

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



BÁO CÁO ĐỒ ÁN CUỐI KỲ
ĐỀ TÀI: XÂY DỰNG KHO DỮ LIỆU
BOOKING HOTEL
Môn học: KHO DỮ LIỆU
Mã lớp học phần: DAWH430784_23_2_03CLC
GVHD: ThS. Nguyễn Văn Thành

Nhóm sinh viên thực hiện: Nhóm 3

Nguyễn Hoàng Yến Nhi	21110267
Lê Thị Khánh Ngân	21110253
Vũ Ngọc Nam	21110252
Trần Văn Tiến	21110319

TP. Hồ Chí Minh, 20 tháng 05 năm 2024

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

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

STT	HỌ VÀ TÊN	MSSV	TỶ LỆ ĐÓNG GÓP
1	Nguyễn Hoàng Yến Nhi	21110267	100%
2	Lê Thị Khánh Ngân	21110253	100%
3	Vũ Ngọc Nam	21110252	100%
4	Trần Văn Tiến	21110319	100%

Nhận xét của giảng viên:

Ngày ... tháng 05 năm 2024

Giảng viên chấm điểm

Ths. Nguyễn Văn Thành

LỜI CẢM ƠN

Lời đầu tiên, nhóm chúng em xin gửi lời cảm ơn chân thành đến Trường Đại học Sư phạm Kỹ thuật TPHCM đã đưa môn học Kho Dữ Liệu vào chương trình giảng dạy. Đặc biệt, chúng em xin gửi lời cảm ơn sâu sắc đến giảng viên bộ môn – Th.S Thầy Nguyễn Văn Thành đã dạy dỗ, truyền đạt những kiến thức quý báu cho chúng em trong suốt thời gian học tập vừa qua. Trong thời gian tham gia lớp học Kho Dữ Liệu của Thầy, nhóm em đã có thêm cho mình nhiều kiến thức bổ ích, tinh thần học tập hiệu quả, nghiêm túc. Đây chắc chắn sẽ là những kiến thức quý báu, là hành trang để chúng em có thể vững bước sau này. Bộ môn Kho Dữ Liệu là môn học thú vị, vô cùng bổ ích và có tính thực tế cao. Đảm bảo cung cấp đủ kiến thức, gắn liền với nhu cầu thực tiễn của sinh viên. Mặc dù chúng em đã cố gắng hết sức nhưng chắc chắn bài đồ án này khó có thể tránh khỏi những thiếu sót và nhiều chỗ còn chưa chính xác, kính mong Thầy xem xét và góp ý để bài đồ án của chúng em được hoàn thiện hơn.

Nhóm chúng em xin chân thành cảm ơn!

MỤC LỤC

CHƯƠNG 1: TỔNG QUAN VỀ ĐỀ TÀI	1
1.1. Lí do chọn đề tài	1
1.2. Tổng quan về tập dữ liệu	1
1.2.1. Nguồn dữ liệu	1
1.2.2. Mô tả tập dữ liệu.....	1
1.2.3. Giới thiệu các công cụ được sử dụng	2
CHƯƠNG 2 : THIẾT KẾ CƠ SỞ DỮ LIỆU TÁC NGHIỆP (OLAP)	3
2.1 Xác định các Business Process và bảng Fact	3
2.1.1. Business Process: Booking_Fact.....	3
2.1.2. Business Process: Cancellation_Fact	4
2.2. Xây dựng các bảng Dimension.....	4
2.2.1. DimBooking	4
2.2.2. DimRoomType	5
2.2.3. DimDate.....	5
2.2.4. DimMarket_Segment_type.....	5
2.2.5. DimMealType.....	6
2.2.6. DimCustomerType	6
CHƯƠNG 3: TÍCH HỢP DỮ LIỆU VÀO KHO (SSIS)	8
3.1. Tạo ETL project và solution	8
3.2. Thiết lập các Connection	9
3.3. Import Dữ liệu vào các bảng dimension.....	14
3.3.1 Booking Dimension	14
3.3.1.1 Extract.....	15
3.3.1.2 Load	17
3.3.1.3 Run.....	19
3.3.2 Date Dimension	20
3.3.2.1 Extract.....	21
3.3.2.2 Load	23

3.3.2.3 Run.....	25
3.3.3 Customer Type Dimension	26
3.3.3.1 Extract.....	27
3.3.3.2 Load	29
3.3.3.3 Run.....	31
3.3.4 Market Segment Type	32
3.3.4.1 Extract.....	32
3.3.4.2 Load	34
3.3.4.3 Run.....	36
3.3.5 Meal Type Dimension	38
3.3.5.1 Extract.....	38
3.3.5.2 Load	39
3.3.5.3 Run.....	41
3.3.6 Room Type Dimension.....	43
3.3.6.1 Extract.....	43
3.3.6.2 Load	45
3.3.6.3 Run.....	47
3.4 Import dữ liệu vào bảng các bảng fact.....	50
3.4.1 Booking Fact.....	50
3.4.1.1 Extract.....	51
3.4.1.2 Load	52
3.4.1.3 Run.....	59
3.4.2 Cancellation Fact	61
3.4.2.1 Extract.....	61
3.4.2.2 Load	63
3.4.2.3 Run.....	65
CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS)	68
4.1. Quá trình xây dựng mô hình	68
4.1.1. Tạo Data Source	68
4.1.2. Tạo Data Source View.....	70

4.2. Quá trình xây dựng khối Cube.....	74
4.3. Cấu hình Hierarchy	78
4.3.1. Tạo Hierarchy cho Dim Date.....	78
4.3.2. Tạo Hierarchy cho Dim Customer Type	78
4.3.3. Tạo Hierarchy cho Dim Market Segment Type	79
4.3.4. Tạo Hierarchy cho Dim Market Segment Type	79
4.3.5. Tạo Hierarchy cho Dim Room Type	80
4.3.4. Tạo Hierarchy cho DimBooking	80
4.4. Deploy Process Cube	81
4.5. Thực hiện phân tích dữ liệu	81
4.5.1. Cho biết loại phòng được đặt nhiều nhất và doanh thu từ việc đặt phòng theo tháng quý năm.....	81
4.5.1.1. Sử dụng công cụ SSAS.....	81
4.5.1.2. Sử dụng Pivot Table trong Excel	83
4.5.1.3. Sử dụng Power BI.....	84
4.5.2. Cho biết khách hàng thường chọn thực đơn nào nhiều nhất	84
4.5.2.1. Sử dụng công cụ SSAS.....	84
4.5.2.2. Sử dụng Pivot Table trong Excel	85
4.5.2.3. Sử dụng Power BI.....	86
4.5.3. Thống kê số lượng khách hàng của mỗi loại phân khúc thị trường	86
4.5.3.1. Sử dụng công cụ SSAS.....	86
4.5.3.2. Sử dụng Pivot Table trong Excel	88
4.5.3.3. Sử dụng Power BI.....	88
4.5.4. Thống kê doanh thu theo loại khách (repeated_guest).....	89
4.5.4.1.Sử dụng công cụ SSAS.....	89
4.5.4.2.Sử dụng Pivot Table trong Excel	90
4.5.4.3. Sử dụng Power BI.....	91
4.5.5. Thống kê lượt đặt phòng mỗi tháng qua đó tiến hành thống kê lượt đặt thành công và lượt hủy mỗi tháng	92
4.5.5.1.Sử dụng công cụ SSAS.....	92
4.5.5.2.Sử dụng Pivot Table trong Excel	94

4.5.5.3. Sử dụng Power BI.....	96
CHƯƠNG 5: KẾT LUẬN	97
5.1. Kết quả đạt được	97
5.2. Hạn chế	97
5.3. Tài liệu tham khảo	97

CHƯƠNG 1: TỔNG QUAN VỀ ĐỀ TÀI

1.1. Lý do chọn đề tài

Trong thời đại công nghệ ngày nay, việc đặt phòng khách sạn thông qua các ứng dụng và trang web trở nên phổ biến, vì vậy việc có một hệ thống quản lý dữ liệu booking hotel là rất quan trọng. Từ việc phân tích xu hướng thị trường đến tối ưu hóa trải nghiệm của khách hàng, một kho dữ liệu về đặt phòng khách sạn có thể cung cấp thông tin giúp các doanh nghiệp trong ngành du lịch và khách sạn nắm bắt được nhu cầu của thị trường và tối ưu hóa hoạt động kinh doanh của mình. Chính vì lí do đó, nhóm sinh viên chúng em chọn đề tài "Xây dựng Kho dữ liệu booking hotel" với mong muốn tạo ra một Kho dữ liệu giúp khách sạn dễ dàng quản lý, thao tác và từ đó đưa ra những quyết định đúng đắn.

1.2. Tổng quan về tập dữ liệu

1.2.1. Nguồn dữ liệu

- Nhóm sử dụng tập dữ liệu “Hotel Reservations” được lấy trên trang web Kaggle ([kaggle.com](https://www.kaggle.com/datasets/ahsan81/hotel-reservations-classification-dataset)).
- Đường dẫn tải tập dữ liệu: <https://www.kaggle.com/datasets/ahsan81/hotel-reservations-classification-dataset>

1.2.2. Mô tả tập dữ liệu

Tập dữ liệu “Hotel Reservations” bao gồm các thông tin khác nhau về đặt phòng khách sạn

Tên cột	Mô tả	Số dòng
Booking_ID	Mã Booking, dùng để phân biệt các thông tin booking được thực hiện	36275
no_of_adults	Số lượng người lớn	36275
no_of_children	Số lượng trẻ em	36275
no_of_weekend_nights	Số đêm cuối tuần (Thứ 7, Chủ nhật) mà khách đã ở hoặc đặt phòng tại khách sạn.	36275
no_of_week_nights	Số đêm trong tuần (Thứ 2 tới Thứ 6) mà khách đã ở hoặc đặt phòng tại khách sạn.	36275

type_of_meal_plan	Loại thực đơn được khách hàng đặt	36275
required_car_parking_space	Yêu cầu về chỗ đậu xe: 0:không cần, 1:cần	36275
room_type_reserved	Loại phòng	36275
lead_time	Khoảng cách ngày giữa ngày đặt và ngày đến nhận phòng	36275
arrival_year	Năm đến	36275
arrival_month	Tháng đến	36275
arrival_date	Ngày đến	36275
market_segment_type	Phân khúc thị trường	36275
repeated_guest	Có phai khách quen không	36275
no_of_previous_cancellations	Số lần đặt trước đó mà khách hàng đã hủy trước khi đặt hiện tại.	36275
no_of_previous_bookings_not_canceled	Số lần đặt trước đó mà không bị khách hàng hủy trước khi đặt hiện tại	36275
avg_price_per_room	Giá trung bình mỗi ngày của đặt phòng, giá phòng biến động	36275
no_of_special_requests	Yêu cầu đặt biệt của khách hàng (ví dụ tầng cao,)	36275
booking_status	Đặt phòng đã bị hủy hay chưa	36275

1.2.3. Giới thiệu các công cụ được sử dụng

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

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

2.1 Xác định các Business Process và bảng Fact

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

Instructions!					Dim Date	Dim Booking	Dim ParkingSpace	Market_Segment	Dim MealType	Dim CustomerType	Dim RoomType
Business Process Name	Fact Table	Fact Grain Type	Granularity	Facts							
Booking	Booking_fact	Transaction	one row per order detail	date, avg_price_per_room, total (avg_price_per_room*(no_of_weekend_nights+ no_of_week_nights))	X	X	X	X	X	X	X
Cancellation	Cancellation_fact	Transaction	one row per order	date, leadtime, bookingStatus	X	X		X		X	X

2.1.1. Business Process: Booking_Fact

- Cho biết loại phòng được đặt nhiều nhất và doanh thu từ việc đặt phòng theo tháng quý năm
- Cho biết khách hàng thường chọn thực đơn nào nhiều nhất
- Thống kê số lượng khách hàng của mỗi loại phân khúc thị trường
- Thống kê doanh thu theo loại khách (repeated_guest)
- Thống kê lượt đặt thành công và lượt hủy mỗi tháng

Table Name	Booking_fact	Home Page									
Table Type	Fact										
Display Name	Booking_Fact										
Database Schema											
Table Description	Booking_Fact										
Comment	Comes from the Hotel Reservations dataset										
Biz Filter Logic											
Size	One row per order detail										
Generate Script?	Y										
Column Name	Display Name	Description	Example Values	Display Folder	ETL Rules	Comments					
DateKey	DateKey										
BookingKey	BookingKey										
ParkingSpaceKey	ParkingSpaceKey										
MarketSegmentKey	MarketSegmentKey										
MealKey	MealKey										
CustomerTypeKey	CustomerTypeKey										
RoomTypeKey	RoomTypeKey										
avg_price_per_room	avg_price_per_room										
total	total										

2.1.2. Business Process: Cancellation_Fact

Table Name	Cancellation_fact	Home Page
Table Type	Fact	
Display Name	Cancellation_Fact	
Database Schema		
Table Description	Cancellation_Fact	
Comment	Comes from the Hotel Reservations dataset	
Biz Filter Logic		
Size	One row per order detail	
Generate Script?	Y	
Column Name	Display Name	Description
DateKey	DateKey	
BookingKey	BookingKey	
MarketSegmentKey	MarketSegmentKey	
BookingStatus	BookingStatus	BookingStatus
Lead_time	Lead_time	Number of days between the date of booking and t 0, 1, 2, ...
RoomTypeKey	RoomTypeKey	

2.2. Xây dựng các bảng Dimension

2.2.1. DimBooking

Table Name	DimBooking	Home Page
Table Type	Dimension	
Display Name	Booking	
Database Schema		
Table Description	Booking dimension	
Comment	comes from the Hotel Reservations dataset	
Biz Filter Logic		
Size	one for each booking	
Generate Script?	Y	
Column Name	Display Name	Description
BookingKey	BookingKey	Surrogate primary key
BookingID	BookingID	unique identifier of each booking
BookingStatus	BookingStatus	Flag indicating if the booking was canceled or not.
SpecialRequests	SpecialRequests	Total number of special requests made by the customer (e.g. high floor, view from the room, etc)
avg_price_per_room	avg_price_per_room	Average price per day of the reservation; prices of the rooms are dynamic. (in euros)
Lead_time	Lead_time	Number of days between the date of booking and the arrival
Weekend_nights	Weekend_nights	Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel
Week_nights	Week_nights	Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel
no_of_childrens	no_of_childrens	
required_car_parking_space	required_car_parking_space	

2.2.2. DimRoomType

Table Name	DimRoomType	Home Page							
Table Type	Dimension								
Display Name	RoomType								
Database Schema	DimRoomType								
Table Description	Comment	Comes from the Hotel Reservations dataset							
Biz Filter Logic									
Size									
Generate Script?	Y								
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Display Folder	ETL Rules	Comments	
RoomTypeKey room_type_reserved	RoomTypeKey room_type_reserved	Surrogate primary key Type of room reserved by the customer. The values are ciphered (encoded) by INN Hotels	-1 -1	1, 2, 3... Room_Type 1, Room_Type 2, Room_Type 4 ...	key				

2.2.3. DimDate

Table Name	DimDate	Home Page						
Table Type	Date							
Display Name	Date							
Database Schema	Date Dimension							
Table Description	Date Dimension							
Comment	Comment							
Biz Filter Logic	Comes from the Hotel Reservations dataset							
Size								
Generate Script?	Y							
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type			
DateKey Arrival_date	DateKey Arrival_date	Surrogate primary key Year of arrival date	-1 1, 2, 3, ...	1, 2, 3, ...	key			
Arrival_month	Arrival_month	Month of arrival date		1, 2, 3, ...	2			
Arrival_year	Arrival_year	Date of the month		1, 2, 3, ...	2			
Quarter	Quarter	Quarter		2020	2			
				1, 2, 3, 4	2			

2.2.4. DimMarket_Segment_type

Table Name	DimMarketSegmentType	Home Page						
Table Type	Dimension							
Display Name	MarketSegmentType							
Database Schema	DimMarketSegmentType							
Table Description	Comment							
Biz Filter Logic	Comes from the Hotel Reservations dataset							
Size								
Generate Script?	Y							
Column Name	Display Name	Description	Unknown Member	Example Values	SCD Type	Display Folder	ETL Rules	Comments
MarketSegmentKey market_segment_type	MarketSegmentKey Market_segment_type	Surrogate primary key Market segment designation	-1 -1	1, 2, 3... Online,Offline,Corporate	key			

2.2.5. DimMealType

Table Name	DimMealType	Home Page		
Table Type	Dimension			
Display Name	Meal			
Database Schema				
Table Description	Meal dimension			
Comment	comes from the Hotel Reservations dataset			
Biz Filter Logic				
Size	one for each mealtype			
Generate Script?	Y			
Column Name	Display Name	Description	Unknown Member	SCD Type
MealKey	MealKey	Surrogate primary key	-1	1, 2, Not Selected, Meal Plan 1, Meal Plan 2...
type_of_meal_plan	type_of_meal_plan	Type of meal plan booked by the customer:		

2.2.6. DimCustomerType

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

3.1. Tạo ETL project và solution

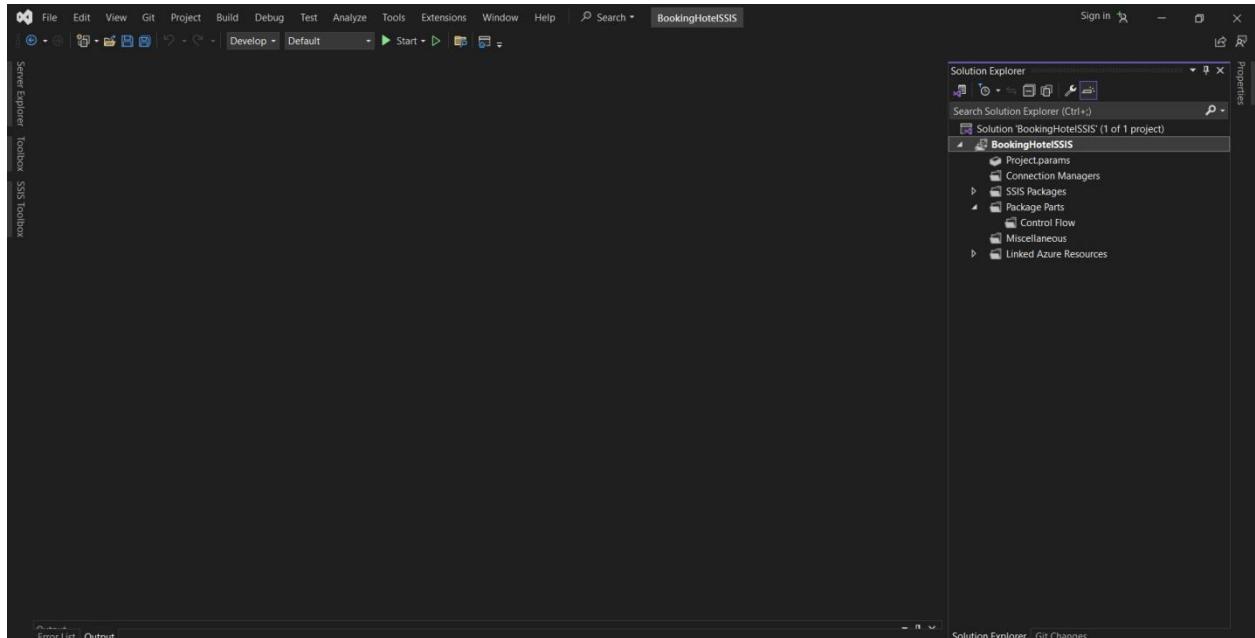
1. Từ **Visual Studio** menu, chọn **File -> New -> Project**.

2. Từ **dialog**:

a. Chọn **Integration Services Project**

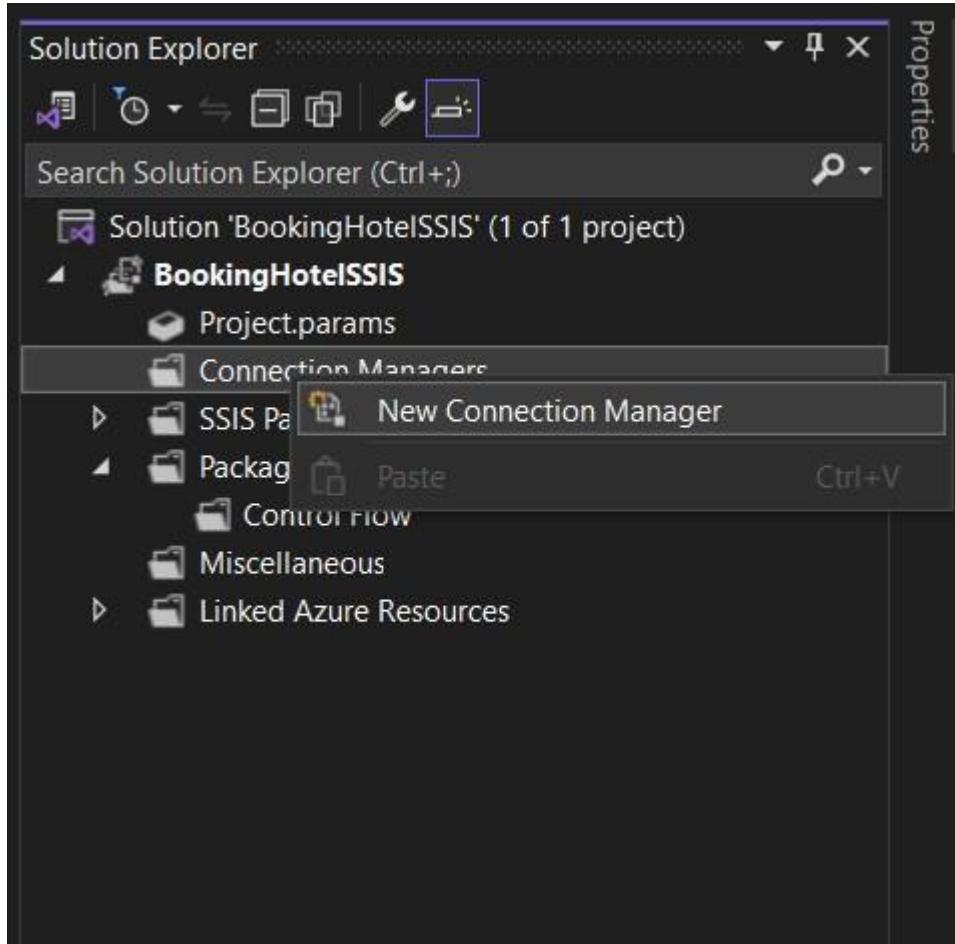
b. Đặt tên: **Hardware Integration Services**

3. Click **OK**

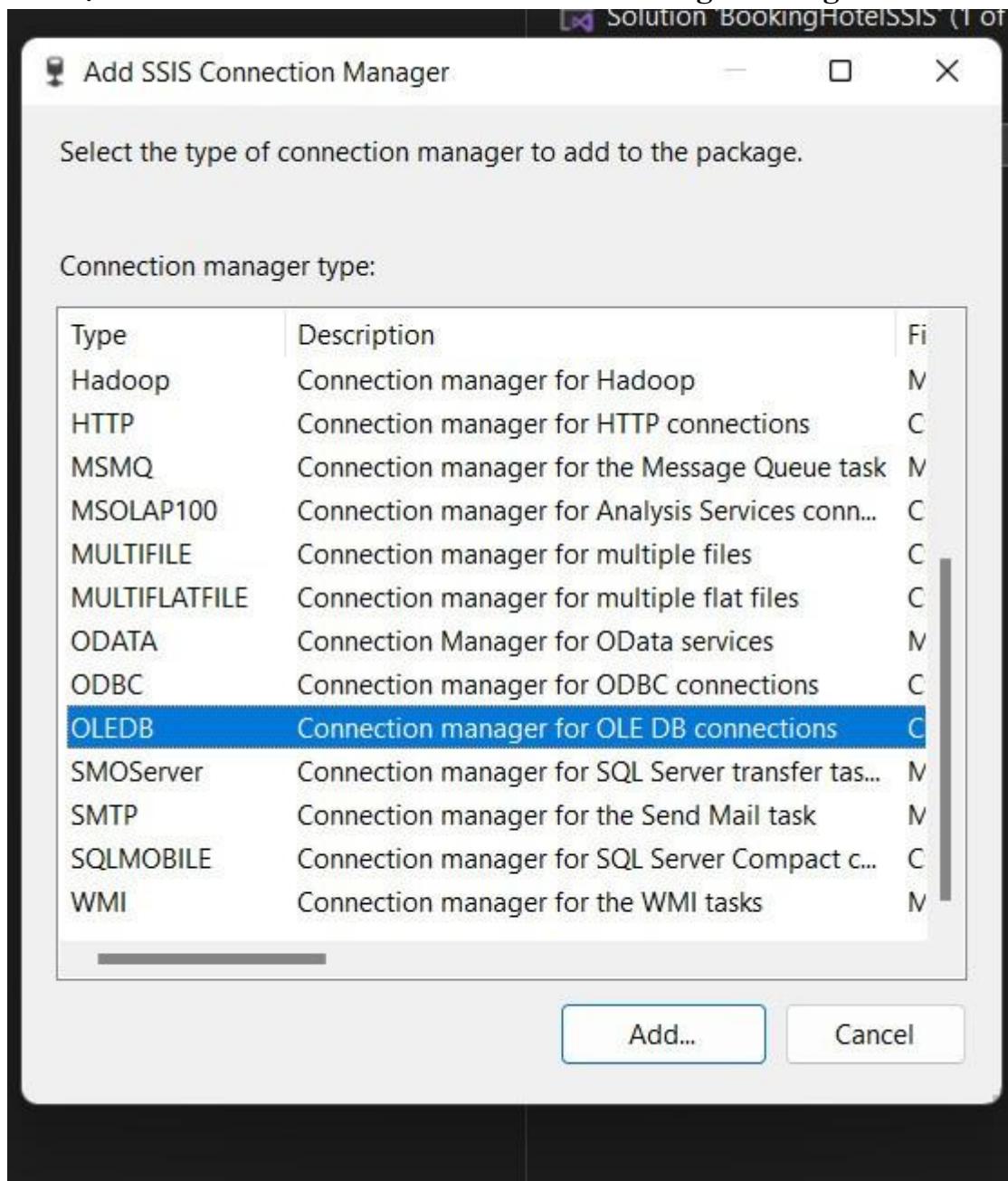


3.2. Thiết lập các Connection

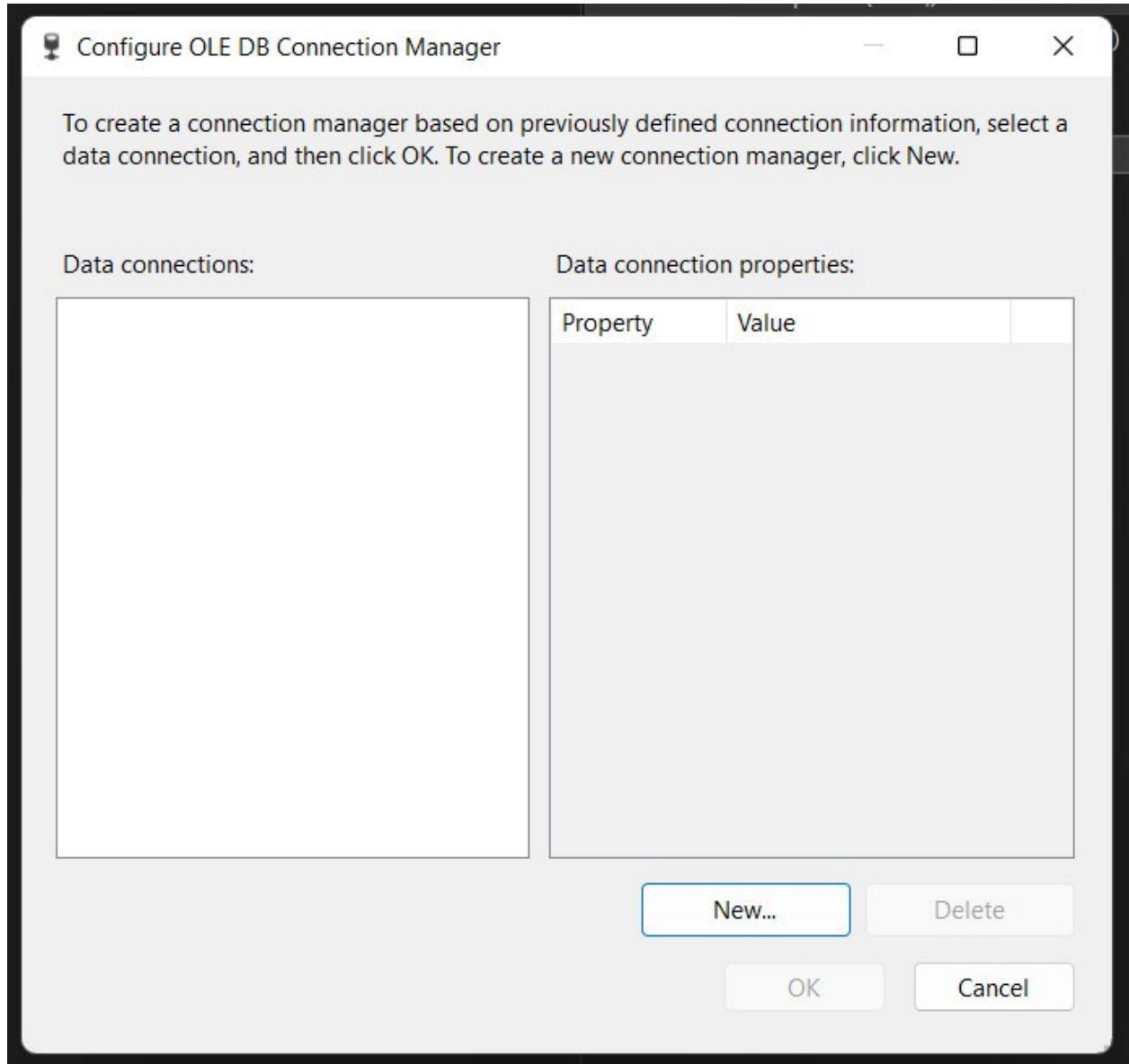
1. Tại **solution explorer**, right-click vào **Connection Managers**, chọn **New Connection Manager**.



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



3. Configure OLE DB Connection Manager dialog xuất hiện, click New...



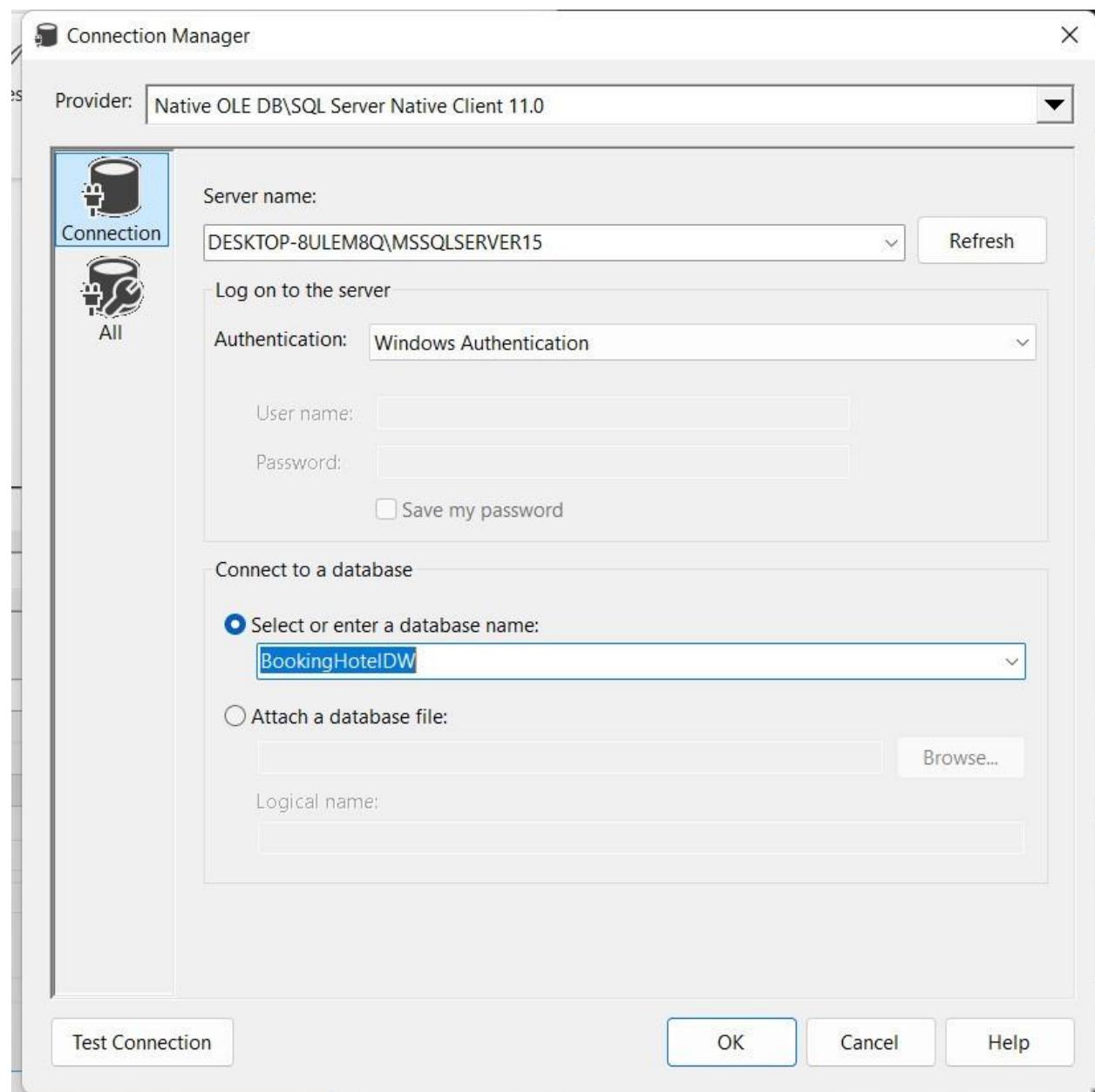
Đặt tên cho các thông tin như bên dưới:

Provider: **SQL Server Native Client**

Server Name: **DESKTOP -8ULEM8Q|MSSQLSERVER15**

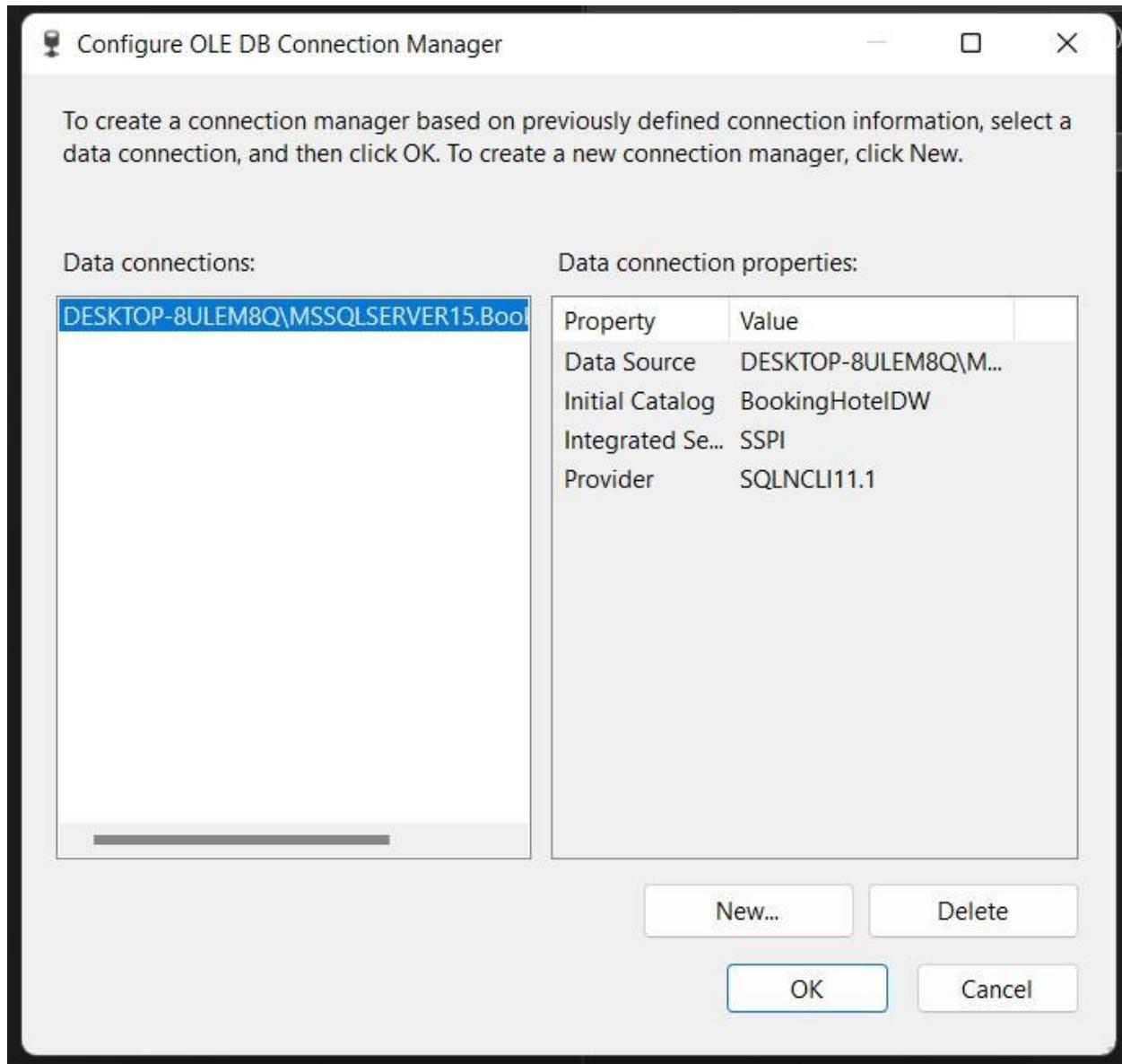
Log on to the Server: **Use Windows Authentication**

Connect to a database: **BookingHotelDW**



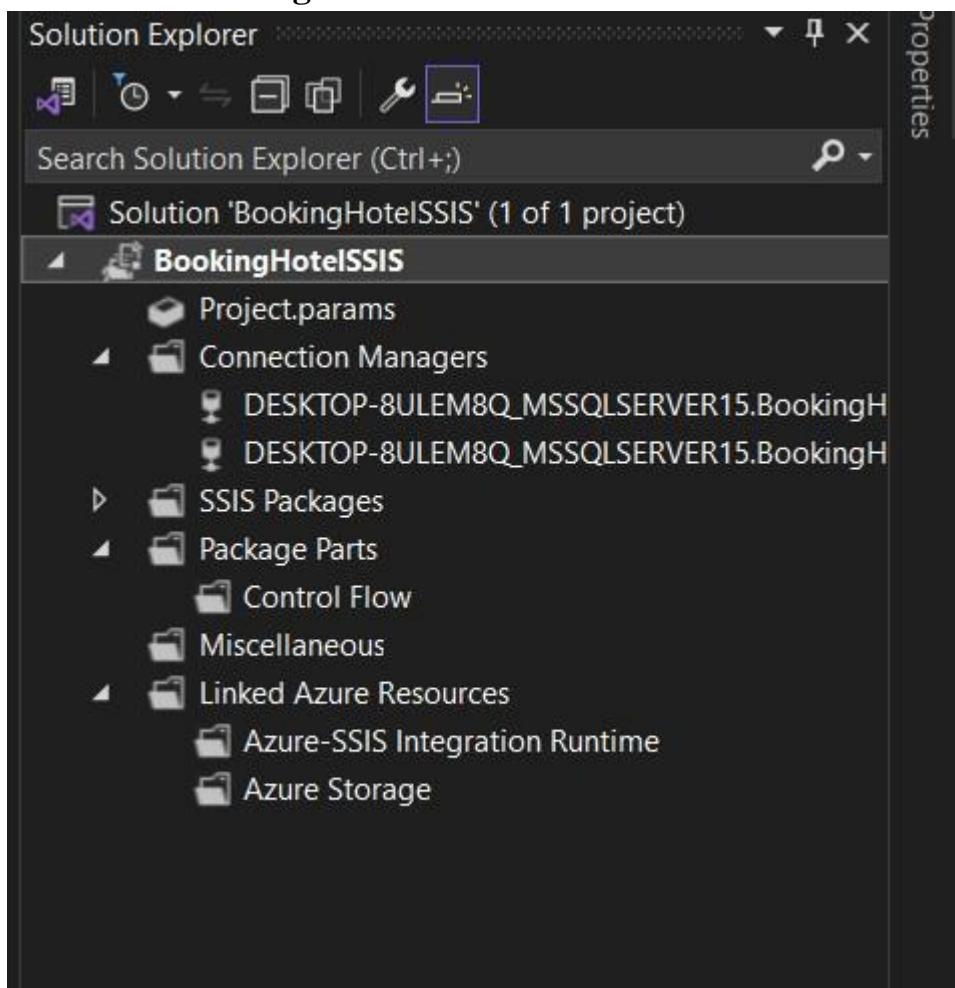
Click **Test Connection** để kiểm tra connection hoạt động, sau đó click **OK** để lưu connection.

4. Trở lại màn hình **Configure OLE DB Connection Manager**



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

6. Lặp lại các bước trên (bước 1->5) cho **BookingHotelStage**. Sau khi tạo xong **Connection Managers** folder sẽ như bên dưới.



3.3. Import Dữ liệu vào các bảng dimension

3.3.1 Booking Dimension

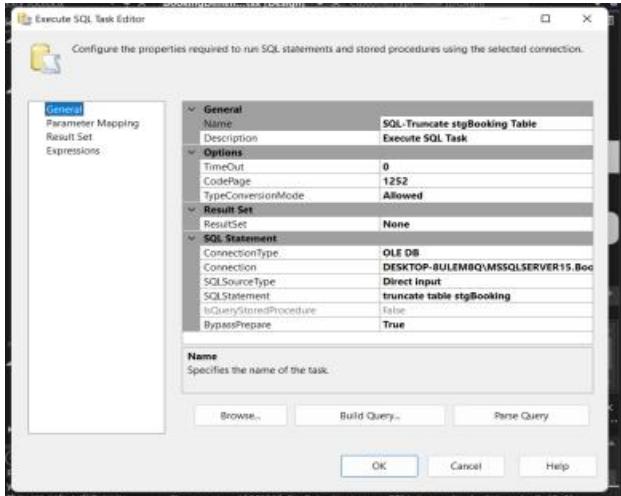
Thiết lập truncate table SQL task

Click **Control Flow** tab:

Double-click vào **SQL - Truncate stgBooking Table** task để cấu hình nó, **Execute SQL**

Task Editor sẽ mở ra.

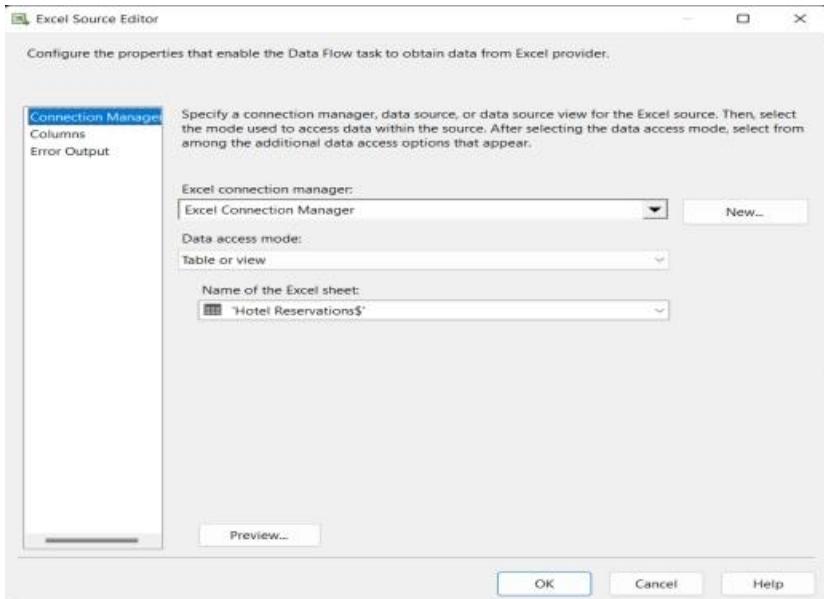
Ở mục **SQL Statement**, chọn **Connection** là **DESKTOP-8ULEM8Q|MSSQLSERVER15.BookingHotelStage** và **SQLStatement** là **truncate table stgBooking;**



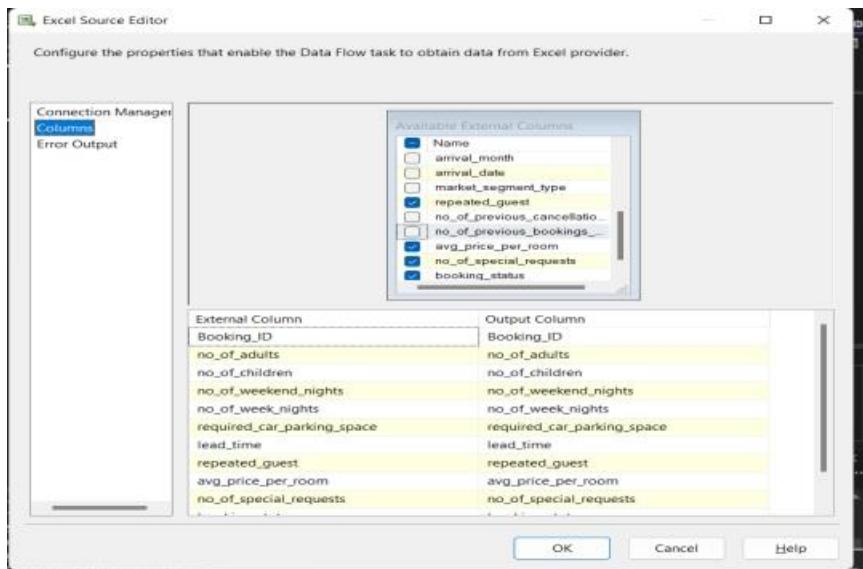
3.3.1.1 Extract

Double-click vào **DF-Extract From Source to Stage stgBooking** task để mở data flow design Surface.

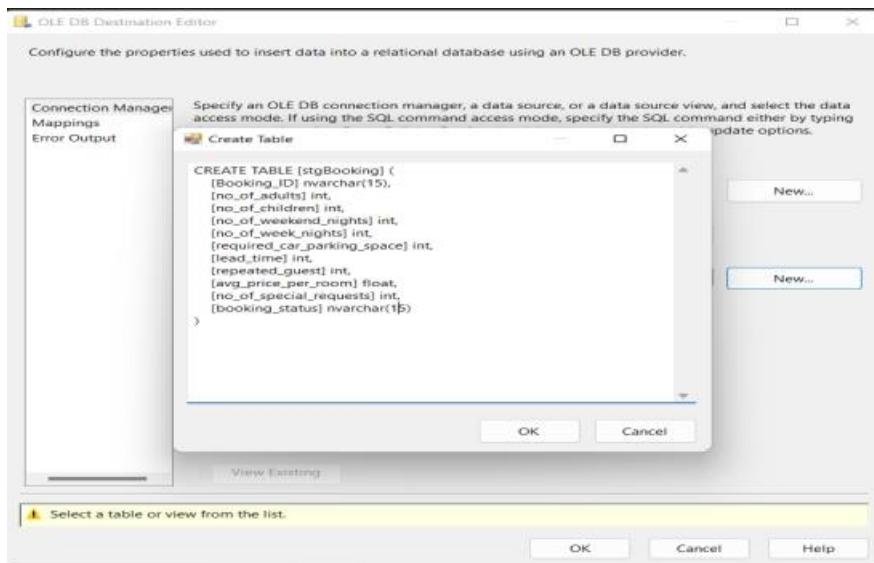
Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, chọn source Type: **Excel**



Chọn các column cần thiết cho bảng stgBooking

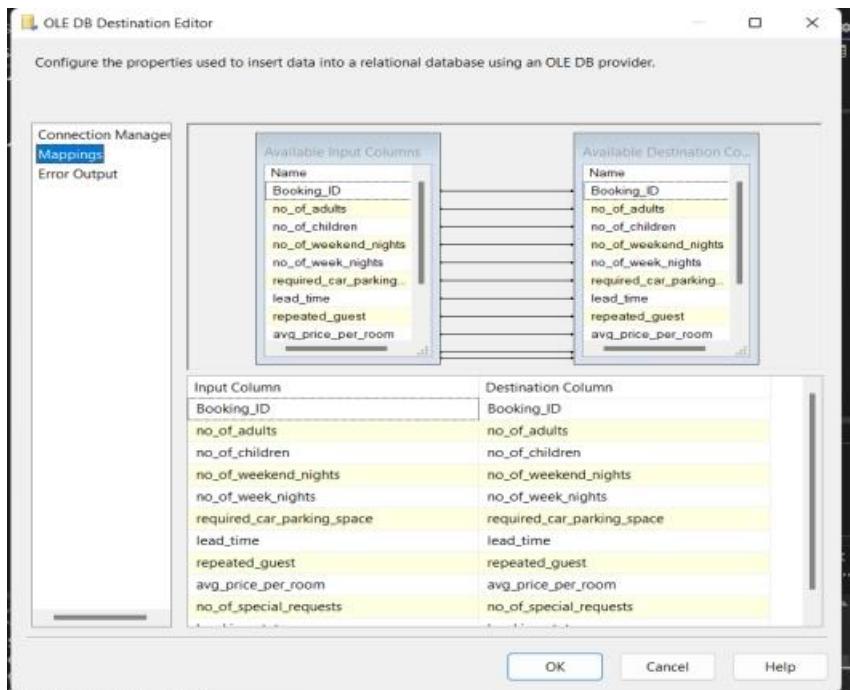


Double click vào destination để cấu hình nó, **OLE DB Destination** sẽ xuất hiện.
Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table**
Đổi tên bảng lại thành **stgBooking**. (**CREATE TABLE [stgBooking]**).
Click OK.



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

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



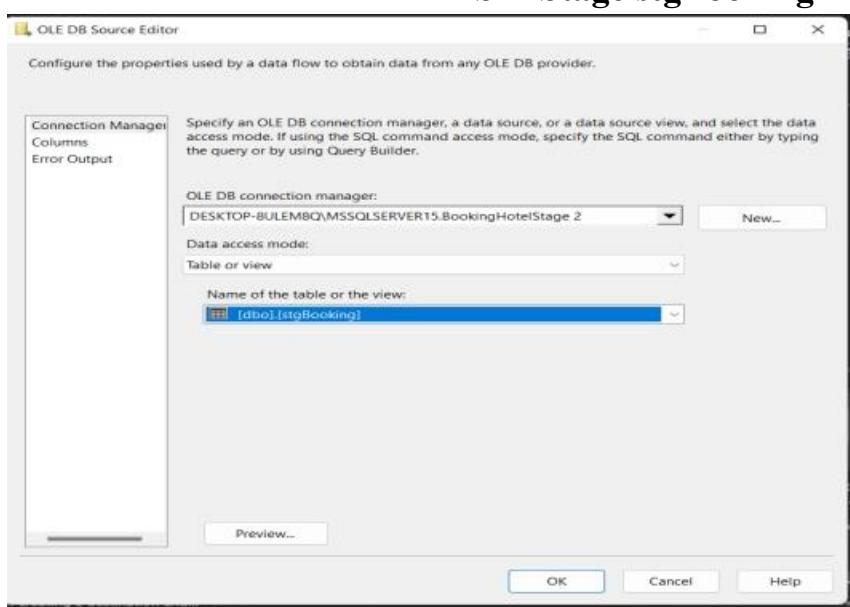
Click **OK** để hoàn tất việc cấu hình.

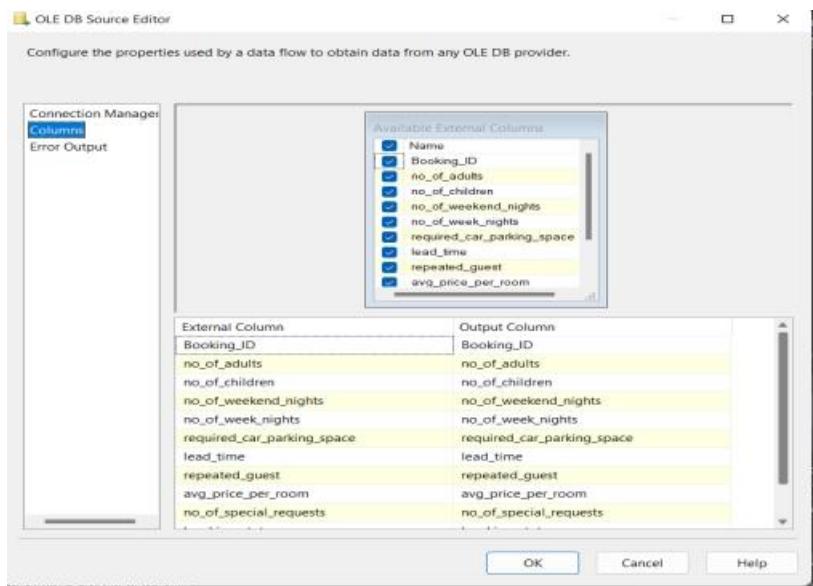
3.3.1.2 Load

Double-click vào **DF - Load from stgBooking to DimBooking** task để mở data flow design Surface.

Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, Chọn source Type: **SQL Server** và **BookingHotelStage** connection manager.

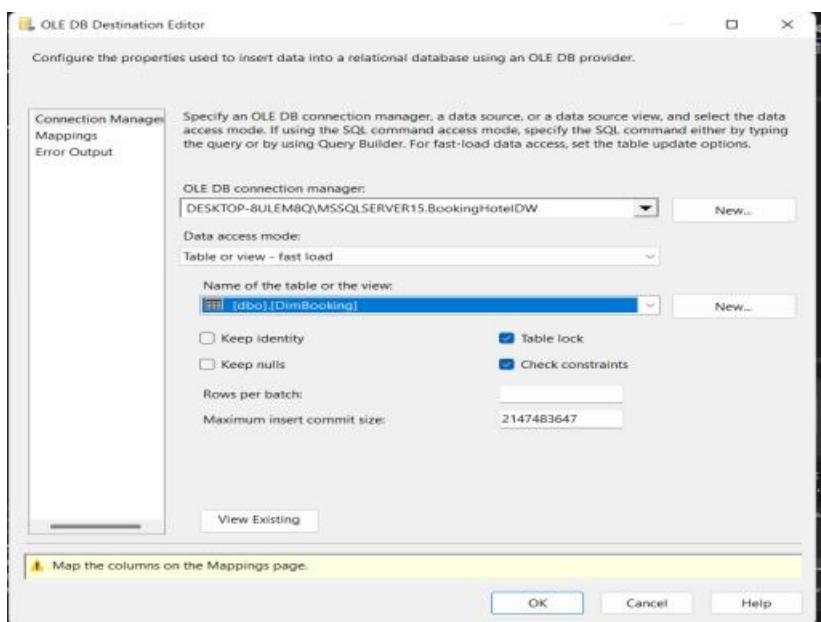
Cấu hình và đổi tên nó thành **DST–Stage stgBooking**

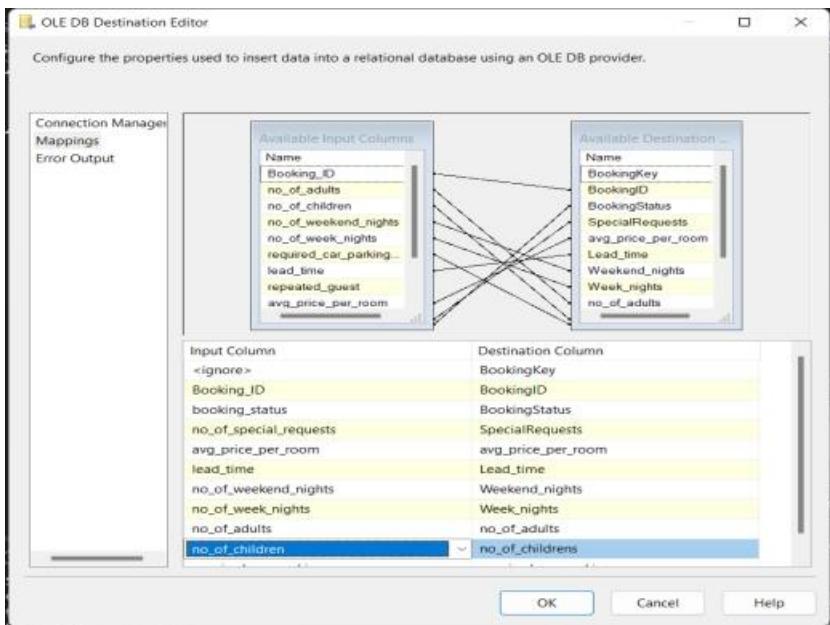




Ta có một **destination** như bên dưới: Đặt tên nó lại thành **SRC-Stage stgBooking** và connect **SRC-Stage stgBooking** -> **OLE DB Destination**

Double click vào **OLE DB Destination** để cấu hình.

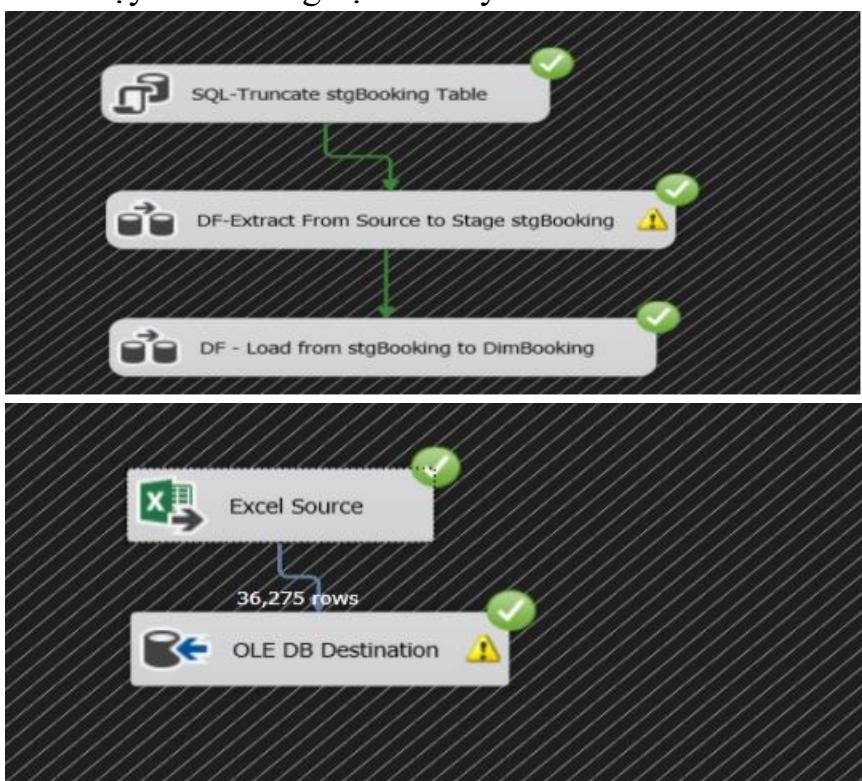


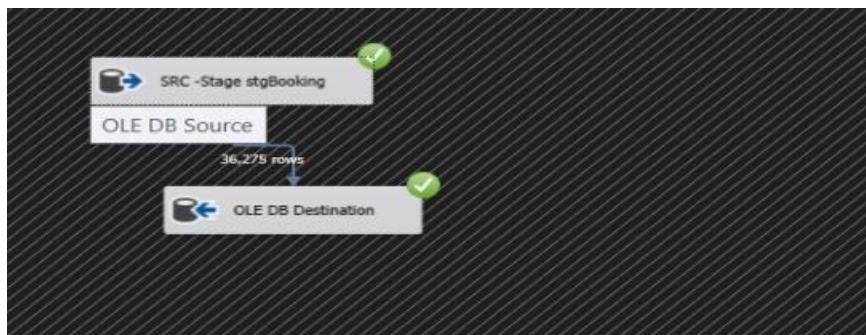


Click vào OK để hoàn tất cấu hình.

3.3.1.3 Run

Bấm Start để chạy. Nếu xảy ra lỗi, có găng đeo hiệu thông báo lỗi để tìm cách sửa. Nếu chạy thành công bạn sẽ thấy màn hình như sau:





Kiểm tra kết quả trong bảng stgBooking trong cơ sở dữ liệu BookingHotelStage.

```
SQLQuery1.sp - DESKTOP-BULEMBQ\YenNhi (75) - Microsoft SQL Server Management Studio
File Edit View Query Project Tools Window Help
BookingHotelStage - Execute New Query Find Replace Home Back Forward Stop Refresh
Object Explorer
SQLQuery1.sp - [BookingID]
SELECT TOP (1000) [Booking_ID]
,[no_of_adults]
,[no_of_children]
,[no_of_weekend_nights]
,[no_of_week_nights]
,[required_car_parking_space]
,[lead_time]
,[repeated_guest]
,[avg_price_per_room]
,[no_of_special_requests]
,[booking_status]
FROM [BookingHotelStage].[dbo].[stgbooking]
```

	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	required_car_parking_space	lead_time	repeated_guest	avg_price_per_room	no_of_special_requests	booking_status
1	INN00001	2	0	2	0	224	0	65	106.88	0	Not_Canceled
2	INN00002	2	0	2	3	0	5	0	106.88	1	Not_Canceled
3	INN00003	1	0	2	1	0	1	0	80	0	Canceled
4	INN00004	2	0	0	2	0	211	0	100	0	Canceled
5	INN00005	2	0	1	1	0	48	0	94.5	0	Canceled
6	INN00006	2	0	0	2	0	348	0	115	1	Canceled
7	INN00007	2	0	1	3	0	34	0	107.95	1	Not_Canceled
8	INN00008	2	0	1	3	0	89	0	109.11	1	Not_Canceled
9	INN00009	3	0	0	4	0	121	0	98.8	1	Not_Canceled
10	INN00010	2	0	0	5	0	44	0	133.44	3	Not_Canceled
11	INN00011	1	0	1	0	0	0	0	85.03	0	Not_Canceled
12	INN00012	1	0	2	1	0	35	0	140.4	1	Not_Canceled
13	INN00013	2	0	2	1	0	30	0	88	0	Canceled
14	INN00014	1	0	2	0	0	96	0	90	2	Canceled
15	INN00015	2	0	0	2	0	47	0	94.5	2	Not_Canceled
16	INN00016	2	0	0	2	0	256	0	115	1	Canceled
17	INN00017	1	0	0	1	0	n	n	n	n	Not_Processed

Kiểm tra kết quả trong bảng DimBooking trong cơ sở dữ liệu BookingHotelDW.

```
SQLQuery1.sp - DESKTOP-BULEMBQ\YenNhi (76) - Microsoft SQL Server Management Studio
File Edit View Query Project Tools Window Help
BookingHotelDW - Execute New Query Find Replace Home Back Forward Stop Refresh
Object Explorer
SQLQuery1.sp - [BookingKey]
SELECT TOP (1000) [BookingKey]
,[BookingID]
,[BookingStatus]
,[SpecialRequests]
,[Avg_Price_per_Room]
,[Lead_Time]
,[Weekend_Nights]
,[Week_Nights]
,[No_of_Adults]
,[No_of_Children]
,[Required_Car_Parking_Space]
FROM [BookingHotelDW].[dbo].[DimBooking]
```

	BookingKey	BookingID	BookingStatus	SpecialRequests	avg_price_per_room	Lead_time	Weekend_nights	Week_nights	no_of_adults	no_of_children	required_car_parking_space
1	1	INN00001	Not_Canceled	0	65	224	1	2	2	0	0
2	2	INN00002	Not_Canceled	1	106.88	5	2	2	0	0	0
3	3	INN00003	Canceled	0	80	1	2	1	0	0	0
4	4	INN00004	Canceled	0	100	211	0	2	2	0	0
5	5	INN00005	Canceled	0	94.5	48	1	1	2	0	0
6	6	INN00006	Canceled	1	115	348	0	2	2	0	0
7	7	INN00007	Not_Canceled	1	107.95	34	1	3	2	0	0
8	8	INN00008	Not_Canceled	1	109.11	89	1	3	2	0	0
9	9	INN00009	Not_Canceled	1	98.8	121	0	4	3	0	0
10	10	INN00010	Not_Canceled	3	133.44	44	0	5	2	0	0
11	11	INN00011	Not_Canceled	0	85.03	0	1	1	0	0	0
12	12	INN00012	Not_Canceled	1	140.4	35	2	1	1	0	0
13	13	INN00013	Canceled	0	88	30	2	1	2	0	0
14	14	INN00014	Cancelled	2	90	95	2	0	1	0	0
15	15	INN00015	Not_Canceled	2	94.5	47	0	2	2	0	0
16	16	INN00016	Cancelled	1	115	256	0	3	2	0	0
17	17	INN00017	Not_Processed	0	n	n	n	n	n	n	n

3.3.2 Date Dimension

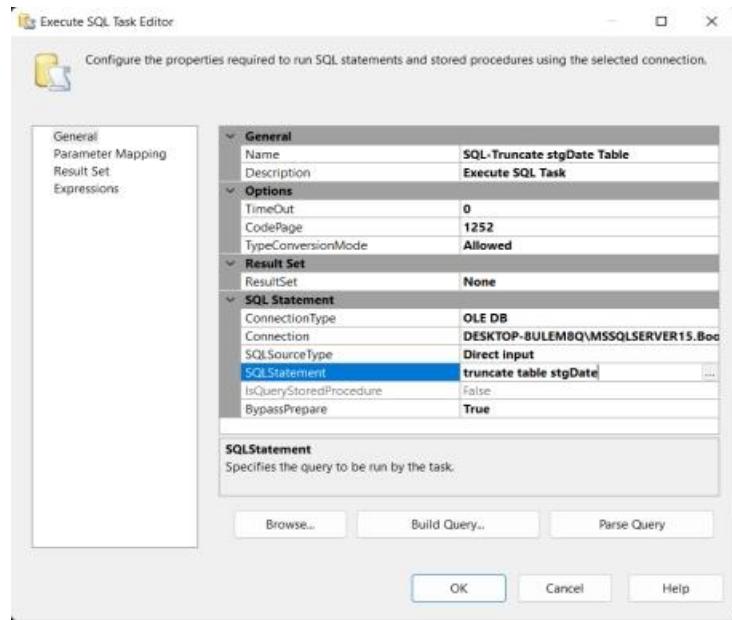
Thiết lập truncate table SQL task

Click **Control Flow** tab:

Double-click vào **SQL - Truncate stgDate Table** task để cấu hình nó, **Execute SQL**

Task Editor sẽ mở ra.

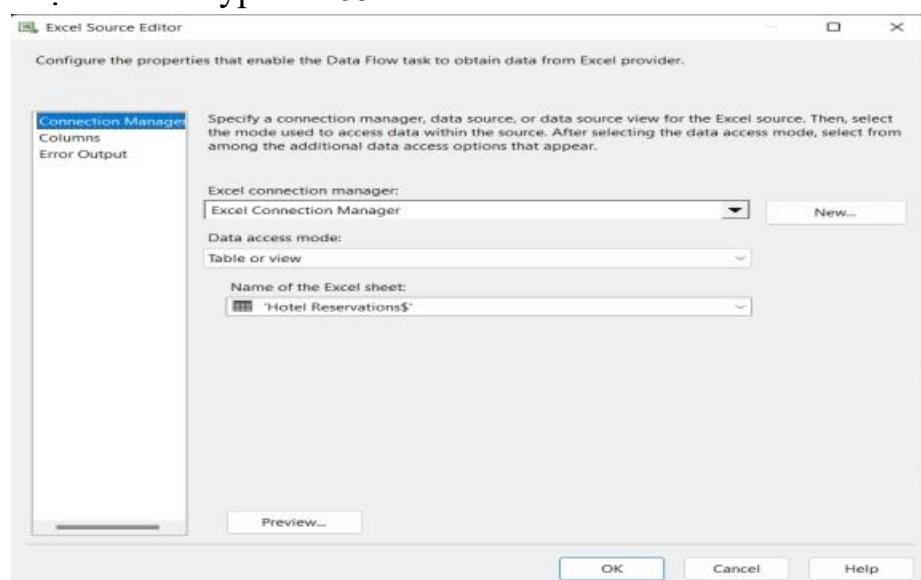
Ở mục **SQL Statement**, chọn **Connection** là **DESKTOP-BULEM8Q\MSSQLSERVER15.BookingHotelStage** và **SQLStatement** là **truncate table stgDate;**



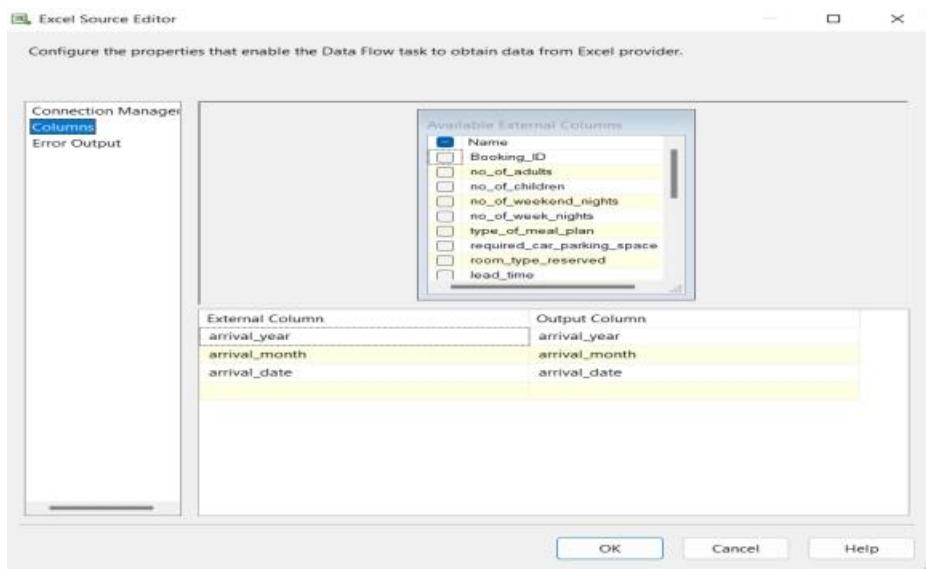
3.3.2.1 Extract

Double-click vào **DF-Extract From Source to Stage stgDate** task để mở data flow design Surface.

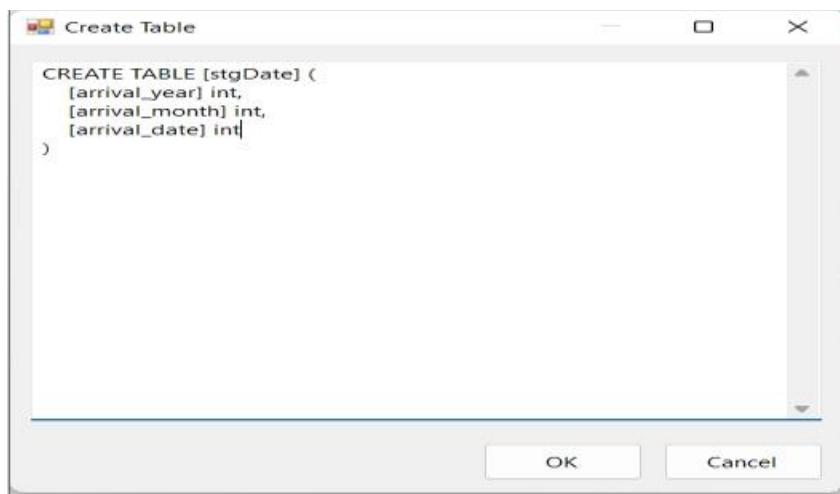
Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, chọn source Type: **Excel**



Chọn các column cần thiết cho bảng stgDate

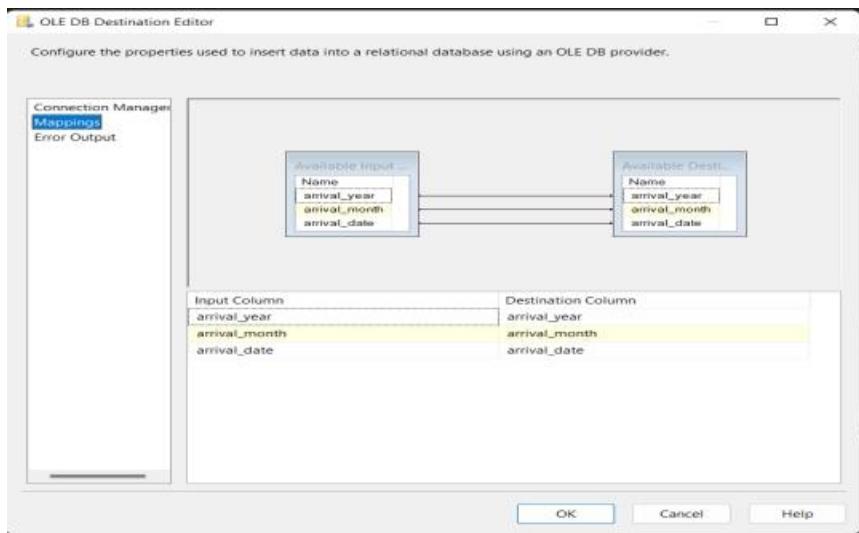


Double click vào destination để cấu hình nó, **OLE DB Destination** sẽ xuất hiện.
Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table**
Đổi tên bảng lại thành **stgDate**. (**CREATE TABLE [stgDate]**).
Click OK.



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

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



Click **Ok** để hoàn tất cấu hình.

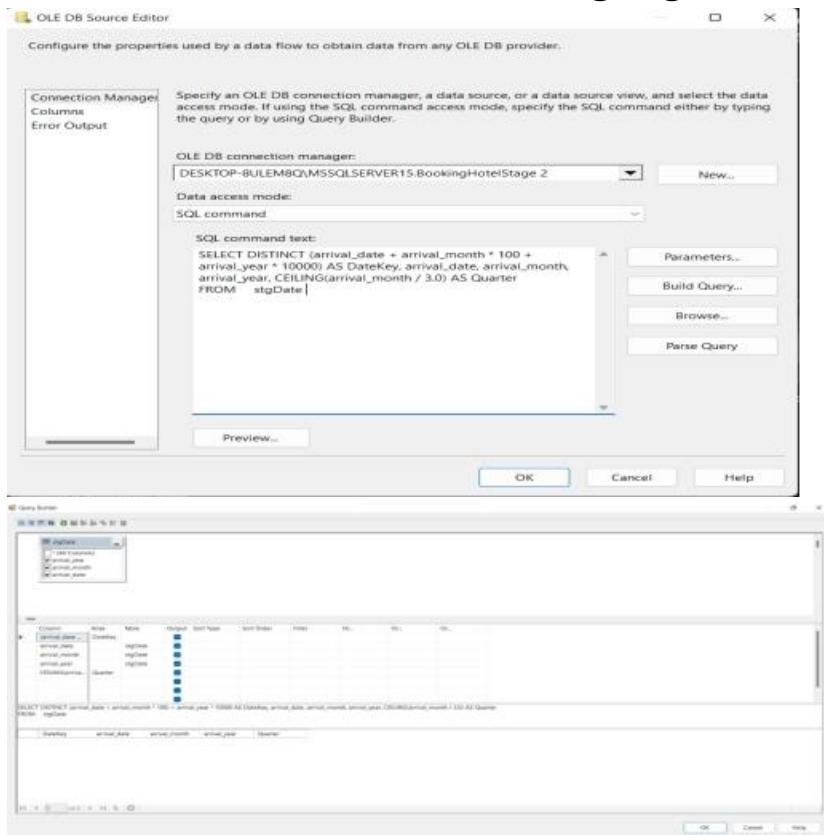
3.3.2.2 Load

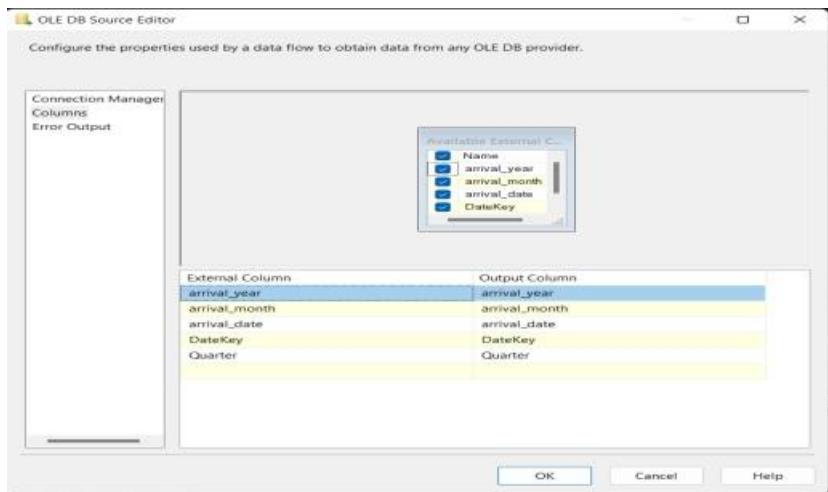
Double-click vào **DF - Load from stgDate to DimDate** task để mở data flow design Surface.

Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra,

Chọn source Type: **SQL Server** và **BookingHotelStage** connection manager.

Cấu hình và đổi tên nó thành **DST–Stage stgDate**

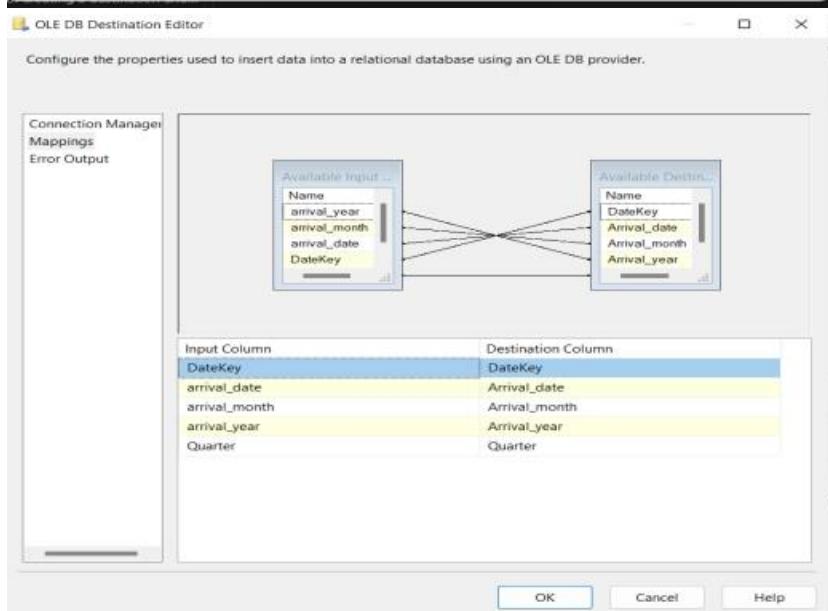
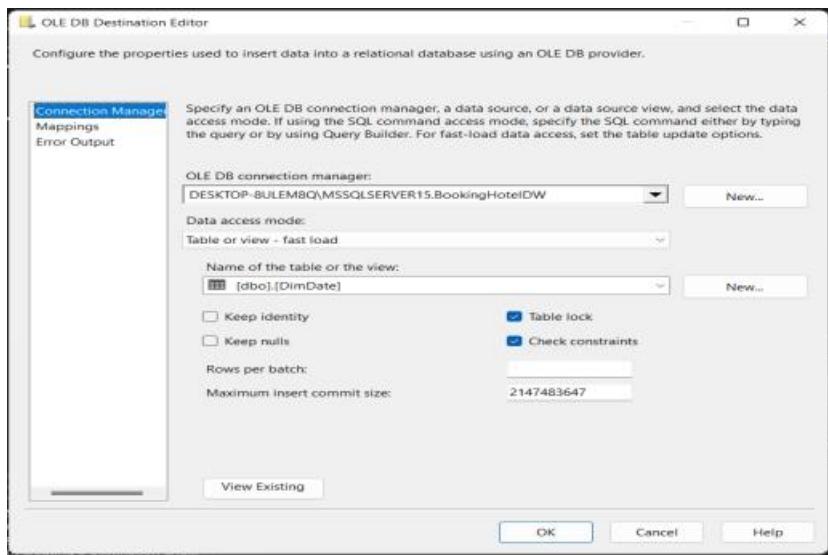




Ta có một **destination** như bên dưới: Đặt tên nó lại thành **SRC-Stage stgDate** và

connect **SRC-Stage stgDate -> OLE DB Destination**

Double click vào **OLE DB Destination** để cấu hình.



Click vào OK để hoàn tất cấu hình.

3.3.2.3 Run

Bấm Start để chạy. Nếu xảy ra lỗi, có găng đeo hiệu thông báo lỗi để tìm cách sửa. Nếu chạy thành công bạn sẽ thấy màn hình như sau:



Kiểm tra kết quả trong bảng stgDate trong cơ sở dữ liệu BookingHotelStage.

```

SELECT TOP (1000) [arrival_year]
      ,[arrival_month]
      ,[arrival_date]
   FROM [BookingHotelStage].[dbo].[stgDate]
  
```

Kiểm tra kết quả trong bảng DimDate trong cơ sở dữ liệu BookingHotelDW.

```

SELECT TOP (1000) [DimKey]
      ,[Arrival_date]
      ,[Arrival_month]
      ,[Arrival_year]
      ,[Quarter]
   FROM [BookingHotelDW].[dbo].[DimDate]
  
```

3.3.3 Customer Type Dimension

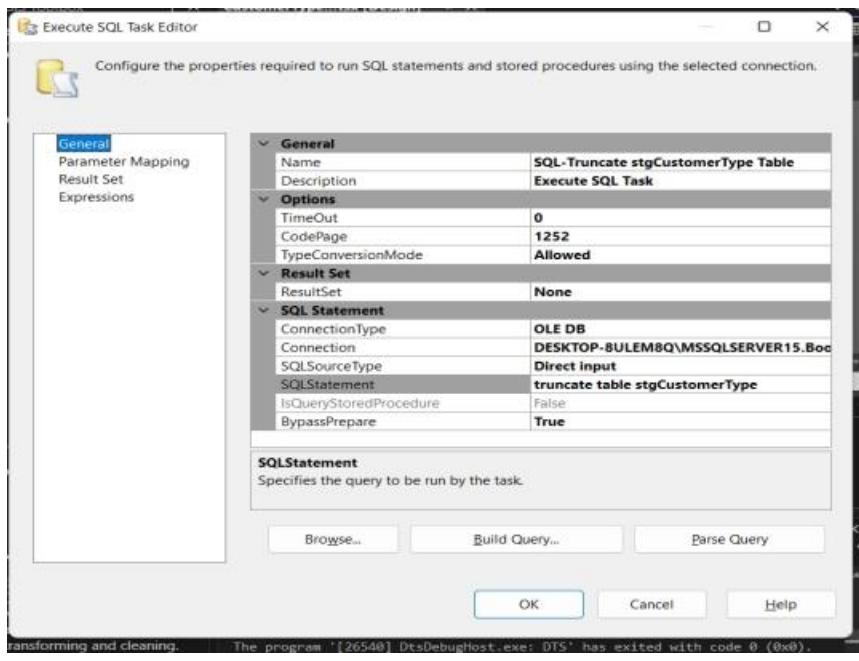
Thiết lập truncate table SQL task

Click **Control Flow** tab:

Double-click vào **SQL - Truncate stgCustomerType Table** task để cấu hình nó, **Execute SQL**

Task Editor sẽ mở ra.

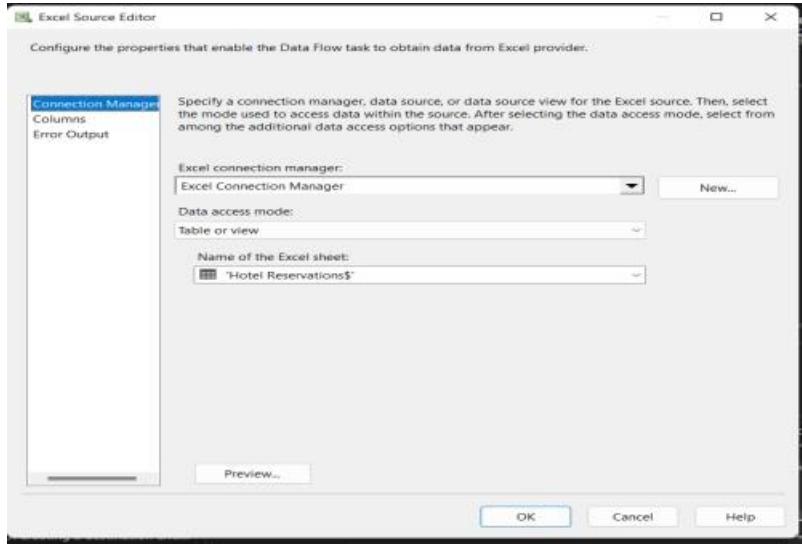
Ở mục **SQL Statement**, chọn **Connection** là **DESKTOP-8ULEM8Q\SQLSERVER15.BookingHotelStage** và **SQLStatement** là **truncate table stgCustomerType;**



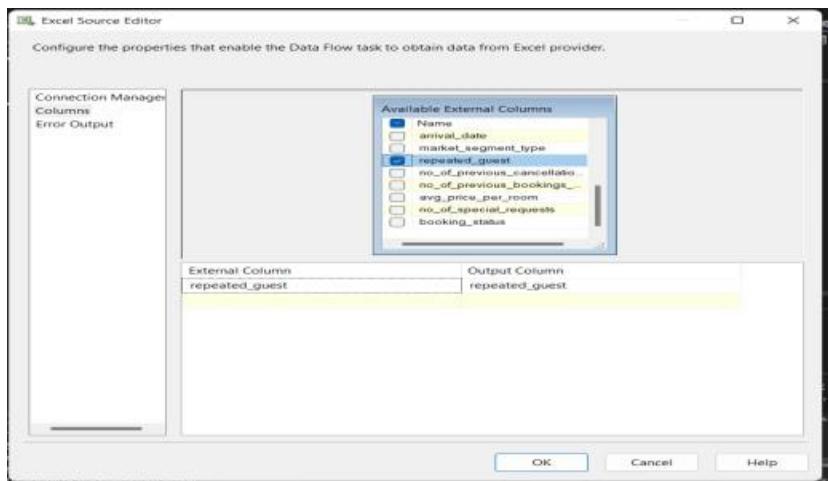
3.3.3.1 Extract

Double-click vào **DF-Extract From Source to Stage stgCustomerType** task để mở data flow design Surface.

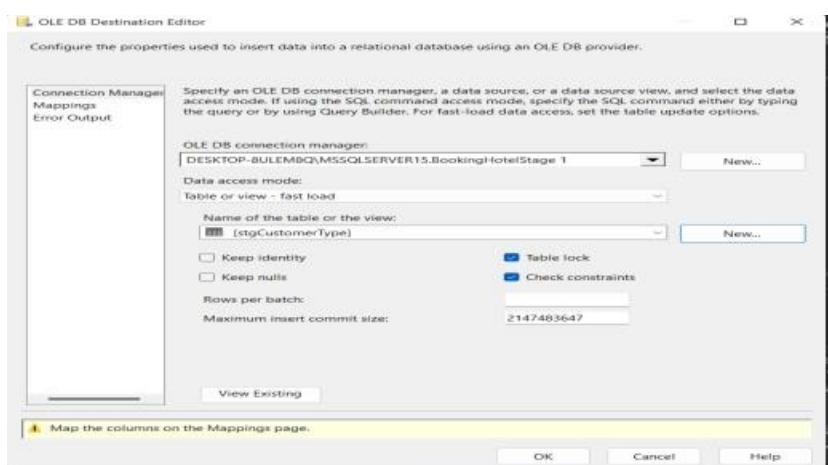
Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, chọn source Type: **Excel**



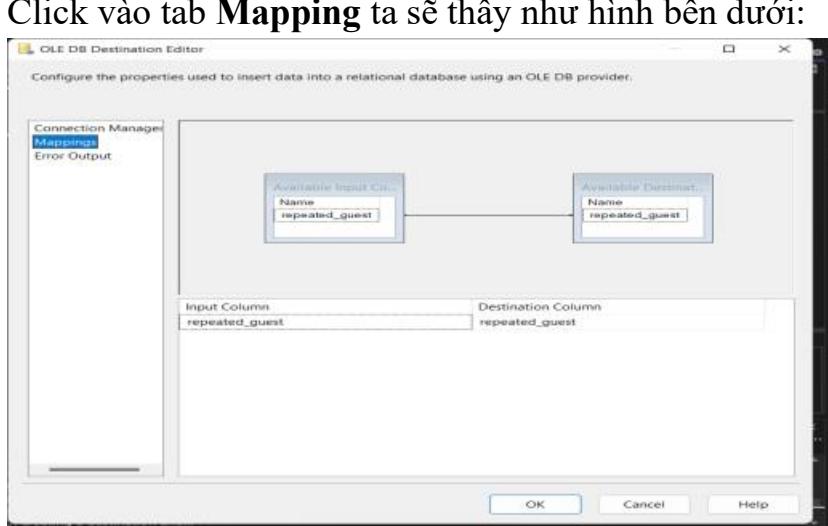
Chọn các column cần thiết cho bảng stgCustomerType



Double click vào destination để cấu hình nó, **OLE DB Destination** sẽ xuất hiện.
Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table**
Đổi tên bảng lại thành **stgCustomerType**(**CREATE TABLE [stgCustomerType]**)
Click OK.



Trở lại màn hình **OLEDB Destination Editor**, tại mục **Name of the table or view** bạn
sẽ thấy **[stgCustomerType]**.
Click vào tab **Mapping** ta sẽ thấy như hình bên dưới:



Click **OK** để hoàn tất việc cấu hình.

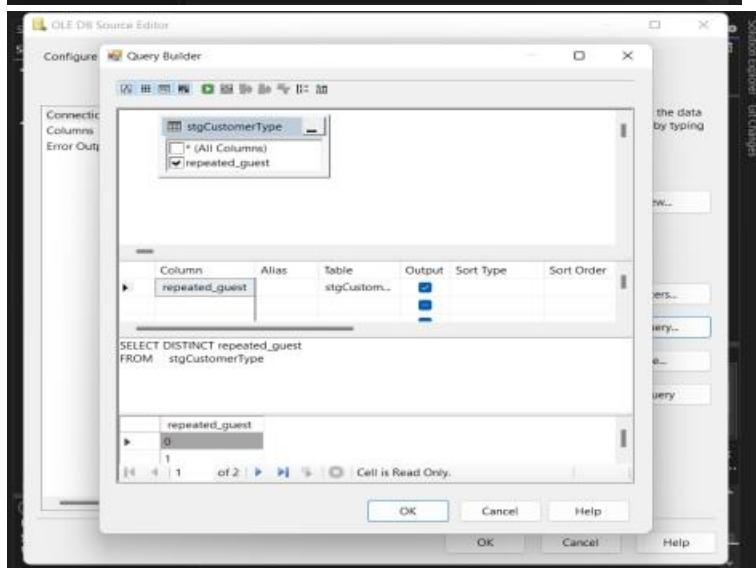
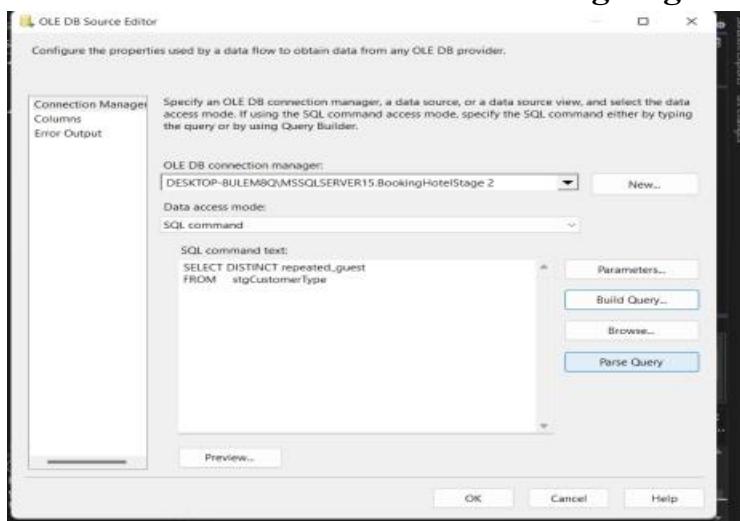
3.3.3.2 Load

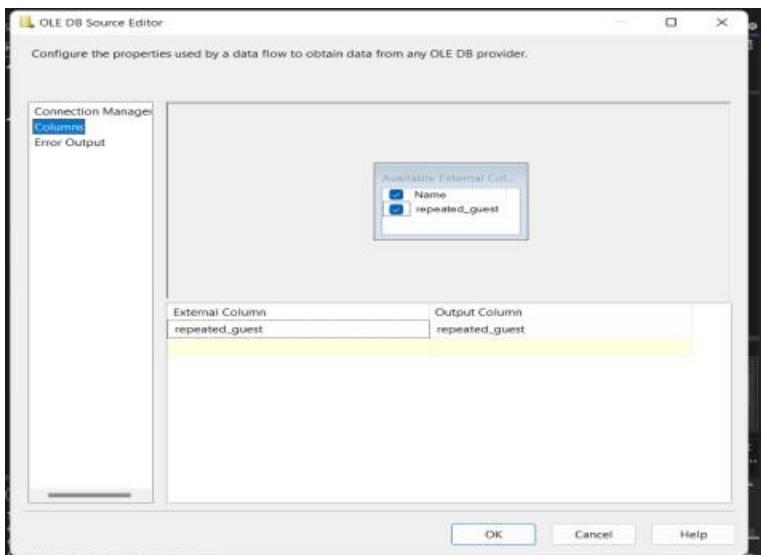
Double-click vào **DF - Load from stgCustomerType to DimCustomerType** task để mở data flow design Surface.

Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra,

Chọn source Type: **SQL Server** và **BookingHotelStage** connection manager.

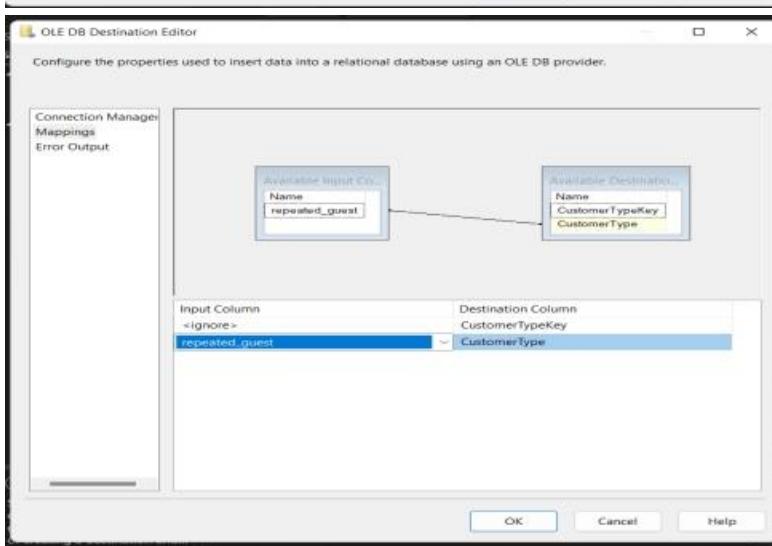
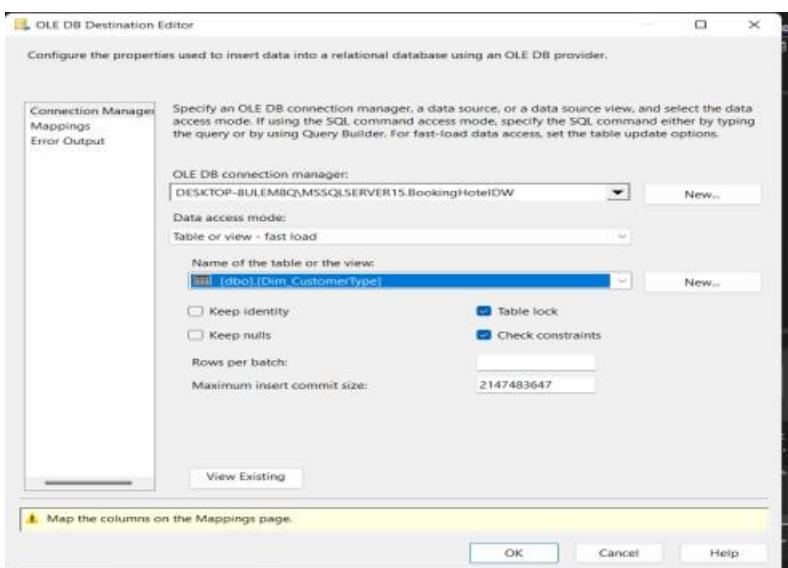
Cấu hình và đổi tên nó thành **DST–Stage stgCustomerType**





Ta có một ***destination*** như bên dưới: Đặt tên nó lại thành **SRC-Stage stgCustomerType** và connect **SRC-Stage stgCustomerType** -> **OLE DB Destination**

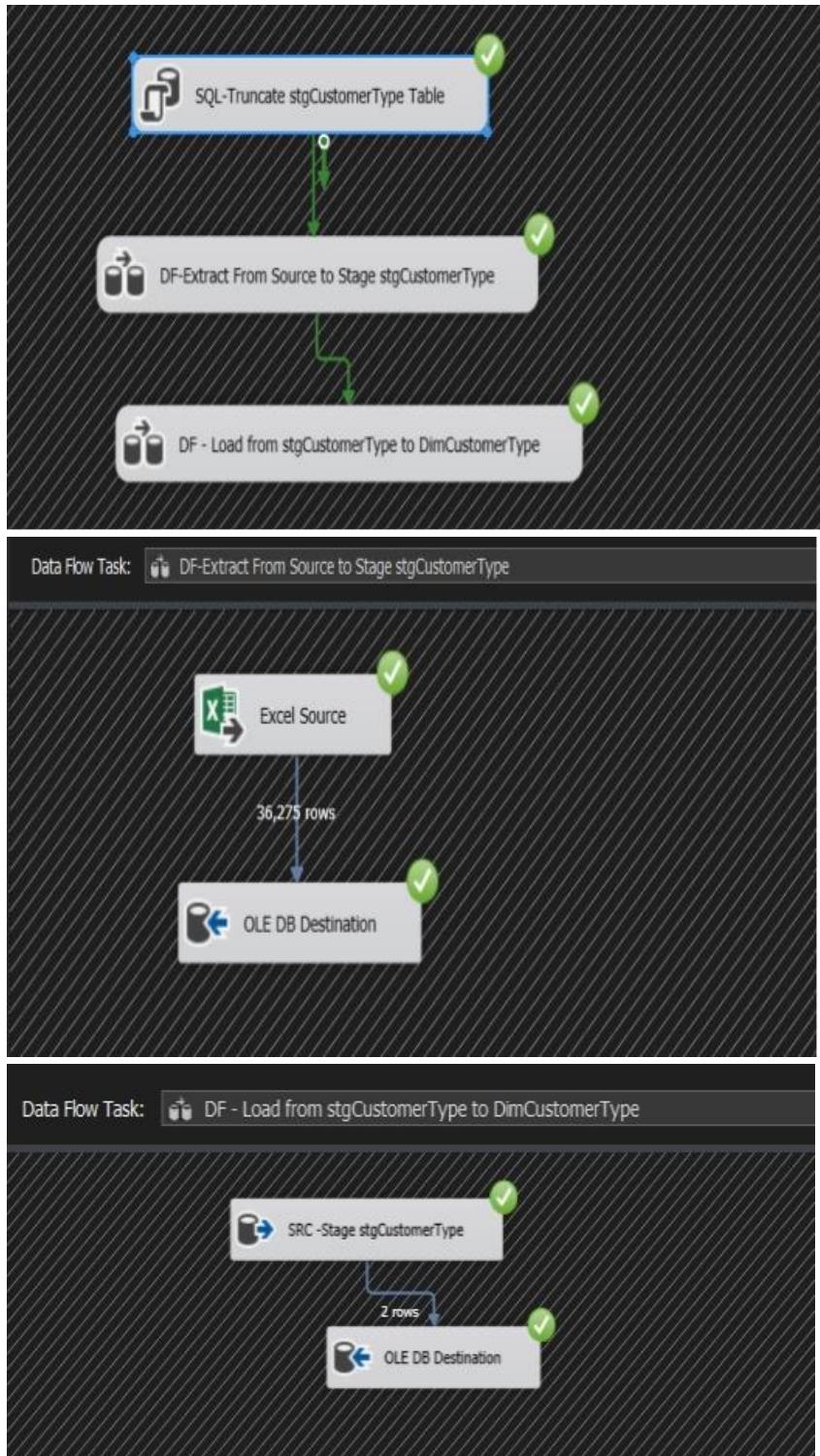
Double click vào **OLE DB Destination** để cấu hình.



Click vào OK để hoàn tất cấu hình.

3.3.3.3 Run

Bấm Start để chạy. Nếu xảy ra lỗi, cố gắng đọc hiểu thông báo lỗi để tìm cách sửa. Nếu chạy thành công bạn sẽ thấy màn hình như sau:



Kiểm tra kết quả trong bảng stgCustomerType trong cơ sở dữ liệu BookingHotelStage.

The screenshot shows a Microsoft SQL Server Management Studio interface with multiple tabs open. The active tab is titled 'SQLQuery6.sql - DESKTOP-BULEMBQ\YenNN (65)'. The query window contains the following SQL code:

```
SELECT TOP (1000) [repeated_guest]
FROM [BookingHotelStage].[dbo].[stgCustomerType]
```

The results pane displays the following data:

repeated_guest
325 0
326 0
327 0
328 0
329 1
330 0
331 0
332 0
333 0
334 1
335 0
336 0
337 0
338 0
339 0
340 0

A status bar at the bottom indicates: 'Query executed successfully.' and 'DESKTOP-BULEMBQ\MSSQLSERVER15\BookingHotelStage 00:00:00 1,000 rows'.

Kiểm tra kết quả trong bảng Dim_CustomerType trong cơ sở dữ liệu BookingHotelDW.

The screenshot shows a Microsoft SQL Server Management Studio interface with multiple tabs open. The active tab is titled 'SQLQuery7.sql - DESKTOP-BULEMBQ\YenNN (55)'. The query window contains the following SQL code:

```
SELECT TOP (1000) [CustomerTypeKey]
,[CustomerType]
FROM [BookingHotelDW].[dbo].[Dim_CustomerType]
```

The results pane displays the following data:

CustomerTypeKey	CustomerType
36276	0
36277	1

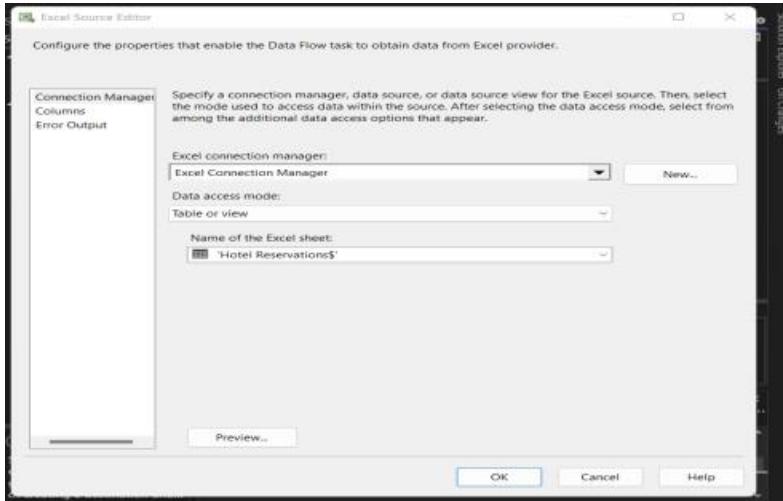
A status bar at the bottom indicates: 'Query executed successfully.' and 'DESKTOP-BULEMBQ\MSSQLSERVER15\BookingHotelDW 00:00:00 2 rows'.

3.3.4 Market Segment Type

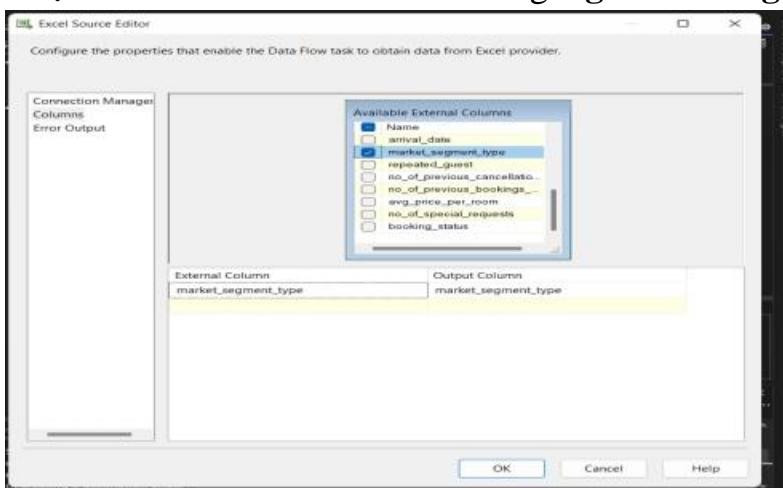
3.3.4.1 Extract

Double-click vào **DF-Extract From Source to Stage stgMarketSegmentType** task để mở data flow design Surface.

Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, chọn source Type: **Excel**

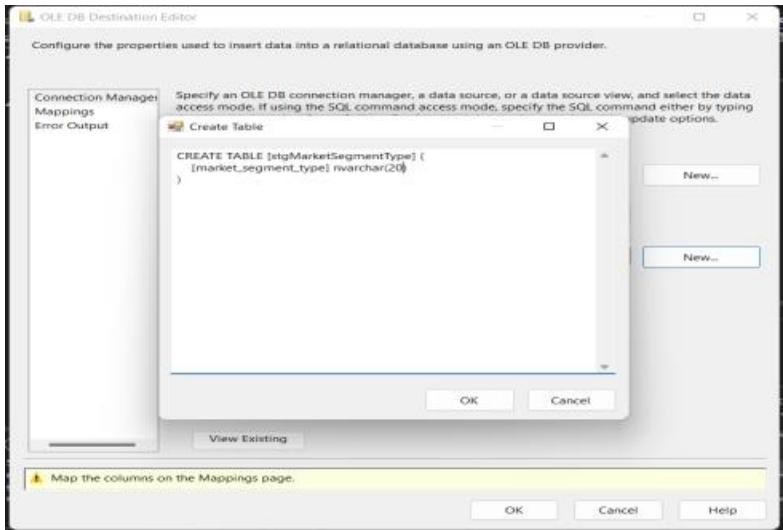


Chọn các column cần thiết cho bảng **stgMarketSegmentType**



Double click vào destination để cấu hình nó, **OLE DB Destination** sẽ xuất hiện.
Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table**
Đổi tên bảng lại thành **stgMarketSegmentType** (**CREATE TABLE [stgMarketSegmentType]**)

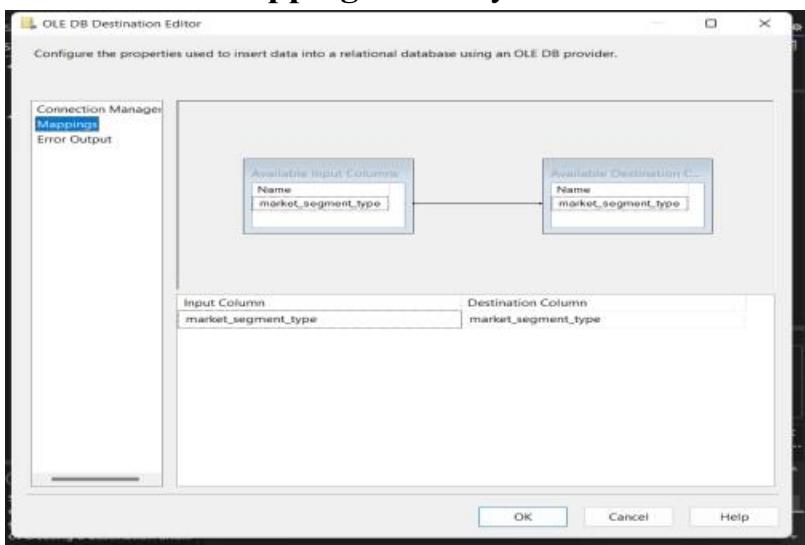
Click OK.



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



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

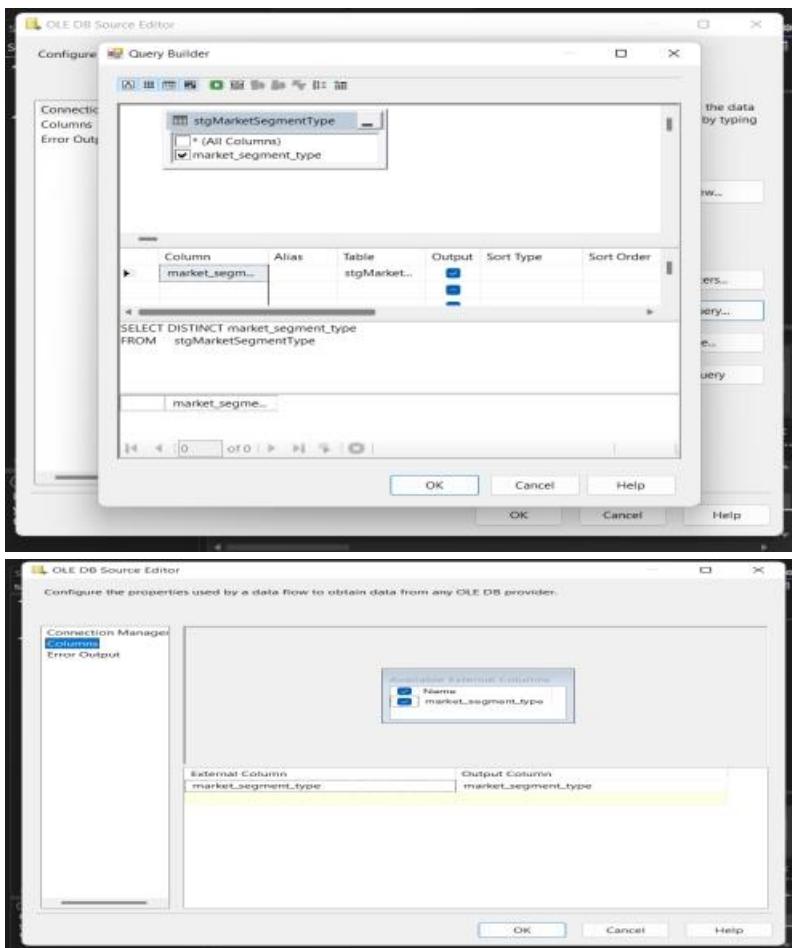


Click **OK** để hoàn tất việc cấu hình.

3.3.4.2 Load

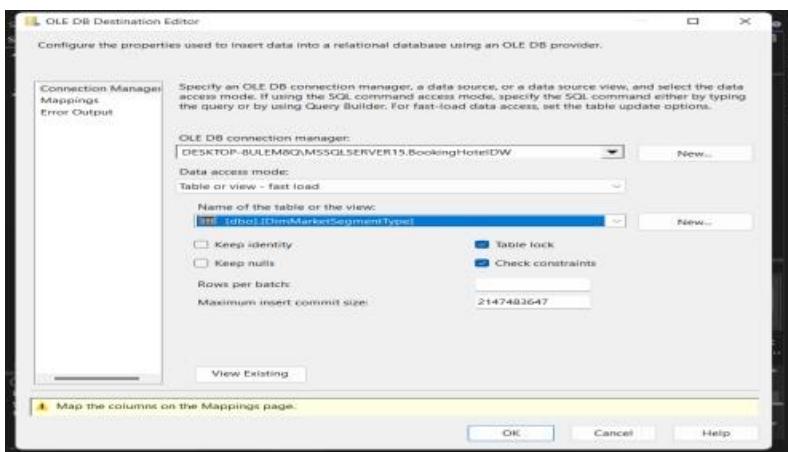
Double-click vào **DF - Load from stgMarketSegmentType to DimMarketSegmentType** task để mở data flow design Surface.

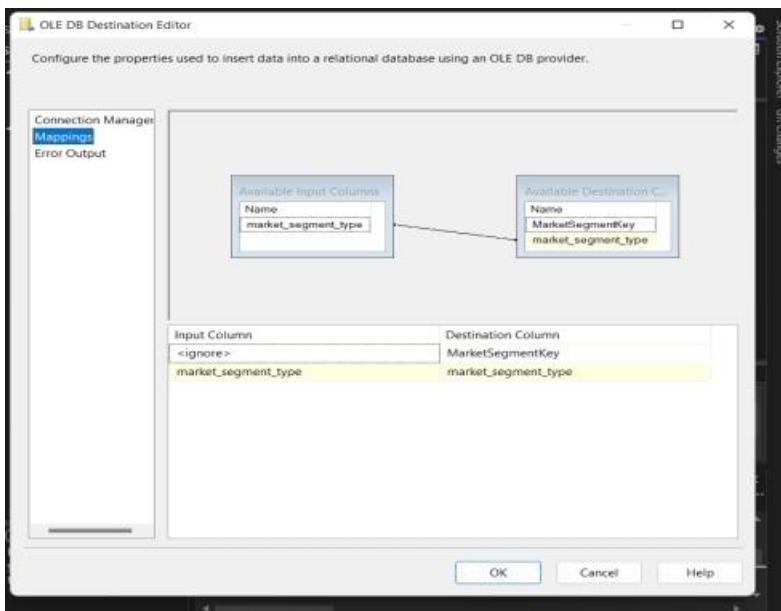
Từ SSIS Toolbox, kéo rê **Source Assistant** vào design surface, một dialog mở ra, Chọn source Type: **SQL Server** và **BookingHotelStage** connection manager. Cấu hình và đổi tên nó thành **DST–Stage stgMarketSegmentType**



Ta có một **destination** như bên dưới: Đặt tên nó lại thành **SRC-Stage stgMarketSegmentType** và connect **SRC-Stage stgMarketSegmentType-> OLE DB Destination**

Double click vào **OLE DB Destination** để cấu hình.

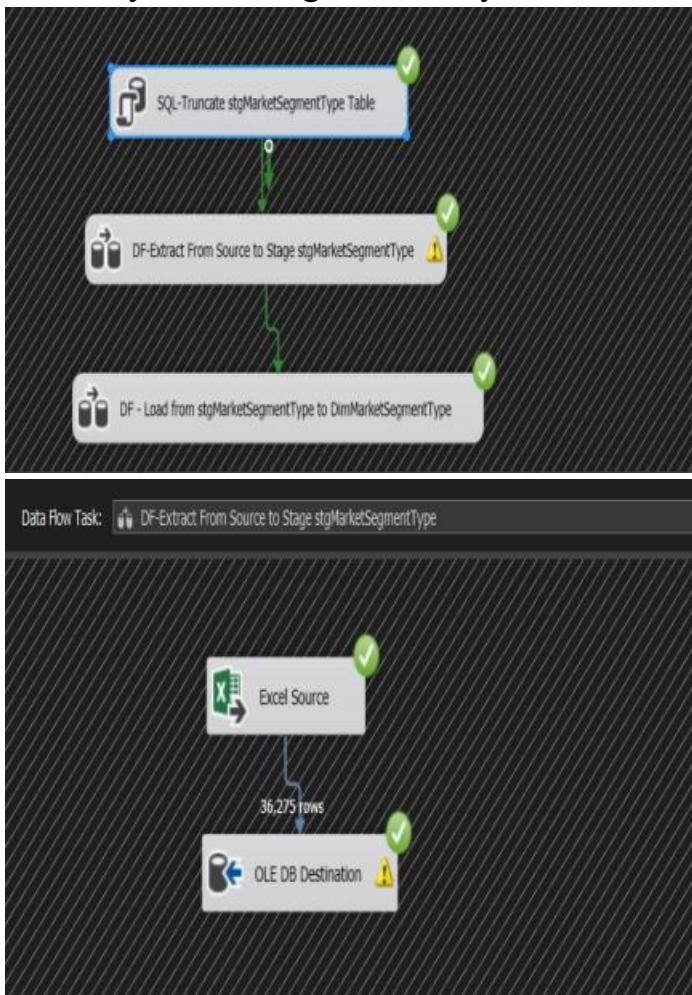


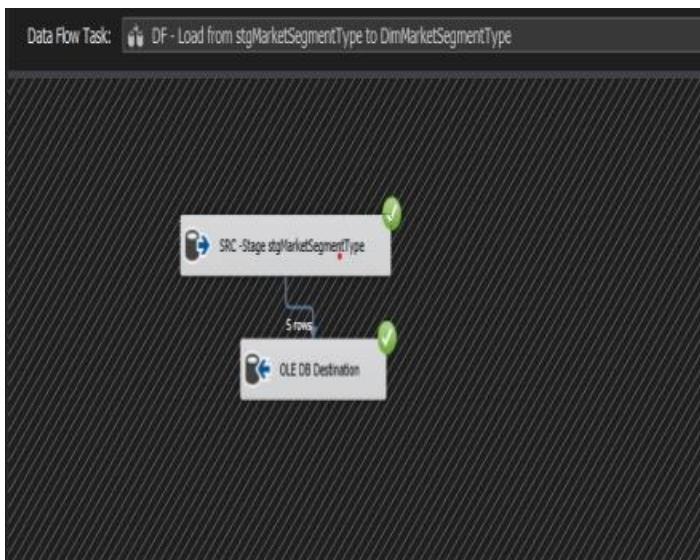


Click vào OK để hoàn tất cấu hình.

3.3.4.3 Run

Bấm Start để chạy. Nếu xảy ra lỗi, có gắng đọc hiểu thông báo lỗi để tìm cách sửa. Nếu chạy thành công bạn sẽ thấy màn hình như sau:





Kiểm tra kết quả trong bảng **stgMarketSegmentType** trong cơ sở dữ liệu BookingHotelStage.

```
SELECT market_segment_type
FROM [BookingHotelStage].[dbo].[MarketSegmentType]
```

Kiểm tra kết quả trong bảng **DimMarketSegmentType** trong cơ sở dữ liệu BookingHotelDW.

The screenshot shows the Microsoft SQL Server Management Studio interface. At the top, there are five tabs labeled: 'SQLQuery1.sql - DESKTOP-BULEM\Q\SQLSERVER15.BookingHotelDW (DESKTOP-BULEM\Q\YenNh1 (74)) - Microsoft SQL Server Management Studio', 'File', 'Edit', 'View', 'Query', 'Project', 'Tools', 'Window', 'Help', 'Quick Launch (Ctrl+Q)', and a maximize/minimize button. Below the tabs, there are several toolbar icons. The main area contains five query windows. The first window (selected) displays the following SQL code:

```
SELECT TOP (1000) [MarketSegmentKey]
      ,[market_segment_type]
  FROM [BookingHotelDW].[dbo].[DimMarketSegmentType]
```

The second window shows the results of the query:

MarketSegmentKey	market_segment_type
1	Corporate
2	Complementary
3	Office
4	Aviation
5	

3.3.5 Meal Type Dimension

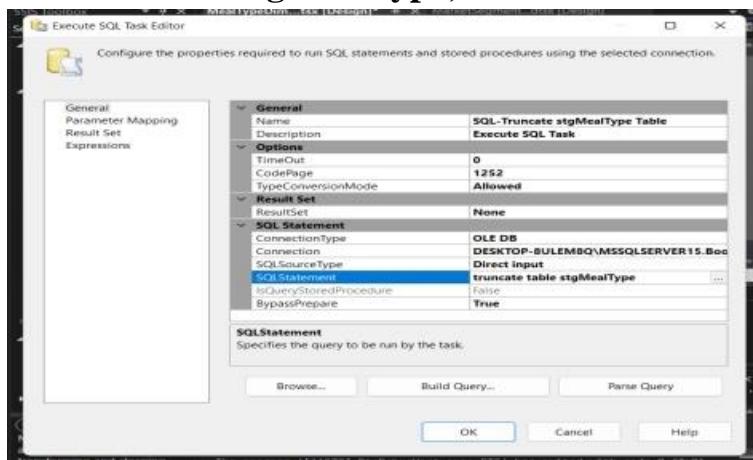
Thiết lập truncate table SQL task

Click **Control Flow tab**:

Double-click vào **SQL - Truncate stgMealType Table** task để cấu hình nó, **Execute SQL**

Task Editor sẽ mở ra.

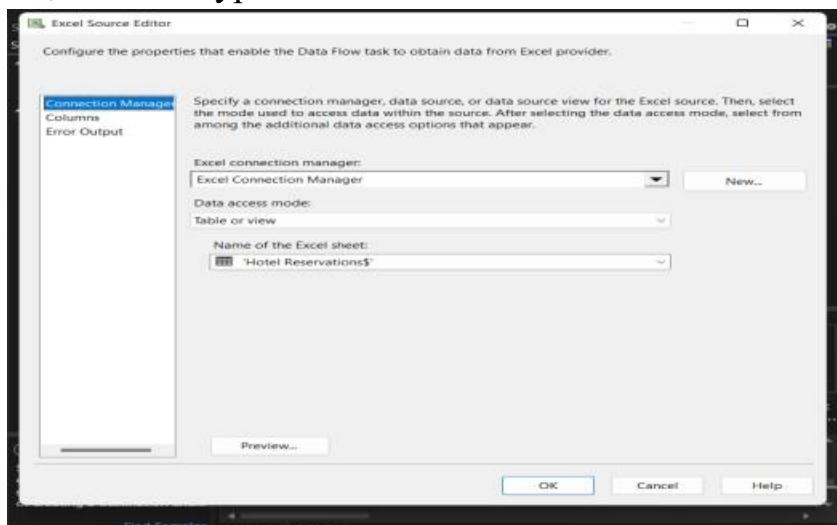
Ở mục **SQL Statement**, chọn **Connection** là **DESKTOP-BULEMBQ\MSSQLSERVER15.BookingHotelStage** và **SQLStatement** là **truncate table stgMealType;**



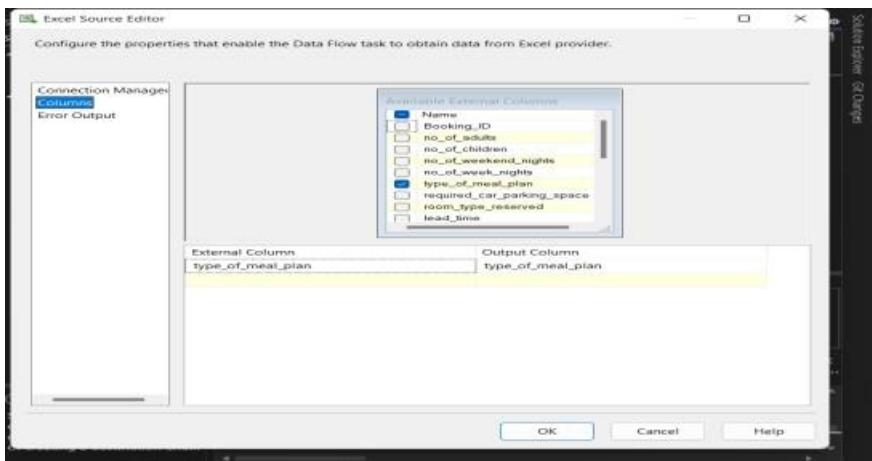
3.3.5.1 Extract

Double-click vào **DF-Extract From Source to Stage stgMealType** task để mở data flow design Surface.

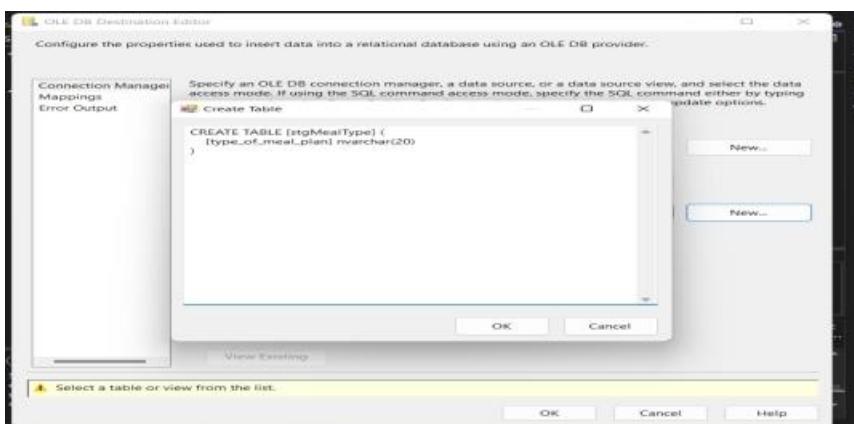
Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, chọn source Type: **Excel**



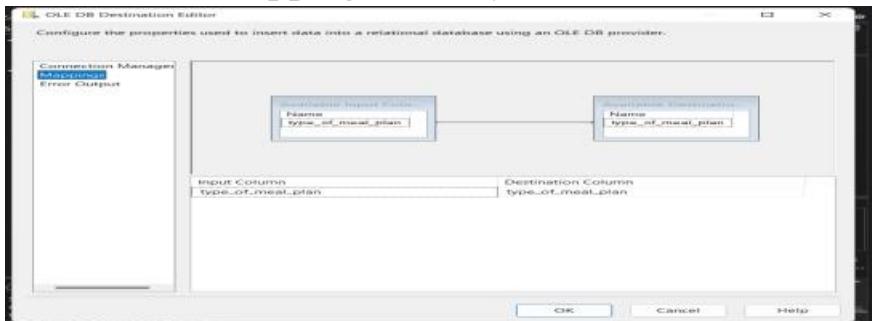
Chọn các column cần thiết cho bảng stgMealType



Double click vào destination để cấu hình nó, **OLE DB Destination** sẽ xuất hiện.
Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table**
Đổi tên bảng lại thành **stgMealType** (**CREATE TABLE [stgMealType]**)
Click OK.



Trở lại màn hình **OLEDB Destination Editor**, tại mục **Name of the table or view** bạn
sẽ thấy **[stgMealType]**.
Click vào tab **Mapping** ta sẽ thấy như hình bên dưới:



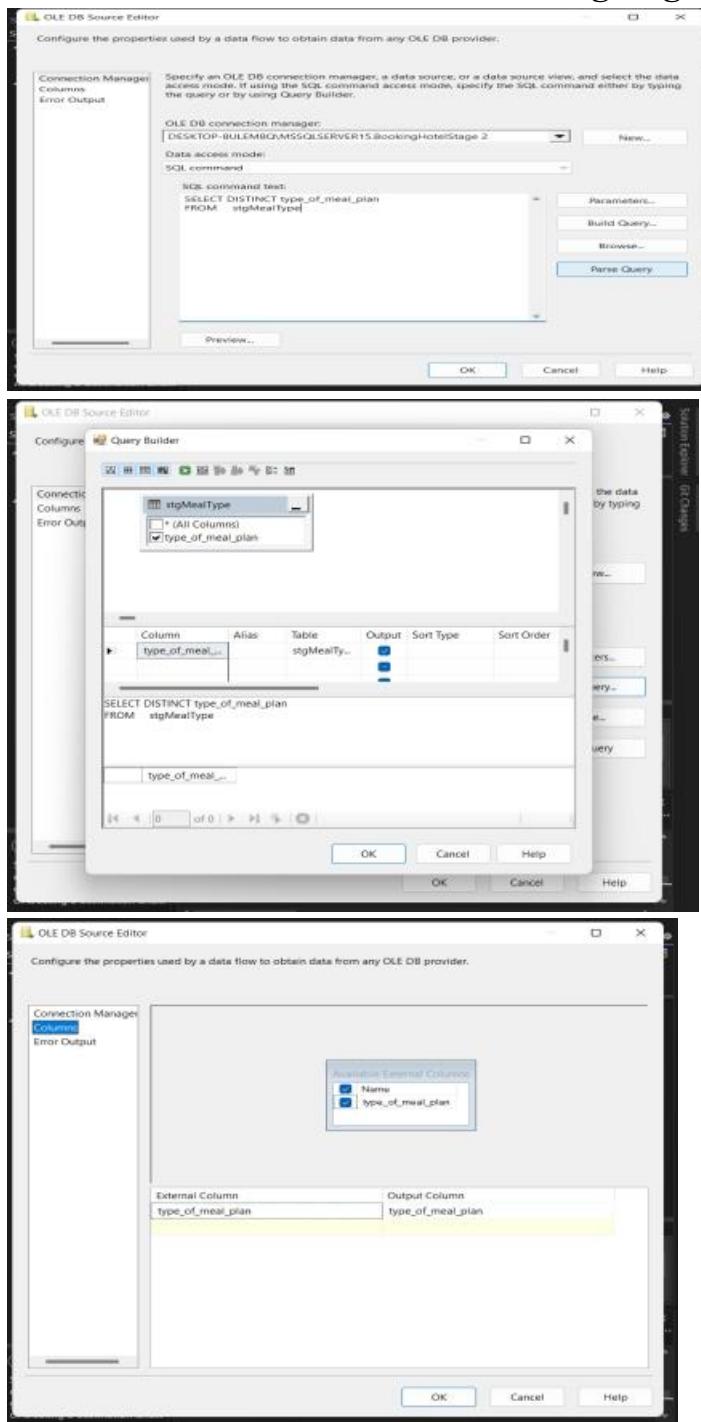
Click **OK** để hoàn tất việc cấu hình.

3.3.5.2 Load

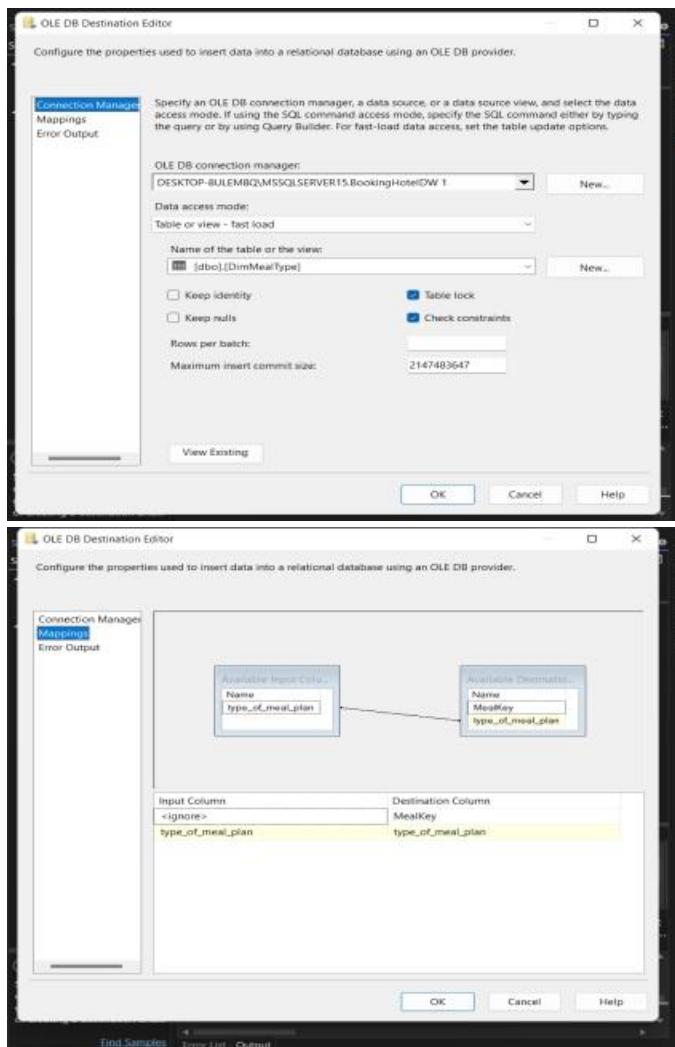
Double-click vào **DF - Load from stgMealType to DimMealType** task để mở data flow design Surface.

Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra,

Chọn source Type: **SQL Server** và **BookingHotelStage** connection manager.
 Câu hình và đổi tên nó thành **DST–Stage stgMealType**



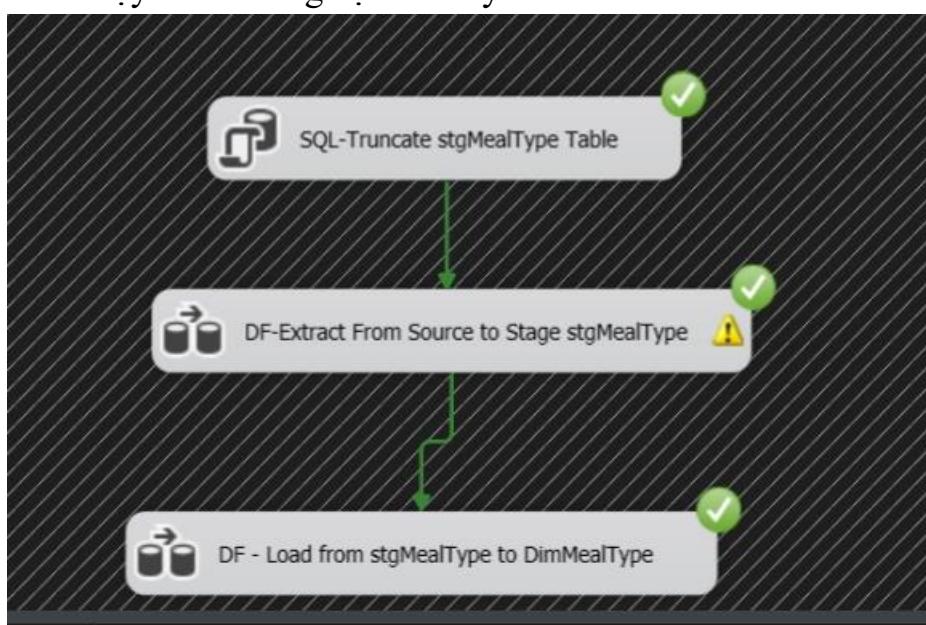
Ta có một **destination** như bên dưới: Đặt tên nó lại thành **SRC–Stage stgMealType** và connect **SRC–Stage stgMealType-> OLE DB Destination**
 Double click vào **OLE DB Destination** để cấu hình.

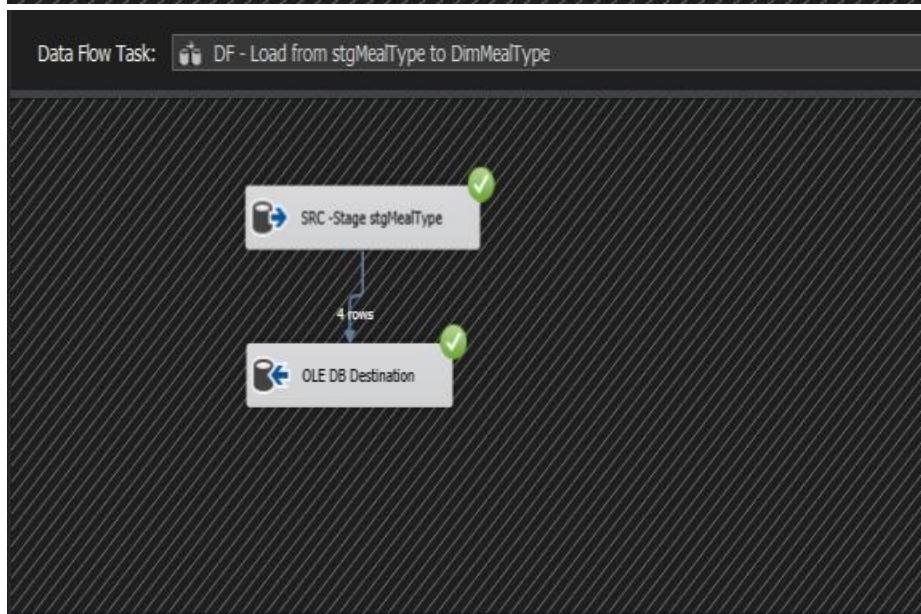
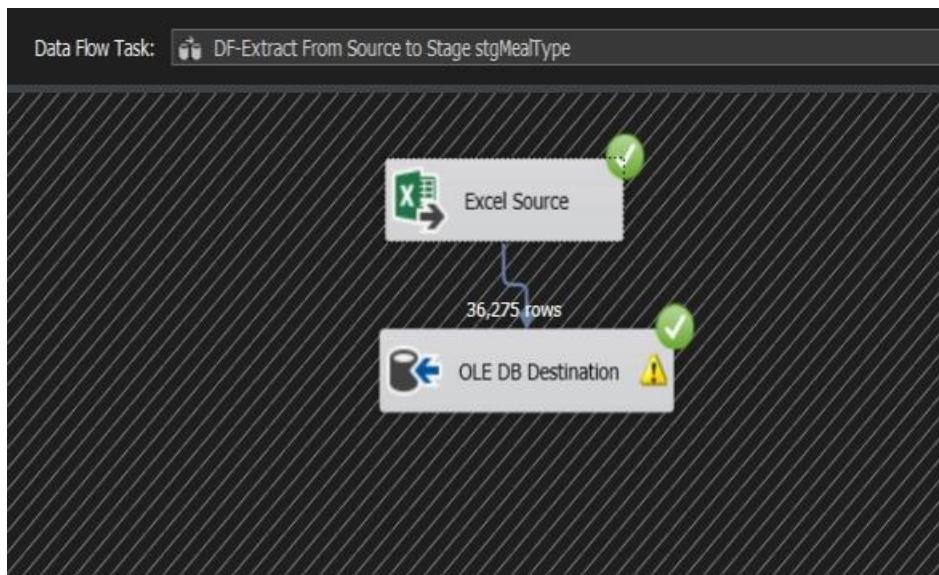


Click vào OK để hoàn tất cấu hình.

3.3.5.3 Run

Bấm Start để chạy. Nếu xảy ra lỗi, có gắng đọc hiểu thông báo lỗi để tìm cách sửa. Nếu chạy thành công bạn sẽ thấy màn hình như sau:





Kiểm tra kết quả trong bảng **stgMealType** trong cơ sở dữ liệu BookingHotelStage.

SQLQuery11.sql - DESKTOP-BULEMBQ\YenNNH (79) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Quick Launch (Ctrl+Q)

Browsing Stage

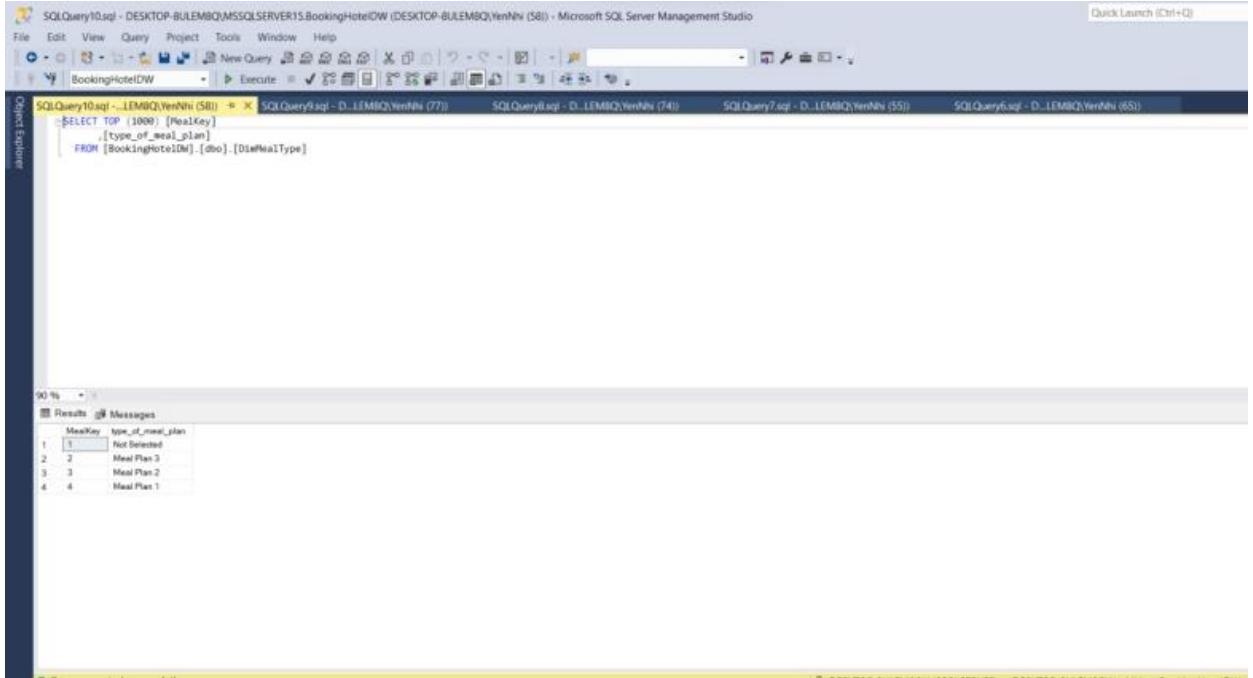
SQLQuery1.sql - [LEMBO]\YenNNH (79) SQLQuery10.sql - [LEMBO]\YenNNH (58) SQLQuery9.sql - D...\LEMBO\YenNNH (77) SQLQuery8.sql - D...\LEMBO\YenNNH (74) SQLQuery7.sql - D...\LEMBO\YenNNH (55)

SELECT TOP (2000) [type_of_meal_plan]
FROM [BookingHotelStage].[dbo].[stgMealType]

Results Messages

type_of_meal_plan
1 Meal Plan 1
2 Not Selected
3 Meal Plan 1
4 Meal Plan 1
5 Not Selected
6 Meal Plan 2
7 Meal Plan 1
8 Meal Plan 1
9 Meal Plan 1
10 Meal Plan 1
11 Not Selected
12 Meal Plan 1
13 Not Selected
14 Meal Plan 1
15 Meal Plan 1
16 Meal Plan 2
17 Meal Plan 1
18 Not Selected
19 Meal Plan 1
20 Meal Plan 1
21 Meal Plan 1
22 Meal Plan 1
23 Meal Plan 1
24 Meal Plan 1
25 Not Selected
26 Meal Plan 1
27 Meal Plan 1
28 Meal Plan 1
29 Meal Plan 1
30 Meal Plan 2

Kiểm tra kết quả trong bảng **DimMealType** trong cơ sở dữ liệu BookingHotelDW.



The screenshot shows the Microsoft SQL Server Management Studio interface. In the top bar, there are tabs for 'SQLQuery10.sql - DESKTOP-BULEMBQ\MSQLSERVER15.BookingHotelDW (DESKTOP-BULEMBQ\YenNh1 (S8)) - Microsoft SQL Server Management Studio' and 'Quick Launch (Ctrl+Q)'. Below the tabs, there are several other tabs labeled 'SQLQuery9.sql - D...LEM8Q\YenNh1 (77)', 'SQLQuery8.sql - D...LEM8Q\YenNh1 (74)', 'SQLQuery7.sql - D...LEM8Q\YenNh1 (55)', and 'SQLQuery6.sql - D...LEM8Q\YenNh1 (65)'. The main area displays a query window with the following SQL code:

```
SELECT TOP (1000) [MealKey]
,[type_of_meal_plan]
FROM [BookingHotelDW].[dbo].[DimMealType]
```

The results pane shows a table with two columns: 'MealKey' and 'type_of_meal_plan'. The data is as follows:

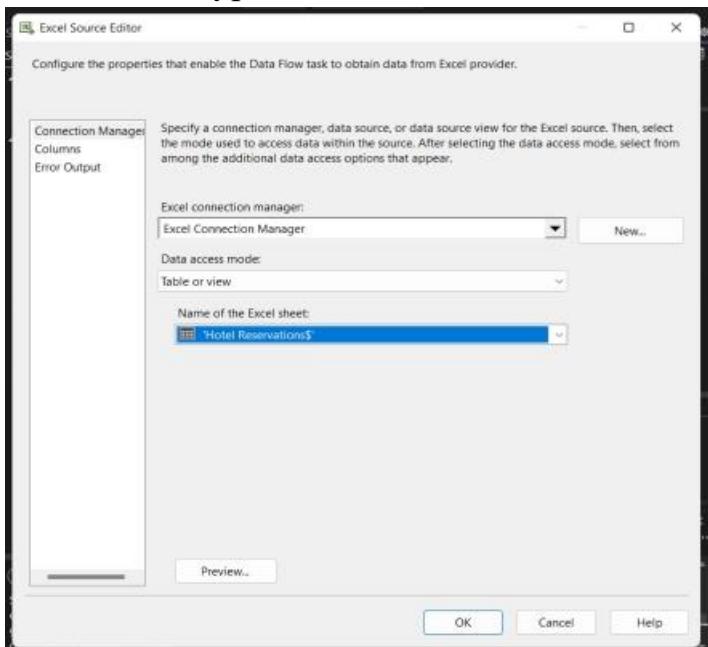
MealKey	type_of_meal_plan
1	Not Specified
2	Meal Plan 3
3	Meal Plan 2
4	Meal Plan 1

3.3.6 Room Type Dimension

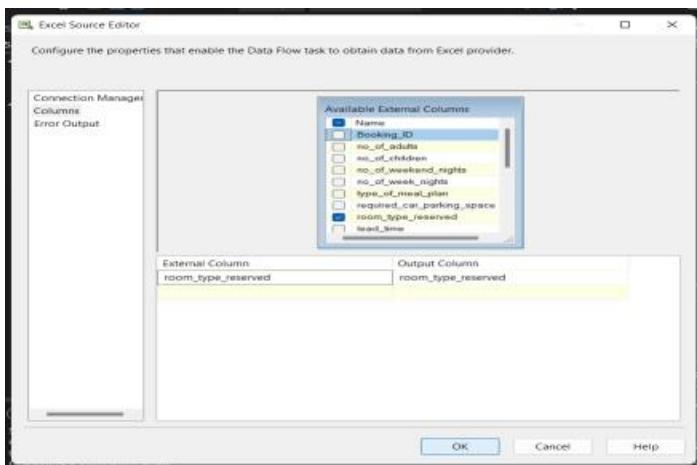
3.3.6.1 Extract

Double-click vào **DF-Extract From Source to Stage stgRoomType** task để mở data flow design Surface.

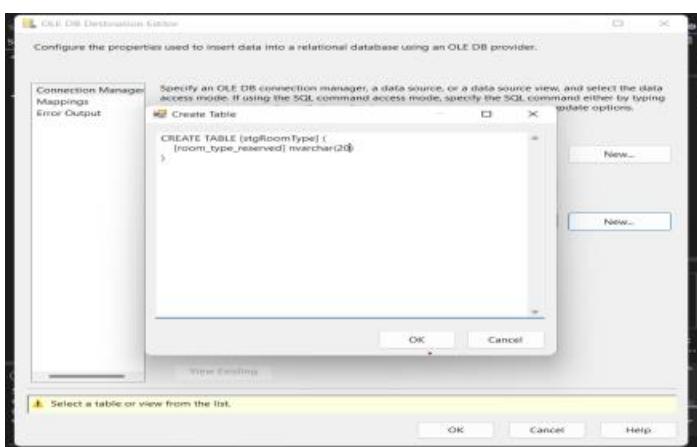
Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, chọn source Type: **Excel**



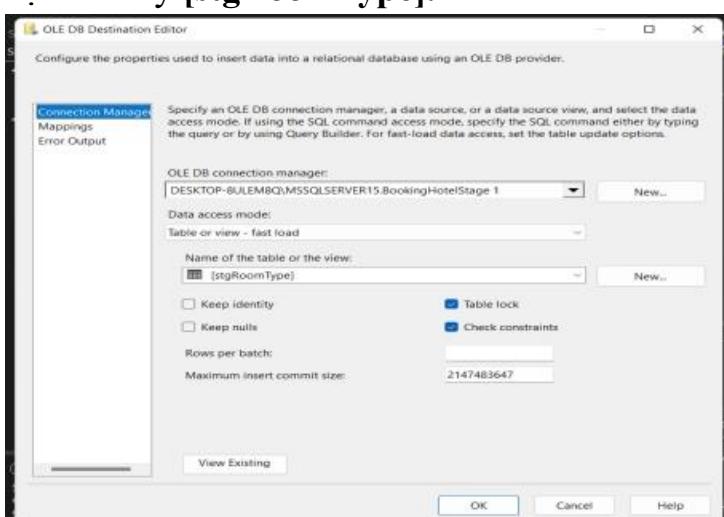
Chọn các column cần thiết cho bảng stgRoomType



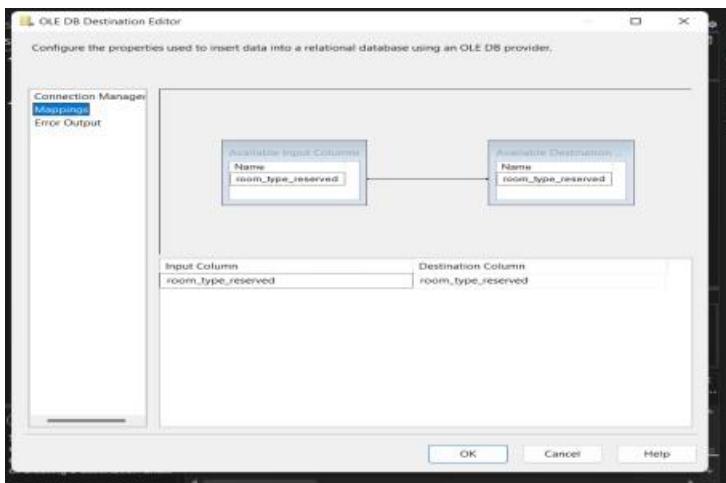
Double click vào destination để cấu hình nó, **OLE DB Destination** sẽ xuất hiện.
Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table**
Đổi tên bảng lại thành **stgRoomType** (**CREATE TABLE [stgRoomType]**)
Click OK.



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



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



Click **OK** để hoàn tất việc cấu hình.

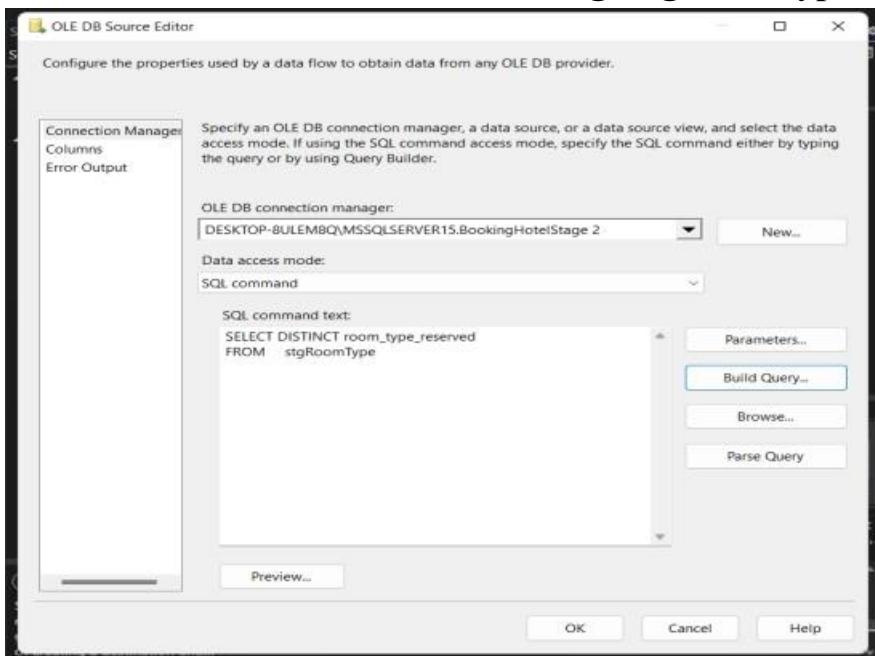
3.3.6.2 Load

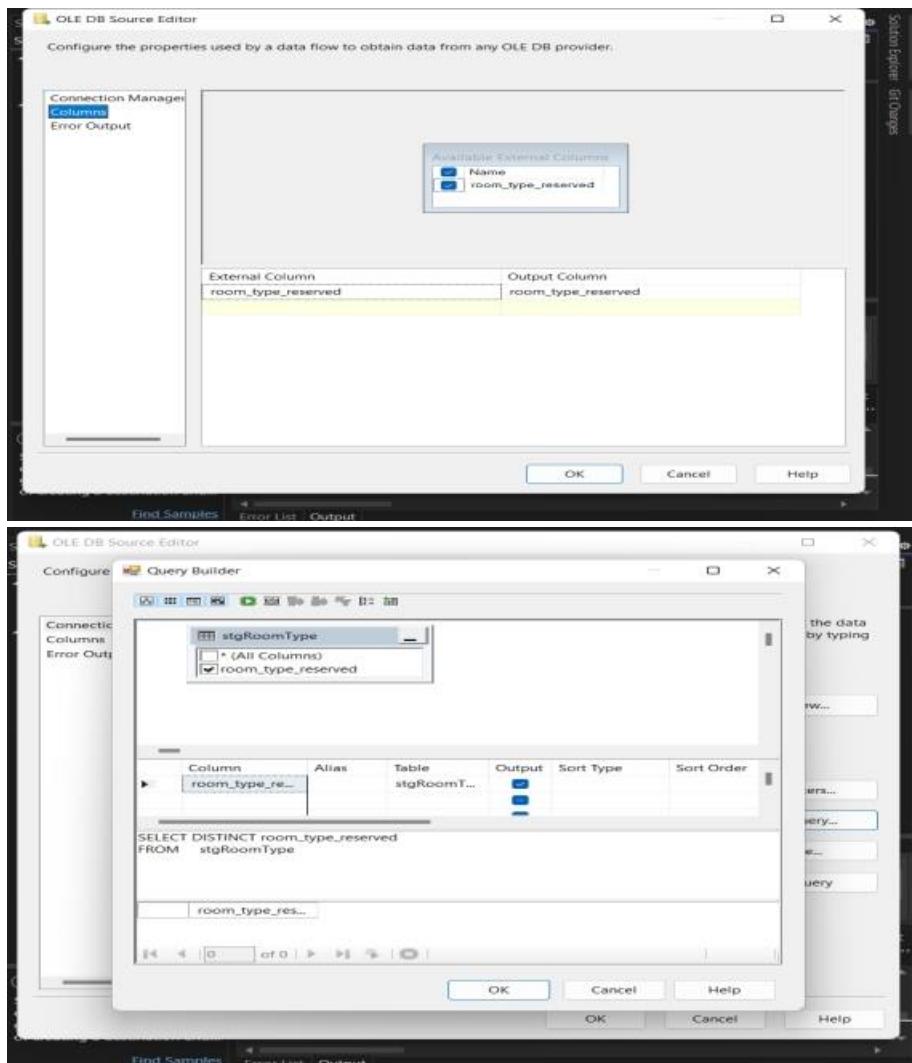
Double-click vào **DF - Load from stgRoomType to DimRoomType** task để mở data flow design Surface.

Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra,

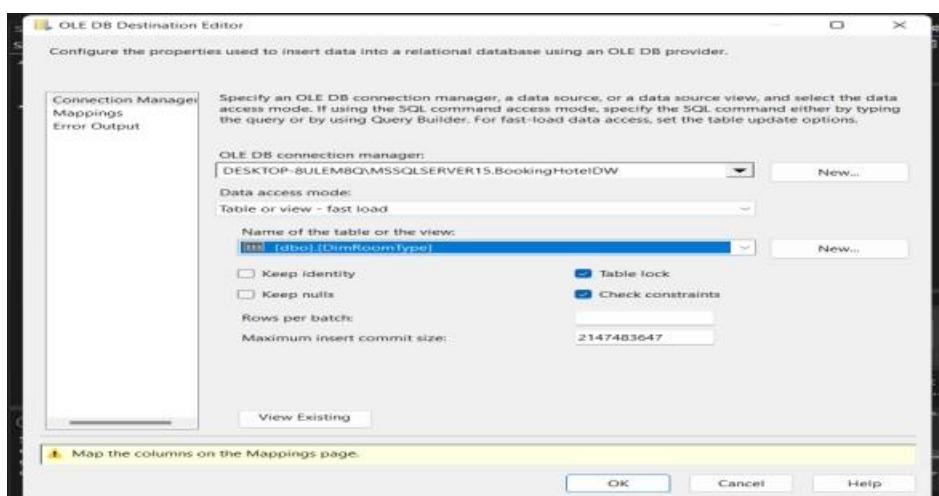
Chọn source Type: **SQL Server** và **BookingHotelStage** connection manager.

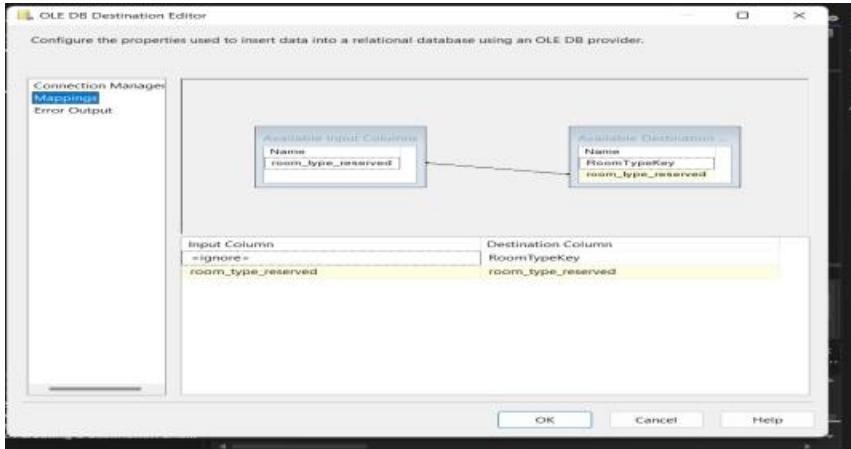
Cấu hình và đổi tên nó thành **DST–Stage stgRoomType**





Ta có một **destination** như bên dưới: Đặt tên nó lại thành **SRC-Stage stgRoomType** và connect **SRC-Stage stgRoomType-> OLE DB Destination**
Double click vào **OLE DB Destination** để cấu hình.

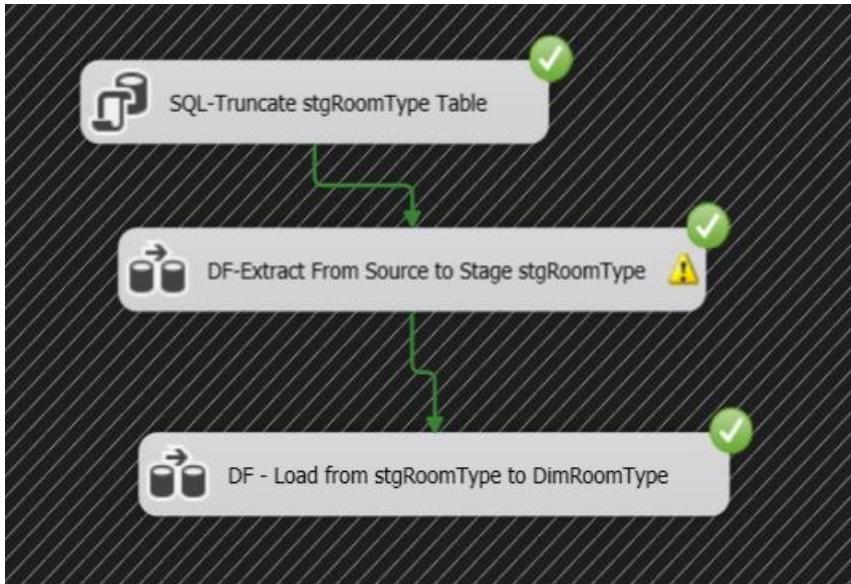


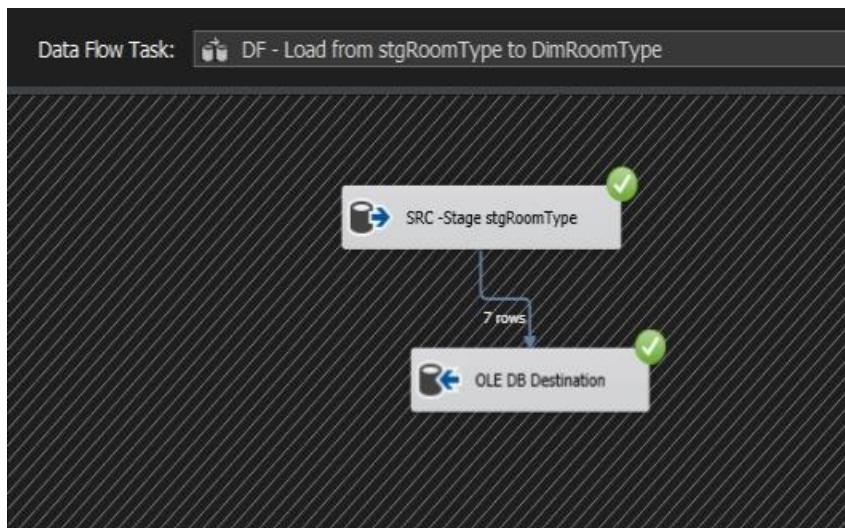


Click vào OK để hoàn tất cấu hình.

3.3.6.3 Run

Bấm Start để chạy. Nếu xảy ra lỗi, cố gắng đọc hiểu thông báo lỗi để tìm cách sửa. Nếu chạy thành công bạn sẽ thấy màn hình như sau:





Kiểm tra kết quả trong bảng **stgRoomType** trong cơ sở dữ liệu BookingHotelStage.

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. At the top, there's a menu bar with File, Edit, View, Query, Project, Tools, Window, Help. Below the menu is a toolbar with various icons. The title bar indicates the connection is to DESKTOP-8ULEM8Q\BOOKINGHOTELSTAGE (YenNhi (68)). There are three tabs open in the query editor: SQLQuery13.sql (active), SQLQuery12.sql, and SQLQuery11.sql. The active query window contains the following SQL code:

```
SELECT TOP (1000) [room_type_reserved]
FROM [BookingHotelStage].[dbo].[stgRoomType]
```

The results grid shows a single column named "room_type_reserved" with 25 rows, all containing the value "Room_Type 1".

At the bottom of the screen, a yellow status bar displays the message "Query executed successfully."

Kiểm tra kết quả trong bảng **stgRoomType** trong cơ sở dữ liệu BookingHotelDW.

The screenshot shows the Microsoft SQL Server Management Studio interface. In the top bar, it says "SQLQuery12.sql - DESKTOP-8ULEMBQ\MSSQLSERVER15.BookingHotelDW (DESKTOP-8ULEMBQ\YenNhi (78)) - Microsoft SQL Server Management Studio". Below the bar are standard menu options: File, Edit, View, Query, Project, Tools, Window, Help. A toolbar follows with icons for New Query, Execute, Save, and others. The Object Explorer sidebar is visible on the left. The main area has four tabs at the top: "SQLQuery12.sql - ...LEM8Q\YenNhi (78)" (selected), "SQLQuery11.sql - ...LEM8Q\YenNhi (79)", "SQLQuery10.sql - ...LEM8Q\YenNhi (58)", and "SQLQuery9.sql - D...LEM8Q\YenNhi (57)". The selected tab contains the following SQL query:

```
SELECT TOP (1000) [RoomTypeKey]
      ,[room_type_reserved]
  FROM [BookingHotelDW].[dbo].[DimRoomType]
```

Below the query is a results grid titled "Results". It displays the following data:

	RoomTypeKey	room_type_reserved
1	1	Room_Type 3
2	2	Room_Type 5
3	3	Room_Type 2
4	4	Room_Type 4
5	5	Room_Type 1
6	6	Room_Type 7
7	7	Room_Type 6

3.4 Import dữ liệu vào bảng các bảng fact

3.4.1 Booking Fact

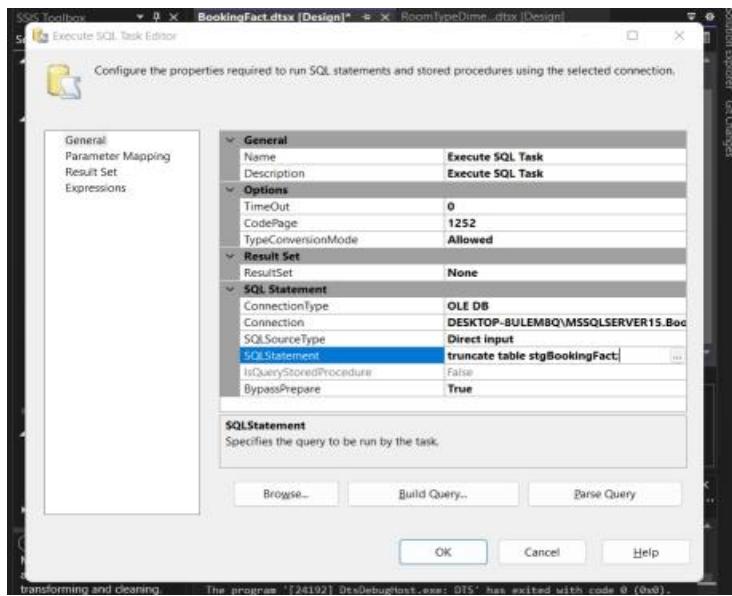
Thiết lập truncate table SQL task

Click **Control Flow** tab:

Double-click vào **SQL - Truncate stgBookingFact Table** task để cấu hình nó, **Execute SQL**

Task Editor sẽ mở ra.

Ở mục **SQL Statement**, chọn **Connection** là **DESKTOP-8ULEMBQ\MSSQLSERVER15.BookingHotelStage** và **SQLStatement** là **truncate table stgBookingFact;**

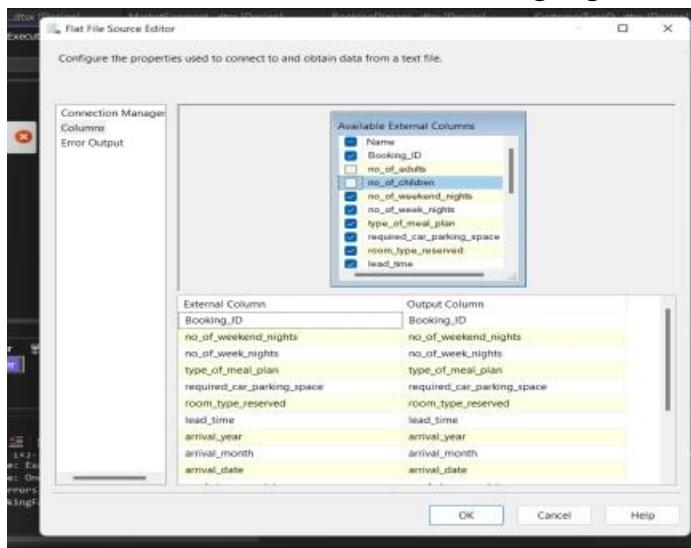


3.4.1.1 Extract

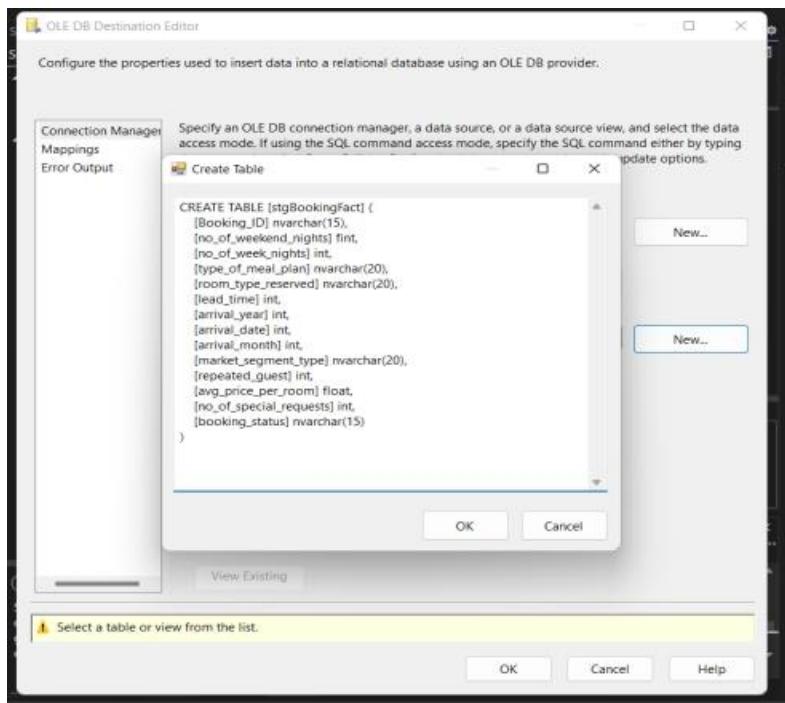
Double-click vào **DF-Extract From Source to Stage stgBookingFact** task để mở data flow design Surface.

Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, chọn source Type: **Excel**

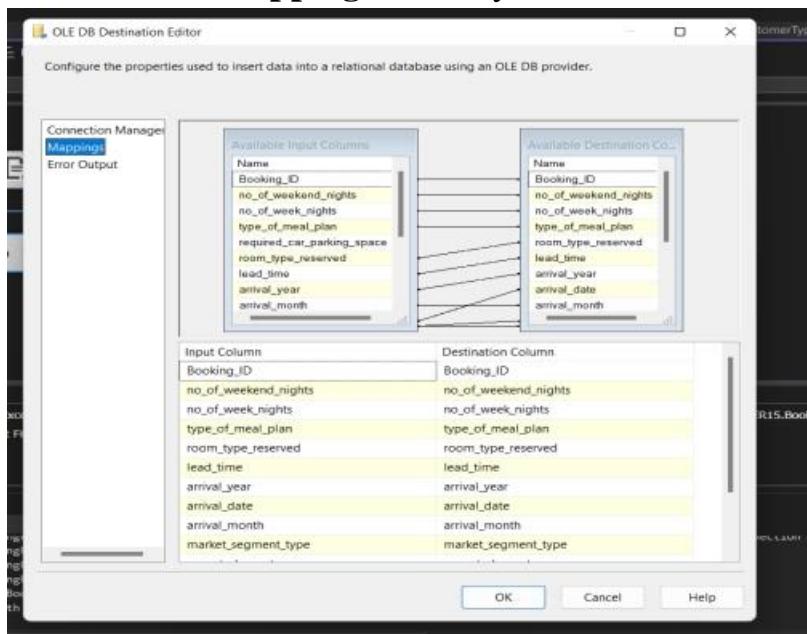
Chọn các column cần thiết cho bảng stgBookingFact



Double click vào destination để cấu hình nó, **OLE DB Destination** sẽ xuất hiện. Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table** Đổi tên bảng lại thành **stgBookingFact (CREATE TABLE [stgBookingFact])** Click OK.



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

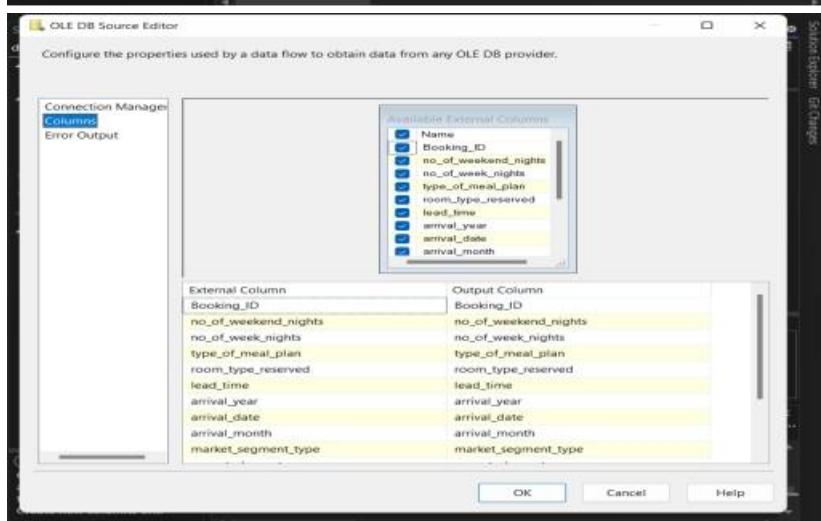
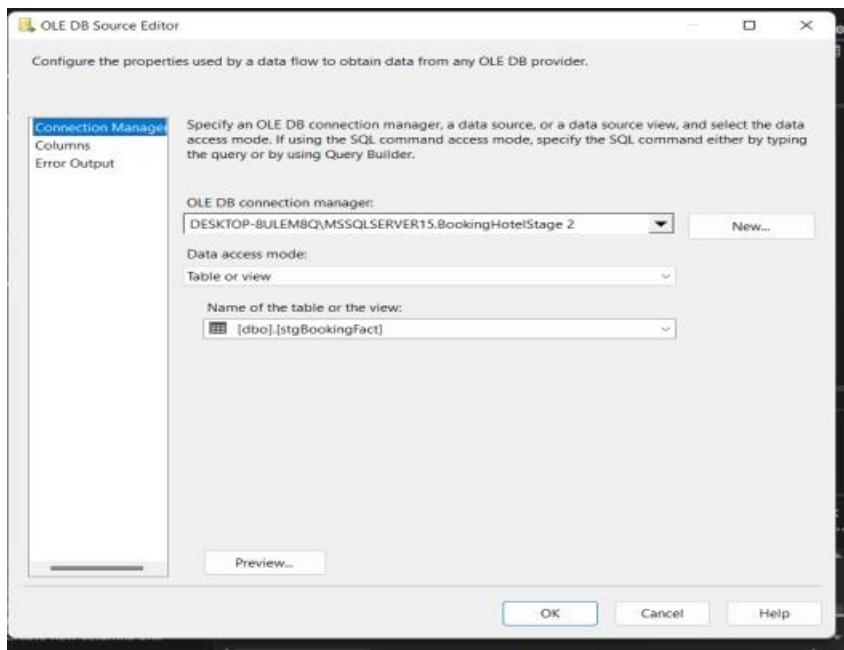


Click Ok để hoàn tất cấu hình.

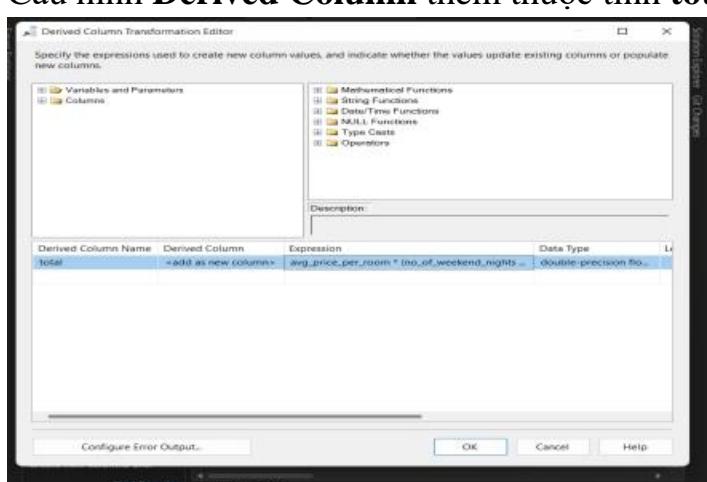
3.4.1.2 Load

Double-click vào **DF - Load from stgBookingFact to FactBooking task** để mở data flow design Surface.

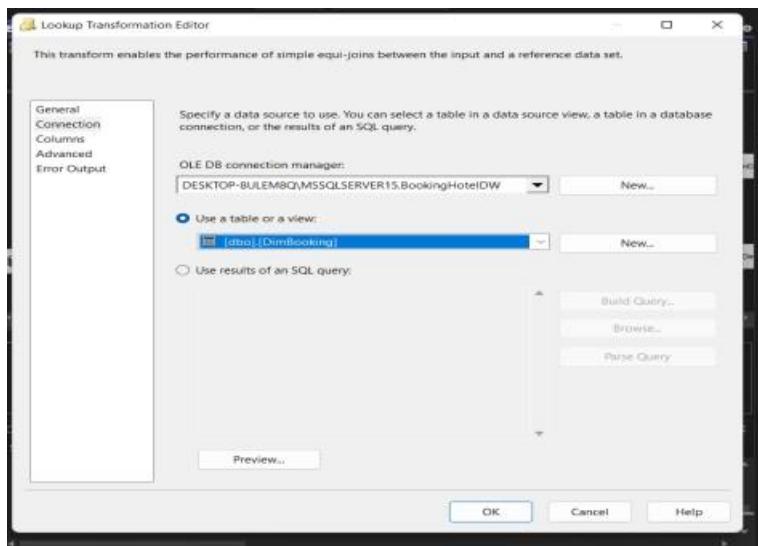
Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, Chọn source Type: **SQL Server** và **BookingHotelStage** connection manager.



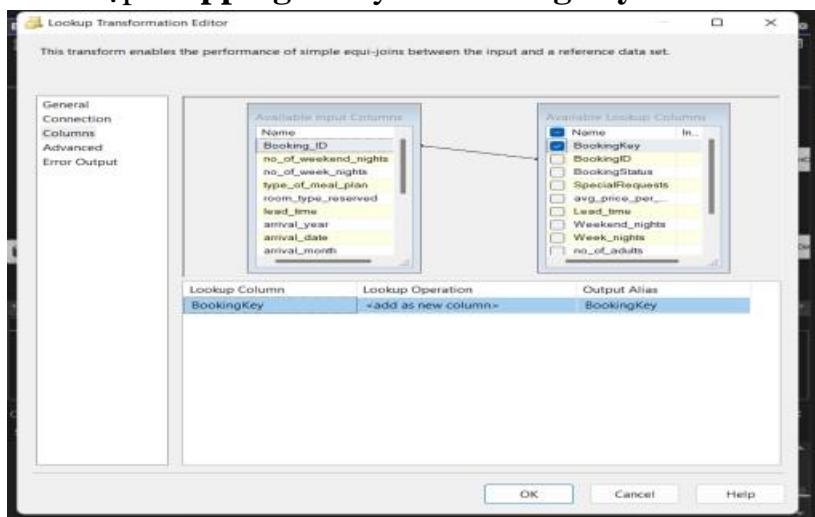
Cấu hình Derived Column thêm thuộc tính total



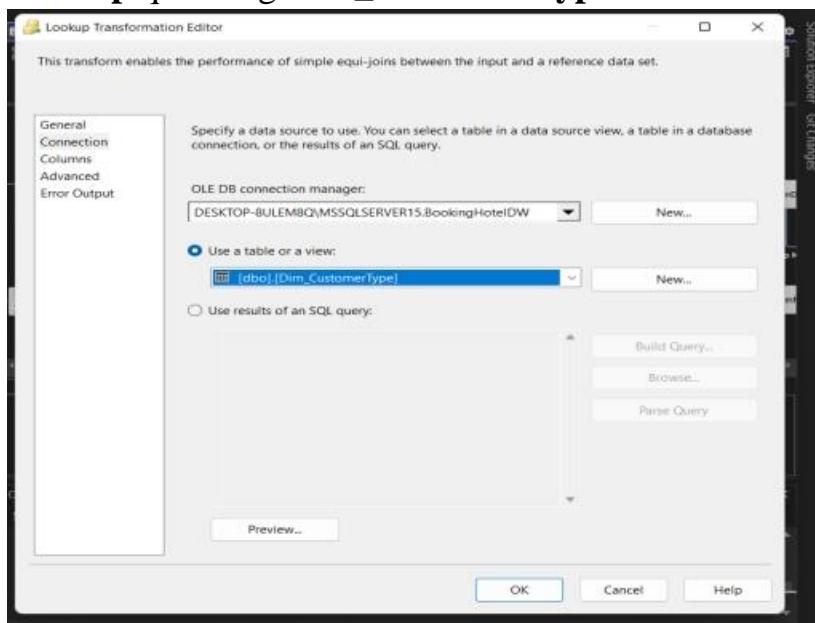
Lookup qua bảng DimBooking



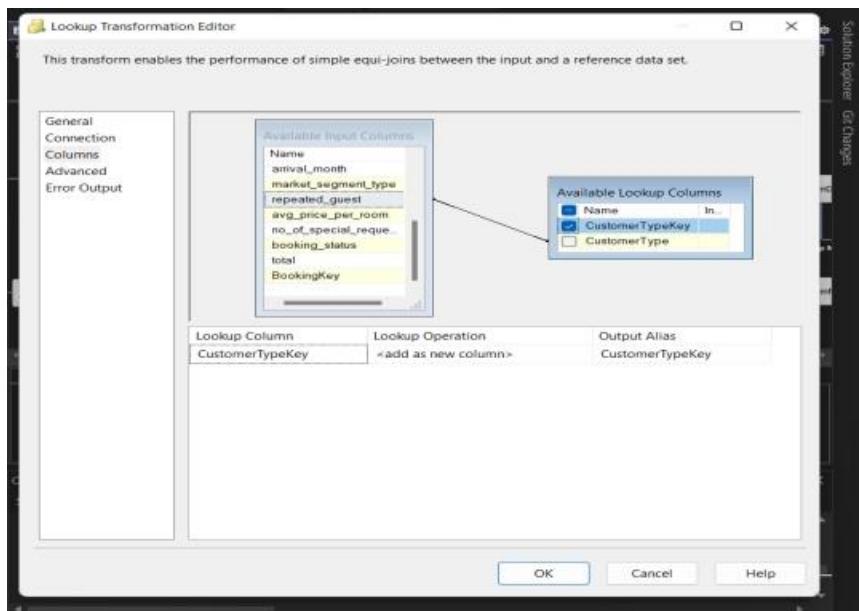
Thiết lập mapping và lấy ra BookingKey



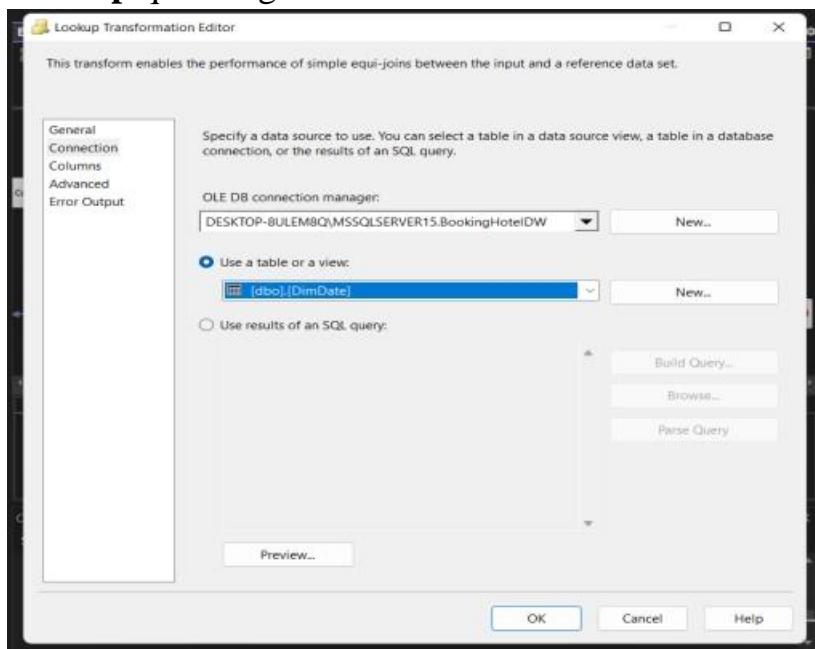
Lookup qua bảng Dim_CustomerType



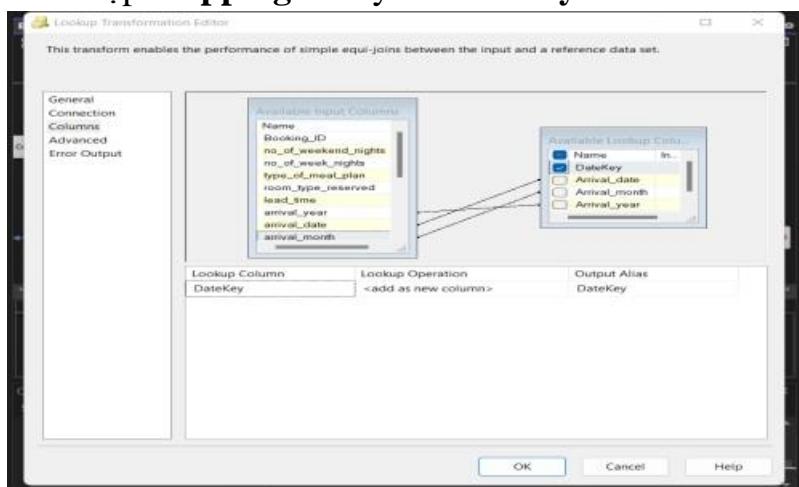
Thiết lập mapping và lấy ra CustomerTypeKey



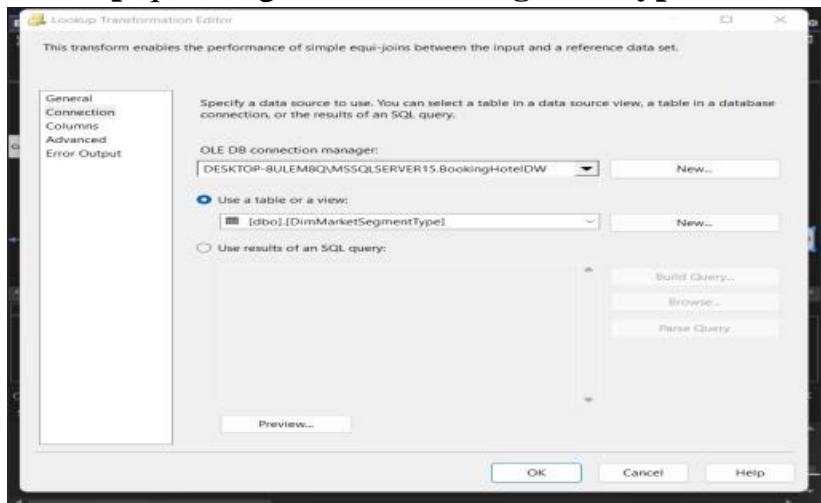
Lookup qua bảng DimDate



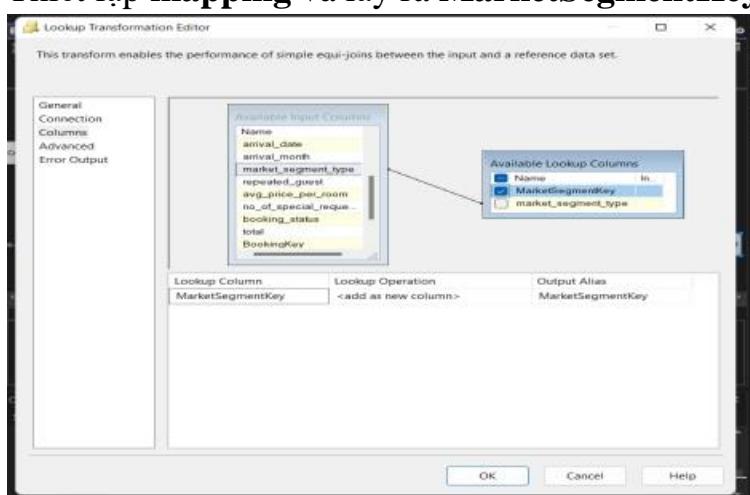
Thiết lập mapping và lấy ra DateKey



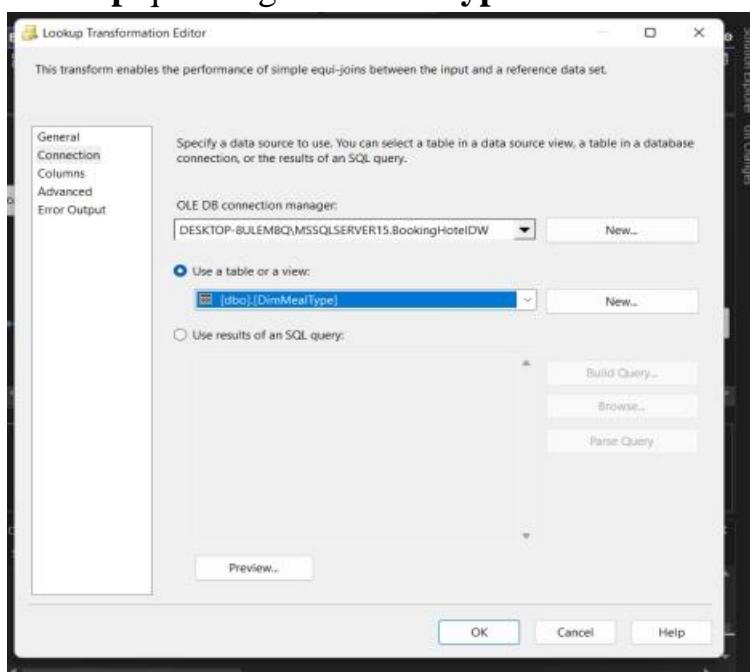
Lookup qua bảng DimMarketSegmentType



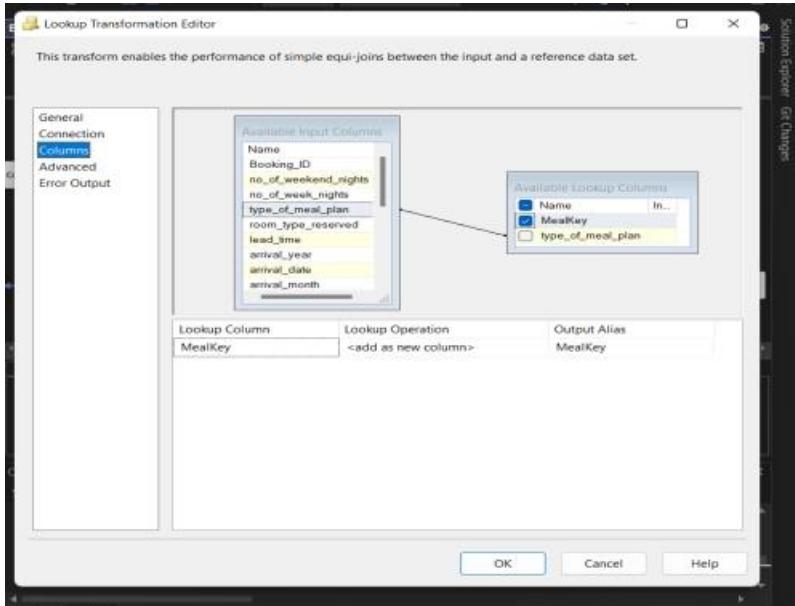
Thiết lập mapping và lấy ra MarketSegmentKey



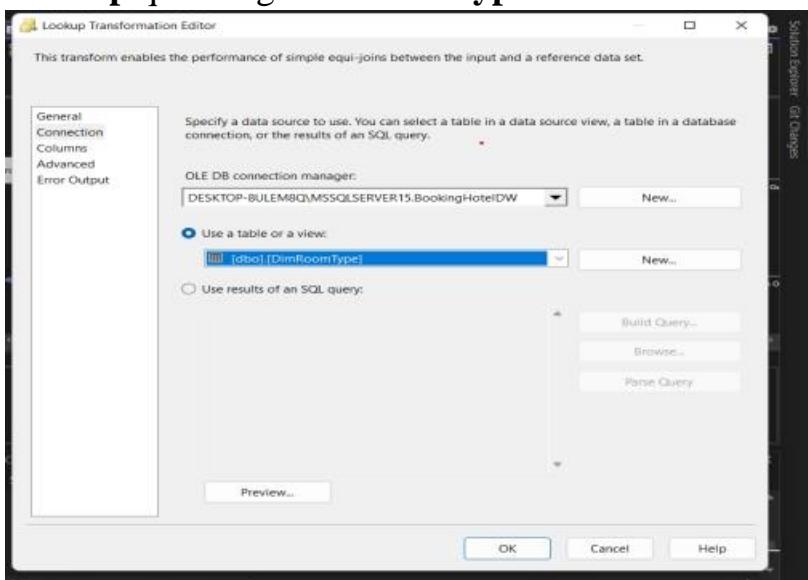
Lookup qua bảng DimMealType



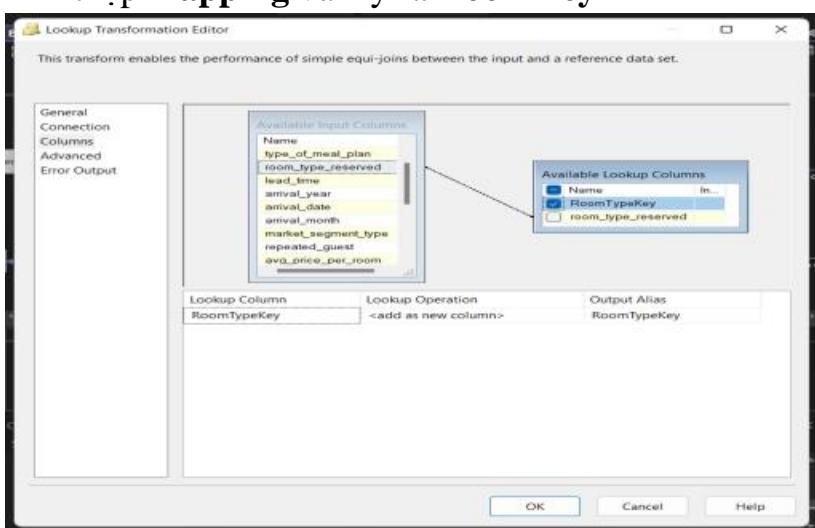
Thiết lập mapping và lấy ra MealKey



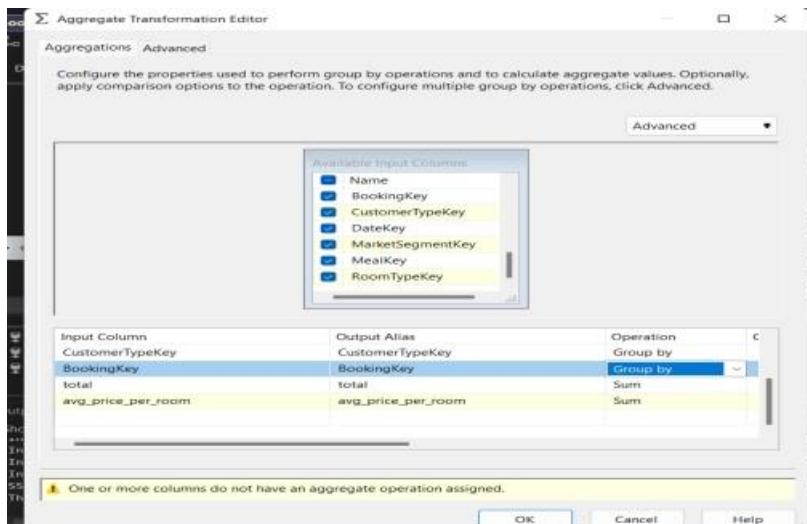
Lookup qua bảng DimRoomType



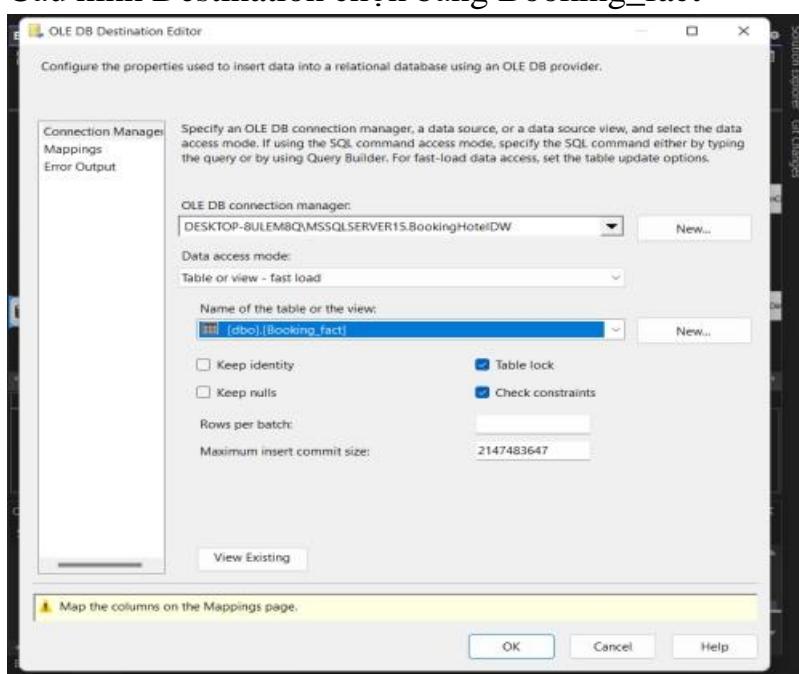
Thiết lập mapping và lấy ra RoomKey



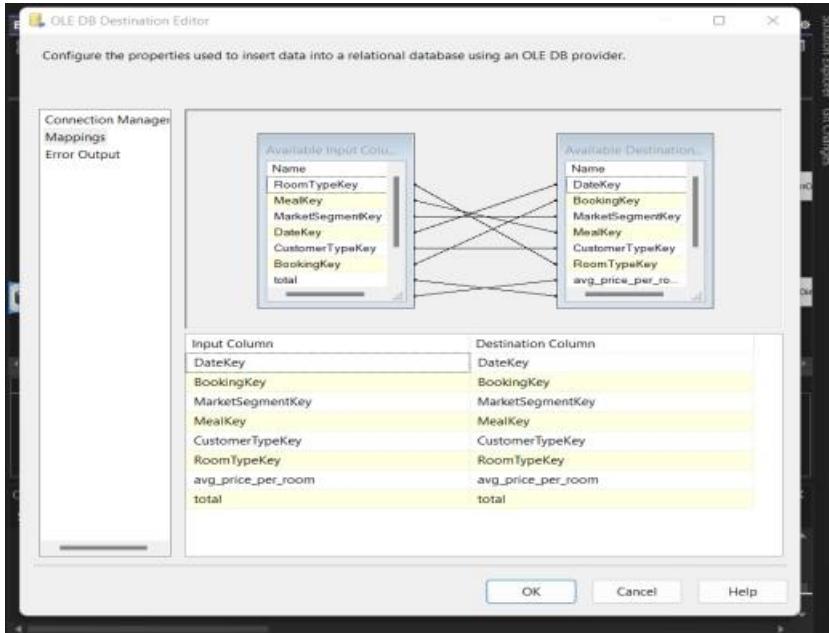
Cấu hình Aggregate chọn các thuộc tính cần thiết và các measures



Cấu hình Destination chọn bảng Booking_fact

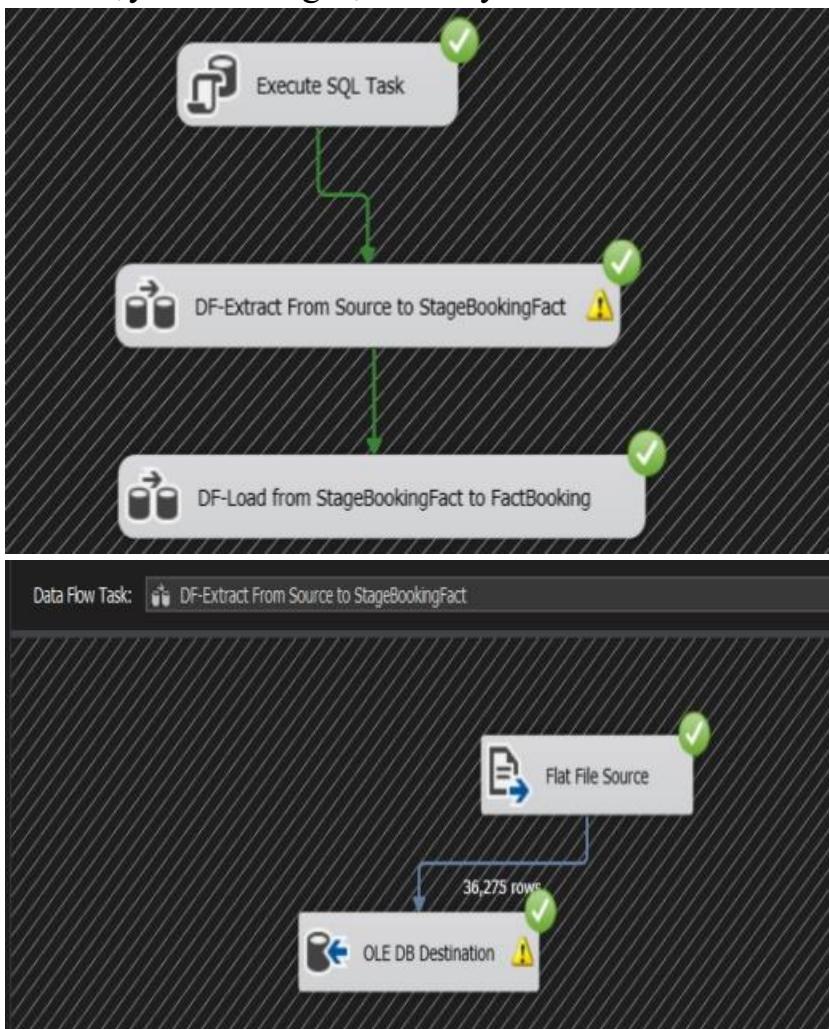


Mapping các cột lại với nhau.



3.4.1.3 Run

Bấm Start để chạy. Nếu xảy ra lỗi, có gắng đọc hiểu thông báo lỗi để tìm cách sửa. Nếu chạy thành công bạn sẽ thấy màn hình như sau:





Kiểm tra kết quả trong bảng **stgBookingFact** trong cơ sở dữ liệu BookingHotelStage.

SQLQuery14.sql - [EMBQ\YenNhi (63)] SQLQuery13.sql - [EMBQ\YenNhi (68)] SQLQuery12.sql - [EMBQ\YenNhi (78)] SQLQuery11.sql - [EMBQ\YenNhi (79)] SQLQuery10.sql - [EMBQ\YenNhi (88)]

```
SELECT TOP (1000) [Booking_ID]
      ,[no_of_weekend_nights]
      ,[no_of_week_nights]
      ,[type_of_meal_plan]
      ,[room_type_reserved]
      ,[lead_time]
      ,[arrival_year]
      ,[arrival_date]
      ,[arrival_month]
      ,[market_segment_type]
      ,[repeated_guest]
      ,[avg_price_per_room]
      ,[no_of_special_requests]
      ,[booking_status]
  FROM [BookingHotelStage].[dbo].[stgBookingFact]
```

Booking_ID	no_of_weekend_nights	no_of_week_nights	type_of_meal_plan	room_type_reserved	lead_time	arrival_year	arrival_date	arrival_month	market_segment_type	repeated_guest	avg_price_per_room	no_of_special_requests	booking_status
INN00001	1	2	Meal Plan 1	Room_Type 1	224	2017	2	10	Online	0	65	0	Not_Cancelled
INN00002	2	3	Not Selected	Room_Type 1	5	2018	8	11	Online	0	105.68	1	Not_Cancelled
INN00003	2	1	Meal Plan 1	Room_Type 1	1	2018	28	2	Online	0	60	0	Cancelled
INN00004	0	2	Meal Plan 1	Room_Type 1	211	2018	20	5	Online	0	100	0	Cancelled
INN00005	1	1	Not Selected	Room_Type 1	48	2018	11	4	Online	0	94.5	0	Cancelled
INN00006	0	2	Meal Plan 2	Room_Type 1	346	2018	13	9	Online	0	115	1	Cancelled
INN00007	1	3	Meal Plan 1	Room_Type 1	34	2017	* 15	10	Online	0	107.95	1	Not_Cancelled
INN00008	1	3	Meal Plan 1	Room_Type 4	63	2018	26	12	Online	0	105.61	1	Not_Cancelled
INN00009	0	4	Meal Plan 1	Room_Type 1	121	2018	6	7	Offline	0	96.9	1	Not_Cancelled
INN00010	0	5	Meal Plan 1	Room_Type 4	44	2018	18	10	Online	0	133.44	3	Not_Cancelled
INN00011	1	0	Not Selected	Room_Type 1	0	2018	11	9	Online	0	85.03	0	Not_Cancelled
INN00012	2	1	Meal Plan 1	Room_Type 4	35	2018	30	4	Online	0	140.4	1	Not_Cancelled
INN00013	2	1	Not Selected	Room_Type 1	30	2018	26	11	Online	0	88	0	Cancelled
INN00014	2	0	Meal Plan 1	Room_Type 1	95	2018	20	11	Online	0	90	2	Cancelled
INN00015	0	2	Meal Plan 1	Room_Type 1	47	2017	20	10	Online	0	94.5	2	Not_Cancelled
INN00016	0	2	Meal Plan 2	Room_Type 1	296	2018	15	6	Online	0	115	1	Cancelled
INN00017	1	0	Meal Plan 1	Room_Type 1	0	2017	5	10	Offline	0	96	0	Not_Cancelled
INN00018	1	3	Not Selected	Room_Type 1	1	2017	10	8	Online	0	98	1	Not_Cancelled
INN00019	2	2	Meal Plan 1	Room_Type 1	99	2017	30	10	Online	0	65	0	Cancelled
INN00020	1	0	Meal Plan 1	Room_Type 1	12	2017	4	10	Offline	0	72	0	Not_Cancelled
INN00021	2	3	Meal Plan 1	Room_Type 1	99	2017	30	10	Online	0	65	0	Cancelled

Query executed successfully.

Kiểm tra kết quả trong bảng **Booking_fact** trong cơ sở dữ liệu BookingHotelDW.

```

SELECT TOP (1000) [DateKey]
    ,[BookingKey]
    ,[MarketSegmentKey]
    ,[MealKey]
    ,[CustomerTypeKey]
    ,[RoomTypeKey]
    ,[avg_price_per_room]
    ,[total]
    FROM [BookingHotelDW].[dbo].[Booking_fact]

```

DateKey	BookingKey	MarketSegmentKey	MealKey	CustomerTypeKey	RoomTypeKey	avg_price_per_room	total	
1	20180213	31480	1	4	36276	5	70.07	70.07
2	20181010	4829	2	4	36277	5	65	260
3	20180415	34865	1	1	36276	5	77.35	386.75
4	20181001	17508	4	4	36276	5	89.75	89.75
5	20180517	24233	2	4	36276	5	130	260
6	20181125	33920	2	4	36276	5	67	67
7	20180706	12330	1	4	36276	5	90.95	363.8
8	20170814	19729	4	3	36276	5	101.5	203
9	20180704	20525	1	4	36276	5	123.25	616.25
10	20180705	16432	1	4	36276	5	114.83	459.32
11	20181125	34018	2	4	36276	5	67	67
12	20171231	34837	4	4	36276	5	118	354
13	20180215	23475	1	4	36276	5	80.82	404.1
14	20180701	32860	1	4	36276	5	169	169
15	20180513	25411	1	4	36276	4	156	156
16	20181002	13590	4	4	36276	5	85	255
17	20180712	32156	4	1	36276	5	84	336
18	20181031	11985	1	1	36276	5	93.75	375
19	20170814	19846	4	3	36276	5	101.5	203
20	20180513	16548	4	4	36276	5	61	305
21	20180828	12753	1	4	36276	5	135.9	135.9
22	20180213	31759	1	4	36276	5	91	91
23	20170814	19867	4	3	36276	5	87	174
24	20180220	3097	1	4	36276	5	93	93
25	20181010	5189	2	4	36276	5	159	477
26	20170813	33497	1	4	36276	5	76.5	382.5
27	20180704	20698	1	4	36276	5	133.88	535.52

Query executed successfully.

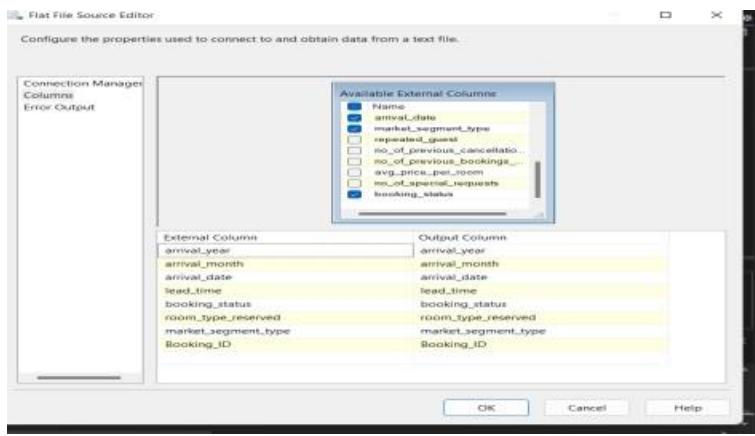
3.4.2 Cancellation Fact

3.4.2.1 Extract

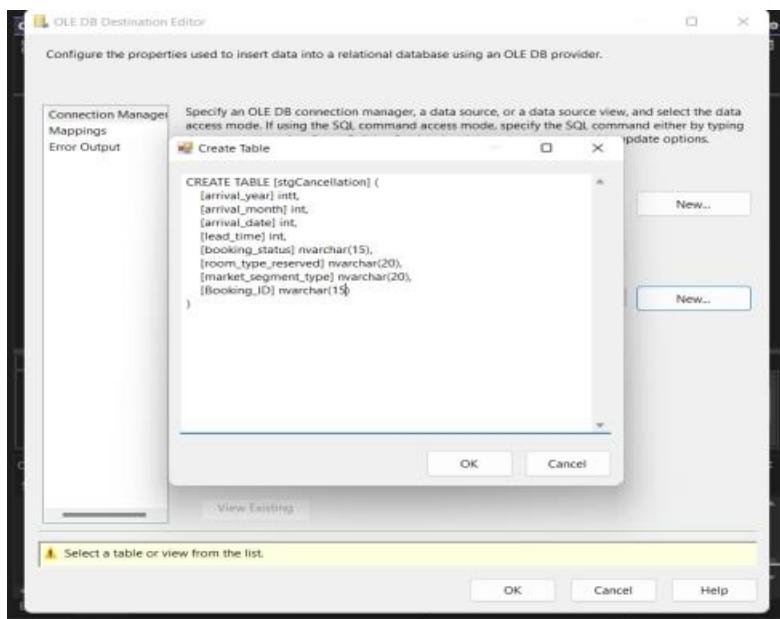
Double-click vào **DF-Extract From Source to Stage stgCancellationFact** task để mở data flow design Surface.

Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, chọn source Type: **Excel**

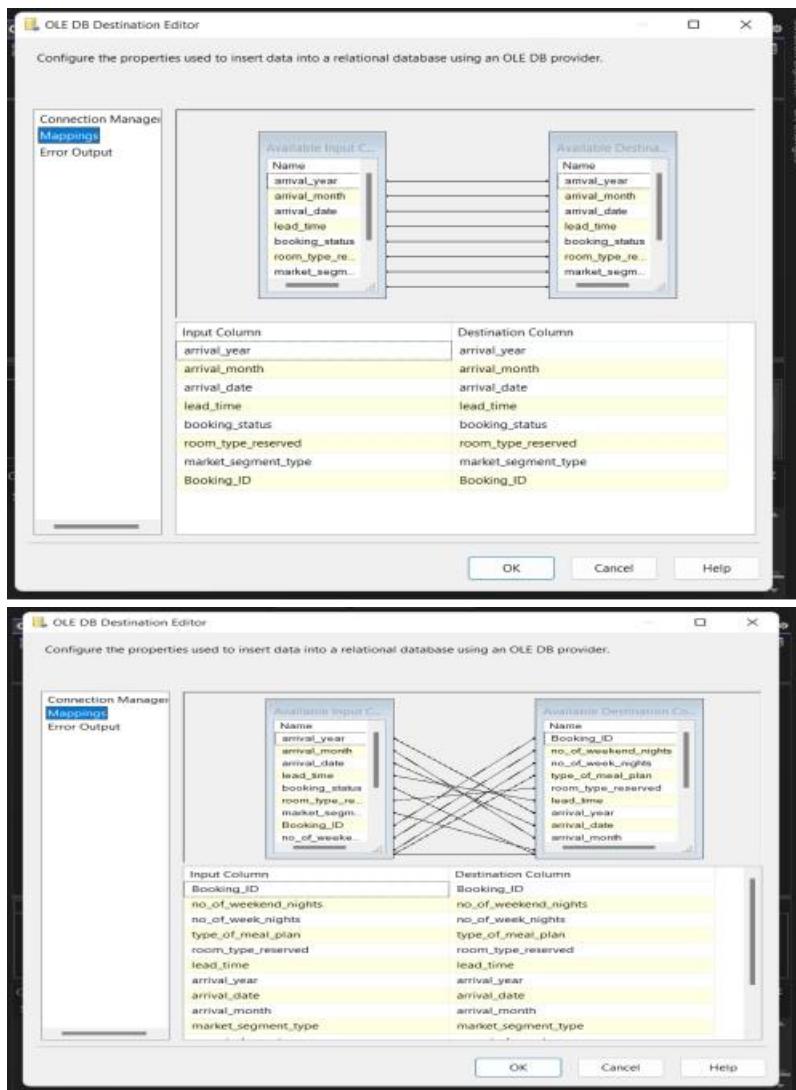
Chọn các column cần thiết cho bảng stgCancellationFact



Double click vào destination để cấu hình nó, **OLE DB Destination** sẽ xuất hiện.
Click nút **New...** để tạo bảng trong CSDL stage, ta sẽ thấy window **Create Table**
Đổi tên bảng lại thành **stgCancellationFact** (**CREATE TABLE [stgCancellationFact]**)
Click OK.



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

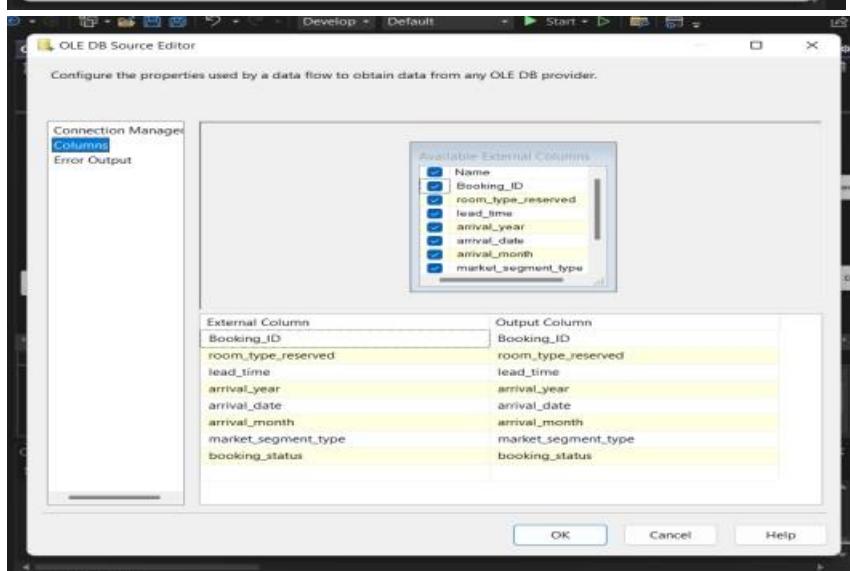
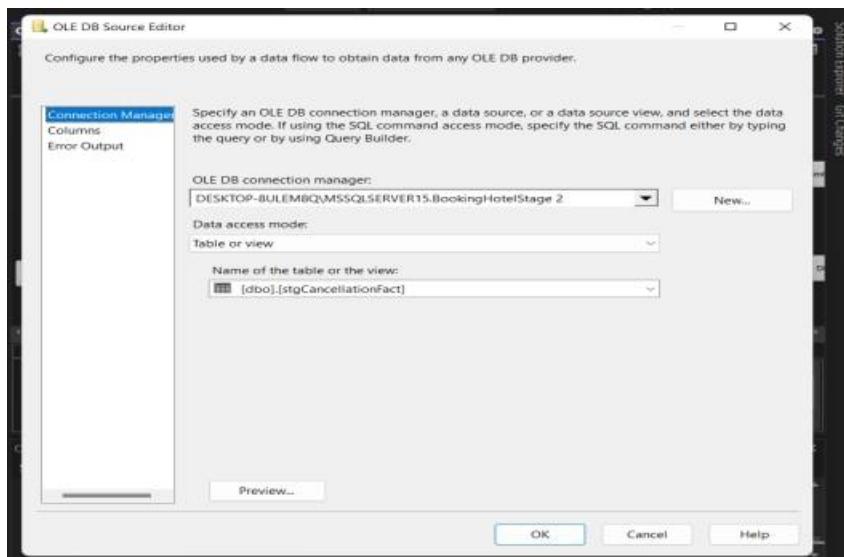


Click **OK** để hoàn tất việc cấu hình.

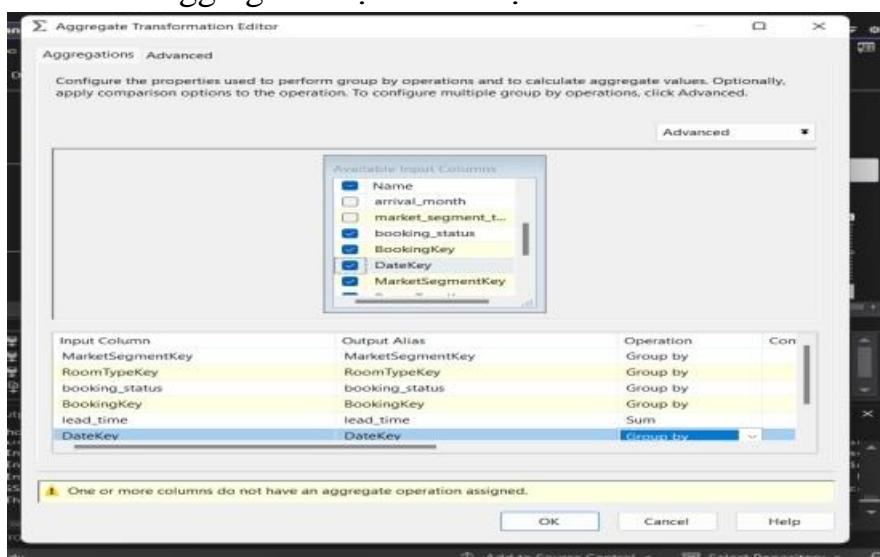
3.4.2.2 Load

Double-click vào **DF - Load from stgCancellationFact to FactCancellation** task để mở data flow design Surface.

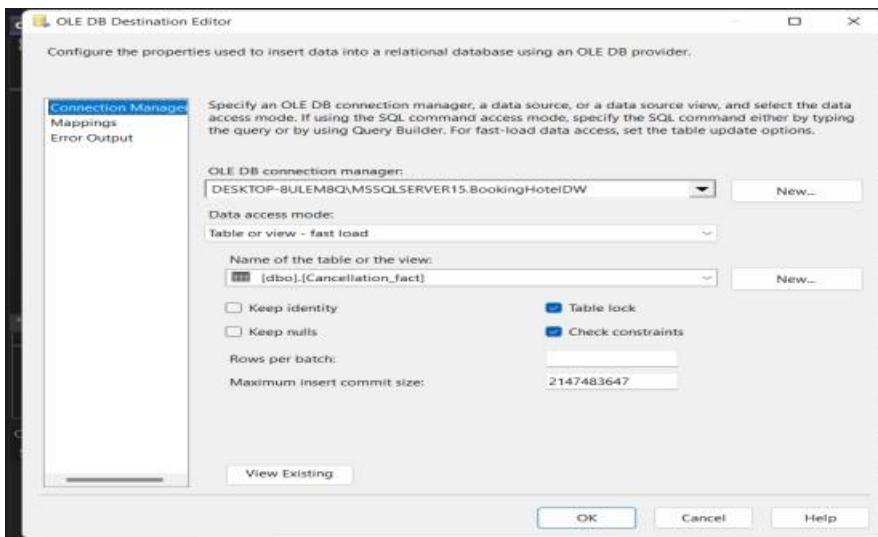
Từ **SSIS Toolbox**, kéo rê **Source Assistant** vào design surface, một dialog mở ra, Chọn source Type: **SQL Server** và **BookingHotelStage** connection manager.



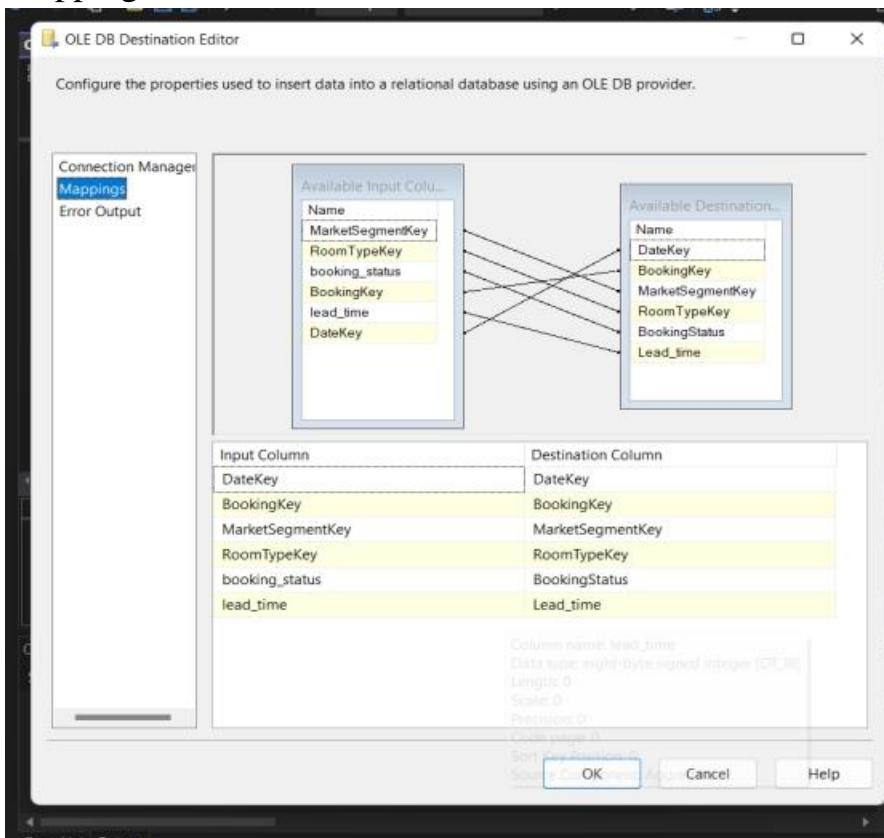
Cấu hình Aggregate chọn các thuộc tính cần thiết và các measures



Cấu hình Destination chọn bảng Cancellation_fact

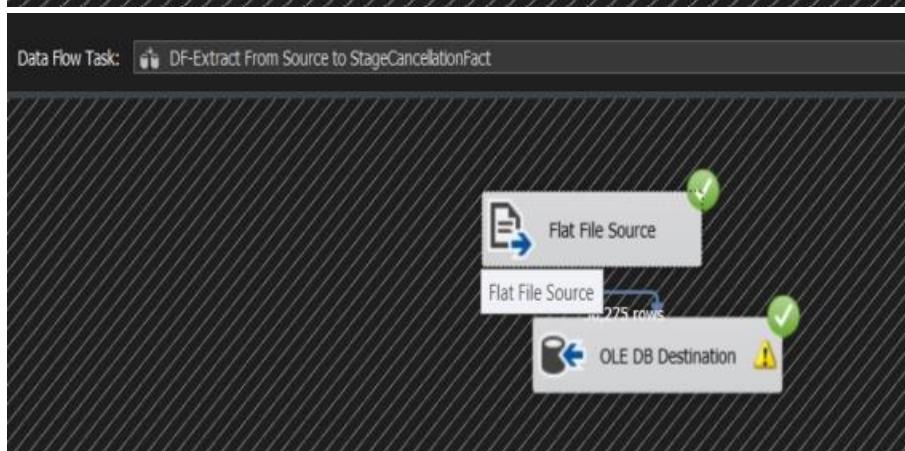
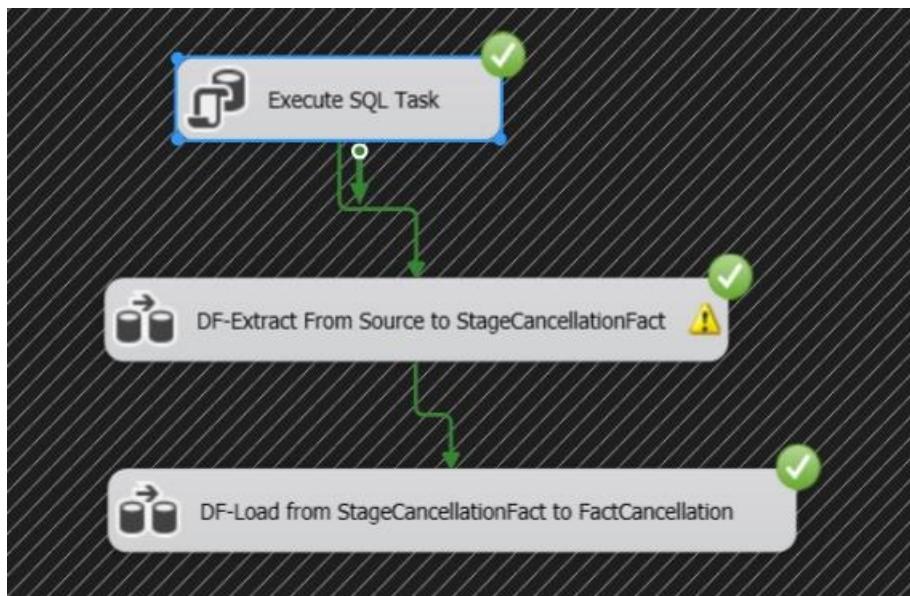


Mapping các cột lại với nhau.



3.4.2.3 Run

Bấm Start để chạy. Nếu xảy ra lỗi, cố gắng đọc hiểu thông báo lỗi để tìm cách sửa.
Nếu chạy thành công bạn sẽ thấy màn hình như sau:



Kiểm tra kết quả trong bảng **stgCancellationFact** trong cơ sở dữ liệu BookingHotelStage.

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. The title bar reads "SQLQuery16.sql - DESKTOP-8ULEM8Q\MSQLSERVER15.BookingHotelStage (DESKTOP-8ULEM8Q\YenNhi (80)) - Microsoft SQL Server Management Studio". The menu bar includes File, Edit, View, Query, Project, Tools, Window, Help. The toolbar has various icons for file operations like New Query, Import, Export, etc. The Object Explorer sidebar is visible on the left. The main area has three tabs at the top: "SQLQuery16.sql - ...LEM8Q\YenNhi (80)" (selected), "SQLQuery15.sql - ...LEM8Q\YenNhi (88)", and "SQLQuery14.sql - ...LEM8Q\YenNhi (63)". Below the tabs is a code editor window containing a T-SQL query:

```

SELECT TOP (1000) [arrival_year]
      ,[arrival_month]
      ,[arrival_date]
      ,[lead_time]
      ,[booking_status]
      ,[room_type_reserved]
      ,[market_segment_type]
      ,[Booking_ID]
  FROM [BookingHotelStage].[dbo].[stgCancellationFact]

```

Below the code editor is a results grid titled "Results". The grid displays the output of the query, showing columns: arrival_year, arrival_month, arrival_date, lead_time, booking_status, room_type_reserved, market_segment_type, and Booking_ID. The data consists of 25 rows, each representing a cancellation fact record.

	arrival_year	arrival_month	arrival_date	lead_time	booking_status	room_type_reserved	market_segment_type	Booking_ID
1	2017	10	2	224	Not_Canceled	Room_Type 1	Offline	INN00001
2	2018	11	6	5	Not_Canceled	Room_Type 1	Online	INN00002
3	2018	2	28	1	Canceled	Room_Type 1	Online	INN00003
4	2018	5	20	211	Canceled	Room_Type 1	Online	INN00004
5	2018	4	11	48	Canceled	Room_Type 1	Online	INN00005
6	2018	9	13	346	Canceled	Room_Type 1	Online	INN00006
7	2017	10	15	34	Not_Canceled	Room_Type 1	Online	INN00007
8	2018	12	26	83	Not_Canceled	Room_Type 4	Online	INN00008
9	2018	7	6	121	Not_Canceled	Room_Type 1	Offline	INN00009
10	2018	10	18	44	Not_Canceled	Room_Type 4	Online	INN00010
11	2018	9	11	0	Not_Canceled	Room_Type 1	Online	INN00011
12	2018	4	30	35	Not_Canceled	Room_Type 4	Online	INN00012
13	2018	11	28	30	Canceled	Room_Type 1	Online	INN00013
14	2018	11	20	95	Canceled	Room_Type 1	Online	INN00014
15	2017	10	20	47	Not_Canceled	Room_Type 1	Online	INN00015
16	2018	6	15	256	Canceled	Room_Type 1	Online	INN00016
17	2017	10	5	0	Not_Canceled	Room_Type 1	Offline	INN00017
18	2017	8	10	1	Not_Canceled	Room_Type 1	Online	INN00018
19	2017	10	30	99	Canceled	Room_Type 1	Online	INN00019
20	2017	10	4	12	Not_Canceled	Room_Type 1	Offline	INN00020
21	2017	10	30	99	Canceled	Room_Type 1	Online	INN00021
22	2018	11	25	122	Not_Canceled	Room_Type 1	Corporate	INN00022
23	2018	3	20	2	Not_Canceled	Room_Type 1	Offline	INN00023
24	2018	10	13	37	Not_Canceled	Room_Type 1	Offline	INN00024
25	2018	5	22	120	Not_Canceled	Room_Type 1	Online	INN00025

Kiểm tra kết quả trong bảng **Cancellation_fact** trong cơ sở dữ liệu BookingHotelDW.

The screenshot shows a Microsoft SQL Server Management Studio window with multiple tabs open. The active tab displays a query result set for a fact table named 'Cancellation_fact' from the 'BookingHotelDW' database. The results show 28 rows of data with columns: DateKey, BookingKey, MarketSegmentKey, BookingStatus, LeadTime, and RoomTypeKey.

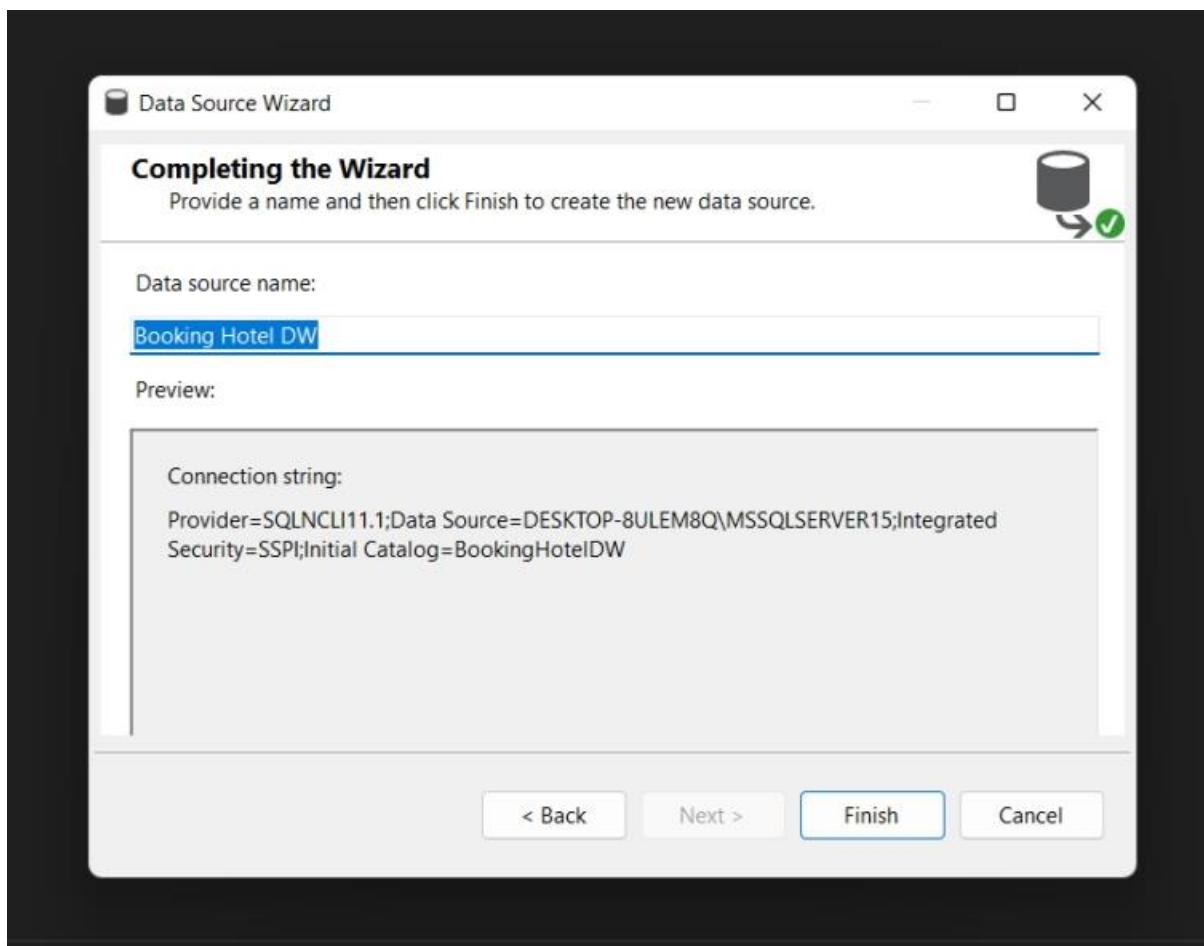
DateKey	BookingKey	MarketSegmentKey	BookingStatus	LeadTime	RoomTypeKey
1	20191010 - 18623	1	Canceled	33	4
2	20191007 - 19267	1	Canceled	13	6
3	20191006 - 24688	1	Not_Canceled	25	5
4	20191004 - 18747	1	Canceled	179	5
5	20190926 - 7288	1	Canceled	213	6
6	20190918 - 18713	1	Canceled	14	6
7	20190914 - 23518	1	Canceled	162	5
8	20190909 - 28929	1	Not_Canceled	20	4
9	20191018 - 23515	1	Canceled	66	5
10	20191014 - 1132	1	Not_Canceled	49	5
11	20190806 - 12079	1	Not_Canceled	35	6
12	20181110 - 23841	4	Canceled	305	5
13	20190907 - 28908	1	Not_Canceled	125	4
14	20190929 - 19708	1	Canceled	132	5
15	20191029 - 19481	1	Canceled	202	4
16	20190602 - 25962	4	Not_Canceled	56	5
17	20190912 - 13074	4	Not_Canceled	48	4
18	20190907 - 17462	4	Not_Canceled	203	5
19	20181030 - 2118	4	Not_Canceled	44	5
20	20190626 - 32777	1	Not_Canceled	43	5
21	20191021 - 17744	4	Not_Canceled	79	5
22	20191022 - 3827	1	Canceled	93	4
23	20190914 - 12077	1	Not_Canceled	58	4
24	20190302 - 32303	1	Not_Canceled	30	5
25	20171010 - 18803	1	Not_Canceled	63	5
26	20181008 - 14476	2	Not_Canceled	14	5
27	20190326 - 1204	1	Not_Canceled	130	3
28	20190617 - 19040	4	Canceled	245	5

CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS)

4.1. Quá trình xây dựng mô hình

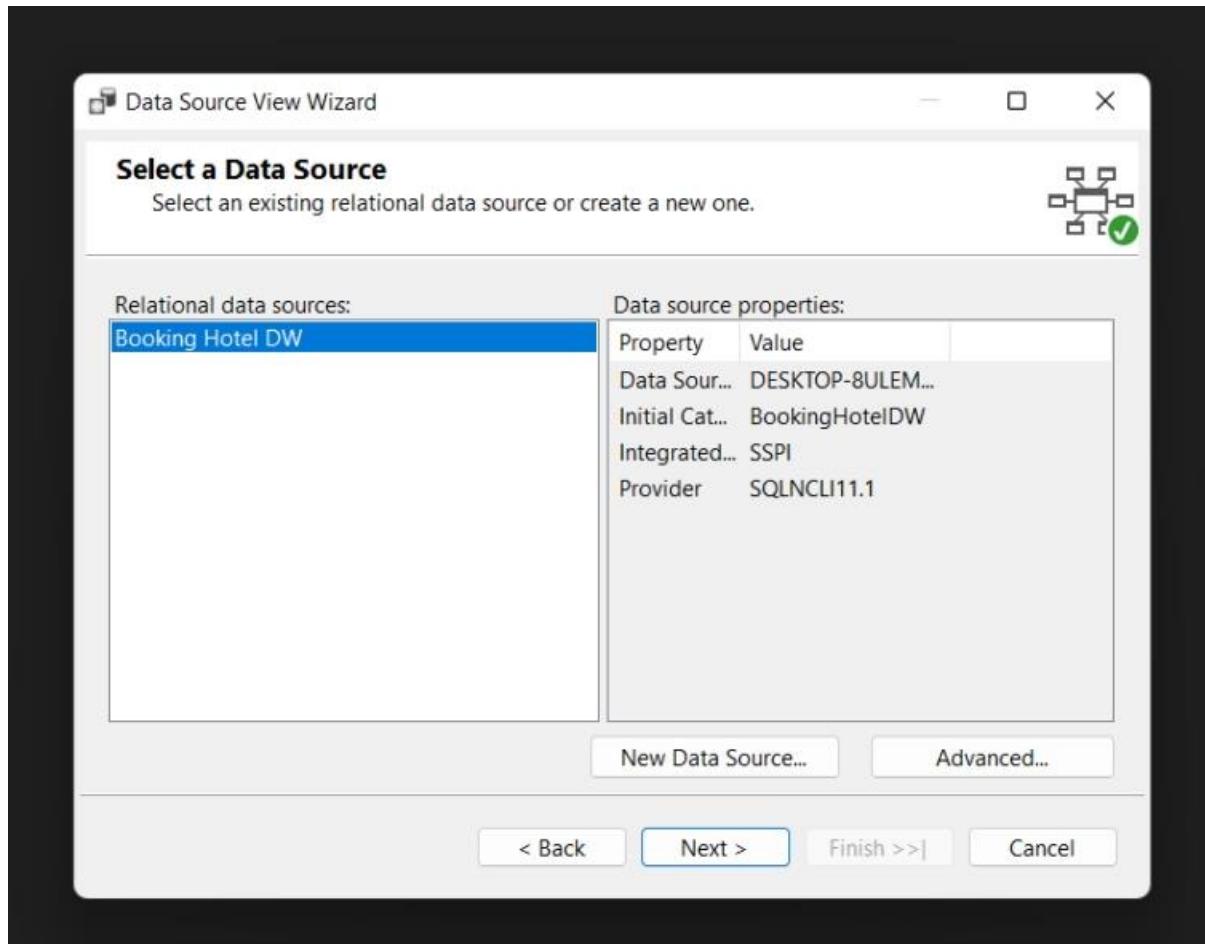
4.1.1. Tạo Data Source

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

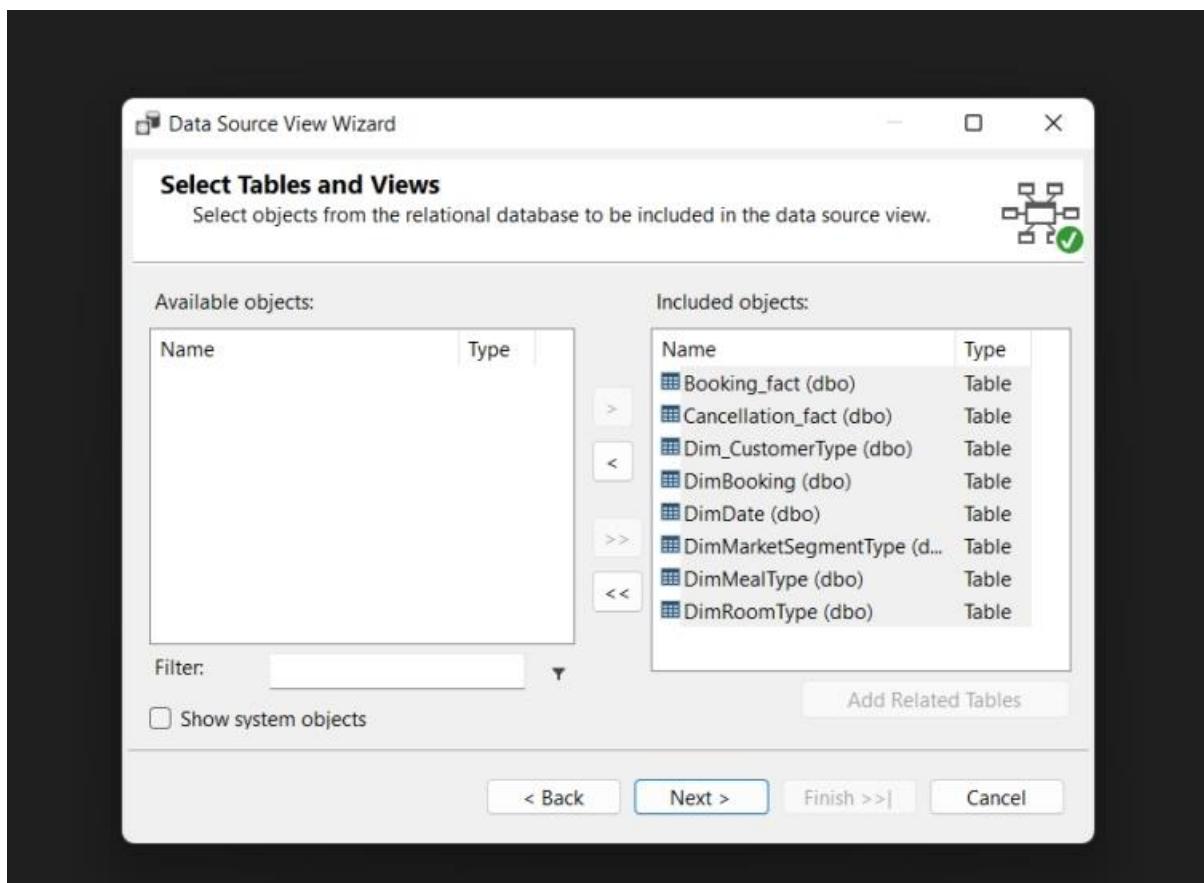


4.1.2. Tạo Data Source View

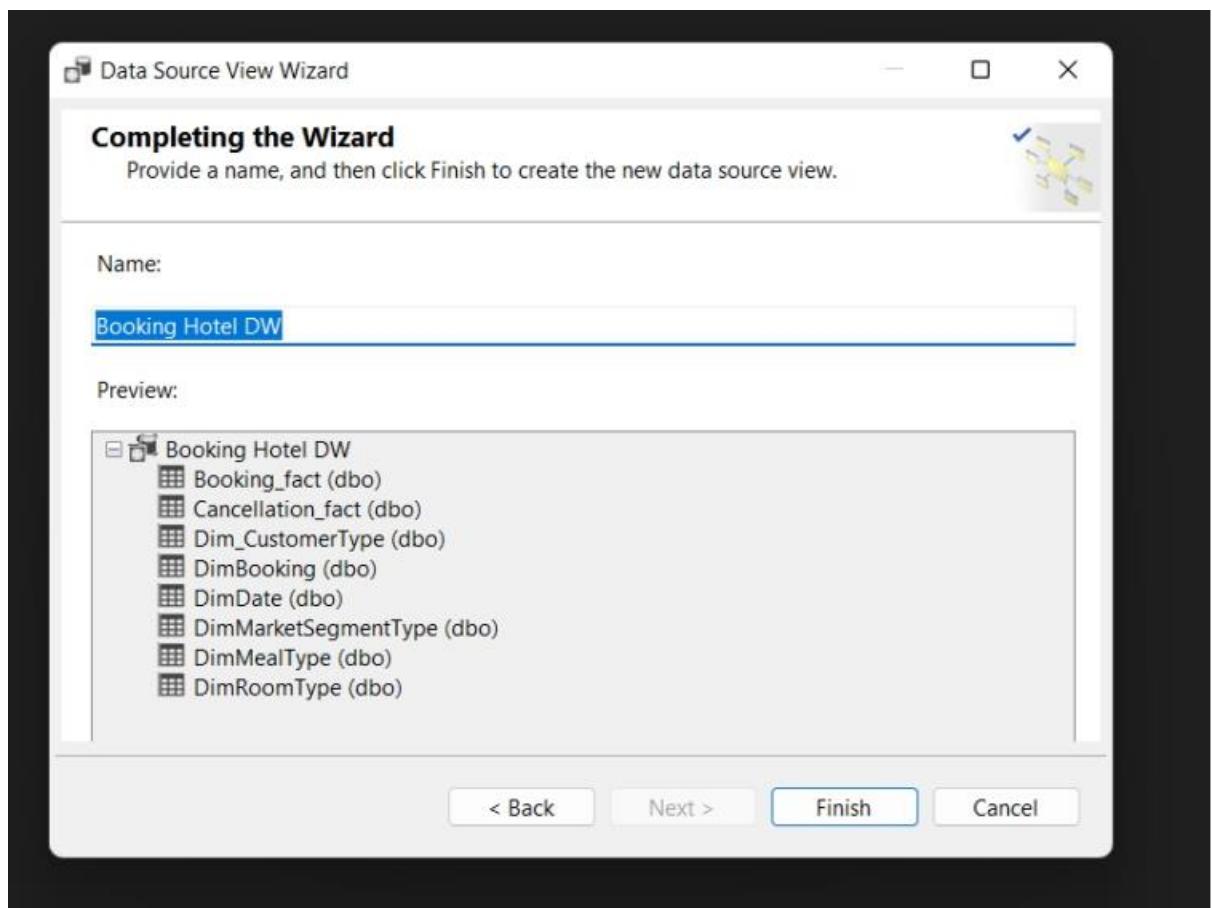
- Right-click vào Data Source Views ở khung Solution Explorer. Chon New Data Source View để tạo Data Source View mới. Chọn Data Soure là **Booking Hotel DW** và chọn Next.



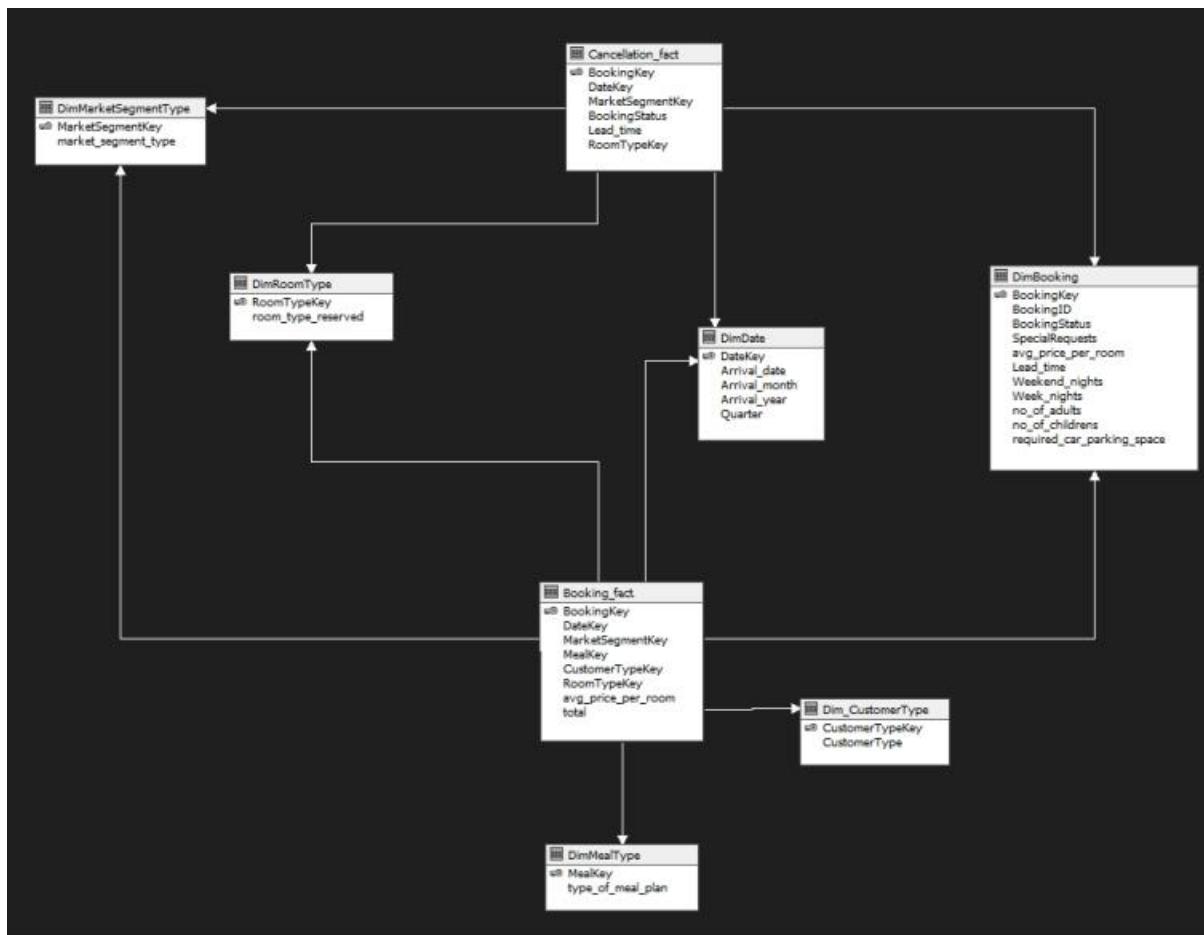
- Trong mục **Included objects**, chọn các bảng Dim và Fact cần thiết, nhấn Next



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

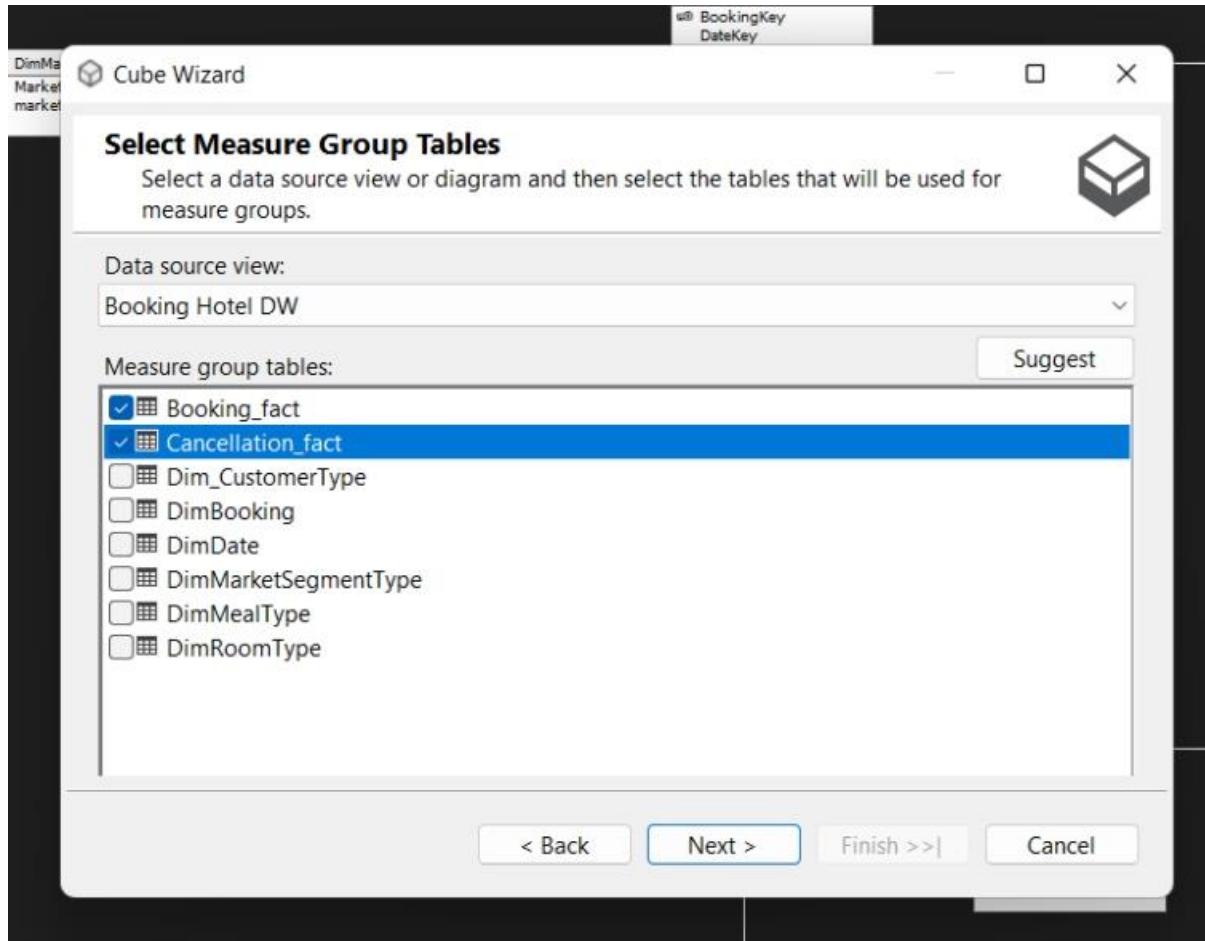


- Double-Click vào Data Source View vừa tạo để hiển thị diagram

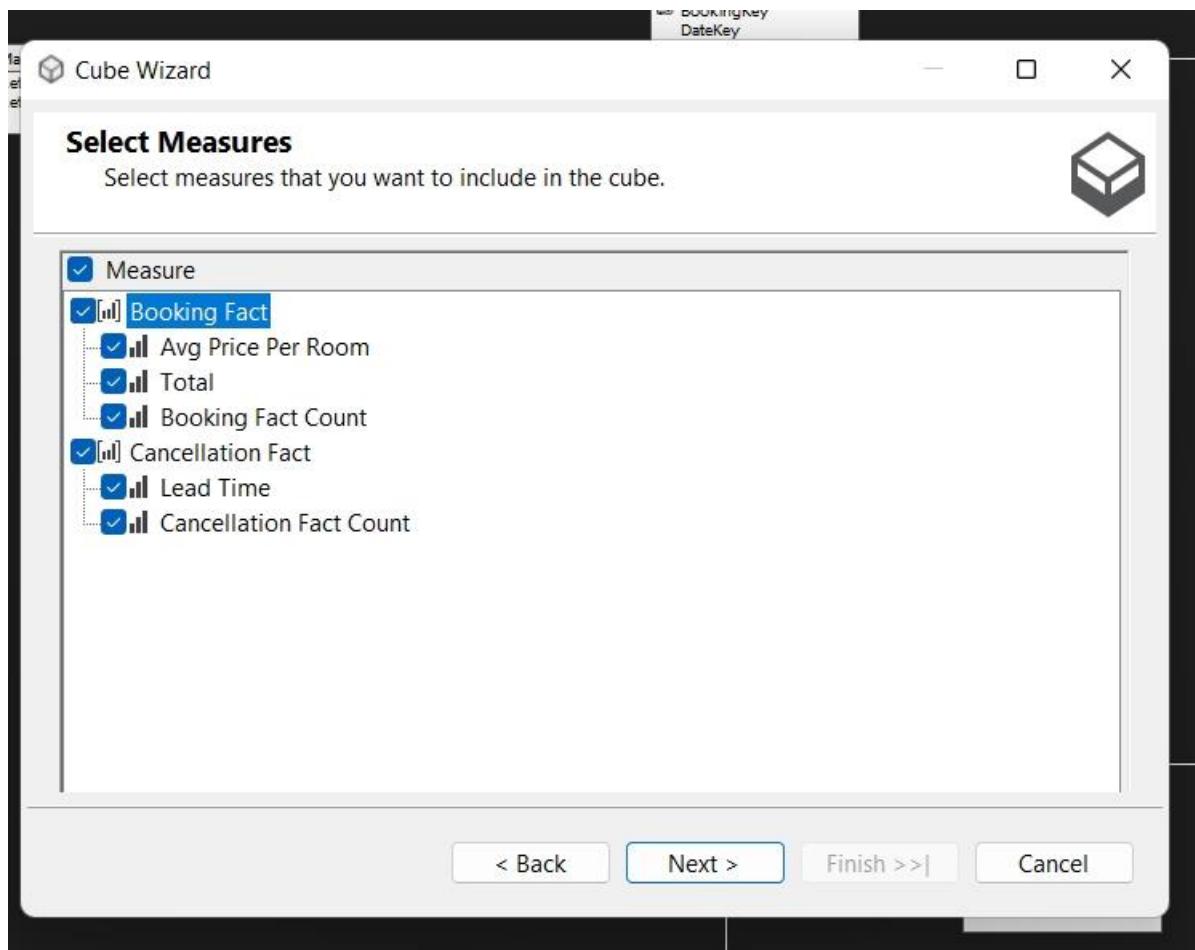


4.2. Quá trình xây dựng khối Cube

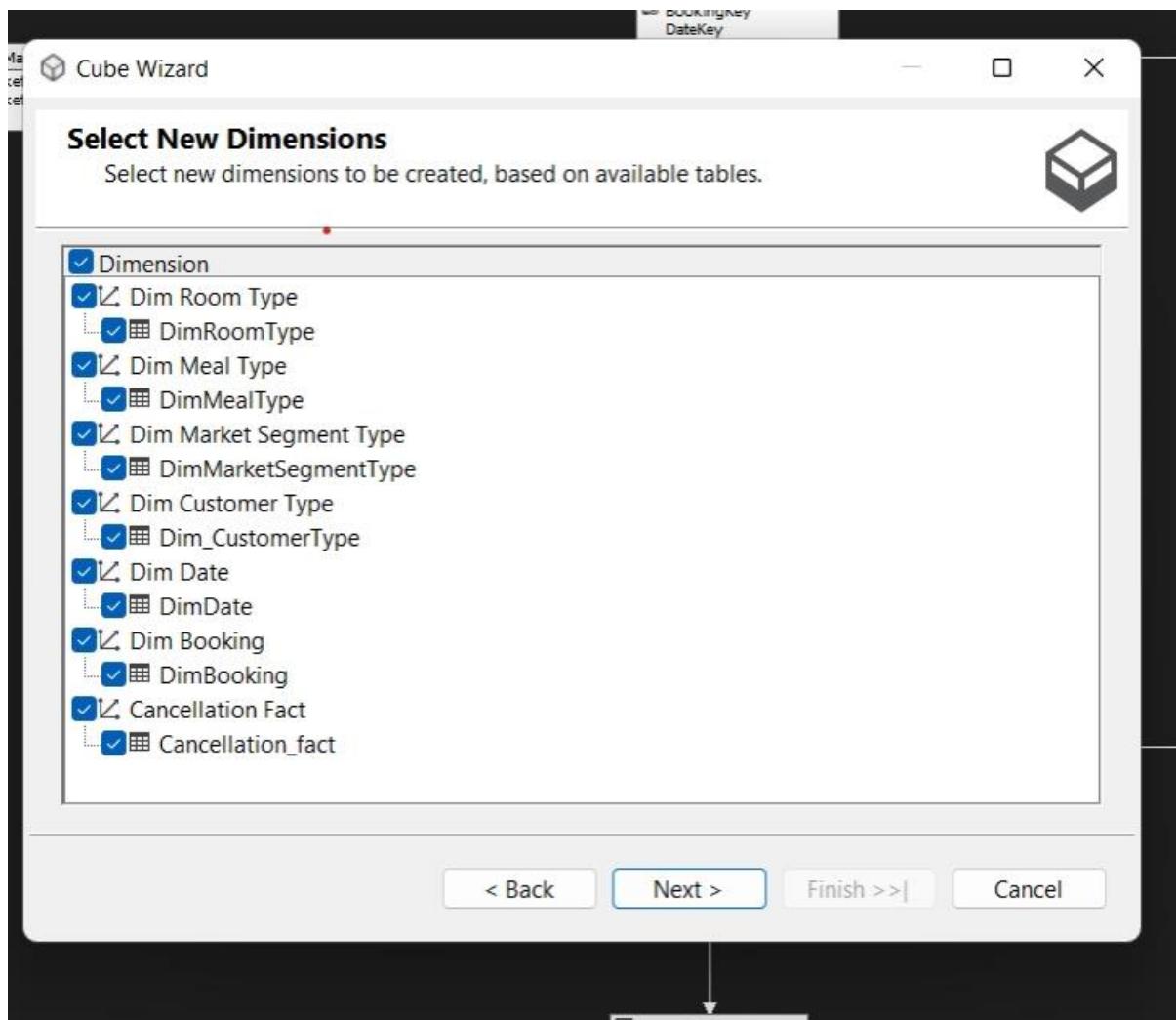
- Right-click vào **Cubes** ở khung Solution Explorer. Chon New Cube để tạo Cube mới.
- Chọn Data Source View là **Booking Hotel DW**. Tích chọn 2 bảng **Booking_fact** và **Cancellation_fact**, nhấn Next



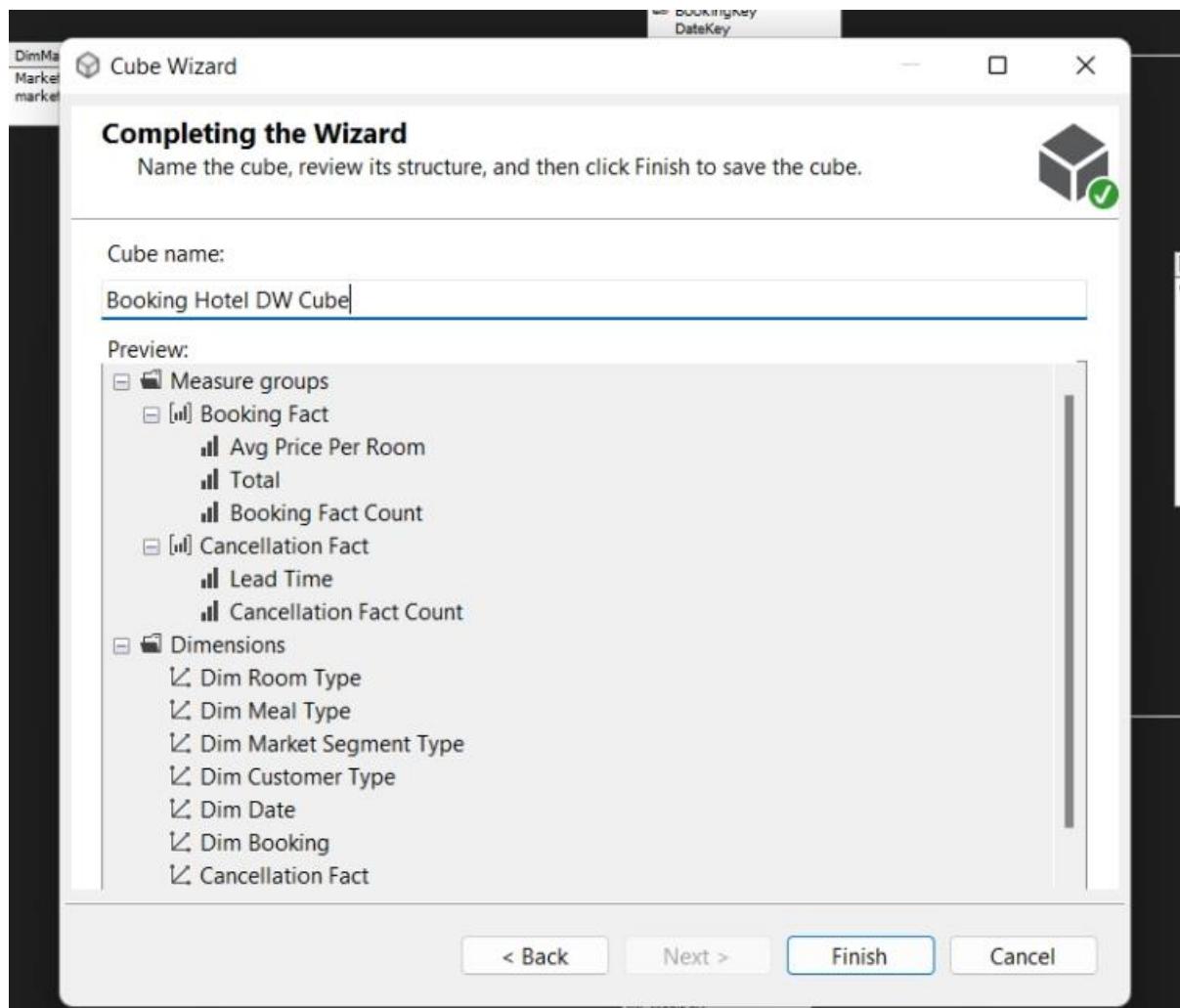
- Ở dialog Select Measures, tích chọn những measure cần thiết các mục sau



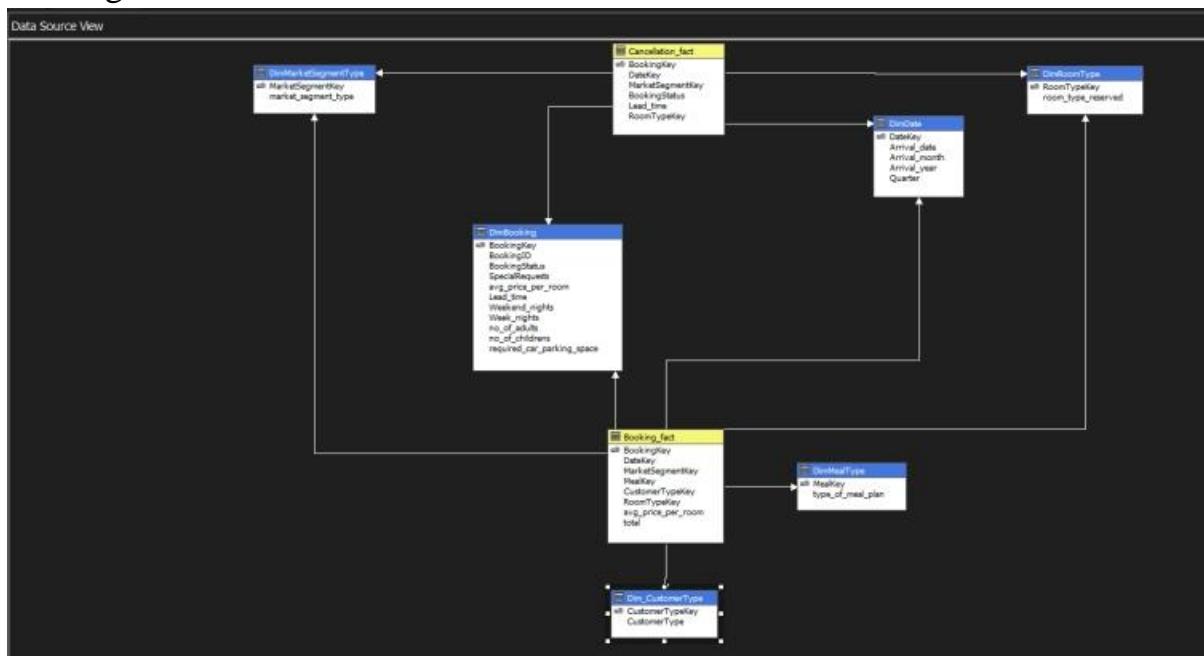
- Khởi tạo các bảng Dimension



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



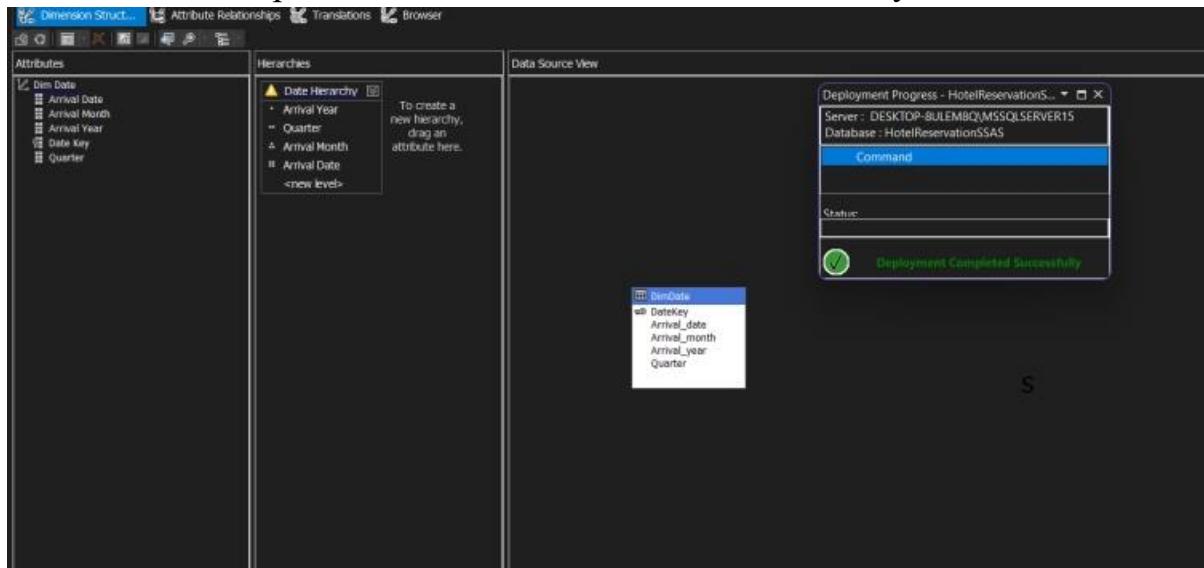
- Các bảng Fact, Dimension và quan hệ giữa chúng sẽ được thể hiện trong Diagram



4.3. Cấu hình Hierarchy

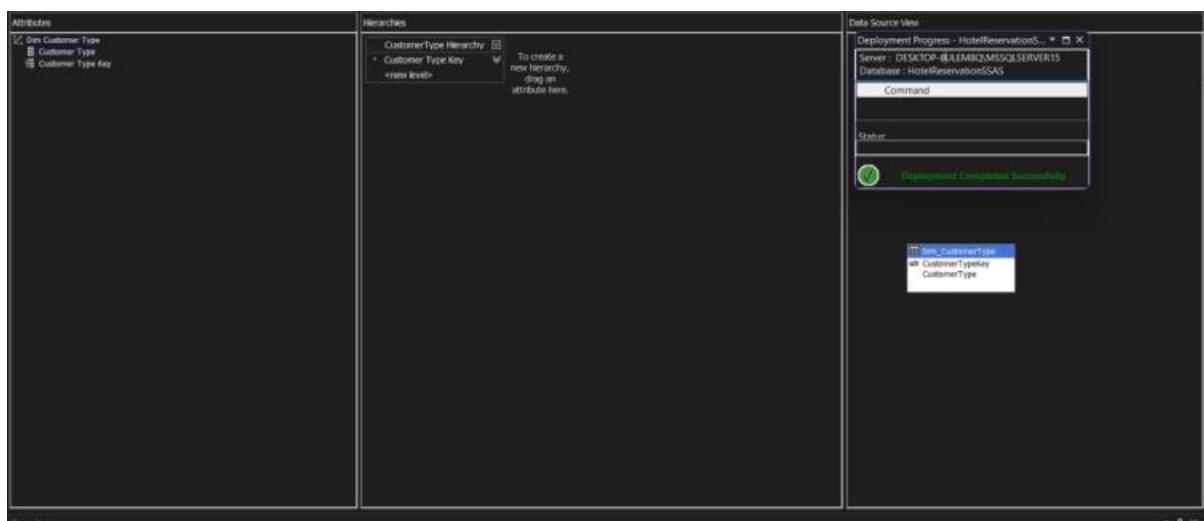
4.3.1. Tạo Hierarchy cho Dim Date

- Thêm thuộc tính vào bảng Dim: Tại table DimDate trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.
- Kéo thả 4 thuộc tính *Arrival_Year*, *Quarter*, *Arrival_Month*, *Arrival_Date* từ Attributes vào panel Hierarchies để tạo ra Date Hierarchy



4.3.2. Tạo Hierarchy cho Dim Customer Type

- Thêm thuộc tính vào bảng Dim: Tại table DimCustomerType trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.
- Kéo thả thuộc tính Customer Type từ Attributes vào panel Hierarchies để tạo ra CustomerType Hierarchy



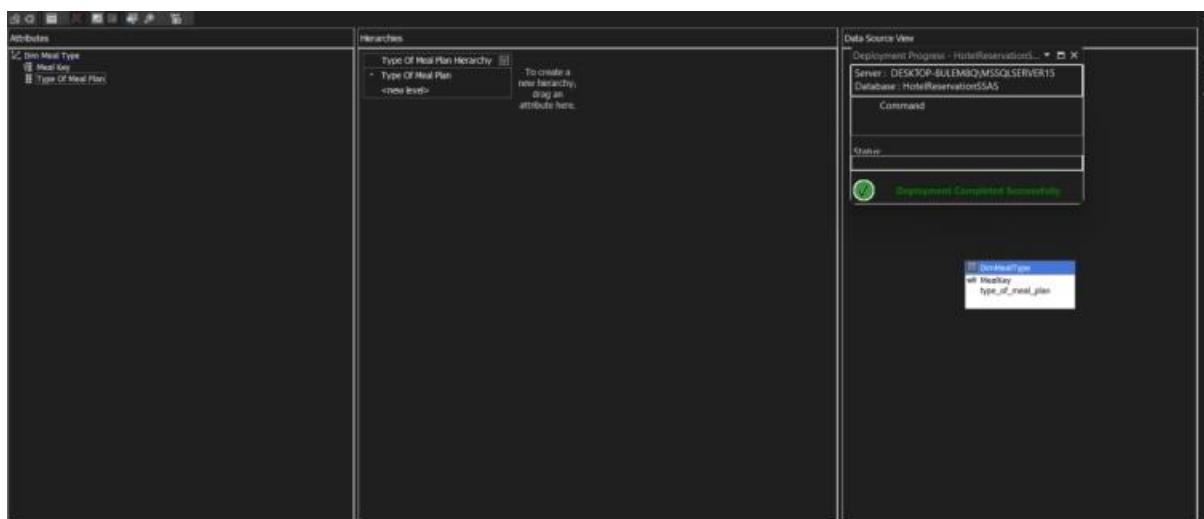
4.3.3. Tạo Hierarchy cho Dim Market Segment Type

- Thêm thuộc tính vào bảng Dim Market Segment Type: Tại table DimMarketSegmentType trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.
- Kéo thả thuộc tính Market Segment Type từ Attributes vào panel Hierarchies để tạo ra Market Segment Type Hierarchy



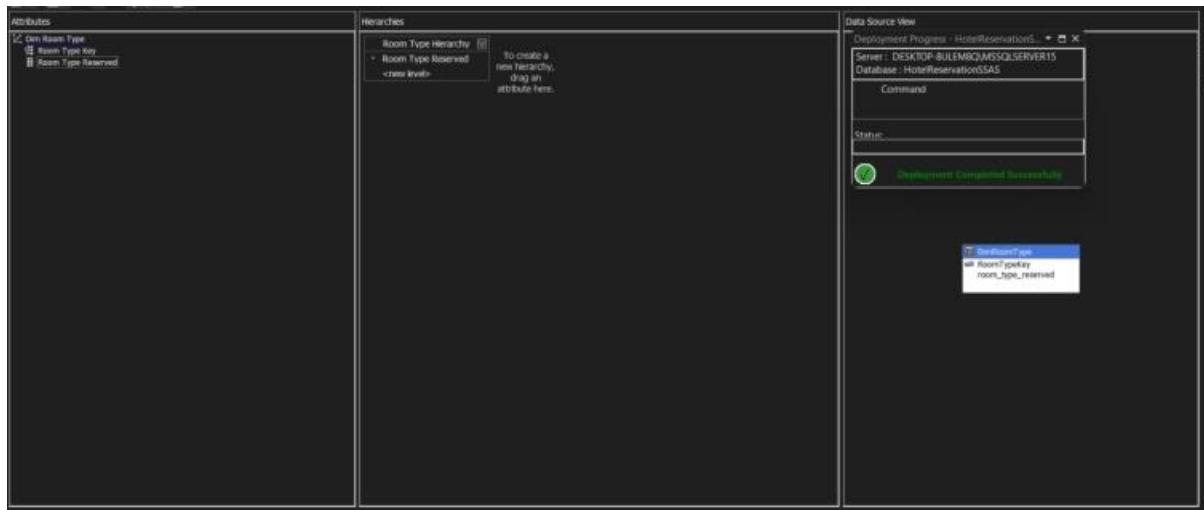
4.3.4. Tạo Hierarchy cho Dim Market Segment Type

- Thêm thuộc tính vào bảng Dim Meal Type: Tại table DimMealType trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.
- Kéo thả thuộc tính Type of Meal Plan từ Attributes vào panel Hierarchies để tạo ra Type of Meal Plan Hierarchy



4.3.5. Tạo Hierarchy cho Dim Room Type

- Thêm thuộc tính vào bảng Dim Room Type: Tại table DimRoomType trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.
- Kéo thả thuộc tính Room Type Reserved từ Attributes vào panel Hierarchies để tạo ra Room Type Hierarchy



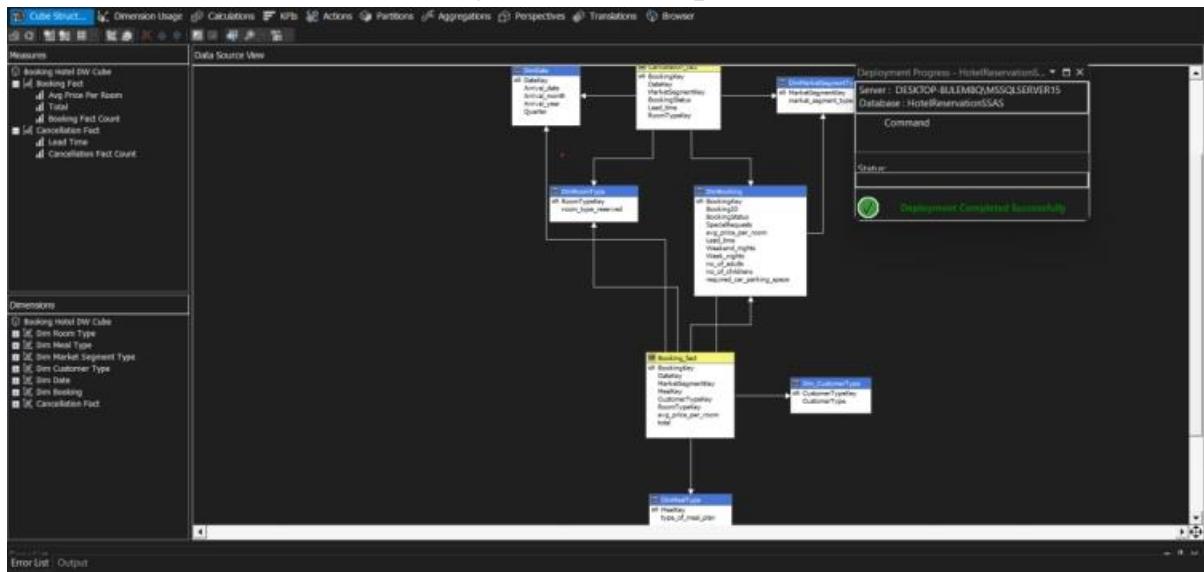
4.3.4. Tạo Hierarchy cho DimBooking

- Thêm thuộc tính vào bảng Dim Booking: Tại table DimBooking trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.
- Kéo thả thuộc tính Booking Status từ Attributes vào panel Hierarchies để tạo ra Booking Hierarchy



4.4. Deploy Process Cube

Sau khi tạo xong các Hierarchy thì tiến hành deploy process Booking Hotel DW Cube.cube để load các hierarchy và tiến hành phân tích.

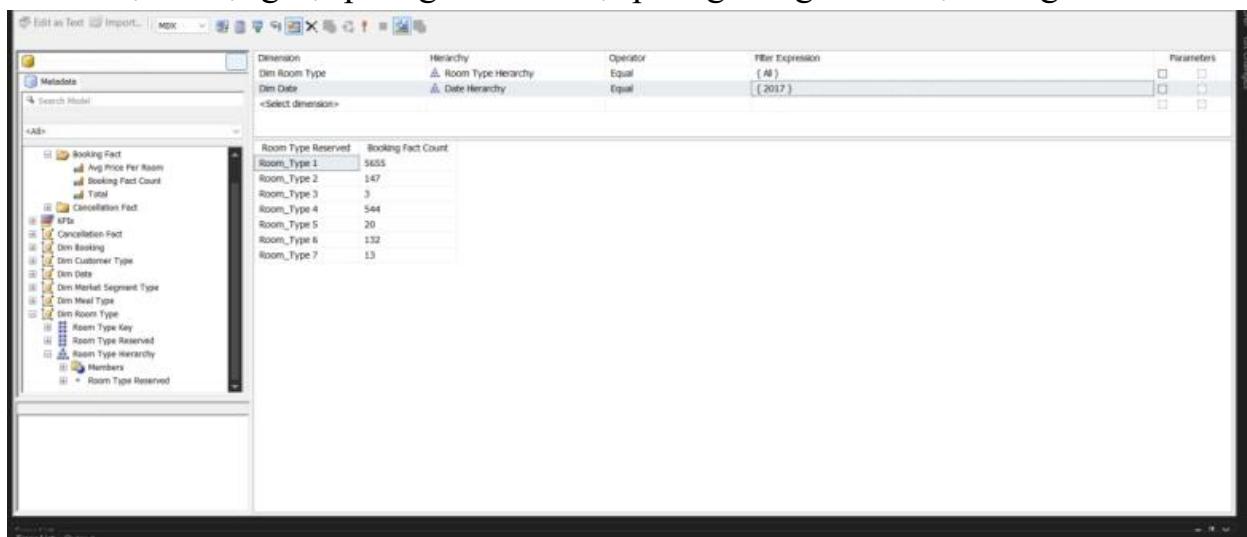


4.5. Thực hiện phân tích dữ liệu

4.5.1. Cho biết loại phòng được đặt nhiều nhất và doanh thu từ việc đặt phòng theo tháng quý năm

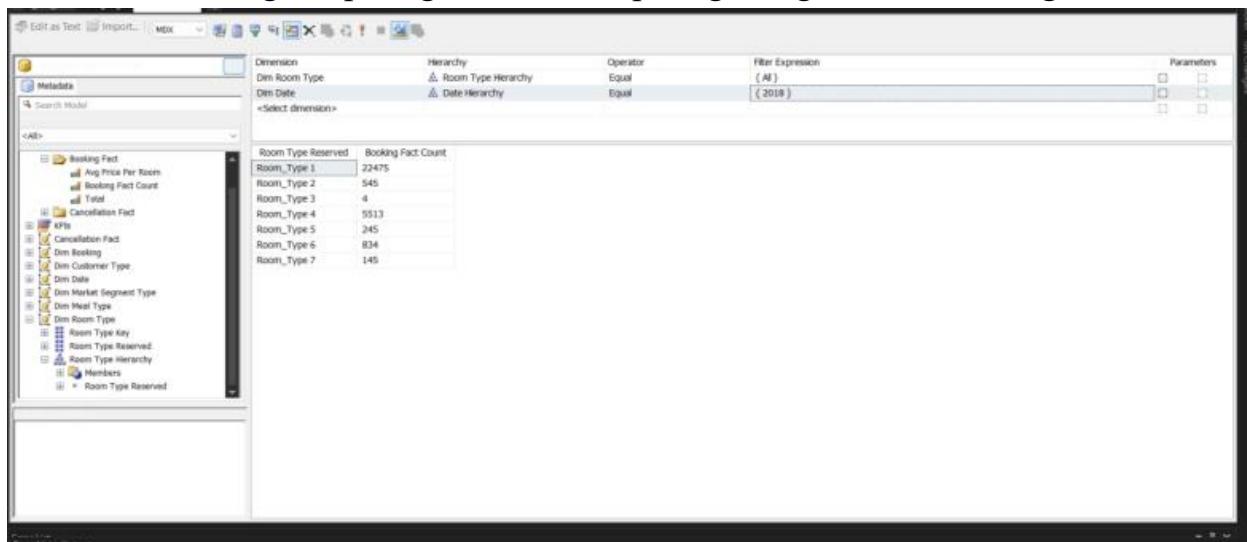
4.5.1.1. Sử dụng công cụ SSAS

-Hiển thị số lượng đặt phòng của các loại phòng trong khách sạn trong năm 2017



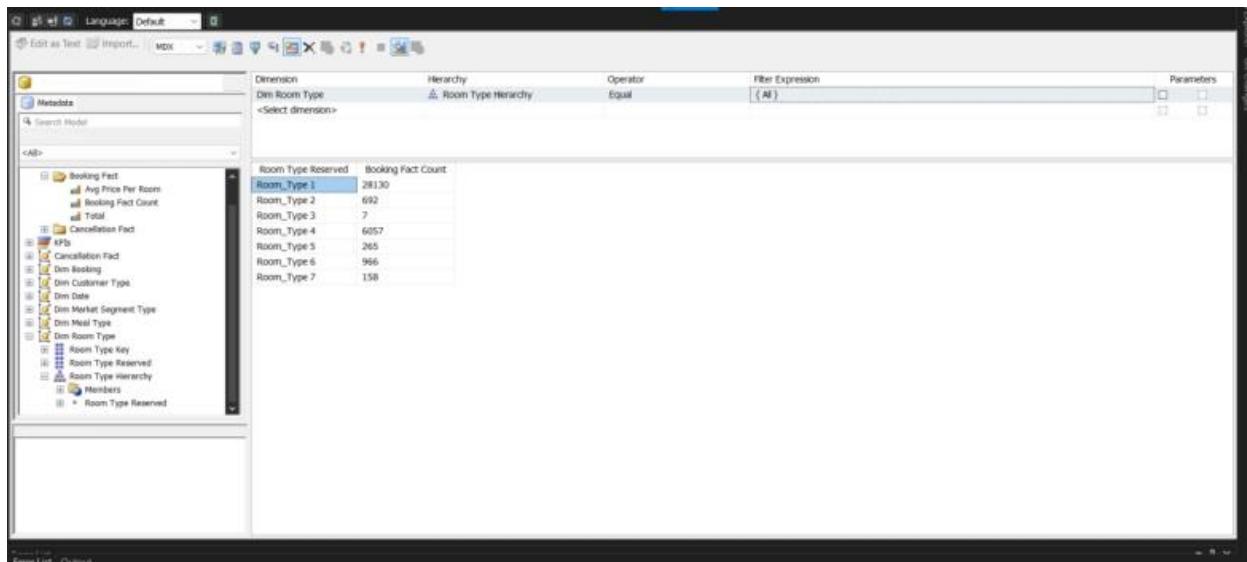
Nhận xét: Trong năm 2017 thì loại phòng được đặt nhiều nhất và chiếm số lượng áp đảo là Room_Type 1 với số lượng đặt là 5655 và đặt ít nhất là Room_Type 3 với số lượng là 3.

-Hiển thị số lượng đặt phòng của các loại phòng trong khách sạn trong năm 2018



Nhận xét: Trong năm 2018, loại phòng được đặt và chiếm số lượng áp đảo tiếp tục là loại phòng Room_Type 1 và loại phòng Room_Type 3 tiếp tục được đặt ít nhất

-Hiển thị số lượng đặt phòng của các loại phòng trong khách sạn trong cả hai năm 2017, 2018

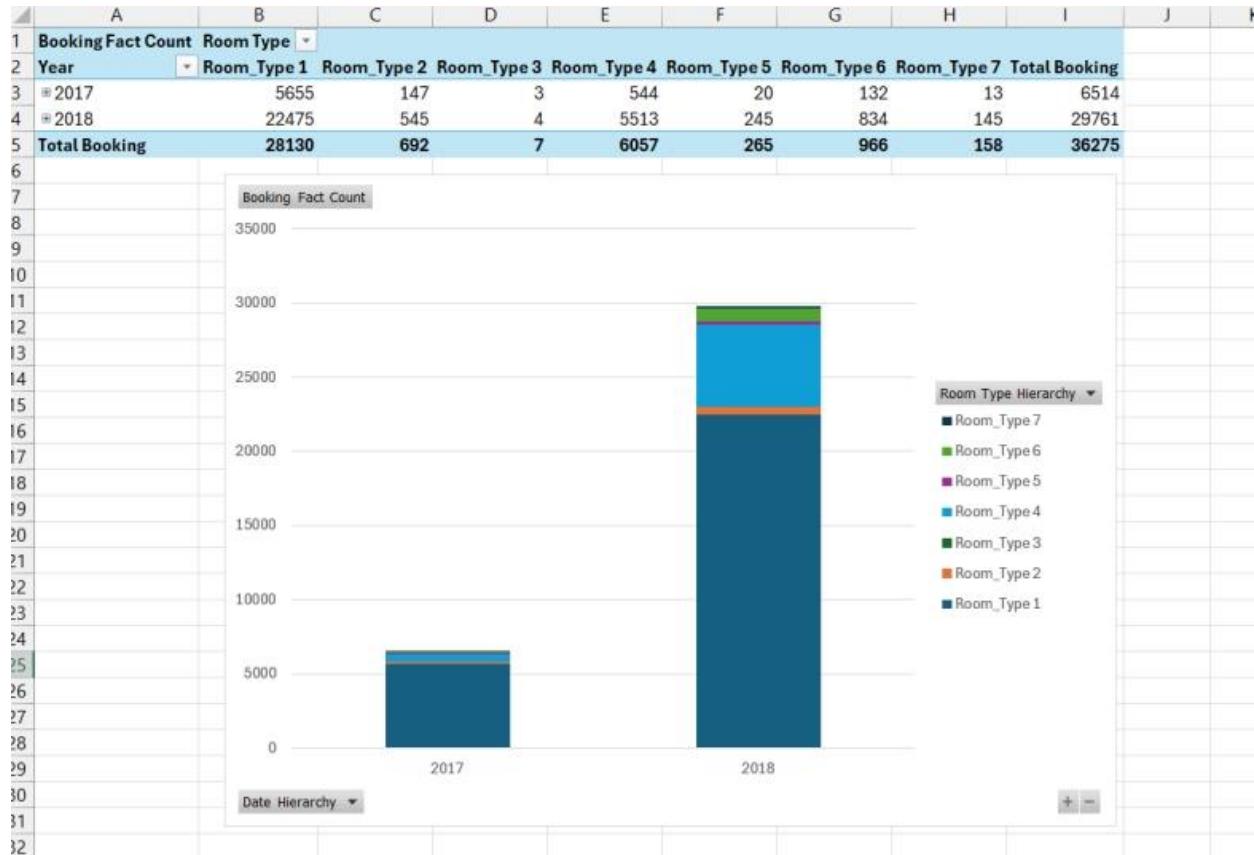


Nhận xét: Theo dữ liệu thống kê của cả 2 năm 2017, 2018 của khách sạn thì loại phòng được đặt nhiều nhất là Room_Type 1 với số lượng đặt là 28130, loại phòng có số lượng đặt thấp nhất là Room_Type 3

Qua những thống kê trên chủ doanh nghiệp có thể tiến hành thay đổi cải tạo để tăng số lượng phòng các loại như Room_Type 1, Room_Type 4 để đáp ứng đủ nhu cầu của khách hàng. Chủ doanh nghiệp cũng có thể đưa ra các chính sách ưu đãi của các loại phòng ít được đặt để thu hút khách hàng.

4.5.1.2. Sử dụng Pivot Table trong Excel

Hiển thị theo thời gian năm



Với hình trên ta thấy được số liệu về số lượng đặt phòng của mỗi loại phòng theo năm (2017, 2018) của khách sạn. Đồng thời biểu đồ dưới thể hiện trực quan kết quả thống kê khi chúng ta có thể thấy được loại phòng nào được đặt chiếm lượng áp đảo (Room_Type 1). Biểu đồ còn cho thấy sự tăng nhanh số lượng đặt phòng của khách sạn.

4.5.1.3. Sử dụng Power BI



Nhận xét: Biểu đồ trên thể hiện doanh thu từ việc đặt phòng từ các loại phòng và doanh thu thể hiện theo từng tháng.

- Với loại phòng, loại phòng với doanh thu nhiều nhất là loại phòng 1, và ít nhất là loại 3. Từ đây ta có thể đưa ra chiến lược tối ưu nhiều loại phòng 1 hơn để doanh thu đạt được nhiều nhất.

- Với doanh thu được thống kê theo tháng, ta thấy được thời gian mà khách sạn có doanh thu nhiều nhất nằm trong khoảng từ tháng 8 đến tháng 10. Trong những tháng này khách sạn có thể đưa ra chiến lược tăng thêm nhân sự để tối ưu đáp ứng được dịch vụ cho khách hàng.

4.5.2. Cho biết khách hàng thường chọn thực đơn nào nhiều nhất

4.5.2.1. Sử dụng công cụ SSAS

- Thống kê số lượng đặt của mỗi loại thực đơn

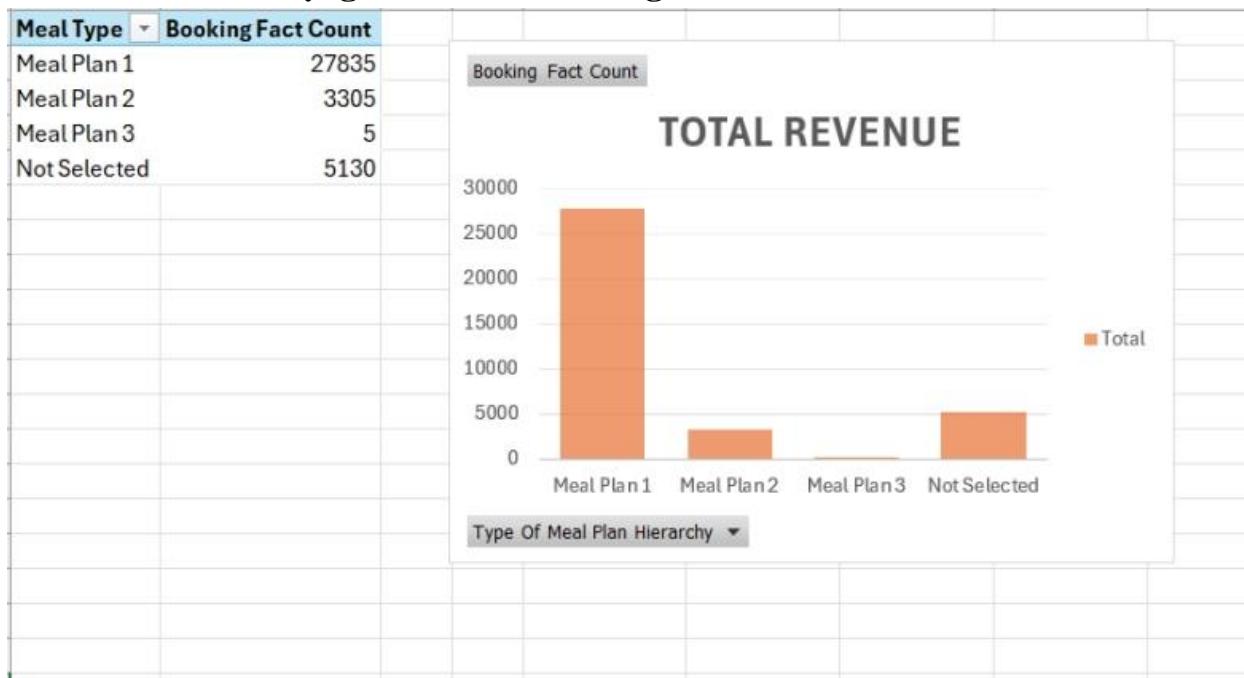
The screenshot shows the SSAS Dimension Manager window. On the left, there's a tree view of the dimension structure under 'Type Of Meal Plan'. The root node 'Total' has three children: 'Cancellation Fact', 'Meal Plan 1', and 'Not Selected'. 'Meal Plan 1' has three children: 'Meal Plan 1', 'Meal Plan 2', and 'Meal Plan 3'. On the right, a table titled 'Booking Fact Count' displays the count for each category:

Type Of Meal Plan	Booking Fact Count
Meal Plan 1	27835
Meal Plan 2	3305
Meal Plan 3	5
Not Selected	5130

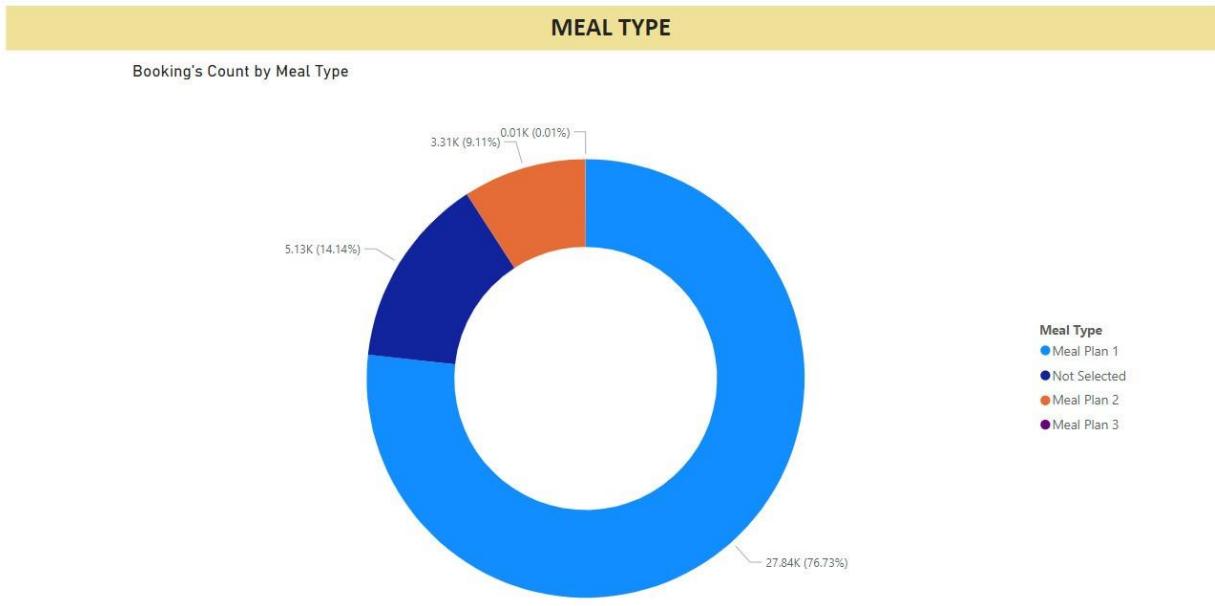
Nhận xét: Theo dữ liệu thống kê thì loại thực đơn được ưu chuộng nhất là Meal Plan 1 và ít nhất là Meal Plan 3 và có số lượng khá lớn khách hàng không lựa chọn thực đơn nào

Qua những thống kê trên khách sạn có thể tìm hiểu nguyên nhân tại sao thực đơn 3 ít được chọn. Thông qua đó, khách sạn có thể đưa ra quyết định tiếp tục giữ hay bỏ và tiến hành cải tiến thực đơn. Đồng thời khách sạn có thể đưa ra nhiều thực đơn khác với nhiều món, nhiều mức giá, thực hiện các yêu cầu bên ngoài của khách hàng về bữa ăn để có thể đa dạng sự lựa chọn cho khách hàng và có thể thu hút khách hàng chọn thực đơn để ăn tại khách sạn.

4.5.2.2. Sử dụng Pivot Table trong Excel



4.5.2.3. Sử dụng Power BI

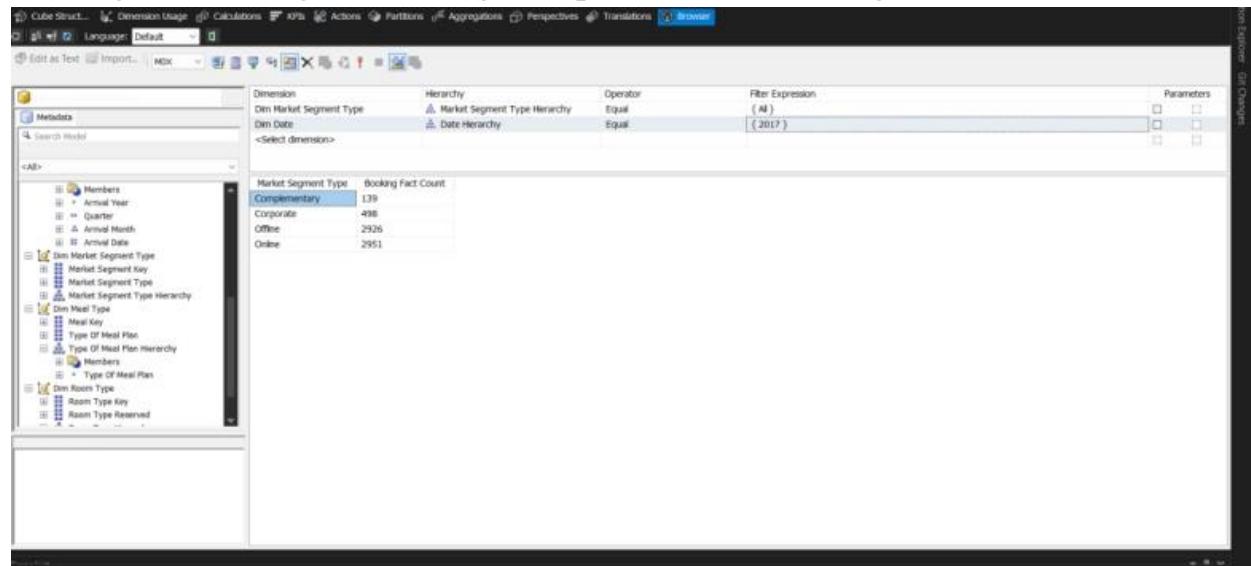


Nhận xét: Biểu đồ trên cho ta thấy được khẩu phần ăn 1 được khách hàng chọn nhiều nhất. Khách sạn có thể cân nhắc tăng nhiều khẩu phần 1 nhiều để đáp ứng đủ nhu cầu cho khách hàng

4.5.3. Thống kê số lượng khách hàng của mỗi loại phân khúc thị trường

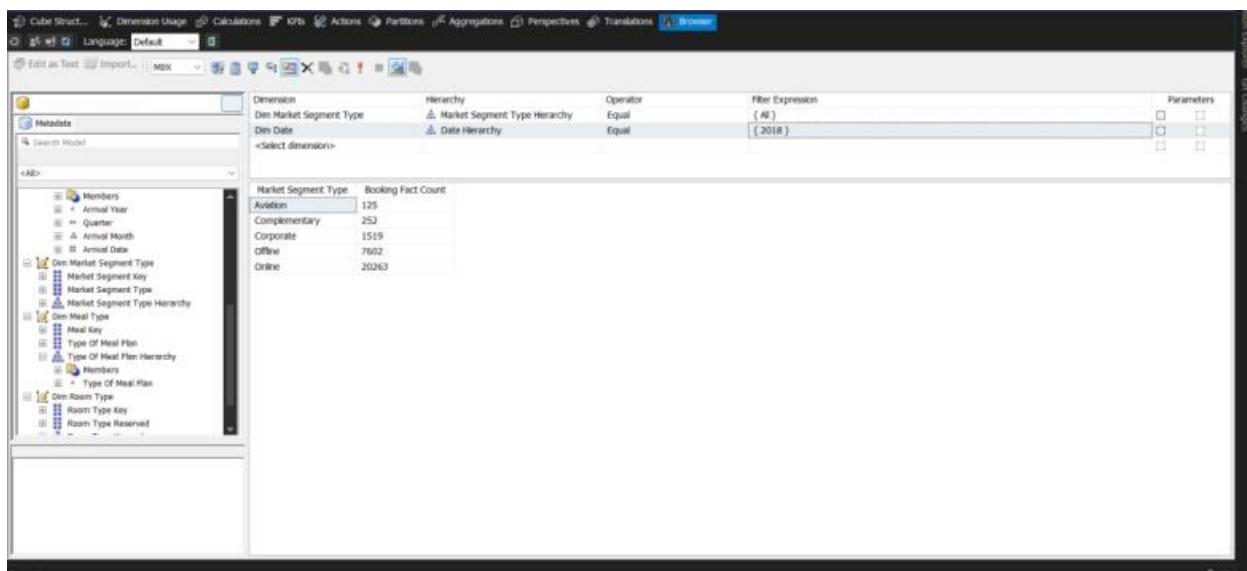
4.5.3.1. Sử dụng công cụ SSAS

+Thống kê số lượng khách hàng theo phân khúc thị trường năm 2017:



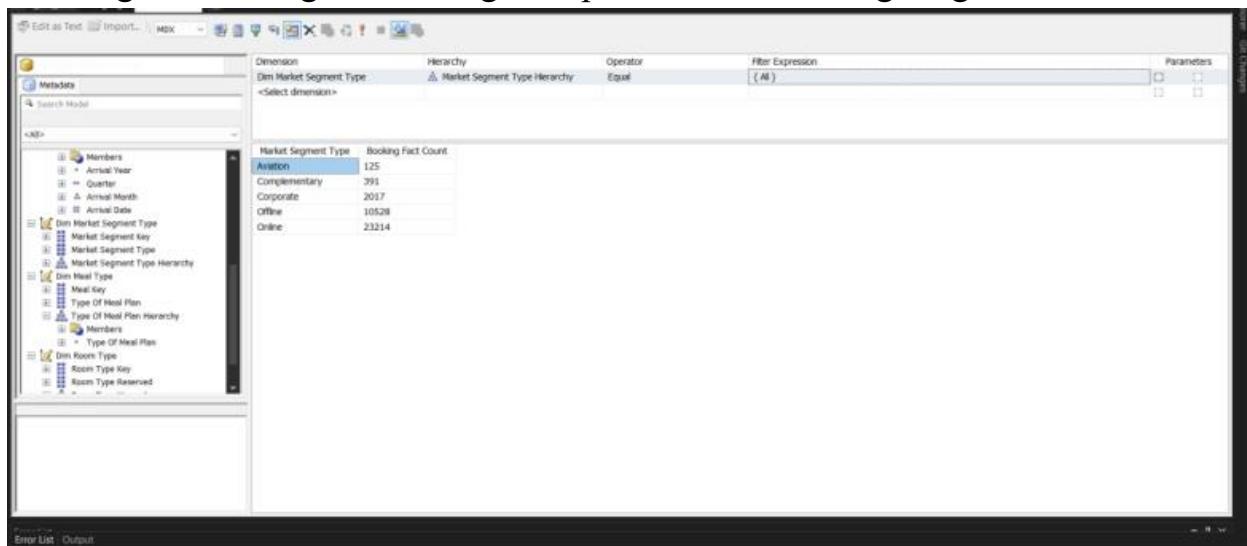
Nhận xét: Ở năm 2017 khi mới thành lập khách sạn chỉ có 4 phân khúc Online được đặt nhiều nhất nhưng có chênh lệch không đáng kể với phân khúc Offline. Tuy nhiên 2 phân khúc còn lại có tỉ lệ đặt rất thấp so với 2 phân khúc trên

+ Thống kê số lượng khách hàng theo phân khúc thị trường năm 2018

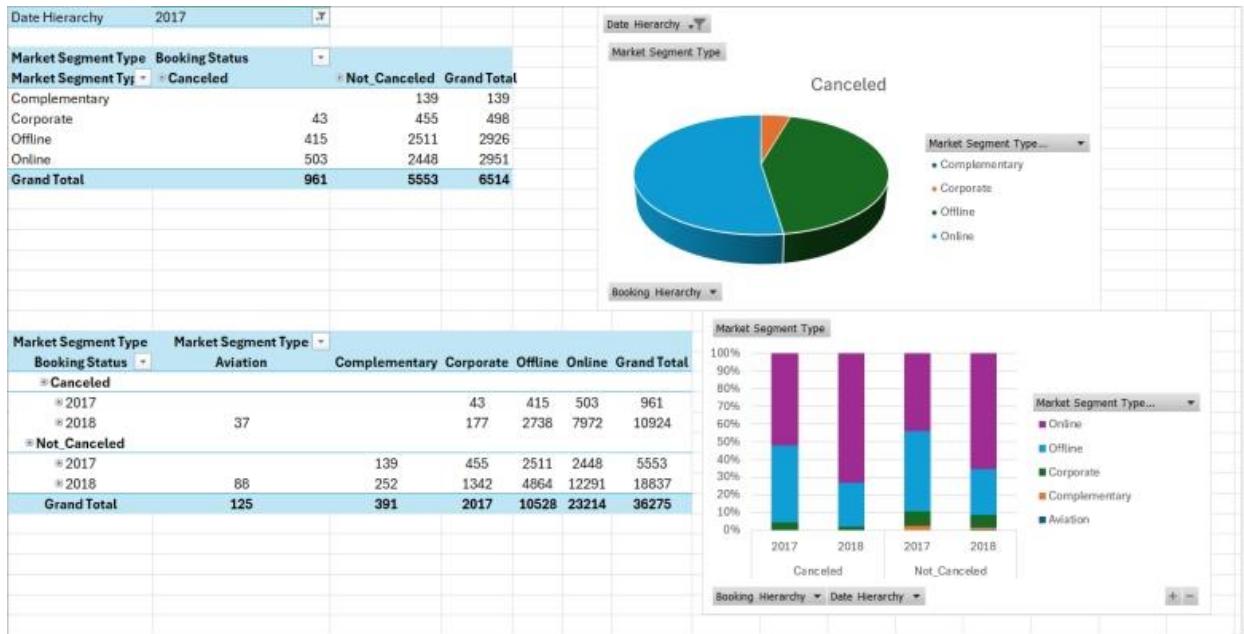


Nhận xét: Năm 2018, phân khúc online đã có sự phát triển rất lớn khi bỏ xa về số lượng đặt so với các phân khúc thị trường khác. Nguyên nhân dẫn đến việc này có thể là sự nổi tiếng của khách sạn và sự tiện lợi khi đặt online. Phân khúc Offline vẫn đứng thứ 2 nhưng không còn chiếm tỉ lệ nhiều như năm 2017. Ngoài ra ở năm này còn thêm 1 hình thức đặt phòng mới là Aviation tuy nhiên tỉ lệ đặt phòng thông qua nó rất thấp nguyên nhân có thể là do hình thức này còn mới chưa được biết đến rộng rãi. Hình thức Corporate có số lượng đặt phòng tăng lên nhưng vẫn chiếm tỉ lệ rất nhỏ.

+ Thống kê số lượng khách hàng theo phân khúc thị trường tổng cả hai năm



4.5.3.2. Sử dụng Pivot Table trong Excel



Qua hình trên, ngoài thống kê số lượt đặt phòng theo các phân khúc thị trường thì còn thống kê số lượng hủy và không hủy của mỗi phân khúc. Theo như hình trên thì tuy tỉ lệ đặt hàng theo hình thức online cao tuy nhiên tỉ lệ hủy của nó cũng cao không kém khi chiếm hơn 50% tổng số lượt hủy tiếp theo là hình thức

4.5.3.3. Sử dụng Power BI

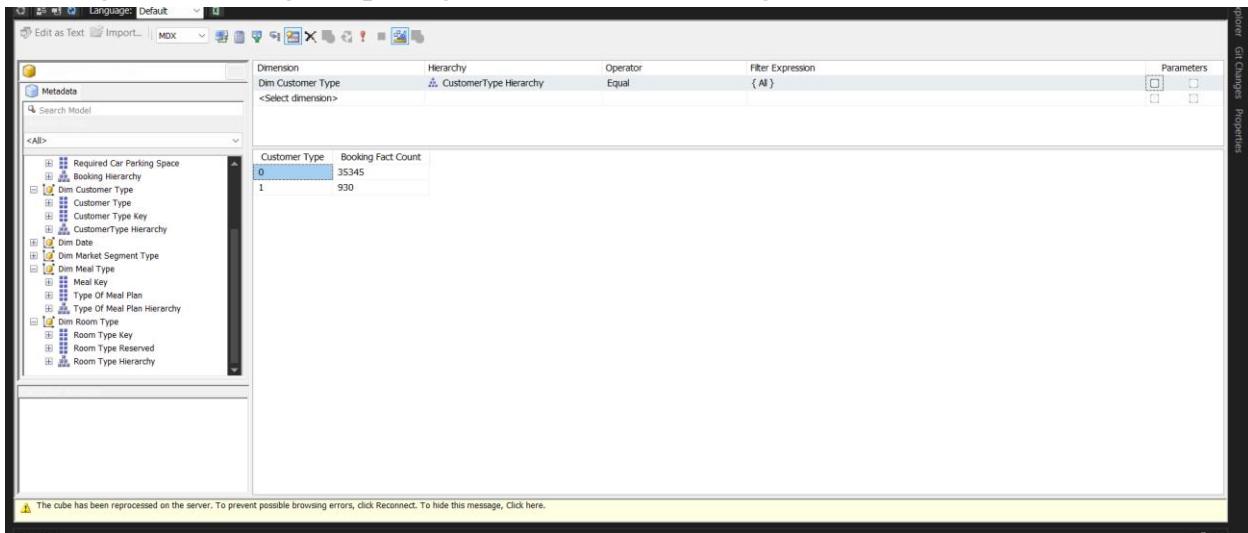


Nhận xét: Phân khúc thị trường nhiều nhất của khách sạn là đặt hàng qua online, khách sạn nên tối ưu những hệ thống đặt phòng trên online tốt nhất và có thêm nhiều nhân sự hỗ trợ khách hàng qua online nhiều nhất để đáp ứng trải nghiệm của khách hàng

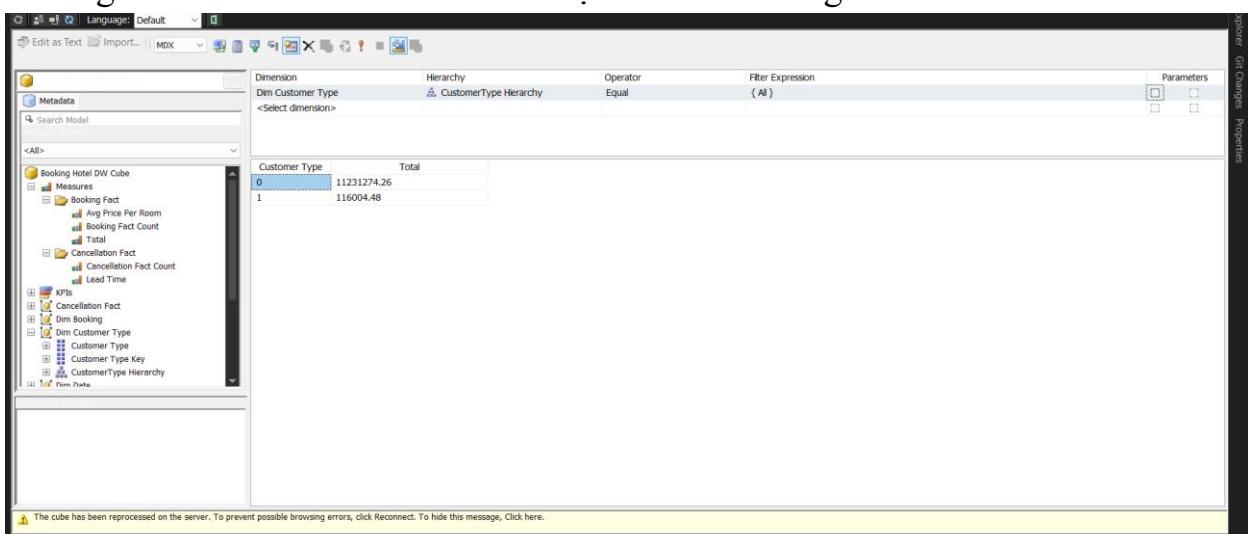
4.5.4. Thống kê doanh thu theo loại khách (repeated_guest)

4.5.4.1. Sử dụng công cụ SSAS

-Thống kê số lượng đặt phòng theo loại khách hàng năm 2017 và 2018



-Thống kê doanh thu theo loại khách hàng năm 2017 và 2018



-Thống kê doanh thu theo loại khách hàng năm 2017

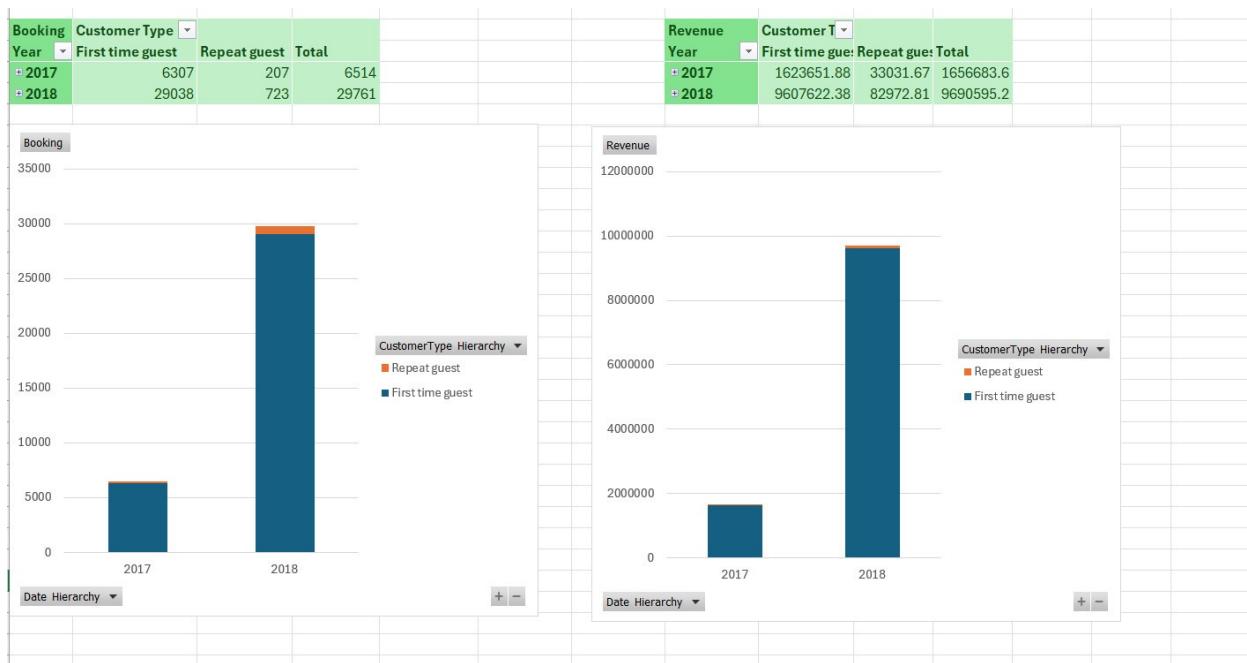
The screenshot shows the Analysis Services Management Studio interface. The main area displays a table titled 'Customer Type' with a single column 'Total' and two rows, 0 and 1. Row 0 has a value of 1623651.88 and row 1 has a value of 33031.67. The left sidebar shows the cube's structure with nodes like 'Booking Hotel DW Cube', 'Measures', 'Booking Fact', 'Cancellation Fact', 'KPIs', etc. At the top, there are tabs for 'Cube Struct...', 'Dimension Usage', 'Calculations', 'KPIs', 'Actions', 'Partitions', 'Aggregations', 'Perspectives', 'Translations', and 'Browser'. A message at the bottom says 'The cube has been reprocessed on the server. To prevent possible browsing errors, click Reconnect. To hide this message, Click here.'

-Thống kê doanh thu theo loại khách hàng năm 2018

This screenshot is similar to the one above, showing the Analysis Services Management Studio interface. The main area displays a table titled 'Customer Type' with a single column 'Total' and two rows, 0 and 1. Row 0 has a value of 9607622.38000004 and row 1 has a value of 82972.81. The left sidebar shows the cube's structure with nodes like 'Booking Hotel DW Cube', 'Measures', 'Booking Fact', 'Cancellation Fact', 'KPIs', etc. At the top, there are tabs for 'Cube Struct...', 'Dimension Usage', 'Calculations', 'KPIs', 'Actions', 'Partitions', 'Aggregations', 'Perspectives', 'Translations', and 'Browser'. A message at the bottom says 'The cube has been reprocessed on the server. To prevent possible browsing errors, click Reconnect. To hide this message, Click here.'

4.5.4.2.Sử dụng Pivot Table trong Excel

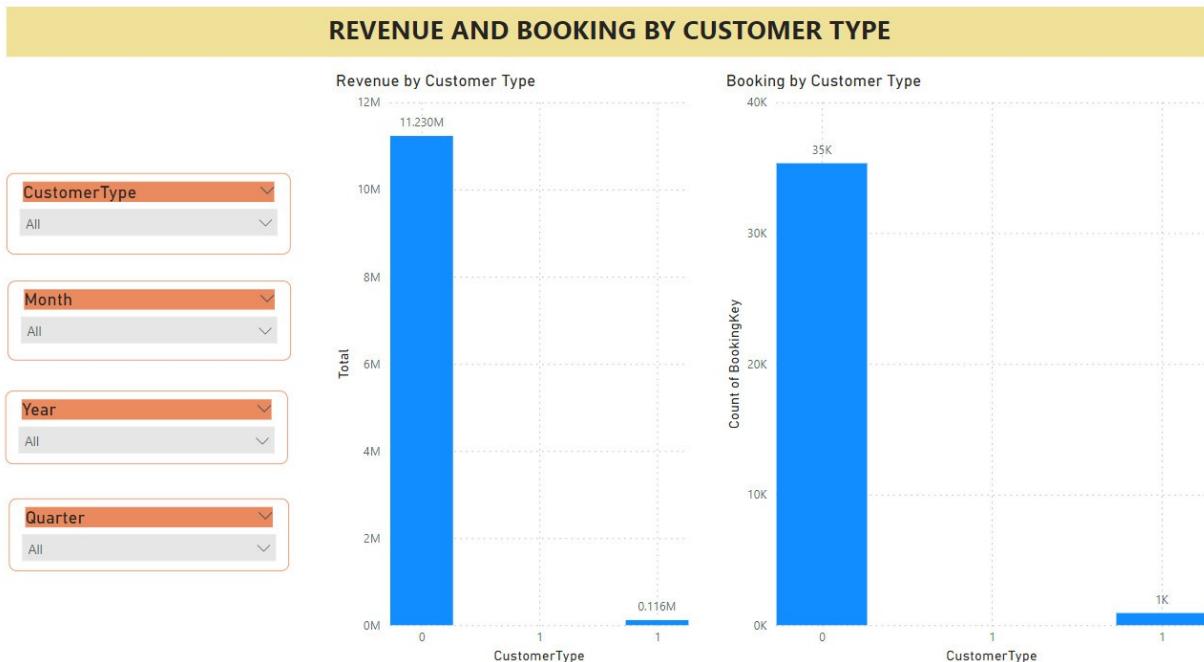
Biểu đồ thống kê số lượng đặt phòng và doanh thu của mỗi loại khách hàng



Nhận xét: Qua hình trên ta thấy từ năm 2017 đến 2018 thì số lượng khách quay lại khách sạn có tăng nhưng vẫn chiếm tỉ lệ rất thấp nên doanh thu thu được từ loại khách hàng này rất thấp

Qua đó, khách sạn có thể tìm hiểu nguyên nhân và tiến hành các chính sách ưu đãi giữ chân khách hàng, cải tiến dịch vụ để khách có thể quay lại khách sạn.

4.5.4.3. Sử dụng Power BI



4.5.5. Thống kê lượt đặt phòng mỗi tháng qua đó tiền hành thống kê lượt đặt thành công và lượt hủy mỗi tháng

4.5.5.1. Sử dụng công cụ SSAS

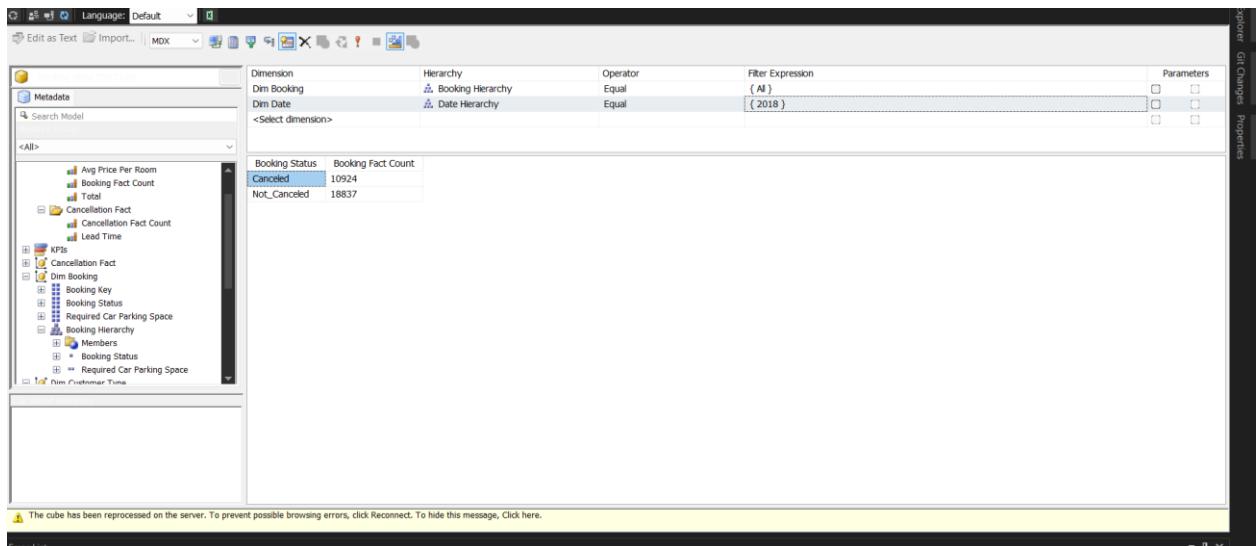
-Thống kê lượt đặt thành công và lượt hủy của 2 năm 2017, 2018

Booking Status	Booking Fact Count
Cancelled	11885
Not_Cancelled	24390

-Thống kê lượt đặt thành công và lượt hủy của 2 năm 2017

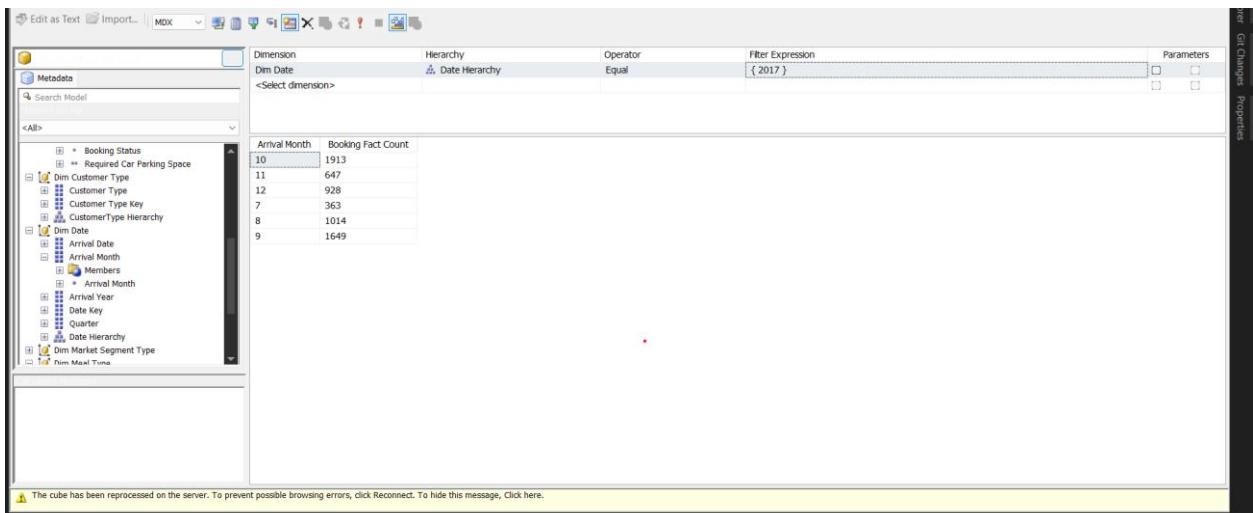
Booking Status	Booking Fact Count
Cancelled	961
Not_Cancelled	5553

-Thống kê lượt đặt thành công và lượt hủy của 2 năm 2018

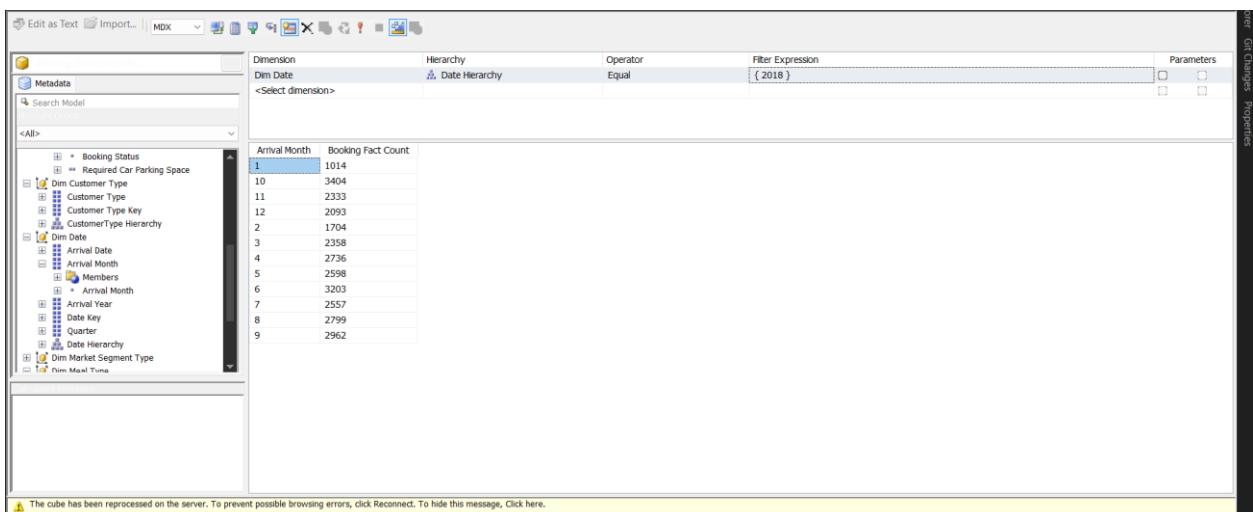


Nhận xét: Theo kết quả thống kê từ ba hình trên ta thấy tỉ lệ đặt phòng thành công từ năm 2017, 2018 bị giảm dần, tỉ lệ hủy tăng khi vào năm 2018 tỉ lệ hủy lên tới 37% so với tỉ lệ 15% vào năm 2017. Đây là tỉ lệ đáng báo động của khách sạn. Qua đó khách sạn cần phải tìm hiểu nguyên nhân và đưa ra hướng giải quyết phù hợp cho tình trạng này

-Thống kê số lượng đặt phòng theo tháng để tìm tháng được đặt phòng nhiều nhất năm 2017



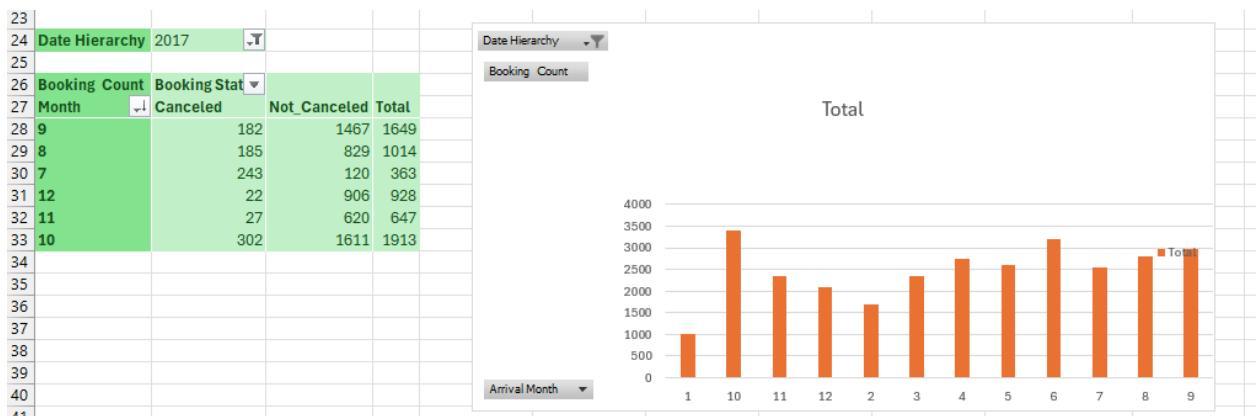
-Thống kê số lượng đặt phòng theo tháng để tìm tháng được đặt phòng nhiều nhất năm 2018



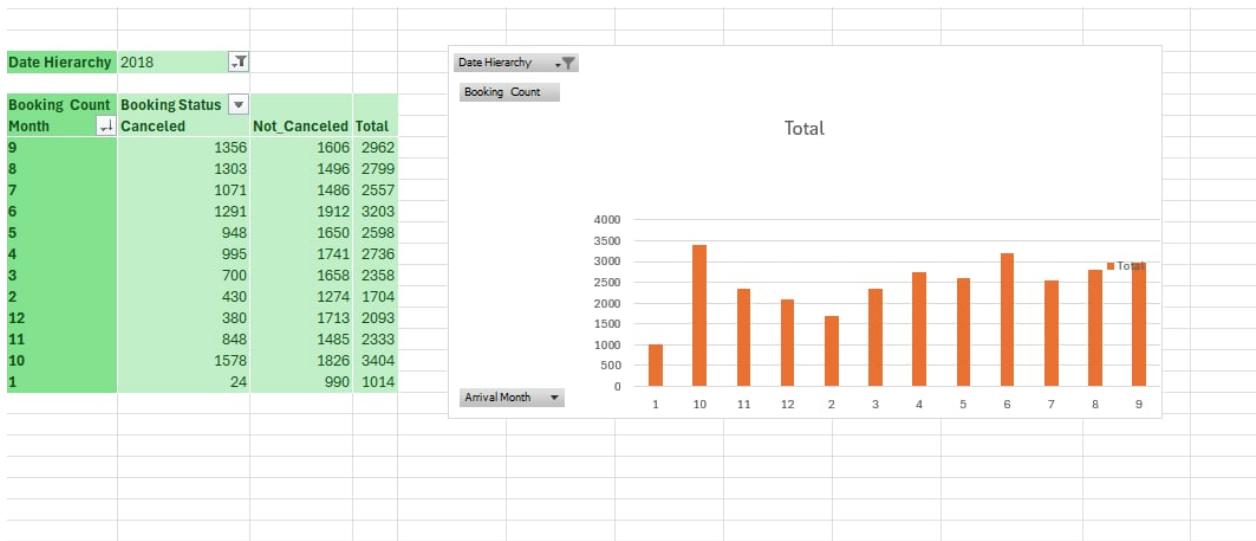
Nhận xét: Qua 2 thống kê ở trên ta thấy vào năm 2017 khách sạn chỉ hoạt động vào quý 3 và quý 4. Tuy nhiên khi thống kê cả hai năm thì đều có điểm chung là tháng đông khách nhất là tháng 10. Vì vậy khách sạn cần phân bổ nguồn lực sao cho hợp lý để phục vụ khách tốt nhất

4.5.5.2. Sử dụng Pivot Table trong Excel

-Biểu đồ thể hiện số lượt booking của khách sạn mỗi tháng và số lượt đặt thành công, hủy mỗi tháng của năm 2017



-Biểu đồ thể hiện số lượt booking của khách sạn mỗi tháng và số lượt đặt thành công, hủy mỗi tháng của năm 2018



4.5.5.3. Sử dụng Power BI



Nhận xét: Biểu đồ thể hiện tất cả lượt đặt và hủy phòng thành công, mặc dù lượt hủy phòng ít hơn với lượt đặt thành công nhưng vẫn tương đối nhiều, nên khách sạn phải đưa ra chiến lược để khách hàng hạn chế hủy phòng làm ảnh hưởng tới khách sạn.

CHƯƠNG 5: KẾT LUẬN

5.1. Kết quả đạt được

Thực hiện xây dựng ‘*KHO DỮ LIỆU BOOKING HOTEL*’ giúp chúng em hiểu rõ hơn về quy trình xây dựng nên 1 kho dữ liệu và phân tích các quy trình doanh nghiệp thực tế của đề tài. Ngoài ra, còn giúp chúng em thành thạo hơn trong việc sử dụng các công cụ hỗ trợ quá trình ETL đưa dữ liệu vào kho dữ liệu như công cụ SSIS, SSAS, các công cụ để trực quan hóa dữ liệu như Excel Pivot, Power BI Desktop. Nâng cao kỹ năng nhìn nhận và tiếp thu, xử lý thông tin đầu vào, đặt những câu hỏi tiền đề và thảo luận hướng đi để giải quyết và trả lời câu hỏi.

5.2. Hạn chế

Trong quá trình thực hiện đề tài cũng có những khó khăn nhất định như thiếu kinh nghiệm về mặt thực tế, mặc dù đã được học lí thuyết nhưng trong quá trình thực hiện nhóm cũng gặp phải những trở ngại khi thực hiện các bước. Ngoài ra do hạn chế về mặt kiến thức chuyên môn trong quy trình doanh nghiệp nên nhóm chưa thể phân tích kĩ và sâu với tập dữ liệu.

5.3. Tài liệu tham khảo

- [1] Tài liệu các file PDF hướng dẫn về Data Integration với SQL Server, SSIS, SSAS của GVHD Ths. Nguyễn Văn Thành.
- [2] “A proposed model for data warehouse ETL processes” - Shaker H. Ali El-Sappagh , Abdeltawab M. Ahmed Hendawi , Ali Hamed El Bastawiss (8/5/2011)
<https://www.sciencedirect.com/science/article/pii/S131915781100019X/?imgSel=Y>
- [3] Sukawirasa, I. Kadek Anom, et al. "Implementasi Data Warehouse Dan Penerapannya Pada PHI-Minimart Dengan Menggunakan Tools Pentaho dan Power BI." Jurnal Elektronik Ilmu Komputer Udayana P-ISSN 2301 (2008)
<https://ojs.unud.ac.id/index.php/JLK/article/download/73555/41172>
- [4] Obisesan, Abiodun Bamidele. "Implementing Data Warehouse On-premises." (2021)
<https://www.thesimus.fi/bitstream/handle/10024/493555/obisesan.abiodunThesisWork.pdf?sequence=2>

