

TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HCM
KHOA CÔNG NGHỆ THÔNG TIN



ĐỒ ÁN HỌC PHẦN

TÊN HỌC PHẦN: **Kho dữ liệu (DWH)**

MÃ SỐ LỚP HP:

Tên đề tài Nhóm: **Xây dựng Kho Dữ Liệu Doanh Số Game Khai Thác Tại Công Ty Phân Phối Game**

Lớp:

NHÓM 10

THÀNH VIÊN:	MSSV
Trần Minh Đức	20133037
Trần Văn Thắng	20133121
Trần Văn Trọng	20133103

TP.HCM, ngày 17 Tháng 4 năm 2023

LỜI CẢM ƠN

Chúng em xin chân thành cảm ơn sự hướng dẫn tận tình của thầy *Nguyễn Văn Thành* đã giúp chúng em hoàn thành bài báo cáo cuối kì bộ môn Kho dữ liệu. Tuy đã cố gắng học tập, tìm hiểu, phân tích thiết kế và xây dựng hệ thống nhưng vẫn còn thiếu sót là điều không thể tránh khỏi. Chúng em rất mong nhận được sự thông cảm và góp ý chân thành của Thầy. Chúng em xin chân thành cảm ơn.

MỤC LỤC

Chương 1. GIỚI THIỆU VỀ ĐỒ ÁN HỌC PHẦN	5
1.1. Tổng quan về ĐỒ ÁN HỌC PHẦN	5
1.2. Nội dung chuyên môn chính của ĐỒ ÁN HỌC PHẦN	5
1.3. Mục tiêu của ĐỒ ÁN HỌC PHẦN	5
1.4. Công cụ và nền tảng kỹ thuật thực hiện ĐỒ ÁN HỌC PHẦN	5
1.5. Sản phẩm của ĐỒ ÁN HỌC PHẦN	5
Chương 2. TỔNG QUAN VỀ KHO DỮ LIỆU GAME SALES KHAI THÁC TẠI CÔNG TY PHÂN PHỐI GAME.....	6
2.1. Ý tưởng hình thành DWH Game Sales khai thác tại công ty phân phối game	6
2.2. Giới thiệu chung về Kho dữ liệu	6
2.2.1. Mô tả tổng quan về DWH	6
2.2.2. Các dữ liệu nguồn hình thành DWH	6
2.3. Xác định nhu cầu tổ chức và phân tích dữ liệu của toàn bộ DWH	6
Chương 3. PHÂN TÍCH, THIẾT KẾ KHO DỮ LIỆU DOANH SỐ GAME KHAI THÁC TẠI CÔNG TY PHÂN PHỐI GAME.....	7
3.1. Mô tả hệ thống các DB gốc liên quan phân hệ DWH	7
3.2. Phân phân tích các DB gốc xác định yêu cầu phân tích DWH của phân hệ	10
3.3. Thiết kế DB mới tổ chức phân tích Kho dữ liệu cho phân hệ	10
Chương 4. TÍCH HỢP DỮ LIỆU VÀO KHO(SSIS)	12
4.1. Xử lí dữ liệu	12
4.2. Quá trình đổ dữ liệu vào Stage	22
4.2.1. Quá trình tạo Control Flow.....	22
4.2.2. Kết quả thực thi	31
4.2.3. Kết quả trong SQL Server	33
4.3. Quá trình đổ dữ liệu từ Kho Stage sang Kho Ware Hosue	41
4.3.1. Quá trình tạo Control Flow.....	41
4.3.2. Kết quả thực thi	50
4.3.3. Kết quả trên SQL Server	52

Chương 5. PHÂN TÍCH DỮ LIỆU(SSAS)	57
5.1. Quá trình xây dựng mô hình.....	57
5.1.1. Tạo SSAS Project	57
5.2. Thực hiện các câu truy vấn (SSAS, Pivot Table, T-SQL)	64
5.2.1. Câu 1: Tổng doanh thu của các thể loại game.....	64
5.2.2. Câu 2: Cho biết Doanh thu của các thể loại game theo từng năm?.....	65
5.2.3. Câu 3: Cho biết doanh thu của các thể loại game theo từng khu vực?	67
5.2.4. Câu 4: Cho biết Doanh thu của các thể loại game theo từng nền tảng?	68
5.2.5. Câu 5: Doanh thu của các thể loại game theo từng năm và tổng doanh thu của toàn thị trường?.....	70
5.2.6. Câu 6: Số lượng doanh thu của các nền tảng game theo từng năm?	71
5.2.7. Câu 7: Số lượng doanh thu của các nền tảng game theo từng khu vực?.....	72
Chương 6. TRỰC QUAN HÓA BẰNG POWER BI.....	73
6.1. Tổng doanh thu của các thể loại game	73
6.2. Tổng doanh thu theo từng năm(1980 - 2017).....	73
6.3. Tổng doanh thu theo khu vực	74
Phân chia công việc	75

Chương 1. GIỚI THIỆU VỀ ĐỒ ÁN HỌC PHẦN

1.1. Tổng quan về ĐỒ ÁN HỌC PHẦN

Đồ án Học phần môn Kho dữ liệu: **Phát Triển Mô Hình Kho Dữ Liệu
Doanh số về game Khai Thác Tại Công Ty Phân Phối Game**

1.2. Nội dung chuyên môn chính của ĐỒ ÁN HỌC PHẦN

Phân tích, thiết kế, chuyển đổi dữ liệu database gốc sang nhà Kho dữ liệu

1.3. Mục tiêu của ĐỒ ÁN HỌC PHẦN

Năm được cách phân tích thiết kế Data Warehouse theo nhu cầu phân tích dữ liệu

Năm được kỹ thuật tích hợp dữ liệu từ các nguồn về Data Warehouse

Năm được kỹ thuật phân tích bằng SSAS

1.4. Công cụ và nền tảng kỹ thuật thực hiện ĐỒ ÁN HỌC PHẦN

Các công cụ chính sử dụng trong đồ án này là

- + Microsoft SQL Server
- + SSAS
- + SSIS
- + Power BI

1.5. Sản phẩm của ĐỒ ÁN HỌC PHẦN

Kho dữ liệu “video_games_DW”

SSIS Project

SSAS Project

Các view phân tích, truy vấn Database gốc, DW

Chương 2. TỔNG QUAN VỀ KHO DỮ LIỆU GAME SALES KHAI THÁC TẠI CÔNG TY PHÂN PHỐI GAME

2.1. Ý tưởng hình thành DWH Game Sales khai thác tại công ty phân phối game

Ý tưởng: Xây dựng 1 Data Warehouse dưới dạng mô hình Star Schema có các Dim table (chứa các đối tượng cần phân tích) và Fact table (chứa các id của các đối tượng và 1 cột chứa giá trị được group by từ các id)

2.2. Giới thiệu chung về Kho dữ liệu

2.2.1. Mô tả tổng quan về DWH

1 tập đoàn chuyên phân phối game cho các thị trường trên thế giới có 1 hệ thống database lưu trữ doanh số bán game qua hằng năm.

2.2.2. Các dữ liệu nguồn hình thành DWH

Nguồn dữ liệu lấy trên Kaggle là Database video_games đặt tại trụ sở của công ty phân phối game.

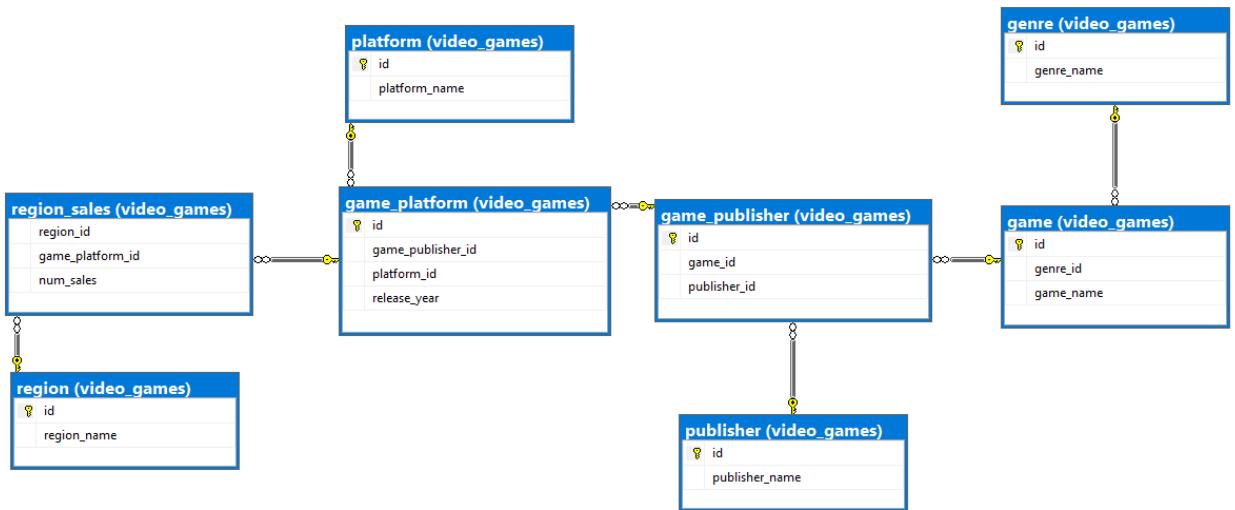
2.3. Xác định nhu cầu tổ chức và phân tích dữ liệu của toàn bộ DWH

Tập đoàn có nhu cầu phân tích các doanh số bán hàng đã thu thập qua các năm để hiểu thị trường, hiểu thể loại game nào được ưa thích, nhà phát hành, thể loại game nào đang làm mưa làm gió trên thị trường...

Chương 3. PHÂN TÍCH, THIẾT KẾ KHO DỮ LIỆU DOANH SỐ GAME KHAI THÁC TẠI CÔNG TY PHÂN PHỐI GAME

3.1. Mô tả hệ thống các DB gốc liên quan phân hệ DWH

Hệ thống Database lưu trữ



Mô tả DB:

region(id, region_name)

genre(id, genre_name)

publisher(id, publisher_name)

platform(id, platform_name)

game(id, genre_id, game_name)

game_publisher(id, game_id, publisher_id)

game_platform(id, game_publisher_id, platform_id, release_year)

region_sales(region_id, game_platform_id, num_sales)

Region				
STT	Thuộc tính	Kiểu dữ liệu	Ràng buộc	Diễn giải
1	<code>id</code>	<code>int</code>	PRIMARY KEY	Khoá chính phân biệt vùng
2	<code>region_name</code>	<code>nvarchar(150)</code>	NOT NULL	Tên vùng

Genre				
STT	Thuộc tính	Kiểu dữ liệu	Ràng buộc	Diễn giải
1	id	int	PRIMARY KEY	Khoá chính phân biệt thể loại
2	genre_name	nvarchar(150)	NOT NULL	Tên thể loại

Publisher				
STT	Thuộc tính	Kiểu dữ liệu	Ràng buộc	Diễn giải
1	id	int	PRIMARY KEY	Khoá chính phân biệt nhà phát hành
2	publisher_name	nvarchar(150)	NOT NULL	Tên nhà phát hành

Platform				
STT	Thuộc tính	Kiểu dữ liệu	Ràng buộc	Diễn giải
1	id	int	PRIMARY KEY	Khoá chính phân biệt nền tảng
2	platform_name	nvarchar(150)	NOT NULL	Tên nền tảng

Game				
STT	Thuộc tính	Kiểu dữ liệu	Ràng buộc	Diễn giải
1	id	int	PRIMARY KEY	Khoá chính phân biệt game
2	game_name	nvarchar(150)	NOT NULL	Tên game

game_publisher				
STT	Thuộc tính	Kiểu dữ liệu	Ràng buộc	Diễn giải

1	id	int	PRIMARY KEY	Khoá chính phân biệt game và nhà phát hành
2	game_id	int	NOT NULL	ID game
3	publisher_id	int	NOT NULL	ID nhà phát hành

game_platform				
STT	Thuộc tính	Kiểu dữ liệu	Ràng buộc	Diễn giải
1	id	int	PRIMARY KEY	Khoá chính phân biệt game và nền tảng
2	game_publisher_id	int	NOT NULL	ID game và nhà phát hành
3	platform_id	int	NOT NULL	ID nền tảng
4	release_year	int	NOT NULL	Năm phát hành

region_sale				
STT	Thuộc tính	Kiểu dữ liệu	Ràng buộc	Diễn giải
1	region_id	int	PRIMARY KEY	ID khu vực
2	game_platform_id	int	PRIMARY KEY	ID game và nền tảng
3	num_sales	Decimal(5,2)	NOT NULL	Đơn vị bán (1.000.000 USD)

3.2. Phân phân tích các DB gốc xác định yêu cầu phân tích DWH của phân hệ

Dim:

DimRegion: Dim về khu vực/thị trường bán game

DimGenre: Dim về thể loại game

DimPublisher: Dim về nhà phát hành

DimPlatform: Dim về nền tảng phát triển

DimGame: Dim về game

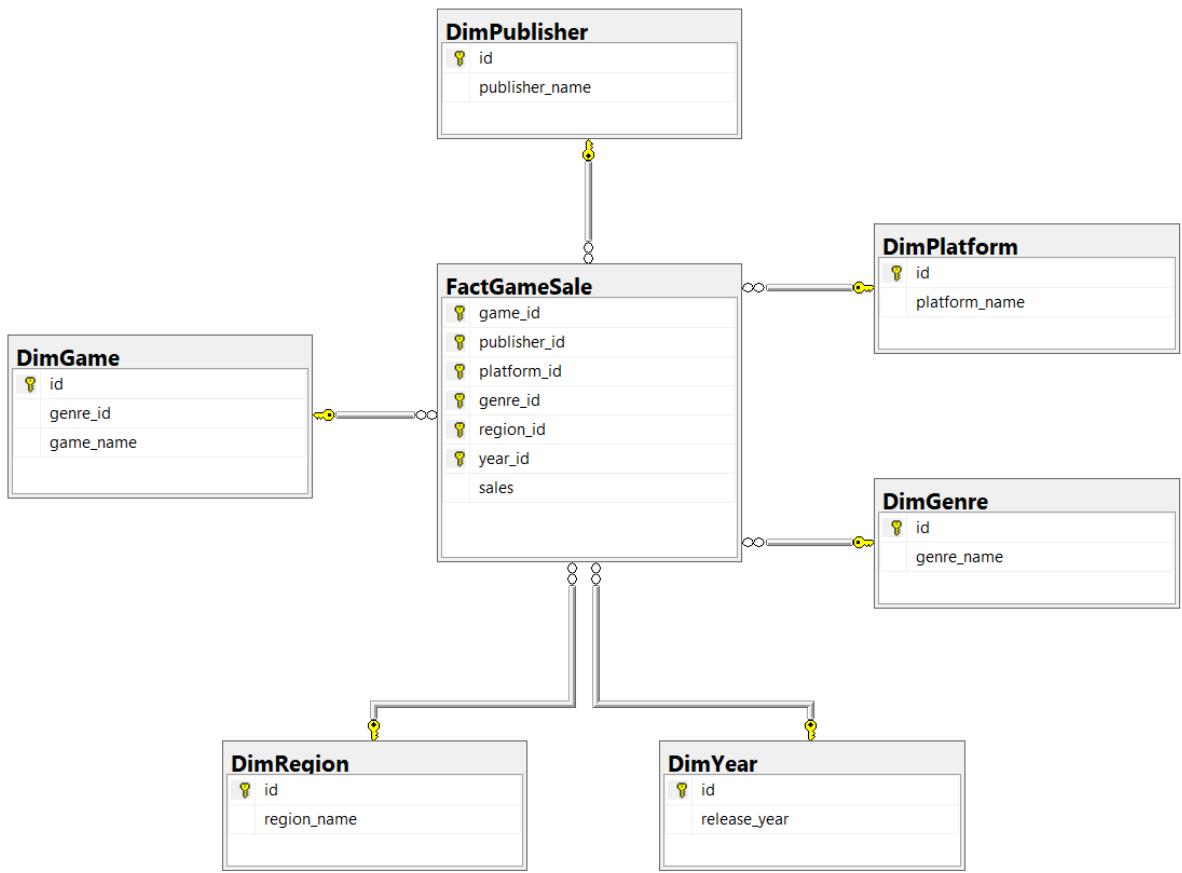
DimYear: Dim về năm phát hành

Có 1 fact:

Fact doanh thu = sum(num_sales) group by (id DimRegion, id DimGenre, id DimPublisher, id DimPlatform, id DimGame, id DimYear)

3.3. Thiết kế DB mới tổ chức phân tích Kho dữ liệu cho phân hệ

DW chung :



DW phân hệ phân tích doanh thu theo nhà phát hành và khu vực



Chương 4. TÍCH HỢP DỮ LIỆU VÀO KHO(SSIS)

4.1. Xử lý dữ liệu

- Dùng python để chuyển dữ liệu từ file csv thành các bảng trong DB Stage
- Sau khi xử lý xong ta tạo được các file csv với nội dung là dữ liệu của các bảng:

*Region

The screenshot shows a Microsoft Excel spreadsheet titled "Region.csv". The table has two columns: "id" (Column A) and "region_name" (Column B). The data is as follows:

	A	B	C	D	E	F	G	H
1	id	region_name						
2	1	North America						
3	2	Europe						
4	3	Japan						
5	4	Other						
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								

*Year

The screenshot shows a Microsoft Excel spreadsheet titled "Year.csv". The ribbon menu is visible at the top, with "Home" selected. The main area displays a table with two columns: "A" and "B". Column A contains row numbers from 1 to 27. Column B contains years from 1980 to 2005. The font is set to Calibri, size 11, and the text color for column B is red.

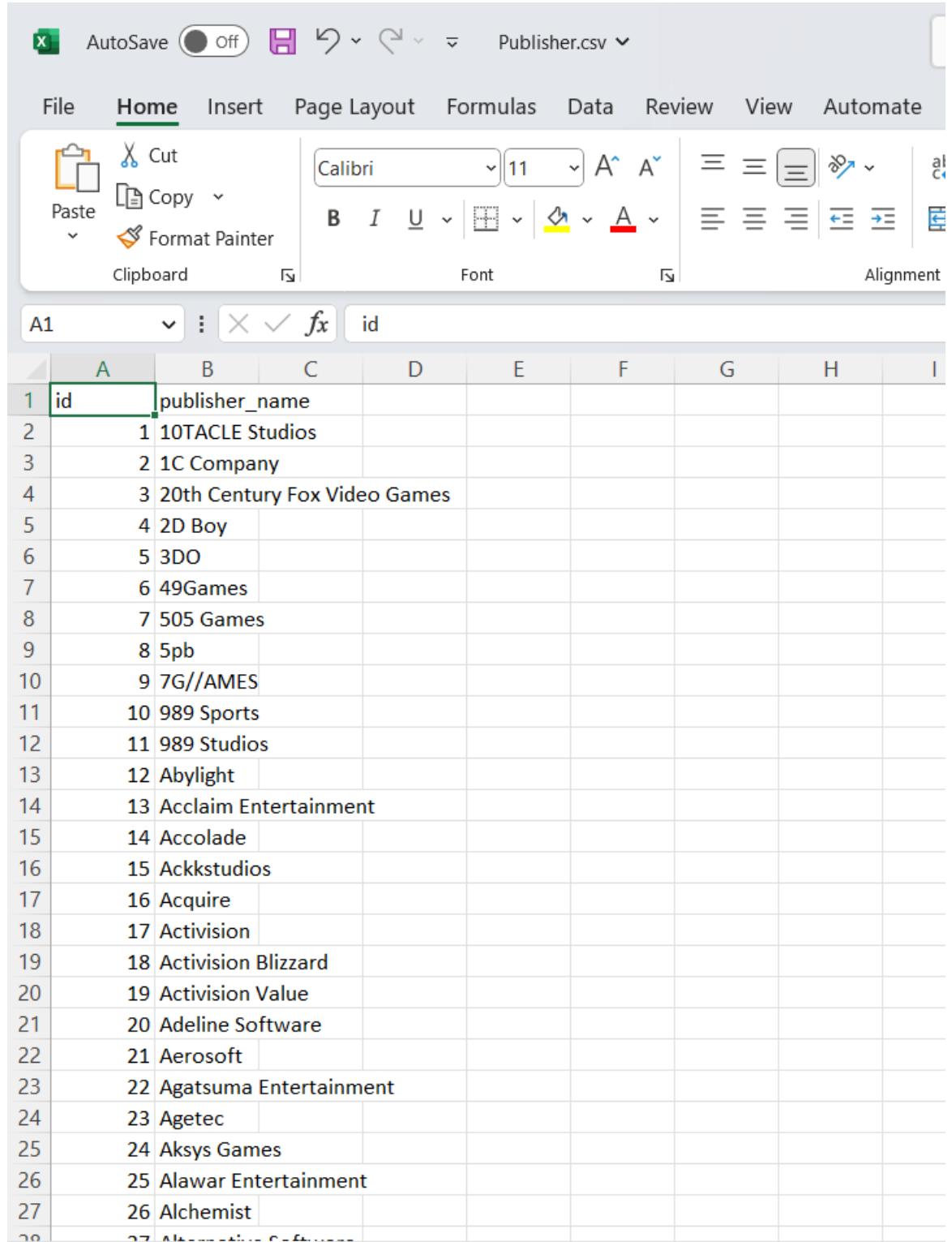
	A	B	C	D	E	F	G	H
1		Year						
2	1	1980						
3	2	1981						
4	3	1982						
5	4	1983						
6	5	1984						
7	6	1985						
8	7	1986						
9	8	1987						
10	9	1988						
11	10	1989						
12	11	1990						
13	12	1991						
14	13	1992						
15	14	1993						
16	15	1994						
17	16	1995						
18	17	1996						
19	18	1997						
20	19	1998						
21	20	1999						
22	21	2000						
23	22	2001						
24	23	2002						
25	24	2003						
26	25	2004						
27	26	2005						
28	--	-----						

*Platform

The screenshot shows a Microsoft Excel spreadsheet titled "Platform.csv". The "Home" tab is selected in the ribbon. The data consists of two columns: "id" and "platform_name". The "platform_name" column lists various gaming platforms, starting with "Wii" and ending with "WS". The cell containing "2600" is currently selected.

	A	B	C	D	E	F	G	H	I
1	id	platform_name							
2	1	Wii							
3	2	NES							
4	3	GB							
5	4	DS							
6	5	X360							
7	6	PS3							
8	7	PS2							
9	8	SNES							
10	9	GBA							
11	10	3DS							
12	11	PS4							
13	12	N64							
14	13	PS							
15	14	XB							
16	15	PC							
17	16	2600							
18	17	PSP							
19	18	XOne							
20	19	GC							
21	20	WiiU							
22	21	GEN							
23	22	DC							
24	23	PSV							
25	24	SAT							
26	25	SCD							
27	26	WS							

*Publisher



The screenshot shows a Microsoft Excel spreadsheet titled "Publisher.csv". The "Home" tab is selected in the ribbon. The data consists of two columns: "id" (in column A) and "publisher_name" (in column B). The first few rows of data are as follows:

	A	B	C	D	E	F	G	H	I
1	id	publisher_name							
2	1	10TACLE Studios							
3	2	1C Company							
4	3	20th Century Fox Video Games							
5	4	2D Boy							
6	5	3DO							
7	6	49Games							
8	7	505 Games							
9	8	5pb							
10	9	7G//AMES							
11	10	989 Sports							
12	11	989 Studios							
13	12	Abylight							
14	13	Acclaim Entertainment							
15	14	Accolade							
16	15	Ackkstudios							
17	16	Acquire							
18	17	Activision							
19	18	Activision Blizzard							
20	19	Activision Value							
21	20	Adeline Software							
22	21	Aerosoft							
23	22	Agatsuma Entertainment							
24	23	Agetec							
25	24	Aksys Games							
26	25	Alawar Entertainment							
27	26	Alchemist							
28	27	Alteonsoft Software							

*Genre

The screenshot shows a Microsoft Excel spreadsheet titled "Genre.csv". The ribbon menu is visible at the top, with "Home" selected. The "Clipboard" group contains "Cut", "Copy", and "Format Painter" buttons. The "Font" group includes "Calibri" font, size "11", bold ("B"), italic ("I"), underline ("U"), and various color and style options. The "Align" group includes horizontal alignment buttons. The formula bar shows "F11" and a dropdown arrow. The main area displays a list of 12 genres in columns A and B:

	A	B
1		Genre
2	1	Action
3	2	Adventure
4	3	Fighting
5	4	Misc
6	5	Platform
7	6	Puzzle
8	7	Racing
9	8	Role-Playing
10	9	Shooter
11	10	Simulation
12	11	Sports
13	12	Strategy
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		

*Game

The screenshot shows a Microsoft Excel spreadsheet titled "Game.csv". The ribbon is visible at the top with tabs for File, Home, Insert, Page Layout, Formulas, Data, Review, View, and Automate. The "Home" tab is selected. The toolbar below the ribbon includes options for Cut, Copy, Paste, Format Painter, and Clipboard. The font section of the toolbar shows Calibri 11pt. The alignment section shows horizontal and vertical alignment tools.

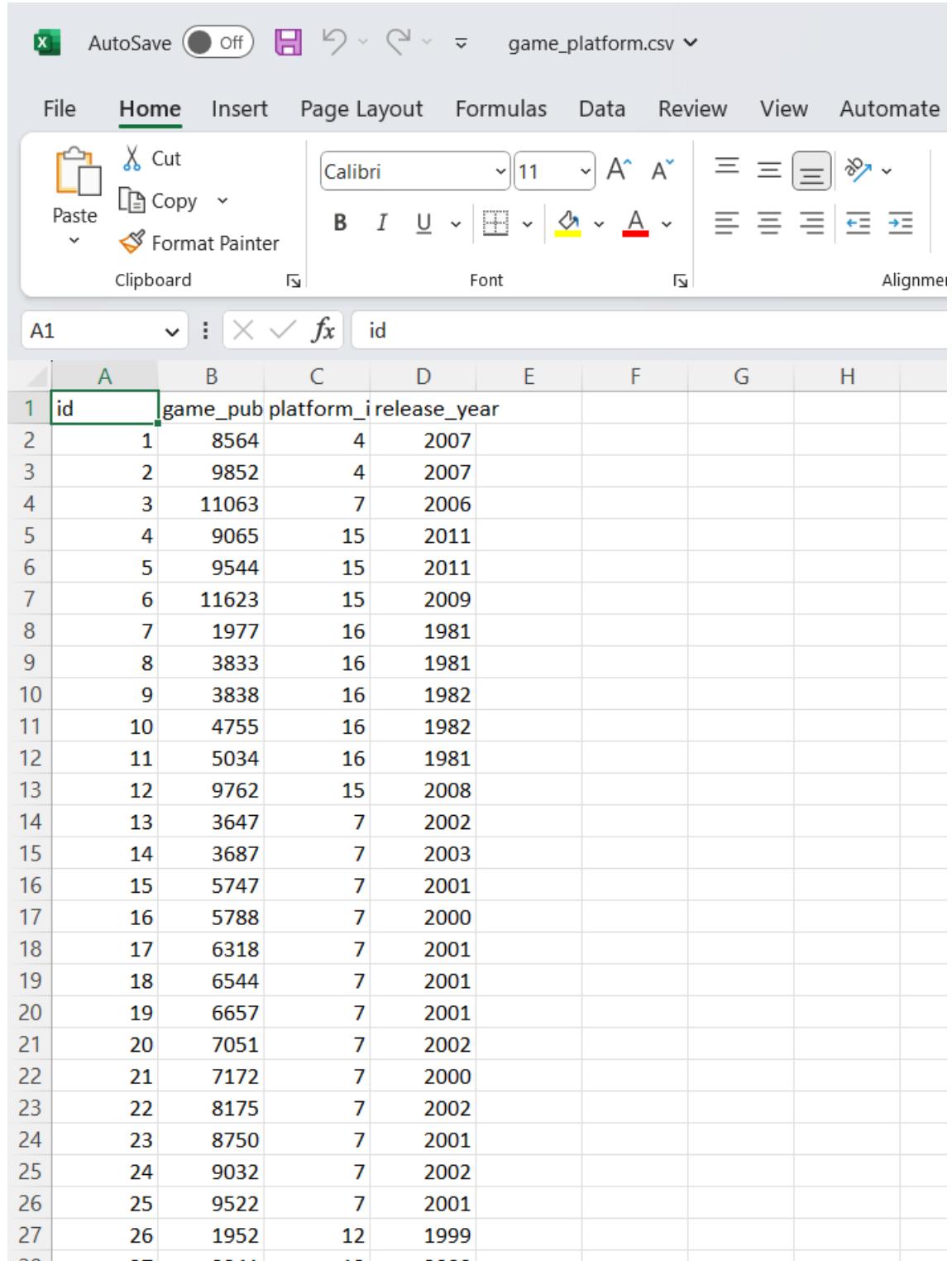
	A	B	C	D	E	F	G	H
1		Name	Genre					
2	1	'98 Koshien	8					
3	2	.hack//G.U.	8					
4	3	.hack//G.U.	8					
5	4	.hack//G.U.	8					
6	5	.hack//G.U.	8					
7	6	.hack//Infection	8					
8	7	.hack//Link	8					
9	8	.hack//Mu	8					
10	9	.hack//Outer	8					
11	10	.hack//Quarantine	1					
12	11	.hack: Sekai	7					
13	12	007 Racing	1					
14	13	007: Quantum	1					
15	14	007: The Venture	9					
16	15	007: Tomcat	4					
17	16	1 vs. 100	2					
18	17	1/2 Summoner	11					
19	18	10 Minute	6					
20	19	100 All-Time	4					
21	20	100 Classic	4					
22	21	100 Classic	4					
23	22	1000 Cook	1					
24	23	1001 Touch	6					
25	24	101-in-1 Edition	4					
26	25	101-in-1 Pack	11					
27	26	101-in-1 Special	11					

*Game_Publisher

The screenshot shows a Microsoft Excel spreadsheet titled "game_publisher.csv". The table contains three columns: "id", "game_id", and "publisher_id". The data consists of 27 rows, each containing a unique ID, a game ID, and a publisher ID. The publisher ID for all entries is 369, except for row 17 where it is 328 and row 24 where it is 498.

	A	B	C	D	E	F	G	H
1	id	game_id	publisher_id					
2	1	10866	369					
3	2	9244	369					
4	3	5464	369					
5	4	10868	369					
6	5	7282	369					
7	6	9603	369					
8	7	6573	369					
9	8	10864	369					
10	9	6576	369					
11	10	2556	369					
12	11	6729	369					
13	12	5463	369					
14	13	7265	369					
15	14	10858	369					
16	15	10859	369					
17	16	4866	328					
18	17	3661	498					
19	18	3665	498					
20	19	9258	369					
21	20	992	369					
22	21	7262	369					
23	22	9251	369					
24	23	9246	369					
25	24	3666	498					
26	25	7283	369					
27	26	7254	369					
28	27					

*Game_Platform



The screenshot shows a Microsoft Excel spreadsheet titled "game_platform.csv". The data is organized into columns A through H. Column A contains numerical IDs from 1 to 27. Column B contains game publication numbers (e.g., 8564, 9852). Column C contains platform identifiers (e.g., 4, 7). Column D contains release years (e.g., 2007, 2011). The rows show a sequence of data points, likely representing game publications across different platforms over time.

	A	B	C	D	E	F	G	H
1	id	game_pub	platform_i	release_year				
2	1	8564	4	2007				
3	2	9852	4	2007				
4	3	11063	7	2006				
5	4	9065	15	2011				
6	5	9544	15	2011				
7	6	11623	15	2009				
8	7	1977	16	1981				
9	8	3833	16	1981				
10	9	3838	16	1982				
11	10	4755	16	1982				
12	11	5034	16	1981				
13	12	9762	15	2008				
14	13	3647	7	2002				
15	14	3687	7	2003				
16	15	5747	7	2001				
17	16	5788	7	2000				
18	17	6318	7	2001				
19	18	6544	7	2001				
20	19	6657	7	2001				
21	20	7051	7	2002				
22	21	7172	7	2000				
23	22	8175	7	2002				
24	23	8750	7	2001				
25	24	9032	7	2002				
26	25	9522	7	2001				
27	26	1952	12	1999				

*Region_sale

AutoSave (Off) region_sales.csv

File Home Insert Page Layout Formulas Data Review View Automate

Cut Copy Format Painter Clipboard

Font Alignment

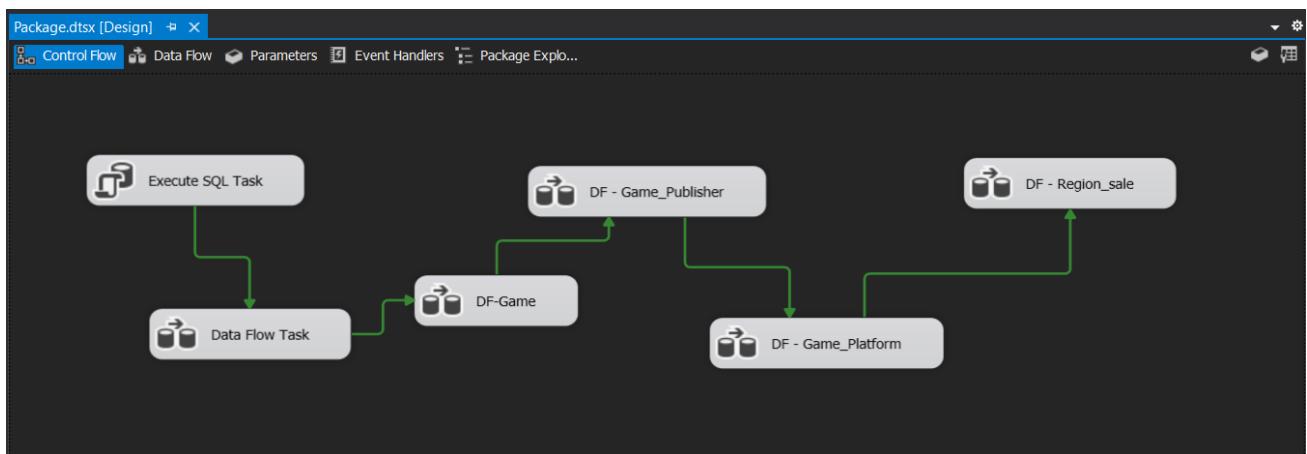
A1 : X ✓ fx region_id

	A	B	C	D	E	F	G	H
1	region_id	game_plat	num_sales					
2	1	50	3.5					
3	1	51	1.43					
4	1	52	0.51					
5	1	53	0.27					
6	1	54	0.48					
7	1	55	0.34					
8	1	56	0.36					
9	1	57	0.36					
10	1	58	0.17					
11	1	59	0.04					
12	1	60	0.21					
13	1	61	0.11					
14	1	62	0.13					
15	1	63	0.12					
16	1	64	0.08					
17	1	65	0.1					
18	1	66	0.09					
19	1	67	0.09					
20	1	68	0.07					
21	1	69	0.07					
22	1	70	0					
23	1	71	0.06					
24	1	72	0.06					
25	1	73	0					
26	1	74	0.04					
27	1	75	0.04					
28	1	76	0.04					

4.2. Quá trình đổ dữ liệu vào Stage

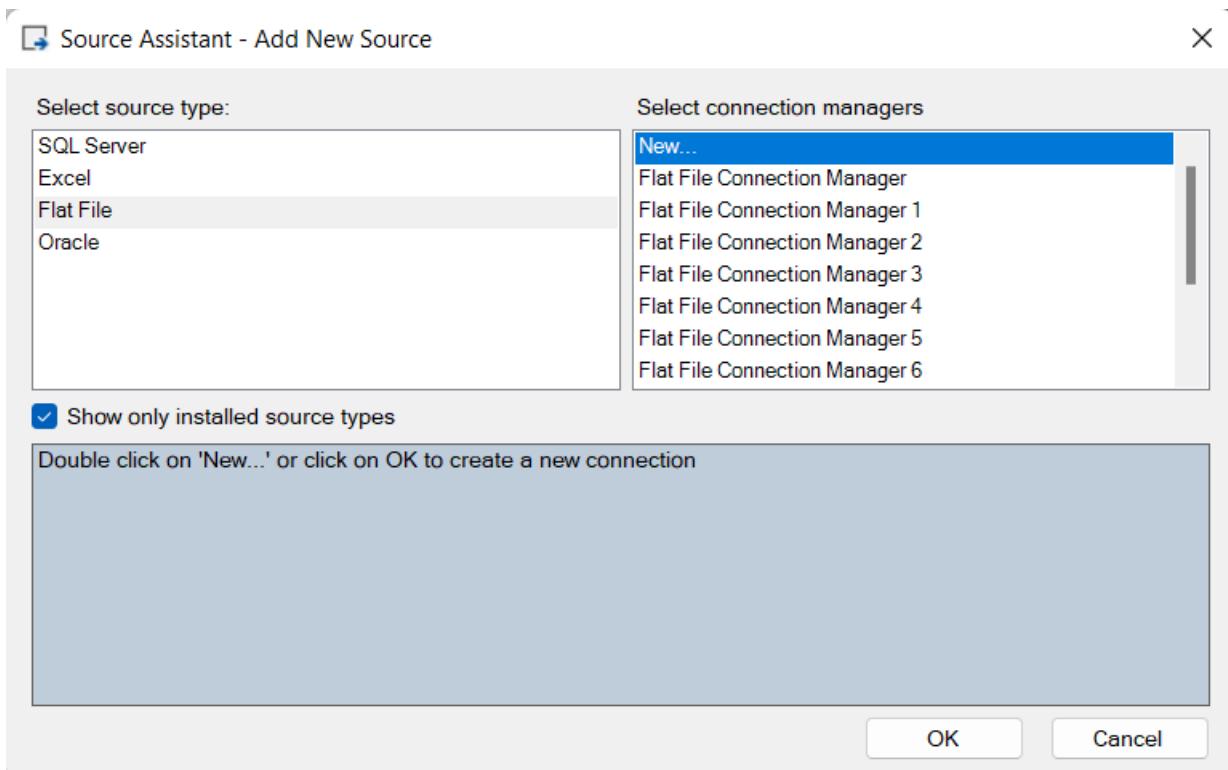
4.2.1. Quá trình tạo Control Flow

Kéo một Execute SQL Task và 5 Data Flow vào Control Flow và đặt tên chúng như sau:

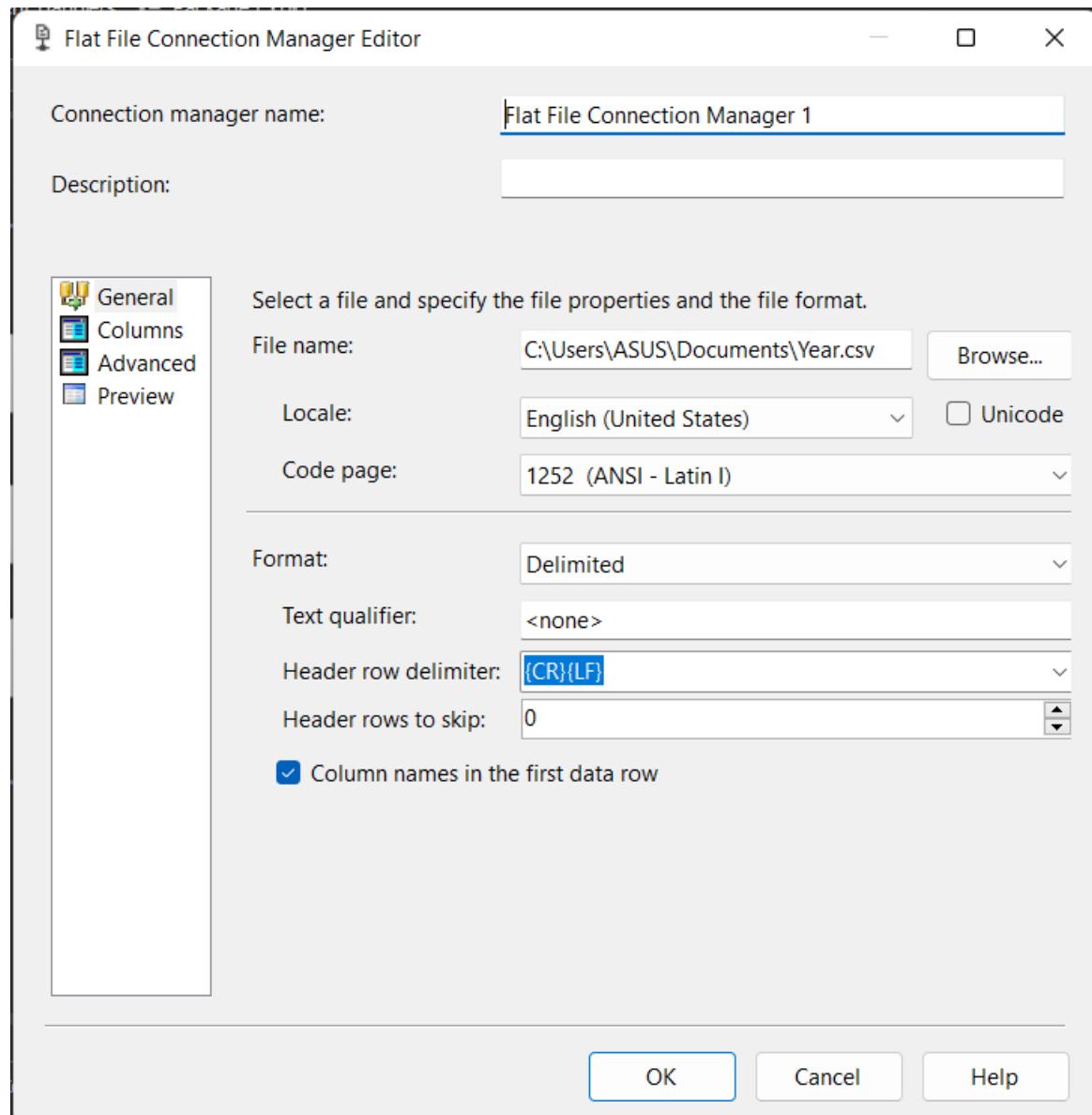


Data Flow Task:

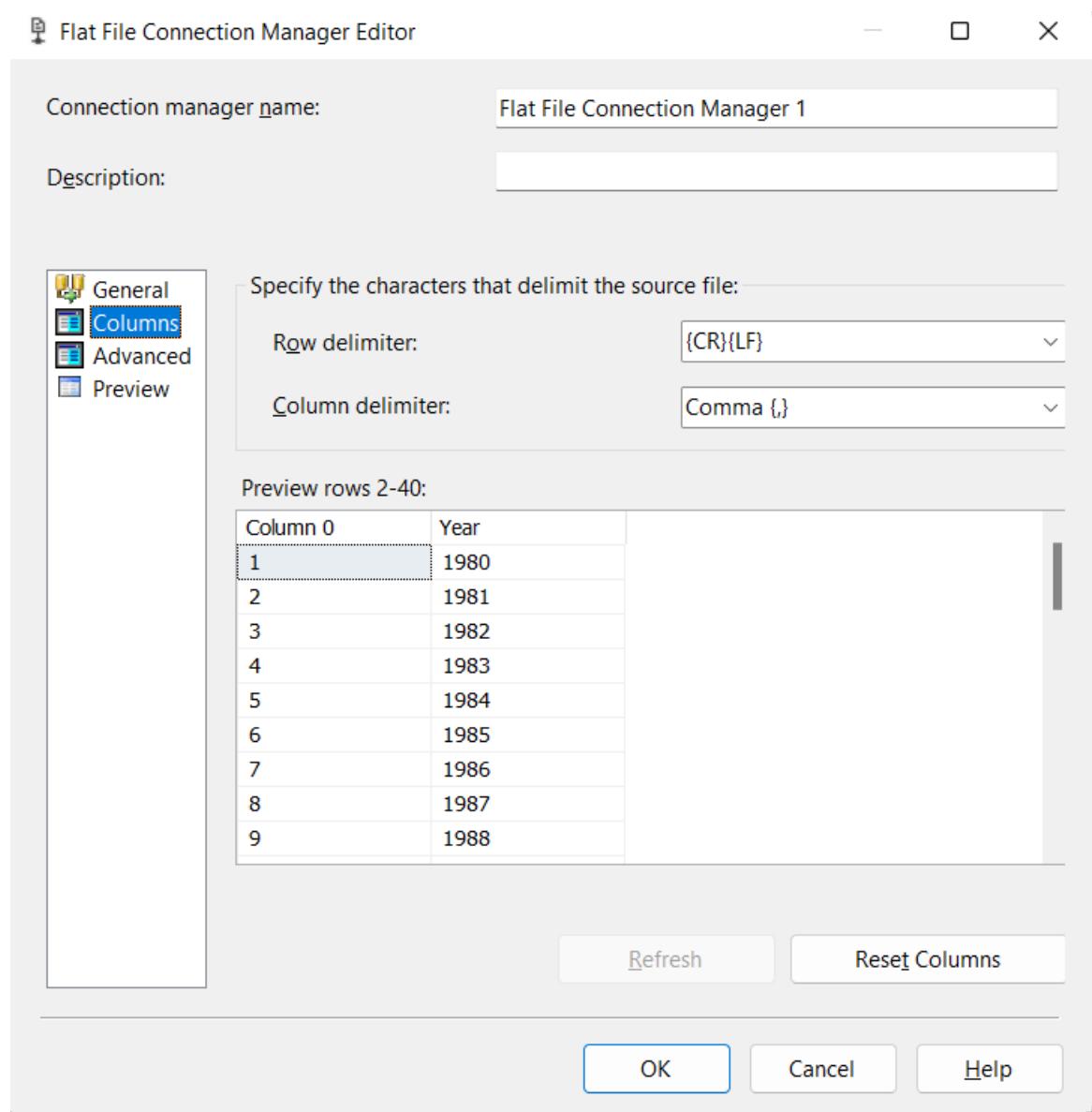
Kéo Source Assistant và chọn Flat File:



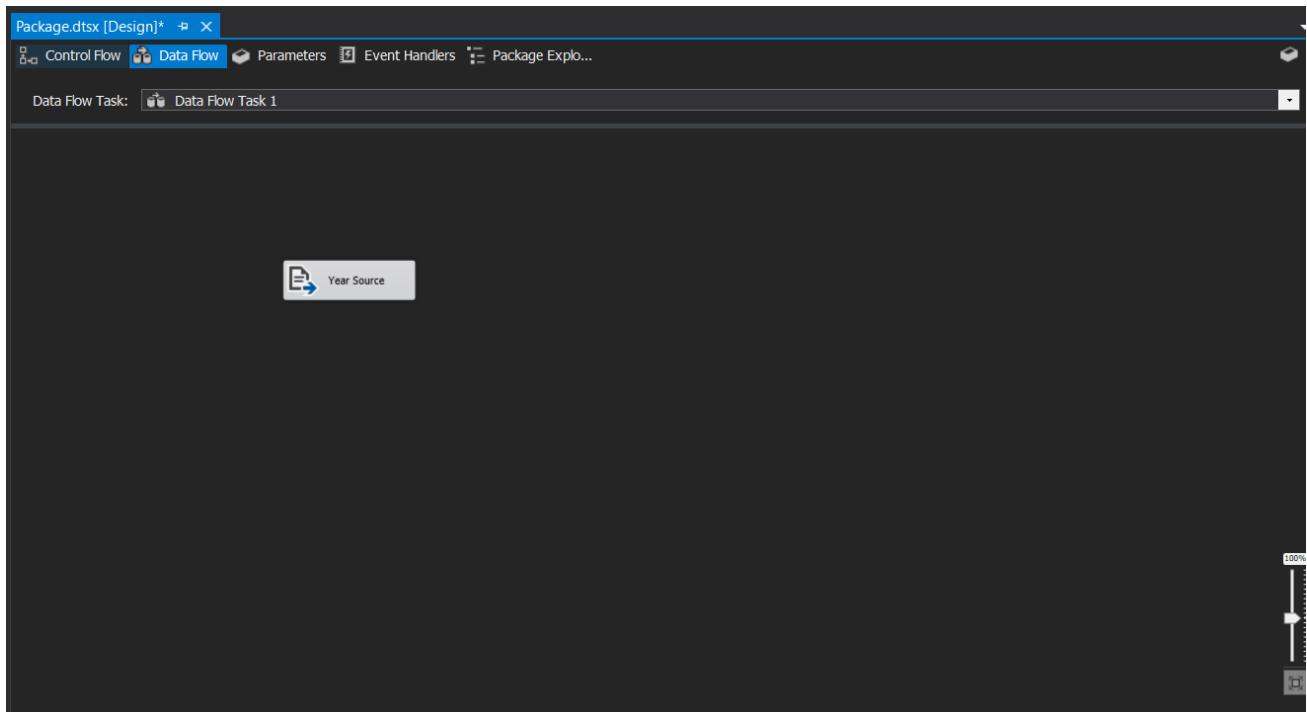
Bấm vào New để tạo một Flat File Connection Manager sau đó chọn file csv của bảng Year



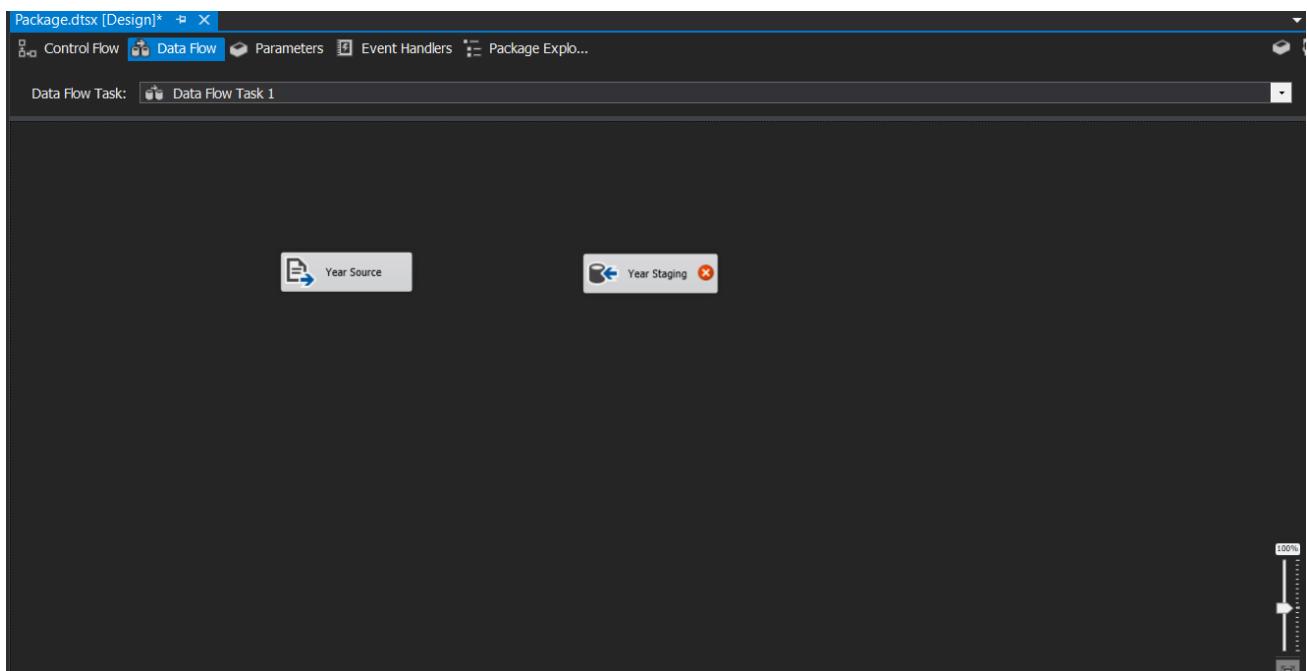
Bấm qua cột Columns để xem thông tin các cột trong bảng



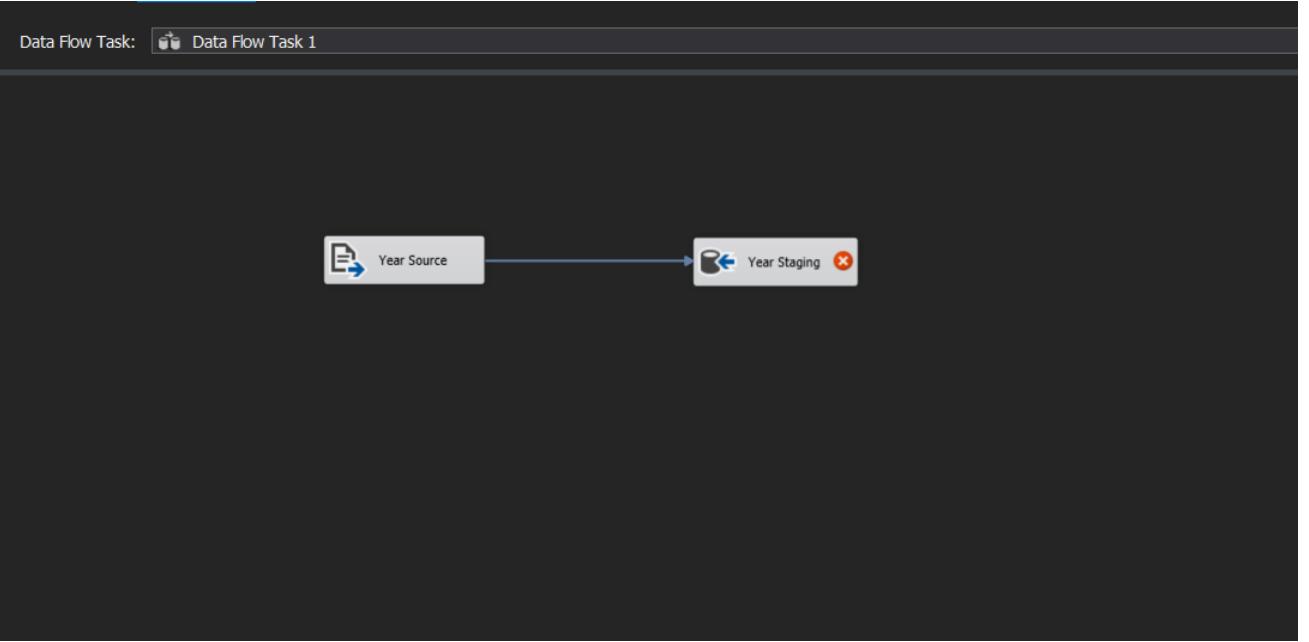
Sau đó bấm OK ta được một Flat File source



Kéo một Destination Assistant vào chọn DB stage đã tạo trước đó sau đó nhấn OK và đổi tên thành Year Staging:



Kéo một kết nối giữa đích và nguồn



Ở Destination ta chọn bảng Year

OLE DB Destination Editor

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

Connection Manager

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. For fast-load data access, set the table update options.

OLE DB connection manager: ./video_games 1

Data access mode: Table or view - fast load

Name of the table or the view: [video_games].[year]

Keep identity Table lock

Keep nulls Check constraints

Rows per batch:

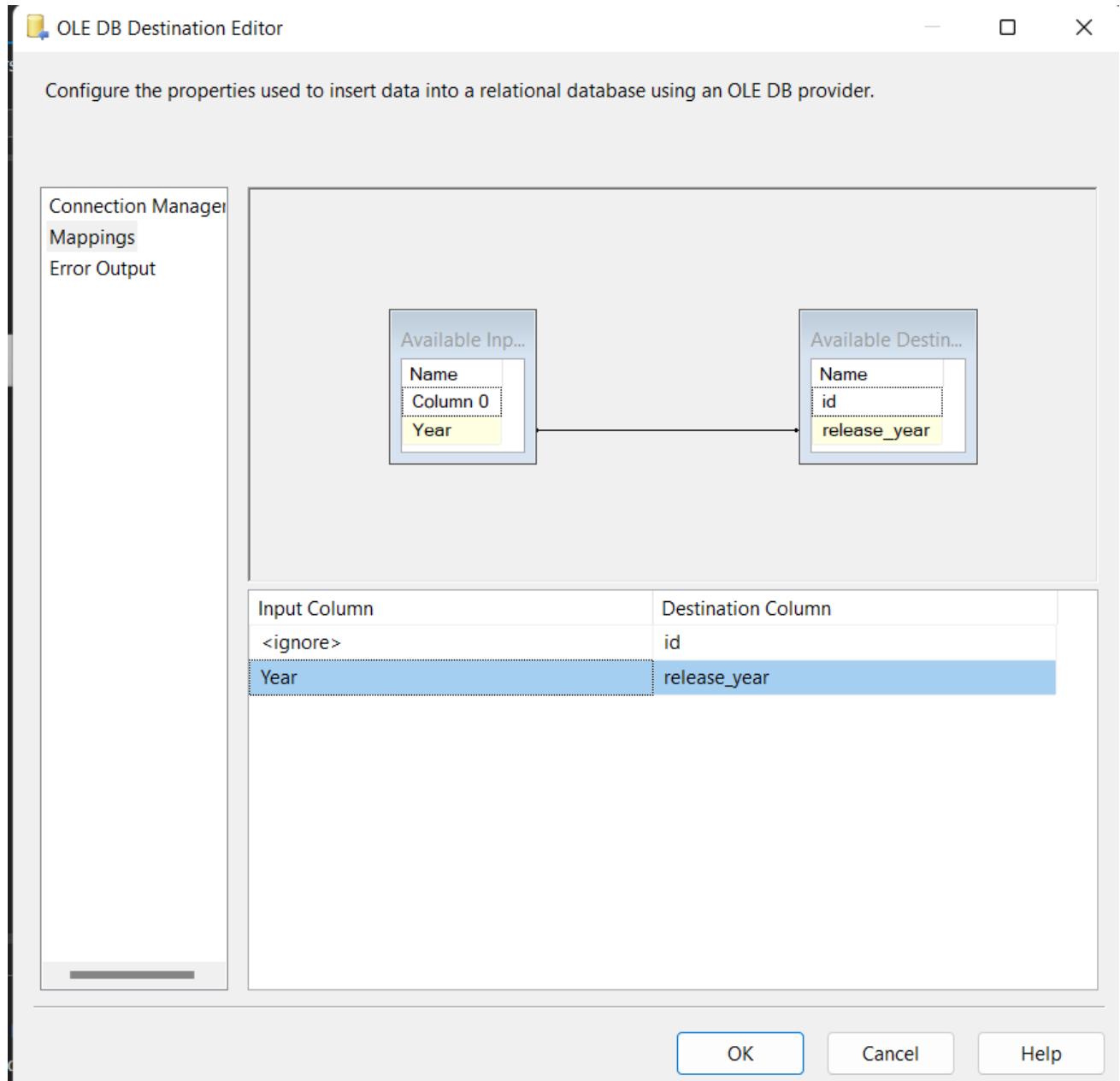
Maximum insert commit size: 2147483647

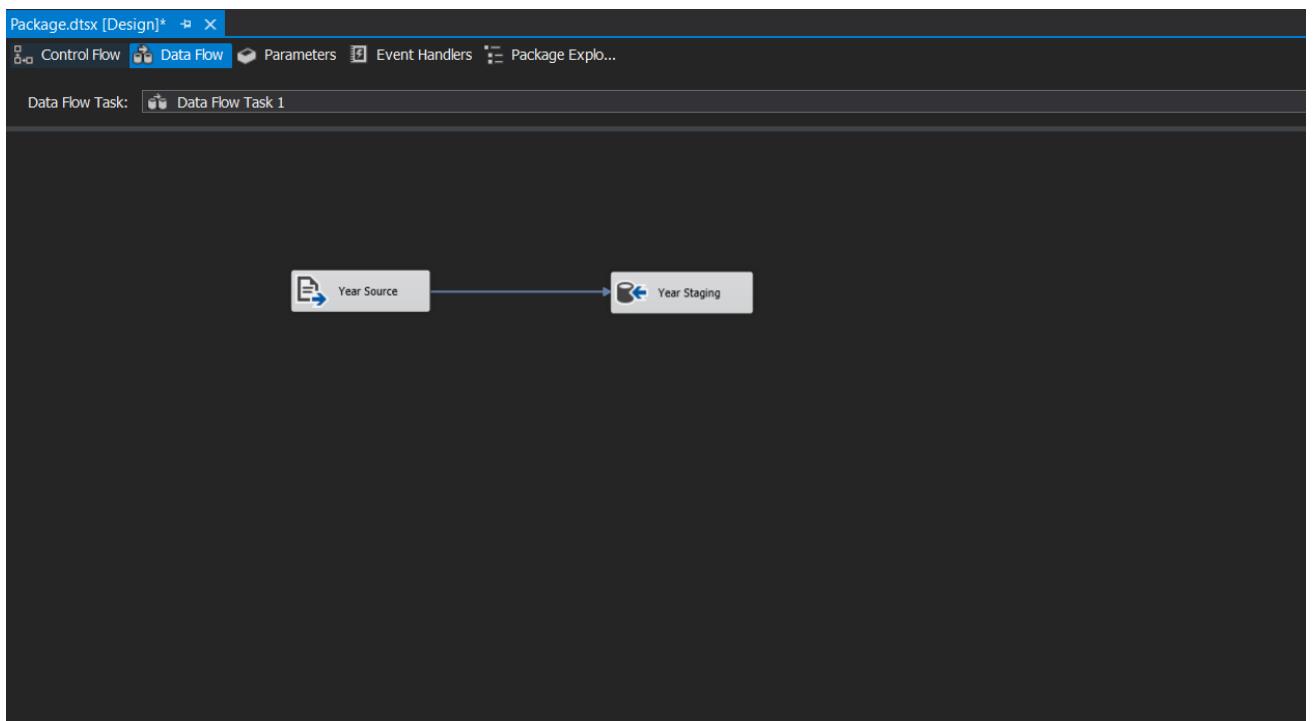
View Existing

Map the columns on the Mappings page.

OK **Cancel** **Help**

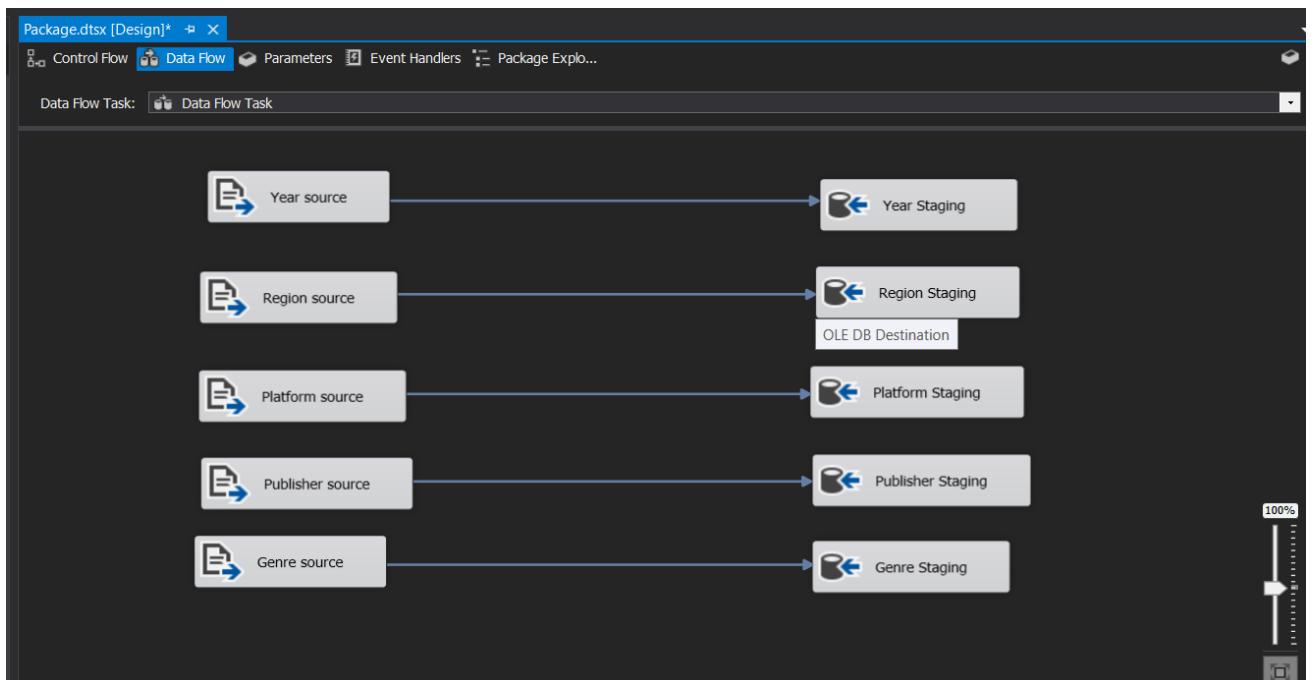
Sau đó Mapping các columns cho phù hợp



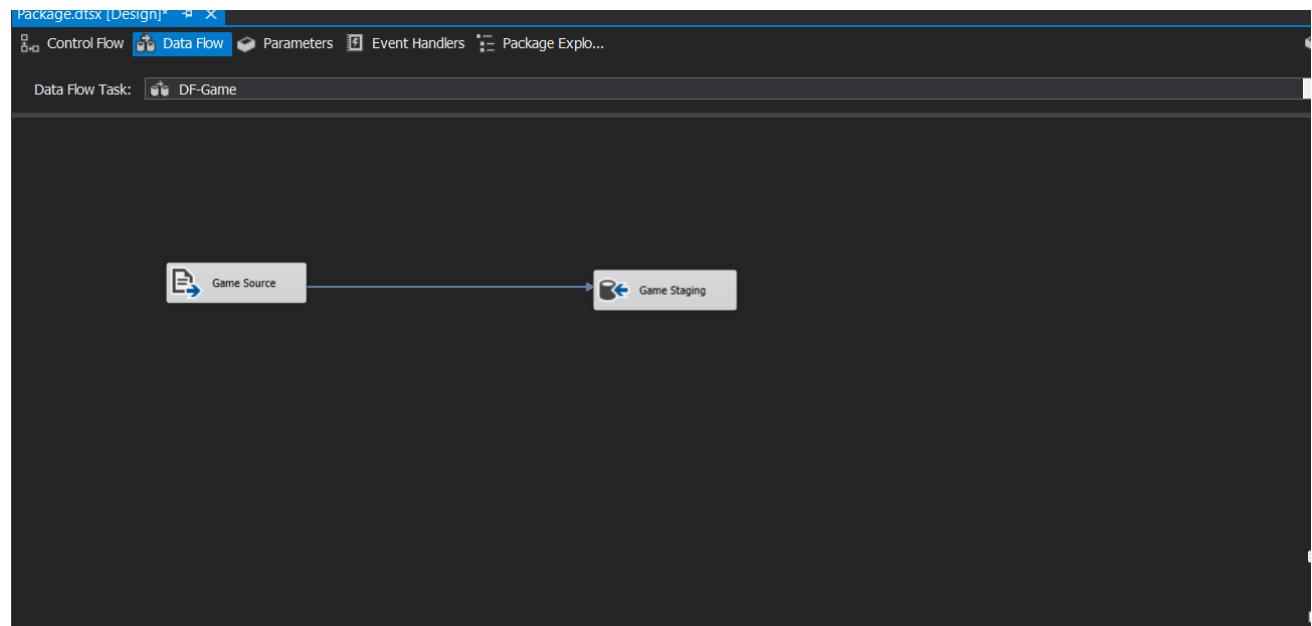


Làm tương tự đối với các file csv khác ta thu được các Data Flow task
như sau:

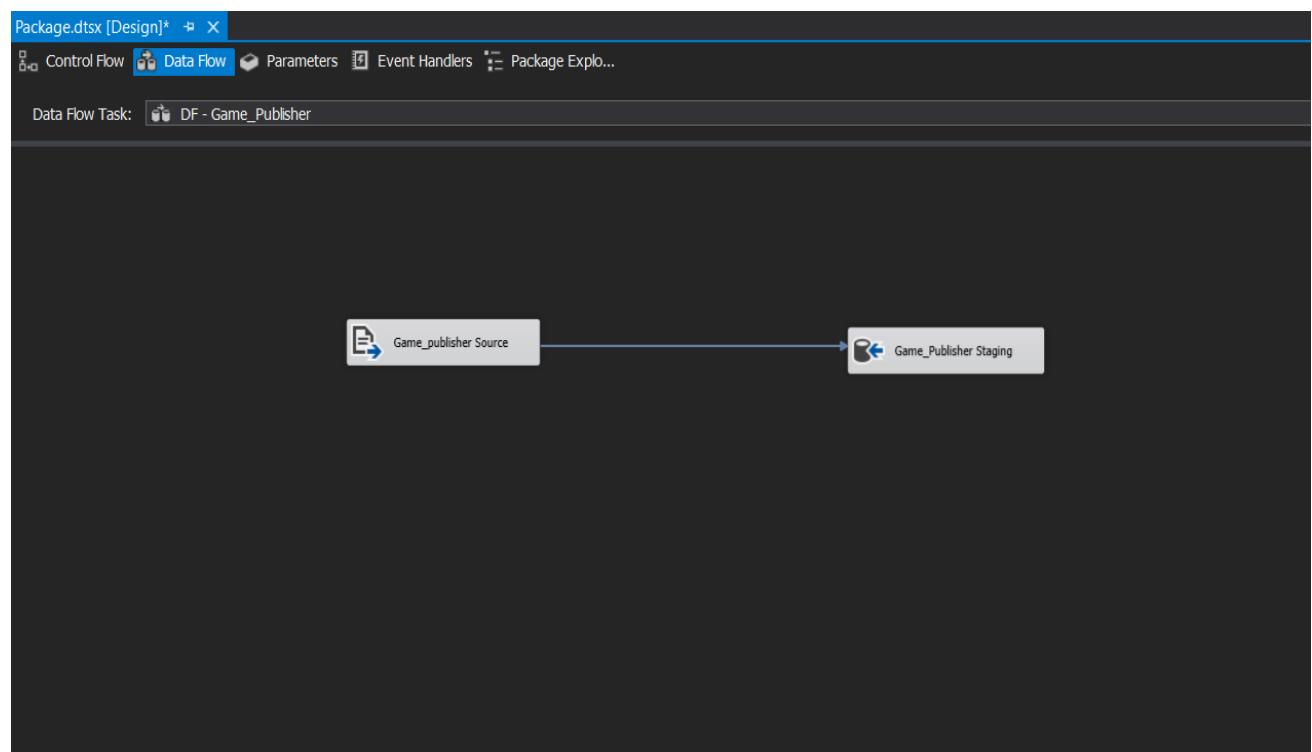
Data Flow task:



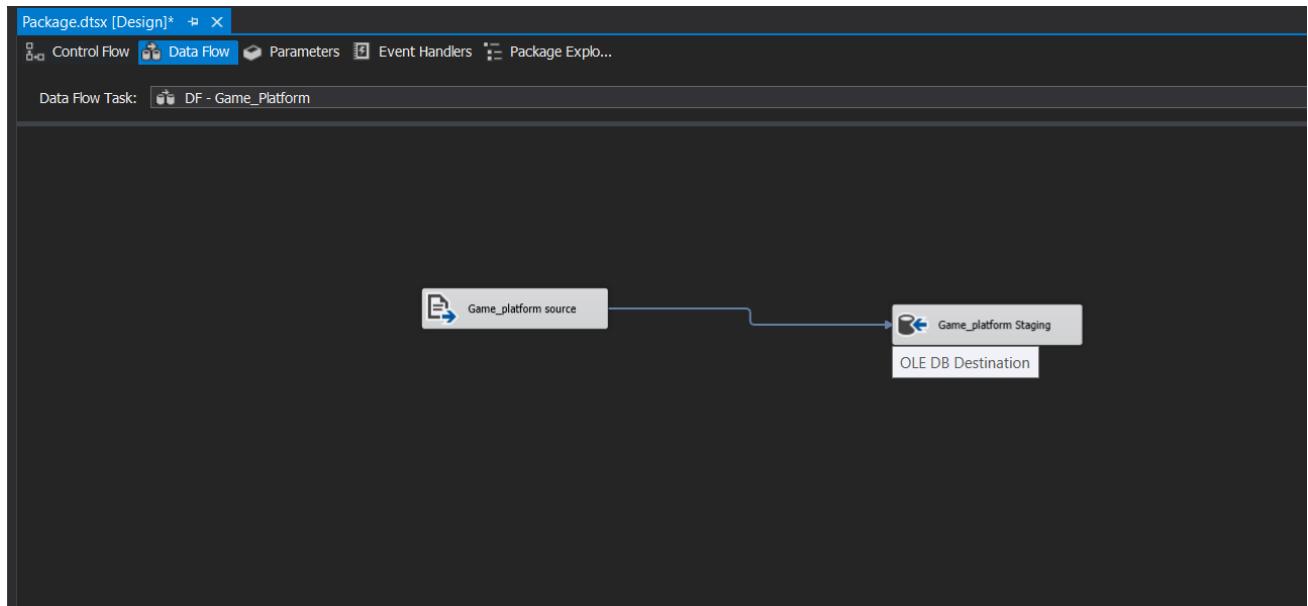
DF - Game



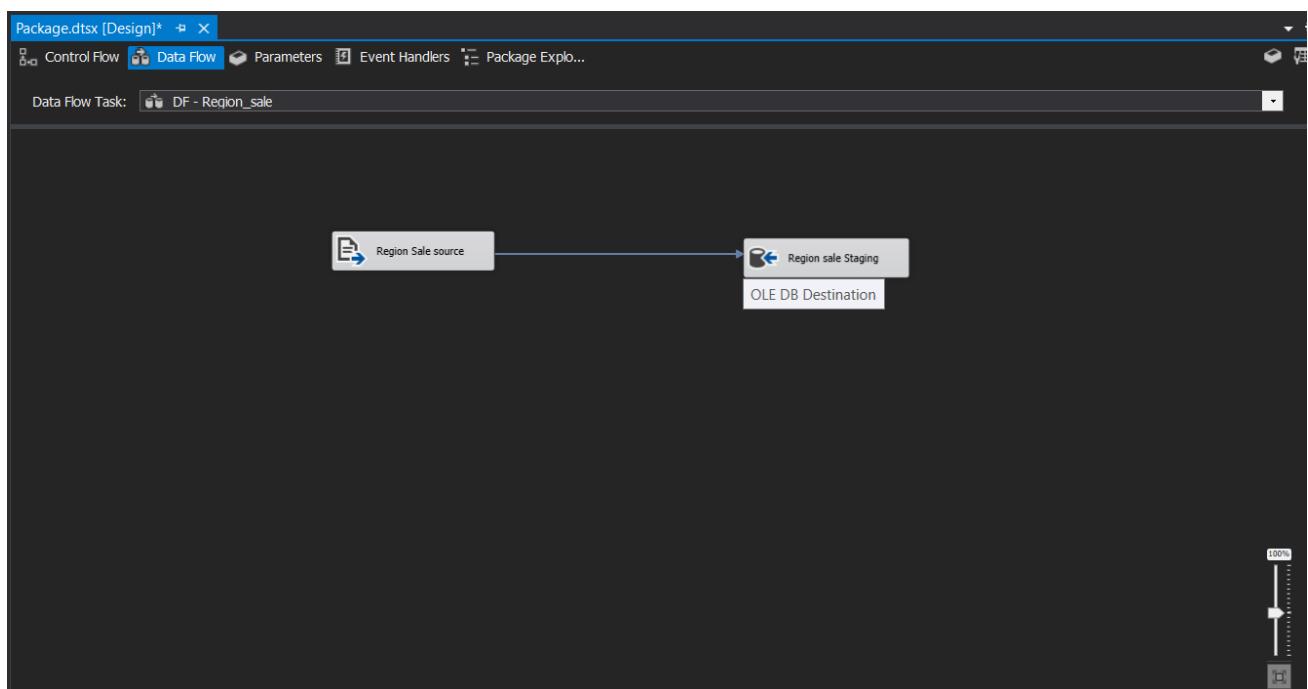
DF – Game_Publisher



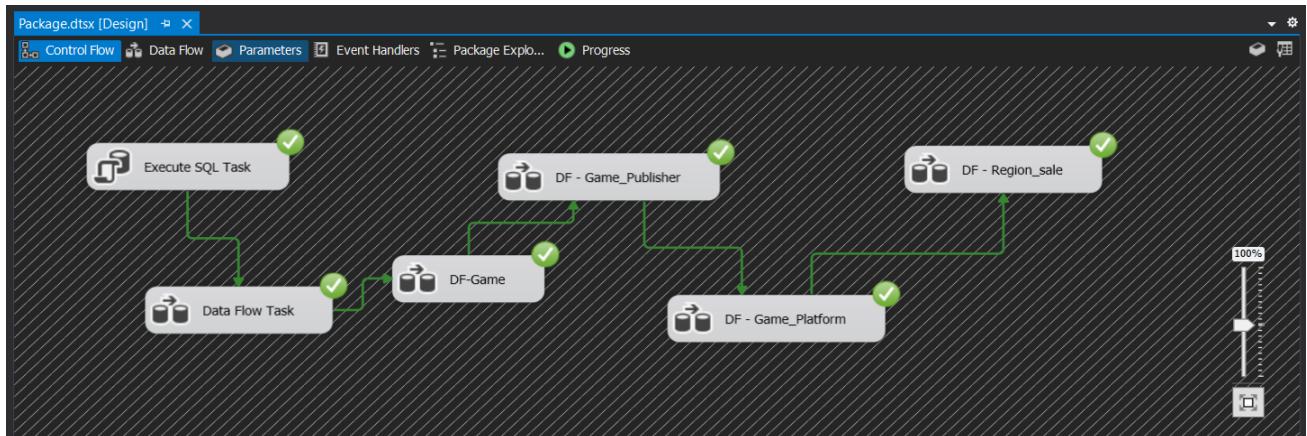
DF – Game_Platform



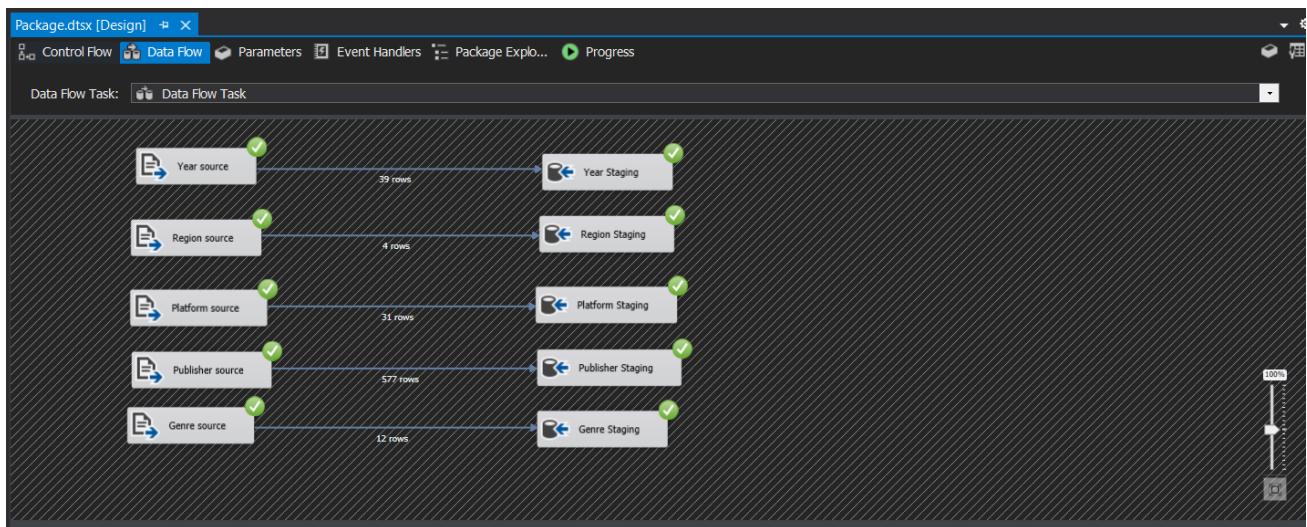
DF – Region_Sale



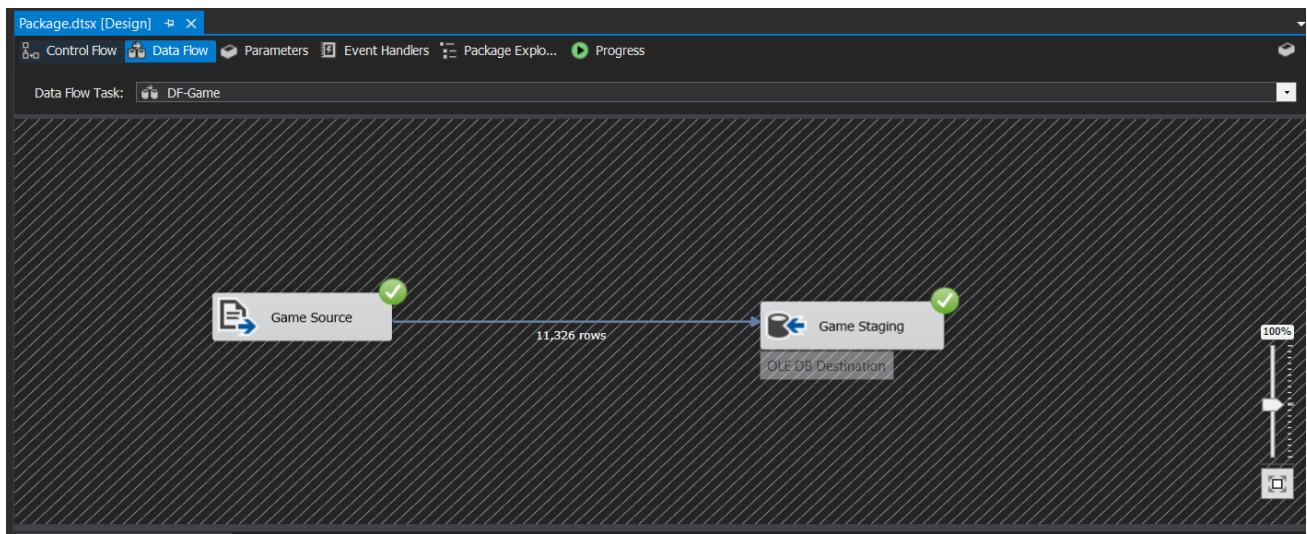
4.2.2. Kết quả thực thi



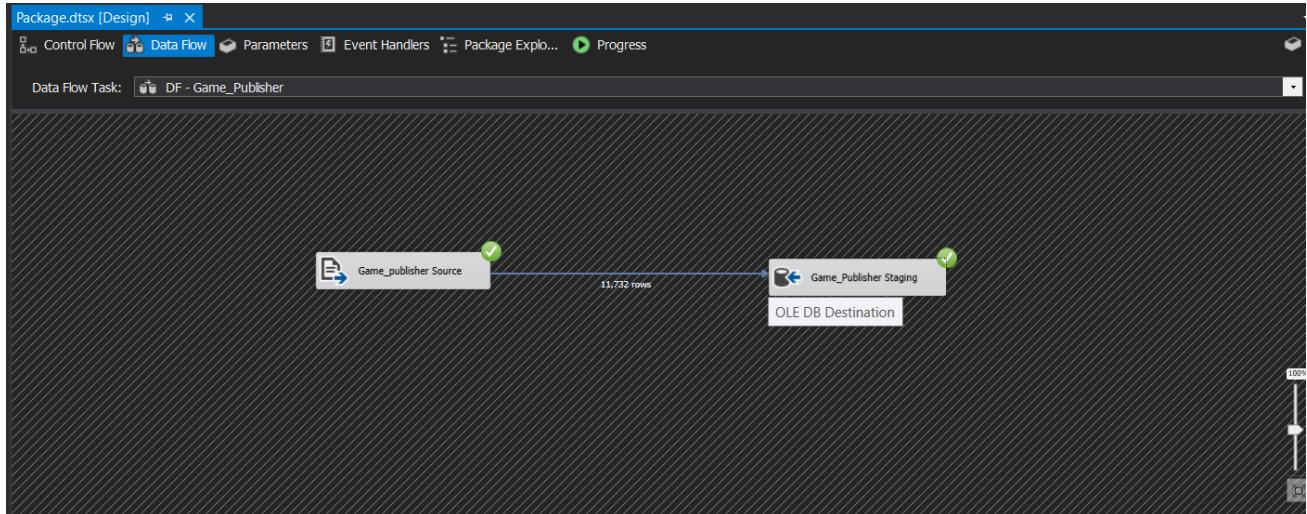
Data Flow Task



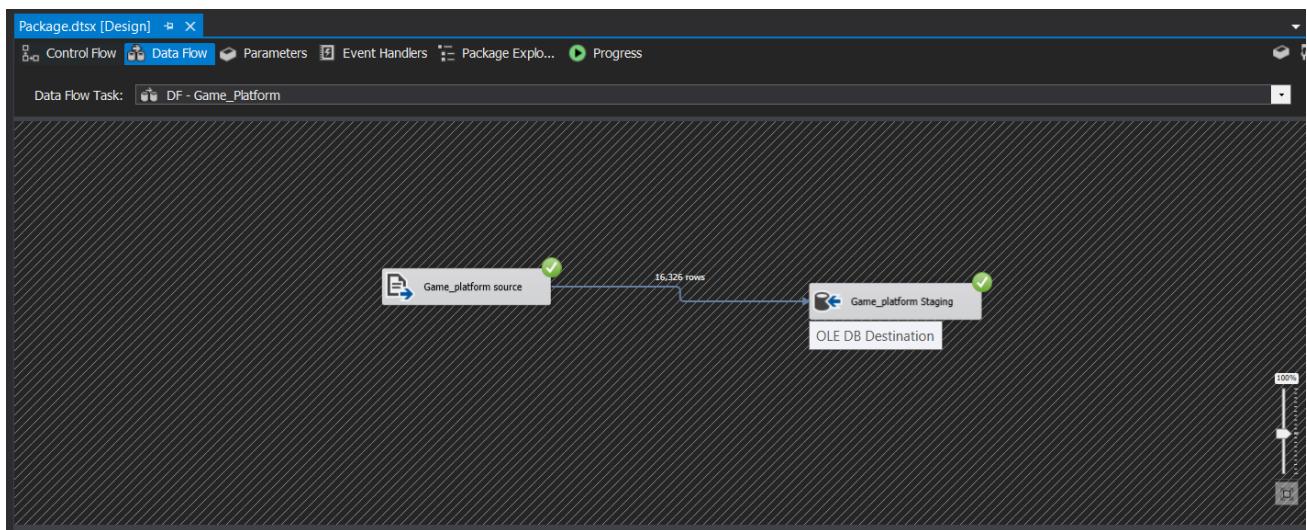
DF- Game



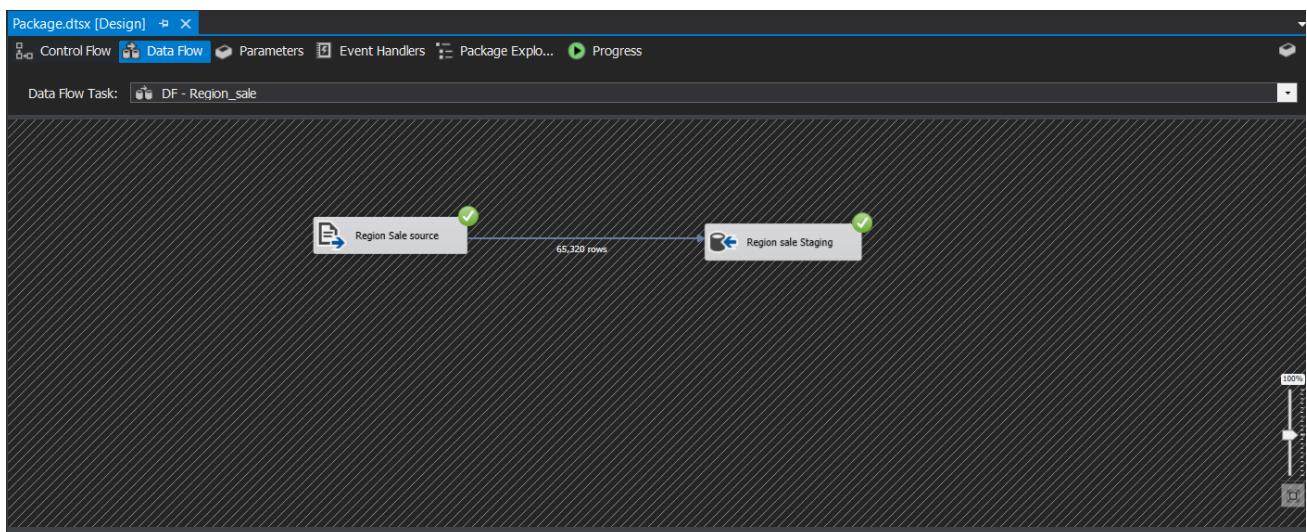
DF – Game_publisher



DF – Game_platform



DF – Region sale



4.2.3. Kết quả trong SQL Server

Bảng Year:

LAPTOP-7MJVALJE.v...video_games.year		
	id	release_year
▶	1	1980
	2	1981
	3	1982
	4	1983
	5	1984
	6	1985
	7	1986
	8	1987
	9	1988
	10	1989
	11	1990
	12	1991
	13	1992
	14	1993
	15	1994
	16	1995
	17	1996
	18	1997
	19	1998
	20	1999
	21	2000
	22	2001
	23	2002
	24	2003
	25	2004
	26	2005
	27	2006
	28	2007
	29	2008

Bảng Region

LAPTOP-7MJVALJE...ideo_games.region ➔ X

	id	region_na...
▶	1	North Amer...
	2	Europe
	3	Japan
	4	Other
*	NULL	NULL

Bảng Genre

LAPTOP-7MJVALJE...video_games.genre ➔ X

	id	genre_name
▶	1	Action
	2	Adventure
	3	Fighting
	4	Misc
	5	Platform
	6	Puzzle
	7	Racing
	8	Role-Playing
	9	Shooter
	10	Simulation
	11	Sports
	12	Strategy
*	NULL	NULL

Bảng Publisher

LAPTOP-7MJVALJE.v...o_games.publisher

	id	publisher_n...
▶	1	10TACLE St...
	2	1C Company
	3	20th Centur...
	4	2D Boy
	5	3DO
	6	49Games
	7	505 Games
	8	5pb
	9	7G//AMES
	10	989 Sports
	11	989 Studios
	12	Abylight
	13	Acclaim Ent...
	14	Accolade
	15	Ackkstudios
	16	Acquire
	17	Activision
	18	Activision Bl...
	19	Activision V...
	20	Adeline Sof...
	21	Aerosoft
	22	Agatsuma E...
	23	Agetec
	24	Aksys Games
	25	Alawar Ente...
	26	Alchemist
	27	Alternative ...
	28	Altron
	29	Alvion
	30	Amplitude S...

Bảng Platform

	id	platform_n...
▶	1	Wii
	2	NES
	3	GB
	4	DS
	5	X360
	6	PS3
	7	PS2
	8	SNES
	9	GBA
	10	3DS
	11	PS4
	12	N64
	13	PS
	14	XB
	15	PC
	16	2600
	17	PSP
	18	XOne
	19	GC
	20	WiiU
	21	GEN
	22	DC
	23	PSV
	24	SAT
	25	SCD
	26	WS
	27	NG
	28	TG16
	29	3DO

Bảng Game

	id	genre_id	game_name
▶	1	11	'98 Koshien
	2	8	.hack//G.U. ...
	3	8	.hack//G.U. ...
	4	8	.hack//G.U. ...
	5	8	.hack//G.U. ...
	6	8	.hack//Infec...
	7	8	.hack//Link
	8	8	.hack//Muta...
	9	8	.hack//Outb...
	10	8	.hack//Quar...
	11	1	.hack: Sekai ...
	12	7	007 Racing
	13	1	007: Quantu...
	14	1	007: The Wo...
	15	9	007: Tomorr...
	16	4	1 vs. 100
	17	2	1/2 Summer...
	18	11	10 Minute S...
	19	6	100 All-Tim...
	20	4	100 Classic ...
	21	4	100 Classic ...
	22	4	1000 Cooki...
	23	1	1001 Touch ...
	24	6	101-in-1 Ex...
	25	4	101-in-1 Par...
	26	11	101-in-1 Sp...
	27	11	101-in-1 Sp...
	28	11	1080°: TenEi...
	29	2	11eyes: Cros...

Bảng Game_Publisher

	id	game_id	publisher_id
▶	1	10866	369
	2	9244	369
	3	5464	369
	4	10868	369
	5	7282	369
	6	9603	369
	7	6573	369
	8	10864	369
	9	6576	369
	10	2556	369
	11	6729	369
	12	5463	369
	13	7265	369
	14	10858	369
	15	10859	369
	16	4866	328
	17	3661	498
	18	3665	498
	19	9258	369
	20	992	369
	21	7262	369
	22	9251	369
	23	9246	369
	24	3666	498
	25	7283	369
	26	7254	369
	27	990	369
	28	3642	464
	29	1199	17
	30	7204	369

Bảng Game_Platform

	id	game_publ...	platform_id	release_year
▶	1	8564	4	2007
	2	9852	4	2007
	3	11063	7	2006
	4	9065	15	2011
	5	9544	15	2011
	6	11623	15	2009
	7	1977	16	1981
	8	3833	16	1981
	9	3838	16	1982
	10	4755	16	1982
	11	5034	16	1981
	12	9762	15	2008
	13	3647	7	2002
	14	3687	7	2003
	15	5747	7	2001
	16	5788	7	2000
	17	6318	7	2001
	18	6544	7	2001
	19	6657	7	2001
	20	7051	7	2002
	21	7172	7	2000
	22	8175	7	2002
	23	8750	7	2001
	24	9032	7	2002
	25	9522	7	2001
	26	1952	12	1999
	27	3341	12	2000
	28	3532	12	1999
	29	3913	12	1998
	--	--	--	--

Bảng Region_Sale

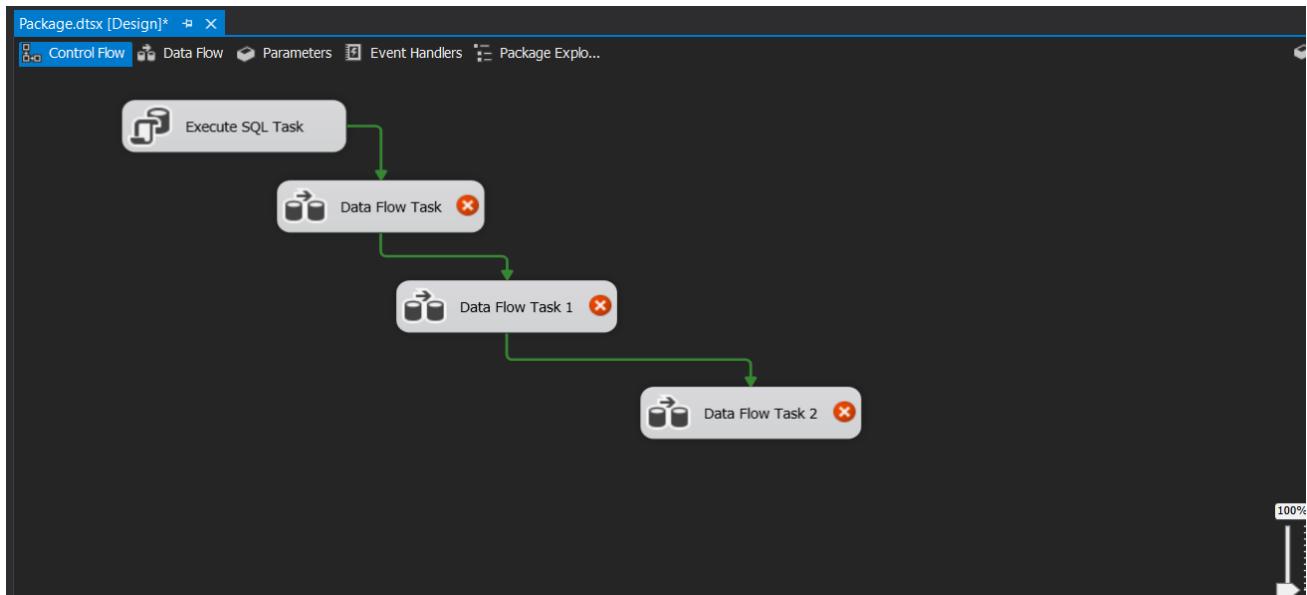
LAPTOP-7MJVALJE.TE... - dbo.region_sale

	region_id	game_platf...	num_sales
▶	1	50	3.50
	1	51	1.43
	1	52	0.51
	1	53	0.27
	1	54	0.48
	1	55	0.34
	1	56	0.36
	1	57	0.36
	1	58	0.17
	1	59	0.04
	1	60	0.21
	1	61	0.11
	1	62	0.13
	1	63	0.12
	1	64	0.08
	1	65	0.10
	1	66	0.09
	1	67	0.09
	1	68	0.07
	1	69	0.07
	1	70	0.00
	1	71	0.06
	1	72	0.06
	1	73	0.00
	1	74	0.04
	1	75	0.04
	1	76	0.04
	1	77	0.03
	1	78	0.02
	1	79	0.02

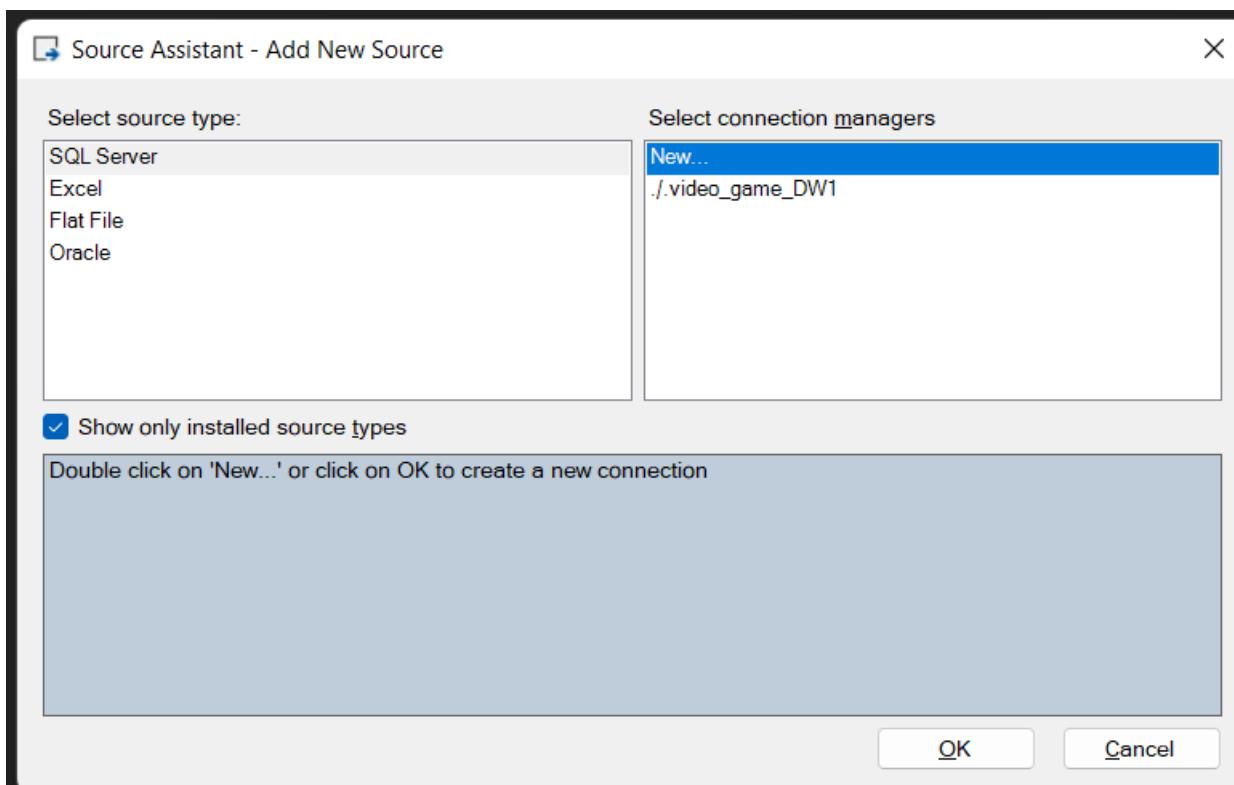
4.3. Quá trình đổ dữ liệu từ Kho Stage sang Kho Ware Hosue

4.3.1. Quá trình tạo Control Flow

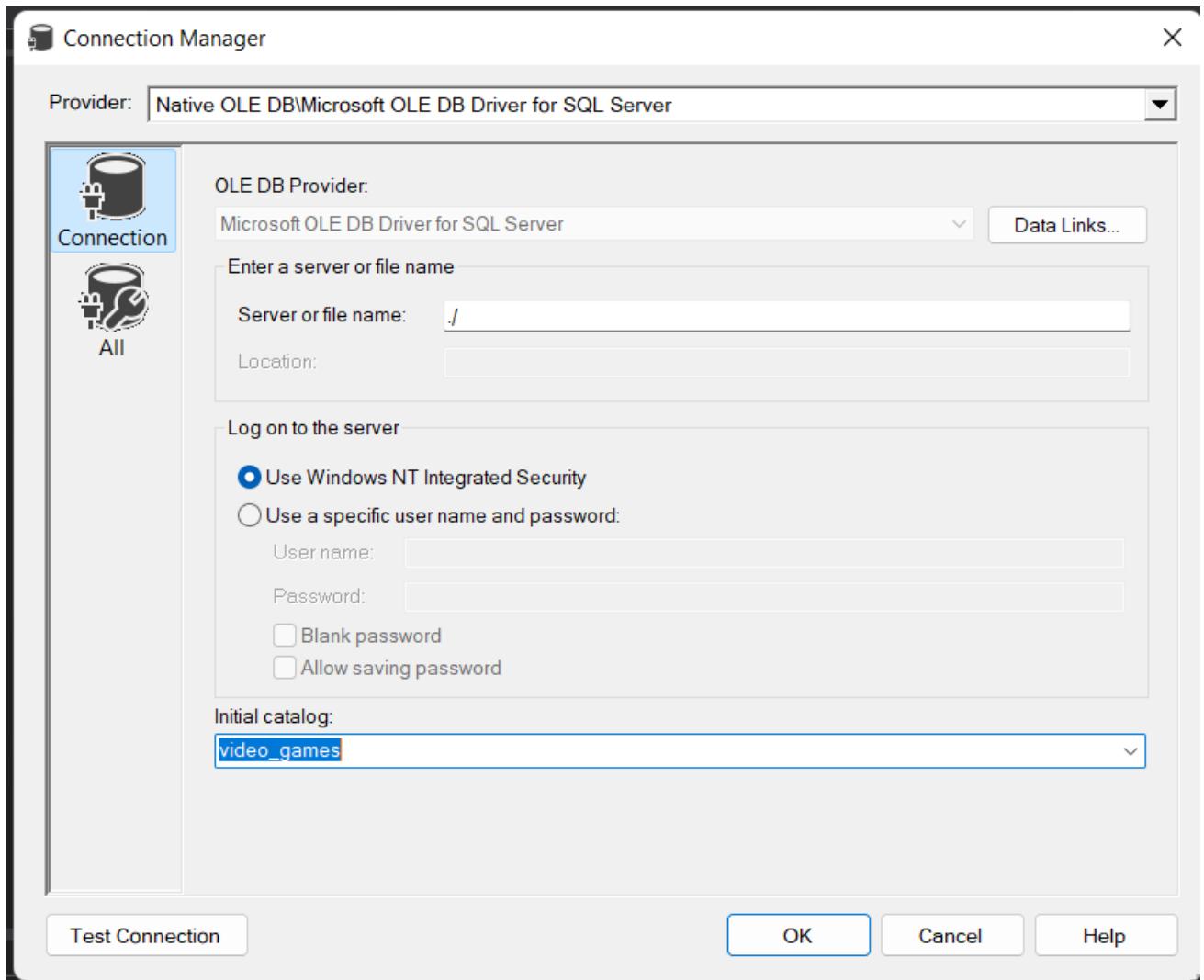
Kéo một Execute SQL task và 3 data flow task vào và setup như sau:



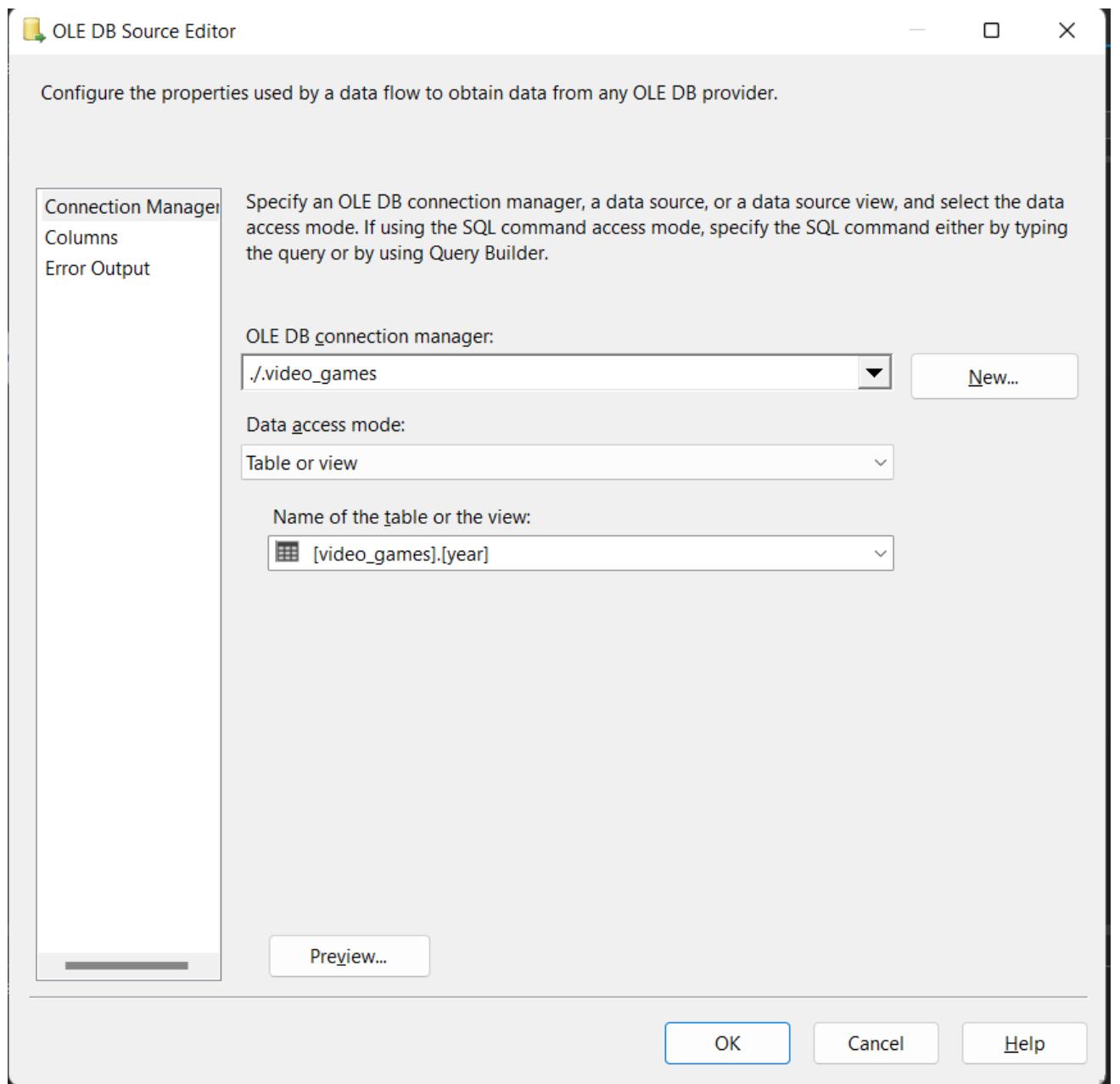
Ở Data Flow Task ta kéo một Source Assistant vào sau đó chọn SQL Server và bấm New:



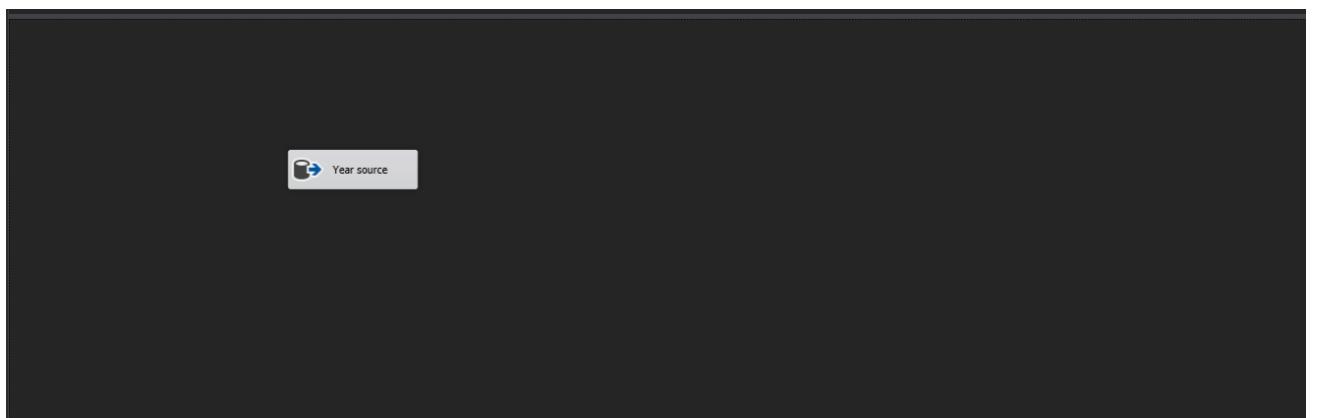
Chọn DB Stage để làm Source



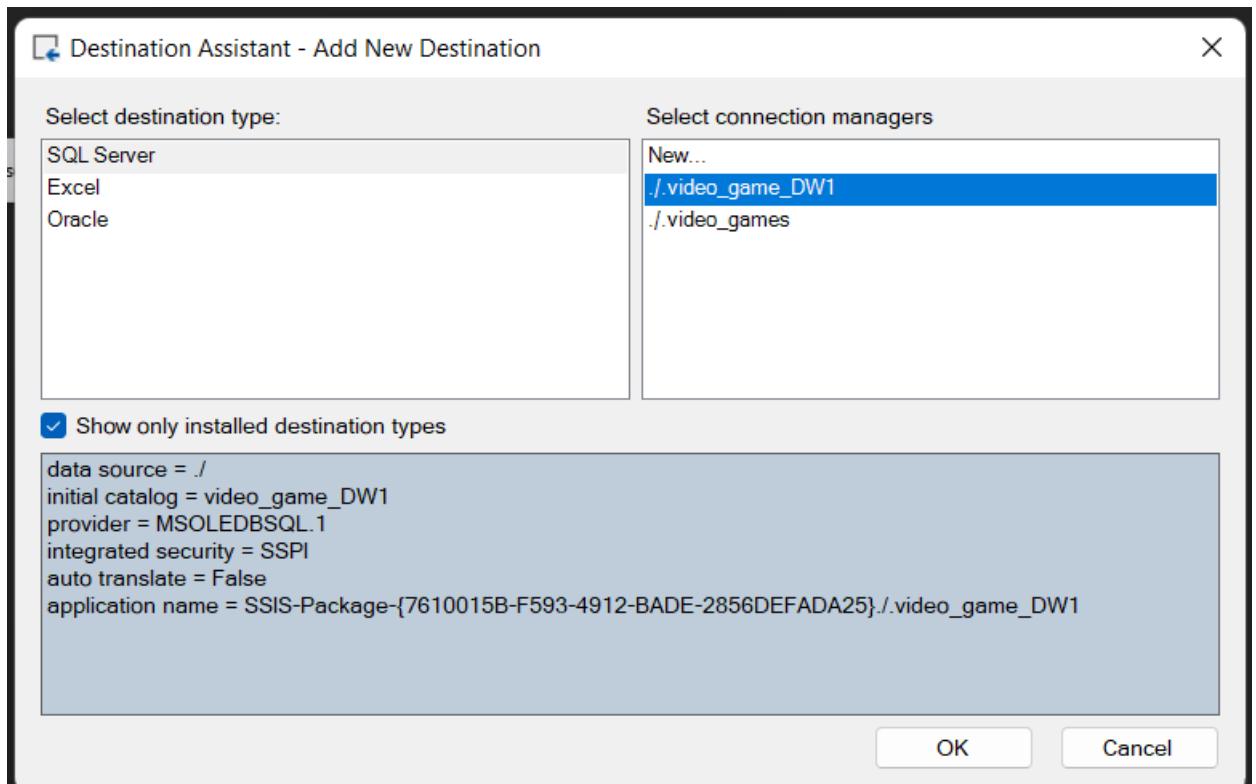
Ở Source Editor chọn bảng year



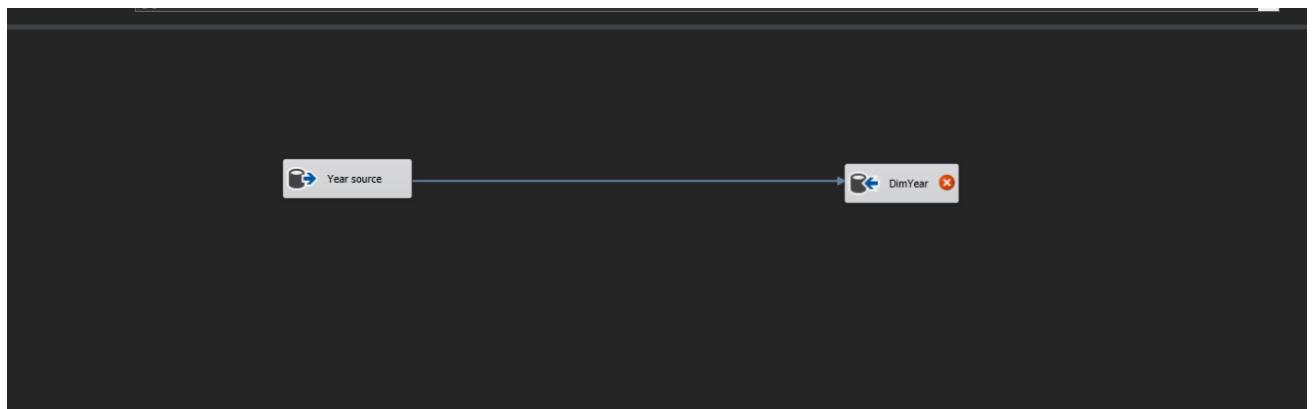
Và đổi tên thành year source



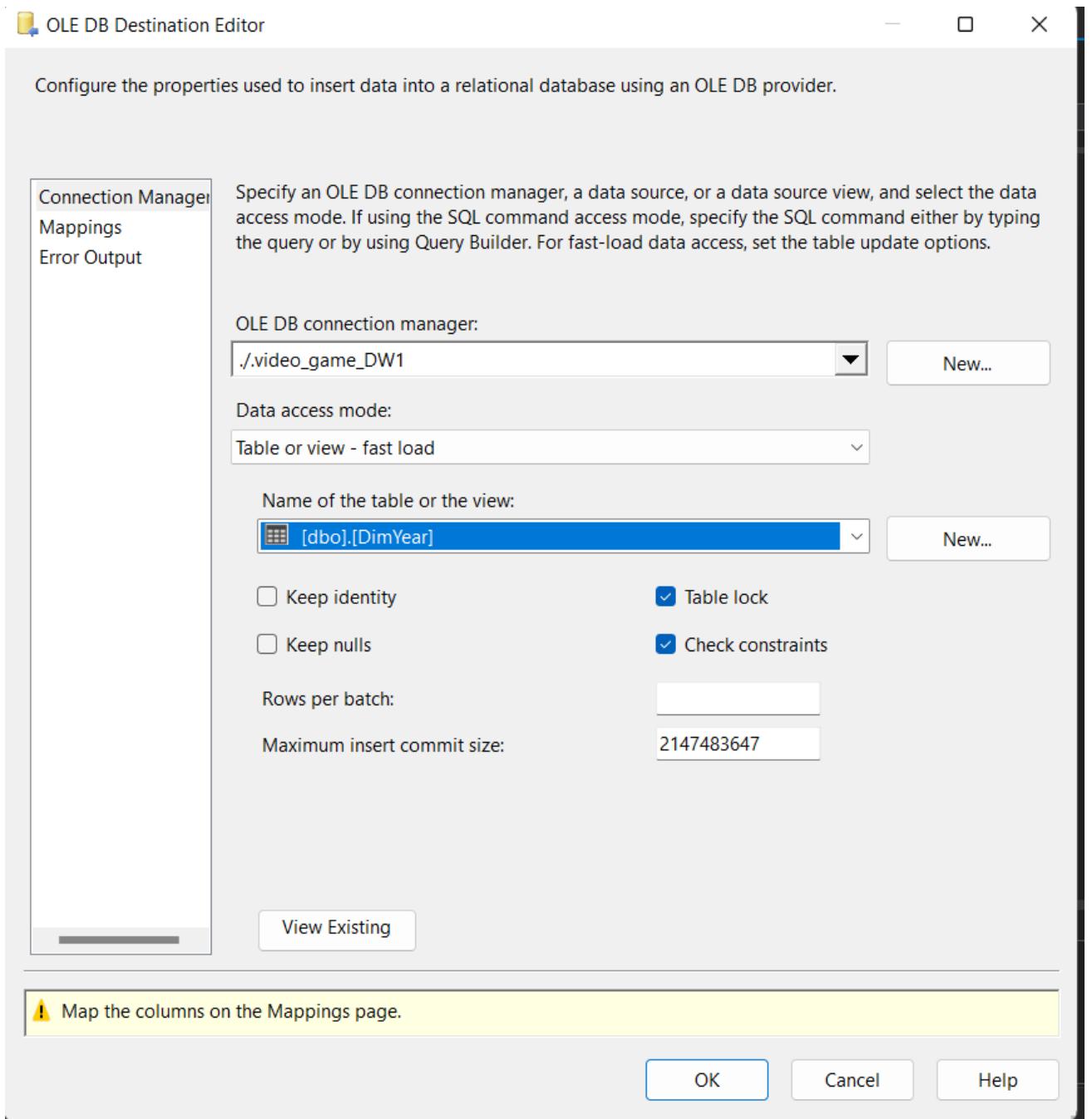
Sau đó kéo một Destination vào và chọn kết nối tới database DWH



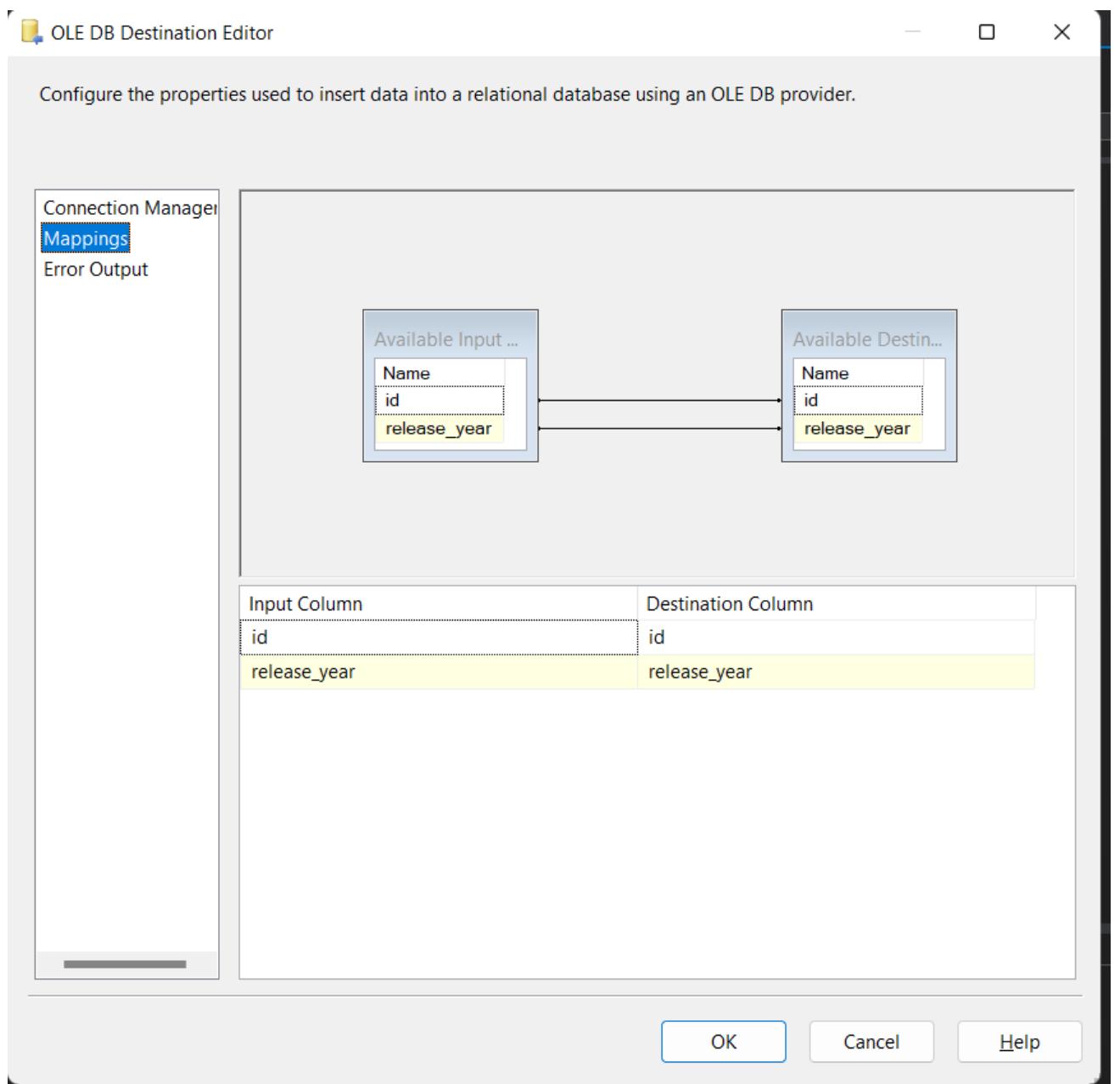
Và đổi tên thành DimYear



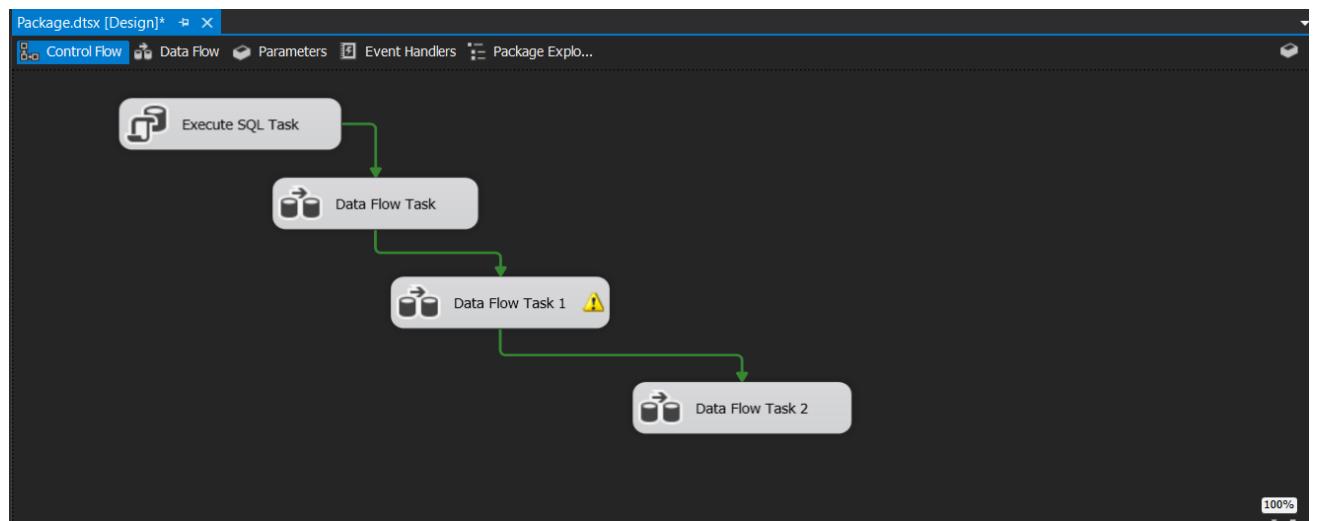
Ở Destination Editor chọn bảng DimYear



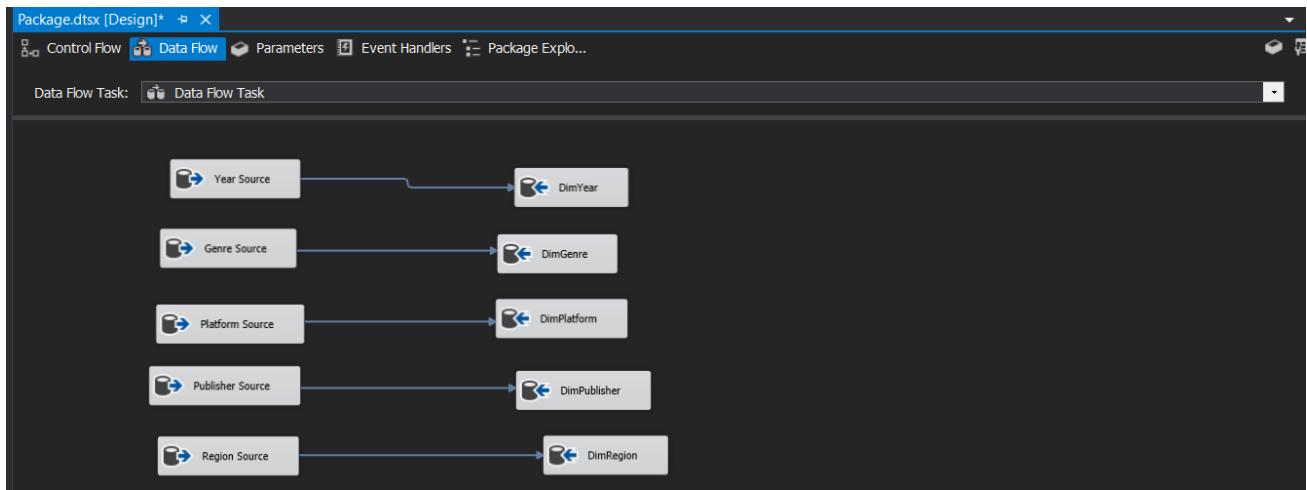
Sau đó Mapping cho phù hợp



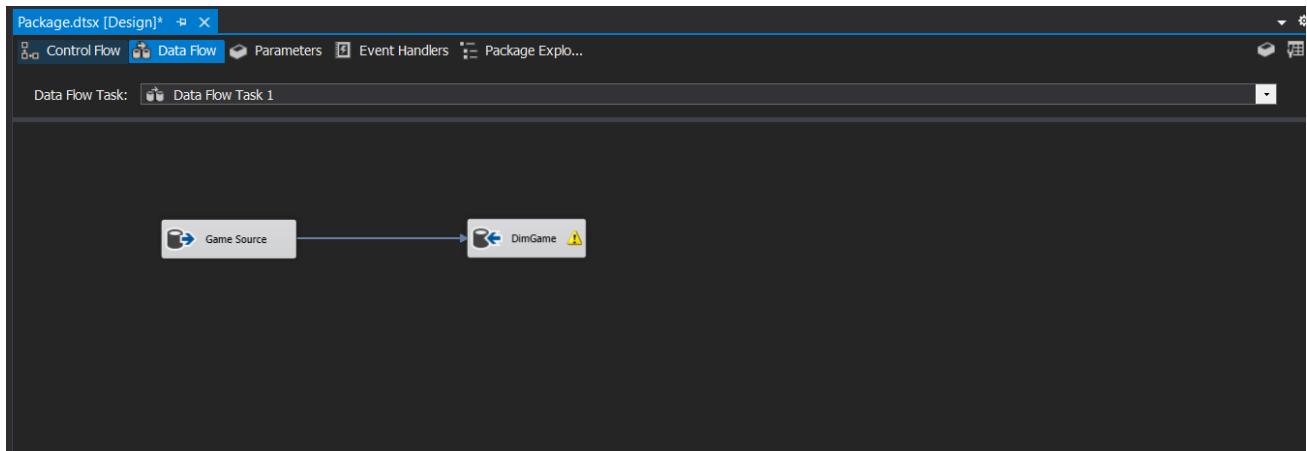
Làm tương tự như vậy ta được như sau



Data Flow Task

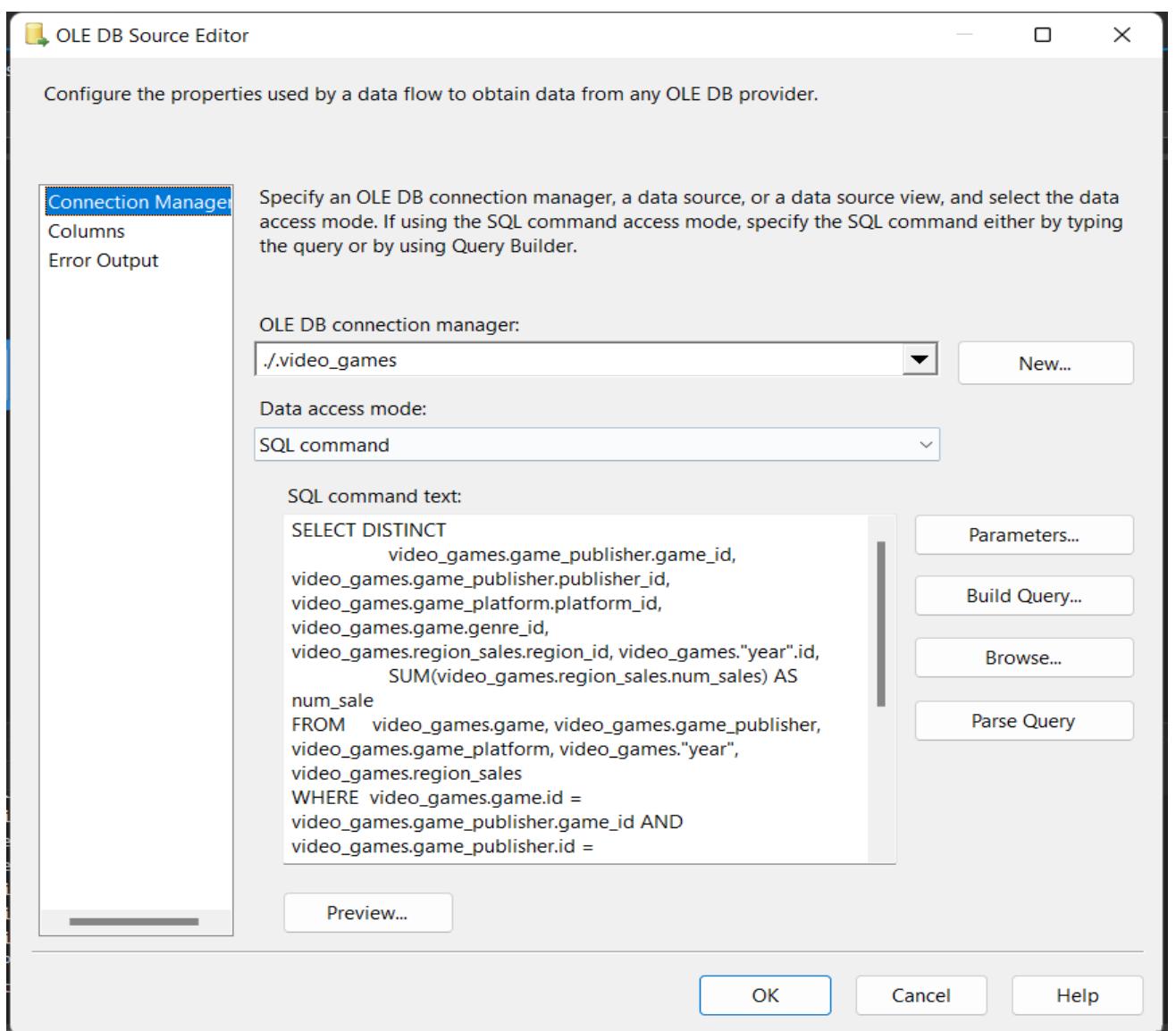
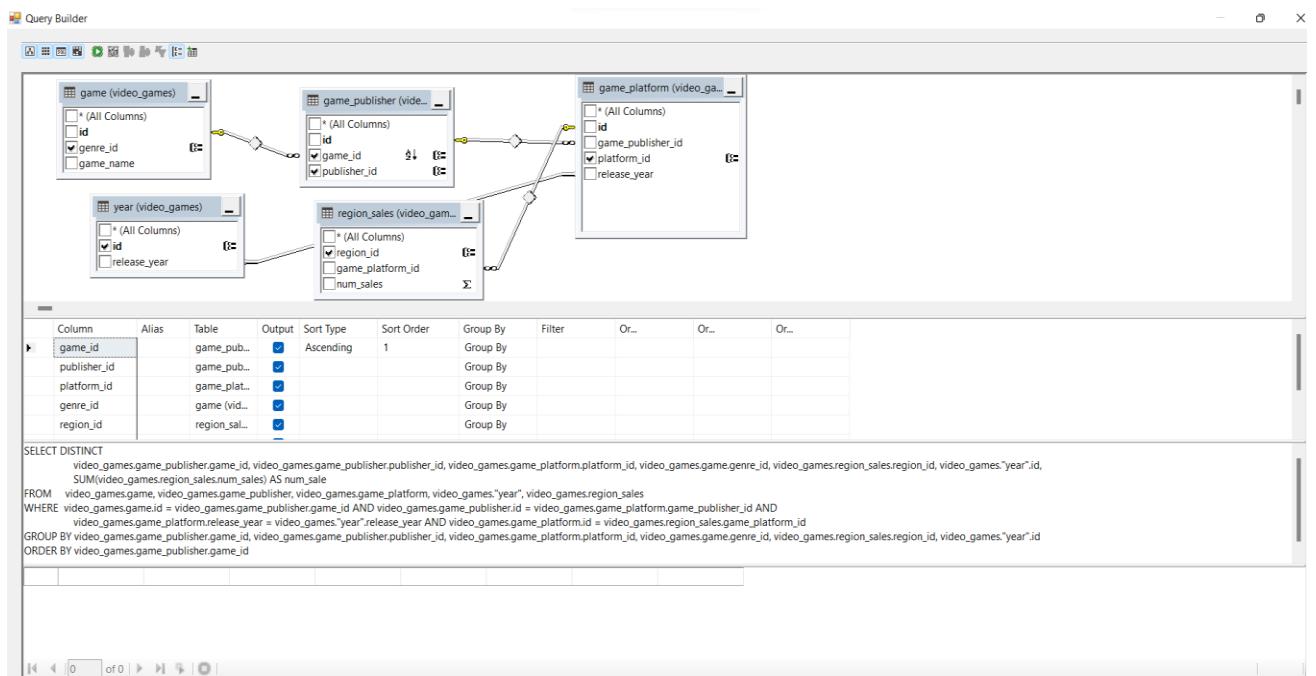


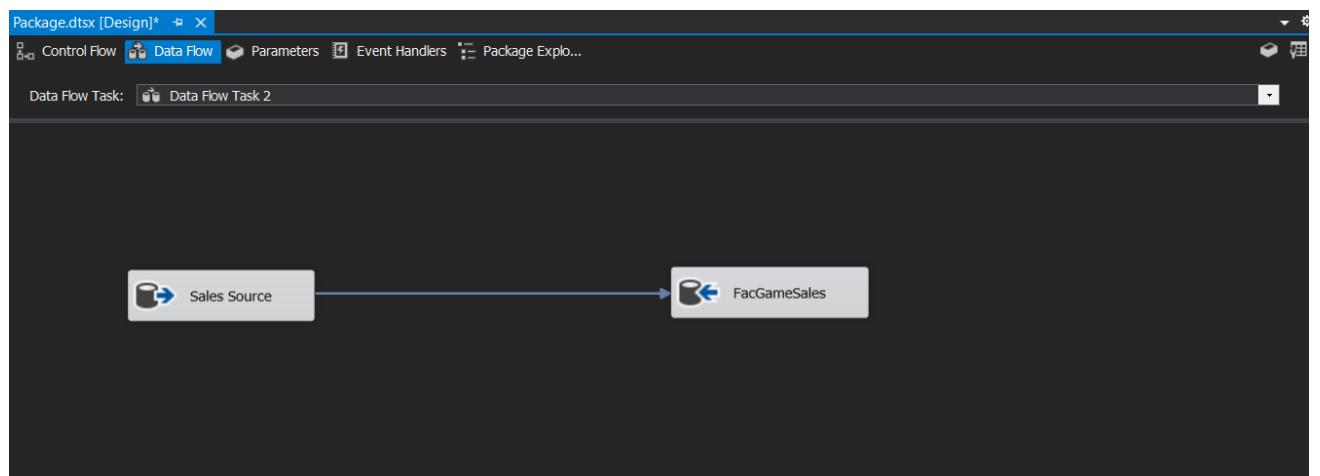
Data Flow Task 1



Data Flow Task 2:

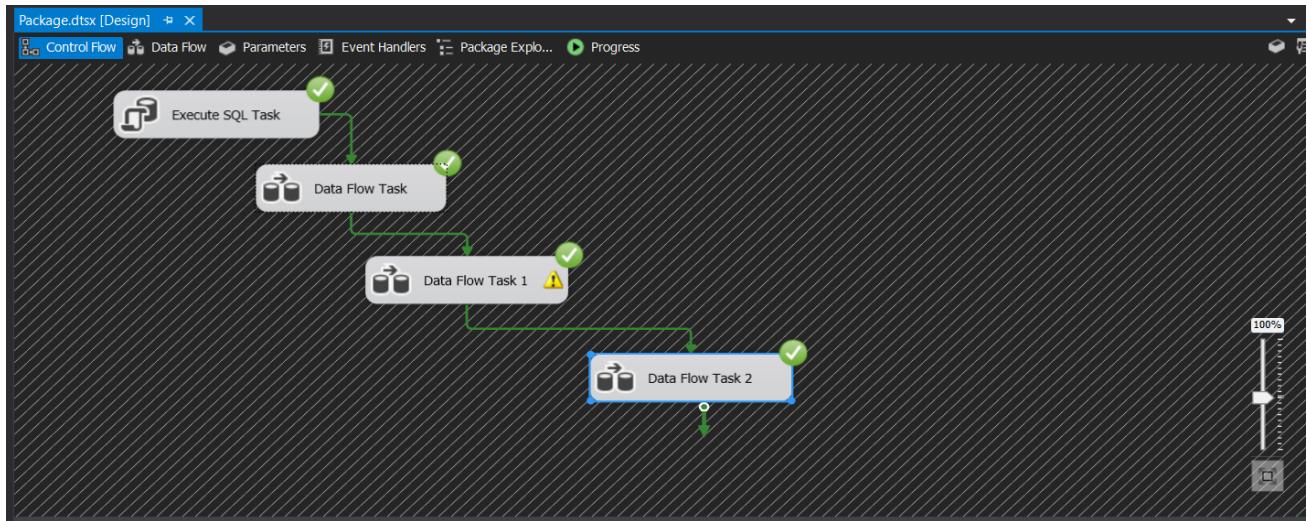
Ở đây Sales Source phải cần dùng đến Sql Command để lấy dữ liệu từ các bảng cần thiết



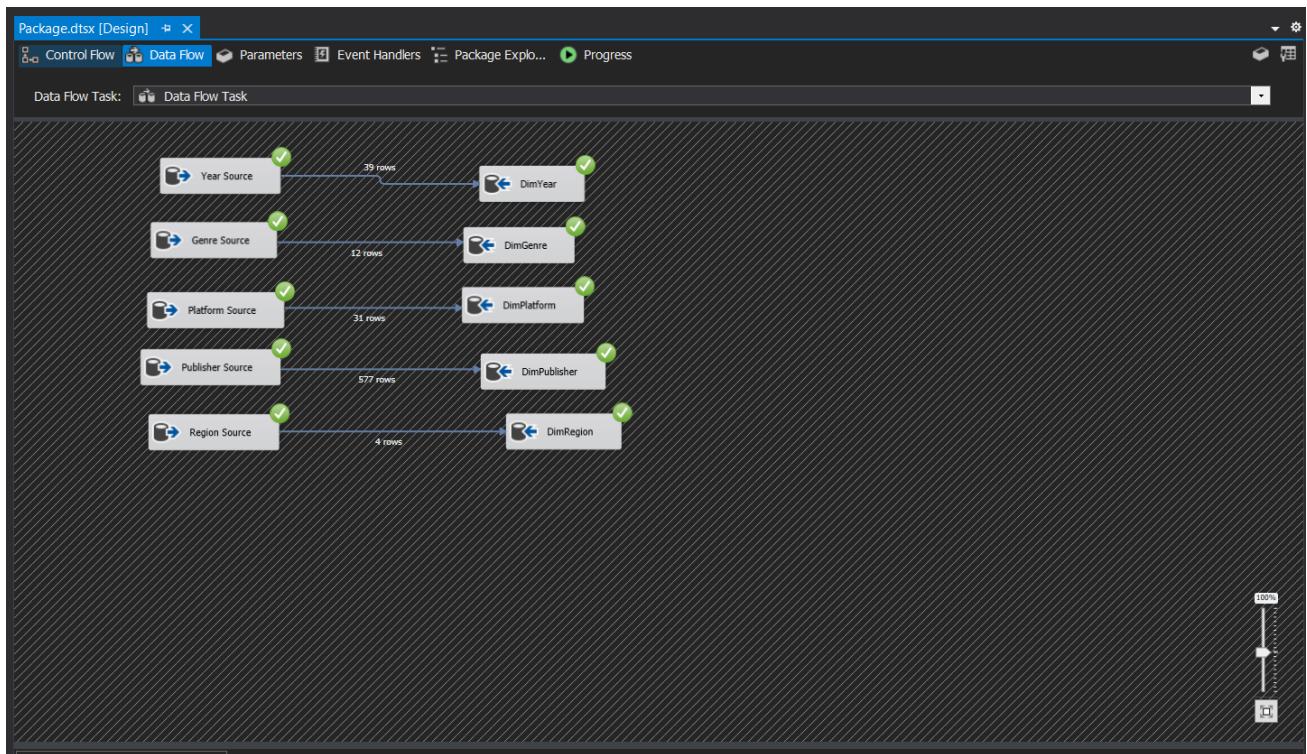


4.3.2. Kết quả thực thi

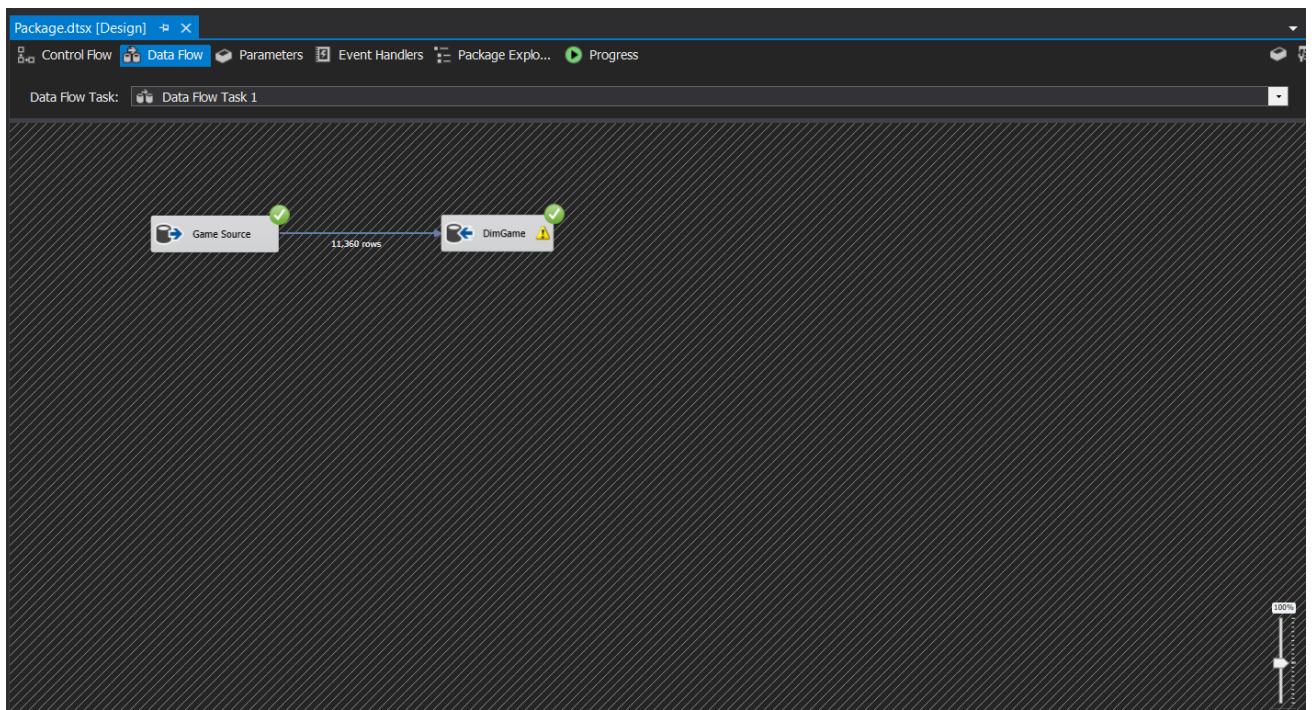
Sau khi chạy SSIS project ta được:



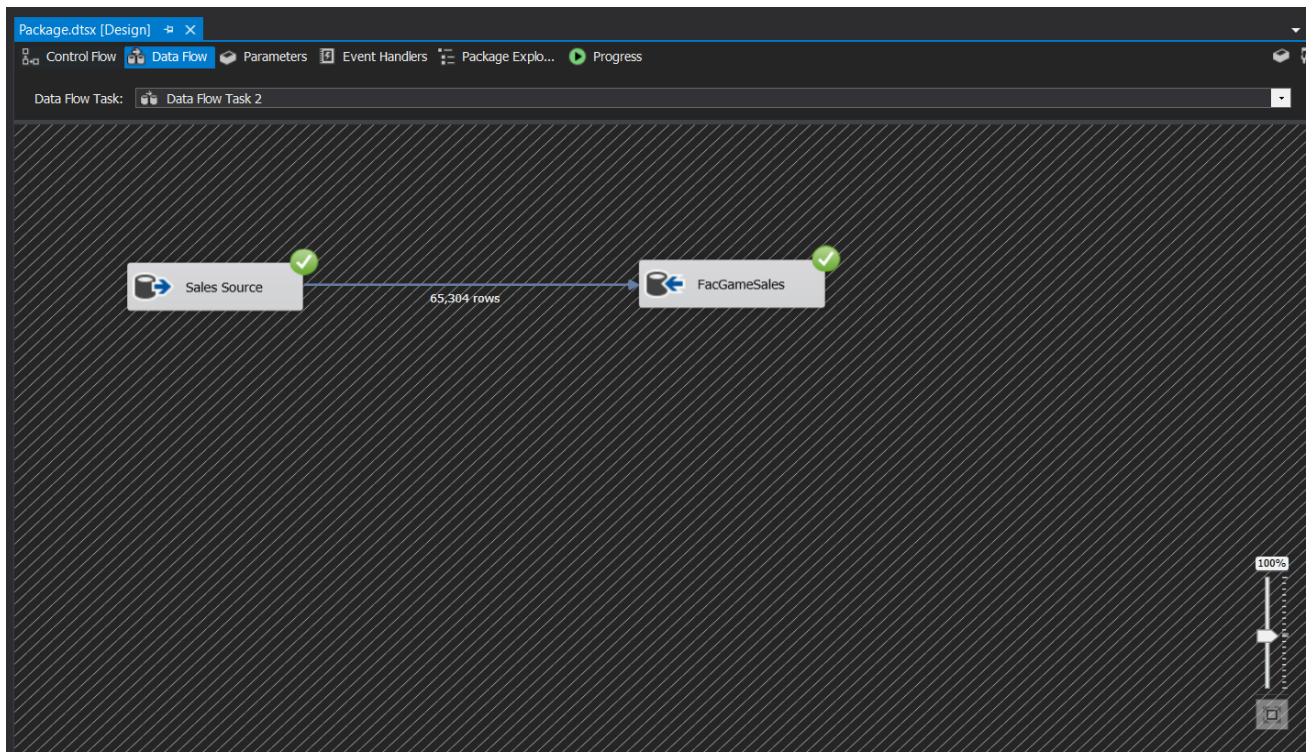
Data Flow Task



Data Flow Task 1



Data Flow Task 2



4.3.3. Kết quả trên SQL Server

DimYear

	id	release_year
▶	1	1980
	2	1981
	3	1982
	4	1983
	5	1984
	6	1985
	7	1986
	8	1987
	9	1988
	10	1989
	11	1990
	12	1991
	13	1992
	14	1993
	15	1994
	16	1995
	17	1996
	18	1997
	19	1998
	20	1999
	21	2000
	22	2001
	23	2002
	24	2003
	25	2004
	26	2005
	27	2006
	28	2007
	29	2008
	30	2009

Dim Region

LAPTOP-7MJVALJE....- dbo.DimRegion

	id	region_na...
▶	1	North Amer...
	2	Europe
	3	Japan
	4	Other
*	NULL	NULL

DimPublisher

LAPTOP-7MJVALJE.v...dbo.DimPublisher

	id	publisher_n...
▶	1	10TACLE St...
	2	1C Company
	3	20th Centur...
	4	2D Boy
	5	3DO
	6	49Games
	7	505 Games
	8	5pb
	9	7G//AMES
	10	989 Sports
	11	989 Studios
	12	Abylight
	13	Acclaim Ent...
	14	Accolade
	15	Ackkstudios
	16	Acquire
	17	Activision
	18	Activision Bl...
	19	Activision V...
	20	Adeline Sof...
	21	Aerosoft

DimPlatform

	id	platform_n...
1	1	Wii
2	2	NES
3	3	GB
4	4	DS
5	5	X360
6	6	PS3
7	7	PS2
8	8	SNES
9	9	GBA
10	10	3DS
11	11	PS4
12	12	N64
13	13	PS
14	14	XB
15	15	PC
16	16	2600
17	17	PSP
18	18	XOne
19	19	GC
20	20	WiiU
21	21	GEN
22	22	DC
23	23	PSV
24	24	SAT
25	25	SCD
26	26	WS
27	27	NG
28	28	TG16
29	29	3DO
30	30	CC

DimGenre

LAPTOP-7MJVALJE....W - dbo.DimGenre		
	id	genre_name
▶	1	Action
	2	Adventure
	3	Fighting
	4	Misc
	5	Platform
	6	Puzzle
	7	Racing
	8	Role-Playing
	9	Shooter
	10	Simulation
	11	Sports
	12	Strategy
*	NULL	NULL

DimGame

LAPTOP-7MJVALJE....W - dbo.DimGenre LAPTO			
	id	genre_id	game_name
▶	1	11	'98 Koshien
	2	8	.hack//G.U. ...
	3	8	.hack//G.U. ...
	4	8	.hack//G.U. ...
	5	8	.hack//G.U. ...
	6	8	.hack//Infec...
	7	8	.hack//Link
	8	8	.hack//Muta...
	9	8	.hack//Outb...
	10	8	.hack//Quar...
	11	1	.hack: Sekai ...
	12	7	007 Racing
	13	1	007: Quantu...
	14	1	007: The Wo...
	15	9	007: Tomorr...
	16	4	1 vs. 100
	17	2	1/2 Summer...
	18	11	10 Minute S...
	19	6	100 All-Tim...
	20	4	100 Classic ...
	21	4	100 Classic ...
	22	4	1000 Cooki...

FactGameSale

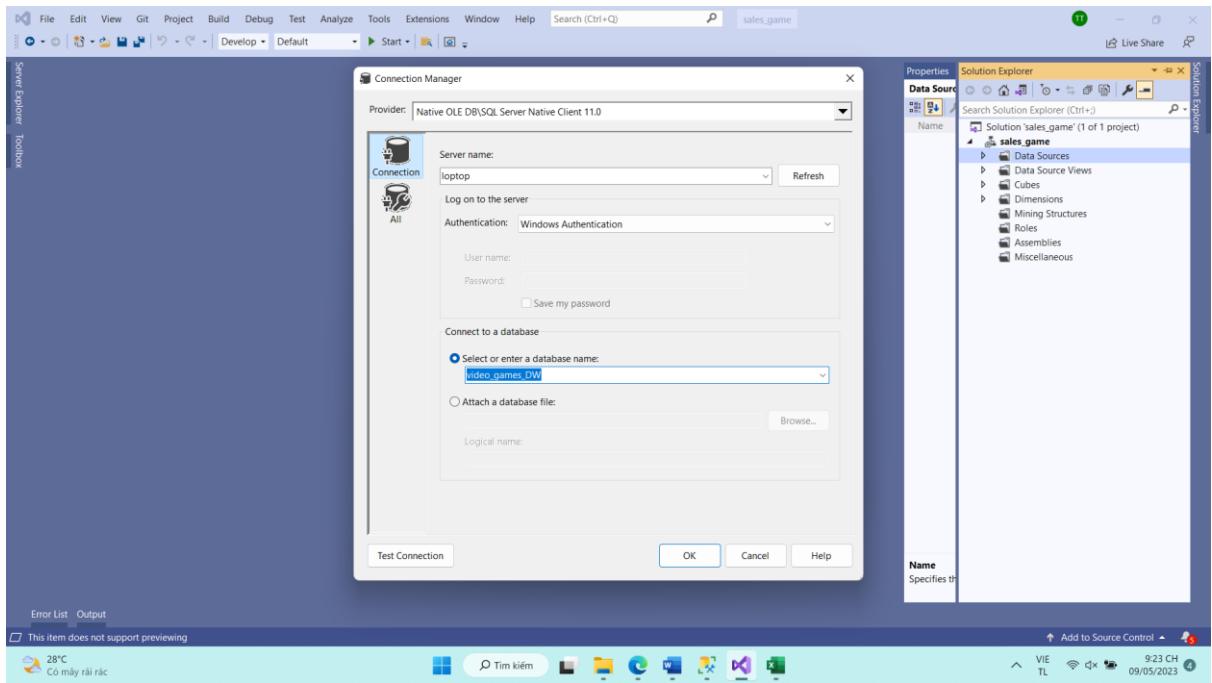
	game_id	publisher_id	platform_id	genre_id	region_id	year_id	num_sales
▶	1	298	13	11	1	19	0.15
	1	298	13	11	2	19	0.10
	1	298	13	11	3	19	0.12
	1	298	13	11	4	19	0.03
	2	352	7	8	1	27	0.00
	2	352	7	8	2	27	0.00
	2	352	7	8	3	27	0.17
	2	352	7	8	4	27	0.00
	3	352	7	8	1	27	0.11
	3	352	7	8	2	27	0.09
	3	352	7	8	3	27	0.00
	3	352	7	8	4	27	0.03
	4	352	7	8	1	27	0.00
	4	352	7	8	2	27	0.00
	4	352	7	8	3	27	0.16
	4	352	7	8	4	27	0.00
	5	352	7	8	1	28	0.00
	5	352	7	8	2	28	0.00
	5	352	7	8	3	28	0.17
	5	352	7	8	4	28	0.00
	6	54	7	8	1	23	0.49
	6	54	7	8	2	23	0.38
	6	54	7	8	3	23	0.26
	6	54	7	8	4	23	0.13
	7	352	17	8	1	31	0.00
	7	352	17	8	2	31	0.00
	7	352	17	8	3	31	0.14
	7	352	17	8	4	31	0.00
	8	54	7	8	1	23	0.23
	8	54	7	8	2	23	0.19

Chương 5. PHÂN TÍCH DỮ LIỆU(SSAS)

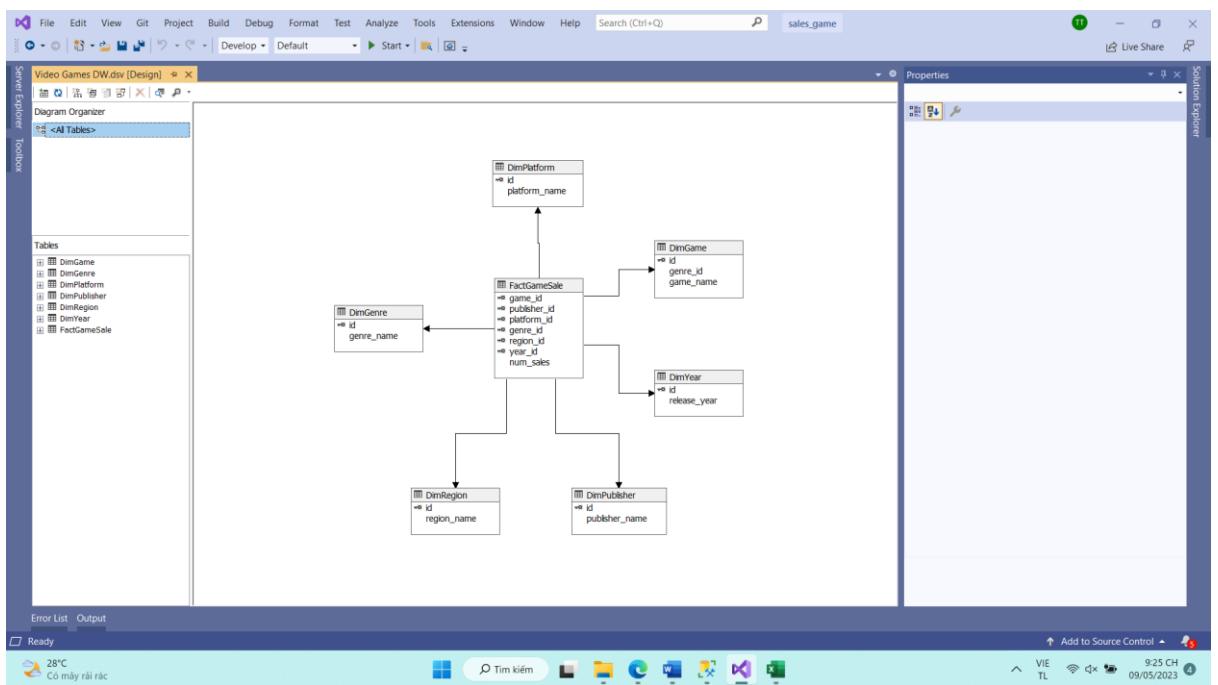
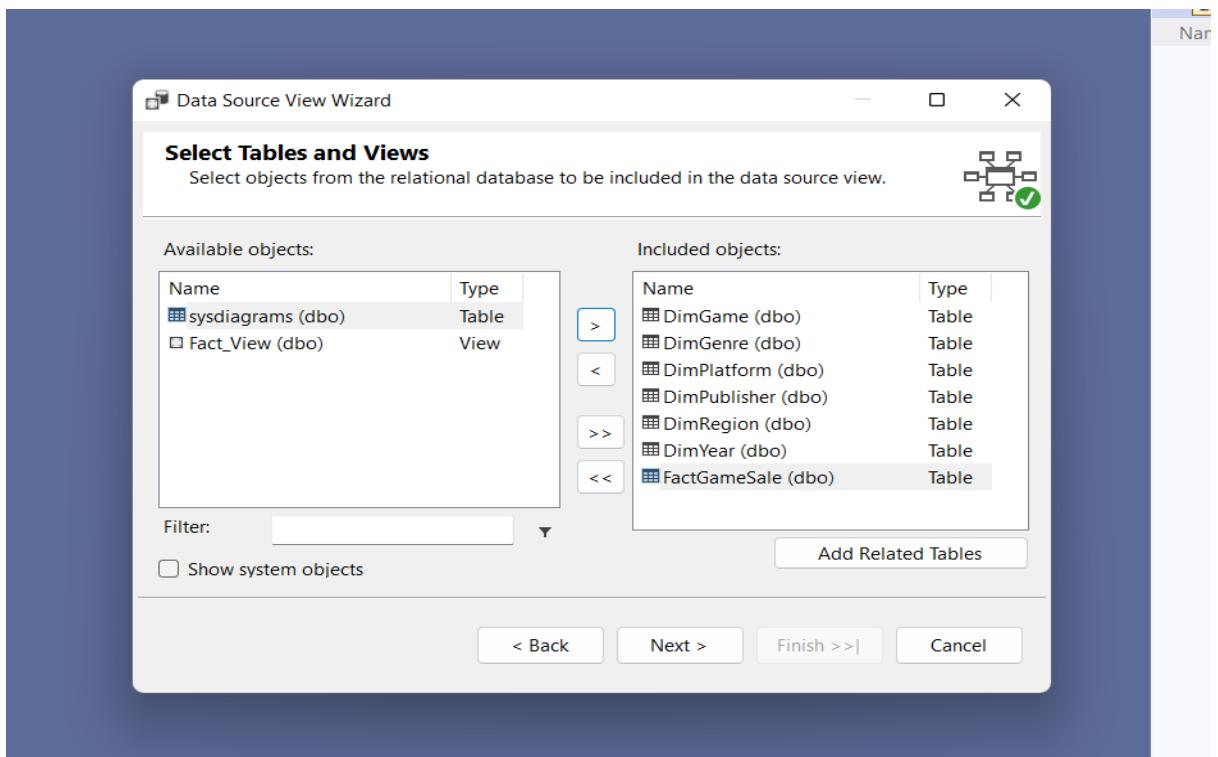
5.1. Quá trình xây dựng mô hình

5.1.1. Tạo SSAS Project

B1: Tạo Data Source



B2: Tạo Data Source View



B4: Tạo Data Cube

Cube Wizard

Select Measure Group Tables

Select a data source view or diagram and then select the tables that will be used for measure groups.

Data source view:
Video Games DW

Measure group tables:

- DimPlatform
- DimPublisher
- DimRegion
- DimGame
- DimGenre
- DimYear
- FactGameSale

Suggest

< Back Next > Finish >>| Cancel

Cube Wizard

Completing the Wizard

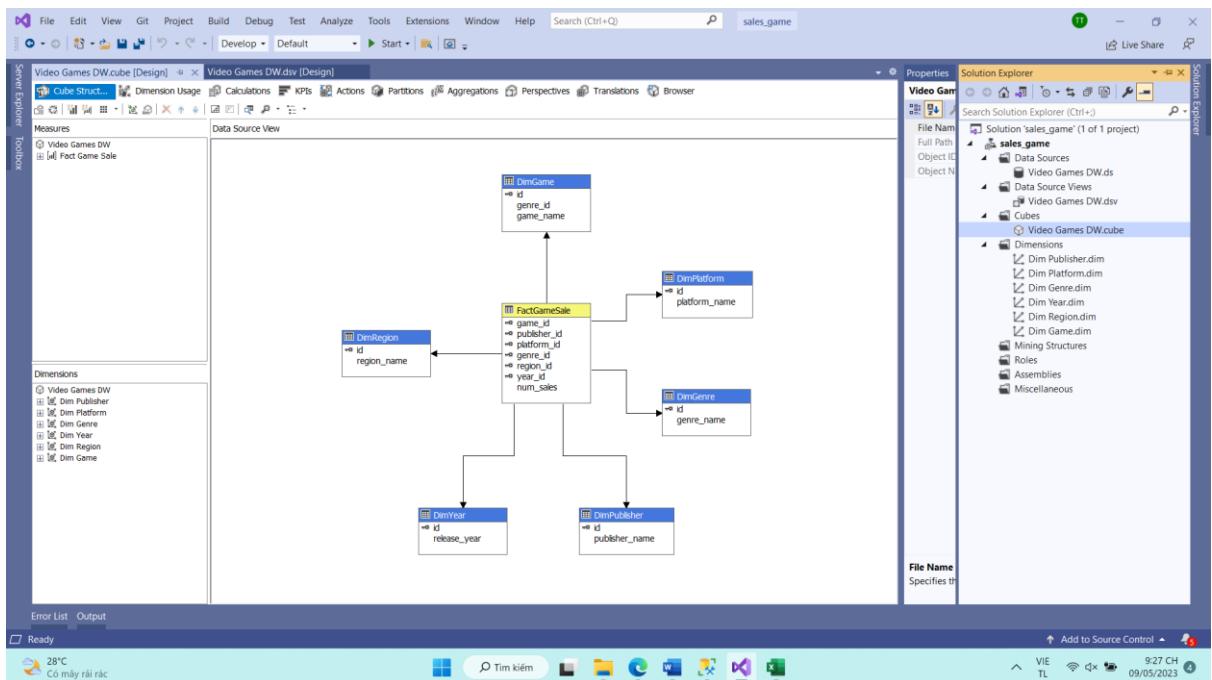
Name the cube, review its structure, and then click Finish to save the cube.

Cube name:
Video Games DW 1

Preview:

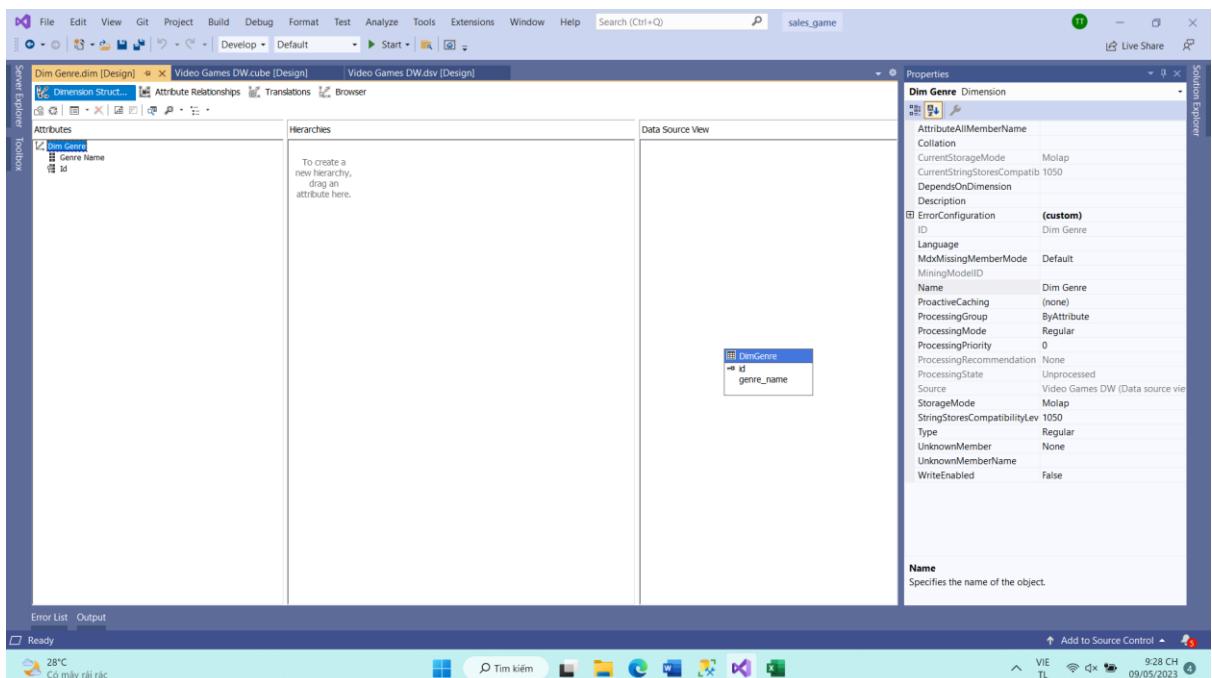
- Measure groups
 - Fact Game Sale
 - Num Sales
 - Fact Game Sale Count
- Dimensions
 - Dim Publisher
 - Dim Platform
 - Dim Genre
 - Dim Year
 - Dim Region
 - Dim Game

< Back Next > **Finish** Cancel

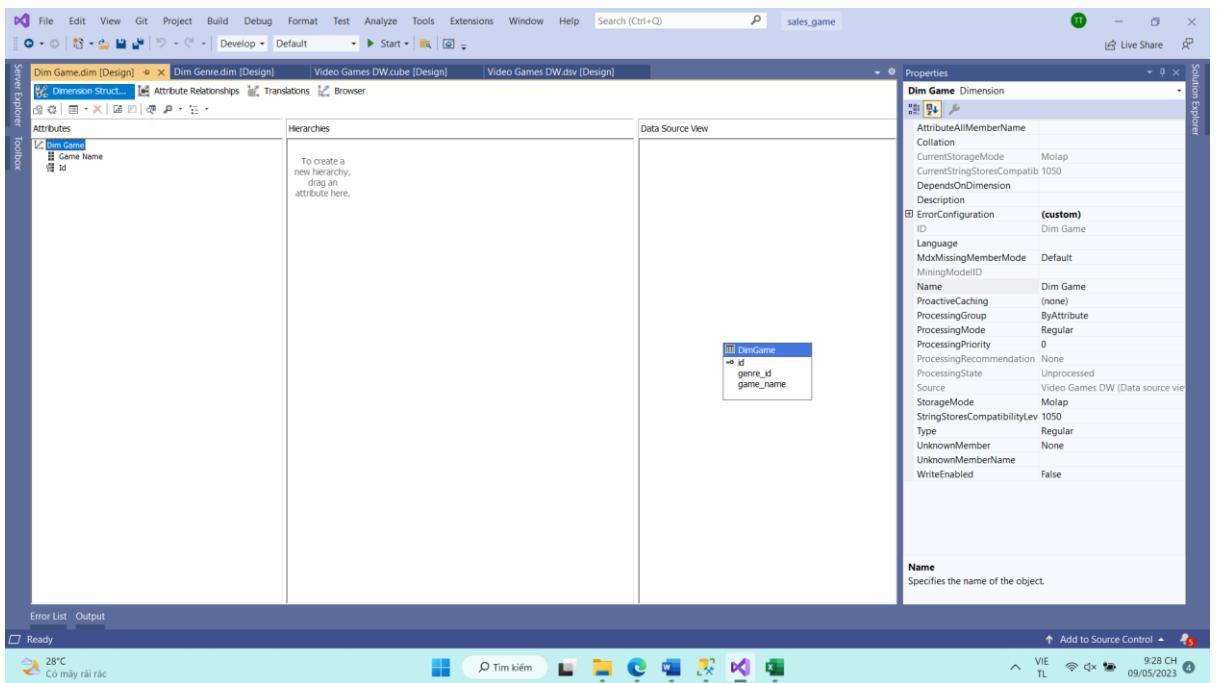


B5: Cấu hình các Dimension

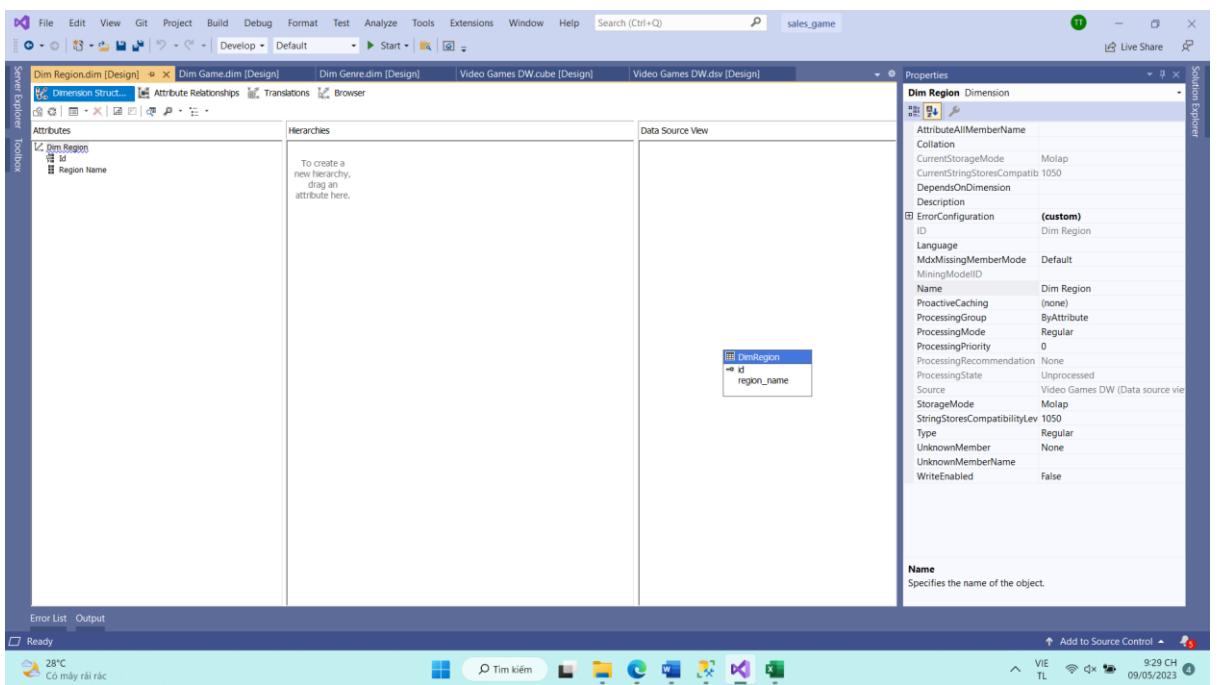
DimGenre



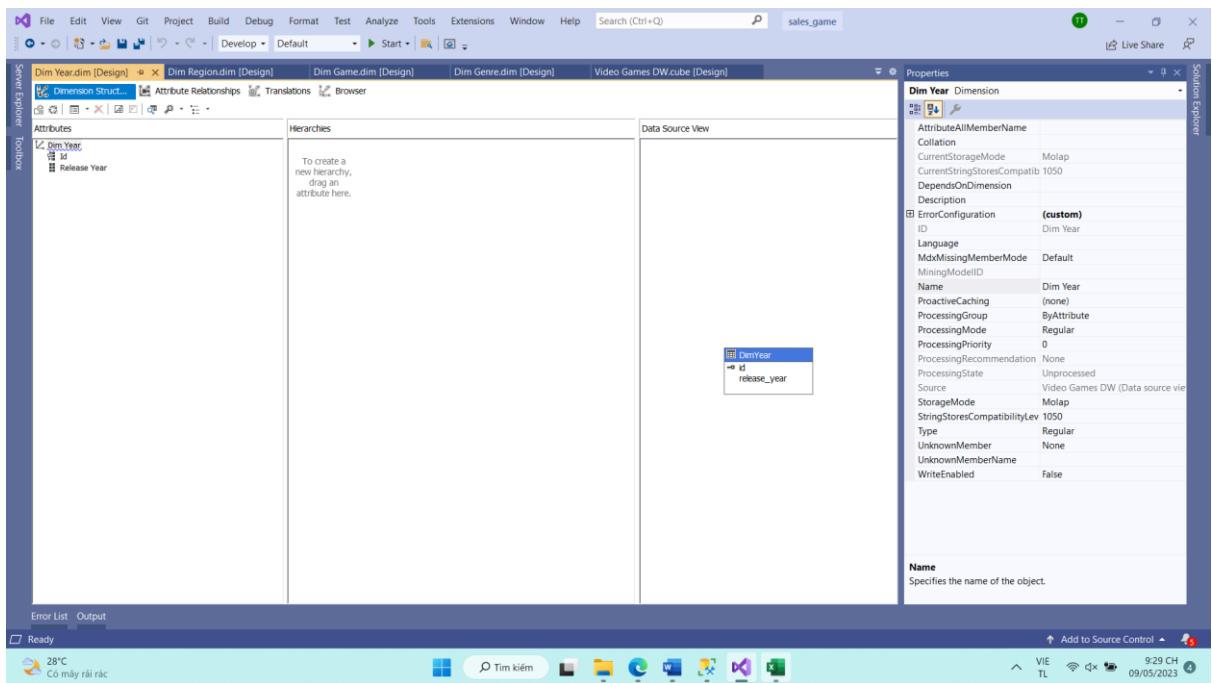
DimGame



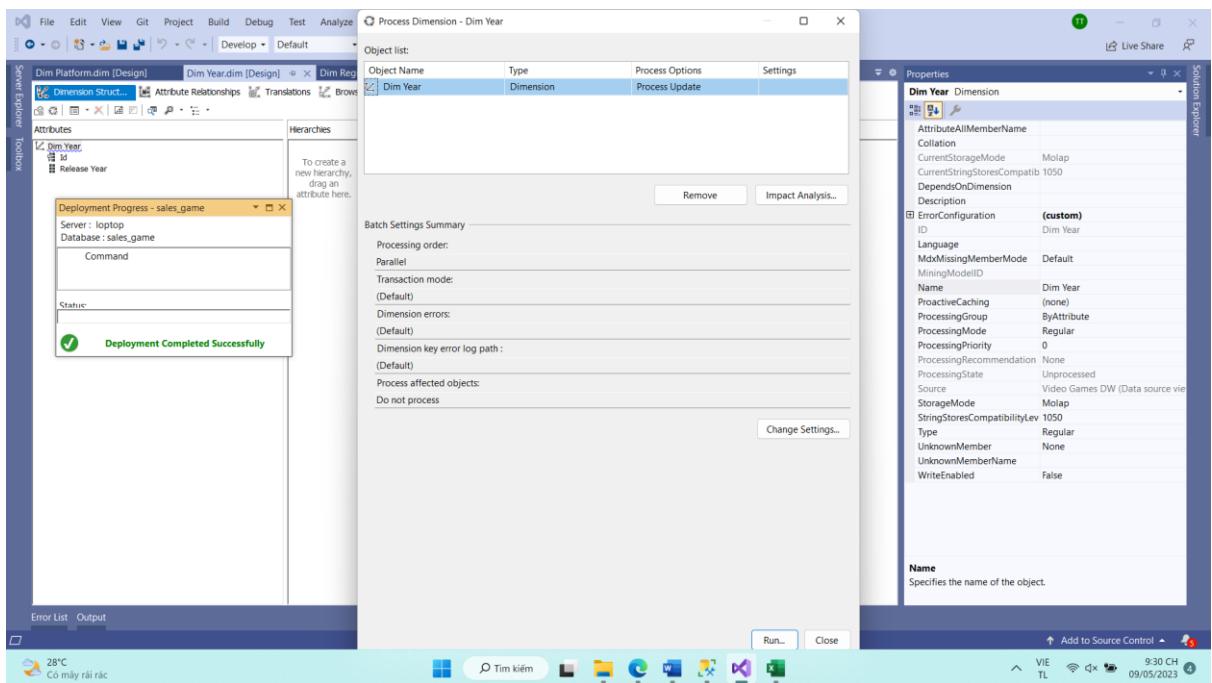
DimRegion

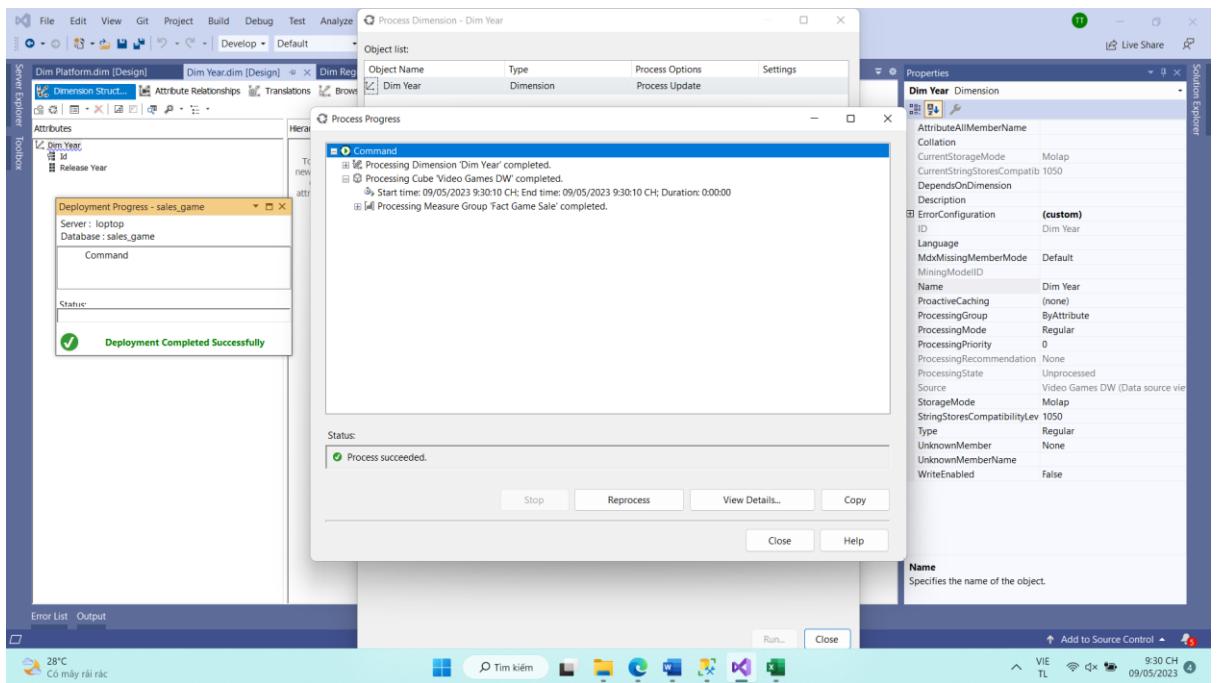


DimYear



B6: Deploy and Process





5.2. Thực hiện các câu truy vấn (SSAS, Pivot Table, T-SQL)

5.2.1. Câu 1: Tổng doanh thu của các thể loại game

-Sử dụng SSAS

The screenshot shows the SSAS Model Explorer interface. On the left, there is a tree view of the model structure under 'Video Games DW'. The 'KPIs' node is expanded, showing 'Num Sales' and 'Dim Game'. The 'Dim Game' node is selected. On the right, a pivot table is displayed with the following data:

Genre Name	Num Sales
Action	1722.05
Adventure	234.37
Fighting	444.08
Misc	796.83
Platform	828.82
Puzzle	241.69
Racing	729.79
Role-Playing	923.75
Shooter	1025.75
Simulation	389.88
Sports	1308.83
Strategy	173.26

-Pivot table trên Excel

	A	B	C
1	Row Labels	Num Sales	
2	Action	1722.05	
3	Adventure	234.37	
4	Fighting	444.08	
5	Misc	796.83	
6	Platform	828.82	
7	Puzzle	241.69	
8	Racing	729.79	
9	Role-Playing	923.75	
10	Shooter	1025.75	
11	Simulation	389.88	
12	Sports	1308.83	
13	Strategy	173.26	
14	Grand Total	8819.1	

-Truy vấn SQL

```
SELECT ge.genre_name, SUM(re.num_sales) AS DoanhSo
FROM     video_games.game AS ga INNER JOIN
          video_games.game_publisher AS pu ON ga.id = pu.game_id INNER JOIN
          video_games.game_platform AS pl ON pu.id = pl.game_publisher_id INNER JOIN
          video_games.genre AS ge ON ga.genre_id = ge.id INNER JOIN
          video_games.region_sales AS re ON pl.id = re.game_platform_id
GROUP BY ge.genre_name
```

100 % ▾

Results Messages

	genre_name	DoanhSo
1	Action	1722.05
2	Adventure	234.37
3	Fighting	444.08
4	Misc	796.83
5	Platform	828.82
6	Puzzle	241.69
7	Racing	729.79
8	Role-Playing	923.75
9	Shooter	1025.75
10	Simulation	389.88
11	Sports	1308.83
12	Strategy	173.26

5.2.2. Câu 2: Cho biết Doanh thu của các thể loại game theo từng năm?

- Sử dụng SSAS

Genre Name	Release Year	Num Sales
Action	1980	0.34
Action	1981	14.79
Action	1982	6.5
Action	1983	2.86
Action	1984	1.85
Action	1985	3.52
Action	1986	13.74
Action	1987	1.11
Action	1988	1.74
Action	1989	4.65
Action	1990	6.39
Action	1991	6.76
Action	1992	3.84
Action	1993	1.81
Action	1994	1.55
Action	1995	3.58
Action	1996	20.6
Action	1997	27.59
Action	1998	39.41
Action	1999	27.73
Action	2000	33.98
Action	2001	59.38
Action	2002	86.65
Action	2003	67.98

-Pivot Table trên excel

-Truy vấn SQL

```
SELECT TOP (100) PERCENT ge.genre_name, pl.release_year, SUM(re.num_sales) AS DoanhSo
FROM      video_games.game AS ga INNER JOIN
                      video_games.game_publisher AS pu ON ga.id = pu.game_id INNER JOIN
                      video_games.game_platform AS pl ON pu.id = pl.game_publisher_id INNER JOIN
                      video_games.genre AS ge ON ga.genre_id = ge.id INNER JOIN
                      video_games.region_sales AS re ON pl.id = re.game_platform_id
GROUP BY ge.genre_name, pl.release_year
ORDER BY ge.genre_name, pl.release_year
```

100 %

Results Messages

	genre_name	release_year	DoanhSo
1	Action	1980	0.34
2	Action	1981	14.79
3	Action	1982	6.50
4	Action	1983	2.86
5	Action	1984	1.85
6	Action	1985	3.52
7	Action	1986	13.74
8	Action	1987	1.11
9	Action	1988	1.74
10	Action	1989	4.65
11	Action	1990	6.39
12	Action	1991	6.76
13	Action	1992	3.84
14	Action	1993	1.81
15	Action	1994	1.55
16	Action	1995	3.58
17	Action	1996	20.60
18	Action	1997	27.59
19	Action	1998	39.41
20	Action	1999	27.73
21	Action	2000	33.98
22	Action	2001	59.38
23	Action	2002	86.65

5.2.3. Câu 3: Cho biết doanh thu của các thể loại game theo từng khu vực?

-Sử dụng SSAS

The screenshot shows the SSAS Management Studio interface. On the left, the cube structure is displayed under the <All> node. The cube is named "Video Games DW" and contains the following components:

- Measures:**
 - Fact Game Sale
 - Fact Game Sale Count
 - Num Sales
- KPIs**
- Dimensions:**
 - Dim Game
 - Dim Genre
 - Genre Name
 - Id
 - Dim Platform
 - Dim Publisher
 - Dim Region
 - Id
 - Region Name
 - Dim Year
 - Id
 - Release Year

Genre Name	Region Name	Num Sales
Action	Europe	516.479999...
Action	Japan	158.85
Action	North America	861.799999...
Action	Other	184.92
Adventure	Europe	63.79
Adventure	Japan	51.82
Adventure	North America	102.06
Adventure	Other	16.699999...
Fighting	Europe	100
Fighting	Japan	87.150000...
Fighting	North America	220.74
Fighting	Other	36.19
Misc	Europe	213.82
Misc	Japan	106.51
Misc	North America	402.48
Misc	Other	74.020000...
Platform	Europe	200.67
Platform	Japan	130.65
Platform	North America	445.99
Platform	Other	51.51
Puzzle	Europe	50.53
Puzzle	Japan	56.68
Puzzle	North America	122.01
Puzzle	Other	12.469999...

-Pivot Table trên Excel

Row Labels	Column Labels	Grand Total				
		Europe	Japan	North America	Other	Grand Total
Action	516.48	158.85		861.8	184.92	1722.05
Adventure	63.79	51.82		102.06	16.7	234.37
Fighting		100	87.15		220.74	36.19
Misc	213.82	106.51		402.48	74.02	796.83
Platform	200.67	130.65		445.99	51.51	828.82
Puzzle		50.53	56.68		122.01	12.47
Racing	237.58	56.64		358.57	77	729.79
Role-Playing	187.58	350.29		326.5	59.38	923.75
Shooter	310.5	38.18		575.16	101.91	1025.75
Simulation	113.2	63.54		181.78	31.36	389.88
Sports	371.33	134.76		670.09	132.65	1308.83
Strategy	44.89	49.26		67.89	11.22	173.26
Grand Total	2410.37	1284.33		4335.07	789.33	8819.1
19						

5.2.4. Câu 4: Cho biết Doanh thu của các thể loại game theo từng nền tảng?

-Sử dụng SSAS

Video Games DW

Metadata

Search Model

Measure Group:

<All>

Video Games DW

- Measures
 - Fact Game Sale
 - Fact Game Sale Count
 - Num Sales
- KPIs
- Dim Game
 - Game Name
 - Id
- Dim Genre
 - Genre Name
 - Id
- Dim Platform
 - Id
 - Platform Name
- Dim Publisher
- Dim Region
 - Id
 - Region Name

Calculated Members

Dimension <Select dimension> Hierarchy

Genre Name	Platform Name	Num Sales
Action	2600	26.31
Action	3DS	56.49
Action	DC	1.26
Action	DS	114.02
Action	GB	7.93
Action	GBA	54.13
Action	GC	37.11
Action	GEN	2.74
Action	N64	29.4
Action	NES	28.75
Action	PC	30.54000...
Action	PS	125.7
Action	PS2	272.43
Action	PS3	304.05
Action	PS4	87.08
Action	PSP	62.77
Action	PSV	19.99
Action	SAT	0.65
Action	SNES	10.08
Action	Wii	110.41
Action	WiiU	19.31
Action	X360	239.68
Action	XB	47.44
Action	XOne	33.78

-Pivot Table

-Truy vấn SQL

```
SELECT CASE WHEN Grouping(g.id) = 1 THEN N'Tổng sale' ELSE p.platform_name END AS TenPlatform, SUM(r.num_sales) AS DoanhThu
FROM video_games.game_platform AS g INNER JOIN
video_games.platform AS p ON g.platform_id = p.id INNER JOIN
video_games.region_sales AS r ON g.id = r.game_platform_id
GROUP BY CUBE (g.id, p.platform_name)
```

% ▾

Results Messages

TenPlatform	DoanhThu
.. PS3	0.99
.. PS3	1.28
.. PS3	1.23
.. PS3	0.38
.. PS3	1.12
.. PS3	1.09
.. PS3	1.06
.. PS3	1.07
.. PS3	0.87
.. PS3	0.11
.. PS3	0.64
.. PS3	0.75
.. PS3	0.71
.. PS3	0.68
.. PS3	0.64
.. PS3	0.40
.. PS3	0.53
.. PS3	0.63
.. PS3	0.61
.. PS3	0.29
.. PS3	0.59
.. PS3	0.46
.. PS3	0.39
.. PS3	0.44
.. PS3	0.26
.. PS3	0.40
PS3	0.12

Query executed successfully.

| laptop (15.0 RTM) | LOPTOP\ThangTran (55)

5.2.5. Câu 5: Doanh thu của các thể loại game theo từng năm và tổng doanh thu của toàn thị trường?

-Truy vấn SQL

```

WITH BANG AS (SELECT ROW_NUMBER() OVER (ORDER BY newid()) AS STT, * FROM dbo.v_DSGenreTheoTungNam)
SELECT CASE WHEN Grouping(BANG.genre_name) = 1 THEN N'Thành Tiên' ELSE BANG.genre_name END AS Genre,
       CASE WHEN Grouping(BANG.release_year) = 1 THEN 0 ELSE BANG.release_year END AS Year,
       SUM(BANG.num_sales) AS DoanhSo
FROM   BANG
GROUP BY ROLLUP (BANG.genre_name), ROLLUP (BANG.release_year)

```

100 %

Results

	Genre	Year	DoanhSo
1	Action	1980	0.34
2	Fighting	1980	0.77
3	Misc	1980	2.70
4	Shooter	1980	7.07
5	Sports	1980	0.50
6	Thành Tiên	1980	11.38
7	Action	1981	14.79
8	Platform	1981	6.92
9	Puzzle	1981	2.25
10	Racing	1981	0.48
11	Shooter	1981	10.02
12	Simulation	1981	0.44
13	Sports	1981	0.78
14	Thành Tiên	1981	35.68
15	Action	1982	6.50
16	Misc	1982	0.87
17	Platform	1982	5.03
18	Puzzle	1982	10.04
19	Racing	1982	1.57
20	Shooter	1982	3.81
21	Sports	1982	1.06
22	Thành Tiên	1982	28.88
23	Action	1983	2.86

	Genre	Year	DoanhSo
416	Platform	2016	2.08
417	Racing	2016	1.67
418	Role-Play...	2016	6.75
419	Shooter	2016	18.17
420	Simulation	2016	0.39
421	Sports	2016	14.63
422	Strategy	2016	0.52
423	Thành Tiên	2016	70.87
424	Action	2017	0.01
425	Role-Play...	2017	0.04
426	Thành Tiên	2017	0.05
427	Simulation	2020	0.29
428	Thành Tiên	2020	0.29
429	Thành Tiên	0	8819.10
430	Action	0	1722.05
431	Adventure	0	234.37
432	Fighting	0	444.08
433	Misc	0	796.83
434	Platform	0	828.82

5.2.6. Câu 6: Số lượng doanh thu của các nền tảng game theo từng năm?

-Sử dụng SSAS

The screenshot shows the Microsoft Analysis Services (SSAS) Management Studio interface. On the left, the cube browser displays the structure of the 'Num Sales' cube, including dimensions like Dim Game, Dim Genre, Dim Platform, Dim Publisher, Dim Region, and Dim Year, along with their respective attributes. Below the cube browser is a 'Calculated Members' section which is currently empty. To the right of the browser is a large fact table grid. The columns of the grid are labeled 'Platform Name', 'Release Year', 'Num Sales', and 'Fact Game Sale Count'. The data in the grid represents sales data from 1980 to 2007, with values such as 11.38 for 1980, 35.68 for 1981, and so on. The bottom of the screen shows a dark blue footer bar with the text 'Error List Output'.

Platform Name	Release Year	Num Sales	Fact Game Sale Count
2600	1980	11.38	36
2600	1981	35.68	184
2600	1982	28.88	144
2600	1983	5.84	44
2600	1984	0.27	4
2600	1985	0.45	4
2600	1986	0.67	8
2600	1987	1.94	24
2600	1988	0.74	8
2600	1989	0.63	8
3DO	1994	0.02	4
3DO	1995	0.08	8
3DS	2011	62.58	464
3DS	2012	51.1	372
3DS	2013	55.86	364
3DS	2014	43.12	316
3DS	2015	26.97	344
3DS	2016	6.55	140
DC	1998	3.38	28
DC	1999	5.16	56
DC	2000	5.99	80
DC	2001	1.07	32
DC	2002	0.29	4
DC	2007	0.02	4

-Pivot Table

A	B	C	D	E
1	Row Labels	Num Sales	Fact Game Sale Count	
2	2600			
3	1980	11.38	36	
4	1981	35.68	184	
5	1982	28.88	144	
6	1983	5.84	44	
7	1984	0.27	4	
8	1985	0.45	4	
9	1986	0.67	8	
10	1987	1.94	24	
11	1988	0.74	8	
12	1989	0.63	8	
13	3DO			
14	1994	0.02	4	
15	1995	0.08	8	
16	3DS			
17	2011	62.58	464	
18	2012	51.1	372	
19	2013	55.86	364	
20	2014	43.12	316	
21	2015	26.97	344	
22	2016	6.55	140	
23	DC			
24	1998	3.38	28	
25	1999	5.16	56	
26	2000	5.99	80	
27	2001	1.07	32	
28	2002	0.29	4	

5.2.7. Câu 7: Số lượng doanh thu của các nền tảng game theo từng khu vực?

-SSAS

The screenshot shows the Microsoft SQL Server Data Tools (BIDS) interface. On the left, there is a 'Data View' pane showing the schema of the 'Video Games DW' database, including measures like 'Fact Game Sale Count' and 'Num Sales', and dimensions like 'Dim Game', 'Dim Genre', 'Dim Platform', 'Dim Publisher', and 'Dim Region'. On the right, there is a large data grid displaying sales data for various platforms across different regions. The columns in the grid are 'Platform Name', 'Region Name', 'Num Sales', and 'Fact Game Sale Count'. The data includes entries for platforms like 2600, 3DO, DS, DC, GB, and PC, and regions like Europe, Japan, North America, and Other.

Platform Name	Region Name	Num Sales	Fact Game Sale Count
2600	Europe	4.86	116
2600	Japan	0	116
2600	North America	80.77999...	116
2600	Other	0.84	116
3DO	Europe	0	3
3DO	Japan	0.1	3
3DO	North America	0	3
3DO	Other	0	3
3DS	Europe	58.3	500
3DS	Japan	97.30000...	500
3DS	North America	78.03	500
3DS	Other	12.55	500
DC	Europe	1.69	52
DC	Japan	8.56	52
DC	North America	5.43	52
DC	Other	0.27	52
DS	Europe	194.07	2133
DS	Japan	175.02	2133
DS	North America	388.5599...	2133
DS	Other	60.28999...	2133
GB	Europe	47.51	97
GB	Japan	85.12	97
GB	North America	113.64	97
GB	Other	8.16	97

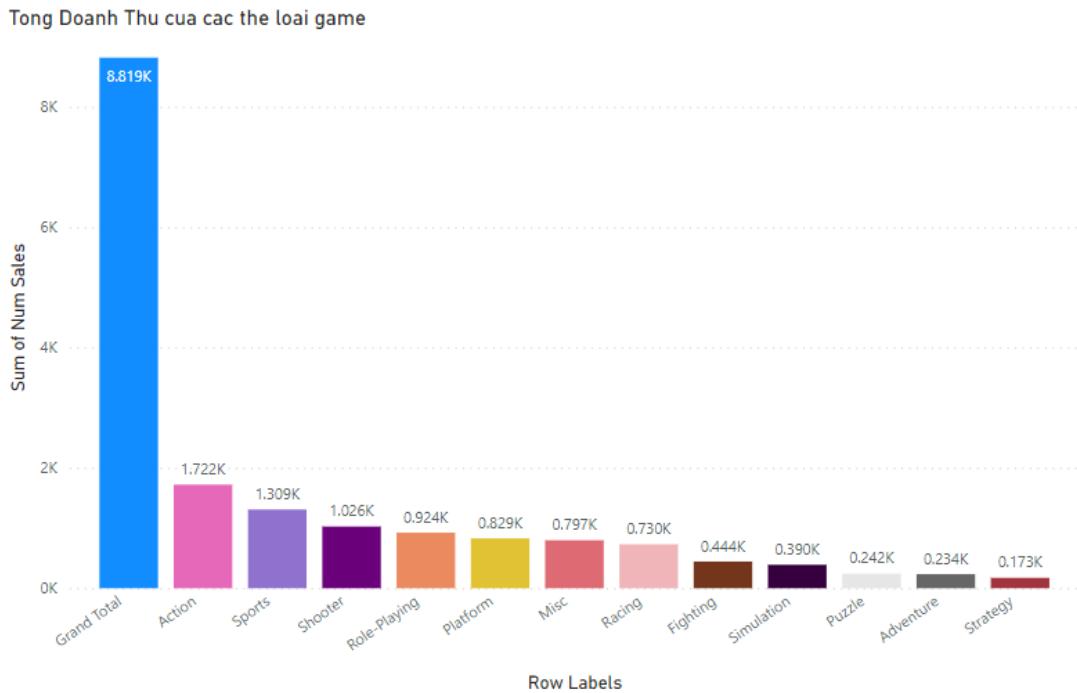
-Pivot Table

The screenshot shows a Microsoft Excel spreadsheet containing a Pivot Table. The table has multiple levels of detail, indicated by the 'Row Labels' and 'Column Labels' dropdown menus at the top. The data is organized into several sections: 'Num Sales' (with sub-sections for Europe, Japan, North America, and Other), 'Fact Game Sale Count' (with sub-sections for Europe, Japan, North America, and Other), and 'Total Num Sales' and 'Total Fact Game Sale Count'. The data includes numerous rows for different platforms like 2600, 3DO, DS, DC, GB, and PC, along with their respective sales figures for different regions.

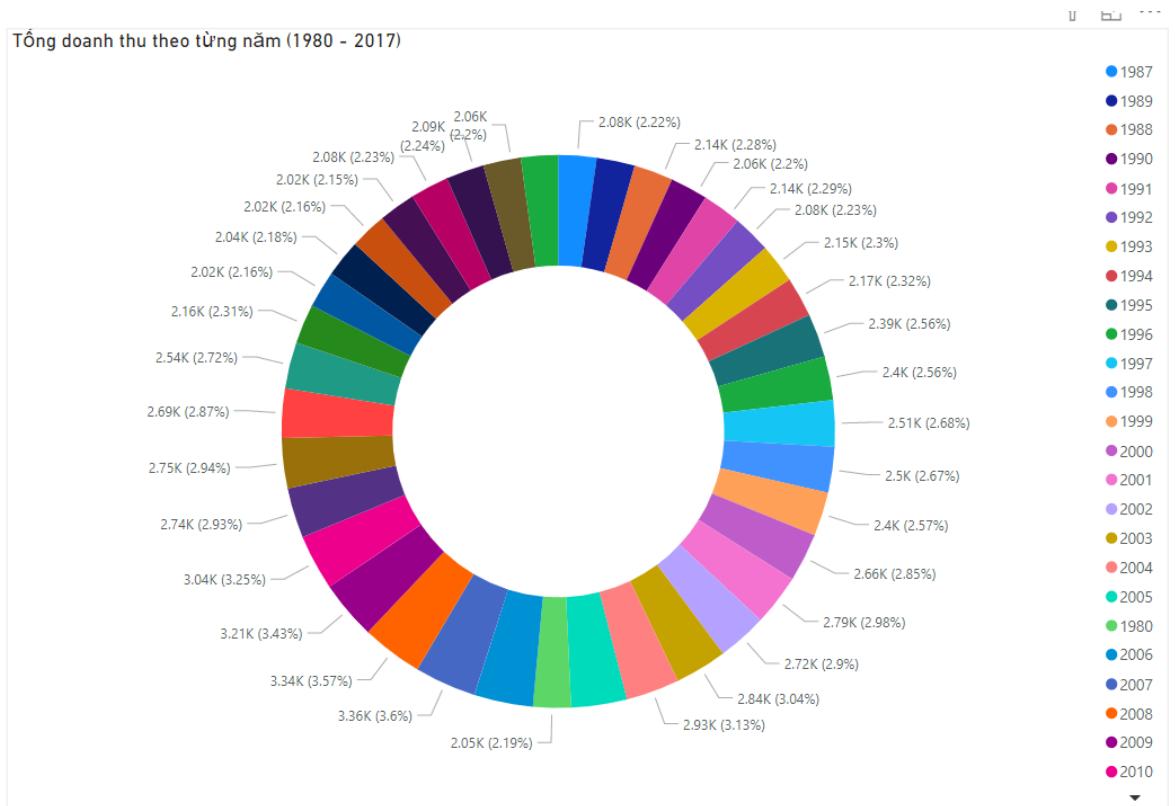
Row Labels	Column Labels				Fact Game Sale Count			Total Num Sales			Total Fact Game Sale Cou	
	Europe	Japan	North America	Other	Europe	Japan	North America	Other	Japan	North America	Other	Total
2600	4.86	0	80.78	0.84	116	116	116	116	86.48	86.48	86.48	4
3DO	0	0.1	0	0	3	3	3	3	0.1	0.1	0.1	20
DS	58.3	97.3	78.03	12.55	500	500	500	500	246.18	246.18	246.18	21
DC	1.69	8.56	5.43	0.27	52	52	52	52	15.95	15.95	15.95	2
GB	47.51	85.12	113.64	8.16	97	97	97	97	254.43	254.43	254.43	3
PC	194.07	175.02	388.56	60.29	2133	2133	2133	2133	817.94	817.94	817.94	85
PS	74.59	46.56	184.12	7.61	811	811	811	811	312.88	312.88	312.88	32
PS2	38.32	21.34	131.94	5.13	542	542	542	542	196.73	196.73	196.73	21
PS3	5.52	2.67	19.27	0.89	27	27	27	27	28.35	28.35	28.35	1
PS4	0	0.04	0	0	1	1	1	1	0.04	0.04	0.04	1
N64	41.03	33.76	138.91	4.31	316	316	316	316	218.01	218.01	218.01	12
NES	21.15	98.65	125.94	5.31	98	98	98	98	251.05	251.05	251.05	3
NG	0	1.44	0	0	12	12	12	12	1.44	1.44	1.44	1
PC	137.93	0.17	92.12	24.4	943	943	943	943	254.62	254.62	254.62	37
PCFX	0	0.03	0	0	1	1	1	1	0.03	0.03	0.03	1
PS	212.38	139.78	334.71	40.69	1189	1189	1189	1189	727.56	727.56	727.56	47
PS2	332.63	137.54	572.92	190.47	2127	2127	2127	2127	1233.56	1233.56	1233.56	85
PS3	340.46	79.21	388.9	140.81	1303	1303	1303	1303	949.38	949.38	949.38	52
PS4	123.7	14.3	96.8	43.36	336	336	336	336	278.16	278.16	278.16	13
PSP	67.16	75.89	107.09	41.52	1197	1197	1197	1197	291.66	291.66	291.66	47
PSV	16.27	20.89	16.07	8.41	412	412	412	412	61.64	61.64	61.64	16
SAT	0.54	32.26	0.72	0.07	173	173	173	173	33.59	33.59	33.59	6
SCD	0.36	0.45	1	0.05	6	6	6	6	1.86	1.86	1.86	1
SNES	19.04	116.55	61.23	3.22	239	239	239	239	200.04	200.04	200.04	9
TG16	0	0.16	0	0	2	2	2	2	0.16	0.16	0.16	1
Total	264.35	68.22	407.37	78.2	1200	1200	1200	1200	800.2	800.2	800.2	51

Chương 6. TRỰC QUAN HÓA BẰNG POWER BI

6.1. Tổng doanh thu của các thể loại game



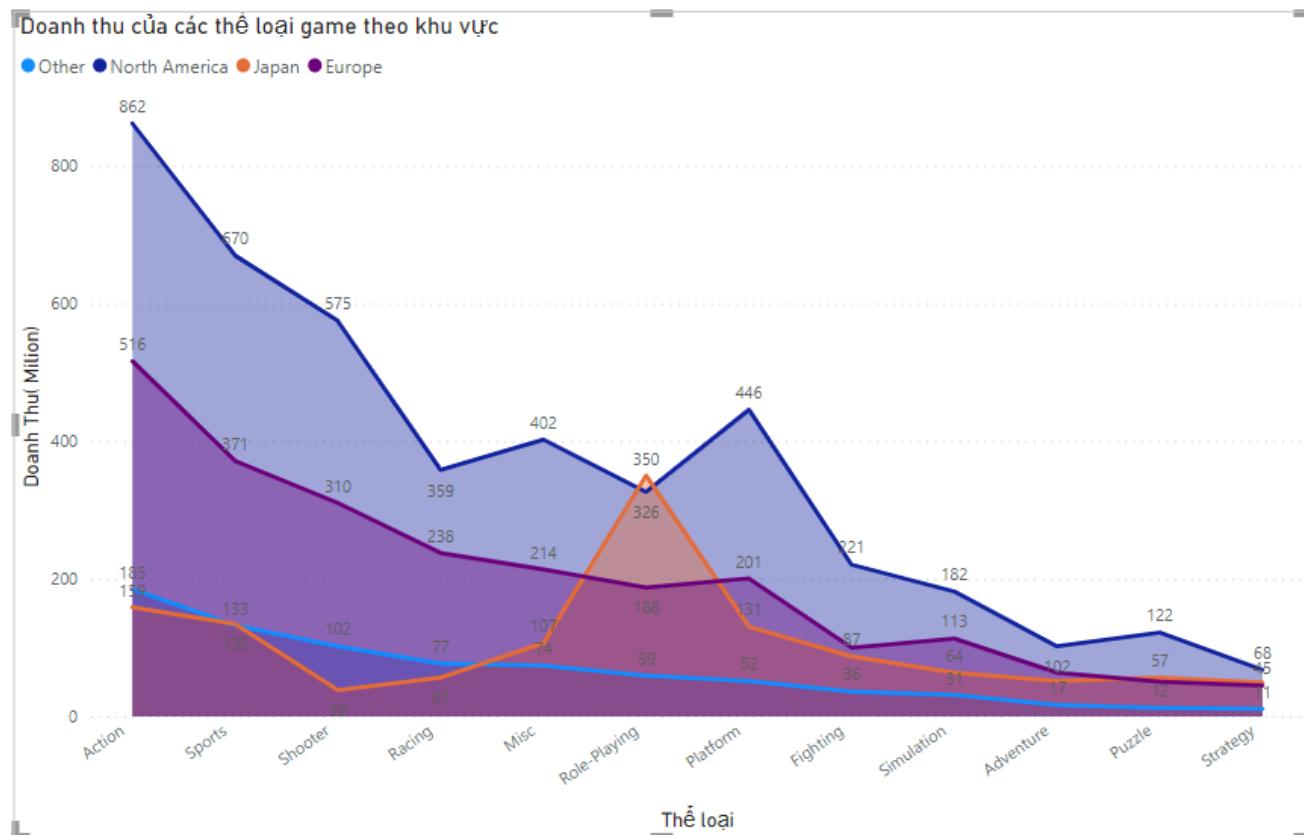
6.2. Tổng doanh thu theo từng năm(1980 - 2017)



Tổng doanh thu của các thể loại game theo từng năm



6.3. Tổng doanh thu theo khu vực



Phân chia công việc

Công Việc	Chi tiết	Trần Minh Đức	Trần Văn Thắng	Trần Văn Trọng
Tìm hiểu về Dataset	Tìm hiểu về nghiệp vụ của tập dữ liệu từ đó đưa ra các câu hỏi cũng như tạo kho	✓	✓	✓
Phân tích thiết kế Kho dữ liệu	Thiết kế Kho Stage, Ware House			✓
	Code tạo kế Kho Stage, Ware House		✓	✓
Tích hợp dữ liệu vào kho (SSIS)	Xử lý dữ liệu	✓		✓
	Đỗ dữ liệu vào kho Stage	✓		
	Đỗ dữ liệu từ Stage sang Ware House	✓		
Phân tích dữ liệu bằng SSAS	Tổng doanh thu của các thể loại game		✓	
	Doanh thu của các thể loại game theo từng năm		✓	
	Doanh thu của các thể loại game theo từng khu vực		✓	
	Doanh thu theo nền tảng		✓	
	Theo từng năm và toàn thị trường		✓	
	Doanh thu theo nền tảng theo từng năm		✓	
	Doanh thu theo		✓	
Trực quan hóa bằng Power BI	Tổng doanh thu của các thể loại game	✓		
	Tổng doanh thu theo từng năm		✓	

	Tổng doanh thu theo khu vực			✓
--	--------------------------------	--	--	---