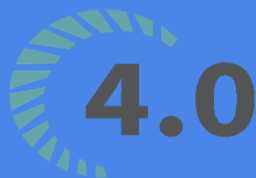


ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN



BÀI TẬP HW#8

MÔN HỆ THỐNG THÔNG TIN PHỤC VỤ TRÍ TUỆ KINH DOANH



Sinh viên: 18120241 - Trần Quốc Thịnh

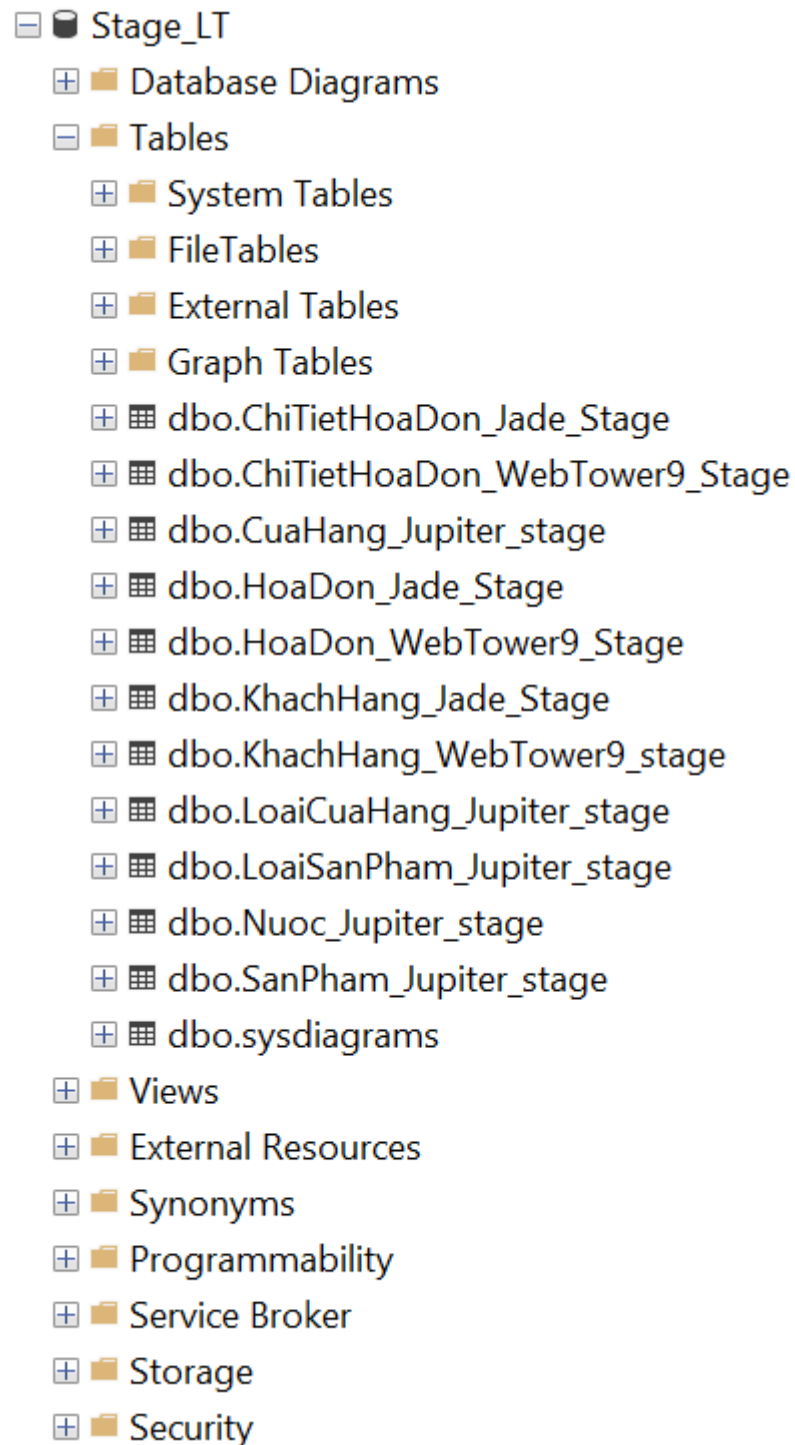
Giảng viên phụ trách: Gv. Hồ Thị Hoàng Vy
BÀI TẬP HW#8 - HỆ THỐNG THÔNG TIN PHỤC VỤ TRÍ TUỆ KINH DOANH
Học kì I - Năm học 2021-2022

MỤC LỤC

Source to Stage.	3
Stage to NDS:	8
NDS to DDS.	12

1. Source to Stage.

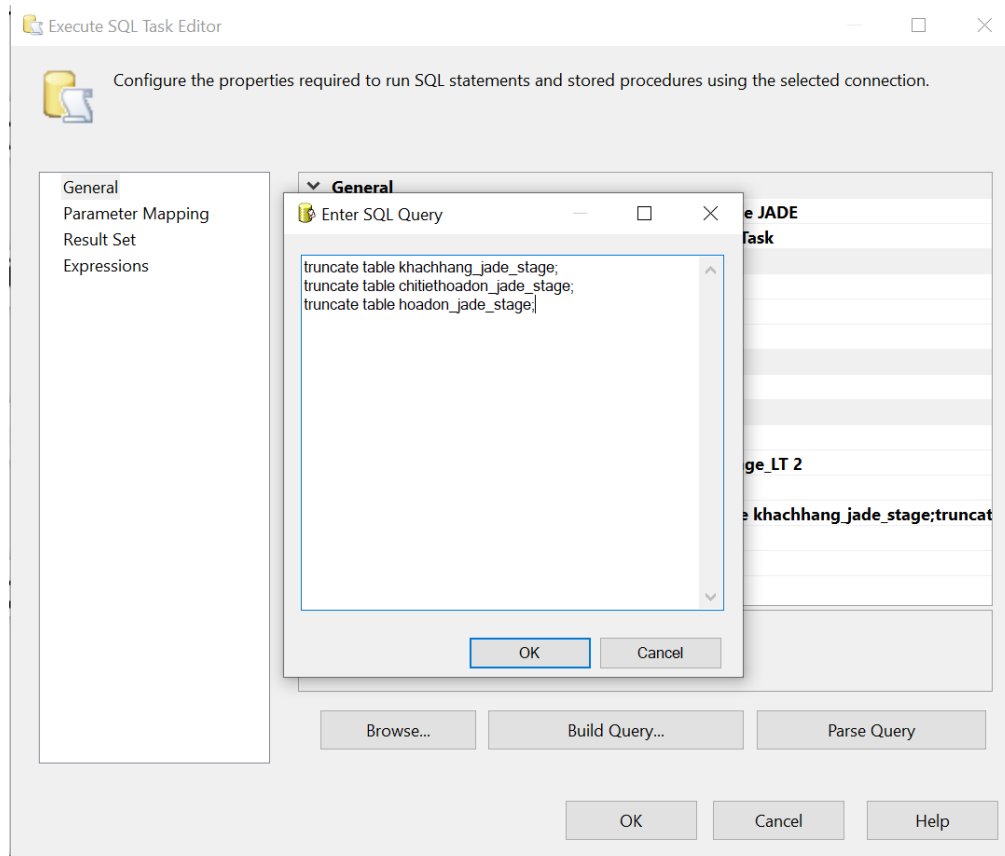
- Tạo sẵn Database Stage_LT:



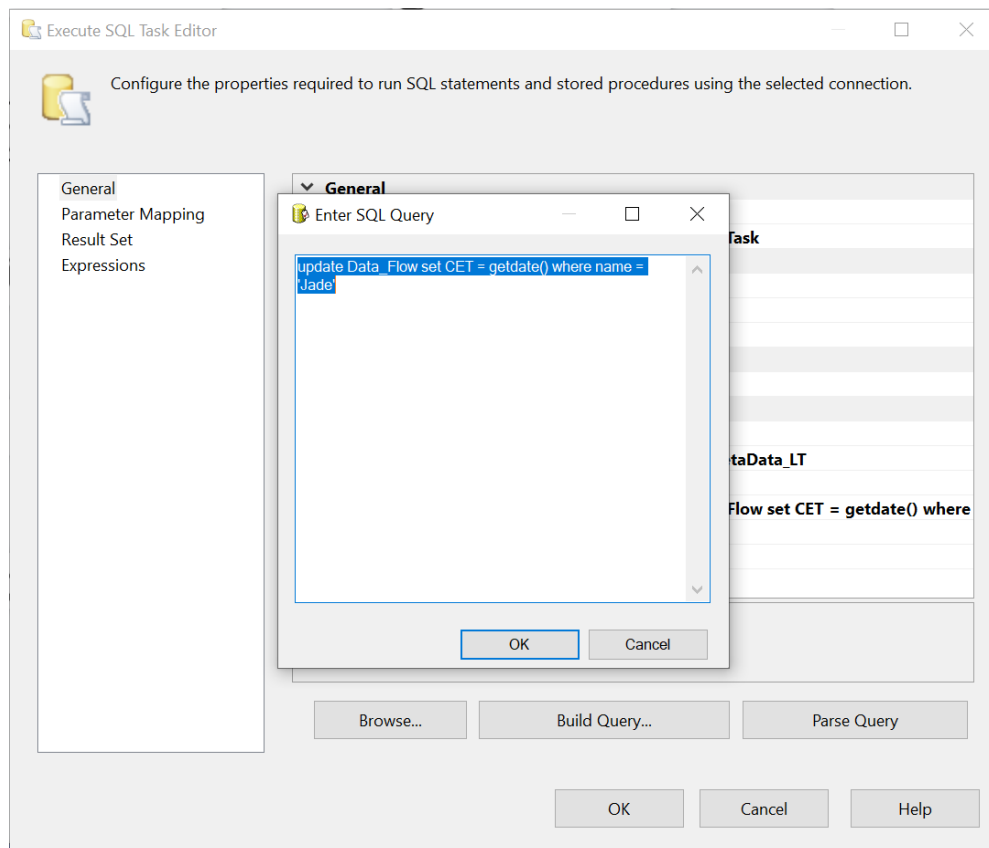
- Tạo Project SSIS và thiết lập package như hình:



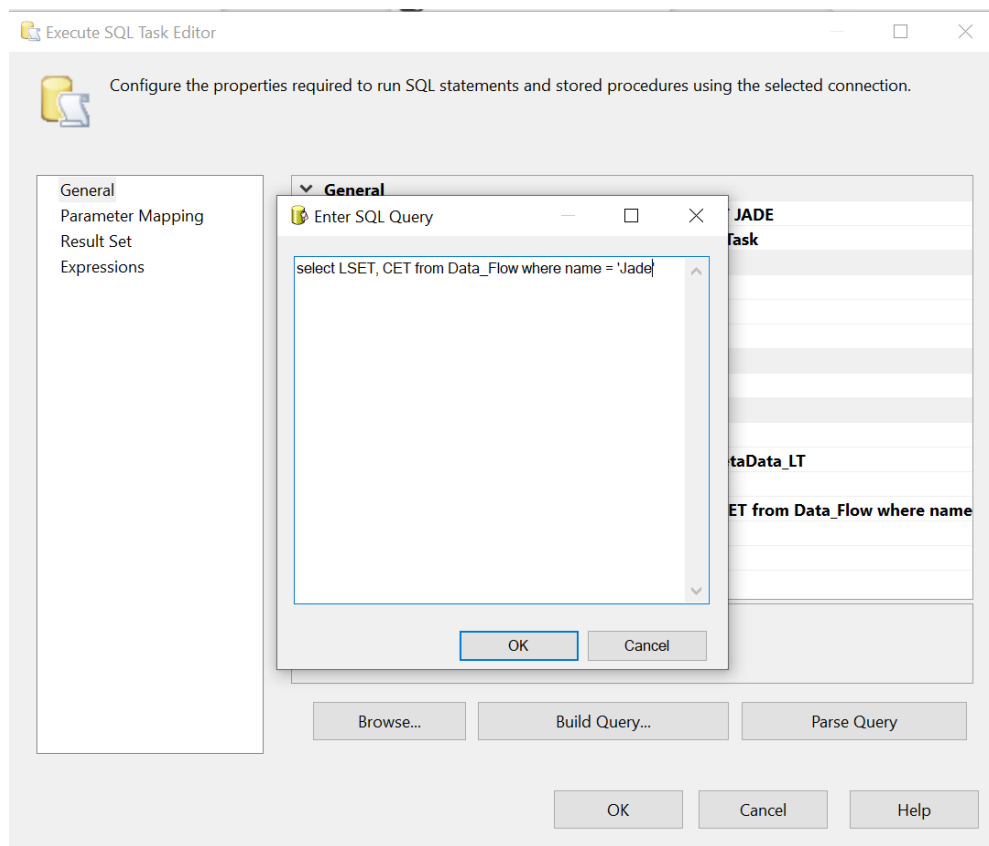
- Truncate Table:



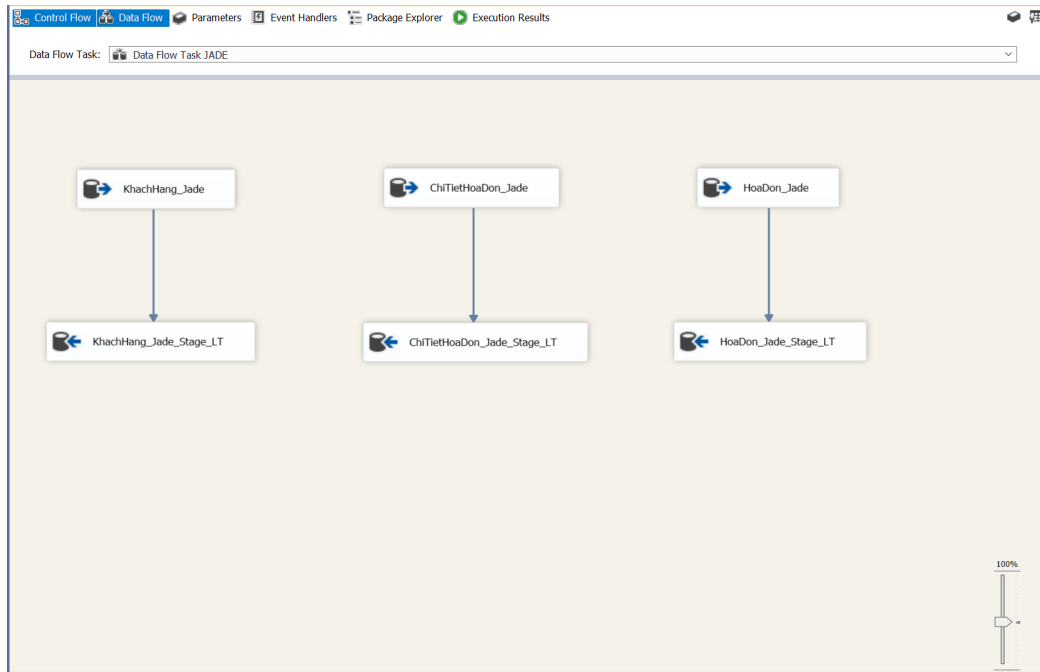
- Set CET:



- Get LSET, CET để so sánh khi nhập dữ liệu từ Source qua Stage:



- Thực hiện việc tải dữ liệu từ Source sang Stage: (Ở đây em thực hiện tải dữ liệu cho các bảng cùng 1 Source nguồn ở cùng 1 Data Flow Task)



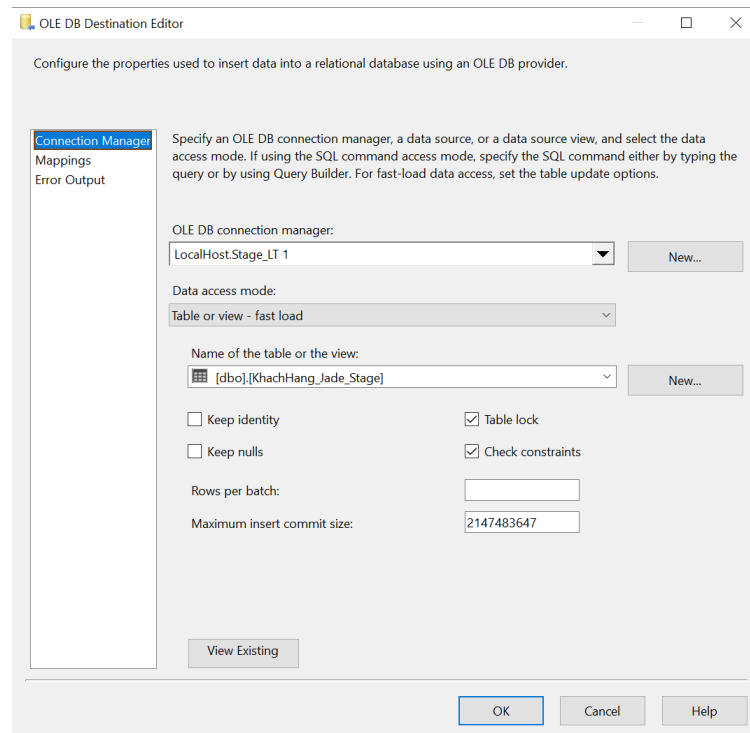
- Thiết lập câu truy vấn cho Source:

The screenshot shows the 'OLE DB Source Editor' dialog box. The 'Connection Manager' tab is selected in the left sidebar. The main area contains the following configuration:

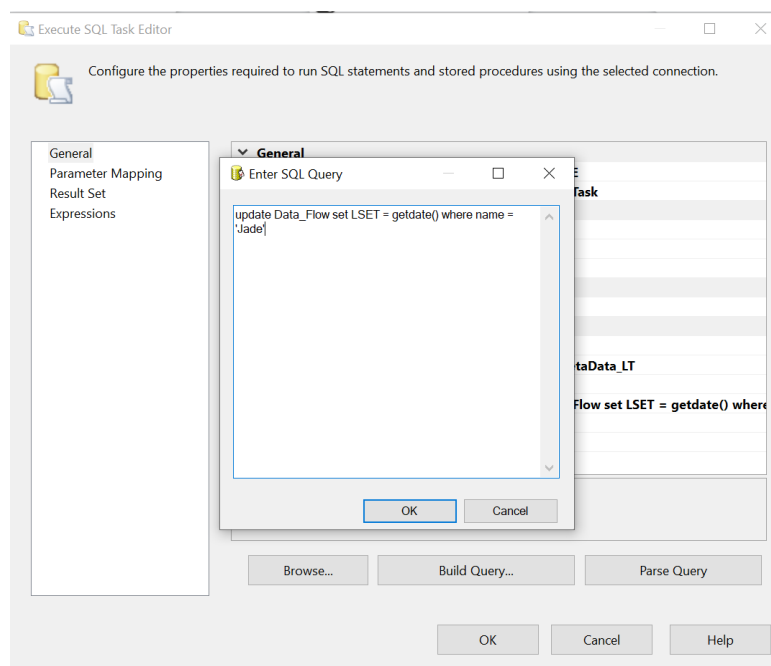
- OLE DB connection manager:** A dropdown menu showing 'LocalHost.Jade' with a 'New...' button to its right.
- Data access mode:** A dropdown menu showing 'SQL command'.
- SQL command text:** A text area containing the query: `select distinct * from KhachHang_Jade where (NgayTao < ? and NgayTao > ?) or (NgayCapNhat < ? and NgayCapNhat > ?)`. To the right of the text area are buttons for 'Parameters...', 'Build Query...', 'Browse...', and 'Parse Query'.
- Preview...** button at the bottom left of the main area.

At the bottom of the dialog are 'OK', 'Cancel', and 'Help' buttons.

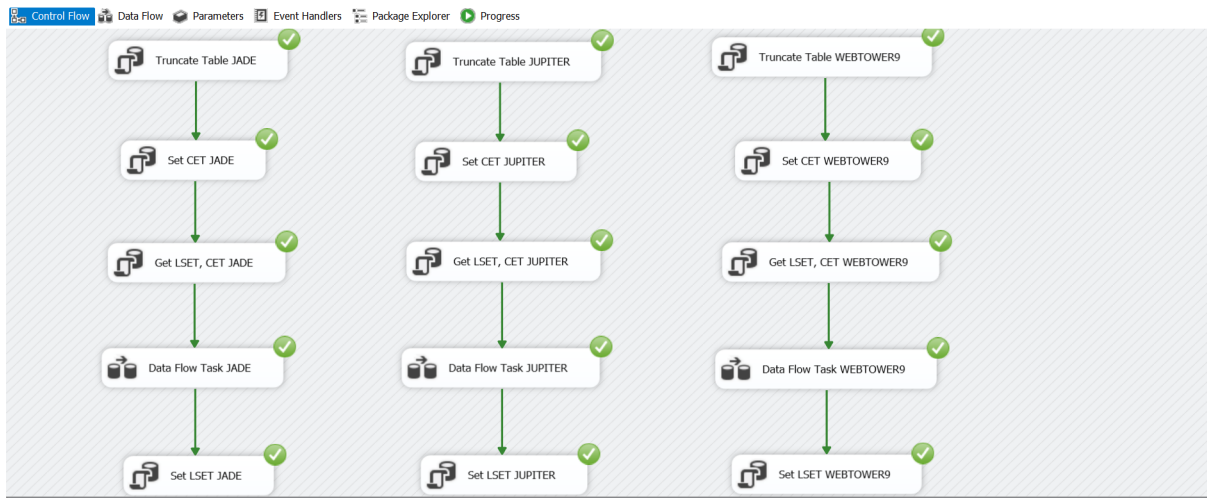
- Thiết lập cho Stage (Đích):



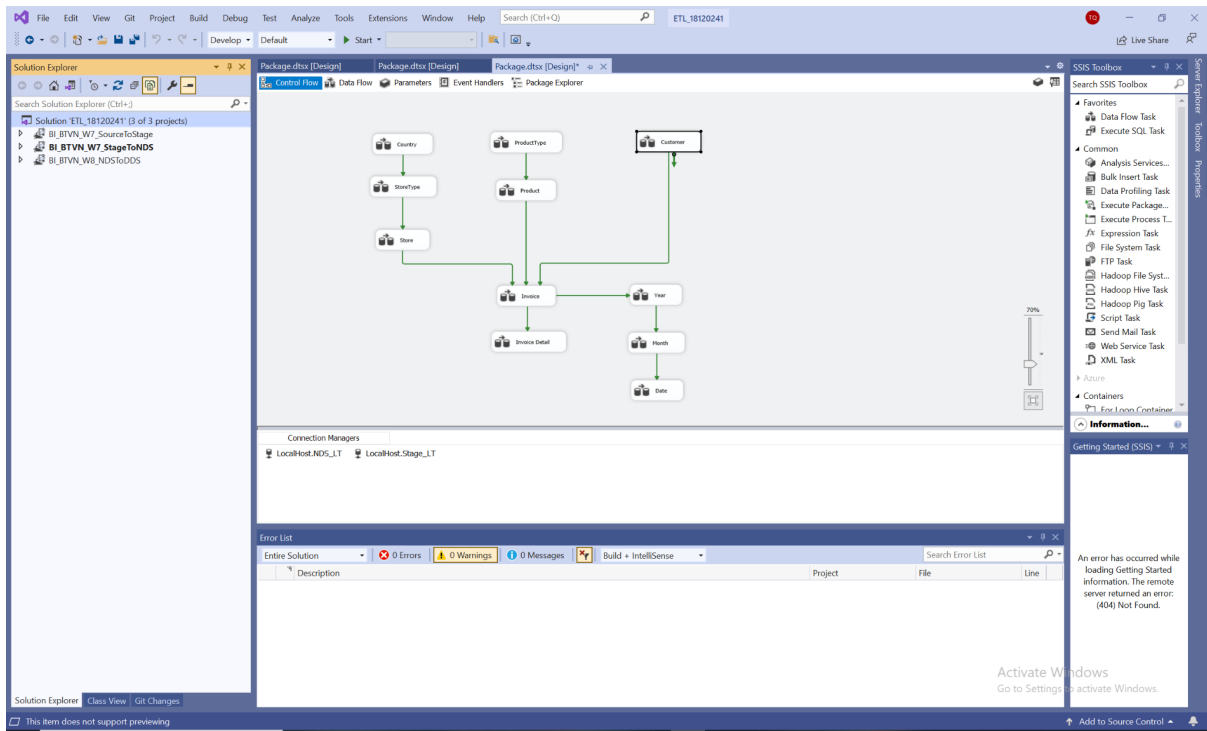
- Cuối cùng của phần tải dữ liệu từ Source → Stage này là Update lại dữ liệu của cột LSET trong Meta Data:



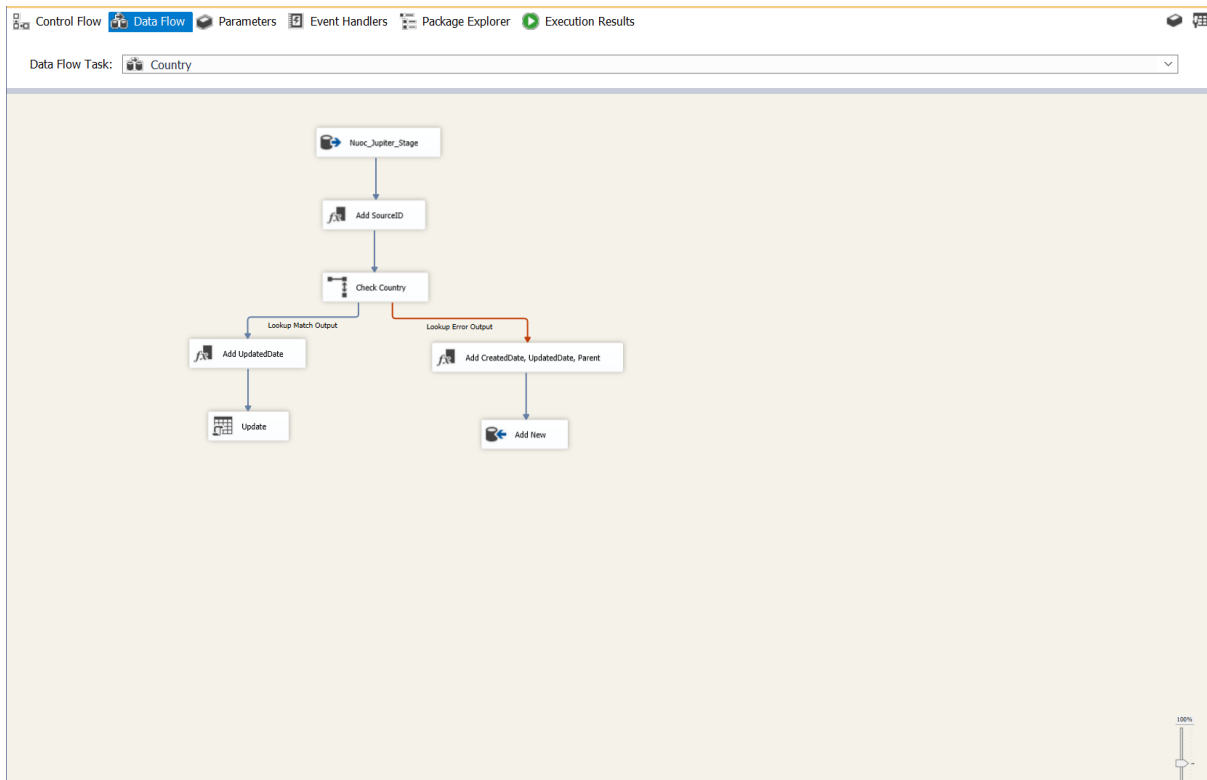
- Chạy Package thành công:



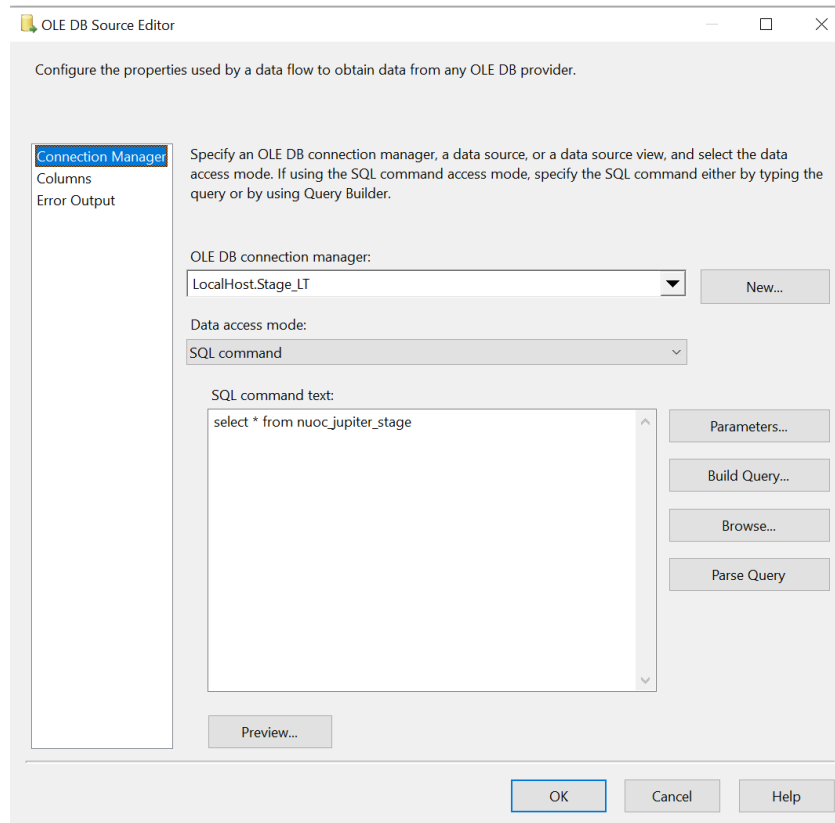
2. Stage to NDS:



- Thiết lập Data Flow Task:



- Thiết lập cho nguồn (Stage):



- “Add Source” để giúp em add thêm 1 cột Source vào dữ liệu nguồn vừa lấy được và thiết lập giá trị cho Source = <giá trị>.

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
Source	<add as new column>	2	four-byte signed inte...	

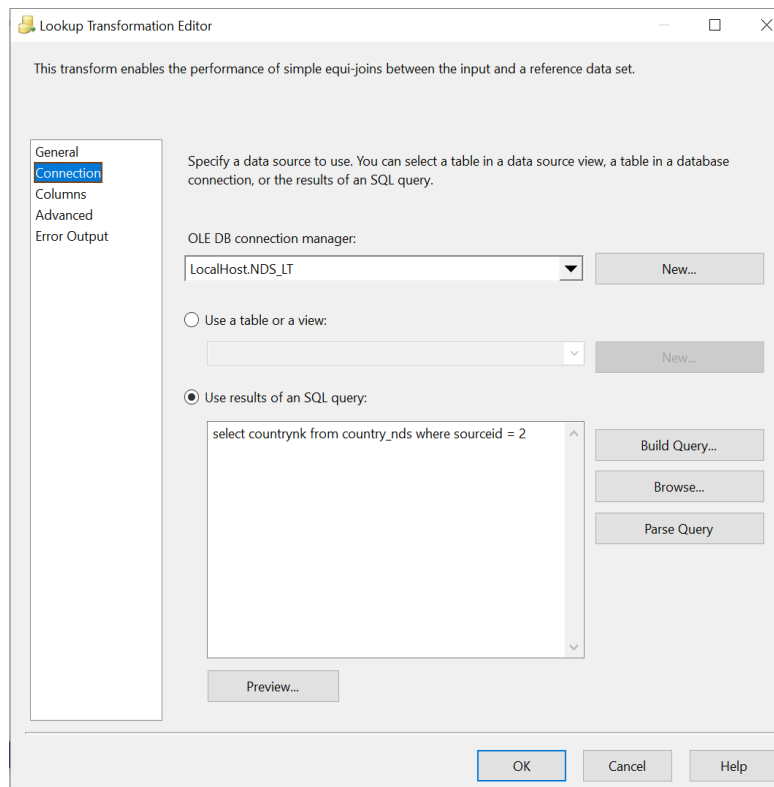
Configure Error Output...

OK

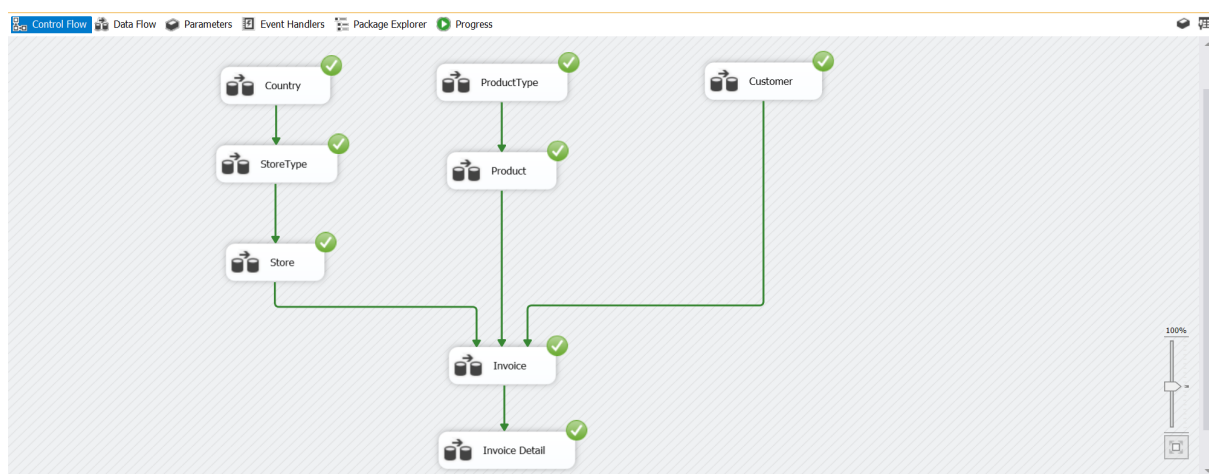
Cancel

Help

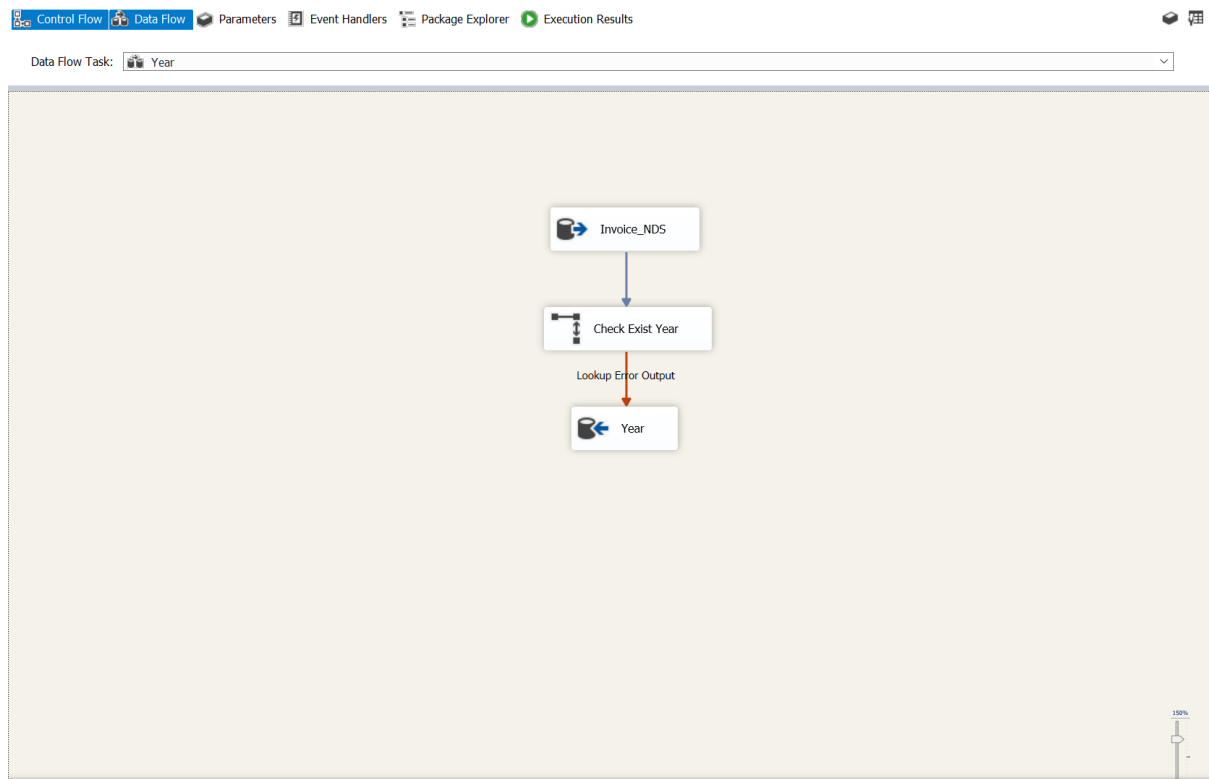
- “Check Country” (Lookup): Em sử dụng để tìm trong bảng Country_NDS của NDS và xác định xem record dữ liệu của Stage đang được thêm vào đã tồn tại hay chưa:



- Ở đây, nếu đã tồn tại thì sẽ thực hiện thêm cột UpdatedDate và tiến hành Update lại dữ liệu cho dòng dữ liệu có cùng mã tự nhiên và source với dòng dữ liệu lấy từ Stage.
- Nếu chưa tồn tại thì sẽ tiến hành insert thêm 1 dòng dữ liệu vào NDS dựa trên dữ liệu từ Stage.
- Chạy Package thành công:



- Data Flow Task Lấy ra năm từ Ngày tạo hoá đơn:



- Ở đây em thực hiện SELECT DISTINCT để lấy ra toàn bộ năm có trong dữ liệu, sau đó thực hiện Lookup để kiểm tra và đưa vào dữ liệu đích (DDS) nếu Năm lấy ra được chưa có trong dữ liệu đích. (Việc thực hiện tương tự như ở Stage to NDS)

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager

Columns

Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:

LocalHost.NDS_LT

New...

Data access mode:

SQL command

SQL command text:

SELECT DISTINCT YEAR(CreatedDate) as Year FROM Invoice_NDS

Parameters...

Build Query...

Browse...

Parse Query

Preview...

OK

Cancel

Help

- Tương tự với Tháng và Ngày:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
LocalHost.NDS_LT New...

Data access mode:
SQL command

SQL command text:
SELECT DISTINCT MONTH(CreatedDate) as Month, YEAR (CreatedDate) as Year FROM Invoice_NDS

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

Preview...

OK Cancel Help

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
LocalHost.NDS_LT New...

Data access mode:
SQL command

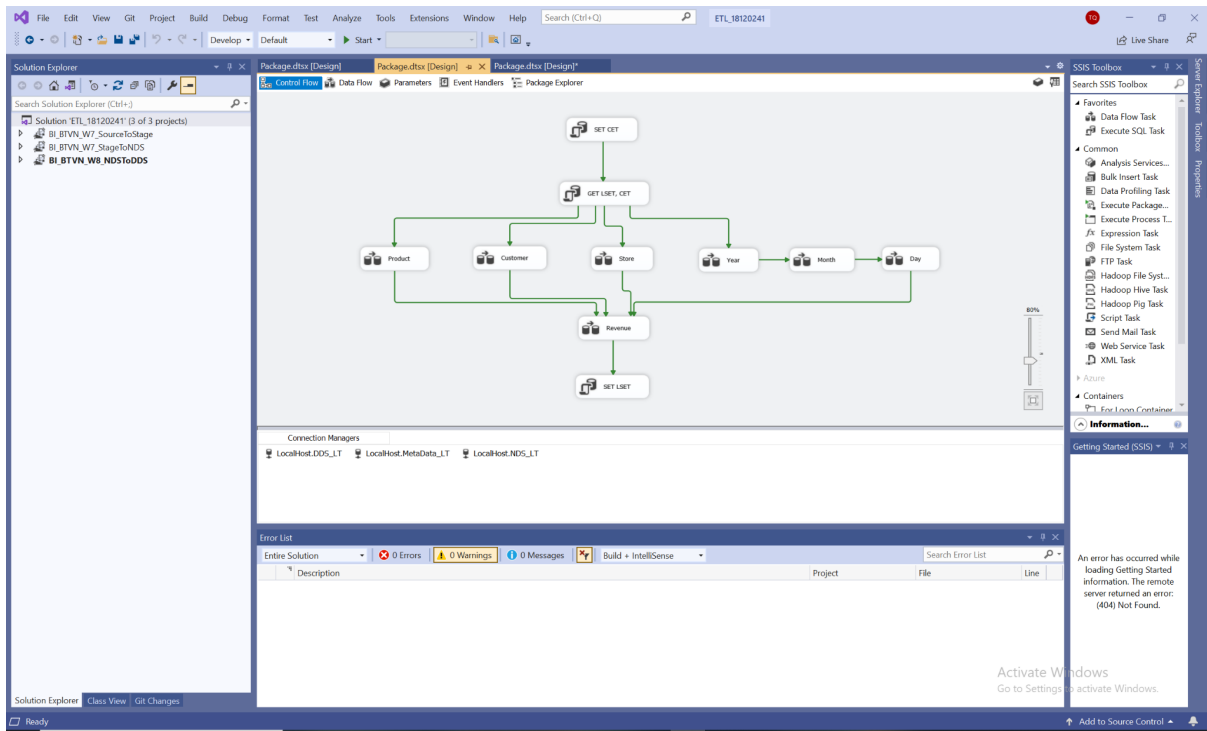
SQL command text:
SELECT DISTINCT DAY(CreatedDate) as Day, MONTH(CreatedDate) as Month, YEAR(CreatedDate) as Year FROM Invoice_NDS

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

Preview...

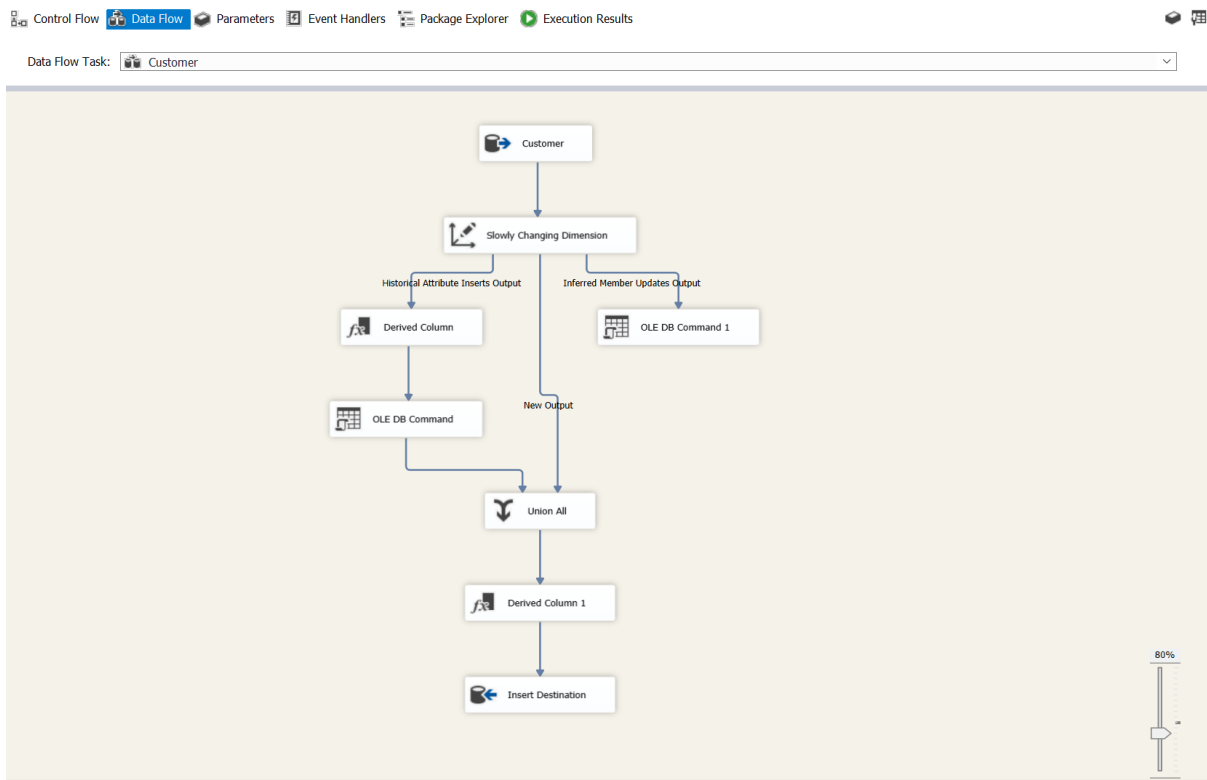
OK Cancel Help

3. NDS to DDS.

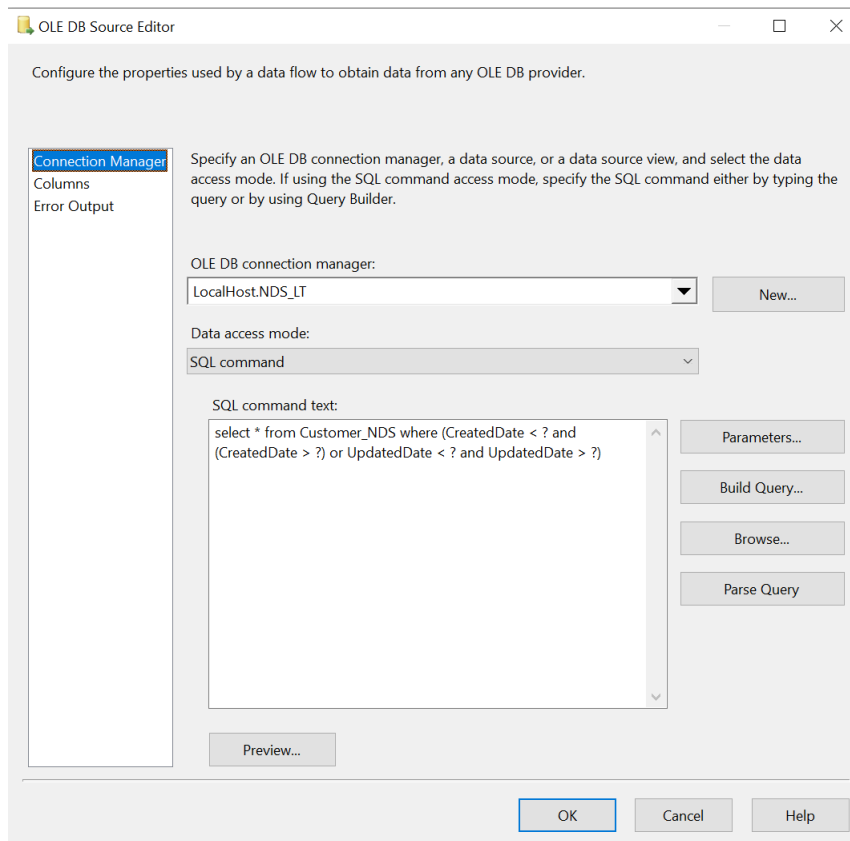


Các bước thực hiện tương tự việc tải dữ liệu từ Source vào Stage.

- Thiết lập Data Flow Task:



- Thiết lập cho dữ liệu đầu vào (NDS):



OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
LocalHost.NDS_LT

Data access mode:
SQL command

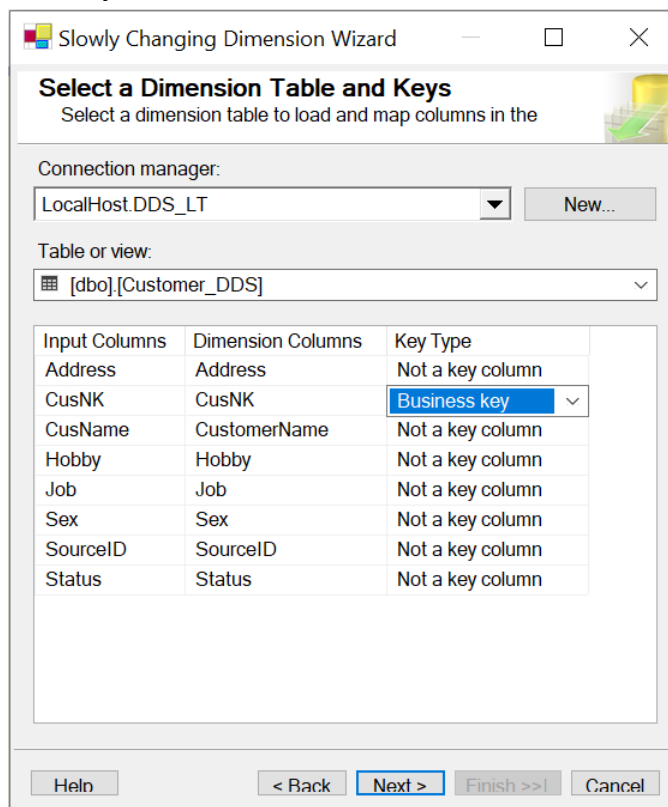
SQL command text:
select * from Customer_NDS where (CreatedDate < ? and (CreatedDate > ? or UpdatedDate < ? and UpdatedDate > ?)

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

Preview...

OK Cancel Help

- Thiết lập chiều thay đổi chậm:



Slowly Changing Dimension Wizard

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

Connection manager:
LocalHost.DDS_LT

Table or view:
[dbo].[Customer_DDS]

Input Columns	Dimension Columns	Key Type
Address	Address	Not a key column
CusNK	CusNK	Business key
CusName	CustomerName	Not a key column
Hobby	Hobby	Not a key column
Job	Job	Not a key column
Sex	Sex	Not a key column
SourceID	SourceID	Not a key column
Status	Status	Not a key column

Help < Back Next > Finish >> Cancel

Slowly Changing Dimension Wizard

Slowly Changing Dimension Columns

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

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
Address	Historical attribute

Remove

Help

< Back

Next >

Finish >>I

Cancel

Slowly Changing Dimension Wizard

Historical Attribute Options

You can record historical attributes using a single column or start and end date

☒ Use a single column to show current and expired records

Column to indicate current record: Status

Value when current: 1

Expiration value: 2

☐ Use start and end dates to identify current and expired records

Start date column:

End date column:

Variable to set date values:

Help

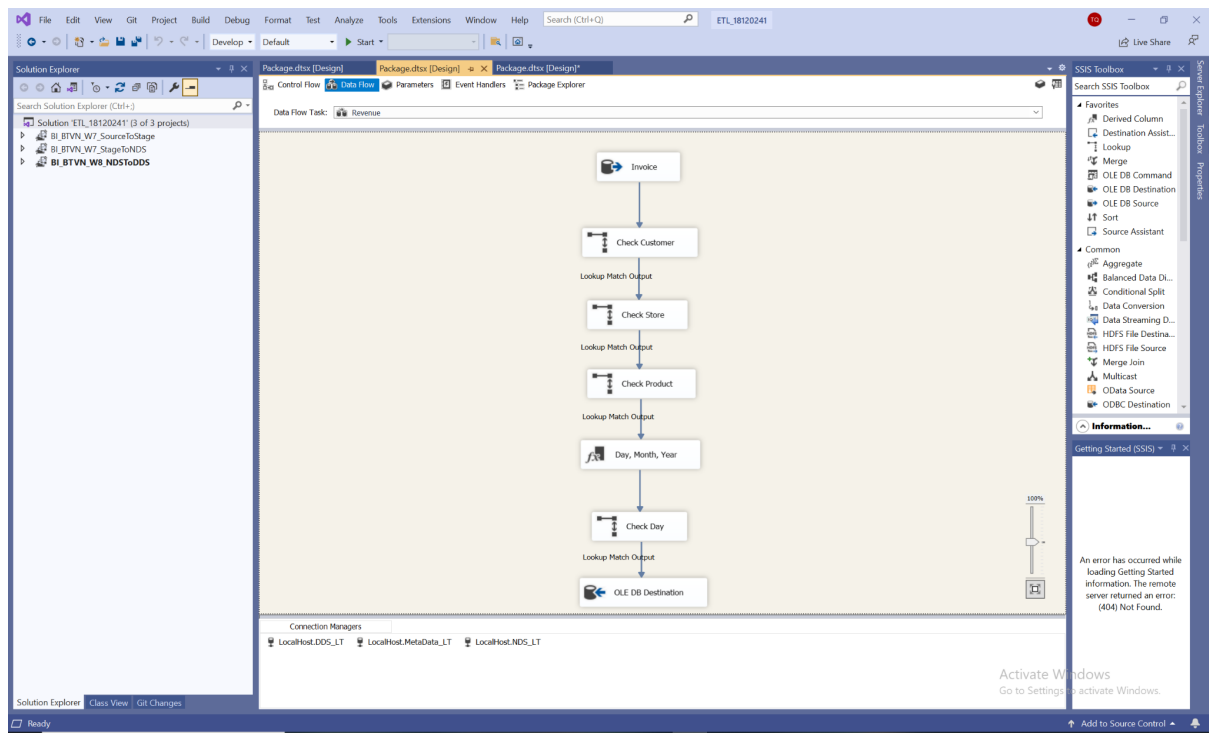
< Back

Next >

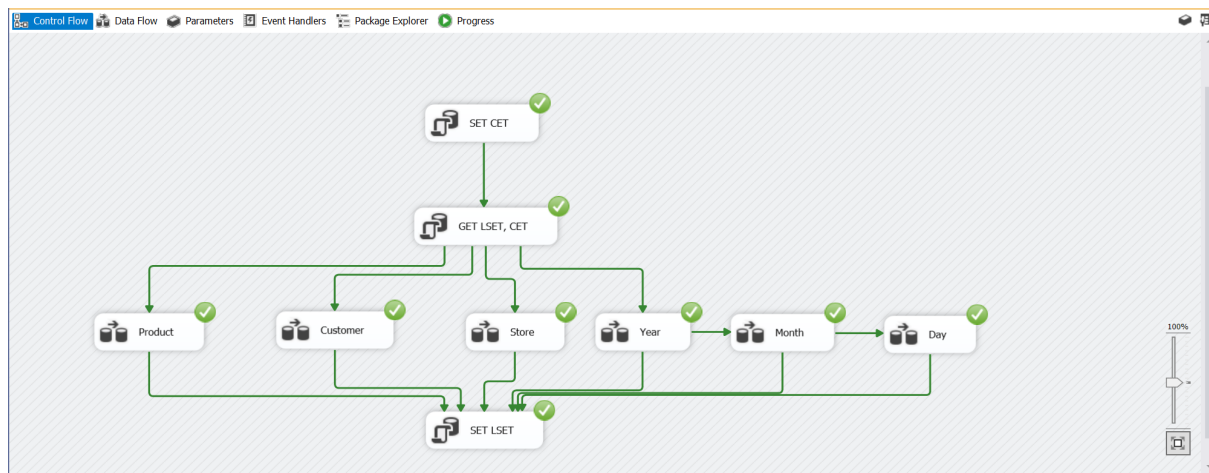
Finish >>I

Cancel

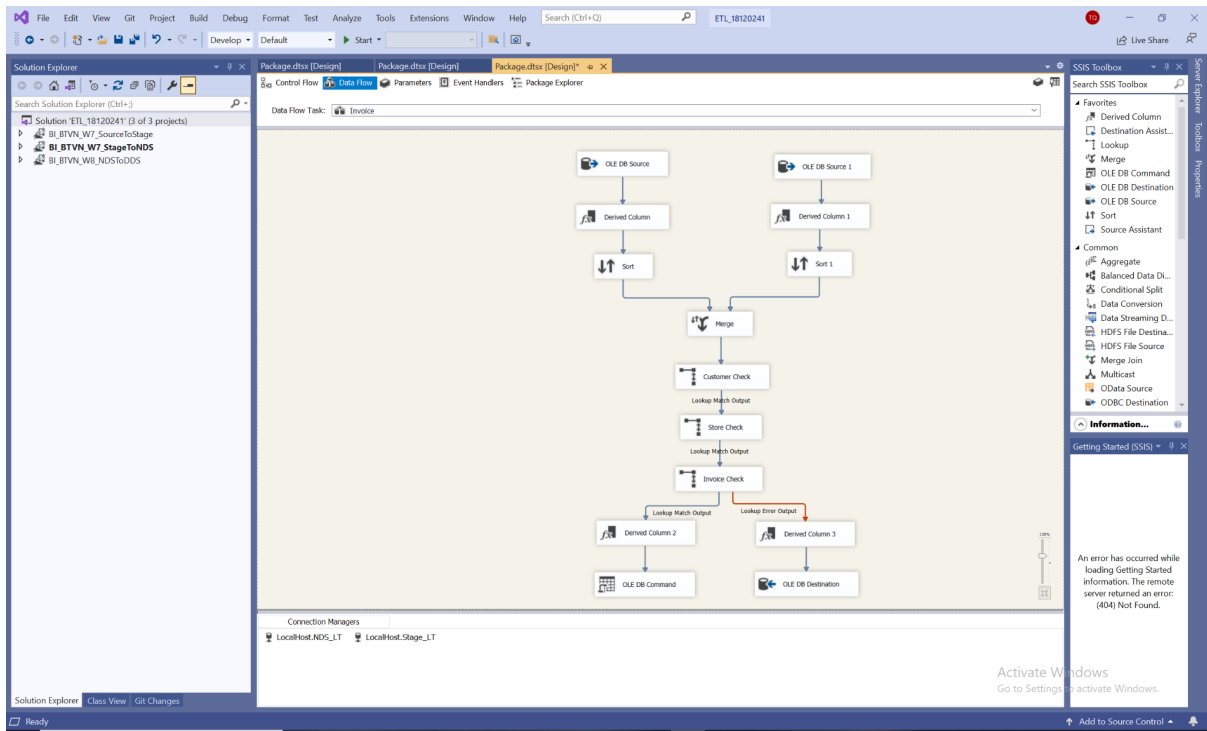
- Dataflow của Revenue:



- Chạy toàn bộ package:



- Lúc này em kiểm tra lại ở SQL Server, dữ liệu của bảng Revenue bị duplicate, em kiểm tra lại thì thấy do phần kiểm tra khoá ngoại của NDS bị lỗi thiếu tính chính xác đối với nguồn dữ liệu của các thuộc tính khoá ngoại.



- Sửa bằng cách mapping SourceID vào giai đoạn kiểm tra khoá ngoại cho các thuộc tính lấy từ 2 nguồn trở lên.

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 In...

Name
MaHD
MaKH
MaCH
NgayBan
TrangT...
Source

Available Lookup ...

Name	I...
<input type="checkbox"/> CusNK	
<input checked="" type="checkbox"/> CusID	
<input type="checkbox"/> SourceID	

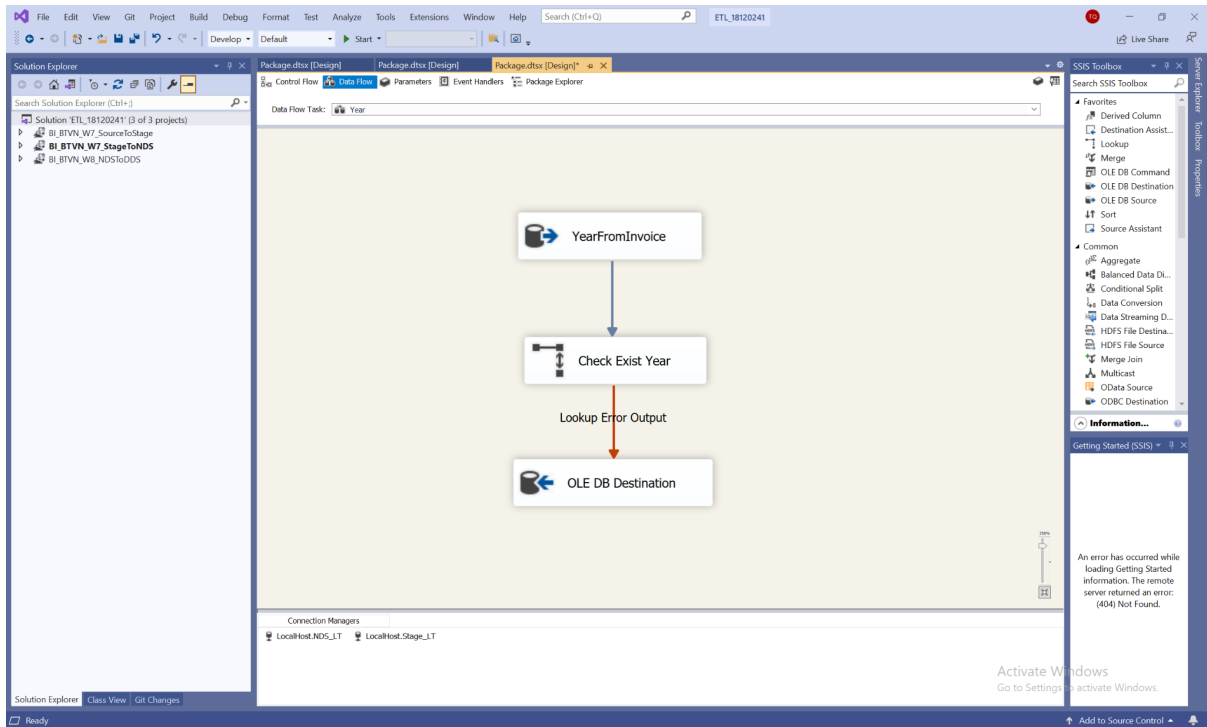
Lookup Column	Lookup Operation	Output Alias
CusID	<add as new column>	CusID

OK

Cancel

Help

- Ngày tháng năm từ NDS vào DDS em sẽ đổ trực tiếp vào, em chỉ kiểm tra nó có tồn tại hay chưa, nếu chưa thì cho vào còn có rồi thì không làm gì.



- Kết quả bảng Revenue ở database:

RevenueID	DayID	StoreID	ProductID	CustomerID	Revenue
1	959	15	443	620	29950
2	939	15	443	623	29950
3	684	985	443	121	29950
4	476	985	443	122	29950
5	1004	551	443	498	29950
6	930	551	443	498	29950
7	226	631	155	119	16353
8	746	631	155	120	16353
9	600	358	468	1612	4623
10	375	358	468	1614	4623
11	463	922	468	1111	4623
12	96	922	468	1112	4623
13	462	678	468	1695	4623
14	156	678	468	1697	4623
15	1052	777	100	434	13342
16	501	777	100	436	13342
17	177	362	100	85	13342
18	756	362	100	86	13342
19	336	58	358	563	27776
20	460	58	358	564	27776
21	557	714	182	209	29504
22	850	714	182	211	29504