

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA HỆ THÔNG THÔNG TIN



BÁO CÁO ĐÖ ÁN MÔN HỌC
KHO DỮ LIỆU VÀ OLAP

**PHÂN TÍCH DỮ LIỆU ĐÁNH GIÁ CỦA KHÁCH HÀNG
HÃNG HÀNG KHÔNG RYANAIR (2012-2024)**

Giảng viên hướng dẫn: ThS. Đỗ Thị Kim Phụng

Trương Vĩnh Thuận - 21522653

Bùi Văn Thái - 21522577

THÀNH PHỐ HỒ CHÍ MINH, 2024

NHẬN XÉT CỦA GIÁO VIÊN

Mục lục

Chương 1. GIỚI THIỆU TỔNG QUAN DỮ LIỆU	1
1.1. Mô tả dữ liệu.....	1
1.2. Thuộc tính của kho dữ liệu	1
Chương 2. XÂY DỰNG KHO DỮ LIỆU.....	4
2.1. Xây dựng lược đồ hình sao:.....	4
2.2. Mô tả chi tiết các bảng dữ liệu	4
2.2.1. Mô tả chi tiết bảng Fact.....	4
2.2.2. Mô tả chi tiết bảng Dim_Route.....	6
2.2.3. Mô tả chi tiết bảng Dim_TravelType.....	6
2.2.4. Mô tả chi tiết bảng Dim_SeatType	6
2.2.5. Mô tả chi tiết bảng Dim_Return.....	7
2.2.6. Mô tả chi tiết bảng Dim_Aircraft.....	7
2.2.7. Mô tả chi tiết bảng Dim_PassengerCountry	7
2.2.8. Mô tả chi tiết bảng Dim_Time	8
Chương 3. QUÁ TRÌNH XÂY DỰNG KHO DỮ LIỆU (SSIS)	9
3.1. Chuẩn bị cơ sở dữ liệu	9
3.2. Tạo mới project SSIS.....	10
3.3. Tạo bảng Dim và bảng Fact.....	12
3.3.1. Bảng Dim_Route.....	16
3.3.2. Bảng Dim_SeatType	23
3.3.3. Bảng Dim_PassengerCountry	26
3.3.4. Bảng Dim_Aircraft.....	29

3.3.5. Bảng Dim_TravelType.....	32
3.3.6. Bảng Dim_Return	35
3.3.7. Bảng Dim_Time.....	38
3.3.8. Bảng Fact.....	42
3.4. Chạy dự án SSIS	112
3.5. Kiểm tra dữ liệu các bảng.....	115
3.6. Lược đồ quan hệ khi hoàn thành	118
Chương 4. PHÂN TÍCH DỮ LIỆU (SSAS)	119
4.1. Tạo mới Project SSAS	119
4.2. Xác định dữ liệu nguồn (Data Sources)	120
4.3. Xác định khung nhìn dữ liệu nguồn (Data Source Views).....	123
4.4. Xây dựng các khối (Cubes) và xác định các độ đo (Measures)	126
4.5. Xác định các chiều (Dimensions).....	130
4.6. Xác định các độ đo (Measures)	133
4.6.1. Đổi tên và thuộc tính các độ đo ban đầu	133
4.6.2. Tạo các New Calculated Number	135
4.7. Phân cấp trong bảng chiều Dim_Time	138
4.8. Chạy dự án SSAS	145
4.9. Thực hiện 15 câu truy vấn – Quá trình phân tích dữ liệu bằng thao tác tay trên các khối CUBE	145
4.9.1. Câu truy vấn 1	145
4.9.2. Câu truy vấn 2	146
4.9.3. Câu truy vấn 3	147
4.9.4. Câu truy vấn 4	148

4.9.5.	Câu truy vấn 5	150
4.9.6.	Câu truy vấn 6	150
4.9.7.	Câu truy vấn 7	151
4.9.8.	Câu truy vấn 8	153
4.9.9.	Câu truy vấn 9	154
4.9.10.	Câu truy vấn 10	156
4.9.11.	Câu truy vấn 11	156
4.9.12.	Câu truy vấn 12	158
4.9.13.	Câu truy vấn 13	159
4.9.14.	Câu truy vấn 14	161
4.9.15.	Câu truy vấn 15	163
4.10.	Thực hiện 15 câu truy vấn – Quá trình phân tích dữ liệu bằng ngôn ngữ truy vấn MDX	163
4.10.1.	Câu truy vấn 1	163
4.10.2.	Câu truy vấn 2	164
4.10.3.	Câu truy vấn 3	165
4.10.4.	Câu truy vấn 4	166
4.10.5.	Câu truy vấn 5	167
4.10.6.	Câu truy vấn 6	168
4.10.7.	Câu truy vấn 7	169
4.10.8.	Câu truy vấn 8	170
4.10.9.	Câu truy vấn 9	171
4.10.10.	Câu truy vấn 10	172
4.10.11.	Câu truy vấn 11	173

4.10.12. Câu truy vấn 12	173
4.10.13. Câu truy vấn 13	174
4.10.14. Câu truy vấn 14	175
4.10.15. Câu truy vấn 15	176
4.11. Thực hiện Report bằng Excel	177
4.11.1. Câu truy vấn 1	177
4.11.2. Câu truy vấn 2	178
4.11.3. Câu truy vấn 3	180
4.11.4. Câu truy vấn 4	182
4.11.5. Câu truy vấn 5	184
4.11.6. Câu truy vấn 6	186
Chương 5. QUÁ TRÌNH LẬP BÁO BIỂU (SSRS)	188
5.1. Report với Visual Studio	188
5.1.1. Thực hiện thống kê số lượng các tuyến bay.....	190
5.1.2. Thực hiện thống kê số lượng chuyến bay theo từng quốc gia	200
5.1.3. Thống kê số lượng theo các quý của từng năm.....	203
5.2. Report với Power BI.....	208
5.2.1. Thống kê các đánh giá trung bình của khách hàng về chất lượng thức ăn, dịch vụ mặt đất, dịch vụ cabin, và độ thoải mái của ghế qua các năm.	212
5.2.2. Thống kê số lượng các chuyến bay theo mục đích chuyến bay từ năm 2016 đến năm 2024	220
5.2.3. Thống kê số lượng hành khách tham gia chuyến bay theo từng quốc gia.	
227	
Chương 6. QUÁ TRÌNH DATA MINING	231

6.1.	Data Mining bằng công cụ SSAS trên Visual Studio.....	231
6.2.	Nhận xét kết quả sau khi thực hiện Mining dữ liệu	240
6.2.1.	Thuật toán Decision Tree	240
6.2.2.	Thuật toán Clustering	246
6.2.3.	Thuật toán Naïve Bayes	247

Chương 1. GIỚI THIỆU TỔNG QUAN DỮ LIỆU

1.1. Mô tả dữ liệu

- Kho dữ liệu Ryan Air Review là một kho dữ liệu thu thập thông tin các đánh giá của khách hàng tham gia các chuyến bay của hãng hàng không Ryan Air.
- Thông qua kho dữ liệu người dùng có thể biết được thông tin các phản hồi của khách hàng đối với các chuyến bay, điểm xếp loại cho nhiều hạng mục của chuyến bay...
- Kho dữ liệu có khoảng hơn 2 nghìn dòng, 20 thuộc tính.
- Link kho dữ liệu: <https://www.kaggle.com/datasets/cristaliss/ryanair-reviews-ratings>

1.2. Thuộc tính của kho dữ liệu

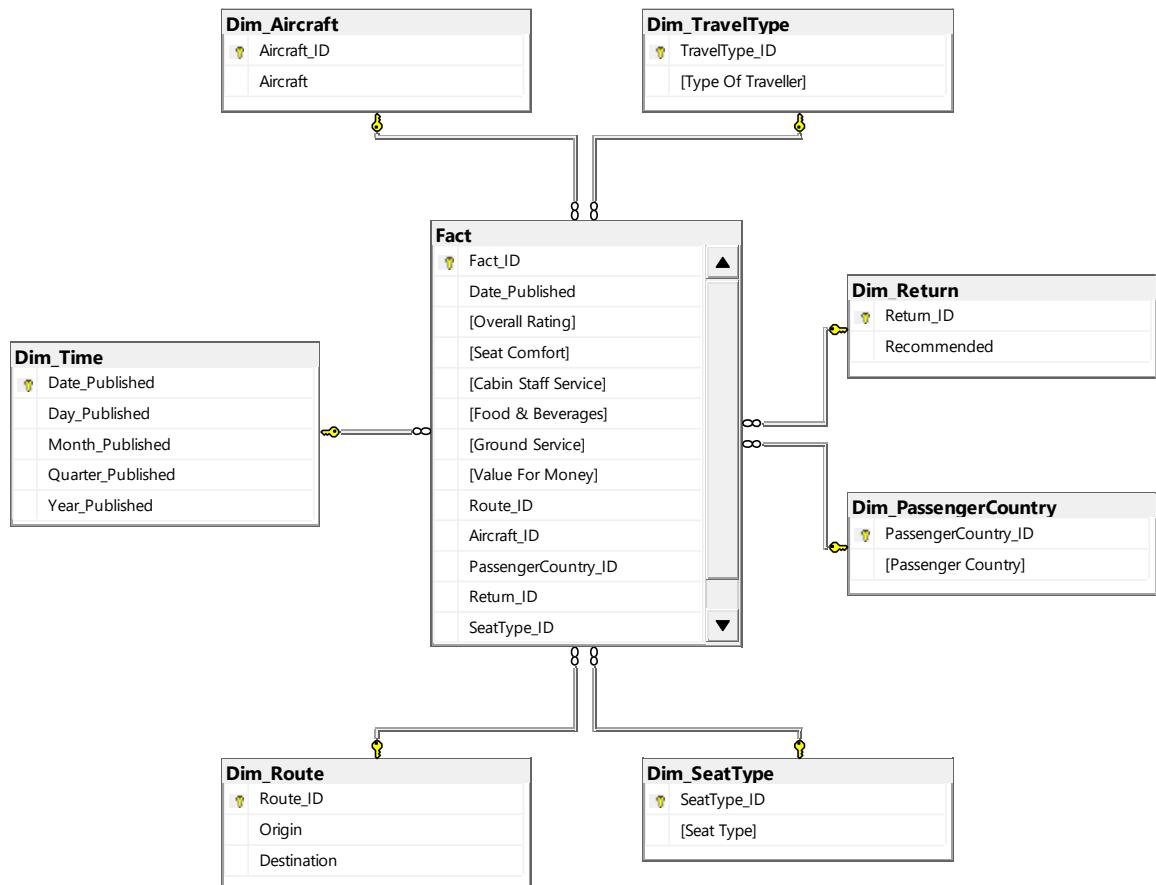
STT	Tên thuộc tính	Kiểu	Ý nghĩa thuộc tính
1	Date Published	Date	Ngày khách hàng đăng phản hồi
2	Overall Rating	Number	Điểm tổng thể khách hàng chấm cho chuyến bay (1-10)
3	Passenger Country	Text	Quốc tịch của khách hàng
4	Trip_verified	Text	Trạng thái xác minh của chuyến bay
5	Comment title	Text	Tiêu đề của bình luận
6	Comment	Text	Bình luận của khách hàng về chuyến bay
7	Aircraft	Text	Tên máy bay

8	Type Of Traveller	Text	Loại hình du lịch
9	Seat Type	Text	Loại ghế ngồi
10	Origin	Text	Điểm xuất phát của chuyến bay
11	Destination	Text	Điểm đến của chuyến bay
12	Date Flown	Date	Ngày và tháng bay của chuyến bay
13	Seat Comfort	Number	Điểm đánh giá mức độ thoải mái của chỗ ngồi (1-5)
14	Cabin Staff Service	Number	Điểm đánh giá sự phục vụ của các nhân viên (1-5)
15	Food & Beverages	Number	Điểm đánh giá dịch vụ ăn uống (1-5)
16	Ground Service	Number	Điểm đánh giá các dịch vụ ở mặt đất (1-5)
17	Value For Money	Number	Điểm đánh giá giá trị mang lại (1-5)
18	Recommended	Text	Đề xuất có nên sử dụng chuyến bay không (yes/no)

19	Inflight Entertainment	Number	Điểm đánh giá các dịch vụ giải trí trong chuyến bay (1-5)
20	Wifi & Connectivity	Number	Điểm đánh giá dịch vụ internet của chuyến bay

Chương 2. XÂY DỰNG KHO DỮ LIỆU

2.1. Xây dựng lược đồ hình sao:



2.2. Mô tả chi tiết các bảng dữ liệu

2.2.1. Mô tả chi tiết bảng Fact

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Ý nghĩa
1	Fact_ID	Integer	PK	Id của bảng Fact
2	Date_published	Date	FK	Mã thời gian đăng bình luận
3	Overall Rating	Double		Điểm đánh giá tổng thể

4	Seat Comfort	Double		Điểm độ thoải mái của chỗ ngồi
5	Cabin Staff Service	Double		Điểm đánh giá dịch vụ nhân viên
6	Food & Beverages	Double		Điểm đánh giá dịch vụ ăn uống
7	Ground Service	Double		Điểm đánh giá dịch vụ mặt đất
8	Value For Money	Double		Điểm đánh giá giá trị mang lại
9	Route_ID	Integer	FK	Mã tuyến
10	Aircraft_ID	Integer	FK	Mã máy bay
11	PassengerCountry_ID	Integer	FK	Mã quốc tịch khách hàng
12	Return_ID	Integer	FK	Mã phản hồi
13	SeatType_ID	Integer	FK	Mã loại chỗ ngồi

2.2.2. Mô tả chi tiết bảng Dim_Route

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Ý nghĩa
1	Route_ID	Integer	PK, FK	Mã tuyến
2	Origin	Text		Điểm xuất phát của chuyến bay
3	Destination	Text		Điểm đến của chuyến bay

2.2.3. Mô tả chi tiết bảng Dim_TravelType

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Ý nghĩa
1	TravelType_ID	Integer	PK, FK	Mã loại hình du lịch
2	Type Of Traveler	Text		Loại hình du lịch

2.2.4. Mô tả chi tiết bảng Dim_SeatType

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Ý nghĩa
1	SeatType_ID	Integer	PK, FK	Mã loại chỗ ngồi
2	Seat Type	Text		Loại chỗ ngồi

2.2.5. Mô tả chi tiết bảng Dim_Return

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Ý nghĩa
1	Return_ID	Integer	PK,FK	Mã phản hồi
2	Recommended	Text		Gợi ý có nên sử dụng chuyến bay này không

2.2.6. Mô tả chi tiết bảng Dim_Aircraft

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Ý nghĩa
1	Aircraft_ID	Integer	PK,FK	Mã máy bay
2	Aircraft	Text		Tên máy bay

2.2.7. Mô tả chi tiết bảng Dim_PassengerCountry

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Ý nghĩa
1	PassengerCountry_ID	Integer	PK,FK	Mã quốc tịch khách hàng
2	Passenger Country	Text		Quốc tịch của hành khách

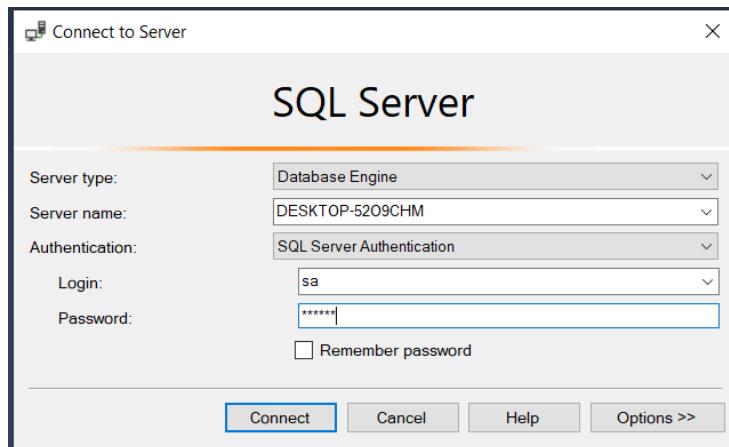
2.2.8. Mô tả chi tiết bảng Dim_Time

STT	Tên thuộc tính	Kiểu dữ liệu	Ràng buộc	Ý nghĩa
1	Date_Published	Integer	PK,FK	Mã thời gian
2	Day_Published	Number		Ngày đăng bình luận
3	Month_Published	Number		Tháng đăng bình luận
4	Quarter_Published	Number		Quý đăng bình luận
5	Year_Published	Number		Năm đăng bình luận

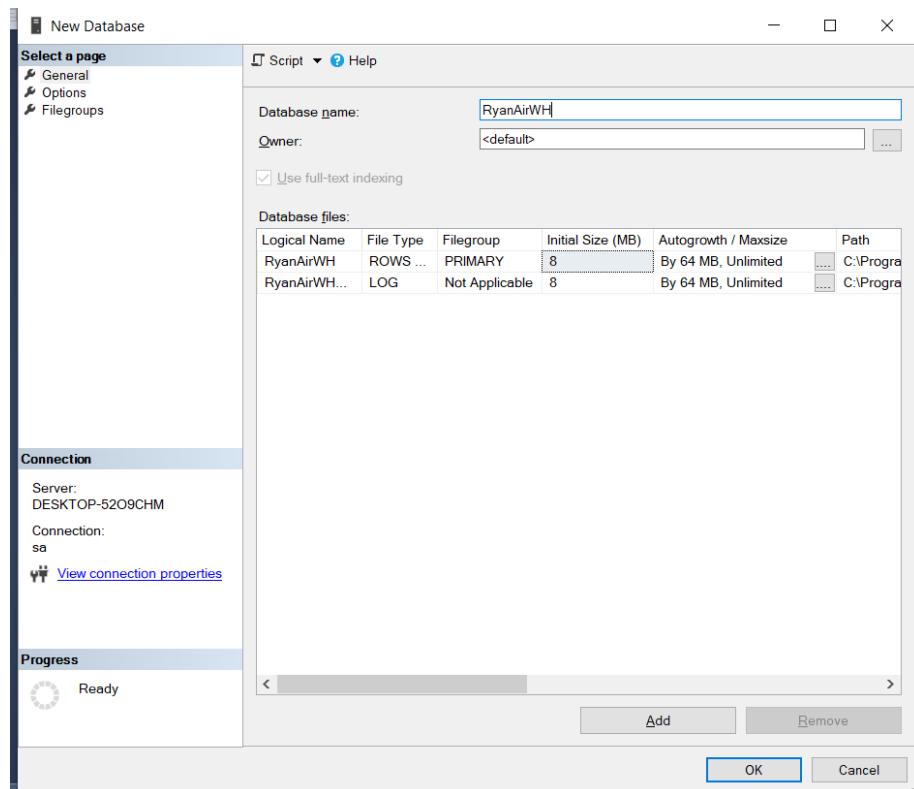
Chương 3. QUÁ TRÌNH XÂY DỰNG KHO DỮ LIỆU (SSIS)

3.1. Chuẩn bị cơ sở dữ liệu

Bước 1: Mở SQL Server 2019 và kết nối với server

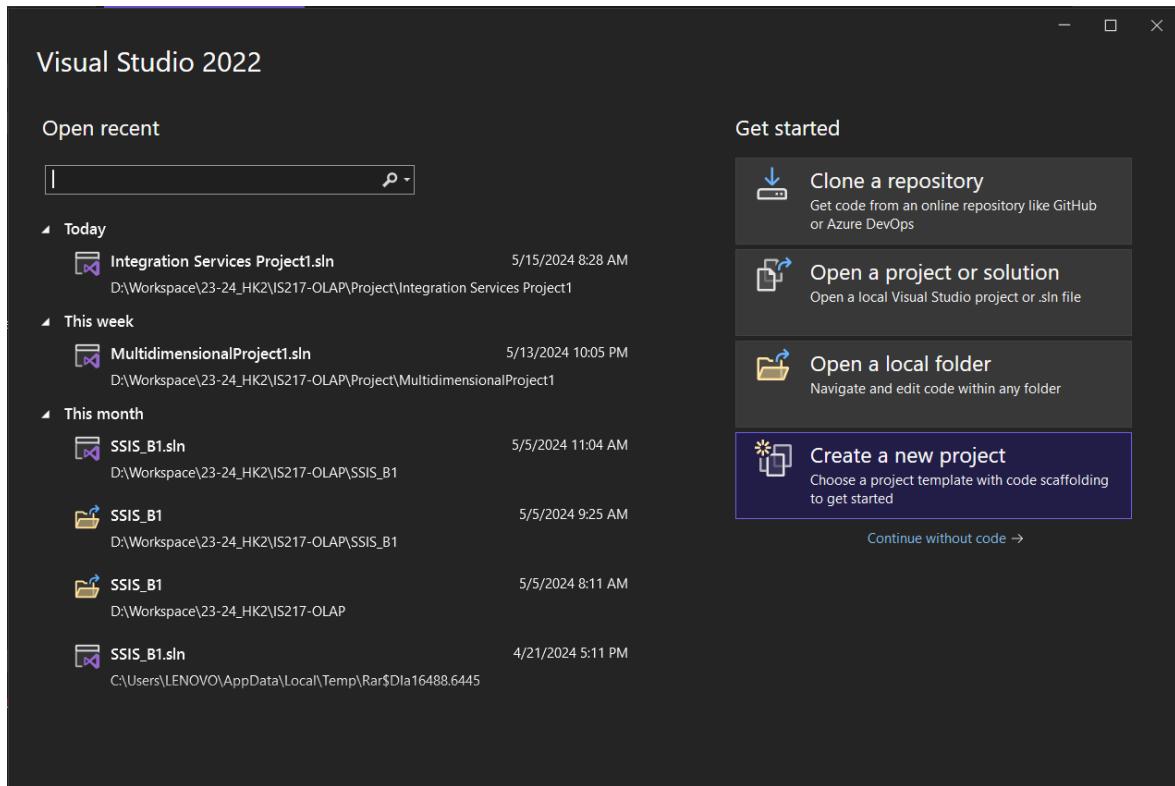


Bước 2: Khởi tạo một cơ sở dữ liệu có tên RyanAirWH, đây là nơi lưu các bảng Dim và bảng Fact cùng dữ liệu của các bảng đó

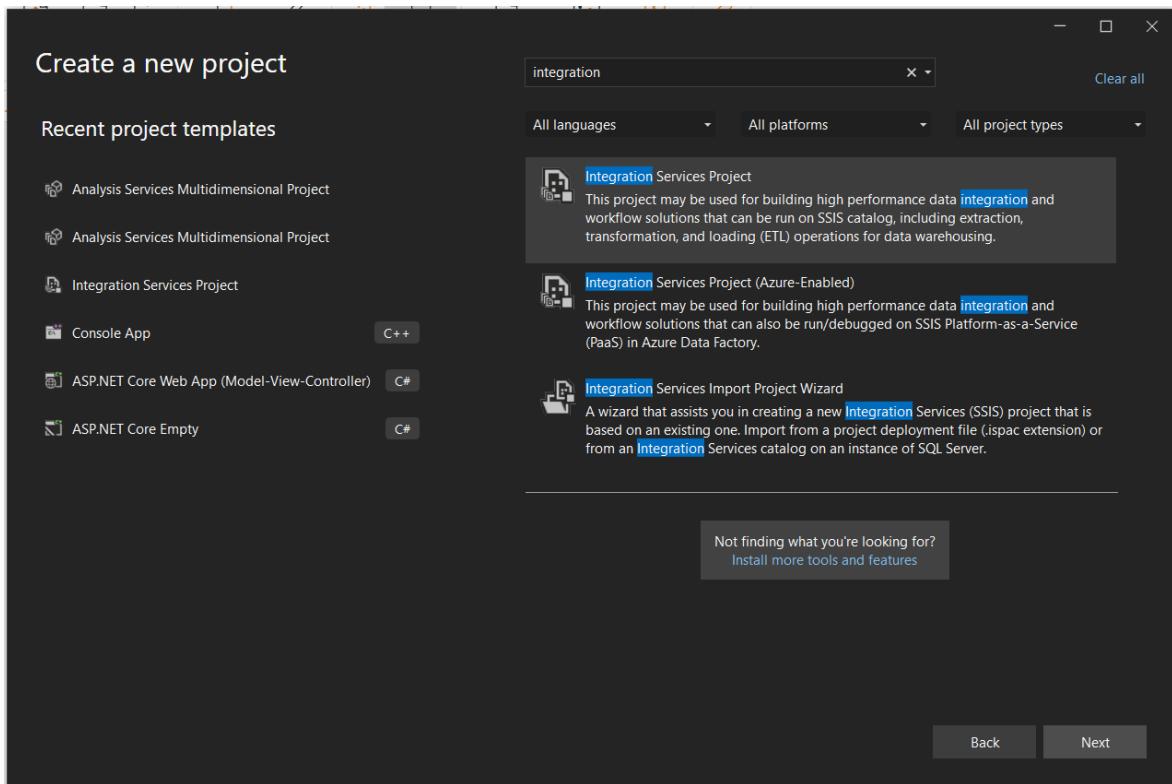


3.2. Tạo mới project SSIS

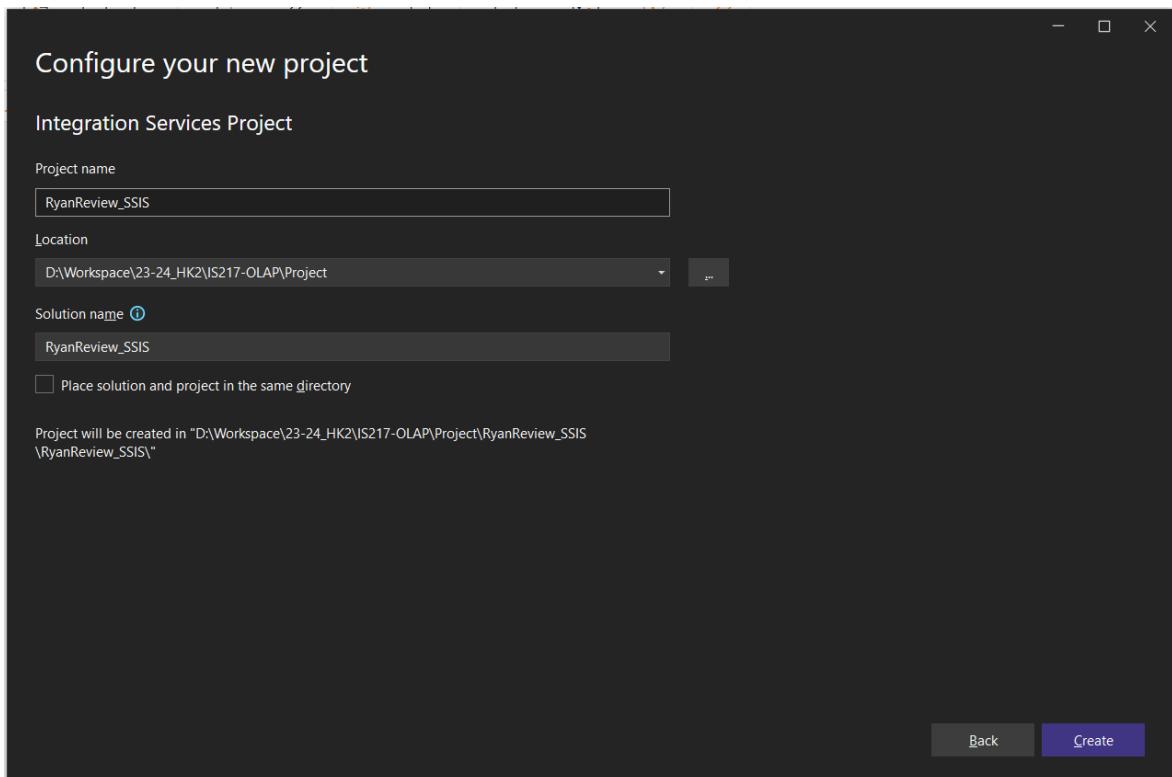
Bước 1: Mở Visual Studio và chọn “Create a new project”.



Bước 2: Chọn Integration Services Project và chọn Next



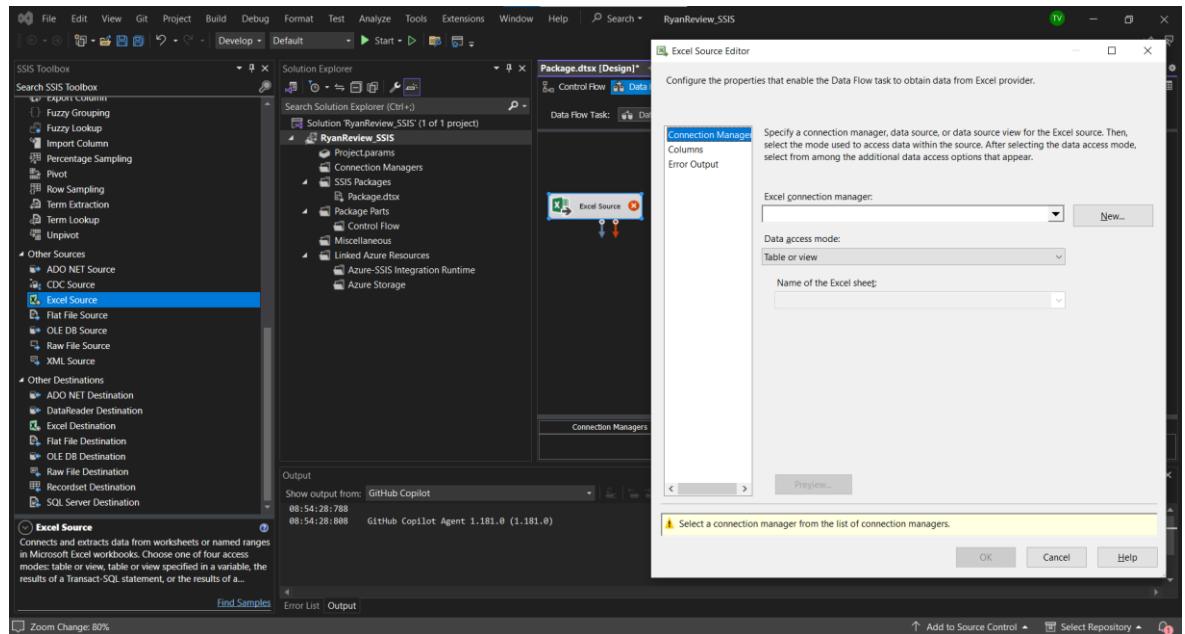
Bước 3: Đặt tên và thiết lập đường dẫn cho Project. Sau đó chọn Create.



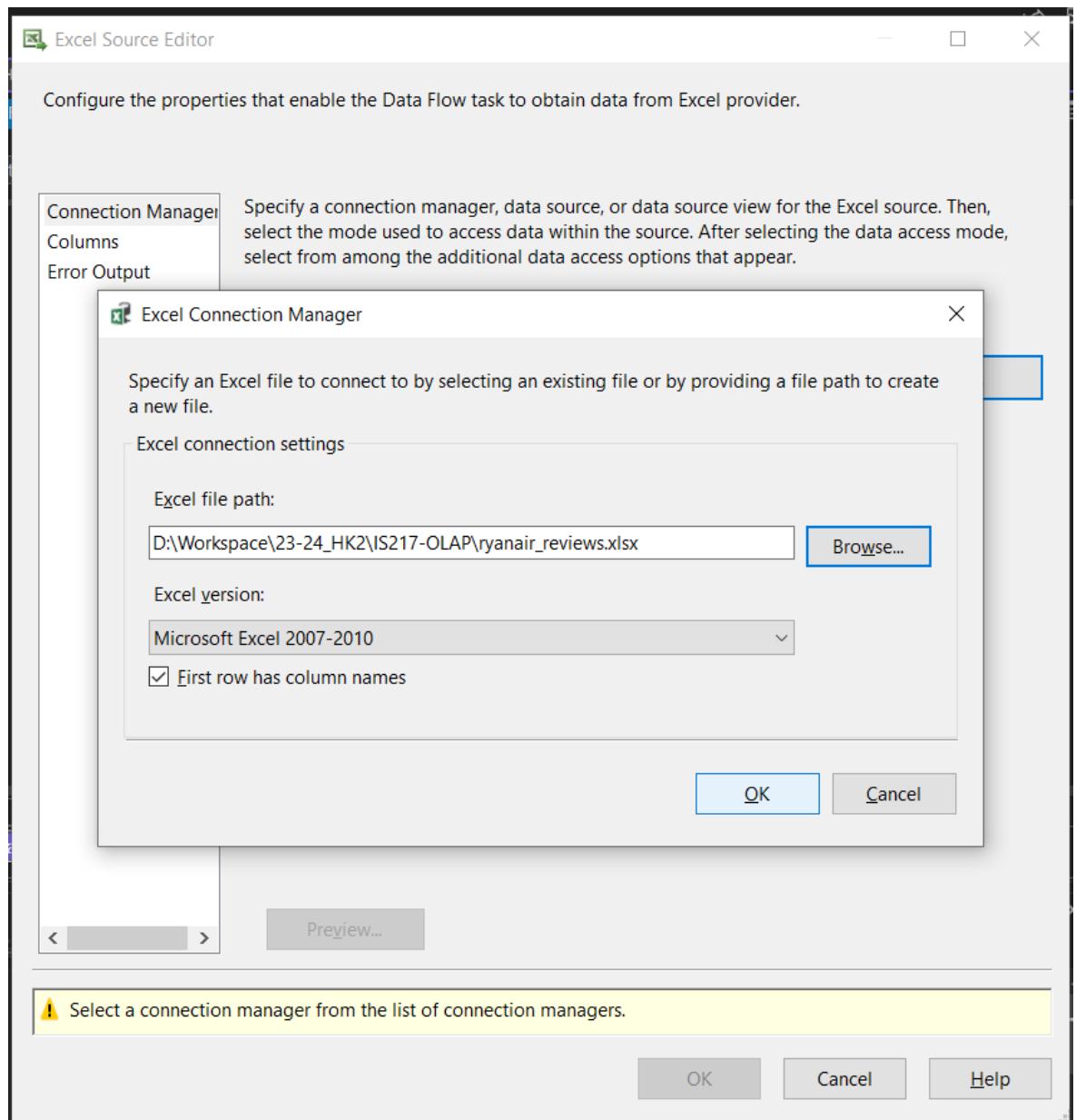
3.3. Tạo bảng Dim và bảng Fact

Trước khi tiến hành chia Dimension và bảng Fact, ta cần load dữ liệu gốc từ file .xlsx vào Data Flow:

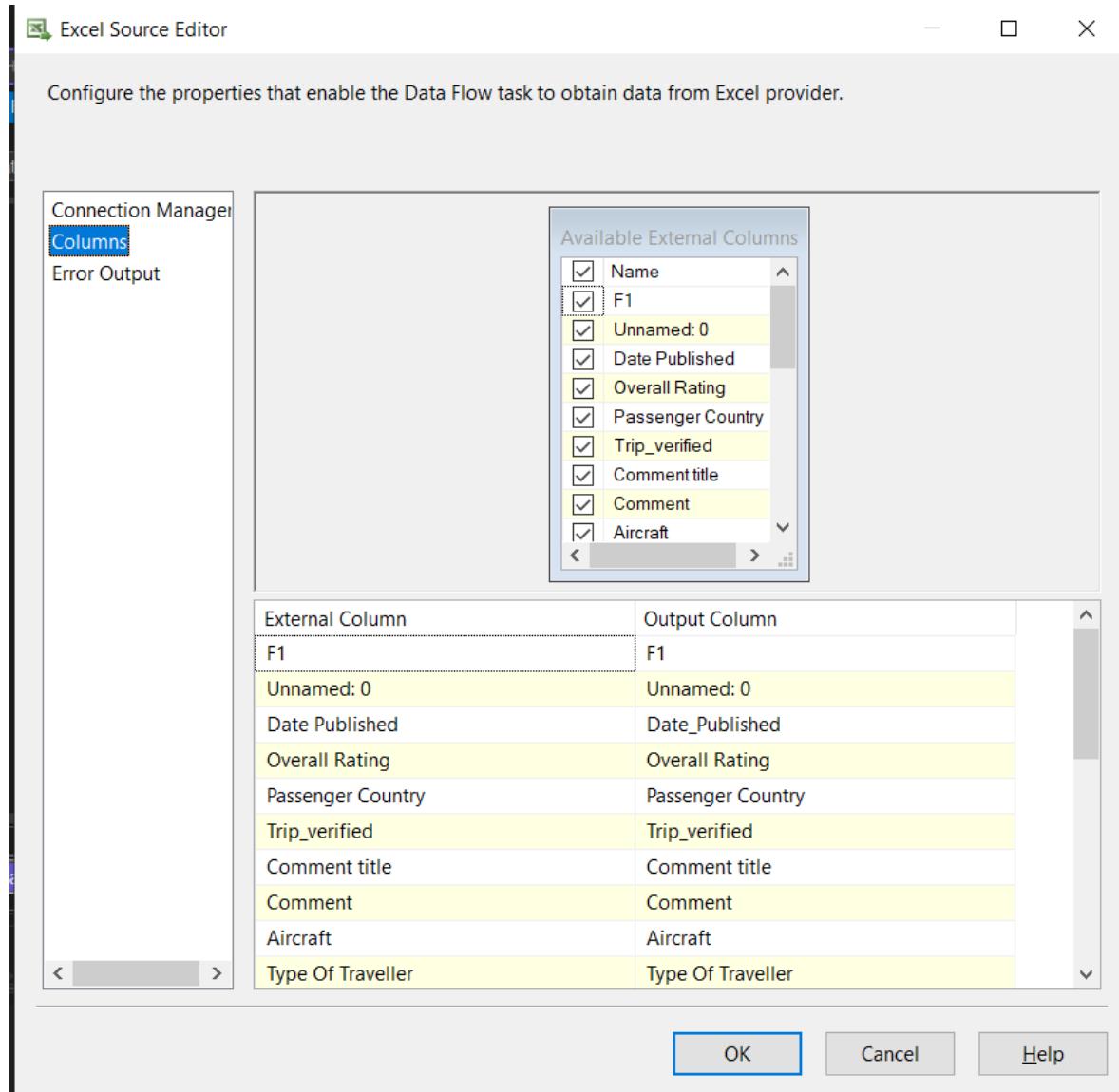
Bước 1. Trong Data Flow, tạo một đối tượng Excel Source để lấy dữ liệu gốc từ file .xlsx. Chọn New để tạo một Excel Connection Manager



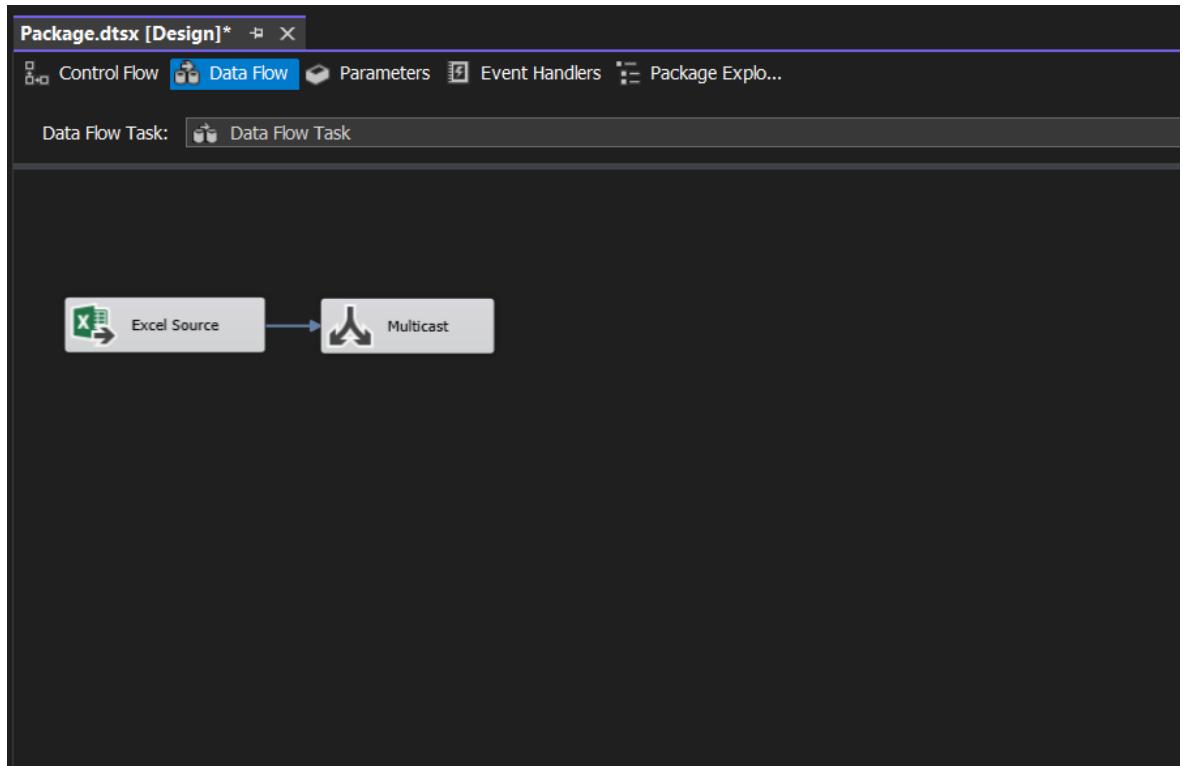
Bước 2: Chọn **Browse** và duyệt đường dẫn đến file cần chọn, sau đó nhấn **OK**



Bước 3. Click chọn **Columns** để kiểm tra các cột dữ liệu ở dạng danh sách. Nhấn **OK** lần nữa để tiến hành hoàn tất quá trình load dữ liệu vào Excel Source



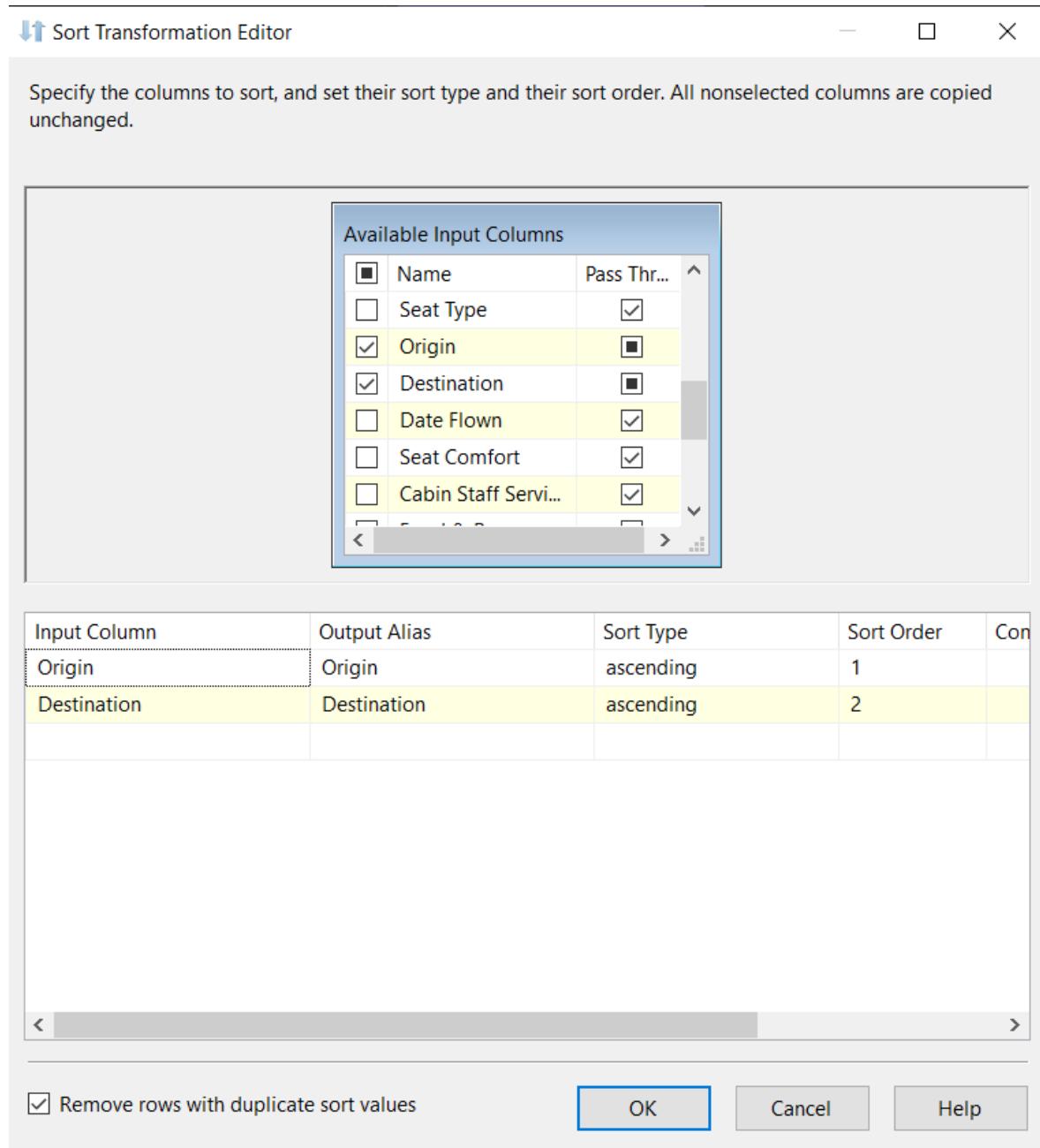
Bước 4. Tạo Multicast để phân tán dữ liệu từ Excel Source đến các Dimension.
Tiến hành kết nối Excel Source và Multicast.



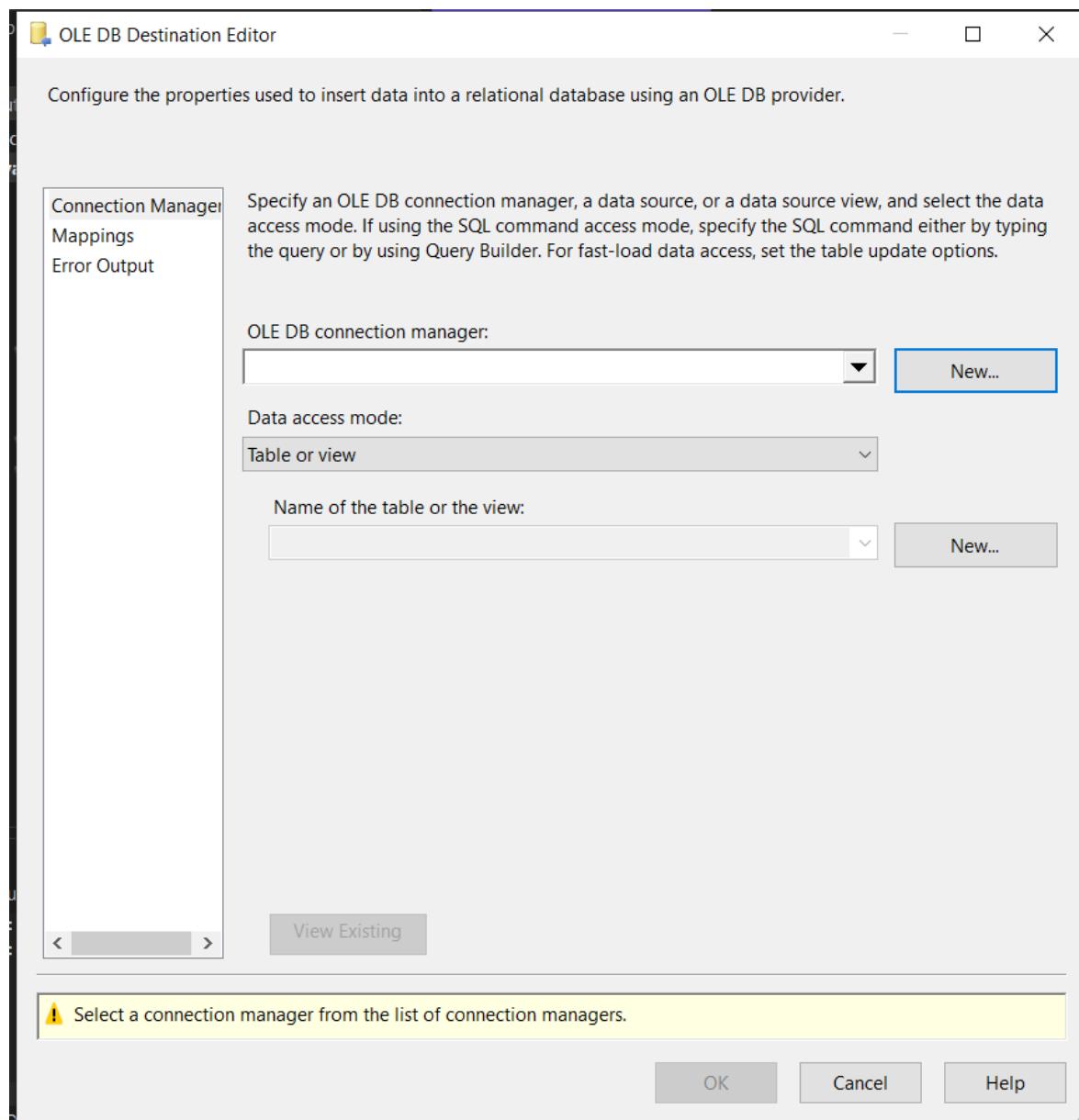
3.3.1. Bảng Dim_Route

Bước 1. Tạo mới một **Sort** có tên là Sort_Dim_Route để lấy ra các cột dữ liệu cần thiết cho Dim_Route. Nhấn chuột phải và nhấn **Edit** để chọn các cột dữ liệu

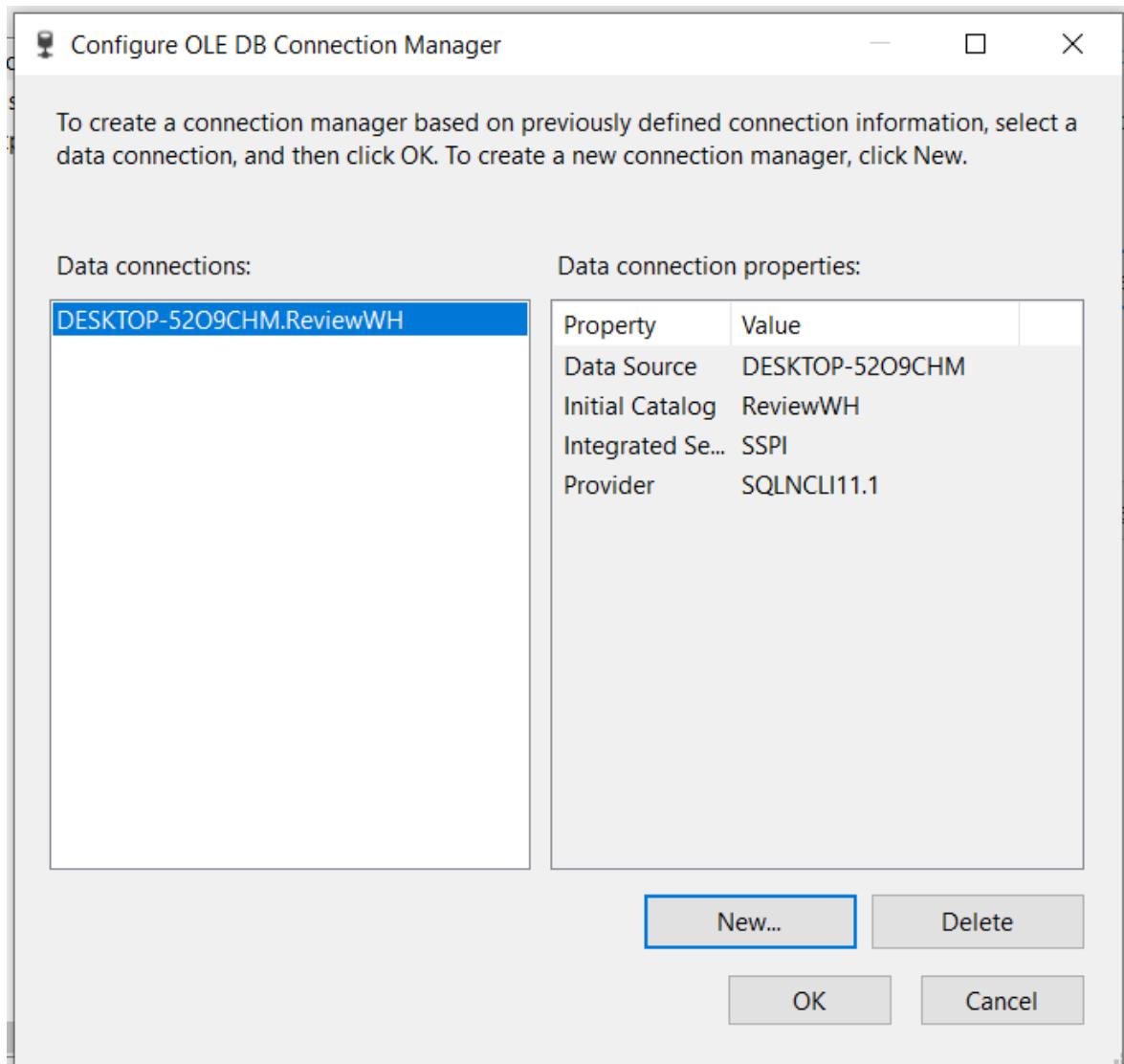
Tick chọn **Remove rows with duplicate sort values** xóa các dòng dữ liệu trùng nhau, sau đó chọn **OK**.



Bước 2. Tạo Dim_Route từ một **OLE DB Destination**. Double click vào OLE DB Destination này để tạo một connection mới đến MS SQL Server.

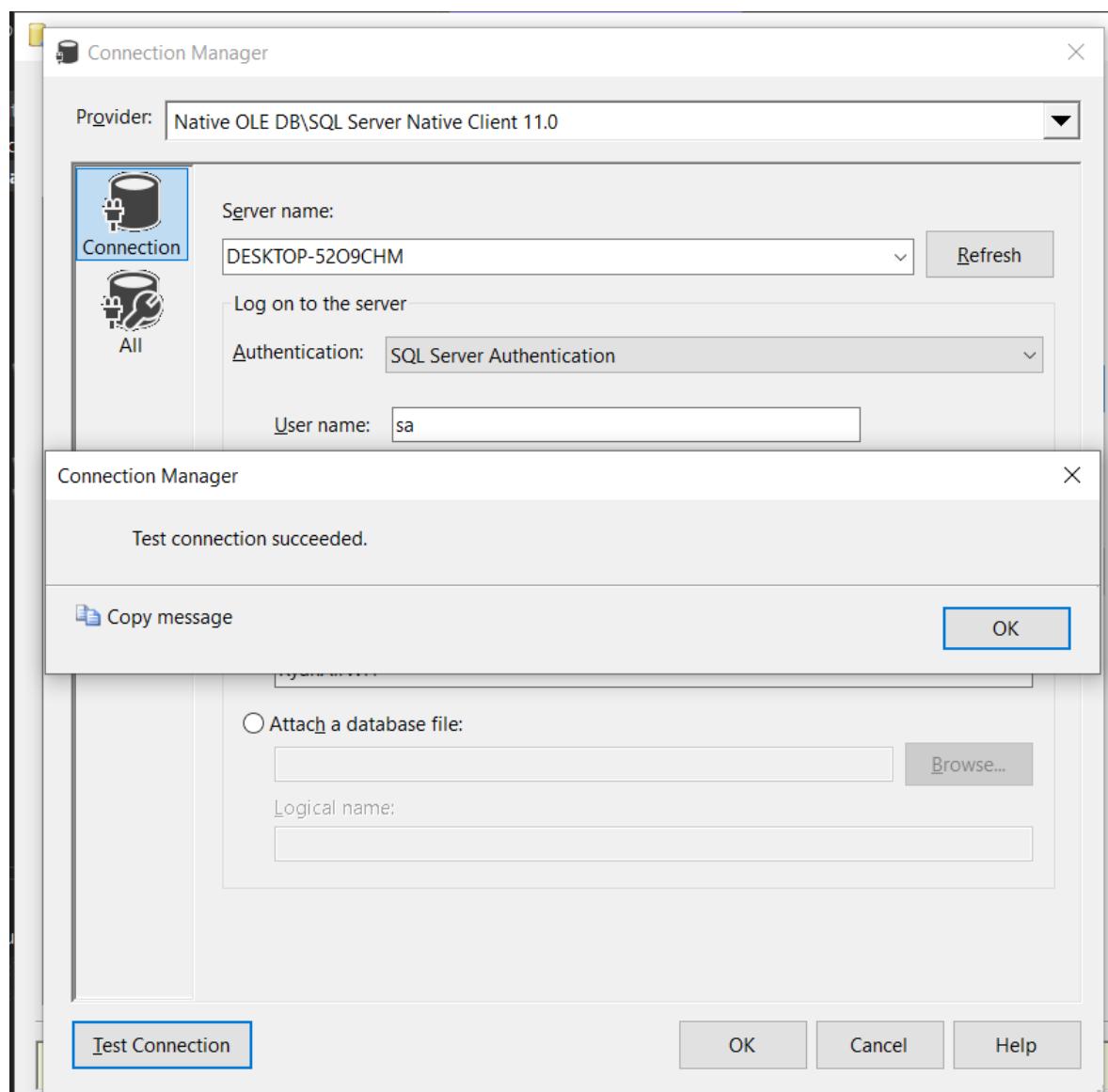


Tiếp tục chọn **New...** để tạo một connection mới:

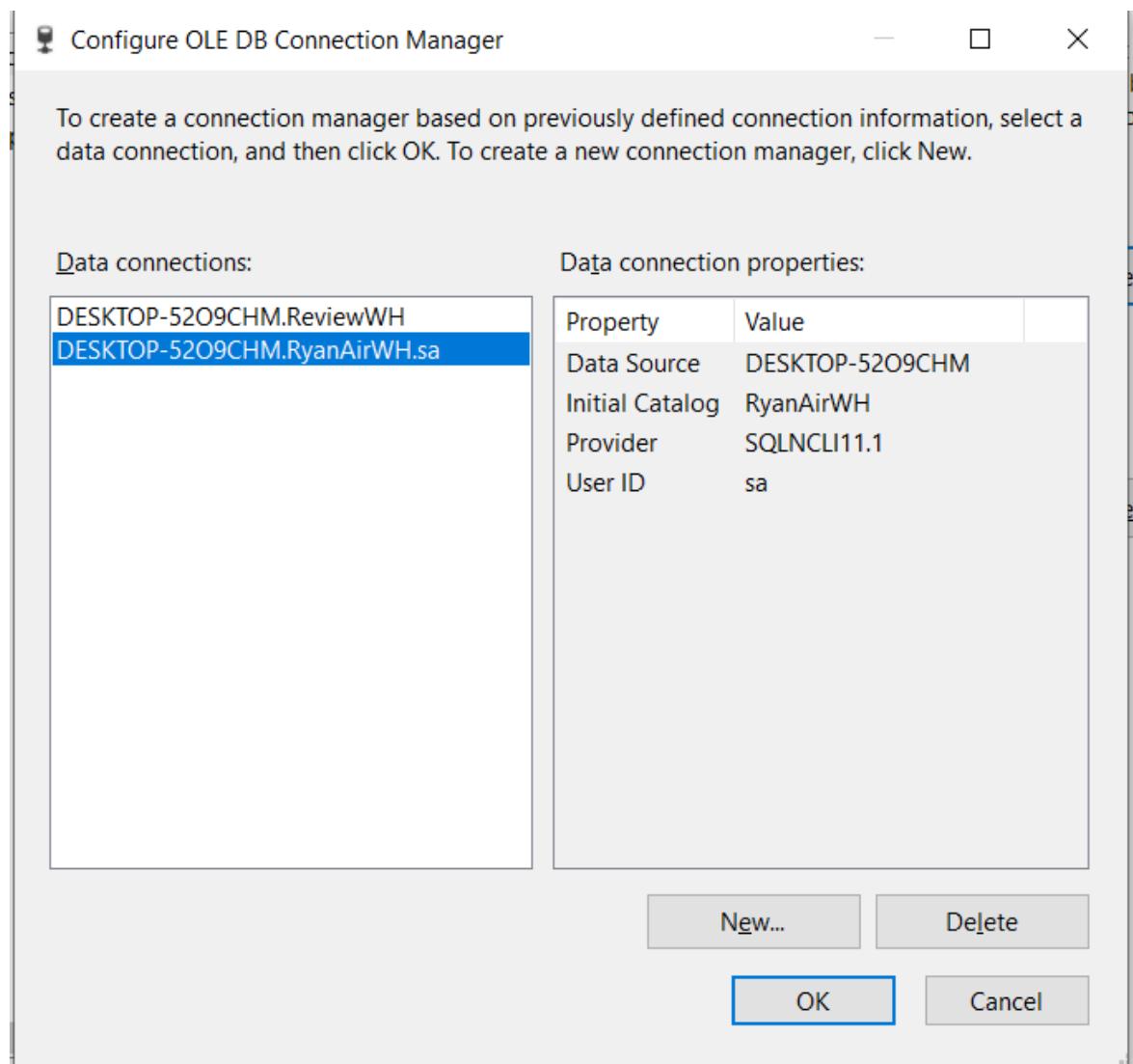


Bước 3. Chọn tên server name trùng với server name MS SQL Server để ta có thể kết nối đến datawarehouse RyanAirWH vừa tạo.

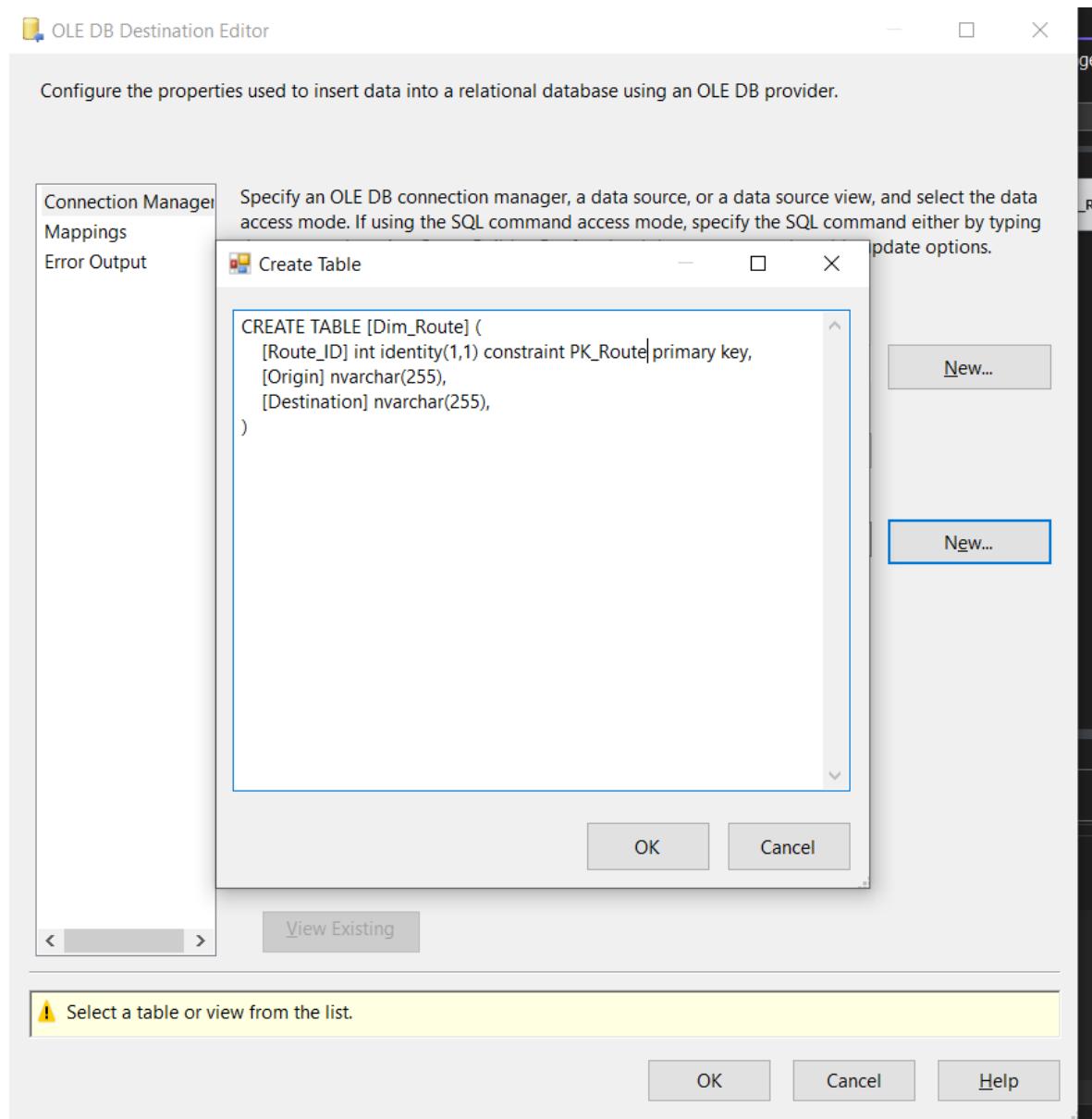
Nhấn **Test Connection** để kiểm tra kết nối



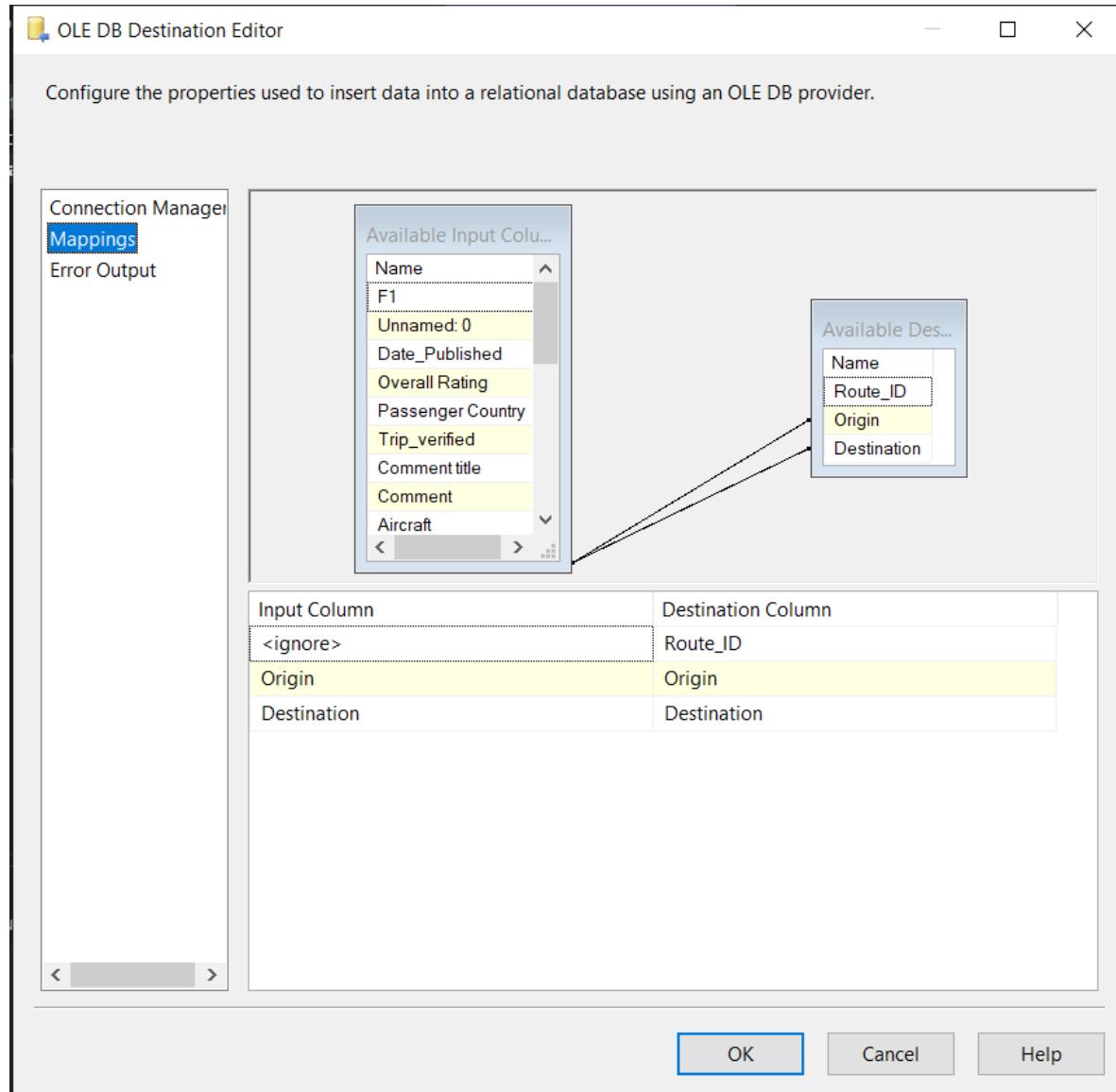
Bước 4. Chọn connection vừa tạo đến MS SQL Server và nhấn OK



Bước 5. Chọn New.. để tạo mới bảng



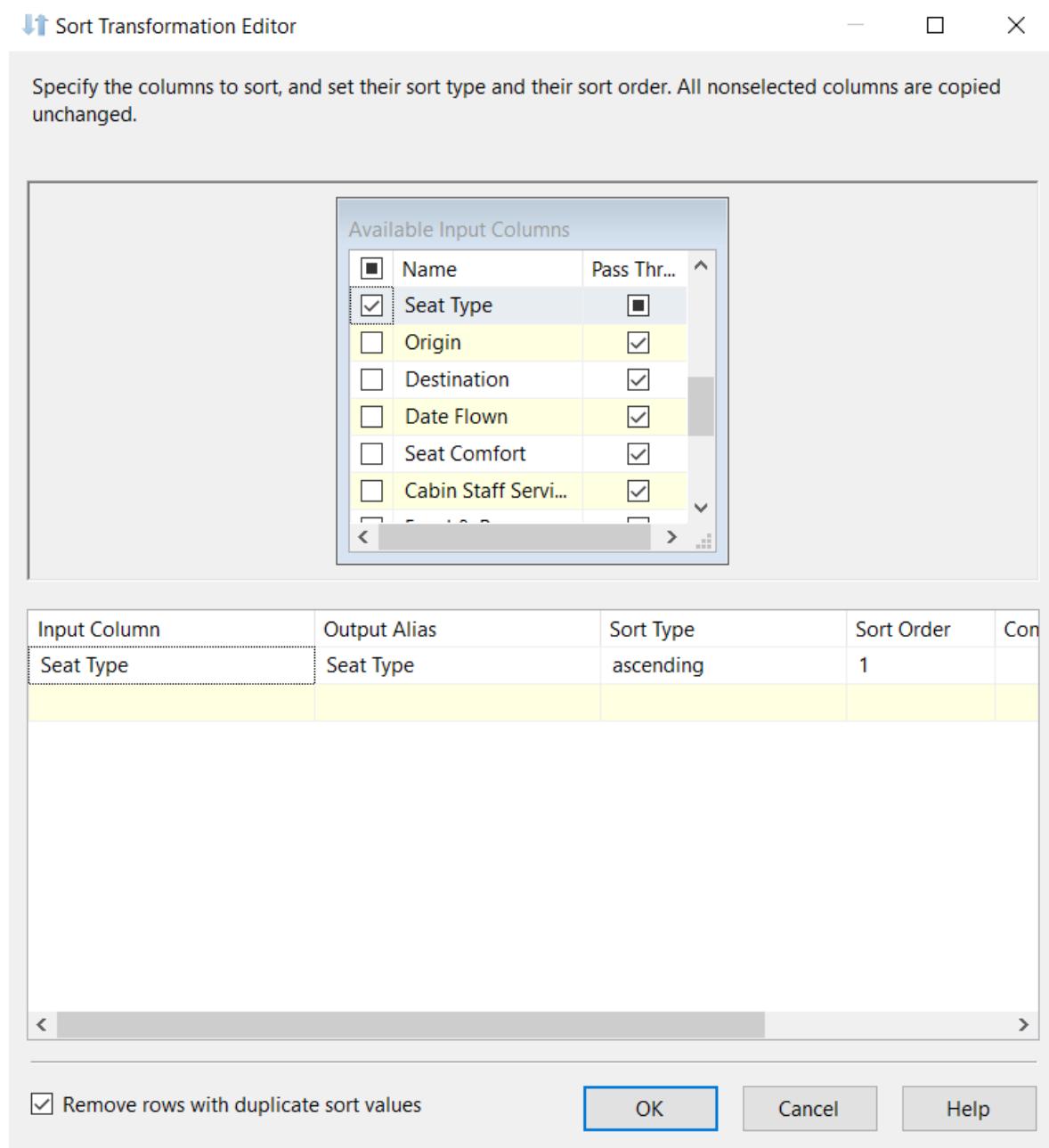
Bước 6. Tiếp đến ta cần chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu



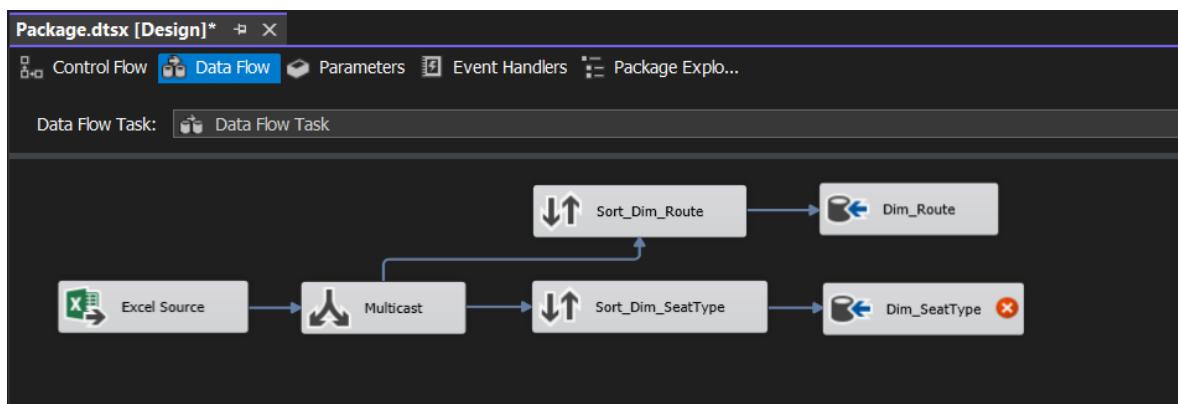
Chọn **OK** để hoàn tất thiết lập

3.3.2. Bảng Dim_SeatType

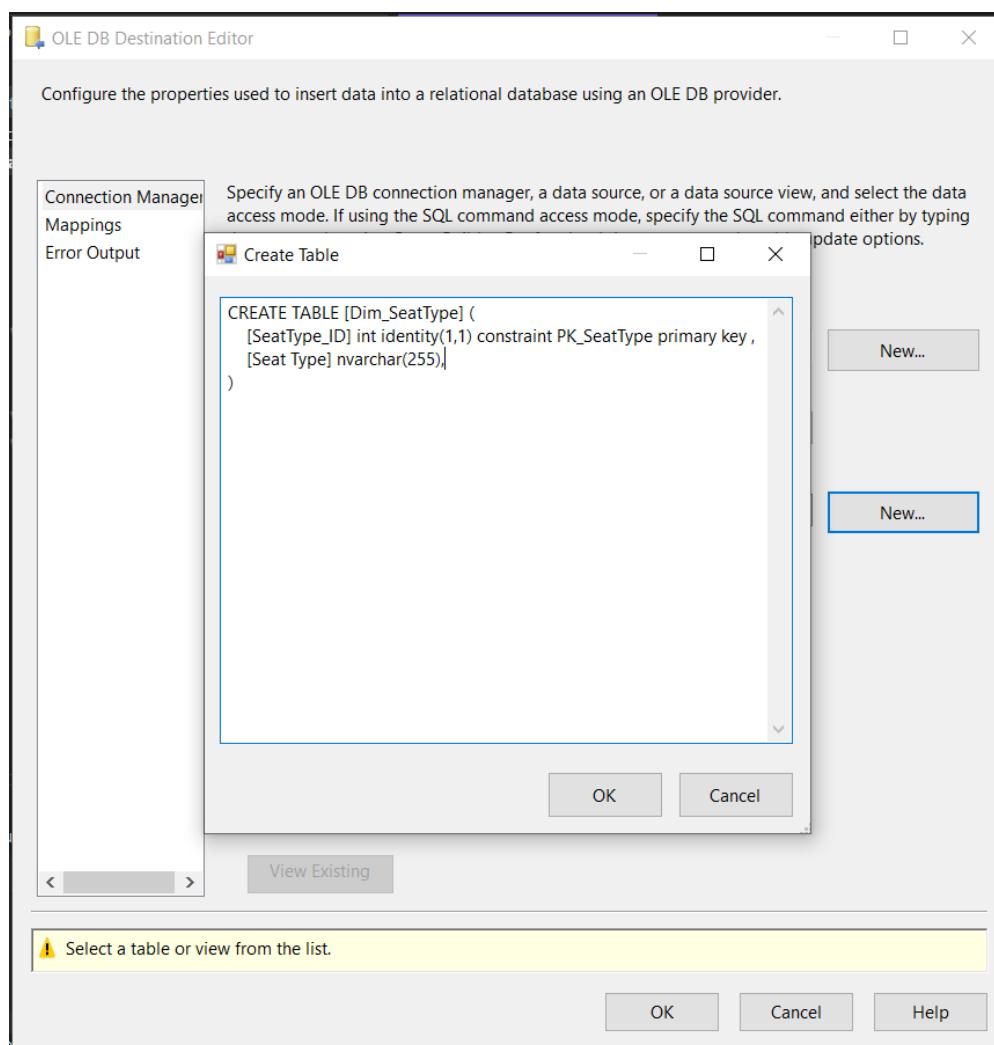
Bước 1. Tạo mới một **Sort** có tên là Sort_Dim_SeatType để lấy ra các cột dữ liệu cần thiết cho Dim_SeatType. Nhấn chuột phải và nhấn **Edit** để chọn các cột dữ liệu Tick chọn **Remove rows with duplicate sort values** xóa các dòng dữ liệu trùng nhau, sau đó chọn **OK**.



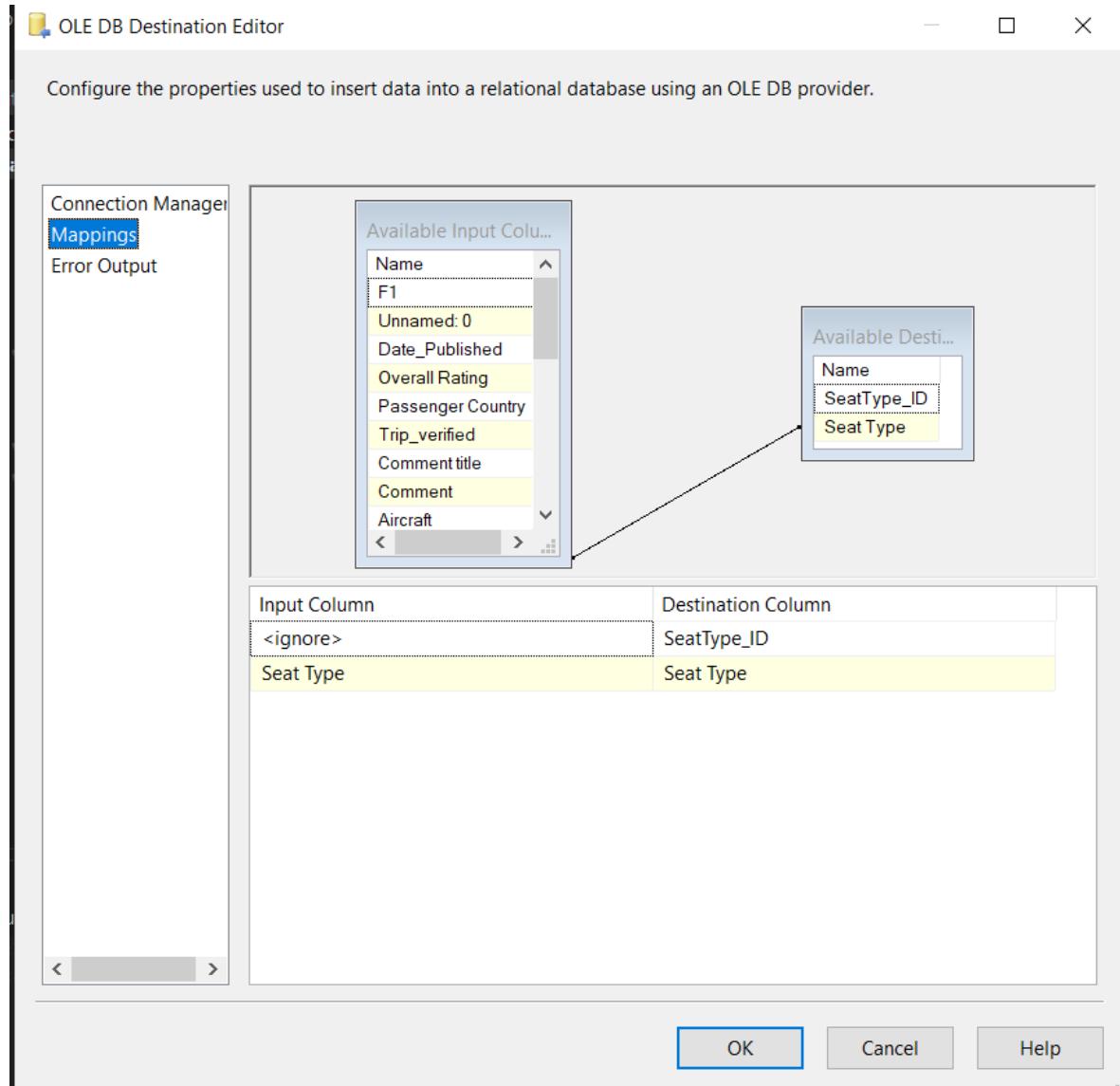
Bước 2. Tạo mới một OLE DB Destination để đổ dữ liệu gốc sau khi đã được xử lý vào trong kho dữ liệu RyanAirWH



Bước 3. Connection đến kho dữ liệu đã được tạo trước đó, vì vậy ta chỉ cần chọn New... để tạo bảng Dim_SeatType



Bước 4. Tiếp đến ta cần chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

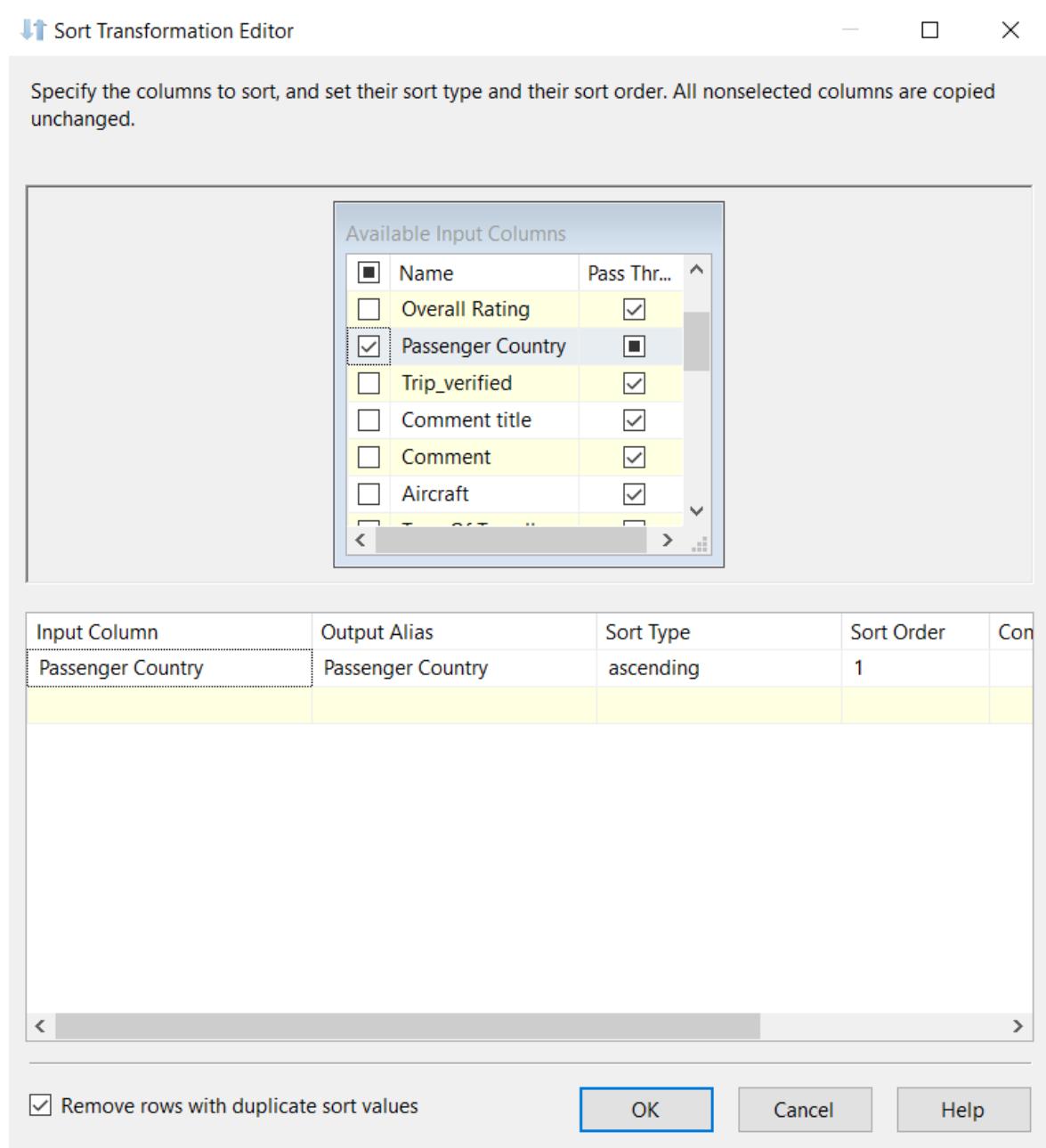


Chọn **OK** để hoàn tất thiết lập

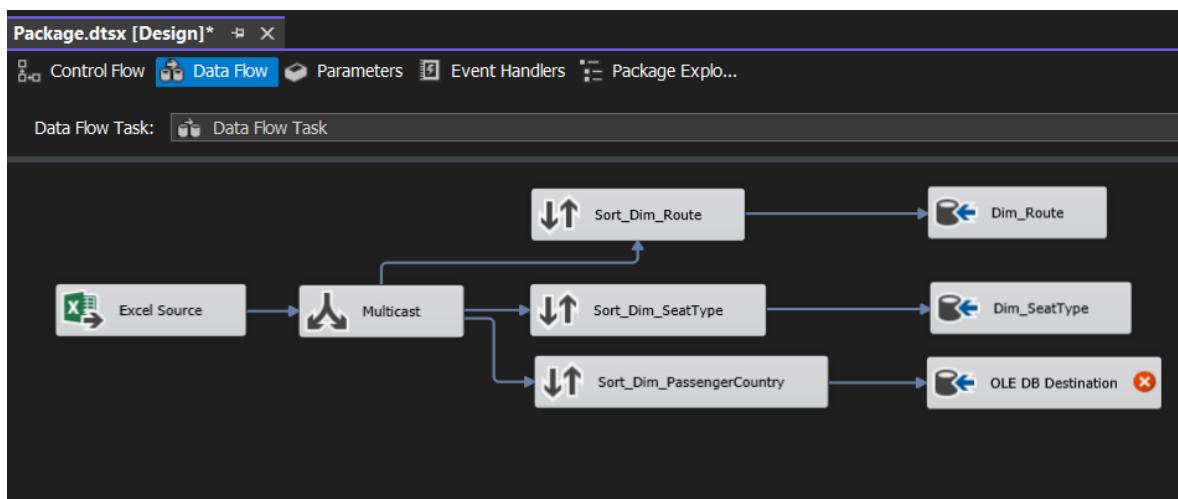
3.3.3. Bảng Dim_PassengerCountry

Bước 1. Tạo mới một **Sort** có tên là Sort_Dim_PassengerCountry để lấy ra các cột dữ liệu cần thiết cho Dim_PassengerCountry. Nhấn chuột phải và nhấn **Edit** để chọn các cột dữ liệu

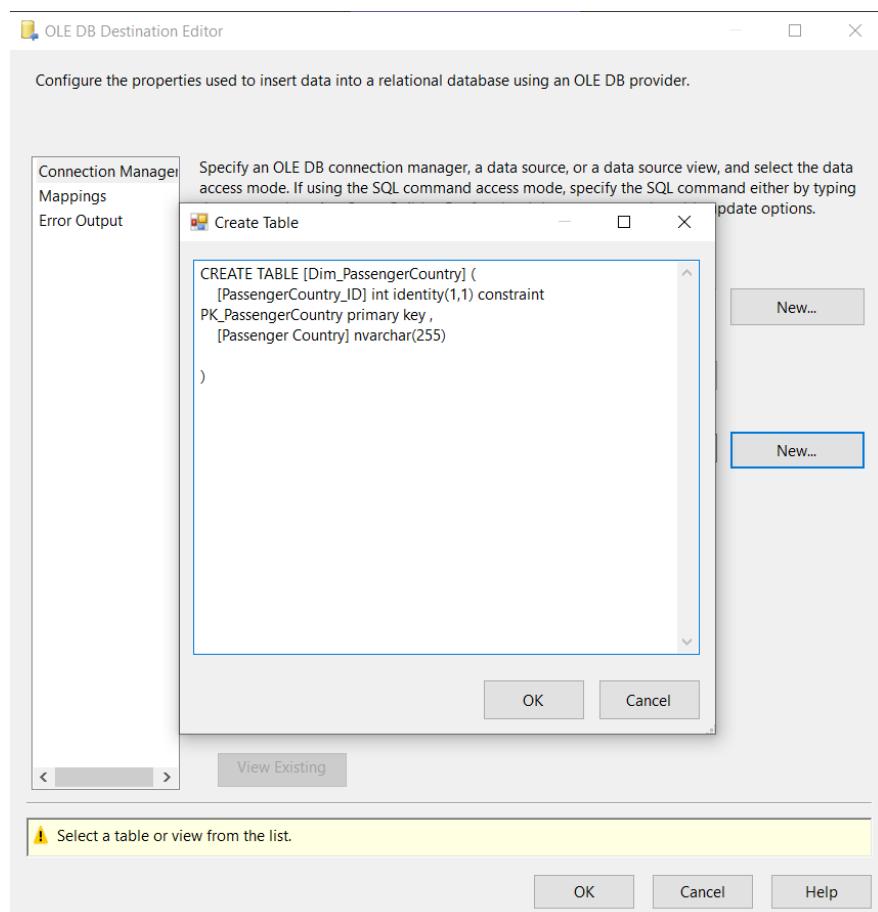
Tick chọn **Remove rows with duplicate sort values** xóa các dòng dữ liệu trùng nhau, sau đó chọn **OK**.



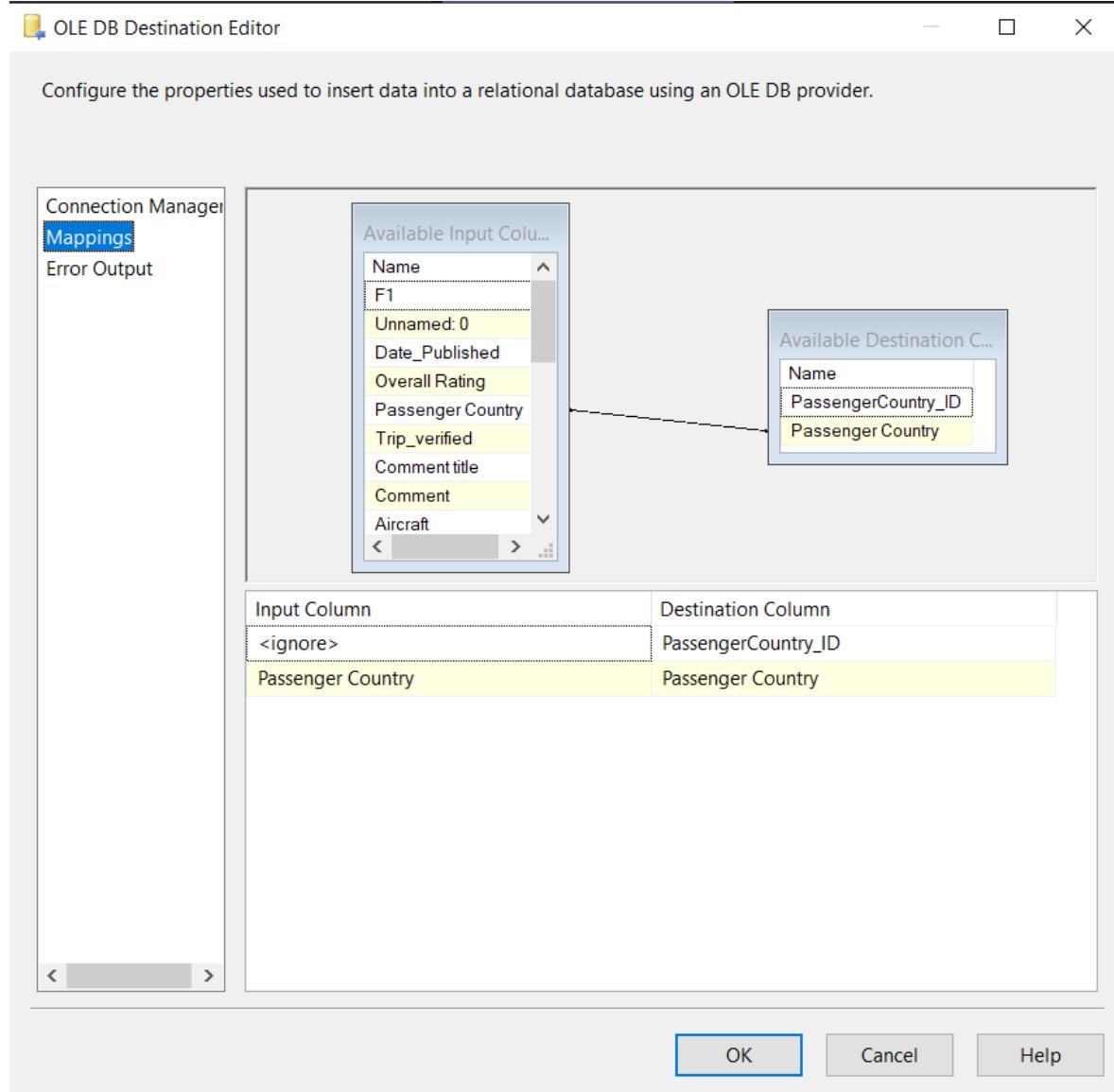
Bước 2. Tạo mới một OLE DB Destination để đổ dữ liệu gốc sau khi đã được xử lý vào trong kho dữ liệu RyanAirWH



Bước 3. Connection đến kho dữ liệu đã được tạo trước đó, vì vậy ta chỉ cần chọn New... để tạo bảng Dim_PassengerCountry



Bước 4. Tiếp đến ta cần chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

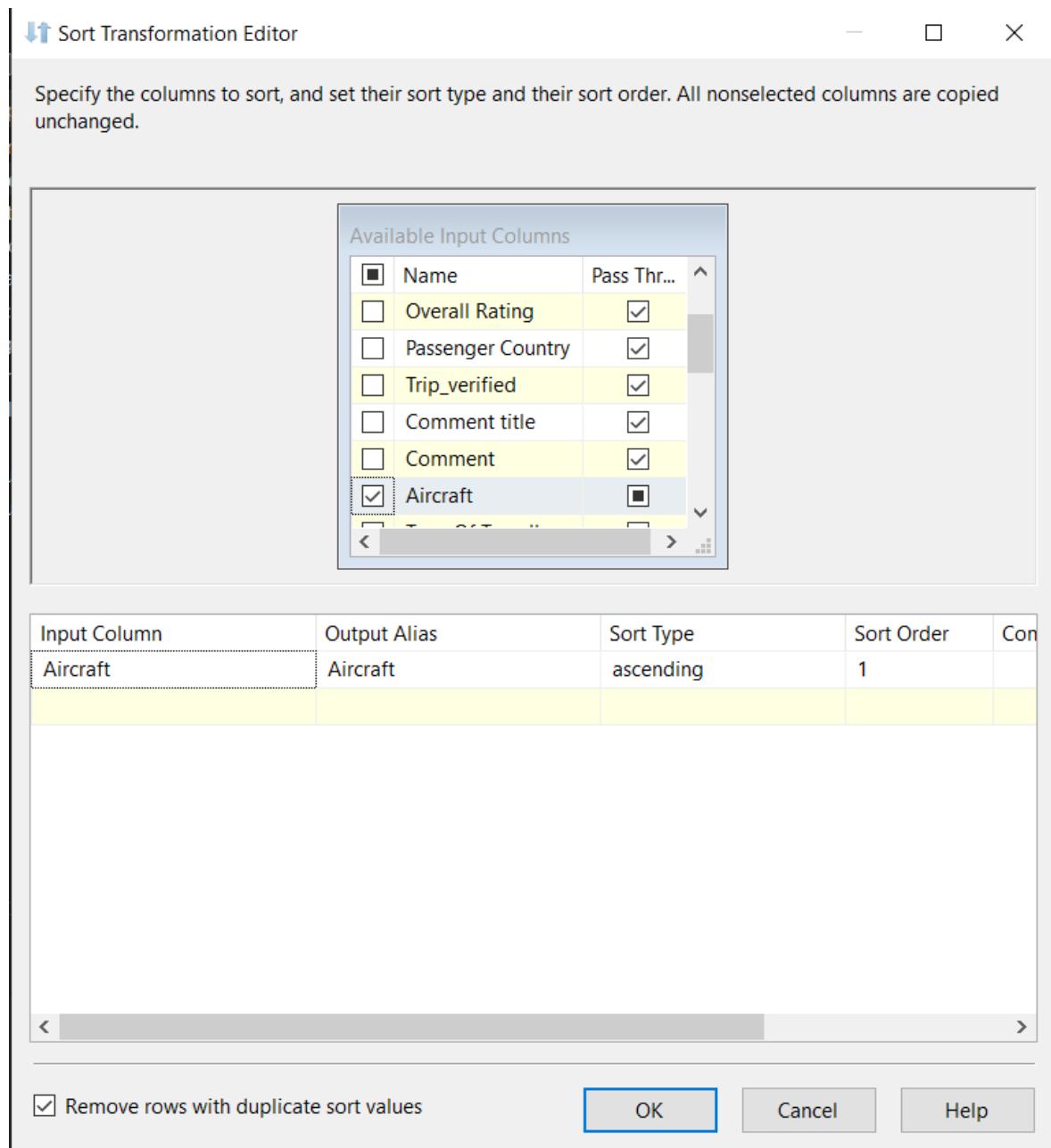


Chọn **OK** để hoàn tất thiết lập

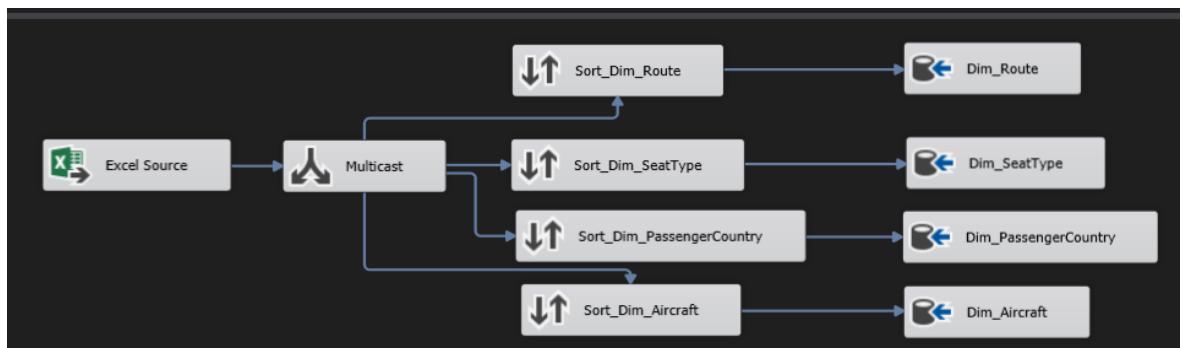
3.3.4. Bảng Dim_Aircraft

Bước 1. Tạo mới một **Sort** có tên là Sort_Dim_Aircraft để lấy ra các cột dữ liệu cần thiết cho Dim_Aircraft. Nhấn chuột phải và nhấp **Edit** để chọn các cột dữ liệu

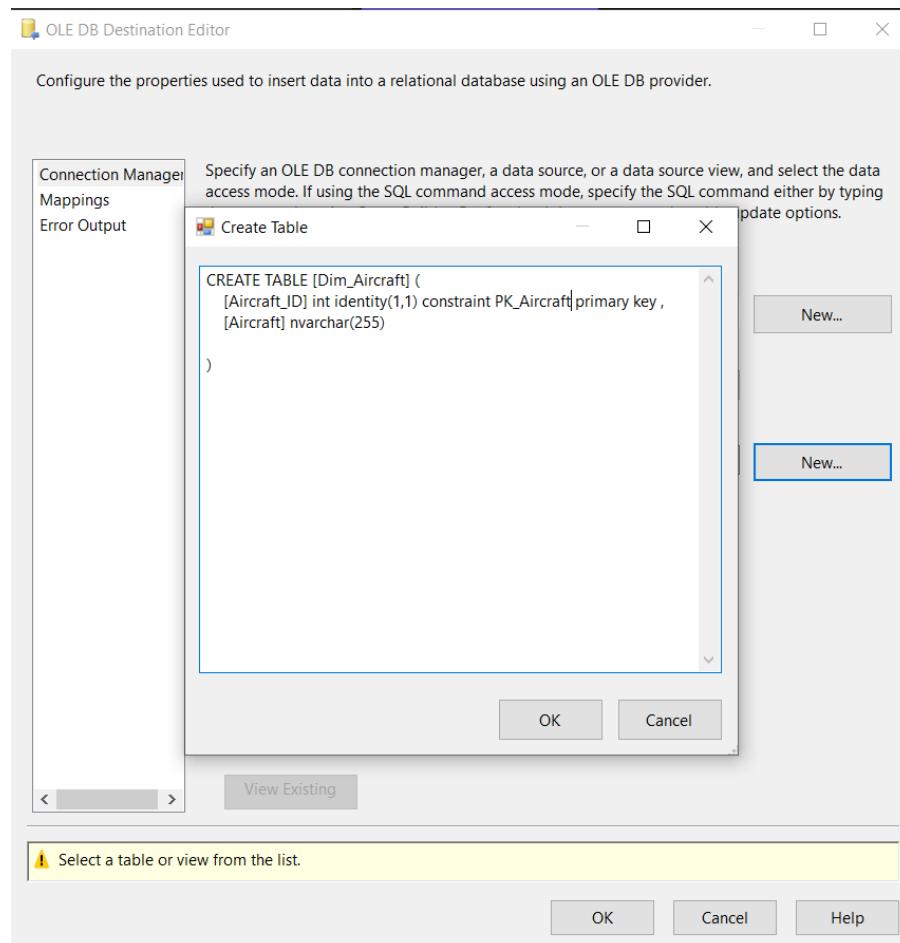
Tick chọn **Remove rows with duplicate sort values** xóa các dòng dữ liệu trùng nhau, sau đó chọn **OK**.



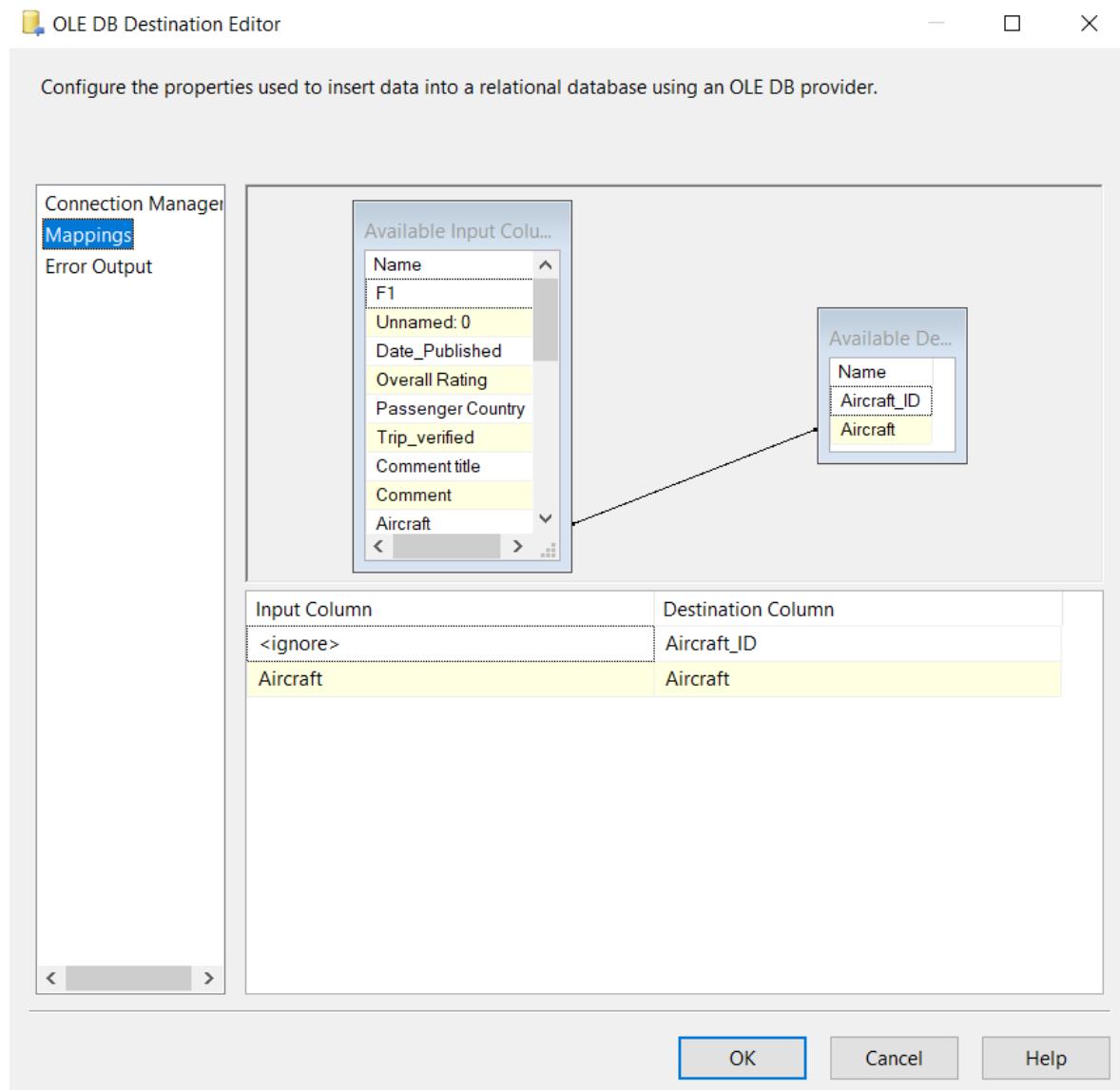
Bước 2. Tạo mới một OLE DB Destination để đổ dữ liệu gốc sau khi đã được xử lý vào trong kho dữ liệu RyanAirWH



Bước 3. Connection đến kho dữ liệu đã được tạo trước đó, vì vậy ta chỉ cần chọn New... để tạo bảng Dim_Aircraft



Bước 4. Tiếp đến ta cần chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

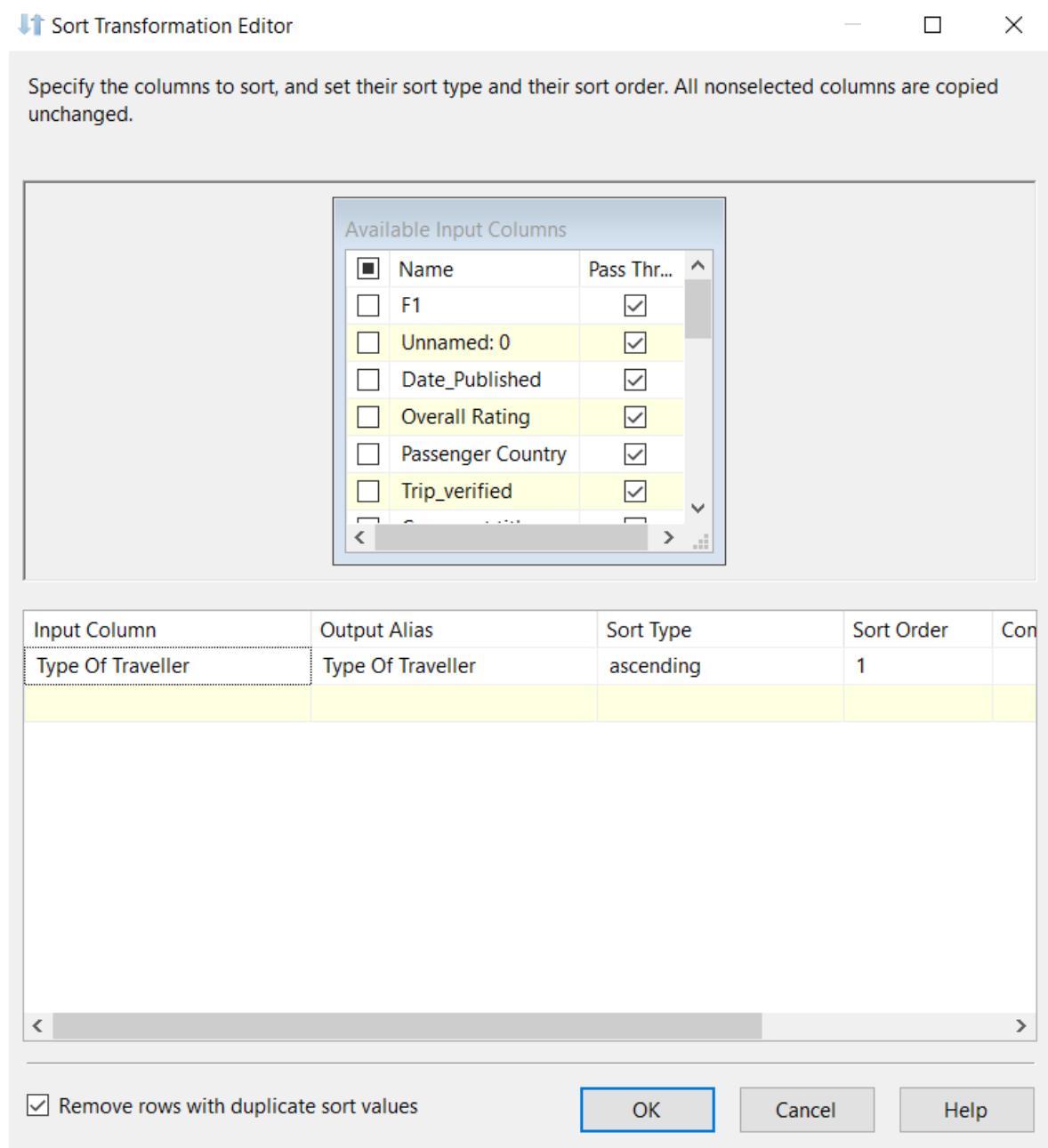


Chọn **OK** để hoàn tất thiết lập

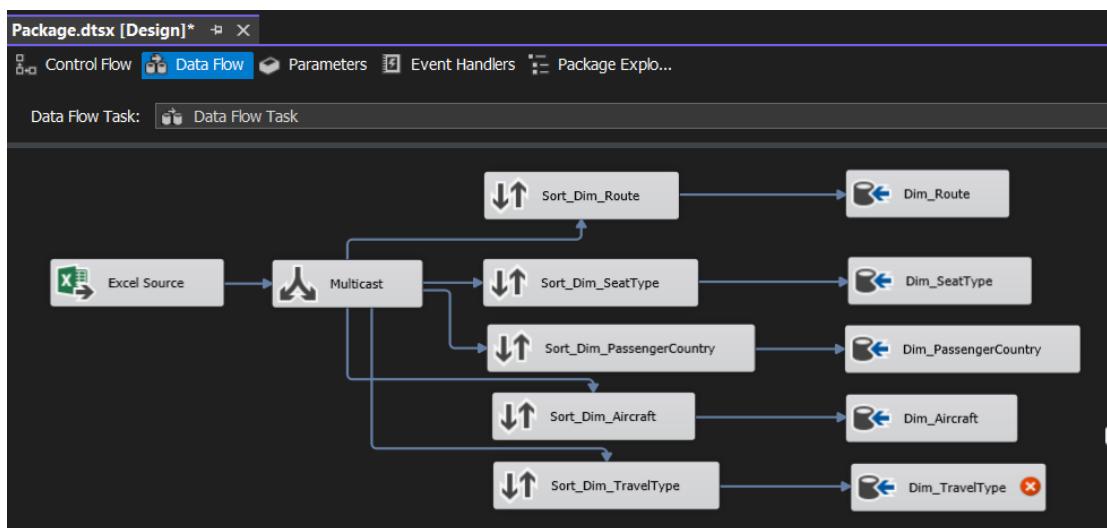
3.3.5. Bảng Dim_TravelType

Bước 1. Tạo mới một **Sort** có tên là Sort_Dim_TravelType để lấy ra các cột dữ liệu cần thiết cho Dim_TravelType. Nhấn chuột phải và nhấp **Edit** để chọn các cột dữ liệu

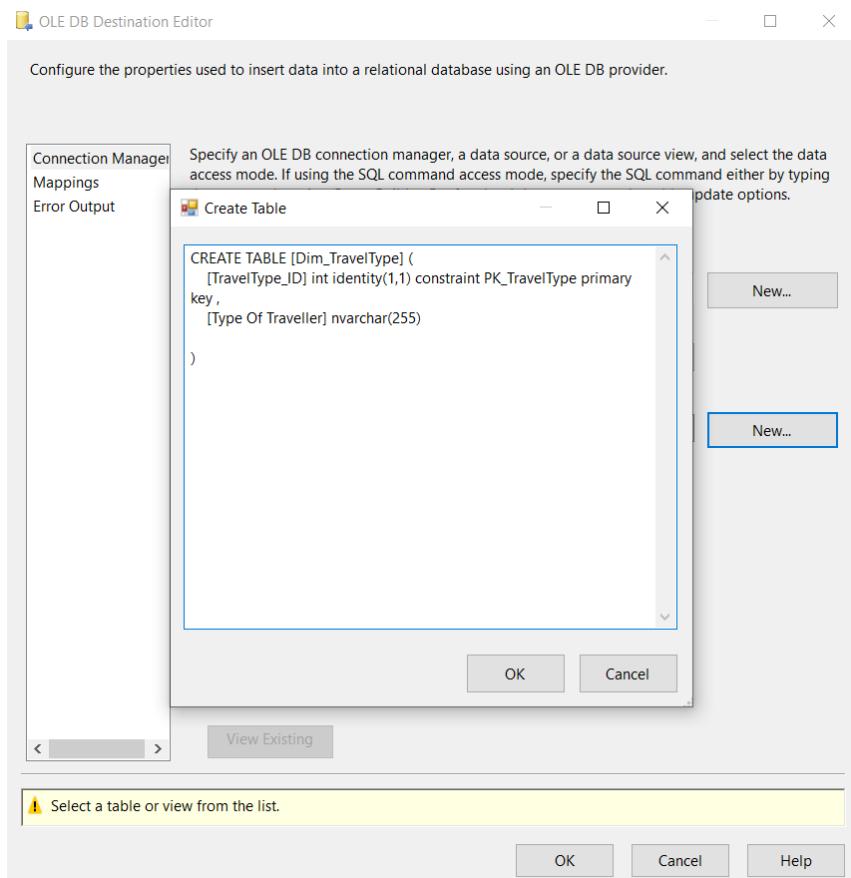
Tick chọn **Remove rows with duplicate sort values** xóa các dòng dữ liệu trùng nhau, sau đó chọn **OK**.



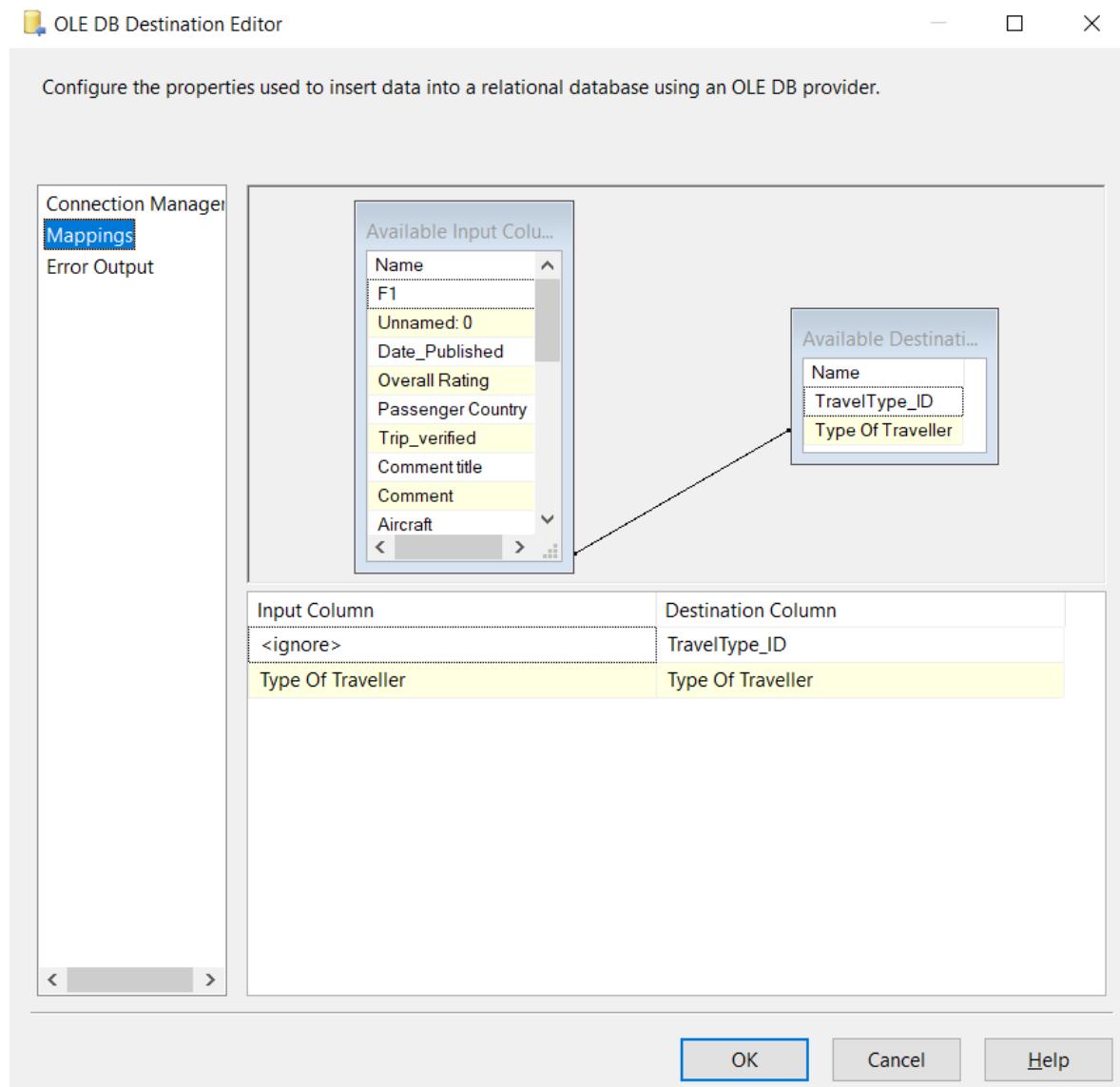
Bước 2. Tạo mới một OLE DB Destination để đổ dữ liệu gốc sau khi đã được xử lý vào trong kho dữ liệu RyanAirWH



Bước 3. Connection đến kho dữ liệu đã được tạo trước đó, vì vậy ta chỉ cần chọn New... để tạo bảng Dim_TravelType



Bước 4. Tiếp đến ta cần chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

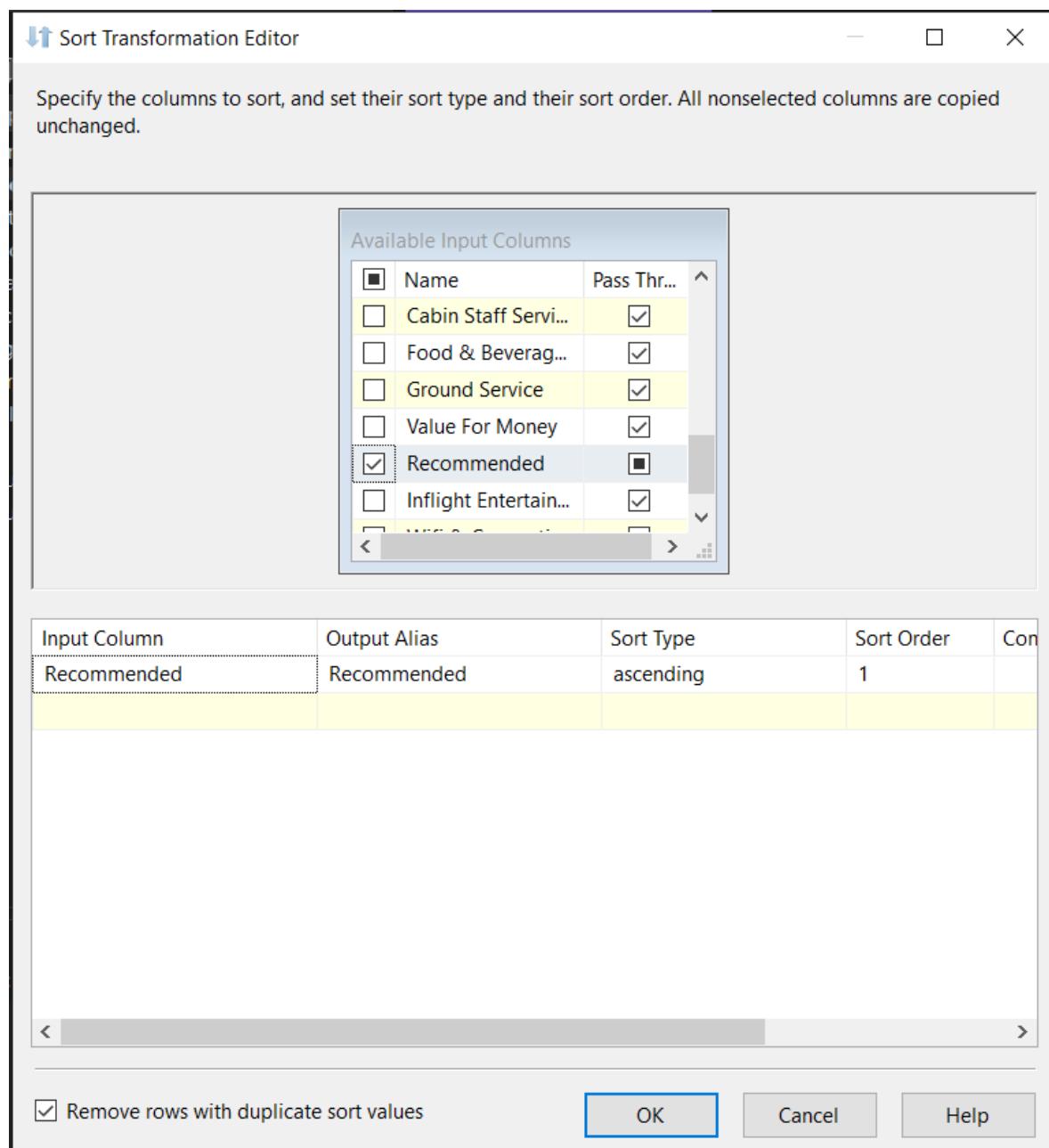


Chọn **OK** để hoàn tất thiết lập

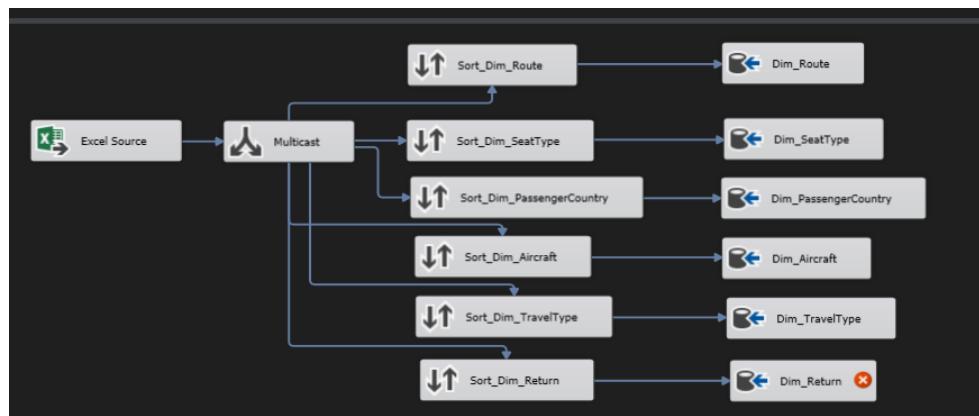
3.3.6. Bảng Dim_Return

Bước 1. Tạo mới một **Sort** có tên là Sort_Dim_Return để lấy ra các cột dữ liệu cần thiết cho Dim_Return. Nhấn chuột phải và nhấp **Edit** để chọn các cột dữ liệu

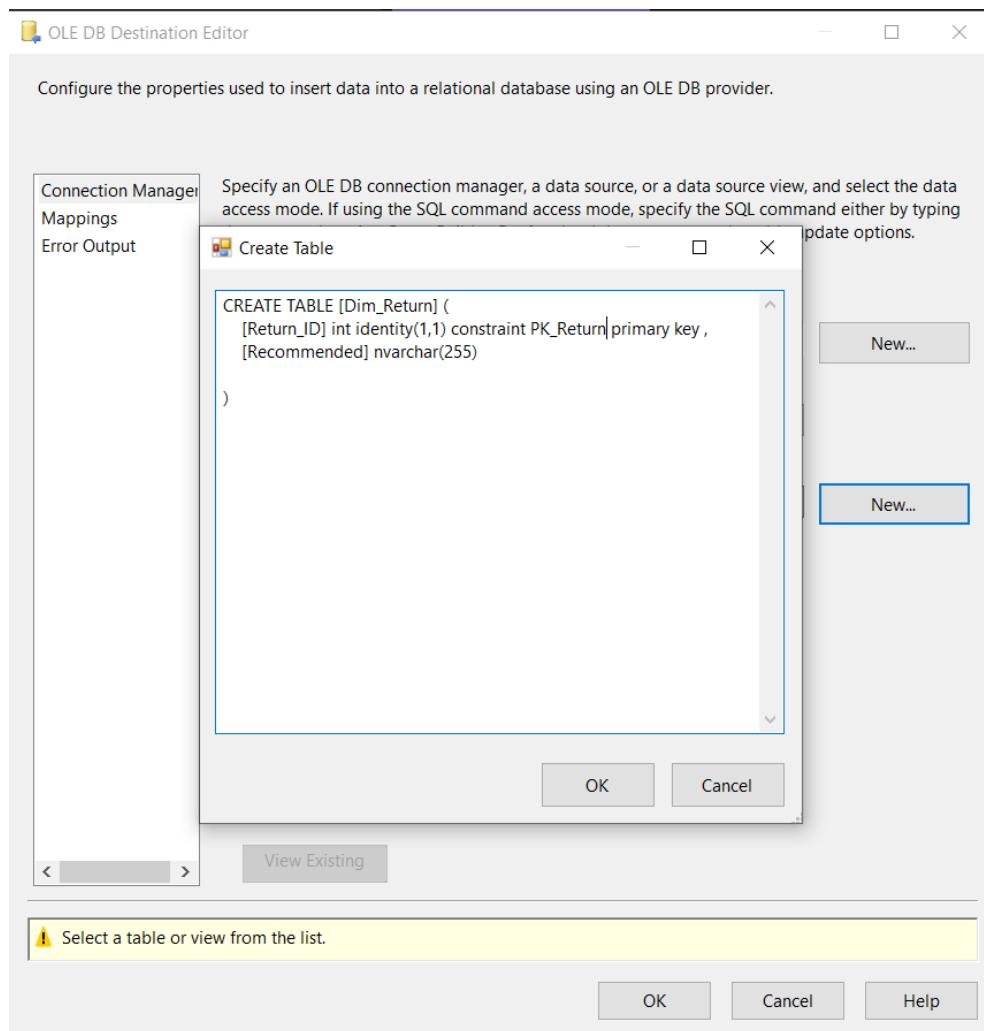
Tick chọn **Remove rows with duplicate sort values** xóa các dòng dữ liệu trùng nhau, sau đó chọn **OK**.



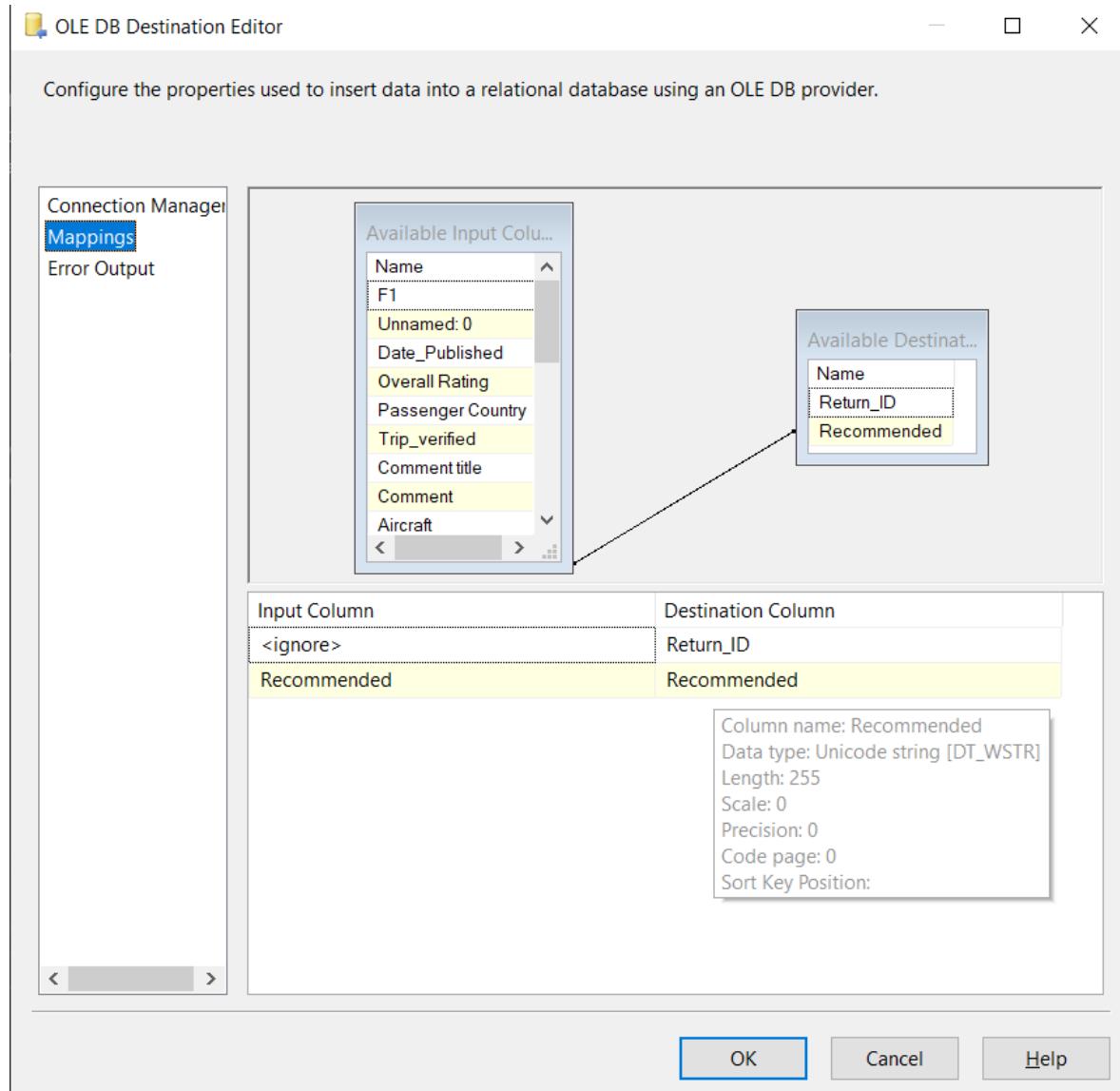
Bước 2. Tạo mới một OLE DB Destination để đổ dữ liệu gốc sau khi đã được xử lý vào trong kho dữ liệu RyanAirWH



Bước 3. Connection đến kho dữ liệu đã được tạo trước đó, vì vậy ta chỉ cần chọn New... để tạo bảng Dim_Return



Bước 4. Tiếp đến ta cần chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

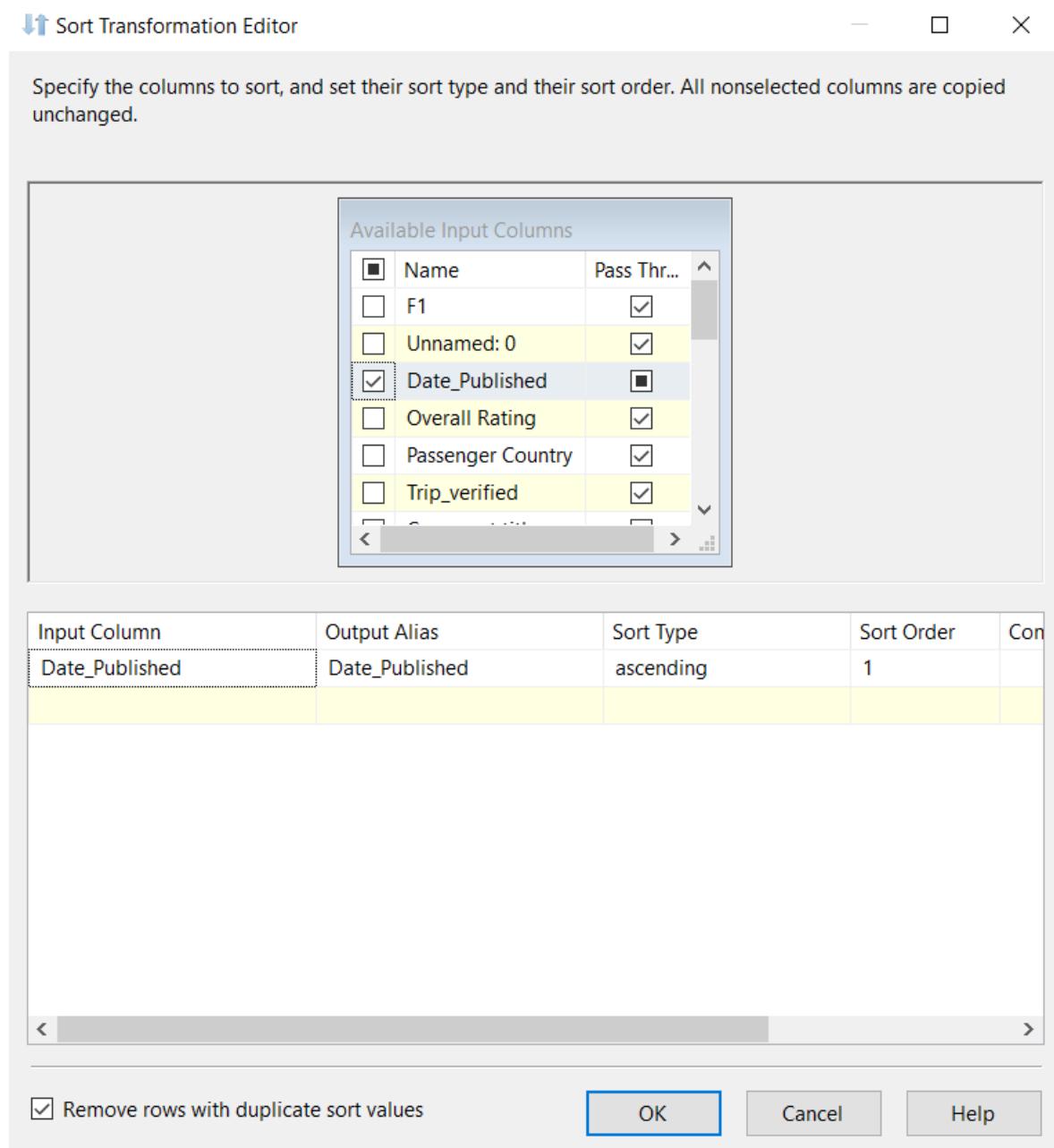


Chọn **OK** để hoàn tất thiết lập

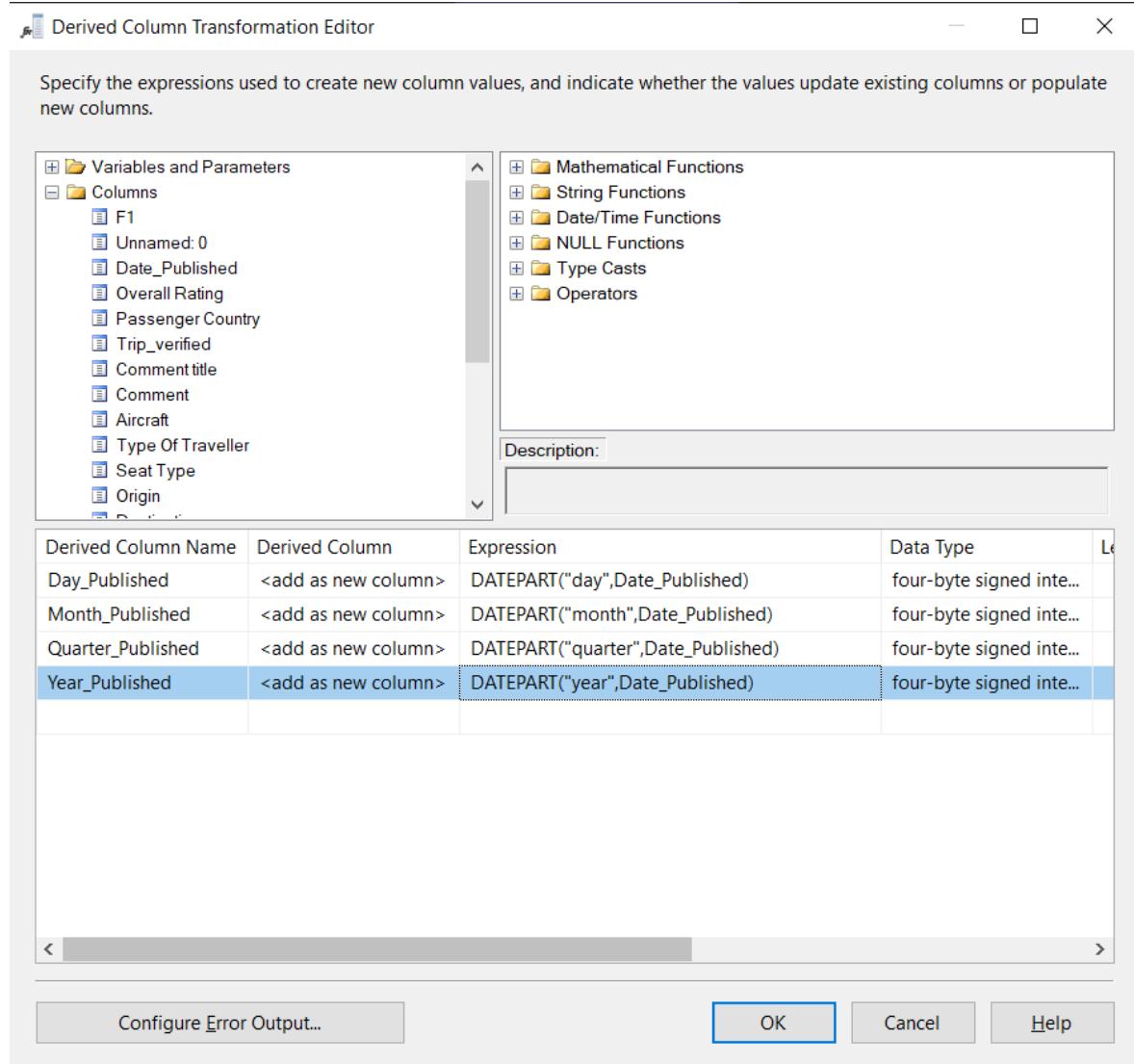
3.3.7. Bảng Dim_Time

Bước 1. Tạo mới một **Sort** có tên là Sort_Dim_Time để lấy ra các cột dữ liệu cần thiết cho Dim_Time. Nhấn chuột phải và nhấn **Edit** để chọn các cột dữ liệu

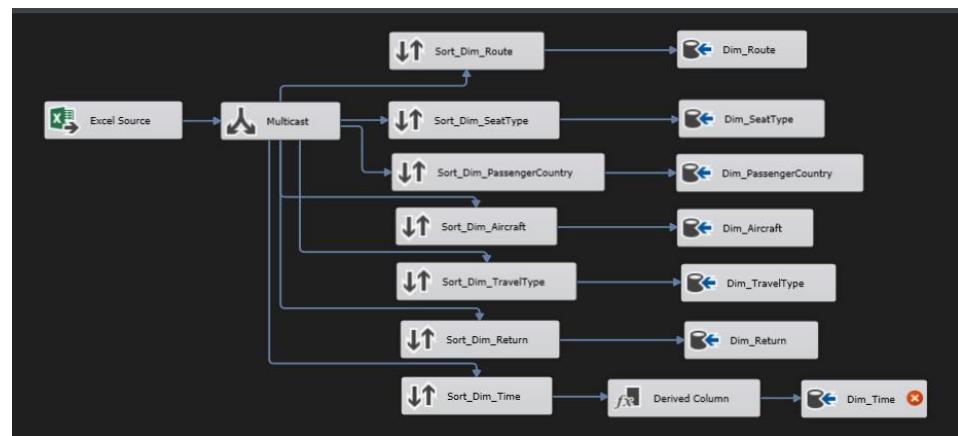
Tick chọn **Remove rows with duplicate sort values** xóa các dòng dữ liệu trùng nhau, sau đó chọn **OK**.



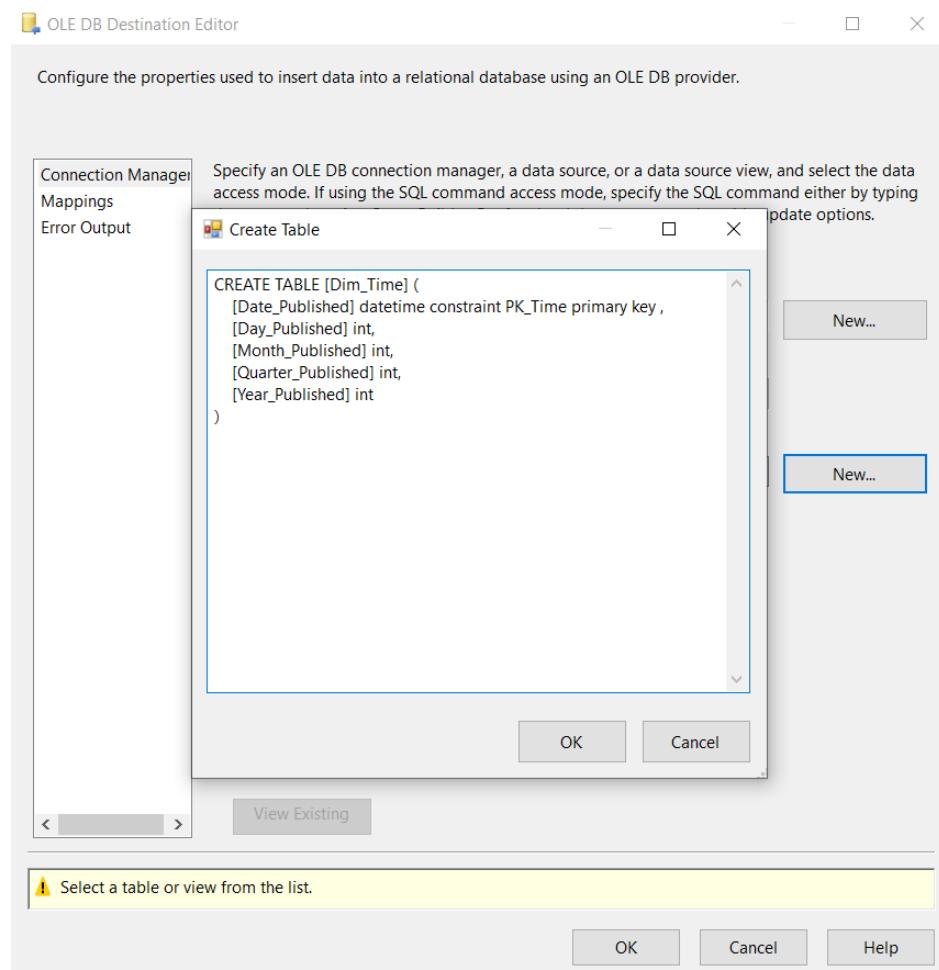
Bước 2. Tạo Derived Column, nhấn Double click sau đó thực hiện chia dữ liệu thời gian thành day, month, quarter và year. Sau đó nhấn **OK**



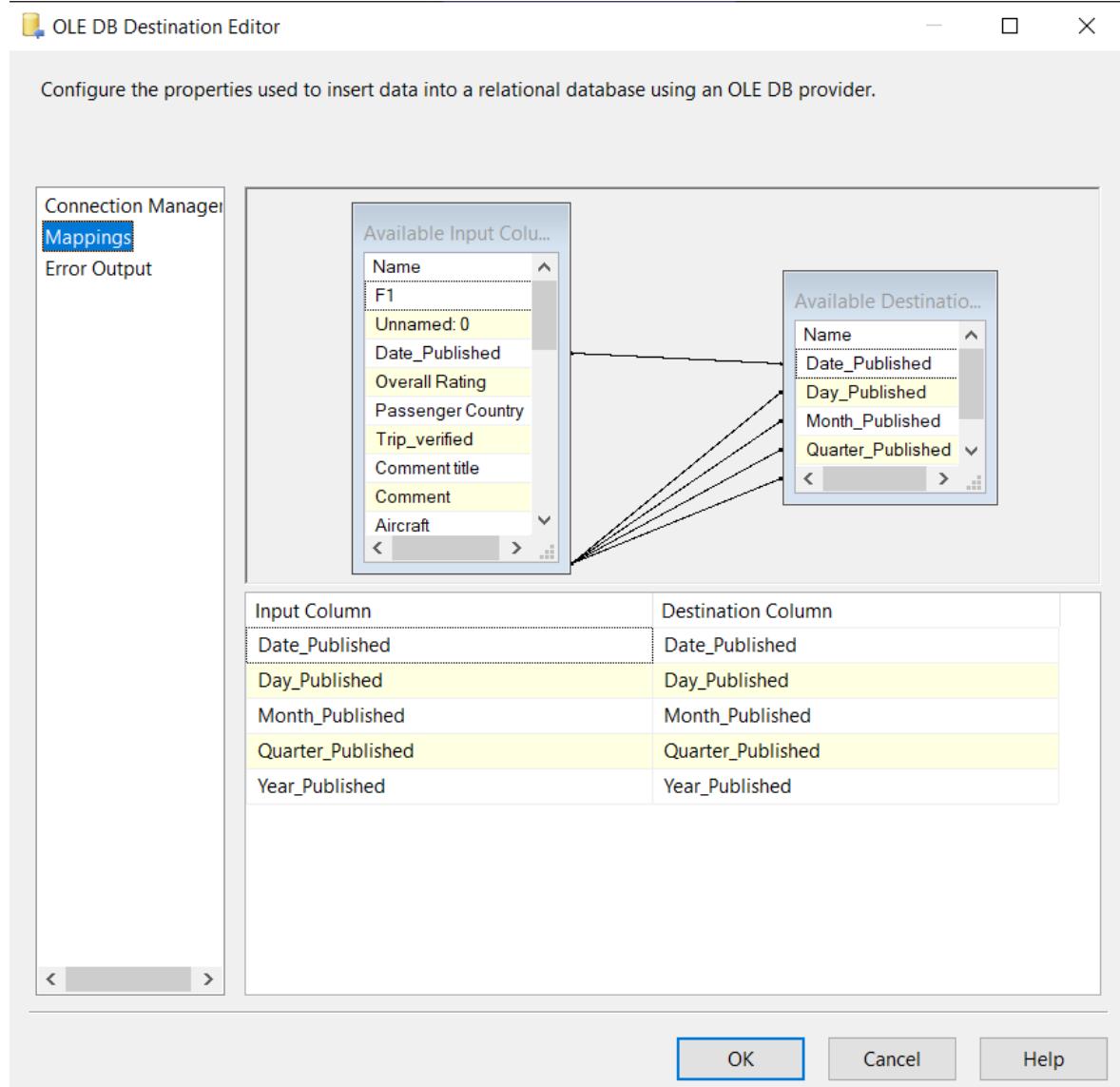
Bước 3. Tạo mới một OLE DB Destination để đổ dữ liệu gốc sau khi đã được xử lý vào trong kho dữ liệu RyanAirWH



Bước 4. Connection đến kho dữ liệu đã được tạo trước đó, vì vậy ta chỉ cần chọn New... để tạo bảng Dim_Time



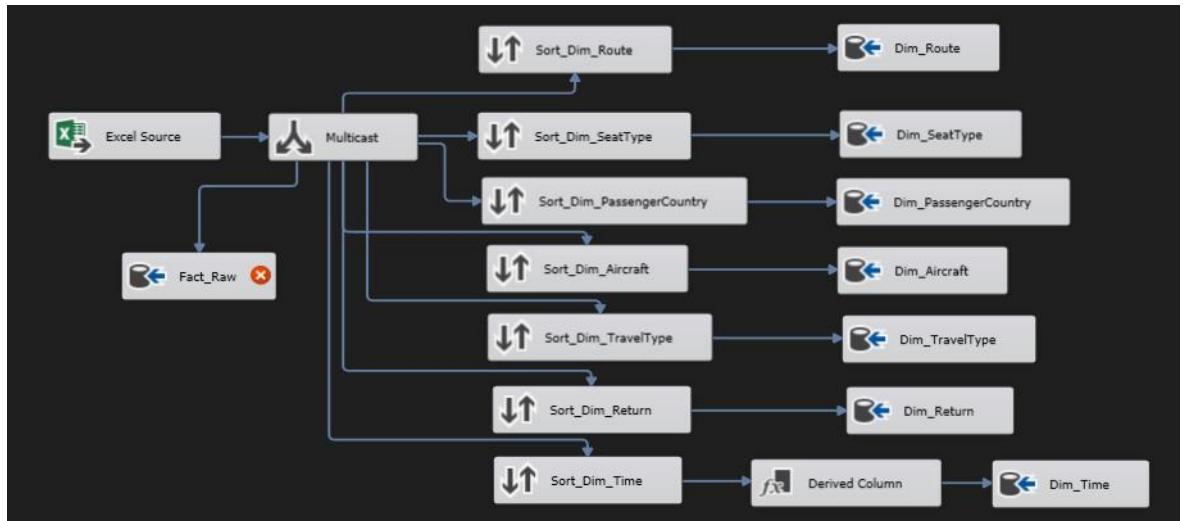
Bước 5. Tiếp đến ta cần chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu



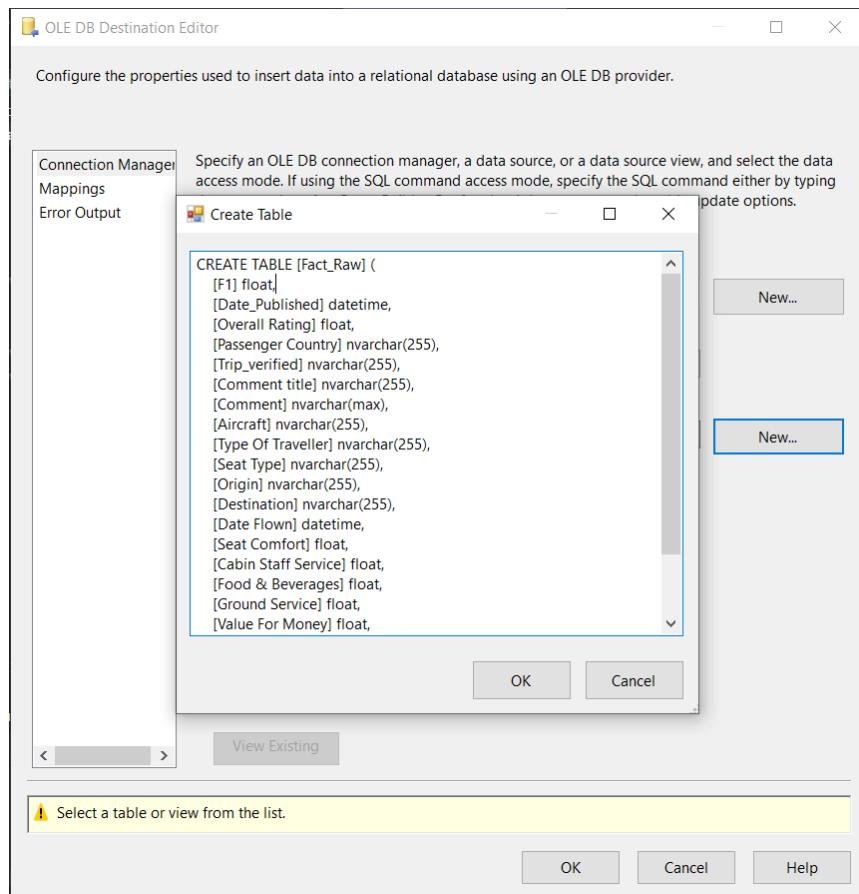
Chọn **OK** để hoàn tất thiết lập

3.3.8. Bảng Fact

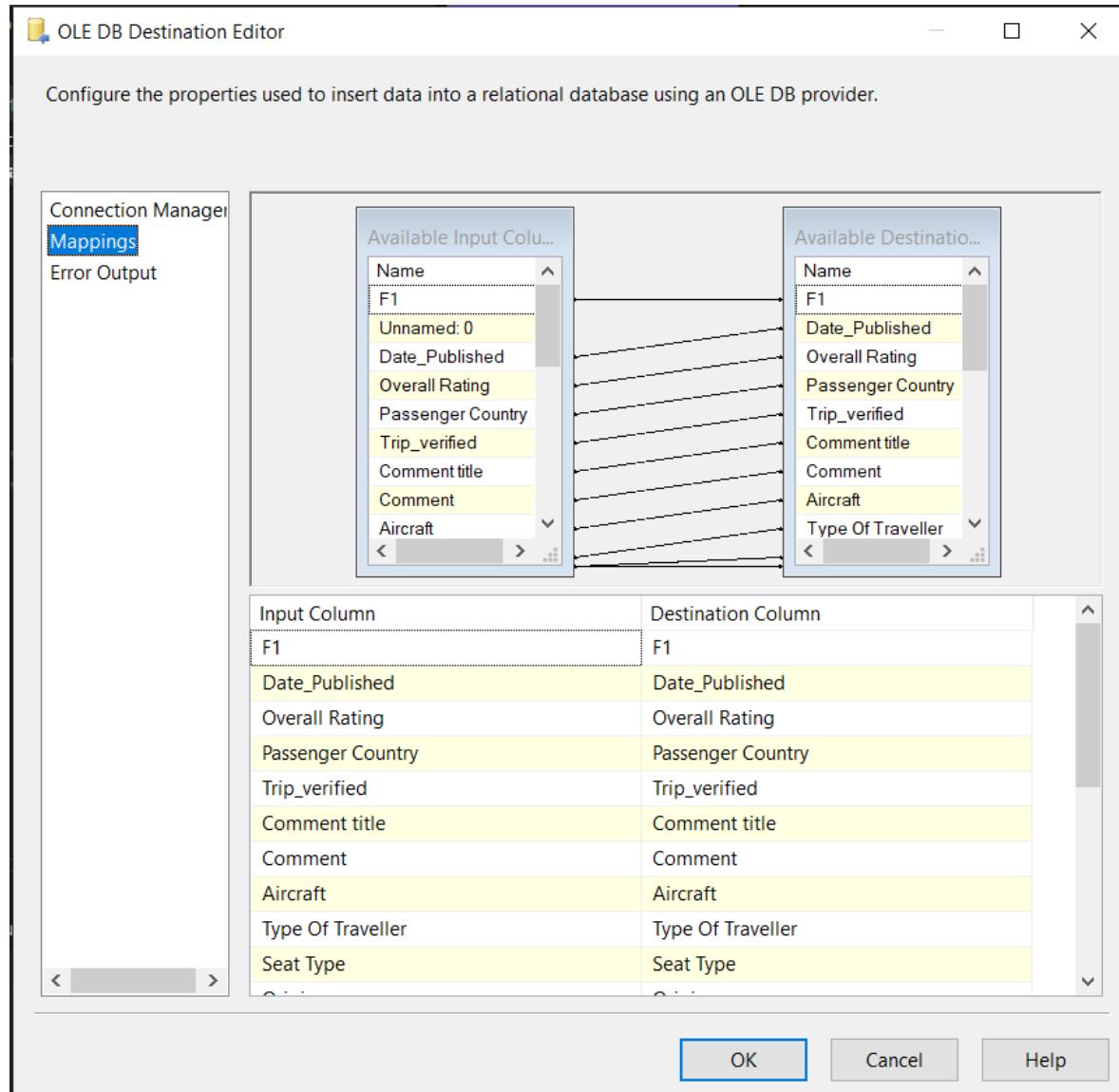
Bước 1. Tiến hành tạo bảng Fact và đặt tên là Fact_Raw từ một OLE DB Destination



Bước 2. Chọn New... để tạo bảng Fact_Raw



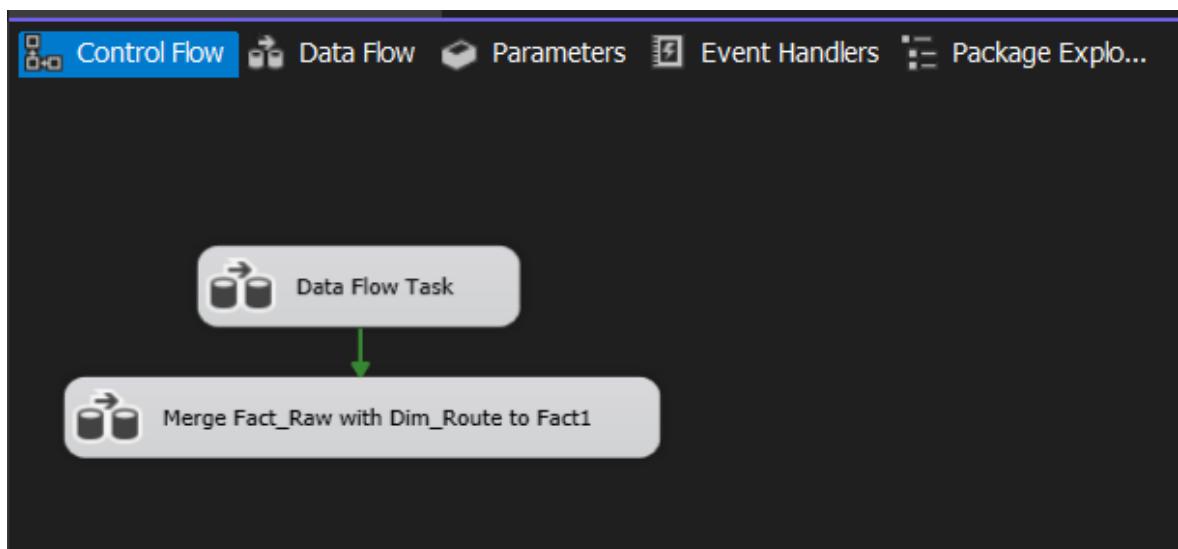
Bước 3. Tiếp đến ta cần chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu



Chọn **OK** để hoàn tất thiết lập

3.3.8.1. Merge Fact_Raw và Dim_Route vào Fact1

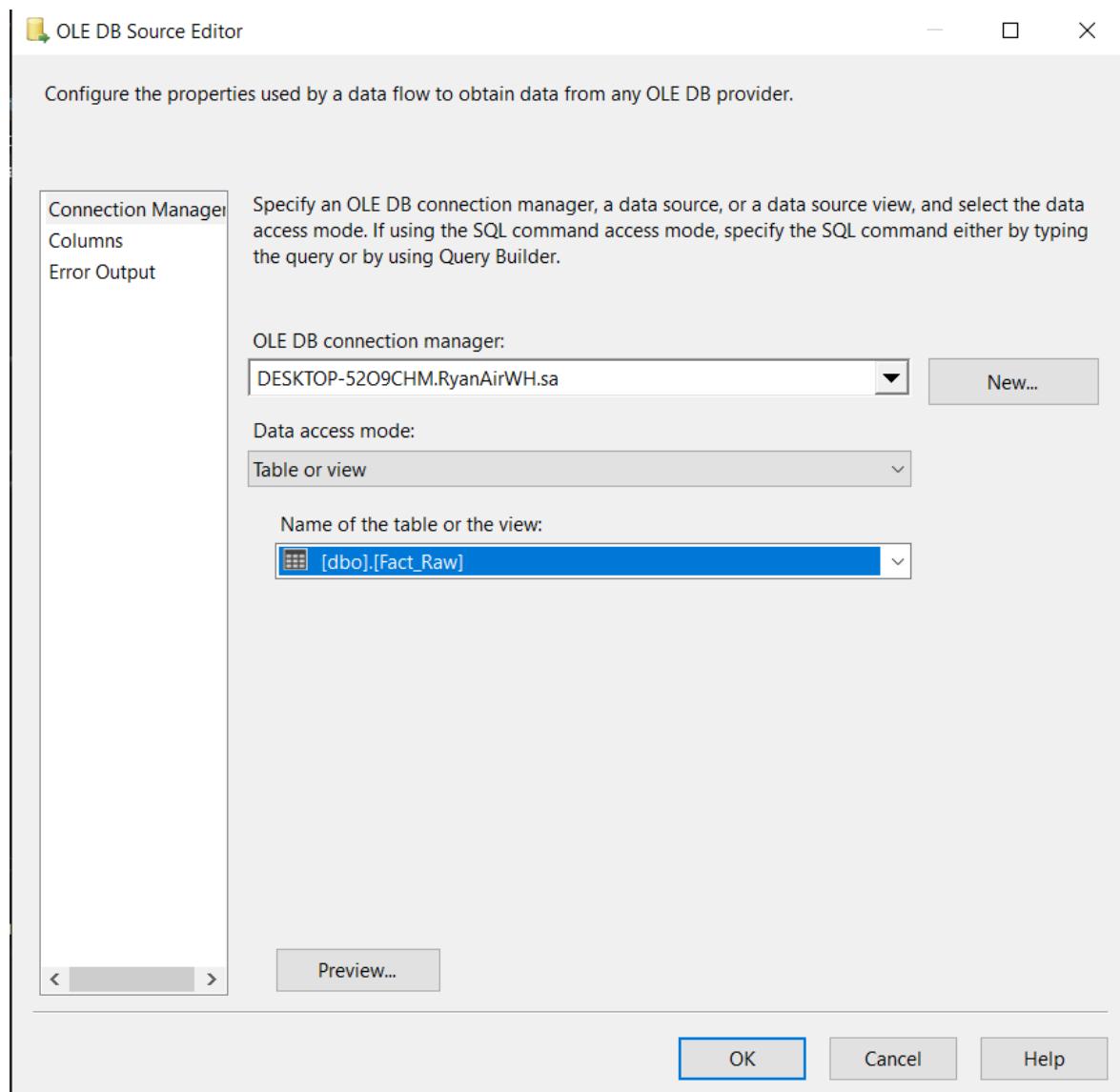
Bước 1. Ở tab Control Flow, tạo **Data Flow Task** và đổi tên Data Flow Task là “Merge Fact_Raw and Dim_Route to Fact1”



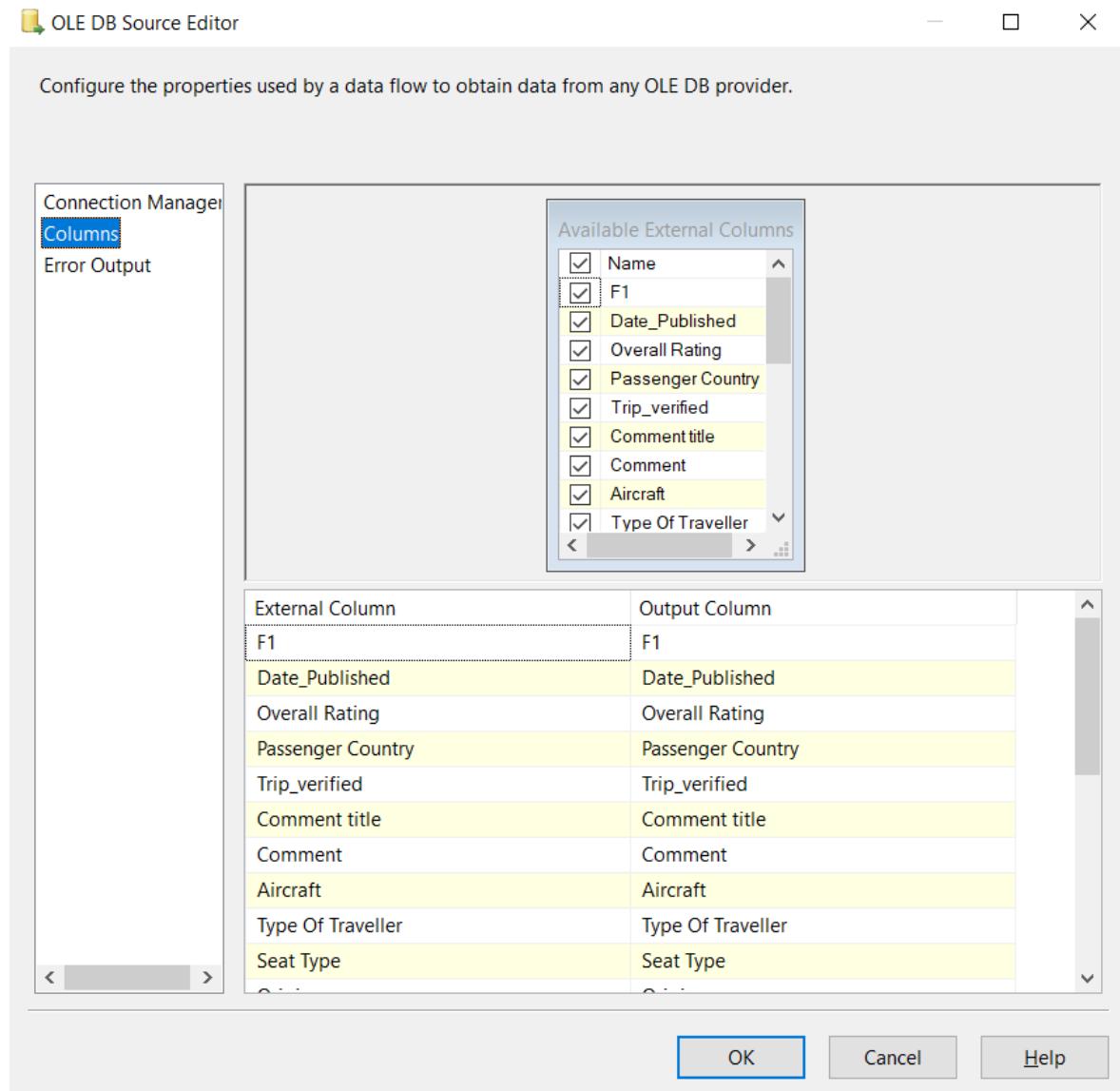
Bước 2. Click chuột phải vào Data Flow Task nói trên và chọn **Edit**, trong tab Data Flow ta tạo 2 OLE DB Source và đổi tên Fact_Raw và Dim_Route



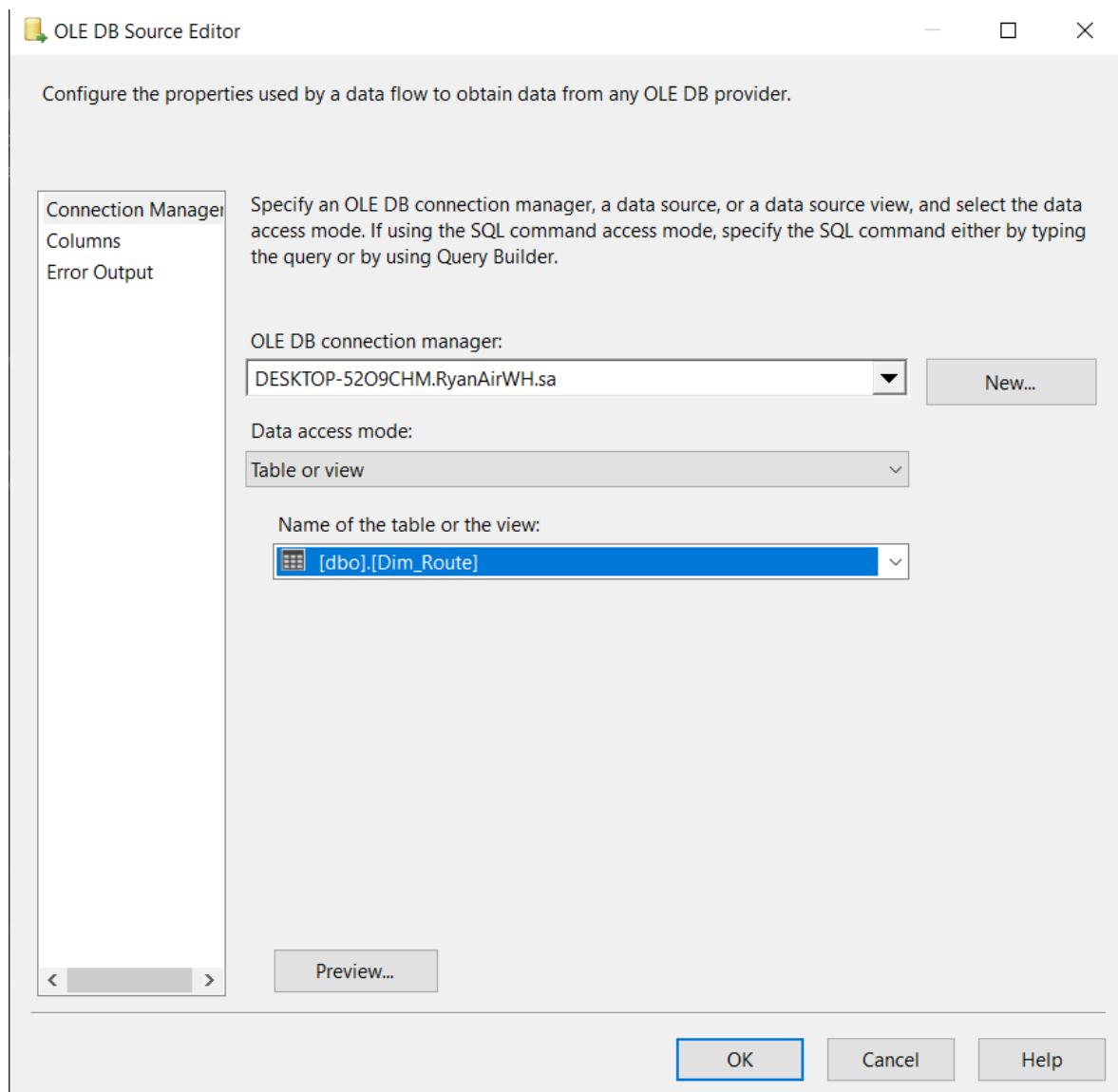
Bước 3. Click chuột phải chọn **Edit**, sau đó chọn bảng Fact_Raw đã tạo trước đó làm data source cho bảng Fact_Raw mới này.

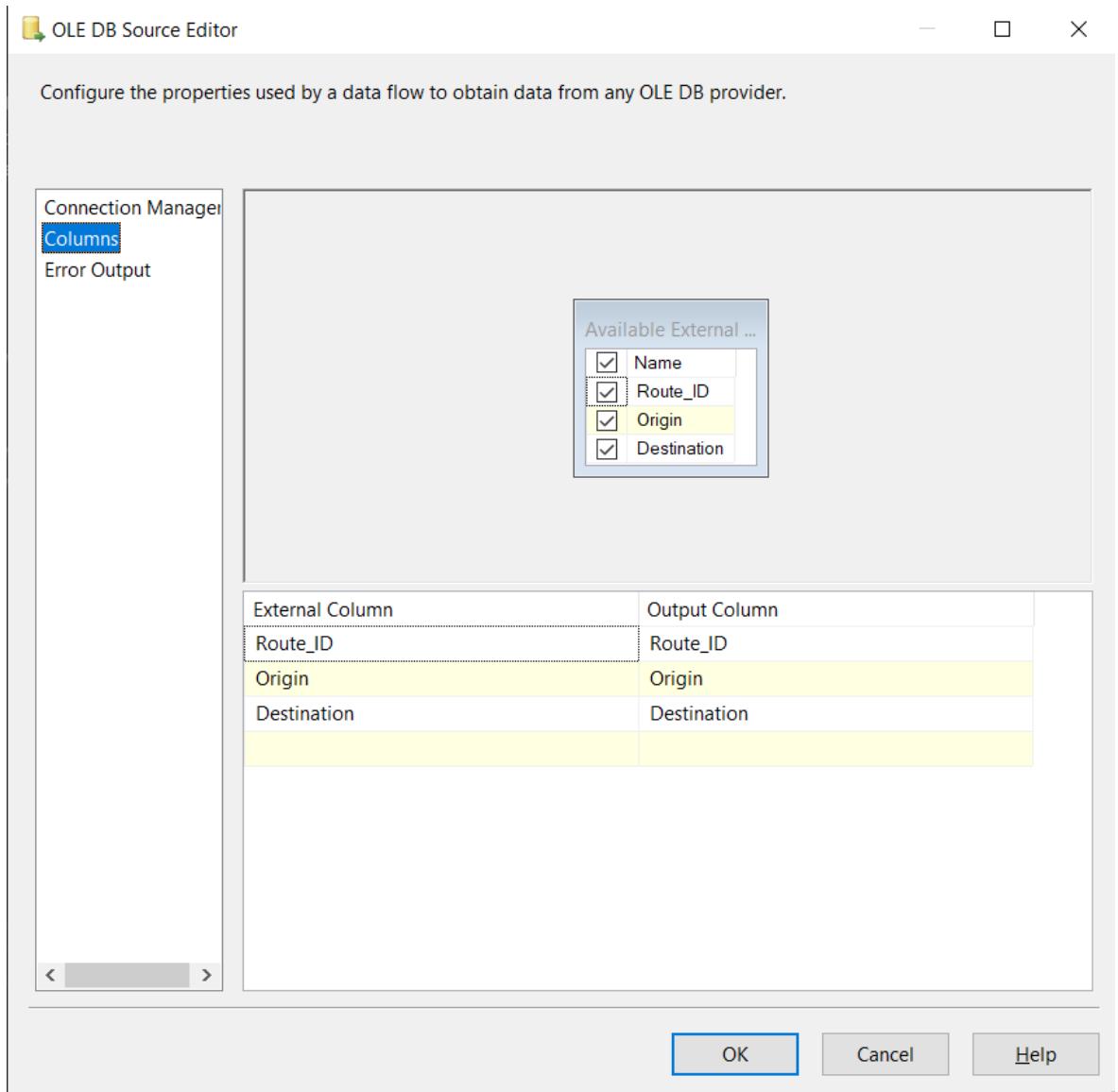


Bước 4. Chọn mục **Columns** để xem xét các cột được ánh xạ. Nhấn **OK**.

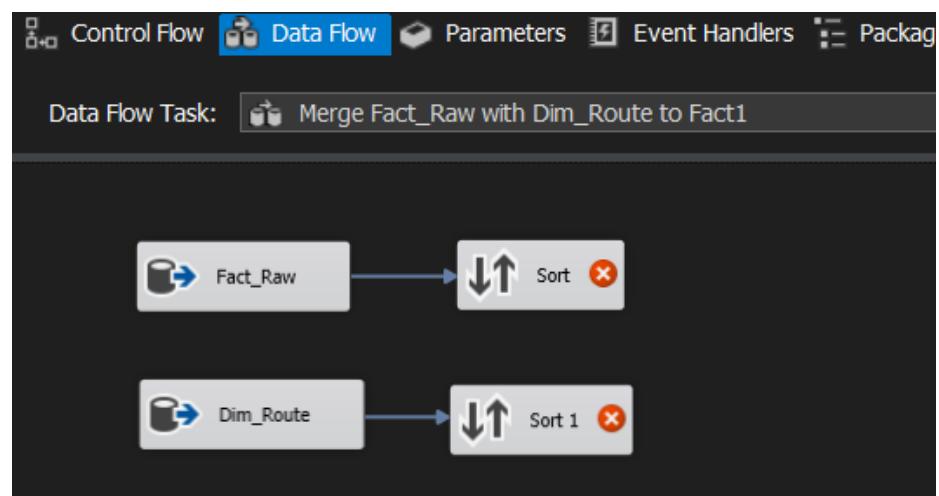


Bước 5. Tương tự thực hiện chọn ánh xạ cột cho Dim_Route

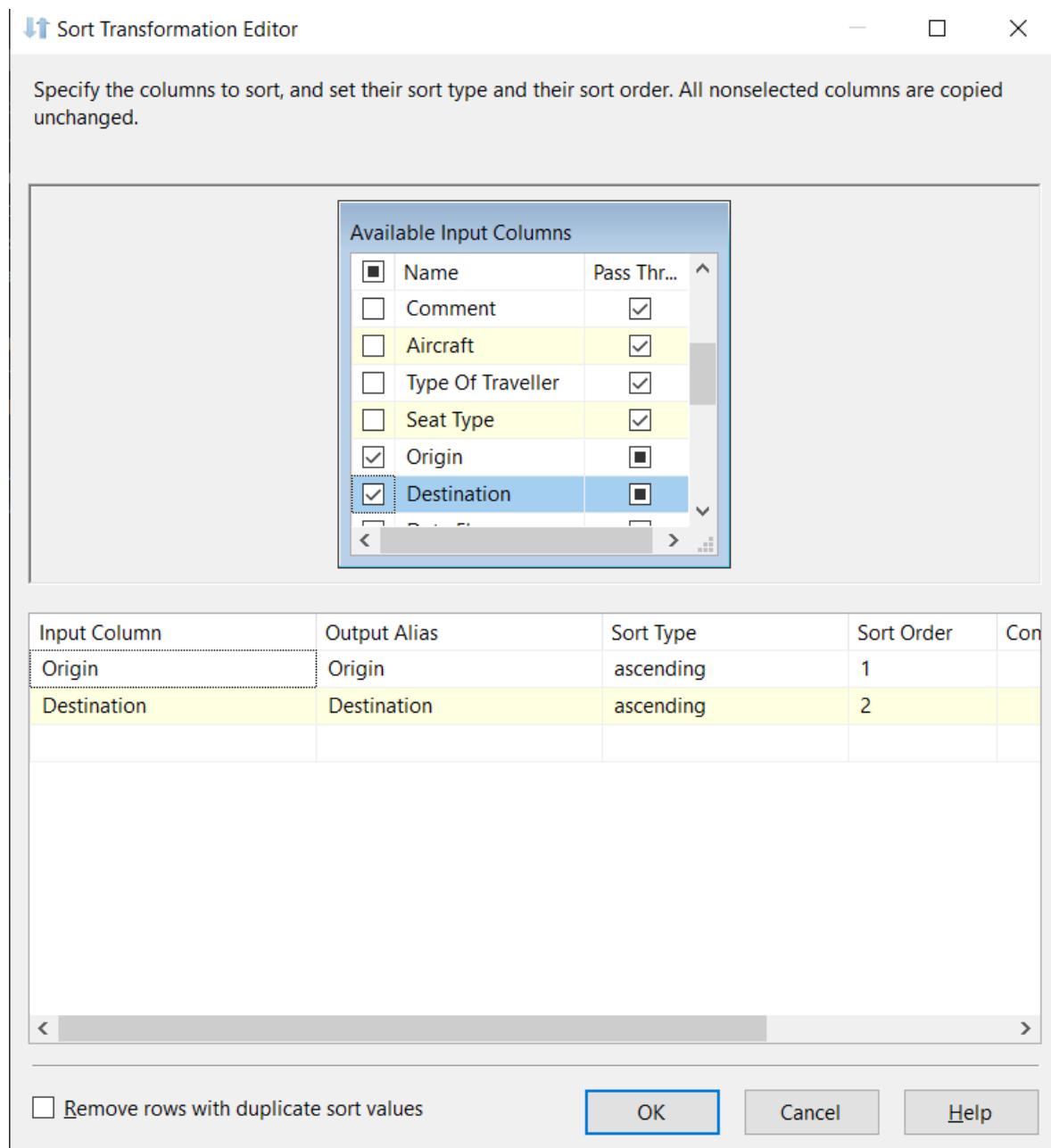




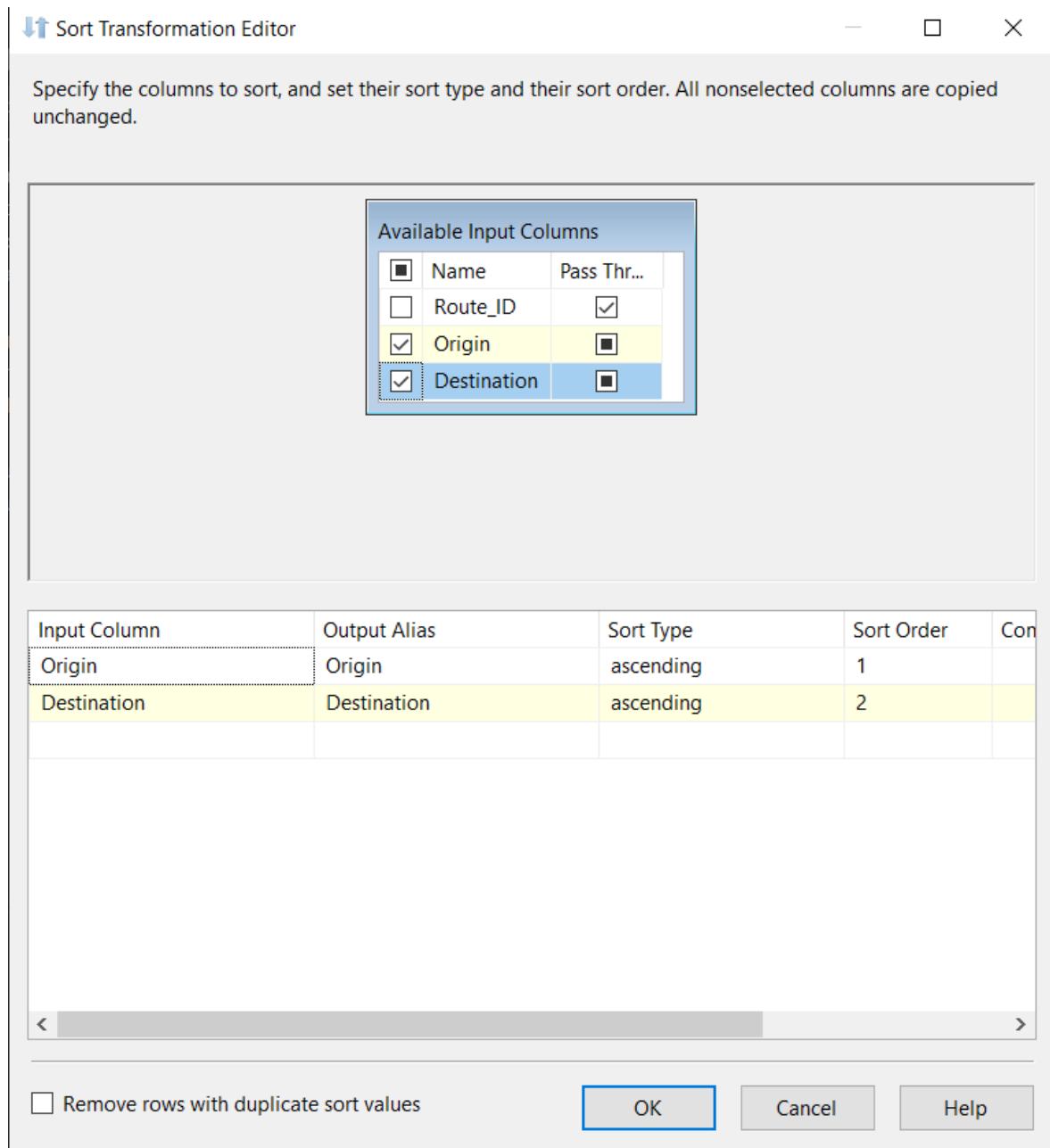
Bước 6. Tạo 2 Sort là Sort và Sort1 tương ứng với mỗi Source.



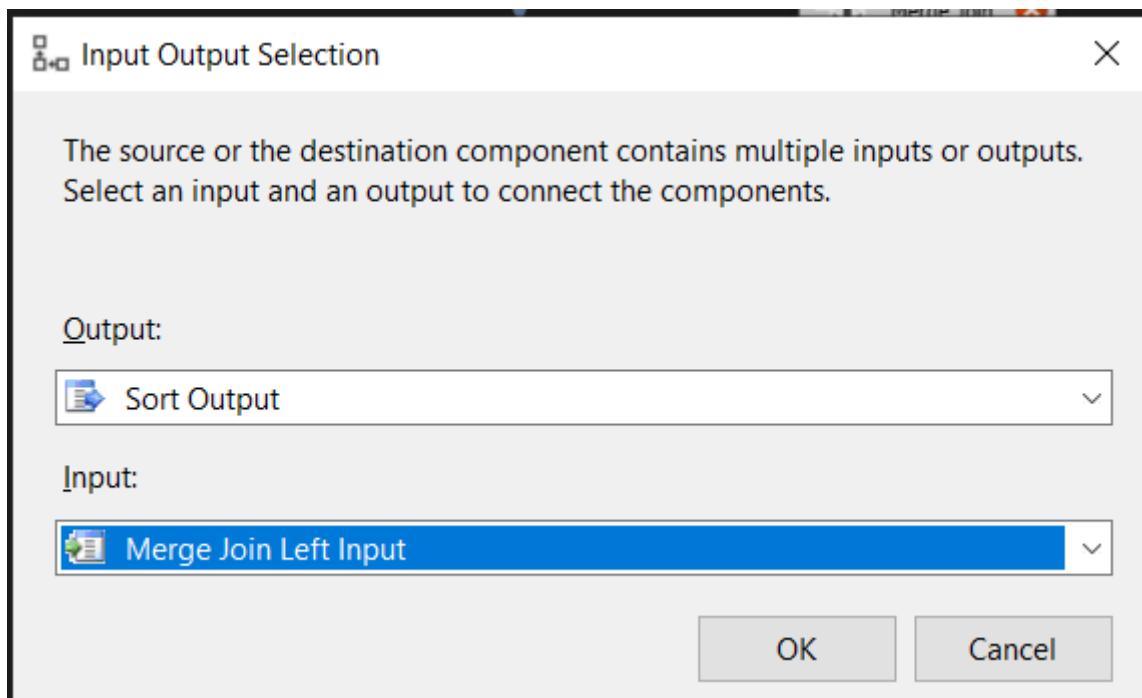
Bước 7. Ở Sort, click chuột phải chọn **Edit** và chọn các cột Origin, Destination theo thứ tự giống với bảng Dim_Route để chuẩn bị cho quá trình merge.



Bước 8. Tương tự ta chọn các cột cho Sort1

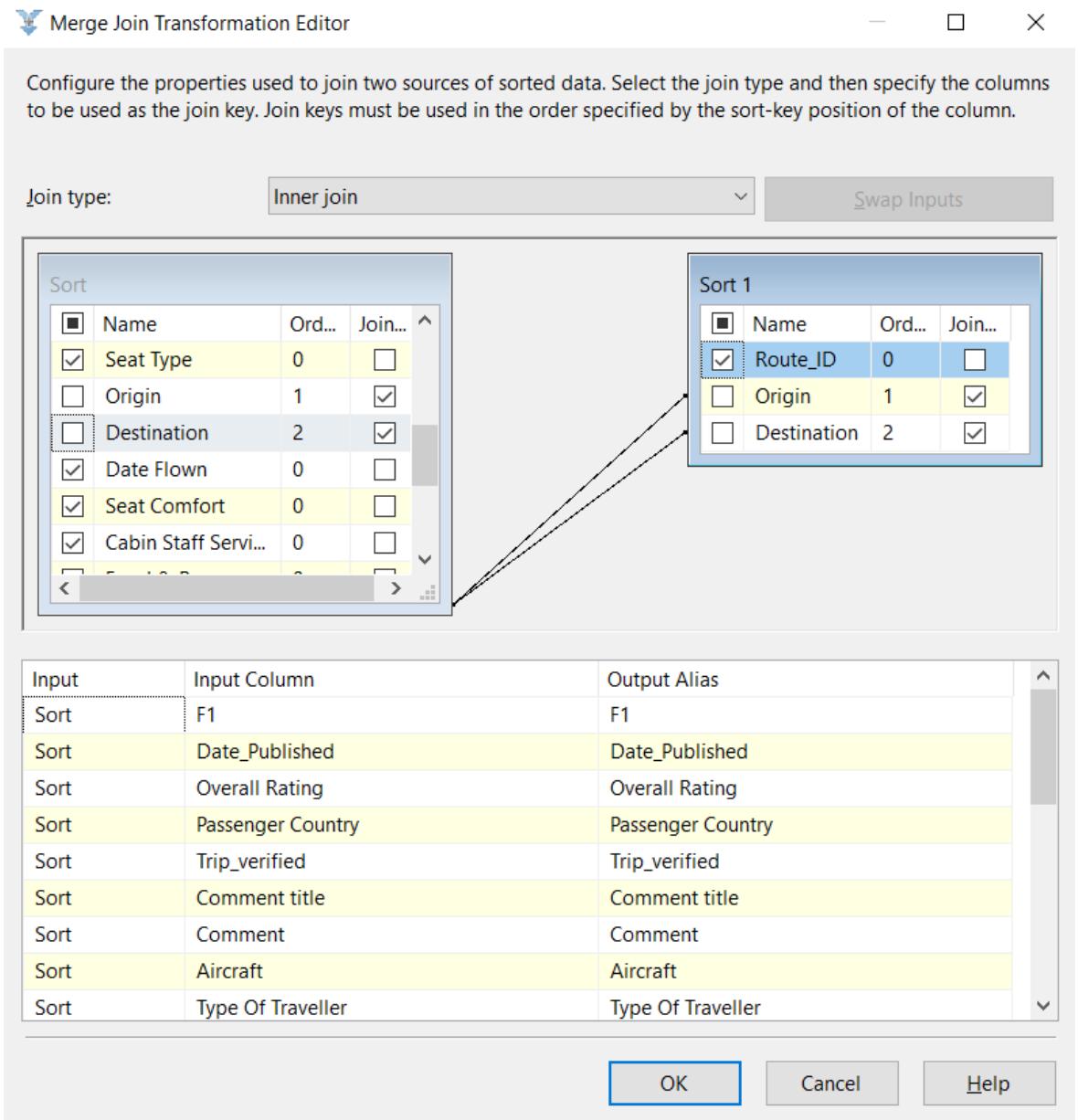


Bước 9. Tạo một **Merge Join** và nối với Sort, tiếp theo ta chọn **Merge Join Left Input** để giữ lại toàn bộ các dòng trong bảng Fact_Raw bất kể có kết quả khi thực hiện phép kết trái với cột ID của bảng Dim_Route hay không.

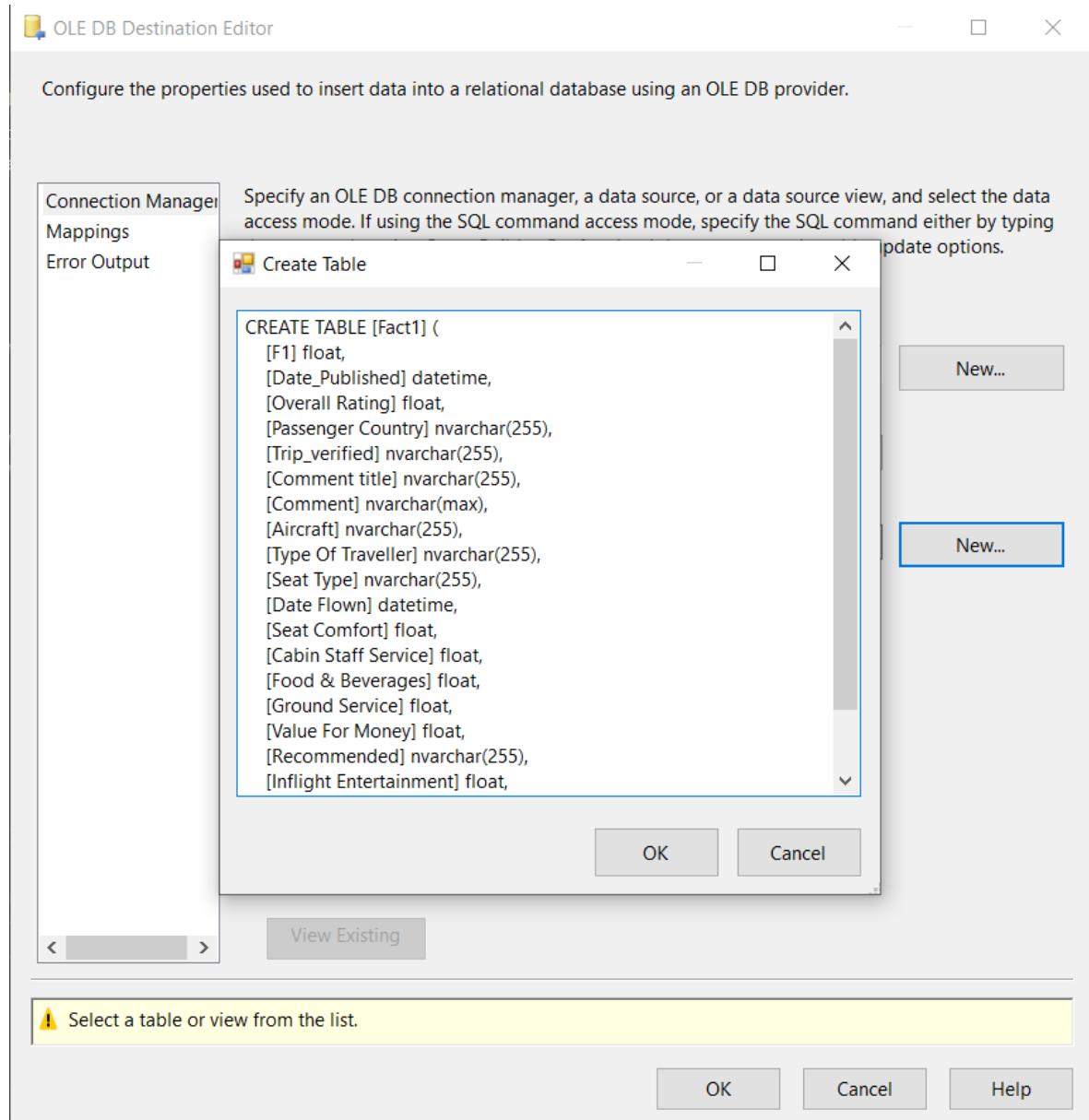


Bước 10.

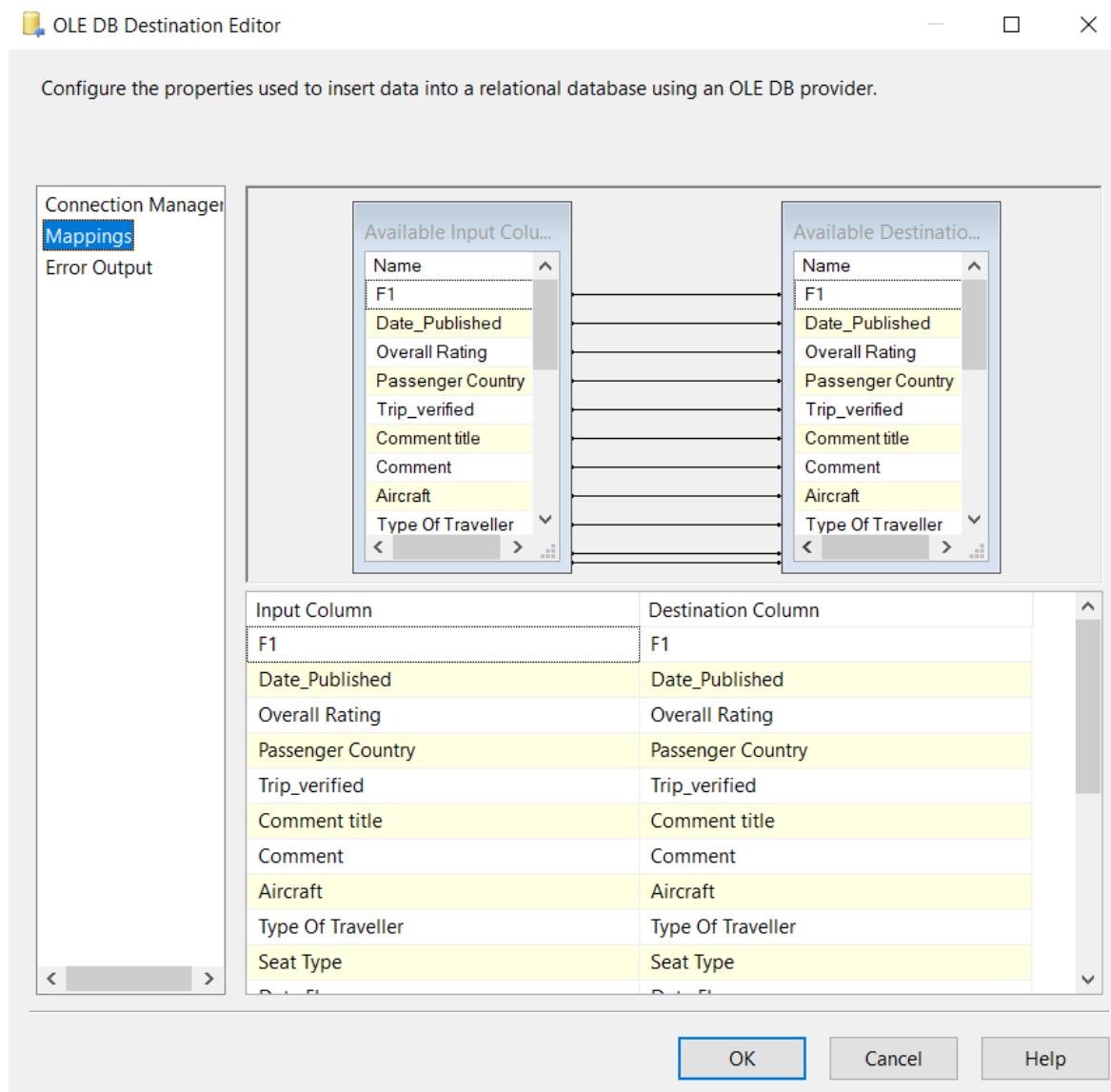
- Chuột phải vào Merge Join và nhấn **Edit**, một hộp thoại merge editor xuất hiện: ở đây ta tick chọn tất cả các cột của Sort nhưng không lấy thuộc tính Origin, Destination
- Tiếp theo ta chọn Route_ID ở Sort1 để merge vào Fact_Raw
- Kết quả sau khi merge là bảng Fact_Raw không còn thuộc tính Origin, Destination và có thêm 1 thuộc tính mới là Route_ID



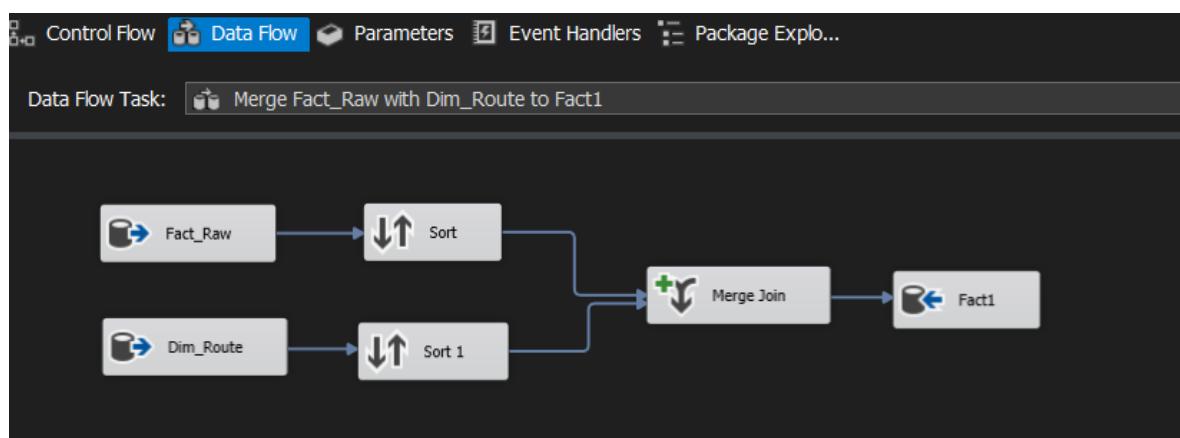
Bước 11. Tạo bảng Fact1 từ một **OLE DB Destination** để chứa tất cả những gì đã merge



Chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

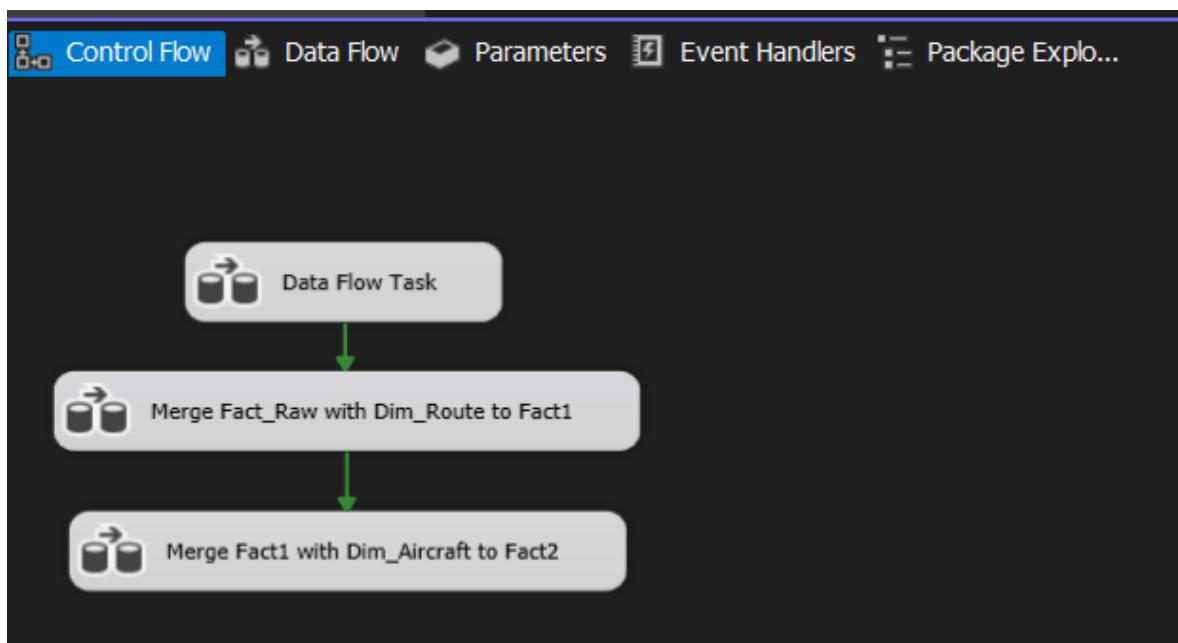


Kết quả:

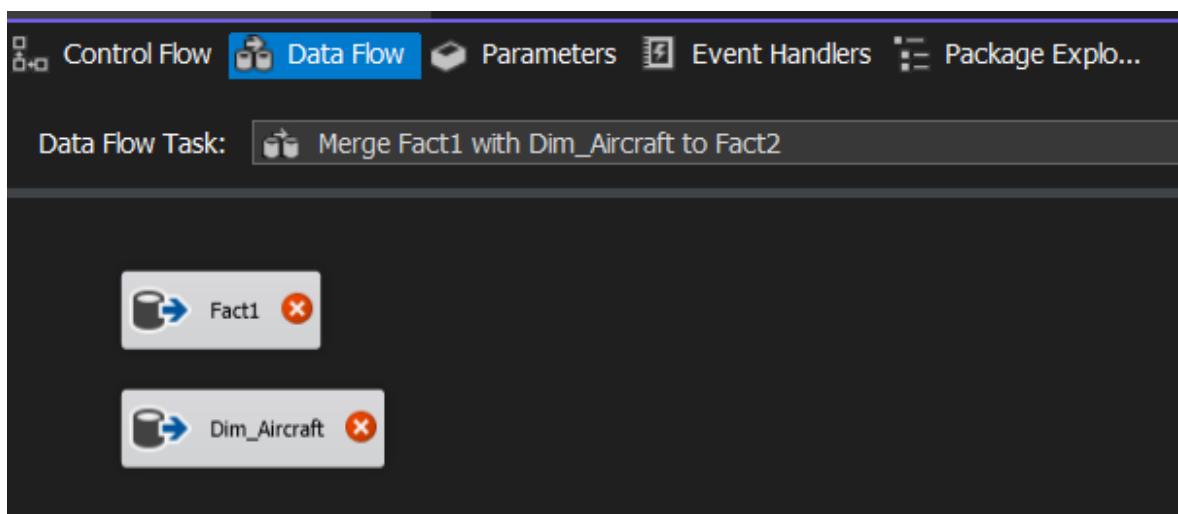


3.3.8.2. Merge Fact1 và Dim_Aircraft vào Fact2

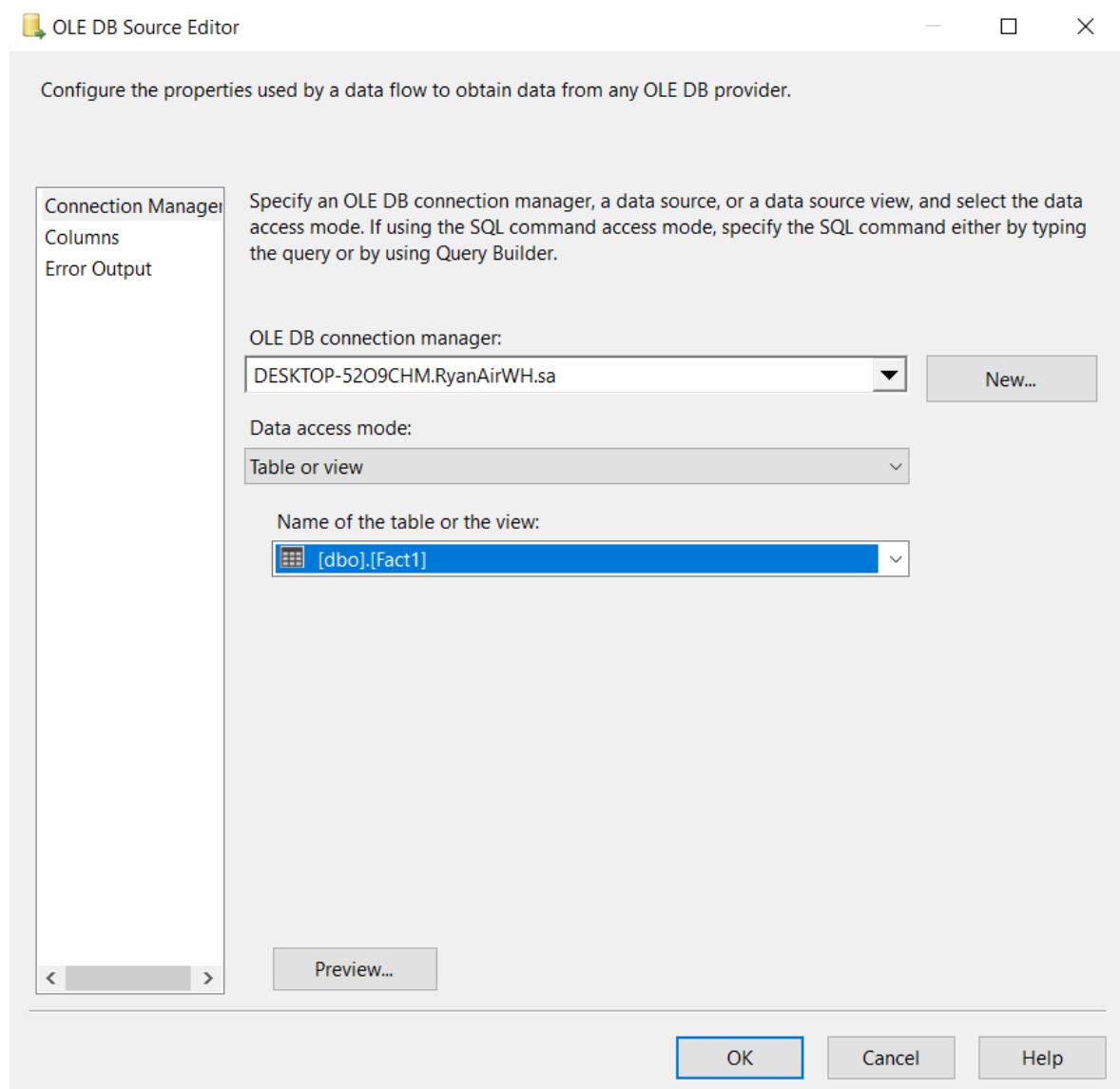
Bước 1. Ở tab Control Flow, tạo **Data Flow Task** và đổi tên Data Flow Task là “Merge Fact1 with Dim_Aircraft to Fact2”



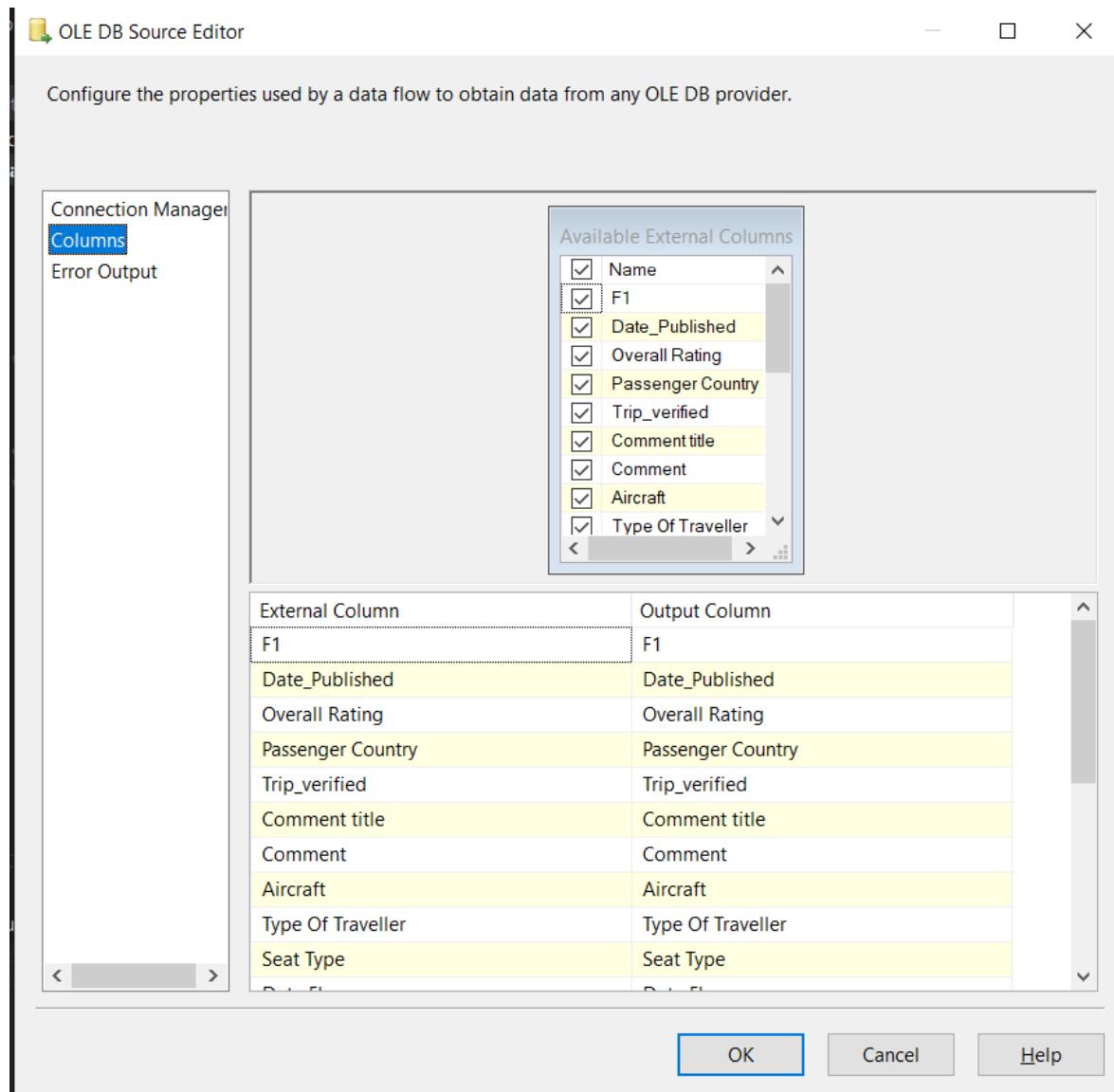
Bước 2. Click chuột phải vào Data Flow Task nói trên và chọn **Edit**, trong tab Data Flow ta tạo 2 OLE DB Source và đổi tên Fact1 và Dim_Aircraft



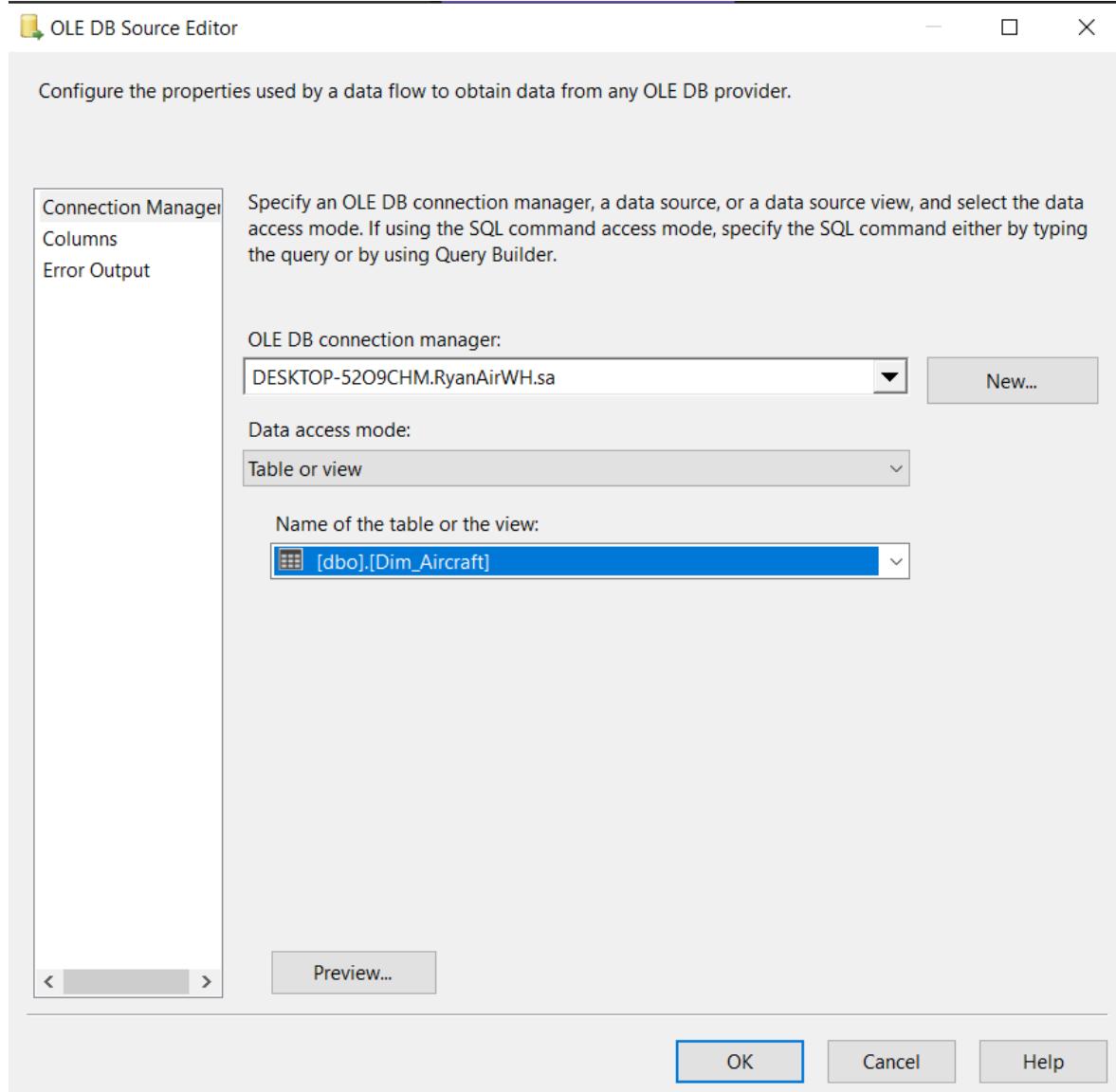
Bước 3. Click chuột phải chọn **Edit**, sau đó chọn bảng Fact1 đã tạo trước đó làm data source cho bảng Fact1 mới này.

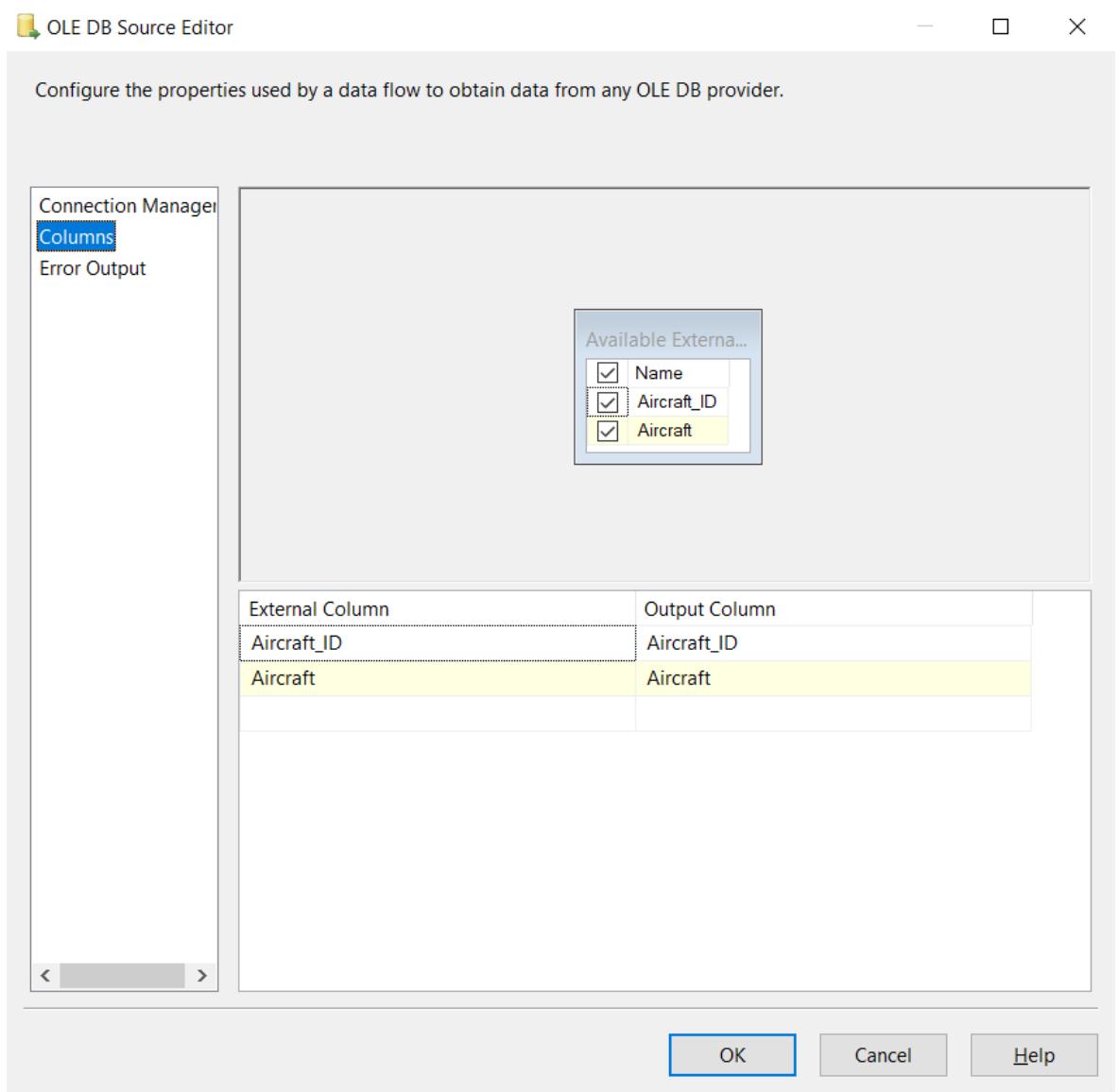


Bước 4. Chọn mục **Columns** để xem xét các cột được ánh xạ. Nhấn **OK**.

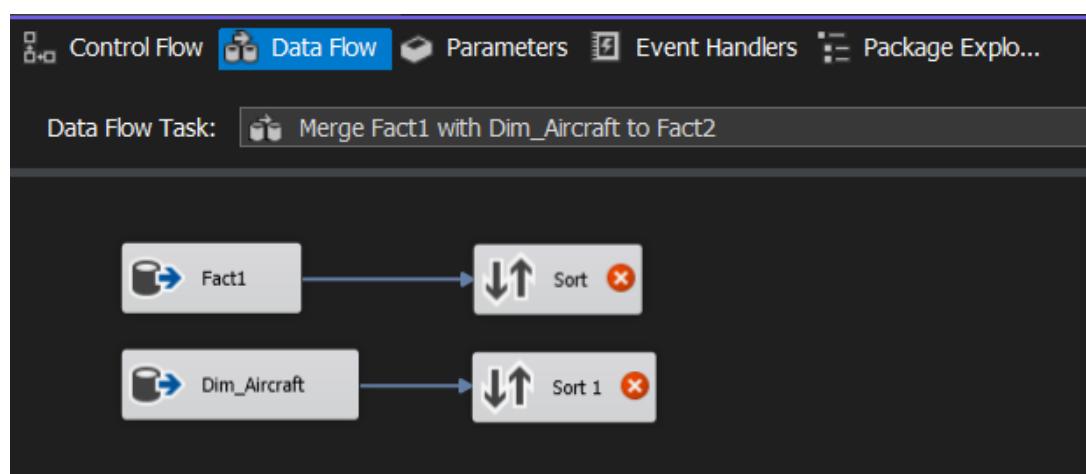


Bước 5. Tương tự thực hiện chọn ánh xạ cột cho Dim_Aircraft

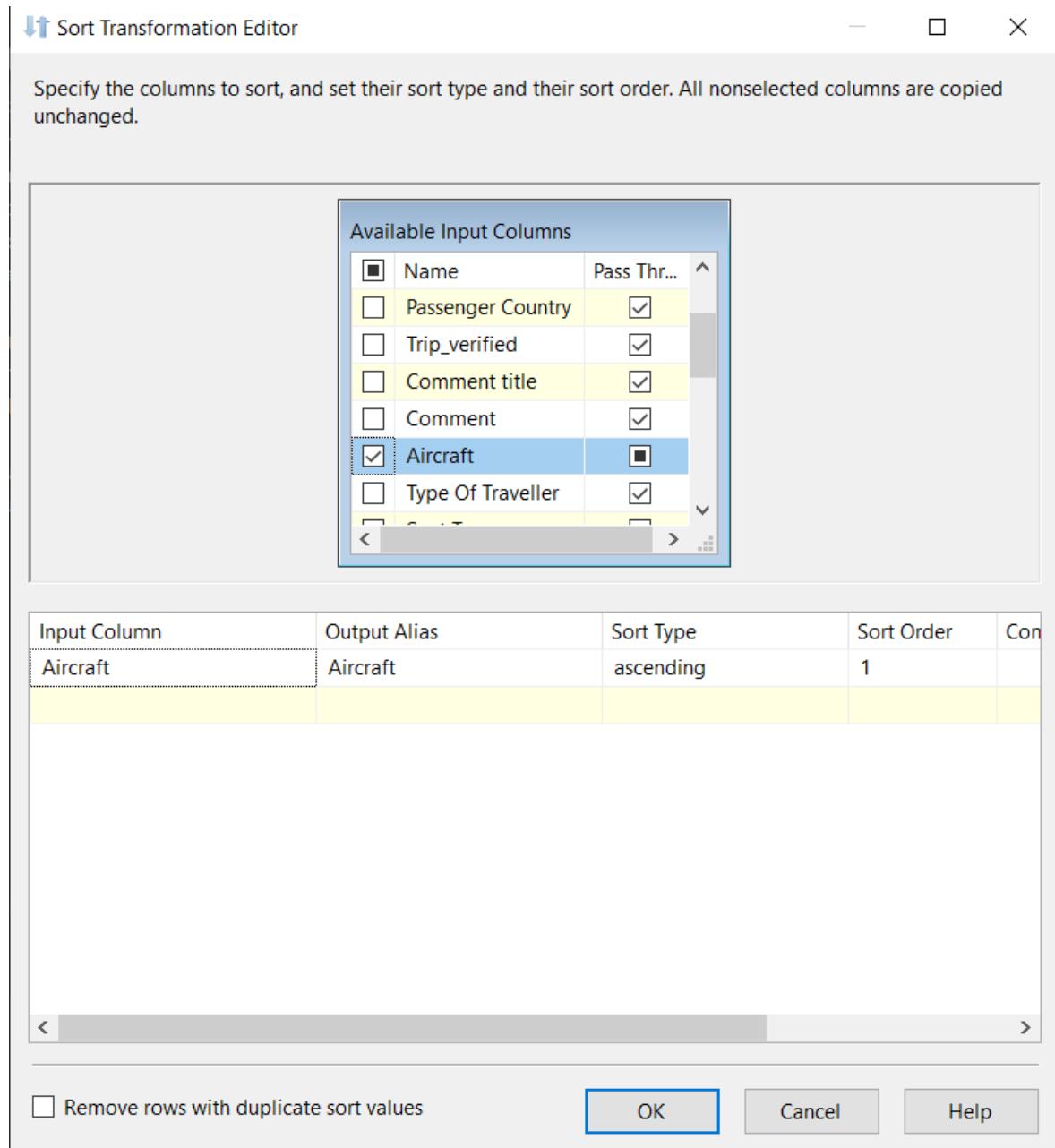




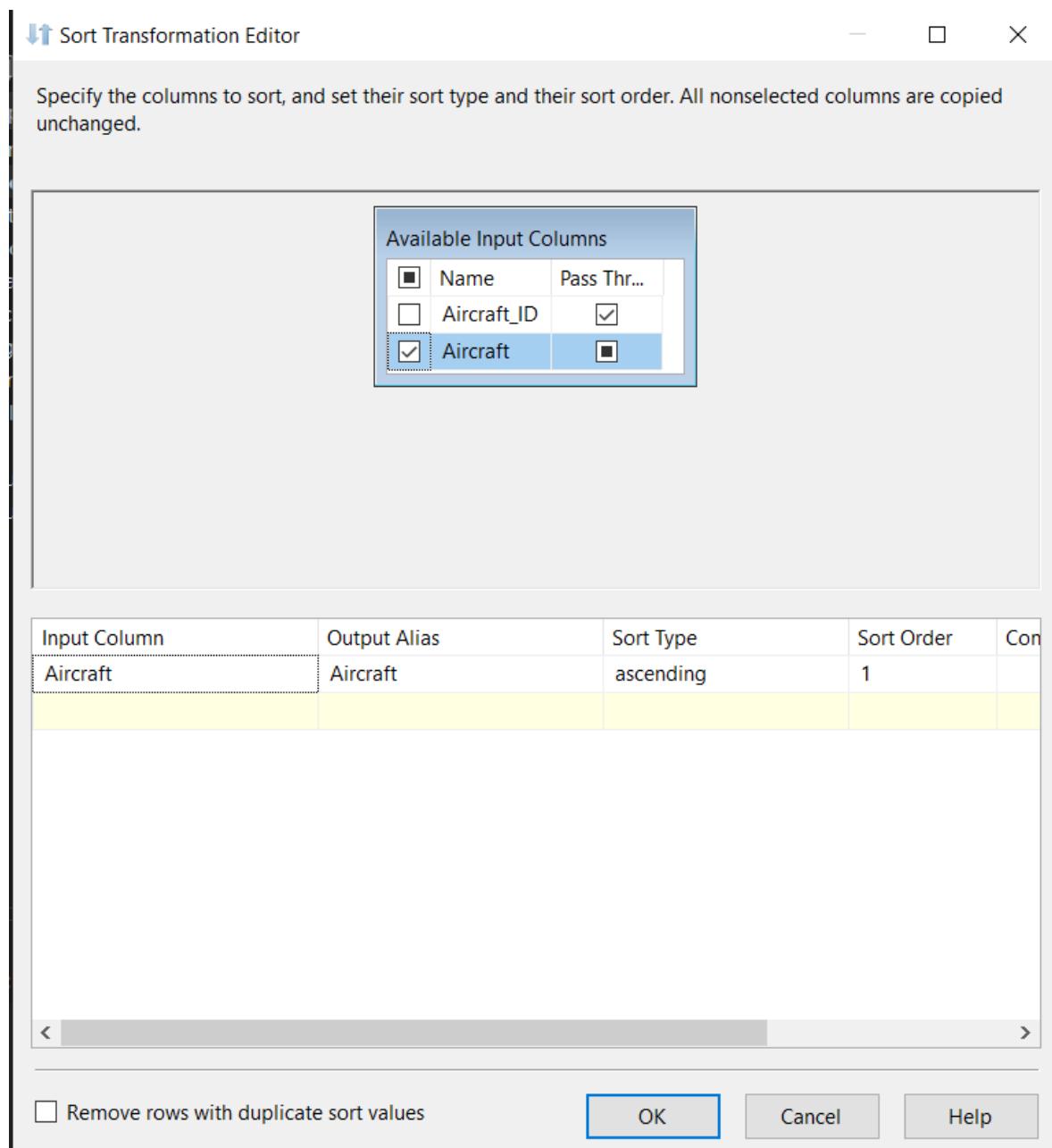
Bước 6. Tạo 2 Sort là Sort và Sort1 tương ứng với mỗi Source.



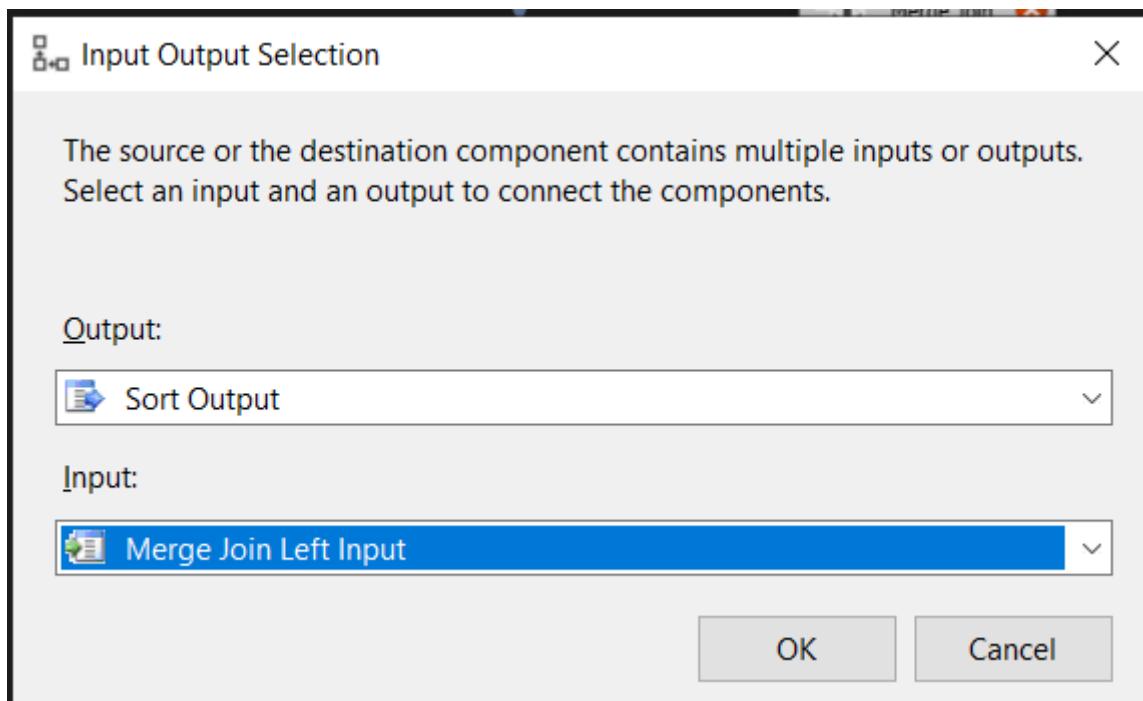
Bước 7. Ở Sort, click chuột phải chọn **Edit** và chọn cột Aircraft theo thứ tự giống với bảng Dim_Aircraft để chuẩn bị cho quá trình merge.



Bước 8. Tương tự ta chọn các cột cho Sort1

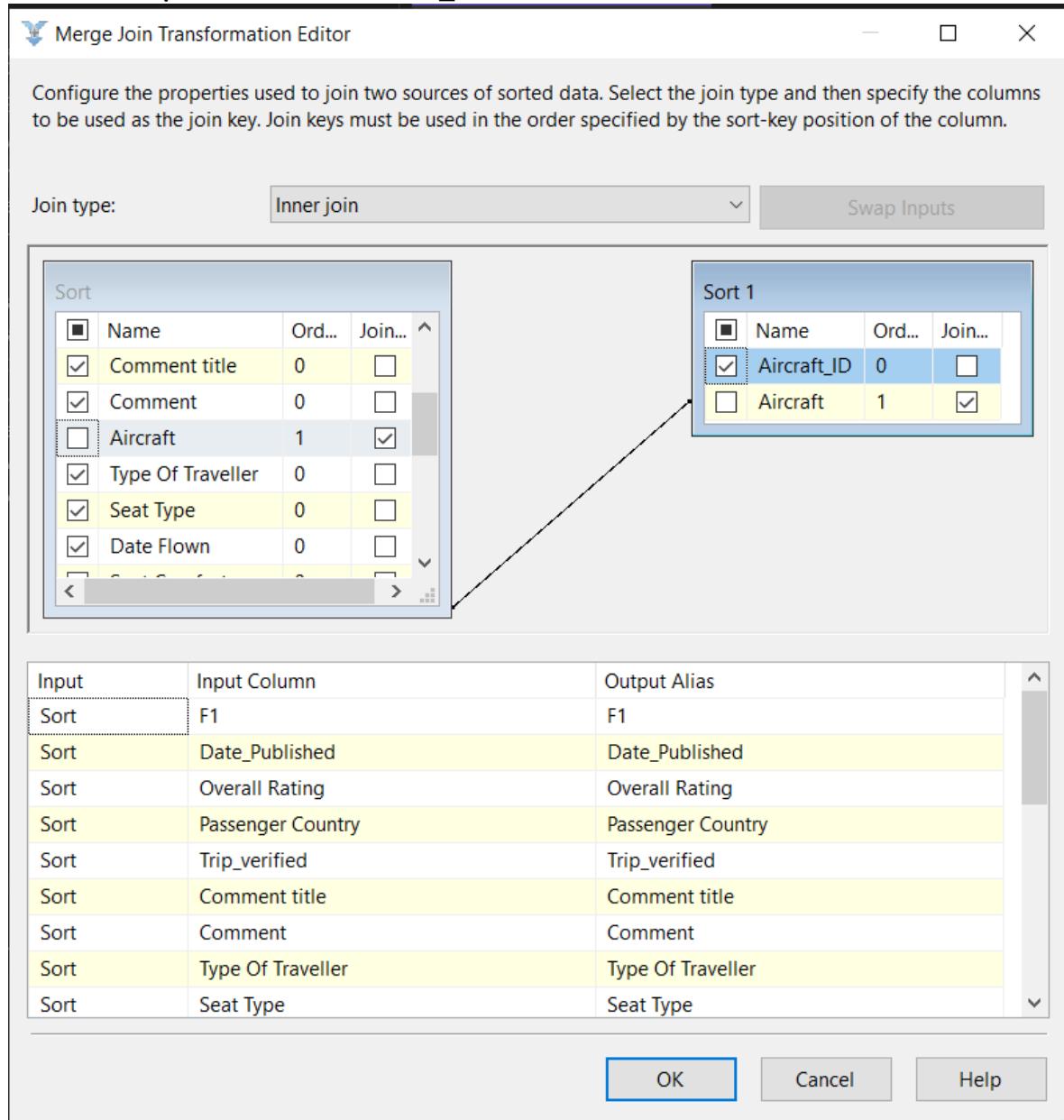


Bước 9. Tạo một **Merge Join** và nối với Sort, tiếp theo ta chọn **Merge Join Left Input** để giữ lại toàn bộ các dòng trong bảng Fact1 bất kể có kết quả khi thực hiện phép kết trái với cột ID của bảng Dim_Aircraft hay không.

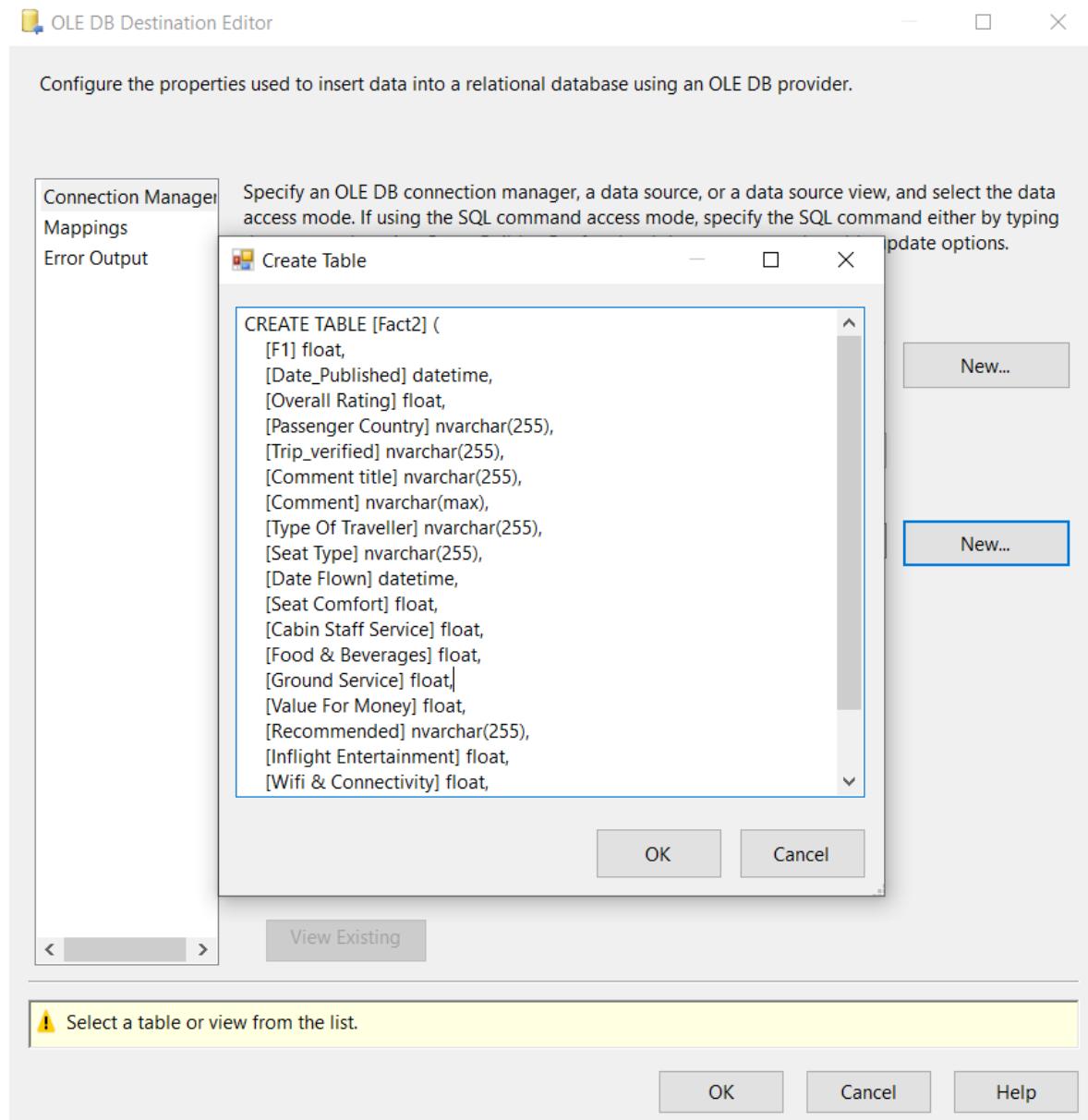


Bước 10.

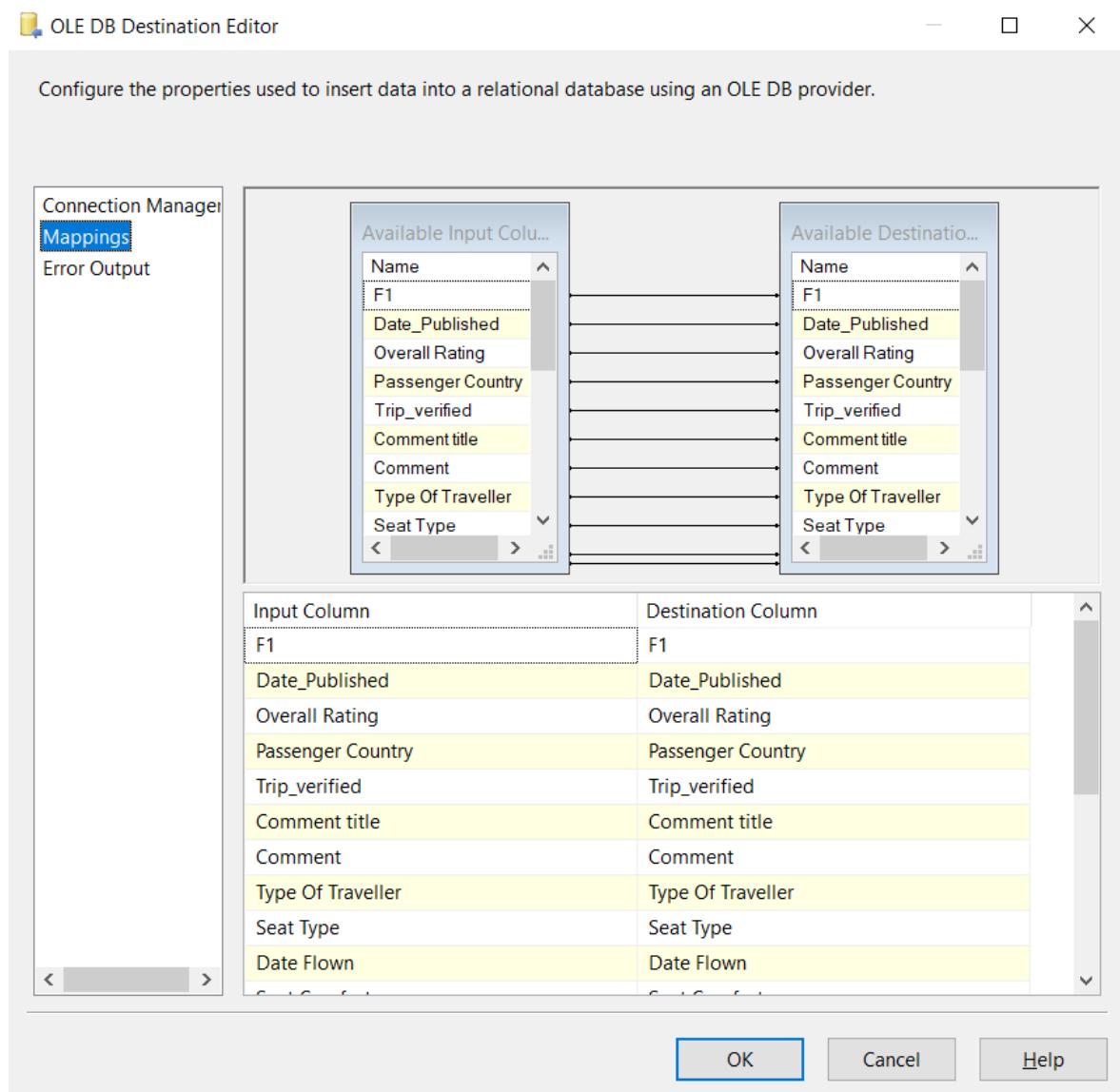
- Chuột phải vào Merge Join và nhấn **Edit**, một hộp thoại merge editor xuất hiện: ở đây ta tick chọn tất cả các cột của Sort nhưng không lấy thuộc tính Aircraft
- Tiếp theo ta chọn Aircraft_ID ở Sort1 để merge vào Fact1
- Kết quả sau khi merge là bảng Fact1 không còn thuộc tính Aircraft và có thêm 1 thuộc tính mới là Aircraft_ID



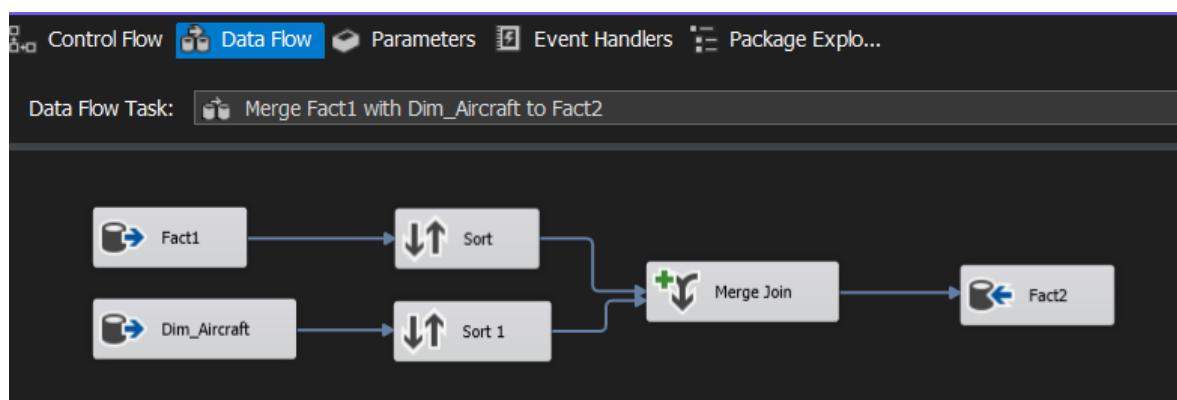
Bước 11. Tạo bảng Fact2 từ một **OLE DB Destination** để chứa tất cả những gì đã merge



Chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

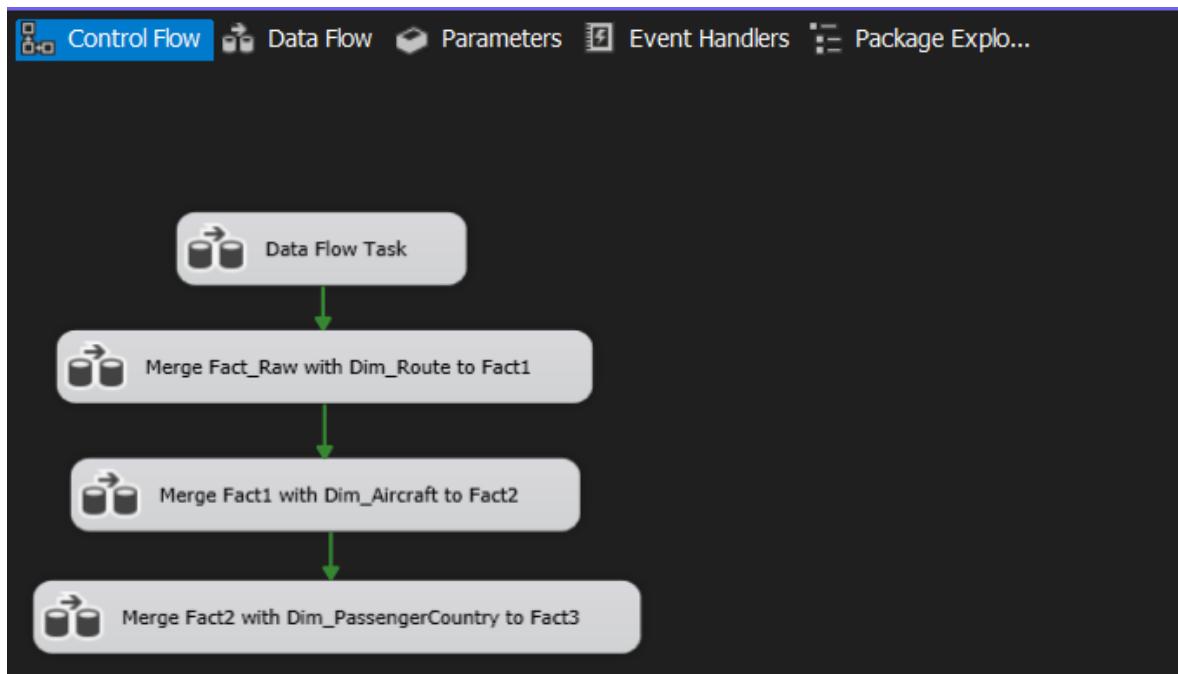


Kết quả:

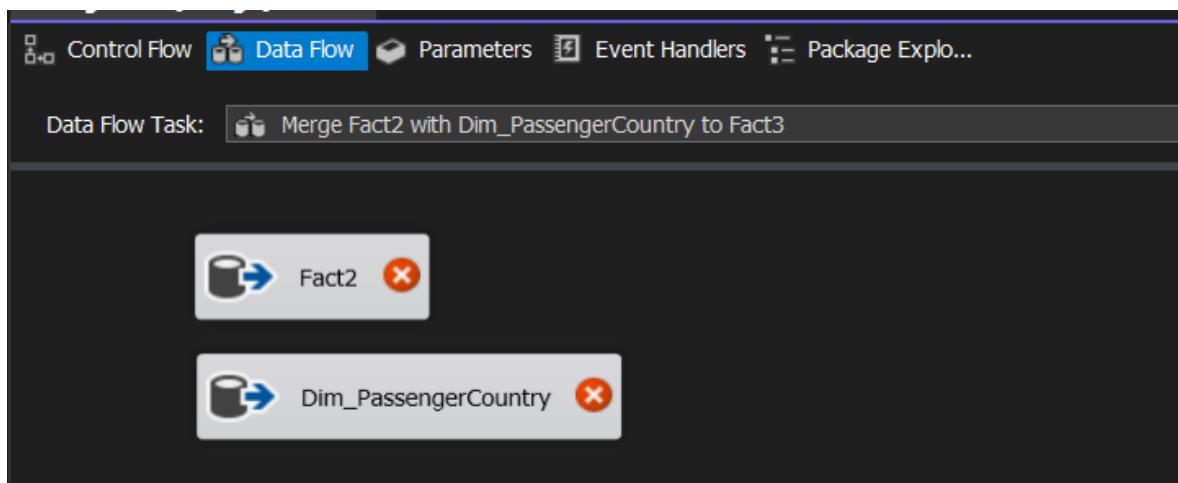


3.3.8.3. Merge Fact2 và Dim_PassengerCountry vào Fact3

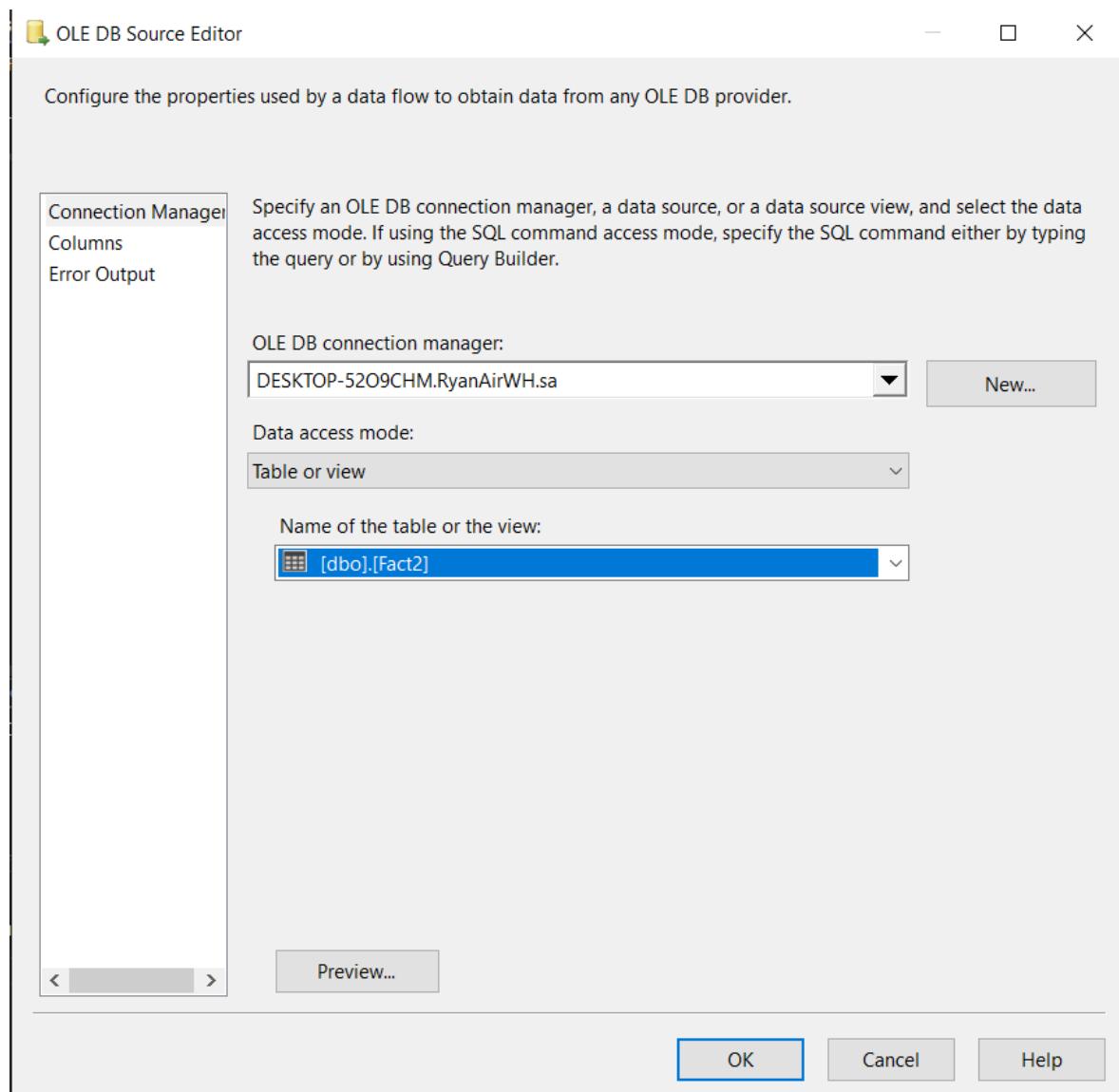
Bước 1. Ở tab Control Flow, tạo **Data Flow Task** và đổi tên Data Flow Task là “Merge Fact2 with Dim_PassengerCountry to Fact3”



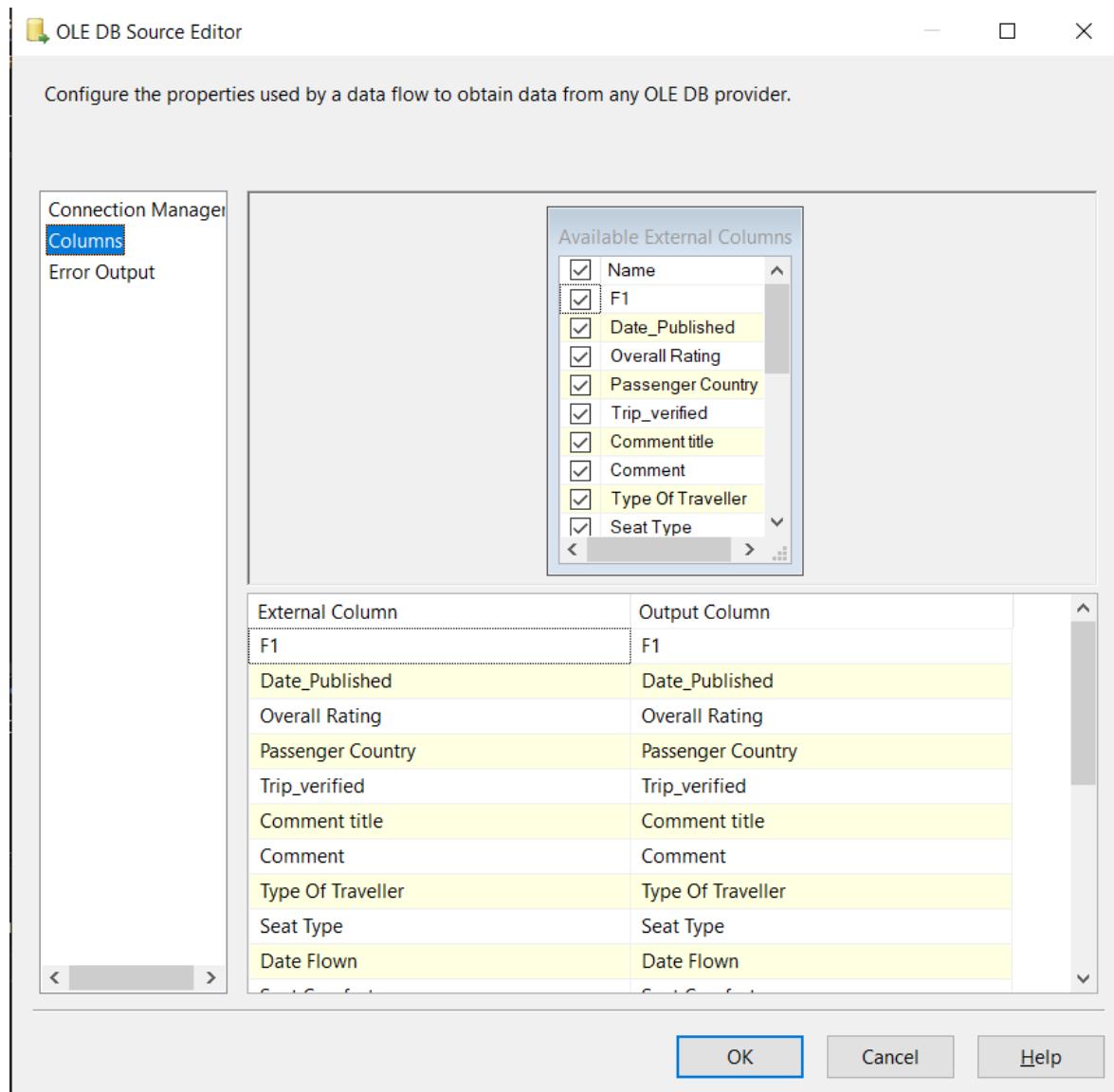
Bước 2. Click chuột phải vào Data Flow Task nói trên và chọn **Edit**, trong tab Data Flow ta tạo 2 OLE DB Source và đổi tên Fact2 và Dim_PassengerCountry



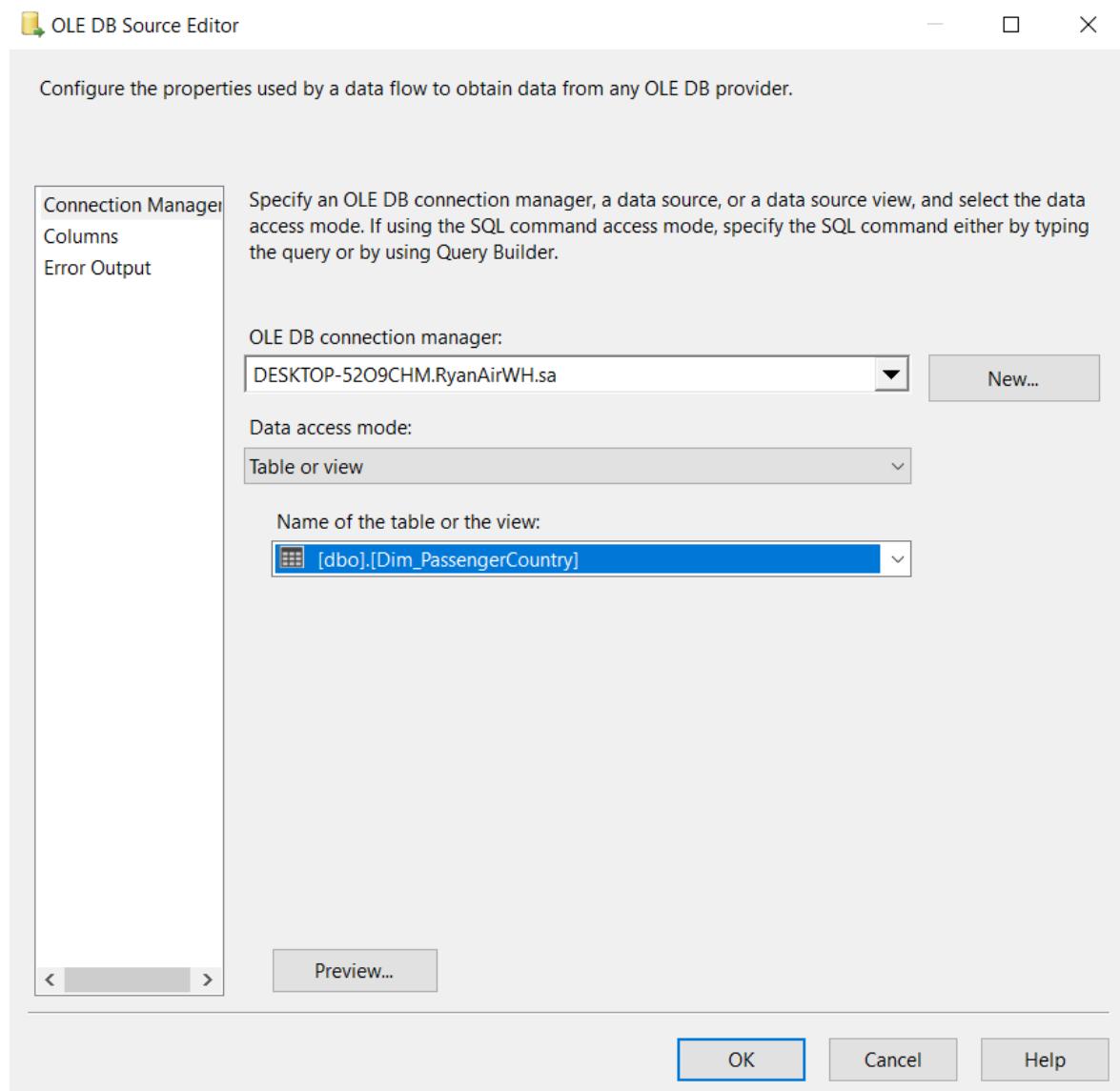
Bước 3. Click chuột phải chọn **Edit**, sau đó chọn bảng Fact2 đã tạo trước đó làm data source cho bảng Fact2 mới này.

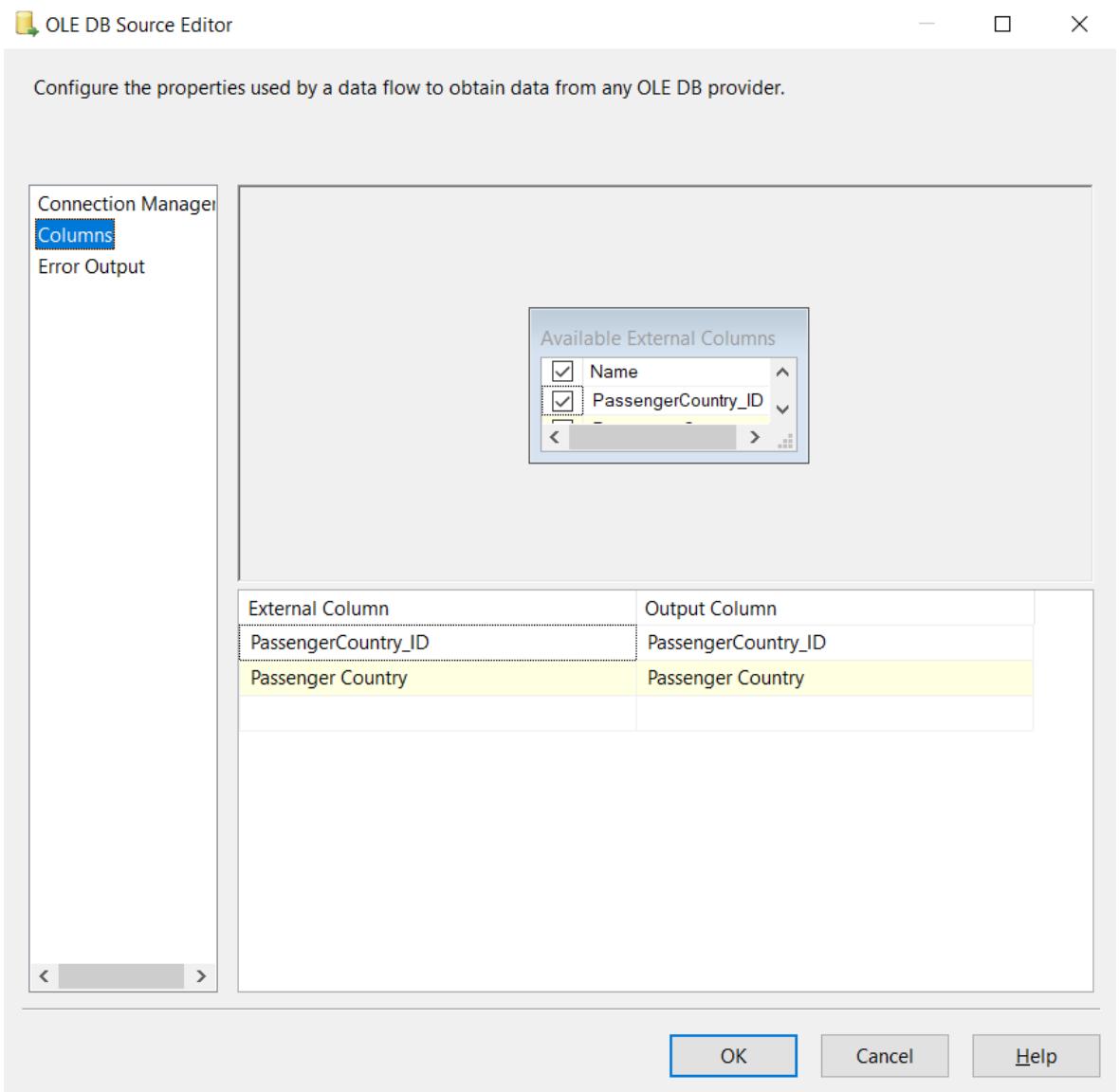


Bước 4. Chọn mục **Columns** để xem xét các cột được ánh xạ. Nhấn **OK**.

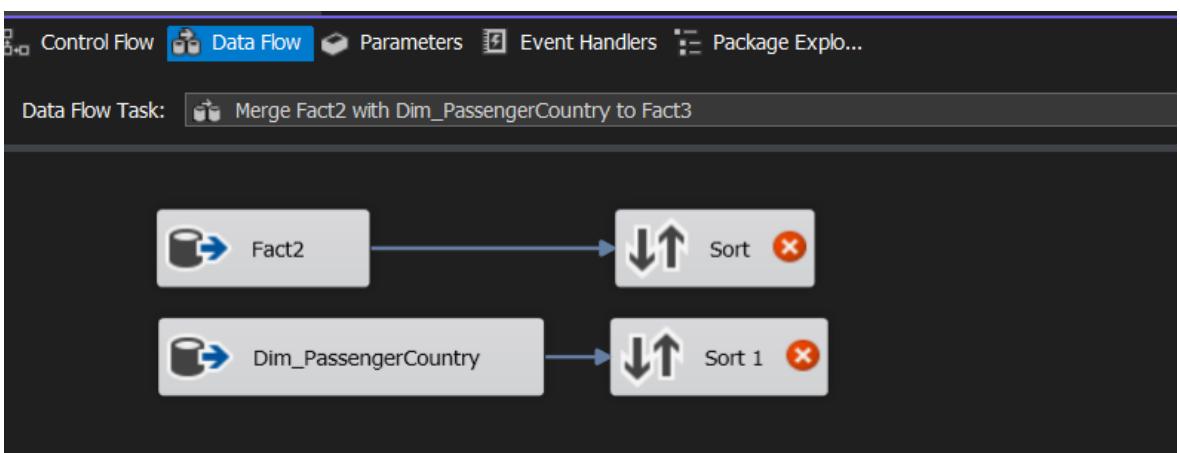


Bước 5. Tương tự thực hiện chọn ánh xạ cột cho Dim_PassengerCountry

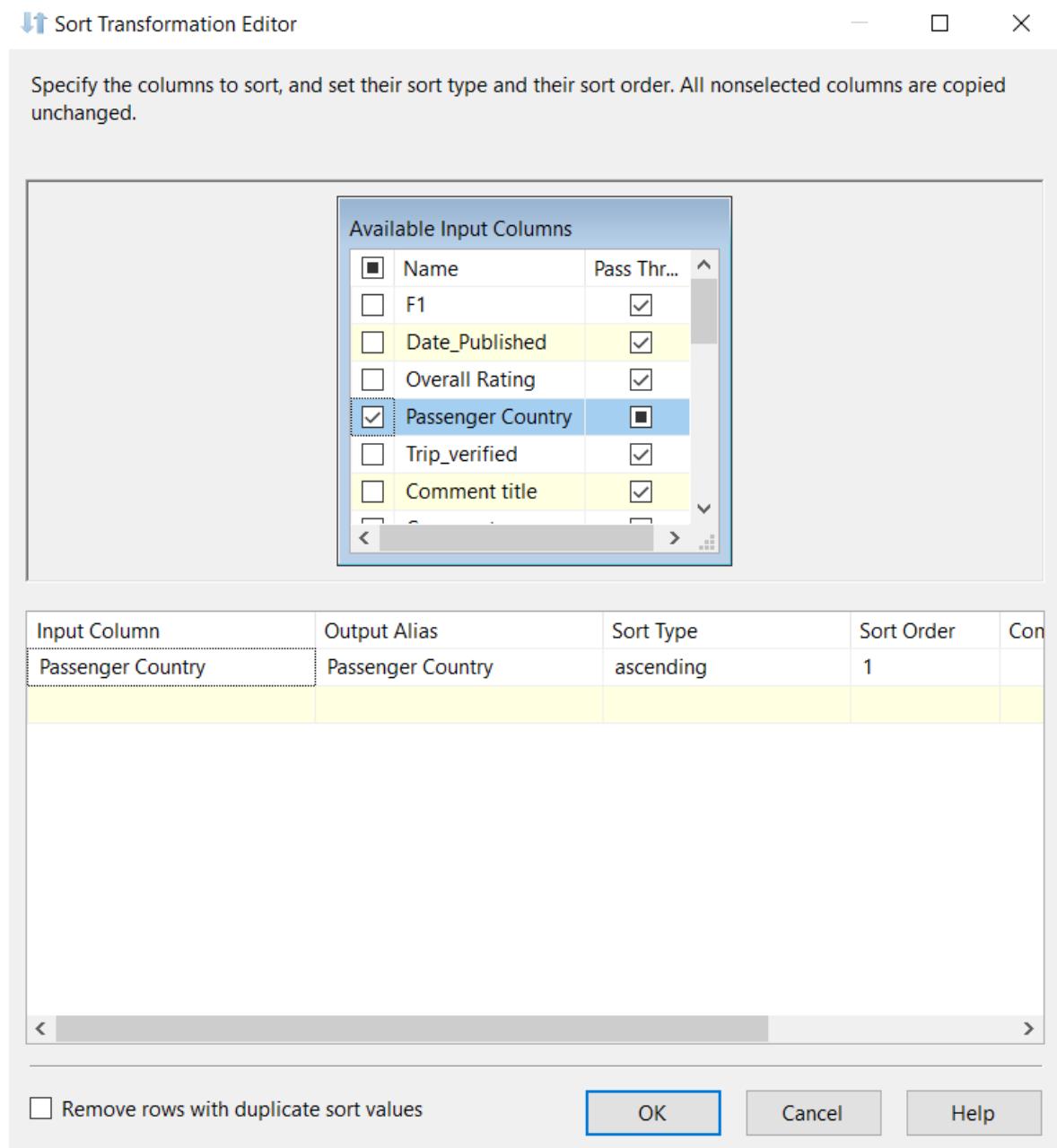




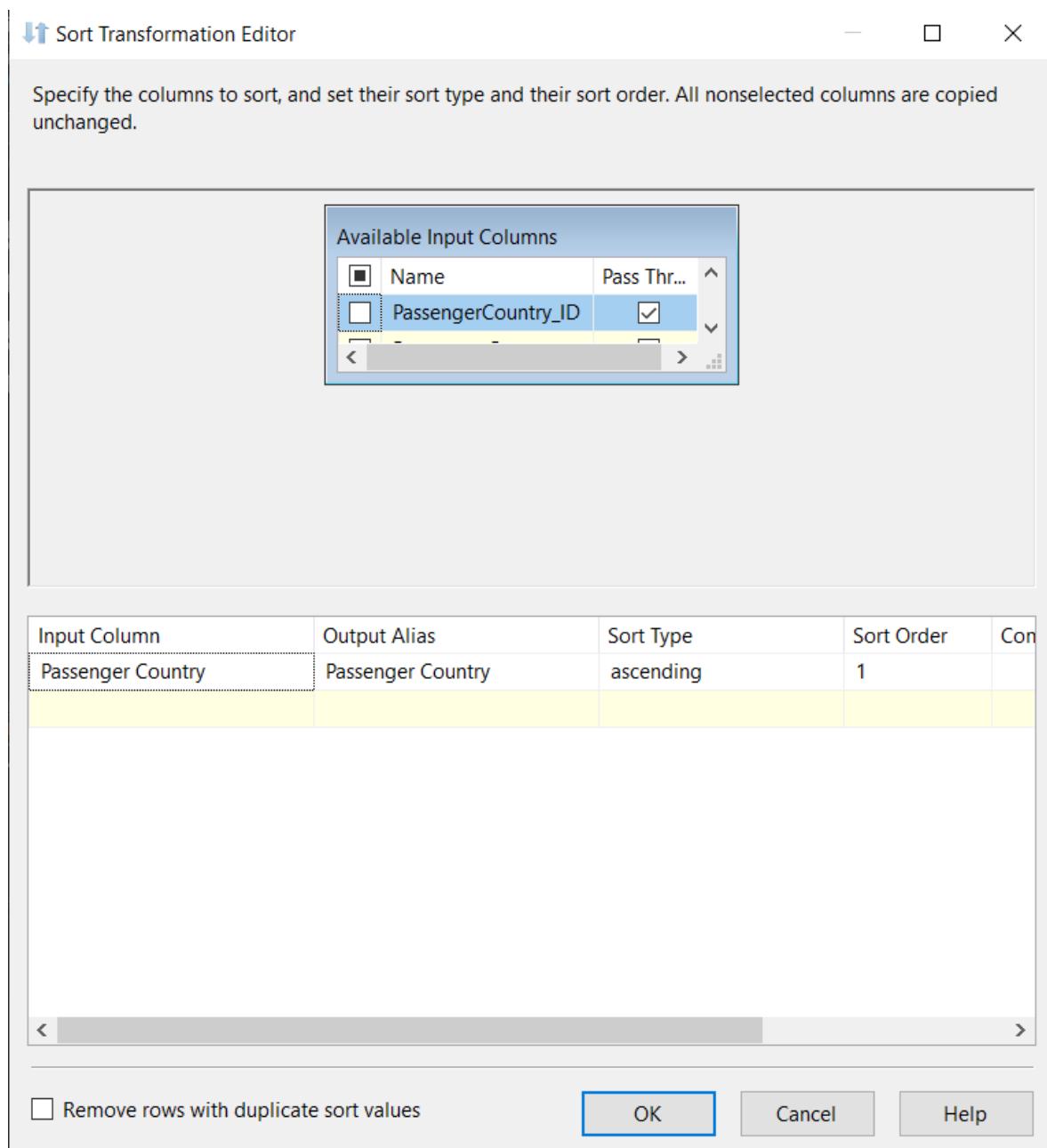
Bước 6. Tạo 2 Sort là Sort và Sort1 tương ứng với mỗi Source.



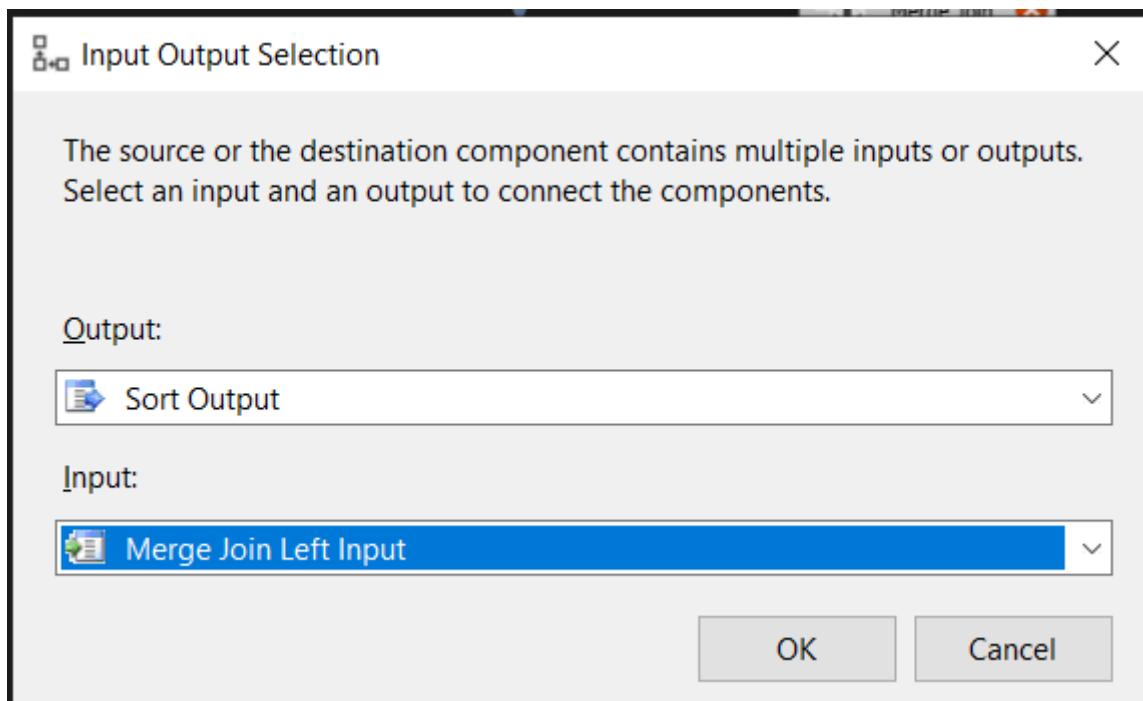
Bước 7. Ở Sort, click chuột phải chọn **Edit** và chọn cột PassengerCountry theo thứ tự giống với bảng Dim_PassengerCountry để chuẩn bị cho quá trình merge.



Bước 8. Tương tự ta chọn các cột cho Sort1

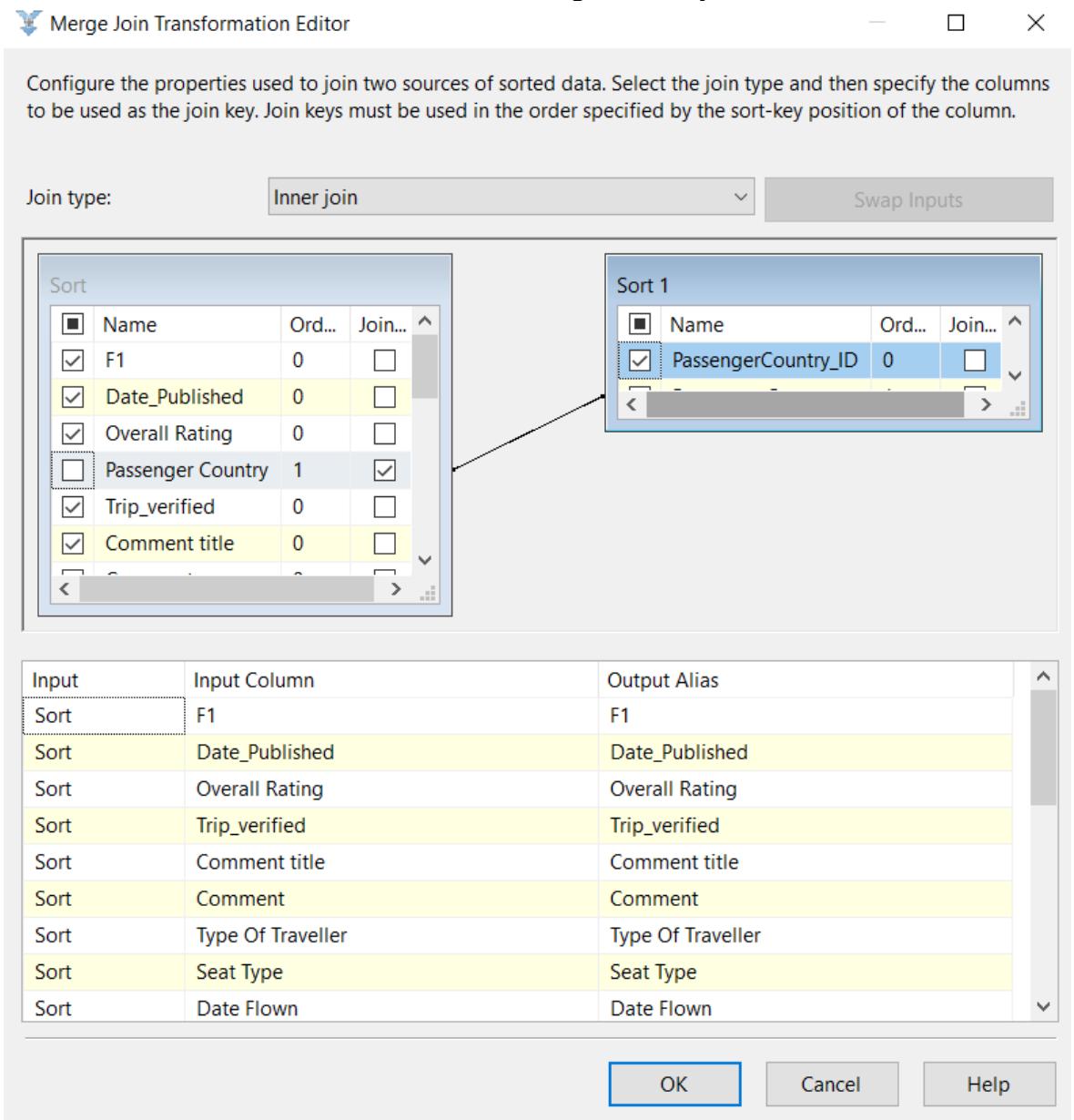


Bước 9. Tạo một **Merge Join** và nối với Sort, tiếp theo ta chọn **Merge Join Left Input** để giữ lại toàn bộ các dòng trong bảng Fact2 bất kể có kết quả khi thực hiện phép kết trái với cột ID của bảng Dim_PassengerCountry hay không.

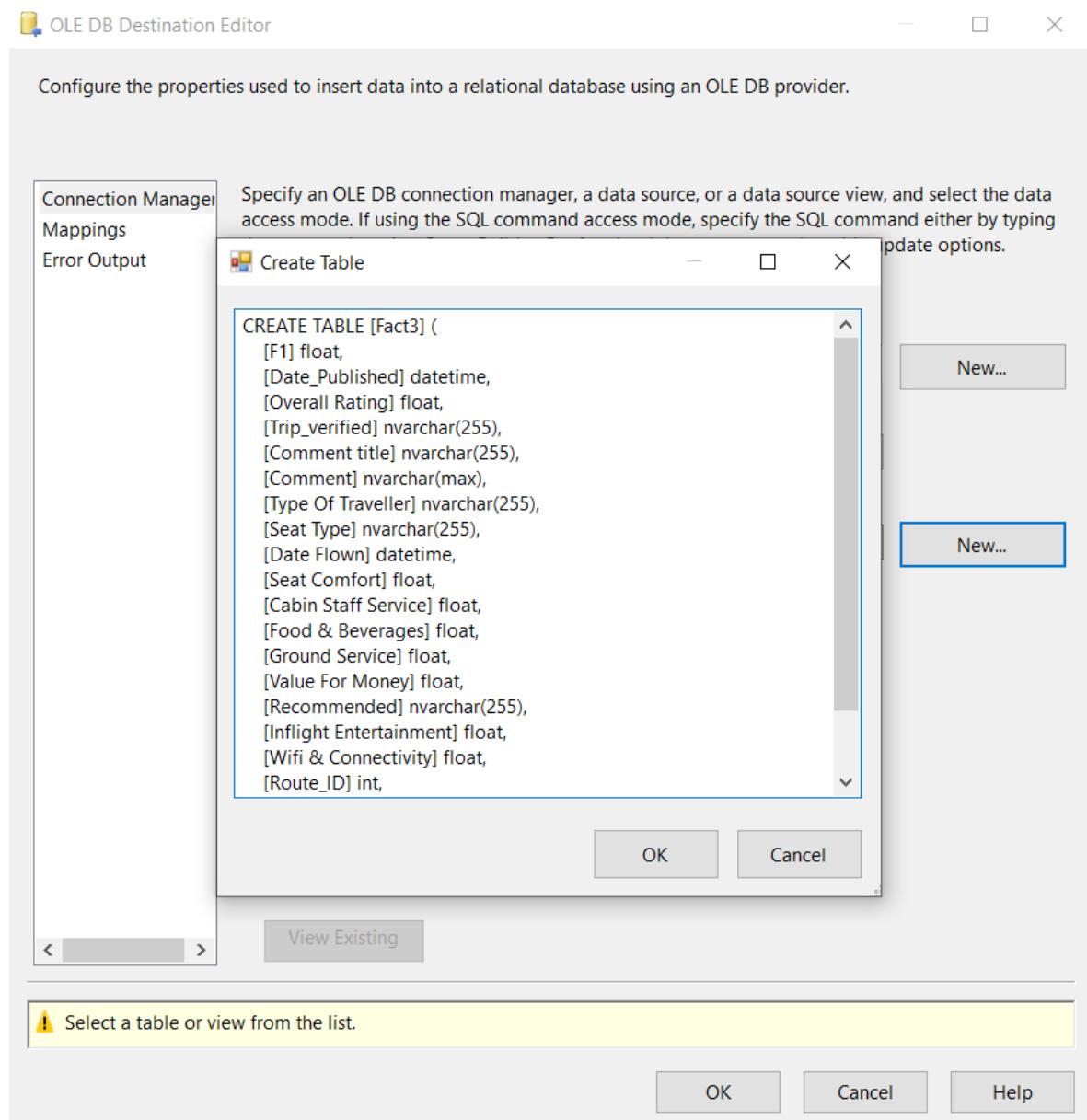


Bước 10.

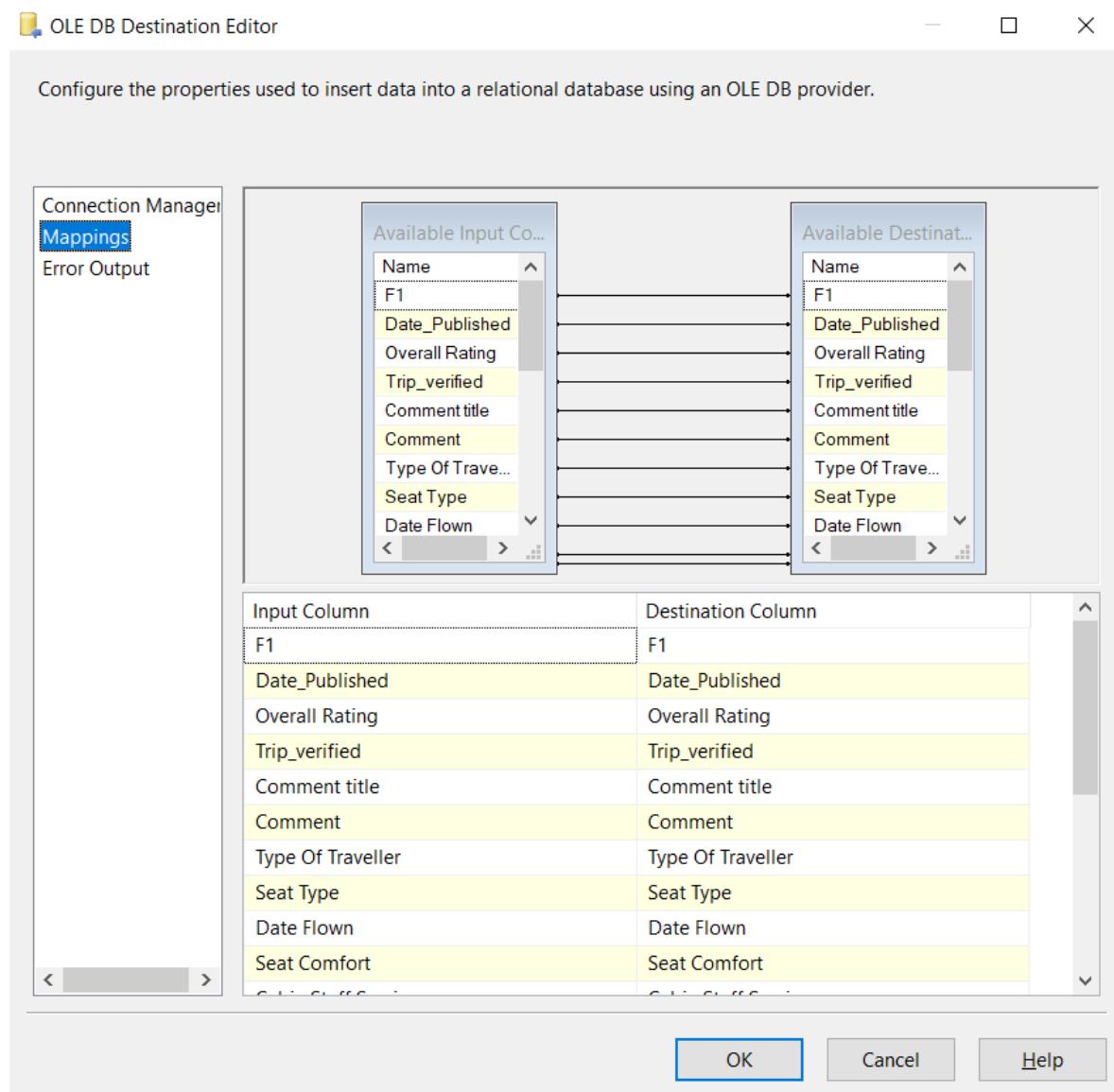
- Chuột phải vào Merge Join và nhấn **Edit**, một hộp thoại merge editor xuất hiện: ở đây ta tick chọn tất cả các cột của Sort nhưng không lấy thuộc tính PassengerCountry
- Tiếp theo ta chọn PassengerCountry_ID ở Sort1 để merge vào Fact2
- Kết quả sau khi merge là bảng Fact2 không còn thuộc tính PassengerCountry và có thêm 1 thuộc tính mới là PassengerCountry_ID



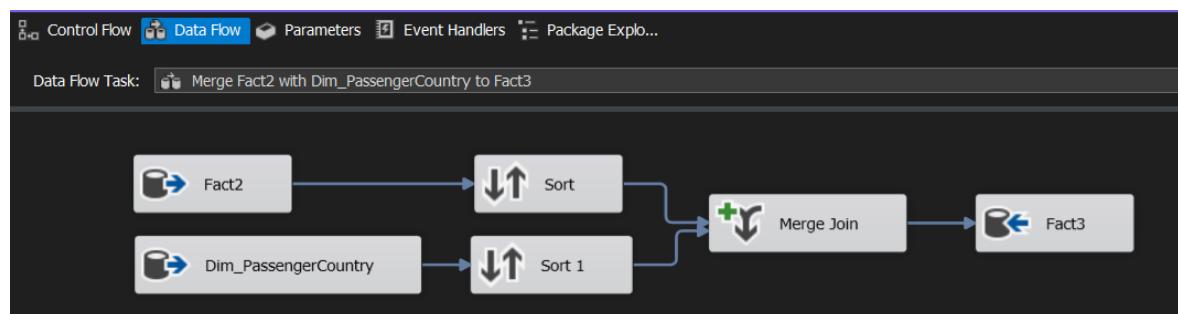
Bước 11. Tạo bảng Fact3 từ một **OLE DB Destination** để chứa tất cả những gì đã merge



Chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

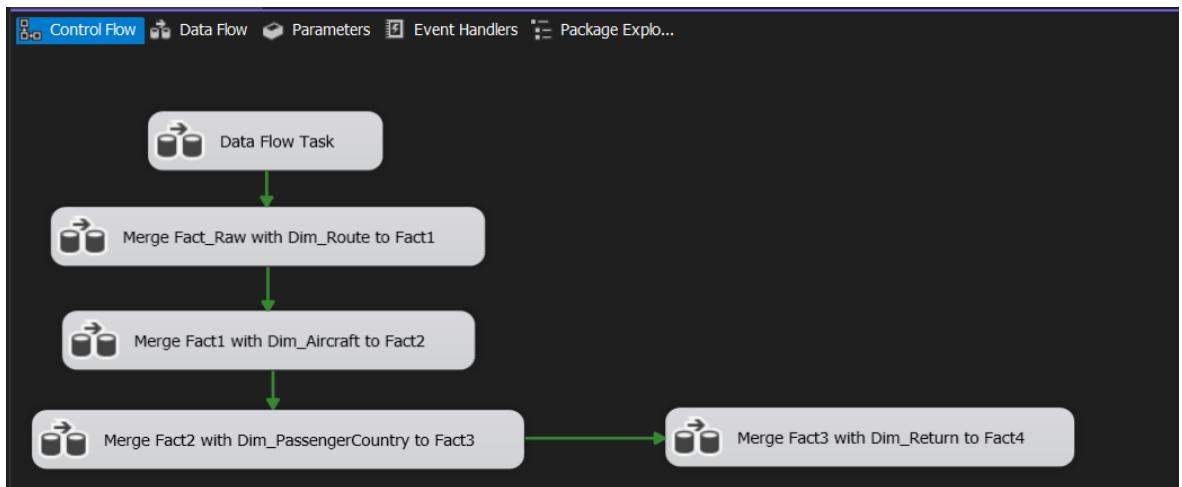


Kết quả:

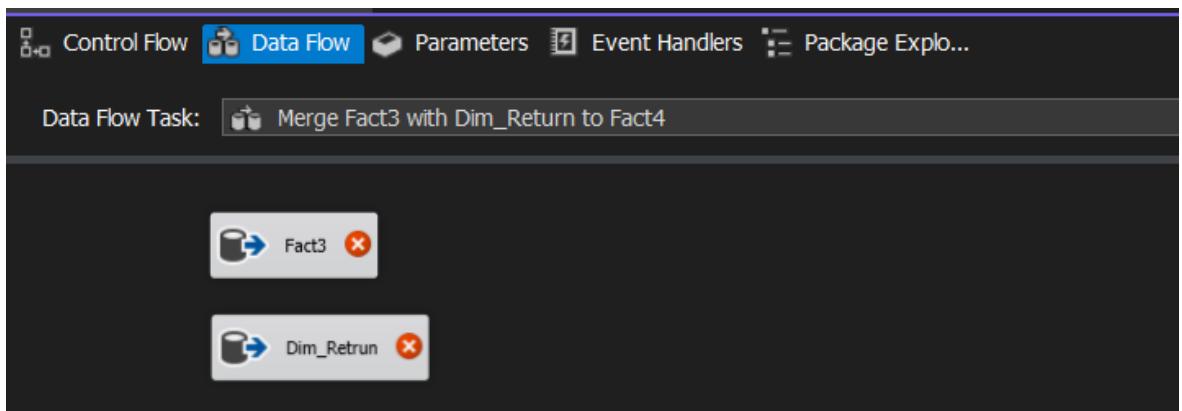


3.3.8.4. Merge Fact3 và Dim_Return vào Fact4

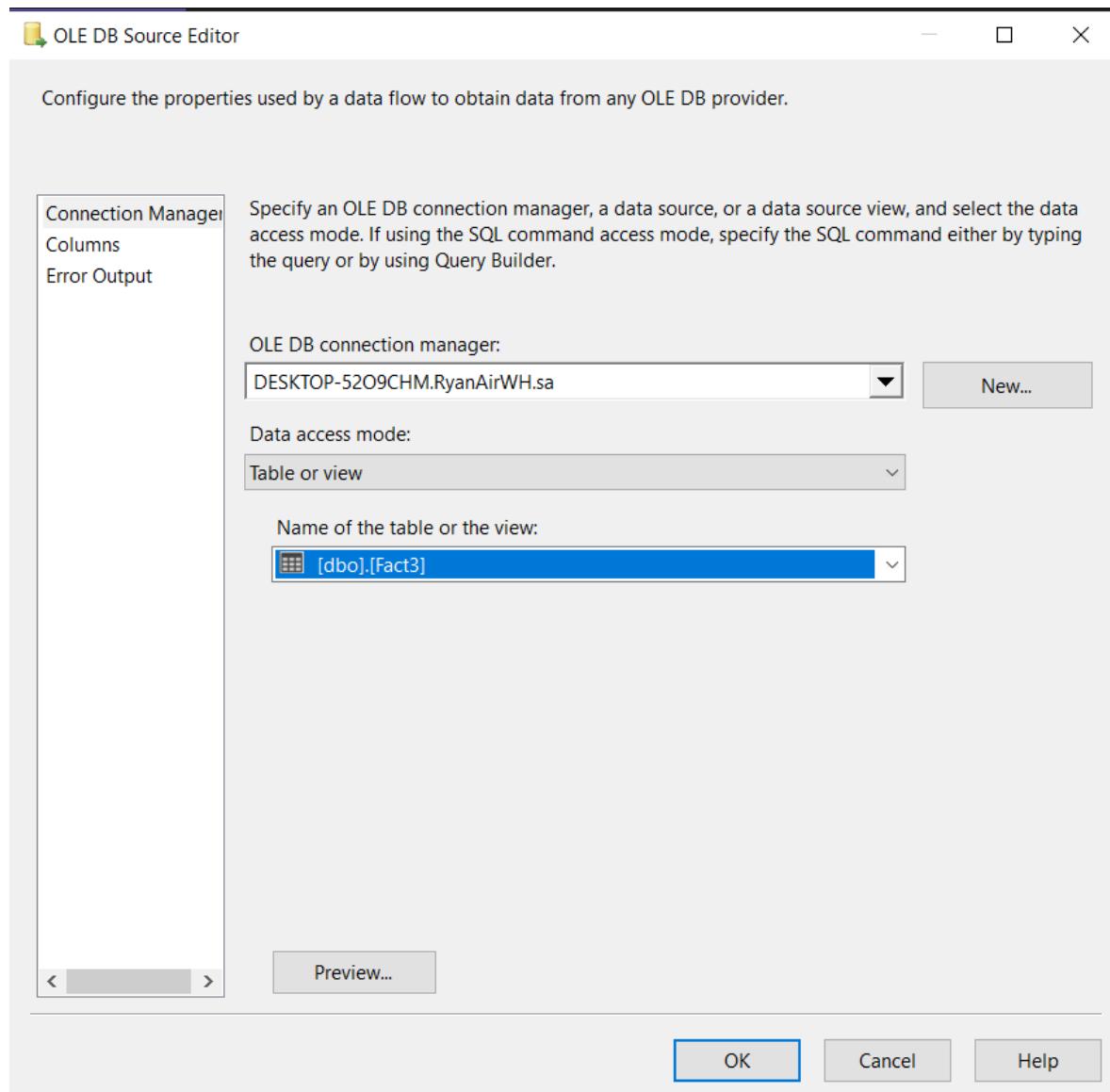
Bước 1. Ở tab Control Flow, tạo Data Flow Task và đổi tên Data Flow Task là “Merge Fact3 with Dim_Return to Fact4”



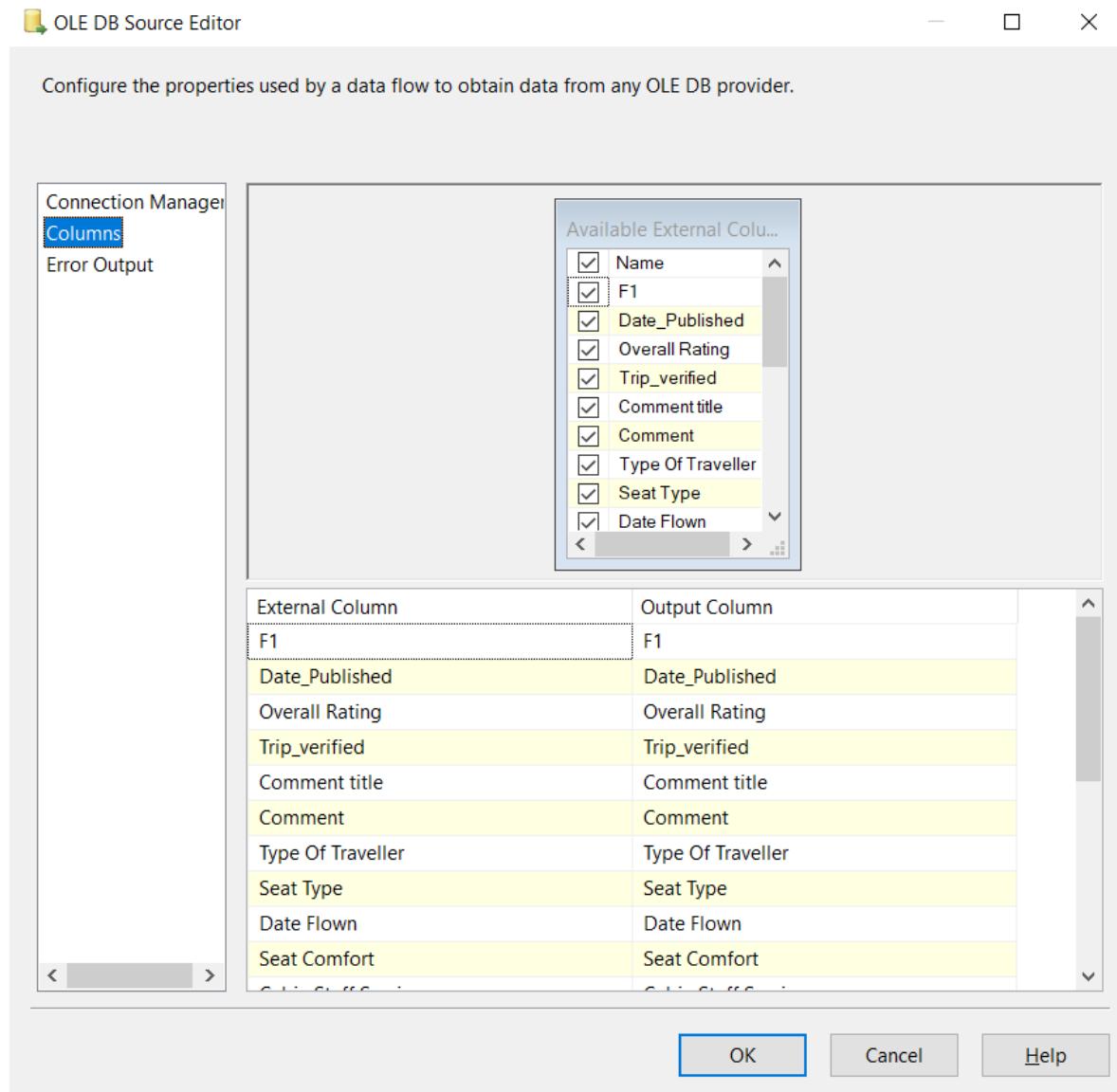
Bước 2. Click chuột phải vào Data Flow Task nói trên và chọn Edit, trong tab Data Flow ta tạo 2 OLE DB Source và đổi tên Fact3 và Dim_Return



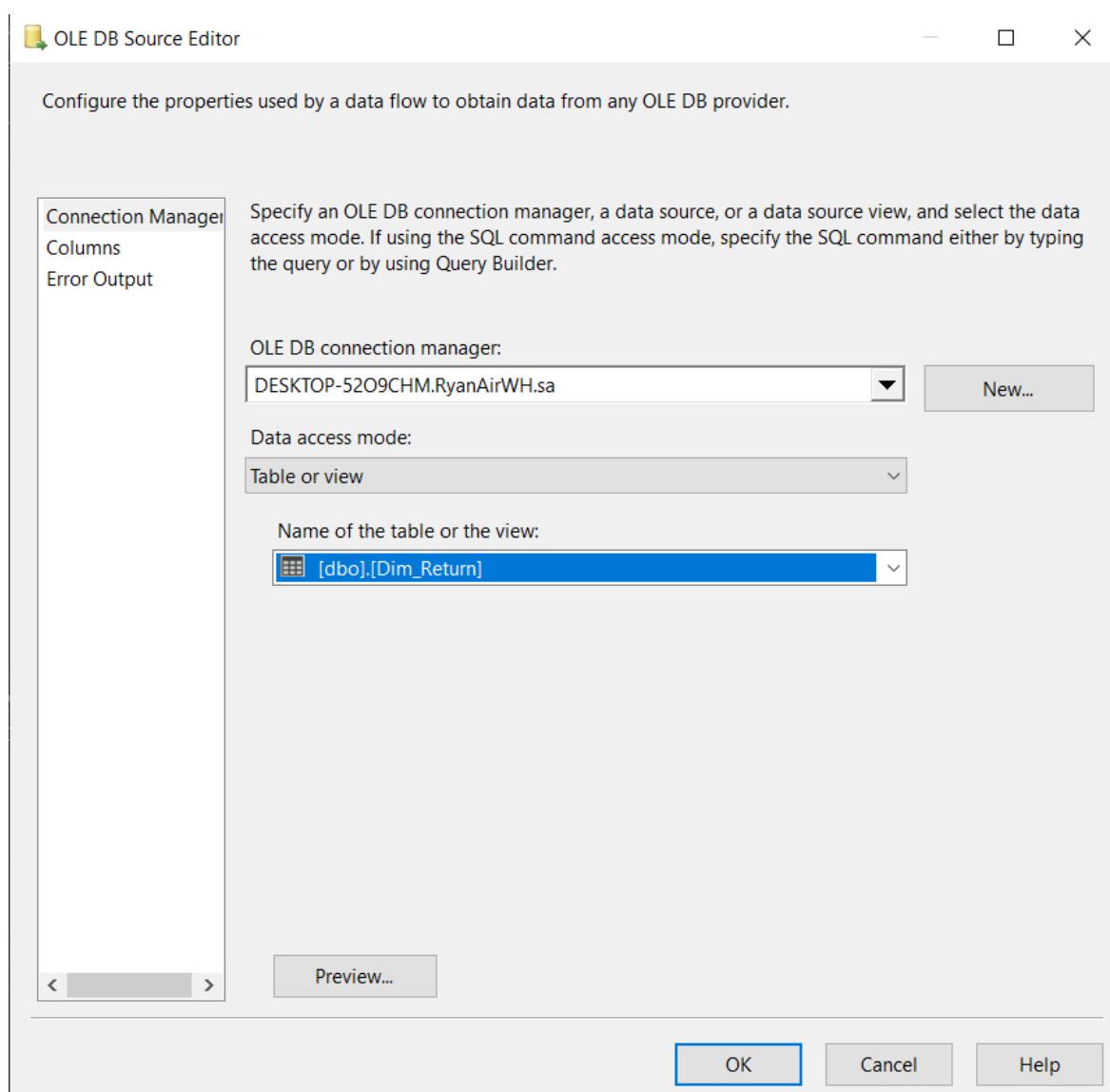
Bước 3. Click chuột phải chọn **Edit**, sau đó chọn bảng Fact3 đã tạo trước đó làm data source cho bảng Fact3 mới này.

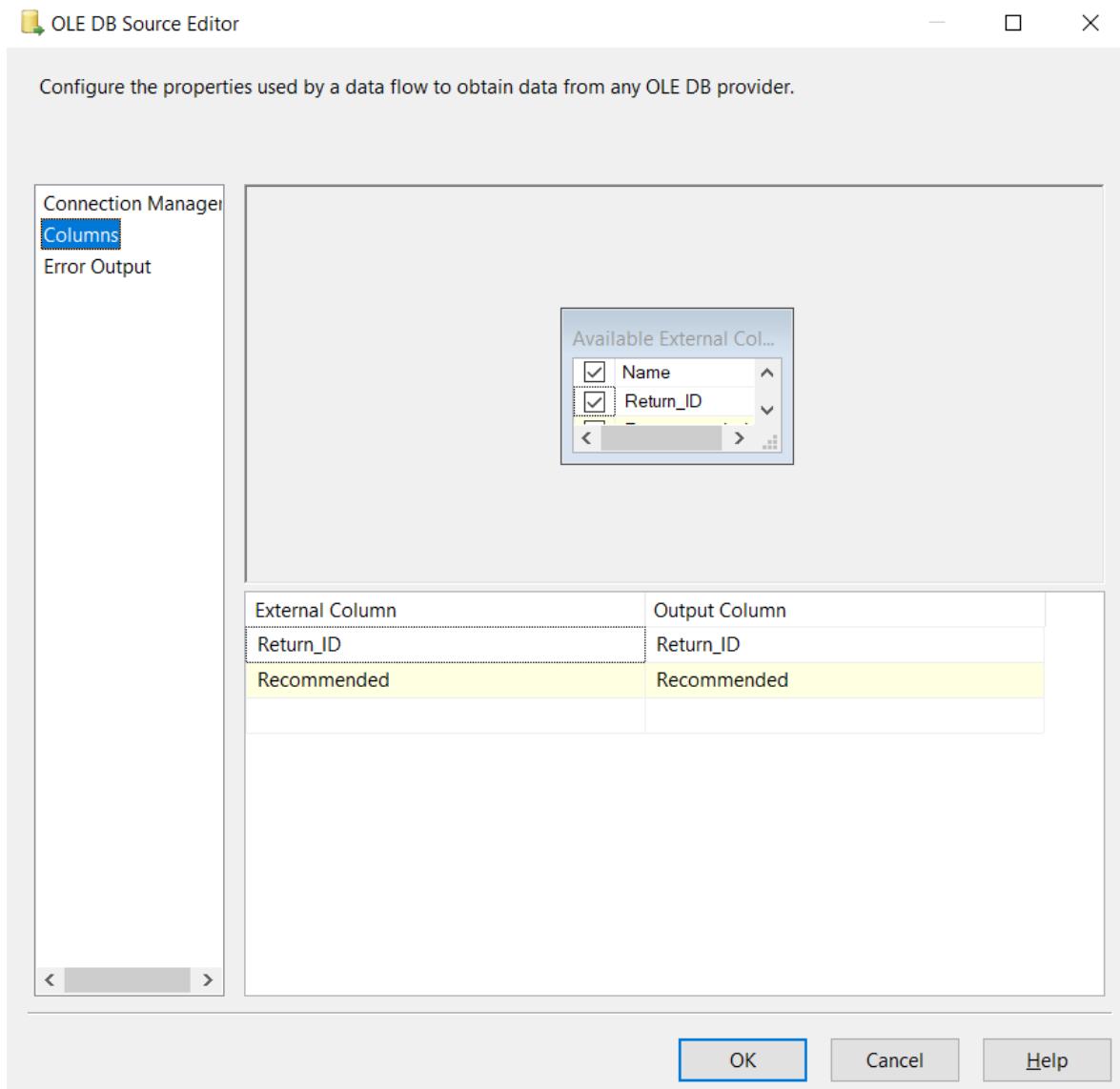


Bước 4. Chọn mục **Columns** để xem xét các cột được ánh xạ. Nhấn **OK**.

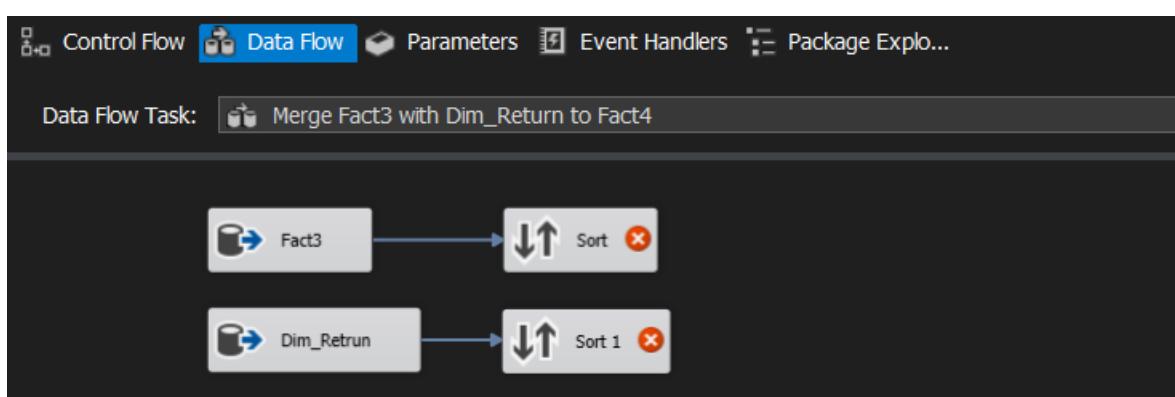


Bước 5. Tương tự thực hiện chọn ánh xạ cột cho Dim_Return

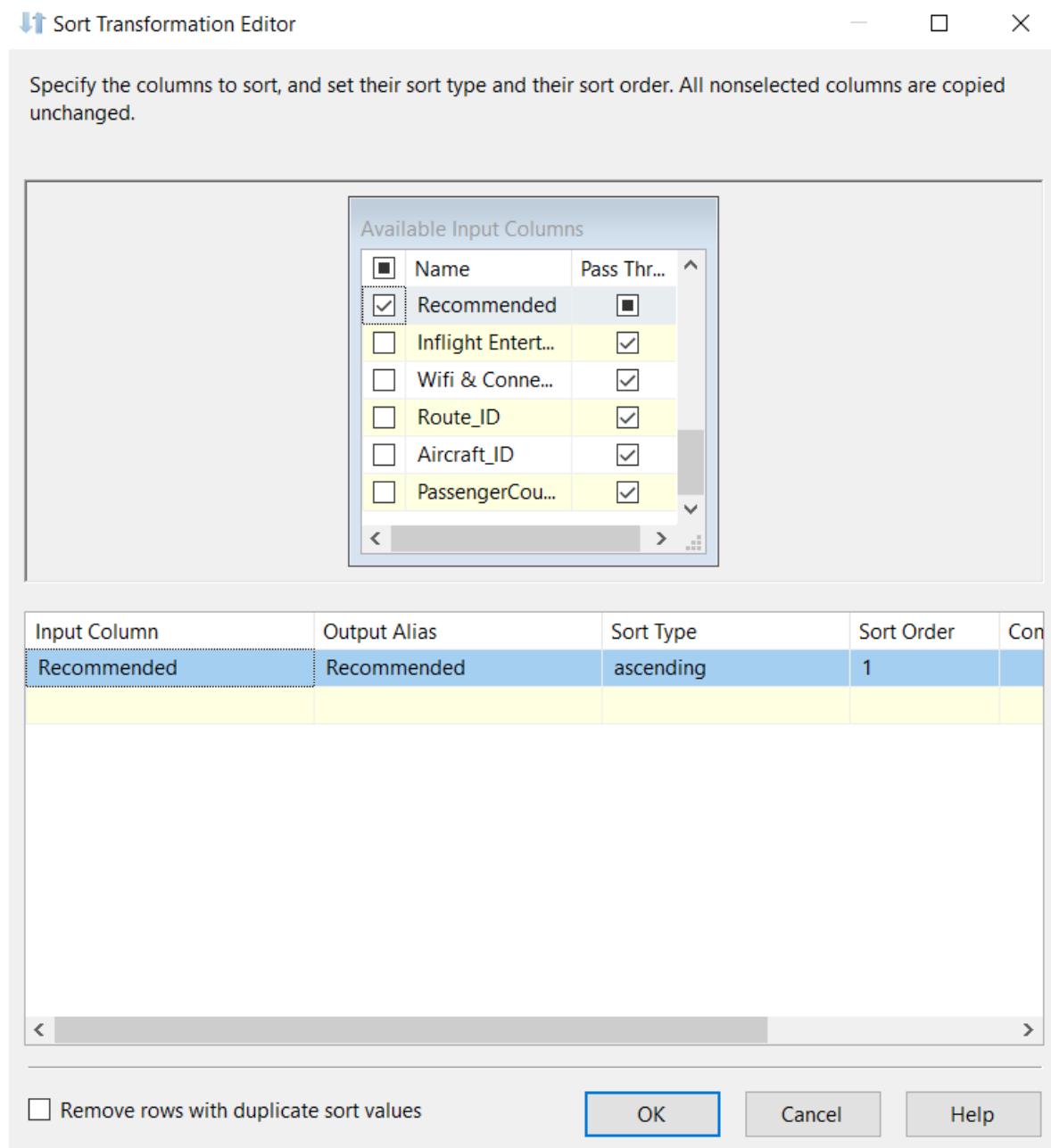




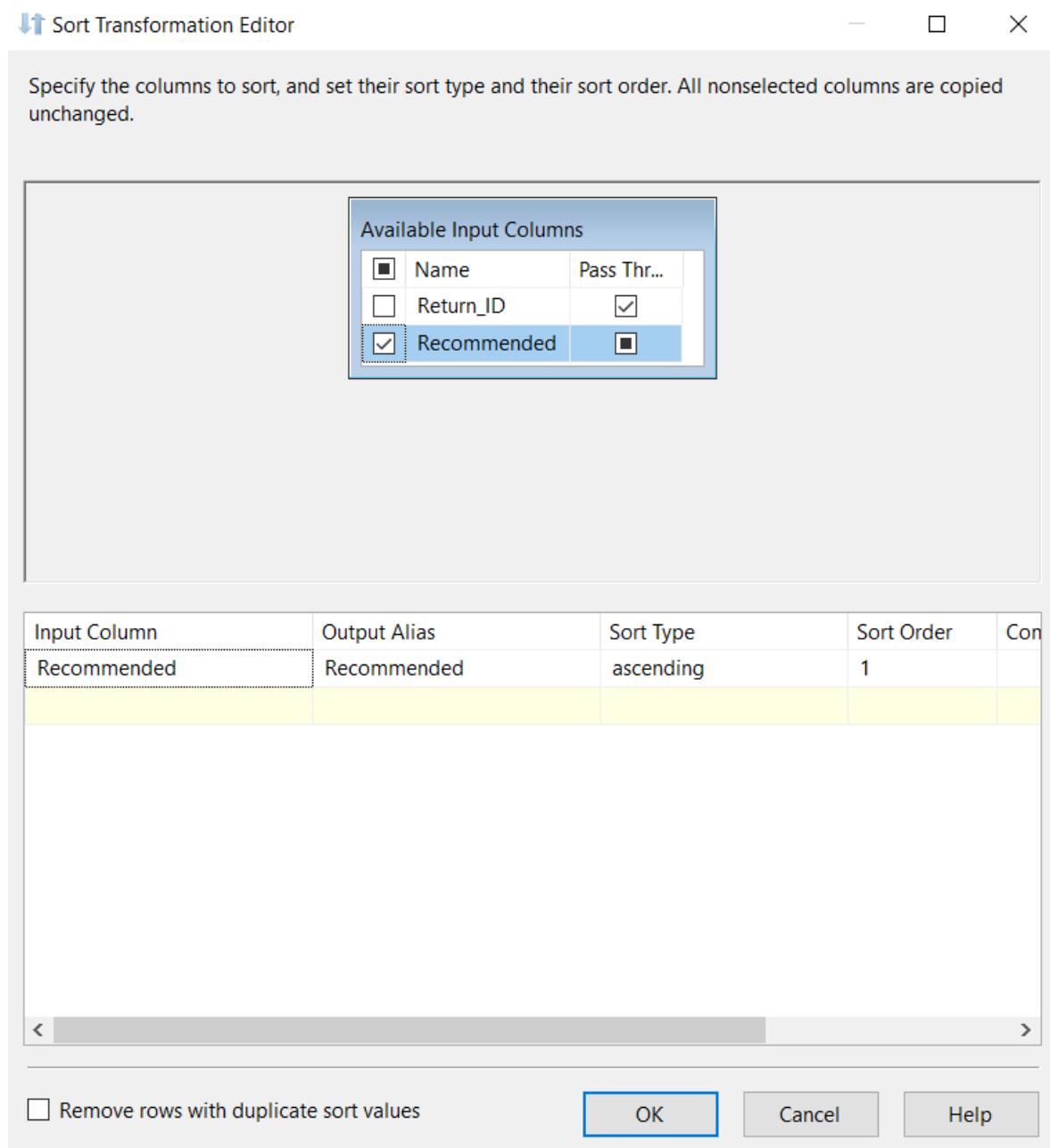
Bước 6. Tạo 2 Sort là Sort và Sort1 tương ứng với mỗi Source.



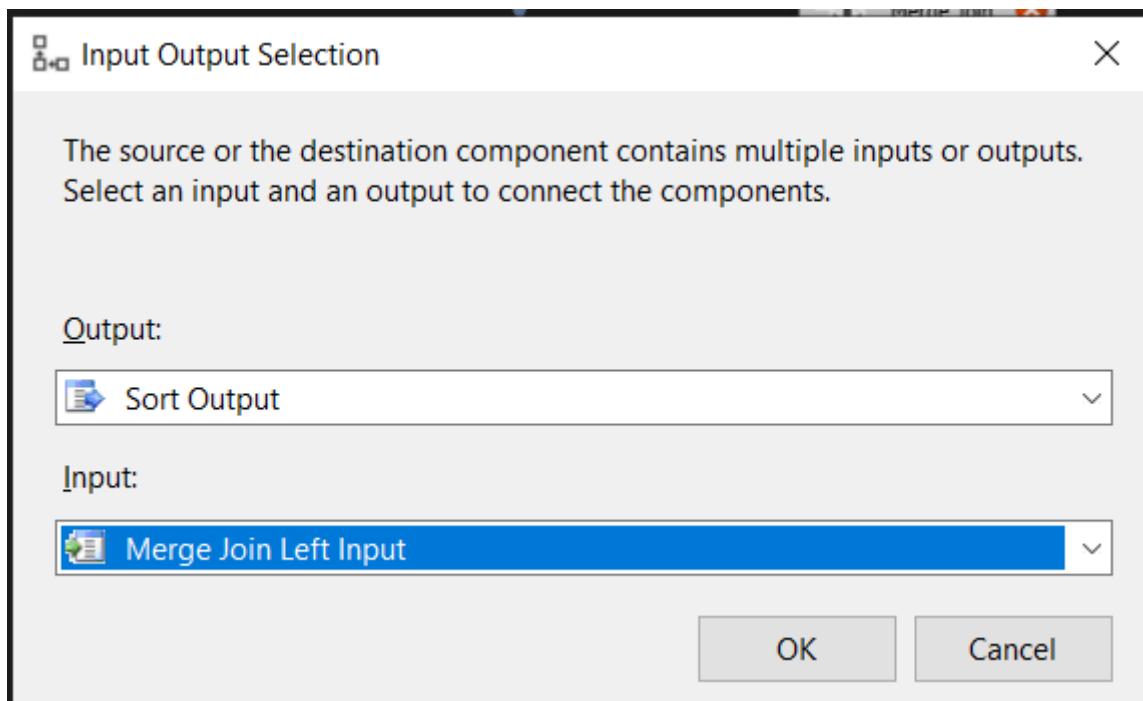
Bước 7. Ở Sort, click chuột phải chọn **Edit** và chọn cột Recommended theo thứ tự giống với bảng Dim_Return để chuẩn bị cho quá trình merge.



Bước 8. Tương tự ta chọn các cột cho Sort1

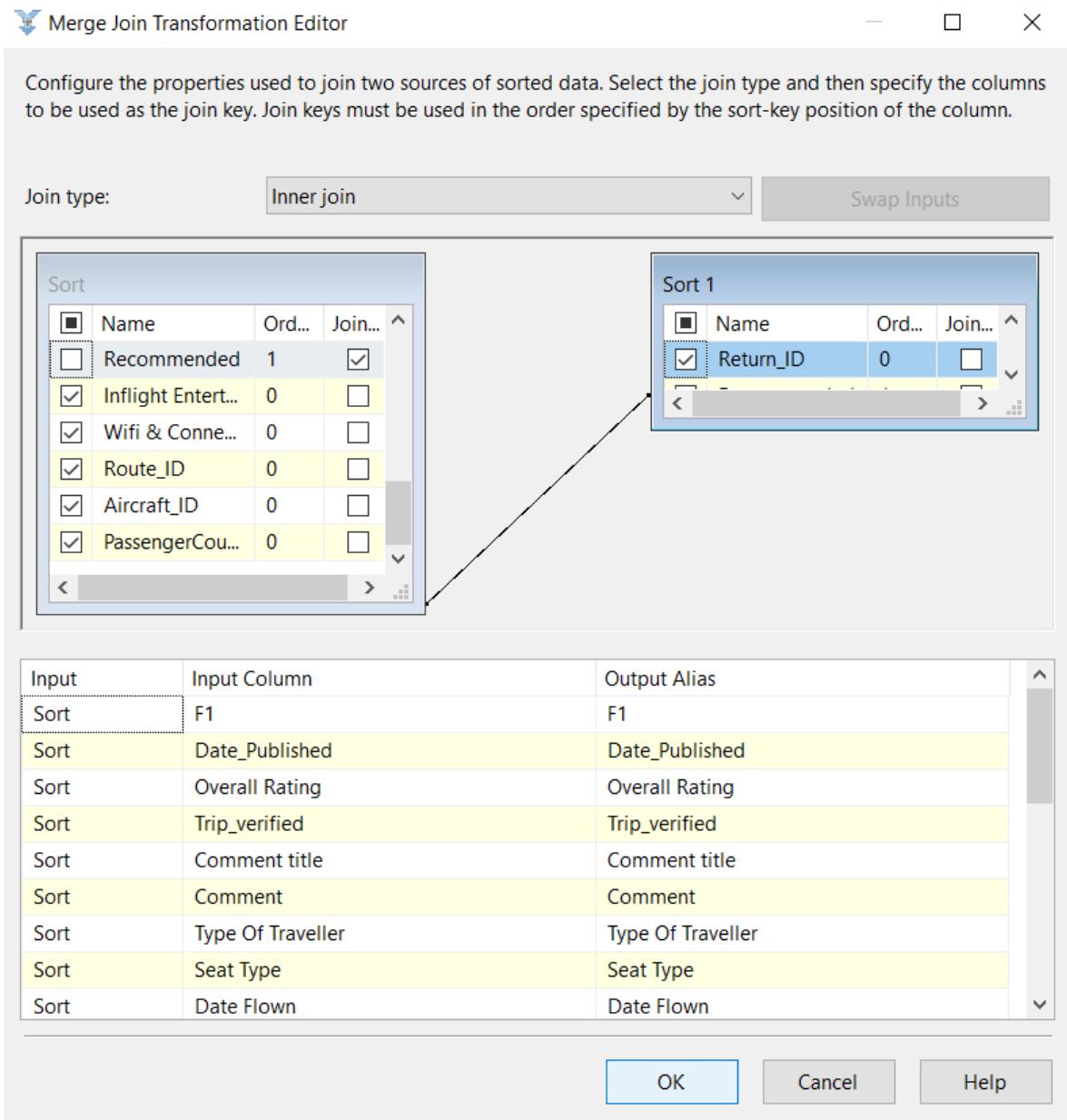


Bước 9. Tạo một **Merge Join** và nối với Sort, tiếp theo ta chọn **Merge Join Left Input** để giữ lại toàn bộ các dòng trong bảng Fact3 bất kể có kết quả khi thực hiện phép kết trái với cột ID của bảng Dim_Return hay không.

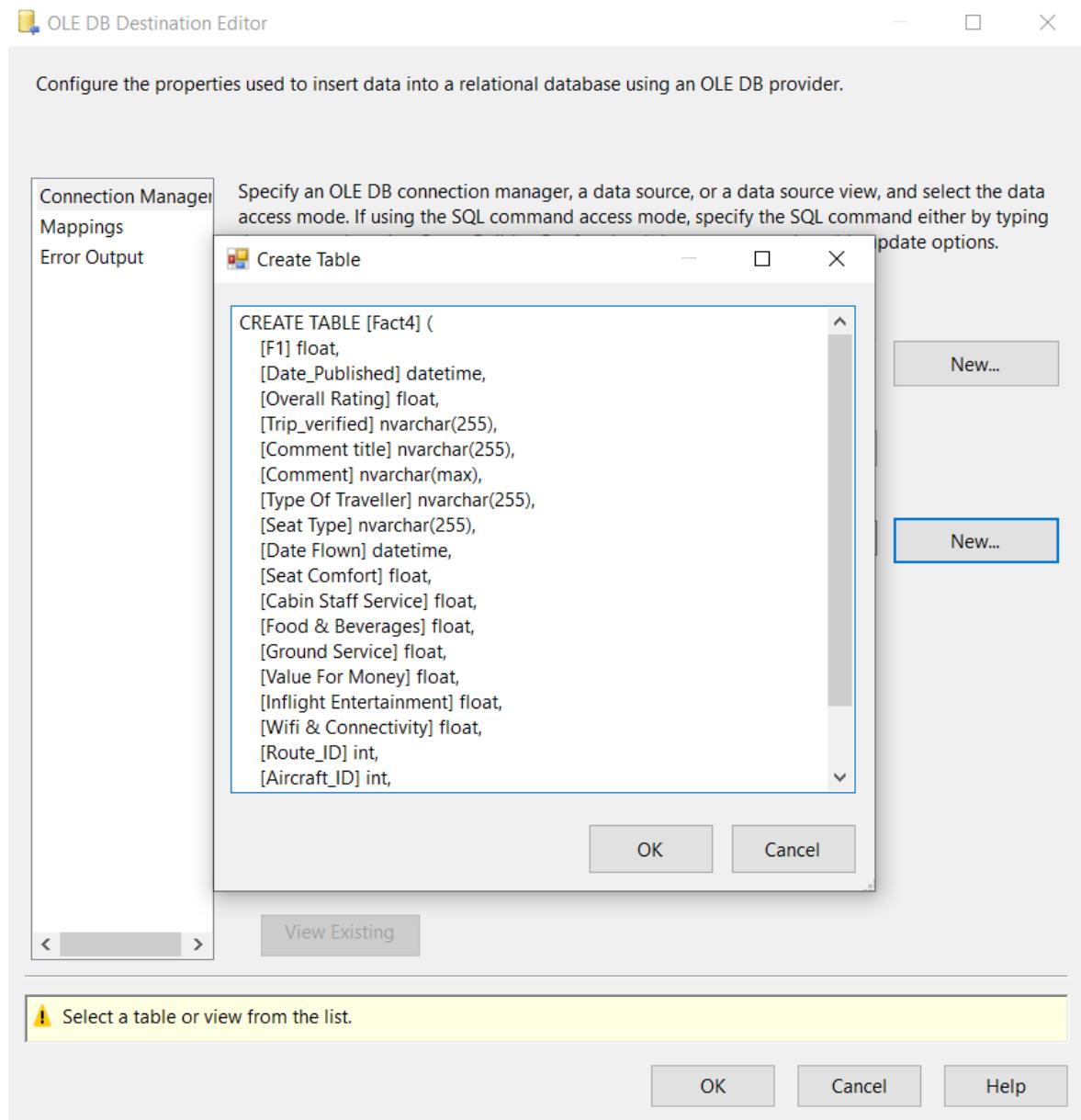


Bước 10.

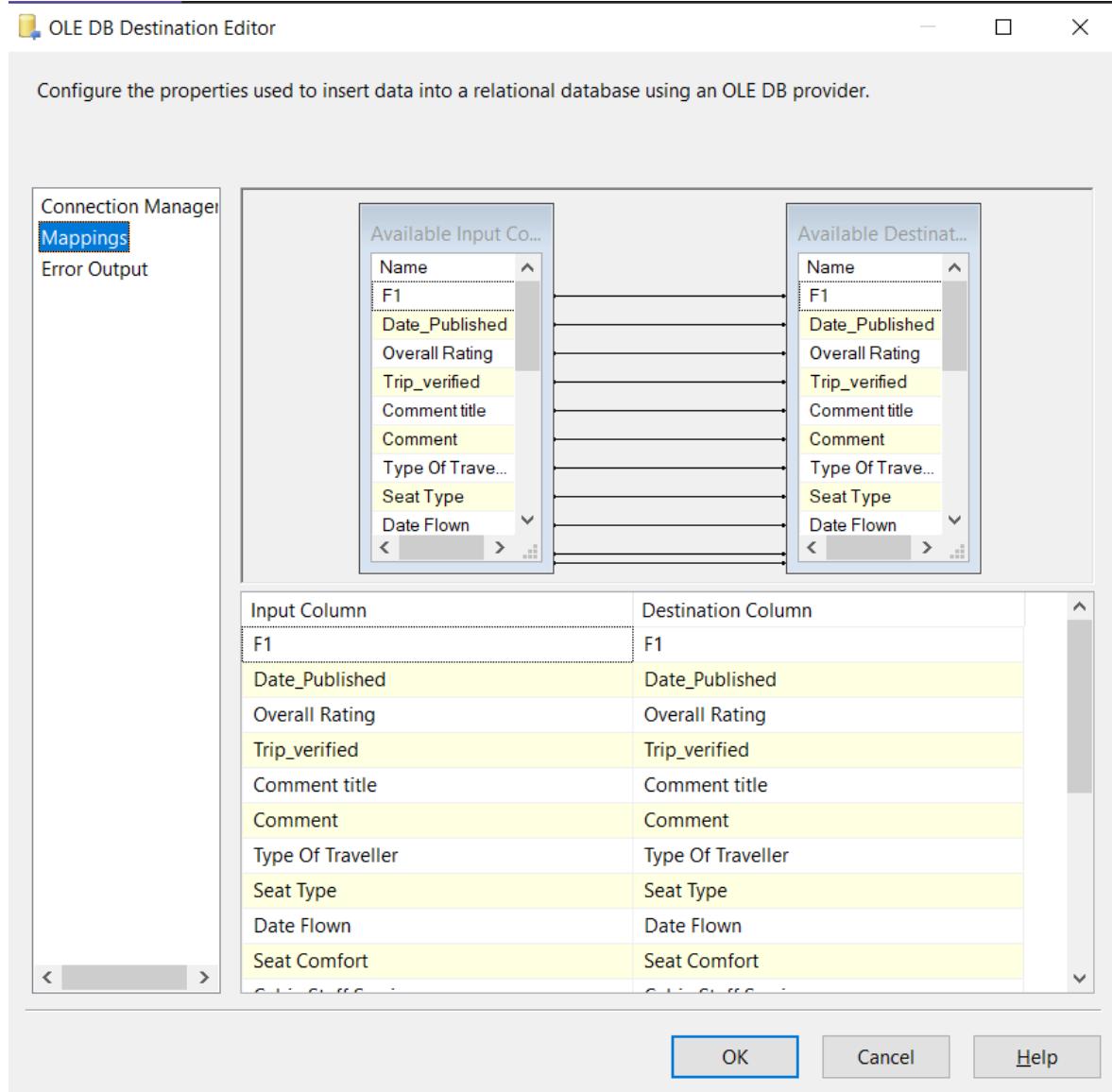
- Chuột phải vào Merge Join và nhấn **Edit**, một hộp thoại merge editor xuất hiện: ở đây ta tick chọn tất cả các cột của Sort nhưng không lấy thuộc tính Recommended
- Tiếp theo ta chọn Return_ID ở Sort1 để merge vào Fact3
- Kết quả sau khi merge là bảng Fact3 không còn thuộc tính Recommended và có thêm 1 thuộc tính mới là Return_ID



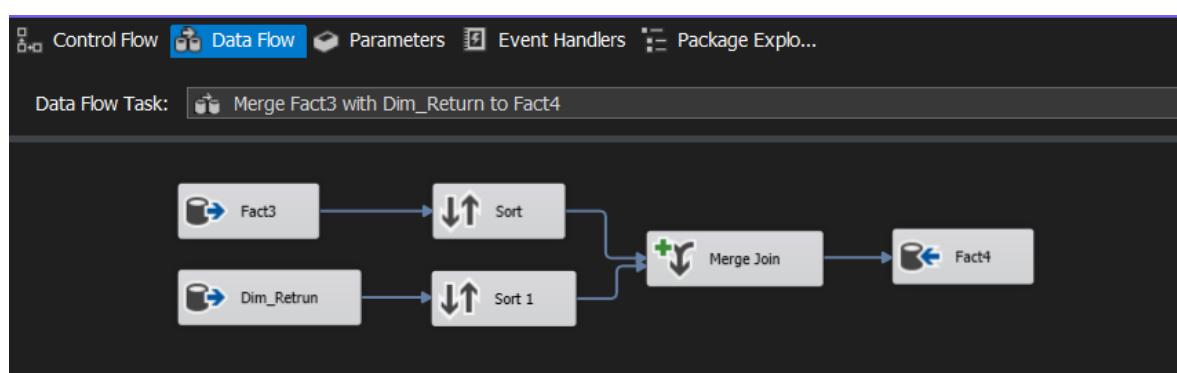
Bước 11. Tạo bảng Fact4 từ một **OLE DB Destination** để chứa tất cả những gì đã merge



Chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu



Kết quả:

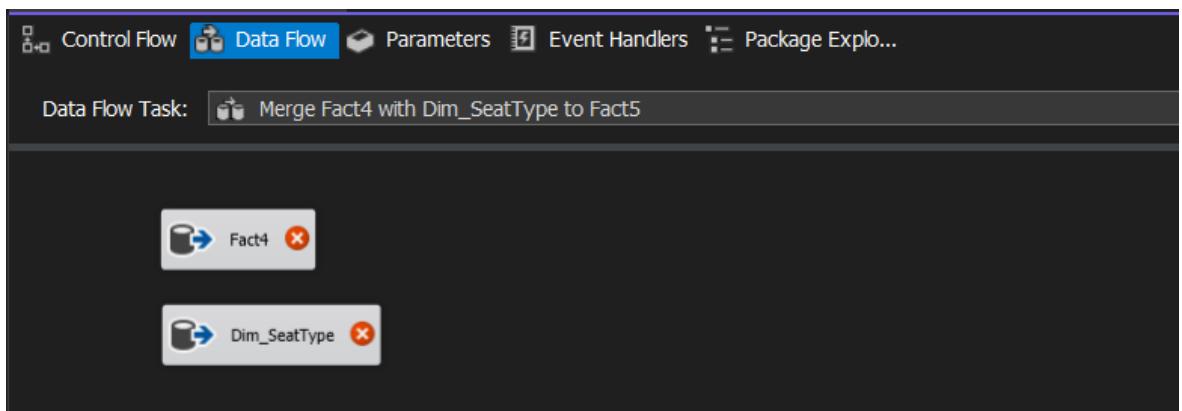


3.3.8.5. Merge Fact4 và Dim_SeatType vào Fact5

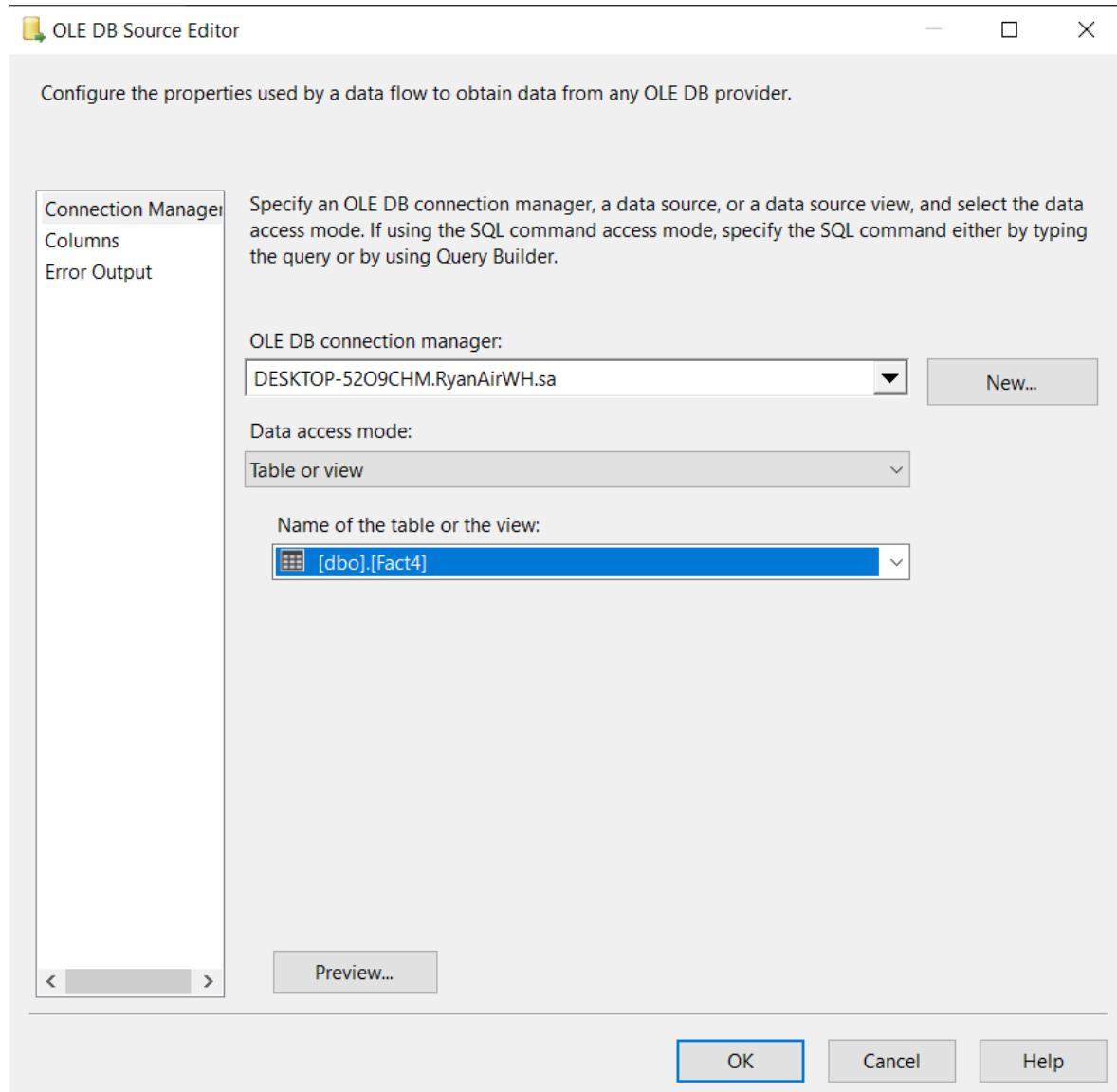
Bước 1. Ở tab Control Flow, tạo **Data Flow Task** và đổi tên Data Flow Task là “Merge Fact4 with Dim_SeatType to Fact5”



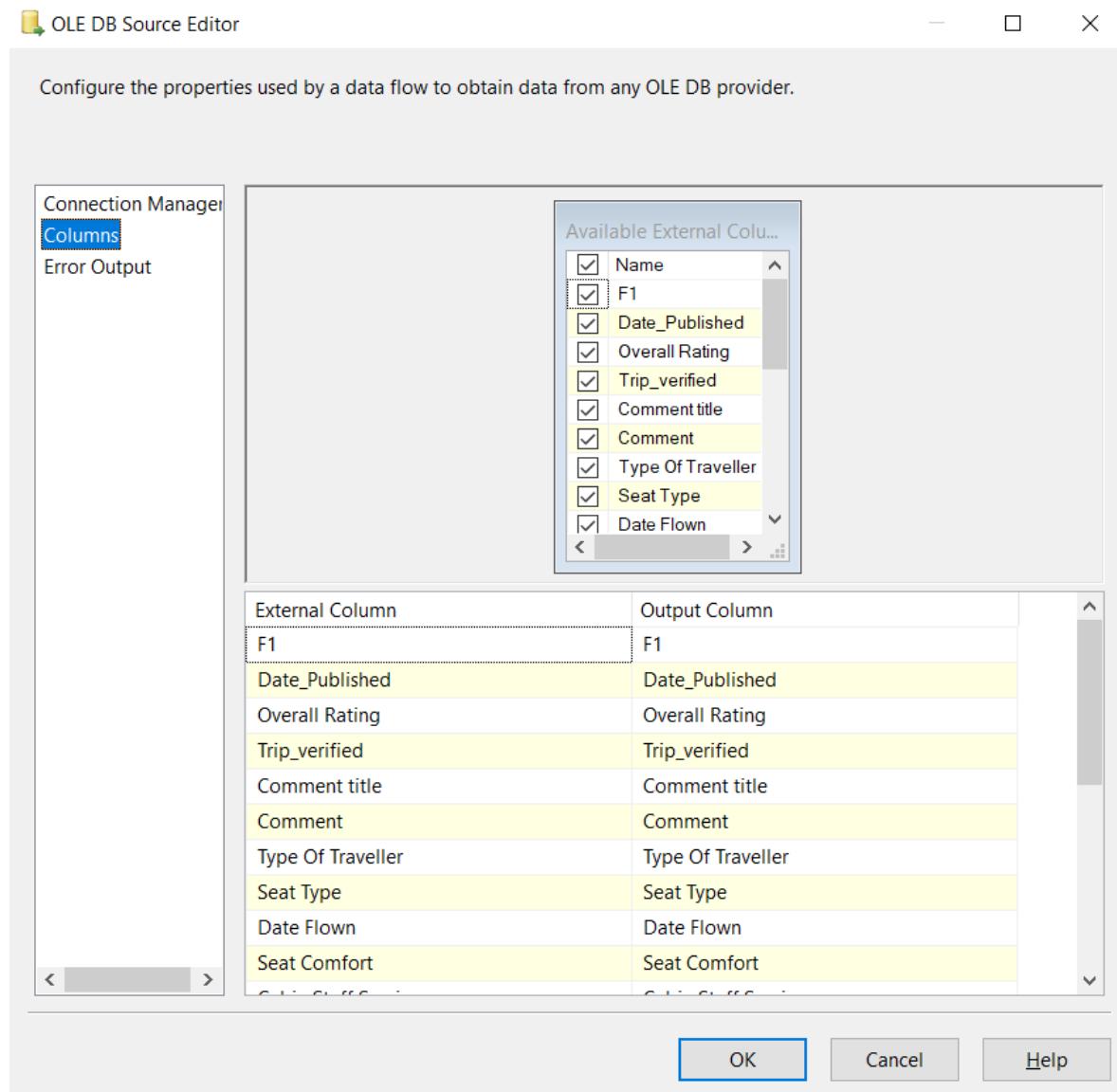
Bước 2. Click chuột phải vào Data Flow Task nói trên và chọn **Edit**, trong tab Data Flow ta tạo 2 OLE DB Source và đổi tên Fact4 và Dim_SeatType



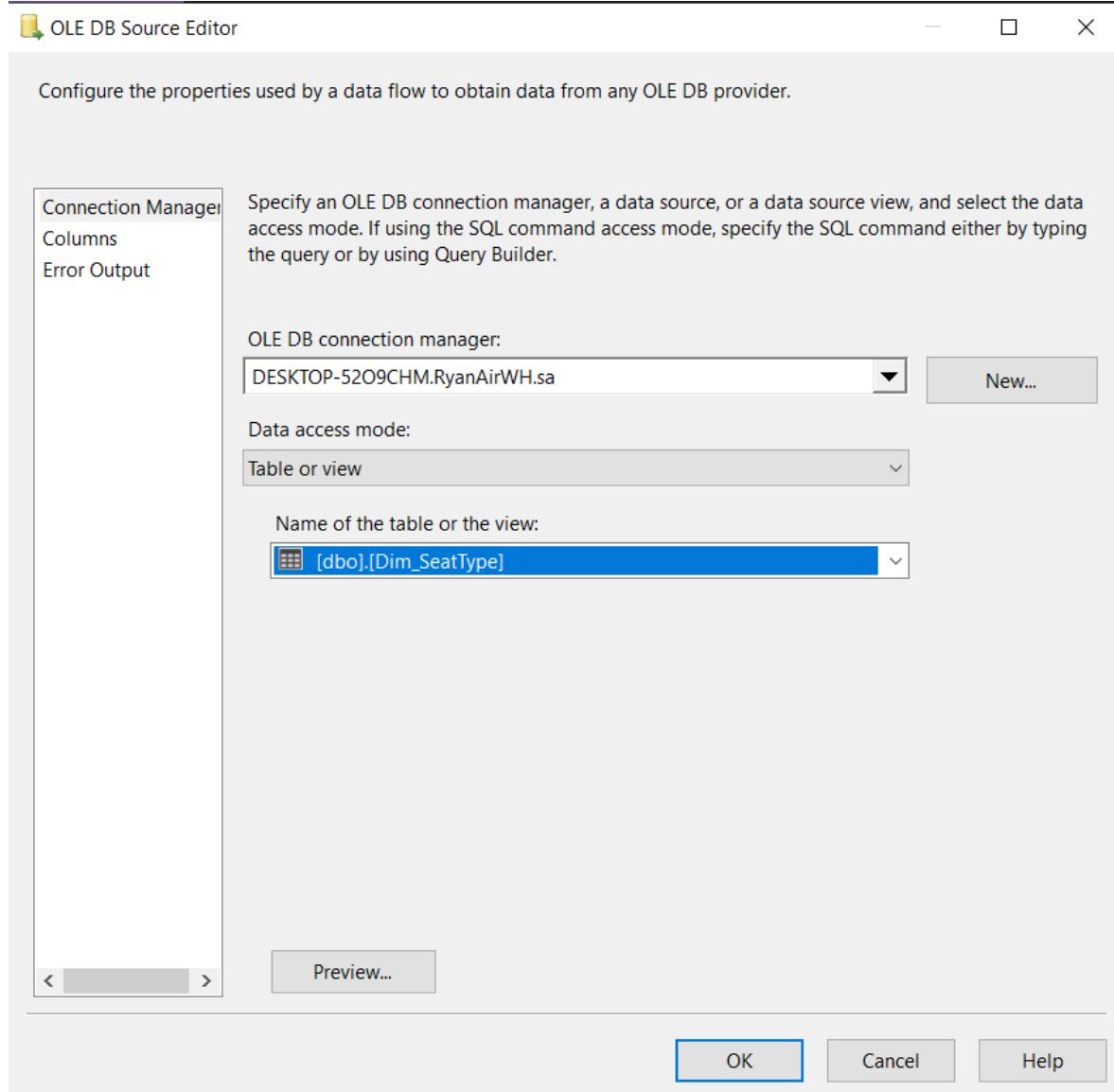
Bước 3. Click chuột phải chọn **Edit**, sau đó chọn bảng Fact4 đã tạo trước đó làm data source cho bảng Fact4 mới này.

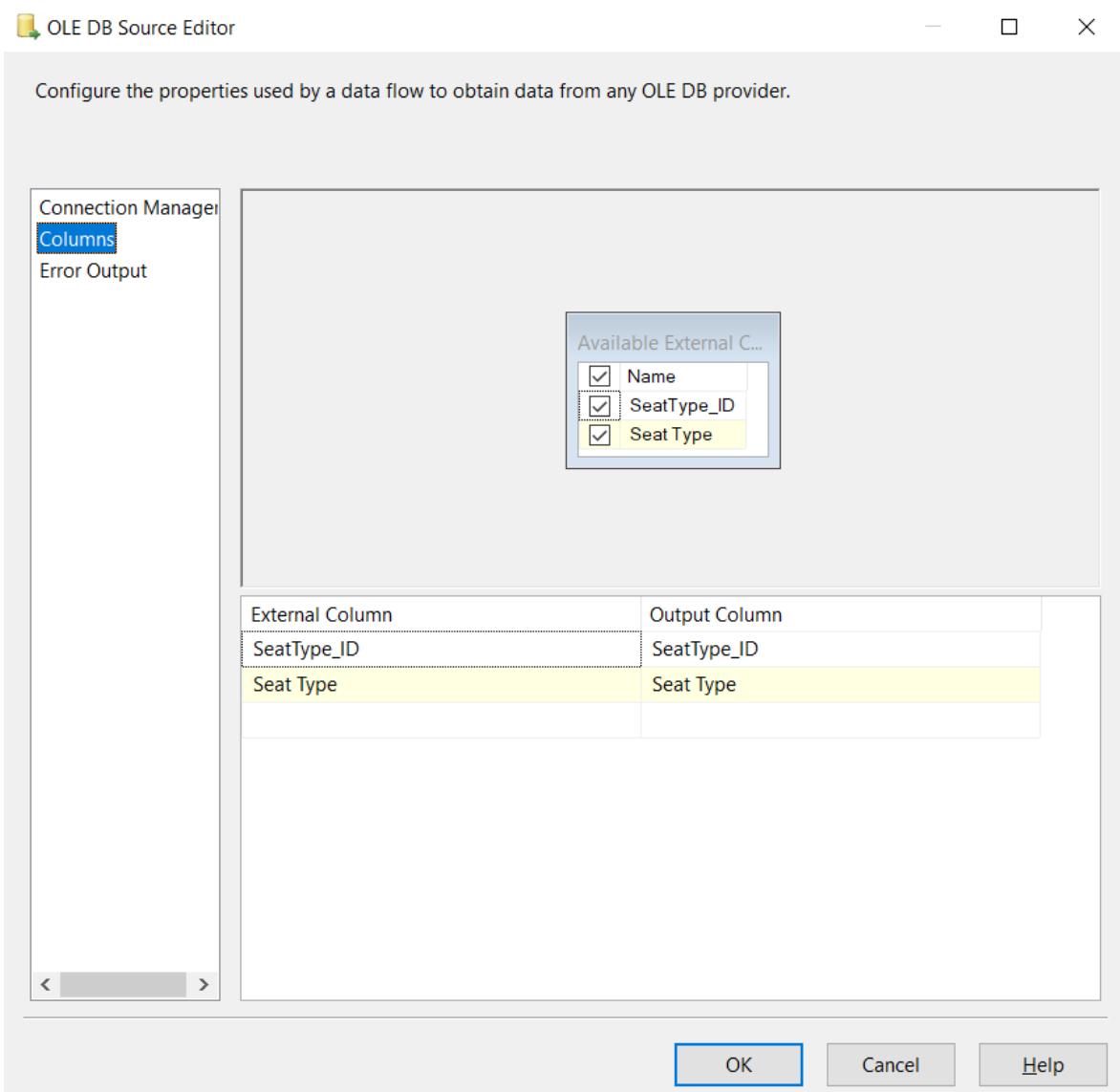


Bước 4. Chọn mục **Columns** để xem xét các cột được ánh xạ. Nhấn **OK**.

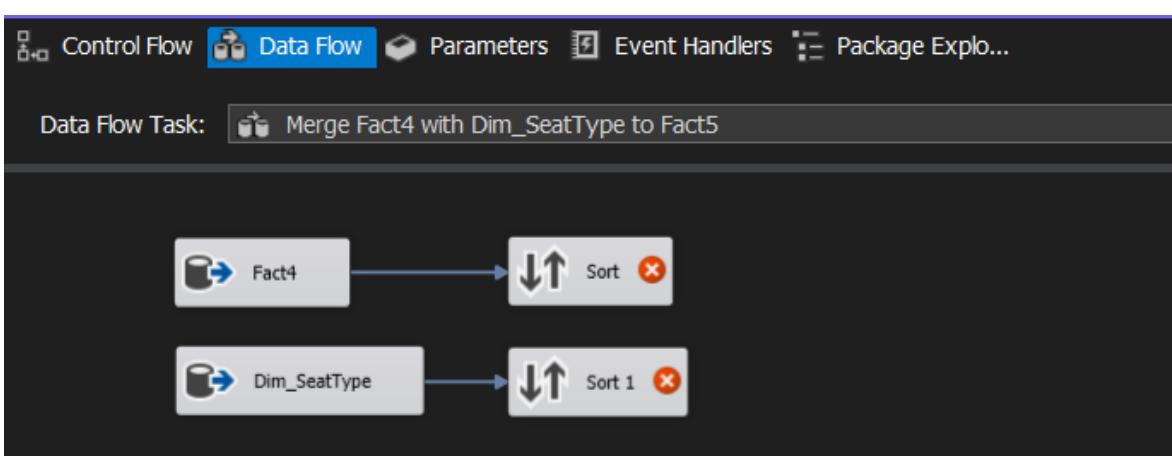


Bước 5. Tương tự thực hiện chọn ánh xạ cột cho Dim_SeatType

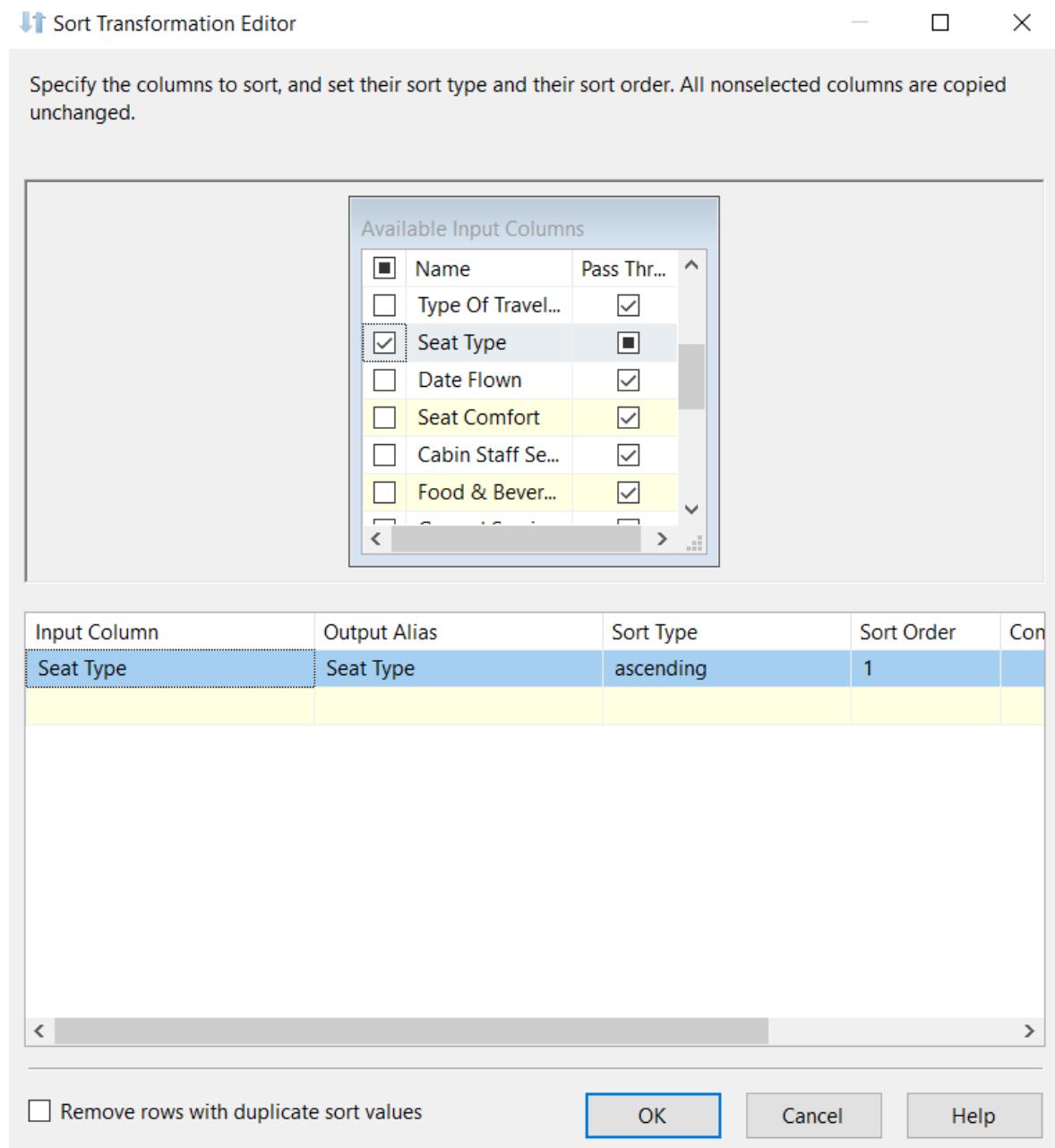




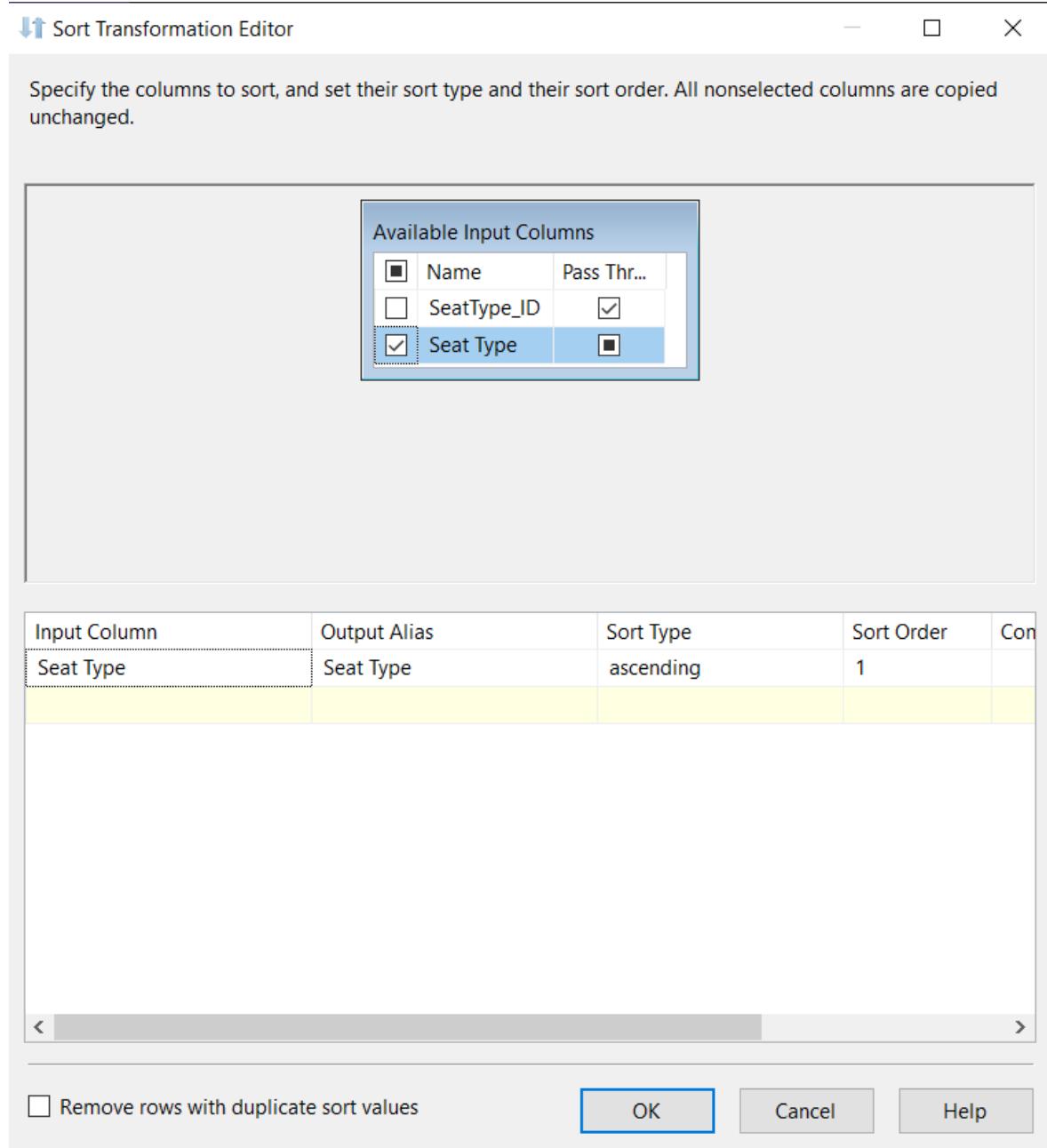
Bước 6. Tạo 2 Sort là Sort và Sort1 tương ứng với mỗi Source.



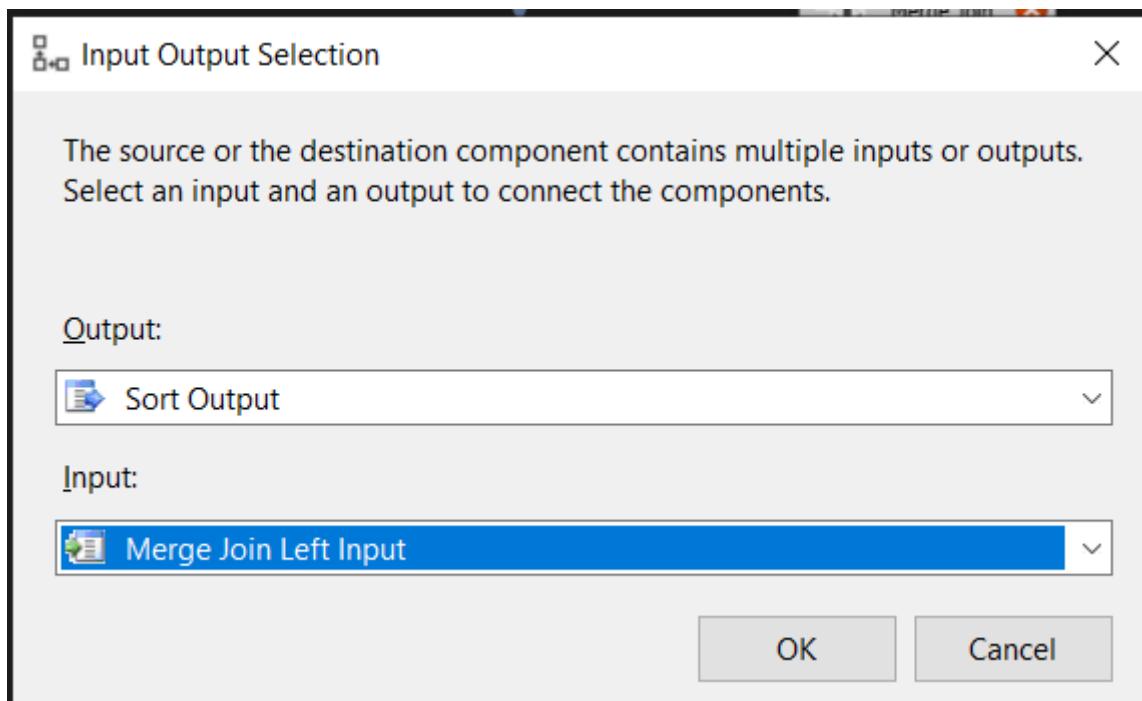
Bước 7. Ở Sort, click chuột phải chọn **Edit** và chọn cột Seat Type theo thứ tự giống với bảng Dim_SeatType để chuẩn bị cho quá trình merge.



Bước 8. Tương tự ta chọn các cột cho Sort1

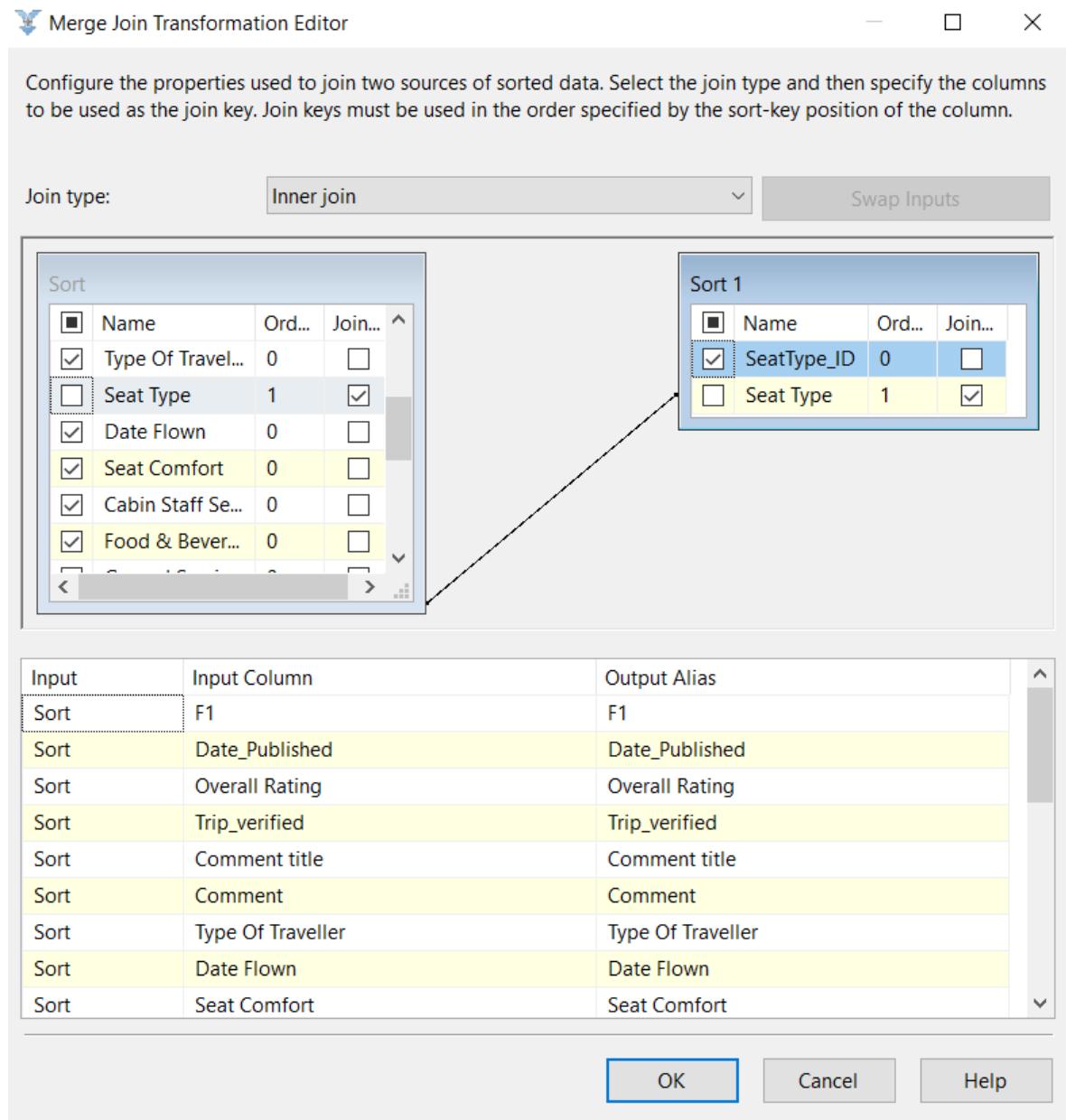


Bước 9. Tạo một **Merge Join** và nối với Sort, tiếp theo ta chọn **Merge Join Left Input** để giữ lại toàn bộ các dòng trong bảng Fact4 bất kể có kết quả khi thực hiện phép kết trái với cột ID của bảng Dim_SeatType hay không.

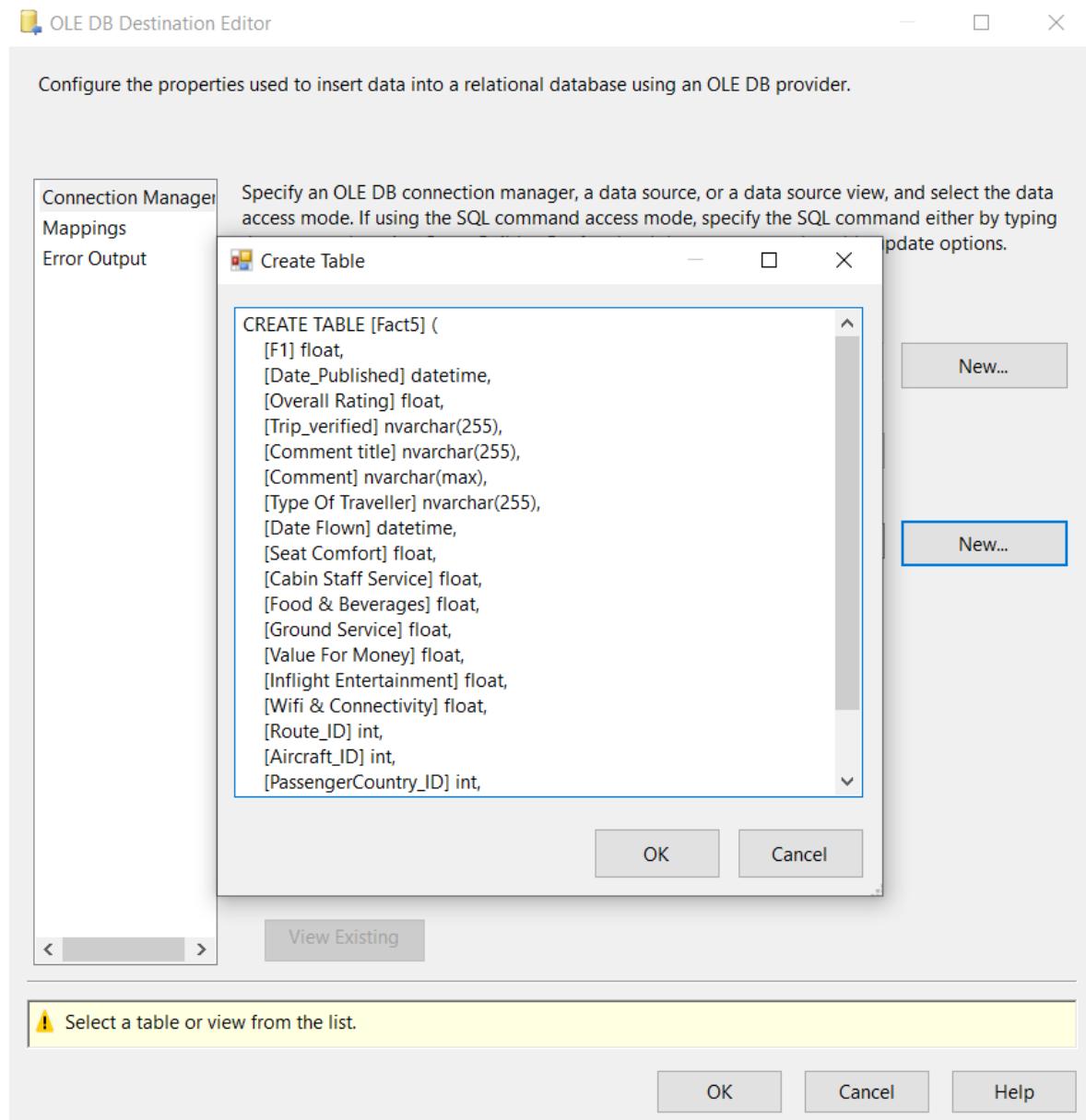


Bước 10.

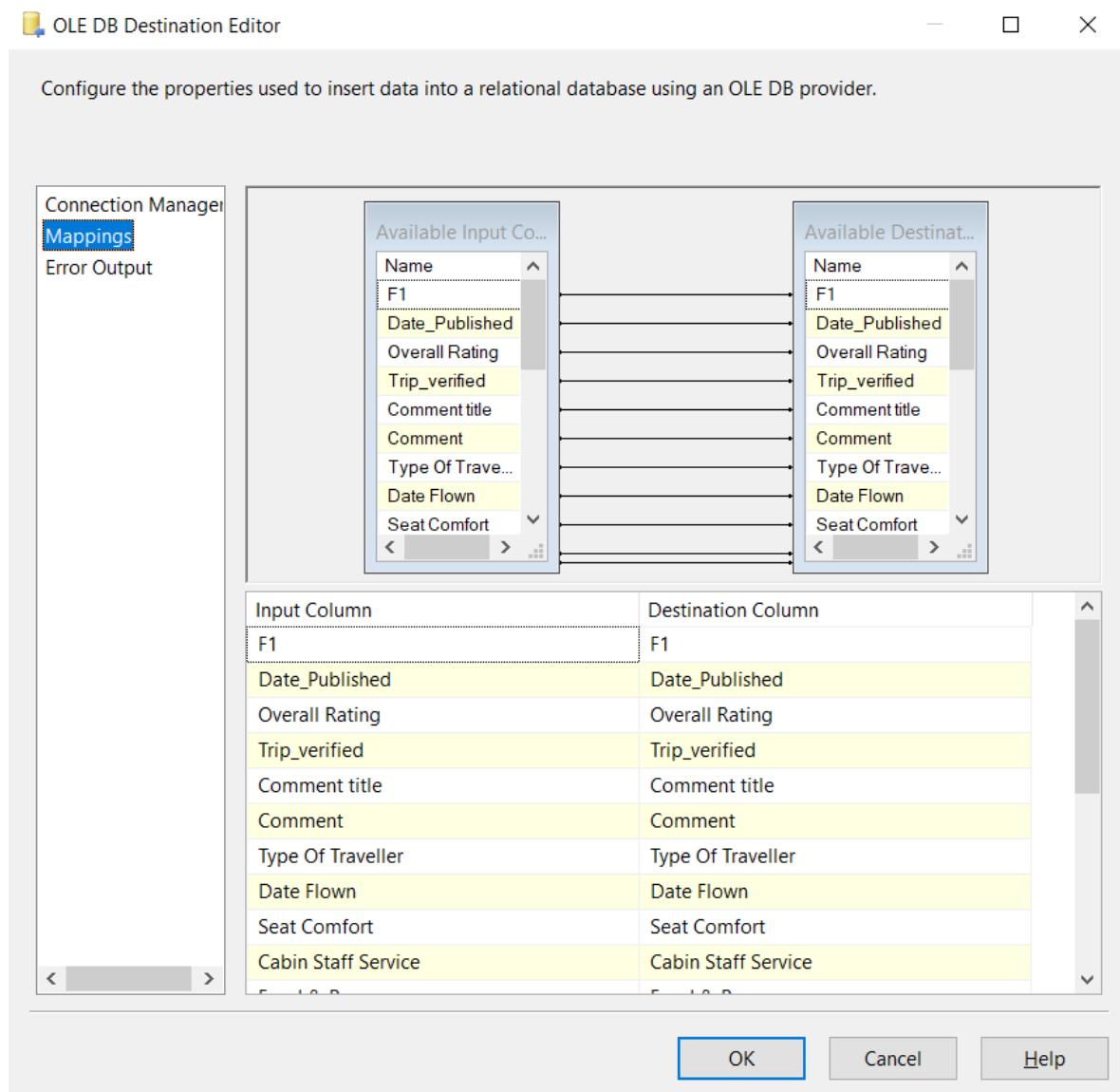
- Chuột phải vào Merge Join và nhấn **Edit**, một hộp thoại merge editor xuất hiện: ở đây ta tick chọn tất cả các cột của Sort nhưng không lấy thuộc tính Seat Type
- Tiếp theo ta chọn Return_ID ở Sort1 để merge vào Fact4
- Kết quả sau khi merge là bảng Fact4 không còn thuộc tính Seat Type và có thêm 1 thuộc tính mới là SeatType_ID



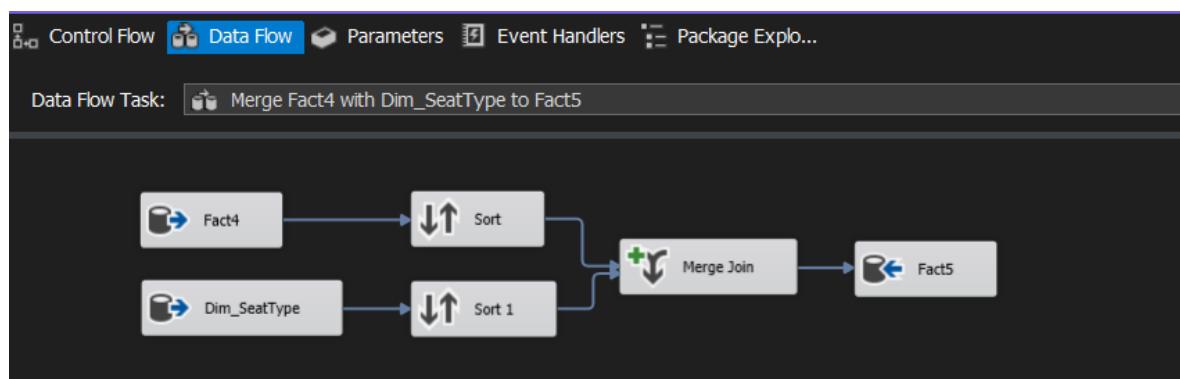
Bước 11. Tạo bảng Fact5 từ một **OLE DB Destination** để chứa tất cả những gì đã merge



Chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

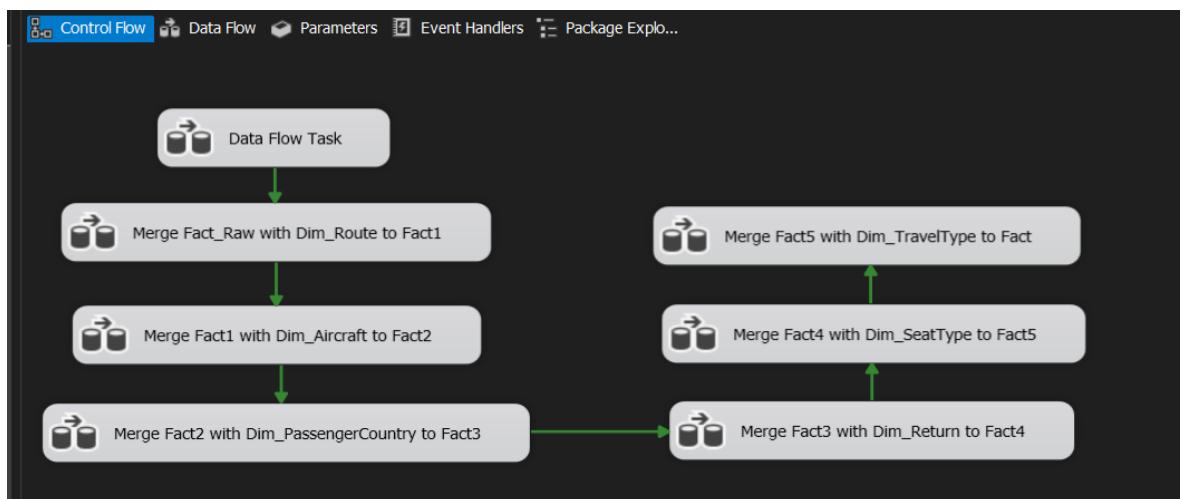


Kết quả:

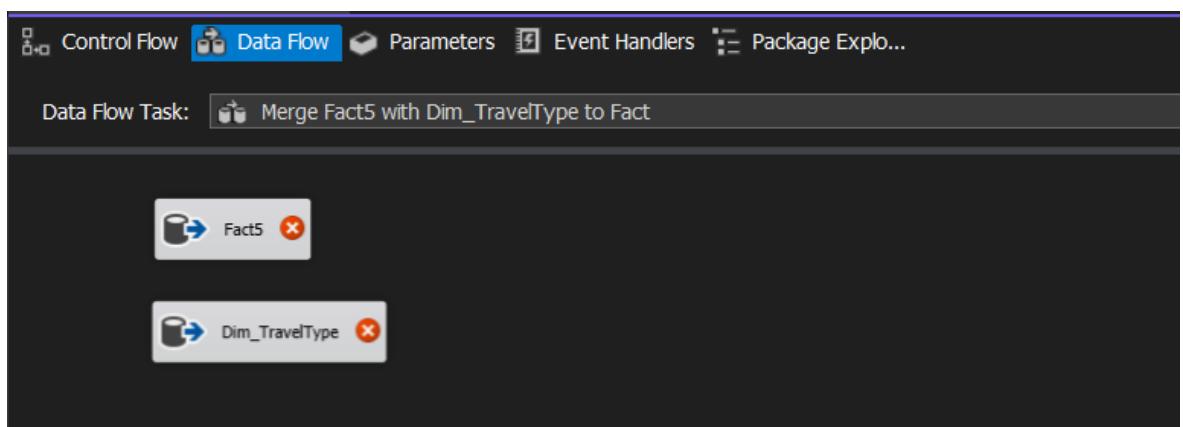


3.3.8.6. Merge Fact5 và Dim_TravelType vào Fact

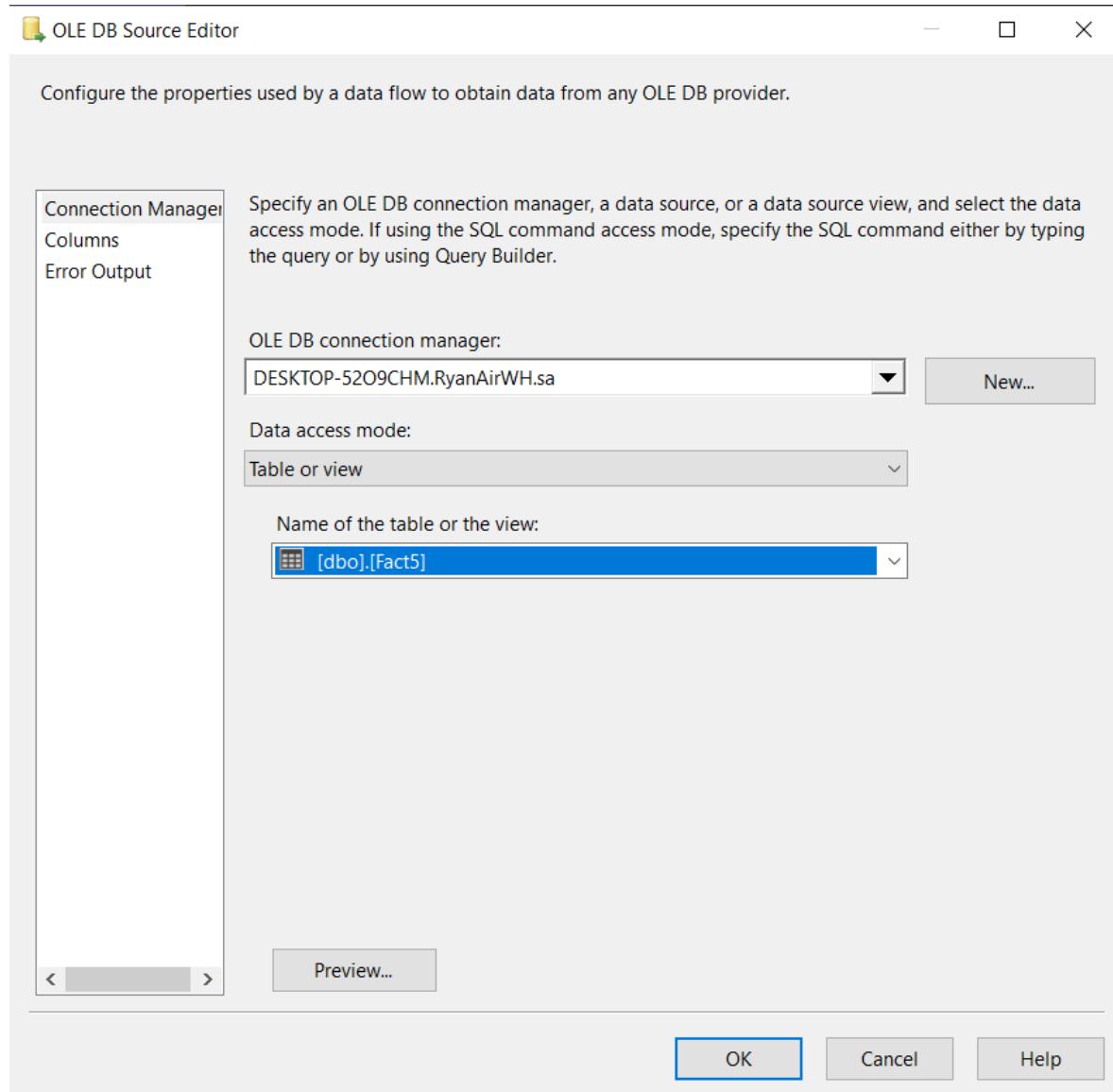
Bước 1. Ở tab Control Flow, tạo **Data Flow Task** và đổi tên Data Flow Task là “Merge Fact5 with Dim_TravelType to Fact”



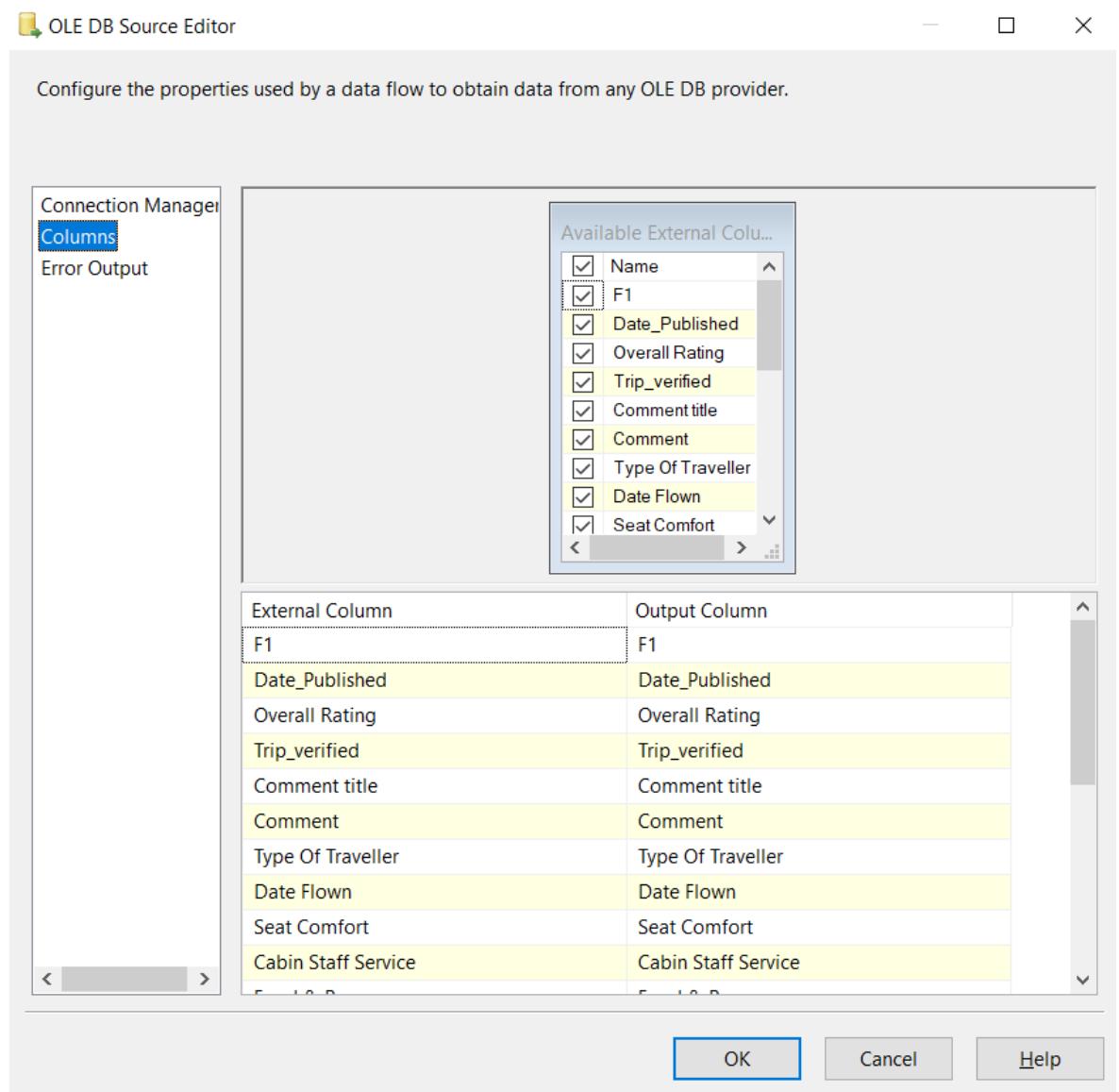
Bước 2. Click chuột phải vào Data Flow Task nói trên và chọn **Edit**, trong tab Data Flow ta tạo 2 OLE DB Source và đổi tên Fact5 và Dim_TravelType



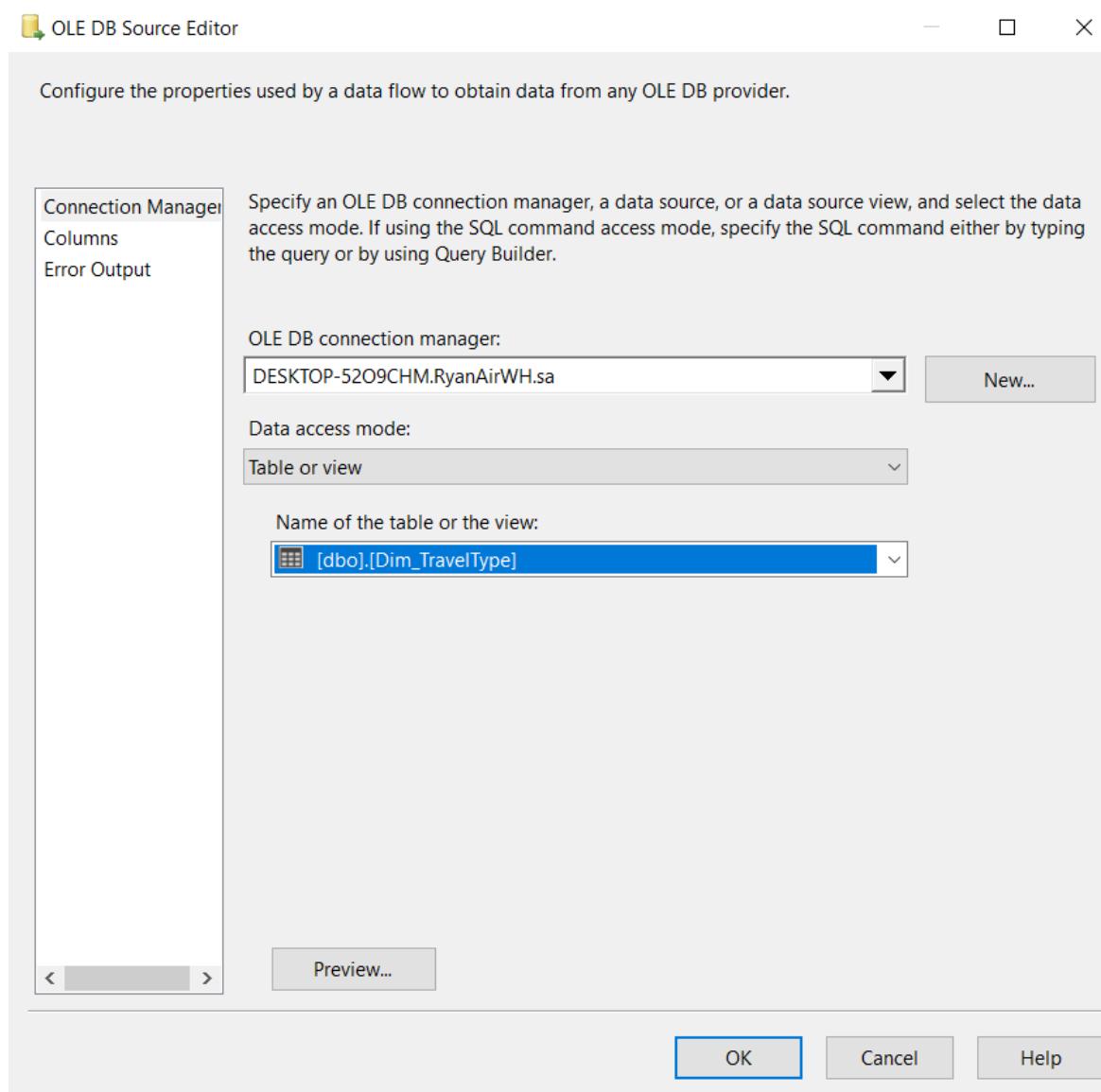
Bước 3. Click chuột phải chọn **Edit**, sau đó chọn bảng Fact5 đã tạo trước đó làm data source cho bảng Fact5 mới này.

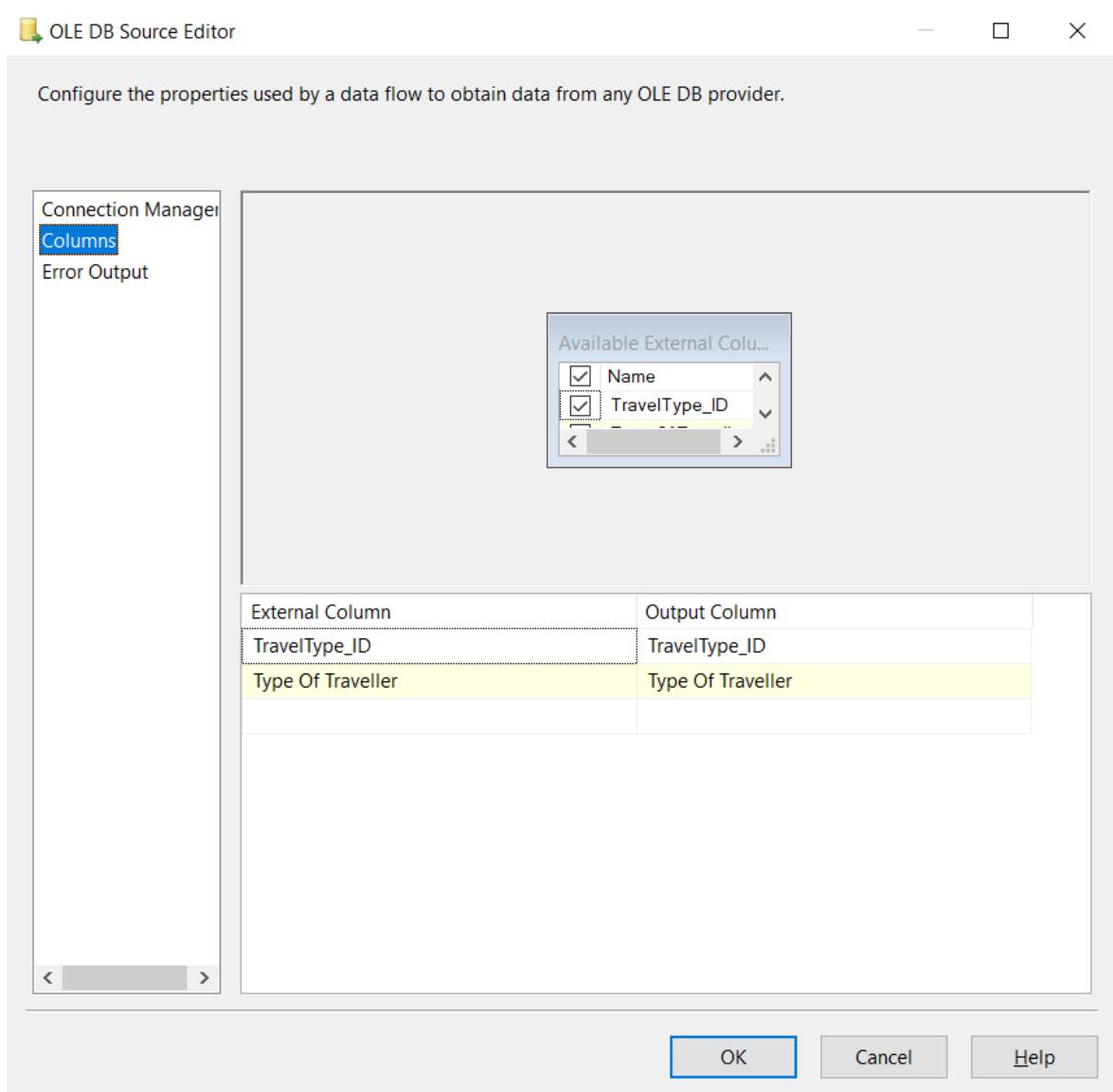


Bước 4. Chọn mục **Columns** để xem xét các cột được ánh xạ. Nhấn **OK**.

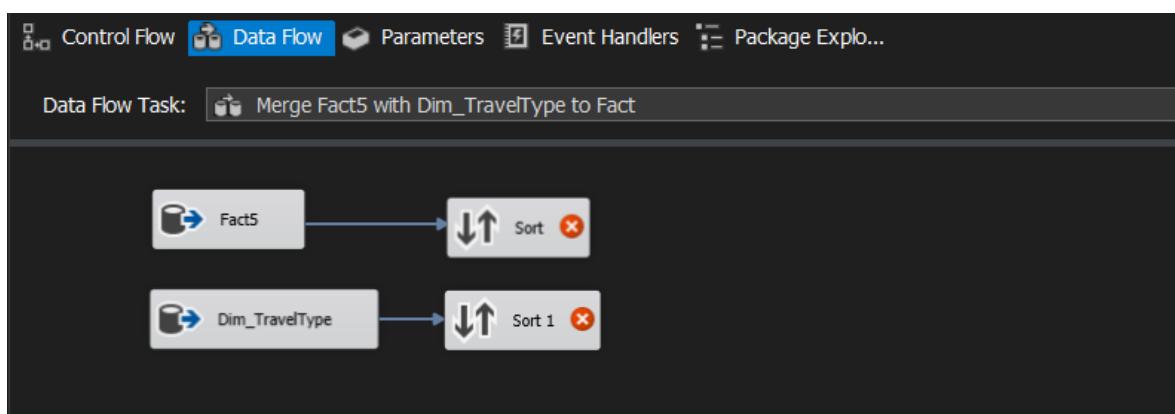


Bước 5. Tương tự thực hiện chọn ánh xạ cột cho Dim_TravelType

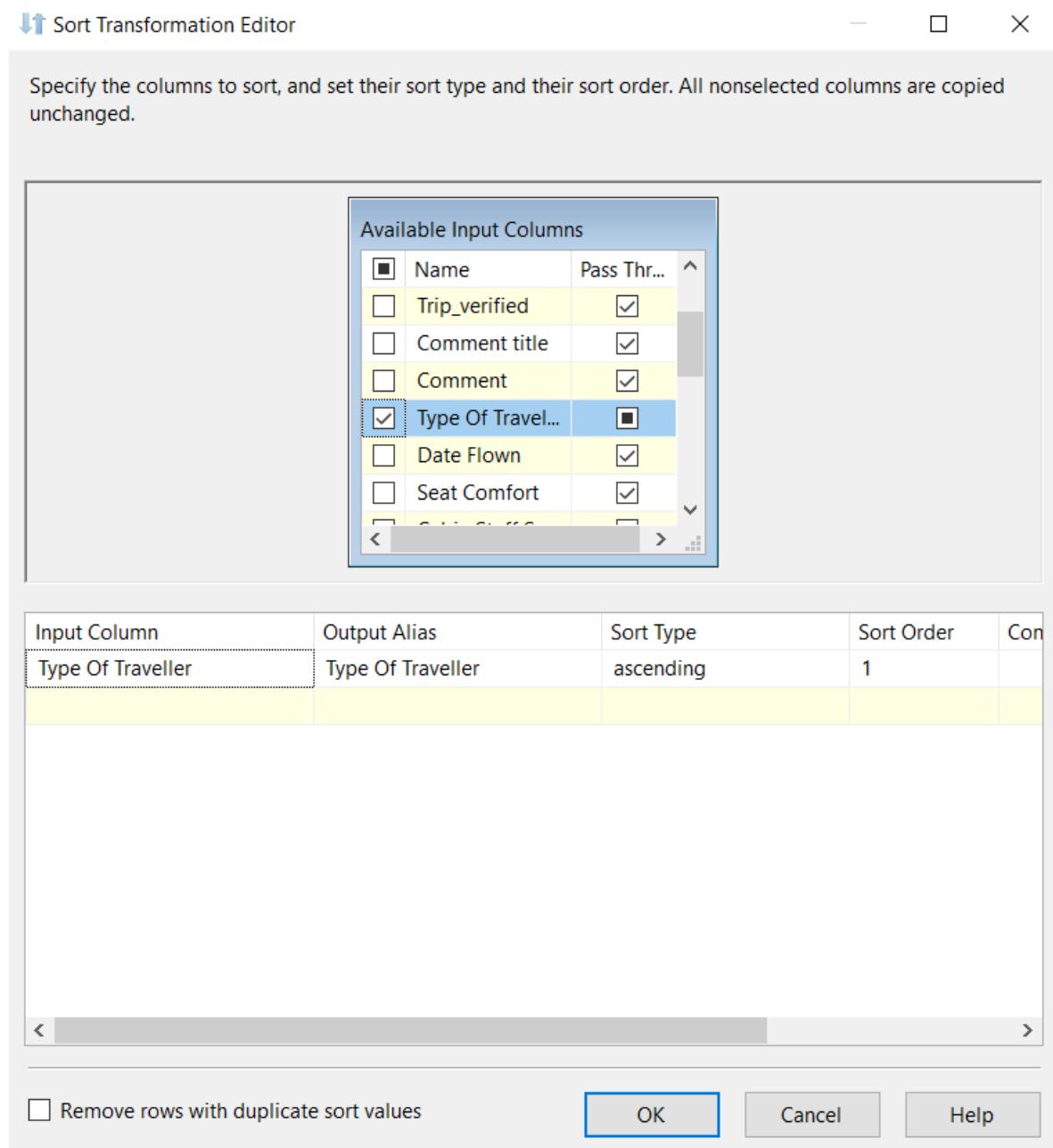




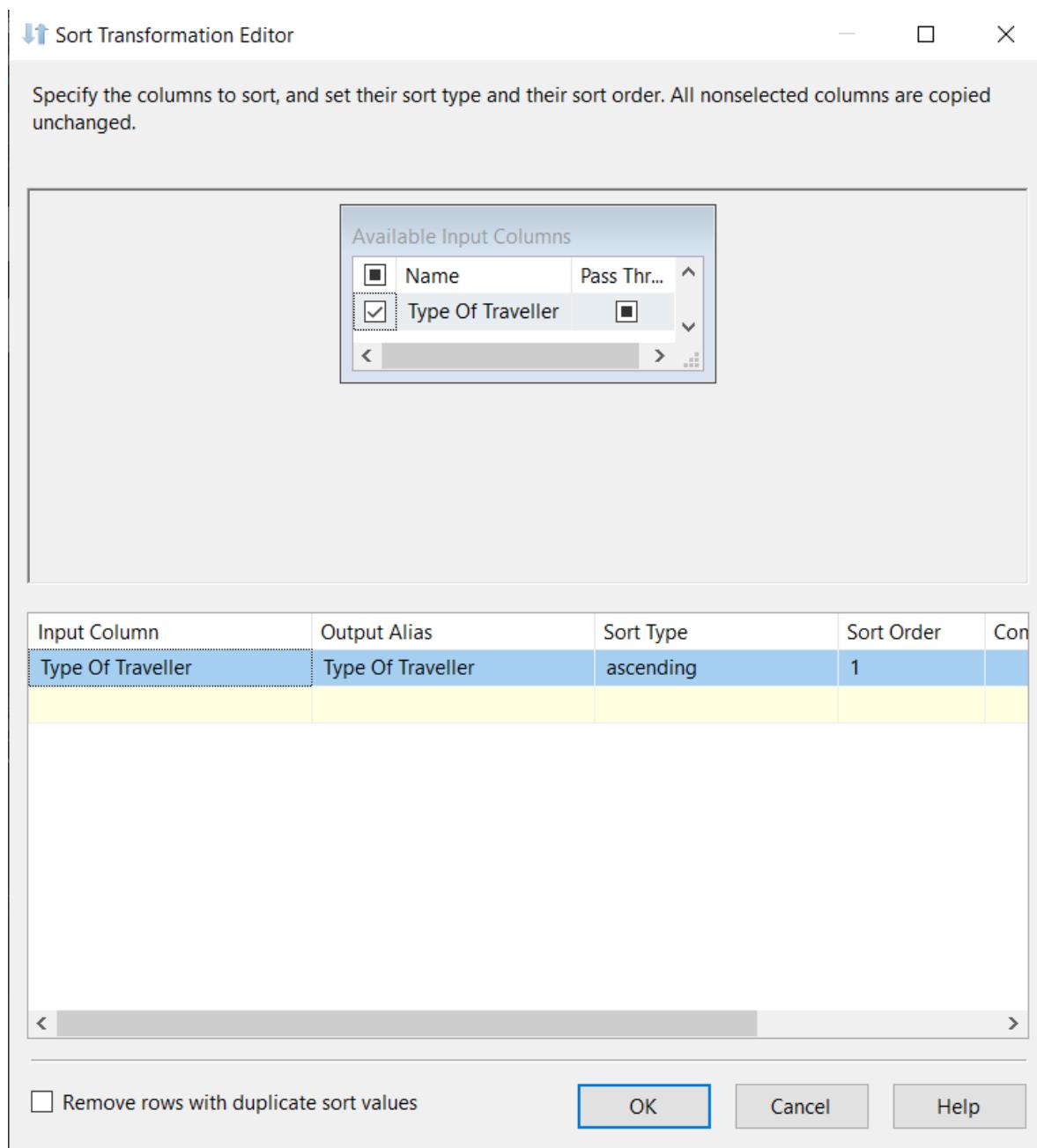
Bước 6. Tạo 2 Sort là Sort và Sort1 tương ứng với mỗi Source.



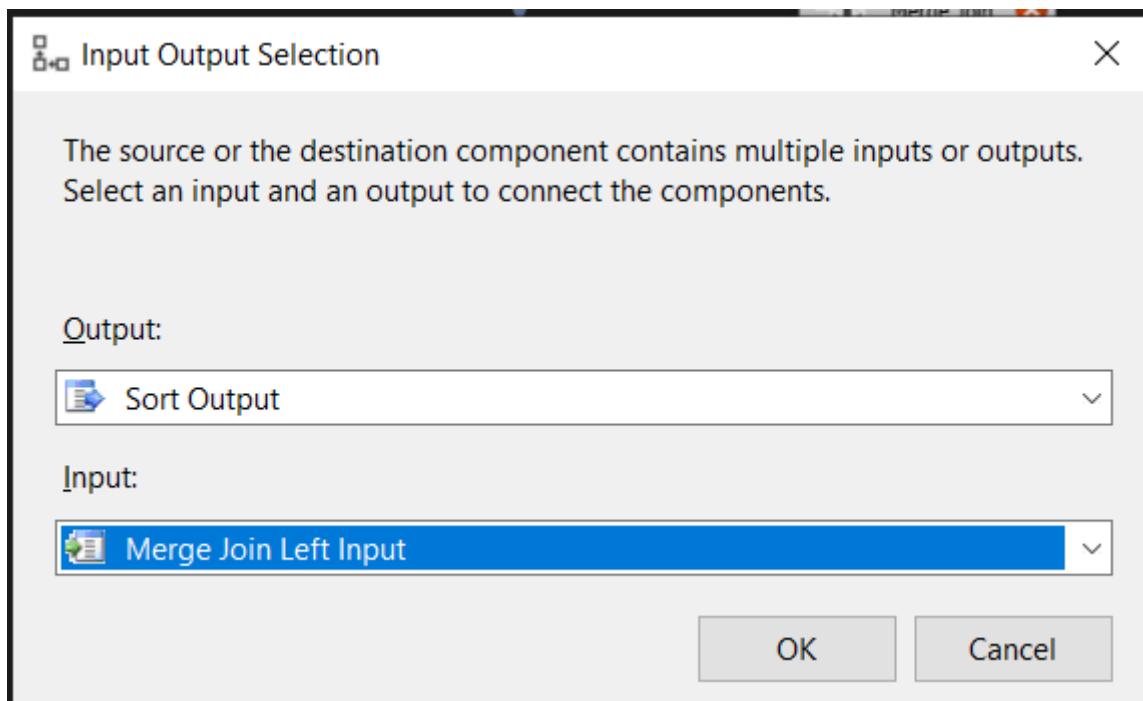
Bước 7. Ở Sort, click chuột phải chọn **Edit** và chọn cột Type Of Traveller theo thứ tự giống với bảng Dim_TravelType để chuẩn bị cho quá trình merge.



Bước 8. Tương tự ta chọn các cột cho Sort1

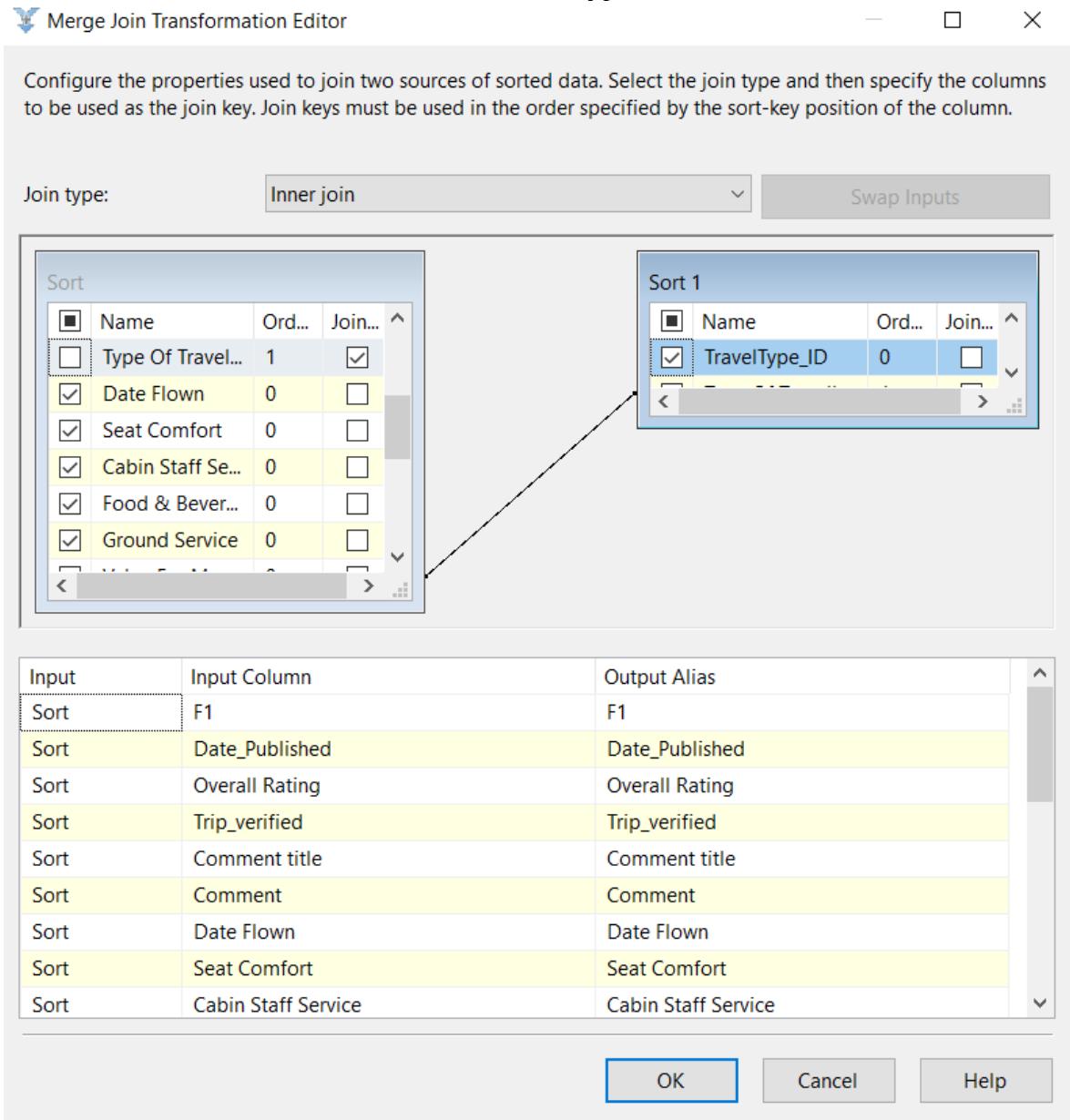


Bước 9. Tạo một **Merge Join** và nối với Sort, tiếp theo ta chọn **Merge Join Left Input** để giữ lại toàn bộ các dòng trong bảng Fact5 bất kể có kết quả khi thực hiện phép kết trái với cột ID của bảng Dim_TravelType hay không.

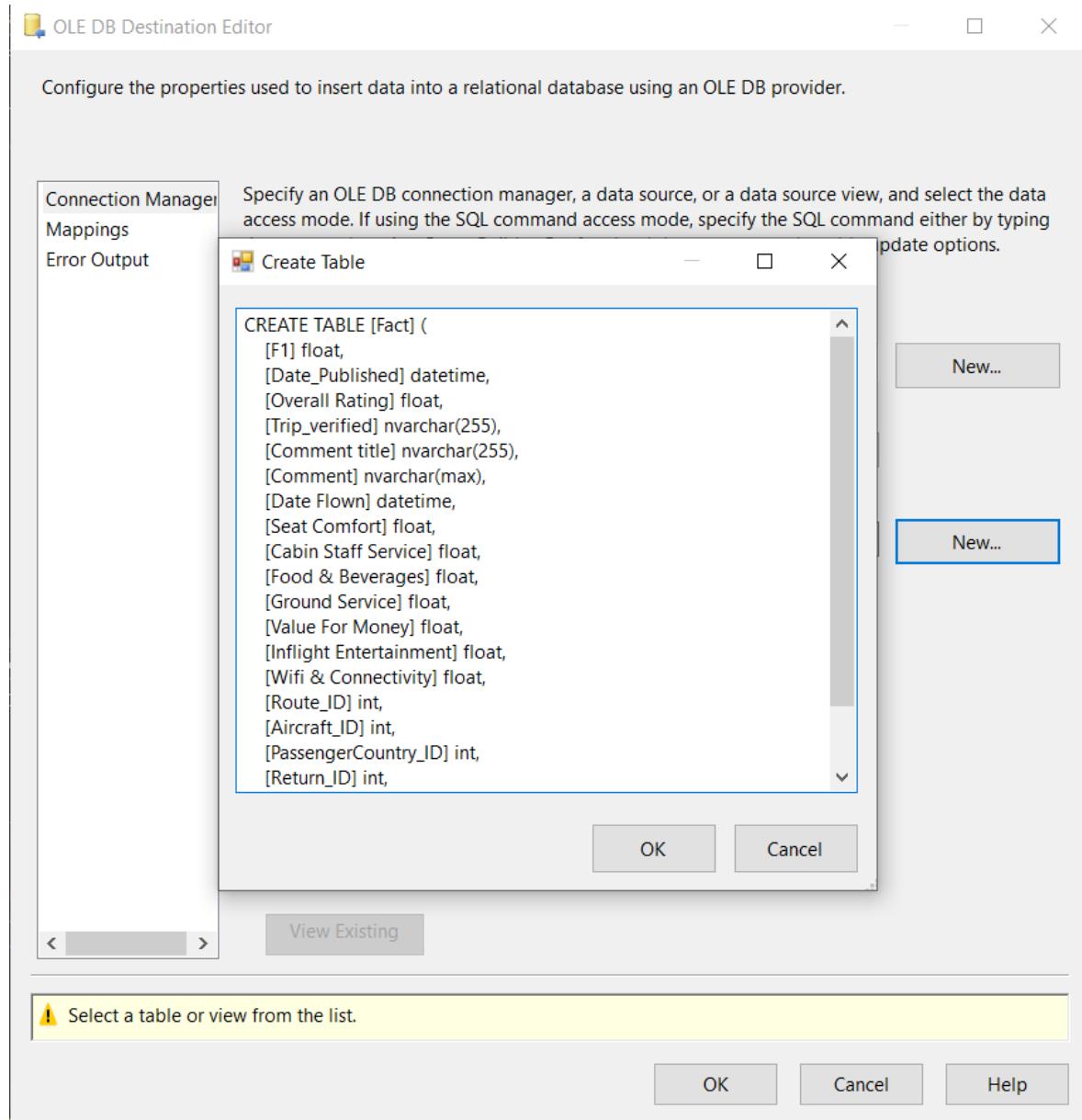


Bước 10.

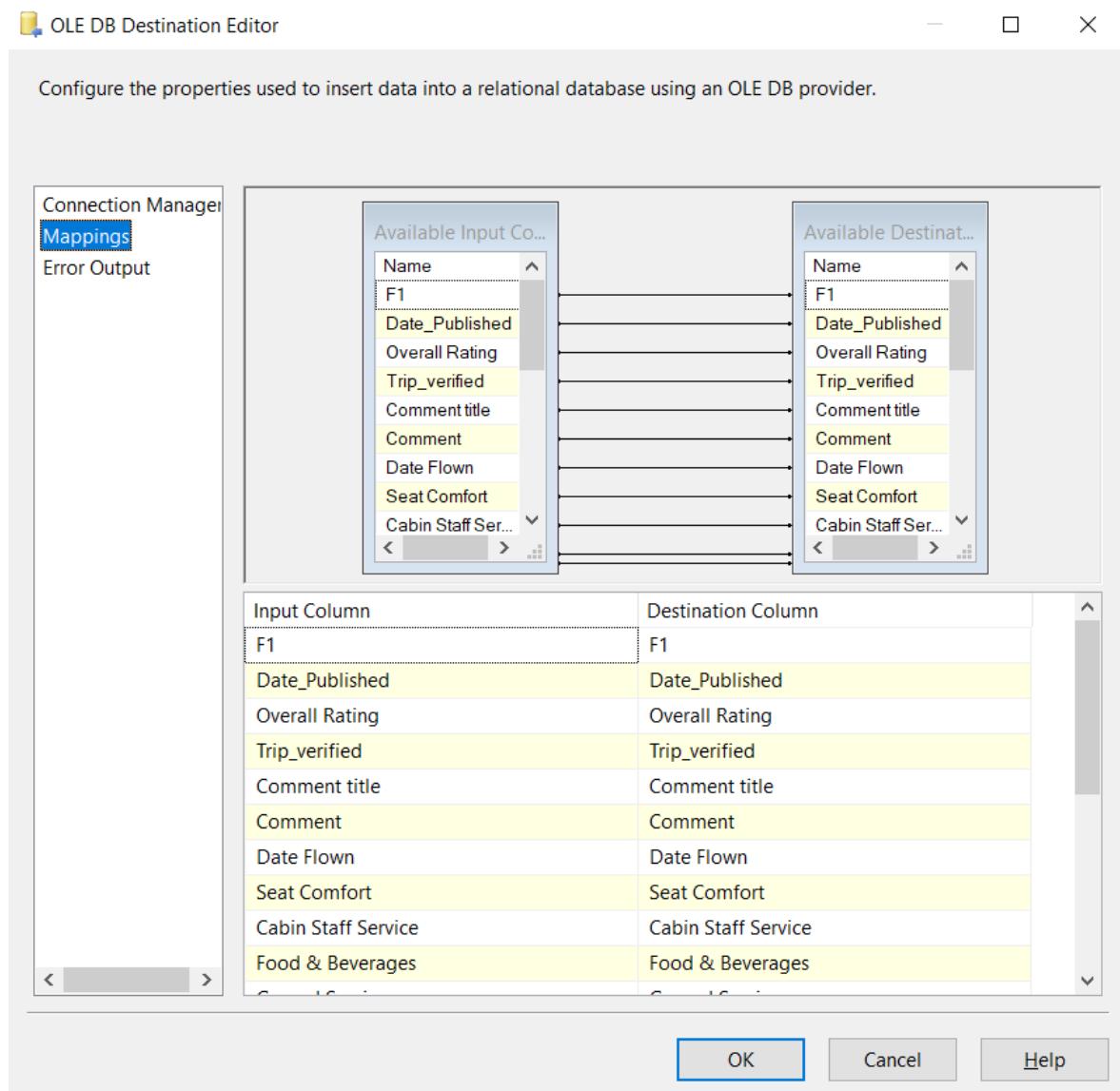
- Chuột phải vào Merge Join và nhấn **Edit**, một hộp thoại merge editor xuất hiện: ở đây ta tick chọn tất cả các cột của Sort nhưng không lấy thuộc tính Type Of Traveller
- Tiếp theo ta chọn TravelType_ID ở Sort1 để merge vào Fact5
- Kết quả sau khi merge là bảng Fact5 không còn thuộc tính Type Of Traveller và có thêm 1 thuộc tính mới là TravelType_ID



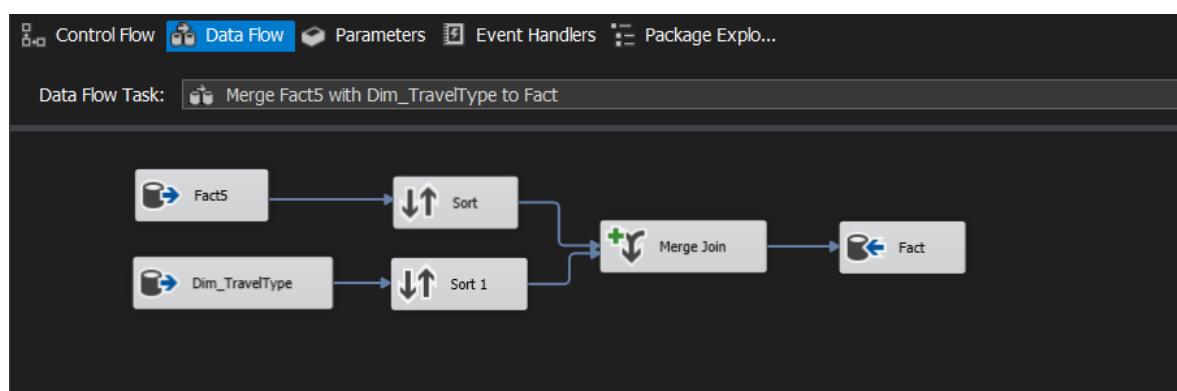
Bước 11. Tạo bảng Fact từ một **OLE DB Destination** để chứa tất cả những gì đã merge



Chọn mục **Mappings** để xem xét việc ánh xạ các cột dữ liệu

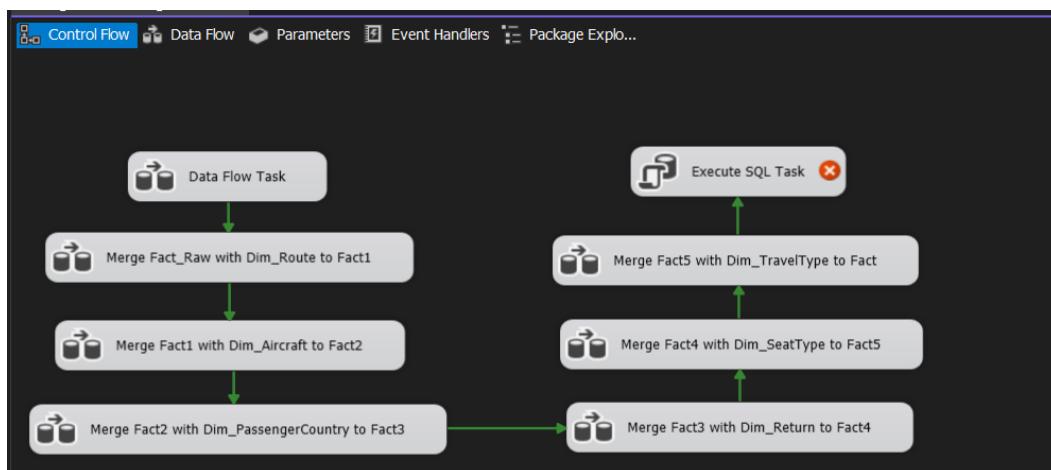


Kết quả:

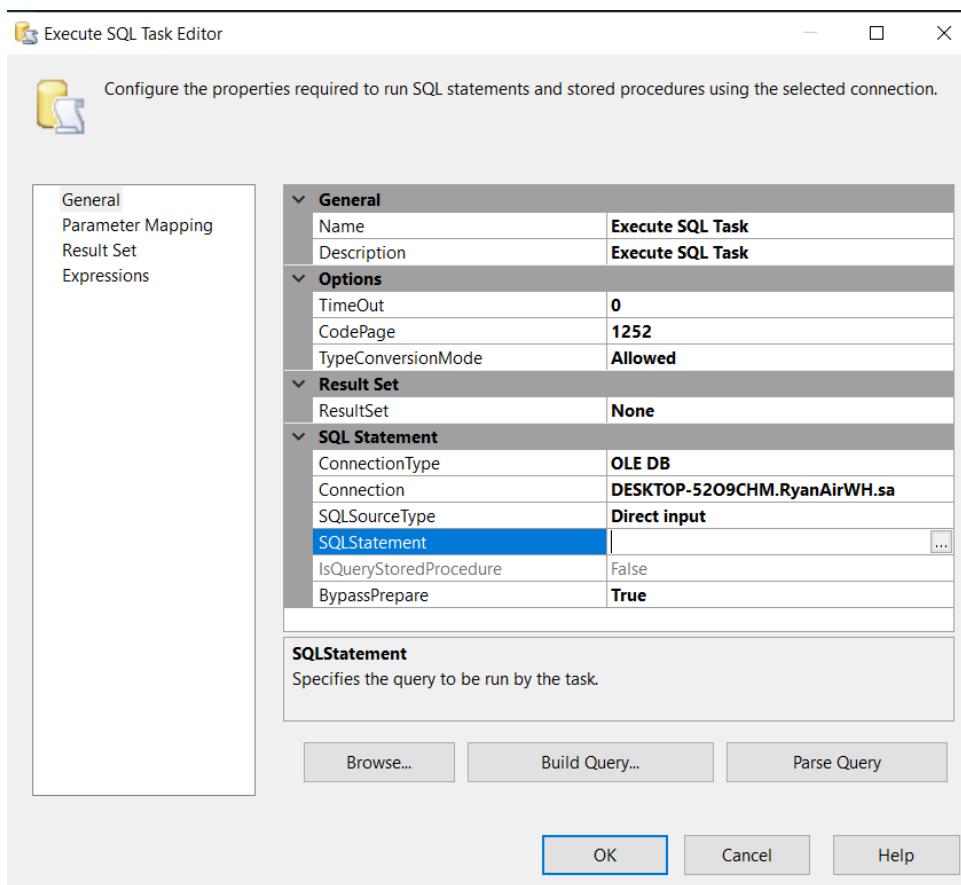


3.3.8.7. Tạo khóa ngoại từ bảng Fact đến các Dimension

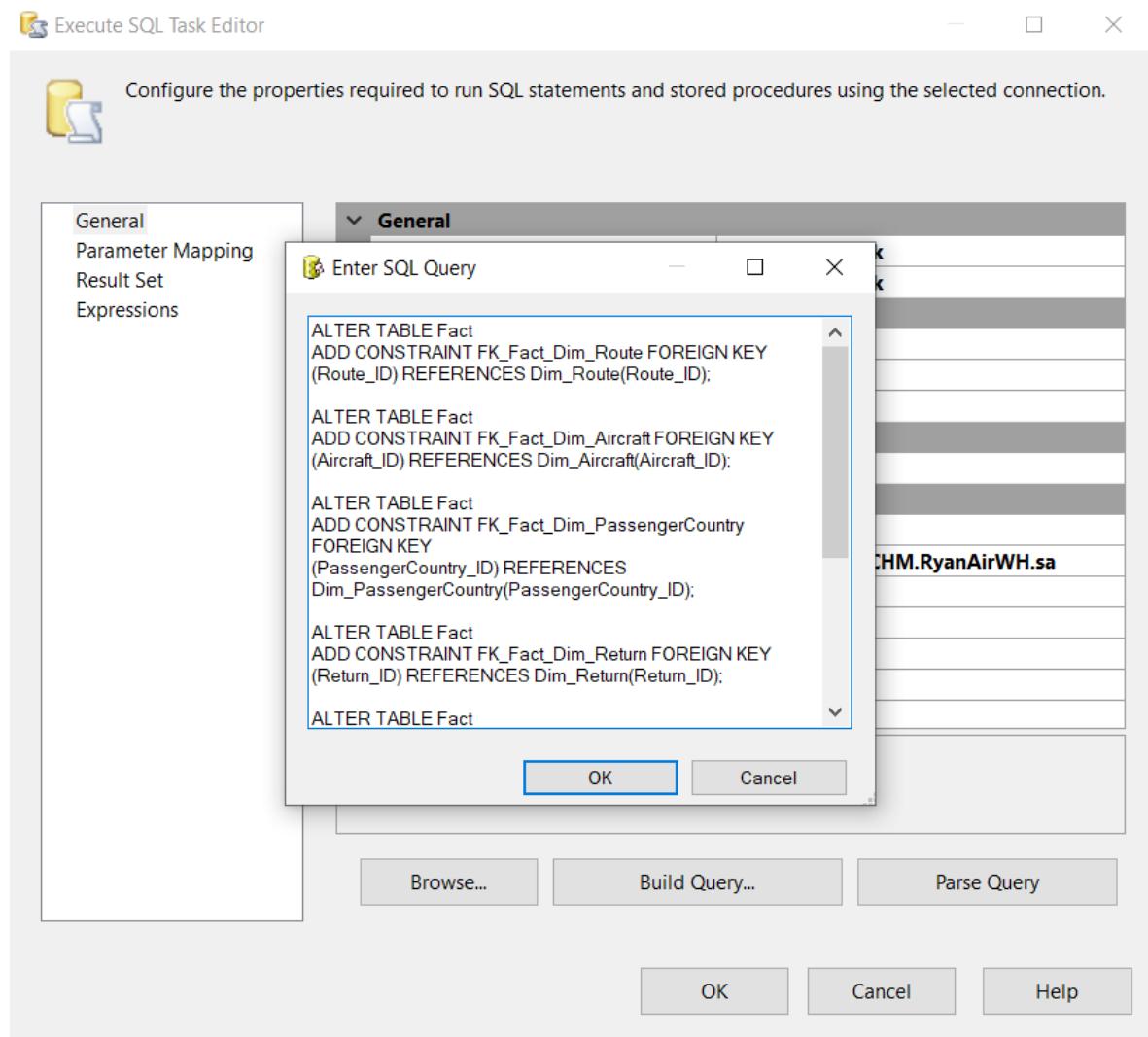
Bước 1. Tạo một **Execute SQL Task** để thực thi các câu lệnh SQL tạo các khóa ngoại từ các Dimension đến bảng Fact



Bước 2. Nhấn chuột phải vào Execute SQL Task này và chọn **Edit**. Ở ô Connection, chọn connection đã thiết lập đến data warehouse trong SQL Server



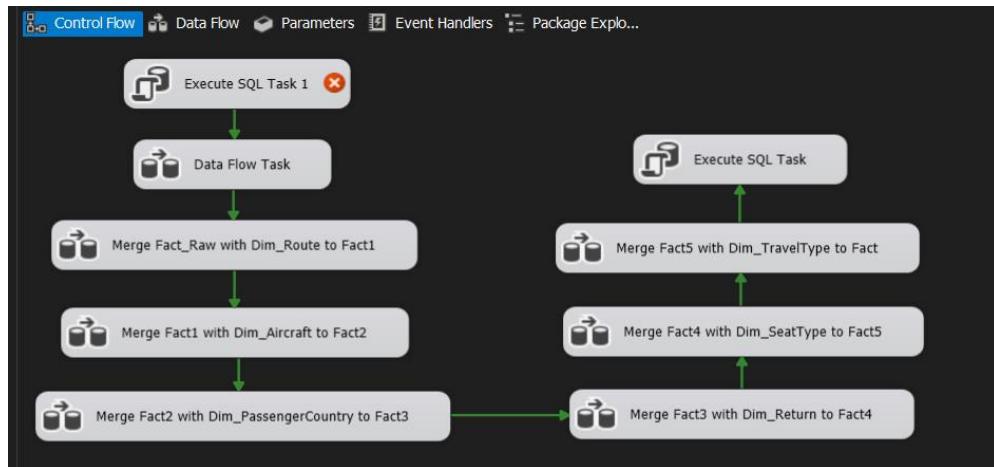
Bước 3. Ở ô SQLStatement, thêm các câu truy vấn SQL thực hiện tạo các khóa ngoại từ các Dimension đến bảng Fact



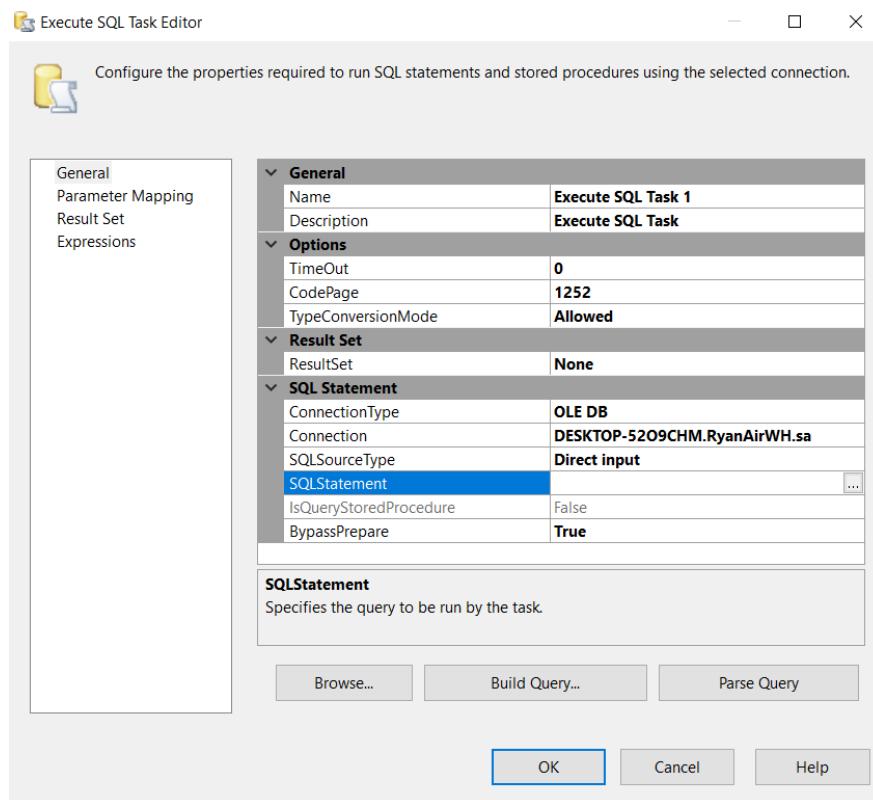
Nhấn **OK** để hoàn tất quá trình.

3.4. Chạy dự án SSIS

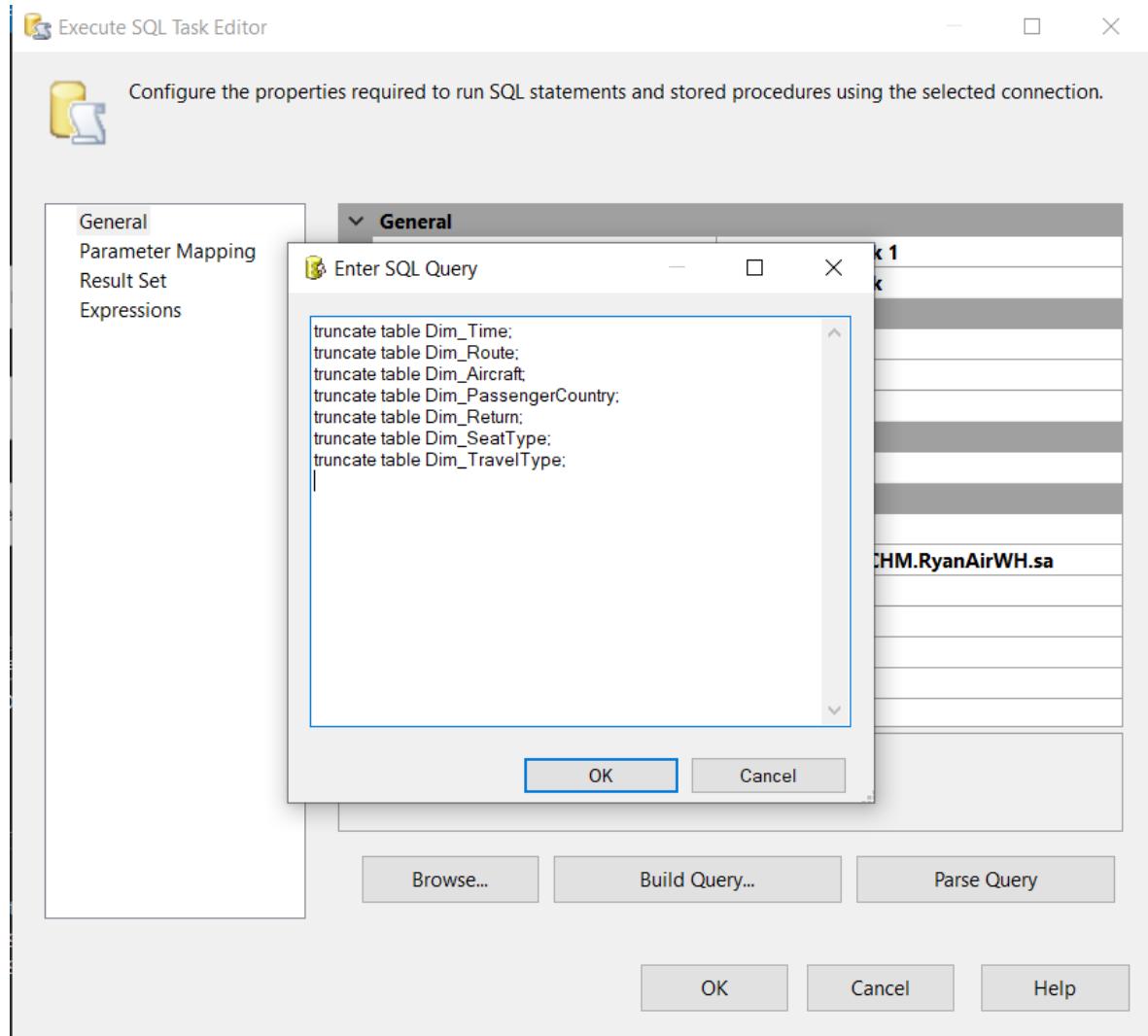
Bước 1. Thêm vào một **Execute SQL Task** nhằm thực hiện nhiệm vụ đảm bảo dữ liệu mới hoàn toàn (không bị chồng chéo dữ liệu cũ) mỗi khi chạy project, trước quá trình chia bảng Fact và các Dimension



Bước 2. Nhấn chuột phải vào Execute SQL Task này và chọn **Edit**. Ở ô Connection, chọn connection đã thiết lập đến data warehouse trong SQL Server

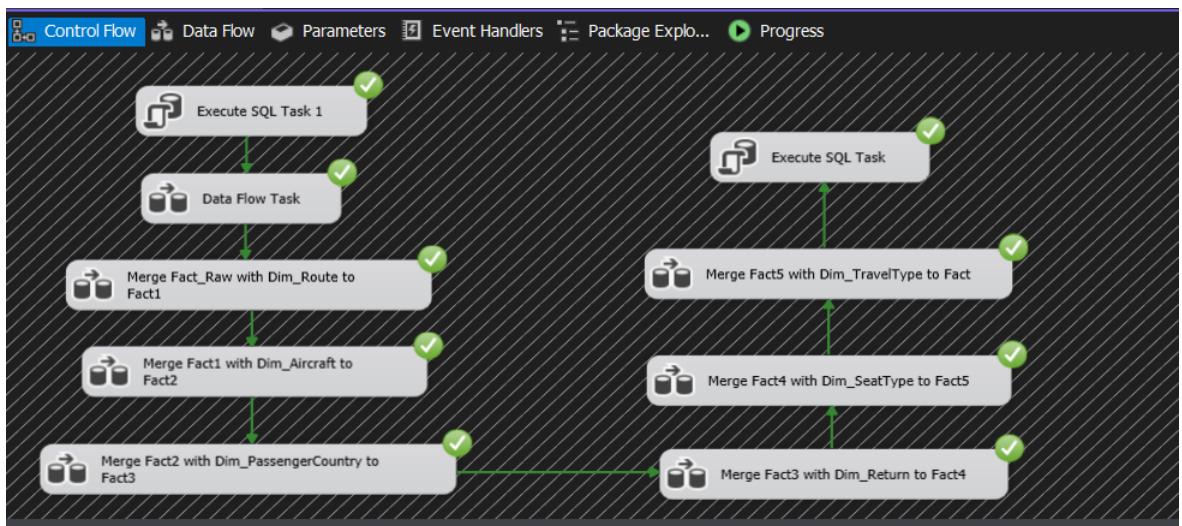


Bước 3. Ở ô SQLStatement, thêm các câu truy vấn SQL thực hiện xóa dữ liệu cũ trong các bảng mỗi khi chạy project.

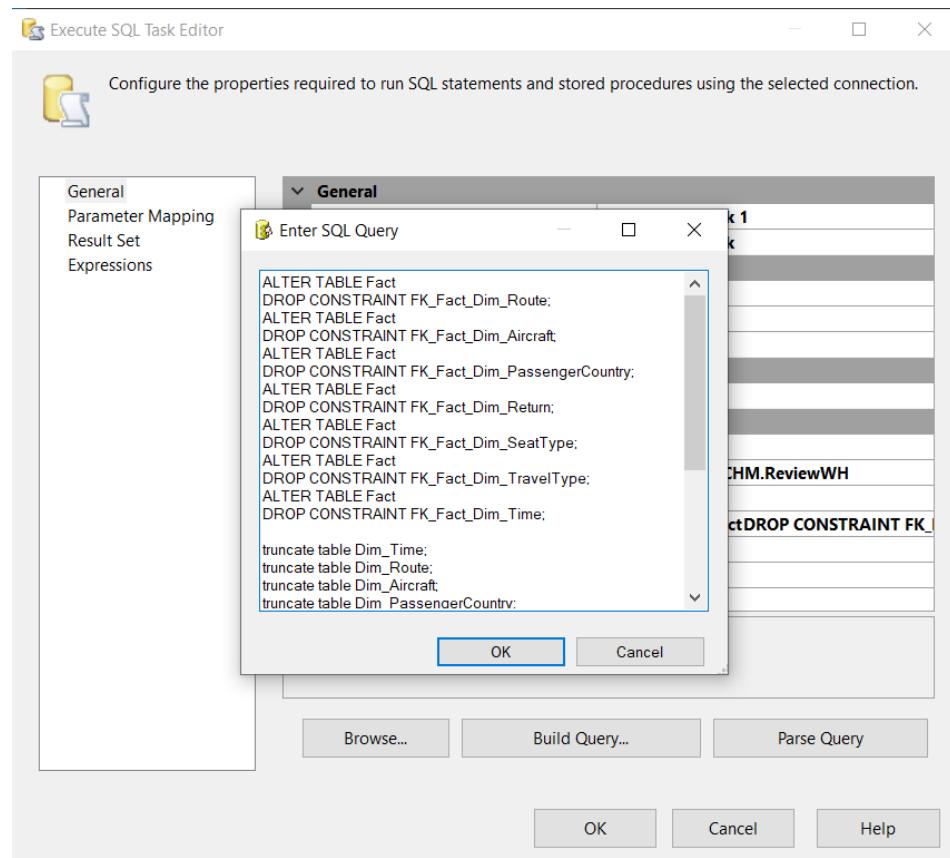


Nhấn OK để hoàn tất quá trình

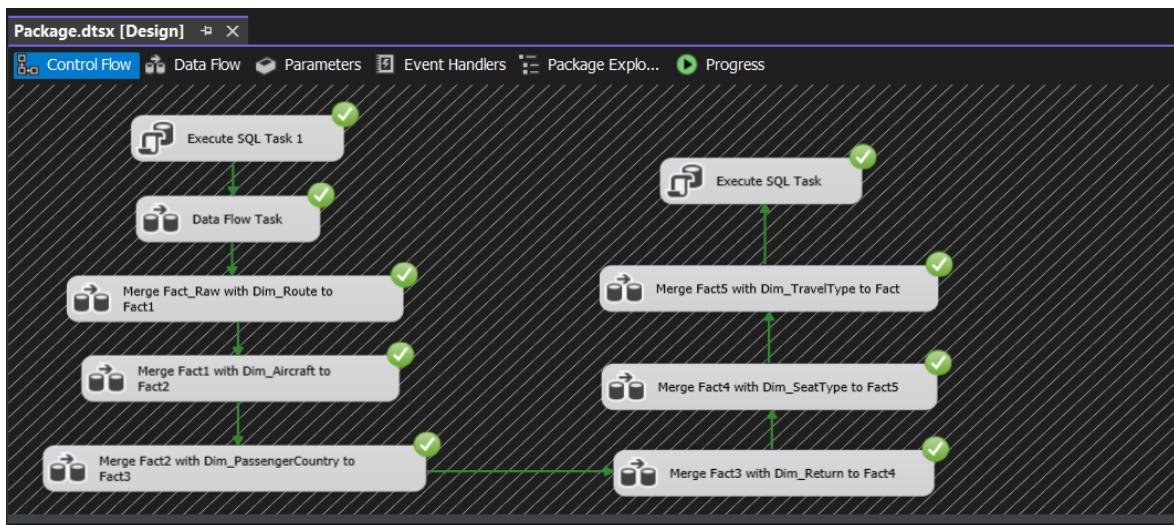
Bước 4. Nhấn nút **Start** trên thanh menu để tiến hành chạy project. Kết quả sau khi chạy:



Bước 5: Sau khi đã chạy project lần đầu thành công, khóa ngoại của bảng Fact tham chiếu đến các Dimension đã được tạo. Ta tiến hành thêm các lệnh SQL để xóa và thêm lại các khóa ngoại vào mỗi lần khởi chạy lại project tiếp theo.



Bước 6. Tiến hành chạy lại project



3.5. Kiểm tra dữ liệu các bảng

Bảng Dim_Route

	Route_ID	Origin	Destination
1	1	NULL	NULL
2	2	Aberdeen	Faro
3	3	AC	MAN
4	4	ACE	BRS
5	5	ACE	LBA
6	6	AGP	BRU
7	7	AGP	MAN
8	8	AGP	PMI
9	9	AGP	STN
10	10	ALC	BHX
11	11	ALC	EDI
12	12	ALC	EMA
13	13	Alicante	Birmingham
14	14	Alicante	Brussels
15	15	Alicante	Dublin
16	16	Alicante	East Midla...
17	17	Alicante	Edinburgh
18	18	Alicante	Gatwick
19	19	Alicante	Kaunas

Bảng Dim_SeatType

	SeatType_ID	Seat Type
1	1	Business Class
2	2	Economy Class
3	3	First Class
4	4	Premium Economy

Bảng Dim_PassengerCountry

	PassengerCountry_ID	Passenger Country
1	1	Albania
2	2	Argentina
3	3	Australia
4	4	Austria
5	5	Belarus
6	6	Belgium
7	7	Brazil
8	8	Brunei
9	9	Bulgaria
10	10	Canada
11	11	Chile
12	12	China
13	13	Costa Rica
14	14	Croatia
...

Bảng Dim_Aircraft

	Aircraft_ID	Aircraft
1	1	737 800
2	2	737300
3	3	A319
4	4	A320
5	5	A340
6	6	A-340
7	7	B733-8022
8	8	B737
9	9	B737 800
10	10	B737-800
11	11	B747
12	12	Boeing
13	13	Boeing 737-800
14	14	Boeing 347

Bảng Dim_TravelType

	TravelType_ID	Type Of Traveller
1	1	Business
2	2	Couple Leisure
3	3	Family Leisure
4	4	Solo Leisure

Bảng Dim_Return

	Return_ID	Recommended
1	1	no
2	2	yes

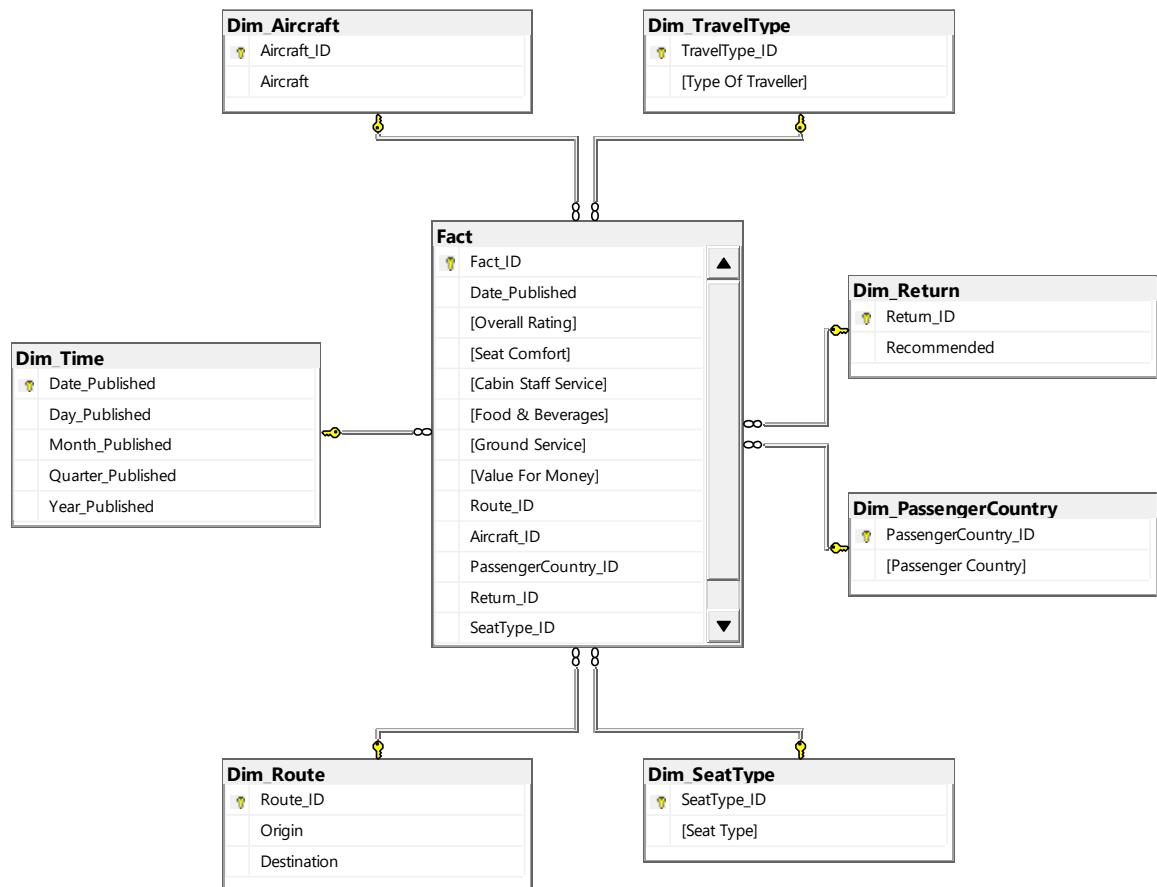
Bảng Dim_Time

	Date_Published	Day_Published	Month_Published	Quarter_Published	Year_Published
1	2012-08-23 00:00:00.000	23	8	3	2012
2	2012-08-28 00:00:00.000	28	8	3	2012
3	2012-08-29 00:00:00.000	29	8	3	2012
4	2012-09-03 00:00:00.000	3	9	3	2012
5	2012-09-05 00:00:00.000	5	9	3	2012
6	2012-09-06 00:00:00.000	6	9	3	2012
7	2013-08-28 00:00:00.000	28	8	3	2013
8	2013-08-29 00:00:00.000	29	8	3	2013
9	2013-09-01 00:00:00.000	1	9	3	2013
10	2013-09-02 00:00:00.000	2	9	3	2013
11	2013-09-04 00:00:00.000	4	9	3	2013
12	2013-09-05 00:00:00.000	5	9	3	2013
13	2013-09-08 00:00:00.000	8	9	3	2013
14	2013-09-09 00:00:00.000	9	9	3	2013

Bảng Fact

Fact_ID	Date_Published	Overall Rating	Seat Comfort	Cabin Staff Service	Food & Beverages	Ground Service	Value For Money	Route_ID	Aircraft_ID	PassengerCountry_ID	Return_ID
1 0	2024-02-03 00:00:00.000	10	4	5	3	4	4	472	21	71	2
2 1	2024-01-26 00:00:00.000	10	3	5	3	5	5	110	33	71	2
3 2	2024-01-20 00:00:00.000	10	5	5	4	5	5	432	33	71	2
4 3	2024-01-07 00:00:00.000	6	3	2	1	3	3	469	16	71	2
5 4	2024-01-06 00:00:00.000	10	4	5	1	4	5	383	33	32	2
6 5	2024-01-06 00:00:00.000	1	2	2	2	1	1	307	33	17	1
7 6	2024-01-03 00:00:00.000	5	2	5	2	1	1	1085	16	71	2
8 7	2024-01-03 00:00:00.000	1	2	2	1	1	1	106	33	3	1
9 8	2023-12-25 00:00:00.000	1	1	2	1	1	1	416	33	71	1
10 9	2023-12-08 00:00:00.000	1	1	1	1	1	1	296	33	22	1
11 10	2023-12-06 00:00:00.000	8	3	5	1	3	5	582	42	1	2
12 11	2023-12-04 00:00:00.000	1	1	1	1	1	1	939	33	58	1
13 12	2023-12-04 00:00:00.000	3	3	2	1	2	3	640	33	53	1

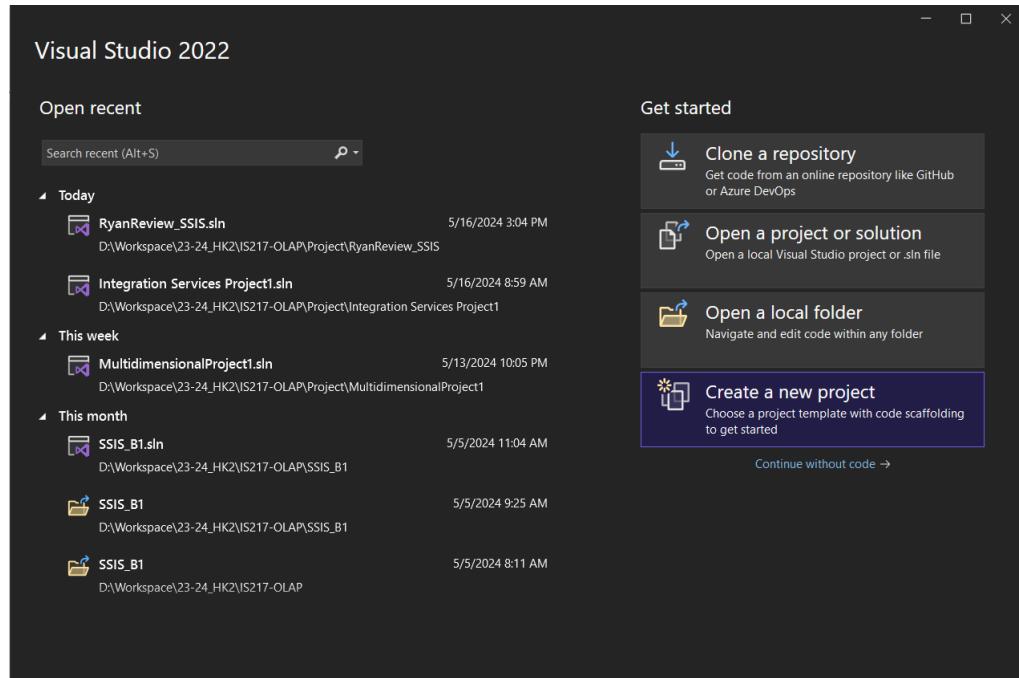
3.6. Lược đồ quan hệ khi hoàn thành



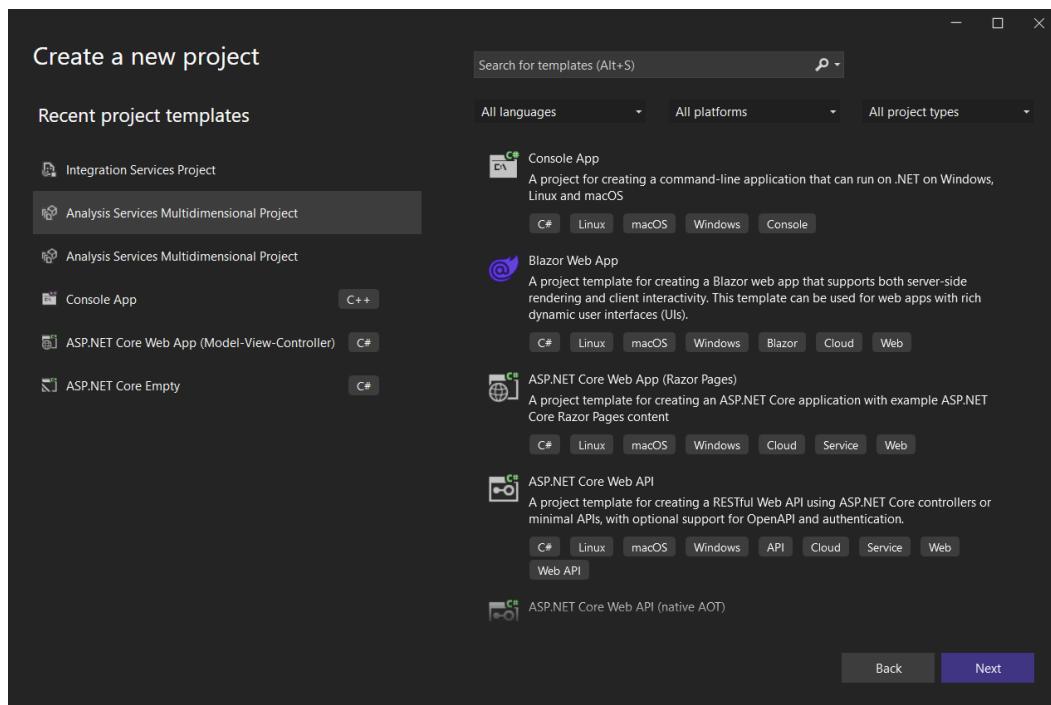
Chương 4. PHÂN TÍCH DỮ LIỆU (SSAS)

4.1. Tạo mới Project SSAS

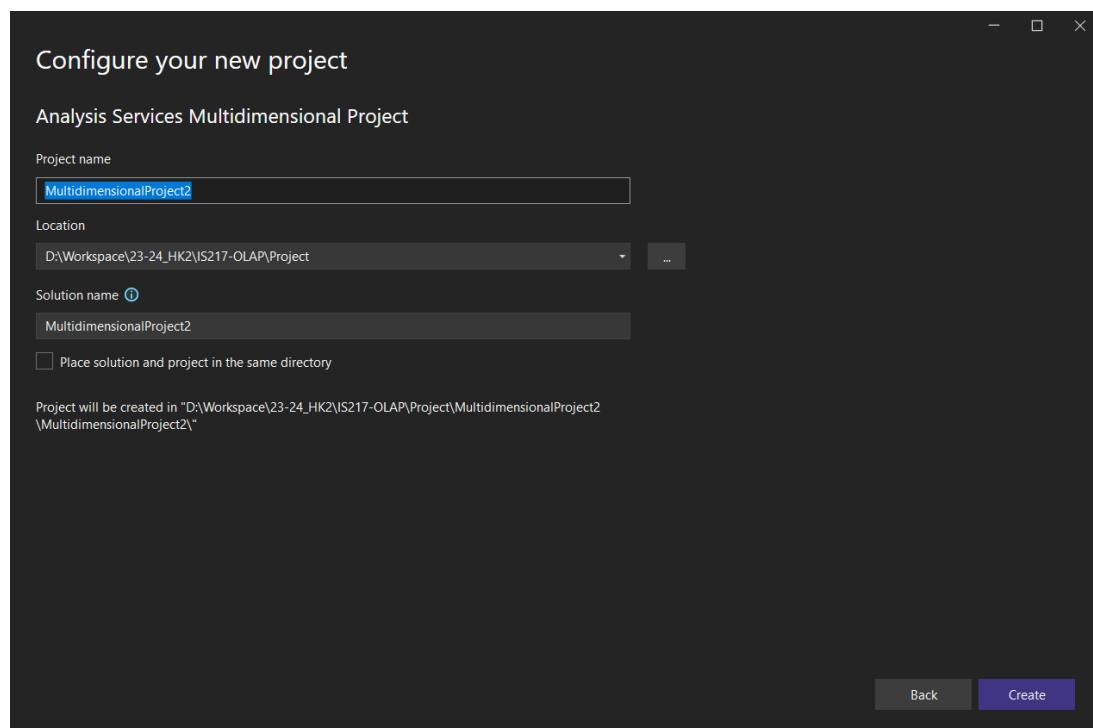
Bước 1: Mở Visual Studio và chọn “Create a new project”



Bước 2: Chọn Analysis Services Multidimensional Project và chọn Next

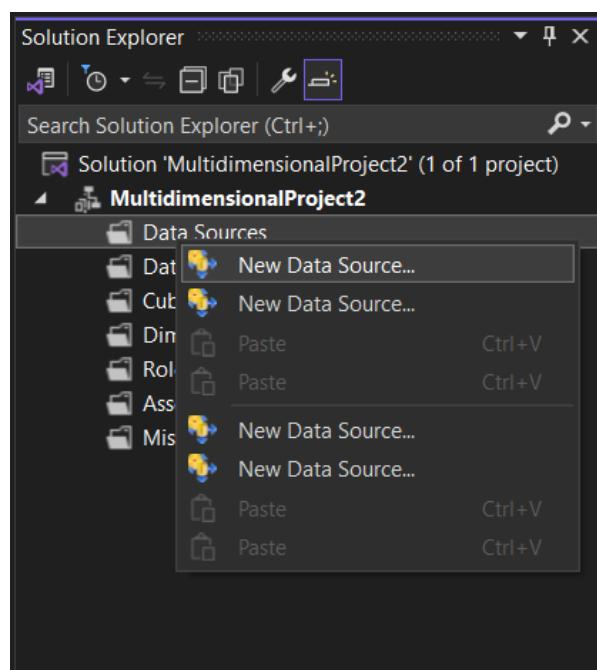


Bước 3: Đặt tên và thiết lập đường dẫn cho Project. Sau đó chọn **Create**.

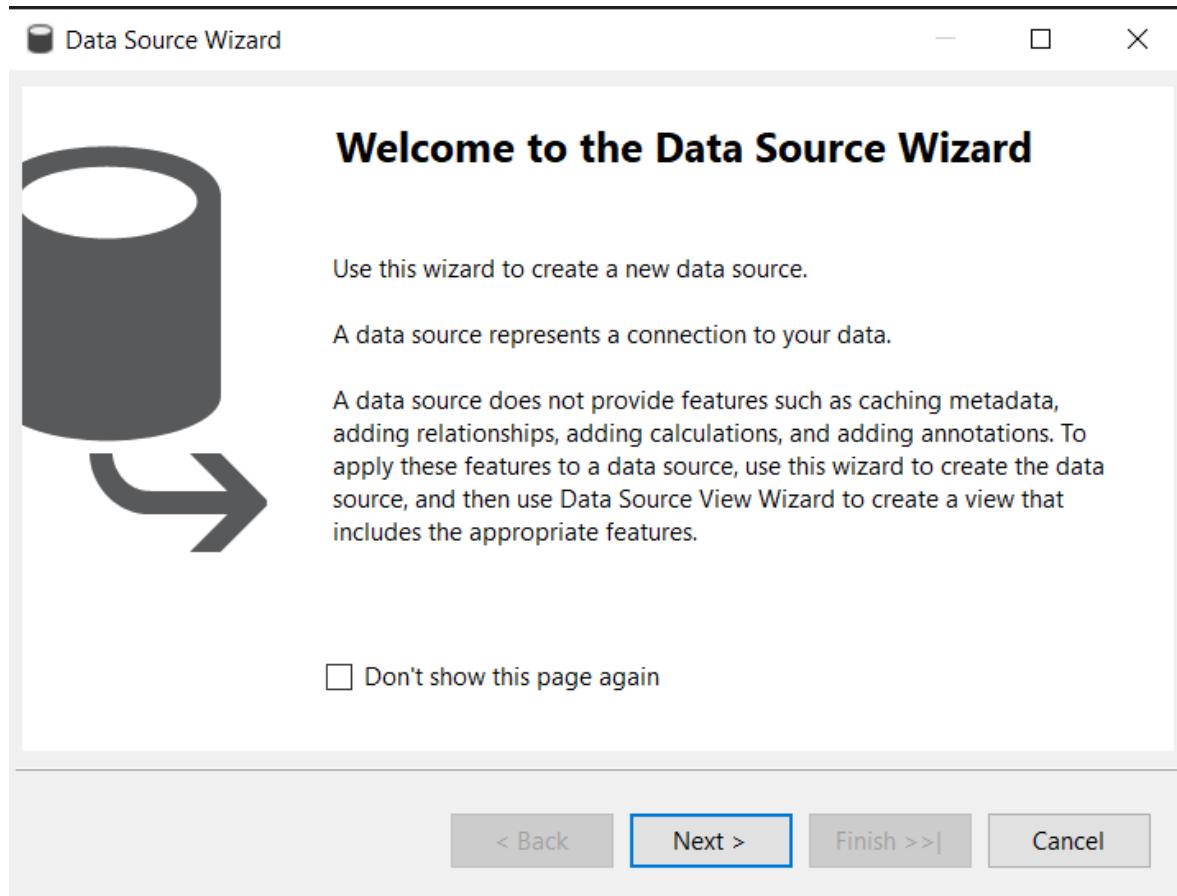


4.2. Xác định dữ liệu nguồn (Data Sources)

Bước 1: Tại Solution Explorer, ta click chuột phải vào thư mục Data Sources và chọn **New Data Source**.

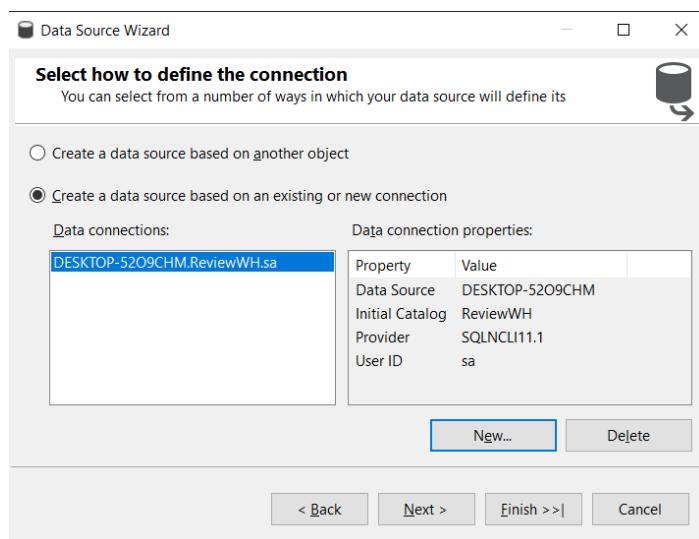


Bước 2: Hộp thoại Data Source Wizard xuất hiện, chọn **Next** để tiếp tục.

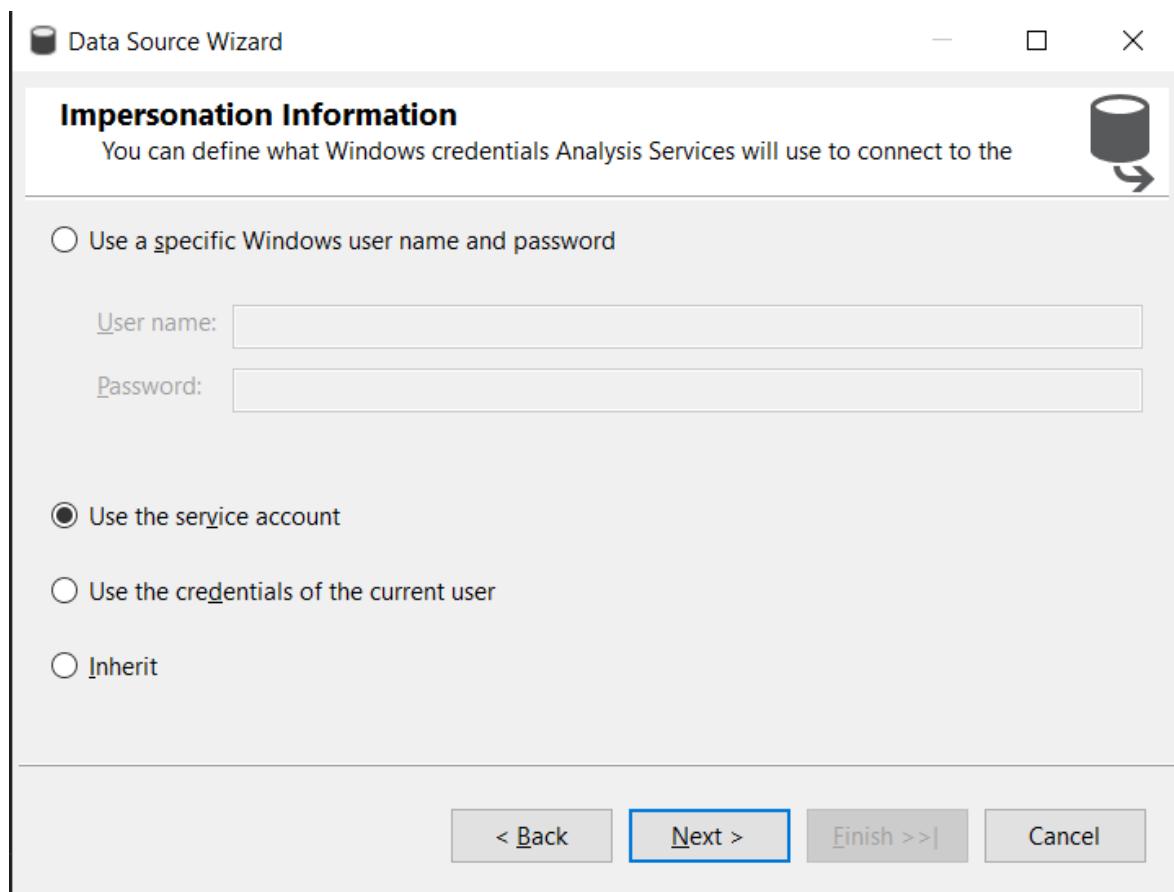


Bước 3: Chọn “Create a data source based on an existing or new connection” sau đó chọn **New...** để tạo kết nối với cơ sở dữ liệu đã được tạo từ quá trình SSIS.

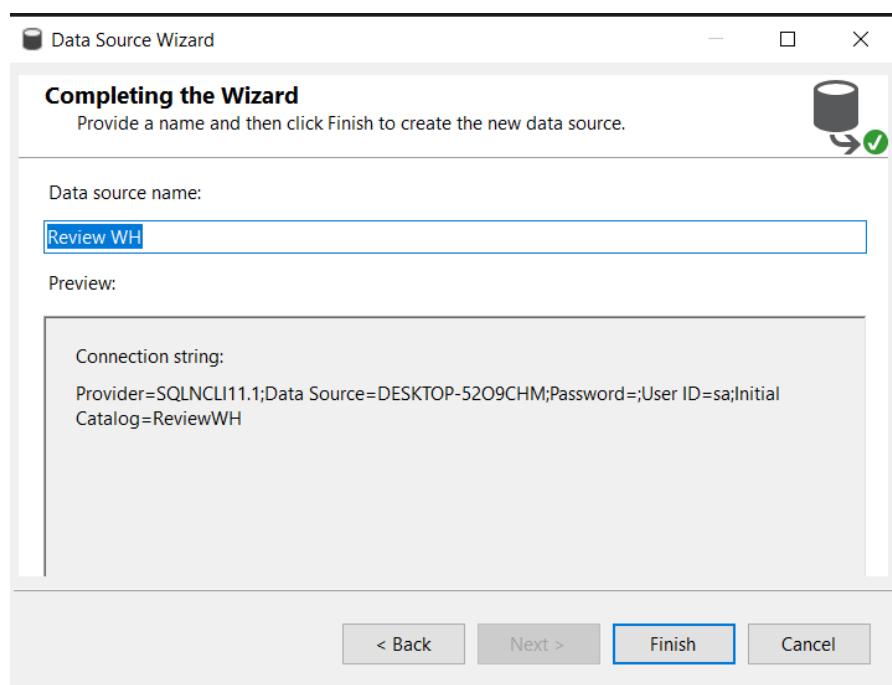
Chọn data source vừa tạo và chọn **Next** để tiếp tục.



Bước 4: Chọn “**Use the service account**”, sau đó chọn **Next** để tiếp tục.

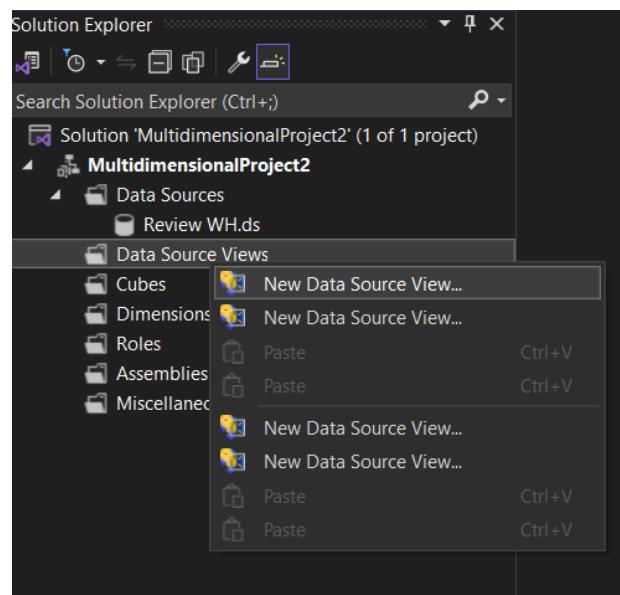


Bước 5: Cuối cùng ta chọn **Finish** để hoàn tất quy trình định nghĩa nguồn dữ liệu.

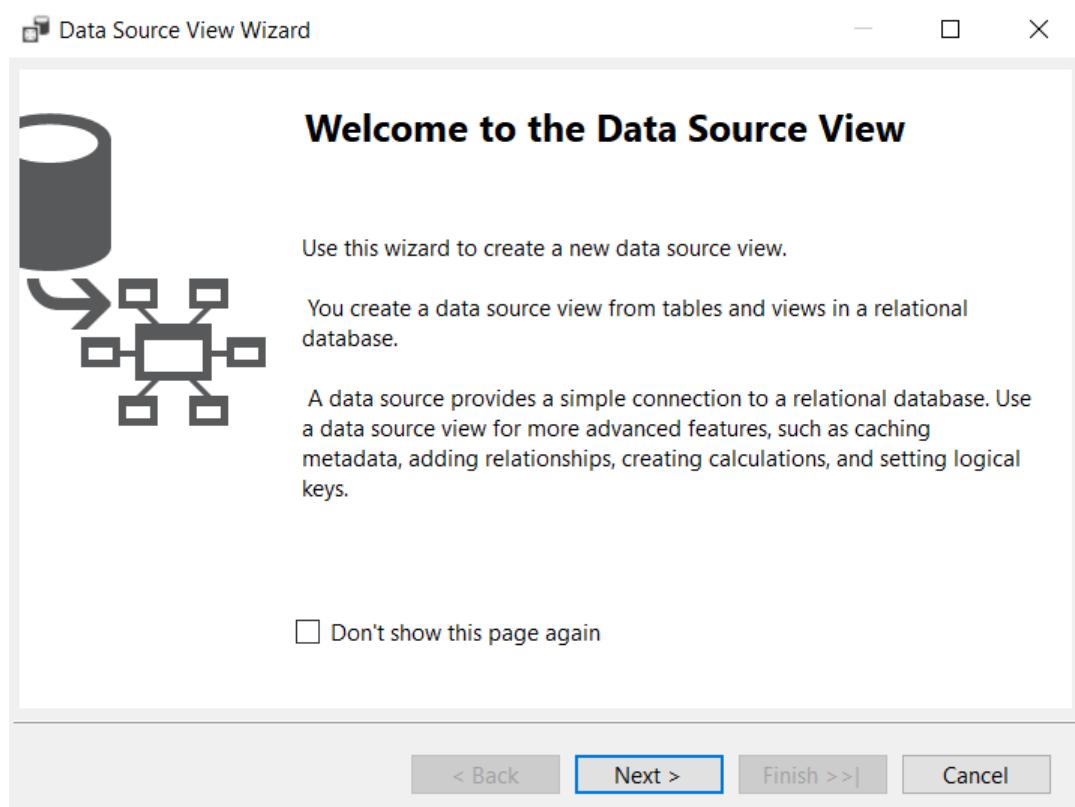


4.3. Xác định khung nhìn dữ liệu nguồn (Data Source Views)

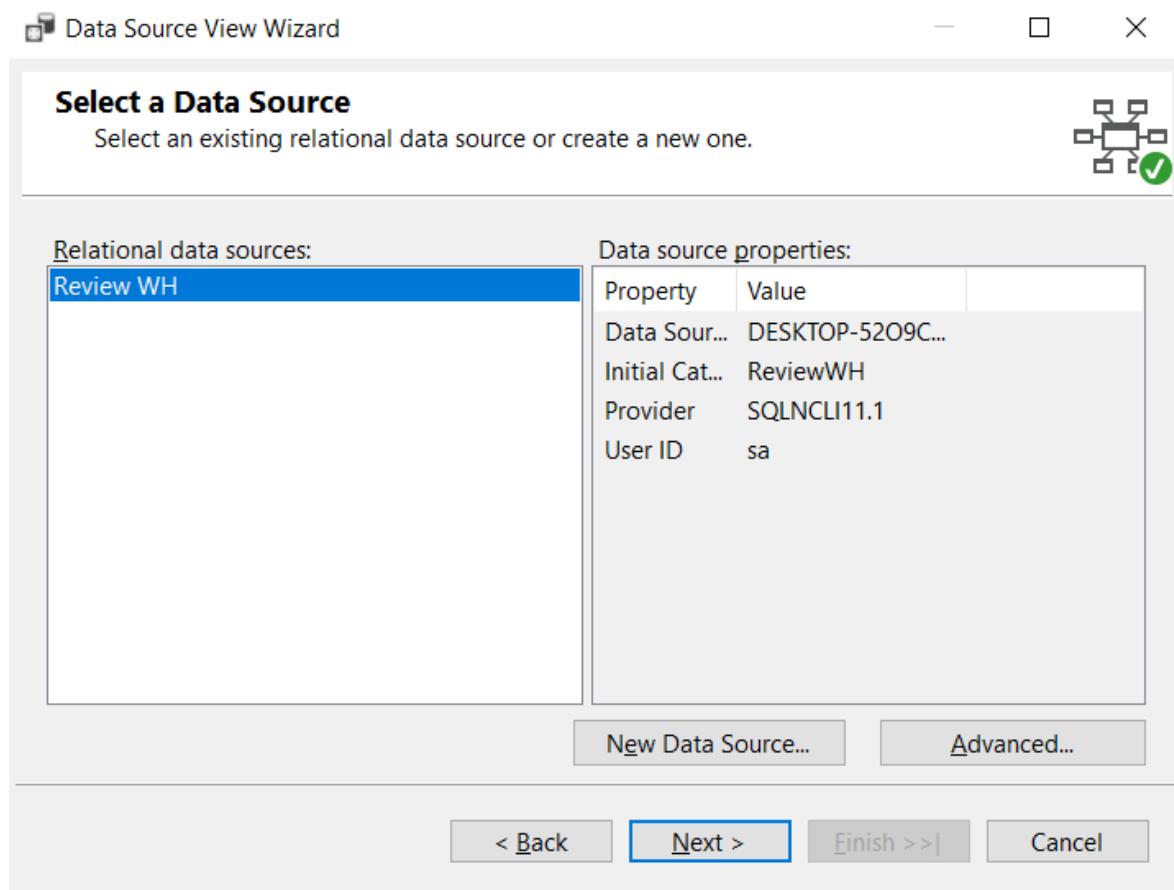
Bước 1: Tại Solution Explorer, ta click chuột phải vào thư mục Data Source Views và chọn **New Data Source View**.



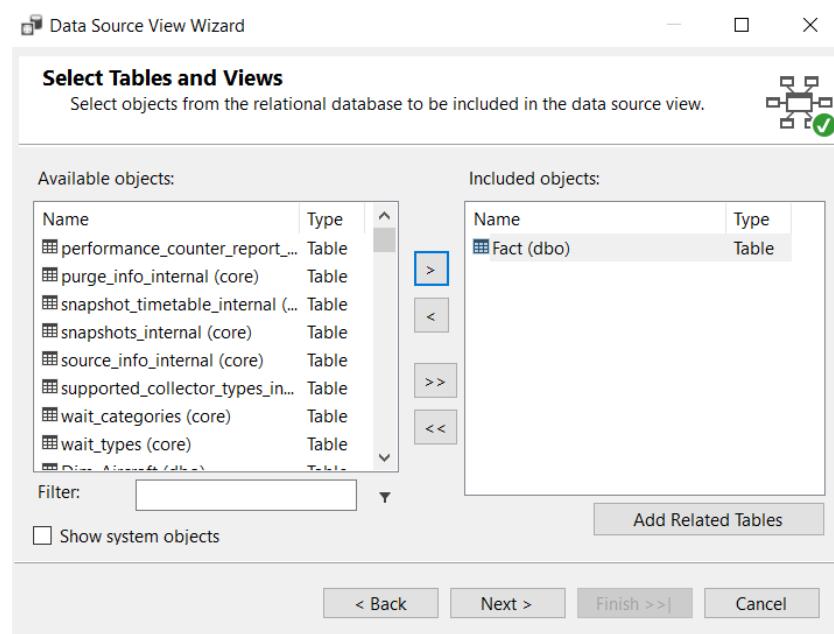
Bước 2: Hộp thoại Data Source View Wizard xuất hiện, chọn **Next** để tiếp tục



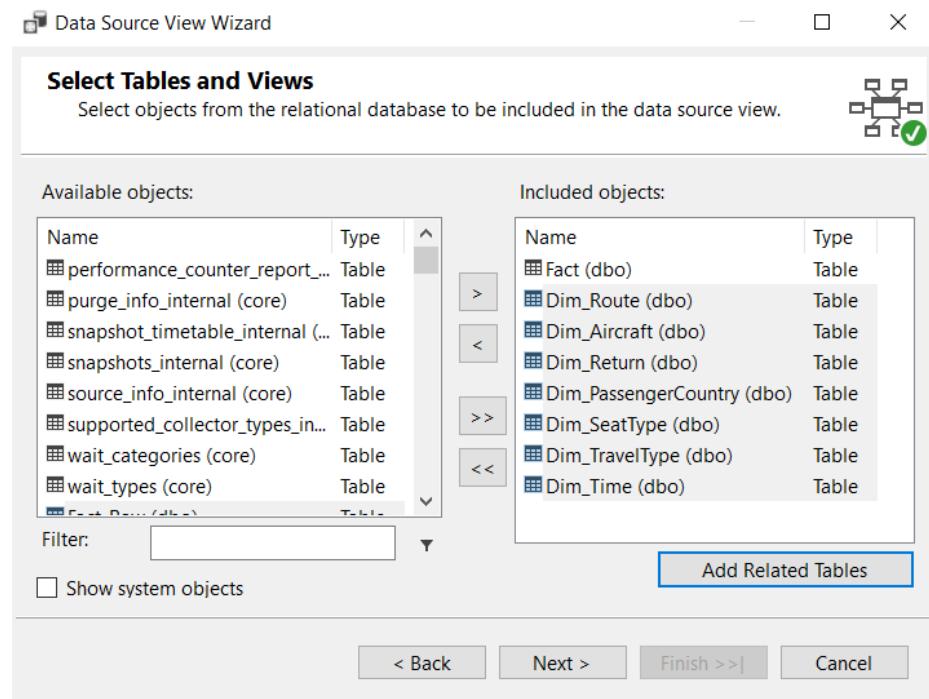
Bước 3: Chọn data source vừa tạo, sau đó chọn **Next** để tiếp tục.



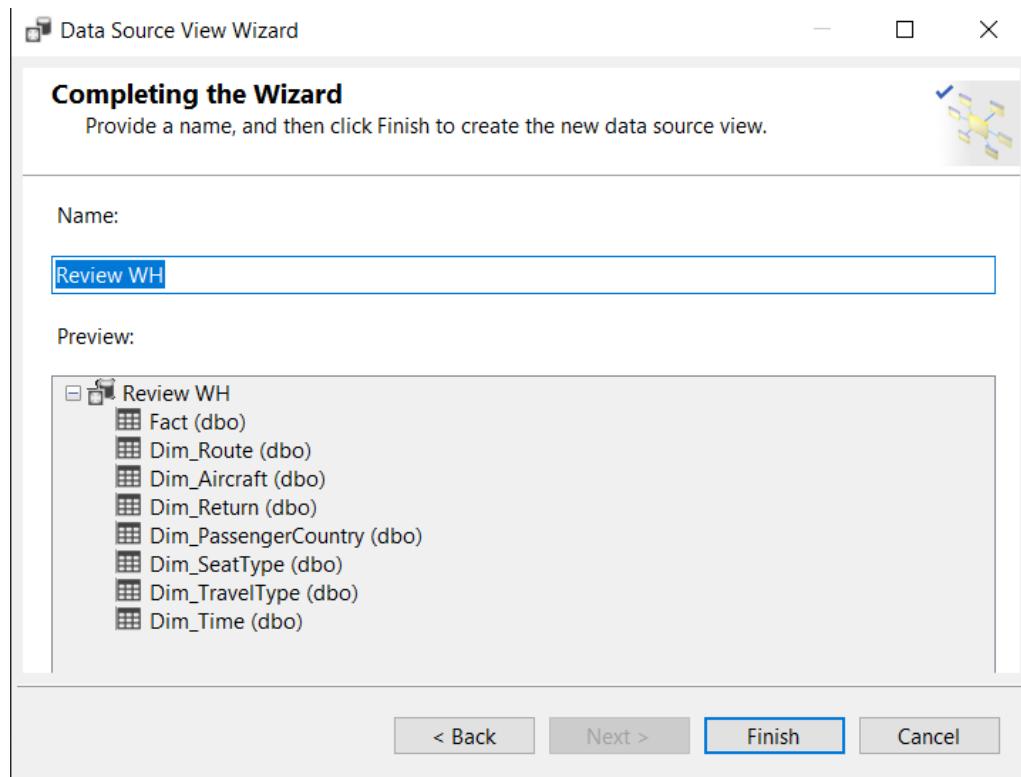
Bước 4: Chọn bảng Fact, sau đó chọn nút > để thêm bảng Fact vào data source view.



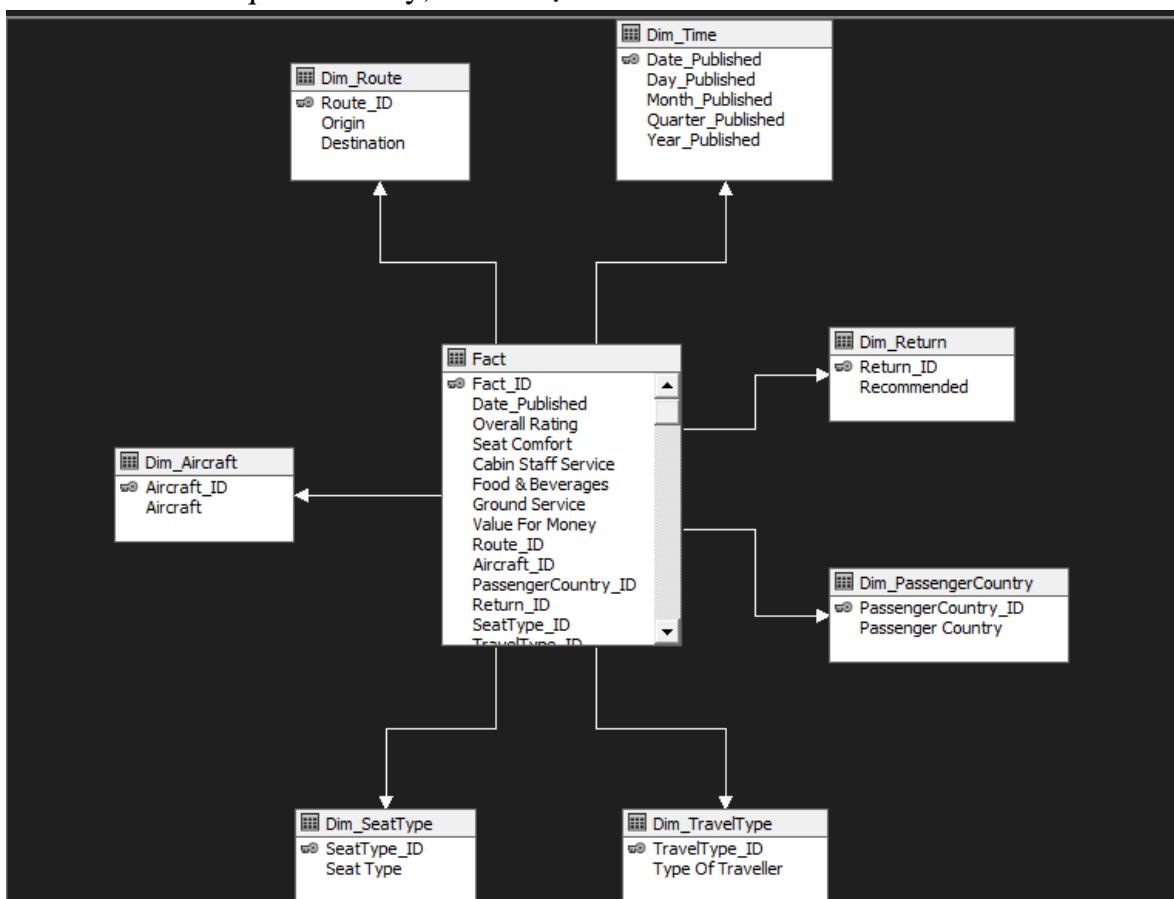
Bước 5: Tiếp theo, chọn nút **Add Related Tables** để thêm tất cả các bảng Dim vào data source view. Sau đó chọn **Next** để tiếp tục.



Bước 6: Chọn **Finish** để hoàn tất quá trình xác định khung nhìn dữ liệu nguồn (Data Source View)

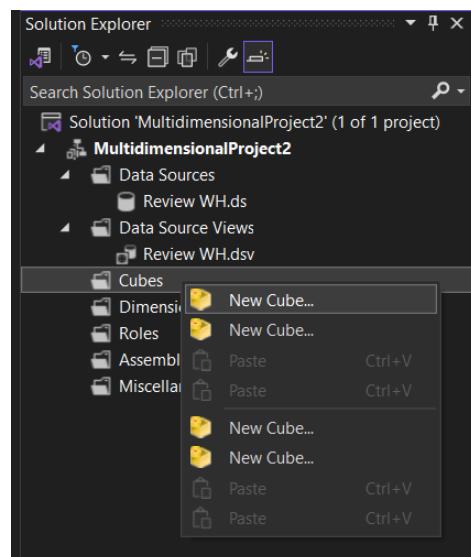


Sau khi kết thúc quá trình này, ta sẽ được data source view như hình sau

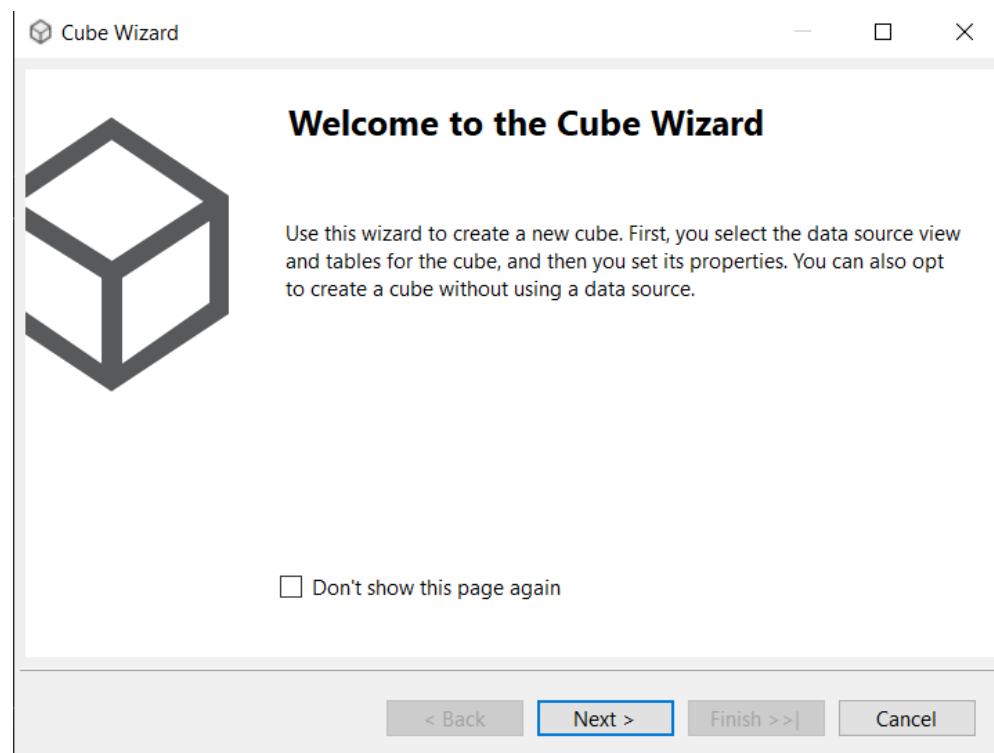


4.4. Xây dựng các khối (Cubes) và xác định các độ đo (Measures)

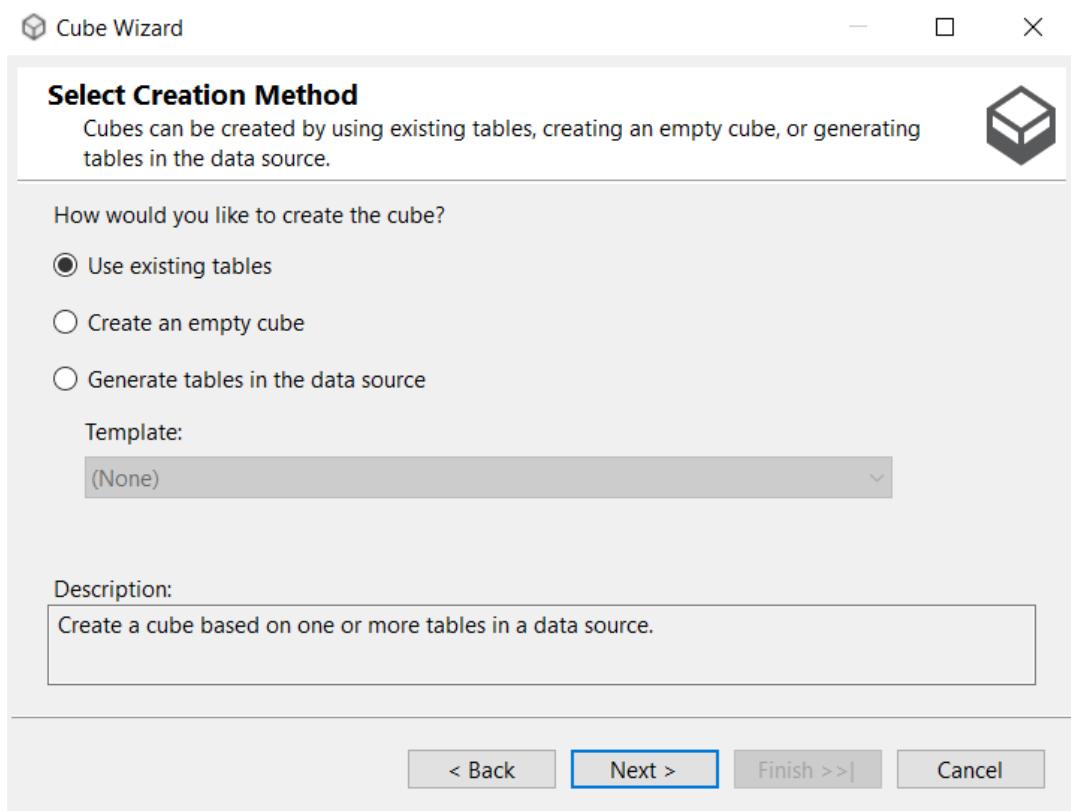
Bước 1: Tại Solution Explorer, ta click chuột phải vào thư mục Cubes và chọn New Cube



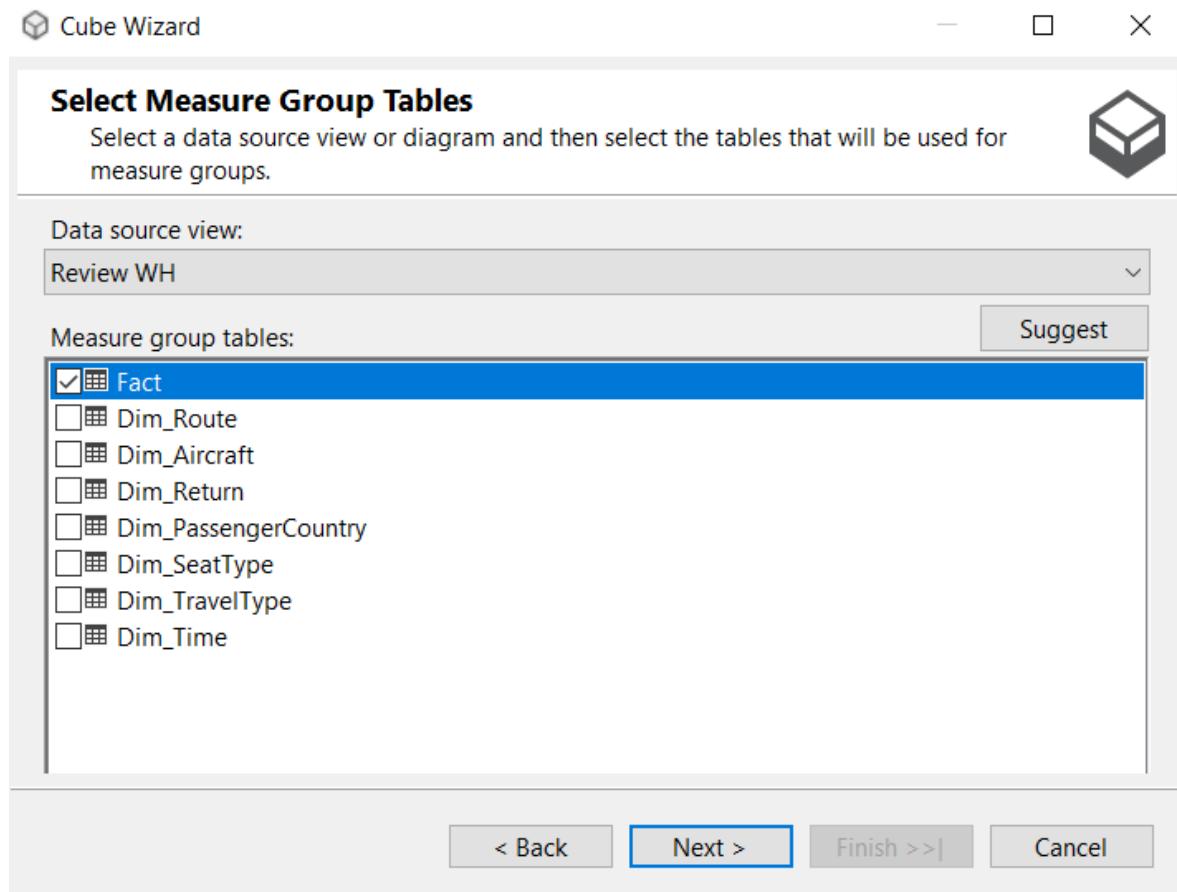
Bước 2: Hộp thoại Cube Wizard xuất hiện, chọn **Next** để tiếp tục.



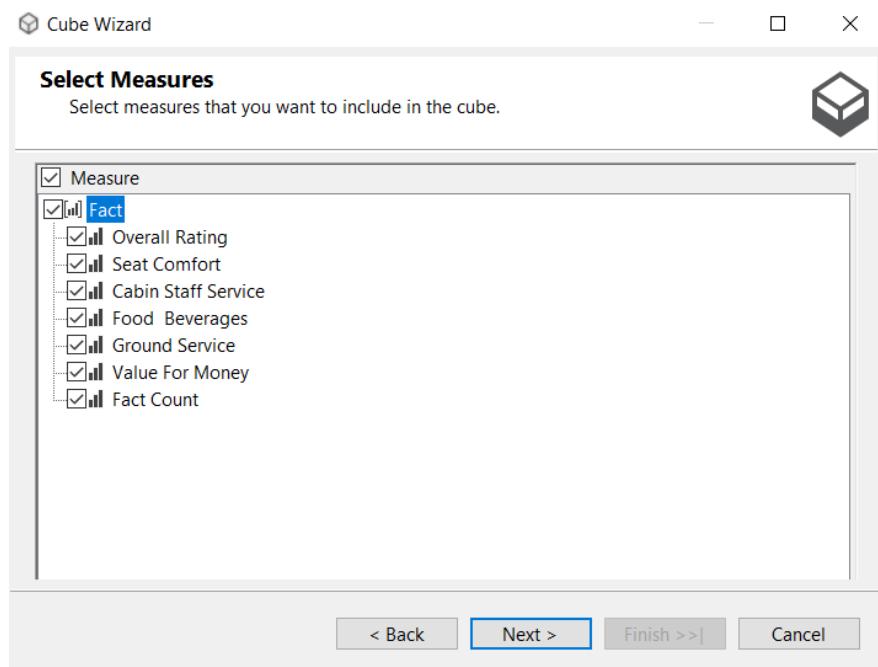
Bước 3: Chọn **use existing tables**, sau đó chọn **Next** để tiếp tục.



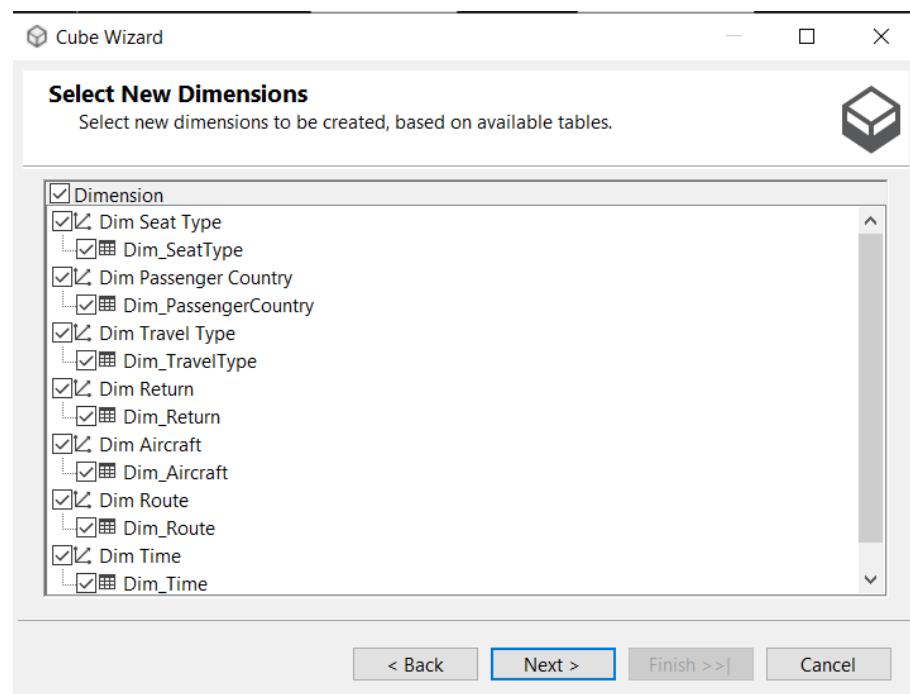
Bước 4: Chọn Fact để phân chia các measure group.



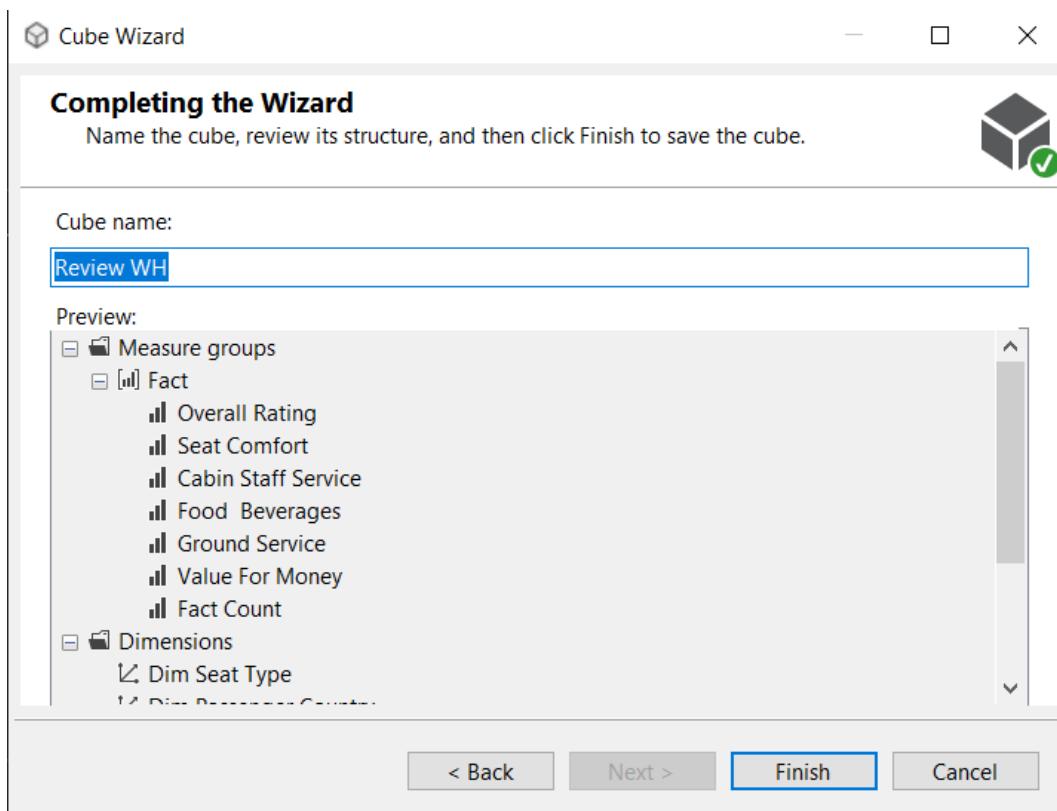
Bước 5: Chọn những độ đo để xuất, sau đó chọn Next để tiếp tục



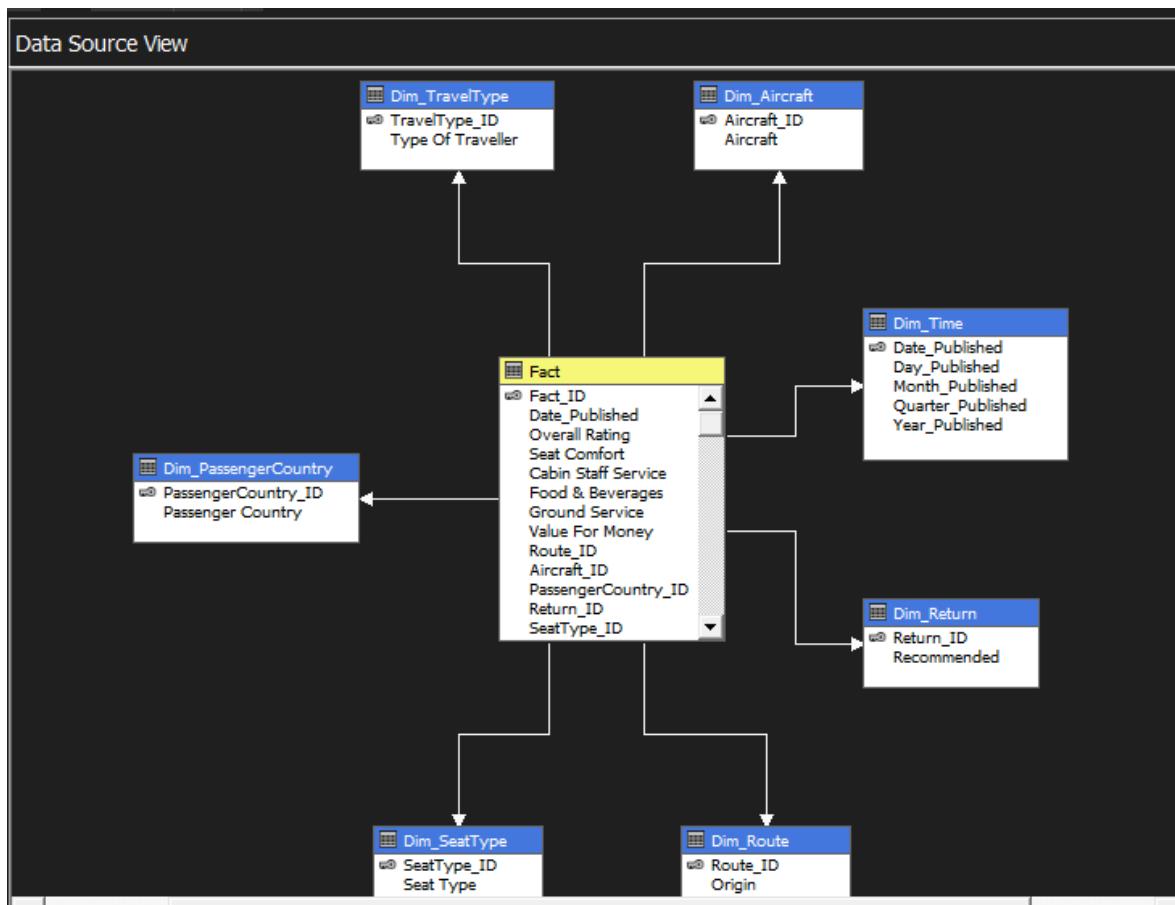
Bước 6: Chọn danh sách các bảng Dimension, sau đó chọn **Next** để tiếp tục.



Bước 7: Chọn **Finish** để hoàn tất quy trình xây dựng các khối (Cubes) và xác định các độ đo (Measures)

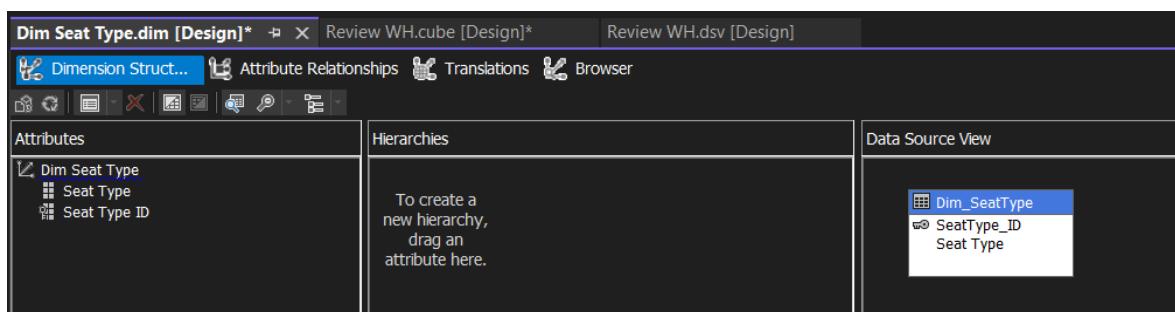


Sau khi kết thúc quá trình này, ta sẽ được kết quả như hình sau:

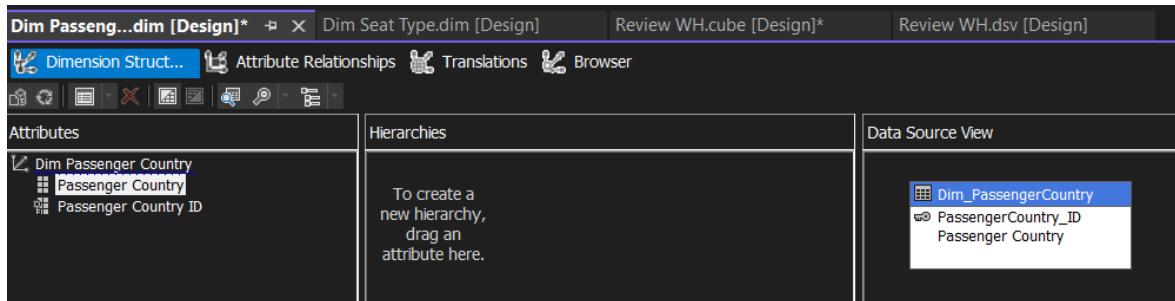


4.5. Xác định các chiều (Dimensions)

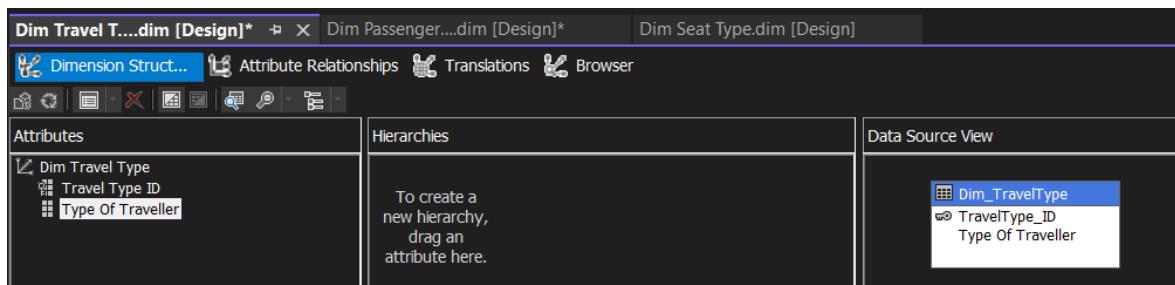
Bước 1: Tại folder Dimensions trong Solution Explorer, ta chọn Dim Seat Type.dim. Sau đó kéo thả các thuộc tính từ Data Source View vào Attributes.



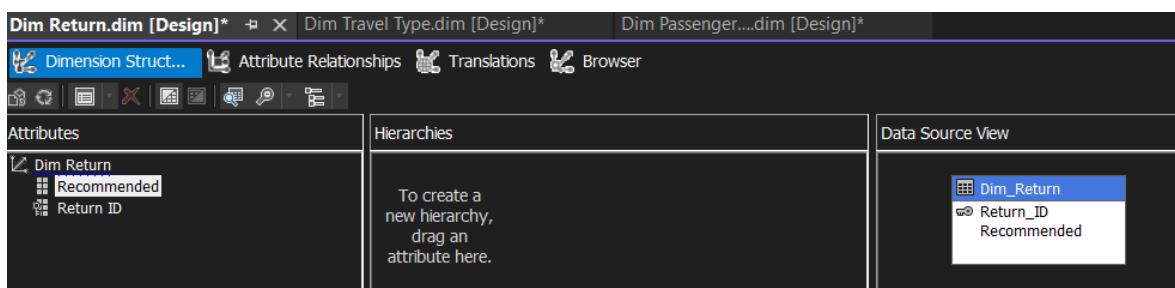
Bước 2: Tại folder Dimensions trong Solution Explorer, ta chọn Dim Passenger Country.dim. Sau đó kéo thả các thuộc tính từ Data Source View vào Attributes.



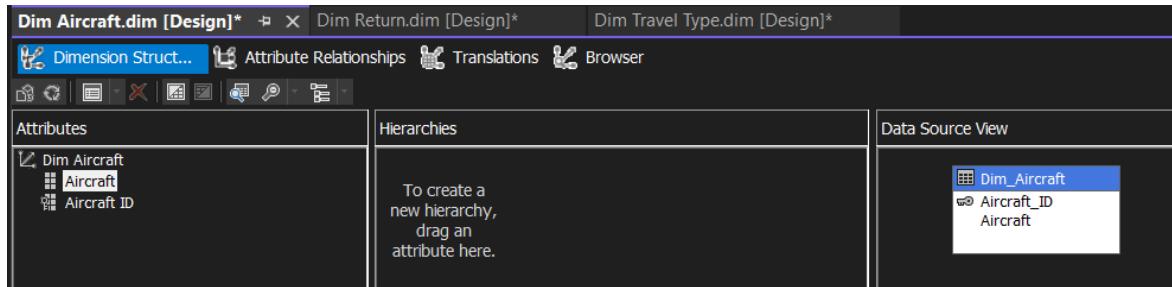
Bước 3: Tại folder Dimensions trong Solution Explorer, ta chọn Dim Travel Type.dim. Sau đó kéo thả các thuộc tính từ Data Source View vào Attributes.



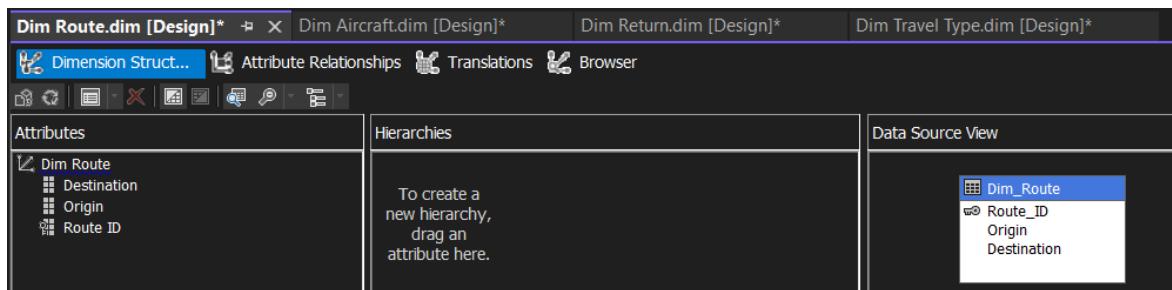
Bước 4: Tại folder Dimensions trong Solution Explorer, ta chọn Dim Return.dim. Sau đó kéo thả các thuộc tính từ Data Source View vào Attributes.



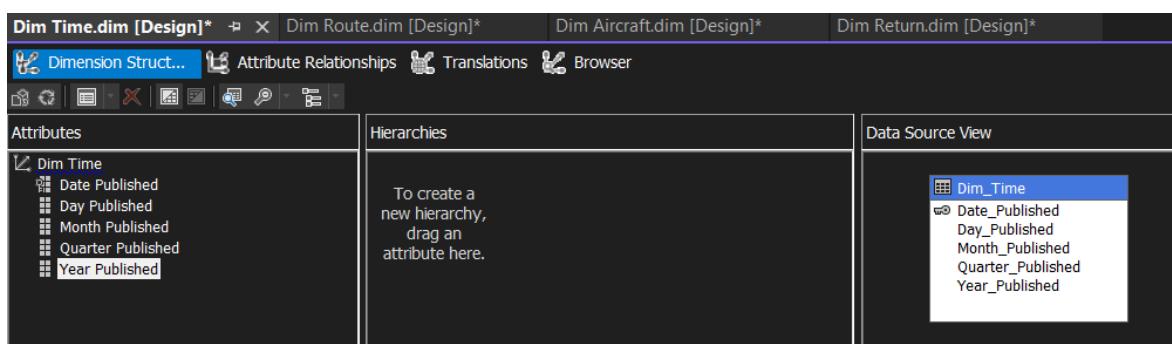
Bước 5: Tại folder Dimensions trong Solution Explorer, ta chọn Dim Aircraft.dim. Sau đó kéo thả các thuộc tính từ Data Source View vào Attributes.



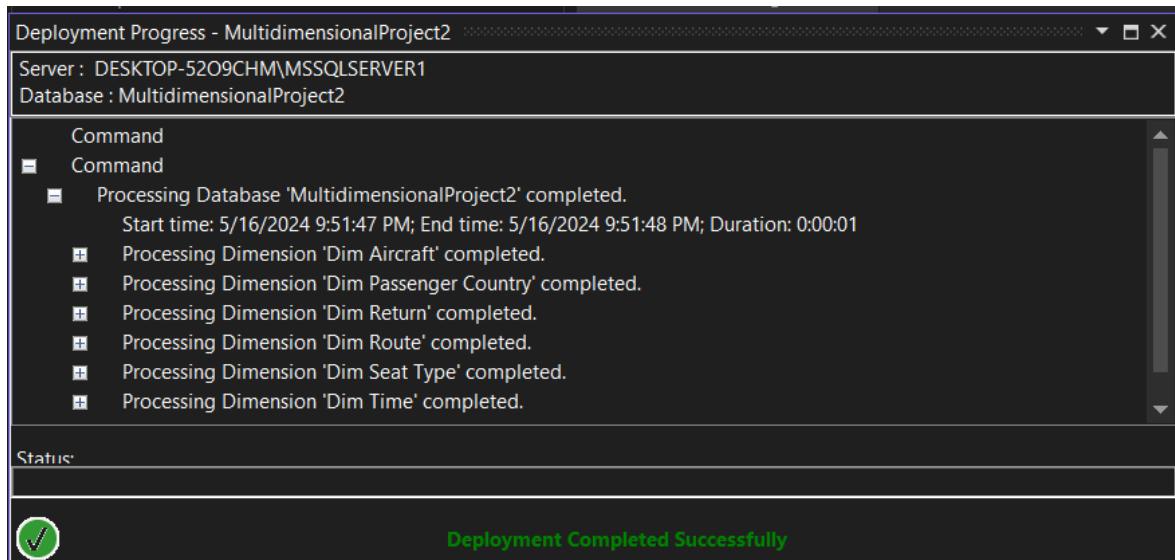
Bước 6: Tại folder Dimensions trong Solution Explorer, ta chọn Dim Route.dim. Sau đó kéo thả các thuộc tính từ Data Source View vào Attributes.



Bước 7: Tại folder Dimensions trong Solution Explorer, ta chọn Dim Time.dim. Sau đó kéo thả các thuộc tính từ Data Source View vào Attributes.



Bước 10: Deploy project. Khi deploy thành công ta sẽ nhận được kết quả như sau:



4.6. Xác định các độ đo (Measures)

4.6.1. Đổi tên và thuộc tính các độ đo ban đầu

Bước 1: Tại khái niệm vừa tạo, chọn **Show Measures Grid** để hiển thị chi tiết các độ đo.

Chi tiết các độ đo sẽ hiển thị dưới dạng bảng, dễ dàng để tương tác

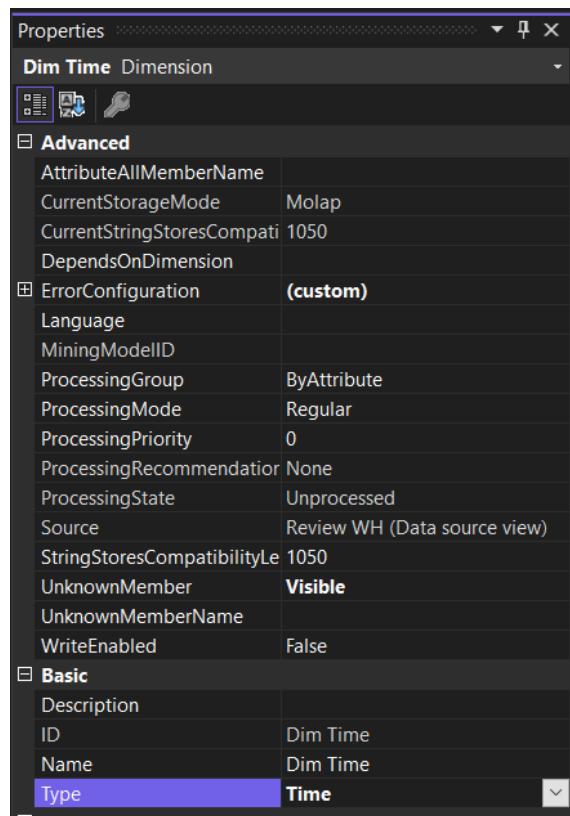
The screenshot shows the 'Review WH.cube [Design]' window. The 'Measures' tab is selected, displaying the following data:

Name	Measure Group	Data Type	Aggregation
Overall Rating	Fact	Double	Sum
Seat Comfort	Fact	Double	Sum
Cabin Staff Service	Fact	Double	Sum
Food Beverages	Fact	Double	Sum
Ground Service	Fact	Double	Sum
Value For Money	Fact	Double	Sum
Fact Count	Fact	Integer	Count
Add new measure...			

Sau khi đổi tên và thuộc tính các độ đo ban đầu. Một thông báo xuất hiện yêu cầu ta phải có một time dimension.

Review WH.cube [Design]*				
Cube Stru... Dimension... Calculations KPIs Actions Part				
Measures				
	Name	Measure Group	Data Type	Aggregation
	Overall Rating	Fact	Double	AverageOfC...
	Seat Comfort	Fact	Semi-additive measure requires a time dimension.	
	Cabin Staff Service	Fact	Double	Sum
	Food Beverages	Fact	Double	Sum
	Ground Service	Fact	Double	Sum
	Value For Money	Fact	Double	Sum
	Fact Count	Fact	Integer	Count
Add new measure...				

Bước 2: Mở Dim Time.dim. Tại cửa sổ Properties, ta đổi kiểu bảng từ Regular sang Time



Quá trình hoàn tất, ta có được các độ đo

	Name	Measure Group	Data Type	Aggregation
1	Overall Rating	Fact	Double	Sum
2	Seat Comfort	Fact	Double	Sum
3	Cabin Staff Service	Fact	Double	Sum
4	Food Beverages	Fact	Double	Sum
5	Ground Service	Fact	Double	Sum
6	Value For Money	Fact	Double	Sum
7	Fact Count	Fact	Integer	Count
	Add new measure...			

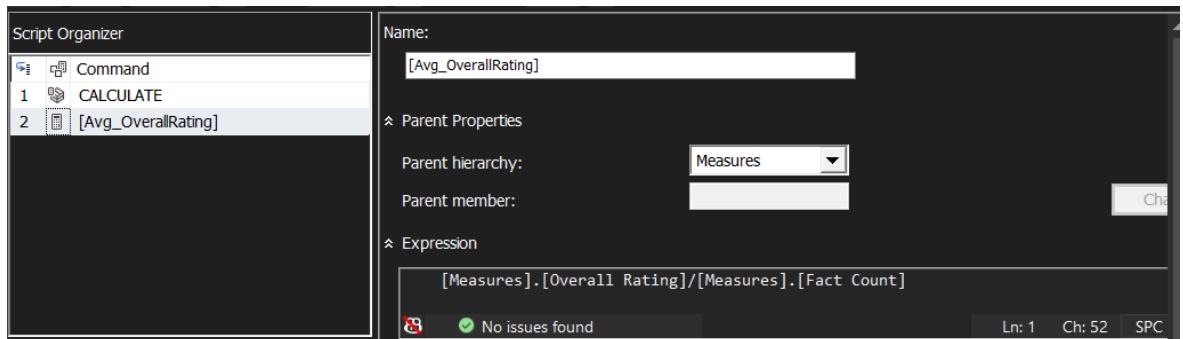
4.6.2. Tạo các New Calculated Number

Bước 1: Chuyển qua tab Calculations và chọn New Calculated Number

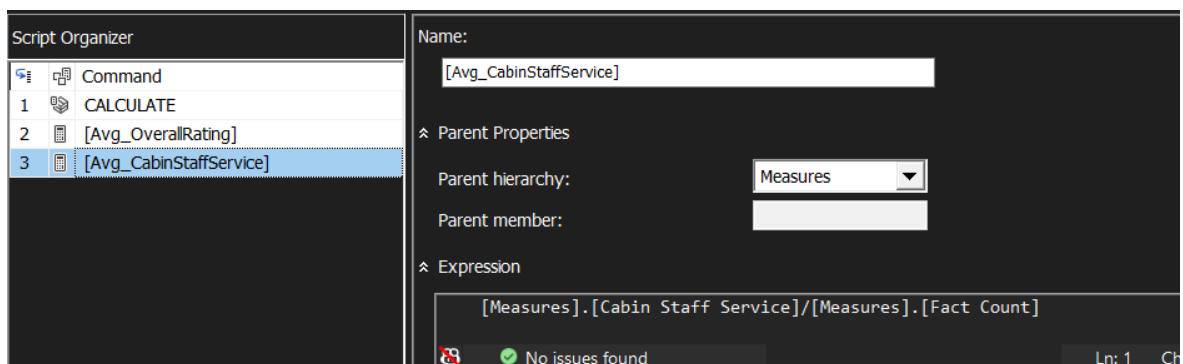
The screenshot shows the 'Script Organizer' pane on the left with a script containing a 'CALCULATE' command and a '[Calculated Member]' entry. The main pane displays the creation of a new calculated member named '[Calculated Member]'. The properties for this member include:

- Name:** [Calculated Member]
- Parent Properties:**
 - Parent hierarchy:** Measures
 - Parent member:** (empty)
- Expression:** (No issues found)
- Additional Properties:**
 - Format string:** (empty)
 - Visible:** True
 - Non-empty behavior:** (empty)
 - Associated measure group:** (Undefined)
 - Display folder:** (empty)

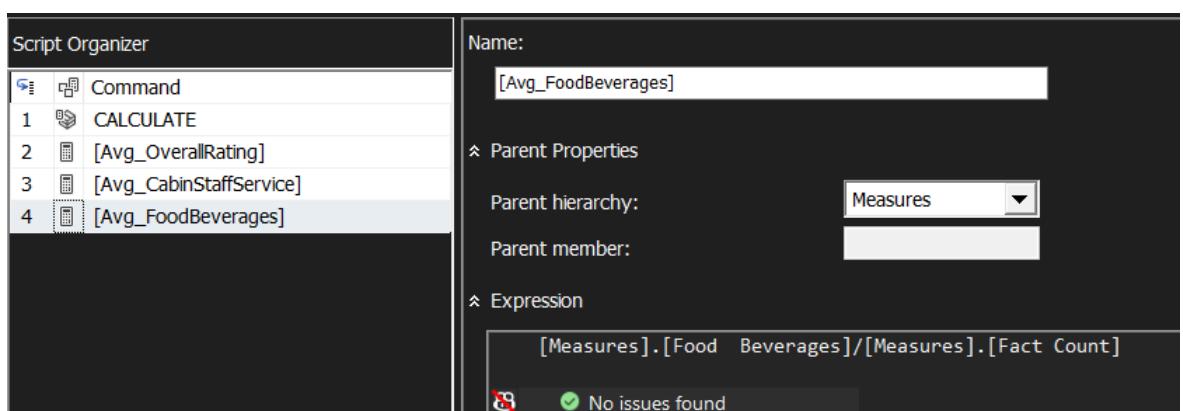
Bước 2: Đặt tên Avg_OverallRating và công thức cho Calculated Number



Bước 3: Đặt tên Avg_CabinStaffService và công thức cho Calculated Number



Bước 4: Đặt tên Avg_FoodBeverages và công thức cho Calculated Number



Bước 5: Đặt tên Avg_GroundService và công thức cho Calculated Number

The screenshot shows the Script Organizer interface. On the left, under the 'Command' section, there is a list of calculated numbers: 1. CALCULATE, 2. [Avg_OverallRating], 3. [Avg_CabinStaffService], 4. [Avg_FoodBeverages], and 5. [Avg_GroundService]. The fifth item, [Avg_GroundService], is selected. On the right, the properties for this selection are displayed: Name: [Avg_GroundService], Parent hierarchy: Measures, and Expression: [Measures].[Ground Service]/[Measures].[Fact Count]. The background of the interface is dark.

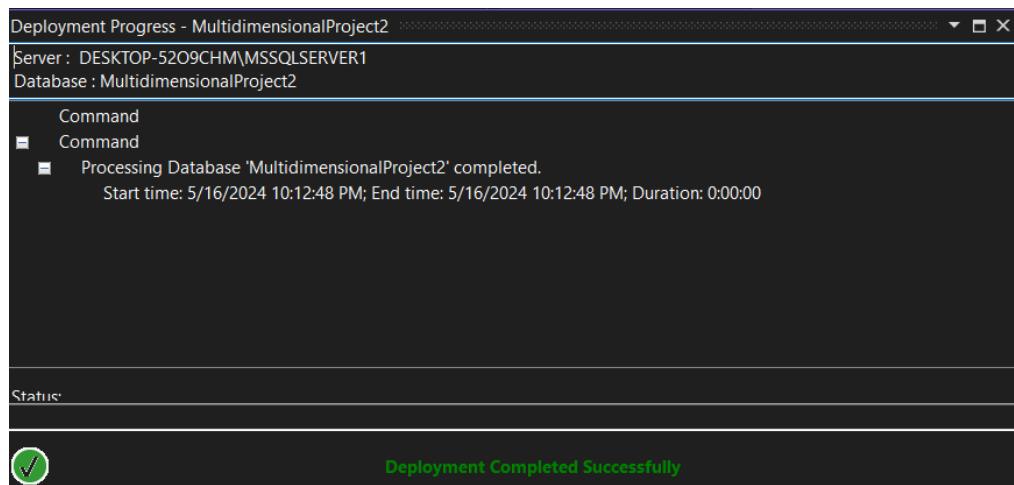
Bước 6: Đặt tên Avg_SeatComfort và công thức cho Calculated Number

The screenshot shows the Script Organizer interface. On the left, under the 'Command' section, there is a list of calculated numbers: 1. CALCULATE, 2. [Avg_OverallRating], 3. [Avg_CabinStaffService], 4. [Avg_FoodBeverages], 5. [Avg_GroundService], and 6. [Avg_SeatComfort]. The sixth item, [Avg_SeatComfort], is selected. On the right, the properties for this selection are displayed: Name: [Avg_SeatComfort], Parent hierarchy: Measures, and Expression: [Measures].[Seat Comfort]/[Measures].[Fact Count]. The background of the interface is dark.

Bước 7: Đặt tên Avg_OverallRating và công thức cho Calculated Number

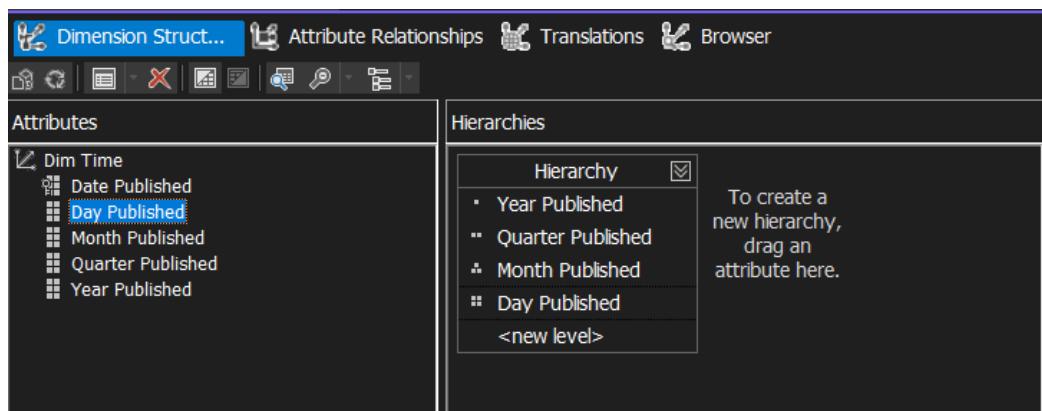
The screenshot shows the Script Organizer interface. On the left, under the 'Command' section, there is a list of calculated numbers: 1. CALCULATE, 2. [Avg_OverallRating], 3. [Avg_CabinStaffService], 4. [Avg_FoodBeverages], 5. [Avg_GroundService], 6. [Avg_SeatComfort], and 7. [Avg_ValueForMoney]. The seventh item, [Avg_ValueForMoney], is selected. On the right, the properties for this selection are displayed: Name: [Avg_ValueForMoney], Parent hierarchy: Measures, and Expression: [Measures].[Value For Money]/[Measures].[Fact Count]. The background of the interface is dark.

Bước 8: Deploy lại project. Kết quả sau khi Deploy

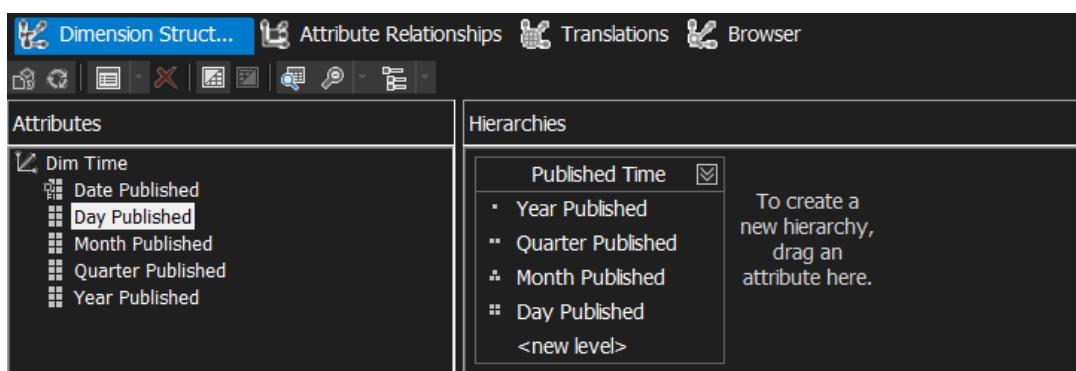


4.7. Phân cấp trong bảng chiều Dim_Time

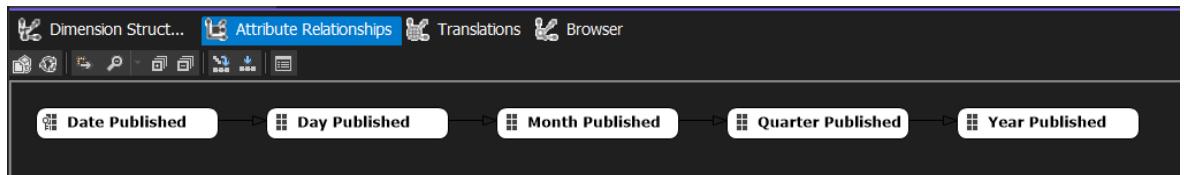
Bước 1: Kéo những thuộc tính cần phân cấp qua cửa sổ Hierarchies



Bước 2: Sắp xếp lại các thuộc tính phân cấp theo thứ tự và đổi tên Hierarchy thành Published Time.



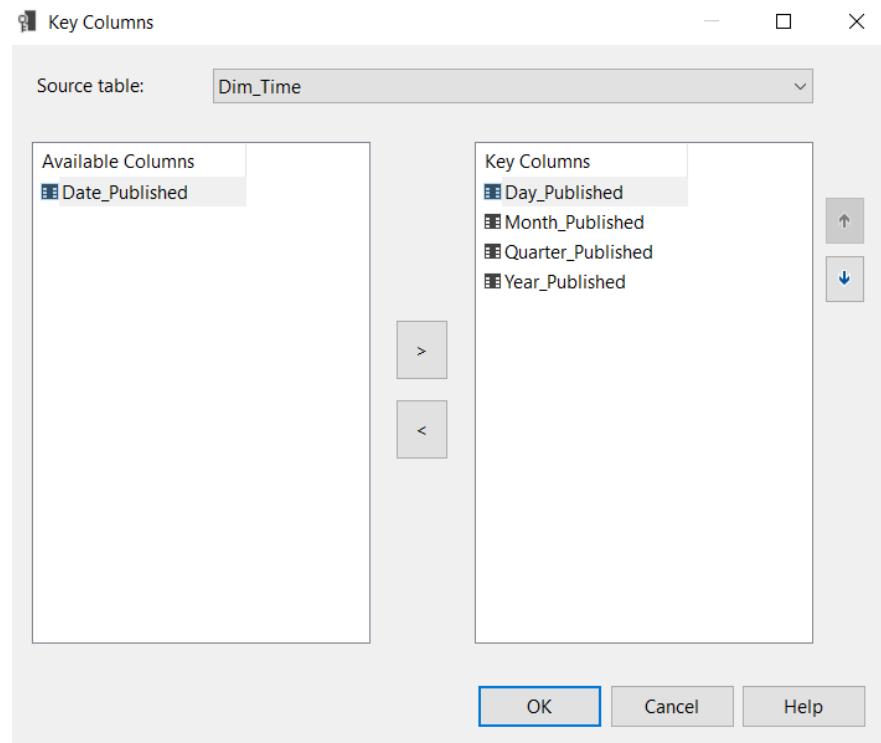
Bước 3: Tại panel Attribute Relationships, tạo mối quan hệ như sau



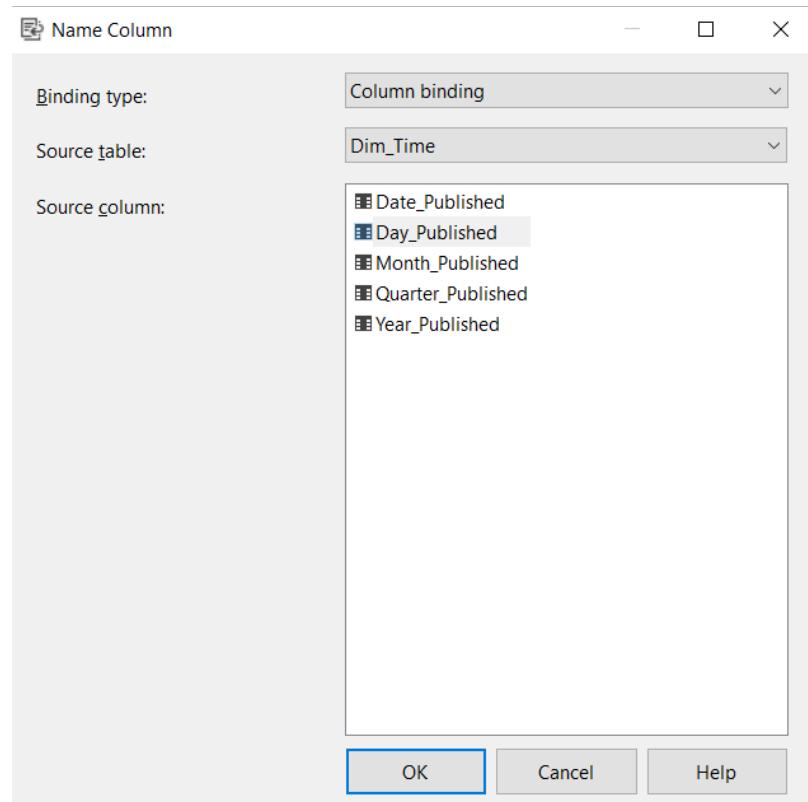
Bước 4: Chính khóa cột (KeyColumns) và tên cột (Name Column) của thuộc tính Day_Published

Day Published DimensionAttribute	
OrderByAttribute	
ProcessingState	Unprocessed
TokenizationBehavior	TokenizationNone
UserEditFlag	0
VertipaqCompressionHi	VertipaqAutomatic
Basic	
Description	
FormatString	
ID	Day Published
Name	Day Published
Type	Regular
Usage	Regular
Misc	
AttributeHierarchyOrder	True
ExtendedType	
GroupingBehavior	EncourageGrouping
InstanceSelection	None
MemberNamesUnique	False
VisualizationProperties	
Parent-Child	
MembersWithData	NonLeafDataVisible
MembersWithDataCaption	
NamingTemplate	
RootMemberIf	ParentIsBlankSelfOrMissing
UnaryOperatorColumn	(none)
Source	
CustomRollupColumn	(none)
CustomRollupProperties	(none)
KeyColumns	dbo_Dim_Time.Day_Published (...)
NameColumn	(none)
ValueColumn	(none)

Thêm các thuộc tính cấp cao hơn vào KeyColumns, sau đó chọn **OK** để hoàn tất.



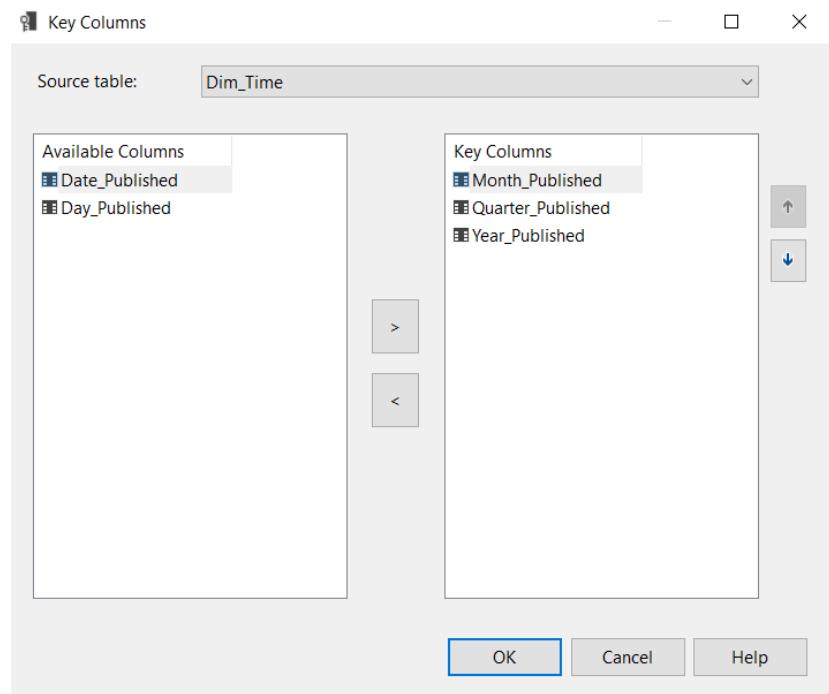
Tại cửa sổ Properties của thuộc tính, ta chọn Name Column và chọn tên thuộc tính



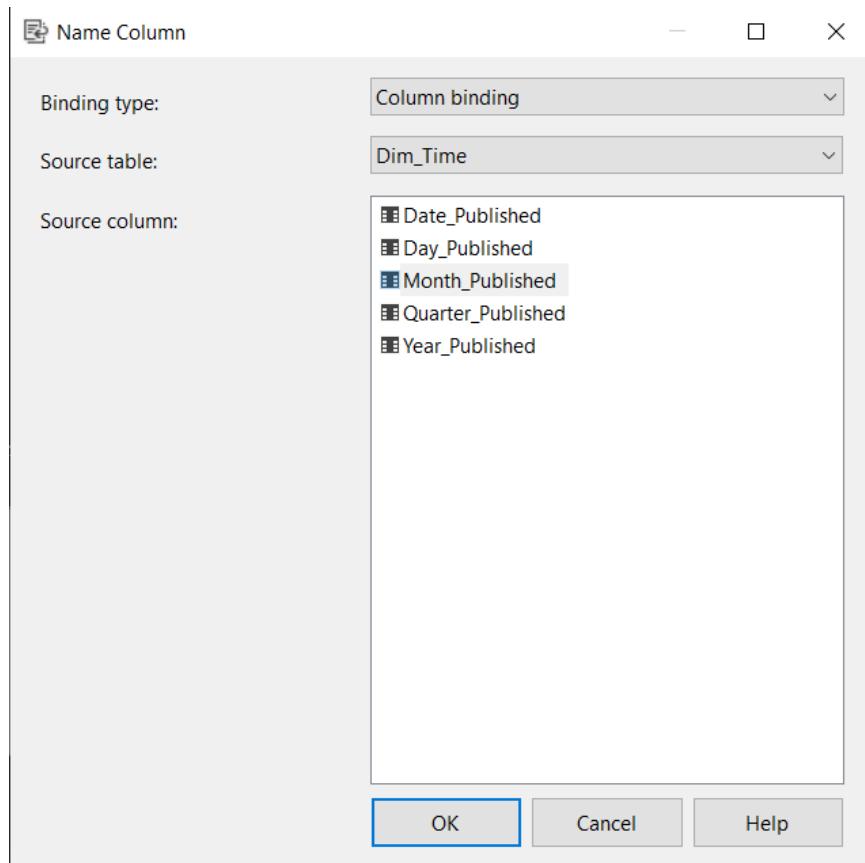
Bước 5: Chỉnh khóa cột (KeyColumns) và tên cột (Name Column) của thuộc tính Month_Published

Properties	
Month Published DimensionAttribute	
OrderByAttribute	
ProcessingState	Unprocessed
TokenizationBehavior	TokenizationNone
UserEditFlag	0
VertipaqCompressionHint	VertipaqAutomatic
Basic	
Description	
FormatString	
ID	Month Published
Name	Month Published
Type	Regular
Usage	Regular
Misc	
AttributeHierarchyOrdered	True
ExtendedType	
GroupingBehavior	EncourageGrouping
InstanceSelection	None
MemberNamesUnique	False
VisualizationProperties	
Parent-Child	
MembersWithData	NonLeafDataVisible
MembersWithDataCaption	
NamingTemplate	
RootMemberIf	ParentIsBlankSelfOrMissing
UnaryOperatorColumn	(none)
Source	
CustomRollupColumn	(none)
CustomRollupPropertiesCol	(none)
KeyColumns	(Collection)
NameColumn	Dim_Time.Month Published
ValueColumn	(none)

Thêm các thuộc tính cấp cao hơn vào KeyColumns, sau đó chọn **OK** để hoàn tất.



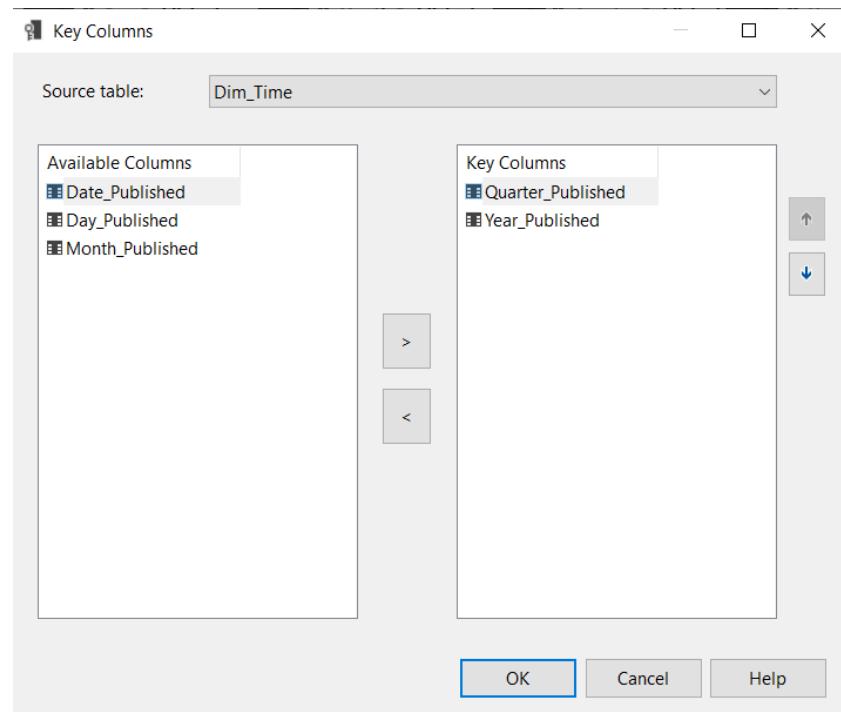
Tại cửa sổ Properties của thuộc tính, ta chọn Name Column và chọn tên thuộc tính



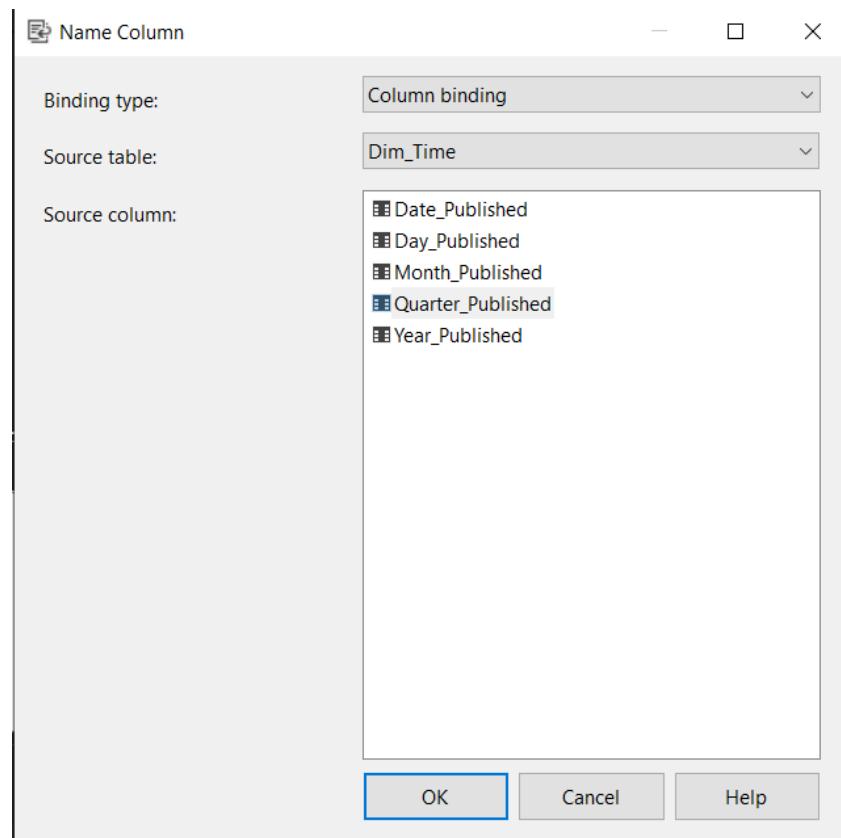
Bước 6: Chỉnh khóa cột (KeyColumns) và tên cột (Name Column) của thuộc tính Quarter_Published

Properties	
Quarter Published DimensionAttribute	
OrderByAttribute	
ProcessingState	Unprocessed
TokenizationBehavior	TokenizationNone
UserEditFlag	0
VertipaqCompressionHint	VertipaqAutomatic
Basic	
Description	
FormatString	
ID	Quarter Published
Name	Quarter Published
Type	Regular
Usage	Regular
Misc	
AttributeHierarchyOrdered	True
ExtendedType	
GroupingBehavior	EncourageGrouping
InstanceSelection	None
MemberNamesUnique	False
VisualizationProperties	
Parent-Child	
MembersWithData	NonLeafDataVisible
MembersWithDataCaption	
NamingTemplate	
RootMemberIf	ParentIsBlankSelfOrMissing
UnaryOperatorColumn	(none)
Source	
CustomRollupColumn	(none)
CustomRollupPropertiesCol	(none)
KeyColumns	(Collection) ...
NameColumn	Dim_Time.Quarter_Published
ValueColumn	(none)

Thêm các thuộc tính cấp cao hơn vào KeyColumns, sau đó chọn **OK** để hoàn tất.



Tại cửa sổ Properties của thuộc tính, ta chọn Name Column và chọn tên thuộc tính



4.8. Chạy dự án SSAS

Sau khi quá trình phân cấp cho các bảng chiều hoàn tất, ta thực hiện delpoy project để đảm bảo không có lỗi xảy ra sau quá trình phân cấp. Sau khi deploy ta được:



4.9. Thực hiện 15 câu truy vấn – Quá trình phân tích dữ liệu bằng thao tác tay trên các khối CUBE

4.9.1. Câu truy vấn 1

Nội dung câu truy vấn: Thống kê trung bình điểm đánh giá tổng thể qua từng năm

Thực hiện câu truy vấn:

Kéo thả thuộc tính [Year Published] trong Dim_Time và Calculated Number [Avg_OverallRating]

The screenshot shows the SSAS Dimension Designer interface for the 'Dim Time.dim [Design]' cube. The top navigation bar includes options like Cube Struc..., Dimension..., Calculations, KPIs, Actions, Partitions, Aggregati..., Perspectives, Translations, and Browser. The main area has tabs for Edit as Text, Import..., MDX, and various icons. On the left, the 'Metadata' pane shows the dimension structure under 'Dim Time'. On the right, a data grid displays the average overall rating by year.

Year Published	Avg_OverallRating
2012	4.15
2013	5.87671232876...
2014	6.31858407079...
2015	5.33024691358...
2016	6.13218390804...
2017	4.03468208092...
2018	3.04545454545...
2019	2.79714285714...
2020	3.10909090909...
2021	2.625
2022	2.88095238095...
2023	3.49612403100...
2024	6.625

4.9.2. Câu truy vấn 2

Nội dung câu truy vấn: Thống kê những hành trình trong năm 2023 có hành khách có quốc tịch United Kingdom tham gia

Thực hiện câu truy vấn:

Bước 1: Kéo thả thuộc tính trong bảng Dim_Route và độ đo [Fact Count]

Bước 2: Kéo thả thuộc tính [Year Published] và [Passenger Country] vào Hierarchy và chọn Filter để lọc theo điều kiện tương ứng

Route ID	Origin	Destination	Fact Count
150	Biarritz	Brussels	1
158	Birmi...	Dublin	1
170	Bodr...	Dublin	1
201	Bristol	Cologne	1
202	Bristol	Dublin	2
274	Cardiff	Faro	1
285	Clam...	Manchester	1
302	Colo...	Stansted	1
337	Dort...	London St...	1
353	Dublin	Bodrum	1
355	Dublin	Bristol	1
402	East ...	Berlin	1
416	Edin...	Tirana	1
424	Edin...	Fuerteven...	1
436	Edin...	Stansted	1

4.9.3. Câu truy vấn 3

Nội dung câu truy vấn: Thống kê những hành trình có hành khách có quốc tịch United States tham gia và có trung bình tổng thẻ ≤ 3

Thực hiện câu truy vấn:

Bước 1: Tạo Named Set [US Avg_OverallRating ≤ 3]

Bước 2: Kéo thả thuộc tính trong bảng Dim_Route và độ đo [Fact Count]

Kéo thả thuộc tính [Passenger Country] vào Hierarchy, chọn Operator là In và chọn Named Set vừa tạo

Route ID	Origin	Destination	Fact Count
1			23
29	Alicante	Venice	1
34	Amsterdam	Malaga	1
36	Ancona Falconara	London Stansted	1
40	ATH	JMK	1
55	Athens	Rome	1
57	Athens	Santorini	2
69	Barcelona	Dublin	1
72	Barcelona	Frankfurt	2
73	Barcelona	Ibiza	1
78	Barcelona	Malaga	1
79	Barcelona	Mallorca	1
80	Barcelona	Manchester	1
83	Barcelona	Palma	1
86	Barcelona	Porto	1
87	Barcelona	Prague	1
88	Barcelona	Rome	1

4.9.4. Câu truy vấn 4

Nội dung câu truy vấn: Thống kê những hành trình có điểm trung bình dịch vụ ăn uống < 3 nhưng dịch vụ nhân viên ≥ 4

Thực hiện câu truy vấn:

Bước 1: Tạo Named Set [Avg_FoodBeverages < 3 Avg_CabinStaffService ≥ 4]

The screenshot shows the Analysis Services Designer application with the following details:

- Title Bar:** Dim Time.dim [Design] | Review WH.cube [Design]*
- Toolbar:** Cube Stru..., Dimension..., Calculations, KPIs, Actions, Partitions, Aggregati..., Perspectives, Translations, Browser.
- Script Organizer:** Shows a list of calculated members, including one named [Avg_FoodBeverages < 3 Avg_CabinStaffService >= 4].
- Expression Editor:** Displays the MDX expression for the calculated member:


```
FILTER(
    [Dim Route].[Route Id].CHILDREN,
    [Measures].[Avg_Food Beverages] < 3 and
    [Measures].[Avg_CabinStaffService] >= 4)
```

No issues found
- Additional Properties:** Type: Dynamic, Display folder: [empty]
- Calculation Tools:** Metadata, Functions, Templates.
- Measure Group:** <All>
- Navigation:** Review WH, Measures, Dim Aircraft, Dim Passenger Country, Dim Return, Dim Route.

Bước 2: Kéo thả thuộc tính trong bảng Dim_Route và độ đo [Fact Count]

Kéo thả thuộc tính [Route ID] vào Hierarchy, chọn Operator là In và chọn Named Set vừa tạo

The screenshot shows the Analysis Services Designer application with the following details:

- Title Bar:** Review WH.cube [Design]
- Toolbar:** Edit as Text, Import..., MDX, etc.
- Table View:** Shows the Dim_Route table with columns: Route ID, Origin, Destination, Fact Count.
- Filter Definition:** A dialog box shows the filter configuration:

Dimension	Hierarchy	Operator	Filter Expression
Dim Route	Route ID	In	Avg_FoodBeverages < 3 Avg_CabinStaffService >= 4
- Navigation:** Shows the Dim_Route dimension tree with the filter applied.

4.9.5. Câu truy vấn 5

Nội dung câu truy vấn: Thống kê những hành trình từ năm 2019 – 2020 mà có khách hàng Australia tham gia và có mong muốn quay lại

Thực hiện câu truy vấn:

Kéo thả các thuộc tính của bảng Dim_Route và độ đo [Fact Count]

Kéo thả các thuộc tính [Year Published], [Passenger Country] và [Recommended] vào Hierarchy và chọn Filter để lọc theo điều kiện tương ứng

Dimension	Hierarchy	Operator	Filter Expression
Dim Time	Year Published	Equal	{ 2019, 2020 }
Dim Passenger C...	Passenger Cou...	Equal	{ Australia }
Dim Return	Recommended	Equal	{ yes }

4.9.6. Câu truy vấn 6

Nội dung câu truy vấn: Thống kê những top 3 loại máy bay có điểm trung bình chỗ ngồi thoải mái nhất theo từng năm

Thực hiện câu truy vấn:

The screenshot shows the Analysis Services Management Studio interface. On the left, there's a navigation pane with 'Metadata' and 'Functions' tabs, and a search bar. Below it is a tree view of the 'Review WH' database, including 'Measures' (Fact, KPIs), 'Dim Aircraft', 'Dim Passenger Country', and 'Dim Return'. The main area contains an MDX query window and a results grid.

```

SELECT
NON EMPTY { [Measures].[Avg_SeatComfort] } ON COLUMNS,
NON EMPTY
GENERATE(
[Dim Time].[Year Published].CHILDREN,
TOPCOUNT(
[Dim Time].[Year Published].currentmember* [Dim Aircraft].[Aircraft].CHILDREN,
3,
[Measures].[Avg_SeatComfort]
)
)
ON ROWS

```

Year Published	Aircraft	Avg_SeatComfort
2012	Boein...	2.65
2013	Boein...	2.85616438356...
2014	Boein...	2.87315634218...
2015	Boein...	4.5
2015	737300	4
2015	A320	4
2016	B747	5
2016	Boein...	5
2016	Boein...	4
2017	Boein...	4
2017	Boein...	3.07142857142...
2017	B737-...	3
2018	B737-...	3.5

4.9.7. Câu truy vấn 7

Nội dung câu truy vấn: Thống kê top 3 loại hình du lịch khách hàng không muộn quay lại nhặt nhưng trung bình giá trị đem lại > 2

Thực hiện truy vấn:

Bước 1: Tạo Named Set [Top 3 TypeOfTraveller Avg_ValueForMoney > 2 Recommended No]

```

Name: [Top 3 TypeOfTraveller Avg_ValueForMoney]

FILTER(
    TOPCOUNT(
        [Dim Travel Type].[Type Of Traveller].[Type Of Traveller].MEMBERS,
        3,
        [Measures].[Fact Count]
    ),
    [Measures].[Avg_ValueForMoney] > 2 AND
    [Dim Return].[Recommended].&[No]
)

```

No issues found

Bước 2: Kéo thả thuộc tính Type Of Traveller độ đo [Fact Count] và [Avg_ValueForMoney]

Kéo thả thuộc tính [Type Of Traveller] vào Hierarchy, chọn Operator là In và chọn Named Set vừa tạo

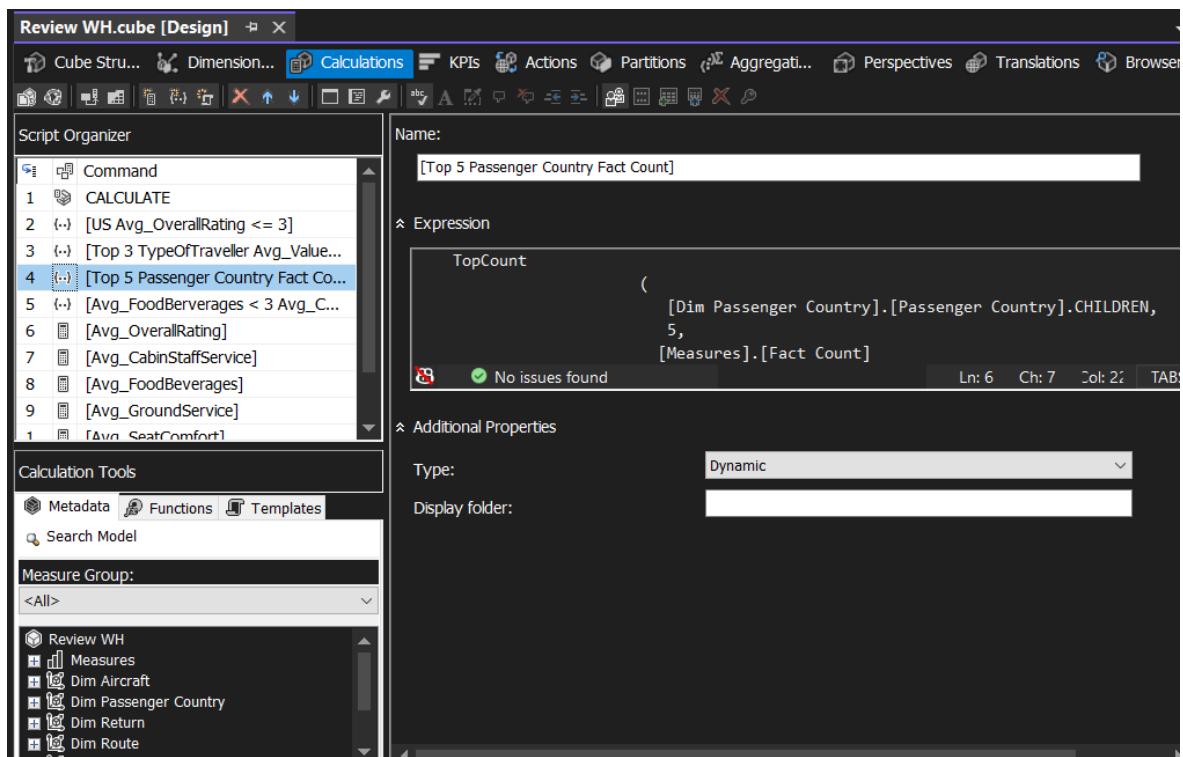
Type Of Traveller	Fact Count	Avg_ValueForMoney
Couple Leisure	1201	2.80183180682764
Family Leisure	374	2.43850267379679
Solo Leisure	548	2.7956204379562

4.9.8. Câu truy vấn 8

Nội dung câu truy vấn: Thông kê top 5 quốc gia có số lượng hành khách tham gia nhiều nhất những năm 2019-2023

Thực hiện truy vấn:

Bước 1: Tạo Named Set [Top 5 Passenger Country Fact Count]



Bước 2: Kéo thả thuộc tính trong bảng Dim_PassengerCountry và độ đo [Fact Count]

Kéo thả thuộc tính [Passenger Country] vào Hierarchy, chọn Operator là In và chọn Named Set vừa tạo

Kéo thả thuộc tính [Year Published] vào Hierarchy và chọn Filter để lọc theo điều kiện tương ứng

The screenshot shows the SSAS Management Studio interface with the title bar "Review WH.cube [Design]". The menu bar includes "File", "Edit", "View", "Tools", "Help", and "Language: Default". The toolbar has icons for "Edit as Text", "Import...", "MDX", and various navigation and search tools. The main area has tabs for "Metadata" and "Search Model". A left sidebar lists dimensions: "Passenger Country ID", "Dim Return", "Dim Route", "Dim Seat Type", "Dim Time" (selected), "Dim Travel Type", and "Dim Traveler". The "Dim Time" section shows hierarchies for Date Published, Day Published, Month Published, Quarter Published, Year Published, and Published Time. The right pane displays a table titled "Passenger Country ID" with columns "Passenger Country ID", "Passenger Country", and "Fact Count". The data shows five rows: Germany (38), Ireland (52), Spain (23), United Kingdom (308), and United States (76). A filter panel at the top right shows "Dimension: Dim Passenger Country", "Hierarchy: Passenger Country", "Operator: In", "Filter Expression: Top 5 Passenger Country Fact Count", and "Parameters: { 2019, 2020, 2021, 2022, 2023 }".

4.9.9. Câu truy vấn 9

Nội dung câu truy vấn: Thống kê những hành trình có trung bình dịch vụ mặt đất, dịch vụ ăn uống, chỗ ngồi thoải mái và dịch vụ nhân viên > 2.5 nhưng khách hàng không có nhu cầu quay lại

Thực hiện truy vấn:

Bước 1: Tạo Named Set [4 Avgs > 2.5]

Bước 2: Kéo thả thuộc tính [Origin] [Destination] và các Calculated Number

Kéo thả thuộc tính [Recommended] vào Hierarchy và chọn Filter để lọc theo điều kiện tương ứng

Kéo thả thuộc tính [Route ID] vào Hierarchy, chọn Operator là In và chọn Named Set vừa tạo

Dimension	Hierarchy	Operator	Filter Expression	Param...
Dim Return	Recommended	Equal	{ no }	<input type="checkbox"/> <input type="checkbox"/>
Dim Route	Route ID	In	4 Avgs > 2.5	<input type="checkbox"/> <input type="checkbox"/>

	Origin	Destination	Avg_CabinStaffService	Avg_GroundService	Avg_Food Beverages	Avg_SeatCo
Athens		Frankfurt	4	4	3	5
Rome Fiumicino		Dublin	5	4	5	5

4.9.10. Câu truy vấn 10

Nội dung câu truy vấn: Theo từng loại hình du lịch liệt kê các quốc gia, trung bình giá trị đem lại và trung bình tổng thể

Thực hiện truy vấn:

Kéo thả các thuộc tính [Type Of Traveller] [Passenger Country] và [Avg_ValueForMoney] [Avg_OverallRating]

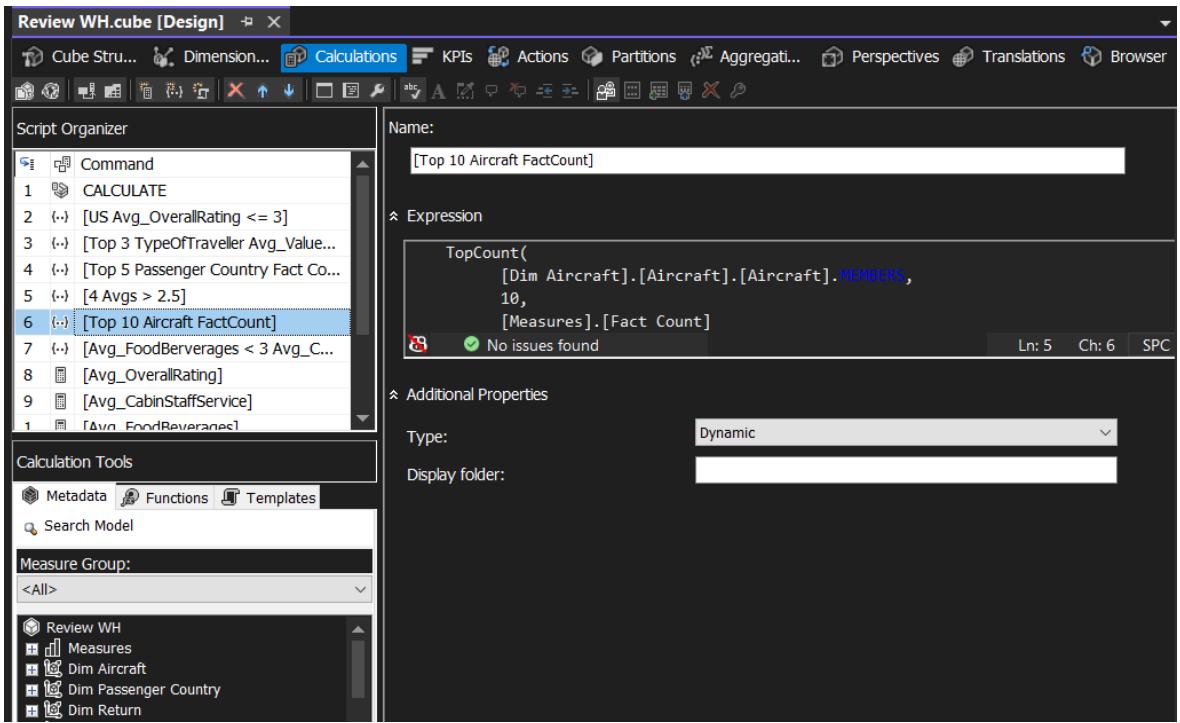
Type Of Traveller	Passenger Country	Avg_ValueForMoney	Avg_OverallRating
Business	Albania	5	8
Business	Austria	2.33333333333333	3.333333333333...
Business	Belgium	3	4.75
Business	Bulgaria	2	1
Business	Canada	4	5
Business	Cyprus	1	1
Business	Czech Republic	1	1
Business	Estonia	1	1
Business	France	1	1
Business	Germany	2.4	3.4
Business	Greece	3	4.5
Business	Hungary	3	2
Business	India	1	1
Business	Indonesia	5	6
Business	Ireland	3.26666666666667	5.133333333333...
Business	Italy	2.14285714285714	3.857142857142...
Business	Jordan	5	9

4.9.11. Câu truy vấn 11

Nội dung câu truy vấn: Thống kê top 10 máy bay được khai thác nhiều nhất năm 2019

Thực hiện truy vấn:

Bước 1: Tạo Named Set [Top 10 Aircraft FactCount]



Bước 2: Kéo thả thuộc tính [Aircraft] và độ đo [Fact Count]

Kéo thả thuộc tính [Year Published] vào Hierarchy và chọn Filter để lọc theo điều kiện tương ứng

Kéo thả thuộc tính [Aircraft] vào Hierarchy, chọn Operator là In và chọn Named Set vừa tạo

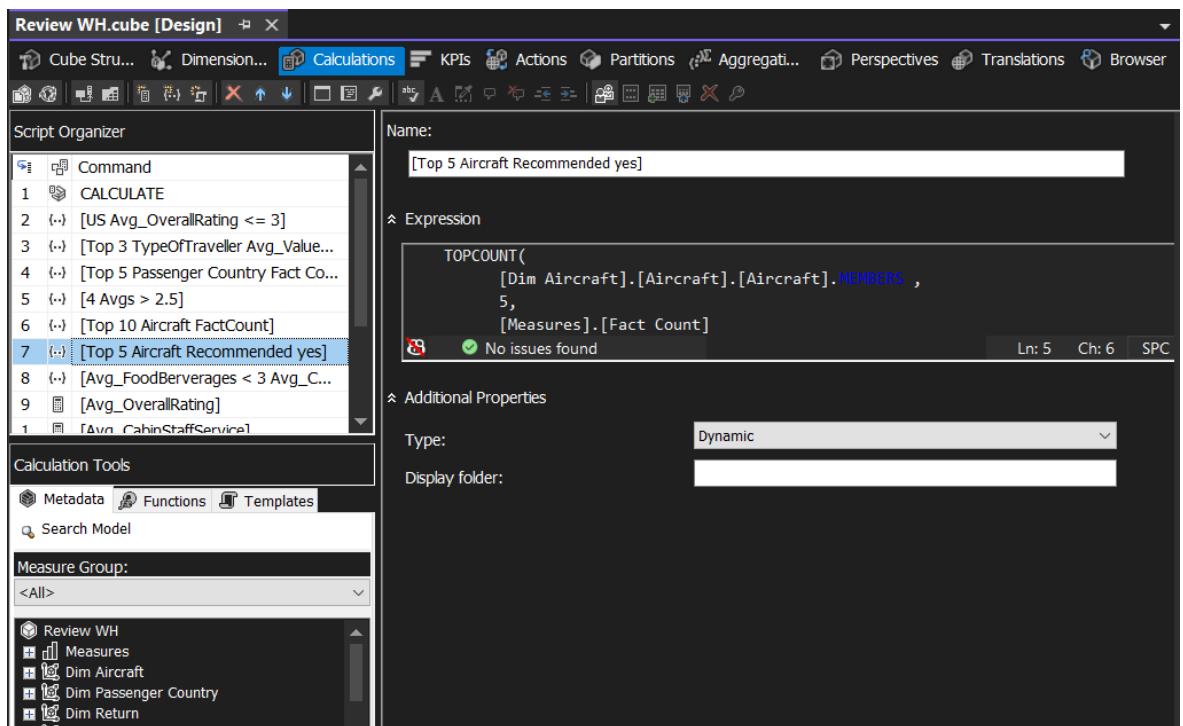
Dimension	Hierarchy	Operator	Filter Expression	Param...																			
Dim Time	# Year Published	Equal	{ 2019 }	<input type="checkbox"/> <input type="checkbox"/>																			
Dim Aircraft	# Aircraft	In	Top 10 Aircraft FactCount	<input type="checkbox"/> <input type="checkbox"/>																			
<Select dimension>																							
<table border="1"> <thead> <tr> <th>Aircraft</th> <th>Fact Count</th> </tr> </thead> <tbody> <tr> <td>A320</td> <td>1</td> </tr> <tr> <td>B737</td> <td>1</td> </tr> <tr> <td>B737 800</td> <td>1</td> </tr> <tr> <td>Boeing 737-800</td> <td>1</td> </tr> <tr> <td>Boeing 737-800</td> <td>33</td> </tr> <tr> <td>Boeing 737 800</td> <td>2</td> </tr> <tr> <td>Boeing 737-800</td> <td>300</td> </tr> <tr> <td>Boeing 738</td> <td>2</td> </tr> <tr> <td>Boeing 747</td> <td>3</td> </tr> </tbody> </table>				Aircraft	Fact Count	A320	1	B737	1	B737 800	1	Boeing 737-800	1	Boeing 737-800	33	Boeing 737 800	2	Boeing 737-800	300	Boeing 738	2	Boeing 747	3
Aircraft	Fact Count																						
A320	1																						
B737	1																						
B737 800	1																						
Boeing 737-800	1																						
Boeing 737-800	33																						
Boeing 737 800	2																						
Boeing 737-800	300																						
Boeing 738	2																						
Boeing 747	3																						

4.9.12. Câu truy vấn 12

Nội dung câu truy vấn: Thông kê top 5 máy bay khách hàng có nhu cầu quay lại nhiều nhất

Thực hiện truy vấn:

Bước 1: Tạo Named Set [Top 5 Aircraft Recommended yes]



Bước 2: Kéo thả thuộc tính [Aircraft] [Recommended và độ đo [Fact Count]]

Kéo thả thuộc tính [Aircraft] vào Hierarchy, chọn Operator là In và chọn Named Set vừa tạo

Kéo thả thuộc tính [Recommended] vào Hierarchy và chọn Filter để lọc theo điều kiện tương ứng

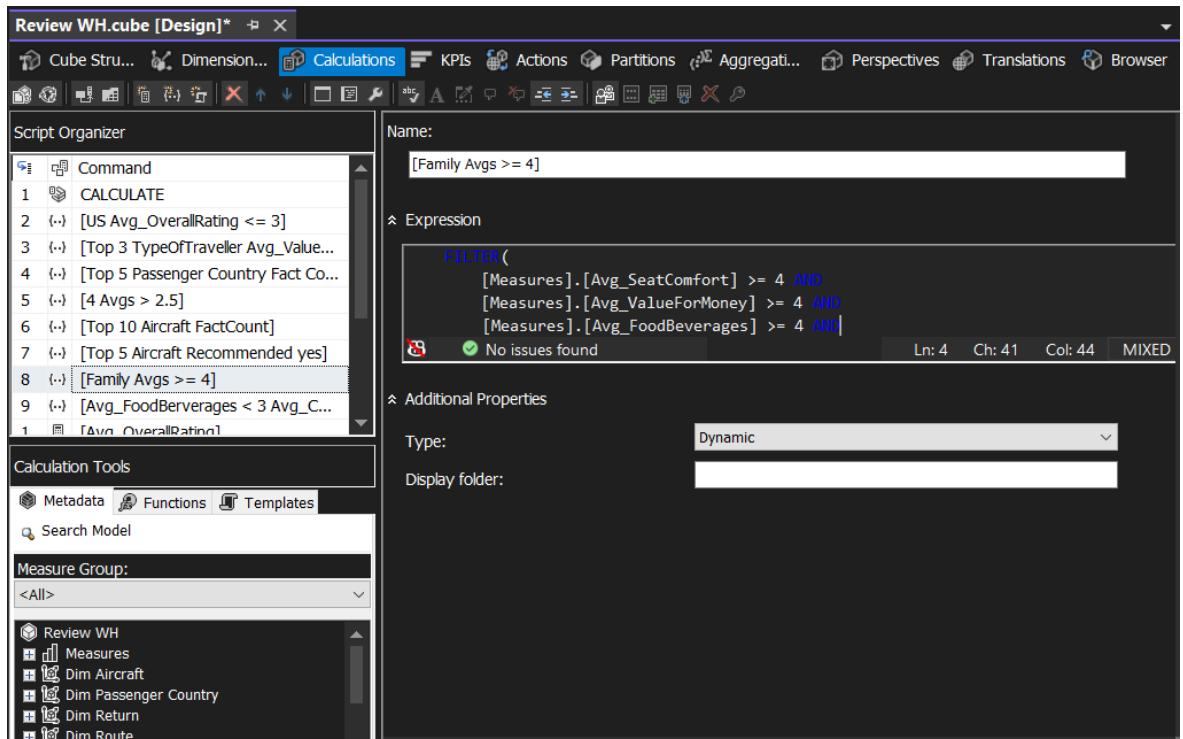
Aircraft	Recommended	Fact Count
B737-800	yes	6
Boeing 737	yes	71
Boeing 737 800	yes	8
Boeing 737-800	yes	741
Boeing 738	yes	6

4.9.13. Câu truy vấn 13

Nội dung câu truy vấn: Thống kê các hành trình với loại hình là gia đình, khách hàng có nhu cầu quay lại, có trung bình thoải mái chỗ ngồi, dịch vụ ăn uống và giá trị mang lại ≥ 4

Thực hiện truy vấn:

Bước 1: Tạo Named Set [3 avg >= 4]



Bước 2: Kéo thả thuộc tính [Origin] [Destination] [Recommended] [Type Of Traveller] [Avg_SeatComfort] [Avg_FoodBeverages] [Avg_ValueForMoney] vào cửa sổ Execute

Kéo thả thuộc tính [Recommended] vào Hierarchy và chọn Filter để lọc theo điều kiện tương ứng

Kéo thả thuộc tính [Route ID] vào Hierarchy và chọn Operator là In và chọn Named Set vừa tạo

Kéo thả thuộc tính [Type Of Traveller] vào Hierarchy và chọn Filter để lọc theo điều kiện tương ứng

Dimension	Hierarchy	Operator	Filter Expression	Para...
Dim Return	Recommended	Equal	{ yes }	<input type="checkbox"/>
Dim Route	Route ID	In	3 avg >= 4	<input type="checkbox"/>
Dim Travel Type	Type Of Traveller	Equal	{ Family Leisure }	<input type="checkbox"/>
<Select dimensio...				

Origin	Destination	Recommended	Type Of Traveller	Avg_SeatComfort	Avg_...
ALC	EMA	yes	Family Leisure	5	5
Athens	Stansted	yes	Family Leisure	4	5
BCN	TSF	yes	Family Leisure	4	5
Birmingham	Palma	yes	Family Leisure	5	5
Bologna	Bucharest	yes	Family Leisure	4	4
Bologna	Stansted	yes	Family Leisure	4	4
Bratislava	Dublin	yes	Family Leisure	4	4
Brussels	Gran Cana...	yes	Family Leisure	4	5
East Midlands	Bergerac	yes	Family Leisure	4	5
Faro	Stansted	yes	Family Leisure	4	3
Frankfurt-Hahn	Rome Cia...	yes	Family Leisure	4	4
KrakÃ³w	Stansted	yes	Family Leisure	5	5
LTN	ACE	yes	Family Leisure	5	5
Manchester	Ibiza	yes	Family Leisure	4	4

4.9.14. Câu truy vấn 14

Nội dung câu truy vấn: Với top 10 quốc gia có số lượng khách hàng tham gia nhiều nhất liệt kê số lượng chuyến bay theo từng loại hình du lịch

Thực hiện truy vấn:

Bước 1: Tạo Named Set [Top 10 Passenger Country FactCount]

Bước 2: Kéo thả thuộc tính [Passenger Country] [Type Of Traveller] [Fact Count] vào cửa sổ Execute

Kéo thả thuộc tính [Passenger Country] vào Hierarchy và chọn Operator là In và chọn Named Set vừa tạo

Passenger Country	Type Of Traveller	Fact Count
Australia	Couple Leisure	35
Australia	Family Leisure	8
Australia	Solo Leisure	12
Canada	Business	1
Canada	Couple Leisure	20
Canada	Family Leisure	17
Canada	Solo Leisure	16
Germany	Business	5
Germany	Couple Leisure	31
Germany	Family Leisure	15
Germany	Solo Leisure	22
Ireland	Business	15
Ireland	Couple Leisure	68
Ireland	Family Leisure	25
Ireland	Solo Leisure	37
Italy	Business	7
Italy	Couple Leisure	13

4.9.15. Câu truy vấn 15

Nội dung câu truy vấn: Với những chuyến bay với hạng thương gia liệt kê quốc tịch khách hàng, hành trình, có mong muốn quay lại hay không

Thực hiện truy vấn:

Kéo thả thuộc tính [Passenger Country] [Origin] [Destination] [Recommended] [Fact Count] vào khung cửa sổ Excute

Passenger Country	Origin	Destination	Recommended	Fact Count
Austria	Salzburg	Stansted	no	1
Cyprus	Thessaloniki	Paphos	no	1
Greece	Knock	Athens via...	no	1
Ireland	Dublin	Copenhagen...	yes	1
Ireland	Dublin	Malta	no	1
Romania	Bologna	Murcia	no	1
United Kingdom	Bristol	Bucharest	no	1
United Kingdom	Corfu	Cologne	no	1
United Kingdom	Lanzarote	Southend	no	1
United Kingdom	LTN	Liverpool	no	1
United Kingdom	Stansted	EMA	no	1
United States	Edinburgh	Toulouse	no	1
United States	London Gatwick	Dublin	no	1

4.10.Thực hiện 15 câu truy vấn – Quá trình phân tích dữ liệu bằng ngôn ngữ truy vấn MDX

4.10.1. Câu truy vấn 1

Nội dung câu truy vấn: Thống kê trung bình điểm đánh giá tổng thể qua từng năm

Thực hiện câu truy vấn:

```
SELECT [Measures].[Avg_OverallRating] ON COLUMNS,  
[Dim Time].[Year Published].[Year Published] ON ROWS  
FROM [Review WH];
```

	Avg_OverallRating
2012	4.15
2013	5.87671232876712
2014	6.31858407079646
2015	5.33024691358025
2016	6.13218390804598
2017	4.03468208092486
2018	3.04545454545455
2019	2.79714285714286
2020	3.10909090909091
2021	2.625
2022	2.88095238095238
2023	3.49612403100775
2024	6.625
Unknown	(null)

4.10.2. Câu truy vấn 2

Nội dung câu truy vấn: Thống kê những hành trình trong năm 2023 có hành khách có quốc tịch United Kingdom tham gia

Thực hiện câu truy vấn:

```

SELECT NON EMPTY { [Measures].[Fact Count] } ON COLUMNS,
NON EMPTY { ([Dim
Route].[Route Id].[Route Id].ALLMEMBERS * [Dim
Route].[Origin].[Origin].ALLMEMBERS * [Dim
Route].[Destination].[Destination].ALLMEMBERS) }

ON ROWS
FROM
(SELECT ({ [Dim Time].[Year Published].&[2023] }) ON COLUMNS
FROM
(SELECT ({ [Dim Passenger Country].[Passenger
Country].&[United Kingdom] }) ON COLUMNS
FROM
(SELECT ({ [Dim Seat Type].[Seat Type].&[Economy Class] })
ON COLUMNS
FROM [Review WH])))
WHERE ([Dim Seat Type].[Seat Type].&[Economy Class],
[Dim Passenger Country].[Passenger Country].&[United
Kingdom],
[Dim Time].[Year Published].&[2023]);

```

			Fact Count
150	Biarritz	Brussels	1
158	Birmingham	Dublin	1
170	Bodrum	Dublin	1
202	Bristol	Dublin	2
274	Cardiff	Faro	1
285	Ciampino Rome	Manchester	1
302	Cologne Bonn	Stansted	1
337	Dortmund	London Stansted	1
353	Dublin	Bodrum	1
355	Dublin	Bristol	1
402	East Midlands	Berlin	1
416	Edinbrough	Tirana	1
424	Edinburgh	Fuerteventura	1
436	Edinburgh	Stansted	1

4.10.3. Câu truy vấn 3

Nội dung câu truy vấn: Thông kê những hành trình có hành khách có quốc tịch United States tham gia và có trung bình tổng thẻ ≤ 3

Thực hiện câu truy vấn:

```

SELECT
    NON EMPTY { [Measures].[Fact Count] } ON COLUMNS,
    NON EMPTY {
        (
            [Dim Route].[Route Id].[Route Id].ALLMEMBERS *
            [Dim Route].[Origin].[Origin].ALLMEMBERS *
            [Dim Route].[Destination].[Destination].ALLMEMBERS
        )
    } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON
ROWS
FROM
(
    SELECT
    (
        FILTER(
            [Dim Passenger Country].[Passenger Country].[United
States],
            [Measures].[Avg_OverallRating] <= 3
        )
    ) ON COLUMNS
) ON COLUMNS
FROM
[Review WH]);

```

			Fact Count
1			23
29	Alicante	Venice	1
34	Amsterdam	Malaga	1
36	Ancona Falconara	London Stansted	1
40	ATH	JMK	1
55	Athens	Rome	1
57	Athens	Santorini	2
69	Barcelona	Dublin	1
72	Barcelona	Frankfurt	2
73	Barcelona	Ibiza	1
78	Barcelona	Malaga	1
79	Barcelona	Mallorca	1
80	Barcelona	Manchester	1
83	Barcelona	Palma	1

4.10.4. Câu truy vấn 4

Nội dung câu truy vấn: Thống kê những hành trình có điểm trung bình dịch vụ ăn uống < 3 nhưng dịch vụ nhân viên >= 4

Thực hiện câu truy vấn:

```

SELECT
    NON EMPTY { [Measures].[Fact Count] } ON COLUMNS,
    NON EMPTY {
        (
            [Dim Route].[Route Id].[Route Id].ALLMEMBERS *
            [Dim Route].[Origin].[Origin].ALLMEMBERS *
            [Dim Route].[Destination].[Destination].ALLMEMBERS
        )
    } ON ROWS
FROM
(
    SELECT
    (
        FILTER(
            [Dim Route].[Route Id].CHILDREN,
            [Measures].[Avg_FoodBeverages] < 3 and
            [Measures].[Avg_CabinStaffService] >= 4
        )
    ) ON COLUMNS
    FROM
        [Review WH]
);

```

Messages Results

			Fact Count
5	ACE	LBA	1
11	ALC	EDI	1
15	Alicante	Dublin	1
19	Alicante	Kaunas	1
31	Amman	Paphos	1
35	Ancona	Stansted	1
36	Ancona Falconara	London Stansted	1
39	ATH	CRL	1
41	ATH	JTR	1
42	ATH	SKG	1
43	ATH	STN	1
46	Athens	Berlin SchÃ¶nefeld	1
48	Athens	Catania	1
84	Barcelona	Palma de Mallorca	2

4.10.5. Câu truy vấn 5

Nội dung câu truy vấn: Thống kê những hành trình từ năm 2019 – 2020 mà có khách hàng Australia tham gia và có mong muốn quay lại

Thực hiện câu truy vấn:

```

SELECT NON EMPTY { [Measures].[Fact Count] } ON COLUMNS,
NON EMPTY { ([Dim
Route].[Route Id].[Route Id].ALLMEMBERS * [Dim
Route].[Origin].[Origin].ALLMEMBERS * [Dim
Route].[Destination].[Destination].ALLMEMBERS) } DIMENSION
PROPERTIES MEMBER_CAPTION ON ROWS

FROM
    (SELECT ([Dim Time].[Year Published].&[2019] : [Dim Time].[Year
Published].&[2020]) ON COLUMNS
    FROM
        (SELECT ({ [Dim Return].[Recommended].&[yes] }) ON COLUMNS
        FROM
            (SELECT ({ [Dim Passenger Country].[Passenger
Country].&[Australia] }) ON COLUMNS
            FROM [Review WH])))
WHERE ([Dim Passenger Country].[Passenger Country].&[Australia],
[Dim Return].[Recommended].&[yes]);

```

			Fact Count
351	Dublin	Bergamo	1
835	Marrakech	Gatwick	1
852	Milan Bergamo	Brindisi	1

4.10.6. Câu truy vấn 6

Nội dung câu truy vấn: Thống kê những top 3 loại máy bay có điểm trung bình chỗ ngồi thoải mái cao nhất theo từng năm

Thực hiện câu truy vấn:

```

SELECT
    NON EMPTY { [Measures].[Avg_SeatComfort] } ON COLUMNS,
    NON EMPTY
        GENERATE(
            [Dim Time].[Year Published].CHILDREN,
            TOPCOUNT(
                [Dim Time].[Year Published].currentmember* [Dim
Aircraft].[Aircraft].CHILDREN,
                3,
                [Measures].[Avg_SeatComfort]
            )
        )
    ON ROWS
FROM
    [Review WH];

```

		Avg_SeatComfort
2012	Boeing 737-800	2.65
2013	Boeing 737-800	2.85616438356164
2014	Boeing 737-800	2.87315634218289
2015	Boeing 737-800	4.5
2015	737300	4
2015	A320	4
2016	B747	5
2016	Boeing 747-800	5
2016	Boeing 737 / 800	4
2017	Boeing 738	4
2017	Boeing 737	3.07142857142857
2017	B737-800	3
2018	B737-800	3.5
2018	A340	3

4.10.7. Câu truy vấn 7

Nội dung câu truy vấn: Thống kê top 3 loại hình du lịch khách hàng không muôn quay lại nhât nhưng trung bình giá trị đem lại > 2

Thực hiện câu truy vấn:

```

SELECT
    NON EMPTY { [Measures].[Fact Count],
    [Measures].[Avg_ValueForMoney] } ON COLUMNS,
    NON EMPTY {
        ORDER(
            FILTER(
                TOPCOUNT(
                    [Dim Travel Type].[Type Of Traveller].[Type Of
Traveller].MEMBERS,
                    3,
                    [Measures].[Fact Count]
                ),
                [Measures].[Avg_ValueForMoney] > 2 AND
                [Dim Return].[Recommended].&[No]
            ),
            [Measures].[Avg_ValueForMoney], DESC
        )
    } ON ROWS
FROM
    [Review WH];

```

	Fact Count	Avg_ValueForMoney
Couple Leisure	1201	2.80183180682764
Solo Leisure	548	2.7956204379562
Family Leisure	374	2.43850267379679

4.10.8. Câu truy vấn 8

Nội dung câu truy vấn: Thống kê top 5 quốc gia có số lượng hành khách tham gia nhiều nhất những năm 2019-2023

Thực hiện câu truy vấn:

```

SELECT NON EMPTY { [Measures].[Fact Count] } ON COLUMNS,
NON EMPTY { ([Dim
Passenger Country].[Passenger Country Id].[Passenger Country
Id].ALLMEMBERS * [Dim Passenger Country].[Passenger
Country].[Passenger Country].ALLMEMBERS) } ON ROWS
FROM
  (SELECT ([Dim Time].[Year Published]&[2019] : [Dim Time].[Year
Published]&[2023]) ON COLUMNS
  FROM
    (SELECT (TopCount
      (
        [Dim Passenger Country].[Passenger
Country].CHILDREN,
        5,
        [Measures].[Fact Count]
      ) ON COLUMNS
    )
  FROM [Review WH]));

```

Messages		Results
		Fact Count
22	Germany	38
31	Ireland	52
62	Spain	23
71	United Kingdom	308
72	United States	76

4.10.9. Câu truy vấn 9

Nội dung câu truy vấn: Thống kê những hành trình có trung bình dịch vụ mặt đất, dịch vụ ăn uống, chỗ ngồi thoải mái và dịch vụ nhân viên > 2.5 nhưng khách hàng không có nhu cầu quay lại

Thực hiện câu truy vấn:

```

SELECT NON EMPTY { [Measures].[Avg_CabinStaffService],
[Measures].[Avg_GroundService],
[Measures].[Avg_FoodBeverages],
[Measures].[Avg_SeatComfort] } ON COLUMNS,
NON EMPTY {
([Dim Route].[Origin].[Origin].ALLMEMBERS * [Dim
Route].[Destination].[Destination].ALLMEMBERS) } ON ROWS
FROM
(SELECT (FILTER([Dim Route].[Route Id].CHILDREN,
[Measures].[Avg_GroundService] > 2.5 AND
[Measures].[Avg_FoodBeverages] > 2.5 AND
[Measures].[Avg_SeatComfort] >2.5 AND
[Measures].[Avg_CabinStaffService] > 2.5
)) ON COLUMNS
FROM
(SELECT ({ [Dim Return].[Recommended].&[no] }) ON COLUMNS
FROM [Review WH]))
WHERE ([Dim Return].[Recommended].&[no]);

```

Messages Results

		Avg_CabinStaffService	Avg_GroundService	Avg_FoodBeverages	Avg_SeatComfort
Athens	Frankfurt	4	4	3	5
Rome Fiumicino	Dublin	5	4	5	5

4.10.10. Câu truy vấn 10

Nội dung câu truy vấn: Theo từng loại hình du lịch liệt kê các quốc gia, trung bình giá trị đem lại và trung bình tổng thể

Thực hiện câu truy vấn:

```
SELECT NON EMPTY { [Measures].[Avg_ValueForMoney],
[Measures].[Avg_OverallRating] } ON COLUMNS,
NON EMPTY { ([Dim Travel Type].[Type Of Traveller].[Type Of Traveller].ALLMEMBERS * [Dim Passenger Country].[Passenger Country].[Passenger Country].ALLMEMBERS) } ON ROWS
FROM [Review WH];
```

Messages Results

		Avg_ValueForMoney	Avg_OverallRating
Business	Albania	5	8
Business	Austria	2.33333333333333	3.33333333333333
Business	Belgium	3	4.75
Business	Bulgaria	2	1
Business	Canada	4	5
Business	Cyprus	1	1
Business	Czech Republic	1	1
Business	Estonia	1	1
Business	France	1	1
Business	Germany	2.4	3.4
Business	Greece	3	4.5
Business	Hungary	3	2
Business	India	1	1
Business	Indonesia	5	6

4.10.11.Câu truy vấn 11

Nội dung câu truy vấn: Thông kê top 10 máy bay được khai thác nhiều nhất năm 2019

Thực hiện câu truy vấn:

```
SELECT NON EMPTY { [Measures].[Fact Count] } ON COLUMNS,
NON EMPTY { (
    TopCount(
        [Dim Aircraft].[Aircraft].[Aircraft].MEMBERS,
        10,
        [Measures].[Fact Count]
    )
) } ON ROWS
FROM [Review WH]
WHERE [Dim Time].[Year Published].[2019];
```

The screenshot shows the SSMS interface with two tabs: 'Messages' and 'Results'. The 'Results' tab contains a table titled 'Fact Count' with the following data:

	Fact Count
Boeing 737-800	300
Boeing 737	33
Boeing 747	3
Boeing 737 800	2
Boeing 738	2
A320	1
B737	1
B737 800	1
Boeing 737-800	1
Boeing 373-800	1

4.10.12.Câu truy vấn 12

Nội dung câu truy vấn: Thông kê top 5 máy bay khách hàng có nhu cầu quay lại nhiều nhất

Thực hiện câu truy vấn:

```
SELECT NON EMPTY { [Measures].[Fact Count] } ON COLUMNS,
NON EMPTY { (
    TOPCOUNT(
        [Dim Aircraft].[Aircraft].[Aircraft].MEMBERS *
        FILTER(
            [Dim Return].[Recommended].MEMBERS,

```

```

        [Dim Return].[Recommended].CurrentMember IS [Dim
Return].[Recommended].&[yes]
    ),
    5,
    [Measures].[Fact Count]
)
) } ON ROWS
FROM [Review WH];

```

		Fact Count
Boeing 737-800	yes	741
Boeing 737	yes	71
Boeing 737 800	yes	8
B737-800	yes	6
Boeing 738	yes	6

4.10.13. Câu truy vấn 13

Nội dung câu truy vấn: Thông kê các hành trình với loại hình là gia đình, khách hàng có nhu cầu quay lại, có trung bình thoải mái chỗ ngồi, dịch vụ ăn uống và giá trị mang lại ≥ 4

Thực hiện câu truy vấn:

```

SELECT NON EMPTY {
[Measures].[Avg_SeatComfort], [Measures].[Avg_FoodBeverages],
[Measures].[Avg_ValueForMoney] } ON COLUMNS,
NON EMPTY { (
    FILTER(
        [Dim Route].[Origin].[Origin].MEMBERS * [Dim
Route].[Destination].[Destination].MEMBERS * [Dim
Return].[Recommended].MEMBERS * [Dim Travel Type].[Type Of
Traveller].&[Family Leisure],
        [Measures].[Avg_SeatComfort] >= 4 AND
        [Measures].[Avg_ValueForMoney] >= 4 AND
        [Measures].[Avg_FoodBeverages] >= 4 AND
        [Dim Return].[Recommended].CurrentMember IS [Dim
Return].[Recommended].&[yes]
    )
}

```

```

        )
    ) } ON ROWS
FROM [Review WH];

```

				Avg_SeatComfort	Avg_FoodBeverages	Avg_ValueForMon
ALC	EMA	yes	Family Leisure	5	5	5
Athens	Stansted	yes	Family Leisure	4	5	5
BCN	TSF	yes	Family Leisure	4	5	5
Birmingham	Palma	yes	Family Leisure	5	5	5
Bologna	Bucharest	yes	Family Leisure	4	4	5
Bologna	Stansted	yes	Family Leisure	4	4	5
Bratislava	Dublin	yes	Family Leisure	4	4	5
Brussels	Gran Canaria	yes	Family Leisure	4	5	5
Dublin	Manchester	yes	Family Leisure	5	5	5
East Midlands	Bergerac	yes	Family Leisure	4	5	5
Frankfurt-Hahn	Rome Ciampino	yes	Family Leisure	4	4	5
Krak&ow	Stansted	yes	Family Leisure	5	5	5
LTN	ACE	yes	Family Leisure	5	5	5

4.10.14. Câu truy vấn 14

Nội dung câu truy vấn: Với top 10 quốc gia có số lượng khách hàng tham gia nhiều nhất liệt kê số lượng chuyến bay theo từng loại hình du lịch

Thực hiện câu truy vấn:

```

SELECT NON EMPTY { [Measures].[Fact Count] } ON COLUMNS,
NON EMPTY { (
TOPCOUNT(
    [Dim Passenger Country].[Passenger Country].[Passenger
Country].MEMBERS,
    10,
    [Measures].[Fact Count]
) *
    [Dim Travel Type].[Type Of Traveller].[Type Of
Traveller].MEMBERS
) } ON ROWS
FROM [Review WH];

```

Messages Results

		Fact Count
United Kingdom	Business	40
United Kingdom	Couple Leisure	732
United Kingdom	Family Leisure	190
United Kingdom	Solo Leisure	205
United States	Business	6
United States	Couple Leisure	73
United States	Family Leisure	42
United States	Solo Leisure	46
Ireland	Business	15
Ireland	Couple Leisure	68
Ireland	Family Leisure	25
Ireland	Solo Leisure	37
Germany	Business	5
Germany	Couple Leisure	31

4.10.15. Câu truy vấn 15

Nội dung câu truy vấn: Với những chuyến bay với hạng thương gia liệt kê quốc tịch khách hàng, hành trình, có mong muốn quay lại hay không

Thực hiện câu truy vấn:

```

SELECT
    NON EMPTY {
        [Dim Passenger Country].[Passenger Country].[Passenger
Country].MEMBERS *
        [Dim Route].[Origin].[Origin].MEMBERS *
        [Dim Route].[Destination].[Destination].MEMBERS *
        [Dim Return].[Recommended].[Recommended].MEMBERS
    } ON ROWS,
    NON EMPTY { [Measures].[Fact Count] } ON COLUMNS
FROM [Review WH]
WHERE (
    [Dim Seat Type].[Seat Type].&[Premium Economy]
);

```

				Fact Count
Austria	Salzburg	Stansted	no	1
Cyprus	Thessaloniki	Paphos	no	1
Greece	Knock	Athens via Stansted	no	1
Ireland	Dublin	Copenhagen	yes	1
Ireland	Dublin	Malta	no	1
Ireland	Dublin	Murcia	no	1
Romania	Bologna	Bucharest	no	1
United Kingdom	Bristol	Cologne	no	1
United Kingdom	Corfu	Southend	no	1
United Kingdom	Lanzarote	Liverpool	no	1
United Kingdom	LTN	LPA	no	1
United Kingdom	Stansted	Toulouse	no	1
United Kingdom	SXF	EMA	no	1
United States	Edinburgh	London Stansted	no	1

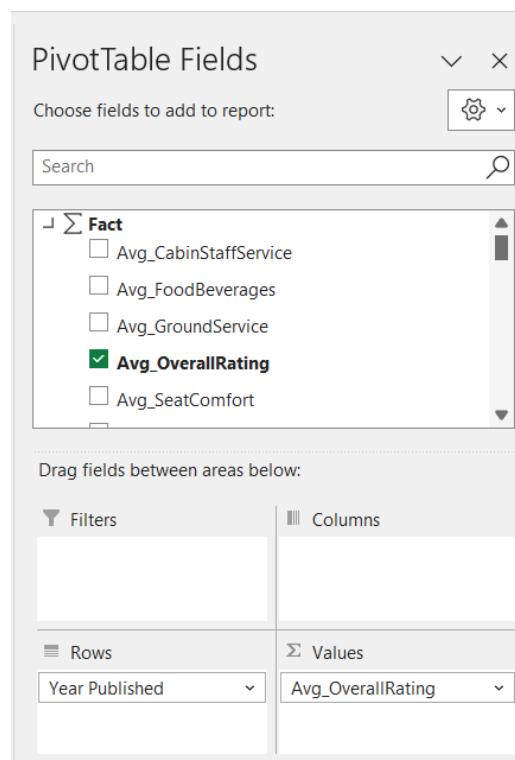
4.11.Thực hiện Report bằng Excel

4.11.1. Câu truy vấn 1

Nội dung câu truy vấn: Thông kê trung bình điểm đánh giá tổng thể qua từng năm

Thực hiện câu truy vấn:

Bước 1: Chọn thuộc tính [Year Published] cho Rows và [Avg_OverallRating] cho Values



Kết quả:

Row Labels	Avg_OverallRating
2012	4.15
2013	5.876712329
2014	6.318584071
2015	5.330246914
2016	6.132183908
2017	4.034682081
2018	3.045454545
2019	2.797142857
2020	3.109090909
2021	2.625
2022	2.880952381
2023	3.496124031
2024	6.625
Grand Total	4.358381503

4.11.2. Câu truy vấn 2

Nội dung câu truy vấn: Thống kê những hành trình trong năm 2023 có hành khách có quốc tịch United Kingdom tham gia

Thực hiện câu truy vấn:

Bước 1: Thêm thuộc tính [Published Year] [Passenger Country] vào Filter để lọc theo điều kiện

Thêm các thuộc tính của bảng Dim_Route vào Rows và thêm Fact Count vào Values

PivotTable Fields

Choose fields to add to report:

Search 

↳ \sum Fact

- Avg_CabinStaffService
- Avg_FoodBeverages
- Avg_GroundService
- Avg_OverallRating
- Avg_SeatComfort

Drag fields between areas below:

Filters	Columns
Published Time	
Passenger Country	

Rows	Values
Route ID	\sum Fact Count
Origin	
Destination	

Kết quả:

Published Time	2023	▼
Passenger Country	United Kingdom	▼
Row Labels		▼ Fact Count
150		
Biarritz		
Brussels	1	
158		
Birmingham		
Dublin	1	
170		
Bodrum		
Dublin	1	
201		
Bristol		
Cologne	1	
202		
Bristol		
Dublin	2	
274		
Cardiff		
Faro	1	
285		
Ciampino Rome		
Manchester	1	
302		

4.11.3. Câu truy vấn 3

Nội dung câu truy vấn: Thống kê những hành trình có hành khách có quốc tịch United States tham gia và có trung bình tổng thẻ ≤ 3

Thực hiện câu truy vấn:

Bước 1: Thêm các thuộc tính của bảng Dim_Route vào Rows và thêm Fact Count vào Values

Thêm Named Set US [Avg_OverallRating <= 3] vào Rows

PivotTable Fields

Choose fields to add to report:

Search 

Avg_ValueForMoney
 Cabin Staff Service
 Fact Count
 Food Beverages
 Ground Service
 Overall Rating

Drag fields between areas below:

Filters	Columns

Rows	Values
US Avg_OverallRati... Route ID Origin	Fact Count

Kết quả:

Row Labels	Fact Count
United States	
1	23
29	1
Alicante	
Venice	1
34	1
Amsterdam	
Malaga	1
36	1
Ancona Falconara	
London Stansted	1
40	1
ATH	
JMK	1
55	1
Athens	
Rome	1
57	2
Athens	
Santorini	2
69	1
Barcelona	
Dublin	1

4.11.4. Câu truy vấn 4

Nội dung câu truy vấn: Thông kê những hành trình có điểm trung bình dịch vụ ăn uống < 3 nhưng dịch vụ nhân viên ≥ 4

Thực hiện câu truy vấn:

Bước 1: Thêm các thuộc tính [Origin] [Destination] vào Rows và thêm Fact Count vào Values

Thêm Named Set US [Avg_FoodBeverages < 3 Avg_CabinStaffService >= 4]
vào Rows

PivotTable Fields

Choose fields to add to report:

Search 

Cabin Staff Service
 Fact Count
 Food Beverages
 Ground Service
 Overall Rating
 Seat Comfort

Drag fields between areas below:

Filters

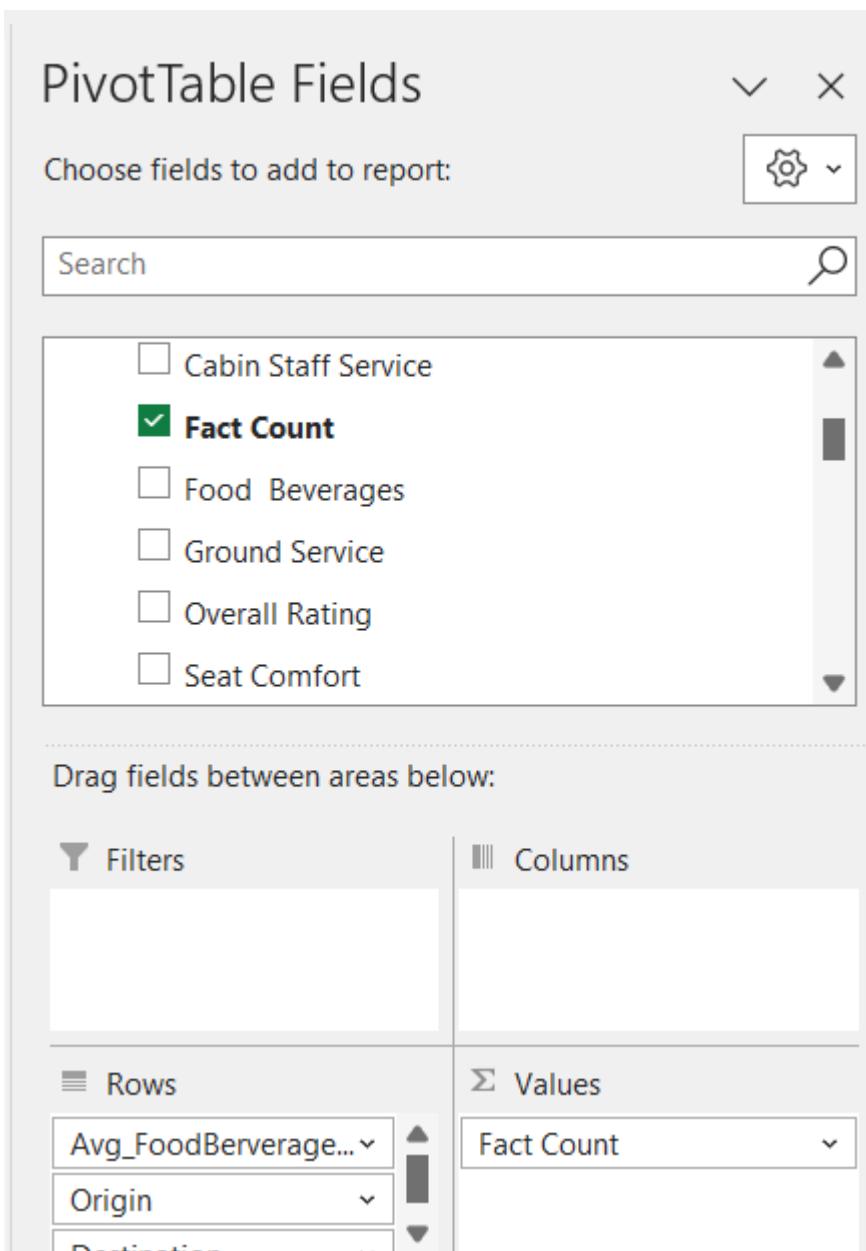
Rows

Avg_FoodBeverage...
Origin
Destination

Columns

Values

Fact Count



Kết quả:

Row Labels	Fact Count
5	
▫ ACE	
LBA	1
11	
▫ ALC	
EDI	1
15	
▫ Alicante	
Dublin	1
19	
▫ Alicante	
Kaunas	1
31	
▫ Amman	
Paphos	1
35	
▫ Ancona	
Stansted	1
36	
▫ Ancona Falconara	
London Stansted	1
39	
▫ ATH	
CRL	1
41	

4.11.5. Câu truy vấn 5

Nội dung câu truy vấn: Thống kê những hành trình từ năm 2019 – 2020 mà có khách hàng Australia tham gia và có mong muốn quay lại

Thực hiện câu truy vấn:

Bước 1: Chọn thuộc tính [Year Published] [Passenger Country] [Recommended] vào Filter để lọc theo điều kiện tương ứng

Bước 2: Thêm các thuộc tính của Dim_Route vào Rows và Fact Count cho Values

PivotTable Fields

Choose fields to add to report:

Search 

Avg_ValueForMoney
 Cabin Staff Service
 Fact Count
 Food Beverages
 Ground Service
 Overall Rating

Drag fields between areas below:

Filters

Published Time
Passenger Country
Recommended ..

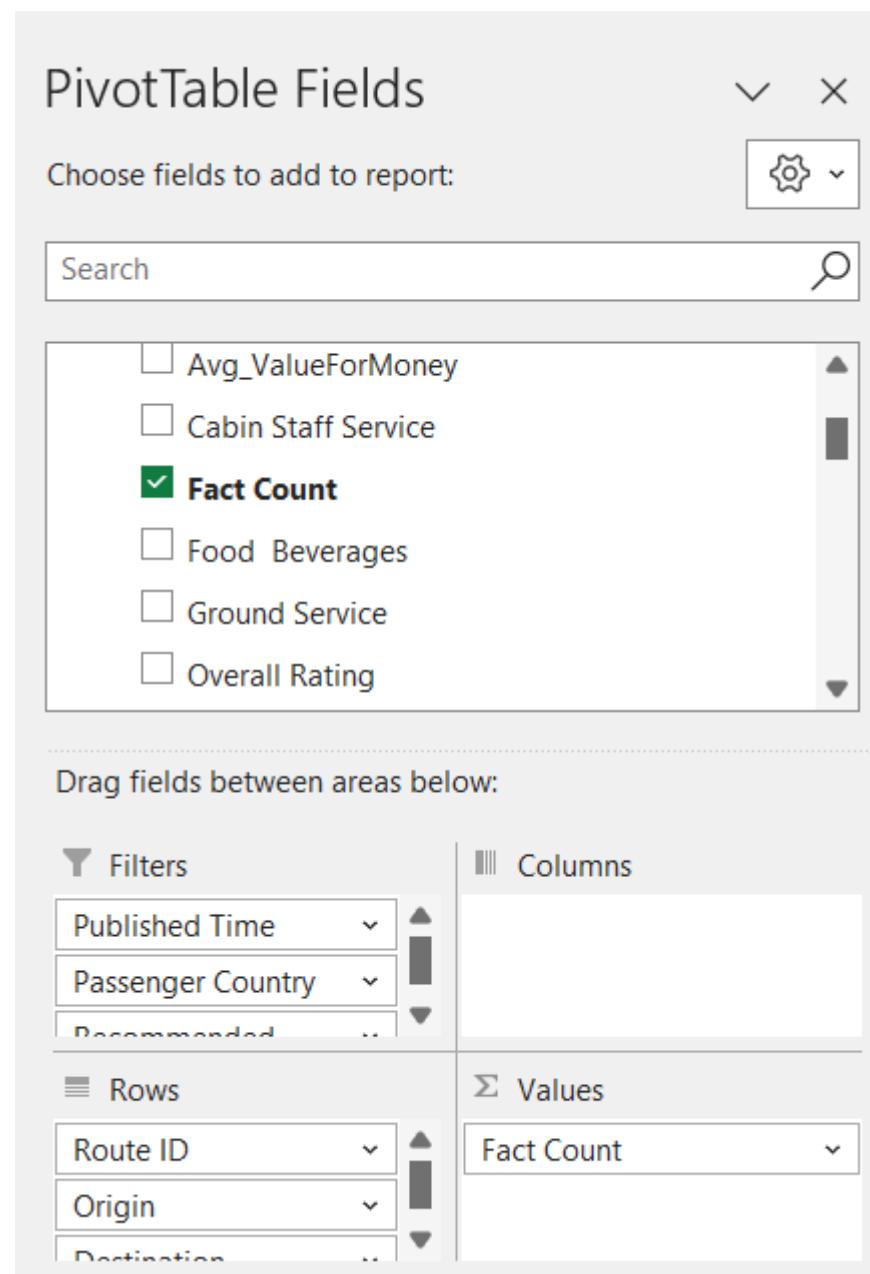
Columns

Rows

Route ID
Origin
Destination ..

Values

Fact Count



Kết quả:

Published Time	(Multiple Items)	<input checked="" type="checkbox"/>
Passenger Country	Australia	<input checked="" type="checkbox"/>
Recommended	yes	<input checked="" type="checkbox"/>
Row Labels		Fact Count
<input checked="" type="checkbox"/> 351		
<input checked="" type="checkbox"/> Dublin		
Bergamo	1	
<input checked="" type="checkbox"/> 835		
<input checked="" type="checkbox"/> Marrakech		
Gatwick	1	
<input checked="" type="checkbox"/> 852		
<input checked="" type="checkbox"/> Milan Bergamo		
Brindisi	1	
Grand Total		3

4.11.6. Câu truy vấn 6

Nội dung câu truy vấn: Thông kê top 3 loại hình du lịch khách hàng không muôn quay lại nhặt nhưng trung bình giá trị đem lại > 2

Thực hiện câu truy vấn:

Chọn Named Set [Top 3 TypeOfTraveller Avg_ValueForMoney > 2 Recommended No] cho Rows

Chọn Fact Count và Avg_ValueForMoney cho Values

PivotTable Fields

Choose fields to add to report:

Avg_OverallRating
 Avg_SeatComfort
 Avg_ValueForMoney
 Cabin Staff Service
 Fact Count
 Food Beverages

Drag fields between areas below:

<input type="checkbox"/> Filters	<input type="checkbox"/> Columns
	Σ Values
<input type="checkbox"/> Rows	Σ Values
Top 3 TypeOfTraveller ...	Fact Count
	Avg_ValueForMoney

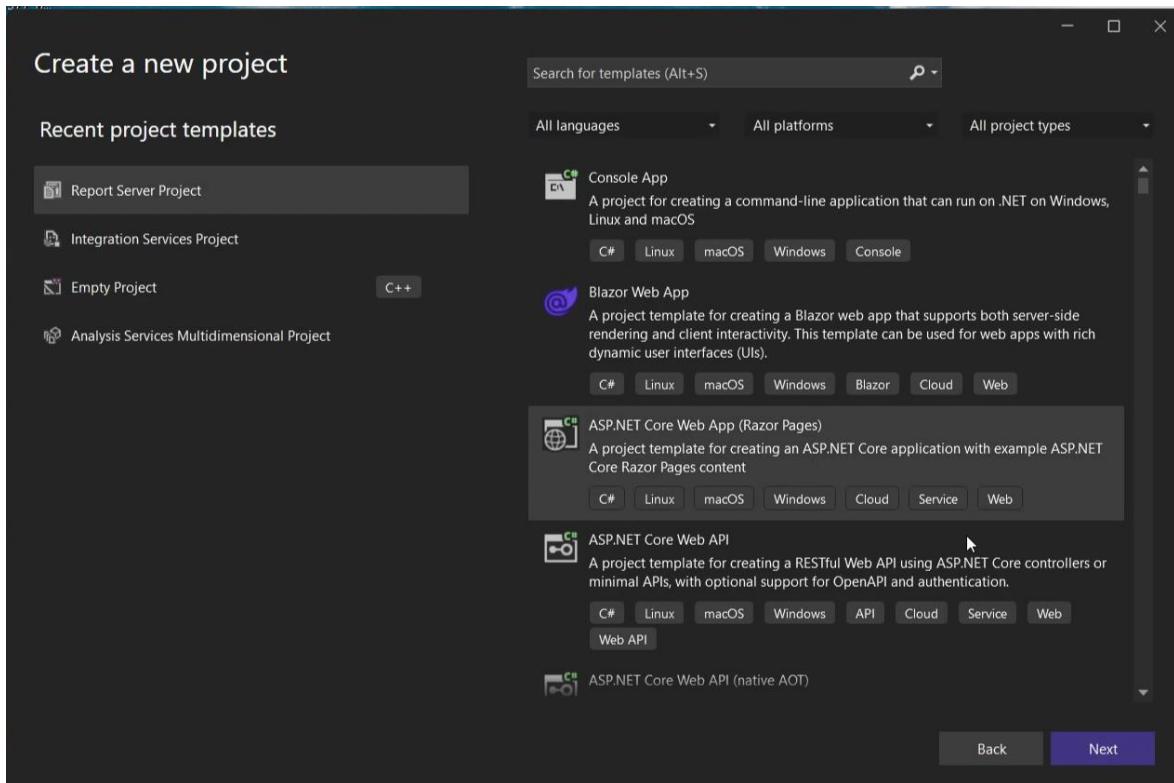
Kết quả:

Row Labels	Fact Count	Avg_ValueForMoney
Couple Leisure	1201	2.801831807
Family Leisure	374	2.438502674
Solo Leisure	548	2.795620438

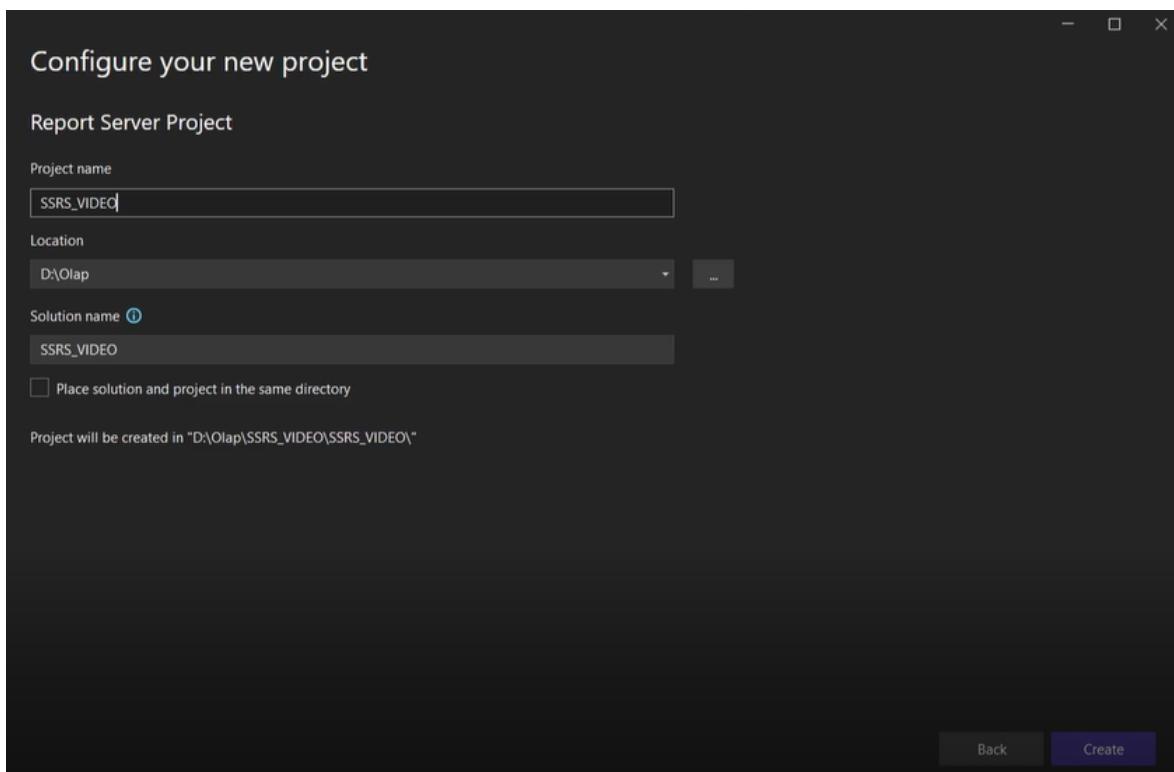
Chương 5. QUÁ TRÌNH LẬP BÁO BIỂU (SSRS)

5.1. Report với Visual Studio

Bước 1: Tạo mới một project với dịch vụ **Report Service Project** của **Visual Studio** và nhấn **Next**.



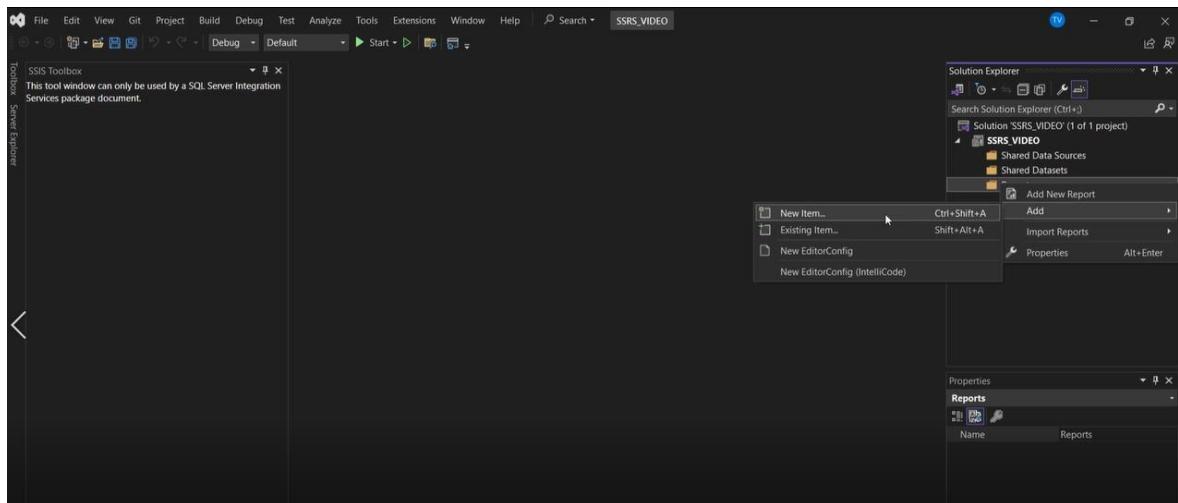
Bước 2: Nhập tên của project và chọn **Create**.



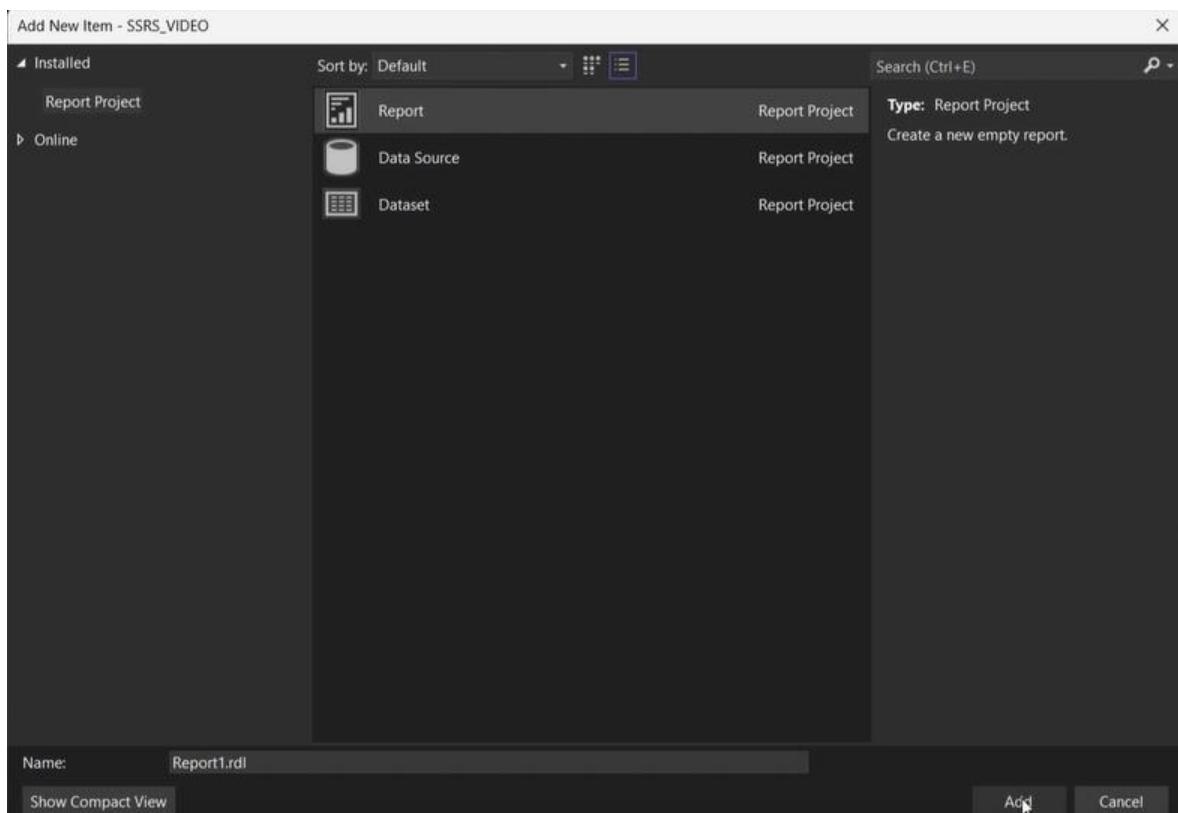
5.1.1. Thực hiện thống kê số lượng các tuyến bay.

Thực hiện tạo mới một báo cáo bằng bảng.

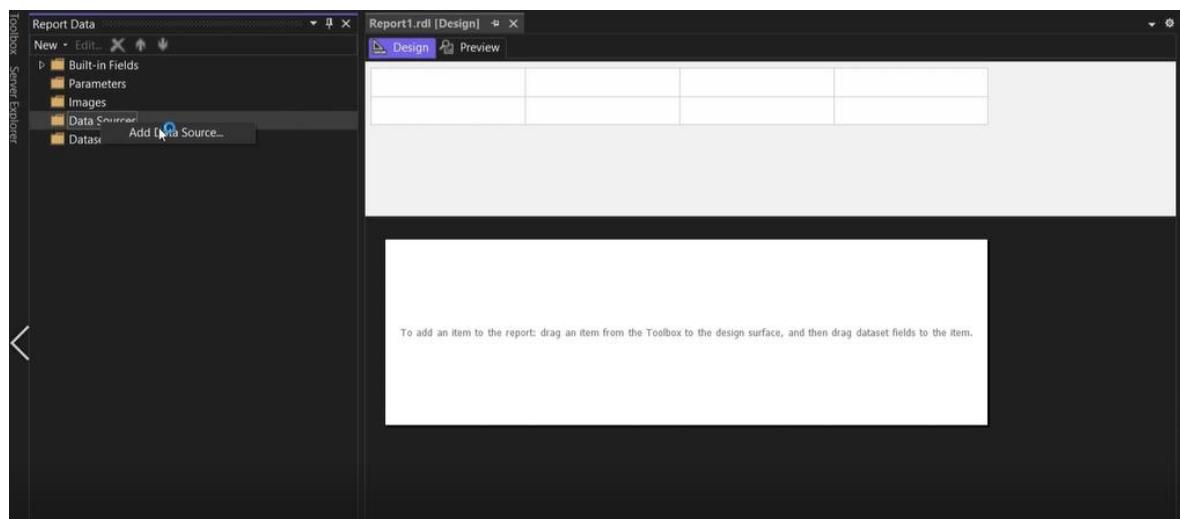
Bước 1: Tại **Solution Explorer**, nhấp chuột phải vào mục **Report** và chọn **New Item**.



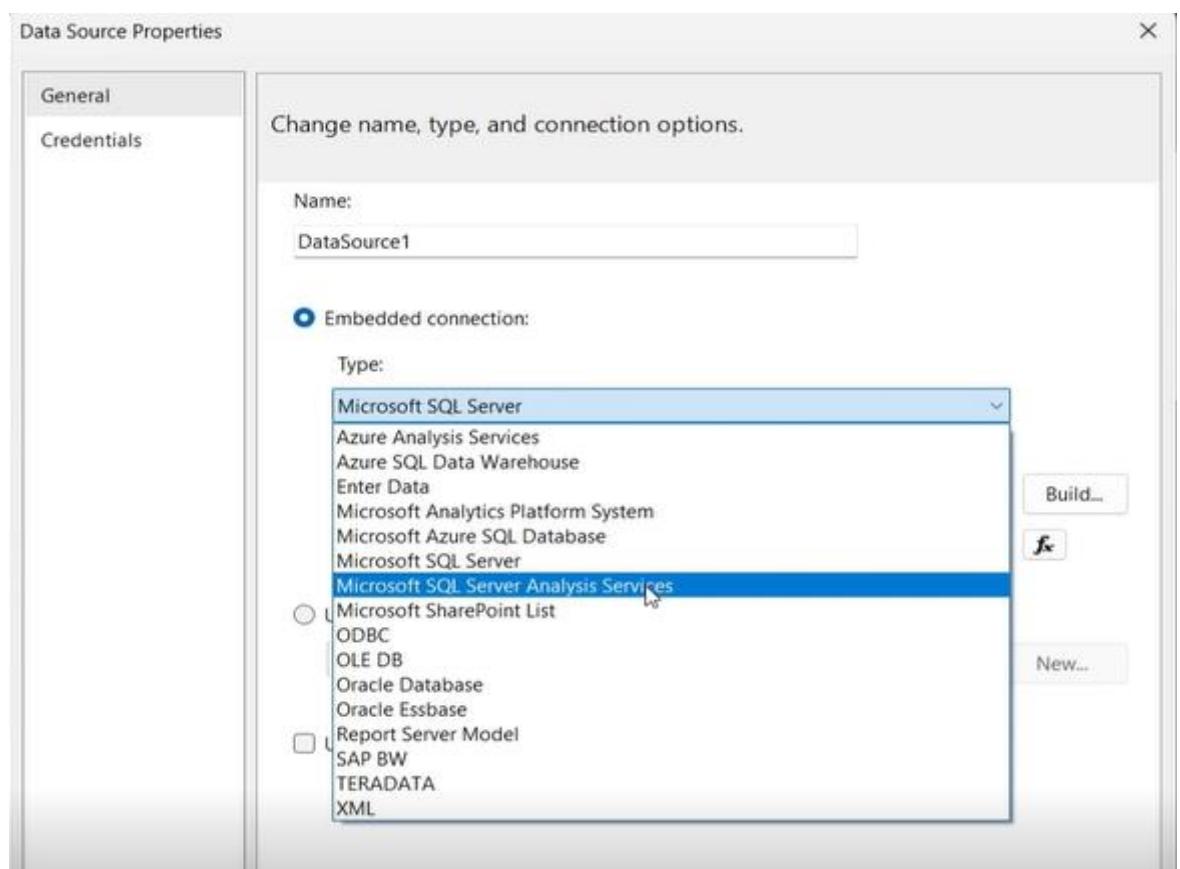
Bước 2: Xuất hiện hộp thoại chọn vào **Report** và nhấn **Add**.



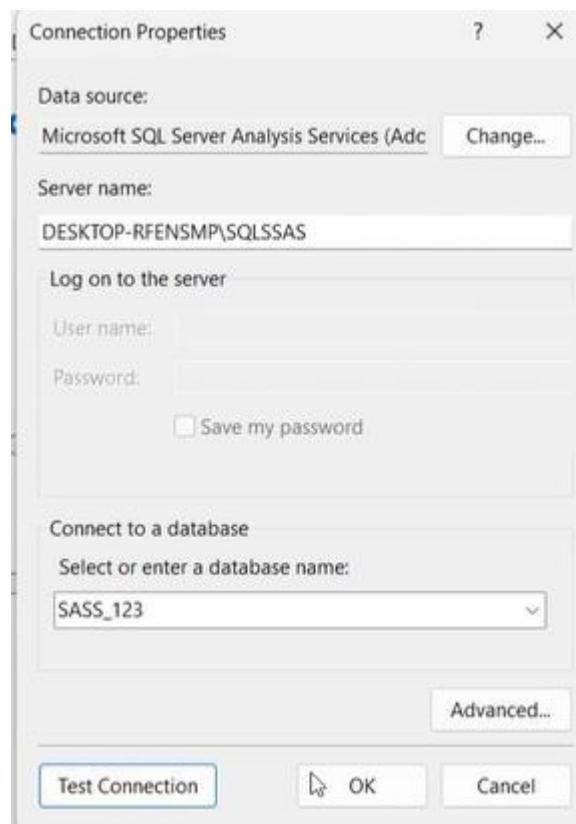
Bước 3: Tại cửa sổ **Report Data**, chuột phải vào **Data Source** chọn **Add Data Source...**



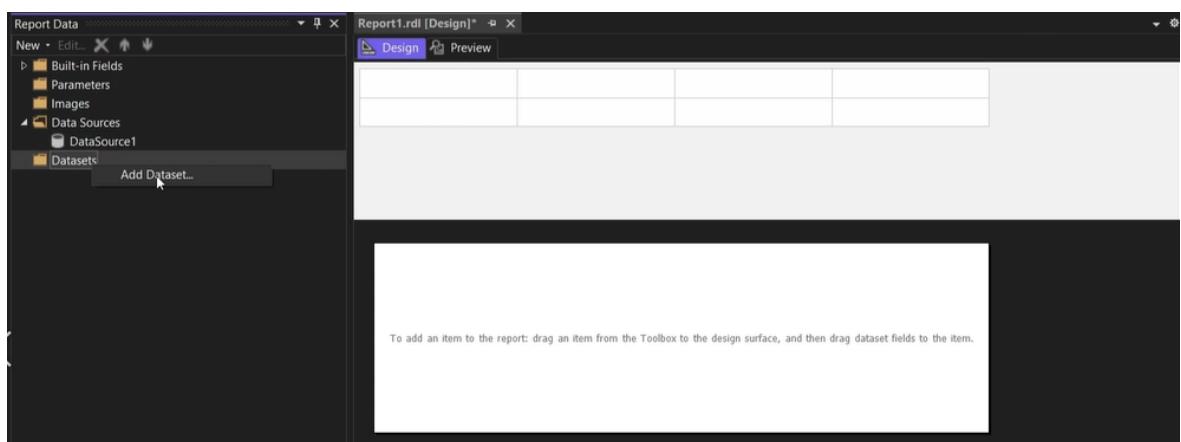
Bước 4: Trong hộp thoại **Data Source Properties**, tại mục **Type** chọn **Microsoft SQL Server Analysis Services**.



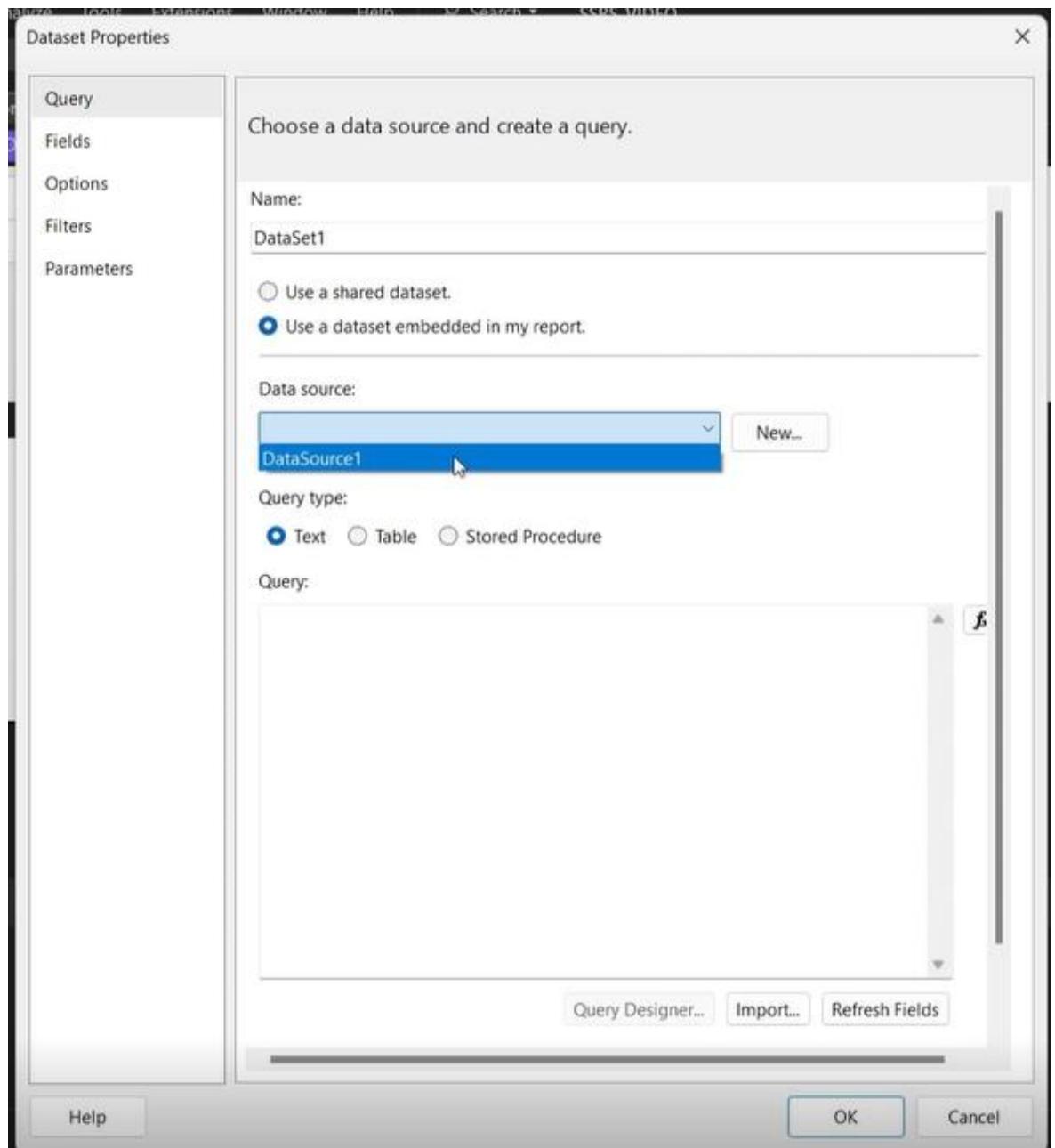
Bước 5: Tiếp tục nhấn vào nút **Builds** và tiến hành điền Server Name và chọn database và nhấn **OK**.



Bước 6: Nhấn chuột phải vào mục **Datasets** và chọn **Add Dataset...**



Bước 7: Click chọn vào mục **Use a dataset embedded im my report** và chọn **Data Source**.



Bước 8: Chọn **Query Designer**, và thực hiện kéo thả các thuộc tính cần thiết cho câu truy vấn tương tự như thực hiện truy vấn manumal. Sau đó nhấn **OK**.

Query Designer

Edit as Text Import... MD Dimension Hierarchy Operator Filter Expression Param...

AIRCRAFT WAREHOUSE

Metadata

Search Model Measure Group: <All>

AIRCRAFT WAREHOUSE

- Measures
 - Fact
 - Avg_Cabin_Staff_Satisfaction
 - Avg_Food_Beverage
 - Avg_Ground_Service
 - Avg_Overall_Rating
 - Avg_Seat_Comfort
 - Fact_Count
 - Sum_Value_For_Month
 - AVG_Route
 - tilediem
- KPIs
- Dim Air Craft
- Dim Passenger Country
- Dim Return
- Dim Route
- 10
- cau3
- cau4

Calculated Members

Dimension Hierarchy Operator Filter Expression Param...

<Select dimension>

Origin	Destination	Fact_Count
		615
Aberdeen	Faro	1
AC	MAN	1
ACE	BRS	1
ACE	LBA	1
AGP	BRU	1
AGP	MAN	1
AGP	PMI	1
AGP	STN	1
ALC	BHX	1
ALC	EDI	1
ALC	EMA	2
Alicante	Birmingham	1
Alicante	Brussels	1
Alicante	Dublin	1
Alicante	East Midlands	1
Alicante	Edinburgh	1
Alicante	Gatwick	1
Alicante	Kaunas	1
Alicante	Leeds	1
Alicante	Liverpool	2
Alicante	London	1
Alicante	London Stansted	1

Help OK Cancel

Bước 9: Vào mục ToolBox chọn Table và kéo thả vào vùng thiết kế.

Toolbox

Report Items

- Pointer
- Text Box
- Line
- Table
- Matrix
- Rectangle
- List
- Image
- Subreport
- Chart
- Gauge
- Map
- Data Bar
- Sparkline
- Indicator

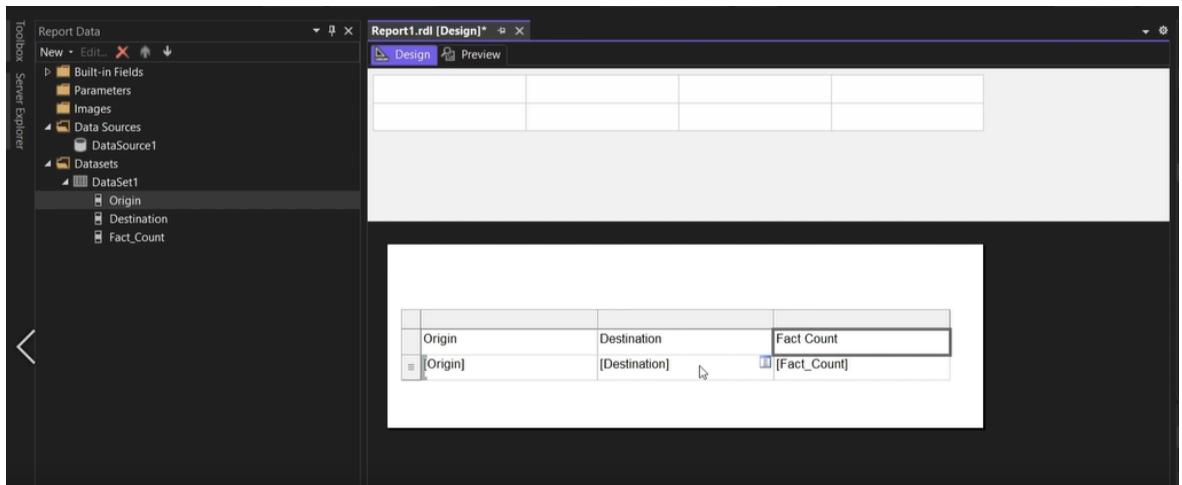
General

There are no usable controls in this group. Drag an item onto this text to add it to the toolbox.

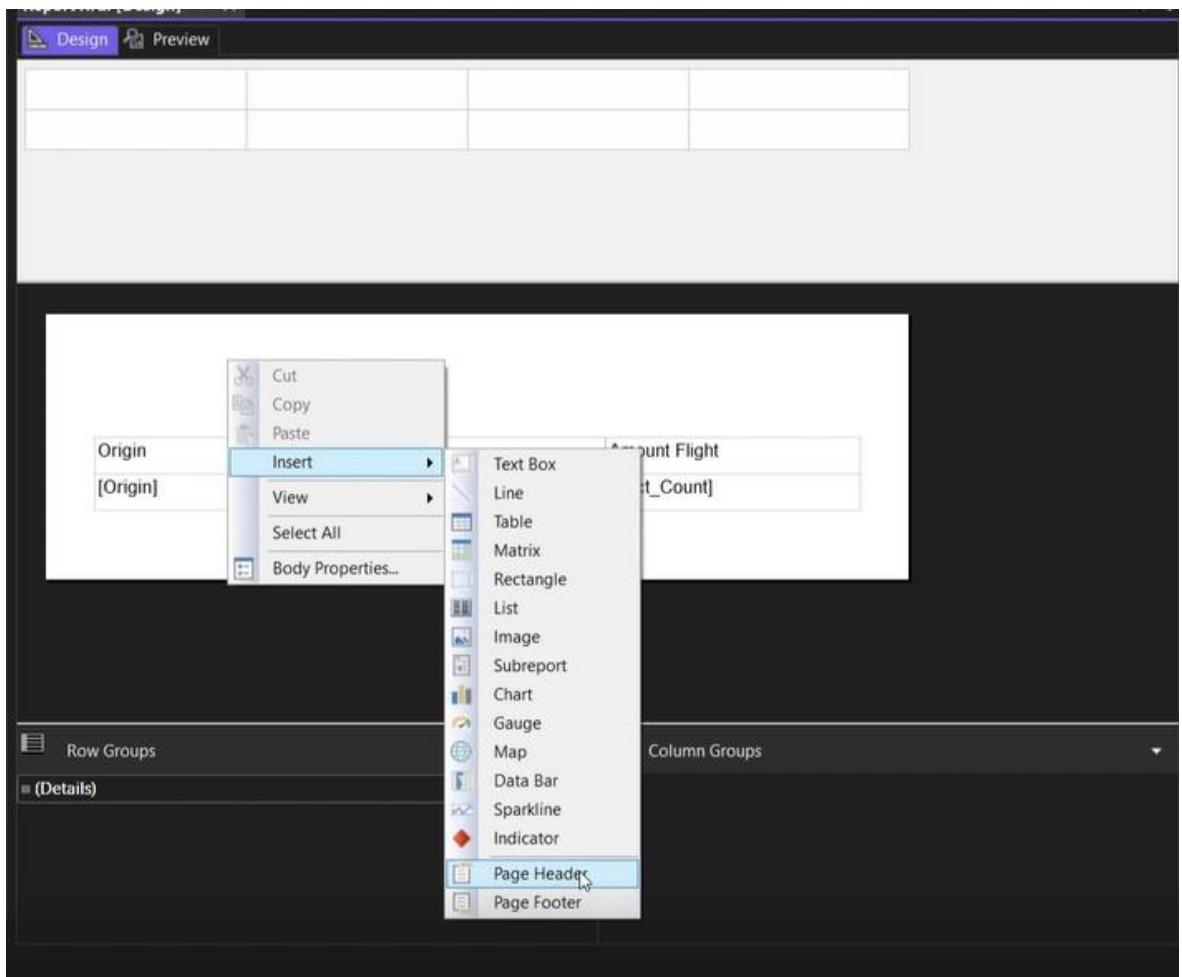
Report1.rdl [Design] Design Preview

To add an item to the report: drag an item from the Toolbox to the design surface, and then drag dataset fields to the item.

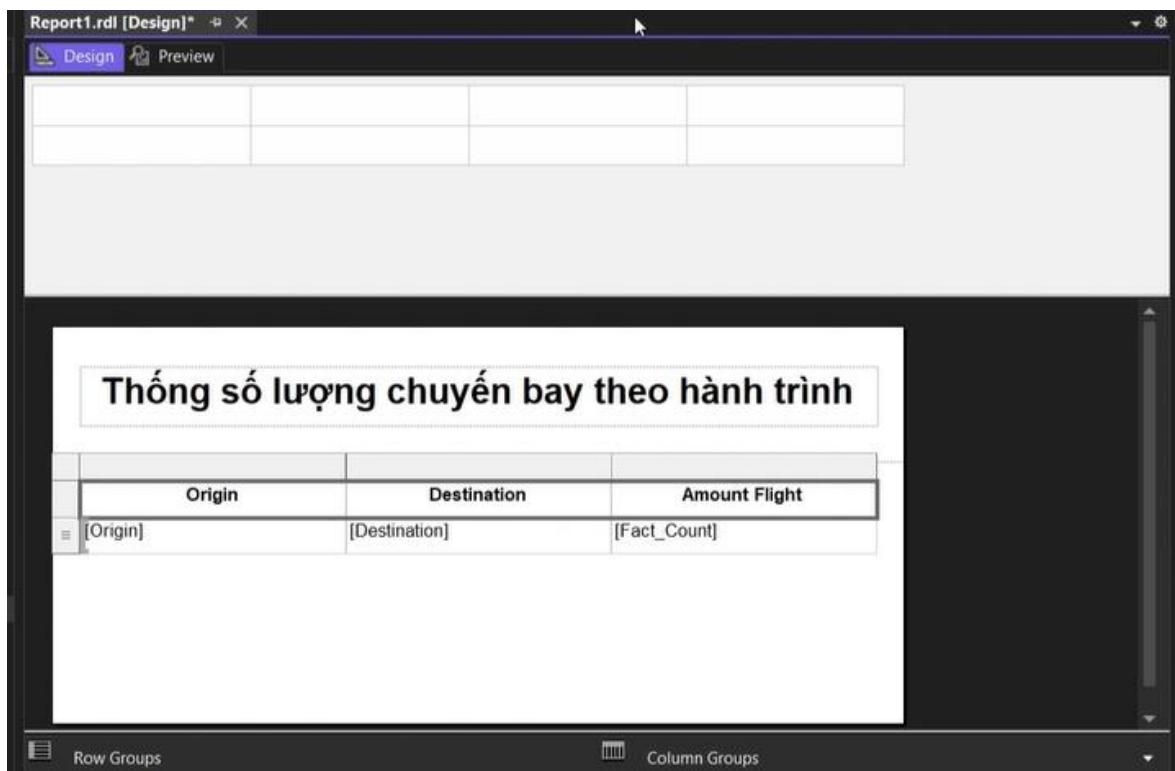
Bước 10: Lần lượt kéo thả các thuộc tính của câu truy vấn vào bảng.



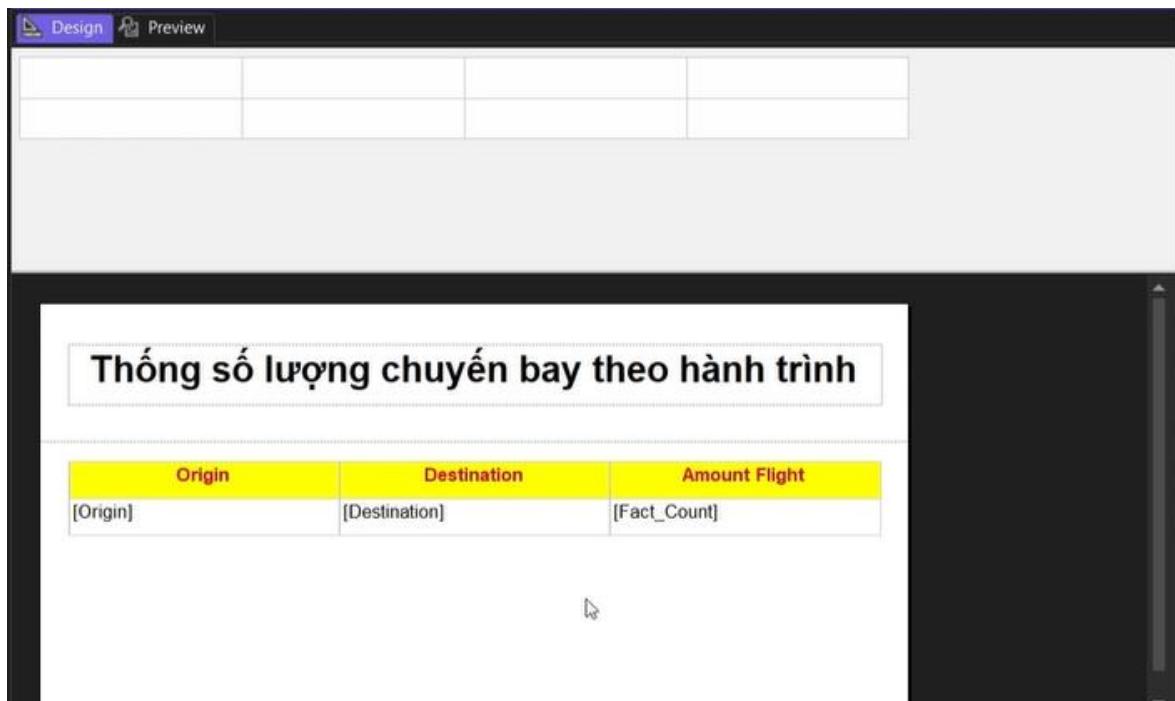
Bước 11: Click chuột phải vào vùng thiết kế chọn **Insert** và chọn **Page Header** để tạo header cho báo cáo.



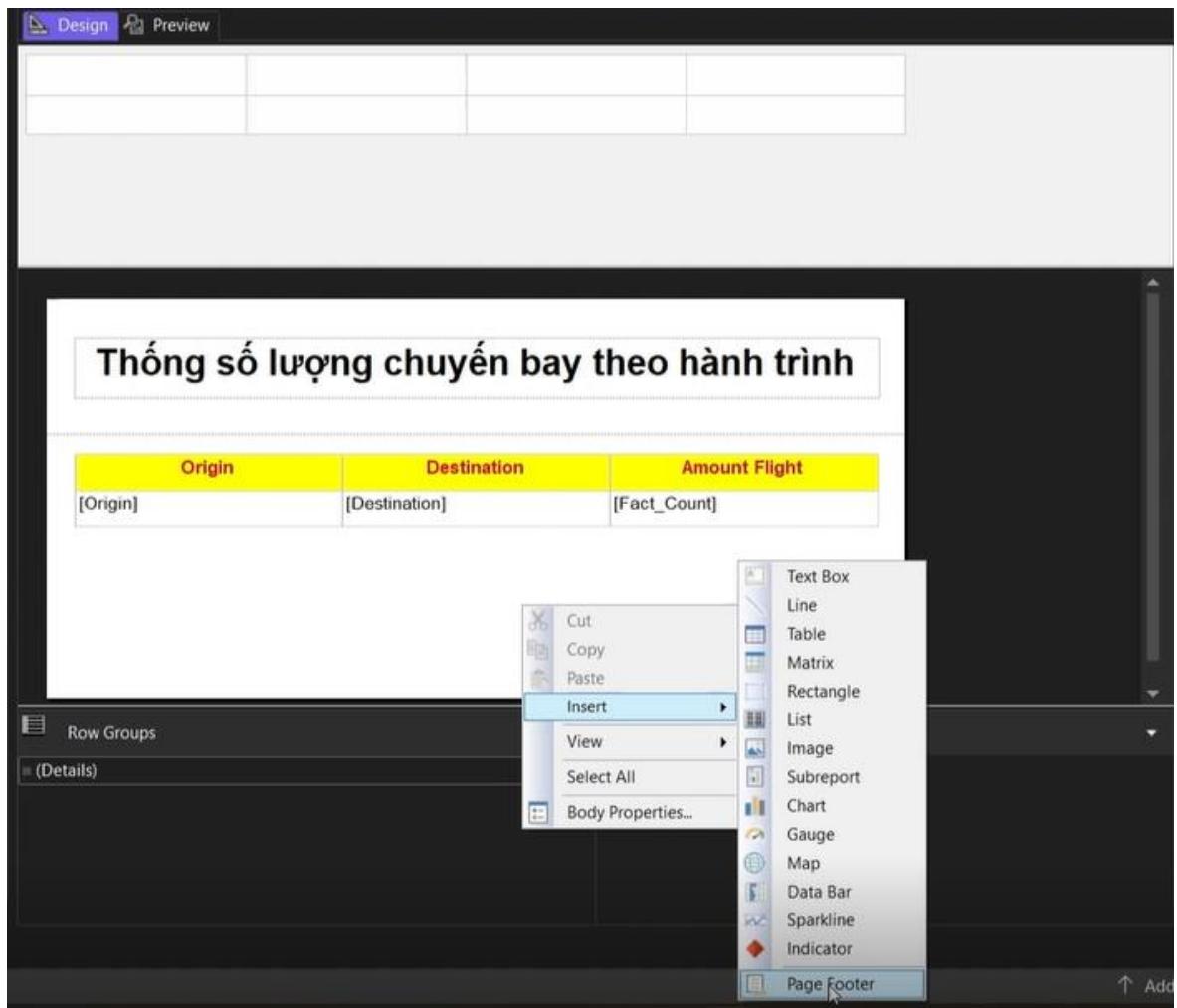
Bước 12: Tại ToolBox kéo thả một TextBox vào phần header và thực hiện điền tiêu đề của báo cáo.



Bước 13: Thực hiện định dạng các cột cho báo cáo màu chữ, cỡ chữ, căn lề.



Bước 14: Click chuột phải chọn **Insert** và chọn **Page Footer** để tạo đánh dấu số trang cho báo cáo.



Bước 16: Tại **ToolBox** kéo thả một **TextBox** vào phần footer và gõ **Page[&PageNumber]/[&TotalPages]**.

Origin	Destination	Amount Flight
[Origin]	[Destination]	 [Fact_Count]

Page
[&PageNumber]
/
[&TotalPages]

Bước 17: Nhấn vào tab Preview để xem báo cáo.

Thống số lượng chuyến bay theo hành trình
First Page

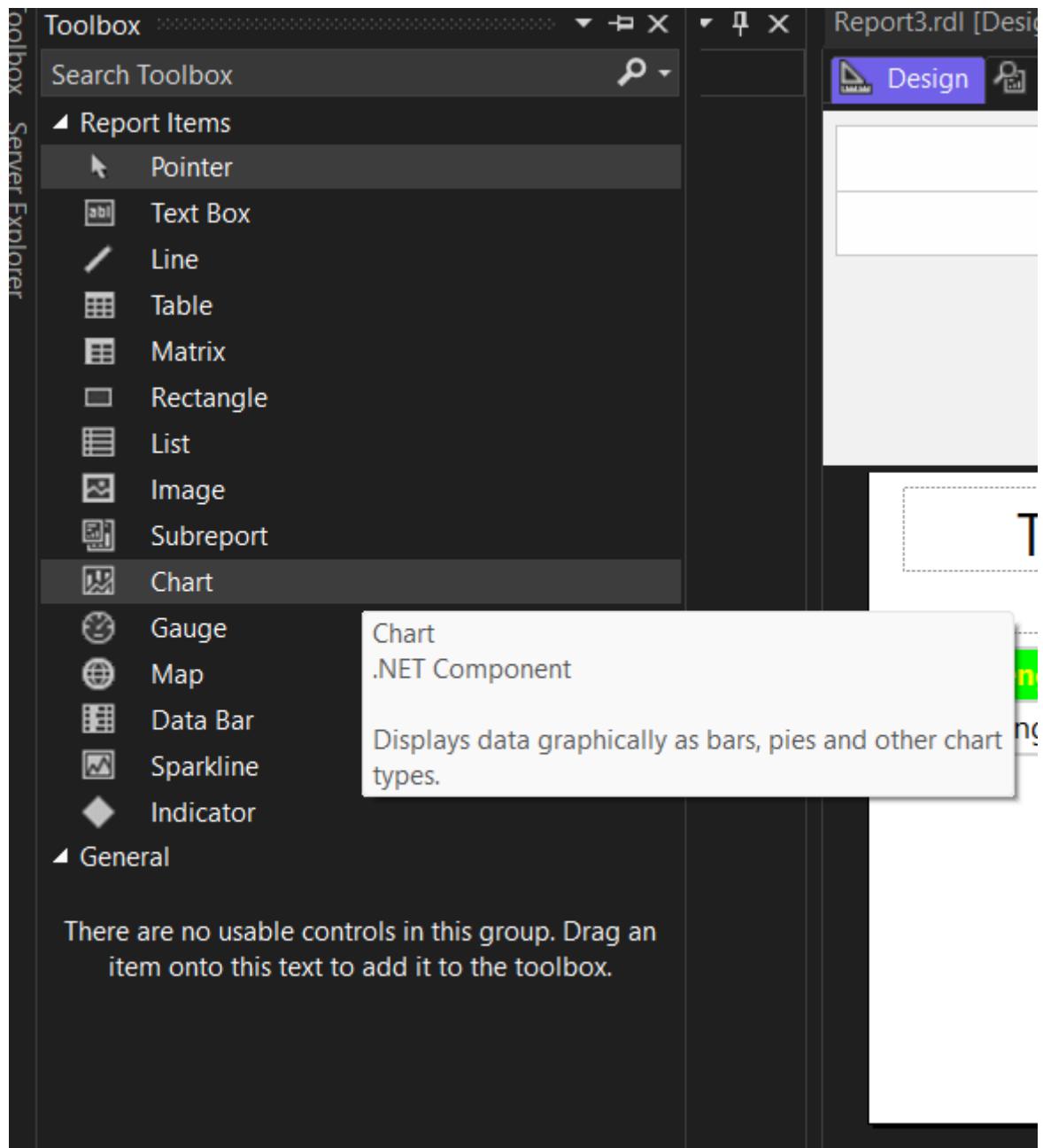
Origin	Destination	Amount Flight
		615
Aberdeen	Faro	1
AC	MAN	1
ACE	BRS	1
ACE	LBA	1
AGP	BRU	1
AGP	MAN	1
AGP	PMI	1
AGP	STN	1
ALC	BHX	1
ALC	EDI	1
ALC	EMA	2
Alicante	Birmingham	1
Alicante	Brussels	1
Alicante	Dublin	1
Alicante	East Midlands	1
Alicante	Edinburgh	1
Alicante	Gatwick	1
Alicante	Kaunas	1
Alicante	Leeds	1
Alicante	Liverpool	2
Alicante	London	1
Alicante	London Stansted	1
Alicante	Newcastle	1
Alicante	Newcastle-upon-Tyne	1
Alicante	Paris	1
Alicante	Shannon	1
Alicante	Stansted	2
Alicante	Venice	1
Amman	Brussels	1
Amman	Paphos	1
AMS	DUB	1
Amsterdam	Dublin	8
Amsterdam	Malaga	2
Ancona	Stansted	1
Ancona Falconara	London Stansted	1
Arrecife	Birmingham	1
Arrecife	Luton	1

Page 1/31

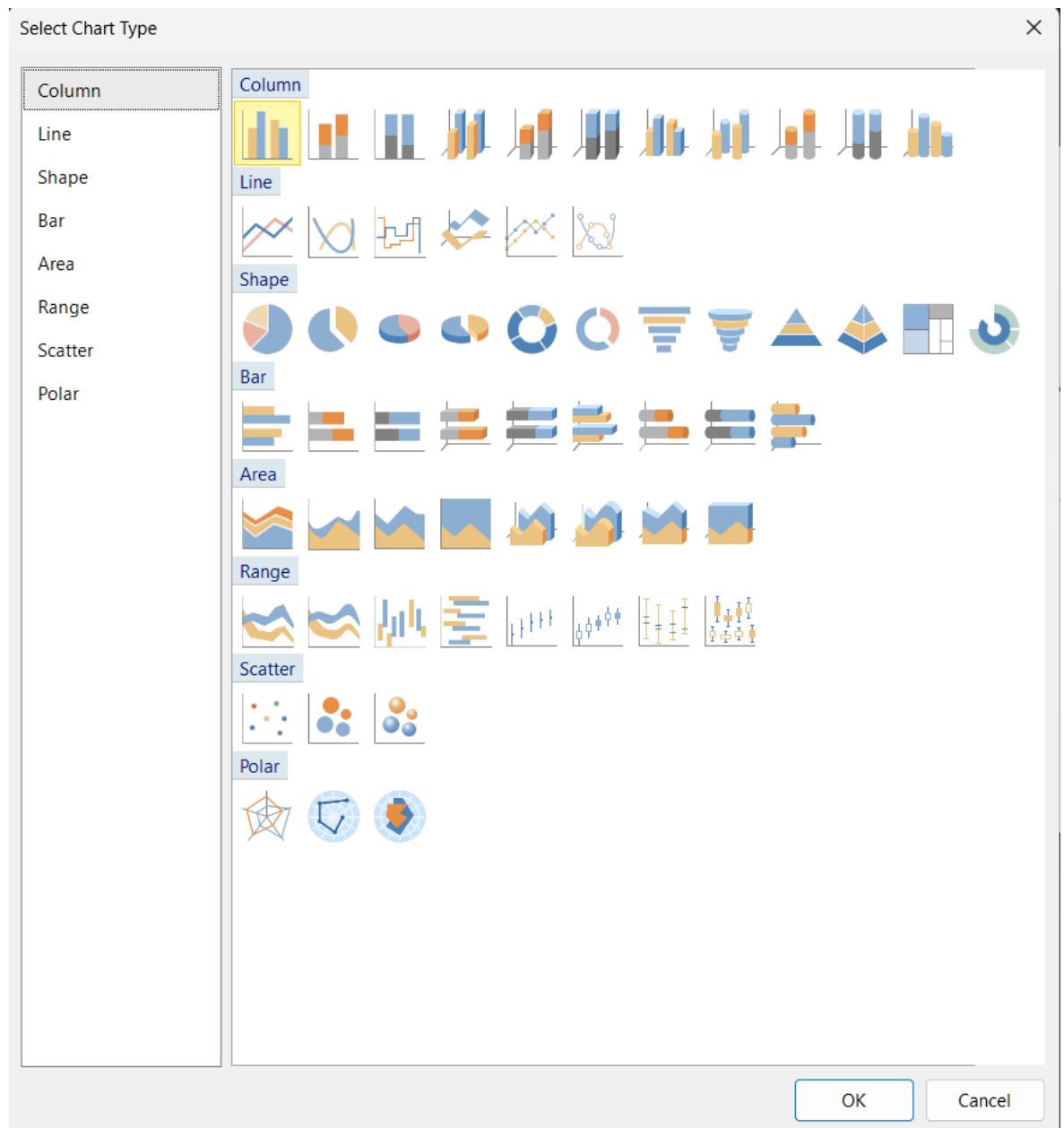
5.1.2. Thực hiện thống kê số lượng chuyến bay theo từng quốc gia

Tiến hành thực hiện tương tự như báo cáo một và thực hiện chèn biểu đồ hình tròn để thể hiện sự so sánh giữa các quốc gia.

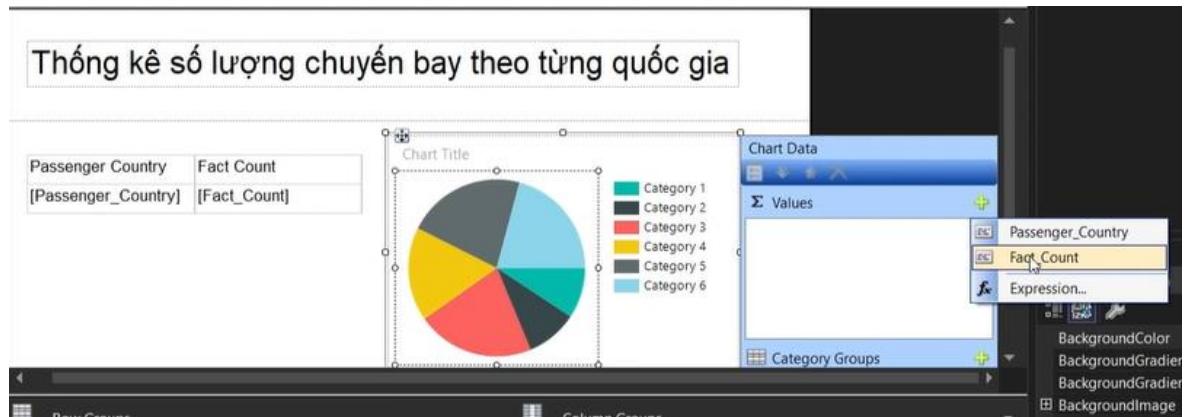
Bước 1: Vào ToolBox chọn Chart.



Bước 2: Chọn biểu đồ tròn.



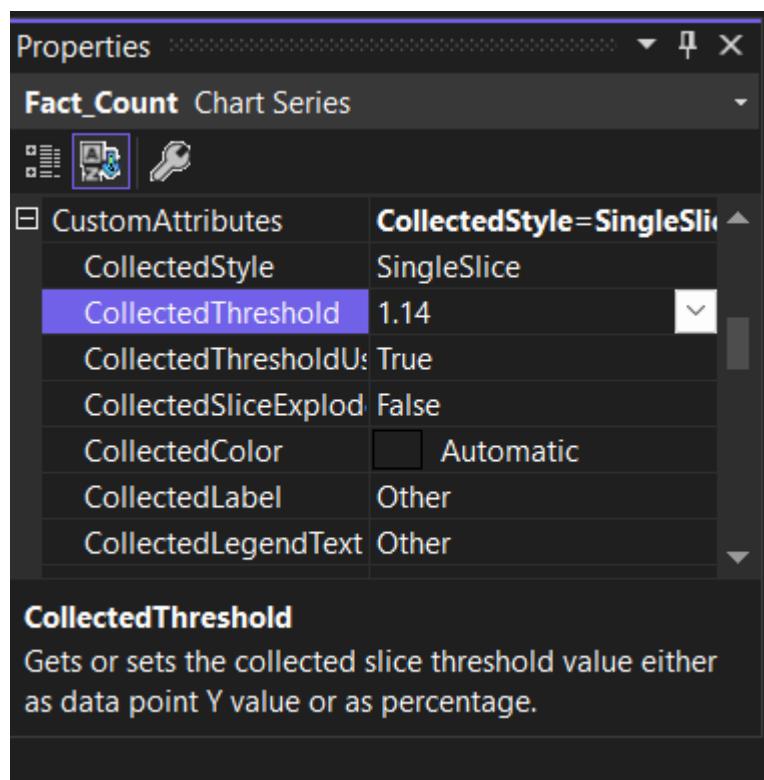
Bước 3: Cài đặt các giá trị mà biểu đồ hình tròn biểu thị.



Bước 4: Gộp các quốc gia có tỉ lệ nhỏ hơn 1.14% thành một giá trị để giảm bớt số quốc gia mà biểu đồ thể hiện.

Bước 4.1: Nhấn chọn vào biểu đồ.

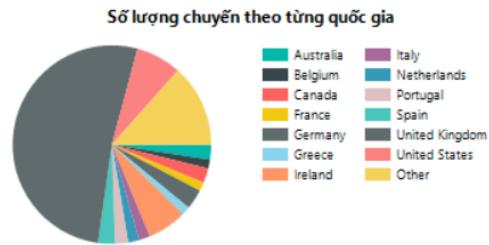
Bước 4.2: Tại tab **Properties** thay đổi các mục **CollectedStyle** chuyển thành **SingleSlice**, **Collected Threshold =1.14**



Bước 5: Nhấn tab **Preview** để xem kết quả.

Thống kê số lượng chuyến bay theo từng quốc gia

Passenger Country	Fact Count
Albania	2
Argentina	1
Australia	55
Austria	13
Belarus	5
Belgium	32
Brazil	1
Brunei	1
Bulgaria	4
Canada	54
Chile	1
China	3
Costa Rica	1
Croatia	4
Cyprus	7
Czech Republic	15
Denmark	12
Egypt	1
Estonia	2
Finland	7

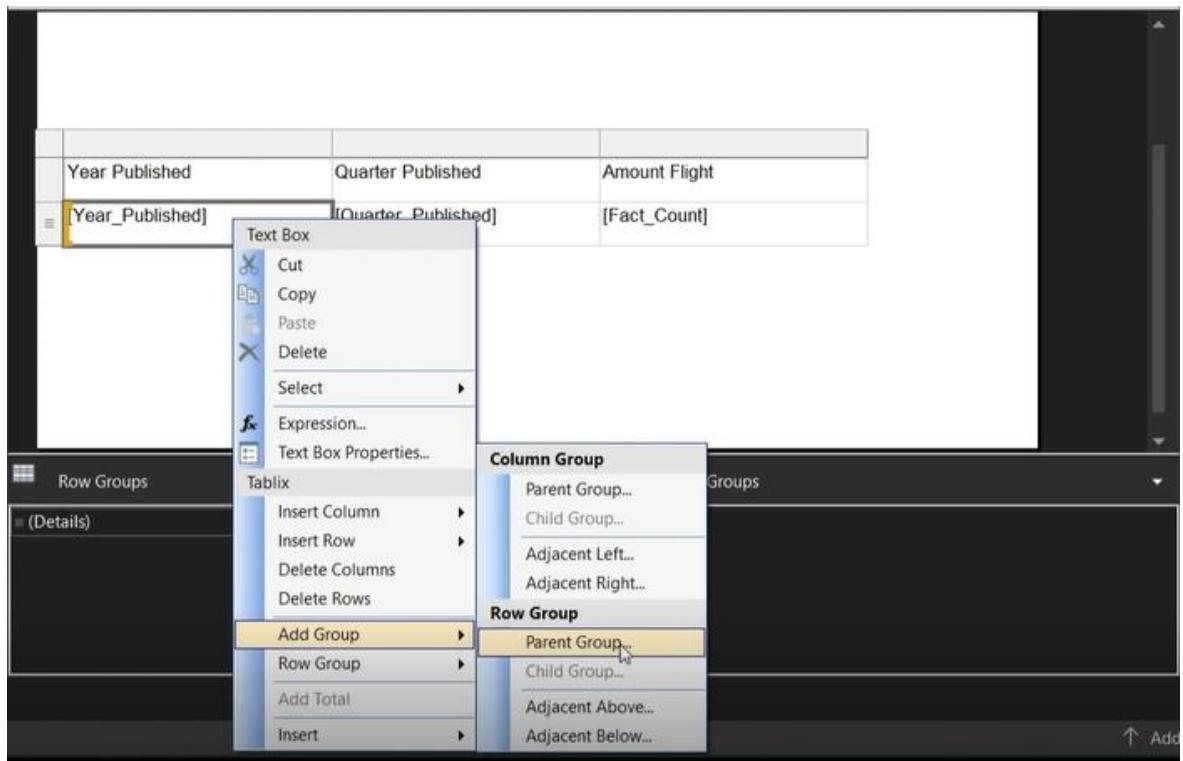


5.1.3. Thống kê số lượng theo các quý của từng năm.

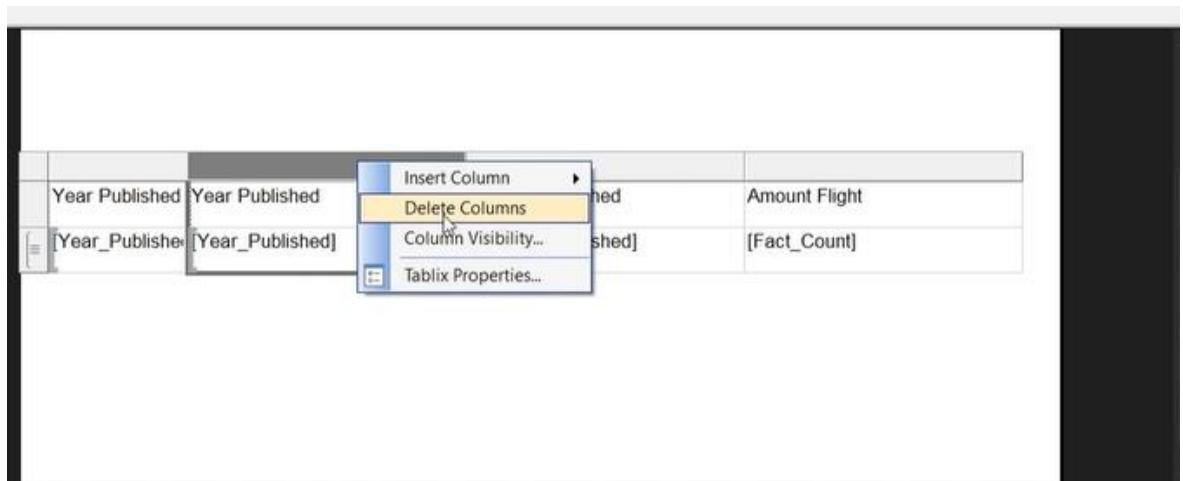
Các bước kết nối cơ sở dữ liệu và truy vấn thực hiện tương tự như báo cáo 1.

Bước 1: Sau khi thực hiện truy vấn vào tạo ra bảng kết quả ở phần **Design** thì tiến hành nhóm các năm lại với nhau.

Bước 1.1: Án chuột vào cột **Year_Published** chọn **Add Group** sau đó chọn **Parent Group**.



Bước 1.2: Tiến hành xóa cột **Year_Publish** ban đầu.



Bước 2: Tiến hành tính tổng các chuyến bay qua của từng năm bằng cách thêm một hàng **Total** ở mỗi năm. Chọn **Quarter_Published**, ấn chuột phải vào **Detail** chọn **Add Total** chọn **After**.

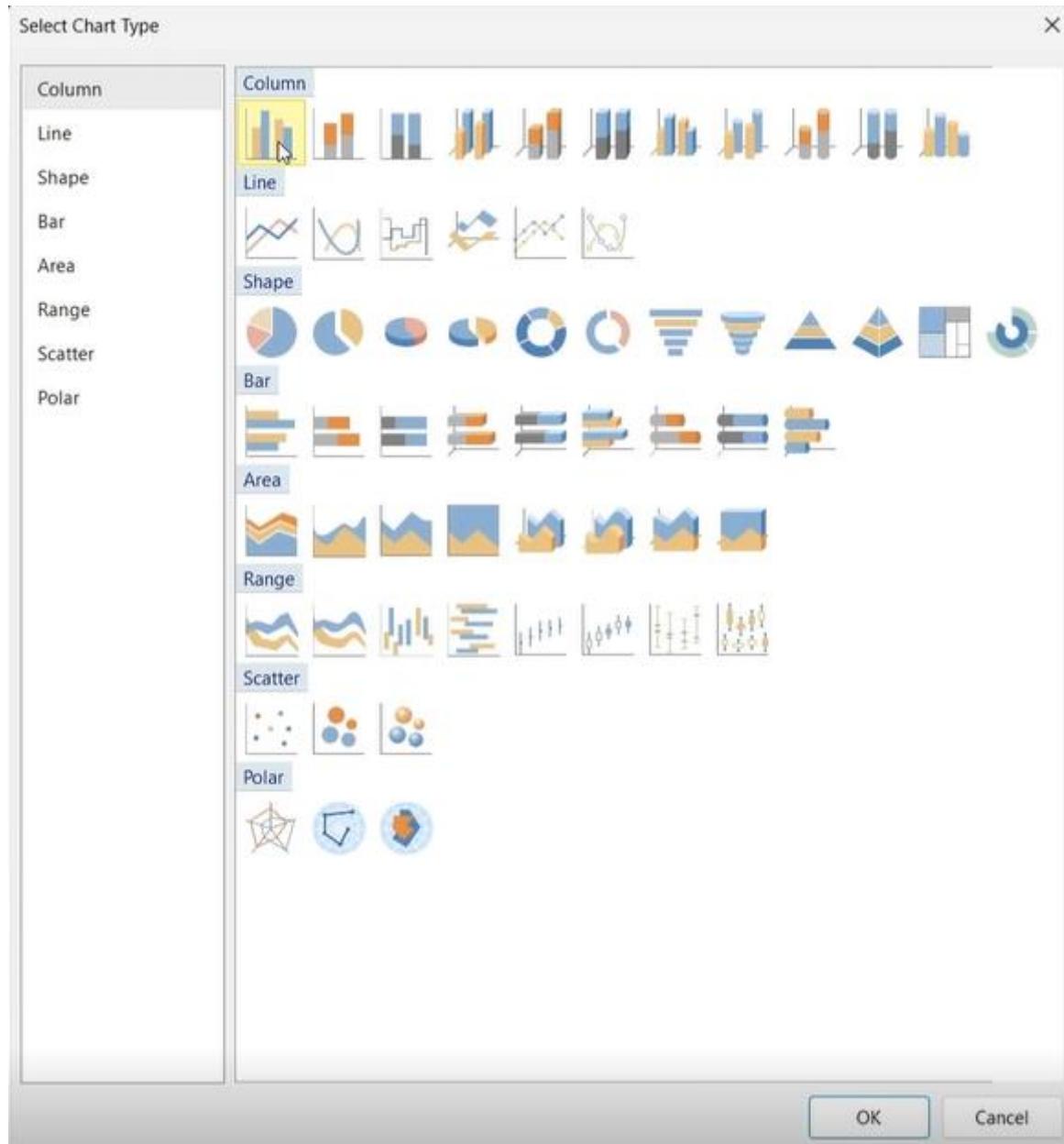
The screenshot shows the Microsoft Power BI Data View. A table is displayed with three columns: 'Year Published', 'Quarter Published', and 'Amount Flight'. The 'Year Published' column is currently selected, indicated by a yellow border around its header and data cells. In the bottom ribbon, the 'Row Groups' section is active, showing a group for 'Year_Published' with '(Details)' selected. A context menu is open over this group, containing options such as 'Add Group', 'Add Total', 'Delete Group', and 'Group Properties...'. The 'Add Total' option is highlighted.

Bước 3: Định dạng màu sắc, cỡ chữ, font chữ cho bảng.

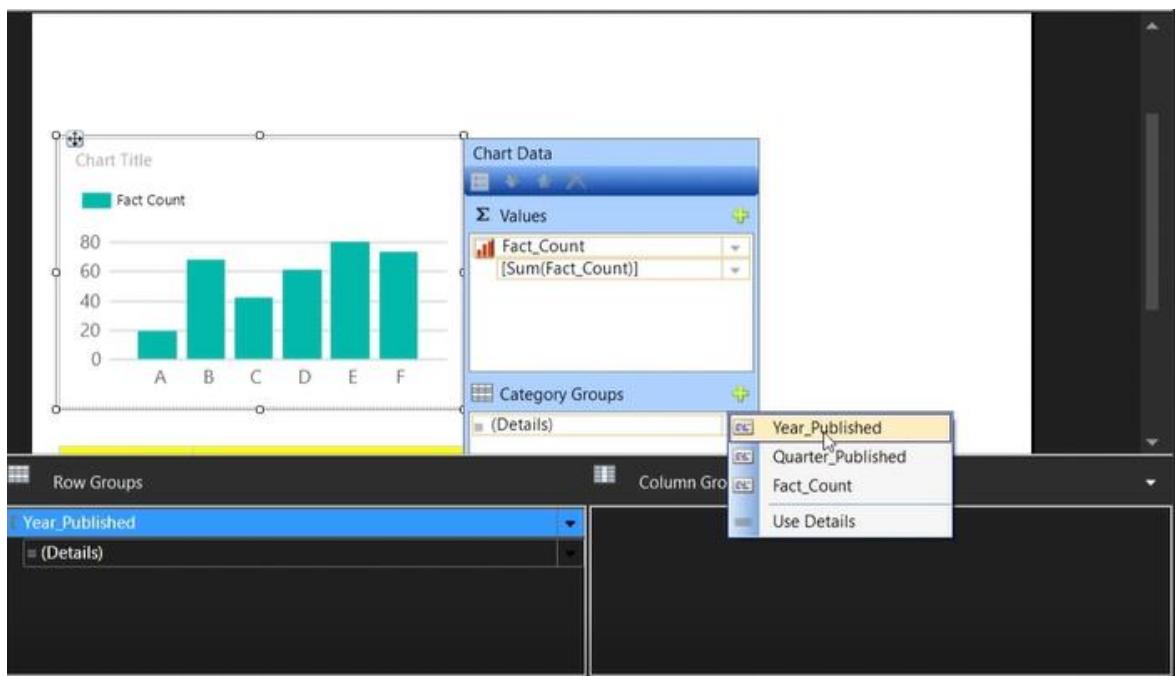
The screenshot shows the Microsoft Power BI Data View after styling. The table columns have been formatted with colors: 'Year Published' is blue, 'Quarter Published' is red, and 'Amount Flight' is green. The 'Quarter Published' column now includes a new row labeled 'Total:' in red. The 'Amount Flight' column displays the formula '[Sum(Fact_Count)]' in green.

Bước 4: Tạo biểu đồ cột.

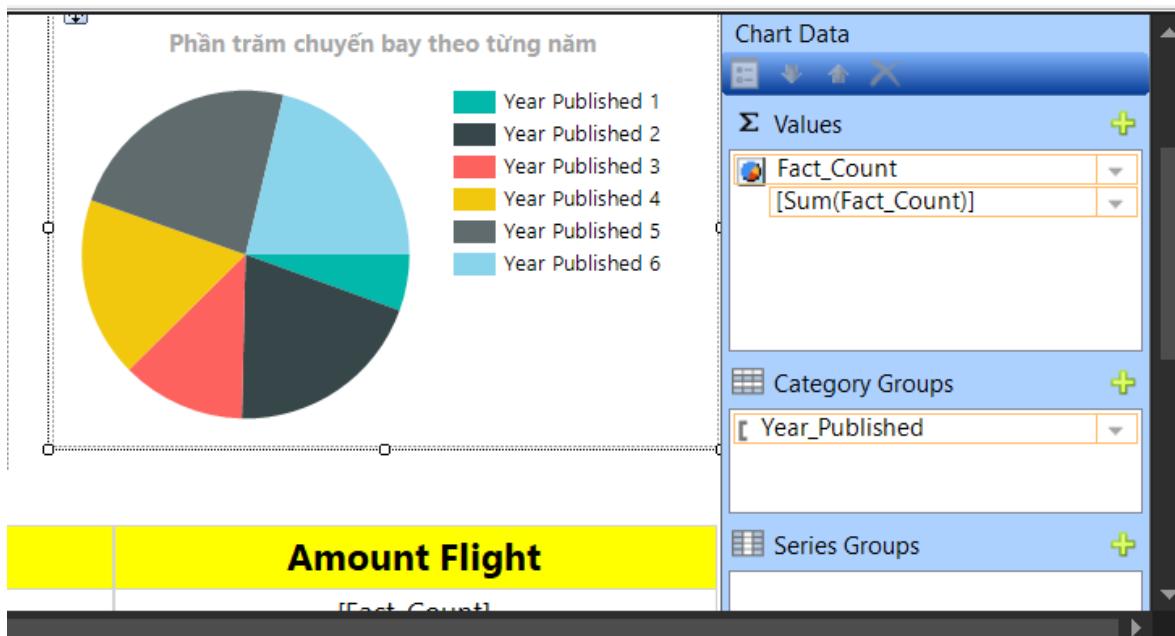
Bước 4.1: Tại ToolBox chọn **Chart** và chọn biểu đồ cột. Sau đó nhấn **OK**.



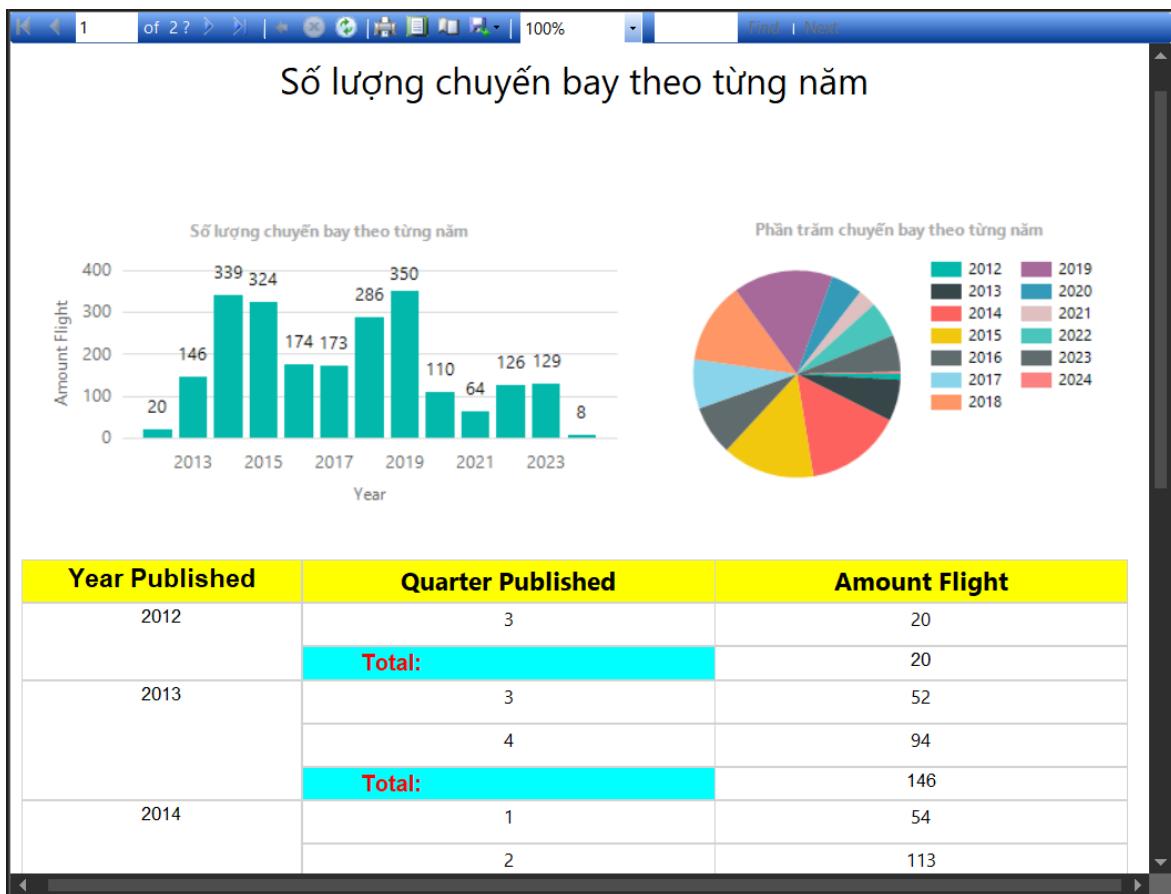
Bước 4.2: Định dạng các giá trị trực x và y cho biểu đồ.



Bước 5: Tạo biểu đồ tròn.



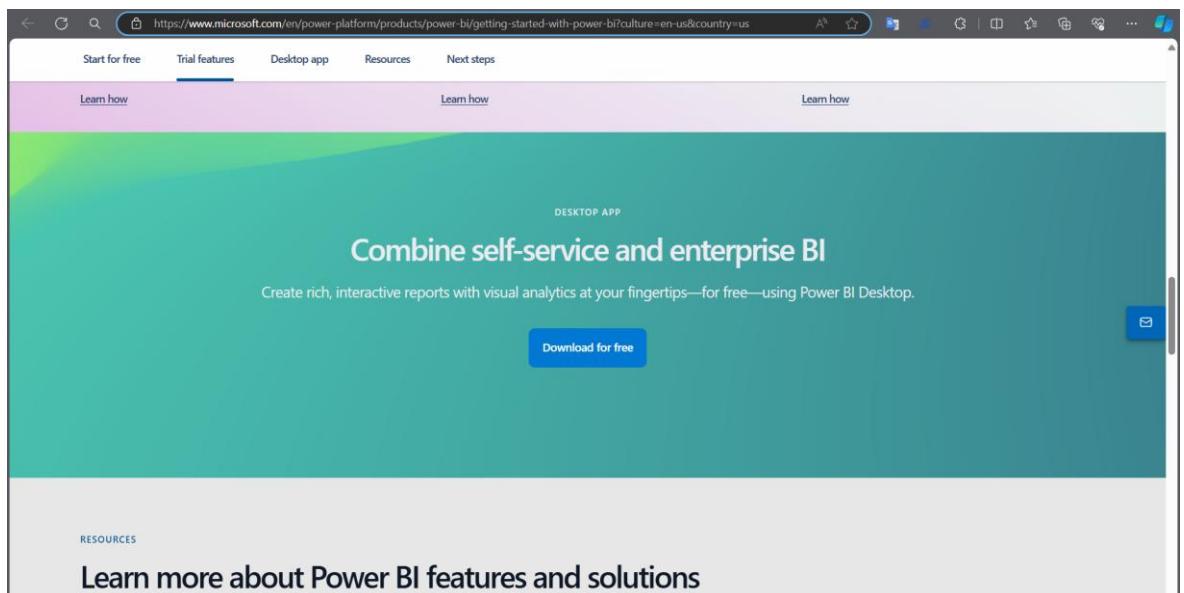
Bước 6: Nhấn vào tab Preview để xem kết quả.



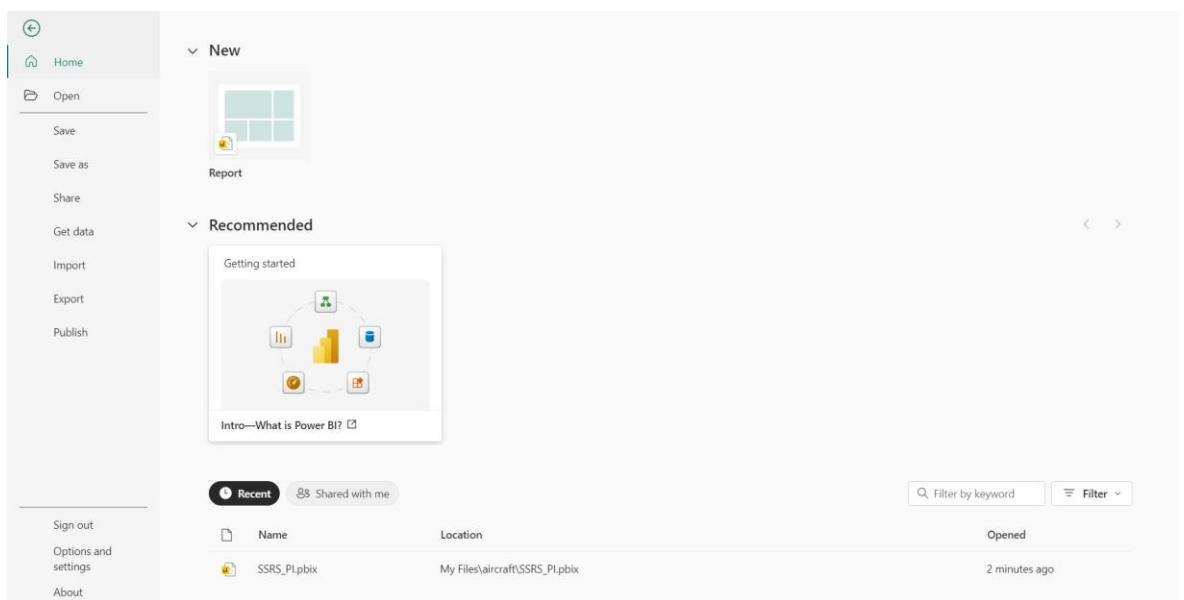
5.2. Report với Power BI

Bước 1: Thực hiện download Power BI tại đường dẫn:

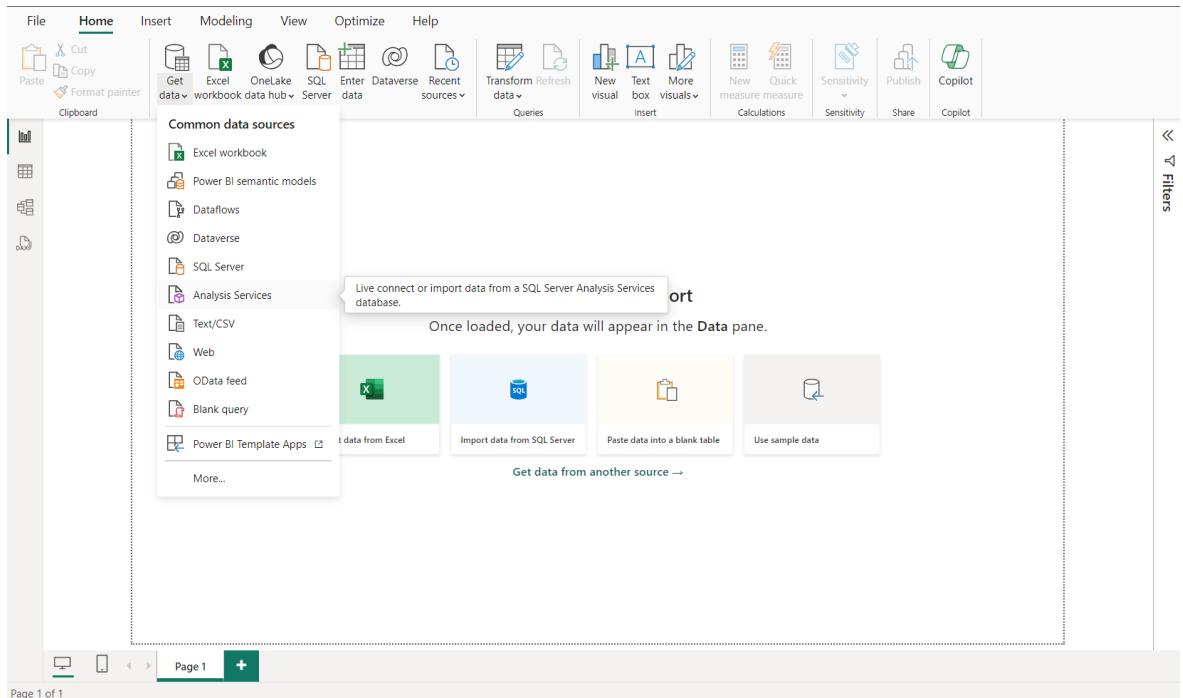
<https://www.microsoft.com/en/power-platform/products/power-bi/getting-started-with-power-bi?culture=en-us&country=us>



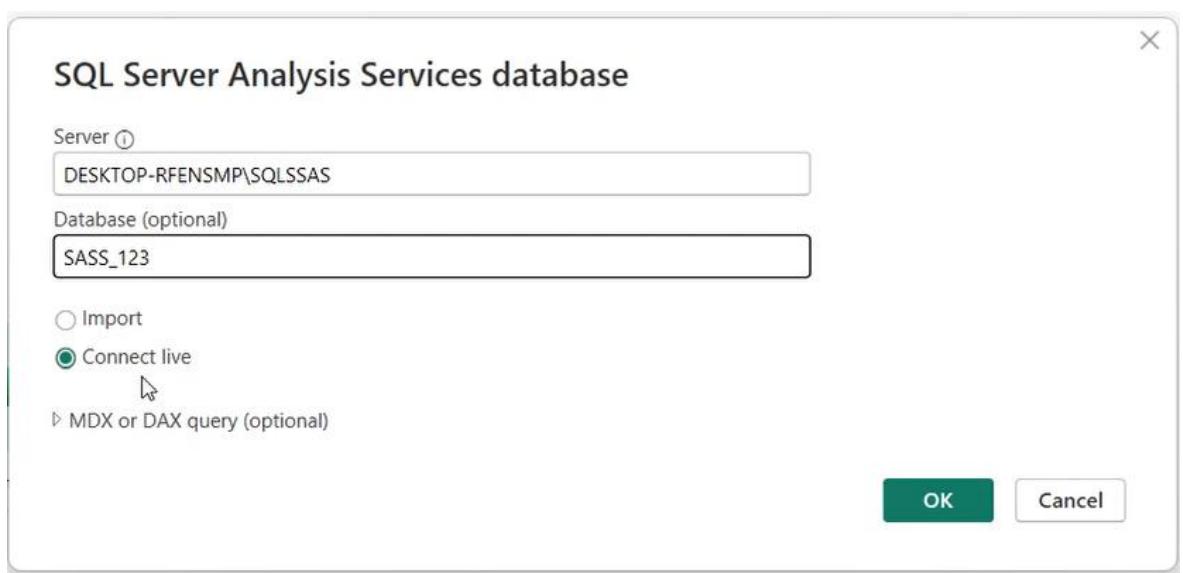
Bước 2: Ở phần New chọn Report để tạo báo cáo mới.



Bước 3: Trên thanh công cụ chọn Get Data, sau đó chọn Analysis Services. Để tiến hành kết nối với cơ sở dữ liệu SSAS.



Bước 4: Tiến hành điền tên Server và tên Database. Và nhấn OK



Bước 5: Sau khi kết nối thành công sẽ hiển thị các bảng của database.

Navigator



AIRCRAFT WAREHOUSE

Last Modified: 05/18/2024 04:48:22

This model contains the following dimensions and measures

Dim Air Craft, Dim Passenger Country, Dim Return, Dim Route, Dim Seat Type, Dim Time, Dim Travel Type, Avg_Overall_Rating, Avg_Seat_Comfort, Avg_Cabin_Staff Service, Avg_Food_Beverages, Avg_Ground_Service, Sum_Value_For_Money, Fact_Count, Sum_Overall_Rating, Avg_OverRating



5.2.1. Thống kê các đánh giá trung bình của khách hàng về chất lượng thức ăn, dịch vụ mặt đất, dịch vụ cabin, và độ thoải mái của ghế qua các năm.

Bước 1: Tại mục Data thực hiện chọn các thuộc tính cần thiết bao gồm:

Year_Published, Avg_Food_Beverages, Avg_Seat_Comfort,
Avg_Ground_Service, Avg_Cabin_Staff_Service.

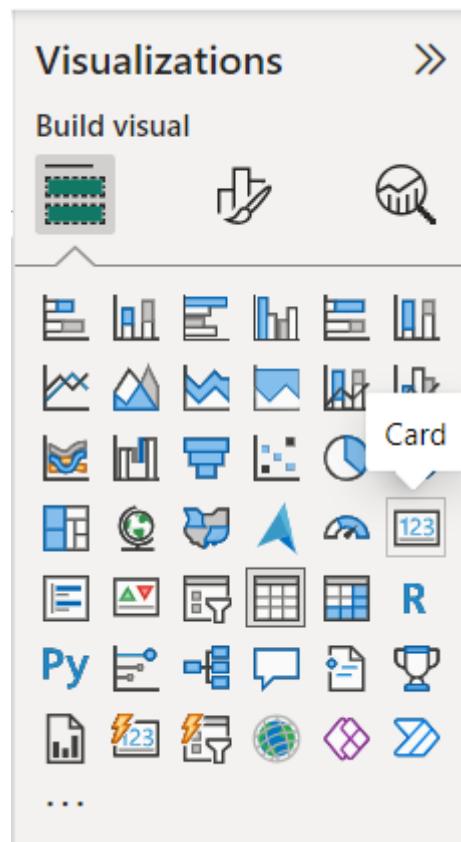
The screenshot shows the Power BI 'Filters' pane on the left, 'Visualizations' pane in the center, and 'Data' pane on the right. In the 'Data' pane, under the 'Fact' section, the following fields are selected with checked checkboxes:

- Avg_Cabin_Staff_Service
- Avg_Food_Beverages
- Avg_Ground_Service
- Avg_Seat_Comfort
- Year_Published

Under the 'Dim Time' section, the 'Hierarchy' field is also selected with a checked checkbox.

Bước 2: Thực hiện tính tổng số lượng các đánh giá của khách hàng.

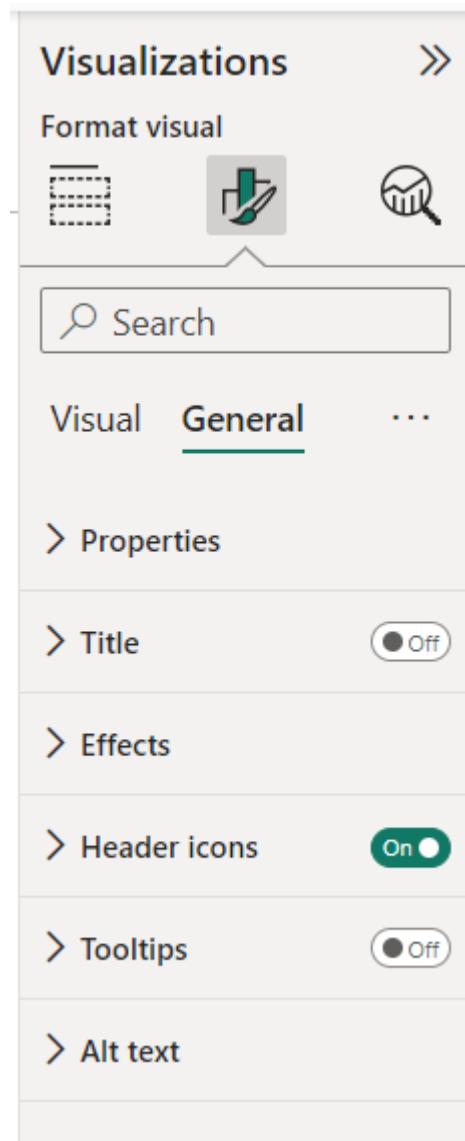
Bước 2.1: Tại mục Visualizations chọn mục Card.



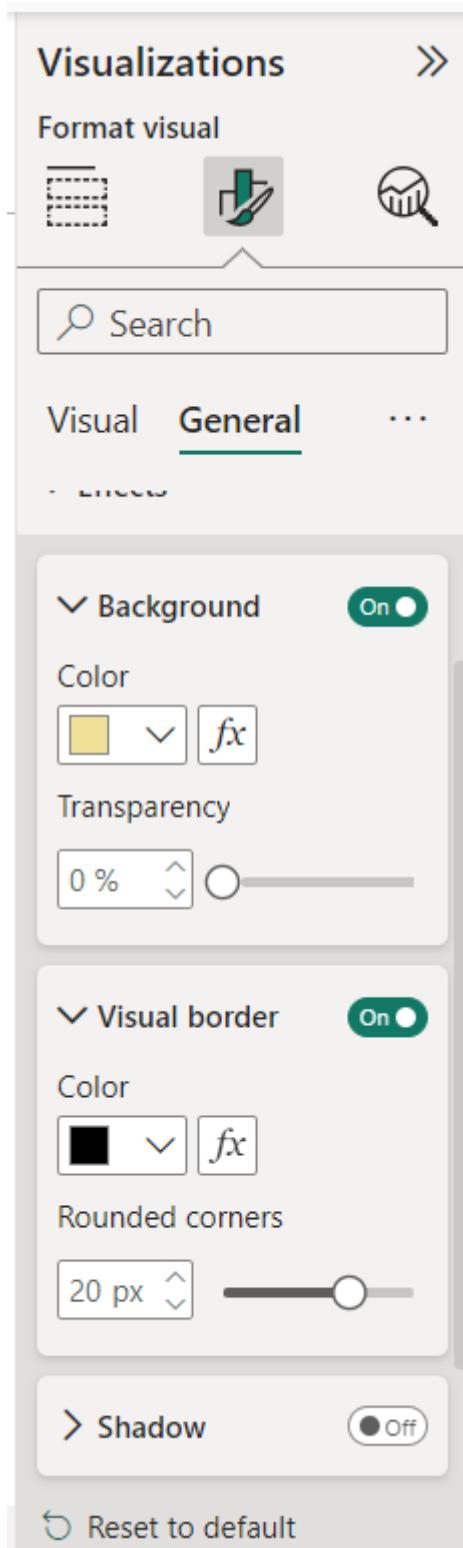
Bước 2.2: Kéo thả thuộc tính Fact_Id tại mục Data vào ô Card.



Bước 2.3: Nhấn chọn vào Card, tại mục Visualization, chọn tab Format your visual và chọn vào tab General.



Bước 2.4: Tại mục Effect, tiến hành định dạng màu nền, và bo tròn góc.



Bước 2.5: Tại mục Fields, chọn Rename for this visual. Để đổi Fact_Count thành Total Review.

The screenshot shows the Power BI Data view interface. On the left, there's a 'Visualizations' pane with various chart icons and a 'Build visual' button. Below it is a 'Fields' section where 'Fact_Count' is listed. A context menu is open over 'Fact_Count', showing options like 'Remove field', 'Rename for this visual', and a list of related dimensions: 'Dim Air Craft', 'Dim Route', 'Dim Seat Type', and 'Dim Time'. Under 'Dim Time', three measures are listed: 'Date Published', 'Day Published', and 'Hierarchy'. The right side of the screen shows a tree view of data sources under 'Fact' and 'Measures'.

Visualizations > Data

Build visual

Fact

- Avg_Cabin_St...
- Avg_Food_Bev...
- Avg_Ground_...
- Avg_Overall_R...
- Avg_Seat_Co...
- Fact_Count
- Sum_Overall_...
- Sum_Value_Fo...

Measures

- Avg_OverRating

Fields

Fact_Count

Remove field

Rename for this visual

Drill through

Cross-report

Keep all filters

Add drill-through fields here

Dim Air Craft

Dim Route

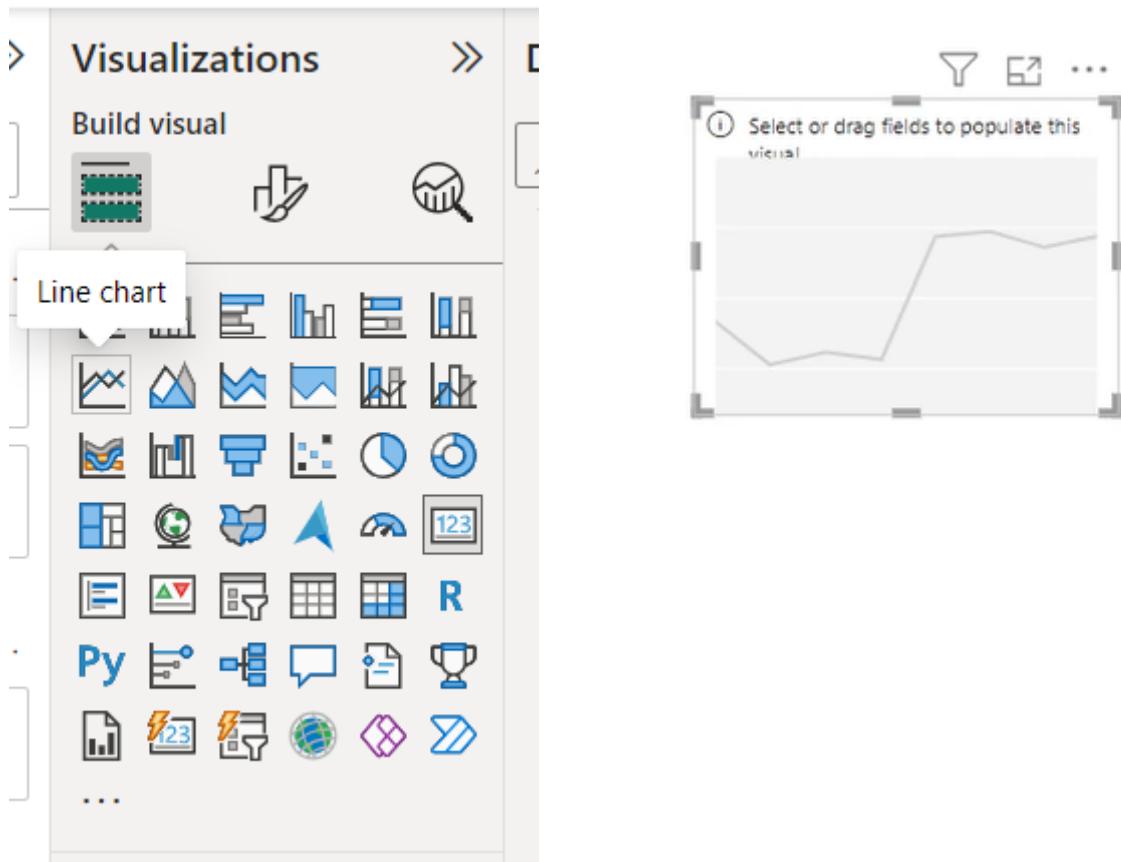
Dim Seat Type

Dim Time

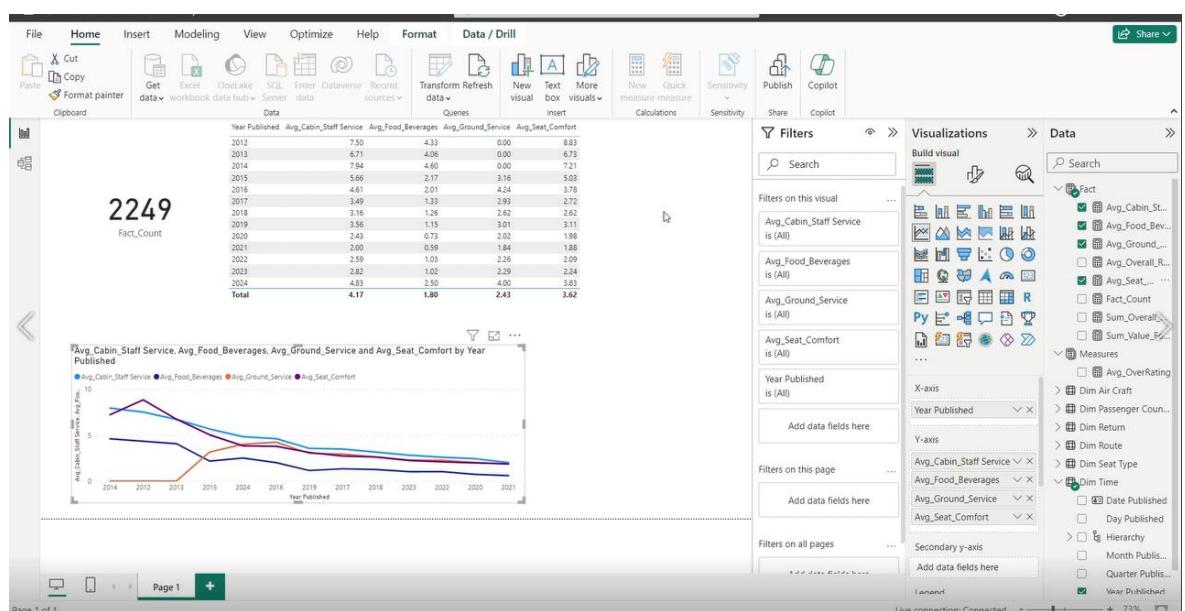
- Date Published
- Day Published
- Hierarchy

Bước 3: Thực hiện vẽ các biểu đồ đường để thể hiện tương quan giữa các giá trị trung bình.

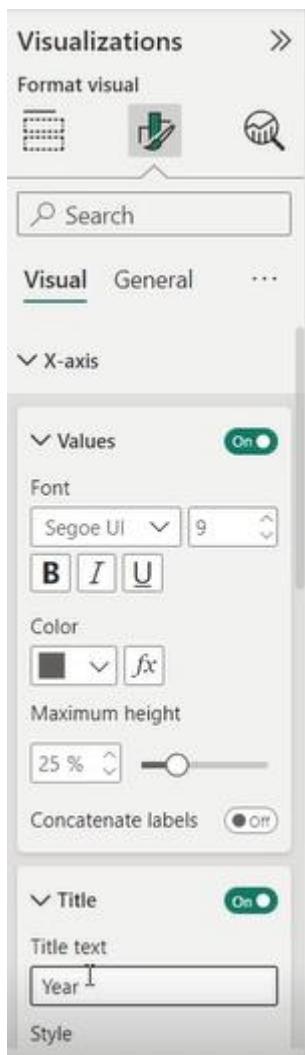
Bước 3.1: Tại mục Visualization chọn vào Line Chart.



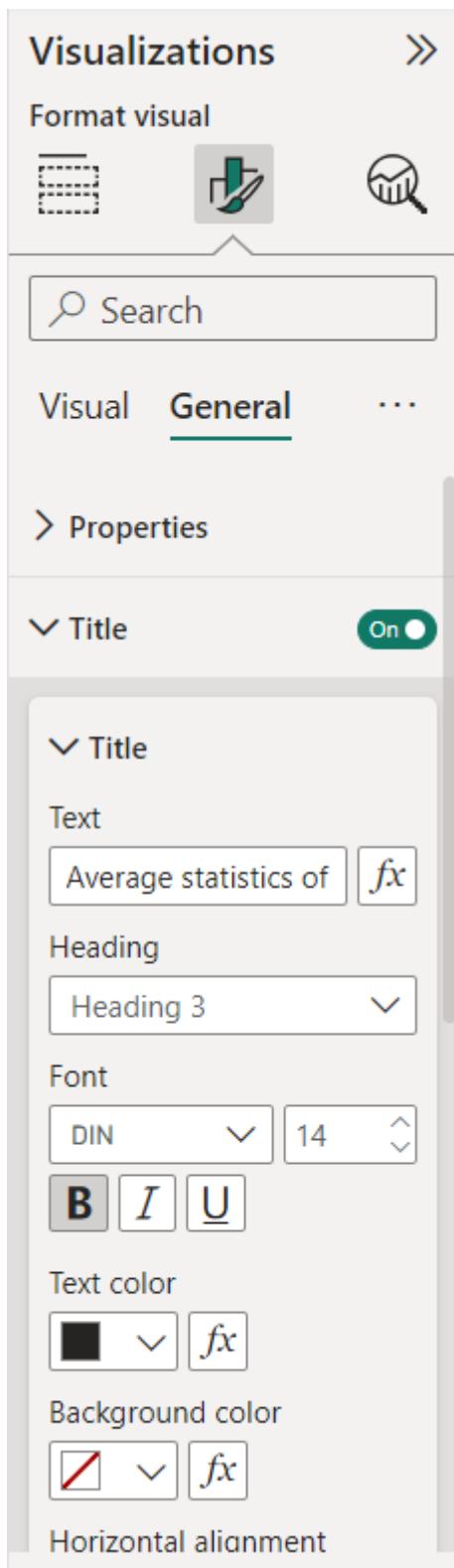
Bước 3.2: Tại mục Data thực hiện chọn các thuộc tính cần thiết bao gồm:
 Year_Published, Avg_Food_Beverages, Avg_Seat_Comfort,
 Avg_Ground_Service, Avg_Cabin_Staff_Service.



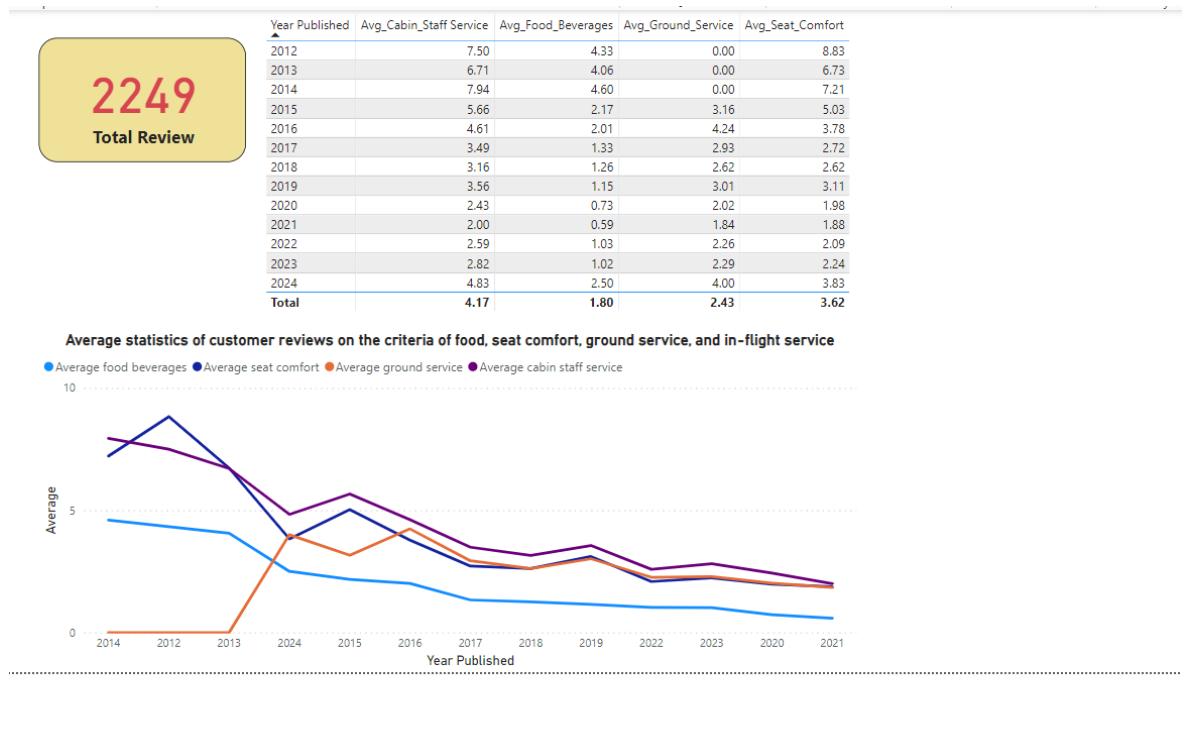
Bước 3.3: Tại mục Format your visual, trong tab Visual chọn X-axis, chọn Title, tại mục Title text nhập tên của trục X. Thực hiện tương tự ở mục Y-axis để đặt tên cột Y.



Bước 3.4: Chuyển sang tab Generate, chọn Title và tiến hành đổi tên cho biểu đồ.



Bước 4: Hoàn thành báo cáo.



5.2.2. Thống kê số lượng các chuyến bay theo mục đích chuyến bay từ năm 2016 đến năm 2024

Thực hiện tung tự các bước kết nối với cơ sở dữ liệu như báo cáo đầu tiên.

Bước 1: Thực hiện kéo thả các thuộc tính cần thiết cho bảng dữ liệu bao gồm:

Year_Published, Type_Of_Travel, Fact_Count.

The screenshot shows the Microsoft Power BI desktop interface. On the left, there is a table visualization with columns: Year Published, Type Of Traveller, and Fact_Count. The data is as follows:

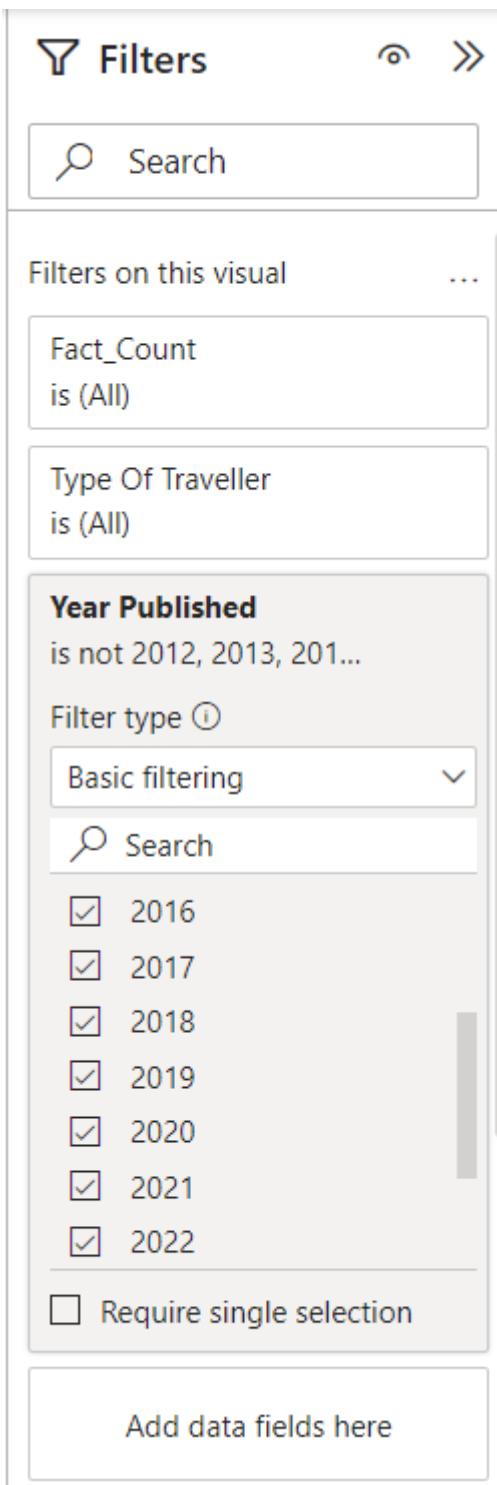
Year Published	Type Of Traveller	Fact_Count
2012		20
2013		146
2014		339
2015		109
2015	Business	14
2015	Couple Leisure	74
2015	Family Leisure	63
2015	Solo Leisure	64
2016	Business	13
2016	Couple Leisure	65
2016	Family Leisure	37
2016	Solo Leisure	59
2017	Business	11
2017	Couple Leisure	76
Total		2249

The Filters pane on the right shows three filters applied to the visual:

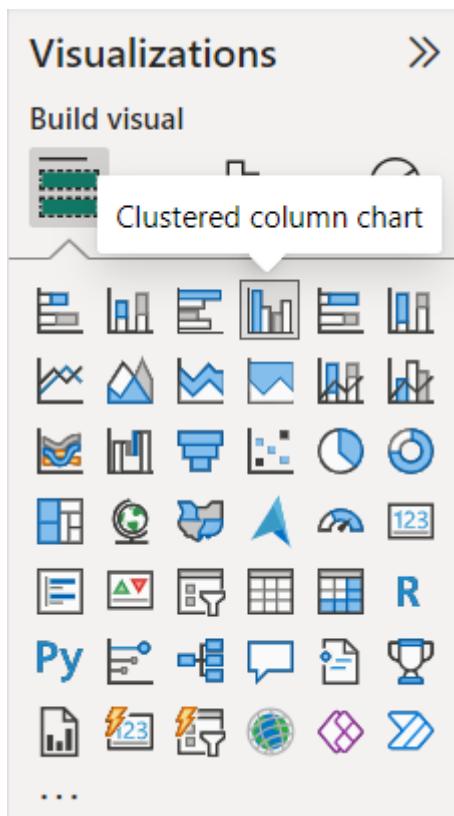
- Fact_Count is (All)
- Type Of Traveller is (All)
- Year Published is (All)

The Data pane on the right lists various dimensions and measures available in the model, such as Fact, Measures, and Dim Time.

Bước 2: Tại mục Filters chọn Year Published và chọn vào các năm từ năm 2016 đến năm 2024.



Bước 3: Tại mục Visualization, chọn vào Clustered column chart để vẽ biểu đồ cột ghép.



Bước 4: Thực hiện kéo thả các thuộc tính cần thiết cho bảng dữ liệu bao gồm:
Year_Published, Type_Of_Travel, Fact_Count.

Bước 5: Tại mục Filters chọn Year Published và chọn vào các năm từ năm 2016 đến
năm 2024.

Filters

Search

Filters on this visual ...

Fact_Count
is (All)

Type Of Traveller
is (All)

Year Published
is not 2012, 2013, 201...
Filter type ⓘ

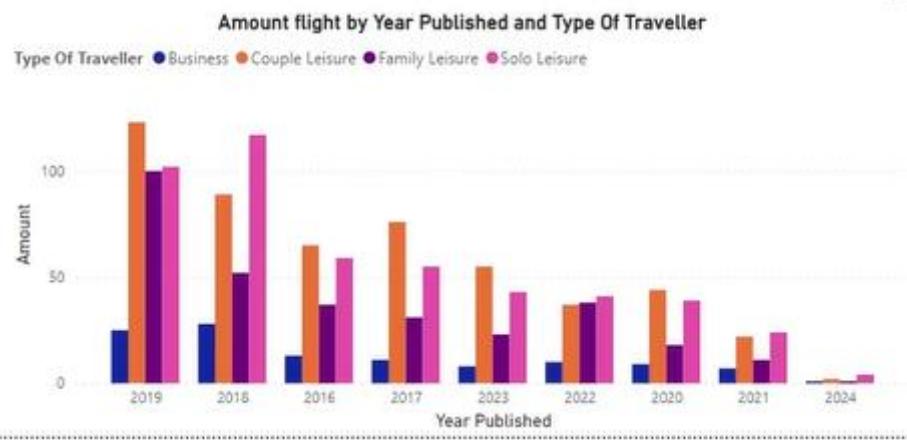
Basic filtering ▾

Search

2019
 2020
 2021
 2022
 2023
 2024
 Unknown

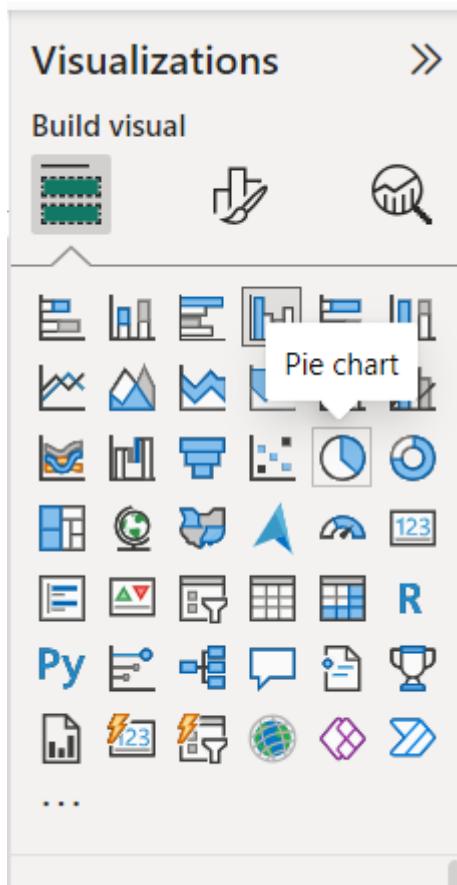
Require single selection

Bước 6: Tiến hành định dạng tên biểu đồ, để hoàn thiện biểu đồ.



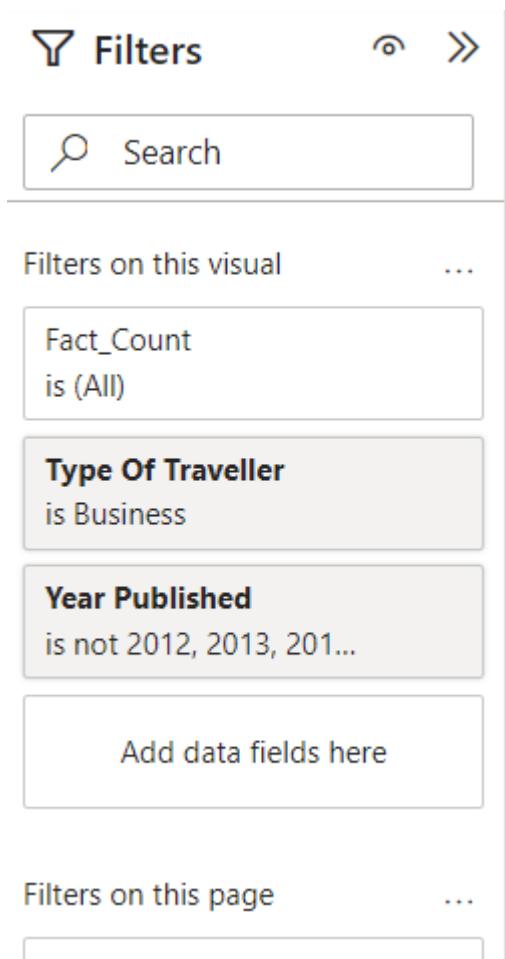
Bước 7: Thực hiện vẽ biểu đồ hình tròn biểu thị phần trăm của từng loại du lịch.

Bước 7.1: Tại mục Visualization, chọn vào Pie Chart để vẽ biểu đồ tròn.



Bước 7.2: Chọn các thuộc tính cho biểu đồ tròn tương tự như biểu đồ cột ch่อง bên trên.

Bước 7.3 Tại mục Filters chọn Year Published và chọn vào các năm từ năm 2016 đến năm 2024. Và Type Of Travel chọn Bussiness.



Bước 7.4: Lần lượt thay đổi Type Of Travel thành các giá trị để được 4 biểu đồ hình tròn thể hiện cho mục đích chuyến bay.



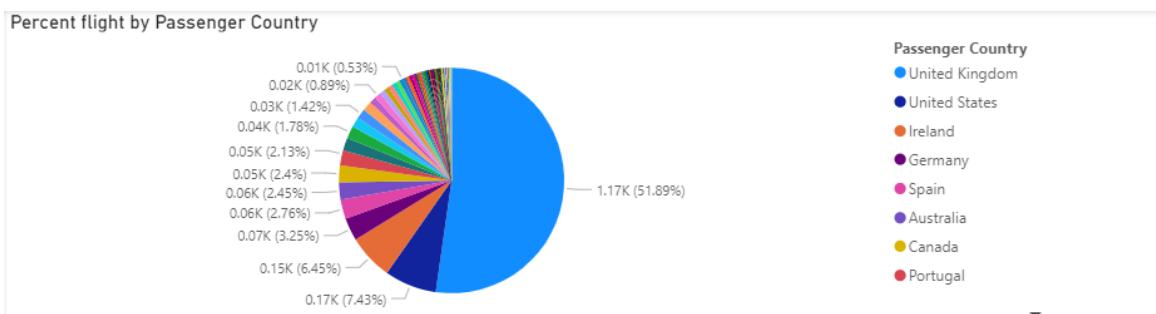
5.2.3. Thông kê số lượng hành khách tham gia chuyến bay theo từng quốc gia.

Thực hiện tung tự các bước kết nối với cơ sở dữ liệu như báo cáo đầu tiên.

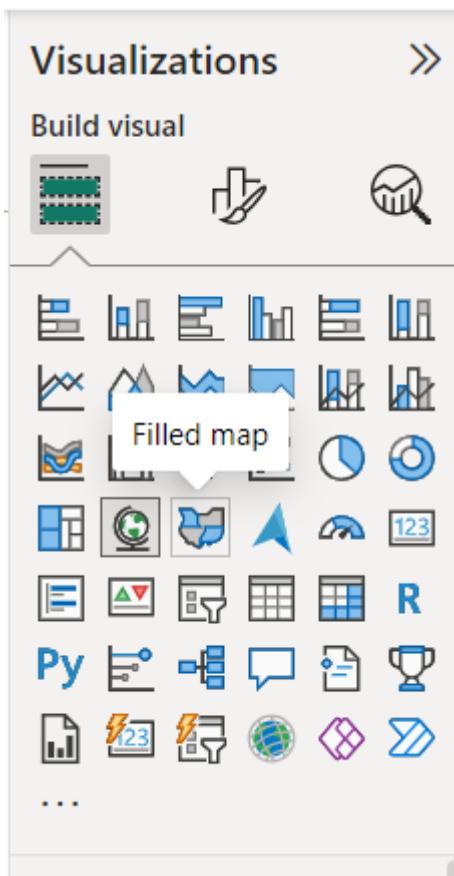
Bước 1: Thực hiện kéo thả các thuộc tính cần thiết cho bảng dữ liệu bao gồm: Passenger_Country, Fact_Count.

Passenger Country	Amount Flight
United Kingdom	1167
United States	167
Ireland	145
Germany	73
Spain	62
Australia	55
Canada	54
Portugal	48
Italy	41
Netherlands	40
Belgium	32
France	32
Greece	31
Romania	21
Sweden	20
Israel	18
Czech Republic	15
Poland	15
Austria	13
India	13
Denmark	12
Ukraine	11
Switzerland	10
South Africa	9
Lithuania	8
Norway	8
Singapore	8
Cyprus	7
Finland	7
Slovakia	7
New Zealand	6
Total	2249

Bước 2: Vẽ biểu đồ hình tròn.



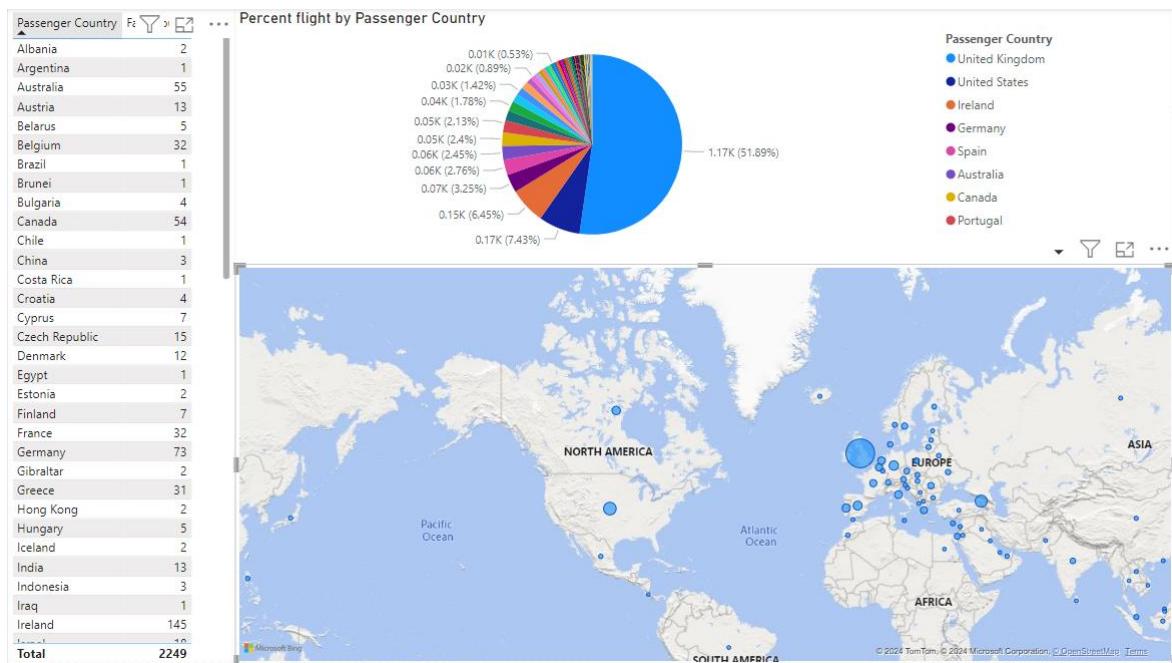
Bước 3: Tại mục Visualization, chọn vào Filled map để vẽ bảng đồ.



Bước 4: Thực hiện chọn các thuộc tính, để vẽ bảng đồ.



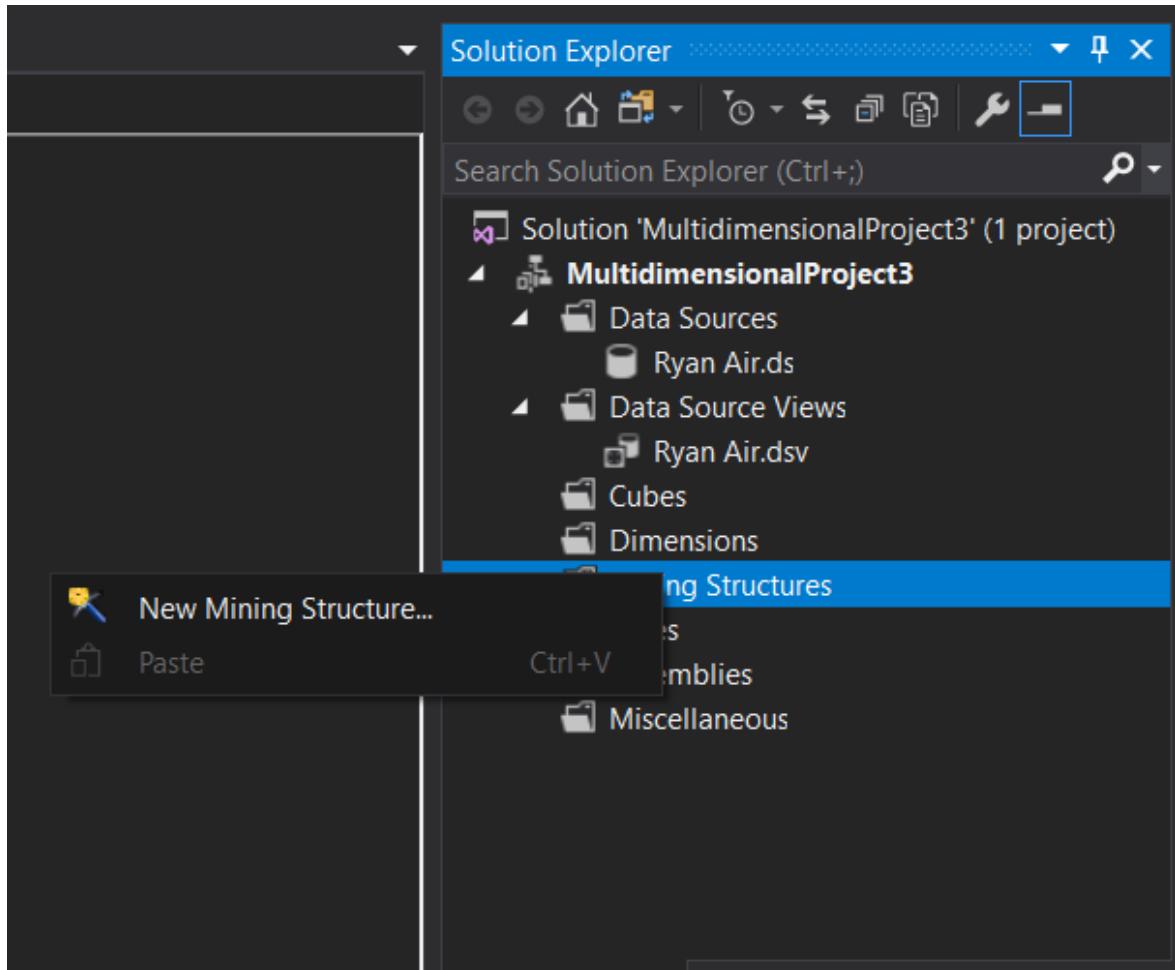
Bước 5: Hoàn thành báo cáo.



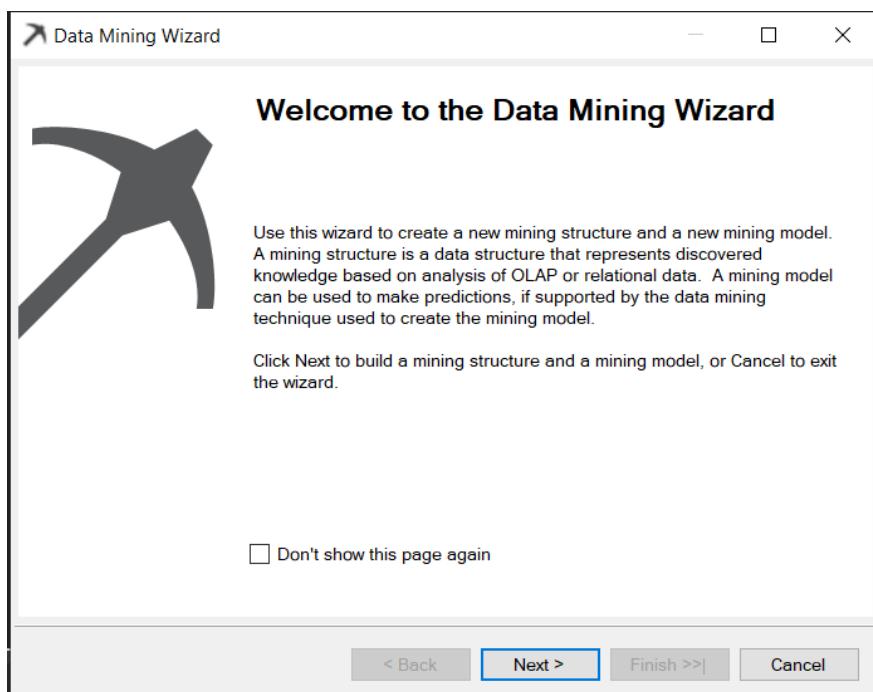
Chương 6. QUÁ TRÌNH DATA MINING

6.1. Data Mining bằng công cụ SSAS trên Visual Studio

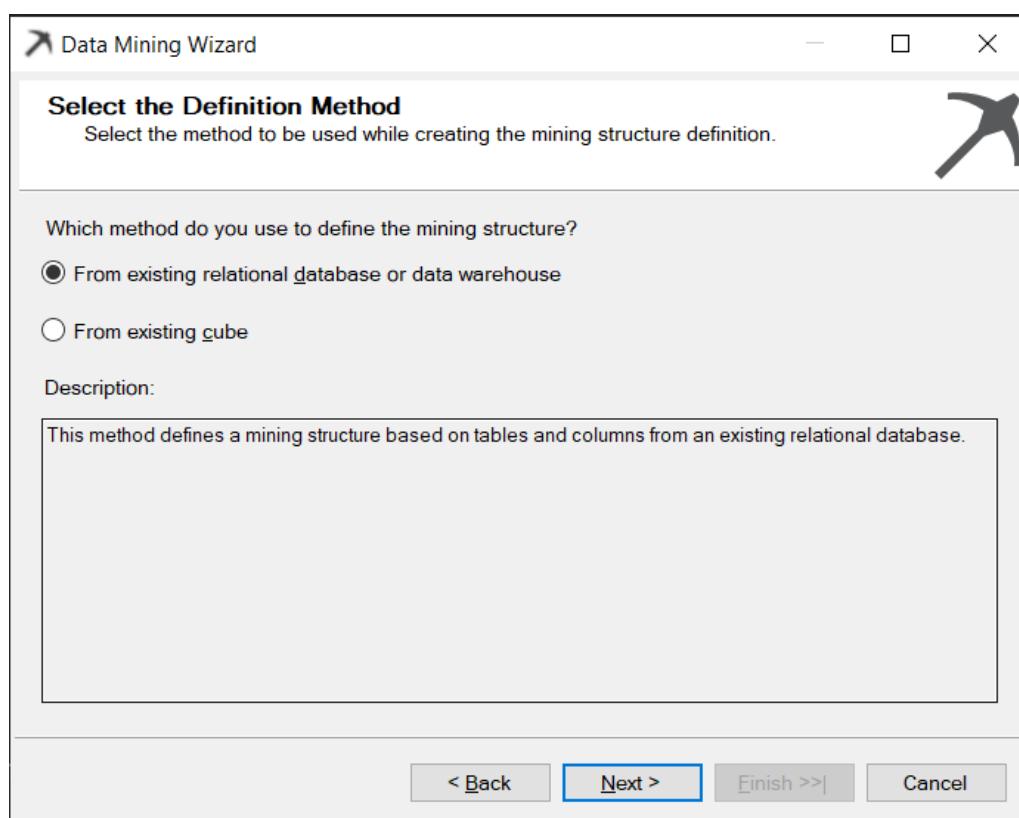
Bước 1: Chọn Mining Structures và chọn New Mining Structure



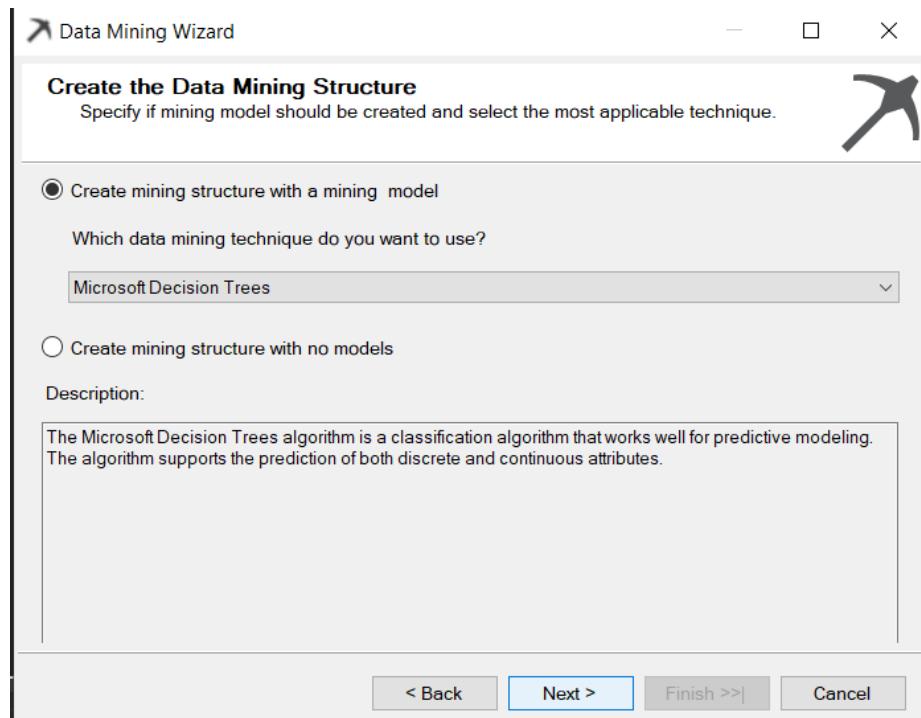
Tiếp tục chọn Next..



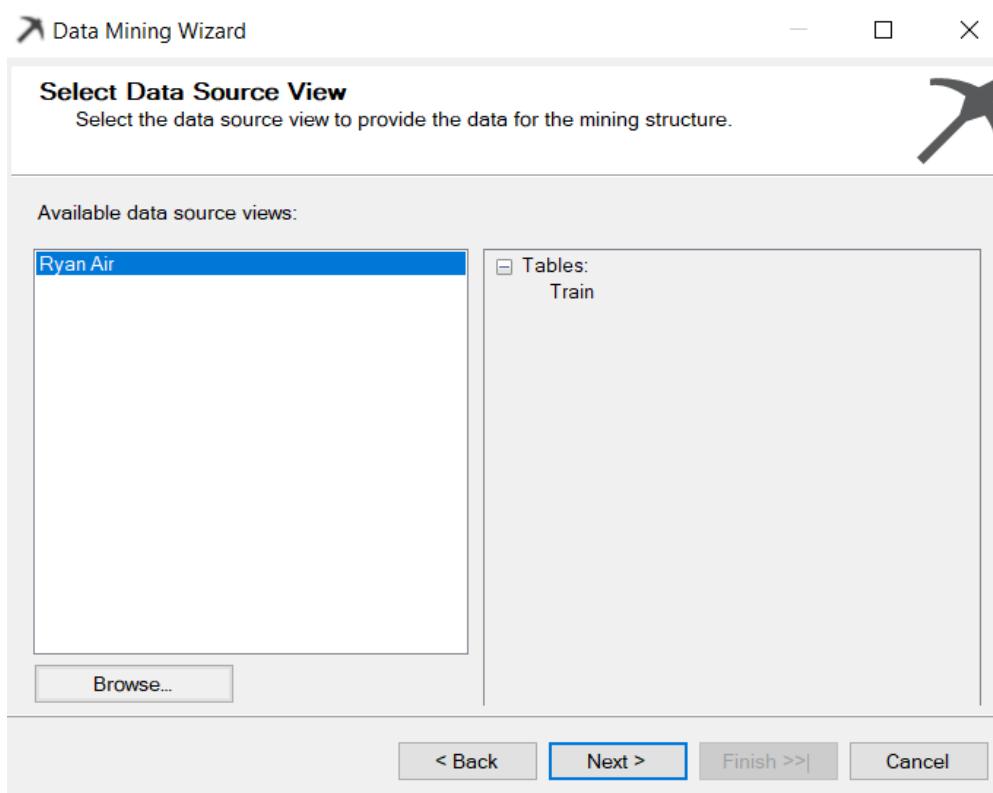
Bước 2: Chọn **From existing relational database or data warehouse** và chọn Next để tiếp tục



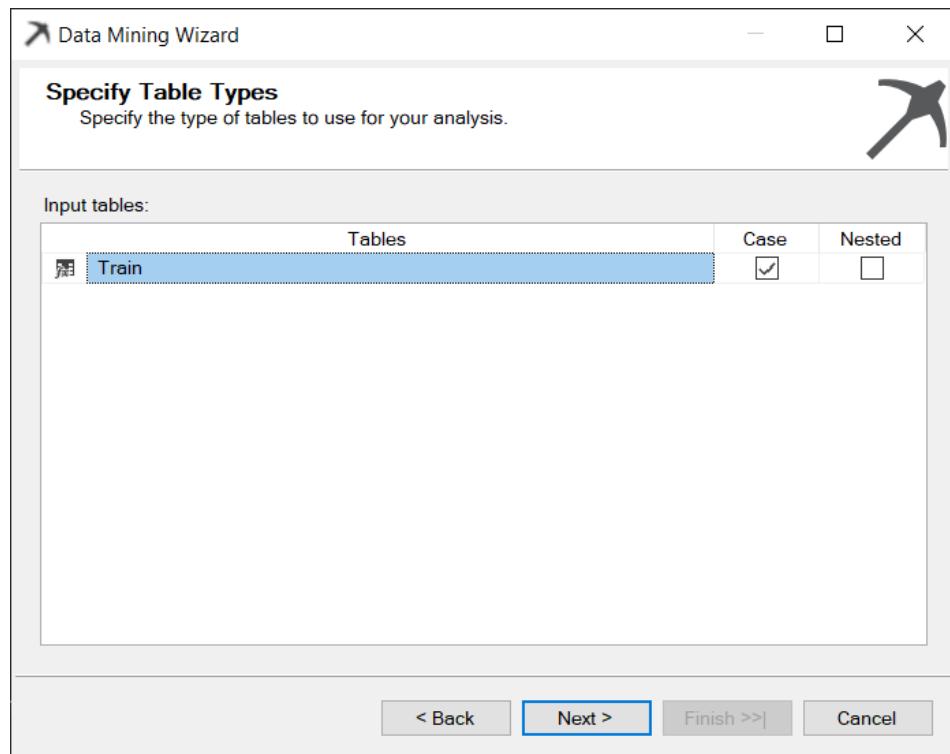
Bước 3: Chọn thuật toán Microsoft Decision Tree và chọn **Next** để tiếp tục



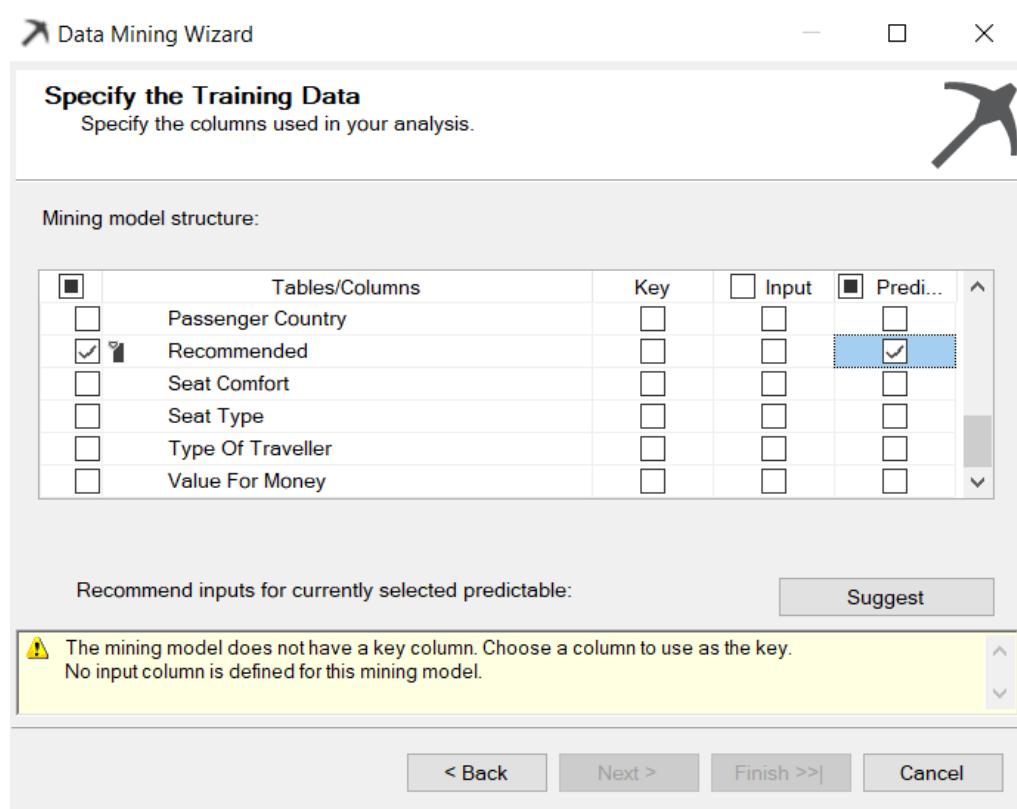
Chọn **Next** để tiếp tục



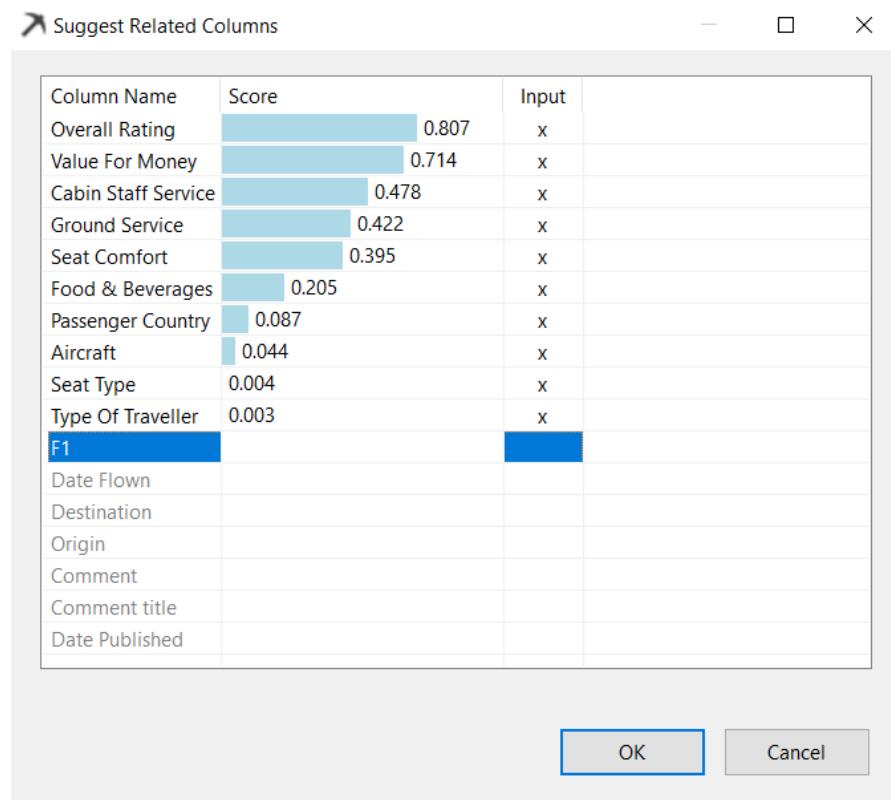
Bước 4: Chọn Case là bảng Train và chọn Next để tiếp tục



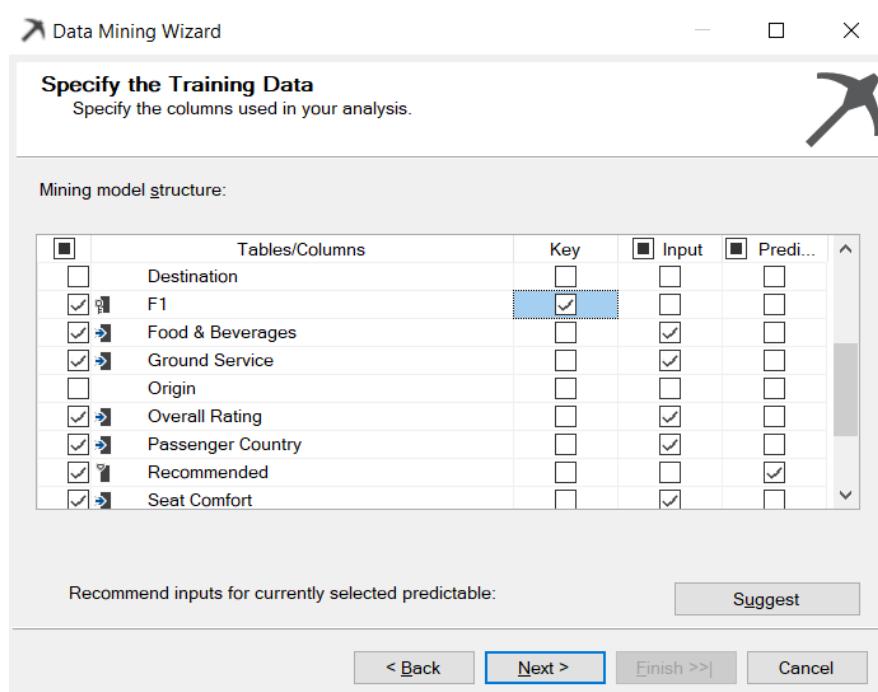
Bước 5: Chọn thuộc tính Recommended là thuộc tính dự đoán



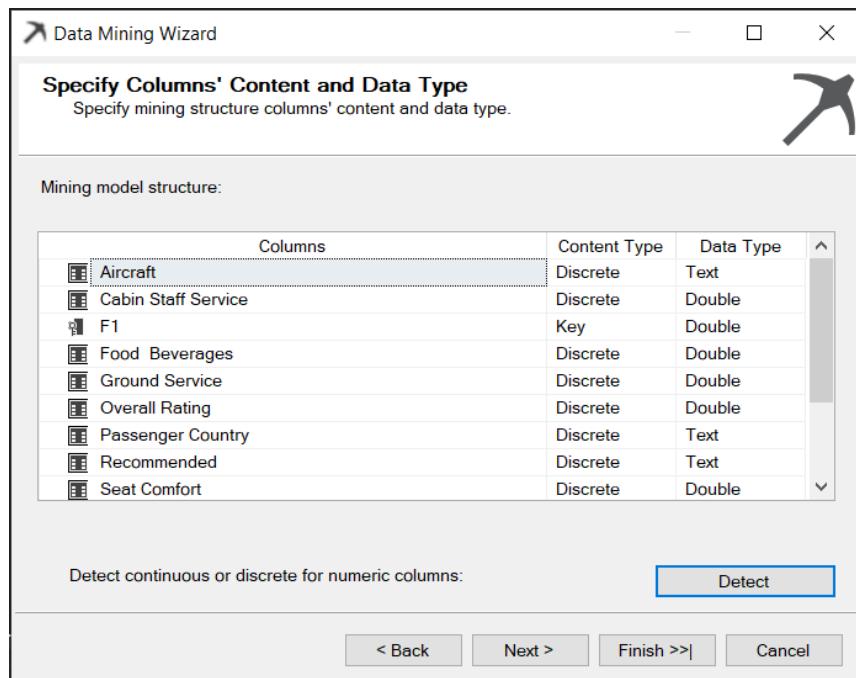
Bước 6: Chọn Suggest để được gợi ý những thuộc tính làm đầu vào, ta có thể thêm hoặc bớt các thuộc tính gợi ý. Chọn **OK** để xác nhận các thuộc tính Input



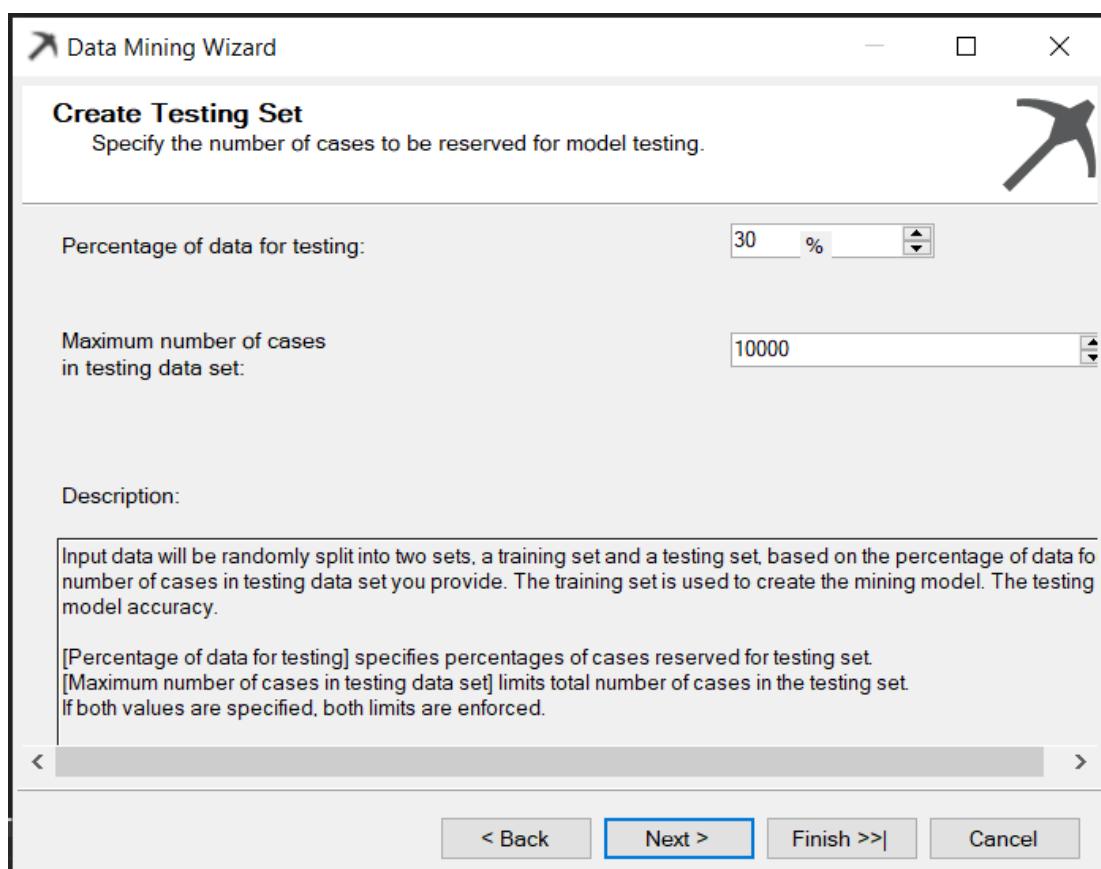
Bước 7: Chọn thuộc tính Key là F1. Chọn Next để tiếp tục



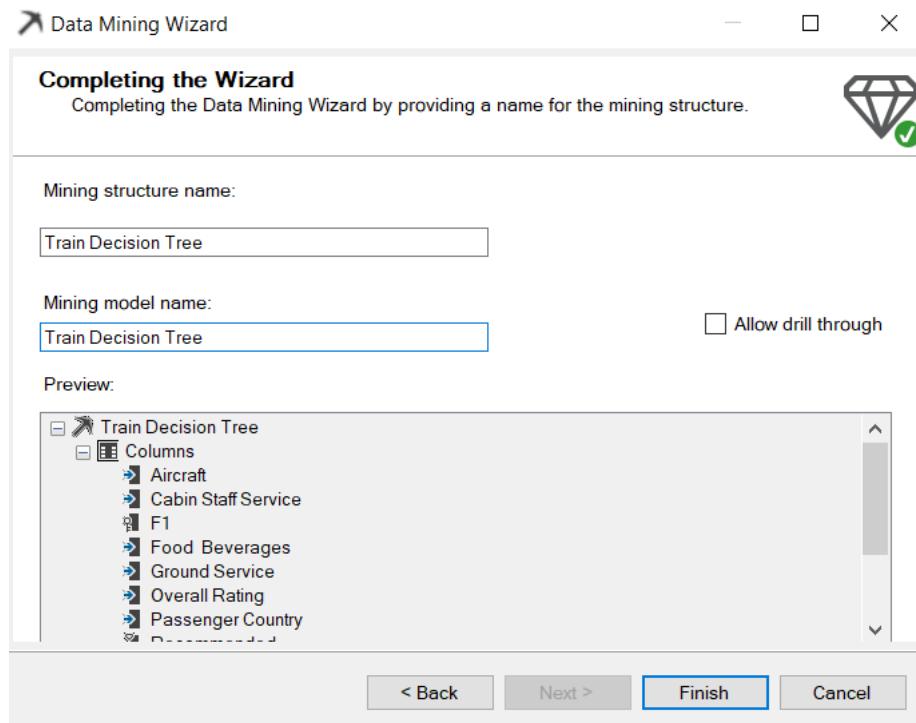
Bước 8: Chọn Detect để chỉnh sửa Content Type là Discrete. Chọn Next để tiếp tục



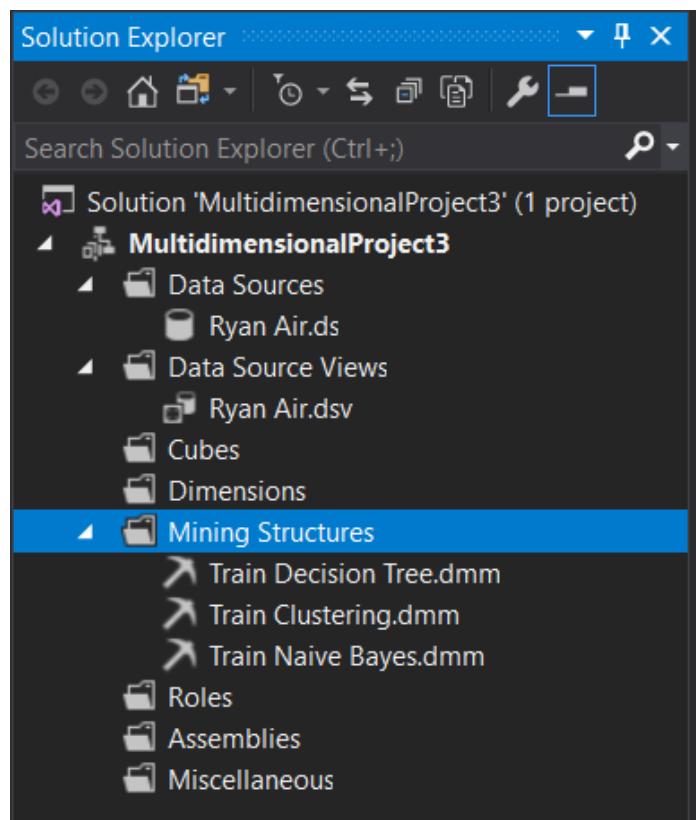
Bước 9: Chọn 30% dữ liệu để Test và tối đa là 10000



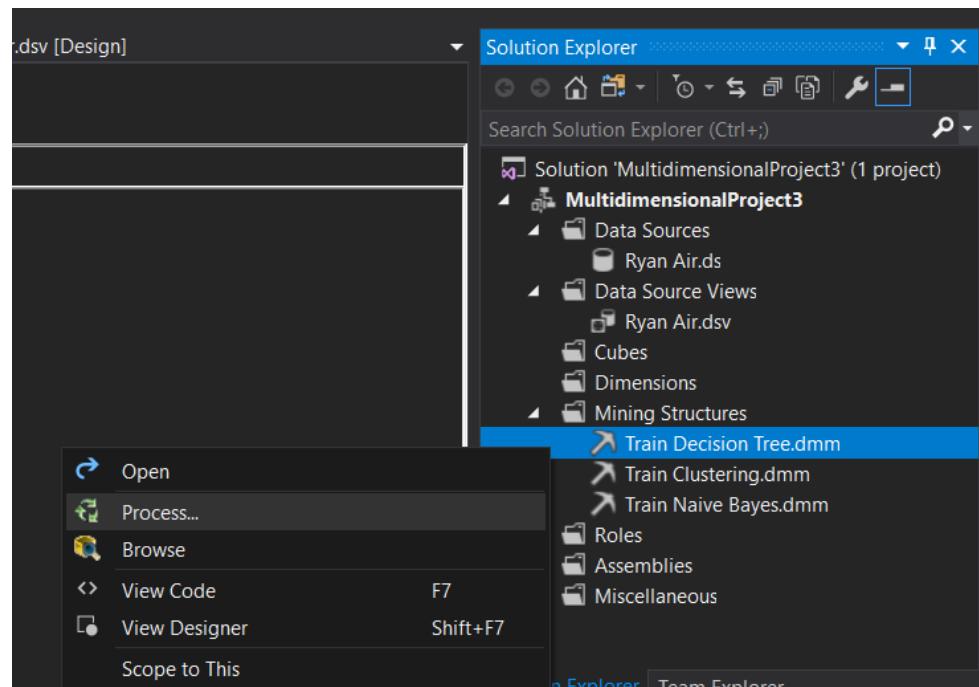
Bước 10: Tiến hành đặt tên và chọn Finish để hoàn thành



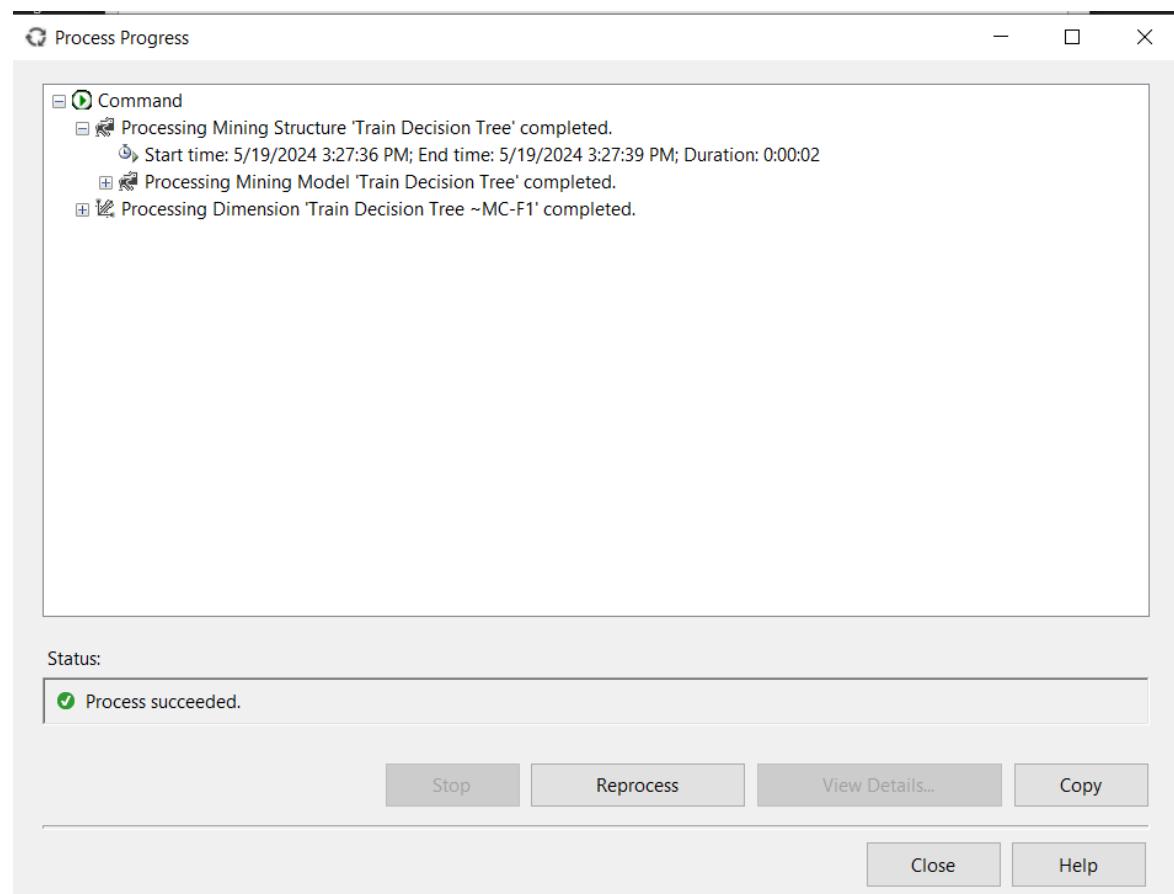
Bước 11: Tiến hành tạo Data Mining bằng thuật toán Cluster và Naïve Bayes



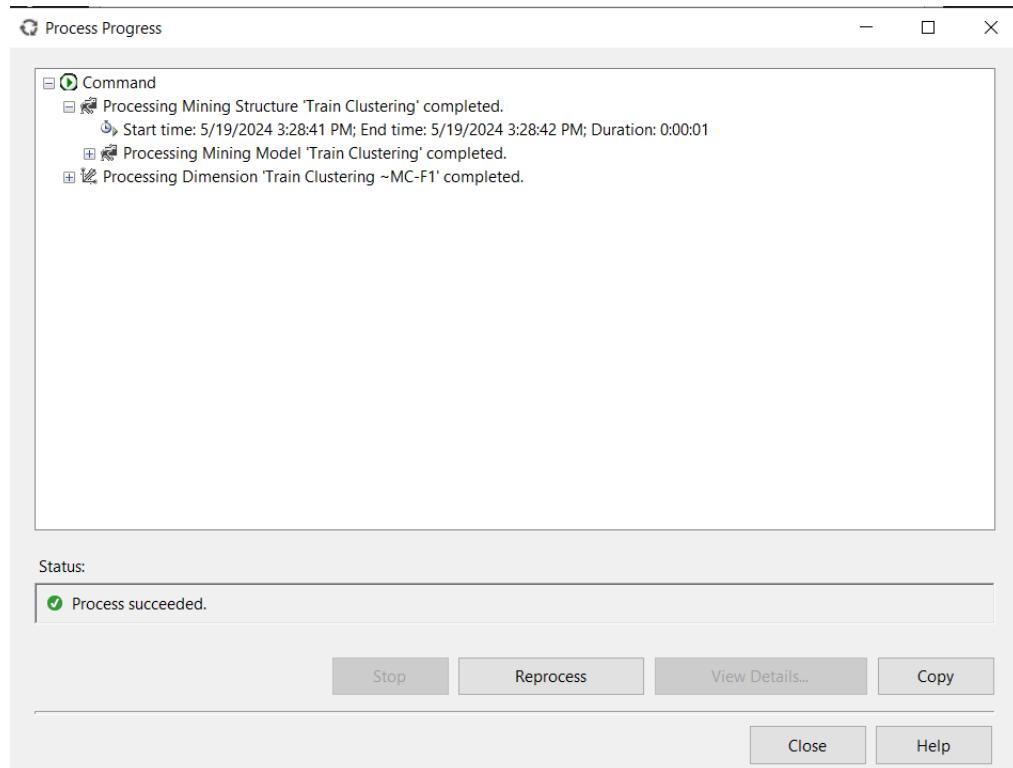
Bước 12: Process để thực thi quá trình Data Mining



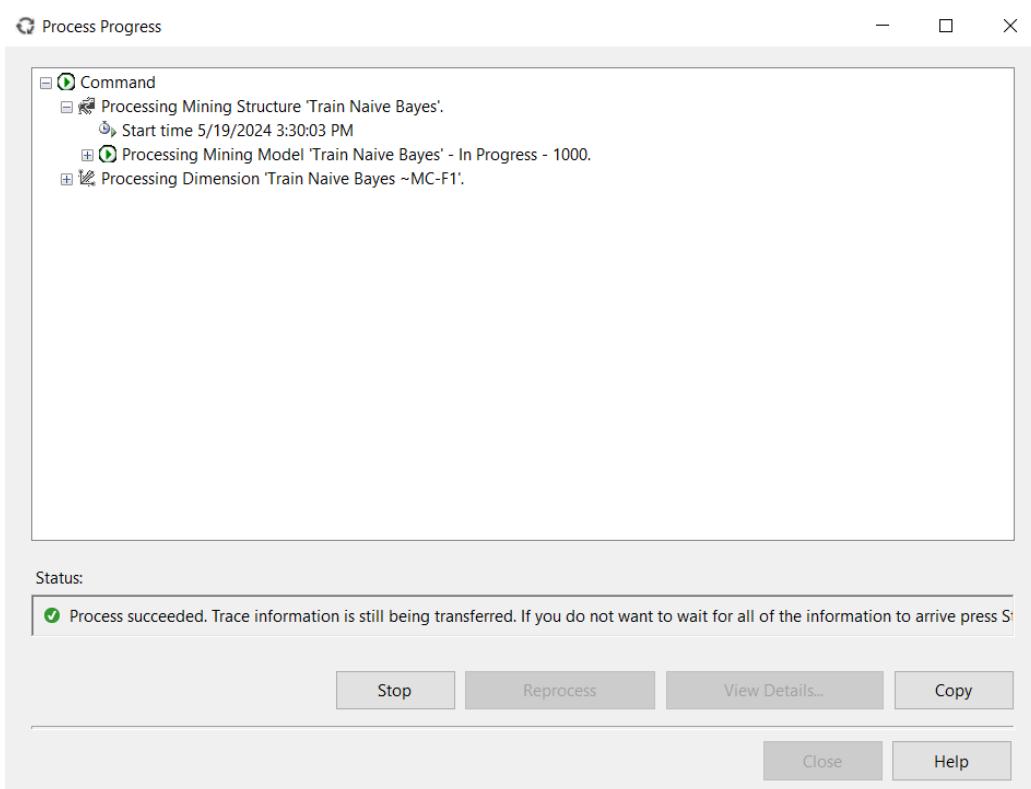
Kết quả:



Process thuật toán Clustering



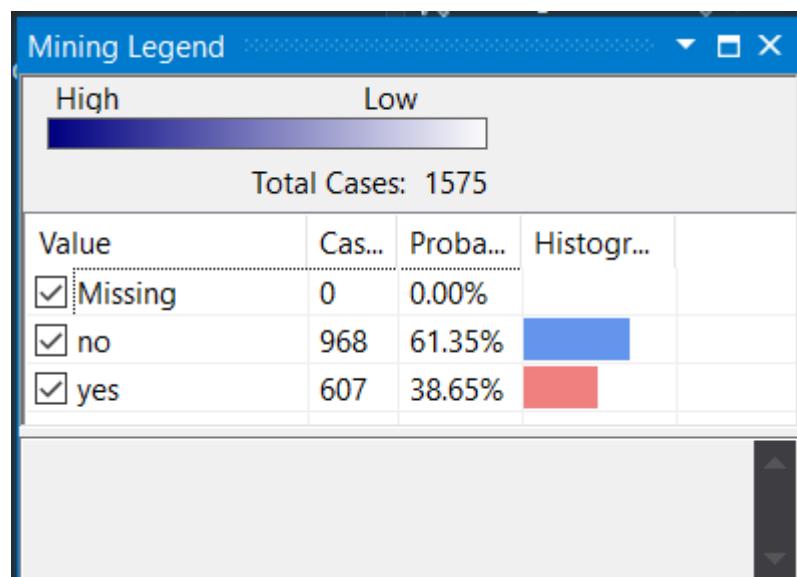
Process thuật toán Naïve Bayes



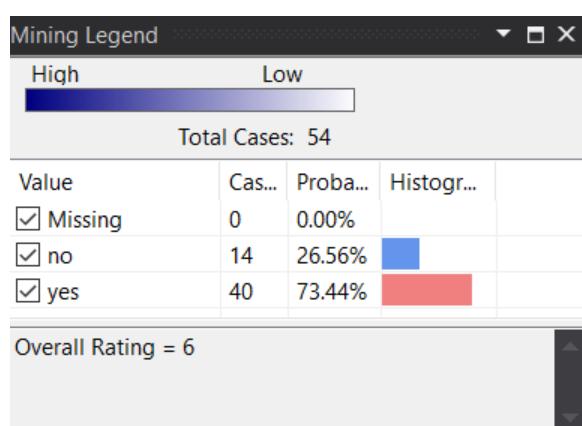
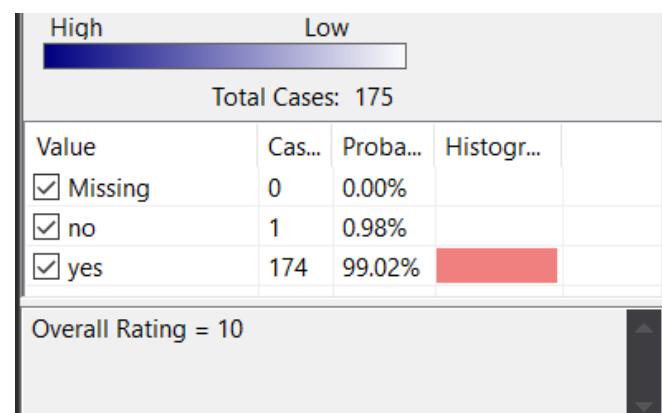
6.2. Nhận xét kết quả sau khi thực hiện Mining dữ liệu

6.2.1. Thuật toán Decision Tree

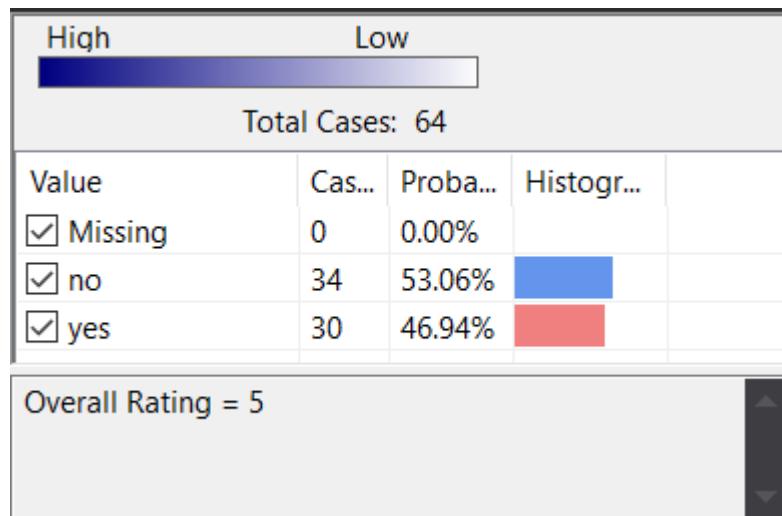




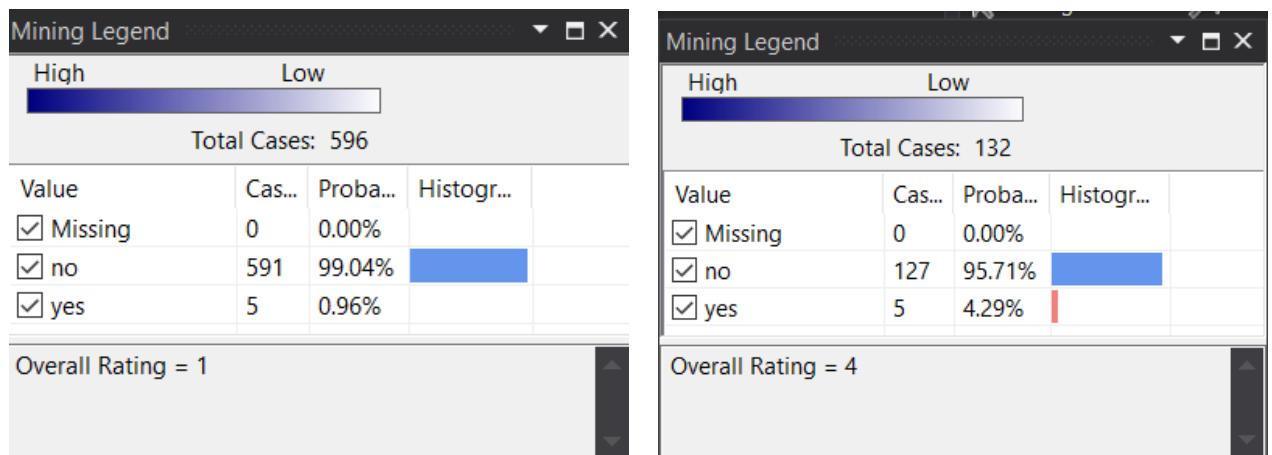
Tổng thể bộ dữ liệu có 61.35% hành khách không hài lòng với chuyến bay và 38.65% hành khách hài lòng với chuyến bay



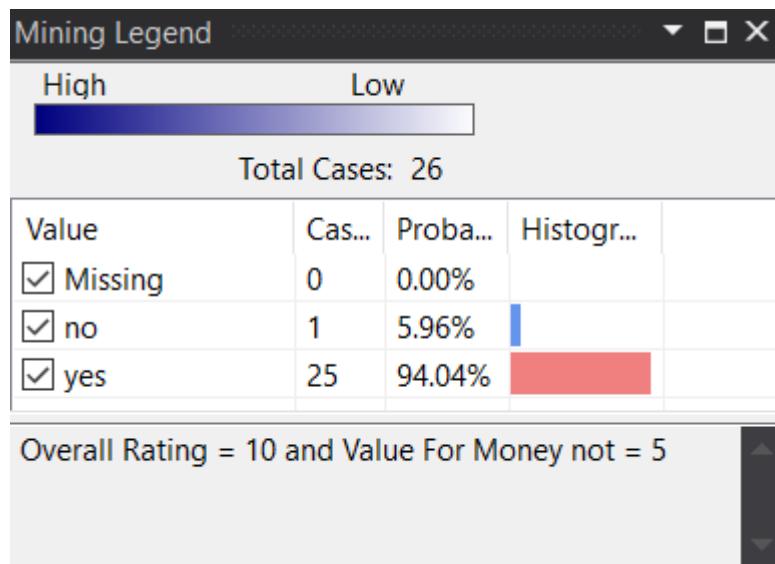
Đối với trung bình tổng thể (Overall Rating) từ 6-10 tỉ lệ hành khách hài lòng với chuyến bay là khá cao từ 73.44% đến 99.02%



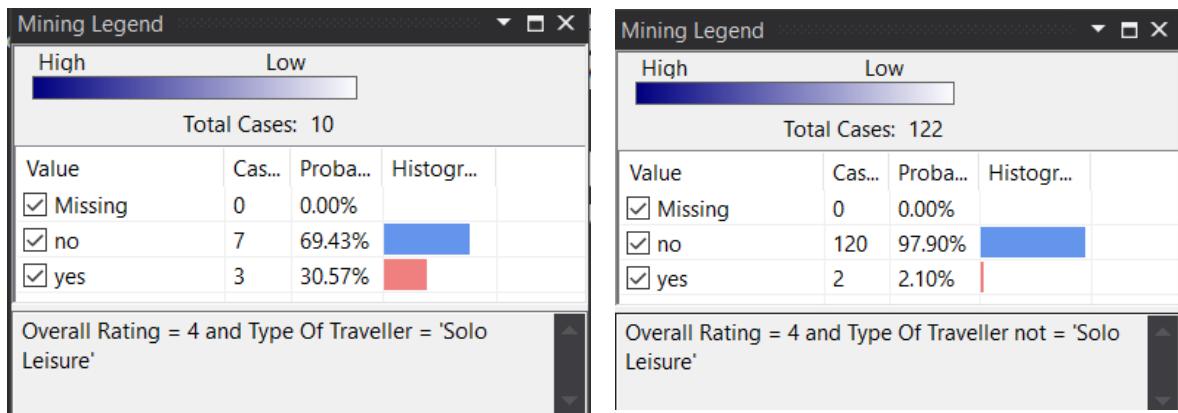
Với mức Overall Rating = 5 thì có khoảng 53.06% khách hàng không hài lòng với chuyến bay và 46.94% khách hàng hài lòng với chuyến bay



Với mức Overall Rating từ (1-4) thì 95.71% đến 99.04% hành khách không hài lòng với chuyến bay

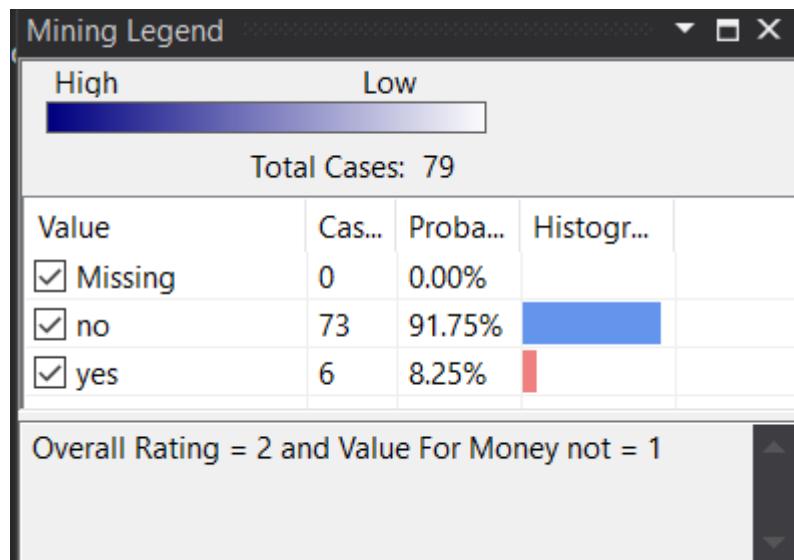


Với mức Overall Rating = 10 nhưng Value For Money khác 5 thì chỉ có khoảng 5.96% hành khách không hài lòng với chuyến bay 94.04% hành khách còn lại hài lòng về chuyến bay



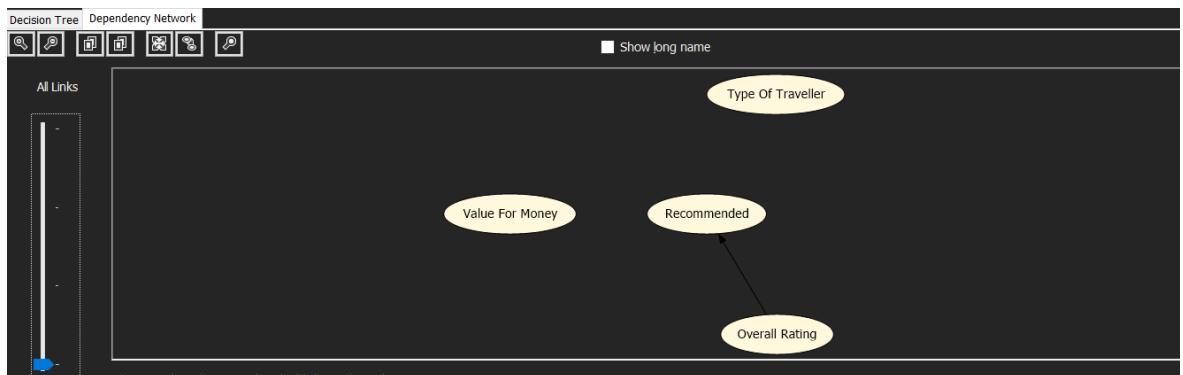
Khi Overall Rating = 4 và loại hình du lịch là 'Solo Leisure' (du lịch một mình) thì có khoảng 69.43% hành khách không hài lòng với chuyến bay và 30.57% hài lòng với chuyến bay

Khi Overall Rating = 4 và loại hình du lịch không phải là 'Solo Leisure' thì có đến 97.9% hành khách không hài lòng với chuyến bay

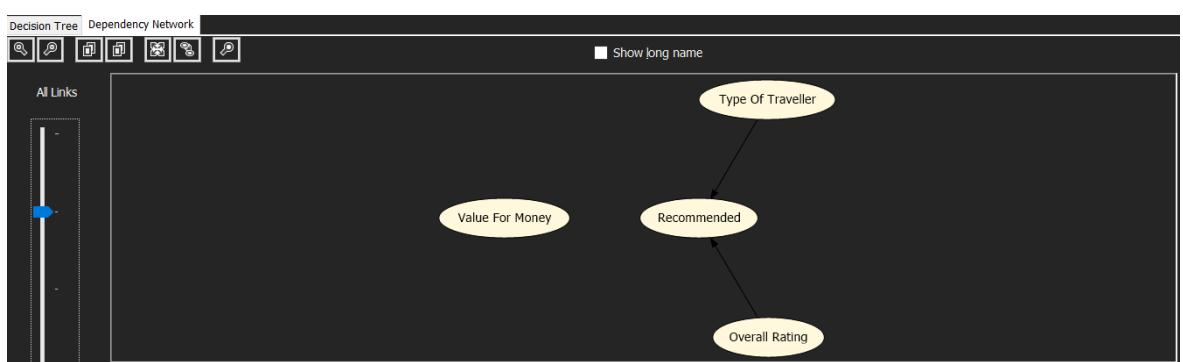


Với mức Overall Rating = 2 và Value For Money khác 1 thì vẫn có 8.25% khách hàng hài lòng với chuyến bay

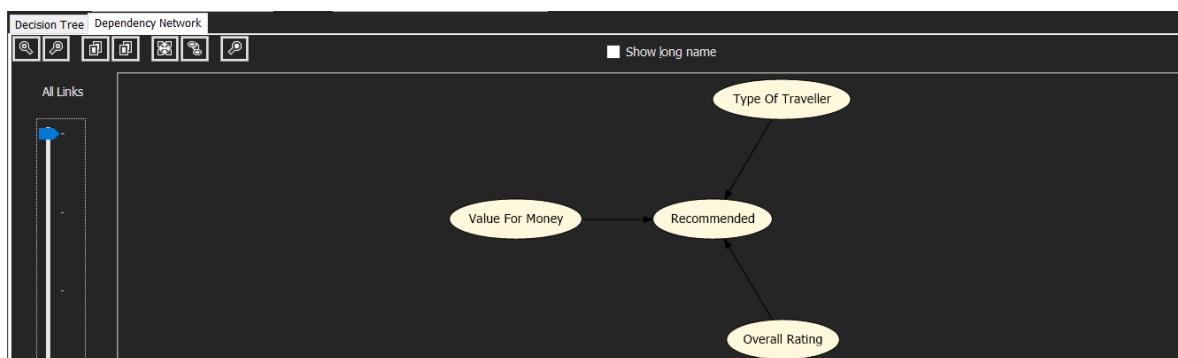
Dependency network: mức độ ảnh hưởng của các thuộc tính theo thứ tự giảm dần
Overall Rating có sự ảnh hưởng lớn nhất đến độ hài lòng của hành khách



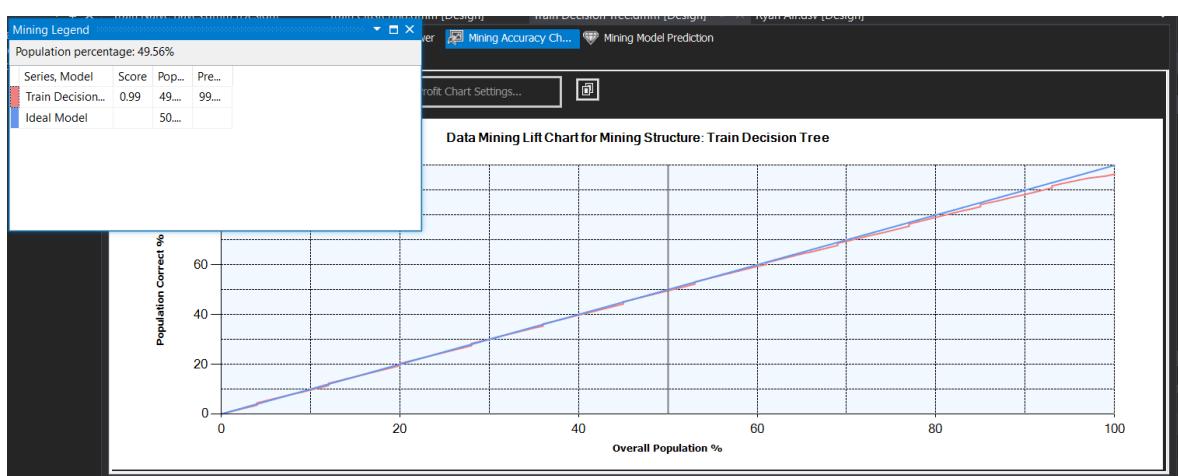
Kế tiếp là thuộc tính Type Of Traveller



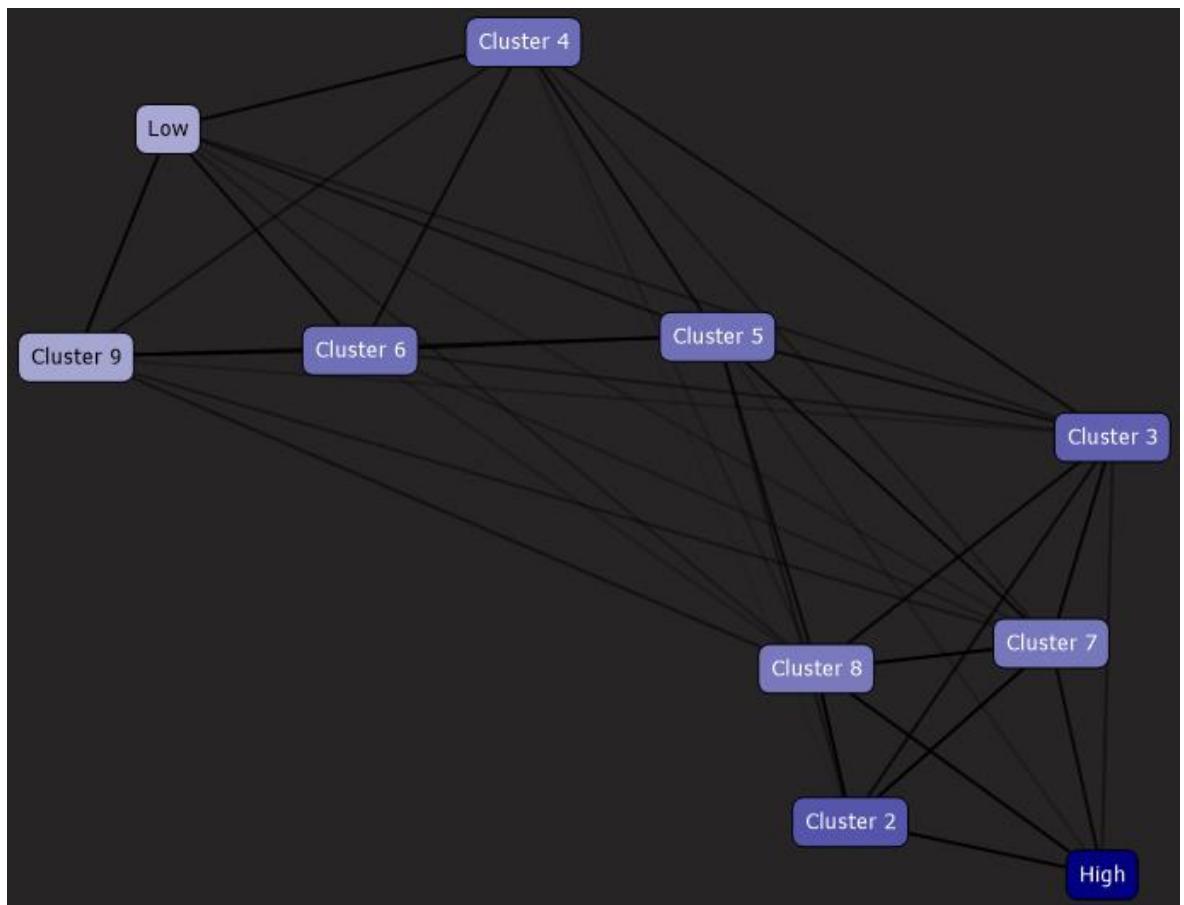
Và cuối cùng là Value For Money



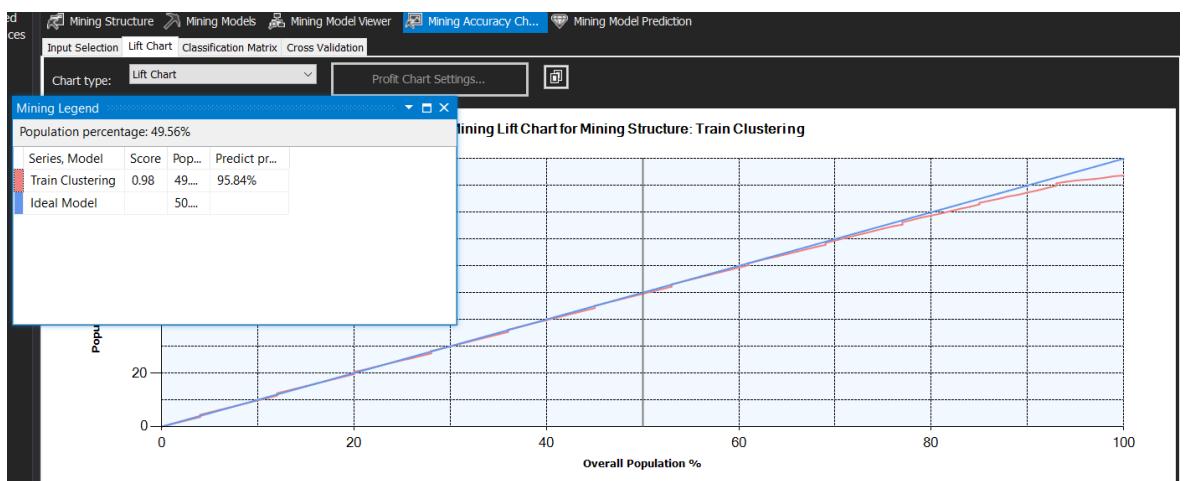
Độ chính xác của thuật toán: thuật toán có độ chính xác là 0.99 điểm



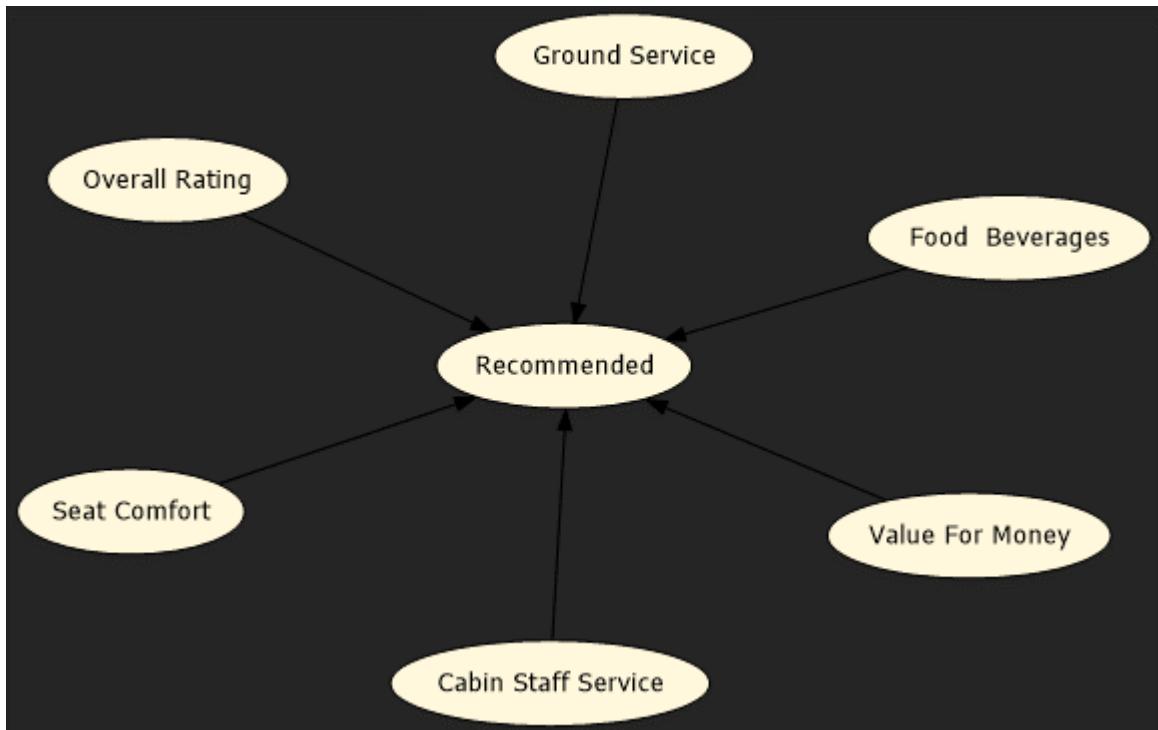
6.2.2. Thuật toán Clustering



Độ chính xác thuật toán: thuật toán có độ chính xác 0.98 điểm

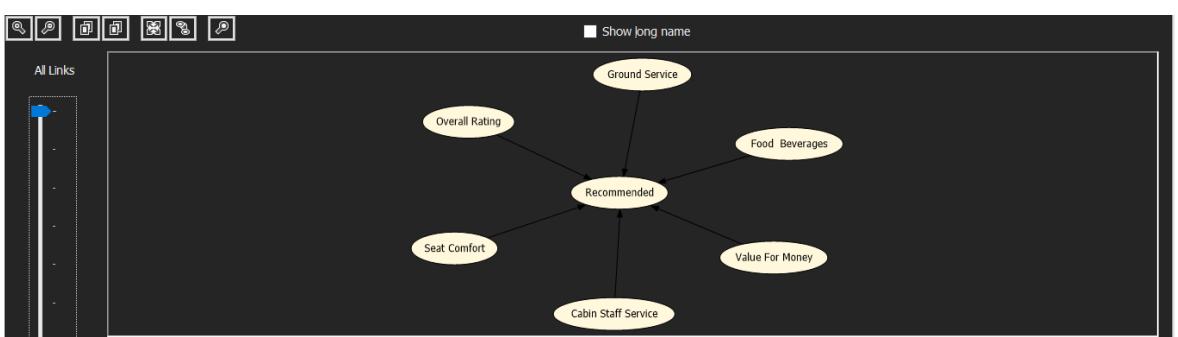
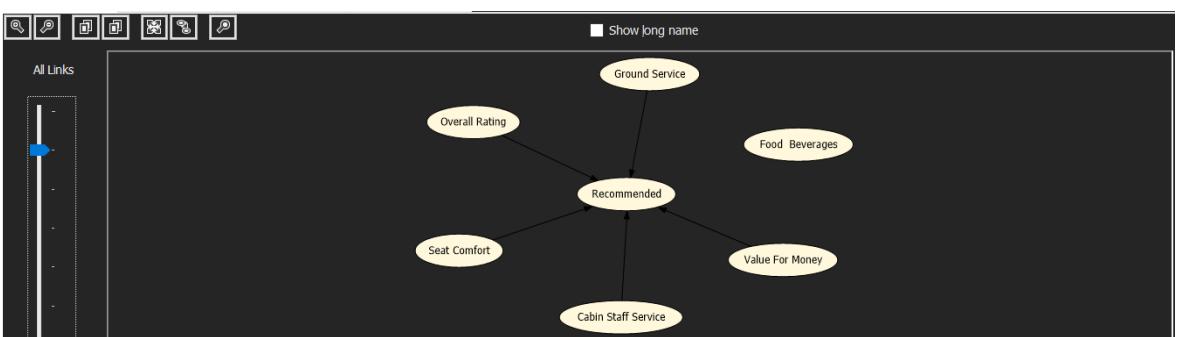
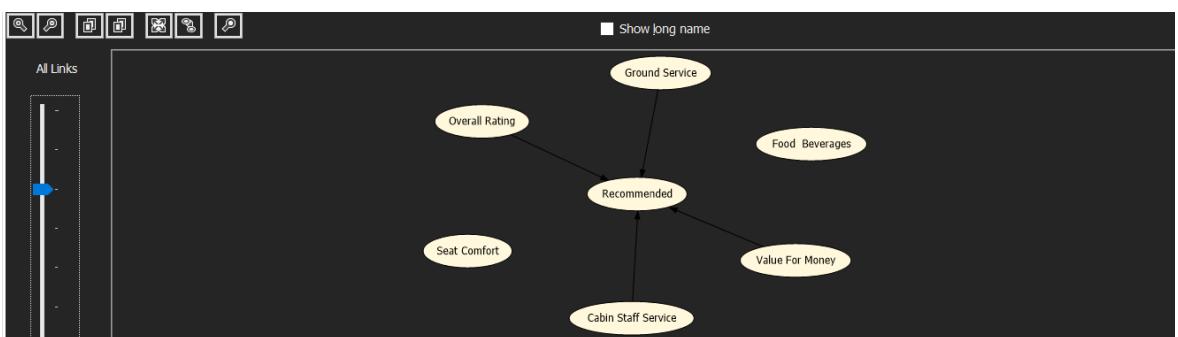
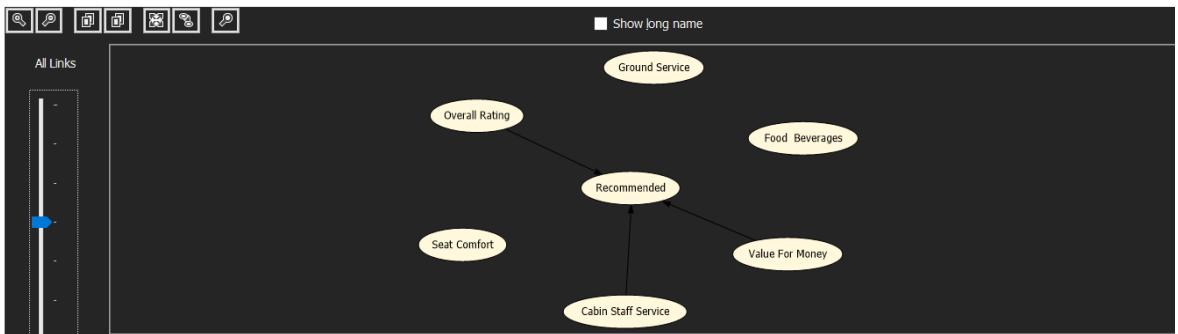


6.2.3. Thuật toán Naïve Bayes

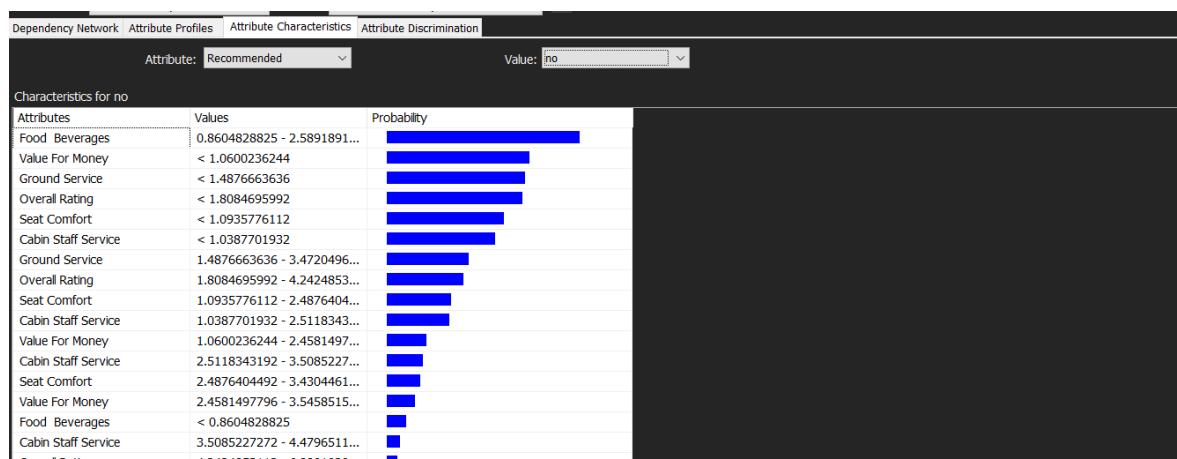


Mức độ ảnh hưởng theo thứ tự giảm dần: Overall Rating, Value For Money, Cabin Staff Service, Ground Service, Seat Comfort, Food Beverages





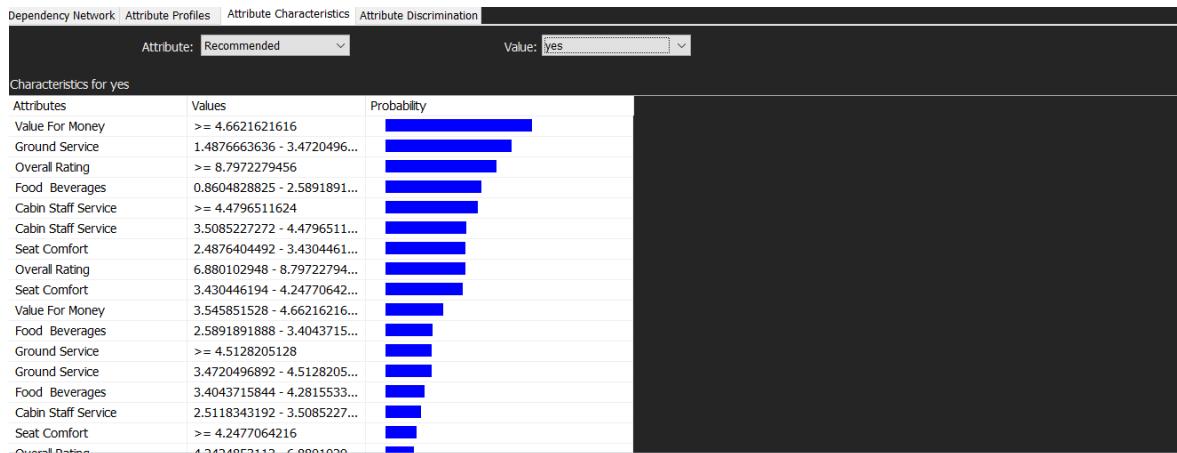
Đặc điểm của các thuộc tính khác khi Recommended là No



Attributes	Values	Probability
Food Beverages	0.8604828825 - 2.5891891888	85.550%
Value For Money	< 1.0600236244	63.560%
Ground Service	< 1.4876663636	61.361%
Overall Rating	< 1.8084695992	60.209%
Seat Comfort	< 1.0935776112	52.251%
Cabin Staff Service	< 1.0387701932	48.063%
Ground Service	1.4876663636 - 3.4720496892	36.649%
Overall Rating	1.8084695992 - 4.2424853112	34.346%
Seat Comfort	1.0935776112 - 2.4876404492	28.796%
Cabin Staff Service	1.0387701932 - 2.5118343192	27.853%
Value For Money	1.0600236244 - 2.4581497796	17.801%
Cabin Staff Service	2.5118343192 - 3.5085227272	16.126%
Seat Comfort	2.4876404492 - 3.430446194	14.974%
Value For Money	2.4581497796 - 3.545851528	12.565%
Food Beverages	< 0.8604828825	8.901%
Cabin Staff Service	3.5085227272 - 4.4796511624	5.969%
Overall Rating	4.2424853112 - 6.880102948	4.712%
Food Beverages	2.5891891888 - 3.4043715844	4.293%
Value For Money	3.545851528 - 4.6621621616	4.188%
Seat Comfort	3.430446194 - 4.2477064216	3.665%
Cabin Staff Service	>= 4.4796511624	1.990%
Value For Money	>= 4.6621621616	1.885%
Ground Service	3.4720496892 - 4.5128205128	1.257%
Food Beverages	3.4043715844 - 4.2815533976	1.047%

Ground Service	≥ 4.5128205128	0.733%
Overall Rating	6.880102948 - 8.7972279456	0.524%

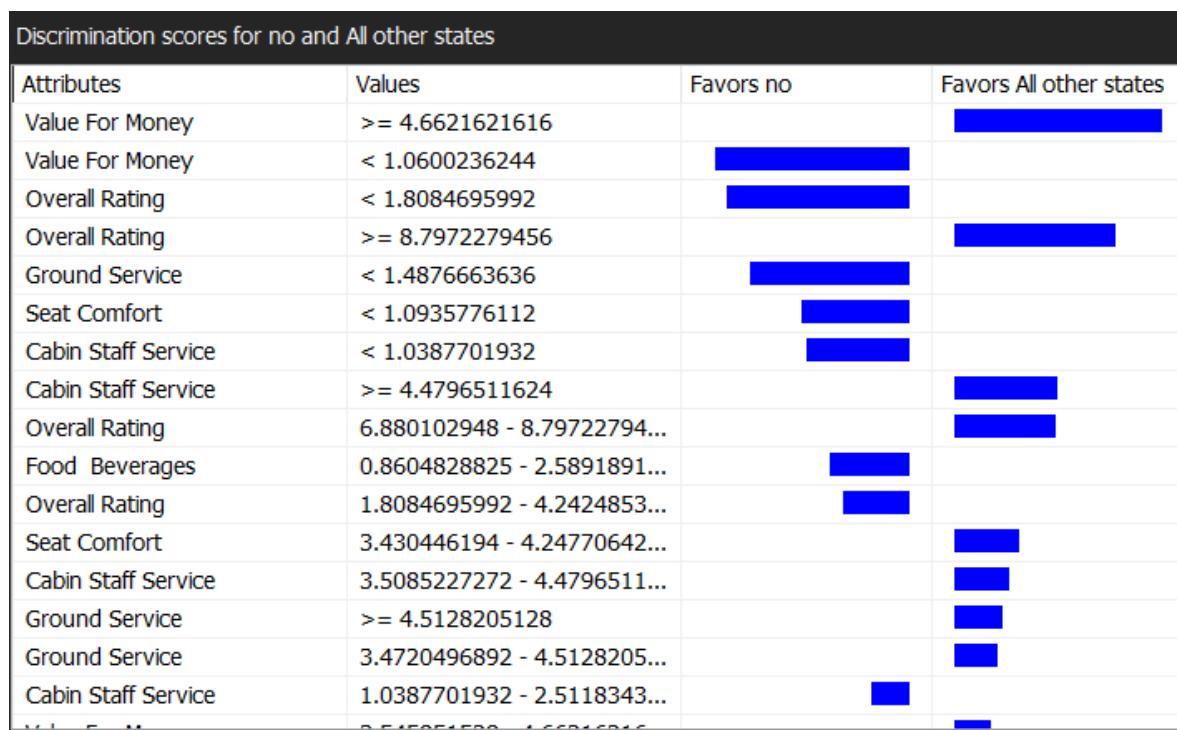
Đặc điểm của các thuộc tính khác khi Recommended là Yes



Attributes	Values	Probability
Value For Money	≥ 4.6621621616	64.839%
Ground Service	1.4876663636 - 3.4720496892	55.806%
Overall Rating	≥ 8.7972279456	49.032%
Food Beverages	0.8604828825 - 2.5891891888	42.258%
Cabin Staff Service	≥ 4.4796511624	40.806%
Cabin Staff Service	3.5085227272 - 4.4796511624	35.645%
Seat Comfort	2.4876404492 - 3.430446194	35.484%
Overall Rating	6.880102948 - 8.7972279456	35.323%
Seat Comfort	3.430446194 - 4.2477064216	34.194%
Value For Money	3.545851528 - 4.6621621616	25.645%
Food Beverages	2.5891891888 - 3.4043715844	20.806%
Ground Service	≥ 4.5128205128	20.645%
Ground Service	3.4720496892 - 4.5128205128	20.484%
Food Beverages	3.4043715844 - 4.2815533976	17.419%
Cabin Staff Service	2.5118343192 - 3.5085227272	15.645%
Seat Comfort	≥ 4.2477064216	13.871%
Overall Rating	4.2424853112 - 6.880102948	12.581%
Seat Comfort	1.0935776112 - 2.4876404492	11.452%
Food Beverages	< 0.8604828825	11.129%
Food Beverages	≥ 4.2815533976	8.387%
Value For Money	2.4581497796 - 3.545851528	8.226%

Seat Comfort	< 1.0935776112	5.000%
Cabin Staff Service	1.0387701932 - 2.5118343192	4.355%
Cabin Staff Service	< 1.0387701932	3.548%
Ground Service	< 1.4876663636	3.065%
Overall Rating	1.8084695992 - 4.2424853112	2.581%
Value For Money	< 1.0600236244	0.806%

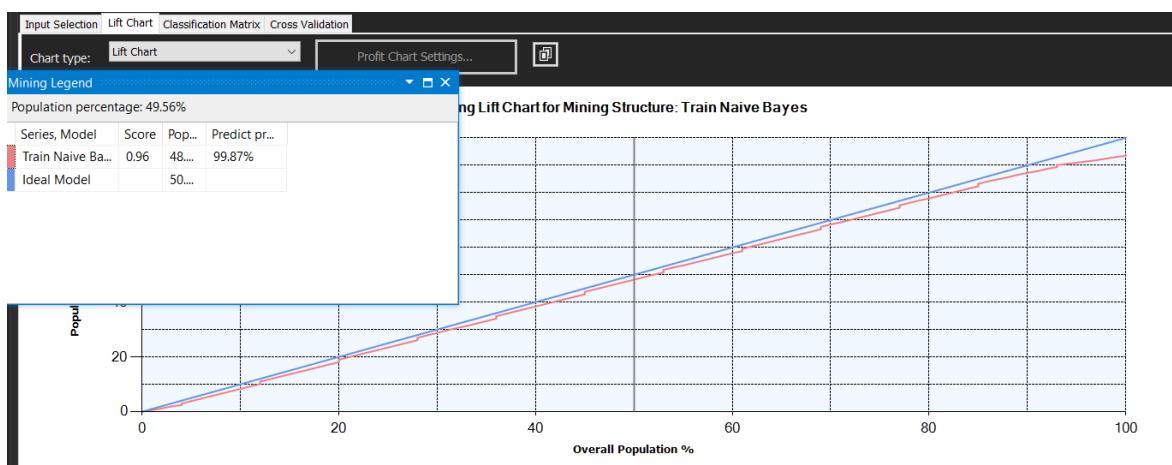
Sự tương quan:



Attributes	Values	Favors no	Favors All other states
Value For Money	≥ 4.6621621616		100.000
Value For Money	< 1.0600236244	93.798	
Overall Rating	< 1.8084695992	88.394	
Overall Rating	≥ 8.7972279456		78.184
Ground Service	< 1.4876663636	77.259	
Seat Comfort	< 1.0935776112	52.182	
Cabin Staff Service	< 1.0387701932	50.017	
Cabin Staff Service	≥ 4.4796511624		49.934
Overall Rating	$6.880102948 - 8.7972279456$		49.345
Food Beverages	$0.8604828825 - 2.5891891888$	38.562	
Overall Rating	$1.8084695992 - 4.2424853112$	32.131	

Seat Comfort	3.430446194 - 4.2477064216		31.731
Cabin Staff Service	3.5085227272 - 4.4796511624		26.598
Ground Service	>= 4.5128205128		23.982
Ground Service	3.4720496892 - 4.5128205128		21.090
Cabin Staff Service	1.0387701932 - 2.5118343192	18.497	
Value For Money	3.545851528 - 4.6621621616		18.073
Value For Money	1.0600236244 - 2.4581497796	17.949	
Food Beverages	3.4043715844 - 4.2815533976		17.679
Seat Comfort	>= 4.2477064216		16.486
Food Beverages	2.5891891888 - 3.4043715844		11.933
Seat Comfort	2.4876404492 - 3.430446194		9.753
Food Beverages	>= 4.2815533976		9.372
Seat Comfort	1.0935776112 - 2.4876404492	7.702	
Ground Service	1.4876663636 - 3.4720496892		5.941
Overall Rating	4.2424853112 - 6.880102948		3.153

Độ chính xác của thuật toán: Thuật toán có độ chính xác là 0.96 điểm



⇒ **Kết luận:** Nên sử dụng thuật toán Decision Tree