

# Báo Cáo ETL

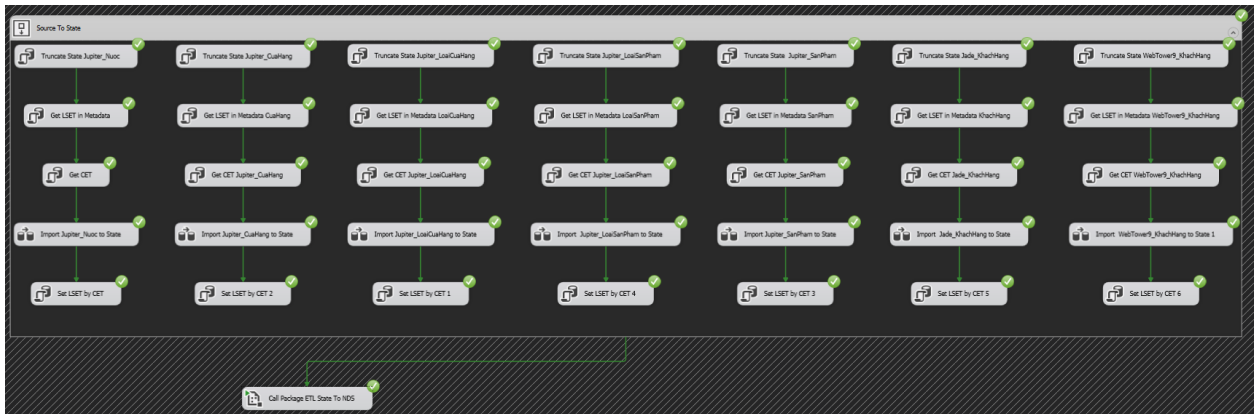
*ETL Source -> Stage, ETL Stage to NDS  
ETL NDS to DDS*

1712391 - Trần Bảo Duy  
13/11/2021

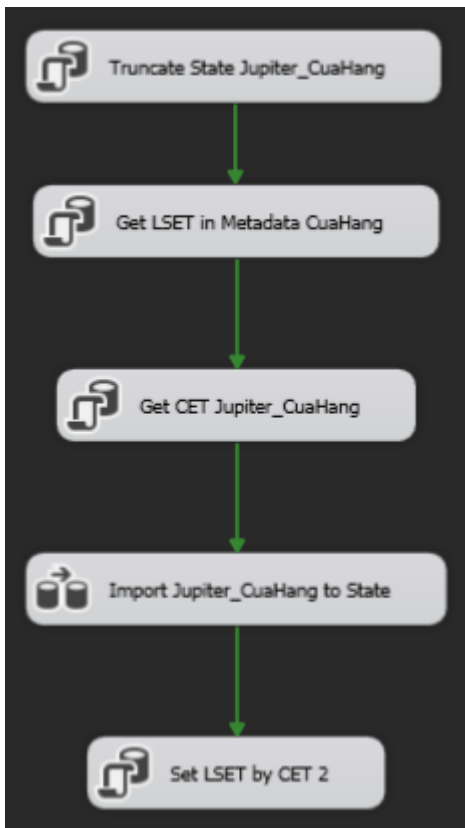
## Giới thiệu

Giải thích project ETL Source to Stage, ETL Stage to NDS.

### ETL Source To Stage



Quá trình này thực hiện đổ dữ liệu từ source vào Stage. Quá trình đó được ghi nhận vào trong metadata.



### Chi tiết một trong nhiều quá trình ETL Source to Stage

Đầu tiên xóa toàn bộ dữ liệu trong Stage

Lấy LSET ( thời điểm rút trích thành công gần nhất )

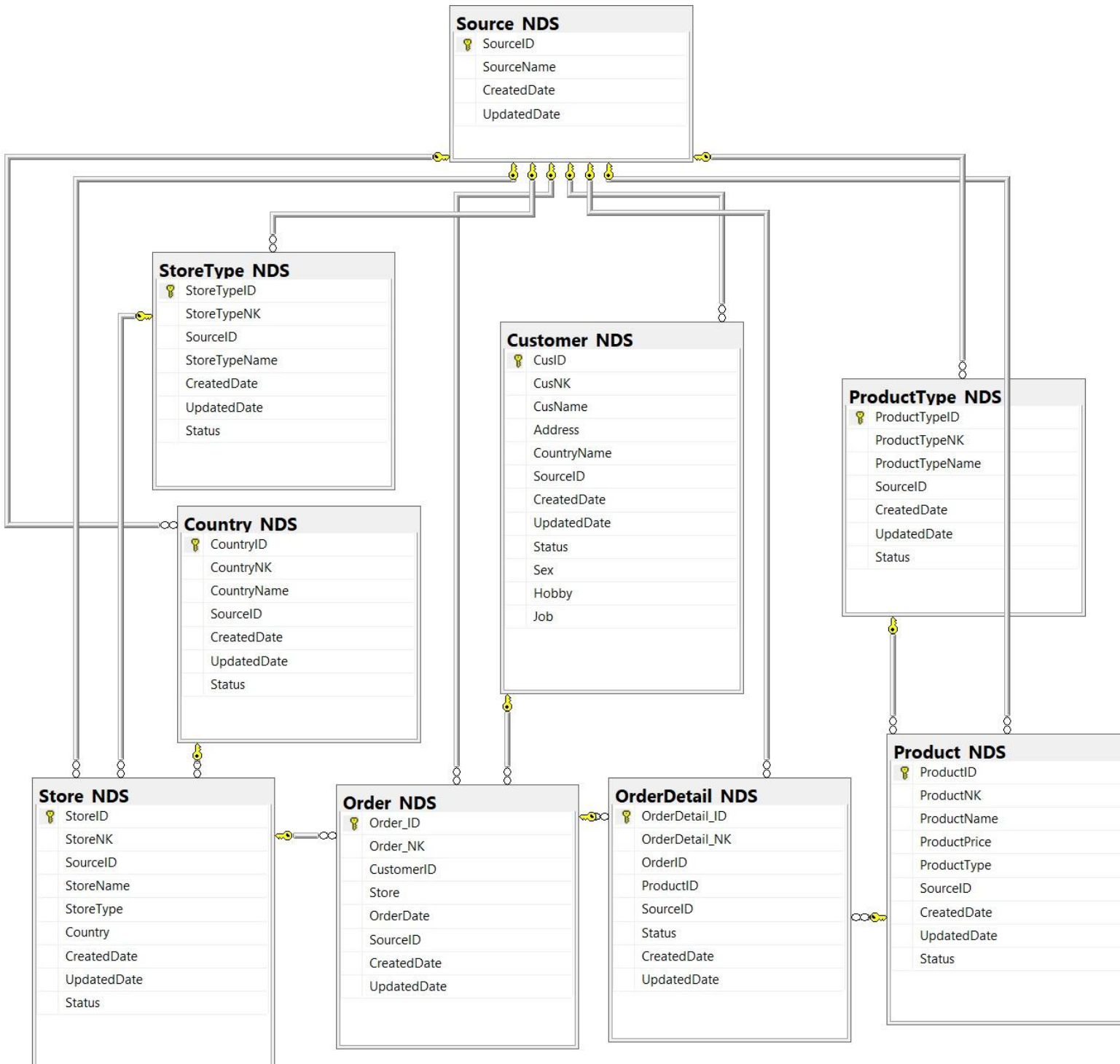
Tạo CET ( thời điểm rút trích hiện tại )

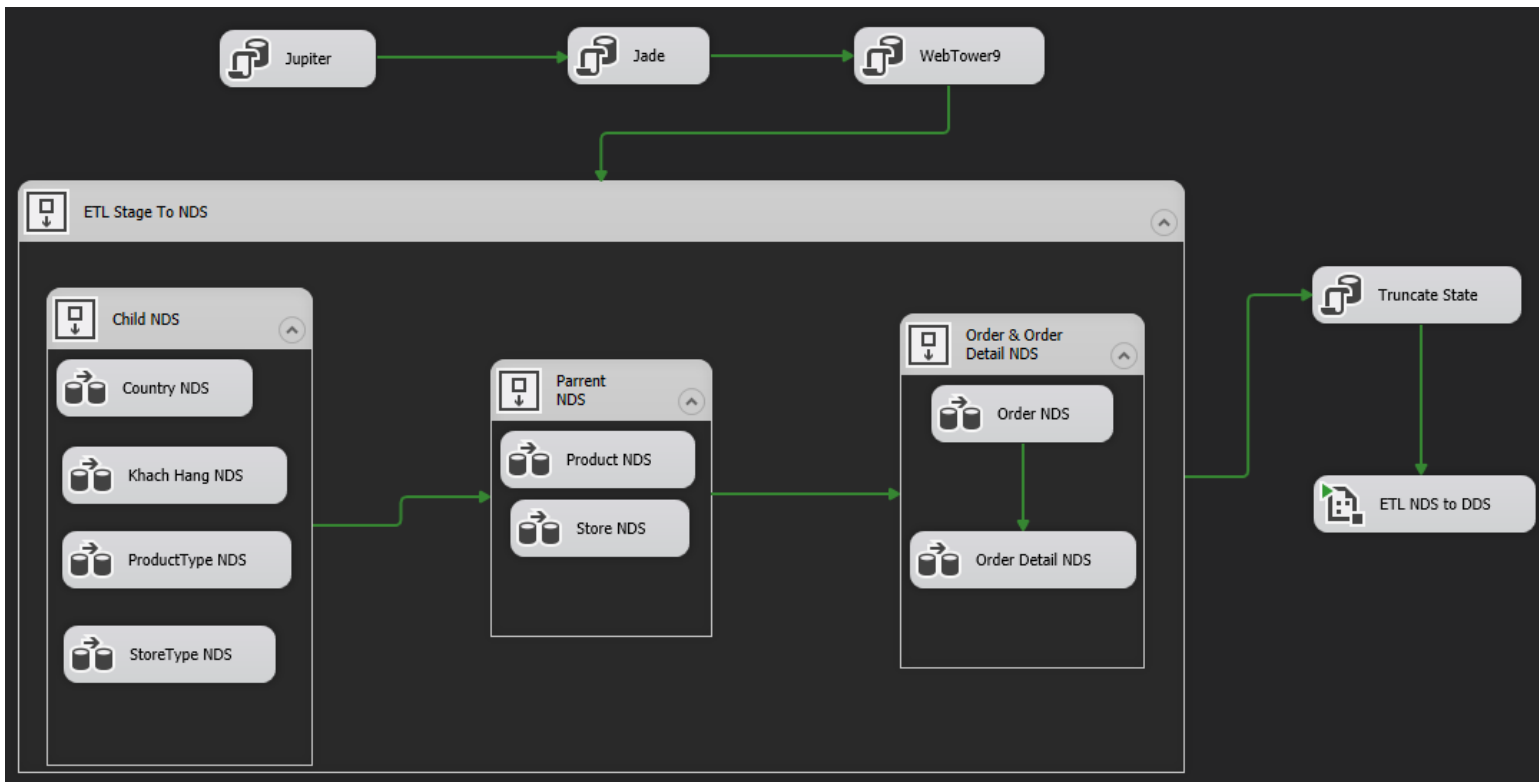
Thêm dữ liệu vào Stage sao cho phù hợp với khoảng thời gian trên

Gán LSET = CET

## ETL Stage To NDS

Mô hình NDS :





### Flow:

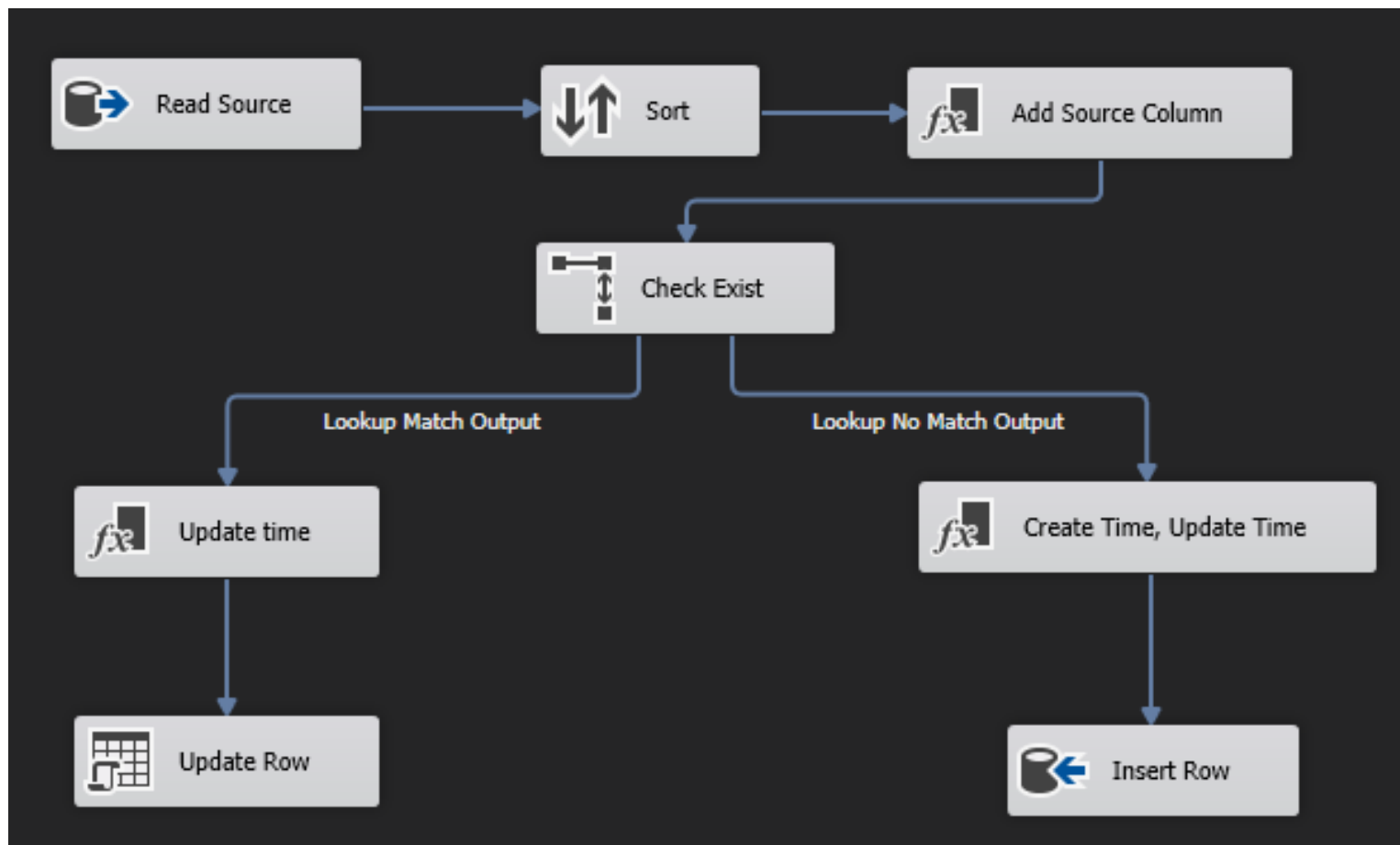
Lấy toàn bộ những source id của bảng Source\_NDS để tiến hành cập nhật cho những bảng khác.

Thực quá trình theo thứ tự như sau:

- A. **Quá trình 1** ETL các bảng Country\_NDS, Customer\_NDS, Product\_Type\_NDS, StoreType\_NDS vì đây là các bảng không có khóa ngoại.
- B. **Quá trình 2** ETL các bảng Product\_NDS, Store\_NDS vì đây là các bảng có 1 khóa ngoại.
- C. **Quá trình 2** ETL các bảng Order\_NDS, Order\_Detail\_NDS vì đây là các bảng có hơn 1 khóa ngoại.

Sử dụng Sequence Container để quyết định các Data Flow Task thực hiện đồng bộ. Ngoài ra để chỉ ra những Component nào thực hiện trước ta kéo mũi tên để định hướng data flow cho các component

Ví dụ đại diện: **Quá trình 1** thực hiện ETL Source to NDS của Country.



Dùng OLE DB Source để đọc dữ liệu từ Stage Nước

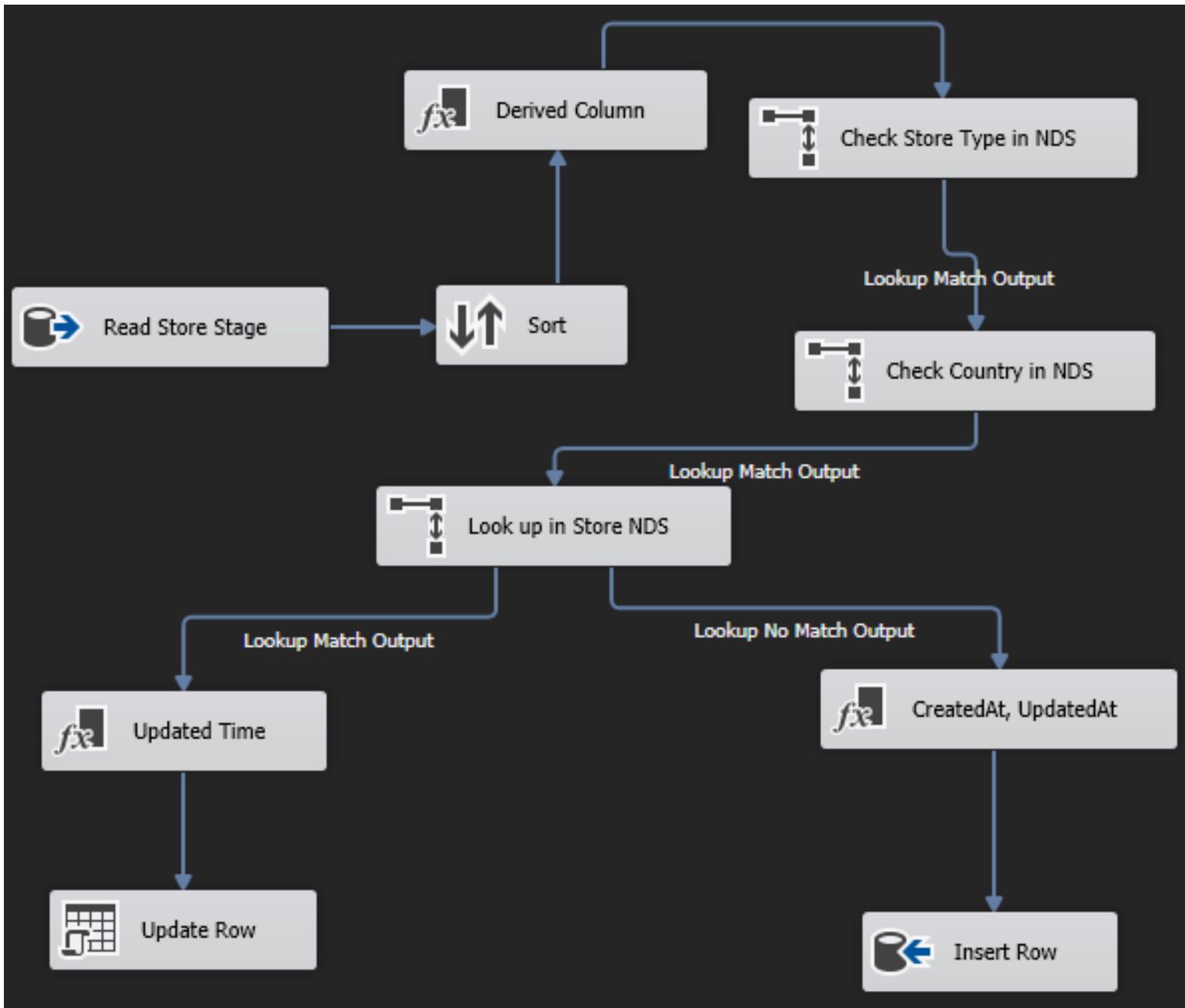
Dùng Sort để sắp xếp mã nước

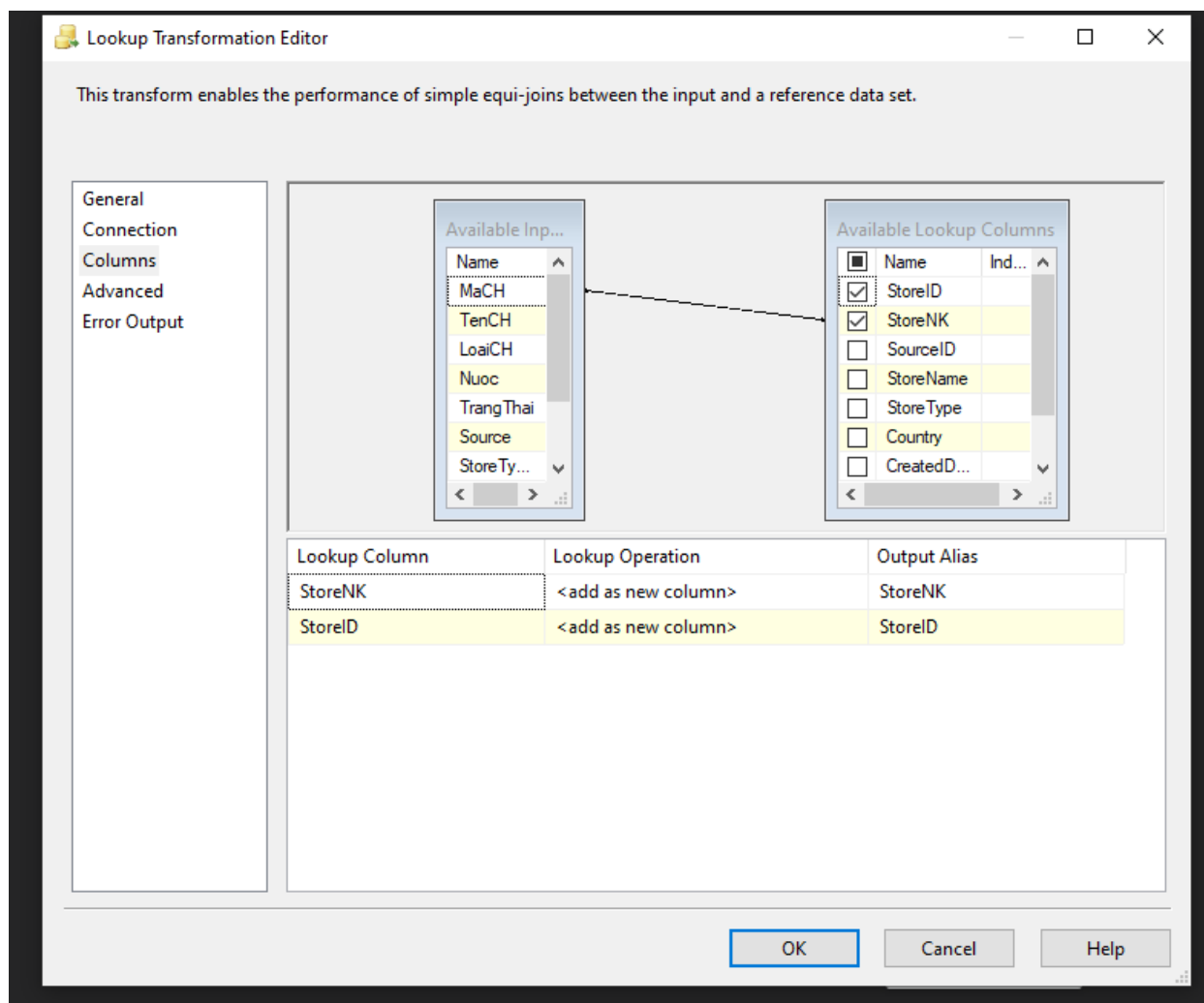
Dùng Derived Column để thêm cột Source

Dùng Look up để kiểm tra xem dòng dữ liệu này có tồn tại trong bảng Country\_NDS hay chưa.

- Nếu có, tạo thời gian cập nhật và Cập nhật dữ liệu tại dòng đó trong NDS\_DB
- Nếu không, tạo thời gian tạo, thời gian cập nhật ( bằng Null ), rồi thêm vào NDS\_DB

Ví dụ đại diện: **Quá trình 2** thực hiện ETL Source to NDS của Country.

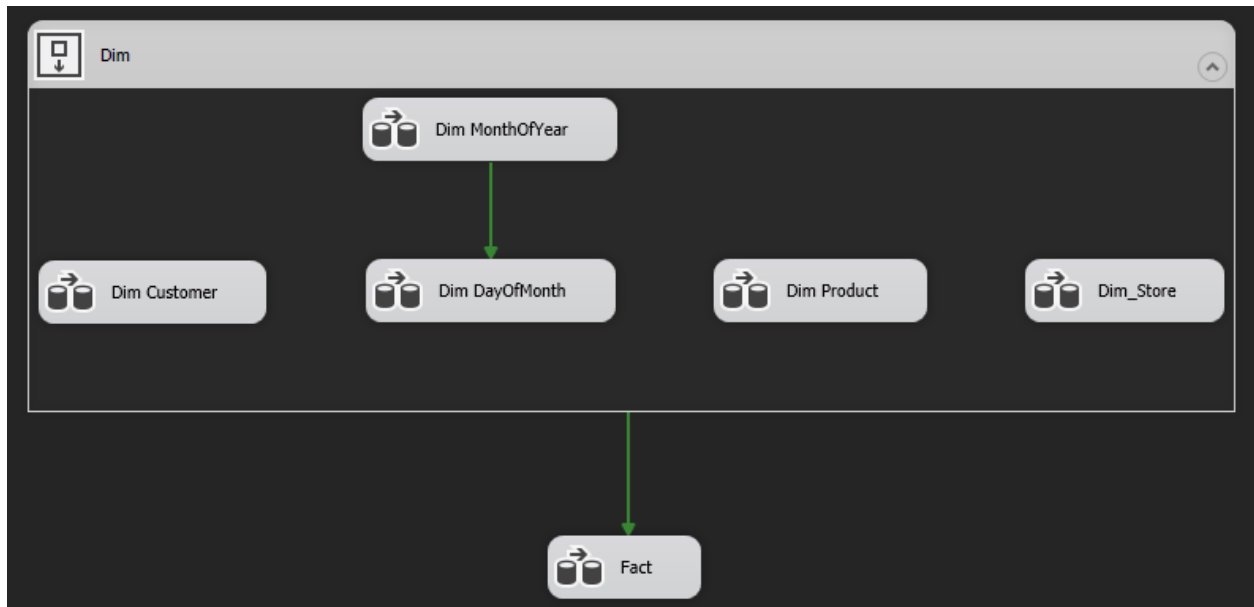




Các bước thực hiện tương tự quá trình trên. Ngoài ra, còn thêm một bước thêm khoá ngoại sử dụng Look Up để lấy Mã Nước và Mã Loại Cửa Hàng trong NDS



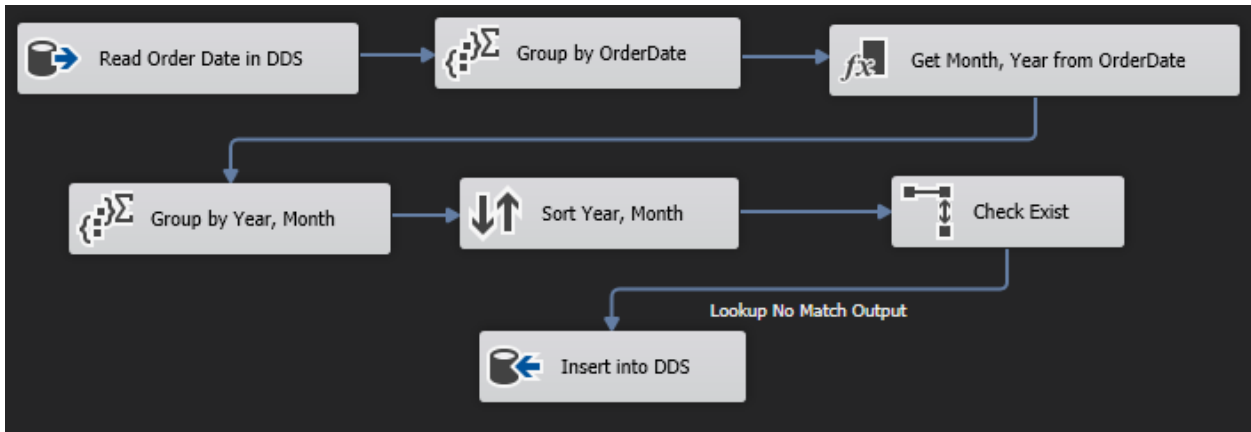
## ETL NDS To DDS



Thực ETL các bảng Dim :

1. Dim\_Customer
2. Dim\_Product
3. Dim\_Store
4. Dim\_MonthOfYear - >Dim DayOfMonth

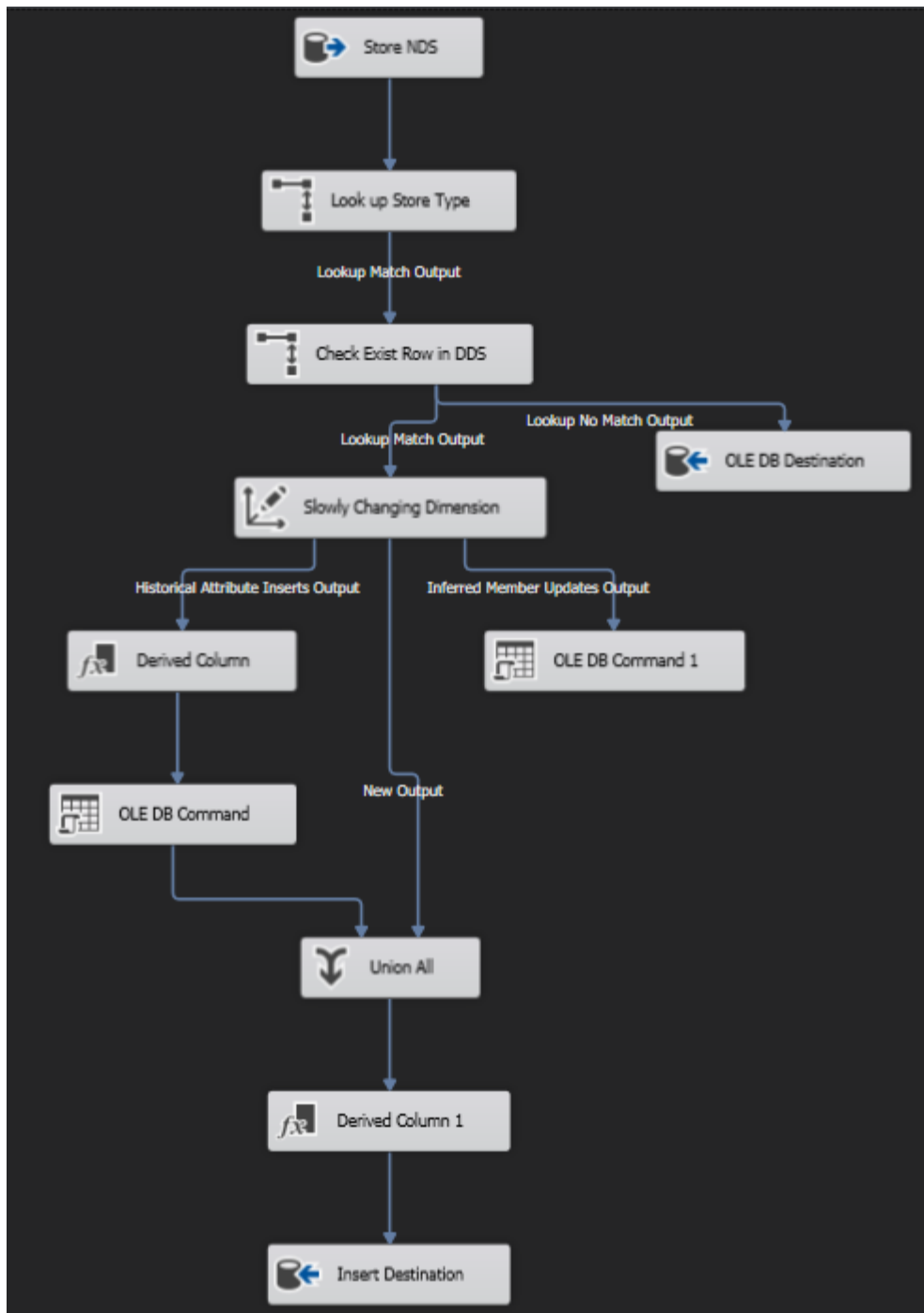
### Ví dụ quá trình ETL NDS to DDS ở DimMonthOfYear (Tháng trong Năm)



Tương tự như cái quá trình trên. Ngoài ra, để tổng hợp cái OrderDate lại với nhau ta sử dụng Component Aggregate để group by các giá trị ngày lại với nhau rồi dùng Derived Column để tạo ra cột Month và Year từ giá trị Order Date.

Sau đó thực hiện quá trình sắp xếp thứ tự tăng dần, kiểm tra tồn tại trong DDS rồi thêm vào DDS

### Xử lý chiều thay đổi chậm tại một số Dim:



1. Lấy dữ liệu từ NDS Store.
2. Dùng look up để điền các khoá ngoại phù hợp. Ở đây cần phải mapping ID\_CuaHang để lấy Mã Loại Cửa Hàng và Tên Loại Cửa Hàng.
3. Kiểm tra xem dòng dữ liệu có tồn tại trong DDS\_DB hay chưa

#### 4. Nếu tồn tại:

Tiến hành xét chiều thay đổi chậm trên bảng Cửa hàng, nhận thấy thuộc tính thay đổi chậm là Tên Cửa Hàng và Nước

**Slowly Changing Dimension Columns**  
Manage the changes to column data in your slowly changing dimensions by setting the change type for dimension columns.

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

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

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

Select a change type for slowly changing dimension columns:

Dimension Columns	Change Type
Nước	Historical att...
TenCH	Historical att...

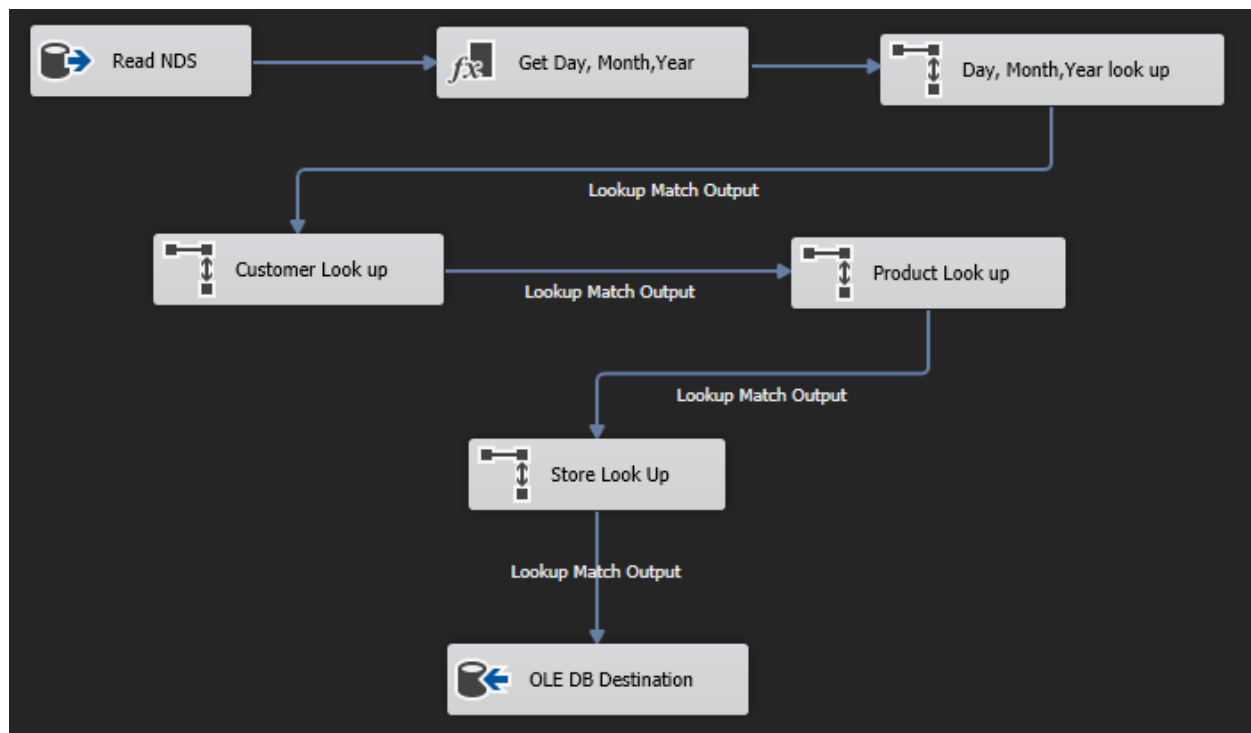
Remove

Help < Back Next > Finish >> Cancel

#### 5. Nếu không tồn tại:

Thêm dòng dữ liệu mới vào DDS

## Fact Doanh Thu



## 1. Đọc dữ liệu từ NDS sao cho phù hợp với DDS:

OLE DB connection manager:  
LocalHost.NDS\_Demo

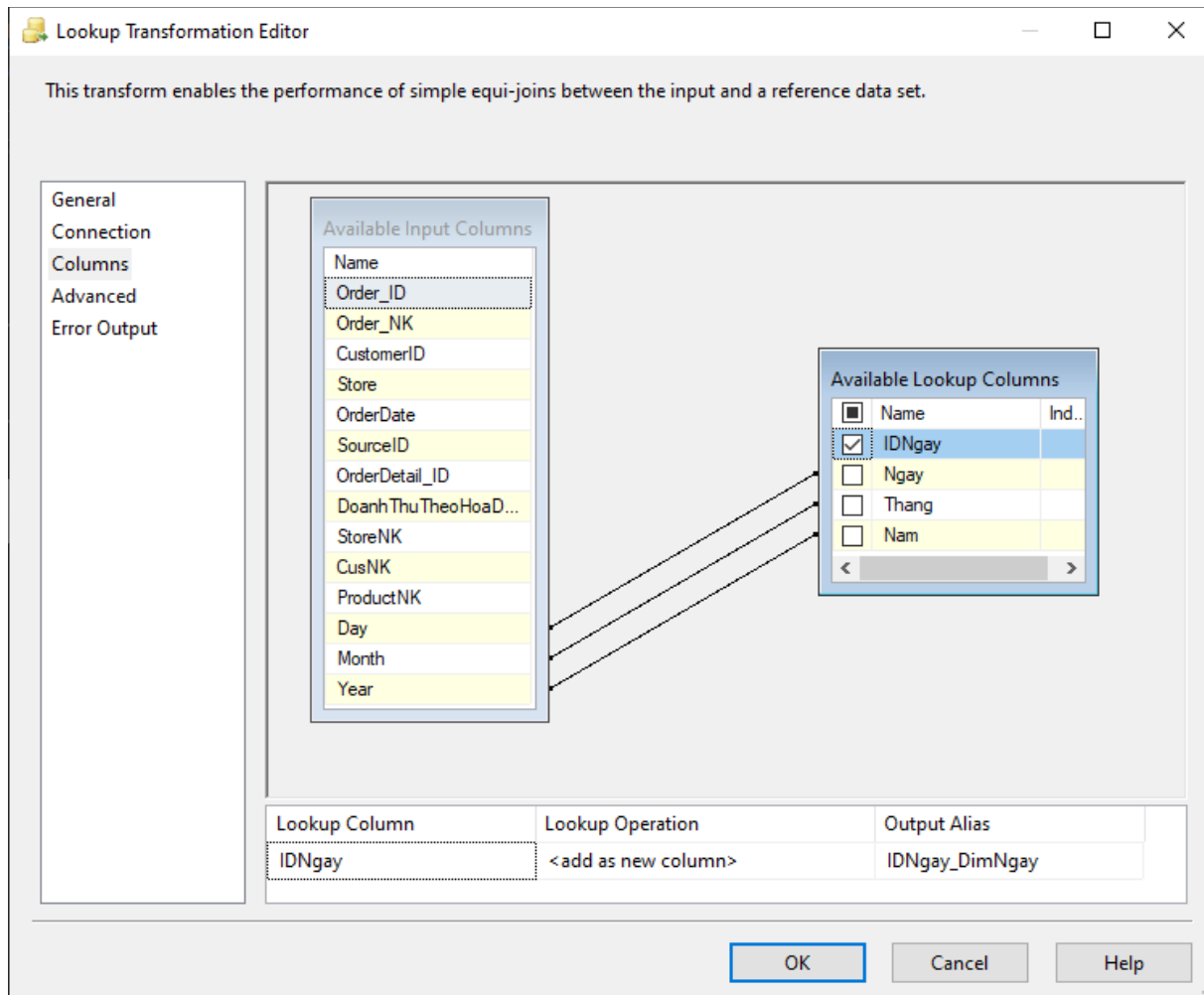
Data access mode:  
SQL command

SQL command text:

```
SELECT Order_NDS.Order_ID,  
       Order_NDS.Order_NK,  
       OrderDetail_NDS.OrderDetail_ID,  
       Order_NDS.CustomerID,  
       Order_NDS.Store,Product_NDS.ProductNK,Store_NDS.StoreNK,dbo.Customer_NDS.CusNK,  
       Order_NDS.OrderDate,  
       Order_NDS.SourceID,  
       SUM(Product_NDS.ProductPrice) AS 'DoanhThuTheoHoaDon'  
FROM dbo.Order_NDS AS Order_NDS  
  INNER JOIN dbo.OrderDetail_NDS OrderDetail_NDS  
    ON OrderDetail_NDS.OrderID = Order_NDS.Order_ID  
  INNER JOIN dbo.Product_NDS Product_NDS  
    ON Product_NDS.ProductID = OrderDetail_NDS.ProductID  
  INNER JOIN dbo.Source_NDS Source_NDS  
    ON Source_NDS.SourceID = Order_NDS.SourceID  
  INNER JOIN dbo.Store_NDS  
    ON Store_NDS.StoreID = Order_NDS.Store  
  INNER JOIN dbo.Customer_NDS  
    ON Customer_NDS.CusID = Order_NDS.CustomerID  
GROUP BY Order_NDS.Order_ID,  
       Order_NDS.Order_NK,  
       OrderDetail_NDS.OrderDetail_ID,  
       Order_NDS.CustomerID,  
       Product_NDS.ProductNK,  
       Store_NDS.StoreNK,  
       dbo.Customer_NDS.CusNK,  
       Order_NDS.Store,  
       Order_NDS.OrderDate,  
       Order_NDS.SourceID
```

Preview...

2. Dùng Derived Column để tạo các cột Day, Month, Year
3. Dùng Look Up để xác định khoá dữ liệu từ bảng DimNgàyTrongThang ( Lưu Ý: trước đó phải look up để tìm khoá cho DimNgàyTrongThang từ bảng DimThangTrongNam)



4. Tiếp tục Look up để tìm các khoá dữ liệu trong bảng DimCuaHang, DimSanPham, DimKhachHang tương ứng
5. Thêm dòng dữ liệu đó vào Fact Doanh Thu