

BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT  
THÀNH PHỐ HỒ CHÍ MINH



**HCMUTE**

**CHỦ ĐỀ: XÂY DỰNG KHO DỮ LIỆU CHO CÁC  
GIAO DỊCH BÁN LẺ  
MÔN: KHO DỮ LIỆU**

**Giảng viên: Ths. Nguyễn Văn Thành  
Sinh viên thực hiện**

STT	Họ và tên	MSSV
1	Phạm Minh Long	20133062
2	Trương Hoàng Anh Khôi	20133058
3	Lê Hoàng Hải Đăng	20110265

Thủ Đức, tháng 5 năm 2023

## **LỜI CẢM ƠN**

Lời đầu tiên, nhóm em muôn gửi lời cảm ơn sâu sắc đến Thầy - ThS. Nguyễn Văn Thành - Giảng viên phụ trách môn Kho Dữ Liệu tại trường Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh.

Trong quá trình thực hiện báo cáo tiểu luận, nhóm em đã nhận được rất nhiều sự hỗ trợ từ Thầy. Thầy đã cung cấp cho chúng em những kiến thức cần thiết, hướng dẫn chi tiết và giúp đỡ nhóm em trong việc giải quyết các vấn đề khó khăn. Nhờ đó, nhóm em đã hoàn thành báo cáo tiểu luận một cách tốt nhất.

Nhóm chúng em rất biết ơn và trân trọng sự hỗ trợ của Thầy trong quá trình học tập và nghiên cứu. Chúng em hy vọng sẽ tiếp tục được học tập và làm việc với Thầy trong những dự án và nghiên cứu tiếp theo. Một lần nữa, chân thành cảm ơn Thầy.

## MỤC LỤC

<b>Chương 1: Lý do chọn Dataset và giới thiệu tổng quan Dataset .....</b>	<b>1</b>
<b>1.1. Lý do hình thành dự án .....</b>	<b>1</b>
<b>1.1.1. Vấn đề nhận thấy .....</b>	<b>1</b>
<b>1.1.2. Giải pháp .....</b>	<b>2</b>
<b>1.1.3. Mục tiêu và ý nghĩa của đồ án .....</b>	<b>2</b>
<b>1.2. Giới thiệu tổng quan dataset.....</b>	<b>2</b>
<b>1.2.1. Nguồn dữ liệu sử dụng.....</b>	<b>2</b>
<b>1.2.2. Mô tả chi tiết cơ sở dữ liệu .....</b>	<b>3</b>
<b>1.3. Giới thiệu các công cụ sử dụng trong đồ án .....</b>	<b>8</b>
<b>1.3.1. Tổng quan về Visual Studio 2019 .....</b>	<b>8</b>
<b>1.3.2. Giới thiệu về SQL Server Data Tools .....</b>	<b>8</b>
<b>1.3.3. Giới thiệu ngôn ngữ truy vấn SQL .....</b>	<b>9</b>
<b>1.3.4. Tổng quan về SQL Server 2019 .....</b>	<b>10</b>
<b>1.3.5. Giới thiệu về SQL Server Analysis Service .....</b>	<b>11</b>
<b>1.3.6. Giới thiệu Microsoft Power BI .....</b>	<b>12</b>
<b>Chương 2: Thiết kế xây dựng cơ sở dữ liệu .....</b>	<b>13</b>
<b>2.1. Yêu cầu khai thác dữ liệu .....</b>	<b>13</b>
<b>2.2. Chi tiết thiết kế .....</b>	<b>13</b>
<b>2.2.1. Lượt đồ.....</b>	<b>13</b>
<b>2.2.2. Chi tiết bảng Fact.....</b>	<b>15</b>
<b>2.2.3. Chi tiết bảng Dim .....</b>	<b>16</b>
<b>Chương 3: Tích hợp dữ liệu vào kho (SISS) .....</b>	<b>20</b>
<b>3.1. Luồng dữ liệu tổng quát.....</b>	<b>20</b>
<b>3.2. Quá trình đổ dữ liệu từ excel vào cơ sở dữ liệu Stage .....</b>	<b>20</b>
<b>3.2.1. Stage Department .....</b>	<b>20</b>
<b>3.2.2. Stage Shipping Mode.....</b>	<b>23</b>
<b>3.2.3. Stage Segment .....</b>	<b>27</b>

<b>3.2.4.</b>	<b>Stage Customer .....</b>	29
<b>3.2.5.</b>	<b>Stage Time .....</b>	36
<b>3.2.6.</b>	<b>Stage Market .....</b>	39
<b>3.2.7.</b>	<b>Stage Region .....</b>	42
<b>3.2.8.</b>	<b>Stage Country.....</b>	47
<b>3.2.9.</b>	<b>Stage City.....</b>	51
<b>3.2.10.</b>	<b>Stage Category .....</b>	57
<b>3.2.11.</b>	<b>Stage Product .....</b>	60
<b>3.2.12.</b>	<b>Stage Sales .....</b>	63
<b>3.2.13.</b>	<b>Stage Delivery.....</b>	75
<b>3.3.</b>	<b>Quá trình đổ dữ liệu từ cơ sở dữ liệu Stage vào kho dữ liệu .....</b>	83
<b>3.3.1.</b>	<b>Dim Department .....</b>	83
<b>3.3.2.</b>	<b>Dim ShippingMode .....</b>	85
<b>3.3.3.</b>	<b>Dim Segment .....</b>	87
<b>3.3.4.</b>	<b>Dim Customer .....</b>	89
<b>3.3.5.</b>	<b>Dim Time .....</b>	90
<b>3.3.6.</b>	<b>Dim Market .....</b>	92
<b>3.3.7.</b>	<b>Dim Region .....</b>	94
<b>3.3.8.</b>	<b>Dim Country.....</b>	96
<b>3.3.9.</b>	<b>Dim City.....</b>	98
<b>3.3.10.</b>	<b>Dim Category .....</b>	99
<b>3.3.11.</b>	<b>Dim Product .....</b>	101
<b>3.3.12.</b>	<b>Fact Sales .....</b>	102
<b>3.3.13.</b>	<b>Fact Delivery .....</b>	104
<b>3.4.</b>	<b>Kết quả đổ dữ liệu vào Data Warehouse .....</b>	106
<b>3.4.1.</b>	<b>Luồng dữ liệu .....</b>	106
<b>3.4.2.</b>	<b>Kết quả dữ liệu .....</b>	118
<b>Chương 4:</b>	<b>Phân tích dữ liệu SSAS .....</b>	125
<b>4.1.</b>	<b>Một số câu hỏi đặt ra để phân tích .....</b>	125

<b>4.2. Quá trình xây dựng mô hình.....</b>	<b>125</b>
<b>4.3. Quá trình xây dựng khối - Cube.....</b>	<b>128</b>
<b>4.4. Thực hiện truy vấn(SSAS và Power BI) .....</b>	<b>135</b>
<b>4.4.1. Báo cáo tổng doanh thu theo quý, năm. ....</b>	<b>135</b>
<b>4.4.2. Báo cáo tổng lợi nhuận theo quý, năm.....</b>	<b>136</b>
<b>4.4.3. Thống kê tình trạng giao hàng trễ của các khu vực .....</b>	<b>138</b>
<b>4.4.4. Thống kê các đơn hàng bị hủy theo quốc gia .....</b>	<b>139</b>
<b>4.4.5. So sánh doanh thu giữa các khu vực theo năm .....</b>	<b>140</b>
<b>4.4.6. Thị trường nào có doanh thu nhiều nhất và thị trường nào có doanh thu ít nhất? .....</b>	<b>142</b>
<b>4.4.7. Sản phẩm nào bán chạy nhất trong một thời gian nhất định? .....</b>	<b>142</b>
<b>4.4.8. Loại mặt hàng nào được ưa chuộng nhất theo từng thị trường .....</b>	<b>143</b>
<b>4.4.9. Tháng nào trong năm thường có mức độ bán hàng cao nhất? .....</b>	<b>145</b>
<b>4.4.10. Danh mục sản phẩm nào được ưa chuộng nhất?.....</b>	<b>146</b>
<b>4.4.11. So sánh các phương thức vận chuyển .....</b>	<b>147</b>
<b>4.5. Lập bảng báo cáo DashBoard.....</b>	<b>148</b>
<b>Chương 5: Kết luận .....</b>	<b>150</b>
<b>5.1. Kết quả đạt được .....</b>	<b>150</b>
<b>5.2. Những hạn chế chưa đạt được.....</b>	<b>150</b>
<b>5.3. Bảng phân công nhiệm vụ.....</b>	<b>151</b>
<b>Tài liệu tham khảo .....</b>	<b>153</b>

# **Chương 1: Lý do chọn Dataset và giới thiệu tổng quan Dataset**

Giới thiệu tổng quan về dataset, lý do hình thành dự án, nguồn dữ liệu thực hiện. Tính cần thiết của xây dựng cơ sở dữ liệu. Khảo sát, nghiên cứu và phân tích các báo cáo nghiệp vụ cần phục vụ cho kho dữ liệu.

## **1.1. Lý do hình thành dự án**

### **1.1.1. Vấn đề nhận thấy**

Từ trước cho đến nay, ngành bán lẻ luôn là một trong những ngành hoạt động sôi nổi và chiếm thị phần lớn trong nền kinh tế của Việt nói riêng và thế giới nói chung.

Thị trường bán lẻ trong những tháng cuối năm 2022 ghi nhận sự tăng trưởng vượt bậc. Theo số liệu của Tổng cục Thống kê, tổng mức bán lẻ hàng hóa và doanh thu dịch vụ tiêu dùng tháng 11/2022 ước đạt 514,1 nghìn tỷ đồng, tăng 2,6% so với tháng trước và tăng 17,5% so với cùng kỳ năm trước. Và cũng trong năm 2022, thị trường bán lẻ toàn cầu đạt giá trị gần 23 nghìn tỷ USD.

Là lợi thế cũng là thử thách cho doanh nghiệp khi thị trường càng lớn thì mức cạnh tranh giữa các doanh nghiệp càng cao. Và việc xác định đúng phân khúc khách hàng, đưa ra các quyết định về chiến lược kinh doanh cũng như tối ưu nguồn lực là những điều cần thiết để doanh nghiệp có thể cạnh tranh và phát triển tốt trong thị trường này.

Trong đó, dữ liệu đóng một vai trò quan trọng trong cuộc cạnh tranh này, bằng cách khai thác và phân tích dữ liệu được thu thập trong quá trình hoạt động, doanh nghiệp bán lẻ có thể sử dụng kết quả để phục vụ tốt cho các chiến lược nhằm đẩy mạnh doanh thu cho mình.

Để quá trình lưu trữ, phân tích dữ liệu cũng như tạo ra các báo cáo chính xác các thông tin về ngành bán lẻ thì việc có một kho dữ liệu để lưu trữ và thao tác là điều cần thiết.

Vì vậy, nhóm chúng em quyết định xây dựng kho dữ liệu cho các giao dịch của ngành bán lẻ.

### **1.1.2. Giải pháp**

Dựa trên nhu cầu thống kê, phân tích và khai thác dữ liệu các đơn vị bán lẻ. Giải pháp là xây dựng kho dữ liệu phục vụ mục đích phân tích, khai thác, và tạo báo cáo tổng. Đưa ra các dự đoán về nhu cầu khách hàng để xây dựng mô hình bán lẻ hiệu quả nhất

### **1.1.3. Mục tiêu và ý nghĩa của đồ án**

**Mục tiêu:** Việc xây dựng và phát triển ứng dụng nhằm tạo ra một nền tảng dữ liệu hoàn chỉnh và chính xác, cung cấp thông tin và giúp cho các người dùng kinh doanh có được cái nhìn tổng thể về hoạt động bán hàng của công ty. Nhờ đó có thể dễ dàng truy xuất, phân tích và đưa ra quyết định dựa trên dữ liệu, từ đó giúp tăng hiệu quả kinh doanh, tối ưu hóa chiến lược bán hàng và đưa ra những quyết định đúng đắn.

**Ý nghĩa của đồ án:** Giúp cho người dùng có thể thu thập, lưu trữ và quản lý dữ liệu bán hàng của mình một cách hiệu quả và có hệ thống. Việc xây dựng data warehouse cũng giúp giảm thiểu tình trạng dữ liệu phân tán, trùng lặp, không chính xác, giúp cho việc phân tích dữ liệu trở nên dễ dàng hơn và cải thiện khả năng dự báo kinh doanh cho người dùng. Ngoài ra, việc xây dựng data warehouse còn giúp tiết kiệm thời gian và chi phí khi thu thập và phân tích dữ liệu, từ đó tăng hiệu quả kinh doanh và cạnh tranh trên thị trường.

## **1.2. Giới thiệu tổng quan dataset**

### **1.2.1. Nguồn dữ liệu sử dụng**

Nguồn dữ liệu được thu thập từ [data.mendeley.com](http://data.mendeley.com), dataset [stackoverflow.com](http://stackoverflow.com) DataCo Smart Supply Chain For Big Data Analysis.

#### **1.2.1.1. Giới thiệu nội cấp dữ liệu**

Kho lưu trữ dữ liệu Mendeley của Elsevier là một thành viên tham gia của dự án GREI Văn phòng chiến lược khoa học dữ liệu (ODSS) của Viện Y tế Quốc gia (NIH). GREI bao gồm bảy kho lưu trữ chung đã được thiết lập do NIH tài trợ để làm việc cùng nhau nhằm thiết lập siêu dữ liệu nhất quán, phát triển các trường hợp sử dụng để chia sẻ dữ liệu, đào tạo và giáo dục các nhà nghiên cứu về dữ liệu FAIR và tầm quan trọng của việc chia sẻ dữ liệu, v.v.

Kho lưu trữ dữ liệu chung của Mendeley Data được cung cấp bởi Digital Commons Data.

Digital Commons Data đã nhận được chứng nhận ISO/IEC 27001:2013 về các phương pháp quản lý thông tin an toàn.

Digital Commons Data sử dụng các giao diện bảo mật như TLS 1.2 trở lên. Tất cả dữ liệu đang truyền và khi lưu trữ được mã hóa bằng ít nhất AES-256 hoặc tương đương. Kho lưu trữ hỗ trợ lưu trữ dữ liệu được mã hóa an toàn cả trong và ngoài trang web. Nó tận dụng công cụ AWS KMS để quản lý khóa an toàn.

Dữ liệu Digital Commons được lưu trữ trong các trung tâm dữ liệu AWS. AWS có chứng nhận tuân thủ ISO/IEC 27001:2013, 27017:2015, 27018:2019 và ISO/IEC 9001:2015.

### **1.2.1.2. Hướng dẫn tải dataset và các dataset khác của nhà cung cấp**

Link tải dataset:

<https://data.mendeley.com/datasets/8gx2fgv2k6/5/files/72784be5-36d3-44fe-b75d-0edbf1999f65>

Bao gồm dữ liệu kết quả khảo sát với định dạng CSV.

### **1.2.2. Mô tả chi tiết cơ sở dữ liệu**

Bộ dữ liệu về chuỗi cung ứng của công ty DataCo Global. Gồm các giao dịch của công ty với khách hàng. Tập dữ liệu gồm 53 thuộc tính khác nhau, từ thông tin đặt hàng và vận chuyển đến thông tin bán hàng, 180.519 hàng và các tính năng bao gồm sự kết hợp giữa dữ liệu văn bản và dữ liệu số, chẳng hạn như vị trí đặt hàng và dữ liệu bán hàng số. Cụ thể có 24 cột ký tự và 28 cột số.

#### **1.2.2.1. Dữ liệu sau khi trích xuất**

Thực hiện trích xuất dữ liệu:

Lấy 180.519 (dòng) \* 40 (cột) phục vụ cho việc xây dựng kho dữ liệu:

Dữ liệu sau khi trích xuất từ dữ liệu gốc:

Type	Days for shipping (n)	Days for shipment (scheduled)	Delivery Status	Late_delivery_risk	Category Id	Category Name	Customer City	Customer Country	Customer Name	Customer Id	Customer Name
DEBIT	3	4	Advance shipping	0	73	Sporting Goods	Caguas	Puerto Rico	Cally	20755	Holloway
TRANSFER	5	4	Late delivery	1	73	Sporting Goods	Caguas	Puerto Rico	Irene	19492	Luna
CASH	4	4	Shipping on time	0	73	Sporting Goods	San Jose	EE. UU.	Gillian	19491	Maldonado
DEBIT	3	4	Advance shipping	0	73	Sporting Goods	Los Angeles	EE. UU.	Tana	19490	Tate
PAYMENT	2	4	Advance shipping	0	73	Sporting Goods	Caguas	Puerto Rico	Orli	19489	Hendricks
TRANSFER	6	4	Shipping canceled	0	73	Sporting Goods	Tonawanda	EE. UU.	Kimberly	19488	Flowers
DEBIT	2	1	Late delivery	1	73	Sporting Goods	Caguas	Puerto Rico	Constance	19487	Terrell
TRANSFER	2	1	Late delivery	1	73	Sporting Goods	Miami	EE. UU.	Erica	19486	Stevens
CASH	3	2	Late delivery	1	73	Sporting Goods	Caguas	Puerto Rico	Nichole	19485	Olsen
CASH	2	1	Late delivery	1	73	Sporting Goods	San Ramon	EE. UU.	Oprah	19484	Delacruz
TRANSFER	6	2	Shipping canceled	0	73	Sporting Goods	Caguas	Puerto Rico	Germane	19483	Short
TRANSFER	5	2	Late delivery	1	73	Sporting Goods	Freepost	EE. UU.	Freya	19482	Robbins
TRANSFER	4	2	Late delivery	1	73	Sporting Goods	Salinas	EE. UU.	Cassandra	19481	Jensen
DEBIT	2	1	Late delivery	1	73	Sporting Goods	Caguas	Puerto Rico	Natalie	19480	McFadden
TRANSFER	2	1	Late delivery	1	73	Sporting Goods	Peabody	EE. UU.	Kimberley	19479	Sharpe
DEBIT	2	1	Late delivery	1	73	Sporting Goods	Caguas	Puerto Rico	Sade	19478	Lancaster
PAYMENT	5	2	Late delivery	1	73	Sporting Goods	Canovanas	Puerto Rico	Brynn	19477	Giles
CASH	2	1	Late delivery	1	73	Sporting Goods	Paramount	EE. UU.	Ciara	19476	Bird
DEBIT	2	1	Late delivery	1	73	Sporting Goods	Caguas	Puerto Rico	Bo	19475	Griffin
PAYMENT	0	0	Shipping on time	0	73	Sporting Goods	Mount Prospect	EE. UU.	Kim	19474	Simon
TRANSFER	0	0	Shipping on time	0	73	Sporting Goods	Long Beach	EE. UU.	Kellie	19473	Farmer
TRANSFER	5	4	Late delivery	1	73	Sporting Goods	Caguas	Puerto Rico	Alma	19472	Conley
TRANSFER	4	2	Late delivery	1	73	Sporting Goods	Rancho Cordova	EE. UU.	Yeo	19471	Bird

Customer Segm	Customer Sta	Customer Street	Customer Zipco	Department	Department Name	Market	Order City	Order Country	order date [DateOrder]	Order Id
Consumer	PR	5365 Noble Nectar Island	725	2	Fitness	Pacific Asia	Bekasi	Indonesia	1/31/2018 22:56	77202
Consumer	PR	2679 Rustic Loop	725	2	Fitness	Pacific Asia	Bikarner	India	1/13/2018 12:27	75939
Consumer	CA	8510 Round Bear Gate	95125	2	Fitness	Pacific Asia	Bikarner	India	1/13/2018 12:06	75938
Home Office	CA	3200 Amber Bend	90027	2	Fitness	Pacific Asia	Townsville	Australia	1/13/2018 11:45	75937
Corporate	PR	8671 Iron Anchor Corners	725	2	Fitness	Pacific Asia	Townsville	Australia	1/13/2018 11:24	75936
Consumer	NY	2122 Hazy Corner	14150	2	Fitness	Pacific Asia	Toowoomba	Australia	1/13/2018 11:03	75935
Home Office	PR	1879 Green Pine Bank	725	2	Fitness	Pacific Asia	Guangzhou	China	1/13/2018 10:42	75934
Corporate	FL	7595 Cotton Log Row	33162	2	Fitness	Pacific Asia	Guangzhou	China	1/13/2018 10:21	75933
Corporate	PR	2051 Dusty Route	725	2	Fitness	Pacific Asia	Guangzhou	China	1/13/2018 10:00	75932
Corporate	CA	9139 Blue Blossom Court	94583	2	Fitness	Pacific Asia	Guangzhou	China	1/13/2018 9:39	75931
Corporate	PR	4058 Quiet Heights	725	2	Fitness	Pacific Asia	Tokio	Japón	1/13/2018 9:18	75930
Consumer	NY	3243 Shady Corner	11520	2	Fitness	Pacific Asia	Manado	Indonesia	1/13/2018 8:57	75929
Corporate	CA	131 Sunny Treasure Green	93905	2	Fitness	Pacific Asia	Manado	Indonesia	1/13/2018 8:36	75928
Corporate	PR	2531 Wishing Square	725	2	Fitness	Pacific Asia	Sangli	India	1/13/2018 8:15	75927
Corporate	MA	6417 Silver Towers	1960	2	Fitness	Pacific Asia	Sangli	India	1/13/2018 7:54	75926
Corporate	PR	257 Harvest Close	725	2	Fitness	Pacific Asia	Sangli	India	1/13/2018 7:33	75925
Corporate	PR	7342 Hazy Beacon Park	729	2	Fitness	Pacific Asia	Seúl	Corea del Sur	1/13/2018 7:12	75924
Corporate	CA	7787 Lazy Corners	90723	2	Fitness	Pacific Asia	Jabalpur	India	1/13/2018 6:51	75923
Consumer	PR	5136 Rustic Pioneer Estates	725	2	Fitness	Pacific Asia	Jabalpur	India	1/13/2018 6:30	75922
Corporate	CA	7111 Silent Fox Gardens	90813	2	Fitness	Pacific Asia	Jabalpur	India	1/13/2018 5:48	75920
Corporate	PR	3732 Old Mountain Bank	725	2	Fitness	Pacific Asia	Jabalpur	India	1/13/2018 5:27	75919

Order Item Discou	Order Item Discount Ra	Order Item	Order Item Product Price	Order Item Profit Ratio	Order Item Quant	Sales	Order Item To	Order Profit Per Ord	Order Region	Order State
13.10999966	0.03999999	180517	327.75	0.28999999	1	327.75	314.6400146	91.25	Southeast Asia	Java Occidental
16.38999939	0.05000001	179254	327.75	-0.80000001	1	327.75	311.3599854	-249.0899963	South Asia	Rajastán
18.03000069	0.05999999	179253	327.75	-0.80000001	1	327.75	309.7200012	-247.7799988	South Asia	Rajastán
22.94000053	0.07	179252	327.75	0.07999998	1	327.75	304.8099976	22.86000061	Oceania	Queensland
29.5	0.09000004	179251	327.75	0.44999998	1	327.75	298.25	134.2100067	Oceania	Queensland
32.77999878	0.10000001	179250	327.75	0.05999999	1	327.75	294.980011	18.57999992	Oceania	Queensland
39.33000183	0.13999997	179249	327.75	0.330000013	1	327.75	288.4200134	95.18000031	Eastern Asia	Guangdong
42.61000061	0.12999995	179248	327.75	0.23999995	1	327.75	285.1400146	68.43000031	Eastern Asia	Guangdong
49.15999985	0.15000006	179247	327.75	0.47999998	1	327.75	278.5699963	133.7200012	Eastern Asia	Guangdong
52.43999863	0.15999996	179246	327.75	0.47999998	1	327.75	275.3099976	132.1499939	Eastern Asia	Guangdong
55.72000122	0.17000002	179245	327.75	0.47999998	1	327.75	272.0299984	130.5800018	Eastern Asia	Tokio
59	0.18000007	179244	327.75	0.17000002	1	327.75	268.7600098	45.68999863	Southeast Asia	Célebes Septentrional
65.55000305	0.20000003	179243	327.75	0.07999998	1	327.75	262.2000122	21.7600023	Southeast Asia	Célebes Septentrional
81.94000244	0.25	179242	327.75	0.10000001	1	327.75	245.8099976	24.57999992	South Asia	Maharashtra
0	0	179241	327.75	0.05000001	1	327.75	327.75	16.38999939	South Asia	Maharashtra
3.27999971	0.01	179240	327.75	-0.800000012	1	327.75	324.4700012	-259.5799866	South Asia	Maharashtra
6.55999943	0.02	179239	327.75	-0.769999981	1	327.75	321.2000122	-246.3600006	Eastern Asia	Seúl
9.82999924	0.02999999	179238	327.75	0.07999998	1	327.75	317.9200134	23.84000015	South Asia	Madhya Pradesh
13.10999966	0.03999999	179237	327.75	0.330000013	1	327.75	314.6400146	102.2600021	South Asia	Madhya Pradesh
16.38999939	0.05000001	179236	327.75	0.28000001	1	327.75	311.3599854	87.18000031	South Asia	Madhya Pradesh
18.03000069	0.05999999	179235	327.75	0.5	1	327.75	309.7200012	154.8600006	South Asia	Madhya Pradesh
22.94000053	0.07	179234	327.75	0.27000001	1	327.75	304.8099976	82.30000305	South Asia	Madhya Pradesh
29.5	0.09000004	179233	327.75	0.07999998	1	327.75	298.25	22.37000084	Oceania	Victoria

Order State	Order Status	Product Card	Product Category	Product Name	Product Pr	shipping date [DateOrder]	Shipping M
Java Occidental	COMPLETE	1360	73	Smart watch	327.75	2/3/2018 22:56	Standard Class
Java Occidental	PENDING	1360	73	Smart watch	327.75	1/18/2018 12:27	Standard Class
Java Occidental	CLOSED	1360	73	Smart watch	327.75	1/1/2018 12:06	Standard Class
Queensland	COMPLETE	1360	73	Smart watch	327.75	1/16/2018 11:45	Standard Class
Queensland	PENDING_PAYMENT	1360	73	Smart watch	327.75	1/15/2018 11:24	Standard Class
Queensland	CANCELED	1360	73	Smart watch	327.75	1/19/2018 11:03	Standard Class
Guangdong	COMPLETE	1360	73	Smart watch	327.75	1/15/2018 10:42	First Class
Guangdong	PROCESSING	1360	73	Smart watch	327.75	1/15/2018 10:21	First Class
Guangdong	CLOSED	1360	73	Smart watch	327.75	1/16/2018 10:00	Second Class
Guangdong	CLOSED	1360	73	Smart watch	327.75	1/15/2018 9:39	First Class
Okio	SUSPECTED_FRAUD	1360	73	Smart watch	327.75	1/19/2018 9:18	Second Class
Elebes Septentrional	PENDING	1360	73	Smart watch	327.75	1/18/2018 8:57	Second Class
Elebes Septentrional	PENDING	1360	73	Smart watch	327.75	1/17/2018 8:36	Second Class
Elebes Septentrional	COMPLETE	1360	73	Smart watch	327.75	1/15/2018 8:15	First Class
Elebes Septentrional	PROCESSING	1360	73	Smart watch	327.75	1/15/2018 7:54	First Class
Elebes Septentrional	COMPLETE	1360	73	Smart watch	327.75	1/15/2018 7:33	First Class
Madhya Pradesh	PENDING_PAYMENT	1360	73	Smart watch	327.75	1/18/2018 7:12	Second Class
Madhya Pradesh	CLOSED	1360	73	Smart watch	327.75	1/15/2018 6:51	First Class
Madhya Pradesh	COMPLETE	1360	73	Smart watch	327.75	1/15/2018 6:30	First Class
Madhya Pradesh	PENDING_PAYMENT	1360	73	Smart watch	327.75	1/13/2018 18:09	Same Day
Madhya Pradesh	PROCESSING	1360	73	Smart watch	327.75	1/13/2018 17:48	Same Day
Madhya Pradesh	PENDING	1360	73	Smart watch	327.75	1/18/2018 5:27	Standard Class

### 1.2.2.2. Mô tả chi tiết các thuộc tính trong tập dữ liệu

Type	Phương thức thanh toán của giao dịch
------	--------------------------------------

Days for shipping (real)	Số ngày giao hàng thực tế
Days for shipment (scheduled)	Số ngày giao hàng dự kiến
Delivery Status	Tình trạng giao hàng
Late_delivery_risk	Xác định đơn hàng có giao trễ hay không
Category Id	Mã danh mục sản phẩm
Category Name	Tên danh mục sản phẩm
Customer City	Thành phố nơi khách hàng thực hiện mua hàng
Customer Country	Quốc gia nơi khách hàng thực hiện giao dịch mua
Customer Fname	Tên của khách hàng
Customer Id	Mã số của khách hàng
Customer Lname	Họ của khách hàng
Customer Segment	Thể loại khách hàng: Consumer, Corporate, Home Office
Customer State	Tiểu bang nơi cửa hàng đăng ký mua hàng trực thuộc
Customer Street	Địa chỉ nhà riêng của khách hàng
Customer Zipcode	Mã bưu điện của bảng khách hàng ở
Department Id	ID của bộ phận của công ty thực hiện giao dịch
Department Name	Tên của bộ phận của công ty thực hiện giao dịch
Market	Thị trường nơi đơn hàng được giao: Africa, Europe, LATAM, Pacific Asia, USCA
Order City	Thành phố nơi đơn hàng được giao tới

Order Country	Đất nước nơi đơn hàng được giao tới
order date (DateOrders)	ngày thực hiện order
Order Id	ID của order
Order Item Discount	Giá trị chiết khấu đơn hàng
Order Item Discount Rate	Tỷ lệ phần trăm giảm giá mặt hàng đặt hàng
Order Item Product Price	Giá niêm yết của sản phẩm
Order Item Profit Ratio	Tỉ lệ lợi nhuận trên một mặt hàng
Order Item Quantity	Số lượng sản phẩm cho mỗi đơn đặt hàng
Sales	Giá trị đơn hàng
Order Item Total	Tổng giá trị đơn hàng sau khi trừ chiết khấu
Order Profit Per Order	Tổng loại nhuật của đơn hàng
Order Region	Khu vực của thế giới nơi đơn đặt hàng được giao
Order State	Bang của khu vực nơi đơn đặt hàng được giao
Order Status	Trạng thái đơn hàng: COMPLETE, PENDING, CLOSED, PENDING_PAYMENT, CANCELED, PROCESSING, SUSPECTED_FRAUD, ON_HOLD, PAYMENT REVIEW
Product Card Id	Mã sản phẩm
Product Category Id	Danh mục của sản phẩm

Product Name	Tên sản phẩm
Product Price	Giá của sản phẩm
Shipping date (DateOrders)	Ngày đơn hàng được tạo
Shipping Mode	Các chế độ vận chuyển: Standard Class, First Class, Second Class, Same Day

### **1.3. Giới thiệu các công cụ sử dụng trong đồ án**

#### **1.3.1. Tổng quan về Visual Studio 2019**

Microsoft Visual Studio là một môi trường phát triển tích hợp (IDE) từ Microsoft. Nó được sử dụng để phát triển chương trình máy tính cho Microsoft Windows, cũng như các trang web, các ứng dụng web và các dịch vụ web. Visual Studio sử dụng nền tảng phát triển phần mềm của Microsoft như Windows API, Windows Forms, Windows Presentation Foundation, Windows Store và Microsoft Silverlight. Nó có thể sản xuất cả hai ngôn ngữ máy và mã số quản lý.

Visual Studio 2019 hỗ trợ nhiều ngôn ngữ lập trình, bao gồm C++, C#, Visual Basic, F#, JavaScript, Python và nhiều ngôn ngữ khác. Nó cũng cung cấp nhiều tính năng hỗ trợ cho phát triển ứng dụng web và di động, bao gồm hỗ trợ cho ASP.NET, Node.js, React Native và Xamarin.

Các tính năng khác của Visual Studio 2019 bao gồm tích hợp Git và tương tác với các kho lưu trữ phổ biến như GitHub, tính năng gỡ lỗi, IntelliSense để giúp hỗ trợ tự động hoàn thành mã, kiểm tra lỗi cú pháp và gợi ý các khả năng trong quá trình nhập mã. Gỡ lỗi mạnh mẽ cho phép gỡ lỗi với cả mã nguồn và máy, bao gồm bộ xử lý, bộ nhớ và tài nguyên. Và có hỗ trợ DevOps: cung cấp các tính năng tích hợp để hỗ trợ quy trình DevOps, bao gồm kiểm thử tự động, triển khai liên tục và giám sát.

Visual Studio 2019 được phát hành với các phiên bản khác nhau, bao gồm: Community, Professional, Enterprise và Test Professional.

- Phiên bản Community được cung cấp miễn phí cho các cá nhân và các doanh nghiệp nhỏ để sử dụng trong các dự án phi thương mại và thương mại nhỏ.
- Phiên bản Professional và Enterprise cung cấp các tính năng và công cụ phát triển phần mềm chuyên nghiệp hơn và phù hợp cho các doanh nghiệp lớn.
- Phiên bản Test Professional là phiên bản dành riêng cho các nhà phát triển.

#### **1.3.2. Giới thiệu về SQL Server Data Tools**

SQL Server Data Tools (SSDT) là một công cụ tích hợp trong Visual Studio và được sử dụng để phát triển, triển khai và quản lý cơ sở dữ liệu SQL Server. SSDT hỗ trợ các tính năng như thiết kế cơ sở dữ liệu, lập trình và triển khai các biểu đồ cơ sở dữ liệu, các truy vấn SQL và các lệnh xử lý dữ liệu.

Với SSDT, người phát triển có thể tạo các dự án cơ sở dữ liệu, quản lý phiên bản và phát triển các mã tùy chỉnh cho cơ sở dữ liệu của họ. SSDT cũng cung cấp các công cụ để thực hiện kiểm tra đơn vị, kiểm tra tích hợp và kiểm tra hiệu suất cho cơ sở dữ liệu. SSDT được tích hợp vào Visual Studio và có sẵn trên nhiều phiên bản SQL Server, bao gồm SQL Server 2019 Developer.

SSDT cũng hỗ trợ các công cụ tạo và quản lý các lược đồ, các trình động và các tài khoản truy cập đến cơ sở dữ liệu. Bên cạnh đó, SSDT còn hỗ trợ tích hợp với các công cụ như Git để quản lý mã nguồn của các dự án cơ sở dữ liệu.

### 1.3.3. Giới thiệu ngôn ngữ truy vấn SQL

Ngôn ngữ truy vấn SQL (Structured Query Language) là một ngôn ngữ được sử dụng để truy vấn và quản lý cơ sở dữ liệu quan hệ (RDBMS). SQL là một ngôn ngữ chung và được sử dụng bởi hầu hết các hệ quản trị cơ sở dữ liệu (DBMS) như MySQL, Oracle, PostgreSQL, SQL Server, SQLite và nhiều hơn nữa.

SQL được thiết kế để cho phép người dùng truy cập và quản lý dữ liệu trong cơ sở dữ liệu quan hệ. Nó cho phép người dùng truy vấn cơ sở dữ liệu để lấy thông tin từ các bảng hoặc chèn, cập nhật hoặc xóa dữ liệu trong cơ sở dữ liệu. SQL cũng được sử dụng để tạo và quản lý các đối tượng cơ sở dữ liệu như bảng, chỉ mục và ràng buộc.

Các câu lệnh SQL phổ biến bao gồm SELECT, INSERT, UPDATE và DELETE. Câu lệnh SELECT được sử dụng để truy vấn dữ liệu từ các bảng trong cơ sở dữ liệu và trả về kết quả dưới dạng một bảng kết quả. Câu lệnh INSERT được sử dụng để chèn dữ liệu mới vào bảng trong cơ sở dữ liệu. Câu lệnh UPDATE được sử dụng để cập nhật dữ liệu hiện có trong bảng và câu lệnh DELETE được sử dụng để xóa dữ liệu khỏi bảng.

SQL cũng hỗ trợ các phép toán logic như AND, OR và NOT để giúp người dùng truy vấn dữ liệu theo nhiều điều kiện khác nhau. Ngoài ra, SQL cũng cung cấp các chức năng toán học và chuỗi để tính toán và xử lý dữ liệu trong cơ sở dữ liệu.

Tổng quan về ngôn ngữ truy vấn SQL là rất quan trọng trong việc phát triển các ứng dụng web hoặc desktop liên quan đến cơ sở dữ liệu. Hiểu biết về SQL cho phép người phát triển tương tác với cơ sở dữ liệu, truy vấn và lưu trữ dữ liệu một cách hiệu quả và nhanh chóng.

### **1.3.4. Tổng quan về SQL Server 2019**

QL Server là một hệ quản trị cơ sở dữ liệu quan hệ (RDBMS) được phát triển bởi Microsoft. Nó cho phép quản lý và lưu trữ dữ liệu trong một hệ thống cơ sở dữ liệu tập trung và cung cấp các công cụ cho việc quản lý cơ sở dữ liệu, truy vấn dữ liệu, bảo mật dữ liệu và phân tích dữ liệu.

Một số phiên bản chính của SQL server 2019 như: Enterprise, Standard, Express, Developer, Web, và các phiên bản Evaluation. Mỗi phiên bản có tính năng và giá cả khác nhau.

Các tính năng của SQL Server 2019 so với các phiên bản trước đó:

- Dịch vụ Big Data Clusters: Dịch vụ này cho phép lưu trữ dữ liệu lớn và xử lý dữ liệu phân tán với sự kết hợp giữa SQL Server và Apache Spark.
- Tính năng Intelligent Query Processing: Tính năng này cải tiến quá trình truy vấn dữ liệu để tăng tốc độ truy vấn và cải thiện hiệu suất.
- Tính năng Accelerated Database Recovery: Tính năng này cho phép phục hồi cơ sở dữ liệu nhanh hơn và giảm thiểu thời gian chết của cơ sở dữ liệu.
- Tính năng Always On Availability Groups: Tính năng này cho phép tạo và quản lý các nhóm sẵn sàng luôn để đảm bảo tính sẵn sàng cao cho cơ sở dữ liệu.
- Hỗ trợ cho các loại dữ liệu mới: SQL Server 2019 hỗ trợ các loại dữ liệu mới như Graph, JSON và Machine Learning.
- Hỗ trợ Linux: SQL Server 2019 hỗ trợ cài đặt và chạy trên hệ điều hành Linux.
- Tính năng Data Discovery and Classification: Tính năng này cho phép phát hiện và phân loại dữ liệu nhạy cảm để đảm bảo tuân thủ quy định về bảo mật dữ liệu.
- Tính năng Machine Learning Services: SQL Server 2019 tích hợp sẵn các dịch vụ Machine Learning để phân tích dữ liệu và tạo các mô hình dự đoán.
- Tính năng Data Virtualization: Tính năng này cho phép truy cập dữ liệu từ nhiều nguồn khác nhau bằng cách ảo hóa dữ liệu và đưa chúng vào cơ sở dữ liệu một cách trực tiếp.

Trong đồ án học phần này sử dụng phiên bản SQL Server 2019 Developer.

SQL Server 2019 Developer là một phiên bản của SQL Server 2019 được thiết kế cho các nhà phát triển và các nhà quản trị cơ sở dữ liệu phát triển, kiểm thử và triển

khai các ứng dụng cơ sở dữ liệu. SQL Server 2019 Developer cung cấp tất cả các tính năng của SQL Server 2019 Enterprise Edition, nhưng chỉ được sử dụng cho mục đích phát triển và kiểm thử. Nó có sẵn miễn phí cho các nhà phát triển và các thành viên của các chương trình MSDN (Microsoft Developer Network).

SQL Server 2019 Developer cung cấp nhiều tính năng mới và cải tiến so với các phiên bản trước đó của SQL Server. Một trong những tính năng mới quan trọng nhất là hỗ trợ cho bộ xử lý đám mây Azure SQL Edge, cho phép các ứng dụng cơ sở dữ liệu được triển khai và chạy trên các thiết bị IoT (Internet of Things). Nó cũng cung cấp tính năng Machine Learning Services được tích hợp sẵn, cho phép các nhà phát triển tạo, huấn luyện và triển khai các mô hình Machine Learning trực tiếp trên cơ sở dữ liệu.

### **1.3.5. Giới thiệu về SQL Server Analysis Service**

SSAS (SQL Server Analysis Services) là một phần của SQL Server của Microsoft, được sử dụng để phân tích dữ liệu và xây dựng các mô hình dữ liệu. Nó cung cấp một giải pháp OLAP (Online Analytical Processing) cho các doanh nghiệp để xử lý và phân tích các bộ dữ liệu lớn.

SSAS hỗ trợ hai loại mô hình dữ liệu: mô hình đa chiều và mô hình phẳng. Mô hình đa chiều cho phép dữ liệu được sắp xếp thành các khối dữ liệu có liên kết với nhau, còn mô hình phẳng làm việc trực tiếp với các bảng dữ liệu.

Các tính năng chính của SSAS bao gồm:

- Khả năng phân tích dữ liệu dưới nhiều góc độ khác nhau.
- Cung cấp các mẫu dữ liệu và hình thức dữ liệu độc đáo để phân tích và truy xuất dữ liệu.
- Tăng tốc quá trình phân tích dữ liệu bằng cách lưu trữ dữ liệu trên các máy chủ tách biệt.
- Tính năng chia sẻ dữ liệu giữa các phân tích và các ứng dụng khác trong doanh nghiệp.
- Hỗ trợ nhiều nguồn dữ liệu khác nhau bao gồm các cơ sở dữ liệu quan hệ, các tệp dữ liệu văn bản và các tệp dữ liệu XML.

SSAS là một công cụ quan trọng cho các doanh nghiệp trong việc phân tích và quản lý dữ liệu lớn. Nó giúp tăng tốc độ phân tích dữ liệu và cung cấp các khía cạnh mới để hiểu và tận dụng dữ liệu của doanh nghiệp.

### 1.3.6. Giới thiệu Microsoft Power BI

Microsoft Power BI là một nền tảng phân tích dữ liệu và trực quan hóa được phát triển bởi Microsoft. Nó cung cấp các công cụ và tính năng cho phép người dùng kết nối, chuẩn bị, trực quan hóa và chia sẻ dữ liệu một cách dễ dàng và hiệu quả.

Power BI cho phép người dùng kết nối và tổ chức dữ liệu từ nhiều nguồn khác nhau như cơ sở dữ liệu, tệp Excel, dịch vụ đám mây và nhiều nguồn dữ liệu trực tuyến khác. Sau đó, người dùng có thể thực hiện các thao tác chuẩn bị dữ liệu để làm sạch, biến đổi và tạo ra mô hình dữ liệu.

Với các công cụ trực quan hóa mạnh mẽ, Power BI cho phép người dùng tạo ra các biểu đồ, đồ thị, bảng, bản đồ và báo cáo tương tác để hiển thị dữ liệu một cách dễ hiểu và hấp dẫn. Người dùng có thể tùy chỉnh giao diện và định dạng biểu đồ để phù hợp với nhu cầu trình bày thông tin.

Power BI cũng cung cấp khả năng phân tích dữ liệu mạnh mẽ bằng cách sử dụng các tính năng như lọc, sắp xếp, nhóm, tính toán cột mới và truy vấn dữ liệu. Người dùng có thể thực hiện các phân tích sâu hơn để khám phá mối quan hệ, xu hướng và thông tin bên trong dữ liệu.

Cuối cùng, Power BI cho phép người dùng chia sẻ báo cáo và tạo bảng điều khiển tương tác với người dùng khác thông qua các ứng dụng di động, trình duyệt web hoặc các báo cáo nhúng trong ứng dụng và trang web khác.

Với các tính năng và khả năng linh hoạt, Microsoft Power BI đã trở thành một công cụ phổ biến trong lĩnh vực phân tích dữ liệu và trực quan hóa, hỗ trợ các tổ chức và cá nhân trong việc hiểu và tận dụng thông tin từ dữ liệu để đưa ra quyết định thông minh.

## **Chương 2: Thiết kế xây dựng cơ sở dữ liệu**

Ở chương này sẽ phân tích các yêu cầu khai thác dữ liệu từ đó xây dựng kho dữ liệu đưa ra được Snowflake Schema và Constellation Schema.

### **2.1. Yêu cầu khai thác dữ liệu**

Dựa vào tập dữ liệu thì chúng ta cần đưa ra các báo cáo về doanh số tổng quan của cũng như chi tiết của từng sản phẩm dựa theo từng thuộc tính khác nhau trong tập dữ liệu và đưa ra các nhận xét về cách vận hành của các phương thức vận chuyển từ đó cải thiện hiệu quả. Ngoài ra, còn có thể so sánh hiệu quả bán hàng của các đất nước và khu vực khác nhau.

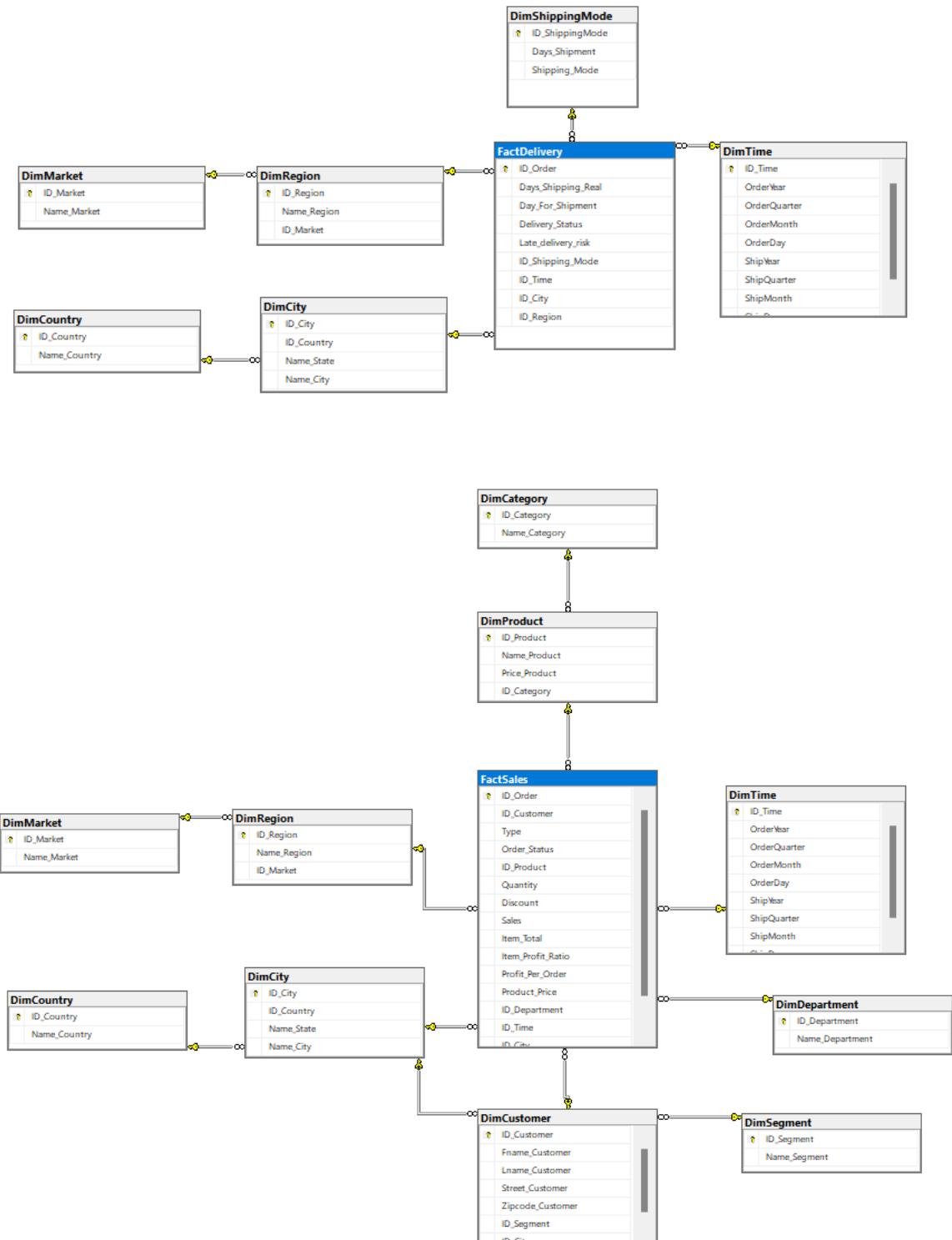
Từ các yêu cầu đó thì kho dữ liệu sẽ có:

- Hai bảng Fact gồm FactSales và FactDelivery cả hai đều thuộc loại Transactions. Business process của mỗi bản Fact:
  - FactSales: Dùng để phân tích và báo cáo doanh số, lợi nhuận và số lượng bán ra của từng sản phẩm theo từng thuộc tính.
  - FactDelivery: Báo cáo về hoạt động của việc vận chuyển đơn hàng đến cho người dùng
- Các bảng Dim gồm có: DimDepartment, DimShippingMode, DimTime, DimMarket, DimRegion, DimSegment, DimCustomer, DimCountry, DimCity, DimCategory, DimProduct. Dùng để cung cấp thông tin cho bảng Fact.

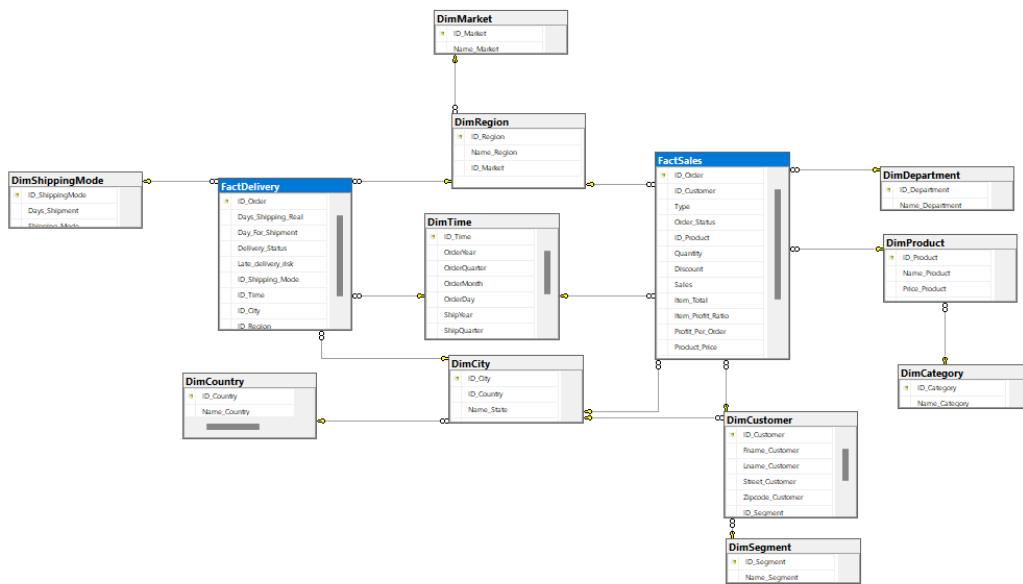
### **2.2. Chi tiết thiết kế**

#### **2.2.1. Lượt đồ**

- **Snowflake Schema**



## 1. Constellation Schema



### 2.2.2. Chi tiết bảng Fact

#### Bảng FactSales

Tên thuộc tính	Mô tả
ID_Line_Order	Mã định danh của mỗi hàng trong 1 đơn hàng
ID_Order	Định danh duy nhất cho mỗi đơn hàng
ID_Customer	Mã định danh khách hàng
Type	Phương thức thanh toán của giao dịch
Order_Status	Trạng thái đơn hàng
ID_Product	ID sản phẩm
Quantity	Số lượng
Discount	Giảm giá mặt hàng
Sales	Giá trị sản phẩm
Item_Total	Tổng tiền của sản phẩm được đặt trong đơn hàng
Item_Profit_Ratio	Tỉ lệ lợi nhuận trên 1 sản phẩm
Profit_Per_Order	Tổng lợi nhuận trên sản phẩm trên đơn hàng
Product_Price	Giá sản phẩm
ID_Time	Định danh thời gian khi đơn hàng được đặt
ID_City	Định danh thành phố

ID_Region	Định danh khu vực
ID_Department	ID của bộ phận của công ty thực hiện giao dịch

### Bảng FactDelivery

Tên thuộc tính	Mô tả
ID_Order	ID đơn hàng
Days_Shipping_Real	Số ngày giao thực tế
Day_For_Shipment	Số ngày giao dự kiến
Delivery_Status	Trạng thái giao hàng
Late_delivery_risk	Xác định đơn hàng có giao trễ hay không
ID_Shipping_Mode	ID xác định hình thức giao hàng
ID_Time	Định danh thời gian khi đơn hàng được giao
ID_City	Định danh thành phố
ID_Region	Định danh khu vực

#### 2.2.3. Chi tiết bảng Dim

### Bảng DimTime

Tên thuộc tính	Mô tả
ID_Time	Id thời gian
OrderYear	Năm đặt hàng
OrderQuarter	Quý của đơn đặt hàng
OrderMonth	Tháng của đơn đặt hàng
OrderDay	Ngày của đơn đặt hàng
ShipYear	Năm giao hàng
ShipQuarter	Quý giao hàng
ShipMonth	Tháng của giao hàng
ShipDay	Ngày giao hàng

### **Bảng DimShippingMode**

Tên thuộc tính	Mô tả
ID_ShippingMode	ID định danh cho mỗi phương thức vận chuyển
Days_Shipment	Số ngày giao dự kiến
Shipping_Mode	Tên của phương thức vận chuyển

### **Bảng DimMarket**

Tên thuộc tính	Mô tả
ID_Market	ID định danh cho mỗi thị trường kinh doanh
Name_Market	Tên của thị trường

### **Bảng DimRegion**

Tên thuộc tính	Mô tả
ID_Region	Mã định danh ID khu vực
Name_Region	Tên khu vực
ID_Market	Định danh thị trường kinh doanh của khu vực

### **Bảng DimDepartment**

Tên thuộc tính	Mô tả
ID_Department	ID phòng ban
Name_Department	Tên phòng ban

### **Bảng DimCategory**

Tên thuộc tính	Mô tả
ID_Category	Định danh cho mỗi danh mục
Name_Category	Tên danh mục

### **Bảng DimProduct**

Tên thuộc tính	Mô tả
ID_Product	Định danh duy nhất cho mỗi sản phẩm
Name_Product	Tên sản phẩm
Price_Product	Giá sản phẩm
ID_Category	Định danh danh mục sản phẩm mà sản phẩm thuộc về

### **Bảng DimCountry**

Tên thuộc tính	Mô tả
ID_Country	Định danh cho mỗi quốc gia
Name_Country	Tên quốc gia

### **Bảng DimCity**

Tên thuộc tính	Mô tả
ID_City	Định danh duy nhất cho mỗi thành phố
ID_Country	Định danh của quốc gia mà thành phố này thuộc về
Name_State	Tên bang
Name_City	Tên thành phố

### **Bảng DimSegment**

Tên thuộc tính	Mô tả
ID_Segment	Định danh duy nhất cho mỗi đối tượng được phân loại
Name_Segment	Tên của phân loại đối tượng

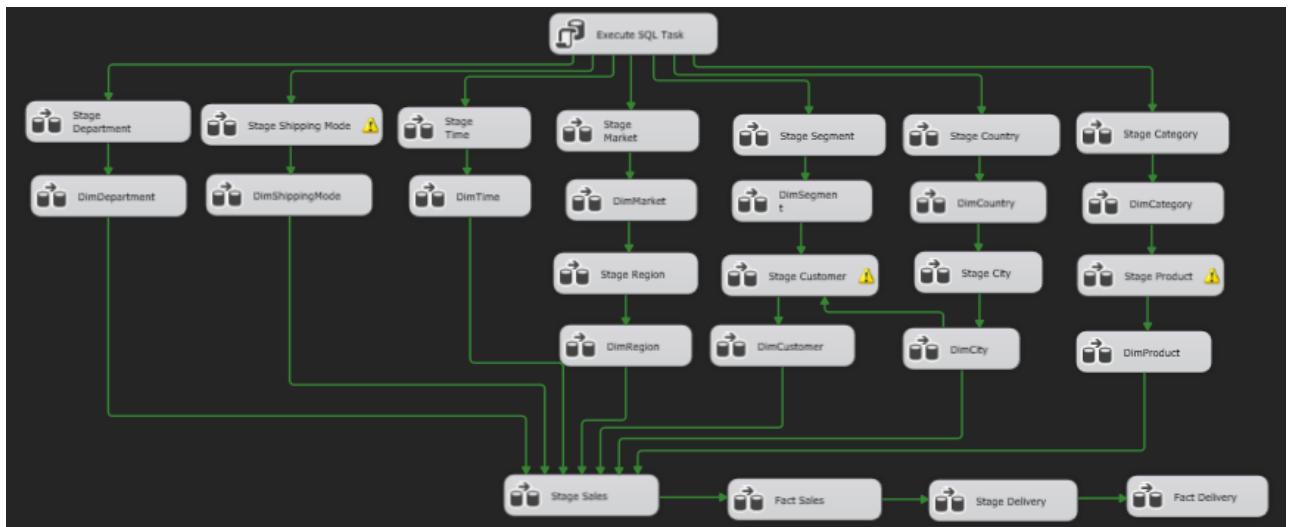
## Bảng DimCustomer

Tên thuộc tính	Mô tả
ID_Customer	Mã định danh của mỗi khách hàng
Fname_Customer	Tên của khách hàng
Lname_Customer	Họ của khách hàng
Street_Customer	Địa chỉ nhà riêng của khách hàng
Zipcode_Customer	Mã bưu điện của khách hàng
ID_Segment	Định danh của phân loại khách hàng
ID_City	Định danh thành phố

# Chương 3: Tích hợp dữ liệu vào kho (SISS)

Ở chương này, sử dụng công cụ SSIS xây dựng các gói tích hợp dữ liệu, cài đặt các gói chạy tự động. Đưa dữ liệu từ excel vào cơ sở dữ liệu Stage và đưa từ cơ sở dữ liệu Stage vào kho dữ liệu.

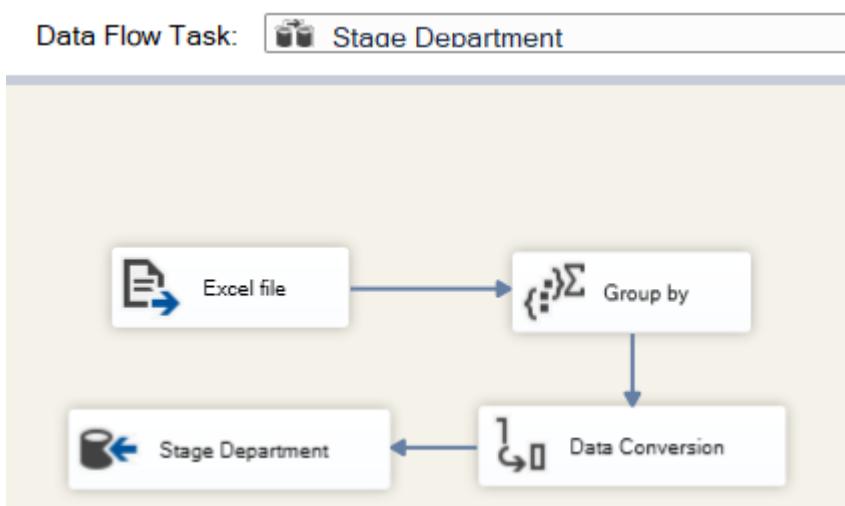
## 3.1. Luồng dữ liệu tổng quát



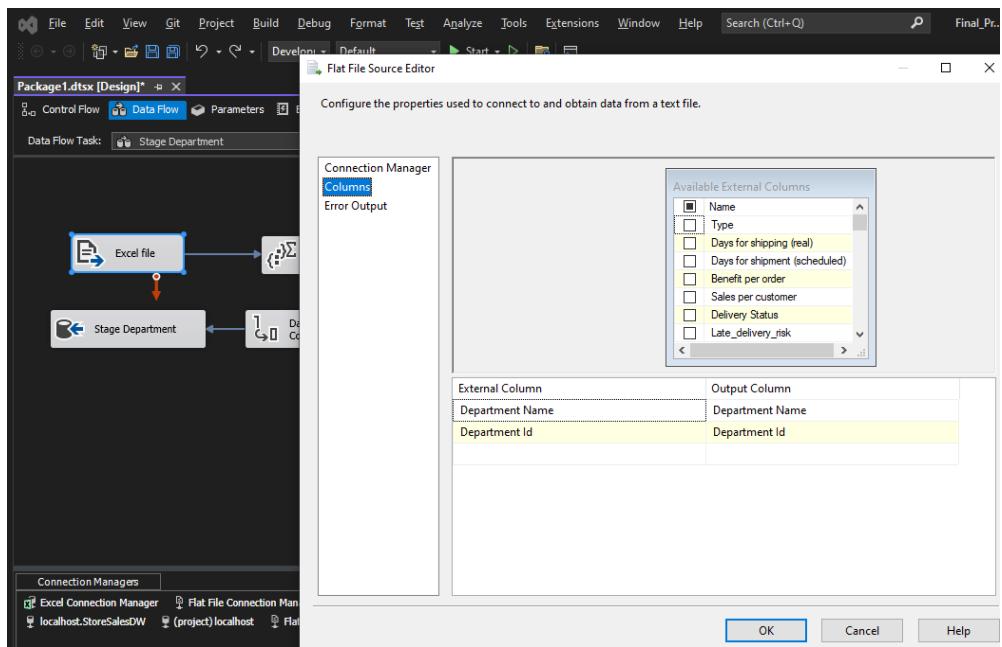
## 3.2. Quá trình đổ dữ liệu từ excel vào cơ sở dữ liệu Stage

### 3.2.1. Stage Department

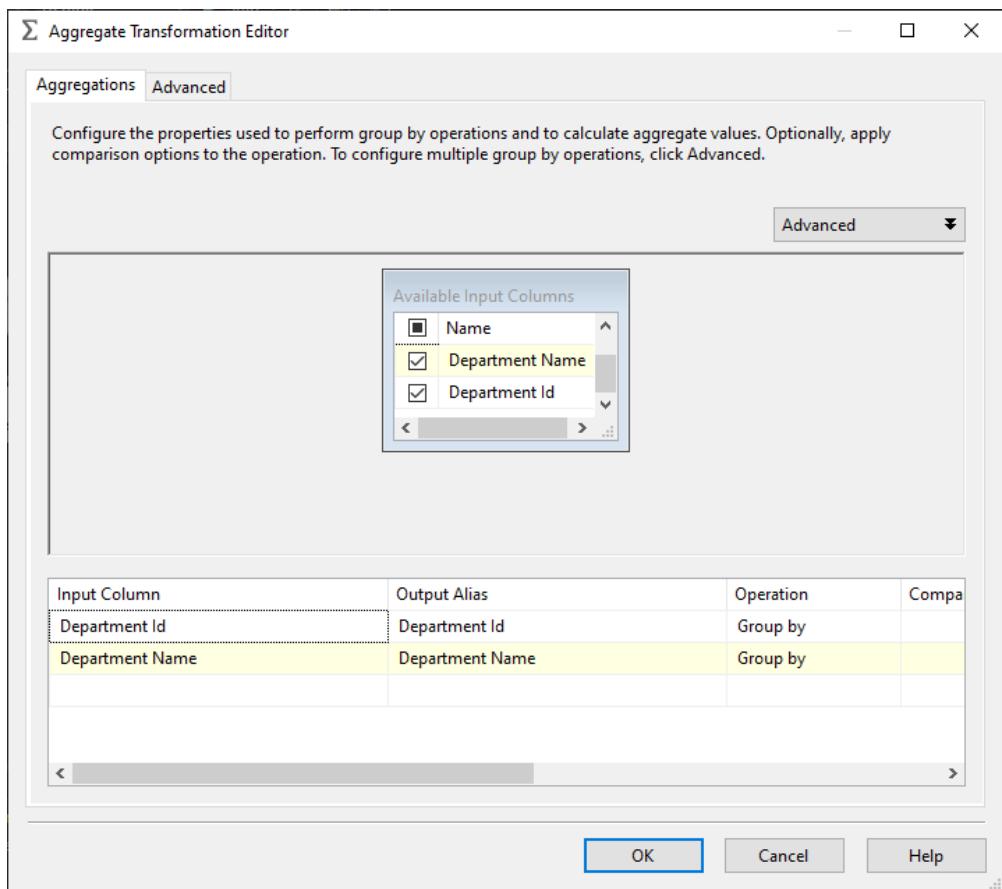
- Các thành phần của Stage Department



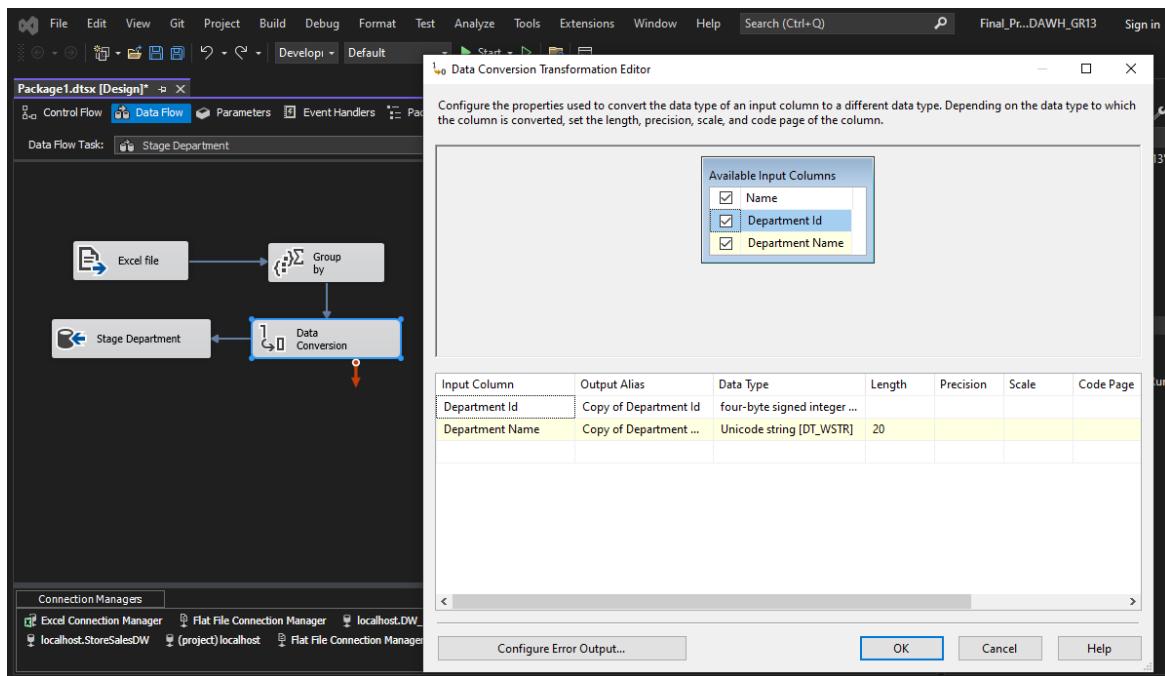
- Từ file excel nguồn chọn ra cột thích hợp. Ở Stage Department chọn ra Department Id và Department Name



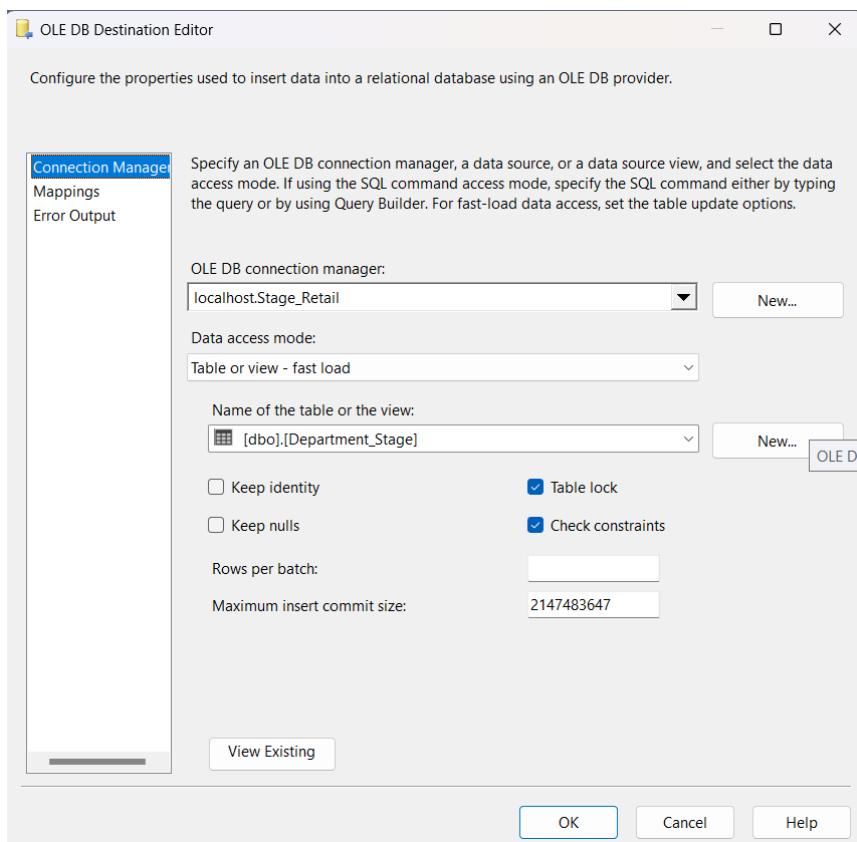
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



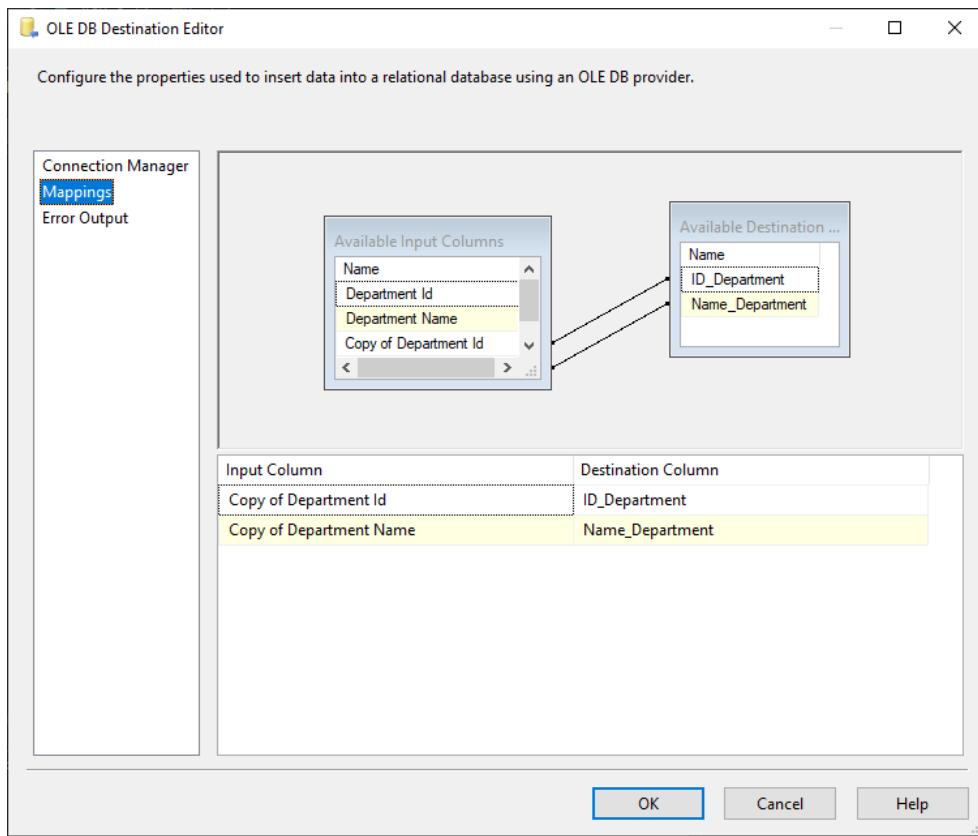
- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào Stage Department.



- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp

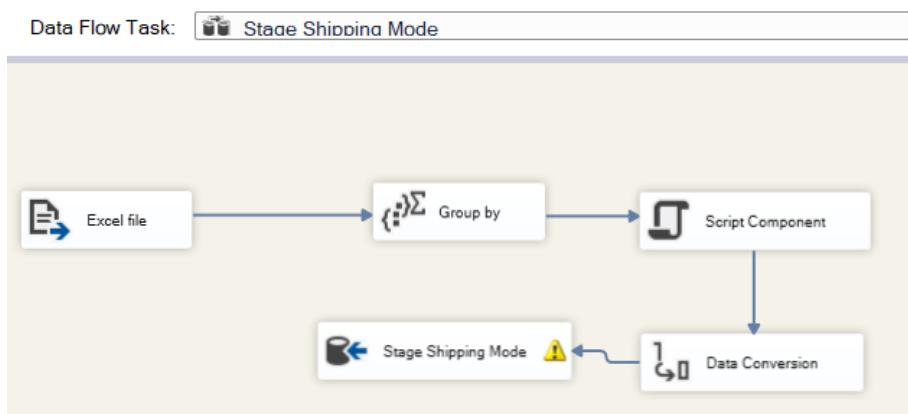


- Mapping các thuộc tính với các cột của bảng vừa chọn

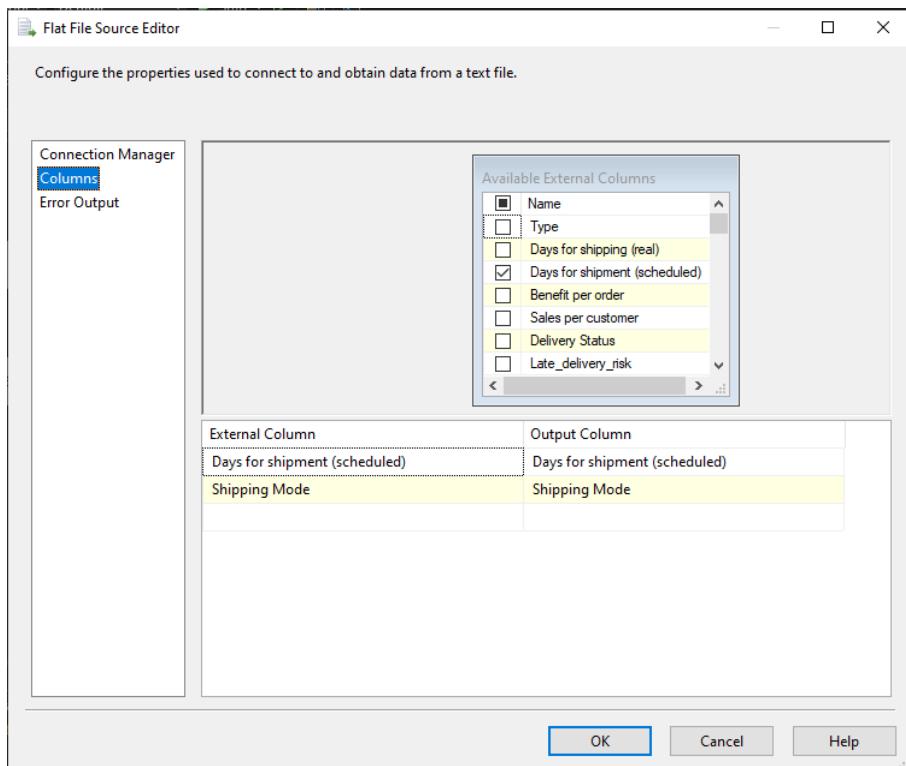


### 3.2.2. Stage Shipping Mode

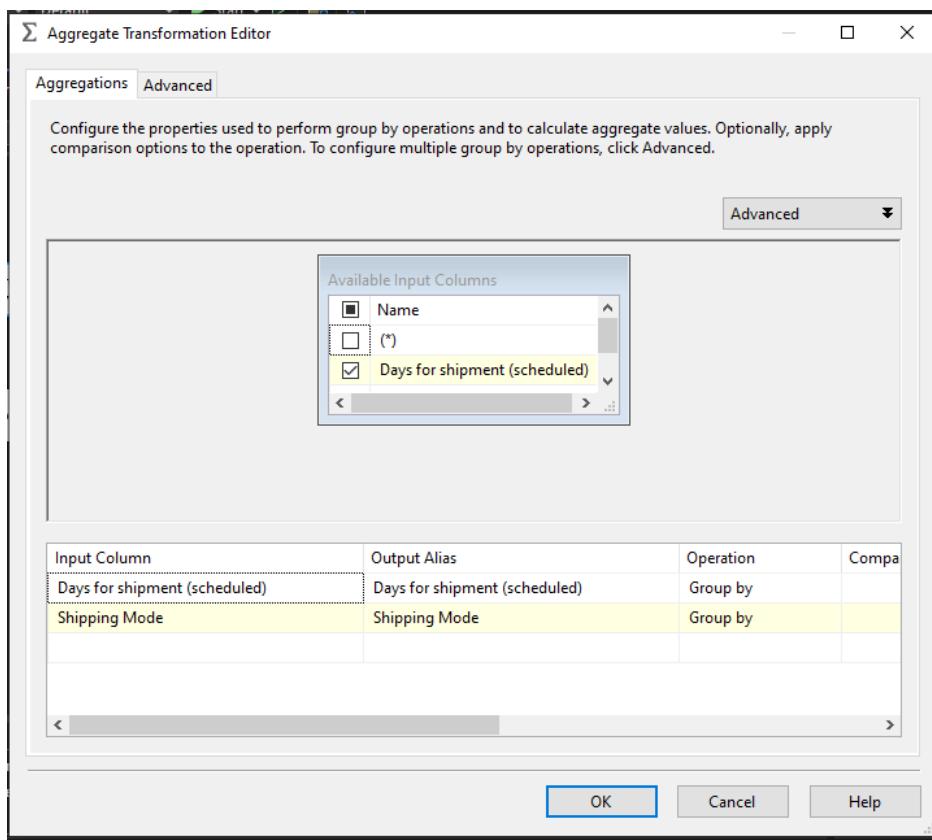
- Cách thành phần của Stage **Shipping Mode**



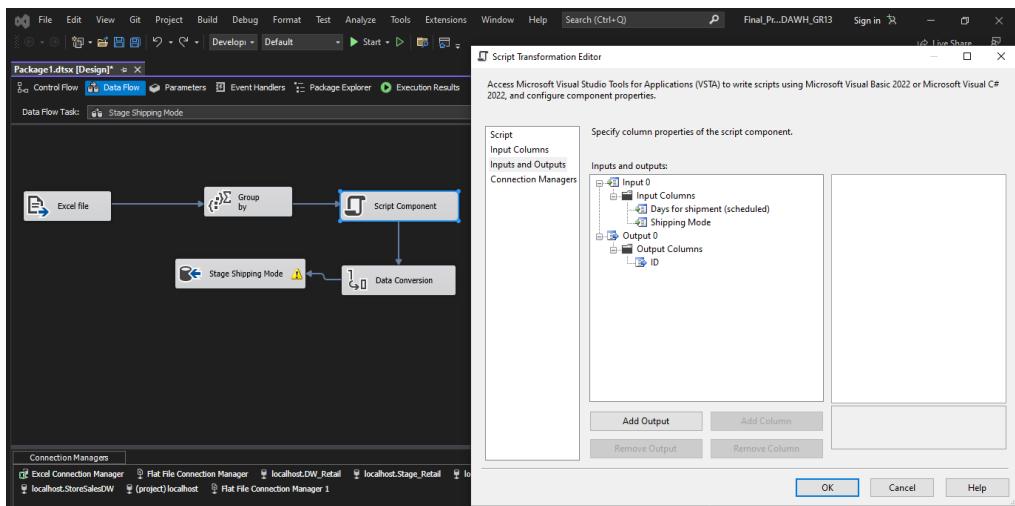
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



- Dùng Script Transformation để tạo thêm ID tự động cho bảng



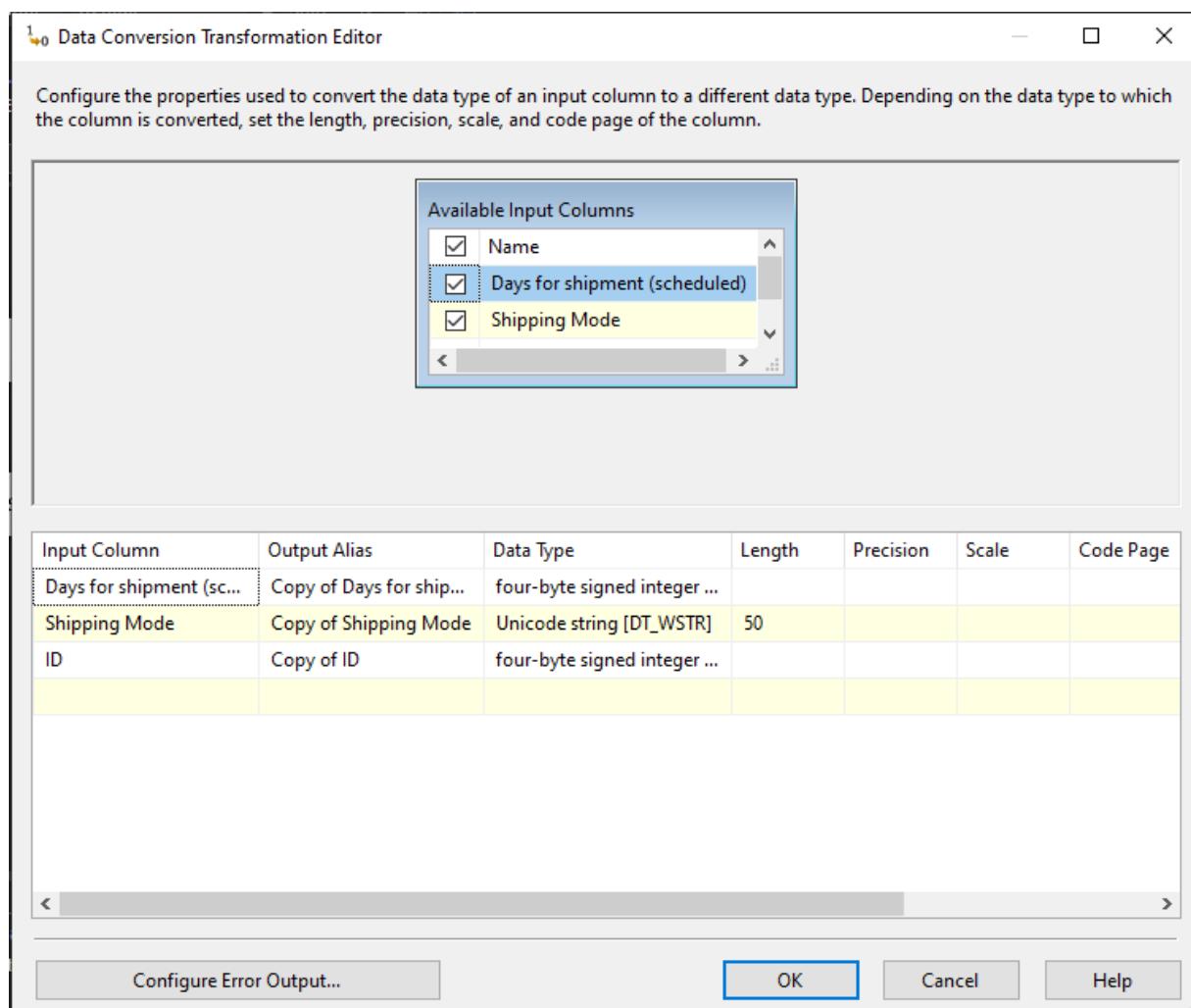
- **Code trong Script**

```

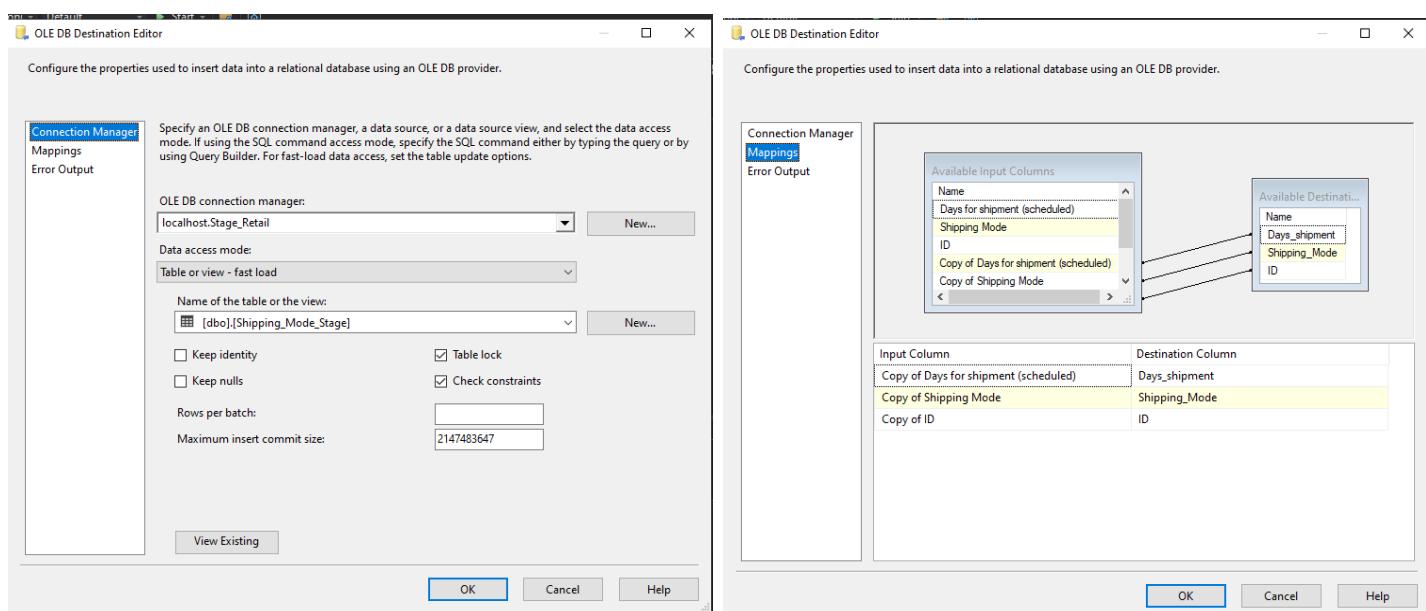
104     int count = 0;
105     /// <summary>
106     /// This method is called once for every row that passes through the component.
107     ///
108     /// Example of reading a value from a column in the the row:
109     /// string zipCode = Row.ZipCode
110     ///
111     /// Example of writing a value to a column in the row:
112     /// Row.ZipCode = zipCode
113     /// </summary>
114     /// <param name="Row">The row that is currently passing through the component</param>
115     public override void Input0_ProcessInputRow(Input0Buffer Row)
116     {
117         /*
118         * Add your code here
119         */
120         count++;
121         Row.ID = count;
122     }

```

- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào **Stage Shipping Mode**

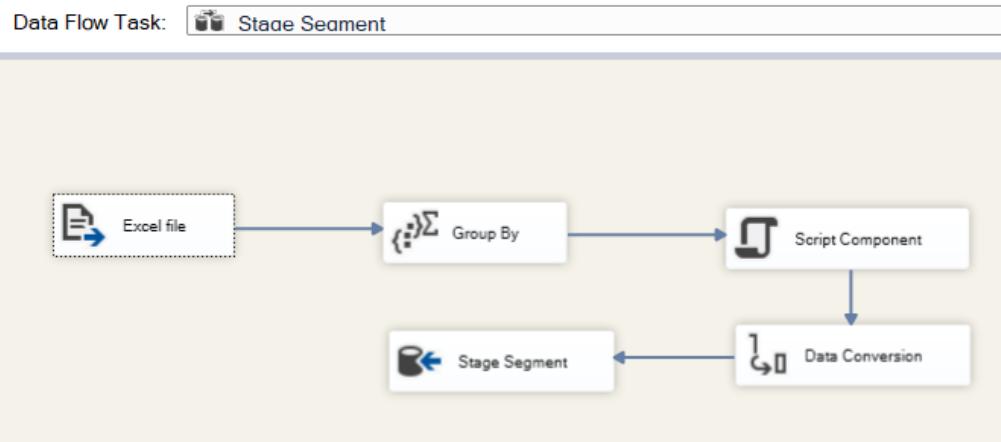


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn

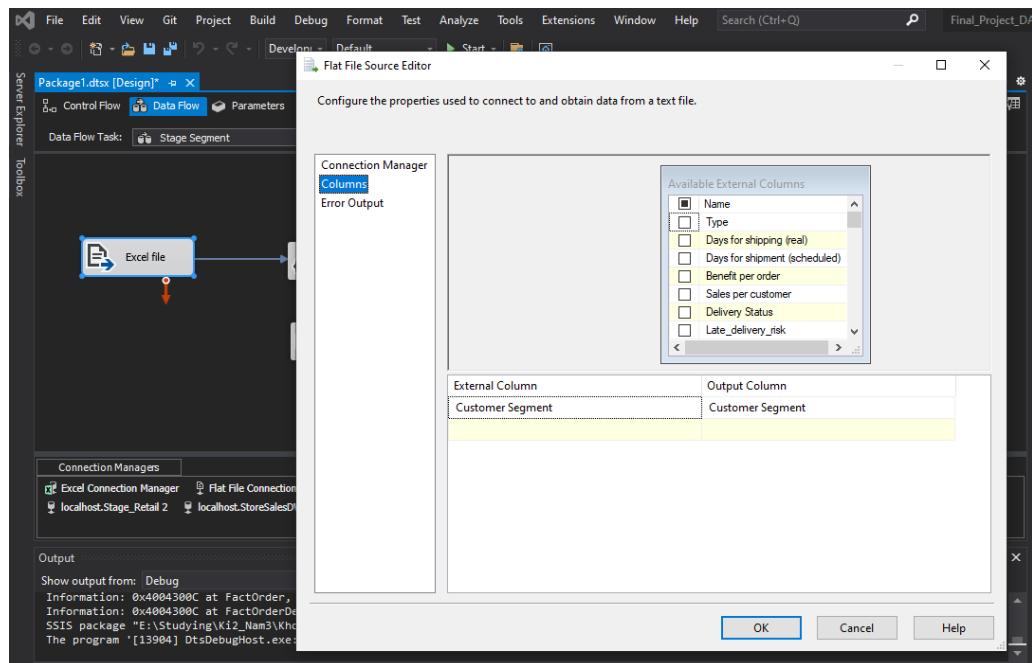


### 3.2.3. Stage Segment

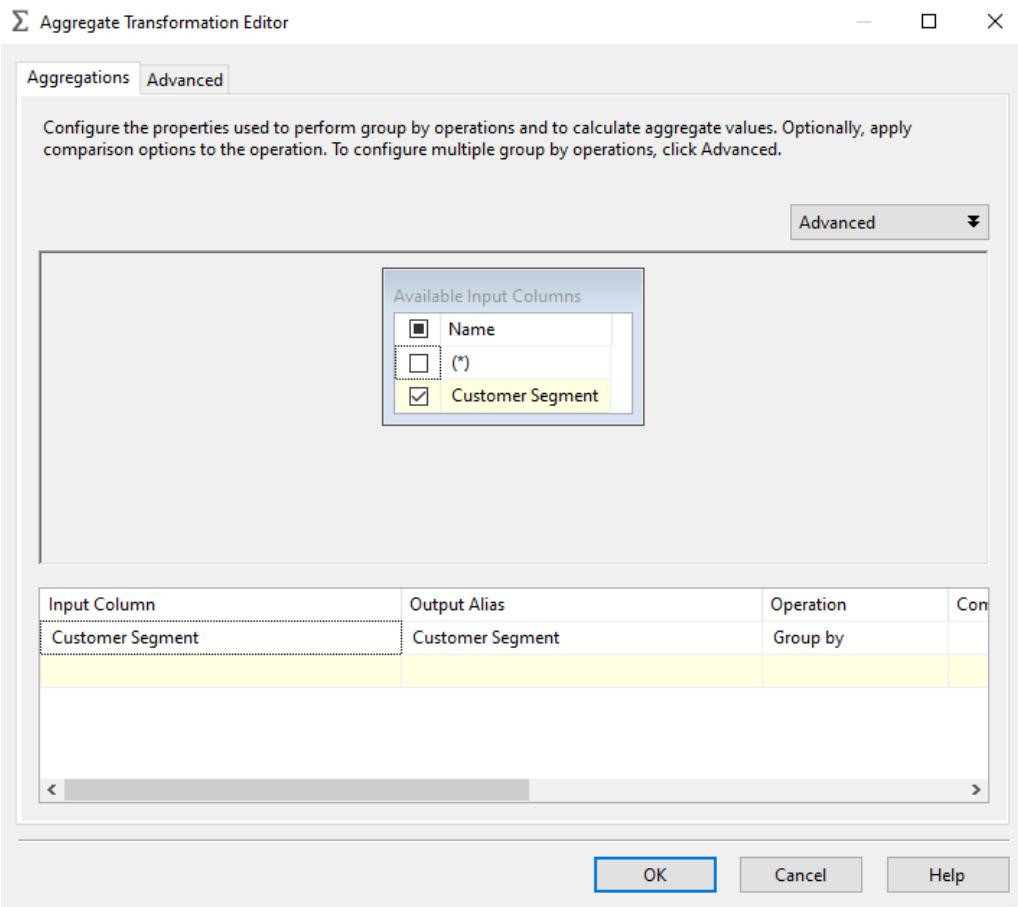
- Cách thành phần của Stage Segment



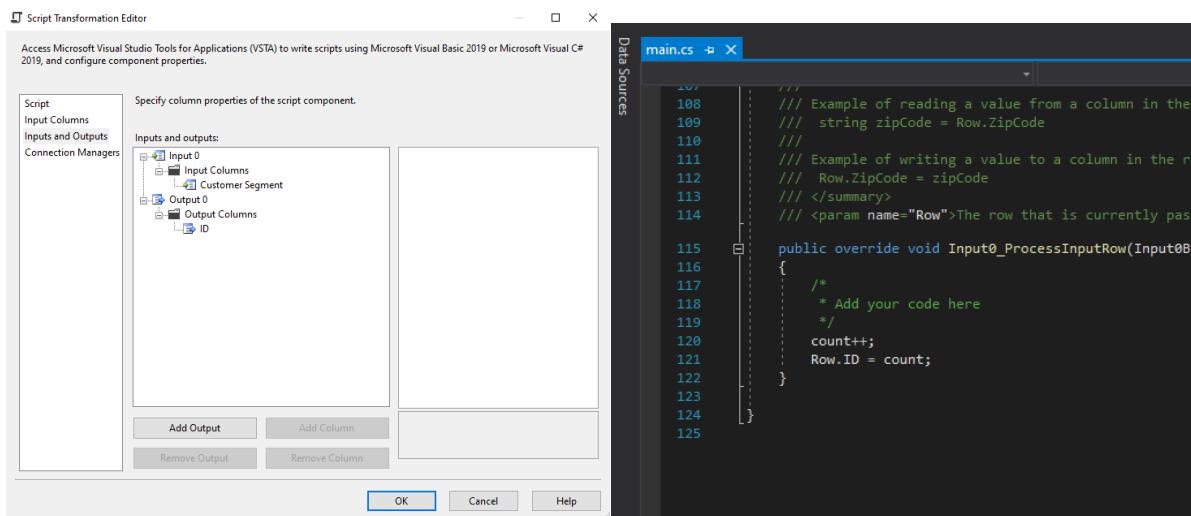
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



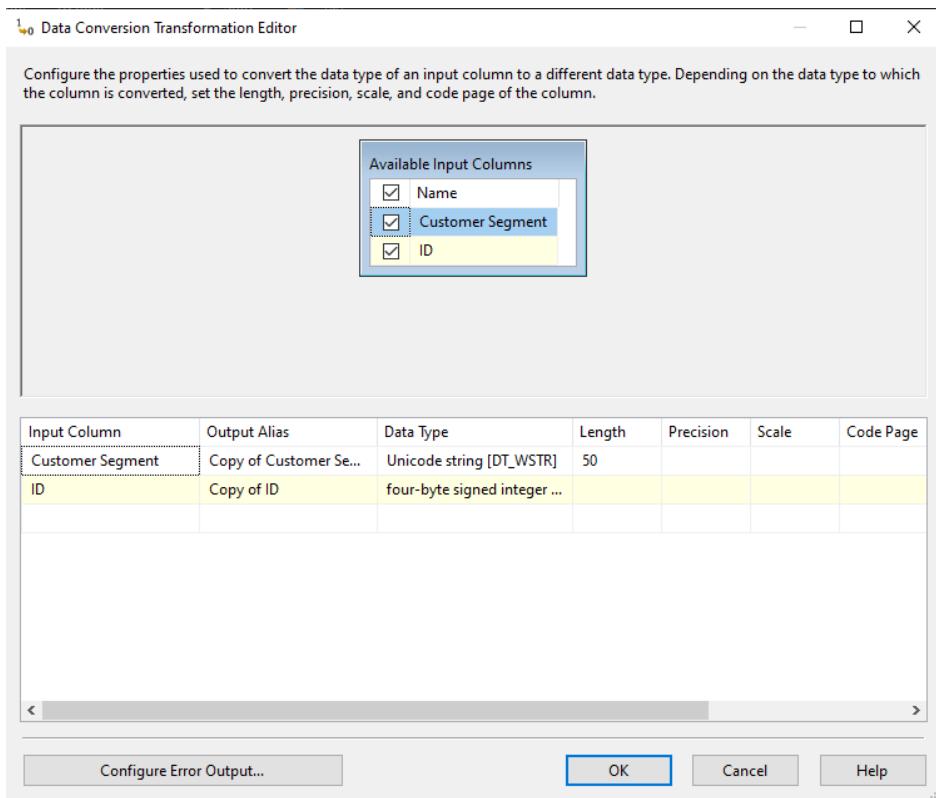
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



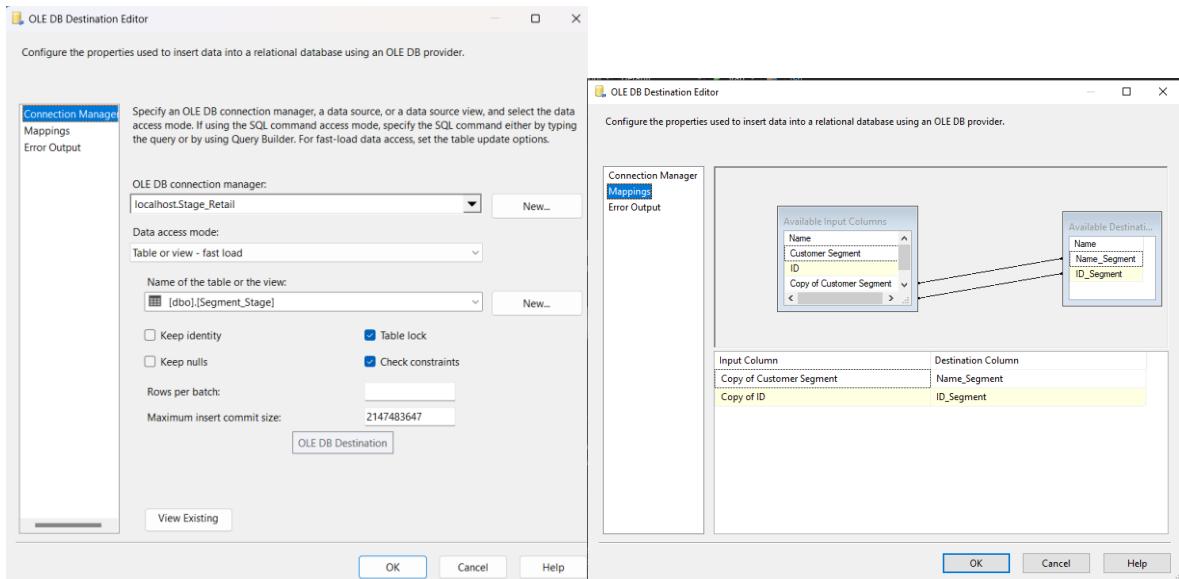
- Dùng Script Transformation để tạo thêm ID tự động cho bảng. Thay đổi code trong Script



- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào Stage Segment

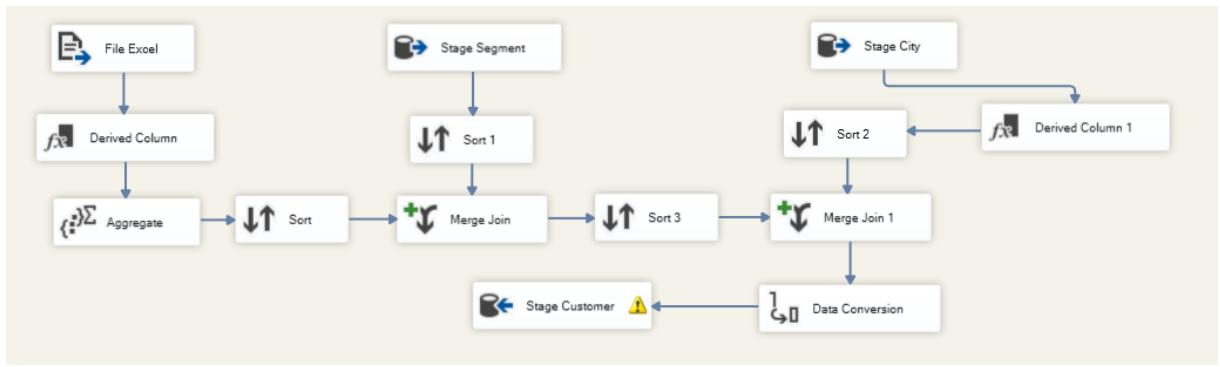


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn

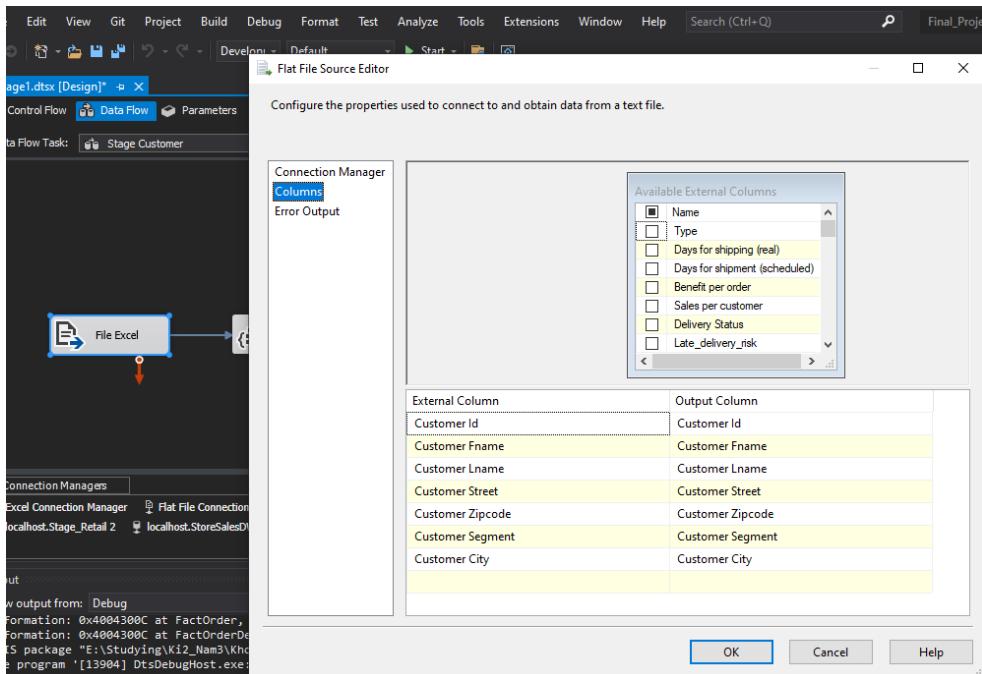


### 3.2.4. Stage Customer

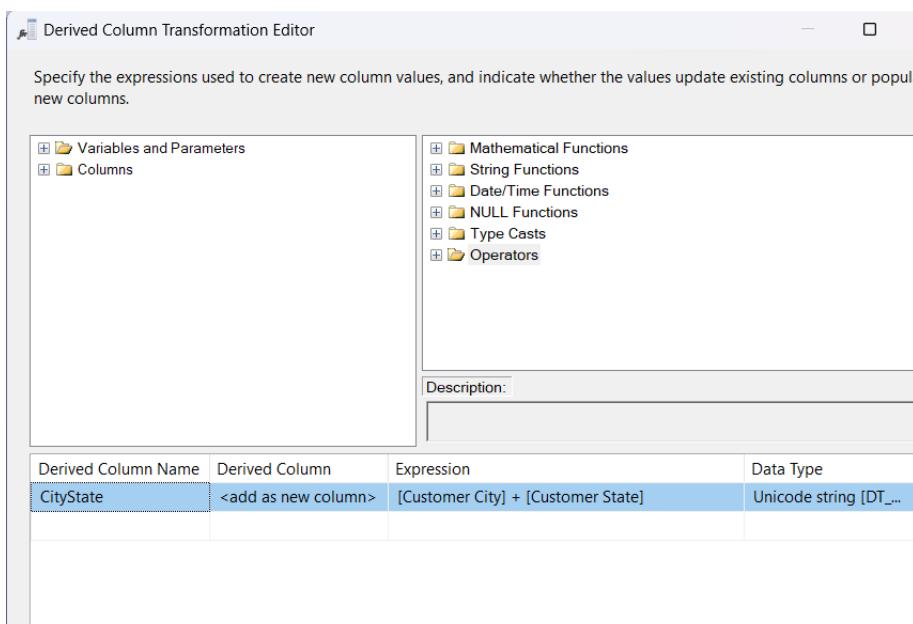
- Cách thành phần của Stage Customer



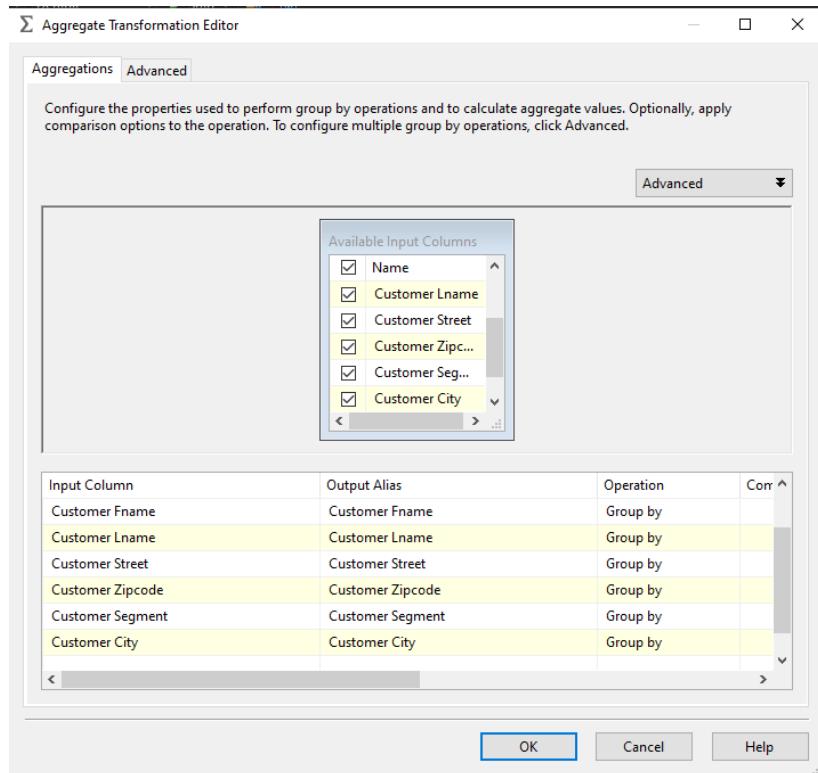
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



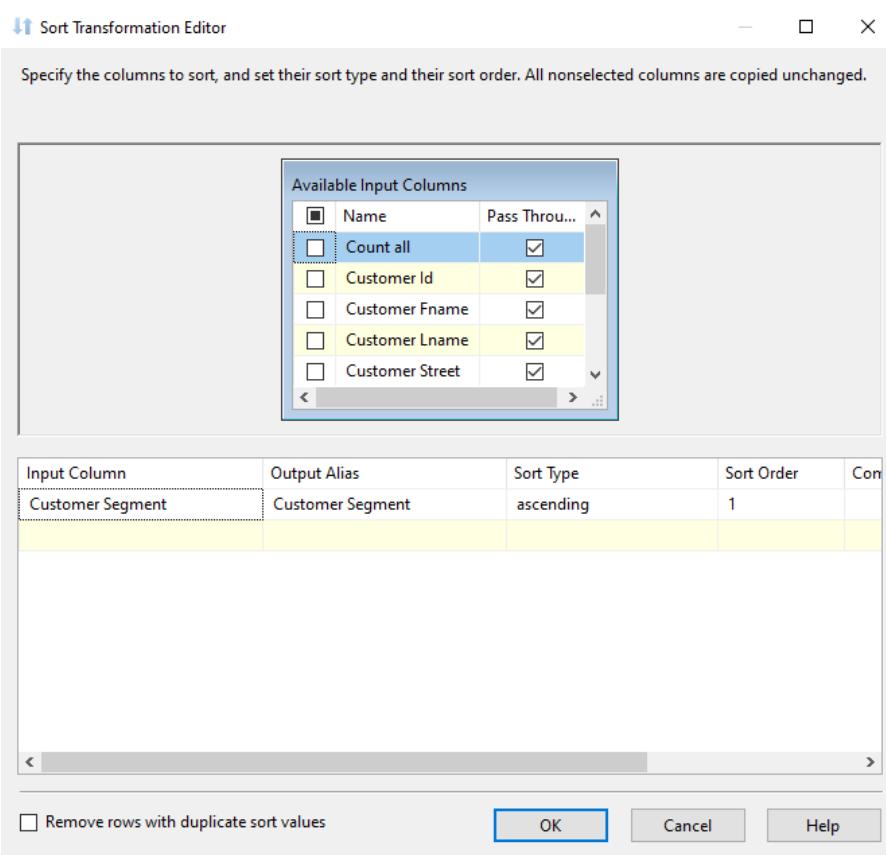
- Sử dụng Derived Column để thêm cột CityStage dùng để inner join



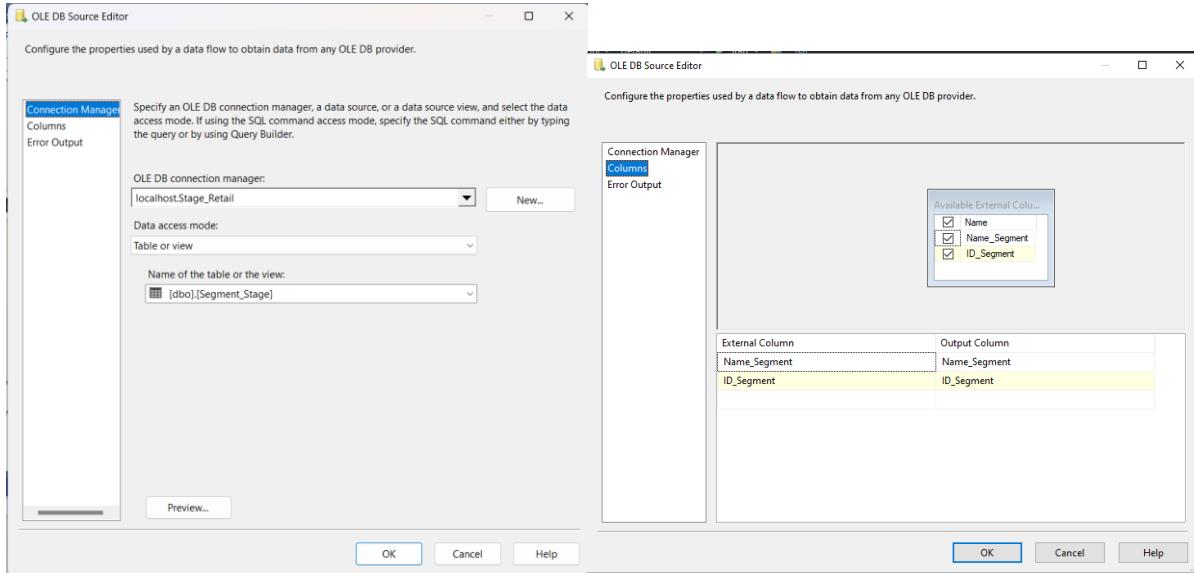
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



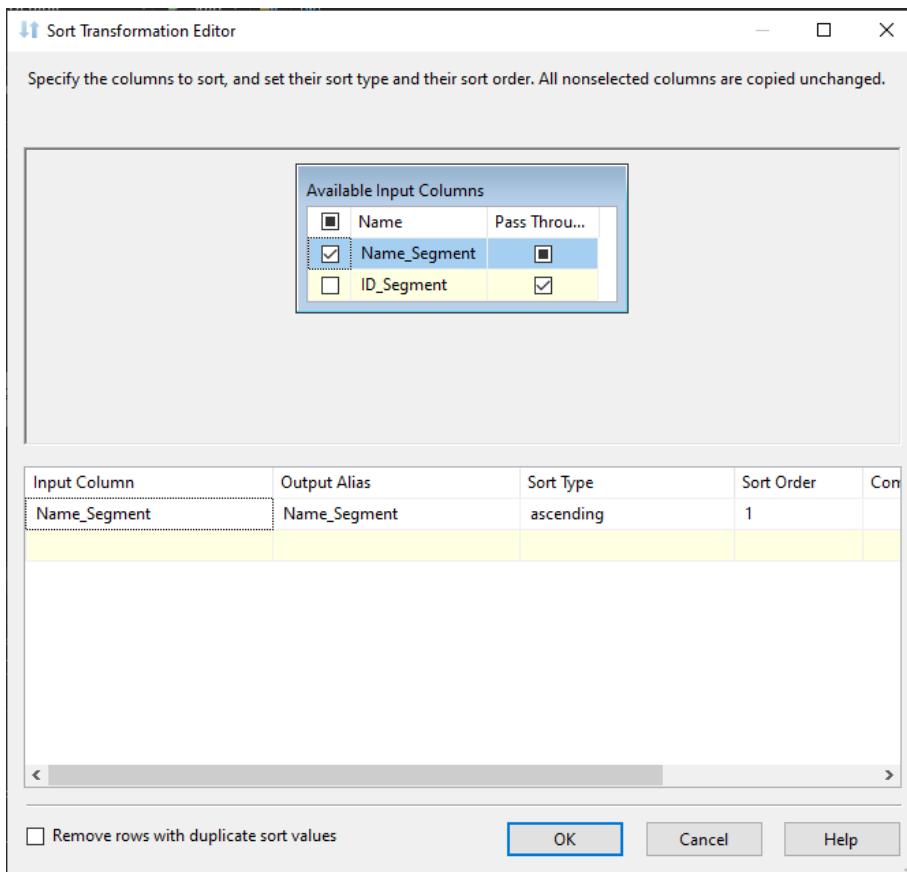
- Dùng sort để chỉ định thuộc tính dùng để inner join



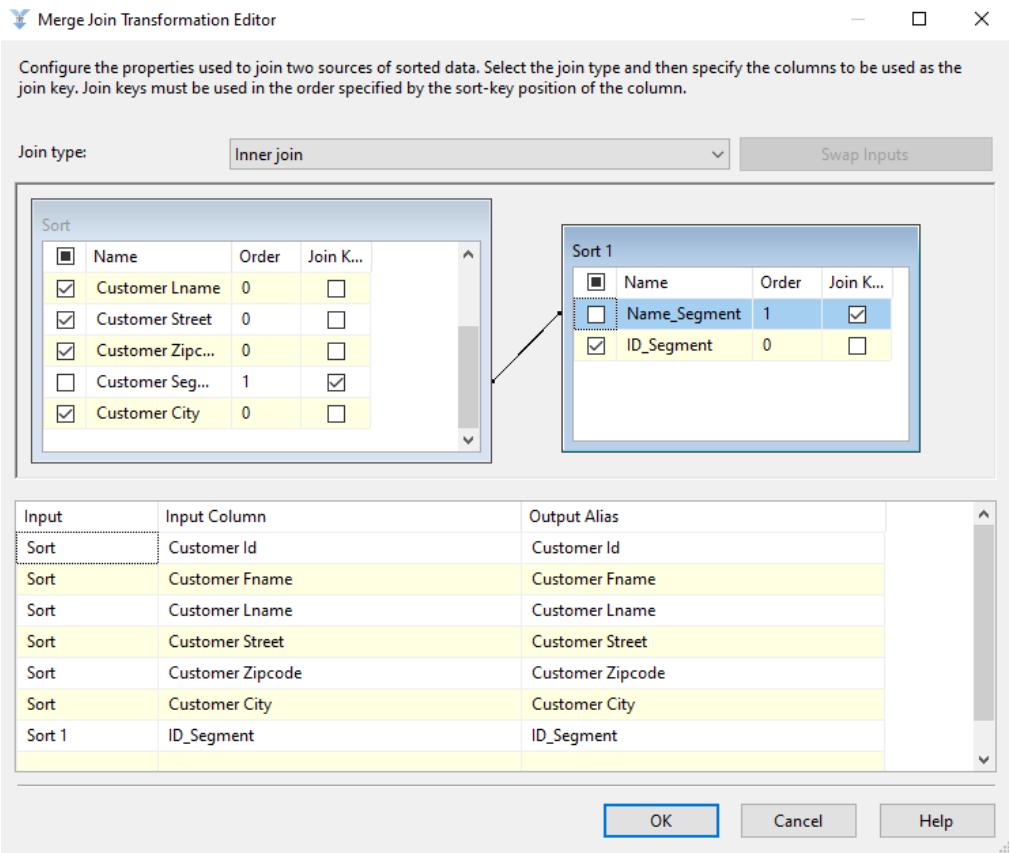
- Chọn nguồn là Stage segment để inner join



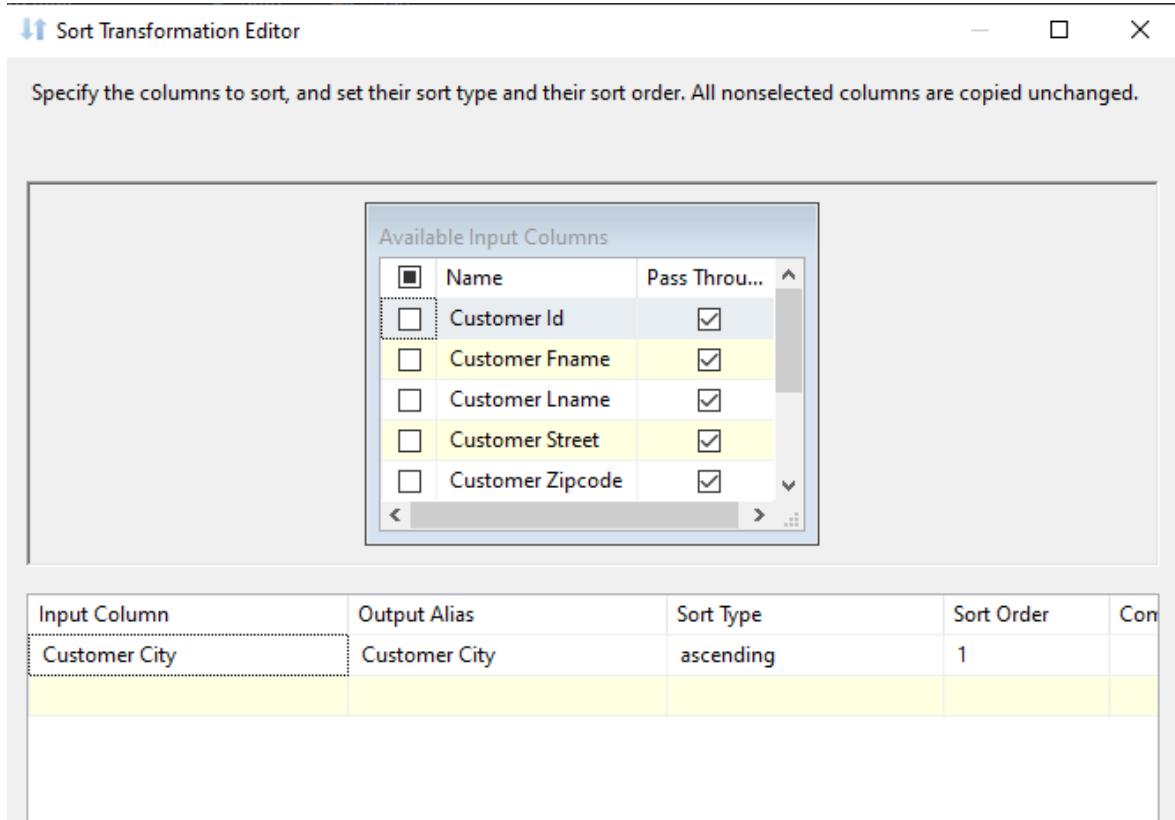
- Dùng sort để chỉ định thuộc tính của Stage segment dùng để inner join



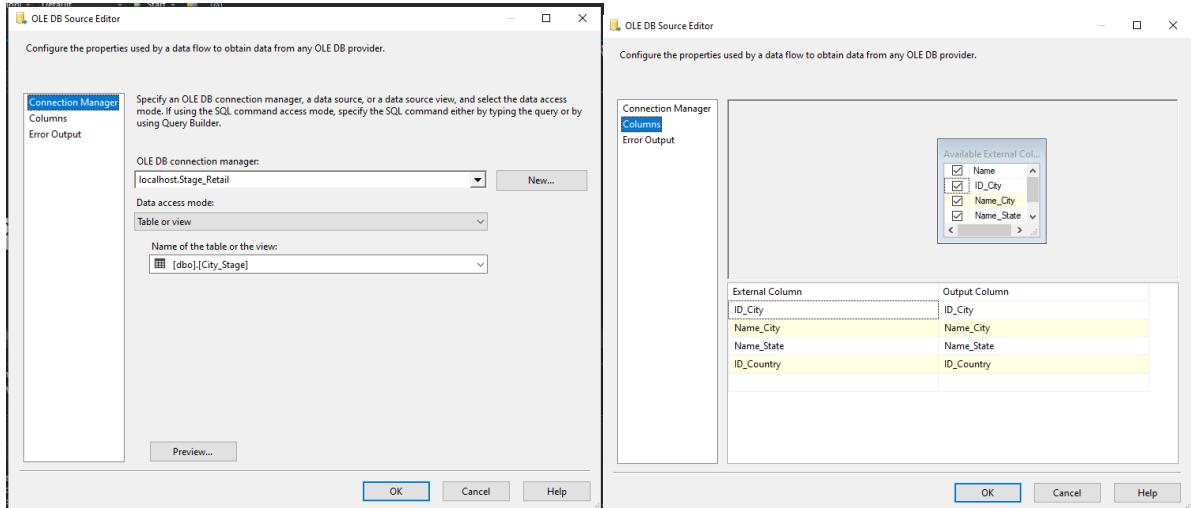
- Sử dụng Merge Join để kết hợp hai bảng lại với nhau để lấy được ID\_Segment và chọn các thuộc tính cần thiết khác



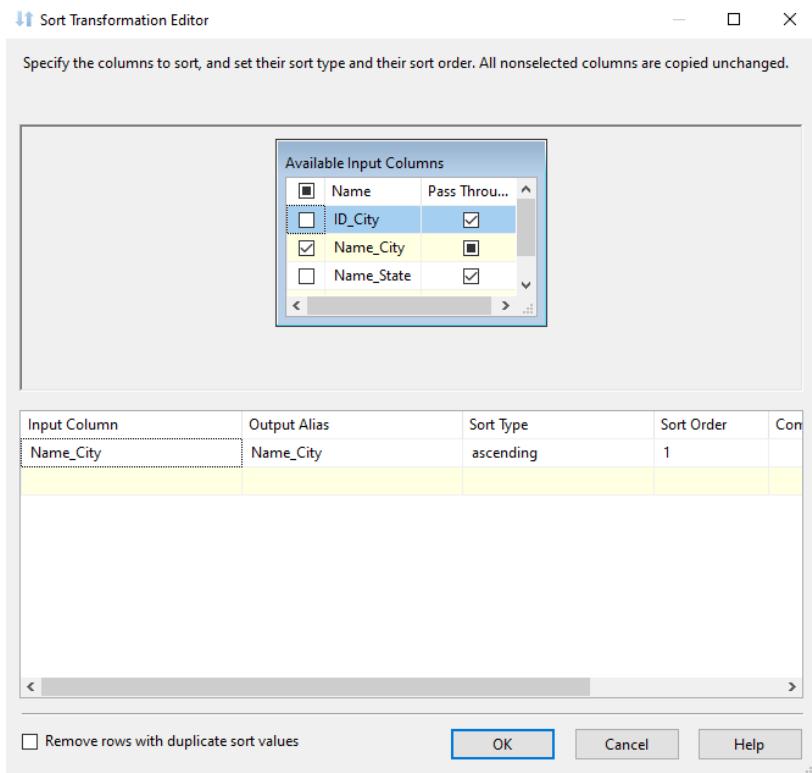
- Tiếp tục dùng Sort để chỉ định thuộc tính của kết quả vừa merge dùng để inner join.



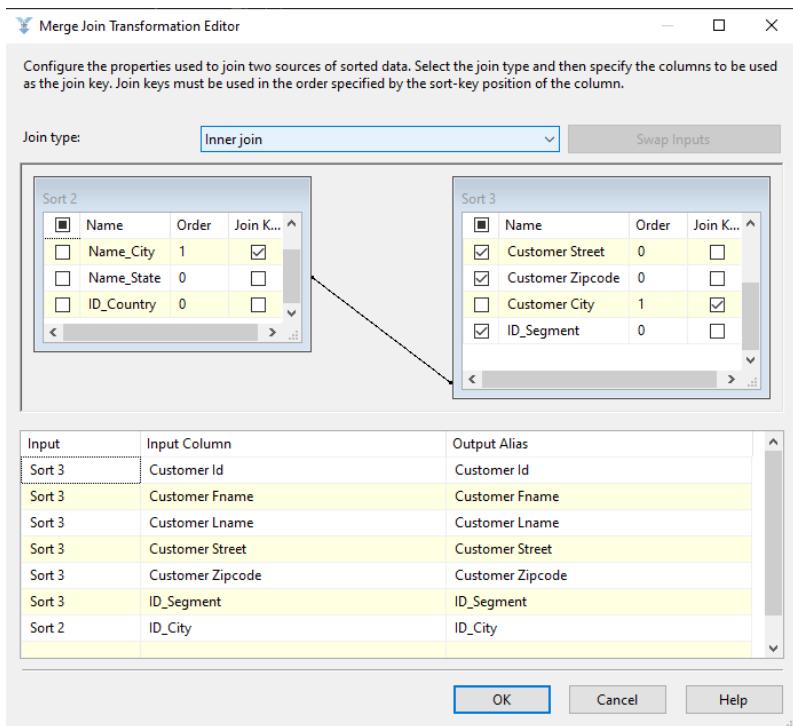
- Chọn nguồn là Stage City để tiếp tục inner join.



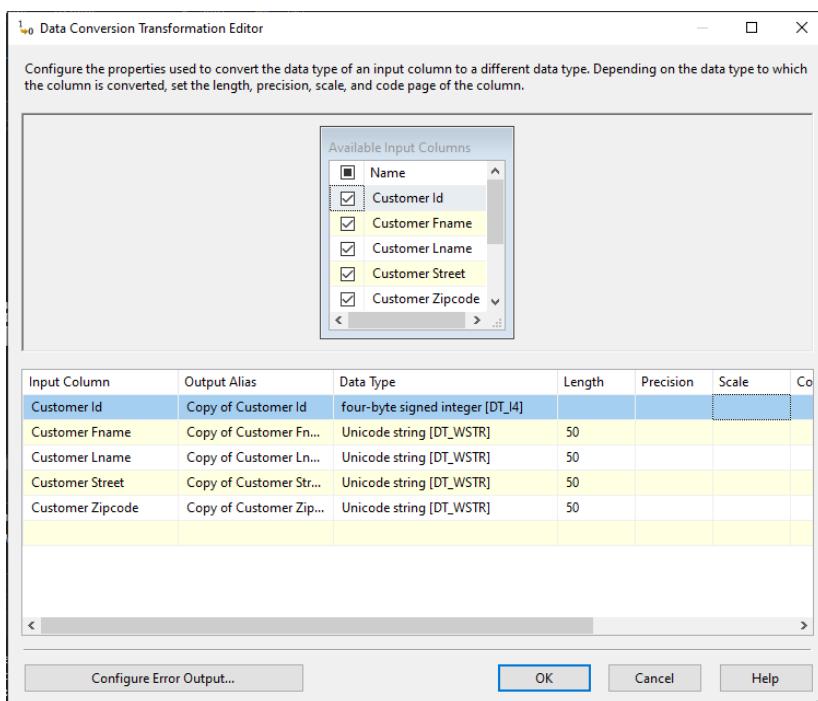
- Dùng sort để chỉ định thuộc tính của Stage City dùng để inner join



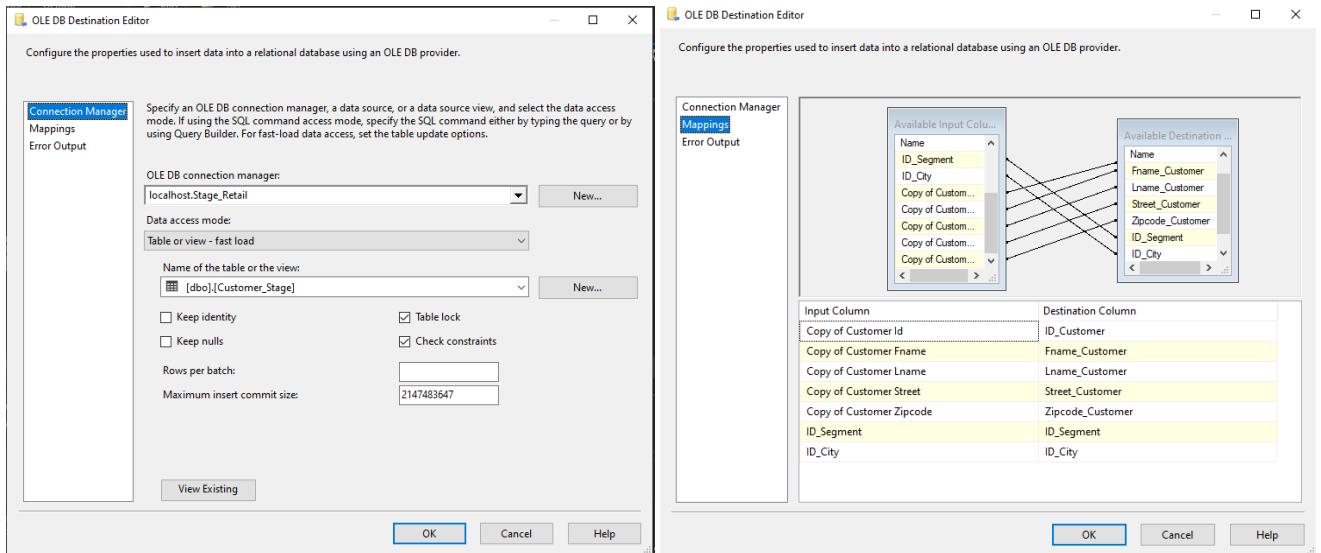
- Sử dụng Merge Join để kết hợp hai bảng lại với nhau để lấy được ID\_City và chọn các thuộc tính cần thiết khác



- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào Stage Customer

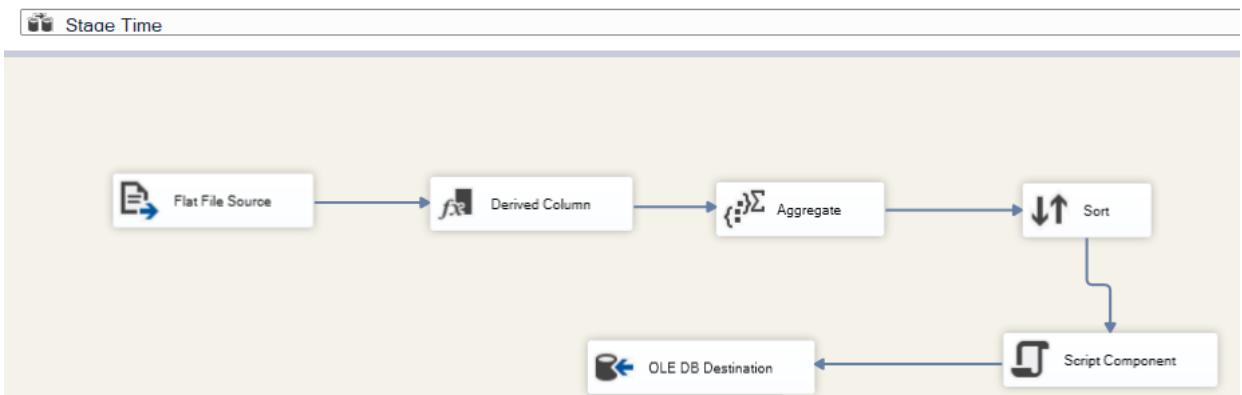


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn

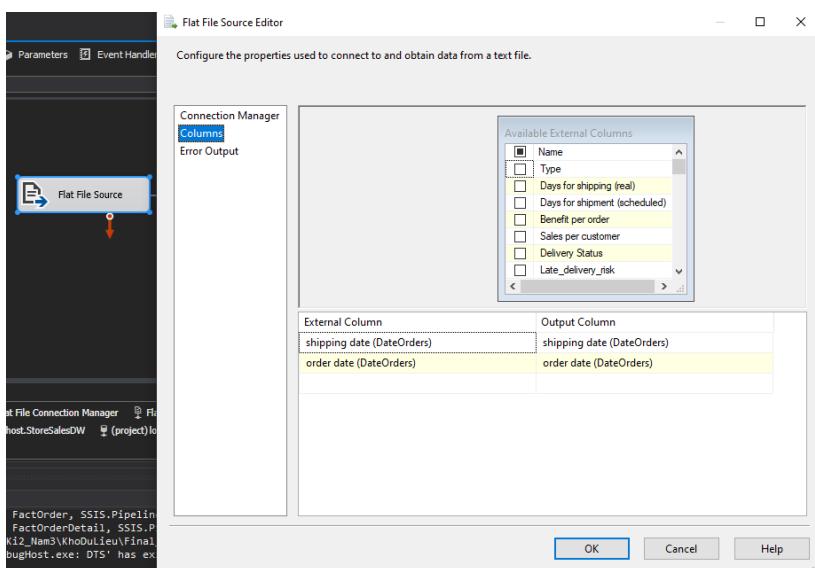


### 3.2.5. Stage Time

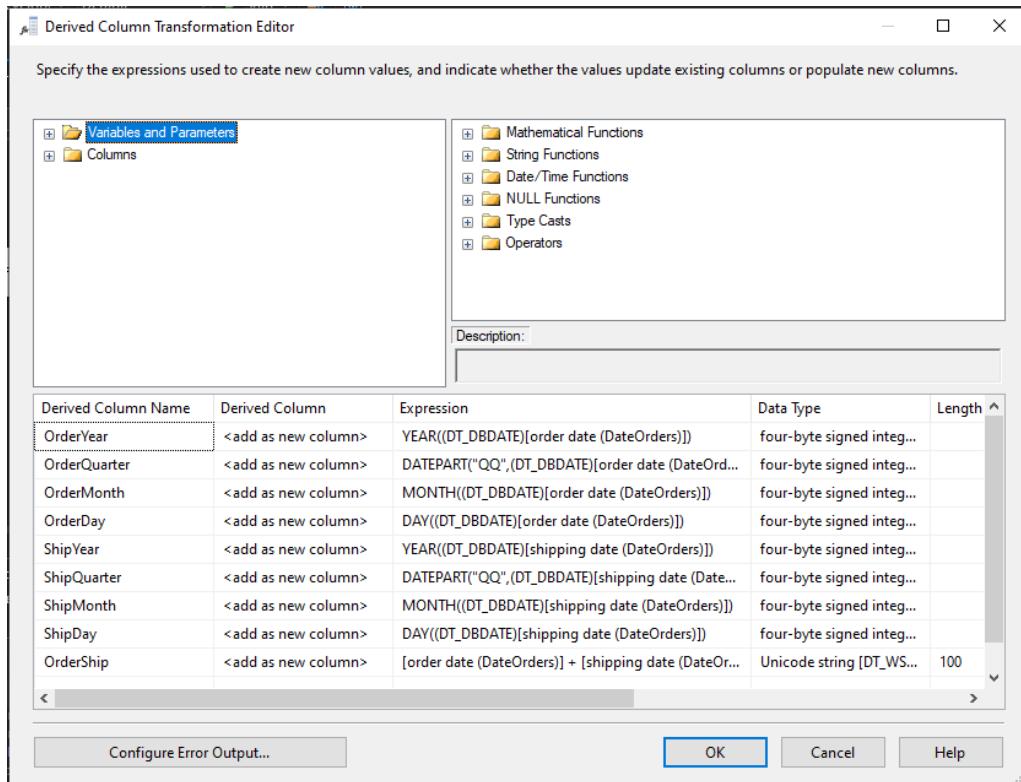
- Cách thành phần của Stage Time



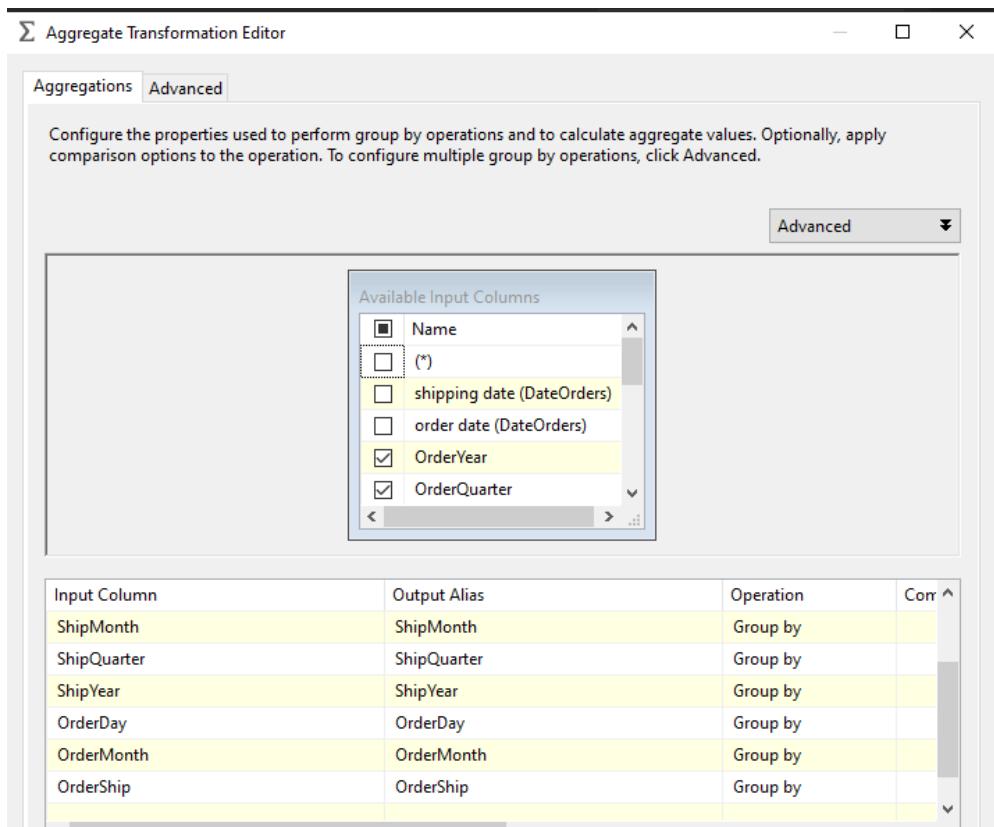
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



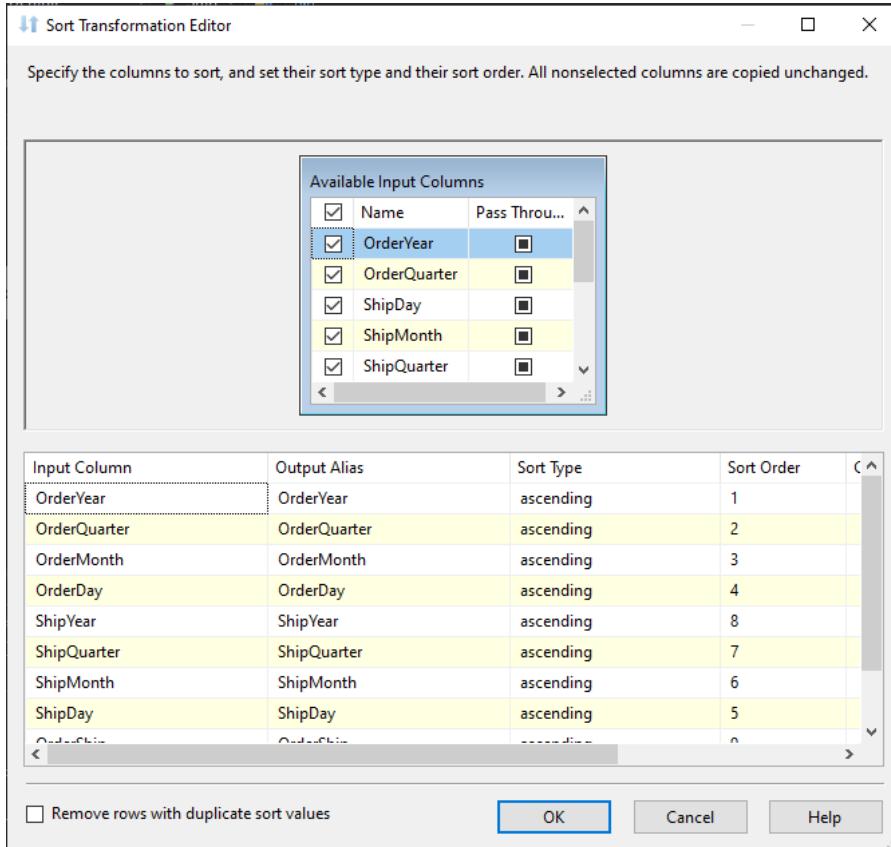
- Sử dụng Derived Column để thêm các thuộc tính mới và dựa trên thuộc tính đã có. Tạo thuộc tính ngày, tháng, năm, quý dựa trên order date và shipping date.



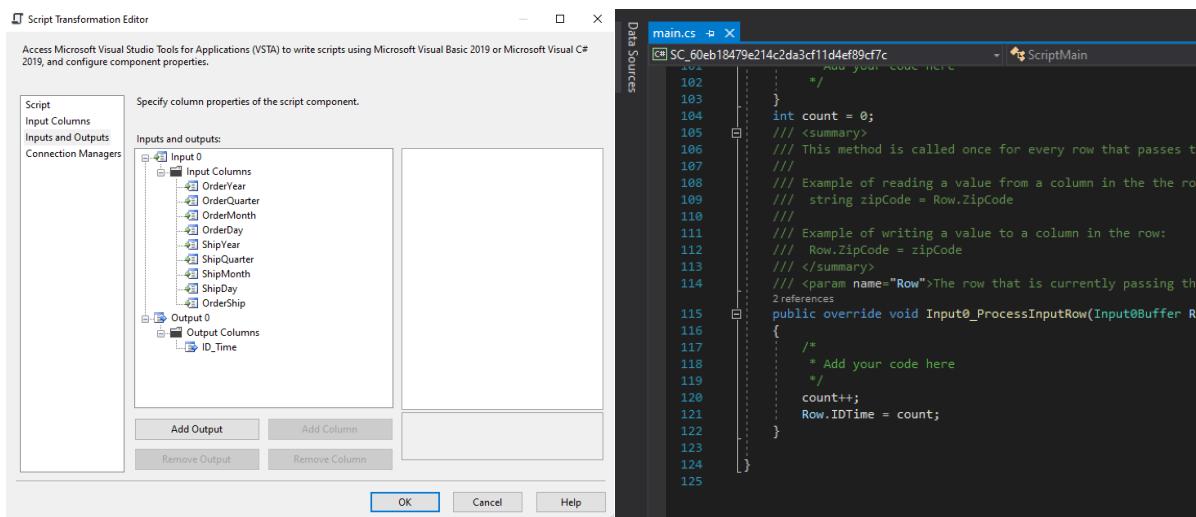
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



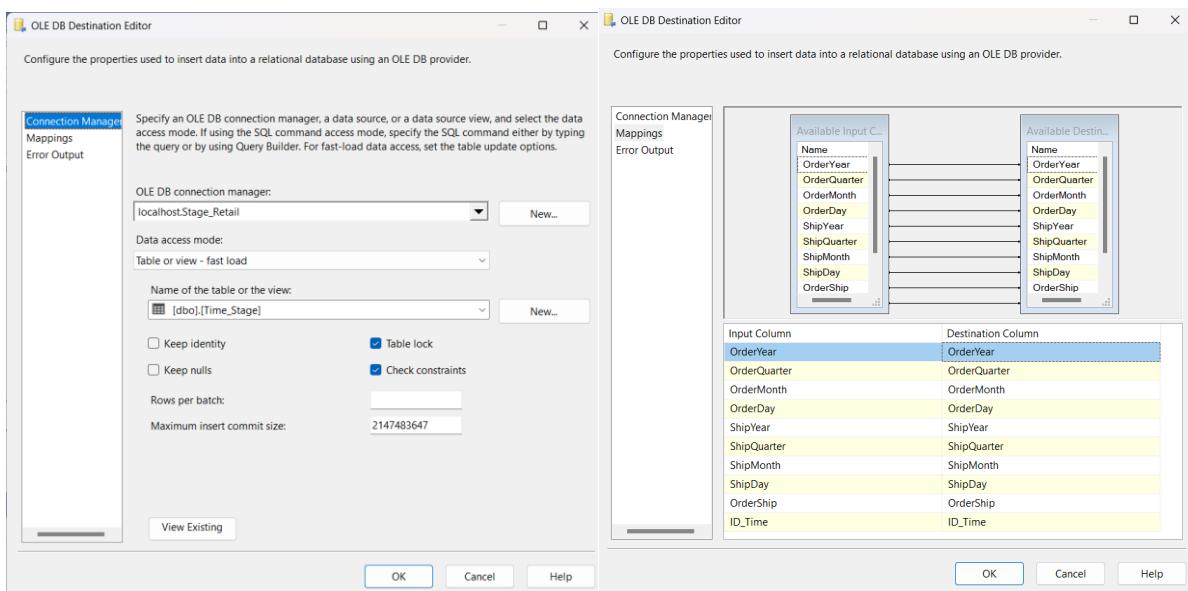
- Dùng sort sắp xếp lại các thuộc tính.



- Dùng Script Transformation để tạo thêm ID tự động cho bảng. Thay đổi code trong Script

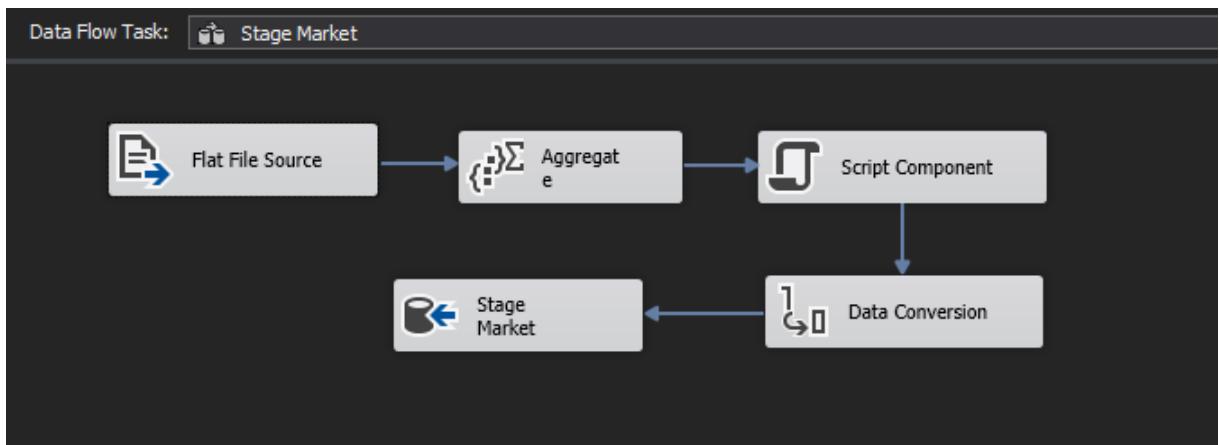


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn

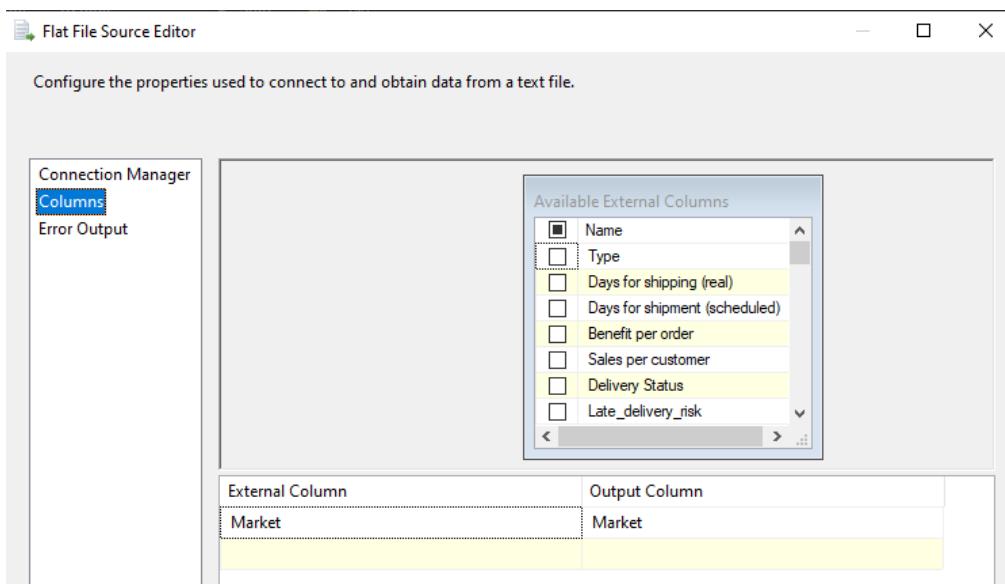


### 3.2.6. Stage Market

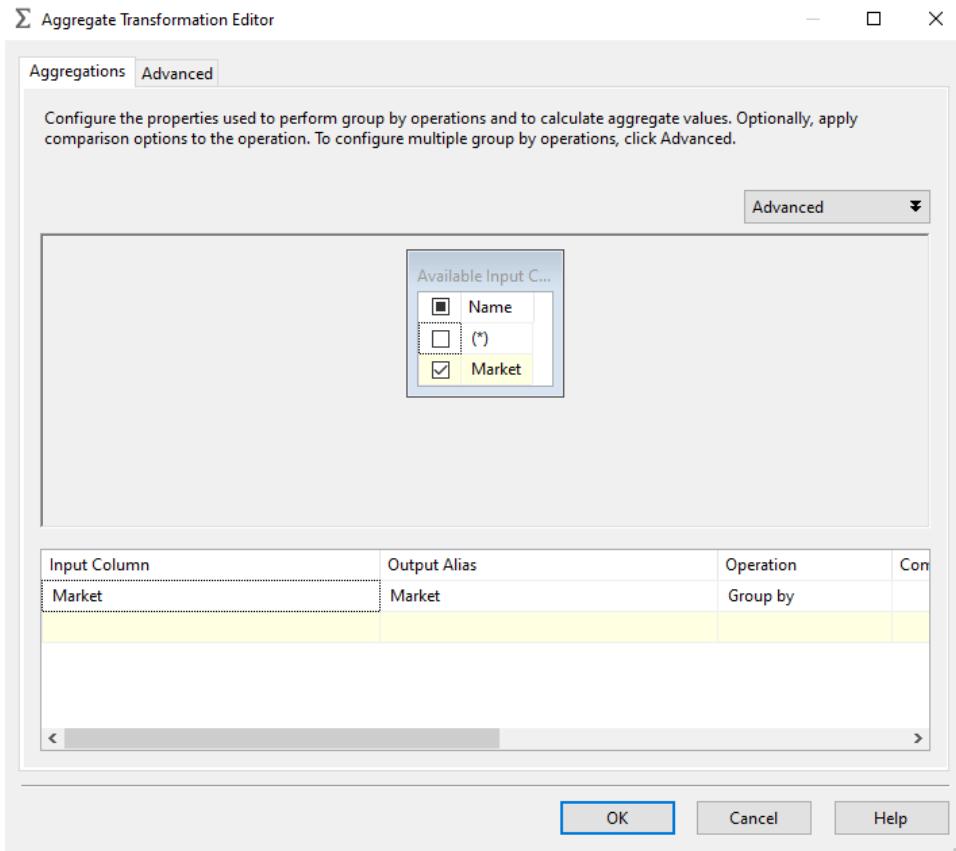
- Cách thành phần của Stage Market



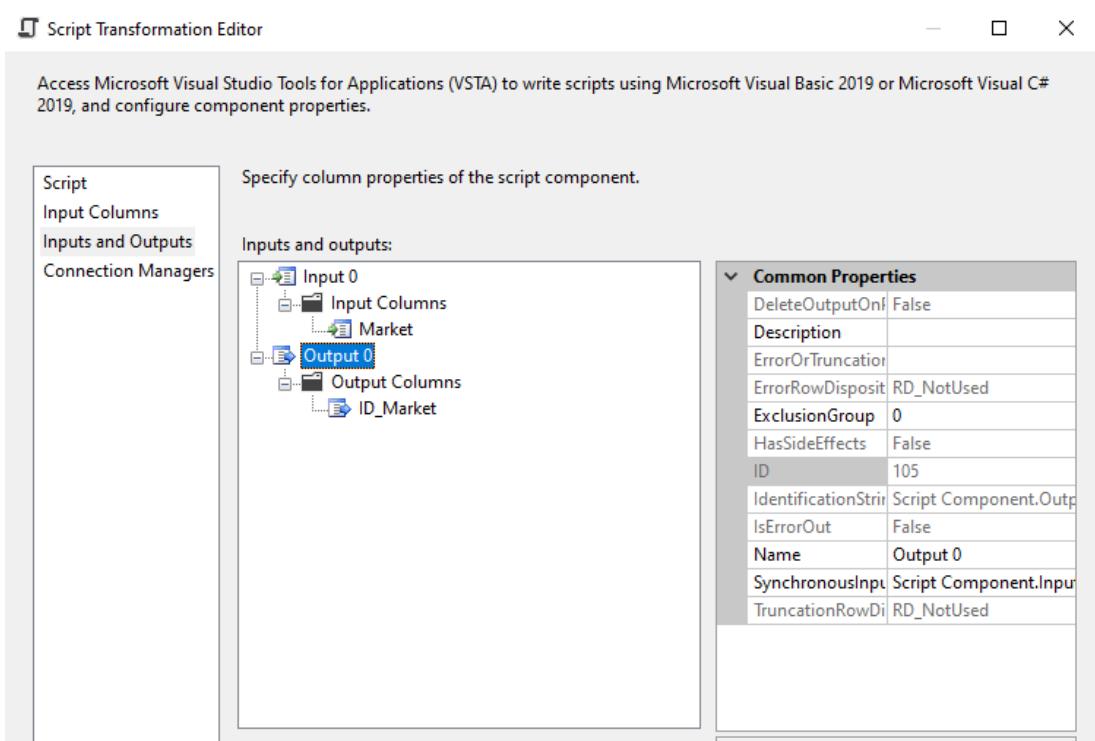
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



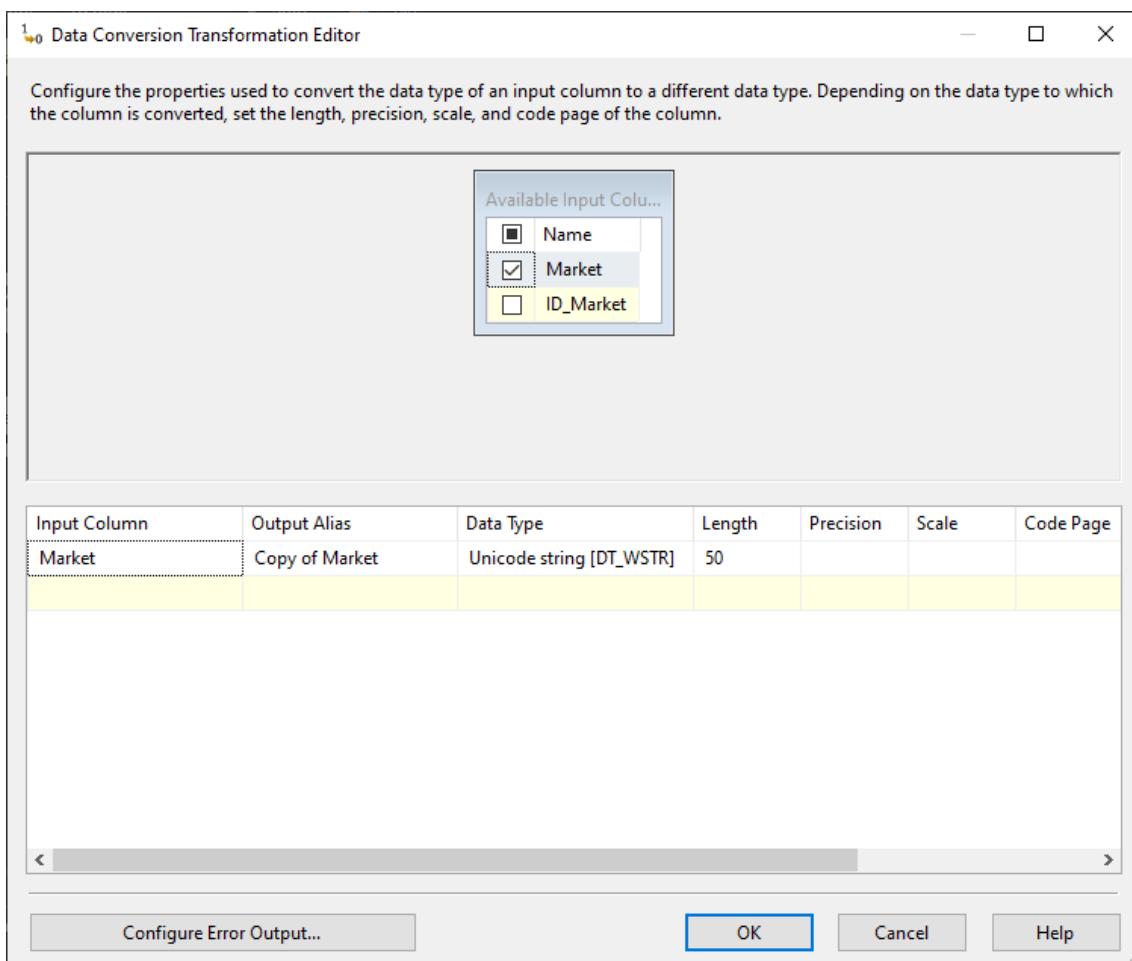
- Dùng Script Transformation để tạo thêm ID tự động cho bảng. Thay đổi code trong Script



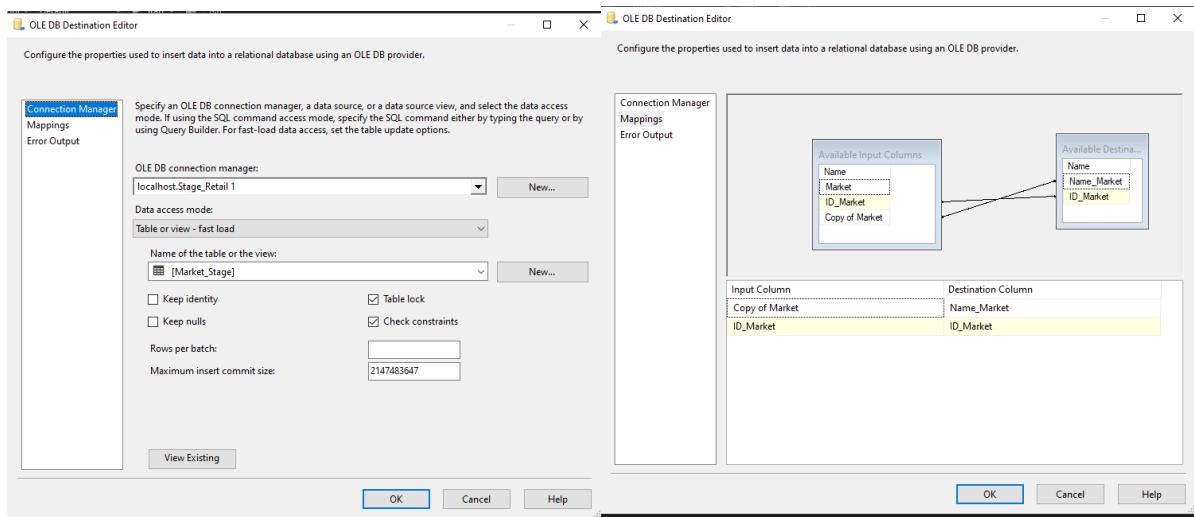
```
main.cs + X
```

```
101     /* Add your code here
102 }
103 */
104 int count = 0;
105 /// <summary>
106 /// This method is called once for every row that passes through the component from Input0.
107 ///
108 /// Example of reading a value from a column in the the row:
109 /// string zipCode = Row.ZipCode
110 ///
111 /// Example of writing a value to a column in the row:
112 /// Row.ZipCode = zipCode
113 /// </summary>
114 /// <param name="Row">The row that is currently passing through the component</param>
115 public override void Input0_ProcessInputRow(Input0Buffer Row)
116 {
117     /*
118     * Add your code here
119     */
120     count++;
121     Row.IDMarket = count;
122 }
123
124 }
```

- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào Stage Market

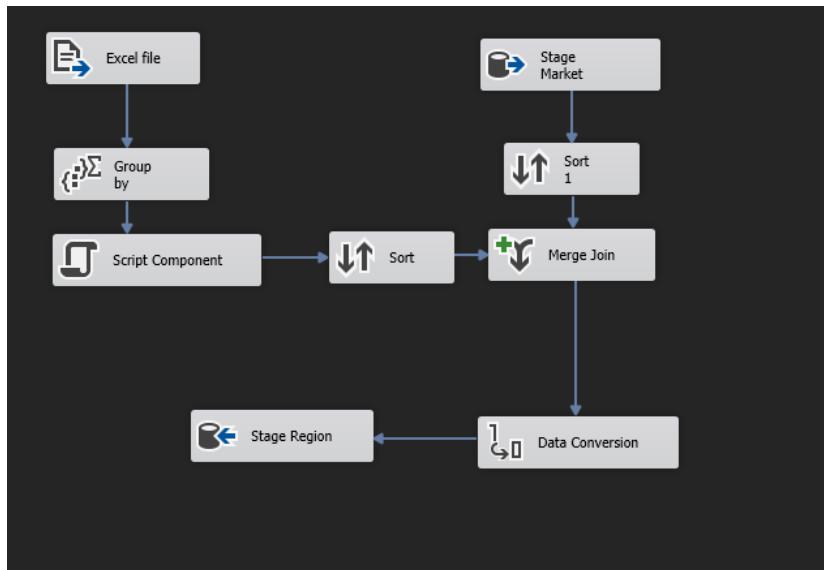


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn

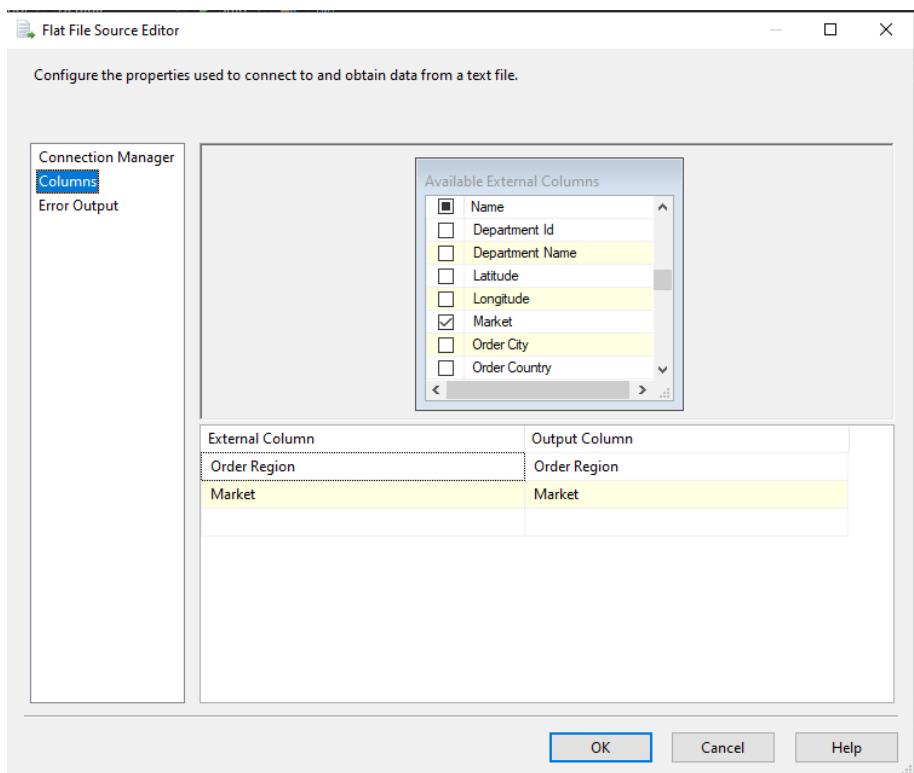


### 3.2.7. Stage Region

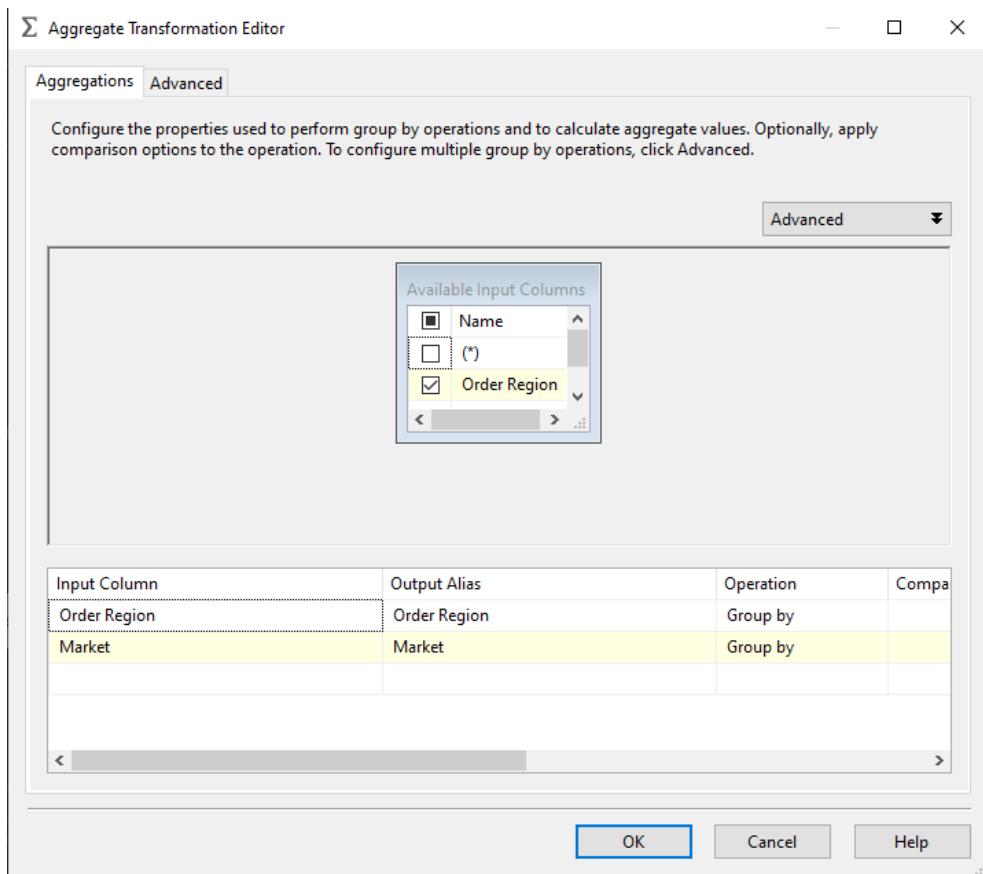
- Cách thành phần của Stage Region



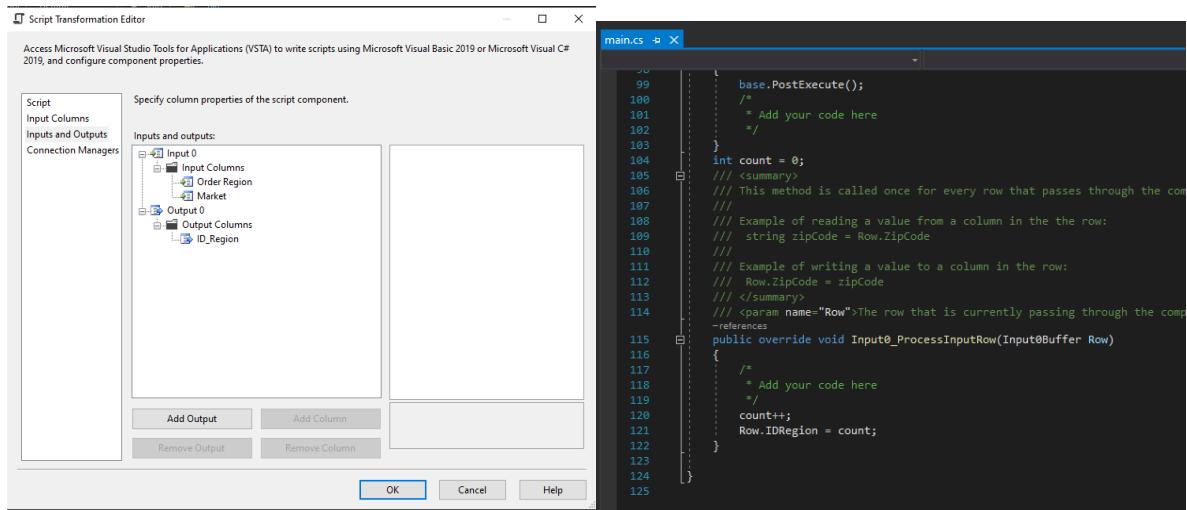
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



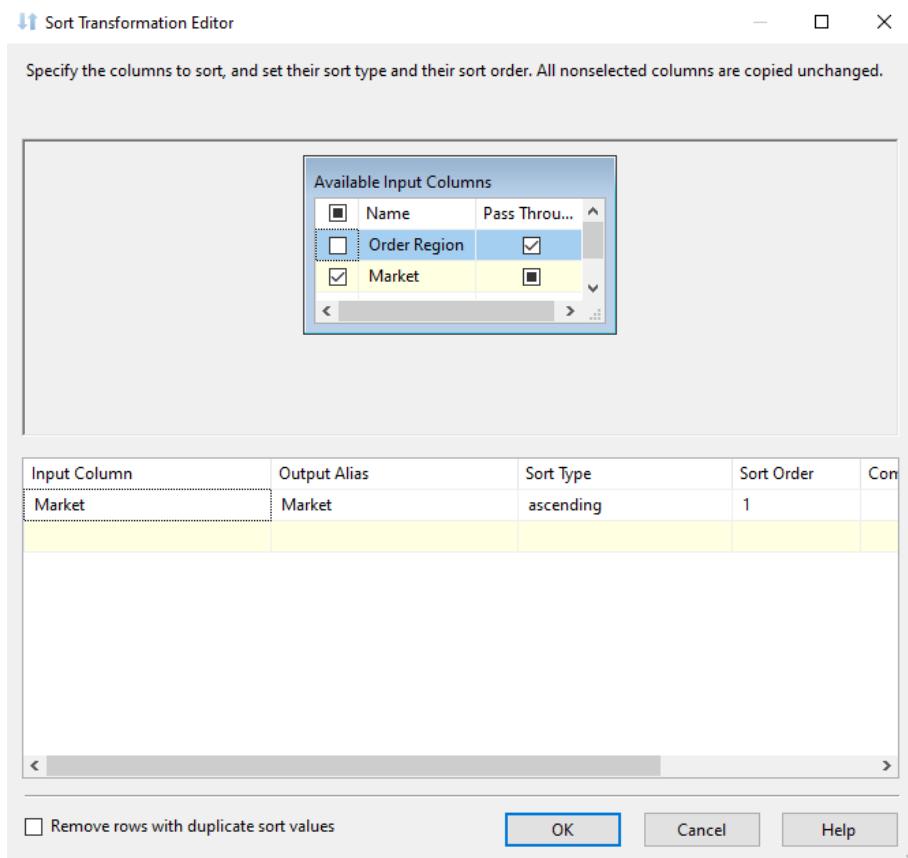
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



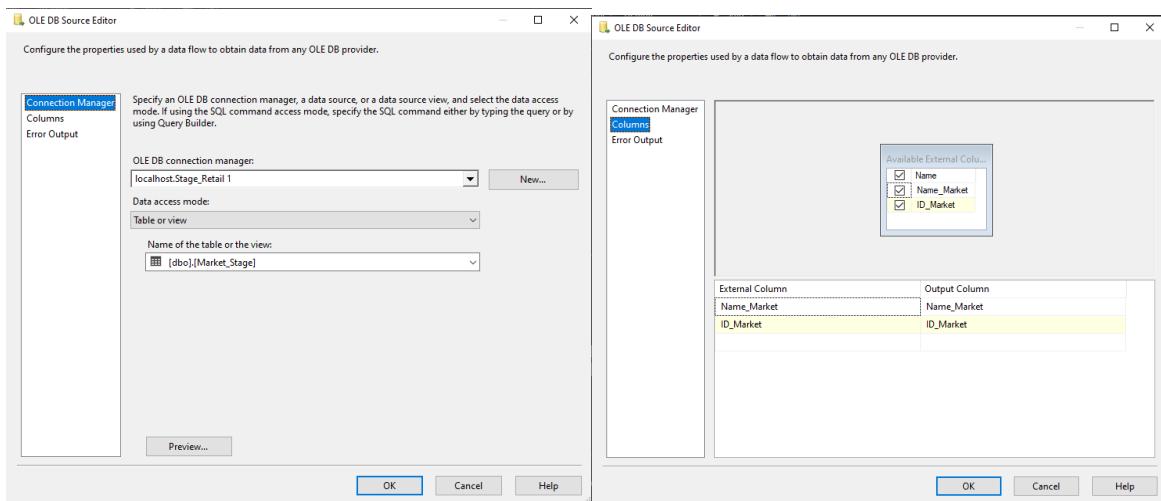
- Dùng Script Transformation để tạo thêm ID tự động cho bảng. Thay đổi code trong Script



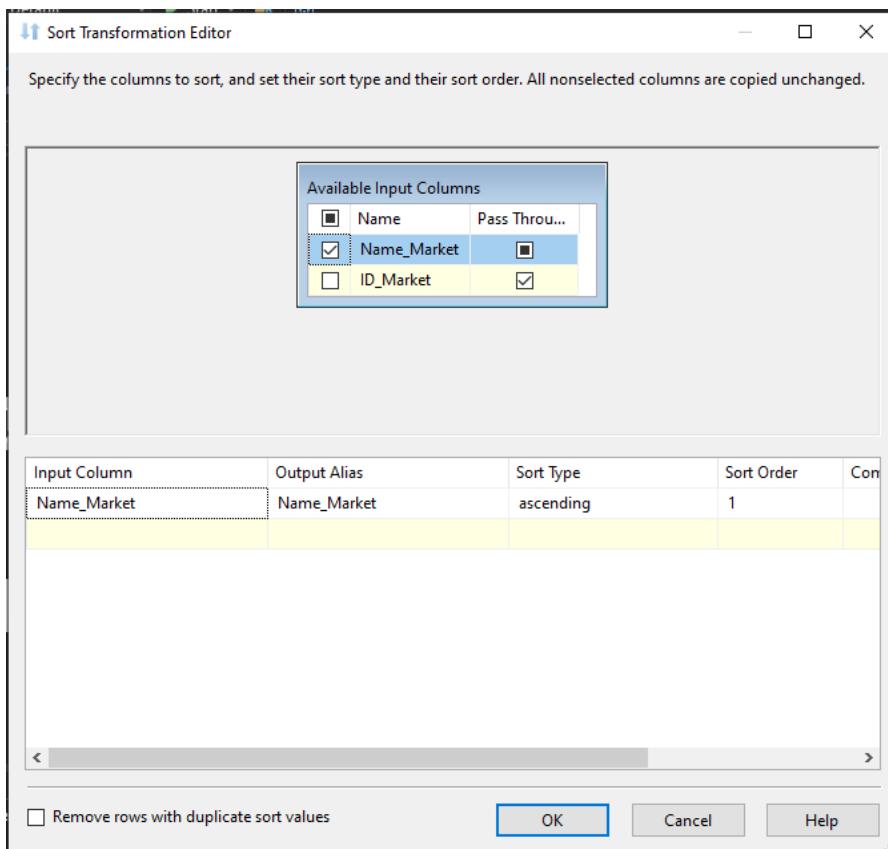
- Dùng sort để chỉ định thuộc tính dùng để inner join



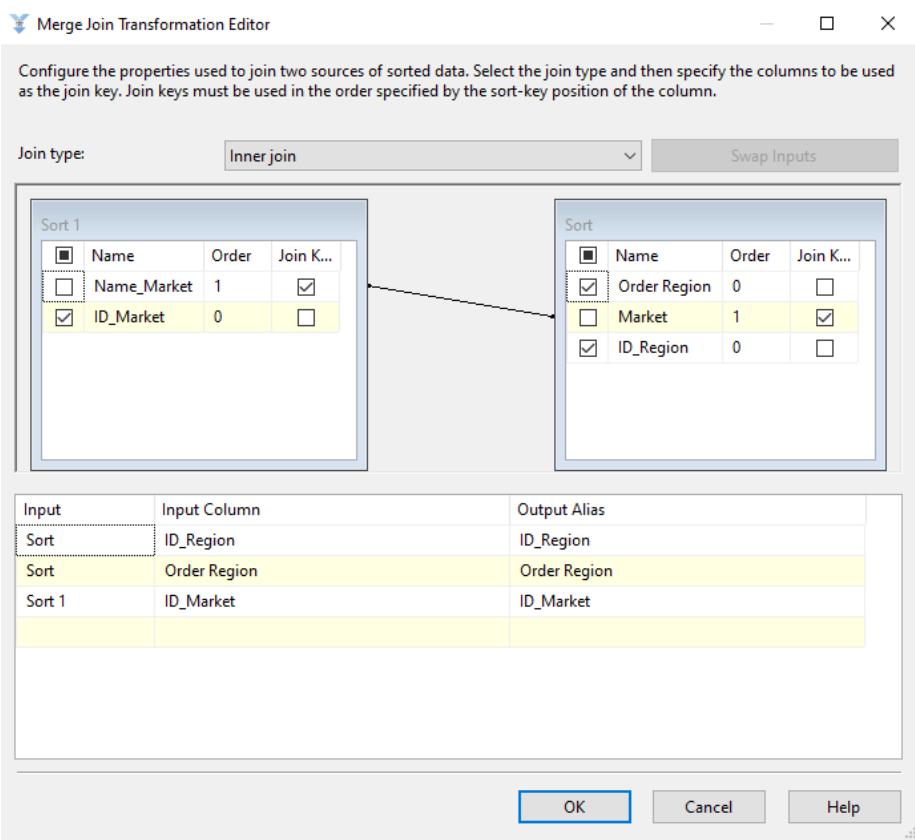
- Chọn nguồn là Stage Market để inner join



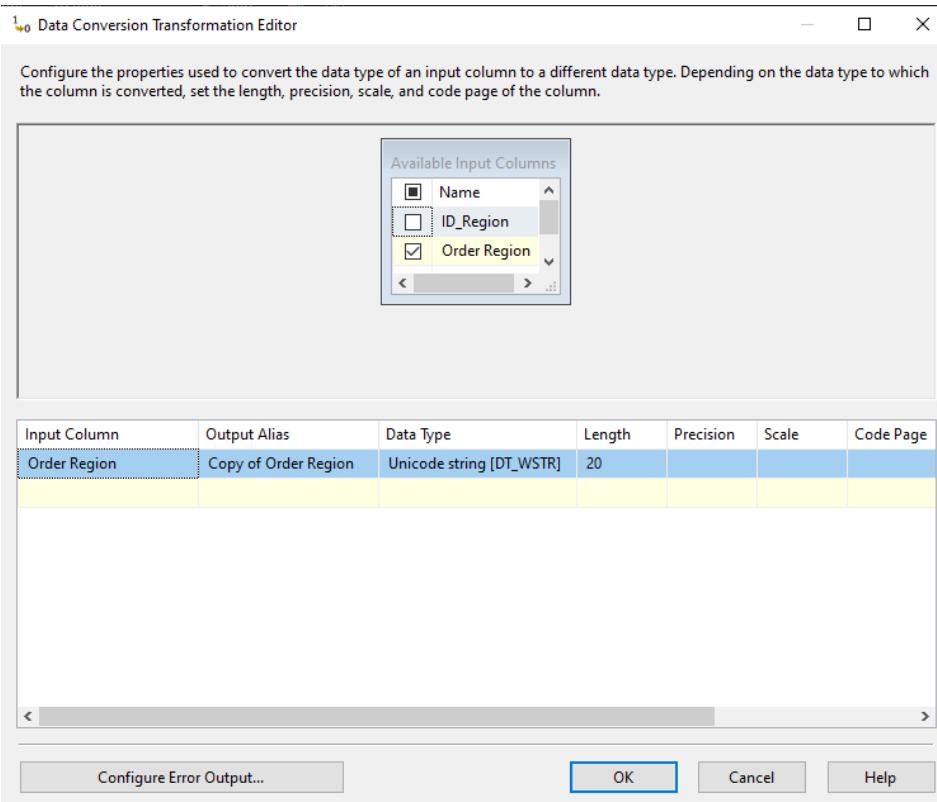
- Dùng **sort** để chỉ định thuộc tính của **Stage Market** dùng để **inner join**



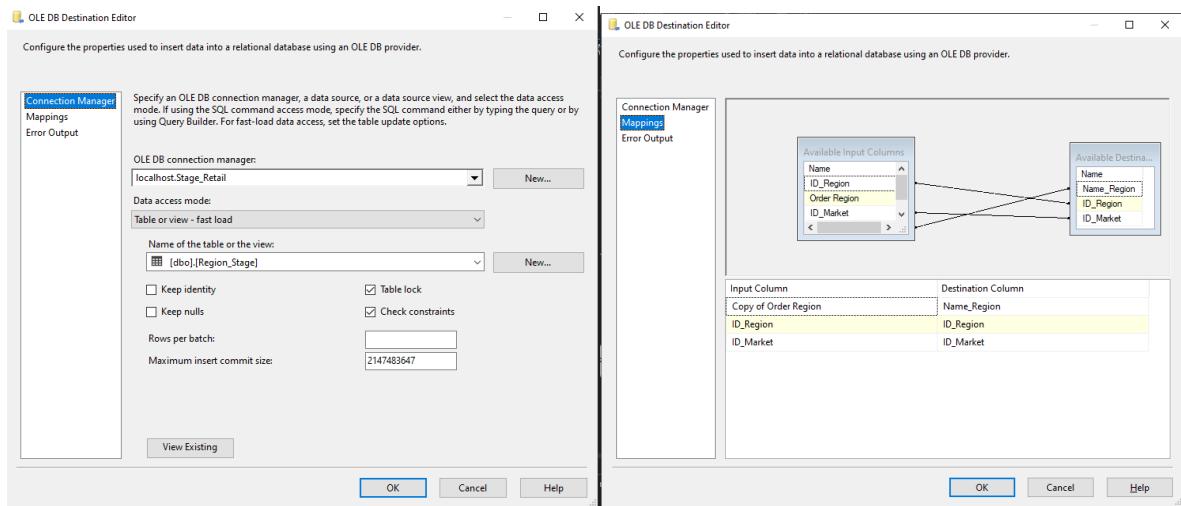
- Sử dụng **Merge Join** để kết hợp hai bảng lại với nhau để lấy được **ID\_Market** và chọn các thuộc tính cần thiết khác



- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào Stage Region

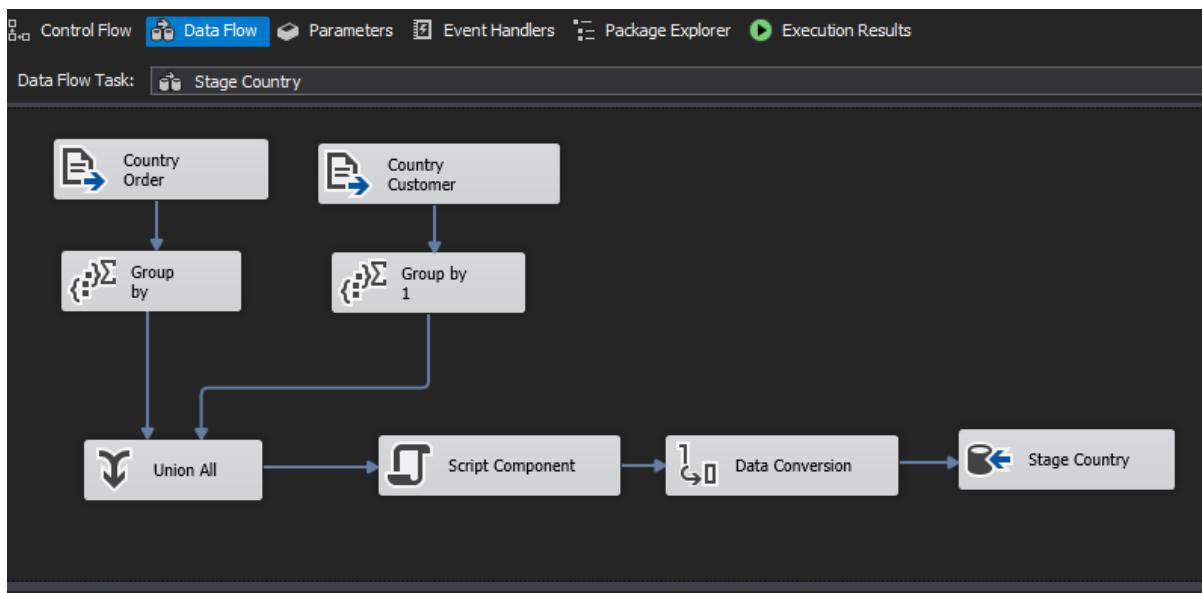


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn

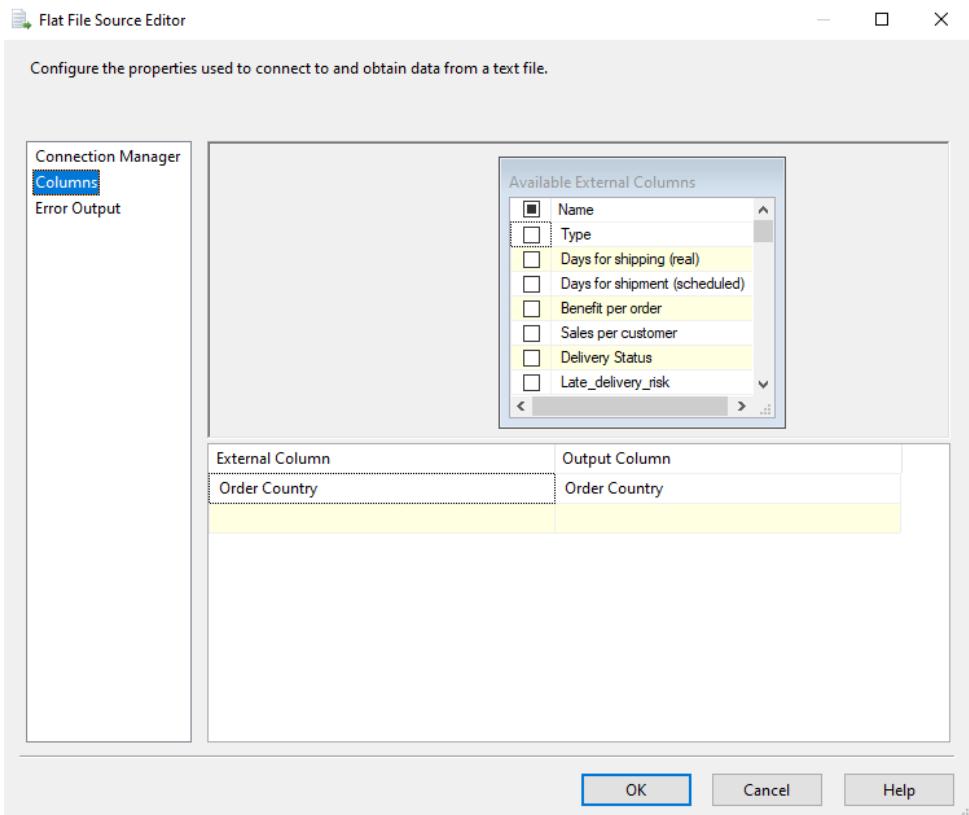


### 3.2.8. Stage Country

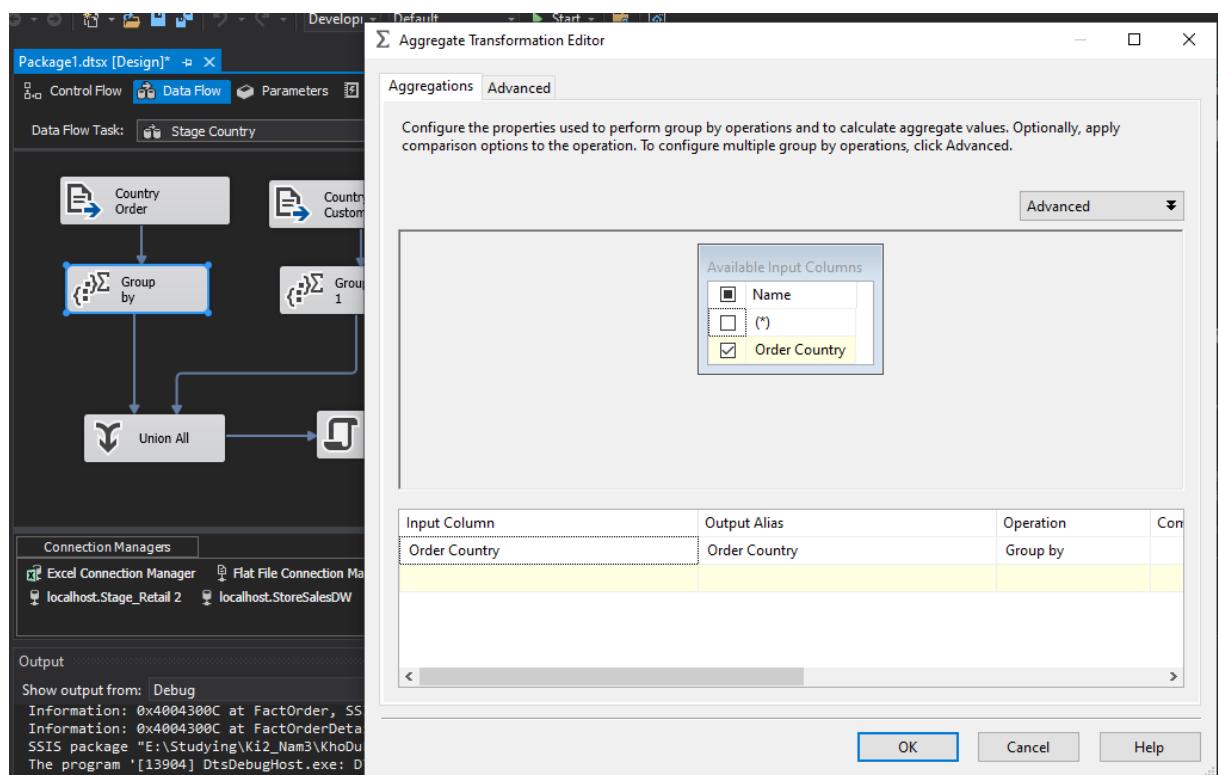
- Cách thành phần của Stage Country



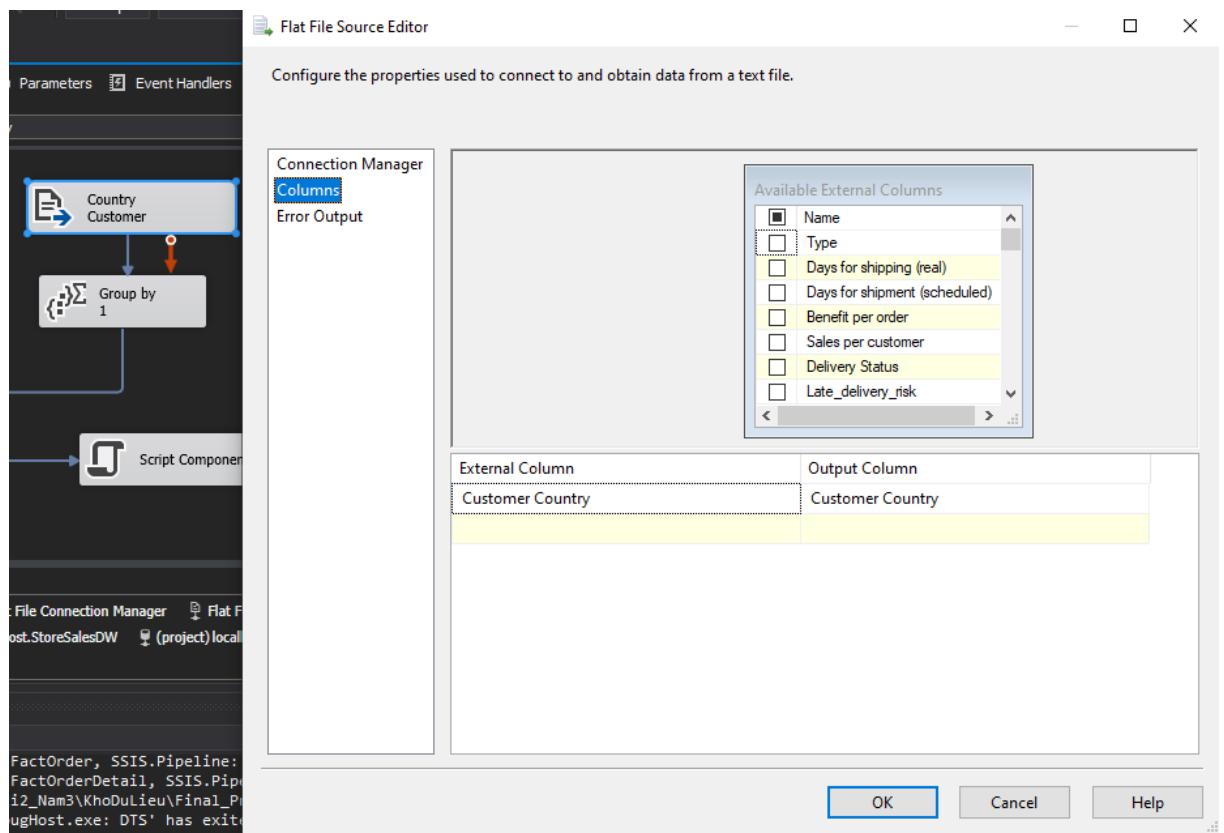
- Từ file excel nguồn đầu tiên sẽ lấy cột Country của order



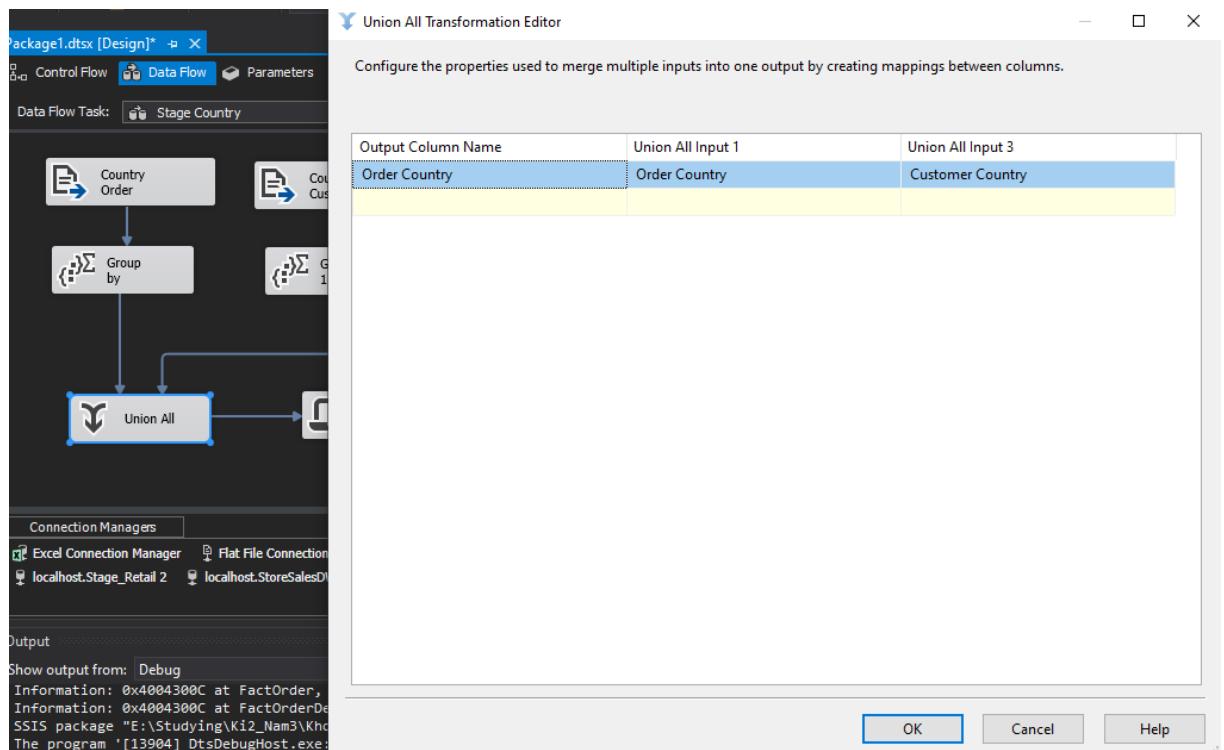
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



- Từ file excel nguồn thứ 2 sẽ lấy sẽ lấy cột Country của customer



- Sau đó dùng Union All để kết hợp hai nguồn này với nhau



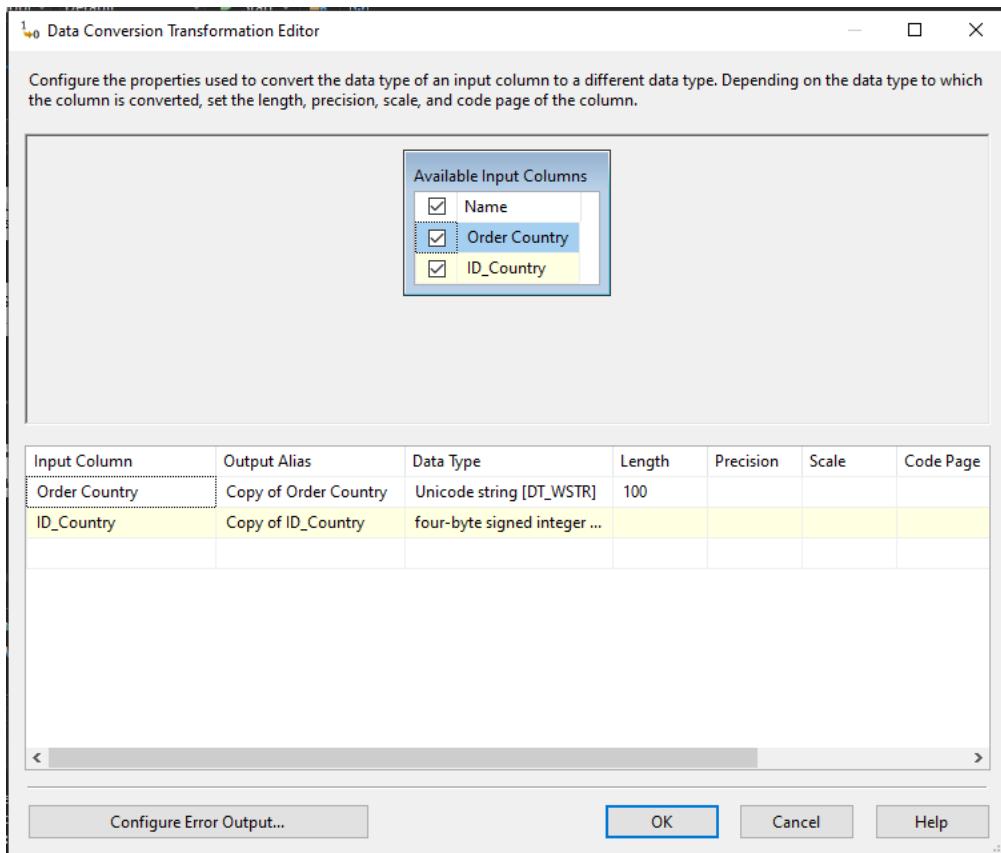
- Dùng Script Transformation để tạo thêm ID tự động cho bảng. Thay đổi code trong Script

```

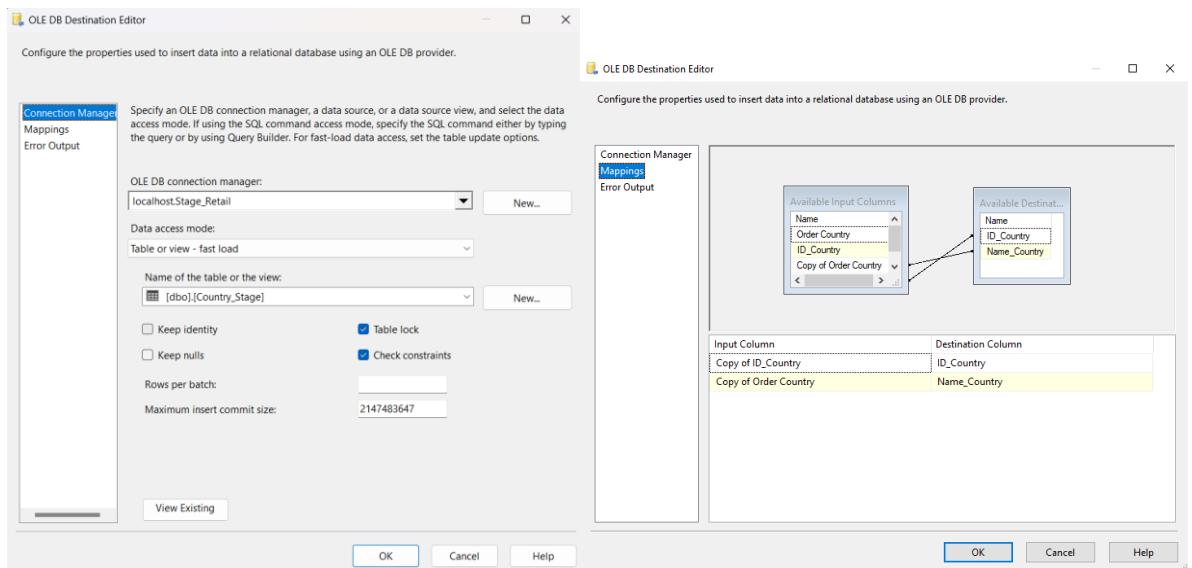
    101     */
    102     int count = 0;
    103     /// <summary>
    104     /// This method is called once for every row that passes through this component.
    105     ///
    106     /// Example of reading a value from a column in the row.
    107     /// string zipCode = Row.ZipCode
    108     ///
    109     /// Example of writing a value to a column in the row.
    110     /// Row.ZipCode = zipCode
    111     ///
    112     /// </summary>
    113     /// <param name="Row">The row that is currently passing through this component.
    114     public override void Input0_ProcessInputRow(Input0Buffer Row)
    115     {
    116         /*
    117         * Add your code here
    118         */
    119         count++;
    120         Row.IDCountry = count;
    121     }
    122 }
    123
    124
    125

```

- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào Stage Country

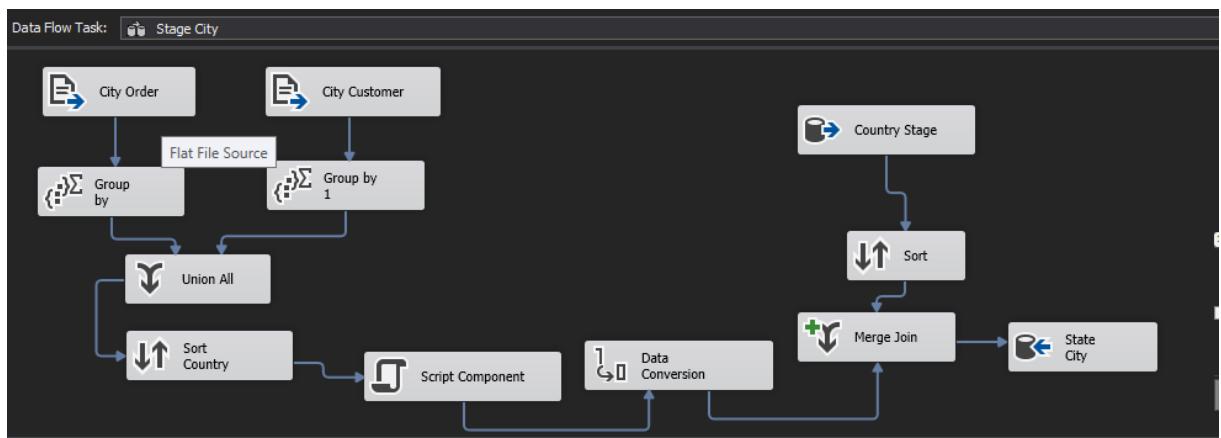


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn

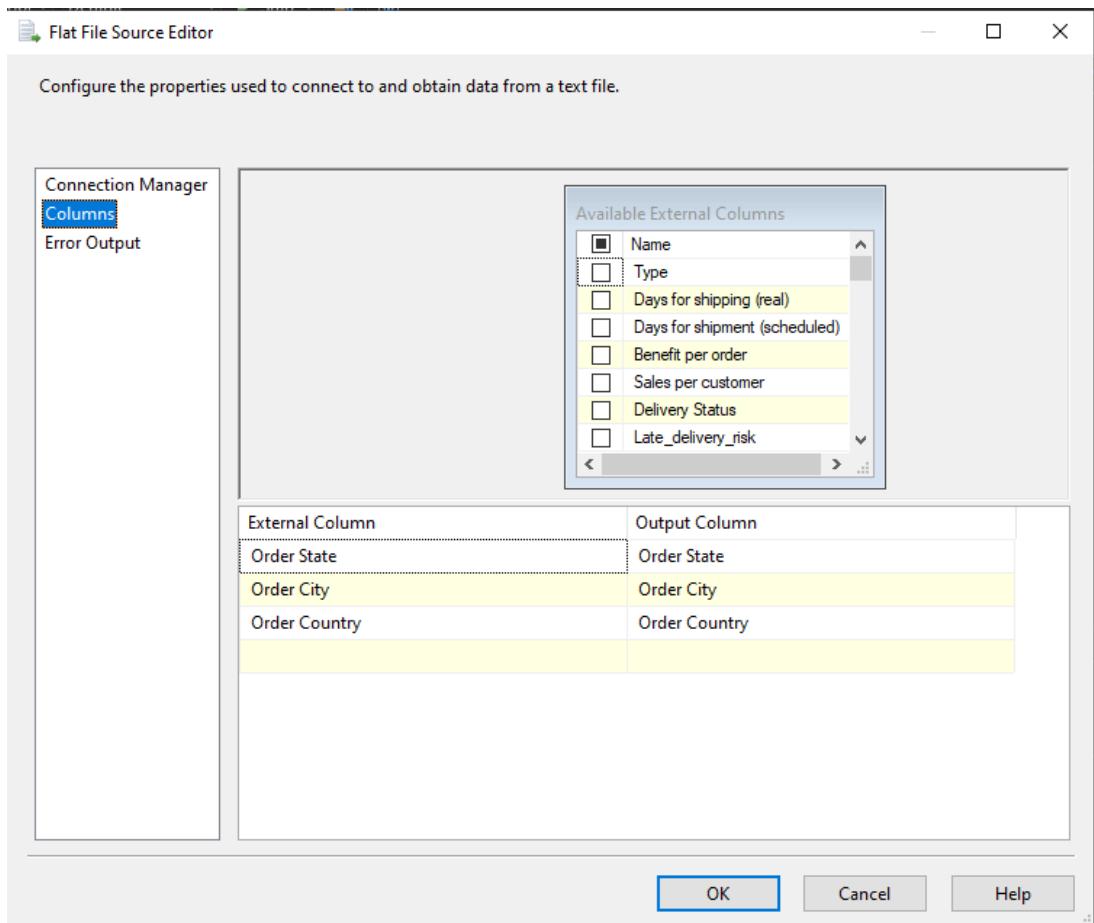


### 3.2.9. Stage City

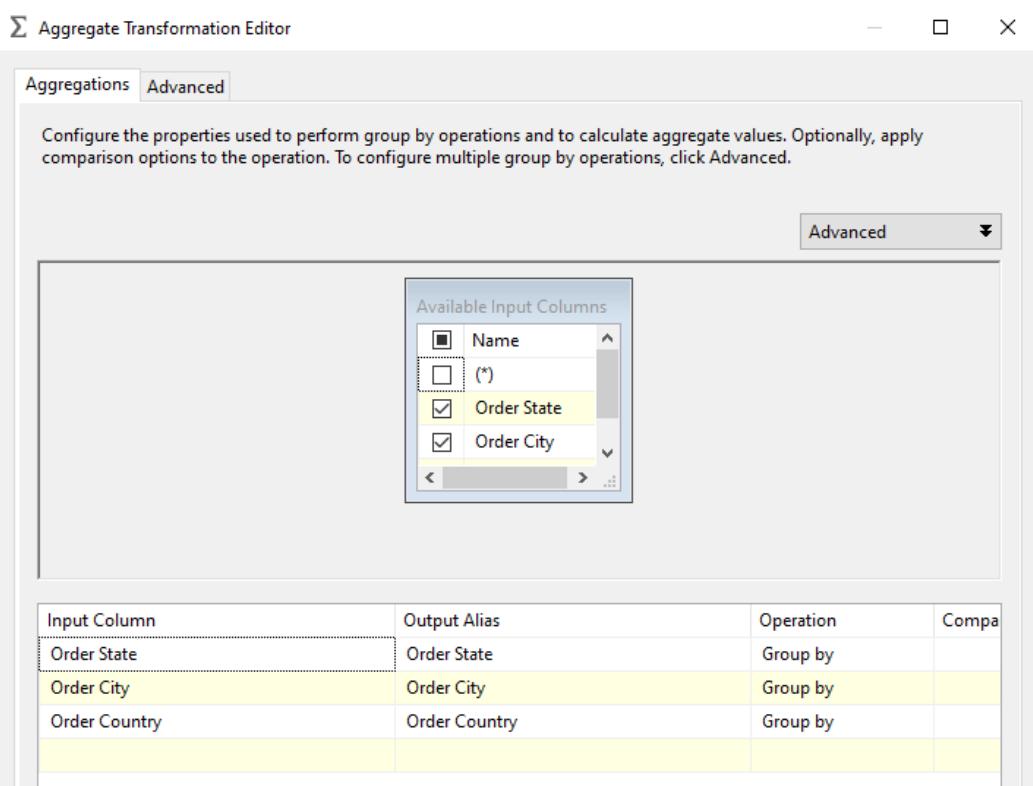
- Cách thành phần của **Stage City**



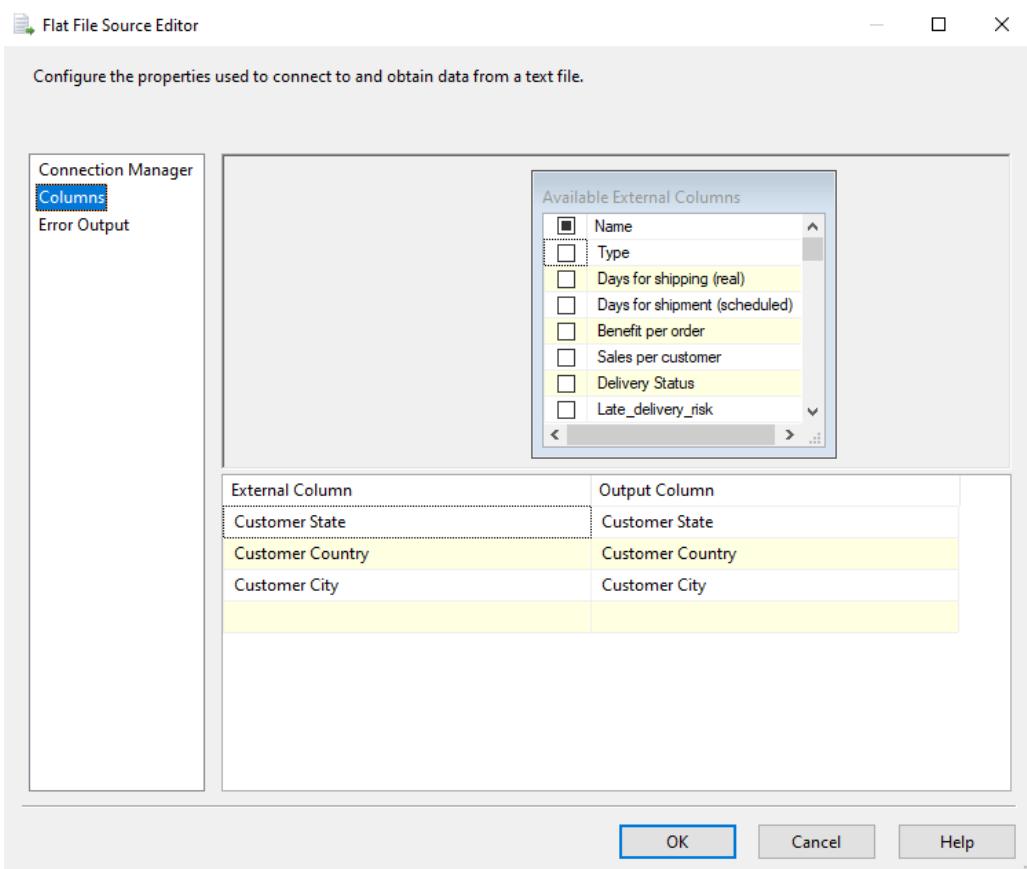
- Từ file excel nguồn đầu tiên sẽ lấy cột liên quan tới City



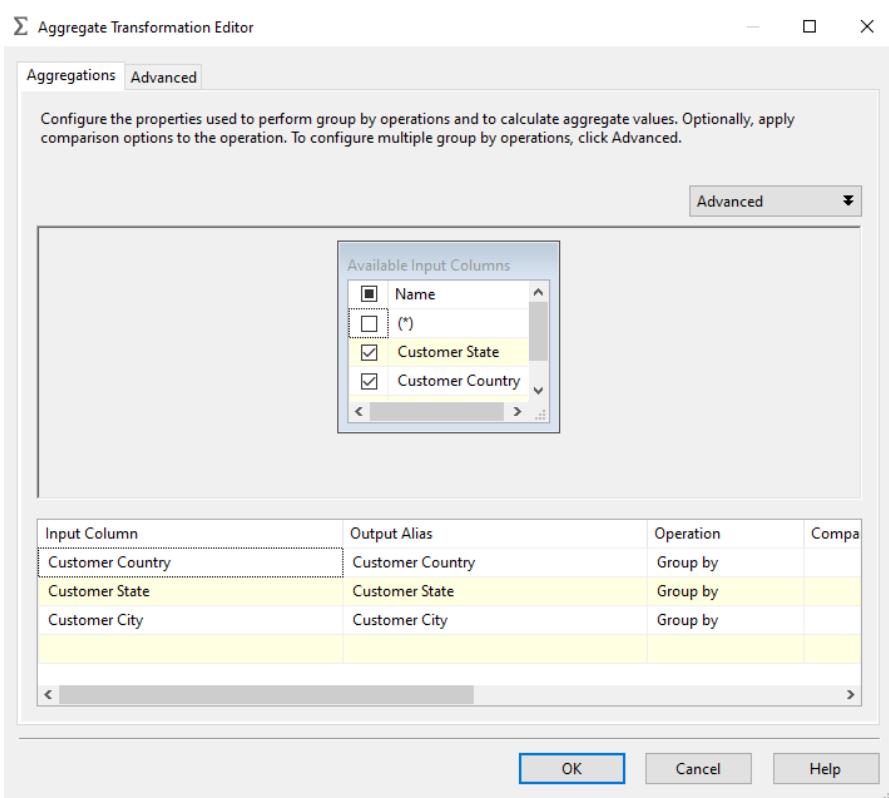
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



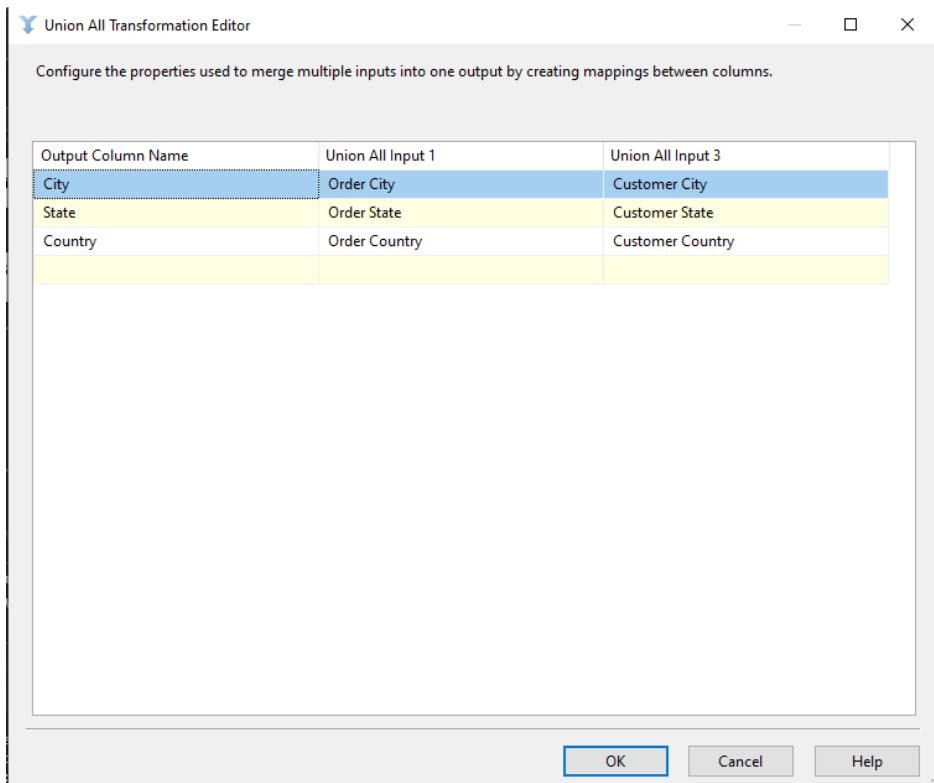
- Từ file excel nguồn thứ 2 sẽ lấy sẽ lấy cột liên quan tới customer



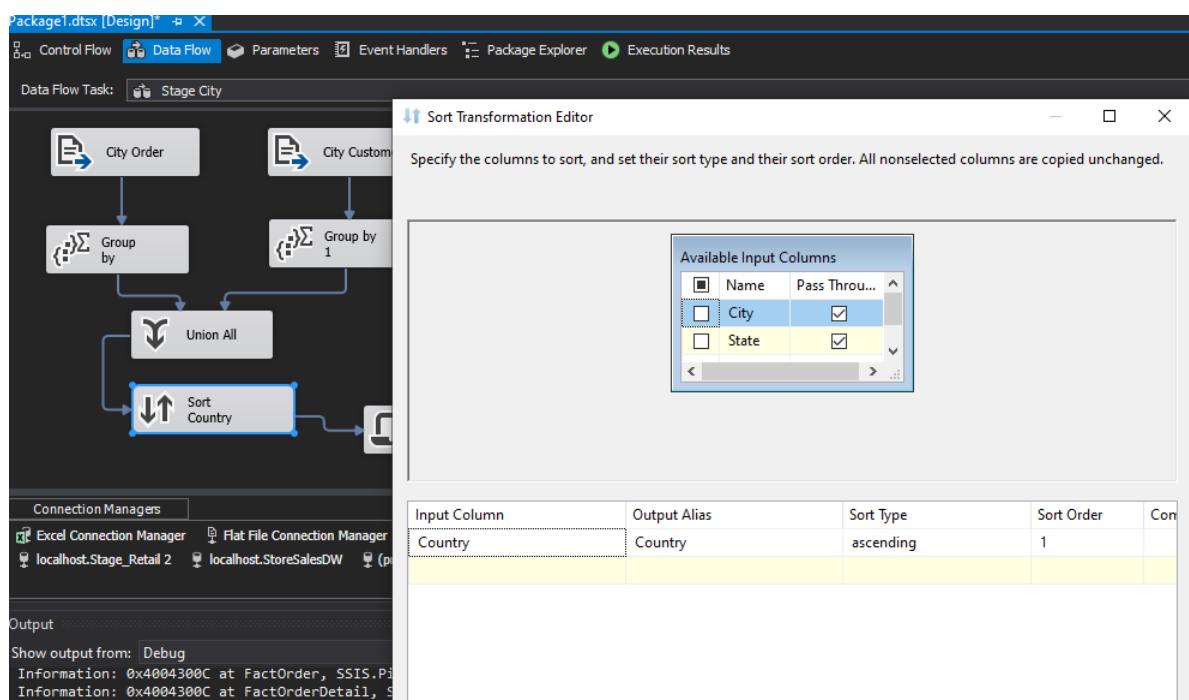
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



- Sau đó dùng Union All để kết hợp hai nguồn này với nhau



- Dùng sort để chỉ định thuộc tính của kết quả của Union All dùng để inner join



- Dùng Script Transformation để tạo thêm ID tự động cho bảng. Thay đổi code trong Script

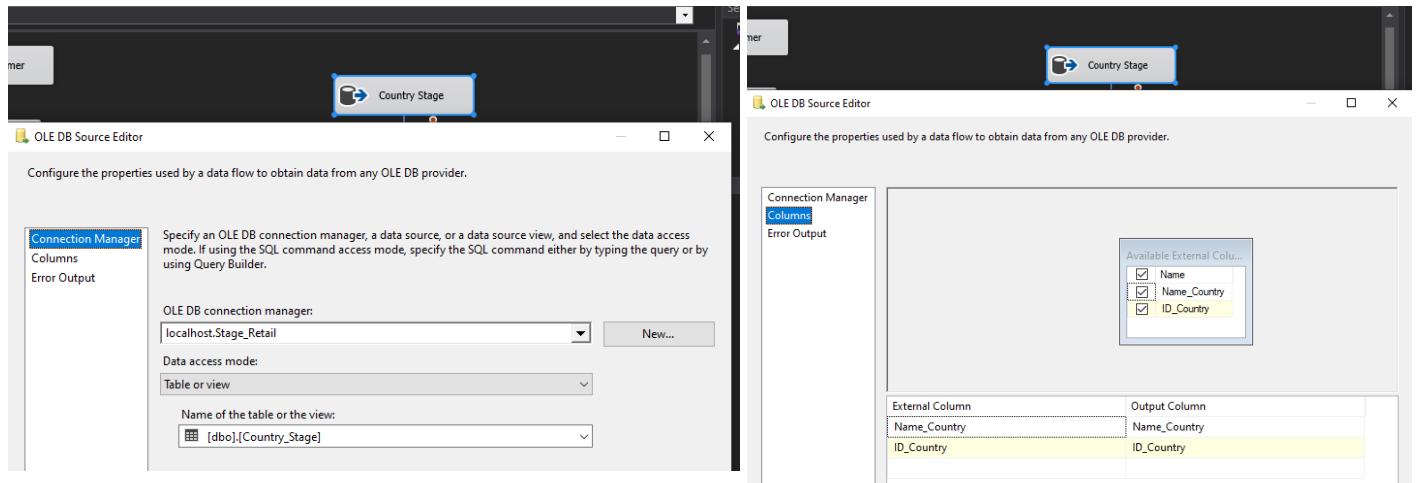
The screenshot shows two windows side-by-side. On the left is the 'Script Transformation Editor' showing the 'Inputs and outputs' configuration for 'Input 0'. It lists 'Input Columns' (City, State, Country) and 'Output Columns' (ID). On the right is the 'main.cs' code editor with the following C# code:

```

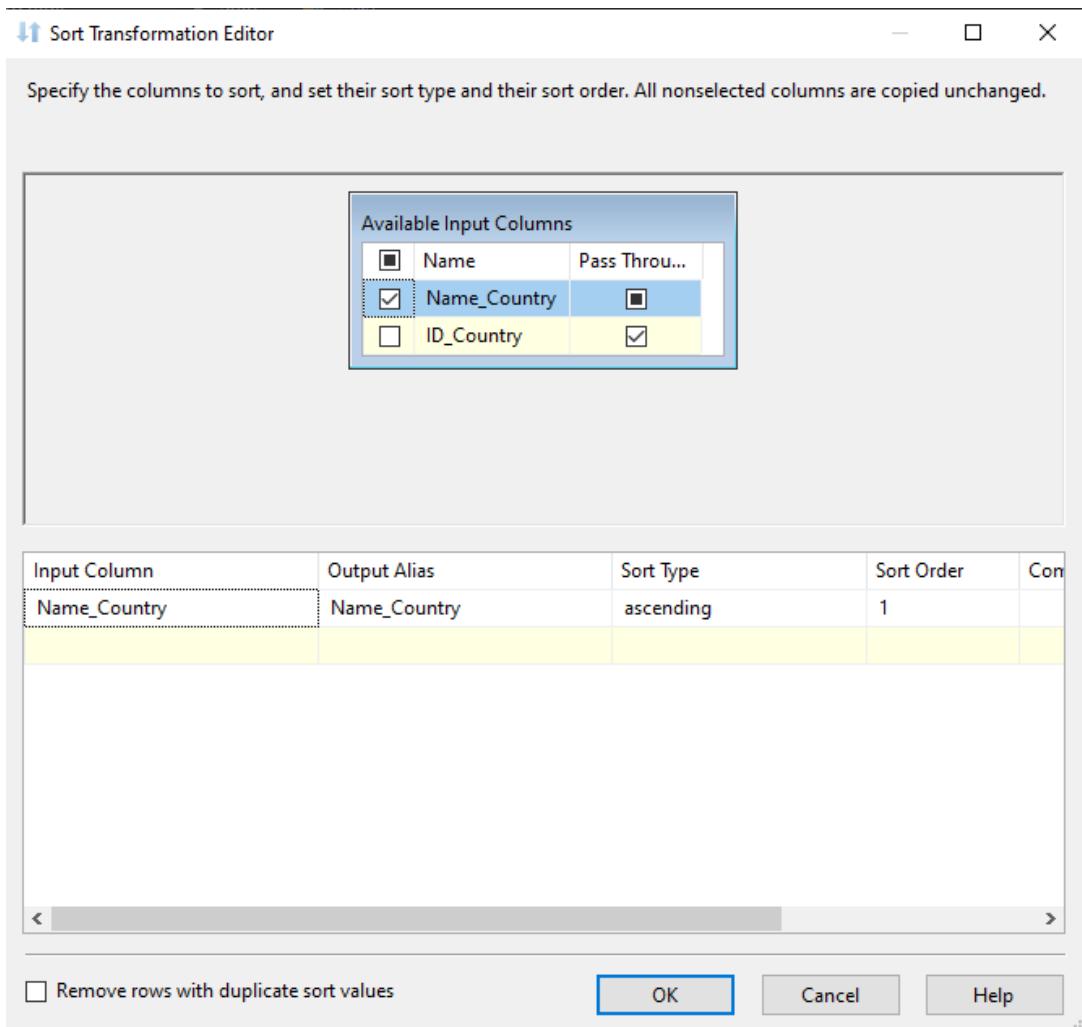
    /*
    * Add your code here
    */
    int count = 0;
    /// <summary>
    /// This method is called once for every row that passes through this component.
    /// Example of reading a value from a column in the row:
    /// string zipCode = Row.ZipCode
    /// 

```

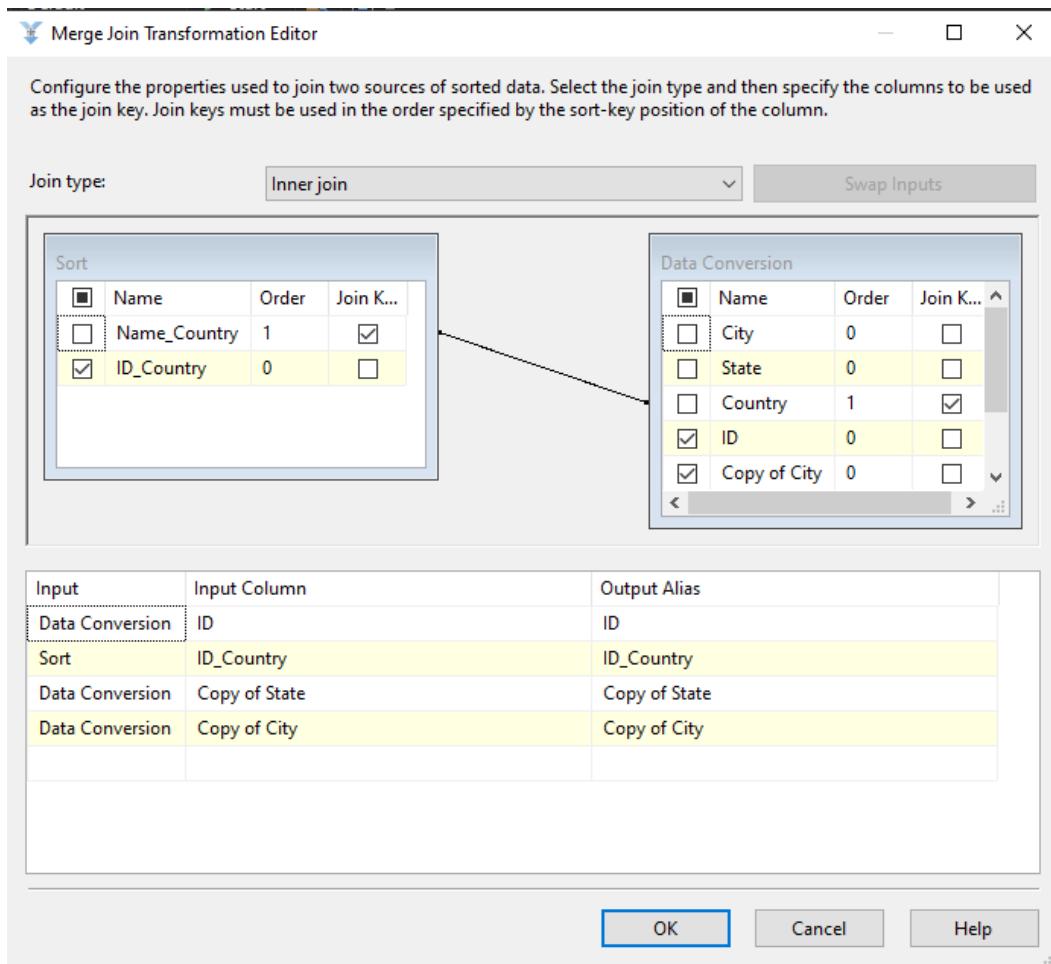
- Chọn nguồn là Stage Country để inner join



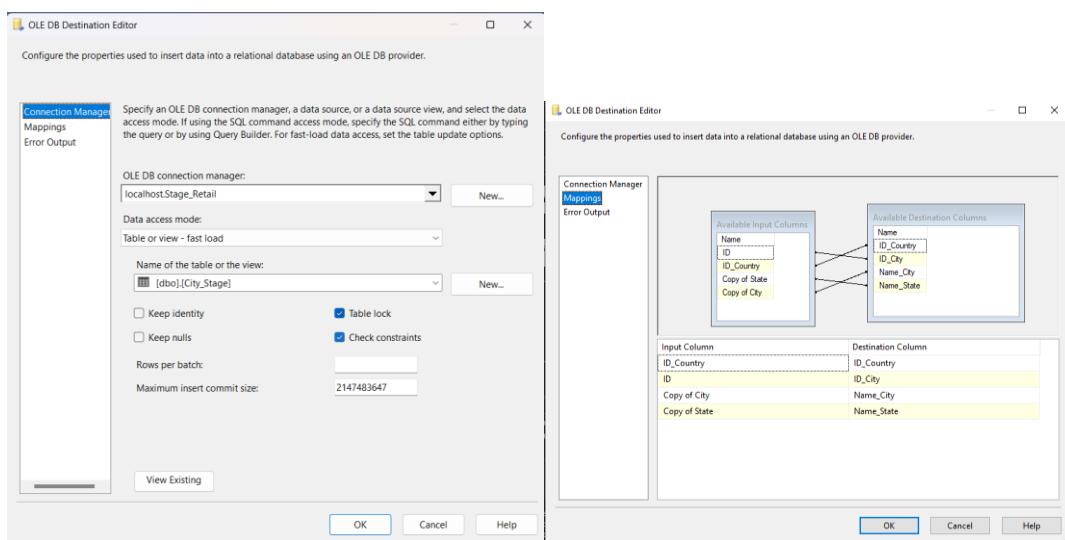
- Dùng sort để chỉ định thuộc tính của Stage Country dùng để inner join



- Sử dụng Merge Join để kết hợp hai bảng lại với nhau để lấy được ID\_Country và chọn các thuộc tính cần thiết khác

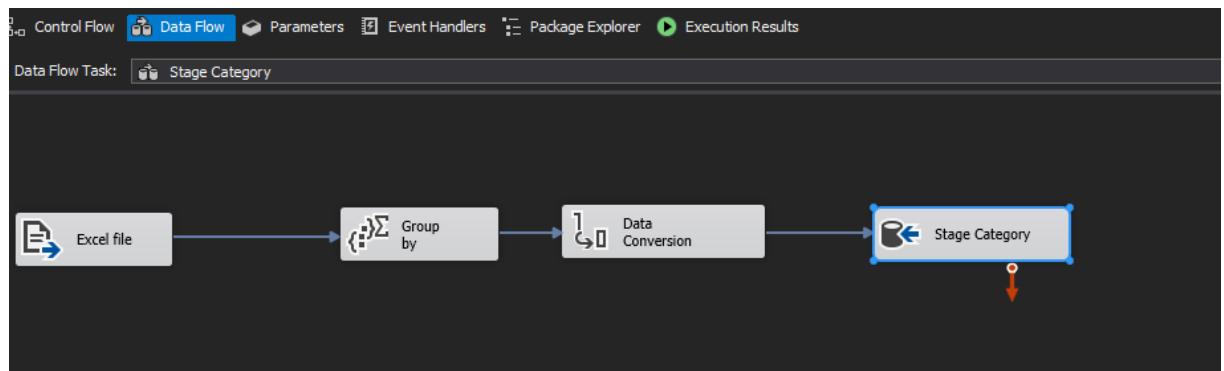


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn

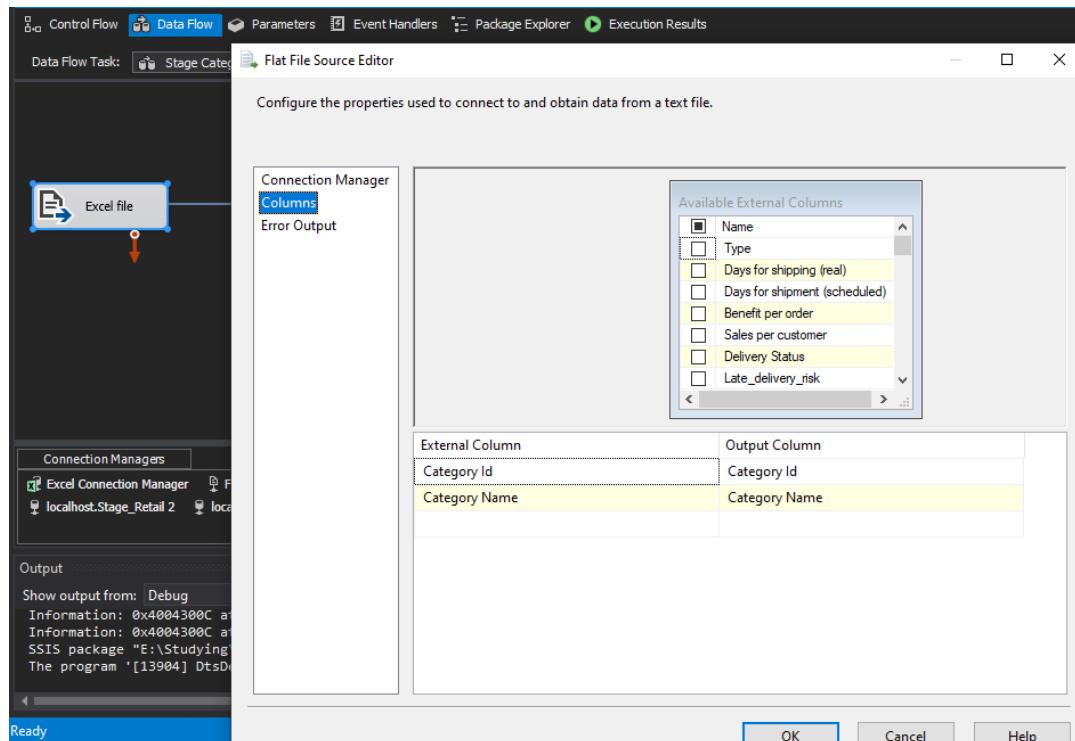


### 3.2.10.Stage Category

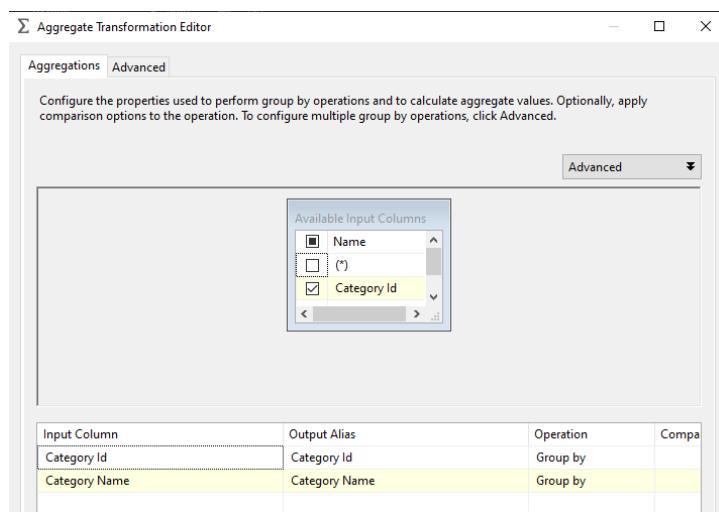
- Cách thành phần của Stage Category



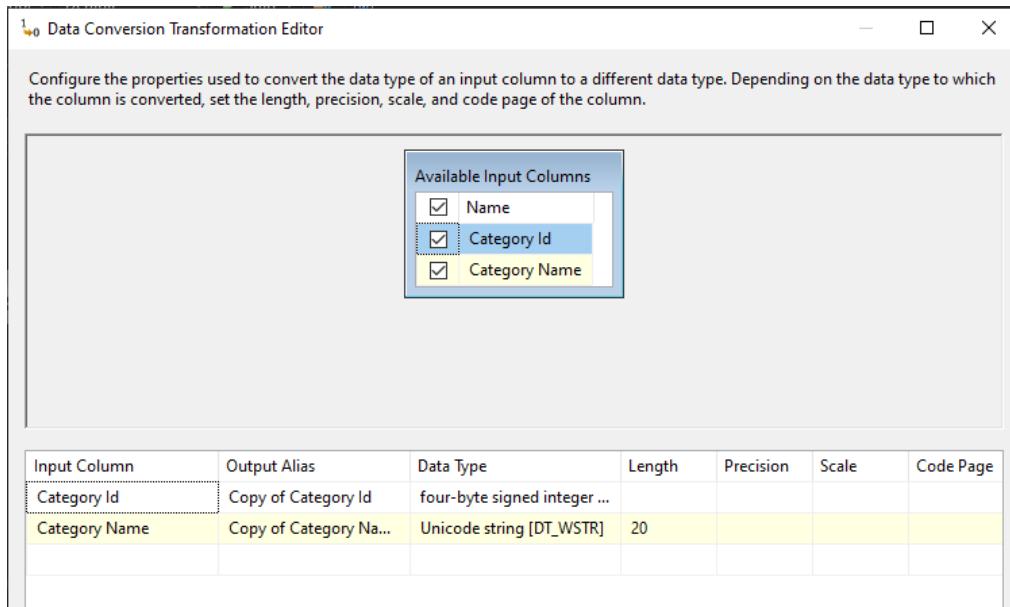
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



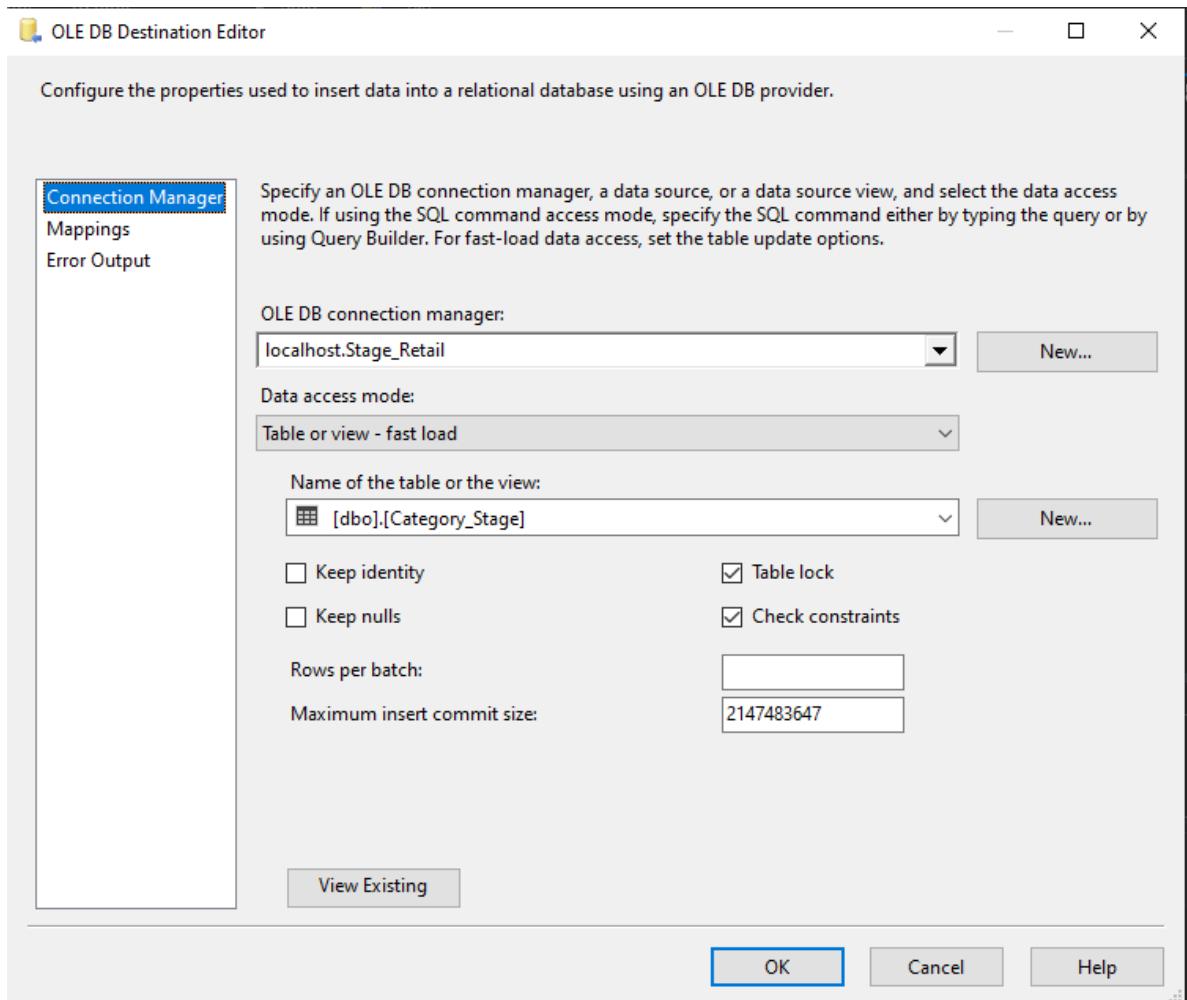
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.

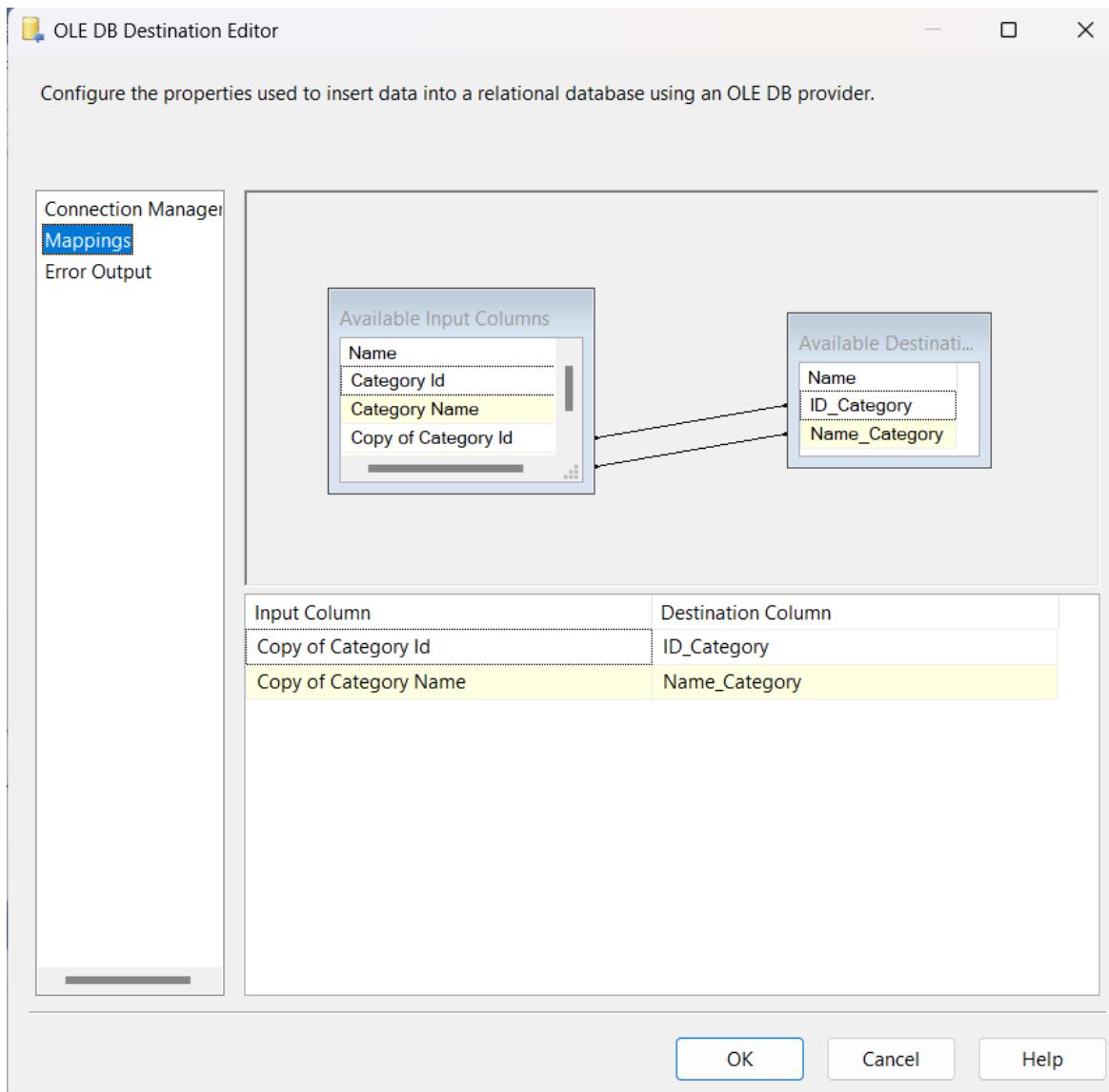


- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào Stage Catagory



- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn



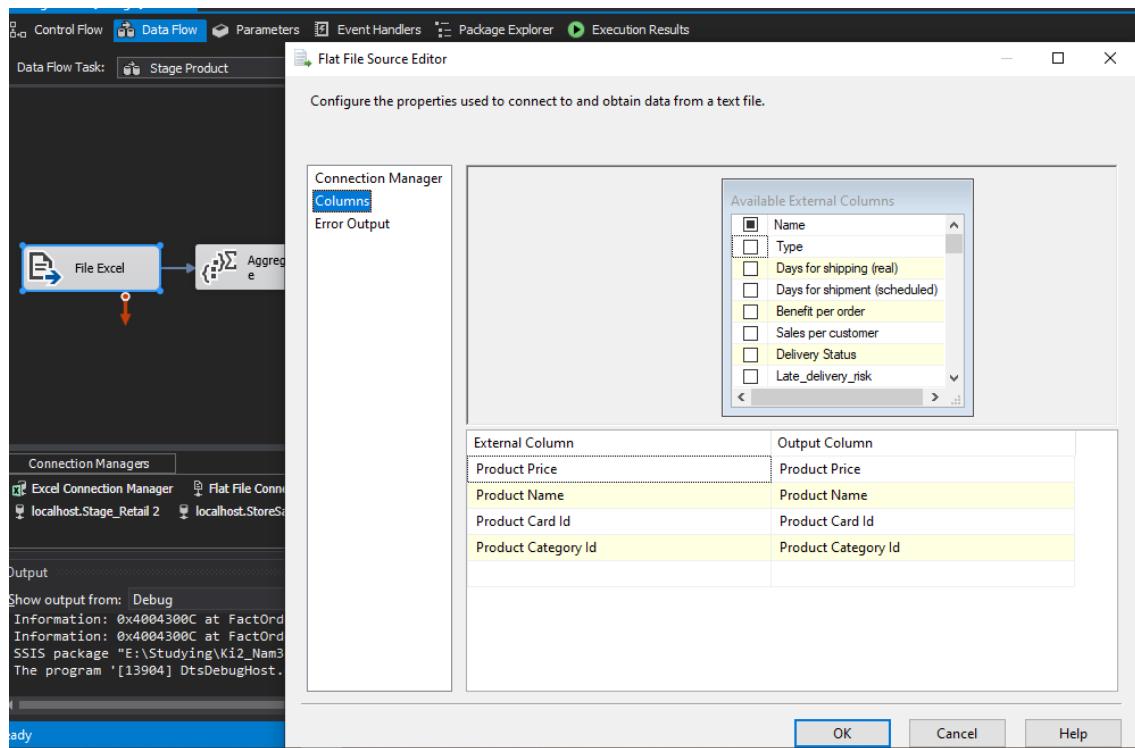


### 3.2.11.Stage Product

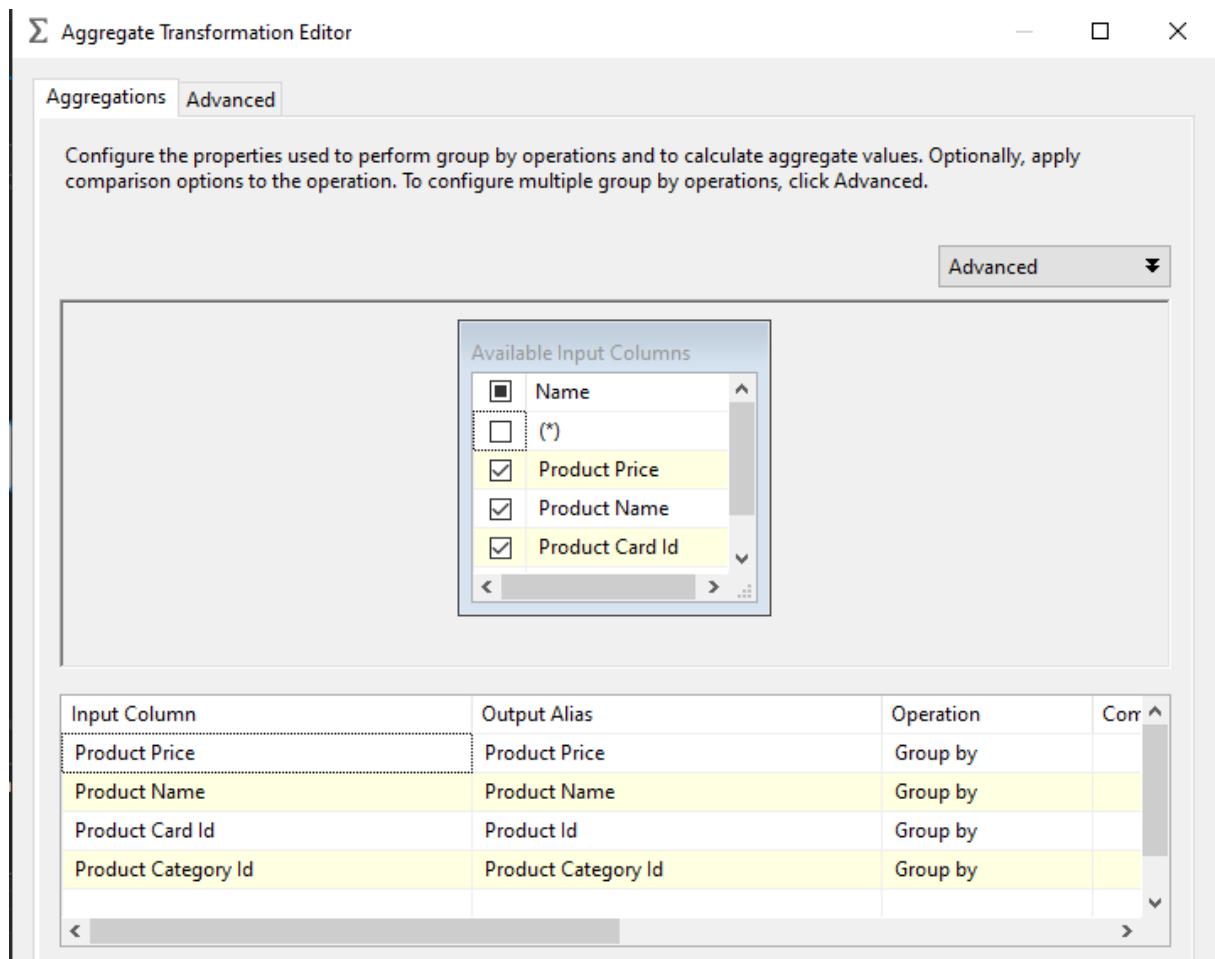
- Cách thành phần của Stage Product



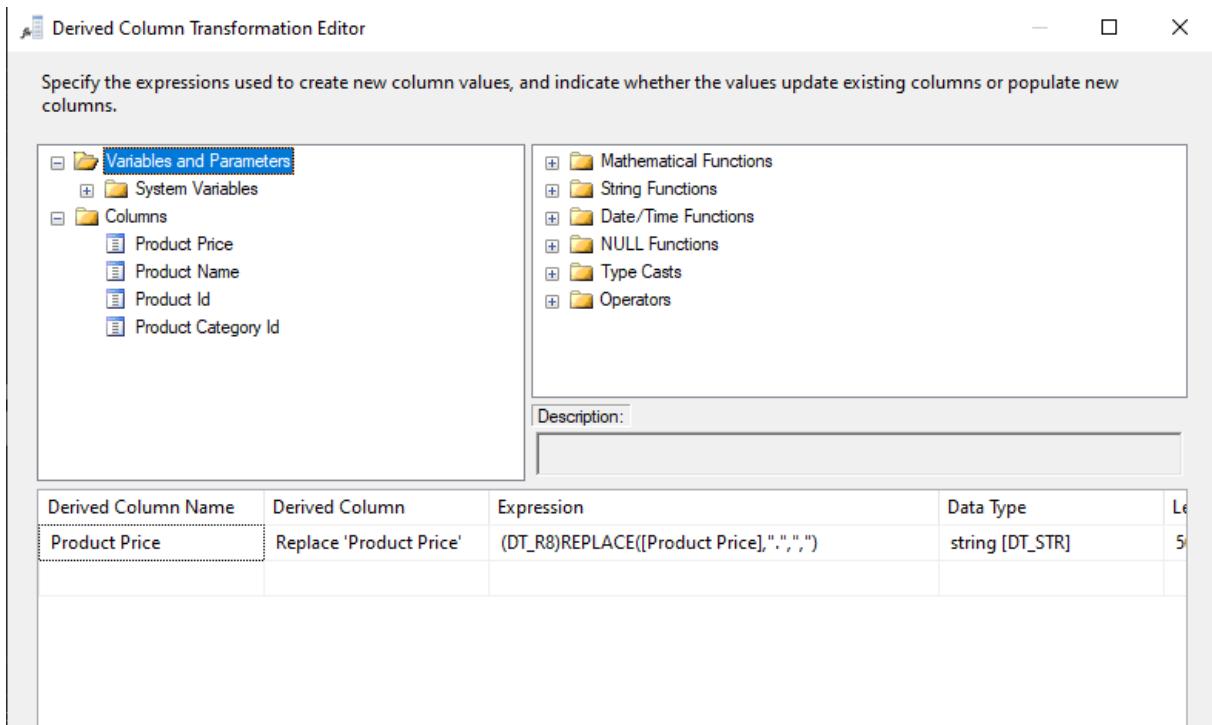
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



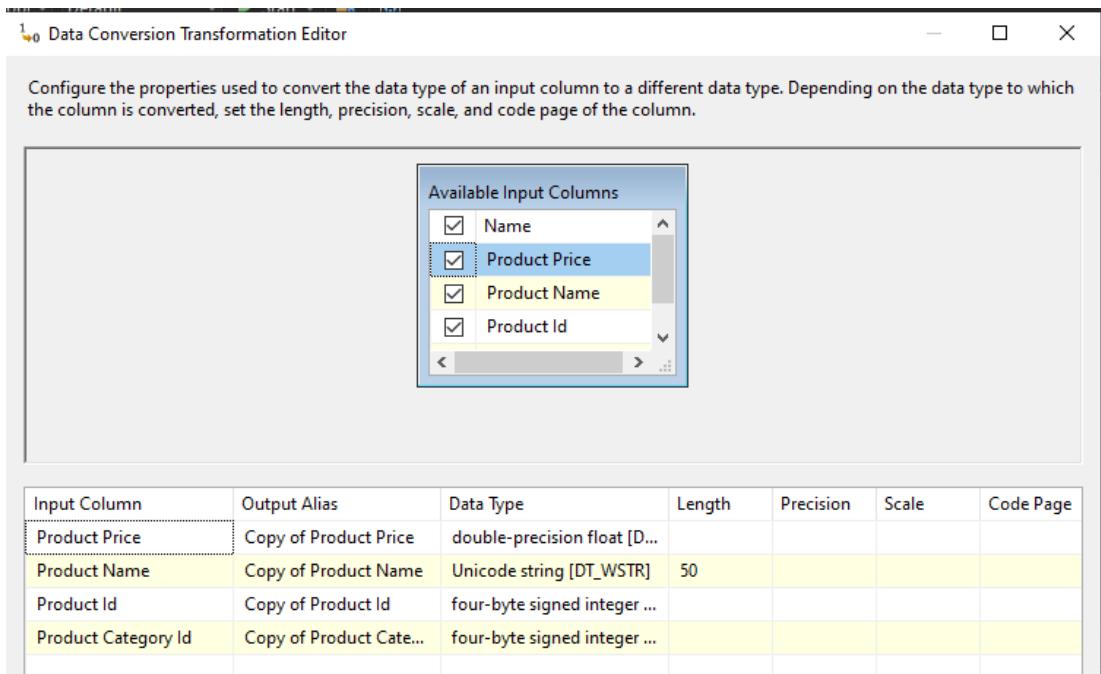
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



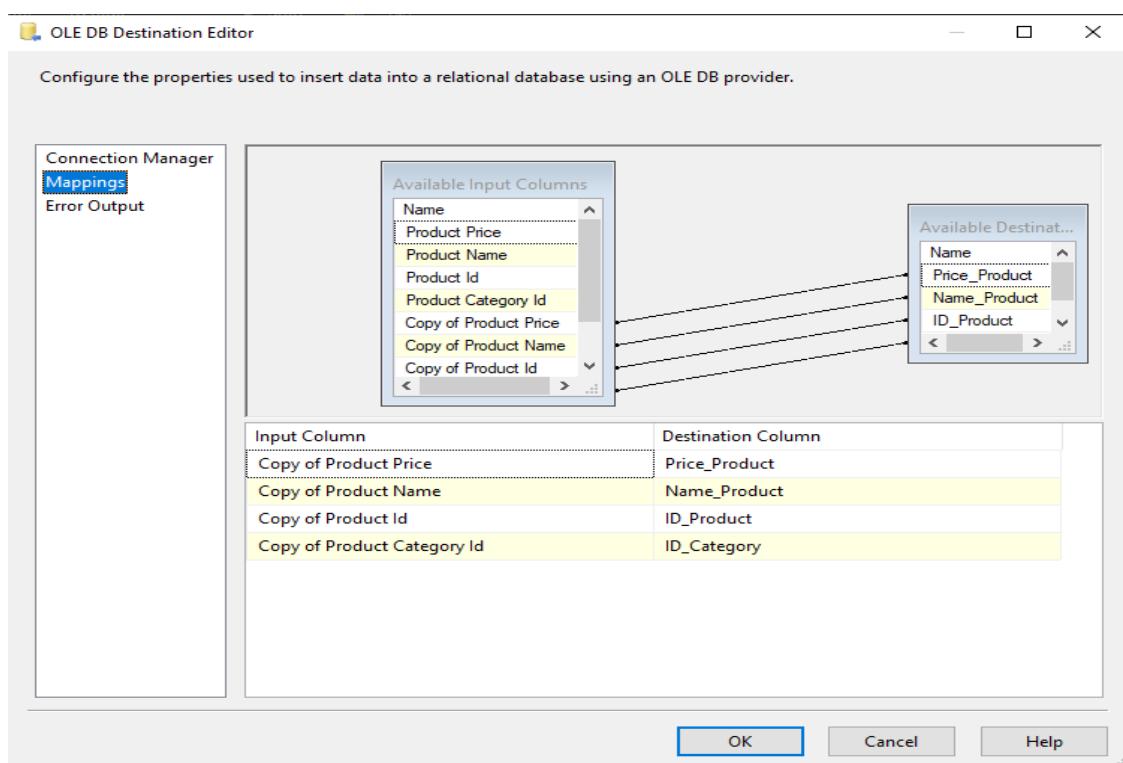
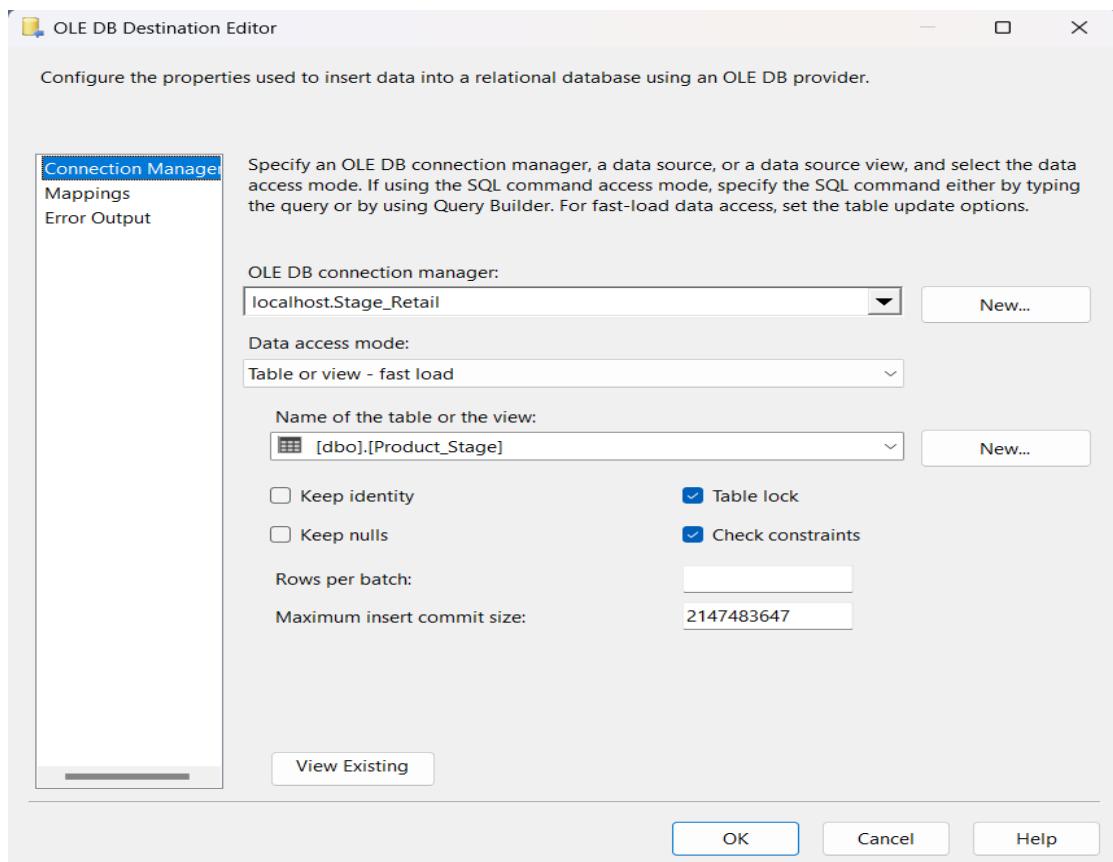
- Dùng Derived Column để định dạng lại cột Product Price



- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp trước khi đưa vào Stage Product

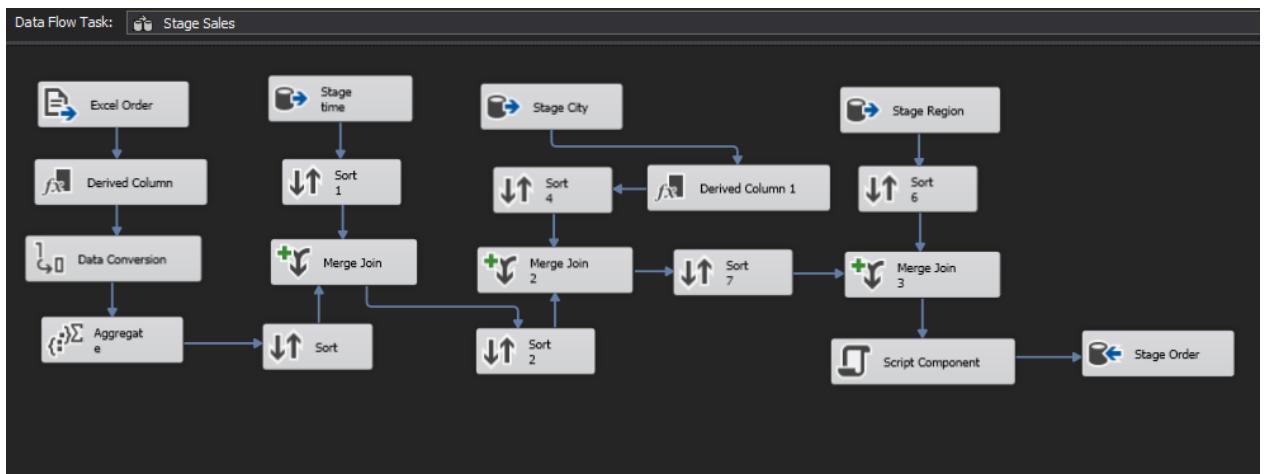


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn



### 3.2.12.Stage Sales

- Cách thành phần của Stage Sales



- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.

Flat File Source Editor

Configure the properties used to connect to and obtain data from a text file.

**Connection Manager**

- Columns
- Error Output

**Available External Columns**

Name	Type
<input checked="" type="checkbox"/> Days for shipping (real)	
<input type="checkbox"/> Days for shipment (scheduled)	
<input type="checkbox"/> Benefit per order	
<input type="checkbox"/> Sales per customer	
<input checked="" type="checkbox"/> Delivery Status	
<input checked="" type="checkbox"/> Late_delivery_risk	
<input type="checkbox"/> Category Id	
<input type="checkbox"/> Category Name	

**External Column**

External Column	Output Column
Order Id	Order Id
Customer Id	Customer Id
order date (DateOrders)	order date (DateOrders)
shipping date (DateOrders)	shipping date (DateOrders)
Type	Type
Delivery Status	Delivery Status
Late_delivery_risk	Late_delivery_risk
Days for shipping (real)	Days for shipping (real)
Order Profit Per Order	Order Profit Per Order
Order City	Order City

**OK**   **Cancel**   **Help**

- Sử dụng Derived Column để thêm các cột và định dạng lại cột.

Derived Column Transformation Editor

Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

Variables and Parameters

Columns

Mathematical Functions

String Functions

Date/Time Functions

NULL Functions

Type Casts

Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Length
ID_Time	<add as new column>	[order date (DateOrders)] + [shipping date (DateOrders)]	Unicode string [DT_WSTR]	1
CityState	<add as new column>	[Order City] + [Order State]	Unicode string [DT_WSTR]	1
Order Item Total	Replace 'Order Item ...'	(DT_R8)REPLACE([Order Item Total],",","")	string [DT_STR]	5
Order Profit Per Order	Replace 'Order Profit Per Order'...	(DT_R8)REPLACE([Order Profit Per Order],",","")	string [DT_STR]	5

- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

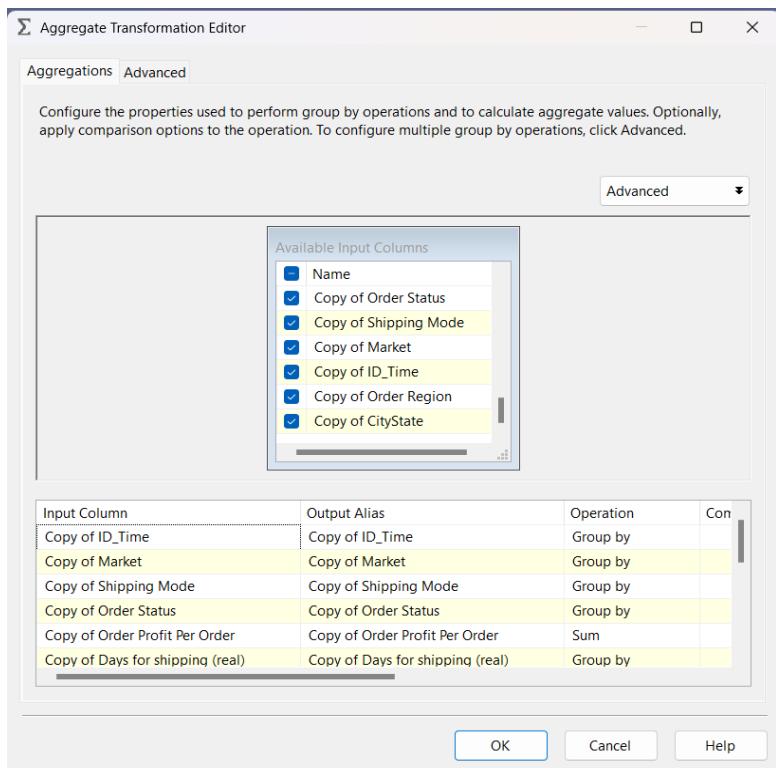
Available Input Columns

- Name
- Order Id**
- Customer Id
- order date (DateOrders)
- shipping date (DateOrders)
- Type
- Delivery Status

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
Order Item Total	Copy of Order Item ...	double-precision float ...				
Customer Id	Copy of Customer Id	four-byte signed integer				
Order Id	Copy of Order Id	four-byte signed integer				
Type	Copy of Type	Unicode string [DT_WSTR]	20			
Delivery Status	Copy of Delivery Sta...	Unicode string [DT_WSTR]	20			
Late_delivery_risk	Copy of Late_deliver...	four-byte signed integer				
Days for shipping (r...	Copy of Days for shi...	four-byte signed integer				
Order Profit Per Order	Copy of Order Profit...	double-precision float ...				
Order City	Copy of Order City	Unicode string [DT_WSTR]	50			
Order Status	Copy of Order Status	Unicode string [DT_WSTR]	20			

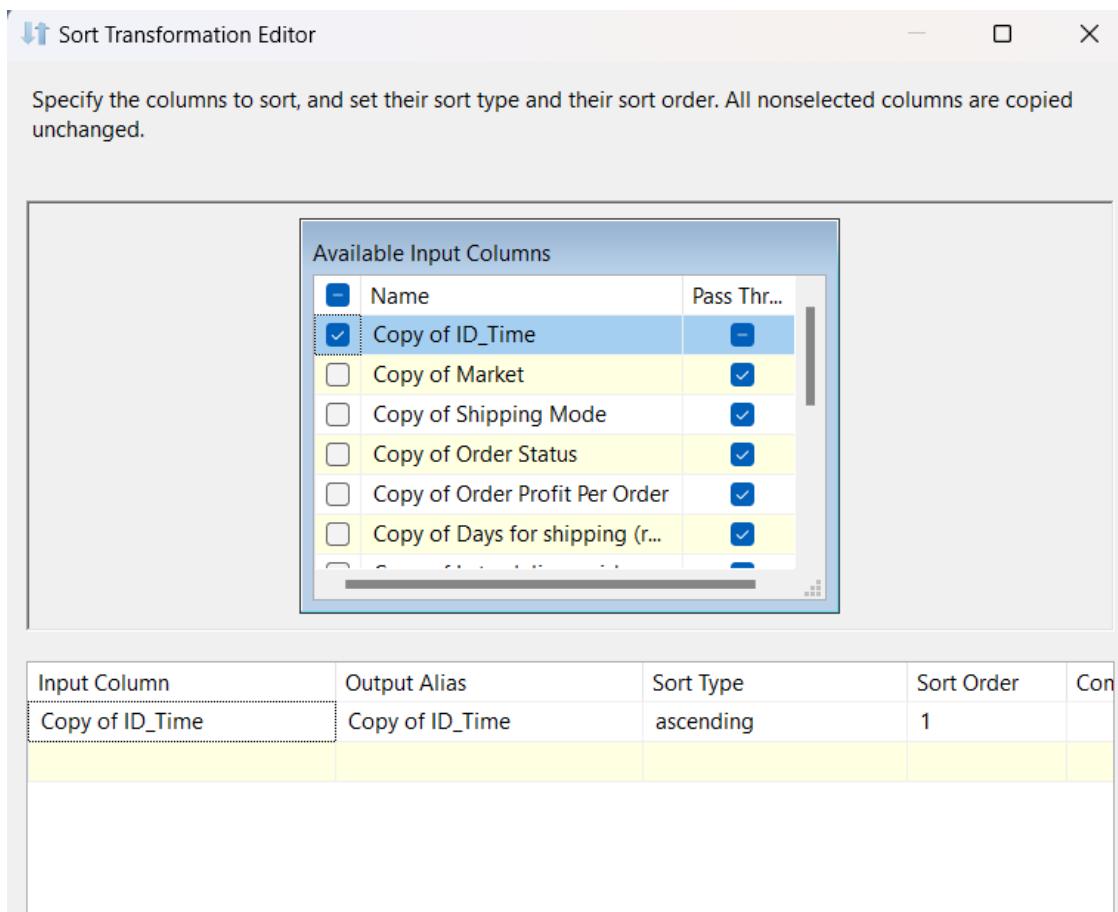
Configure Error Output...      OK      Cancel      Help

- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.

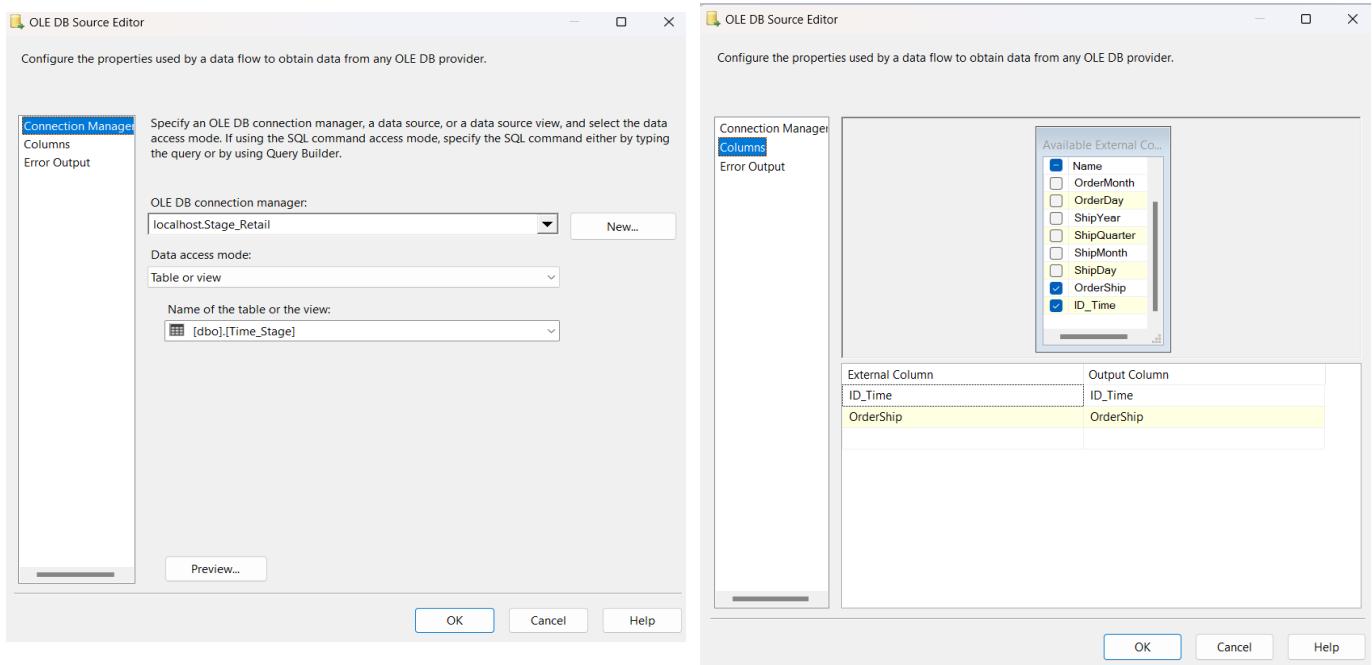


### 3.2.12.1. Kết hợp với Stage Time

- Dùng sort để chỉ định thuộc tính dùng để inner join của bảng được lấy từ file excel nguồn



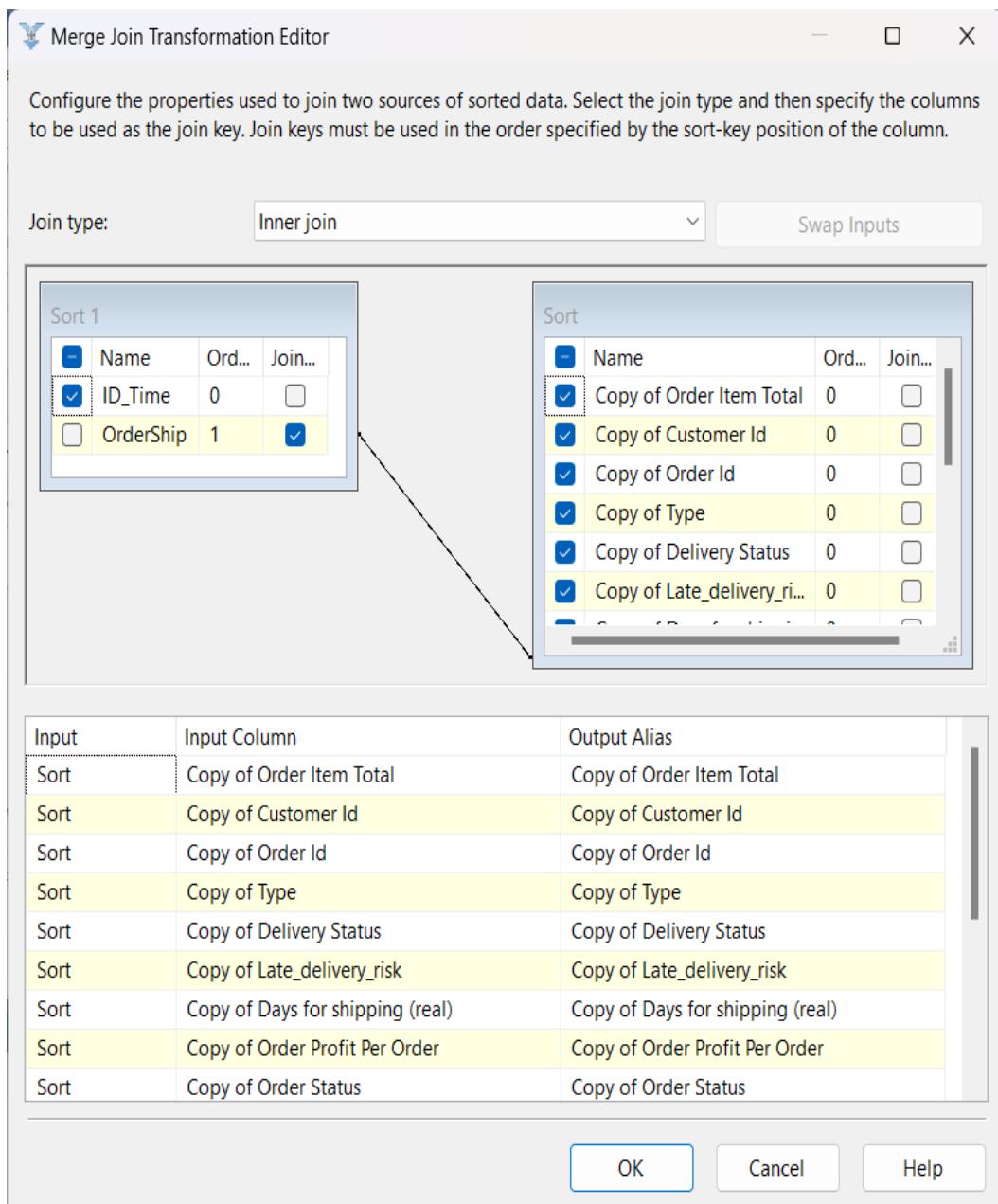
- Chọn nguồn là Stage Time để inner join và chọn các thuộc tính phù hợp



- Dùng sort để chỉ định thuộc tính dùng để inner join của Stage Time

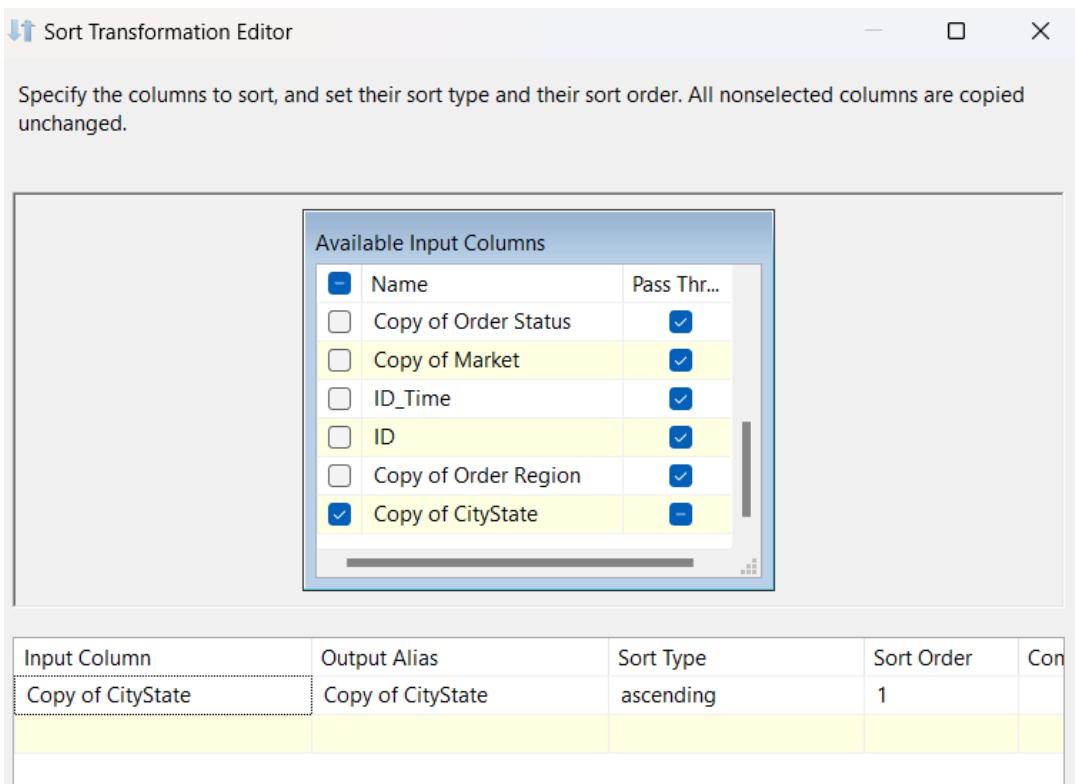
Input Column	Output Alias	Sort Type	Sort Order	Con
OrderShip	OrderShip	ascending	1	

- Sử dụng Merge Join để kết hợp hai bảng với nhau theo thuộc tính đã chọn và chọn các thuộc tính cần thiết cho bước tiếp theo

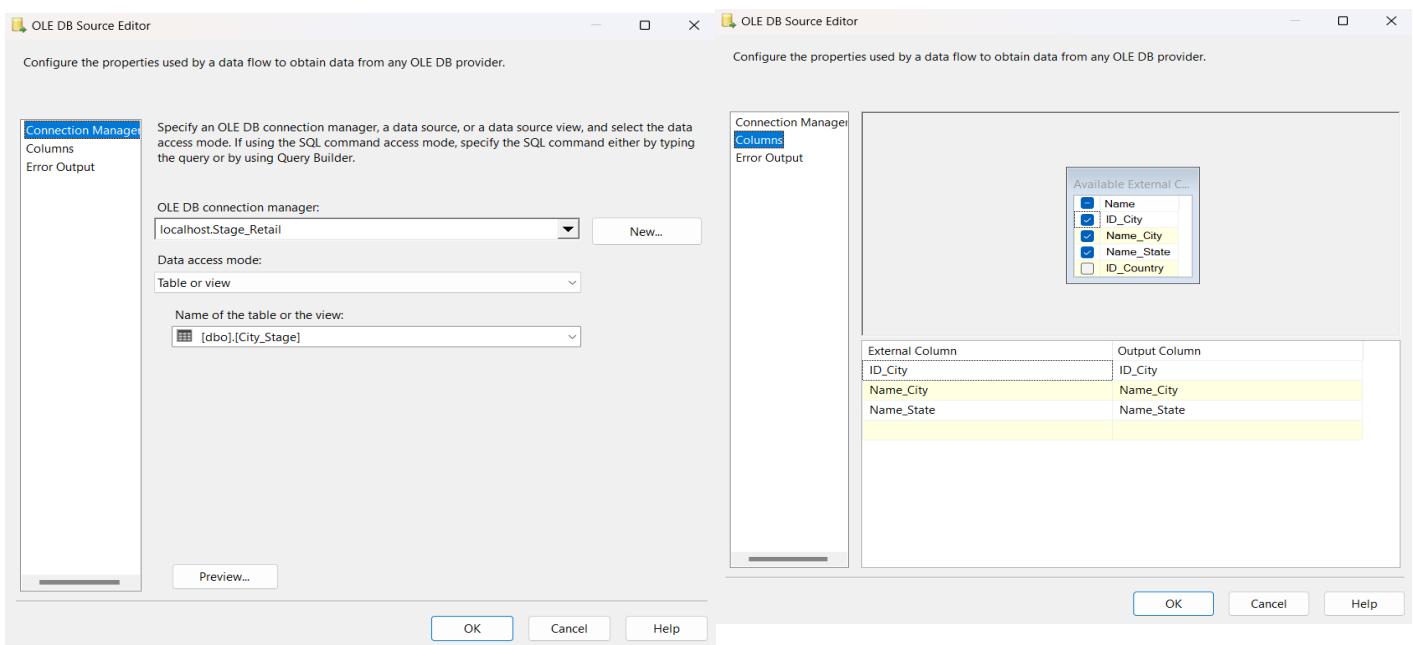


### 3.2.12.2. Kết hợp với Stage City

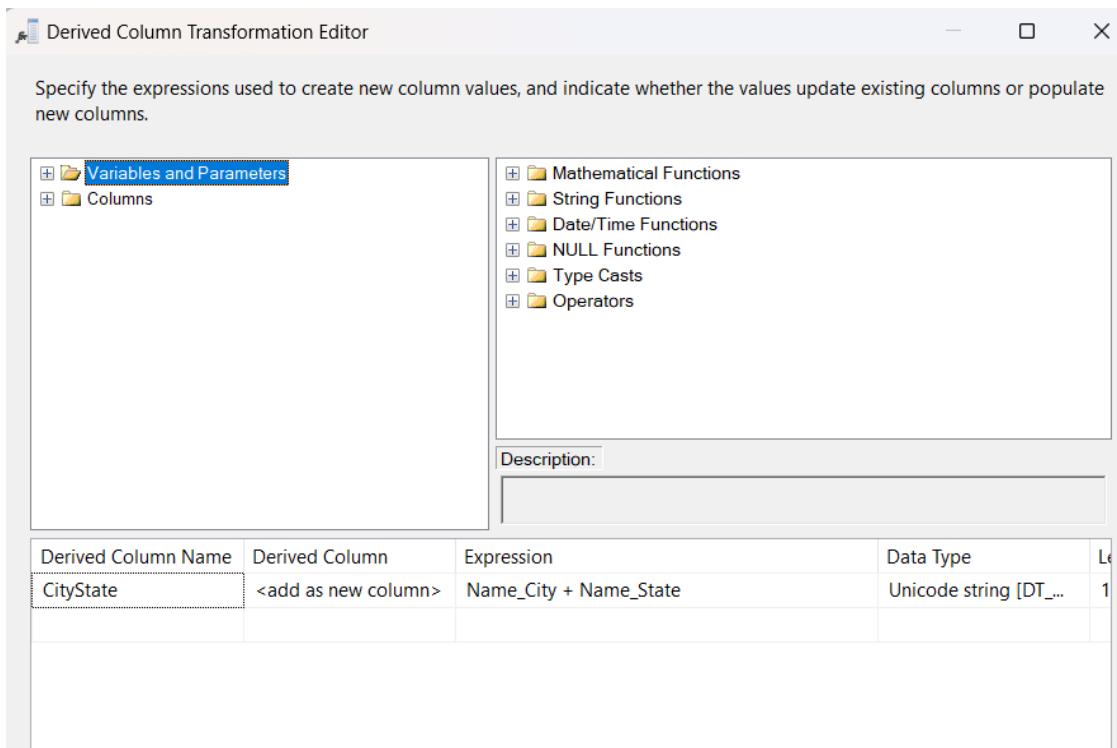
- Dùng sort để chỉ định thuộc tính dùng để inner join của bảng được lấy từ kết quả merge trước.



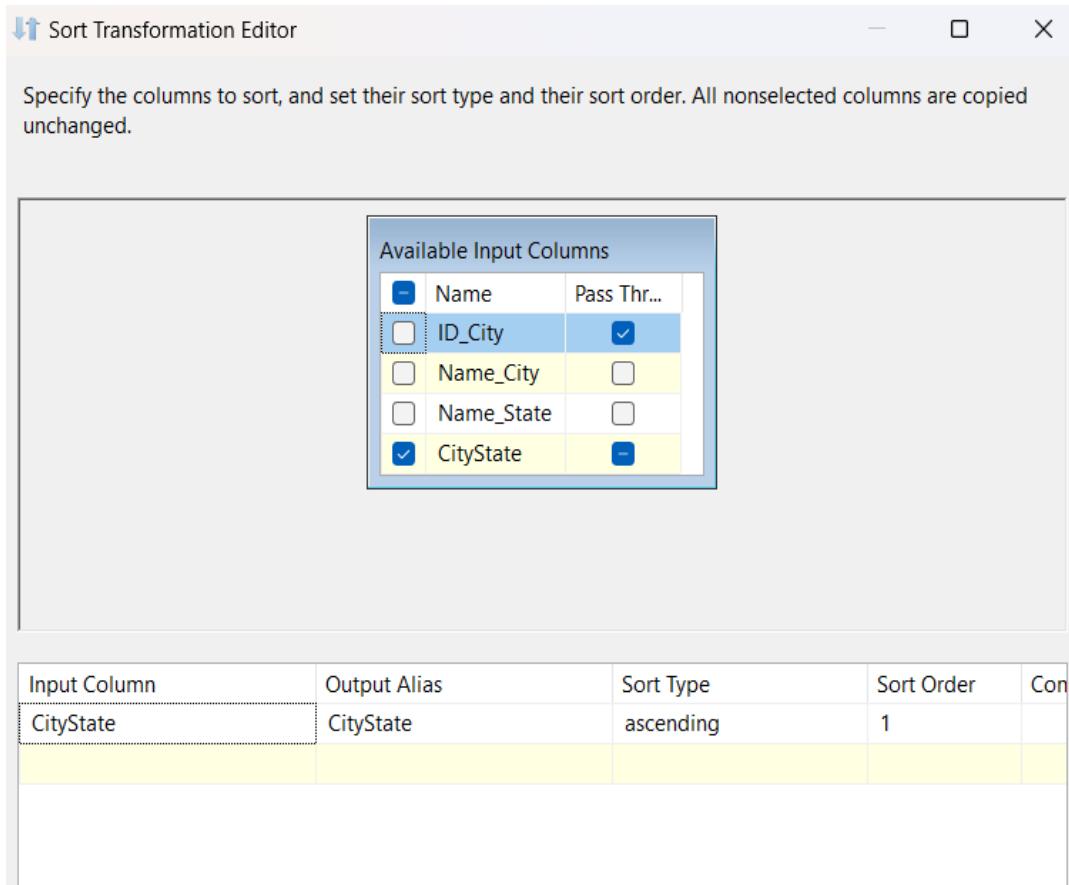
- Chọn nguồn là Stage City để inner join và chọn các thuộc tính phù hợp



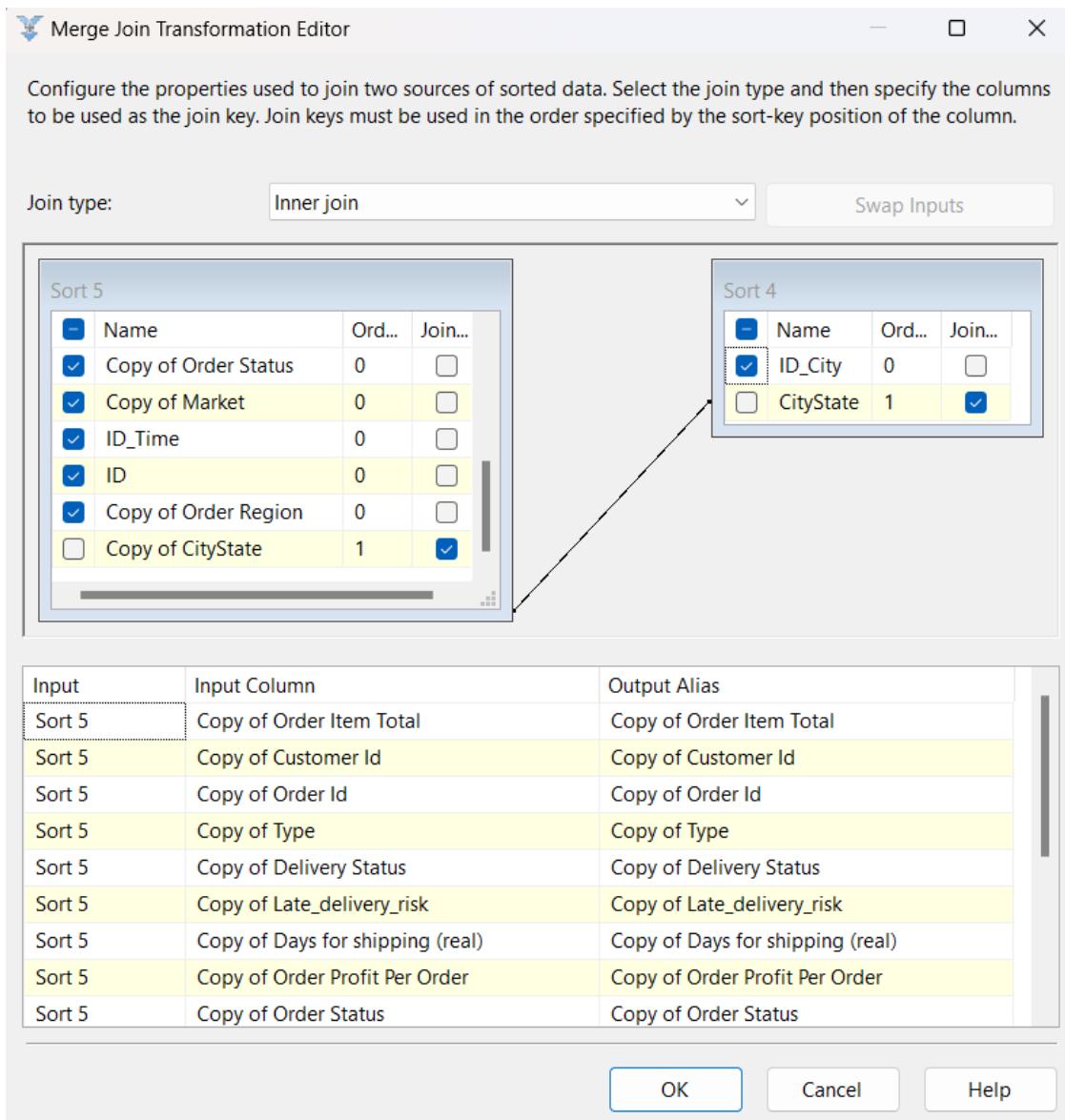
- Dùng Derived Column để thêm cột CityStage vào để làm thuộc tính dùng để inner join



- Dùng sort để chỉ định thuộc tính dùng để inner join của Stage City

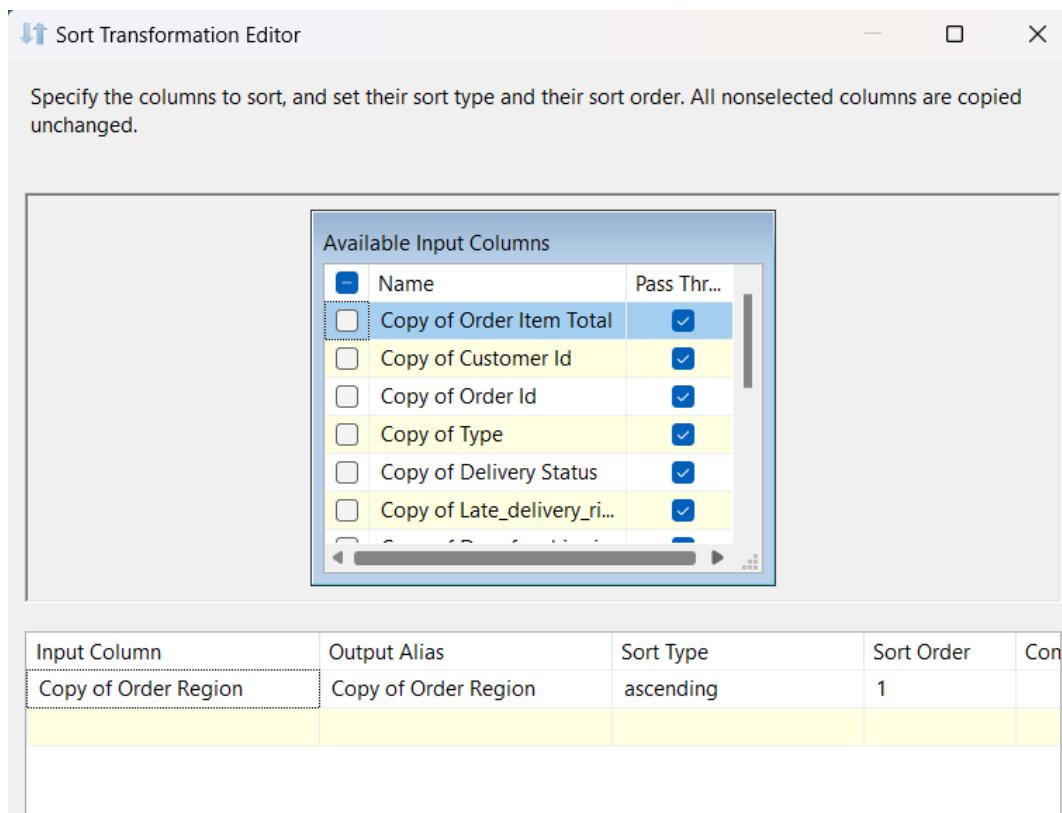


- Sử dụng Merge Join để kết hợp hai bảng với nhau theo thuộc tính đã chọn và chọn các thuộc tính cần thiết cho bước tiếp theo



### 3.2.12.3. Kết hợp với Stage Region

- Dùng sort để chỉ định thuộc tính dùng để inner join của bảng được lấy từ kết quả merge trước.

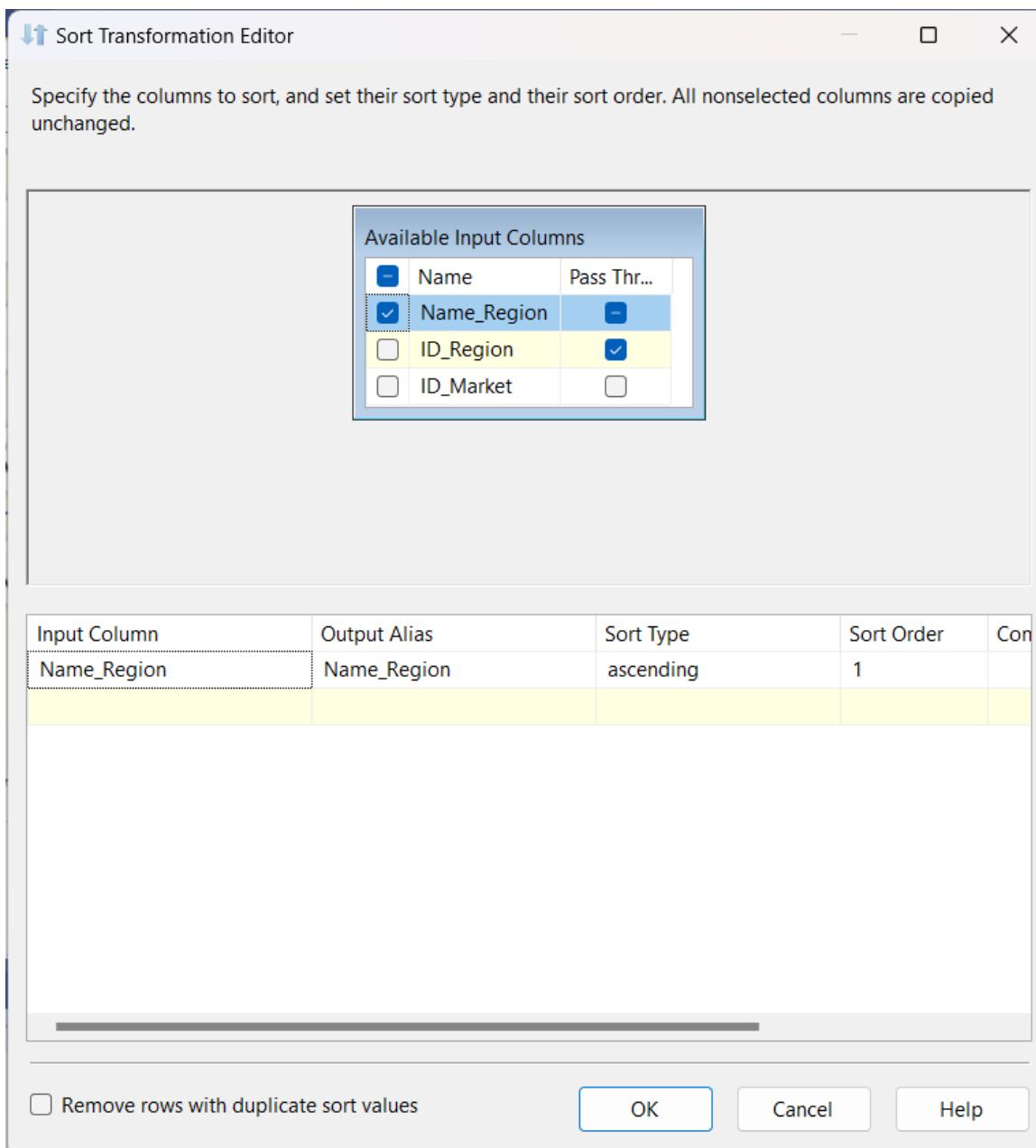


- Chọn nguồn là Stage Region để inner join và chọn các thuộc tính phù hợp

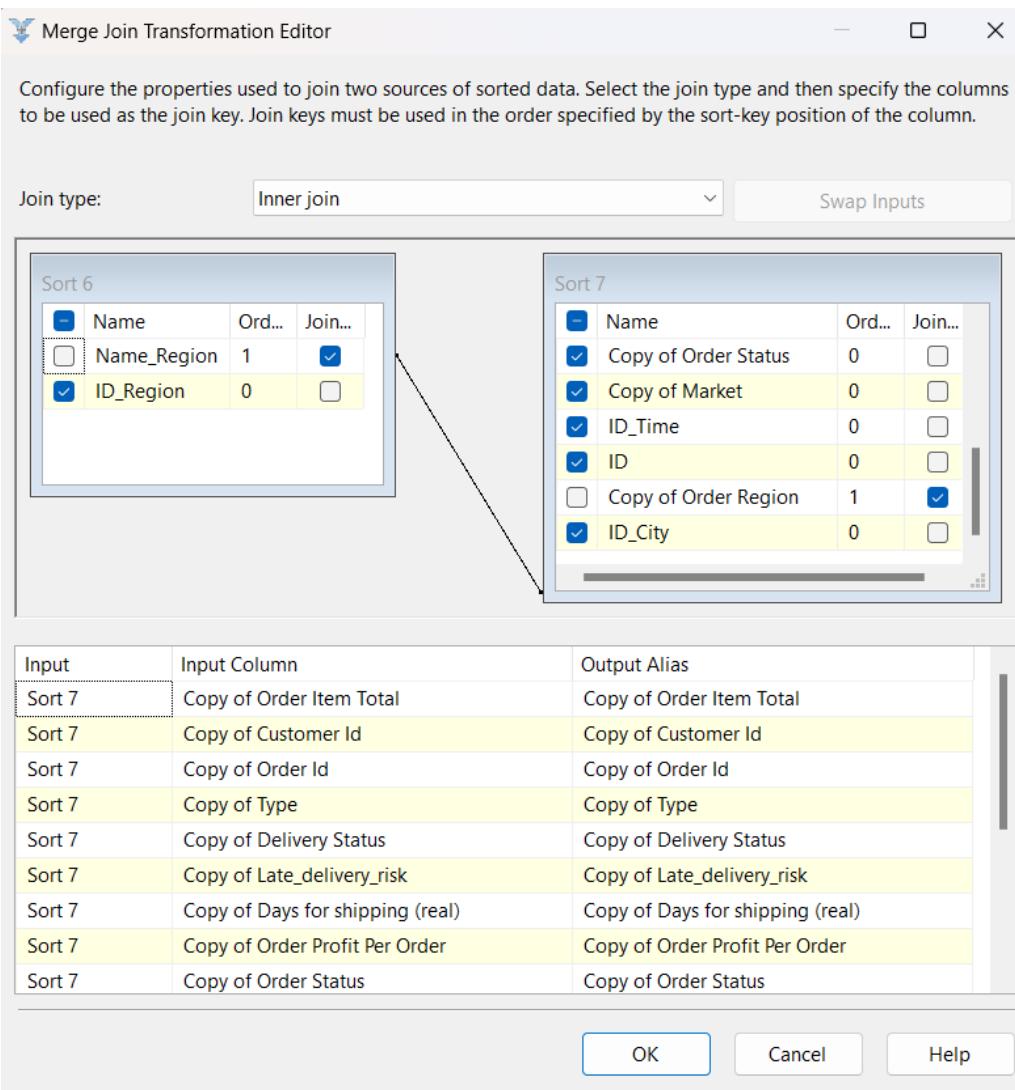
This screenshot displays two windows side-by-side. On the left is the 'Sort Transformation Editor' showing the 'Available Input Columns' list with 'Name', 'Name\_Region', 'ID\_Region', and 'ID\_Market'. The 'Name\_Region' column is selected. A mapping grid below shows 'Name\_Region' mapped to 'Name\_Region' with 'ascending' sort type and sort order 1. At the bottom are buttons for 'OK', 'Cancel', and 'Help', and a checkbox for 'Remove rows with duplicate sort values'. On the right is the 'OLE DB Source Editor' titled 'Configure the properties used by a data flow to obtain data from any OLE DB provider'. It shows a 'Connection Manager' section with 'Columns' selected, and an 'Available External Columns' list with 'Name', 'Name\_Region', and 'ID\_Region'. Below this is a mapping grid:

External Column	Output Column
Name_Region	Name_Region
ID_Region	ID_Region
ID_Market	ID_Market

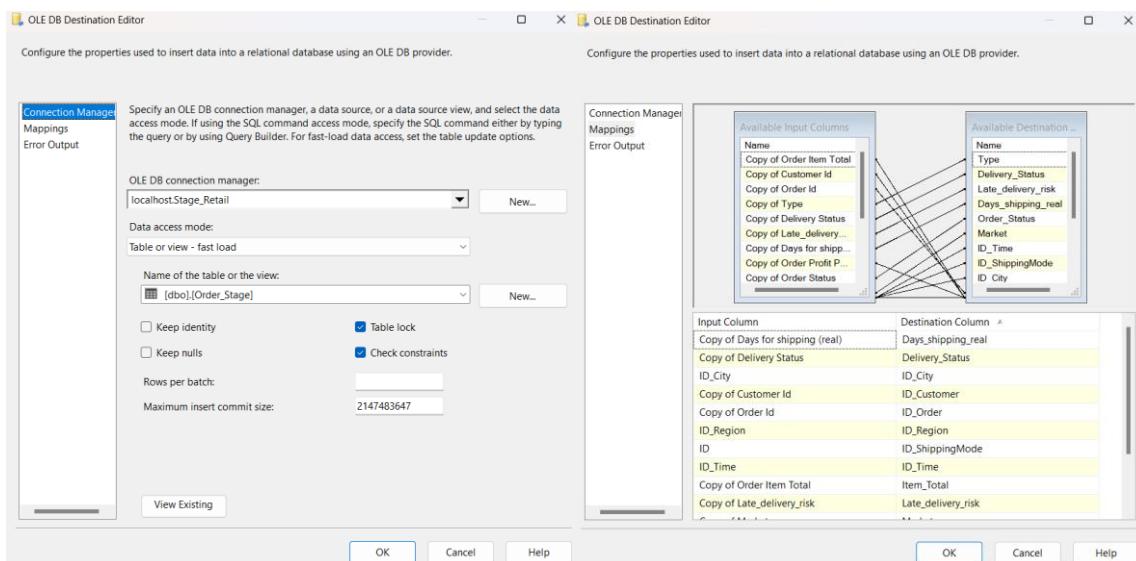
- Dùng sort để chỉ định thuộc tính dùng để inner join của Stage Region



- Sử dụng Merge Join để kết hợp hai bảng với nhau theo thuộc tính đã chọn và chọn các thuộc tính cần thiết cho bước tiếp theo

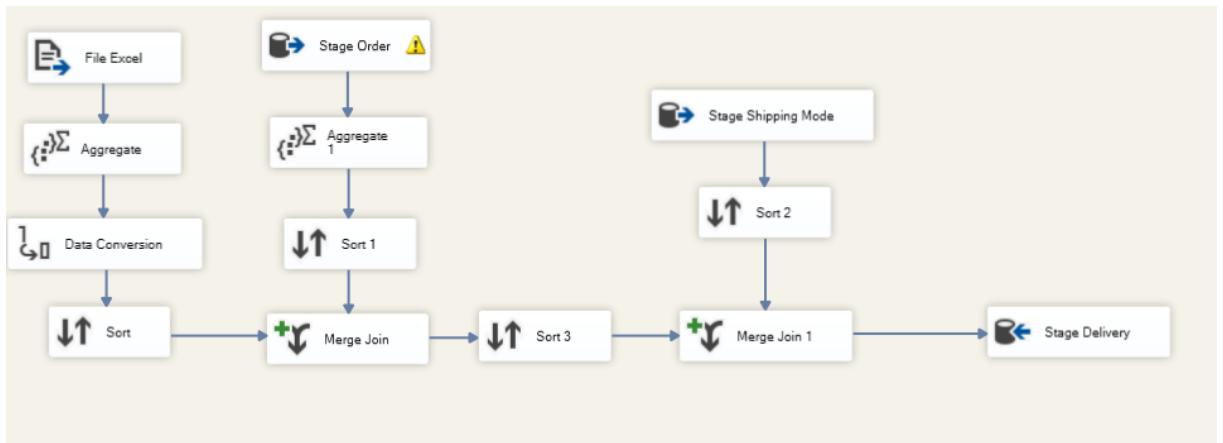


- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn.

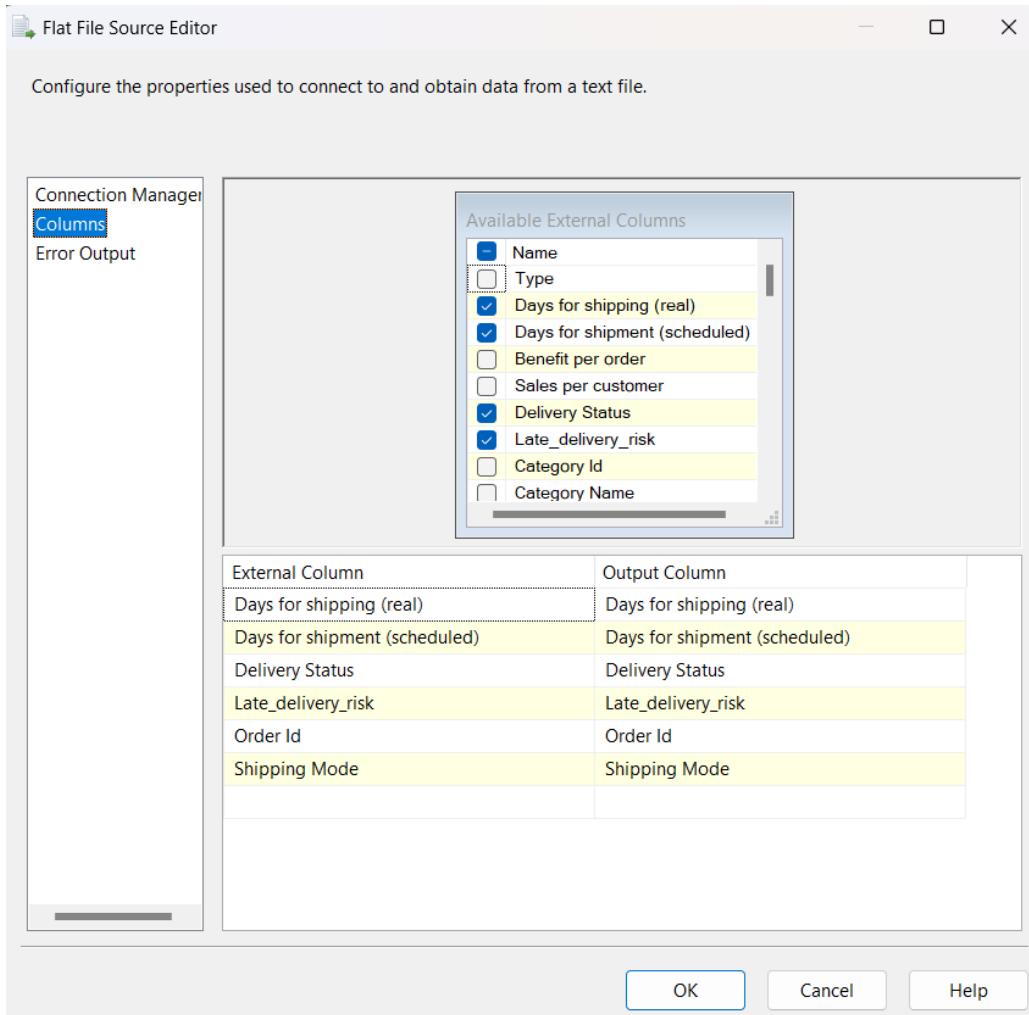


### 3.2.13. Stage Delivery

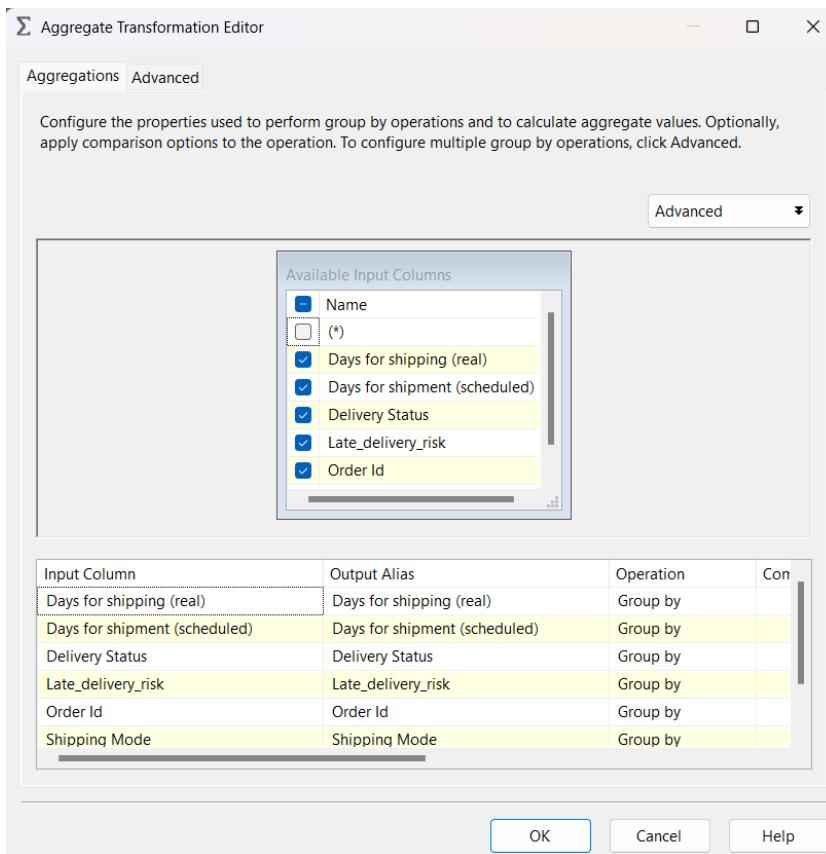
- Cách thành phần của Stage Delivery



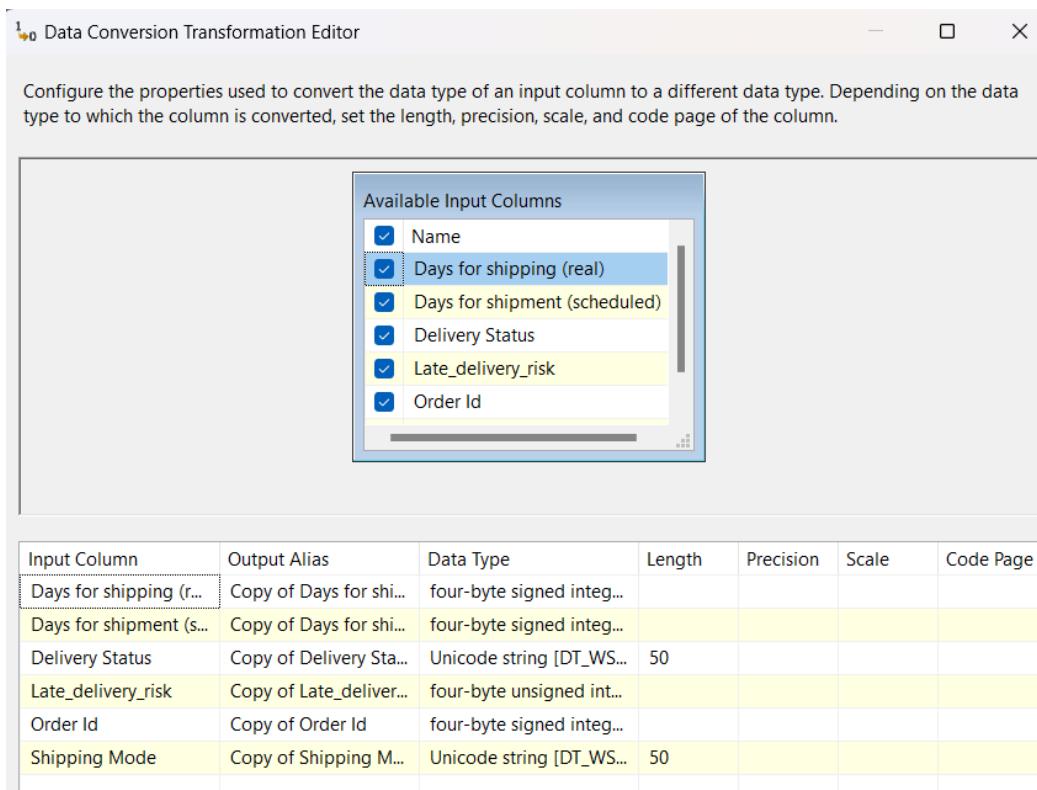
- Từ file excel nguồn chọn ra cột thích hợp để đưa vào cơ sở dữ liệu Stage.



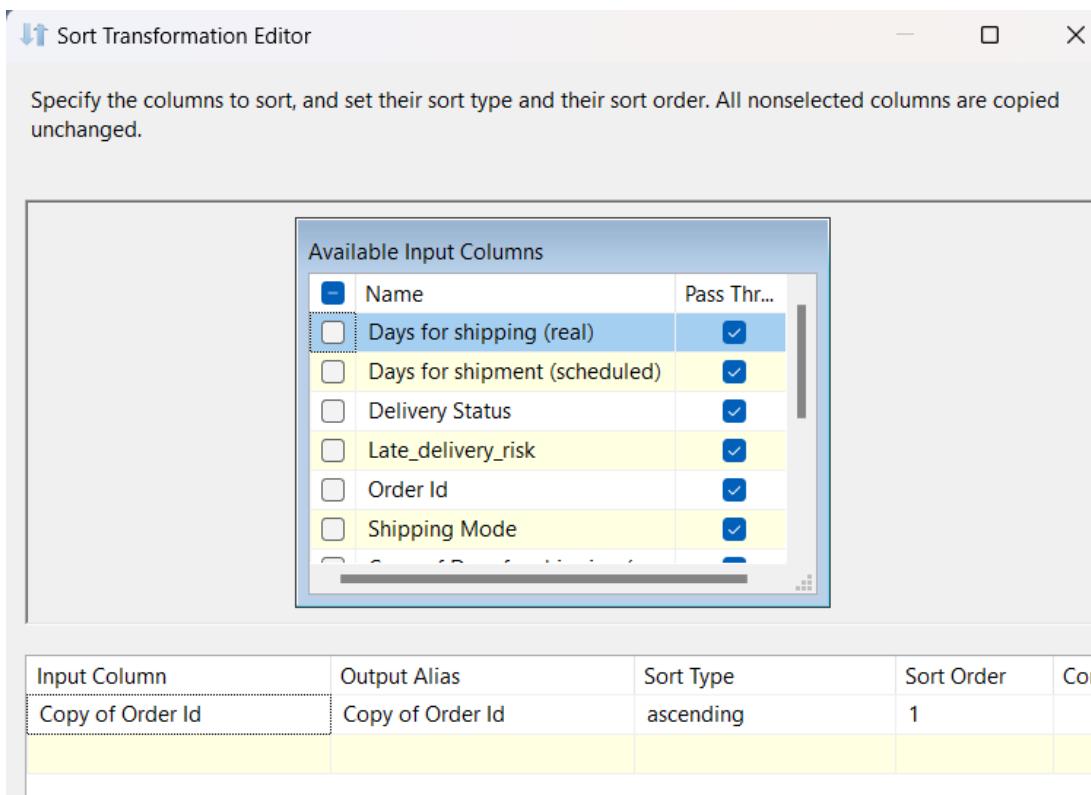
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



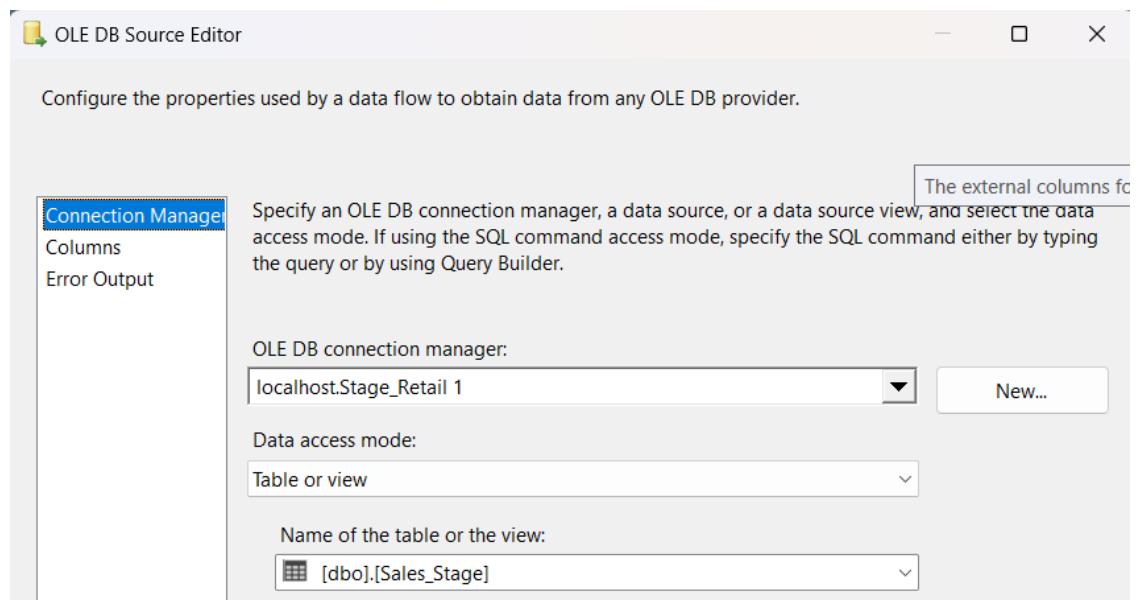
- Chuyển đổi kiểu dữ liệu của các thuộc tính sao cho phù hợp để tiếp tục các bước tiếp theo

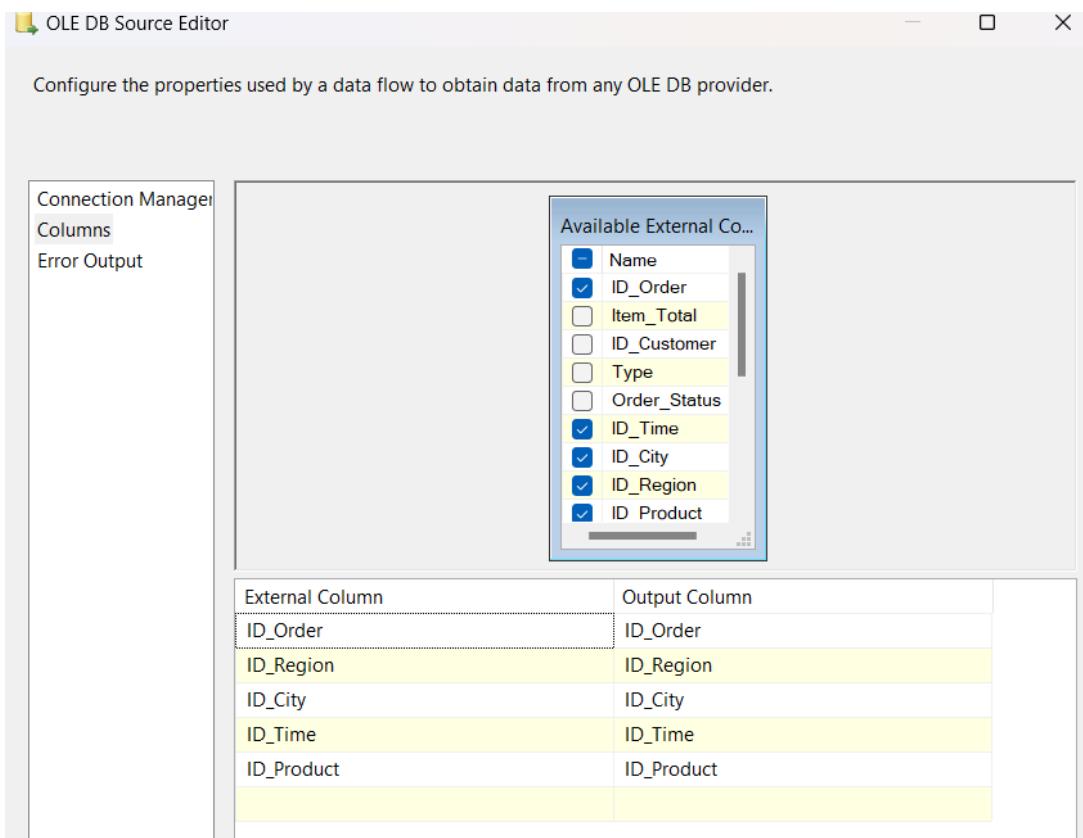


- Dùng sort để chỉ định thuộc tính dùng để inner join

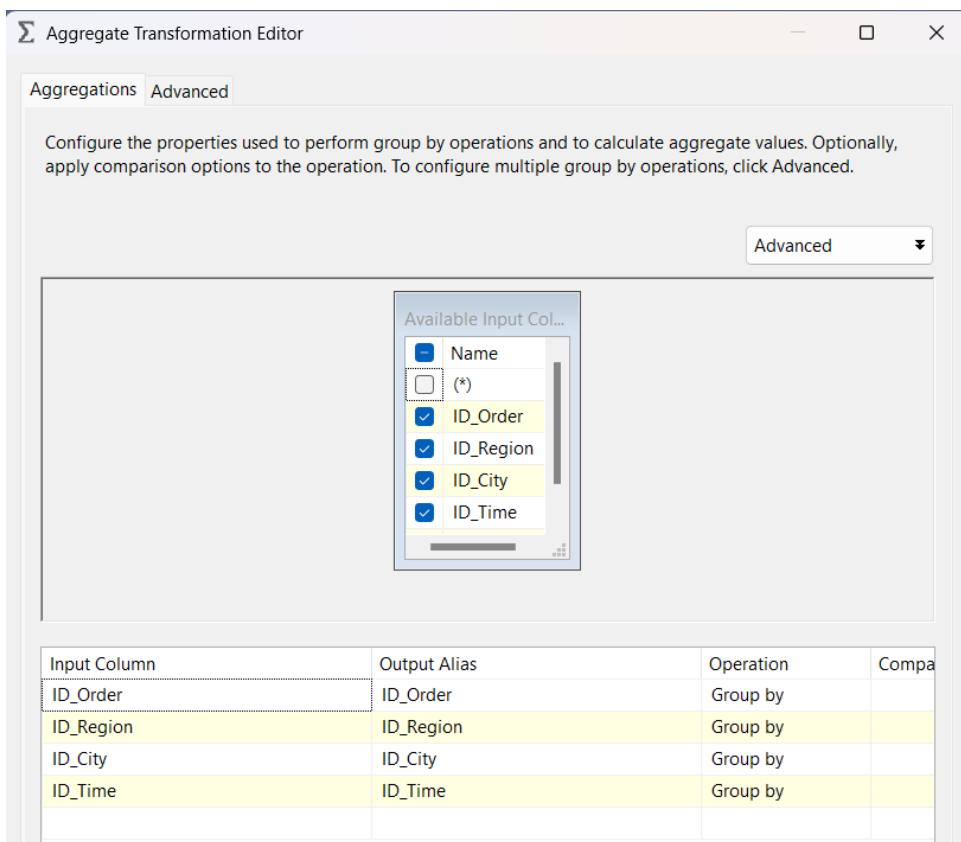


- Chọn nguồn là Stage Sales để inner join

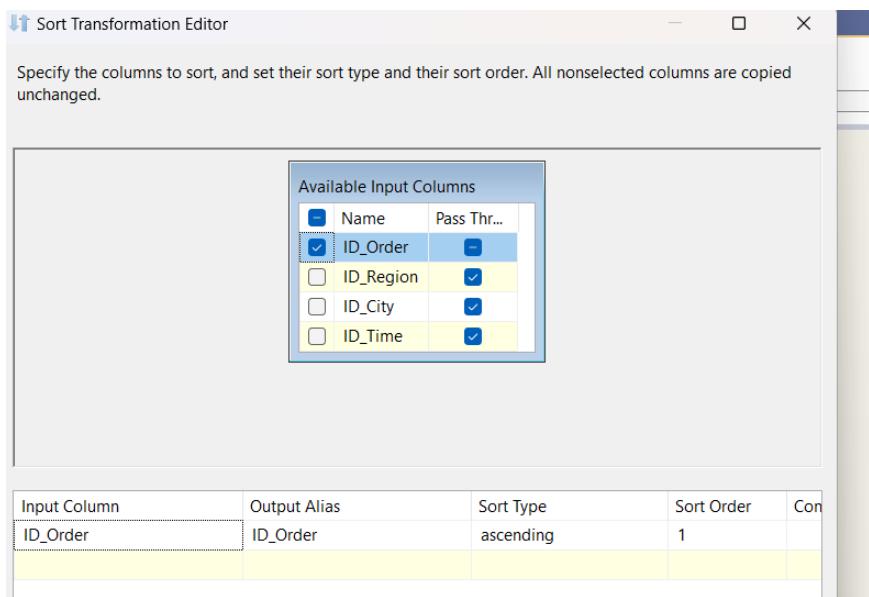




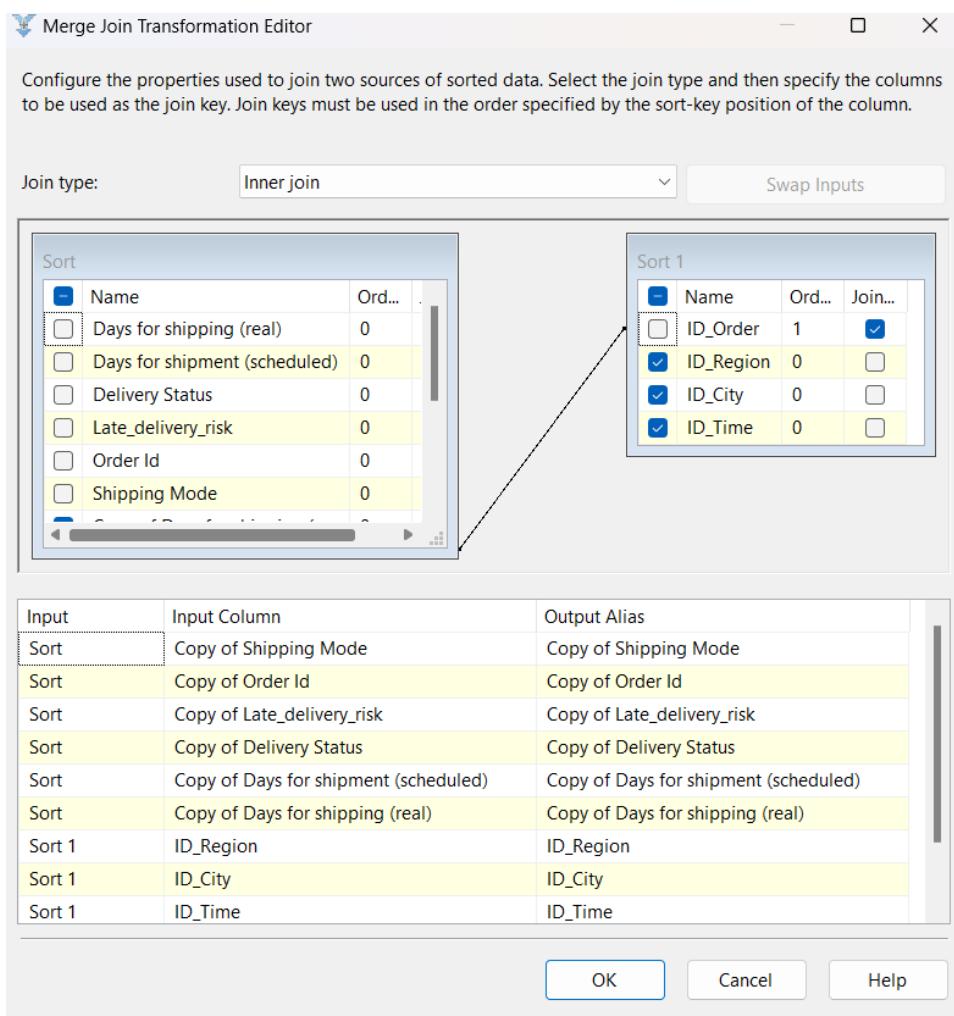
- Sử dụng Aggregate để group by các thuộc tính lại với nhau chỉ còn các giá trị duy nhất.



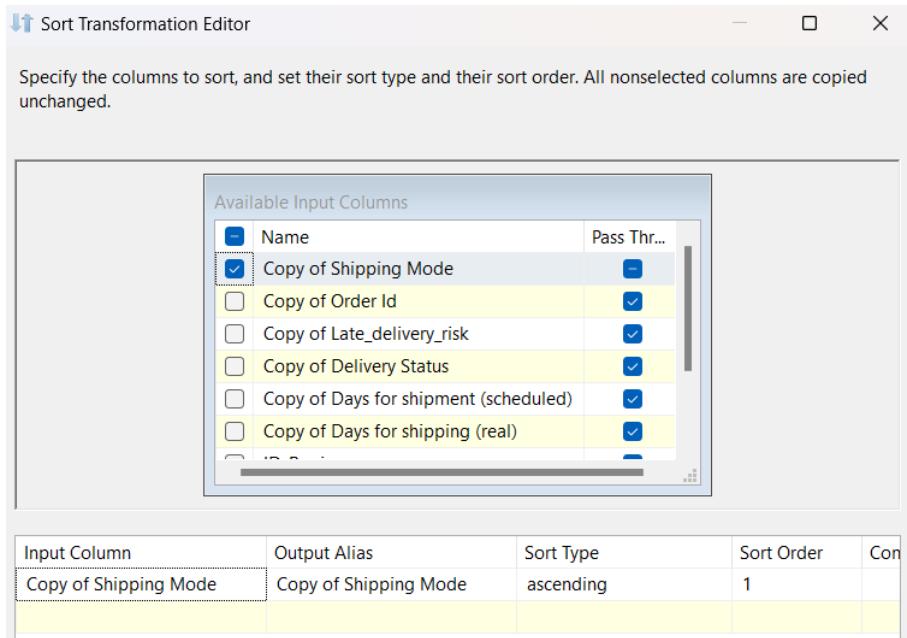
- Dùng sort để chỉ định thuộc tính của Stage Sales dùng để inner join



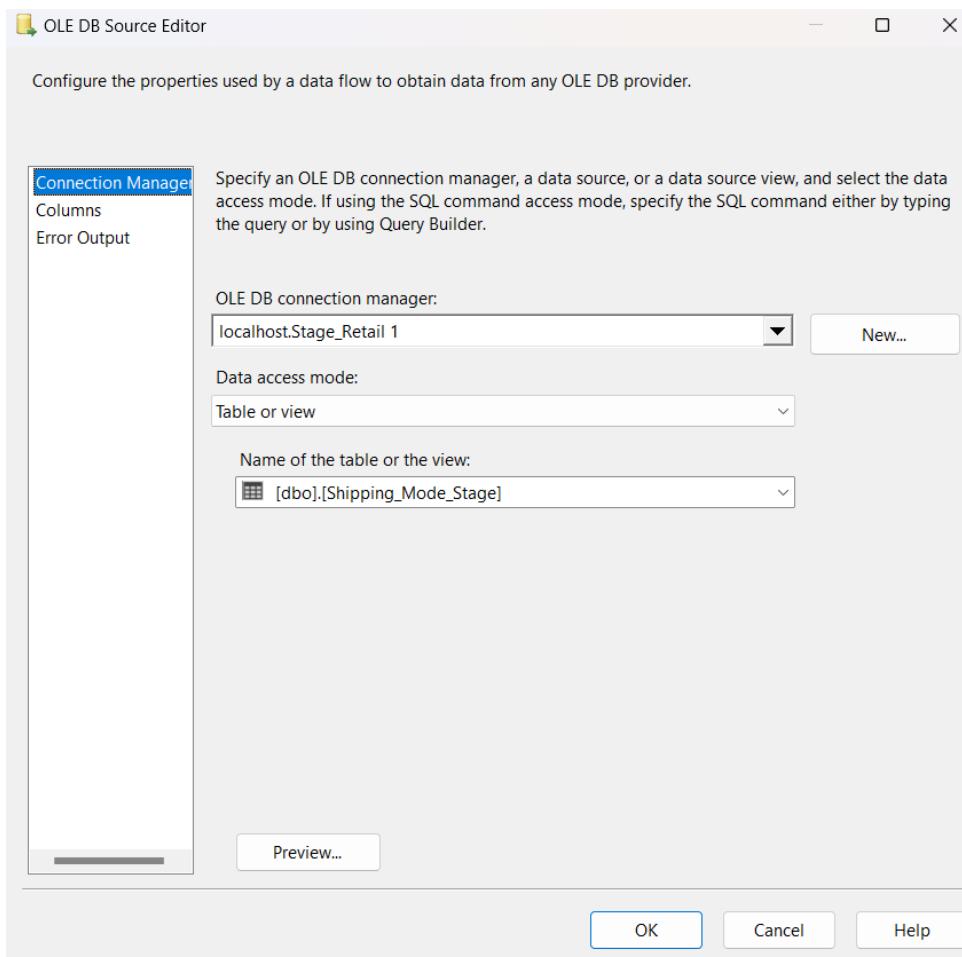
- Sử dụng Merge Join để kết hợp hai bảng lại với nhau để lấy được các thuộc tính cần thiết khác



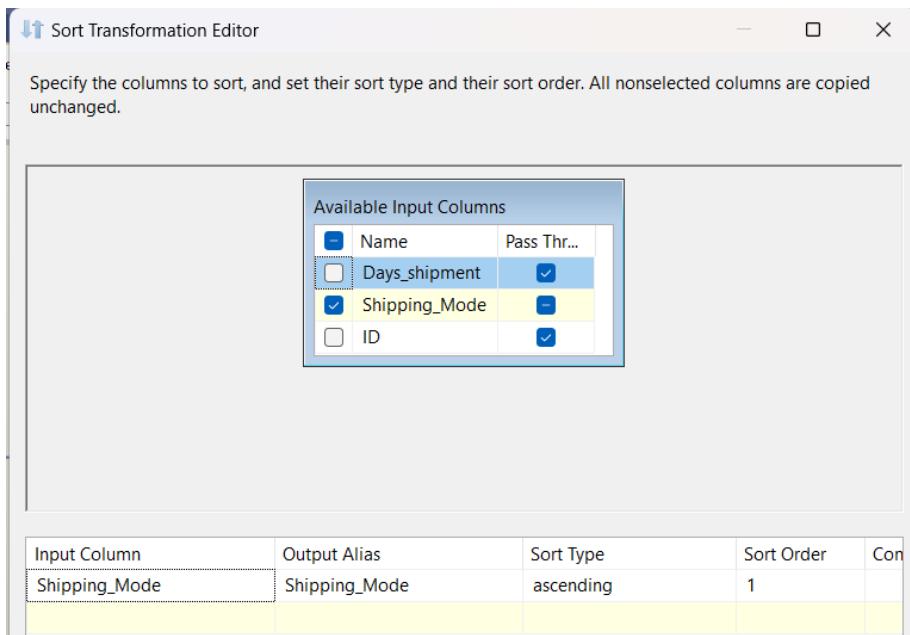
- Tiếp tục dùng Sort để chỉ định thuộc tính của kết quả vừa merge dùng để inner join



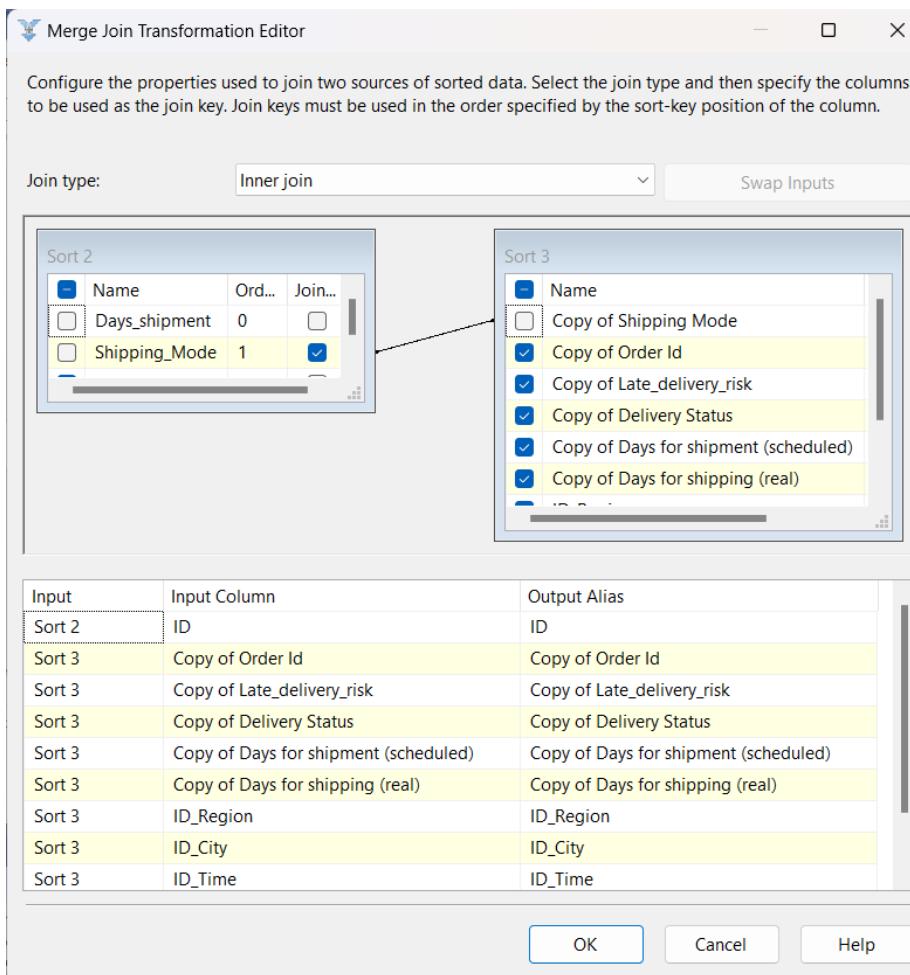
- Chọn nguồn là Stage Shipping Mode để tiếp tục inner join



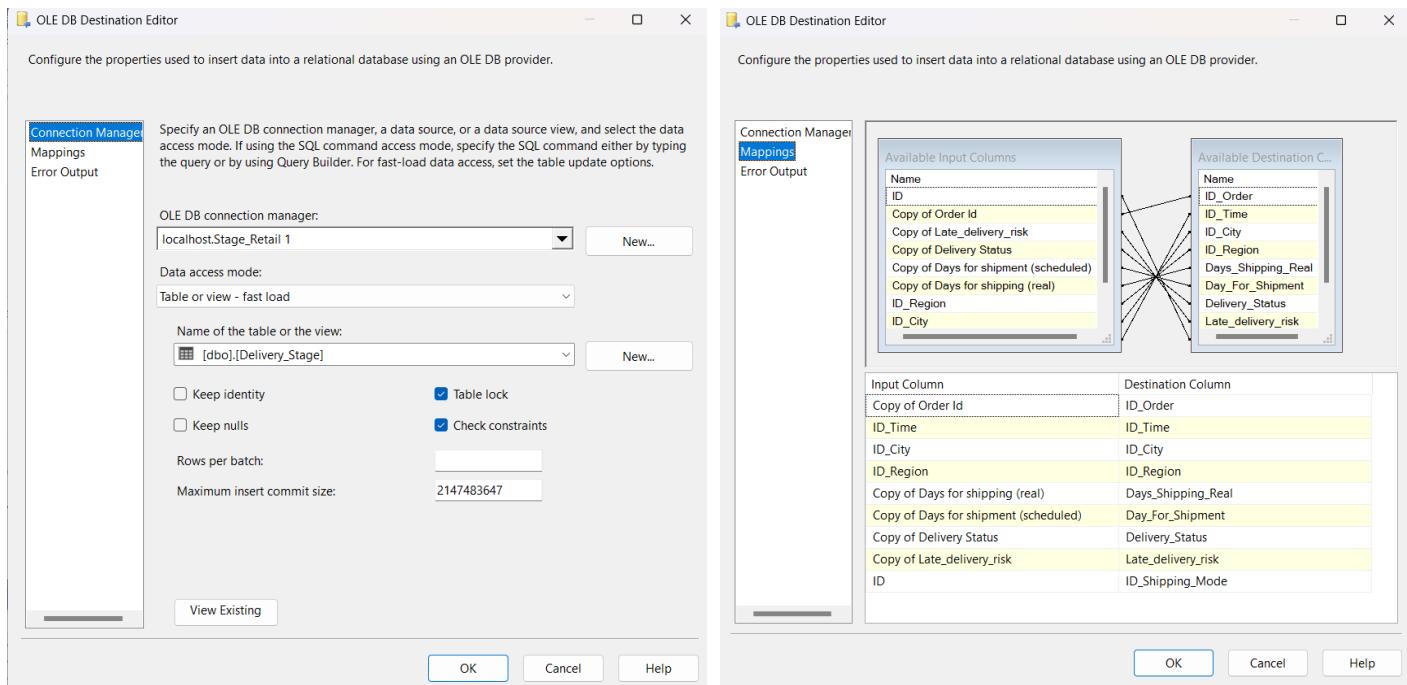
- Dùng sort để chỉ định thuộc tính của Stage Shipping Mode dùng để inner join



- Sử dụng Merge Join để kết hợp hai bảng lại với nhau để lấy được các thuộc tính cần thiết khác



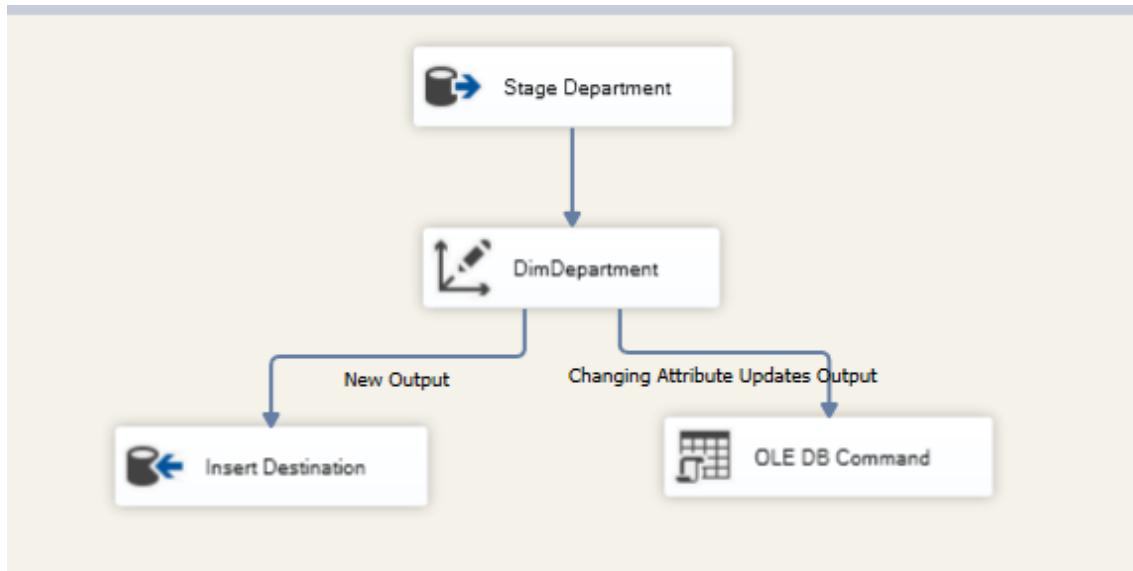
- Ở OLE DB Destination tạo kết nối với database và chọn bảng phù hợp và Mapping các thuộc tính với các cột của bảng vừa chọn



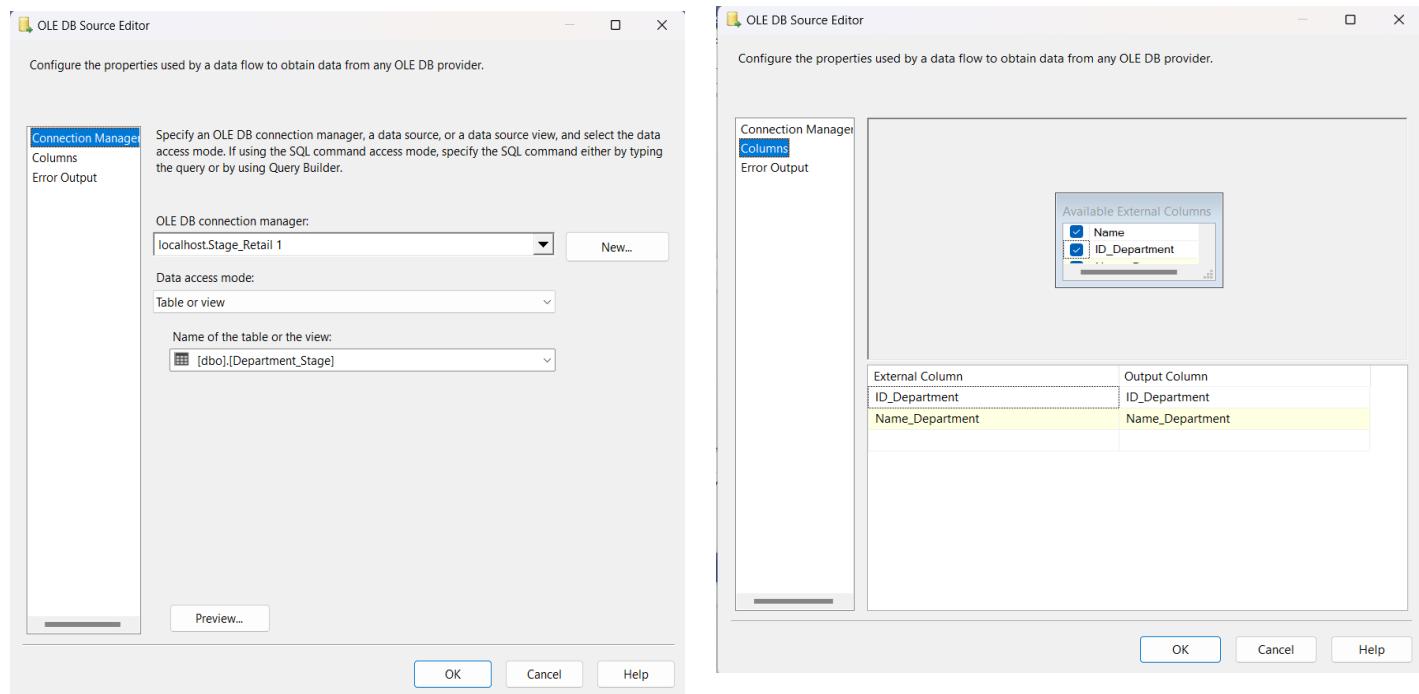
### 3.3. Quá trình đổ dữ liệu từ cơ sở dữ liệu Stage vào kho dữ liệu

#### 3.3.1. Dim Department

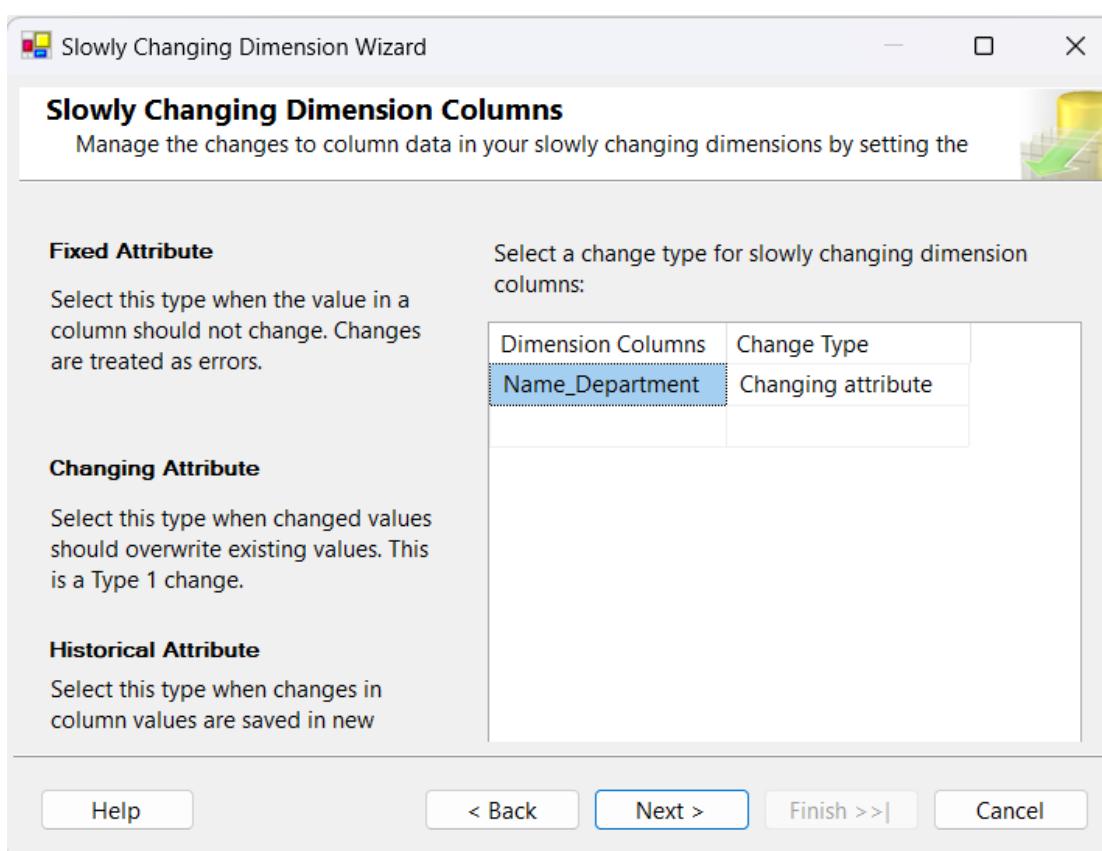
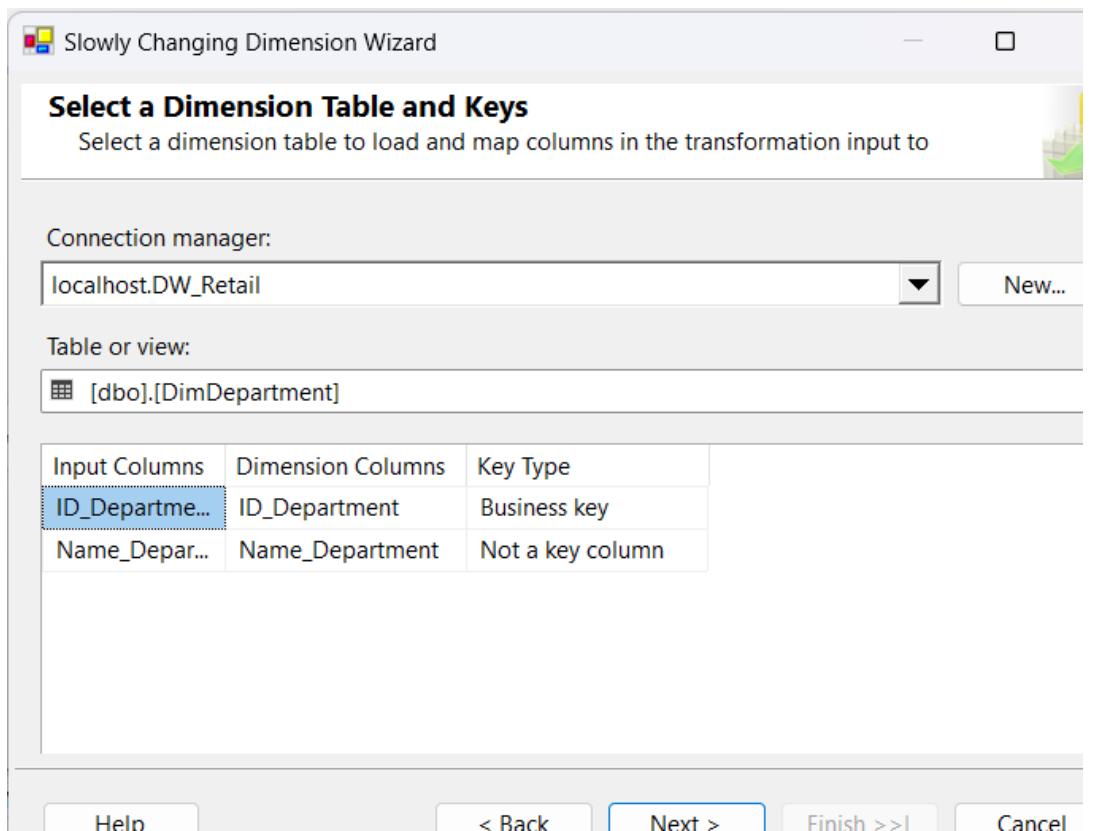
- Các thành phần của DimDepartment



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Department.Stage và chọn các cột thích hợp



- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute

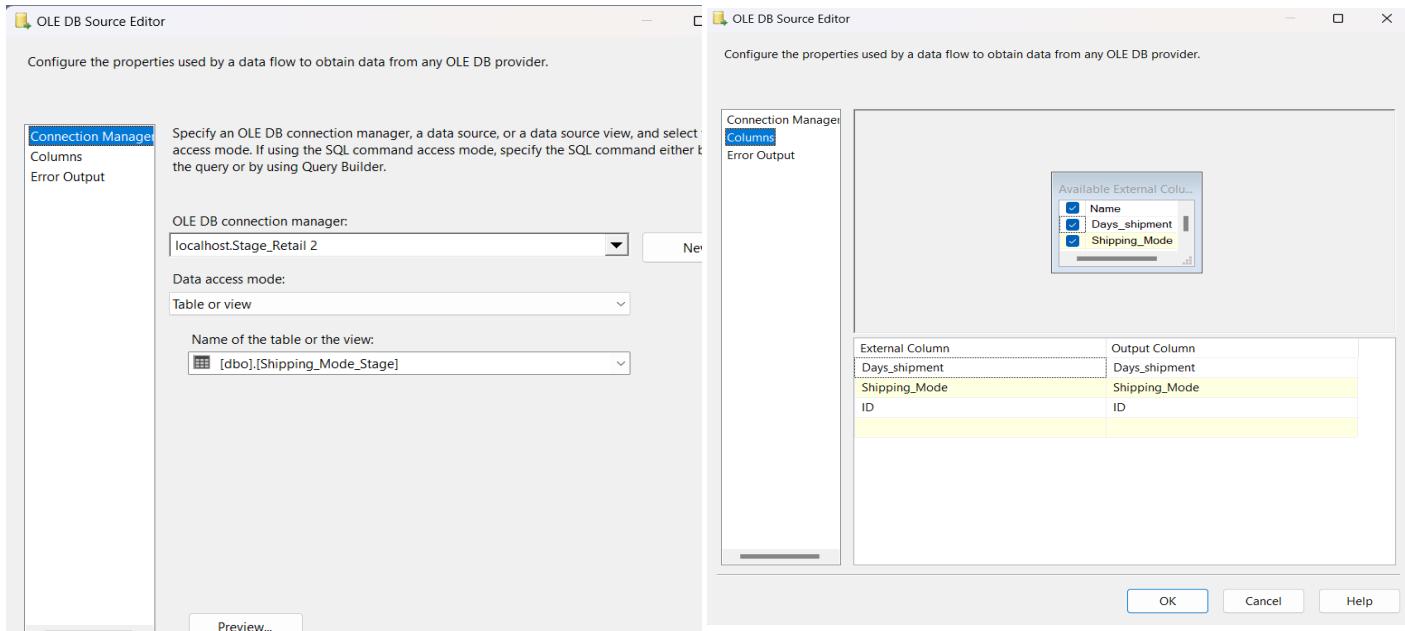


### 3.3.2. Dim ShippingMode

- Các thành phần của DimShippingMode



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng ShippingMode\_Stage và chọn các cột thích hợp.



- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute

**Slowly Changing Dimension Wizard**

### Select a Dimension Table and Keys

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

Connection manager:

localhost.DW\_Retail

Table or view:

[dbo].[DimShippingMode]

Input Columns	Dimension Columns	Key Type
Days_shipme...	Days_Shipment	Not a key column
ID	ID_ShippingMode	Business key
Shipping_Mo...	Shipping_Mode	Not a key column

Help      < Back      Next >      Finish >>      Cancel

**Slowly Changing Dimension Wizard**

### Slowly Changing Dimension Columns

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

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

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

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

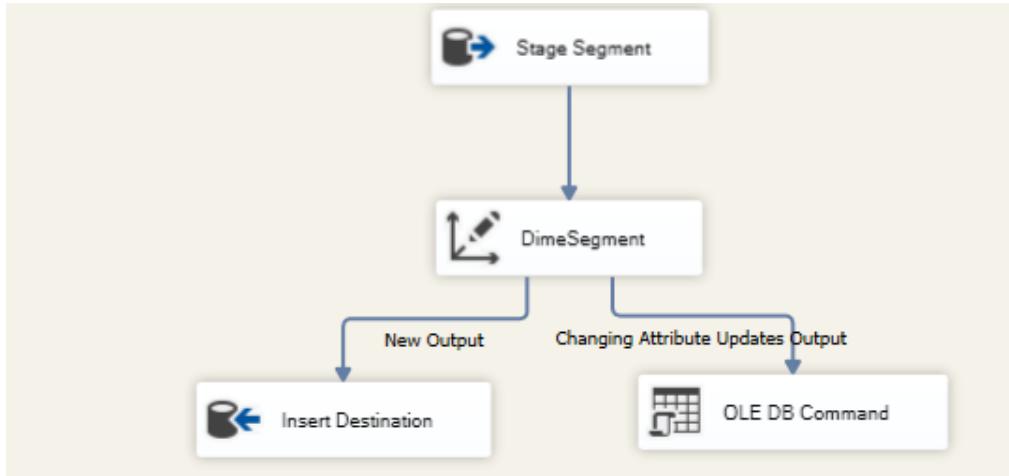
Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
Days_Shipment	Changing a...
Shipping_Mode	Changing a...

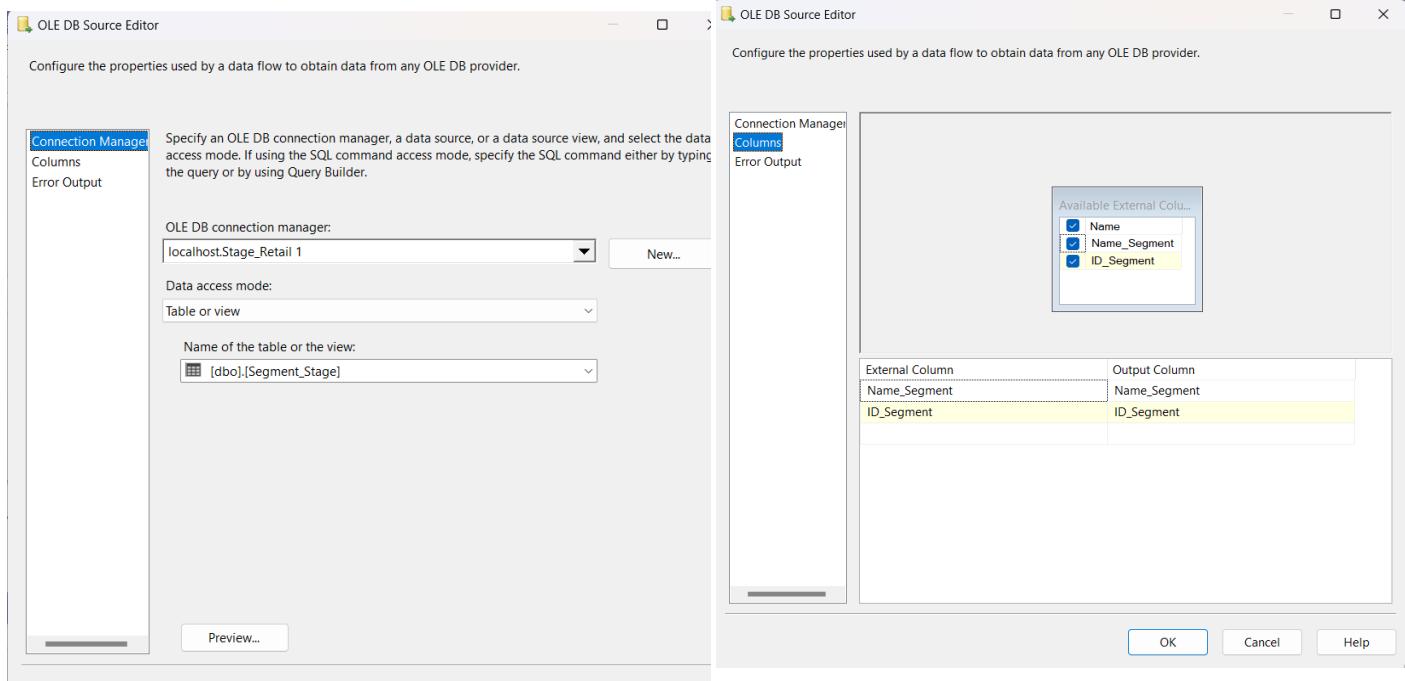
Help      < Back      Next >      Finish >>      Cancel

### 3.3.3. Dim Segment

- Các thành phần của DimSegment



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng DimSegment\_Stage và chọn các cột thích hợp.



- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute

Slowly Changing Dimension Wizard

## Select a Dimension Table and Keys

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

Connection manager:

localhost.DW\_Retail

Table or view:

[dbo].[DimSegment]

Input Columns	Dimension Columns	Key Type
ID_Segment	ID_Segment	Business key
Name_Segm...	Name_Segment	Not a key column

Help < Back Next > Finish >> Cancel

Slowly Changing Dimension Wizard

## Slowly Changing Dimension Columns

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

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

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

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

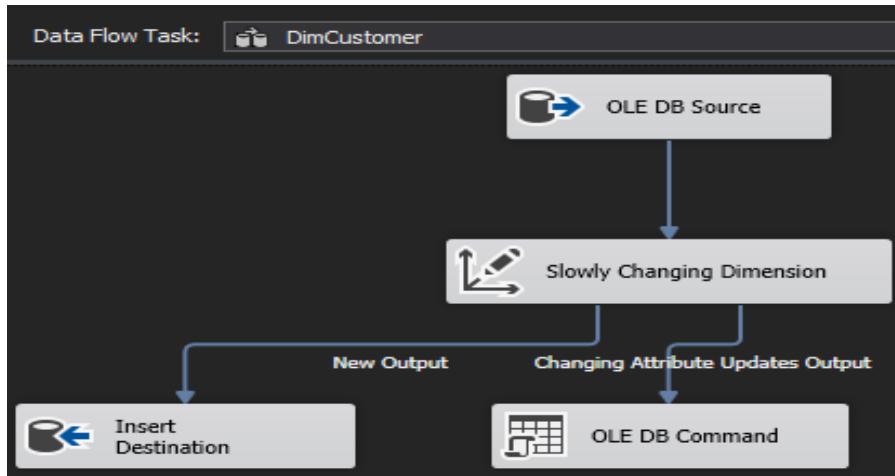
Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
Name_Segment	Changing attribute

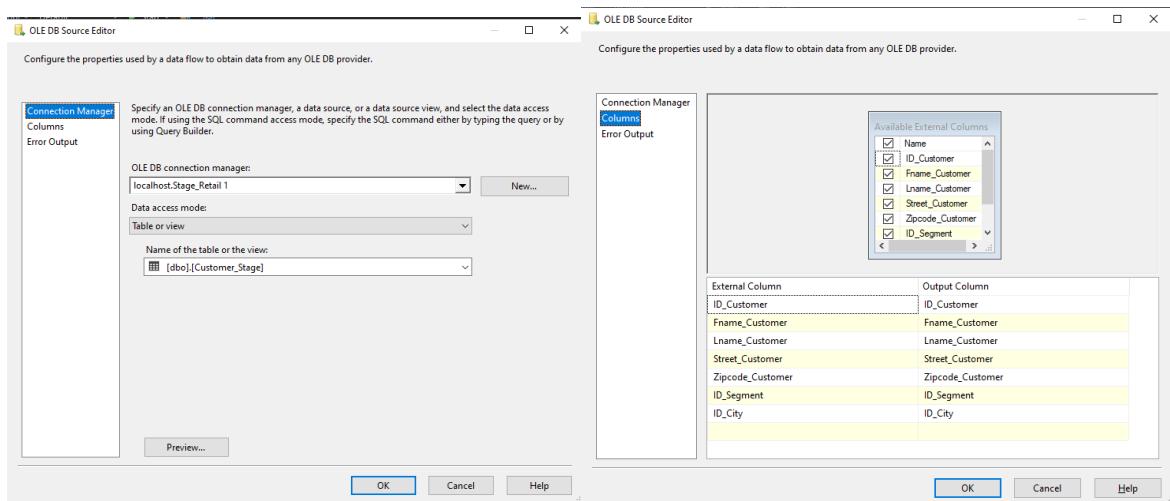
Help < Back Next > Finish >> Cancel

### 3.3.4. Dim Customer

- Các thành phần của DimCustomer



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Customer\_Stage và chọn các cột thích hợp.



- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute

Slowly Changing Dimension Wizard

Select a Dimension Table and Keys

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

Connection manager: localhost.DW\_Retail

Table or view: [dbo].[DimCustomer]

Input Columns	Dimension Columns	Key Type
Fname_Cust...	Fname_Customer	Not a key column
ID_City	ID_City	Not a key column
ID_Customer	ID_Customer	Business key
ID_Segment	ID_Segment	Not a key column
Lname_Cust...	Lname_Customer	Not a key column
Street_Custo...	Street_Customer	Not a key column
Zipcode_Cus...	Zipcode_Customer	Not a key column

**Slowly Changing Dimension Columns**

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

Dimension Columns	Change Type
Fname_Customer	Changing a...
ID_City	Changing a...
ID_Segment	Changing a...
Lname_Customer	Changing a...
Street_Customer	Changing a...
Zipcode_Customer	Changing a...

**Fixed Attribute**

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

**Changing Attribute**

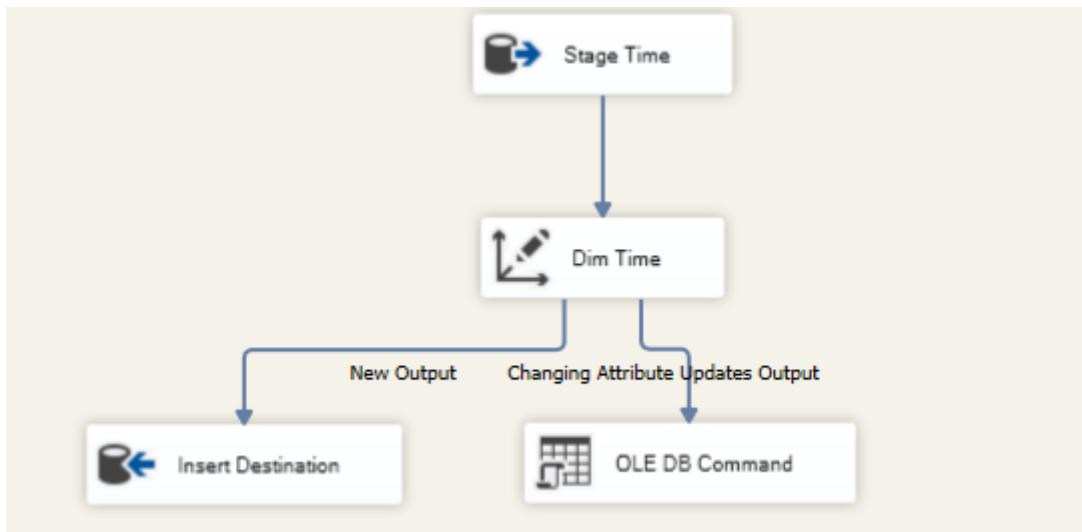
Select this type when changed values should overwrite existing values. This is a Type 1 change.

**Historical Attribute**

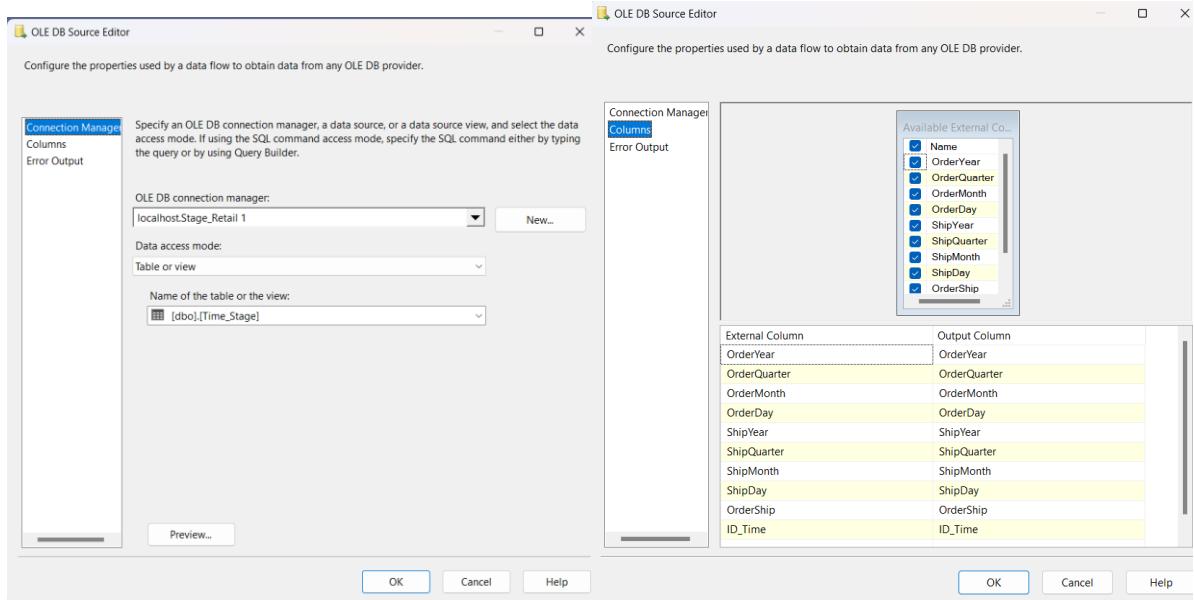
Select this type when changes in column values are saved in new records. Previous values are saved in records marked as outdated. This is a Type 2 change.

### 3.3.5. Dim Time

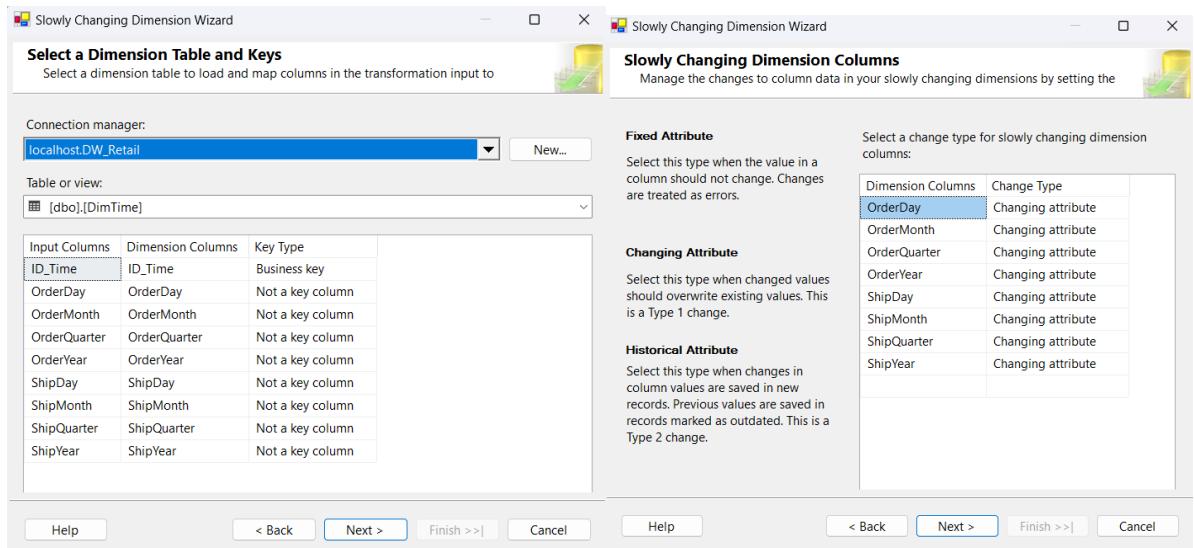
- Các thành phần của DimTime



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Time\_Stage và chọn các cột thích hợp.

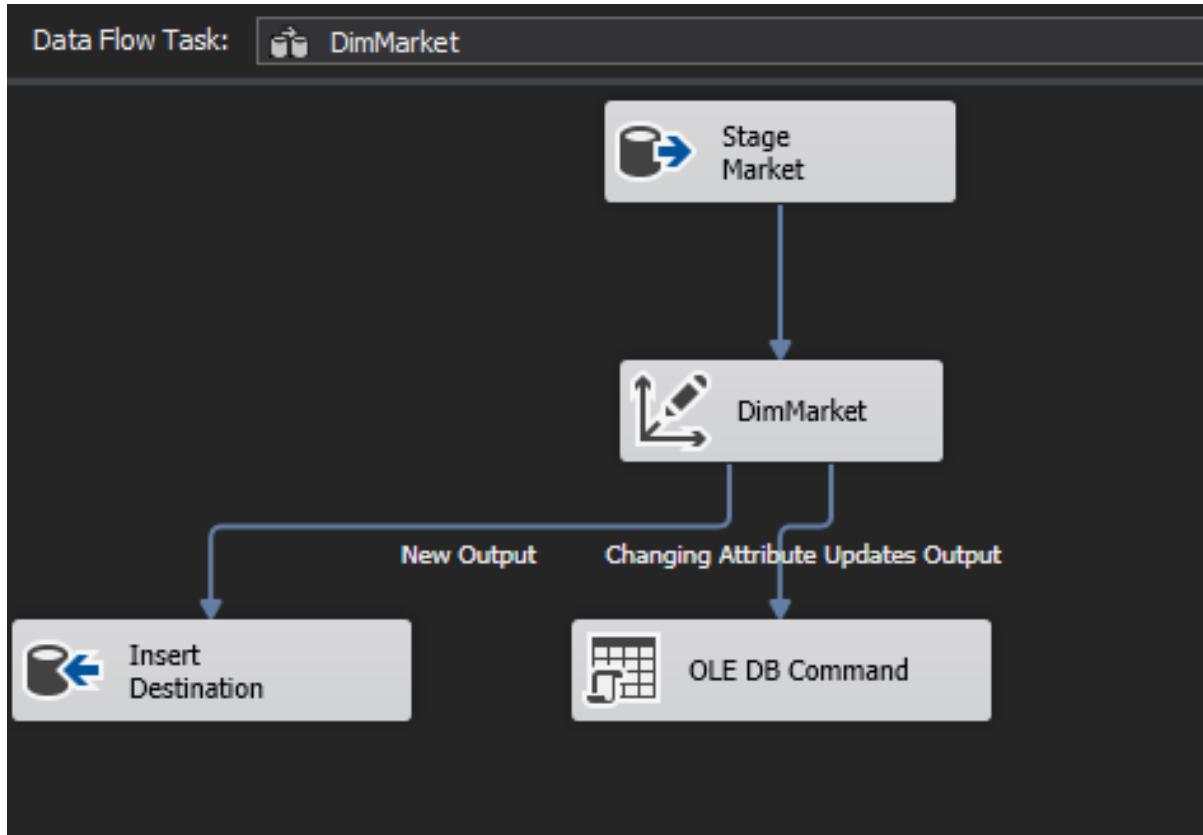


- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu.  
Chọn business key. Thuộc tính còn lại chọn changing attribute

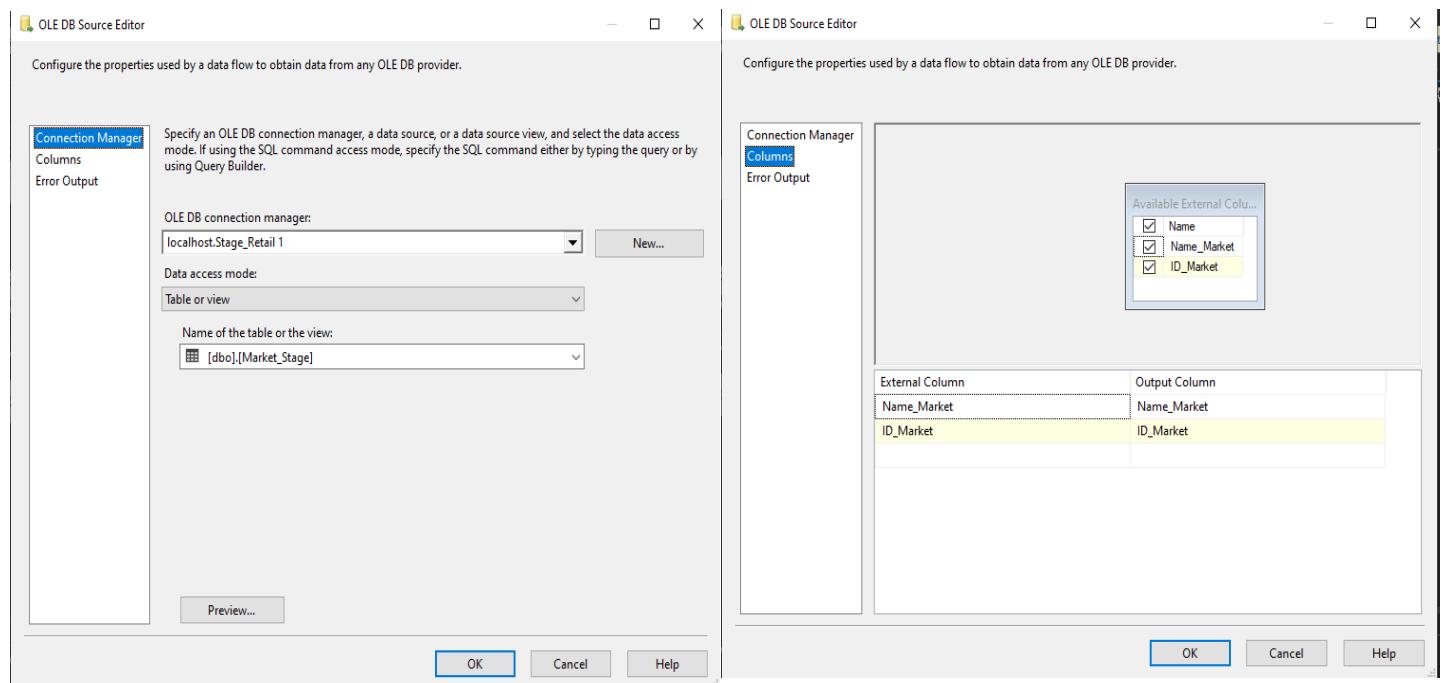


### 3.3.6. Dim Market

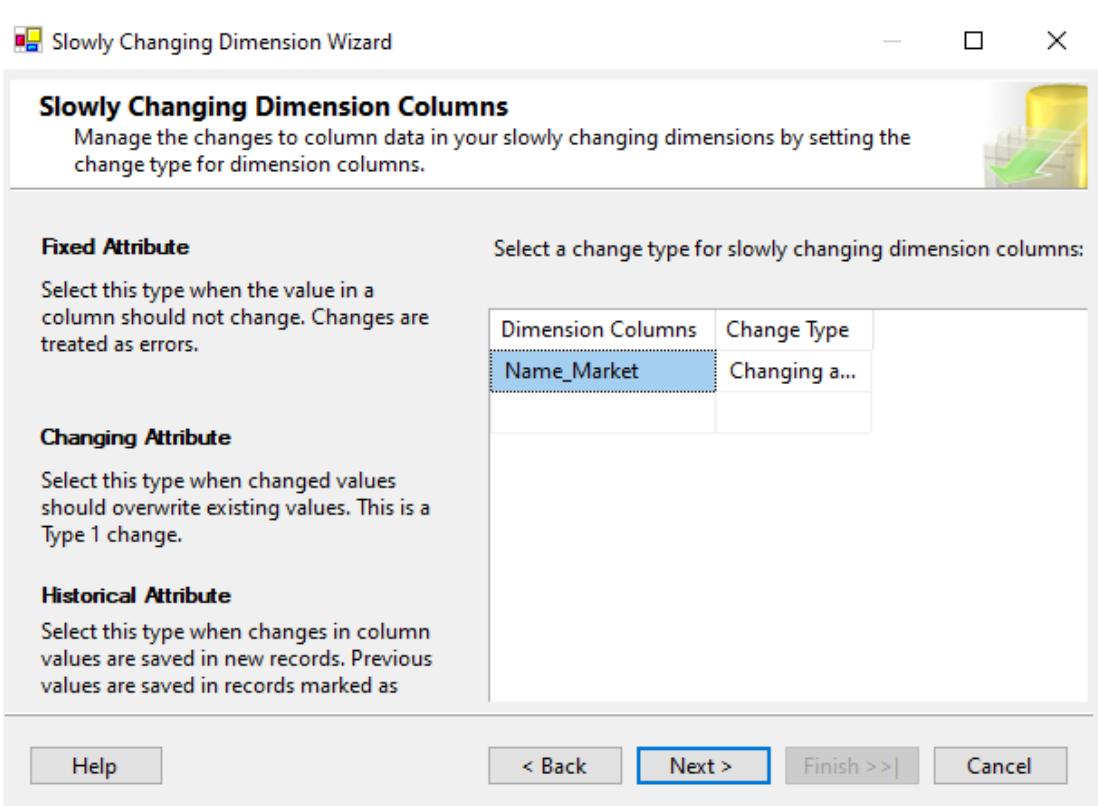
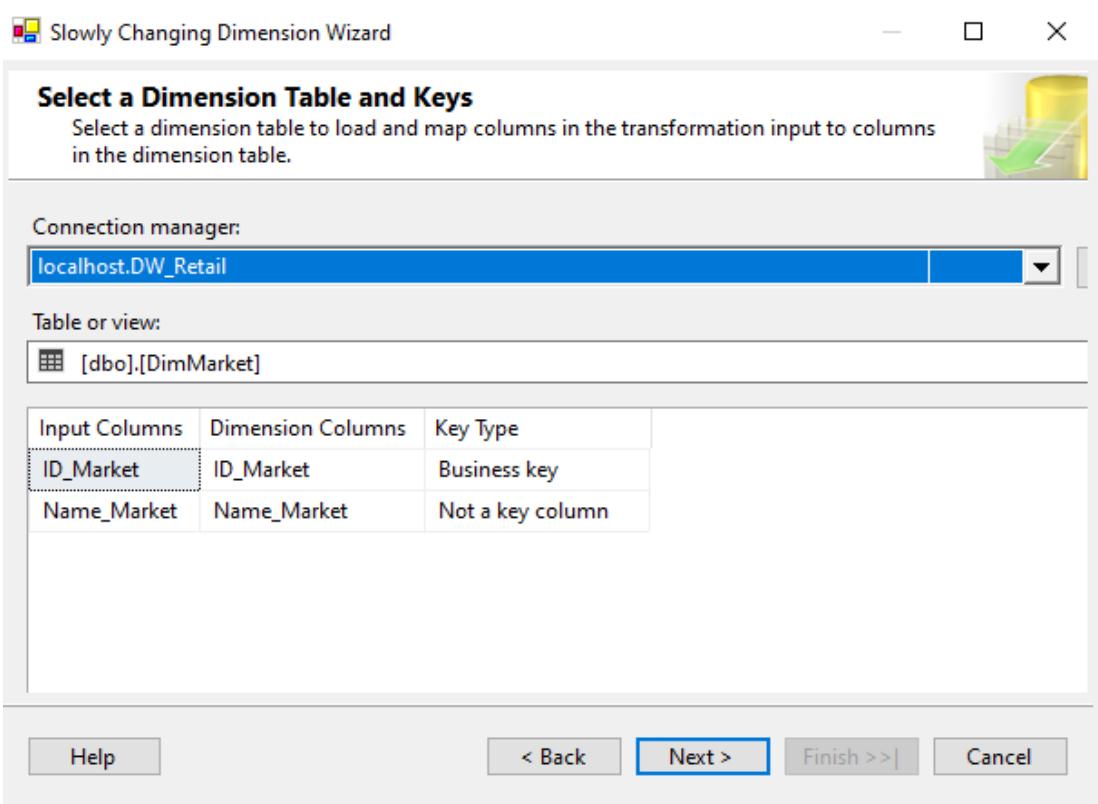
- Các thành phần của DimMarket



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Market\_Stage và chọn các cột thích hợp.

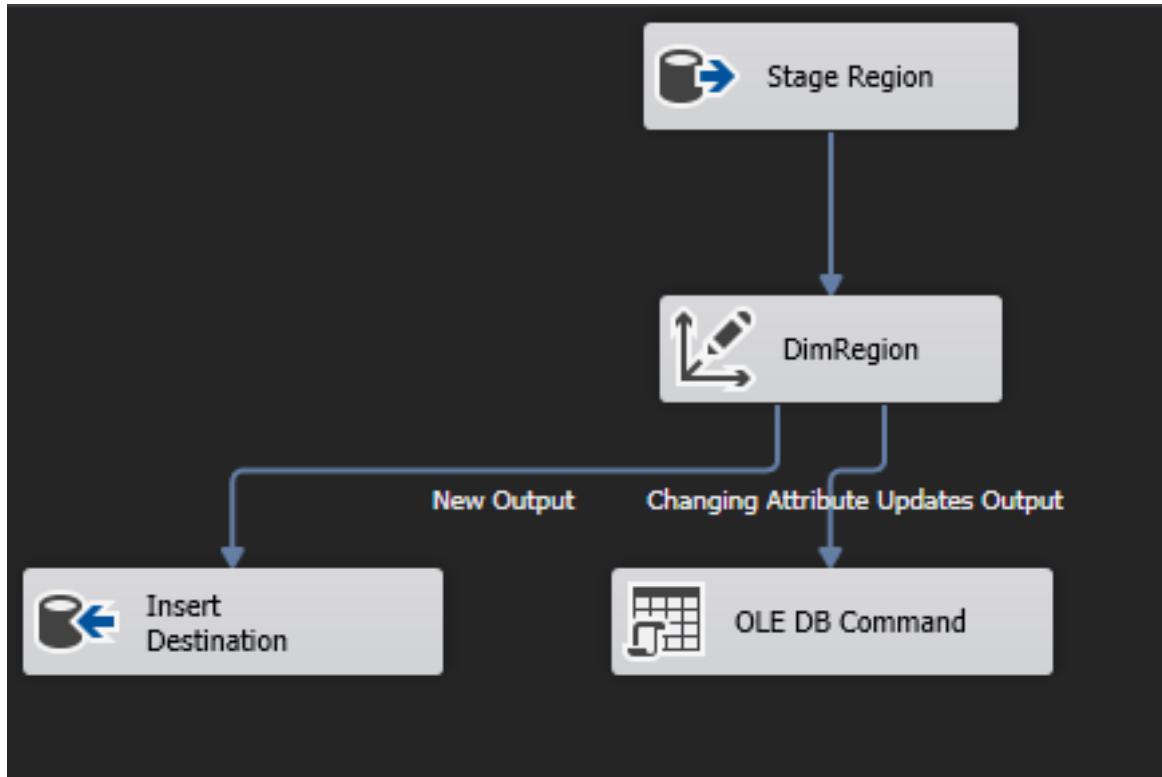


- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu.  
Chọn business key. Thuộc tính còn lại chọn changing attribute

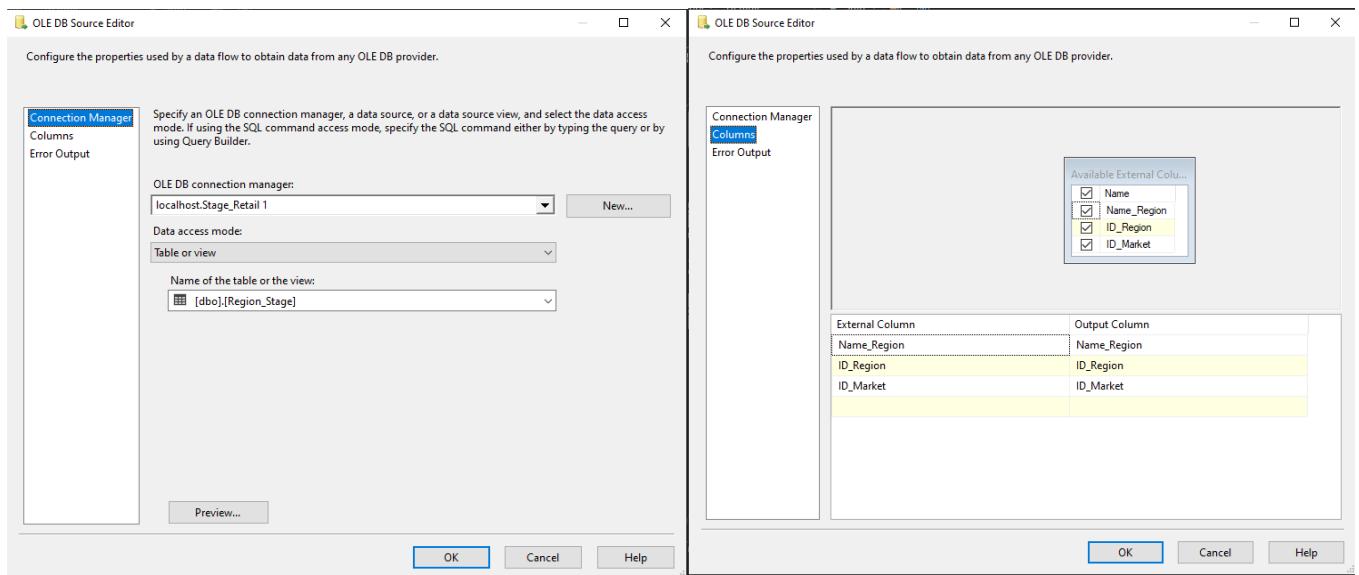


### 3.3.7. Dim Region

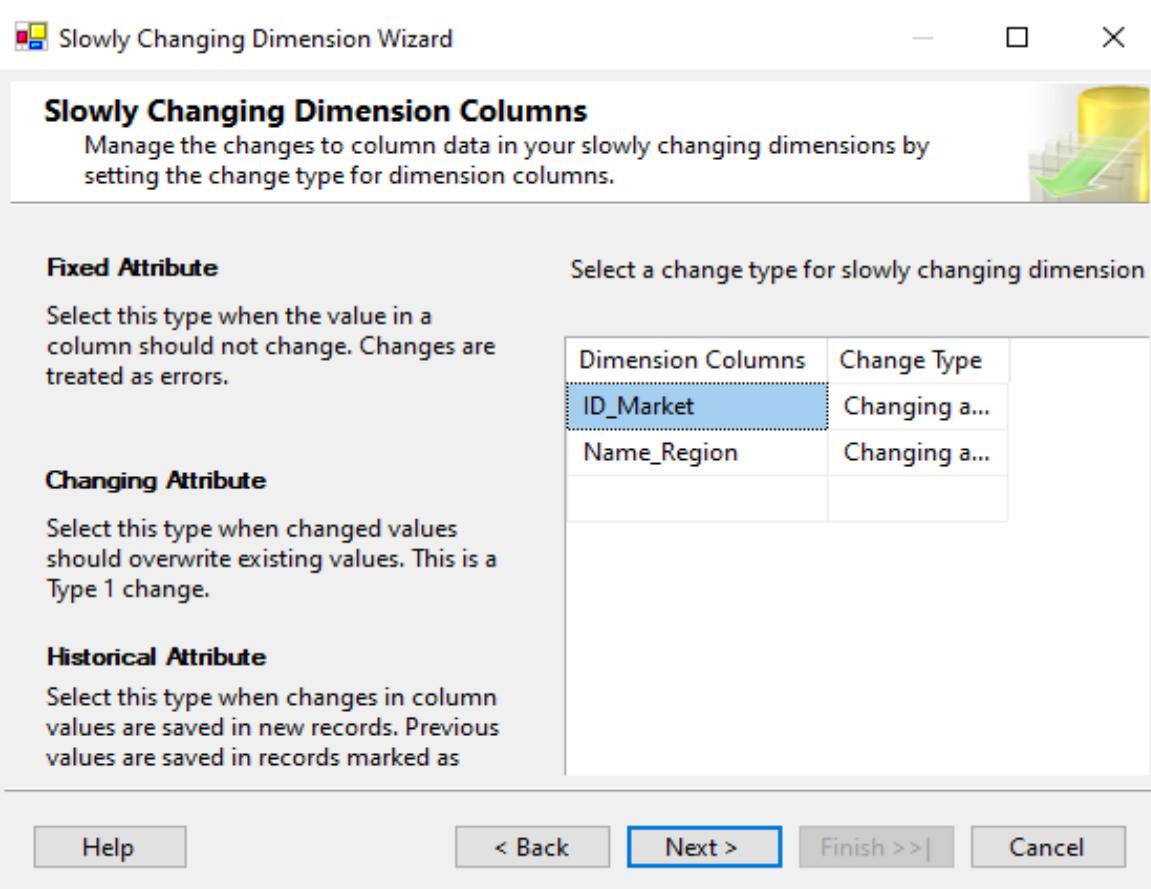
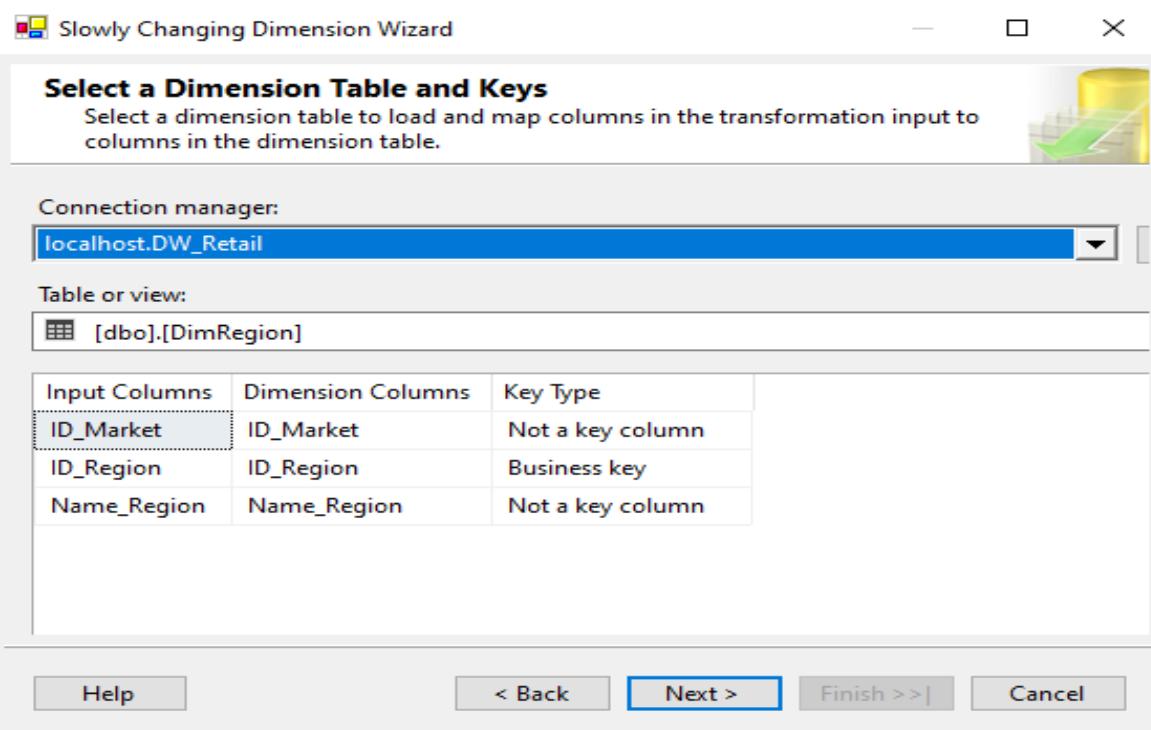
- Các thành phần của DimRegion



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Region\_Stage và chọn các cột thích hợp.

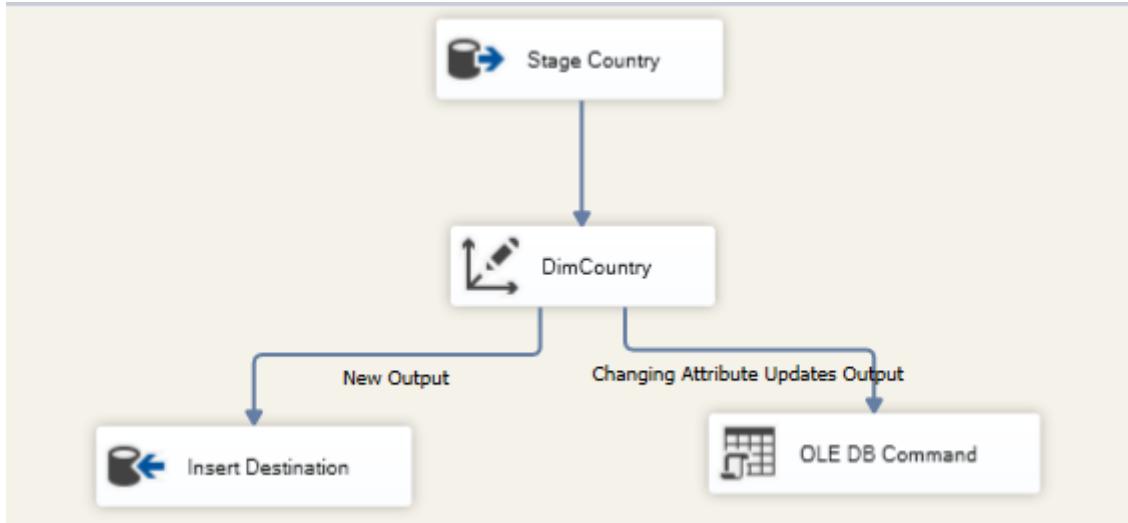


- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu.  
Chọn business key. Thuộc tính còn lại chọn changing attribute

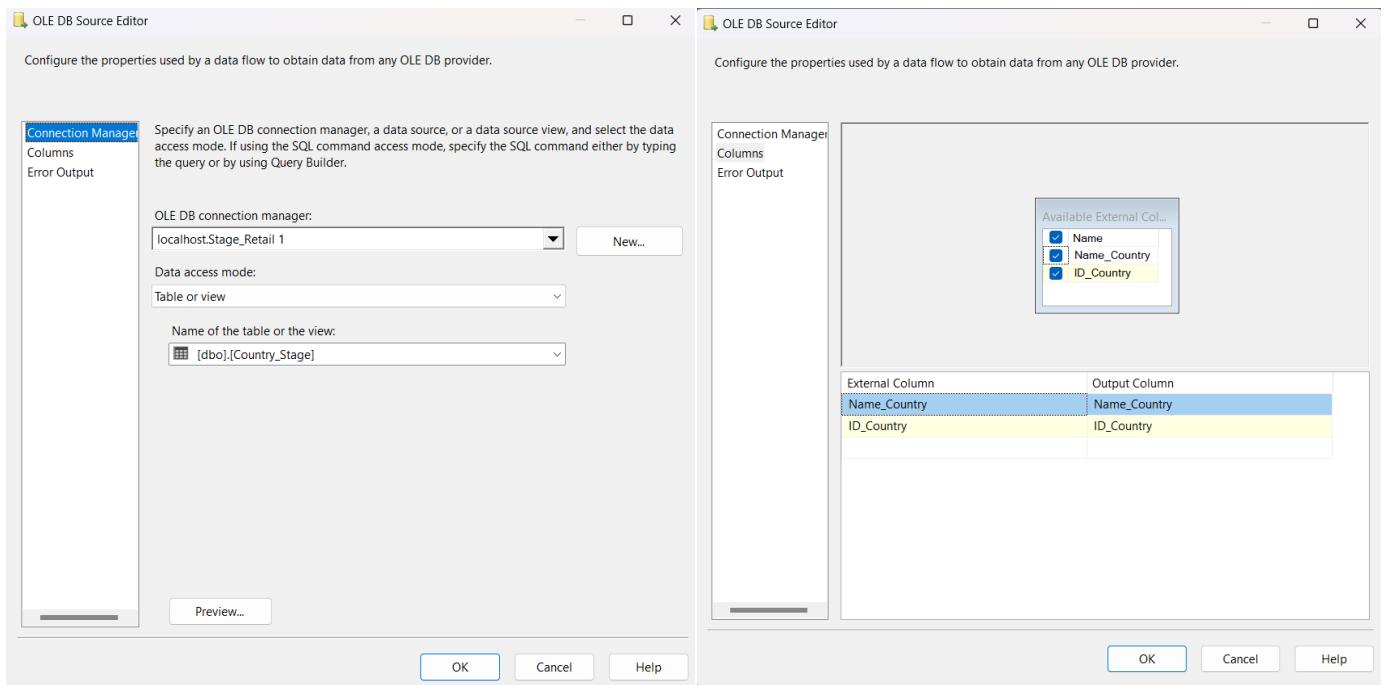


### 3.3.8. Dim Country

- Các thành phần của DimCountry



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Country\_Stage và chọn các cột thích hợp



- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute

**Slowly Changing Dimension Wizard**

### Select a Dimension Table and Keys

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

Connection manager: localhost.DW\_Retail New...

Table or view: [dbo].[DimCountry]

Input Columns	Dimension Columns	Key Type
ID_Country	ID_Country	Business key
Name_Country	Name_Country	Not a key column

[Help](#) [< Back](#) [Next >](#) [Finish >>](#) [Cancel](#)

**Slowly Changing Dimension Wizard**

### Slowly Changing Dimension Columns

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

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

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

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

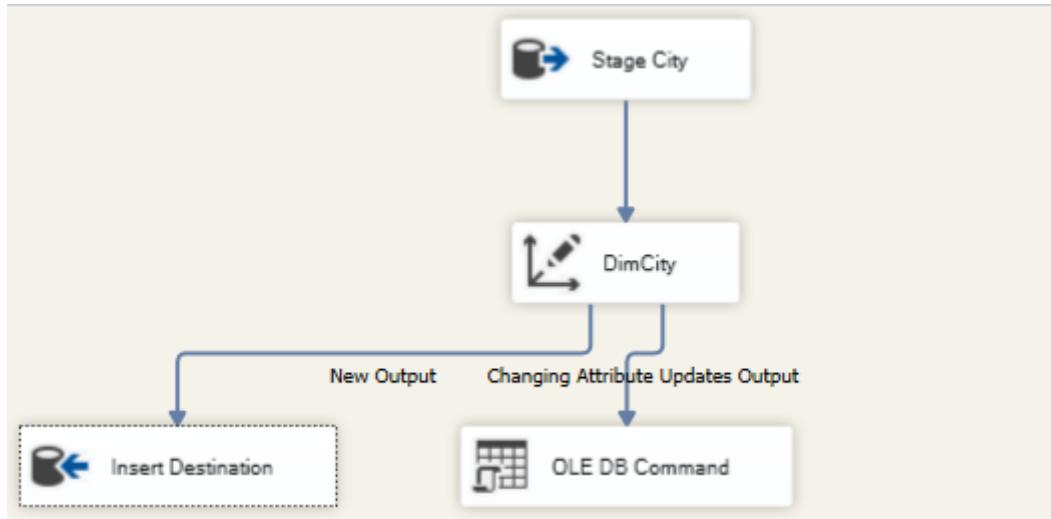
Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
Name_Country	Changing attribute

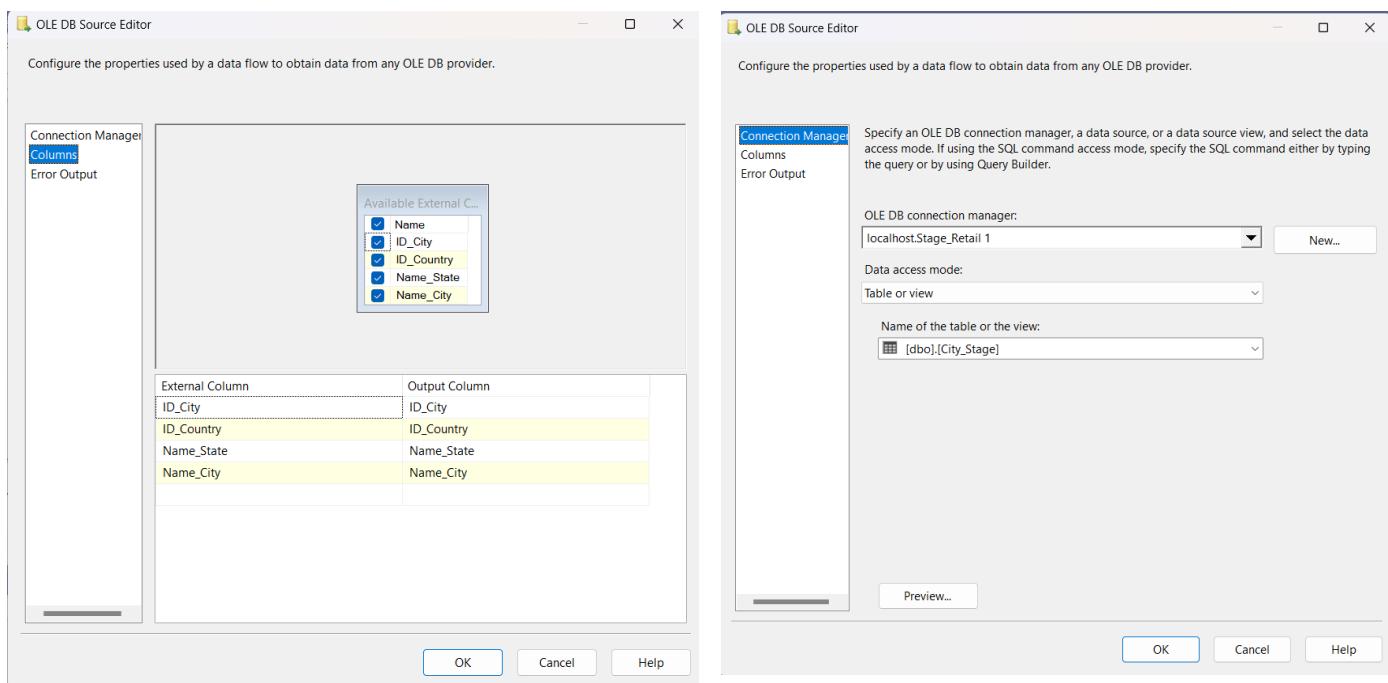
[Help](#) [< Back](#) [Next >](#) [Finish >>](#) [Cancel](#)

### 3.3.9. Dim City

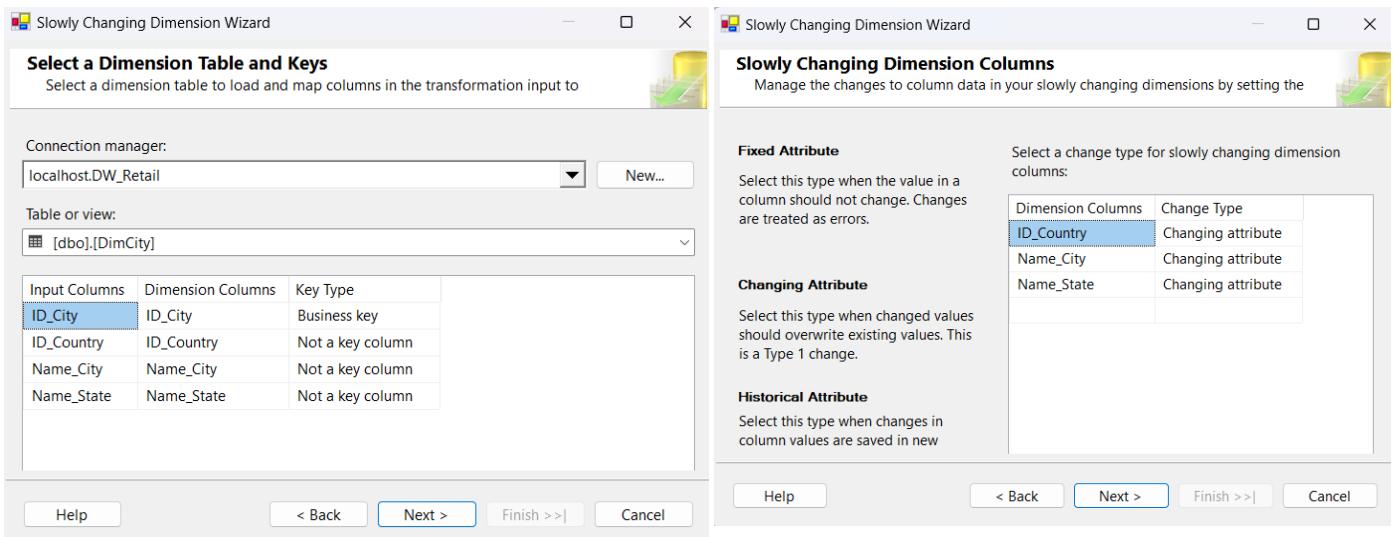
- Các thành phần của DimCity



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng City\_Stage và chọn các cột thích hợp

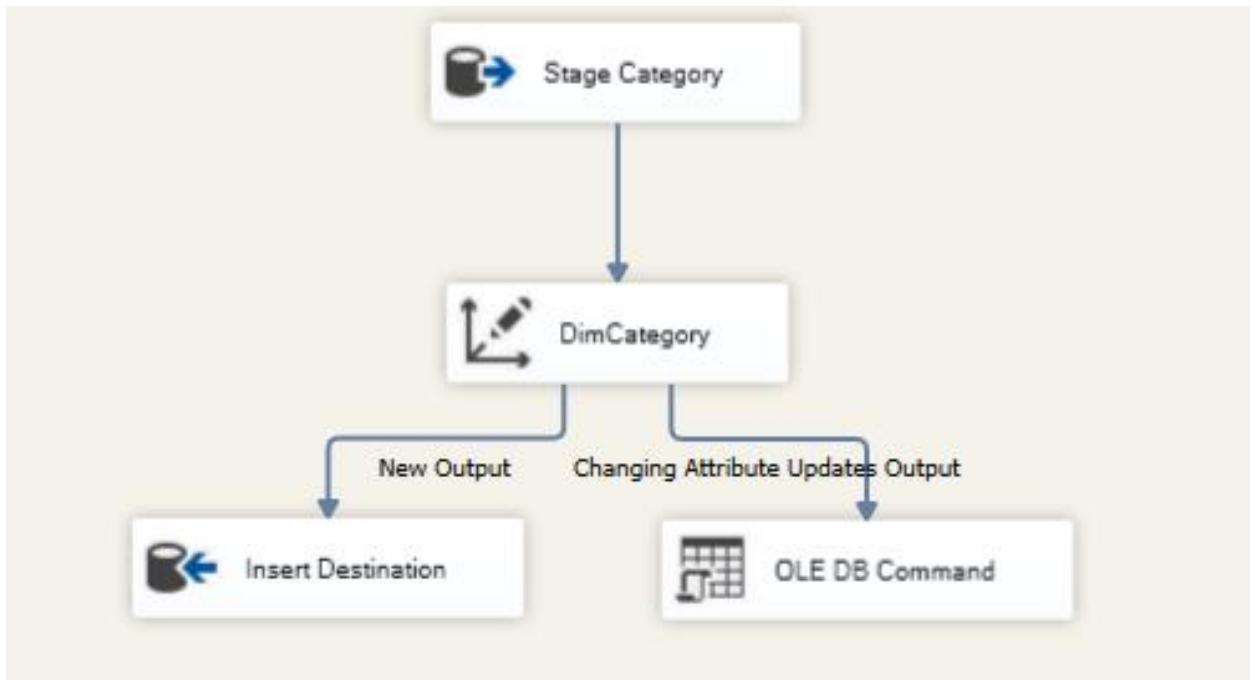


- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute

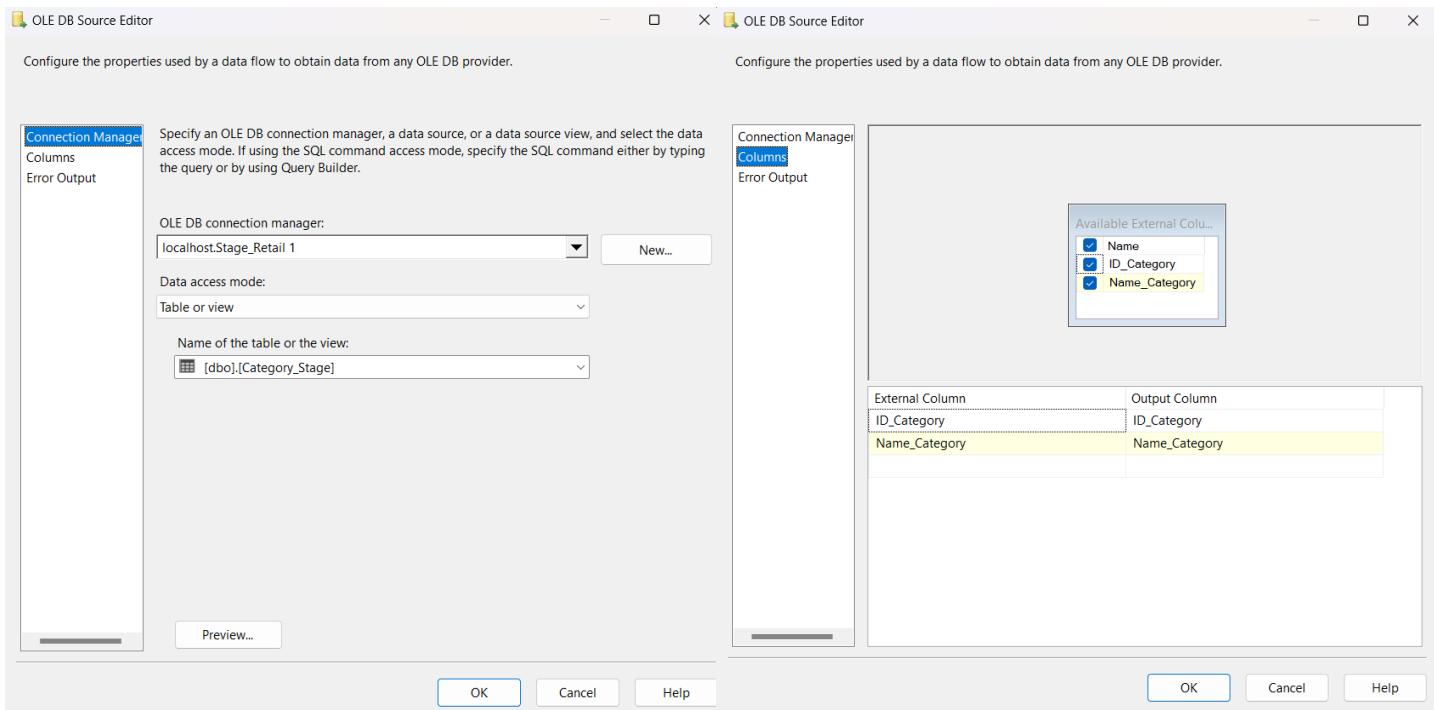


### 3.3.10. Dim Category

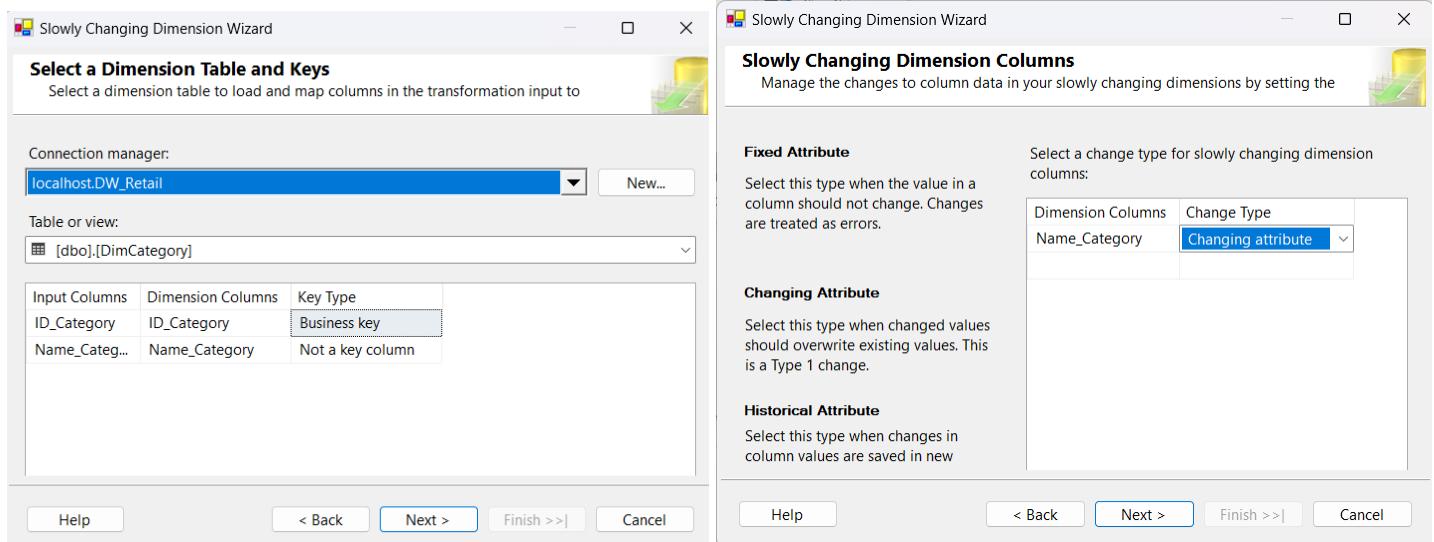
- Các thành phần của DimCategory



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Category\_Stage và chọn các cột thích hợp

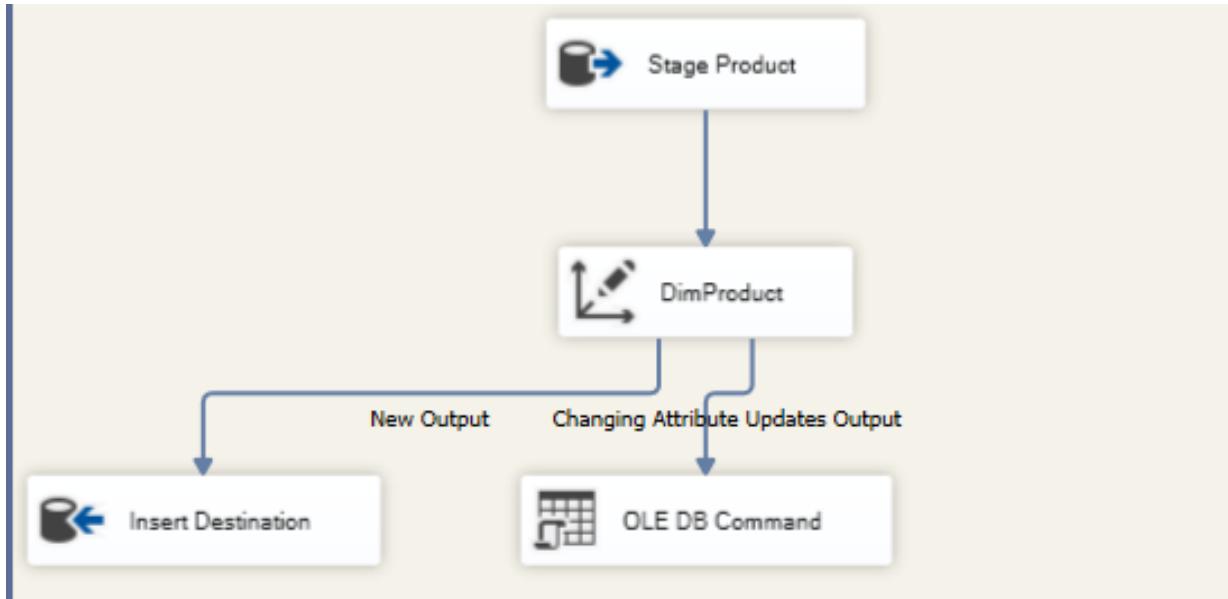


- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu.  
Chọn business key. Thuộc tính còn lại chọn changing attribute

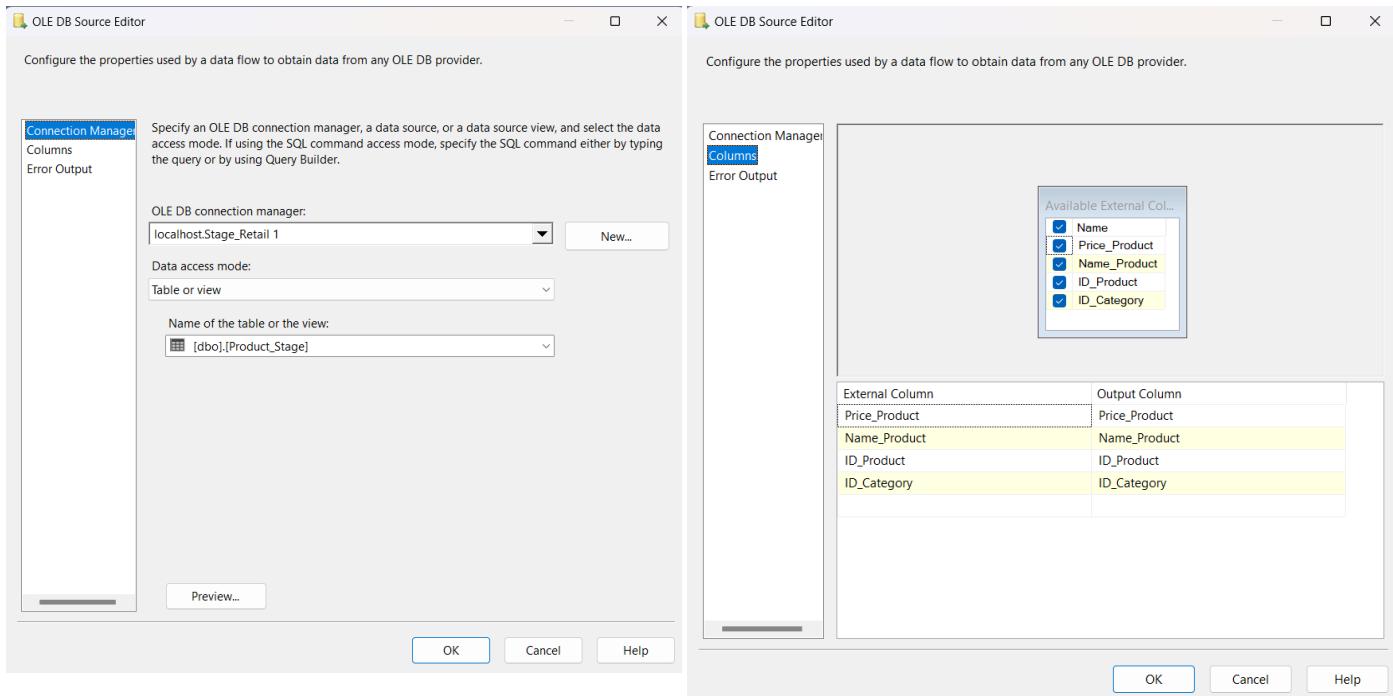


### 3.3.11. Dim Product

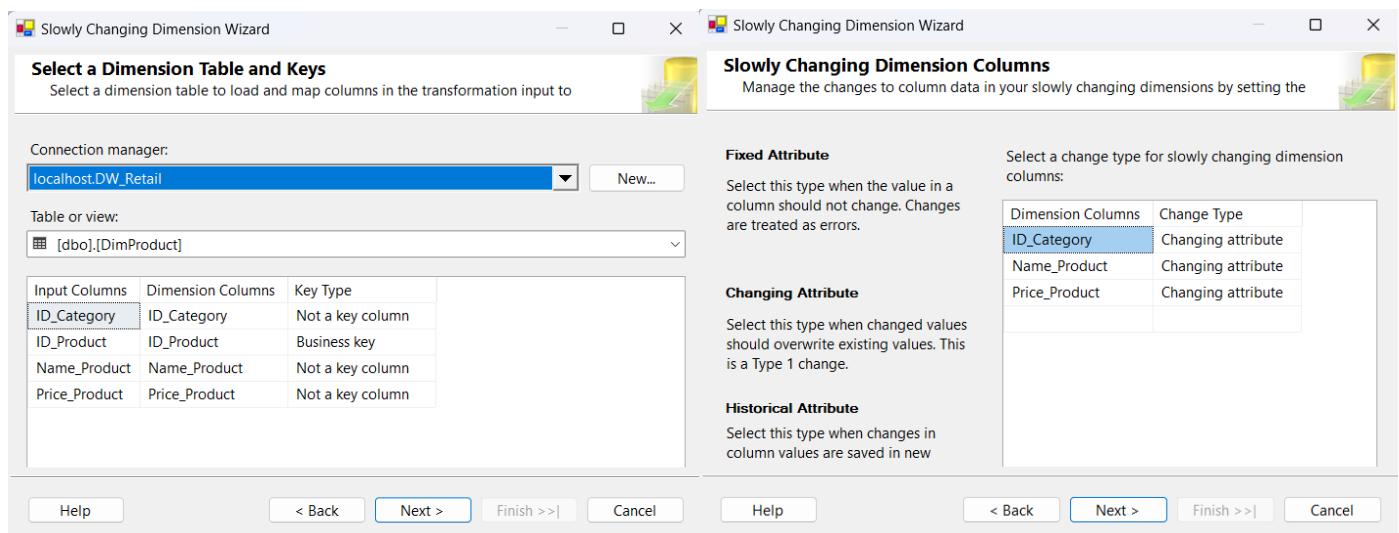
- Các thành phần của DimProduct



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Product\_Stage và chọn các cột thích hợp

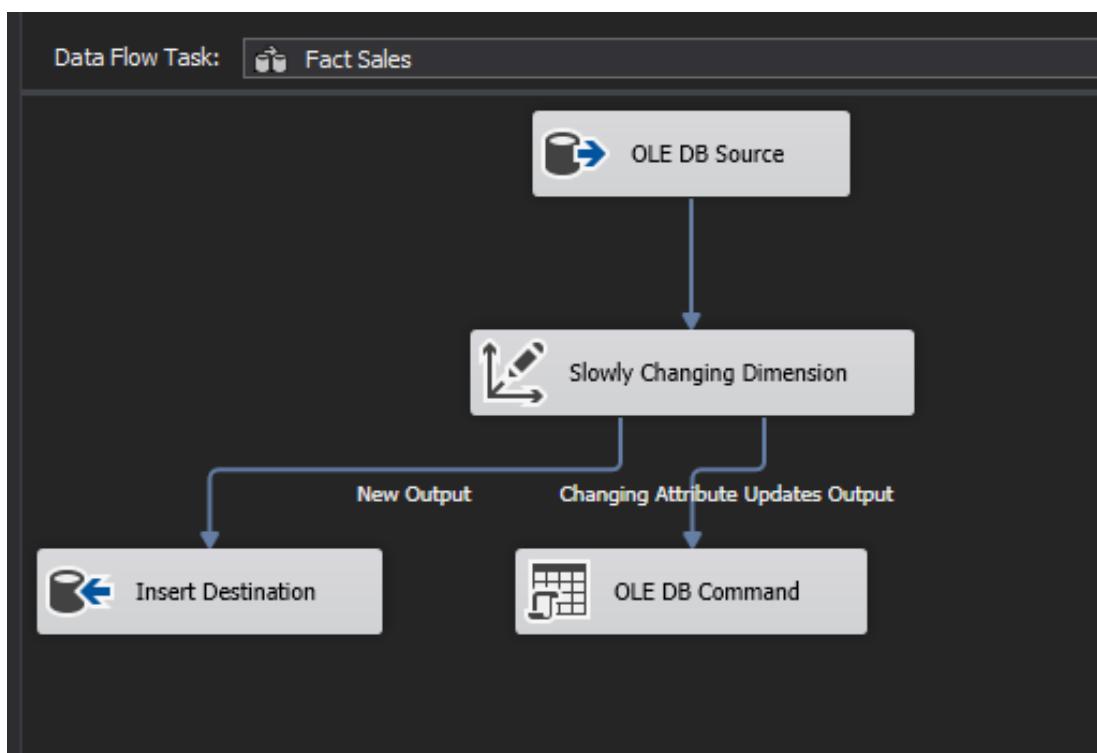


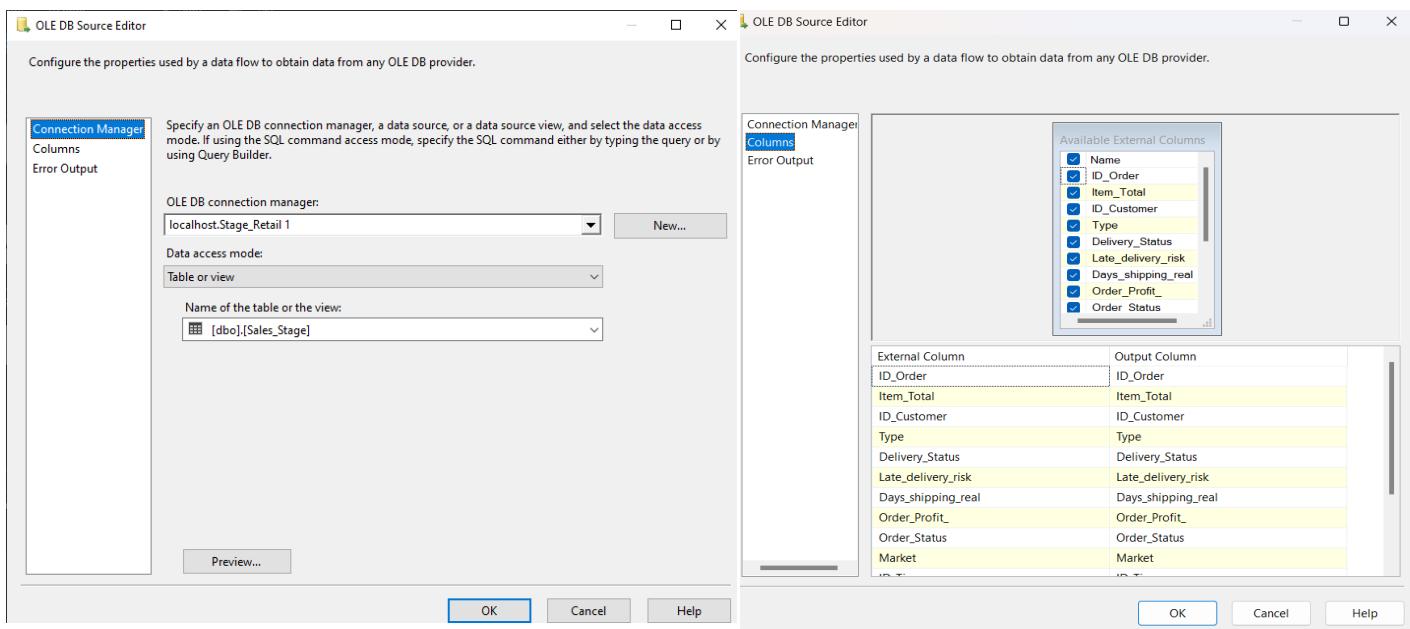
- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute



### 3.3.12. Fact Sales

- Các thành phần của FactSales





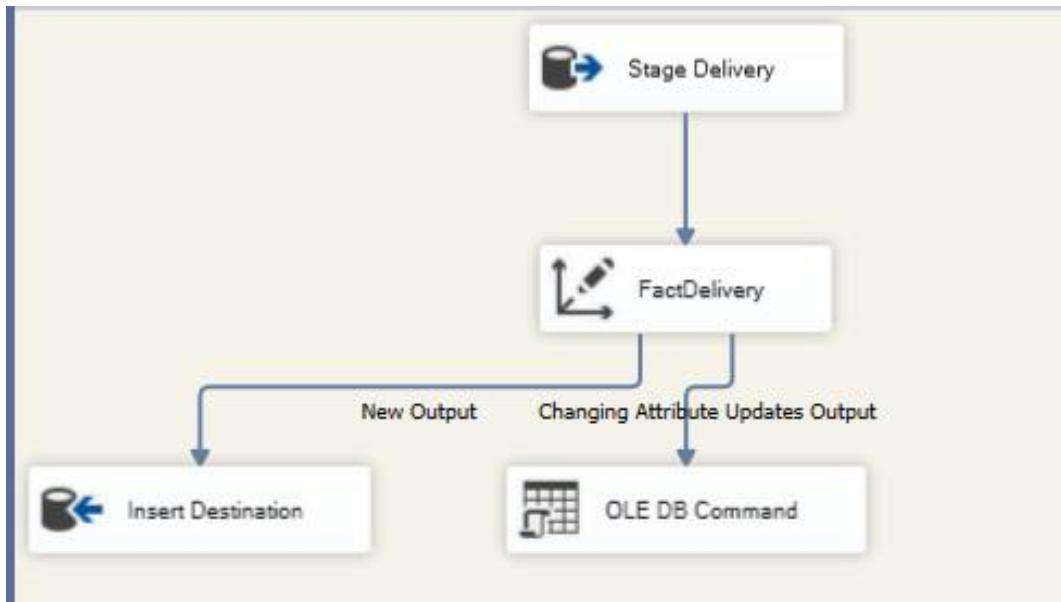
- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Sales\_Stage và chọn các cột thích hợp
- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute .

Input Columns	Dimension Columns	Key Type
Discount	Discount	Not a key column
ID_City	ID_City	Not a key column
ID_Customer	ID_Customer	Not a key column
ID_Department	ID_Department	Not a key column
ID_Line_Order	ID_Line_Order	Business key
ID_Order	ID_Order	Business key
ID_Product	ID_Product	Not a key column
ID_Region	ID_Region	Not a key column
ID_Time	ID_Time	Not a key column
Item_Profit_R...	Item_Profit_Ratio	Not a key column
Item_Total	Item_Total	Not a key column
Order_Status	Order_Status	Not a key column
Product_Price	Product_Price	Not a key column
Profit_Per_Or...	Profit_Per_Order	Not a key column
Quantity	Quantity	Not a key column
Sales	Sales	Not a key column
Type	Type	Not a key column

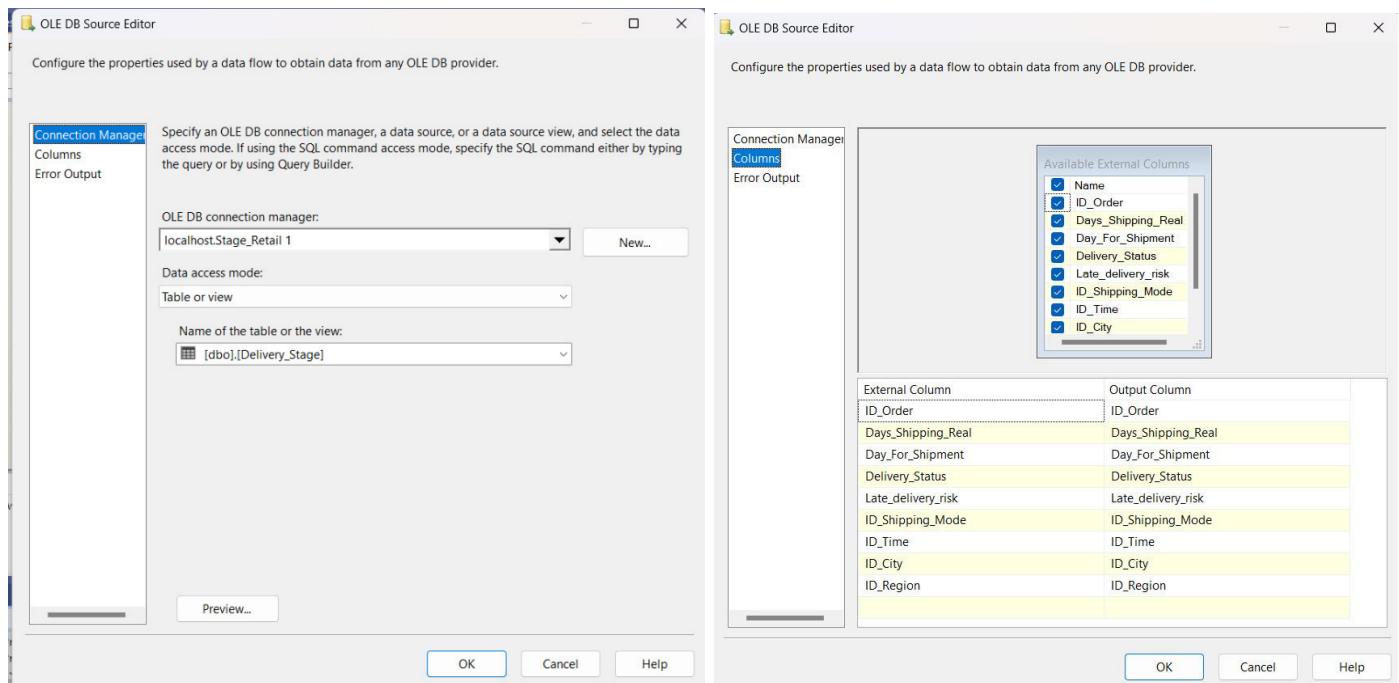
Dimension Columns	Change Type
Discount	Changing a...
ID_City	Changing a...
ID_Customer	Changing a...
ID_Department	Changing a...
ID_Product	Changing a...
ID_Region	Changing a...
ID_Time	Changing a...
Item_Profit_Ratio	Changing a...
Item_Total	Changing a...
Order_Status	Changing a...
Product_Price	Changing a...
Profit_Per_Order	Changing a...
Quantity	Changing a...
Sales	Changing a...
Type	Changing a...

### 3.3.13. Fact Delivery

- Các thành phần của FactDelivery



- Ở OLE DB Source chọn nguồn từ cơ sở dữ liệu Stage và chọn bảng Delivery\_Stage và chọn các cột thích hợp.



- Sử dụng Slowly Changing Dimension để đưa dữ liệu vào kho dữ liệu. Chọn business key. Thuộc tính còn lại chọn changing attribute .

**Slowly Changing Dimension Wizard**

### Select a Dimension Table and Keys

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

Connection manager:

localhost.DW\_Retail ▼ New...

Table or view:

[dbo].[FactDelivery] ▼

Input Columns	Dimension Columns	Key Type
Day_For_Ship...	Day_For_Shipment	Not a key column
Days_Shippin...	Days_Shipping_Real	Not a key column
Delivery_Stat...	Delivery_Status	Not a key column
ID_City	ID_City	Not a key column
ID_Order	ID_Order	Business key
ID_Region	ID_Region	Not a key column

Help < Back Next > Finish >> Cancel

**Slowly Changing Dimension Wizard**

### Slowly Changing Dimension Columns

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

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

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

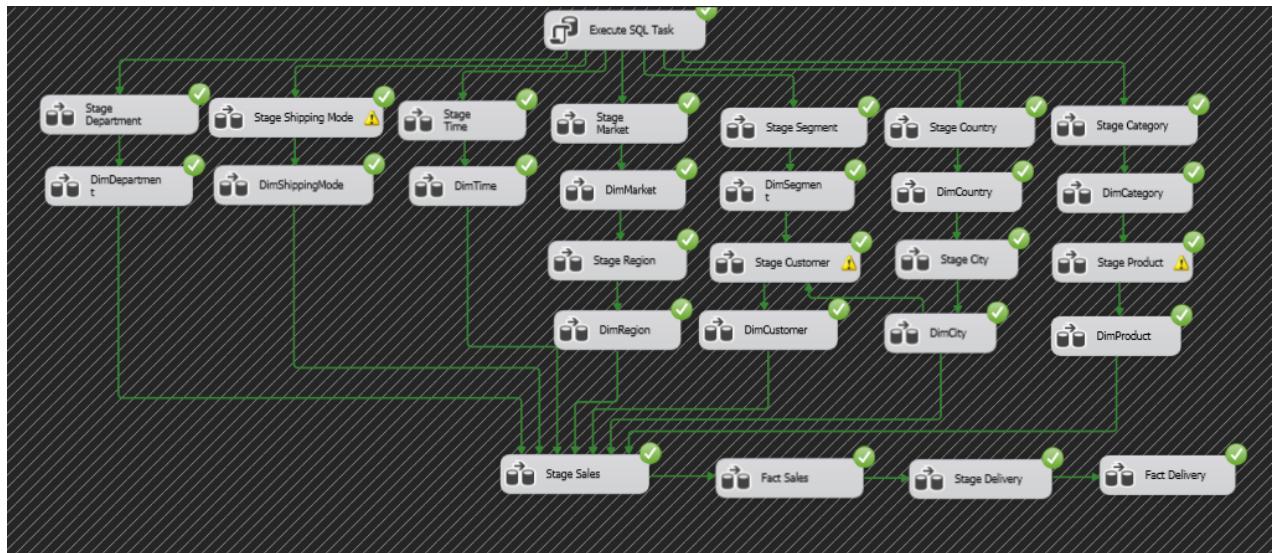
**Historical Attribute**  
Select this type when changes in column values are saved in new

Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
Day_For_Shipment	Changing a...
Days_Shipping_Real	Changing a...
Delivery_Status	Changing a...
ID_City	Changing a...
ID_Region	Changing a...
ID_Shipping_Mode	Changing a...
ID_Time	Changing a...
Late_delivery_risk	Changing a...

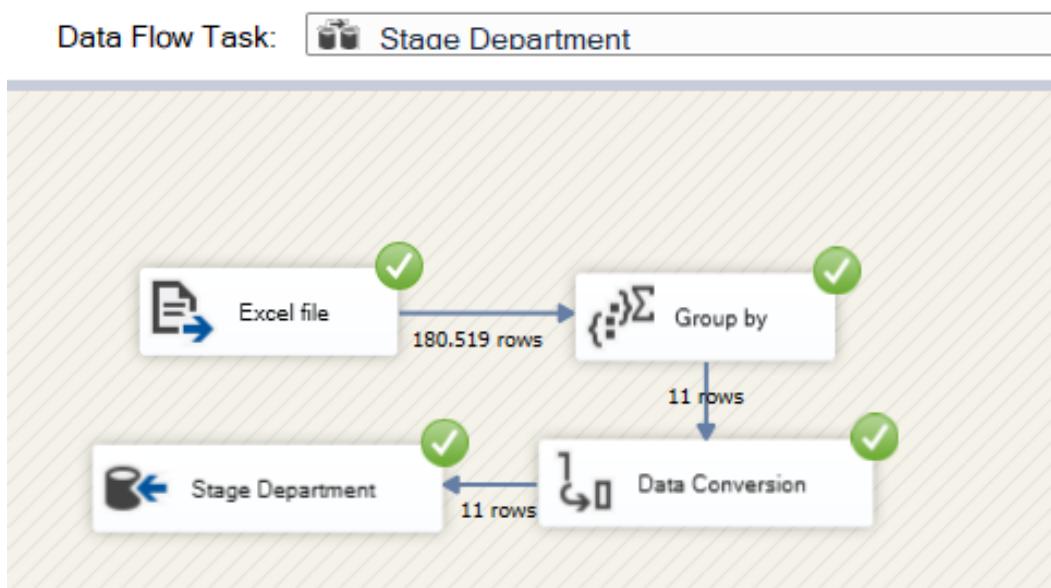
Help < Back Next > Finish >> Cancel

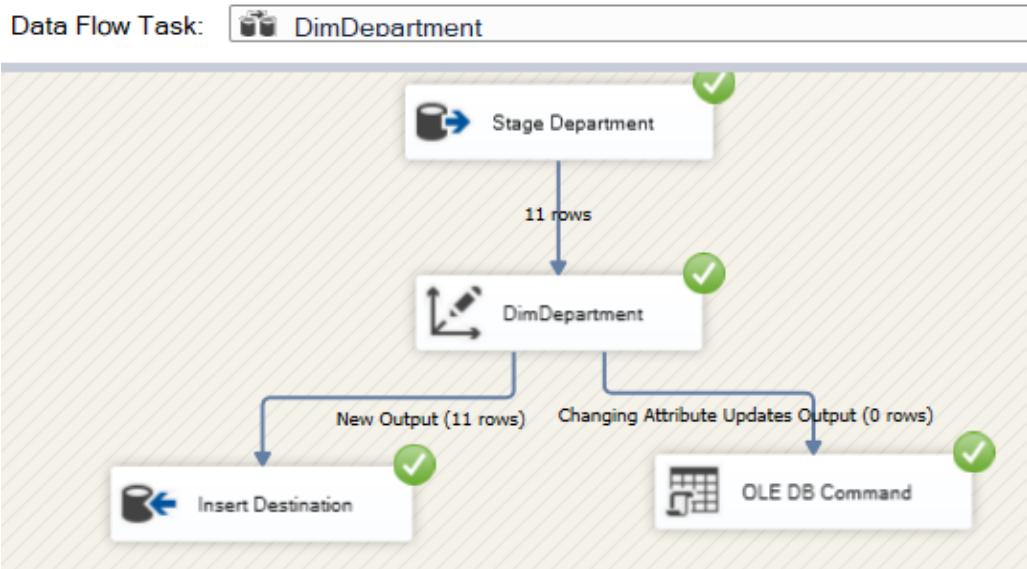
### 3.4. Kết quả đổ dữ liệu vào Data Warehouse



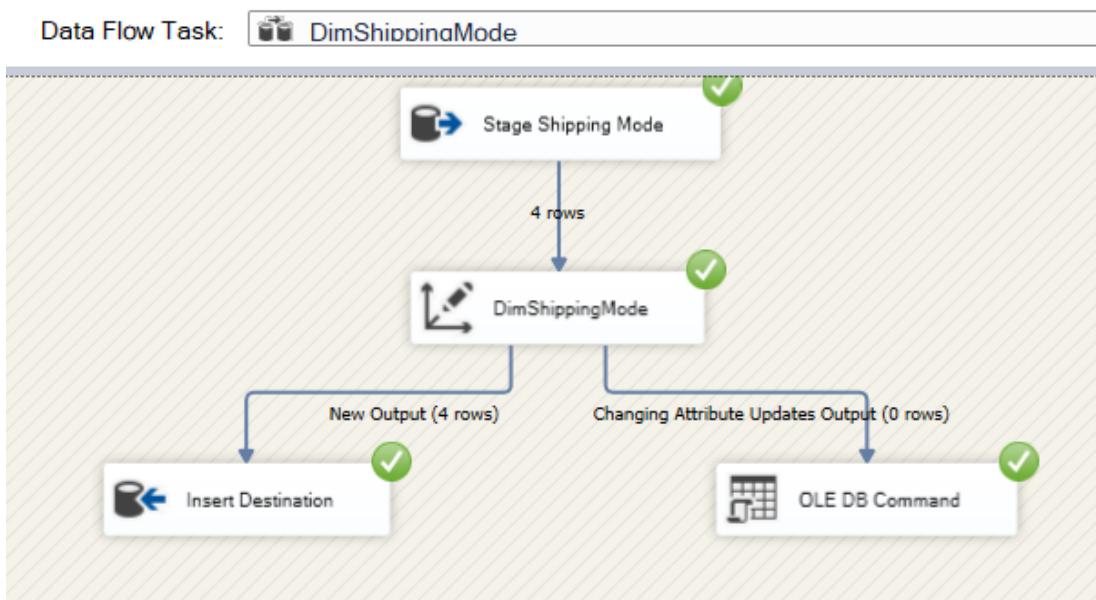
#### 3.4.1. Luồng dữ liệu

##### 3.4.1.1. Department



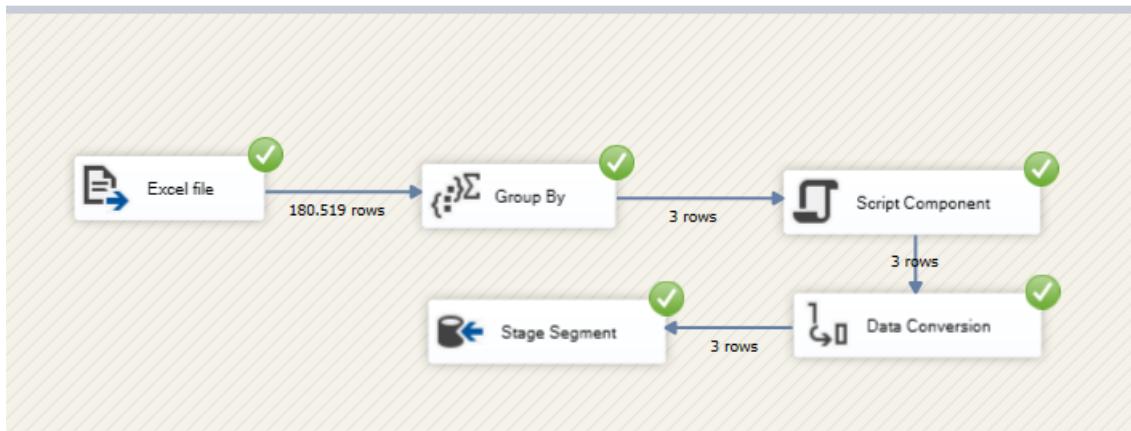


### 3.4.1.2. Shipping Mode

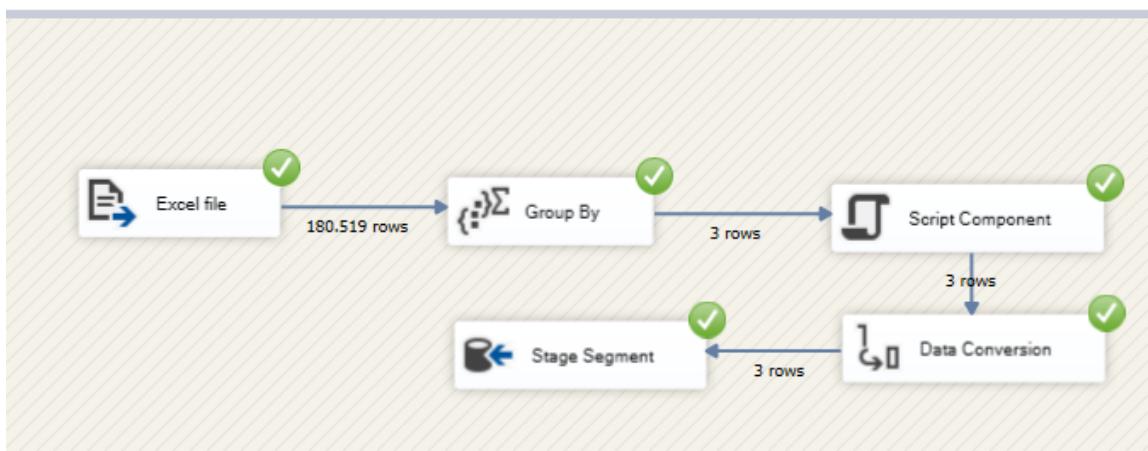


### 3.4.1.3. Segment

Data Flow Task: Stage Segment

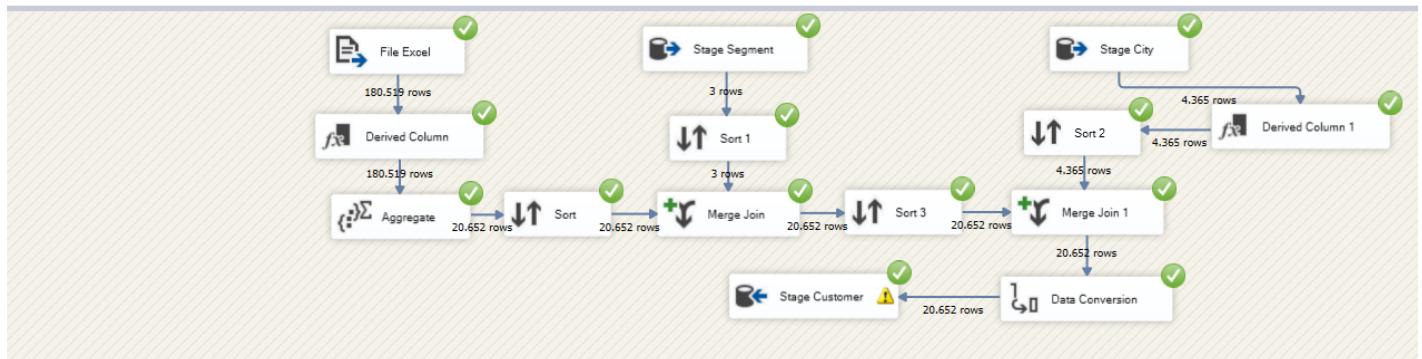


Data Flow Task: DimSegment

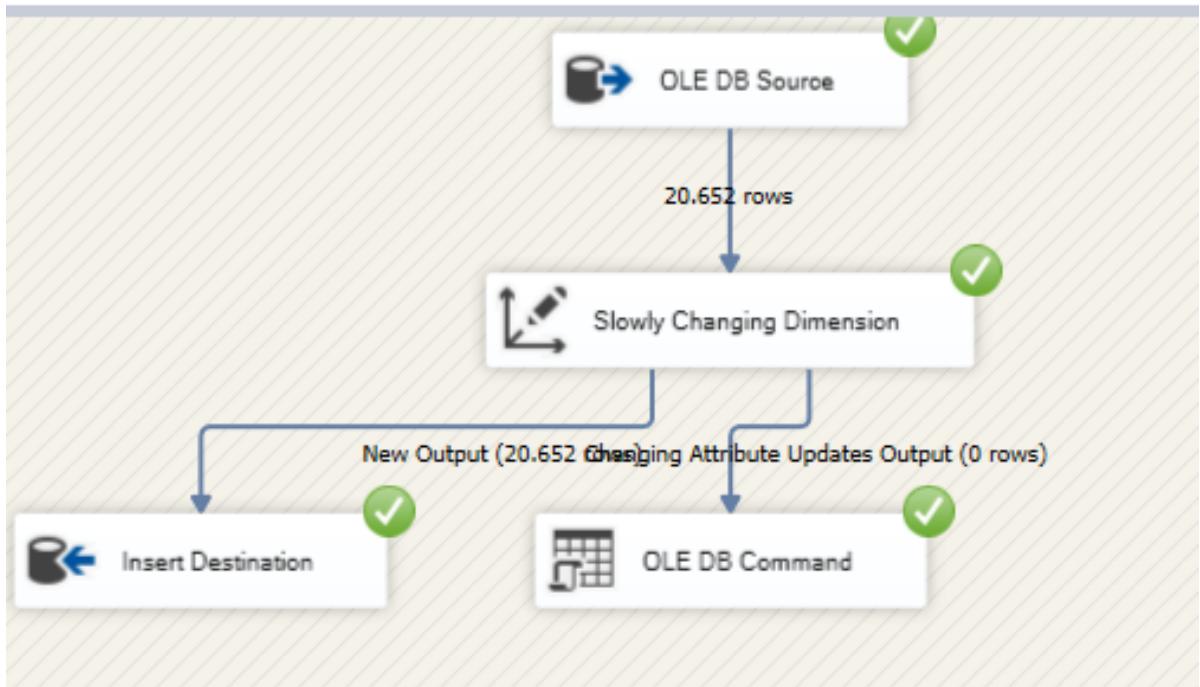


### 3.4.1.4. Customer

Data Flow Task: Stage Customer

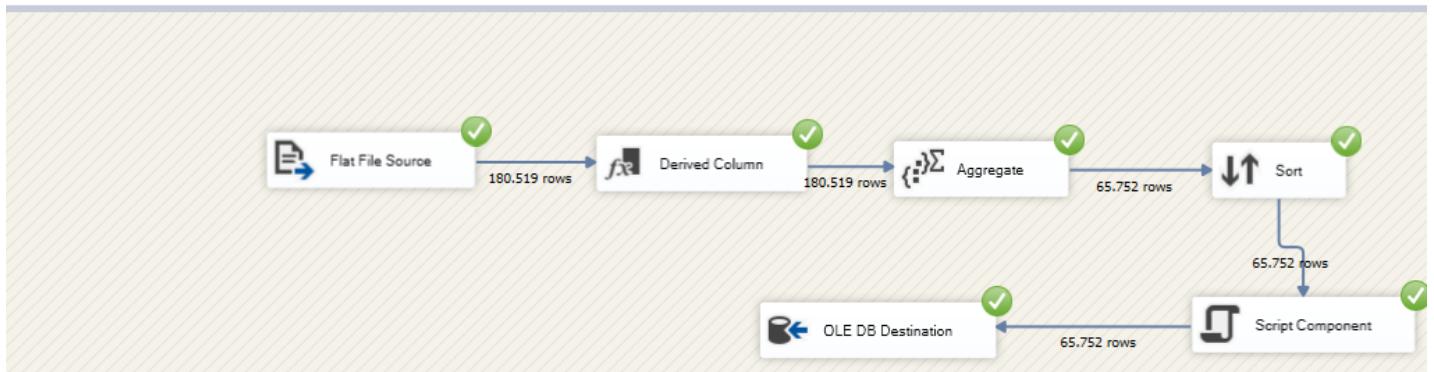


Data Flow Task:  DimCustomer

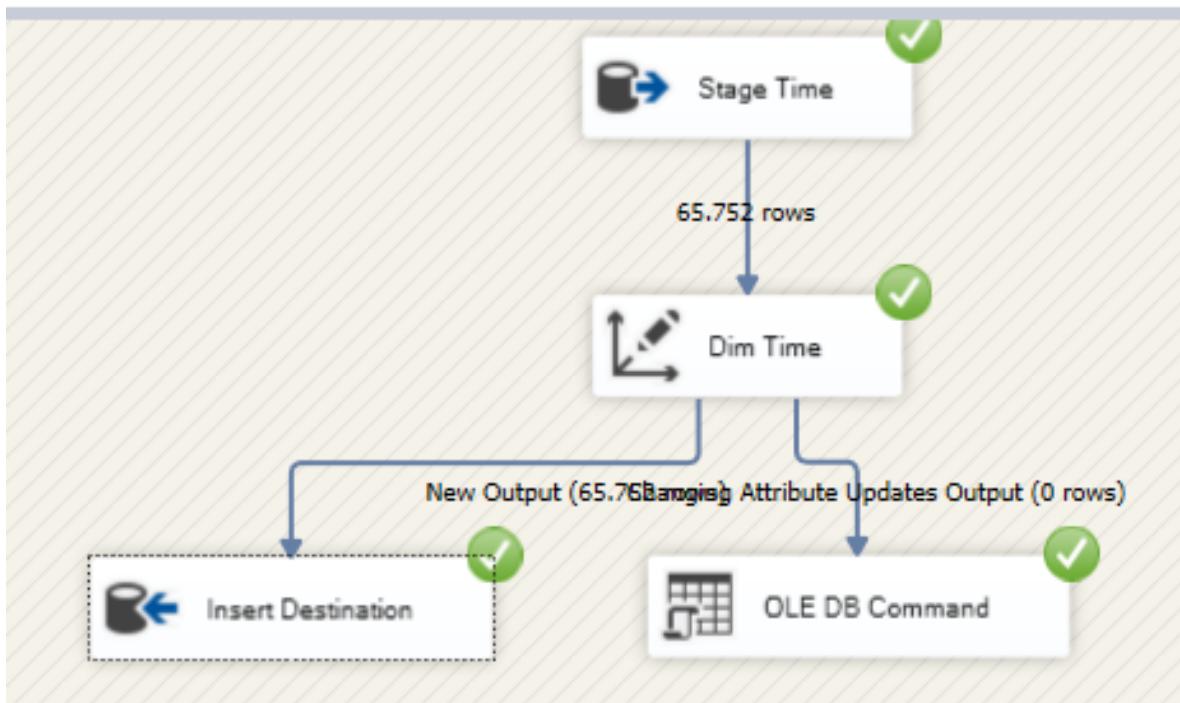


### 3.4.1.5. Time

Data Flow Task:  Stage Time

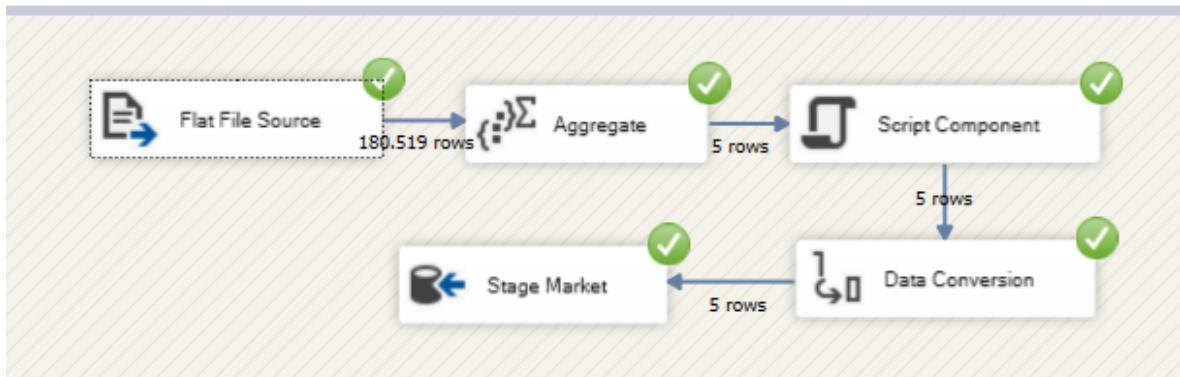


Data Flow Task:  DimTime

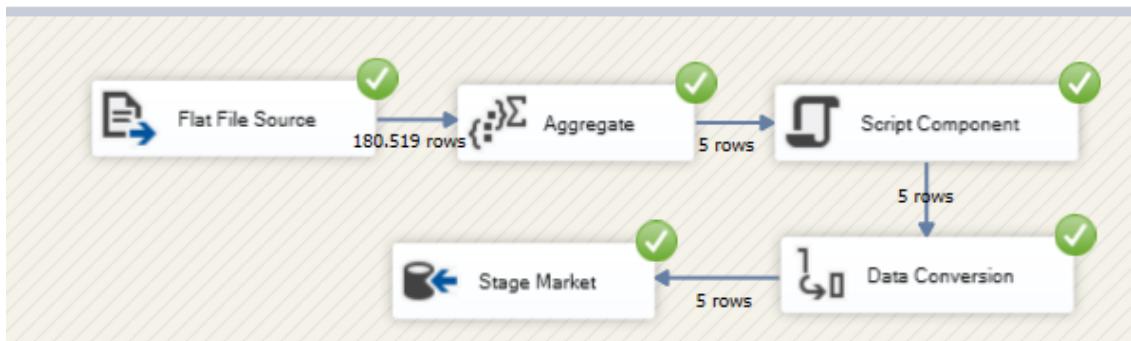


### 3.4.1.6. Market

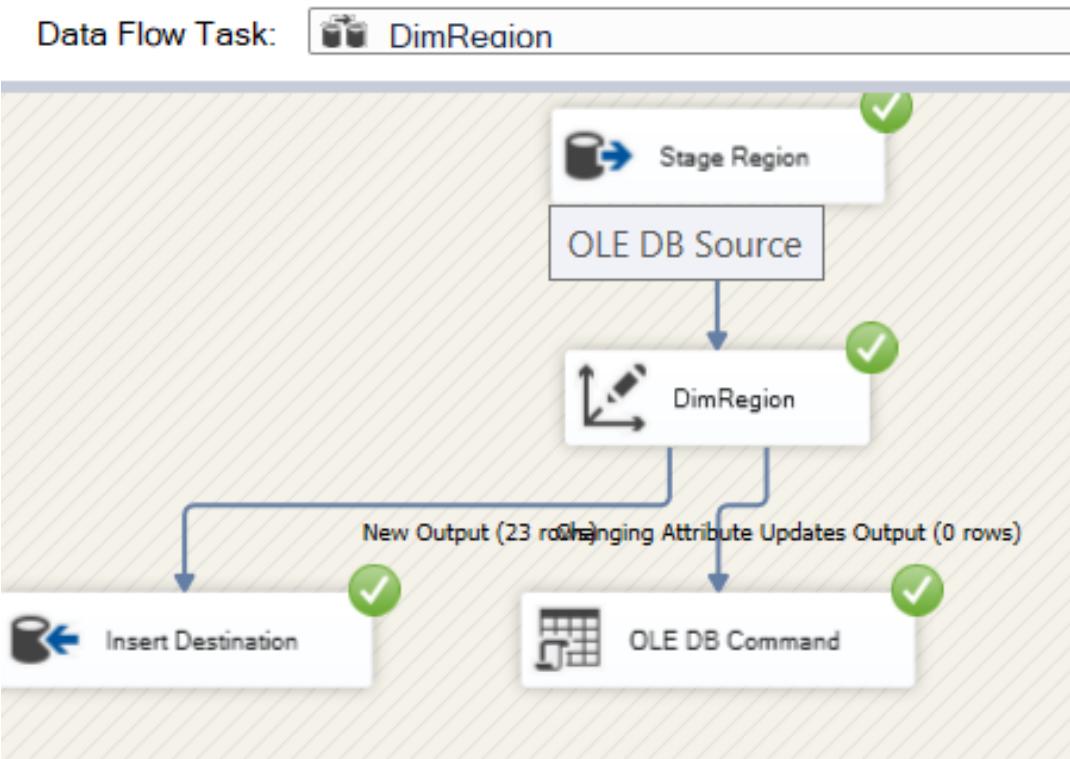
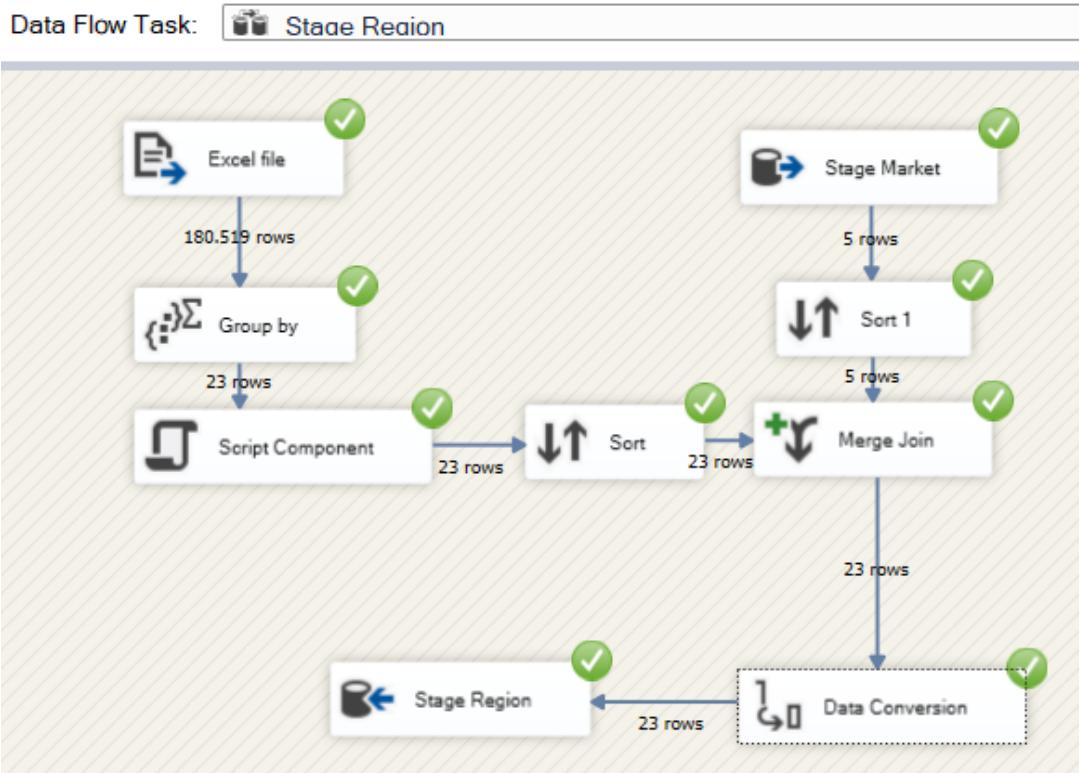
Data Flow Task:  Stage Market



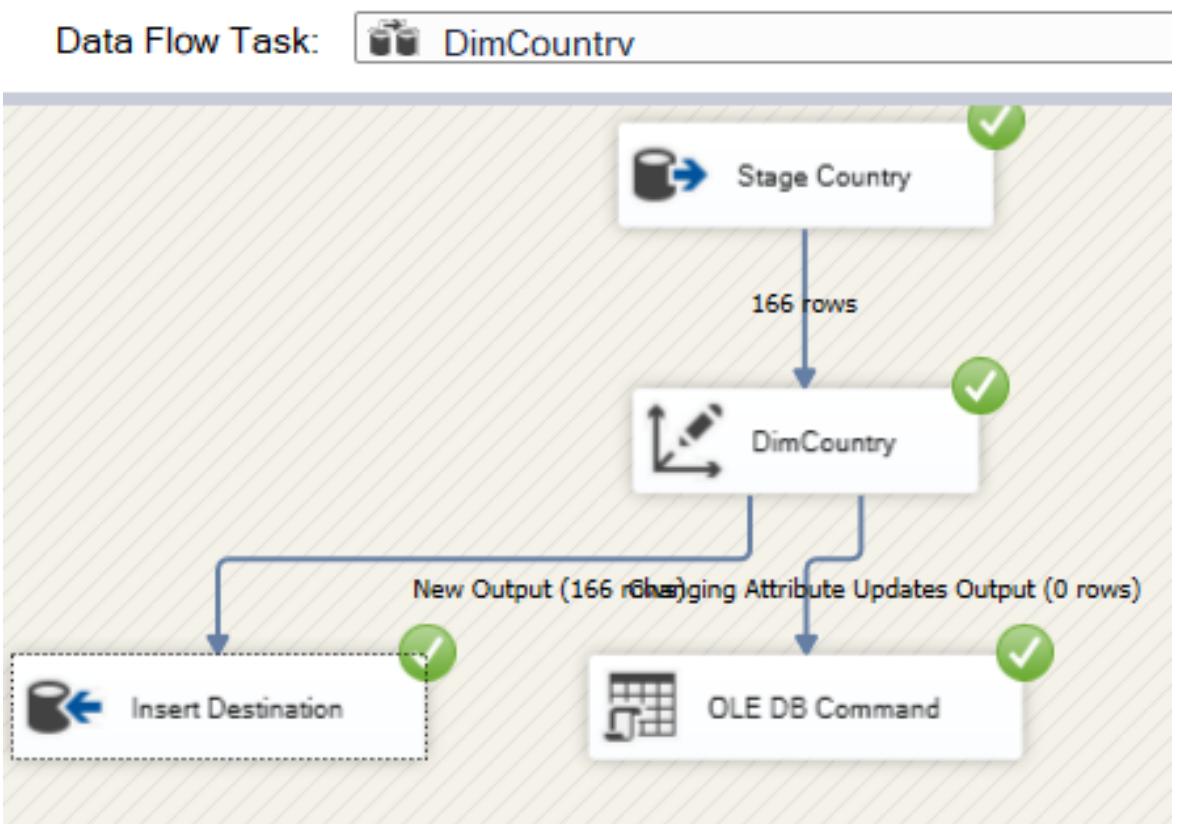
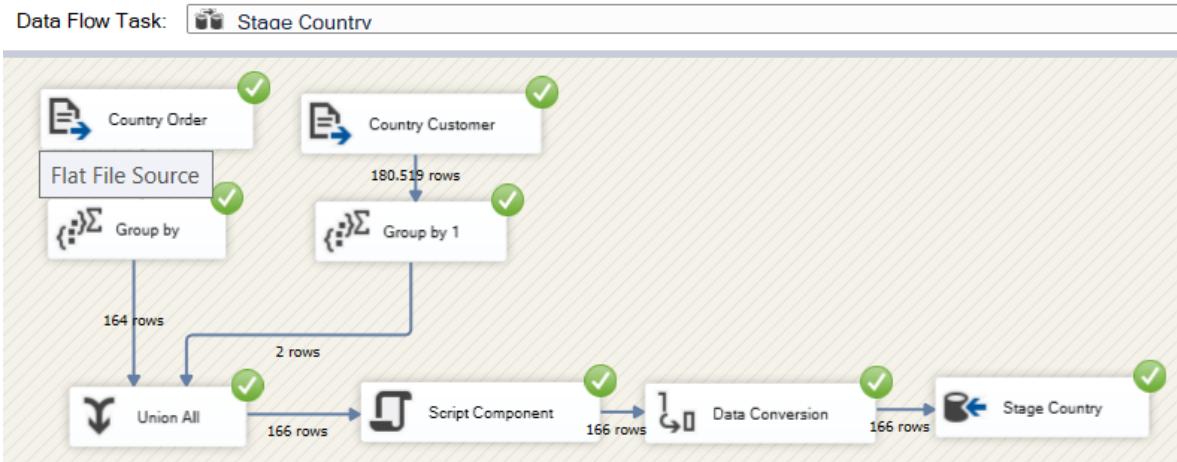
Data Flow Task:  DimMarket



### 3.4.1.7. Region

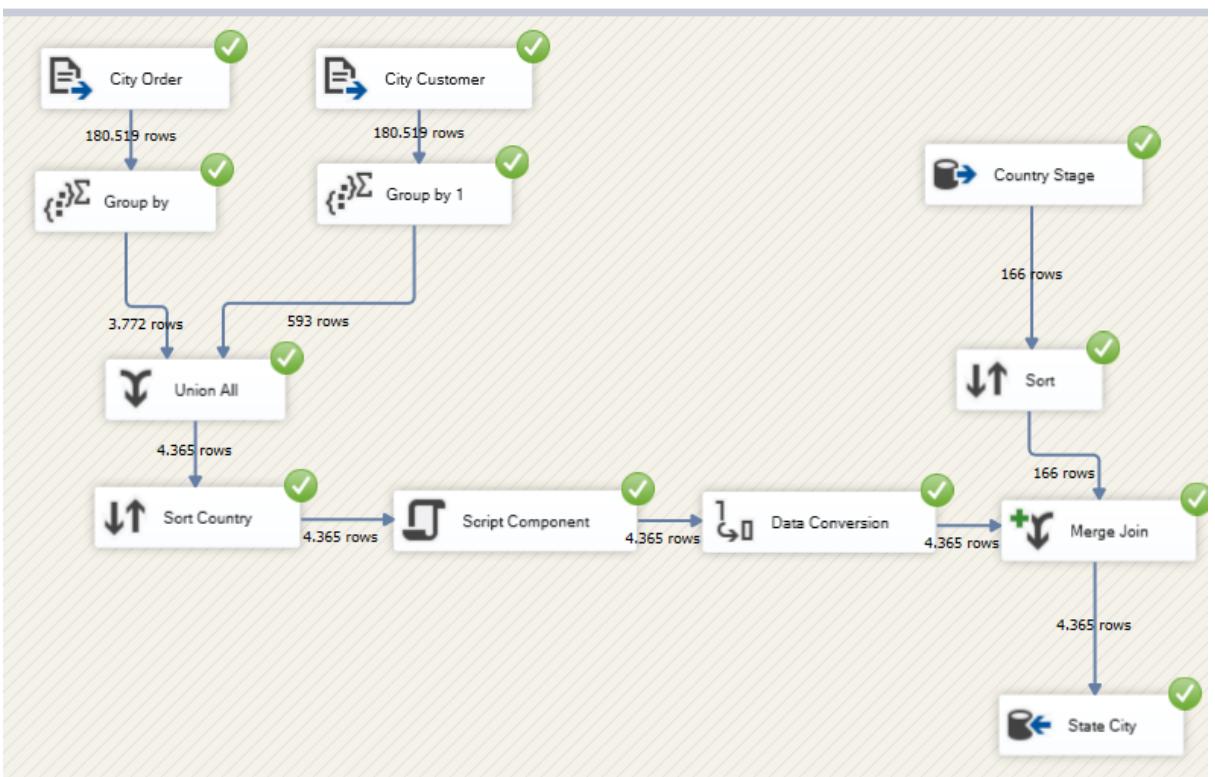


### 3.4.1.8. Country

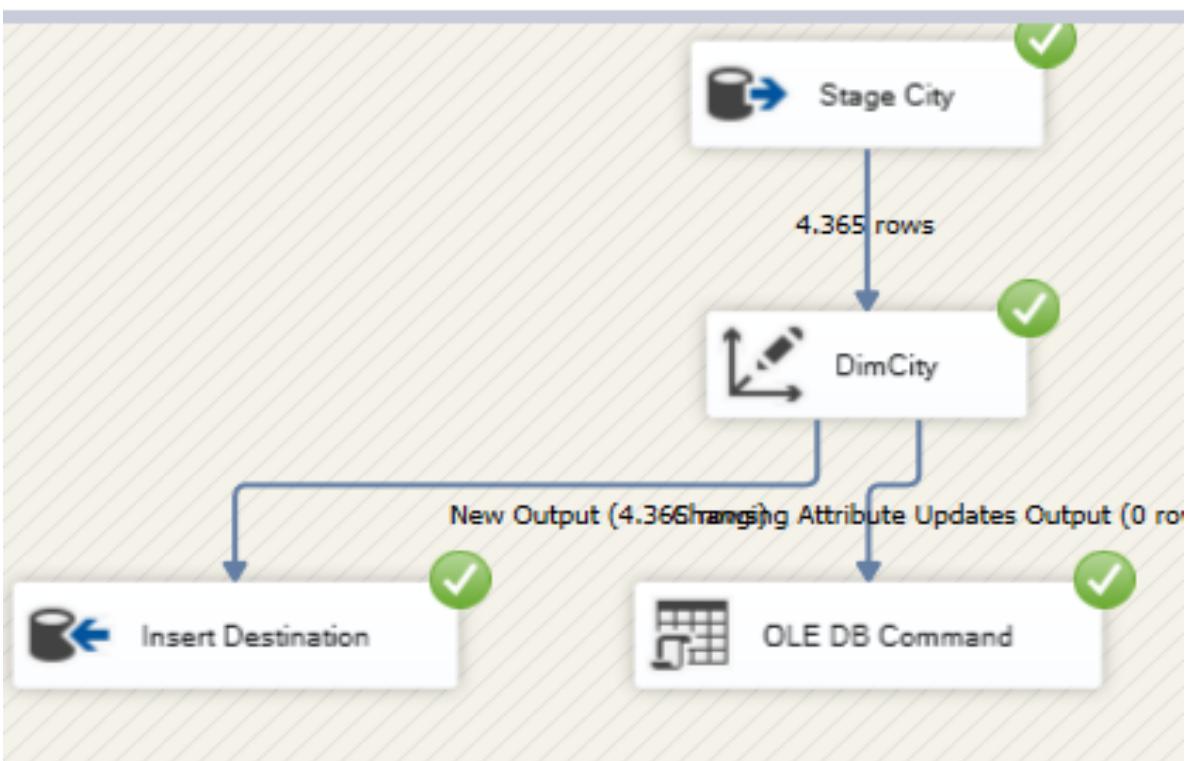


### 3.4.1.9. City

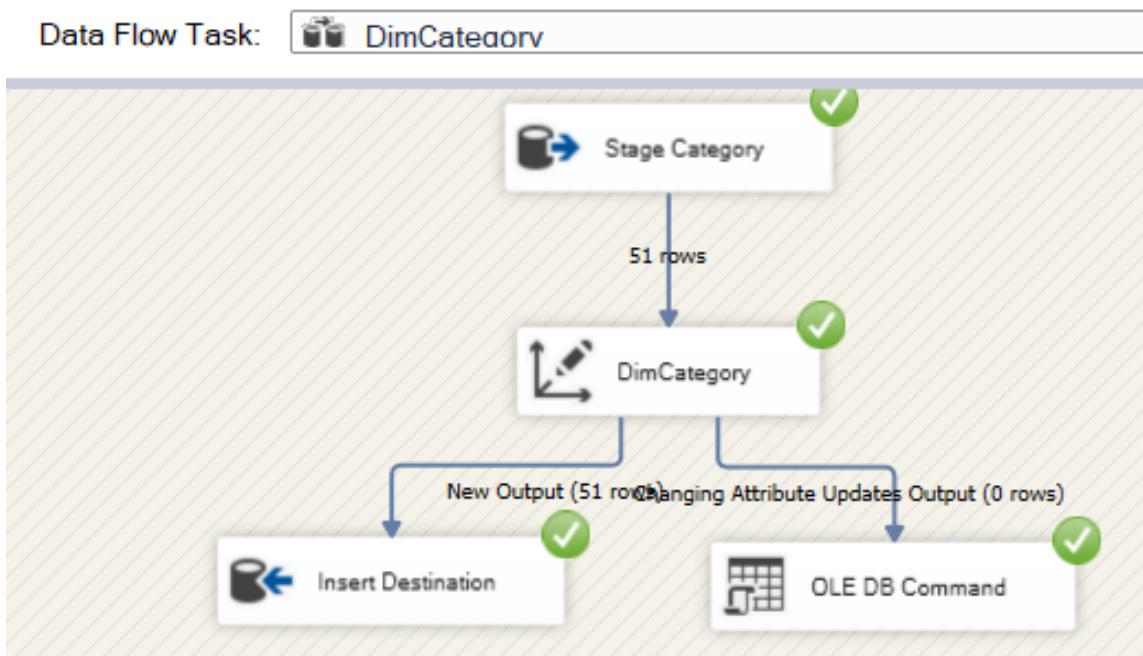
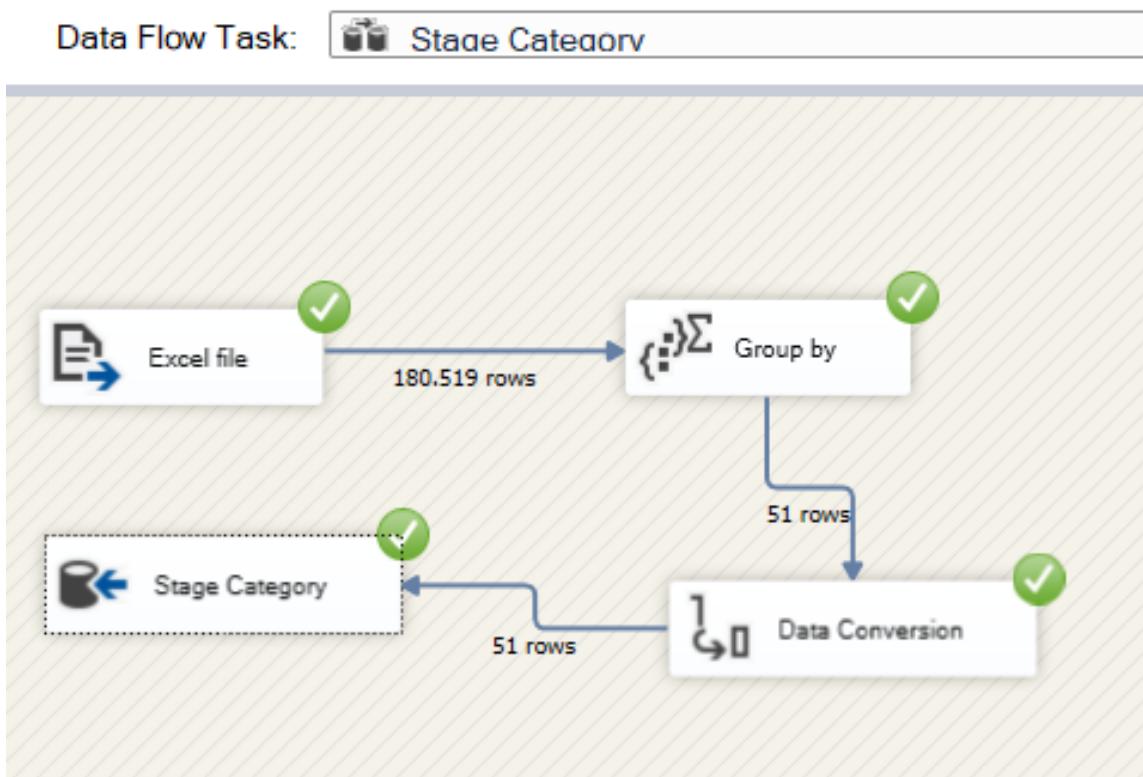
Data Flow Task: Stage City



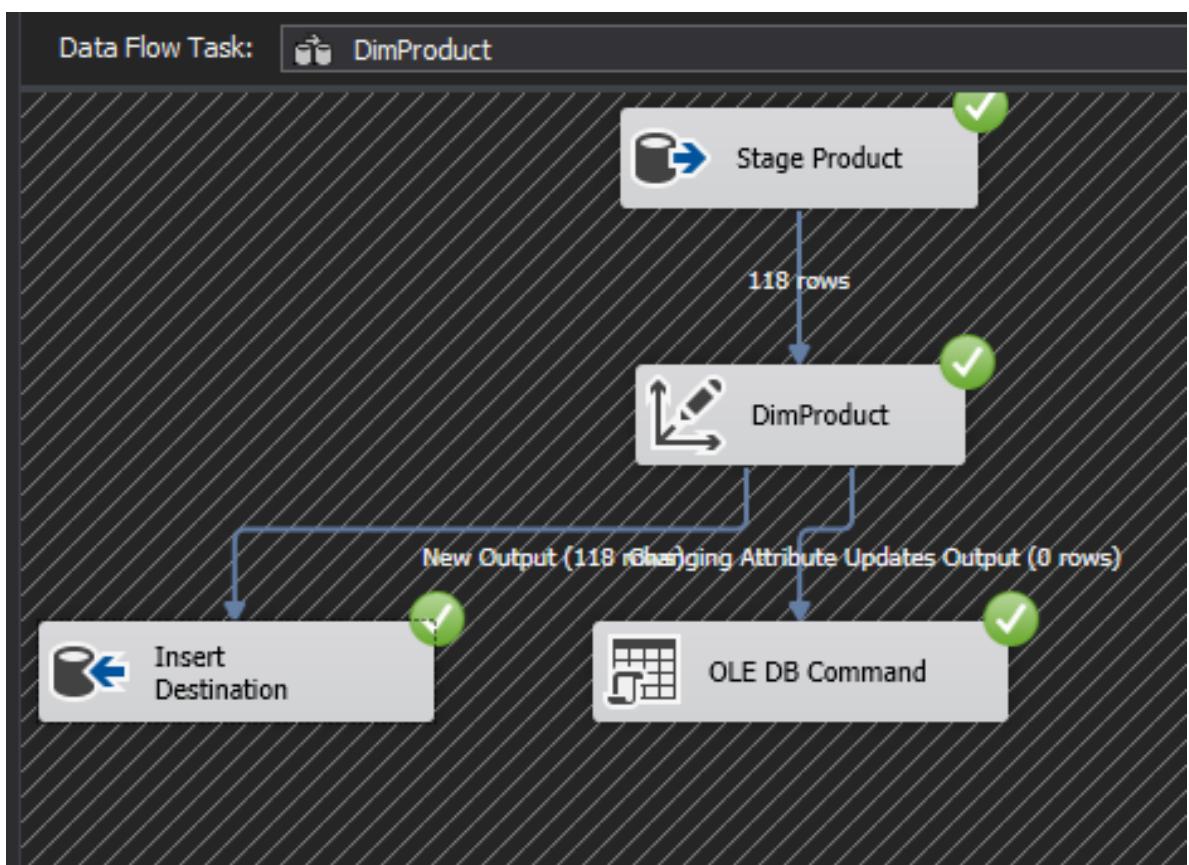
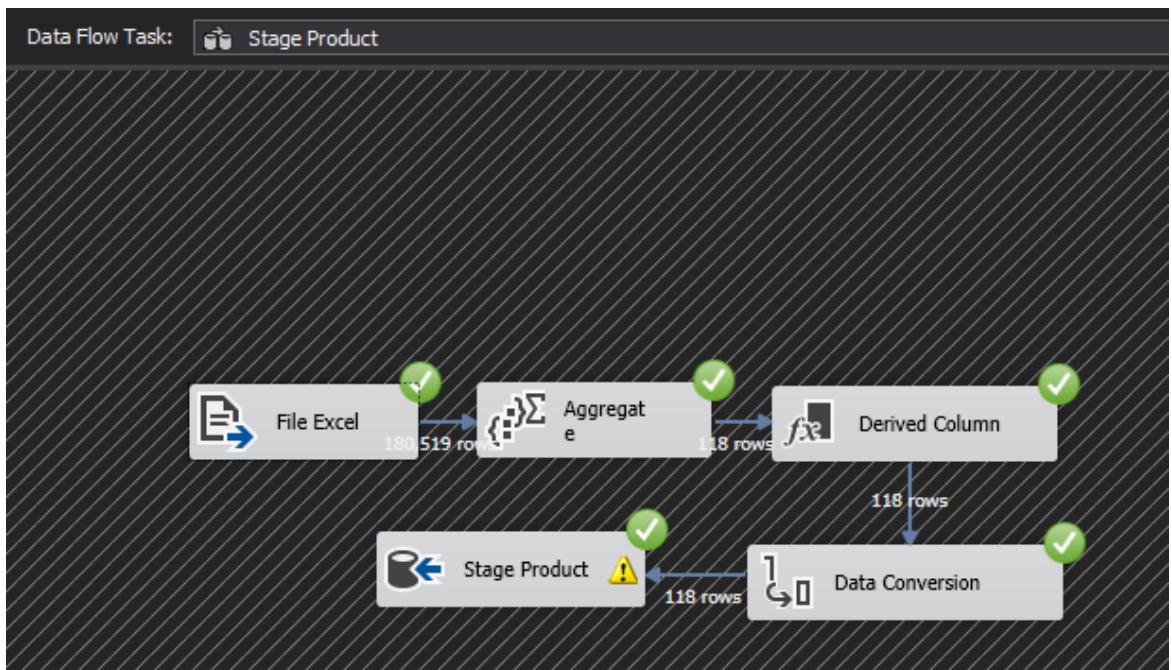
Data Flow Task: DimCity



### 3.4.1.10. Category



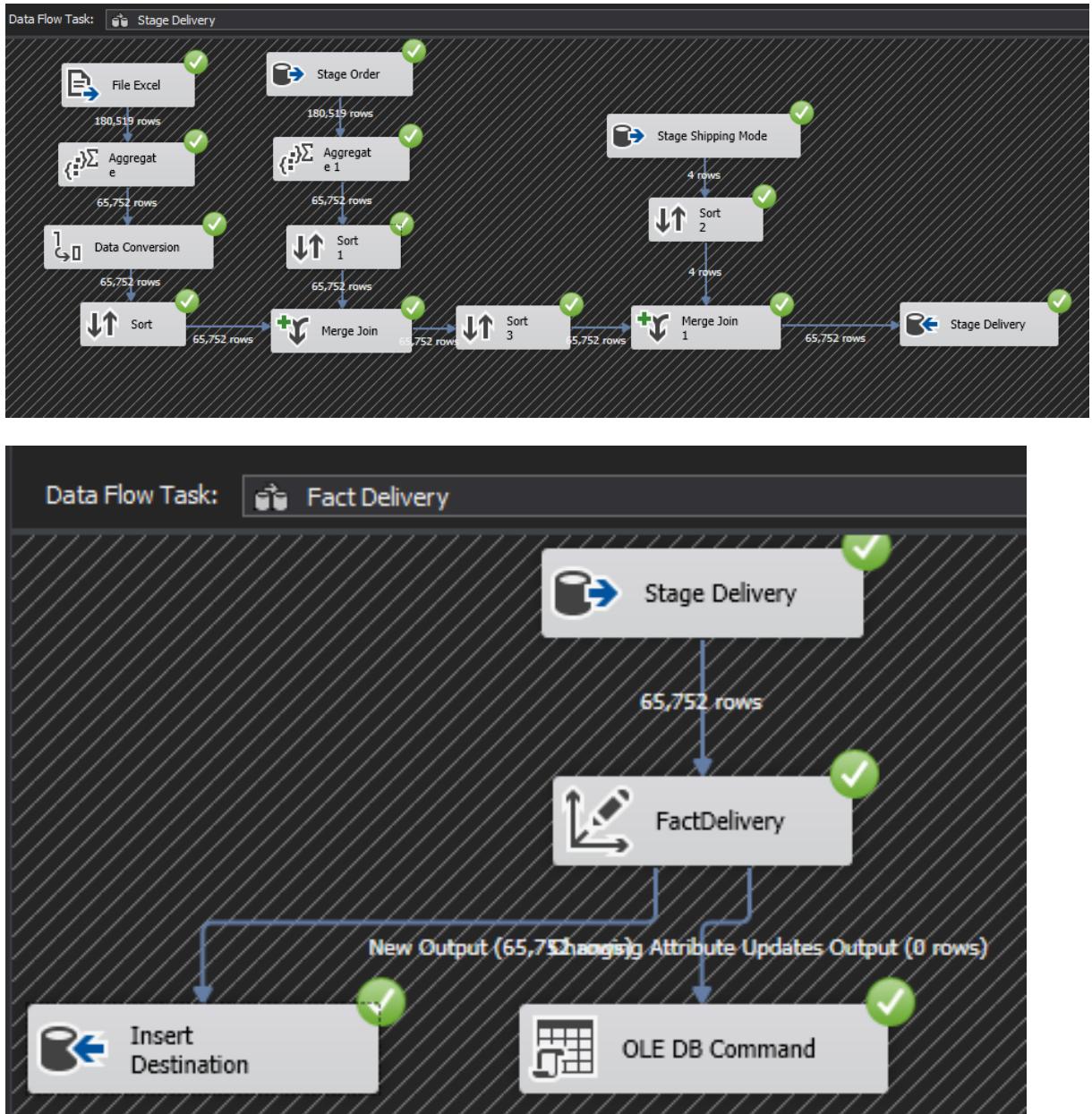
### 3.4.1.11. Product



### **3.4.1.12. Sales**



### 3.4.1.13. Delivery



### 3.4.2. Kết quả dữ liệu

#### 3.4.2.1. Department

	ID_Depart...	Name_Dep...
▶	2	Fitness
	3	Footwear
	4	Apparel
	5	Golf
	6	Outdoors
	7	Fan Shop
	8	Book Shop
	9	Discs Shop
	10	Technology
	11	Pet Shop
	12	Health and ...
*	NULL	NULL

#### 3.4.2.2. ShippingMode

	ID_Shippin...	Days_Ship...	Shipping_...
▶	1	1	First Class
	2	4	Standard Cl...
	3	2	Second Class
	4	0	Same Day
*	NULL	NULL	NULL

### 3.4.2.3. Segment

LAPTOP-ER9IVPEV....- dbo.DimSegment		
	ID_Segment	Name_Seg...
▶	1	Corporate
	2	Home Office
	3	Consumer
*	NULL	NULL

### 3.4.2.4. Customer

	ID_Customer	Fname_Cus...	Lname_Cus...	Street_Customer	Zipcode_C...	ID_Segment	ID_City
▶	1	Richard	Hernandez	6303 Heather Plaza	78521	3	1139
	2	Mary	Barrett	9526 Noble Embers Ridge	80126	3	1096
	3	Ann	Smith	3422 Blue Pioneer Bend	725	3	3741
	4	Mary	Jones	8324 Little Common	92069	3	1150
	5	Robert	Hudson	10 Crystal River Mall	725	2	3741
	6	Mary	Smith	3151 Sleepy Quail Promenade	7055	3	1148
	7	Melissa	Wilcox	9453 High Concession	725	1	3741
	8	Megan	Smith	3047 Foggy Forest Plaza	1841	1	1328
	9	Mary	Perez	3616 Quaking Street	725	3	3741
	10	Melissa	Smith	8598 Harvest Beacon Plaza	22554	1	1106
	11	Mary	Huffman	3169 Stony Woods	725	3	3741
	12	Christopher	Smith	5594 Jagged Embers By-pass	78227	1	1081
	13	Mary	Baldwin	7922 Iron Oak Gardens	725	2	3741
	14	Katherine	Smith	5666 Hazy Pony Square	90660	1	1373
	15	Jane	Luna	673 Burning Glen	92336	1	1642
	16	Tiffany	Smith	6651 Iron Port	725	1	3741
	17	Mary	Robinson	1325 Noble Pike	48180	3	1391
	18	Robert	Smith	2734 Hazy Butterfly Circle	94553	3	1212
	19	Stephanie	Mitchell	3543 Red Treasure Bay	725	2	3741
	20	Mary	Ellis	4703 Old Route	7093	3	1422
	21	William	Zimmerman	3323 Old Willow Mall	725	3	3741

### 3.4.2.5. Time

	ID_Time	OrderYear	OrderQuart...	OrderMonth	OrderDay	ShipYear	ShipQuarter	ShipMonth	ShipDay
▶	1	2015	1	1	1	2015	1	3	1
	2	2015	1	1	1	2015	1	3	1
	3	2015	1	1	1	2015	1	3	1
	4	2015	1	1	1	2015	1	3	1
	5	2015	1	1	1	2015	1	3	1
	6	2015	1	1	1	2015	1	3	1
	7	2015	1	1	1	2015	1	3	1
	8	2015	1	1	1	2015	1	3	1
	9	2015	1	1	1	2015	1	3	1
	10	2015	1	1	1	2015	1	3	1
	11	2015	1	1	1	2015	1	3	1
	12	2015	1	1	1	2015	1	3	1
	13	2015	1	1	1	2015	1	3	1
	14	2015	1	1	1	2015	1	3	1
	15	2015	1	1	1	2015	1	3	1
	16	2015	1	1	1	2015	1	3	1
	17	2015	1	1	1	2015	1	3	1
	18	2015	1	1	1	2015	1	3	1
	19	2015	1	1	1	2015	2	4	1
	20	2015	1	1	1	2015	2	4	1
	21	2015	1	1	1	2015	2	4	1
	22	2015	1	1	1	2015	2	4	1
	23	2015	1	1	1	2015	2	4	1
	24	2015	1	1	1	2015	2	4	1
	25	2015	1	1	1	2015	2	4	1

### 3.4.2.6. Market

LAPTOP-ER9IVPEV.D...I - dbo.DimMarket		
	ID_Market	Name_Mar...
▶	1	Pacific Asia
	2	Africa
	3	Europe
	4	LATAM
	5	USCA
*	NULL	NULL

### 3.4.2.7. Region

## LAPTOP-ER9IVPEV.D...I - dbo.DimRegion

	ID_Region	Name_Regi...	ID_Market
▶	1	South Asia	1
	2	West of USA	5
	3	South Amer...	4
	4	Central Asia	1
	5	Caribbean	4
	6	Canada	5
	7	Central Africa	2
	8	Oceania	1
	9	Southern Eu...	3
	10	Western Eur...	3
	11	Eastern Asia	1

### 3.4.2.8. Country

## LAPTOP-ER9IVPEV.D...- dbo.DimCountry

	ID_Country	Name_Cou...
▶	1	EE. UU.
	2	Puerto Rico
	3	Estonia
	4	Burundi
	5	Camerún
	6	Namibia
	7	Bután
	8	Irak
	9	Togo
	10	Hong Kong
	11	MÁvico

### 3.4.2.9. City

	ID_City	ID_Country	Name_State	Name_City
▶	1	94	Herat	Herat
	2	94	Nangarhar	Jalalabad
	3	94	Kandahar	Kandahar
	4	94	Kabul	Kabul
	5	109	Shkodër	Shkoder
	6	109	Vlorë	Vlore
	7	109	Durrës	Durres
	8	109	Korçë	Korce
	9	109	Elbasan	Elbasan
	10	142	Renania del...	Gummersba...
	11	142	Baden-Wurt...	Heilbronn
	12	142	Bavaria	Rosenheim
	13	142	Bremen	Bremerhaven
	14	142	Hesse	Wiesbaden
	15	142	Baja Sajonia	Garbsen
	16	142	Sajonia-Anh...	Naumburg
	17	142	Baden-Wurt...	Freiburg
	10	142	Renania del...	Potzdam

### 3.4.2.10. Category

	ID_Category	Name_Cate...
▶	2	Soccer
	3	Baseball & ...
	4	Basketball
	5	Lacrosse
	6	Tennis & Ra...
	7	Hockey
	9	Cardio Equi...
	10	Strength Tra...
	11	Fitness Acce...
	12	Boxing & M...
	13	Electronics
	16	As Seen on ...
	17	Closets

### 3.4.2.11. Product

LAPTOP-ER9IVPEV.D... - dbo.DimProduct				
	ID_Product	Name_Product	Price_Prod...	ID_Category
►	19	Nike Men's Fingertra	124,9899979	2
	24	Elevation Training M	79,98999786	2
	35	adidas Brazuca 2014	159,9900055	3
	37	adidas Kids' F5 Mess	34,99000168	3
	44	adidas Men's F10 Mes	59,99000168	3
	58	Diamondback Boys' In	299,9899902	4
	60	SOLE E25 Elliptical	999,9899902	4
	61	Diamondback Girls' C	299,9899902	4
	78	Nike Kids' Grade Sch	99,98999786	5
	93	Under Armour Men's T	24,98999977	5
	116	Nike Men's Comfort 2	44,99000168	6
	127	Stiga Master Series	329,9899902	7
	134	Nike Women's Legend	25	7
	135	Nike Dri-FIT Crew So	22	7
	172	Nike Women's Tempo S	30	9

### 3.4.2.12. Sales

	ID_Order	ID_Customer	Type	Order_Status	ID_Product	Quantity	Discount	Sales	Item_Total
►	41385	3527	DEBIT	COMPLETE	191	5	10	499.9500122	489.9500122
	41385	3527	DEBIT	COMPLETE	365	5	53.99000168	299.9500122	245.9600062
	41385	3527	DEBIT	COMPLETE	365	2	15.60000038	119.9800034	104.3799972
	41400	1430	TRANSFER	SUSPECTED_FR...	403	1	26	129.9900055	103.9899972
	41400	1430	TRANSFER	SUSPECTED_FR...	403	1	32.5	129.9900055	97.48999782
	41400	1430	TRANSFER	SUSPECTED_FR...	502	2	9	100	91
	41401	11720	TRANSFER	PROCESSING	282	5	20.79000092	159.9499969	139.1600032
	41401	11720	TRANSFER	PROCESSING	365	5	47.99000168	299.9500122	251.9600062
	41401	11720	TRANSFER	PROCESSING	502	2	7	100	93
	41401	11720	TRANSFER	PROCESSING	502	5	17.5	250	232.5
	41401	11720	TRANSFER	PROCESSING	1004	1	4	399.980011	395.980011
	41402	5410	PAYMENT	PENDING_PAY...	191	5	0	499.9500122	499.9500122
	41402	5410	PAYMENT	PENDING_PAY...	502	1	2	50	48
	41402	5410	PAYMENT	PENDING_PAY...	1014	2	24.98999977	99.95999908	74.97000122
	41403	7278	TRANSFER	PENDING	249	3	11.53999996	164.9100037	153.3699952
	41403	7278	TRANSFER	PENDING	502	5	13.75	250	236.25
	41403	7278	TRANSFER	PENDING	1004	1	0	399.980011	399.980011
	41403	7278	TRANSFER	PENDING	1073	1	14	199.9900055	185.9900052
	41437	10418	TRANSFER	PROCESSING	502	2	5	100	95
	41437	10418	TRANSFER	PROCESSING	957	1	16.5	299.980011	283.480011
	41437	10418	TRANSFER	PROCESSING	1073	1	0	199.9900055	199.9900052
	41479	4582	CASH	CLOSED	403	1	0	129.9900055	129.9900052
	41479	4582	CASH	CLOSED	502	2	25	100	75
	41479	4582	CASH	CLOSED	705	2	20.00000005	20.0000001	20.0000002
<	1	of 200		▶	▶	▶	▶	▶	▶

### 3.4.2.13. Delivery

	ID_Order	Days_Shipping...	Day_For_Ship...	Delivery_Status	Late_delivery_r...	ID_Shipping_...	ID_Time	ID_City	ID_Region
►	1	2	4	Advance shippi...	0	2	1	3461	12
	2	3	4	Advance shippi...	0	2	19	712	3
	4	5	4	Late delivery	1	2	39	712	3
	5	6	4	Late delivery	1	2	50	712	3
	7	3	2	Late delivery	1	3	23	413	3
	8	4	4	Shipping on time	0	2	32	413	3
	9	5	4	Late delivery	1	2	43	413	3
	10	6	4	Late delivery	1	2	55	413	3
	11	2	4	Advance shippi...	0	2	16	413	3
	12	3	4	Advance shippi...	0	2	26	3518	12
	13	4	4	Shipping on time	0	2	35	3518	12
	14	5	4	Late delivery	1	2	46	3518	12
	15	6	4	Late delivery	1	2	56	398	3
	16	2	4	Advance shippi...	0	2	17	398	3
	17	3	4	Advance shippi...	0	2	27	3667	12
	18	4	4	Shipping on time	0	2	36	3667	12
	19	5	4	Late delivery	1	2	47	3667	12
	20	6	4	Late delivery	1	2	57	3667	12
	21	2	1	Late delivery	1	1	18	3455	12
	23	4	2	Late delivery	1	3	37	789	3
	24	5	2	Late delivery	1	3	48	3666	12
	25	6	2	Late delivery	1	3	58	3468	12
	27	3	2	Late delivery	1	3	28	3387	12
	29	·	·	Shipping on time	0	2	29	3112	12
<	1	of 200		▶	▶	▶	▶	▶	▶

# Chương 4: Phân tích dữ liệu SSAS

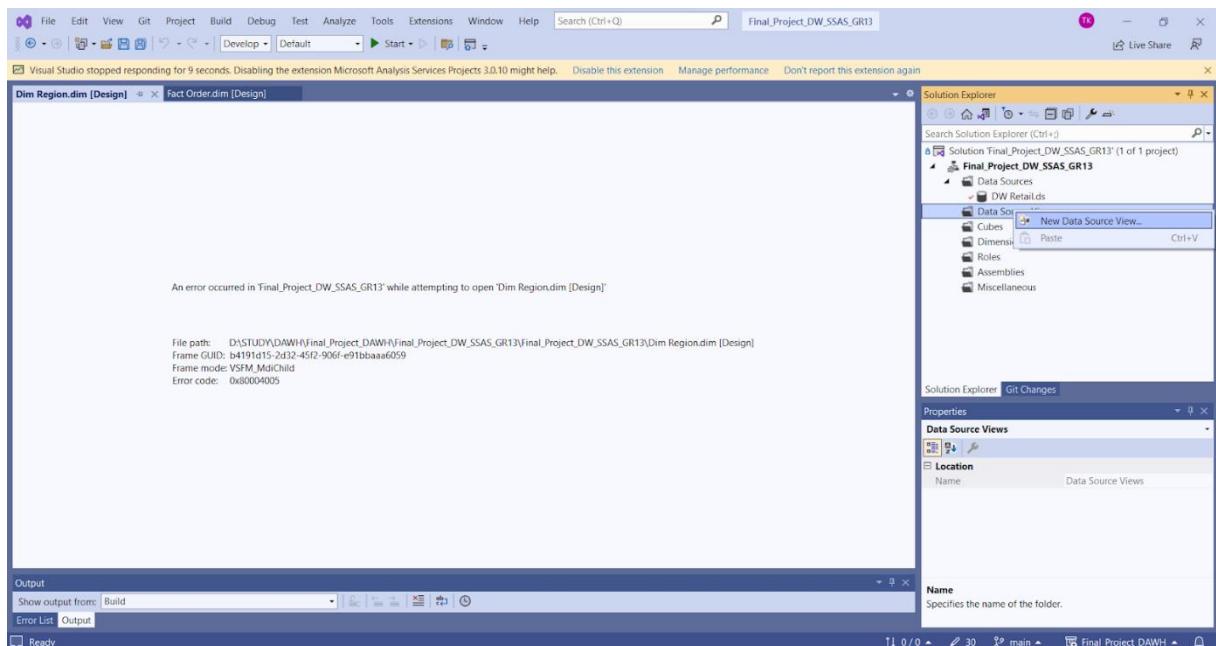
Chương này, dùng công cụ SSAS xây dựng mô hình và xây dựng khối - cube và phân tích dữ liệu trên Cube. Sử dụng Power BI để trực quan các kết quả phân tích trên SSAS

## 4.1. Một số câu hỏi đặt ra để phân tích

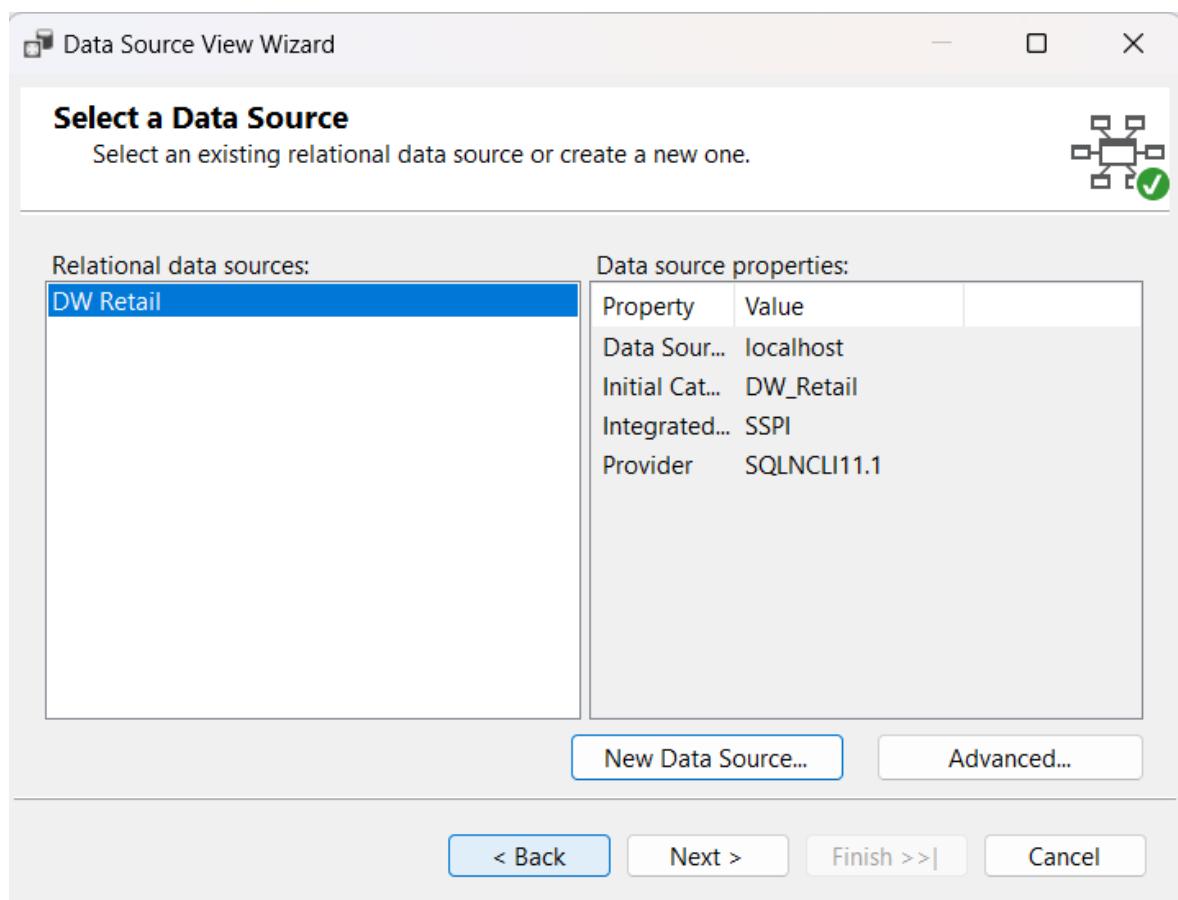
- Báo cáo tổng doanh thu theo quý, năm.
- Báo cáo tổng lợi nhuận theo quý, năm.
- Thông kê tình trạng giao hàng trễ của các khu vực
- Thông kê các đơn hàng bị hủy theo quốc gia
- So sánh doanh thu giữa các khu vực theo năm
- Thị trường nào có doanh thu nhiều nhất và thị trường nào có doanh thu ít nhất?
- Sản phẩm nào bán chạy nhất trong một thời gian nhất định?
- Loại mặt hàng nào được ưa chuộng nhất theo từng thị trường
- Tháng nào trong năm thường có mức độ bán hàng cao nhất?
- Danh mục sản phẩm nào được ưa chuộng nhất?
- So sánh các phương thức vận chuyển?

## 4.2. Quá trình xây dựng mô hình

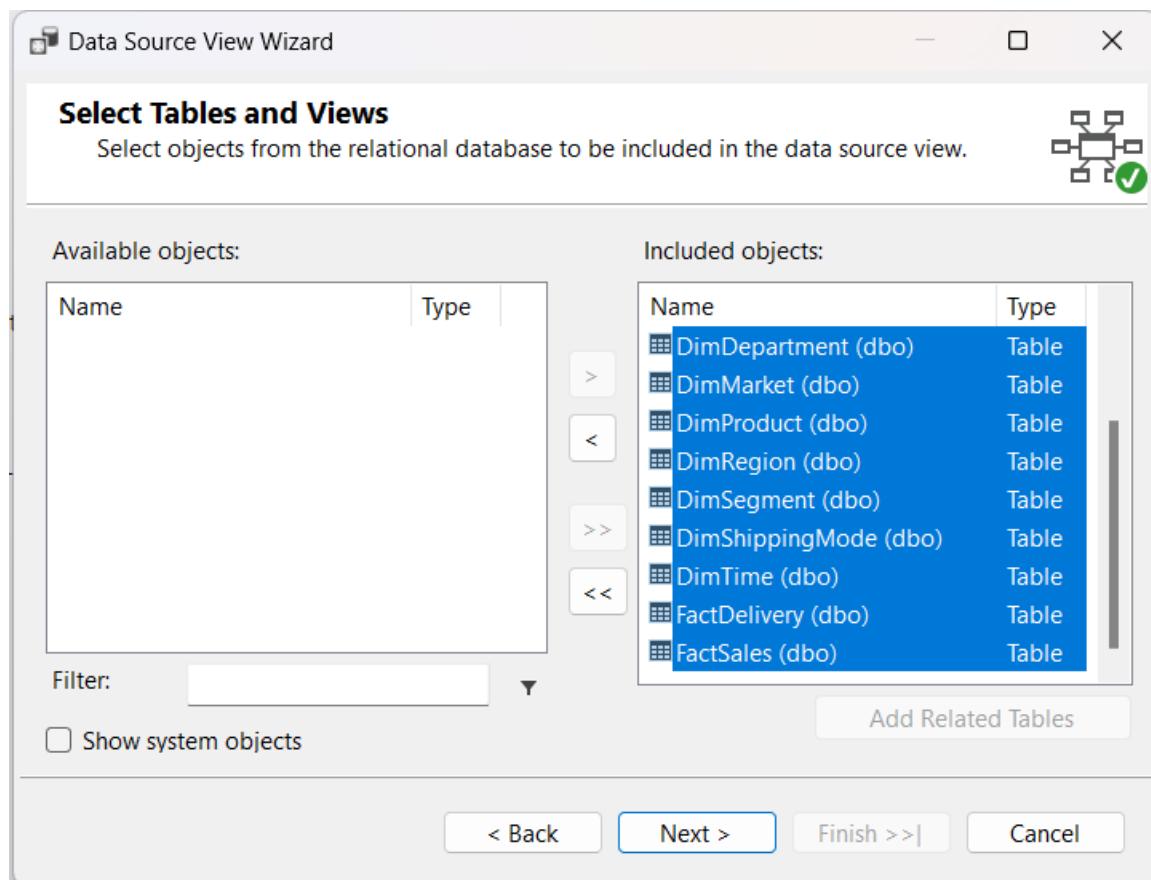
Chọn New Data Source View ...



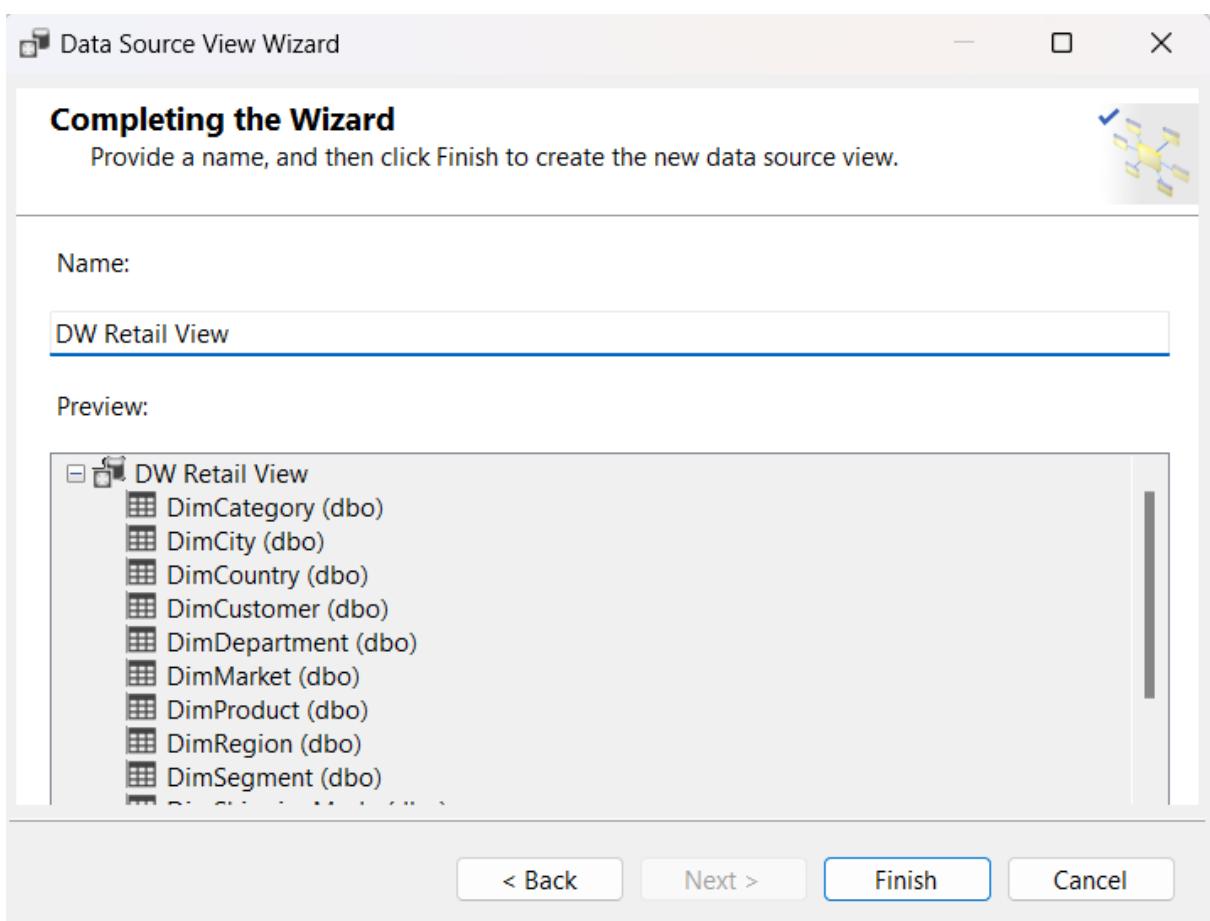
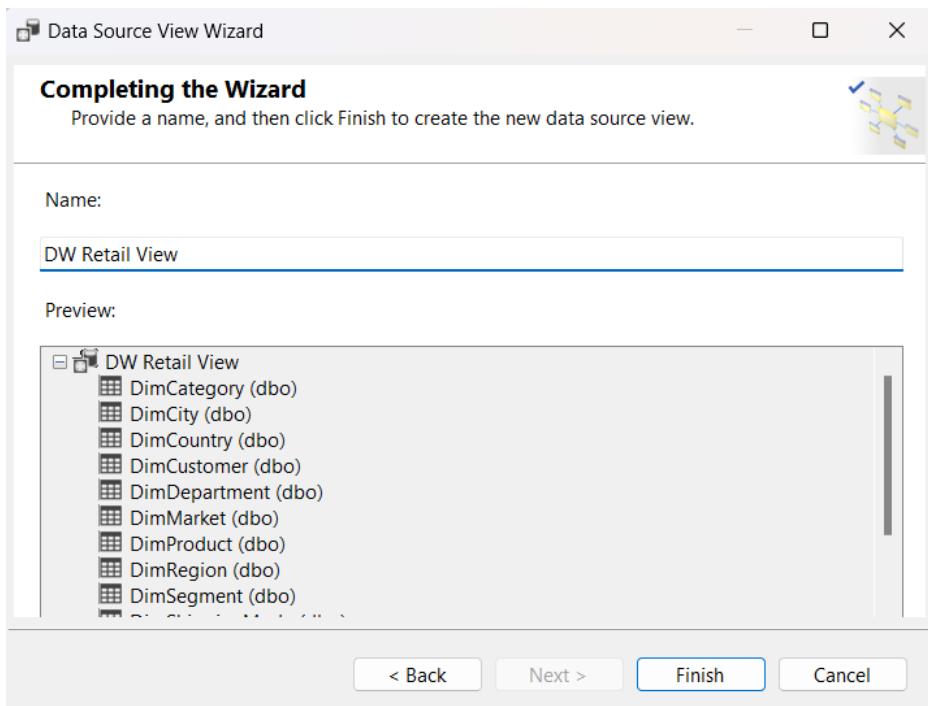
Nhấn chọn **Data Source** đã tạo ở trước và nhấn **Next:**

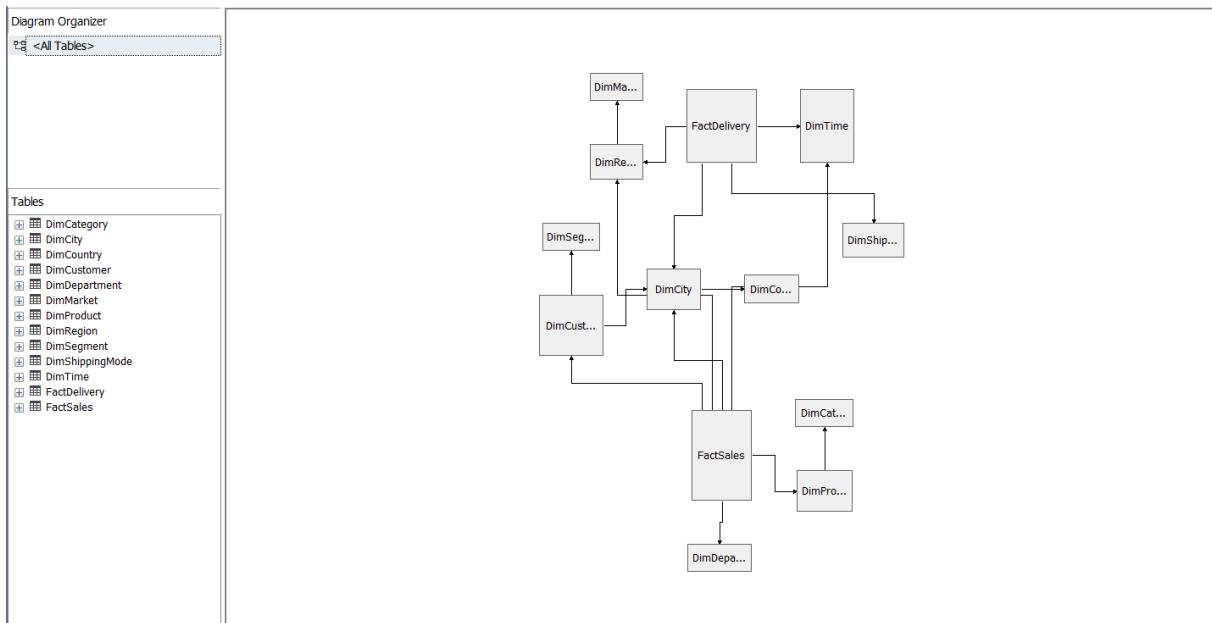


Chọn bảng fact qua và nhấn **Next::**



Nhập tên **Data Source View** và nhấn Finish để hoàn tất:

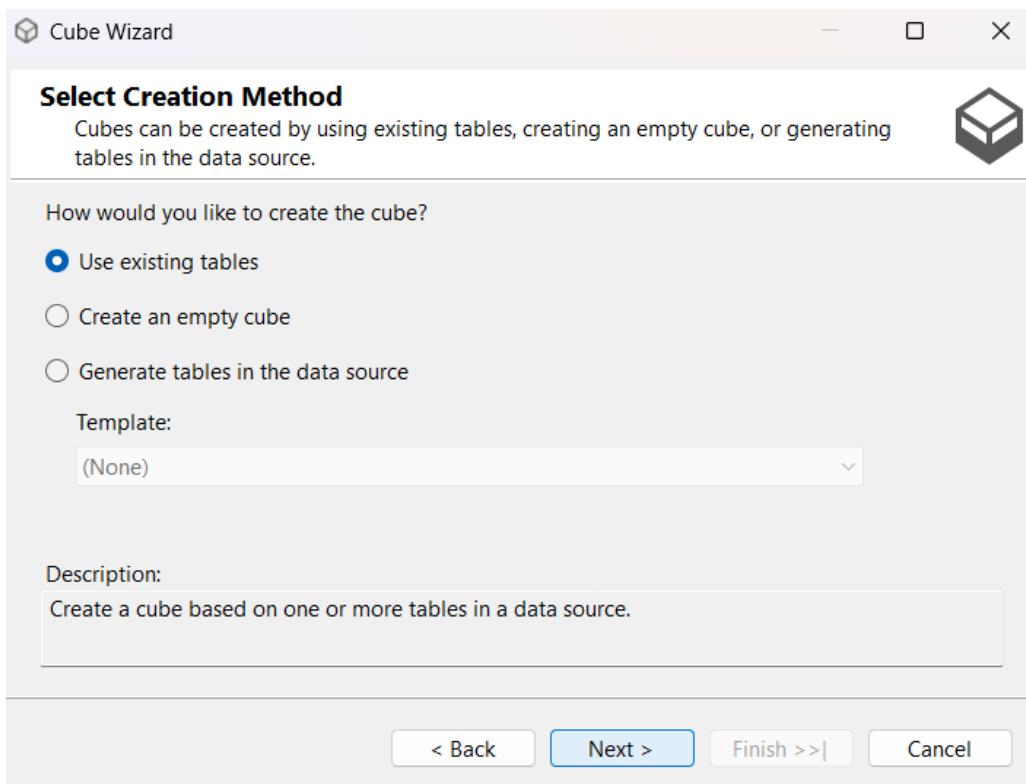




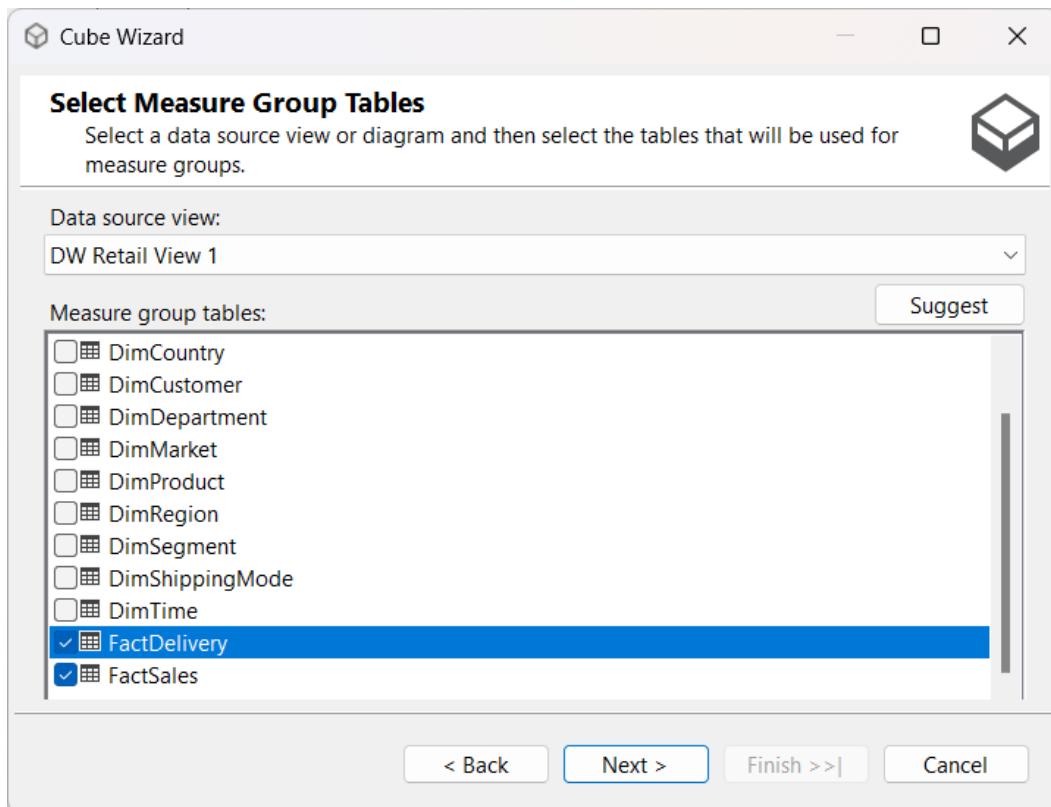
### 4.3. Quá trình xây dựng khối - Cube

Nhấn chuột phải vào **Cubes** chọn **New Cube...** dialog hiện ra và nhấn **Next**

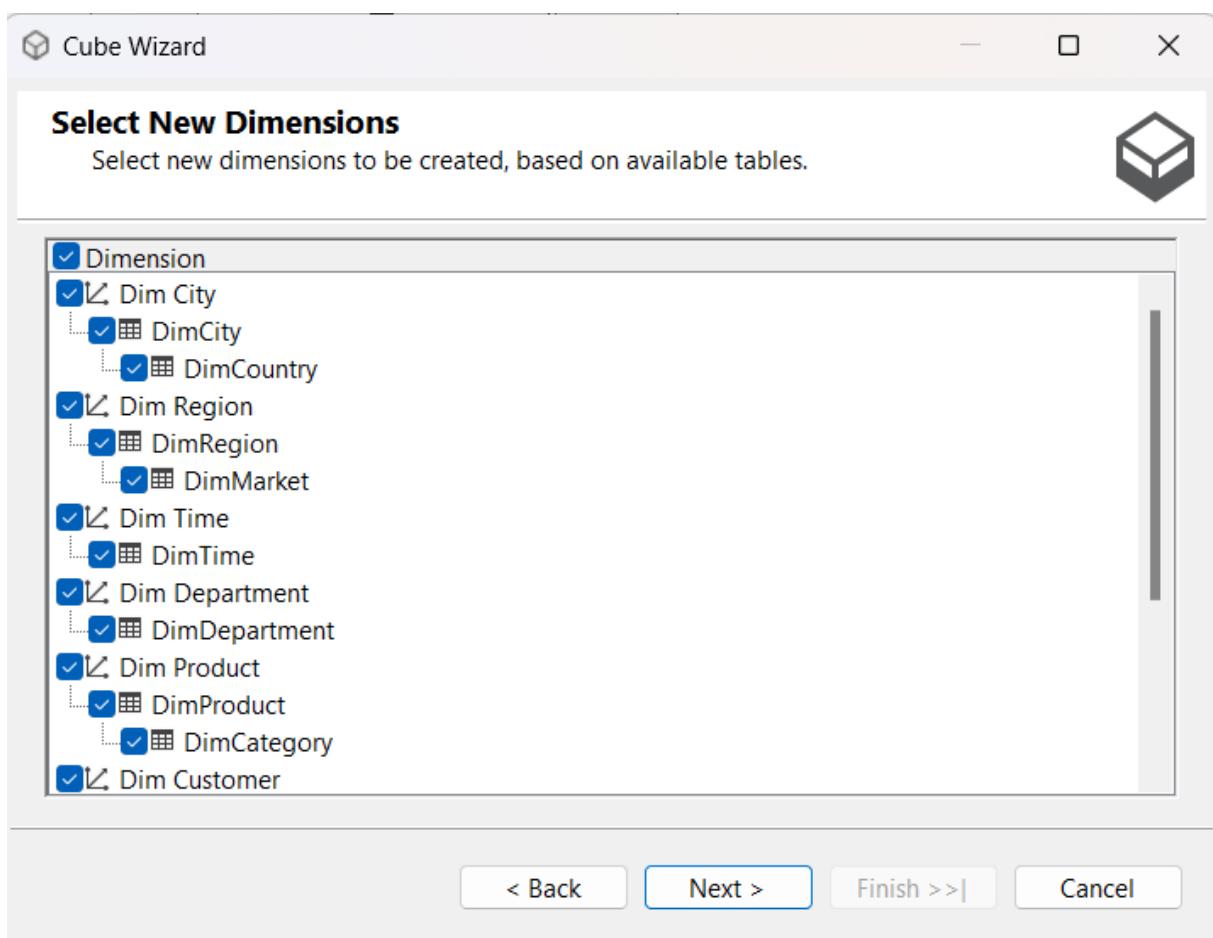
Chọn sử dụng bảng có sẵn(**Use existing table**) và nhấn **Next**:



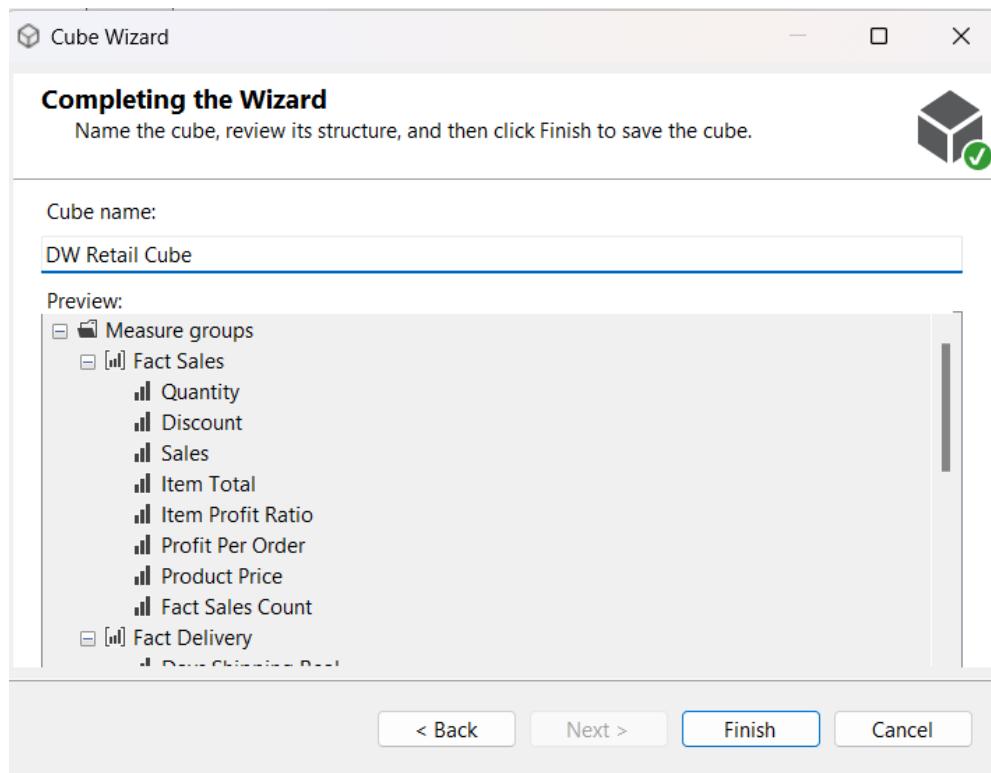
Chọn **Measure Group Table** và nhấn **Next**:



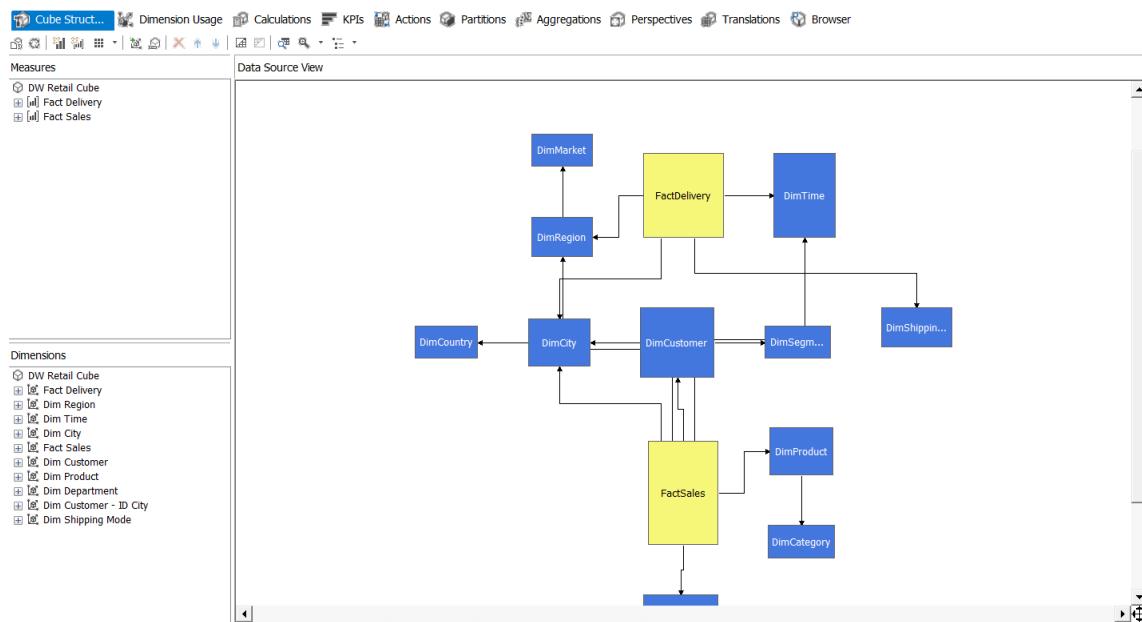
Chọn Dimension:



Nhập tên **Cube** và nhấn **Finish** để hoàn tất:



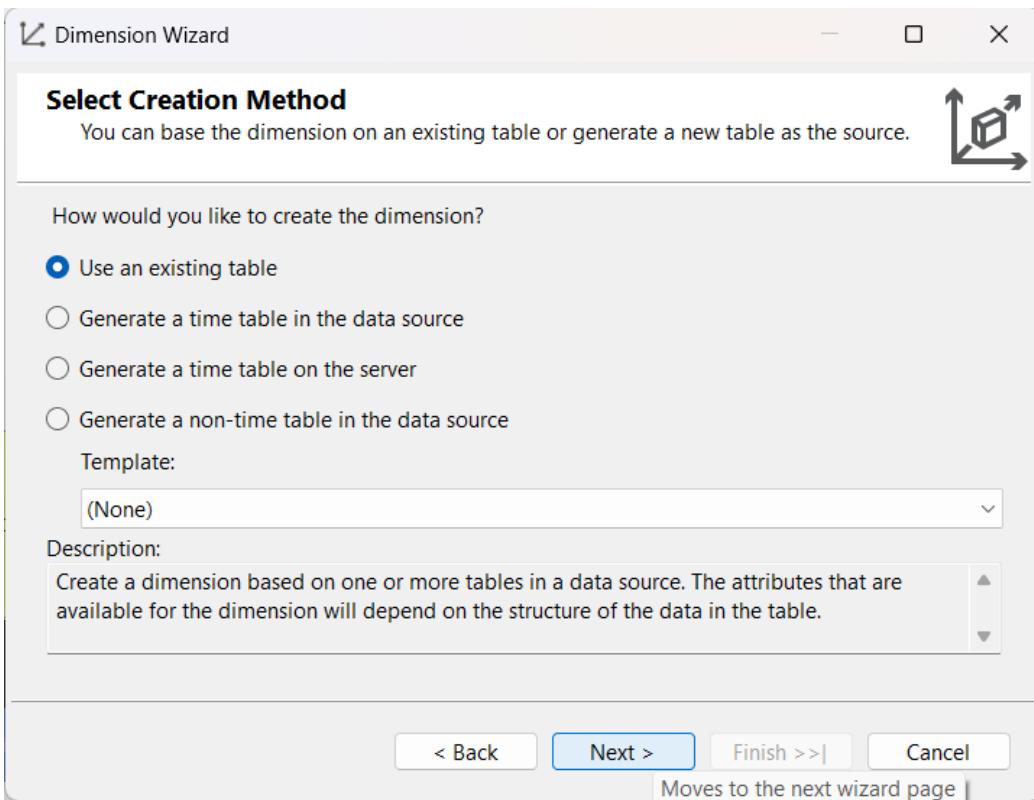
**Cube** đã tạo:



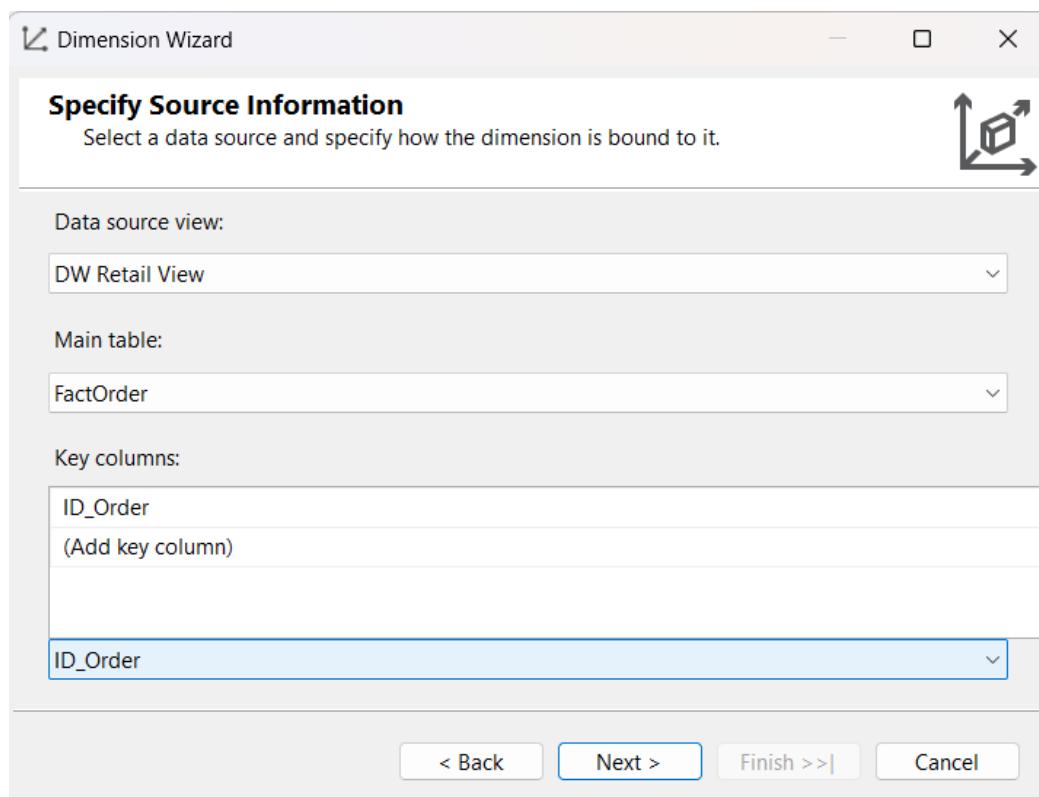
Thêm Dimensions DimDeliveryStatus

Chuột phải vào Dimensions chọn NewDimensions nhấn next:

Chọn sử dụng bảng có sẵn(Use existing table) và nhấn Next:



Chọn Main Table là Fact Order, nhấn next:



Chọn các thuộc tính cần thiết, nhấn next:

Dimension Wizard

### Select Dimension Attributes

Specify dimension attributes and select Enable Browsing to surface them as hierarchies.



Available attributes:

Attribute Name	Enable Browsing	Attribute Type
<input checked="" type="checkbox"/> Attribute Name	<input checked="" type="checkbox"/>	Regular
<input checked="" type="checkbox"/> ID Order	<input checked="" type="checkbox"/>	Regular
<input type="checkbox"/> Item Total	<input type="checkbox"/>	Regular
<input type="checkbox"/> Type	<input type="checkbox"/>	Regular
<input checked="" type="checkbox"/> Delivery Status	<input checked="" type="checkbox"/>	Regular
<input type="checkbox"/> Late Delivery Risk	<input type="checkbox"/>	Regular
<input type="checkbox"/> Days Shipping Real	<input type="checkbox"/>	Regular
<input type="checkbox"/> Order Profit	<input type="checkbox"/>	Regular
<input type="checkbox"/> Order Status	<input type="checkbox"/>	Regular
<input type="checkbox"/> Market	<input type="checkbox"/>	Regular
<input type="checkbox"/> ID Customer	<input type="checkbox"/>	Regular
<input type="checkbox"/> Enname Customer	<input type="checkbox"/>	Regular

< Back Next > Finish >> Cancel

Đổi tên Dimensions, nhấn Finish:

Dimension Wizard

### Completing the Wizard

Type a name for the new dimension, verify the dimension structure, and then click Finish to save the dimension.



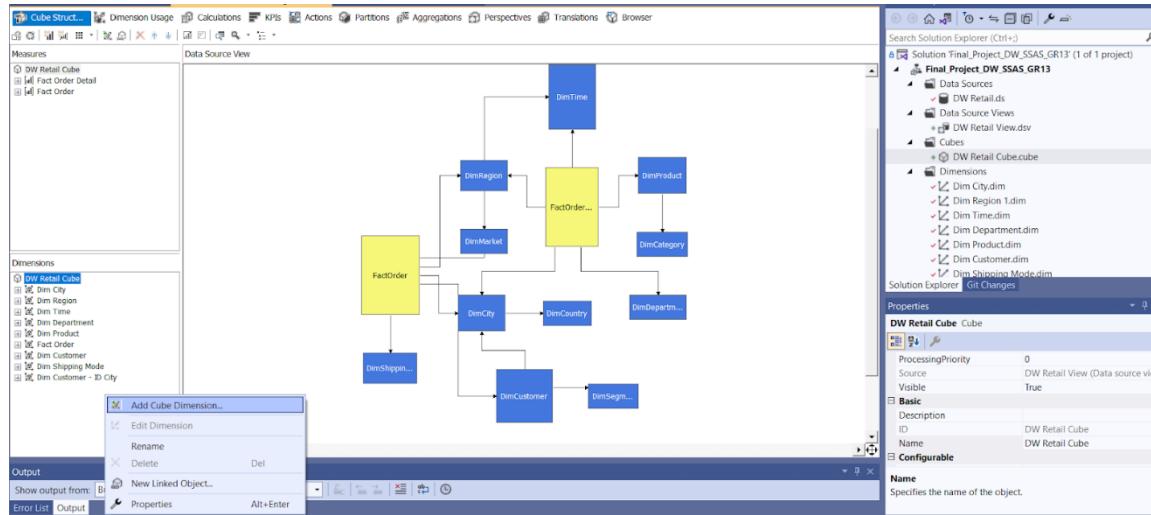
Name:

Preview:

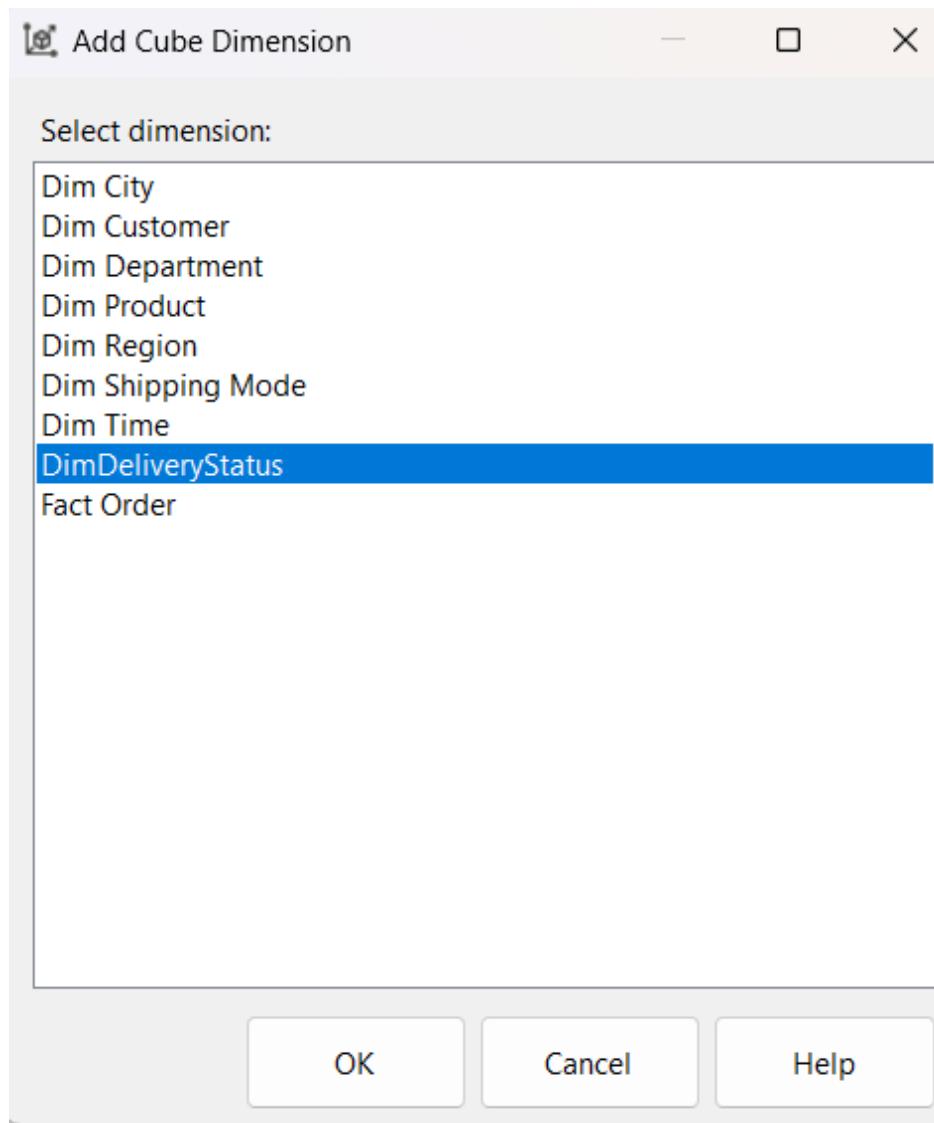
- Fact Order 1
  - Attributes
    - ID Order
    - Delivery Status

## Thêm Dimension vào Cube

Chuột phải vào tab Dimensions của Cube, nhấn Add Cube Dimension:

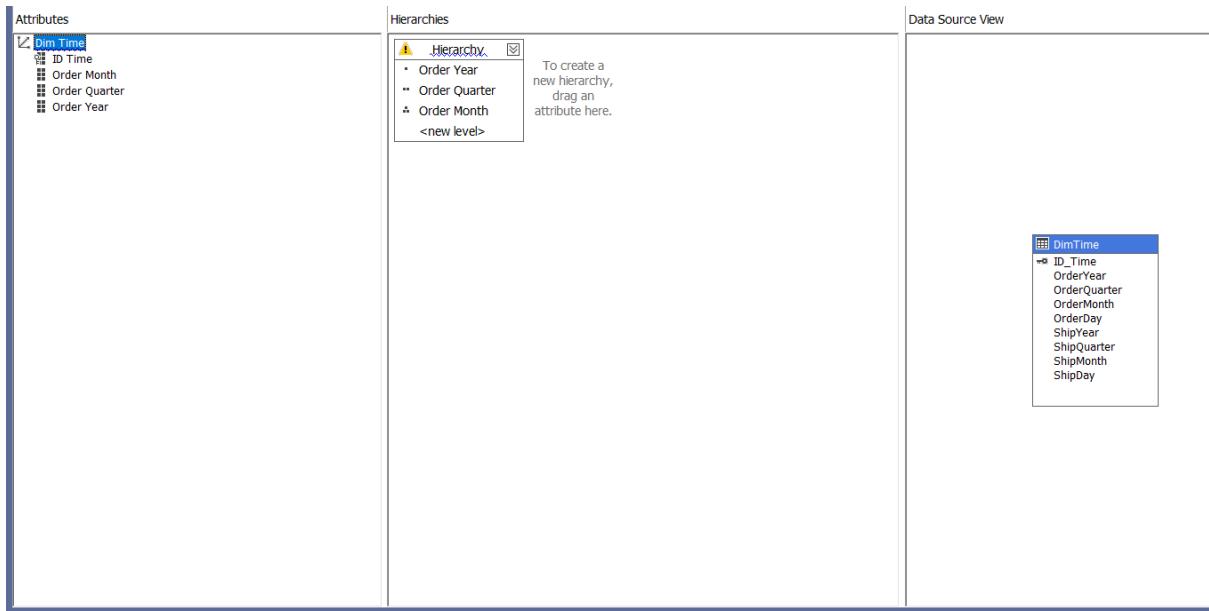


Chọn Dimensions, nhấn OK:



## Cấu hình Dimension Structure của các Dimensions

### - Dim Time



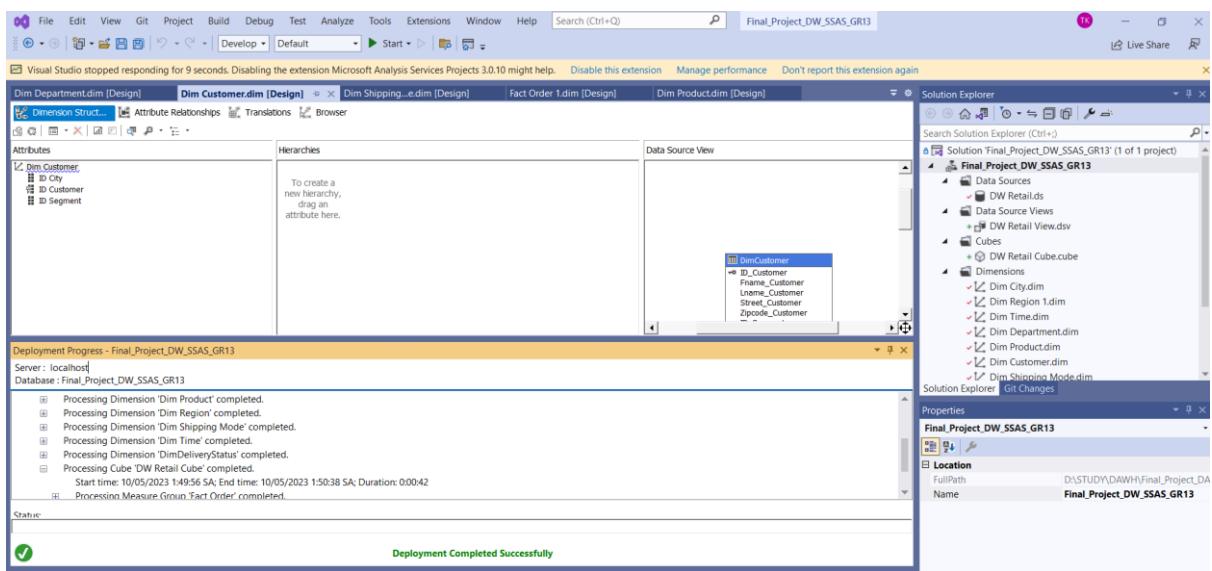
### - Dim Product



### - Dim Region



## Deploy dự án:



## 4.4. Thực hiện truy vấn(SSAS và Power BI)

### 4.4.1. Báo cáo tổng doanh thu theo quý, năm.

- SSAS

Screenshot of the SSAS cube browser interface showing the DW Retail Cube structure and a query results grid.

**Measure Group:**

- <All>
- DW Retail Cube
  - Measures
  - KPIs
  - Dim City
  - Dim Customer
  - Dim Customer - ID City
  - Dim Delivery
    - Delivery Status
    - Members
    - Delivery Status
    - ID Order
  - Dim Department
  - Dim Product
  - Dim Region
  - Dim Shipping Mode
  - Dim Time
  - Fact Delivery
  - Fact Sales

**Calculated Members**

**Dimension Hierarchy Operator Filter Expression**

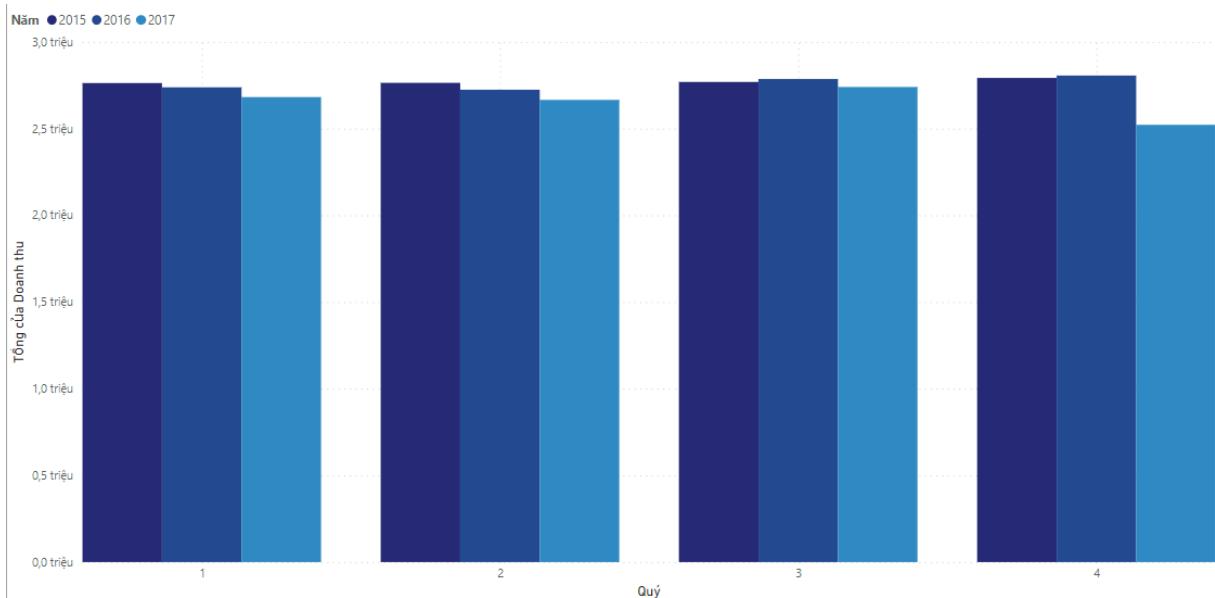
Dimension	Hierarchy	Operator	Filter Expression
Dim Delivery	Delivery Status	Not Equal	{ Shipping canceled }

**<Select dimension>**

**Order Year Order Quarter Item Total**

Order Year	Order Quarter	Item Total
2015	1	2763081,1006932
2015	2	2764260,60041693
2015	3	2769484,25112291
2015	4	2792717,60073736
2016	1	2738489,12065065
2016	2	2724646,42027475
2016	3	2786446,57105227
2016	4	2806414,65185902
2017	1	2682156,58775214
2017	2	2666054,55836372
2017	3	2740445,37748518
2017	4	2522253,48881328
2018	1	182822,780351235
2018	2	45816,18998918
2018	3	15584,45993419
2018	4	53728,62071996

- Power BI: doanh số bán hàng ở các quý trong năm tương đối đồng đều không có sự chênh lệch quá lớn chỉ có quý 4 năm 2017 doanh thu có sự sụt giảm nhẹ



#### 4.4.2. Báo cáo tổng lợi nhuận theo quý, năm.

- SSAS

DW Retail Cube

Metadata

Search Model

Measure Group:

<All>

DW Retail Cube

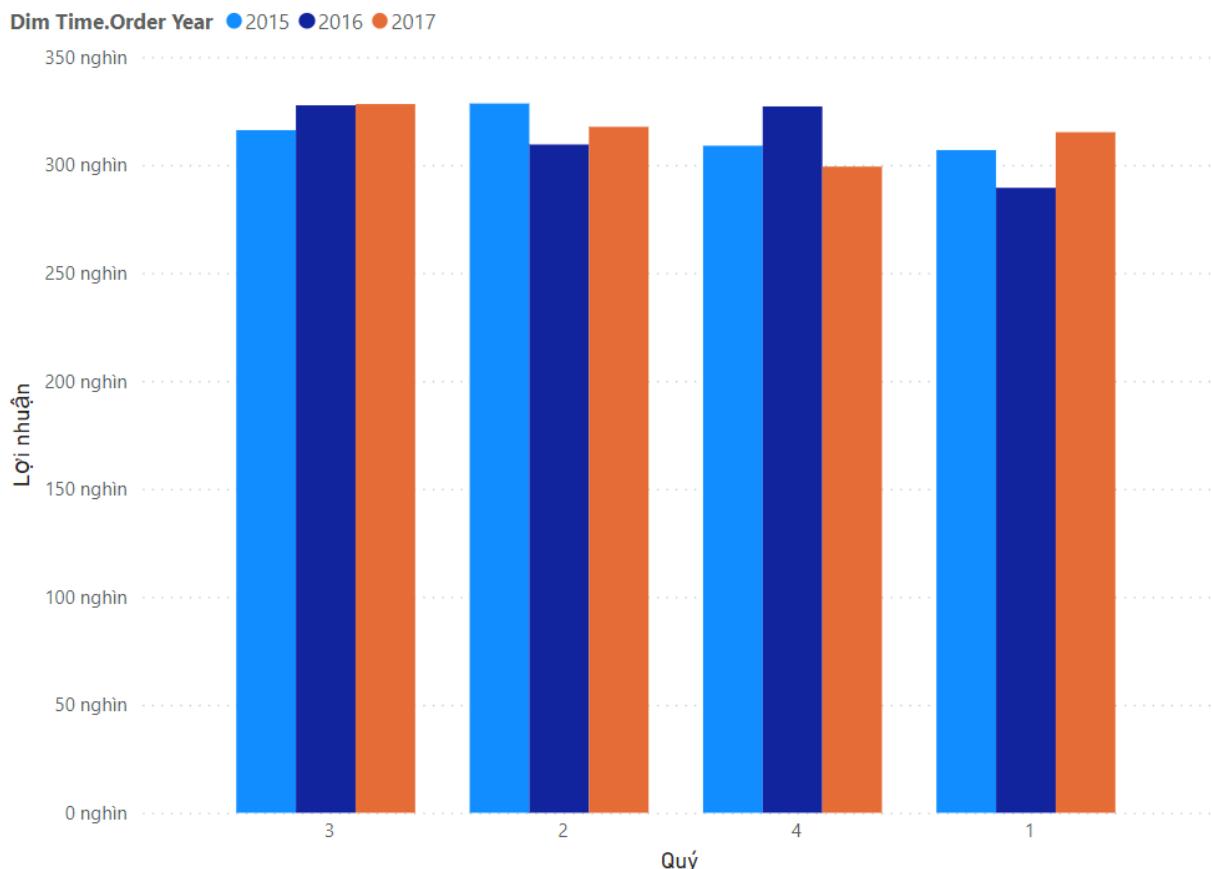
- Measures
  - Fact Delivery
  - Fact Sales
    - Discount
    - Fact Sales Count
    - Item Profit Ratio
    - Item Total
    - Product Price
    - Profit Per Order
    - Quantity
    - Sales
- KPIs
- Dim City
- Dim Customer
- Dim Customer - ID City
- Dim Delivery
  - Delivery Status

Calculated Members

Dimension	Hierarchy	Operator	Filter Expression
Dim Delivery	Delivery Status	Not Equal	{ Shipping canceled }
<Select dimension>			

Order Year	Order Quarter	Profit Per Order
2015	1	323620,030441477
2015	2	344170,610419629
2015	3	328459,78021791
2015	4	322606,480191493
2016	1	303979,291071538
2016	2	322680,75017461
2016	3	340660,740382006
2016	4	342798,289843101
2017	1	326766,660335845
2017	2	333480,640552219
2017	3	338751,820399159
2017	4	305085,990044306
2018	1	24820,919995837
2018	2	3909,839928751
2018	3	1688,379987851
2018	4	3422,750064647

- Power BI: tuy danh số không có quá nhiều sự chênh lệch như lợi nhuận ở quý 1 năm 2017 có sự tăng trưởng hơn hai năm trước nhưng ở quý 4 năm 2017 lại bị sụt giảm đáng kể so với các 2 năm trước.



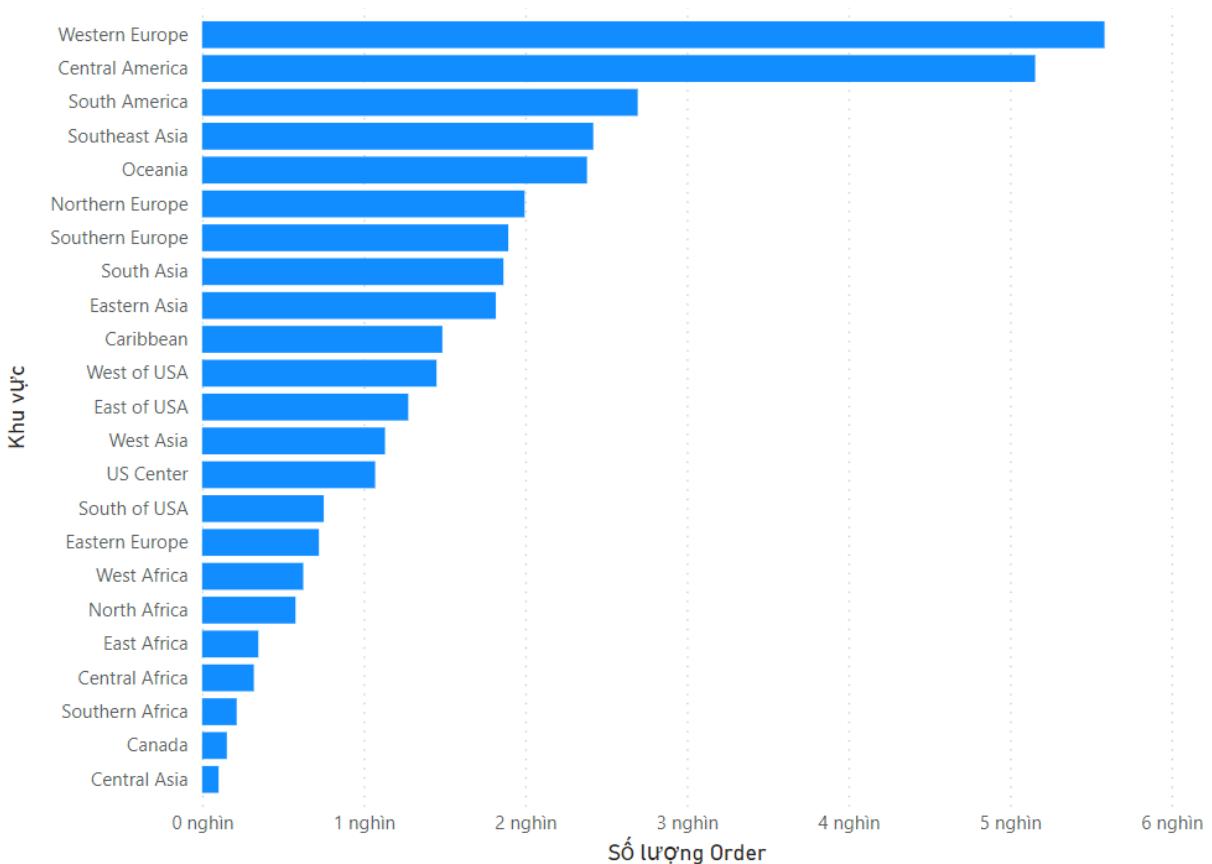
#### 4.4.3. Thống kê tình trạng giao hàng trễ của các khu vực

- SSAS

Screenshot of the SSAS Dimension browser interface showing the DW Retail Cube. The left pane displays the dimension hierarchy for Fact Sales, KPIs, Dim City, Dim Customer, Dim Customer - ID City, Dim Delivery, Dim Department, Dim Product, Dim Region, Dim Shipping Mode, and Dim Time. The right pane shows a table titled "Fact Delivery Count" with data for various regions:

Name Region	Fact Delivery Count
Canada	153
Caribbean	1487
Central Africa	320
Central Amer...	5156
Central Asia	102
East Africa	348
East of USA	1276
Eastern Asia	1818
Eastern Eur...	723
North Africa	578
Northern Eu...	1996
Oceania	2382
South America	2696
South Asia	1865
South of USA	753
Southeast Asia	2421
Southern Afr...	214
Southern Eu...	1895
US Center	1071
West Africa	626
West Asia	1132
West of USA	1451

- Power BI: Tình trạng trễ hàng xảy ra nhiều ở hai khu vực Western Europe và Central America



#### 4.4.4. Thống kê các đơn hàng bị hủy theo quốc gia

- SSAS

Dimension	Hierarchy	Operator	Filter Expression
Dim Delivery	Delivery Status	Equal	{ Shipping canceled }
<Select dimension>			
Name Country			Fact Delivery Count
Afganistán			3
Alemania			150
Angola			4
Arabia Saudí			12
Argelia			10
Argentina			32
Australia			165
Austria			21
Bangladés			5
Barbados			1
Bélgica			13
Benín			1
Bielorrusia			3
Bolivia			3
Bosnia y Herz...			1
Brasil			130
Bulgaria			5
Camboya			3
Camerún			7
Canada			18
Chile			9
China			121

- Power BI: Mỹ và Úc số lượng hủy hàng cao nhất

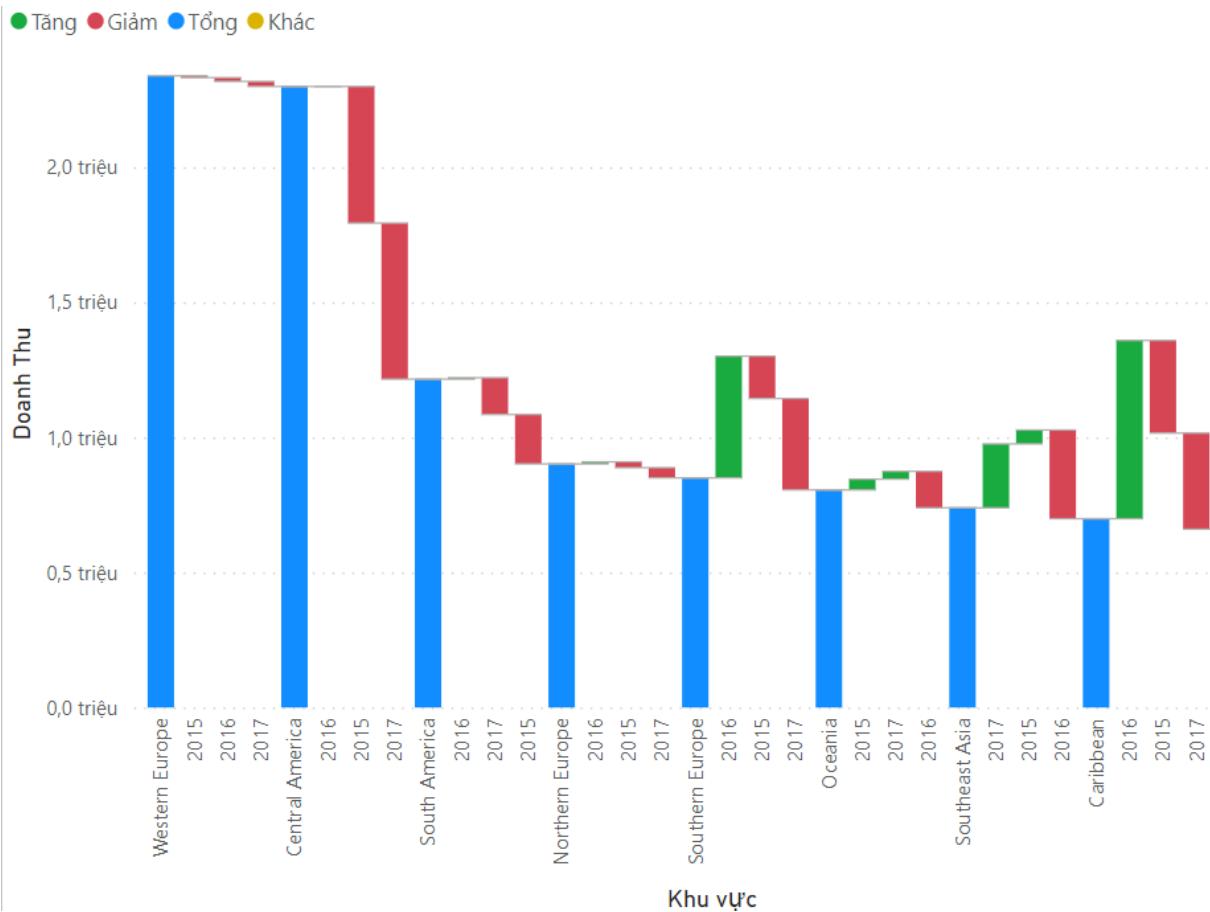


#### 4.4.5. So sánh doanh thu giữa các khu vực theo năm

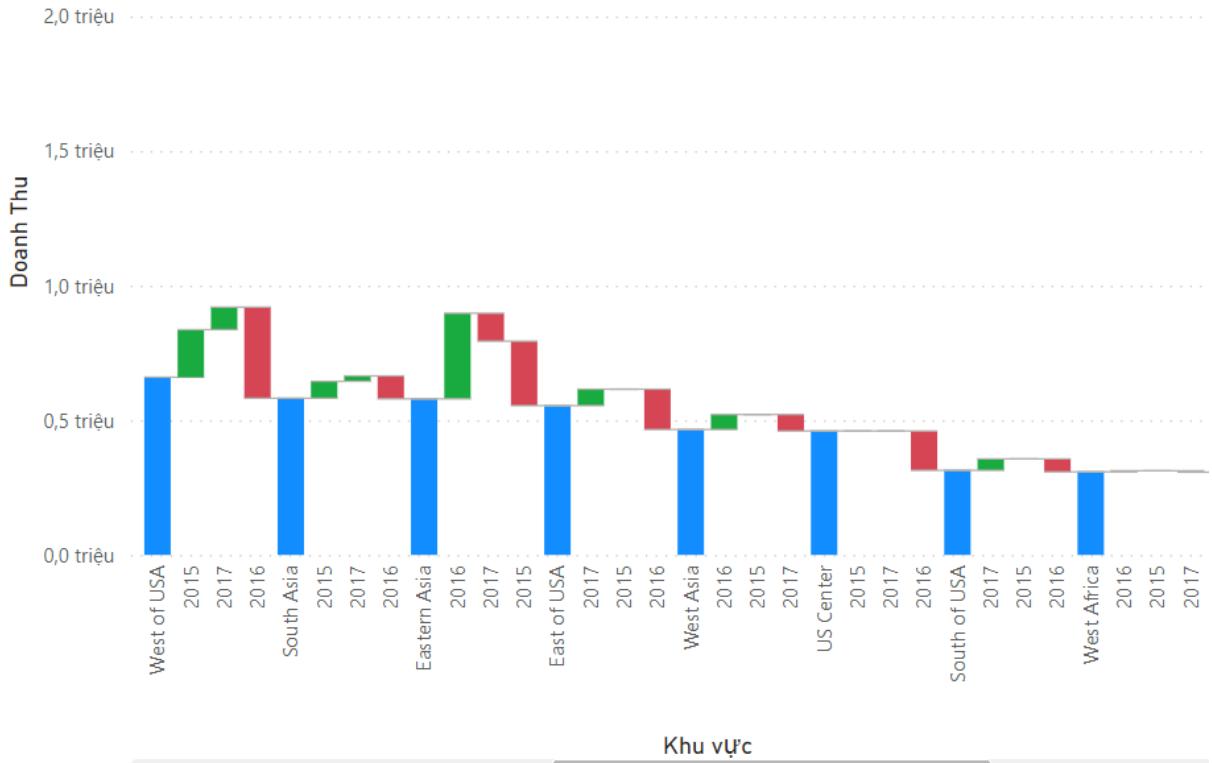
- SSAS

<All>			
	Order Year	Name Region	Item Total
	2015	Caribbean	745760,138289641
	2015	Central Amer...	2474553,79714731
	2015	Eastern Asia	519296,736050802
	2015	Northern Eu...	944920,010250129
	2015	Oceania	526772,055885413
	2015	South America	1337843,65488311
	2015	South Asia	410583,794463544
	2015	Southeast Asia	660071,227311609
	2015	Southern Eu...	882611,239500755
	2015	Western Eur...	2587130,89918825
	2016	Canada	147743,021538451
	2016	Central Africa	266170,482843222
	2016	Central Asia	87173,9008543302
	2016	East Africa	300984,573508454
	2016	East of USA	1231955,13389652
	2016	Eastern Asia	530380,665983082
	2016	Eastern Eur...	611010,476822799
	2016	North Africa	507204,195447343
	2016	Northern Eu...	16895,350239
	2016	Oceania	1016808,22115733
	2016	South Asia	728209,048134587
	2016	South of USA	706904,337914258

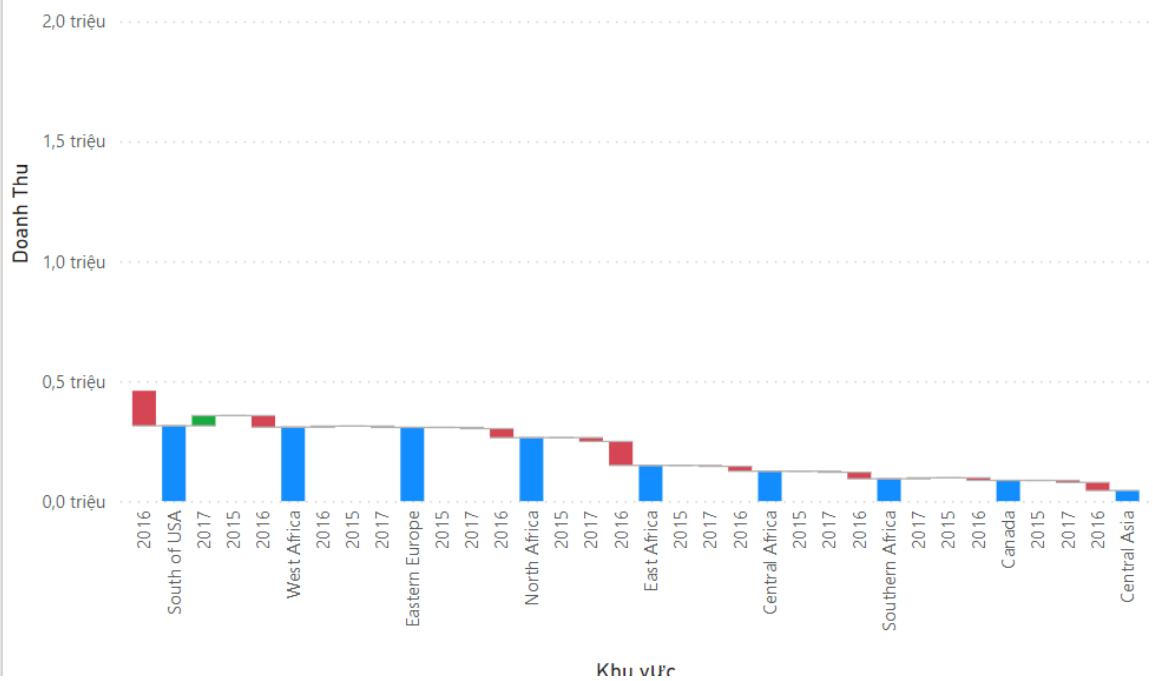
- Power BI



● Tăng ● Giảm ● Tổng ● Khác



● Tăng ● Giảm ● Tổng ● Khác

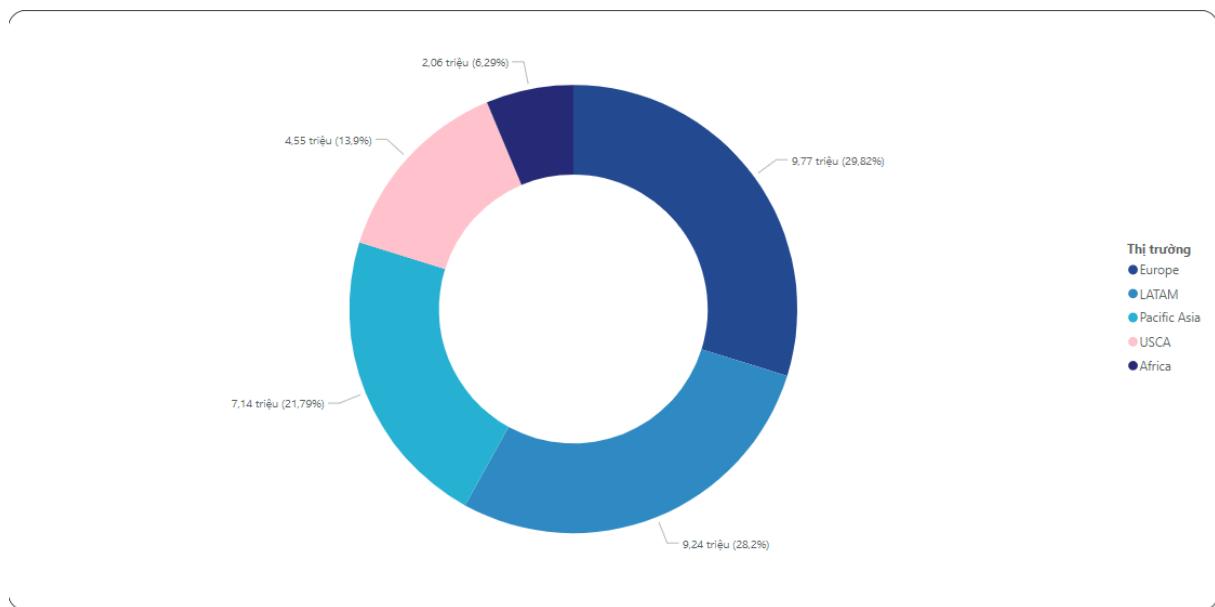


#### 4.4.6. Thị trường nào có doanh thu nhiều nhất và thị trường nào có doanh thu ít nhất?

- SSAS

Name Market	Item Total
Africa	2061679,39285038
Europe	9769198,32166946
LATAM	9235762,09184097
Pacific Asia	7434262,65283859
USCA	4553499,92101627

- Power BI: thị trường Europe là thị trường lớn nhất và Africa là thị trường nhỏ nhất.



#### 4.4.7. Sản phẩm nào bán chạy nhất trong một thời gian nhất định?

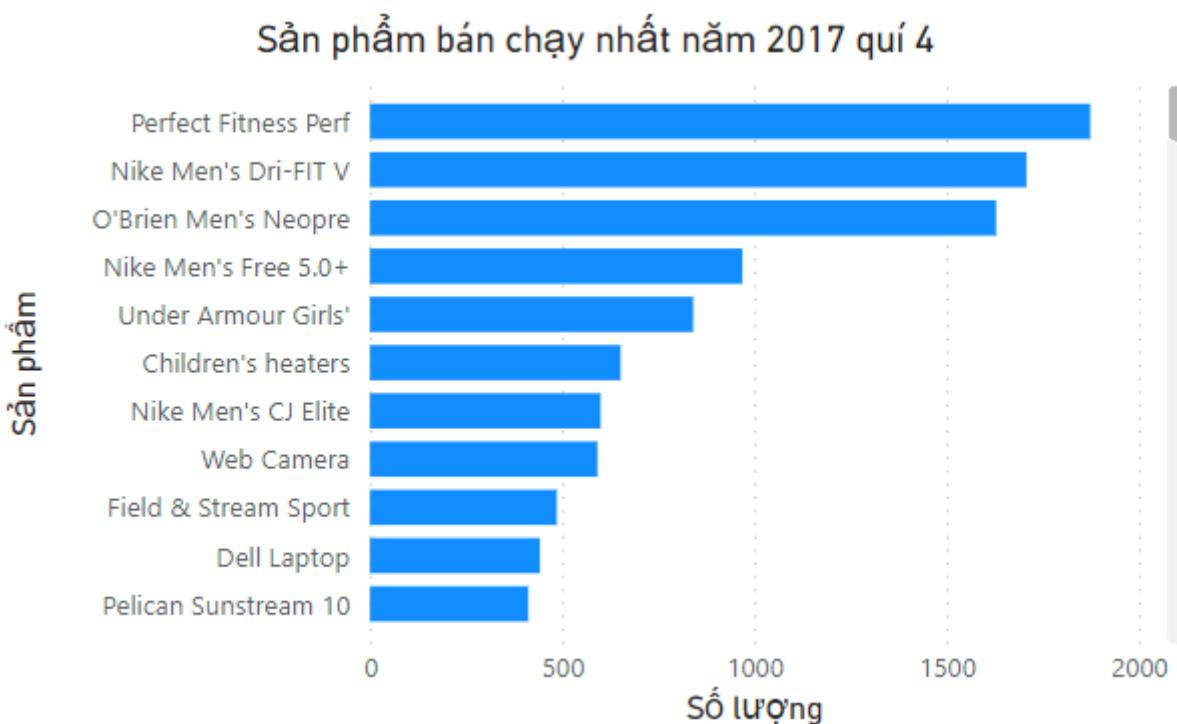
Mẫu ví dụ ta lấy thử: số sản phẩm bán ra năm 2017 quý 4

- SSAS

Dimension Hierarchy Operator Filter Expression

Dim Time	Order Year	Equal	{ 2017 }	
Dim Time	Order Quarter	Equal	{ 4 }	
<Select dimension>				
Name	Product	Order Year	Order Quarter	Quantity
	adidas Brazuca 2014	2017	4	8
	adidas Kids' F5 Mess	2017	4	9
	adidas Men's F10 Mes	2017	4	12
	adidas Men's Germany	2017	4	25
	adidas Youth Germany	2017	4	12
	Adult dog supplies	2017	4	127
	Baby sweater	2017	4	125
	Bag Boy Beverage Hol	2017	4	7
	Bag Boy M330 Push Ca	2017	4	20
	Bridgestone e6 Strai	2017	4	31
	Brooks Women's Ghost	2017	4	23
	CDs of rock	2017	4	149
	Children's heaters	2017	4	652
	Cleveland Golf Colle	2017	4	10
	Cleveland Golf Women	2017	4	6
	Clicgear 8.0 Shoe Br	2017	4	6
	Clicgear Rovic Coole	2017	4	9
	Columbia Men's PFG A	2017	4	5
	Dell Laptop	2017	4	442
	Diamondback Boys' In	2017	4	6
	Diamondback Girls' C	2017	4	2
	Diamondback Women's	2017	4	382
	DVDs	2017	4	326

- Power BI: theo như biểu đồ thì ở quý 4 năm 2017, 2 sản phẩm là Perfect Fitness Perf và Nike Men's Dri-FIT V là bán chạy nhất



#### 4.4.8. Loại mặt hàng nào được ưa chuộng nhất theo từng thị trường

- SSAS

**DW Retail Cube**

**Metadata**

**Measure Group:**

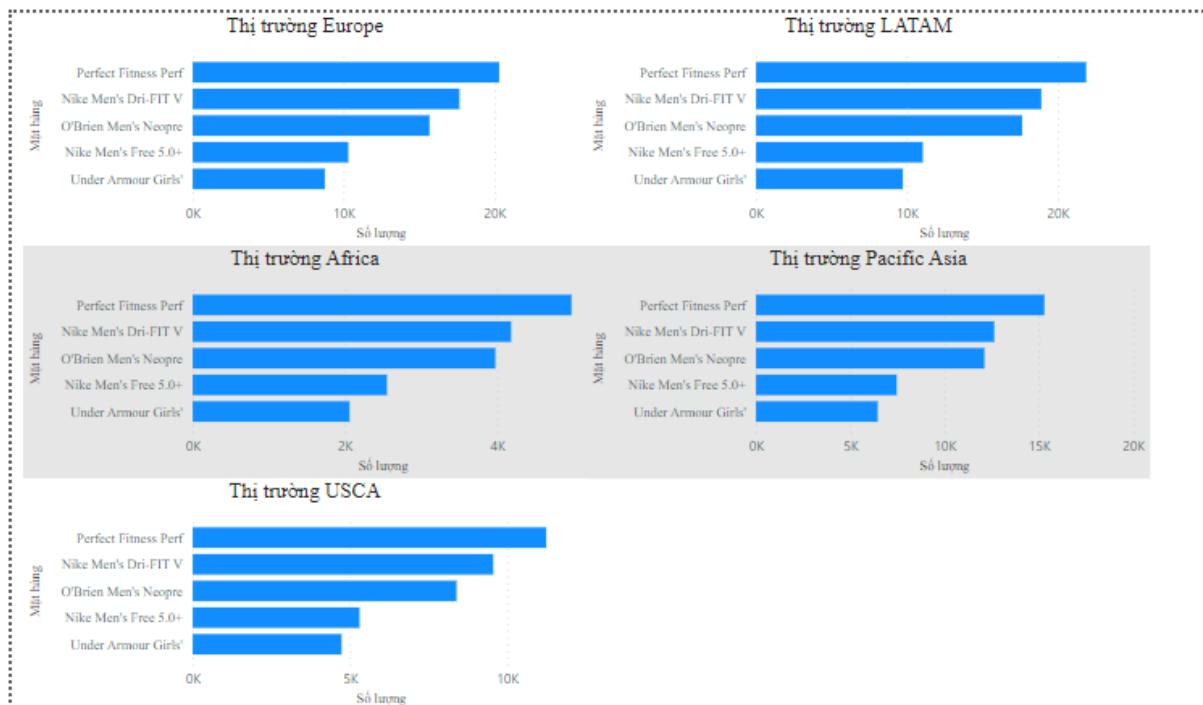
<All>

- DW Retail Cube
  - Measures
  - KPIs
  - Dim City
  - Dim Customer
  - Dim Customer - ID City
  - Dim Delivery
  - Dim Department
  - Dim Product
  - Dim Region
    - ID Market
    - ID Region
    - Name Market
    - Name Region
  - Dim Shipping Mode
  - Dim Time
  - Fact Delivery
  - Fact Sales

**Calculated Members**

Dimension	Hierarchy			
<Select dimension>				
Name	Market	Name	Category	Quantity
Africa	Accessories	462		
Africa	Baseball & Softball	166		
Africa	Boxing & MMA	95		
Africa	Camping & Hiking	884		
Africa	Cardio Equipment	2621		
Africa	Cleats	4975		
Africa	Electronics	758		
Africa	Fishing	1185		
Africa	Fitness Accessories	76		
Africa	Girls' Apparel	236		
Africa	Golf Apparel	93		
Africa	Golf Balls	366		
Africa	Golf Gloves	179		
Africa	Golf Shoes	78		
Africa	Hockey	124		
Africa	Hunting & Shooting	79		
Africa	Indoor/Outdoor Games	3975		
Africa	Lacrosse	70		
Africa	Men's Footwear	1500		
Africa	Shop By Sport	2135		
Africa	Tennis & Racquet	83		
Africa	Trade-In	220		
Africa	Water Sports	1061		

- Power BI: Ta thấy cả các thị trường thì Perfect Fitness Perf và Nike Men's Dri-FIT V đều bán chạy nhất

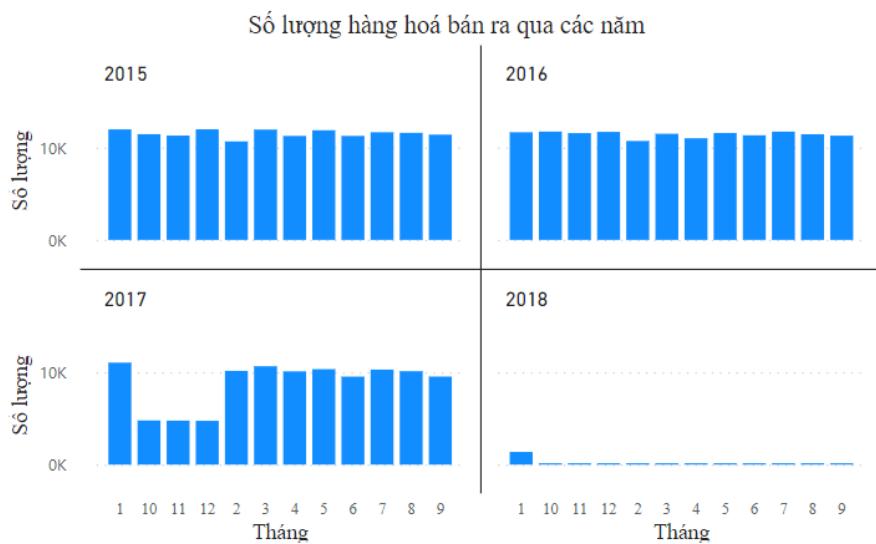


#### 4.4.9. Tháng nào trong năm thường có mức độ bán hàng cao nhất?

- SSAS

Order Year	Order Month	Quantity
2015	1	11979
2015	10	11461
2015	11	11309
2015	12	11987
2015	2	10674
2015	3	11958
2015	4	11271
2015	5	11883
2015	6	11273
2015	7	11669
2015	8	11604
2015	9	11412
2016	1	11668
2016	10	11738
2016	11	11576
2016	12	11706
2016	2	10723
2016	3	11504
2016	4	11020
2016	5	11591
2016	6	11337
2016	7	11734
2016	8	11455

- Power BI: ở năm 2015 và 2016 thì giữa các tháng đều có doanh số ổn định đến năm 2017 thì doanh số ở ba tháng cuối năm có sự sụt giảm đáng kể

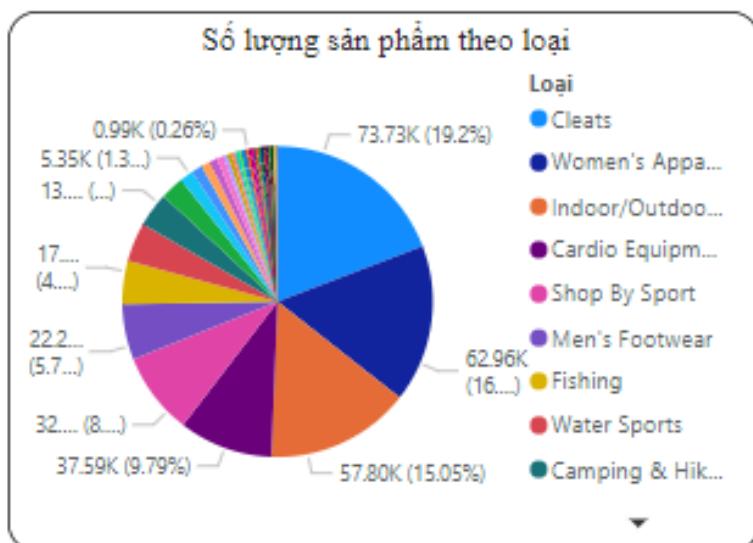


#### 4.4.10. Danh mục sản phẩm nào được ưa chuộng nhất?

- SSAS

Name Category	Delivery Status	Quantity
Accessories	Advance shipp...	5349
Accessories	Late delivery	5349
Accessories	Shipping on time	5349
As Seen on TV!	Advance shipp...	206
As Seen on TV!	Late delivery	206
As Seen on TV!	Shipping on time	206
Baby	Advance shipp...	207
Baby	Late delivery	207
Baby	Shipping on time	207
Baseball & Soft...	Advance shipp...	1785
Baseball & Soft...	Late delivery	1785
Baseball & Soft...	Shipping on time	1785
Basketball	Advance shipp...	67
Basketball	Late delivery	67

- Power BI: Ta thấy danh mục sản phẩm Cleats có số lượng sản phẩm cao nhất chiếm 19.2% trong tổng số lượng sản phẩm bán ra(Không tính sản phẩm hoàn trả).



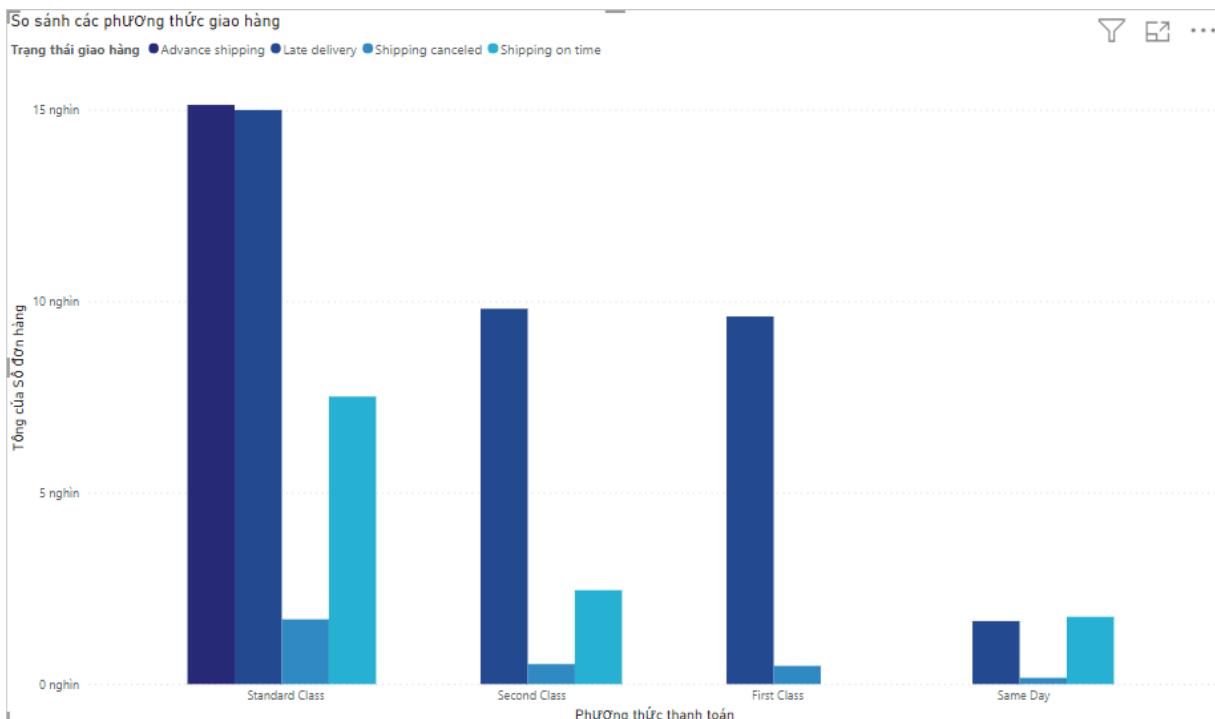
#### 4.4.11. So sánh các phương thức vận chuyển

- SSAS

The screenshot shows the SSAS Dimension browser interface. On the left, there is a tree view of dimensions and measures. The 'Fact Delivery' dimension is expanded, showing sub-items like 'Day For Shipment', 'Days Shipping Real', 'Fact Delivery Count', and 'Late Delivery Risk'. Other dimensions listed include 'KPIs', 'Dim City', 'Dim Customer', 'Dim Customer - ID City', 'Dim Delivery', 'Dim Department', 'Dim Product', 'Dim Region', 'Dim Shipping Mode', 'Dim Time', and 'Fact Delivery'. On the right, a table displays the 'Fact Delivery Count' for each combination of 'Shipping Mode' and 'Delivery Status'. The table has columns for 'Shipping Mode', 'Delivery Status', and 'Fact Delivery Count'. The data includes:

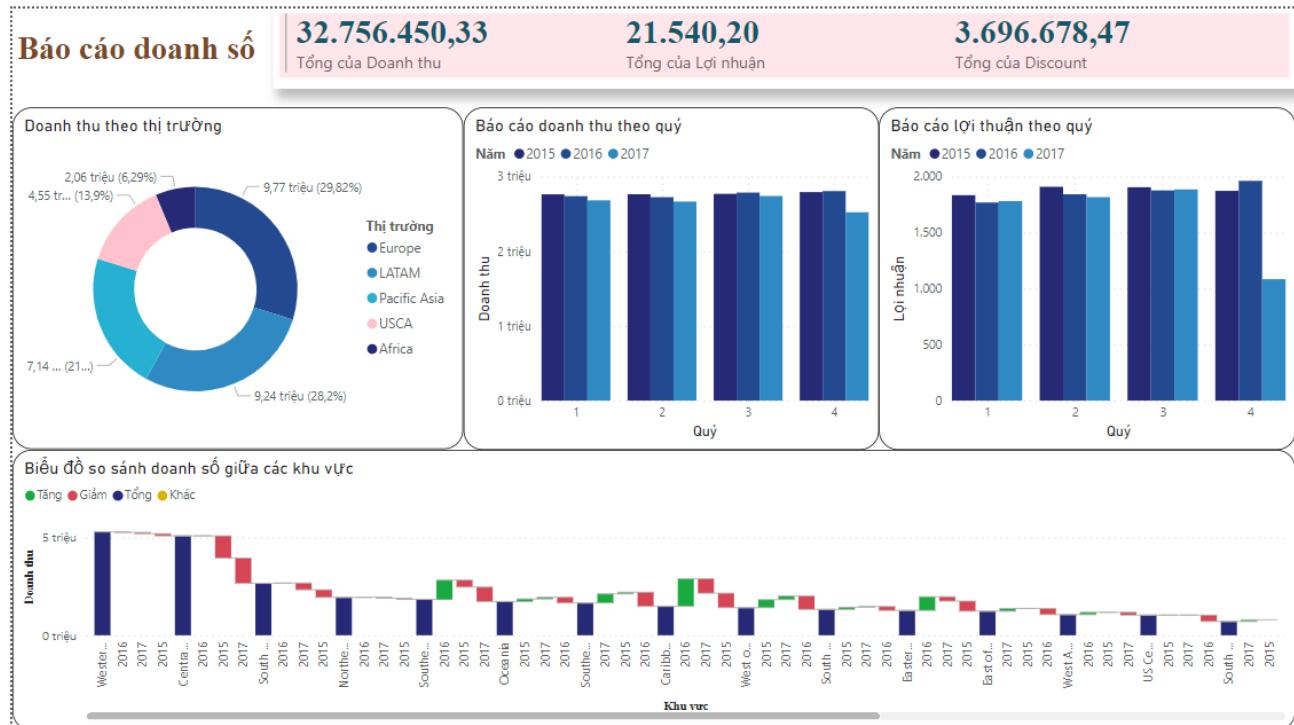
Shipping Mode	Delivery Status	Fact Delivery Count
First Class	Late delivery	9602
First Class	Shipping canceled	477
Same Day	Late delivery	1648
Same Day	Shipping canceled	164
Same Day	Shipping on time	1759
Second Class	Late delivery	9803
Second Class	Shipping canceled	522
Second Class	Shipping on time	2453
Standard Class	Advance shipping	15127
Standard Class	Late delivery	14995
Standard Class	Shipping canceled	1692
Standard Class	Shipping on time	7510

- Power BI:

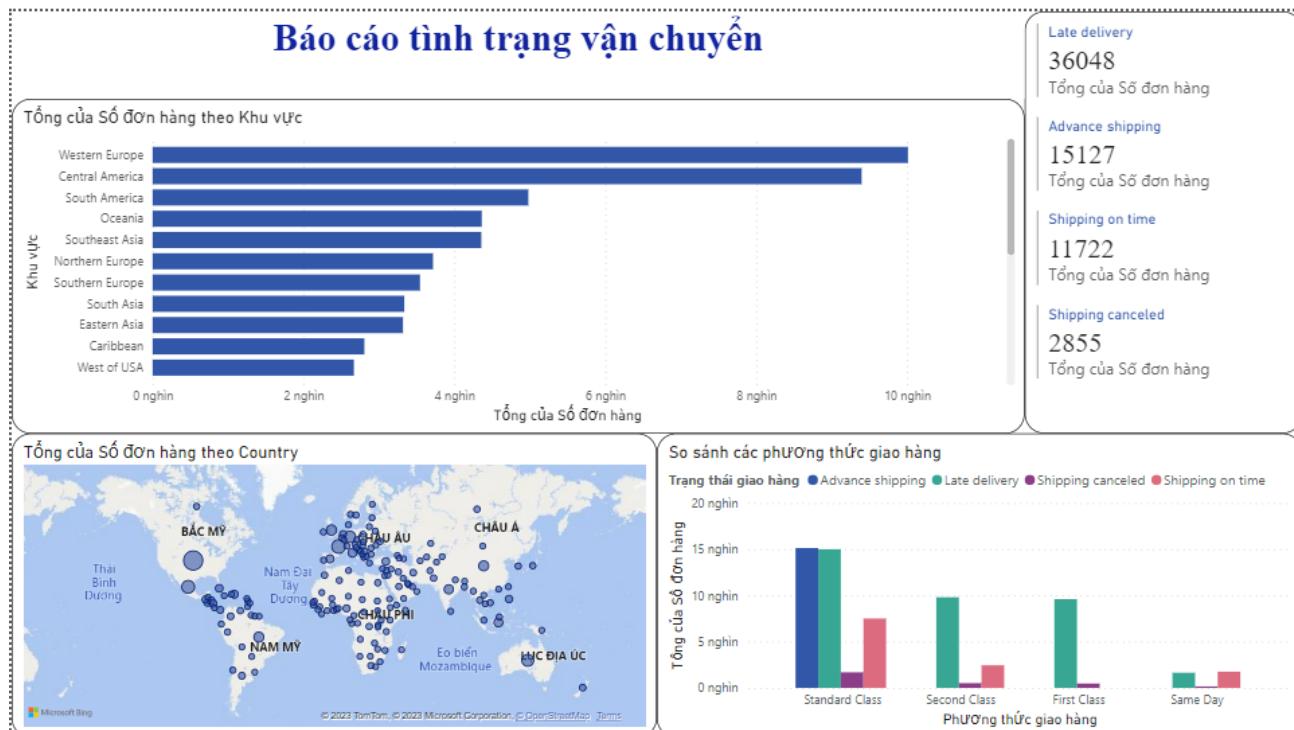


## 4.5. Lập bảng báo cáo DashBoard

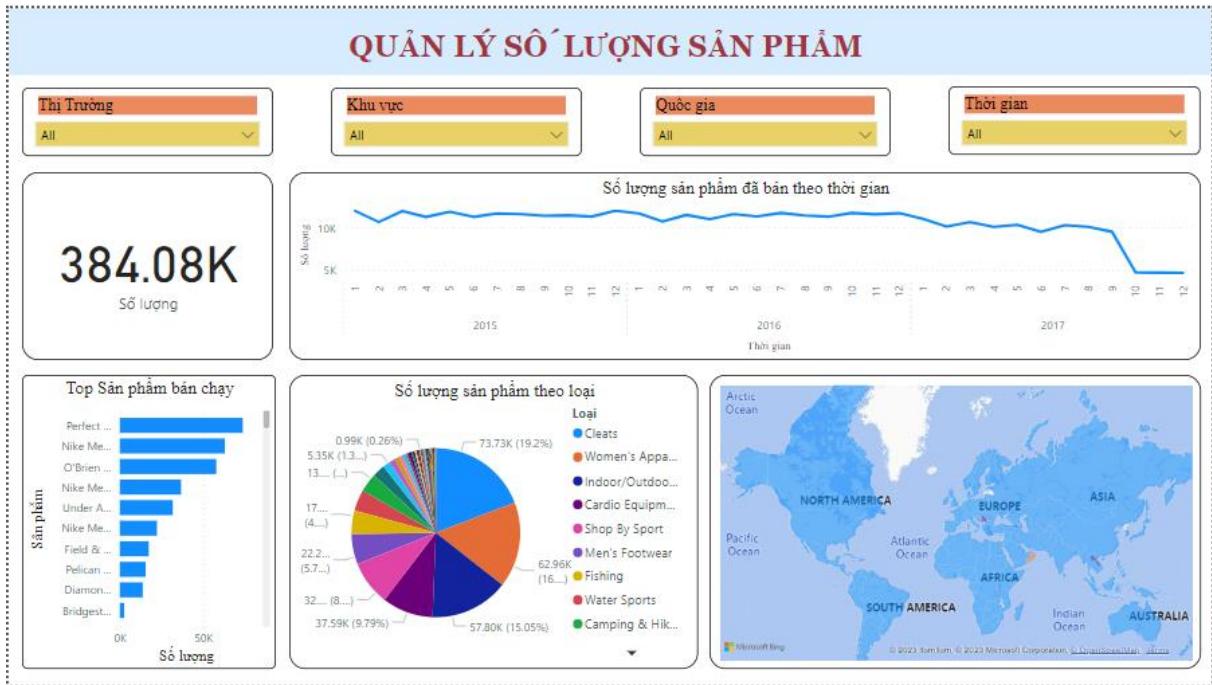
### - Báo cáo doanh thu



### - Báo cáo tình trạng vận chuyển hàng



- Báo cáo số lượng hàng hóa bán ra theo thời gian, khu vực và thị trường và theo loại sản phẩm.



Dashboard cho ta biết tổng số lượng sản phẩm: có thể lựa chọn để xem số lượng theo thị trường, khu vực, quốc gia và cả theo thời gian nhất định. Ngoài ra có biểu đồ map để theo dõi số lượng sản phẩm đã bán trong khu vực và so sánh giữa các khu vực hoặc giữa các thị trường hoặc các quốc gia.

## **Chương 5: Kết luận**

### **5.1. Kết quả đạt được**

Xây dựng kho dữ liệu quản lý bán hàng là đề tài mà nhóm em thực hiện. Mục tiêu của chúng tôi là tạo ra một kho dữ liệu hoàn chỉnh để quản lý dữ liệu bán hàng của một cửa hàng bán lẻ. Để làm được điều này, chúng em đã tiến hành thiết kế cơ sở dữ liệu, xử lý và biến đổi dữ liệu để phù hợp với các yêu cầu phân tích dữ liệu. Sau khi hoàn thành, kho dữ liệu sẽ giúp cho người quản lý bán hàng có thể nhanh chóng truy xuất thông tin và phân tích dữ liệu để đưa ra các quyết định kinh doanh hợp lý.

### **5.2. Những hạn chế chưa đạt được**

Tuy nhiên, nhóm đã nhận ra những điểm yếu và chưa hoàn thiện trong đồ án của mình và đang cố gắng hoàn thiện các vấn đề trên để đạt được một đồ án hoàn chỉnh và chất lượng hơn.

Để tối ưu hóa quá trình xây dựng kho dữ liệu, nhóm đang tiếp tục nghiên cứu và tìm hiểu thêm về các phương pháp và công cụ mới nhất để cải thiện quá trình xây dựng kho dữ liệu. Đồng thời, nhóm cũng đang tập trung vào việc tối ưu hóa các truy vấn trong kho dữ liệu để tăng tốc độ truy xuất dữ liệu và giảm thiểu tải cho hệ thống.

Để nâng cao khả năng phân tích bằng SSAS, nhóm đang tìm hiểu và áp dụng các kỹ thuật và công cụ mới nhất trong việc phân tích dữ liệu. Ngoài ra, nhóm cũng đang tập trung vào việc phân tích các mô hình dữ liệu để đưa ra các phân tích vĩ mô hơn và giúp người dùng có cái nhìn toàn cảnh về dữ liệu.

Tất cả những nỗ lực trên đều nhằm mục đích hoàn thiện đồ án và đưa ra sản phẩm cuối cùng có chất lượng tốt nhất. Nhóm sẽ tiếp tục nghiên cứu và hoàn thiện đồ án để đạt được kết quả tốt nhất và đáp ứng được yêu cầu của khách hàng.

### 5.3. Bảng phân công nhiệm vụ

Nhiệm vụ	Người thực hiện	Tỷ lệ hoàn thành
Vấn đề nhìn thấy	Trương Hoàng Anh Khôi	100%
Giải pháp	Lê Hoàng Hải Đăng	100%
Mục tiêu và ý nghĩa	Phạm Minh Long	100%
Nguồn dữ liệu sử dụng	Lê Hoàng Hải Đăng	100%
Mô tả chi tiết	Trương Hoàng Anh Khôi	100%
Giới thiệu các công cụ	Phạm Minh Long	100%
Thiết kế xây dựng cơ sở dữ liệu	Trương Hoàng Anh Khôi Lê Hoàng Hải Đăng Phạm Minh Long	100%
Xây dựng SSIS cho Department, Shipping Mode, Market, Region	Phạm Minh Long	100%
Xây dựng SSIS cho Time, Segment, Customer, Country, City	Lê Hoàng Hải Đăng	100%
Xây dựng SSIS cho Category, Product, Sales, Delivery	Trương Hoàng Anh Khôi	100%
Xây dựng mô hình và khối – Cube	Trương Hoàng Anh Khôi Phạm Minh Long Lê Hoàng Hải Đăng	100%
Phân tích dữ liệu bằng SSAS và trực quan hóa bằng Power BI - Báo cáo tổng doanh thu theo quý, năm - Báo cáo tổng lợi nhuận theo quý, năm - So sánh doanh thu giữa các khu vực theo năm - Thị trường nào có doanh thu nhiều nhất và ít nhất	Trương Hoàng Anh Khôi	100%

<ul style="list-style-type: none"> <li>- Tạo dashboard cho báo cáo doanh thu</li> </ul>		
<ul style="list-style-type: none"> <li>- Thống kê tình trạng giao hàng trễ của các khu vực</li> <li>- Thống kê các đơn hàng bị hủy theo quốc gia</li> <li>- So sánh các phương thức vận chuyển</li> <li>- Tạo dashboard cho báo cáo tình trạng vận chuyển hàng</li> </ul>	Lê Hoàng Hải Đăng	100%
<ul style="list-style-type: none"> <li>- Sản phẩm nào bán chạy nhất trong một thời gian nhất định</li> <li>- Loại mặt hàng nào được ưa chuộng nhất theo từng thị trường</li> <li>- Tháng nào trong năm thường có mức độ bán hàng cao nhất</li> <li>- Danh mục sản phẩm nào được ưa chuộng nhất</li> <li>- Tạo dashboard cho báo cáo quản lý số lượng sản phẩm</li> </ul>	Phạm Minh Long	100%

## Tài liệu tham khảo

- [1] Tổng cục thống kê Việt Nam, “BÁO CÁO TÌNH HÌNH KINH TẾ – XÃ HỘI THÁNG 11 VÀ 11 THÁNG NĂM 2022”, 2022. Truy cập vào ngày 23 tháng 4 năm 2023 từ: <https://www.gso.gov.vn/bai-top/2022/11/bao-cao-tinh-hinh-kinh-te-xa-hoi-thang-11-va-11-thang-nam-2022/>
- [2] Genk.vn. (2019, April 3). Microsoft chính thức tung ra Visual Studio 2019 với nhiều tính năng mới. Truy cập vào ngày 6 tháng 5 năm 2023, từ <https://genk.vn/microsoft-chinh-thuc-tung-ra-visual-studio-2019-voi-nhieu-tinh-nang-moi-20190403003258636.chn>
- [3] Mendeley Data. (n.d.). Retail dataset for sales forecasting. Truy cập vào ngày 10 tháng 5 năm 2023, từ:  
<https://data.mendeley.com/datasets/8gx2fg2k6/5/files/72784be5-36d3-44fe-b75d-0edbf1999f65>.
- [4] Microsoft. (n.d.). SQL Server Data Tools (SSDT). Truy cập vào ngày 2 tháng 10, 2023, từ <https://learn.microsoft.com/en-us/sql/ssdt/sql-server-data-tools?view=sql-server-ver16>
- [5] QL Server 2012 Tutorials - Analysis Services Multidimensional Modeling: Microsoft Corporation. (2012). SQL Server 2012 Tutorials - Analysis Services Multidimensional Modeling. Truy cập từ <https://docs.microsoft.com/en-us/sql/analysis-services/multidimensional-models/multidimensional-modeling-tutorials-analysis-services?view=sql-server-ver15>
- [6] SSRS Tutorials - SQL Server Reporting Service: Microsoft Corporation. (2022). SSRS Tutorials - SQL Server Reporting Service. Truy cập từ <https://docs.microsoft.com/en-us/sql/reporting-services/tutorial-ssrs?view=sql-server-ver15>
- [7] Microsoft. (2017). SQL Server Integration Services Truy cập vào ngày 10 tháng 5 năm 2023, từ: <https://learn.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver16>