

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP.HCM
KHOA CÔNG NGHỆ THÔNG TIN
MÔN KHO DỮ LIỆU



HCMUTE

BÁO CÁO ĐỀ TÀI CUỐI KỲ

XÂY DỰNG KHO DỮ LIỆU
QUẢN LÝ KHÁCH SẠN

LỚP HỌC PHẦN: DAWH430784_23_2_03CLC

HỌC KỲ II – NĂM HỌC 2023-2024

Thực hiện: Nhóm 05

Giảng viên hướng dẫn : Ths. Nguyễn Văn Thành

Thành phố Hồ Chí Minh, tháng 5 năm 2024

BỘ GIÁO DỤC VÀ ĐÀO TẠO
ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP HCM
KHOA CÔNG NGHỆ THÔNG TIN
MÔN KHO DỮ LIỆU

CỘNG HÒA XÃ HỘI CHỦ NGHĨA
VIỆT NAM
Độc lập – Tự do – Hạnh phúc



HCMUTE

BÁO CÁO ĐỀ TÀI CUỐI KỲ

XÂY DỰNG KHO DỮ LIỆU
QUẢN LÝ KHÁCH SẠN

LỚP HỌC PHẦN: DAWH430784_23_2_03CLC

HỌC KỲ II – NĂM HỌC 2023-2024

Thực hiện: Nhóm 05

Giảng viên hướng dẫn : Ths. Nguyễn Văn Thành

Thành phố Hồ Chí Minh, tháng 5 năm 2024

DANH SÁCH NHÓM THỰC HIỆN ĐỀ TÀI
MÔN KHO DỮ LIỆU

HỌC KỲ II NĂM HỌC 2023-2024

- 1. Giảng viên hướng dẫn:** Ths. Nguyễn Văn Thành
- 2. Tên đề tài :** Xây dựng kho dữ liệu quản lý khách sạn
- 3. Danh sách nhóm viết tiểu luận cuối kỳ:**

STT	HỌ VÀ TÊN SINH VIÊN	Mã số sinh viên	Tỉ lệ % tham gia
01	Trương Thị Thuỷ Dung	21110820	100%
02	Trần Văn Bảo Duy	21110155	100%
03	Bùi Đặng Thuỳ Thương	21110316	100%
04	Đinh Thị Thúy Quỳnh	21110284	100%

Nhận xét của giảng viên:

.....
.....
.....

Ngày ... tháng 05 năm 2024

Giảng viên chấm điểm

LỜI CẢM ƠN

Chúng em xin gửi lời cảm ơn sâu sắc đến Thầy Nguyễn Văn Thành, Thầy đã tận tình giảng dạy trong suốt quá trình tìm hiểu và học tập môn Kho dữ liệu, Trường đại học Sư phạm kỹ thuật Thành phố Hồ Chí Minh. Trong quá trình nghiên cứu và học, Thầy đã nhiệt huyết giảng dạy và những kiến thức mà Thầy đã truyền đạt không chỉ là nền tảng cho quá trình nghiên cứu, thực hiện đồ án mà còn là tư liệu quý báu để chúng em trong quá trình học tập và tìm hiểu sau này cũng như là trên con đường nghề nghiệp phía trước.

Một lần nữa, em xin bày tỏ lòng biết ơn đến Thầy. Xin kính chúc Thầy luôn dồi dào sức khỏe và đạt được nhiều thành công trong công việc.

Thành phố Hồ Chí Minh, tháng 5 năm 2024

KẾ HOẠCH PHÂN CÔNG NHIỆM VỤ THỰC HIỆN ĐỒ ÁN CUỐI KÌ MÔN KHO DỮ LIỆU HỌC KỲ II NĂM HỌC 2023-2024

1. Tên đề tài: Xây dựng kho dữ liệu quản lý khách sạn
2. Bảng phân công nhiệm vụ:

	Trương Thị Thùy Dung	Đinh Thị Thúy Quỳnh	Bùi Đăng Thùy Thương	Trần Văn Bảo Duy
NHIỆM VỤ				
Tìm kiếm tập dữ liệu	X	X	X	X
Mô hình hóa dữ liệu vào workbook	X	X	X	X
- Dim Customers - Dim Employees				X
- Dim Date			X	
- Fact Sales		X		
- Fact Bookings	X			
Thực hiện ETL đưa dữ liệu vào các dim và fact dùng SSIS	X		X	
- Tạo và đưa dữ liệu vào các dim stages và dim	X		X	
- Tạo và đưa dữ liệu vào stgHotelFactBooking và FactBookings	X			
- Tạo và đưa dữ liệu vào stgHotelFactSales và FactSales			X	
- Tạo và đưa dữ liệu vào SSAS, Tạo data source view và cube				X

Tạo cây phân cấp hierarchy cho các bảng dim và các attribute relationships.		X		X
- Hierarchy cho DimDate		X		
- Hierarchy cho DimCustomers				X
- Hierarchy cho DimEmployees				X
Đặt và thực hiện trả lời các câu hỏi cho Business process: Booking Transaction: Quản lý lượng booking của khách sạn	X	X	X	X
Thực hiện trả lời các câu hỏi đã đặt ra cho Business process: Booking Transaction: Quản lý lượng booking của khách sạn	X	X	X	X
Đặt và thực hiện trả lời các câu hỏi cho Business process: Sales Management: Quản lý doanh số của khách sạn	X	X	X	X
Thực hiện trả lời các câu hỏi đã đặt ra cho Business process: Sales Management: Quản lý doanh số của khách sạn	X	X	X	X
Trả lời câu hỏi dùng SSAS	X			X
Trả lời câu hỏi dùng Power BI Desktop		X	X	
Trả lời câu hỏi dùng Excel Pivot	X	X		

MỤC LỤC



CHƯƠNG 1: TỔNG QUAN VỀ ĐỀ TÀI	1
1.1. Lý do chọn đề tài	1
1.2. Tổng quan về tập dữ liệu	1
1.2.1. Nguồn dữ liệu	1
1.2.2. Mô tả chi tiết tập dữ liệu	1
CHƯƠNG 2: THIẾT KẾ XÂY DỰNG CƠ SỞ DỮ LIỆU TÁC NGHIỆP (OLAP)	5
2.1. Xác định các Business Process và bảng Fact	5
2.2. Xây dựng các bảng Dimension	6
2.3. Star schema (Lược đồ hình sao)	7
CHƯƠNG 3: TÍCH HỢP DỮ LIỆU VÀO KHO (SSIS)	8
3.1. Nhập dữ liệu vào bảng Date Dimension	8
3.1.1. Tạo package mới và thiết lập control flow	8
3.1.2. Load dữ liệu từ nguồn vào bảng stgDate	9
3.1.3. Thiết lập the truncate table SQL task	10
3.1.4. Load dữ liệu từ bảng stgDate vào bảng DimDate	11
3.2. Cấu hình cho bảng FactBookings	13
3.2.1. Cấu hình staging cho FactBookings	13
3.2.2. Load dữ liệu vào bảng FactBookings	16
3.3. Cấu hình cho bảng FactSales	24
3.3.1. Cấu hình staging cho FactSales	24
3.3.2. Load dữ liệu vào bảng FactSales	28
CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS)	35
4.1. Xây dựng mô hình	35
4.1.1. Tạo Data Source	35
4.1.2. Tạo Data SourceView	37
4.2. Tạo Data cube	40
4.3. Cấu hình Hierarchy	42

4.3.1.	Tạo Hierarchy Dim Date	42
4.3.2.	Tạo Hierarchy cho Dim Employees	43
4.3.3.	Tạo Hierarchy cho Dim Customers	44
4.4.	Thực hiện tạo Cube	44
4.5.	Thực hiện phân tích dữ liệu	45
4.5.1.	Sử dụng SSAS	45
4.5.2.	Sử dụng Pivot Table trong Excel và Power BI	47
CHƯƠNG 5. KẾT LUẬN		57
5.1.	Kết quả đạt được.....	57
5.2.	Những hạn chế	57

CHƯƠNG 1: TỔNG QUAN VỀ ĐỀ TÀI

1.1. Lý do chọn đề tài

Việc chọn đề tài xây dựng kho dữ liệu cho nhà hàng khách sạn với ứng dụng ETL và các công cụ như SSAS và SSIS mang lại nhiều lợi ích quan trọng. Đầu tiên, ngành nhà hàng và khách sạn thường phải quản lý lượng dữ liệu lớn từ nhiều nguồn khác nhau như đặt phòng, dịch vụ ăn uống và phản hồi của khách hàng.

Việc sử dụng quy trình ETL giúp thu thập, chuyển đổi và nạp dữ liệu vào kho một cách hiệu quả, đảm bảo chất lượng và tính toàn vẹn của dữ liệu. SSIS hỗ trợ xây dựng và thực thi các quy trình ETL, trong khi SSAS cung cấp khả năng phân tích dữ liệu mạnh mẽ, tạo ra các khối dữ liệu phân tích để đưa ra quyết định dựa trên dữ liệu.

Điều này không chỉ tối ưu hóa hoạt động kinh doanh mà còn cải thiện trải nghiệm khách hàng thông qua các chiến dịch tiếp thị tùy chỉnh và dịch vụ cá nhân hóa. Hơn nữa, việc ứng dụng các công cụ công nghệ này giúp nâng cao kỹ năng chuyên môn, đáp ứng nhu cầu ngày càng tăng về phân tích dữ liệu trong ngành Hệ thống thông tin.

1.2. Tổng quan về tập dữ liệu

1.2.1. Nguồn dữ liệu

Nhóm sử dụng Tập dữ liệu Hotel Database được lấy từ trang web Kaggle ([kaggle.com](https://www.kaggle.com)), trang web này cung cấp cơ sở dữ liệu được tổng hợp bởi các nguồn dữ liệu lớn và đáng tin cậy.

Đường dẫn tải tập dữ liệu: [Hotel booking analysis Database](#)

1.2.2. Mô tả chi tiết tập dữ liệu

Tập dữ liệu gồm thông tin của một khách sạn và việc đặt phòng của khách, dữ liệu gồm 7 bảng

Tên bảng	Mô tả	Số dòng
BedTypes	Lưu trữ thông tin loại giường	11 dòng
Bookings	Lưu trữ thông tin đặt phòng của khách	400 dòng
Customers	Lưu trữ thông tin của khách	400 dòng
Employees	Lưu trữ thông tin nhân viên	59 dòng
Payments	Lưu trữ thông tin thanh toán của khách	99 dòng

Rooms	Lưu trữ thông tin về phòng	200 dòng
RoomTypes	Lưu trữ thông tin về loại phòng	10 dòng

1.2.2.1. Bảng BedTypes

Thuộc tính	Mô tả
BedTypeID	Mã loại giường (Khóa chính)
BedType	Tên loại giường
Notes	Ghi chú
BedCost	Giá của loại giường

1.2.2.2. Bảng Bookings

Thuộc tính	Mô tả
BookingID	Mã đặt phòng của khách (Khóa chính)
ArrvialDate	Ngày đến
DepartureDate	Ngày đi
RoomNumber	Số phòng
CustomerID	Mã khách hàng (Khóa ngoại tham chiếu đến Customers(CustomerID))
ReservationStatus	Trạng thái đặt phòng
ReservationStatusDate	Ngày cập nhật trạng thái
EmployeeID	Mã nhân viên (Khóa ngoại tham chiếu đến Employees(EmployeeID))
GuestCount	Số lượng khách

1.2.2.3. Bảng Customers

Thuộc tính	Mô tả
CustomerID	Mã khách hàng (Khóa chính)
FirstName	Tên của khách
LastName	Họ của khách
BirthDate	Ngày sinh của khách
Email	Email của khách

Phone	Số điện thoại của khách
-------	-------------------------

1.2.2.4. Bảng Employees

Thuộc tính	Mô tả
EmployeeID	Mã nhân viên (Khóa chính)
EmployeeName	Tên nhân viên
Title	Vị trí công việc của nhân viên
City	Tên thành phố
Country	Tên quốc gia
HireDate	Ngày tuyển dụng làm việc của nhân viên

1.2.2.5. Bảng Payments

Thuộc tính	Mô tả
PaymentID	Mã thanh toán (Khóa chính)
BookingId	Mã đặt phòng của khách (Khóa ngoại tham chiếu đến Bookings(BookingID))
PaymentDate	Ngày thanh toán
FirstDateOccupied	Ngày đến của khách
LastDateOccupied	Ngày rời khách sạn của khách
TaxRate	Thuế
Duration	Số ngày ở khách sạn
Discount	Giảm giá

1.2.2.6. Bảng Rooms

Thuộc tính	Mô tả
RoomNumber	Số phòng (Khóa chính)
RoomTypeID	Mã loại phòng (Khóa ngoại tham chiếu đến RoomTypes(RoomTypeID))
BedTypeID	Mã loại giường (Khóa ngoại tham chiếu đến BedTypes(BedTypeID))
Rate	
Notes	Ghi chú

1.2.2.7. Bảng RoomTypes

Thuộc tính	Mô tả
RoomTypeID	Mã loại phòng (Khóa chính)
RoomType	Tên loại phòng
Notes	Ghi chú
TypeCost	Giá loại phòng

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

2.1. Xác định các Business Process và bảng Fact

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

Instructions!												
Business Process Name	Fact Table	Fact Grain Type	Granularity	Facts	DimDate	Bookings	Customers	Employees	Payments	Rooms	BedTypes	RoomTypes
Booking Transaction	FactBookings	Transaction	one row per booking detail	BookingID, CustomerKey, EmployeeKey, ArrivalDateKey, DepartureDateKey, RoomNumberKey, RoomType, BedType, GuestCount, BookingStatus, RoomCost, BedCost, TotalPrice	x	x	x	x		x	x	x
Sales Management	FactSales	Transaction	one row per sales detail	BookingID, CustomerKey, EmployeeKey, SettlementDateKey, RoomNumber, RoomType, BedType, RoomCost, BedCost, GuestNumber, Discount, Tax, Total	x	x	x	x	x	x	x	x

Hình 1. Detailed Bus Matrix

Business Process: Booking Transaction

Instructions!		
Dimension / Fact Table	Attribute / Fact Name	Description
FactBookings	BookingID	A unique identifier for each booking
FactBookings	CustomerKey	ForeignKey to DimCustomers (CustomerKey)
FactBookings	EmployeeKey	ForeignKey to DimEmployees (EmployeeKey)
FactBookings	ArrivalDateKey	ForeignKey to DimDate (DateKey)
FactBookings	DepartureDateKey	ForeignKey to DimDate (DateKey)
FactBookings	RoomNumberKey	Number of each room
FactBookings	RoomType	Name of room type
FactBookings	BedType	Name of bed type
FactBookings	GuestCount	Number of guests included in this booking
FactBookings	BookingStatus	Indicates if it's confirmed, cancelled, pending, etc
FactBookings	RoomCost	Price of room type
FactBookings	BedCost	price of bed type
FactBookings	TotalPrice	The total price charged for this specific booking

Hình 2. Bảng FactBookings

Business Process : Sales Management

Instructions!		
Dimension / Fact Table	Attribute / Fact Name	Description
FactSales	BookingID	A unique identifier for each booking
FactSales	CustomerKey	ForeignKey to DimCustomers (CustomerKey)
FactSales	EmployeeKey	ForeignKey to DimEmployees (EmployeeKey)
FactSales	SettlementDateKey	ForeignKey to DimDate (DateKey)
FactSales	RoomNumber	Number of each room
FactSales	RoomType	Name of room type
FactSales	BedType	Name of bed type
FactSales	RoomCost	The price associated with the type of room
FactSales	BedCost	The cost related to the type of bed in the room
FactSales	GuestNumber	The number of guests that will be staying in the room
FactSales	Discount	discount
FactSales	Tax	VAT
FactSales	Total	Total amount

Hình 3. Bảng FactSales

2.2. Xây dựng các bảng Dimension

DimDate

Instructions!		
Dimension / Fact Table	Attribute / Fact Name	Description
DimDate	DateKey	A unique identifier for each date in the table
DimDate	Date	Full date
DimDate	DayOfWeek	Number of the day of week
DimDate	DayName	Day name of the week
DimDate	DayOfMonth	Number of the day in the month
DimDate	DayOfYear	Number of the day in the year
DimDate	WeekOfYear	Number of the week in the year
DimDate	MonthName	Month name of year
DimDate	MonthOfYear	Number of the month in the year
DimDate	Quarter	Number of the quarter
DimDate	Year	Number of the year
DimDate	IsWeekDay	Weekend of Weekday

Hình 4. DimDate

DimCustomers

Instructions!		
Dimension / Fact Table	Attribute / Fact Name	Description
DimCustomers	CustomerKey	A unique identifier for each customer in the table
DimCustomers	CustomerID	ID of each customer
DimCustomers	FirstName	First name of customer
DimCustomers	LastName	Last name of customer
DimCustomers	BirthDate	BirthDate of customer
DimCustomers	Email	Email of customer
DimCustomers	Phone	Phone of customer

Hình 5. DimCustomers

DimEmployees

Instructions!		
Dimension / Fact Table	Attribute / Fact Name	Description
DimEmployees	EmployeeKey	A unique identifier for each employee in the table
DimEmployees	EmployeeID	ID of each employee
DimEmployees	EmployeeName	Name of each employee
DimEmployees	Title	Title of each employee
DimEmployees	City	City of each employee
DimEmployees	Country	Country of each employee
DimEmployees	HireDate	Hire date of each employee

Hình 6. DimEmployees

2.3. Star schema (Lược đồ hình sao)

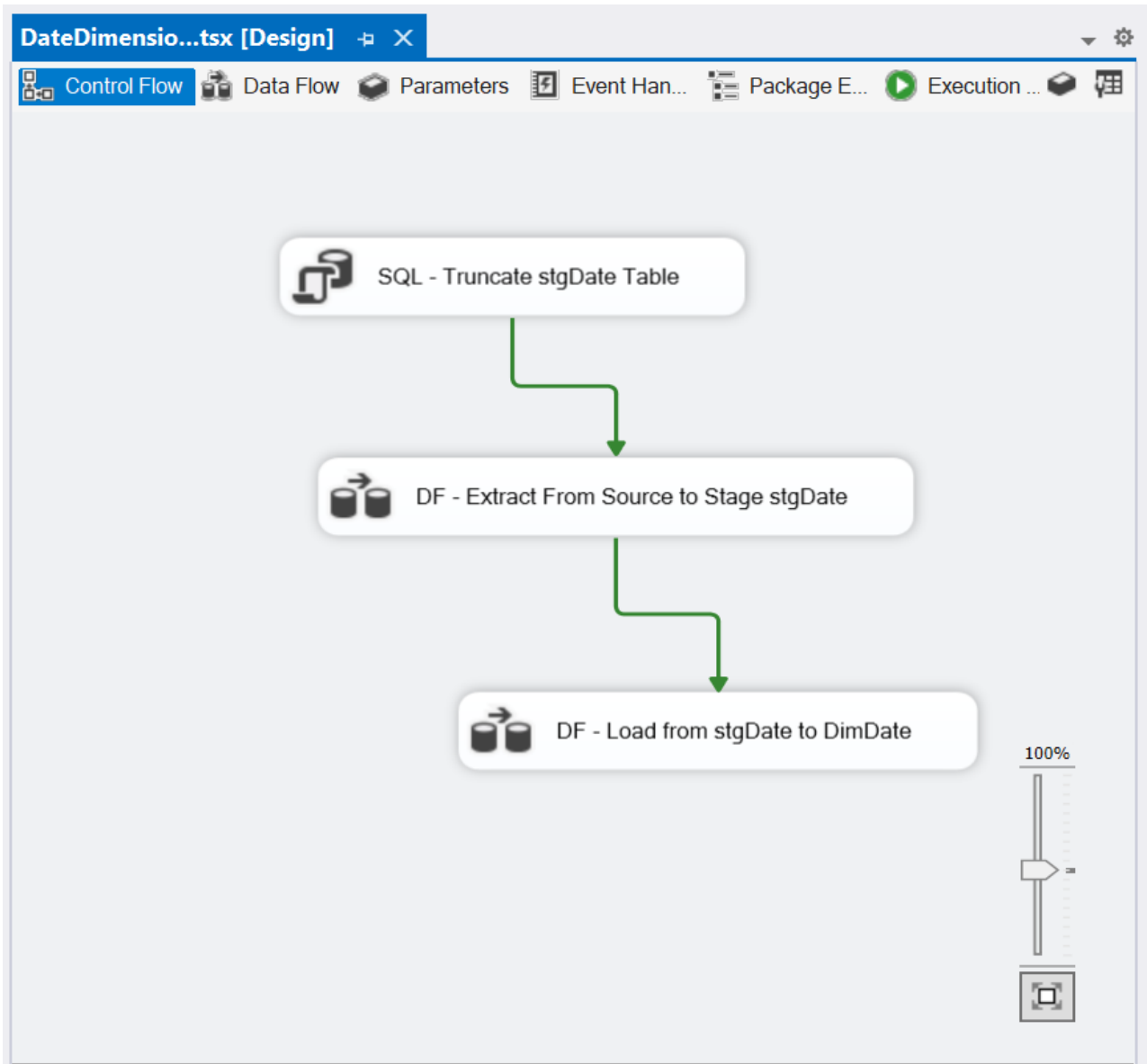


Hình 7. Lược đồ hình sao

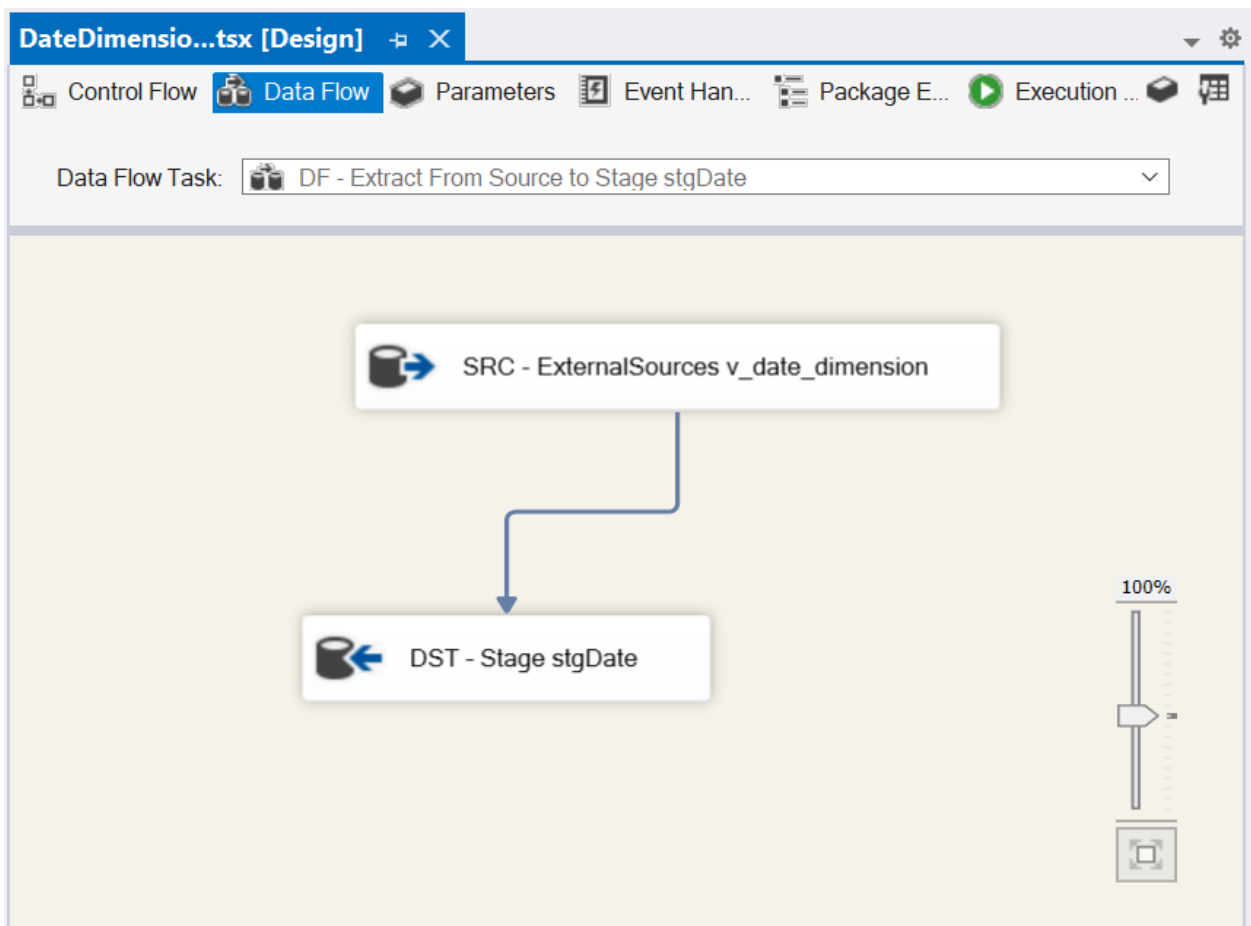
CHƯƠNG 3: TÍCH HỢP DỮ LIỆU VÀO KHO (SSIS)

3.1. Nhập dữ liệu vào bảng Date Dimension

3.1.1. Tạo package mới và thiết lập control flow



3.1.2. Load dữ liệu từ nguồn vào bảng stgDate



3.1.3. Thiết lập the truncate table SQL task

Execute SQL Task Editor

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

General
Parameter Mapping
Result Set
Expressions

General	
Name	SQL - Truncate stgDate Table
Description	Execute SQL Task
Options	
TimeOut	0
CodePage	1252
TypeConversionMode	Allowed
Result Set	
ResultSet	None
SQL Statement	
ConnectionType	OLE DB
Connection	DESKTOP-UQMD91R.hotel_stage
SQLSourceType	Direct input
SQLStatement	truncate table stgDate;
IsQueryStoredProcedure	False
BypassPrepare	True

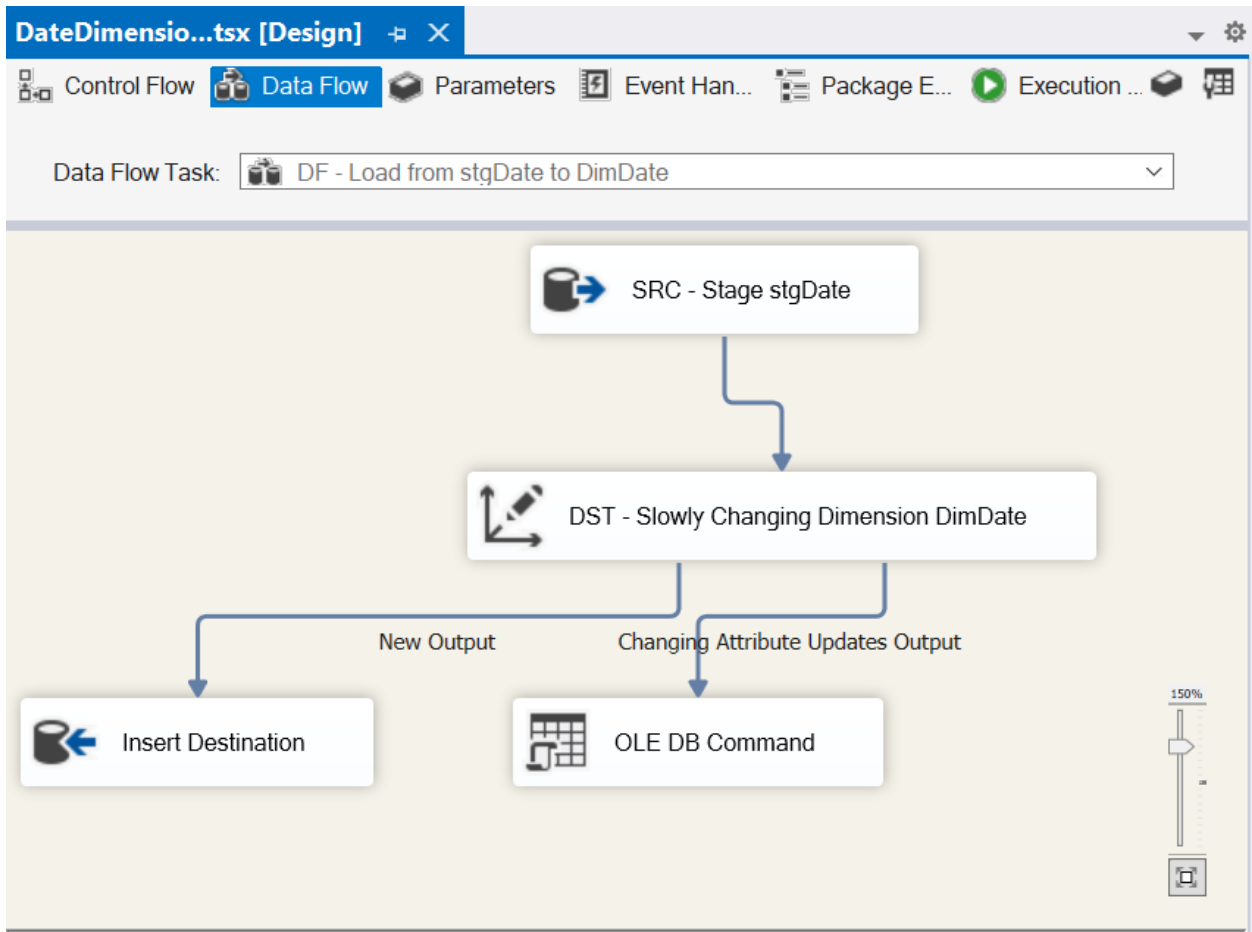
Name
Specifies the name of the task.

Browse... Build Query... Parse Query

OK Cancel Help

3.1.4. Load dữ liệu từ bảng stgDate vào bảng DimDate

Bước cuối cùng là định cấu hình giai đoạn cho luồng dữ liệu bảng DimDate. Trong luồng dữ liệu này, sử dụng Type 1 Slowly Changing Dimension để đảm bảo rằng không vô tình đưa cùng một khóa business key nhiều lần vào bảng dimension



Select a Dimension Table and Keys

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

Connection manager:

DESKTOP-UQMD91R.HotelDW

New...

Table or view:

[hotel].[DimDate]

Input Columns	Dimension Columns	Key Type
full date	Date	Not a key column
date key	DateKey	Business key
day name	DayName	Not a key column
day num in month	DayOfMonth	Not a key column
day of week	DayOfWeek	Not a key column
day num overall	DayOfYear	Not a key column
weekday flag	IsWeekday	Not a key column
month name	MonthName	Not a key column

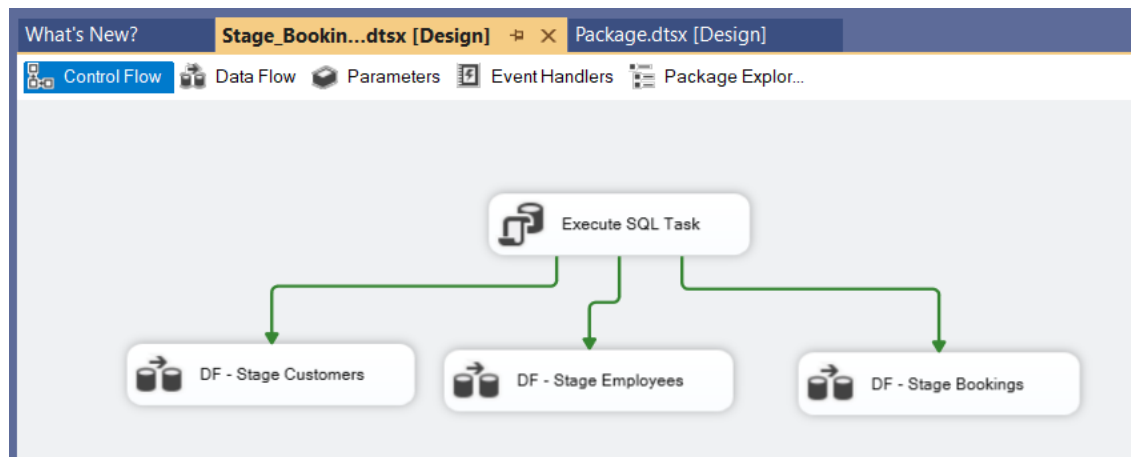
Help

3.2. Cấu hình cho bảng FactBookings

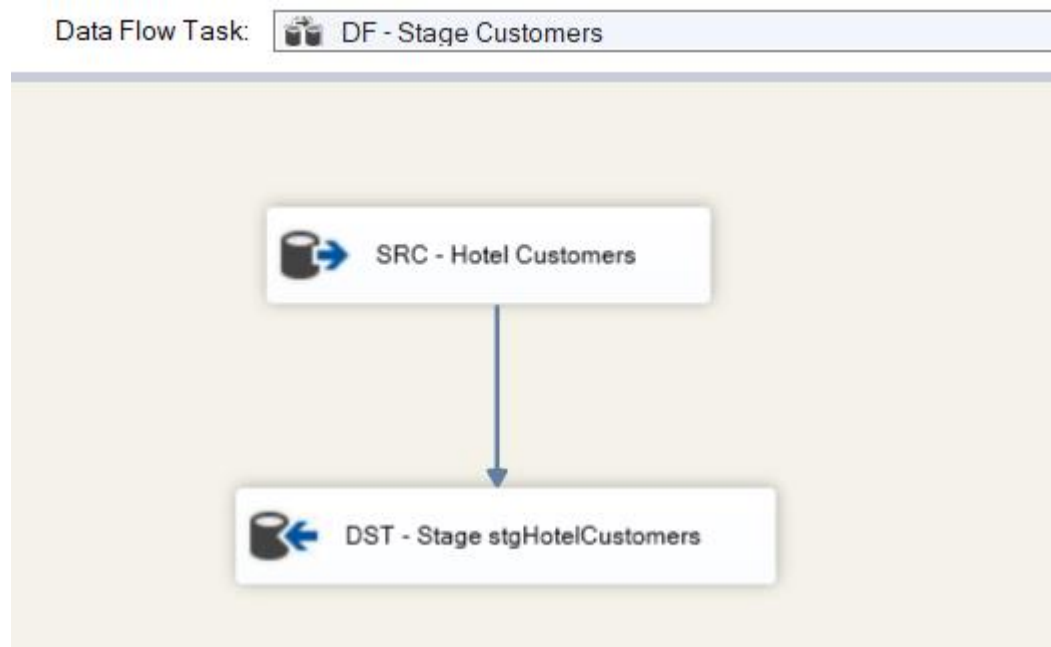
3.2.1. Cấu hình staging cho FactBookings

3.2.1.1. Tạo package mới và thiết lập control flow

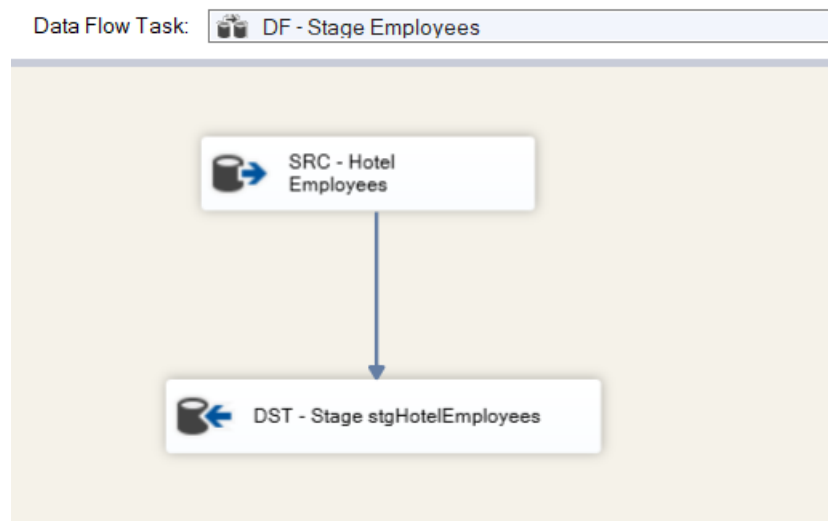
❖ Tạo một Execute SQL Task và 3 Data Flow Task



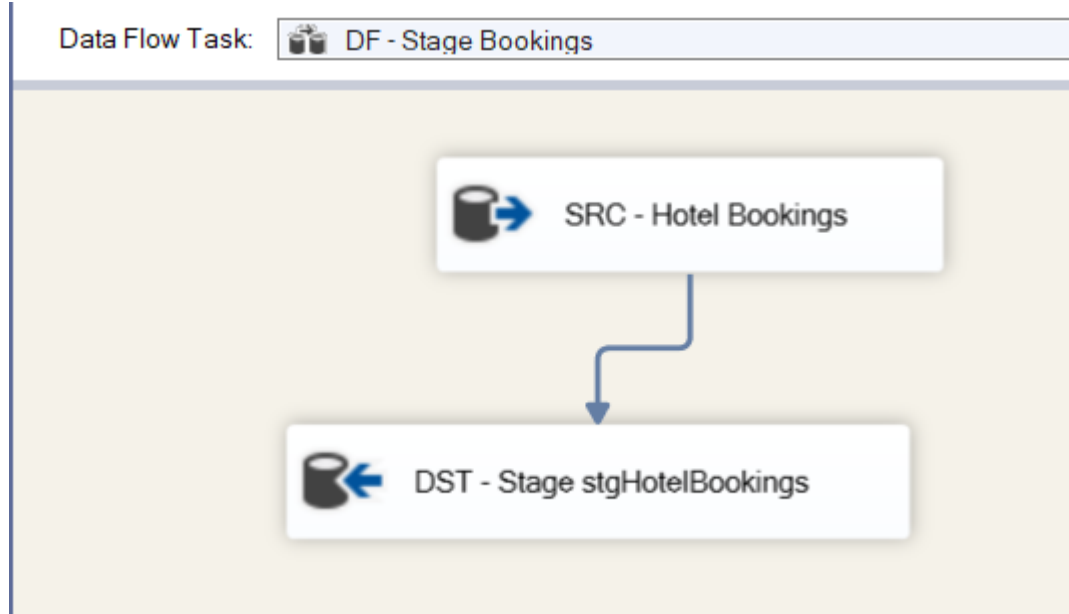
3.2.1.2. Load dữ liệu từ nguồn vào bảng stgHotelCustomers



3.2.1.3. Load dữ liệu từ nguồn vào bảng stgHotelEmployees



3.2.1.4. Load dữ liệu từ nguồn vào bảng stgHotelBookings



❖ **Dùng SQL command để kết bảng cho stgHotelBookings**

OLE DB connection manager:
 DESKTOP-KC6BO6E\SQL2019.Hotel

Data access mode:
 SQL command

SQL command text:

```
SELECT Bookings.BookingID, Bookings.ArrivalDate,
Bookings.DepartureDate, Bookings.GuestCount,
Rooms.RoomNumber, BedTypes.BedType,
RoomTypes.RoomType, BedTypes.BedCost,
RoomTypes.TypeCost, Bookings.ReservationStatus,
Bookings.EmployeeID,
Bookings.CustomerID
FROM Bookings INNER JOIN
Rooms ON Bookings.RoomNumber =
Rooms.RoomNumber INNER JOIN
BedTypes ON Rooms.BedTypeID = BedTypes.BedTypeID
INNER JOIN
RoomTypes ON Rooms.RoomTypeID =
RoomTypes.RoomTypeID
```

Parameters...
 Build Query...
 Browse...
 Parse Query

❖ **Kiểm tra mapping giữa source và destination**

Available Input...

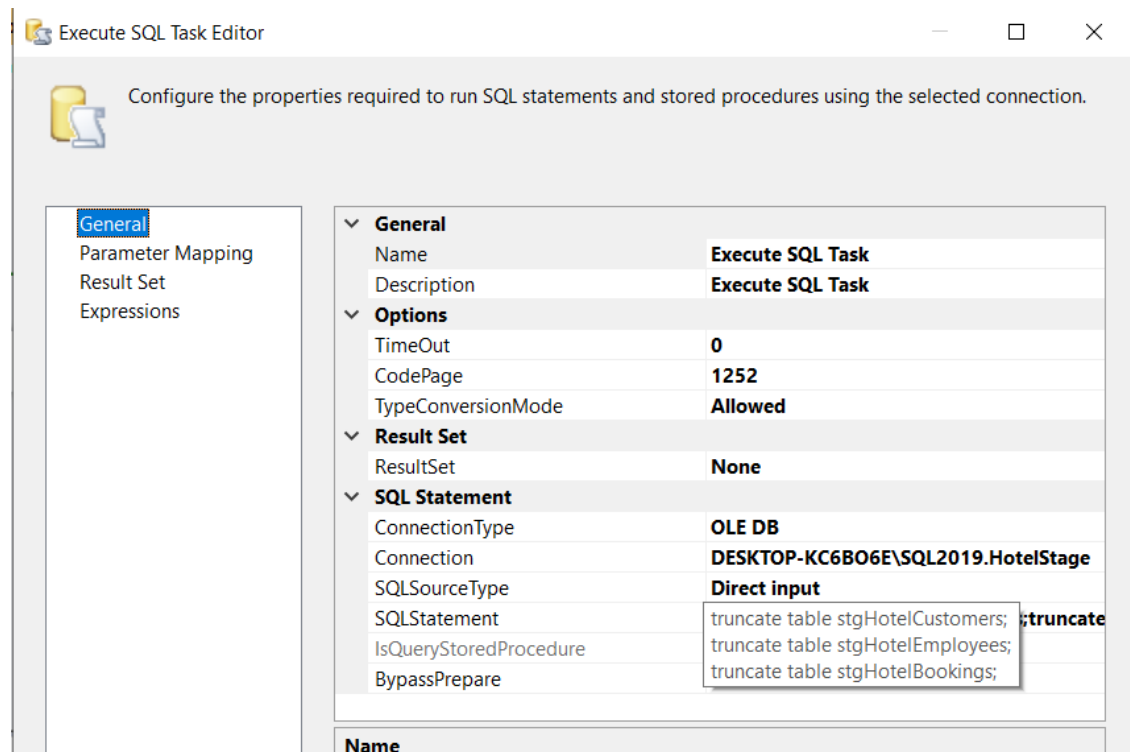
Name
CustomerID
EmployeeID
BookingID
GuestCount
ArrivalDate
Departure...
Reservatio...
BedType
RoomType

Available Desti...

Name
CustomerID
EmployeeID
BookingID
GuestCount
ArrivalDate
Departure...
Reservatio...
BedType
RoomType

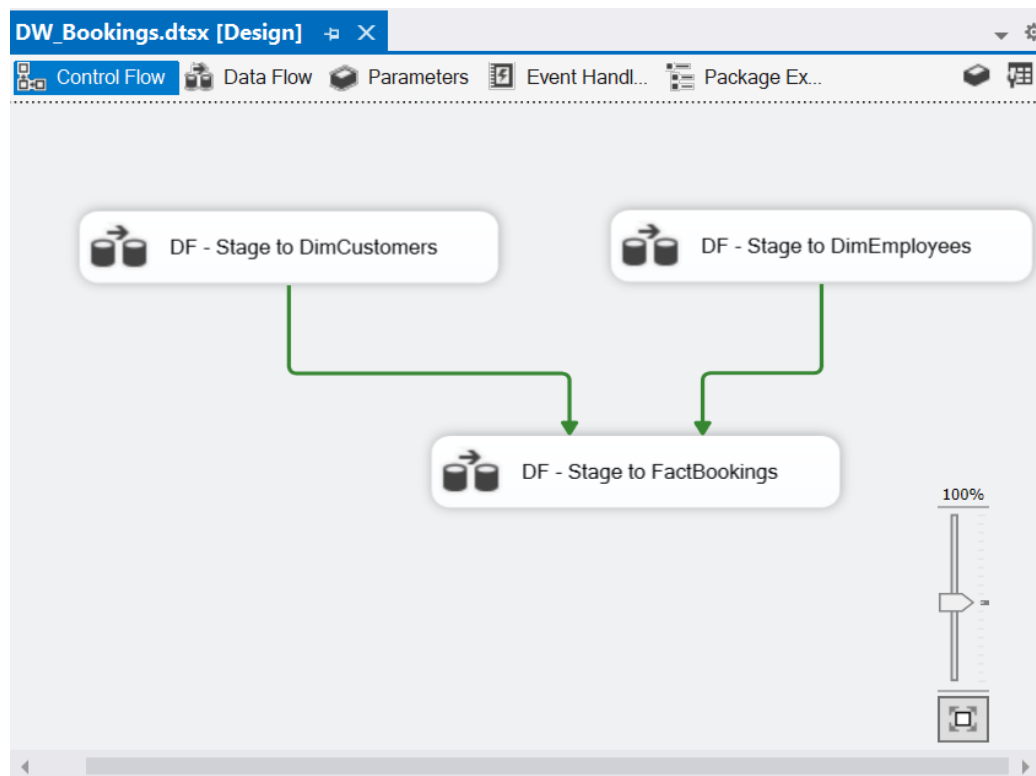
Input Column	Destination Column
CustomerID	CustomerID
EmployeeID	EmployeeID
BookingID	BookingID
GuestCount	GuestCount
ArrivalDate	ArrivalDate
DepartureDate	DepartureDate
ReservationStatus	ReservationStatus
BedType	BedType

3.2.1.5. Thiết lập the truncate table SQL task

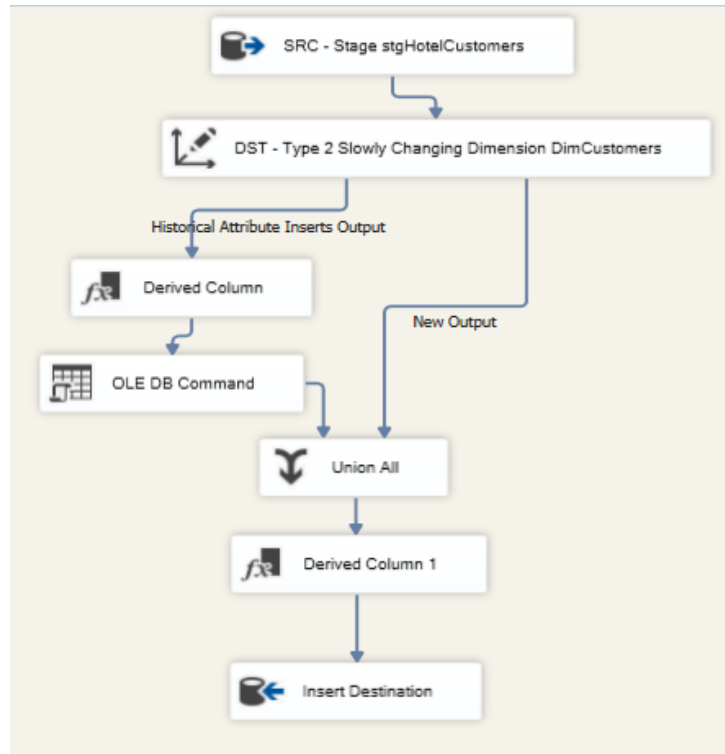


3.2.2. Load dữ liệu vào bảng FactBookings

3.2.2.1. Tạo package mới và thiết lập control flow



3.2.2.2. Load dữ liệu từ stgCustomers vào DimCustomers



Slowly Changing Dimension Wizard

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

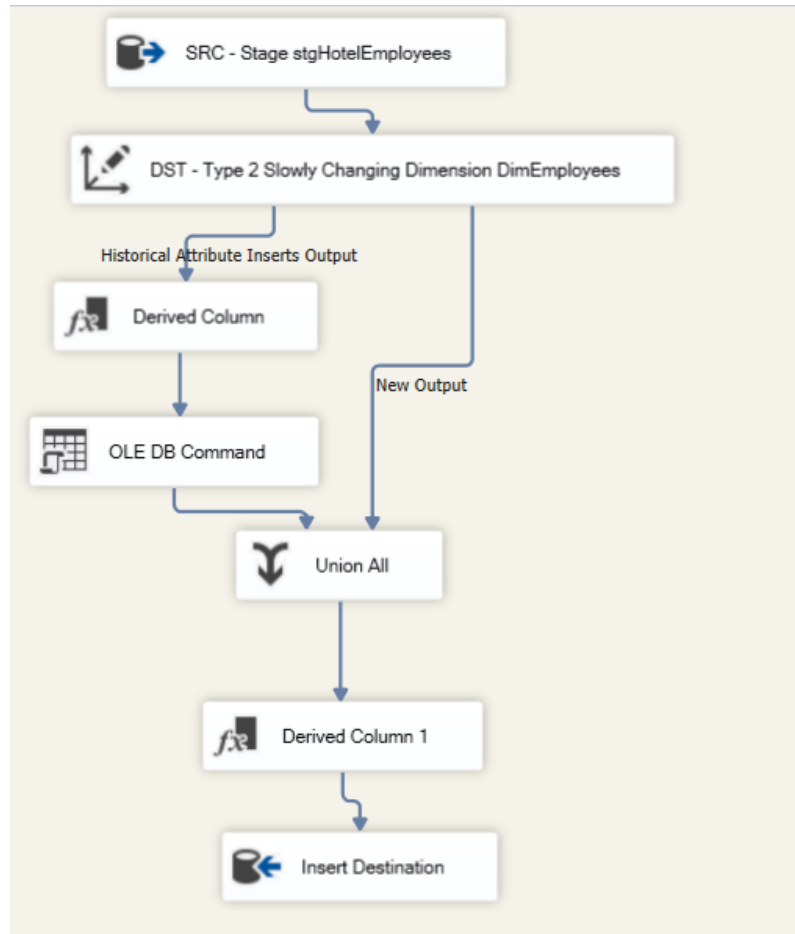
Connection manager:
DESKTOP-UQMD91R.HotelDW

Table or view:
[hotel].[DimCustomers]

Input Columns	Dimension Columns	Key Type
BirthDate	BirthDate	Not a key column
CustomerID	CustomerID	Business key
Email	Email	Not a key column
FirstName	FirstName	Not a key column
LastName	LastName	Not a key column
Phone	Phone	Not a key column
	RowChangeReason	
	RowEndDate	

Help < Back Next > Finish >>| Cancel

3.2.2.3. Load dữ liệu từ stgEmployees vào DimEmployees



Slowly Changing Dimension Wizard

Slowly Changing Dimension Columns

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

Fixed Attribute

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

Changing Attribute

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

Historical Attribute

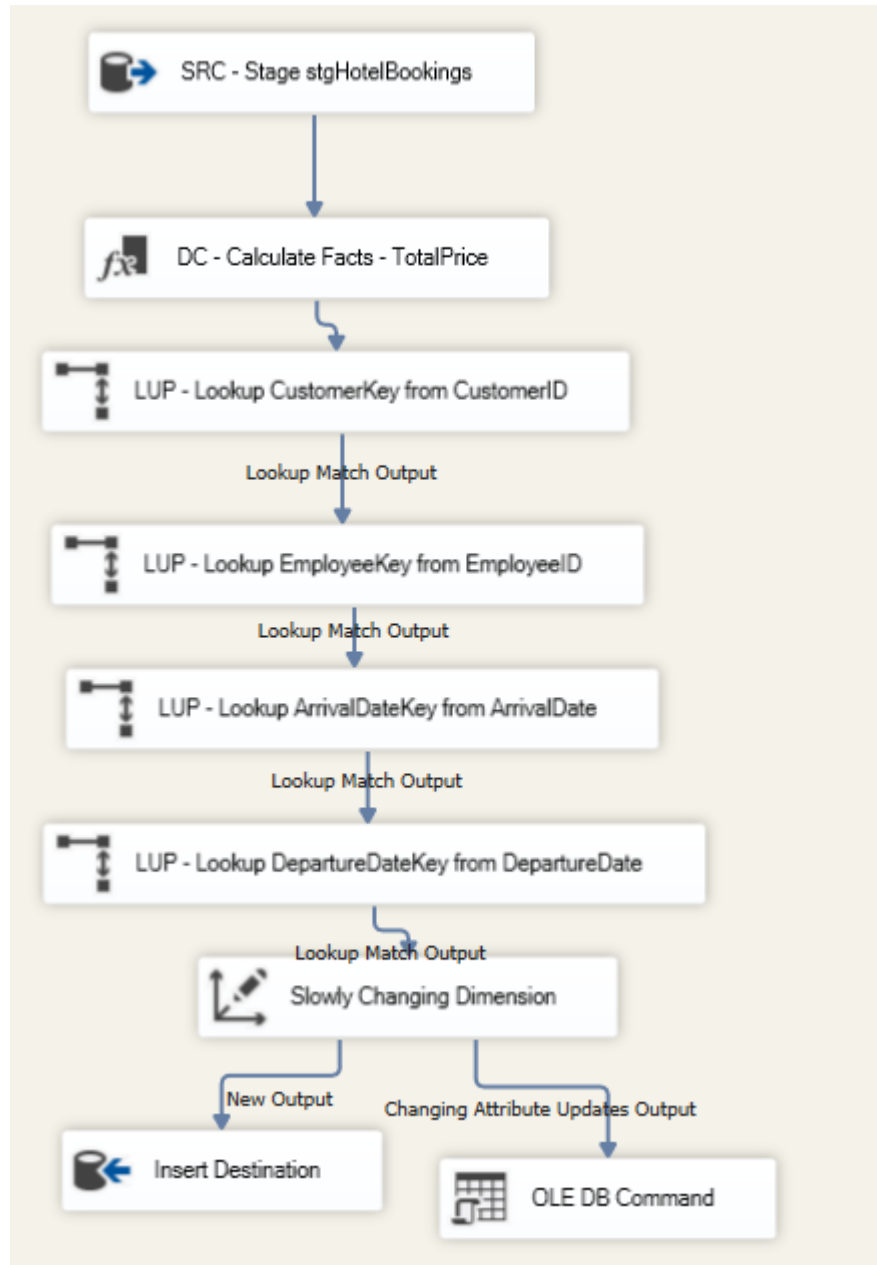
Select this type when changes in column values are saved in new records. Previous values are saved in

Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
City	Historical a...
Country	Historical a...
EmployeeName	Historical a...
HireDate	Historical a...
Title	Historical a...

Help < Back Next > Finish >> Cancel

3.2.2.4. Load dữ liệu từ stgBookings vào bảng FactBookings



❖ Tính toán tổng chi phí đặt phòng (TotalPrice)

Derived Column Transformation Editor

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

Variables and Parameters

Columns

Mathematical Functions

String Functions

Date/Time Functions

NULL Functions

Type Casts

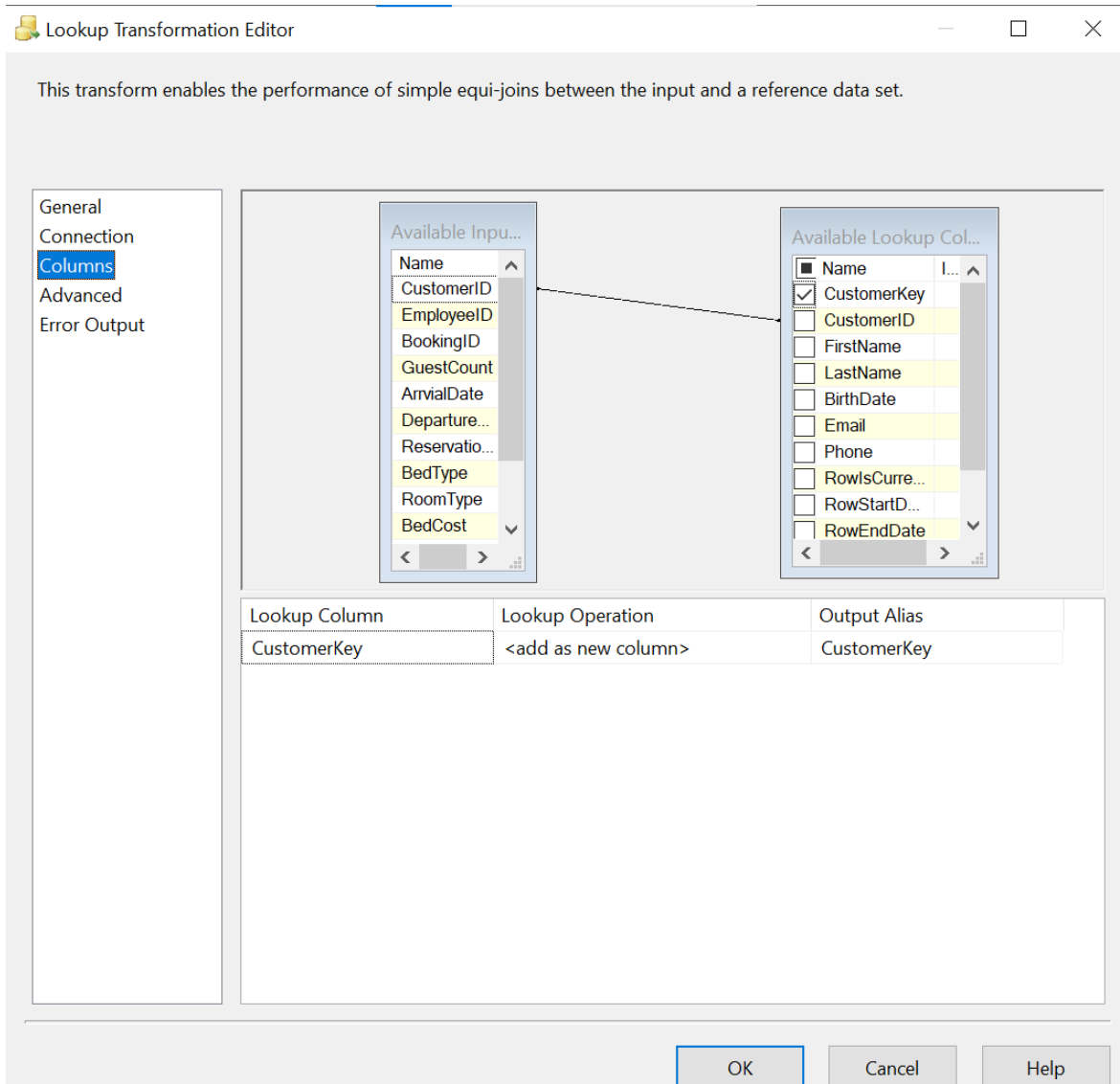
Operators

Description:

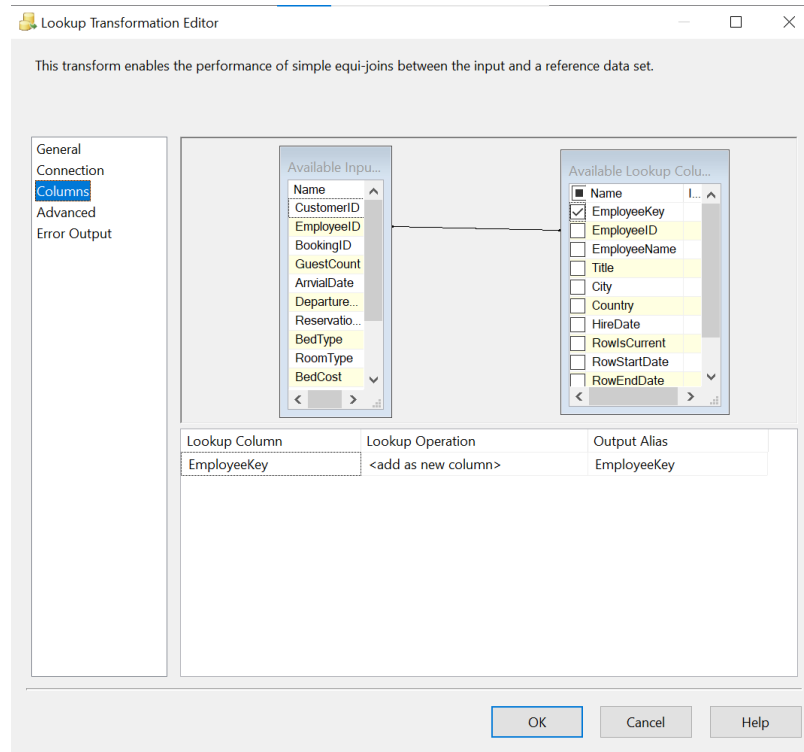
Derived Column Name	Derived Column	Expression	Data Type	Le
TotalPrice	<add as new column>	(DT_NUMERIC,18,4)(TypeCost + BedCost)	numeric [DT_NUMERIC]	

Configure Error Output... OK Cancel Help

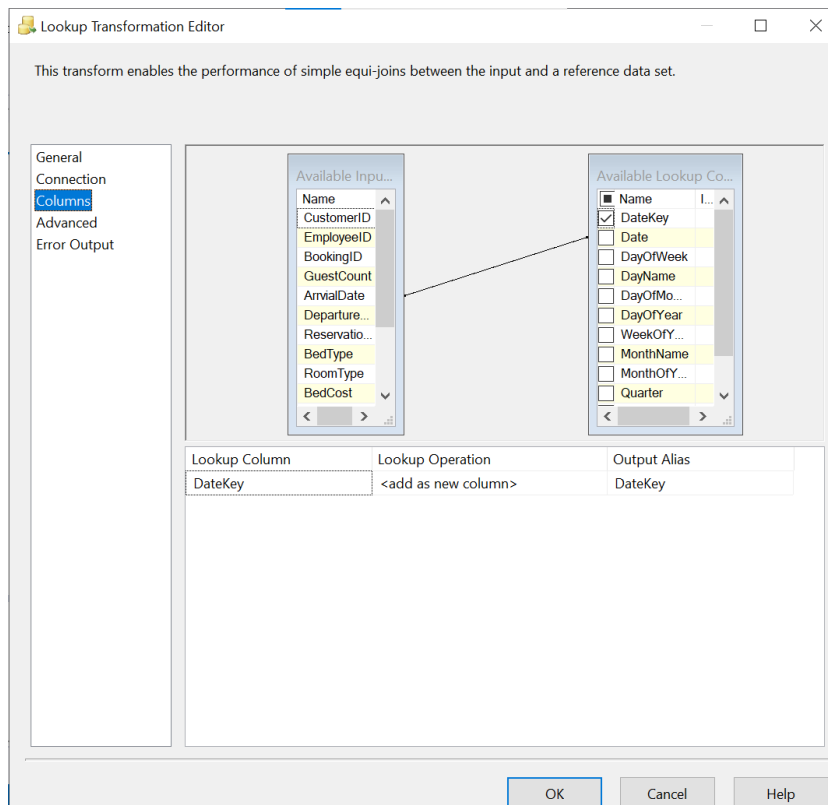
❖ Tìm CustomerKey từ CustomerID



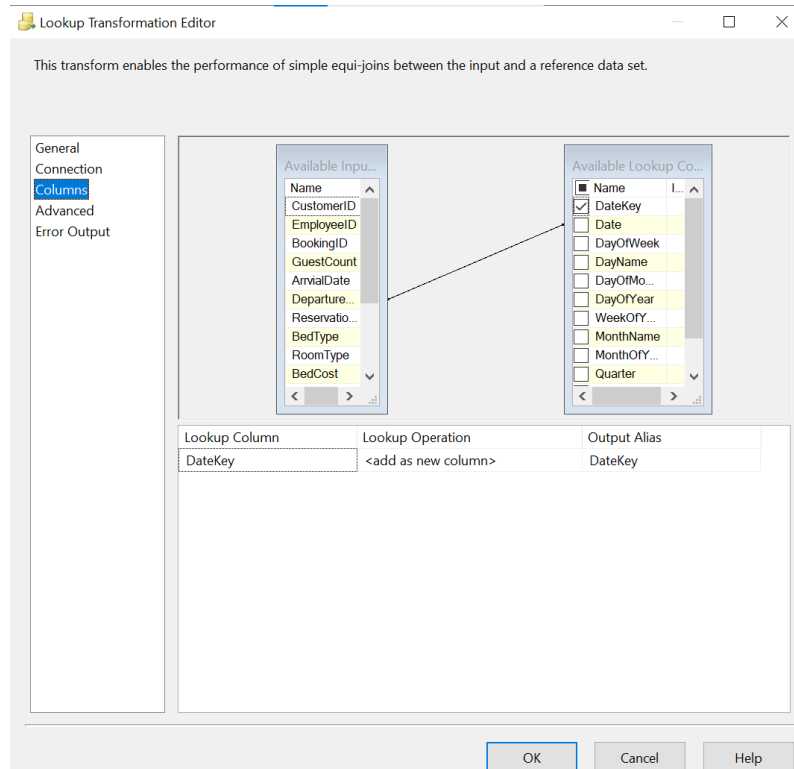
❖ Tìm EmployeeKey từ EmployeeID



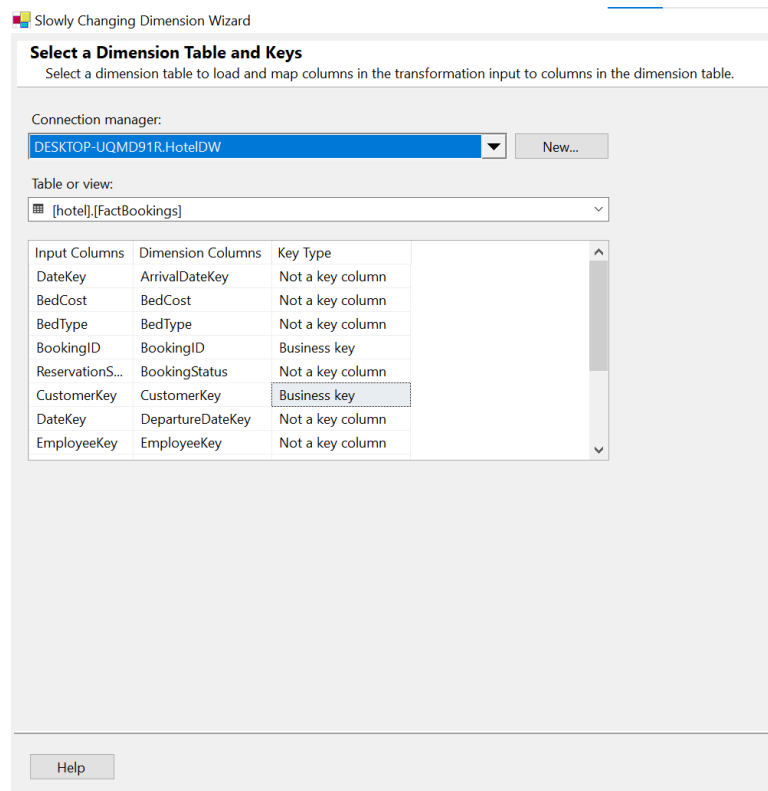
❖ Tìm ArrivalDateKey từ ArrivalDate



❖ Tım DepartureDateKey tır DepartureDate



❖ Slow Changing Dimension

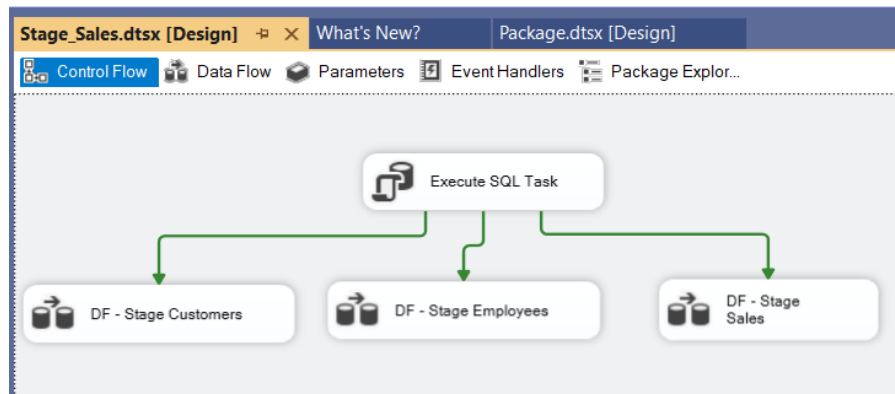


3.3. Cấu hình cho bảng FactSales

3.3.1. Cấu hình staging cho FactSales

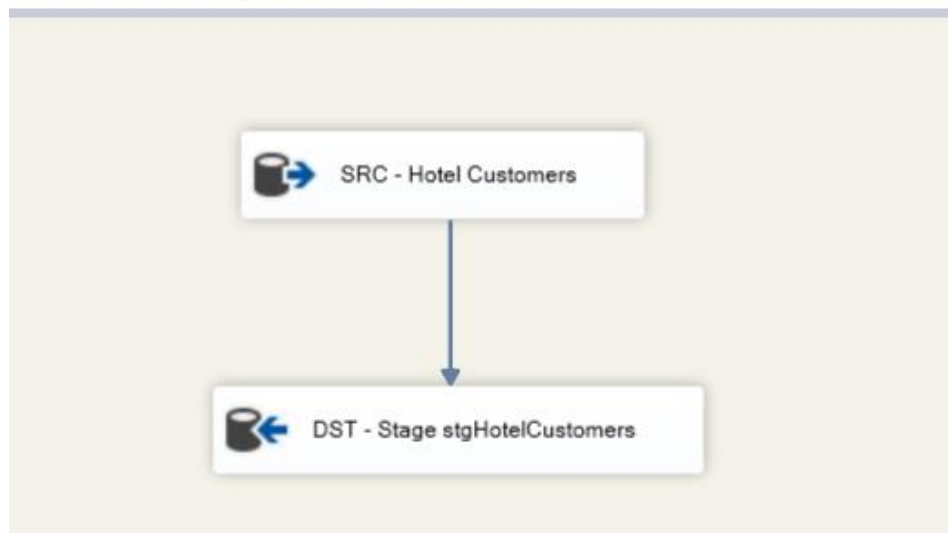
3.3.1.1. Tạo package mới và thiết lập control flow

❖ Tạo một Execute SQL Task và 3 Data Flow Task

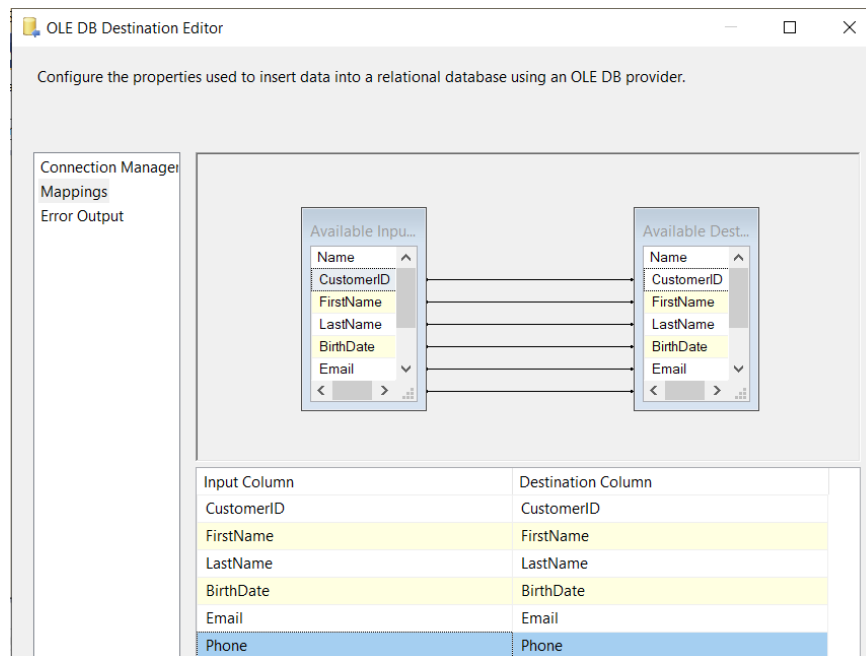


3.3.1.2. Load dữ liệu từ nguồn vào bảng stgHotelCustomers

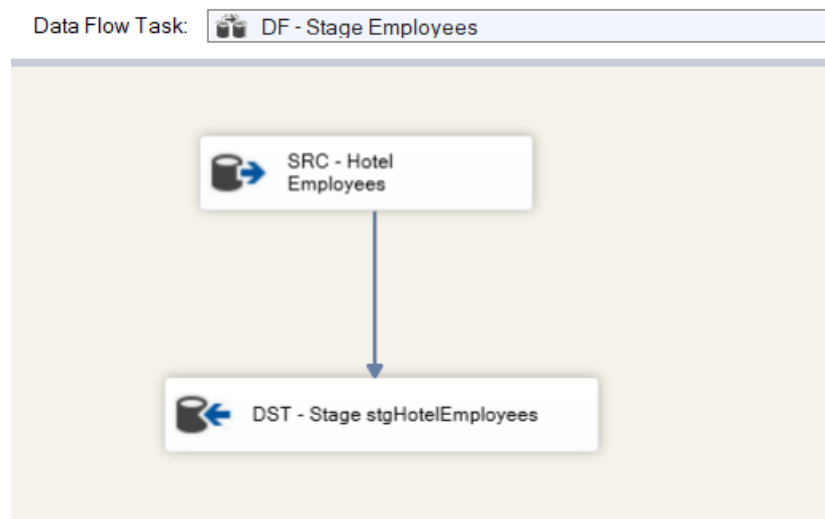
Data Flow Task: DF - Stage Customers



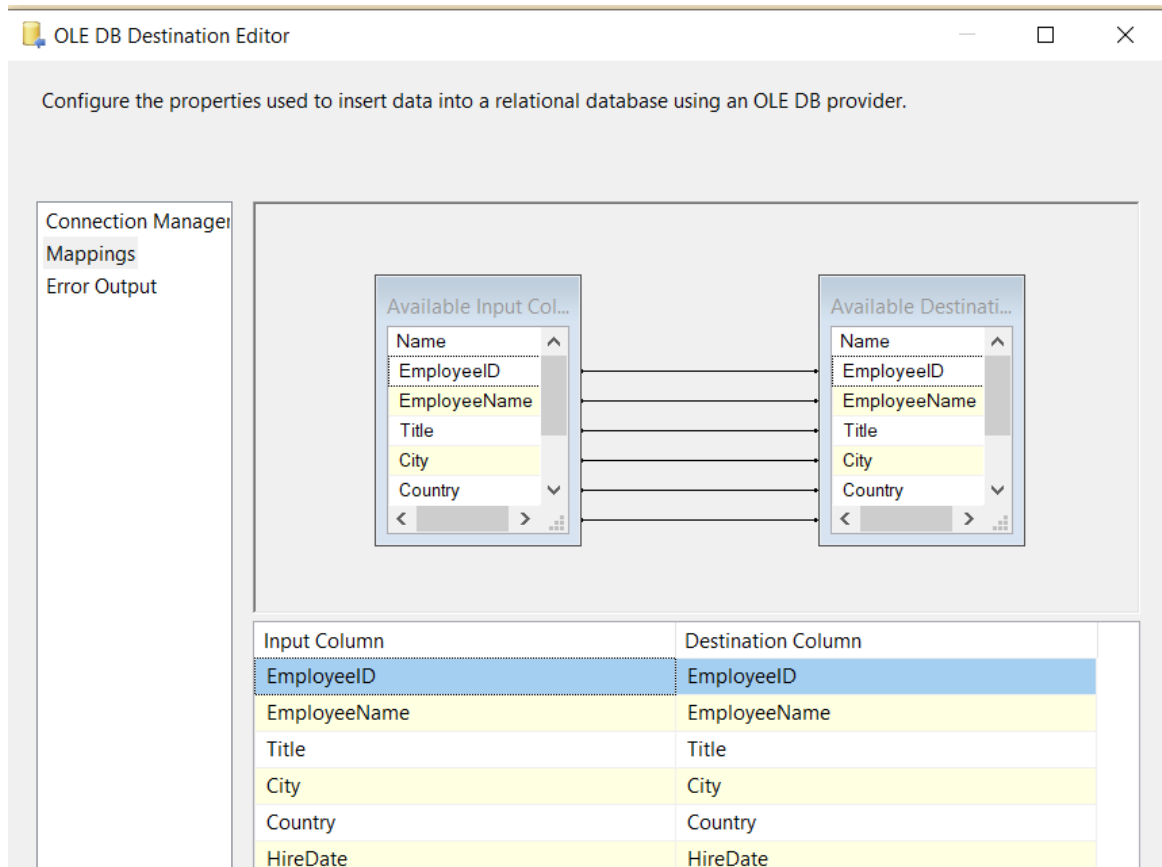
❖ Kiểm tra mapping giữa source và destination



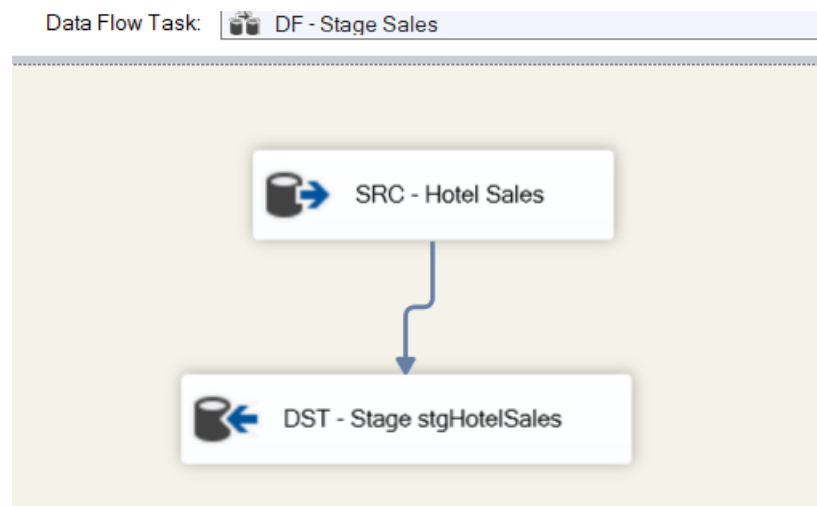
3.3.1.3. Load dữ liệu từ nguồn vào bảng stgHotelEmployees



❖ Kiểm tra mapping giữa source và destination



3.3.1.4. Load dữ liệu từ nguồn vào bảng stgHotelSales



❖ Dùng SQL command để kết bảng cho stgHotelSales

OLE DB connection manager:

DESKTOP-KC6BO6E\SQL2019.Hotel 2

New...

Data access mode:

SQL command

SQL command text:

```
SELECT b.CustomerID, b.EmployeeID, b.BookingID,
p.PaymentDate, b.RoomNumber, r.RoomType, r.BedType,
r.TypeCost, r.BedCost, b.GuestCount, p.Discount, p.TaxRate
FROM [Hotel].[dbo].[Payments] p JOIN [Hotel].[dbo].[Bookings]
b ON p.BookingID=b.BookingID
      JOIN (SELECT r.RoomNumber, b.BedType, b.BedCost,
rt.RoomType, rt.TypeCost FROM [Hotel].[dbo].[Rooms] r JOIN
[Hotel].[dbo].[BedTypes] b ON
r.BedTypeID=b.BedTypeID JOIN
[Hotel].[dbo].[RoomTypes] rt ON
r.RoomTypeID=rt.RoomTypeID) r ON
b.RoomNumber=r.RoomNumber
```

Parameters...

Build Query...

Browse...

Parse Query

❖ Kiểm tra mapping giữa source và destination

OLE DB Destination Editor

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

Connection Manager

Mappings

Error Output

Available Input ...

Name
CustomerID
EmployeeID
BookingID
RoomNumber
RoomType
BedType
TypeCost
BedCost
GuestCount

Available Destin...

Name
CustomerID
EmployeeID
BookingID
RoomNumber
RoomType
BedType
TypeCost
BedCost
GuestCount

Input Column	Destination Column
CustomerID	CustomerID
EmployeeID	EmployeeID
BookingID	BookingID
RoomNumber	RoomNumber

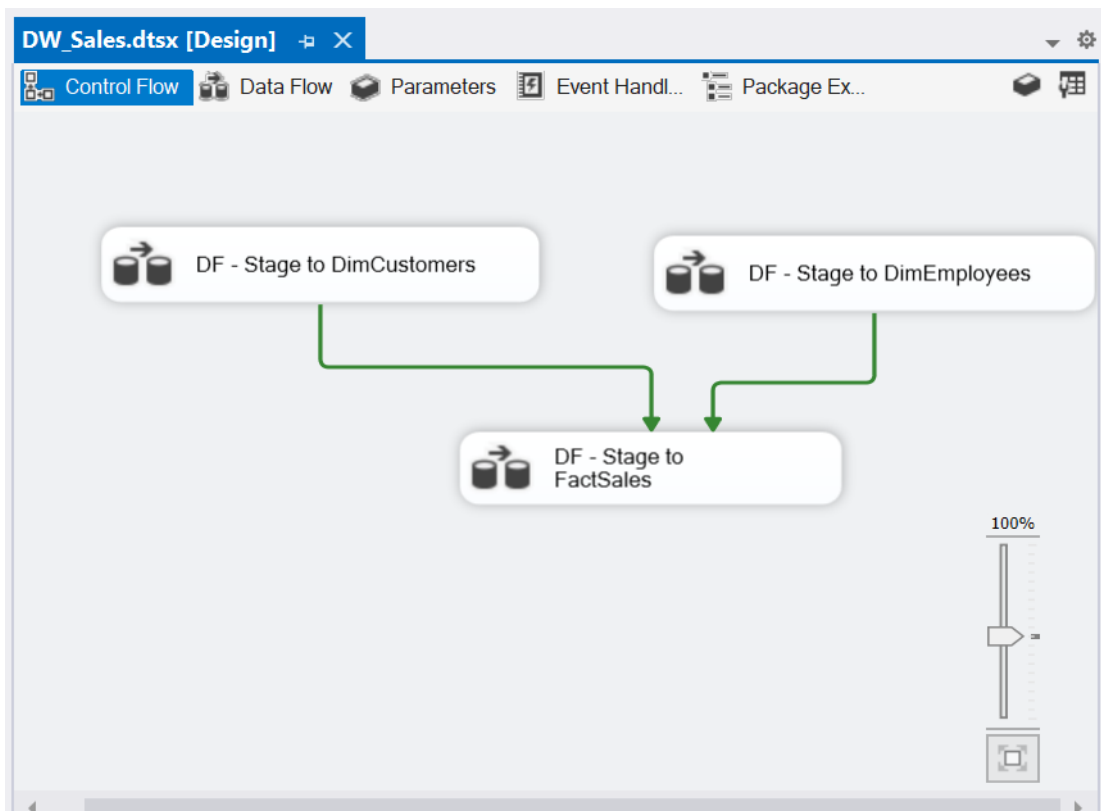
3.3.1.5. Thiết lập truncate table SQL task

▼ Options	
TimeOut	0
CodePage	1252
TypeConversionMode	Allowed
▼ Result Set	
ResultSet	None
▼ SQL Statement	
ConnectionType	OLE DB
Connection	DESKTOP-KC6BO6E\SQL2019.HotelStage
SQLSourceType	Direct input
SQLStatement	truncate table stgHotelCustomers; ;truncate
IsQueryStoredProcedure	truncate table stgHotelEmployees;
BypassPrepare	truncate table stgHotelSales;

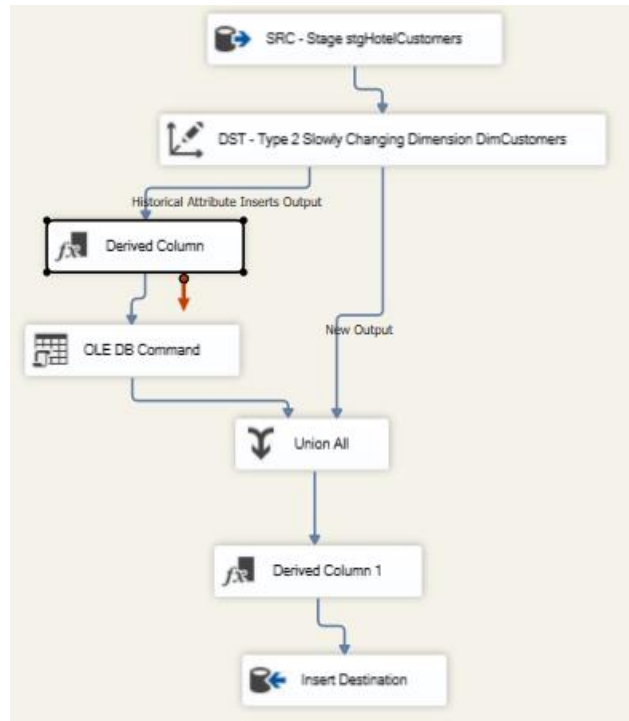
Name
Specifies the name of the task.

3.3.2. Load dữ liệu vào bảng FactSales

3.3.2.1. Tạo package mới và thiết lập control flow



3.3.2.2. Load dữ liệu từ stgCustomers vào DimCustomers



❖ DST - Type 2 Slowly Changing Dimension DimCustomers

Slowly Changing Dimension Wizard

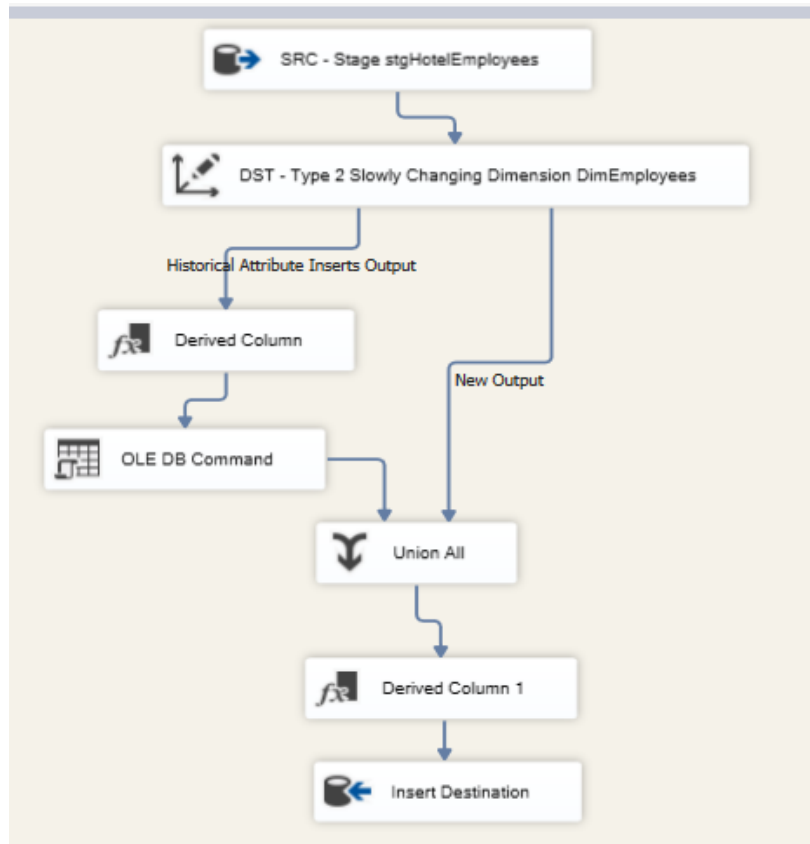
Select a Dimension Table and Keys
Select a dimension table to load and map columns in the transformation input to columns in the dimension table.

Connection manager:
DESKTOP-UQMD91R.HotelDW [New...]

Table or view:
[hotel].[DimCustomers]

Input Columns	Dimension Columns	Key Type
BirthDate	BirthDate	Not a key column
CustomerID	CustomerID	Business key
Email	Email	Not a key column
FirstName	FirstName	Not a key column
LastName	LastName	Not a key column
Phone	Phone	Not a key column
	RowChangeReason	
	RowEndDate	

3.3.2.3. Load dữ liệu từ stgEmployees vào DimEmployees



❖ DST - Type 2 Slowly Changing Dimension DimEmployees

Slowly Changing Dimension Wizard

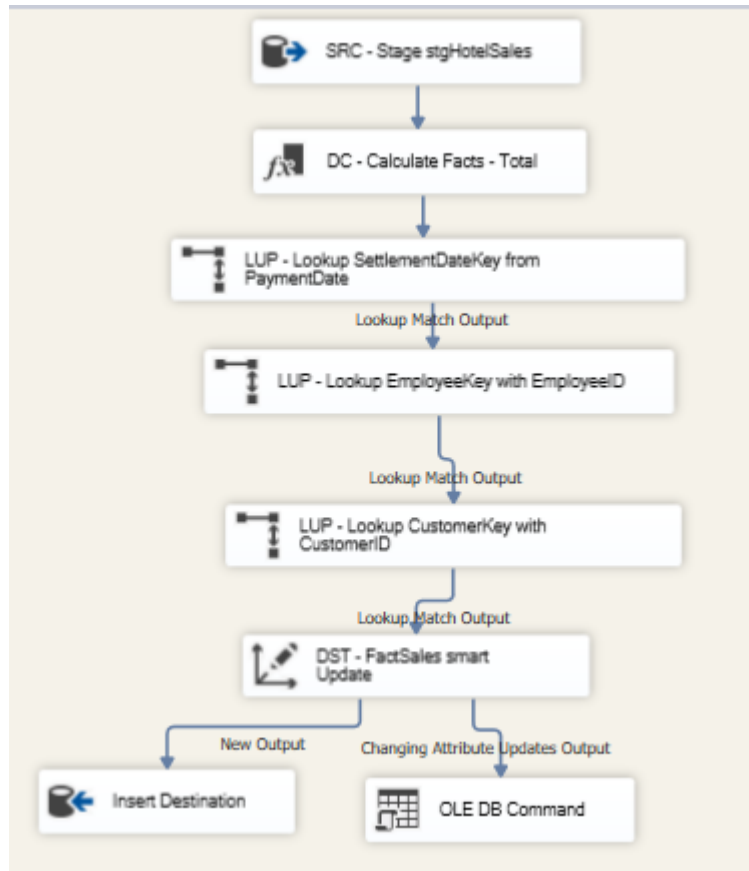
Select a Dimension Table and Keys
Select a dimension table to load and map columns in the transformation input to columns in the dimension table.

Connection manager:
DESKTOP-UQMD91R.HotelDW ▼ New...

Table or view:
[hotel].[DimEmployees] ▼

Input Columns	Dimension Columns	Key Type
City	City	Not a key column
Country	Country	Not a key column
EmployeeID	EmployeeID	Business key ▼
EmployeeNa...	EmployeeName	Not a key column
HireDate	HireDate	Not a key column
	RowChangeReason	
	RowEndDate	
	RowIsCurrent	

3.3.2.4. Load dữ liệu từ stgSales vào bảng FactSales



❖ Tính toán tổng doanh thu của khách sạn (Total)

Derived Column Transformation Editor

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

Variables and Parameters

Columns

Mathematical Functions

String Functions

Date/Time Functions

NULL Functions

Type Casts

Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Le
Total	<add as new column>	(DT_CY)(((BedCost + TypeCost) * (1 + TaxRate)) - ...	currency [DT_CY]	

❖ Tìm CustomerKey từ CustomerID

Lookup Transformation Editor

This transform enables the performance of simple equi-joins between the input and a reference data set.

General
Connection
Columns
Advanced
Error Output

Available Input ...

Name
CustomerID
EmployeeID
BookingID
RoomNumber
RoomType
BedType
TypeCost
BedCost
GuestCount
Discount

Available Lookup Col...

<input type="checkbox"/> Name	I...
<input checked="" type="checkbox"/> CustomerKey	
<input type="checkbox"/> CustomerID	
<input type="checkbox"/> FirstName	
<input type="checkbox"/> LastName	
<input type="checkbox"/> BirthDate	
<input type="checkbox"/> Email	
<input type="checkbox"/> Phone	
<input type="checkbox"/> RowIsCurre...	
<input type="checkbox"/> RowStartD...	
<input type="checkbox"/> RowEndDate	

Lookup Column	Lookup Operation	Output Alias
CustomerKey	<add as new column>	CustomerKey

OK Cancel Help

❖ Tìm EmployeeKey từ EmployeeID

Lookup Transformation Editor

This transform enables the performance of simple equi-joins between the input and a reference data set.

General
Connection
Columns
Advanced
Error Output

Available Input ...

Name
CustomerID
EmployeeID
BookingID
RoomNumber
RoomType
BedType
TypeCost
BedCost
GuestCount
Discount

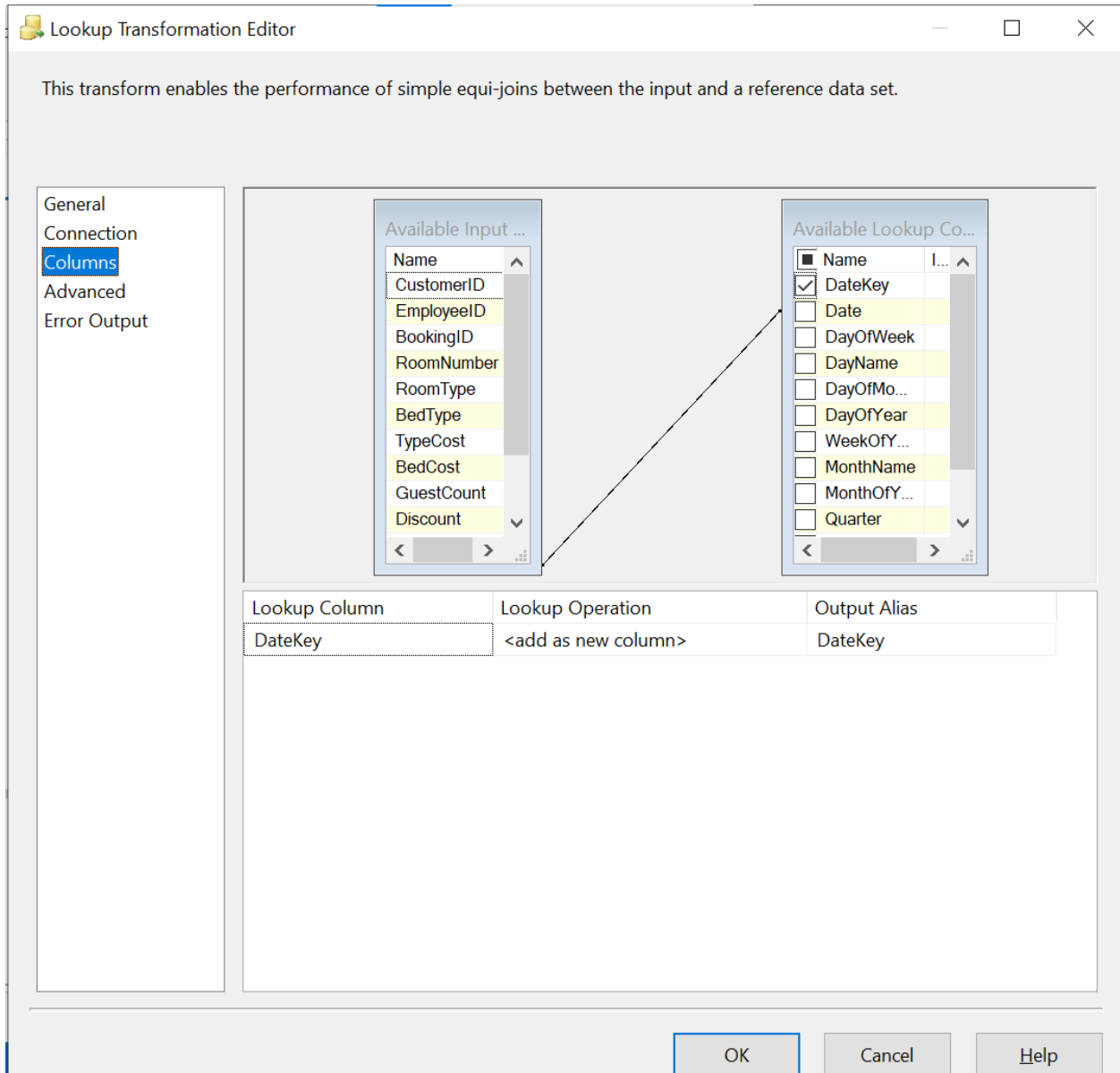
Available Lookup Colu...

<input type="checkbox"/> Name	I...
<input checked="" type="checkbox"/> EmployeeKey	
<input type="checkbox"/> EmployeeID	
<input type="checkbox"/> EmployeeName	
<input type="checkbox"/> Title	
<input type="checkbox"/> City	
<input type="checkbox"/> Country	
<input type="checkbox"/> HireDate	
<input type="checkbox"/> RowsCurrent	
<input type="checkbox"/> RowStartDate	
<input type="checkbox"/> RowEndDate	

Lookup Column	Lookup Operation	Output Alias
EmployeeKey	<add as new column>	EmployeeKey

OK Cancel Help

❖ TÌm SettlementDateKey từ PaymentDate

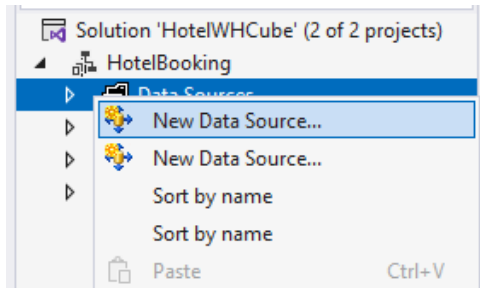


CHƯƠNG 4: PHÂN TÍCH DỮ LIỆU (SSAS)

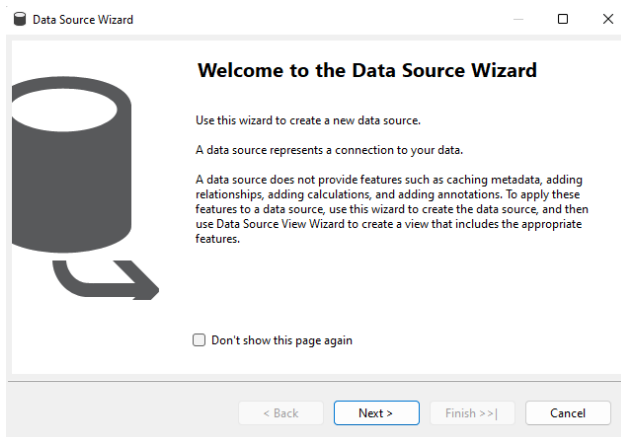
4.1. Xây dựng mô hình

4.1.1. Tạo Data Source

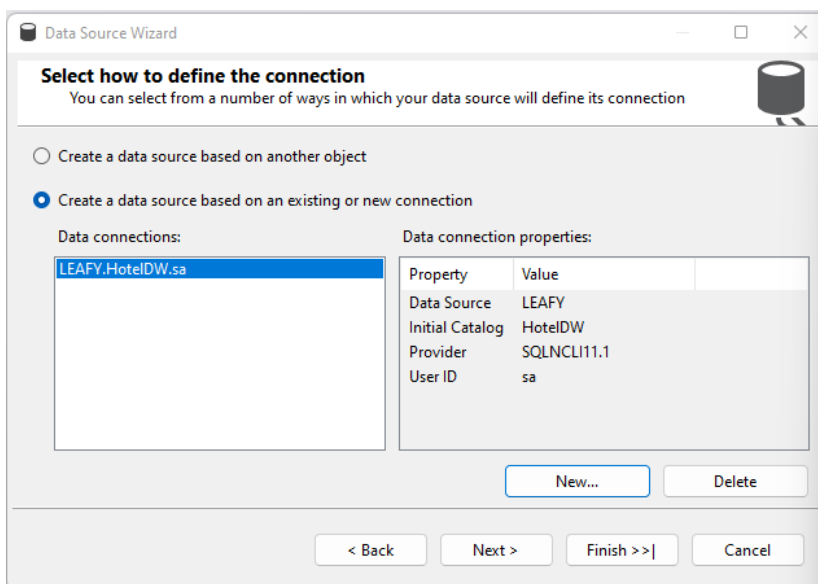
Right Click vào Data Source chọn New Data Source



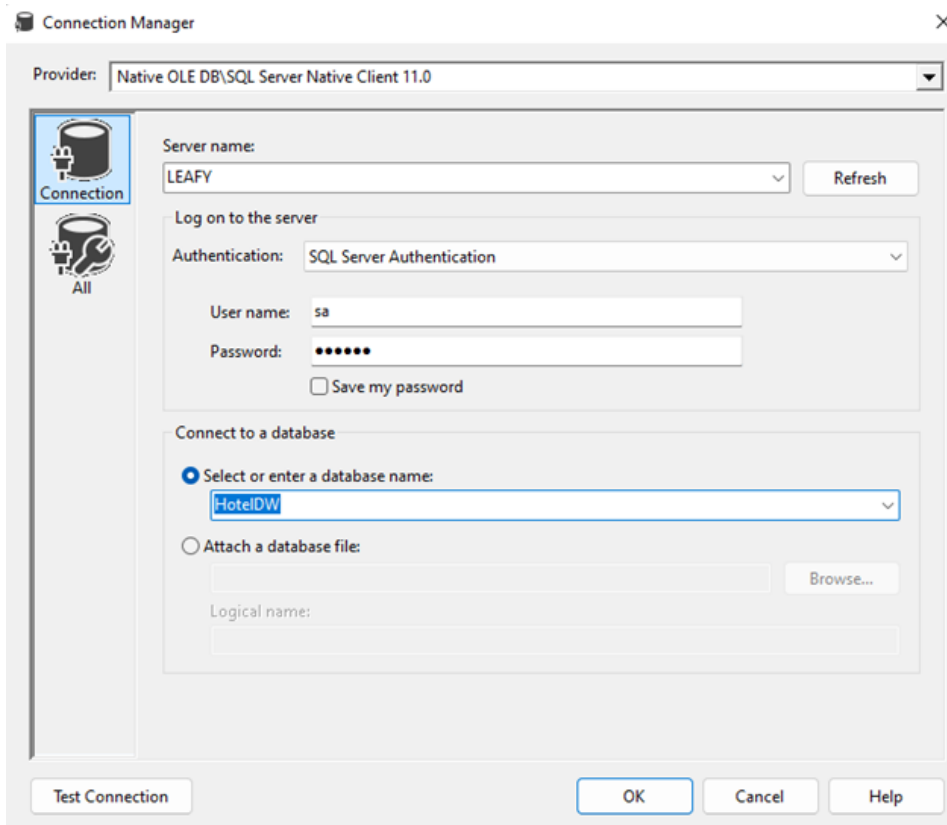
Một màn hình thông báo hiện lên, chọn Next để thực hiện các bước tiếp theo



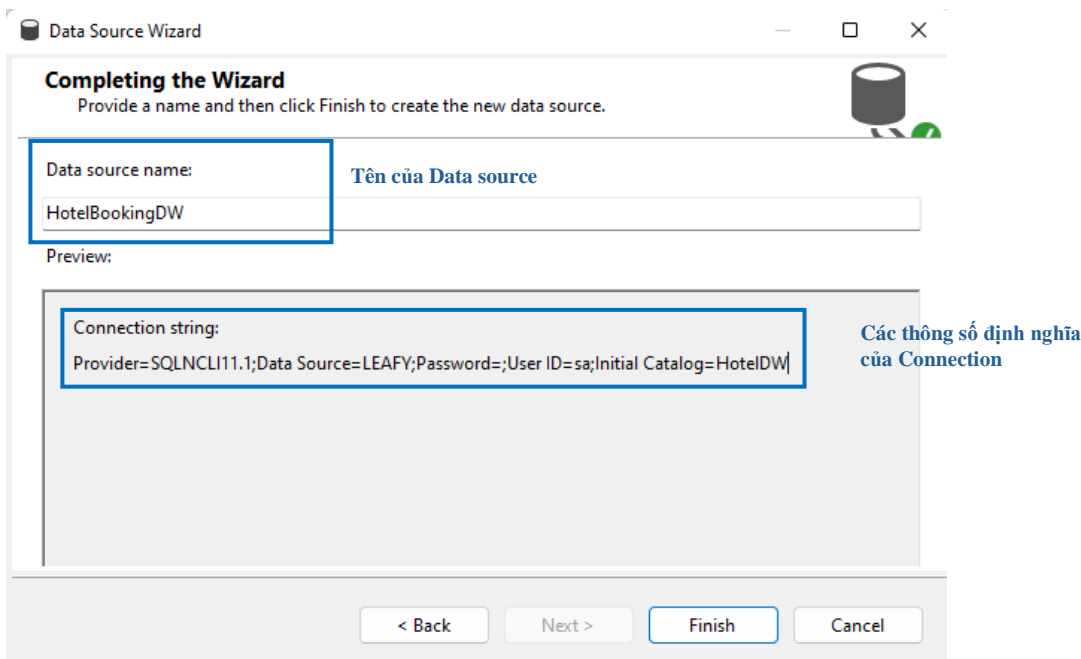
Tạo connection với data source



Nếu cần tạo một connection mới thì chọn New để tạo connection với các tùy chọn thích hợp để kết nối vào SQL Server

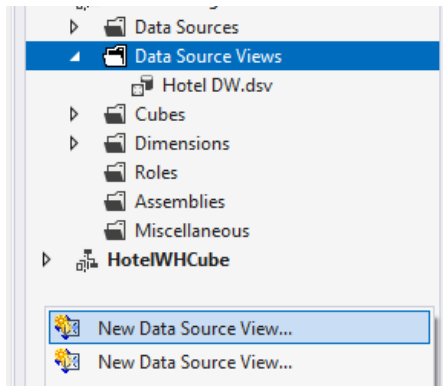


Tiếp theo khi tạo xong Connection chọn Next. Tại đây tiến hành đặt tên cho Data Source sau đó chọn Finish để kết thúc cấu hình và tạo một Data Source mới

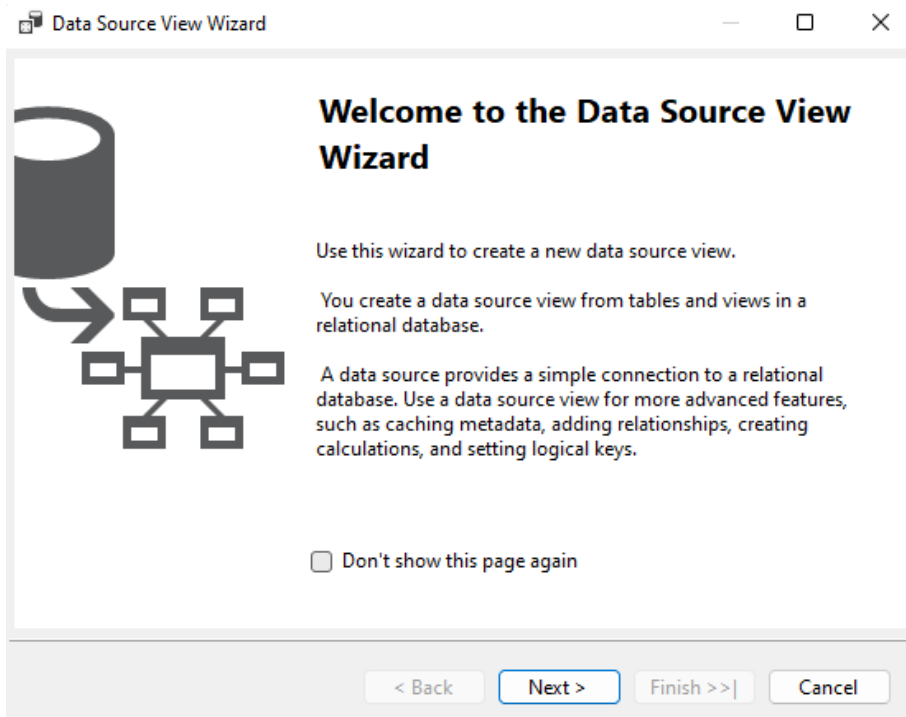


4.1.2. Tạo Data Source View

Right click Data Source Views chọn New Data Source View

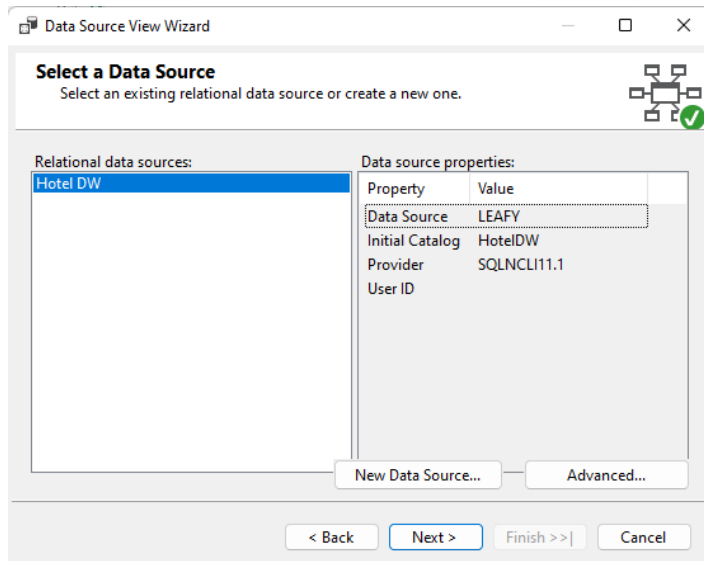


Một màn hình thông báo hiện lên, chọn Next

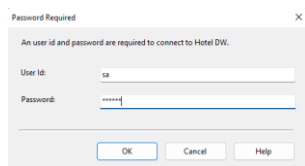


Tiếp theo chọn một Data Source để tạo View, Data source ở trên vừa tạo sẽ được lựa chọn, nếu chưa có chọn New để thực hiện tiến trình tạo một Data source mới.

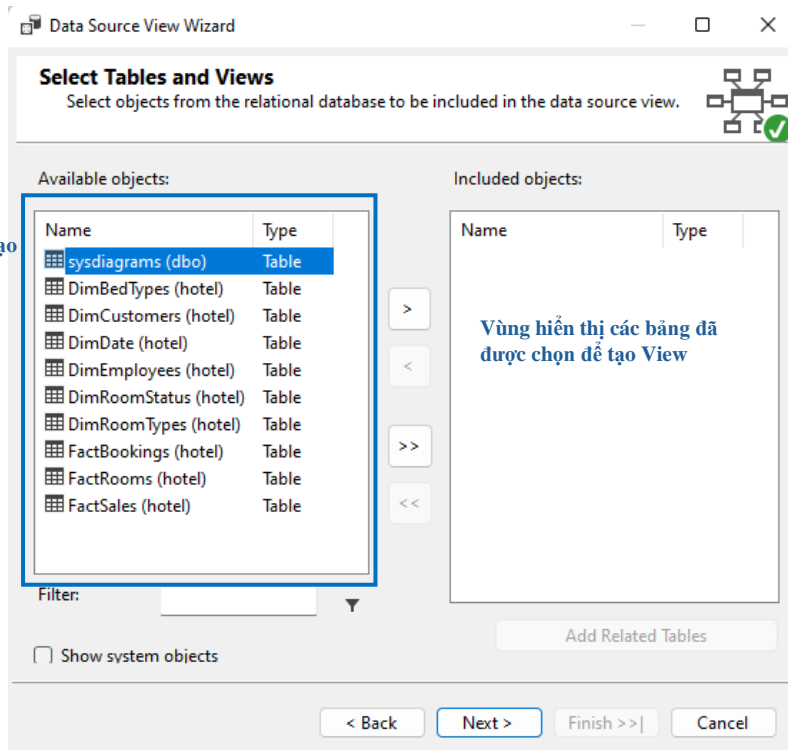
Sau đó chọn Next.



Có thể sẽ yêu cầu nhập Password khi tạo kết nối nếu cơ sở dữ liệu của bạn có tạo mật khẩu bảo mật



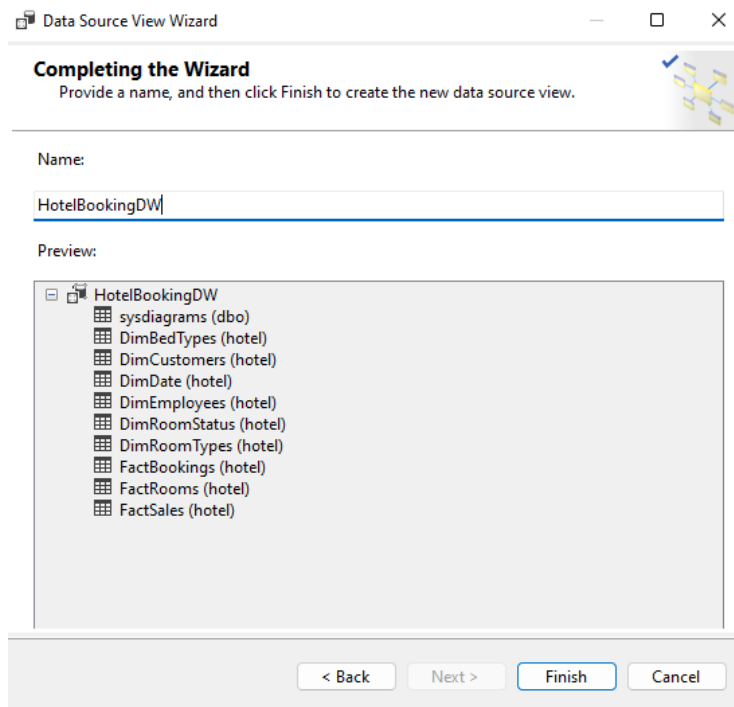
Tiếp theo chọn các Table cần thiết để tạo View. Trong mục **Included objects**, chọn các bảng Dim và Fact cần thiết



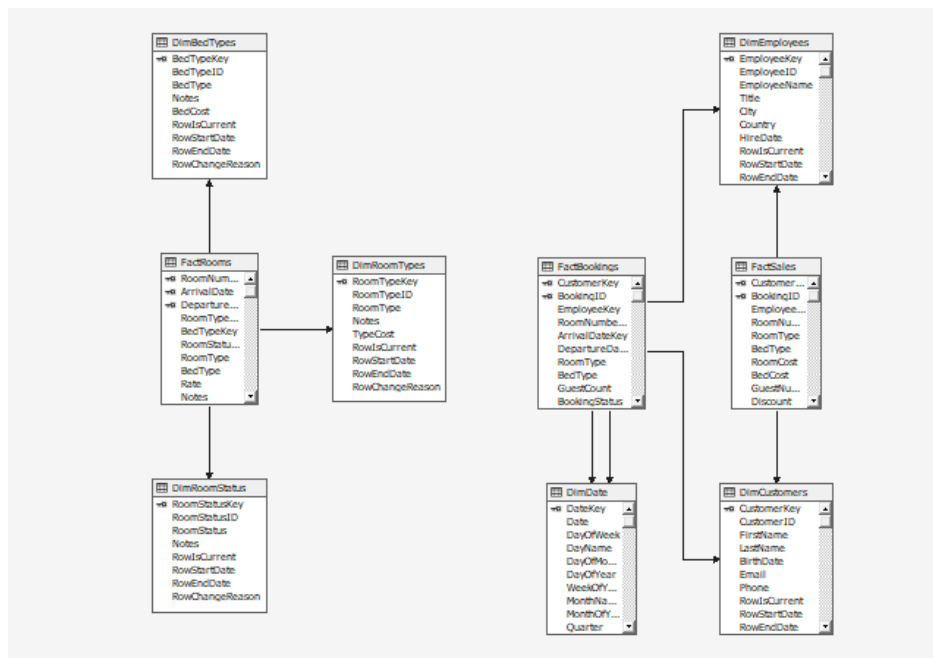
Các bảng mà đã kết nối được và có thể chọn để tạo view

Vùng hiển thị các bảng đã được chọn để tạo View

Tiếp theo tiến hành đặt tên DataView, sau đó chọn Finish để kết thúc tiến trình



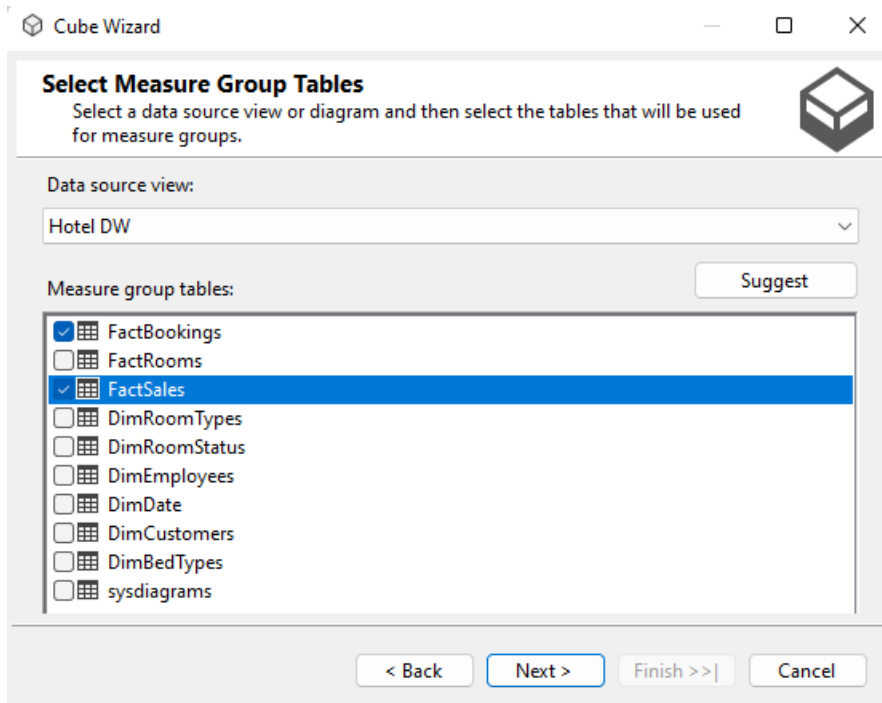
Double-Click vào Data Source View vừa tạo để hiển thị Diagram



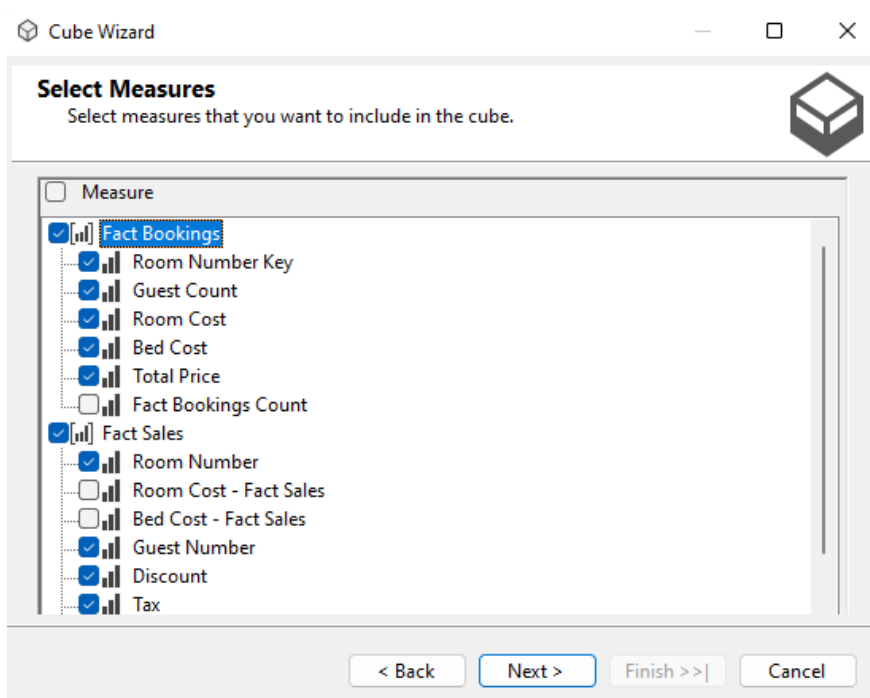
4.2. Tạo Data cube

Right-click vào Cubes ở khung Solution Explorer. Chọn New Cube để tạo Cube mới.

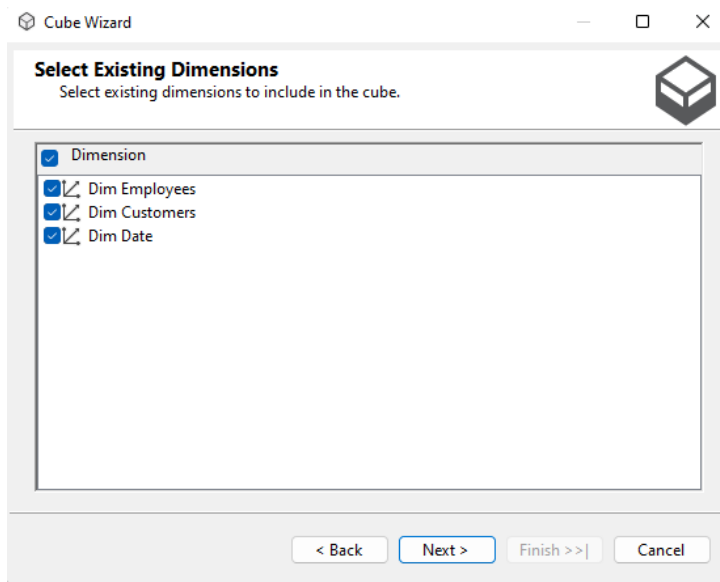
Chọn Data Source View là Hotel DW View . Tích chọn 2 bảng FactSales và FactBooking, nhấn Next



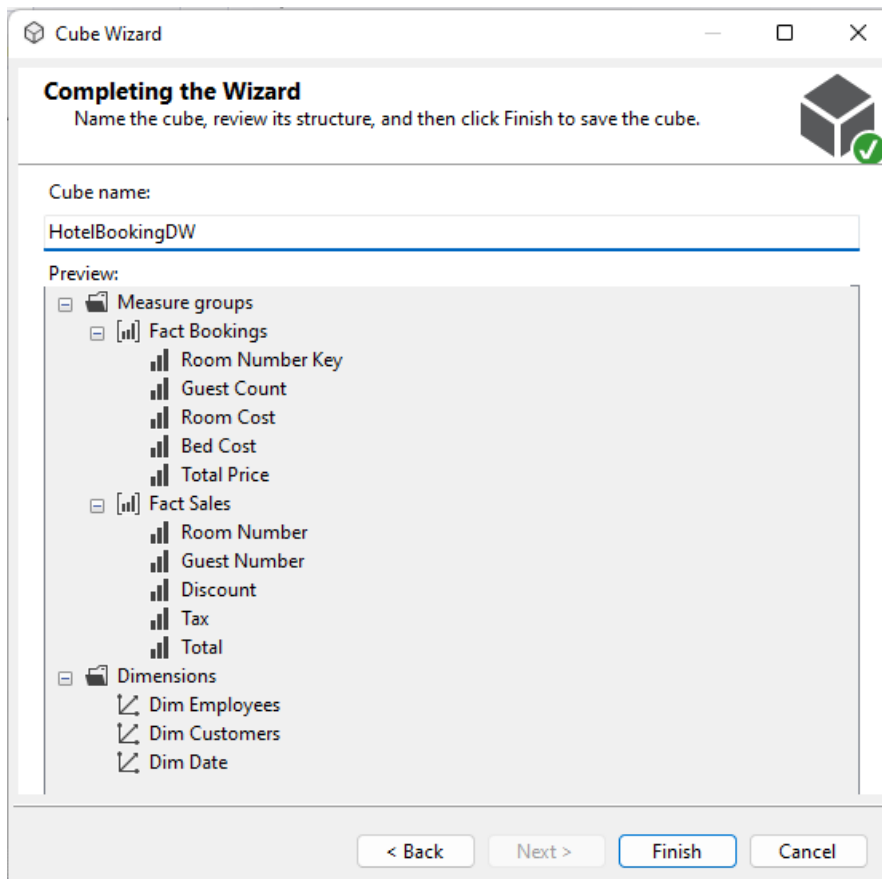
Ở dialog Select Measures, tích chọn những Measure cần thiết



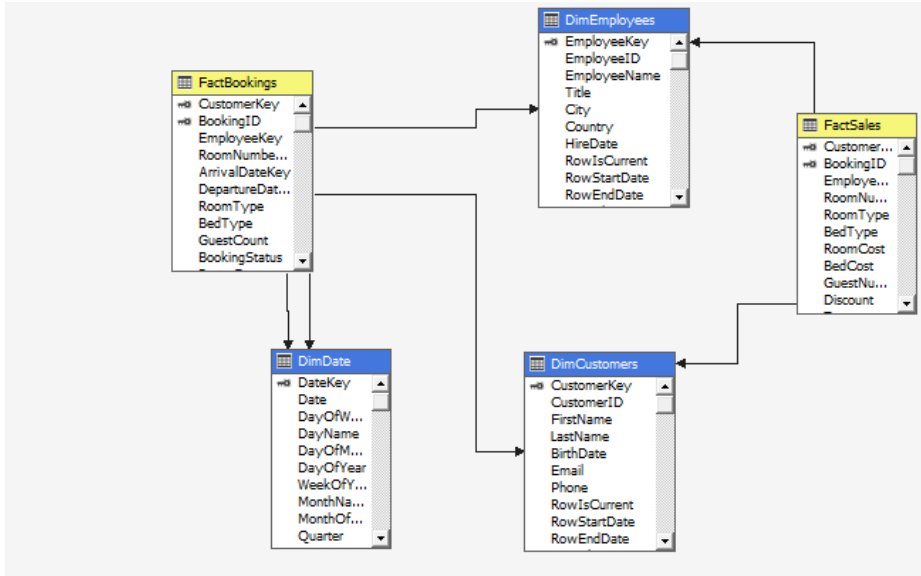
Khởi tạo các Dimension



Cuối cùng đặt tên Data cube để kết thúc tiến trình



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

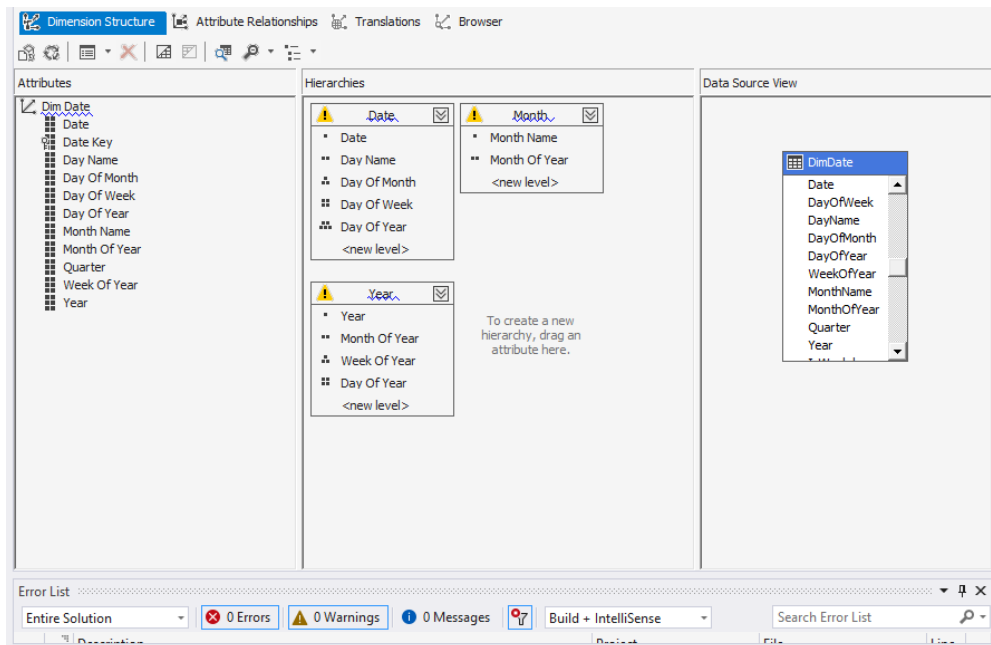


4.3. Cấu hình Hierarchy

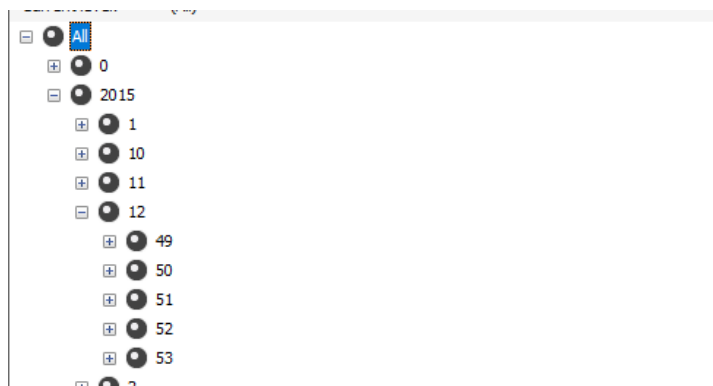
4.3.1. Tạo Hierarchy Dim Date

Thêm thuộc tính vào bảng Dim: Tại table DimDate trong panel Data Source View, kéo thuộc tính cần thiết vào panel Attributes.

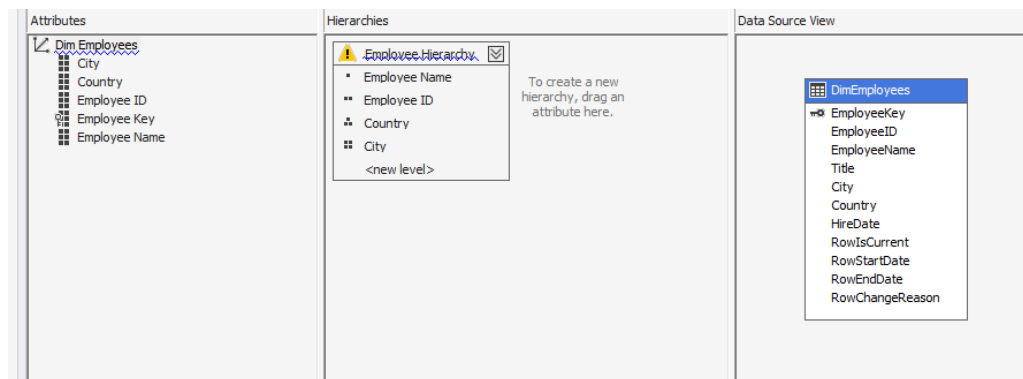
Kéo thả các thuộc tính Date, Month và Year từ Attributes vào panel Hierarchies để lần lượt tạo ra các Date, Month và Year Hierarchy.



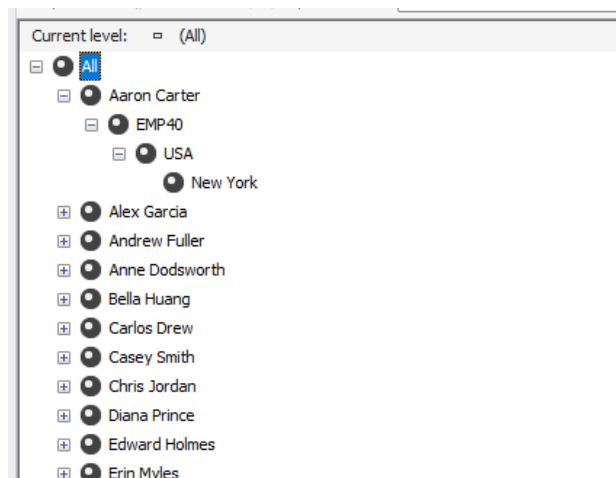
Quan sát thấy trong Year Hierarchy, dữ liệu được sắp xếp theo thứ tự: Năm → tháng trong năm → số tuần trong tháng giống với Year Hierarchy đã cấu hình.



4.3.2. Tạo Hierarchy cho Dim Employees



Vào Browser kiểm tra thấy các thông tin đã được sắp xếp theo thứ tự, Tên nhân viên → Mã nhân viên → Quốc gia → Thành phố



4.3.3. Tạo Hierarchy cho Dim Customers

The screenshot shows the SQL Server Enterprise Manager interface. On the left, the 'Hierarchy' pane shows a tree structure for 'DimCustomers' with attributes: Customer ID, First Name, Last Name, Birth Date, Email, Phone, and a '<new level>' button. Below this is a table of attributes with columns: Name, Usage, Type, Key Column, and Name Column. The table lists attributes like Birth Date, Customer ID, Customer Key, Email, First Name, Last Name, and Phone, each with its usage, type, and key column. On the right, the 'DimCustomers' table is shown with its attributes. Below the main interface, the 'Hierarchy' pane shows a tree structure for 'Customer Hierarchy' with levels: (All), AB30861, Aluinwec, Braisheroa, Jan 15 1988 12:00AM, ngon_huynh@colpal.com, and 3783522281.

Name	Usage	Type	Key Column	Name Column
Birth Date	Regular	Regular	WChar	Same as key
Customer ID	Regular	Regular	WChar	Same as key
Customer Key	Key	Regular	Integer	Same as key
Email	Regular	Regular	WChar	Same as key
First Name	Regular	Regular	WChar	Same as key
Last Name	Regular	Regular	WChar	Same as key
Phone	Regular	Regular	Double	Same as key

Vào Browser kiểm tra thấy các thông tin đã được sắp xếp theo thứ tự, ID Customer → Name Customer → Country → City như đã cấu hình trong Hierarchy Customers

4.4. Thực hiện tạo Cube

Kết thúc các quá trình cấu hình và nhấn Process để tiến hành tạo Cube với các cấu hình. Và vào Browser để thực hiện các phân tích dữ liệu.

The screenshot shows the 'Deployment Progress - HotelWHCube' window. It displays the server 'LEAFY\LEAFYCUBE' and the database 'HotelWHCube'. The 'Command' pane shows a list of tasks completed, including processing dimensions and the cube itself. The status bar at the bottom indicates 'Deployment Completed Successfully' with a green checkmark.

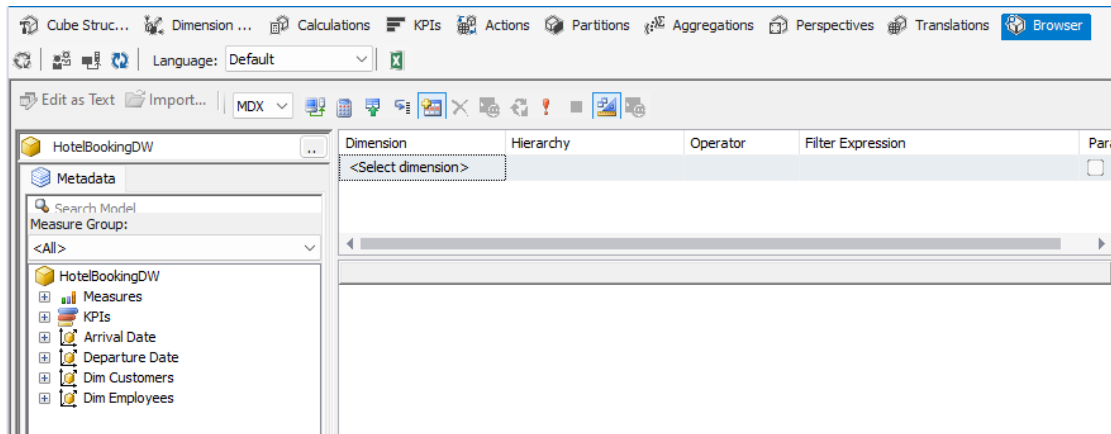
Server: LEAFY\LEAFYCUBE
Database: HotelWHCube

Command

- Processing Database 'HotelWHCube' completed.
- Processing Dimension 'Dim Bed Types' completed.
- Processing Dimension 'Dim Customers' completed.
- Processing Dimension 'Dim Date' completed.
- Processing Dimension 'Dim Employees' completed.
- Processing Dimension 'Dim Room Types' completed.
- Processing Cube 'Hotel DW' completed.
Start time: 5/16/2024 3:00:06 PM; End time: 5/16/2024 3:00:09 PM; Duration: 0:00:02
- Processing Measure Group 'Fact Sales' completed.
Start time: 5/16/2024 3:00:06 PM; End time: 5/16/2024 3:00:10 PM; Duration: 0:00:03

Status:

Deployment Completed Successfully



4.5. Thực hiện phân tích dữ liệu

4.5.1. Sử dụng SSAS

Câu hỏi: Cho biết tổng Doanh thu tháng 5, 6, 7 năm 2015 tại khách sạn

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Arrival Date	Arrival Date.Year Hierarchy	Equal	{ 2015 }	<input type="checkbox"/> <input type="checkbox"/>
Arrival Date	Arrival Date.Month Hierarchy	Equal	{ June, May, July }	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>

Room Cost	Bed Cost	Tax	Discount	Total
15580	8242.349999999999	7.920000000000001	22.35485	18387.6989

Khi dùng Cube, tất cả doanh thu trong điều kiện đã chọn là tháng 5, 6, 7 và năm 2015 được lọc ra và tổng lại thành 1 giá trị duy nhất, giá tổng sẽ bao gồm giá phòng + loại giường * phí thuế * discount để ra total cuối cùng

Từ doanh thu có thể

Câu hỏi: Cho biết doanh số bán hàng theo từng khách hàng trong tháng 10 năm 2017

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Arrival Date	Arrival Date.Year Hierarchy	Equal	{ 2017 }	<input type="checkbox"/> <input type="checkbox"/>
Arrival Date	Arrival Date.Month Hierarchy	Equal	{ October }	<input type="checkbox"/> <input type="checkbox"/>
Dim Customers	First Name	Equal	{ All }	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>

Các điều kiện được cấu hình từ các bảng Dim, Hierachy:

- Year = 2017
- Month = October

- Gôm nhóm theo tên của từng khách hàng

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Arrival Date	Arrival Date.Year Hierarchy	Equal	{ 2017 }	<input type="checkbox"/> <input type="checkbox"/>
Arrival Date	Arrival Date.Month Hierarchy	Equal	{ October }	<input type="checkbox"/> <input type="checkbox"/>
Dim Customers	First Name	Equal	{ All }	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>

Customer ID	First Name	Last Name	Total
AB30861	Aluinwec	Braisheroa	149.5931
AB41486	Aluin	Braisher	398.4361
AC79218	Ambrosio	Catto	210.1638
AD54630	Adrianolif...	Davidwitzlevoa	165.7631
AD68400	Adrianolif	Davidwitzlev	128.0286
AD68525	Adrianowec	Davidwitzoa	128.7574
AE71356	Alisonlifwec	Elwynlevoa	133.4197
AH11563	Audrielifwec	Herreroslevoa	201.3042
AJ34729	Adamolifwec	Joannidevoa	148.1768
AK83797	Aldrich	Kale	220.5098
AM43455	Antinlif	McKielev	128.1003
AM46198	Antin	McKie	201.2188
AR17884	Arleen	Rozier	270.6397
AV44498	Arawec	Velaoa	185.6653
BM87898	Bobbywec	MacLoughlinoa	140.5936
BN25232	Berk	Dasee	164.7362

Từ truy vấn, có thể thấy được kết quả trả về theo cấu trúc đã yêu cầu, bao gồm Customer ID, First Name, Last Name và Total Sales

Từ truy vấn trên chúng ta có thể tương tự và xác định được nhiều vấn đề cần thiết như:

- So sánh doanh số tháng 10 với các tháng trước có thể cho thấy xu hướng tăng hay giảm doanh số. Nếu doanh số của tháng 10 cao hơn các tháng trước, điều này có thể cho thấy hiệu quả của các chiến dịch marketing hoặc sự gia tăng nhu cầu của thị trường
- Đề xuất các chương trình ưu đãi đặc biệt hoặc dịch vụ khách hàng cao cấp cho những khách hàng có doanh số cao,...

4.5.2. Sử dụng Pivot Table trong Excel và Power BI

4.5.2.1. Câu hỏi: Cho biết xếp hạng các nhân viên nhận booking khách hàng theo quý/ tháng

➤ Sử dụng Pivot Table trong Excel

Hiển thị theo thời gian quý 3

M	N	O	P	Q
Tổng của GuestCount	Nhân cột			Tổng Cuối
	3			
Nhân Hàng	August	July	September	
Aaron Carter	7	2	5	14
Alex Garcia	4	4	7	15
Andrew Fuller	5	3	6	14
Anne Dodsworth	4	5	7	16
Bella Huang		17		17
Carlos Drew	17			17
Casey Smith	12	1	5	18
Chris Jordan	4	9	8	21
Diana Prince	5	8	8	21
Edward Holmes	6	8	3	17
Erin Myles	12	3	2	17
Fiona Barrett	10	7	5	22
Fiona Graham	7	11	4	22
George Elton	3	5	6	14
Hana Sato	6	5	5	16
Hugh Densmore	4	5	3	12
Ian Vance	4	1	6	11
Irene Schultz	5	10	3	18
Jake Fletcher	8	4		12
Jamie O'Neil	13	8	7	28
Janet Leverling	9	5	4	18
Jordan Pike	1	11	4	16
Julia Estevez	11	4	4	19
Kim Lee		16	3	19

Trường PivotTable

Hiện hoạtTất cả

Chọn các trường để thêm vào báo cáo:

Tim kiếm

Cus

Date

Emp

F_Booking

F_Room

F_Gate

Kéo trường giữa các vùng bên dưới:

Bộ lọc

Cột

Quarter

MonthName

Hàng

EmployeeName

Giá trị

Tổng của GuestCount

	M	N	O	P	Q
Nadia Singh		12	1		13
Nancy Davolio		26	8	8	42
Nina Patel		14		5	19
Omar Little		2	4	5	11
Oscar Wallace		9			9
Paige Sinclair		3	5	2	10
Pat Lee		15	5		20
Petra Ivanov		3	7	4	14
Quincy Adams		9	3		12
Quinn Johanson		10	4	3	17
Rachel Norman		16	8		24
Robert King		4	9	5	18
Rosa Martinez		4	10	10	24
Sam Patel		9	5		14
Seth Harper		6	3	12	21
Steve Jobs		4	14	2	20
Steven Buchanan		7	10		17
Tanya West		1	8	14	23
Tara Clark		8	8	6	22
Taylor Brooks		4	3	4	11
Uma Khan		11	4	1	16
Victor Dean		5	5	11	21
Wendy Miles		7		2	9
Xavier Neil		11	5	10	26
Yara Flint		3	5		8
Zoe Castello		4	5	4	13
Tổng Cuối		423	318	242	983

Trường PivotTable

Hiện hoạt Tất cả

Chọn các trường để thêm vào báo cáo:

Tìm kiếm

> Cus

> Date

> Emp

> F.Booking

> F.Sale

Kéo trường giữa các vùng bên dưới:

Bộ lọc

Cột

Quarter

MonthName

Hàng

Giá trị

EmployeeName

Tổng của GuestCount

Với hình trên thì ta cũng có thể thấy được số liệu nhận booking của các nhân viên theo quý 3 và theo các tháng 7, 8 và 9. Nhìn vào đây có thể thấy rằng doanh thu tháng 9 vào năm 2015 còn khá ít so với tháng 7 và tháng 8.

➤ Sử dụng Power BI

Biểu đồ xếp hạng các nhân viên nhận booking cao nhất theo quý/ tháng

MonthName

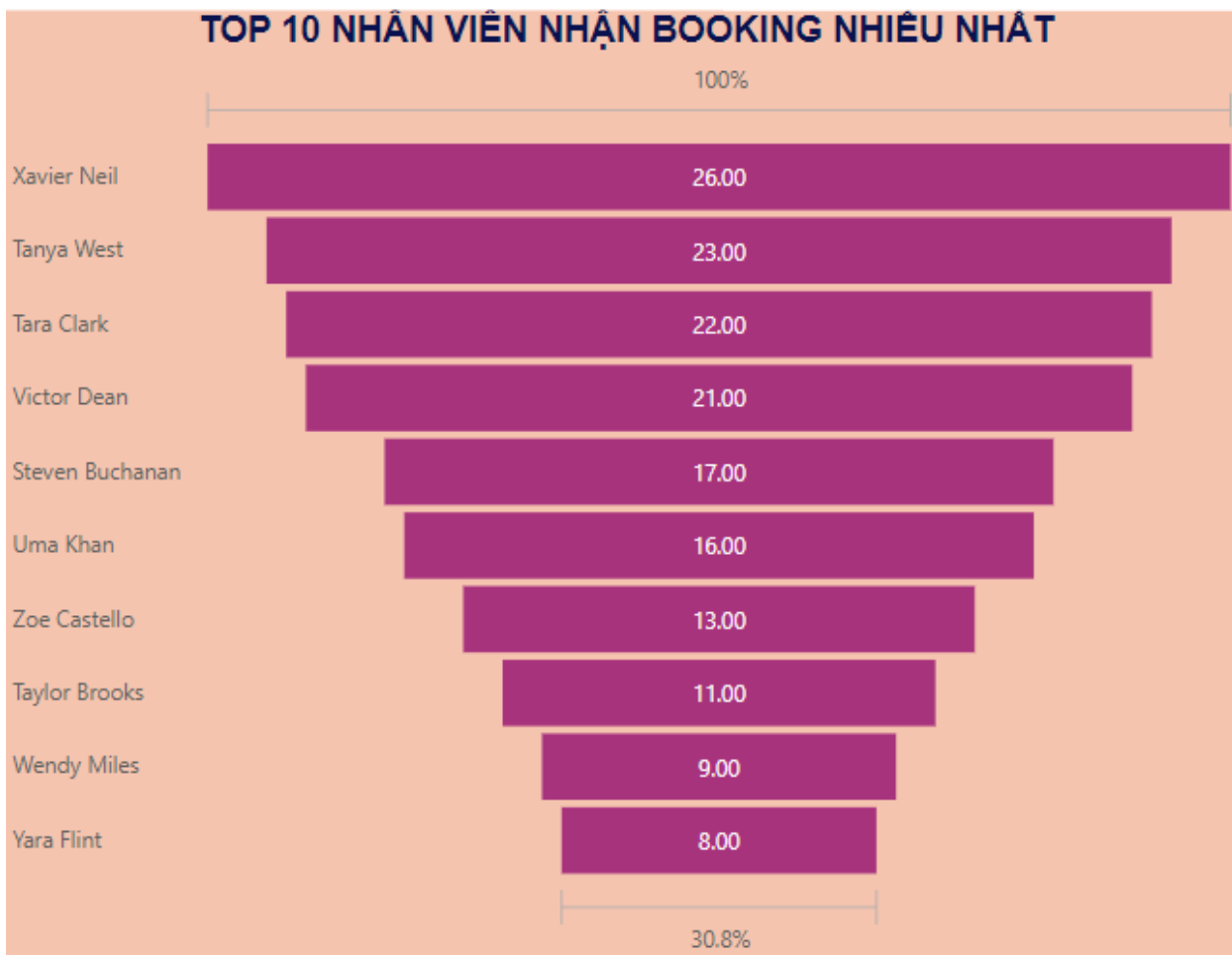
All

Quarter

All

983.00

Sum of GuestCount



Tổng số Khách của 10 Nhân viên dao động từ 8 đến 26.

Ở mức 26, Xavier Neil có Tổng số Khách cao nhất và cao hơn 225.00% so với Yara Flint, người có Tổng số Khách thấp nhất là 8. Xavier Neil chiếm 15.66% tổng số Khách.

4.5.2.2. Câu hỏi: Cho biết doanh thu trên mỗi loại phòng được đặt nhiều nhất theo quý/ tháng

➤ Sử dụng Pivot Table trong Excel

	I	J	K	L	M
Tổng của Total	Nhân cột				
Nhân Hàng	August	July	September	Tổng Cuối	
Apartment	194.0744	751.8506	261.7935	1207.7185	
Bungalow	329.444	381.0003	785.7845	1496.2288	
Chalet	1019.3864	1786.5076	456.9606	3262.8546	
Cottage	990.4014	568.3351	506.8925	2065.629	
Deluxe room	713.7728	603.8605	878.9703	2196.6036	
Penthouse			303.5771	303.5771	
Standard room	1447.5795	1182.9196	981.208	3611.7071	
Studio	682.734	509.5478	255.9616	1448.2434	
Suite	203.8296		551.766	755.5956	
Villa	338.6564	398.4361	1302.4487	2039.5412	
Tổng Cuối	5919.8785	6182.4576	6285.3628	18387.6989	

Trường PivotTable		Hiện hoạt	Tất cả
Chọn các trường để thêm vào báo cáo:			
Tìm kiếm			
>	Cus		
>	Date		
>	Emp		
>	F_Booking		
>	F_Sale		
Kéo trường giữa các vùng bên dưới:			
Bộ lọc	Cột		
	MonthName		
Hàng	Giá trị		
RoomType	Tổng của Total		

Hình trên hiển thị doanh thu của mỗi loại phòng vào tháng 7, 8 và 9. Ta thấy phòng Standard room được khách hàng ưa chuộng và đặt nhiều nhất. Phòng Penthouse được đặt ít nhất với tổng doanh thu là 303.5771. Tổng doanh thu của tất cả loại phòng theo tháng 7, 8 và 9 khá đều nhau. Điều này cho thấy rằng khách sạn có một lượng khách ổn định trong các tháng mùa hè (7, 8, 9). Sự đều đặn trong doanh thu có thể phản ánh hiệu quả của các chiến lược marketing, chất lượng dịch vụ ổn định hoặc không có biến động lớn về nhu cầu du lịch trong giai đoạn này.

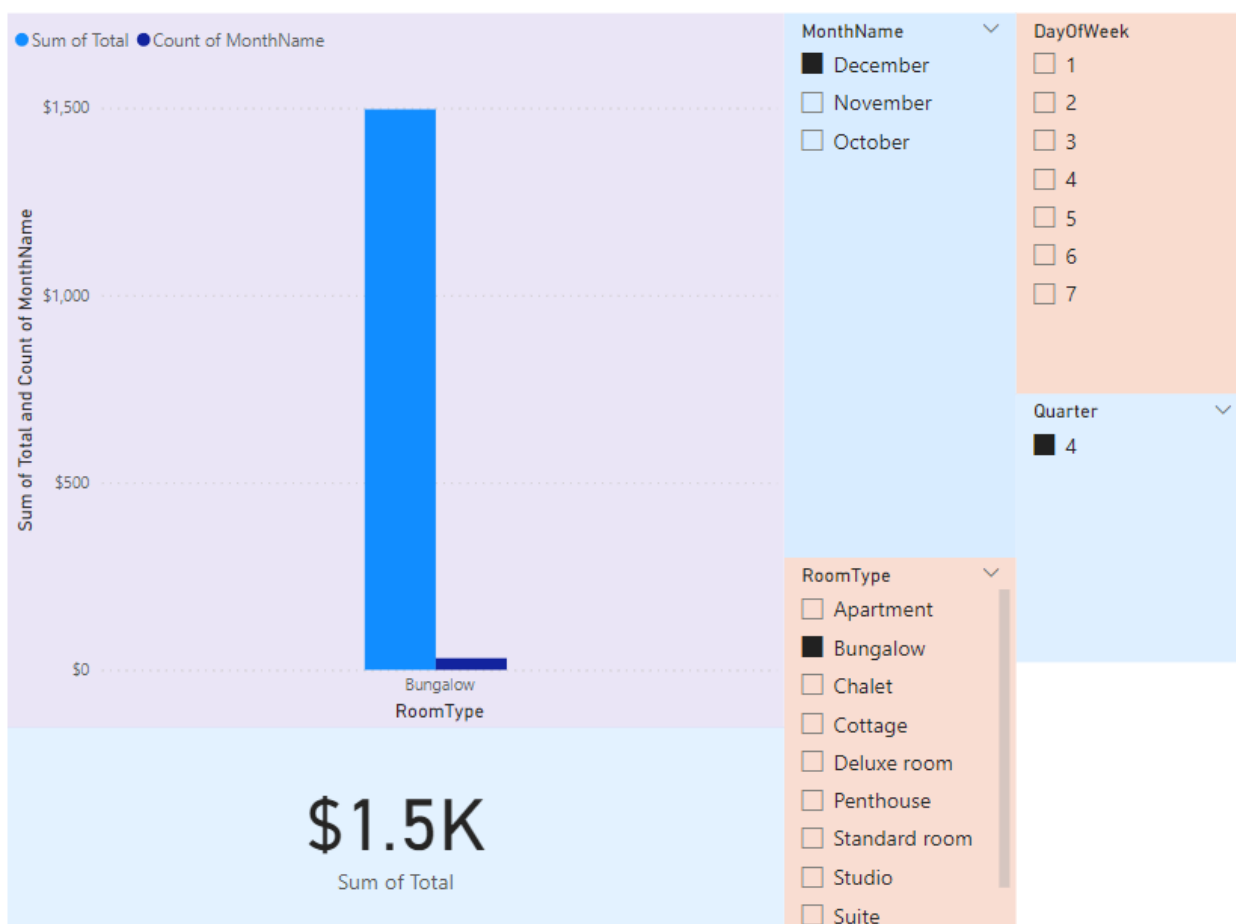
➤ Sử dụng Power BI

Xếp hạng loại phòng được đặt nhiều theo tháng.



Ở mức \$3,611.7071, phòng Standard có tổng số cao nhất và cao hơn 1,089.72% so với Penthouse, phòng có Tổng số thấp nhất là \$303.5771. Loại phòng theo tháng tương đối bằng nhau.

Ta cũng có thể chọn theo quý 4 tháng 12 và chọn loại phòng Bungalow để xem chi tiết về doanh thu theo tháng.



4.5.2.3. *Câu hỏi: Tỷ lệ trạng thái phòng mà khách hàng đặt tổng chi phí bao nhiêu*

Đã rời đi (Checked-out): Khách hàng đã rời khỏi khách sạn sau khi hoàn thành thời gian lưu trú.

Đã hủy (Cancelled): Đặt phòng đã bị hủy bởi khách hàng hoặc bởi hệ thống do không hoàn thành các yêu cầu cần thiết.

No-show : Trạng thái đặt phòng mà khách hàng đã không xuất hiện tại khách sạn

➤ **Sử dụng Pivot Table trong Excel**

K	L	M	N
Nhân Hàng	Weekday	Weekend	Tổng Cuối
Canceled			
Apartment	1065.89	604.23	1670.12
Bungalow	425.58	461.77	887.35
Chalet	2261.51	1699.91	3961.42
Cottage	2081.86	934.51	3016.37
Deluxe room	2246.75	1050.68	3297.43
Standard room	1287.67	658.26	1945.93
Studio	995.46	958.43	1953.89
Suite	340.32		340.32
Villa	1052.46	526.23	1578.69
Check-Out			
Apartment	5390.73	2173.98	7564.71
Bungalow	6628.17	3034.79	9662.96
Chalet	11533.49	1902.24	13435.73
Cottage	6717.98	1489.01	8206.99
Deluxe room	9021.24	2198.07	11219.31
Penthouse	2557.38	852.46	3409.84
Standard room	7358.61	2081.13	9439.74
Studio	6104.22	379.78	6484
Suite	3582.75	2149.65	5732.4
Villa	3157.38		3157.38
No-Show			
Apartment	222.57		222.57
Cottage	190.91		190.91
Standard room	198.98		198.98
Tổng Cuối	74421.91	23155.13	97577.04

Trường PivotTable

Hiện hoạt **Tất cả**

Chọn các trường để thêm vào báo cáo:

Tìm kiếm

☐ RoomCost

☐ BedCost

☒ TotalPrice

> F_Sale

Kéo trường giữa các vùng bên dưới:

Bộ lọc

Cột

IsWeekday

Hàng

Giá trị

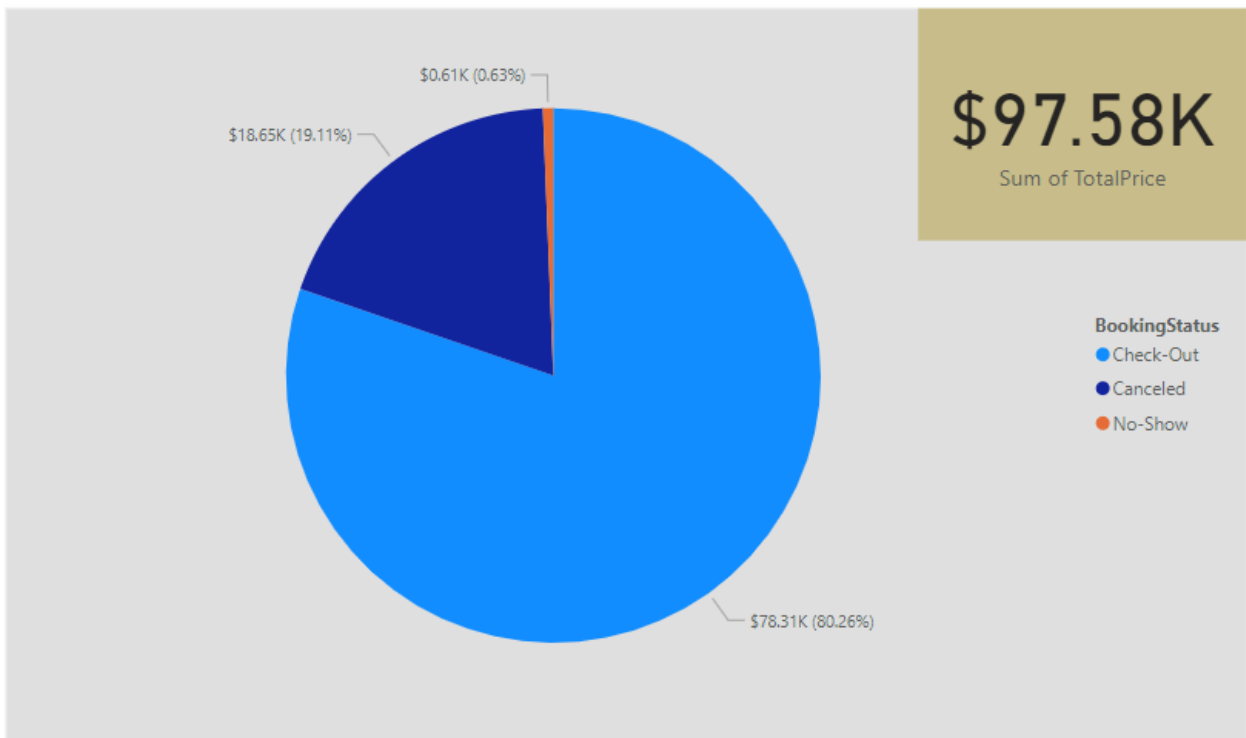
BookingStatus

Tổng của TotalPrice

RoomType

Bảng dữ liệu cho thấy chi phí cho ngày trong tuần (Weekday) khách hàng đã Check-out, Canceled và No-show lớn hơn ngày cuối tuần (Weekend). Tổng chi phí khách hàng đã Check-out lớn hơn khách hàng Canceled (hủy) và No-show (không đến).

➤ **Sử dụng Power BI**



Nhìn hình ta thấy, Check-Out có Tổng Số Tiền cao nhất là \$78,313.06, tiếp theo là Canceled với \$18,651.52 và No-Show với \$612.46. Check-Out chiếm 80.26% Tổng Số Tiền.

Lượng khách thường đến khách sạn nhiều nhất vào ngày nào

➤ **Sử dụng Pivot Table trong Excel**

	J	K	L	M
Tổng của GuestCount	Nhãn cột			
		3		Tổng Cuối
		2015		
Nhãn Hàng	Weekday	Weekend		
Friday	158			158
Monday	153			153
Saturday		124		124
Sunday		119		119
Thursday	173			173
Tuesday	134			134
Wednesday	122			122
Tổng Cuối	740	243		983

Dựa trên thông tin hình ảnh:

Tổng số khách trong năm 2015 là **983** người.

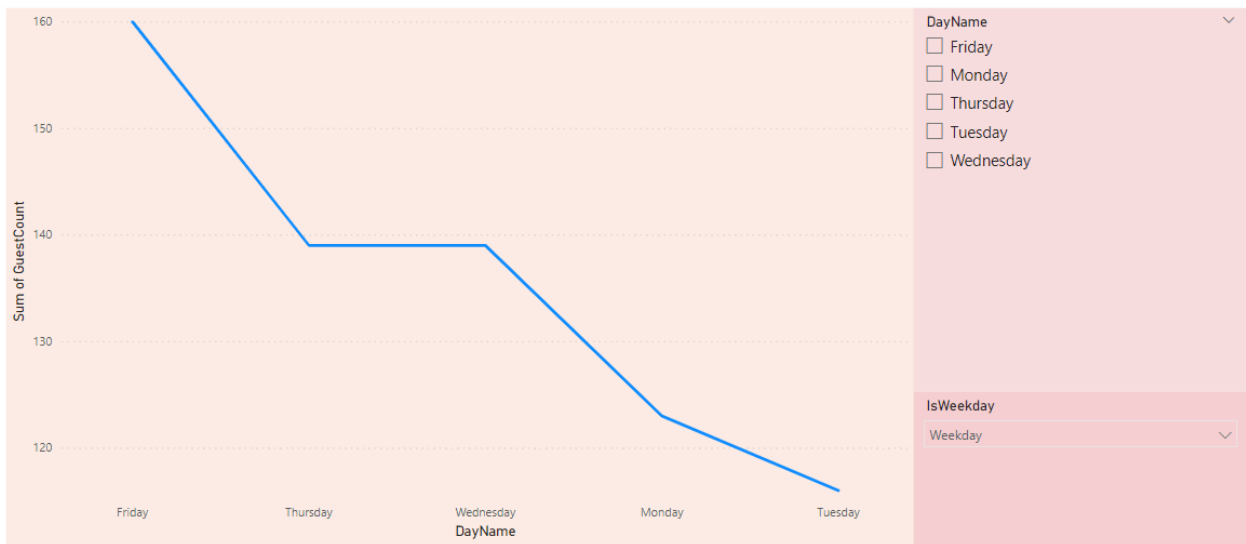
Trong đó, số khách vào **ngày thường** là **740** người, chiếm khoảng **75%** tổng số khách.

Số khách vào **cuối tuần** là **243** người, chiếm khoảng **25%** tổng số khách.

Thứ Sáu và **Thứ Hai** có lượng khách đến nhiều nhất trong tuần, mỗi ngày là **158** và **153** người tương ứng.

Thứ Bảy và **Chủ Nhật** có lượng khách ít hơn, với **124** và **119** người tương ứng.

➤ Sử dụng Power BI



Vào ngày thứ Sáu, tổng số khách cao nhất trong tuần và cao hơn 37,93% so với ngày thứ Ba, ngày có tổng số khách thấp nhất là 116. Trong suốt 5 ngày, tổng số khách dao động từ 116 đến 160.

CHƯƠNG 5. KẾT LUẬN

5.1. Kết quả đạt được

Việc xây dựng 'KHO DỮ LIỆU VỀ KHÁCH SẠN' đã giúp chúng em nâng cao hiểu biết thêm về kiến thức môn học Kho Dữ liệu. Qua quá trình thực hiện, chúng em đã trở nên thành thạo hơn trong việc sử dụng các công cụ hỗ trợ quá trình ETL để đưa dữ liệu vào kho dữ liệu, sử dụng thông thạo trong các công cụ SSIS và SSAS, cũng như các công cụ trực quan hóa và phân tích dữ liệu như Excel Pivot trong Excel và Power BI Desktop.

Dự án này còn giúp chúng em cải thiện kỹ năng phân tích và xử lý thông tin đầu vào, đặt câu hỏi tiền đề, và thảo luận để tìm ra hướng giải quyết và trả lời các câu hỏi.

Những kinh nghiệm thu được từ dự án này là nền tảng vững chắc để chúng em áp dụng vào các dự án thực tế trong tương lai.

5.2. Những hạn chế

Trong quá trình thực hiện xây dựng kho dữ liệu, nhóm chúng em đã gặp phải một số khó khăn do hạn chế về mặt khai thác dữ liệu. Tập dữ liệu còn nhiều thông tin chưa được khai thác hết, dẫn đến việc chưa tận dụng được tối đa các dữ liệu sẵn có. Quá trình ETL gặp một số lỗi phải mất nhiều thời gian để giải quyết.

Bên cạnh đó, những bước thực hiện của chúng em chưa đạt đến mức độ nâng cao trong toàn bộ quá trình xây dựng một kho dữ liệu hoàn chỉnh. Điều này ảnh hưởng đến khả năng tạo ra các phân tích và báo cáo chi tiết, chính xác, phục vụ cho việc ra quyết định quản lý và kinh doanh.

Chúng em nhận thấy cần phải tiếp tục học hỏi và nâng cao kiến thức để có thể hoàn thiện hơn trong các dự án tương lai.