



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



HCMUTE

**BÁO CÁO ĐỒ ÁN CUỐI KỲ
ĐỀ TÀI: XÂY DỰNG KHO DỮ LIỆU
KHÁCH HÀNG QUA CHIẾN DỊCH
MARKETING**

**Môn học: KHO DỮ LIỆU
Mã lớp học phần: DAWH430784_23_2_01
GVHD: ThS. Nguyễn Văn Thành**

Nhóm sinh viên thực hiện: Nhóm 8

Nguyễn Nhật Tân	21133079
Nguyễn Trung Đức	21142261
Huỳnh Phi Hùng	21122114
Lương Tường Vy	21133093

TP. Hồ Chí Minh, 16 tháng 05 năm 2024

DANH SÁCH THÀNH VIÊN THAM GIA

THỰC HIỆN ĐỀ TÀI VÀ VIẾT BÁO CÁO

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

STT	HỌ VÀ TÊN	MSSV	TỶ LỆ ĐÓNG GÓP
1	Nguyễn Nhật Tân	21133079	100%
2	Nguyễn Trung Đức	21142261	100%
3	Huỳnh Phi Hùng	21133114	100%
4	Lương Tường Vy	21133093	100%

Nhận xét của giảng viên:

Ngày ... tháng 05 năm 2024

Giảng viên chấm điểm

Ths. Nguyễn Văn Thành

LỜI CẢM ƠN

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

Nhóm chúng em xin chân thành cảm!

PHÂN CÔNG NHIỆM VỤ NHÓM 8

Nhiệm vụ	Tân	Đức	Hùng	Vy
Tìm kiếm tập dữ liệu	x	x		
Tiền xử lý dữ liệu	x	x		
Thiết kế datawarehouse	x	x		
Thực hiện ETL đưa dữ liệu vào các bảng dim và fact dùng SSIS				x
+ Tạo và đưa dữ liệu và bảng Dim_Customer				x
+ Tạo và đưa dữ liệu và bảng Dim_Date				x
+ Tạo và đưa dữ liệu và bảng Fact_MarketingAnalytic				x
Nhập dữ liệu vào SSAS, tạo data source, data source view và cube			x	
Viết báo cáo			x	x

MỤC LỤC

LỜI CẢM ƠN.....	8
CHƯƠNG 1. GIỚI THIỆU TỔNG QUAN.....	3
1. Phát biểu về dữ liệu.....	3
1.1 Lý do chọn đề tài	3
1.2 Mô tả về dữ liệu	4
1.3 Thuộc tính kho dữ liệu	4
1.4 Kho dữ liệu đã thực hiện tiền xử lý	9
2. Xây dựng kho dữ liệu.....	14
2.1 Lược đồ kho dữ liệu (Star Schema)	14
2.2 Chi tiết các bảng.....	14
2.3 Nội dung 15 câu hỏi truy vấn	16
CHƯƠNG 2: XÂY DỰNG KHO DỮ LIỆU.....	17
2.1 Tổng quan mô hình SSIS.....	17
2.2 Tạo project SSIS và thiết lập kết nối.....	17
2.2.1 Tạo project SSIS	17
2.2.2 Tạo cơ sở dữ liệu và thiết lập kết nối	19
2.3 Làm sạch và nạp dữ liệu vào Data Lake	24
2.4 Định nghĩa và nạp dữ liệu vào các bảng Dimension	38
2.5 Định nghĩa và nạp dữ liệu vào bảng Fact	50
2.6 Định nghĩa và Set các khóa ngoại cho bảng Fact	60
2.7 Thực thi Package.....	61
2.8 Dữ liệu sau khi hoàn thành	62
CHƯƠNG 3: QUÁ TRÌNH SSAS.....	64
3.1 Xây dựng khôi dữ liệu.....	64
3.1.1 Tạo project trong visual studio 2022.....	64
3.1.2 Tạo và deploy Cube	70
3.1.3 Tạo Calculations và Named set	91
3.2 Truy vấn dữ liệu	93
Câu 1 - Cho biết top 5 khách hàng có số lượng đơn hàng kèm khuyến mãi nhiều nhất của từng quốc gia.....	93

Câu 2 - Quốc gia nào có số đơn mua hàng tại store nhiều nhất	95
Câu 3 - Cho biết thu nhập hằng năm của khách hàng đã chấp nhận 4 chiến dịch marketing	97
Câu 4 - Cho biết nơi ở của 10 khách hàng có tổng chi cho các mặt hàng là trái cây theo chiều giảm dần lấy từ vị trí thứ 3	98
Câu 5 - In ra những khách hàng chi cho mặt hàng kẹo nhiều hơn 100\$ nhưng số trẻ em trong nhà ít hơn 2	100
Câu 6 - Mỗi tháng trong năm 2014, liệt kê 5 khách hàng có tổng số đơn mua hàng nhiều nhất.....	102
Câu 7 - Thống kê những khách hàng vừa mua hàng ở web hơn 10 lần vừa mua hàng ở cửa hàng hơn 10 lần	104
Câu 8 - Truy vấn drill down các tháng của năm 2012 có lượng tiêu thụ các mặt hàng về rượu nhiều hơn 2000\$	105
Câu 9 - Cho biết top 10 khách hàng có thu nhập hằng năm lớn nhất	107
Câu 10 - Cho biết tên 2 quốc gia nào có số lượng mua hàng ở store ít nhất	108
Câu 11 - Tính trung bình số tiền đã chi cho tất cả các mặt hàng của 5% khách hàng có thu nhập hàng năm cao nhất theo từng quốc gia	110
Câu 12 - Thống kê số tiền đã mua hàng của top 10 khách hàng có thu nhập hàng năm lớn hơn 30000\$	112
Câu 13 - Cho biết 3 tháng có số lượng đơn hàng nhiều nhất tại South Africa	113
Câu 14 - In ra ngày có ít lượng truy cập vào web nhất trong tháng 9 năm 2012	115
Câu 15 - Cho biết số lượng đơn hàng đã thanh toán qua catalog và thu nhập hằng năm của top 15 khách hàng độc thân và mới tốt nghiệp (tăng dần theo số lượng đơn hàng)	116
3.3 Report với Power BI	118
3.3.1 Tạo project với Power BI.....	118
3.3.2 Report 1 – Câu 1	120
3.3.3 Report 2 – Câu 7	123
3.3.4 Report 3 – Câu 13	126
3.3.5 Publish Report lên Power BI Server	130
CHƯƠNG 4: KẾT LUẬN.....	132
4.1. Kết quả đạt được	132
4.2. Những hạn chế	132
4.3. Tài liệu tham khảo	132

CHƯƠNG 1. GIỚI THIỆU TỔNG QUAN

1. Phát biểu về dữ liệu

1.1 Lý do chọn đề tài

Bên cạnh tốc độ phát triển vượt bậc về công nghệ hiện nay, nhiều lĩnh vực kinh tế được thúc đẩy mạnh, thị trường dần xuất hiện đa dạng các doanh nghiệp, cá nhân buôn bán với quy mô lớn nhỏ khác nhau. Cùng với đó, nhiều thử thách được đặt ra với các doanh nghiệp hoạt động trong lĩnh vực buôn bán và thương mại, đặc biệt nhất có thể kể tới là việc dẫn đầu thị trường từ việc đẩy mạnh Marketing trên nhiều nền tảng khác nhau nhằm tối đa hóa được lợi nhuận

Xác định vấn đề chính yếu có thể giúp doanh nghiệp tăng lợi nhuận một cách đáng kể, nhóm nhận thấy việc tổ chức xây dựng một hệ thống giúp quản lý và phân tích dữ liệu là vô cùng cần thiết. Từ hệ thống này, doanh nghiệp có thể đưa ra được cái nhìn tổng thể, phân loại được nhóm khách hàng có khả năng phản hồi tích cực với các chiến dịch marketing. Từ đó, doanh nghiệp có thể giảm thiểu chi phí bằng cách thu hẹp phạm vi marketing, lựa chọn được những khách hàng có tiềm năng cao dựa trên tập dữ liệu và các thuật toán phân loại

1.2 Mô tả về dữ liệu

The screenshot shows the Kaggle website interface. On the left, there's a sidebar with navigation links: Create, Home, Competitions, Datasets (which is selected), Code, Discussions, Courses, and More. The main area displays a dataset titled 'Marketing Analytics' by Jack Daoud, updated 9 months ago (Version 1). The page includes a search bar at the top right, a 'Sign In' button, and a 'Register' button. Below the title, there are sections for Data, Tasks (2), Code (50), Discussion (2), Activity, and Metadata. A large preview image shows various charts and graphs related to marketing data. At the bottom, there are sections for Usability (10.0), License (CC0: Public Domain), and Tags (business, exploratory data analysis, regression, statistical analysis, marketing).

- Dữ liệu về các chiến dịch marketing được public vào tháng 12 năm 2020 và thu nhập trong 2 năm (từ năm 2012-2014)
- Thông qua dữ liệu, người dùng có thể biết được thông tin của khách hàng, số tiền đã chi trả cho từng loại hàng hóa, số hóa đơn đã đặt theo các hình thức khác nhau, phản hồi của khách hàng,...
- Dữ liệu bao gồm 28 cột, 2240 dòng
- Link dữ liệu gốc: <https://www.kaggle.com/jackdaoud/marketing-data>

1.3 Thuộc tính kho dữ liệu

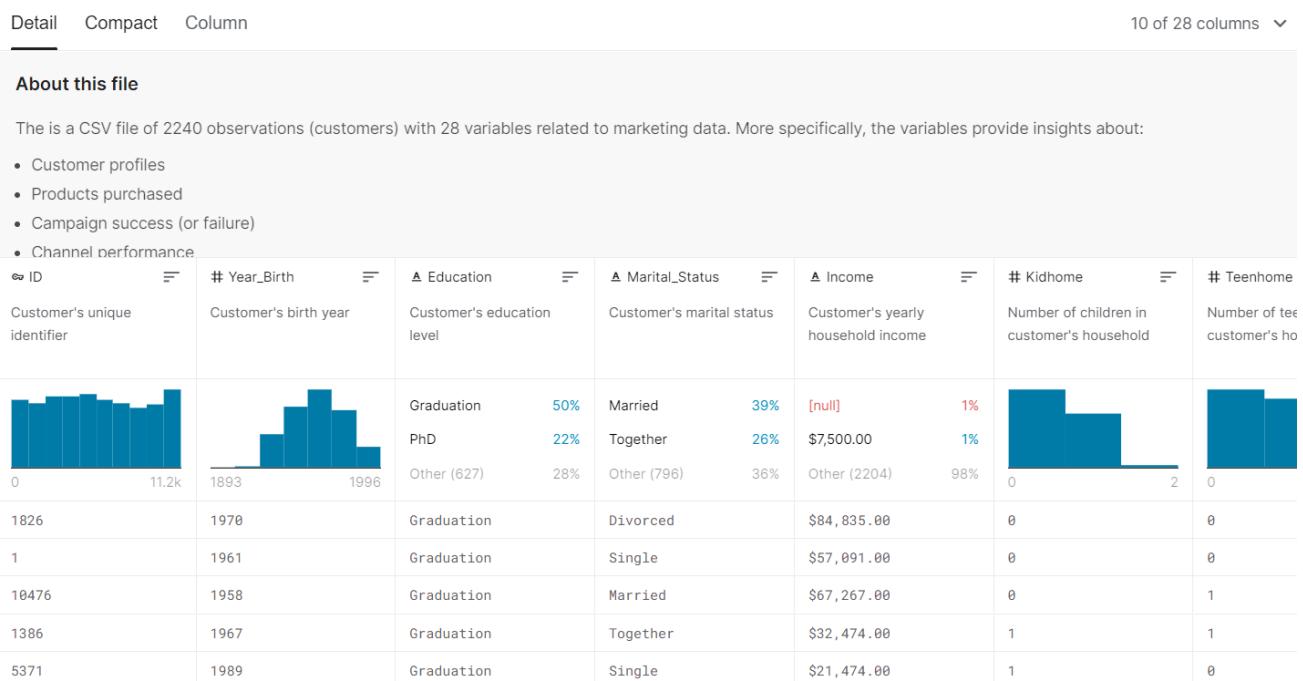
STT	Thuộc tính	Kiểu dữ liệu	Ý nghĩa	Giá trị
1	ID	Int	Mã khách hàng	Random number
2	Year_Birth	Int	Năm sinh	Trong khoảng [1893, 1996]
3	Education	Varchar(255)	Trình độ học vấn	1. Graduation 2. Master 3. PhD 4. 2n Cycle 5. Basic

4	Marital_Status	Varchar(255)	Tình trạng hôn nhân	1. Married 2. Single 3. Widow 4. Divorced 5. Together 6. Alone 7. YOLO 8. Absurd
5	Income	Object	Thu nhập hằng năm	Trong khoảng [1730, 666666]
6	Kidhome	Int	Số lượng trẻ em trong nhà	Trong khoảng [0, 2]
7	Teenhome	Int	Số lượng trẻ vị thành niên trong nhà	Trong khoảng [0, 2]
8	Dt_Customer	Date	Ngày ghi danh vào hệ thống	Trong khoảng [2012, 2014]
9	Recency	Int	Số lượng ngày kể từ khi khách hàng mua hàng lần cuối	Trong khoảng [0, 99]
10	MntWines	Int	Tổng tiền khách hàng chi trả cho mặt hàng về rượu (trong 2 năm gần nhất)	Trong khoảng [0, 1493]
11	MntFruits	Int	Tổng tiền khách hàng chi trả cho mặt hàng về hoa	Trong khoảng [0, 199]

			quả (trong 2 năm gần nhất)	
12	MntMeatProducts	Int	Tổng tiền khách hàng chi trả cho mặt hàng về thịt (trong 2 năm gần nhất)	Trong khoảng [0, 1725]
13	MntFishProducts	Int	Tổng tiền khách hàng chi trả cho mặt hàng về cá (trong 2 năm gần nhất)	Trong khoảng [0, 259]
14	MntSweetProducts	Int	Tổng tiền khách hàng chi trả cho mặt hàng về kẹo (trong 2 năm gần nhất)	Trong khoảng [0, 262]
15	MntGoldProds	Int	Tổng tiền khách hàng chi trả cho mặt hàng về vàng (trong 2 năm gần nhất)	Trong khoảng [0, 321]
16	NumDealsPurchases	Int	Số lượng đơn hàng có khuyến mãi đã thanh toán	Trong khoảng [0, 15]
17	NumWebPurchases	Int	Số lượng đơn hàng đã thanh toán thông qua website bán hàng	Trong khoảng [0, 27]

18	NumCatalogPurchases	Int	Số lượng đơn hàng đã thanh toán thông qua catalouge	Trong khoảng [0, 28]
19	NumStorePurchases	Int	Số lượng đơn hàng đã thanh toán trực tiếp tại cửa hàng	Trong khoảng [0, 13]
20	NumWebVisits Month	Int	Số lần mà khách hàng ghé thăm website mua hàng của cửa hàng (tháng gần nhất)	Trong khoảng [1, 19]
21	AcceptedCmp1	Int	Khách hàng phản hồi Marketing Campaign lần 1	0: Không chấp nhận 1. Chấp nhận
22	AcceptedCmp2	Int	Khách hàng phản hồi Marketing Campaign lần 2	0: Không chấp nhận 1. Chấp nhận
23	AcceptedCmp3	Int	Khách hàng phản hồi Marketing Campaign lần 3	0: Không chấp nhận 1. Chấp nhận
24	AcceptedCmp4	Int	Khách hàng phản hồi Marketing Campaign lần 4	0: Không chấp nhận 1. Chấp nhận
25	AcceptedCmp5	Int	Khách hàng phản hồi Marketing Campaign lần 5	0: Không chấp nhận 1. Chấp nhận

26	Response	Int	Khách hàng phản hồi Marketing Campaign lần cuối (lần gần nhất)	0: Không chấp nhận 1. Chấp nhận
27	Complain	Int	Phản hồi của khách hàng	0: Không phàn nàn 1. Phàn nàn
28	Country	Varchar(255)	Quốc gia	SA: South Africa CA: Canada, SP: Spain GER: Germany IND: India US: United States AUS: Australia ME: Montenegro



Hình 1.1 Kho dữ liệu khi chưa xử lý

1.4 Kho dữ liệu đã thực hiện tiền xử lý

- Sau khi tiền xử lý, ta được 2240 dòng và 28 cột

User_ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Date_register	Recency	MntWines	MntFruits	MntMeatP	MntFishPr	MntSweet	MntGoldPi	NumDeals	NumWebP	NumCatali	NumStore	NumWebA
0	1985	Graduation	Married	70951	0	0	5/4/2013	66	239	10	554	254	87	54	1	3	4	9	1
1	1961	Graduation	Single	57091	0	0	6/15/2014	0	464	5	64	7	0	37	1	7	3	7	5
9	1975	Master	Single	46098	1	1	8/18/2012	86	57	0	27	0	0	36	4	3	2	2	8
13	1947	PhD	Widow	25358	0	1	7/22/2013	57	19	0	5	0	0	8	2	1	0	3	6
17	1971	PhD	Married	60491	0	1	9/6/2013	81	637	47	237	12	19	76	4	6	11	7	5
20	1965	2n Cycle	Married	46891	0	1	9/1/2013	91	43	12	23	29	15	61	1	2	1	4	4
22	1976	Graduation	Divorced	46310	1	0	12/3/2012	99	185	2	88	15	5	14	2	6	1	5	8
24	1960	Master	Together	17144	1	1	2/15/2014	96	18	2	19	0	2	6	5	3	0	4	7
25	1958	Graduation	Married	65148	0	1	11/16/2012	9	460	35	422	33	12	153	2	6	6	7	4
35	1987	Graduation	Married	25545	1	0	9/8/2012	35	32	1	64	16	12	85	3	2	2	3	6
48	1964	Graduation	Together	55761	0	1	4/24/2014	97	136	1	12	0	3	32	2	4	1	3	6
49	1970	Graduation	Single	20587	1	0	5/11/2014	39	2	3	6	4	1	9	1	1	1	2	7
55	1963	Graduation	Together	56253	0	1	12/7/2012	83	509	0	65	7	11	5	4	7	2	9	6
67	1972	Master	Single	46423	1	1	9/18/2013	6	68	0	16	0	0	8	3	2	0	4	7
73	1953	PhD	Single	51411	1	2	5/26/2013	81	14	0	3	0	0	1	1	0	0	3	5
75	1982	2n Cycle	Together	41039	0	0	7/25/2013	32	34	11	137	179	61	45	1	8	1	5	8
78	1969	Graduation	Married	25293	1	0	9/15/2013	51	15	0	11	0	2	9	1	1	1	2	8
87	1981	2n Cycle	Married	27733	1	0	8/26/2013	16	0	7	5	26	2	17	2	2	0	3	7
89	1975	2n Cycle	Married	22212	1	0	12/9/2013	49	5	9	20	6	8	21	2	2	0	4	6
92	1988	Graduation	Alone	34176	1	0	5/12/2014	12	5	7	24	19	14	20	4	3	0	4	6
113	1951	Graduation	Married	30833	1	1	7/16/2013	33	11	0	7	3	1	6	2	1	0	3	5
115	1966	Master	Single	43456	0	1	3/26/2013	0	275	11	68	25	7	7	3	5	1	8	5
123	1951	Graduation	Widow	67046	0	1	1/6/2013	92	544	35	133	23	151	107	2	4	5	4	2
125	1958	2n Cycle	Together	53083	1	1	5/15/2013	65	215	7	33	3	0	13	6	5	1	5	7
143	1970	Graduation	Single	61209	0	0	8/25/2013	73	466	0	224	119	49	99	1	5	3	4	2
146	1960	PhD	Single	76045	0	0	11/15/2013	78	760	53	400	17	93	0	1	4	5	11	2
155	1972	Graduation	Divorced	22222	0	0	10/17/2012	30	10	2	6	4	1	12	1	0	2	1	4

Hình 1.2 Kho dữ liệu sau khi xử lý

STT	Thuộc tính	Kiểu dữ liệu	Ý nghĩa	Giá trị
1	User_ID	Int	Mã khách hàng	Random number
2	Year_Birth	Int	Năm sinh	Trong khoảng [1893, 1996]
3	Education	Varchar(255)	Trình độ học vấn	1. Graduation 2. Master 3. PhD 4. 2n Cycle 5. Basic
4	Marital_Status	Varchar(255)	Tình trạng hôn nhân	1. Married 2. Single 3. Widow 4. Divorced 5. Together 6. Alone

				7. YOLO 8. Absurd
5	Income	Float	Thu nhập hằng năm	Trong khoảng [1730, 666666]
6	Kidhome	Int	Số lượng trẻ em trong nhà	Trong khoảng [0, 2]
7	Teenhome	Int	Số lượng trẻ vị thành niên trong nhà	Trong khoảng [0, 2]
8	Date_Enroll	Date	Ngày ghi danh vào hệ thống	Trong khoảng [2012, 2014]
9	Recency	Int	Số lượng ngày kể từ khi khách hàng mua hàng lần cuối	Trong khoảng [0, 99]
10	MntWines	Float	Tổng tiền khách hàng chi trả cho mặt hàng về rượu (trong 2 năm gần nhất)	Trong khoảng [0, 1493]
11	MntFruits	Float	Tổng tiền khách hàng chi trả cho mặt hàng về hoa quả (trong 2 năm gần nhất)	Trong khoảng [0, 199]
12	MntMeats	Float	Tổng tiền khách hàng chi trả cho mặt hàng về thịt (trong 2 năm gần nhất)	Trong khoảng [0, 1725]

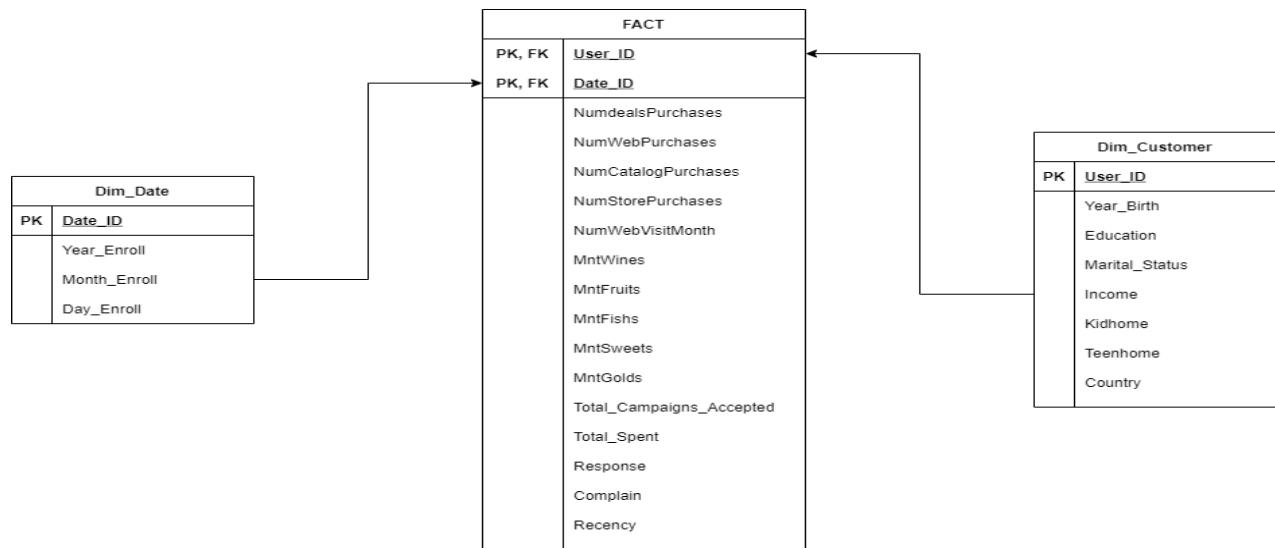
13	MntFishes	Float	Tổng tiền khách hàng chi trả cho mặt hàng về cá (trong 2 năm gần nhất)	Trong khoảng [0, 259]
14	MntSweets	Float	Tổng tiền khách hàng chi trả cho mặt hàng về kẹo (trong 2 năm gần nhất)	Trong khoảng [0, 262]
15	MntGolds	Float	Tổng tiền khách hàng chi trả cho mặt hàng về vàng (trong 2 năm gần nhất)	Trong khoảng [0, 321]
16	NumDealsPurchases	Int	Số lượng đơn hàng có khuyến mãi đã thanh toán	Trong khoảng [0, 15]
17	NumWebPurchases	Int	Số lượng đơn hàng đã thanh toán thông qua website bán hàng	Trong khoảng [0, 27]
18	NumCatalogPurchases	Int	Số lượng đơn hàng đã thanh toán thông qua catalouge	Trong khoảng [0, 28]
19	NumStorePurchases	Int	Số lượng đơn hàng đã thanh toán	Trong khoảng [0, 13]

			trực tiếp tại cửa hàng	
20	NumWebVisits Month	Int	Số lần mà khách hàng ghé thăm website mua hàng của cửa hàng (tháng gần nhất)	Trong khoảng [1, 19]
21	Response	Int	Khách hàng phản hồi Marketing Campaign lần cuối (lần gần nhất)	0: Không chấp nhận 1. Chấp nhận
22	Complain	Int	1 nếu khách hàng có phàn nàn (trong 2 năm gần nhất), 0 là trường hợp ngược lại	0: Không phàn nàn 1. Phàn nàn
23	Country	Varchar(255)	Quốc gia	SA: South Africa CA: Canada, SP: Spain GER: Germany IND: India US: United States AUS: Australia ME: Montenegro
24	AcceptedCmp1	Int	Phản hồi của khách hàng về chiến dịch marketing lần 1	0: Không chấp nhận 1. Chấp nhận

25	AcceptedCmp2	Int	Phản hồi của khách hàng về chiến dịch marketing lần 2	0: Không chấp nhận 1. Chấp nhận
26	AcceptedCmp3	Int	Phản hồi của khách hàng về chiến dịch marketing lần 3	0: Không chấp nhận 1. Chấp nhận
27	AcceptedCmp4	Int	Phản hồi của khách hàng về chiến dịch marketing lần 4	0: Không chấp nhận 1. Chấp nhận
28	AcceptedCmp5	Int	Phản hồi của khách hàng về chiến dịch marketing lần 5	0: Không chấp nhận 1. Chấp nhận

2. Xây dựng kho dữ liệu

2.1 Lược đồ kho dữ liệu (Star Schema)



2.2 Chi tiết các bảng

- Bảng FACT

Tên thuộc tính	Kiểu dữ liệu	Ràng buộc
User_ID	Int	Khóa chính, khóa ngoại
Date_ID	Date	Khóa chính, khóa ngoại
NumdealsPurchases	Int	
NumWebPurchases	Int	
NumCatalogPurchases	Int	
NumStorePurchases	Int	
NumWebVisitMonth	Int	
MntWines	Int	
MntFruits	Int	
MntMeats	Int	

MntFishes	Int	
MntSweets	Int	
MntGolds	Int	
Total_campaigns_accepted	Int	
Total_Spent	Int	
Response	Int	
Complain	Int	
Recency	Int	

- Bảng Dim_Date

Tên thuộc tính	Kiểu dữ liệu	Ràng buộc
Date_ID	Date	Khóa chính
Year_Enroll	Int	
Month_Enroll	Int	
Day_Enroll	Int	
Date_Enroll	Date	

- Dim_Customer

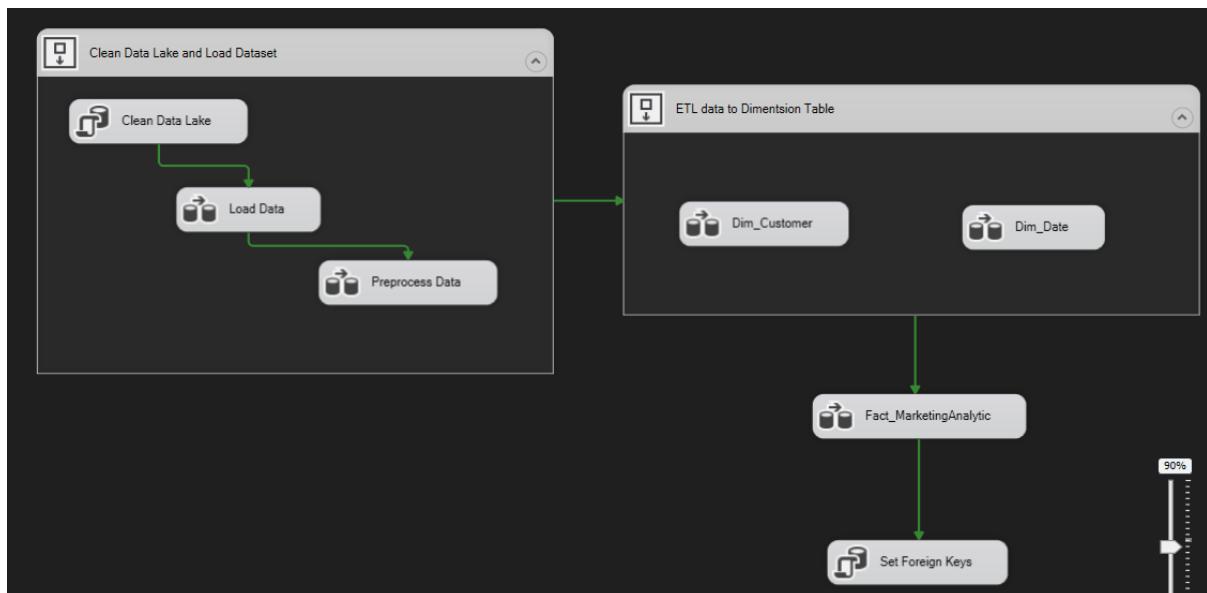
Tên thuộc tính	Kiểu dữ liệu	Ràng buộc
User_ID	Int	Khóa chính
Year_Birth	Int	
Education	Varchar(255)	
Marital_Status	Varchar(255)	
Income	Int	
Kidhome	Int	
Teenhome	Int	
Country	Varchar(255)	

2.3 Nội dung 15 câu hỏi truy vấn

1. Cho biết top 5 khách hàng có số lượng đơn hàng kèm khuyến mãi nhiều nhất của từng quốc gia
2. Quốc gia nào có số đơn mua hàng tại store nhiều nhất
3. Cho biết thu nhập hằng năm của khách hàng đã chấp nhận 4 chiến dịch marketing
4. Cho biết nơi ở của 10 khách hàng có tổng chi cho các mặt hàng là trái cây theo chiều giảm dần lấy từ vị trí thứ 3
5. In ra những khách hàng chi cho mặt hàng kẹo nhiều hơn 100\$ nhưng số trẻ em trong nhà ít hơn 2
6. Mỗi tháng trong năm 2014, liệt kê 5 khách hàng có tổng số đơn mua hàng nhiều nhất
7. Thống kê những khách hàng vừa mua hàng ở web hơn 10 lần vừa mua hàng ở cửa hàng hơn 10 lần
8. Truy vấn drill down các tháng của năm 2012 có lượng tiêu thụ các mặt hàng về rượu nhiều hơn 2000\$
9. Cho biết top 10 khách hàng có thu nhập hằng năm lớn nhất
10. Cho biết tên 2 quốc gia nào có số lượng mua hàng ở store ít nhất
11. Tính trung bình số tiền đã chi cho tất cả các mặt hàng của 5% khách hàng có thu nhập hàng năm cao nhất theo từng quốc gia
12. Thống kê số tiền đã mua hàng của top 10 khách hàng có thu nhập hàng năm lớn hơn 30000\$
13. Cho biết 3 tháng có số lượng đơn hàng nhiều nhất tại South Africa
14. In ra ngày có ít lượng truy cập vào web nhất trong tháng 9 năm 2012
15. Cho biết số lượng đơn hàng đã thanh toán qua catalog và thu nhập hằng năm của top 15 khách hàng độc thân và mới tốt nghiệp (tăng dần theo số lượng đơn hàng)

CHƯƠNG 2: XÂY DỰNG KHO DỮ LIỆU

2.1 Tổng quan mô hình SSIS



Bước 1 - Clean Data Lake and Load Dataset: Thực hiện xóa toàn bộ dữ liệu trong Data Lake, tiến hành đưa toàn bộ dữ liệu thô vào bảng tương ứng trong Data Lake, thực hiện tiền xử lý dữ liệu và đổ vào các bảng dữ liệu tương ứng

Bước 2 – ETL data to Dimension Table: Thực hiện đổ dữ liệu vào các bảng Dimension trong Data warehouse

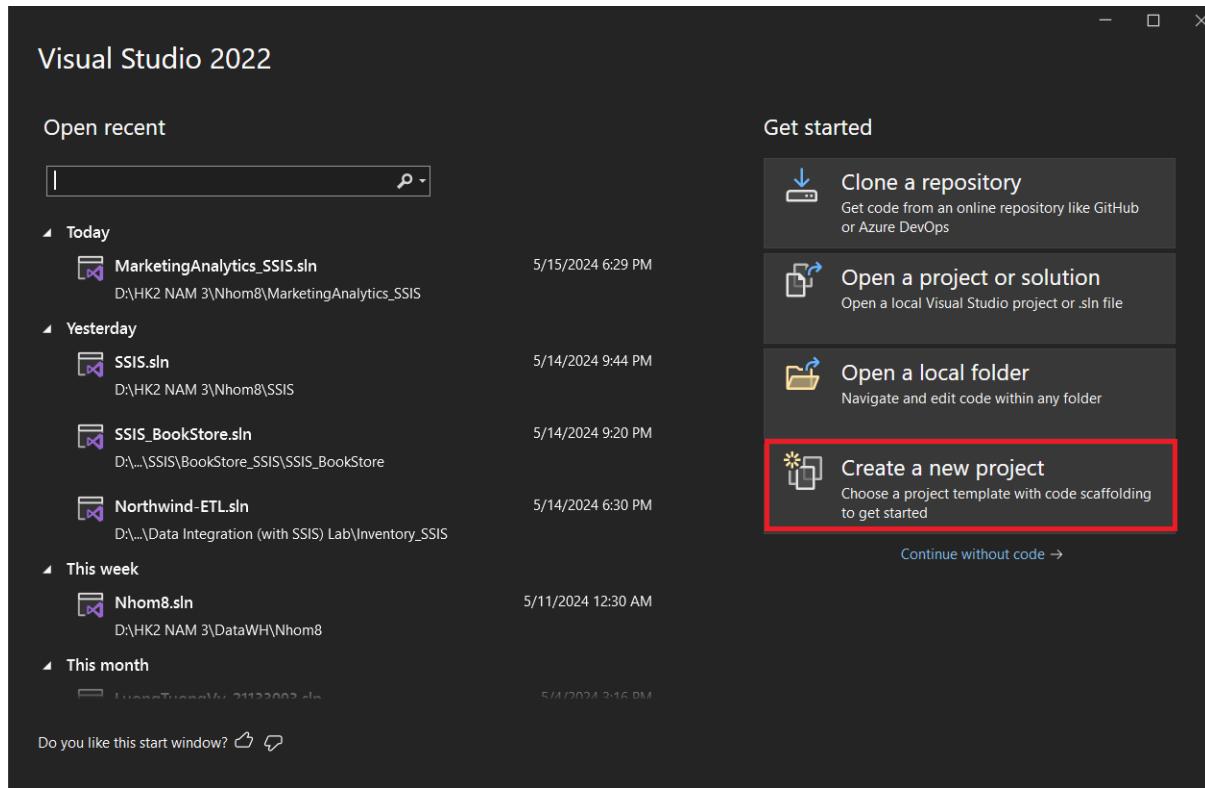
Bước 3 – Fact MarketingAnalytic: Thực hiện đổ dữ liệu vào bảng Fact trong Data warehouse

Bước 4 – Set Foreign Keys: Cài đặt các khóa ngoại tương ứng theo thiết kế lược đồ

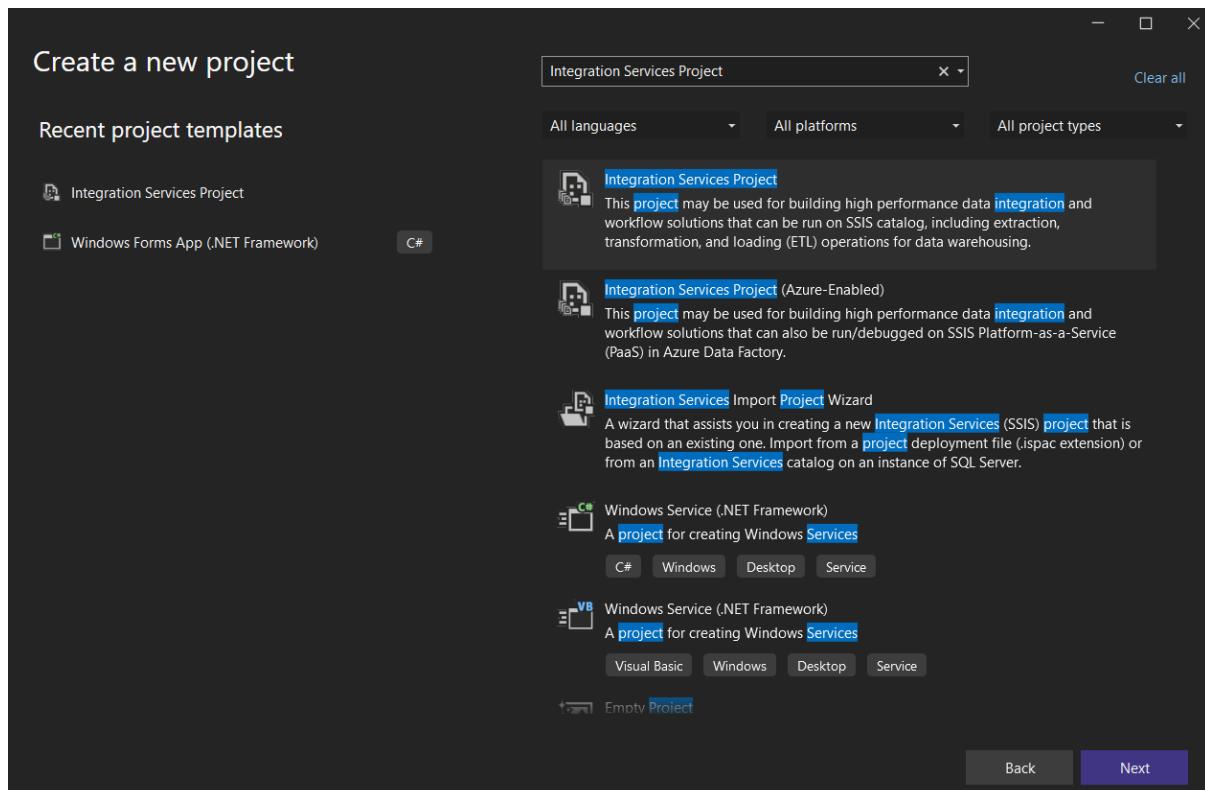
2.2 Tạo project SSIS và thiết lập kết nối

2.2.1 Tạo project SSIS

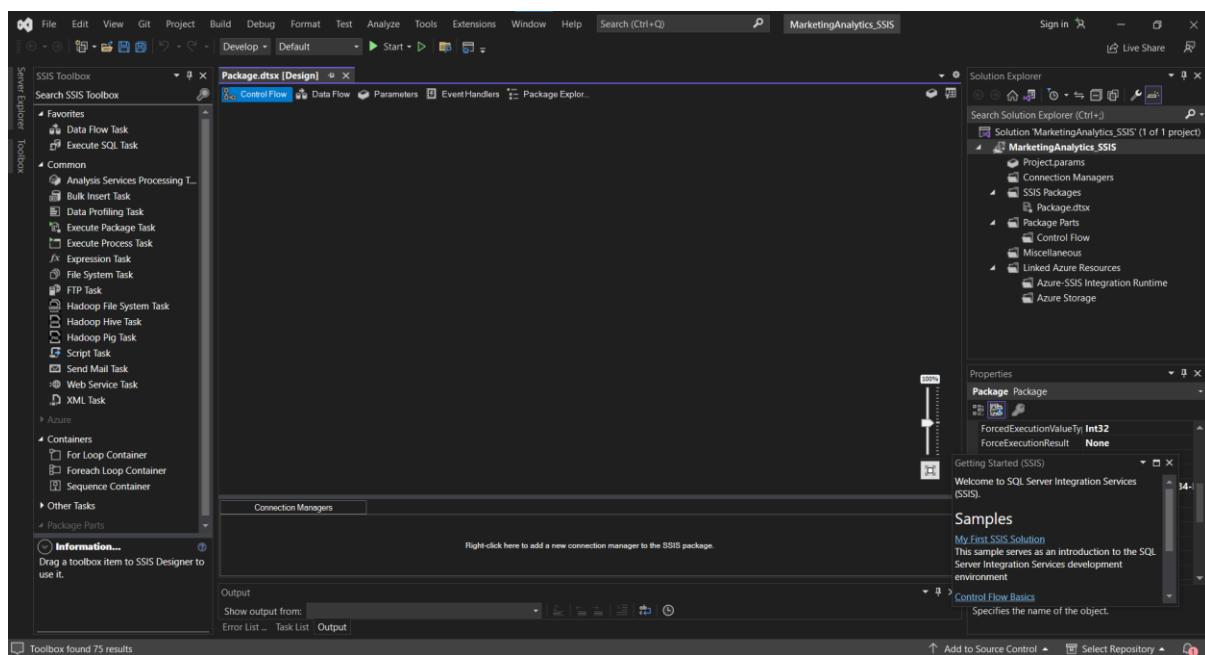
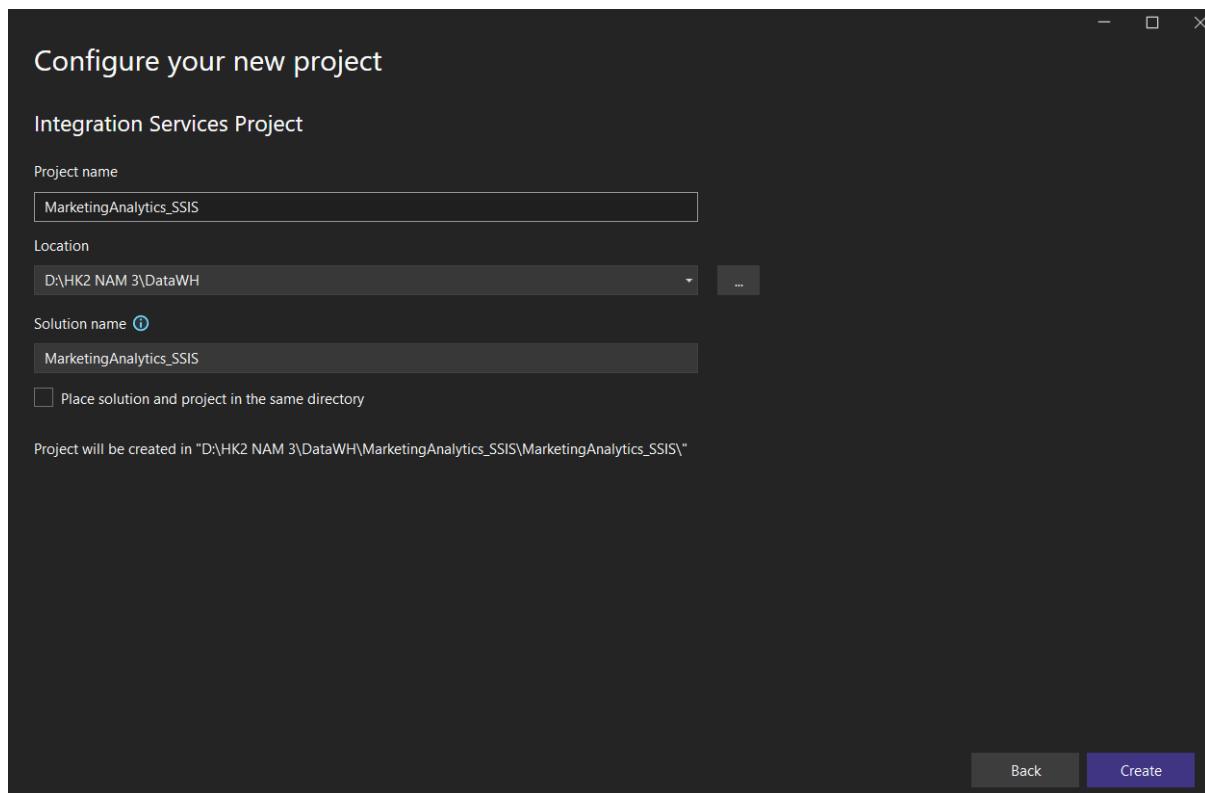
Bước 1: Mở visual studio 2022 -> Create a new project



Bước 2: Tìm Integration Services Project -> Click chọn và nhấn Next



Bước 3: Đặt tên cho project -> chọn location đặt project -> Create



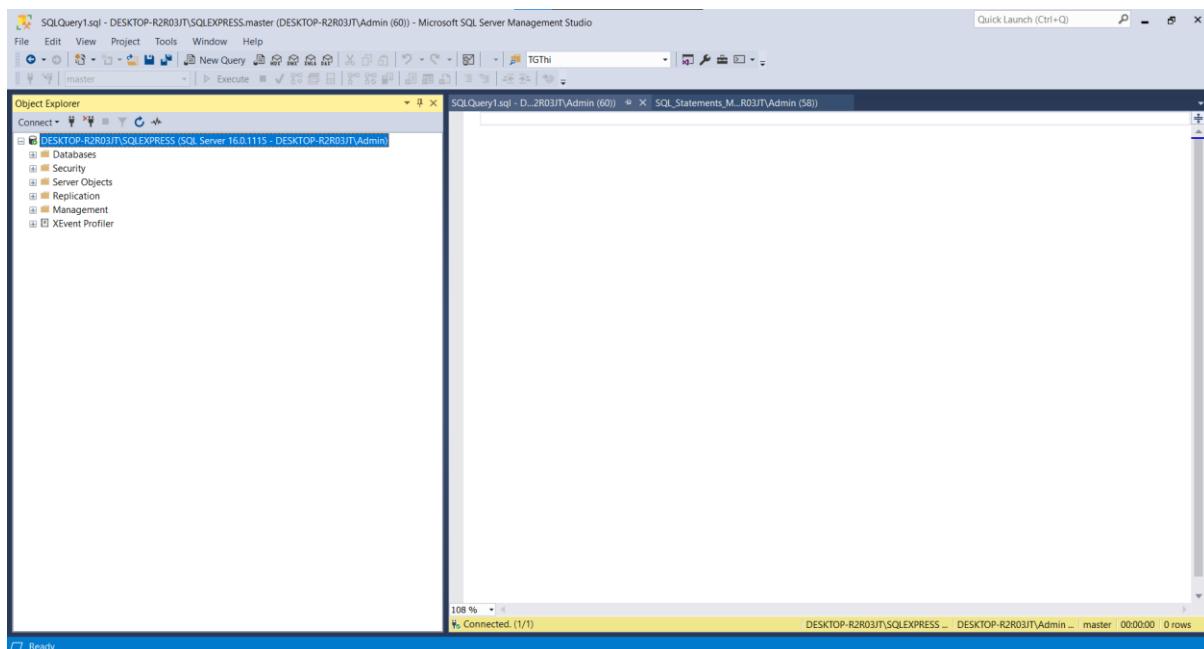
Hình 2.1.1 Màn hình giao diện của SSIS project

Giao diện chính của SSIS project, các chức năng chính nằm ở cột bên trái SSIS Toolbox

2.2.2 Tạo cơ sở dữ liệu và thiết lập kết nối

a. Tạo cơ sở dữ liệu

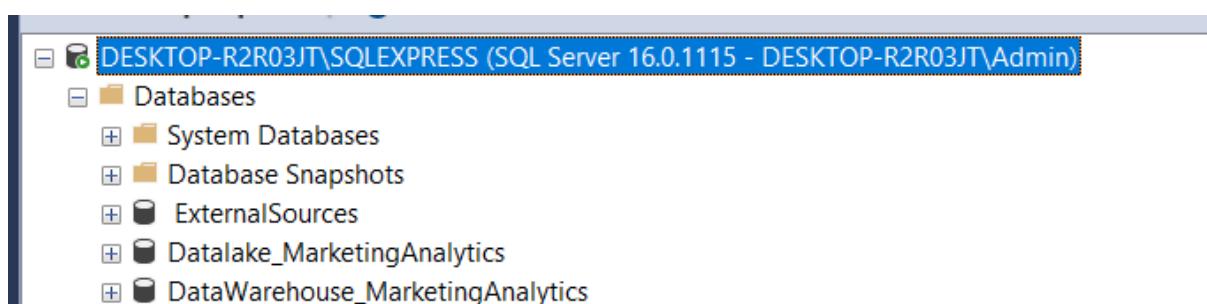
Bước 1: Vào MS SQL Server Management -> tạo New Query



Bước 2: Viết lệnh tạo database cần thiết -> Run

```
CREATE DATABASE Datalake_MarketingAnalytics;
CREATE DATABASE DataWarehouse_MarketingAnalytics;
```

Kết quả sau khi chạy lệnh tạo database



Datalake_MarketingAnalytics là database chứa các table sau

- Original_Data: Table chứa dữ liệu gốc, dữ liệu chưa qua các bước tiền xử lý
- Clean_Data: Table chứa dữ liệu đã trải qua các bước tiền xử lý (lọc Null, đổi tên column, thêm column,...)

- Null_Data: Table chứa những dòng dữ liệu chứa giá trị Null được lọc từ bảng Original_Data

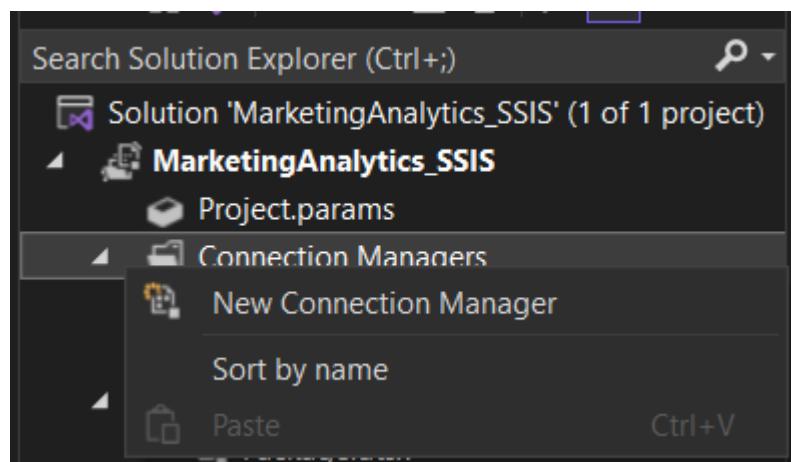
DataWarehouse_MarketingAnalytics là database đóng vai trò là một kho dữ liệu, chứa các table theo đúng thiết kế star schema ở mục Xây dựng lược đồ kho dữ liệu

- Dim_Customer: Bảng chiết chứa dữ liệu thông tin về khách hàng
- Dim_Date: Bảng chiết chứa dữ liệu thông tin về ngày tháng năm
- Fact_MarketingAnalytic: Bảng fact chứa thông tin và các thuộc tính dùng để tính toán

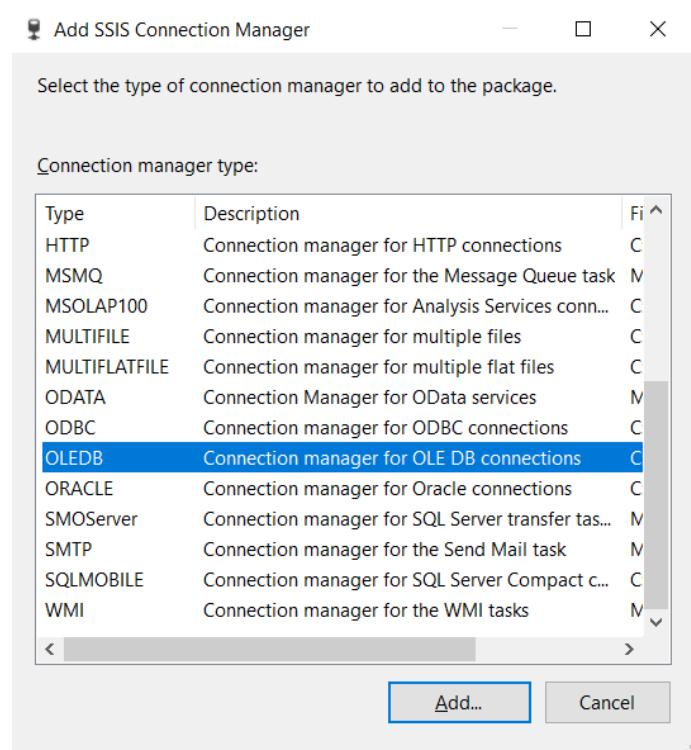
b. Thiết lập kết nối

b.1 Thiết lập kết nối đến database Datalake_MarketingAnalytics

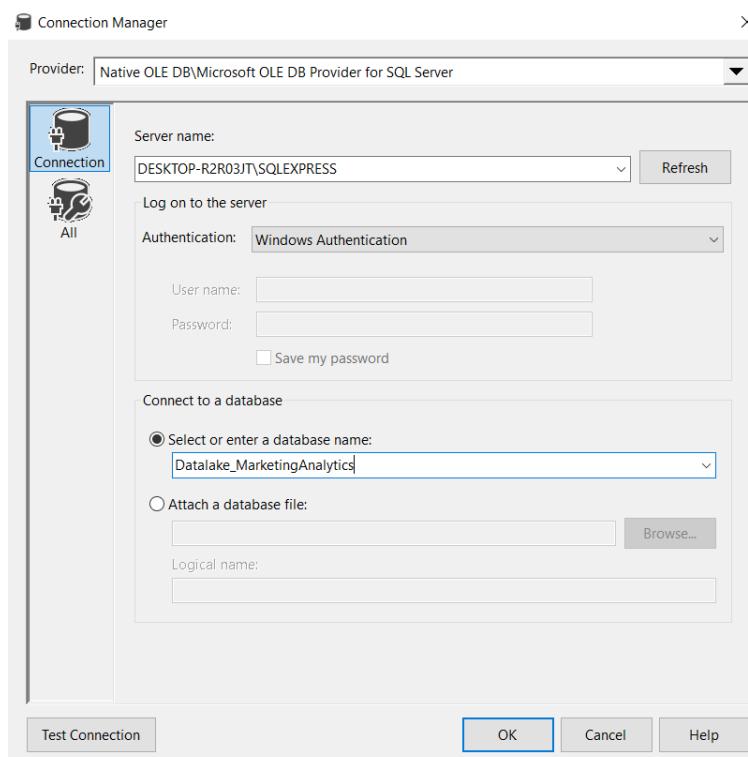
Bước 1: Mở Project SSIS trong Visual Studio -> Chuột phải Connection Managers -> New Connection Manager



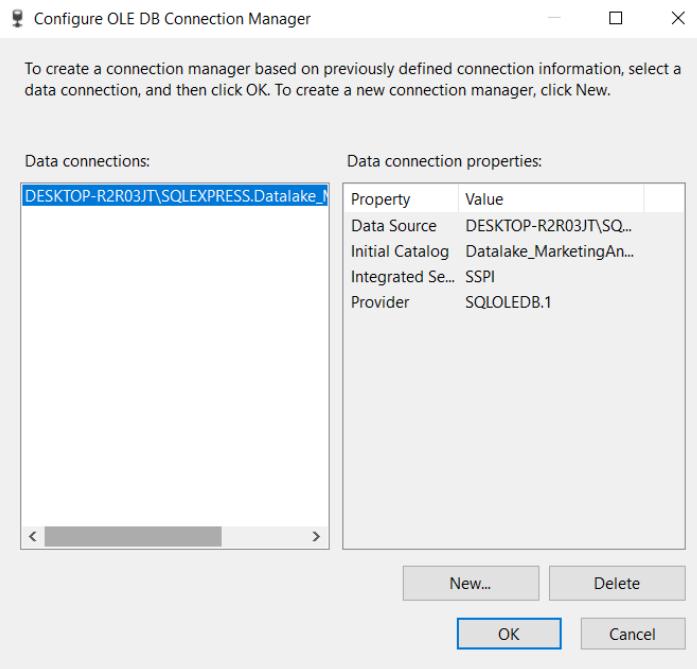
Bước 2: Chọn OLEDB -> Nhấn Add



Bước 3: Chọn New -> Vào MS SQL Server Management để lấy được Server Name -> Điền Server Name -> Tại bước Select database name, chọn Datalake_MarketingAnalytics -> Chọn Test Connection -> Chọn OK nếu test connection thành công



Kết quả thu được

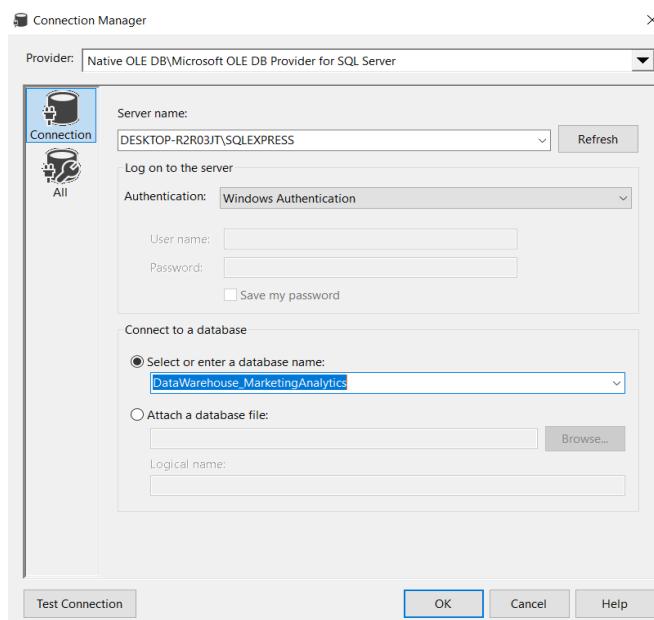


b.2 Thiết lập kết nối đến database DataWarehouse_MarketingAnalytics

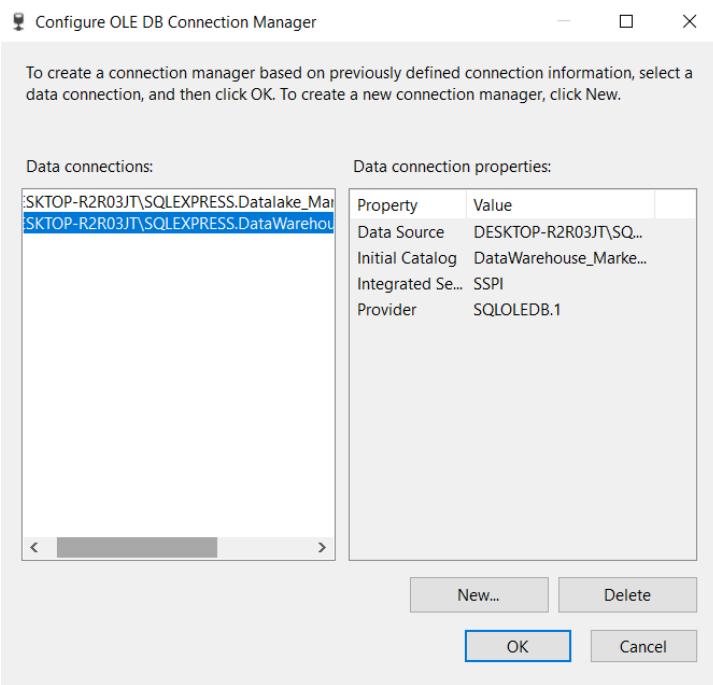
Bước 1: Thực hiện tương tự 2.2.2.1

Bước 2: Thực hiện tương tự 2.2.2.1

Bước 3: Chọn New -> Vào MS SQL Server Management để lấy được Server Name -> Điền Server Name -> Tại bước Select database name, chọn DataWarehouse_MarketingAnalytics -> Chọn Test Connection -> Chọn OK nếu test connection thành công

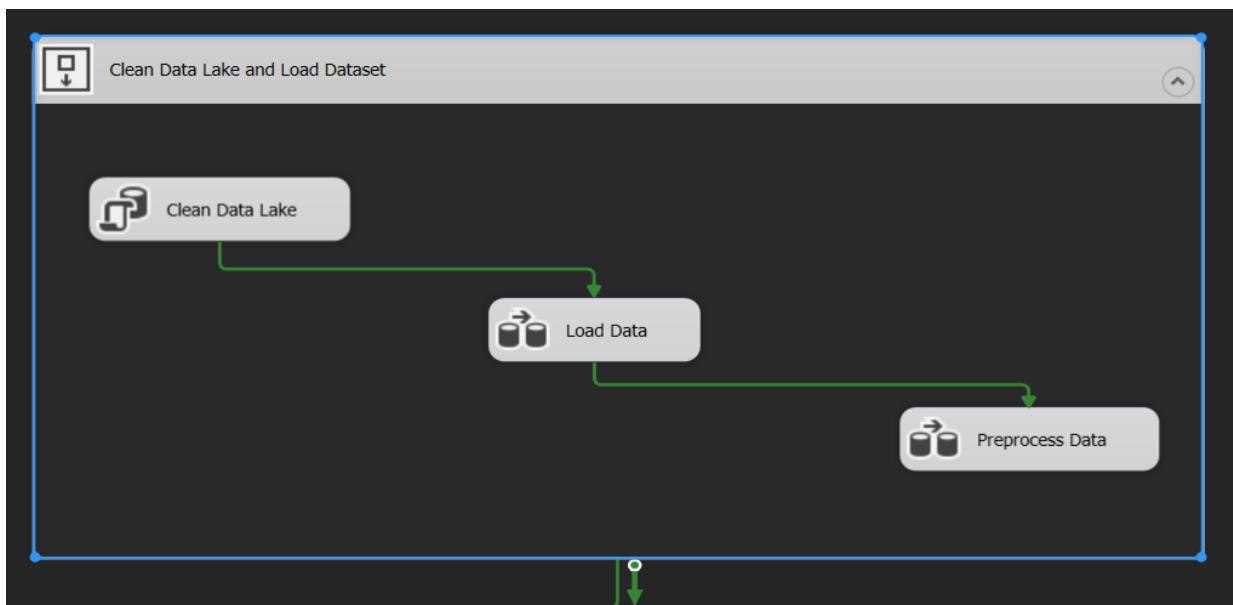


Kết quả thu được



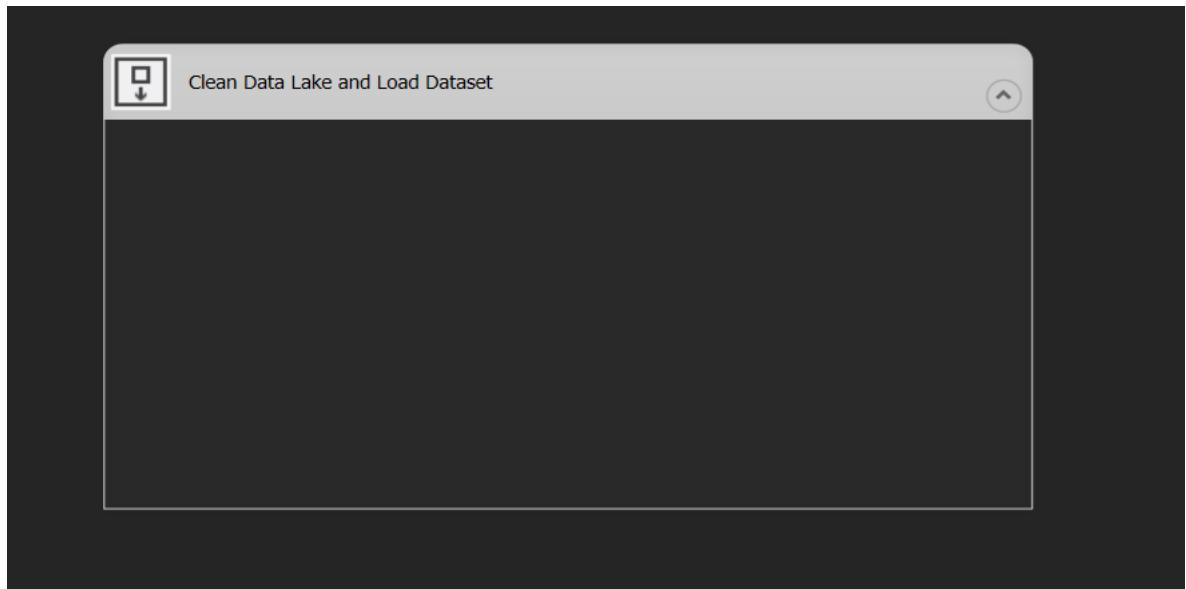
2.3 Làm sạch và nạp dữ liệu vào Data Lake

Mục tiêu: Xóa tất cả dữ liệu đã được insert vào 3 bảng (Original_Data, Clean_Data, Null_Data) trong Datalake_MarketingAnalytics, tiến hành nạp dữ liệu vào bảng Original_Data, làm sạch dữ liệu và nạp dữ liệu vào 2 bảng (Clean_Data, Null_Data) tương ứng

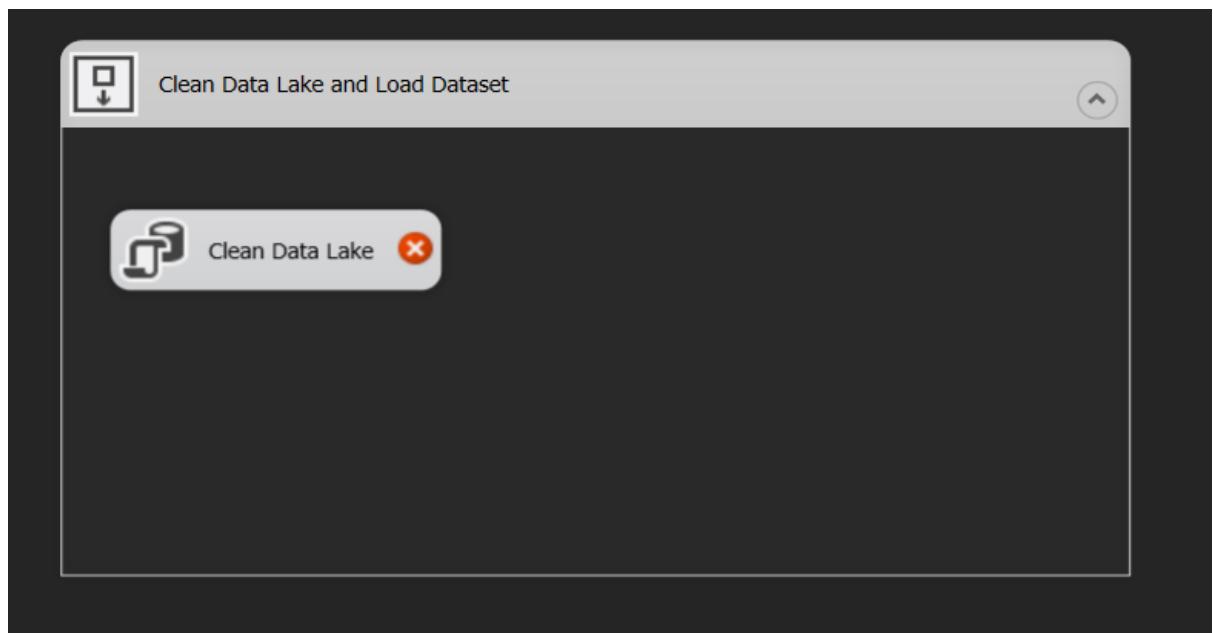


Tổng quan Sequence Container Clean Data Lake and Load Dataset

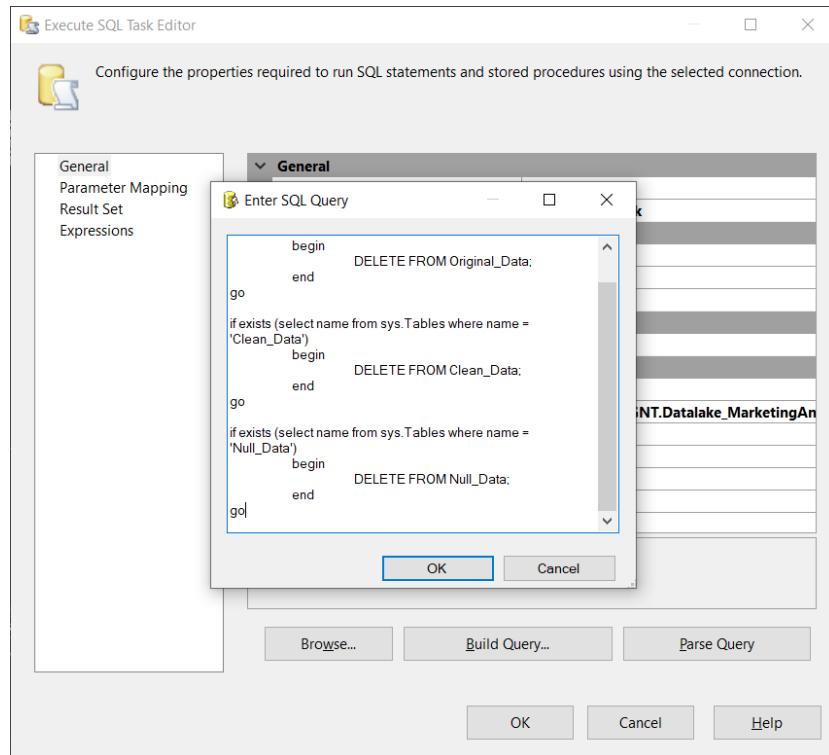
Bước 1: Kéo thả để tạo mới Sequence Container từ thanh công cụ SSIS Toolbox -> Đổi tên thành Clean Data Lake and Load Dataset



Bước 2: Kéo thả công cụ Execute SQL Task vào Container -> Đổi tên thành Clean Data Lake



Bước 3: Chuột phải vào Clean Data Lake -> Edit -> Định nghĩa Connection là tới DataLake_MarketingAnalytics -> Chính sửa SQLStatement -> OK



Chỉnh sửa connection và add SQLStatement

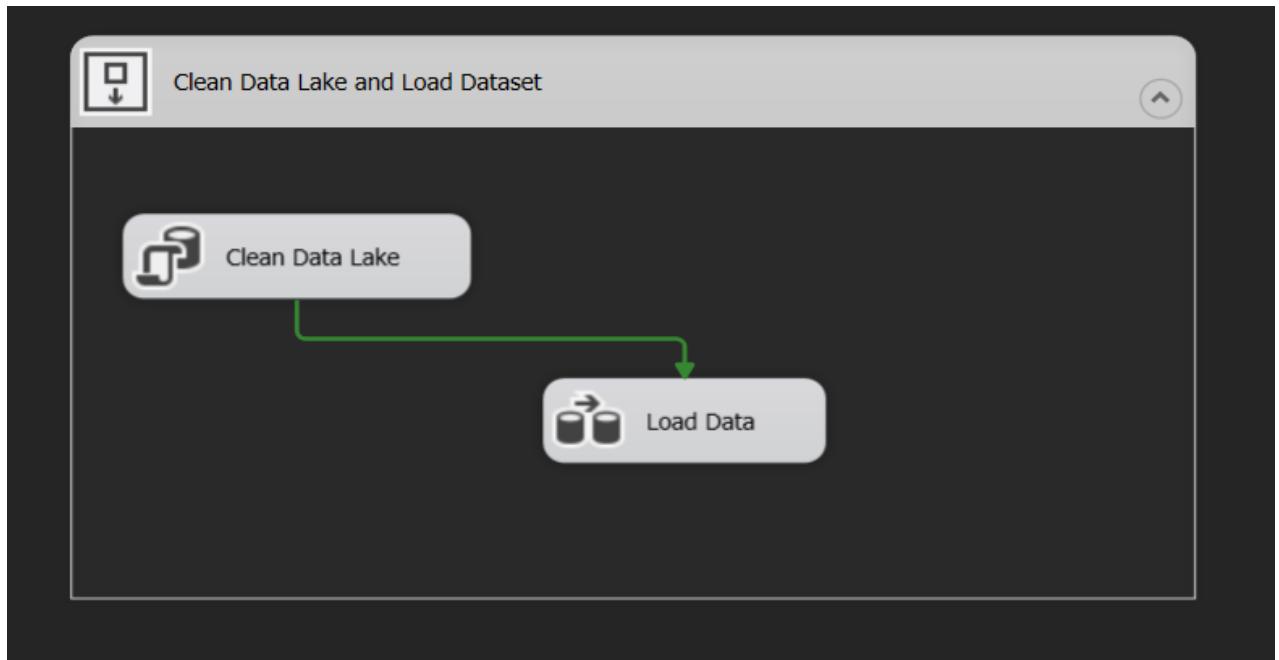
```
if exists (select name from sys.Tables where name = 'Original_Data')
begin
    DELETE FROM Original_Data;
end
go

if exists (select name from sys.Tables where name = 'Clean_Data')
begin
    DELETE FROM Clean_Data;
end
go

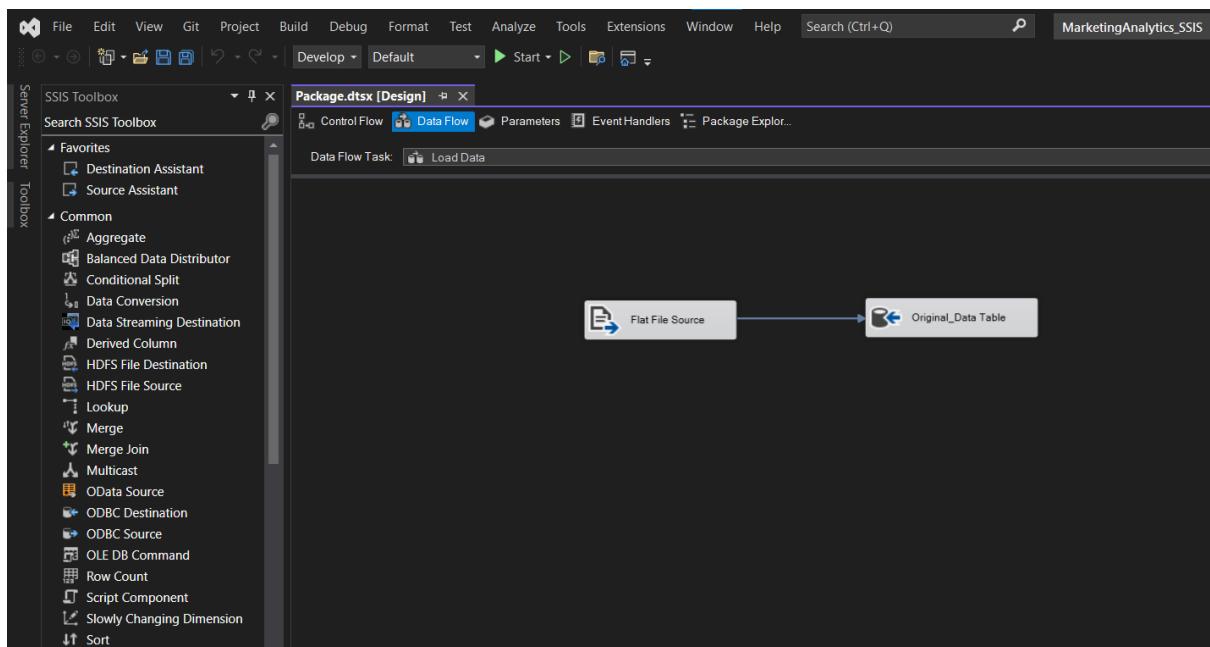
if exists (select name from sys.Tables where name = 'Null_Data')
begin
    DELETE FROM Null_Data;
end
go
```

Câu lệnh SQL để add vào SQLStatement, thực hiện việc clean các table trong DataLake

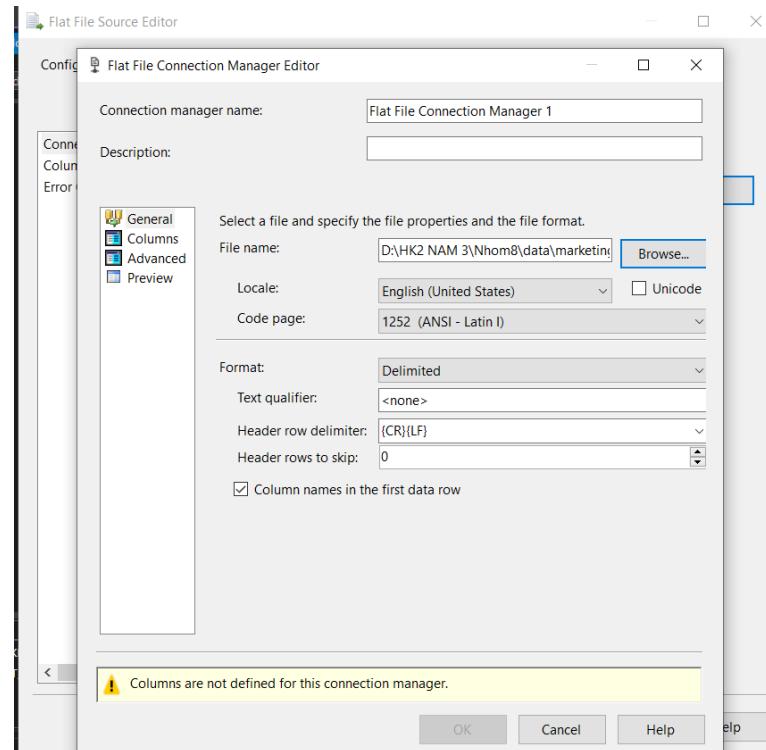
Bước 4: Kéo thả công cụ Data Flow Task vào Container -> Đổi tên thành Load Data



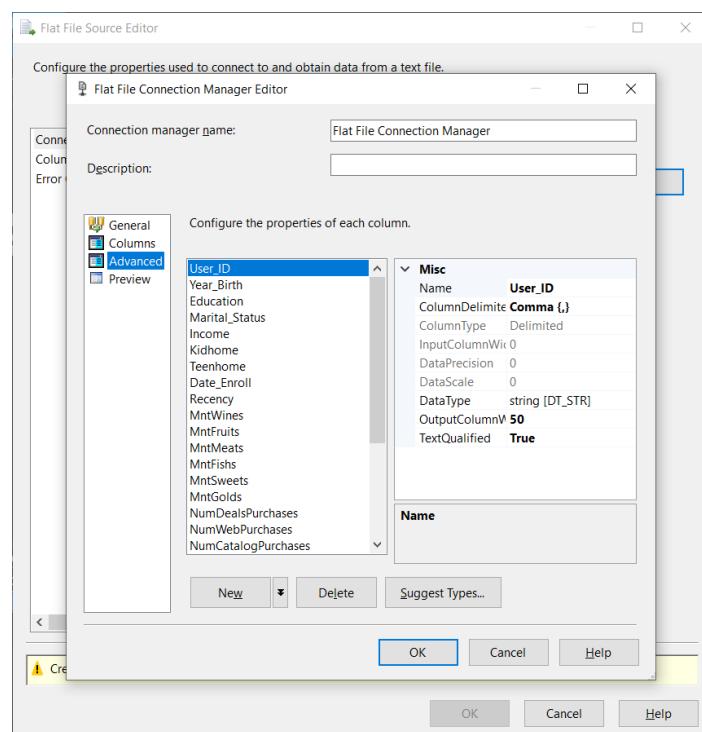
Bước 5: Click chuột trái vào Load Data để chỉnh sửa Data Flow Task -> Kéo thả công cụ Flat File Source để định nghĩa dữ liệu nguồn



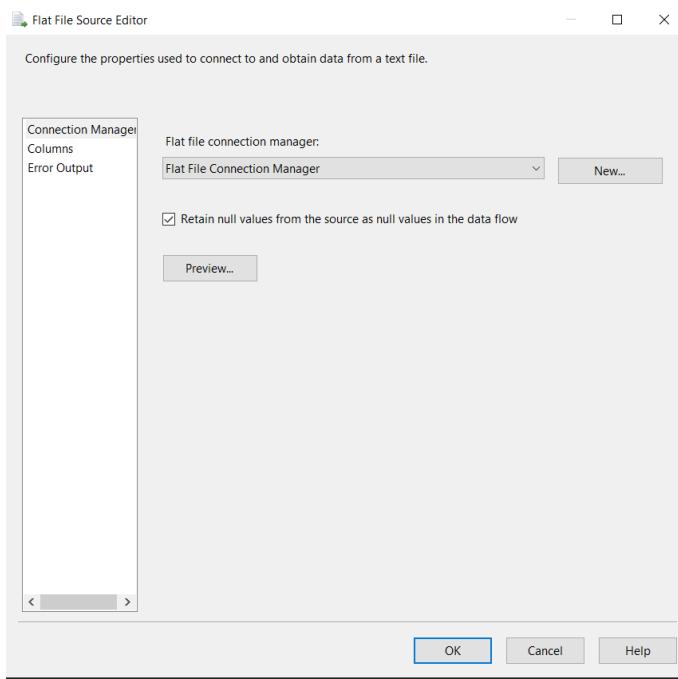
Bước 6: Click chuột phải và chọn Edit -> Chọn New -> Browse đến đường dẫn chứa file dữ liệu với định dạng csv



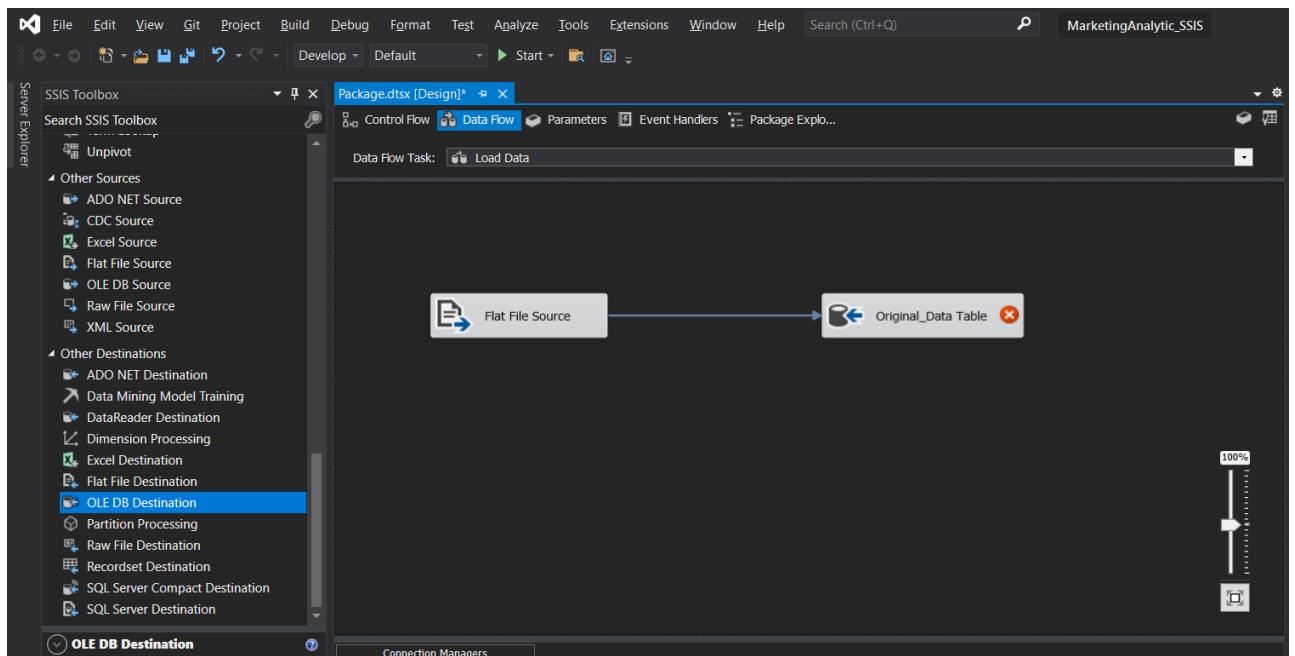
Bước 7: Chọn Columns -> Kiểm tra xem tình trạng dữ liệu đã đúng theo mong muốn hay chưa -> Chọn tiếp Advanced -> Chọn Suggest Types để Visual Studio tự động format lại kiểu dữ liệu của các cột -> OK



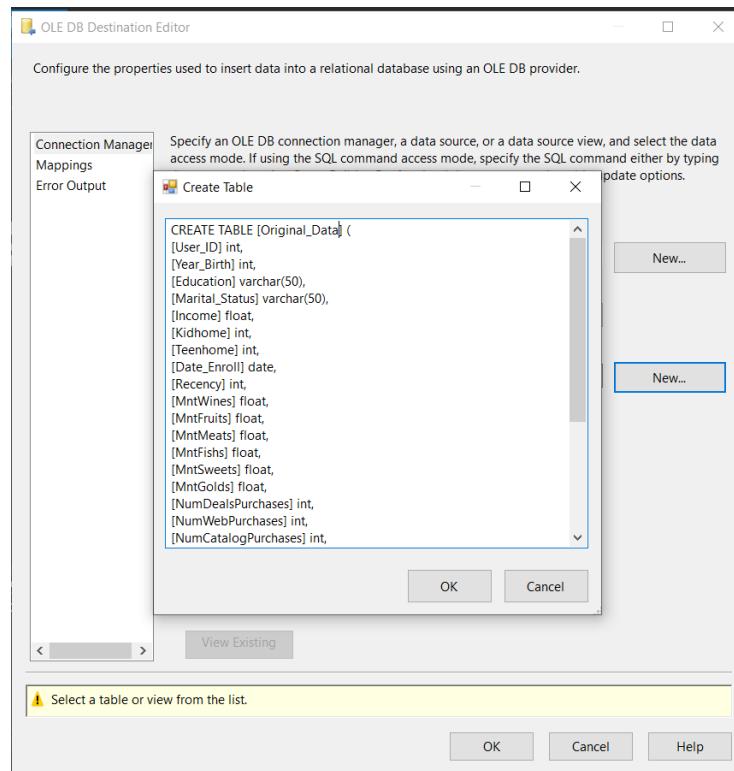
Bước 8: Click chọn Retain Null để giữ lại các dòng dữ liệu có giá trị Null cho các bước xử lý sau -> OK



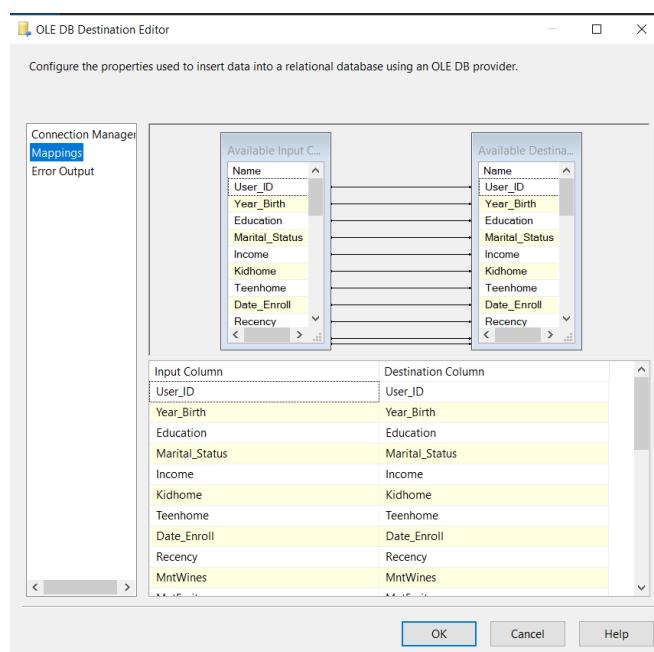
Bước 9: Kéo thả công cụ OLE DB Destination để đưa dữ liệu vào bảng Original_Data -> Đổi tên thành Original_Data Table



Bước 10: Click chuột phải vào Original_Data Table -> Chọn Edit -> Chọn connection đến Data Lake -> Chọn New để SSIS tự sinh ra đoạn mã insert data vào table -> Chính sửa data type của các cột thuộc tính theo đúng báo cáo ngay trên đoạn mã đã được tự động sinh ra -> Click chọn OK



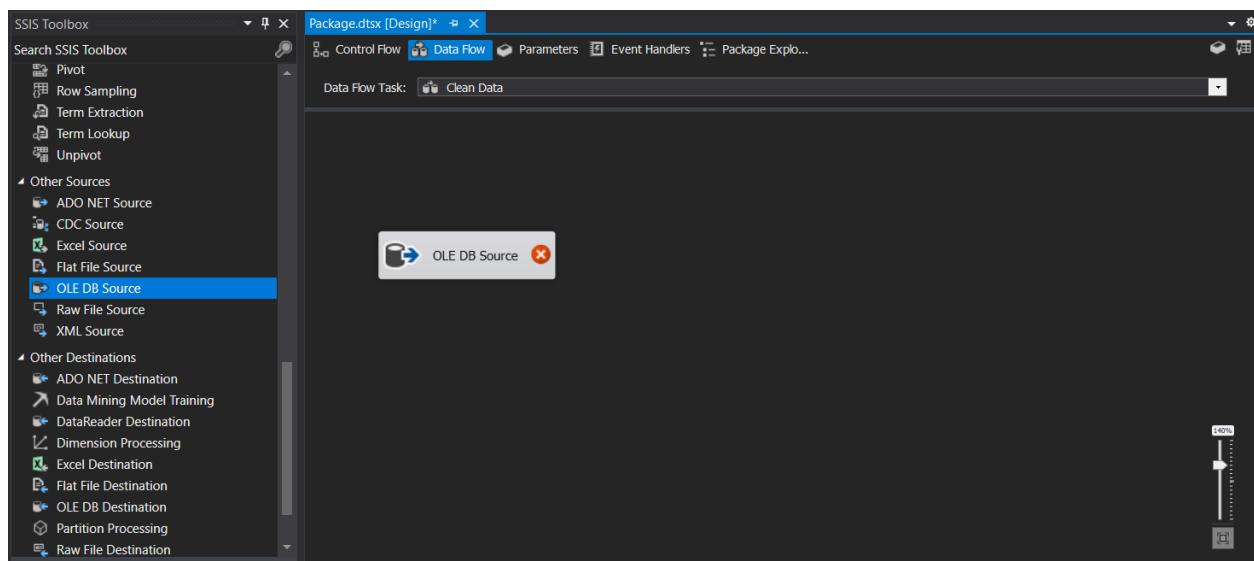
Bước 11: Click chọn Mappings -> Check xem những columns đã được map đúng hay chưa -> OK



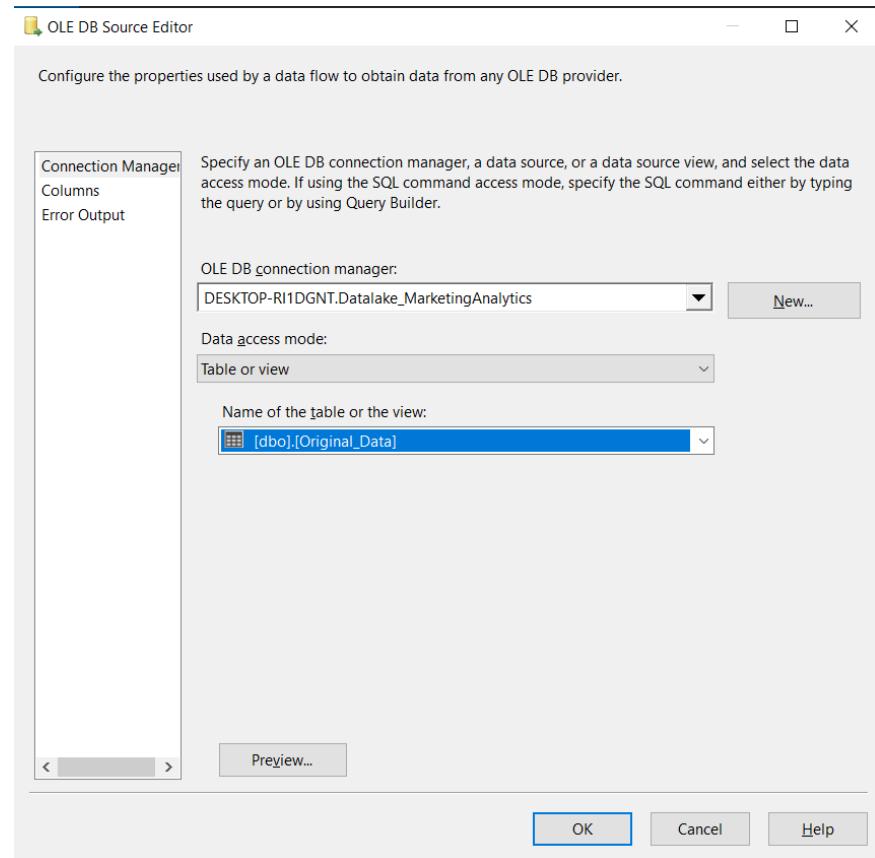
Bước 12: Kéo thả công cụ Data Flow Task vào Container -> Đổi tên thành Preprocess Data



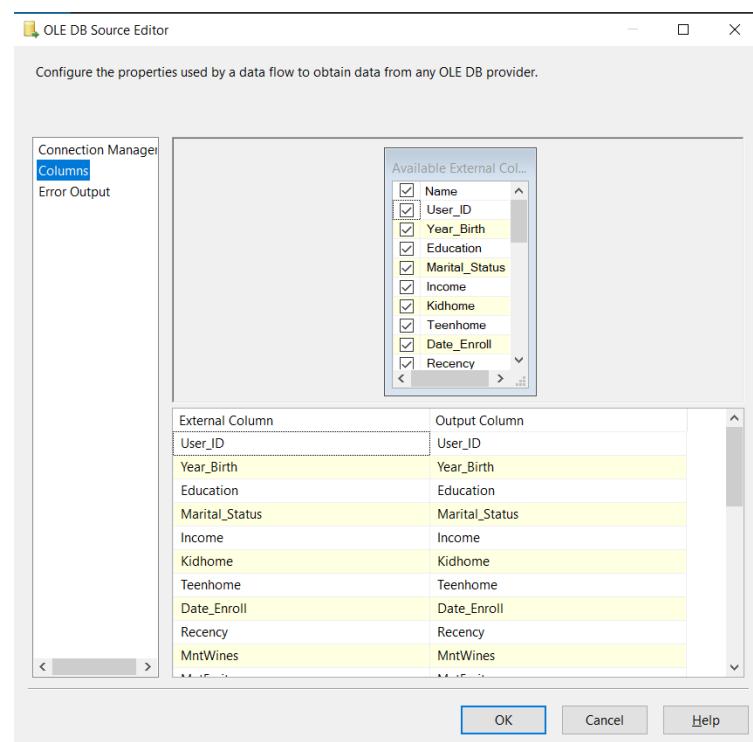
Bước 13: Click chuột trái vào Preprocess Data để chỉnh sửa Data Flow Task -> Kéo thả công cụ OLE DB Source để định nghĩa dữ liệu nguồn



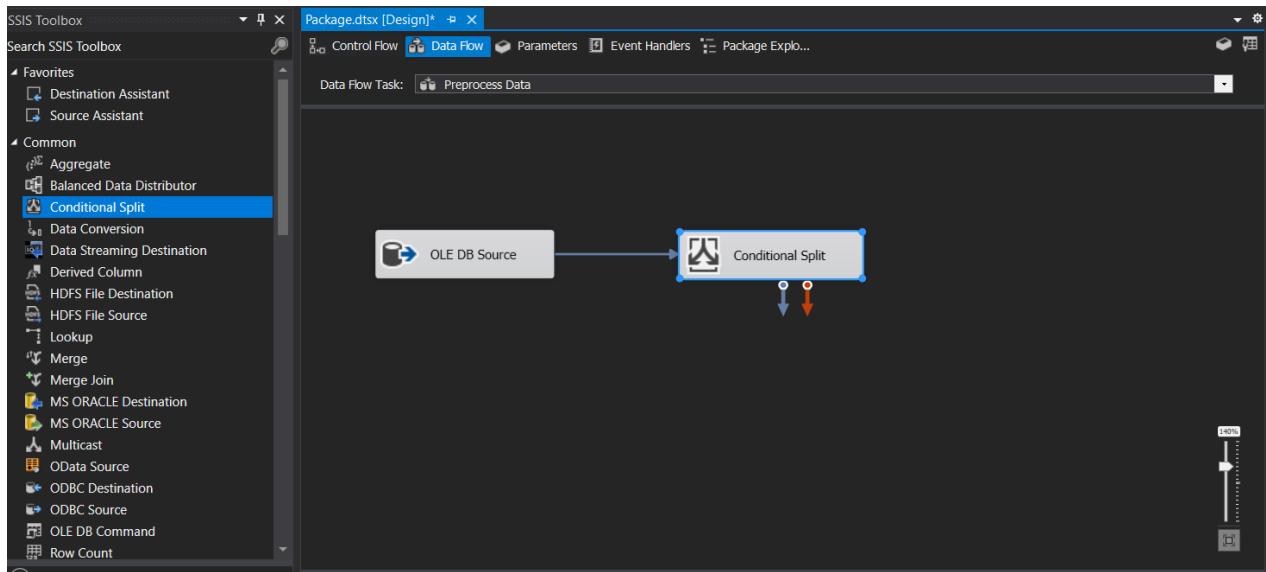
Bước 14: Click chuột phải và chọn Edit -> Chọn connection đến Data Lake -> Chọn table Original_Data



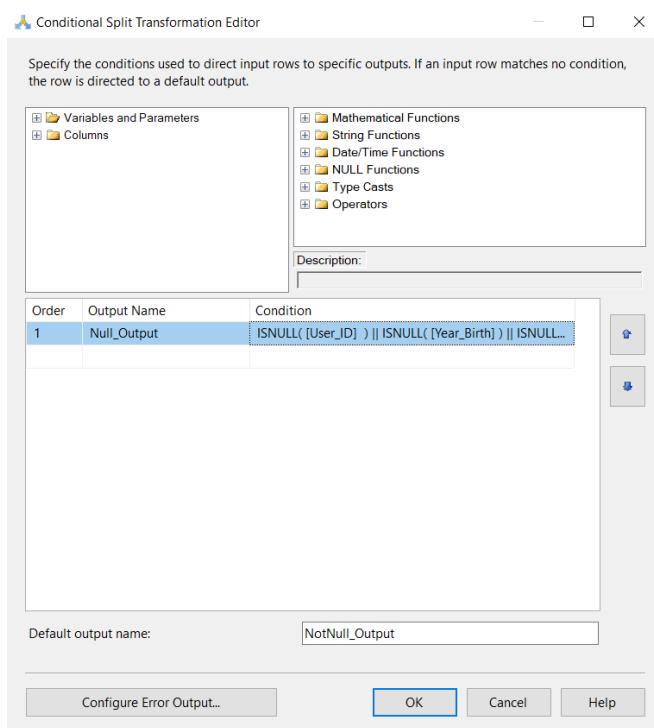
Bước 15: Click chọn Columns để lấy ra những cột cần thiết -> OK



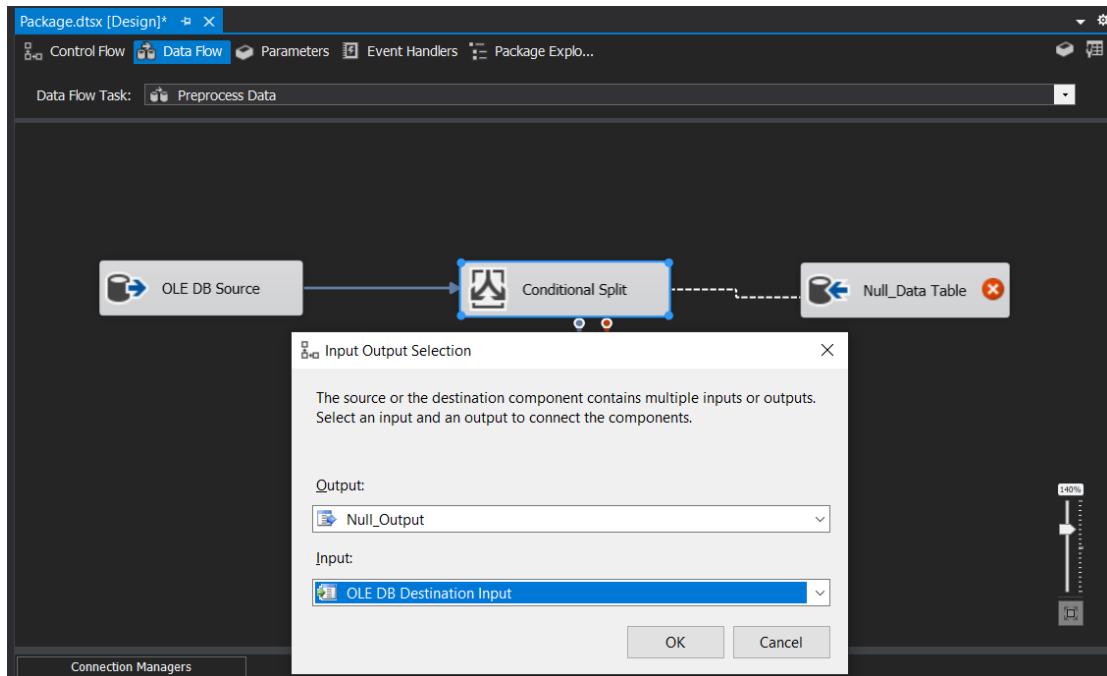
Bước 16: Kéo thả công cụ Conditional Split để lọc dữ liệu



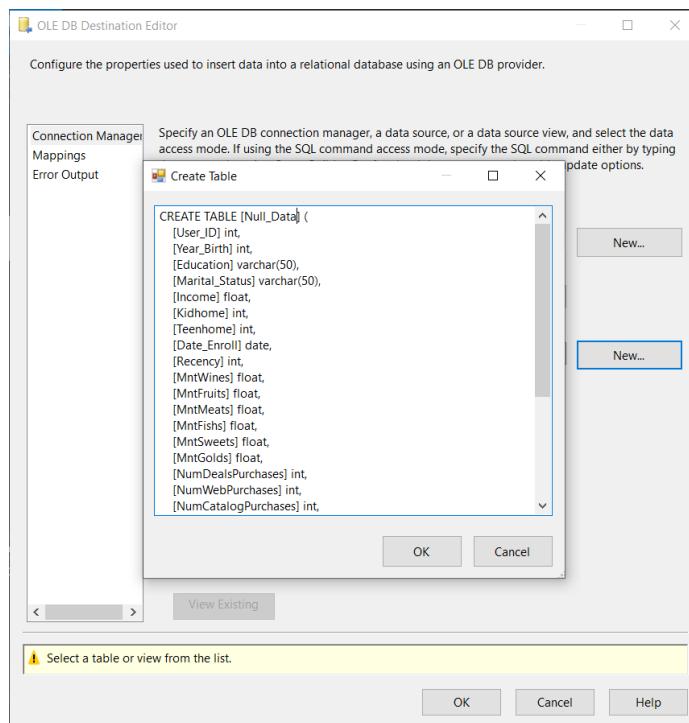
Bước 17: Click chuột phải vào Conditional Split -> Chọn Edit -> Chính sửa Output Name thành Null_Output -> Thêm điều kiện lọc Null trên toàn bộ cột -> Chính sửa Default Output Name thành NotNull_Output -> OK



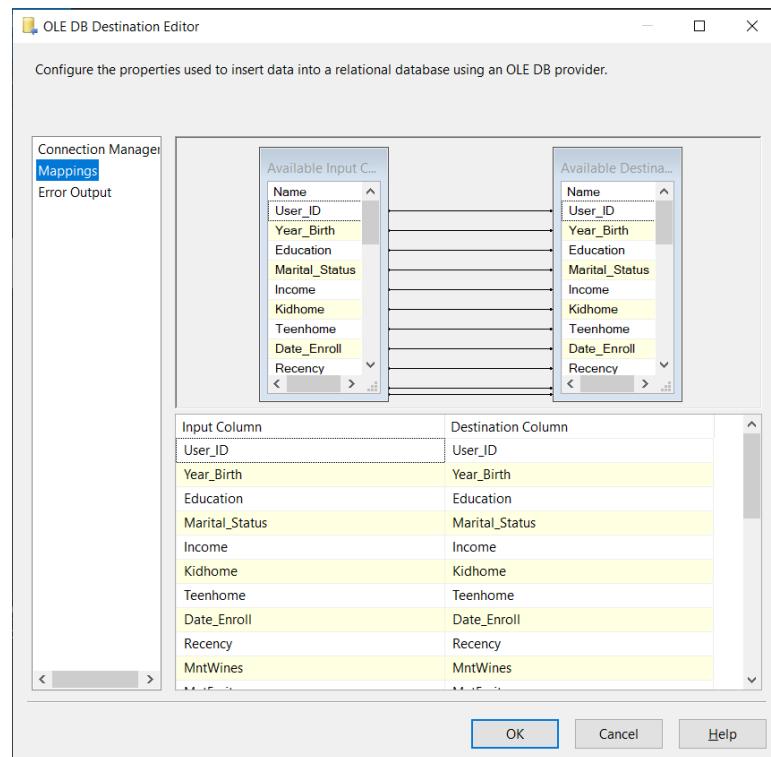
Bước 18: Kéo thả công cụ OLE DB Destination để định nghĩa bảng sẽ đưa dữ liệu Null vào -> Đổi tên thành Null_Data Table -> Nối Data Flow -> Chọn Null_Output -> OK



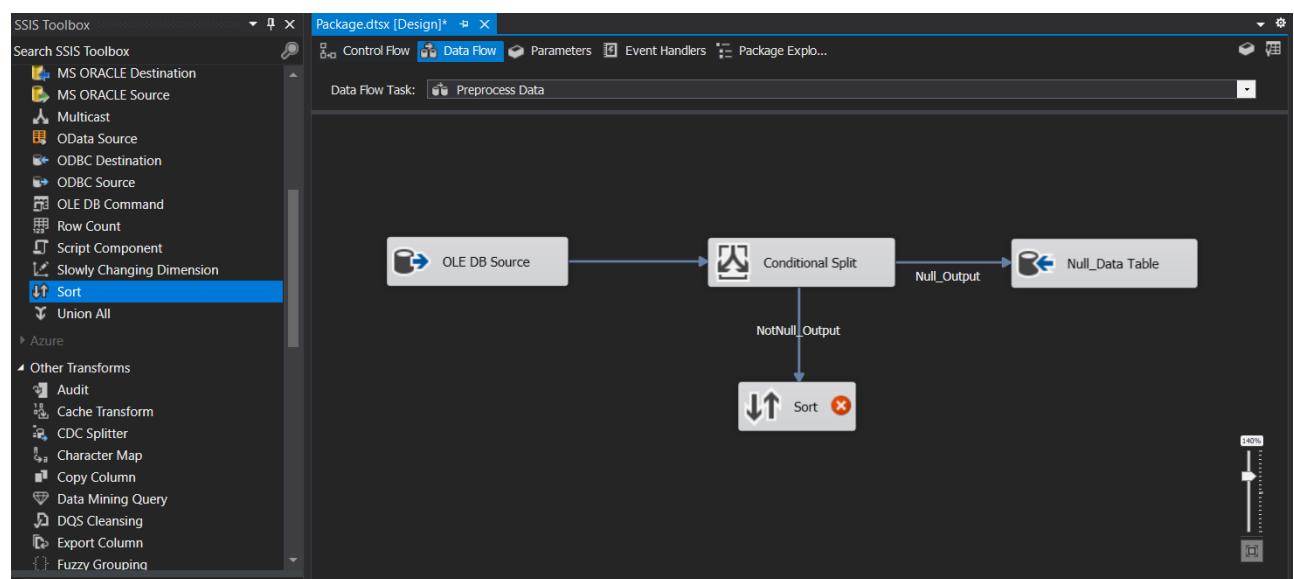
Bước 19: Click chuột phải vào Null_Data Table -> Chọn Edit -> Chọn connection đến Data Lake -> Chọn New để SSIS tự sinh ra đoạn mã insert data vào table -> Click chọn OK



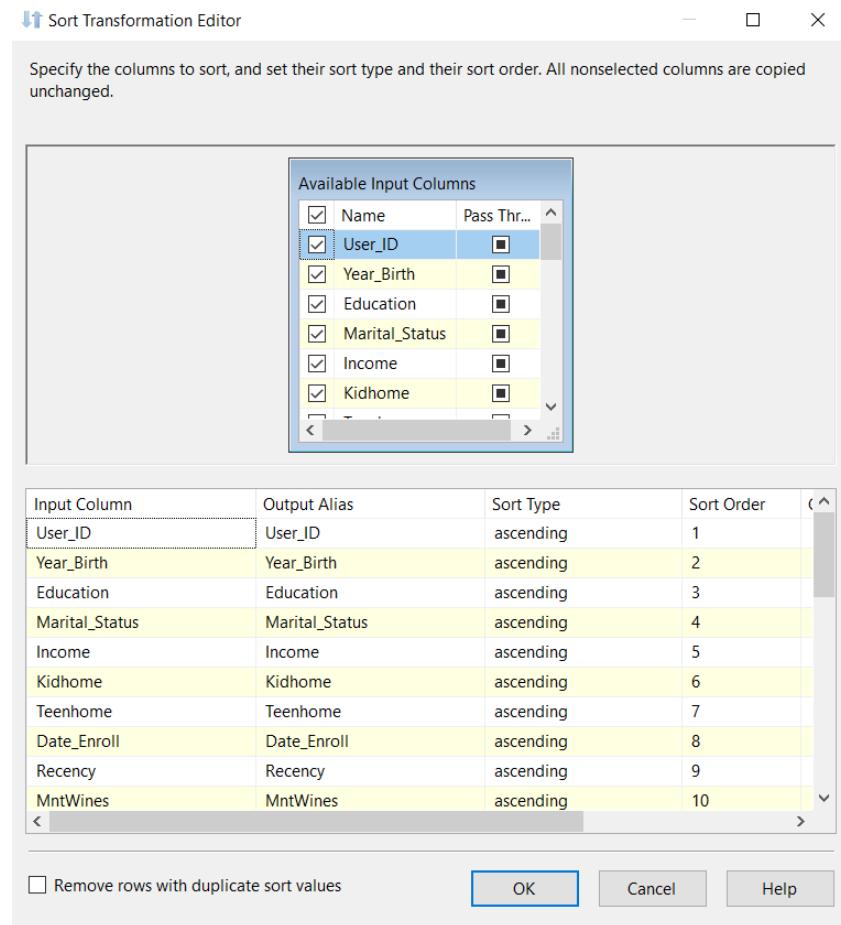
Bước 20: Click chọn Mappings -> Check xem những columns đã được map đúng hay chưa -> OK



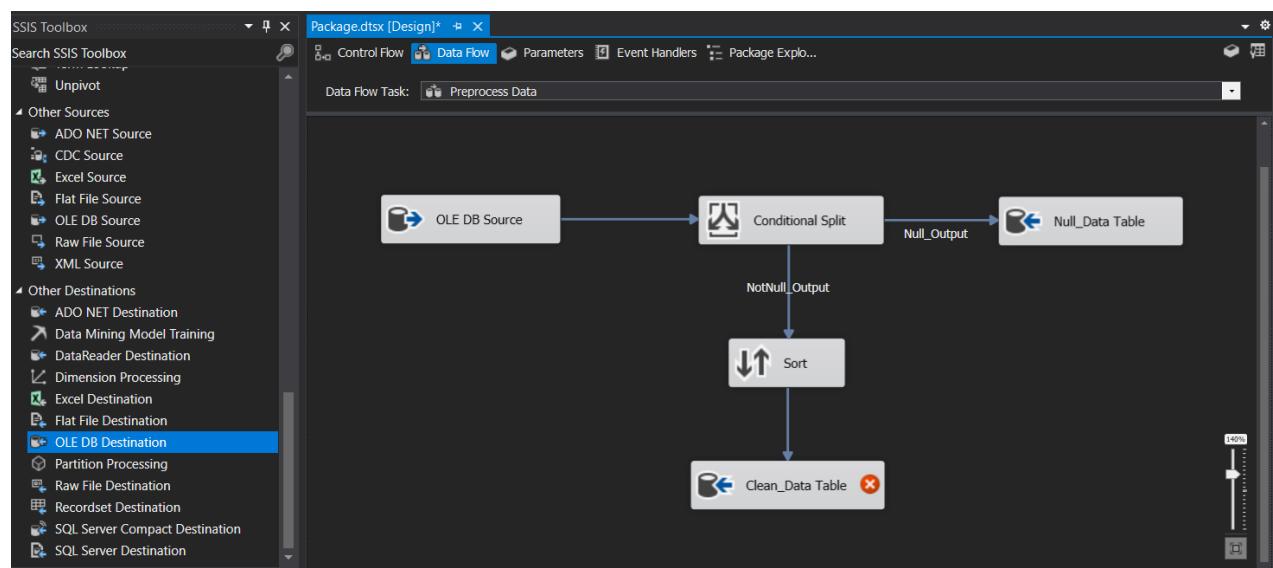
Bước 21: Kéo thả công cụ Sort để sắp xếp lại các dòng dữ liệu



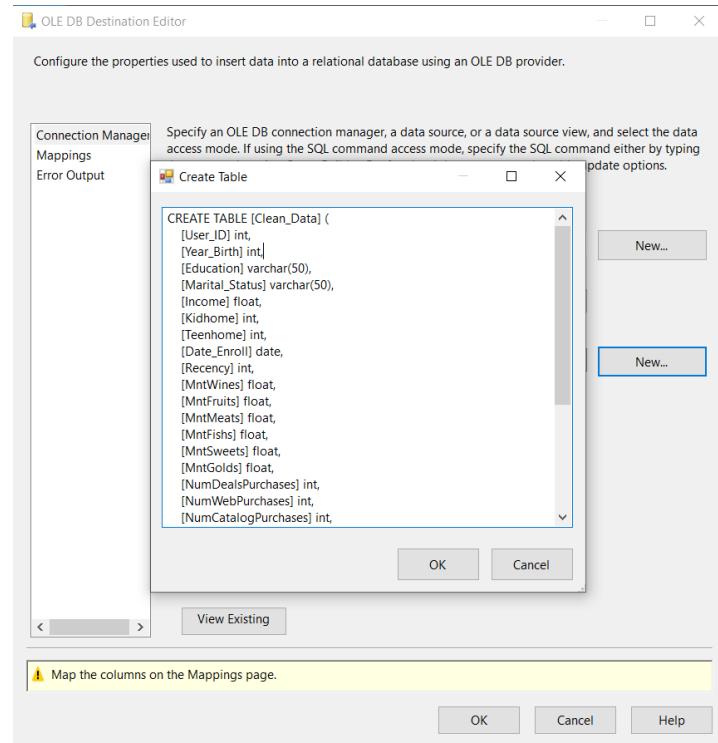
Bước 22: Click chuột phải vào Sort -> Chọn Edit -> Click chọn sort theo tất cả các cột và phương thức là Ascending -> OK



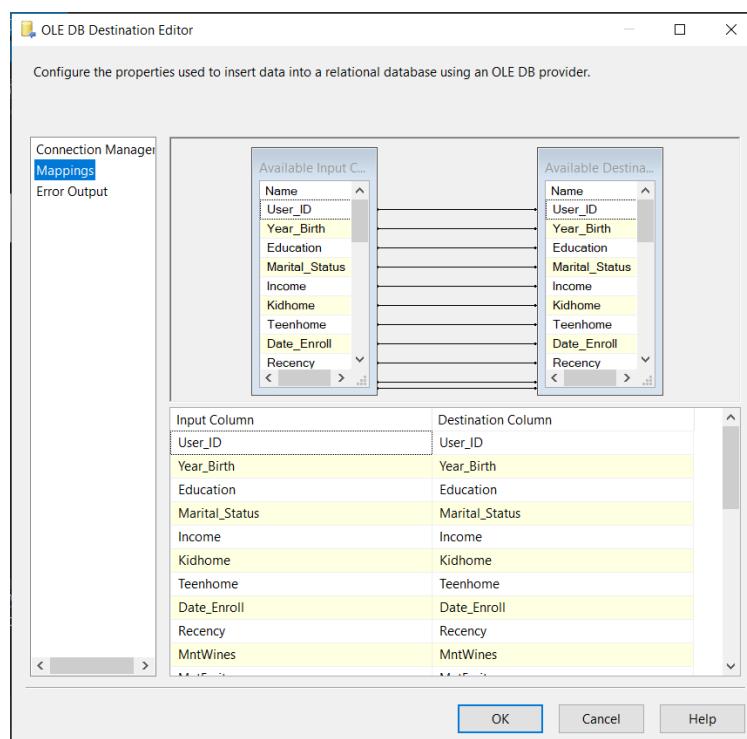
Bước 23: Kéo thả công cụ OLE DB Destination để định nghĩa bảng sẽ đưa dữ liệu đã Clean vào -> Đổi tên thành Clean_Data Table



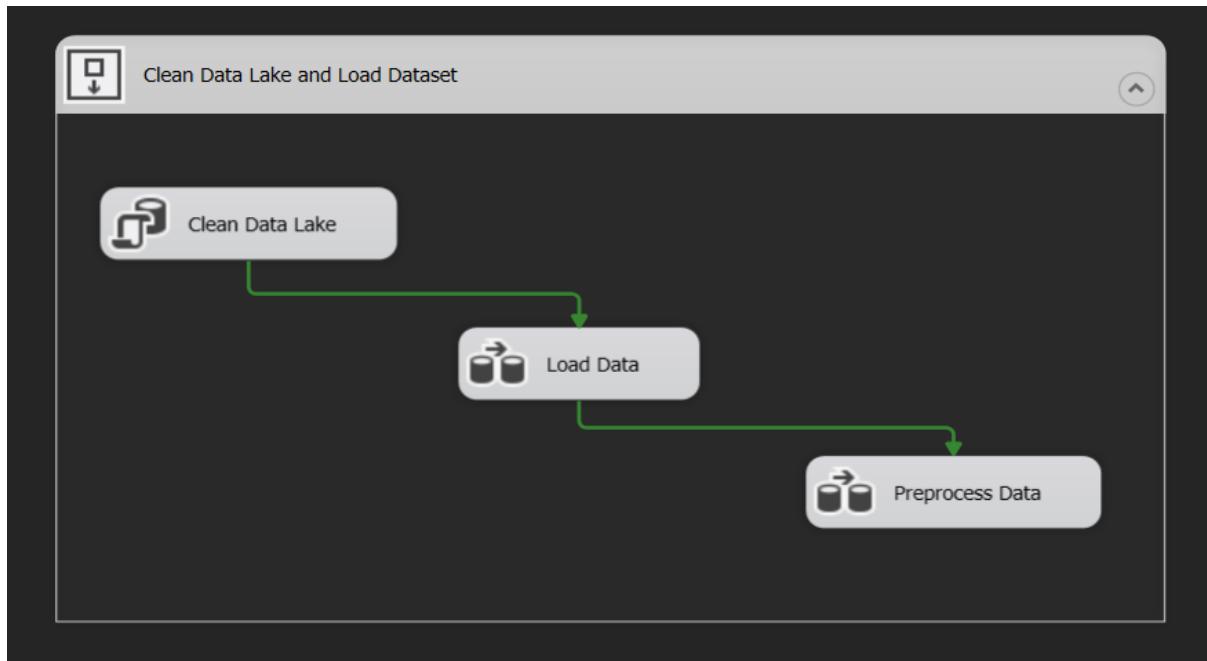
Bước 24: Click chuột phải vào Clean_Data Table -> Chọn Edit -> Chọn connection đến Data Lake -> Chọn New để SSIS tự sinh ra đoạn mã insert data vào table -> Click chọn OK



Bước 25: Click chọn Mappings -> Check xem những columns đã được map đúng hay chưa -> OK



Kết quả thu được sau khi hoàn thành định nghĩa các Data Flow Task cho Sequence Container Clean Data Lake and Load Dataset



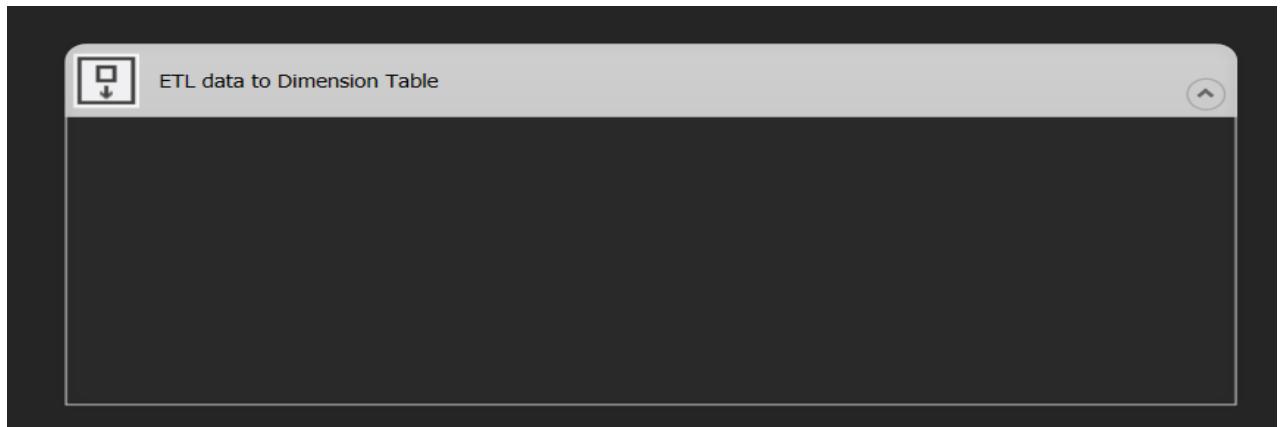
2.4 Định nghĩa và nạp dữ liệu vào các bảng Dimension

Mục tiêu: Nạp dữ liệu đã được tiền xử lý nằm trong bảng Clean_Data của Data Lake vào các bảng Dimension tương ứng theo đúng thiết kế lược đồ kho dữ liệu

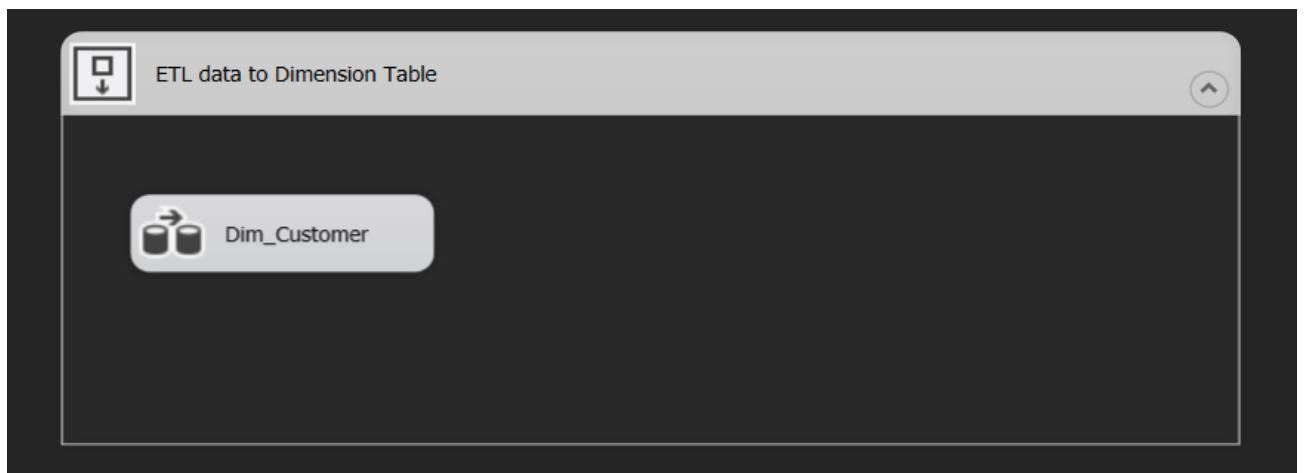


Tổng quan Sequence Container ETL data to Dimension Table

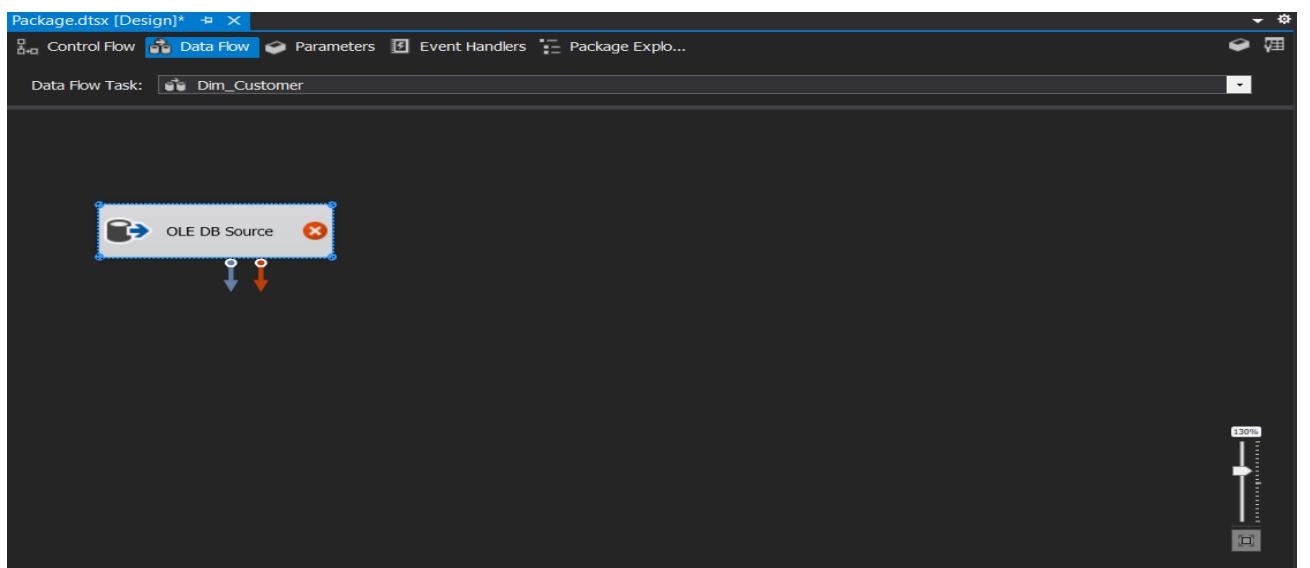
Bước 1: Kéo thả để tạo mới Sequence Container từ thanh công cụ SSIS Toolbox -> Đổi tên thành ETL data to Dimentzion Table



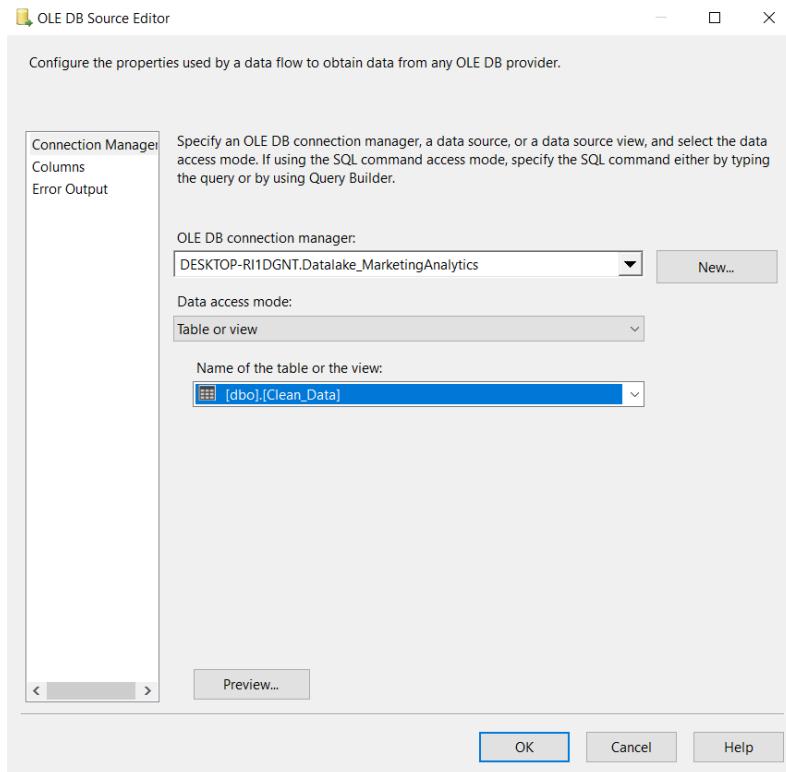
Bước 2: Kéo thả công cụ Data Flow Task vào Container -> Đổi tên thành Dim_Customer



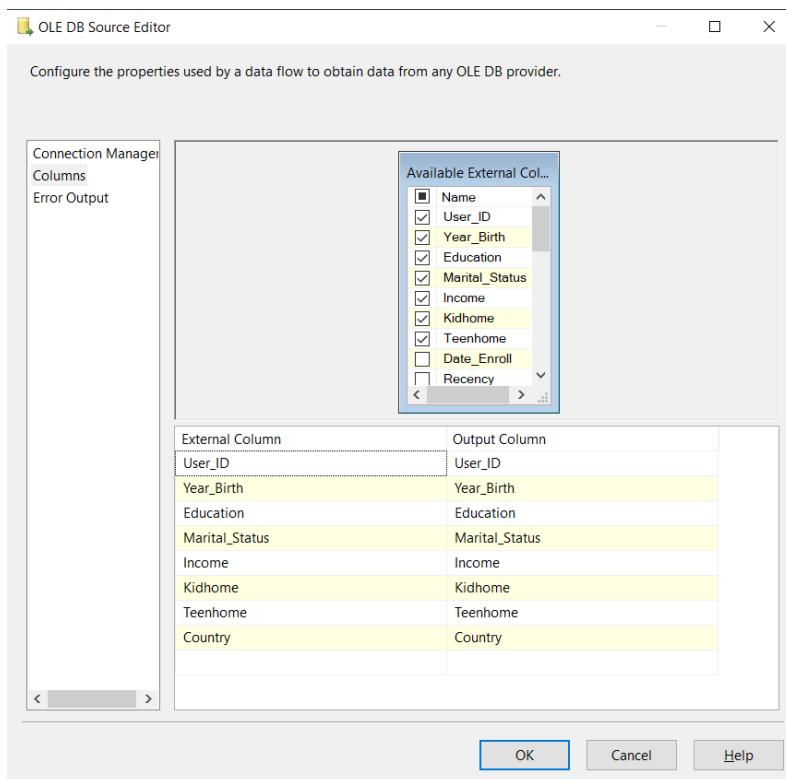
Bước 3: Click chuột trái vào Dim_Customer để chỉnh sửa Data Flow Task -> Kéo thả công cụ OLE DB Source để định nghĩa dữ liệu nguồn



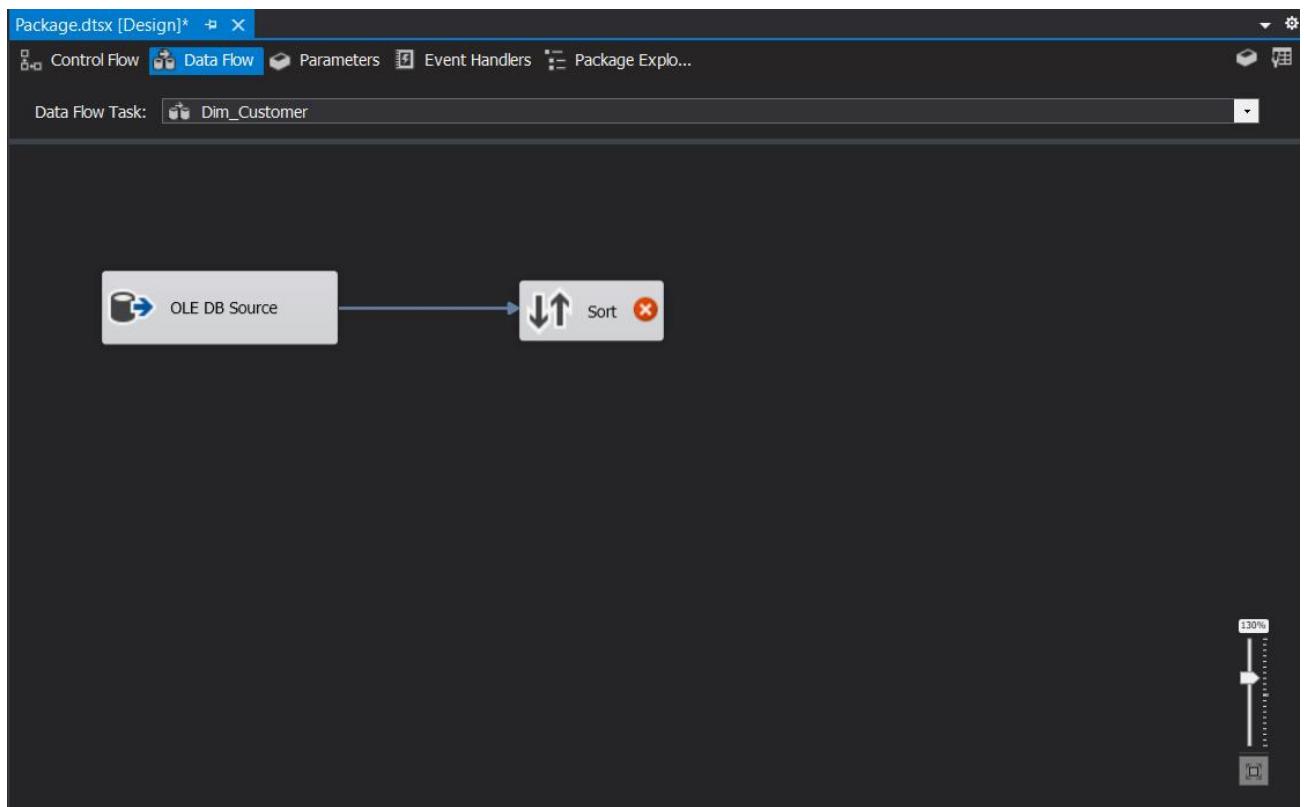
Bước 4: Click chuột phải và chọn Edit -> Chọn connection đến Data Lake -> Chọn table Clean_Data



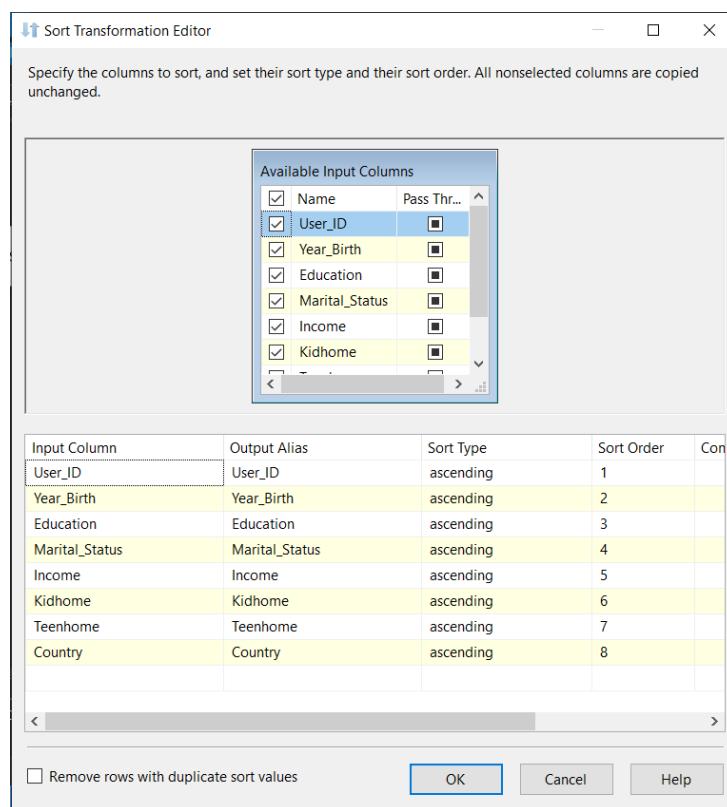
Bước 5: Click chọn Columns để lấy ra những column cần thiết cho bảng Dim_Customer -> OK



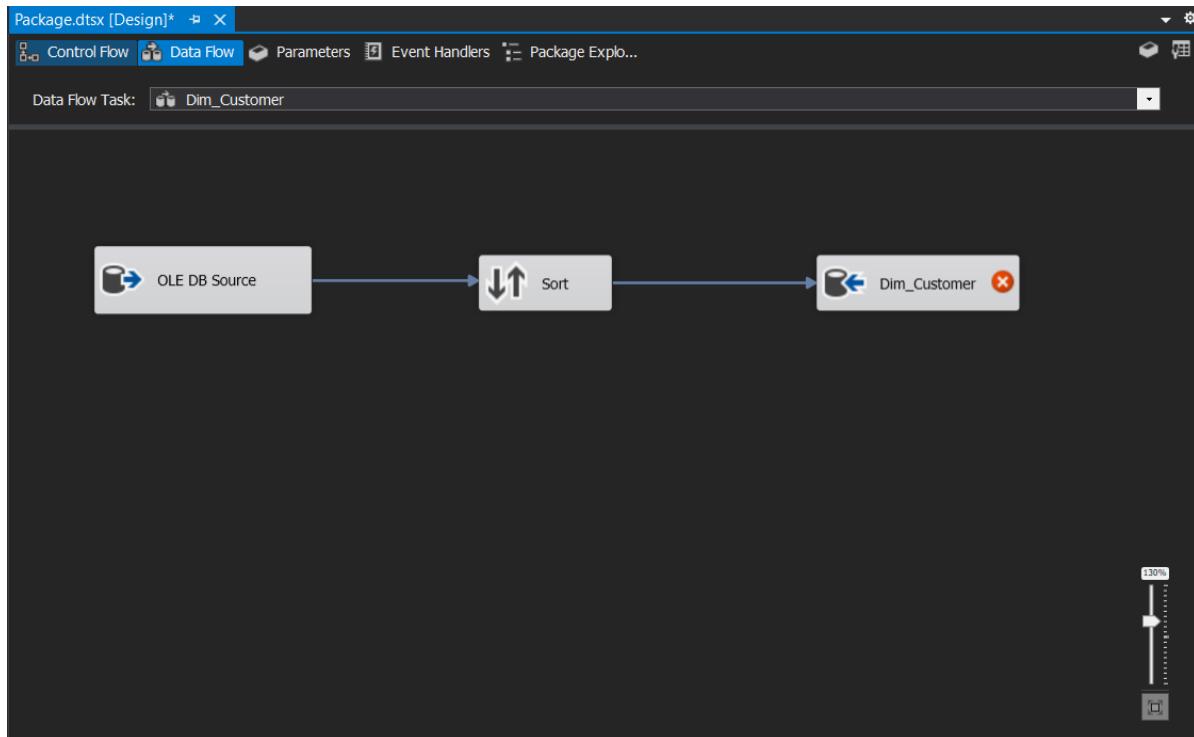
Bước 6: Kéo thả công cụ Sort để sắp xếp lại các dòng dữ liệu



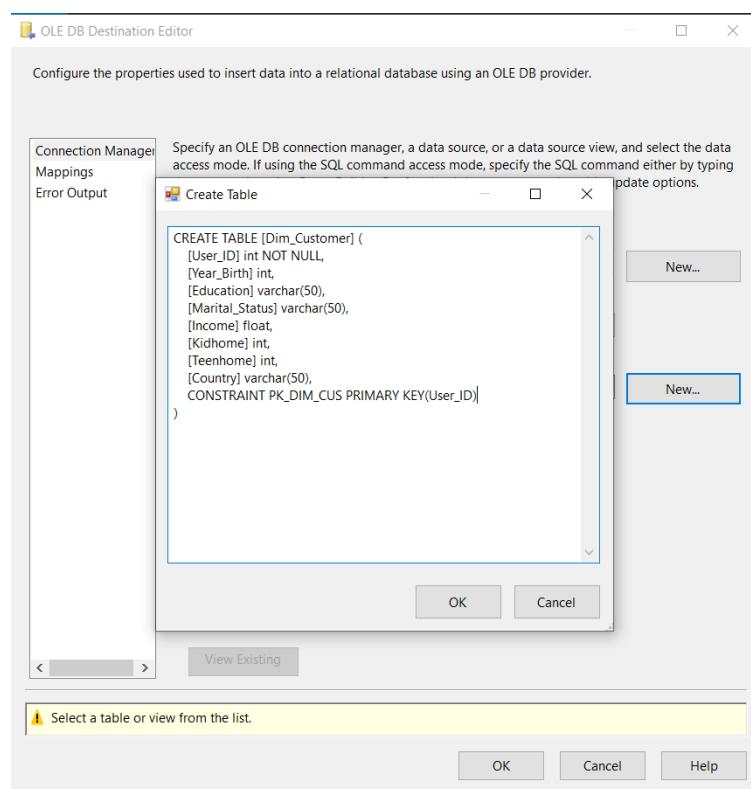
Bước 7: Click chuột phải vào Sort -> Chọn Edit -> Click chọn sort theo tất cả các cột và phương thức là Ascending -> OK



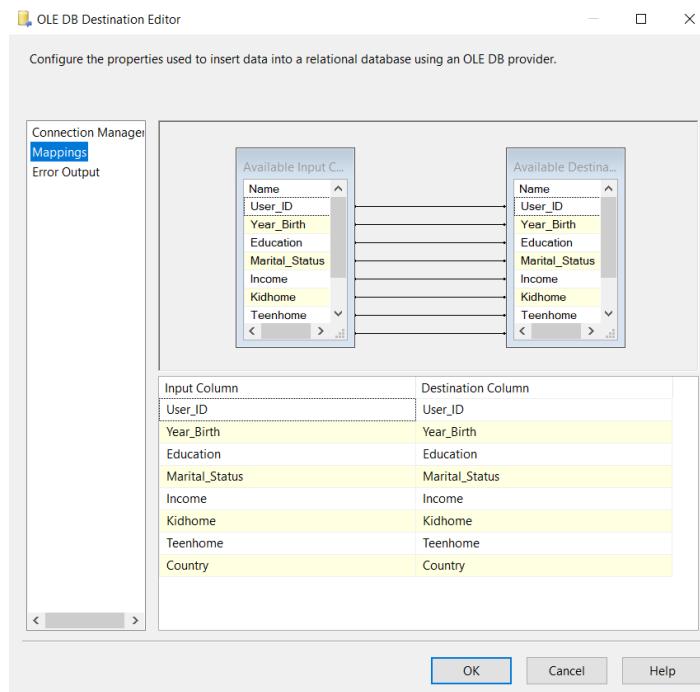
Bước 8: Kéo thả công cụ OLE DB Destination để định nghĩa Dim_Customer-> Đổi tên thành Dim_Customer



Bước 9: Click chuột phải vào Dim_Customer -> Chọn Edit -> Chọn connection đến Data Warehouse -> Chọn New để SSIS tự sinh ra đoạn mã insert data vào table -> Thêm primary key ngay trên đoạn mã được sinh tự động -> OK



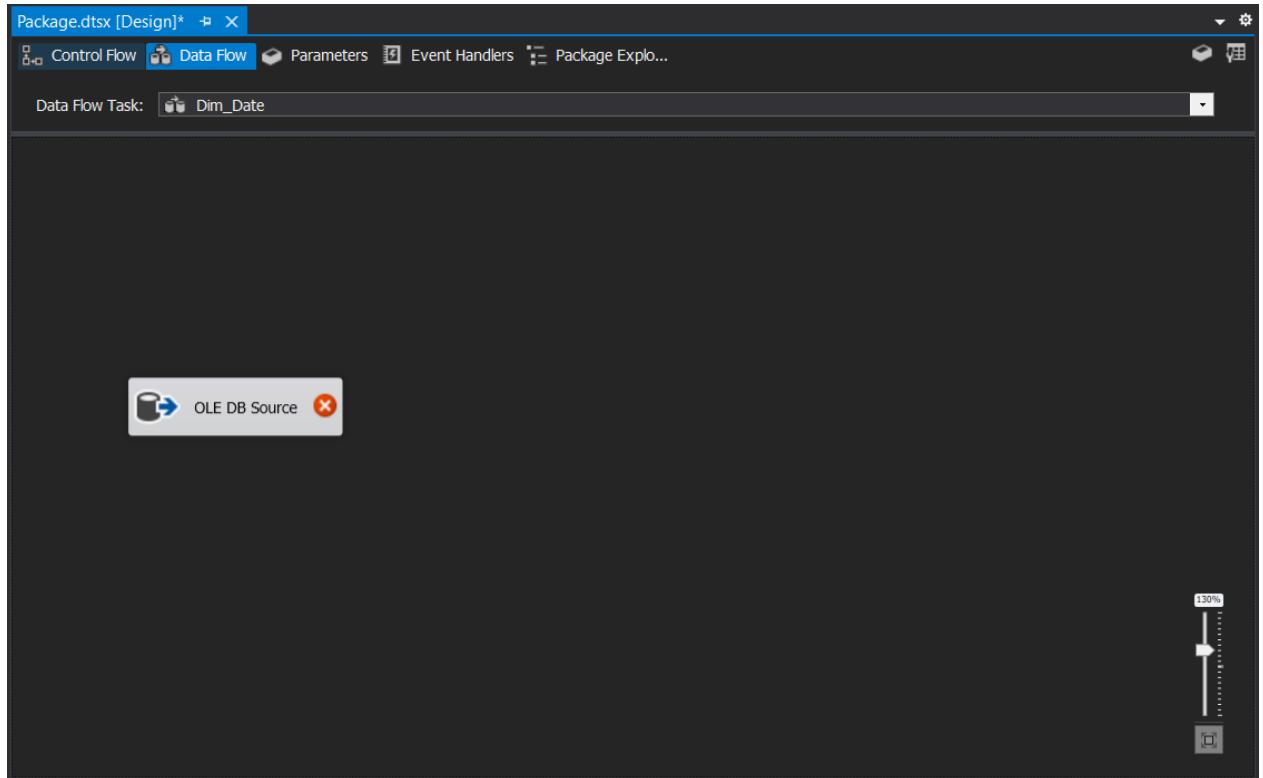
Bước 10: Click chọn Mappings -> Check xem những columns đã được map đúng hay chưa -> OK



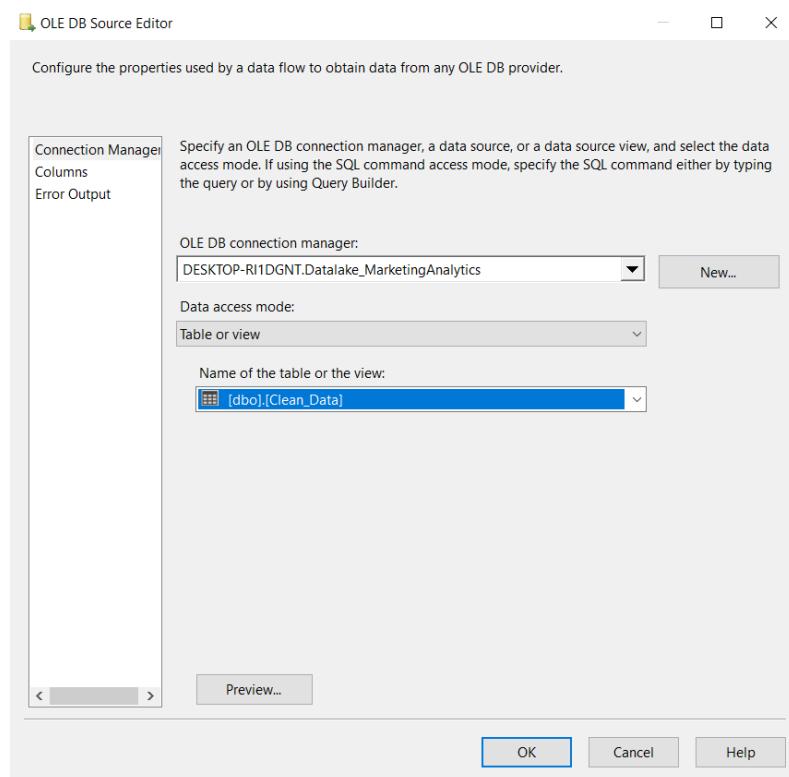
Bước 11: Kéo thả công cụ Data Flow Task vào Container -> Đổi tên thành Dim_Date



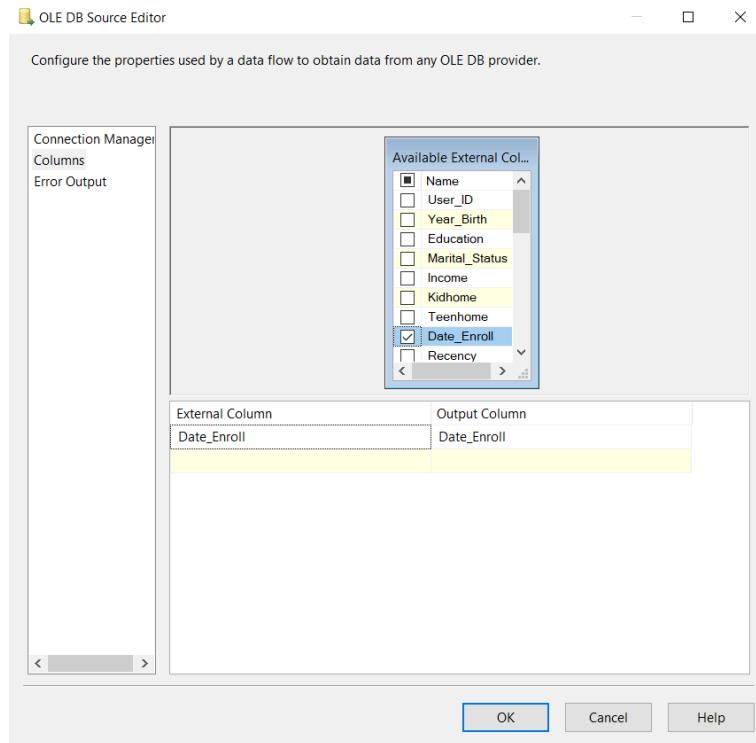
Bước 12: Click chuột trái vào Dim_Date để chỉnh sửa Data Flow Task -> Kéo thả công cụ OLE DB Source để định nghĩa dữ liệu nguồn



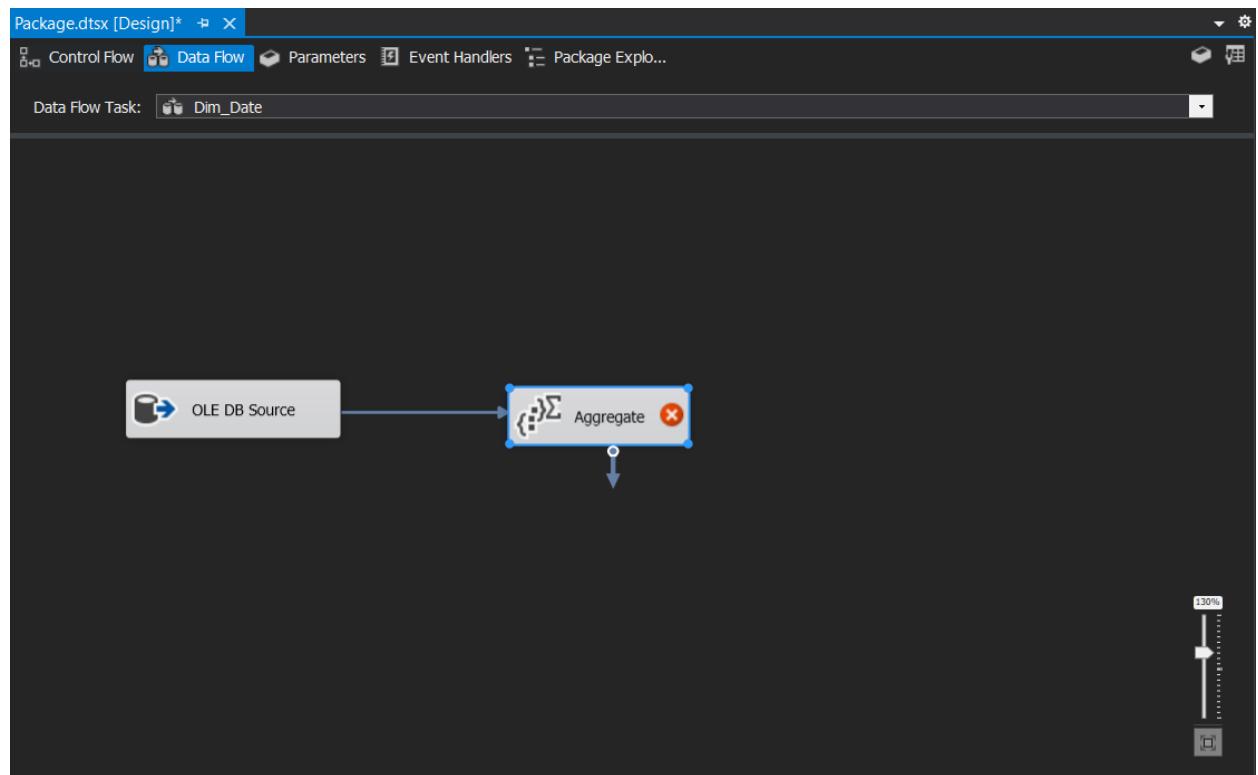
Bước 13: Click chuột phải và chọn Edit -> Chọn connection đến Data Lake -> Chọn table Clean_Data



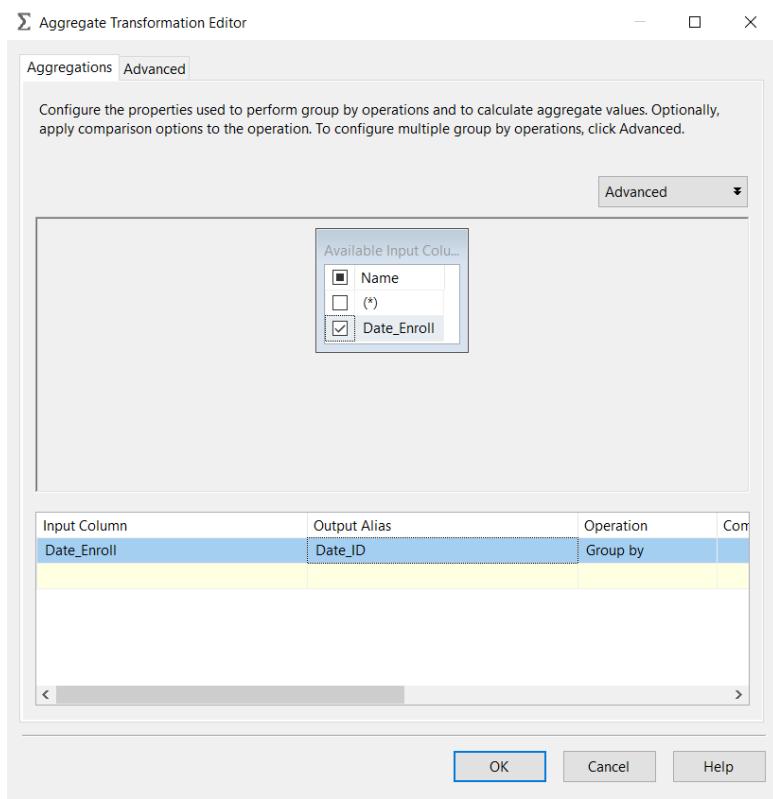
Bước 14: Click chọn Columns để lấy ra những column cần thiết cho bảng Dim_Date -> OK



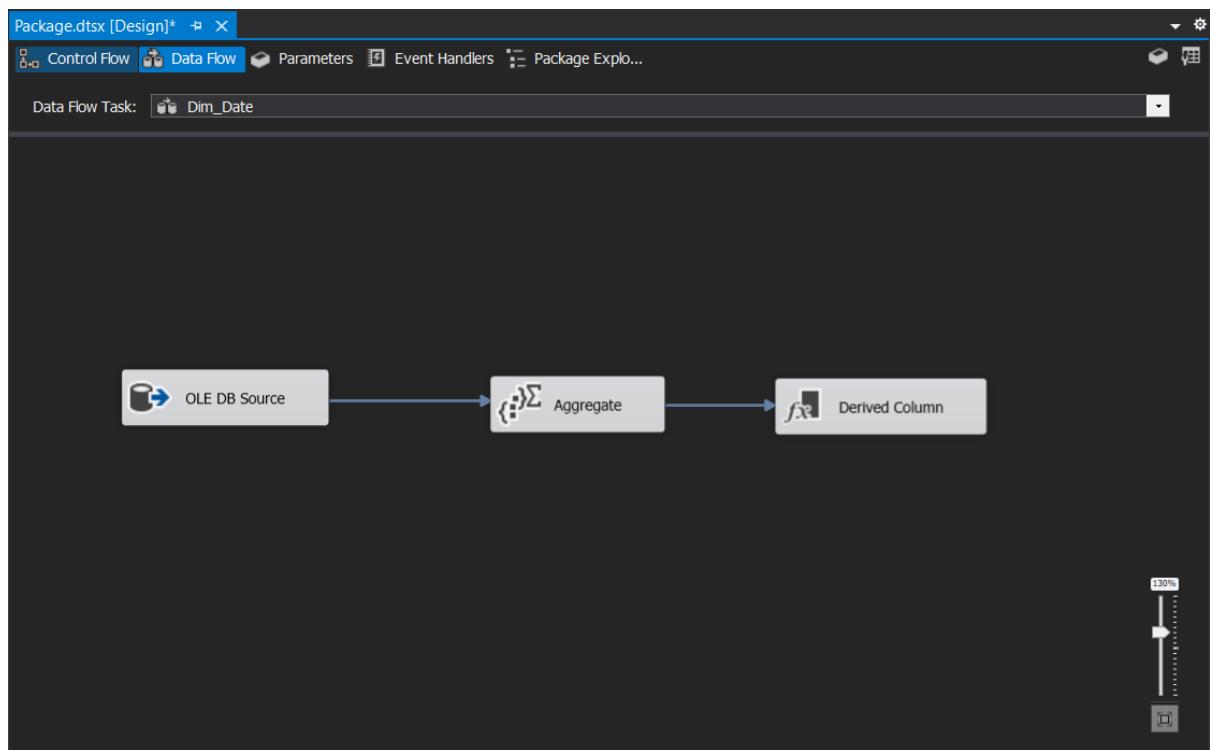
Bước 15: Kéo thả công cụ Aggregate để Group by dữ liệu, xóa những dữ liệu bị trùng



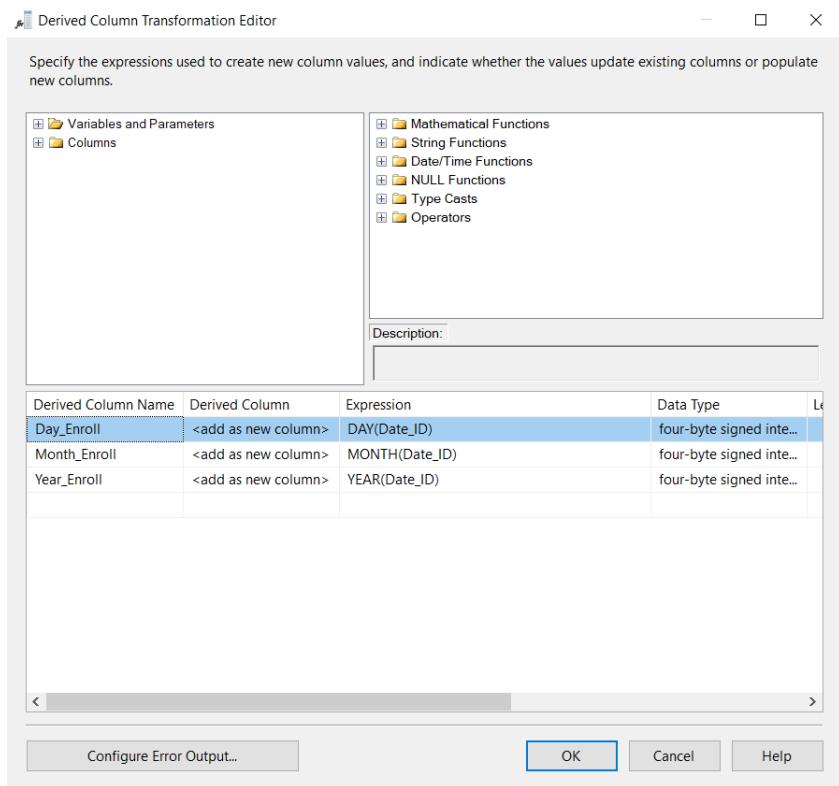
Bước 16: Click chuột phải vào Aggerate -> Chọn Edit -> Click chọn group by theo cột Date_Enroll -> Đổi tên Output Alias thành Date_ID theo đúng thiết kế kho dữ liệu -> OK



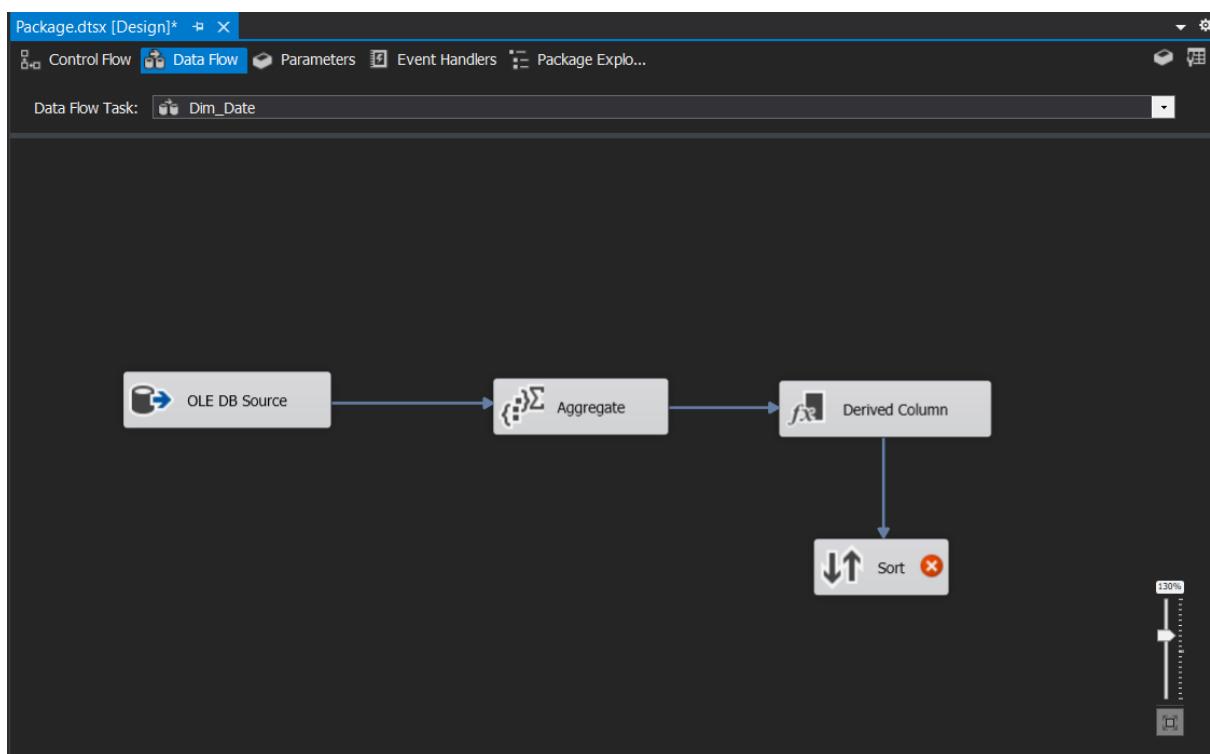
Bước 17: Kéo thả công cụ Derived Column để tách dữ liệu Date_Enroll



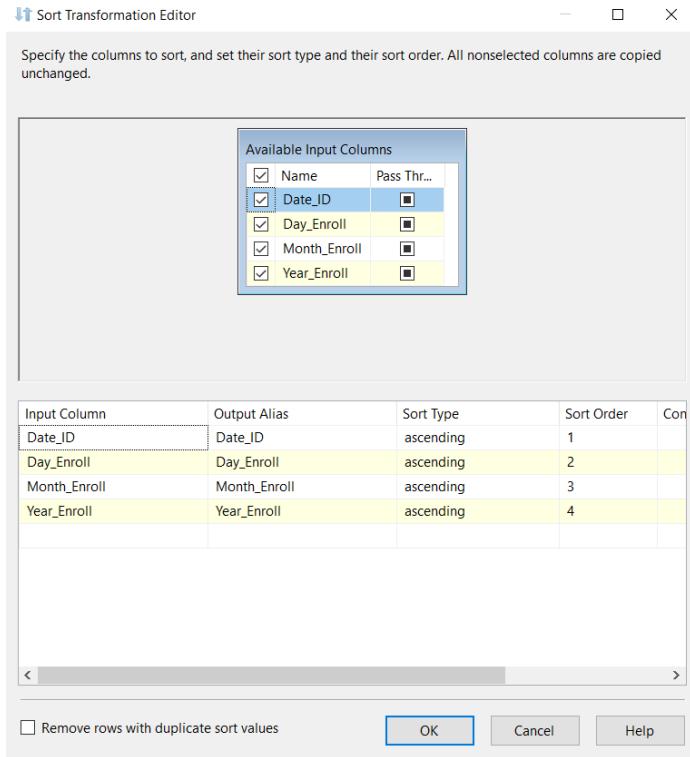
Bước 18: Click chuột phải vào Derived Column -> Chọn Edit -> Chính sửa Derived Column Name theo thiết kế kho dữ liệu tương ứng-> Thêm các function cần thiết để split Date_Enroll -> OK



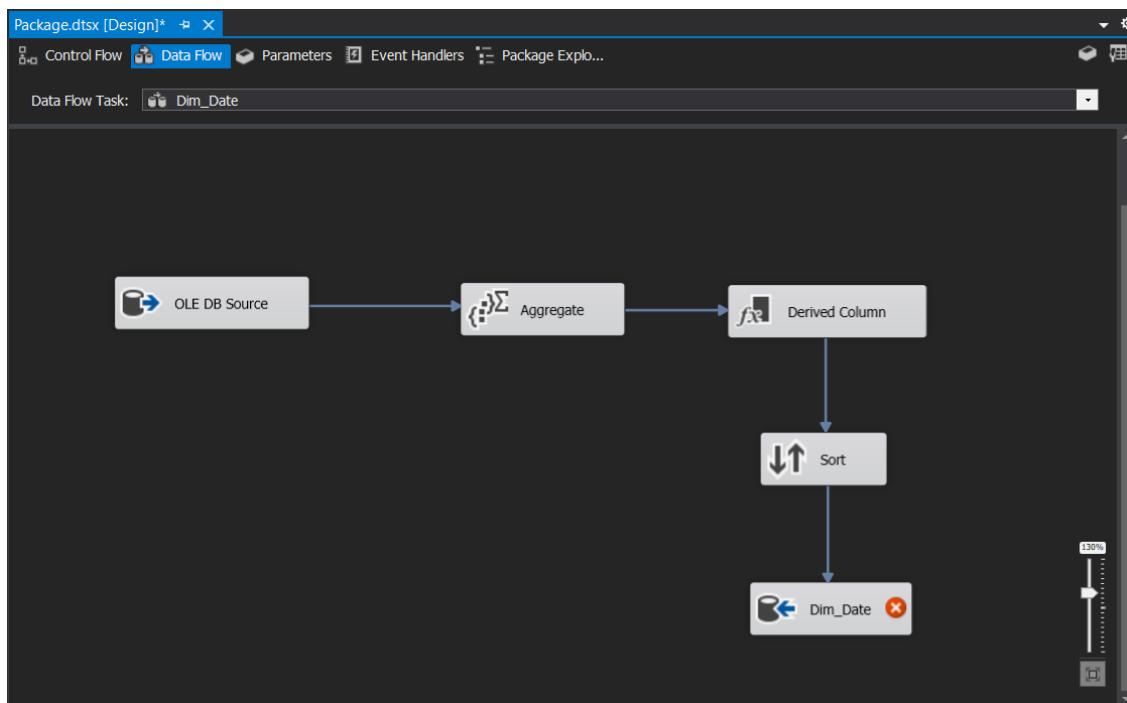
Bước 19: Kéo thả công cụ Sort để sắp xếp lại các dòng dữ liệu



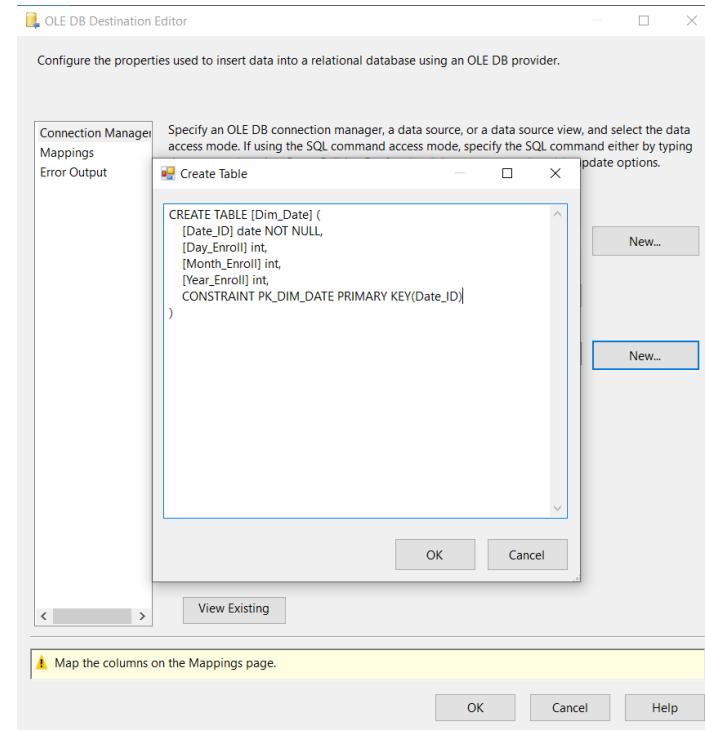
Bước 20: Click chuột phải vào Sort -> Chọn Edit -> Click chọn sort theo tất cả các cột và phương thức là Ascending -> OK



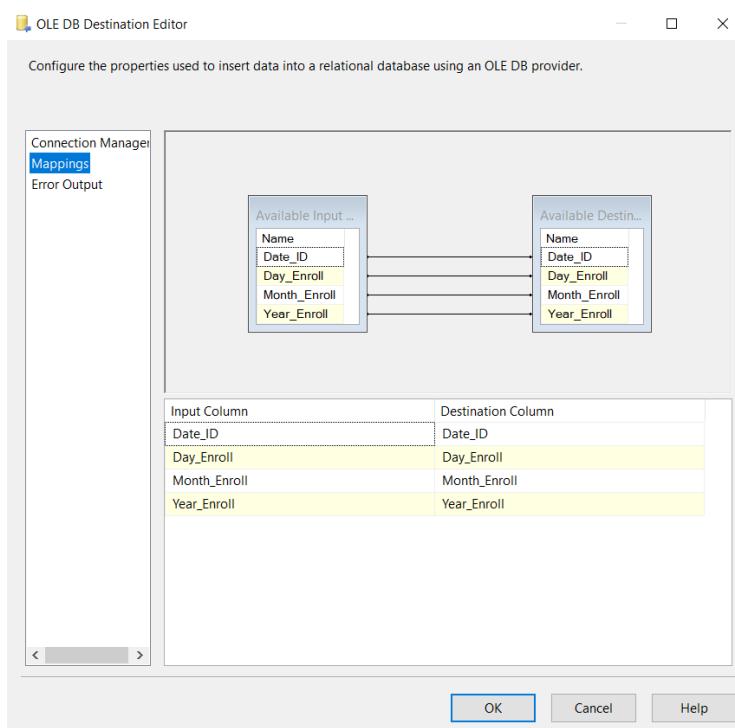
Bước 21: Kéo thả công cụ OLE DB Destination để định nghĩa Dim_Date -> Đổi tên thành Dim_Date



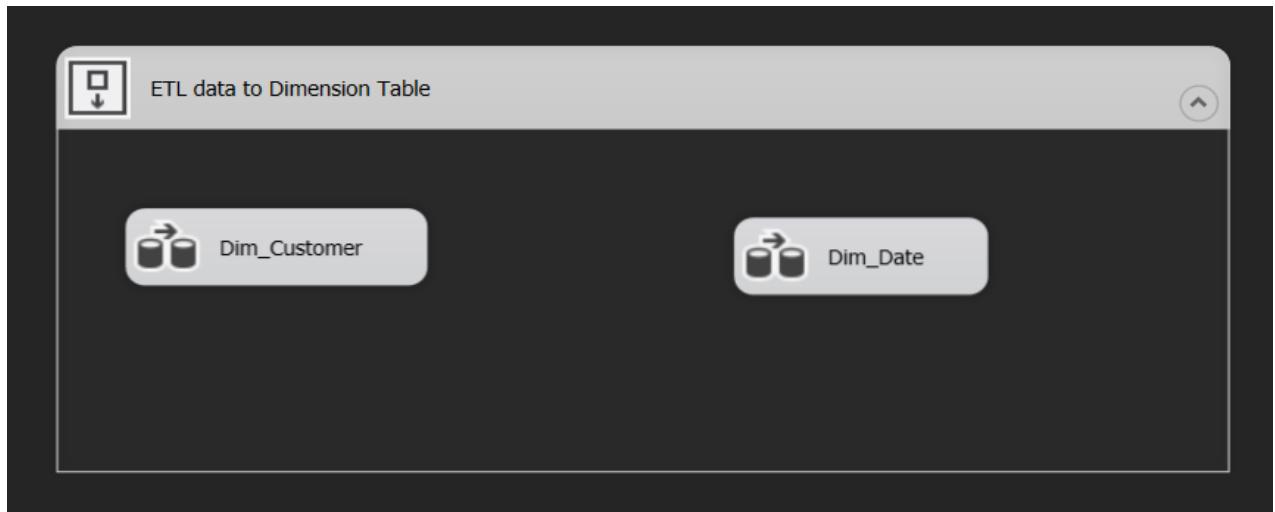
Bước 22: Click chuột phải vào Dim_Date -> Chọn Edit -> Chọn connection đến Data Warehouse -> Chọn New để SSIS tự sinh ra đoạn mã insert data vào table -> Thêm primary key ngay trên đoạn mã được sinh tự động -> OK



Bước 23: Click chọn Mappings -> Check xem những columns đã được map đúng hay chưa -> OK



Kết quả thu được sau khi hoàn thành định nghĩa các Data Flow Task cho Sequence Container ETL data to Dimension Table



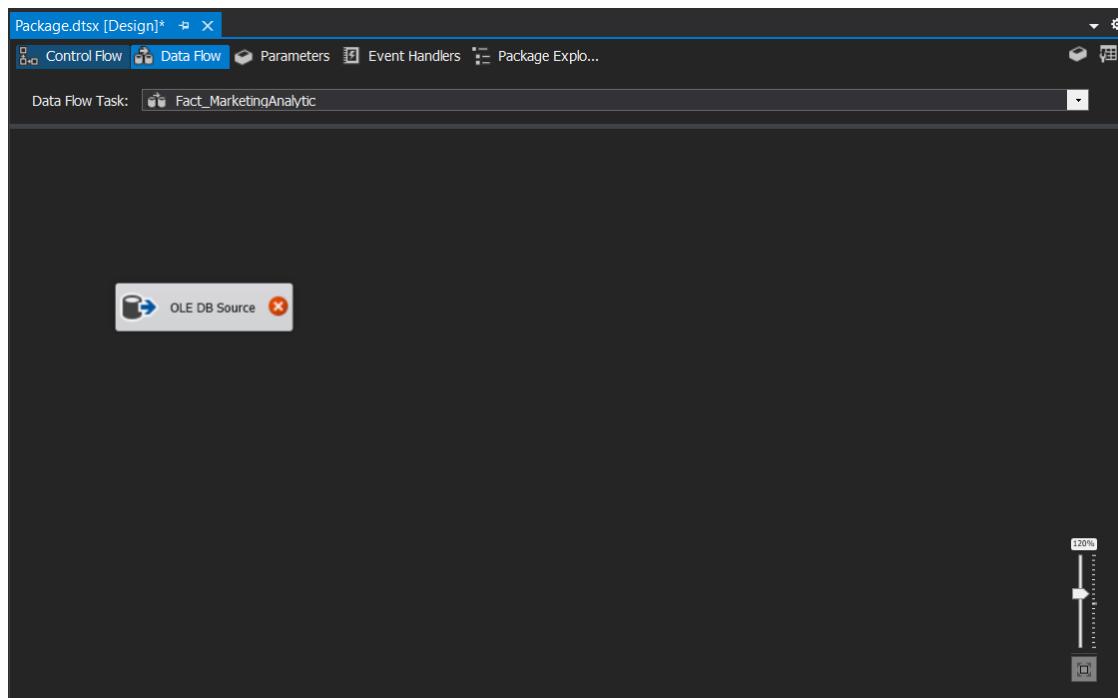
2.5 Định nghĩa và nạp dữ liệu vào bảng Fact

Mục tiêu: Tiến hành nạp dữ liệu vào bảng Fact_MarketingAnalytic theo đúng thiết kế tương ứng

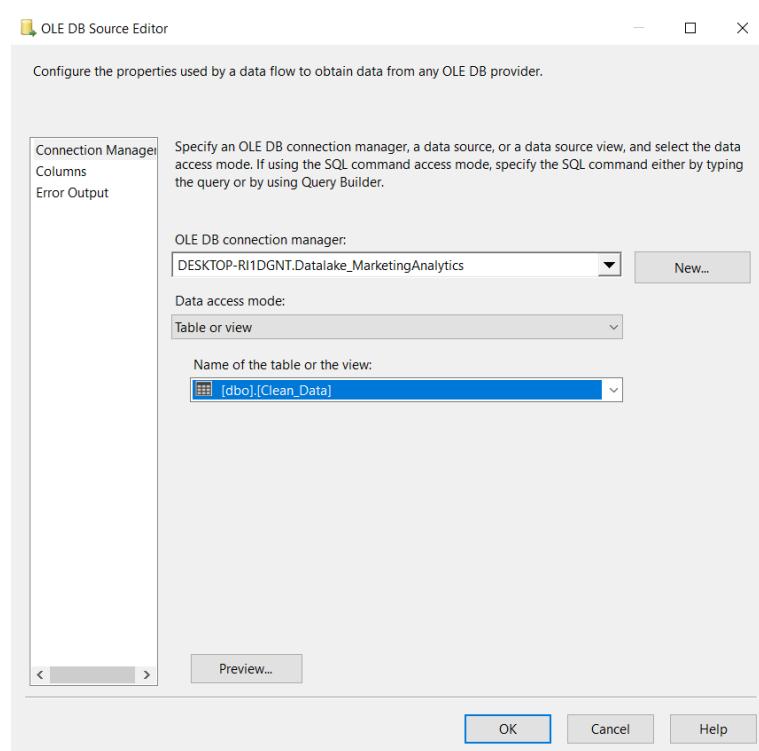
Bước 1: Kéo thả công cụ Data Flow Task vào Container -> Đổi tên thành Fact_MarketingAnalytic



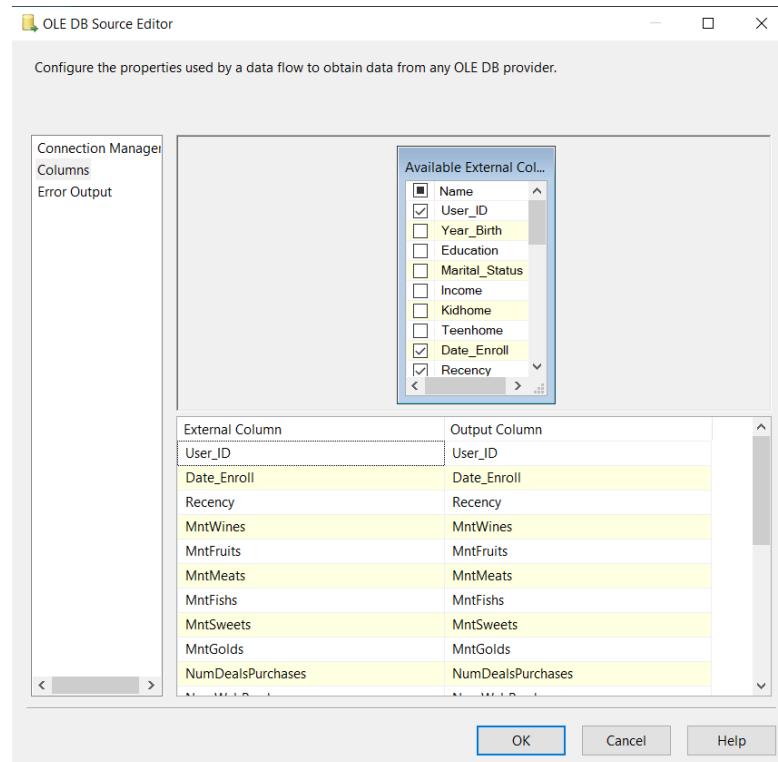
Bước 2: Click chuột trái vào Fact_MarketingAnalytic để chỉnh sửa Data Flow Task -> Kéo thả công cụ OLE DB Source để định nghĩa dữ liệu nguồn



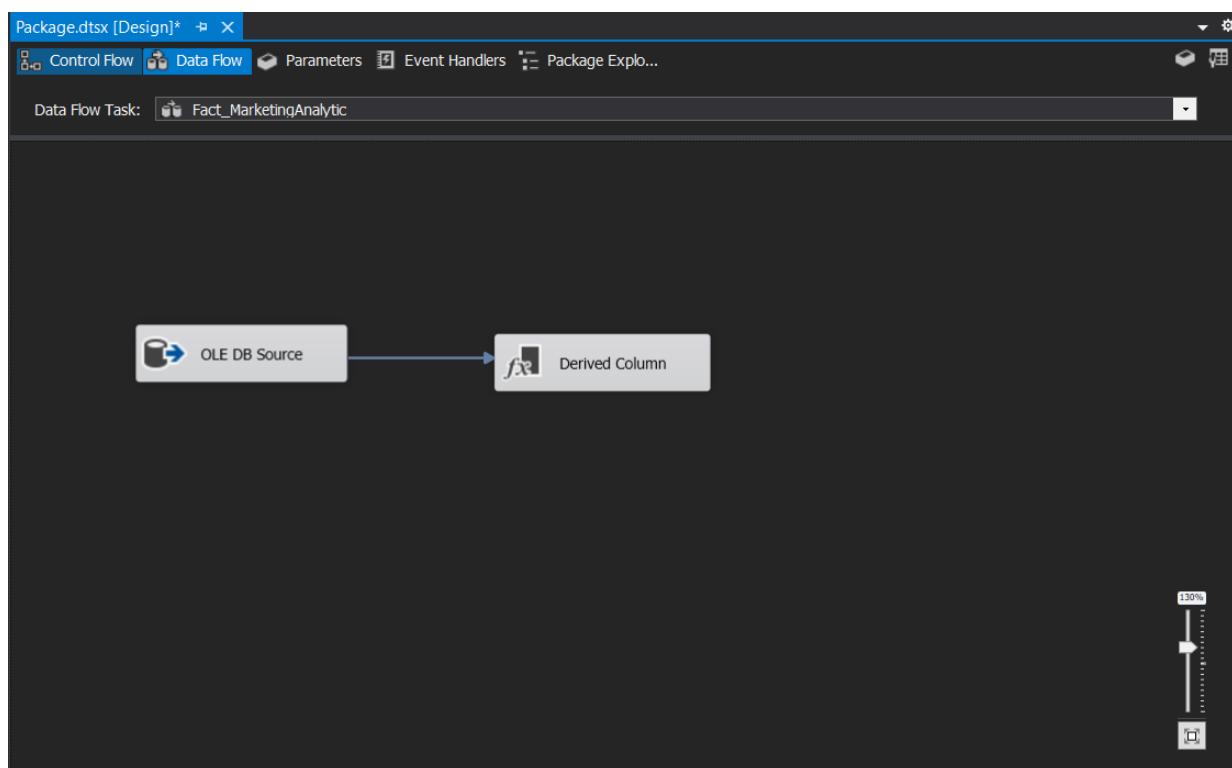
Bước 3: Click chuột phải và chọn Edit -> Chọn connection đến Data Lake -> Chọn table Clean_Data



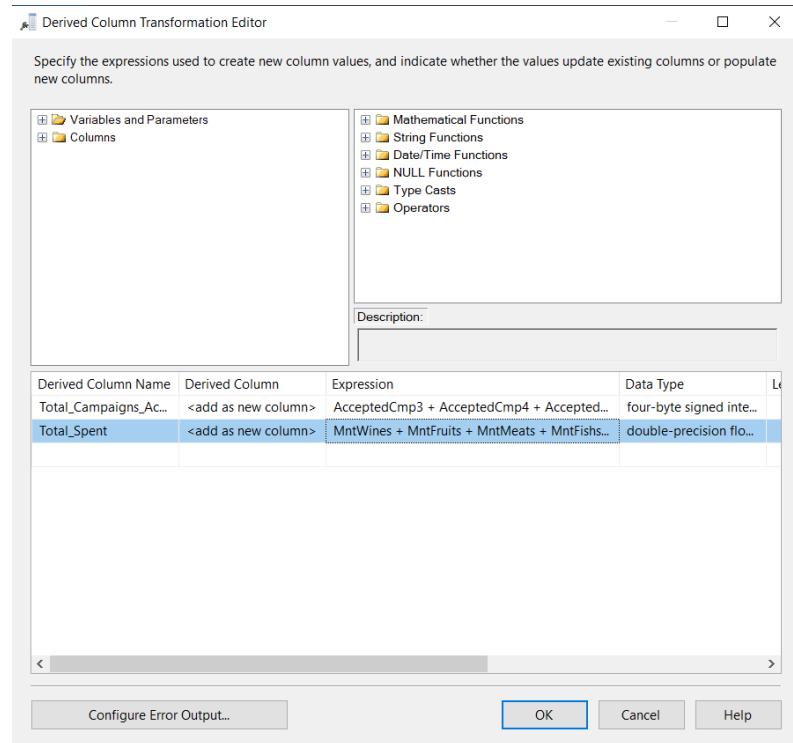
Bước 4: Click chọn Columns để lấy ra những column cần thiết cho bảng Fact_MarketingAnalytic -> OK



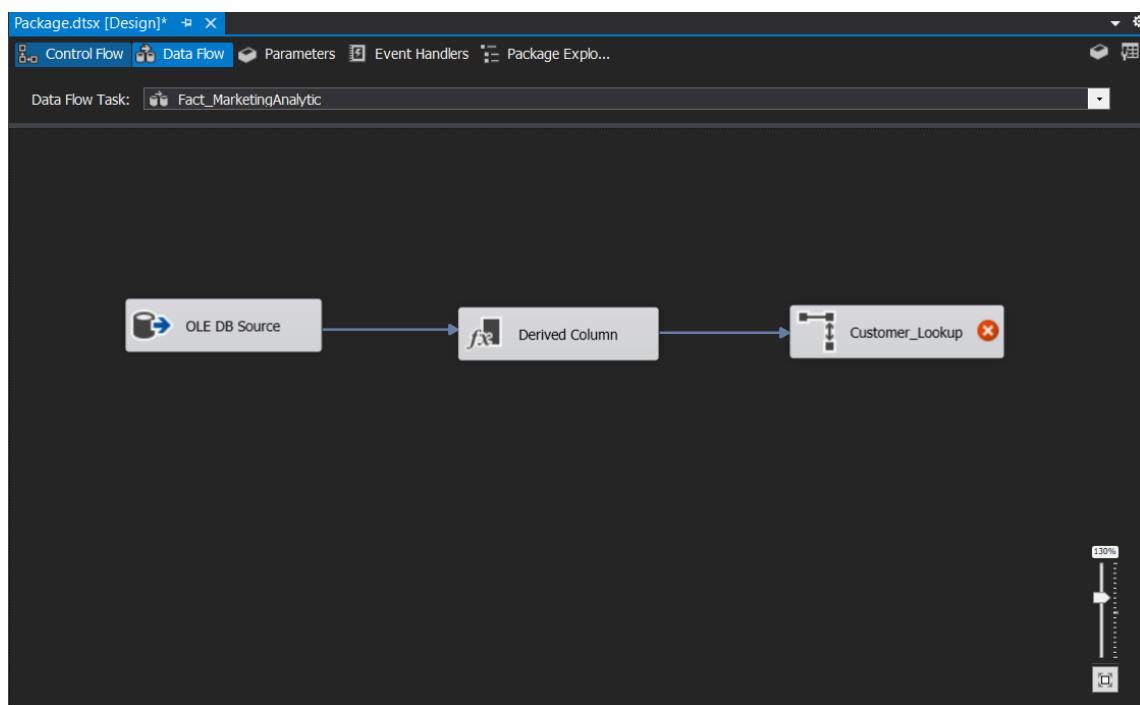
Bước 5: Kéo thả công cụ Derived Column để thêm 2 cột dữ liệu Total_Campaigns_Accepted và Total_Spent như trong thiết kế kho dữ liệu



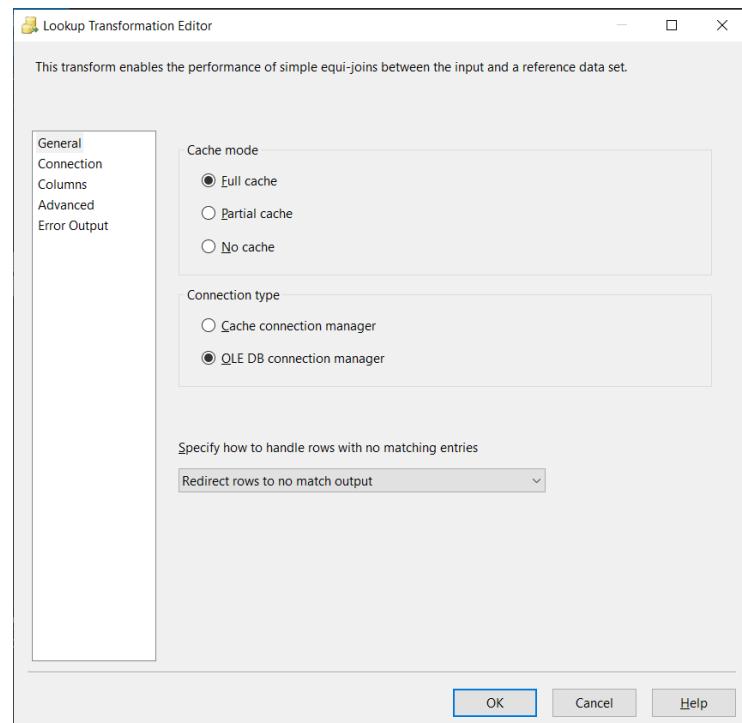
Bước 6: Click chuột phải vào Derived Column -> Chọn Edit -> Chính sửa Derived Column Name theo thiết kế kho dữ liệu tương ứng-> Thêm các function cần thiết để thêm 2 cột Total_Campaigns_Accepted và Total_Spent -> OK



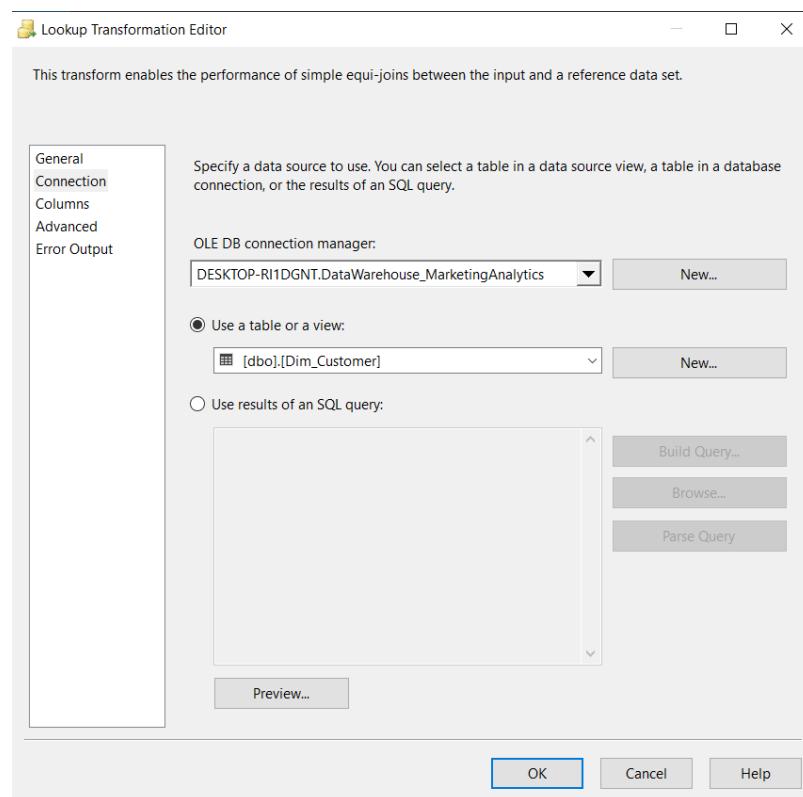
Bước 7: Kéo thả công cụ Lookup để link bảng Fact và các bảng Dim_Customer -> Đổi tên thành Customer_Lookup



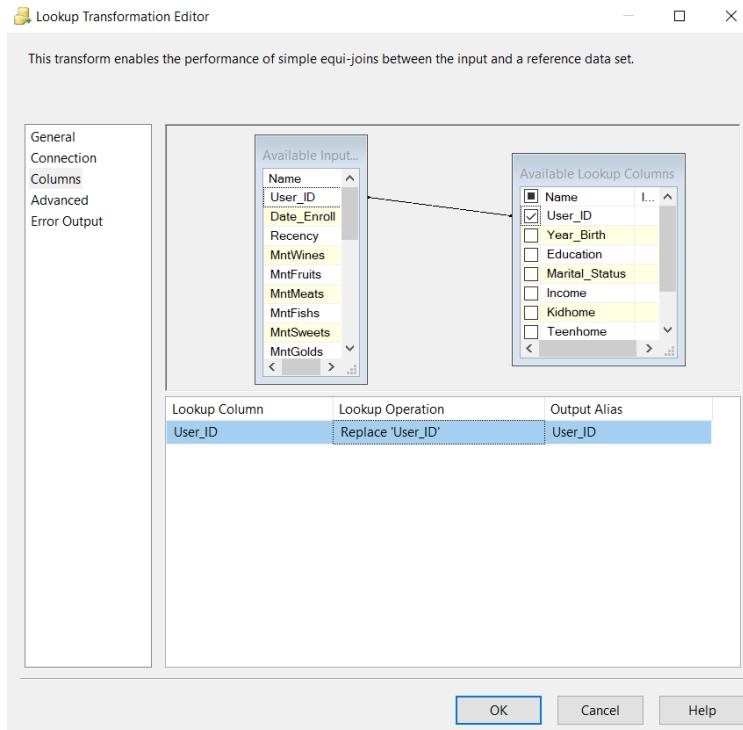
Bước 8: Click chuột phải vào Customer_Lookup -> Chuyển specify how to handle rows thành Redirect rows to no match output



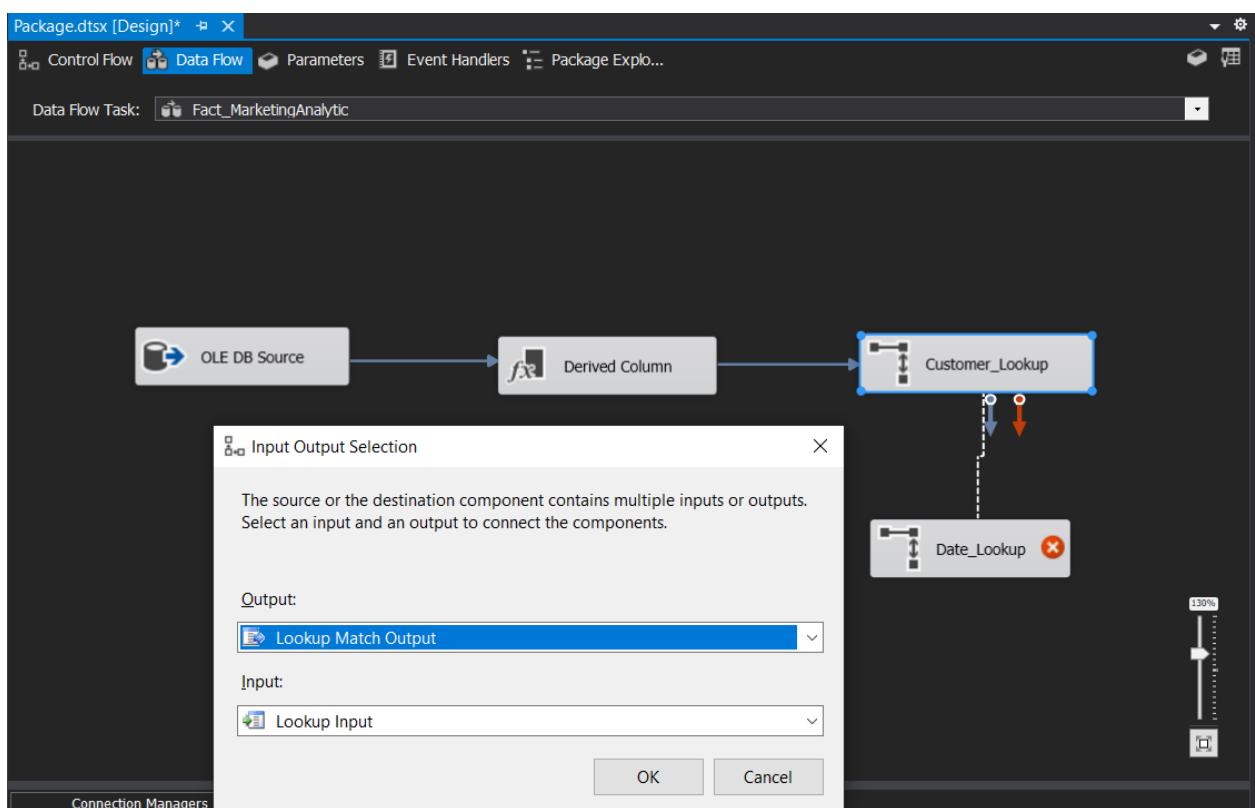
Bước 9: Click chọn Connection -> Đổi connection đến Data Warehouse -> Chọn table Dim_Customer



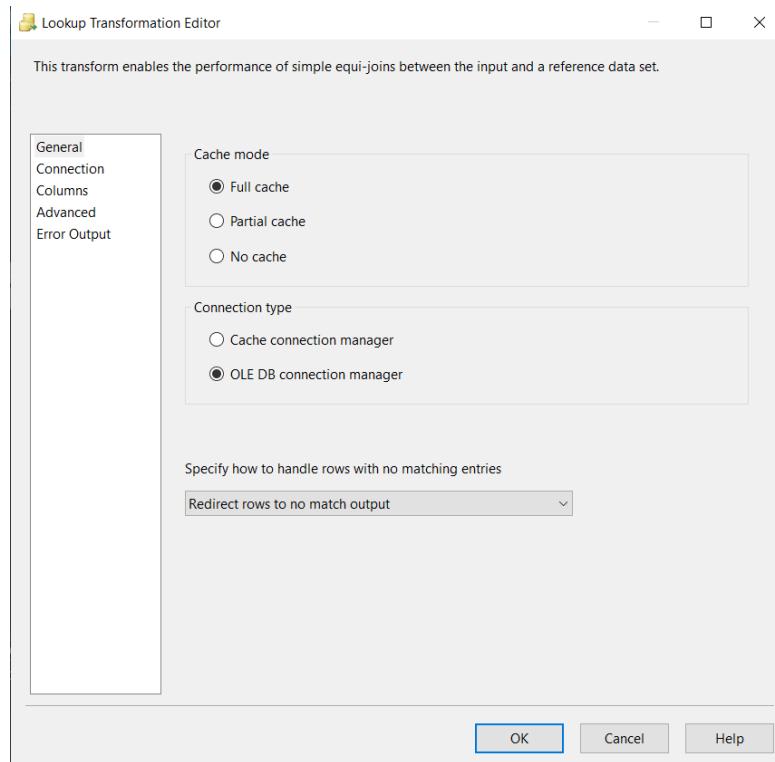
Bước 10: Chọn columns -> Nối cột User_ID -> Đổi Lookup Operation thành Replace User_ID -> OK



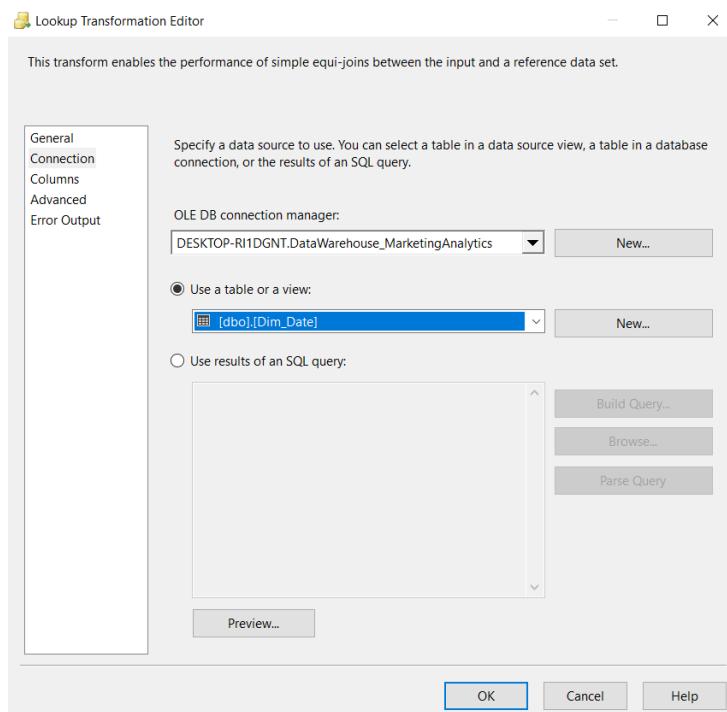
Bước 11: Kéo thả công cụ Lookup để link bảng Fact và các bảng Dim_Date -> Đổi tên thành Date_Lookup -> Kéo Data Flow -> Chọn Lookup Match Output -> OK



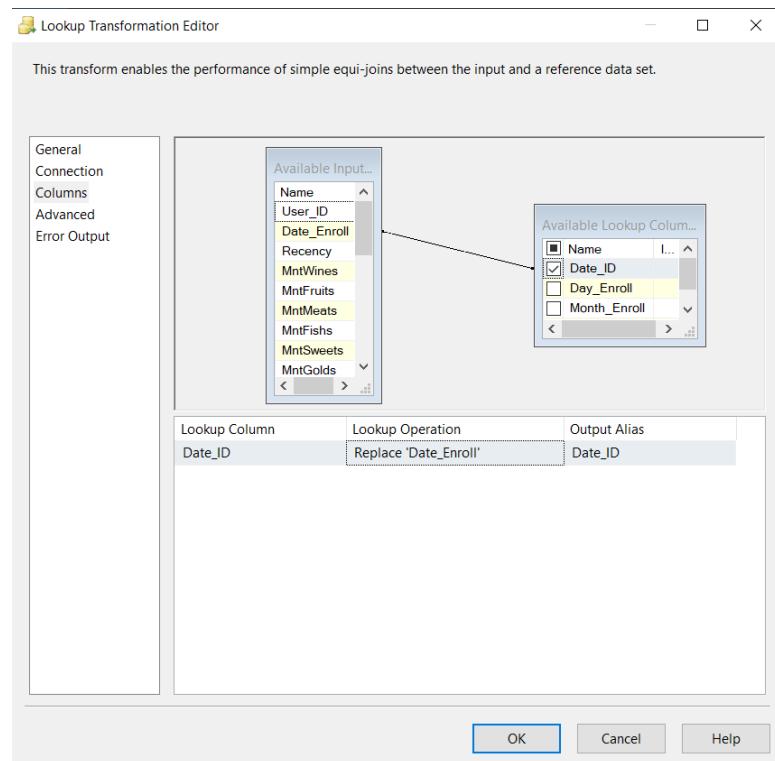
Bước 12: Click chuột phải vào Date_Lookup -> Chuyển specify how to handle rows thành Redirect rows to no match output



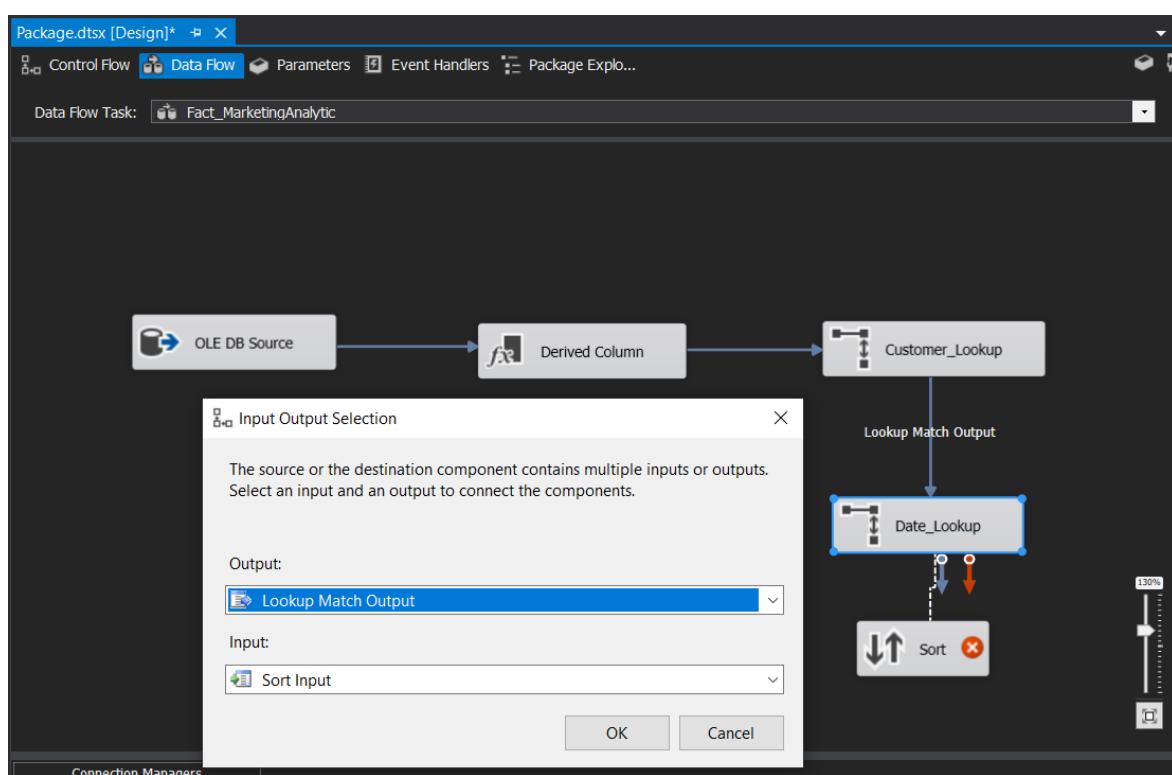
Bước 13: Click chọn Connection -> Đổi connection đến Data Warehouse -> Chọn table Dim_Date



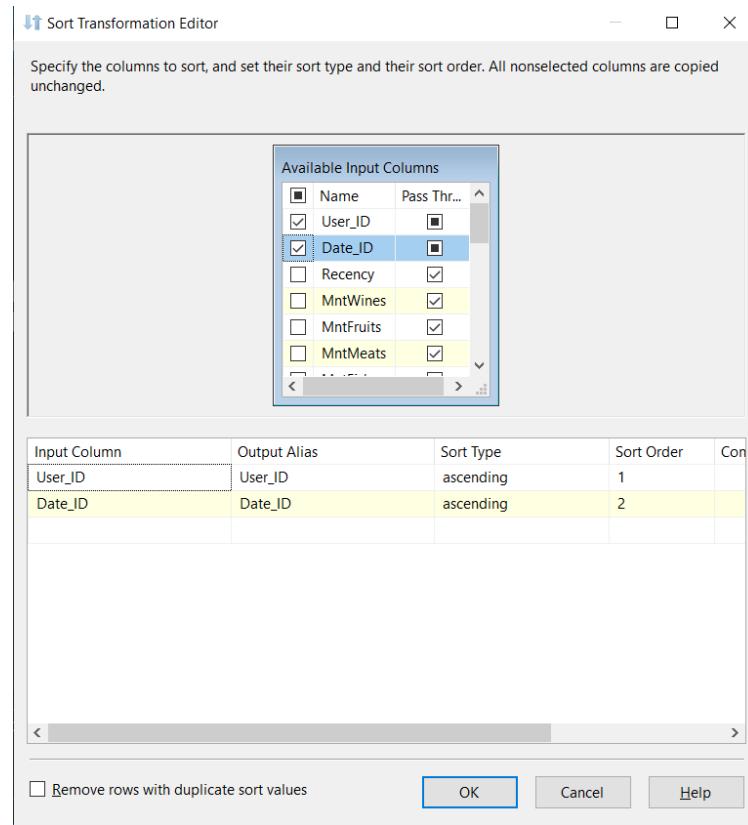
Bước 14: Chọn columns -> Nối cột Date_ID -> Đổi Lookup Operation thành Replace Date_ID -> OK



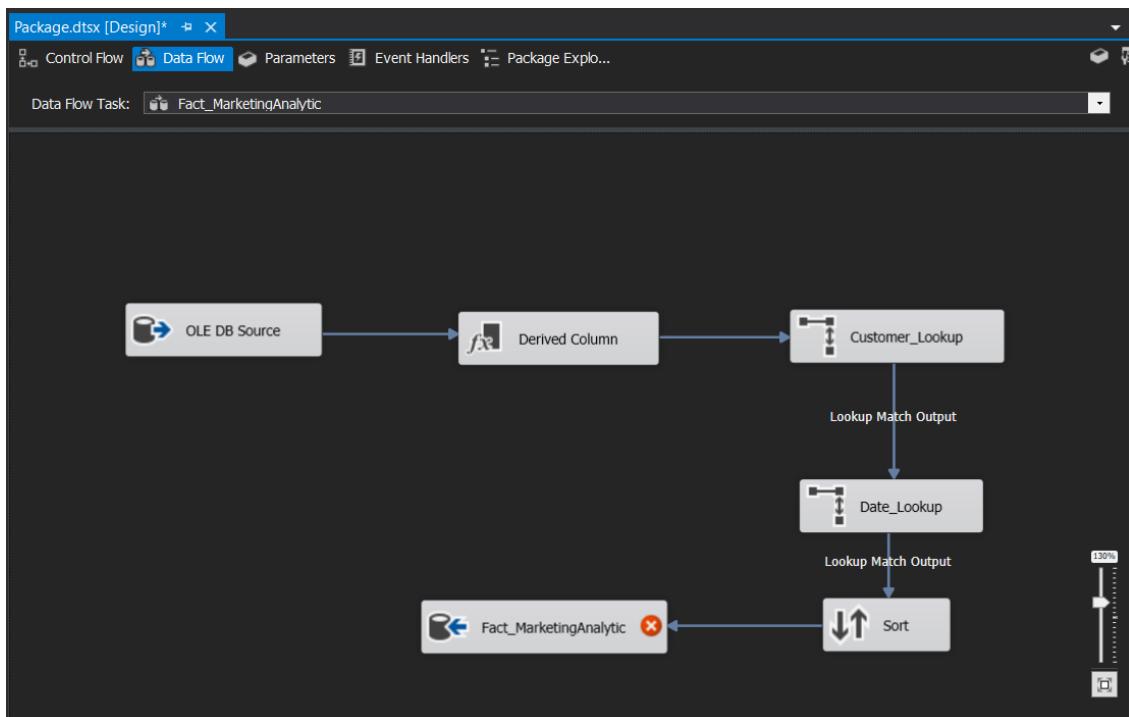
Bước 15: Kéo thả công cụ Sort để sắp xếp lại các dòng dữ liệu -> Kéo Data Flow -> Chọn Lookup Match Output -> OK



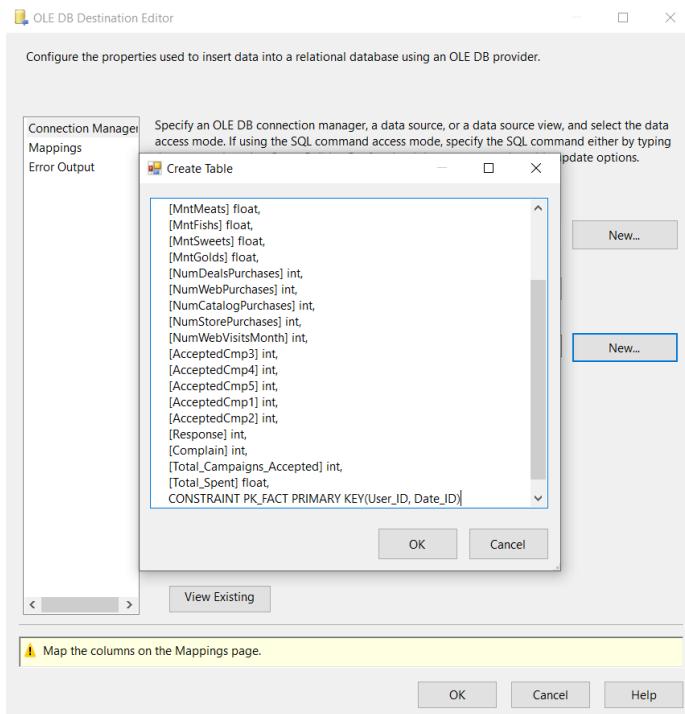
Bước 16: Click chuột phải vào Sort -> Chọn Edit -> Click chọn sort theo cột User_ID, Date_ID và phương thức là Ascending -> OK



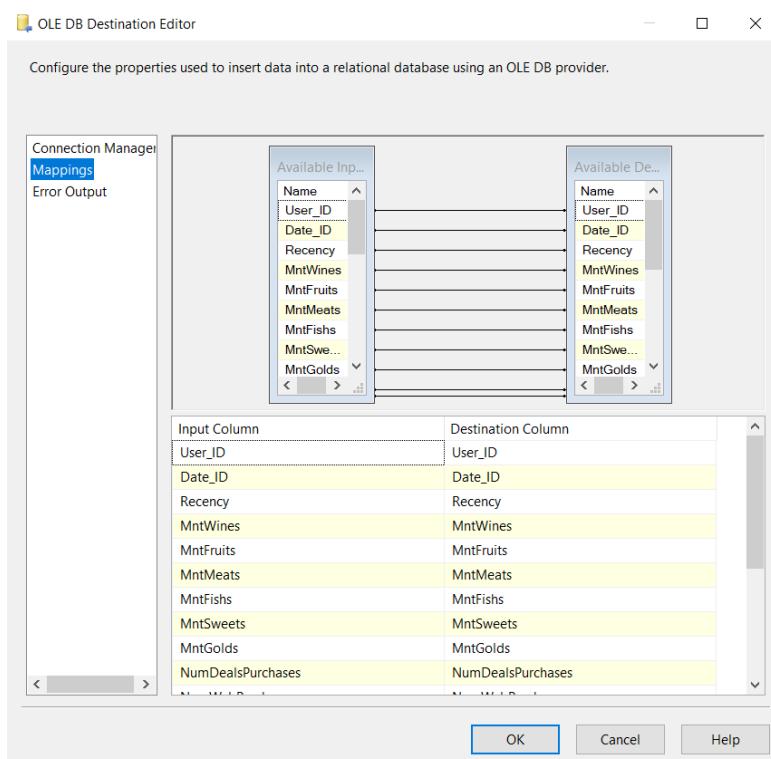
Bước 17: Kéo thả công cụ OLE DB Destination để định nghĩa Fact_MarketingAnalytic
-> Đổi tên thành Fact_MarketingAnalytic



Bước 18: Click chuột phải vào Fact_MarketingAnalytic -> Chọn Edit -> Chọn connection đến Data Warehouse -> Chọn New để SSIS tự sinh ra đoạn mã insert data vào table -> Thêm primary key ngay trên đoạn mã được sinh tự động -> OK



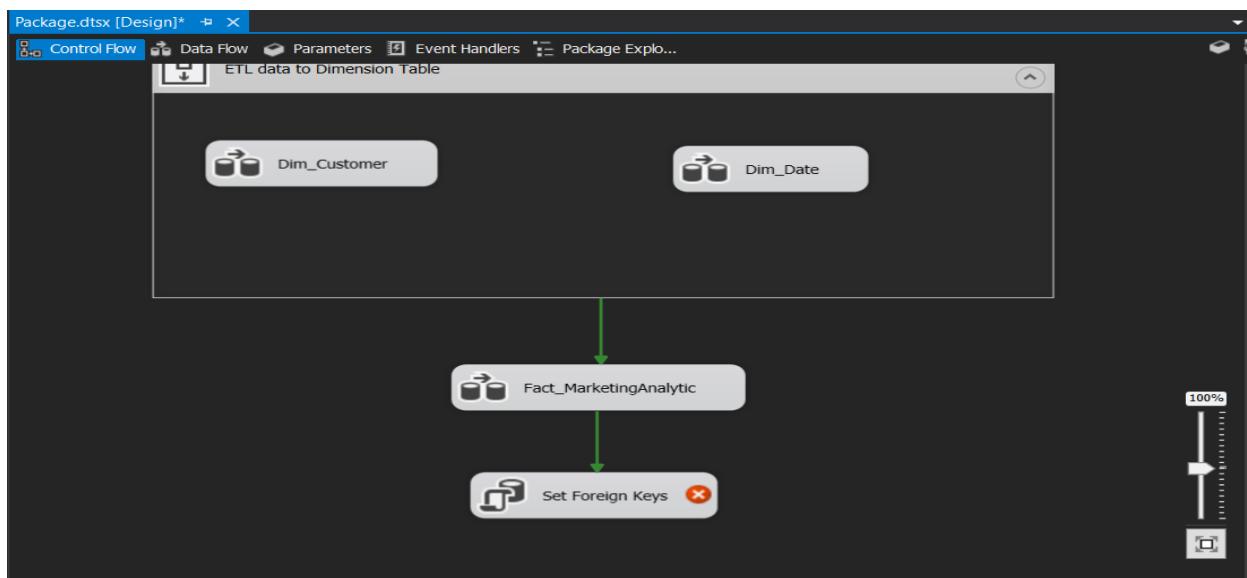
Bước 19: Click chọn Mappings -> Check xem những columns đã được map đúng hay chưa -> OK



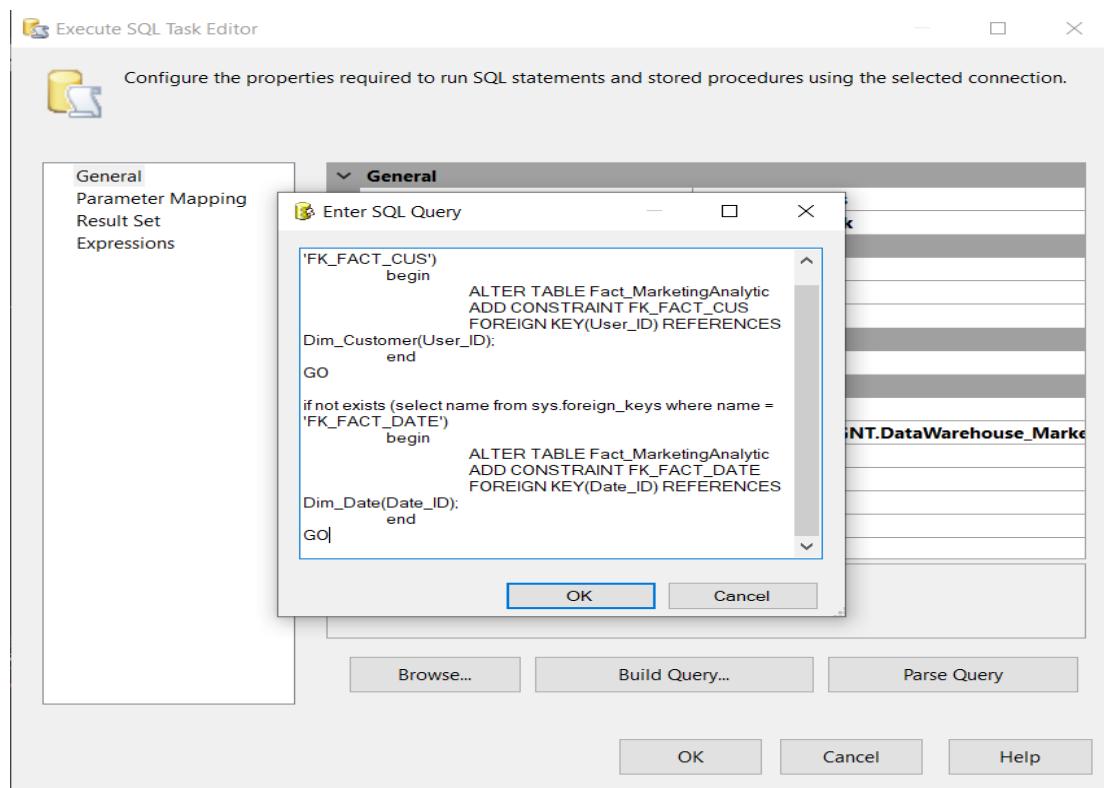
2.6 Định nghĩa và Set các khóa ngoại cho bảng Fact

Mục tiêu: Sau khi đã nạp dữ liệu vào bảng Fact, tiến hành tạo các khóa ngoại theo đúng thiết kế

Bước 1: Kéo thả công cụ Execute SQL Task vào Container -> Đổi tên thành Set Foreign Keys



Bước 2: Chuột phải vào Set Foreign Keys -> Edit -> Định nghĩa Connection là tới DataWarehouse_MarketingAnalytics -> Chính sửa SQLStatement -> OK



```

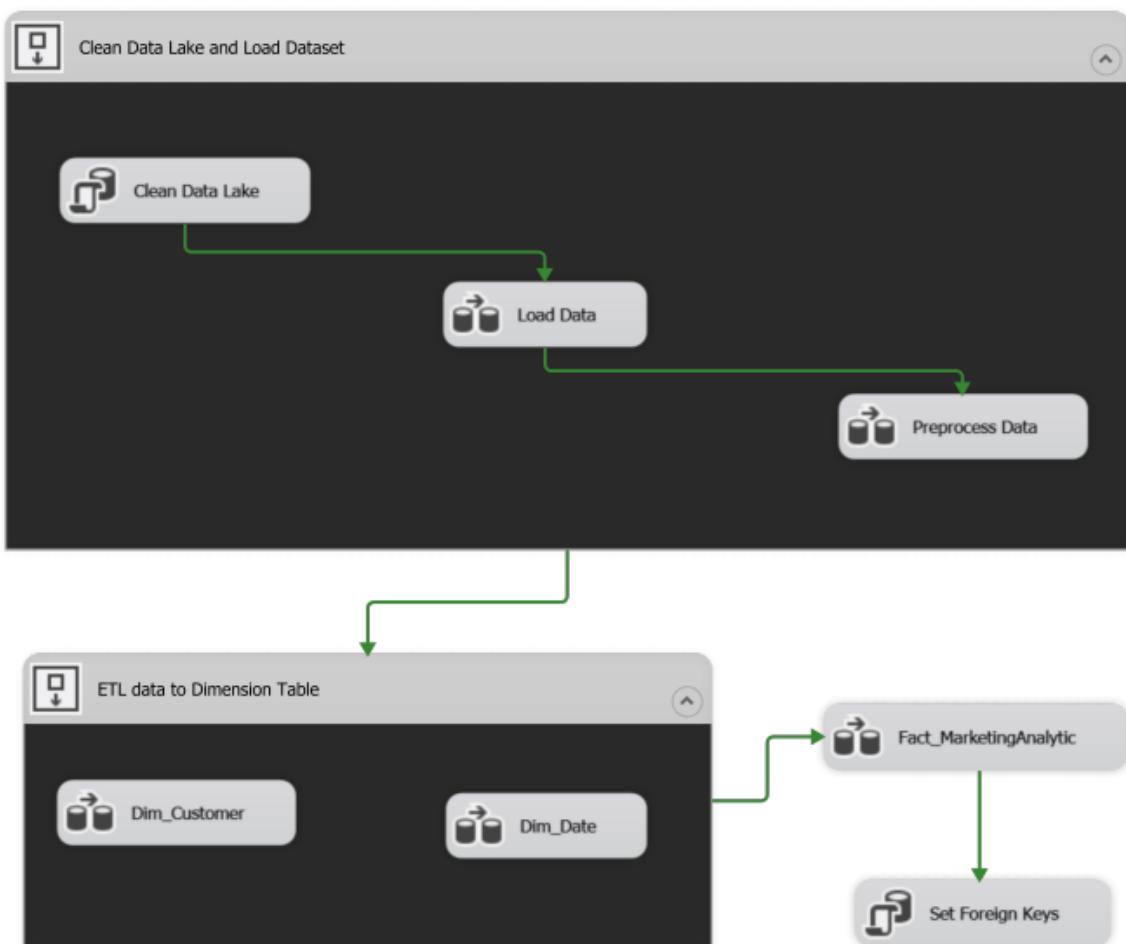
if not exists (select name from sys.foreign_keys where name = 'FK_FACT_CUS')
begin
    ALTER TABLE Fact_MarketingAnalytic
    ADD CONSTRAINT FK_FACT_CUS
    FOREIGN KEY(User_ID) REFERENCES Dim_Customer(User_ID);
end
GO

if not exists (select name from sys.foreign_keys where name = 'FK_FACT_DATE')
begin
    ALTER TABLE Fact_MarketingAnalytic
    ADD CONSTRAINT FK_FACT_DATE
    FOREIGN KEY(Date_ID) REFERENCES Dim_Date(Date_ID);
end
GO

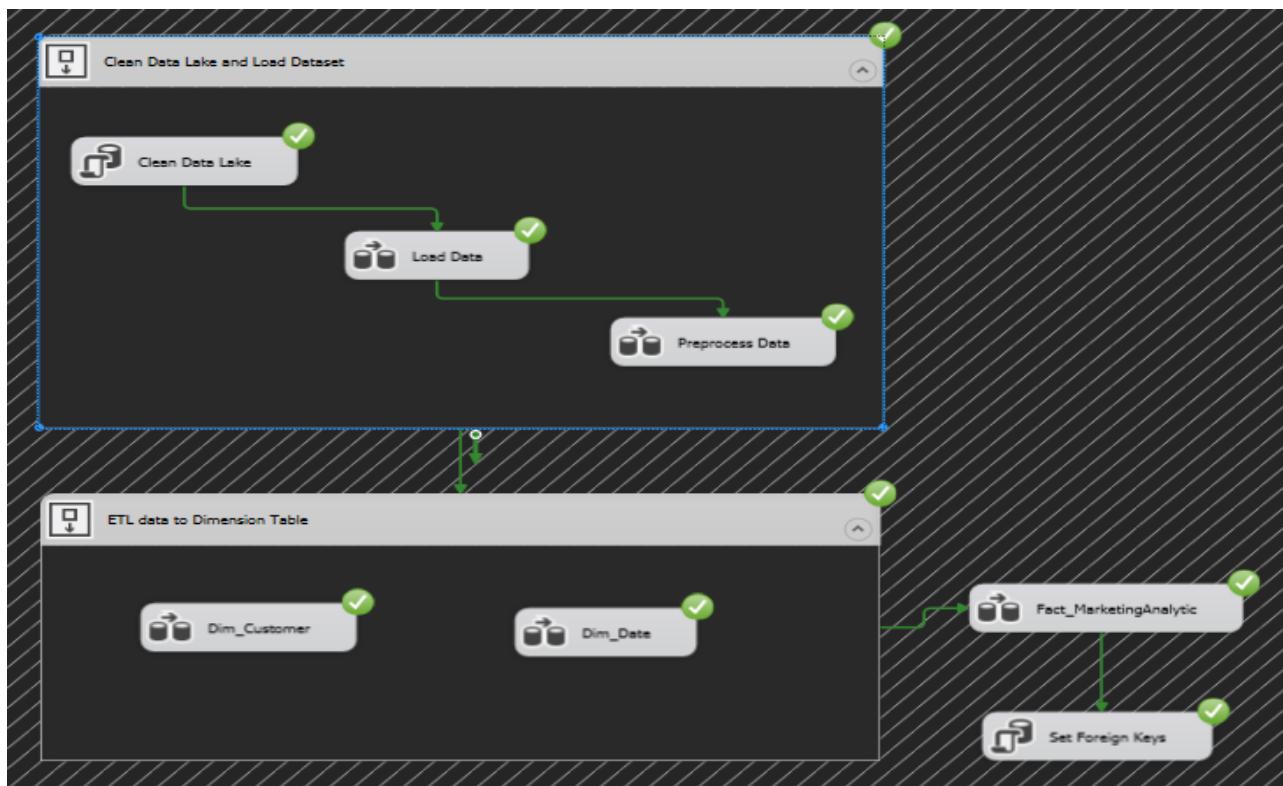
```

Câu lệnh SQL để add vào SQLStatement, thực hiện việc set foreign keys cho bảng Fact

2.7 Thực thi Package



Control Flow



Kết quả thực thi Package

2.8 Dữ liệu sau khi hoàn thành

Dữ liệu bảng Fact_MarketingAnalytic

Results																
Messages																
User_ID	Date_ID	Recency	MntWines	MntFruits	MntMeats	MntFishes	MntSweets	MntGolds	NumDealsPurchases	NumWebPurchases	NumCatalogPurchases	NumStorePurchases	NumWebVisitsMonth	Response		
1	0	2013-05-04	66	239	10	554	254	87	54	1	3	4	1	0		
2	1	2014-06-15	0	464	5	64	7	0	37	1	7	3	7	5	1	
3	9	2012-08-18	86	57	0	27	0	0	36	4	3	2	2	8	0	
4	13	2013-07-22	57	19	0	5	0	0	8	2	1	0	3	6	0	
5	17	2013-09-06	81	637	47	237	12	19	76	4	6	11	7	5	0	
6	20	2013-09-01	91	43	12	23	29	15	61	1	2	1	4	4	0	
7	22	2012-12-03	99	185	2	88	15	5	14	2	6	1	5	8	0	
8	24	2014-02-15	96	18	2	19	0	2	6	5	3	0	4	7	0	
9	25	2012-11-16	9	460	35	422	33	12	153	2	6	6	7	4	0	
10	35	2012-09-08	35	32	1	64	16	12	85	3	2	2	3	6	1	
11	48	2014-04-24	97	136	1	12	0	3	32	2	4	1	3	6	0	
12	49	2014-05-11	39	2	3	6	4	1	9	1	1	1	2	7	0	
13	55	2012-12-07	83	509	0	65	7	11	5	4	7	2	9	6	0	
14	67	2013-09-18	6	68	0	16	0	0	8	3	2	0	4	7	0	
15	73	2013-05-26	81	14	0	3	0	0	1	1	0	0	3	5	0	
16	75	2013-07-25	32	34	11	137	179	61	45	1	8	1	5	8	0	
17	78	2013-09-15	51	15	0	11	0	2	9	1	1	1	2	8	0	
18	87	2013-08-26	16	0	7	5	26	2	17	2	2	0	3	7	0	
19	89	2013-12-09	49	5	9	20	6	8	21	2	2	0	4	6	0	
20	92	2014-05-12	12	5	7	24	19	14	20	4	3	0	4	6	0	
21	113	2013-07-16	33	11	0	7	3	1	6	2	1	0	3	5	0	
22	115	2013-03-26	0	275	11	68	25	7	7	3	5	1	8	5	0	
23	123	2013-01-06	92	544	35	133	23	151	107	2	4	5	4	2	0	
24	125	2013-05-15	65	215	7	33	3	0	13	6	5	1	5	7	0	
25	143	2013-08-25	73	466	0	224	119	49	99	1	5	3	4	2	0	
26	146	2013-11-15	78	760	53	400	17	93	0	1	4	5	11	2	0	
27	153	2013-10-17	29	19	3	6	4	2	12	1	1	0	3	4	0	
28	158	2013-11-17	3	345	53	528	98	75	97	1	8	3	5	4	1	
29	164	2012-10-13	59	206	0	46	3	2	12	2	5	1	5	8	0	
30	175	2013-01-10	93	656	80	455	52	107	93	2	8	4	8	4	0	
31	176	2014-01-02	73	90	38	67	33	36	67	1	4	1	6	3	0	

Query executed successfully.

Dữ liệu bảng Dim_Customer

The screenshot shows a SQL Server Management Studio window with two tabs: 'SQLQuery1.sql - D...\nghoanglong (65)' and 'SQL_Statements_M...\nghoanglong (61)*'. The code in the editor is:

```
SQLQuery1.sql - D...\nghoanglong (65) SQL_Statements_M...\nghoanglong (61)*
DROP TABLE dbo.Fact_MarketingAnalytic;
SELECT * FROM dbo.Fact_MarketingAnalytic;
SELECT * FROM dbo.Dim_Customer;
```

The results pane displays the data from the Dim_Customer table:

User_ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Country	
1	0	1985	Graduation	Married	70951	0	0	South Africa
2	1	1961	Graduation	Single	57091	0	0	Canada
3	9	1975	Master	Single	46098	1	1	Canada
4	13	1947	PhD	Widow	25358	0	1	South Africa
5	17	1971	PhD	Married	60491	0	1	Spain
6	20	1965	2n Cycle	Married	46891	0	1	Germany
7	22	1976	Graduation	Divorced	46310	1	0	Spain
8	24	1960	Master	Together	17144	1	1	South Africa
9	25	1958	Graduation	Married	65148	0	1	Canada
10	35	1987	Graduation	Married	25545	1	0	Spain
11	48	1964	Graduation	Together	55761	0	1	Spain
12	49	1970	Graduation	Single	20587	1	0	Canada
13	55	1963	Graduation	Together	56253	0	1	Spain
14	67	1972	Master	Single	46423	1	1	Spain
15	73	1953	PhD	Single	51411	1	2	India
16	75	1982	2n Cycle	Together	41039	0	0	United States
17	78	1969	Graduation	Married	25293	1	0	Spain
18	87	1981	2n Cycle	Married	27733	1	0	Germany
19	89	1975	2n Cycle	Married	22212	1	0	South Africa
20	92	1988	Graduation	Alone	34176	1	0	Canada
21	113	1951	Graduation	Married	30833	1	1	Spain
22	115	1966	Master	Single	43456	0	1	India
23	123	1951	Graduation	Widow	67046	0	1	Australia
24	125	1958	2n Cycle	Together	53083	1	1	India
25	143	1970	Graduation	Single	61209	0	0	United States
26	146	1960	PhD	Single	76045	0	0	Spain
27	153	1953	Graduation	Divorced	23272	0	0	Spain
28	158	1945	PhD	Together	71604	0	0	Spain
29	164	1977	PhD	Married	28973	0	0	Spain
30	175	1986	Graduation	Married	71952	1	0	Spain

Dữ liệu bảng Dim_Date

The screenshot shows a SQL Server Management Studio window with two tabs: 'SQLQuery1.sql - D...\nghoanglong (65)' and 'SQL_Statements_M...\nghoanglong (61)*'. The code in the editor is:

```
SQLQuery1.sql - D...\nghoanglong (65) SQL_Statements_M...\nghoanglong (61)*
DROP TABLE dbo.Fact_MarketingAnalytic;
SELECT * FROM dbo.Fact_MarketingAnalytic;
SELECT * FROM dbo.Dim_Customer;
SELECT * FROM dbo.Dim_Date;
```

The results pane displays the data from the Dim_Date table:

Date_ID	Day_Enroll	Month_Enroll	Year_Enroll	
1	2012-07-30	30	7	2012
2	2012-07-31	31	7	2012
3	2012-08-01	1	8	2012
4	2012-08-02	2	8	2012
5	2012-08-03	3	8	2012
6	2012-08-04	4	8	2012
7	2012-08-05	5	8	2012
8	2012-08-06	6	8	2012
9	2012-08-07	7	8	2012
10	2012-08-08	8	8	2012
11	2012-08-09	9	8	2012
12	2012-08-10	10	8	2012
13	2012-08-11	11	8	2012
14	2012-08-12	12	8	2012
15	2012-08-13	13	8	2012
16	2012-08-14	14	8	2012
17	2012-08-15	15	8	2012
18	2012-08-16	16	8	2012
19	2012-08-17	17	8	2012
20	2012-08-18	18	8	2012
21	2012-08-19	19	8	2012
22	2012-08-20	20	8	2012
23	2012-08-21	21	8	2012
24	2012-08-22	22	8	2012
25	2012-08-23	23	8	2012
26	2012-08-24	24	8	2012
27	2012-08-25	25	8	2012
28	2012-08-26	26	8	2012
29	2012-08-27	27	8	2012

Query executed successfully.

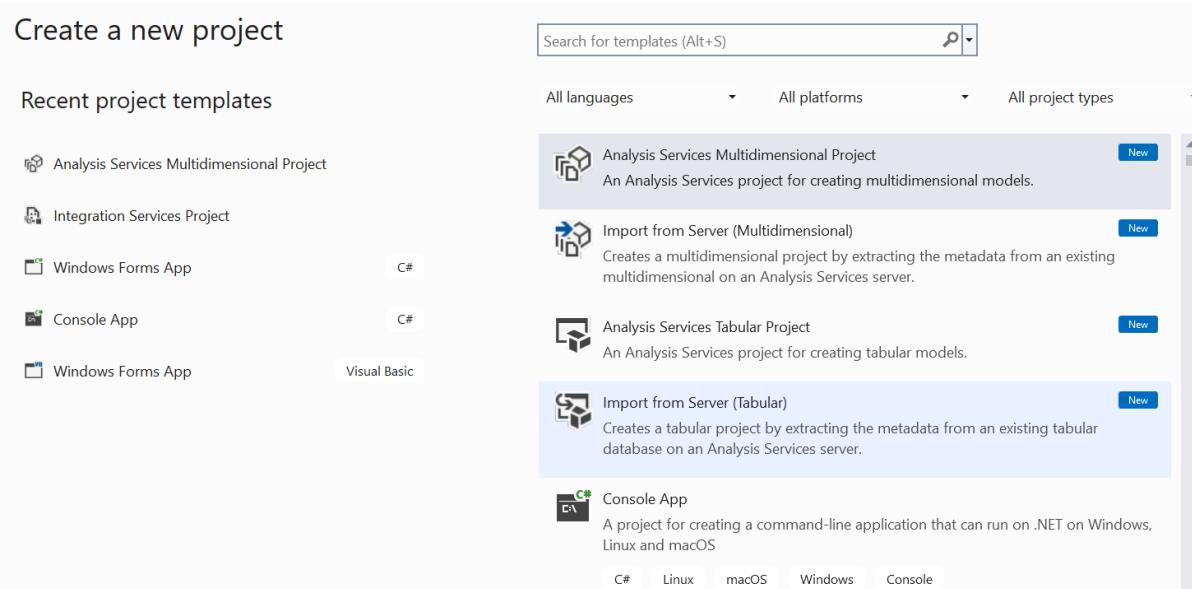
DESKTOP-R1DGNT (14.0 RTM) DESKTOP-R1DGNT\nghoan... DataWarehouse_Marketin... 00:00:00 662 rows

CHƯƠNG 3: QUÁ TRÌNH SSAS

3.1 Xây dựng khối dữ liệu

3.1.1 Tạo project trong visual studio 2022

Bước 1: Tạo project mới với Analysis Service Multidimensional and Data Mining Project



Configure your new project

Analysis Services Multidimensional Project

Project name

SSAS_MarketingAnalysis

Location

C:\Users\ADMIN\Desktop\SPKT\HKII_23_24\DAWH430784_01_Thu\



Solution

Create new solution

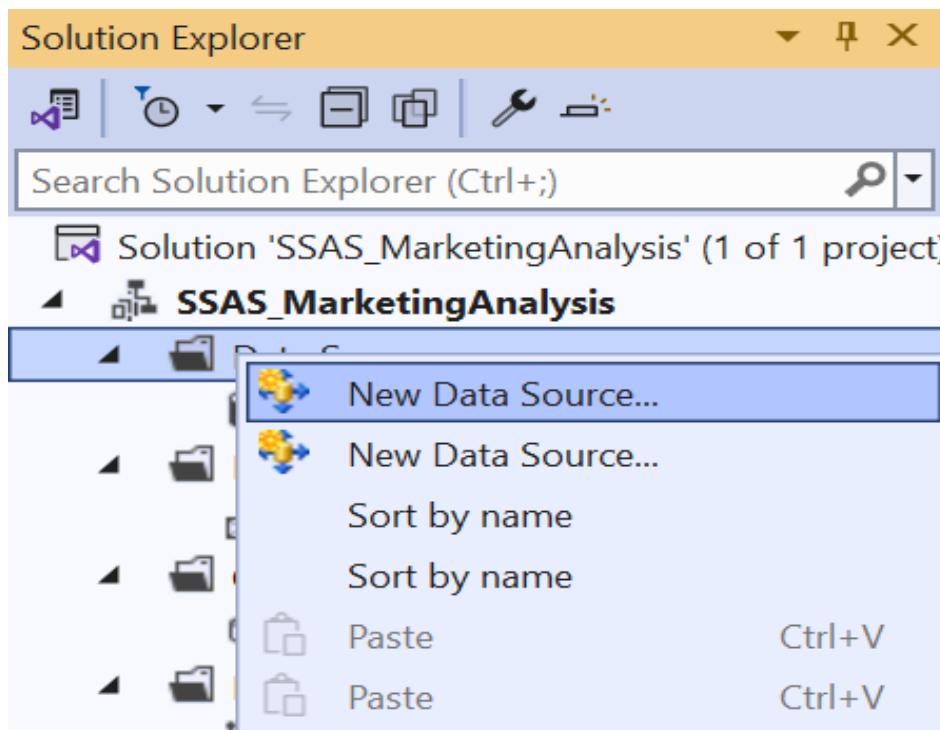
Solution name ?

SSAS_MarketingAnalysis

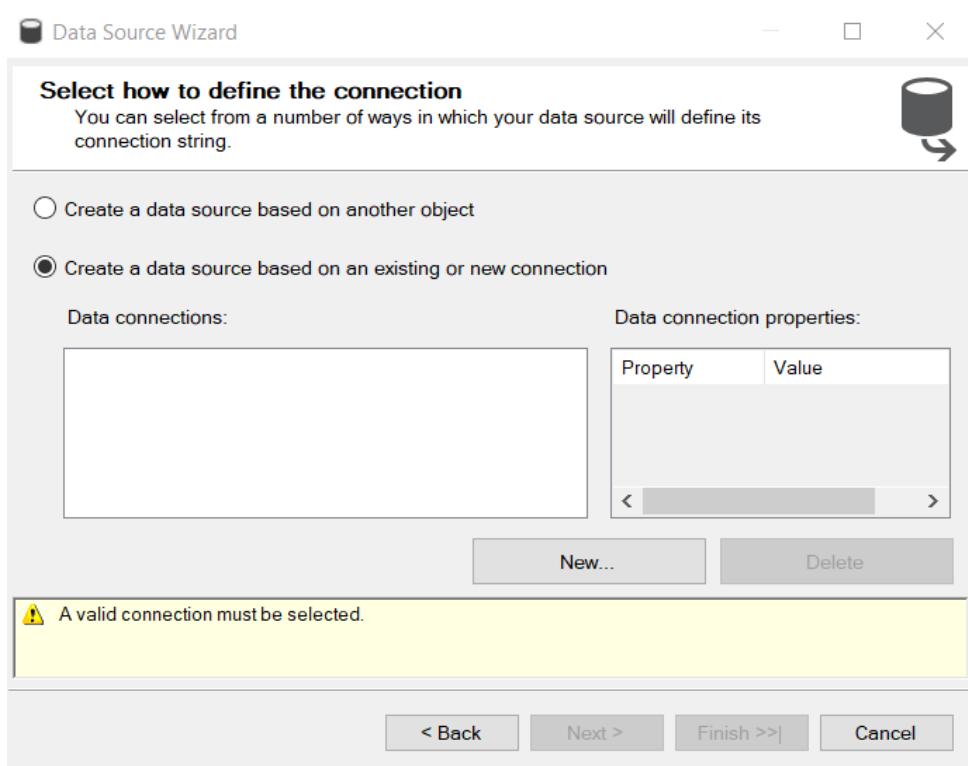
Place solution and project in the same directory

Project will be created in "C:\Users\ADMIN\Desktop\SPKT\HKII_23_24\DAWH430784_01_Thu\\SSAS_MarketingAnalysis\"

Bước 2: Vào giao diện chính, chuột phải vào Data Source -> New Data Sources

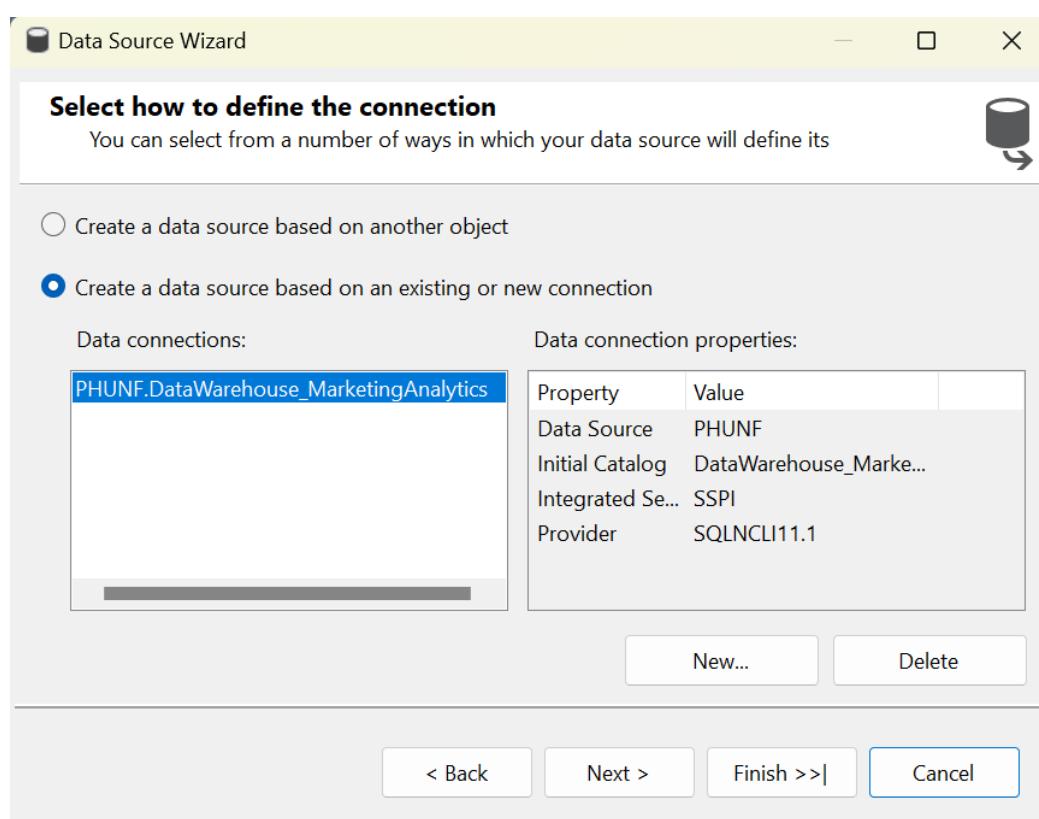
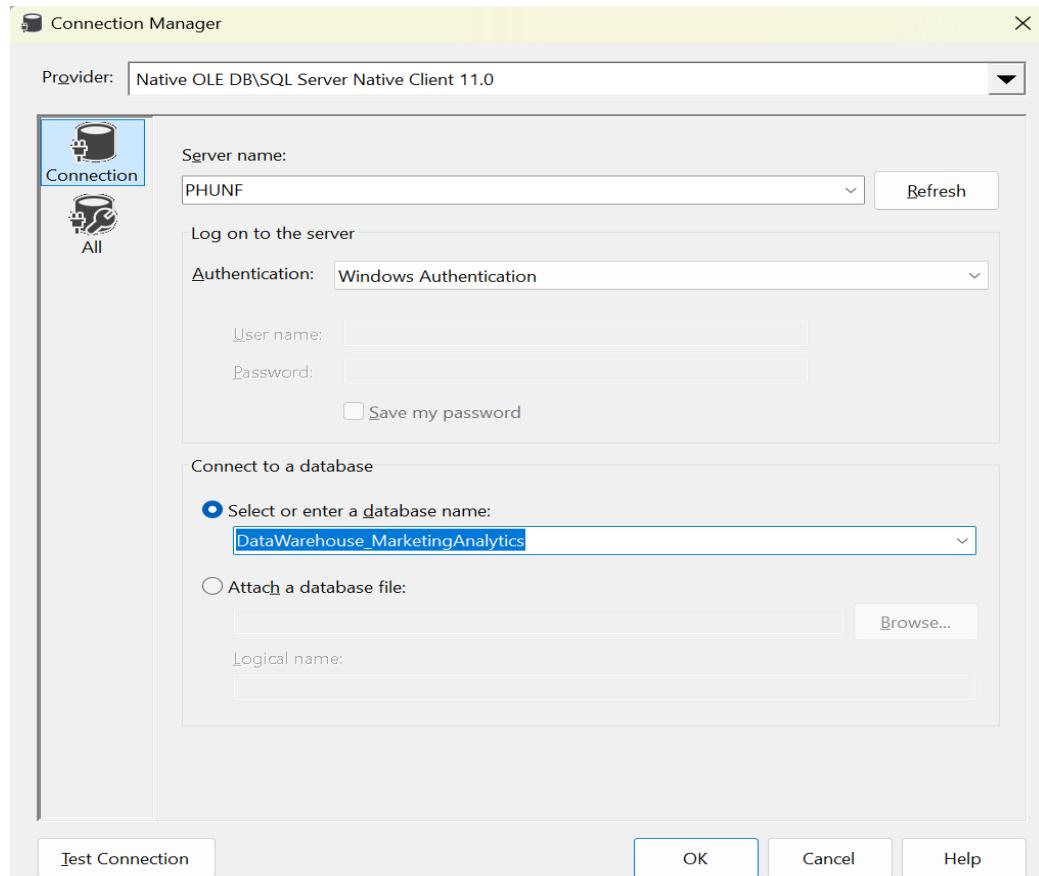


Bước 3: Click chuột vào lựa chọn thứ 2, rồi nhấn vào New để thêm database vào

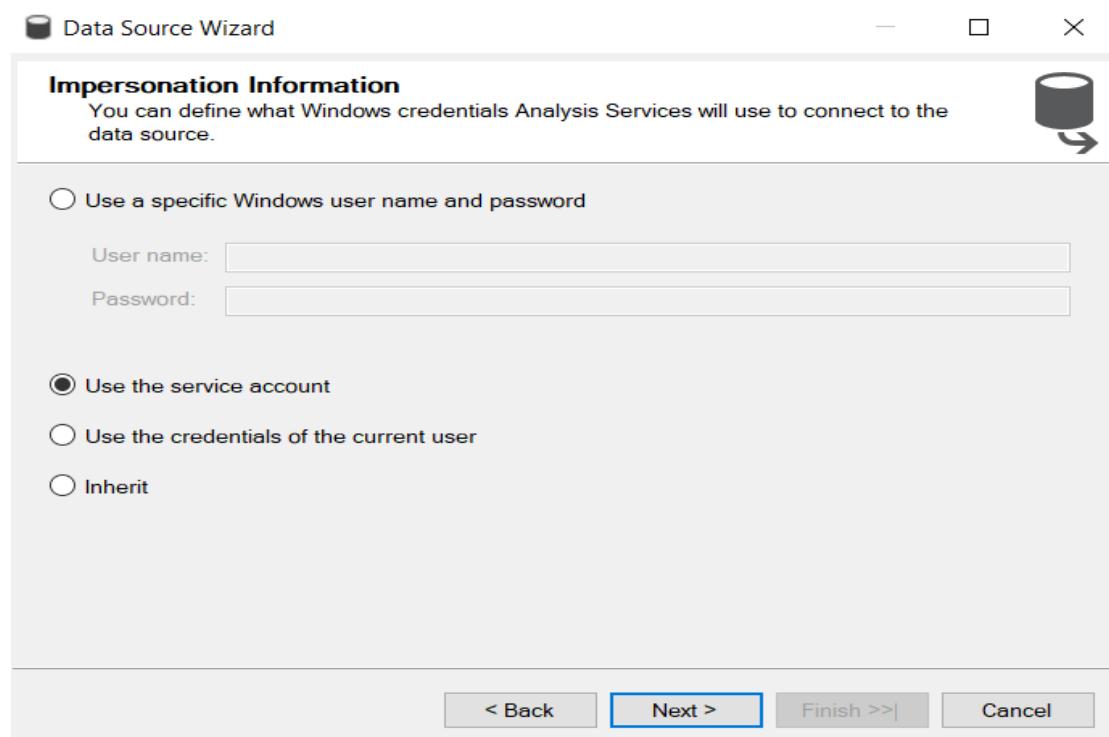


Bước 4: Đến đây, lấy tên Server name là tên Server name của database, rồi chọn database ở phần Connect to a database, xong nhấn OK

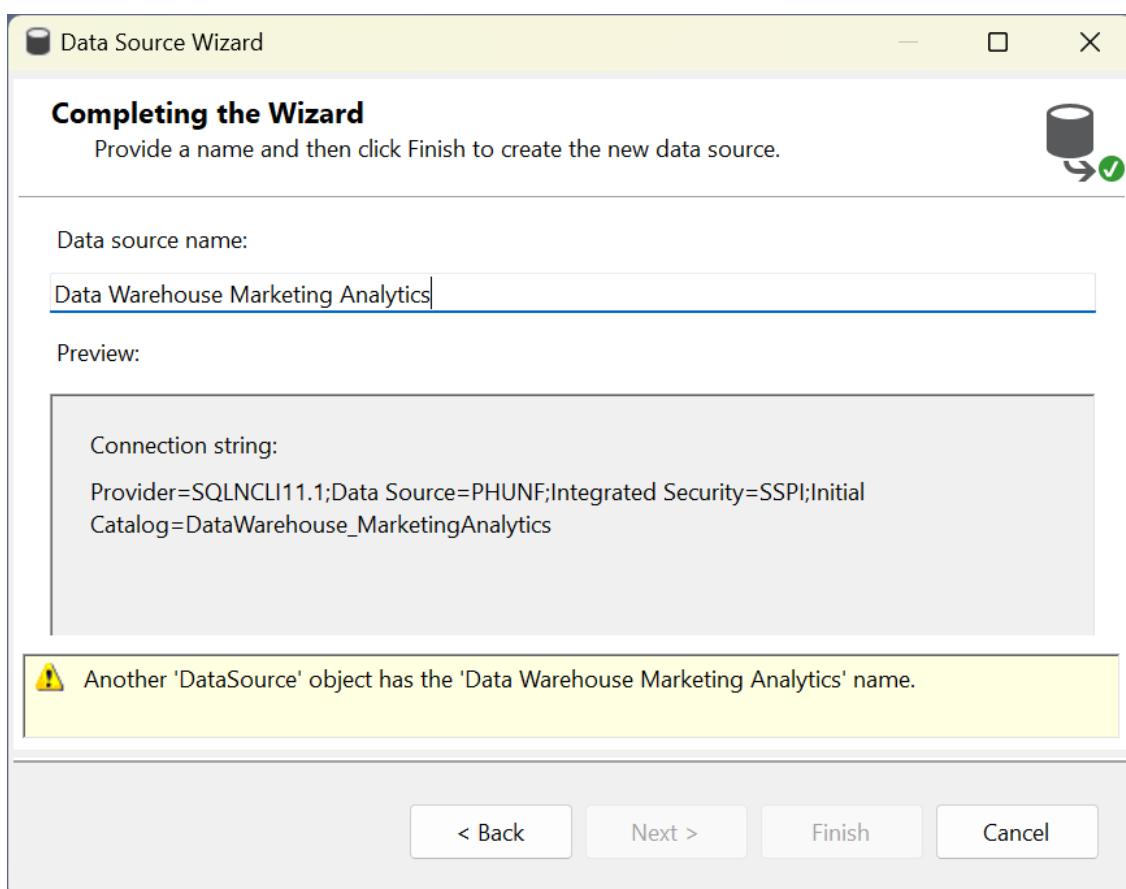
Bước 5: Khi đã chọn được database, sẽ hiển thị bảng như hình, nhấn Next để tiếp tục các bước tiếp theo



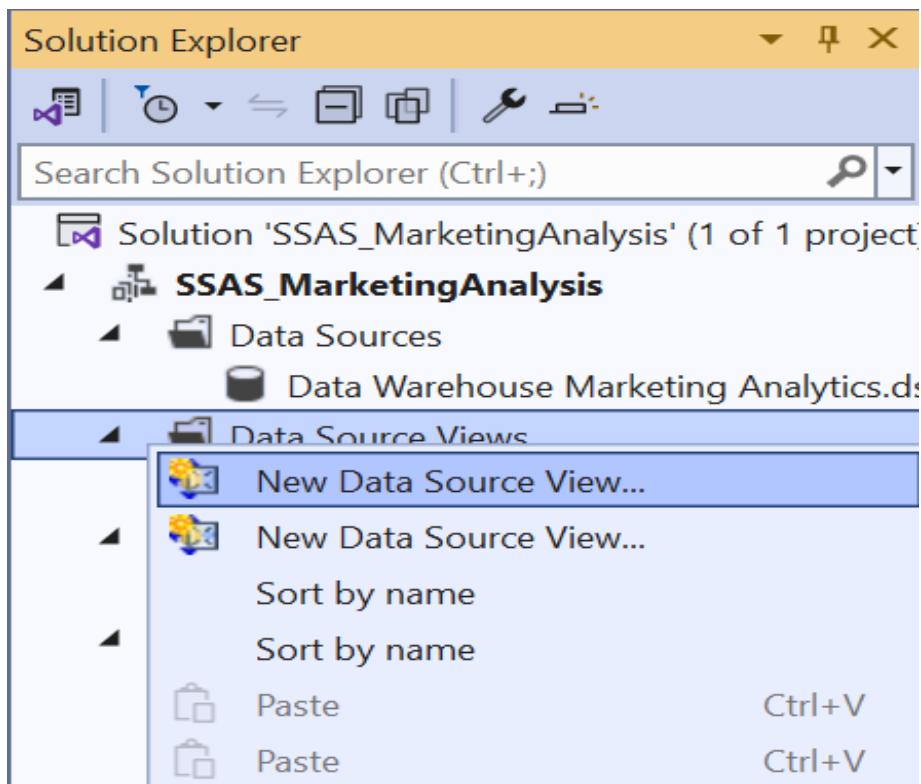
Bước 6: Click và “Use the service account” rồi nhấn Next



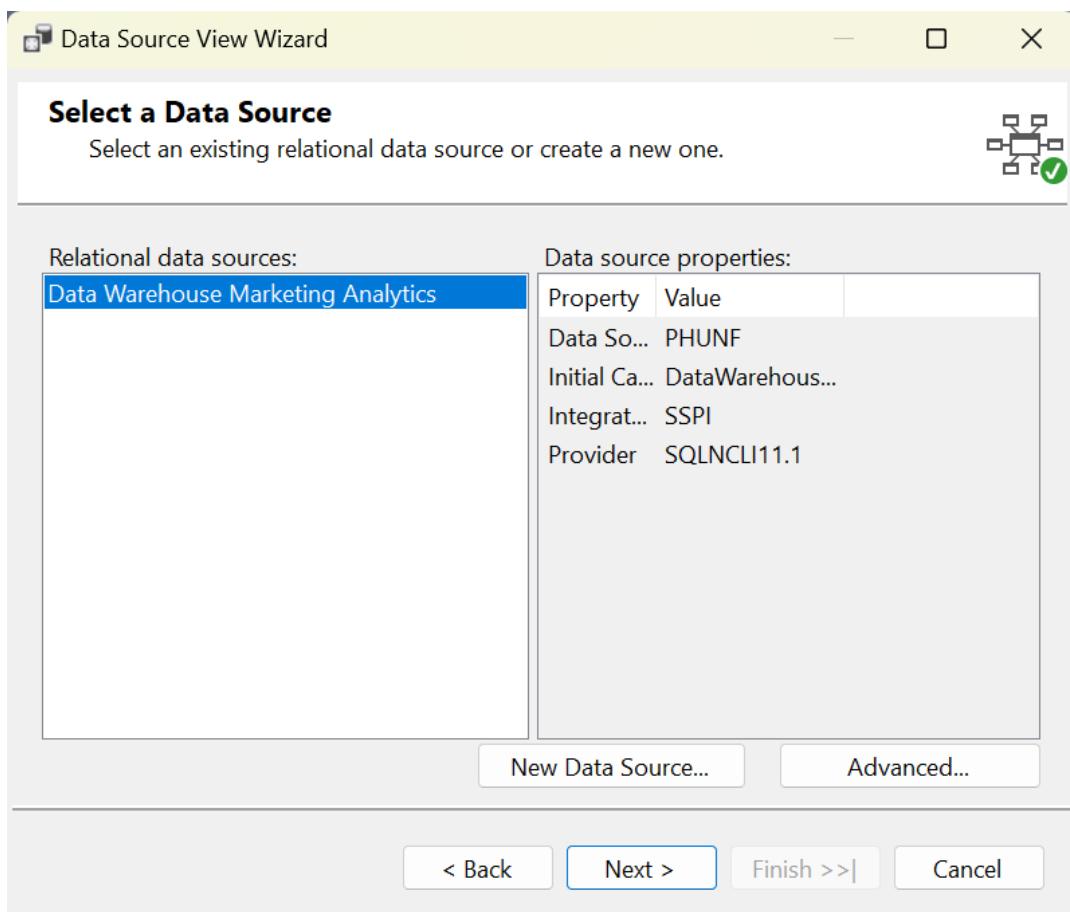
Bước 7: Cuối cùng sẽ báo thêm data source thành công, nhấn Finish để kết thúc



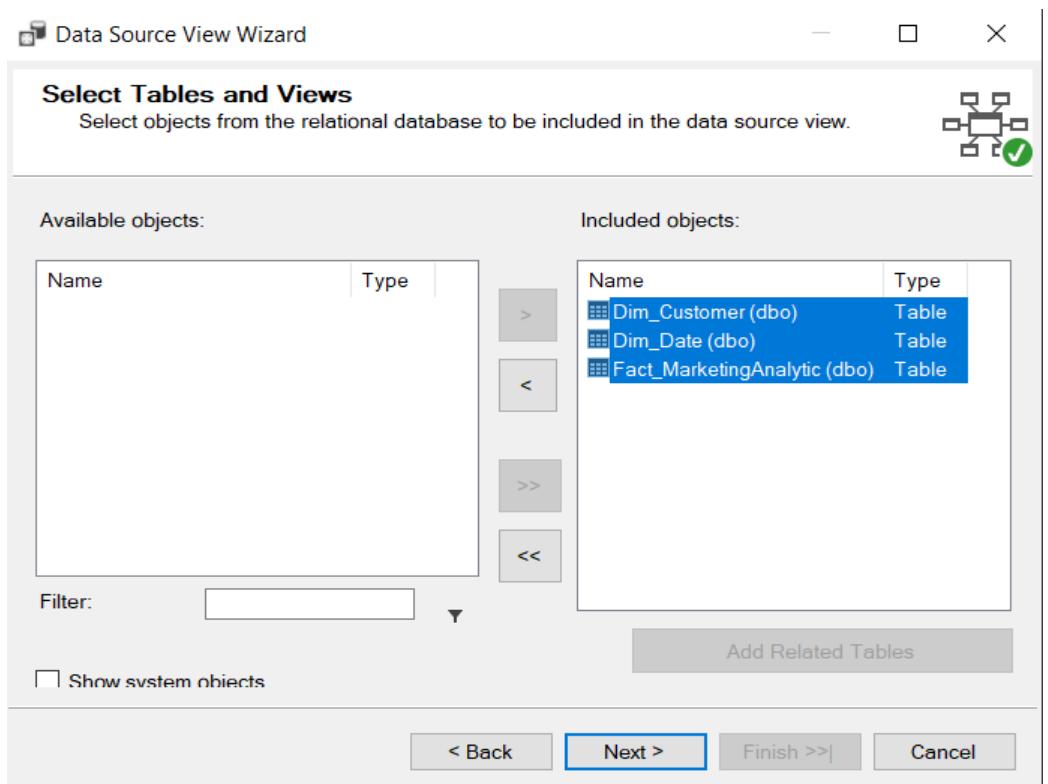
Bước 8: Tiếp theo, chuột phải vào Data Source Views, chọn New Data Source View



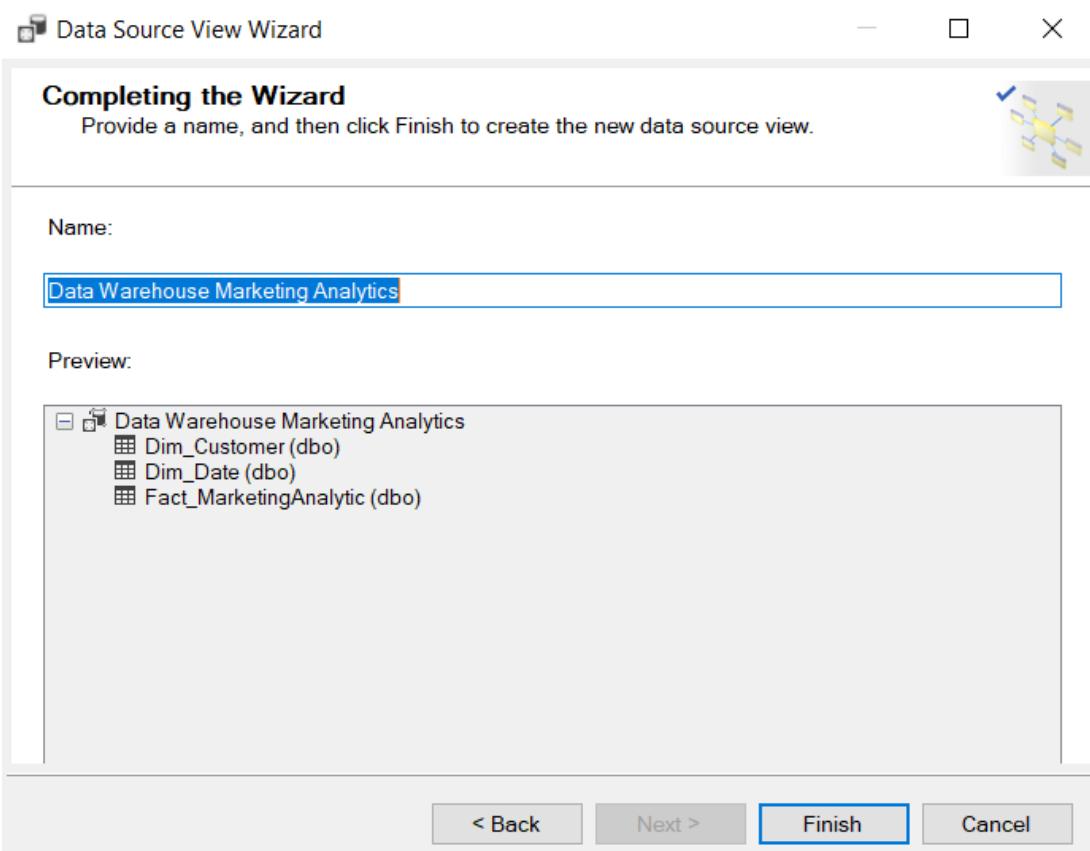
Bước 9: Vì đã thêm data source vào ở phía trên nên ở đây đã hiển thị sẵn, chọn Next

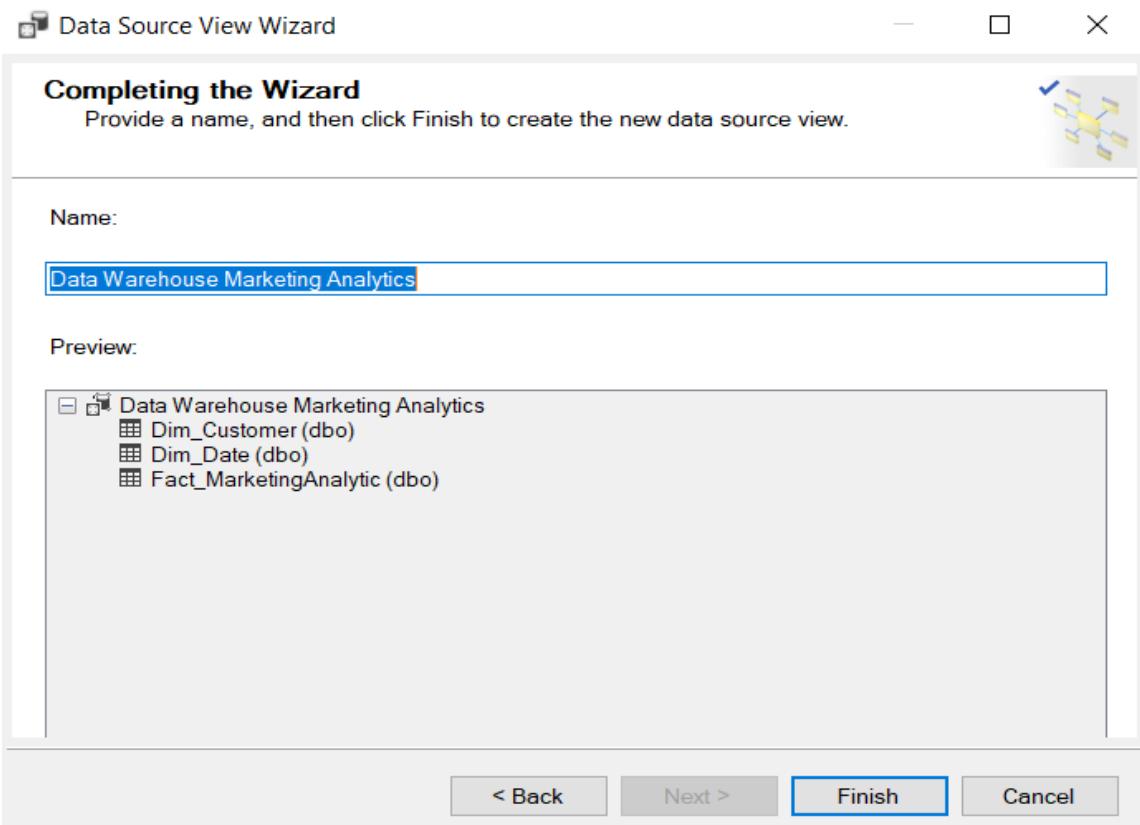


Bước 10: Chuyển những bảng cần sử dụng sang bên phải, rồi chọn Next



Bước 11: Đến đây là đã hoàn thành thêm database vào công cụ Visual Studio 2022, chọn Finish để kết thúc

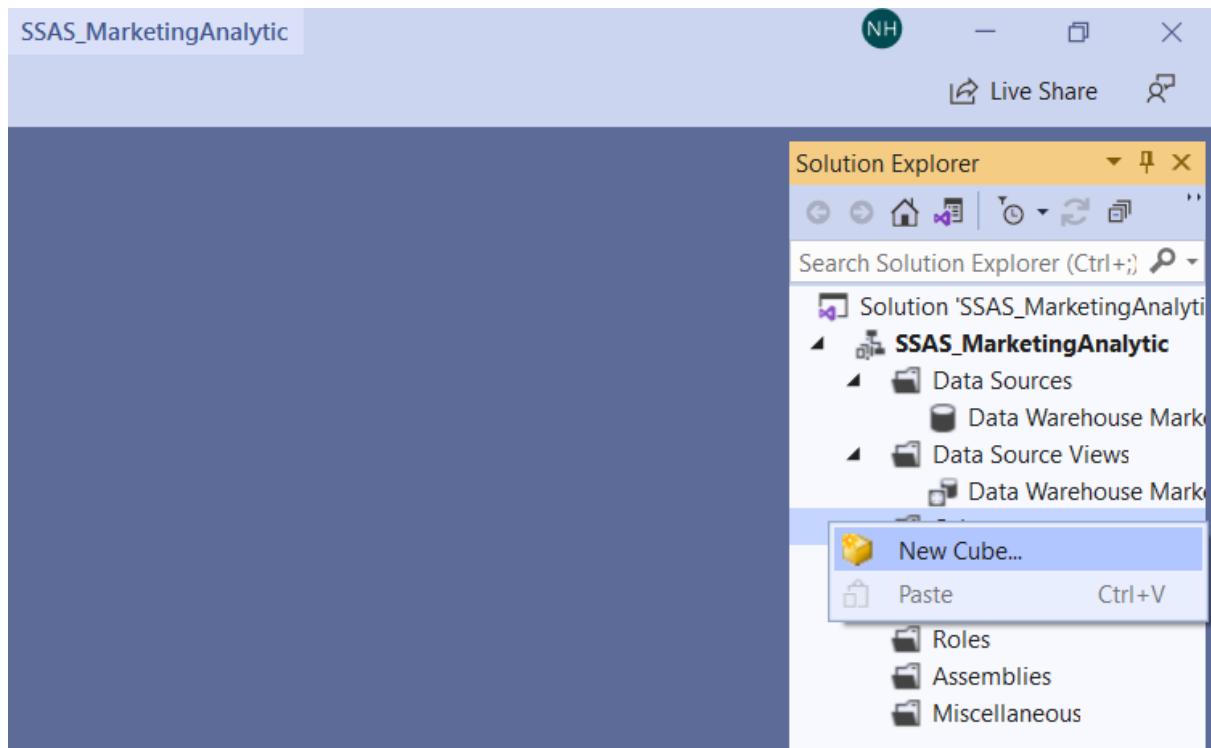




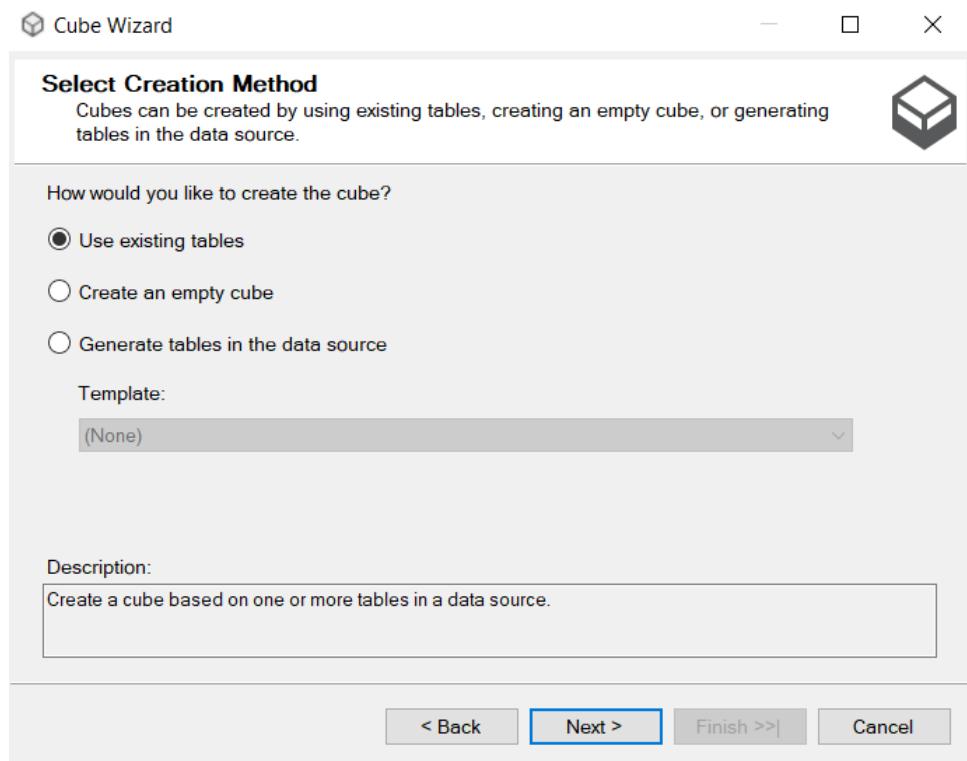
3.1.2 Tạo và deploy Cube

1. Tạo Cube

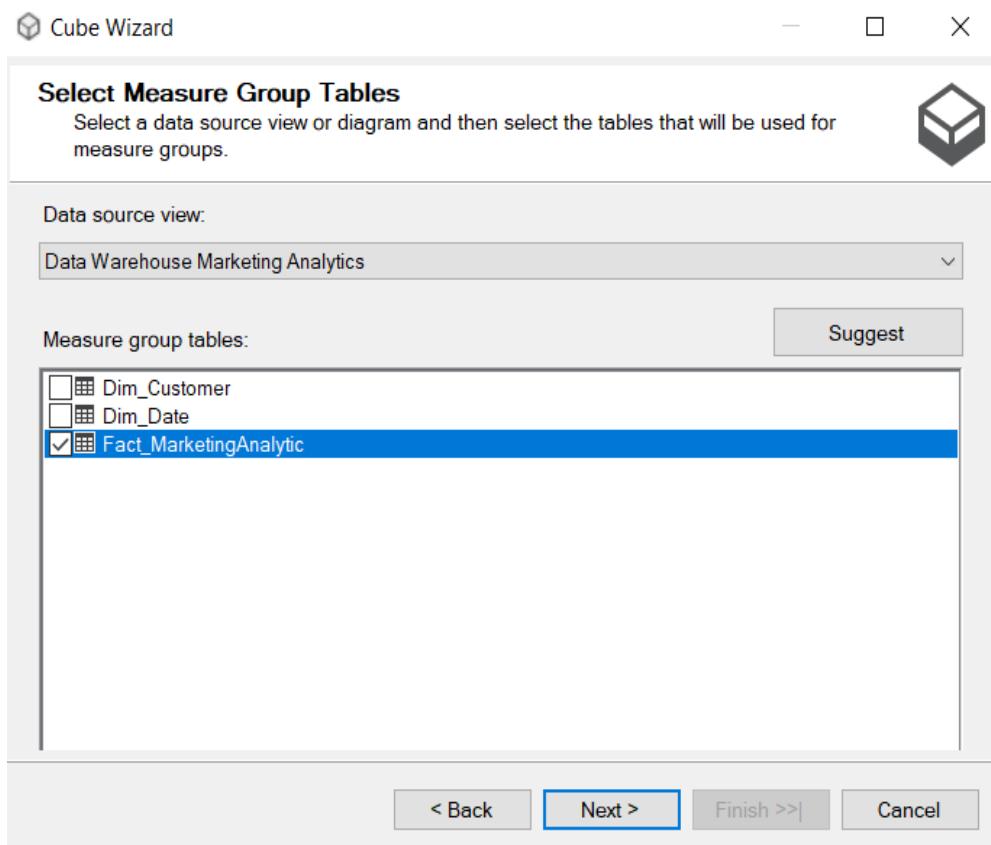
Bước 1: Chuột phải vào Cubes, chọn New Cube

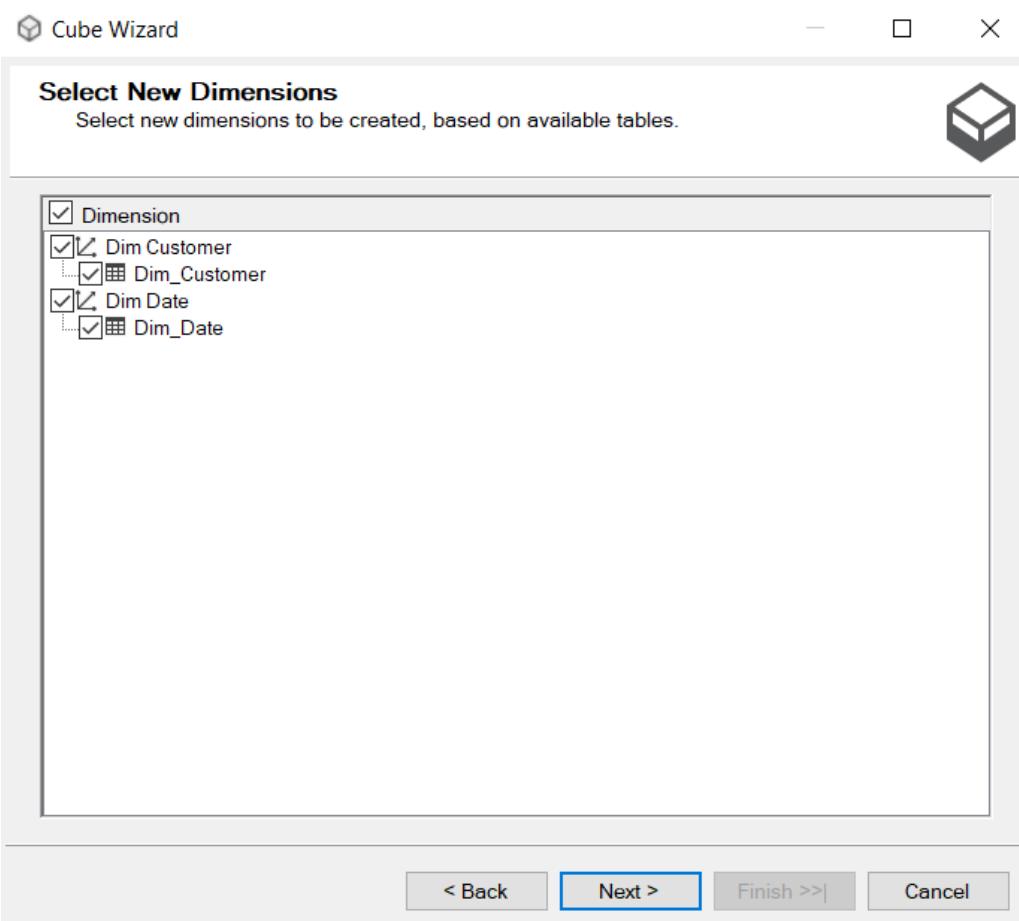
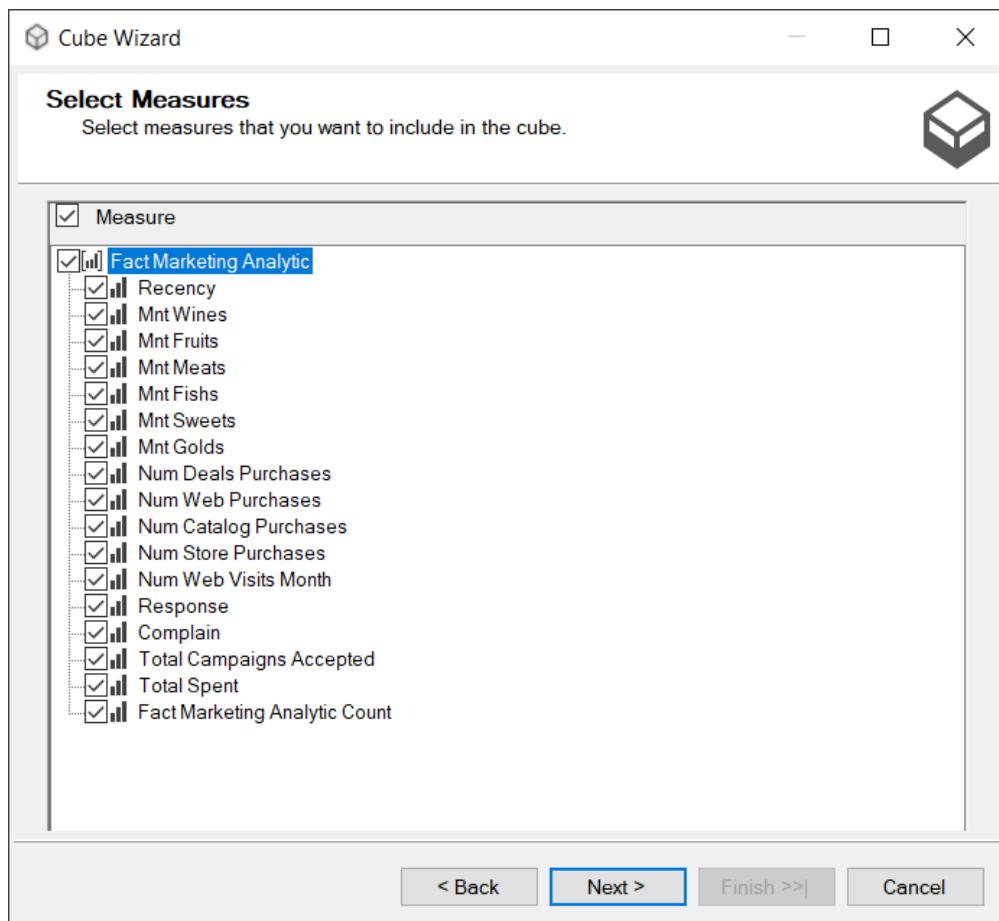


Bước 2: Chọn “ Using existing tables” rồi nhấn Next

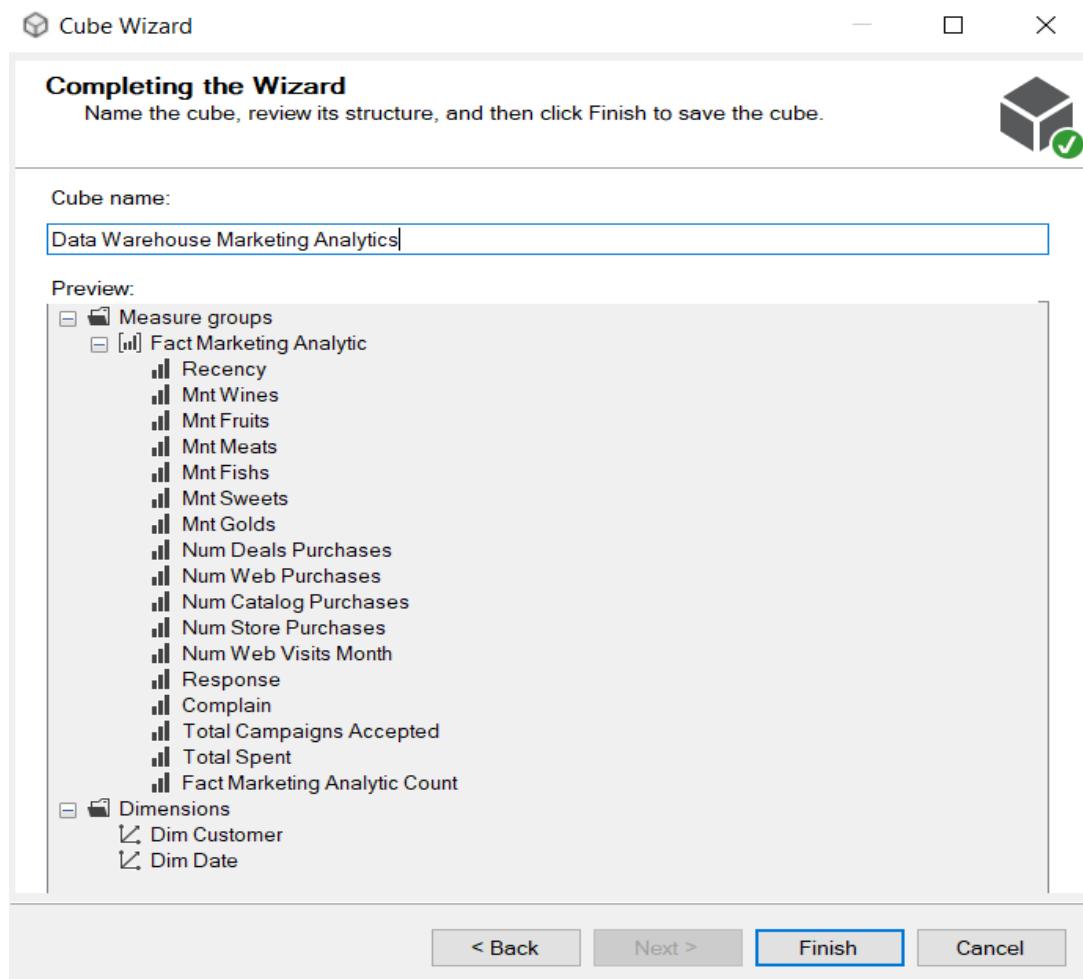


Bước 3: Tick vào bảng Fact và nhấn Next

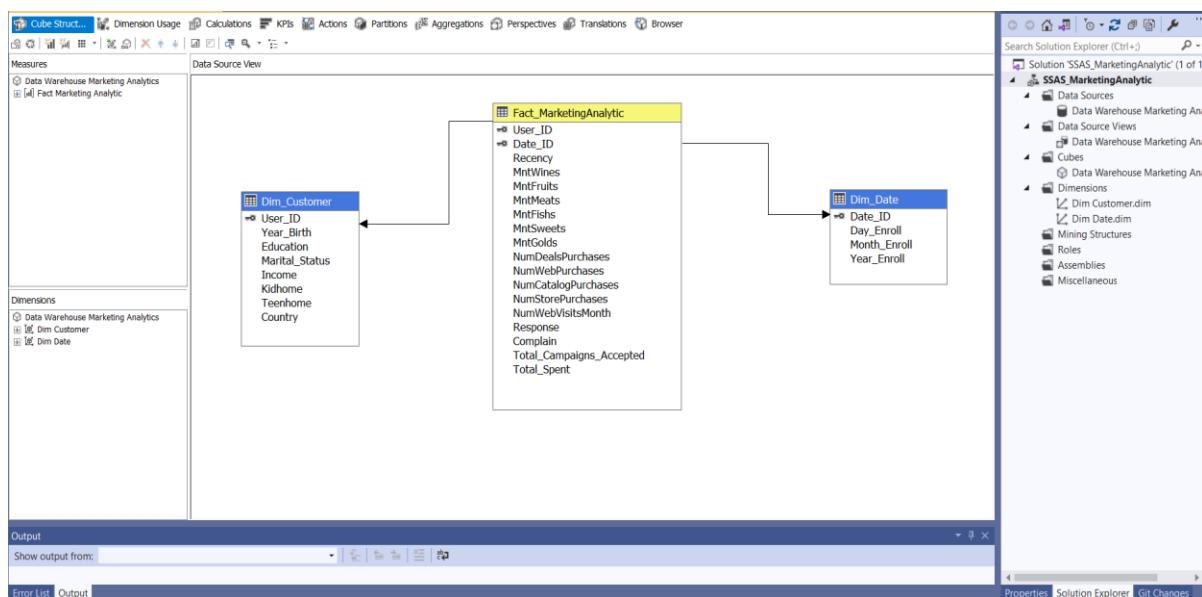




Bước 4: Nhấn Finish để hoàn thành tạo cube



Bước 5: Kết thúc quá trình Data Source Views kết quả nhận được như hình



Bước 6: Kéo thả các thuộc tính từ data source sang attributes

The screenshot shows the Data Warehouse Mar..tics.cube [Design] interface. The top navigation bar includes tabs for Dim Date.dim [Design], Dim Customer.dim [Design], and Data Warehouse Mar..tics.cube [Design]. Below the tabs are buttons for Dimension Structure, Attribute Relationships, Translations, and Browser.

The main area is divided into three columns:

- Attributes:** A tree view showing the Dim Customer dimension with attributes: Country, Education, Income, Kidhome, Marital Status, Teenhome, User ID, and Year Birth.
- Hierarchies:** A placeholder area with the instruction: "To create a new hierarchy, drag an attribute here."
- Data Source View:** A preview window titled "Dim_Customer" showing the same attributes: User_ID, Year_Birth, Education, Marital_Status, Income, Kidhome, Teenhome, and Country.

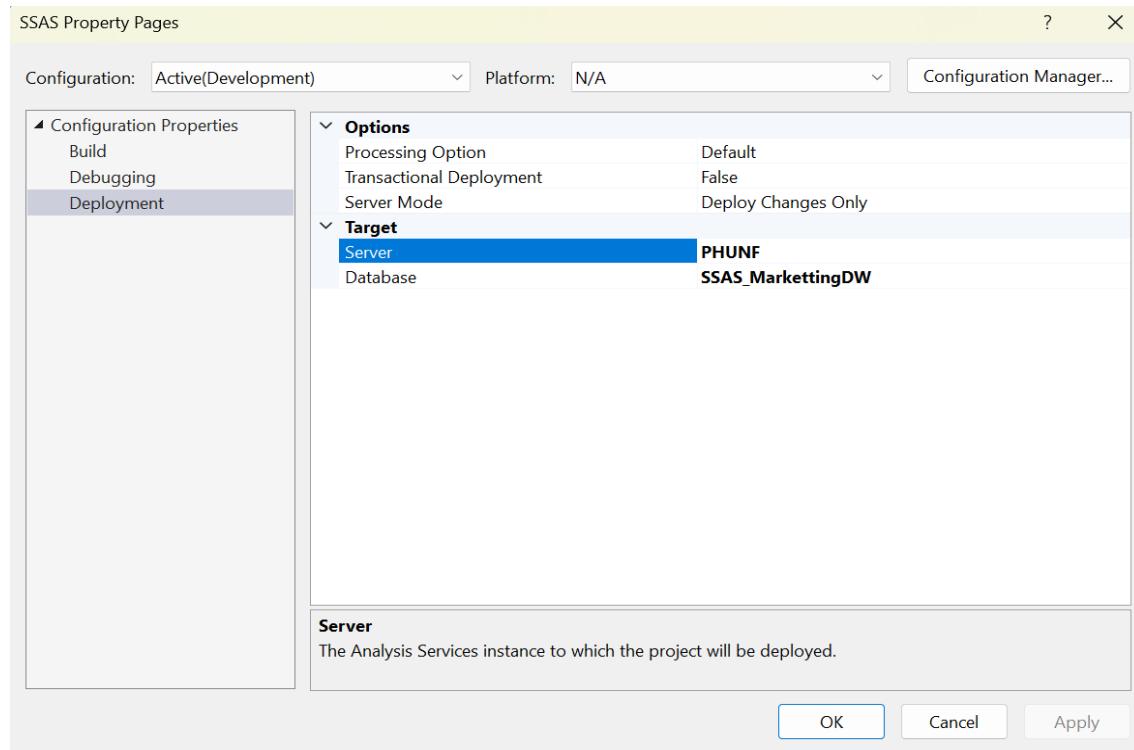
The screenshot shows the Data Warehouse Mar..tics.cube [Design] interface. The top navigation bar includes tabs for Dim Date.dim [Design], Dim Customer.dim [Design], and Data Warehouse Mar..tics.cube [Design]. Below the tabs are buttons for Dimension Structure, Attribute Relationships, Translations, and Browser.

The main area is divided into three columns:

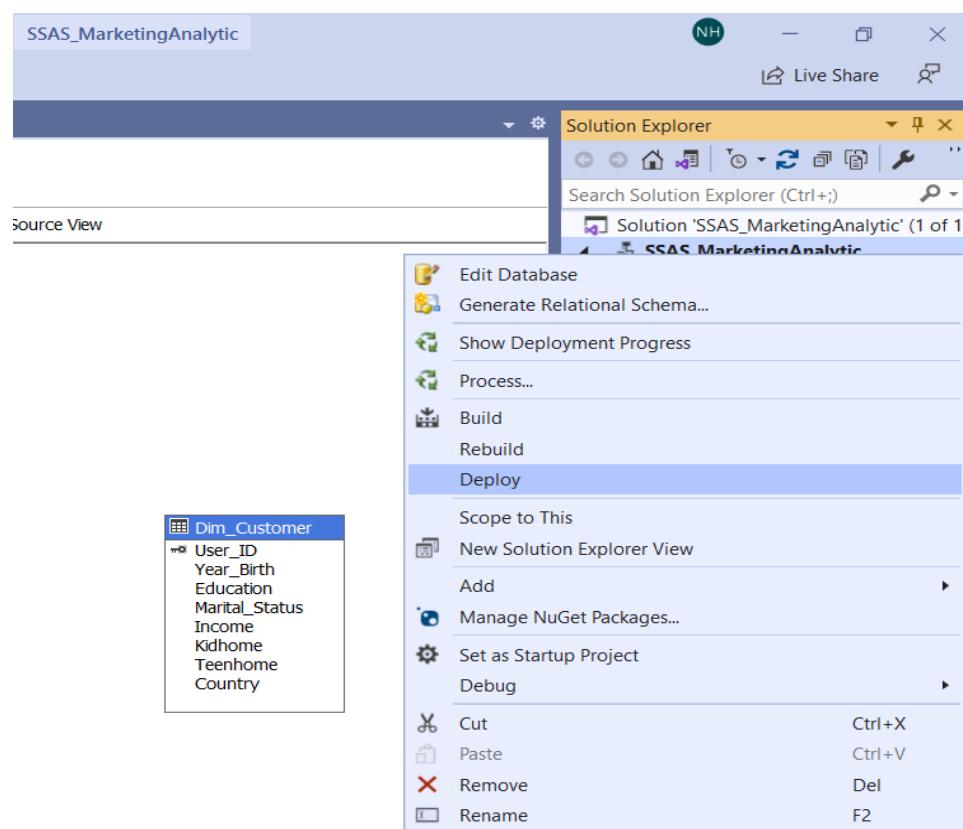
- Attributes:** A tree view showing the Dim Date dimension with attributes: Date_ID, Day_Enroll, Month_Enroll, and Year_Enroll.
- Hierarchies:** A placeholder area with the instruction: "To create a new hierarchy, drag an attribute here."
- Data Source View:** A preview window titled "Dim_Date" showing the attributes: Date_ID, Day_Enroll, Month_Enroll, and Year_Enroll.

2. Deploy Cube

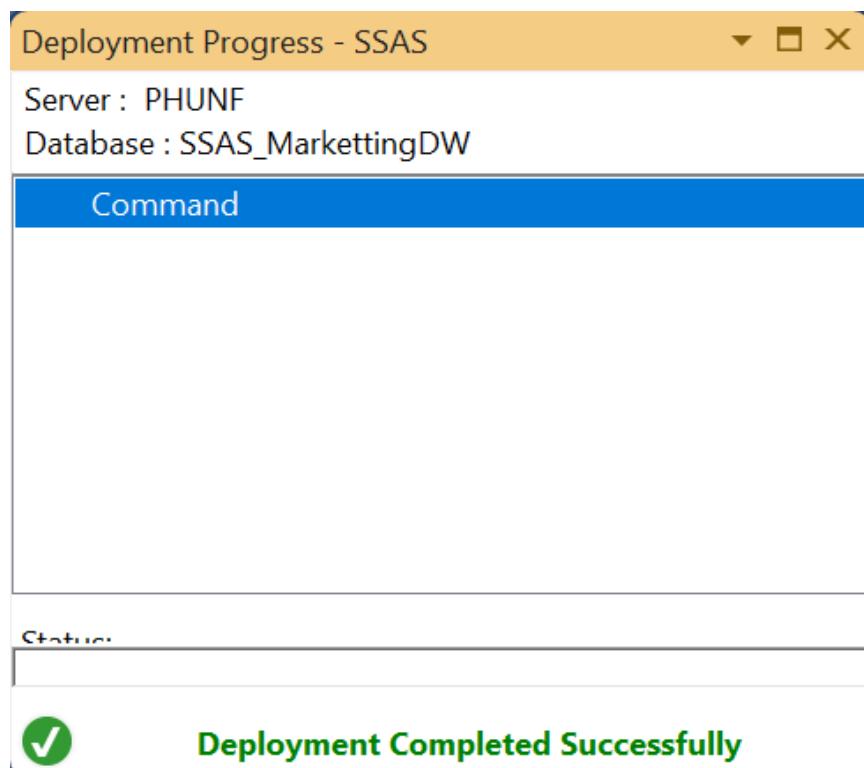
Bước 1: Chuột phải vào solution, chọn Properties ở mục Deployment đổi tên Server từ localhost sang server name của database



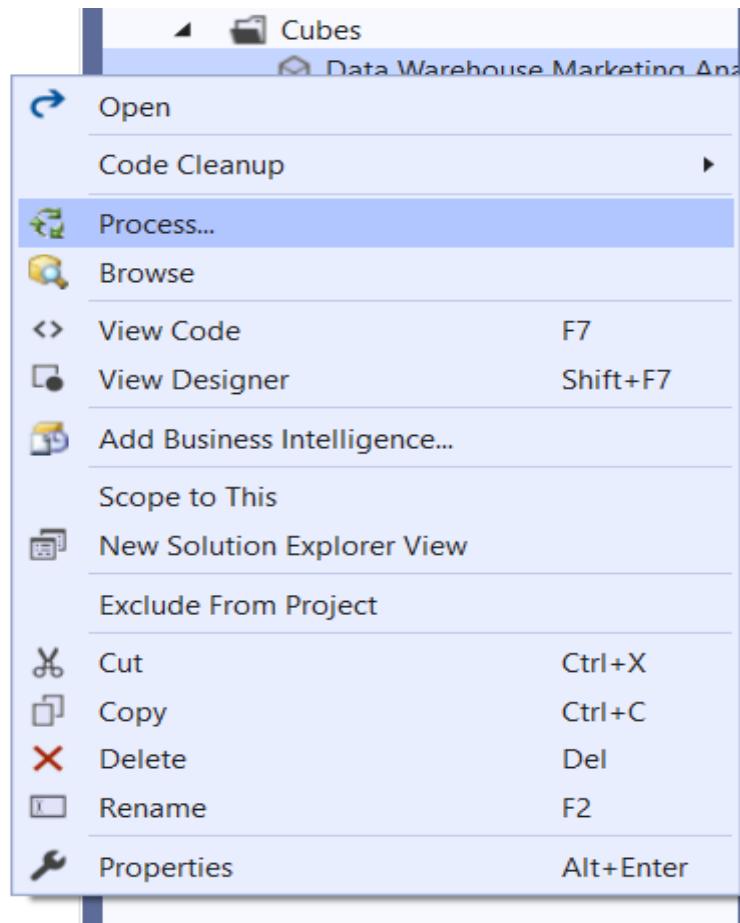
Bước 2: Tiếp tục chuột phải vào solution, chọn deploy



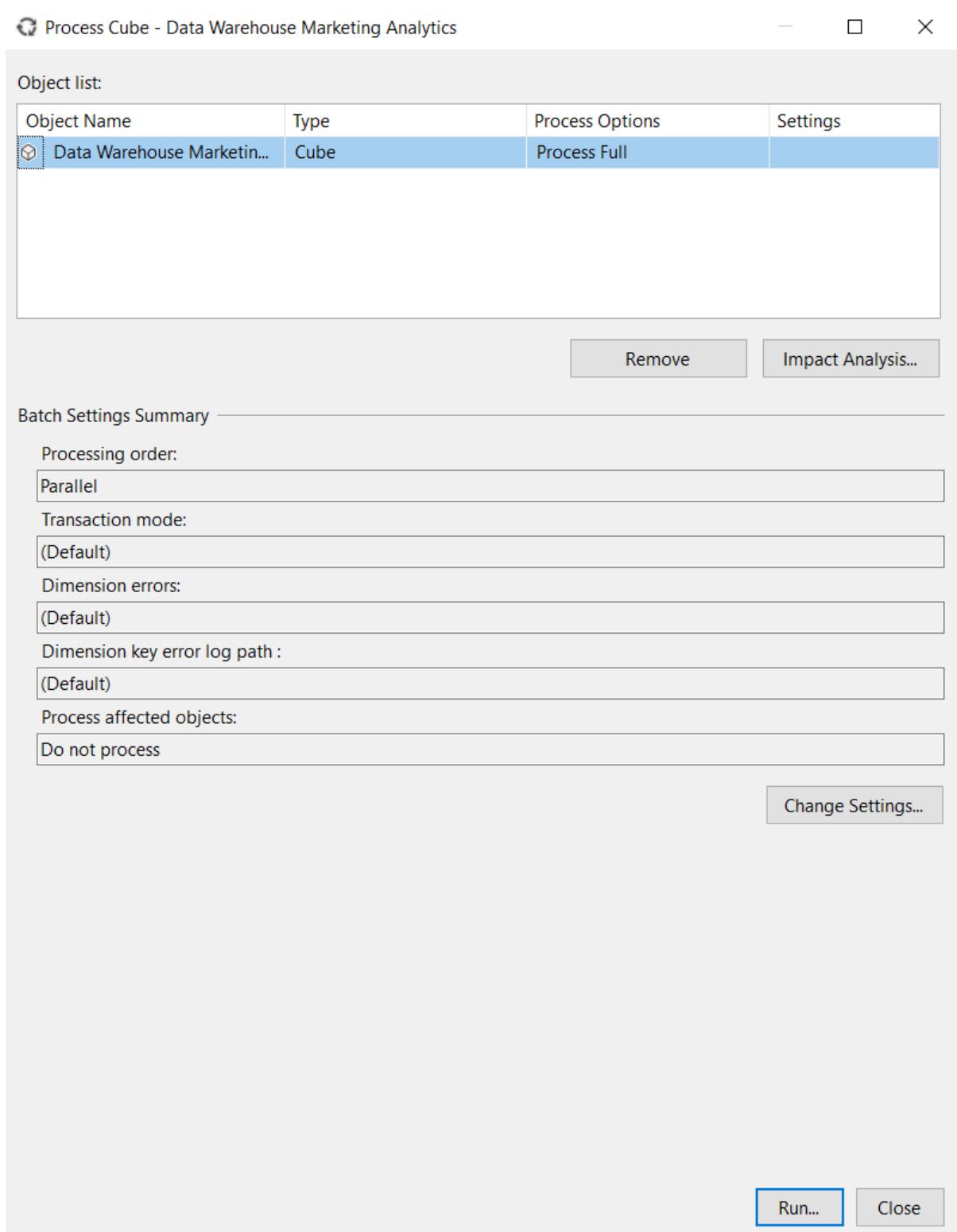
Bước 3: Sau khi deploy thành công, sẽ hiển thị như hình dưới

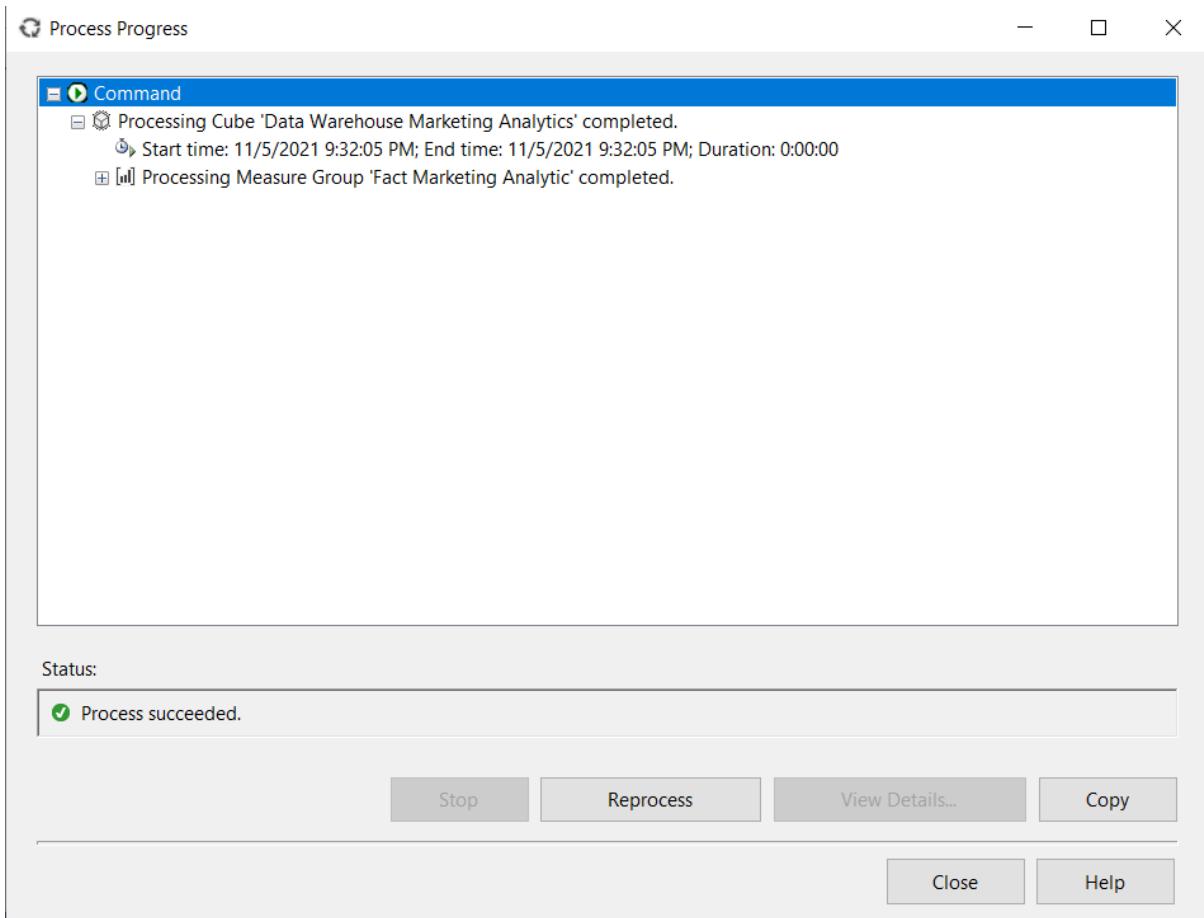


Bước 4: Sau đó chuột phải vào cube vừa tạo, chọn process để chạy quá trình



Bước 5: Nhấn Run để bắt đầu chạy, hiển thị như hình dưới là process thành công





3.1.3 Tạo thêm các Measures và Hierachies

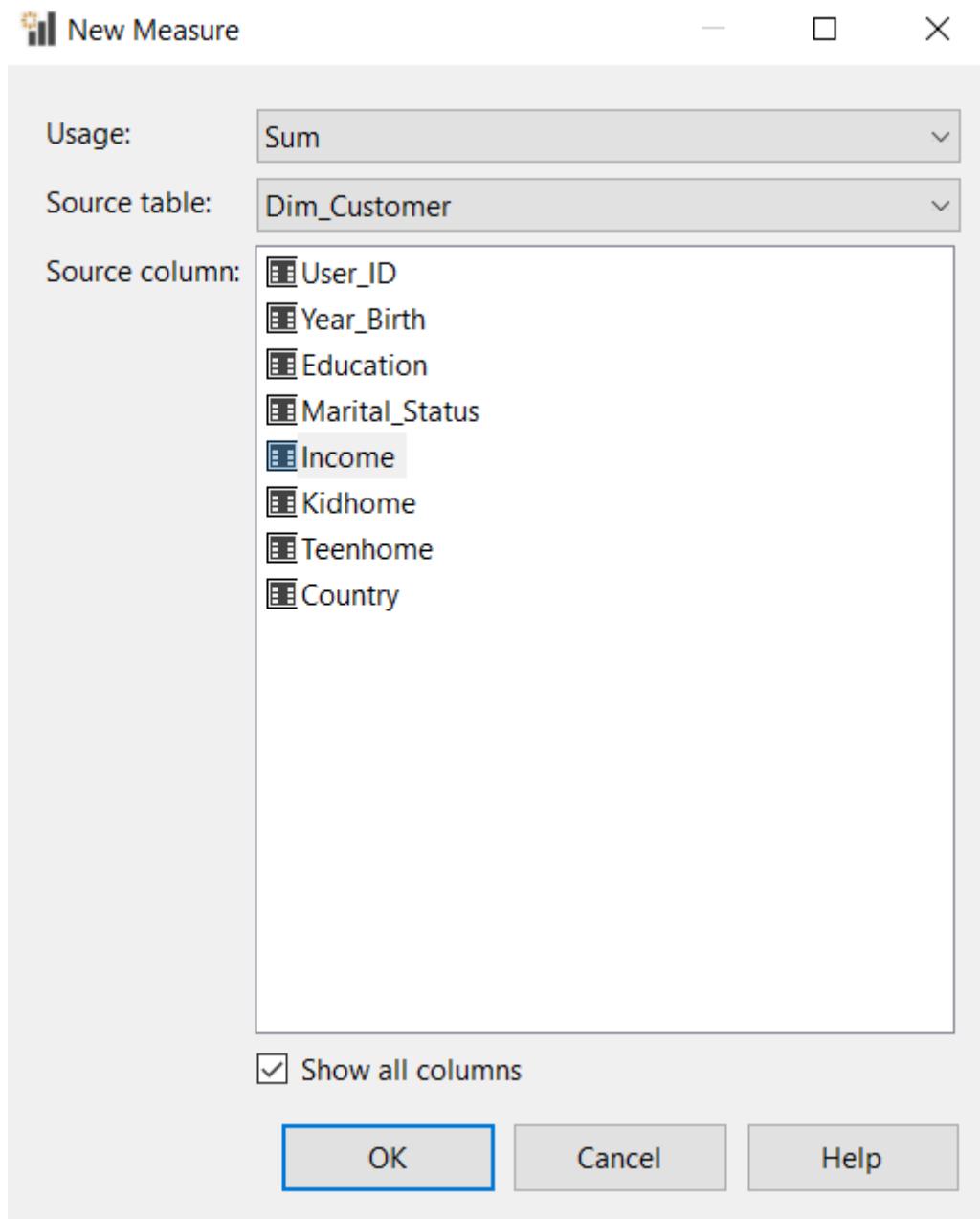
1. Tạo Measures

Bước 1: Ở Cube Struct, chọn vào icon Show Measures Grid, chuột phải vào rồi chọn New Measure

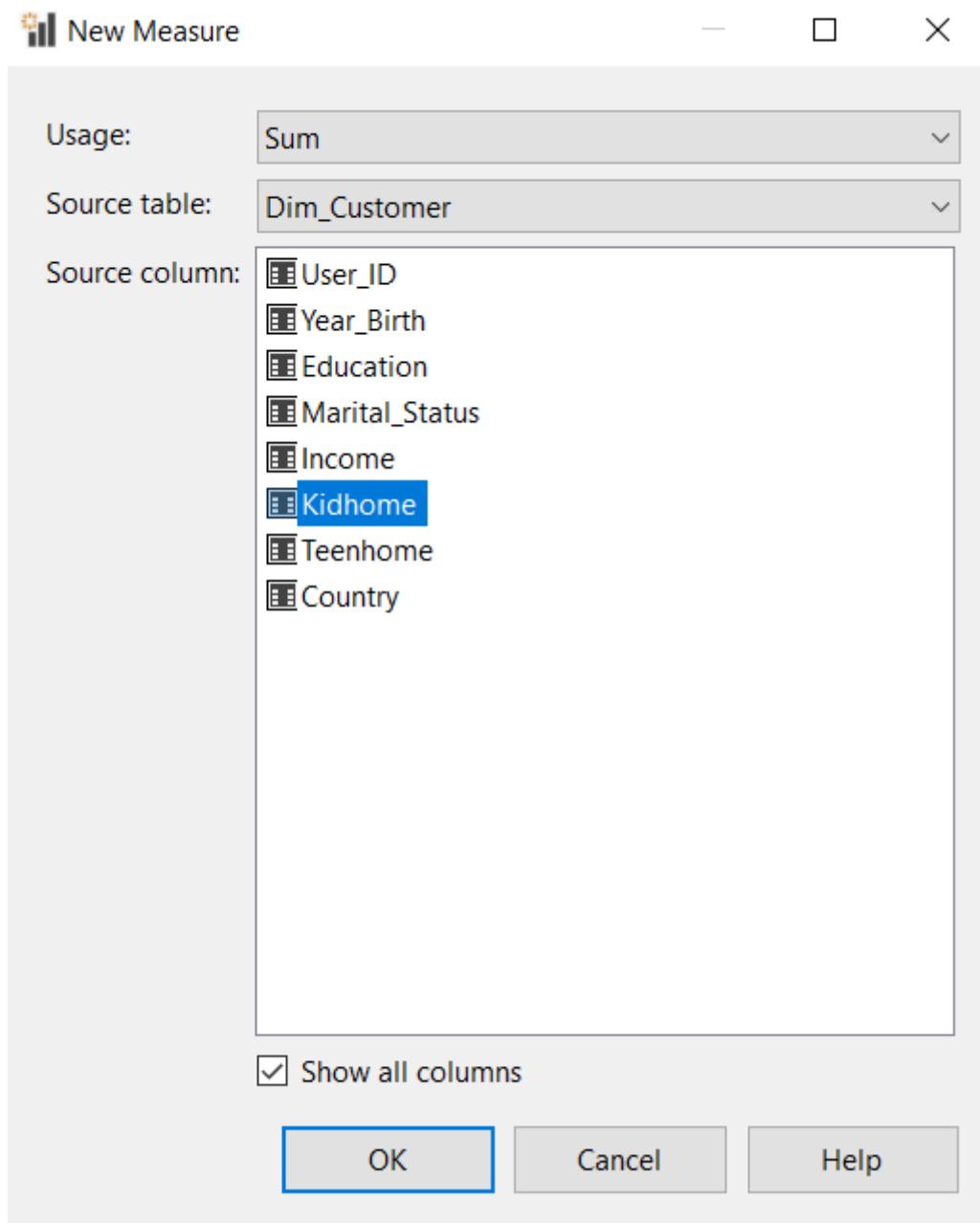
The screenshot shows the 'Cube Struct...' tab selected in the top navigation bar. Below it is a toolbar with various icons. The main area is titled 'Measures' and contains a grid of 18 rows. Each row has a small icon, a name, a 'Measure Group' column, a 'Data Type' column, and an 'Aggregation' column. A context menu is open over the last row, listing options: 'New Measure...', 'Edit Measure...', 'New Measure Group...', 'Show Measures in', and 'New Linked Object...'. The 'New Measure...' option is highlighted.

	Name	Measure Group	Data Type	Aggregation
1	Recency		Integer	Sum
2	Mnt Wines		Double	Sum
3	Mnt Fruits			
4	Mnt Meats			
5	Mnt Fishs			
6	Mnt Sweets			
7	Mnt Golds			
8	Num Deals Purchases			
9	Num Web Purchases		Integer	Sum
10	Num Catalog Purchases		Integer	Sum
11	Num Store Purchases		Integer	Sum
12	Num Web Visits Month		Integer	Sum
13	Response		Integer	Sum
14	Complain		Integer	Sum
15	Total Campaigns Accepted		Integer	Sum
16	Total Spent		Double	Sum
17	Fact Marketing Analytic Count		Integer	Count
18				

Bước 2: Tiếp tục, chọn measure là Income từ bảng Dim_Customer, để Aggregation là Sum, xong nhấn OK



Bước 3: Chọn measure là Kidhome từ bảng Dim_Customer, để Aggregation là Sum, xong nhấn OK



Bước 4: Kết thúc việc tạo sẽ thấy 2 measure mới ở trong bảng như hình dưới

Measures				
	Name	Measure Group	Data Type	Aggregation
Bar chart icon	Recency	Fact Marketing Analytic	Integer	Sum
Bar chart icon	Mnt Wines		Double	Sum
Bar chart icon	Mnt Fruits		Double	Sum
Bar chart icon	Mnt Meats		Double	Sum
Bar chart icon	Mnt Fishs		Double	Sum
Bar chart icon	Mnt Sweets		Double	Sum
Bar chart icon	Mnt Golds		Double	Sum
Bar chart icon	Num Deals Purchases		Integer	Sum
Bar chart icon	Num Web Purchases		Integer	Sum
Bar chart icon	Num Catalog Purchases		Integer	Sum
Bar chart icon	Num Store Purchases		Integer	Sum
Bar chart icon	Num Web Visits Month		Integer	Sum
Bar chart icon	Response		Integer	Sum
Bar chart icon	Complain		Integer	Sum
Bar chart icon	Total Campaigns Accepted		Integer	Sum
Bar chart icon	Total Spent		Double	Sum
Bar chart icon	Fact Marketing Analytic Count		Integer	Count
Bar chart icon	Income		Double	Sum
Bar chart icon	Kidhome		Integer	Sum

Bước 5: Tiếp theo, Format lại dữ liệu hiển thị của các thuộc tính Mnt Wines, Mnt Fruits, Mnt Meats, Mnt Fishs, Mnt Sweets, Mnt Gold và Income

Measures					Properties
Name	Measure Group	Data Type	Aggregation		
Recency		Integer	Sum		
Mnt Wines	Fact Marketing Analytic	Double	Sum	Mnt Wines Measure	
Mnt Fruits		Double	Sum	AggregateFunction	Sum
Mnt Meats		Double	Sum	DataType	Inherited
Mnt Fishs		Double	Sum	Description	
Mnt Sweets		Double	Sum	DisplayFolder	
Mnt Golds		Double	Sum	FormatString	Currency
Num Deals Purchases		Integer	Sum	ID	Mnt Wines
Num Web Purchases		Integer	Sum	MeasureExpression	
Num Catalog Purchases		Integer	Sum	Name	Mnt Wines
Num Store Purchases		Integer	Sum	Source	Fact_MarketingAnalytic.MntWines
Num Web Visits Month		Integer	Sum	Visible	True
Response		Integer	Sum		
Complain		Integer	Sum		
Total Campaigns Accepted		Integer	Sum		
Total Spent		Double	Sum		
Fact Marketing Analytic Count		Integer	Count		
Income		Double	Sum		
Kidhome		Integer	Sum		

Measures					Properties
Name	Measure Group	Data Type	Aggregation		
Recency		Integer	Sum	Mnt Fruits Measure	
Mnt Wines		Double	Sum	AggregateFunction	Sum
Mnt Fruits	Fact Marketing Analytic	Double	Sum	DataType	Inherited
Mnt Meats		Double	Sum	Description	
Mnt Fishs		Double	Sum	DisplayFolder	
Mnt Sweets		Double	Sum	FormatString	Currency
Mnt Golds		Double	Sum	ID	Mnt Fruits
Num Deals Purchases		Integer	Sum	MeasureExpression	
Num Web Purchases		Integer	Sum	Name	Mnt Fruits
Num Catalog Purchases		Integer	Sum	Source	Fact_MarketingAnalytic.MntFruits
Num Store Purchases		Integer	Sum	Visible	True
Num Web Visits Month		Integer	Sum		
Response		Integer	Sum		
Complain		Integer	Sum		
Total Campaigns Accepted		Integer	Sum		
Total Spent		Double	Sum		
Fact Marketing Analytic Count		Integer	Count		
Income		Double	Sum		
Kidhome		Integer	Sum		

Measures

Name	Measure Group	Data Type	Aggregation
Recency		Integer	Sum
Mnt Wines		Double	Sum
Mnt Fruits		Double	Sum
Mnt Meats	Fact Marketing Analytic	Double	Sum
Mnt Fishs		Double	Sum
Mnt Sweets		Double	Sum
Mnt Golds		Double	Sum
Num Deals Purchases		Integer	Sum
Num Web Purchases		Integer	Sum
Num Catalog Purchases		Integer	Sum
Num Store Purchases		Integer	Sum
Num Web Visits Month		Integer	Sum
Response		Integer	Sum
Complain		Integer	Sum
Total Campaigns Accepted		Integer	Sum
Total Spent		Double	Sum
Fact Marketing Analytic Count		Integer	Count
Income		Double	Sum
Kidhome		Integer	Sum

Properties

Mnt Meats Measure

AggregateFunction	Sum
DataType	Inherited
Description	
DisplayFolder	
FormatString	Currency
ID	Mnt Meats
MeasureExpression	
Name	Mnt Meats
Source	Fact_MarketingAnalytic.MntMeats
Visible	True

Measures

Name	Measure Group	Data Type	Aggregation
Recency		Integer	Sum
Mnt Wines		Double	Sum
Mnt Fruits		Double	Sum
Mnt Meats	Fact Marketing Analytic	Double	Sum
Mnt Fishs		Double	Sum
Mnt Sweets		Double	Sum
Mnt Golds		Double	Sum
Num Deals Purchases		Integer	Sum
Num Web Purchases		Integer	Sum
Num Catalog Purchases		Integer	Sum
Num Store Purchases		Integer	Sum
Num Web Visits Month		Integer	Sum
Response		Integer	Sum
Complain		Integer	Sum
Total Campaigns Accepted		Integer	Sum
Total Spent		Double	Sum
Fact Marketing Analytic Count		Integer	Count
Income		Double	Sum
Kidhome		Integer	Sum

Properties

Mnt Fishs Measure

AggregateFunction	Sum
DataType	Inherited
Description	
DisplayFolder	
FormatString	Currency
ID	Mnt Fishs
MeasureExpression	
Name	Mnt Fishs
Source	Fact_MarketingAnalytic.MntFishes
Visible	True

Measures

Name	Measure Group	Data Type	Aggregation
Recency		Integer	Sum
Mnt Wines		Double	Sum
Mnt Fruits		Double	Sum
Mnt Meats		Double	Sum
Mnt Fishs		Double	Sum
Mnt Sweets	Fact Marketing Analytic	Double	Sum
Mnt Golds		Double	Sum
Num Deals Purchases		Integer	Sum
Num Web Purchases		Integer	Sum
Num Catalog Purchases		Integer	Sum
Num Store Purchases		Integer	Sum
Num Web Visits Month		Integer	Sum
Response		Integer	Sum
Complain		Integer	Sum
Total Campaigns Accepted		Integer	Sum
Total Spent		Double	Sum
Fact Marketing Analytic Count		Integer	Count
Income		Double	Sum
Kidhome		Integer	Sum

Properties

Mnt Sweets Measure

AggregateFunction	Sum
DataType	Inherited
Description	
DisplayFolder	
FormatString	Currency
ID	Mnt Sweets
MeasureExpression	
Name	Mnt Sweets
Source	Fact_MarketingAnalytic.MntSweets
Visible	True

Measures

Name	Measure Group	Data Type	Aggregation
Recency		Integer	Sum
Mnt Wines		Double	Sum
Mnt Fruits		Double	Sum
Mnt Meats		Double	Sum
Mnt Fishs		Double	Sum
Mnt Sweets		Double	Sum
Mnt Golds	Fact Marketing Analytic	Double	Sum
Num Deals Purchases		Integer	Sum
Num Web Purchases		Integer	Sum
Num Catalog Purchases		Integer	Sum
Num Store Purchases		Integer	Sum
Num Web Visits Month		Integer	Sum
Response		Integer	Sum
Complain		Integer	Sum
Total Campaigns Accepted		Integer	Sum
Total Spent		Double	Sum
Fact Marketing Analytic Count		Integer	Count
Income		Double	Sum
Kidhome		Integer	Sum

Properties

Mnt Golds Measure	
AggregateFunction	Sum
DataType	Inherited
Description	
DisplayFolder	
FormatString	Currency
ID	Mnt Golds
MeasureExpression	
Name	Mnt Golds
Source	Fact_MarketingAnalytic.MntGolds
Visible	True

Measures

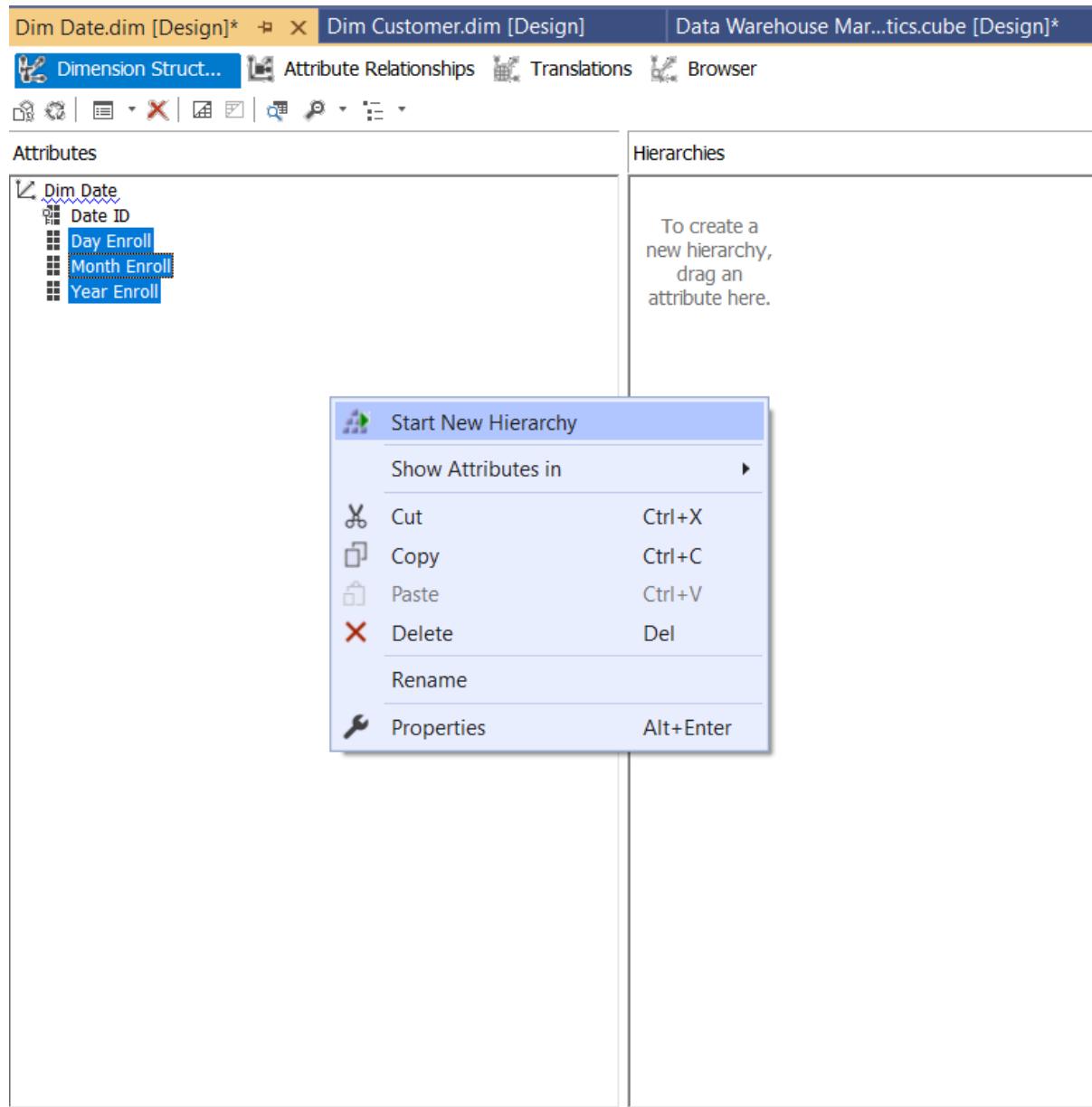
Name	Measure Group	Data Type	Aggregation
Recency		Integer	Sum
Mnt Wines		Double	Sum
Mnt Fruits		Double	Sum
Mnt Meats		Double	Sum
Mnt Fishs		Double	Sum
Mnt Sweets		Double	Sum
Mnt Golds		Double	Sum
Num Deals Purchases		Integer	Sum
Num Web Purchases		Integer	Sum
Num Catalog Purchases		Integer	Sum
Num Store Purchases		Integer	Sum
Num Web Visits Month		Integer	Sum
Response		Integer	Sum
Complain		Integer	Sum
Total Campaigns Accepted		Integer	Sum
Total Spent		Double	Sum
Fact Marketing Analytic Count		Integer	Count
Income	Dim Customer	Double	Sum
Kidhome		Integer	Sum

Properties

Income Measure	
AggregateFunction	Sum
DataType	Inherited
Description	
DisplayFolder	
FormatString	Currency
ID	Income
MeasureExpression	
Name	Income
Source	Dim_Customer.Income (D)
Visible	True

2. Tạo hierarchies

Bước 1: Chuyển sang tab Dim Date để thực hiện phân cấp ngày tháng năm. Chuột phải vào bảng Attributes chọn Start New Hierarchy

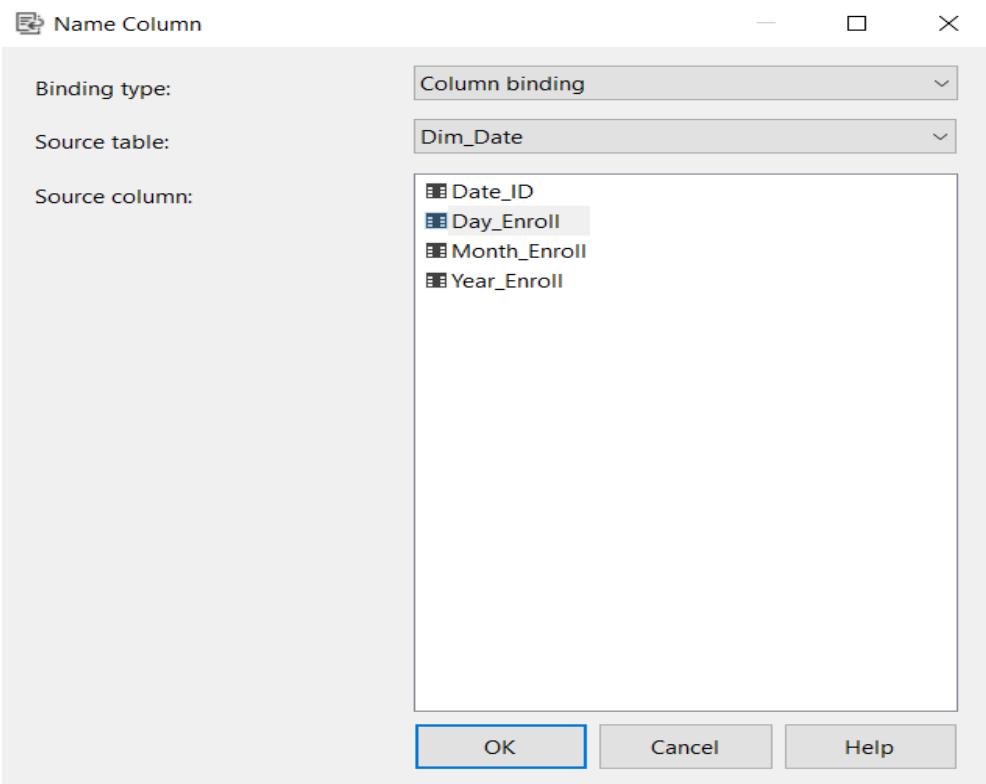


Bước 2: Chỉnh các thuộc tính theo level mong muốn: Year Enroll -> Month Enroll -> Day Enroll

Bước 3: Chuột phải vào thuộc tính Day_Enroll, chọn Properties , ở mục NameColumn chọn vào 3 dấu chấm ở cuối hàng

Properties	
Day Enroll	DimensionAttribute
HasLineage	False
ID	Day Enroll
InstanceSelection	None
IsAggregatable	True
KeyColumns	Dim_Date.Day_Enroll (Integer)
MemberNamesUnique	False
MembersWithData	NonLeafDataVisible
MembersWithDataCaption	
Name	Day Enroll
NameColumn	Dim_Date.Day_Enroll (WChar)
NamingTemplate	
OrderBy	
OrderByAttribute	
ProcessingState	Unprocessed
RootMemberIf	ParentIsBlankSelfOrMissing
TokenizationBehavior	TokenizationNone
Type	Regular
UnaryOperatorColumn	(none)

Bước 4: Chọn Day_Enroll rồi nhấn OK, tiếp tục ở mục OrderBy chọn Key



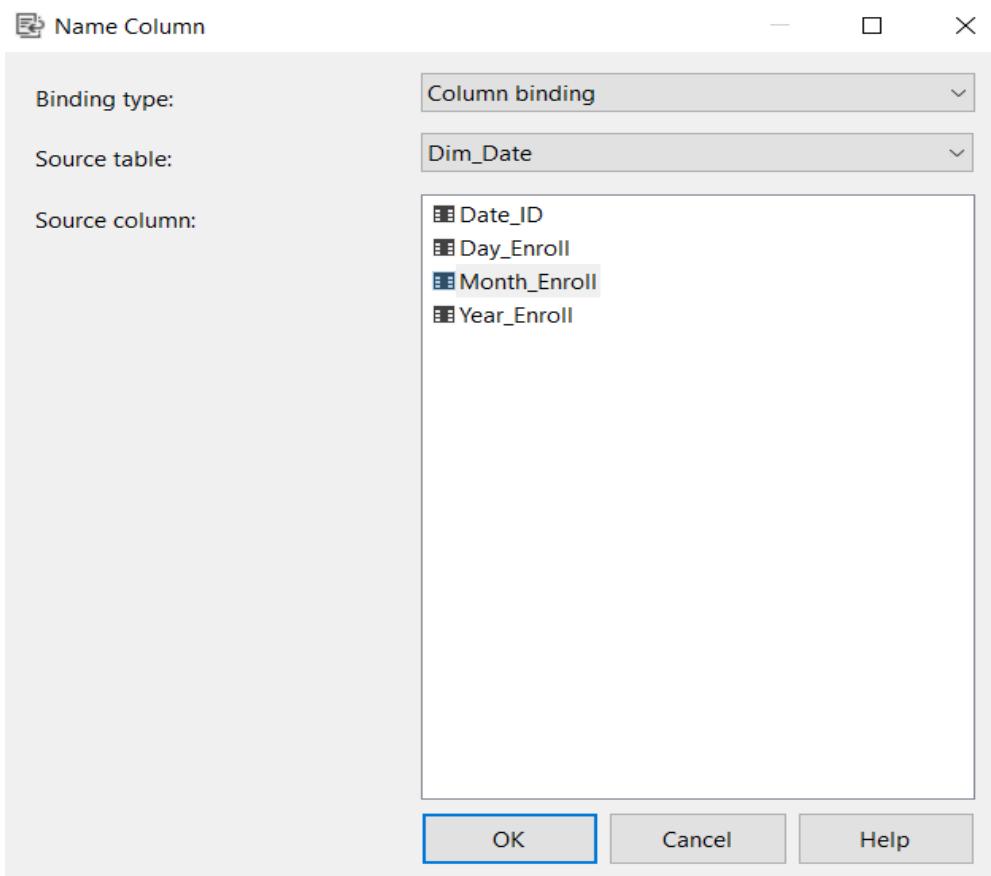
Bước 5: Chuột phải vào thuộc tính Month Enroll, chọn Properties , ở mục NameColumn chọn vào 3 dấu chấm ở cuối hàng

The screenshot shows the 'Properties' window for the 'Month Enroll' dimension attribute. On the left, there's an 'Attributes' pane showing 'Dim Date' expanded with 'Date ID', 'Day Enroll', 'Month Enroll' (selected), and 'Year Enroll'. The main area is a properties grid:

Properties	
Month Enroll DimensionAttribute	
HasLineage	False
ID	Month Enroll
InstanceSelection	None
IsAggregatable	True
KeyColumns	Dim_Date.Month_Enroll (Integer)
MemberNamesUnique	False
MembersWithData	NonLeafDataVisible
MembersWithDataCaption	
Name	Month Enroll
NameColumn	Dim_Date.Month_Enroll (WChar)
NamingTemplate	
OrderBy	Key
OrderByAttribute	
ProcessingState	Unprocessed
RootMemberIf	ParentIsBlankSelfOrMissing
TokenizationBehavior	TokenizationNone
Type	Regular
UserOperatorColumn	(none)

The 'NameColumn' row is highlighted with a red box. The 'Name' and 'OrderBy' columns are also highlighted with red boxes.

Bước 6: Chọn Day_Enroll rồi nhấn OK, tiếp tục ở mục OrderBy chọn Key

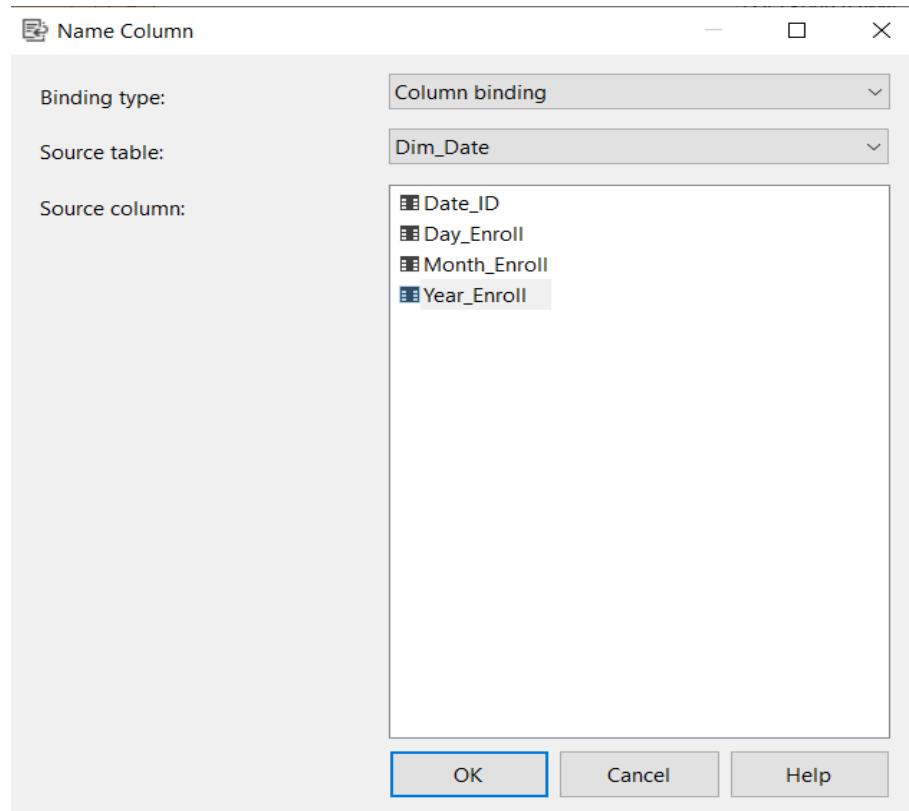


Bước 7: Chuột phải vào thuộc tính Year Enroll, chọn Properties , ở mục NameColumn
chọn vào 3 dấu chấm ở cuối hàng

The Properties window shows the following settings for the 'Year Enroll' attribute:

Property	Value
ExtendedType	
FormatString	
GroupingBehavior	EncourageGrouping
HasLineage	False
ID	Year Enroll
InstanceSelection	None
IsAggregatable	True
KeyColumns	Dim_Date.Year_Enroll (Integer)
MemberNamesUnique	False
MembersWithData	NonLeafDataVisible
MembersWithDataCaption	
Name	Year Enroll
NameColumn	Dim_Date.Year_Enroll (WChar)
NamingTemplate	
OrderBy	Key
OrderByAttribute	
ProcessingState	Unprocessed
RootMemberIf	ParentIsBlankSelfOrMissing

Bước 8: Chọn Day_Enroll rồi nhấn OK, tiếp tục ở mục OrderBy chọn Key



Bước 9: Khi hoàn thành, deploy lại khối dữ liệu , sang mục Brower để kiểm tra tính phân cấp

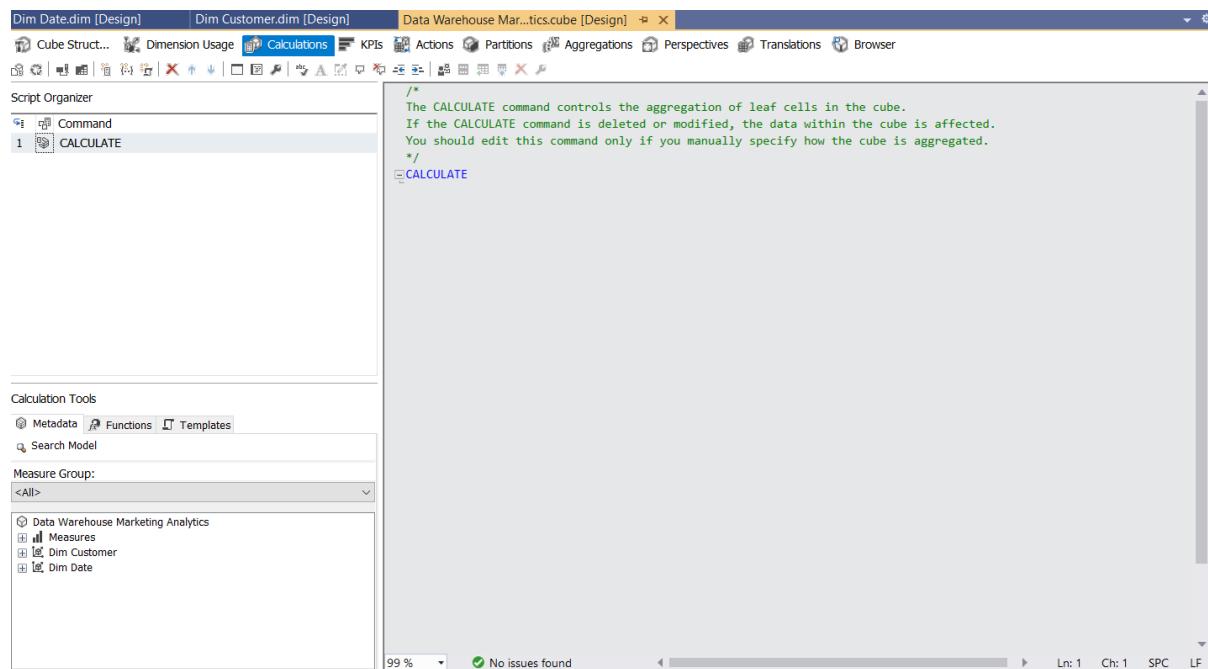
The screenshot shows the Analysis Services Browser interface. The top navigation bar includes tabs for 'Dim Date.dim [Design]', 'Dim Customer.dim [Design]', and 'Data Warehouse Mar...tics.cube [Design]'. Below the tabs, there are buttons for 'Dimension Structure', 'Attribute Relationships', 'Translations', and 'Browser' (which is highlighted). A search bar and language selection dropdown ('Language: Default') are also present. The main area displays a date hierarchy. The 'Current level:' dropdown is set to '(All)'. The hierarchy tree shows levels for years (2012, 2013, 2014), months (7, 8, 9, 10, 11, 12), and days (30, 31).

- Current level: (All)
- [-] All
 - [-] 2012
 - [-] 7
 - [-] 30
 - [-] 31
 - [-] 8
 - [-] 9
 - [-] 10
 - [-] 11
 - [-] 12
 - [-] 2013
 - [-] 2014

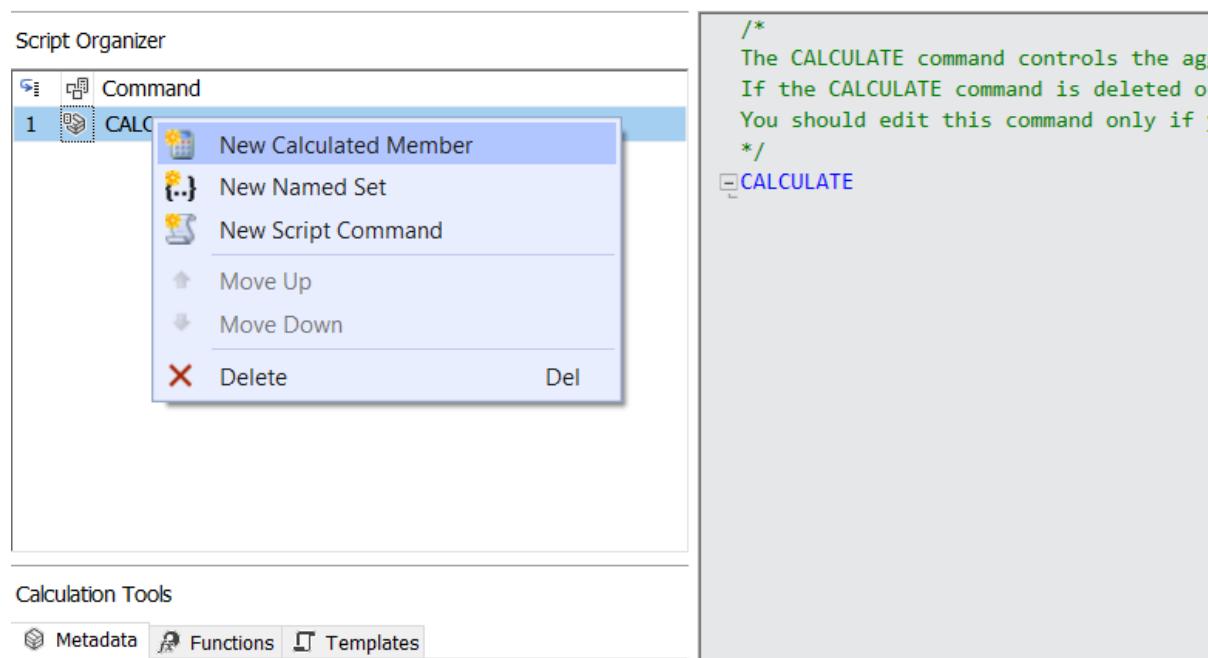
3.1.3 Tạo Calculations và Named set

1. Tạo Calculations

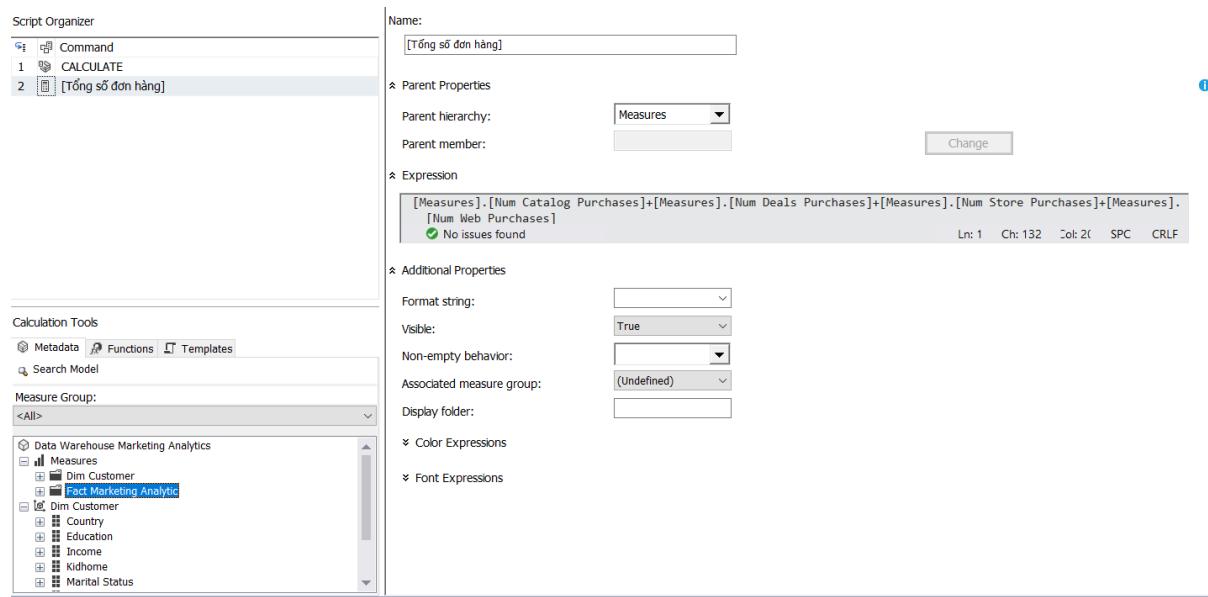
Bước 1: Nhấn vào mục Calculations trên thanh công cụ của cube



Bước 2: Chuột phải vào bảng Script Organizer chọn New Calculated Member

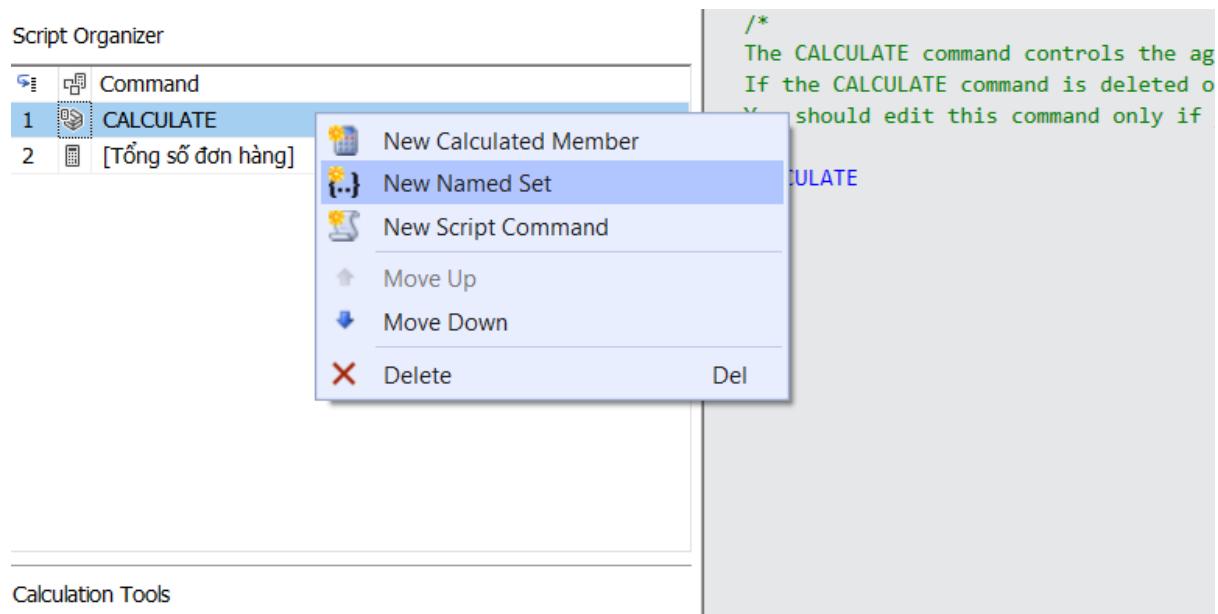


Bước 3: Thay đổi tên của calculate thành Tổng số đơn hàng và thêm phương thức tính toán vào phần Expression. Để lưu, thực hiện deploy trên cube

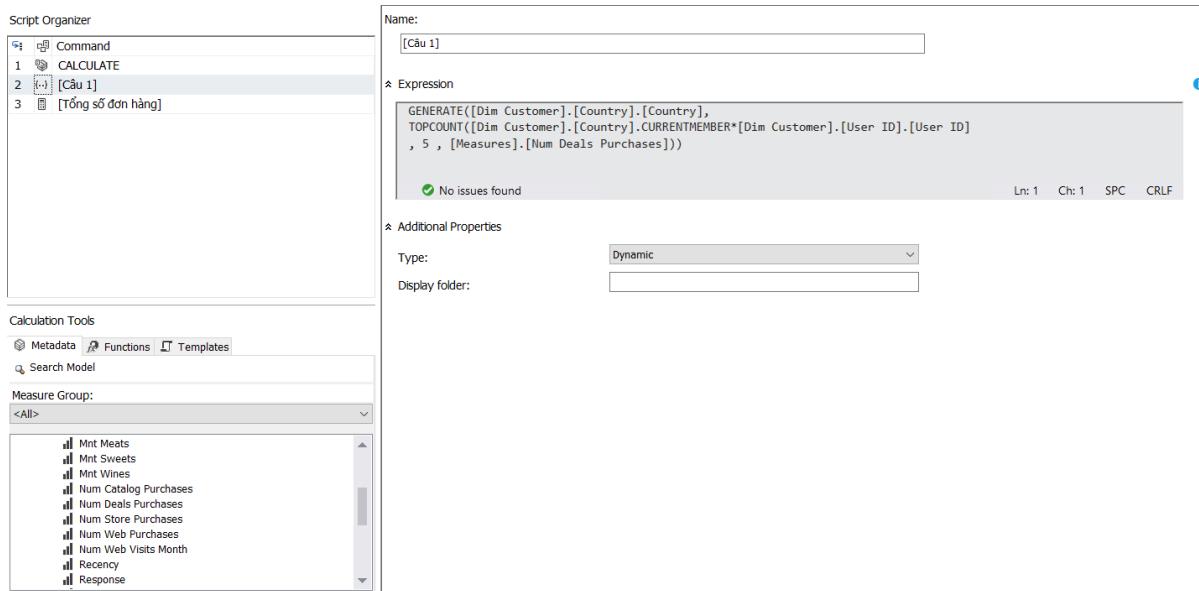


2. Tạo Named set

Bước 1: Chuột phải vào bảng Script Organizer chọn New Named Set



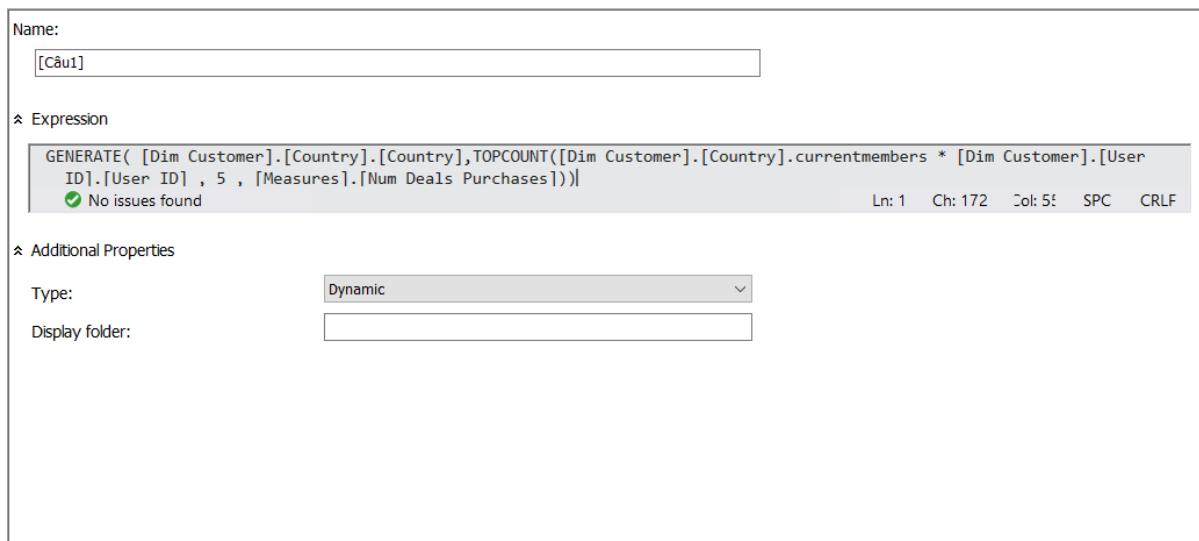
Bước 2: Thay đổi tên của Named set thành Câu 1 và thêm các hàm cần dùng vào phần Expression. Để lưu, thực hiện deploy trên cube



3.2 Truy vấn dữ liệu

Câu 1 - Cho biết top 5 khách hàng có số lượng đơn hàng kèm khuyến mãi nhiều nhất của từng quốc gia

Truy vấn bằng Visual Studio 2019



Screenshot of the Analysis Services Dimension Usage tool showing a custom dimension named "Dim Customer". The dimension uses a "Country" hierarchy and a custom operator. The filter expression is set to "GENERATE([Dim Customer].[Country],TOPCOUNT([Dim Customer].[Country].currentMember, 10))".

The interface includes tabs for Cube Structure, Dimension Usage, Calculations, KPIs, Actions, Partitions, Aggregations, Perspectives, Translations, and Browser. There are also buttons for Edit as Text, Import, MDX, and various navigation icons.

The main pane displays a table titled "Data Warehouse Marketing Analytics" with columns "Country", "User ID", and "Num Deals Purchases". The data shows purchases from Australia, Canada, Germany, India, and Montenegro.

Country	User ID	Num Deals Purchases
Australia	610	8
Australia	4865	7
Australia	6566	9
Australia	9381	7
Australia	9384	7
Canada	851	9
Canada	2350	7
Canada	2849	7
Canada	3478	10
Canada	3594	13
Germany	273	5
Germany	1357	7
Germany	1916	5
Germany	8897	8
Germany	9729	6
India	773	7
India	1456	7
India	2408	12
India	4478	8
Montenegro	2920	2
Montenegro	5080	1

Truy vấn bằng Pivot Excel

Screenshot of an Excel PivotTable setup. The PivotTable is titled "Trường PivotTable".

The data is organized by country and user ID. The table shows the number of deals purchased for each user in Australia, Canada, Germany, India, and Montenegro.

	A	B	C	D
3	Nhân Hàng	Num Deals Purchases		
4	Australia			
5	610	8		
6	4865	7		
7	6566	9		
8	9381	7		
9	9384	7		
10	Canada			
11	851	9		
12	2350	7		
13	2849	7		
14	3478	10		
15	3594	13		
16	Germany			
17	273	5		
18	1357	7		
19	1916	5		
20	8897	8		
21	9729	6		
22	India			
23	773	7		
24	1456	7		
25	2408	12		
26	4478	8		
27	8475	15		
28	Montenegro			
29	2920	2		
30	5080	1		

The PivotTable settings include the following fields:

- Hiển thị trường:** Tất cả
- Tìm kiếm:** (Tất cả)
- Đối tượng:**
 - Marital Status
 - Teenhome
 - User ID
 - Year Birth
 - query
 - Các tập hợp
- Bộ lọc:** (Không)
- Cột:** (Không)
- Hàng:** Country, User ID
- Giá trị:** Num Deals Purchases

Truy vấn bằng MDX

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. On the left is the Object Explorer pane, which lists various cubes and dimensions under the 'Data Warehouse Marketing Analytics' cube. In the center is the 'Messages' tab of the query editor, containing an MDX query:

```
--câu 1:  
SELECT [Measures].[Num Deals Purchases] ON 0,  
GENERATE([Dim Customer].[Country].[Country],  
TOPCOUNT([Dim Customer].[Country].CURRENTMEMBER*  
[Dim Customer].[User ID].[User ID]  
, 5 , [Measures].[Num Deals Purchases])) ON 1  
FROM [Data Warehouse Marketing Analytics]
```

To the right of the query is the 'Results' tab, displaying a table with three columns: Country, Num Catalog Purchases, and Num Deals Purchases. The data is as follows:

Country	Num Catalog Purchases	Num Deals Purchases
Australia	6566	9
Australia	610	8
Australia	4865	7
Australia	9381	7
Australia	9384	7
Canada	3594	13
Canada	3478	10
Canada	851	9
Canada	2350	7
Canada	2849	7
Germany	8897	8
Germany	1357	7
Germany	9729	6
Germany	273	5
Germany	1916	5
India	8475	15
India	2408	12
India	4478	8
India	773	7
India	1456	7
Montenegro	9323	4
Montenegro	2920	2
Montenegro	5080	1
South Africa	8148	13
South Africa	1081	11

Câu 2 - Quốc gia nào có số đơn mua hàng tại store nhiều nhất

Truy vấn bằng Visual Studio 2019

The screenshot shows the Microsoft Visual Studio 2019 interface, specifically the MDX query editor. The 'Name:' field is set to '[Câu2]'. The 'Expression' pane contains the following MDX code:

```
TOPSUM( [Dim Customer].[Country].[Country].members, 1, [Measures].[Num Store Purchases] )
```

The status bar at the bottom right indicates 'Ln: 1 Ch: 89 SPC CRLF'. Below the expression pane are sections for 'Additional Properties' where 'Type:' is set to 'Dynamic' and 'Display folder:' is set to 'query'.

Truy vấn bằng Pivot Excel

Truy vấn bằng MDX

--Câu 2:

```

SELECT [Measures].[Num Store Purchases] ON 0,
TOPSUM([Dim Customer].[Country].[Country],1,
[Measures].[Num Store Purchases]) ON 1
FROM [Data Warehouse Marketing Analytics]

```

Câu 3 - Cho biết thu nhập hàng năm của khách hàng đã chấp nhận 4 chiến dịch marketing

Truy vấn bằng Visual Studio 2019

The screenshot shows the MDX Editor in Visual Studio 2019. The query is:

```
FILTER( [Dim Customer].[User ID].[User ID], [Measures].[Total Campaigns Accepted] = 4 )
```

Additional Properties:

- Type: Dynamic
- Display folder: query

The results pane displays a table with columns: User ID, Income, and Total Campaigns Accepted. The data is as follows:

User ID	Income	Total Campaigns Accepted
477	83512	4
1509	84460	4
3138	91249	4
3174	87771	4
3520	91172	4
3725	84865	4
5067	80134	4
5536	87771	4
5758	65169	4
8545	85683	4
10089	102692	4

Truy vấn bằng Pivot Excel

The screenshot shows a PivotTable in Excel with the following data:

Nhận Hàng	Income	Total Campaigns Accepted
477	\$83,512.00	4
1509	\$84,460.00	4
3138	\$91,249.00	4
3174	\$87,771.00	4
3520	\$91,172.00	4
3725	\$84,865.00	4
5067	\$80,134.00	4
5536	\$87,771.00	4
5758	\$65,169.00	4
8545	\$85,683.00	4
10089	\$102,692.00	4

The PivotTable Fields pane on the right shows the following fields:

- Hiển thị trường: (Tất cả)
- Tìm kiếm
- Num Web Purchases
- Num Web Visits Month
- Recency
- Response
- Total Campaigns Accepted** (checked)
- Total Spent

Truy vấn bằng MDX

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface for writing and executing MDX queries against a Data Warehouse Marketing Analytics cube.

Cube: Data Warehouse Marketing Analytics

Measure Group: <All>

Results:

```
--Câu 3:  
SELECT {[Measures].[Income],  
[Measures].[Total Campaigns Accepted]} ON 0,  
FILTER([Dim Customer].[User ID].[User ID],  
[Measures].[Total Campaigns Accepted] = 4) ON 1  
FROM [Data Warehouse Marketing Analytics];
```

	Income	Total Campaigns Accepted
477	\$83,512.00	4
1509	\$84,460.00	4
3138	\$91,249.00	4
3174	\$87,771.00	4
3520	\$91,172.00	4
3725	\$84,865.00	4
5067	\$80,134.00	4
5536	\$87,771.00	4
5758	\$65,169.00	4
8545	\$85,683.00	4
10089	\$102,692.00	4

Câu 4 - Cho biết nơi ở của 10 khách hàng có tổng chi cho các mặt hàng là trái cây theo chiều giảm dần lấy từ vị trí thứ 3

Truy vấn bằng Visual Studio 2019

Name: [Câu 4]

Expression:

```
GENERATE([Dim Customer].[Country].[Country],[Dim Customer].[Country].CURRENTMEMBER *  
(SUBSET(ORDER([Dim Customer].[User ID].CHILDREN,[Measures].[Mnt Fruits],DESC),3,10)))|  
No issues found
```

Additional Properties:

Type: Dynamic

Display folder: query

Cube Structure... Dimension Usage Calculations KPIs Actions Partitions Aggregations Perspectives Translations Browser

Language: Default

MDX

Dimension: Dim Customer Hierarchy: Country Operator: Custom Filter Expression: GENERATE([Dim Customer].[Country].[Country],[Dim Customer].[Country],CURRENTMEMBER)

Country	User ID	Mnt Fruits
Canada	10991	194
India	202	194
India	6422	189
India	8091	184
India	9589	194
South Africa	9242	185
Spain	1137	190
Spain	3283	193
Spain	3698	185
Spain	10676	193

Truy vấn bằng Pivot Excel

Nhân Hàng Mnt Fruits

Canada

10991	\$194.00
-------	----------

India

202	\$194.00
6422	\$189.00
8091	\$184.00
9589	\$194.00

South Africa

9242	\$185.00
------	----------

Spain

1137	\$190.00
3283	\$193.00
3698	\$185.00
10676	\$193.00

Trường PivotTable

Hiển thị trường: (Tất cả)

Tìm kiếm

- query
 - Câu 1
 - Câu 11
 - Cau 15
 - Câu 4
 - Câu 6
- Dim Customer
 - Country

Kéo trường giữa các vùng bên dưới:

Bộ lọc	Cột
Hàng	Giá trị
Câu 4	Mnt Fruits

Truy vấn bằng MDX

Cube: Data Warehouse Marketing Analyst

```
--Câu 4:
SELECT [Measures].[Mnt Fruits] ON 0,
GENERATE([Dim Customer].[Country].[Country],
[Dim Customer].[Country].CURRENTMEMBER *
(SUBSET(
    ORDER([Dim Customer].[User ID].CHILDREN,
    [Measures].[Mnt Fruits],DESC),3,10))) ON 1
FROM [Data Warehouse Marketing Analytics];
```

	Messages	Results
		Mnt Fruits
Canada	10991	\$194.00
India	202	\$194.00
India	9589	\$194.00
India	6422	\$189.00
India	8091	\$184.00
South Africa	9242	\$185.00
Spain	3283	\$193.00
Spain	10676	\$193.00
Spain	1137	\$190.00
Spain	3698	\$185.00

Câu 5 - In ra những khách hàng chi cho mặt hàng kẹo nhiều hơn 100\$ nhưng số trẻ em trong nhà ít hon 2

Truy vấn bằng Visual Studio 2019

Name: [Câu5]

Expression:

```
EXCEPT( FILTER( [Dim Customer].[User ID].children, [Measures].[Mnt Sweets] > 100 ), FILTER( [Dim Customer].[User ID].children,
[Measures].[Kidhome] = 2 ) )
```

No issues found

Ln: 3 Ch: 1 SPC CRLF

Additional Properties:

Type: Dynamic

Display folder:

Cube Structure: MDX

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim Customer	User ID	In	Câu5	

User ID	Kidhome	Mnt Sweets
123	0	151
175	1	107
202	0	125
291	0	149
313	0	163
367	1	103
486	0	137
500	0	194
528	0	165
830	0	176
843	1	147
851	1	109
961	0	195
1020	0	133
1077	0	172
1137	0	156
1215	0	102
1340	0	103
1399	0	165
1411	0	153
1440	0	188
1513	1	150

Truy vấn bằng Pivot Excel

Trường PivotTable

Hiển thị trường: (Tất cả)

Tìm kiếm

- Mnt Golds
- Mnt Meats
- Mnt Sweets**
- Mnt Wines
- Num Catalog Purchases
- Num Deals Purchases
- Num Store Purchases
- Num Web Purchases

Kéo trường giữa các vùng bên dưới:

Bộ lọc	Cột
	Σ Giá trị
Hàng	Giá trị
Câu5	Kidhome
	Mnt Sweets

Truy vấn bằng MDX

--Câu 5:

```

SELECT {[Measures].[Kidhome],
[Measures].[Mnt Sweets]} ON 0,
EXCEPT (
    {FILTER([Dim Customer].[User ID].CHILDREN,
    [Measures].[Mnt Sweets]>100)},
    {FILTER([Dim Customer].[User ID].CHILDREN,
    [Measures].[Kidhome] = 2)}) ON 1
FROM [Data Warehouse Marketing Analytics];

```

146 %

Messages Results

	Kidhome	Mnt Sweets
123	0	\$151.00
175	1	\$107.00
202	0	\$125.00
291	0	\$149.00
313	0	\$163.00
367	1	\$103.00
486	0	\$137.00
500	0	\$194.00
528	0	\$165.00
830	0	\$176.00
843	1	\$147.00
851	1	\$109.00
961	0	\$195.00
1020	0	\$133.00
1077	0	\$172.00
1137	0	\$156.00
1215	0	\$102.00
1340	0	\$103.00
1399	0	\$165.00
1411	0	\$153.00
1440	0	\$188.00
1513	1	\$150.00
1524	0	\$160.00
1553	0	\$105.00
1577	0	\$175.00
1627	0	\$122.00
1763	0	\$148.00

Câu 6 - Mỗi tháng trong năm 2014, liệt kê 5 khách hàng có tổng số đơn mua hàng nhiều nhất

Truy vấn bằng Visual Studio 2019

Name: [Câu 6]

Expression

```
GENERATE(
    DESCENDANTS([Dim Date].[Time].[Year Enroll].&[2014].CHILDREN,[Dim Date].[TIME].[Month Enroll]),
    TOPCOUNT([Dim Date].[Month Enroll].CURRENTMEMBER*[Dim Customer].[User ID].CHILDREN,5,[Measures].[Tổng số đơn hàng]))
```

No issues found

Ln: 6 Ch: 1 SPC CRLF

Additional Properties

Type: Dynamic

Display folder: query

Name: [Tổng số đơn hàng]

Parent Properties

Parent hierarchy: Measures

Parent member:

Expression

```
[Measures].[Num Catalog Purchases] + [Measures].[Num Store Purchases] + [Measures].[Num Web Purchases]+[Measures].[Num Deals Purchases]
```

No issues found

Ln: 1 Ch: 136 Col: 11 SPC CRLF

MDX Editor

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim Customer	User ID	Custom	GENERATE(DESCENDANTS([Dim Date].[Time].[Year Enroll].&[2014].CHILDREN,[Dim Date].[TIME].[Month Enroll]),TOPCOUNT([Dim Date].[Month Enroll].CURRENTMEMBER*[Dim Customer].[User ID].CHILDREN,5,[Measures].[Tổng số đơn hàng]))	
Month Enroll	User ID			

Calculated Members

Month Enroll	User ID	Tổng số đơn hàng
1	4073	28
1	5136	28
1	5758	28
1	8659	27
1	9559	27
2	535	27
2	3595	28
2	4608	27
2	7646	24
2	10133	25
3	1958	25
3	3924	28
3	4676	25
3	8475	37
3	9369	32
4	1277	28
4	1340	27
4	1446	28
4	4686	24
4	7875	26
5	716	27
5	4619	27

Truy vấn bằng Pivot Excel

A screenshot of a PivotTable in Microsoft Excel. The table has three columns: Category (Nhãn Hàng), Sub-category (Tổng số đơn hàng), and Value (Giá trị). The data is grouped into five main categories (1, 2, 3, 4, 5) and further subdivided. The values range from 24 to 28.

Nhãn Hàng	Tổng số đơn hàng	
1	4073	28
	5136	28
	5758	28
	8659	27
	9559	27
2	535	27
	3595	28
	4608	27
	7646	24
	10133	25
3	1958	25
	3924	28
	4676	25
	8475	37
	9369	32
4	1277	28
	1340	27
	1446	28
	4686	24
	7875	26
5	716	27
	4619	27

Trường PivotTable

Hiển thị trường: (Tất cả)

Tìm kiếm

Các tập nộp:

- query
 - Câu 1
 - Câu 11
 - Cau 15
 - Câu 4
 - Câu 6

Dim Customer

Kéo trường giữa các vùng bên dưới:

Bộ lọc

Cột

Hàng

Giá trị

Truy vấn bằng MDX

A screenshot of the SSMS MDX Editor. On the left, the cube browser shows the Data Warehouse Marketing Analytics cube with various measure groups and dimensions selected. The central pane displays the MDX query for generating a report. The right pane shows the results of the query, which is identical to the PivotTable data above.

Cube: Data Warehouse Marketing Analytics

Measure Group: <All>

-- Câu 6:

```

WITH MEMBER [Measures].[Tổng số đơn hàng] AS
[Measures].[Num Catalog Purchases]+[Measures].[Num Deals Purchases]
+[Measures].[Num Store Purchases]+[Measures].[Num Web Purchases]
SELECT {[Measures].[Tổng số đơn hàng]} ON 0,
GENERATE(
{DESCENDANTS([Dim Date].[Time].[Year Enroll].[2014].CHILDREN,
[Dim Date].[TIME].[Month Enroll])},
{TOPCOUNT([Dim Date].[Month Enroll].CURRENTMEMBER*
[Dim Customer].[User ID].CHILDREN,
5,[Measures].[Tổng số đơn hàng]))} ON 1
FROM [Data Warehouse Marketing Analytics];
  
```

Messages Results

	Tổng số đơn hàng
1	4073
1	5136
1	5758
1	8659
1	9559
2	535
2	4608
2	10133
2	7646
3	8475
3	9369
3	3924
3	1958
3	4676
4	1277
4	1446
4	1340
4	7875
4	4686
5	4619
5	716

Câu 7 - Thống kê những khách hàng vừa mua hàng ở web hơn 10 lần và mua hàng ở cửa hàng hơn 10 lần

Truy vấn bằng Visual Studio 2019

Name: [Câu7]

Expression:

```
[INTERSECT( FILTER( [Dim Customer].[User ID].[User ID], [Measures].[Num Web Purchases] >10 ), FILTER( [Dim Customer].[User ID].[User ID], [Measures].[Num Store Purchases] >10 ) )]
```

No issues found

Additional Properties:

Type: Dynamic

Display folder: query

Ln: 1 Ch: 1 SPC CRLF

MDX Editor:

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim Customer	User ID	In	Câu7	

Data Warehouse Marketing Analytic

Fact Marketing Analytic

User ID	Num Web Purchases	Num Store Purchases
454	11	11
1204	11	11
1911	11	12
3968	11	12
4127	11	11
4910	11	13
5153	11	11
6658	11	12
7030	11	12
7966	11	11
8920	11	12

Truy vấn bằng Pivot Excel

Nhân Hàng	Num Web Purchases	Num Store Purchases
454	11	11
1204	11	11
1911	11	12
3968	11	12
4127	11	11
4910	11	13
5153	11	11
6658	11	12
7030	11	12
7966	11	11
8920	11	12

Trường PivotTable

Hiển thị trường: (Tất cả)

Tìm kiếm

Num Store Purchases

Num Web Purchases

Num Web Visits Month

Recency

Response

Total Campaigns Accepted

Total Spent

Kéo trường giữa các vùng bên dưới:

Truy vấn bằng MDX

Cube: Data Warehouse Marketing Analyst

Measure Group: <All>

```
--Câu 7:
SELECT { [Measures].[Num Web Purchases],
          [Measures].[Num Store Purchases]} ON 0,
INTERSECT({FILTER([Dim Customer].[User ID].[User ID],
                  [Measures].[Num Web Purchases]>10)},
          {FILTER([Dim Customer].[User ID].[User ID],
                  [Measures].[Num Store Purchases]>10)}) ON 1
FROM [Data Warehouse Marketing Analytics]
```

	Num Web Purchases	Num Store Purchases
454	11	11
1204	11	11
1911	11	12
3968	11	12
4127	11	11
4910	11	13
5153	11	11
6658	11	12
7030	11	12
7966	11	11
8920	11	12

Câu 8 - Truy vấn drill down các tháng của năm 2012 có lượng tiêu thụ các mặt hàng về rượu nhiều hơn 2000\$

Truy vấn bằng Visual Studio 2019

Name: [Câu8]

Expression:

```
FILTER(DRILLDOWNMEMBER( [Dim Date].[Time].[Day Enroll],[Dim Date].[Time].[Month Enroll] ), [Measures].[Mnt Wines] >= 2000)
```

No issues found

Additional Properties:

Type: Dynamic

Display folder:

Dimension: Dim Date

Hierarchy: Time

Operator: In

Filter Expression: Câu 8

Parameters: {2012}

Month Enroll	Day Enroll	Mnt Wines
8	1	2051
8	3	2236
8	8	2773
8	12	4090
8	17	2431
8	20	2325
8	22	2083
8	29	2415
8	31	4439
9	7	2051
9	8	2254
9	12	4147
9	21	2675
9	22	2615
9	28	2862
9	29	2091
10	1	4359
10	5	2164
10	12	2201
10	14	3822
10	16	2863
10	29	3254

Truy vấn bằng Pivot Excel

The screenshot shows a PivotTable in Excel with the following structure:

	Year Enroll	Mnt Wines
8	1	\$2,051.00
	3	\$2,236.00
	8	\$2,773.00
	12	\$4,090.00
	17	\$2,431.00
	20	\$2,325.00
	22	\$2,083.00
	29	\$2,415.00
	31	\$4,439.00
9	7	\$2,051.00
	8	\$2,254.00
	12	\$4,147.00
	21	\$2,675.00
	22	\$2,615.00
	28	\$2,862.00
	29	\$2,091.00
10	1	\$4,359.00
	5	\$2,164.00
	12	\$2,201.00
	14	\$3,822.00
	16	\$2,863.00
	29	\$3,254.00

To the right is the "Trường PivotTable" (PivotTable Fields) pane:

- Câu 8** (Row Labels): Contains Year Enroll (checked), Month Enroll (checked), and Câu 8 (checked).
- Bộ lọc** (Filters): Contains Year Enroll.
- Cột** (Columns): Empty.
- Hàng** (Rows): Contains Month Enroll and Câu 8.
- Giá trị** (Values): Contains Mnt Wines.

Truy vấn bằng MDX

The screenshot shows the SSMS MDX query editor with the following details:

--Câu 8:

```

SELECT [Measures].[Mnt Wines] ON 0,
[Dim Date].[Month Enroll].[Month Enroll]*

(FILTER (
    {DRILLDOWNMEMBER([Dim Date].[Time].[Day Enroll],
    [Dim Date].[Time].[Month Enroll].MEMBERS)},
    [Measures].[Mnt Wines]>= 2000)) ON 1
FROM [Data Warehouse Marketing Analytics]
WHERE [Dim Date].[Year Enroll].&[2012];

```

Results:

	Mnt Wines
8 1	\$2,051.00
8 3	\$2,236.00
8 8	\$2,773.00
8 12	\$4,090.00
8 17	\$2,431.00
8 20	\$2,325.00
8 22	\$2,083.00
8 29	\$2,415.00
8 31	\$4,439.00
9 7	\$2,051.00
9 8	\$2,254.00
9 12	\$4,147.00
9 21	\$2,675.00
9 22	\$2,615.00
9 28	\$2,862.00
9 29	\$2,091.00
10 1	\$4,359.00
10 5	\$2,164.00
10 12	\$2,201.00
10 14	\$3,822.00
10 16	\$2,863.00
10 29	\$3,254.00

The left pane shows the cube structure:

- Measure Group:** <All>
 - Total Campaigns Acce
 - Total Spent
 - Tổng số đơn hàng
 - Trung bình
- KPIs**
- Dim Customer**
- query**
- Country**
 - Members
- Education**
- Income**
- Kidhome**
- Marital Status**
- Teenhome**
- User ID**
 - Members
- Year Birth**
- Dim Date**
- query**
- Date ID**
- Day Enroll**
- Month Enroll**
- Year Enroll**
 - Members
- Time**
 - Members

Câu 9 - Cho biết top 10 khách hàng có thu nhập hàng năm lớn nhất

Truy vấn bằng Visual Studio 2019

Name: [Câu9]

Expression

```
HEAD( order( [Dim Customer].[User ID].[User ID].members,[Measures].[Income],DESC),10 )
```

No issues found

Ln: 1 Ch: 87 SPC CRLF

Additional Properties

Type: Dynamic

Display folder:

Data Warehouse Marketing Analytics [Design] Dim Date.dim [Design] Dim Customer.dim [Design]

Cube Struct... Dimension Usage Calculations KPIs Actions Partitions Aggregations Perspectives Translations Browser

Edit as Text Import... MDX

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim Customer	User ID	In	Câu9	

<Select dimension>

User ID	Income
1501	160803
1503	162397
4611	105471
4619	113734
4931	157146
5336	157733
5555	153924
8475	157243
9432	666666
11181	156924

Truy vấn bằng Pivot Excel

Nhận Hàng	Income
1501	160803
1503	162397
4611	105471
4619	113734
4931	157146
5336	157733
5555	153924
8475	157243
9432	666666
11181	156924

Trường Pivot Table

Hiển thị trường: (Tất cả)

Tìm kiếm

Σ Dim Customer
Income

Σ Fact Marketing Analytic
Complain
Fact Marketing Analytic Count
Mnt Fishs
Mnt Fruits
Mnt Golds

Kéo trường giữa các vùng bên dưới:

Bộ lọc Cột

Truy vấn bằng MDX

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. On the left, the Object Explorer displays a cube named "Data Warehouse Marketing Analyst" with various dimensions like Dim Customer, Dim Date, and Time, and measure groups like KPIs and Income. In the center, the Results pane shows the output of an MDX query:

```
--Câu 9:
SELECT [Measures].[Income] ON 0,
HEAD(
    ORDER([Dim Customer].[User ID].[User ID].MEMBERS,
        [Measures].[Income],DESC),10) ON 1
FROM [Data Warehouse Marketing Analytics];
```

The results table has two columns: "Income" and "Results". The data is as follows:

Income	Results
9432	\$666,666.00
1503	\$162,397.00
1501	\$160,803.00
5336	\$157,733.00
8475	\$157,243.00
4931	\$157,146.00
11181	\$156,924.00
5555	\$153,924.00
4619	\$113,734.00
4611	\$105,471.00

Câu 10 - Cho biết tên 2 quốc gia nào có số lượng mua hàng ở store ít nhất

Truy vấn bằng Visual Studio 2019

The screenshot shows the Microsoft Visual Studio 2019 interface. The top navigation bar includes "File", "Edit", "View", "Project", "Server Explorer", "Toolbox", "Task List", "Output", "Properties", "Toolbox", "MDX", "Help", and "Language: Default".

The main area shows an MDX query in the "Query" editor:

```
BOTTOMCOUNT([Dim Customer].[Country].[Country].members,2, [Measures].[Num Store Purchases] )
```

A status bar at the bottom right indicates "Ln: 1 Ch: 93 SPC CRLF".

Below the editor, the "Properties" window shows:

- Name: [Câu10]
- Type: Dynamic
- Display folder: query

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. The top navigation bar includes "File", "Edit", "View", "Project", "Server Explorer", "Toolbox", "Task List", "Output", "Properties", "Toolbox", "MDX", "Help", and "Language: Default".

The main area shows an MDX query in the "Query" editor:

```
Dimension Hierarchy Operator Filter Expression
Dim Customer Country In Câu 10
```

The results table shows the number of store purchases by country:

Country	Num Store Purchases
Montenegro	19
United States	650

The left sidebar shows the cube structure with dimensions like Dim Customer, Dim Date, and Time, and measure groups like KPIs and Income. A calculated member named "Câu 10" is visible under the Income group.

Truy vấn bằng Pivot Excel

The screenshot shows a PivotTable in Excel with the following data:

Nhãn Hàng	Num Store Purchases
Montenegro	19
United States	650

To the right is the "Trường PivotTable" (PivotTable Fields) pane, which includes:

- Hiển thị trường: (Tất cả)
- Tìm kiếm:
- Checkboxes for fields: Kidhome, Marital Status, Teenhome, User ID, Year Birth.
- A folder named "query" containing:
 - Câu 10
 - Câu 12

Truy vấn bằng MDX

The screenshot shows the SSMS MDX Query Editor with the following details:

Cube: Data Warehouse Marketing Analytics

Measure Group: <All>

Query:

```

[Measures].[Income],DESC),10) ON 1
FROM [Data Warehouse Marketing Analytics];

--Câu 10:
SELECT [Measures].[Num Store Purchases] ON 0,
BOTTOMCOUNT([Dim Customer].[Country].[Country].MEMBERS,
2,[Measures].[Num Store Purchases]) ON 1
FROM [Data Warehouse Marketing Analytics];
  
```

Results:

	Num Store Purchases
Montenegro	19
United States	650

Câu 11 - Tính trung bình số tiền đã chi cho tất cả các mặt hàng của 5% khách hàng có thu nhập hàng năm cao nhất theo từng quốc gia

Truy vấn bằng Visual Studio 2019

Name: [Câu11]

Expression

```
GENERATE([Dim Customer].[Country].[Country],
    [TOPPERCENT( [Dim Customer].[Country].currentmember * FILTER( [Dim Customer].[User ID].[User ID], [Measures].[Tổng số đơn hàng] <> 0 ),5,[Measures].[Income]))
```

No issues found

Additional Properties

Type: Dynamic

Display folder: query

Ln: 2 Ch: 5 SPC CRLF

Name: [Trung bình]

Parent Properties

Parent hierarchy: Measures

Parent member: Change

Expression

```
[Measures].[Total Spent]/[Measures].[Tổng số đơn hàng]
```

No issues found

Ln: 1 Ch: 55 SPC CRLF

Dimension Hierarchy Operator Filter Expression Parameters

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim Customer	Country	Custom	GENERATE([Dim Customer].[Country].[Country],[TOPPERCENT([Dim Customer].[Country].currentmember * FILTER([Dim Customer].[User ID].[User ID], [Measures].[Tổng số đơn hàng] <> 0),5,[Measures].[Income]))	
	Country			
	User ID			
	Income			
	Trung bình			

Country User ID Income Trung bình

Australia	2802	87000	56.5294117647059
Australia	3619	87305	72
Australia	10430	89694	133.818181818182
Australia	10909	92344	94.95
Australia	11101	89891	33.44
Canada	821	92859	98.9047619047619
Canada	1592	90765	114.933333333333
Canada	3520	91172	78
Canada	4070	94871	63.4117647058824
Canada	6303	91820	68.0454545454545
Canada	7215	101970	42.037037037037
Canada	9499	93404	115.428571428571
Canada	10133	93790	93.96
Germany	2088	84906	70.9130434782609
Germany	3560	83844	82.8421052631579
Germany	9349	92533	51.4090909090909
Germany	9369	88194	41.0967741935484
India	4278	87188	106.833333333333
India	6422	86718	79.375
India	8475	157243	73.09090909090909
India	10469	88328	95.4705882352941
Montenegro	5080	70515	78.625
South Africa	4931	157146	61.7857142857143

Truy vấn bằng Pivot Excel

Nhãn Hàng Income Trung bình

Nhãn Hàng	Income	Trung bình
Australia		
2802	87000	56.52941176
3619	87305	72
10430	89694	133.8181818
10909	92344	94.95
11101	89891	33.44
Canada		
821	92859	98.9047619
1592	90765	114.9333333
3520	91172	78
4070	94871	63.41176471
6303	91820	68.04545455
7215	101970	42.03703704
9499	93404	115.4285714
10133	93790	93.96
Germany		
2088	84906	70.91304348
3560	83844	82.84210526
9349	92533	51.40909091
9369	88194	41.09677419
India		
4278	87188	106.8333333
6422	86718	79.375
8475	157243	73.09090909
10469	88325	95.47058824
Montenegro		
5080	70515	78.625

Truy vấn bằng MDX

-- Câu 11:

```

WITH MEMBER [Measures].[Trung bình] AS
[Measures].[Total Spent]/[Measures].[Tổng số đơn hàng]
SELECT {[Measures].[Income],[Measures].[Trung bình]} ON 0,
GENERATE([Dim Customer].[Country].[Country],
TOPPERCENT([Dim Customer].[Country].CURRENTMEMBER*
FILTER([Dim Customer].[User ID].[User ID],
[Measures].[Tổng số đơn hàng] <>0)),
5,[Measures].[Income])) ON 1
FROM [Data Warehouse Marketing Analytics];

```

	Income	Trung bình
Australia	\$92,344.00	94.95
Australia	\$89,891.00	32.1538461538462
Australia	\$89,694.00	105.142857142857
Australia	\$87,305.00	67.7647058823529
Australia	\$87,000.00	53.38888888888889
Canada	\$101,970.00	42.037037037037
Canada	\$94,871.00	59.8888888888889
Canada	\$93,790.00	93.96
Canada	\$93,404.00	95.0588235294118
Canada	\$92,859.00	94.4090909090909
Canada	\$91,820.00	68.0454545454545
Canada	\$91,172.00	78
Canada	\$90,765.00	114.933333333333
Germany	\$92,533.00	49.1739130434783
Germany	\$88,194.00	39.8125
Germany	\$84,906.00	67.9583333333333
Germany	\$83,844.00	78.7
India	\$157,243.00	43.4594594594595
India	\$88,325.00	90.16666666666667
India	\$87,188.00	101.210526315789
India	\$86,718.00	74.7058823529412
Montenegro	\$70,515.00	74

Câu 12 - Thống kê số tiền đã mua hàng của top 10 khách hàng có thu nhập hàng năm lớn hơn 30000\$

Truy vấn bằng Visual Studio 2019

Name: [Câu12]

Expression

```
HEAD( FILTER( ORDER( [Dim Customer].[User ID].[User ID], [Measures].[Total Spent], DESC ), [Measures].[Income]>30000 ), 10 )
```

No issues found

Ln: 1 Ch: 120 SPC CRLF

Additional Properties

Type: Dynamic

Display folder: query

User ID	Income	Total Spent
1763	87679	2524
4475	69098	2440
4580	75759	2486
5350	90638	2525
5386	94384	2302
5453	90226	2352
5735	90638	2525
6024	94384	2302
9010	83151	2346
10133	93790	2349

Truy vấn bằng Pivot Excel

Nhân Hàng	Income	Total Spent
1763	87679	2524
4475	69098	2440
4580	75759	2486
5350	90638	2525
5386	94384	2302
5453	90226	2352
5735	90638	2525
6024	94384	2302
9010	83151	2346
10133	93790	2349

Trường Pivot Table

Hiển thị trường: (Tất cả)

Tìm kiếm

- query
 - Câu10
 - Câu12
 - Cau14
 - Câu2
 - Câu3
 - Câu5
 - Câu8

Kéo trường giữa các vùng bên dưới:

Bộ lọc

Cột

Σ Giá trị

Truy vấn bằng MDX

Cube: Data Warehouse Marketing Analytics

Measure Group: <All>

```
-- Câu 12:
SELECT {[Measures].[Income],[Measures].[Total Spent]} ON 0,
HEAD(
    {FILTER({ORDER([Dim Customer].[User ID].[User ID],
[Measures].[Total Spent],DESC)},
[Measures].[Income] > 30000)},10) ON 1
FROM [Data Warehouse Marketing Analytics];
```

Messages Results

	Income	Total Spent
5350	\$90,638.00	2525
5735	\$90,638.00	2525
1763	\$87,679.00	2524
4580	\$75,759.00	2486
4475	\$69,098.00	2440
5453	\$90,226.00	2352
10133	\$93,790.00	2349
9010	\$83,151.00	2346
5386	\$94,384.00	2302
6024	\$94,384.00	2302

Câu 13 - Cho biết 3 tháng có số lượng đơn hàng nhiều nhất tại South Africa

Truy vấn bằng Visual Studio 2019

Name: [Cau13]

Expression:

```
HEAD(ORDER( [Dim Date].[Month Enroll].[Month Enroll],[Measures].[Tổng số đơn hàng],DESC),3 )
```

No issues found

Additional Properties:

Type: Dynamic
Display folder: query

Cube Structure | Dimension Usage | Calculations | KPIs | Actions | Partitions | Aggregations | Perspectives | Translations | Browser

Language: Default

Dimension: Dim Date | Hierarchy: Month Enroll | Operator: In | Filter Expression: Câu 13 | Parameters: (South Africa)

Month Enroll | Tổng số đơn hàng

1	510
8	552
10	653

Calculated Members

Truy vấn bằng Pivot Excel

The screenshot shows a Microsoft Excel spreadsheet with a PivotTable. The PivotTable has 'Country' in the Row Labels and 'Tổng số đơn hàng' (Total Sales) in the Value field. The data shows:

Nhãn Hàng	Tổng số đơn hàng
1	510
8	552
10	653

To the right of the spreadsheet is the 'Trường PivotTable' (PivotTable Fields) pane. It shows the fields used in the PivotTable:

- Hiển thị trường: (Tất cả)
- Tìm kiếm
- Trung bình
- Các tập hợp
- Dim Customer
 - Country
 - Education
 - Income
 - Kidhome

Truy vấn bằng MDX

The screenshot shows the SQL Server Management Studio interface. On the left is the Object Explorer showing the cube 'Data Warehouse Marketing Analytics'. The 'Dim Customer' dimension is expanded, showing the 'Country' hierarchy with members: All, Australia, Canada, Germany, India, Montenegro, South Africa, Spain, and United States.

In the center, the 'Messages' tab displays an MDX query:

```
-- Câu 13:  
SELECT [Measures].[Tổng số đơn hàng] ON 0,  
HEAD(ORDER([Dim Date].[Month Enroll].[Month Enroll],  
[Measures].[Tổng số đơn hàng],DESC),3) ON 1  
FROM [Data Warehouse Marketing Analytics]  
WHERE[Dim Customer].[Country].&[South Africa];
```

Below the query is the 'Results' tab showing the results of the query:

	Tổng số đơn hàng
10	653
8	552
1	510

Câu 14 - In ra ngày có ít lượng truy cập vào web nhất trong tháng 9 năm 2012

Truy vấn bằng Visual Studio 2019

Name: [Cau14]

Expression

```
BOTTOMSUM( DESCENDANTS( [Dim Date].[Time].[Year Enroll].&[2012].&[10], [Dim Date].[Time].[Day Enroll]),1,[Measures].[Num Web Visits Month] )
```

No issues found

Ln: 1 Ch: 141 Col: 17 SPC CRLF

Additional Properties

Type: Dynamic

Display folder: query

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim Date	Time	In	Cau14	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				

Day Enrol	Num Web Visits Month
19	2

Truy vấn bằng Pivot Excel

Truy vấn bằng MDX

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. On the left, the 'Cube' browser displays the 'Data Warehouse Marketing Analytics' cube with various measure groups and dimensions. The main pane contains an MDX query:

```
--Câu 14:  
SELECT [Measures].[Num Web Visits Month] ON 0,  
BOTTOMSUM(  
{DESCENDANTS([Dim Date].[Time].[Year Enroll].&[2012].&[10],  
[Dim Date].[Time].[Day Enroll])},1,  
[Measures].[Num Web Visits Month]) ON 1  
FROM [Data Warehouse Marketing Analytics];
```

The results pane shows the output for the query:

Messages	Results
Num Web Visits Month	19 2

Câu 15 - Cho biết số lượng đơn hàng đã thanh toán qua catalog và thu nhập hàng năm của top 15 khách hàng độc thân và mới tốt nghiệp (tăng dần theo số lượng đơn hàng)

Truy vấn bằng Visual Studio 2019

The screenshot shows the Microsoft Visual Studio 2019 interface with the 'MDX Editor' tool. The 'Name' field is set to '[Cau15]'. The 'Expression' pane contains the following MDX query:

```
TAIL(  
ORDER (  
    DRILLUPLEVEL( ( [Dim Customer].[Marital Status].&[Single], [Dim Customer].[Education].&[Graduation]), [Dim Customer].  
    [Marital Status].[Marital Status])* [Dim Customer].[User ID].[User ID] )  
, [Measures].[Num Catalog Purchases], ASC) , 15 )
```

The status bar at the bottom indicates 'No issues found'.

The 'Additional Properties' section shows:

- Type: Dynamic
- Display folder: query

The screenshot shows the Analysis Services Management Studio interface. At the top, there's a ribbon with various tabs like Cube Structure, Dimension Usage, Calculations, KPIs, Actions, Partitions, Aggregations, Perspectives, Translations, and Browser. Below the ribbon, there are buttons for Edit as Text, Import..., MDX, and other common actions. A search bar labeled 'Search Model' and a dropdown for 'Language: Default' are also present.

The main area is divided into several sections:

- Measure Group:** Shows 'Data Warehouse Marketing Analytics' and a tree view of dimensions and measures. Dimensions include Dim Customer, Fact Marketing Analytic, and Complain. Measures include Income, Num Catalog Purchases, and various Mnt (Meats, Sweets, Wines) and Num (Catalog Purchases, Deals Purchases, Store Purchases, Web Purchases) metrics.
- Data Grid:** Displays a table with columns 'User ID', 'Income', and 'Num Catalog Purchases'. The data is as follows:

User ID	Income	Num Catalog Purchases
2109	96843	11
3005	83528	10
3138	91249	9
3262	68281	9
4138	76624	10
4216	91065	9
4483	72354	10
4508	75127	11
4843	75127	11
5147	90842	9
5524	58138	10
6246	73892	11
6945	84574	11
9145	72504	9
10736	72258	9
- Calculated Members:** A section where users can define calculated members.
- Parameters:** A section for defining parameters used in the MDX query.

Truy vấn bằng Pivot Excel

The screenshot shows an Excel spreadsheet with a PivotTable on the right. The PivotTable is titled 'Trường PivotTable' and has the following structure:

- Hiển thị trường:** (Tất cả)
- Tìm kiếm:** (empty)
- Cột:** Σ Giá trị
- Bộ lọc:** (empty)
- Hàng:**
 - Câu 15
 - Income
 - Num Catalog Purchases

The data in the PivotTable is as follows:

Nhân Hàng	Income	Num Catalog Purchases
Single		
Graduation		
2109	\$96,843.00	11
3005	\$83,528.00	10
3138	\$91,249.00	9
3262	\$68,281.00	9
4138	\$76,624.00	10
4216	\$91,065.00	9
4483	\$72,354.00	10
4508	\$75,127.00	11
4843	\$75,127.00	11
5147	\$90,842.00	9
5524	\$58,138.00	10
6246	\$73,892.00	11
6945	\$84,574.00	11
9145	\$72,504.00	9
10736	\$72,258.00	9

Truy vấn bằng MDX

The screenshot shows the Microsoft Analysis Services Management Studio interface. On the left, there's a tree view of the cube structure under 'Data Warehouse Marketing Analytics'. The 'Measures' node is expanded, showing various measures like Income, Kidhome, Complain, etc. In the center, an MDX query is displayed:

```
-- Câu 15:  
SELECT {[Measures].[Income],[Measures].[Num Catalog Purchases]} ON 0,  
TAIL(  
    ORDER (  
        (DRILLUPLEVEL( ([Dim Customer].[Marital Status].&[Single],  
            [Dim Customer].[Education].&[Graduation]),  
            [Dim Customer].[Marital Status].[Marital Status])*  
            [Dim Customer].[User ID].[User ID] )  
        , [Measures].[Num Catalog Purchases], ASC) , 15 ) ON 1  
FROM [Data Warehouse Marketing Analytics];
```

Below the query, a table titled 'Results' displays the data:

Income	Num Catalog Purchases
\$91,249.00	9
\$68,281.00	9
\$91,065.00	9
\$90,842.00	9
\$72,504.00	9
\$72,258.00	9
\$83,528.00	10
\$76,624.00	10
\$72,354.00	10
\$58,138.00	10
\$96,843.00	11
\$75,127.00	11
\$75,127.00	11
\$73,892.00	11
\$84,574.00	11

3.3 Report với Power BI

3.3.1 Tạo project với Power BI

Bước 1: Tải Power BI tại [link](#)

Microsoft Power BI Desktop

The page has a message: 'Important! Selecting a language below will dynamically change the complete page content to that language.' A dropdown menu shows 'English' and a red 'Download' button.

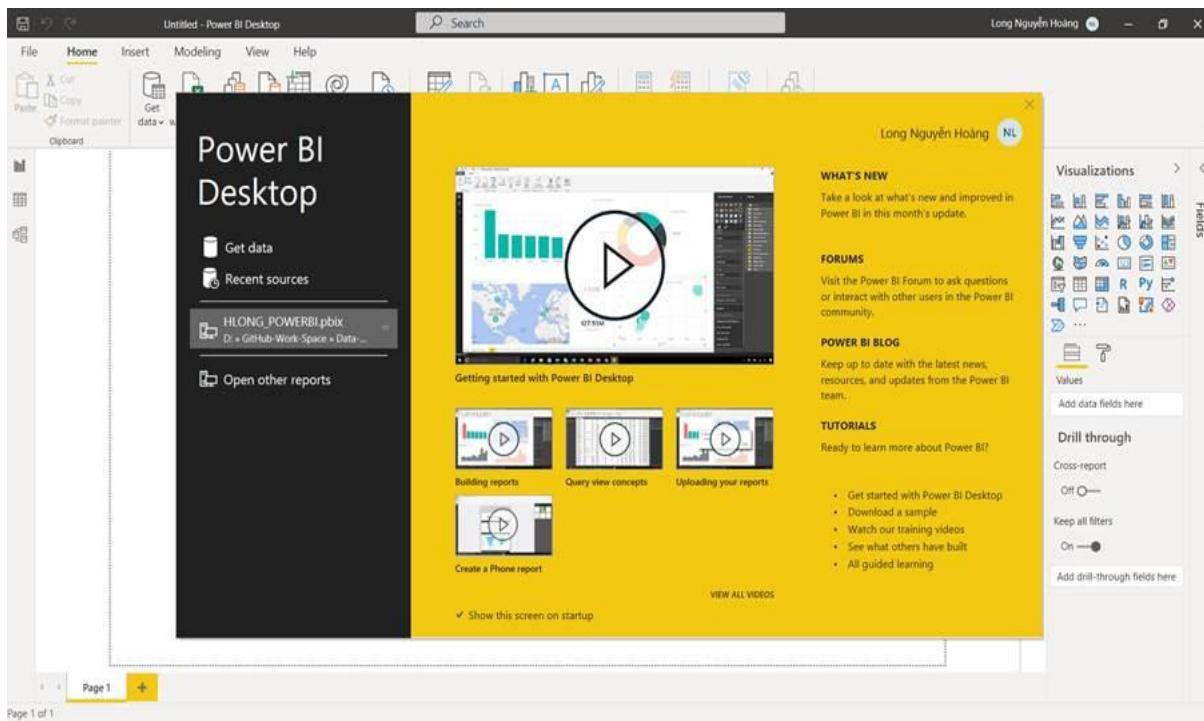
Microsoft Power BI Desktop is built for the analyst. It combines state-of-the-art interactive visualizations, with industry-leading data query and modeling built-in. Create and publish your reports to Power BI. Power BI Desktop helps you empower others with timely critical insights, anytime, anywhere.

Details

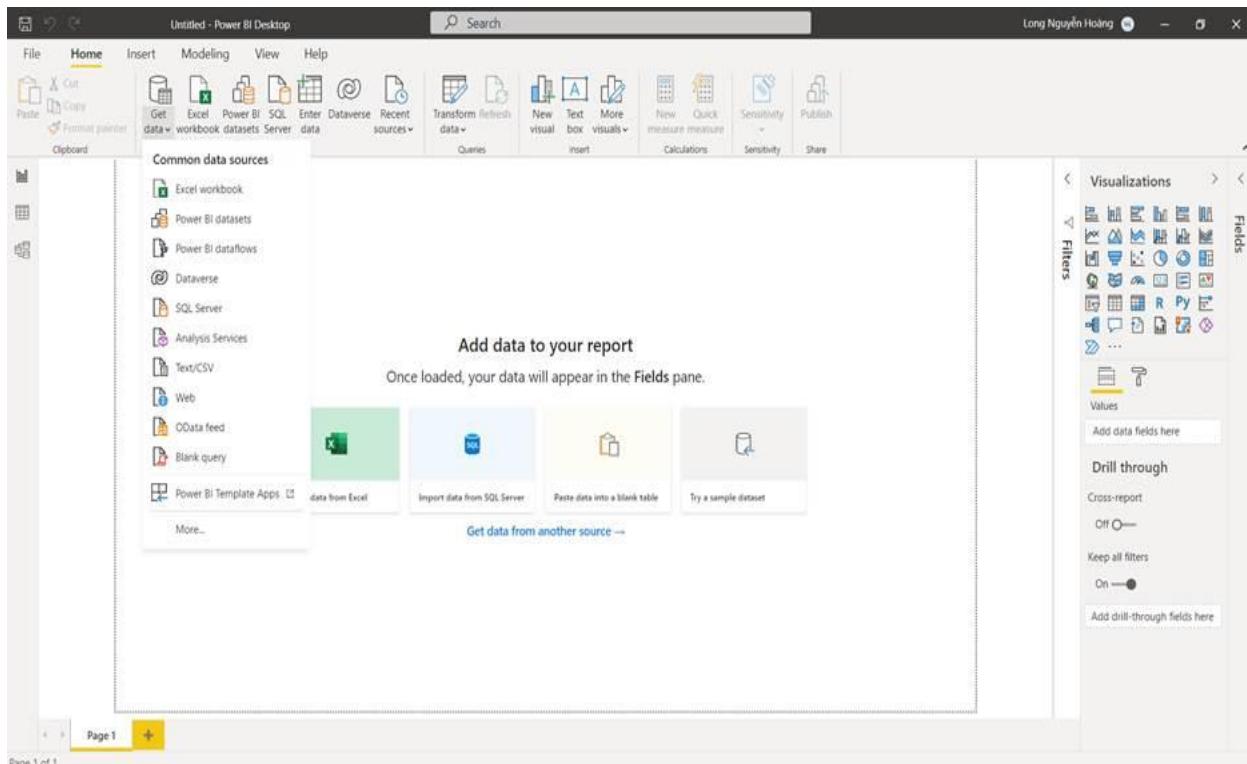
System Requirements

Install Instructions

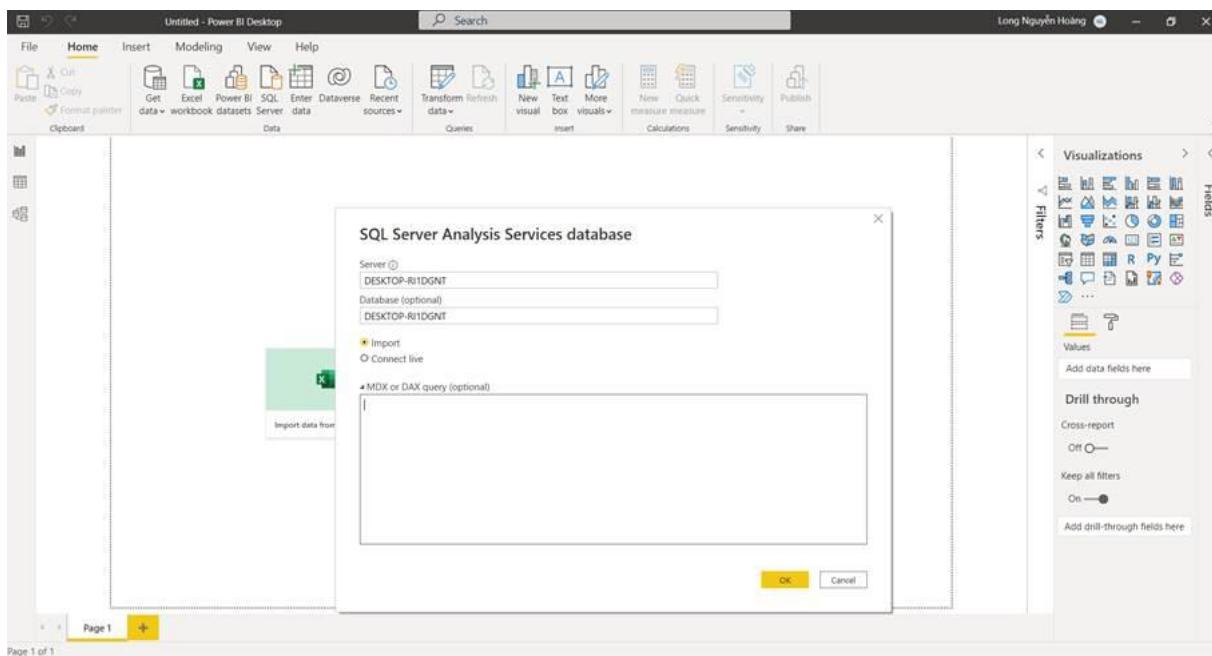
Bước 2: Giao diện chính của Power BI



Bước 3: Chọn Get Data -> Analysis Services



Bước 4: Điền tên Server trên máy -> Chọn Import -> Điền câu lệnh MDX tương ứng



3.3.2 Report 1 – Câu 1

Bước 1: Điền câu truy vấn MDX

The screenshot shows the 'SQL Server Analysis Services database' import dialog box. It has fields for 'Server' (set to 'DESKTOP-RI1DGNT') and 'Database' (set to 'DESKTOP-RI1DGNT'). The 'Import' radio button is selected. The 'MDX or DAX query (optional)' section contains the following MDX query:

```
SELECT [Measures].[Num Deals Purchases] ON 0,
GENERATE([Dim Customer].[Country].[Country],
TOPCOUNT([Dim Customer].[Country].CURRENTMEMBER*
[Dim Customer].[User ID].[User ID]
, 5 , [Measures].[Num Deals Purchases])) ON 1
FROM [Data Warehouse Marketing Analytics];
```

At the bottom right of the dialog box are 'OK' and 'Cancel' buttons.

Truy vấn MDX câu 1

Bước 2: Click vào Transform data để chỉnh sửa tên cột và kiểu dữ liệu

The screenshot shows the 'DESKTOP-RI1DGNT: SSAS' dialog box. It displays a preview of data from the 'Dim Customer' dimension and a measure from the 'Measures' cube. The data is presented in a table with three columns: 'Country' (containing values like Australia, Canada, Germany, India, Montenegro), 'User_ID' (containing values like 6566, 3594, 8996, 8475, 9323), and 'Num Deals Purchases' (containing values like 9, 13, 12, 15, 4). A note at the bottom left states: 'The data in the preview has been truncated due to size limits.' At the bottom right, there are three buttons: 'Load' (highlighted in yellow), 'Transform Data', and 'Cancel'.

Bước 3: Đổi tên column -> Click chuột phải vào column -> Change type -> Close and Apply

The screenshot shows the 'Transform Data' interface in Power BI. It displays a table with columns: 'Country', 'User_ID', and 'Num Deals Purchases'. The 'User_ID' column is currently selected, as indicated by a green border around its header. The formula bar at the top of the interface shows the command: = Table.TransformColumnTypes(#"Renamed Columns",{{"User_ID", type number}, {"Num Deals Purchases", type number}}). At the bottom right, there are two buttons: 'Close' and 'Apply'.

Bước 4: Chọn Matrix Visualization -> Chọn các thuộc tính cần thiết -> Stack các thuộc tính theo mong muốn

Country

	Num Deals Purchases
Australia	38.00
Canada	46.00
Germany	38.00
273.00	5.00
1,357.00	7.00
8,897.00	8.00
8,996.00	12.00
9,729.00	6.00
India	49.00
773.00	7.00
1,456.00	7.00
2,408.00	12.00
4,478.00	8.00
8,475.00	15.00
Montenegro	7.00
2,920.00	2.00
5,080.00	1.00
9,323.00	4.00
South Africa	52.00
1,081.00	11.00
1,165.00	9.00
2,877.00	10.00
6,097.00	9.00
8,146.00	13.00
Spain	73.00
United States	56.00
1,501.00	15.00
4,095.00	8.00
4,974.00	10.00
5,376.00	15.00
9,967.00	8.00
Total	359.00

Search

Filters on this visual

- Country is (All)
- Num Deals Purchases is (All)
- User_ID is (All)

Add data fields here

Rows

- Country
- User_ID

Columns

Add data fields here

Values

- Num Deals Purchases

Drill through

Bước 5: Chọn format -> Chính sửa report cho vừa mắt

Search

General

Style

Grid

Column headers

Row headers

Values

Subtotals

Grand total

Field formatting

Conditional formatting

Bước 6: Kết quả cuối cùng

Country	Num Deals Purchases
Australia	38.00
Canada	46.00
Germany	38.00
India	49.00
Montenegro	7.00
South Africa	52.00
Spain	73.00
10749	15.00
1513	13.00
238	15.00
6862	15.00
9931	15.00
United States	56.00
1501	15.00
4095	8.00
4974	10.00
5376	15.00
9967	8.00
Total	359.00

Matrix Visualization

3.3.3 Report 2 – Câu 7

SQL Server Analysis Services database

Server ①
DESKTOP-RI1DGNT

Database
SSAS

Import
 Connect live

MDX or DAX query (optional)

```
SELECT { [Measures].[Num Web Purchases],  
         [Measures].[Num Store Purchases]} ON 0,  
INTERSECT({FILTER([Dim Customer].[User ID].[User ID],  
                  [Measures].[Num Web Purchases]>10)},  
          {FILTER([Dim Customer].[User ID].[User ID],  
                  [Measures].[Num Store Purchases]>10)}) ON 1  
FROM [Data Warehouse Marketing Analytics];
```

OK Cancel

Truy vấn MDX câu 7

Bước 2: Thực hiện đổi tên column, đổi kiểu dữ liệu các cột tương tự như câu 1

Bước 3: Chọn Table Visualization -> Chọn các thuộc tính cần thiết -> Stack các thuộc tính theo mong muốn -> Chọn format và chỉnh sửa report cho vừa mắt

The screenshot shows the Power BI Data View. On the left is a table with three columns: User_ID, Num Store Purchases, and Num Web Purchases. The table contains data for individual users and a total row. On the right, there are several interface elements: a 'Values' section with dropdowns for User_ID, Num Store Purchases, and Num Web Purchases; a 'Drill through' section; and a sidebar with a tree view showing 'User_ID', 'Query2', and other metrics.

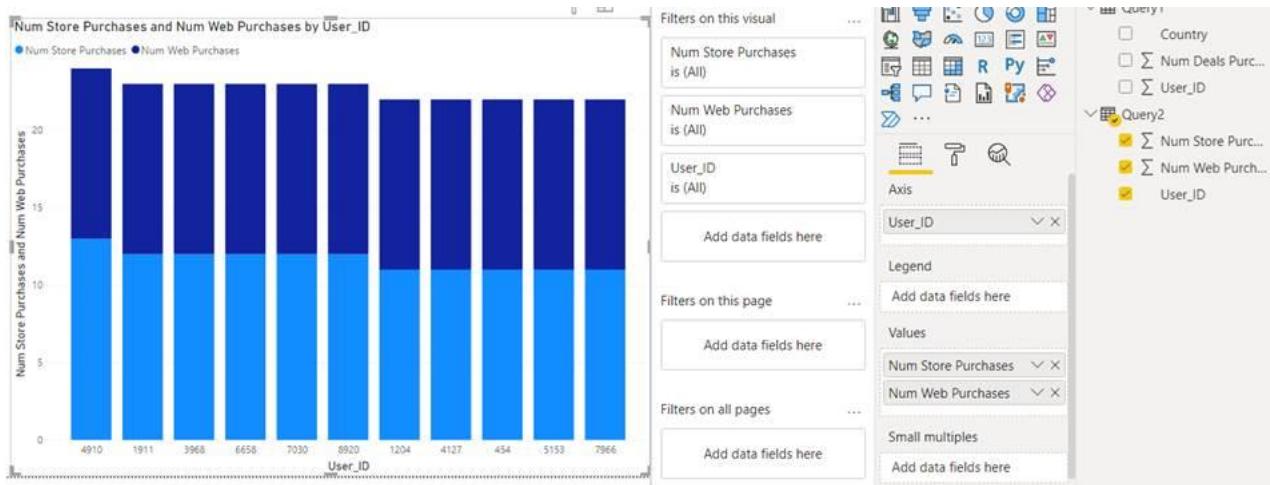
Kết quả thu được

Back to report | NUM STORE PURCHASES AND NUM WEB PURCHASES BY CUSTOMER_ID

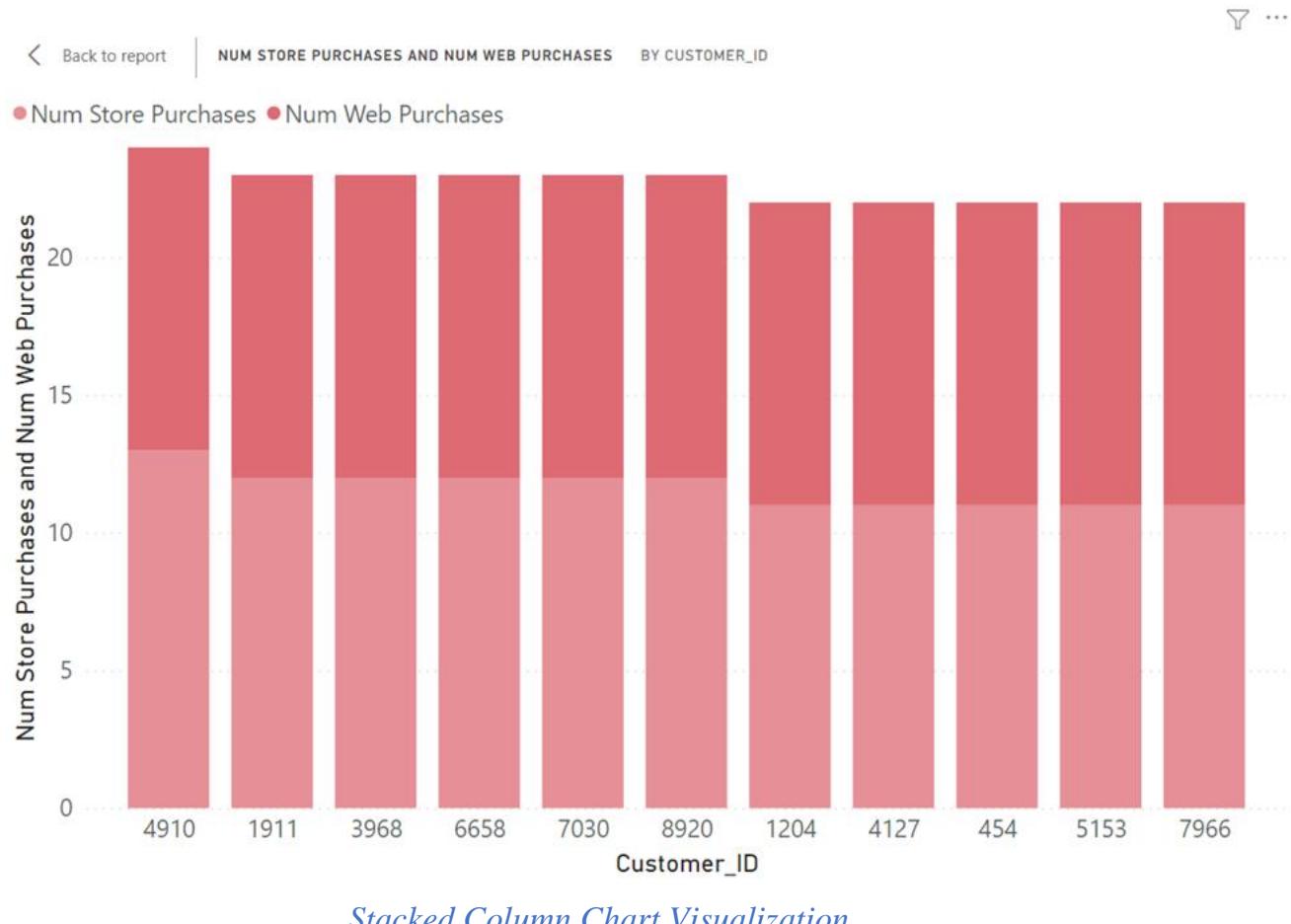
Customer_ID	Num Web Purchases	Num Store Purchases
1204	11.00	11.00
1911	11.00	12.00
3968	11.00	12.00
4127	11.00	11.00
454	11.00	11.00
4910	11.00	13.00
5153	11.00	11.00
6658	11.00	12.00
7030	11.00	12.00
7966	11.00	11.00
8920	11.00	12.00
Total	121.00	128.00

Matrix Visualization

Bước 4: Chọn Stacked Column Chart -> Chọn các thuộc tính cần thiết -> Stack các thuộc tính theo mong muốn -> Chọn format và chỉnh sửa report cho vừa mắt



Kết quả thu được



Kết quả report cuối cùng



Matrix and Stacked Column Chart

3.3.4 Report 3 – Câu 13

SQL Server Analysis Services database

Server ①

DESKTOP-RI1DGNT

Database

SSAS

Import

Connect live

MDX or DAX query (optional)

```
SELECT [Measures].[Tổng số đơn hàng] ON 0,
HEAD(ORDER([Dim Date].[Month Enroll].[Month Enroll],
[Measures].[Tổng số đơn hàng],DESC),3) ON 1
FROM [Data Warehouse Marketing Analytics]
WHERE[Dim Customer].[Country].&[South Africa];
```

OK

Cancel

Truy vấn MDX câu 13

Bước 2: Thực hiện đổi tên column, đổi kiểu dữ liệu các cột tương tự như câu 1

Bước 3: Chọn Matrix Visualization -> Chọn các thuộc tính cần thiết -> Stack các thuộc tính theo mong muốn -> Chọn format và chỉnh sửa report cho vừa mắt

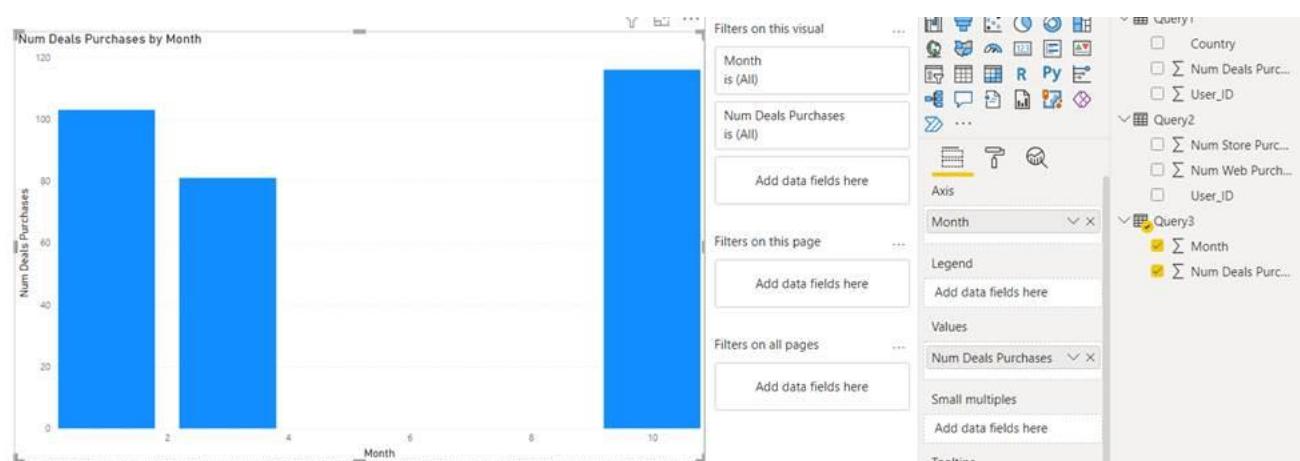
The screenshot shows a 'Matrix Visualization' titled 'Total Deal Purchase by Month'. The matrix structure is defined by 'Month' in the Rows and 'Total Deal Purchase' in the Values. The data points are 10.00, 8.00, 1.00, and a total of 1,715.00. The Power BI interface includes a ribbon at the top and a sidebar on the right containing 'Filters on this visual', 'Filters on this page', 'Filters on all pages', and sections for 'Rows', 'Columns', and 'Values'. A query pane on the right lists 'Query 1', 'Query 13' (selected), and 'Query 7'.

Kết quả thu được

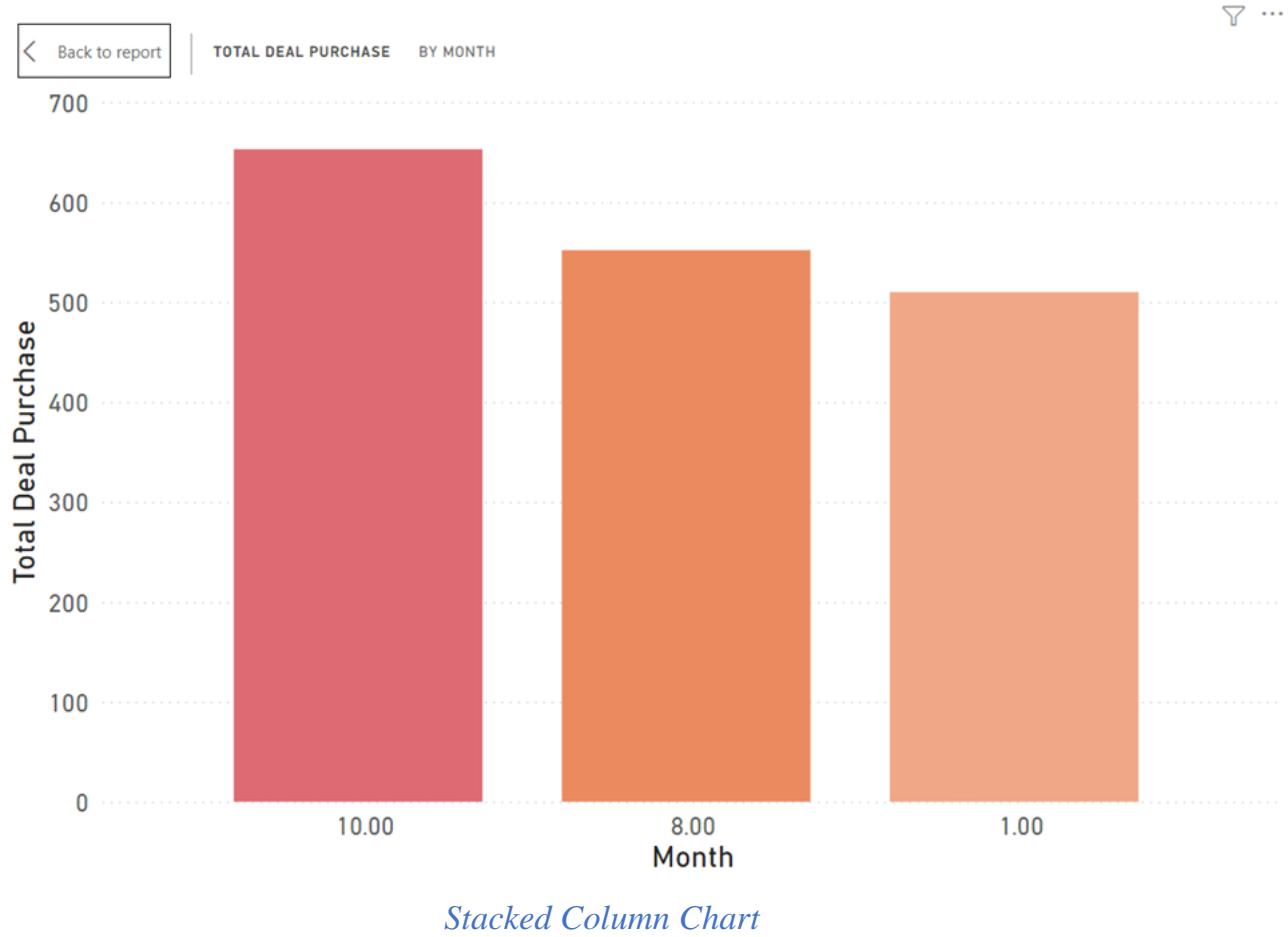
Total Deal Purchase by Month	
Month	Total Deal Purchase
10.00	653.00
8.00	552.00
1.00	510.00
Total	1,715.00

Matrix Visualization

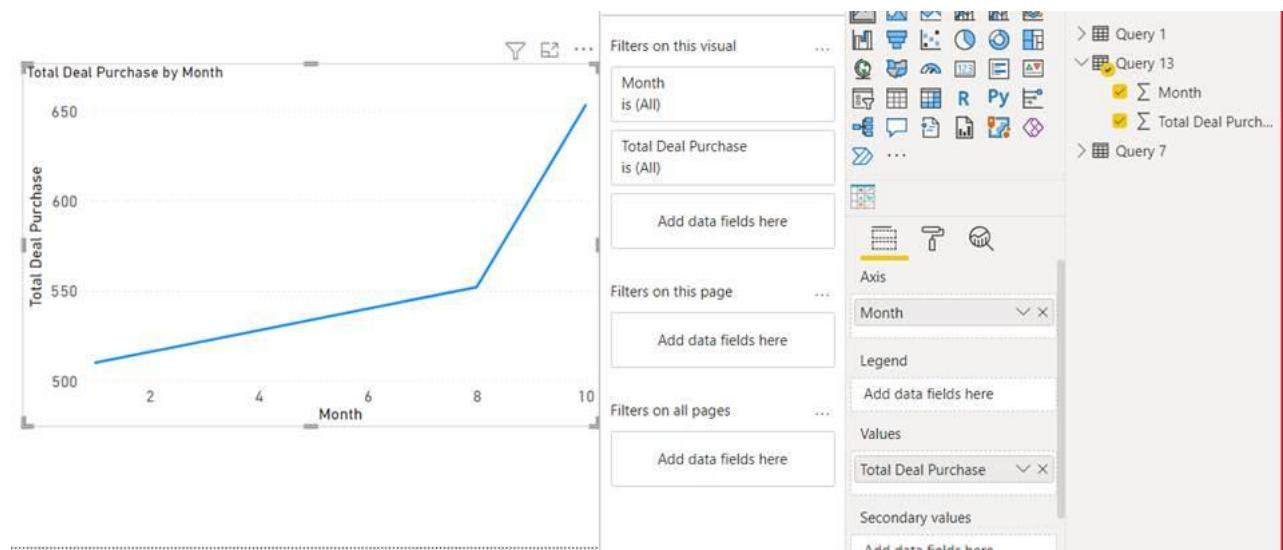
Bước 4: Chọn Stacked Column Chart -> Chọn các thuộc tính cần thiết -> Stack các thuộc tính theo mong muốn -> Chọn format và chỉnh sửa report cho vừa mắt



Kết quả thu được

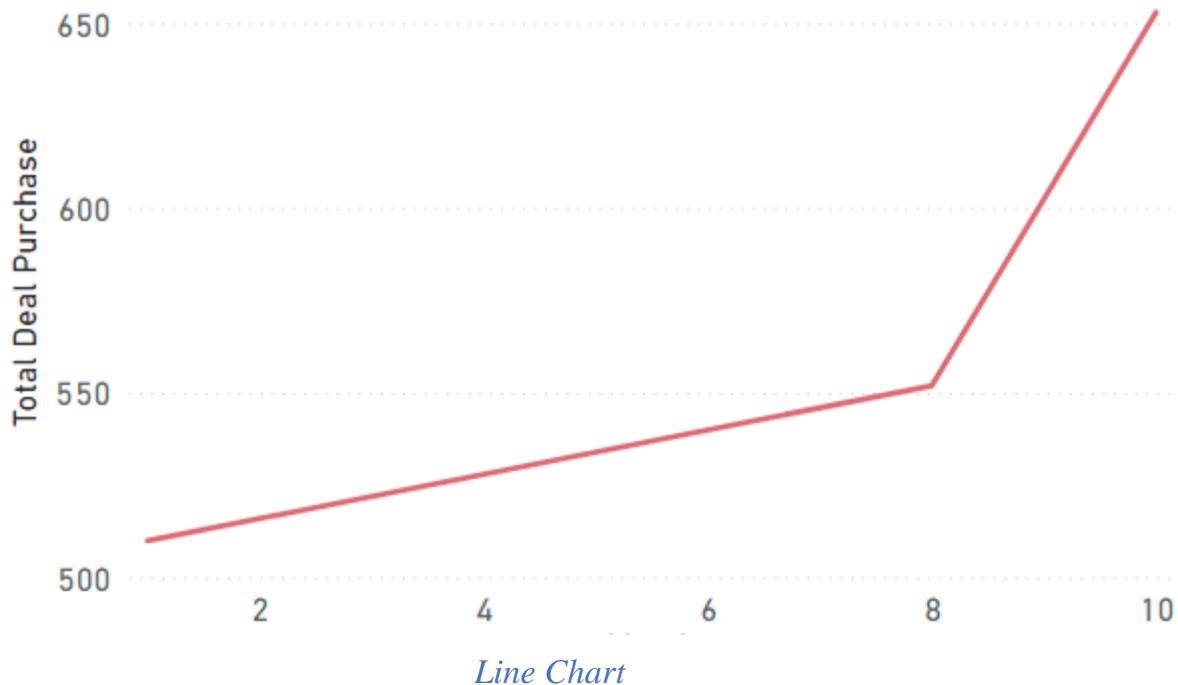


Bước 5: Chọn Line Chart -> Chọn các thuộc tính cần thiết -> Stack các thuộc tính theo mong muốn -> Chọn format và chỉnh sửa report cho vừa mắt
Kết quả cuối cùng



Kết quả thu được

Total Deal Purchase by Month



Kết quả cuối cùng

Total Deal Purchase by Month	
Month	Total Deal Purchase
10.00	653.00
8.00	552.00
1.00	510.00
Total	1,715.00



Stacked Column, Matrix and Line Chart

3.3.5 Publish Report lên Power BI Server

Bước 1: Click chọn Publish -> My WorkSpace

The screenshot shows the Microsoft Power BI desktop interface. A bar chart titled "Total Deal Purchase by Month" is displayed. The chart has three bars: one red bar at 10.00, one orange bar at 8.00, and one small orange bar at 1.00. The Y-axis ranges from 0 to 700. The X-axis is labeled "Month". In the top ribbon, the "Home" tab is selected. A "Publish" button is highlighted with a red circle. A "Publish to Power BI" dialog box is open, showing "My workspace" as the selected destination. The "Visualizations" and "Fields" panes are visible on the right side of the interface.

Bước 2: Hoàn thành Publish

The screenshot shows the Microsoft Power BI desktop interface after publishing. A "Publishing to Power BI" dialog box is open, indicating "Success!" and prompting to "Open 'THLONG_POWERBI.pbix' in Power BI". It also suggests "Get Quick Insights". The report title "Total Deal Purchase by Month" is visible above the chart. The chart shows the same data as before. The Power BI ribbon is visible at the top.

Bước 3: Truy cập Power BI trên web -> Chọn My WorkSpace

The screenshot shows the Microsoft Power BI web interface. On the left, there is a navigation menu with options like Home, Favorites, Recent, Create, Datasets, Goals, Apps, Shared with me, Learn, Workspaces, and My workspace. The "My workspace" option is selected. The main area displays a "Good evening, Long" greeting and a "Find and share actionable insights to make data-driven decisions" message. Below this is a "Recommended" section with four cards: "Getting started with Power BI" (Explore basic Power BI concepts), "Explore this data story" (Explore the 100 most useful productivity tips), "Explore this data story" (Cancer statistics in the USA), and "Getting started with Power BI" (Intro—What is Power BI?). At the top of the page, there is a search bar and buttons for "Open in Teams" and "Learn more".

Kết quả thu được

Power BI My workspace

HLONG_POWERBI | Data updated 11/13/21

Pages RP_1 RP_13 RP_7

Total Deal Purchase by Month

700
600
500
400
300
200
100
0

Total Deal Purchase

Month	Total Deal Purchase
1.00	510.00
8.00	552.00
10.00	653.00
Total	1,715.00

Filters

Search

Month is (All)

Total Deal Purchase is (All)

Get data

CHƯƠNG 4: KẾT LUẬN

4.1. Kết quả đạt được

Thực hiện xây dựng ‘*KHO DỮ LIỆU KHÁCH HÀNG QUA CHIẾN DỊCH MARKETING*’ giúp chúng em hiểu sâu hơn về kiến thức môn học Kho Dữ liệu và thành thạo hơn trong việc sử dụng các công cụ hỗ trợ quá trình ETL đưa dữ liệu vào kho dữ liệu như công cụ SSIS, SSAS, các công cụ để trực quan hóa dữ liệu như Excel Pivot, Power BI Desktop. Nâng cao kỹ năng nhìn nhận và tiếp thu, xử lý thông tin đầu vào, đặt những câu hỏi tiền đề và thảo luận hướng đi để giải quyết và trả lời câu hỏi.

4.2. Những hạn chế

Trong quá trình thực hiện xây dựng kho dữ liệu, nhóm chúng em không tránh khỏi gặp một số những khó khăn vì còn hạn chế về mặt kiến thức. Tập dữ liệu còn nhiều dữ kiện chưa được khai thác hết. Những thực hiện của chúng em chưa đạt đến mức nâng cao trong tổng thể quá trình xây dựng một kho dữ liệu hoàn chỉnh.

4.3. Tài liệu tham khảo

- [1] Tài liệu các file PDF hướng dẫn về Data Integration với SQL Server, SSIS, SSAS của GVHD Ths. Nguyễn Văn Thành.
- [2] Minewiskan(no date), *Lesson 3: Modifying Measures, Attributes and Hierarchies*. [online] learn.microsoft.com. Available at: <https://learn.microsoft.com/en-us/analysis-services/multidimensional-tutorial/lesson-3-modifying-measures-attributes-and-hierarchies?view=asallproducts-allversions>. [Accessed 12 May 2024].
- [3]Minewiskan(no date), *Lesson 4: Defining Advanced Attribute and Dimension Properties*. [online] learn.microsoft.com. Available at: <https://learn.microsoft.com/en-us/analysis-services/multidimensional-tutorial/lesson-4-defining-advanced-attribute-and-dimension-properties?view=asallproducts-allversions>. [Accessed 12 May 2024].
- [4] Nguyễn Văn Chúc (no date). *ETL Project From Excel Data Source to Star Schema with SSIS*. [online] youtube.com. Available at: https://www.youtube.com/watch?v=Yp8fXLnVCp8&ab_channel=ChucNguyenVan. [Accessed 10 May 2024].
- [5] lakshmi sowbhagya (no date). *SSIS -ENGLISH*. [online] youtube.com. Available at: https://youtube.com/playlist?list=PLMRFTMKnLG07hIPv5EPWnWUvWac9_DkSd&si=UQk1bdo629rU99Ij [Accessed 15 May 2024].