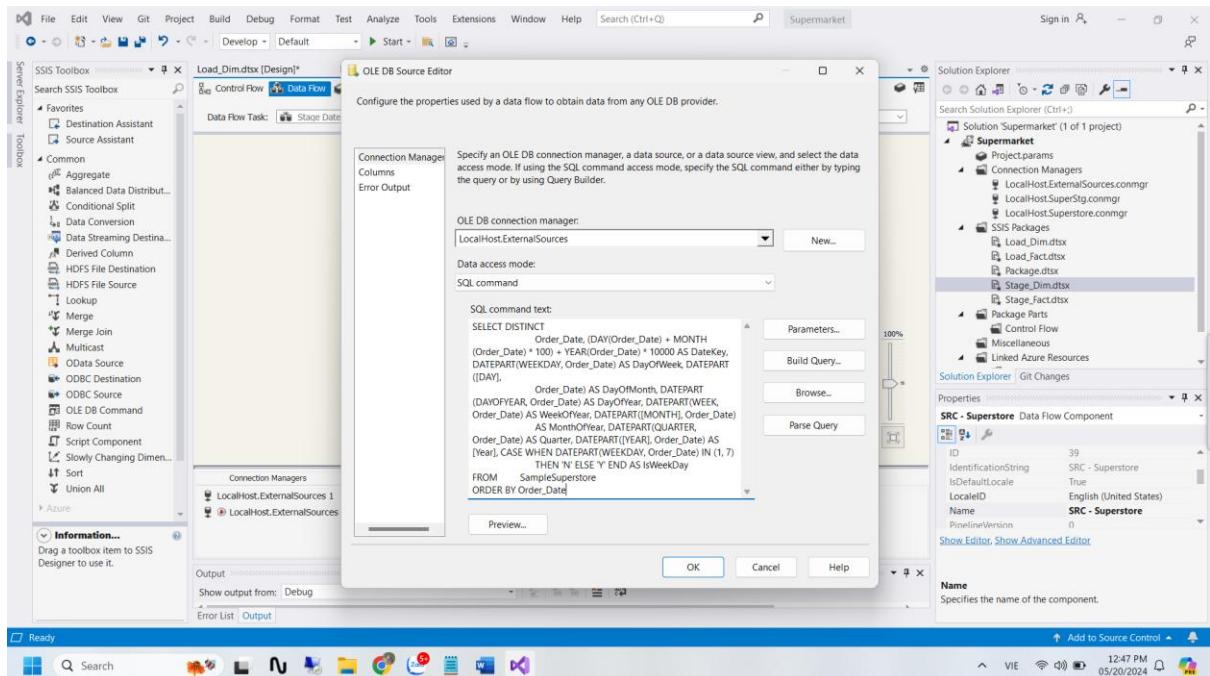


### 3.2.4. ShipDate Dimension

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



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



## Script :

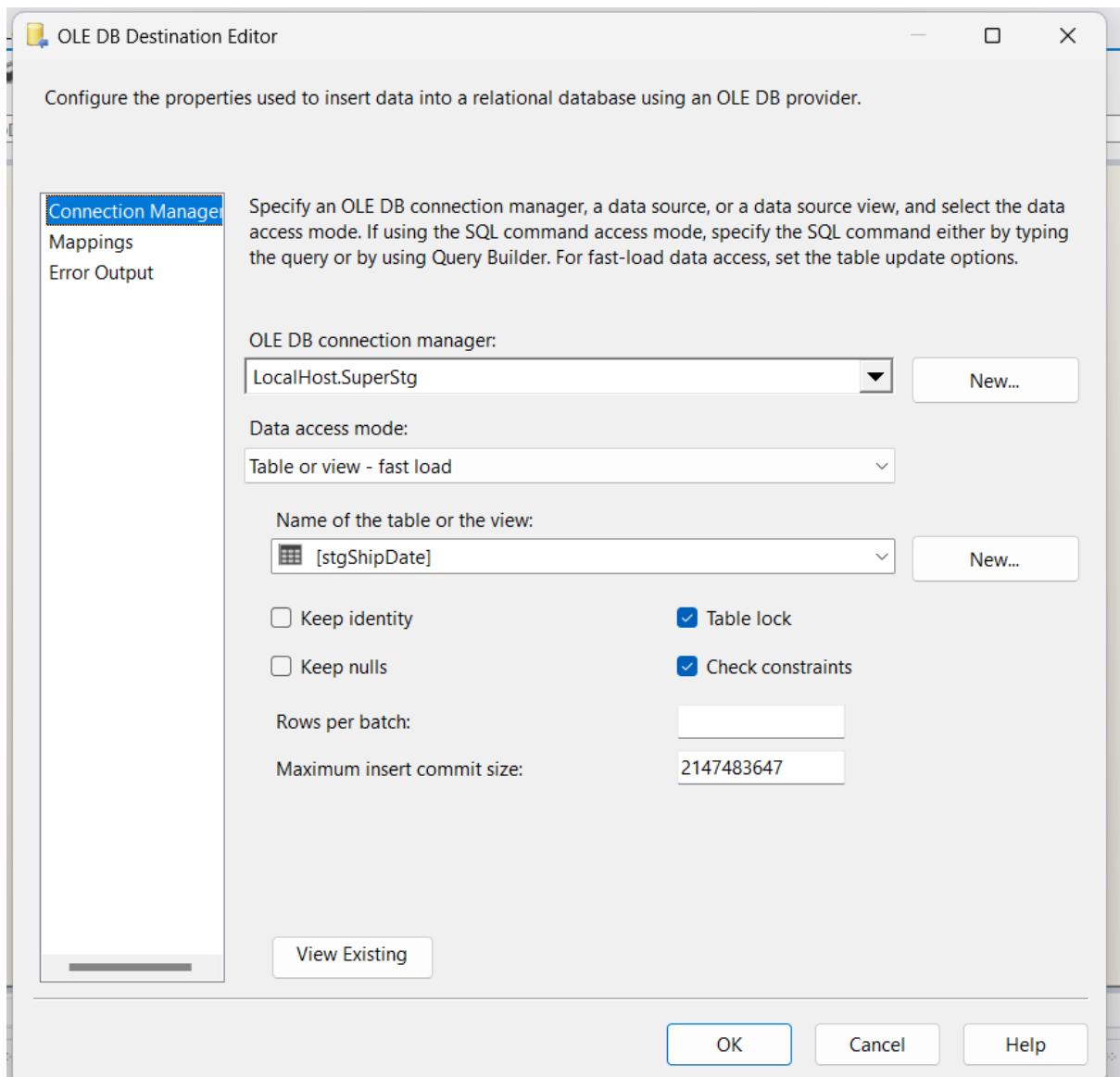
```

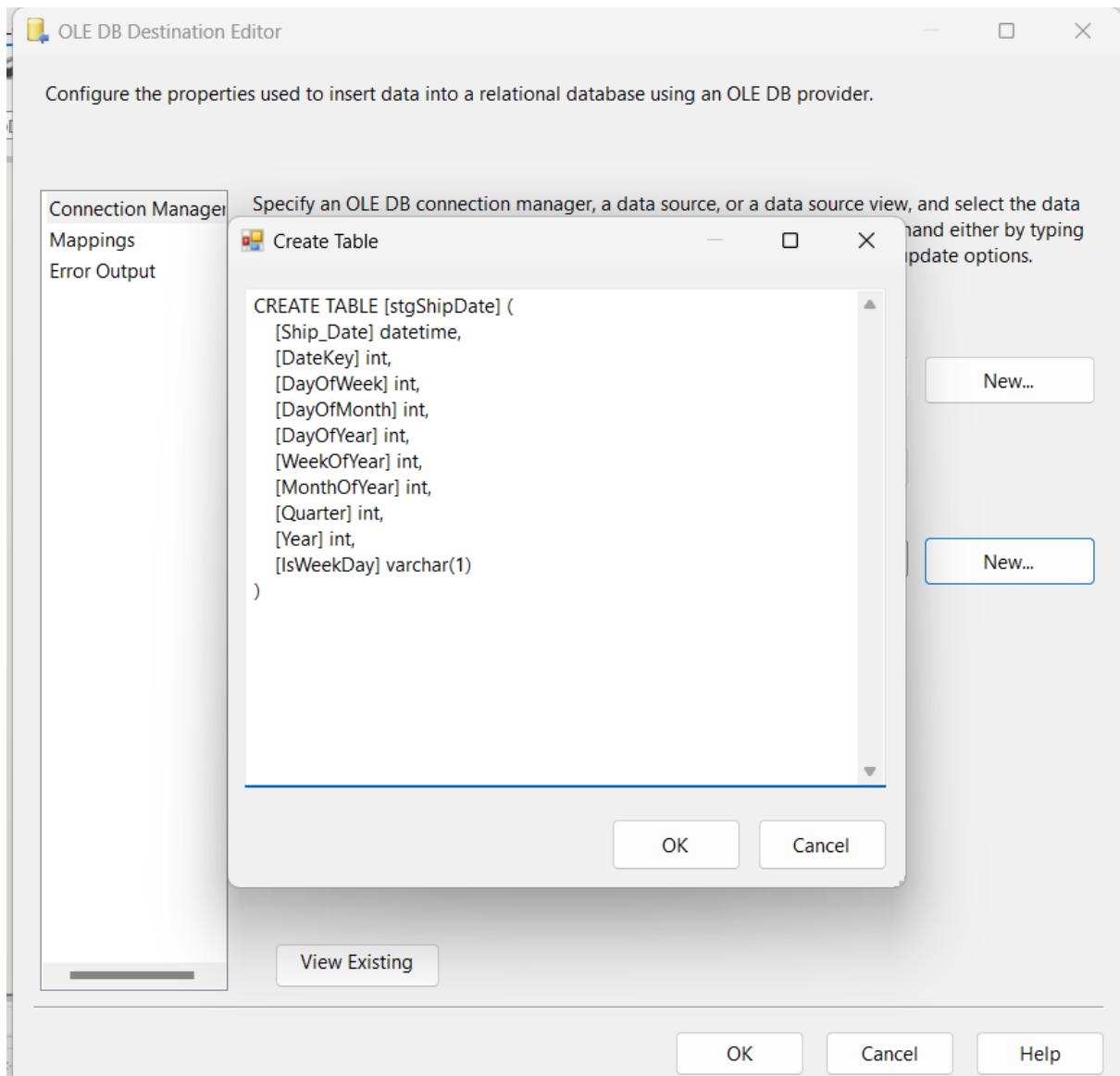
SELECT DISTINCT Ship_Date,
    (Day("Ship_Date") + MONTH("Ship_Date") * 100 + YEAR("Ship_Date") * 10000 ) as DateKey,
    DATEPART(WEEKDAY, Ship_Date) AS DayOfWeek,
    DATEPART(DAY, Ship_Date) AS DayOfMonth,
    DATEPART(DAYOFYEAR, Ship_Date) AS DayOfYear,
    DATEPART(WEEK, Ship_Date) AS WeekOfYear,
    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

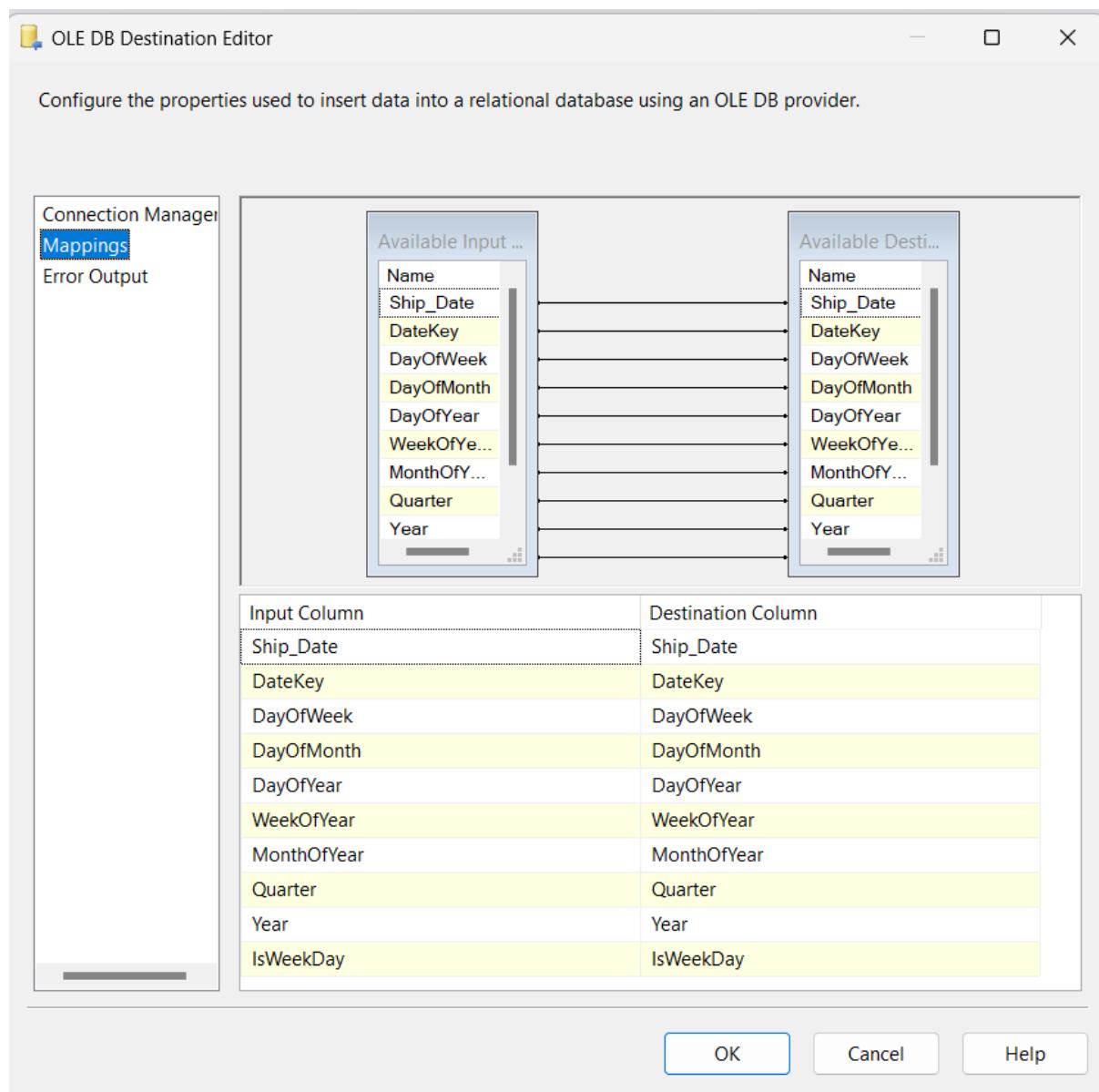
```

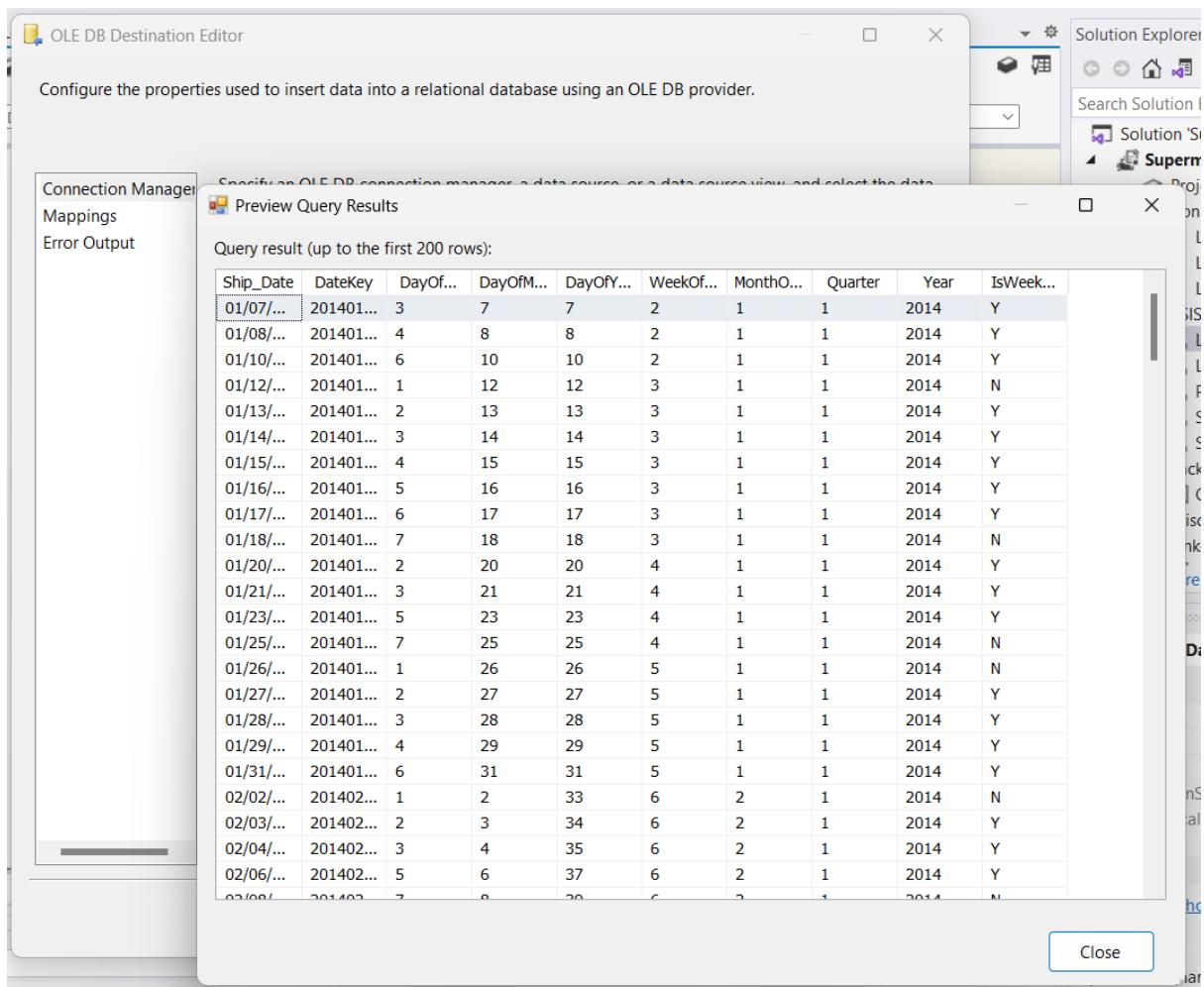
Ta sẽ lấy nguồn dữ liệu này để có thể tạo ra stgShipDate chứa thời gian thông kê kho và thời gian giao hàng

## DST – stgShipDate:

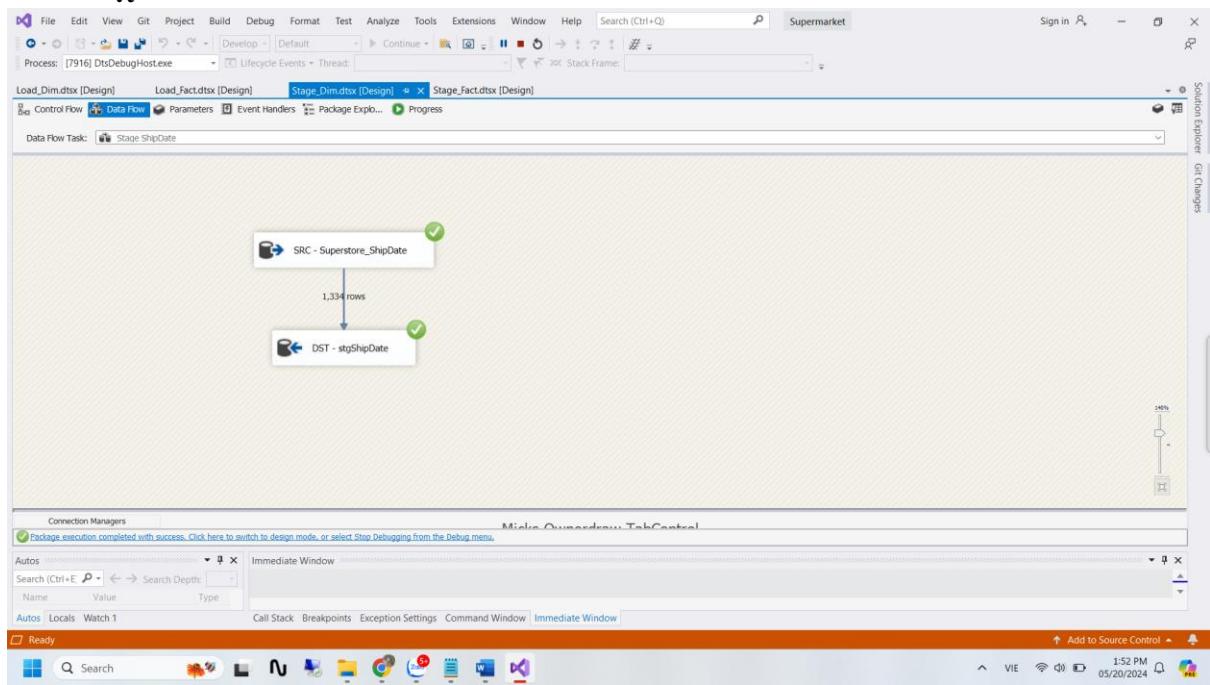




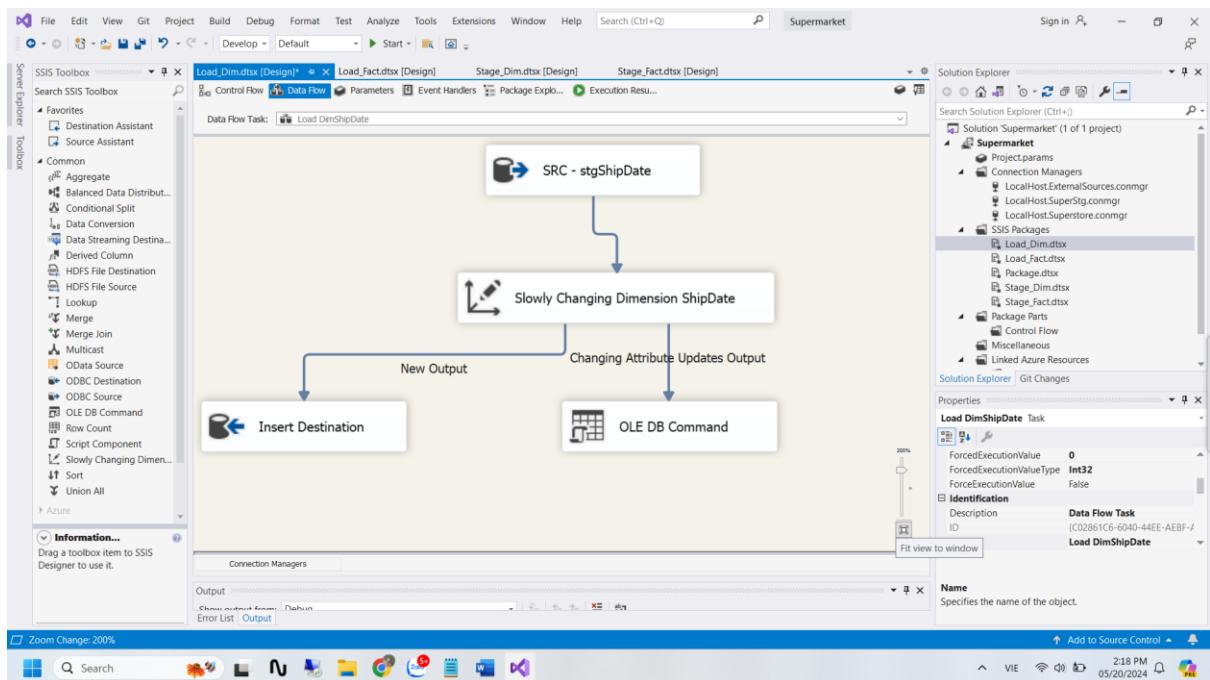




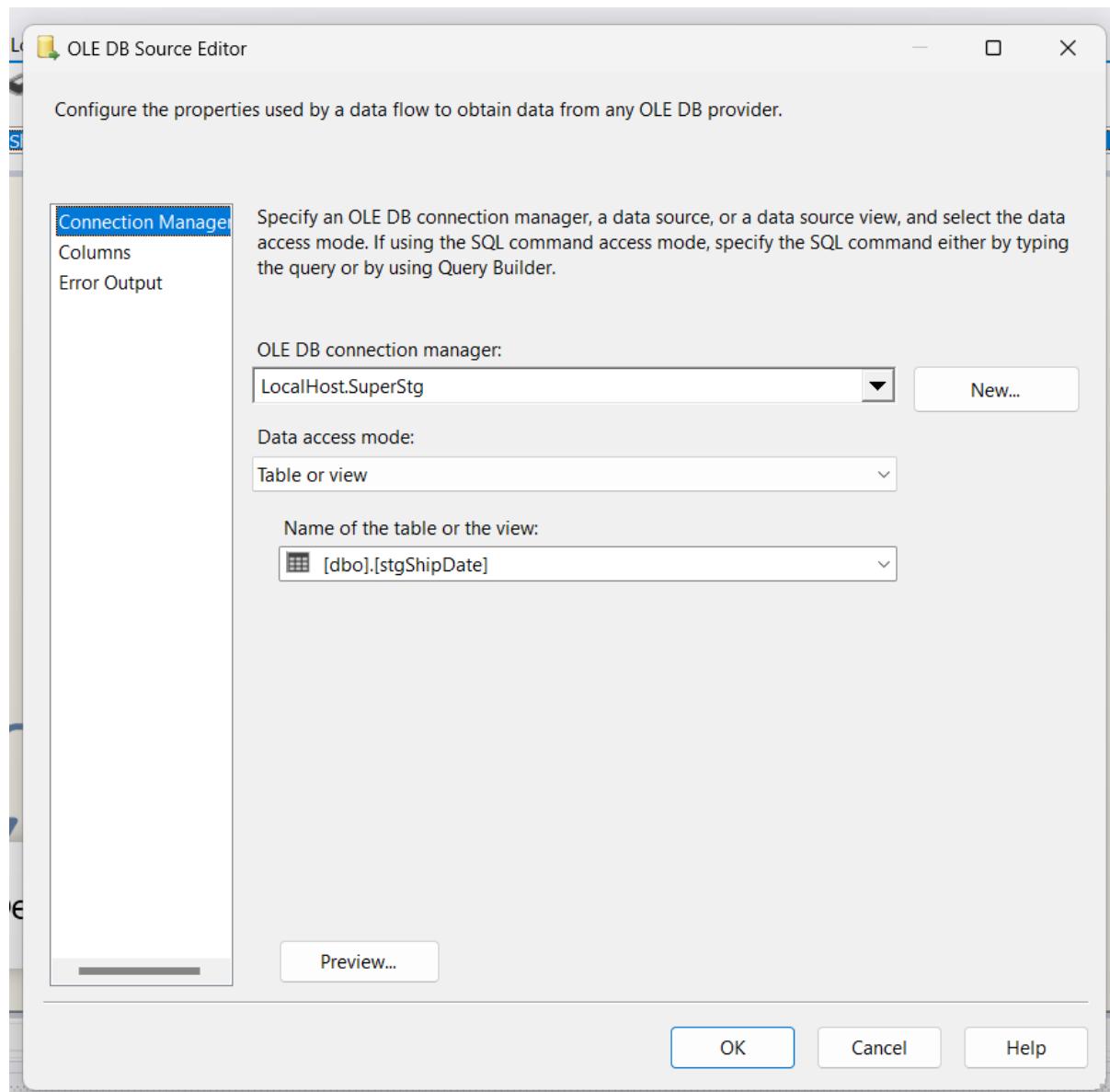
## Khởi chạy



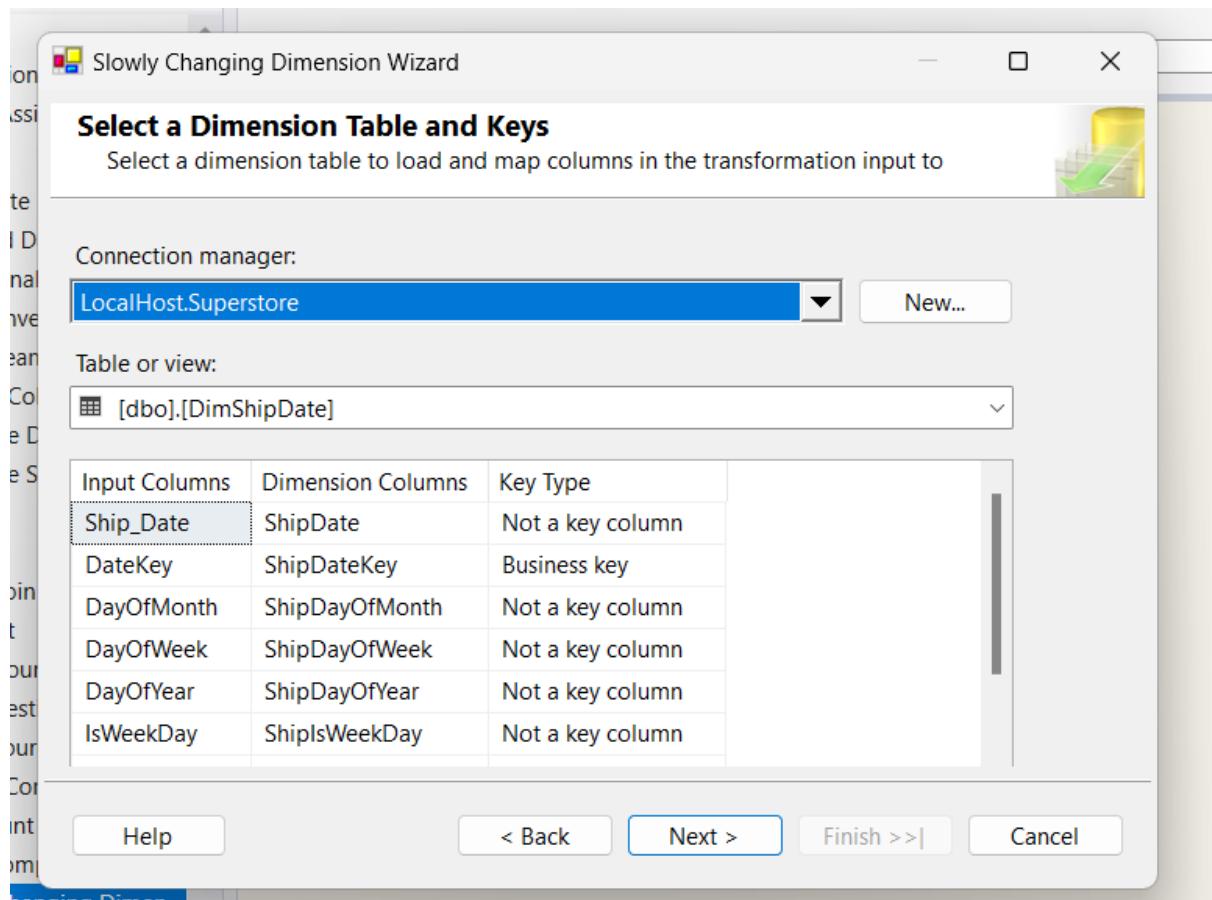
## b) Load dữ liệu từ stgShipDate vào DimDate



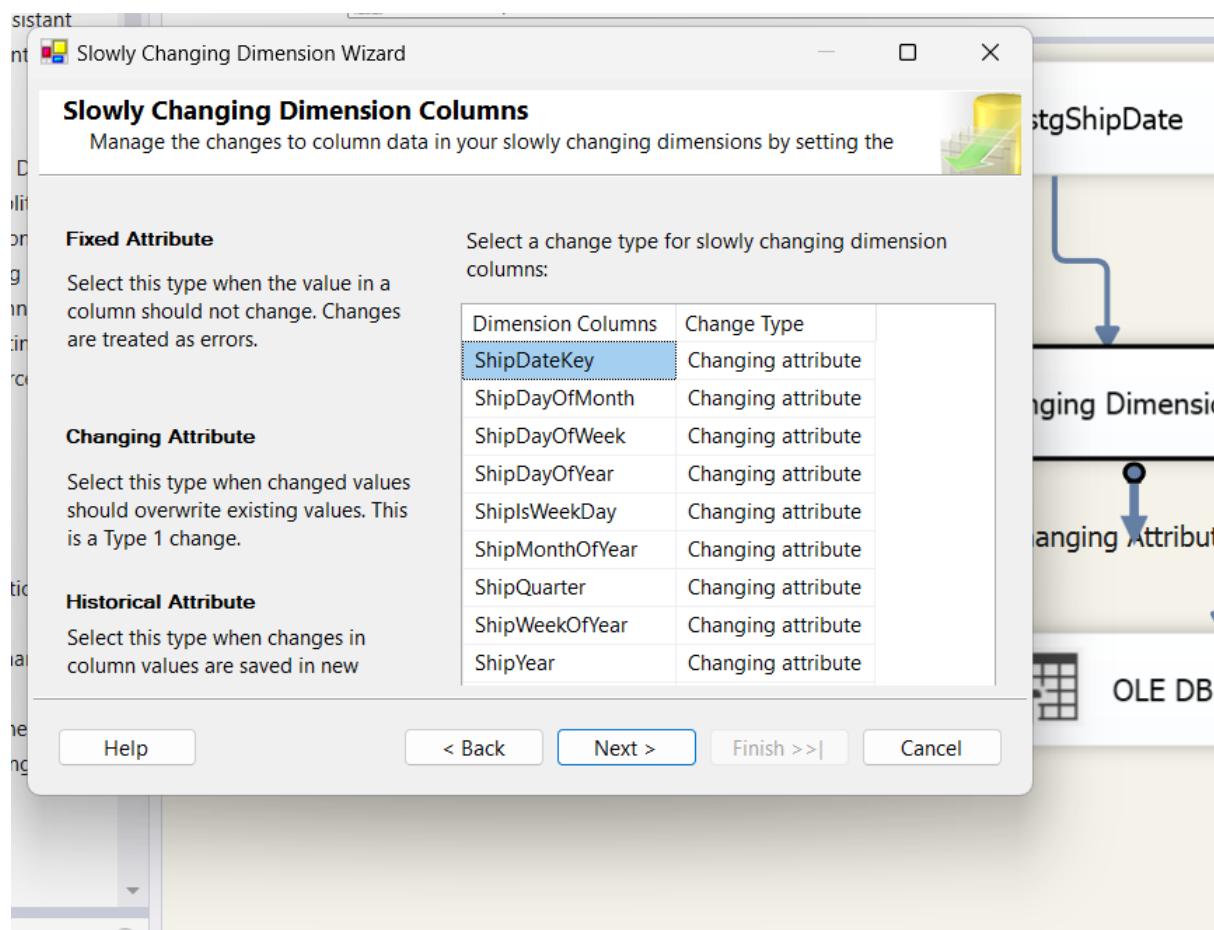
**SRC – stgShipDate** là dữ liệu từ bảng stgShipDate đã load ở trên



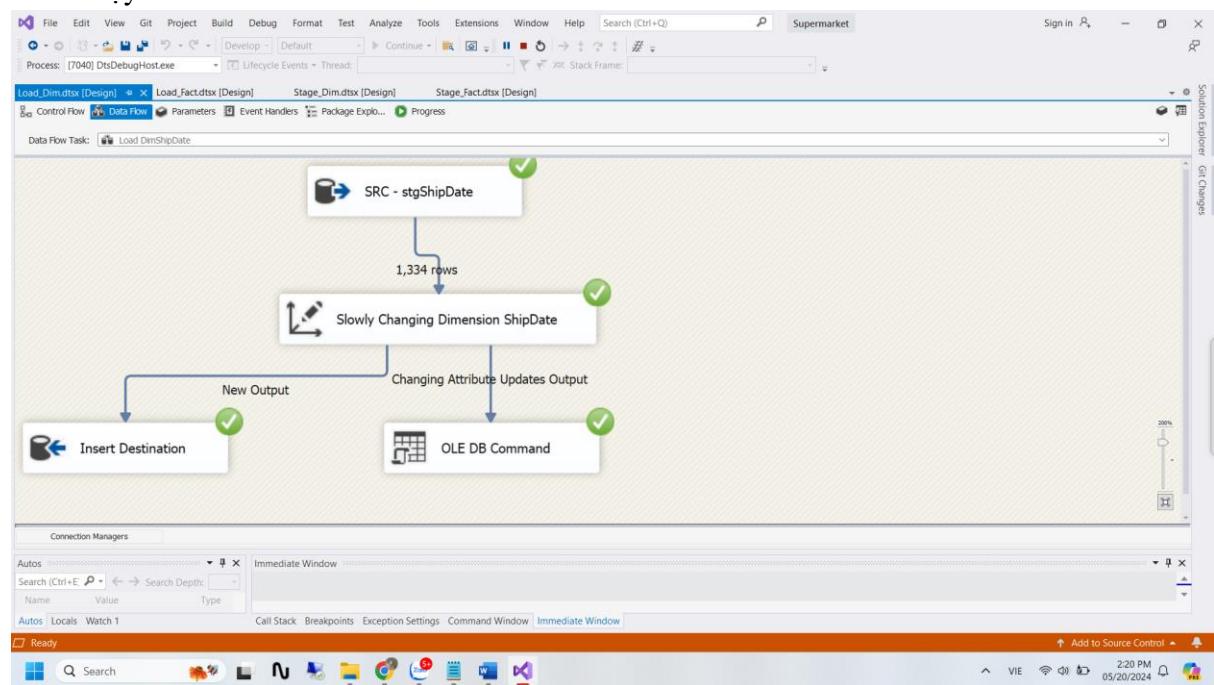
**Slowly Changing Dimension – DimShipDate** load dữ liệu từ stgShipDate vào DimShipDate



Chọn loại SCD là Changing attribute (Update dòng)

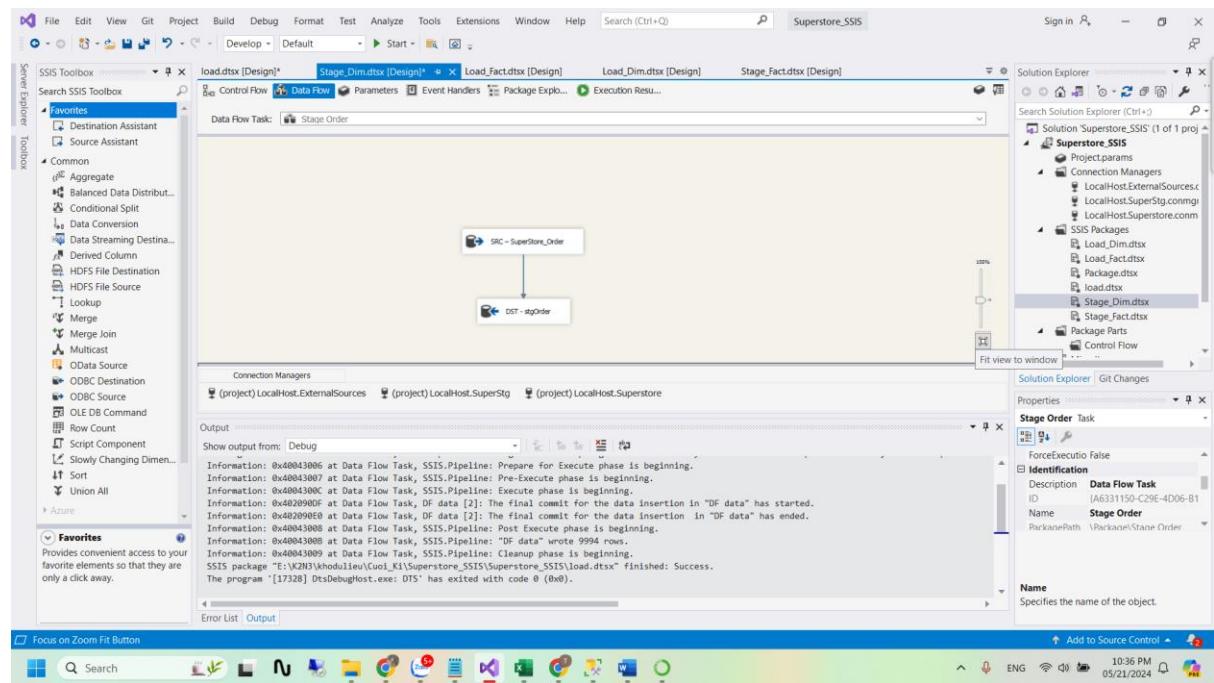


## Khởi chạy

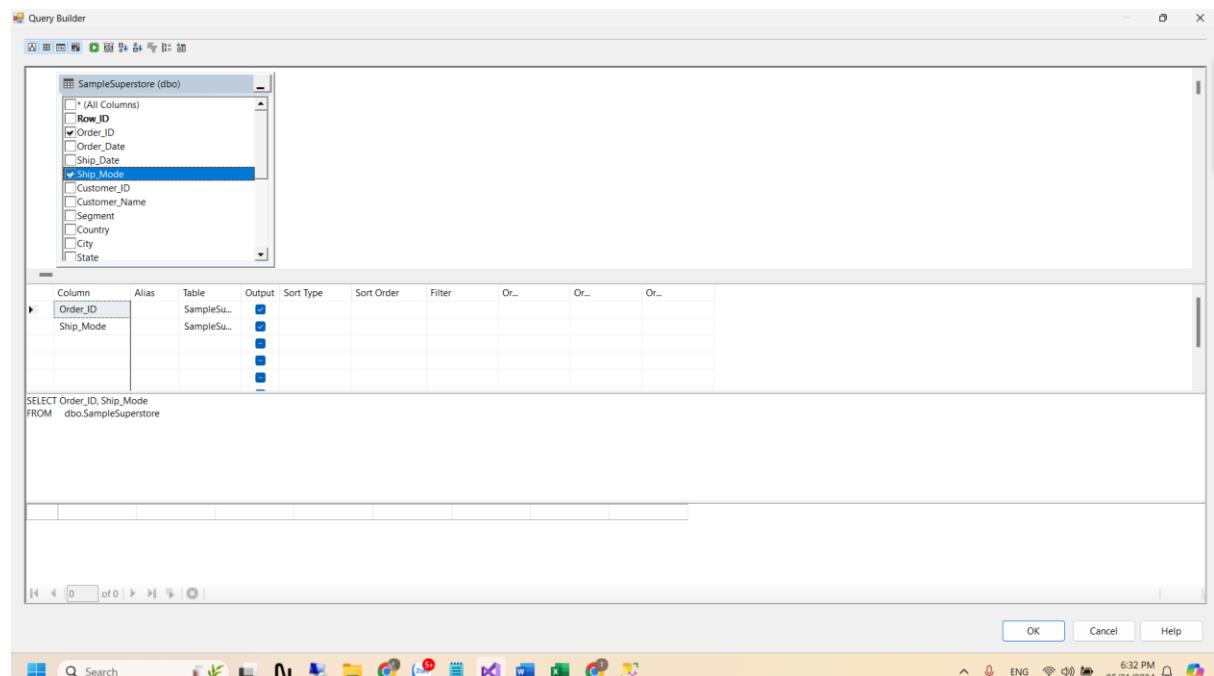


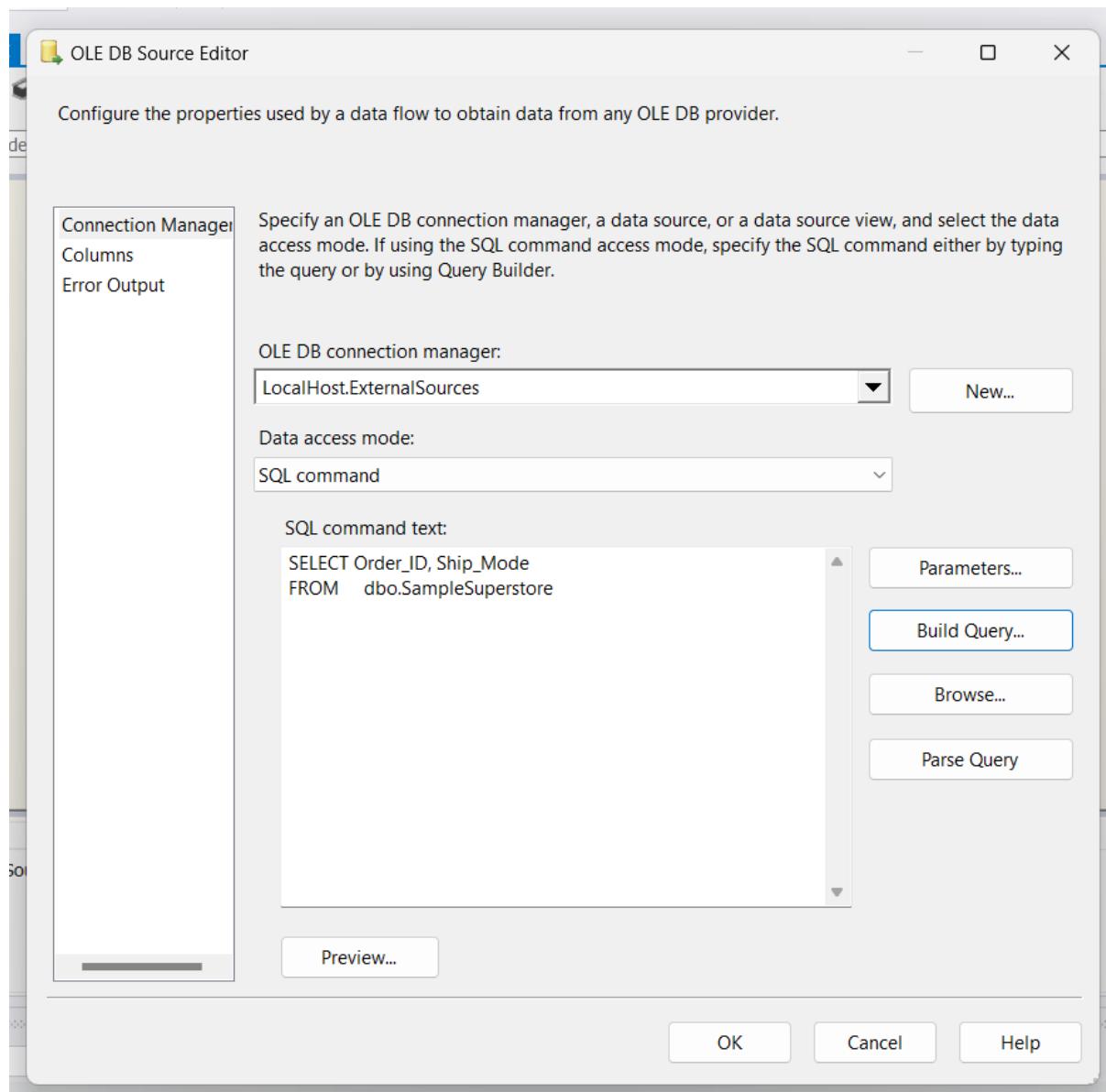
### 3.2.5. Order Dimension

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



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





## DST – stgProduct

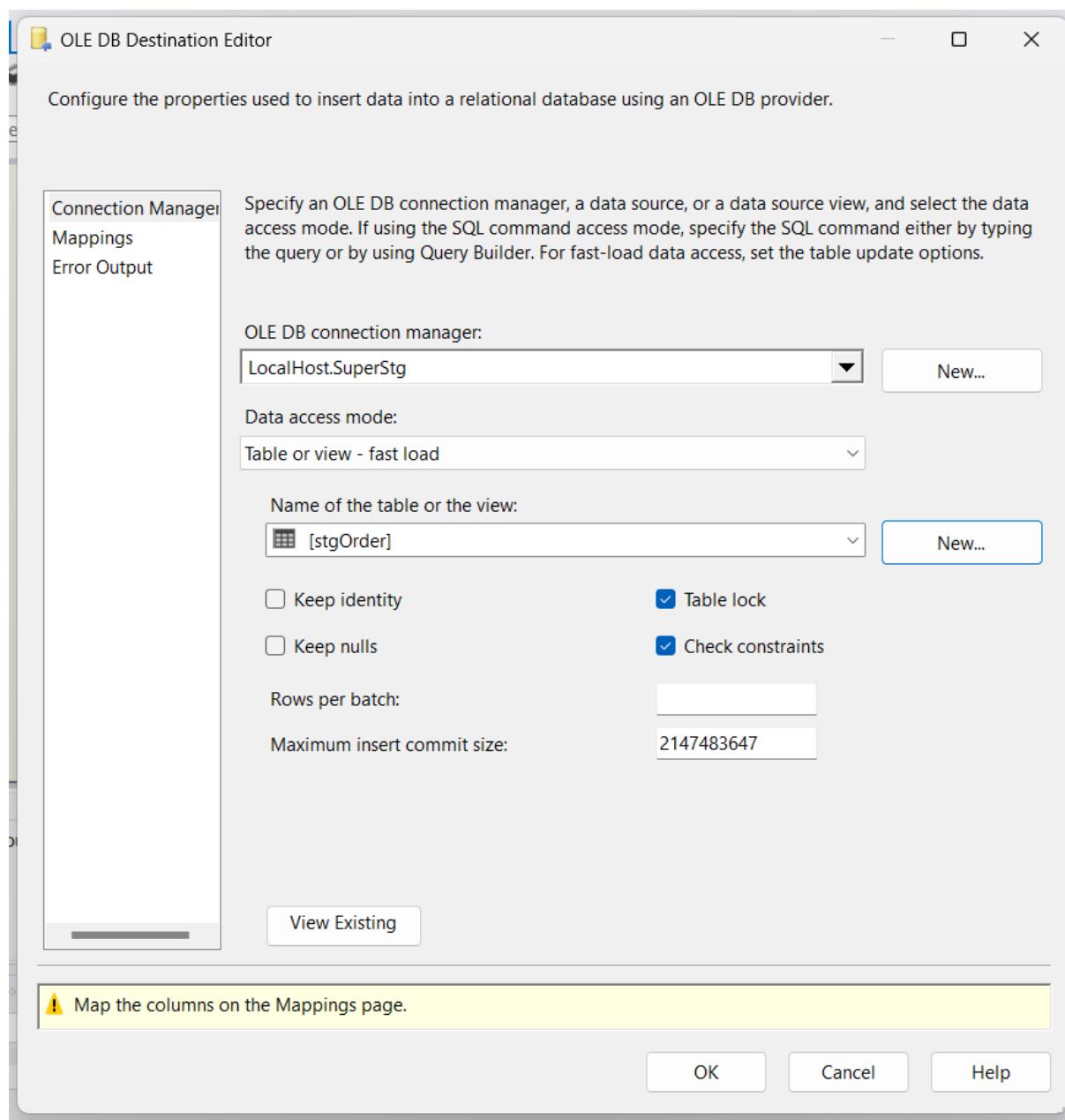
access mode. If using the SQL command access mode, specify the SQL command either in the SQL command text box or in the SQL command file box.

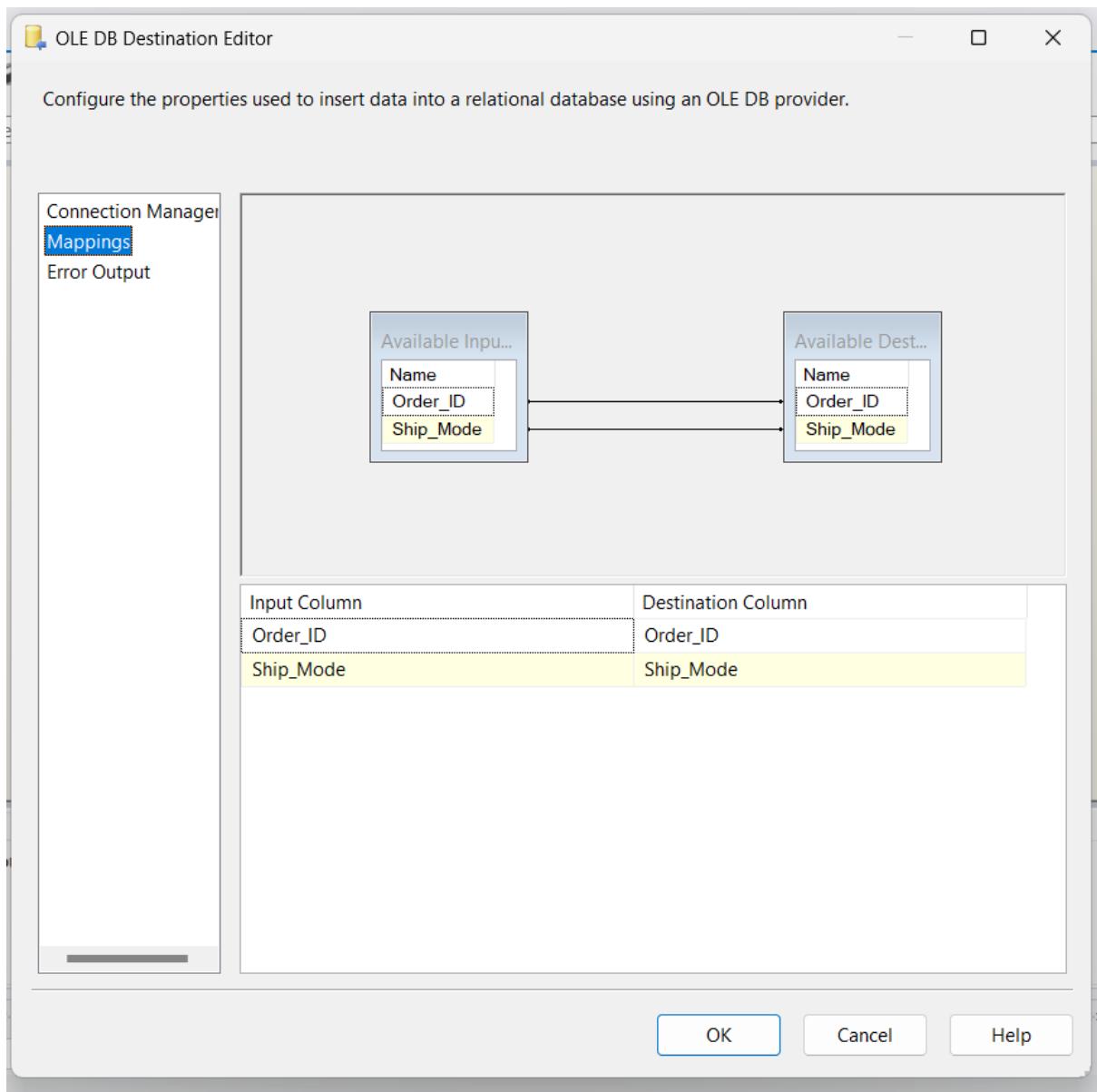
### Create Table

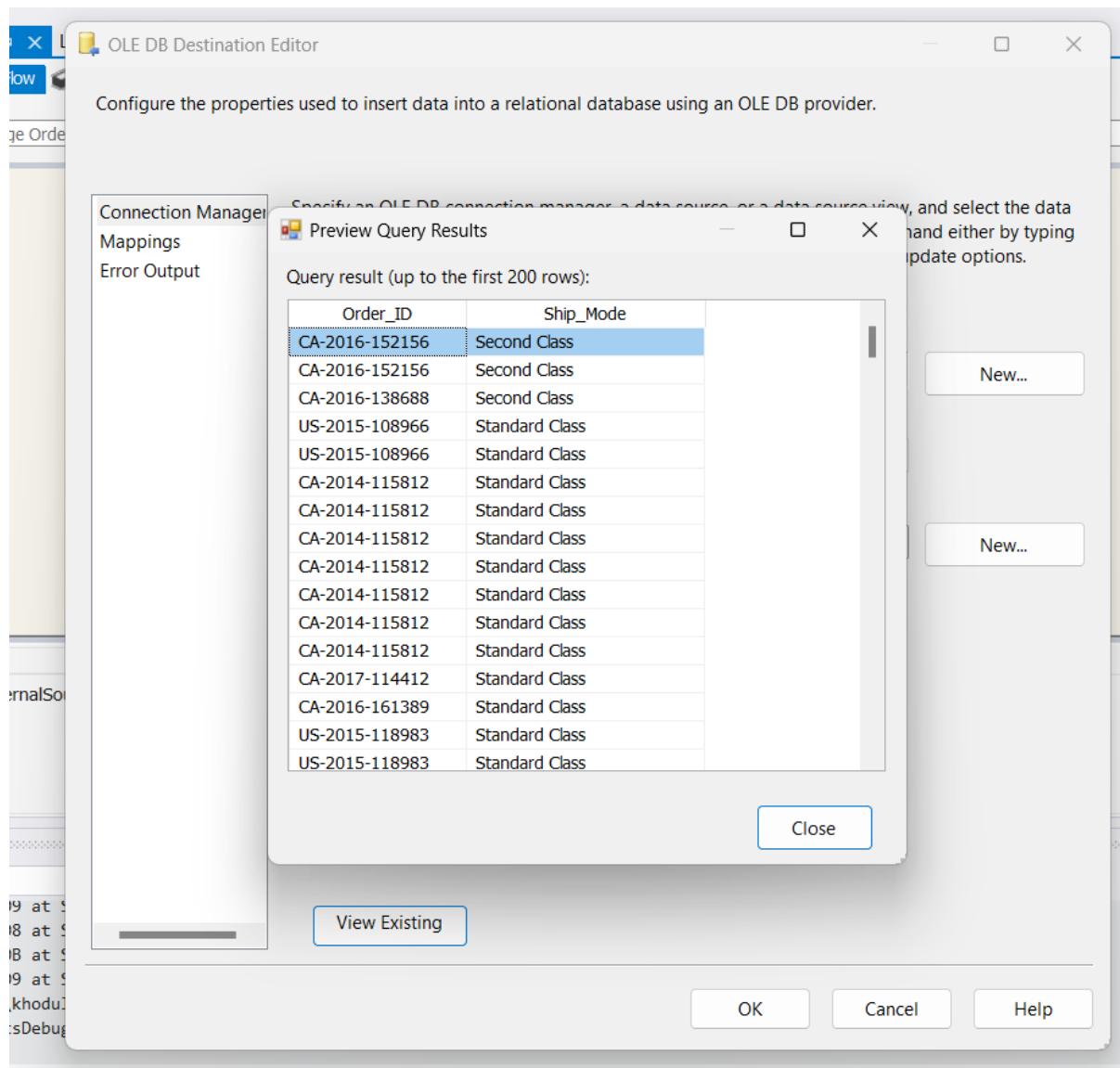
```
CREATE TABLE [stgOrder] (
    [Order_ID] nvarchar(30),
    [Ship_Mode] nvarchar(20)
)
```

OK

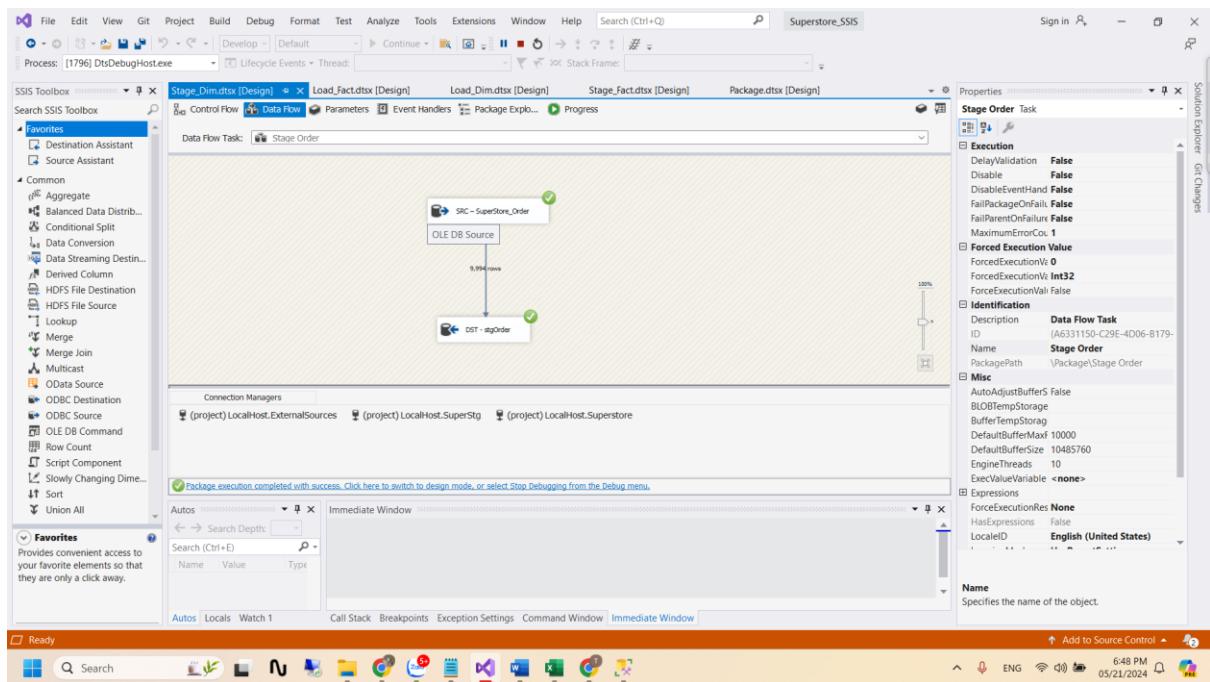
Cancel



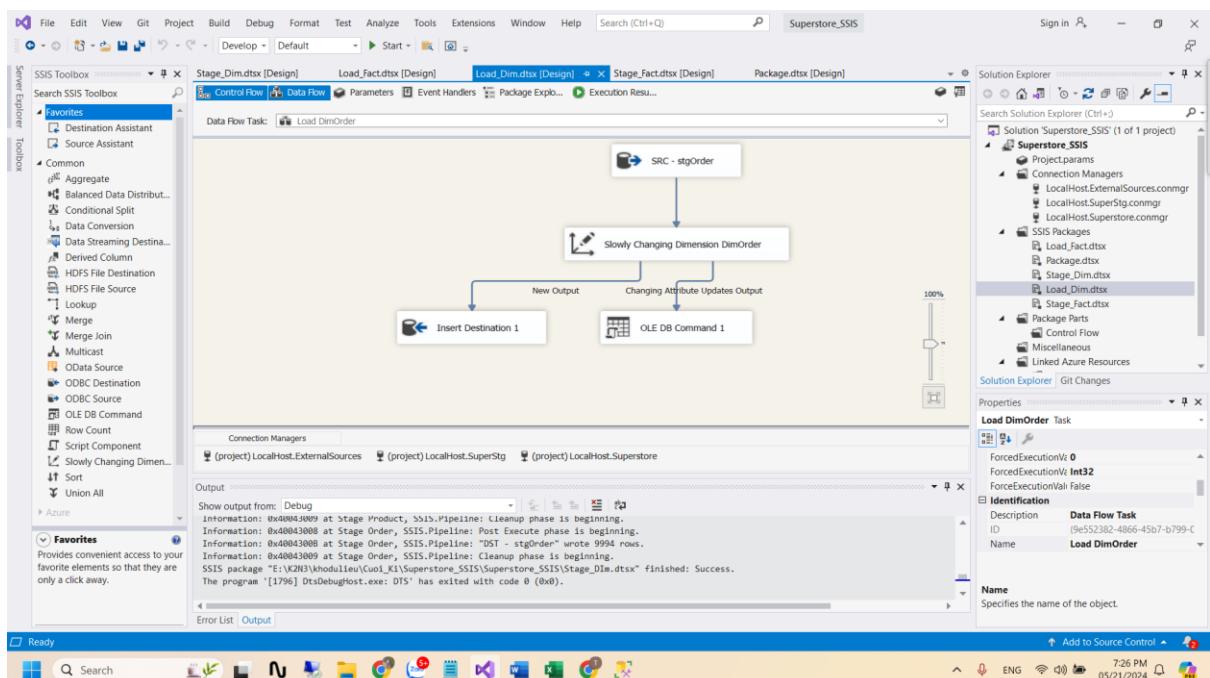




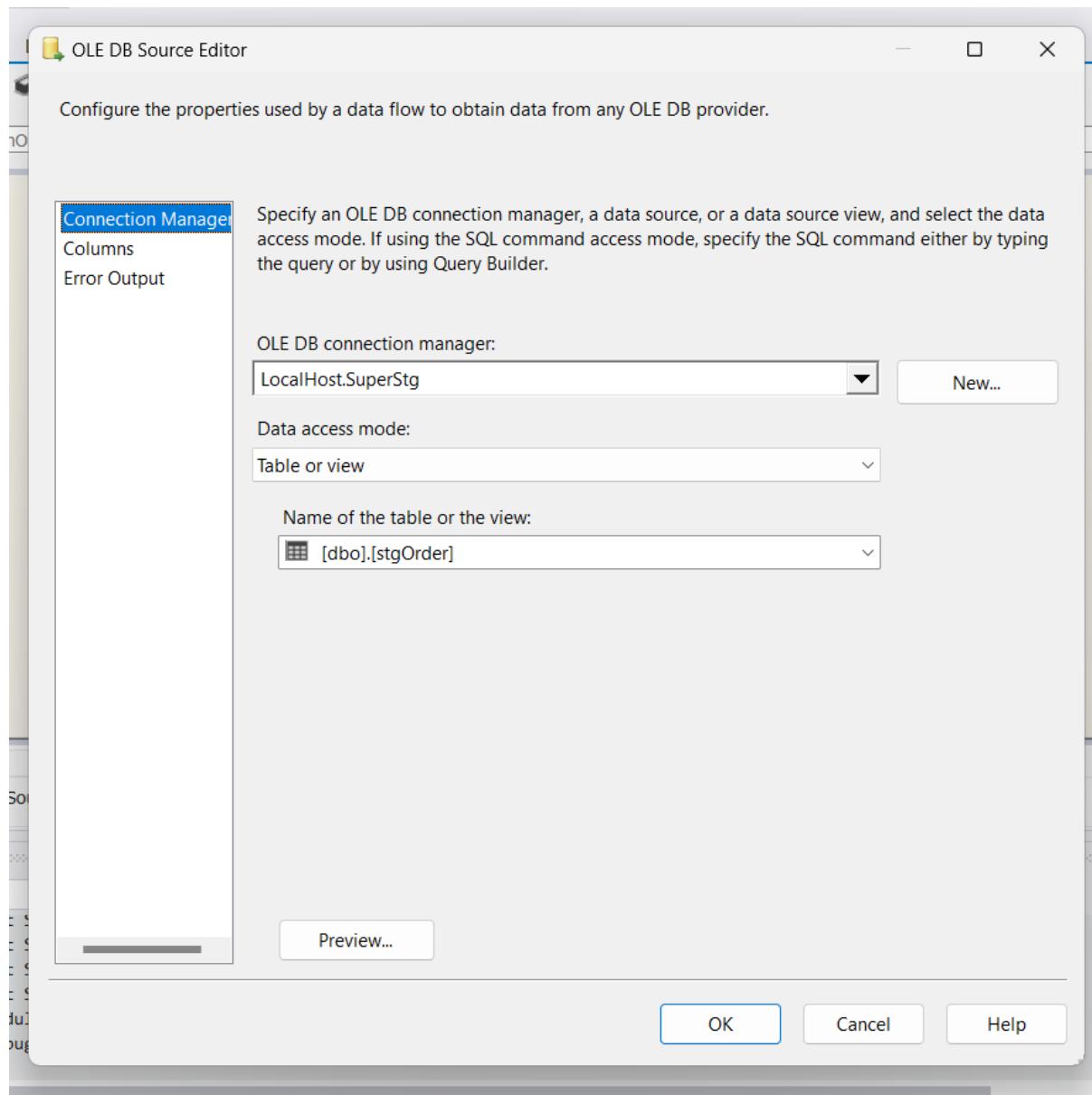
## **Khởi chạy**



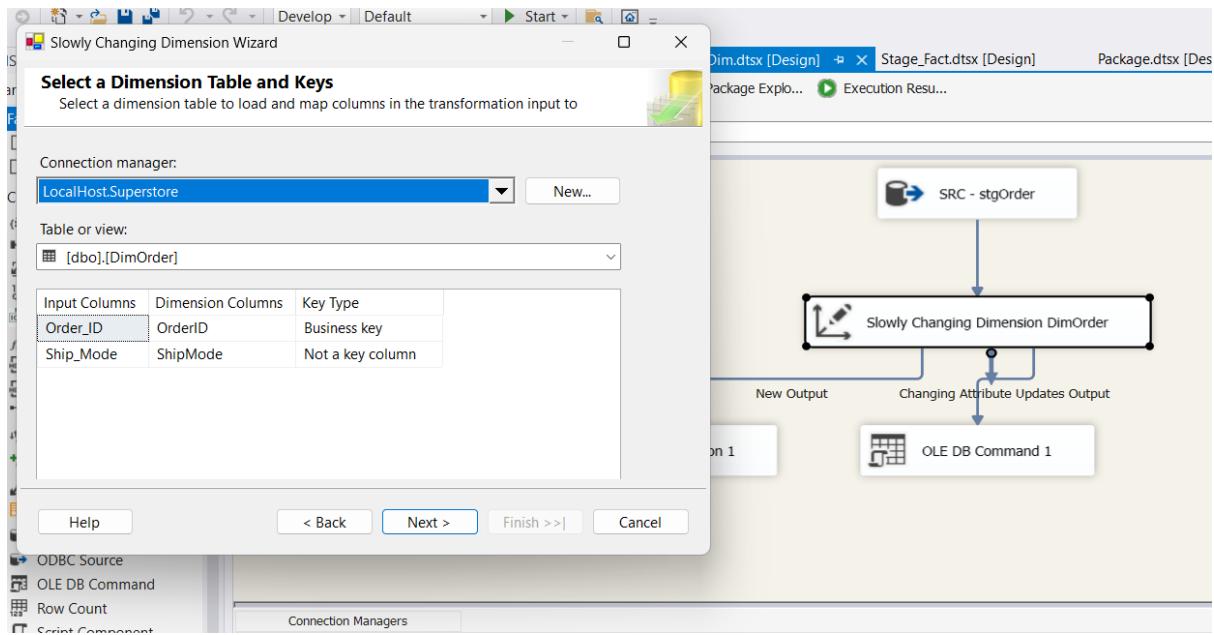
### b) Load dữ liệu từ stgOrder vào DimOrder



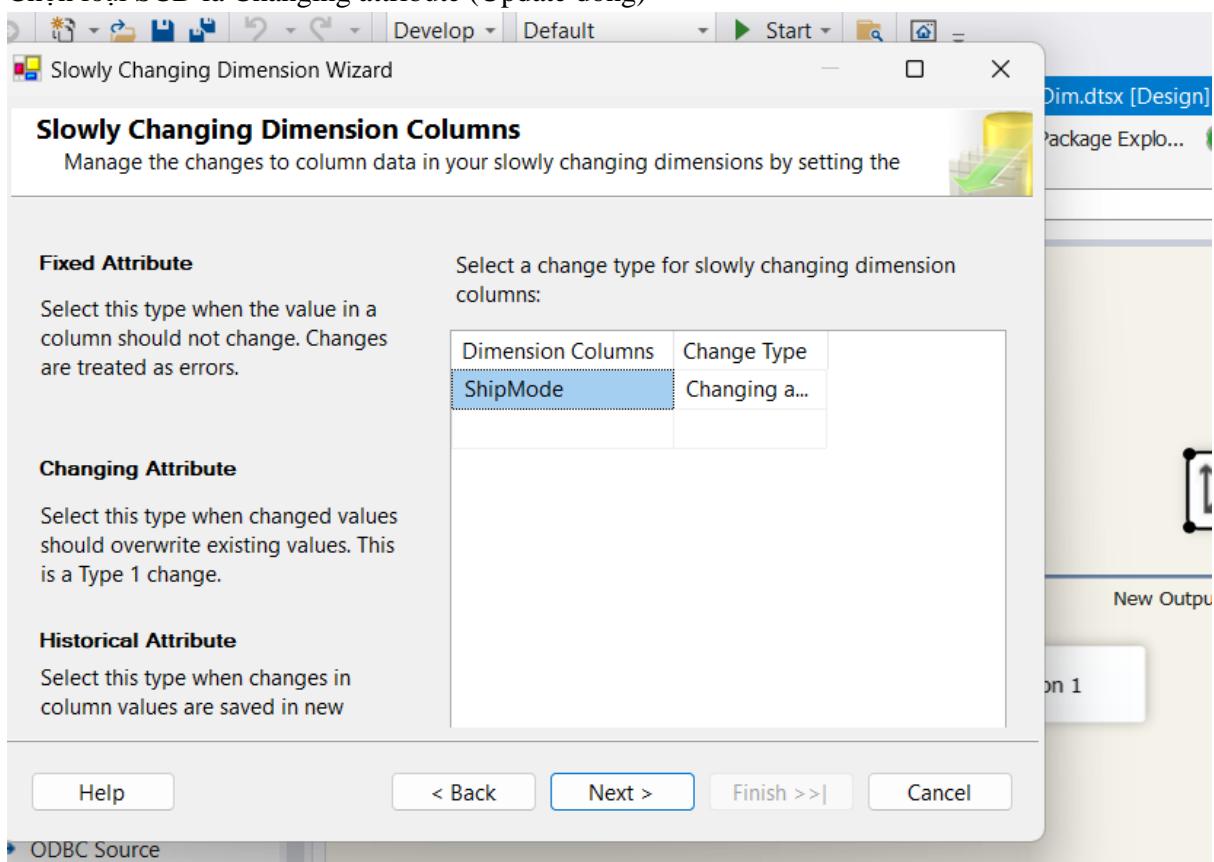
**SRC – stgOrder** là dữ liệu từ bảng stgOrder đã load ở trên



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

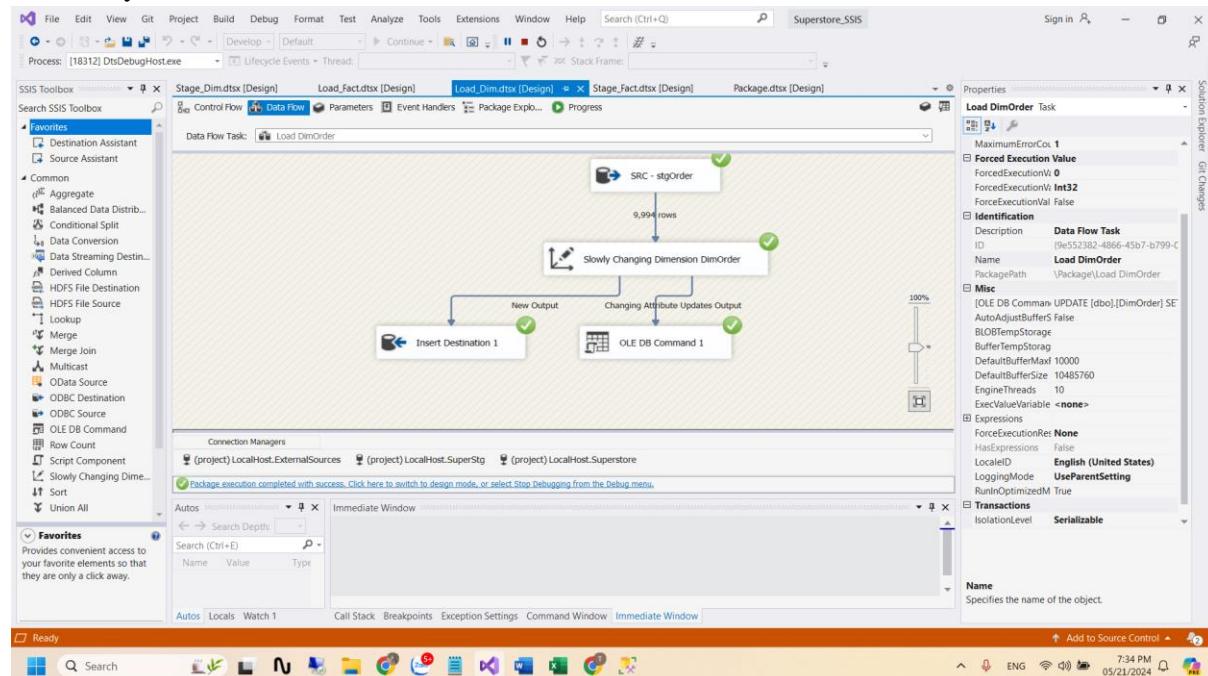


Chọn loại SCD là Changing attribute (Update dòng)



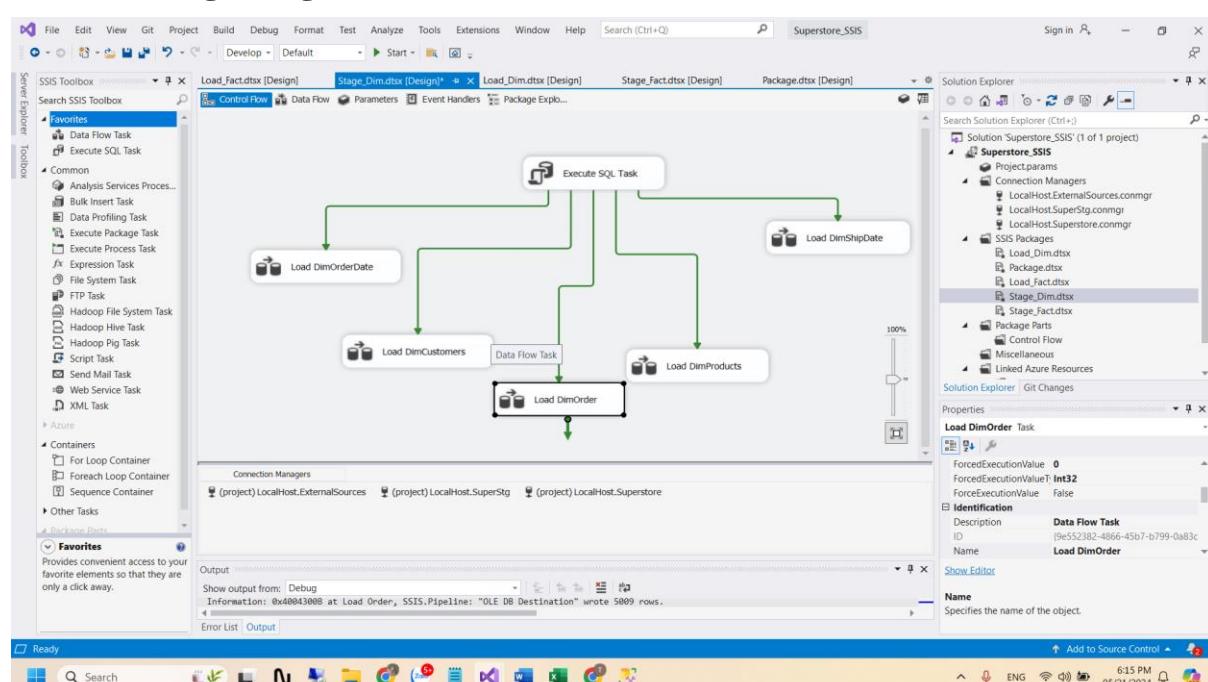
Khi thực hiện xong Slowly Changing Dimension – DimOrder thì Insert Destination và OLE DB Command sẽ xuất hiện.

Khởi chạy

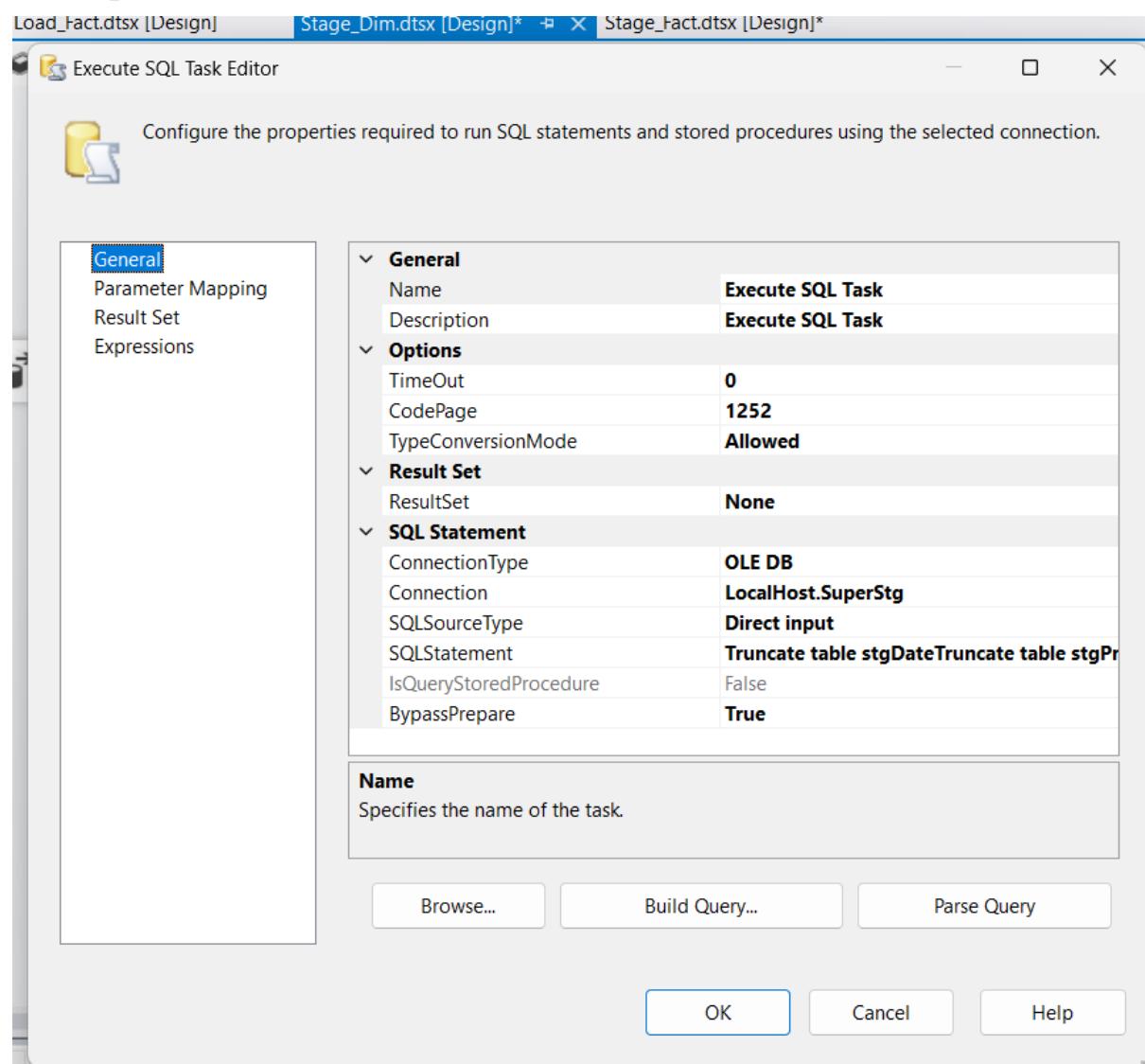


Khi chạy hoàn tất dữ liệu đã được Load vào DimOrder

### 3.2.6. Package Stage\_Dim



## a) Setup Execute SQL Task

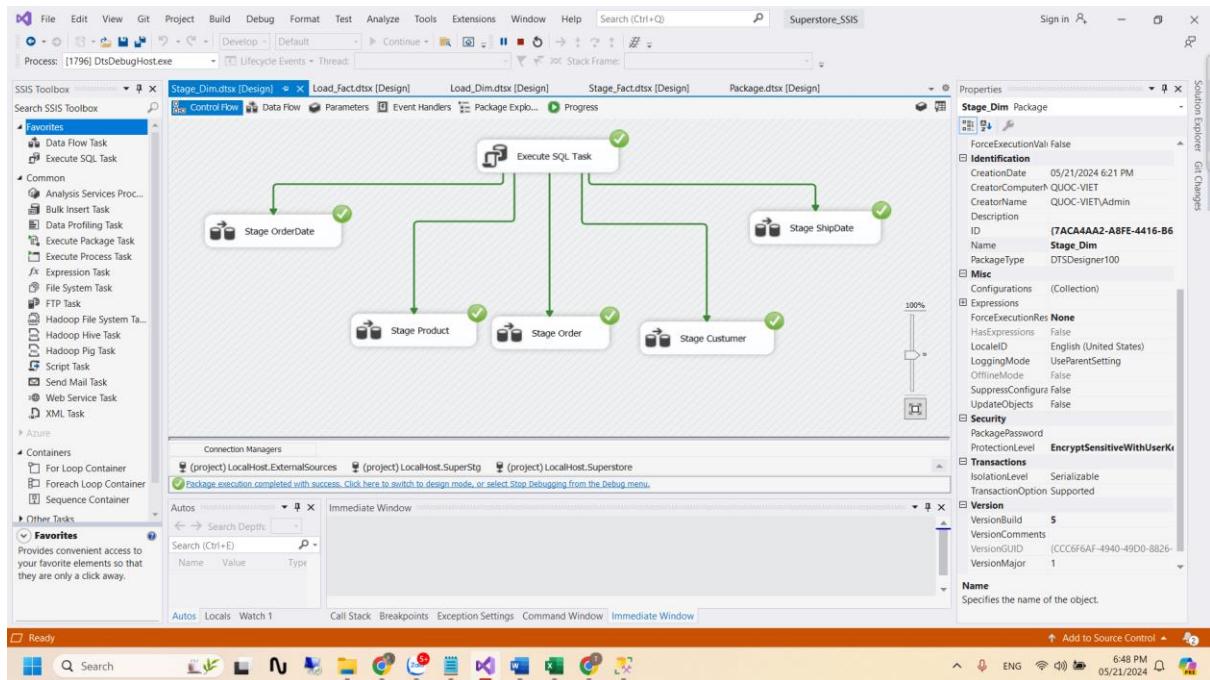


Connection đến database SuperStg

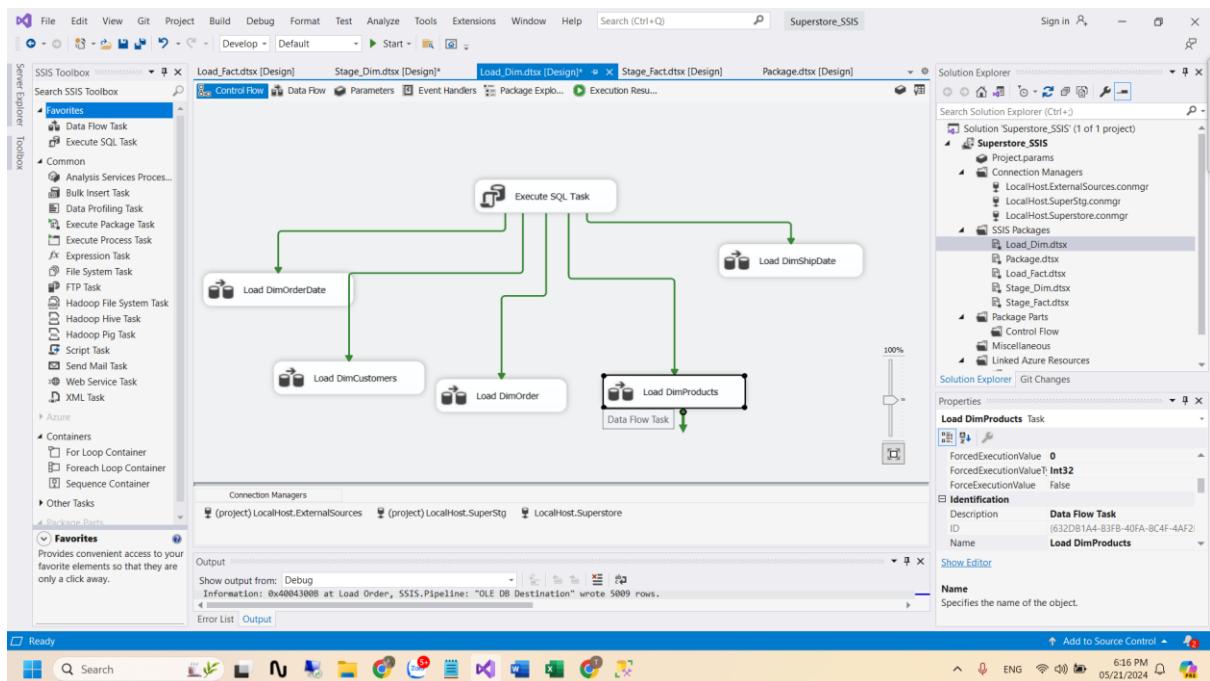
SQLStatement :

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

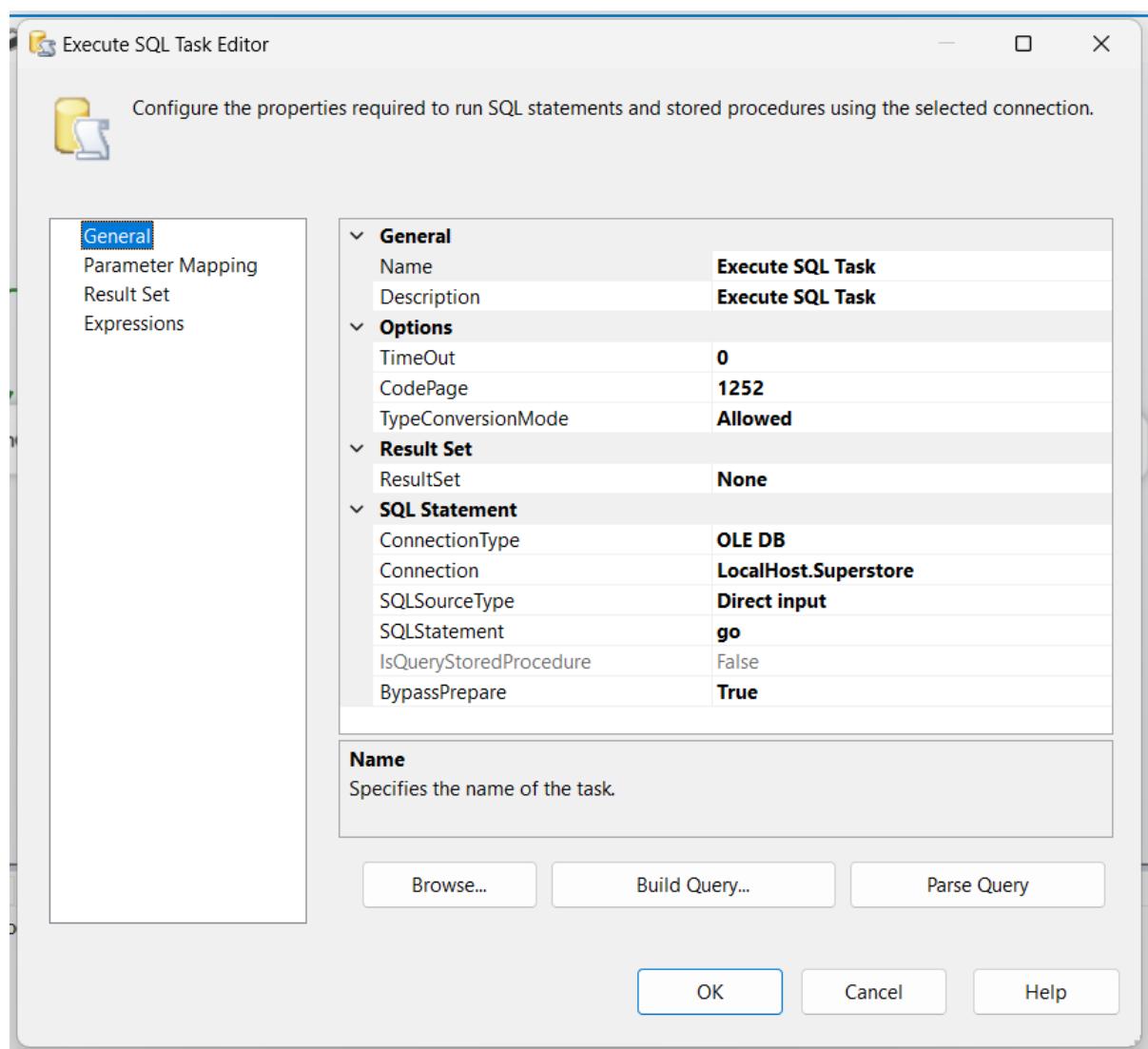
**Khởi chạy:**



### 3.2.7. Package Load\_Dim



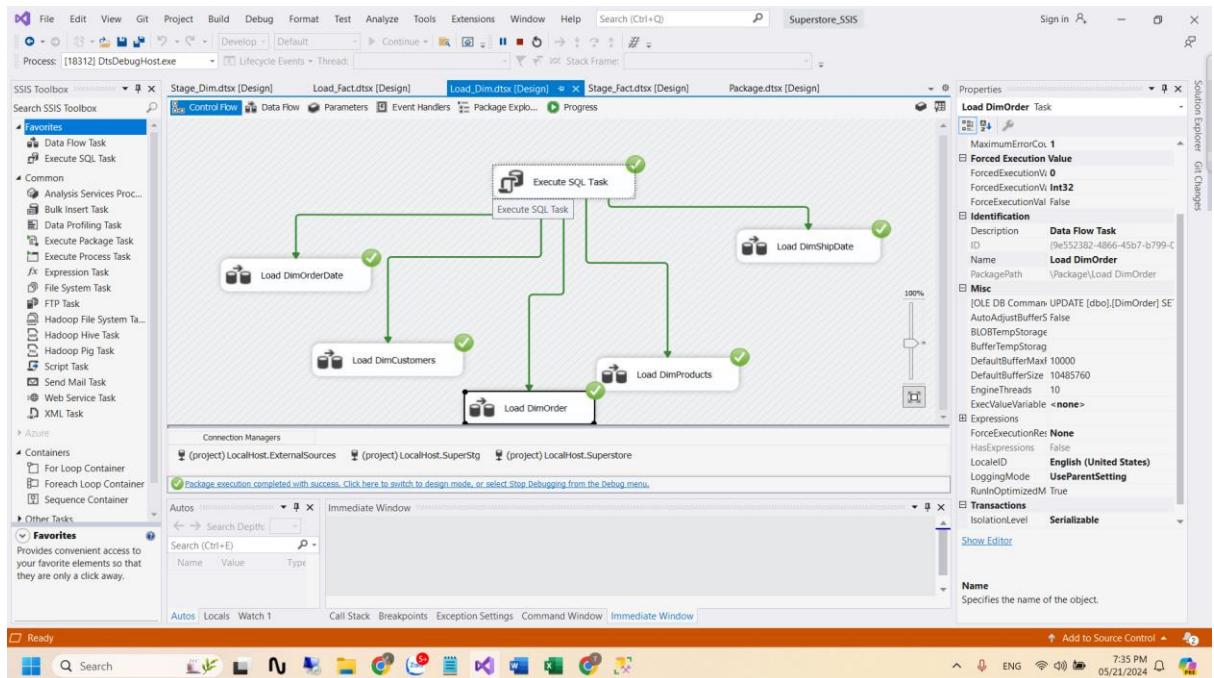
### a) Setup Execute SQL Task



Connection : Kết nối đến database Superstore

SQLStatement : go

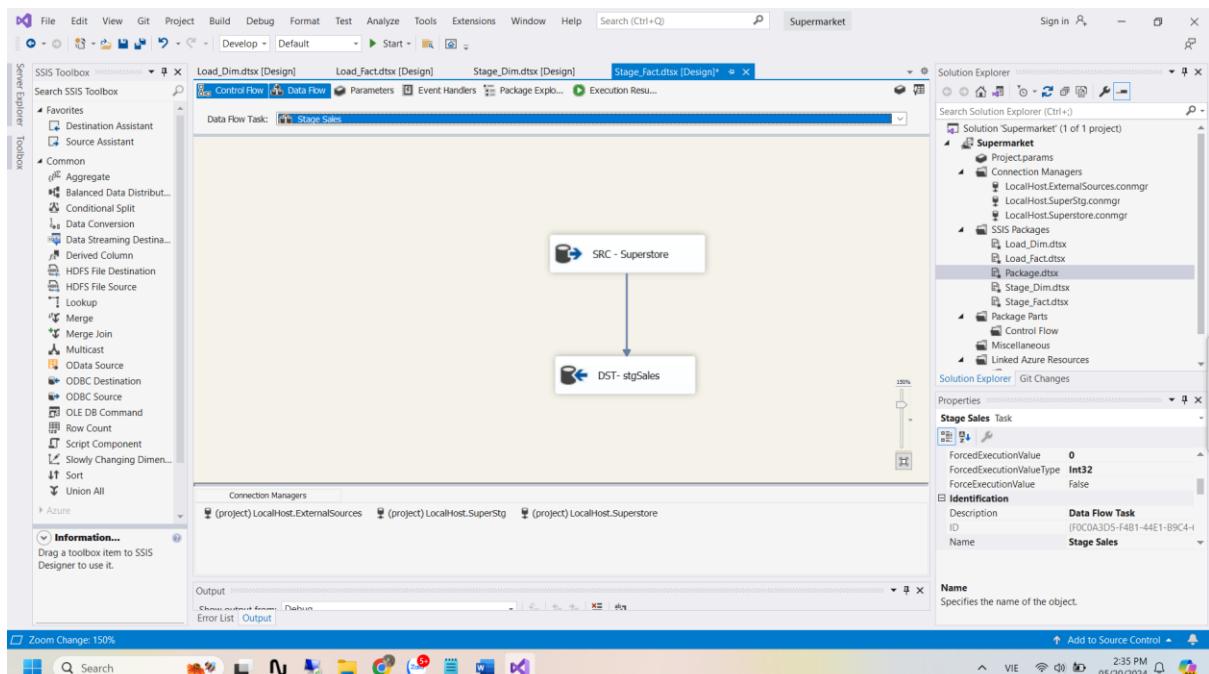
**Khởi chạy**



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

#### 3.3.1. Fact Sales

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



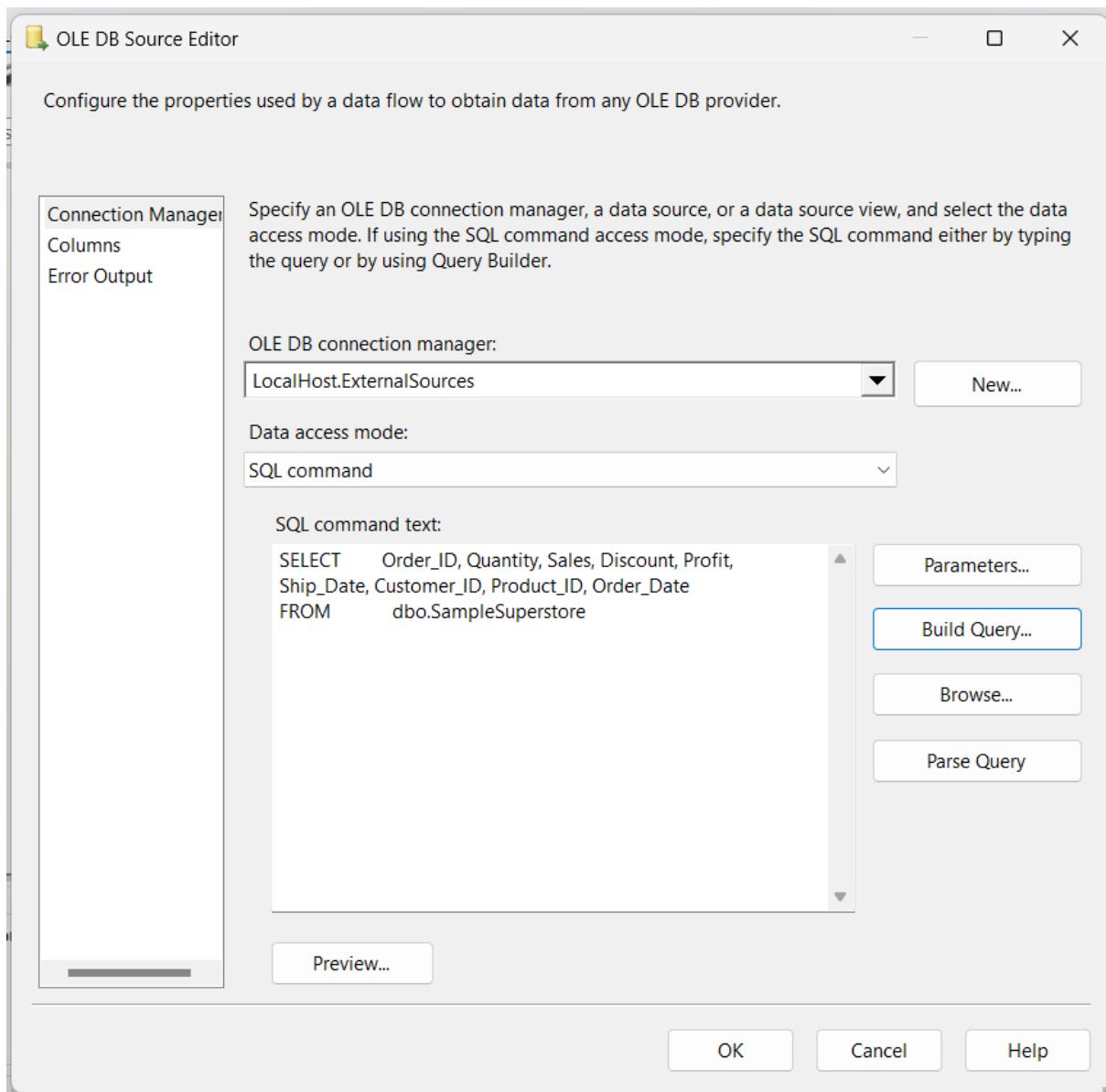
SRC – SuperStore chứa dữ liệu cần thiết để đưa vào bảng stgSales

The screenshot shows the Query Builder interface. The top pane displays the "SampleSuperstore (dbo)" table with a list of columns: 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. The bottom pane shows a grid for selecting columns for the output:

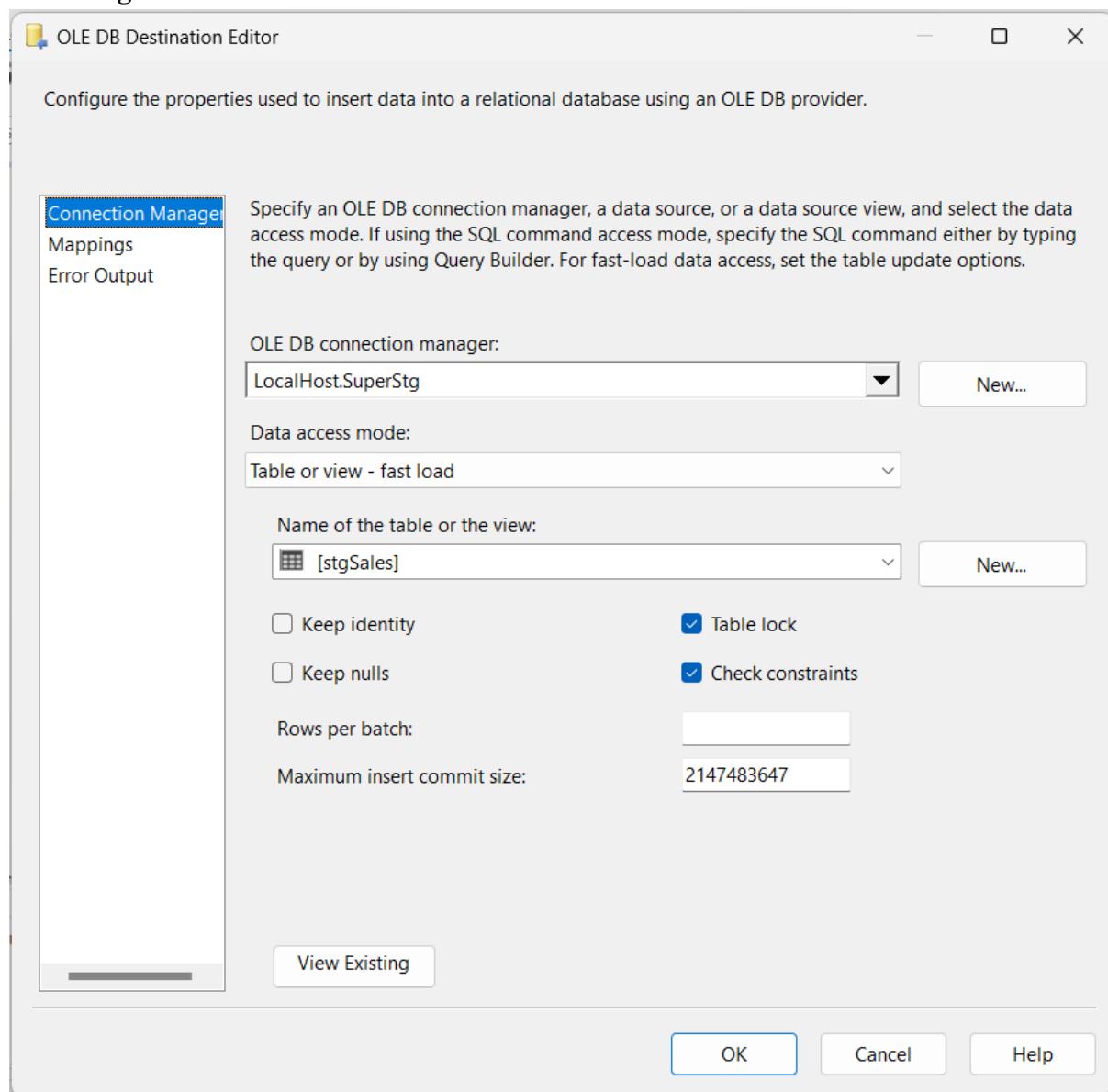
Column	Alias	Table	Output	Sort Type	Sort Order	Filter	Or...	Or...	Or...
Order_ID	Quantity	SampleSu...	<input checked="" type="checkbox"/>						
Quantity	Sales	SampleSu...	<input checked="" type="checkbox"/>						
Sales	Discount	SampleSu...	<input checked="" type="checkbox"/>						
Discount	Profit	SampleSu...	<input checked="" type="checkbox"/>						
Profit	Ship_Date	SampleSu...	<input checked="" type="checkbox"/>						
Ship_Date	Customer_ID	SampleSu...	<input checked="" type="checkbox"/>						
Customer_ID	Product_ID	SampleSu...	<input checked="" type="checkbox"/>						
Product_ID	Order_Date	SampleSu...	<input checked="" type="checkbox"/>						

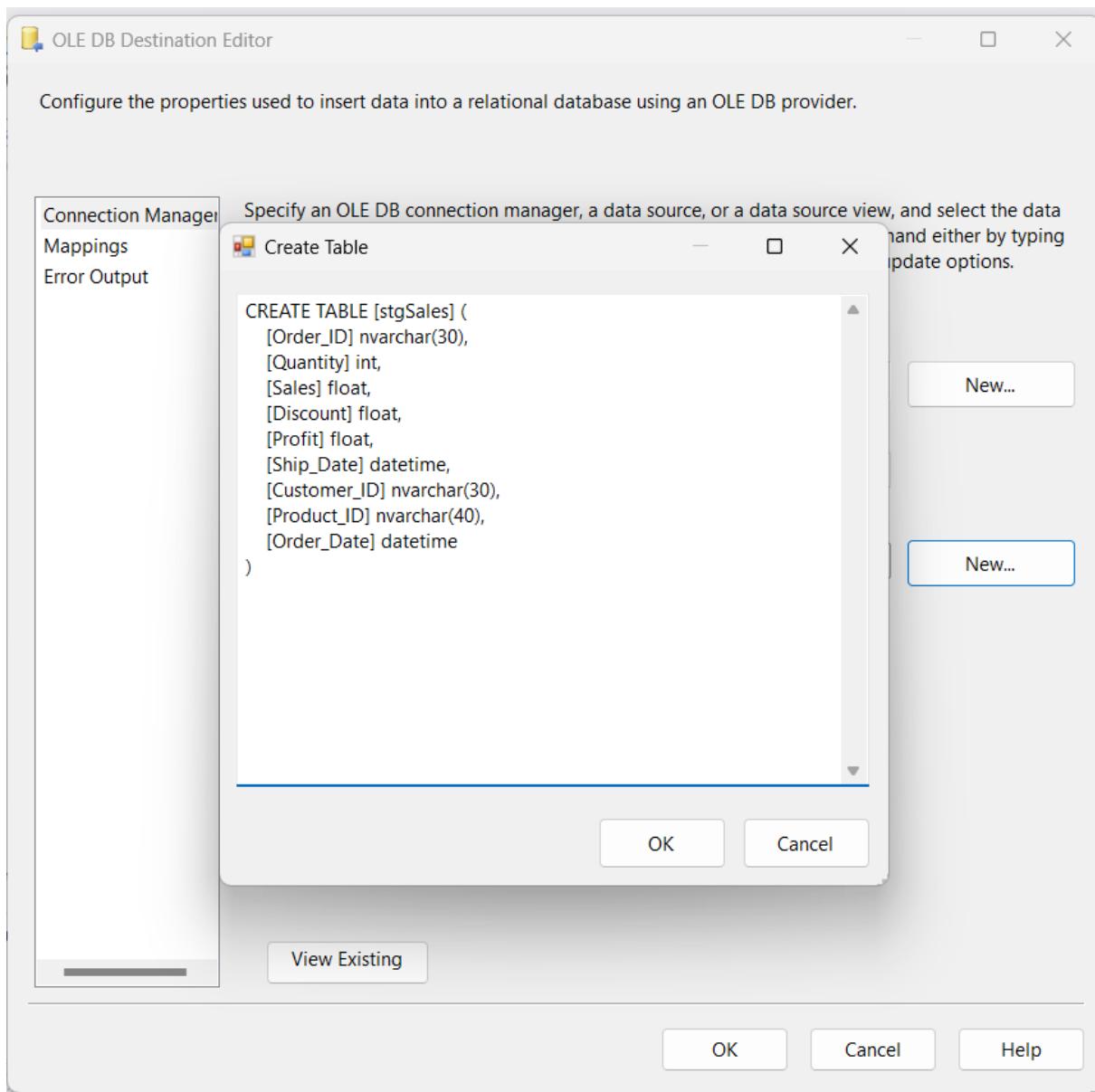
The SQL query at the bottom is:

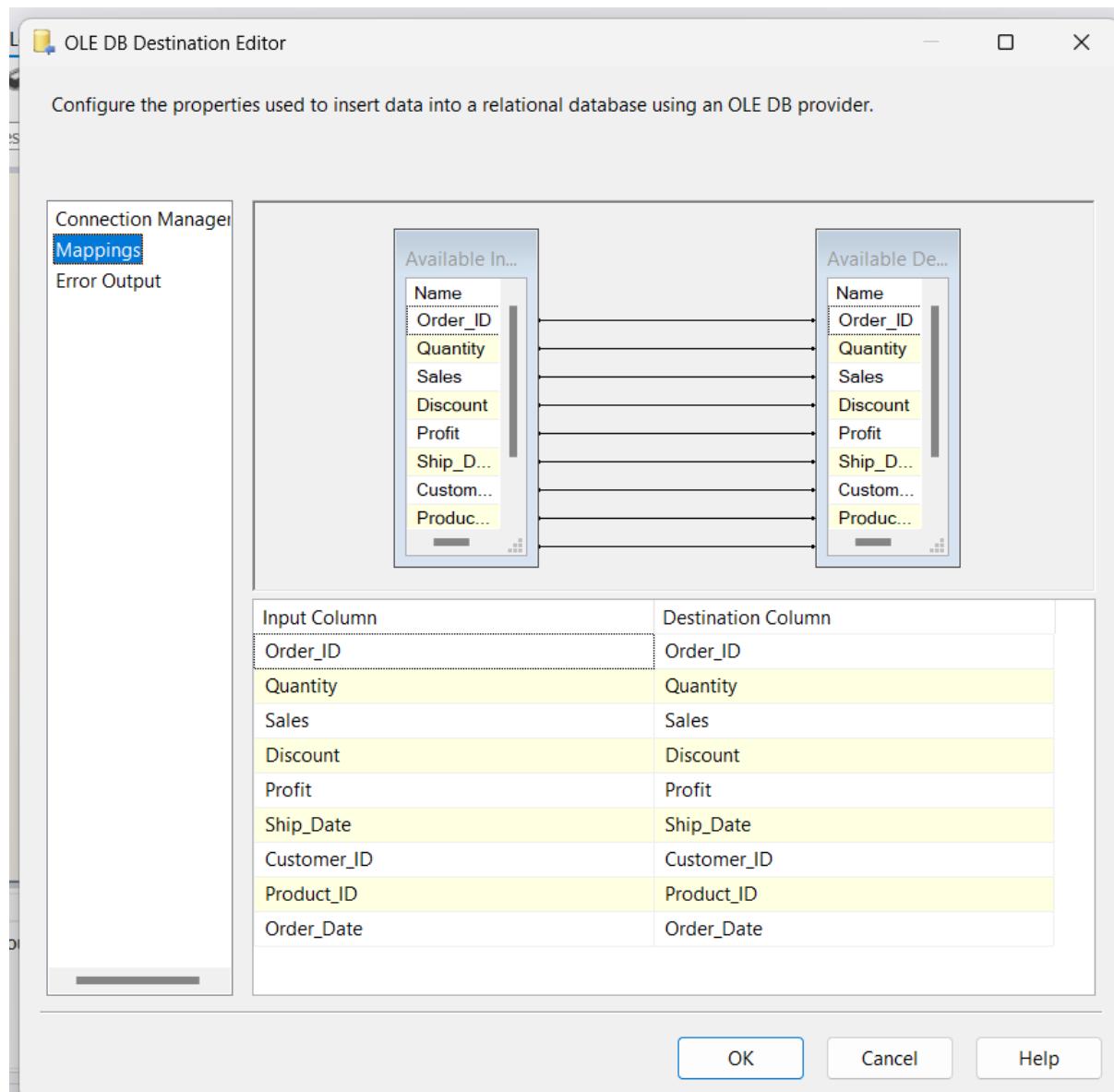
```
SELECT Order_ID, Quantity, Sales, Discount, Profit, Ship_Date, Customer_ID, Product_ID, Order_Date
FROM
dbo.SampleSuperstore
```



## 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 Default 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

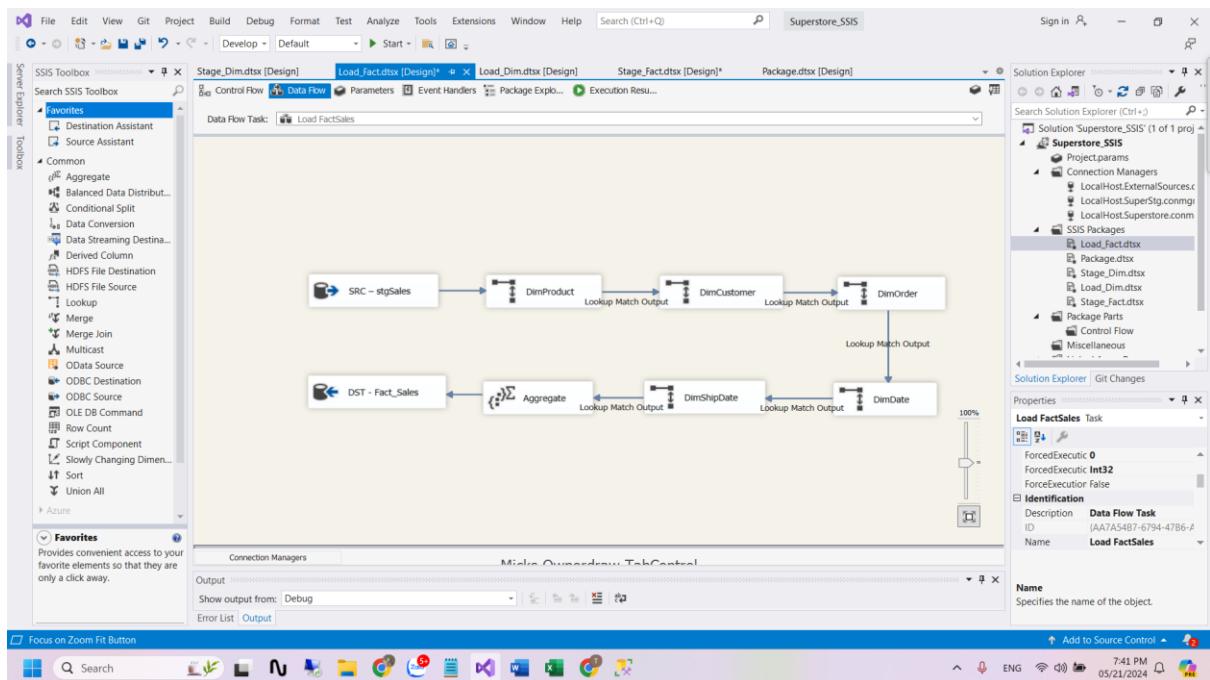
Connection Managers: (project) localhost.ExternalSources, (project) localhost.Superstore, (project) localhost.Superstore

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

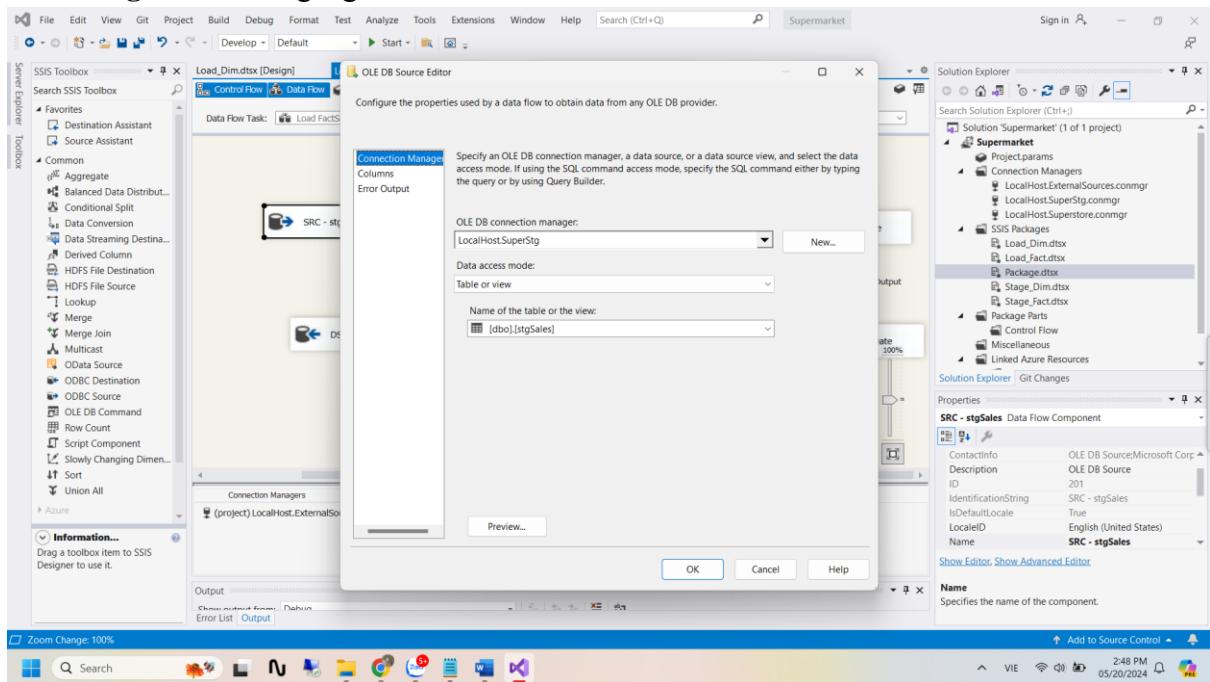
Autos, Immediate Window, Call Stack, Breakpoints, Exception Settings, Command Window, Immediate Window

Ready

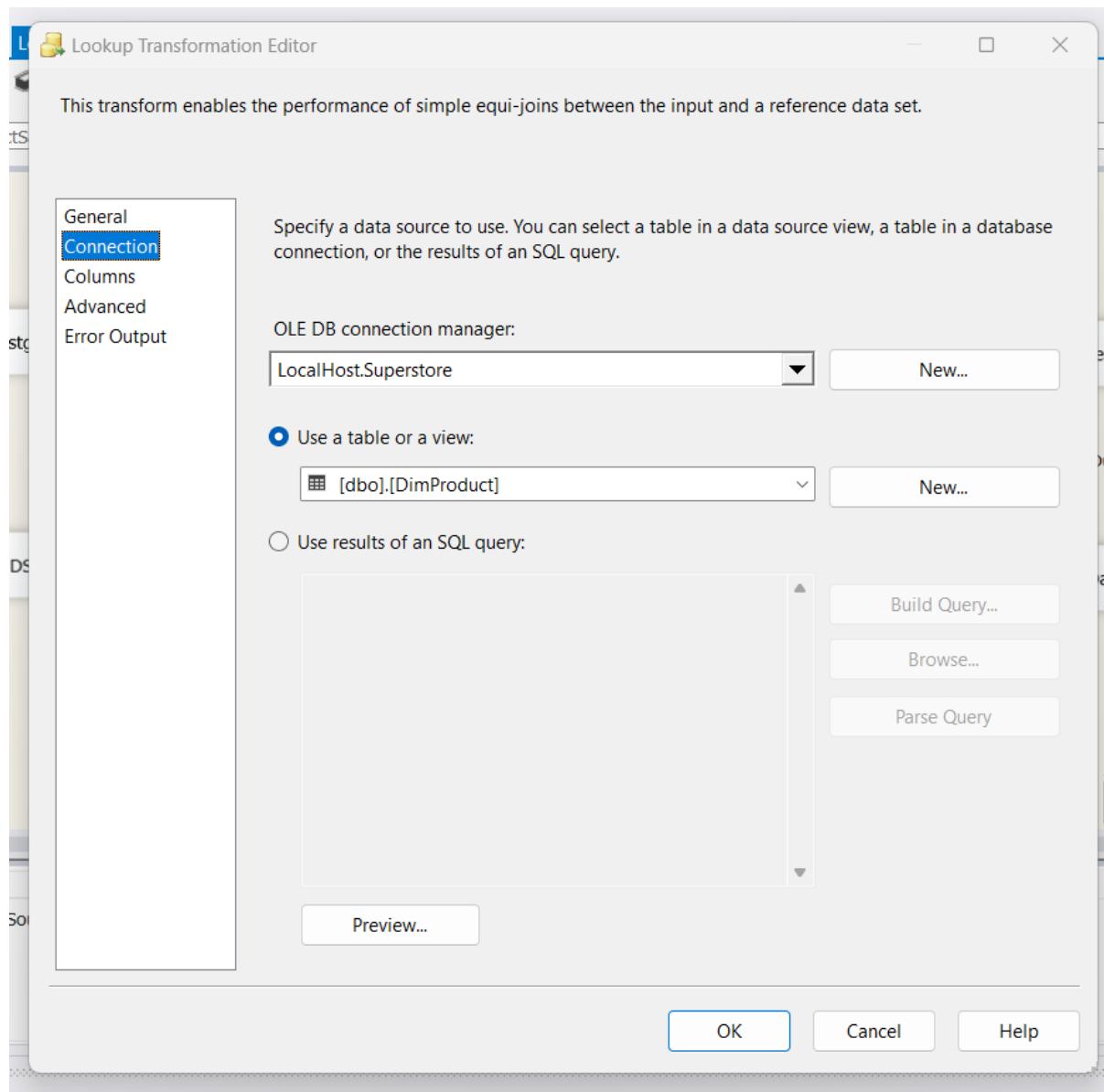
## b) Load dữ liệu vào Sales\_fact

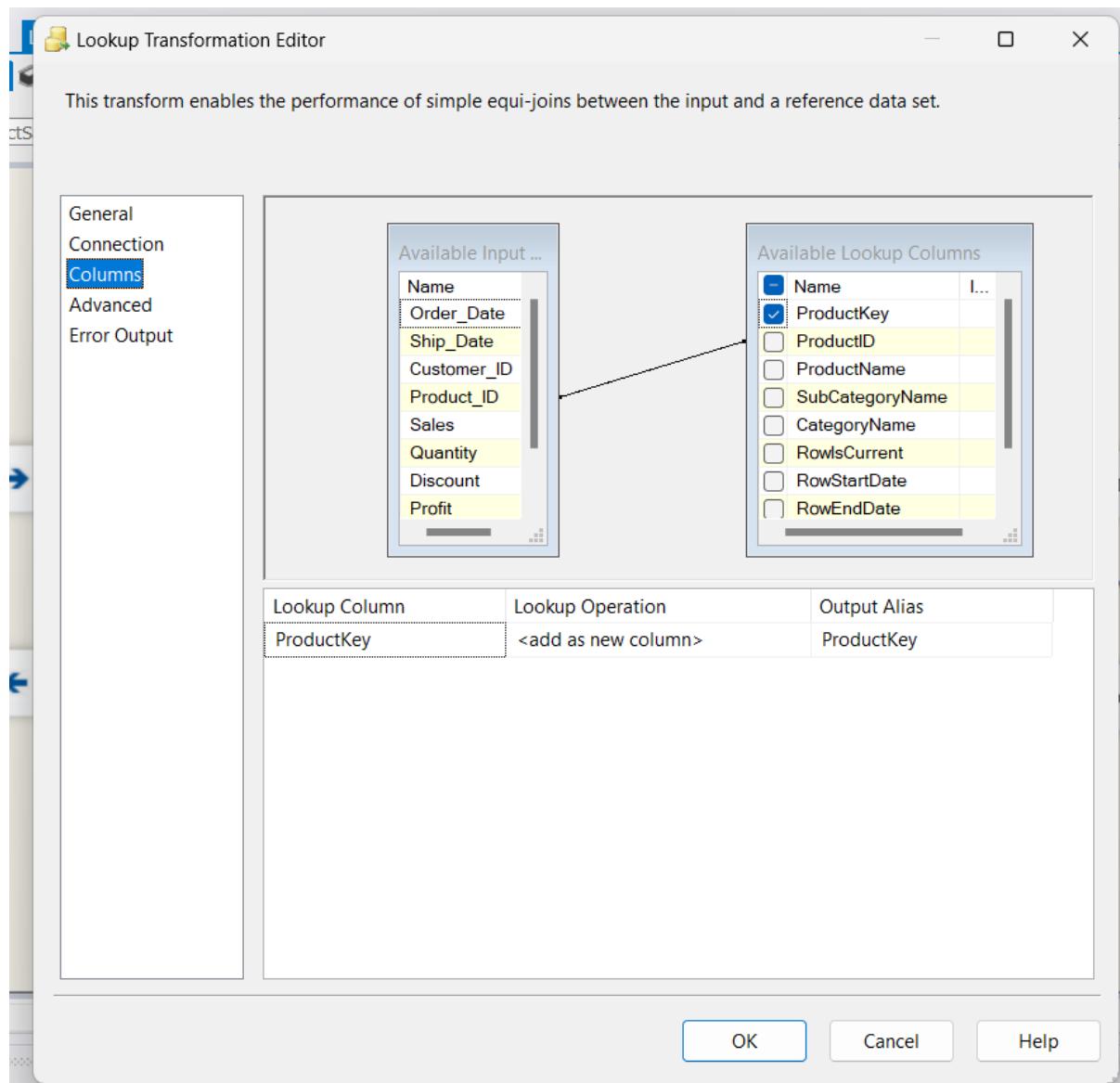


SRC – stgSales là bảng stgSales đã load từ trước

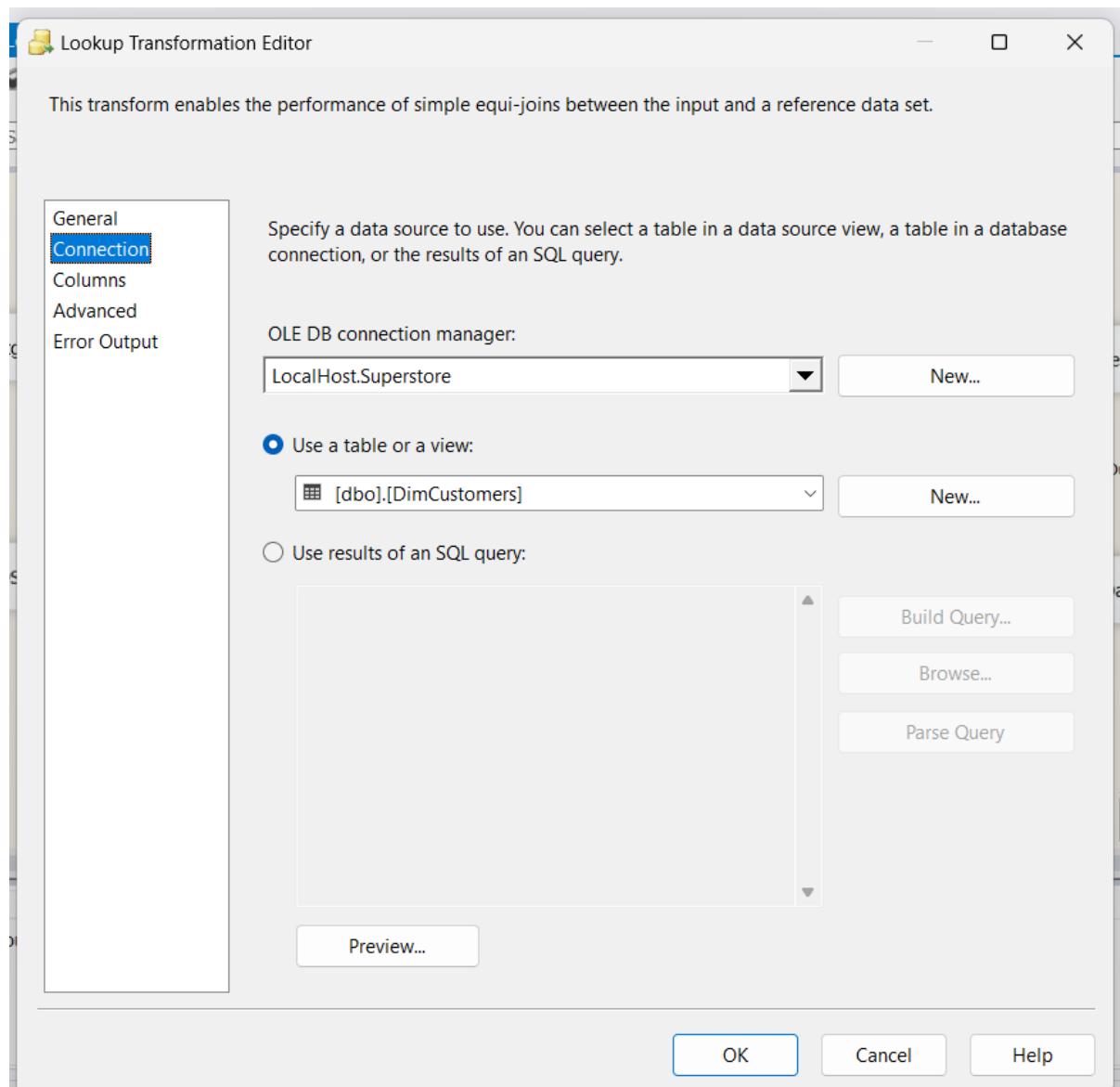


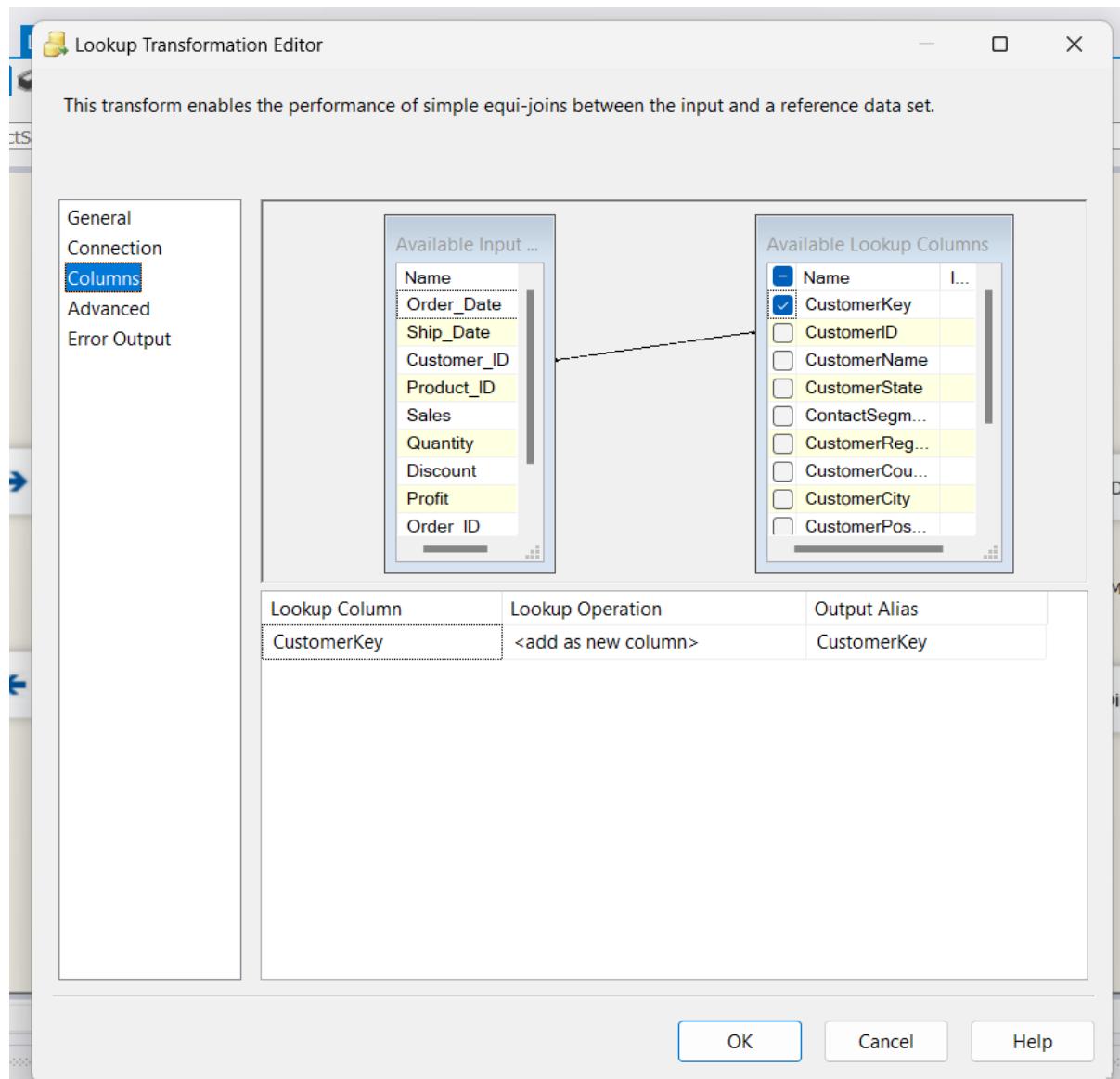
Lookup **DimProduct** ta sẽ map hai **ProductID** từ hai bảng lại với nhau và lấy ra thuộc tính **ProductKey** trong bảng DimProduct



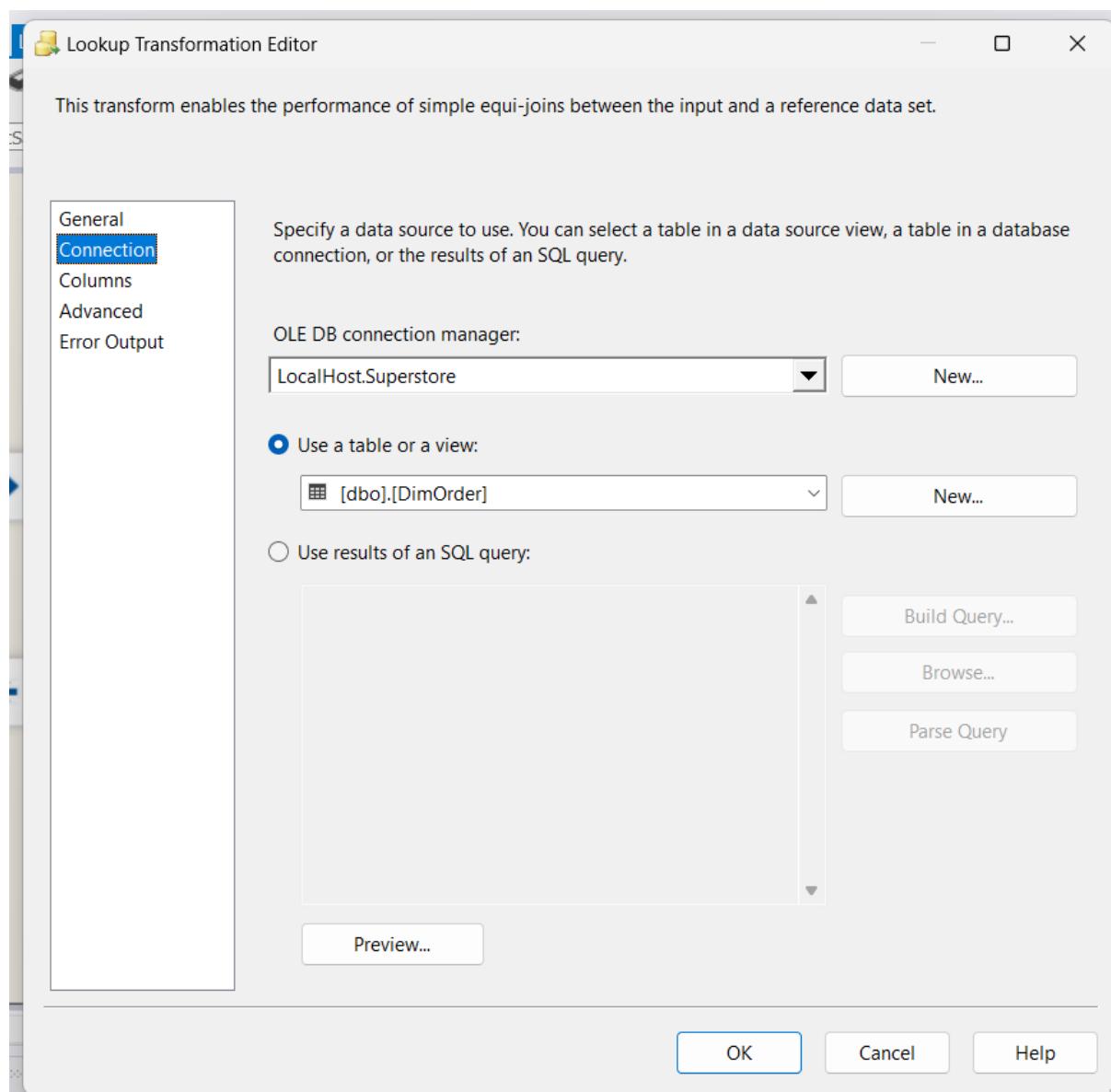


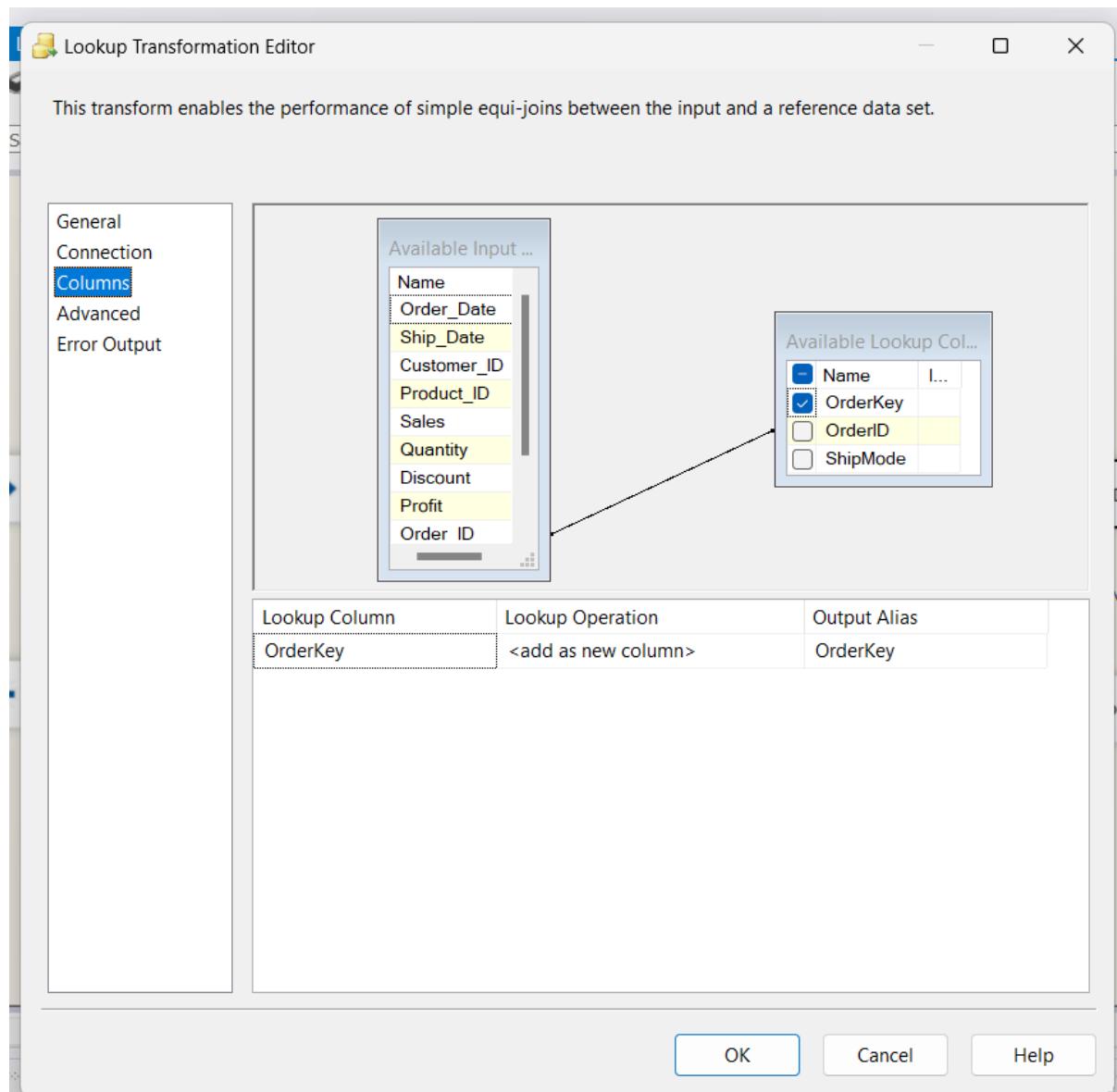
Lookup **DimCustomer** ta sẽ map cột customer\_id trong bảng stgSales với cột CustomerID trong bảng DimCustomer để có thể lấy ra cột CustomerKey



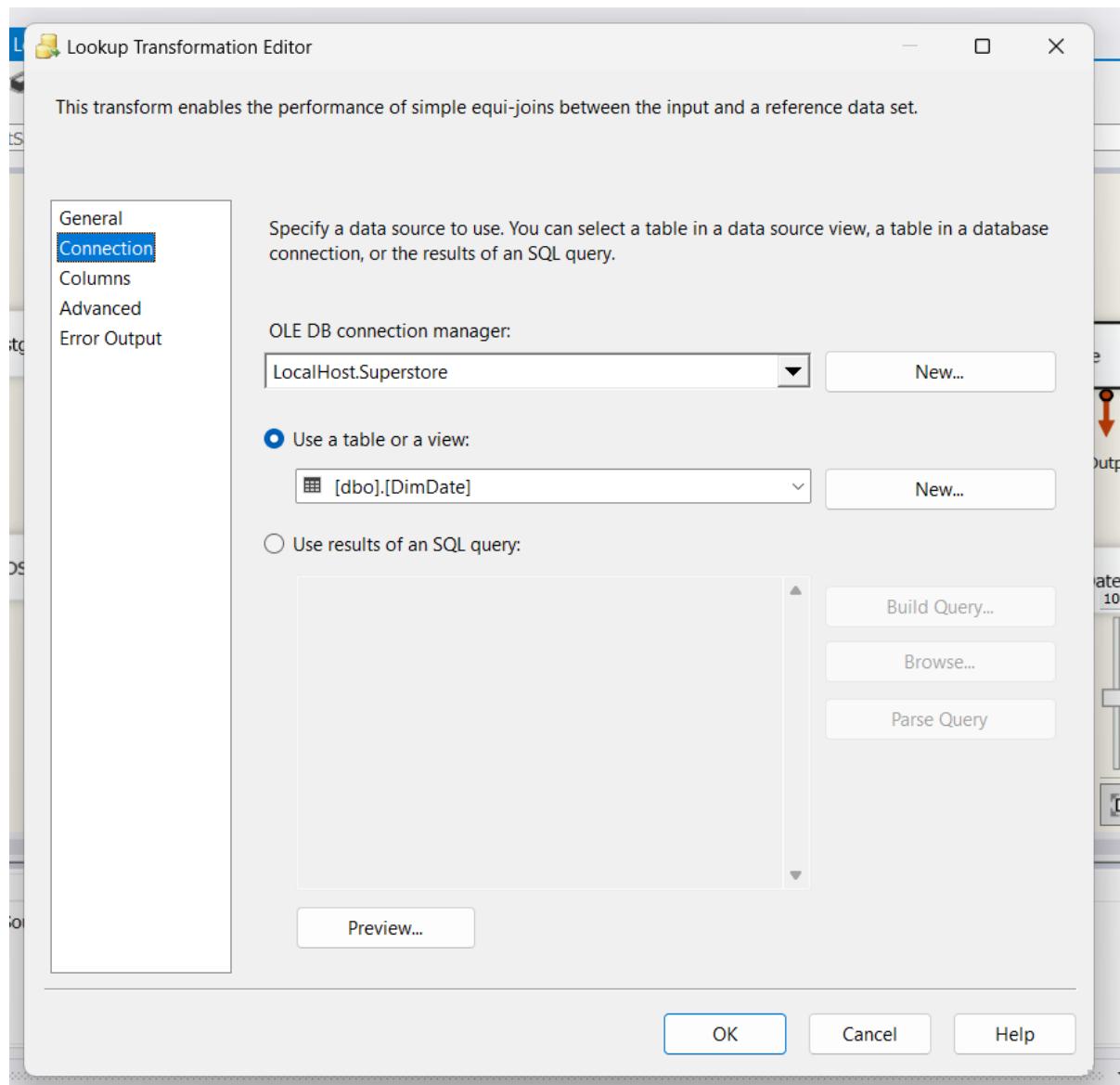


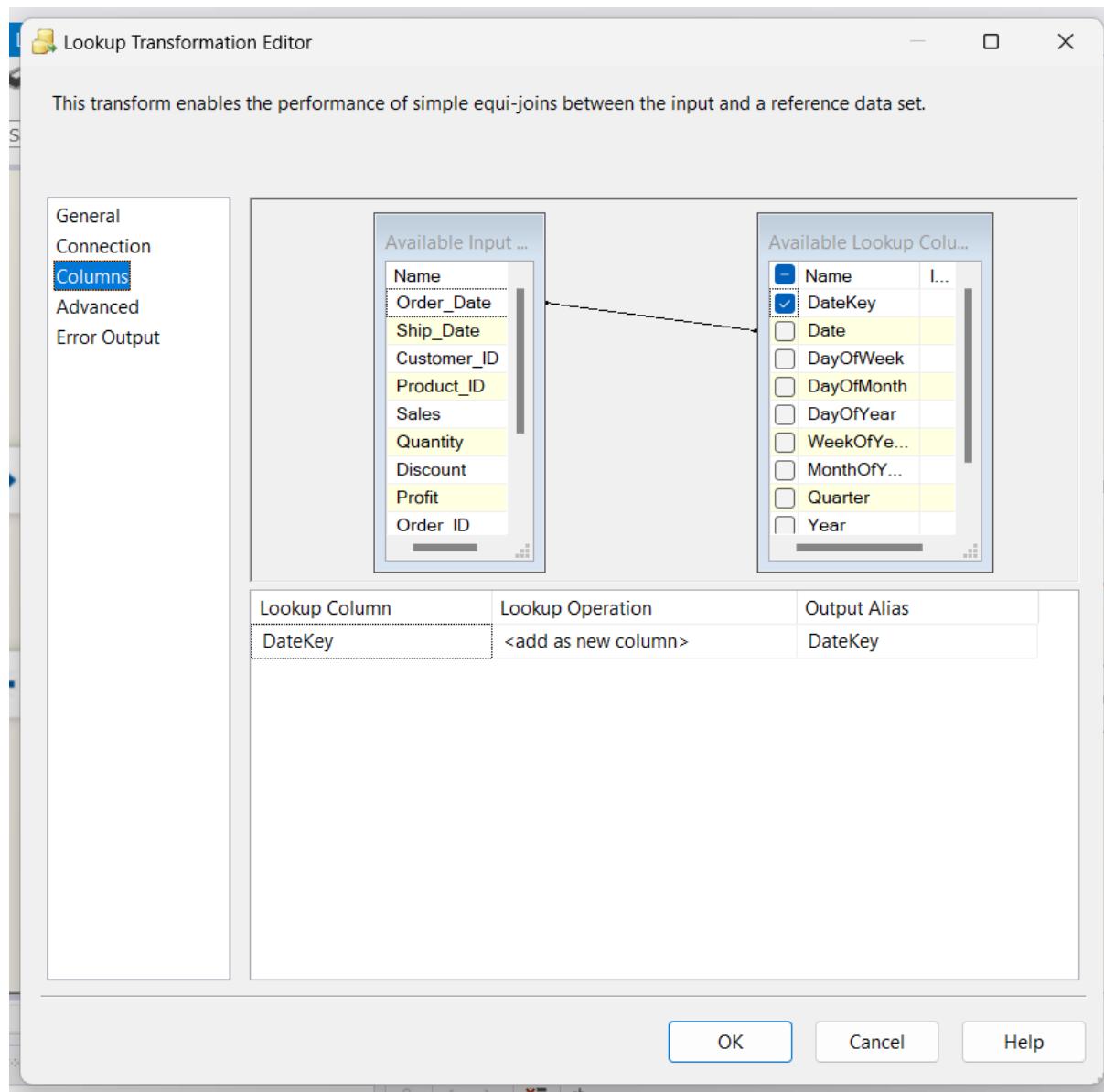
Lookup **DimOrder** ta sẽ map cột order\_id trong bảng stgOrder với cột OrderID trong bảng DimOrder để có thể lấy ra cột OrderKey



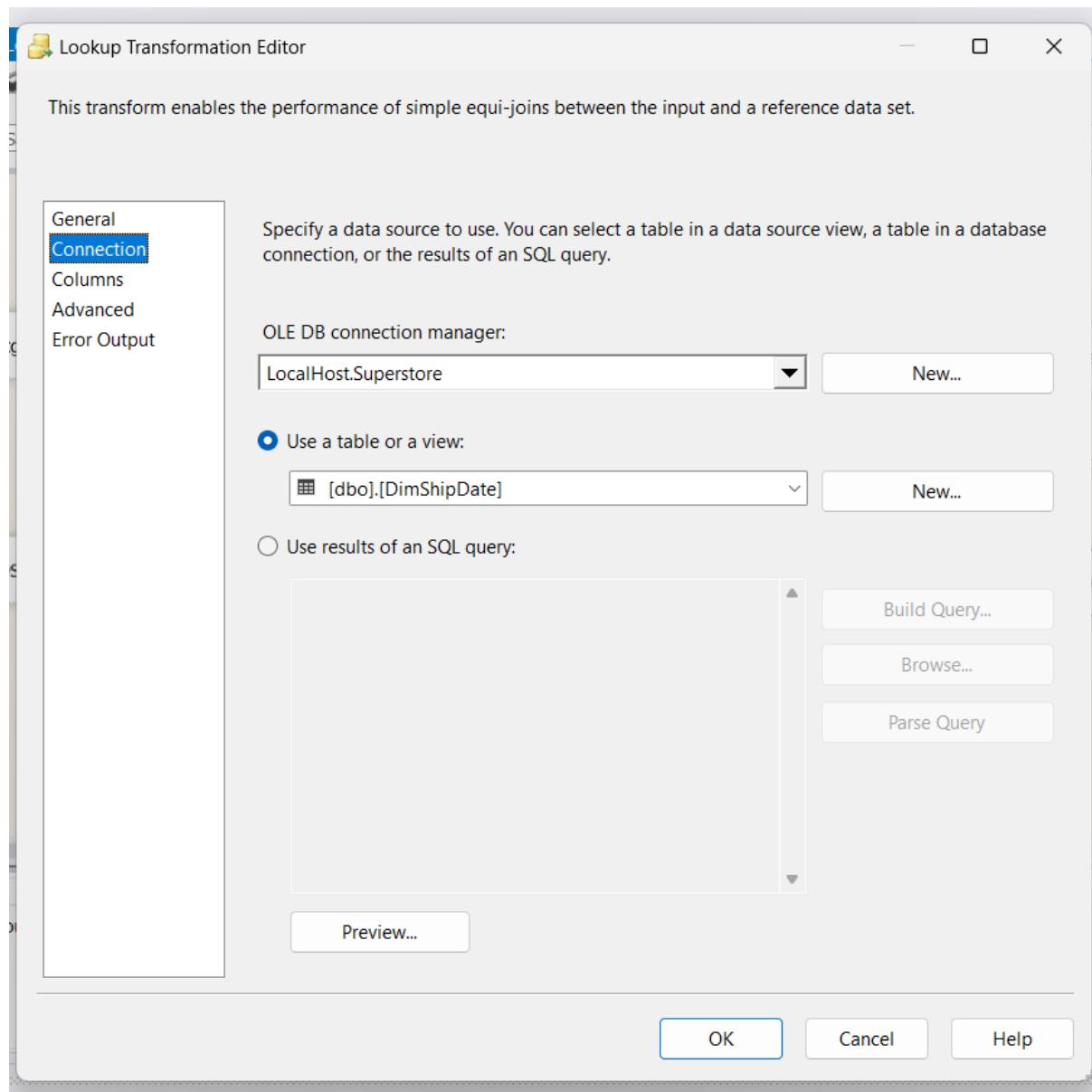


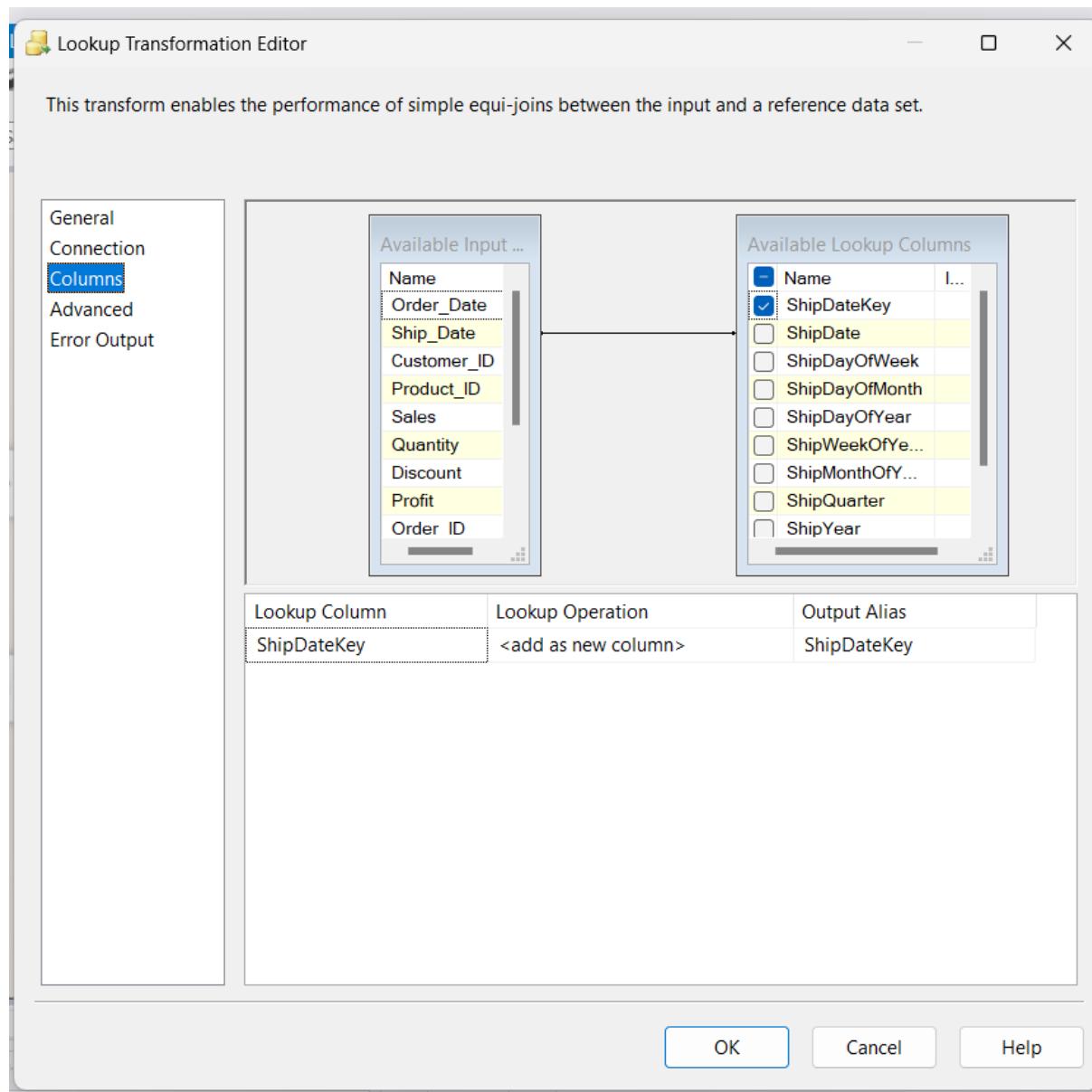
Lookup **DimDate** ta sẽ map cột order\_date ở bảng stgSales với cột Date trong bảng DimDate để lấy ra DateKey



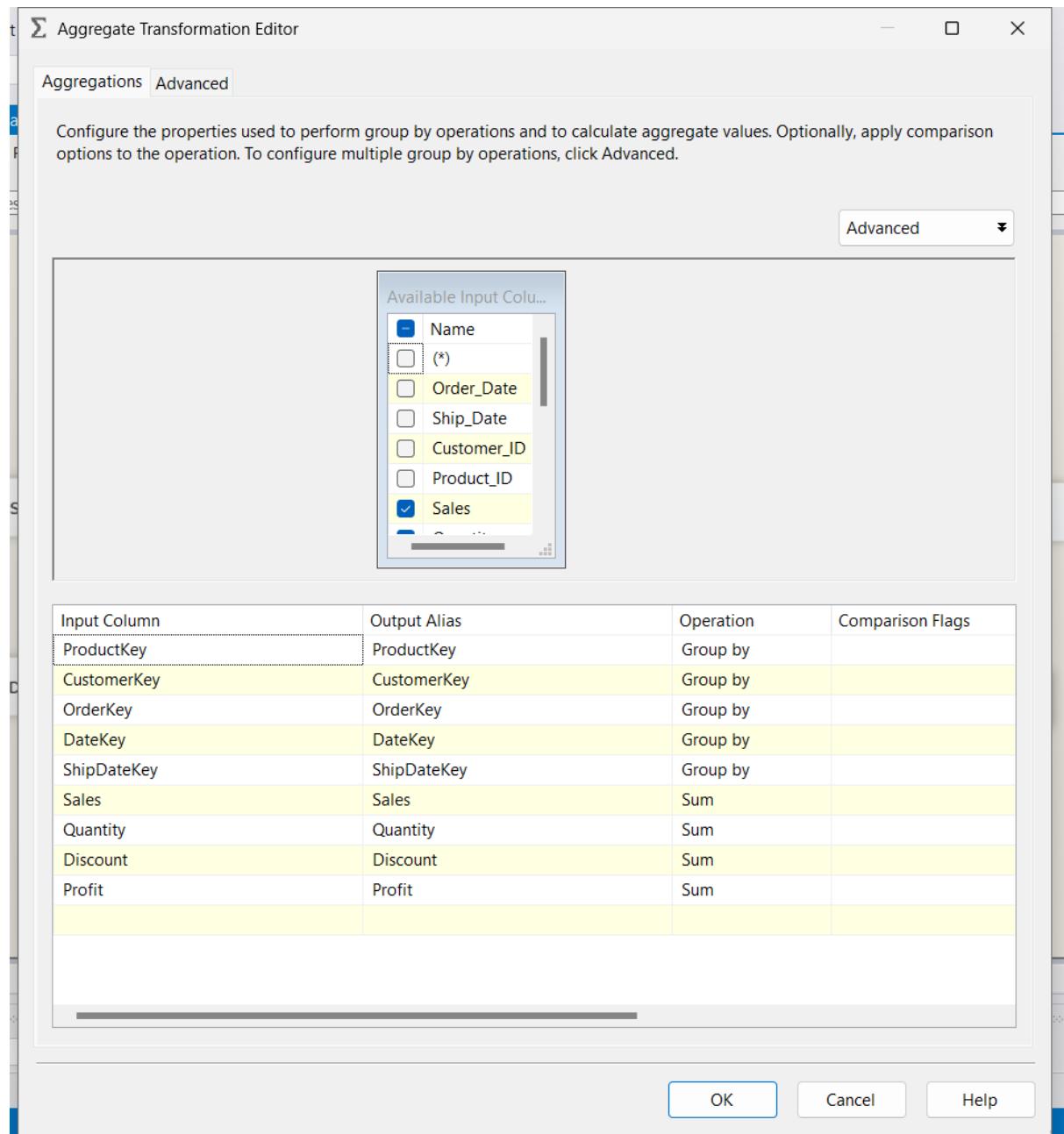


Lookup **DimShipDate** ta sẽ map cột ship\_date ở bảng stgSales với cột ShipDate trong bảng DimShipDate để lấy ra ShipDateKey

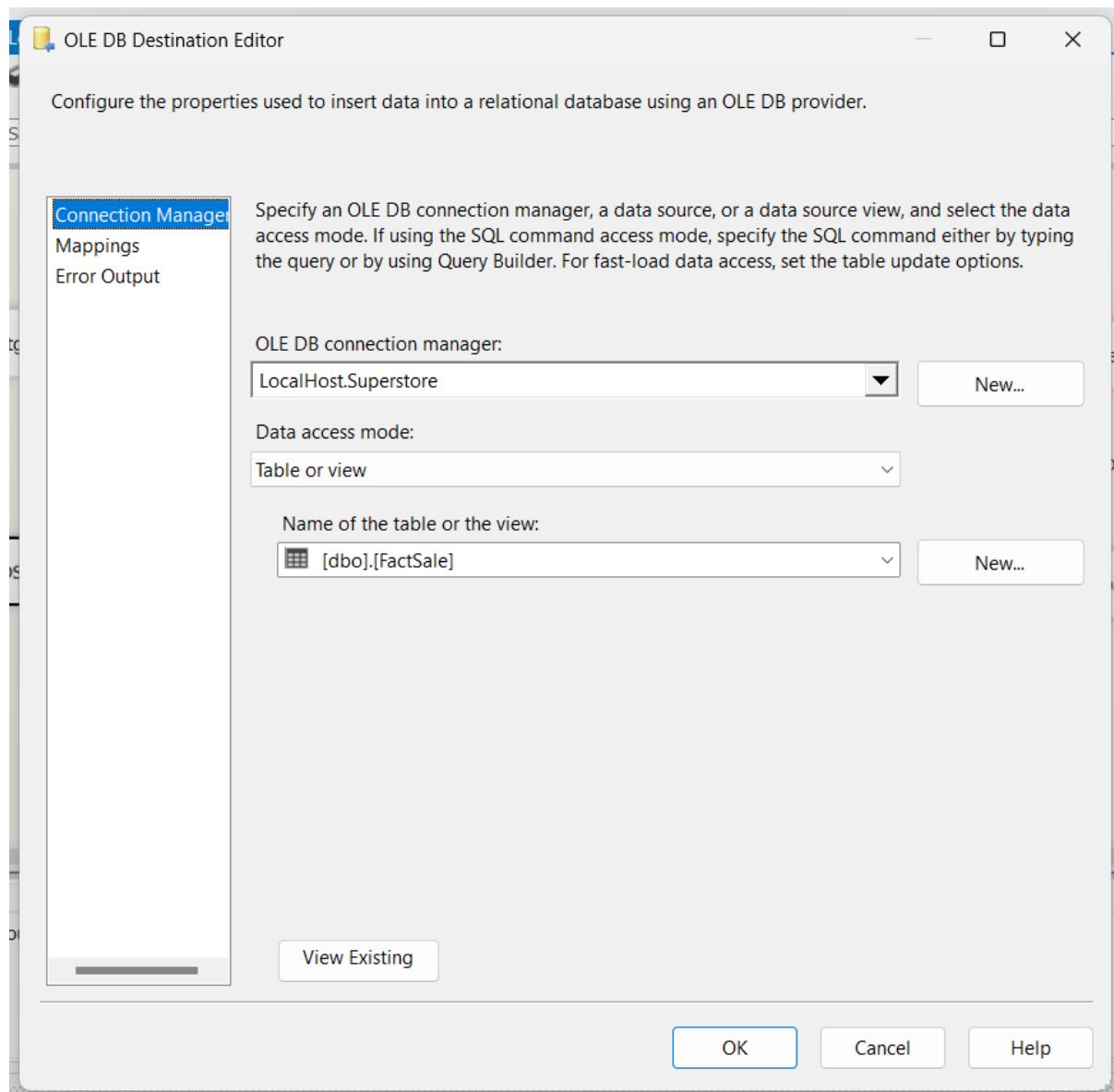


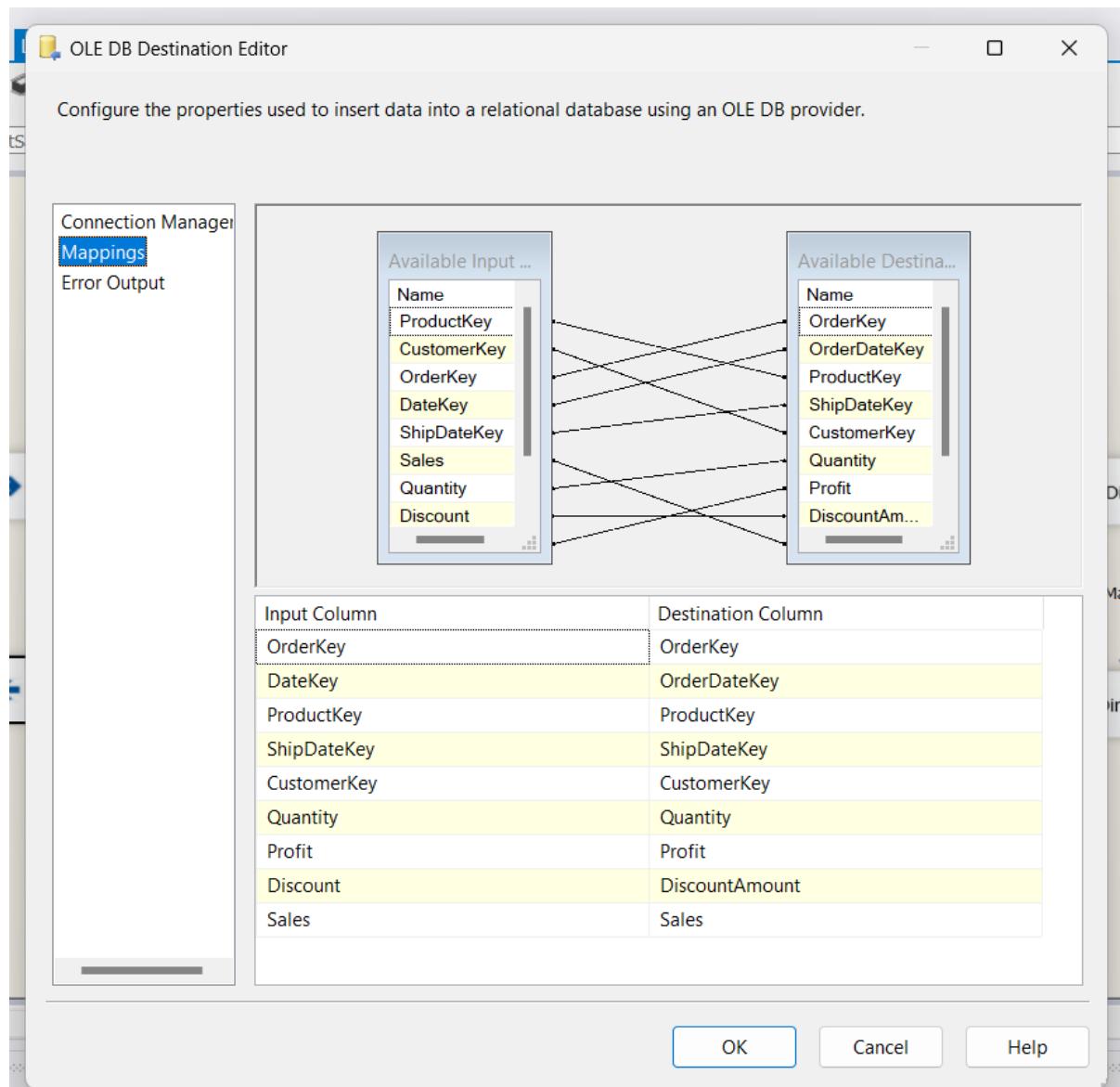


**Aggregate** Ta sẽ lấy những cột cần thiết để đưa vào bảng fact

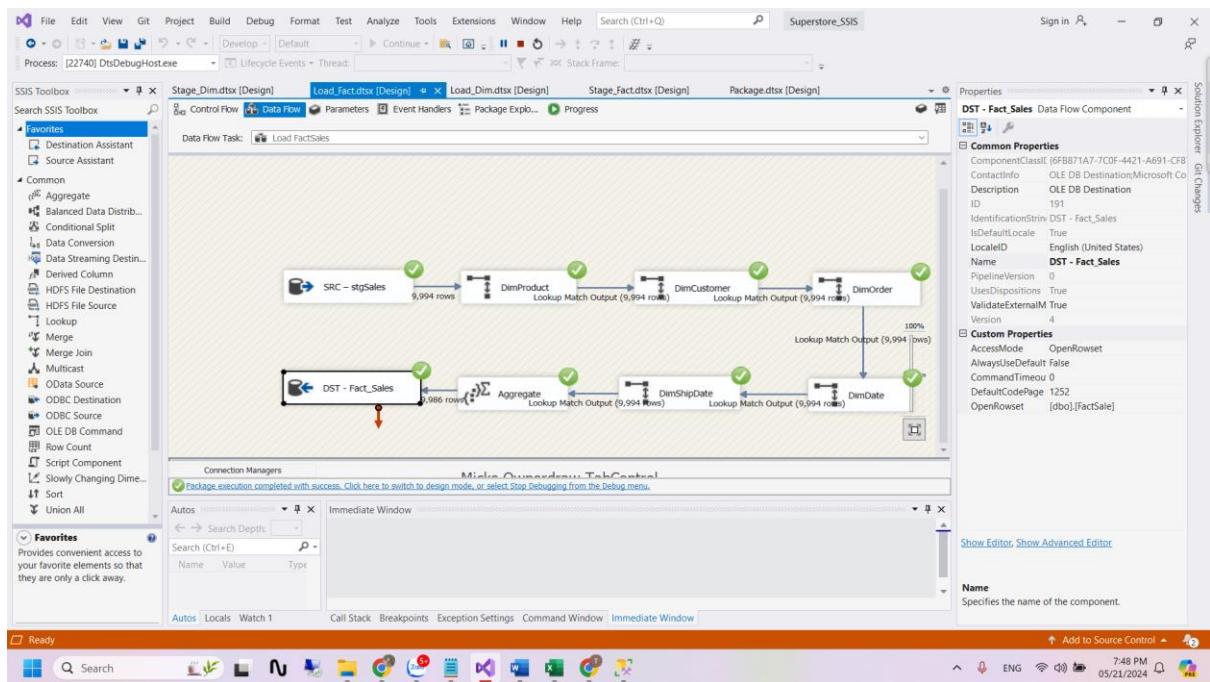


## DST - Fact\_Sales



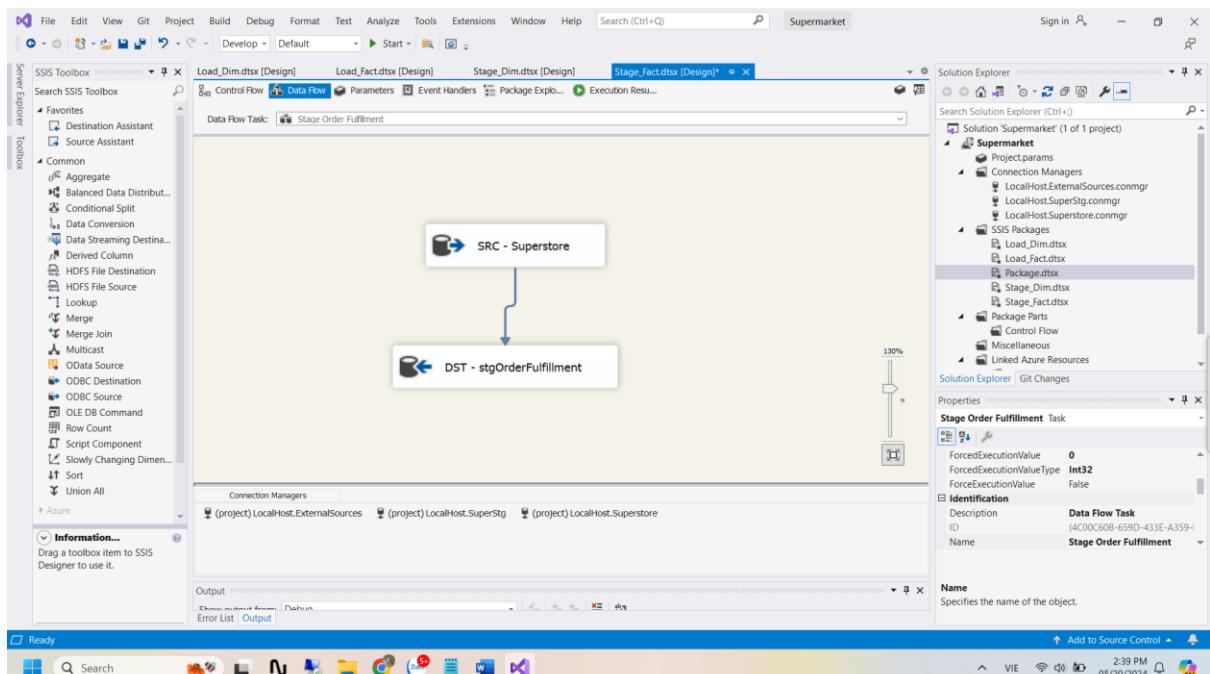


**Khởi chạy**

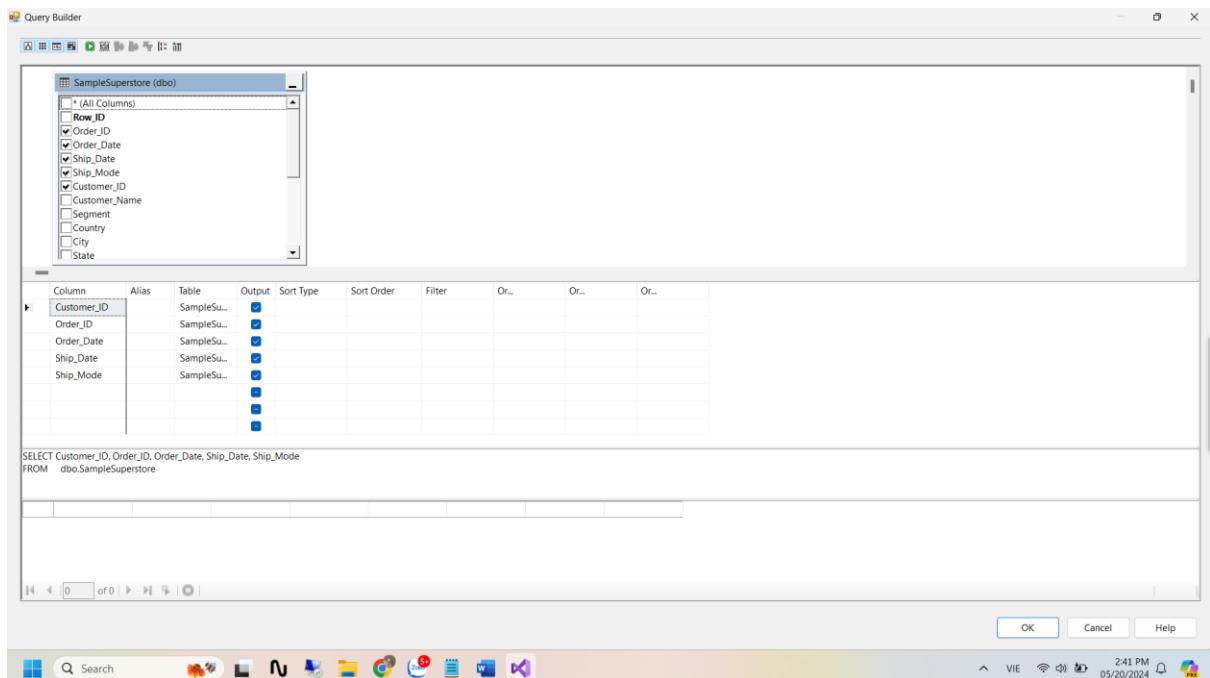


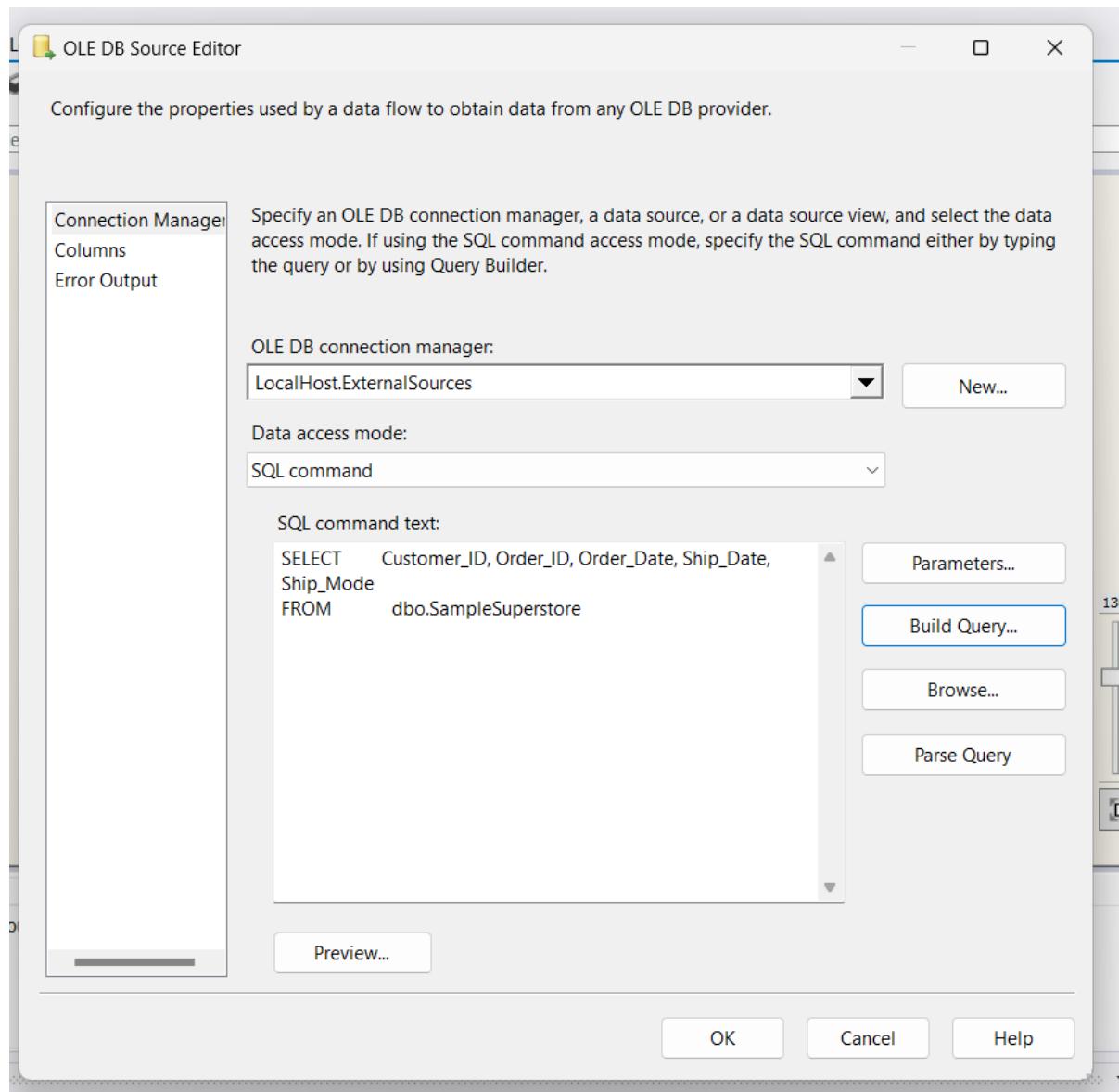
### 3.3.2. Fact OrderFulfillment

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

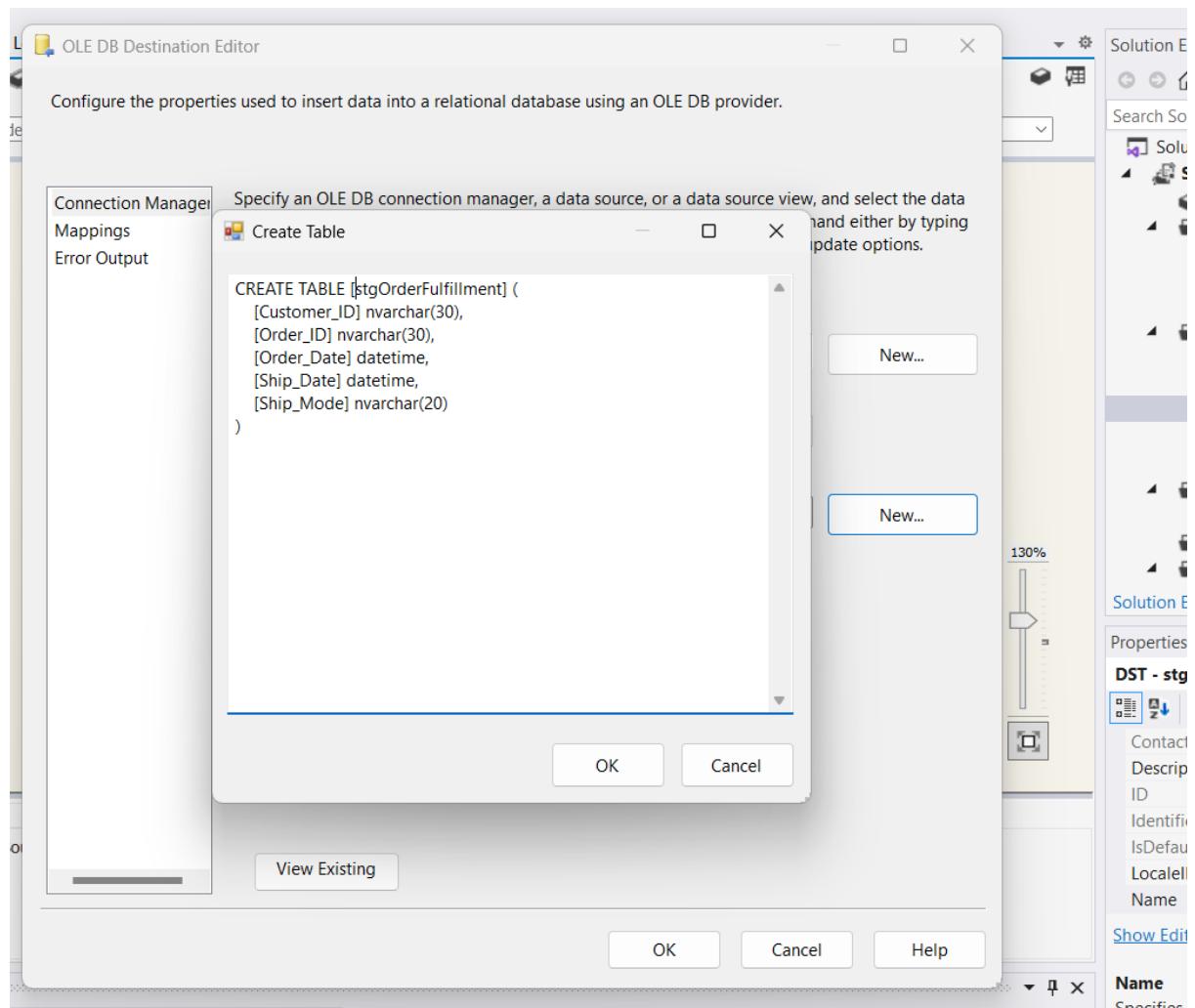


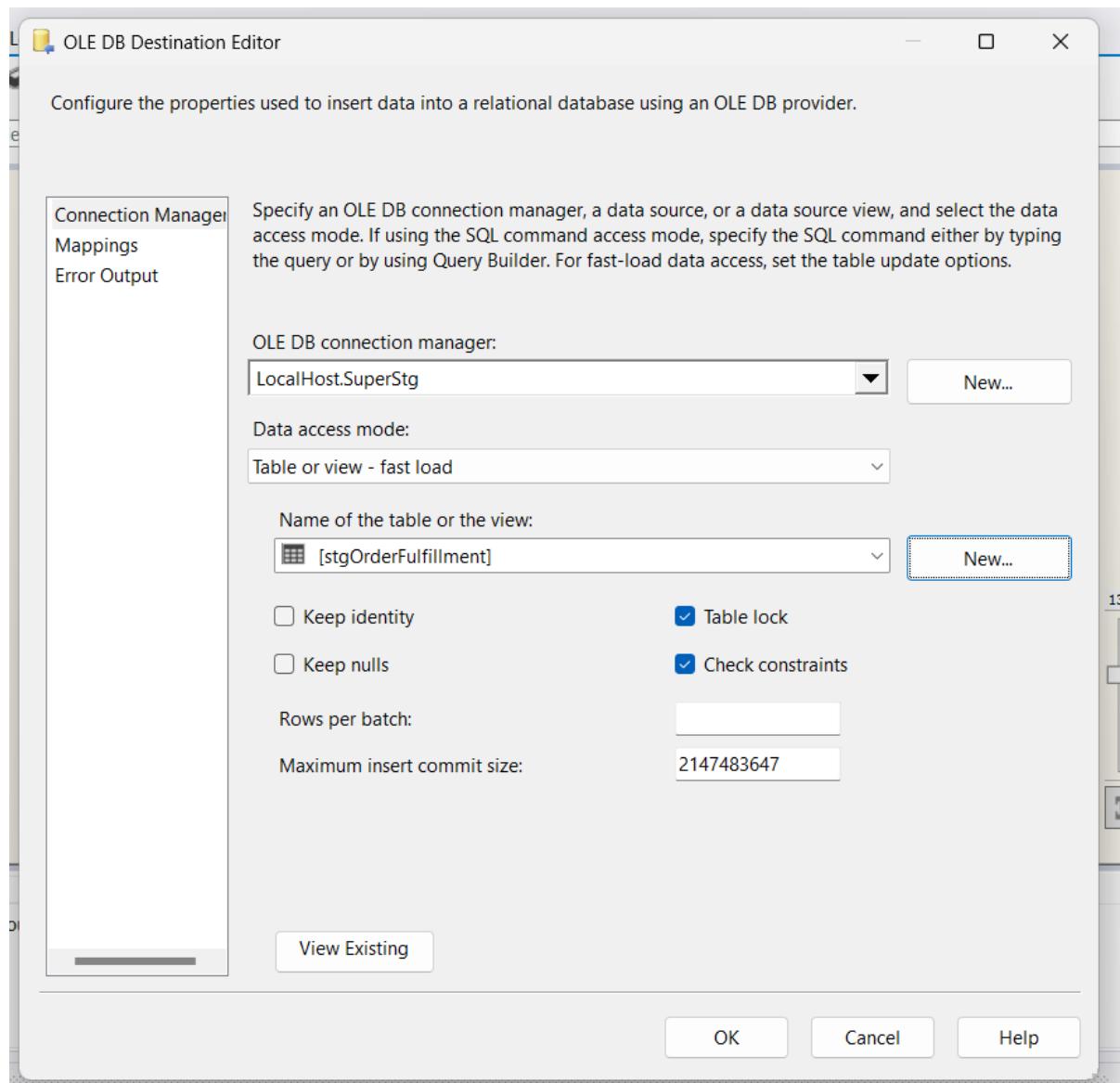
SRC – Superstore là dữ liệu nguồn ta lấy những thuộc tính cần thiết để đưa vào bảng stgOrderFulfillment

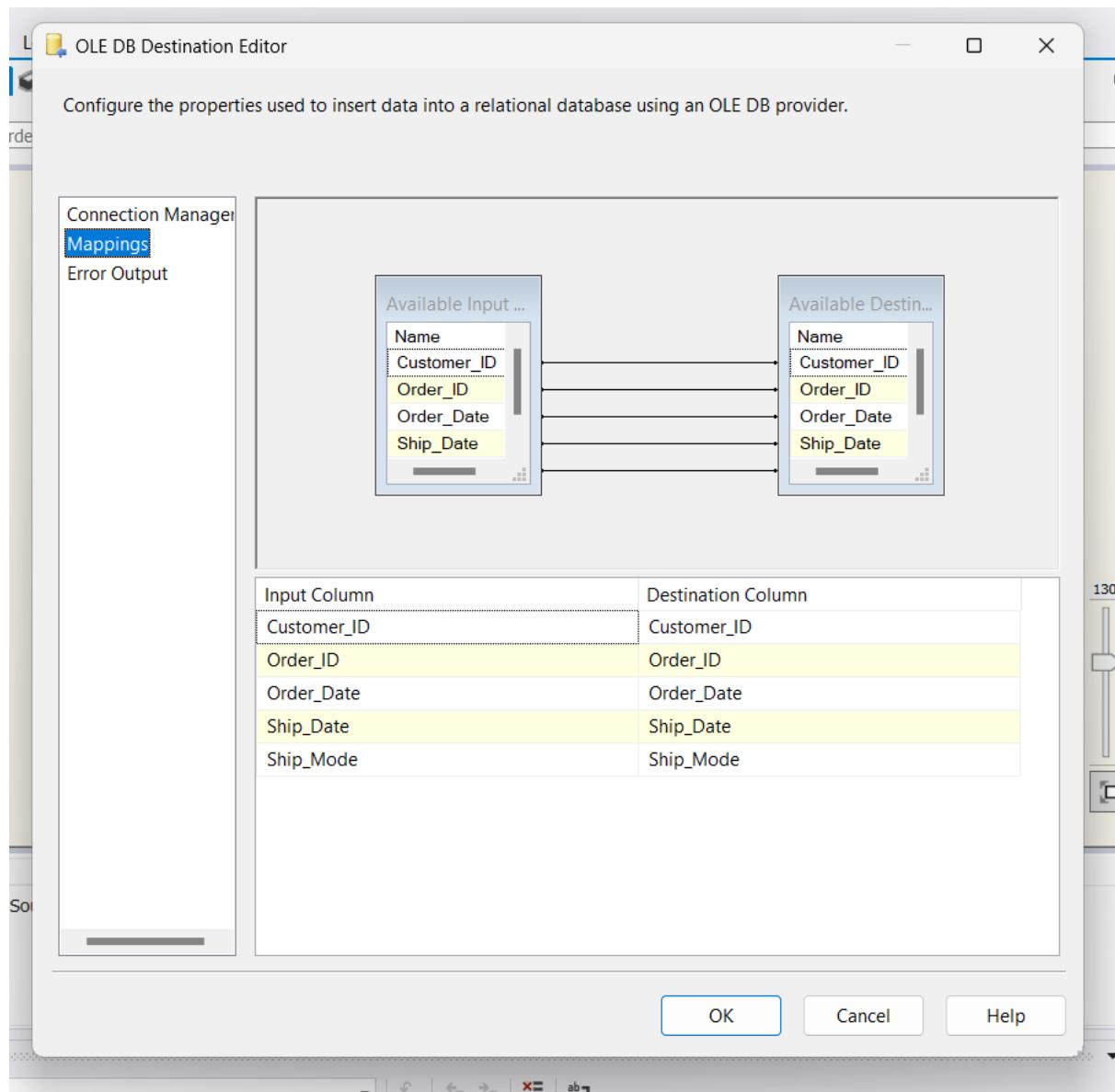


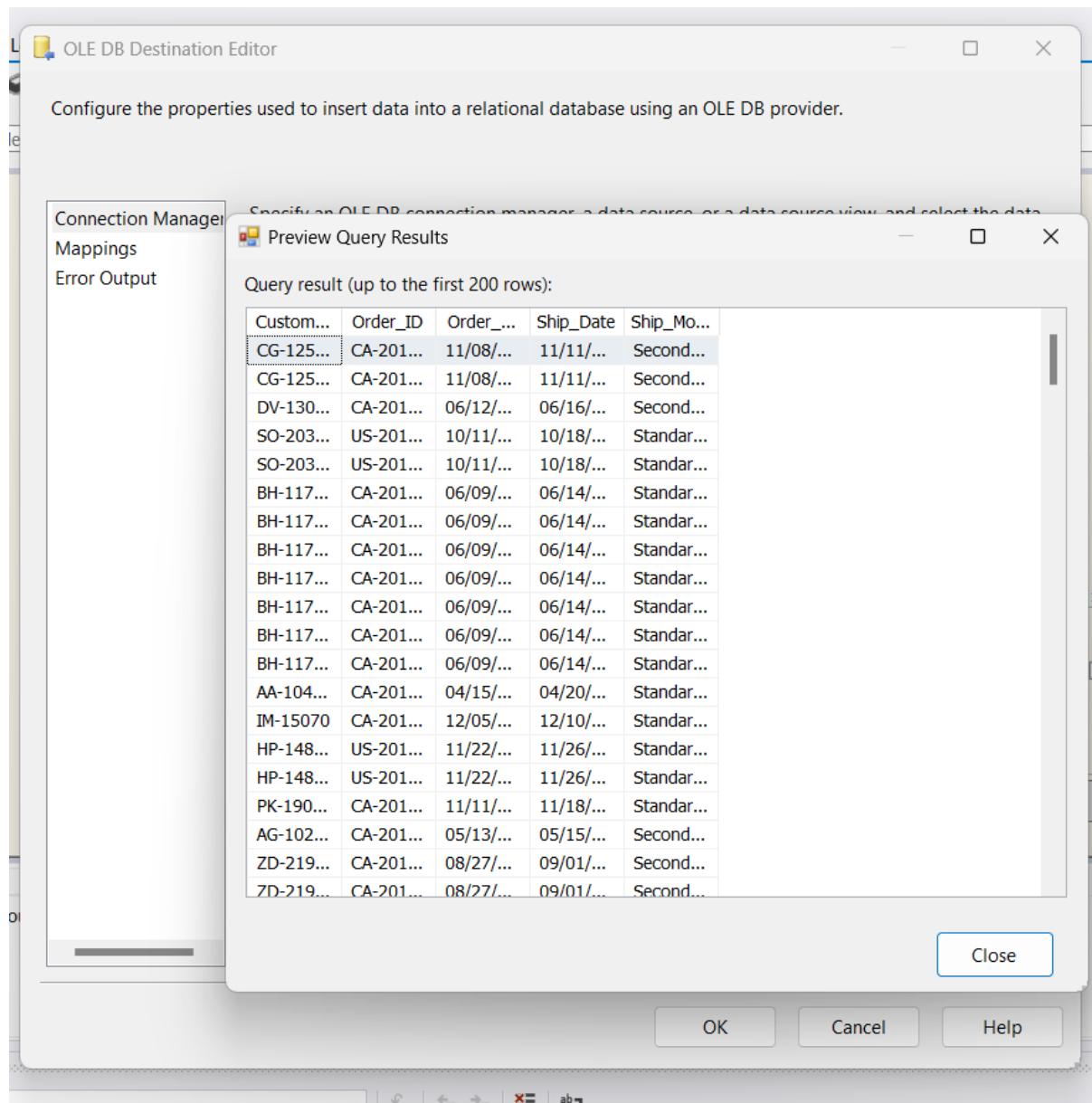


## DST – stgOrderFulfillment

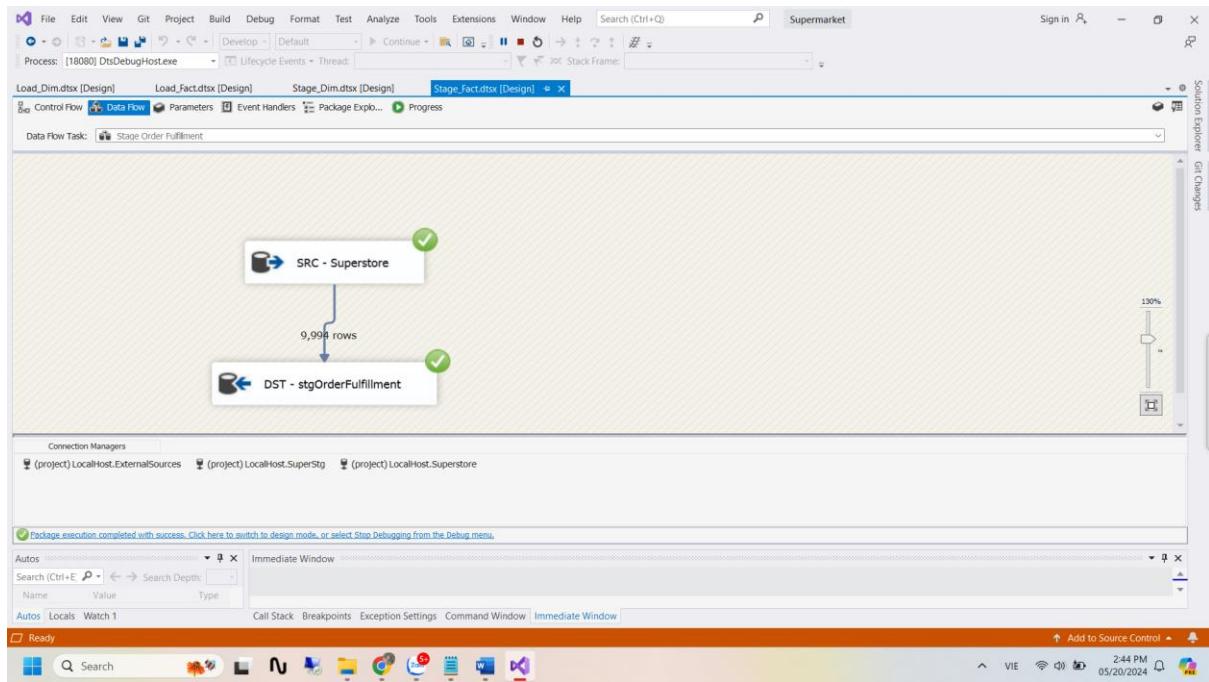




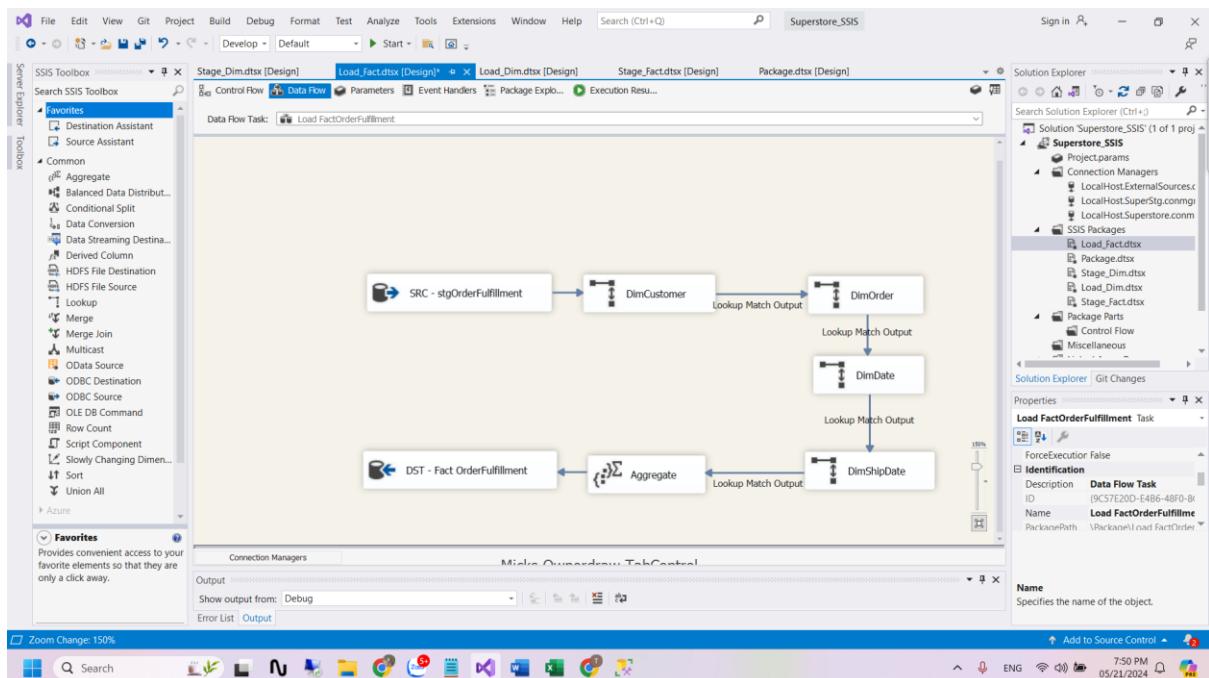




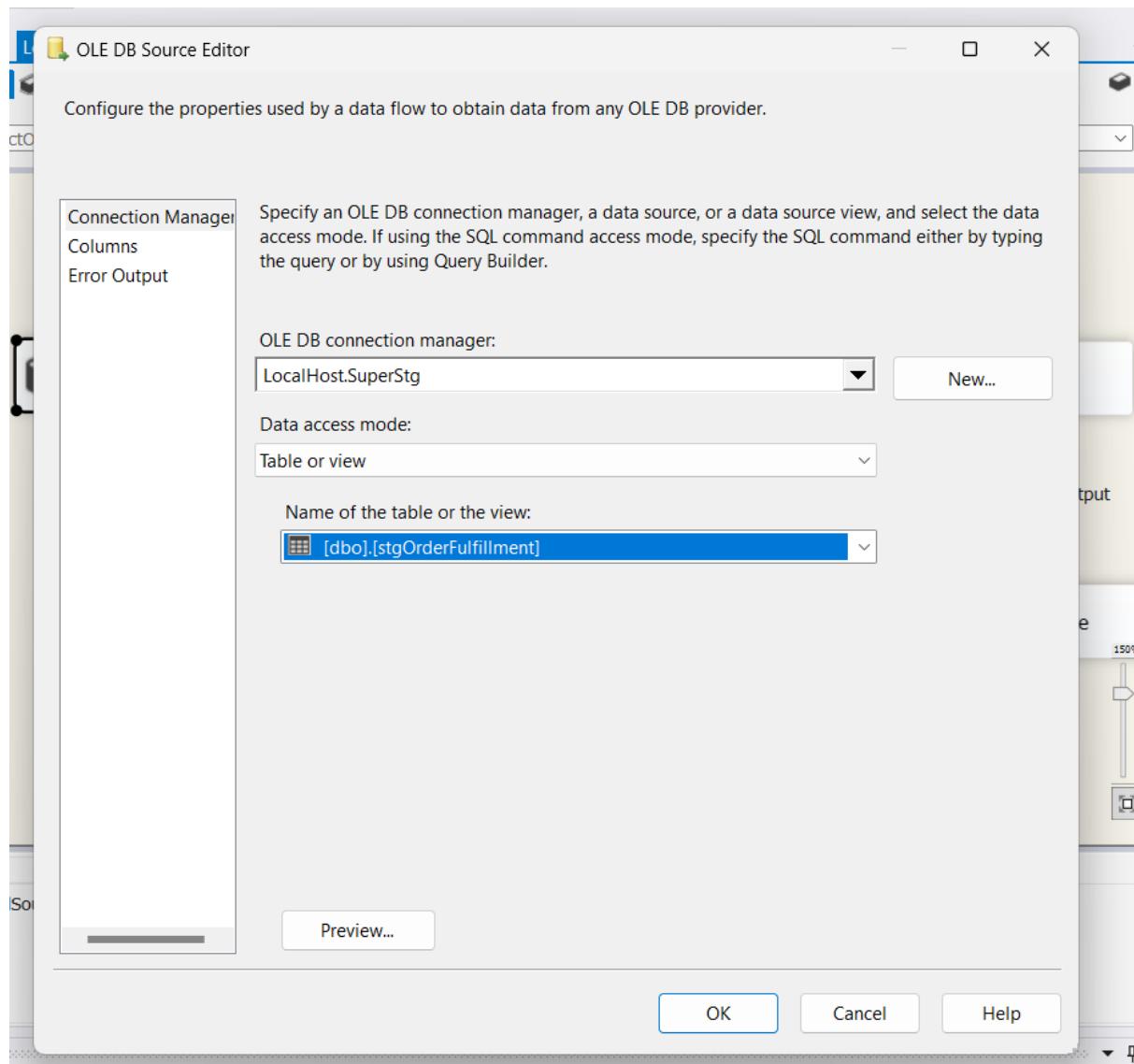
## Khởi chạy



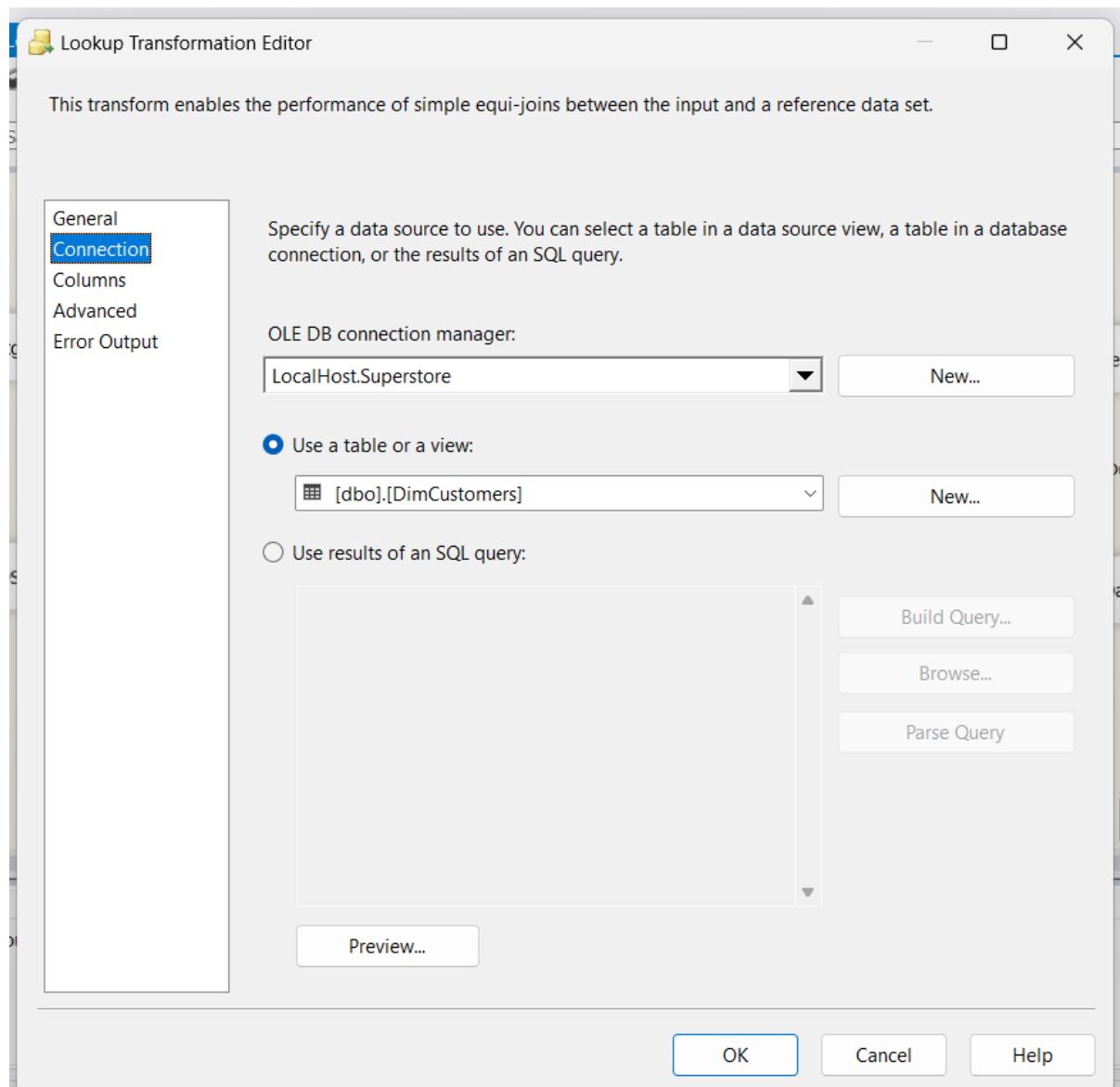
## b) Load dữ liệu vào Fact OrderFulfillment

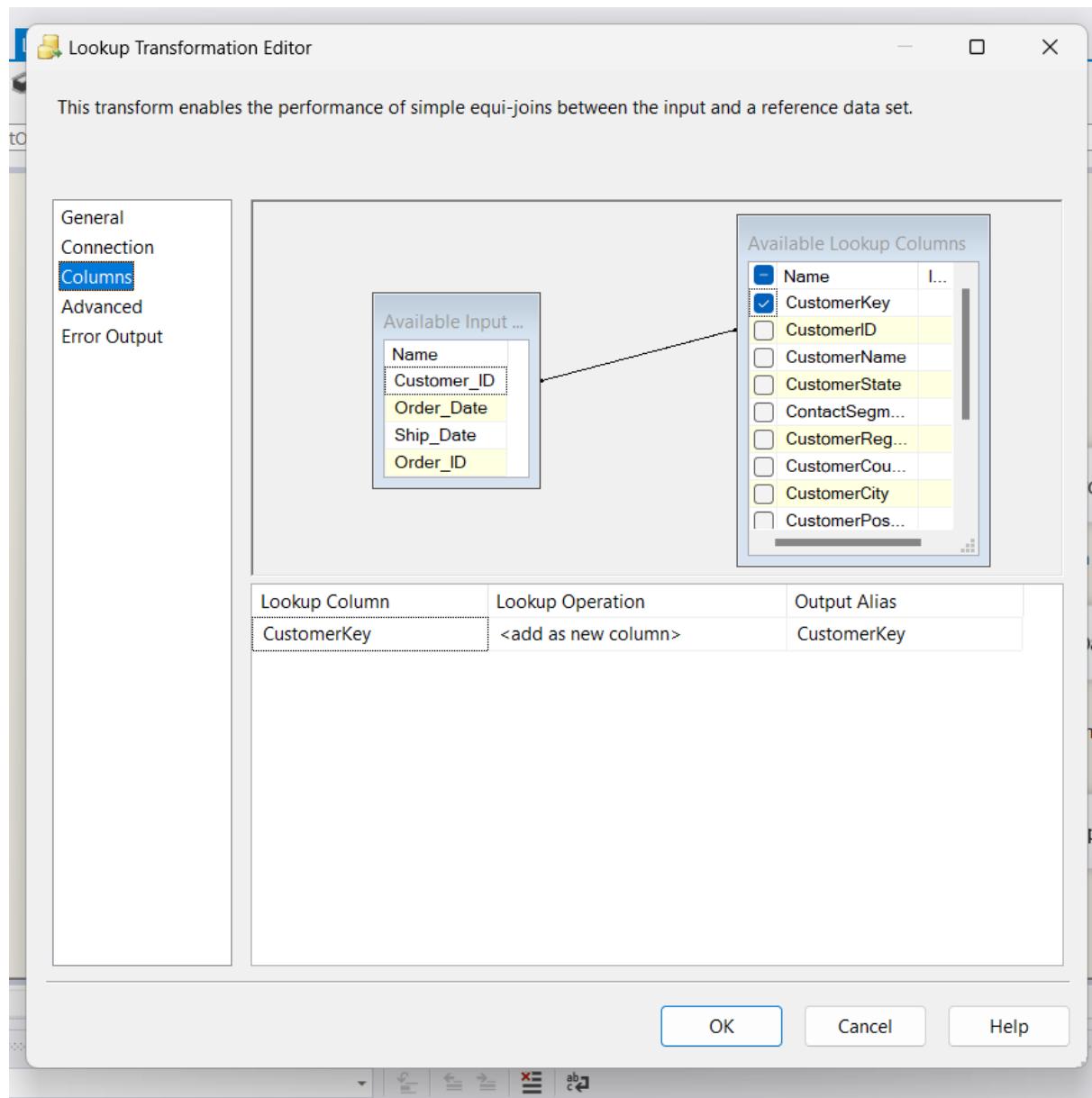


**SRC - stgOrderFulfillment** là bảng stgOrderFulfillment

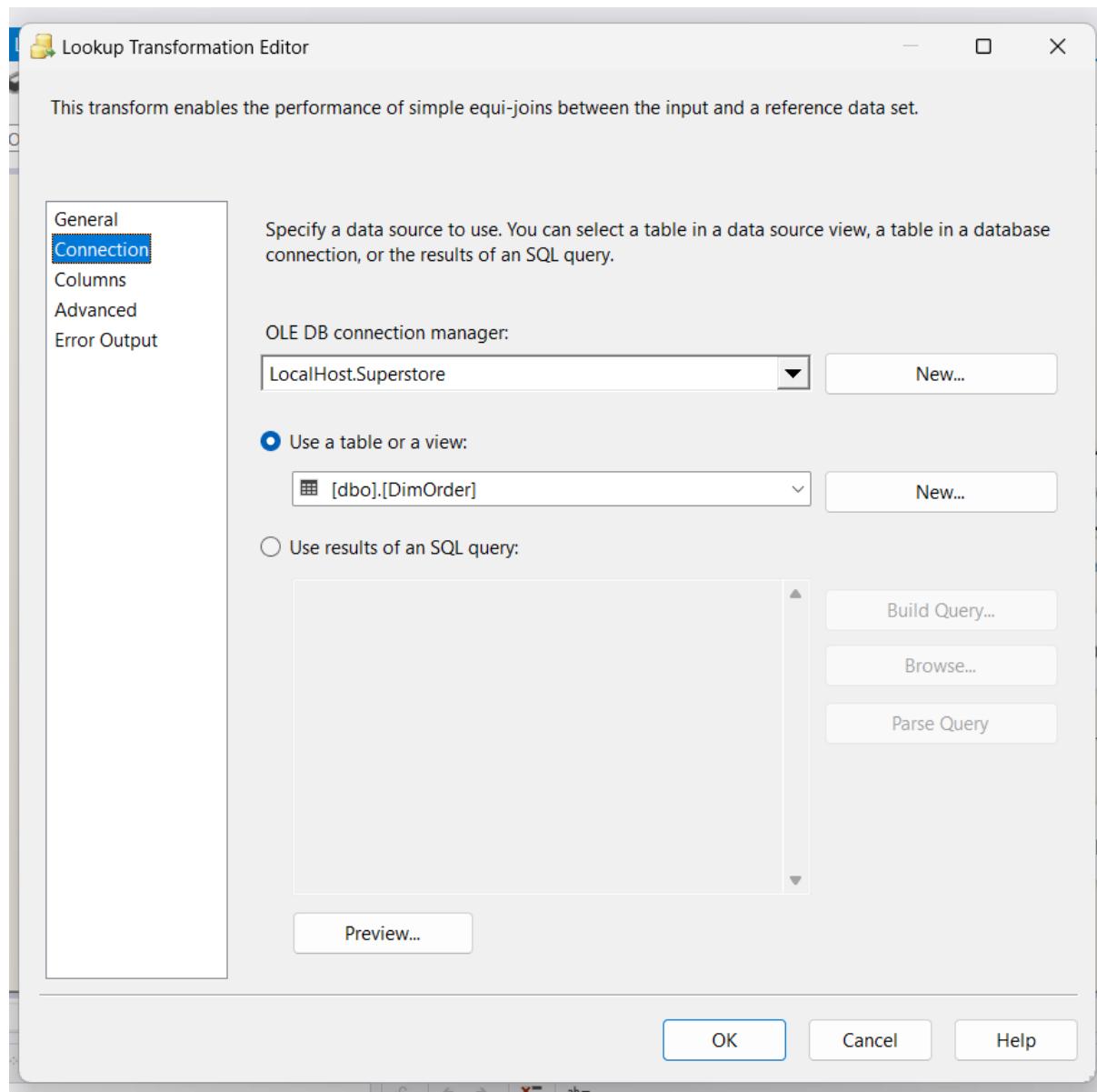


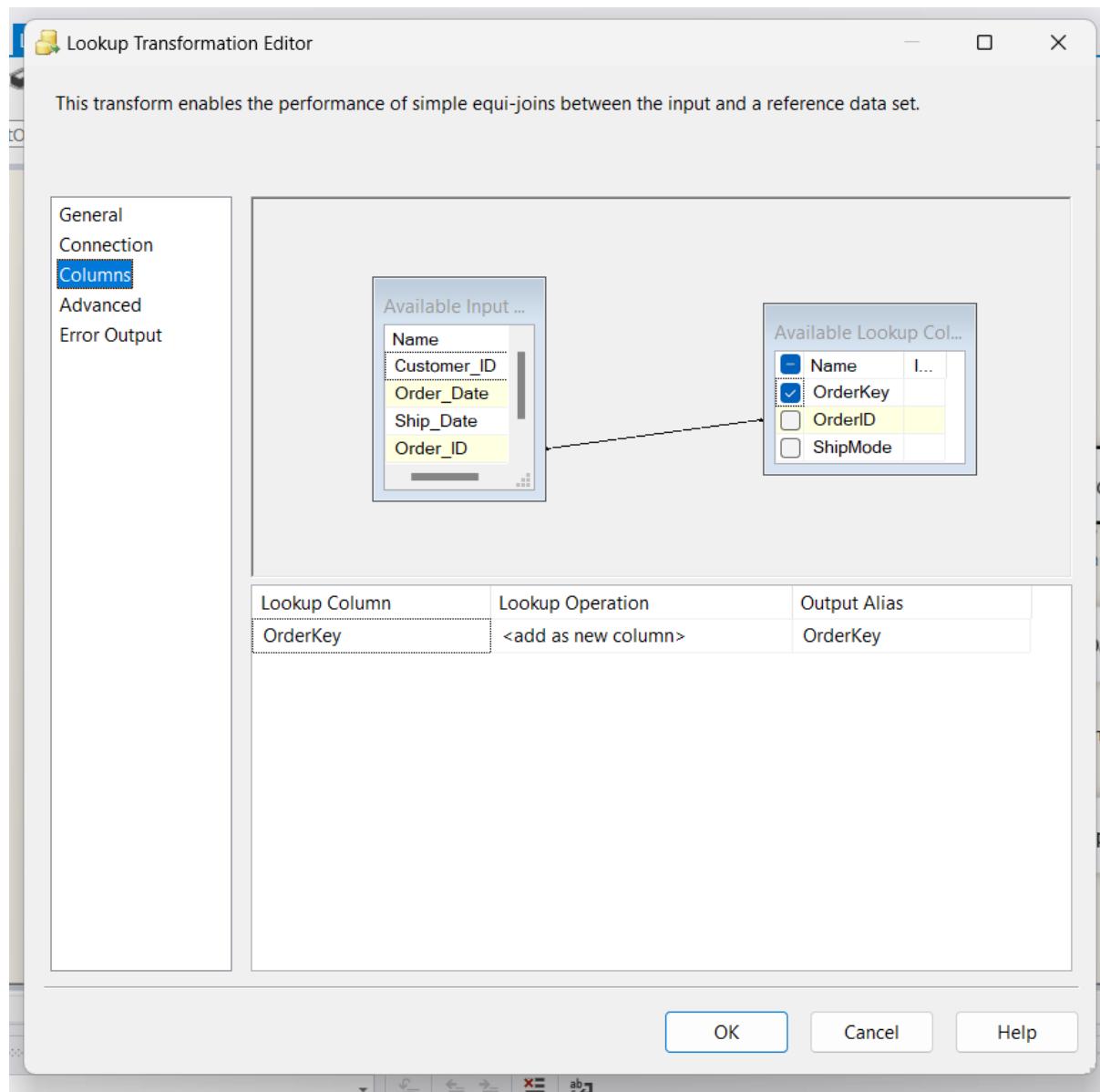
Lookup **DimCustomer** ta sẽ map cột customer\_id trong bảng stgSales với cột CustomerID trong bảng DimCustomer để có thể lấy ra cột CustomerKey



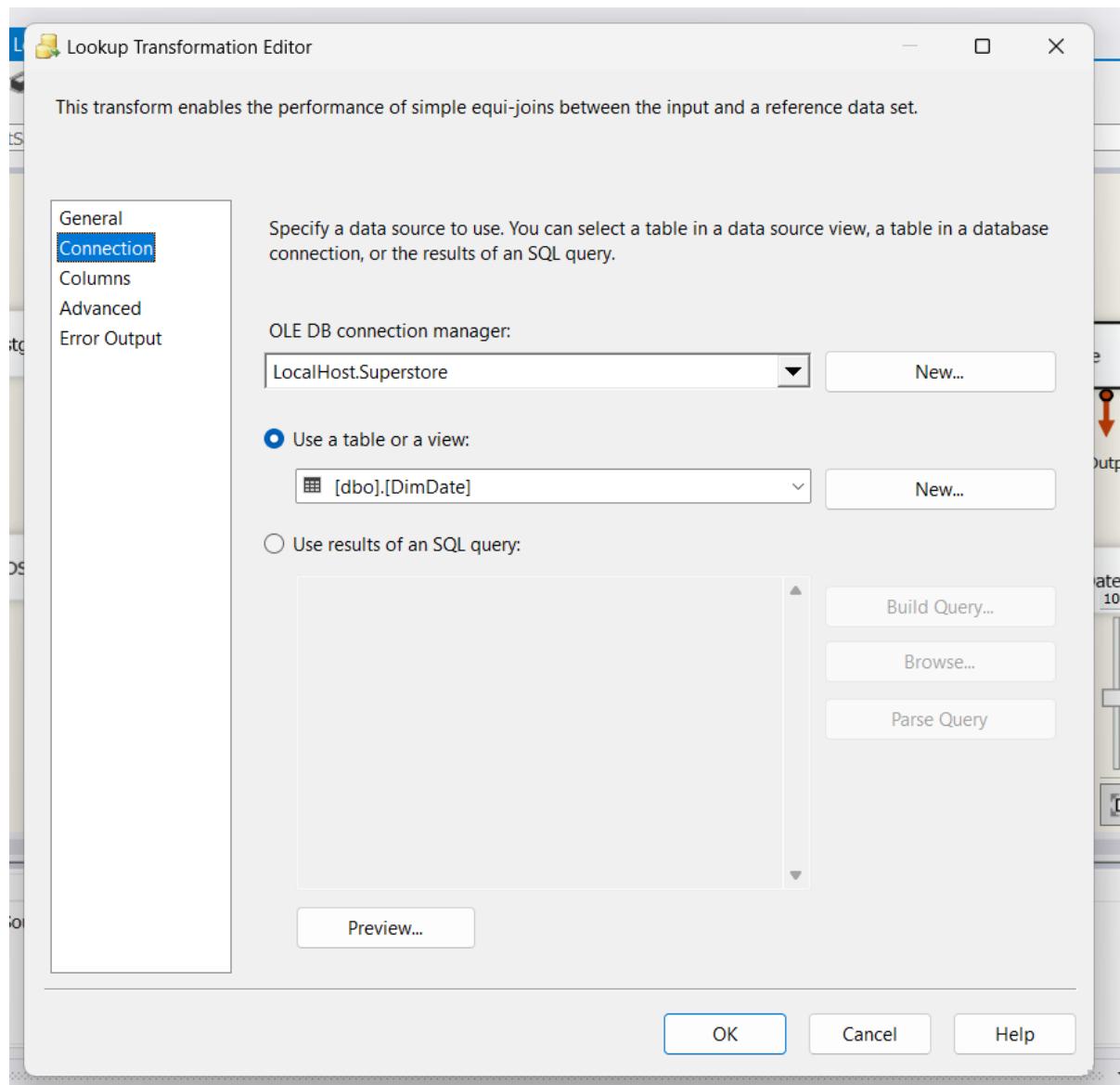


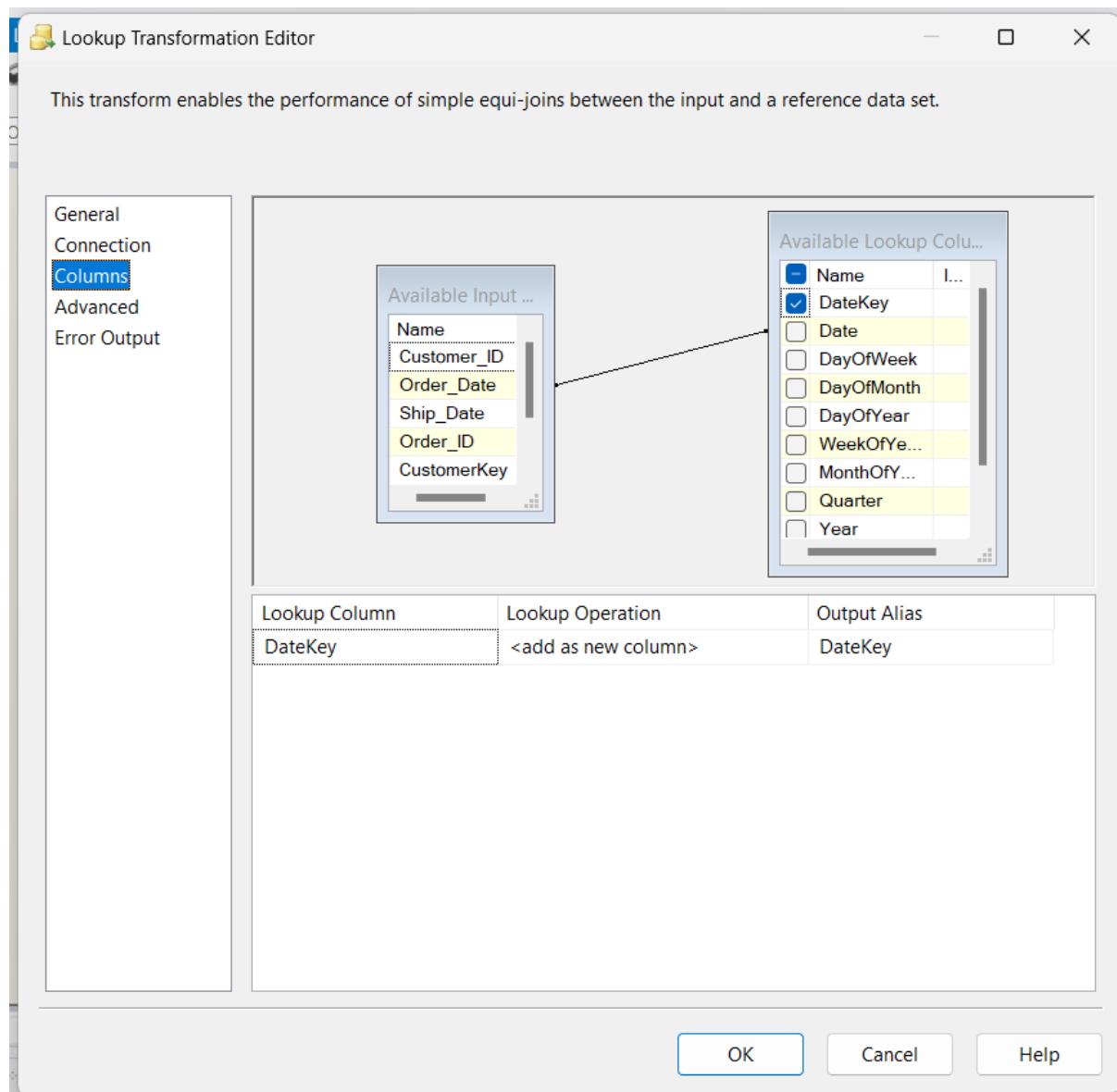
Lookup **DimOrder** ta sẽ map cột order\_id trong bảng stgOrder với cột OrderID trong bảng DimOrder để có thể lấy ra cột OrderKey



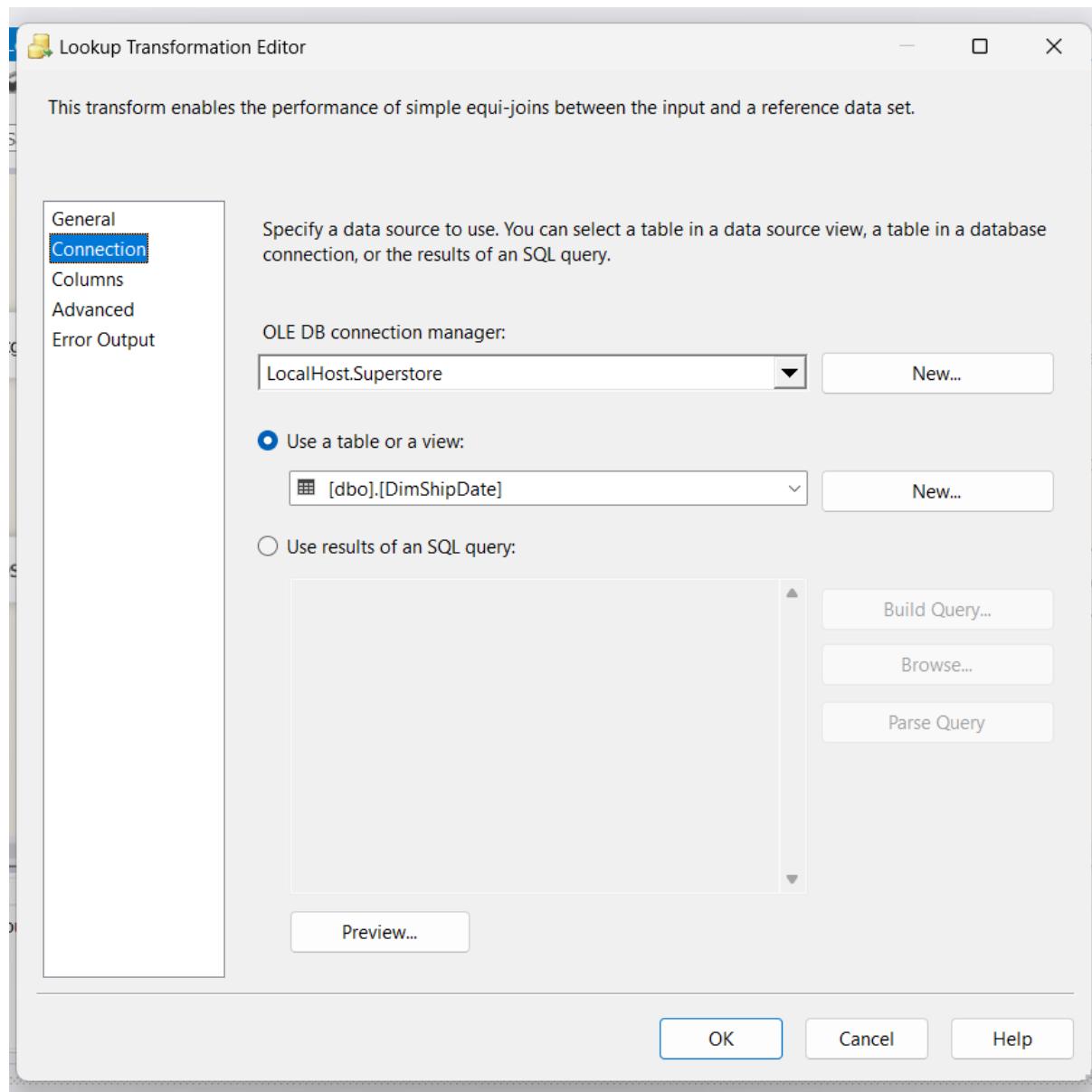


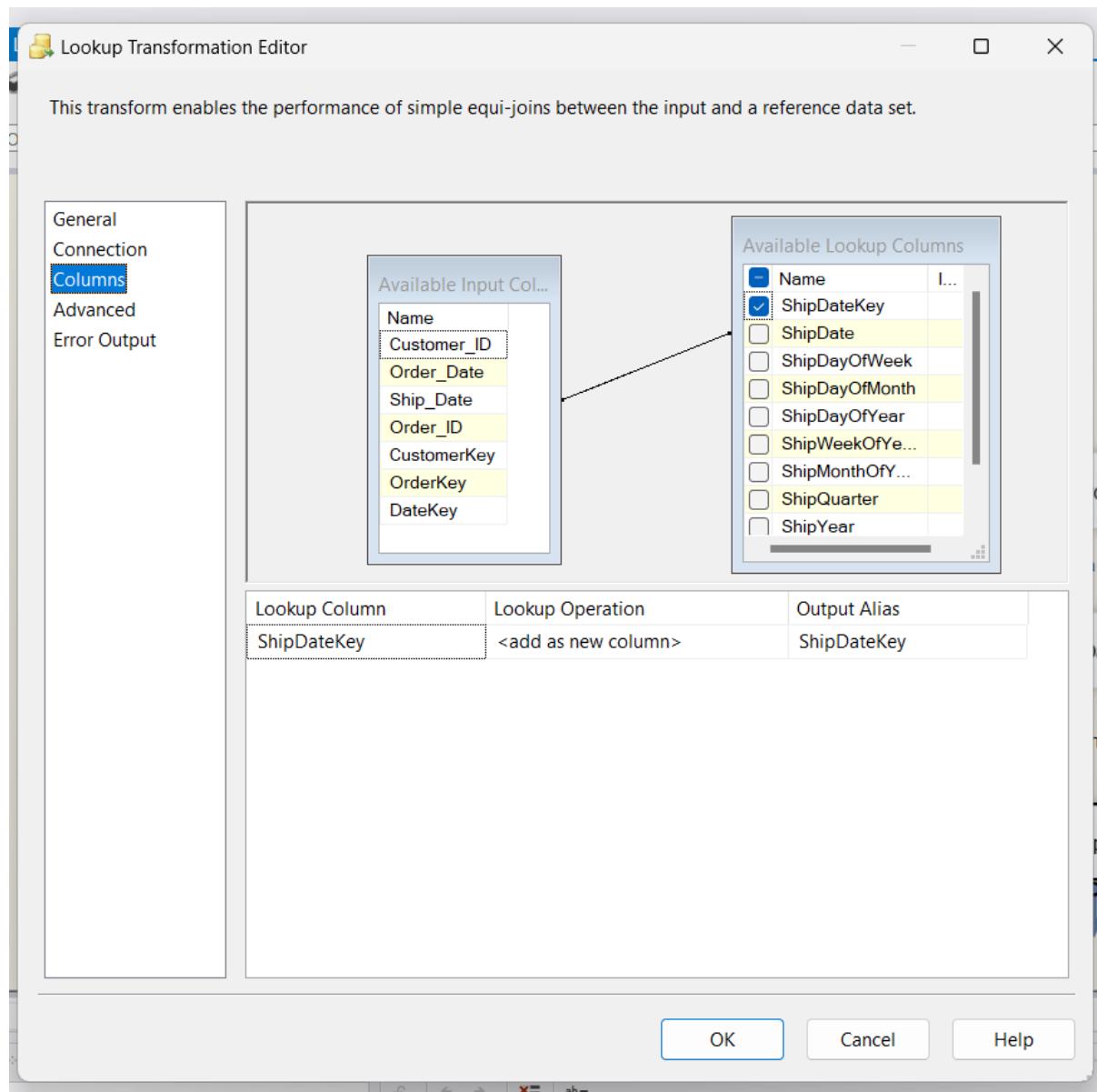
Lookup **DimDate** ta sẽ map cột order\_date ở bảng stgSales với cột Date trong bảng DimDate để lấy ra DateKey



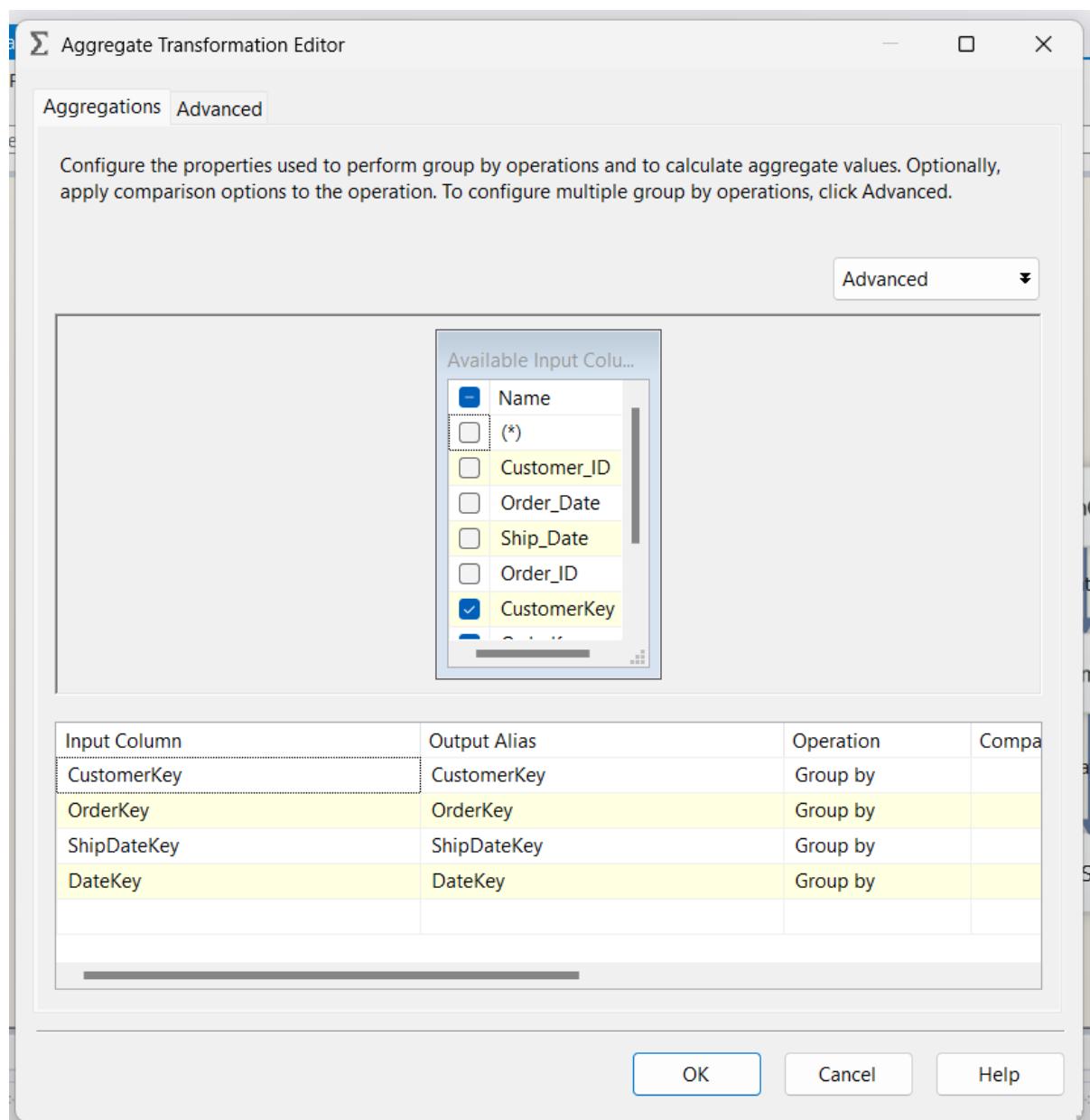


Lookup **DimShipDate** ta sẽ map cột ship\_date ở bảng stgSales với cột ShipDate trong bảng DimShipDate để lấy ra ShipDateKey

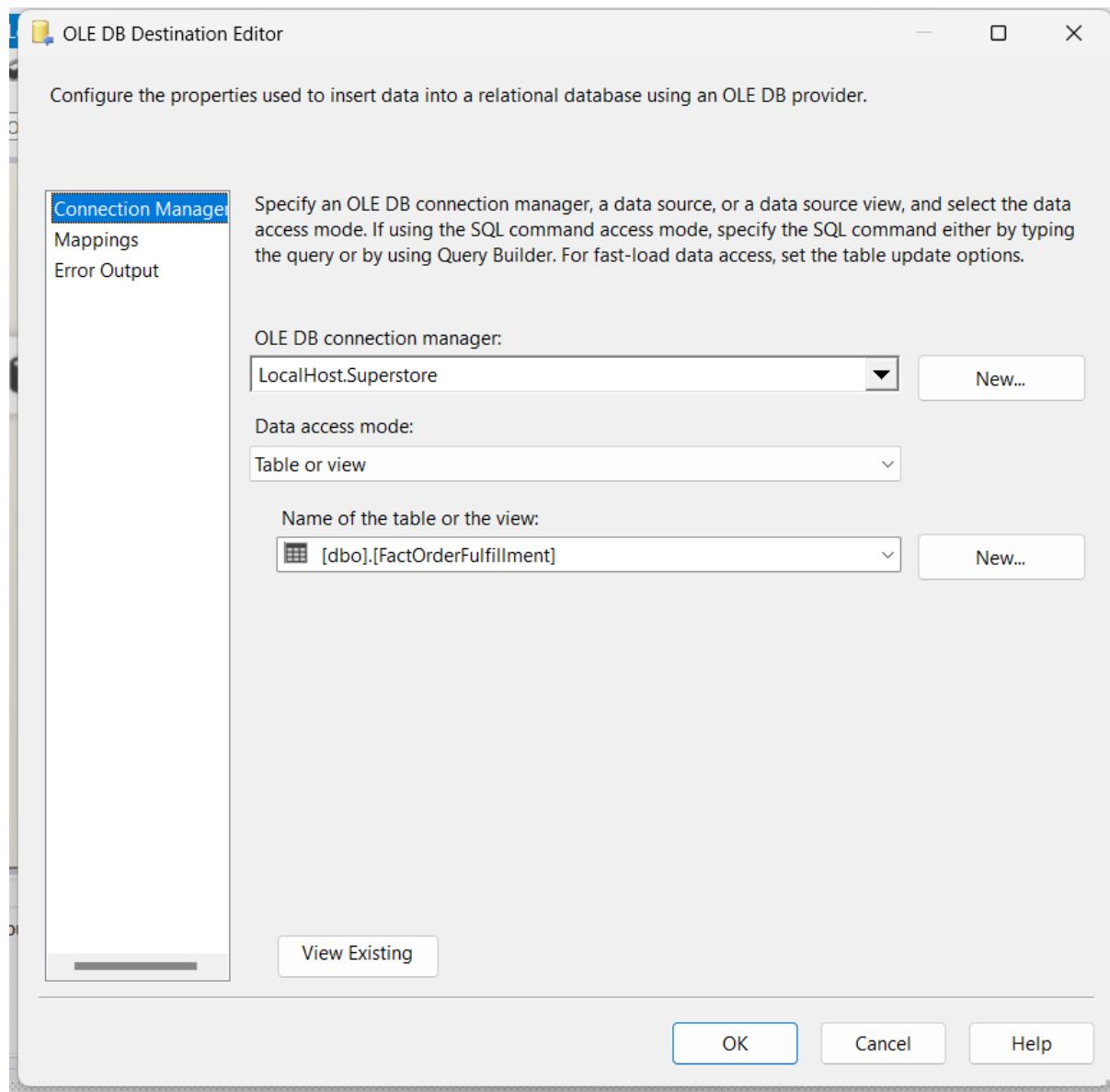


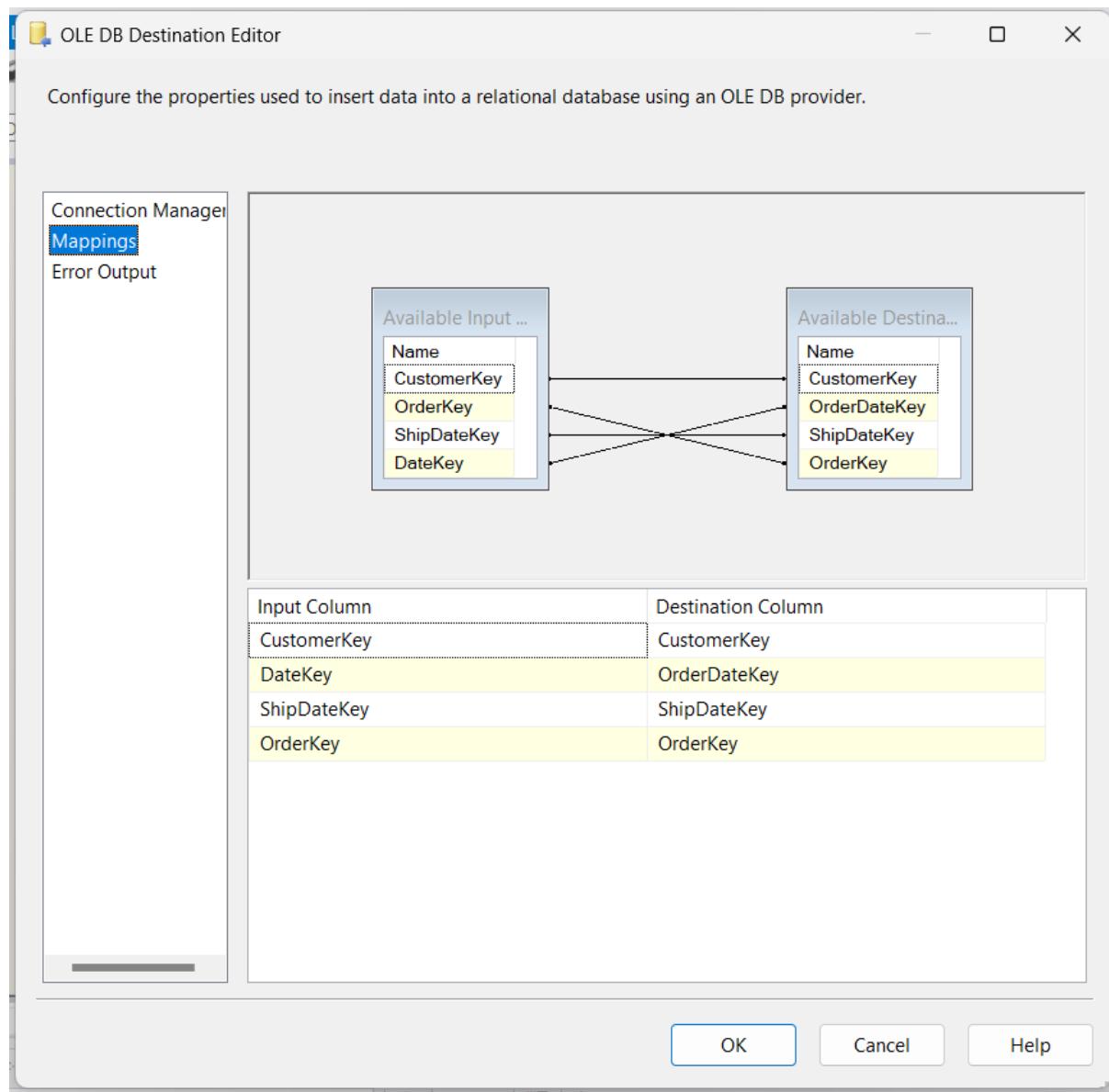


**Aggregate** ta sẽ chỉ lấy những cột cần thiết để đưa vào bảng Fact OrderFulfillment

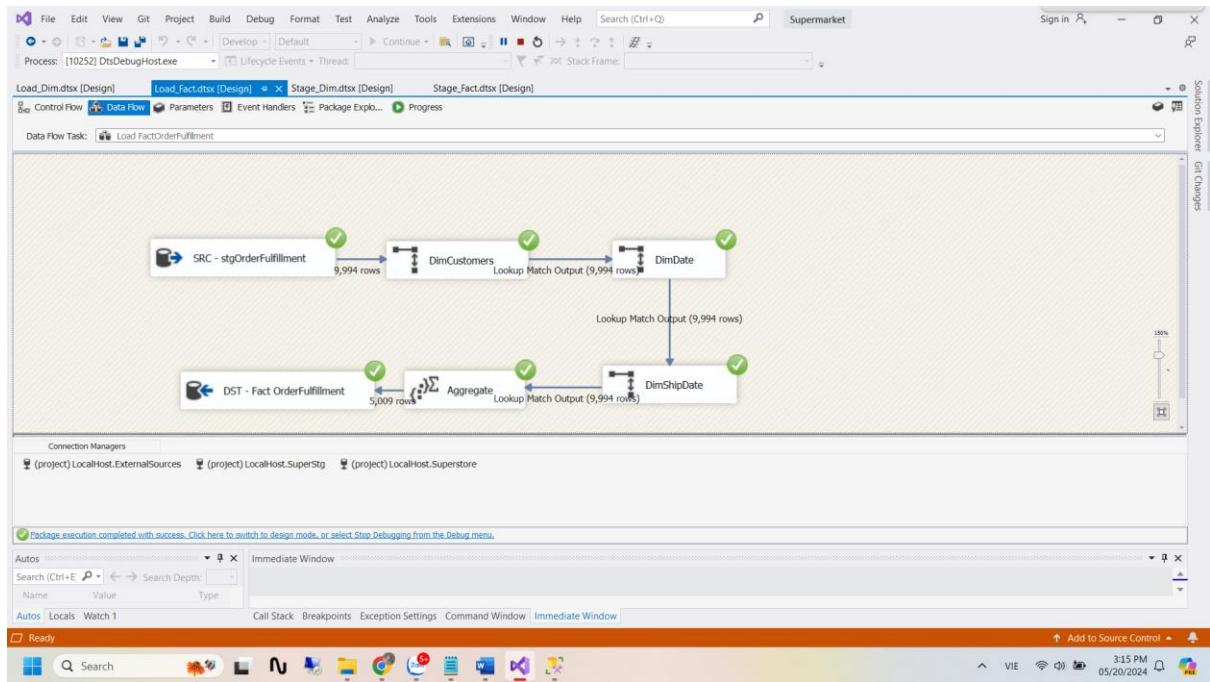


## DST - Fact OrderFulfillment

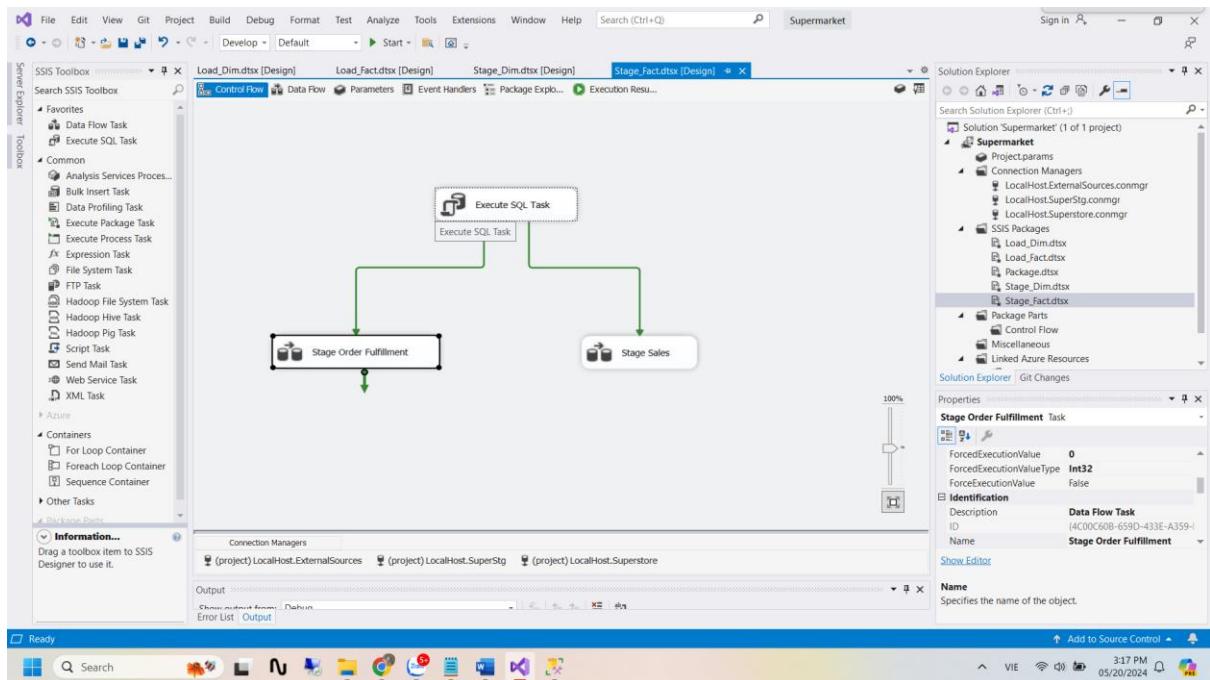




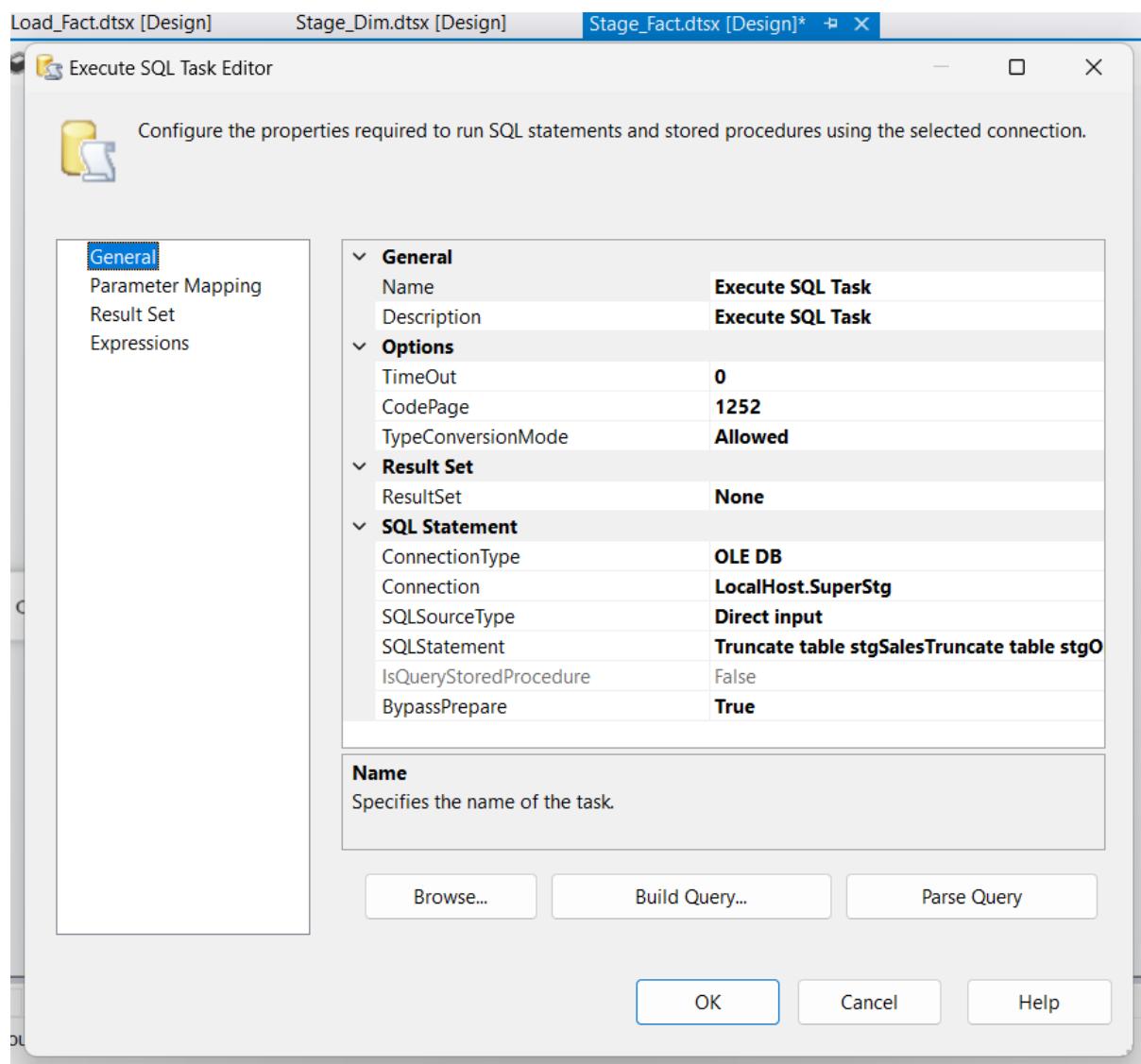
**Khởi chạy**



### 3.3.3. Package Stage\_Fact



## a) Setup Execute SQL Task



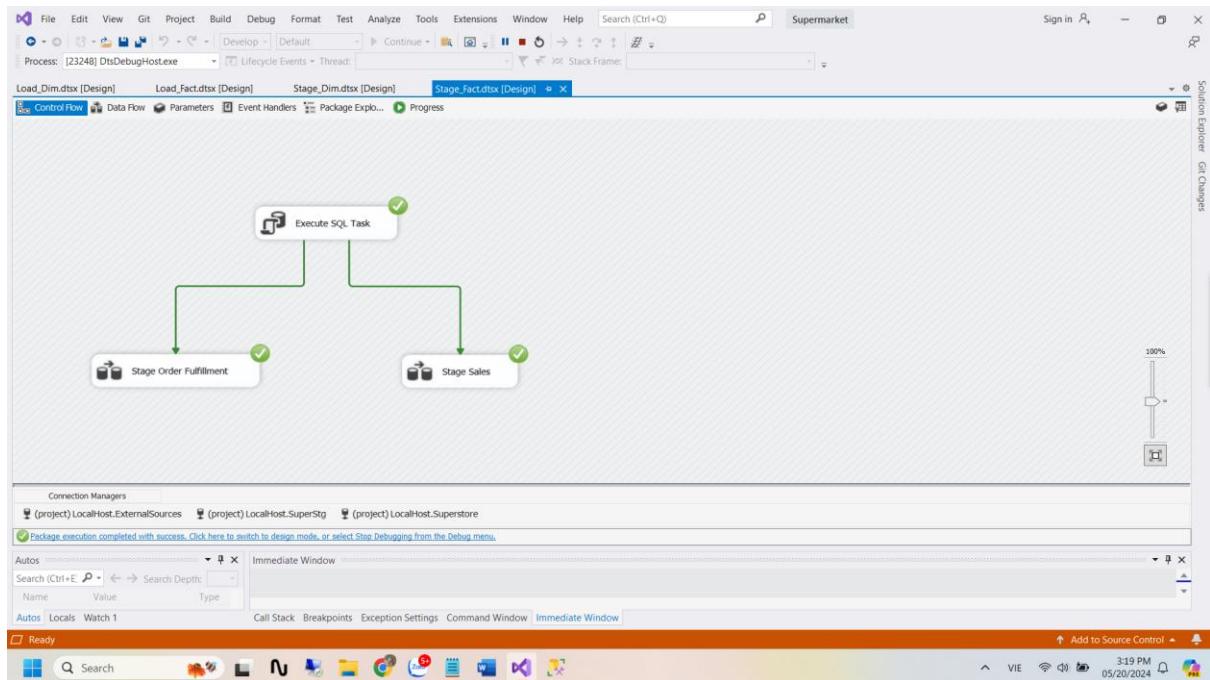
Connection : kết nối đến database SuperStg

SQLStatement :

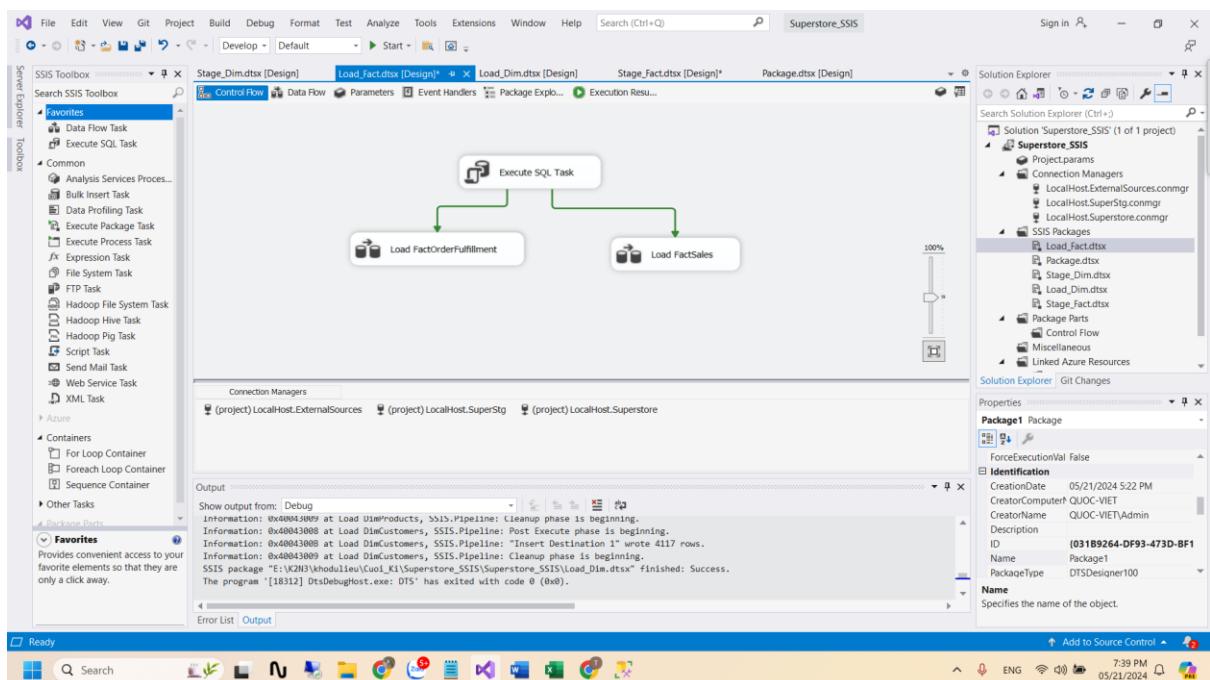
Truncate table stgSales

Truncate table stgOrderFulfillment

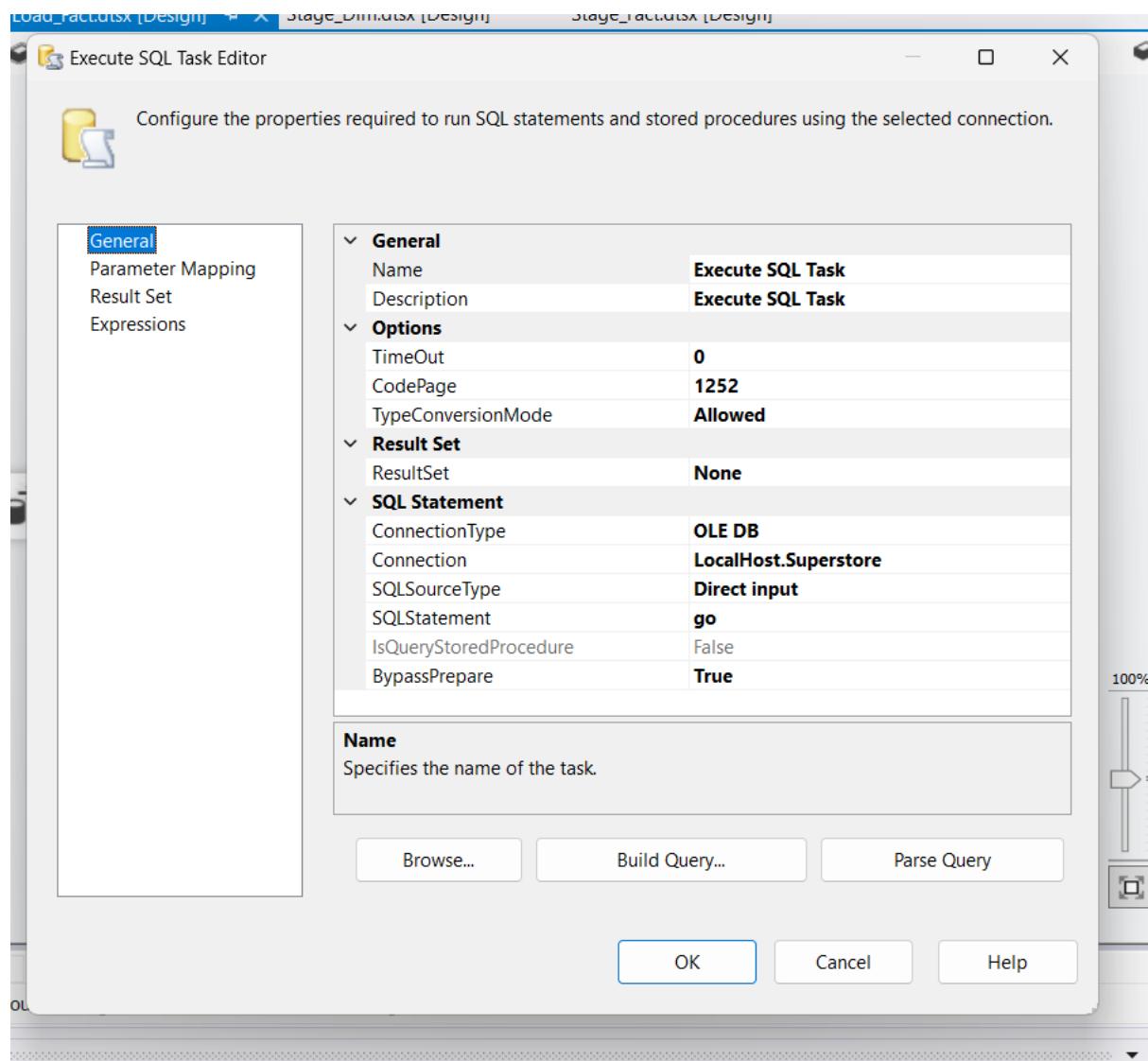
**Khởi chạy**



### 3.3.4. Package Load\_Fact



### a) Setup Execute SQL Task

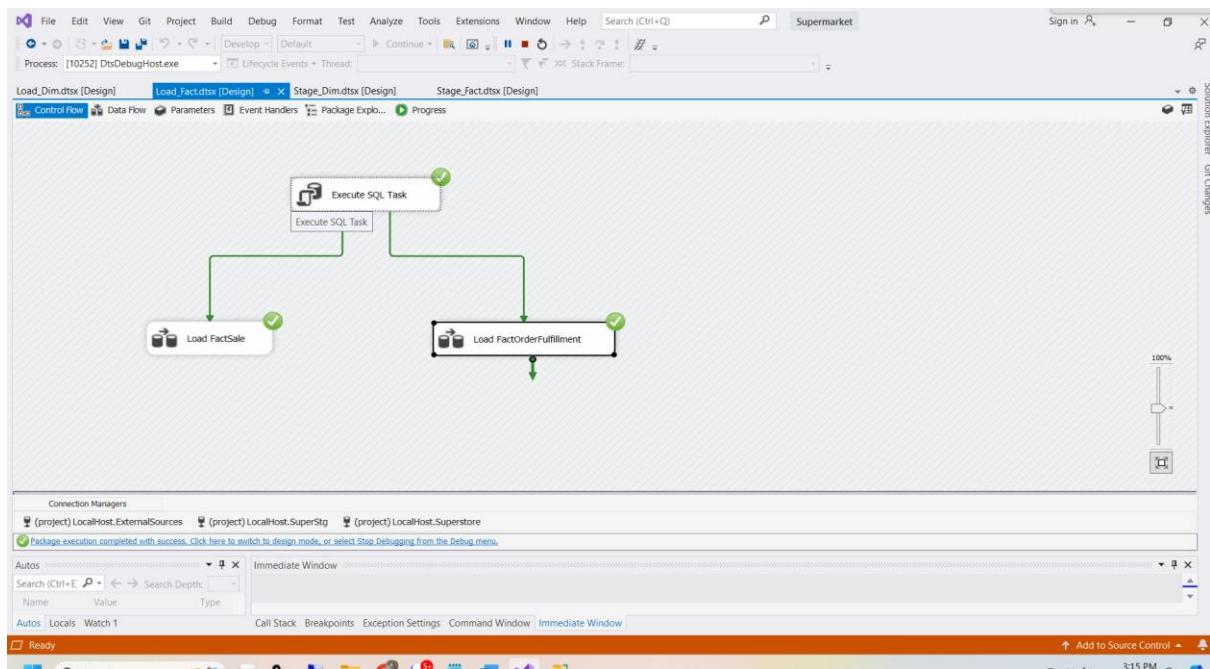


Connection : kết nối đến database Superstore

SQLStatement :

go

**Khởi chạy**



## SSIS hoàn tất

Kiểm tra dữ liệu data base Superstore ở sql :

### DimDate

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

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

Object Explorer

Properties

qry1.dtt

Identity

Query Designer

Top Specification Yes

Dim Order

Properties

qry1.dtt

Identity

Query Designer

Top Specification Yes

DimCustomer

Properties

qry1.dtt

Identity

Query Designer

Top Specification Yes

## Dim Order

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

Object Explorer

Properties

qry1.dtt

Identity

Query Designer

Top Specification Yes

DimCustomer

Properties

qry1.dtt

Identity

Query Designer

Top Specification Yes

## DimCustomer

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

File Edit View Project Query Designer Tools Window Help

New Query Execute

OrderDate

Properties [Query1.dts]

Identity (Name) Query1.dts Database Name Superstore Server Name quoc-viet

Query Designer Destination Table

Distinct Values No GROUP BY Extent <None> Output All Column No

Query Parameter No parameters have SQL Comment \*\*\*\* Script for Select

Top Specification Yes

Object Explorer

Connect to Server

QUOC-VIET (Microsoft Analysis Server 16.0.42.209 - QUOC-VIET) QUOC-VIET (SQL Server 16.0.1000.6 - QUOC-VIET\Admin)

Databases System Databases Database Snapshots BankLoanDB quanhangQA External Sources QLNH QLVATTU QLVLT StoreSalesDW Super SupermarketDW SuperStg Superstore Database Diagrams Tables System Tables File Tables External Tables Graph Tables dbo.DimAudit dbo.DimCustomers dbo.DimDate dbo.DimProduct dbo.DimShipFulfillment dbo.FactSale Dropped Ledger Tables Views External Resources Synonyms Programmability Oracle Store

CustomerK\_ CustomerID CustomerN\_ CustomerSL\_ ContactSug\_ CustomerR\_ CustomerC\_ CustomerC\_ CustomerP\_ RowScurr\_ 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 Butter... 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 Dianne Vitt... New York Consumer East United States Long Beach 11561 False 1899-12-31...

10014 DV-13465 Dianne 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 Gute... Kentuky Consumer South United States Henderson 42429 False 1899-12-31...

Properties [Query1.dts]

Identity (Name) Query1.dts Database Name Superstore Server Name quoc-viet

Query Designer Destination Table

Distinct Values No GROUP BY Extent <None> Output All Column No

Query Parameter No parameters have SQL Comment \*\*\*\* Script for Select

Top Specification Yes

Object Explorer

Connect to Server

QUOC-VIET (Microsoft Analysis Server 16.0.42.209 - QUOC-VIET) QUOC-VIET (SQL Server 16.0.1000.6 - QUOC-VIET\Admin)

Databases System Databases Database Snapshots BankLoanDB quanhangQA External Sources QLNH QLVATTU QLVLT StoreSalesDW Super SupermarketDW SuperStg Superstore Database Diagrams Tables System Tables File Tables External Tables Graph Tables dbo.DimAudit dbo.DimCustomers dbo.DimDate dbo.DimProduct dbo.DimShipFulfillment dbo.FactSale Dropped Ledger Tables Views External Resources Synonyms Programmability Oracle Store

CustomerK\_ CustomerID CustomerN\_ CustomerSL\_ ContactSug\_ CustomerR\_ CustomerC\_ CustomerC\_ CustomerP\_ RowScurr\_ RowStartD...

1 of 200 > >> (i) Cell is Read Only.

Ready

Search

3:26 PM 05/20/2024

## DimShipDate

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

File Edit View Project Query Designer Tools Window Help

New Query Execute

OrderDate

Properties [Query1.dts]

Identity (Name) Query1.dts Database Name Superstore Server Name quoc-viet

Query Designer Destination Table

Distinct Values No GROUP BY Extent <None> Output All Column No

Query Parameter No parameters have SQL Comment \*\*\*\* Script for Select

Top Specification Yes

Object Explorer

Connect to Server

QUOC-VIET (Microsoft Analysis Server 16.0.42.209 - QUOC-VIET) QUOC-VIET (SQL Server 16.0.1000.6 - QUOC-VIET\Admin)

Databases System Databases Database Snapshots BankLoanDB quanhangQA External Sources QLNH QLVATTU QLVLT StoreSalesDW Super SupermarketDW SuperStg Superstore Database Diagrams Tables System Tables File Tables External Tables Graph Tables dbo.DimAudit dbo.DimCustomers dbo.DimDate dbo.DimProduct dbo.DimShipDate dbo.DimShipFulfillment dbo.FactSale Dropped Ledger Tables Views External Resources Synonyms Programmability Oracle Store

ShipDateKey ShipDate ShipDayOf\_ ShipDayOf\_ ShipDayOf\_ ShipWeekOf\_ ShipMonth\_ ShipQuarter\_ ShipYear\_ ShipsWeek...

20140107 2014-01-07 3 8 8 2 1 1 2014 Y

20140108 2014-01-08 4 8 8 2 1 1 2014 Y

20140110 2014-01-10 6 10 10 2 1 1 2014 Y

20140112 2014-01-12 1 12 12 3 1 1 2014 N

20140113 2014-01-13 2 13 13 3 1 1 2014 Y

20140114 2014-01-14 3 14 14 3 1 1 2014 Y

20140115 2014-01-15 4 15 15 3 1 1 2014 Y

20140116 2014-01-16 5 16 16 3 1 1 2014 Y

20140117 2014-01-17 6 17 17 3 1 1 2014 Y

20140118 2014-01-18 7 18 18 3 1 1 2014 N

20140120 2014-01-20 2 20 20 4 1 1 2014 Y

20140121 2014-01-21 3 21 21 4 1 1 2014 Y

20140123 2014-01-23 5 23 23 4 1 1 2014 Y

20140125 2014-01-25 7 25 25 4 1 1 2014 N

20140126 2014-01-26 1 26 26 5 1 1 2014 N

20140127 2014-01-27 2 27 27 5 1 1 2014 Y

20140128 2014-01-28 3 28 28 5 1 1 2014 Y

20140129 2014-01-29 4 29 29 5 1 1 2014 Y

20140131 2014-01-31 6 31 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 [Query1.dts]

Identity (Name) Query1.dts Database Name Superstore Server Name quoc-viet

Query Designer Destination Table

Distinct Values No GROUP BY Extent <None> Output All Column No

Query Parameter No parameters have SQL Comment \*\*\*\* Script for Select

Top Specification Yes

Object Explorer

Connect to Server

QUOC-VIET (Microsoft Analysis Server 16.0.42.209 - QUOC-VIET) QUOC-VIET (SQL Server 16.0.1000.6 - QUOC-VIET\Admin)

Databases System Databases Database Snapshots BankLoanDB quanhangQA External Sources QLNH QLVATTU QLVLT StoreSalesDW Super SupermarketDW SuperStg Superstore Database Diagrams Tables System Tables File Tables External Tables Graph Tables dbo.DimAudit dbo.DimCustomers dbo.DimDate dbo.DimProduct dbo.DimShipDate dbo.DimShipFulfillment dbo.FactSale Dropped Ledger Tables Views External Resources Synonyms Programmability Oracle Store

ShipDateKey ShipDate ShipDayOf\_ ShipDayOf\_ ShipDayOf\_ ShipWeekOf\_ ShipMonth\_ ShipQuarter\_ ShipYear\_ ShipsWeek...

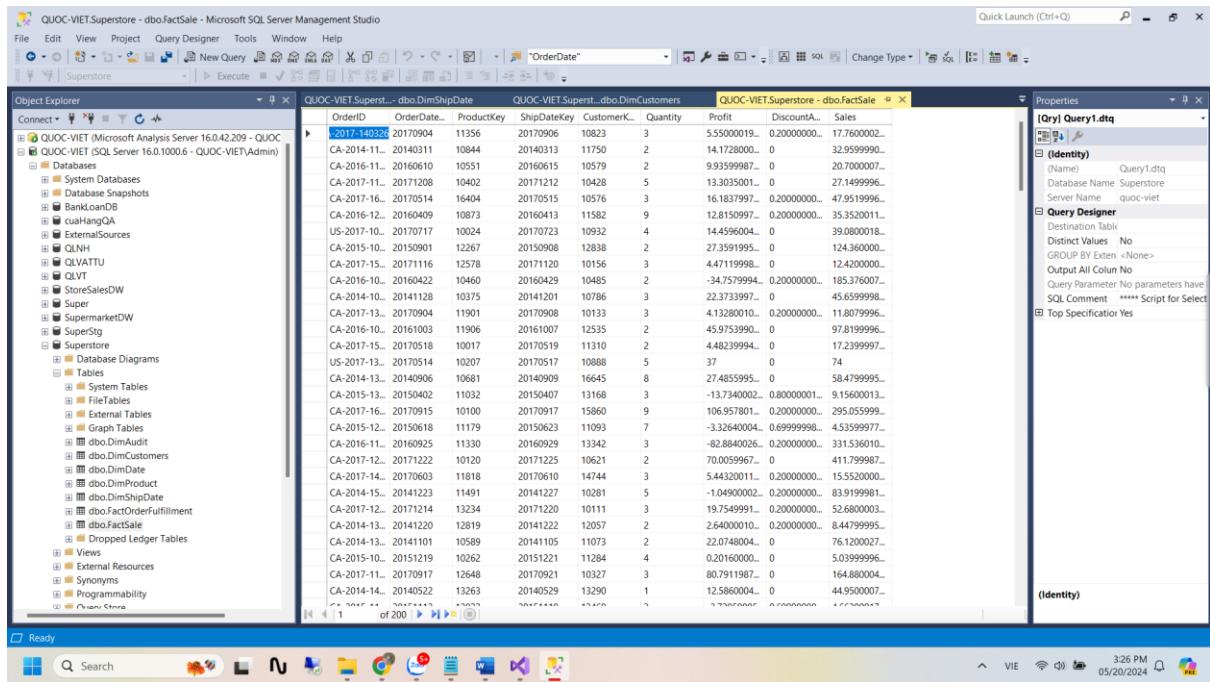
1 of 200 > >> (i) Cell is Read Only.

Ready

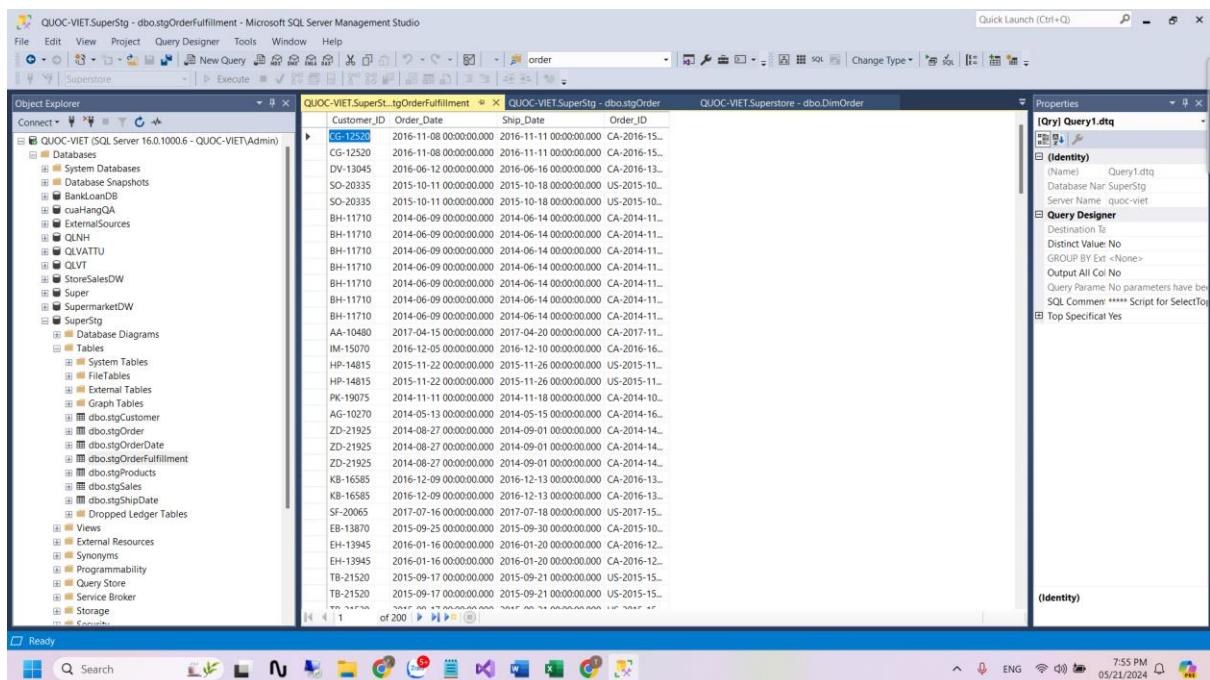
Search

3:26 PM 05/20/2024

## FactSales



## Fact OrderFulfillment

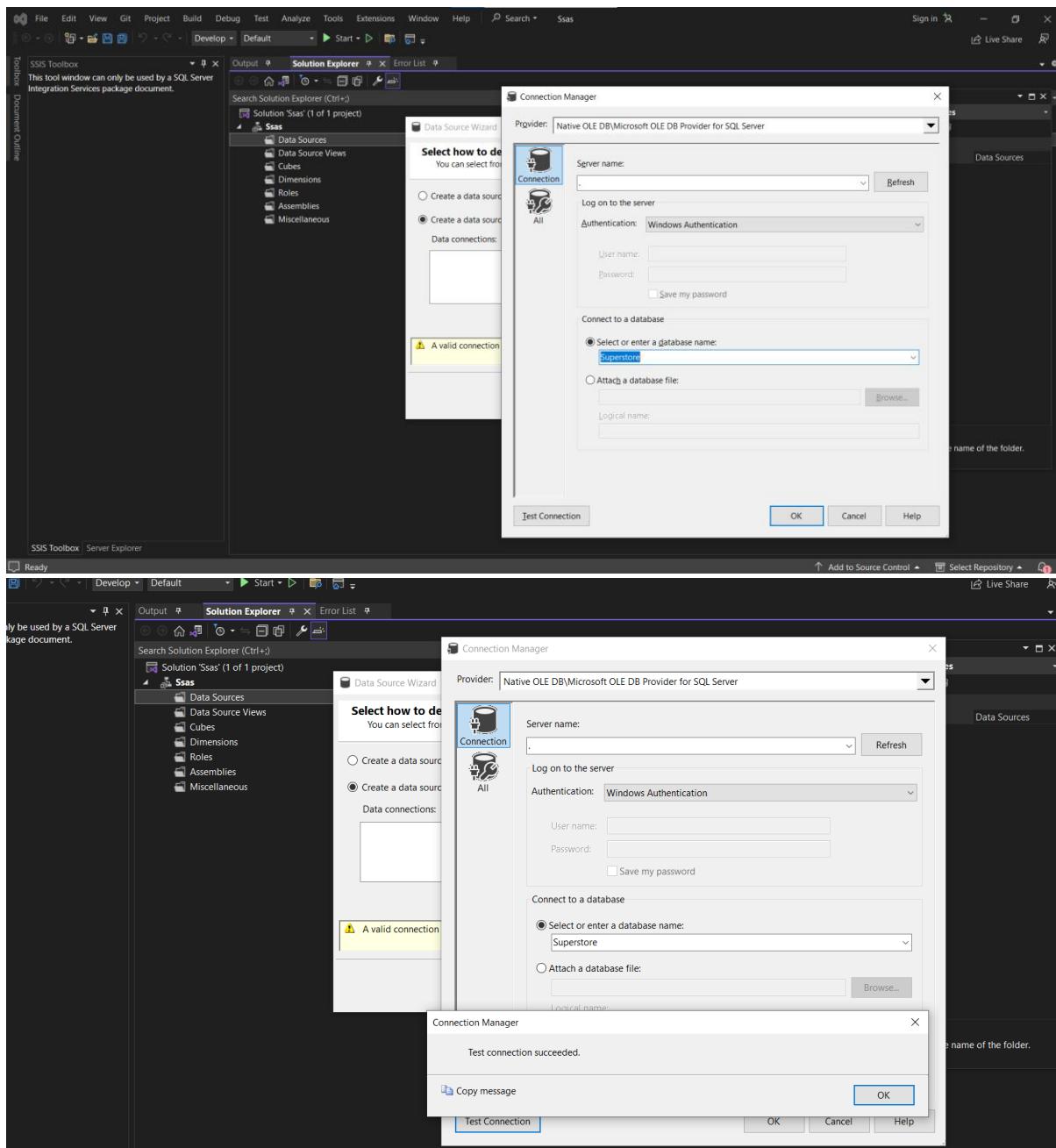


# CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS)

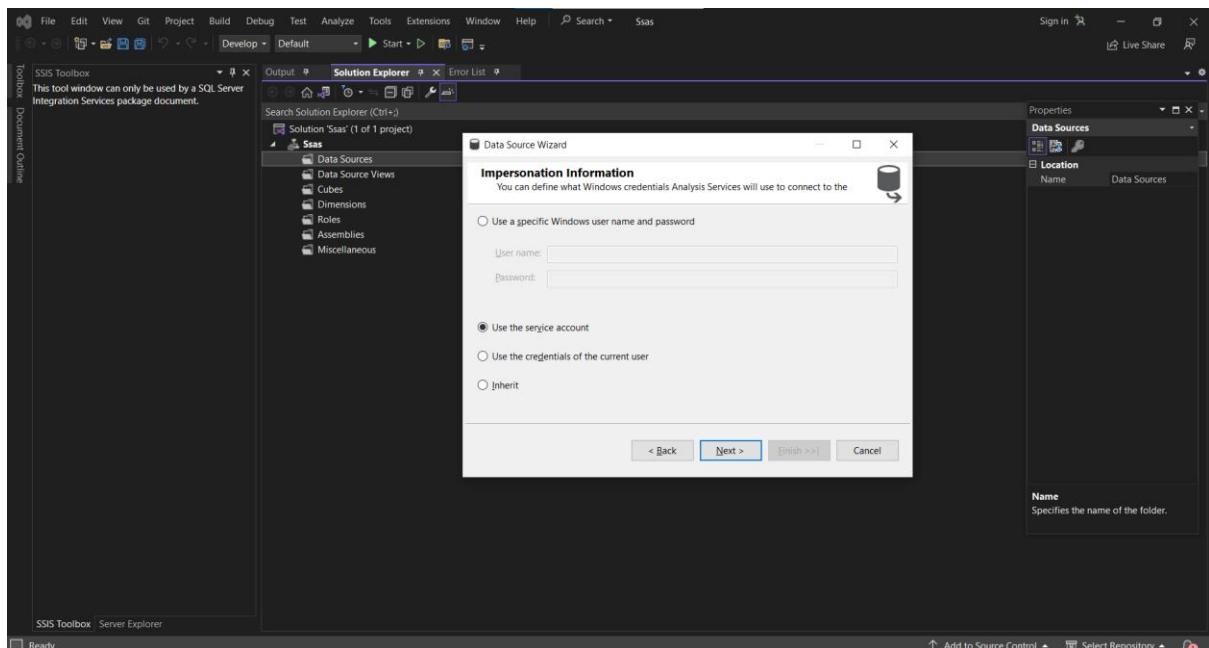
## 4.1. Quá trình xây dựng mô hình

### 4.1.1. Tạo Data Source

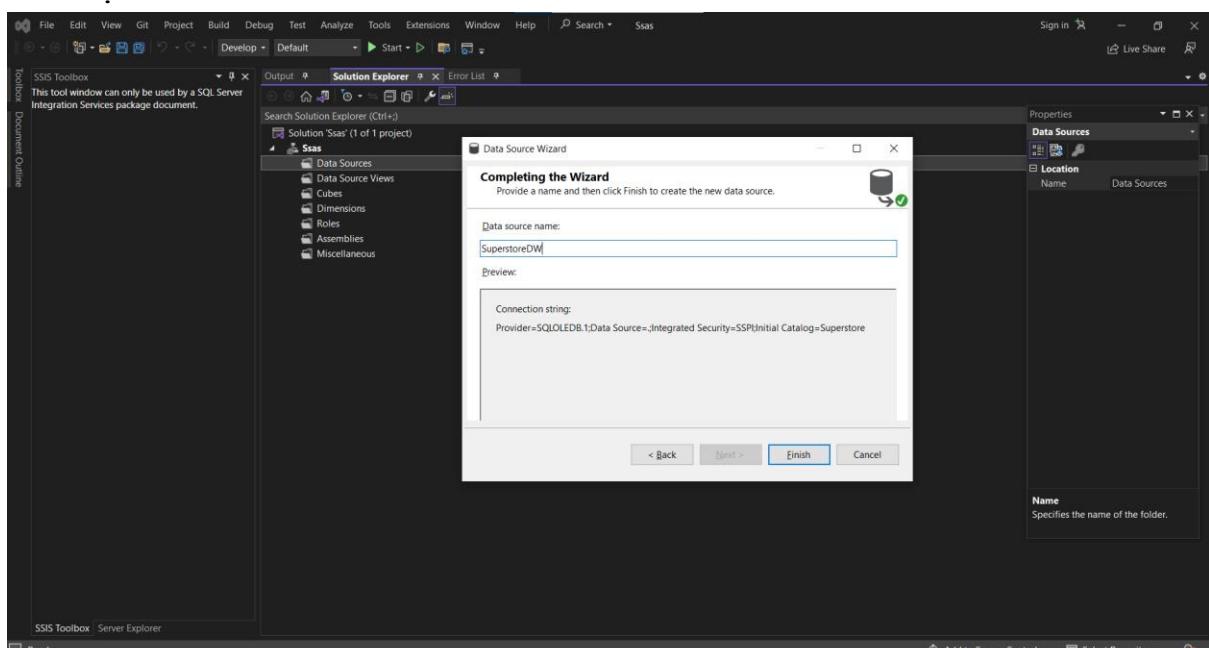
- Right-click vào Data Sources ở khung Solution Explorer. Chon New Data Source để tạo Data Source mới. Kết nối đến Database Superstore



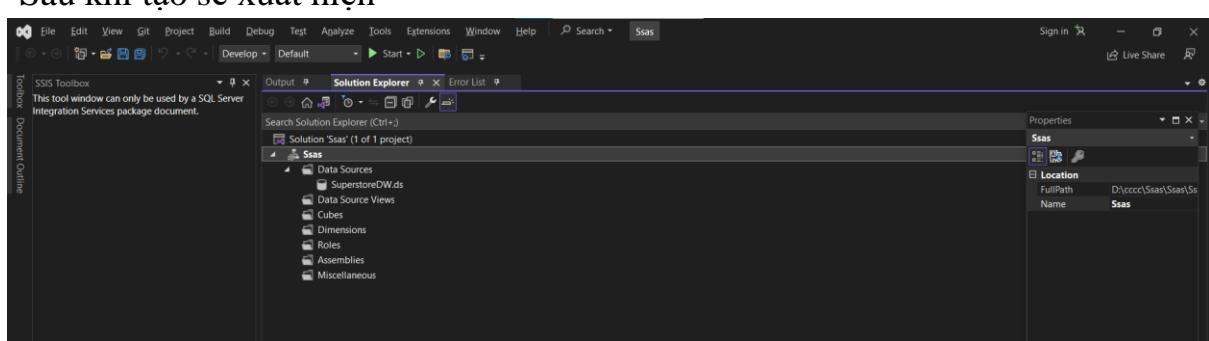
- Chọn Use the service account



- Chọn tên source và bấm Finish

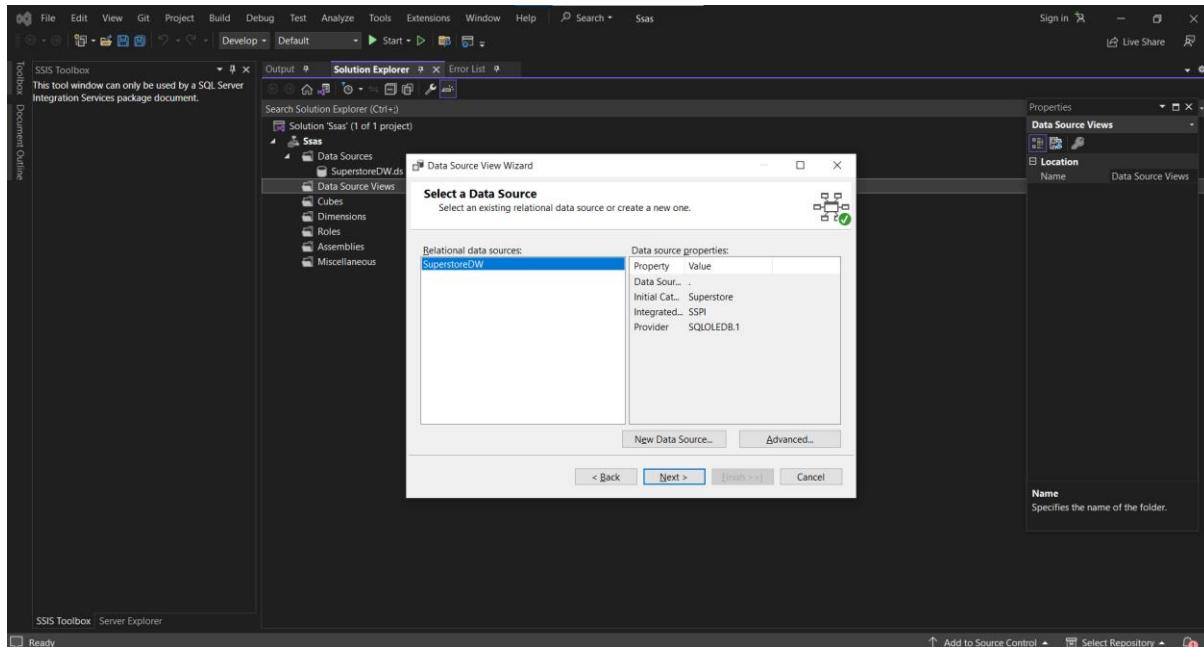


- Sau khi tạo sẽ xuất hiện

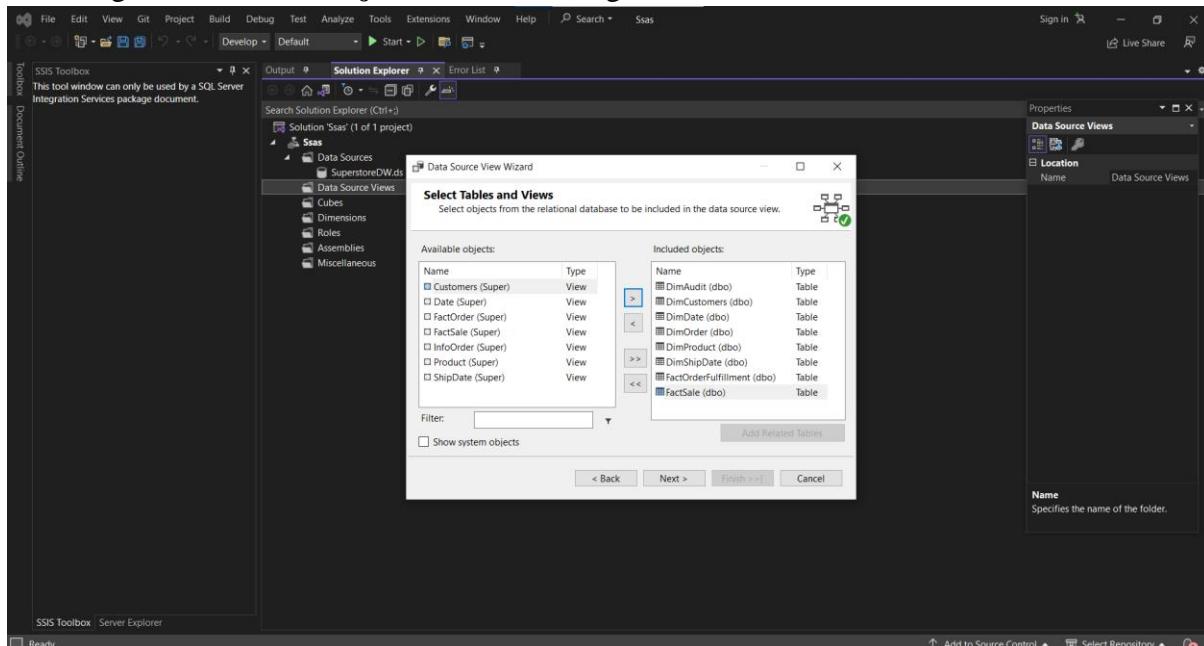


#### 4.1.2. Tạo Data Source View

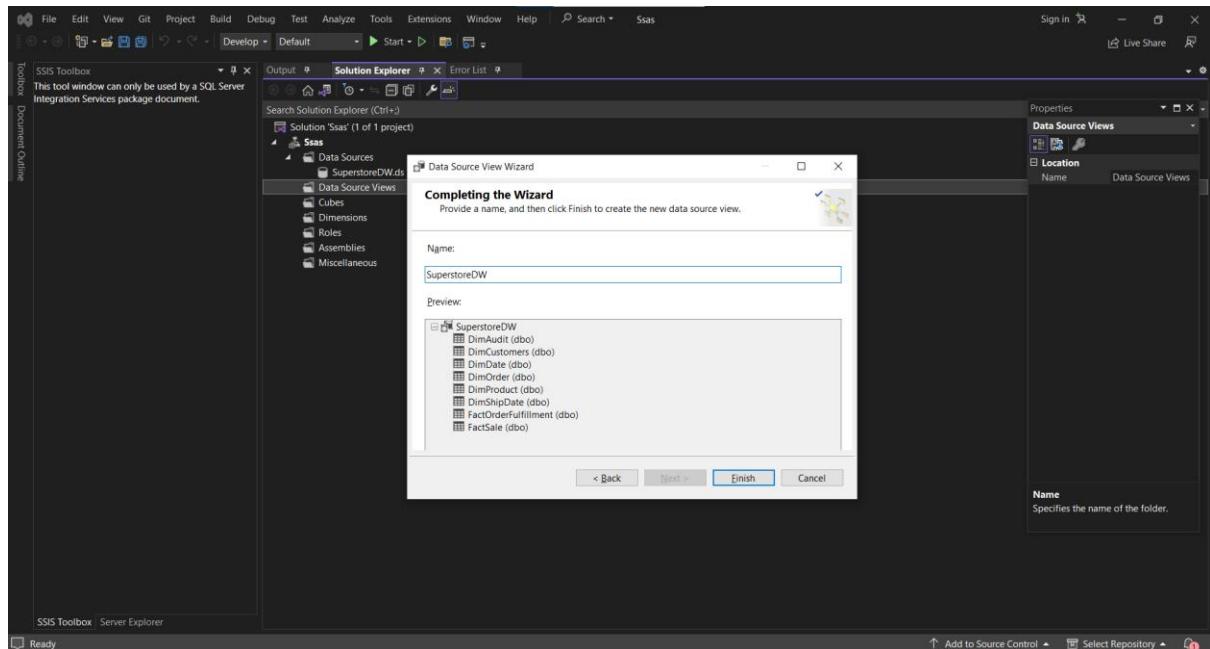
- Right-click vào Data Source Views ở khung Solution Explorer. Chon New Data Source View để tạo Data Source View mới. Chọn Data Soure là **SuperstoreDW** và chọn **Next**.



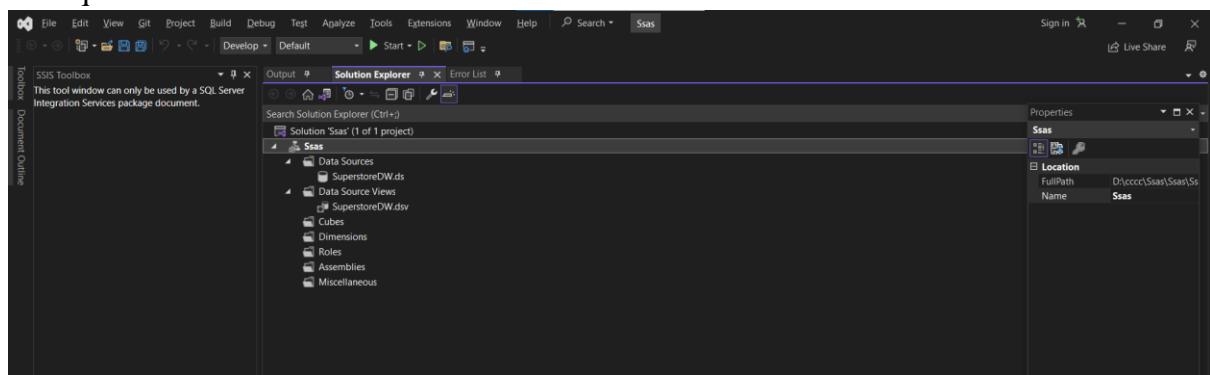
- Trong mục **Included objects**, chọn các bảng Dim và Fact cần thiết, nhấn **Next**



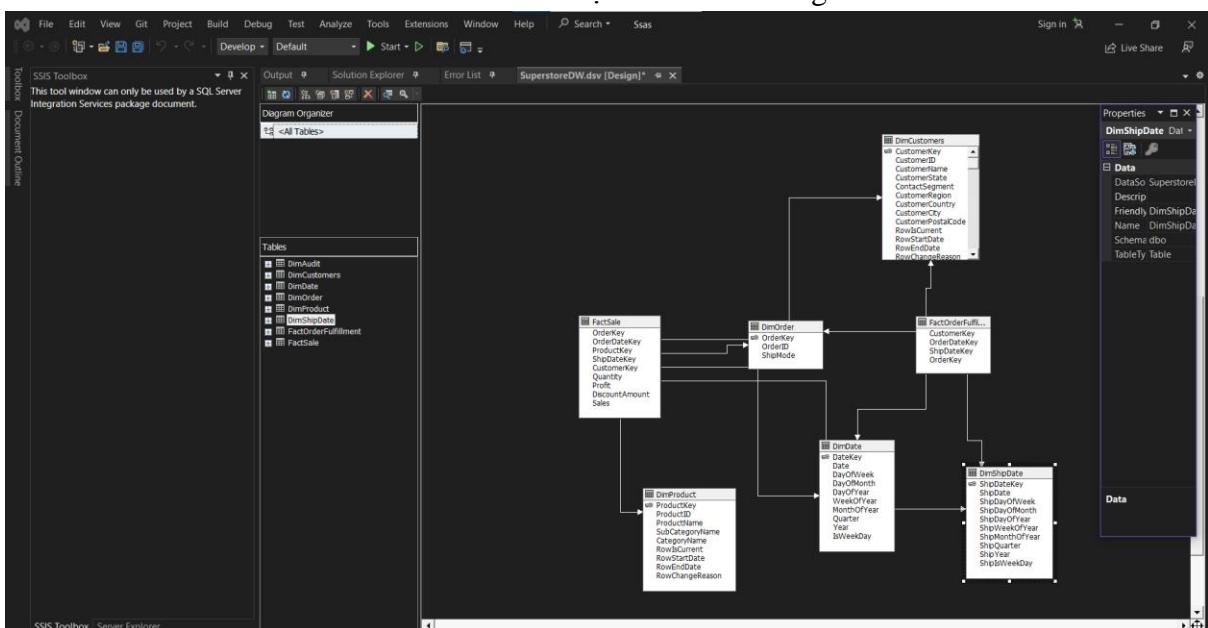
- Đặt tên cho Data Source View và nhấn **Finish** để hoàn tất việc tạo Data Source View



- Kết quả sau Finish

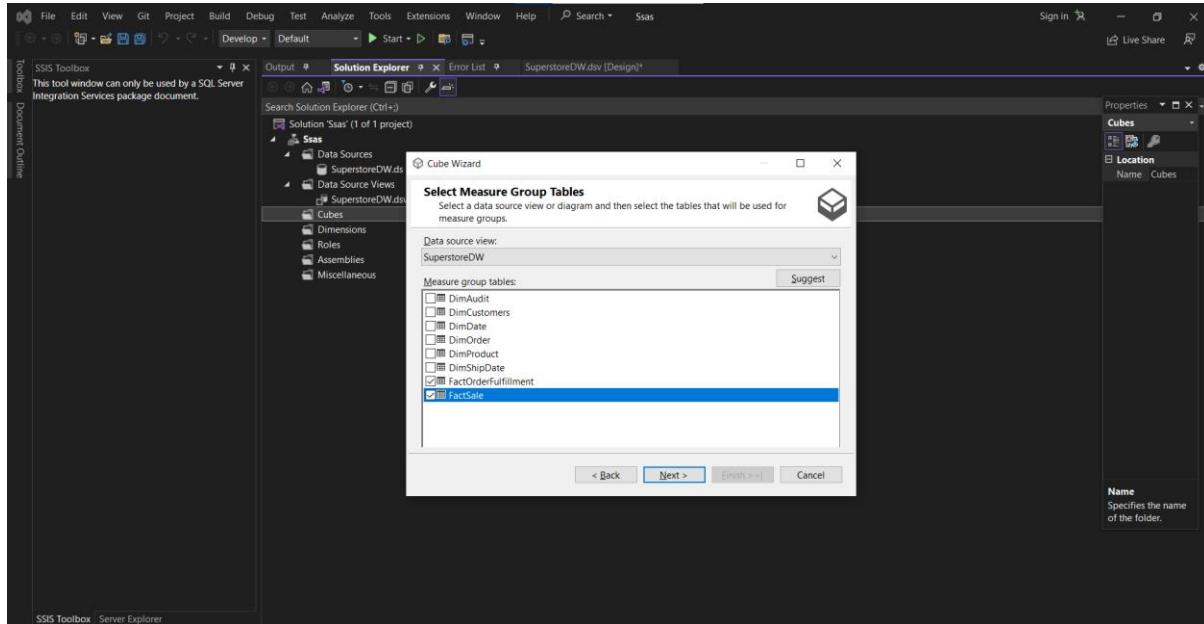


- Double-Click vào Data Source View vừa tạo để hiển thị diagram

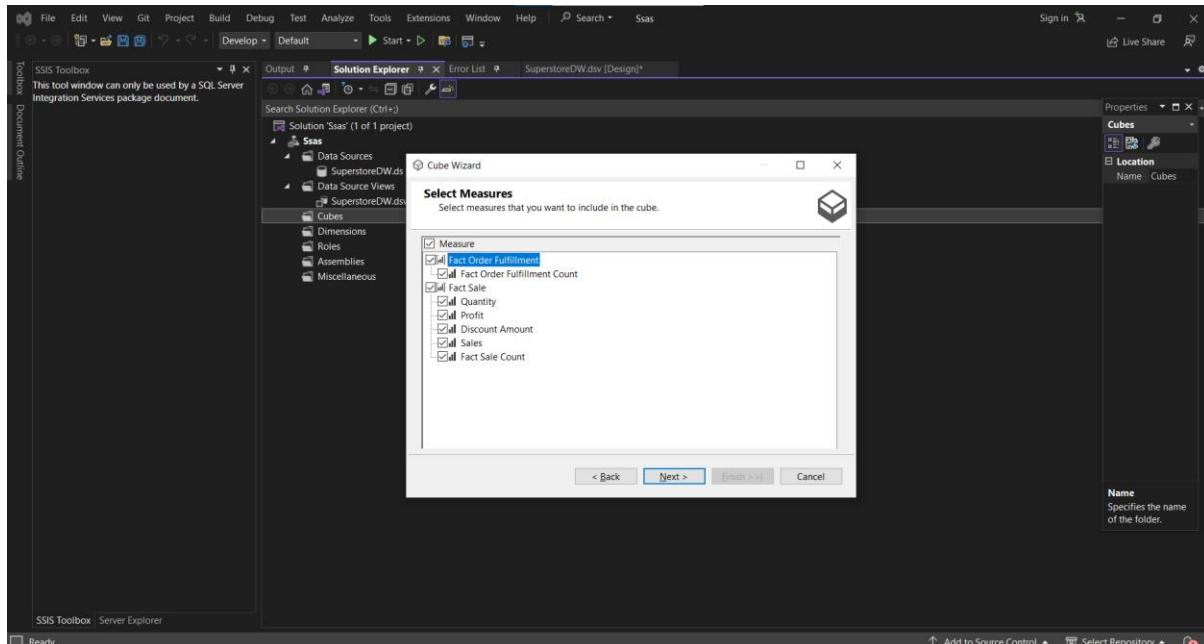


## 4.2. Quá trình xây dựng khối Cube

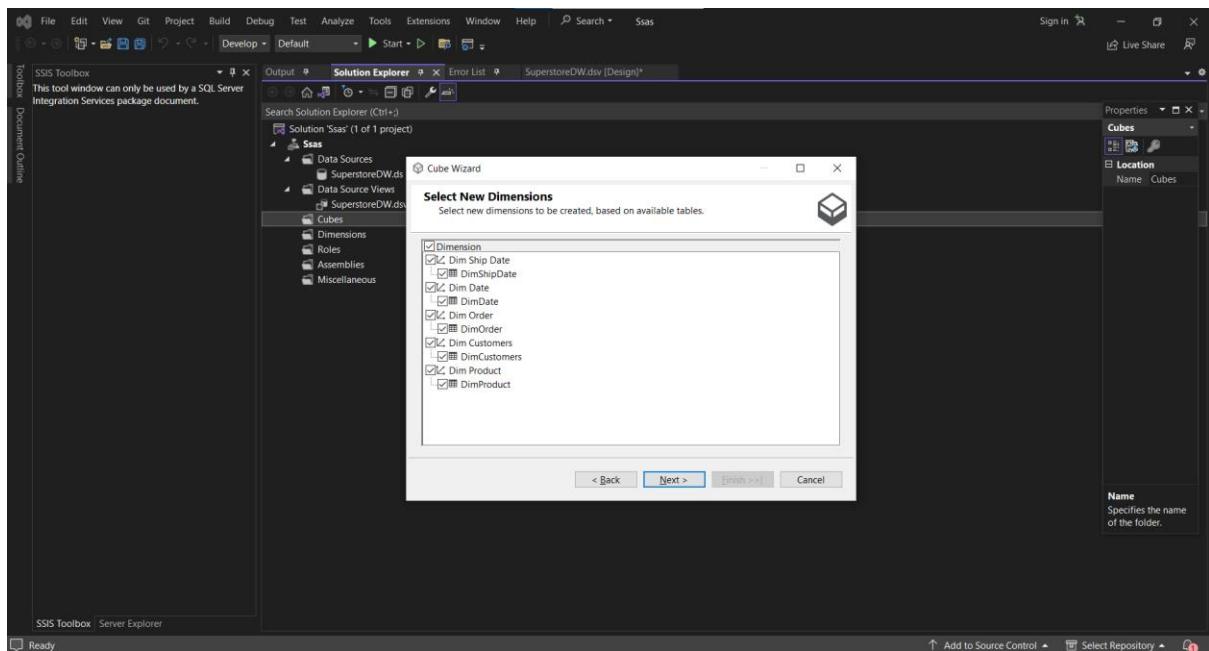
- Right-click vào **Cubes** ở khung Solution Explorer. Chon New Cube để tạo Cube mới.
- Chọn Data Source View là **SuperstoreDW**. Tích chọn 2 bảng **FactSale** và **FactOrderFulfilment**, nhấn Next



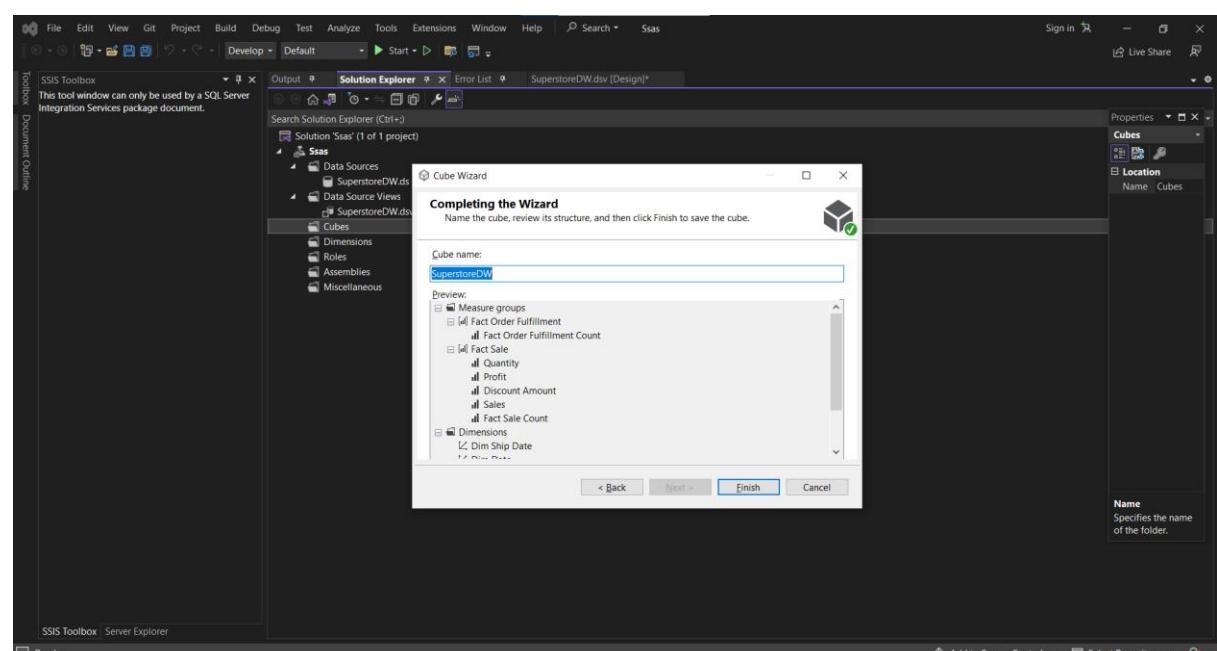
- Ở dialog Select Measures, tích chọn những measure cần thiết các mục sau



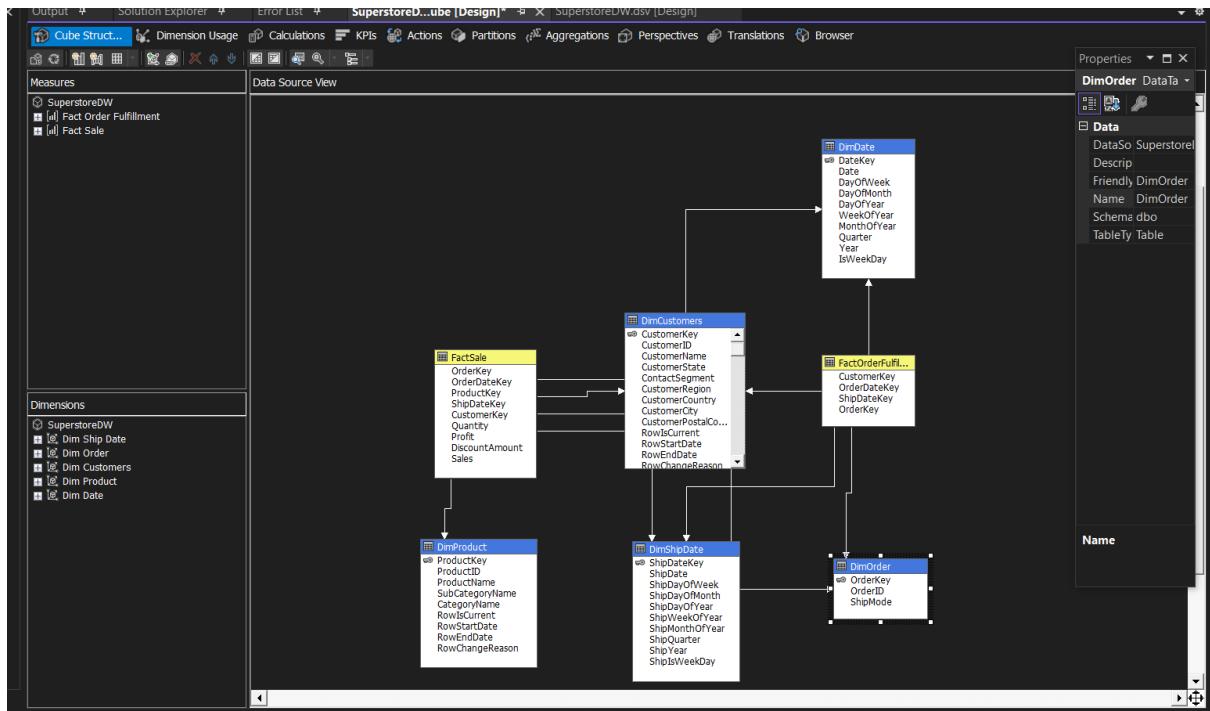
- Khởi tạo các bảng Dimension



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



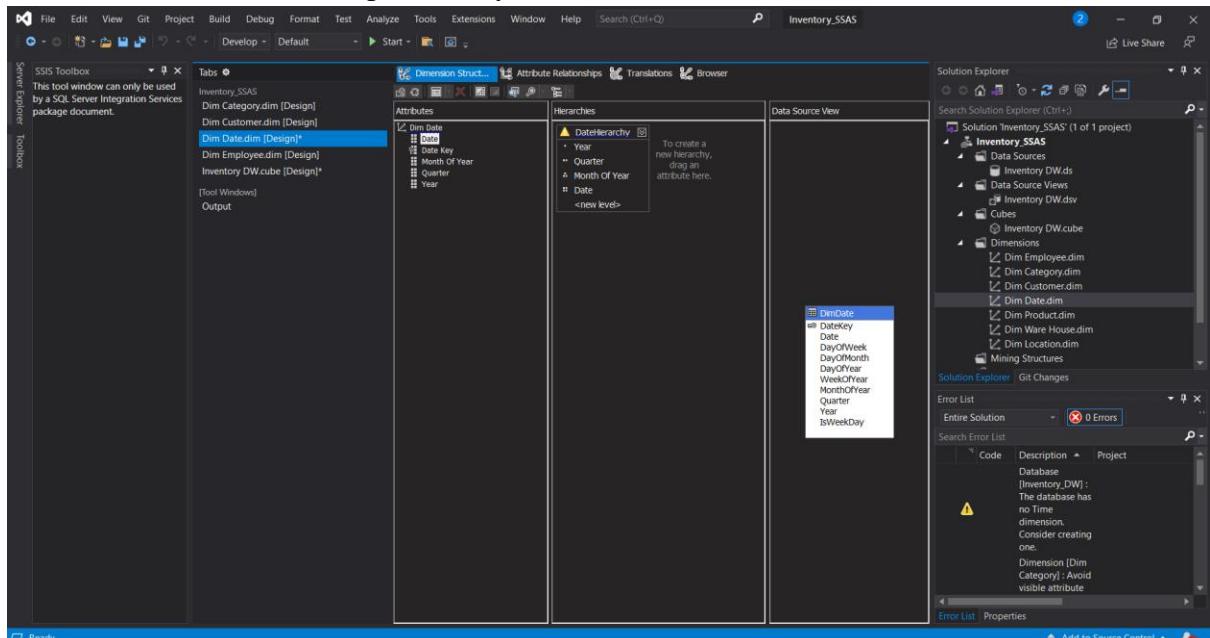
- Các bảng Fact, Dimension và quan hệ giữa chúng sẽ được thể hiện trong Diagram



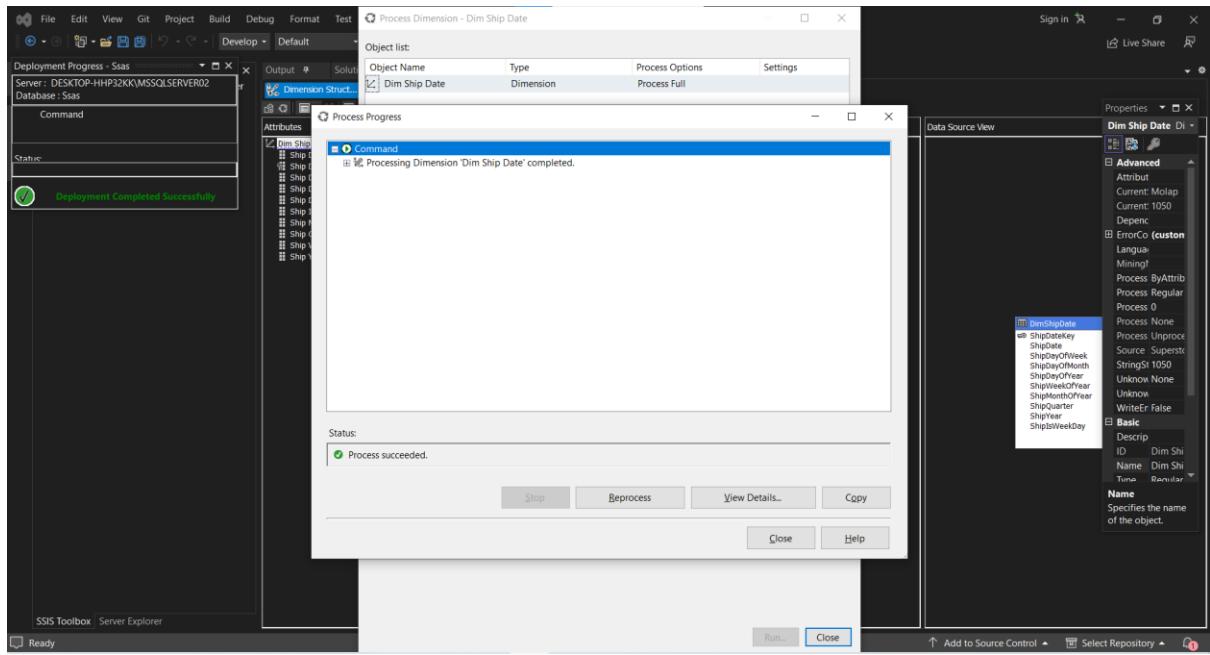
## 4.3. Cấu hình Hierarchy

### 4.3.1. Tạo Hierarchy cho Dim ShipDate

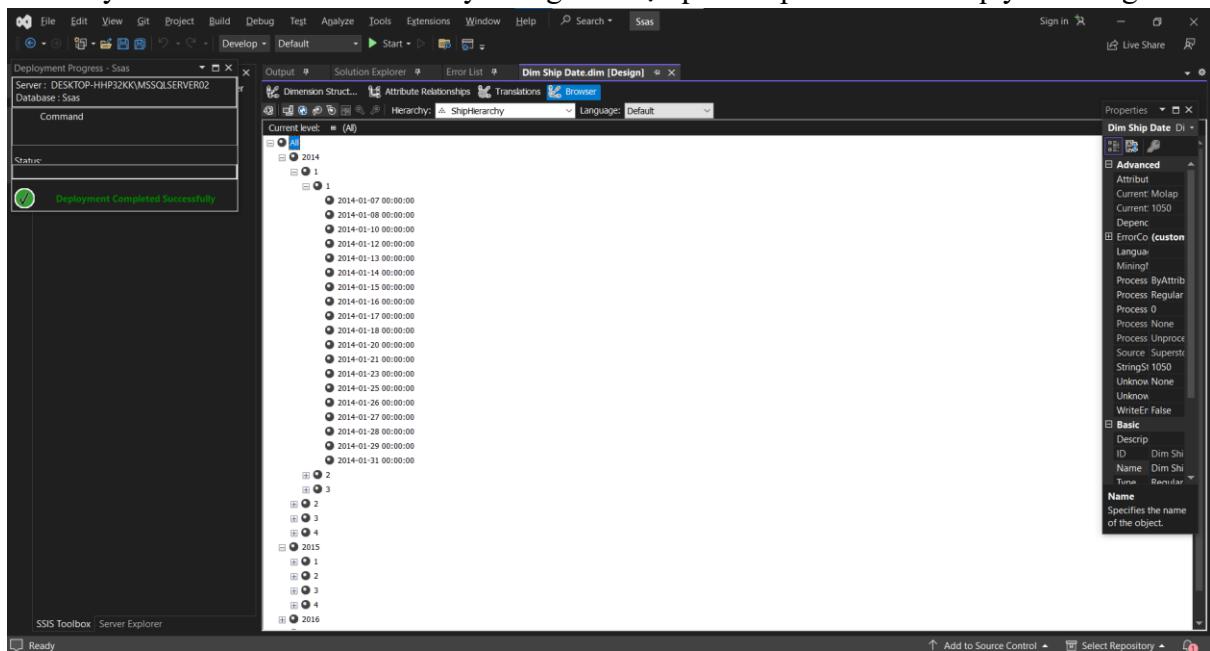
- Thêm thuộc tính vào bảng Dim: Tại table DimShipDate trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.
- Kéo thả 4 thuộc tính *Year*, *Quarter*, *Month Of Year*, *Date* từ Attributes vào panel Hierarchies để tạo ra ShipHierarchy



- Process để load dữ liệu cho bảng Dim ShipDate

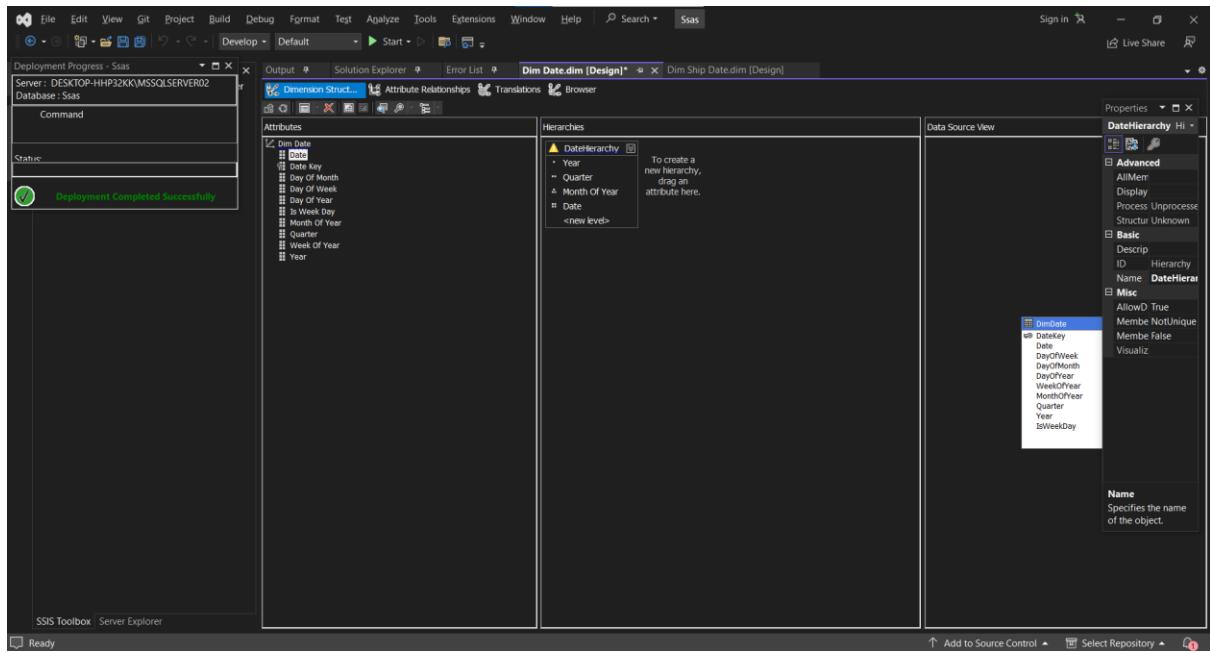


- Chuyển đến tab Browser. Ta thấy thời gian được phân cấp theo năm → quý → tháng.

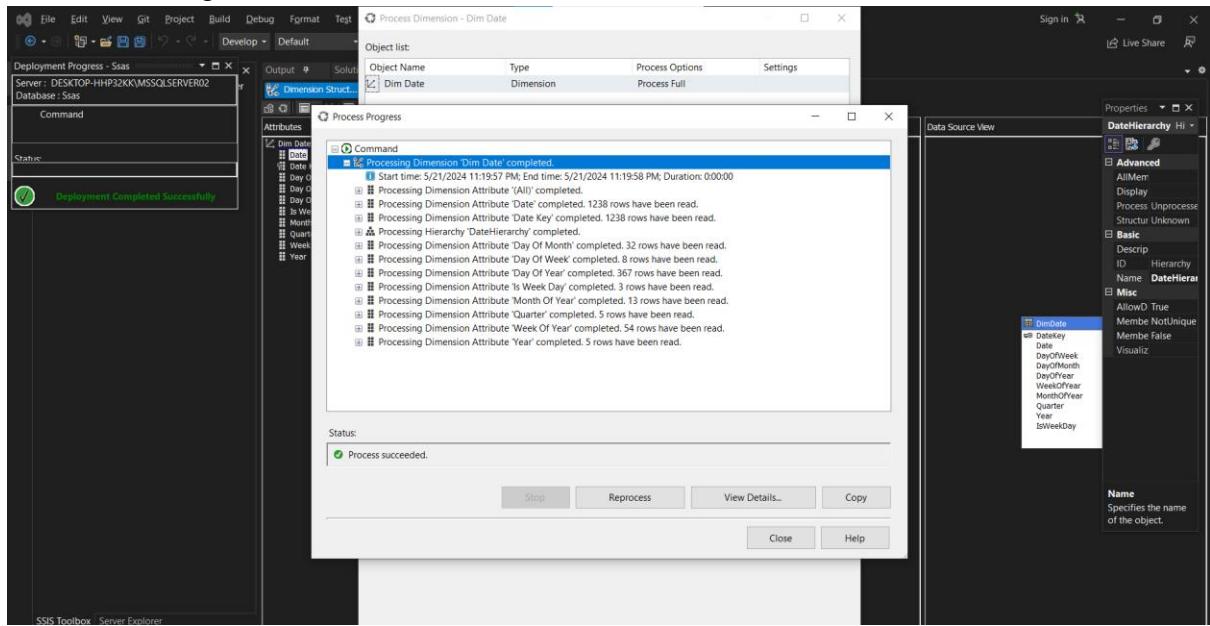


### 4.3.2. Tạo Hierarchy cho Dim Date

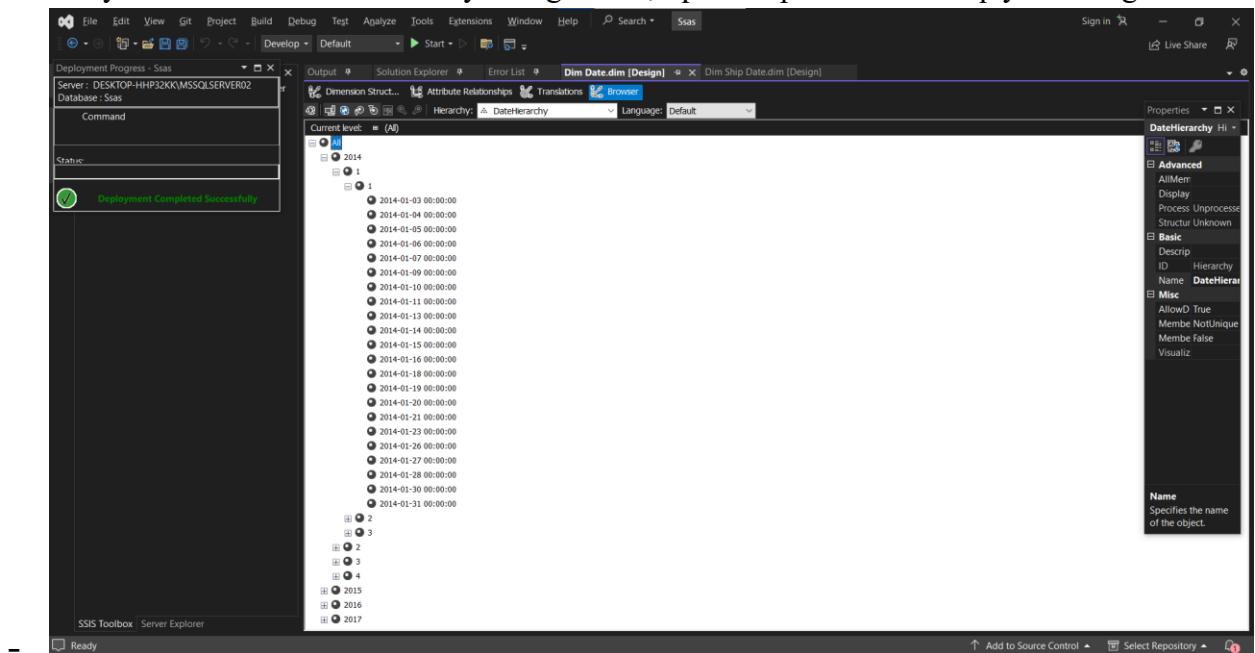
- Tại table Dim Date trong panel Data Source View, kéo 4 thuộc tính như hình vào panel Attributes. Sau đó kéo thả lần lượt Year, Quarter, Month of year, Date vào Hierarchies để tạo DateHierarchy



### - Process bảng Dim Date

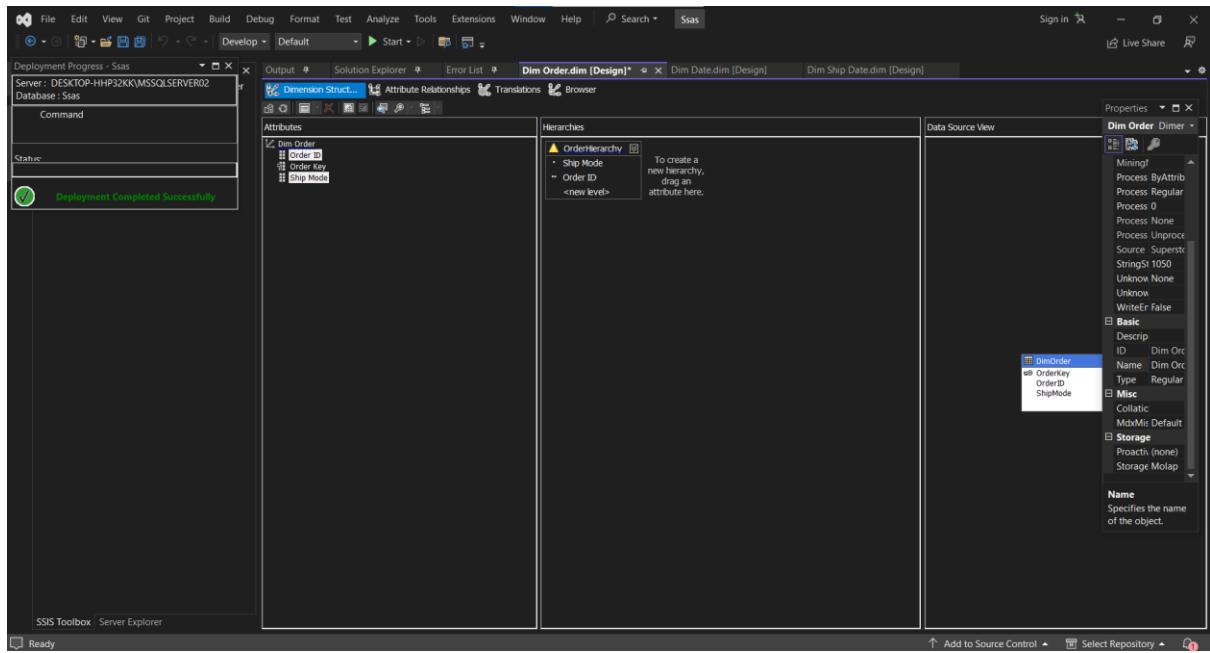


- Chuyển đến tab Browses. Ta thấy thời gian được phân cấp theo năm → quý → tháng.

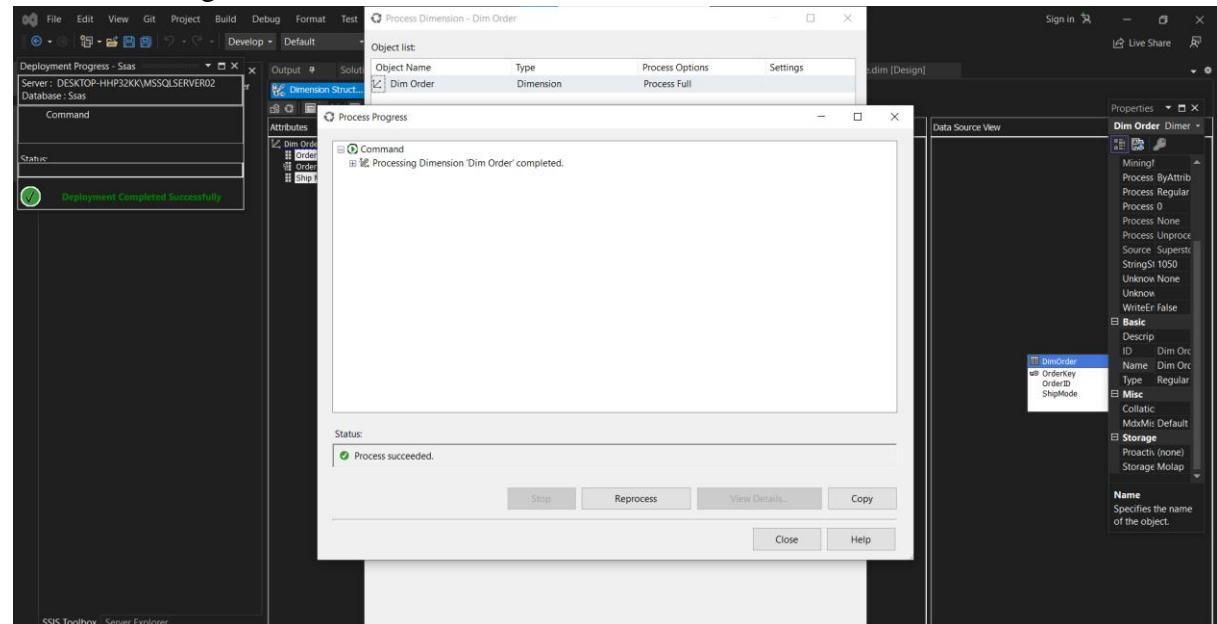


### 4.3.3. Tạo Hierarchy cho Dim Order

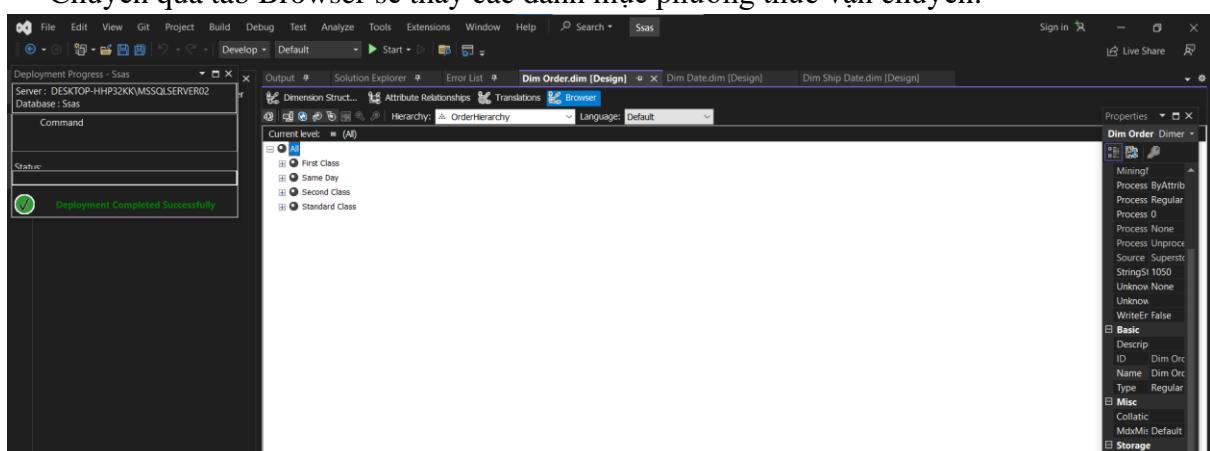
- Tạo OrderHierarchy



### - Process bảng Dim Order

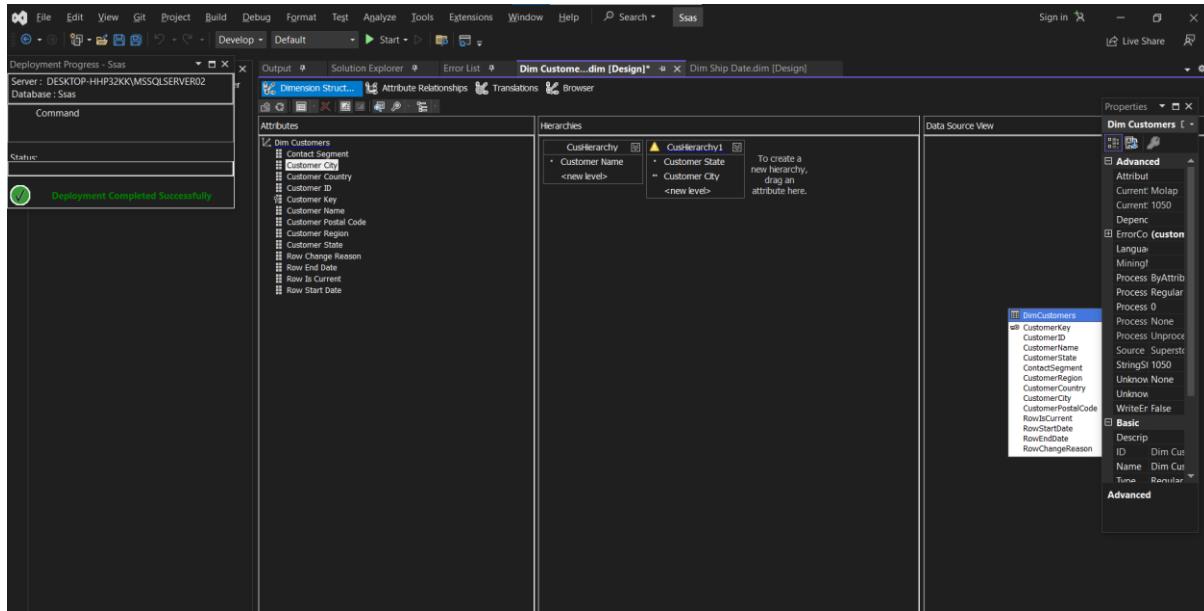


### - Chuyển qua tab Browser sẽ thấy các danh mục phương thức vận chuyển:

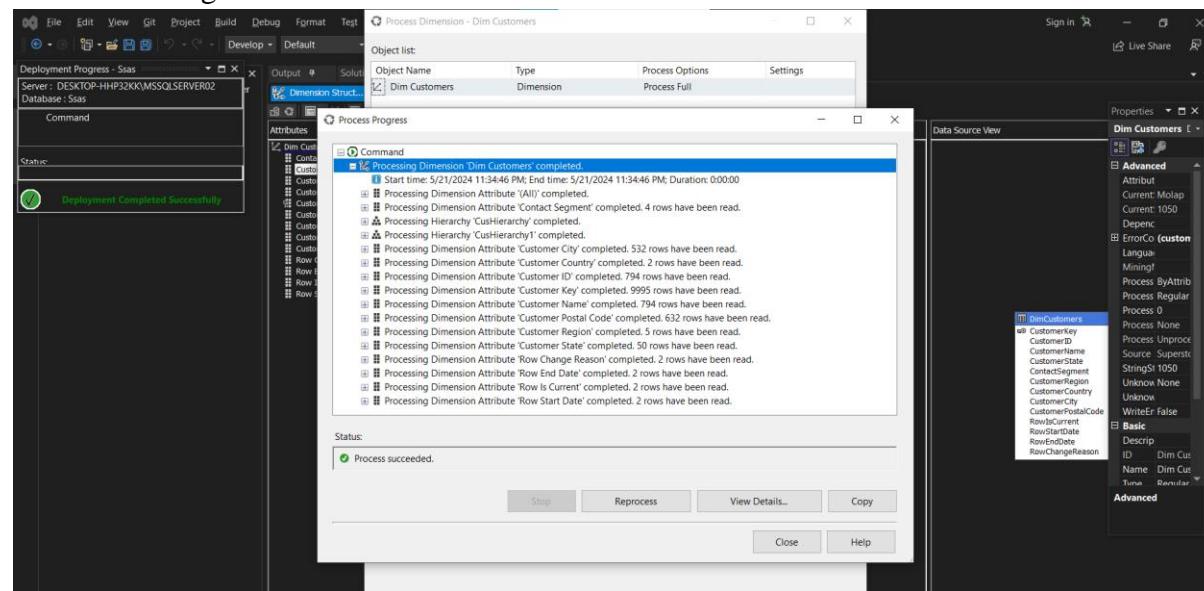


#### 4.3.4. Tạo Hierarchy cho Dim Customer

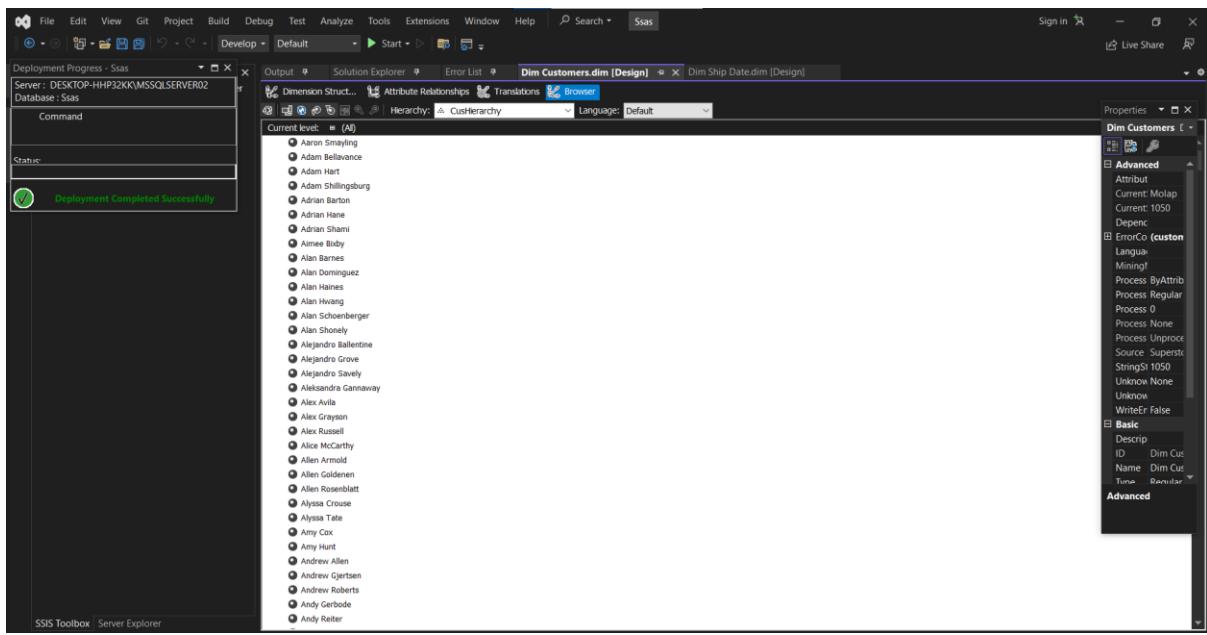
- Tạo CusHierarchy và CusHierarchy1



- Process bảng Dim Customer

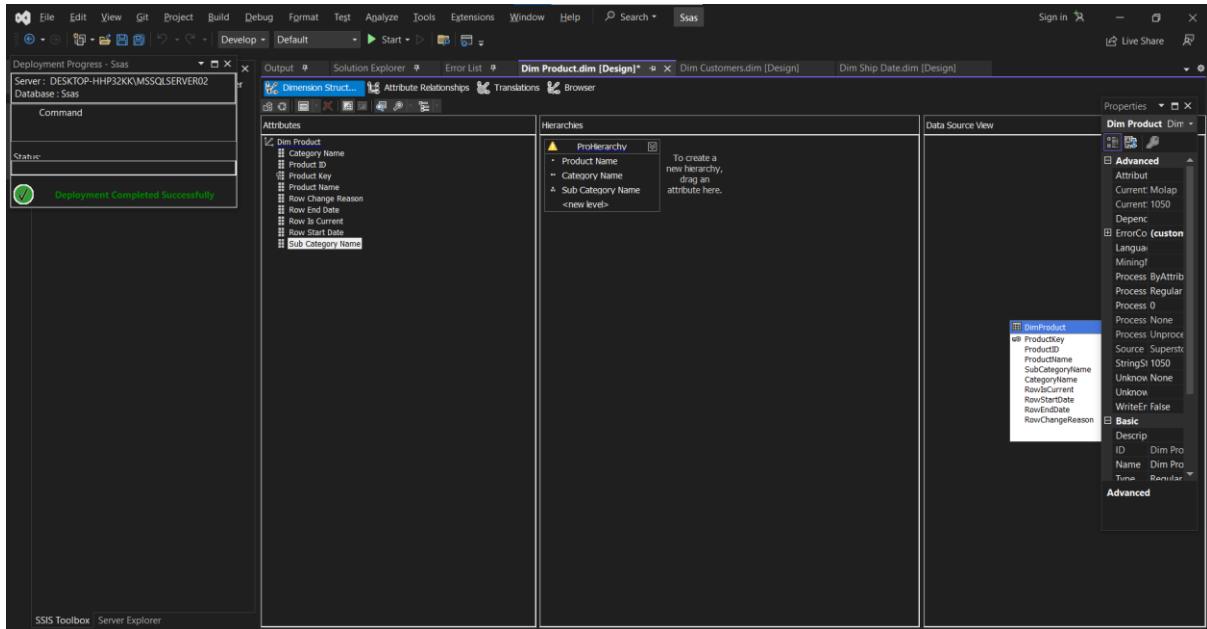


- Chuyển qua tab Browser Danh sách các khách hàng từng mua sản phẩm:

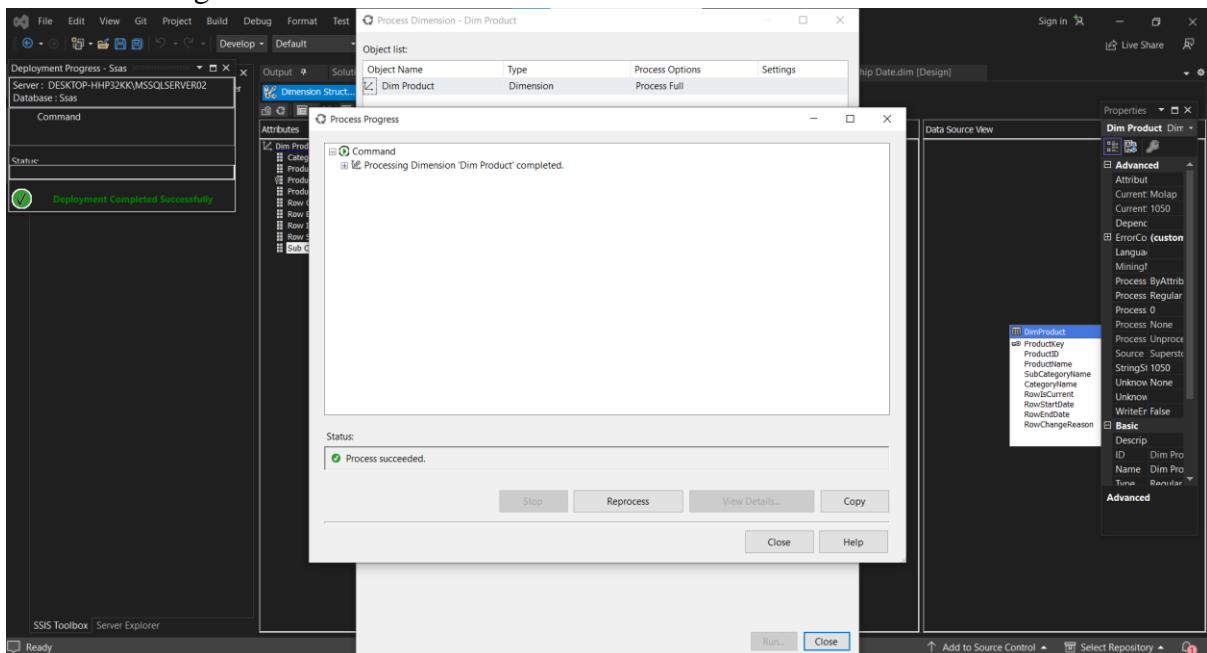


### 4.3.5. Tạo Hierarchy cho Dim Product

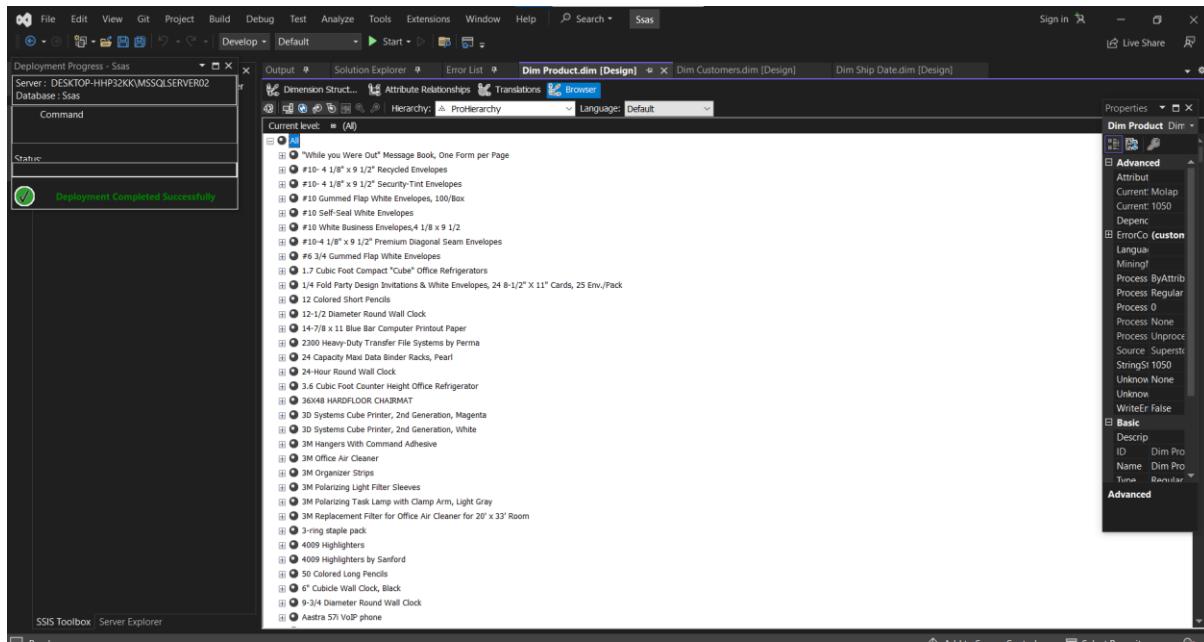
- Tạo ProHierarchy



- Process bảng Dim Product



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

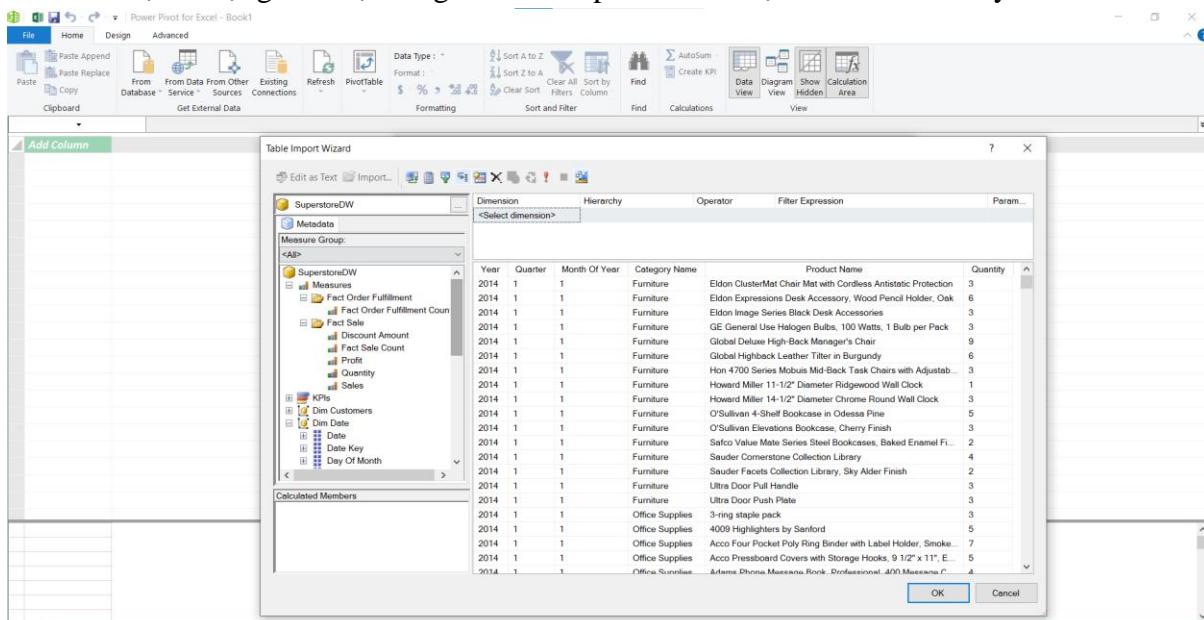


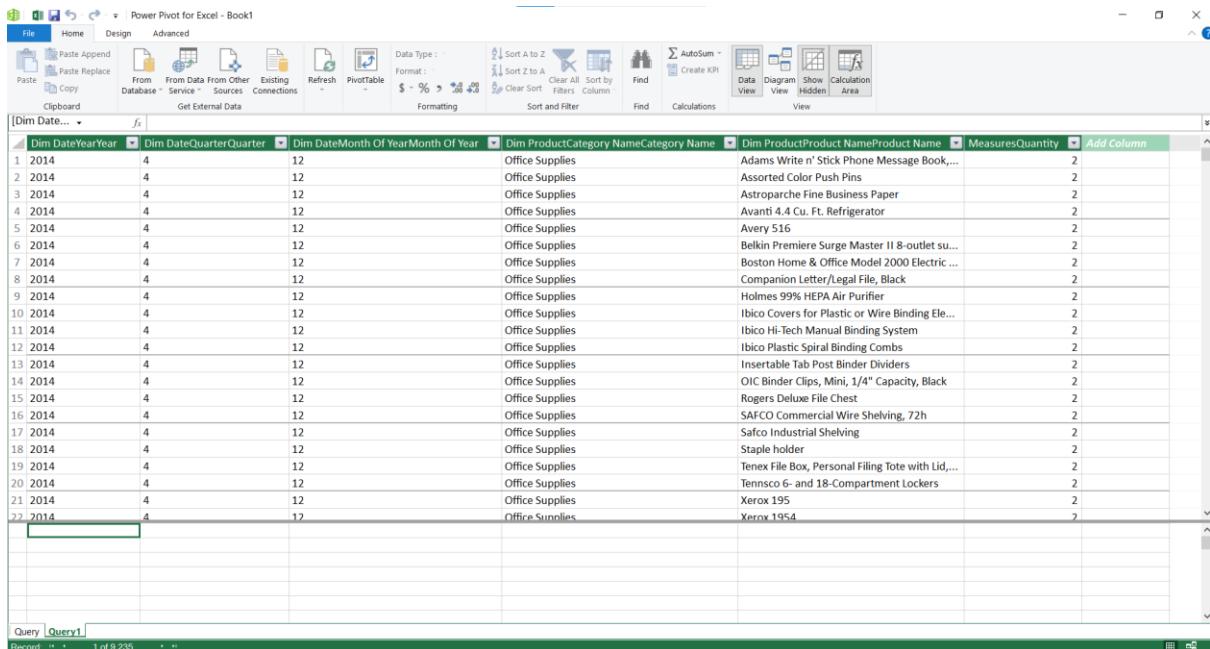
## 4.4. Thực hiện phân tích dữ liệu

### 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





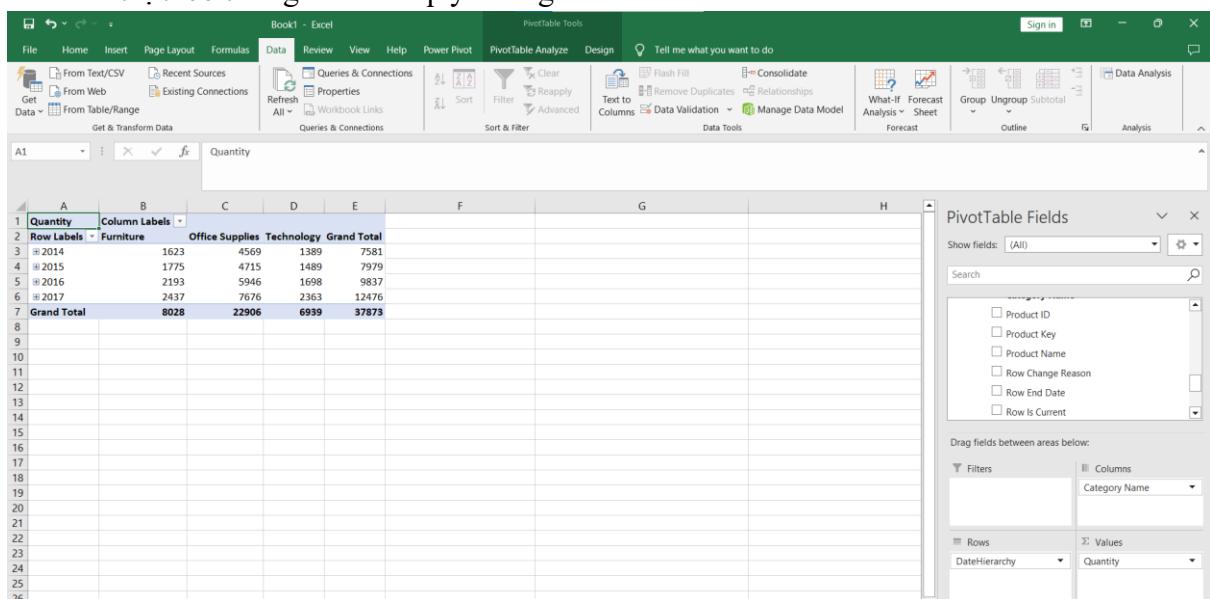
The screenshot shows a Power Pivot data grid with the following columns and data:

Dim DateYear	Dim DateQuarterQuarter	Dim DateMonth Of YearMonth Of Year	Dim ProductCategory NameCategory Name	Dim ProductProduct NameProduct Name	MeasuresQuantity
1 2014	4	12	Office Supplies	Adams Write n' Stick Phone Message Book,...	2
2 2014	4	12	Office Supplies	Assorted Color Push Pins	2
3 2014	4	12	Office Supplies	Astroprobe Fine Business Paper	2
4 2014	4	12	Office Supplies	Avanti 4.4 Cu. Ft. Refrigerator	2
5 2014	4	12	Office Supplies	Avery 516	2
6 2014	4	12	Office Supplies	Belkin Premiere Surge Master II 8-outlet su...	2
7 2014	4	12	Office Supplies	Boston Home & Office Model 2000 Electric ...	2
8 2014	4	12	Office Supplies	Companion Letter/Legal File, Black	2
9 2014	4	12	Office Supplies	Holmes 99% HEPA Air Purifier	2
10 2014	4	12	Office Supplies	Ibico Covers for Plastic or Wire Binding Ele...	2
11 2014	4	12	Office Supplies	Ibico Hi-Tech Manual Binding System	2
12 2014	4	12	Office Supplies	Ibico Plastic Spiral Binding Combs	2
13 2014	4	12	Office Supplies	Insertable Tab Post Binder Dividers	2
14 2014	4	12	Office Supplies	OIC Binder Clips, Mini, 1/4" Capacity, Black	2
15 2014	4	12	Office Supplies	Rogers Deluxe File Chest	2
16 2014	4	12	Office Supplies	SAFCO Commercial Wire Shelving, 72h	2
17 2014	4	12	Office Supplies	Safco Industrial Shelving	2
18 2014	4	12	Office Supplies	Staple holder	2
19 2014	4	12	Office Supplies	Texex File Box, Personal Filing Tote with Lid,...	2
20 2014	4	12	Office Supplies	Tenneco 6- and 18-Compartment Lockers	2
21 2014	4	12	Office Supplies	Xerox 195	2
22 2014	4	12	Office Supplies	Xerox 1954	2

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 a PivotTable in Excel with the following data:

Quantity	Column Labels	Office Supplies	Technology	Grand Total
Row Labels	Furniture			
# 2014		1623	4569	7581
# 2015		1775	4715	7979
# 2016		2193	5946	9837
# 2017		2437	7676	12476
Grand Total		8028	22906	37873

The PivotTable Fields pane on the right shows the following fields:

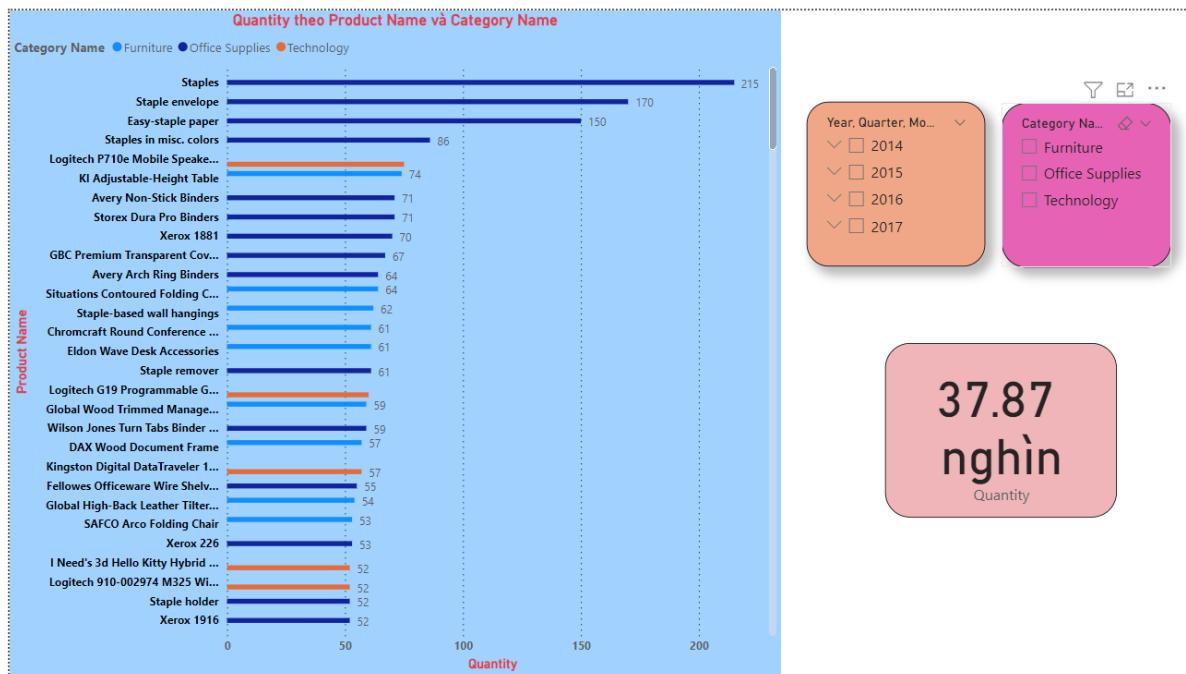
- Show fields: (All)
- Search:
- Fields:
  - Product ID
  - Product Key
  - Product Name
  - Row Change Reason
  - Row End Date
  - Row Is Current

The PivotTable Fields pane also shows the following field mappings:

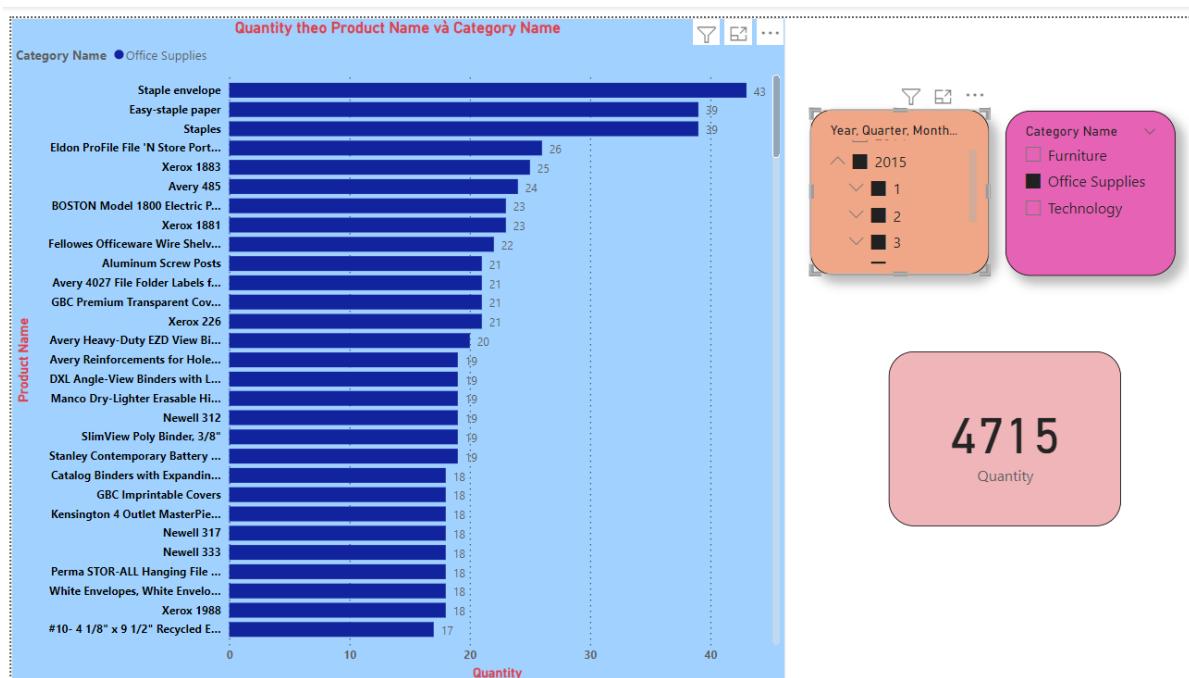
- Filters: DateHierarchy
- Columns: Category Name
- Rows: DateHierarchy
- Values: Quantity

- 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

Year	Quarter	Month	Date	Category Name	Product Name	Sales
2014	1	1	2014-01-03	Office Supplies	Message Book, Wirebound, Four 5 1/2"...	16.479999542...
2014	1	1	2014-01-04	Office Supplies	Avery 508...	11.7840003967...
2014	1	1	2014-01-04	Office Supplies	GBC Standard Plastic Binding Systems...	3.5399996185...
2014	1	1	2014-01-04	Office Supplies	SAFCO Boltless Steel Shelving...	272.735992431...
2014	1	1	2014-01-05	Office Supplies	Avery Hi-liter EverBold Pen Style Fluor...	19.8359992980...
2014	1	1	2014-01-06	Office Supplies	Global Deluxe High-Back Manager's Chair...	2571.82006835...
2014	1	1	2014-01-06	Office Supplies	Allenco Super-Size Bands, Assorted Si...	31.1200005392...
2014	1	1	2014-01-06	Office Supplies	Dixon Prang Watercolor Pencils, 10-Col...	12.7799997329...
2014	1	1	2014-01-06	Office Supplies	Ibico Hi-Tech Manual Binding System...	609.97998468...
2014	1	1	2014-01-06	Office Supplies	Rogers Handheld Barrel Pencil Sharpener...	5.4800001907...
2014	1	1	2014-01-06	Office Supplies	Southworth 25% Cotton Granite Paper...	6.5399996185...
2014	1	1	2014-01-06	Office Supplies	Xerox 225...	19.4400005340...
2014	1	1	2014-01-06	Technology	GE 30524EE4...	391.980010986...
2014	1	1	2014-01-06	Technology	Wireless Extenders zBoost YX545 SOH...	755.960021972...
2014	1	1	2014-01-07	Furniture	Howard Miller 14-1/2" Diameter Chrome...	76.7279965261...
2014	1	1	2014-01-07	Office Supplies	Acco Four Pocket Poly Ring Binder with...	10.4300003051...
2014	1	1	2014-01-09	Office Supplies	Newell 312...	9.34399986267...
2014	1	1	2014-01-09	Technology	Memorex Micro Travel Drive 8 GB...	31.2000007629...
2014	1	1	2014-01-10	Furniture	Howard Miller 11-1/2" Diameter Ridgew...	51.3399986267...
2014	1	1	2014-01-10	Office Supplies	Avery 482...	2.89000010490...
2014	1	1	2014-01-11	Furniture	DAY VALUE U-Channel Document Frame...	0.93999950938...

Dim Date	Dim Date	Dim Date	Dim Date	Dim Date	Dim Date	Dim Date	Dim Date	Dim Date	Dim Date
1 2017	4	11	2017-11-00:00:00:00	Office Supplies	GBC Standard Plastic Binding Systems' Com...	25.1200008392...			
2 2017	4	11	2017-11-01:00:00:00	Office Supplies	Xerox 1893...	40.9900016784...			
3 2017	4	11	2017-11-01:00:00:00	Office Supplies	Xerox 1939...	189.699996948...			
4 2017	4	11	2017-11-02:00:00:00	Office Supplies	Adams Telephone Message Book W/Divider...	11.7600002288...			
5 2017	4	11	2017-11-02:00:00:00	Office Supplies	Binney & Smith Crayola Metallic Colored Pe...	27.7800006866...			
6 2017	4	11	2017-11-02:00:00:00	Office Supplies	Black & Decker Filter for Double Action Dus...	104.036001205...			
7 2017	4	11	2017-11-02:00:00:00	Office Supplies	Colored Envelopes...	8.85599994659...			
8 2017	4	11	2017-11-02:00:00:00	Office Supplies	Fellowes 8 Outlet Superior Workstation Sur...	168.100006103...			
9 2017	4	11	2017-11-02:00:00:00	Office Supplies	GBC Therma-A-Bind 250T Electric Binding S...	590.351989746...			
10 2017	4	11	2017-11-02:00:00:00	Office Supplies	Letter or Legal Size Expandable Poly String ...	5.32000017166...			
11 2017	4	11	2017-11-02:00:00:00	Office Supplies	Newell 318...	5.55999994277...			
12 2017	4	11	2017-11-02:00:00:00	Office Supplies	Nontoxic Chalk...	8.80000019073...			
13 2017	4	11	2017-11-02:00:00:00	Office Supplies	Premier Automatic Letter Opener...	384.592010498...			
14 2017	4	11	2017-11-02:00:00:00	Office Supplies	Space Solutions Commercial Steel Shelving...	452.549987792...			
15 2017	4	11	2017-11-02:00:00:00	Office Supplies	Stanley Bostitch Contemporary Electric Pen...	54.3359985351...			
16 2017	4	11	2017-11-02:00:00:00	Office Supplies	Staples...	18.239999711...			
17 2017	4	11	2017-11-02:00:00:00	Office Supplies	Super Bands, 12/Pack...	5.57999992370...			
18 2017	4	11	2017-11-02:00:00:00	Office Supplies	Super Decoflex Portable Personal File...	59.9199981689...			
19 2017	4	11	2017-11-02:00:00:00	Office Supplies	White Business Envelopes with Contempor...	76.5800018310...			
20 2017	4	11	2017-11-02:00:00:00	Office Supplies	Wirebound Message Books, Four 2 3/4 x 5 ...	23.8500003814...			
21 2017	4	11	2017-11-02:00:00:00	Office Supplies	Xerox 220...	19.4400005340...			
22 2017	4	11	2017-11-03:00:00:00	Office Supplies	Avery 476...	12.3900003433...			

- 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

	Sales	Column Labels					
1	Sales	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	

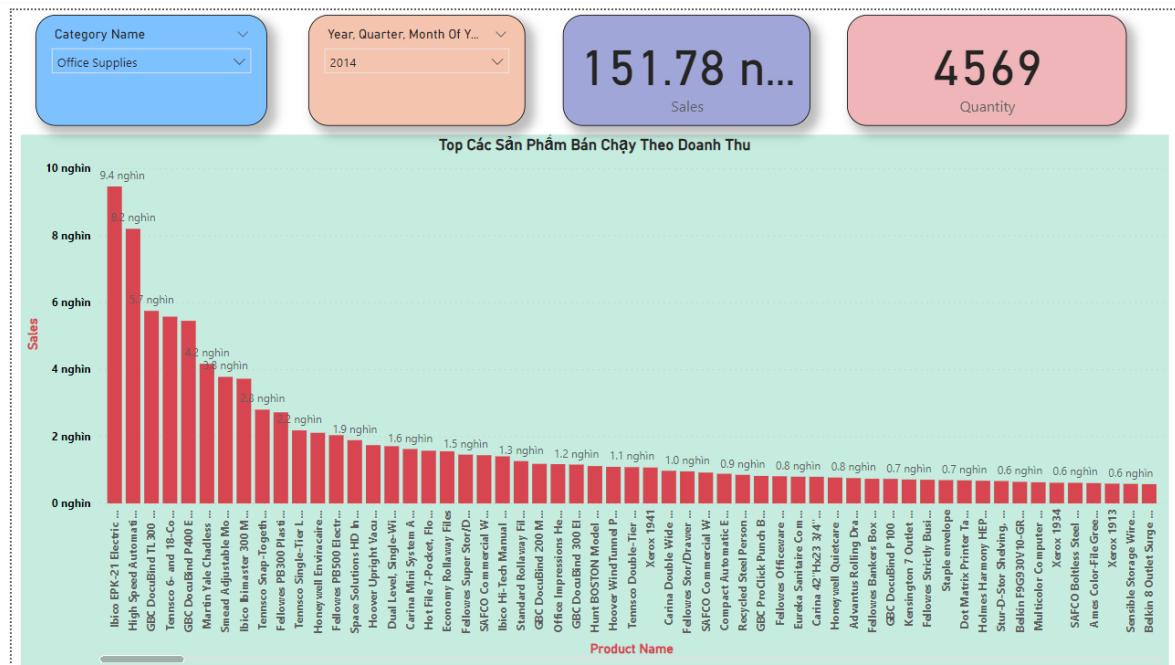
Giải thích: Hình trên thể hiện doanh thu các mặt hàng theo năm có thể thấy được Năm 2017 là năm có tổng doanh thu cao nhất với hơn 700 nghìn đô la và mặt hàng có doanh số cao nhất chính là Furniture(đồ nội thất) với hơn 740 nghìn đô la.

#### 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

Edit as Text Import...

Dimension Hierarchy Operator Filter Expression Param...

<Select dimension>

Measure Group: **<All>**

Quarter Week Of Year Year DateHierarchy

Dim Order Dim Product Category Name Product ID Product Key Product Name Row Change Reason Row End Date Row Is Current Row Start Date Sub Category Name ProductHierarchy

Dim Ship Date

Calculated Members

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

Power Pivot for Excel - Book1

File Home Design Advanced

Paste Append From Database Refresh PivotTable

Clipboard Get External Data

Formatting Sort and Filter Find Calculations

AutoSum Data View Diagram View Show Hidden Calculation Area

Dim Product Category Name Category Name Dim Product Sub-Category Name Sub-Category Name Measures Sales

1 2014 1 1 Office Supplies Appliances 100.0679965019... 312.58000183...

2 2014 1 1 Office Supplies Art 52.548899769783 176.9899452...

3 2014 1 1 Office Supplies Binders 190.7386970520... 814.50798058...

4 2014 1 1 Office Supplies Envelopes 78.47200012207... 194.3599965...

5 2014 1 1 Office Supplies Fasteners 0.31199992895... 37.060000896...

6 2014 1 1 Office Supplies Labels 15.77929949760... 44.746001005...

7 2014 1 1 Office Supplies Paper 101.5603990554... 242.02000427...

8 2014 1 1 Office Supplies Storage 249.4720983505... 3028.8160018...

9 2014 1 2 Office Supplies Appliances 9.01879882125 89.919999122...

10 2014 1 2 Office Supplies Art 12.21359968185... 73.663999574...

11 2014 1 2 Office Supplies Binders 61.55069839954... 339.26200163...

12 2014 1 2 Office Supplies Fasteners 5.804500102996... 12.350000381...

13 2014 1 2 Office Supplies Labels 16.41329956054... 36.875999450...

14 2014 1 2 Office Supplies Paper 56.50009965896... 138.48800182...

15 2014 1 2 Office Supplies Storage 9.817600488662... 142.54399776...

16 2014 1 2 Office Supplies Supplies 4.772399902343... 238.61999511...

17 2014 1 3 Office Supplies Appliances -361.546808242... 502.96199417...

18 2014 1 3 Office Supplies Art 108.3423999845... 413.26800179...

19 2014 1 3 Office Supplies Binders 381.3023970127... 1525.6750056...

20 2014 1 3 Office Supplies Envelopes 121.5210962295... 353.20000076...

21 2014 1 3 Office Supplies Fasteners 4.085399985313... 48.240000724...

Record 1 of 782

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

Book1 - Excel

File Home Insert Page Layout Formulas Data Review View Help Power Pivot PivotTable Analyze Design

From Text/CSV From Web Existing Connections Get From Table/Range Get & Transform Data

Queries & Connections Refresh Properties All Edit Links Sort & Filter Advanced Text to Columns Remove Duplicates Relationships Data Tools

What-If Analysis Forecast Sheet Group Ungroup Subtotal Outline

K24 286817.020819017

Column Labels: Sales Profit Total Sales Total Profit

Row Labels: 2014 2015 2016 2017 2014 2015 2016 2017

Furniture

Bookcases 20036.67757 38543.57256 26275.46651 30024.27937 -346.1707034 -2755.23026 212.471014 -583.6260874 114879.996 -3472.556036

Chairs 72241.57597 71734.52907 83198.64464 95554.35265 6955.047001 6228.416157 5763.153843 7643.54924 328449.1023 26590.16624

Furnishings 13826.23397 21089.71206 28784.12382 28915.09389 1972.891899 3051.818881 3935.269995 4099.162791 91705.16374 13059.14357

Tables 46088.36541 39150.42334 60833.20027 60893.54258 -3124.04273 -3509.801855 -2950.941785 -7720.69468 206965.5316 -17305.48106

Office Supplies

Appliances 15313.62506 23241.28909 26050.31515 42926.93201 2459.499009 2511.89575 5301.341505 7865.268199 107532.1613 18138.00536

Art 6057.98205 6236.833989 5960.90802 8863.06801 1406.803094 1485.058202 1413.962595 2221.963091 27118.79207 6527.786982

Binders 43488.2647 37453.09808 49683.32523 72788.04432 4739.716022 7596.617575 10215.63411 7669.741668 203412.7324 30221.76338

Envelopes 3855.750205 4512.187984 4729.889987 3378.573991 1495.478793 1960.19059 2066.748289 1441.758993 16476.40199 6964.176666

Fasteners 661.328005 545.2239981 960.1339962 857.5939962 178.5853007 171.9376004 294.0733986 304.9488965 3024.279991 949.5181961

Labels 2841.393973 2956.461994 2827.240016 3861.21598 1286.004204 1323.028609 1192.611899 1744.609317 12486.31196 5546.254028

Paper 14834.95589 15287.63795 20661.89395 27694.71788 6371.021357 6570.170006 9071.534431 12040.84334 78479.20567 34053.56913

Storage 50329.04208 45048.24814 58788.70024 60949.61831 4166.27889 3505.460093 6204.286775 7402.800276 23843.6088 17278.82639

Supplies 14394.06969 1952.481981 14277.57584 16049.40957 490.0556121 -24.87840691 -698.9639013 -955.312794 46673.5371 -1189.09949

Technology

Accessories 25014.27192 40523.95995 41895.85402 59946.23199 6402.71501 10197.27522 9664.288473 15672.357 167380.3179 41936.6357

Copiers 10849.7818 26179.45013 49599.40869 62899.38702 2912.943268 9930.298569 17742.79219 25031.95276 149528.0276 55617.8246

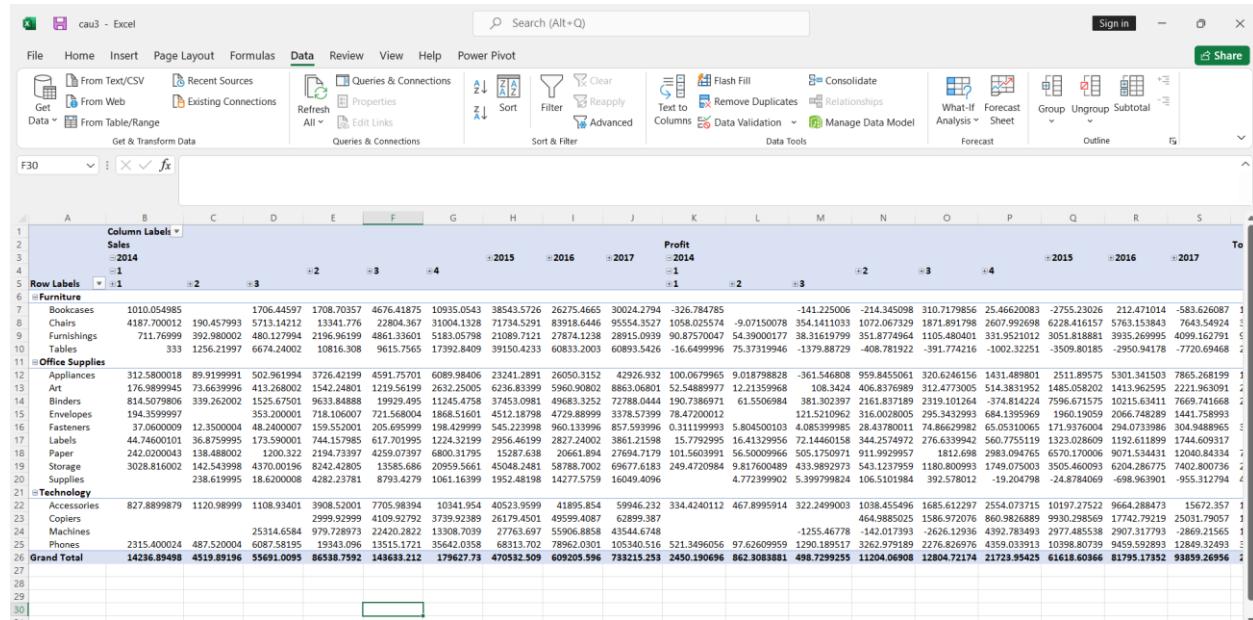
Machines 62023.37354 27763.69696 55906.88575 43544.67476 369.1689606 2977.485538 2907.317793 -2869.215652 189238.631 3384.576639

Phones 77390.80586 68313.70201 78962.03007 105340.5164 11808.0053 10398.80739 9459.592893 12849.32493 330007.0543 44515.73052

Grand Total 484247.4975 470532.5093 60925.5962 73215.2528 49543.97408 61618.60366 81795.17352 93859.26956 229720.856 286817.0208

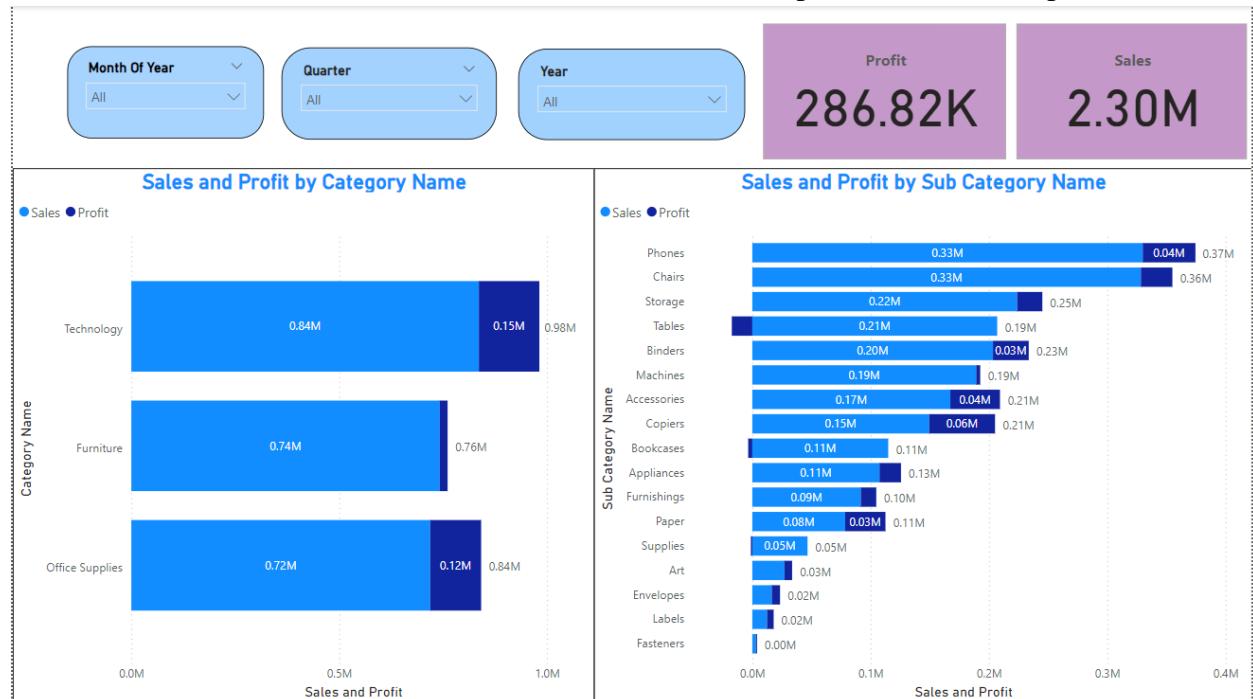
- Nhìn vào hình, ta có thể thấy được lợi nhuận và doanh thu tăng dần theo thời gian, cao nhất đều đạt ở năm 2017.
- Về doanh thu, danh mục có doanh thu cao nhất là Office Supplies vì danh mục này có số lượng mặt hàng nhiều nhất, danh mục phụ có doanh thu cao nhất là Phones(điện thoại) với doanh thu hơn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



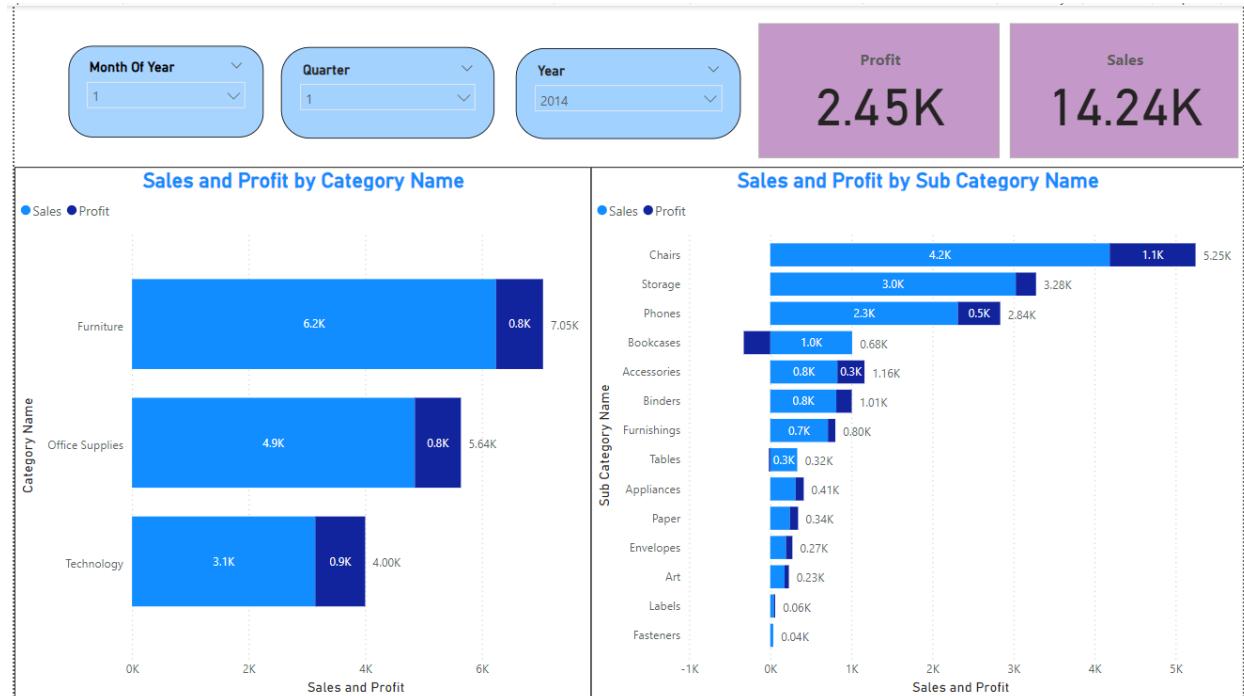
#### 4.4.3.3. Sử dụng Power BI

- 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

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Fact Order Fulfillment Count	Column Labels	C	D	E	Grand Total	G	H	I	J	K	L	M	N	O	P	Q
2	Row Labels	First Class	Same Day	Second Class	Standard Class		G	H	I	J	K	L	M	N	O	P	Q
3	2014																
4	Consumer	82	30	108	304	524											
5	Corporate	43	10	58	176	287											
6	Home Office	20	8	24	106	158											
7	2015																
8	Consumer	68	37	101	329	535											
9	Corporate	50	8	61	193	312											
10	Home Office	25	8	44	114	191											
11	2016																
12	Consumer	105	42	121	383	654											
13	Corporate	71	17	78	256	422											
14	Home Office	39	15	45	143	242											
15	2017																
16	Consumer	141	48	168	519	876											
17	Corporate	86	23	104	280	493											
18	Home Office	57	18	52	191	318											
19	Grand Total	787	264	964	2994	5009											
20																	
21																	
22																	
23																	
24																	
25																	
26																	
27																	
28																	
29																	
30																	
31																	

- 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 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

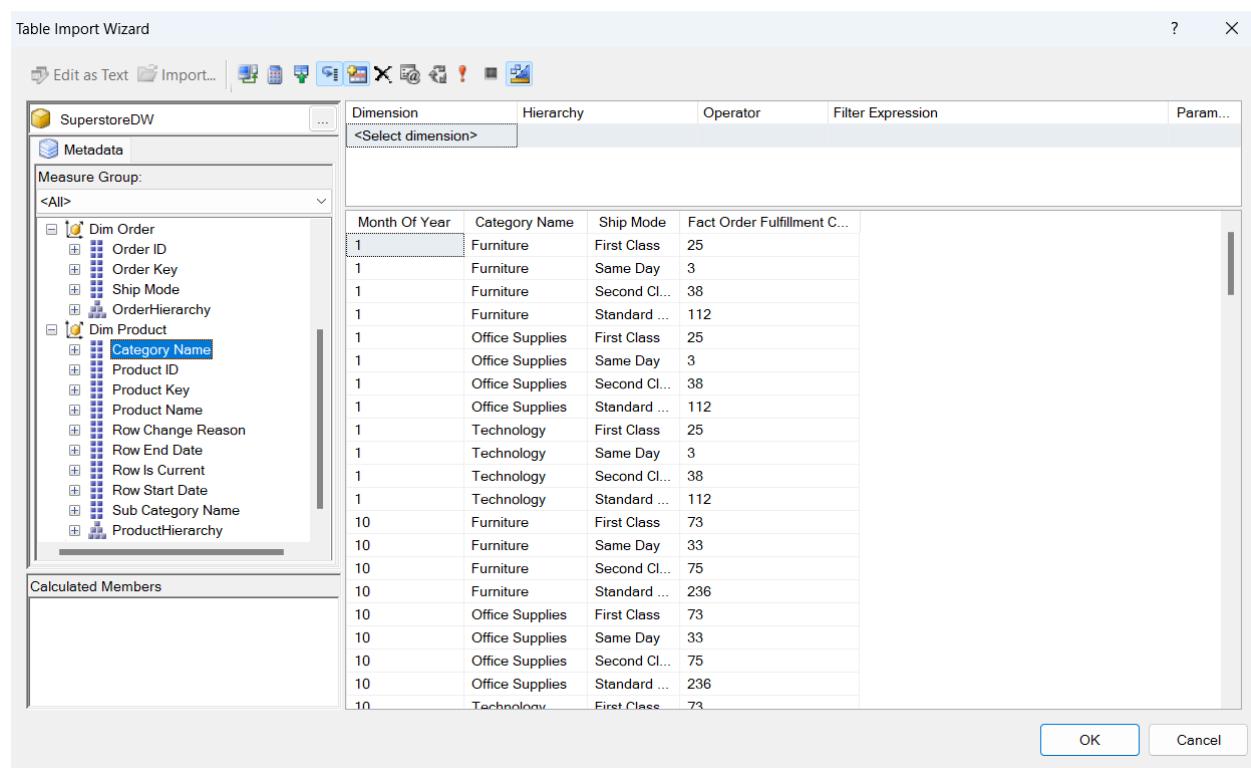
Standard Class trong mọi phân khúc khách hàng là một xu hướng rõ ràng và liên tục trong suốt bốn năm qua.

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.



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

- 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

Book2 - Excel

File Home Insert Page Layout Formulas Data Review View Help Power Pivot PivotTable Analyze Design

Subtotals Grand Report Blank Layout Row Totals Layout Rows

Row Headers Banded Rows

Column Headers Banded Columns

Layout PivotTable Style Options PivotTable Styles

N18 5009

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Fact Order Fulfillment Count	Column Labels	1	10	11	12	2	3	4	5	6	7	8	9	Grand Total						
2	Row Labels		10	11	12	2	3	4	5	6	7	8	9	Grand Total							
3	■ Furniture																				
4	First Class		25	73	109	114	25	50	62	53	57	52	56	111	787						
5	Same Day		3	33	40	32	3	25	8	24	18	25	16	37	264						
6	Second Class		38	75	143	140	34	78	55	59	74	63	71	134	964						
7	Standard Class		112	236	461	416	100	201	218	233	215	198	198	406	2994						
8	■ Office Supplies																				
9	First Class		25	73	109	114	25	50	62	53	57	52	56	111	787						
10	Same Day		3	33	40	32	3	25	8	24	18	25	16	37	264						
11	Second Class		38	75	143	140	34	78	55	59	74	63	71	134	964						
12	Standard Class		112	236	461	416	100	201	218	233	215	198	198	406	2994						
13	■ Technology																				
14	First Class		25	73	109	114	25	50	62	53	57	52	56	111	787						
15	Same Day		3	33	40	32	3	25	8	24	18	25	16	37	264						
16	Second Class		38	75	143	140	34	78	55	59	74	63	71	134	964						
17	Standard Class		112	236	461	416	100	201	218	233	215	198	198	406	2994						
18	Grand Total		178	417	753	702	162	354	343	369	364	338	341	688	5009						
19																					
20																					
21																					
22																					
23																					
24																					
25																					

Sheet1

PivotTable Fields

Show fields: (All)

Search

Dim Product

Dim Ship Date

Ship Date Hierarchy

More Fields

Drag fields between areas below:

Filters

Columns

Month Of Year

Rows

Category Name

Ship Mode

Values

Fact Order Fulfillment

Defer Layout Update

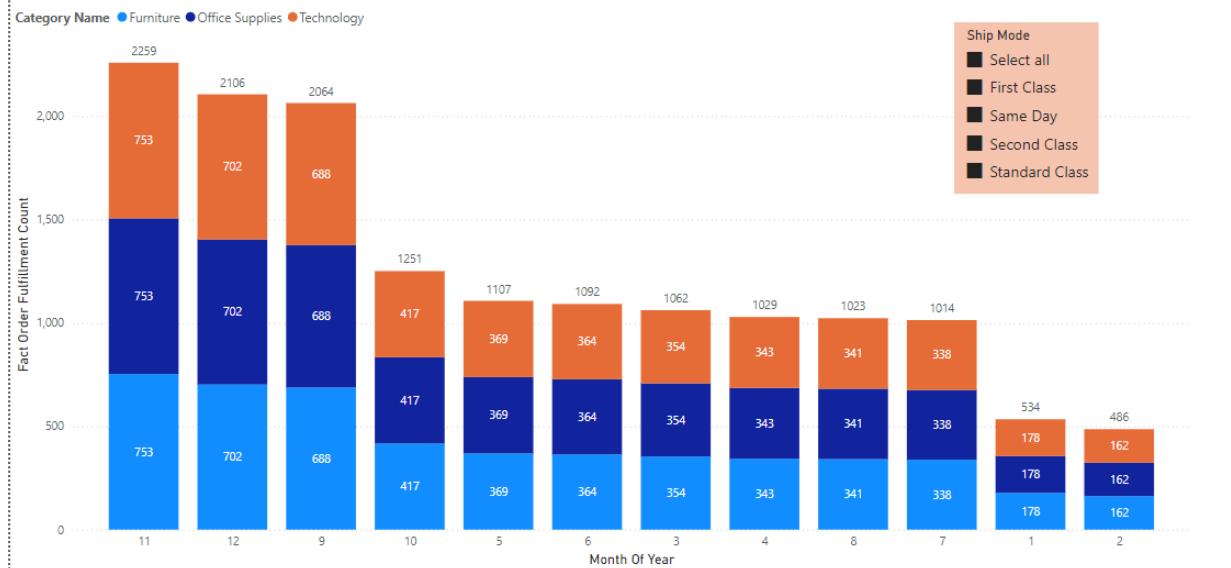
Update

- 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.

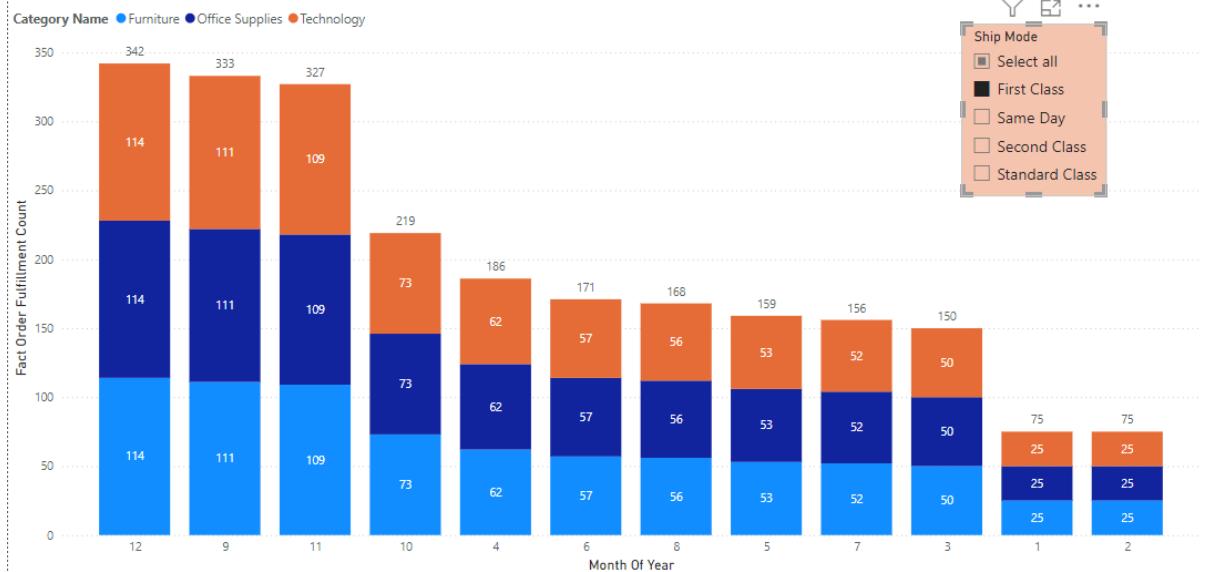
### Fact Order Fulfillment Count by Month Of Year and Category Name



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.