

NATIONAL UNIVERSITY OF HO CHI MINH CITY
UNIVERSITY OF INFORMATION TECHNOLOGY



UIT
TRƯỜNG ĐẠI HỌC
CÔNG NGHỆ THÔNG TIN

**BÁO CÁO ĐỀ TÀI
NGHIÊN CỨU TỶ LỆ TỰ TỬ TỪ NĂM
1985 – 2016**

Giảng viên hướng dẫn:

Ths. Đỗ Thị Minh Phụng

Nhóm thực hiện:

Phan Phạm Quỳnh Hoa – 19521520

Huỳnh Minh Thư – 19522304

Ho Chi Minh, ngày 3 tháng 3 năm 2022

Contents

1	Tổng quan đề tài.....	5
1.1	Lý do chọn đề tài	5
1.2	Mô tả dữ liệu.....	5
1.3	Thuộc tính dữ liệu gốc	6
2	Xây dựng kho dữ liệu.....	7
2.1.	Lược đồ kho dữ liệu (lược đồ hình sao)	7
2.2.	Mô tả chi tiết bảng Fact và Dim.....	8
3	Nội dung 15 câu truy vấn.....	10
4	Quá trình xây dựng kho dữ liệu SSIS	11
4.1.	Tạo project SSIS và thiết lập kết nối.....	11
4.2.	Chuẩn bị dữ liệu gốc, Import dữ liệu gốc	16
4.3.	Quá trình làm sạch dữ liệu.....	20
4.4.	Tạo các bảng Dimension	29
4.5.	Quá trình tạo bảng Fact	44
4.6.	Tạo ràng buộc khóa ngoại	53
4.7.	Thực thi Package.....	54
4.8.	Kiểm tra database	57
4.9.	Tạo và cấu hình ExecuteSQLTask	59
5	Quá trình phân tích kho dữ liệu SSAS và ngôn ngữ truy vấn MDX.....	67
5.1.	Tạo project SSAS	67
5.2.	Tạo Data Source	68
5.3.	Tạo Data Source View	72
5.4.	Tạo các khối (Cube) và xác định các chiều (Dimension)	75
5.5.	Định nghĩa Calculation	86
5.6.	Thực hiện câu truy vấn.....	87
1.	Thống kê tổng số vụ tự tử theo từng nhóm tuổi.....	87
2.	Thông tin quốc gia có tổng số vụ tự tử cao nhất và thấp nhất trong năm 2016.	89
3.	Tìm nhóm tuổi có tỉ lệ người tự tử cao nhất ở United States qua các năm.....	91
4.	Liệt kê top 10 quốc gia có tỉ lệ tự tử cao nhất trong năm 2016, sắp xếp theo thứ tự tỉ lệ tự tử giảm dần.....	93
5.	Liệt kê những quốc gia và năm từng có tổng số vụ tự tử cao hơn 10 000 vụ.....	96

6.	Cho biết tên quốc gia, tổng số vụ tự tử, tỉ lệ tự tử của quốc gia có GDP bình quân đầu người cao nhất và thấp nhất trong năm 2016.....	99
7.	Mỗi quốc gia, tìm năm có tổng số vụ tự tử cao nhất.....	101
8.	Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính.....	104
9.	Mỗi quốc gia, cho biết nhóm tuổi có tỷ lệ tử tử cao nhất và thấp nhất.....	106
10.	Cho biết Top10 nước có tỷ lệ tự tử ở độ tuổi 15-24 và 25-34 nhiều nhất từ năm 2000-2016.	110
11.	Mỗi năm, tìm quốc gia có tỷ lệ tự tử cao nhất và thấp nhất, đồng thời hiển thị GDP bình quân đầu người của quốc gia đó.....	113
12.	Tỉ lệ tự tử ở từng giới tính của Top 10 quốc gia có số dân đông nhất năm 2012.....	117
13.	Trong khoảng 10 năm gần nhất, năm nào thế giới ghi nhận tổng số vụ tự tử cao nhất.	119
14.	Mỗi năm, thống kê số vụ tự tử ở các nước theo GDP For Year giảm dần.....	120
15.	Top 10 quốc gia có tổng số vụ tự tử cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính.	123
6	Quá trình lập báo biểu SSRS.....	125
6.1.	Tạo project SSRS	125
6.2.	Tạo kết nối kho dữ liệu Data Source	128
6.3.	Lập báo biểu trên Report Builder	131
6.3.1.	Report1 (Report Totals) Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính (Câu truy vấn 8).	131
6.3.2.	Report2 (Report Grouping) Mỗi quốc gia, cho biết nhóm tuổi có tỷ lệ tử tử cao nhất và thấp nhất (Câu truy vấn 9).	142
6.3.3.	Report3 (Report Totals + Grouping) Top 10 quốc gia có tổng số vụ tự tử cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính (Câu truy vấn 15).	153
7	Quá trình lập báo biểu bằng công cụ Power BI	161
7.1.	Đăng ký dịch vụ Power BI.....	161
7.2.	Tải Power BI Desktop trên Microsoft Store	163
7.3.	Tạo Report trên Power Bi Desktop	165
7.3.1.	Report1 (Report Totals) Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính.	165
7.3.2.	Report2 (Report Grouping) Mỗi quốc gia, cho biết nhóm tuổi có tỷ lệ tử tử cao nhất và thấp nhất.	172
7.3.3.	Report3 (Report Totals + Grouping) Top 10 quốc gia có tổng số vụ tự tử cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính.	177
8	Quá trình Data mining.....	179
8.1.	Chọn lại kho dữ liệu Marketing Analytics.....	179
8.2.	Chuẩn bị dữ liệu gốc, Import dữ liệu	181

8.3. Quá trình Data mining sử dụng thuật toán Cây quyết định (Microsoft Decision Trees) .	187
8.3.1. Tạo project SSAS	187
8.3.2. Tạo Data Source	188
8.3.3. Tạo Data Source Views.....	191
8.3.4. Tạo cấu trúc data mining (Mining Structures)	193
8.4. Phân tích kết quả thuật toán Cây quyết định(Decision Trees) và đưa ra tập luật.....	205
8.5. So sánh độ chính xác với các thuật toán khác	209
Tài liệu tham khảo	213
Link Đồ án cuối kỳ OLAP IS217M22HTCL_19522304_19521520.....	213

1 Tổng quan đề tài

1.1 Lý do chọn đề tài

Hàng năm, trên thế giới có trên 800.000 người chết vì tự tử, và con số toan tự tử còn cao hơn thế. Tổ chức Y tế thế giới (WHO) công bố số liệu cho thấy, cứ 40 giây có 1 người tự tử. Mặc dù xảy ra ở tất cả độ tuổi, tự tử là nguyên nhân gây tử vong thứ hai đối với nhóm tuổi từ 15 đến 29 trên thế giới trong năm 2012 (sau tai nạn giao thông). Những người có xu hướng tự tử rất đa dạng (giới tính, độ tuổi, trình độ học vấn,...), kể cả những người thuộc các tầng lớp khác nhau, sống ở đất nước có mức độ phát triển khác nhau,...

Mức độ của vấn nạn tử tự quan ngại sẽ ngày càng tăng và ảnh hưởng lớn đến đời sống xã hội, vì vậy cần có những hành động cụ thể để đối phó với vấn đề này. Trước khi muốn đưa ra giải pháp, cần tìm hiểu và đặt vấn đề. Ví dụ ở độ tuổi nào có xu hướng tự tử tăng nhanh nhất? Ở những nước phát triển tỷ lệ tự tử cao hơn những nước kém phát triển hay không? Cần đưa ra những số liệu cụ thể, được phân tích dựa trên bộ dữ liệu thực tế để có cái nhìn trực quan hơn về vấn đề và đưa ra các giải pháp, chính sách và hướng giải quyết giúp giảm tỷ lệ người tự tử.

Nhóm thực hiện mong muốn sau khi đề tài được hoàn thành sẽ để lại được những thông tin và kiến thức có thể đóng góp một phần vào công trình nghiên cứu của đất nước và thế giới.

1.2 Mô tả dữ liệu

Tên bộ dữ liệu: Suicide Rates Overview 1985 to 2016.

Ngày cập nhật gần nhất: 02/12/2018.

Kho dữ liệu thu thập thông tin về tỉ lệ tự tử trên 101 quốc gia từ năm 1985 đến năm 2016. Được xây dựng để tìm các tín hiệu liên quan đến việc gia tăng tỷ lệ tự tử giữa các nhóm khác nhau trên toàn cầu, trên phạm vi kinh tế xã hội.

Kho dữ liệu được xây dựng với hướng chủ đề phân tích tỉ lệ tự tử.

Kho dữ liệu gồm: 27820 dòng, 12 thuộc tính.

Link dataset: <https://www.kaggle.com/datasets/russellyates88/suicide-rates-overview-1985-to-2016>

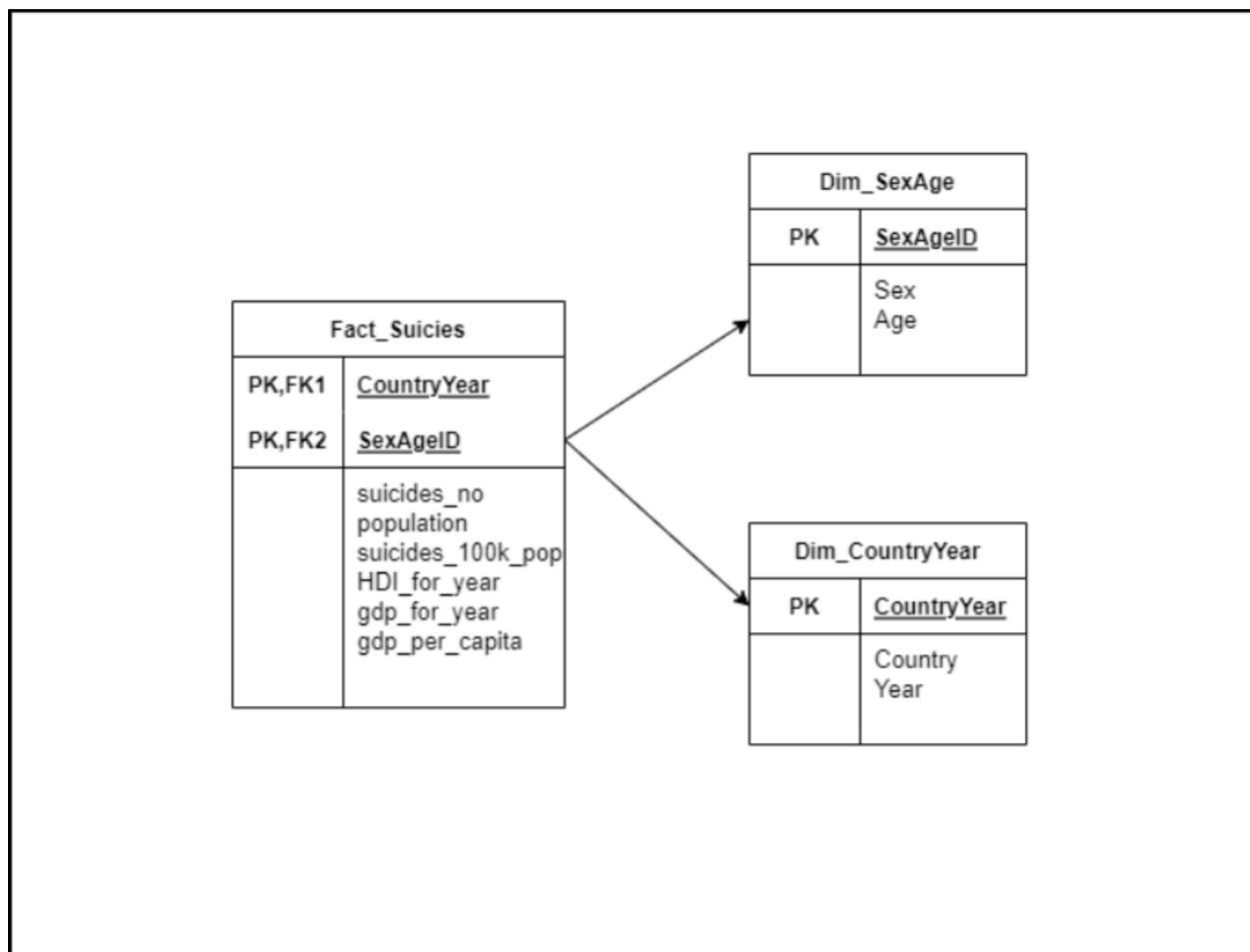
1.3 Thuộc tính dữ liệu gốc

STT	Tên thuộc tính	Kiểu dữ liệu	Ý nghĩa
1	country	Text	Tên quốc gia
2	year	Number	Năm thu thập dữ liệu
3	sex	Text	Giới tính 1. Male 2. Female
4	age	Number	Nhóm tuổi - 5-14 years - 15-24 years - 25-34 years - 35-54 years - 55-74 years - 75+ years
5	suicides_no	Number	Số vụ tự tử
6	population	Number	Dân số
7	suicides/100k pop	Number	Tỉ lệ tự tử trên 100k người
8	country-year	Text	Tổng hợp tên quốc gia và năm
9	HDI for year	Number	Human development index: Chỉ số phát triển con người theo năm.
10	gdp_for_year	Number	Chỉ số GDP theo năm (Đơn vị \$)
11	gdp_per_capita	Number	GDP bình quân đầu người (Đơn vị \$)

12	generation	Text	<p>Thé hệ (dựa trên mức trung bình của nhóm tuổi).</p> <ul style="list-style-type: none"> - Generation X - Millenials - Boomers - Silent - Generation Z - G.I. Generation
----	------------	------	---

2 Xây dựng kho dữ liệu

2.1. Lược đồ kho dữ liệu (lược đồ hình sao)



2.2. Mô tả chi tiết bảng Fact và Dim

Fact: Fact_Suicides

Measures: Suicides_no, Population, Suicides_rate, HDI, gdp_for_year, gdp_per_capita

Dimention: Dim_CountryYear, Dim_SexAge

2.2.1. Bảng Fact_Suicides

Tên thuộc tính	NULL	Kiểu dữ liệu	Mô tả thuộc tính
CountryYear (PK)		Varchar	Tên quốc gia và năm
SexAgeID (PK)		Varchar	Mã
suicides_no		Number	Số vụ tự tử
population		Number	Dân số
Suicides_100k_pop		Number	Tỉ lệ tự tử trên 100k người
HDI_for_year		Number	Human development index: Chỉ số phát triển con người theo năm.
gdp_for_year		Number	Chỉ số GDP theo năm (Đơn vị \$)
gdp_per_capita		Number	GDP bình quân đầu người (Đơn vị \$)

2.2.2. Bảng Dim_CountryYear

Tên thuộc tính	NULL	Kiểu dữ liệu	Mô tả thuộc tính
CountryYear (PK)		Varchar	Tên quốc gia và năm
Country		Varchar	Tên quốc gia
Year		Int	Năm

2.2.3. Bảng Dim_SexAge

Tên thuộc tính	NULL	Kiểu dữ liệu	Mô tả thuộc tính
SexAgeID (PK)		Varchar	Mã giới tính nhóm tuổi
Sex		Varchar	Giới tính - Male - Female
Age		Varchar	Nhóm tuổi - 5-14 years - 15-24 years - 25-34 years - 35-54 years - 55-74 years - 75+ years

3 Nội dung 15 câu truy vấn

1. Thống kê tổng số vụ tự tử theo từng nhóm tuổi.
2. Thông tin quốc gia có tổng số vụ tự tử cao nhất và thấp nhất trong năm 2016.
3. Tìm nhóm tuổi có tỉ lệ người tự tử cao nhất ở United States qua các năm.
4. Liệt kê top 10 quốc gia có tỉ lệ tự tử cao nhất trong năm 2016, sắp xếp theo thứ tự tỉ lệ tự tử giảm dần.
5. Liệt kê những quốc gia và năm từng có tổng số vụ tự tử cao hơn 10 000 vụ.
6. Cho biết tên quốc gia, tổng số vụ tự tử, tỉ lệ tự tử của quốc gia có GDP bình quân đầu người cao nhất và thấp nhất trong năm 2016.
7. Mỗi quốc gia, tìm năm có tổng số vụ tự tử cao nhất.
8. Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính, sắp xếp theo nhóm tuổi tăng dần.
9. Mỗi quốc gia, cho biết nhóm tuổi có tỷ lệ tử tử cao nhất và thấp nhất.
10. Cho biết Top10 nước có tỷ lệ tự tử ở độ tuổi 15-24 và 25-34 nhiều nhất từ năm 2000-2016.
11. Mỗi năm, tìm quốc gia có tỷ lệ tự tử cao nhất và thấp nhất, đồng thời hiển thị GDP bình quân đầu người của quốc gia đó.
12. Tỉ lệ tự tử ở từng giới tính của Top 10 quốc gia có số dân đông nhất năm 2012.
13. Trong khoảng 10 năm gần nhất, năm nào thế giới ghi nhận tổng số vụ tự tử cao nhất.
14. Thống kê số vụ tự tử ở các nước theo GDP For Year giảm dần.
15. Top 10 quốc gia có tổng số vụ tự tử cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính.

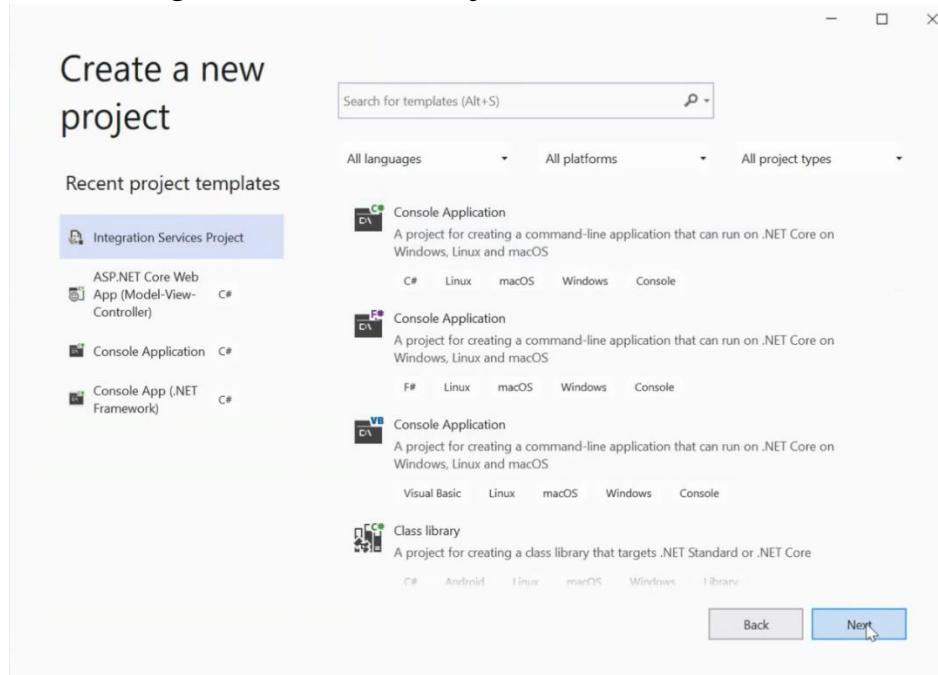
4 Quá trình xây dựng kho dữ liệu SSIS

4.1. Tạo project SSIS và thiết lập kết nối

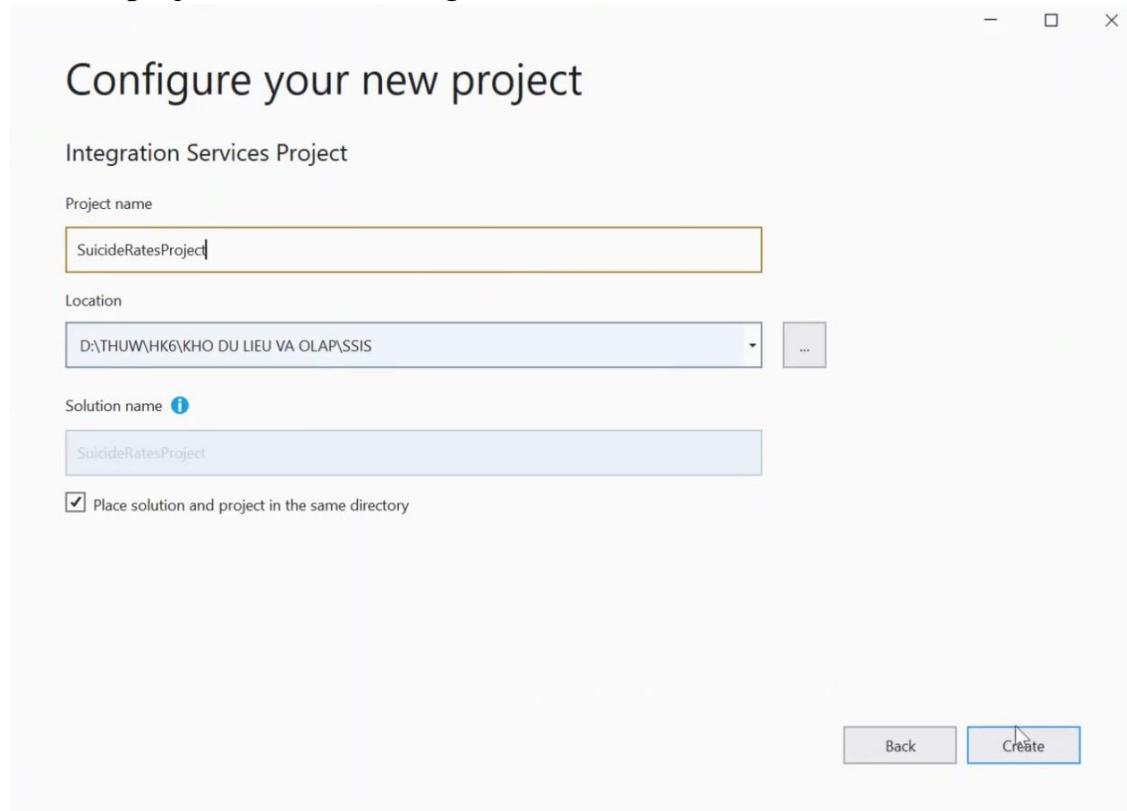
4.1.1. Tạo project SSIS

Mở Visual Studio (phiên bản 2019) và tạo mới project.

Chọn Integration Services Project.



Đặt tên project và chọn đường dẫn lưu trữ.



4.1.2. Tạo cơ sở dữ liệu và thiết lập kết nối

⊕ Tạo cơ sở dữ liệu

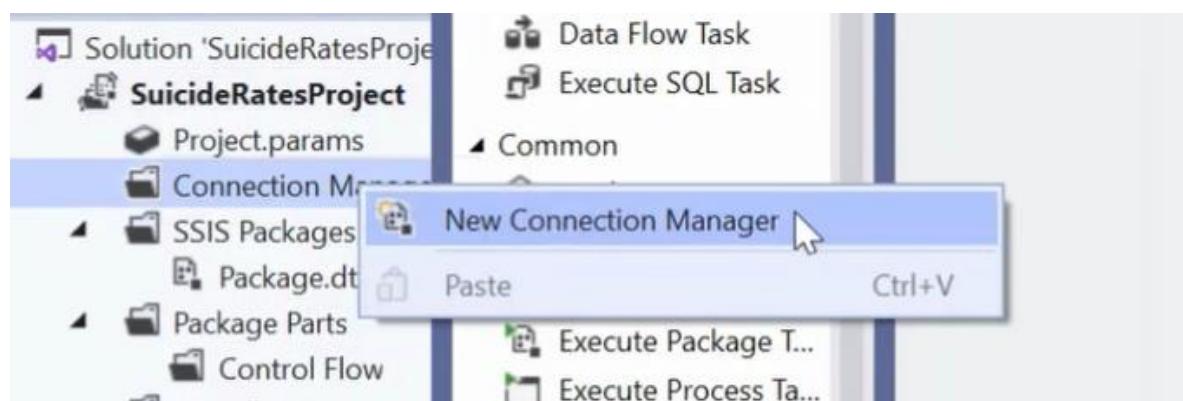
- OriginalData: chứa bảng dữ liệu gốc, bảng dữ liệu null và bảng dữ liệu sau khi làm sạch.
- SuicideRates_DW: kho dữ liệu chứa các bảng Dimension và bảng Fact

```
create database OriginalData;
create database SuicideRates_DW;
```

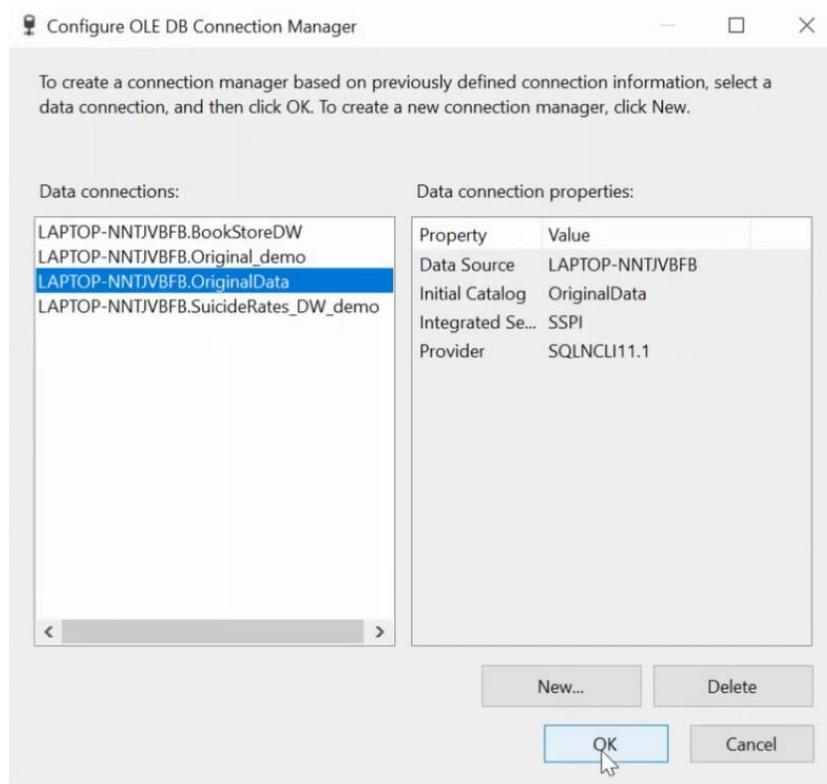
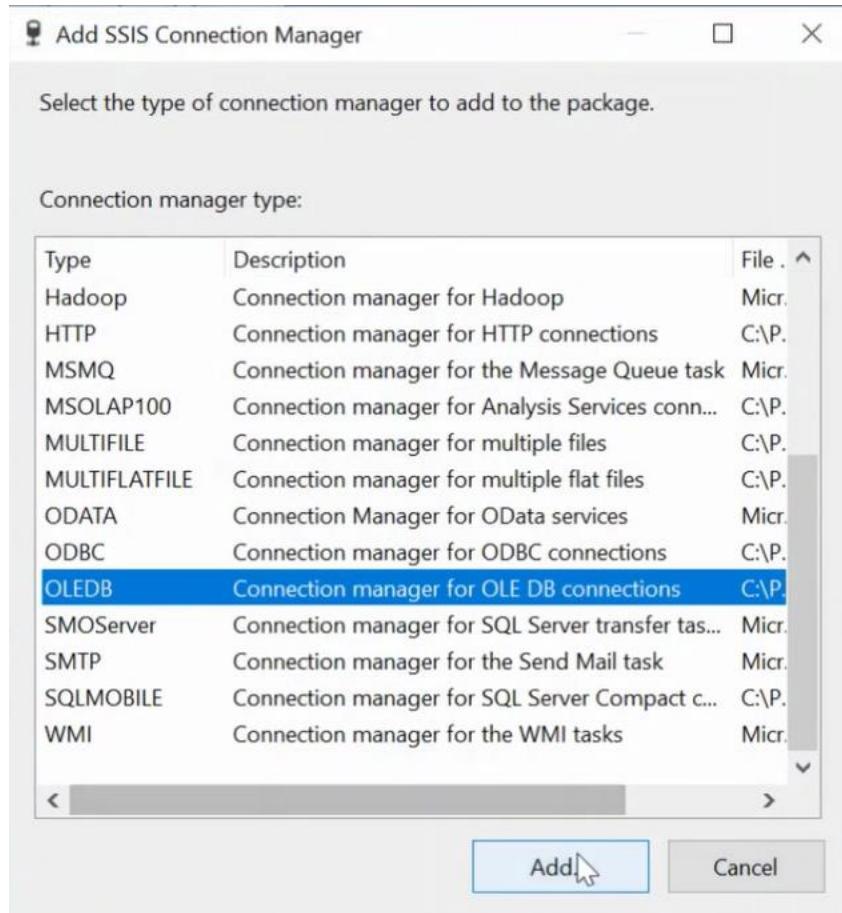
⊕ Thiết lập kết nối

- a. Kết nối tới database OriginalData

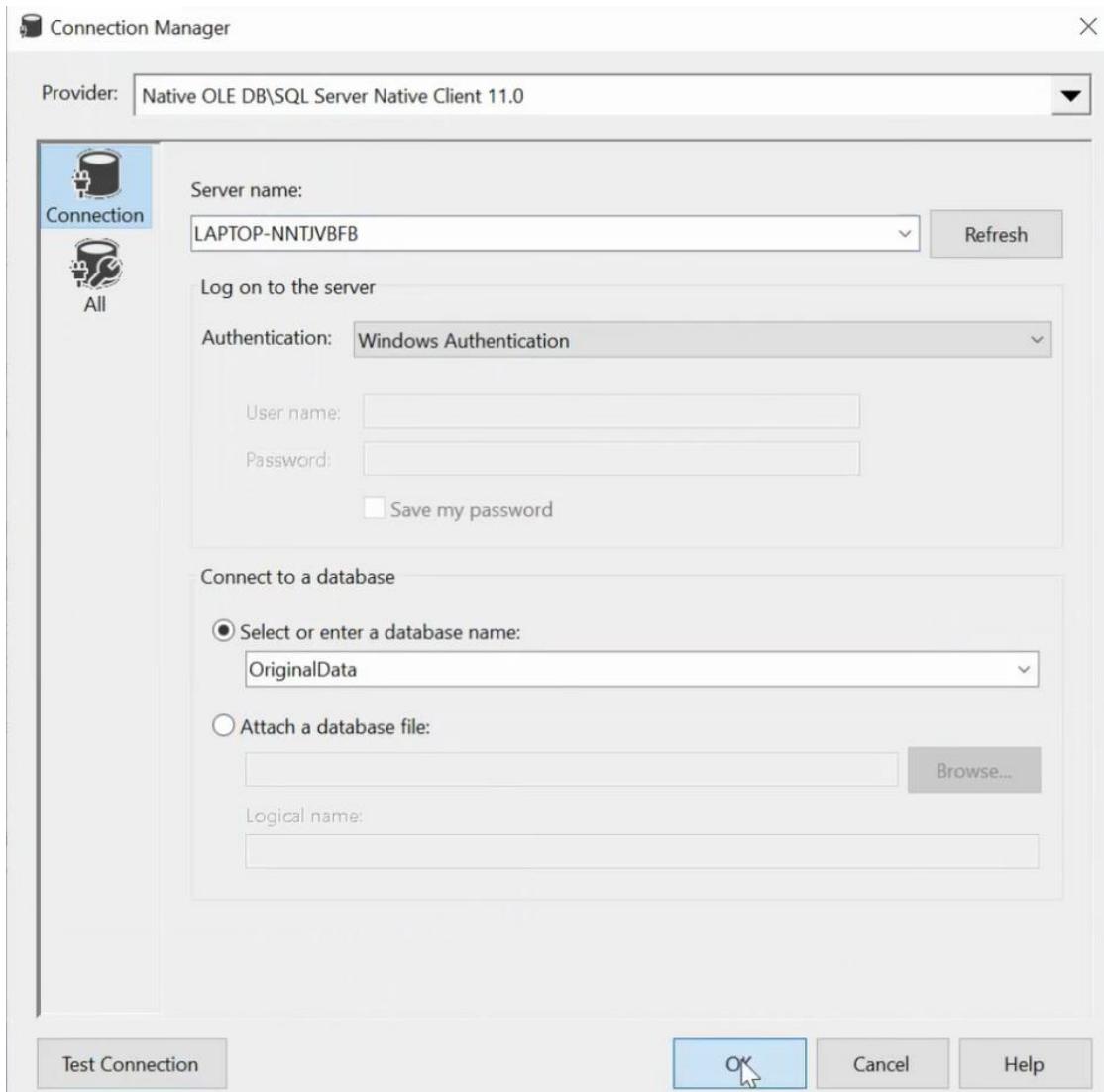
Tại cửa sổ Solution Explorer, click phải vào mục Connection Manager chọn New Connection Manager.



Chọn loại Connection là OLEDB và nhấn Add.

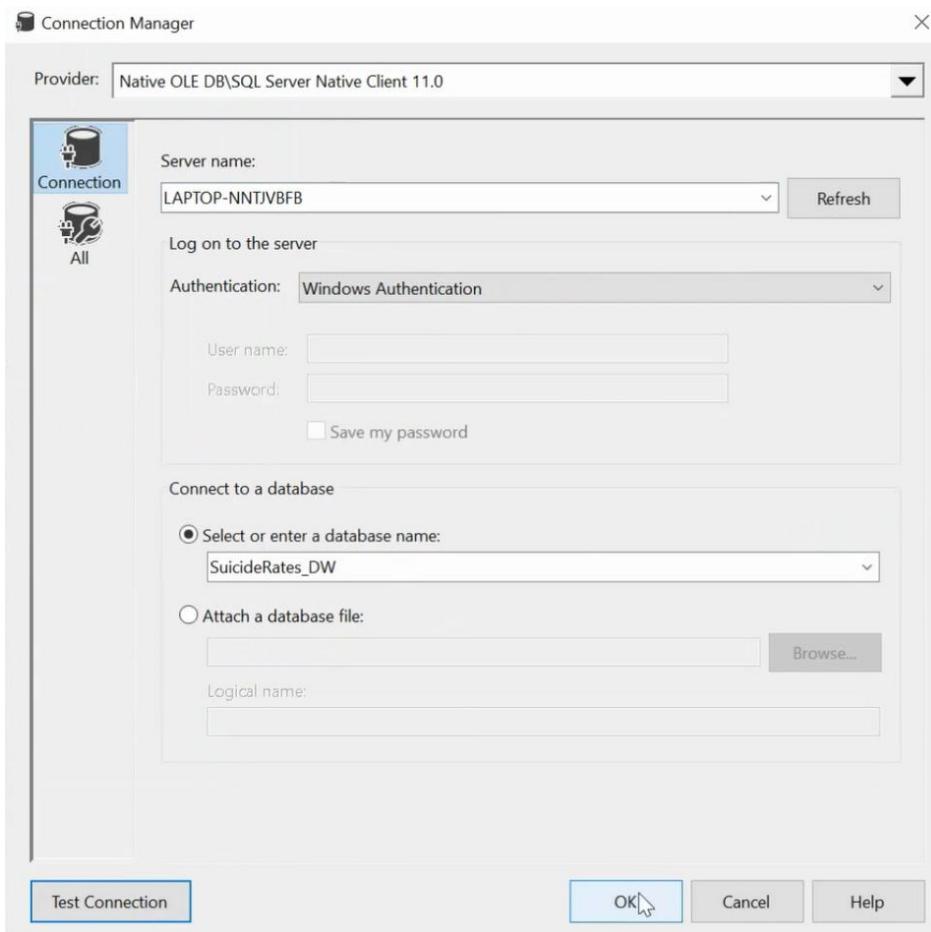
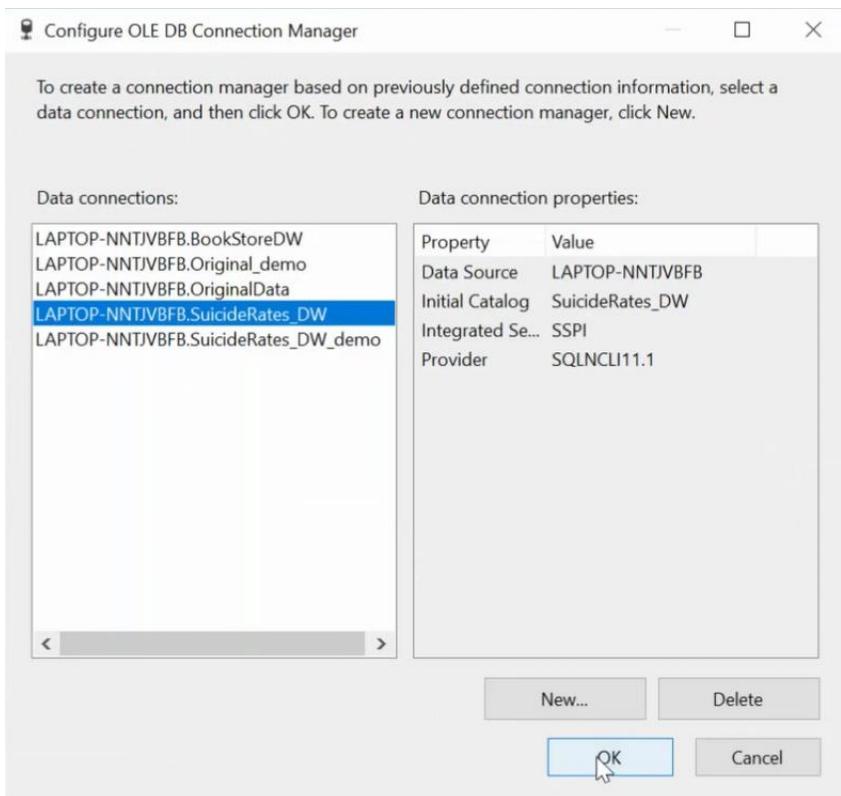


Ở mục Server name chọn tên server cục bộ ở laptop. Chọn database OriginalData ở phần Connect to a database sau đó chọn OK.

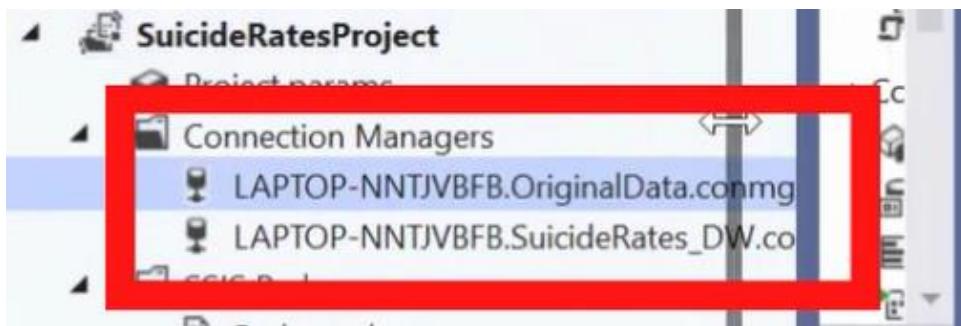


Nhấn OK.

- b. Kết nối tới database SuicideRates_DW
 - Tương tự các bước kết nối tới OriginalData
 - Đến bước chọn Data Connection, chọn database SuicideRates_DW

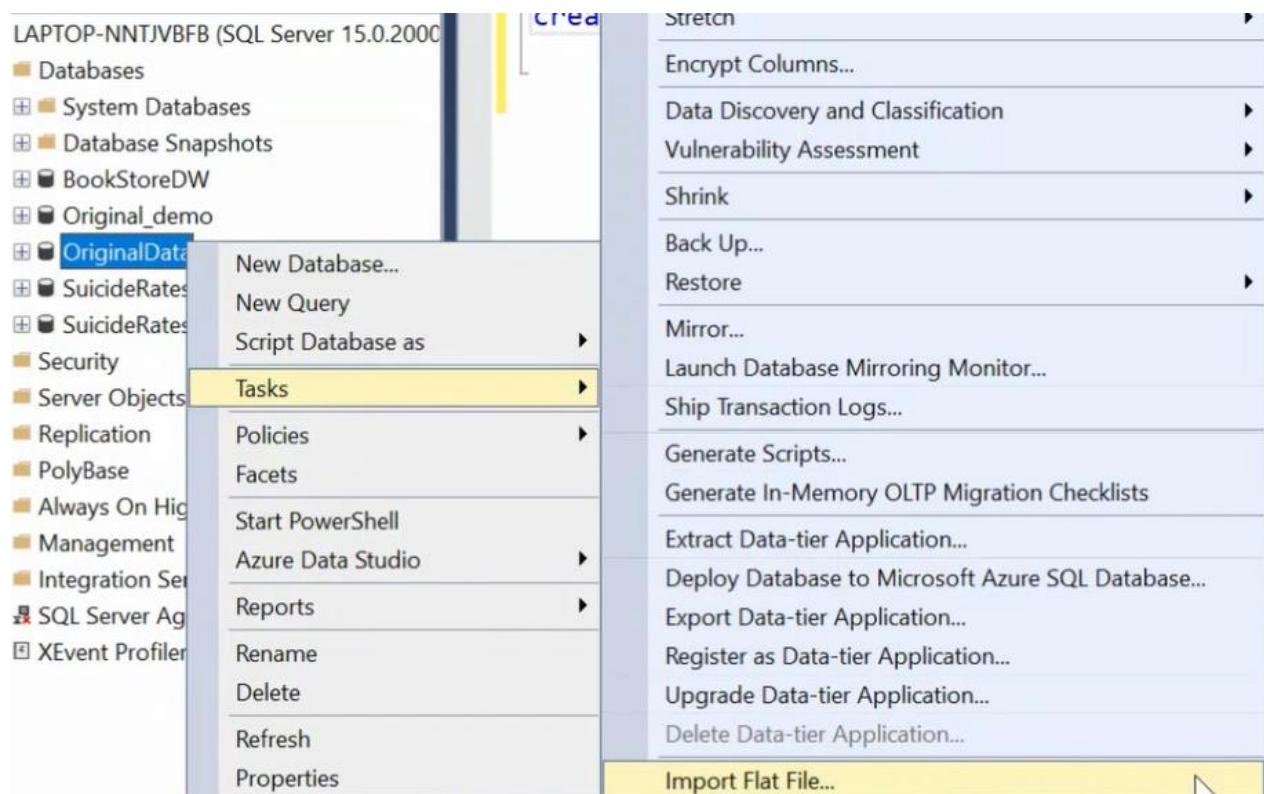


- ➡ Tạo thành công 2 kết nối đến OriginalData và SuicideRates_DW

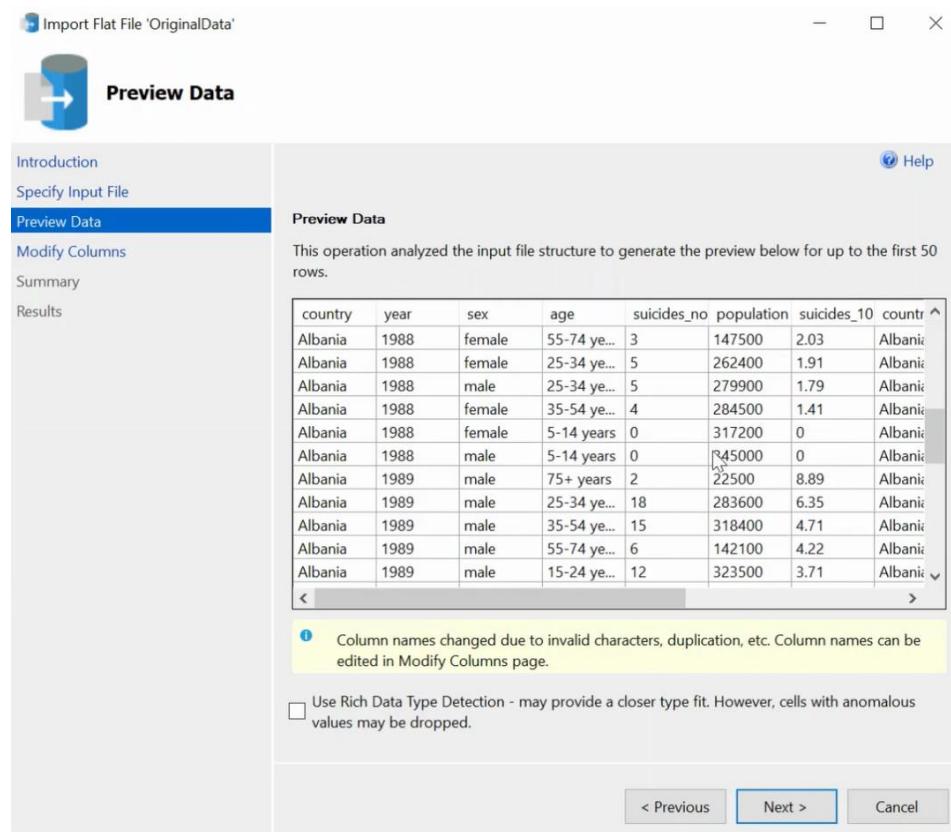
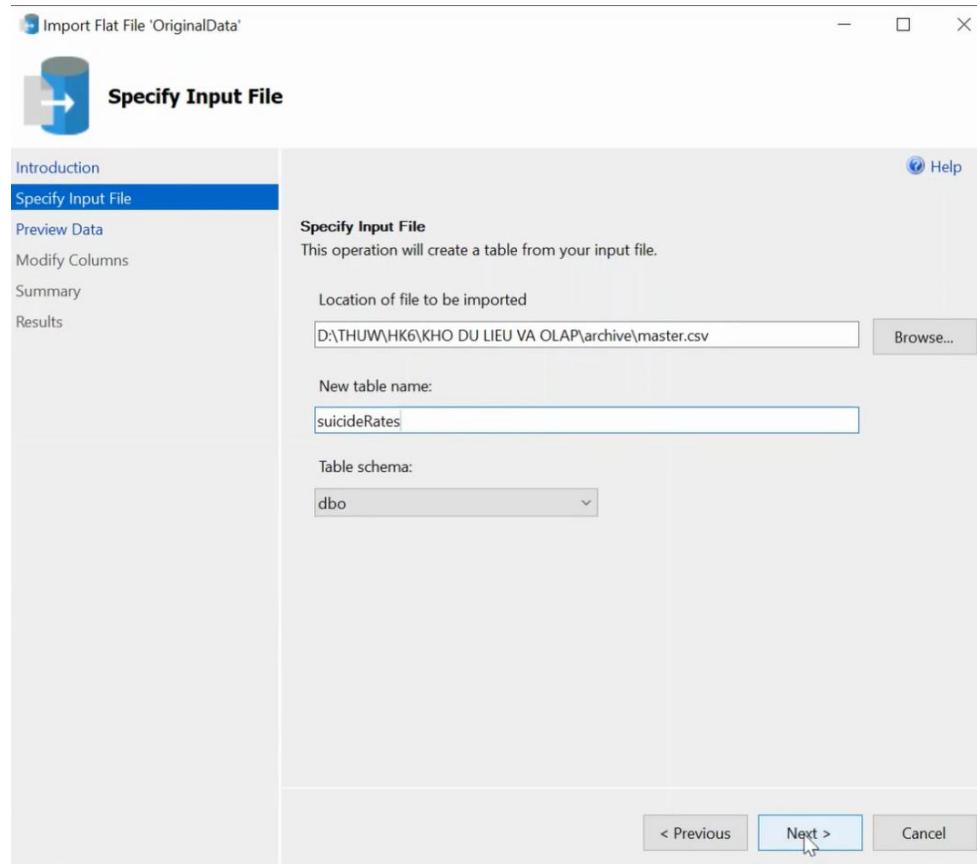


4.2. Chuẩn bị dữ liệu gốc, Import dữ liệu gốc

Click phải vào database OriginalData, chọn Tasks → Import Flat File..



Chọn đường dẫn lưu trữ và đặt tên cho bảng. Chọn Next.



Điều chỉnh kiểu dữ liệu cho phù hợp

 **Modify Columns**

Introduction
Specify Input File
Preview Data
Modify Columns
Summary
Results

Help

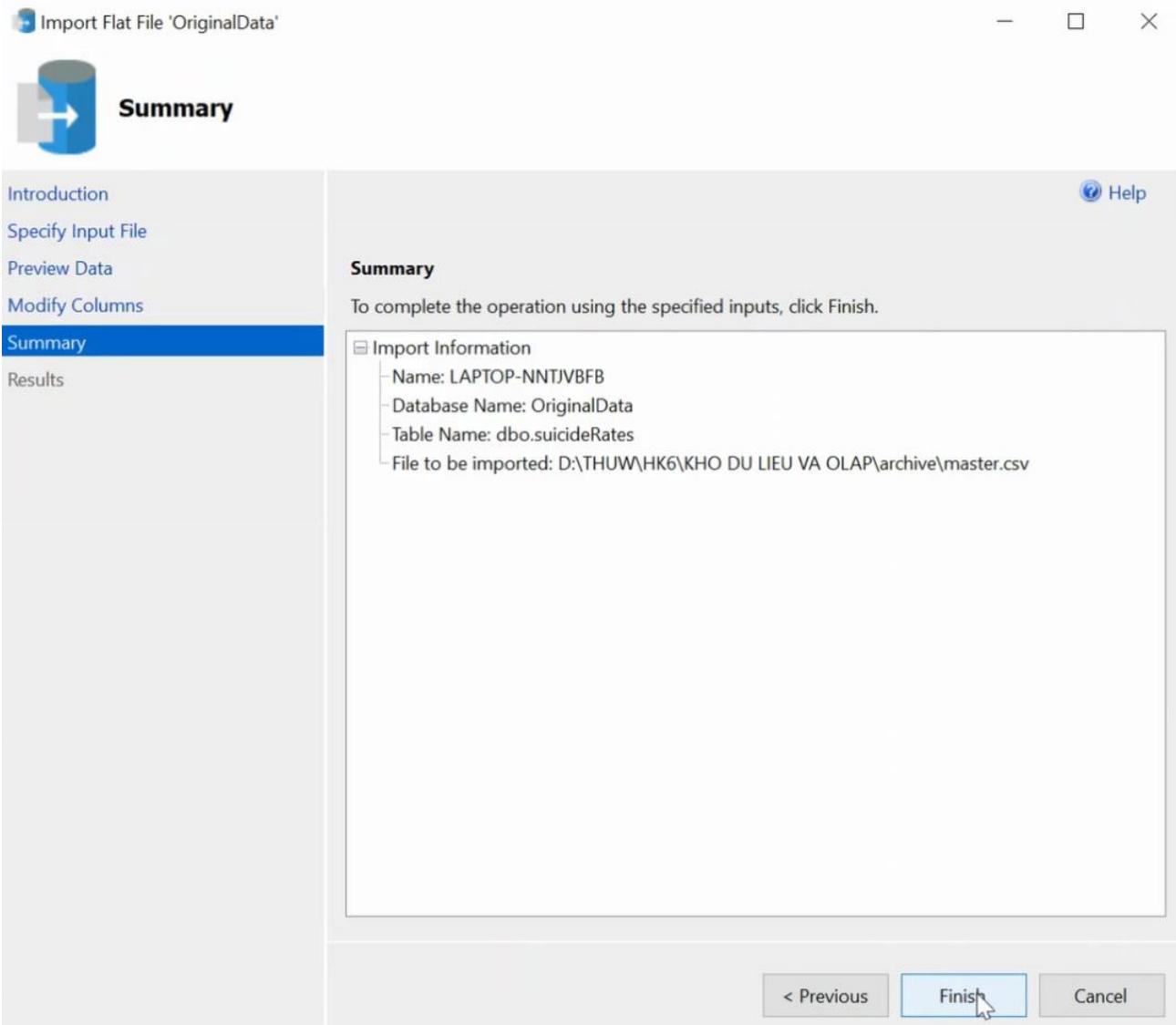
Modify Columns

This operation generated the following table schema. Please verify if schema is accurate, and if not, please make any changes.

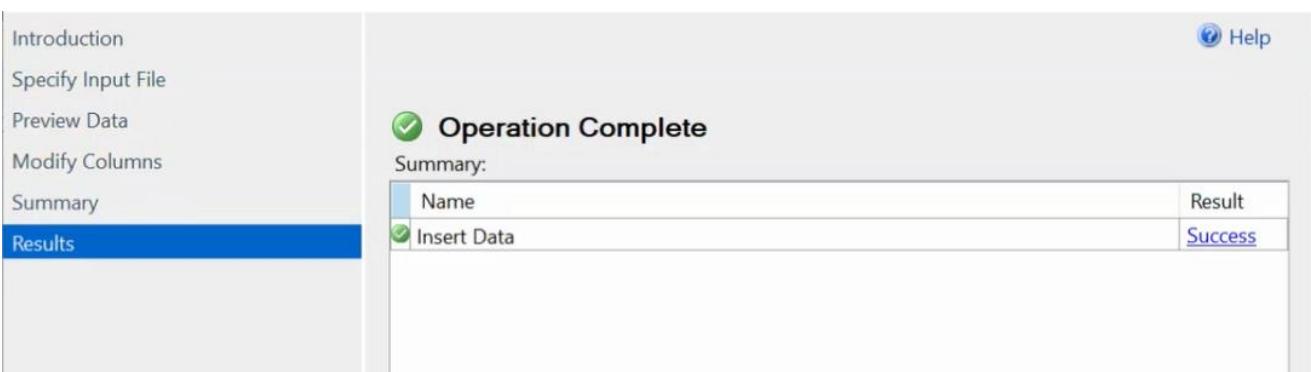
Column Name	Data Type	Primary Key	<input type="checkbox"/> Allow Nulls
country	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
year	int	<input type="checkbox"/>	<input type="checkbox"/>
sex	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
age	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
suicides_no	int	<input type="checkbox"/>	<input type="checkbox"/>
population	int	<input type="checkbox"/>	<input type="checkbox"/>
suicides_100k_pop	float	<input type="checkbox"/>	<input type="checkbox"/>
country_year	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>
HDI_for_year	float	<input type="checkbox"/>	<input checked="" type="checkbox"/>
gdp_for_year	float	<input type="checkbox"/>	<input type="checkbox"/>
gdp_per_capita	int	<input type="checkbox"/>	<input type="checkbox"/>
generation	nvarchar(50)	<input type="checkbox"/>	<input type="checkbox"/>

Error Reporting - Selecting a smaller range may have a performance impact

< Previous Cancel

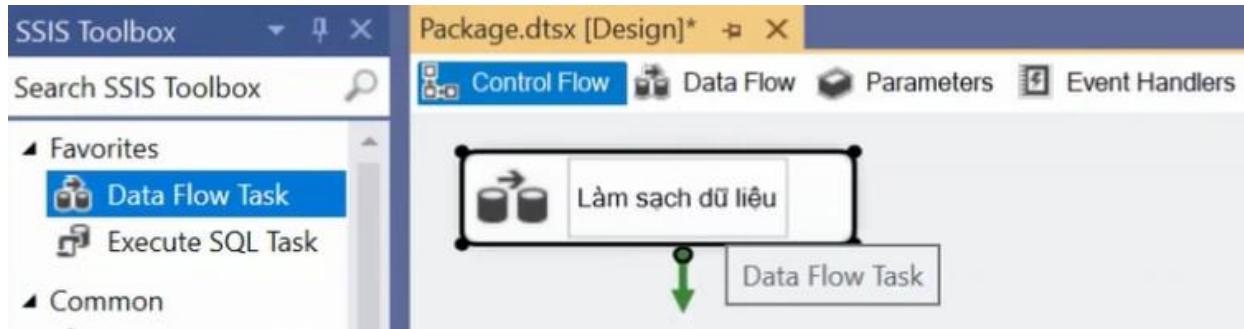


Import dữ liệu thành công.

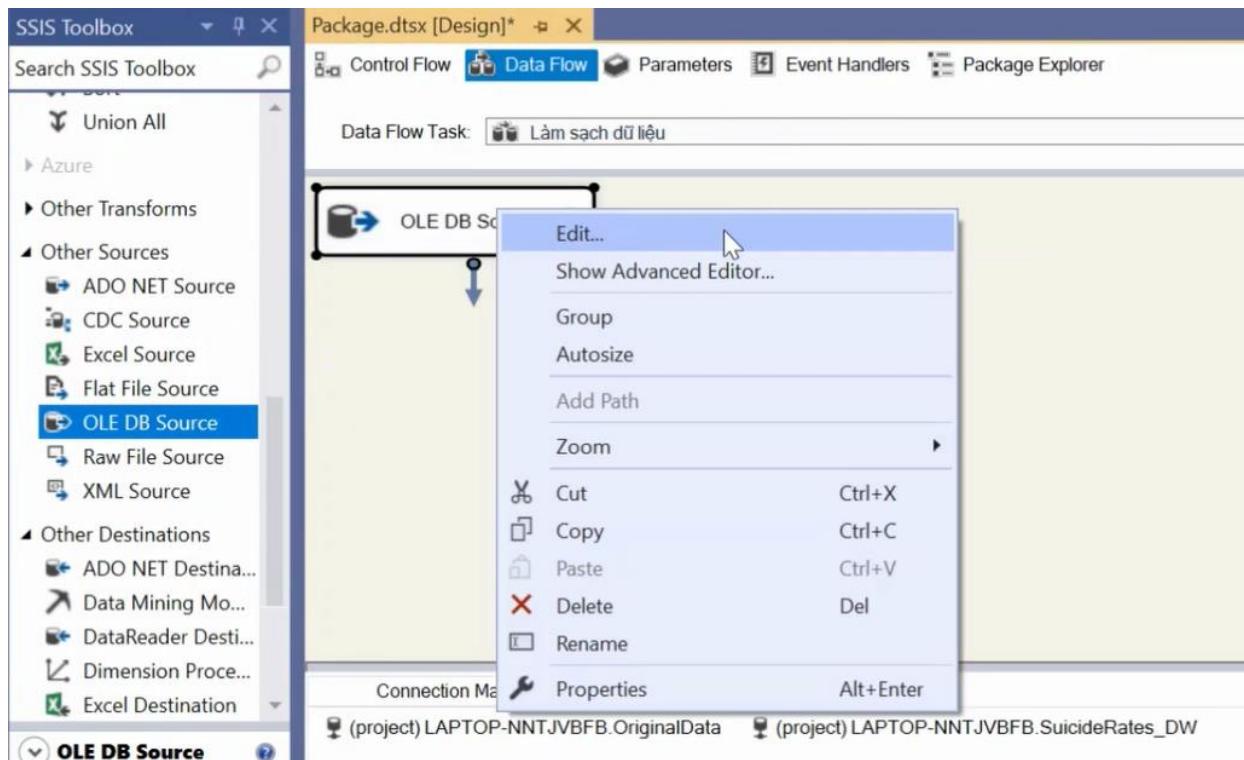


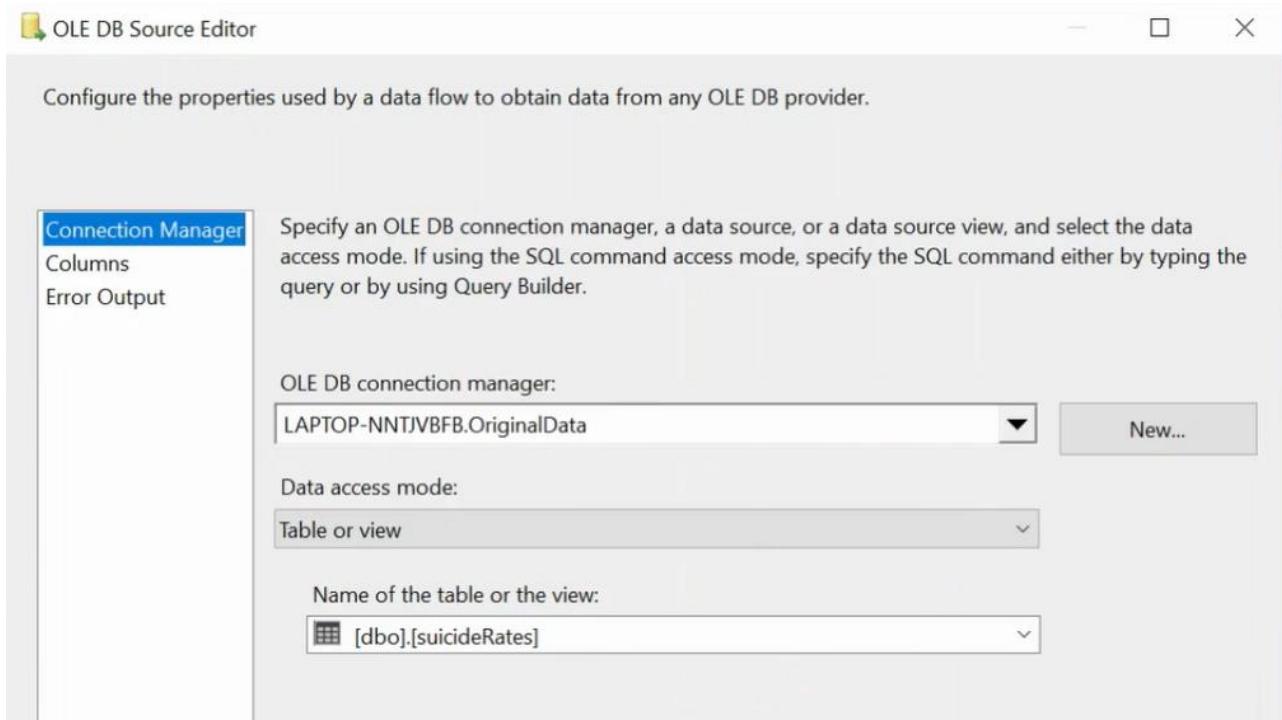
4.3. Quá trình làm sạch dữ liệu

- Mở visual studio
- Kéo Data Flow Task ở SSIS Toolbox để tạo Quá trình làm sạch dữ liệu.

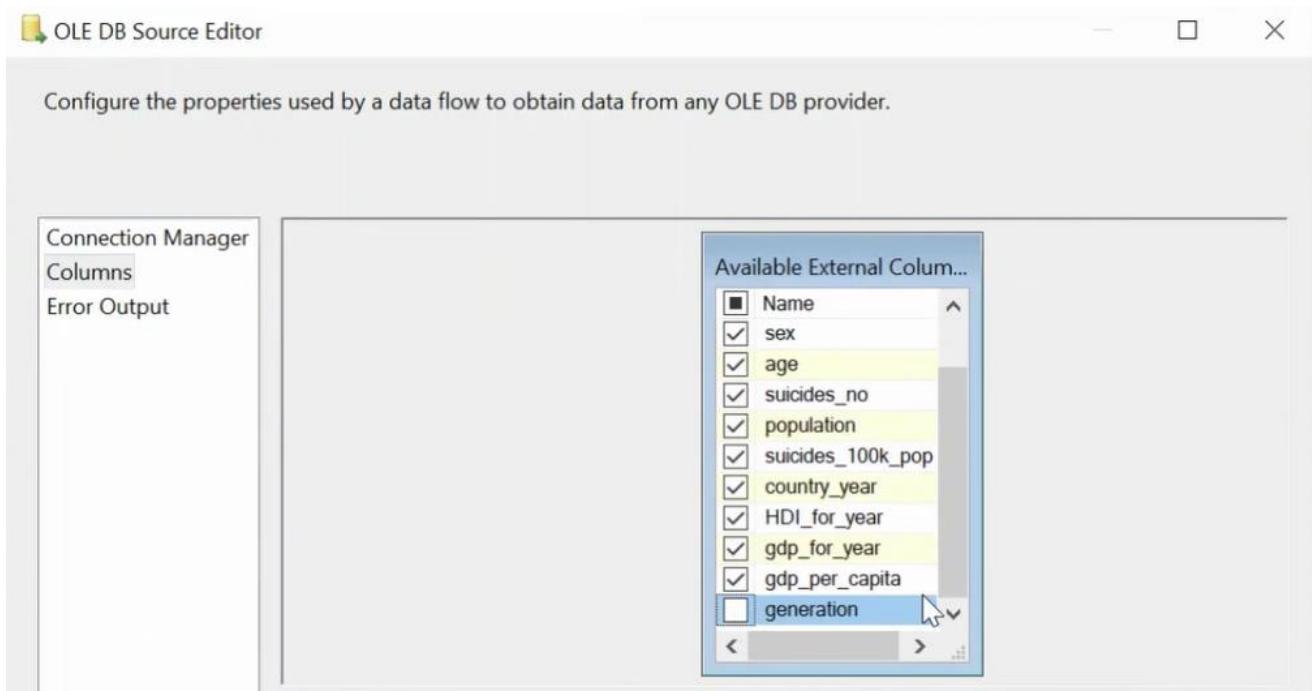


- Kéo và tạo OLE DB Source mới và bấm chuột phải chọn Edit.

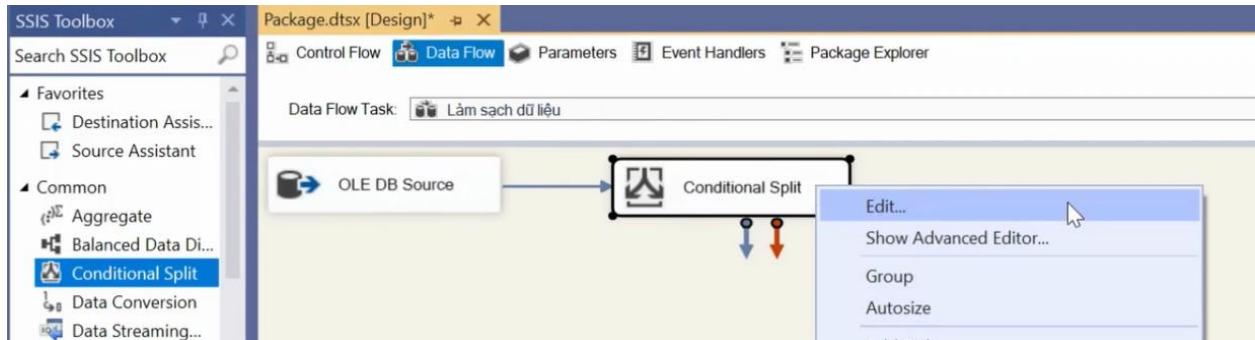




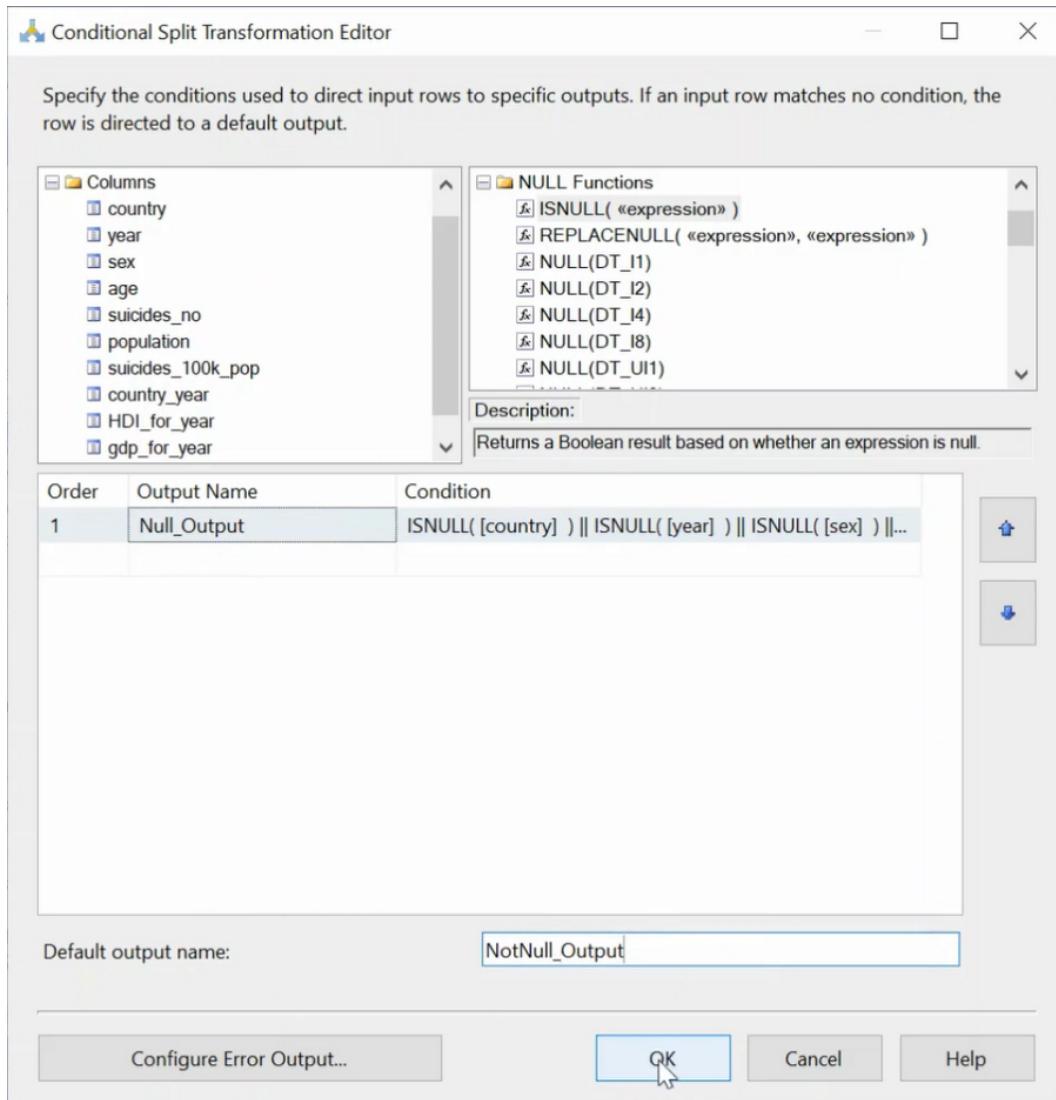
- Tích bỏ chọn các cột không cần sử dụng và nhấn OK.



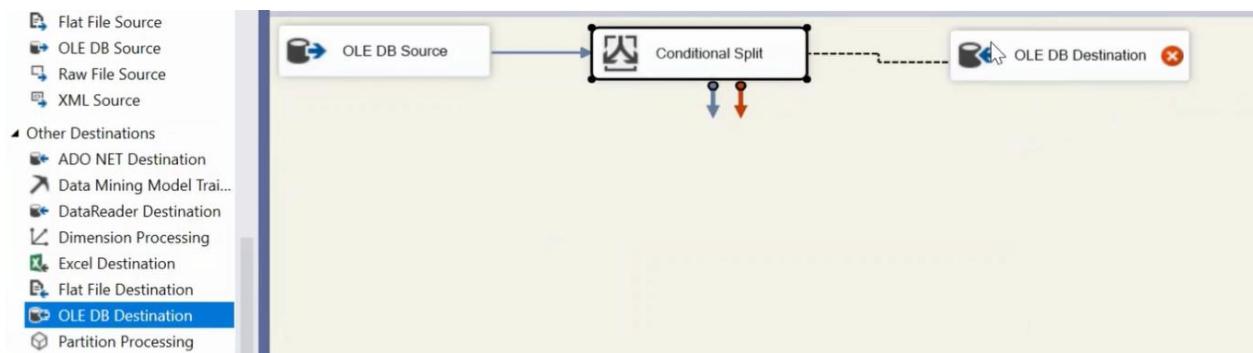
- Kéo thả Condition Split để tạo điều kiện lọc các hàng chứa giá trị Null.
- Nối mũi tên từ OLE DB Source với Condition Split. Nhấn chuột phải và chọn Edit



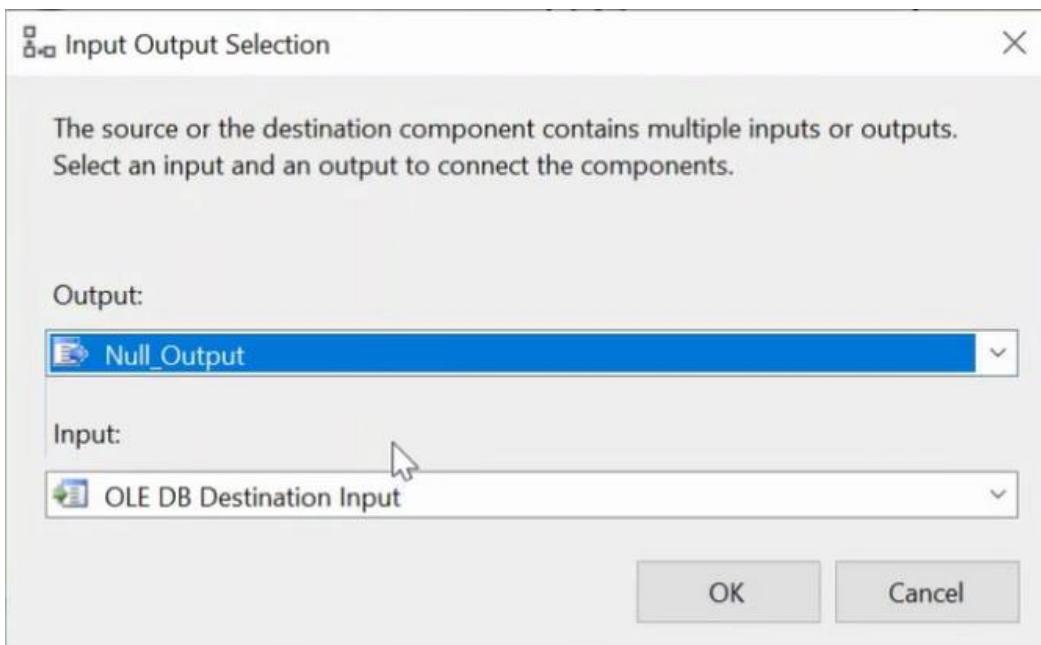
- Sử dụng hàm ISNULL(<<expression>>) để lọc giá trị NULL.
- Đổi tên output name là Null_Output và đặt tên output mặc định là NotNull_Output, nhấn OK.



- Kéo thả OLE DB Destination.
- Nối mũi tên từ OLE DB Source với OLE DB Destination.



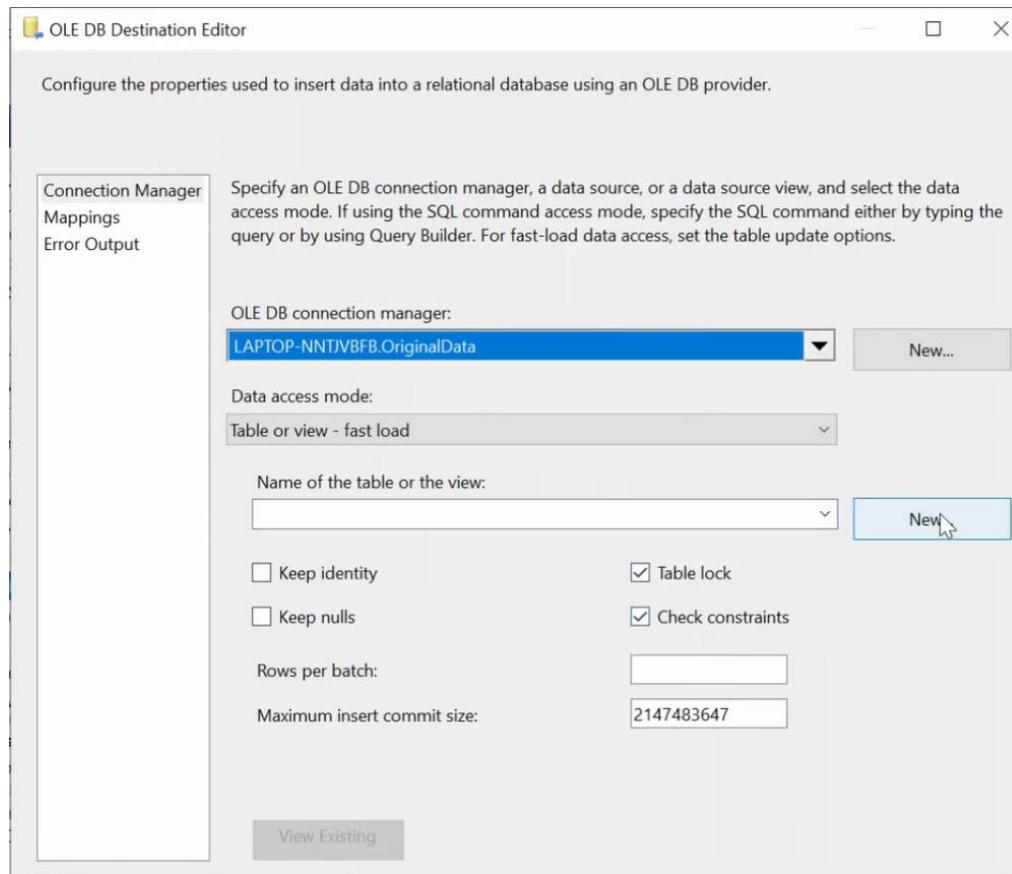
- Hộp thoại Input output selection xuất hiện. Chọn Null_Output và nhấn OK.



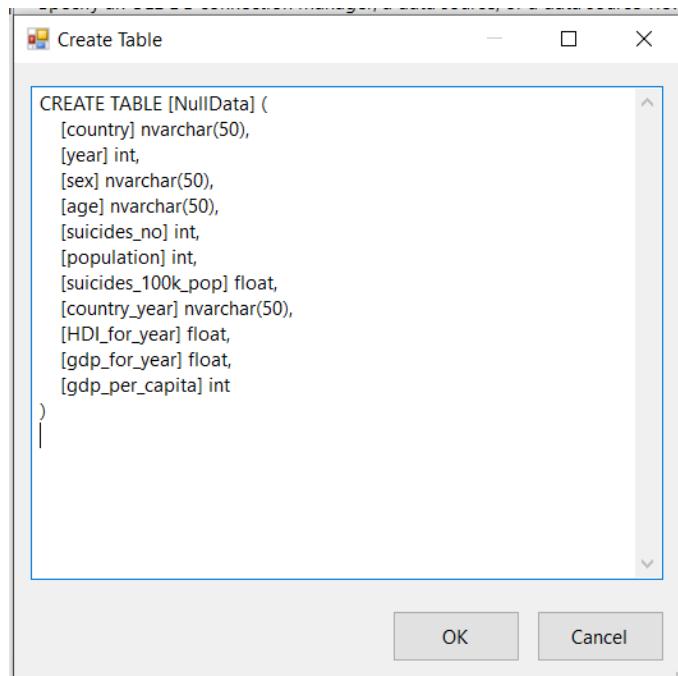
- Nhấn chuột phải vào OLE DB Destination và chọn Edit.



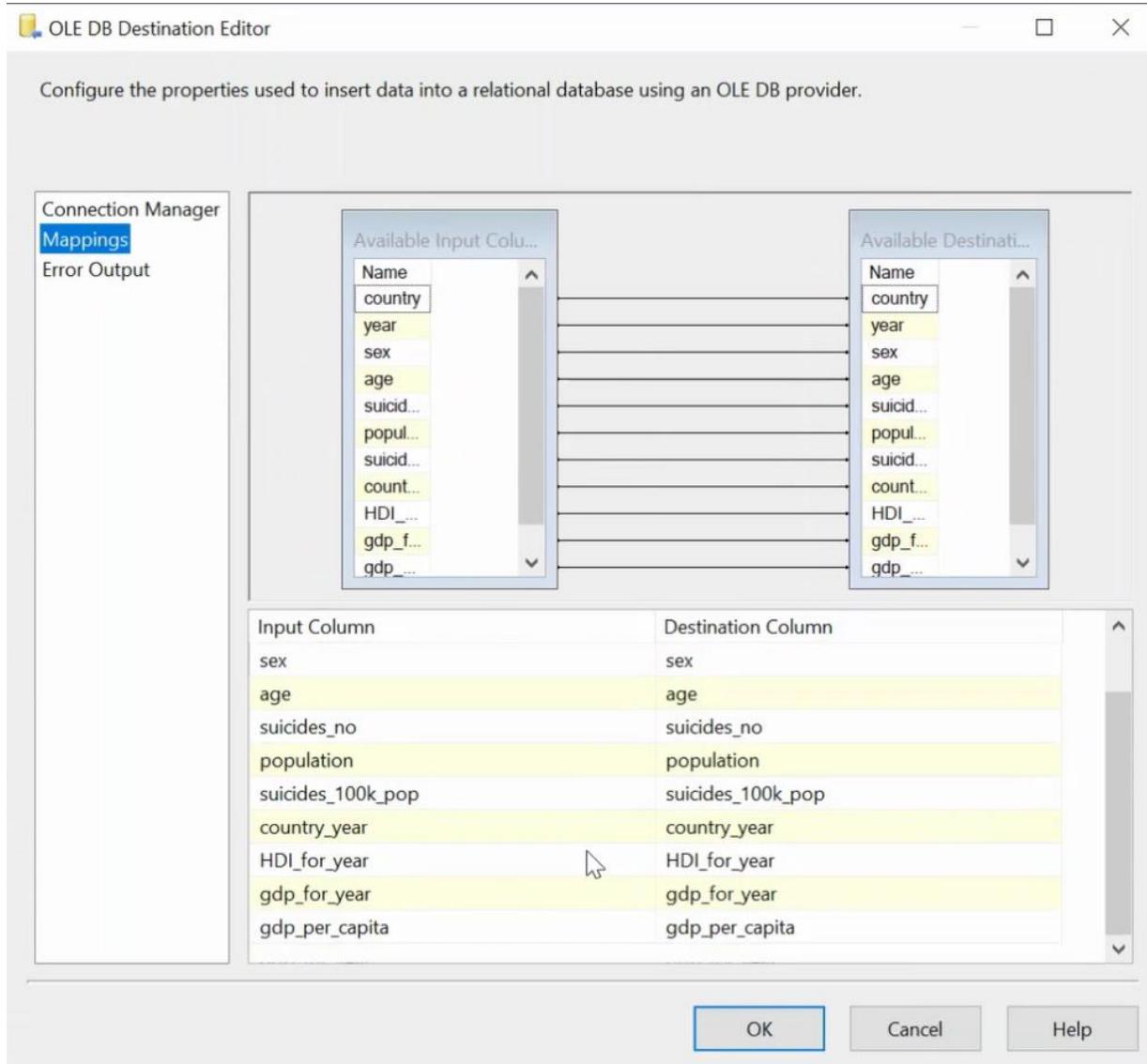
- Chọn kết nối với OriginalData và chọn New để tạo bảng NullData, giữ các dòng dữ liệu có cột chứa giá trị Null.



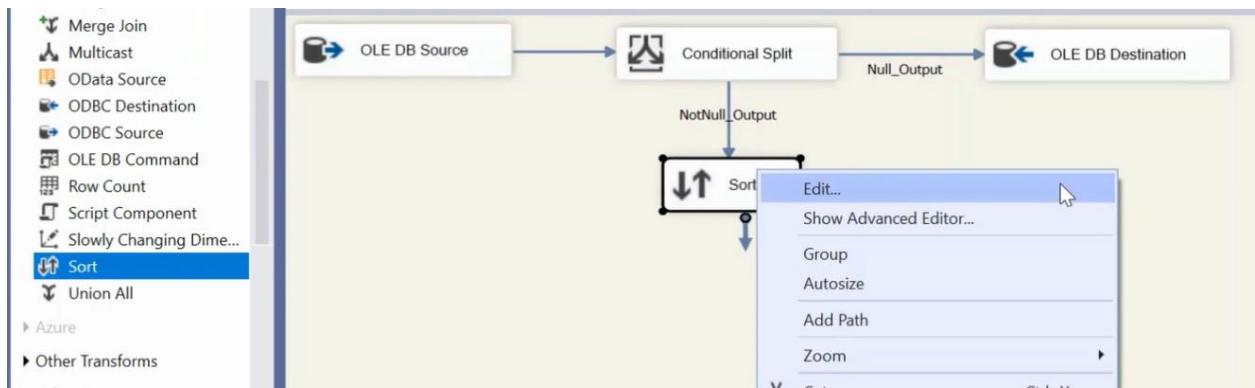
- Đổi tên và chỉnh sửa bảng NullData

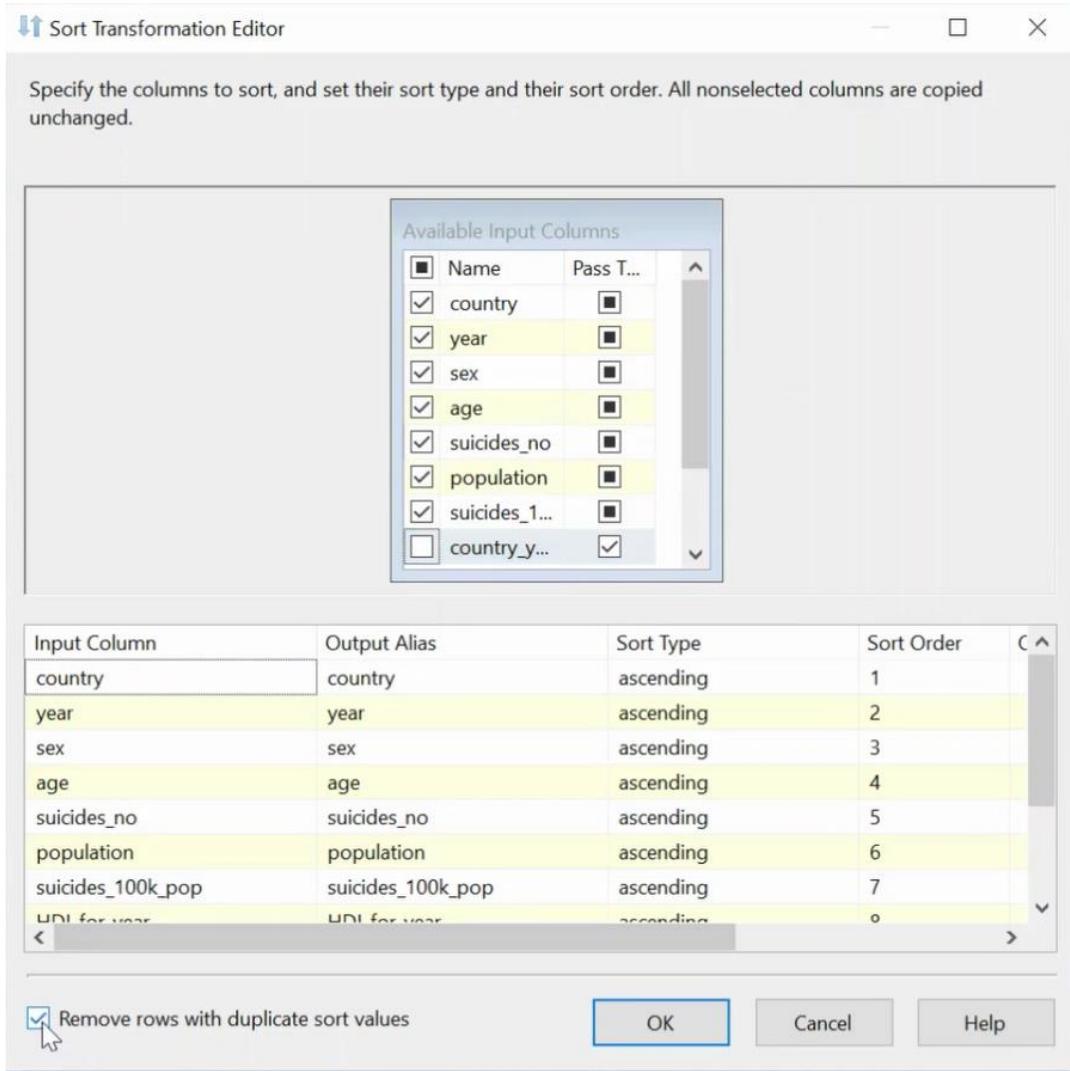


- Kiểm tra mapping các thuộc tính và nhấn OK.

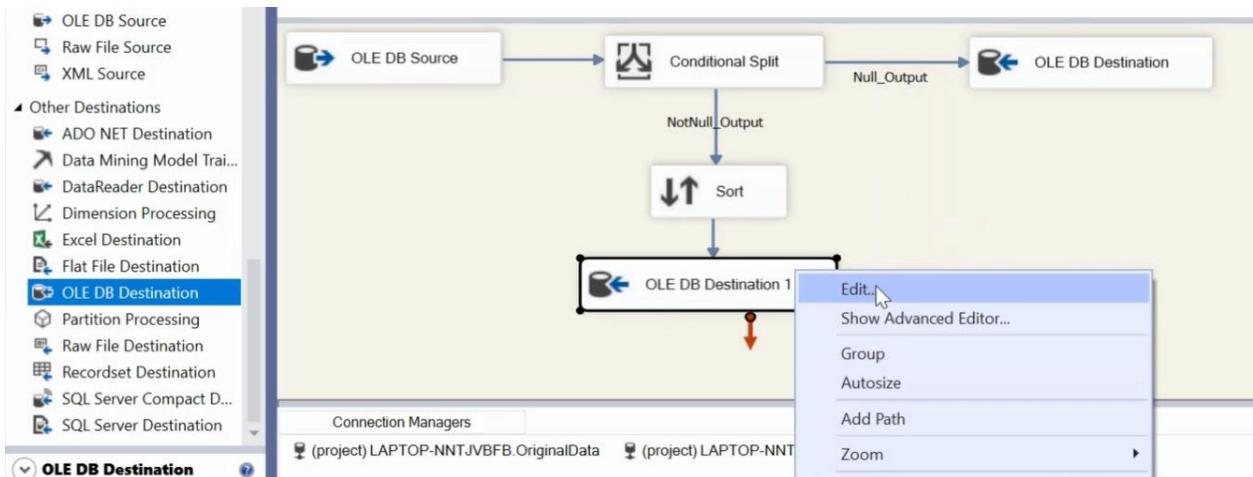


- Kéo thả Sort và nối tạo liên kết với Condition Split
- Chuột phải vào Sort và chọn Edit.

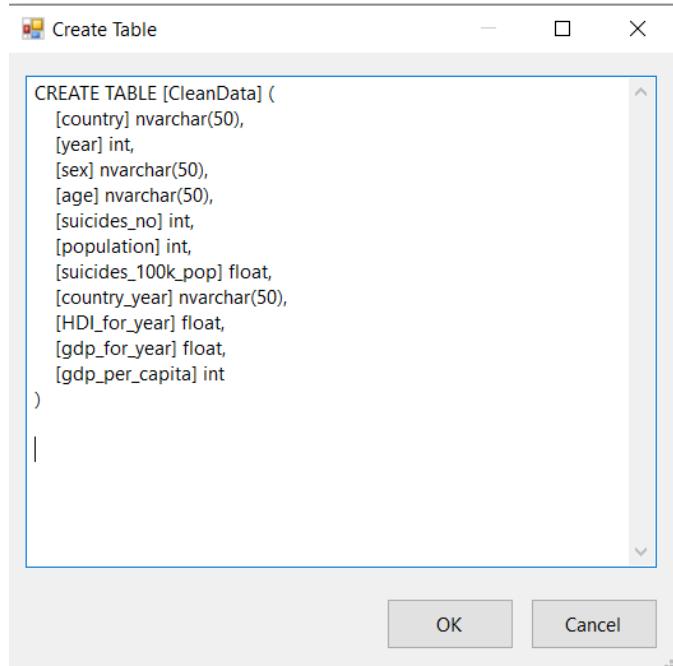
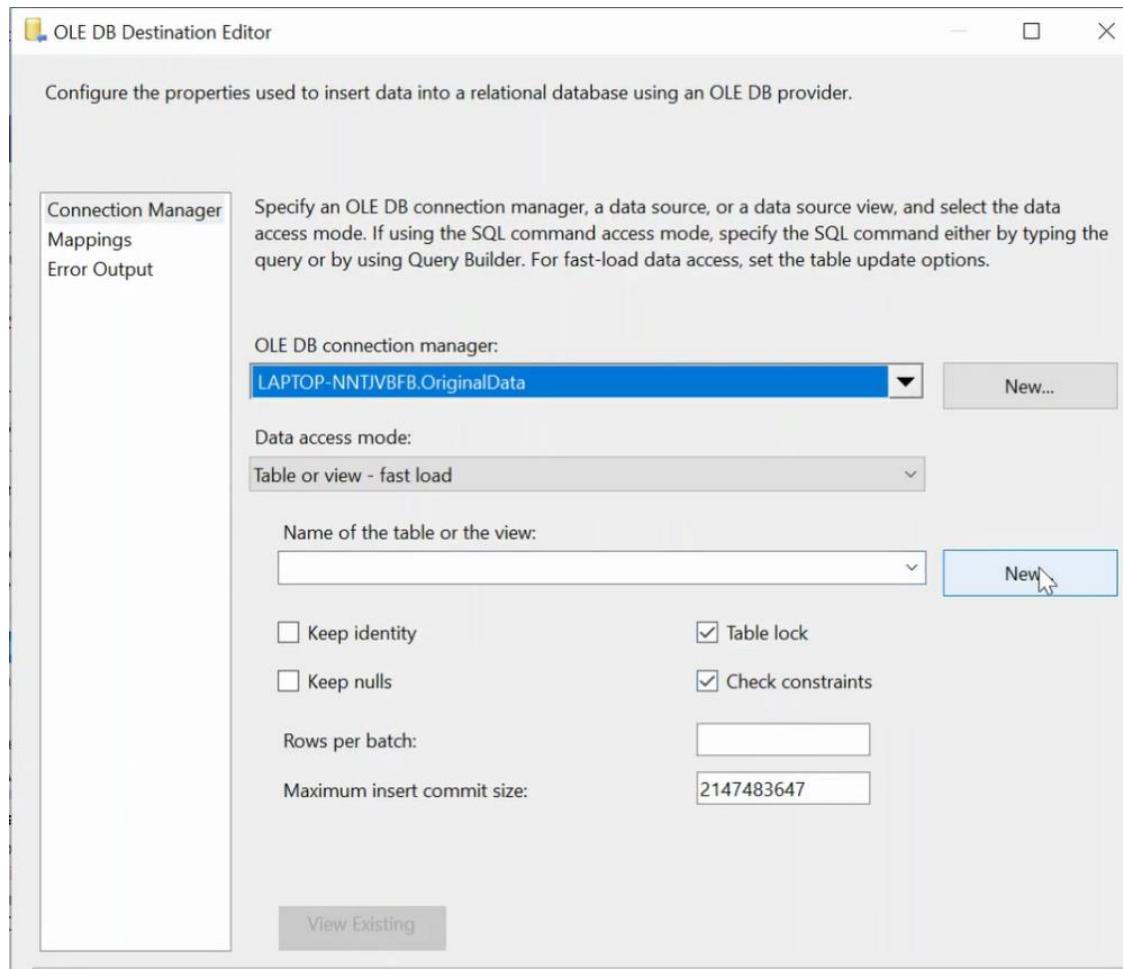


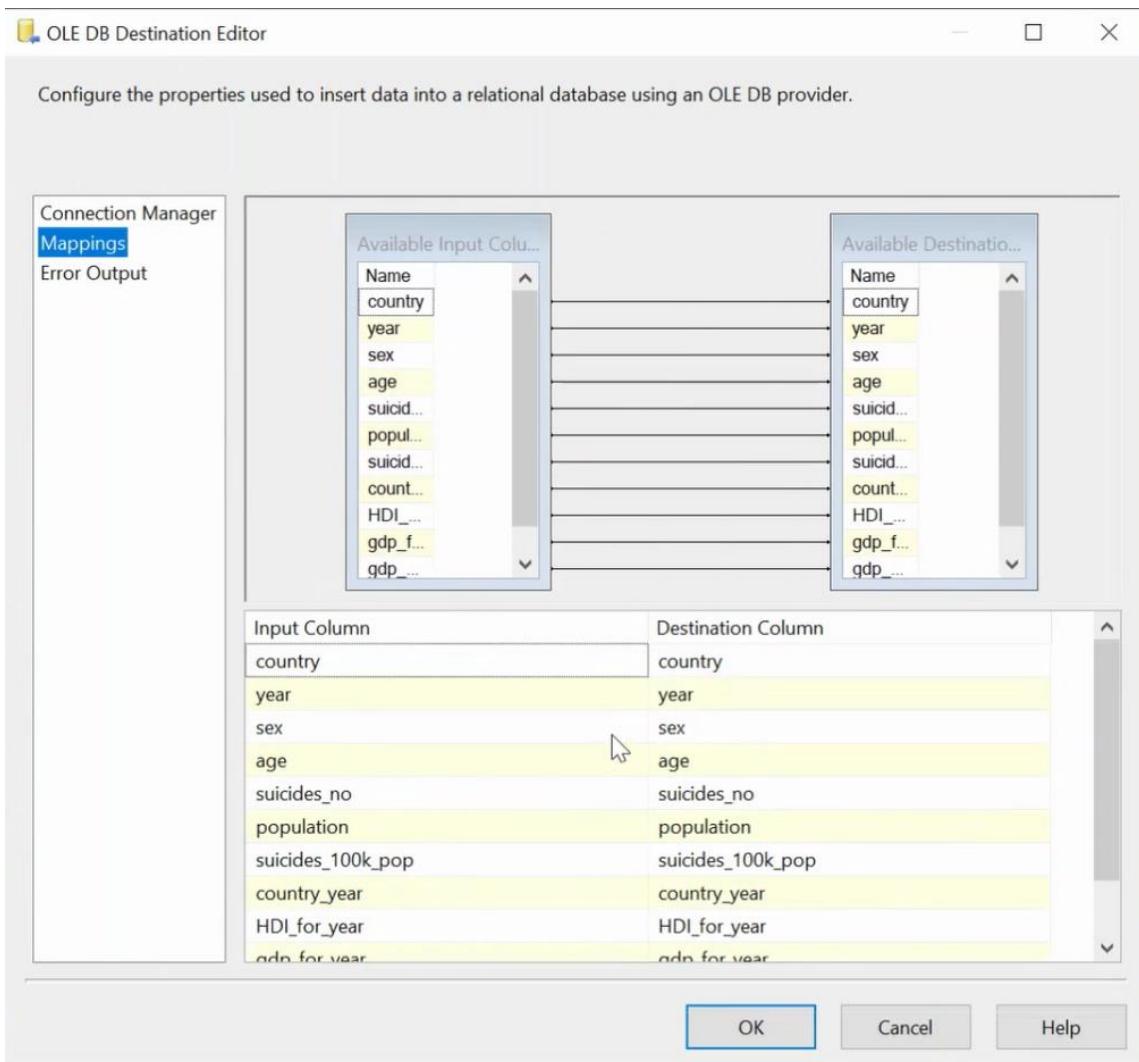


- Kéo thả tạo OLE DB Destination.
- Nối tạo liên kết Sort với OLE DB Destination1, nhấn chuột phải chọn Edit

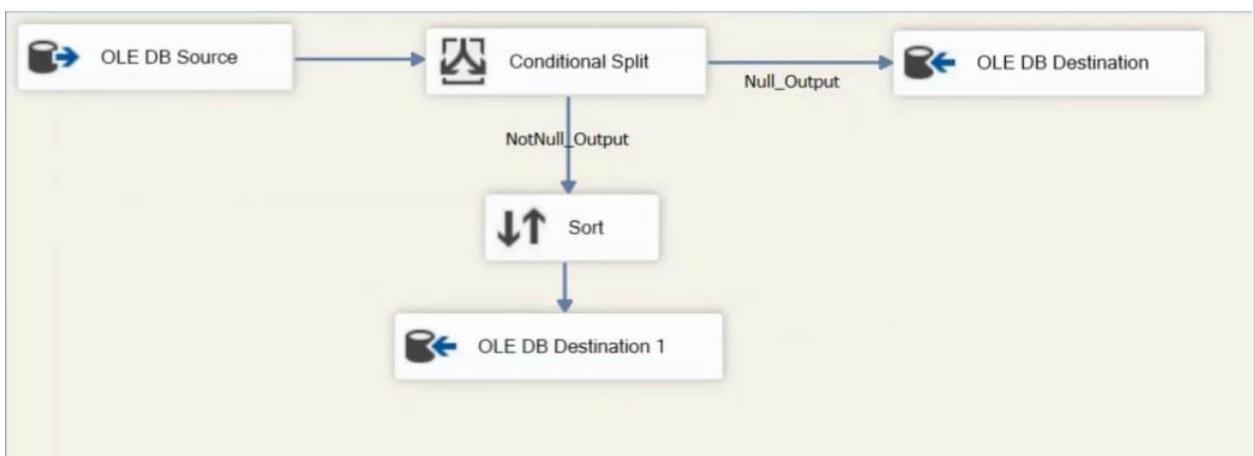


- Tạo bảng CleanData tương tự như tạo bảng NullData để chứa dữ liệu đã được làm sạch.



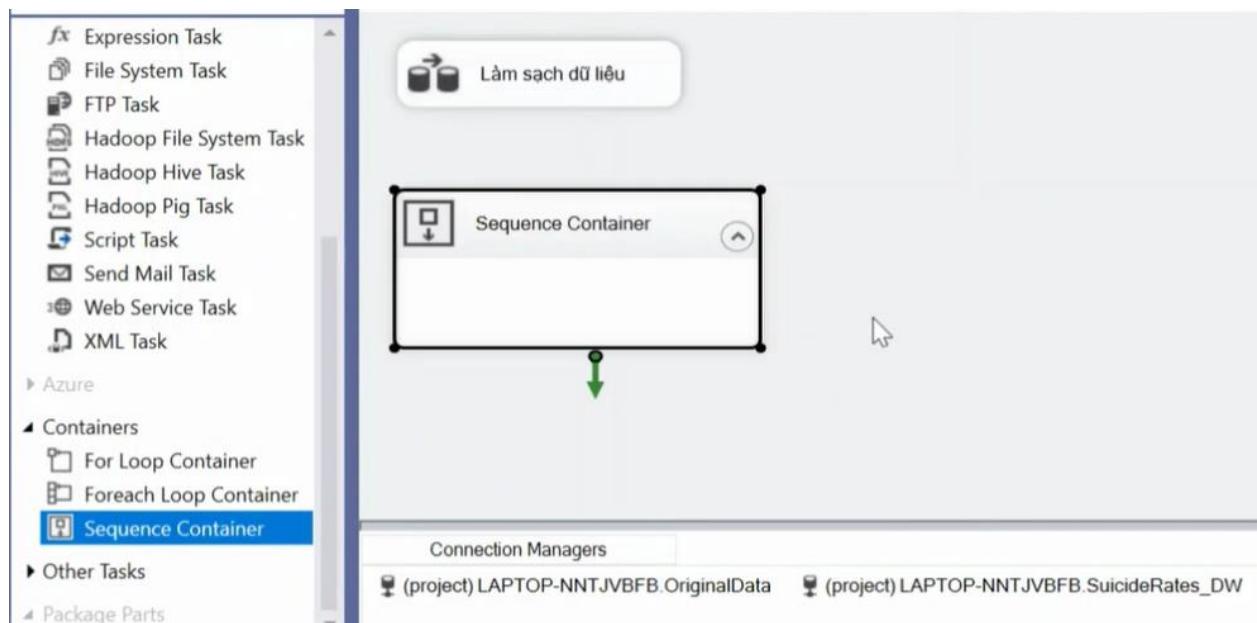


- Hoàn thành DataFlow làm sạch dữ liệu

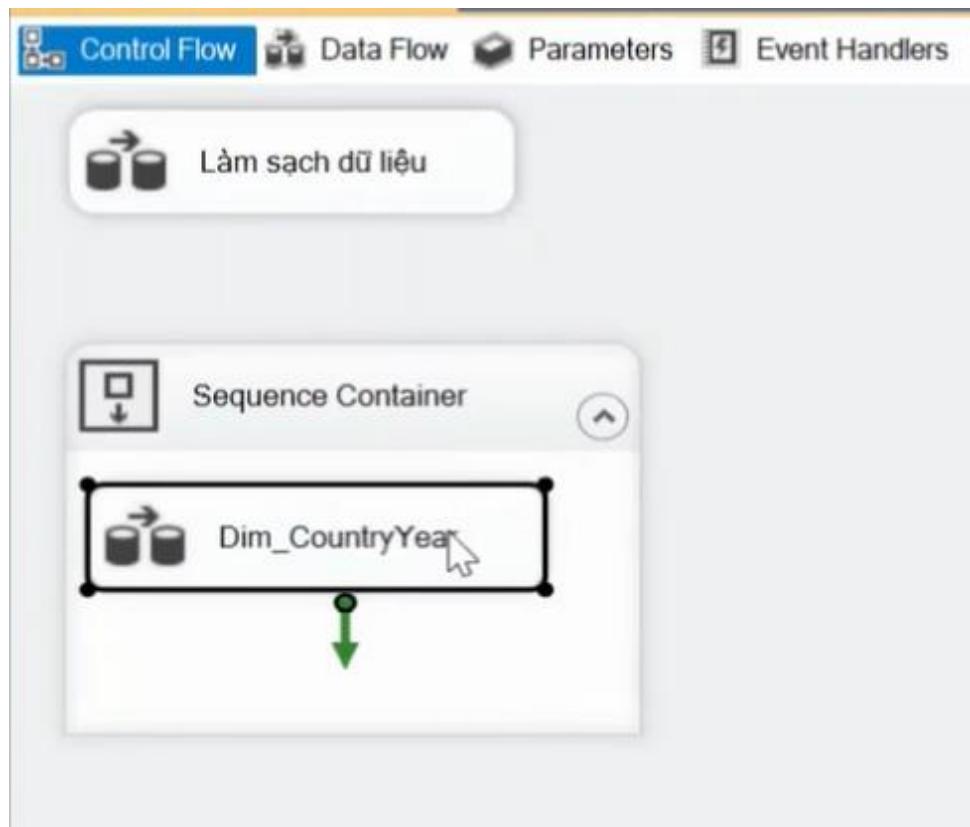


4.4. Tạo các bảng Dimension

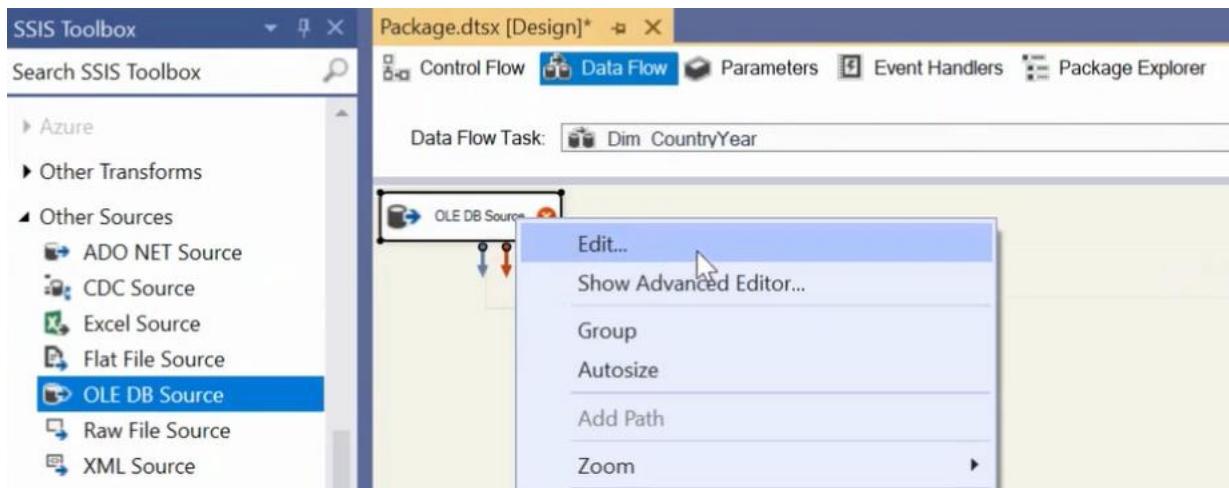
- Kéo thả Sequence Container để chứa Control Flow tạo bảng Dim_CountryYear và Dim_SexAge và thực thi cùng lúc.



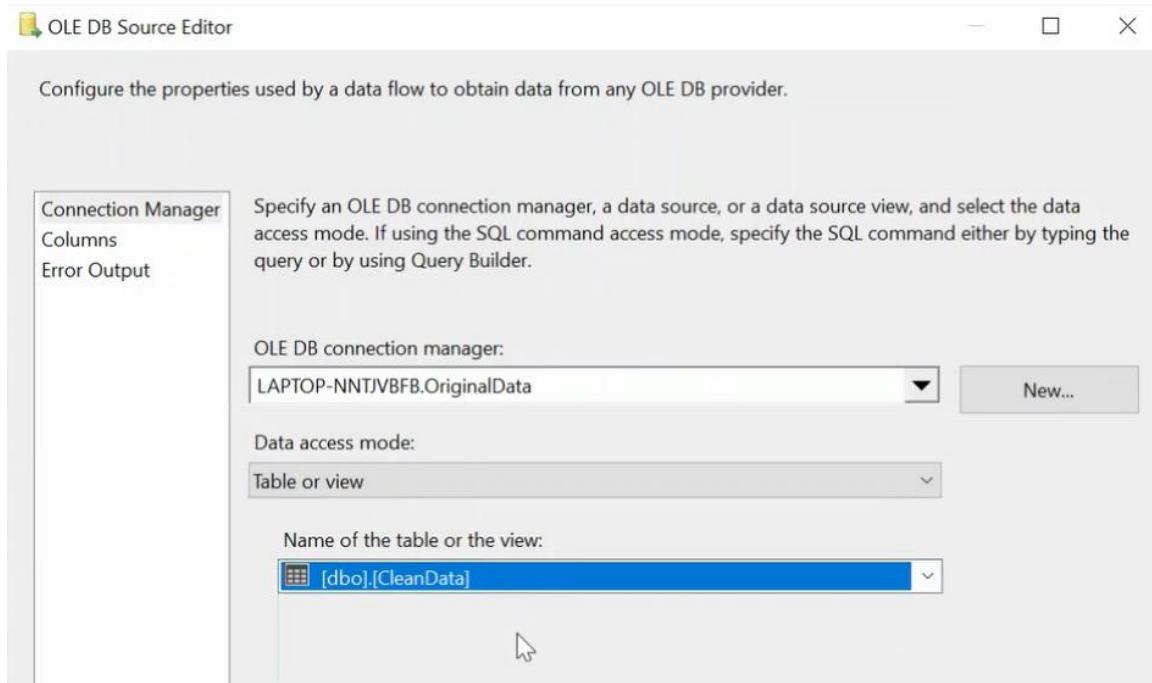
- Tạo bảng Dim_CountryYear
- Kéo thả Data Flow Task vào Sequence Container để tạo Dim_CountryYear

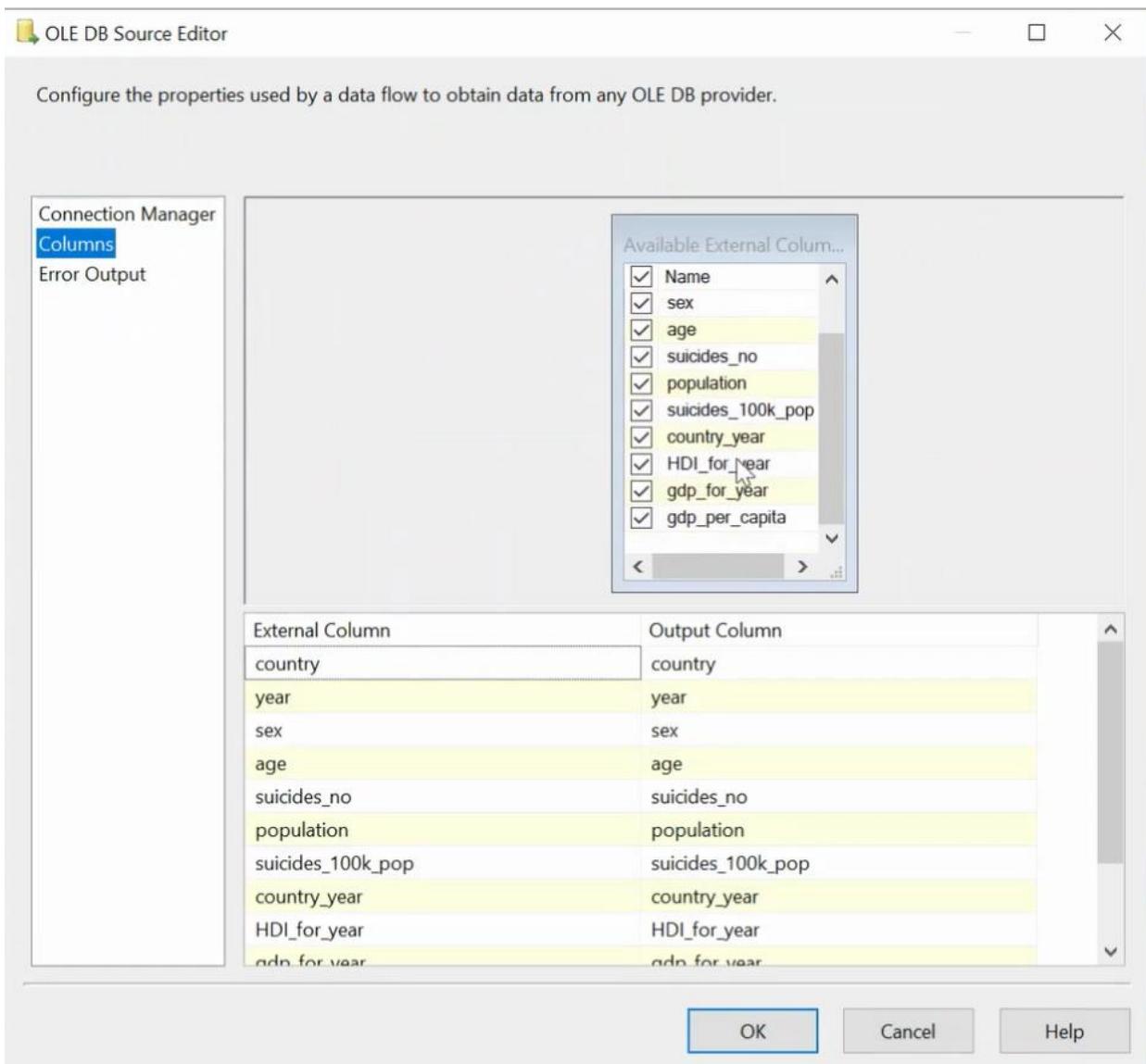


- Kéo thả OLE DB Source và chuột phải chọn Edit.

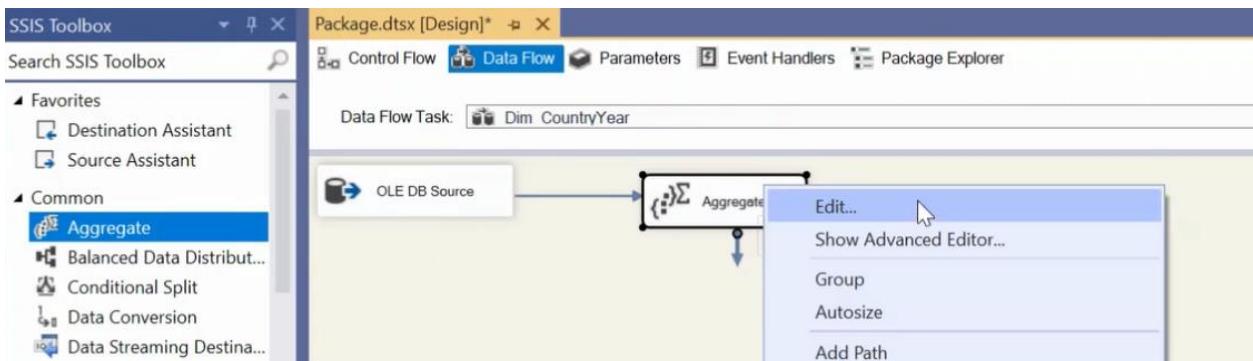


- Chọn kết nối với OriginalData và nguồn dữ liệu là bảng CleanData, nhấn OK.

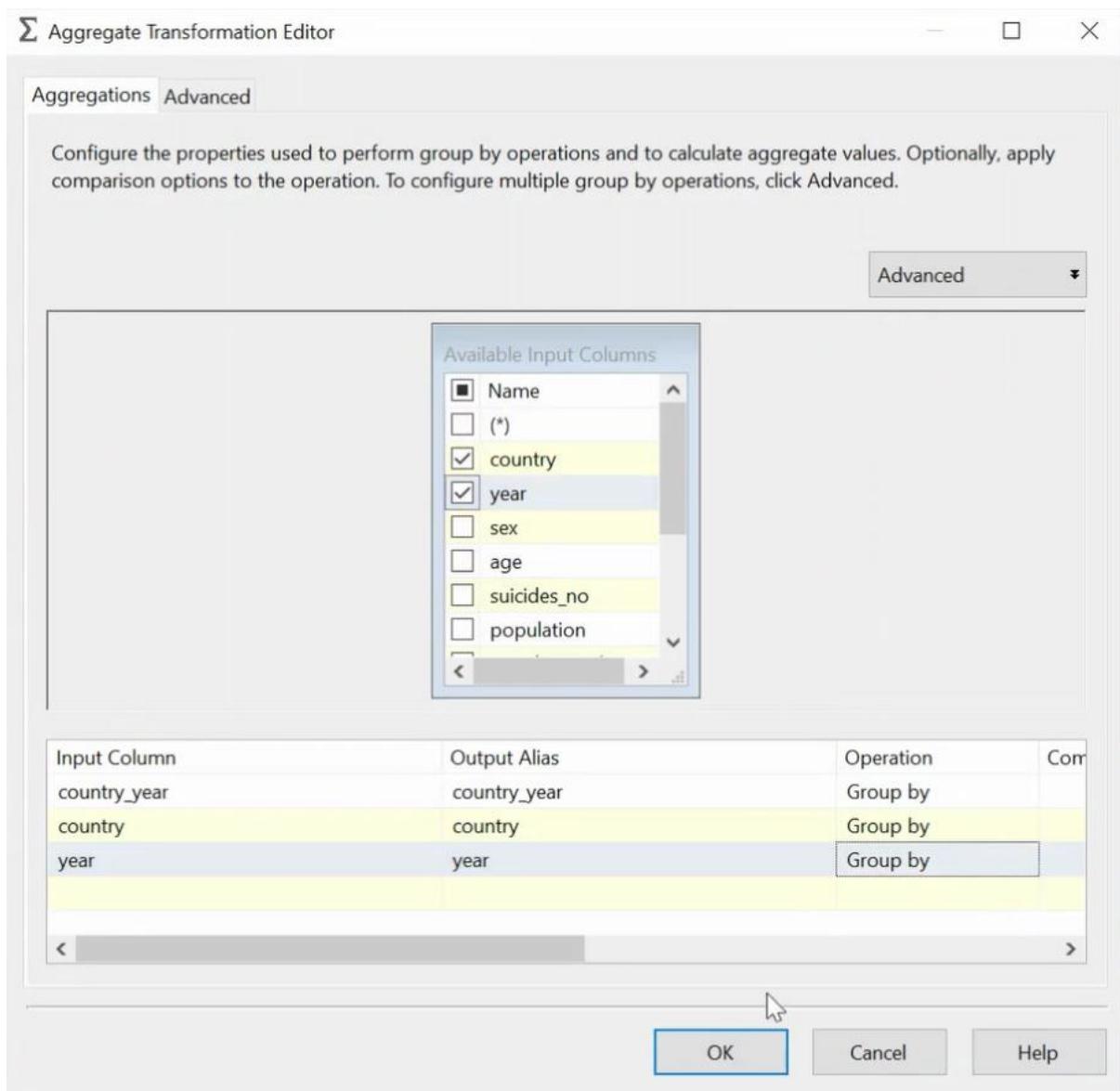




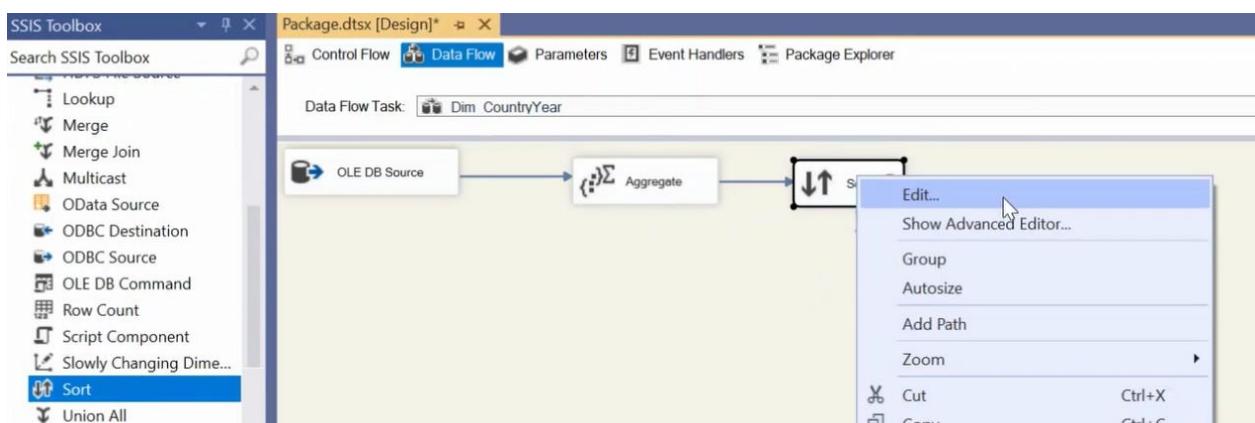
- Kéo thả Aggregate để chọn các thuộc tính cần tạo bảng Dim_CountryYear.
- Nối OLE DB Source với Aggregate, chuột phải chọn Edit.



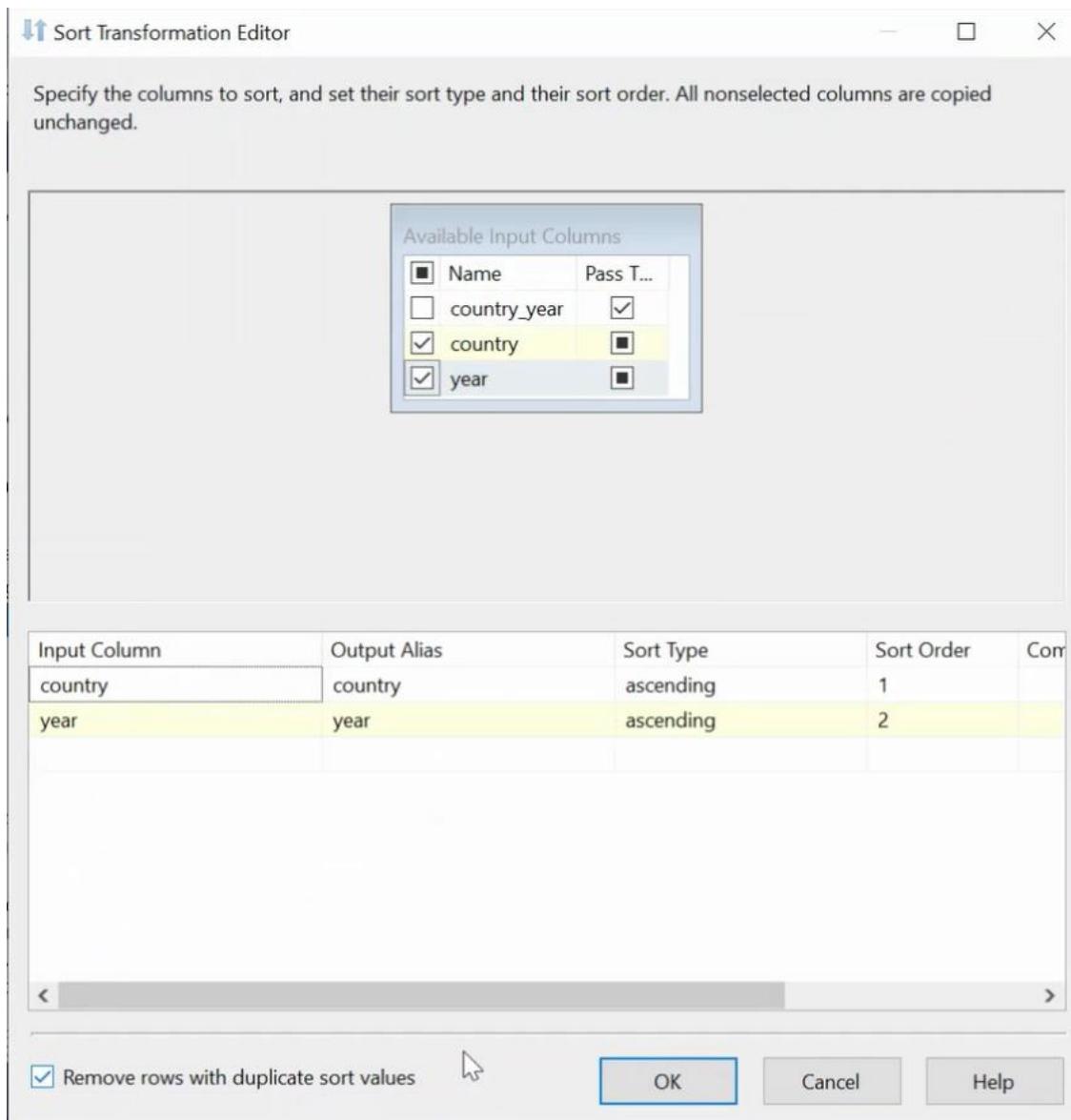
- Chọn các thuộc tính Year, Country, CountryYear



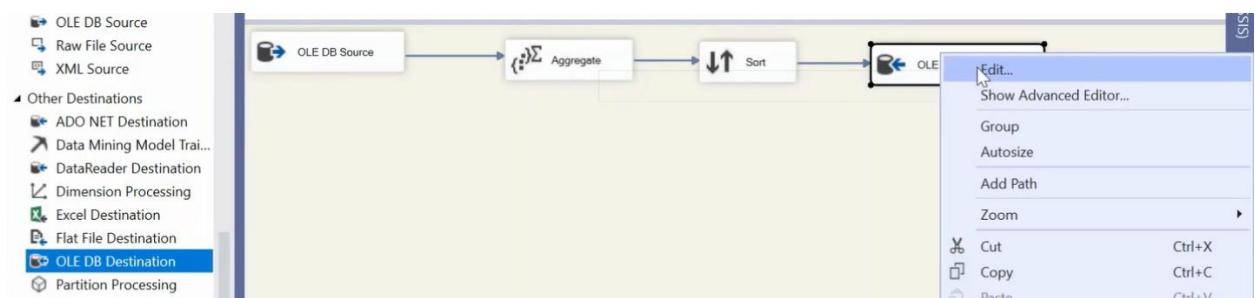
- Kéo thả Sort và tạo liên kết Aggregate đến Sort, chuột phải chọn Edit.



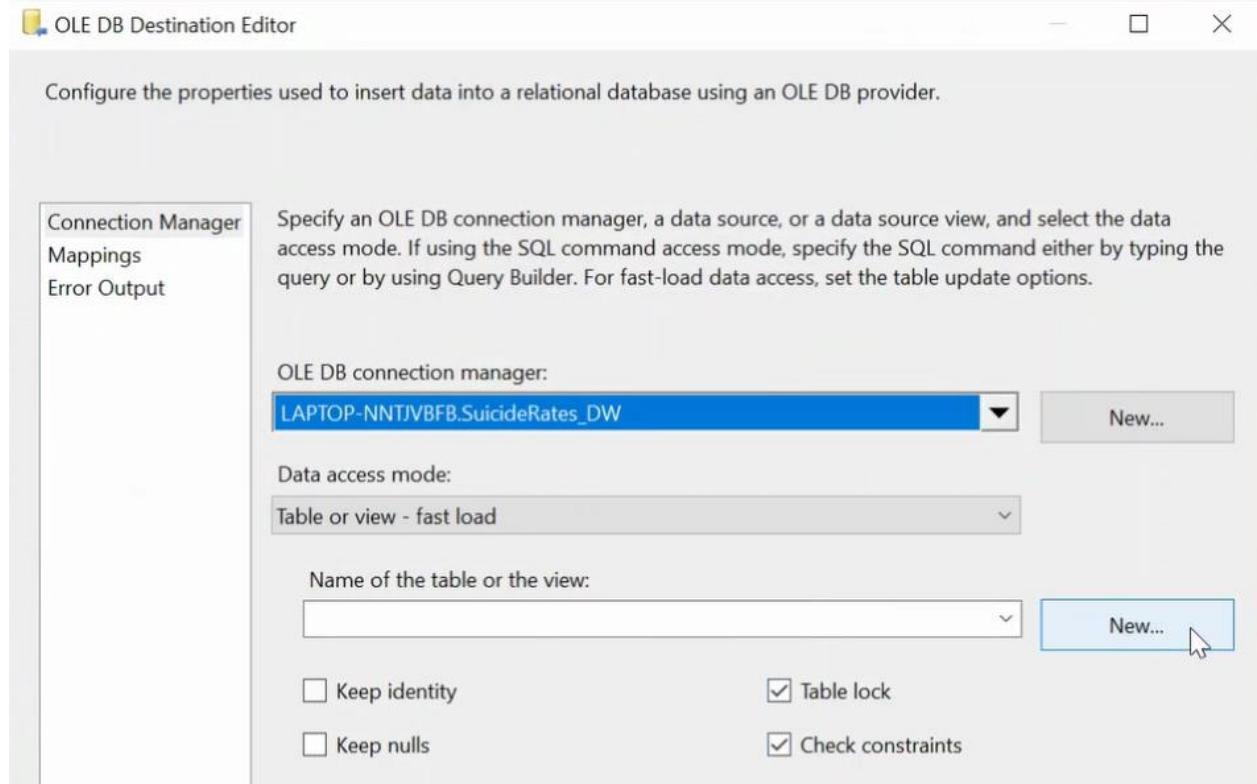
- Chọn thuộc tính cài Sort



- Kéo thả OLE DB Destination và tạo liên kết đến Sort, nhấn chuột phải chọn Edit.



- Chọn kết nối đến SuicideRates_DW và chọn New để tạo bảng Dim_CountryYear

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

Connection Manager
Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

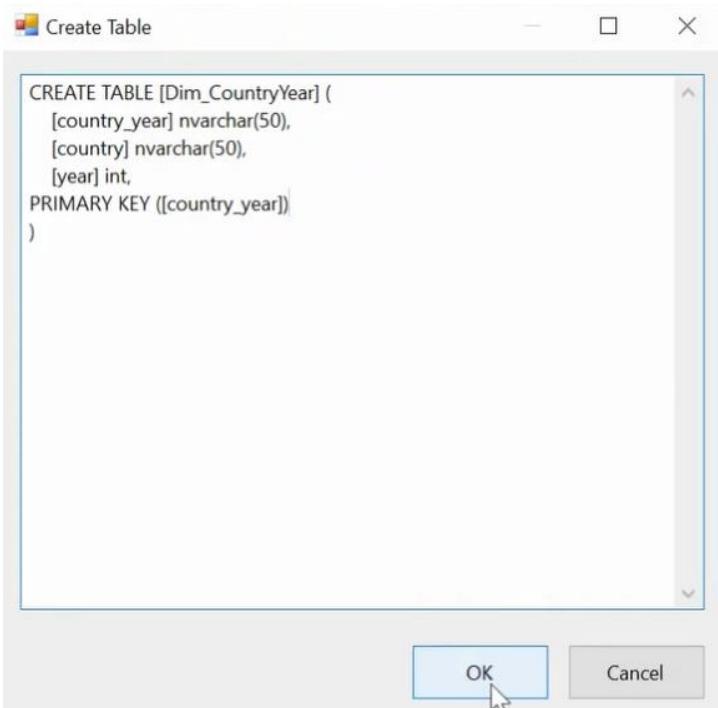
OLE DB connection manager: LAPTOP-NNTJVBF.B.SuicideRates_DW

Data access mode: Table or view - fast load

Name of the table or the view: New...

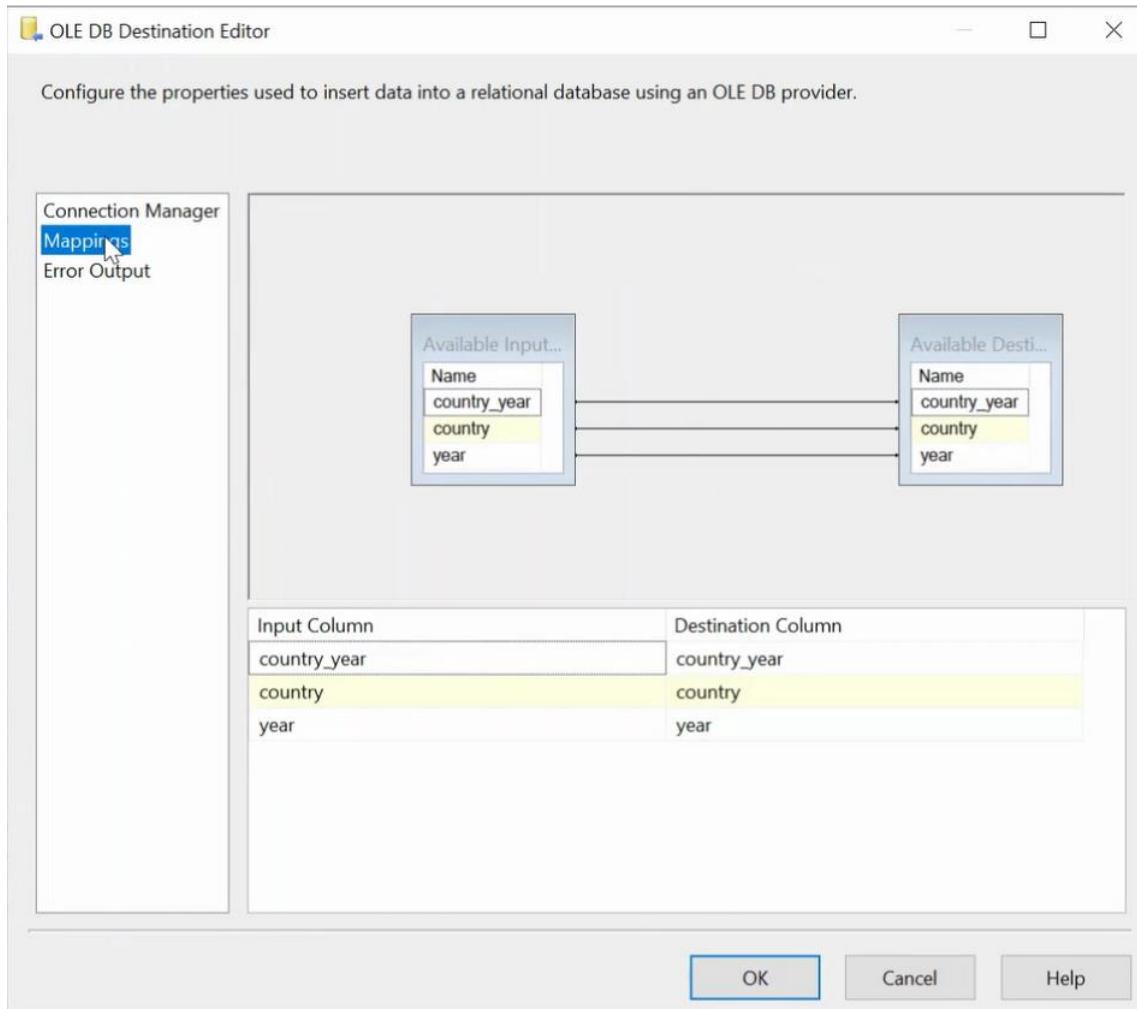
Keep identity Table lock

Keep nulls Check constraints

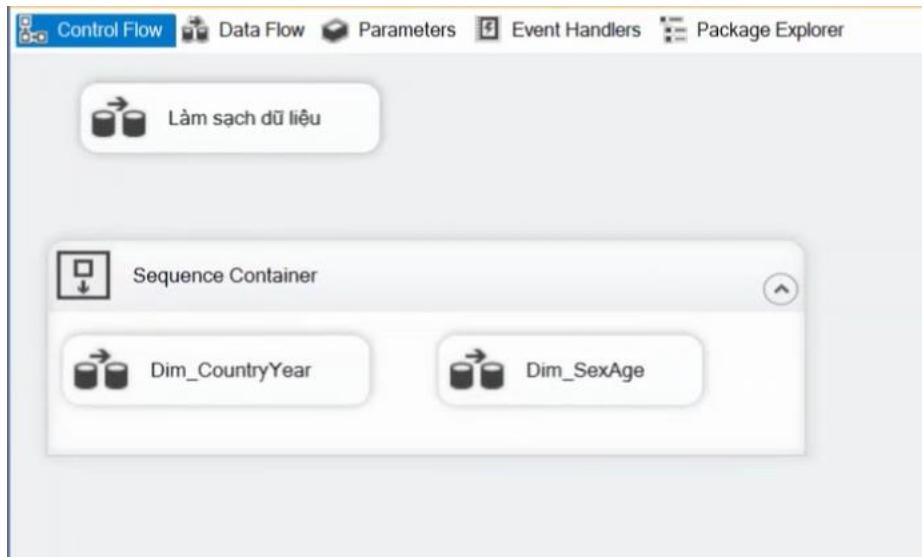
CREATE TABLE [Dim_CountryYear] (
 [country_year] nvarchar(50),
 [country] nvarchar(50),
 [year] int,
 PRIMARY KEY ([country_year])
)

OK Cancel

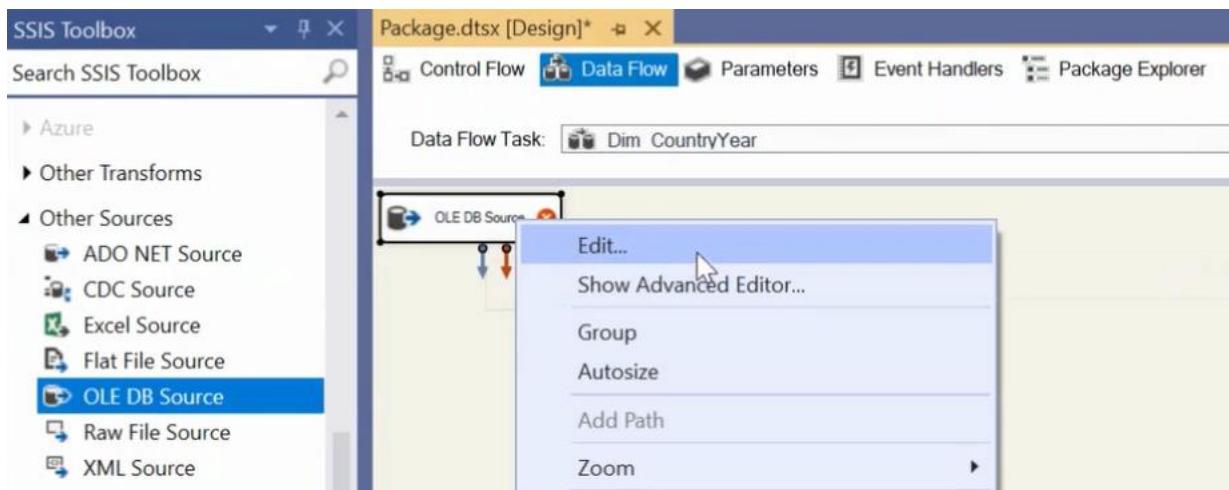
- Kiểm tra mapping các thuộc tính và nhấn OK



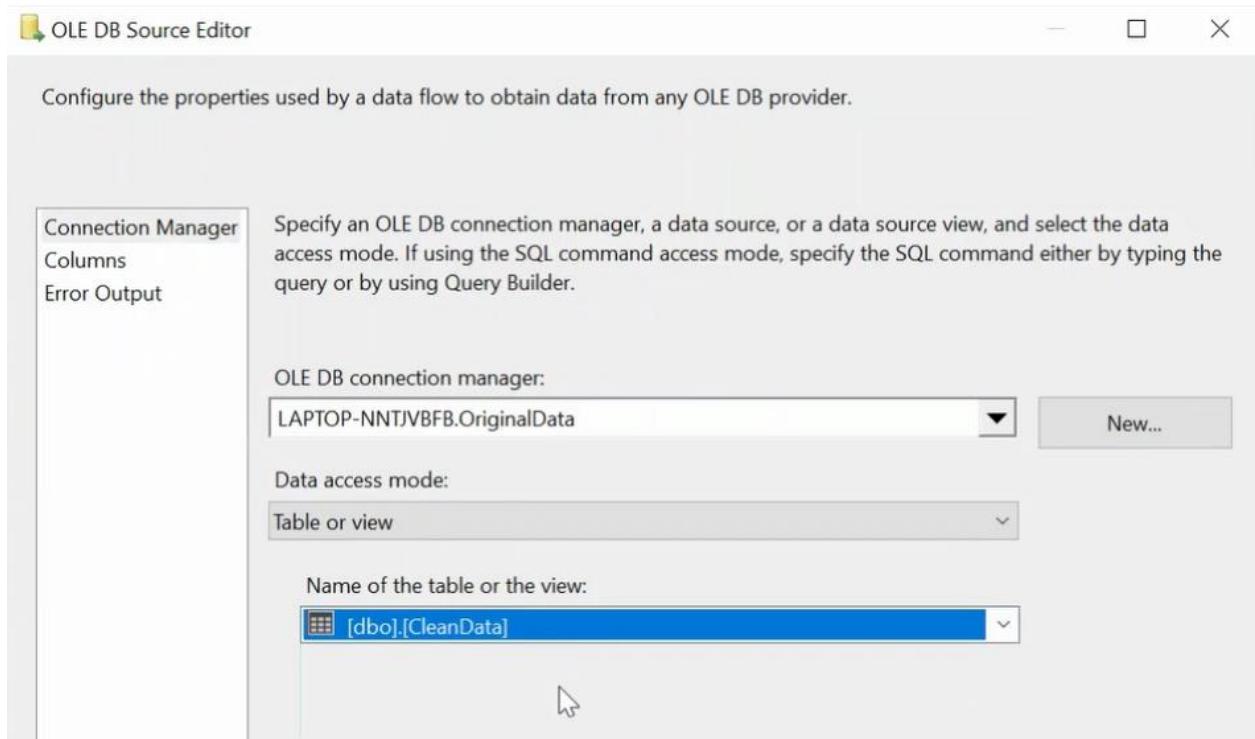
- b. Tạo bảng Dim_SexAge
- Kéo thả Data Flow Task vào Sequence Container để tạo Dim_SexAge

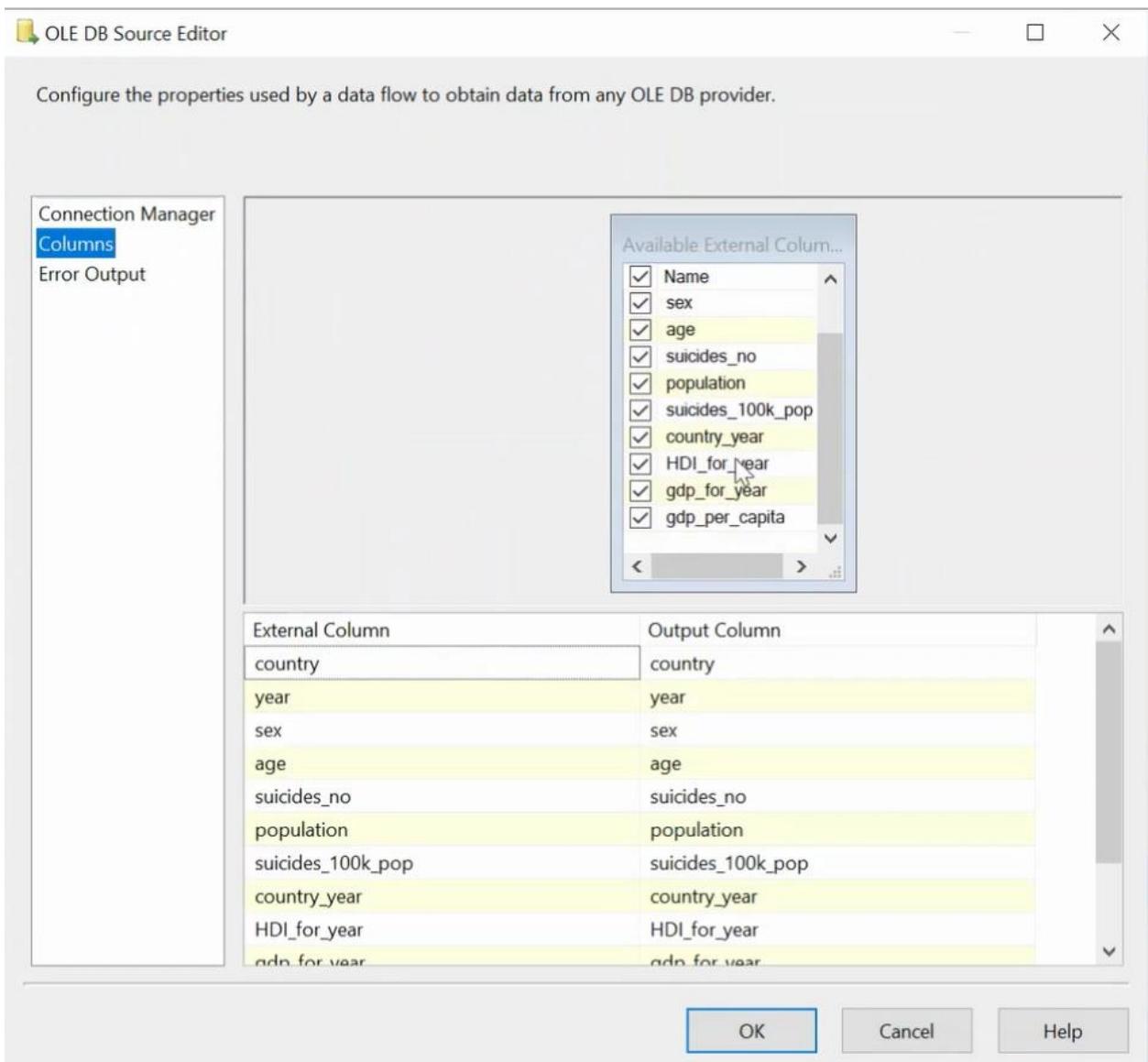


- Kéo thả OLE DB Source và chuột phải chọn Edit.

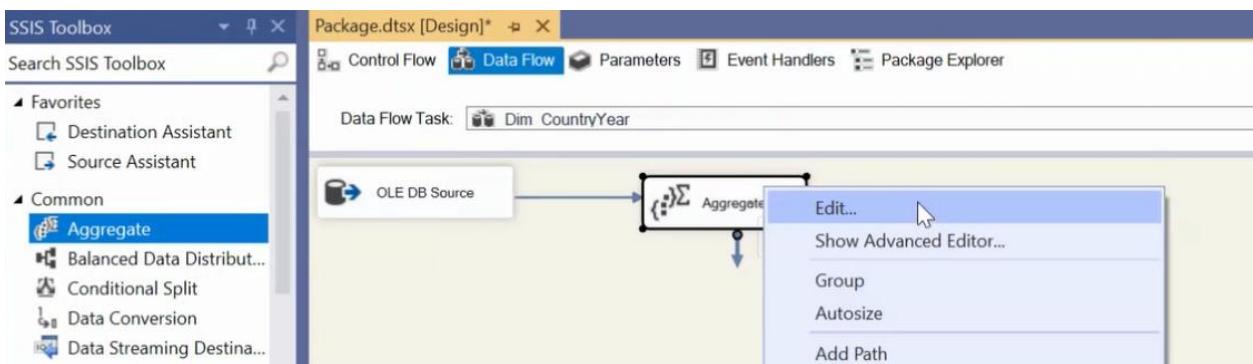


- Chọn kết nối với OriginalData và nguồn dữ liệu là bảng CleanData, nhấn OK.

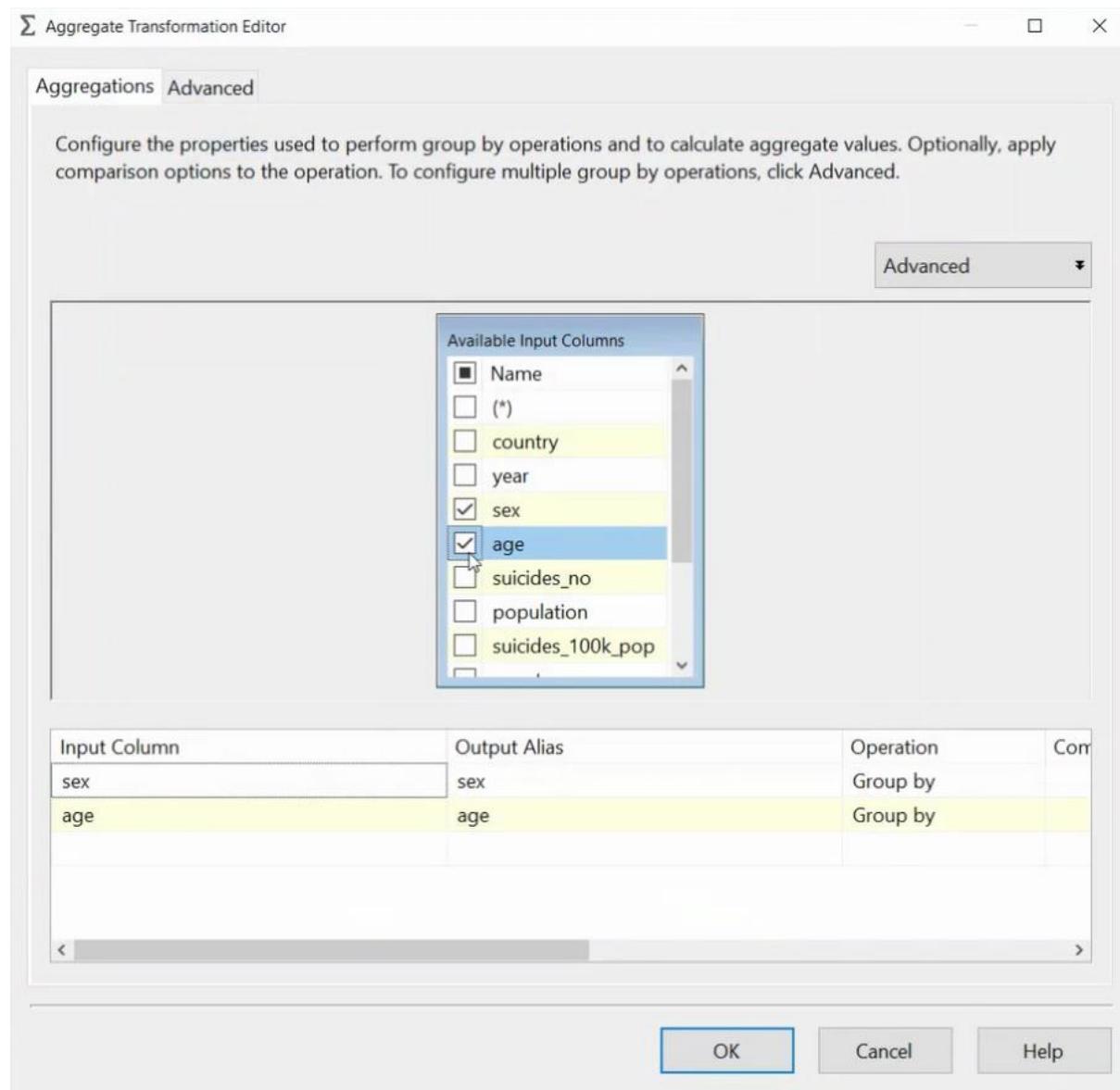




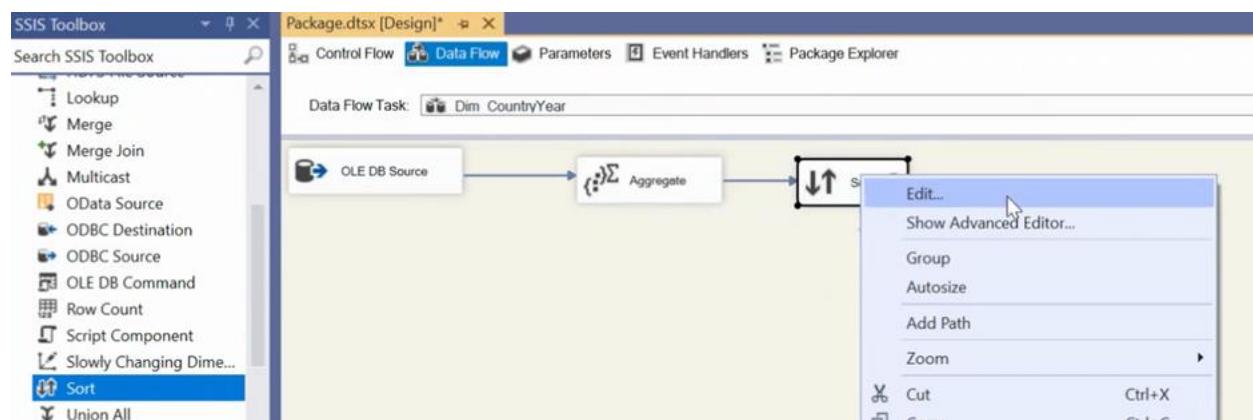
- Kéo thả Aggregate để chọn các thuộc tính cần tạo bảng Dim_SexAge.
- Nối OLE DB Source với Aggregate, chuột phải chọn Edit.



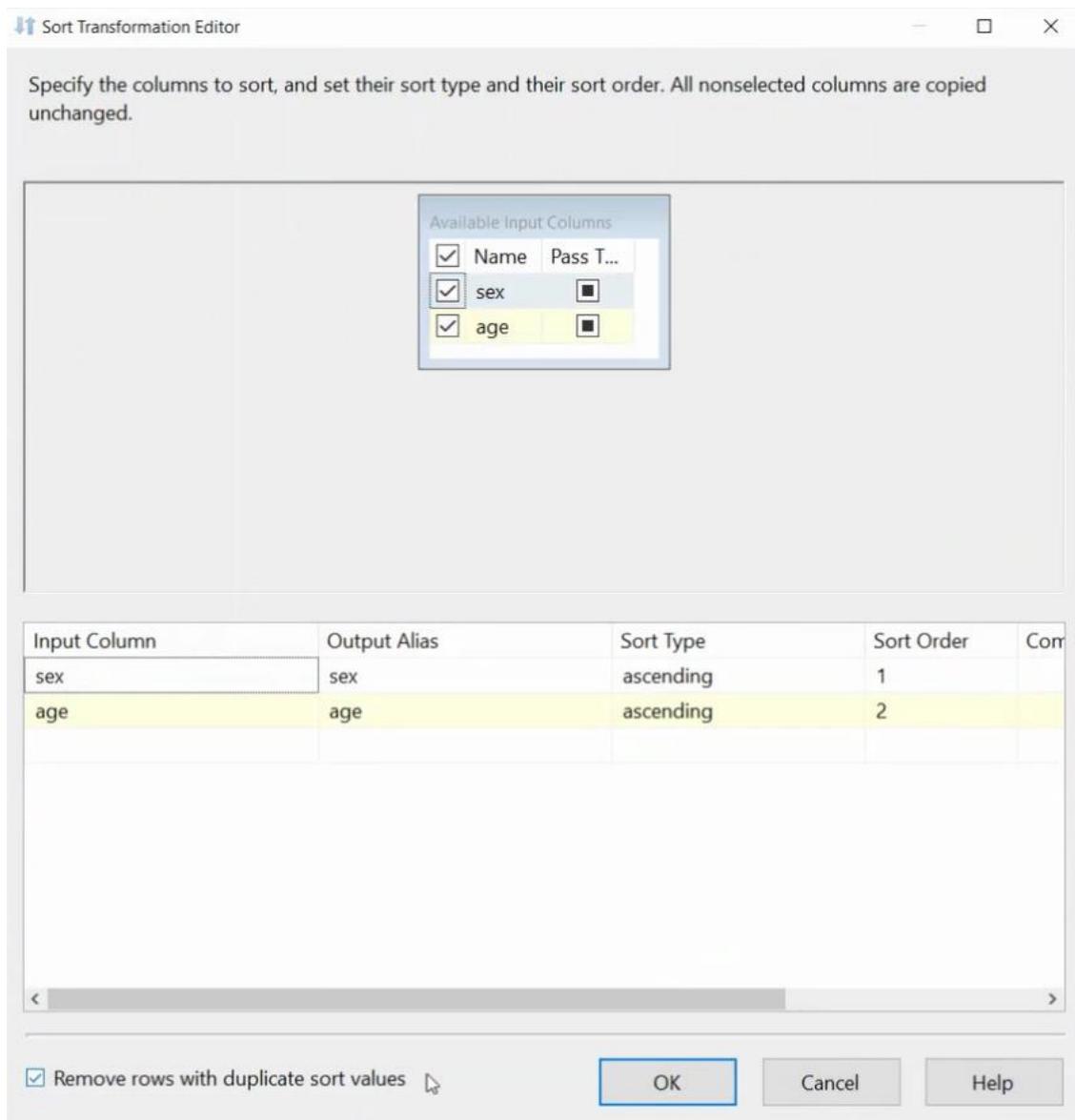
Chọn các thuộc tính Sex, Age



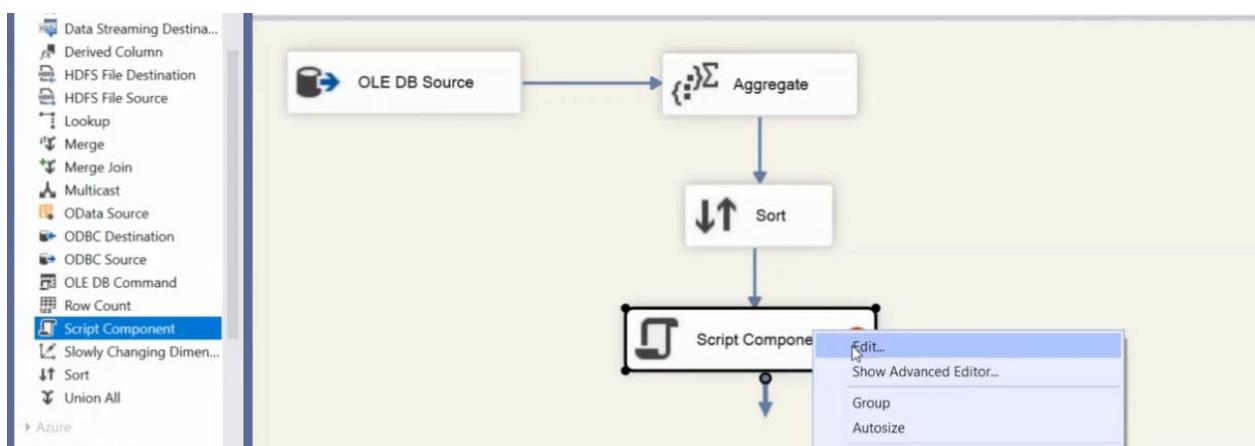
- Kéo thả Sort và tạo liên kết Aggregate đến Sort, chuột phải chọn Edit.



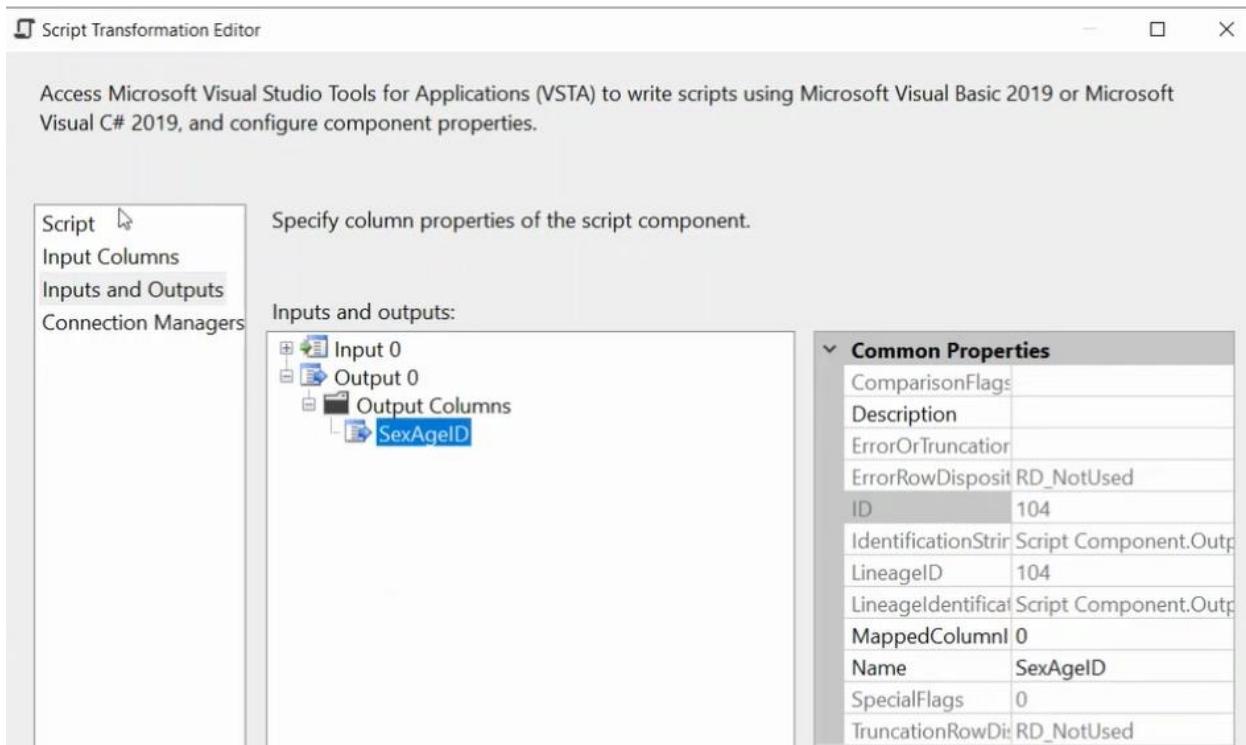
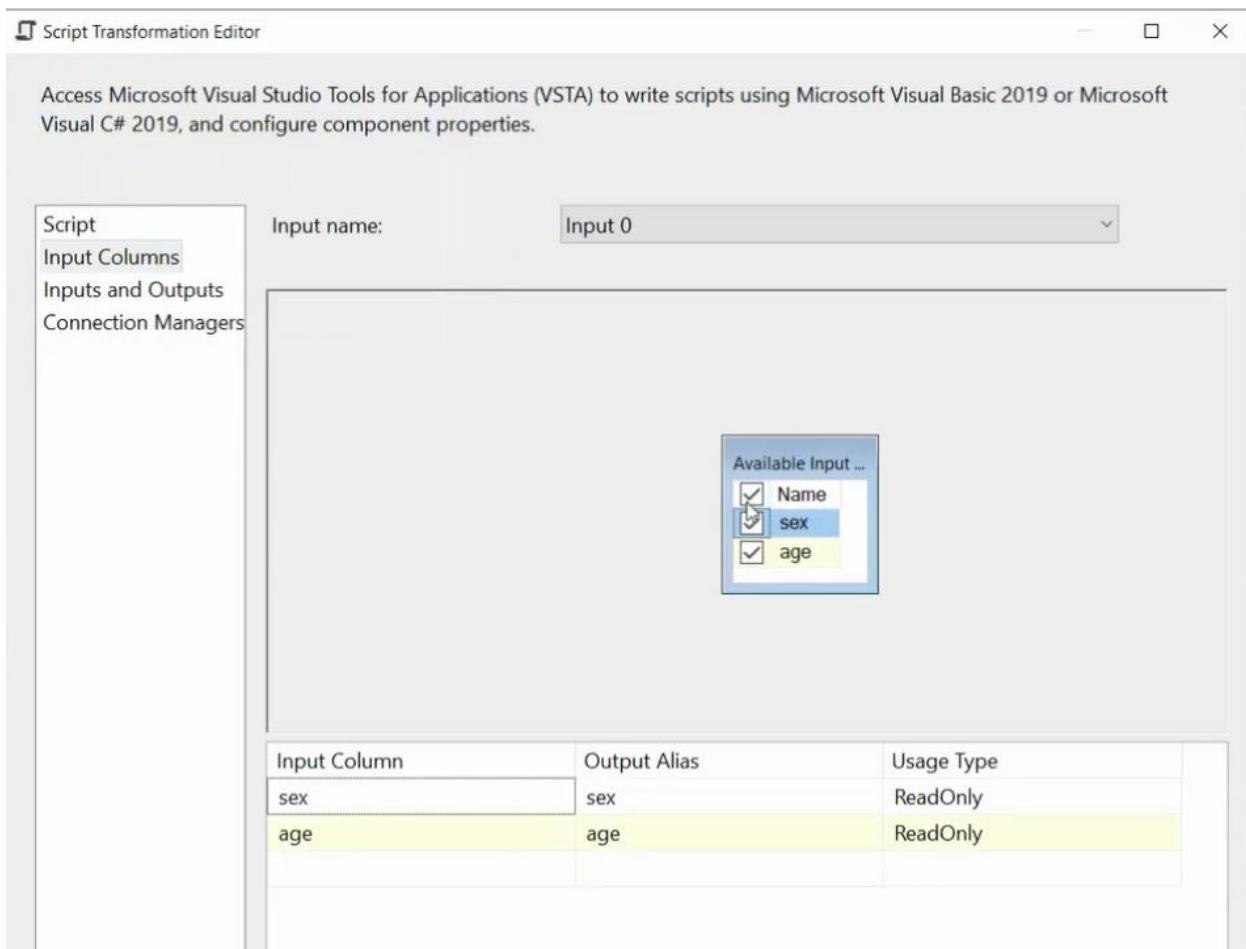
- Chọn thuộc tính cần Sort



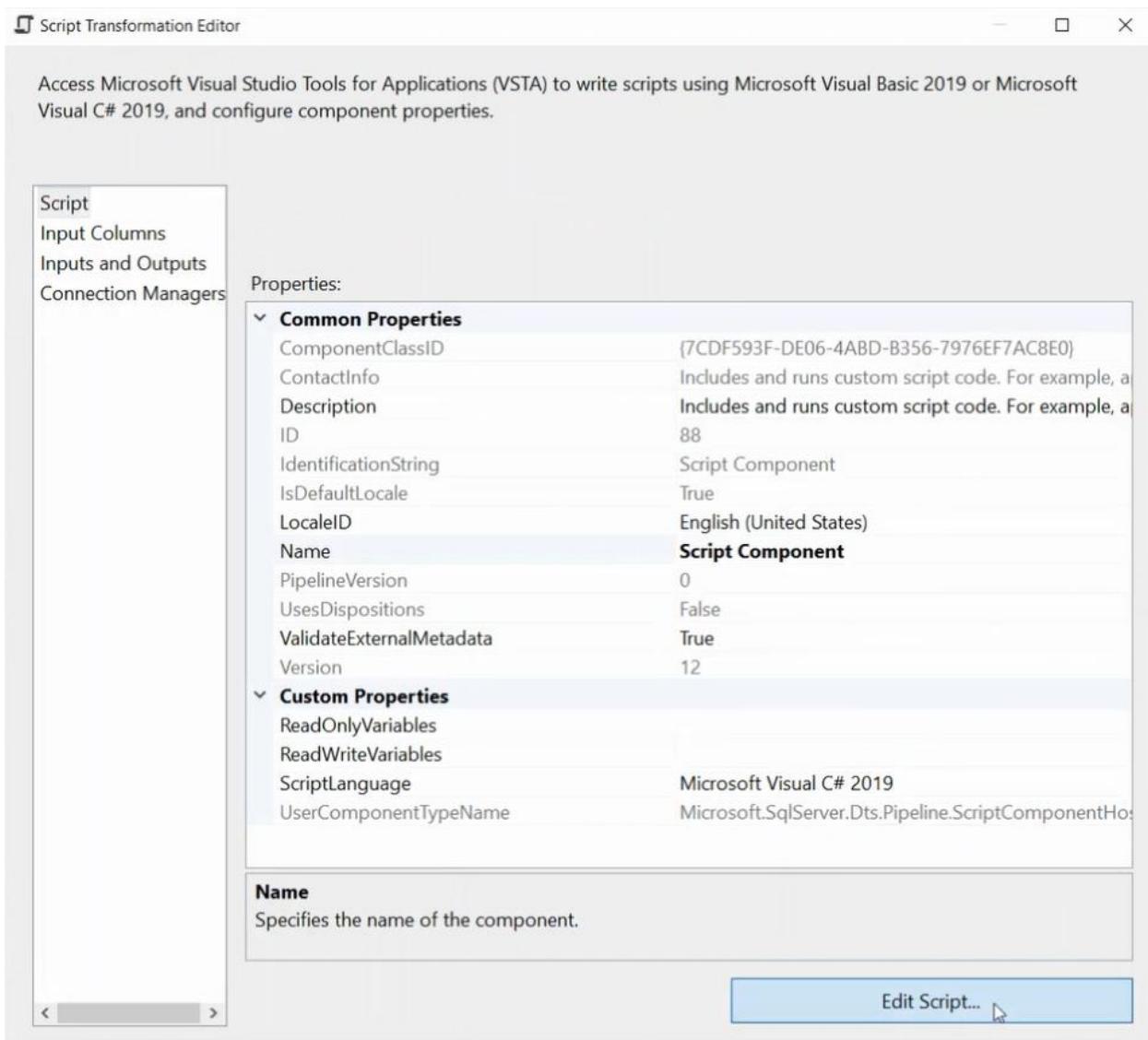
- Kéo thả Script Component , nhấn chuột phải chọn Edit



- Cấu hình Script Component



- Chọn Edit Script

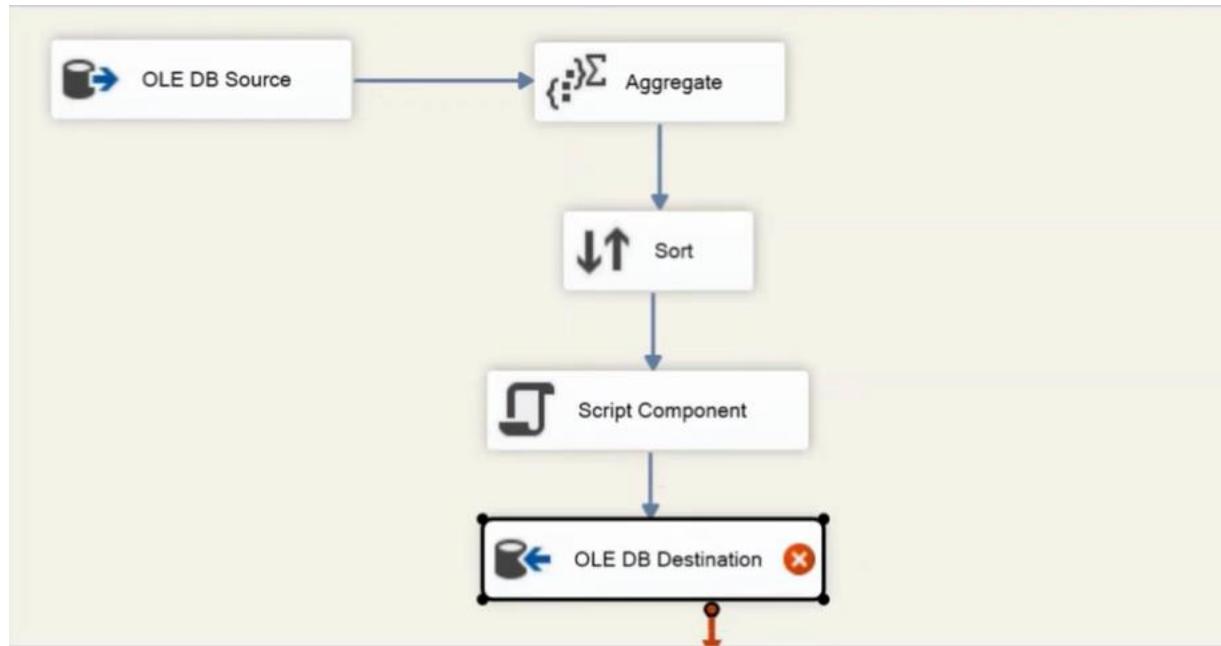


- Tạo cột ID tự tăng

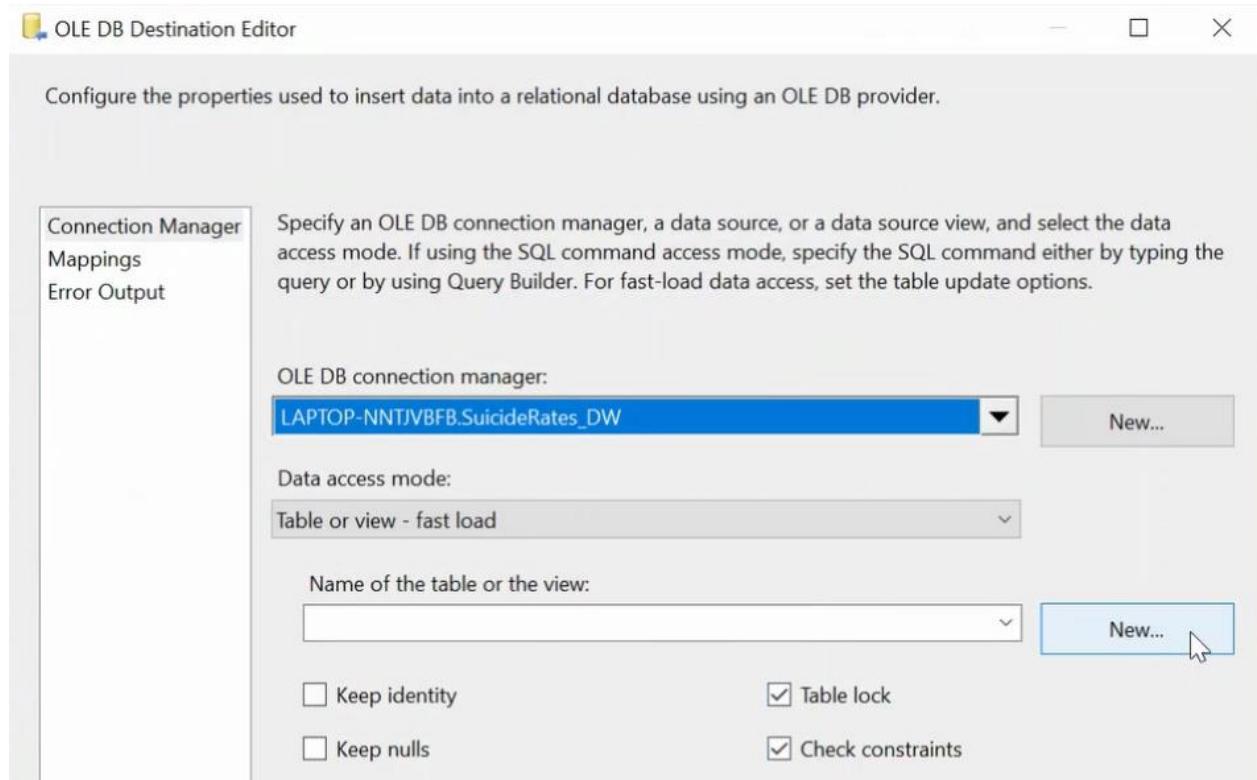
```
[Microsoft.SqlServer.Dts.Pipeline.SSISScriptComponentEntryPointAttribute]
public class ScriptMain : UserComponent
{
    int count = 1;           I
    Help: Using Integration Services variables and parameters
    Help: Using Integration Services Connection Managers
    Help: Firing Integration Services Events

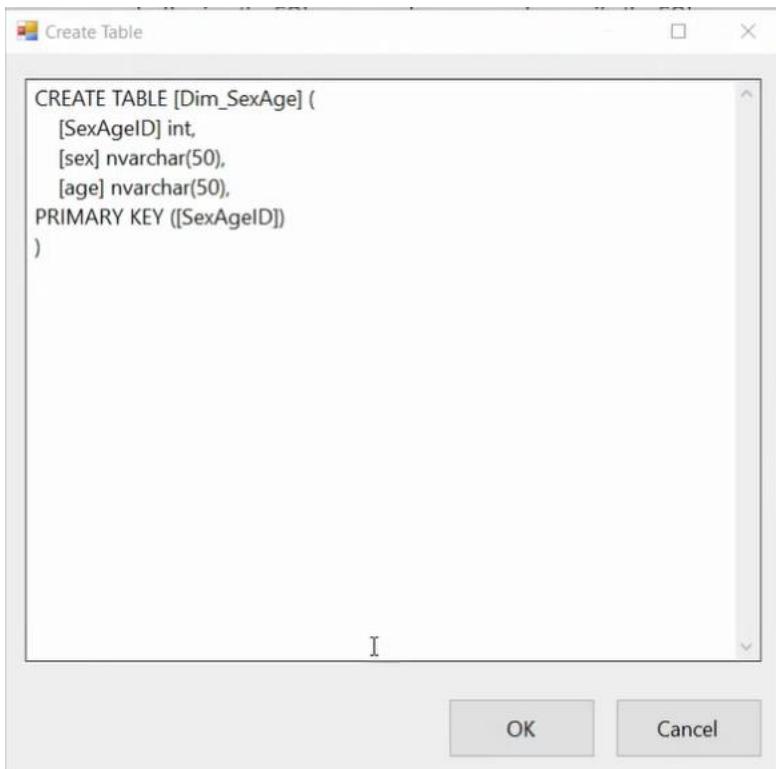
    public override void Input0_ProcessInputRow(Input0Buffer Row)
    {
        Row.SexAgeID = count;
        count++;
    }
}
```

- Kéo thả OLE DB Destination và tạo liên kết đến Script Component, nhấn chuột phải chọn Edit.

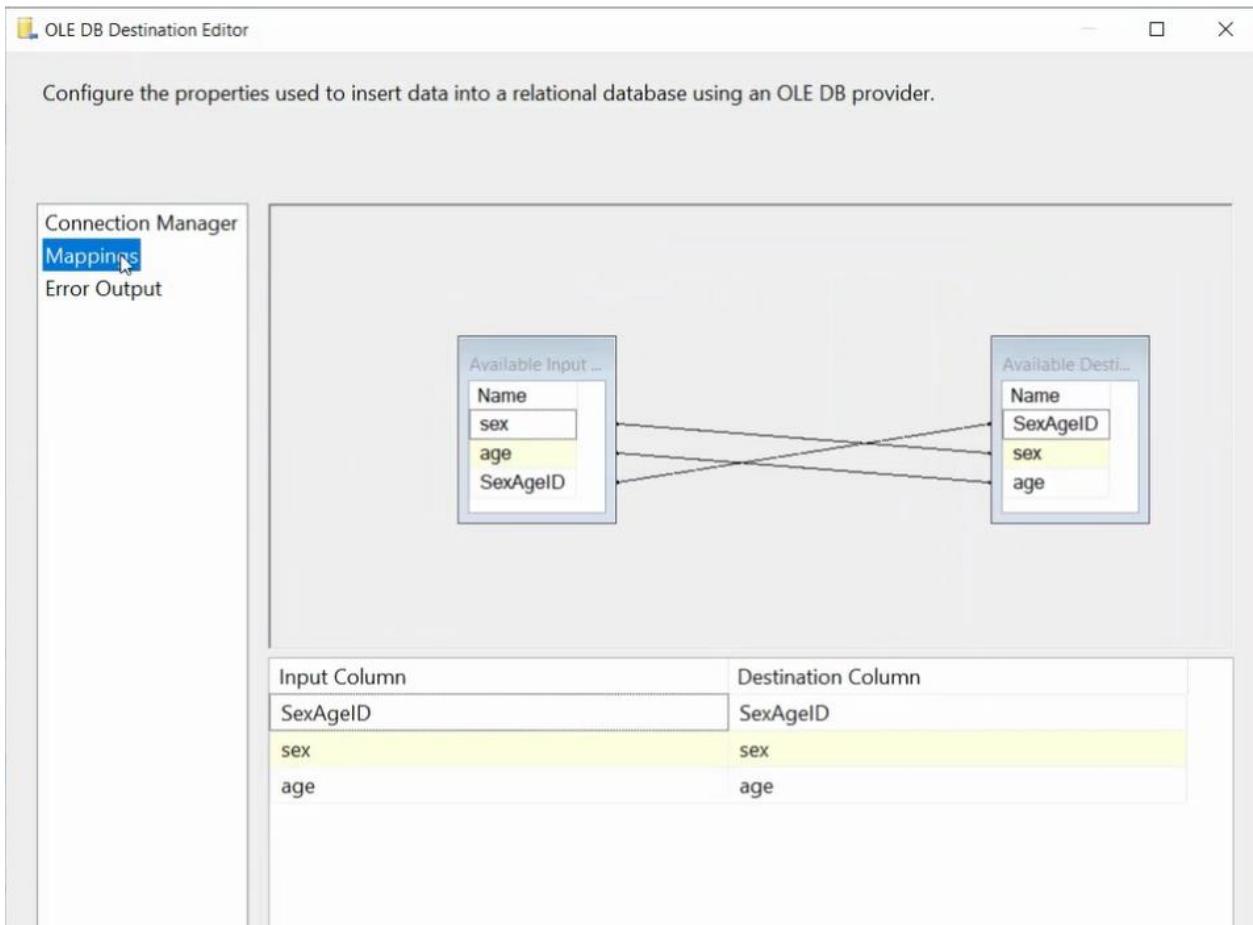


- Chọn kết nối đến SuicideRates_DW và chọn New để tạo bảng Dim_SexAge





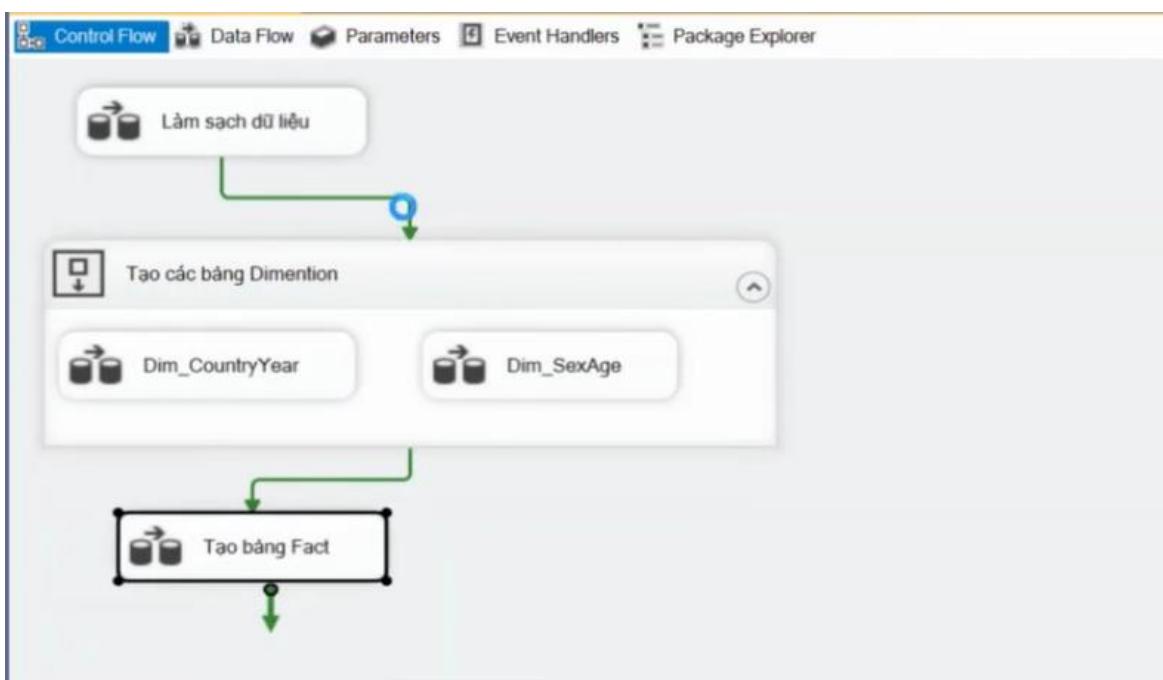
- Kiểm tra mapping các thuộc tính và nhấn OK



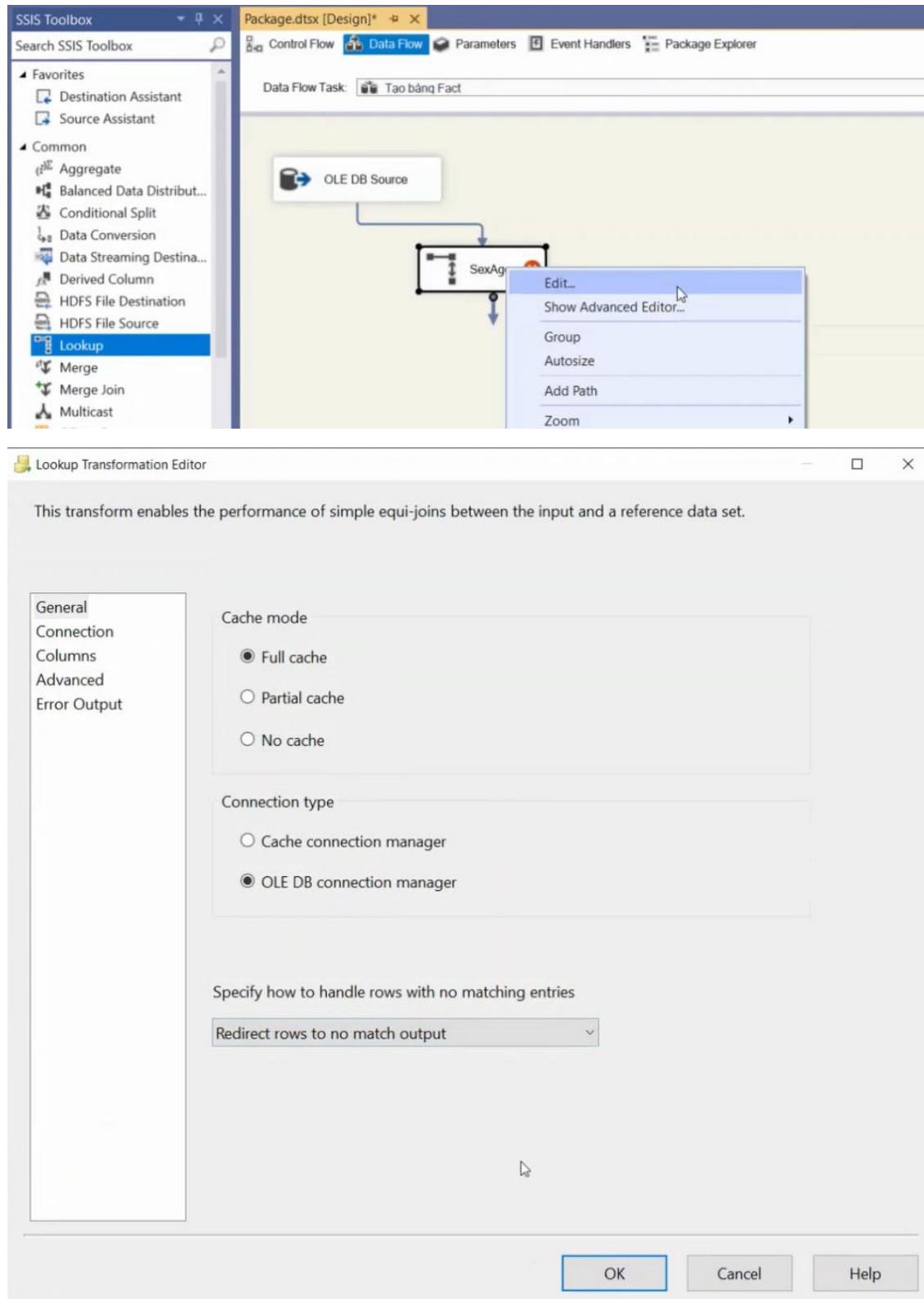


4.5. Quá trình tạo bảng Fact

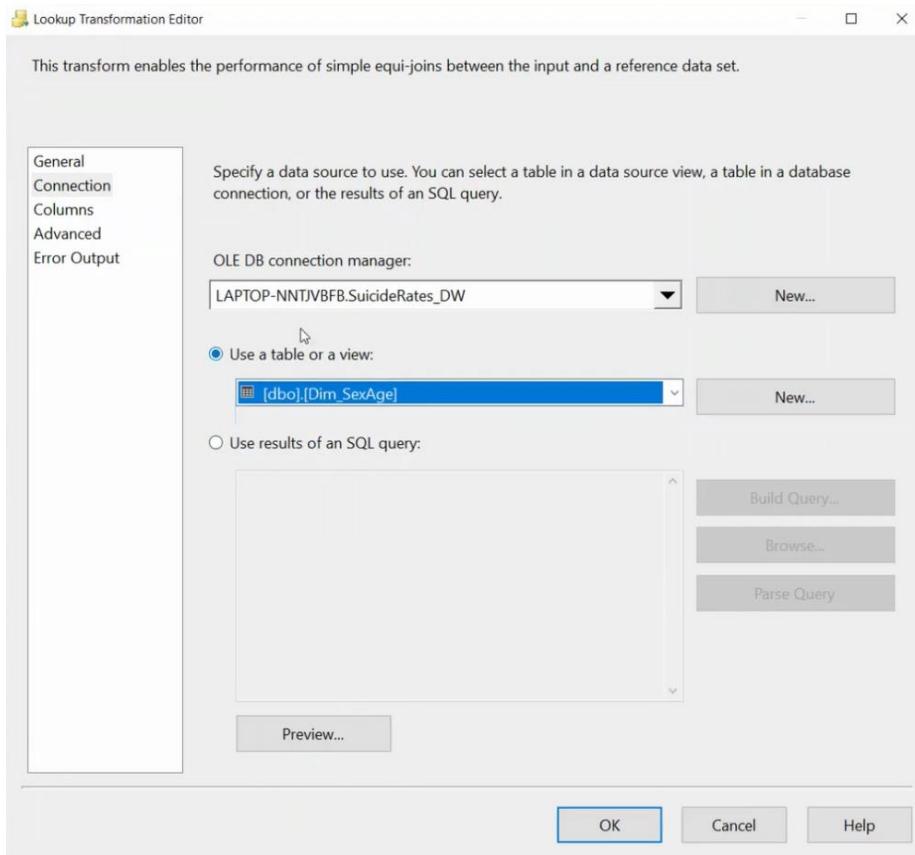
- Kéo thả Data Flow Task vào Sequence Container để tạo bảng Fact



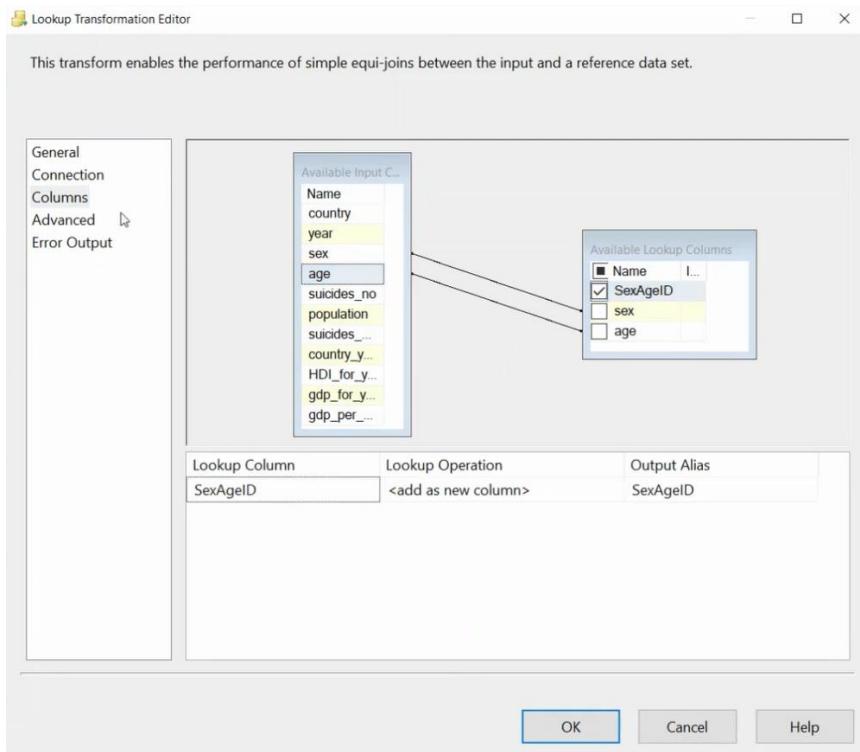
- Kéo thả OLE DB Source.
- Kéo thả Lookup, đổi tên SexAge và tạo liên kết đến OLE DB Source. Nhấn chuột phải chọn Edit



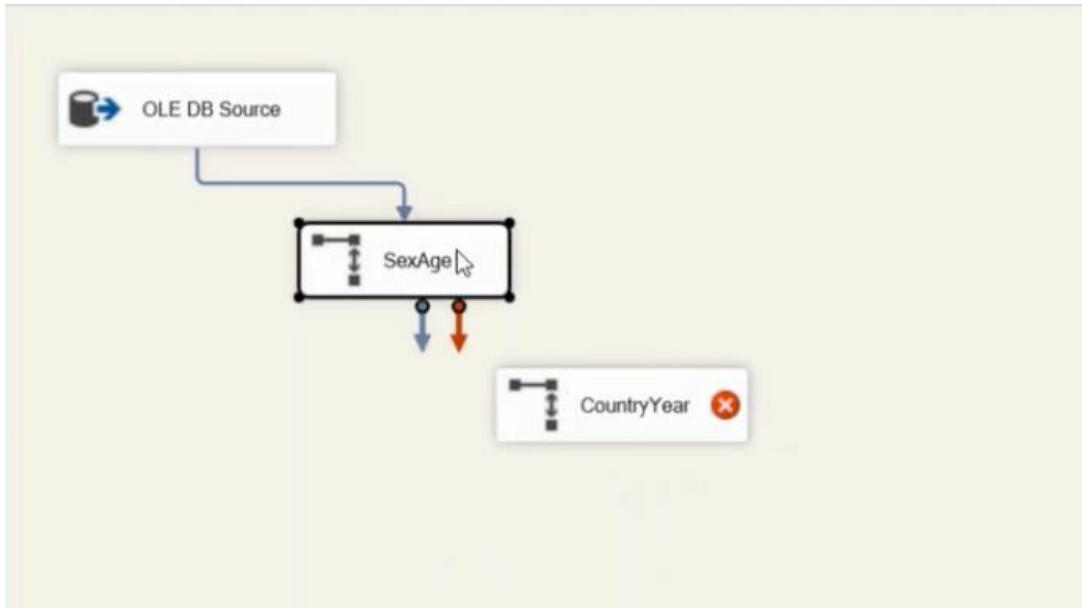
- Chọn connection → chọn bảng [Dim_SexAge]



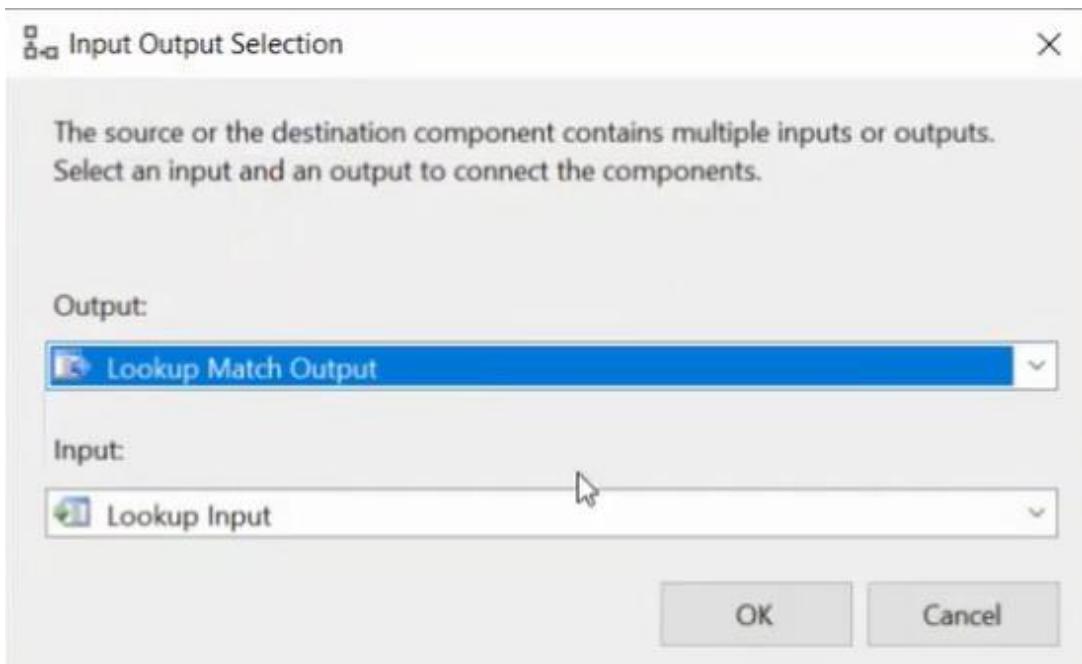
Kiểm tra Mapping dữ liệu



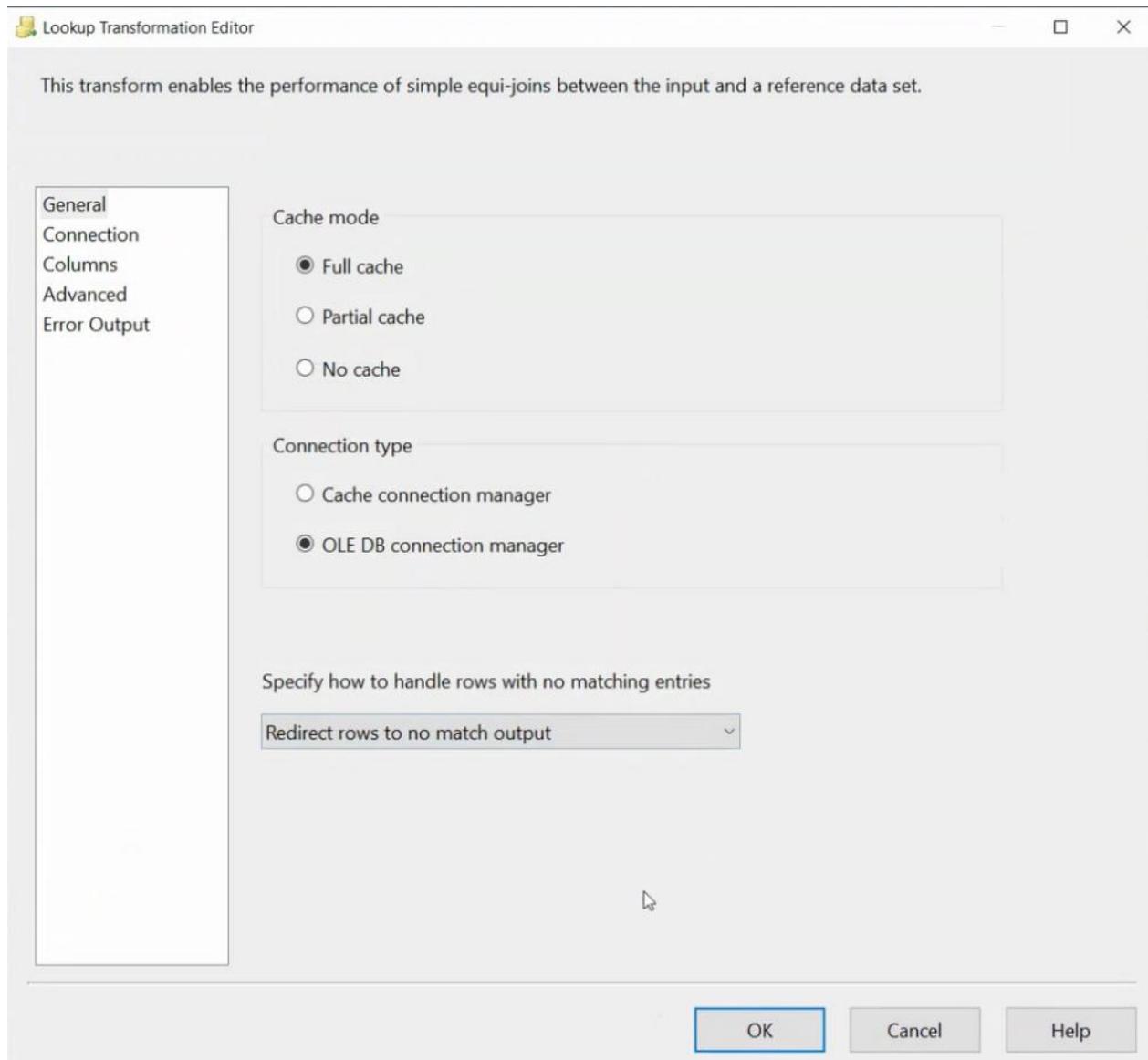
- Kéo thả Lookup, đổi tên CountryYear và tạo liên kết đến SexAge.



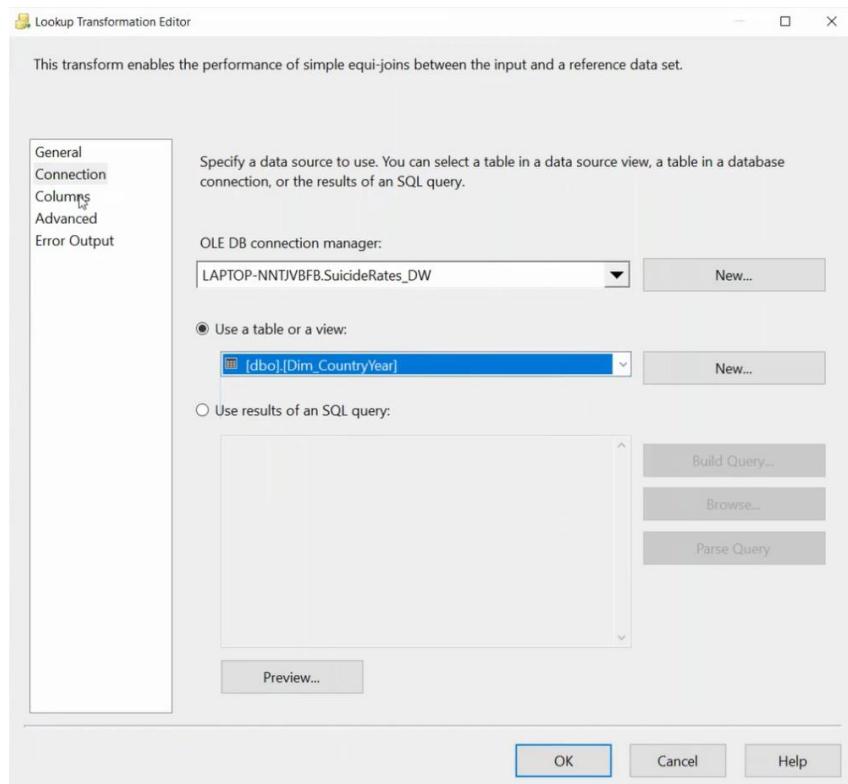
- Hộp thoại xuất hiện, chọn Lookup Match Output



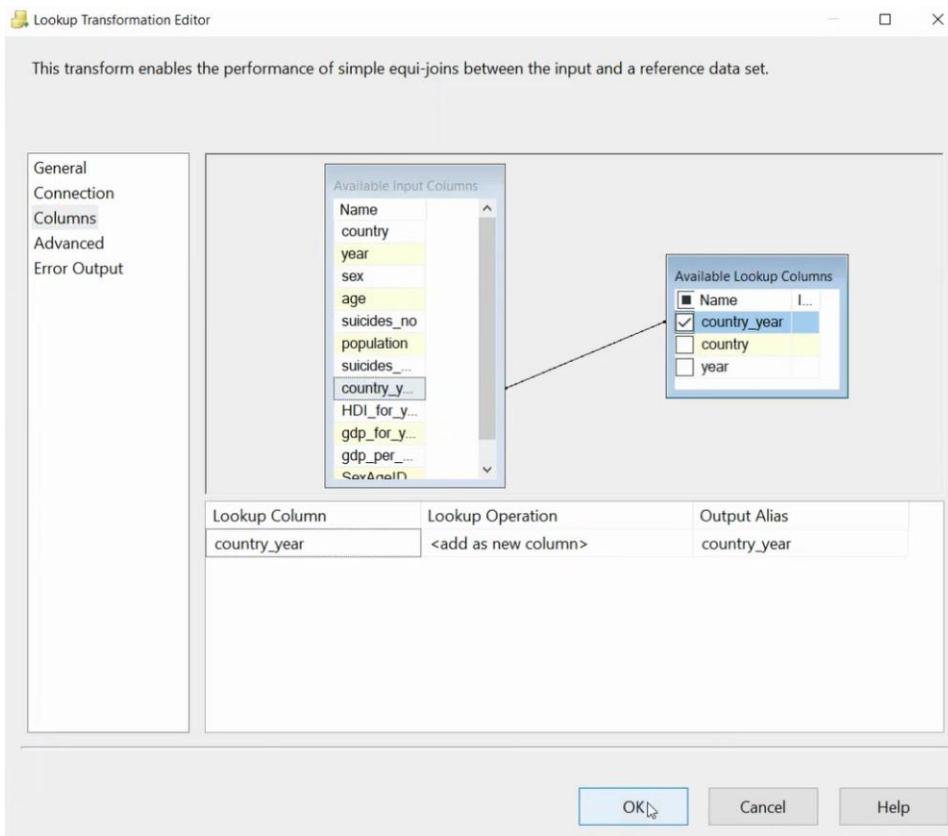
- Nhấn chuột phải chọn Edit



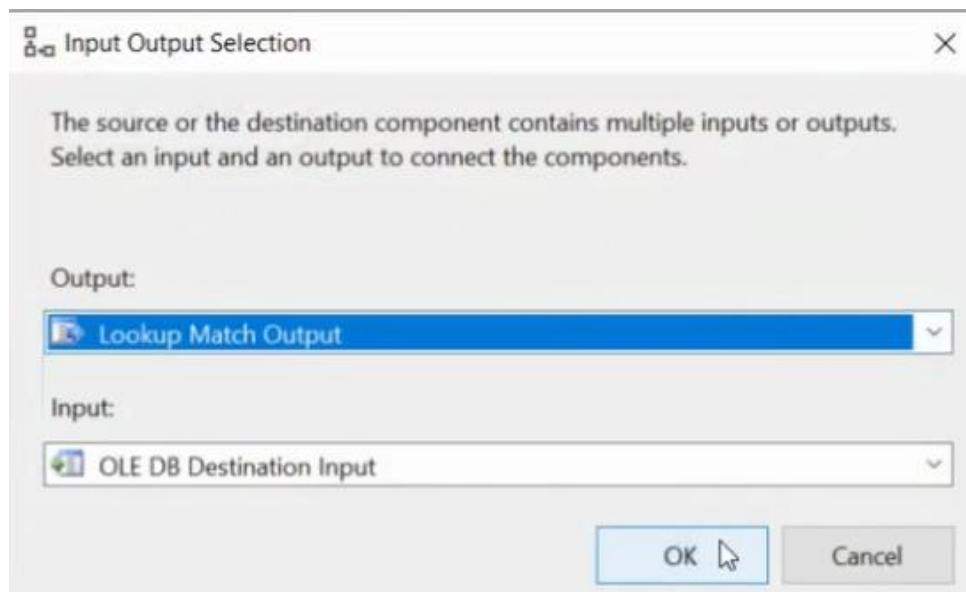
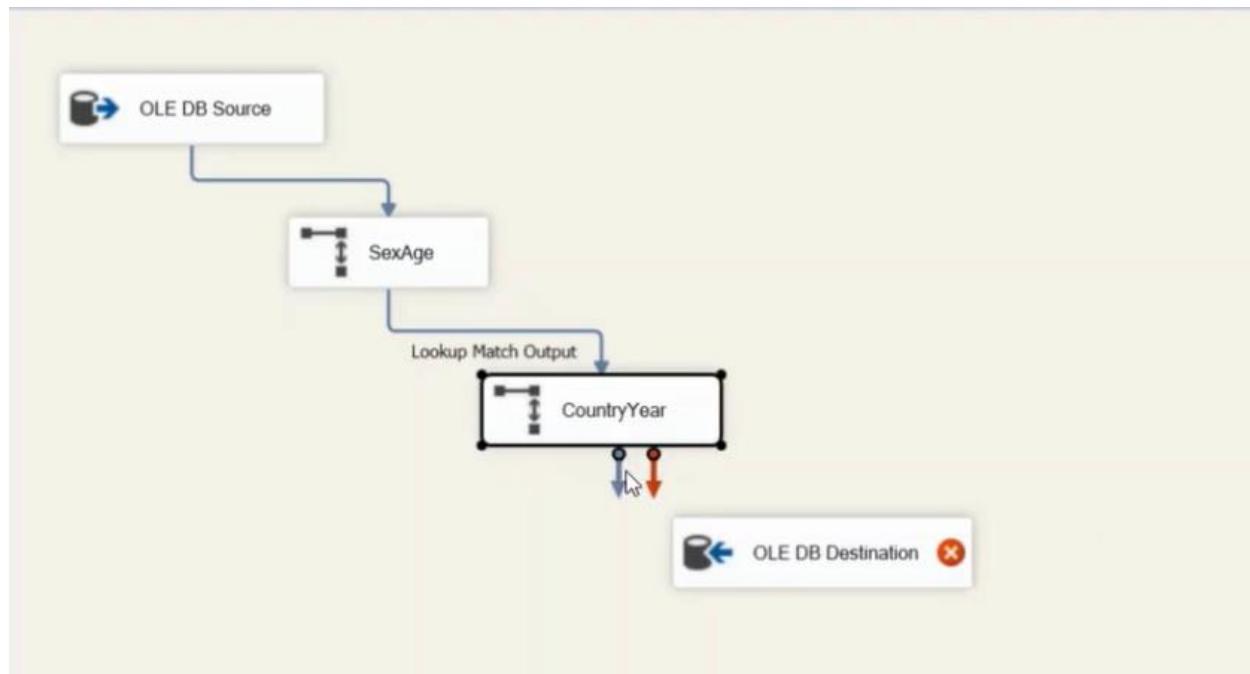
- Chọn connection → chọn bảng [Dim_SexAge]



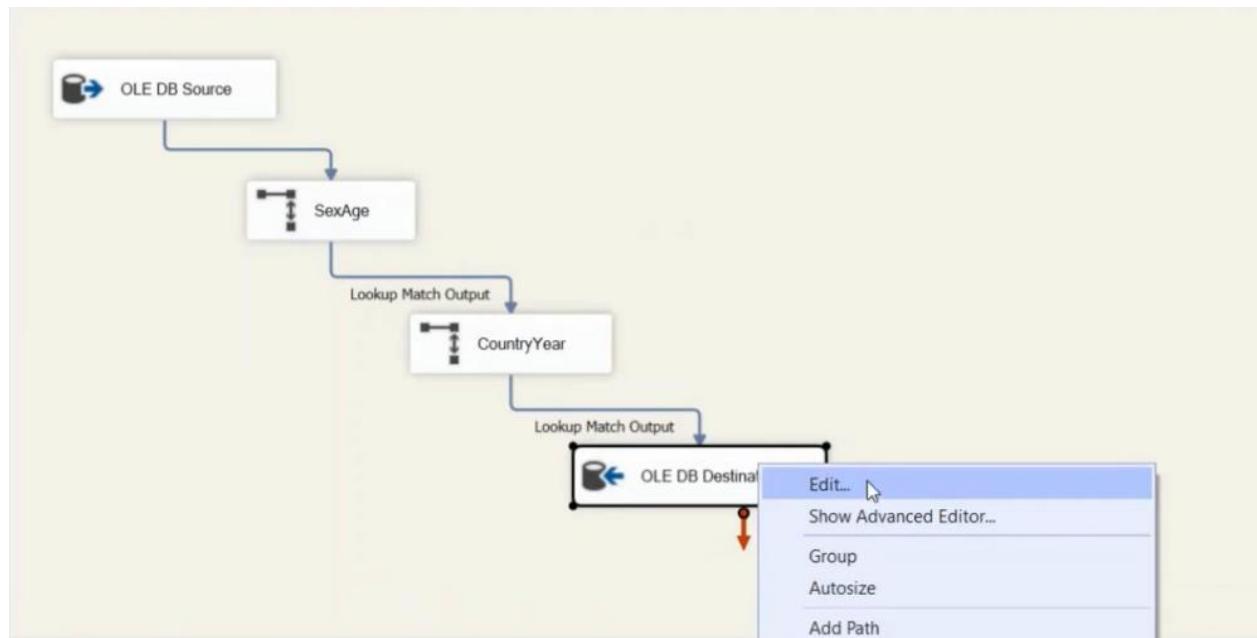
Kiểm tra Mapping dữ liệu



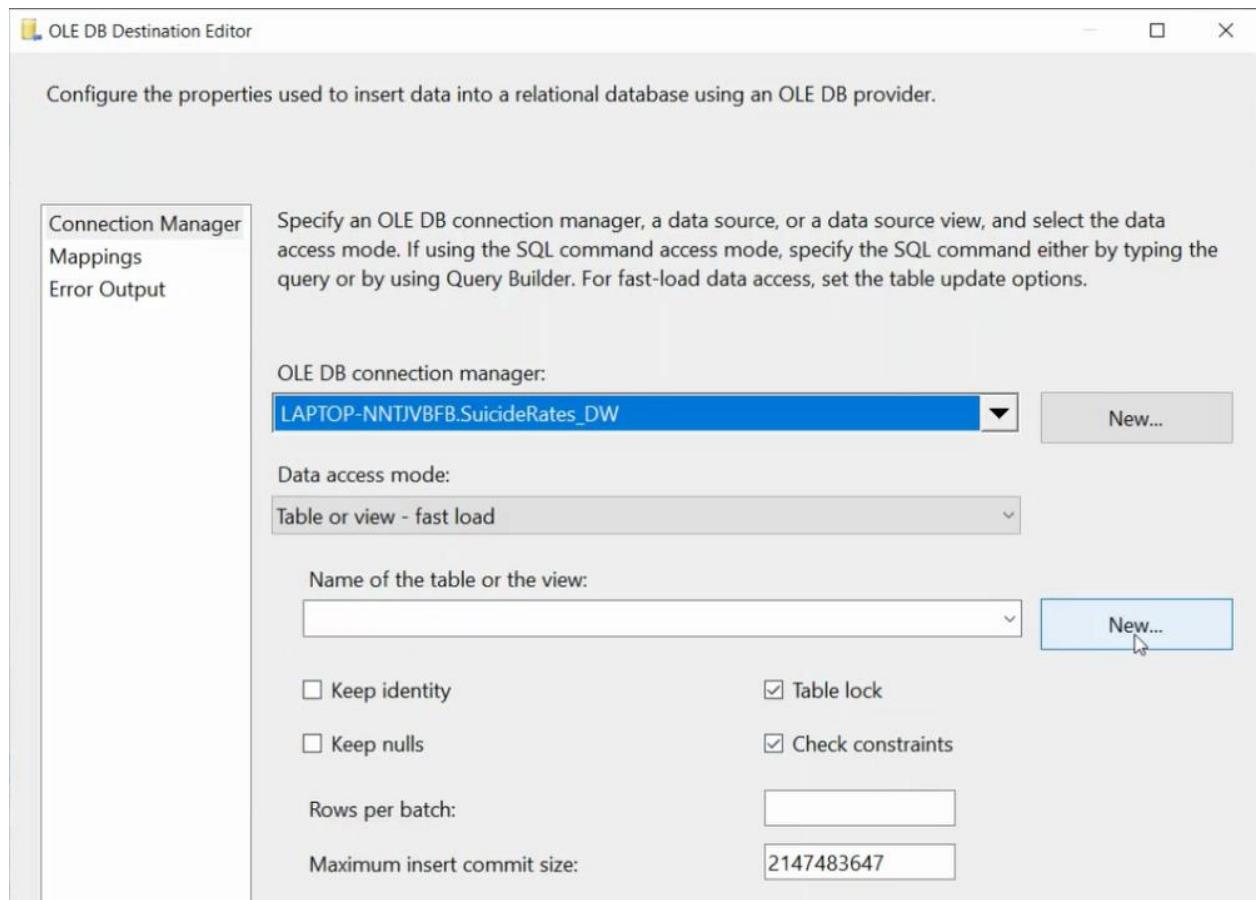
- Kéo thả OLE DB Destination và tạo liên kết đến CountryYear

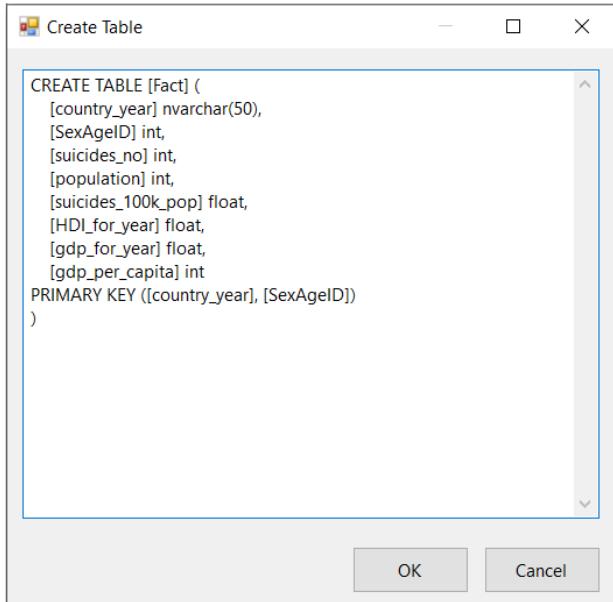


- Nhấn chuột phải vào OLE DB Destination chọn Edit

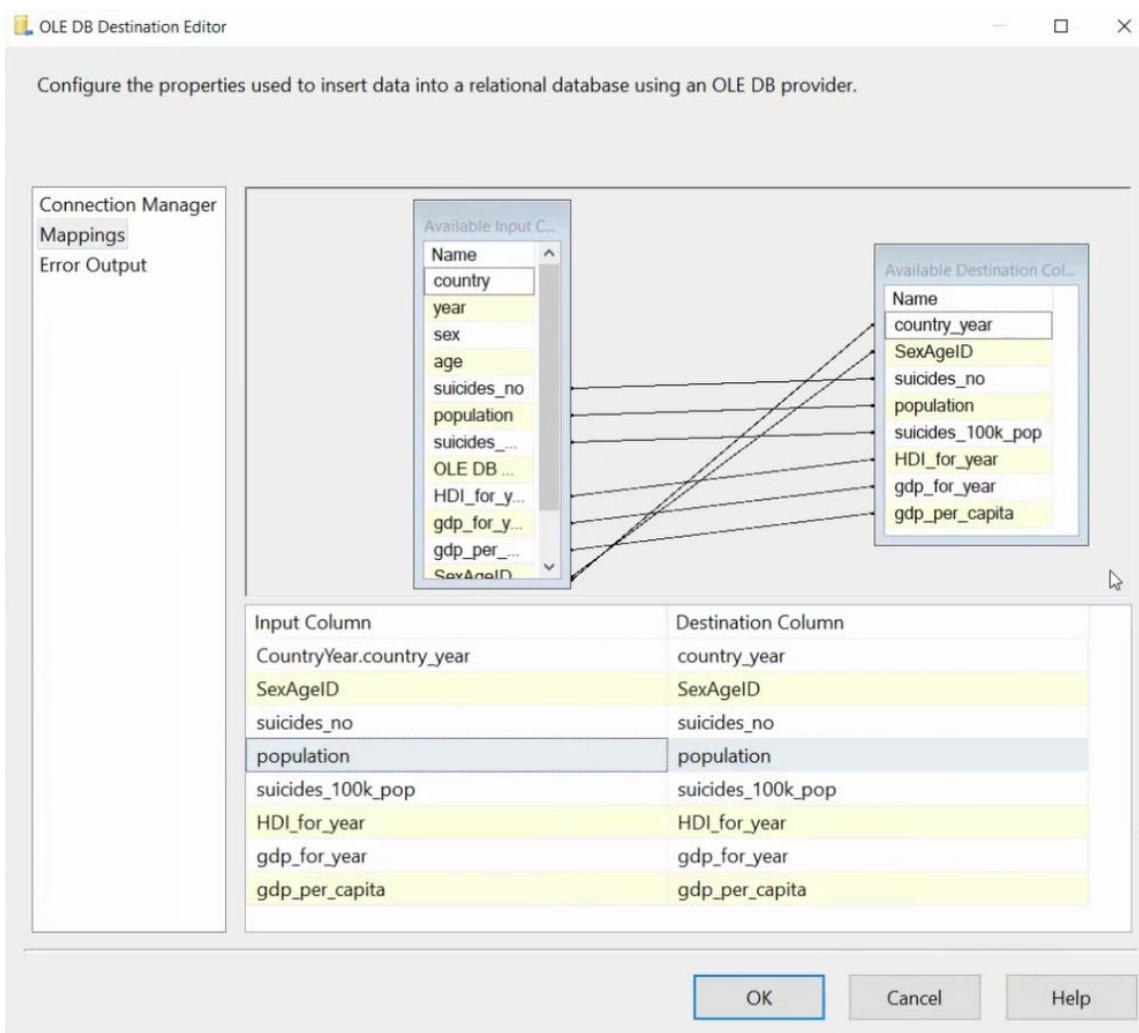


- Kết nối đến SuicideRates_DW và chọn New để tạo bảng Fact



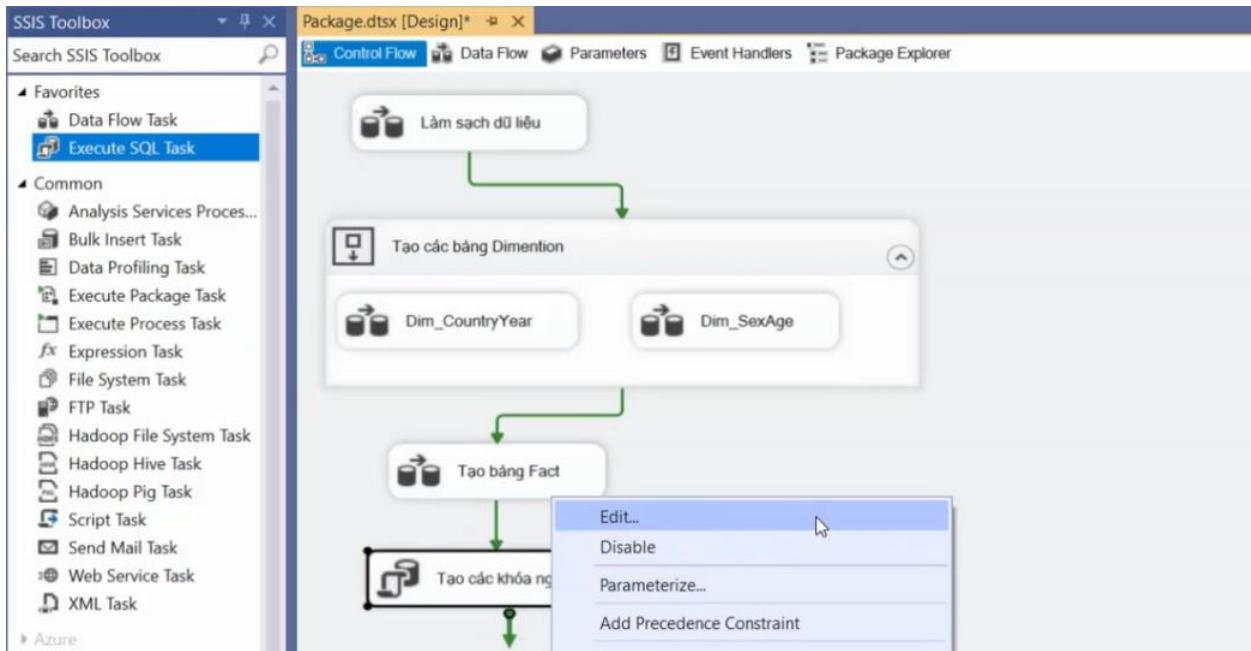


- Kiểm tra mapping các thuộc tính, Input column chọn CountryYear.country_year và nhấn OK.



4.6. Tạo ràng buộc khóa ngoại

- Kéo thả Execute SQL Task, đổi tên thành Tạo các khóa ngoại. Tạo liên kết với bảng Fact, nhấn chuột phải và chọn tiếp Edit.

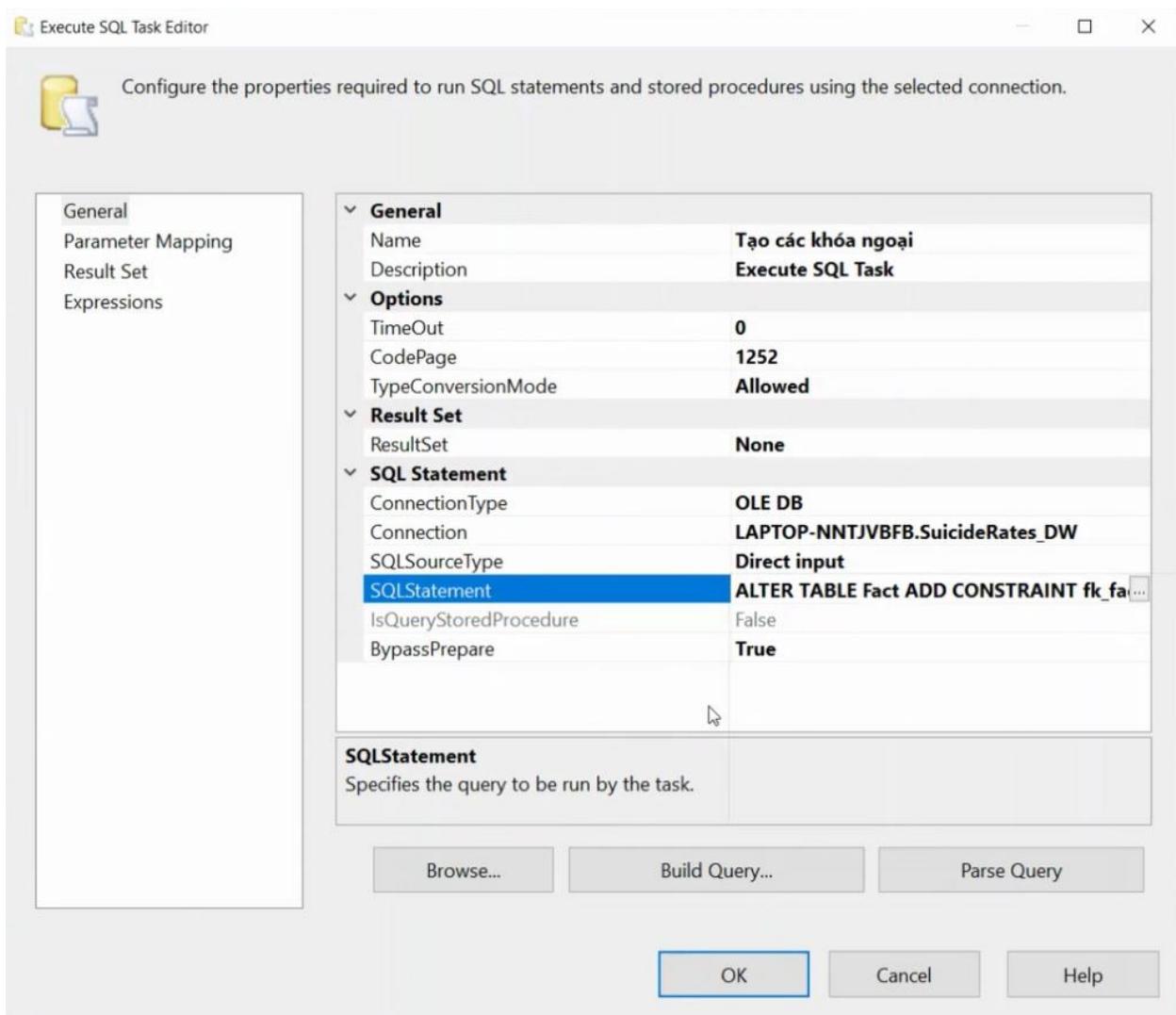


- Chọn kết nối đến SuicideRates_DW và viết câu lệnh tạo khóa ngoại

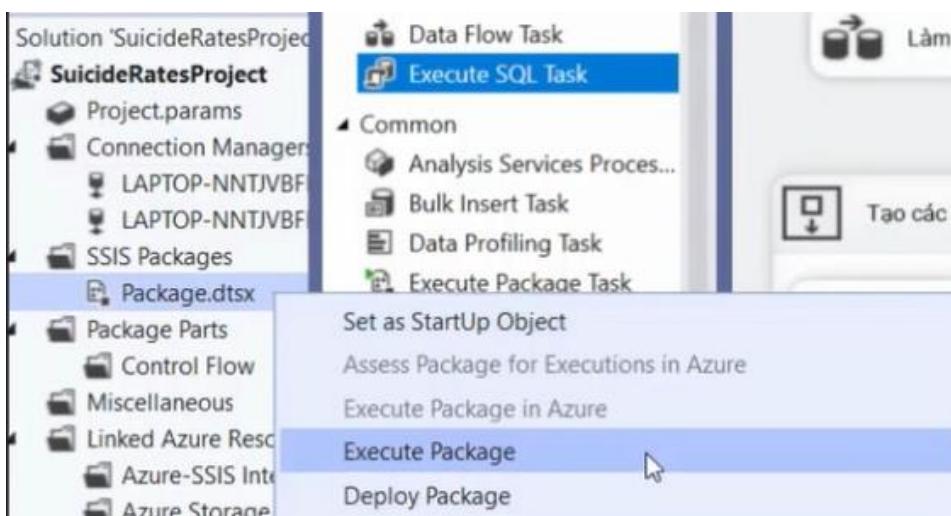
```
ALTER TABLE Fact ADD CONSTRAINT fk_fact_sexage
FOREIGN KEY(SexAgeID)
REFERENCES Dim_SexAge(SexAgeID);

GO

ALTER TABLE Fact ADD CONSTRAINT
fk_fact_countryyear
FOREIGN KEY(country_year)
REFERENCES Dim_CountryYear(country_year);
```



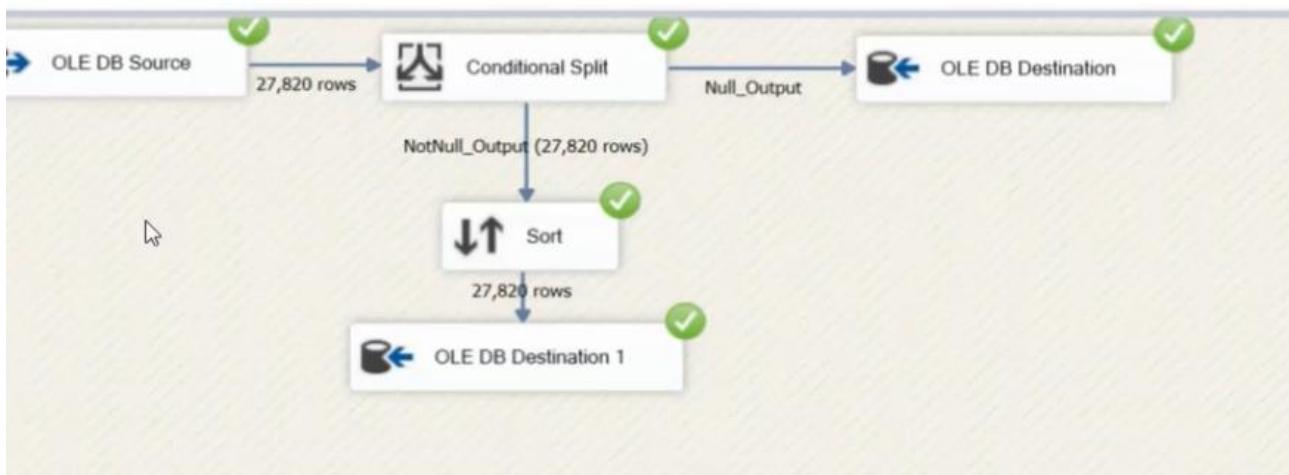
4.7. Thực thi Package



- Control Flow:



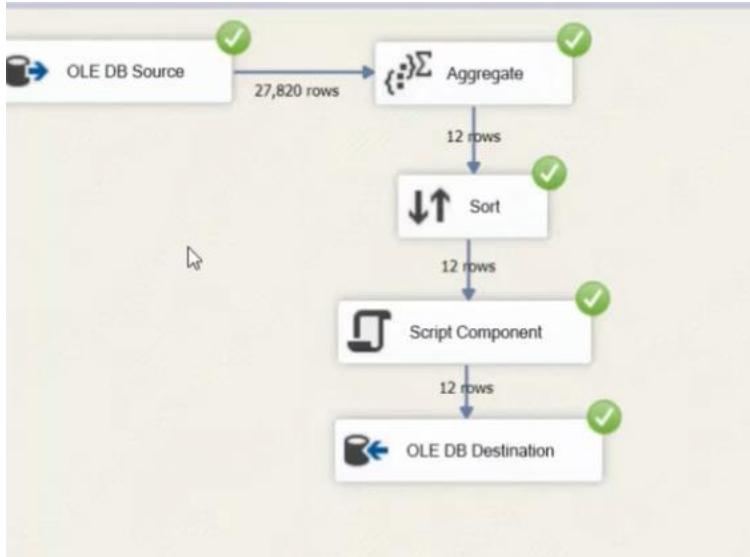
- Data Flow quá trình làm sạch dữ liệu:



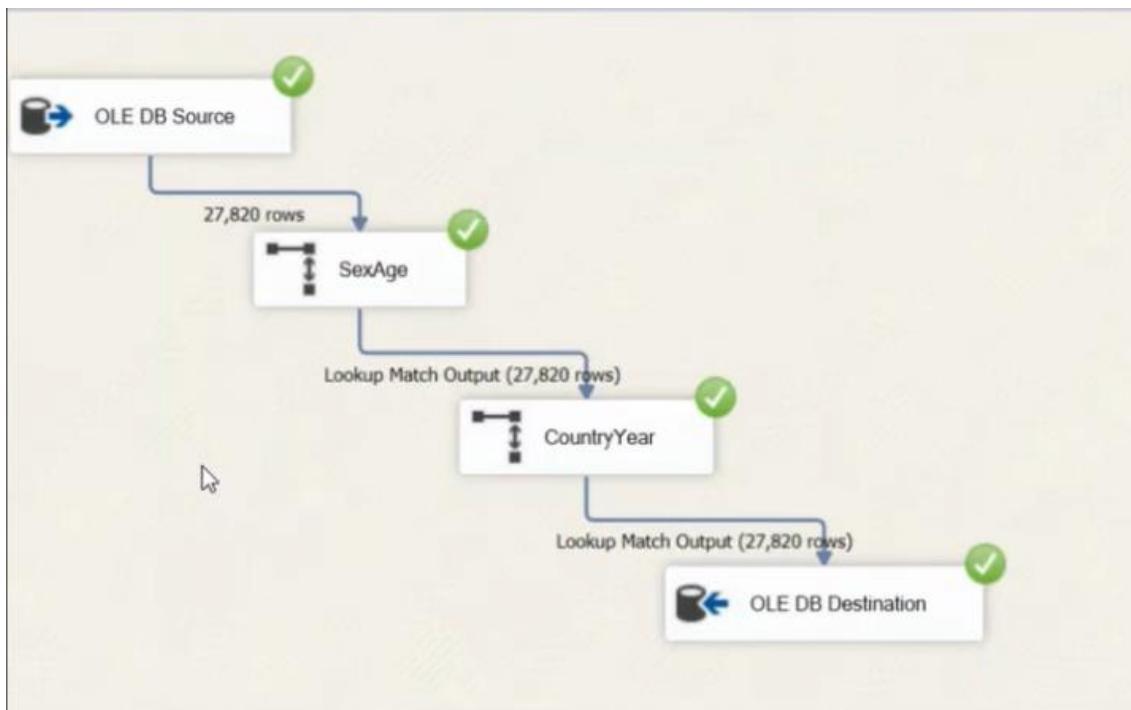
- Data Flow quá trình tạo bảng Dim_CountryYear:



- Data Flow quá trình tạo bảng Dim_SexAge:



- Data Flow quá trình tạo bảng Fact:



4.8. Kiểm tra database

a. Bảng Dim_CountryYear

The screenshot shows the SQL Server Management Studio interface. On the left, the Object Explorer displays a tree view of database objects, including External Tables, Graph Tables, and various tables like dbo.CleanData, dbo.NullData, dbo.suicideRates, and the target table dbo.Dim_CountryYear. On the right, the Results tab shows the output of a T-SQL script titled "***** Script for SelectTopNRows command". The script lists 34 rows from the Dim_CountryYear table, mapping country codes to country names and years.

	country_year	country	year
1	Albania1987	Albania	1987
2	Albania1988	Albania	1988
3	Albania1989	Albania	1989
4	Albania1992	Albania	1992
5	Albania1993	Albania	1993
6	Albania1994	Albania	1994
7	Albania1995	Albania	1995
8	Albania1996	Albania	1996
9	Albania1997	Albania	1997
10	Albania1998	Albania	1998
11	Albania1999	Albania	1999
12	Albania2000	Albania	2000
13	Albania2001	Albania	2001
14	Albania2002	Albania	2002
15	Albania2003	Albania	2003
16	Albania2004	Albania	2004
17	Albania2005	Albania	2005
18	Albania2006	Albania	2006
19	Albania2007	Albania	2007
20	Albania2008	Albania	2008
21	Albania2009	Albania	2009
22	Albania2010	Albania	2010
23	Antigua an...	Antig...	1985
24	Antigua an...	Antig...	1986
25	Antigua an...	Antig...	1987
26	Antigua an...	Antig...	1988
27	Antigua an...	Antig...	1989
28	Antigua an...	Antig...	1990
29	Antigua an...	Antig...	1991
30	Antigua an...	Antig...	1992
31	Antigua an...	Antig...	1993
32	Antigua an...	Antig...	1994
33	Antigua an...	Antig...	1995
34	Antigua an...	Antig...	1998

b. Bảng Dim_SexAge

The screenshot shows the SSMS interface with the Object Explorer on the left and a query results grid on the right.

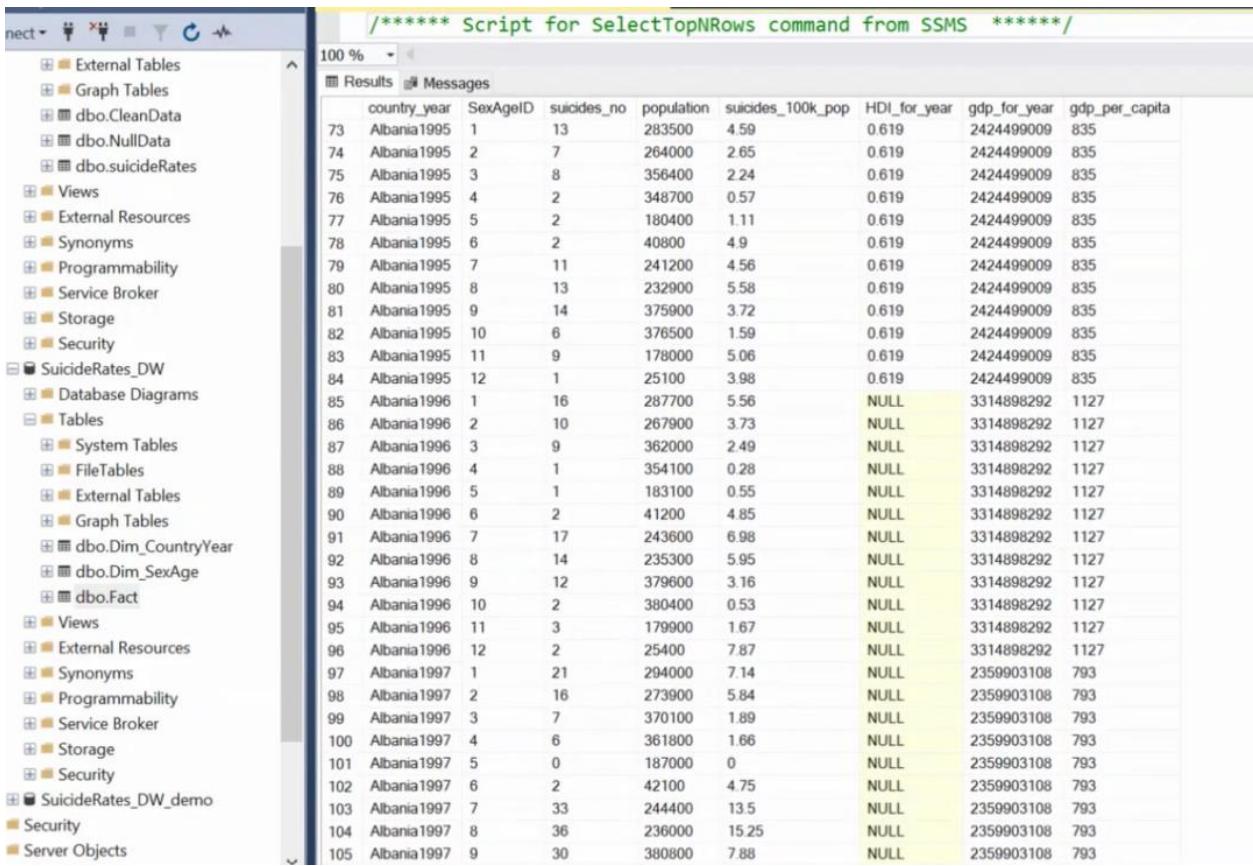
Object Explorer:

- Database: SuicideRates_DW
 - Tables
 - Dim_CountryYear
 - Dim_SexAge
 - Fact
 - Views
 - External Resources
 - Synonyms
 - Programmability
 - Service Broker
 - Storage
 - Security

SQL Query Results Grid:

	SexAgeID	sex	age
1	1	female	15-24 years
2	2	female	25-34 years
3	3	female	35-54 years
4	4	female	5-14 years
5	5	female	55-74 years
6	6	female	75+ years
7	7	male	15-24 years
8	8	male	25-34 years
9	9	male	35-54 years
10	10	male	5-14 years
11	11	male	55-74 years
12	12	male	75+ years

c. Bảng Fact



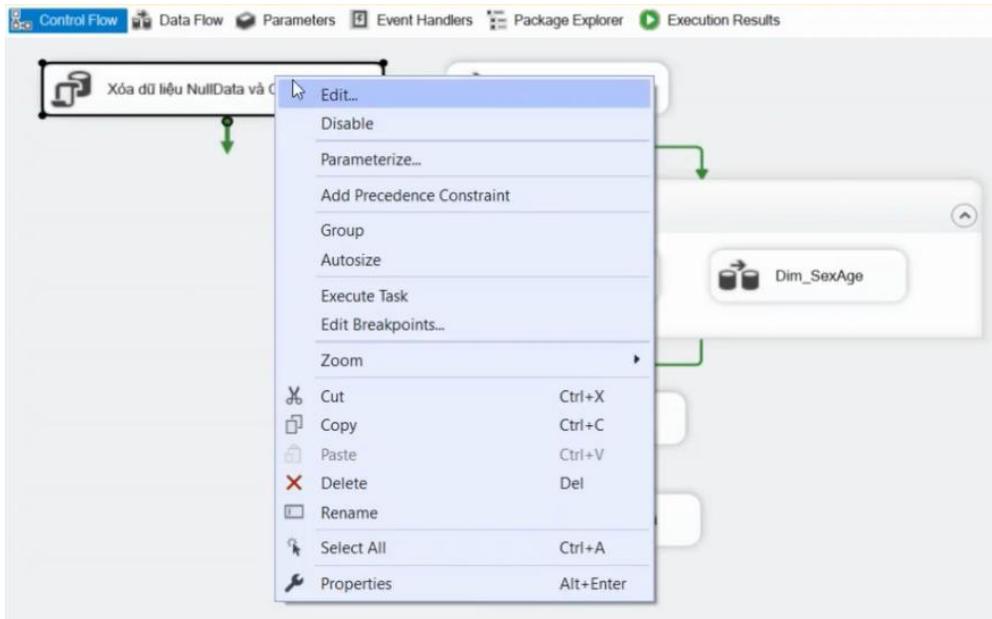
The screenshot shows the SSMS interface with the following details:

- Object Explorer:** Shows the database structure including External Tables, Graph Tables, dbo.CleanData, dbo.NullData, dbo.suicideRates, Views, External Resources, Synonyms, Programmability, Service Broker, Storage, Security, and two SuicideRates_DW databases.
- Results Tab:** Displays the output of the query: `***** Script for SelectTopNRows command from SSMS *****`. The results show data for Dim_SexAge with columns: country_year, SexAgeID, suicides_no, population, suicides_100k_pop, HDI_for_year, gdp_for_year, and gdp_per_capita. The data spans from 1995 to 1997 for Albania.

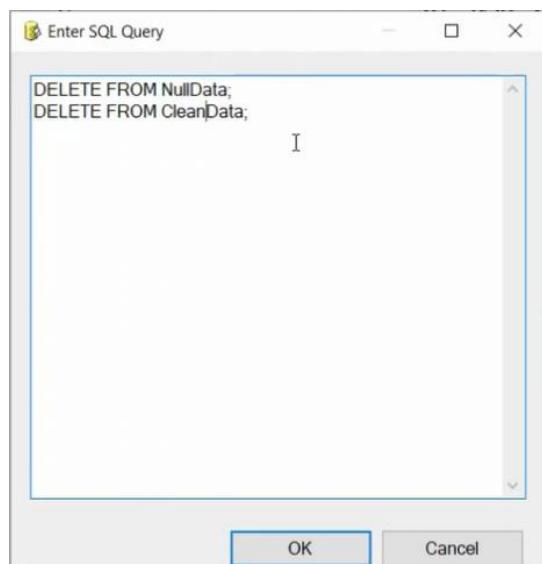
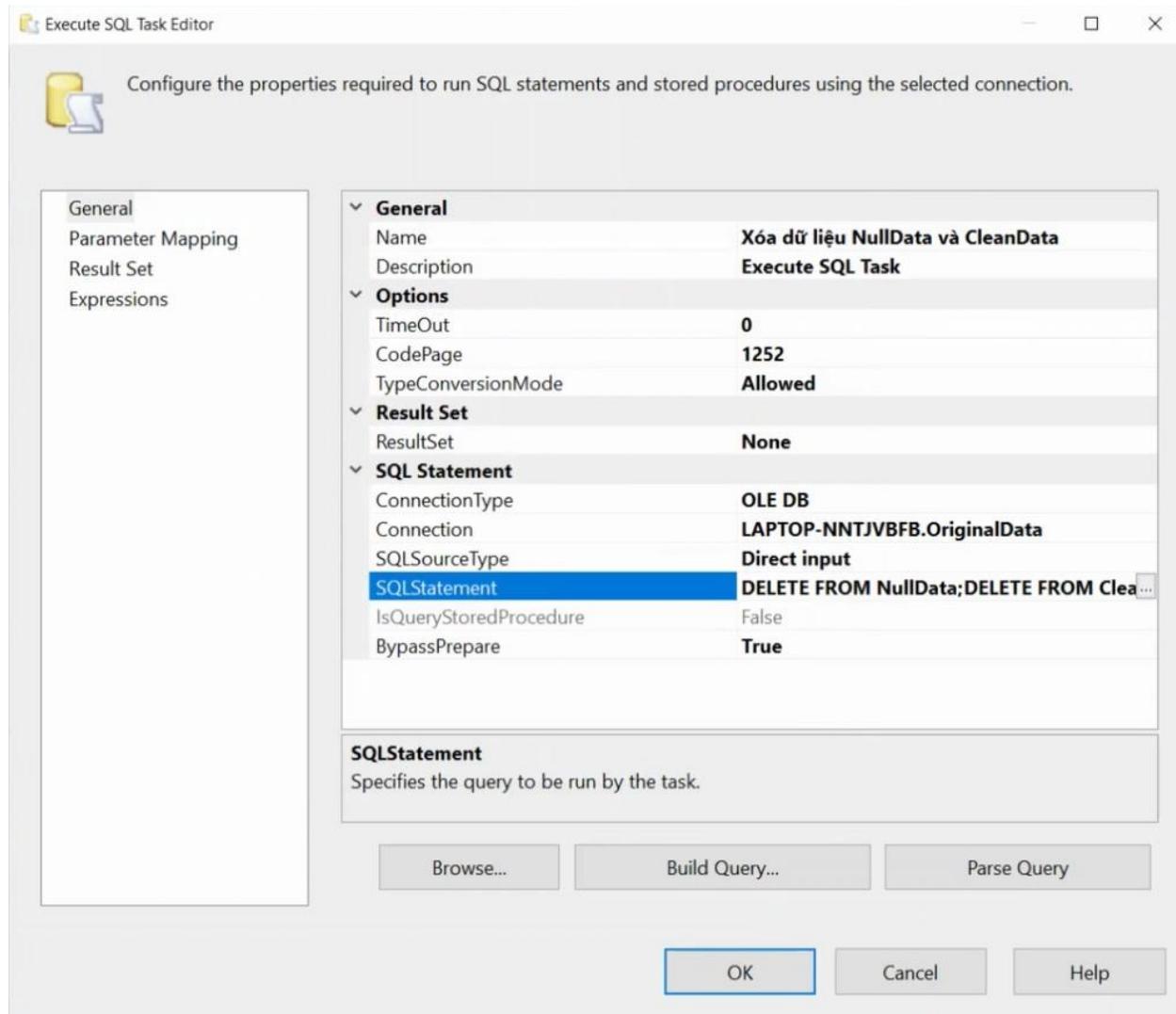
country_year	SexAgeID	suicides_no	population	suicides_100k_pop	HDI_for_year	gdp_for_year	gdp_per_capita	
73	Albania1995	1	13	283500	4.59	0.619	2424499009	835
74	Albania1995	2	7	264000	2.65	0.619	2424499009	835
75	Albania1995	3	8	356400	2.24	0.619	2424499009	835
76	Albania1995	4	2	348700	0.57	0.619	2424499009	835
77	Albania1995	5	2	180400	1.11	0.619	2424499009	835
78	Albania1995	6	2	408000	4.9	0.619	2424499009	835
79	Albania1995	7	11	241200	4.56	0.619	2424499009	835
80	Albania1995	8	13	232900	5.58	0.619	2424499009	835
81	Albania1995	9	14	375900	3.72	0.619	2424499009	835
82	Albania1995	10	6	376500	1.59	0.619	2424499009	835
83	Albania1995	11	9	178000	5.06	0.619	2424499009	835
84	Albania1995	12	1	25100	3.98	0.619	2424499009	835
85	Albania1996	1	16	287700	5.56	NULL	3314898292	1127
86	Albania1996	2	10	267900	3.73	NULL	3314898292	1127
87	Albania1996	3	9	362000	2.49	NULL	3314898292	1127
88	Albania1996	4	1	354100	0.28	NULL	3314898292	1127
89	Albania1996	5	1	183100	0.55	NULL	3314898292	1127
90	Albania1996	6	2	41200	4.85	NULL	3314898292	1127
91	Albania1996	7	17	243600	6.98	NULL	3314898292	1127
92	Albania1996	8	14	235300	5.95	NULL	3314898292	1127
93	Albania1996	9	12	379600	3.16	NULL	3314898292	1127
94	Albania1996	10	2	380400	0.53	NULL	3314898292	1127
95	Albania1996	11	3	179900	1.67	NULL	3314898292	1127
96	Albania1996	12	2	25400	7.87	NULL	3314898292	1127
97	Albania1997	1	21	294000	7.14	NULL	2359903108	793
98	Albania1997	2	16	273900	5.84	NULL	2359903108	793
99	Albania1997	3	7	370100	1.89	NULL	2359903108	793
100	Albania1997	4	6	361800	1.66	NULL	2359903108	793
101	Albania1997	5	0	187000	0	NULL	2359903108	793
102	Albania1997	6	2	42100	4.75	NULL	2359903108	793
103	Albania1997	7	33	244400	13.5	NULL	2359903108	793
104	Albania1997	8	36	236000	15.25	NULL	2359903108	793
105	Albania1997	9	30	380800	7.88	NULL	2359903108	793

4.9. Tạo và cấu hình ExecuteSQL Task

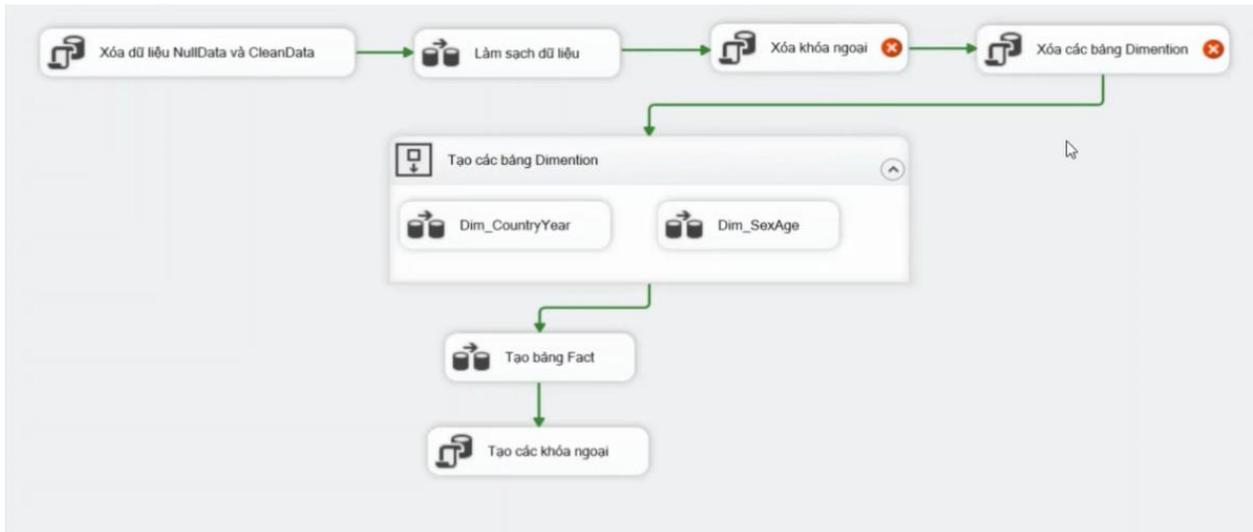
- Kéo thả ExecuteSQL Task và đổi tên thành xóa dữ liệu NullData và CleanData, chuột phải chọn Edit



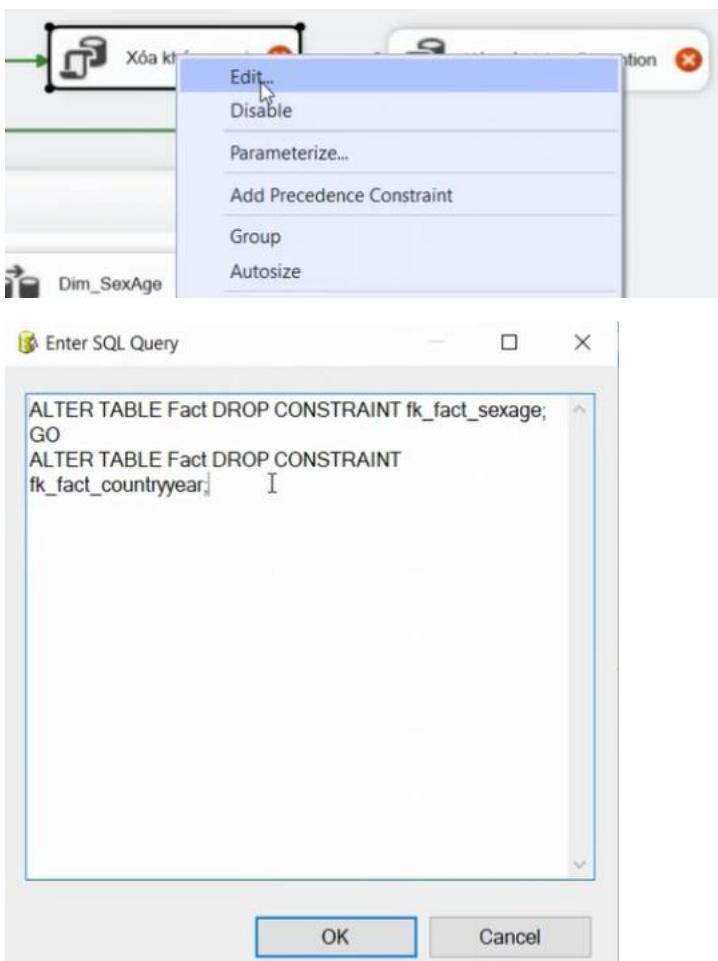
- Chọn kết nối đến OriginalData và viết câu truy vấn xóa NullData và CleanData

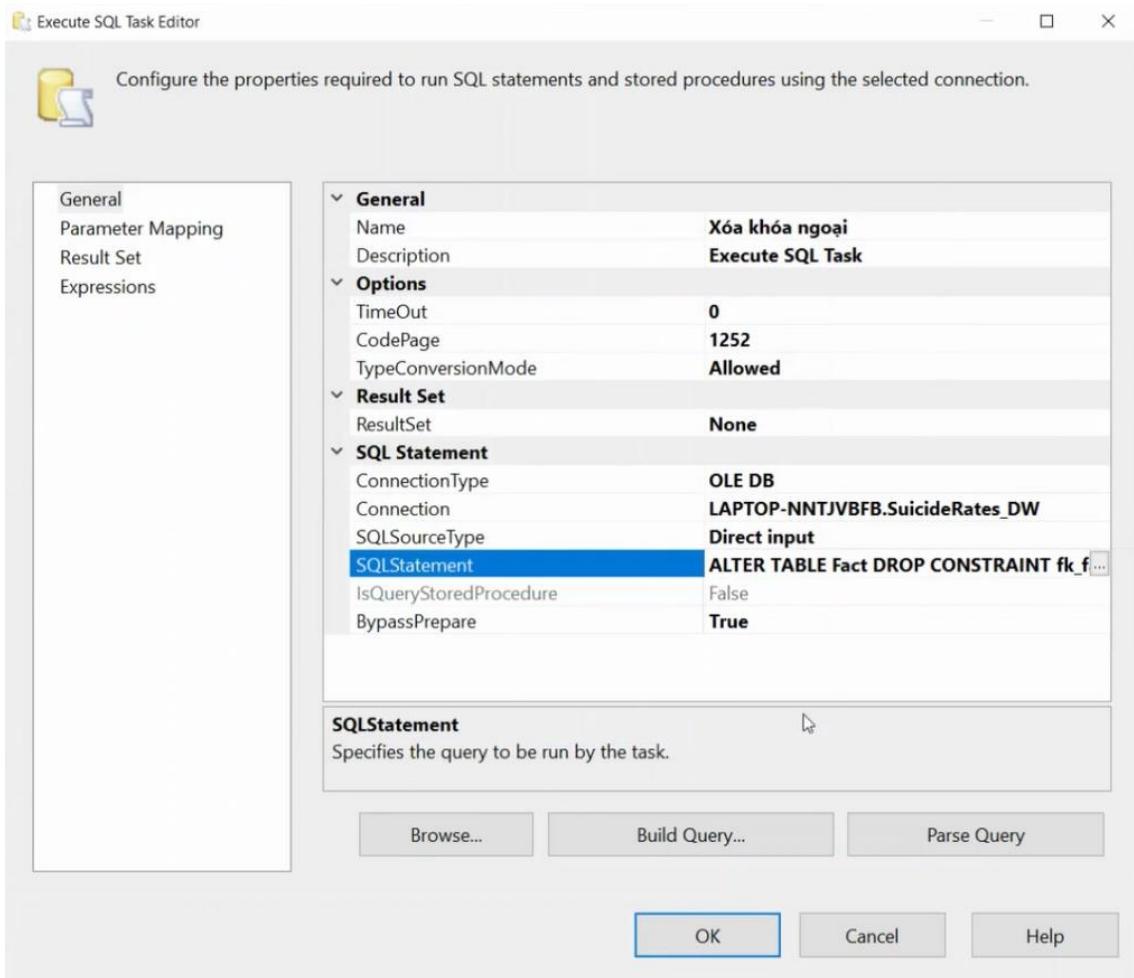


- Kéo thả ExecuteSQL Task và đổi tên thành xóa khóa ngoại và xóa các bảng Dimension, chuột phải chọn Edit.

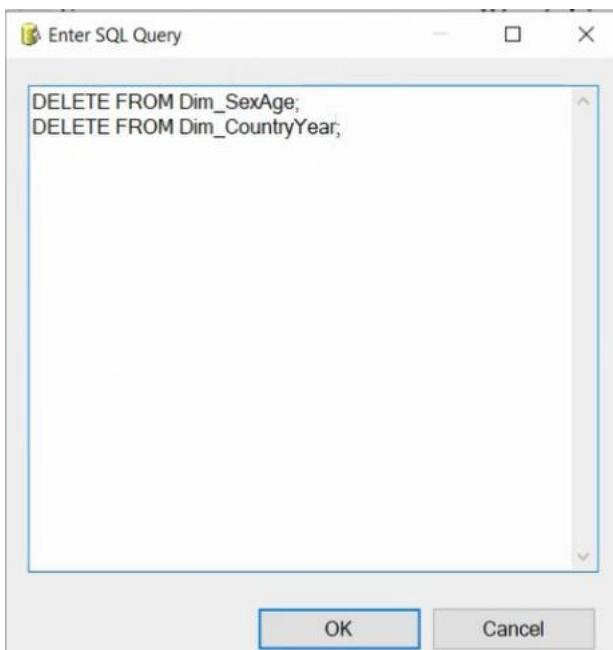


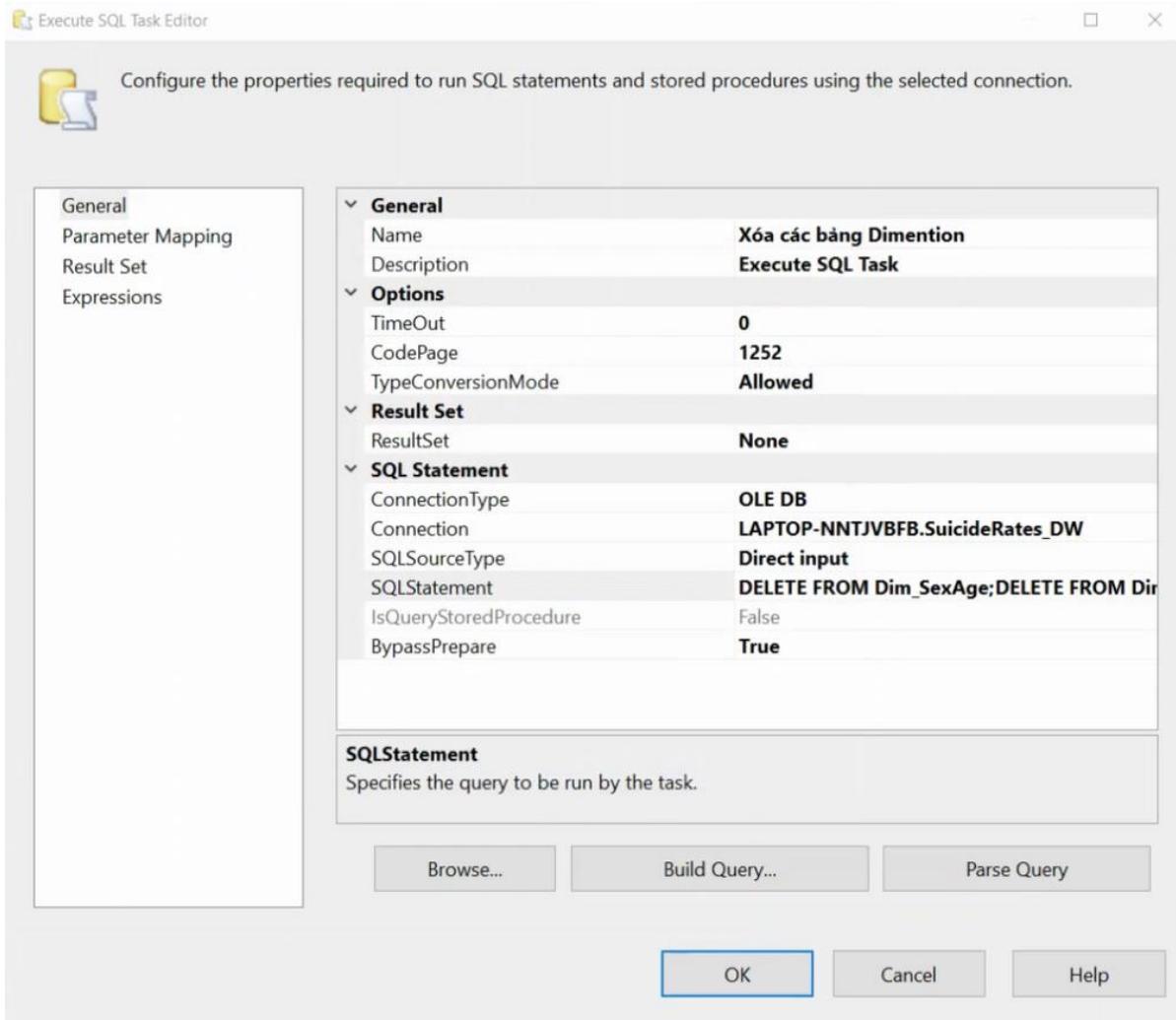
- Chuột phải Xóa khóa ngoại chọn Edit



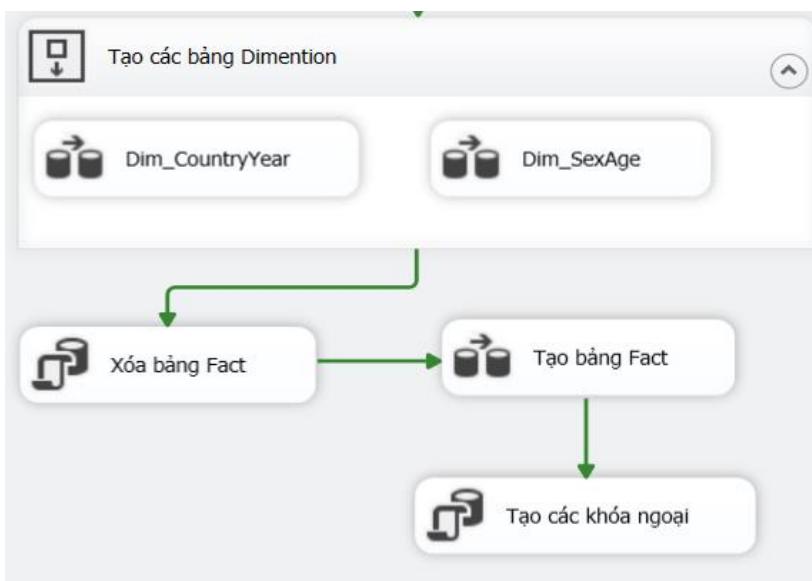


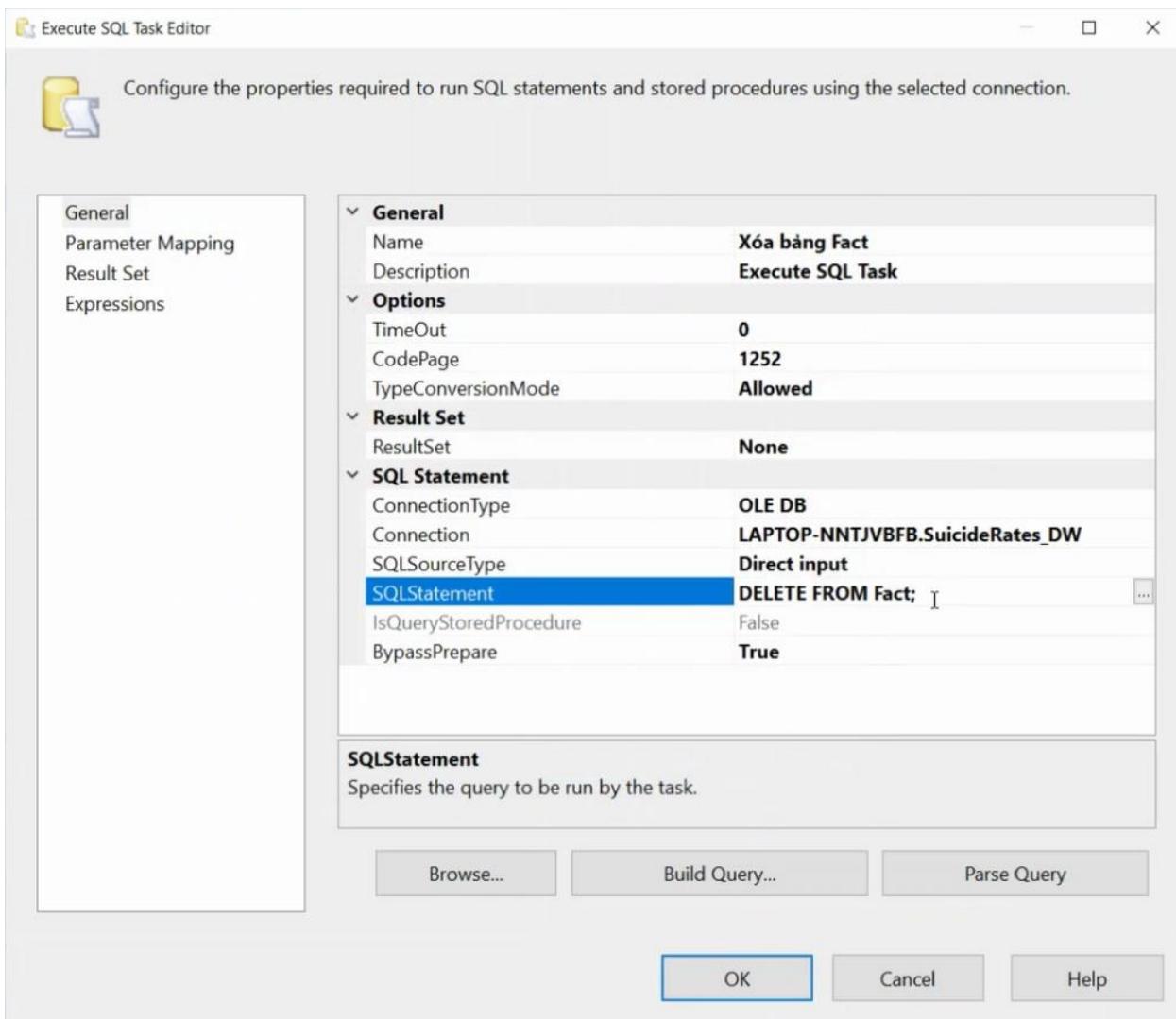
- Chuột phải Xóa các Dimension chọn Edit



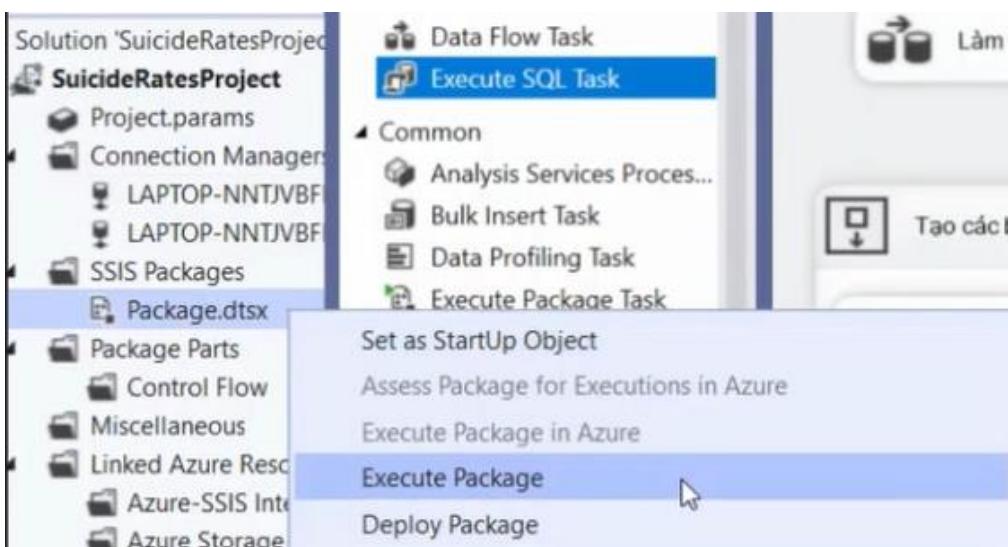


- Kéo thả ExecuteSQL Task và đổi tên thành xóa bảng Fact, tạo liên kết với các bảng Dimention và bảng Fact, chuột phải chọn Edit.

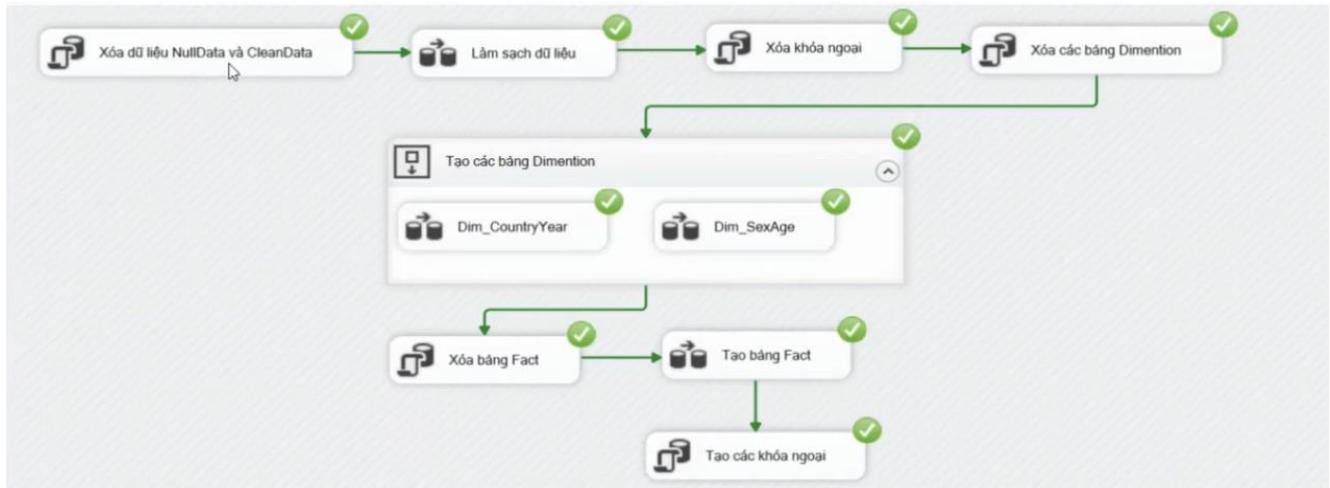




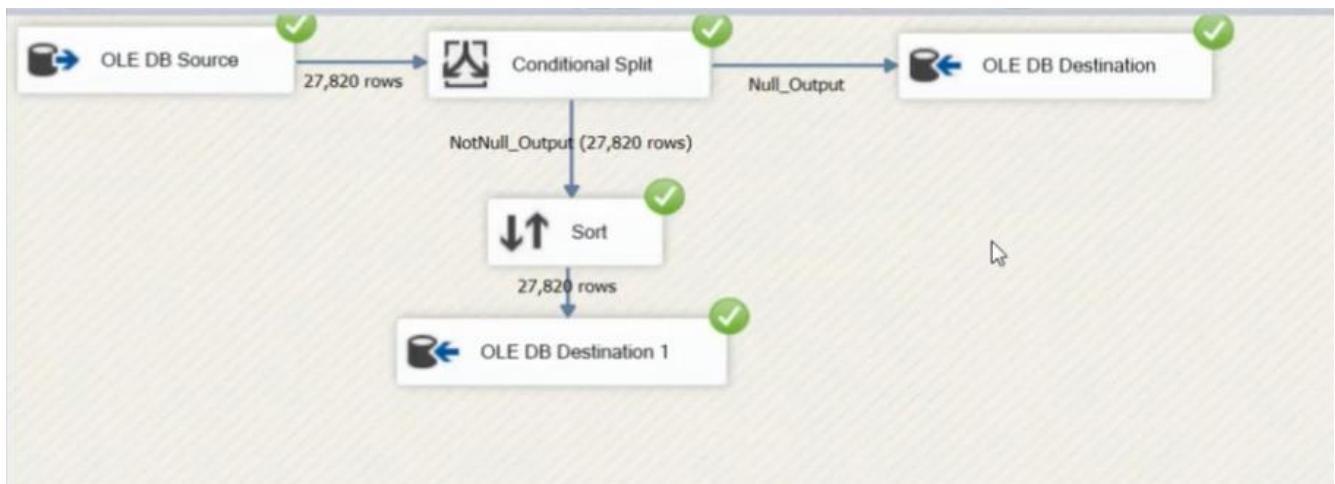
- Thực thi Package



- Control Flow:



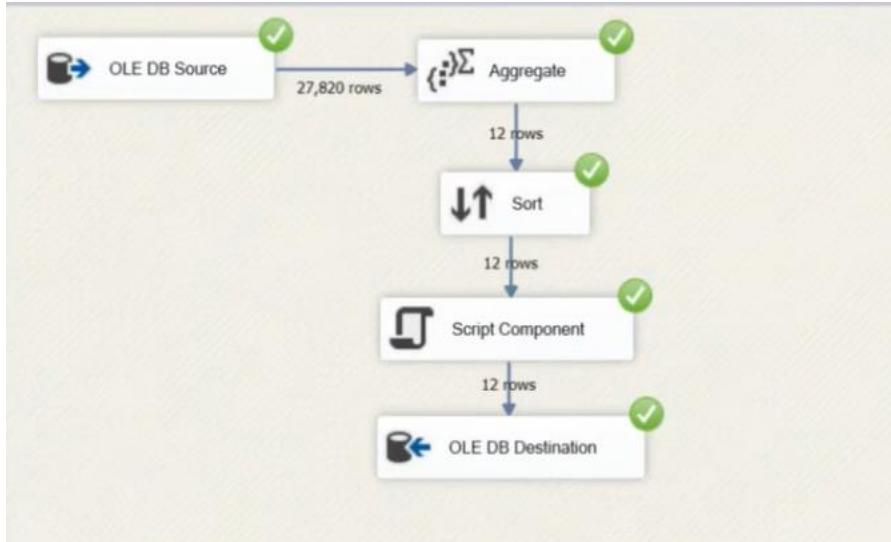
- Data Flow quá trình làm sạch dữ liệu:



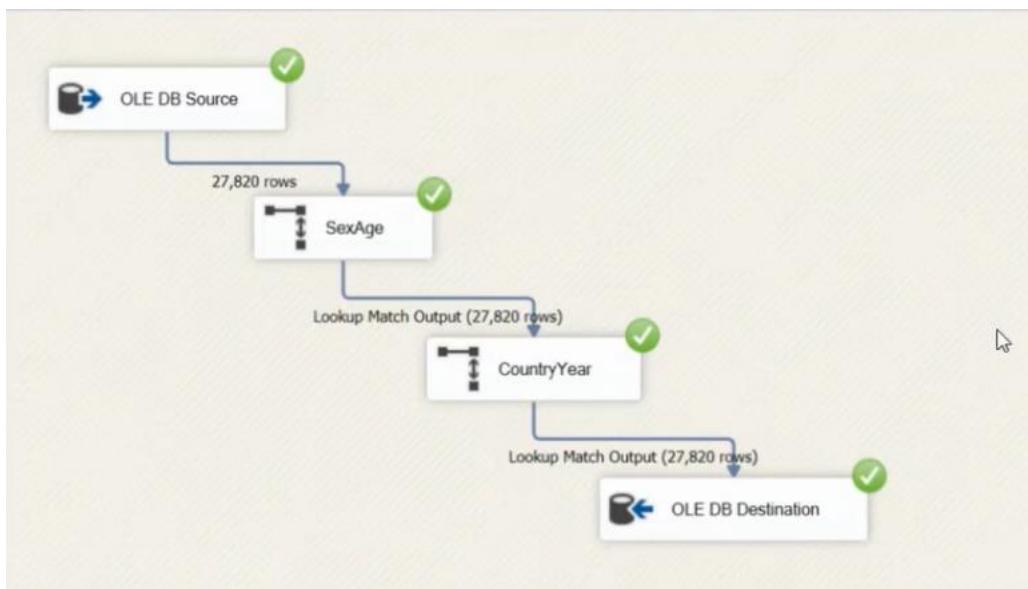
- Data Flow quá trình tạo bảng Dim_CountryYear:



- Data Flow quá trình tạo bảng Dim_SexAge:



- Data Flow quá trình tạo bảng Fact:

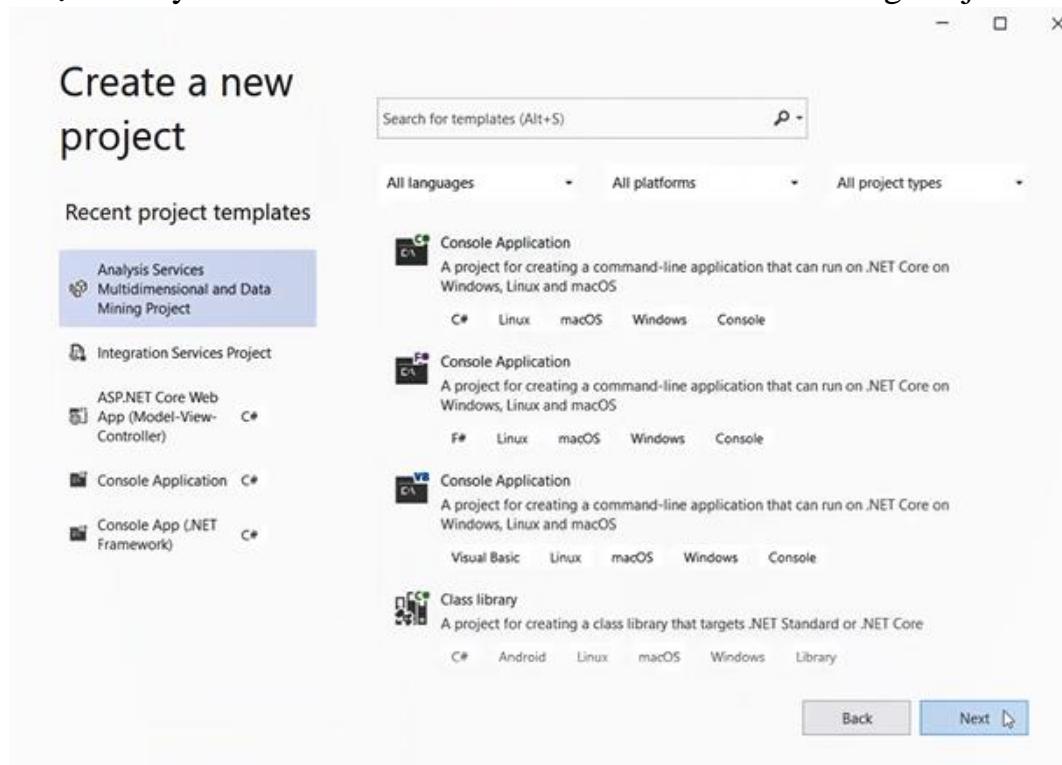


5 Quá trình phân tích kho dữ liệu SSAS và ngôn ngữ truy vấn MDX

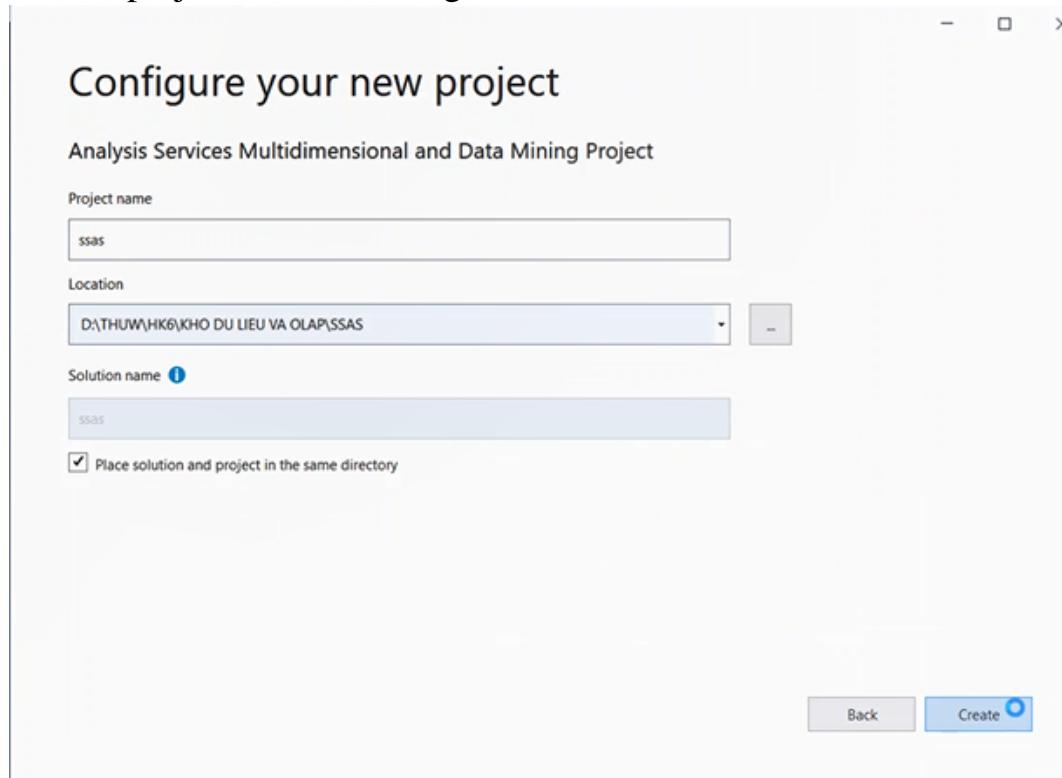
5.1. Tạo project SSAS

Mở Visual Studio (phiên bản 2019) và tạo mới project.

Chọn Analysis Services Multidimensional and Data Mining Project.



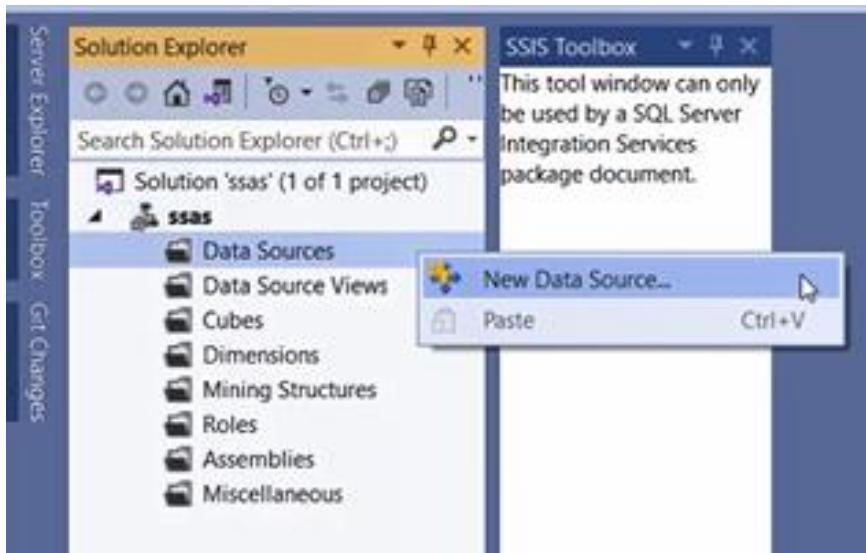
Đặt tên project và chọn đường dẫn lưu trữ.



5.2. Tạo Data Source

Tại phần Solution Explorer nhấn chuột phải chọn Data Source.

Chọn New Data Source để tiến hành tạo một Data Source.



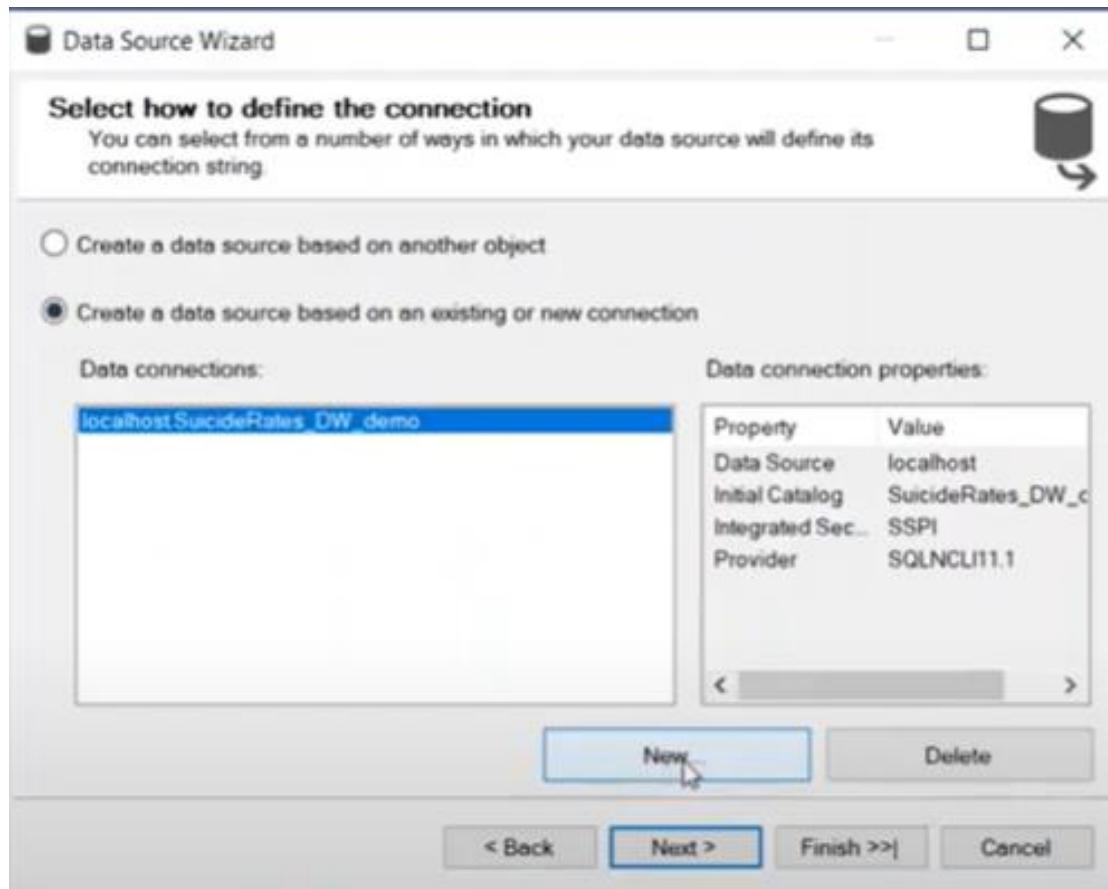
Cửa sổ Data Source Wizard hiện ra nhấn Next để tiếp tục.



Cửa sổ Select how to define the connection được hiển thị.

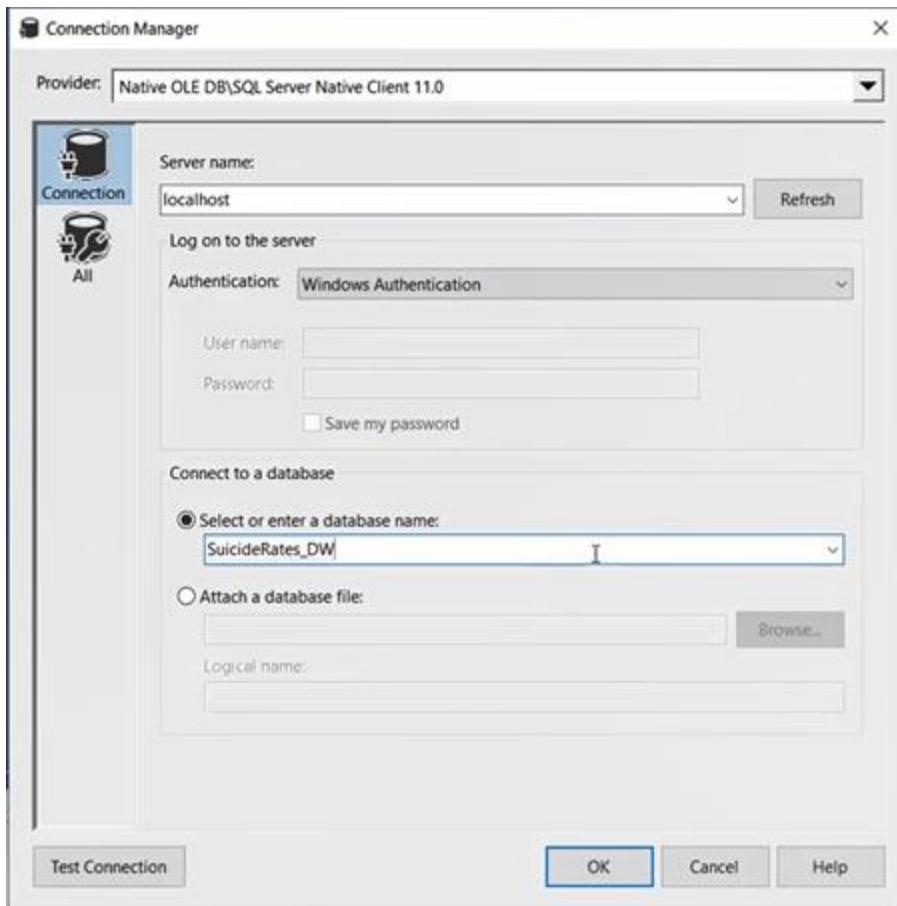
Tại đây chọn Create a data source based on an existing or new connection để tạo 1 data source dựa trên 1 kết nối có sẵn.

Chọn New để tạo kết nối mới

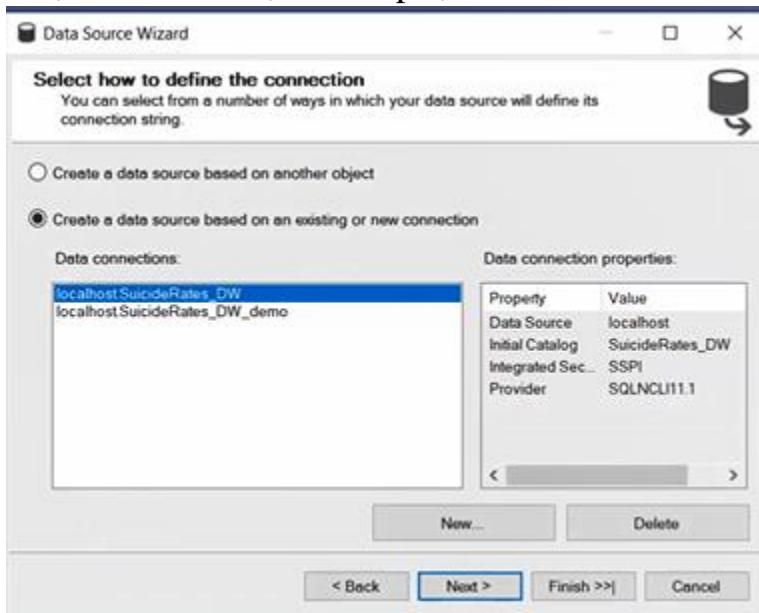


Cửa sổ Connection manager hiển thị.

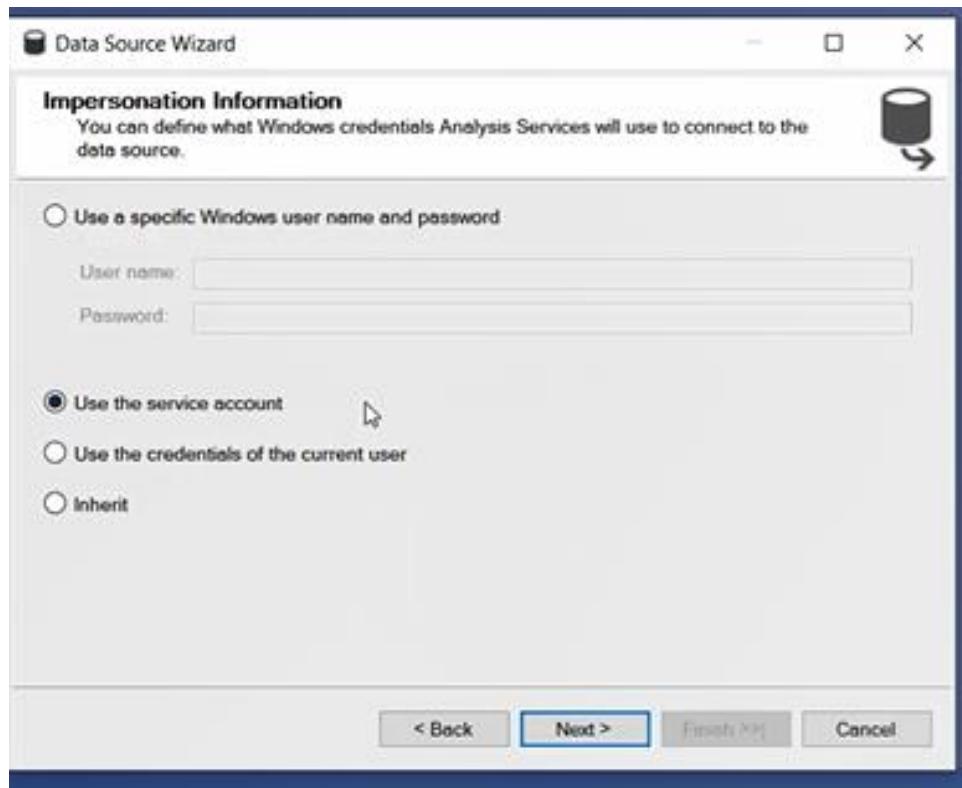
Chọn database cần kết nối, click OK.



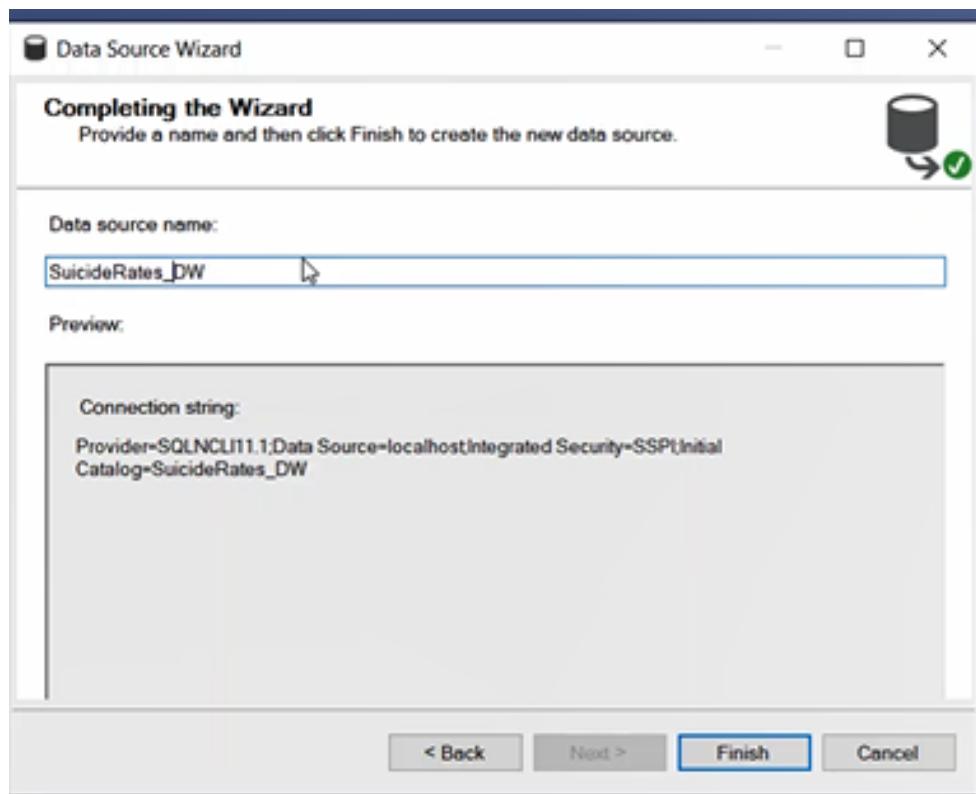
Chọn kết nối đã tạo rồi tiếp tục nhấn Next.



Chọn Use the service account, sau đó chọn Next

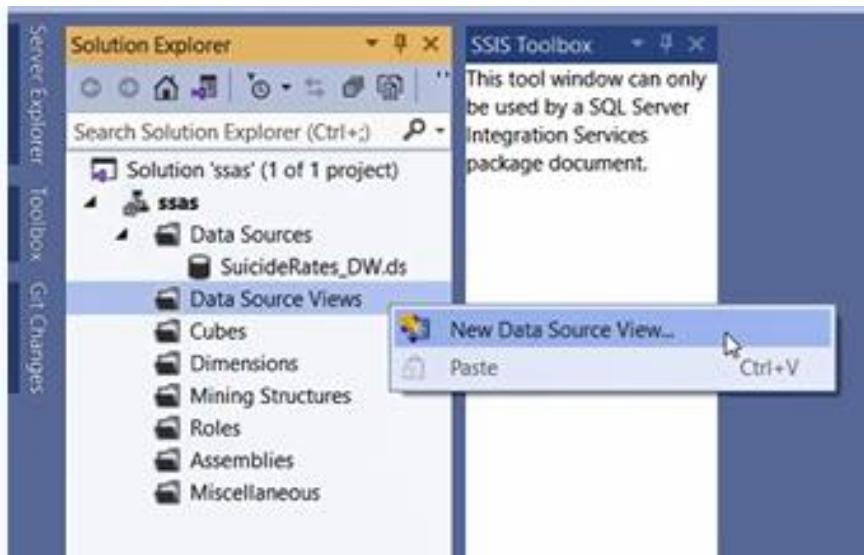


Nhấn Finish để kết thúc



5.3. Tạo Data Source View

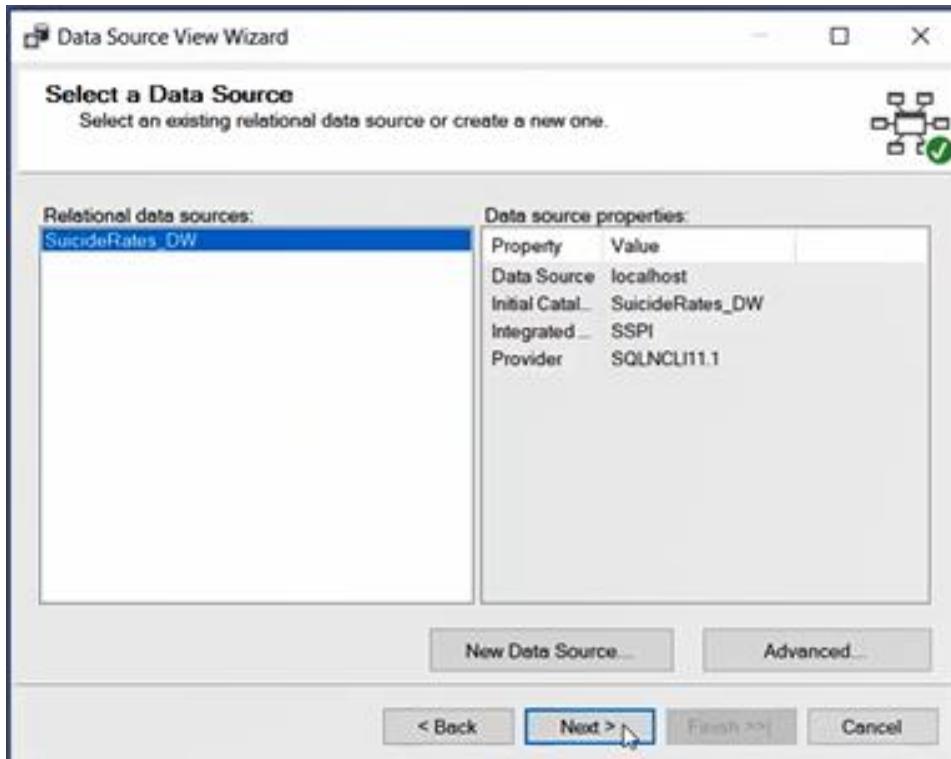
Tại phần Solution Explorer nhấn chuột phải chọn Data Source View.
Chọn New Data Source View...



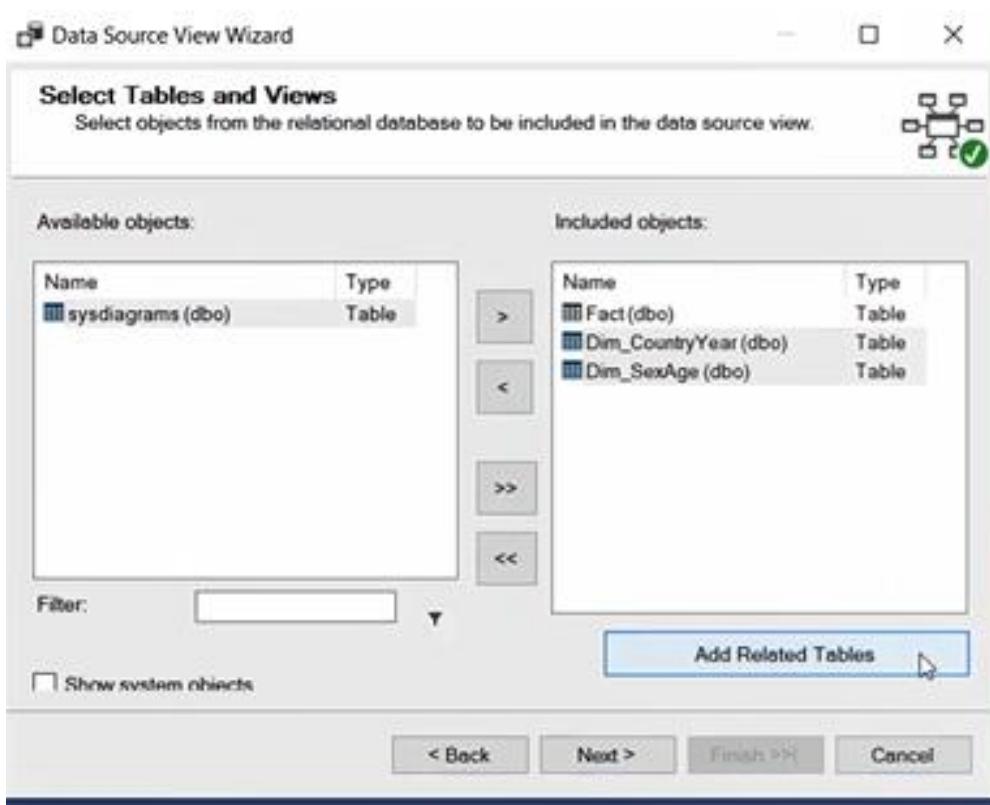
Cửa sổ Data Source View Wizard hiện ra nhấn Next để tiếp tục.



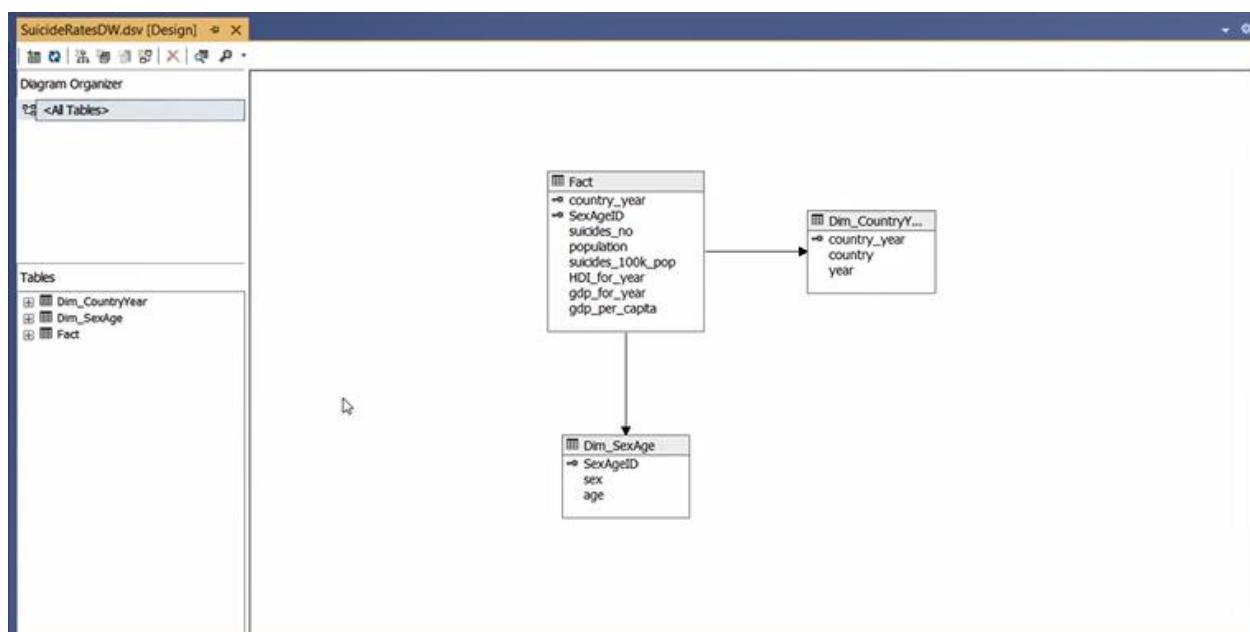
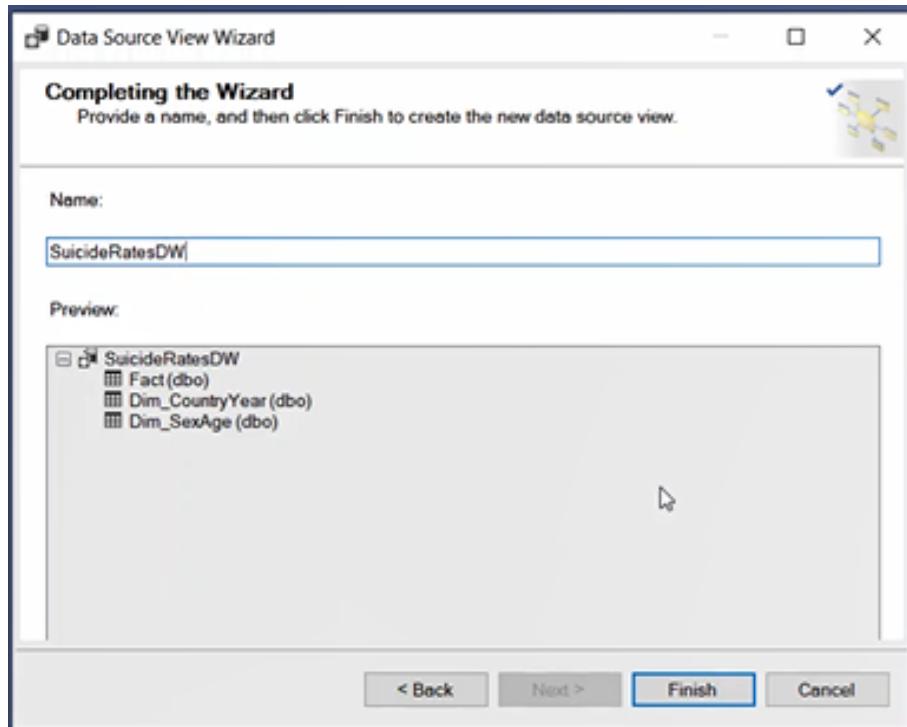
Chọn kho dữ liệu SuicideRates_DW rồi nhấn Next để tiếp tục



Chọn bảng Fact và các bảng Dim có ở Available objects cho quá trình phân tích. Rồi nhấn Next để tiếp tục



Nhấn Finish để kết thúc quá trình tạo Data Source View

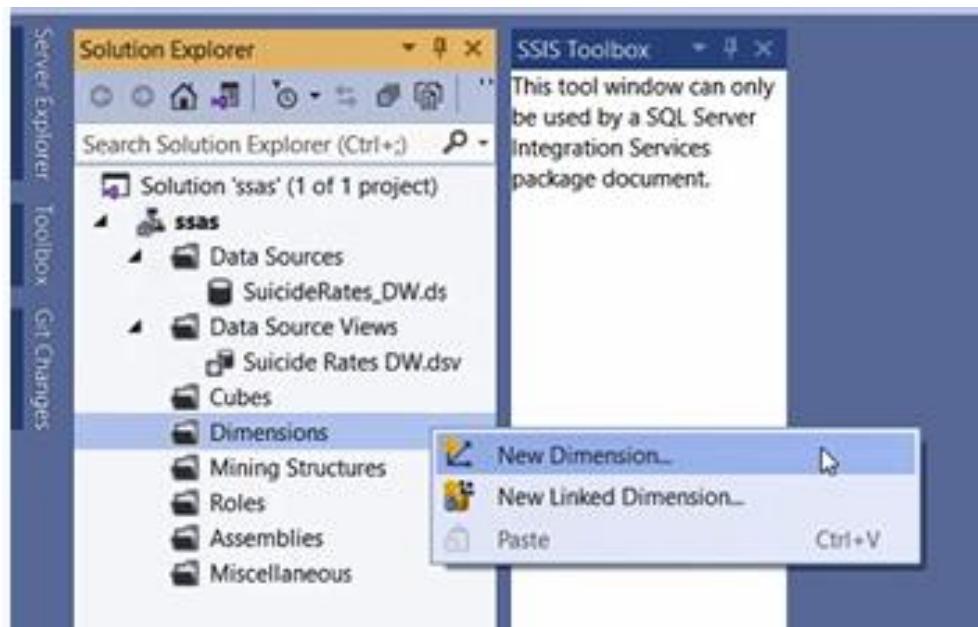


5.4. Tạo các khối (Cube) và xác định các chiều (Dimentions)

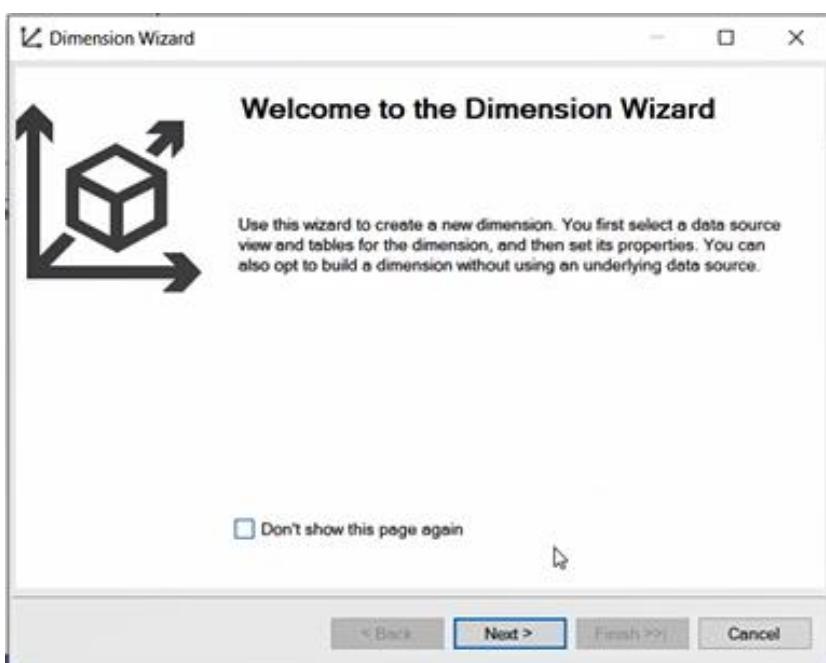
5.4.1. Xác định các chiều (Dimentions)

Tại phần Solution Explorer nhấn chuột phải chọn Dimentions.

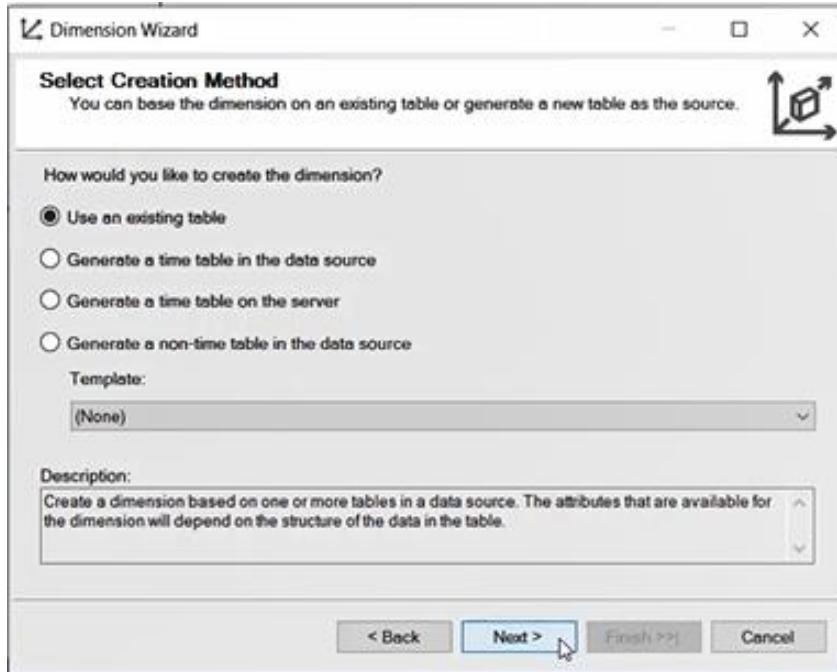
Chọn New Dimension...



Cửa sổ mới hiện ra nhấn Next để tiếp tục.

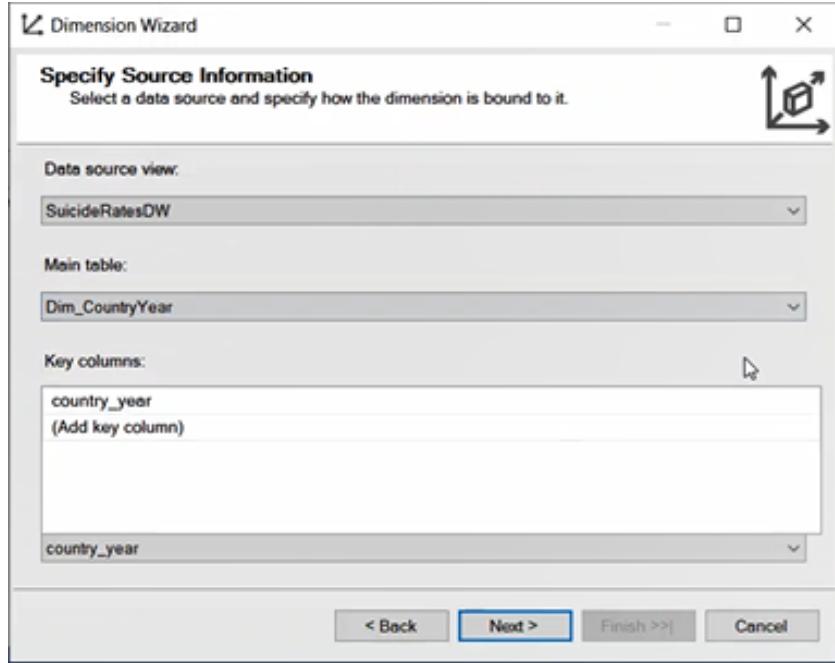


Trong cửa sổ Cube Wizard chọn Use an existing table sau đó nhấn Next để tiếp tục.

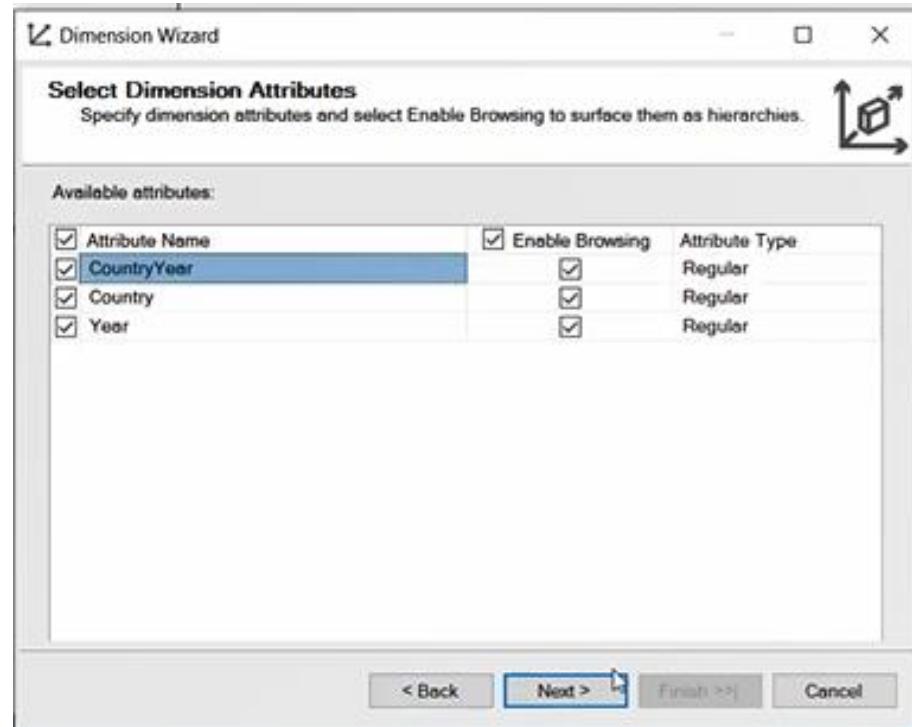


Tiến hành tạo các bảng chiều Dim_CountryYear, Dim_SexAge

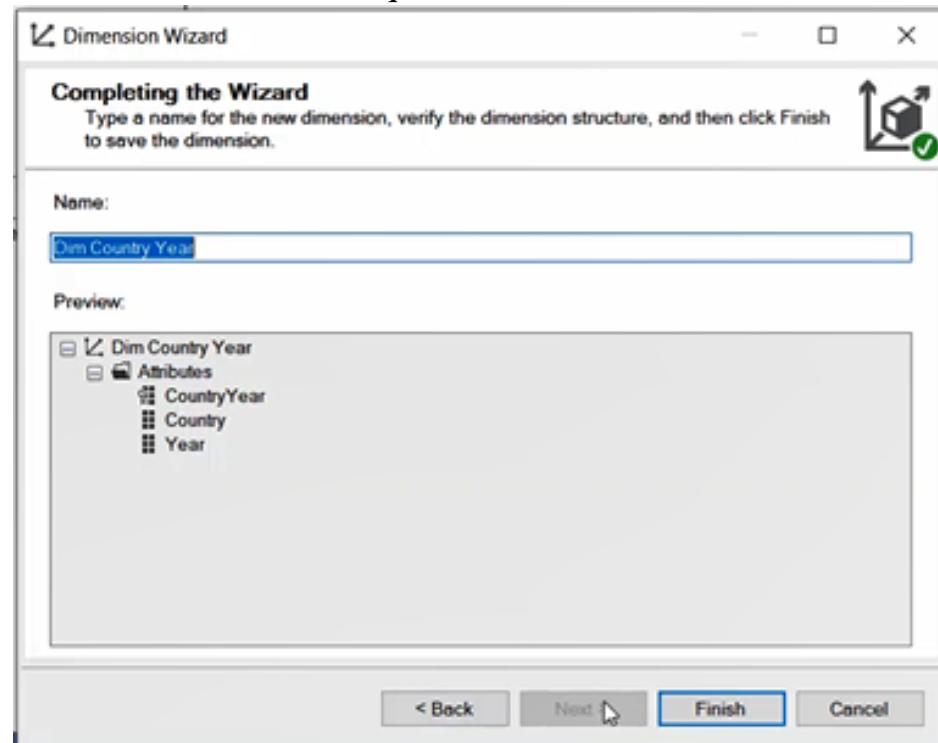
❖ Dim_CountryYear



Chọn các thuộc tính



Nhấn Finish để hoàn tất quá trình



Chiều Dim_CountryYear đã được tạo

The screenshot shows the SSAS Dimension Designer interface. The title bar indicates the current window is 'Dim CountryYear.dim [Design]'. The left pane, titled 'Attributes', lists the attributes of the dimension: Dim_CountryYear, which contains Country, CountryYear, and Year. The middle pane, titled 'Hierarchies', has a placeholder message: 'To create a new hierarchy, drag an attribute here.' The right pane, titled 'Data Source View', shows a single item: 'Dim_CountryY...' with attributes 'country_year', 'country', and 'year'.

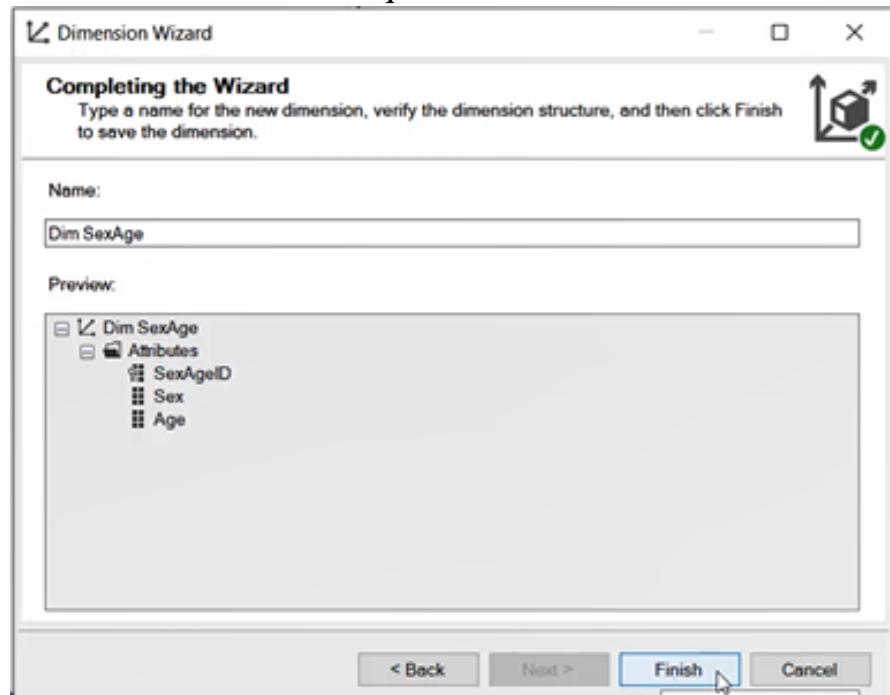
❖ Dim_SexAge

The screenshot shows the 'Specify Source Information' step of the Dimension Wizard. It asks to select a data source and specify how the dimension is bound to it. The 'Data source view' dropdown is set to 'SuicideRatesDW'. The 'Main table' dropdown is set to 'Dim_SexAge'. Under 'Key columns', 'SexAgeID' is selected. At the bottom are buttons for '< Back', 'Next >', 'Finish >>', and 'Cancel'.

Chọn các thuộc tính

The screenshot shows the 'Select Dimension Attributes' step of the Dimension Wizard. It asks to specify dimension attributes and select 'Enable Browsing' to surface them as hierarchies. In the 'Available attributes' grid, four attributes are listed: SexAgeID, Sex, and Age, all with the 'Enable Browsing' checkbox checked. The 'Attribute Type' column shows 'Regular' for all. At the bottom are buttons for '< Back', 'Next >', 'Finish >>', and 'Cancel', with a note: 'Moves to the next wizard page'.

Nhấn Finish để hoàn tất quá trình



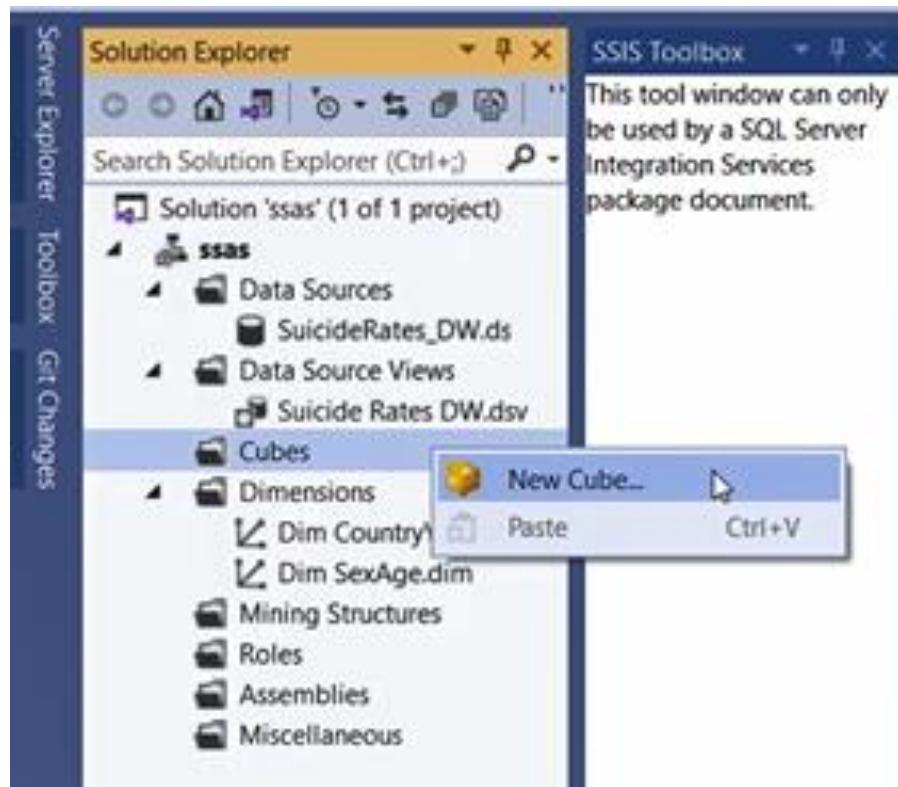
Chiều Dim_SexAge đã được tạo

The screenshot shows the SSAS Dimension Designer interface for the 'Dim SexAge.dim [Design]' tab. The top navigation bar includes tabs for 'Dim CountryYear.dim [Design]', 'SuicideRatesDW.dsv [Design]', and other options like 'Dimension Structure...', 'Attribute Relationships', 'Translations', and 'Browser'. The main area is divided into three panes: 'Attributes' (listing SexAgeID, Sex, and Age), 'Hierarchies' (with a note to drag attributes to create hierarchies), and 'Data Source View' (showing the dimension structure: Dim_SexAge with attributes SexAgeID, sex, and age).

5.4.2. Tạo các khối (Cube)

Tại phần Solution Explorer nhán chuột phải chọn Cube.

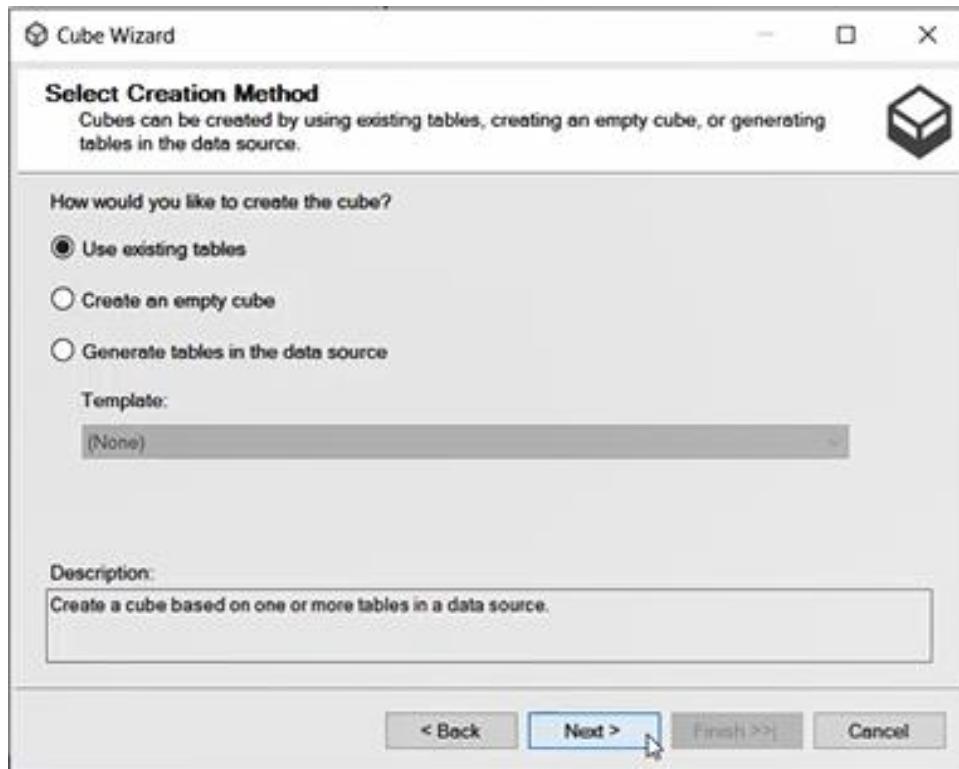
Chọn New Cube...



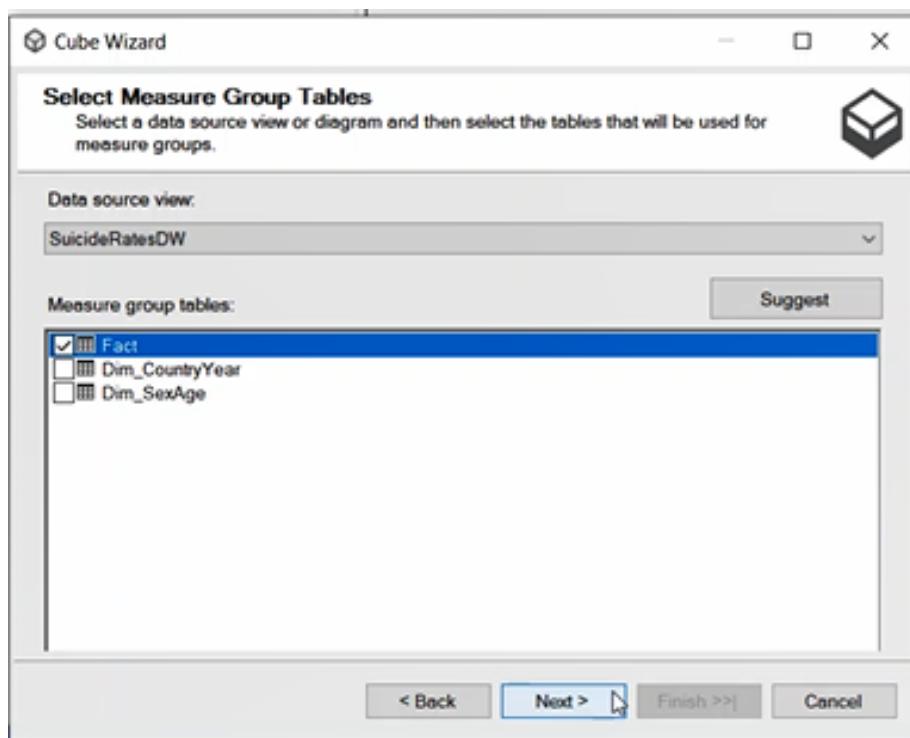
Cửa sổ mới hiện ra nhán Next để tiếp tục.



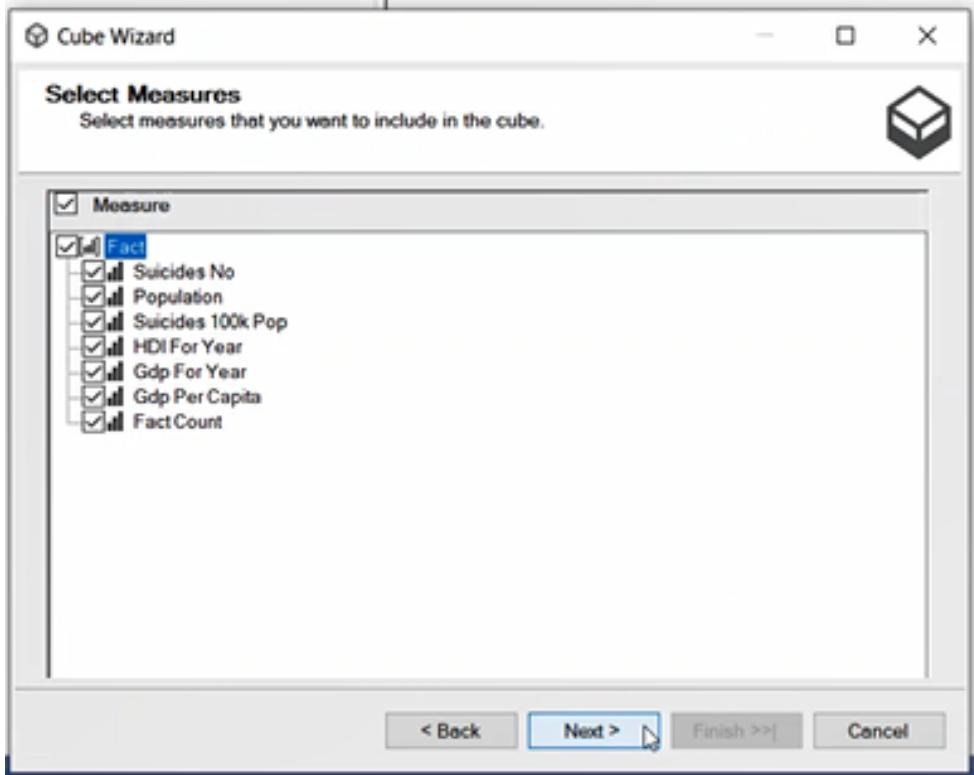
Trong cửa sổ Cube Wizard chọn Use an existing table sau đó nhấn Next để tiếp tục.



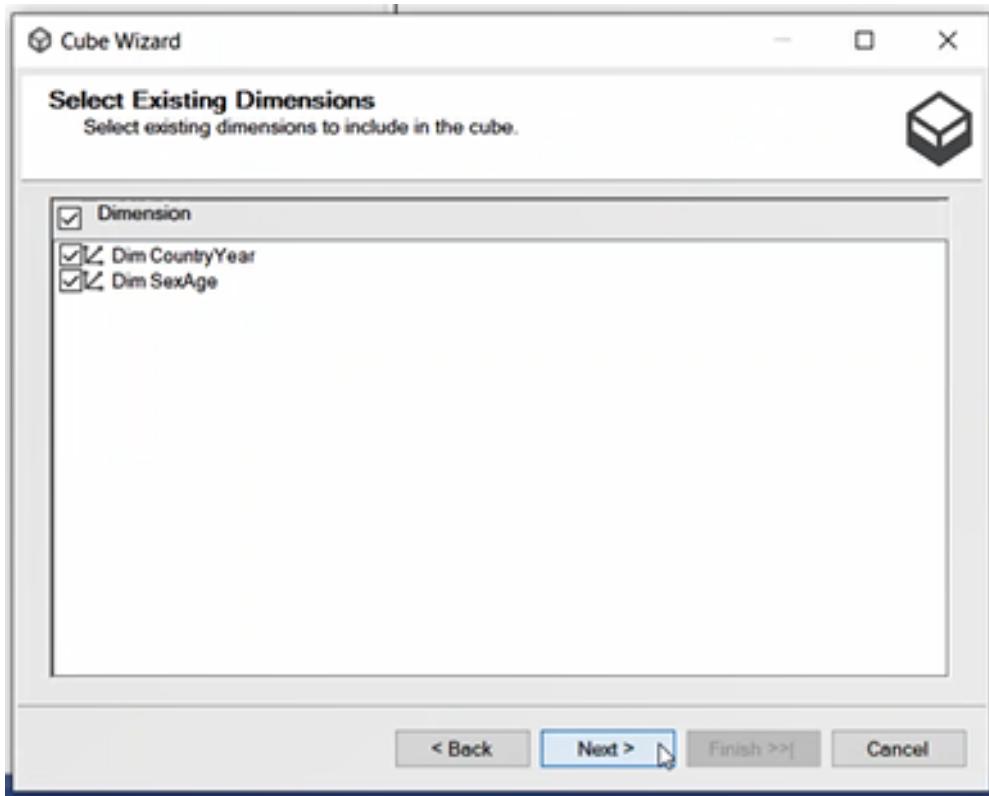
Chọn bảng Fact để phân chia các measure group



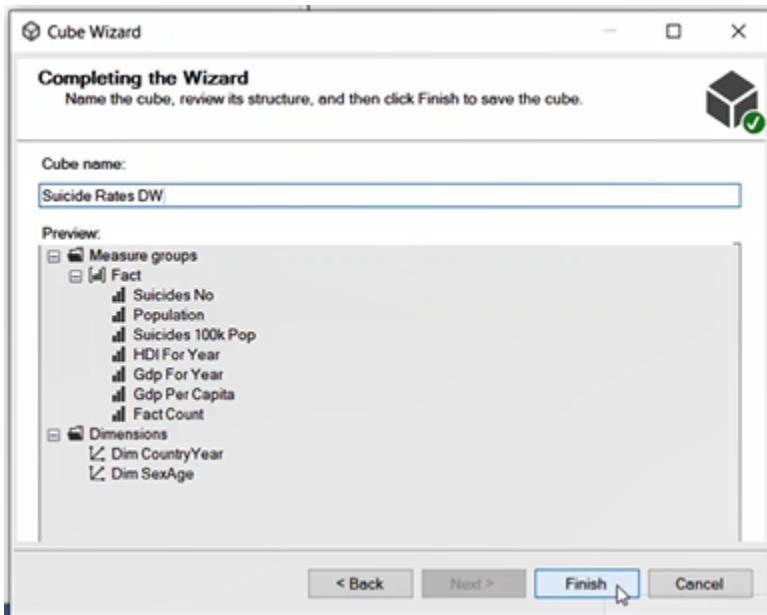
Chọn những thuộc tính định lượng cần thiết. Nhấn Next để tiếp tục.



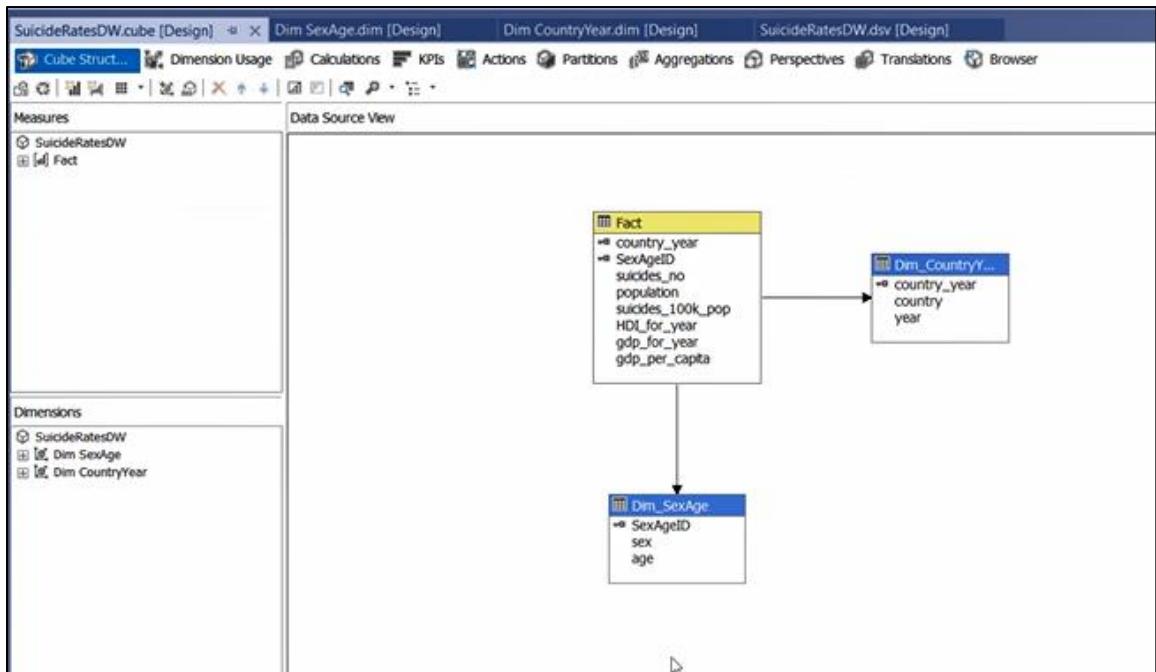
Chọn các Dimension cần thiết, nhấn Next để tiếp tục.



Nhấn Finish để hoàn tất quá trình



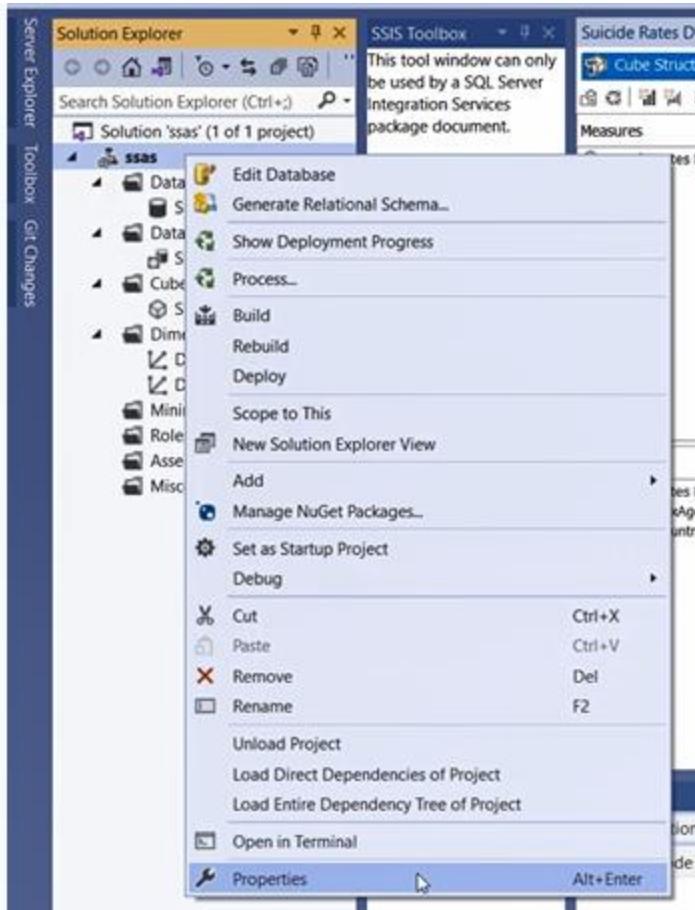
Hoàn tất quá trình tạo Cube



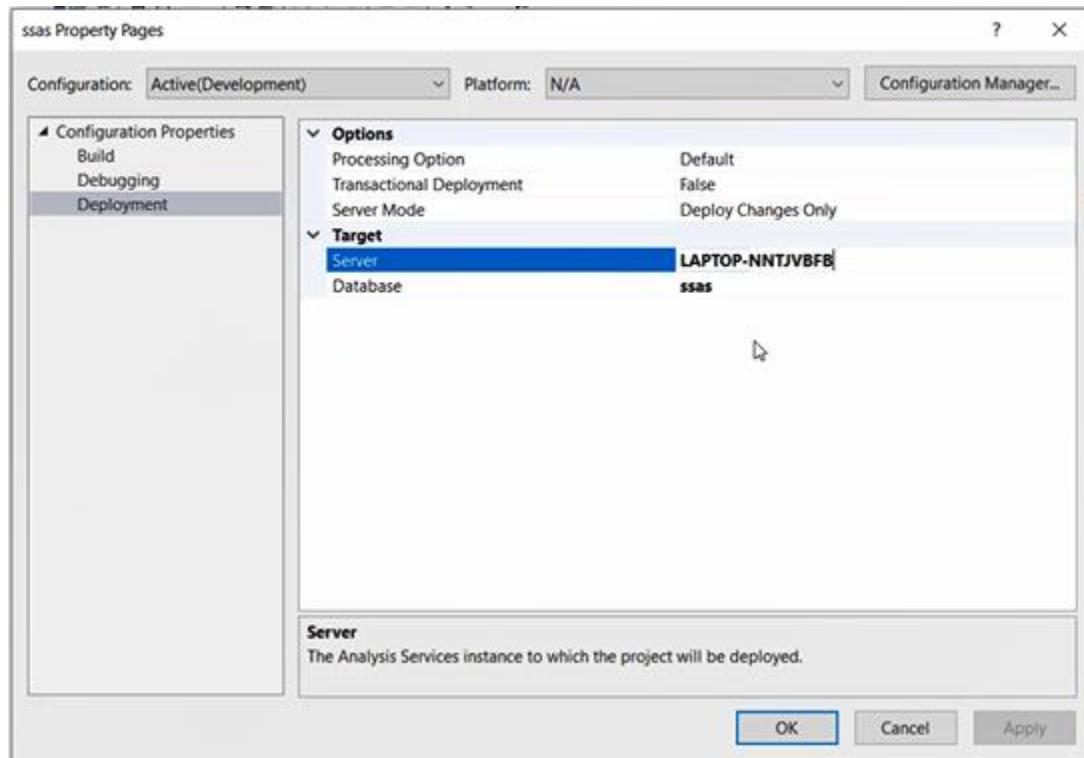
5.4.3. Deploy SSAS Project

Tại phần Solution Explorer nhấn chuột phải chọn ssas_SucideRate.

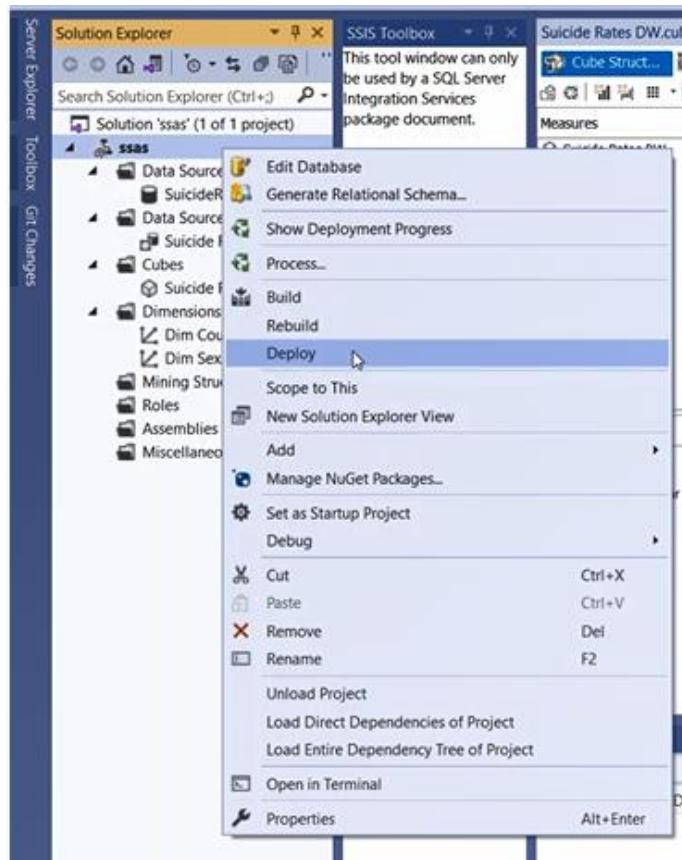
Chọn Properties



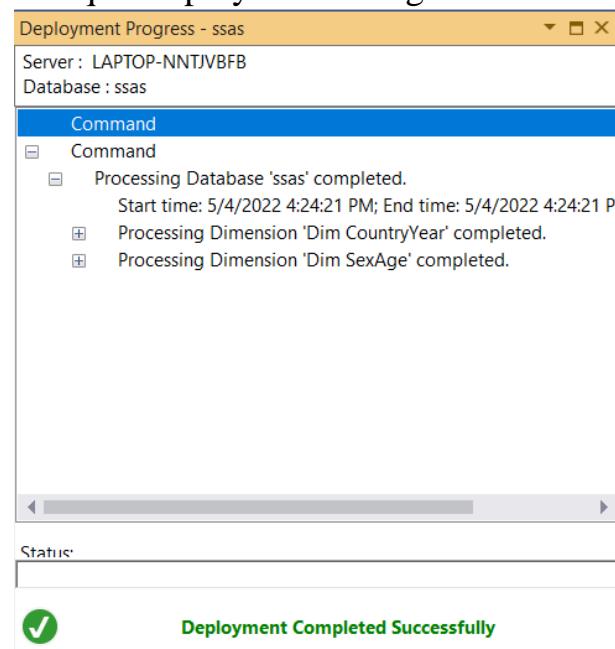
Tại Deployment đổi tên Server tương ứng để kết nối đến SQL Server.



Tại phần Solution Explorer nhấn chuột phải chọn ssas_SucideRate.
Chọn Deploy

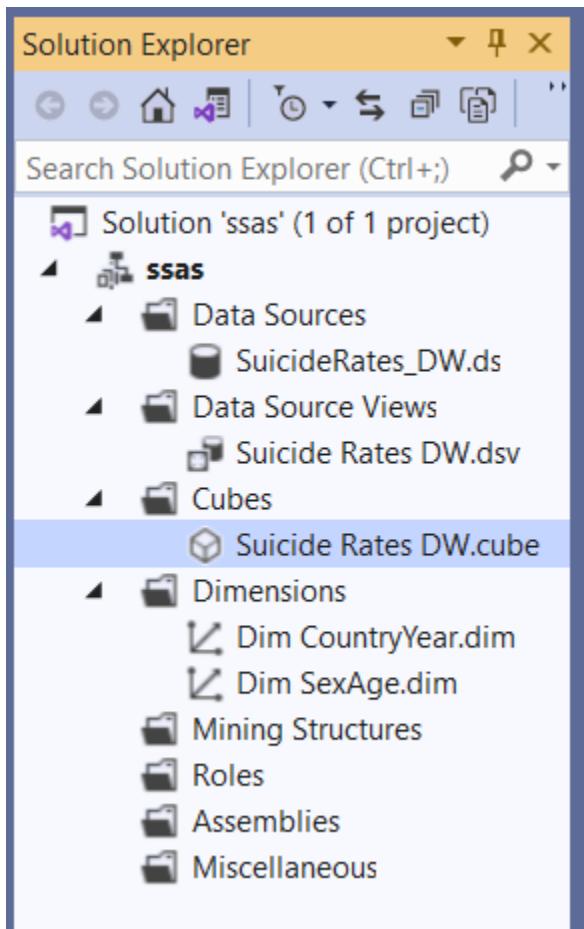


Kết quả Deploy thành công

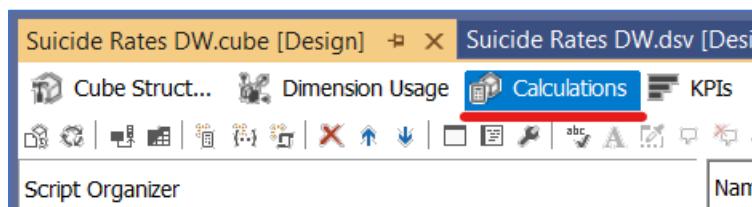


5.5. Định nghĩa Calculation

Chọn ‘Suicide Rates DW.cube’ tại Solution Explorer



Chọn tab Calculation



Nháy chuột phải tại vùng Script Organizer → New Calculated Member

- Tạo Calculation [Suicide_Rates] – tỉ lệ tự tử
Định nghĩa: Suicide_Rates: Cho biết số người tự tử tính trên 100 000 dân số.

$$\text{Suicide_Rates} = \frac{\text{Total Suicides No}}{\text{Total Population}} \times 100000$$

Lấy thuộc tính

[Measures].[Suicides No] / [Measures].[Population] * 100000

Name:

Parent Properties

Parent hierarchy:

Parent member:

Change

Expression

```
[Measures].[Suicides No]/[Measures].[Population]*100000
No issues found
```

Ln: 1 Ch: 56 SPC CRLF

Additional Properties

Format string:

Visible:

Non-empty behavior:

Associated measure group:

Display folder:

Color Expressions

Font Expressions

5.6. Thực hiện câu truy vấn

1. Thống kê tổng số vụ tự tử theo từng nhóm tuổi.

a. Phân tích dữ liệu trên Cube

Age	Suicides No
15-24 years	808542
25-34 years	1123912
35-54 years	2452141
5-14 years	52264
55-74 years	1658443
75+ years	653118

b. Pivot table trong Excel

A B C

Row Labels	Suicides No
15-24 years	808542
25-34 years	1123912
35-54 years	2452141
5-14 years	52264
55-74 years	1658443
75+ years	653118
Grand Total	6748420

PivotTable Fields

Choose fields to add to report:

Search

Σ Fact

- Fact Count
- Gdp For Year
- Gdp Per Capita
- HDI For Year
- Population
- Suicides 100k Pop
- Suicides No**

Drag fields between areas below:

Filters Columns

Rows Values

Age Suicides No

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX

```
SELECT [Measures].[Suicides No] on Columns, [Dim SexAge].[Age].Members on Rows
FROM [Suicide Rates DW];
```

Kết quả sau khi chạy câu truy vấn:

--1.

```
SELECT [Measures].[Suicides No] on Columns, [Dim SexAge].[Age].Members on Rows
FROM [Suicide Rates DW];
```

100 %

	Suicides No
All	6748420
15-24 years	808542
25-34 years	1123912
35-54 years	2452141
5-14 years	52264
55-74 years	1658443
75+ years	653118

2. Thông tin quốc gia có tổng số vụ tự tử cao nhất và thấp nhất trong năm 2016.

a. Phân tích dữ liệu trên Cube

Tạo Nameset [cau2]

Expression: UNION (
 {topcount ([Dim CountryYear].[Country].Children, 1,[Measures].[Suicides No])},
 {bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicides No])}
)

The screenshot shows the 'Script Organizer' window in SSAS Management Studio. A new command named '[cau2]' has been created under the 'CALCULATE' category. The expression for this nameset is defined as:

```
UNION (
    {topcount ([Dim CountryYear].[Country].Children, 1,[Measures].[Suicides No])},
    {bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicides No])}
```

Kéo thả Nameset [cau2] sang trường Filter ta thu được kết quả:

The screenshot shows the 'Edit as Text' MDX editor with the following query:

```
SELECT * FROM [Suicide Rates DW]
```

The 'Filter' pane on the left shows the application of the [cau2] Nameset as a filter on the Dim CountryYear dimension:

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim CountryYear	Year	Equal	{ 2016 }	<input type="checkbox"/> <input type="checkbox"/>
Dim CountryYear	Country	In	cau2	<input type="checkbox"/> <input type="checkbox"/>

The results pane displays the data for the countries included in the [cau2] Nameset:

Country	Suicides No
Grenada	0
Thailand	4117

b. Pivot table trong Excel

A screenshot of a Microsoft Excel spreadsheet titled "tmp7504". The PivotTable Fields pane is open on the right side. The table in the spreadsheet has columns A, B, C, D, and E. Row 1 contains "Year" in A1 and "2016" in B1. Row 2 contains an empty cell in A2 and "Cau2" in B2. Row 3 contains "Row Labels" in A3 and "Suicides No" in B3. Row 4 contains "Grenada" in A4 and "0" in B4. Row 5 contains "Thailand" in A5 and "4117" in B5. The PivotTable Fields pane shows the following settings:

- Choose fields to add to report:**
 - Suicides No
 - Dim CountryYear
 - Country
 - CountryYear
 - Year
 - Sets
 - cau2
 - topcounttutucaonhat
- Drag fields between areas below:**
 - Filters:** Year
 - Columns:** (empty)
 - Rows:** cau2
 - Values:** Suicides No
- Defer Layout Update **Update**

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX

```

SELECT [Measures].[Suicides No] on Columns,
UNION (
{topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicides No])},
{bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicides No])}
)
on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2016];
  
```

Kết quả sau khi chạy câu truy vấn:

```
--2.
SELECT [Measures].[Suicides No] on Columns,
UNION (
{topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicides No])},
{bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicides No])} )
on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2016];
```

	Suicides No
Thailand	4117
Grenada	0

3. Tìm nhóm tuổi có tỉ lệ người tự tử cao nhất ở United States qua các năm.

a. Phân tích dữ liệu trên Cube

Tại Dim SexAge thuộc tính Age ta chọn Operator là Custom, tại trường Filter Expression ta lọc như sau:

```
Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember * topcount ([Dim
SexAge].[Age].Children,1,[Measures].[Suicide_Rates]))
```

The screenshot shows the Expression Builder dialog box. The 'Expression' field contains the generated DAX code. The 'Filter Expression' dropdown menu is open, showing the filter '{ United States }'. The 'OK' button at the bottom right of the dialog is highlighted.

Ta thu được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim CountryYear	Country	Equal	{ United States }	<input type="checkbox"/> <input type="checkbox"/>
Dim SexAge	Age	Custom	Generate ([Dim CountryYear].[Year].Children,[Dim Country...]	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>

Year	Age	Suicide_Rates
1999	75+ years	18.4906337097977
2000	75+ years	18.1557876211382
2001	75+ years	17.4405054059382
2002	75+ years	17.8038611236504
2003	75+ years	16.5679500088129
2004	75+ years	16.3708133355445
2005	75+ years	16.8592254673476
2006	35-54 years	16.1216631308732
2007	35-54 years	16.7471174184363
2008	35-54 years	17.4512696578686
2009	35-54 years	17.8117907805625
2010	35-54 years	17.9757304366901
2011	35-54 years	18.1737099326103
2012	35-54 years	18.4710595805076
2013	35-54 years	18.0469659561449
2014	35-54 years	18.5082286378792
2015	35-54 years	18.8568747817566

b. Pivot table trong Excel

The screenshot shows an Excel spreadsheet with a PivotTable and its corresponding PivotTable Fields pane.

PivotTable Fields pane:

- Choose fields to add to report:**
 - Search:** Search bar.
 - Sets:** cau11, cau14, **cau3** (selected), cau7, cau9.
 - Dim CountryYear:** **Country** (selected).
- Drag fields between areas below:**
 - Filters:** Country (selected).
 - Columns:** None.
 - Rows:** cau3 (selected).
 - Values:** Suicide_Rates (selected).
- Defer Layout Update:** Checkboxes for Defer Layout Update and Update.

PivotTable Data:

	A	B	C	D
1	Country	United States		
2				
3	Row Labels	Suicide_Rates		
4	1985			
5	75+ years	22.92		
6	1986			
7	75+ years	24.15		
8	1987			
9	75+ years	24.95		
10	1988			
11	75+ years	24.59		
12	1989			
13	75+ years	23.06		
14	1990			
15	75+ years	24.28		
16	1991			
17	75+ years	23.63		
18	1992			
19	75+ years	22.60		
20	1993			
21	75+ years	22.41		
22	1994			
23	75+ years	21.76		
24	1995			
25	75+ years	20.96		
26	1996			
27	75+ years	20.07		
28	1997			
29	75+ years	19.68		

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX

```
with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides
No]/[Measures].[Population]*100000', FORMAT_STRING = 'General Number'

SELECT [Measures].[Suicide_Rates] on Columns,
Generate (
[Dim CountryYear].[Year].Children,
[Dim CountryYear].[Year].Currentmember * topcount ([Dim
SexAge].[Age].Children,1,[Measures].[Suicide_Rates]))
on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Country].&[United States];
```

Kết quả sau khi chạy câu truy vấn:

The screenshot shows the SQL Server Management Studio interface. On the left, the query is displayed:

```
--3. Tìm nhóm tuổi có tỉ lệ người tự tử cao nhất ở United States qua các năm.
with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides No]/[Measures].[Population]*100000', FORMAT_STRING = 'General Number'
SELECT [Measures].[Suicide_Rates] on Columns,
Generate (
[Dim CountryYear].[Year].Children,
[Dim CountryYear].[Year].Currentmember * topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]))
on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Country].&[United States];
```

On the right, the results pane displays a table titled "Results" with two columns: "Year" and "Suicide_Rates". The data is as follows:

Year	Suicide_Rates	
1995	75+ years	20.9568678417675
1996	75+ years	20.070302009005
1997	75+ years	19.6764460422418
1998	75+ years	20.030113895501
1999	75+ years	18.4906337097977
2000	75+ years	18.1557876211382
2001	75+ years	17.4405054059382
2002	75+ years	17.8038611236504
2003	75+ years	16.5679500088129
2004	75+ years	16.3708133355445
2005	75+ years	16.8592254673476
2006	35-54 years	16.1216631308732
2007	35-54 years	16.7471174184363
2008	35-54 years	17.4512696578686
2009	35-54 years	17.8117907805625
2010	35-54 years	17.9757304366901
2011	35-54 years	18.1737099326103
2012	35-54 years	18.4710595805076
2013	35-54 years	18.0469659561449
2014	35-54 years	18.5062286378792
2015	35-54 years	18.8568747817566

4. Liệt kê top 10 quốc gia có tỉ lệ tự tử cao nhất trong năm 2016, sắp xếp theo thứ tự tỉ lệ tự tử giảm dần.

a. Phân tích dữ liệu trên Cube

Tạo Nameset [cau4]

Expression:

```
Head( Order(
[Dim CountryYear].[CountryYear].Children,[Measures].[Suicide_Rates],
desc), 10)
```

Name: [cau4]

Expression

```
Head( Order(
    [Dim CountryYear].[CountryYear].Children.[Measures].[Suicide Rates], desc)
    ✓ No issues found
```

Additional Properties

Type: Dynamic

Display folder:

Kéo thả Nameset [cau4] sang trường Filter ta thu được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim CountryYear	# Year	Equal	{ 2016 }	<input type="checkbox"/> <input type="checkbox"/>
Dim CountryYear	# CountryYear	In	cau4	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>

Country	Suicide_Rates
Austria	16.0462195253766
Croatia	19.1475123793013
Czech Republic	14.7593343831349
Hungary	20.988347355361
Iceland	14.882447269629
Lithuania	33.6155622877558
Mongolia	19.5147696357566
Netherlands	13.2447968524857
Romania	11.7333201961802
Sweden	13.9062340785926

Chuyển Design Mode sang Script Mode trong VS Studio để có thể sắp xếp được dữ liệu tăng dần khi ta dùng Name Sets.

Sử dụng Name Sets [cau4] trong truy vấn MDX:

```
SELECT [Measures].[Suicide_Rates] on Columns,
{cau4} on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2016];
```

```

SELECT [Measures].[Suicide_Rates] on Columns,
{cau4} on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2016];

```

CountryYear	Suicide_Rates
Lithuania2016	33.6155622877558
Hungary2016	20.988347355361
Mongolia2016	19.5147696357566
Croatia2016	19.1475123793013
Austria2016	16.0462195253766
Iceland2016	14.882447269629
Czech Republic2016	14.7593343831349
Sweden2016	13.9062340785926
Netherlands2016	13.2447968524857
Romania2016	11.7333201961802

b. Pivot table trong Excel

The screenshot shows the configuration of a PivotTable in Excel. The PivotTable Fields pane on the right is used to define the data structure:

- Choose fields to add to report:**
 - Row Labels: Suicide_Rates
 - Column Labels: Year (set to 2016)
 - Values: Suicide_Rates (selected)
- Drag fields between areas below:**
 - Filters: Year
 - Columns: None
 - Rows: cau4
 - Values: Suicide_Rates

The resulting PivotTable in the main area shows the following data:

Year	Suicide_Rates
Austria2016	16.05
Croatia2016	19.15
Czech Republic2016	14.76
Hungary2016	20.99
Iceland2016	14.88
Lithuania2016	33.62
Mongolia2016	19.51
Netherlands2016	13.24
Romania2016	11.73
Sweden2016	13.91

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX

```
SELECT [Measures].[Suicide_Rates] on Columns,
Head( Order(
    [Dim CountryYear].[CountryYear].Children,[Measures].[Suicide_Rates],
    desc), 10)
on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2016];
```

Kết quả sau khi chạy câu truy vấn:

The screenshot shows the SQL Server Management Studio interface. In the top pane, there is a comment: '--4. Liệt kê top 10 quốc gia có tỉ lệ tự tử cao nhất trong năm 2016, --sắp xếp theo thứ tự tỉ lệ tự tử giảm dần.' Below it is the MDX query:

```
--4. Liệt kê top 10 quốc gia có tỉ lệ tự tử cao nhất trong năm 2016,
--sắp xếp theo thứ tự tỉ lệ tự tử giảm dần.
SELECT [Measures].[Suicide_Rates] on Columns,
Head( Order(
    [Dim CountryYear].[CountryYear].Children,[Measures].[Suicide_Rates], desc),
, 10)
on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2016];
```

In the bottom pane, under the 'Results' tab, the output is displayed in a table:

	Suicide_Rates
Lithuania2016	33.62
Hungary2016	20.99
Mongolia2016	19.51
Croatia2016	19.15
Austria2016	16.05
Iceland2016	14.88
Czech Republic2016	14.76
Sweden2016	13.91
Netherlands2016	13.24
Romania2016	11.73

5. Liệt kê những quốc gia và năm từng có tổng số vụ tự tử cao hơn 10 000 vụ.

a. Phân tích dữ liệu trên Cube

Tạo Nameset [cau5]

Expression: Filter ([Dim
CountryYear].[CountryYear].Members,[Measures].[Suicides No] > 10000)

Name: [cau5]

Expression

```
Filter ([Dim CountryYear].[CountryYear].Members,[Measures].[Suicides No] > 10000)
No issues found
```

Ln: 1 Ch: 82 SPC CRLF

Additional Properties

Type: Dynamic

Display folder:

Chuyển Design Mode sang Script Mode

Sử dụng Name Sets [cau5] thay thế vào vị trí tô xanh như trên hình :

```
SELECT NON EMPTY { [Measures].[Suicides No] } ON COLUMNS, NON EMPTY { {[Dim CountryYear].[CountryYear].[CountryYear].ALLMEMBERS} } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM ( SELECT ( [cau5] ) ON COLUMNS FROM [Suicide Rates DW] ) CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

CountryYear | Suicides No |

Ta thu được kết quả

```
SELECT NON EMPTY { [Measures].[Suicides No] } ON COLUMNS, NON EMPTY { [cau5] } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM ( SELECT ( [cau5] ) ON COLUMNS FROM [Suicide Rates DW] ) CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

CountryYear	Suicides No
(null)	6748420
Brazil2012	10295
Brazil2013	10513
Brazil2014	10631
Brazil2015	11163
France1985	12501
France1986	12529
France1987	12161
France1988	11599
France1989	11715
France1990	11403
France1991	11502
France1992	11644
France1993	12251
France1994	12041

b. Pivot table trong Excel

	A	B	C	D
1	Row Labels	Suicides No		
2	Brazil2012	10295		
3	Brazil2013	10513		
4	Brazil2014	10631		
5	Brazil2015	11163		
6	France1985	12501		
7	France1986	12529		
8	France1987	12161		
9	France1988	11599		
10	France1989	11715		
11	France1990	11403		
12	France1991	11502		
13	France1992	11644		
14	France1993	12251		
15	France1994	12041		
16	France1995	11819		
17	France1996	11279		
18	France1997	11139		
19	France1998	10534		
20	France1999	10268		
21	France2000	10849		
22	France2001	10449		
23	France2002	10643		
24	France2003	10871		
25	France2004	10804		
26	France2005	10713		
27	France2006	10423		
28	France2007	10127		
29	France2008	10353		

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```
SELECT [Measures].[Suicides No] on Columns,
{Filter ([Dim CountryYear].[CountryYear].Members,[Measures].[Suicides No]
> 10000)} on Rows
FROM [Suicide Rates DW];
```

Kết quả sau khi chạy câu truy vấn:

```
--5. Liệt kê những quốc gia và năm từng có tổng số vụ tự tử cao hơn 10 000 vụ.
SELECT [Measures].[Suicides No] on Columns,
{Filter ([Dim CountryYear].[CountryYear].Members,[Measures].[Suicides No] > 10000)} on Rows
FROM [Suicide Rates DW];
```

100 %

	Suicides No
All	6748420
Brazil2012	10295
Brazil2013	10513
Brazil2014	10631
Brazil2015	11163
France1985	12501
France1986	12529
France1987	12161
France1988	11599
France1989	11715
France1990	11403
France1991	11502
France1992	11644
France1993	12251
France1994	12041
France1995	11819
France1996	11279
France1997	11139

6. Cho biết tên quốc gia, tổng số vụ tự tử, tỉ lệ tự tử của quốc gia có GDP bình quân đầu người cao nhất và thấp nhất trong năm 2016.

- Phân tích dữ liệu trên Cube

Tạo Nameset [cau6]

Expression:

```
UNION (
{topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Gdp Per Capita])},
{bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Gdp Per Capita])} )
```

Name: [cau6]

Expression

```
UNION (
{topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Gdp Per Capita])}.
✓ No issues found
```

Additional Properties

Type: Dynamic

Display folder:

Kéo thả Nameset [cau6] và điều kiện Dimension sang trường Filter ta thu được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim CountryYear	Year	Equal	{ 2016 }	<input type="checkbox"/> <input checked="" type="checkbox"/>
Dim CountryYear	Country	In	cau6	<input checked="" type="checkbox"/> <input type="checkbox"/>
<Select dimension>				
Country	Suicides No	Suicide_Rates	Gdp Per Capita	
Armenia	67	2.79162828...	37880	
Iceland	40	14.8824472...	647080	

b. Pivot table trong Excel

The screenshot shows a Microsoft Excel spreadsheet with a PivotTable Fields pane open on the right side. The PivotTable Fields pane allows users to choose fields to add to the report, including Fact, Filters, Rows, and Columns. In the Fact section, 'Gdp Per Capita' and 'Suicide_Rates' are selected. In the Rows section, 'cau6' is selected. The main worksheet displays data for Armenia and Iceland across four columns: Year (2016), Suicides No, Suicide_Rates, and Gdp Per Capita.

	A	B	C	D
1	Year	2016		
2				
3	Row Labels	Suicides No	Suicide_Rates	Gdp Per Capita
4	Armenia	67	2.79	37880
5	Iceland	40	14.88	647080
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```
SELECT {[Measures].[Suicides No], [Measures].[Suicide_Rates],
[Measures].[Gdp Per Capita]} on Columns,
UNION (
{topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Gdp Per Capita])},
{bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Gdp Per Capita])}
)
on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2016];
```

Kết quả sau khi chạy câu truy vấn:

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. In the top pane, there is a code editor containing an MDX query. The query selects measures for suicides, suicide rates, and GDP per capita across all countries, ordered by GDP per capita, for the year 2016. The results are displayed in a table below, showing data for Iceland and Armenia.

	Suicides No	Suicide_Rates	Gdp Per Capita
Iceland	40	14.88	647080
Armenia	67	2.79	37880

7. Mỗi quốc gia, tìm năm có tổng số vụ tự tử cao nhất.

a. Phân tích dữ liệu trên Cube

Tại browser kéo thả những thuộc tính cần thiết cho câu truy vấn:

The screenshot shows the Microsoft Analysis Services browser interface. On the left, the cube structure is visible under the 'Measure Group' section, including dimensions like 'Country' and 'Year', and measures like 'Suicides No'. On the right, a detailed data grid displays the number of suicides for each year from 1987 to 2004, grouped by country (Albania).

Country	Year	Suicides No
Albania	1987	73
Albania	1988	63
Albania	1989	68
Albania	1992	47
Albania	1993	73
Albania	1994	50
Albania	1995	88
Albania	1996	89
Albania	1997	170
Albania	1998	154
Albania	1999	139
Albania	2000	54
Albania	2001	119
Albania	2002	133
Albania	2003	124
Albania	2004	146

Dùng hàm Generate

```
Generate (
    [Dim CountryYear].[Country].Children,
    topcount ([Dim CountryYear].[Country].Currentmember*[Dim
CountryYear].[Year].Children
    ,1,[Measures].[Suicides No])
)
```

Chuyển Design Mode sang Script Mode.

Sử dụng đoạn code trên thay thế vào vị trí tô xanh như trên hình :

```
SELECT NON EMPTY { [Measures].[Suicides No] } ON COLUMNS, NON EMPTY { {[Dim CountryYear].[Country].[Country].ALLMEMBERS * [Dim CountryYear].[Year].ALLMEMBERS } } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

```
SELECT NON EMPTY { [Measures].[Suicides No] } ON COLUMNS, NON EMPTY {Generate (
    [Dim CountryYear].[Country].Children,
    topcount ([Dim CountryYear].[Country].Currentmember*[Dim CountryYear].[Year].Children
    ,1,[Measures].[Suicides No])
)} DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

Nhấn Excute query Filter ta được kết quả:

```
SELECT NON EMPTY { [Measures].[Suicides No] } ON COLUMNS, NON EMPTY {Generate (
    [Dim CountryYear].[Country].Children,
    topcount ([Dim CountryYear].[Country].Currentmember*[Dim CountryYear].[Year].Children
    ,1,[Measures].[Suicides No])
)} DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

Country	Year	Suicides No
Albania	1997	170
Antigua...	2000	2
Argentina	2003	3289
Armenia	1995	130
Aruba	2000	10
Australia	2015	3027
Austria	1986	2139
Azerbaij...	2001	317
Bahamas	2000	11
Bahrain	2008	43
Barbados	1991	17
Belarus	1996	3627
Belgium	1985	2281
Belize	2011	33
Bosnia ...	2014	313
Brazil	2015	11163
Bulgaria	1996	1526

b. Pivot table trong Excel

Tạo Nameset [cau7]

Expression:

```
Generate (
    [Dim CountryYear].[Country].Children,
    topcount ([Dim CountryYear].[Country].Currentmember*[Dim
CountryYear].[Year].Children
    ,1,[Measures].[Suicides No])
)
```

Name: [cau7]

Expression

```
Generate (
    [Dim CountryYear].[Country].Children,
    topcount ([Dim CountryYear].[Country].Currentmember*[Dim CountryYear].[Year].Children
    ,1,[Measures].[Suicides No])
)
No issues found
```

Ln: 5 Ch: 3 Col: 6 TABS CRLF

Additional Properties

Type: Dynamic

Display folder:

Trong Pivot table:

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

	A	B	C	D
1	Row Labels	Suicides No		
2	Albania			
3	1997	170		
4	Antigua and Barbuda			
5	2000	2		
6	Argentina			
7	2003	3289		
8	Armenia			
9	1995	130		
10	Aruba			
11	2000	10		
12	Australia			
13	2015	3027		
14	Austria			
15	1986	2139		
16	Azerbaijan			
17	2001	317		
18	Bahamas			
19	2000	11		
20	Bahrain			
21	2008	43		
22	Barbados			
23	1991	17		
24	Belarus			
25	1996	3627		
26	Belgium			
27	1985	2281		
28	Belize			

PivotTable Fields pane:

- Choose fields to add to report:
- Search: Population, Suicides 100k Pop, **Suicides No** (checked)
- Sets:
 - cau11
 - cau7**
 - cau9
- Drag fields between areas below:
 - Filters: (empty)
 - Columns: (empty)
 - Rows: cau7
 - Values: Suicides No

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```
SELECT [Measures].[Suicides No] on Columns,
Generate (
    [Dim CountryYear].[Country].Children,
    topcount ([Dim CountryYear].[Country].Currentmember*[Dim
    CountryYear].[Year].Children,1,[Measures].[Suicides No])
)
on Rows
FROM [Suicide Rates DW];
```

Kết quả sau khi chạy câu truy vấn:

```
--7.Mỗi quốc gia, tìm năm có tổng số vụ tự tử cao nhất.
SELECT [Measures].[Suicides No] on Columns,
Generate (
    [Dim CountryYear].[Country].Children,
    topcount ([Dim CountryYear].[Country].Currentmember*[Dim CountryYear].[Year].Children
    ,1,[Measures].[Suicides No])
)
on Rows
FROM [Suicide Rates DW];
```

The screenshot shows the SSMS interface with the query above and its results below. The results table has columns for Country, Year, and Suicides No.

		Suicides No
Albania	1997	170
Antigua and Barbuda	2000	2
Argentina	2003	3289
Armenia	1995	130
Aruba	2000	10
Australia	2015	3027
Austria	1986	2139
Azerbaijan	2001	317
Bahamas	2000	11
Bahrain	2008	43
Barbados	1991	17
Belarus	1996	3627
Belgium	1985	2281
Belize	2011	33
Bosnia and Herzegovina	2014	313
Brazil	2015	11163
Bulgaria	1996	1526
Cabo Verde	2011	42
Canada	1999	4074
Chile	2009	2148
...
	2004	2110

8. Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính.

a. Phân tích dữ liệu trên Cube

The screenshot shows the Analysis Services Management Studio interface. On the left, there's a tree view of the cube structure under 'Suicide Rates DW'. The 'Measure Group' section shows '<All>'. The 'Calculated Members' section is empty. On the right, there's a table with columns: Dimension, Hierarchy, Operator, and Filter Expression. A dropdown menu is open under 'Dimension'. The data grid below shows a list of years, ages, sexes, and suicide counts for the year 2012.

Dimension	Hierarchy	Operator	Filter Expression
Dim CountryYear	Year	Equal	{ 2012 }

Year	Age	Sex	Suicides No
2012	15-24 years	female	5940
2012	15-24 years	male	20340
2012	25-34 years	female	7080
2012	25-34 years	male	29545
2012	35-54 years	female	17986
2012	35-54 years	male	64496
2012	5-14 years	female	608
2012	5-14 years	male	1144
2012	55-74 years	female	14071
2012	55-74 years	male	44372
2012	75+ years	female	7326
2012	75+ years	male	17252

b. Pivot table trong Excel

The screenshot shows an Excel spreadsheet with a PivotTable. The PivotTable has 'Year' in Row 1, 'Suicides No' in Row 2, and 'Column Labels' in Row 3. Row 4 defines 'Row Labels' as 'female' and 'male', and 'Grand Total'. The data rows (5-11) show counts for different age groups: 15-24 years, 25-34 years, 35-54 years, 5-14 years, 55-74 years, and 75+ years. The Grand Total row shows 53011, 177149, and 230160 respectively. The PivotTable Fields pane on the right lists 'Dim CountryYear' (with 'Year' checked), 'Dim SexAge' (with 'Sex' checked), and 'Values' (with 'Suicides No'). The 'Filters' pane shows 'Year' and 'Age'.

	A	B	C	D	E
1	Year	2012			
2					
3	Suicides No	Column Labels			
4	Row Labels	female	male	Grand Total	
5	15-24 years	5940	20340	26280	
6	25-34 years	7080	29545	36625	
7	35-54 years	17986	64496	82482	
8	5-14 years	608	1144	1752	
9	55-74 years	14071	44372	58443	
10	75+ years	7326	17252	24578	
11	Grand Total	53011	177149	230160	
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```
SELECT {[Measures].[Suicides No]} on Columns,
{[Dim SexAge].[Age].Children* [Dim SexAge].[Sex].Children} on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2012];
```

Kết quả sau khi chạy câu truy vấn:

```
--8. Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính.
SELECT {[Measures].[Suicides No]} on Columns,
{[Dim SexAge].[Age].Children*[Dim SexAge].[Sex].Children} on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2012];
```

100 %

		Suicides No
15-24 years	female	5940
15-24 years	male	20340
25-34 years	female	7080
25-34 years	male	29545
35-54 years	female	17986
35-54 years	male	64496
5-14 years	female	608
5-14 years	male	1144
55-74 years	female	14071
55-74 years	male	44372
75+ years	female	7326
75+ years	male	17252

9. Mỗi quốc gia, cho biết nhóm tuổi có tỷ lệ tử vong cao nhất và thấp nhất.

a. Phân tích dữ liệu trên Cube

Tại browser kéo thả những thuộc tính cần thiết cho câu truy vấn:

Dimension	Hierarchy	Operator	Filter Expression	Parameters
<Select dimension>				
Country	Age			
Albania	15-24 years	4.47358629...		
Albania	25-34 years	4.26208062...		
Albania	35-54 years	3.70044595...		
Albania	5-14 years	0.43319962...		
Albania	55-74 years	3.16430244...		
Albania	75+ years	4.74648496...		
Antigua...	15-24 years	0		
Antigua...	25-34 years	0.26968570...		
Antigua...	35-54 years	1.11642446...		
Antigua...	5-14 years	0		
Antigua...	55-74 years	1.88657887...		
Antigua...	75+ years	0		
Argentina	15-24 years	9.20328249...		
Argentina	25-34 years	8.02535393...		
Argentina	35-54 years	8.40717242...		
Argentina	5-14 years	0.69221450...		
Argentina	55-74 years	12.2768316...		
Argentina	75+ years	17.5605617...		

Dùng hàm Generate

```
Generate (
    [Dim CountryYear].[Country].Children,
    [Dim CountryYear].[Country].Currentmember *
UNION(
    topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]),
    bottomcount ([Dim SexAge].[Age].Children,
    1,[Measures].[Suicide_Rates])
)
)
```

Thực hiện tương tự như câu 3. Hoặc thực hiện theo cách sau:

Chuyển Design Mode sang Script Mode.

Sử dụng đoạn code trên thay thế vào vị trí tô xanh như trên hình :

```
SELECT NON EMPTY { [Measures].[Suicide_Rates] } ON COLUMNS, NON EMPTY {[Dim CountryYear].[Country].[Country].ALLMEMBERS * [Dim SexAge].[Age].ALLMEMBERS } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

```
SELECT NON EMPTY { [Measures].[Suicide_Rates] } ON COLUMNS, NON EMPTY {Generate (
    [Dim CountryYear].[Country].Children,
    [Dim CountryYear].[Country].Currentmember *
UNION(
    topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]),
    bottomcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates])
)
)} DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

Nhấn Excute query Filter ta được kết quả:

```

SELECT NON EMPTY { [Measures].[Suicide_Rates] } ON COLUMNS, NON EMPTY {Generate (
    [Dim CountryYear].[Country].Children,
    [Dim CountryYear].[Country].Currentmember *
    UNION(
        topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates])
    )
) } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR,
FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS

```

Country	Age	Suicide_Rates
Albania	75+ years	4.74648496...
Albania	5-14 years	0.43319962...
Antigua...	55-74 years	1.88657887...
Antigua...	75+ years	0
Argentina	75+ years	17.5605617...
Argentina	5-14 years	0.69221450...
Armenia	75+ years	6.94013932...
Armenia	5-14 years	0.17646727...
Aruba	75+ years	14.5847006...
Aruba	5-14 years	0
Australia	25-34 years	17.0514509...
Australia	5-14 years	0.40103292...
Austria	75+ years	46.1379506...
Austria	5-14 years	0.51099352...
Azerbaij...	75+ years	2.82245419...
Azerbaij...	5-14 years	0.23144844...

b. Pivot table trong Excel

Tạo Nameset [cau9]

Expression:

```

Generate (
    [Dim CountryYear].[Country].Children,
    [Dim CountryYear].[Country].Currentmember *
    UNION(
        topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates])
    )
)

```

Name: [cau9]

Expression:

```

Generate (
    [Dim CountryYear].[Country].Children,
    [Dim CountryYear].[Country].Currentmember *
    UNION(
        topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates])
    )
)

```

No issues found

Type: Dynamic

Display folder:

Trong Pivot table:

	A	B	C
1	Row Labels	Suicide_Rates	
2	Albania		
3	5-14 years	0.43	
4	75+ years	4.75	
5	Antigua and Barbuda		
6	55-74 years	1.89	
7	75+ years	0.00	
8	Argentina		
9	5-14 years	0.69	
10	75+ years	17.56	
11	Armenia		
12	5-14 years	0.18	
13	75+ years	6.94	
14	Aruba		
15	5-14 years	0.00	
16	75+ years	14.58	
17	Australia		
18	25-34 years	17.05	
19	5-14 years	0.40	
20	Austria		
21	5-14 years	0.51	
22	75+ years	46.14	
23	Azerbaijan		
24	5-14 years	0.23	
25	75+ years	2.82	
26	Bahamas		
27	35-54 years	2.39	
28	5-14 years	0.15	
29	Bahrain		

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```

with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides No]/
[Measures].[Population]* 100000', FORMAT_STRING = 'General Number'

SELECT [Measures].[Suicide_Rates] on Columns,
Generate (
    [Dim CountryYear].[Country].Children,
    [Dim CountryYear].[Country].Currentmember *
UNION(
    topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]),
    bottomcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates])
)
)on Rows
FROM [Suicide Rates DW];

```

Kết quả sau khi chạy câu truy vấn:

```
--9. Mỗi quốc gia, cho biết nhóm tuổi có tỷ lệ tử tử cao nhất và thấp nhất.
with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides No]/[Measures].[Population]*100000', FORMAT_STRING =
    SELECT [Measures].[Suicide_Rates] on Columns,
    Generate (
        [Dim CountryYear].[Country].Children,
        [Dim CountryYear].[Country].Currentmember *
        UNION(
            topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]),
            bottomcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates])
        )
    )
    on Rows
FROM [Suicide Rates DW];
```

100 % ▶ Results

		Suicide_Rates
Albania	75+ years	4.74648496419041
Albania	5-14 years	0.433199623783881
Antigua and Barbuda	55-74 years	1.88657887786288
Antigua and Barbuda	75+ years	0
Argentina	75+ years	17.5605617645497
Argentina	5-14 years	0.692214508485246
Armenia	75+ years	6.94013932059982
Armenia	5-14 years	0.176467279473804
Aruba	75+ years	14.5847006490192
Aruba	5-14 years	0
Australia	25-34 years	17.051450944187
Australia	5-14 years	0.401032925467392
Austria	75+ years	46.1379506723148
Austria	5-14 years	0.510993525426162
Azerbaijan	75+ years	2.82245419112047
Azerbaijan	5-14 years	0.2314484429306

10. Cho biết Top10 nước có tỷ lệ tử tử ở độ tuổi 15-24 và 25-34 nhiều nhất từ năm 2000-2016.

a. Phân tích dữ liệu trên Cube

Tạo Nameset [cau10]

Expression:

topcount ([Dim CountryYear].[Country].Children,10,[Measures].[Suicide_Rates])

Name: [cau10]

Expression

```
topcount ([Dim CountryYear].[Country].Children,10,[Measures].[Suicide_Rates])
✓ No issues found
```

Additional Properties

Type: Dynamic

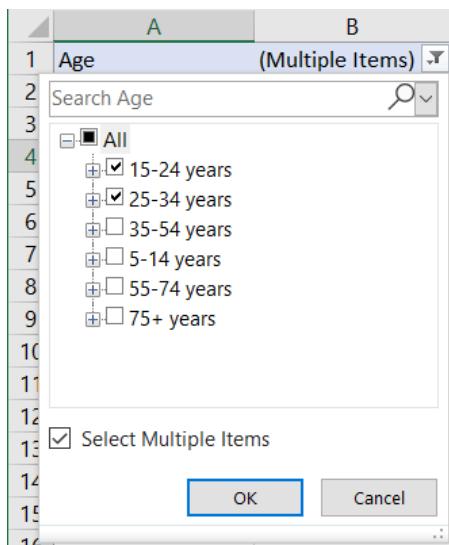
Display folder:

Kéo thả Nameset [cau10] và điều kiện Dimension sang trường Filter ta thu được kết quả:

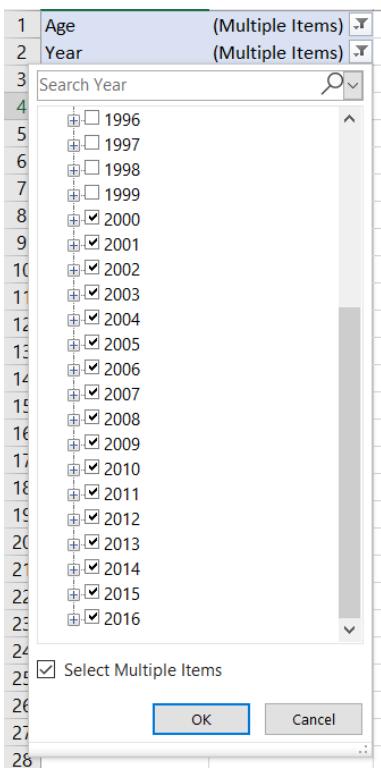
Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim CountryYear	# Year	Range (Inclusive)	2000 : 2016	<input type="checkbox"/> <input type="checkbox"/>
Dim SexAge	# Age	Equal	{ 15-24 years, 25-34 years }	<input type="checkbox"/> <input type="checkbox"/>
Dim CountryYear	# Country	In	cau10	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>
Country	Suicide_Rates			
Belarus	25.9961501480116			
Finland	19.8068186937001			
Guyana	35.2245714629804			
Kazakh...	31.0755358565787			
Lithuania	31.5663885031748			
Mongolia	22.6316863931835			
Russian...	31.2928137167629			
Sri Lanka	27.7481517454698			
Suriname	27.2226164868348			
Ukraine	19.7882764140869			

b. Pivot table trong Excel

Lọc thuộc tính Age



Lọc thuộc tính Year



Kết quả cuối cùng

Screenshot of a PivotTable report in Excel.

PivotTable Fields

Choose fields to add to report:

Search

Sets

- cau10
- cau12
- cau13
- cau2
- cau4
- cau5
- cau6

Drag fields between areas below:

Filters	Columns
Year	
Age	

Rows	Values
cau10	Suicide_Rates

Defer Layout Update Update

PivotTable Data (Approximate Values)

Year	Age	Suicide_Rates
Belarus	cau10	26.00
Finland	cau10	19.81
Guyana	cau10	35.22
Kazakhstan	cau10	31.08
Lithuania	cau10	31.57
Mongolia	cau10	22.63
Russian Federation	cau10	31.29
Sri Lanka	cau10	27.75
Suriname	cau10	27.22
Ukraine	cau10	19.79

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```
with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides No]/[Measures].[Population]*100000', FORMAT_STRING = 'General Number'

SELECT [Measures].[Suicides 100k Pop] on Columns,
topcount ([Dim CountryYear].[Country].Children,10,[Measures].[Suicide_Rates])
on Rows
FROM [Suicide Rates DW]
WHERE ({[Dim SexAge].[Age].&[15-24 years],[Dim SexAge].[Age].&[25-34 years]}
,{[Dim CountryYear].[Year].&[2000]:[Dim CountryYear].[Year].&[2016]});
```

Kết quả sau khi chạy câu truy vấn:

Suicides 100k Pop	Country
1859.65	Guyana
2116.66	Lithuania
1985.94	Russian Federation
2034.49	Kazakhstan
561.25	Sri Lanka
1629.41	Suriname
1133.31	Belarus
89.73	Mongolia
1251.45	Finland
1174.4	Ukraine

11. Mỗi năm, tìm quốc gia có tỷ lệ tự tử cao nhất và thấp nhất, đồng thời hiển thị GDP bình quân đầu người của quốc gia đó.

a. Phân tích dữ liệu trên Cube

Tại browser kéo thả những thuộc tính cần thiết cho câu truy vấn:

Dimension	Hierarchy	Operator	Filter Expression
<Select dimension>			
Year	Country	Suicide_Rates	Gdp Per Capita
1985	Antigua...	0	46200
1985	Argentina	7.33828458...	39168
1985	Australia	12.7793991...	148488
1985	Austria	29.4088070...	117108
1985	Bahamas	0.49091801...	136716
1985	Bahrain	3.00628587...	119760
1985	Barbados	3.03030303...	73644
1985	Belgium	24.6073185...	112272
1985	Brazil	3.59894415...	22776
1985	Bulgaria	17.3488233...	24528
1985	Canada	13.8638887...	186264
1985	Chile	6.39788673...	19896
1985	Colombia	3.99562516...	16716
1985	Costa R...	5.72399606...	25740
1985	Dominica	0	17820
1985	Ecuador	5.12934297...	26856
1985	France	24.3292476...	129180
1985	Greece	4.36638851...	61872

Dùng hàm Generate

```
Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    UNION(
        topcount ([Dim CountryYear].[Country].Children,
        1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim CountryYear].[Country].Children,
        1,[Measures].[Suicide_Rates])
    )
)
```

Chuyển Design Mode sang Script Mode.

Sử dụng đoạn code trên thay thế vào vị trí tó xanh như trên hình :

```
SELECT NON EMPTY { [Measures].[Suicide_Rates], [Measures].[Gdp Per Capita] } ON COLUMNS, NON EMPTY {{[Dim CountryYear].[Year].[Year].ALLMEMBERS * [Dim CountryYear].[Country].[Country].ALLMEMBERS }} DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

```
SELECT NON EMPTY { [Measures].[Suicide_Rates], [Measures].[Gdp Per Capita] } ON COLUMNS, NON EMPTY {Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    UNION(
        topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicide_Rates])
    )
)} DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

Nhấn Excute query Filter ta được kết quả:

```

SELECT NON EMPTY { [Measures].[Suicide_Rates], [Measures].[Gdp Per Capita] } ON COLUMNS, NON EMPTY {Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    UNION(
        topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicide_Rates])
    )
)} DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR,
FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS

```

Year	Country	Suicide_Rates	Gdp Per Capita
1985	Sri Lanka	40.9035144692213	5172
1985	Dominica	0	17820
1986	Sri Lanka	37.9040378969464	5448
1986	Saint Vincent and Grenadines	0	21252
1987	Finland	29.5398886023277	238332
1987	Saint Vincent and Grenadines	0	22896
1988	Finland	30.2572513812155	282552
1988	Antigua and Barbuda	0	80316
1989	Finland	30.359022839095	306984
1989	Jamaica	0	25044
1990	Finland	32.3415541913542	363252
1990	Saint Kitts and Nevis	0	64068
1991	Hungary	40.9212518521145	42732
1991	Saint Kitts and Nevis	0	64416
1992	Hungary	41.0047979714106	47640
1992	Saint Kitts and Nevis	0	70416

b. Pivot table trong Excel

Tạo Nameset [cau11]

Expression:

```

Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    UNION(
        topcount ([Dim CountryYear].[Country].Children,
        1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim CountryYear].[Country].Children,
        1,[Measures].[Suicide_Rates])
    )
)

```

Name:

[cau11]

Expression

```

Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    UNION(
        topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicide_Rates])
    )
)

```

No issues found

Ln: 8 Ch: 3 Col: 6 TABS CRLF

Trong Pivot table:

The screenshot shows a Microsoft Excel PivotTable with the following data structure:

	A	B	C
Row Labels		Suicide_Rates	Gdp Per Capita
1985			
3 Dominica		0.00	17820
4 Sri Lanka		40.90	5172
1986			
6 Saint Vincent and Grenadines		0.00	21252
7 Sri Lanka		37.90	5448
1987			
9 Finland		29.54	238332
10 Saint Vincent and Grenadines		0.00	22896
1988			
12 Antigua and Barbuda		0.00	80316
13 Finland		30.26	282552
1989			
15 Finland		30.36	306984
16 Jamaica		0.00	25044
1990			
18 Finland		32.34	363252
19 Saint Kitts and Nevis		0.00	64068
1991			
21 Hungary		40.92	42732
22 Saint Kitts and Nevis		0.00	64416
1992			
24 Hungary		41.00	47640
25 Saint Kitts and Nevis		0.00	70416
1993			
27 Belize		0.59	39336
28 Russian Federation		40.56	37920
1994			

PivotTable Fields pane details:

- Choose fields to add to report: Suicides No
- Sets:
 - cau11
 - cau14
 - cau3
 - cau7
 - cau9
- Drag fields between areas below:
- Filters: None
- Columns:
- Rows: cau11
- Values:
 - Suicide_Rates
 - Gdp Per Capita
- Defer Layout Update:
- Update button: Update

c. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```

with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides
No]/[Measures].[Population]*100000', FORMAT_STRING = 'General Number'
SELECT {[Measures].[Suicide_Rates],[Measures].[Gdp Per Capita]} on Columns,
Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    UNION(
        topcount ([Dim
CountryYear].[Country].Children,1,[Measures].[Suicide_Rates]),
        bottomcount ([Dim
CountryYear].[Country].Children,1,[Measures].[Suicide_Rates])
    )
)
on Rows
FROM [Suicide Rates DW];
  
```

Kết quả sau khi chạy câu truy vấn:

```
--11. Mỗi năm, tìm quốc gia có tỷ lệ tự tử cao nhất và thấp nhất,
-- đồng thời hiển thị GDP bình quân đầu người của quốc gia đó.
with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides No]/[Measures].[Population]*100000', FORMAT_STR:
    SELECT {[Measures].[Suicide_Rates]},[Measures].[Gdp Per Capita]} on Columns,
    Generate (
        [Dim CountryYear].[Year].Children,
        [Dim CountryYear].[Year].Currentmember *
        UNION(
            topcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicide_Rates]),
            bottomcount ([Dim CountryYear].[Country].Children,1,[Measures].[Suicide_Rates])
        )
    )
    on Rows
FROM [Suicide Rates DW];
```

		Suicide_Rates	Gdp Per Capita
1985	Sri Lanka	40.9035144692213	5172
1985	Dominica	0	17820
1986	Sri Lanka	37.9040378969464	5448
1986	Saint Vincent and Grenadines	0	21252
1987	Finland	29.5398886023277	238332
1987	Saint Vincent and Grenadines	0	22896
1988	Finland	30.2572513812155	282552
1988	Antigua and Barbuda	0	80316
1989	Finland	30.359022839095	306984
1989	Jamaica	0	25044
1990	Finland	32.3415541913542	363252
1990	Saint Kitts and Nevis	0	64068
1991	Hungary	40.9212518521145	42732
1991	Saint Kitts and Nevis	0	64416
1992	Hungary	41.0047979714106	47640

12. Tỉ lệ tự tử ở từng giới tính của Top 10 quốc gia có số dân đông nhất năm 2012.

- Phân tích dữ liệu trên Cube

Tạo Nameset [cau12]

Expression:

topcount ([Dim CountryYear].[Country].Children,10,[Measures].[Suicide_Rates])

Name:

Expression Ln: 1 Ch: 78 SPC CRLF

```
topcount ([Dim CountryYear].[Country].Children,10,[Measures].[Suicide_Rates])
✓ No issues found
```

Additional Properties

Type: <input type="text" value="Dynamic"/>	
Display folder: <input type="text"/>	

Kéo thả Nameset [cau12] và điều kiện Dimension sang trường Filter ta thu được kết quả:

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim CountryYear	Country	In	cau12	<input type="checkbox"/> <input type="checkbox"/>
Dim CountryYear	Year	Equal	{ 2012 }	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>

Country	Sex	Population	Suicide_Rates
Brazil	female	94517324	2.38474800...
Brazil	male	91073891	8.82909460...
Estonia	female	670869	7.00583869...
Estonia	male	580762	32.7156391...
Hungary	female	4972457	10.9804871...
Hungary	male	4480724	40.2390327...
Kazakhstan	female	7831781	7.18865862...
Kazakhstan	male	7191768	37.2231139...
Latvia	female	1053575	6.92879007...
Latvia	male	877946	42.2577242...
Lithuania	female	1538559	11.3092835...
Lithuania	male	1299211	57.9582531...
Slovenia	female	985250	8.11976655...
Slovenia	male	960738	37.4711940...
Ukraine	female	23233556	7.30409068...
Ukraine	male	19682379	37.3989343...

b. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```

with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides
No]/[Measures].[Population]*100000', FORMAT_STRING = 'General Number'
SELECT {[Measures].[Population],[Measures].[Suicide_Rates]} on Columns,
(topcount (
    [Dim CountryYear].[Country].Children,10,[Measures].[Suicide_Rates]))
*[Dim SexAge].[Sex].Children
on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2012];

```

Kết quả sau khi chạy câu truy vấn:

```

with member [Measures].[Suicide_Rates] AS '[Measures].[Suicides_No]/[Measures].[Population]*100000', FORMAT_STRING = 'General Number'
SELECT {[Measures].[Population],[Measures].[Suicide_Rates]} on Columns,
    (topcount ([Dim CountryYear].[Country].Children,10,[Measures].[Suicide_Rates]))*[Dim SexAge].[Sex].Children
    on Rows
FROM [Suicide Rates DW]
WHERE [Dim CountryYear].[Year].&[2012];

```

		Population	Suicide_Rates
Lithuania	female	1538559	11.3092835568867
Lithuania	male	1299211	57.9582531243963
Guyana	female	339529	14.7262825855818
Guyana	male	339627	45.9057108469898
Republic of Korea	female	24029767	18.8849105361696
Republic of Korea	male	23991218	40.1021740538559
Suriname	female	242989	13.5808616850969
Suriname	male	243975	40.9878061276777
Hungary	female	4972457	10.9804871113013
Hungary	male	4480724	40.2390327991637
Latvia	female	1053575	6.92879007189806
Latvia	male	877946	42.2577242791698
Slovenia	female	985250	8.11976655671149
Slovenia	male	960738	37.4711940195974
Russian Federation	female	72781138	7.49369800997
Russian Federation	male	61887384	39.0855105460589
Japan	female	62082000	12.7830933281789
Japan	male	58653000	31.3743542529794
Kazakhstan	female	7831781	7.18865862056153
Kazakhstan	male	7191768	37.2231139825423

13. Trong khoảng 10 năm gần nhất, năm nào thế giới ghi nhận tổng số vụ tự tử cao nhất.

- a. Phân tích dữ liệu trên Cube

Tạo Nameset [cau13]

Expression:

```
topcount ({[Dim CountryYear].[Year].&[2007]:[Dim CountryYear].[Year].&[2016]},  
1,[Measures].[Suicides_No])
```

Name: [cau13]

Expression:

```
topcount ({[Dim CountryYear].[Year].&[2007]:[Dim CountryYear].[Year].&[2016]},1,[Measures].[Suicides_No])
```

No issues found

Type: Dynamic

Kéo thả Nameset [cau13] và điều kiện Dimension sang trường Filter ta thu được kết quả:

Dimension	Hierarchy	Operator	Filter Expression
Dim CountryYear	Year	In	cau13
<Select dimension>			
Year	Suicides No		
2009 243487			

b. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```
SELECT [Measures].[Suicides No] on Columns,
topcount ([Dim CountryYear].[Year].&[2007]:[Dim
CountryYear].[Year].&[2016]),1,[Measures].[Suicides No]) on Rows
FROM [Suicide Rates DW];
```

Kết quả sau khi chạy câu truy vấn:

--13. Trong khoảng 10 năm gần nhất, cho biết 5 quốc gia có tỷ lệ tự tử có xu hướng tăng nhanh nhất. -- Trong khoảng 10 năm gần nhất, năm nào thế giới ghi nhận tổng số vụ tự tử cao nhất SELECT [Measures].[Suicides No] on Columns, topcount ([Dim CountryYear].[Year].&[2007]:[Dim CountryYear].[Year].&[2016]),1,[Measures].[Suicides No]) on Rows FROM [Suicide Rates DW];
100 % Messages Results Suicides No 2009 243487

14. Mỗi năm, thống kê số vụ tự tử ở các nước theo GDP For Year giảm dần.

a. Phân tích dữ liệu trên Cube

Tạo Nameset [cau14]

Expression:

```
Generate ([Dim CountryYear].[Year].Children,
[Dim CountryYear].[Year].Currentmember *
Order ([Dim CountryYear].[CountryYear].Children ,[Measures].[Gdp For Year],
desc))
```

Name: [cau14]

Expression

```
Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    Order(
        [Dim CountryYear].[CountryYear].Children ,[Measures].[Gdp For Year], desc
    )
)
```

No issues found

Type: Dynamic

Display folder:

Chuyển Design Mode sang Script Mode.

Sử dụng đoạn code trên thay thế vào vị trí tô xanh như trên hình :

```
SELECT NON EMPTY { [Measures].[Suicides No], [Measures].[Gdp For Year] } ON COLUMNS, NON EMPTY { ([Dim CountryYear].[CountryYear].[CountryYear].ALLMEMBERS
) } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR,
FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

```
SELECT NON EMPTY { [Measures].[Suicides No], [Measures].[Gdp For Year] } ON COLUMNS, NON EMPTY { Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    Order(
        [Dim CountryYear].[CountryYear].Children ,[Measures].[Gdp For Year], desc
    )
) } DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR,
FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

Nhấn Execute query Filter ta được kết quả:

```

SELECT NON EMPTY { [Measures].[Suicides No], [Measures].[Gdp For Year] } ON COLUMNS, NON EMPTY {Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    Order(
        [Dim CountryYear].[CountryYear].Children ,[Measures].[Gdp For Year], desc
    ))} DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS FROM [Suicide Rates DW] CELL PROPERTIES VALUE, BACK_COLOR,
FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS

```

Year	CountryYear	Suicides No	Gdp For Year
1985	United States1985	29446	52160808000000
1985	Japan1985	23257	16786712937852
1985	France1985	12501	6637660972404
1985	United Kingdom1985	5105	5871421971252
1985	Italy1985	4759	5408709792516
1985	Canada1985	3258	4377077993412
1985	Brazil1985	4228	2675313485220
1985	Mexico1985	1544	2213670688980
1985	Spain1985	2514	2163628946772
1985	Australia1985	1861	2162291938332
1985	Netherlands1985	1638	1704119067672
1985	Republic of Korea1...	3689	1203277166040
1985	Argentina1985	1988	1061000026800
1985	Belgium1985	2281	1040760465516

b. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```

SELECT {[Measures].[Suicides No],[Measures].[Gdp For Year]} on Columns,
Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    Order(
        [Dim CountryYear].[CountryYear].Children ,[Measures].[Gdp For Year], desc
    )) on Rows
FROM [Suicide Rates DW];

```

Kết quả sau khi chạy câu truy vấn:

```

--14. Mỗi năm thống kê số vụ tự tử ở các nước theo GDP For Year giảm dần.
SELECT {[Measures].[Suicides No],[Measures].[Gdp For Year]} on Columns,
Generate (
    [Dim CountryYear].[Year].Children,
    [Dim CountryYear].[Year].Currentmember *
    Order(
        [Dim CountryYear].[CountryYear].Children ,[Measures].[Gdp For Year], desc
    ) on Rows
) on Columns
FROM [Suicide Rates DW];

```

100 %

		Suicides No	Gdp For Year
1985	United States1985	29446	52160808000000
1985	Japan1985	23257	16786712937852
1985	France1985	12501	6637660972404
1985	United Kingdom1985	5105	5871421971252
1985	Italy1985	4759	5408709792516
1985	Canada1985	3258	4377077993412
1985	Brazil1985	4228	2675313485220
1985	Mexico1985	1544	2213670688980
1985	Spain1985	2514	2163628946772
1985	Australia1985	1861	2162291938332
1985	Netherlands1985	1638	1704119067672
1985	Republic of Korea1985	3689	1203277166040
1985	Argentina1985	1988	1061000026800
1985	Belgium1985	2281	1040760465516
1985	Austria1985	2091	832641293700
1985	Greece1985	405	573850211700
1985	Thailand1985	2982	466808312544
1985	Colombia1985	1001	418732936224
1985	Israel1985	234	329923097808
1985	Portugal1985	983	325421714088
1985	New Zealand1985	338	299410741692

15. Top 10 quốc gia có tổng số vụ tự tử cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính.

a. Phân tích dữ liệu trên Cube

```

{Head( Order( [Dim CountryYear].[Country].Children,[Measures].[Suicides No],
desc), 10)}
    * { [Dim SexAge].[Sex].children}

```

Dimension	Hierarchy	Operator	Filter Expression	Parameters
Dim CountryYear	Year	Range (Inclusive)	2010 : 2016	<input type="checkbox"/> <input type="checkbox"/>
Dim CountryYear	Country	Custom	Head(Order([Dim CountryYear].[Country].Children,[Measu...))	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				<input type="checkbox"/> <input type="checkbox"/>

Country	Sex	Suicides No
Brazil	female	13268
Brazil	male	48585
France	female	12368
France	male	36643
Germany	female	15576
Germany	male	44899
Japan	female	47820
Japan	male	110135
Mexico	female	6541
Mexico	male	27379
Poland	female	4784
Poland	male	31626
Republic of Korea	female	27235
Republic of Korea	male	60158
Russian Federation	female	31019
Russian Federation	male	143681
Ukraine	female	7922
Ukraine	male	34740
United States	female	54370
United States	male	192197

b. Ngôn ngữ MDX

Câu lệnh truy vấn bằng ngôn ngữ MDX:

```

SELECT [Measures].[Suicides No] on Columns,
NON EMPTY DRILLDOWNLEVEL (
    {Head( Order( [Dim CountryYear].[Country].Children,[Measures].[Suicides
No], desc), 10)} *
    { [Dim SexAge].[Sex].children}) on Rows
FROM [Suicide Rates DW]
WHERE {[Dim CountryYear].[Year].&[2010]:[Dim CountryYear].[Year].&[2016]};

```

Kết quả sau khi chạy câu truy vấn:

```
--15. Top 10 quốc gia có tổng số vụ tự tử cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính.
SELECT [Measures].[Suicides No] on Columns,
NON EMPTY DRILLDOWNLEVEL (
    { Head( Order( [Dim CountryYear].[Country].Children,[Measures].[Suicides No], desc), 10) }
    * { [Dim SexAge].[Sex].children } on Rows
)
FROM [Suicide Rates DW]
WHERE {[Dim CountryYear].[Year].&[2010]:[Dim CountryYear].[Year].&[2016]};
```

100 % ▶ Messages Results

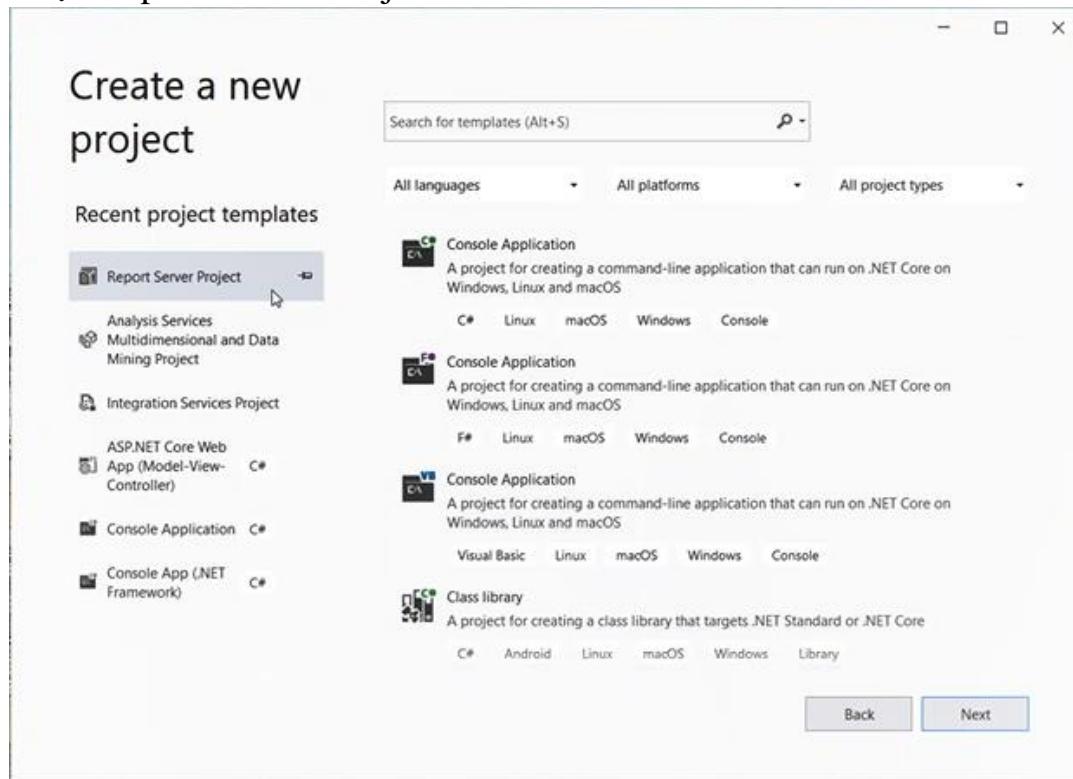
		Suicides No
United States	female	54370
United States	male	192197
Russian Federation	female	31019
Russian Federation	male	143681
Japan	female	47820
Japan	male	110135
Republic of Korea	female	27235
Republic of Korea	male	60158
Brazil	female	13268
Brazil	male	48585
Germany	female	15576
Germany	male	44899
France	female	12368
France	male	36643
Ukraine	female	7922
Ukraine	male	34740
Poland	female	4784
Poland	male	31626
Mexico	female	6541
Mexico	male	27379

6 Quá trình lập báo biểu SSRS

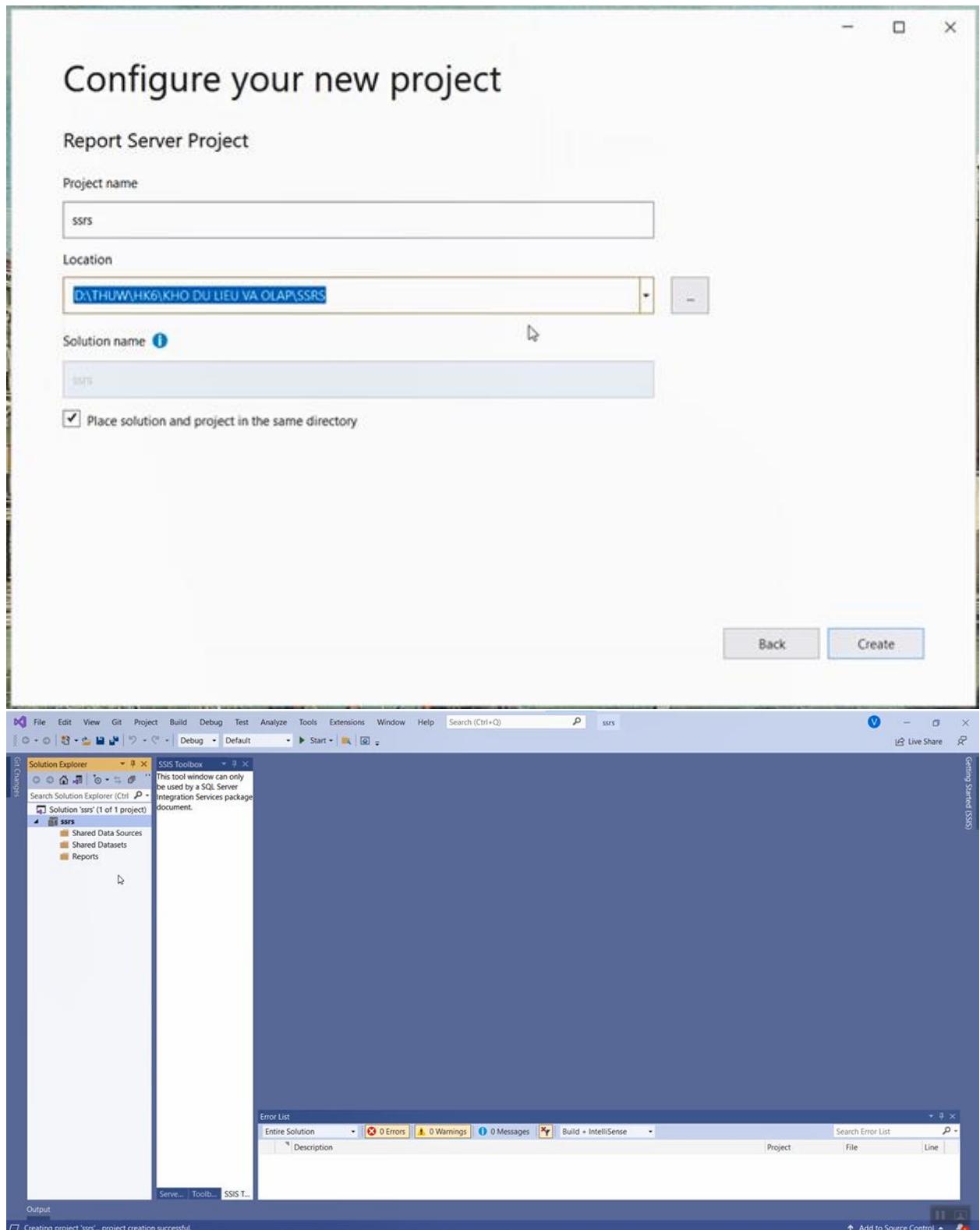
6.1. Tạo project SSRS

Mở Visual Studio (phiên bản 2019) và tạo mới project.

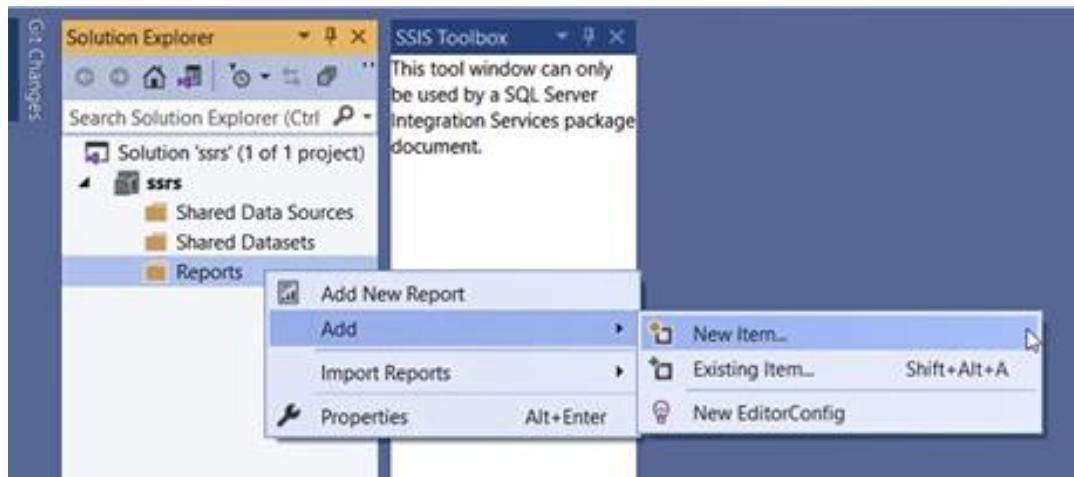
Chọn Report Service Project.



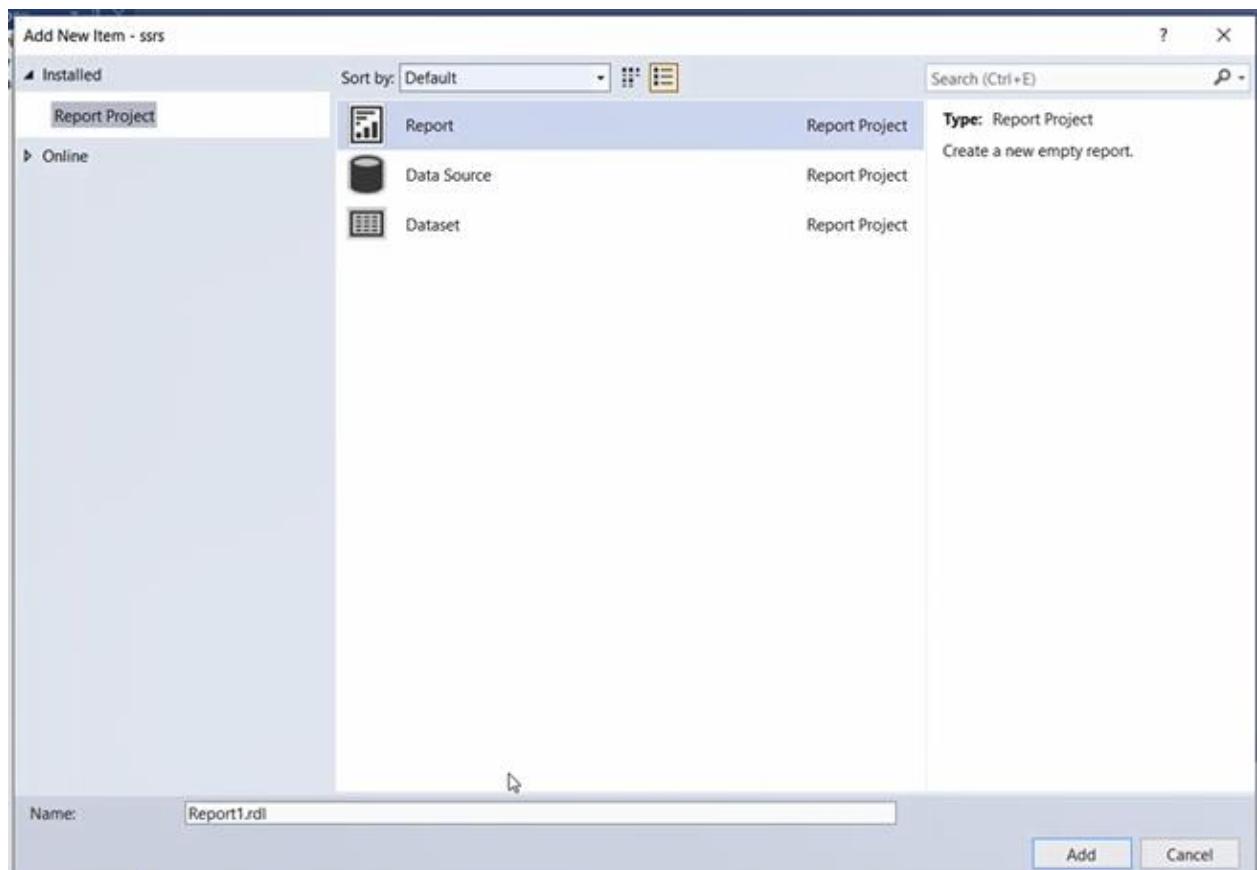
Đặt tên project và chọn đường dẫn lưu trữ.



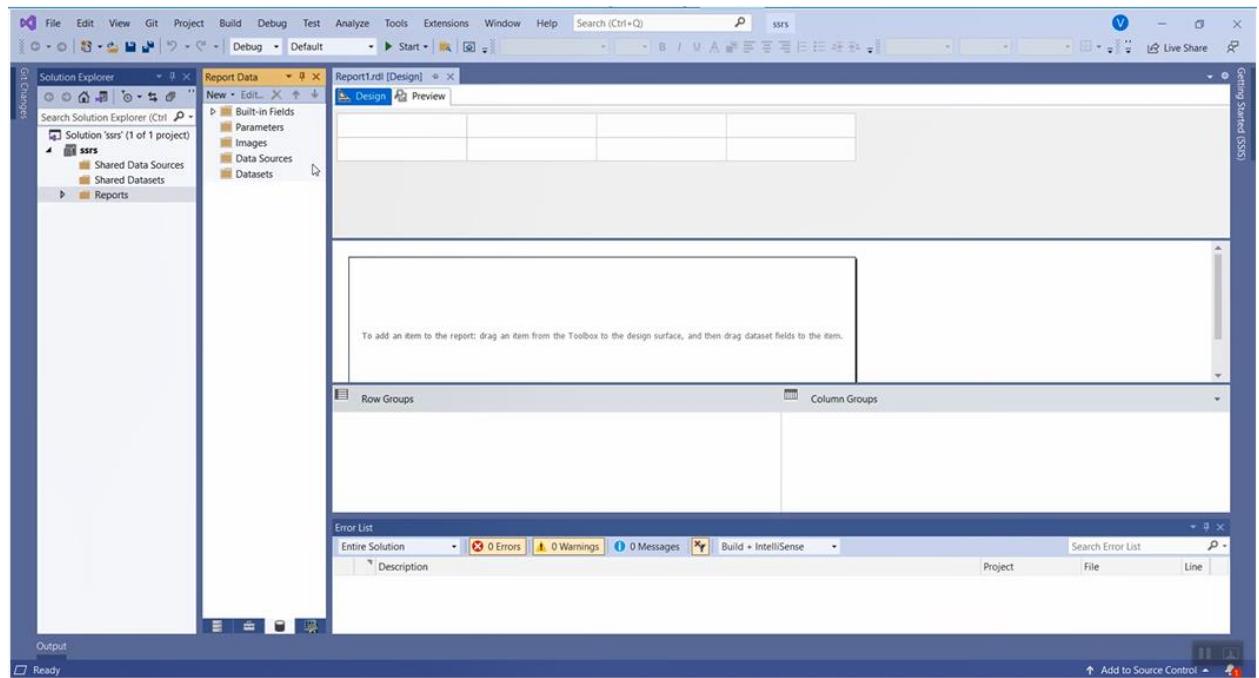
- Tạo một Report mới, tại Solution Explorer, ở mục Report nháy chuột phải chọn Add → New Item



Cửa sổ mới xuất hiện, chọn Report và đặt tên cho report sau đó nhấn Add

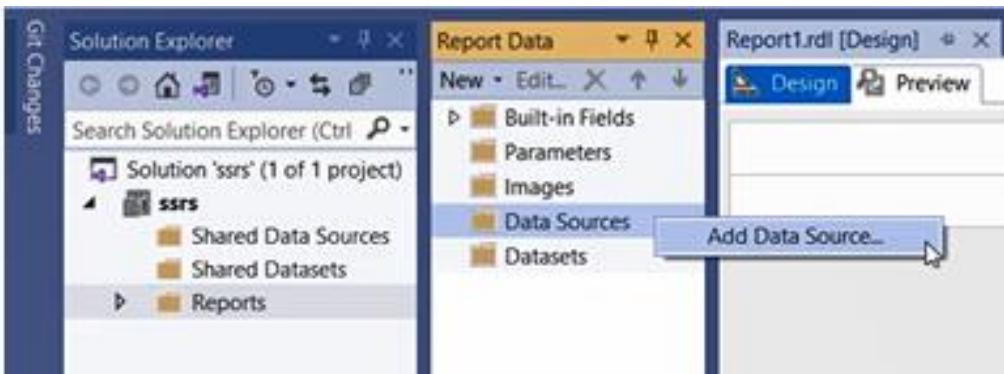


Màn hình làm việc chính



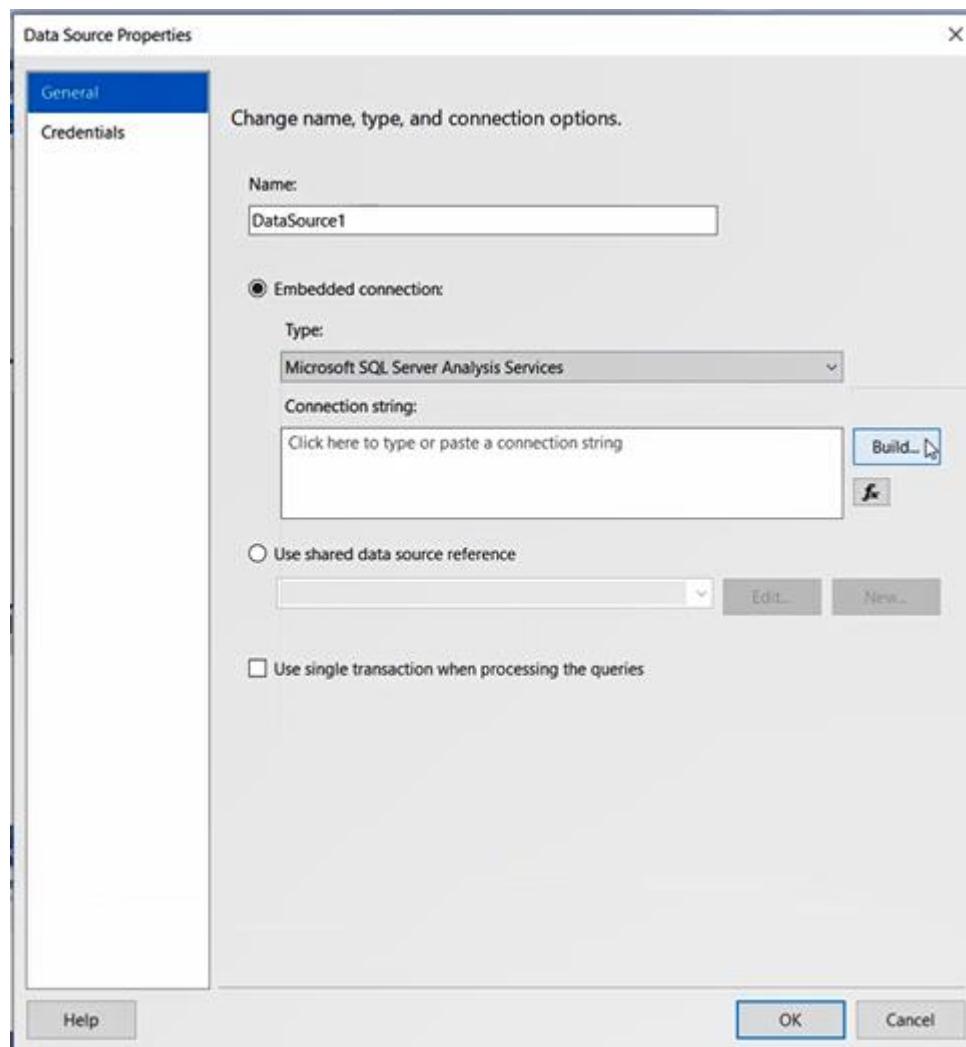
6.2. Tạo kết nối kho dữ liệu Data Source

Tại Report Data, ở mục Data Source nhấp chuột phải chọn Add Data Source



Cửa sổ mới xuất hiện

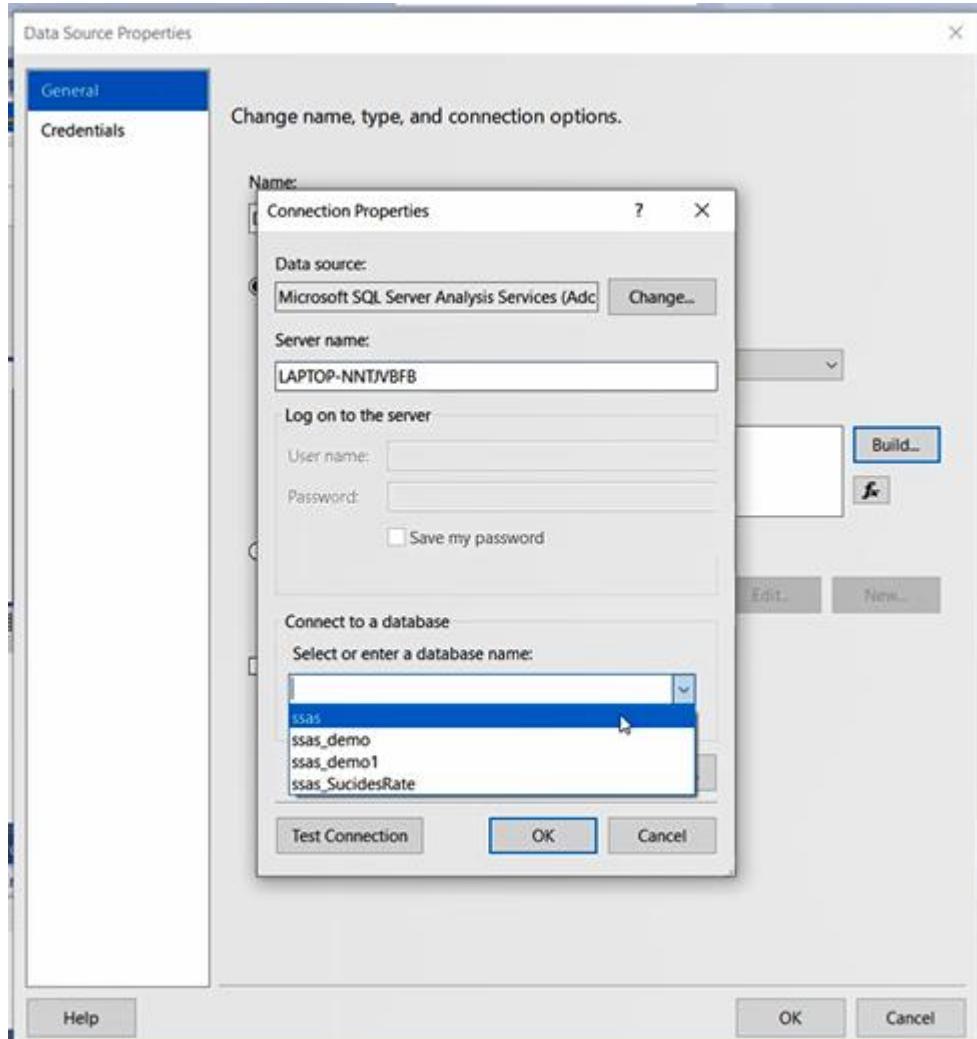
Tại Type chọn Microsoft SQL Server Analysis Services, sau đó nhấn Build...



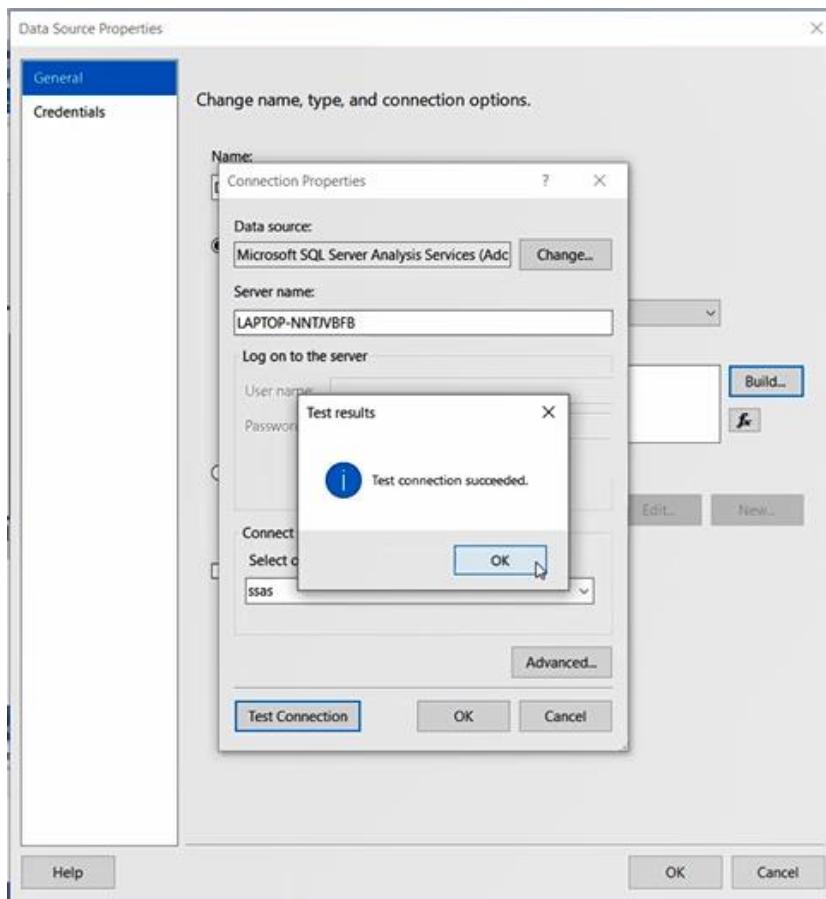
Tại cửa sổ Connection Properties,

Nhập tên Server name.

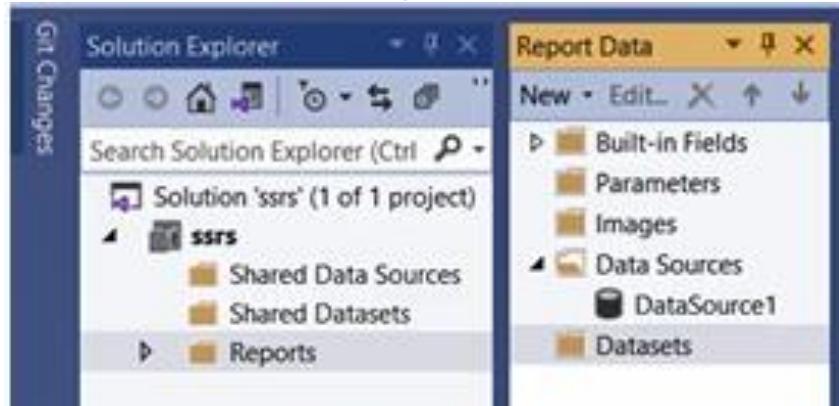
Tại mục Connect to a database nhập tên database “ssas” đã hoàn thành tại phần 5.Quá trình SSAS.



Nhấn Test Connection, nếu thành công nhấn OK để hoàn thành



Data Source tạo thành công

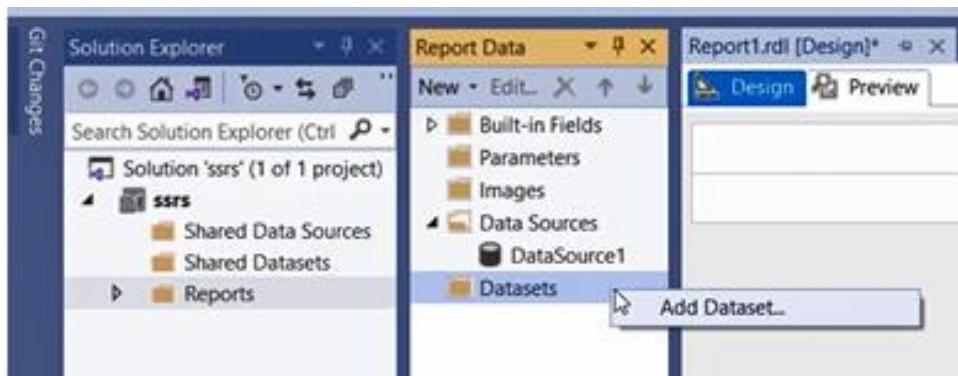


6.3. Lập báo biểu trên Report Builder

6.3.1. Report1 (Report Totals) Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính (Câu truy vấn 8).

Thực hiện phần 6.1 và 6.2 tạo Report1.

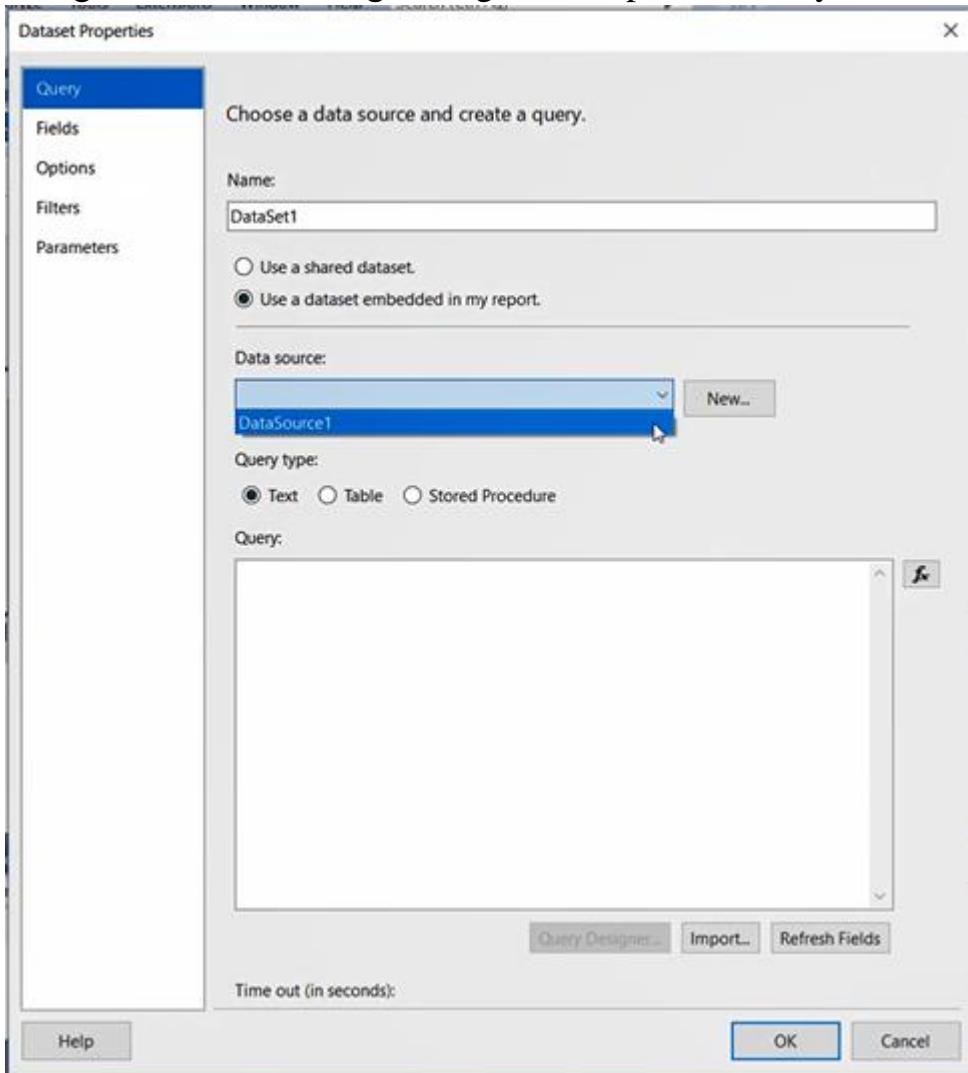
Thêm Datasets cho Report1 bằng cách tại Report Data, ở mục Datasets nháy chuột phải chọn Add Dataset...

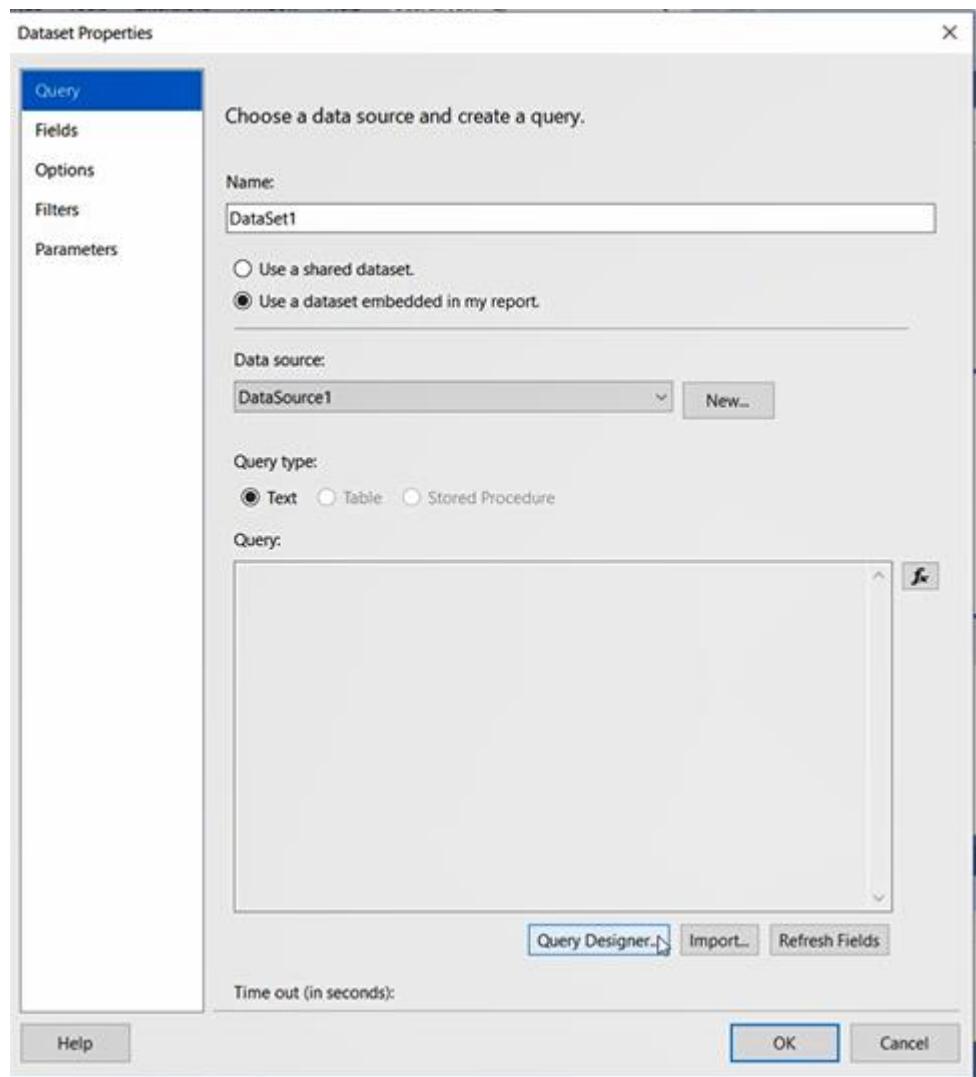


Cửa sổ Dataset Properties xuất hiện.

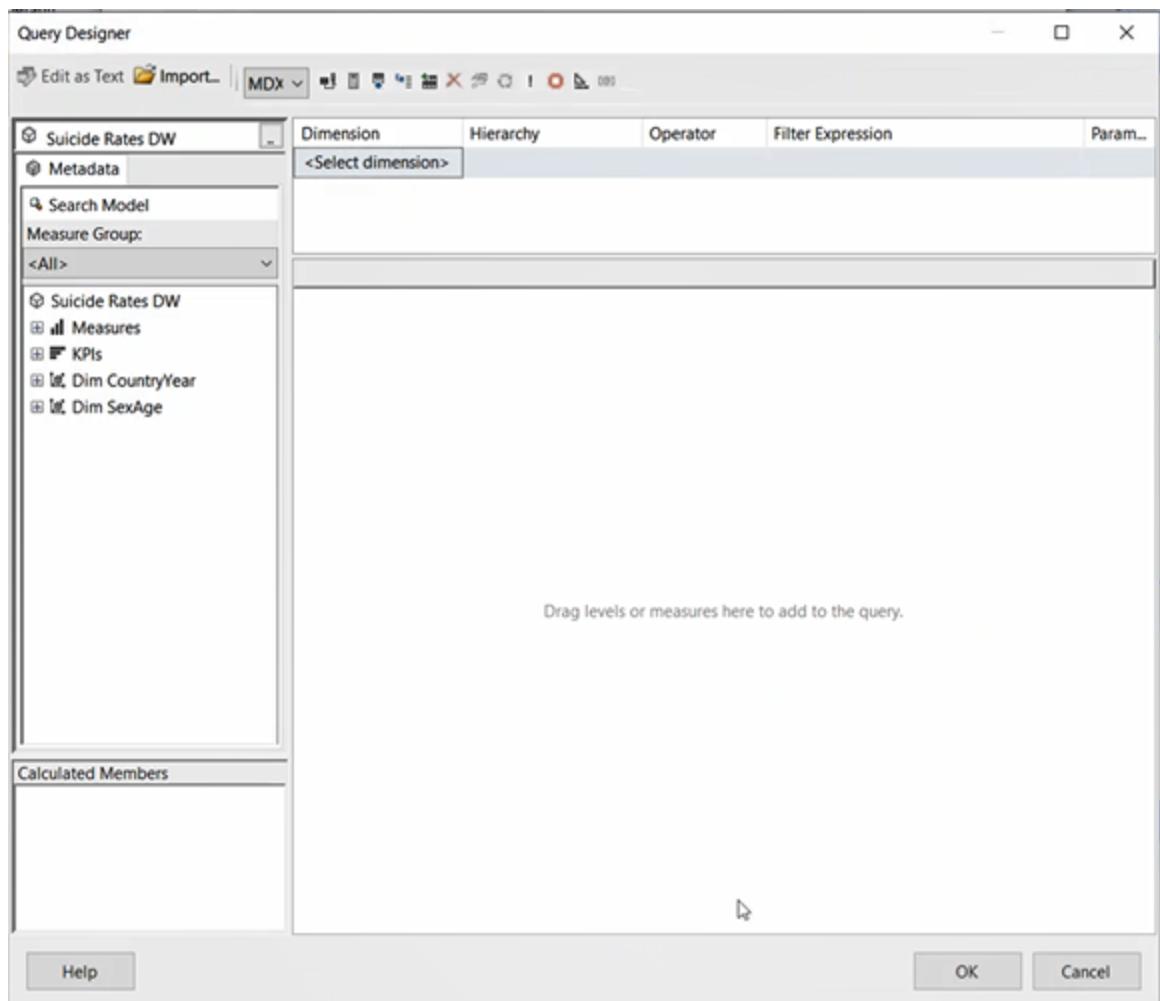
Đặt tên cho Dataset tại mục Name. Chọn Use a dataset embedded in my report.

Chọn Data source đã tạo ở bước trước, sau đó nháy chọn Query Designer... để tìm những trường thích hợp với câu truy vấn.





Cửa sổ Query Designer xuất hiện



Tương tự như ở Cube Browser trong quá trình SSAS, thực hiện kéo thả các thuộc tính Year, Age, Sex, Suicides No và lọc theo năm 2012 để thực hiện câu truy vấn.

Dataset Properties

Query Designer

MDX

Suicide Rates DW

Dimension Hierarchy Operator Filter Expression

Dim CountryYear Year Equal {2012}

Measure Group:

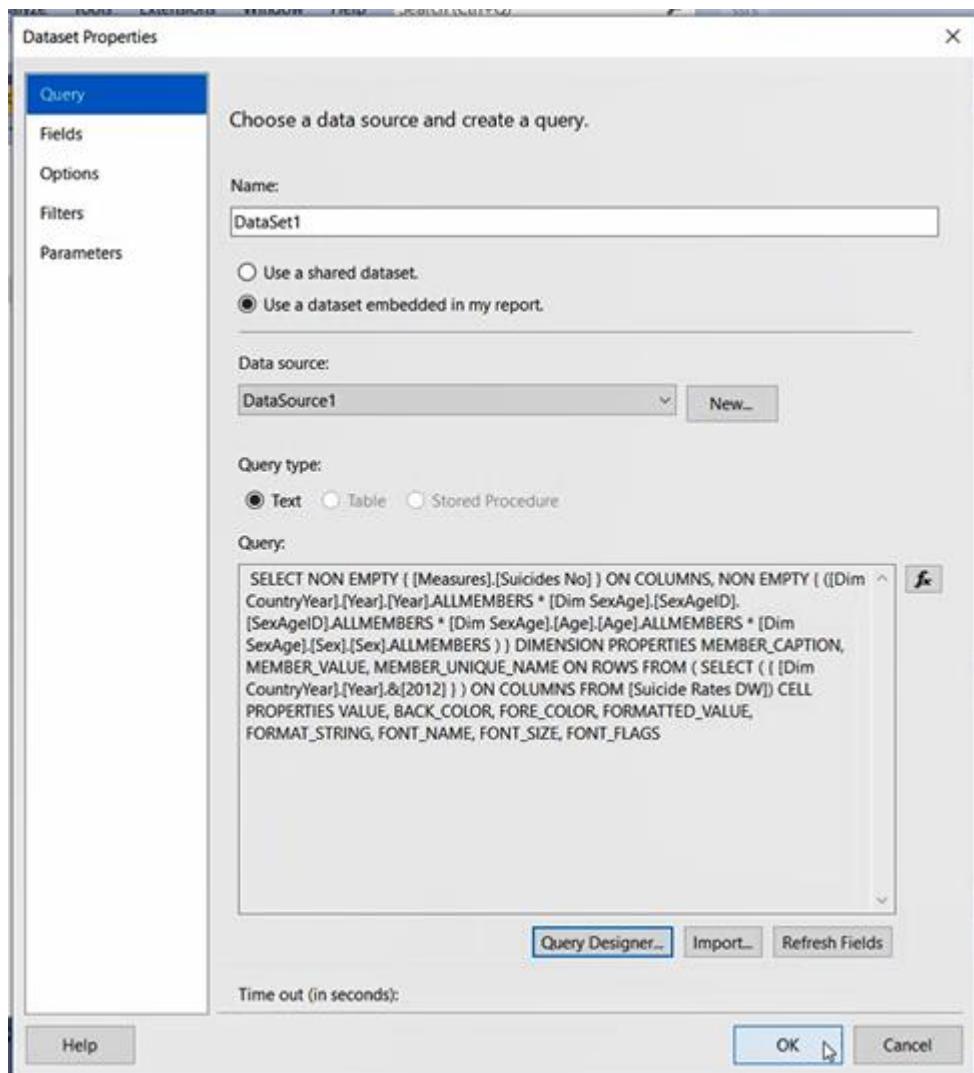
<Select dimension>

Year	SexAgeID	Age	Sex	Suicides No
2012	1	15-24 years	female	5940
2012	2	25-34 years	female	7080
2012	3	35-54 years	female	17886
2012	4	5-14 years	female	608
2012	5	55-74 years	female	14071
2012	6	75+ years	female	7326
2012	7	15-24 years	male	20340
2012	8	25-34 years	male	29545
2012	9	35-54 years	male	64496
2012	10	5-14 years	male	1144
2012	11	55-74 years	male	44372
2012	12	75+ years	male	17252

OK Cancel

Nháy Chọn OK.

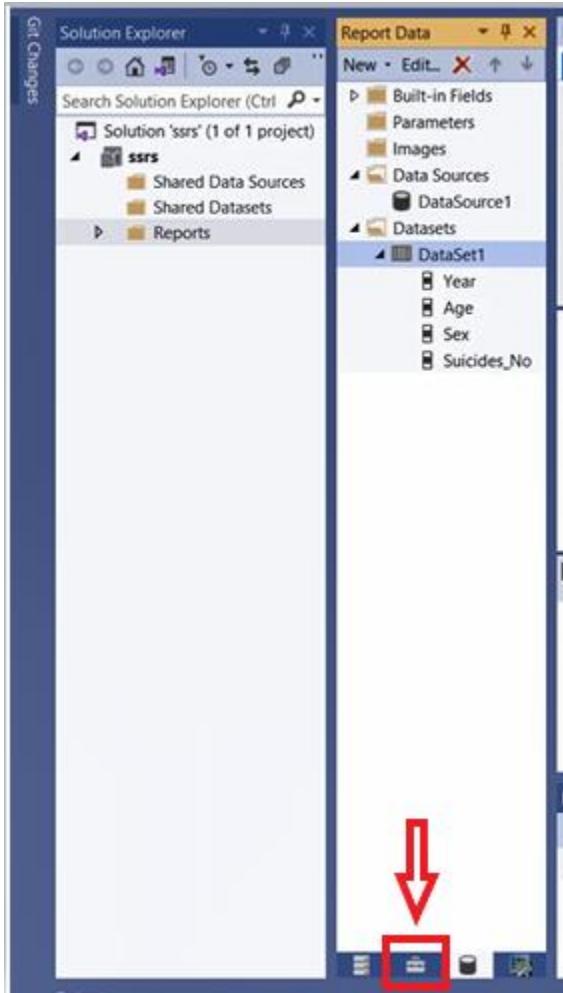
Trở về cửa sổ Dataset Properties nháy chọn OK.



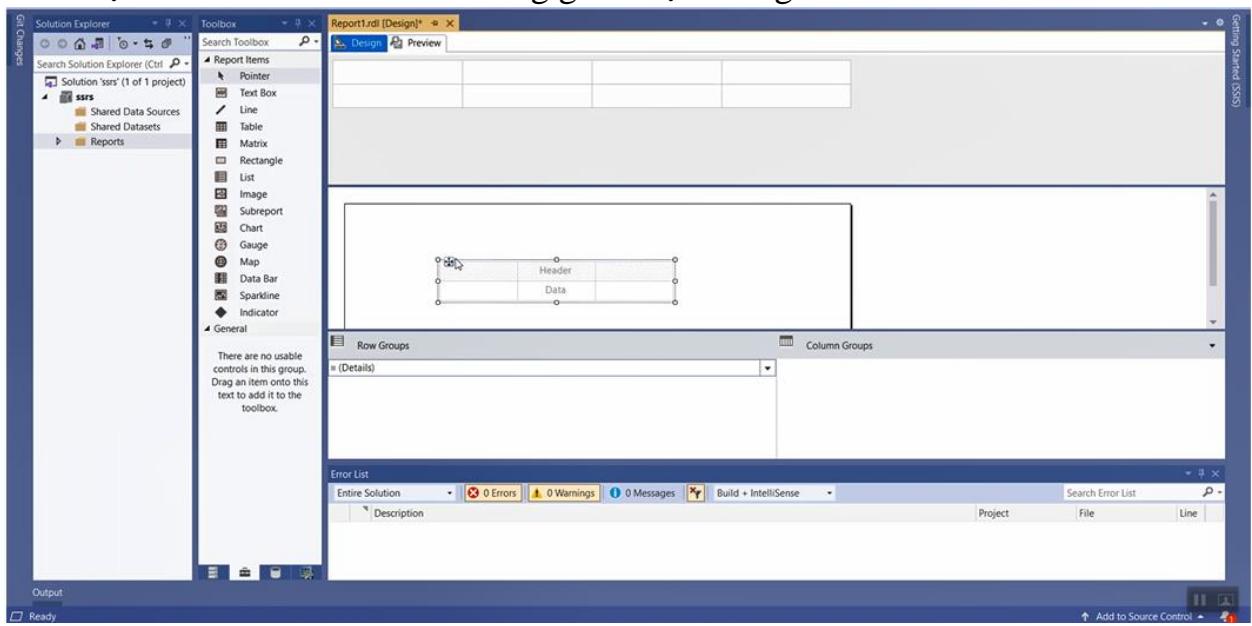
Dataset mới đã được thêm vào mục Dataset.

- Thêm dữ liệu vào Report

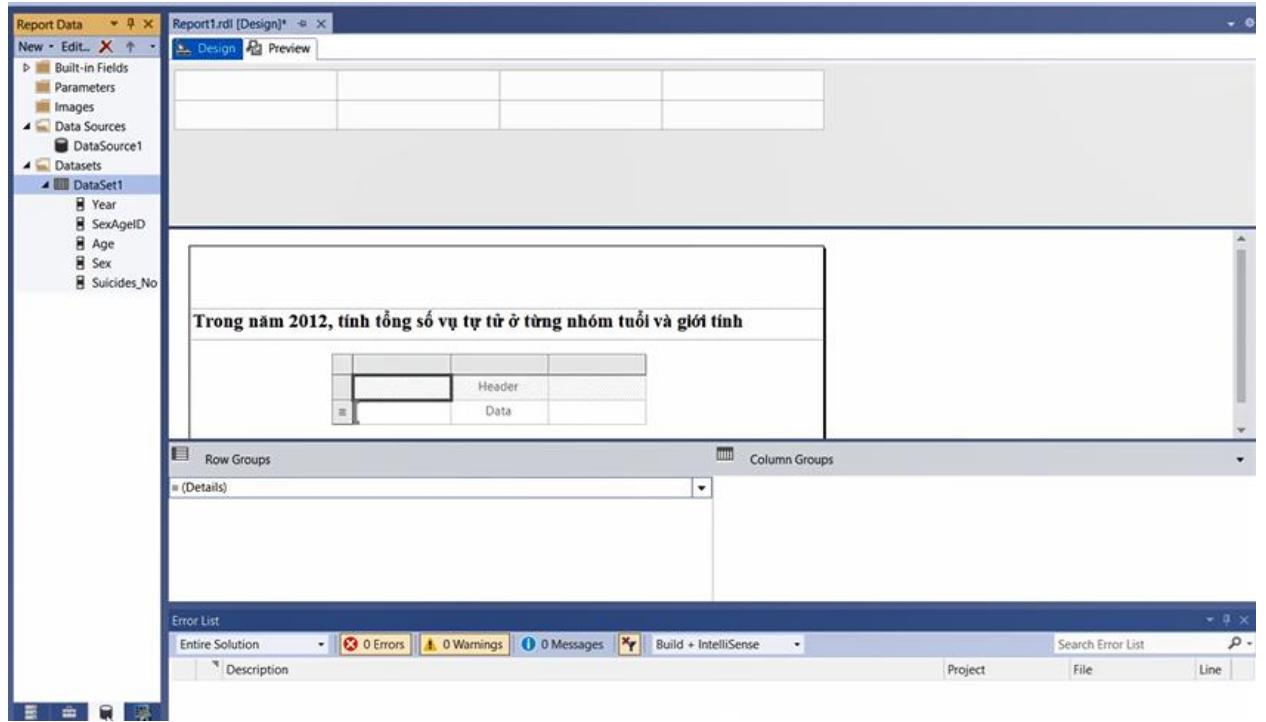
Sử dụng SSIS Toolbox sử dụng tổ hợp phím Ctrl + Alt + X hoặc chuyển tab tại Report Data



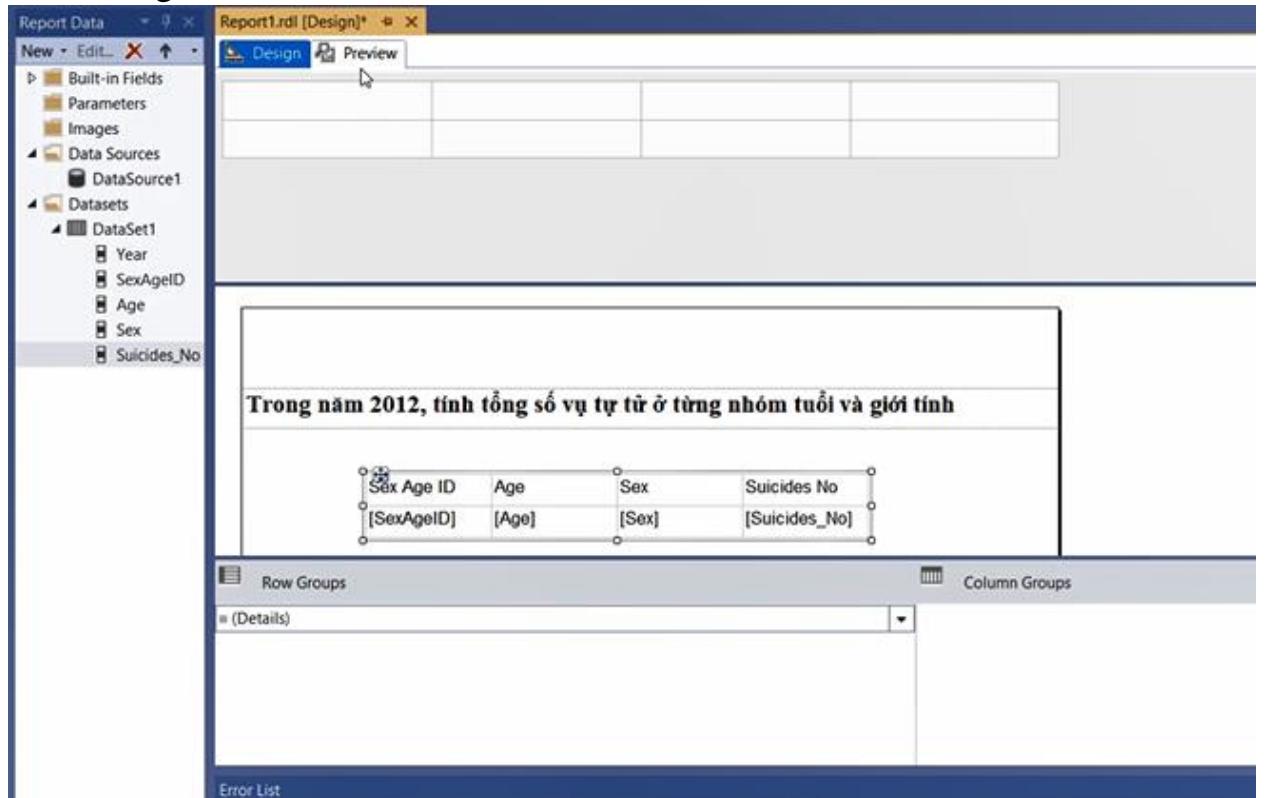
Kéo một Table mới từ Toolbox sang giao diện Design



Kéo một Text Box mới từ Toolbox sang giao diện Design và nhập tiêu đề cho report



Kéo những thuộc tính cần thiết vào Table như hình minh họa:



Chuyển sang tab Preview để xem Report

Report1.rdl [Design] X

Design Preview

1 of 1 100% Find | Next

Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính

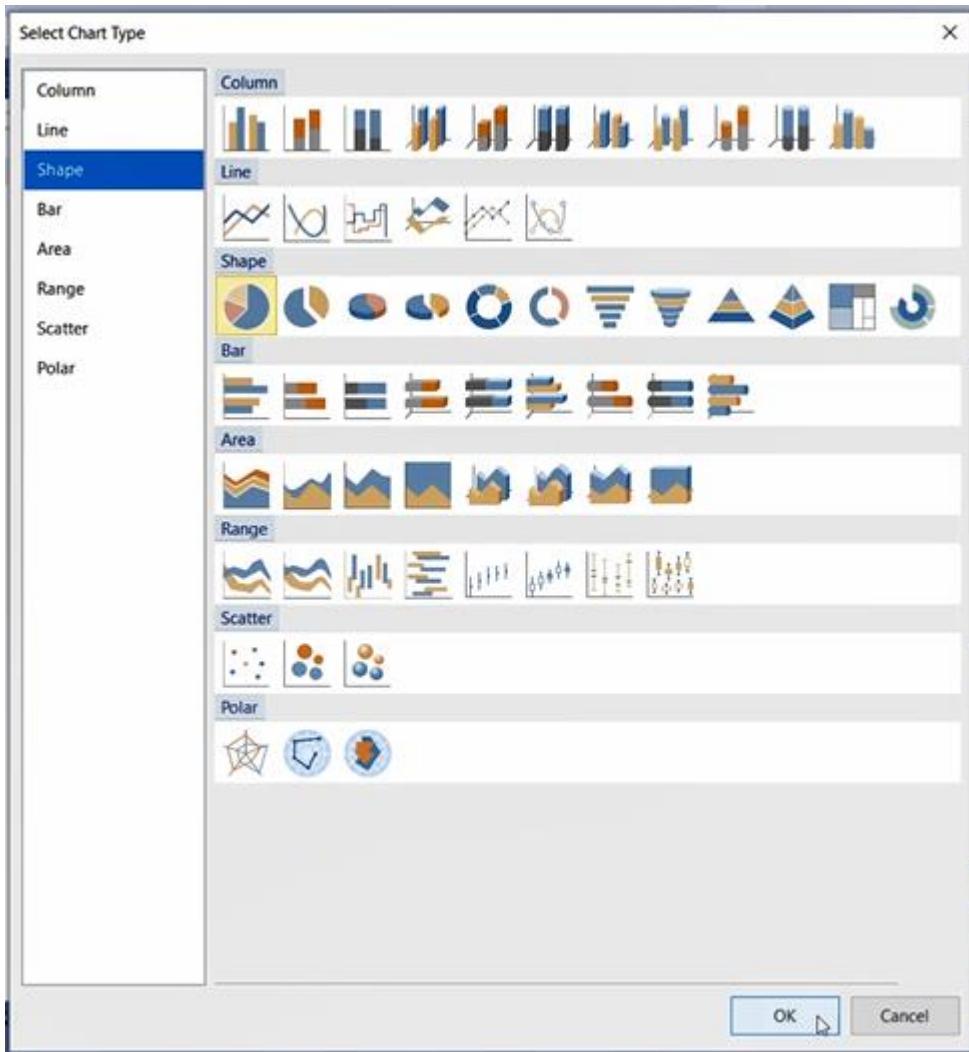
Sex Age ID	Age	Sex	Suicides No
1	15-24 years	female	5940
2	25-34 years	female	7080
3	35-54 years	female	17986
4	5-14 years	female	608
5	55-74 years	female	14071
6	75+ years	female	7326
7	15-24 years	male	20340
8	25-34 years	male	29545
9	35-54 years	male	64496
10	5-14 years	male	1144
11	55-74 years	male	44372
12	75+ years	male	17252

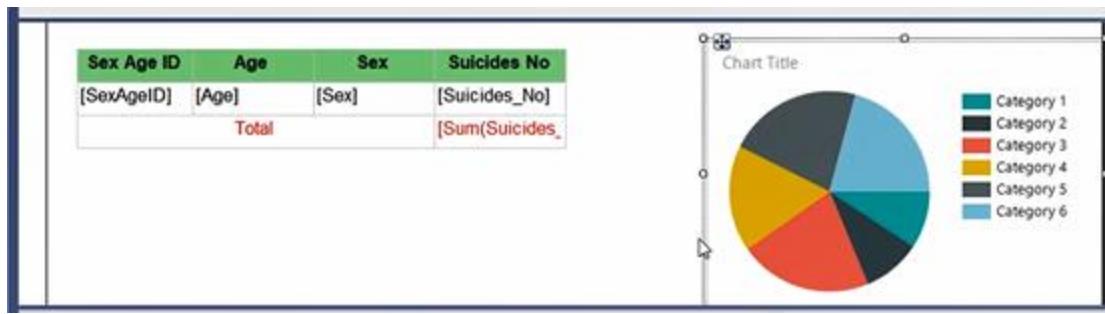
Tiếp tục định dạng lại Report
 Tại ô [Suicides_No] nháy chuột phải chọn Add Total

The screenshot shows the SSRS Design View with a context menu open over the 'Suicides No' column header of the table. The menu includes options like Text Box, Cut, Copy, Paste, Delete, Select, Expression..., Text Box Properties..., Tablix, Insert Column, Insert Row, Delete Columns, Delete Rows, Add Group, Row Group, Add Total (which is highlighted), and Insert.

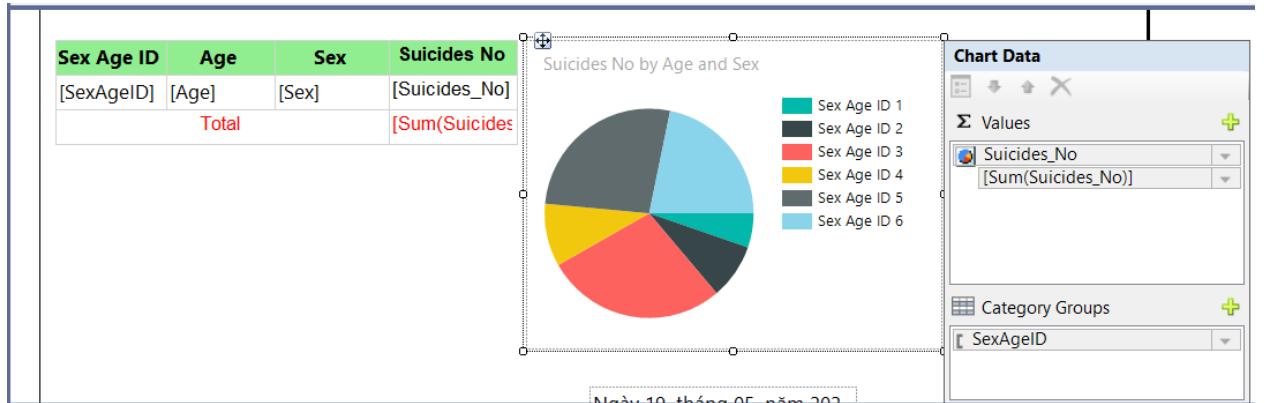
Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính			
Sex Age ID	Age	Sex	Suicides No
[SexAgeID]	[Age]	[Sex]	[Suicides_No]
Total			[Sum(Suicides_

Kéo một Chart mới từ Toolbox sang giao diện Design
Tại cửa sổ Select Chart Type, chọn đồ thị Stacked Column, sau đó nháy chọn OK.





Nháy chọn vào đồ thị Chart Data xuất hiện, kéo thả những thuộc tính cần thiết vào Chart Data như hình minh họa:



Chỉnh sửa báo cáo cho phù hợp

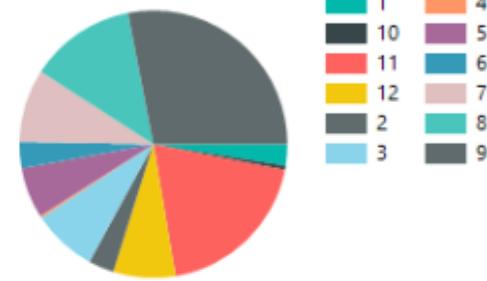
Tại Preview, chọn Print Layout để hiển thị báo cáo trước khi in.



Năm 2012 tổng số vụ tự tử ở từng nhóm tuổi và giới tính

Sex Age ID	Age	Sex	Suicides No
1	15-24 years	female	5940
2	25-34 years	female	7080
3	35-54 years	female	17986
4	5-14 years	female	608
5	55-74 years	female	14071
6	75+ years	female	7326
7	15-24 years	male	20340
8	25-34 years	male	29545
9	35-54 years	male	64496
10	5-14 years	male	1144
11	55-74 years	male	44372
12	75+ years	male	17252
Total			230160

Suicides No by Age and Sex



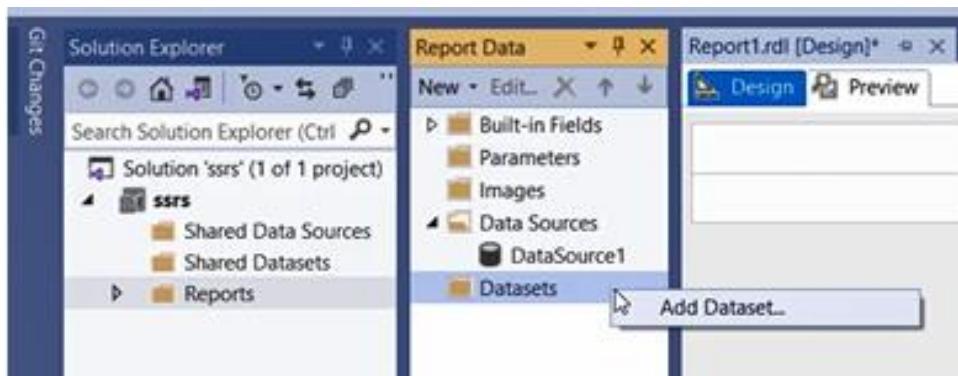
Ngày 19, tháng 05, năm 2022

Người tạo

6.3.2. Report2 (Report Grouping) Mỗi quốc gia, cho biết nhóm tuổi có tỷ lệ tử tử cao nhất và thấp nhất (Câu truy vấn 9).

Thực hiện phần 6.1 và 6.2 tạo Report2.

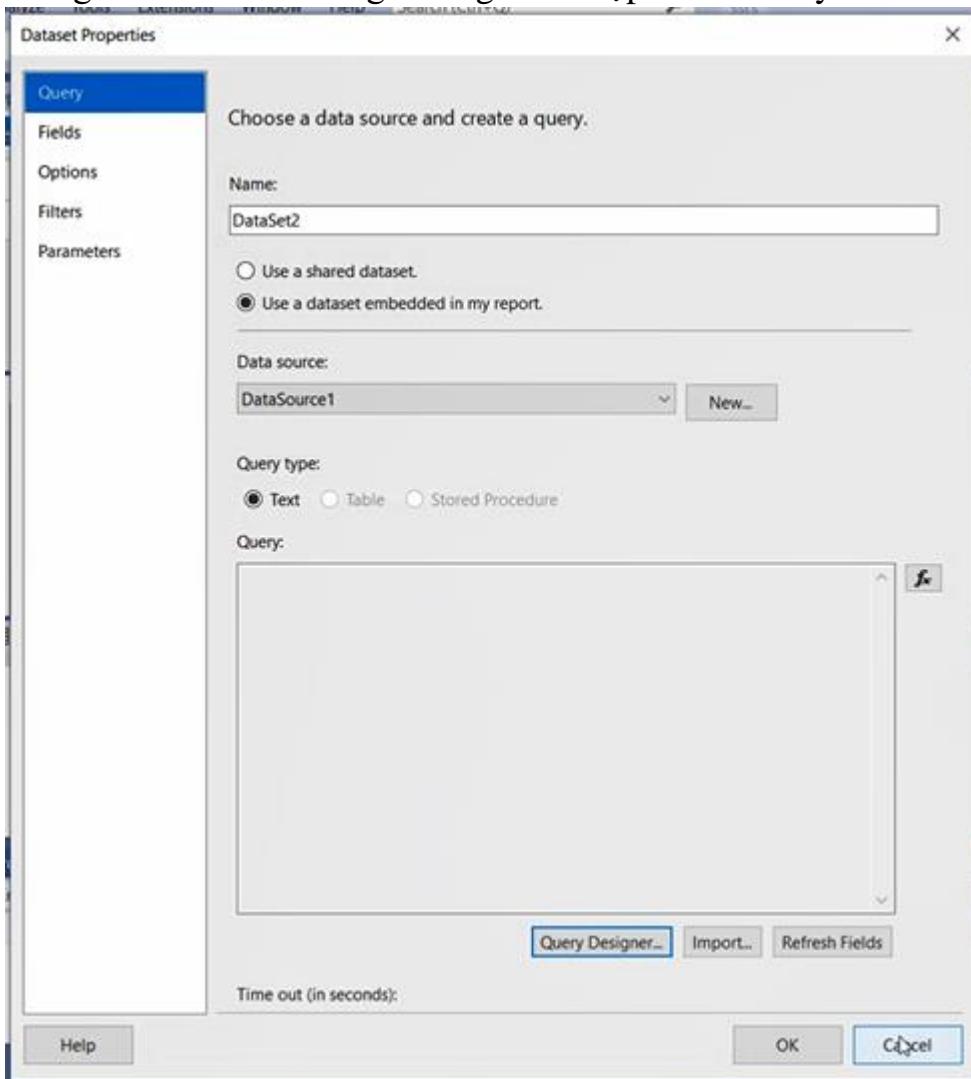
Thêm Datasets cho Report2 bằng cách tại Report Data, ở mục Datasets nháy chuột phải chọn Add Dataset...



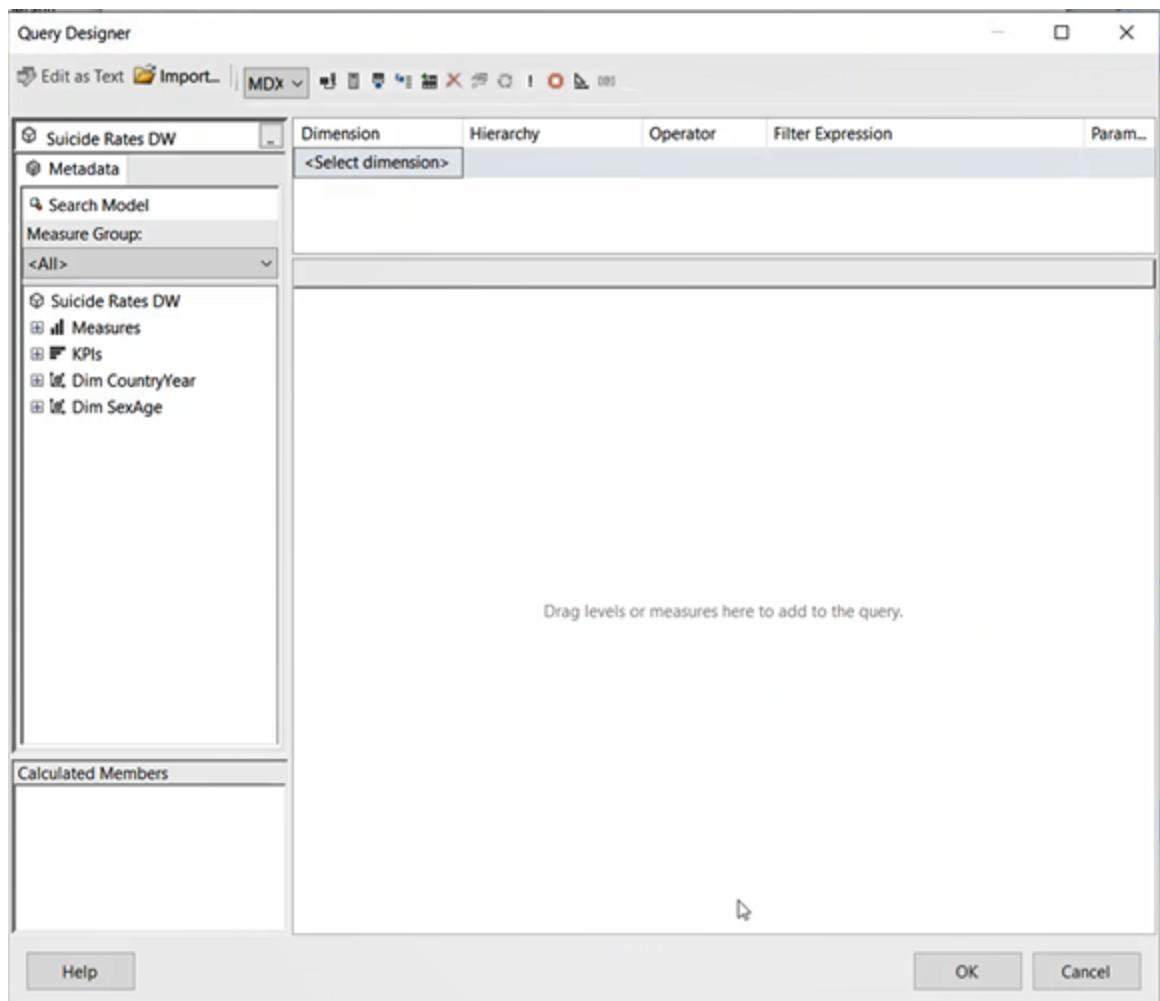
Cửa sổ Dataset Properties xuất hiện.

Đặt tên cho Dataset tại mục Name. Chọn Use a dataset embedded in my report.

Chọn Data source đã tạo ở bước trước, sau đó nháy chọn Query Designer... để tìm những trường thích hợp với câu truy vấn.



Cửa sổ Query Designer xuất hiện



Tương tự như ở Cube Browser trong quá trình SSAS, thực hiện kéo thả các thuộc tính để thực hiện câu truy vấn.

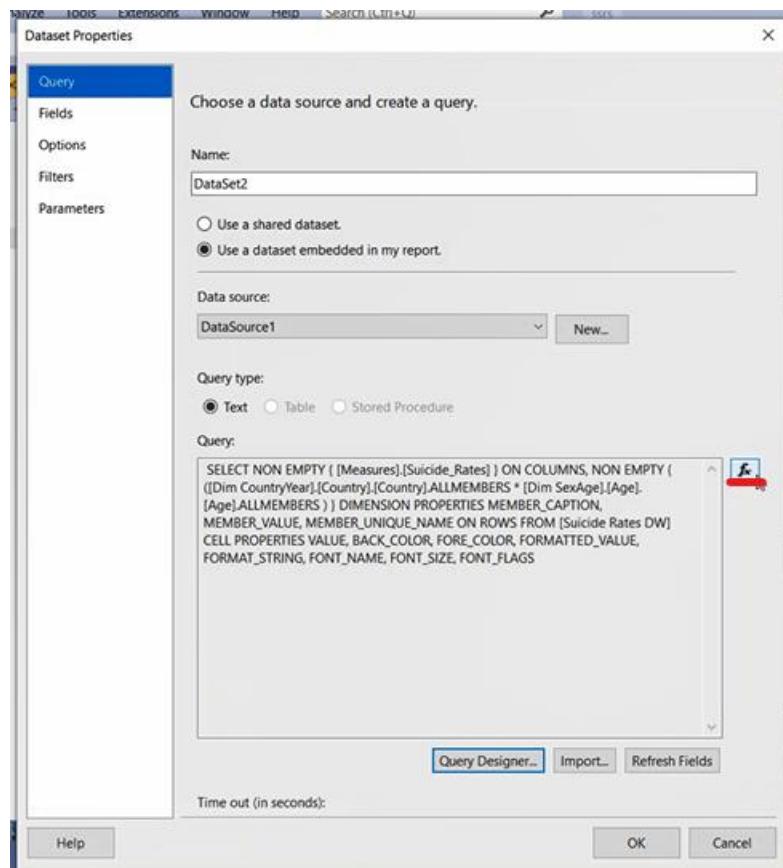
Screenshot of the Dataset Properties dialog box in Analysis Services. The 'Query Designer' tab is selected. In the left pane, under 'Measure Group: <All>', the 'Suicide Rates DW' measure group is expanded, showing 'Measures' (Fact, Suicide_Rates), 'KPIs', and dimensions (Dim CountryYear, Dim SexAge). Under 'Calculated Members', there are none. The main area shows a table with columns 'Country', 'Age', and 'Suicide_Rates'. The table data is as follows:

Country	Age	Suicide_Rates
Albania	15-24 years	4.473586299...
Albania	25-34 years	4.262080627...
Albania	35-54 years	3.700445956...
Albania	5-14 years	0.433199623...
Albania	55-74 years	3.164302449...
Albania	75+ years	4.746484964...
Antigu...	15-24 years	0
Antigu...	25-34 years	0.269685708...
Antigu...	35-54 years	1.116424464...
Antigu...	5-14 years	0
Antigu...	55-74 years	1.886578877...
Antigu...	75+ years	0
Argenti...	15-24 years	9.203282494...
Argenti...	25-34 years	8.025353931...
Argenti...	35-54 years	8.407172426...
Argenti...	5-14 years	0.692214508...
Argenti...	55-74 years	12.27683163...
Argenti...	75+ years	17.56056176...
Armenia	15-24 years	1.221403455...
Armenia	25-34 years	2.172955249...
Armenia	35-54 years	3.081554356...
Armenia	5-14 years	0.176467279...

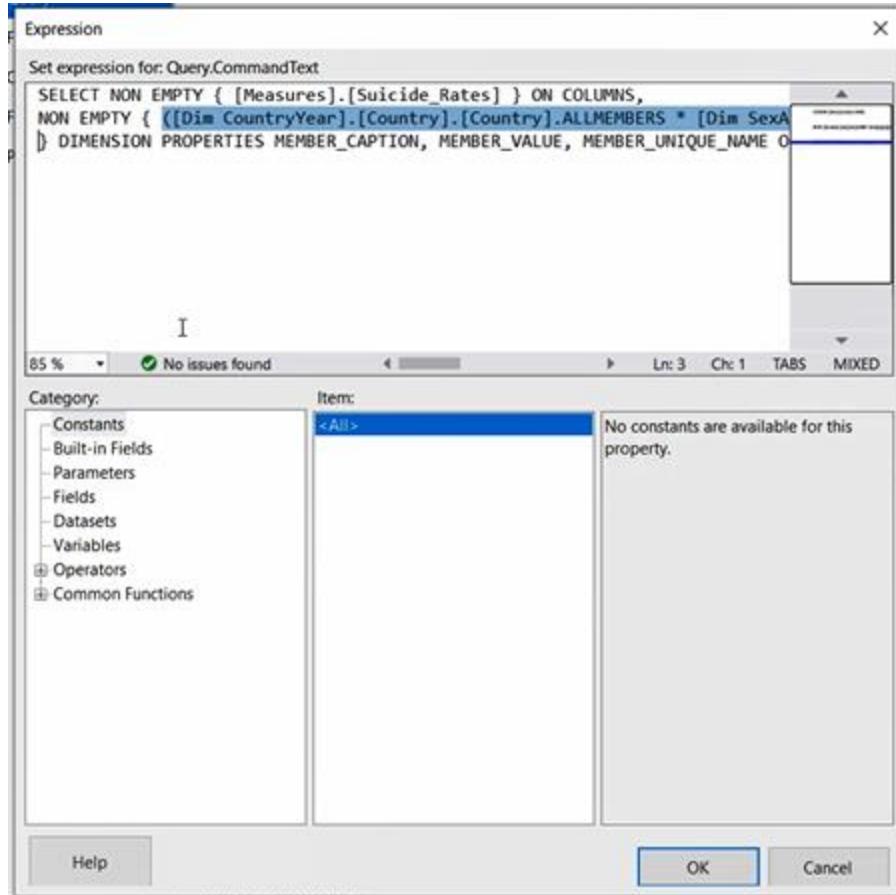
Buttons at the bottom include 'Help', 'OK' (with a cursor arrow pointing to it), and 'Cancel'.

Nháy Chọn OK.

Trở về cửa sổ Dataset Properties nháy chọn nút fx.

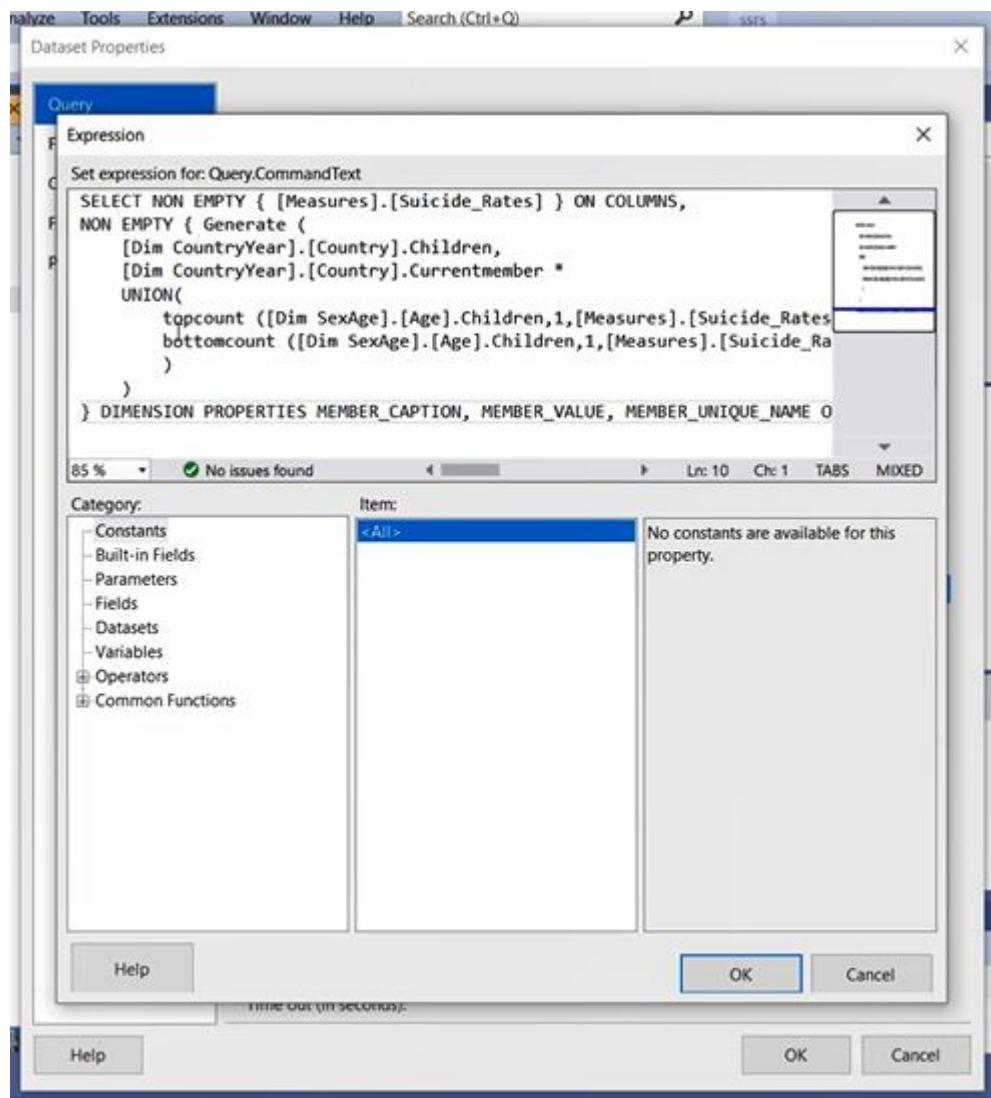


Thay đổi câu truy vấn

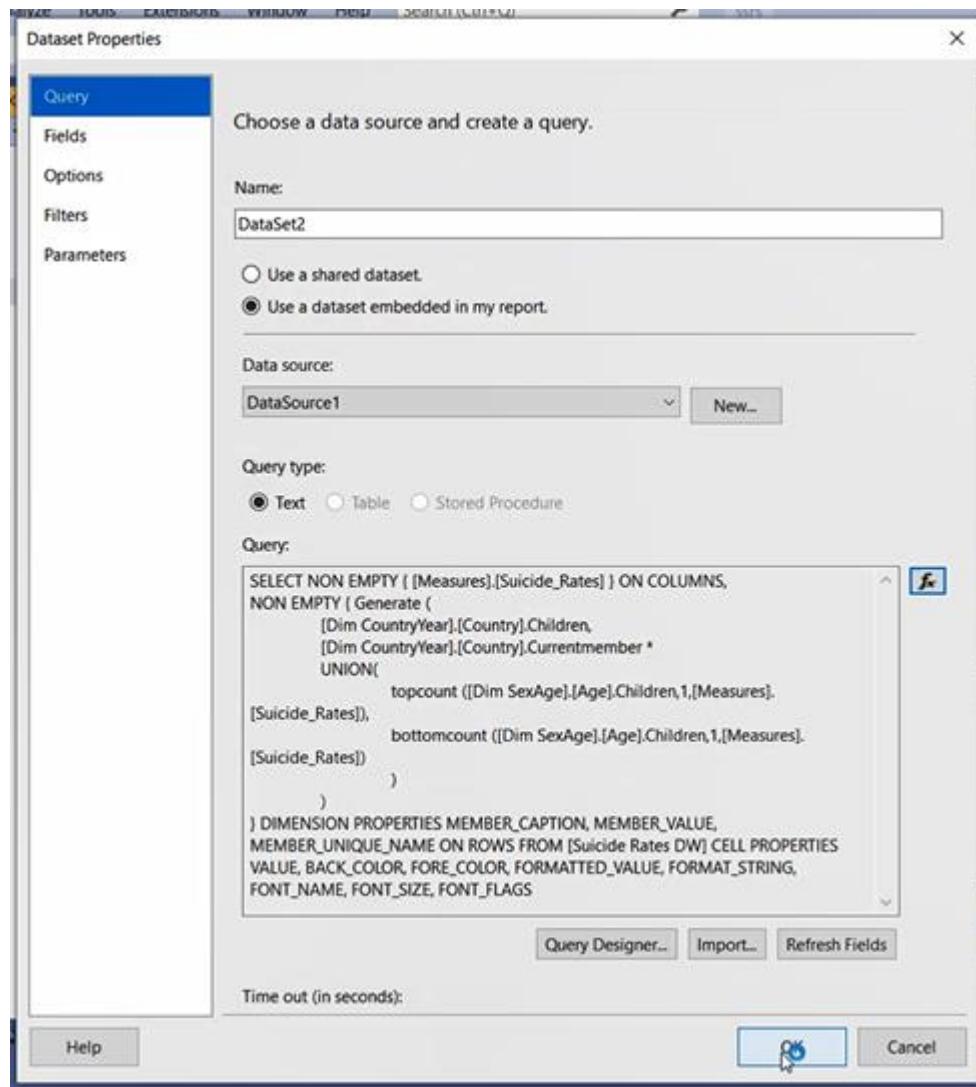


Thay thế bằng câu truy vấn:

```
Generate (  
    [Dim_CountryYear].[Country].Children,  
    [Dim_CountryYear].[Country].Currentmember *  
    UNION(  
        topcount ([Dim_SexAge].[Age].Children, 1, [Measures].[Suicide_Rates]),  
        bottomcount ([Dim_SexAge].[Age].Children, 1, [Measures].[Suicide_Rates])  
    )  
)
```



Nhấn chọn OK để kết thúc

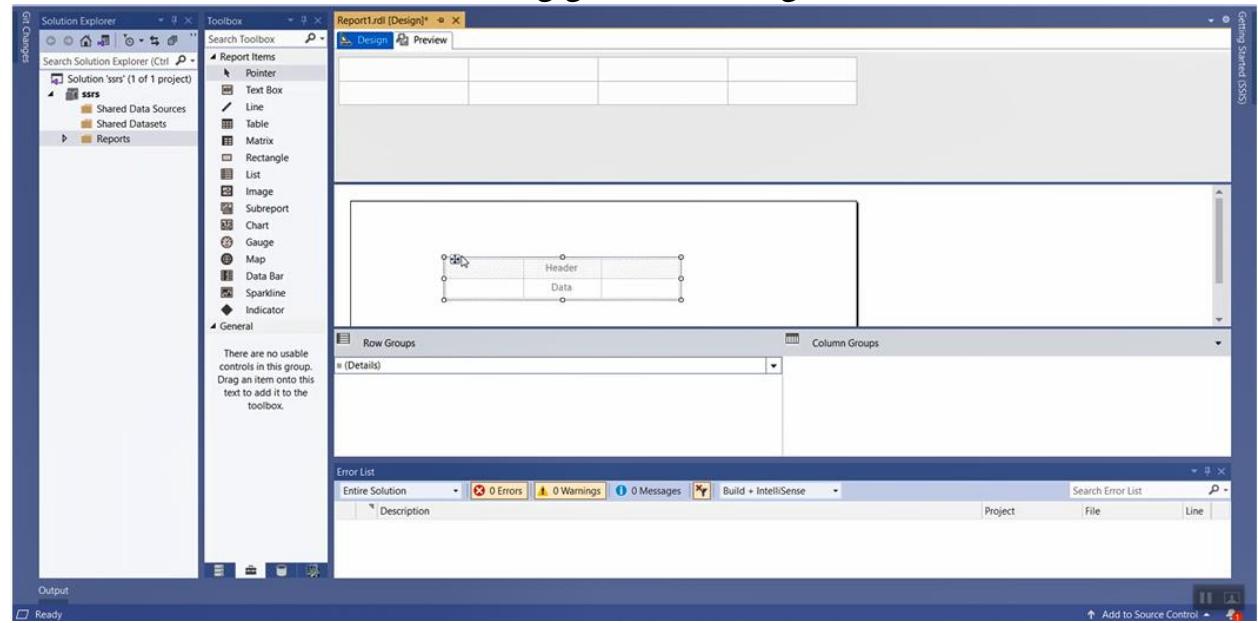


Dataset mới đã được thêm vào mục Dataset.

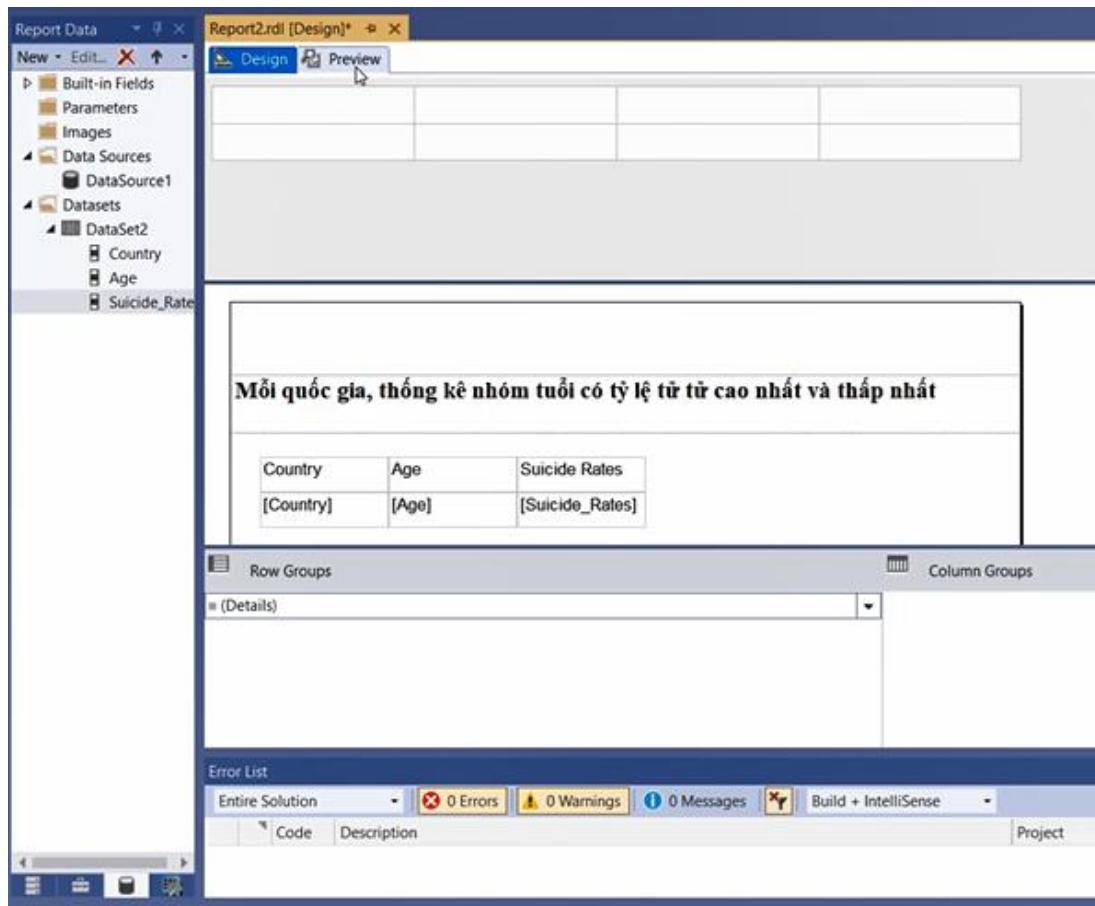
- Thêm dữ liệu vào Report

Sử dụng SSIS Toolbox sử dụng tổ hợp phím Ctrl + Alt + X hoặc chuyển tab tại Report Data

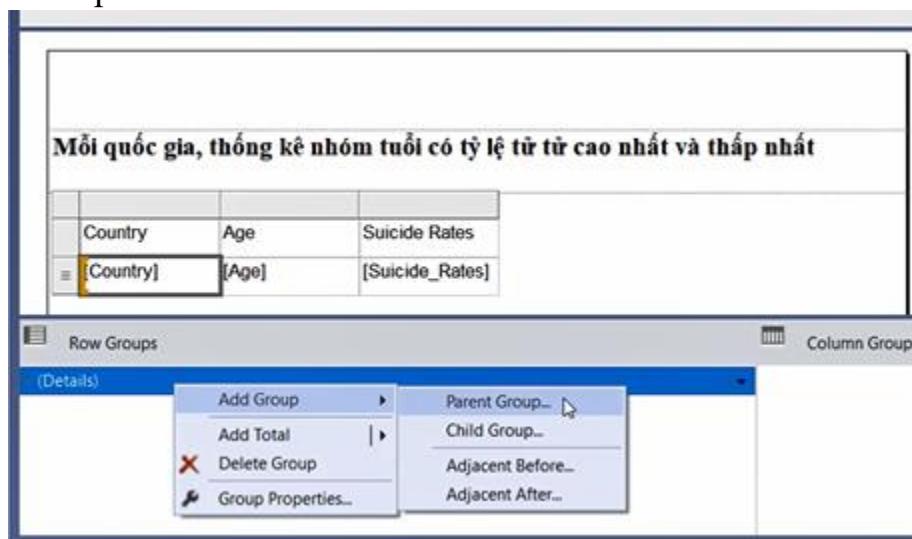
Kéo một Table mới từ Toolbox sang giao diện Design



Kéo những thuộc tính cần thiết vào Table như hình minh họa:



- Tùy chọn Grouping như sau:
Tại Row Groups nháy chuột phải tại (Details) → Add Group → Parent Group...



Cửa sổ mới xuất hiện, chọn Group by [Country], chọn OK

The screenshot shows the Microsoft Report Designer interface. A 'Tablix group' dialog box is open, prompting the user to 'Group by' [Country]. Other options include 'Show detail data', 'Add group header', and 'Add group footer'. The 'OK' button is highlighted. In the background, a report preview window displays a title 'Mỗi quốc gia, thống kê nhóm tuổi có tỷ lệ tử tự cao nhất và thấp nhất' and a table with three columns: Country, Age, and Suicide Rates. The table has three rows: the first row contains the column headers; the second row contains the field names [Country], [Age], and [Suicide_Rates]; and the third row contains the actual data. The 'Row Groups' pane at the bottom shows a single group named 'Country' under the '(Details)' category.

Chỉnh sửa báo cáo cho phù hợp

Tại Preview, chọn Print Layout để hiển thị báo cáo trước khi in.

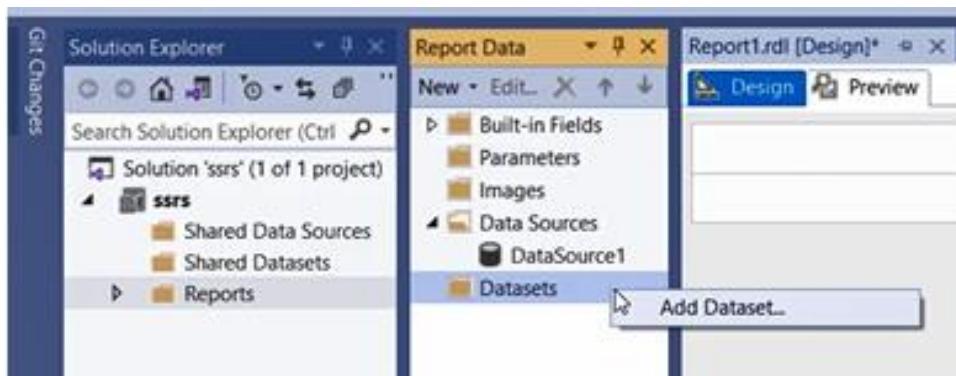
Mỗi quốc gia, thống kê nhóm tuổi có tỷ lệ tử tự cao nhất và thấp nhất

Country	Age	Suicide Rates
Albania	75+ years	4.74648496419041
	5-14 years	0.433199623783881
Antigua and Barbuda	55-74 years	1.88657887786288
	75+ years	0
Argentina	75+ years	17.5605617645497
	5-14 years	0.692214508485246
Armenia	75+ years	6.94013932059992
	5-14 years	0.176467279473804
Aruba	75+ years	14.5847006490192
	5-14 years	0
Australia	25-34 years	17.051450944187
	5-14 years	0.401032925467392
Austria	75+ years	46.1379506723148
	5-14 years	0.510993525426162
Azerbaijan	75+ years	2.82245419112047
	5-14 years	0.2314484429306
Bahamas	35-54 years	2.39234579836113
	5-14 years	0.154885075274147
Bahrain	25-34 years	4.80794263260339

6.3.3. Report3 (Report Totals + Grouping) Top 10 quốc gia có tổng số vụ
tử tự cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính
(Câu truy vấn 15).

Thực hiện phần 6.1 và 6.2 tạo Report4.

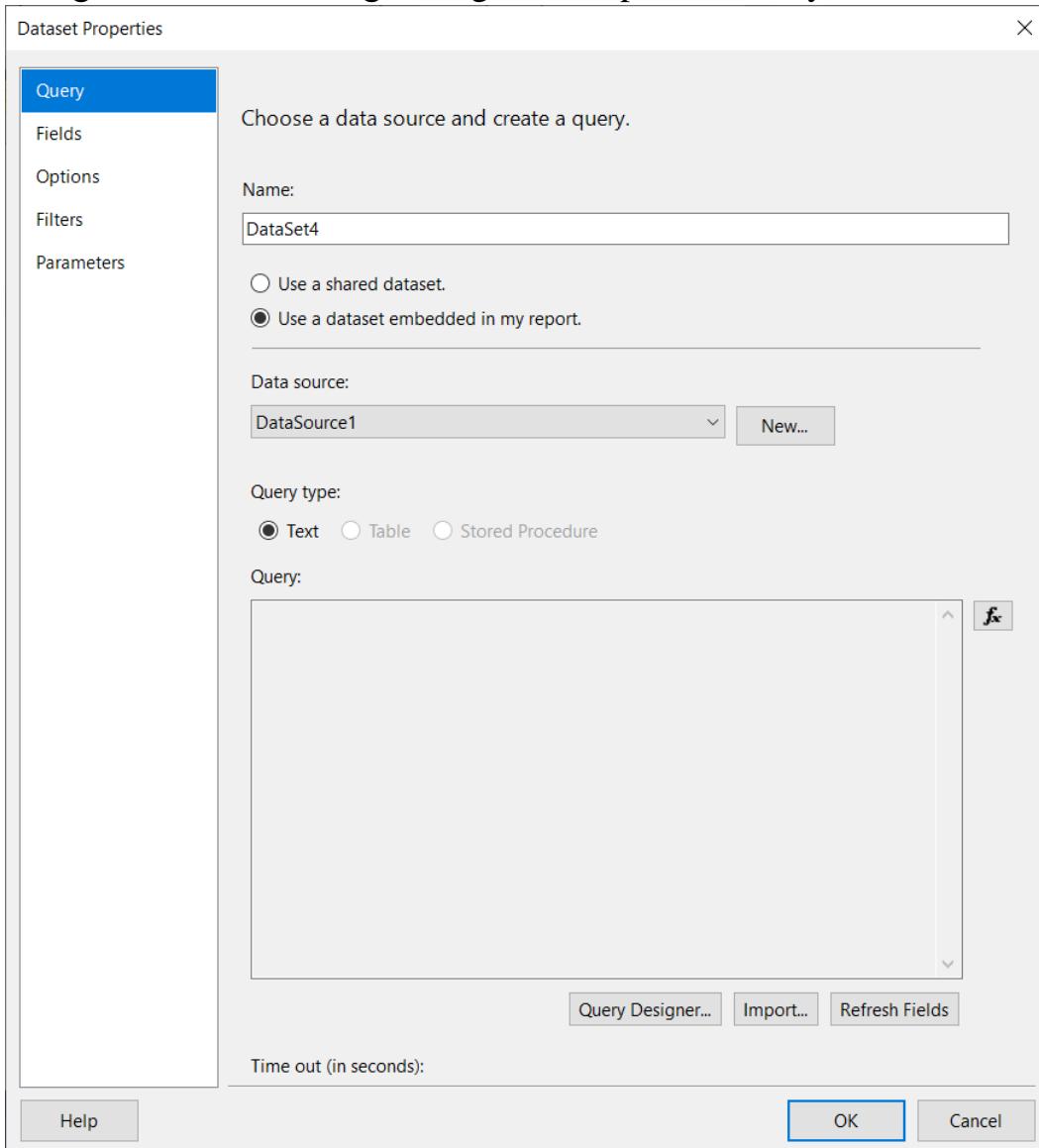
Thêm Datasets cho Report4 bằng cách tại Report Data, ở mục Datasets
nháy chuột phải chọn Add Dataset...



Cửa sổ Dataset Properties xuất hiện.

Đặt tên cho Dataset tại mục Name. Chọn Use a dataset embedded in my report.

Chọn Data source đã tạo ở bước trước, sau đó nháy chọn Query Designer... để tìm những trường thích hợp với câu truy vấn.



Tương tự như ở Cube Browser trong quá trình SSAS, thực hiện kéo thả các thuộc tính để thực hiện câu truy vấn.

The screenshot shows the Microsoft Analysis Services Query Designer window. On the left, there's a navigation pane with 'Suicide Rates DW' selected, followed by 'Metadata', 'Search Model', and 'Measure Group'. Below these are sections for 'All' (containing 'Suicide_Rates', 'KPIs', 'Dim CountryYear' with children 'cau10' through 'cau13', 'cau2', 'cau4', 'cau5', 'cau6', 'Country', 'CountryYear', 'Year', 'Dim SexAge' with children 'Age', 'Sex', and 'SexAgeID'), 'Calculated Members', and 'Help'. The main area displays an MDX query:

Dimension	Hierarchy	Operator	Filter Expression	Param...
Dim CountryYear	Year	Range (Inclu...)	2010 : 2016	<input type="checkbox"/>
Dim CountryYear	Country	MDX	{Head(Order([Dim CountryYear].[Country].Children, [Measures].[Suicides No], desc), 10)}	<input type="checkbox"/>

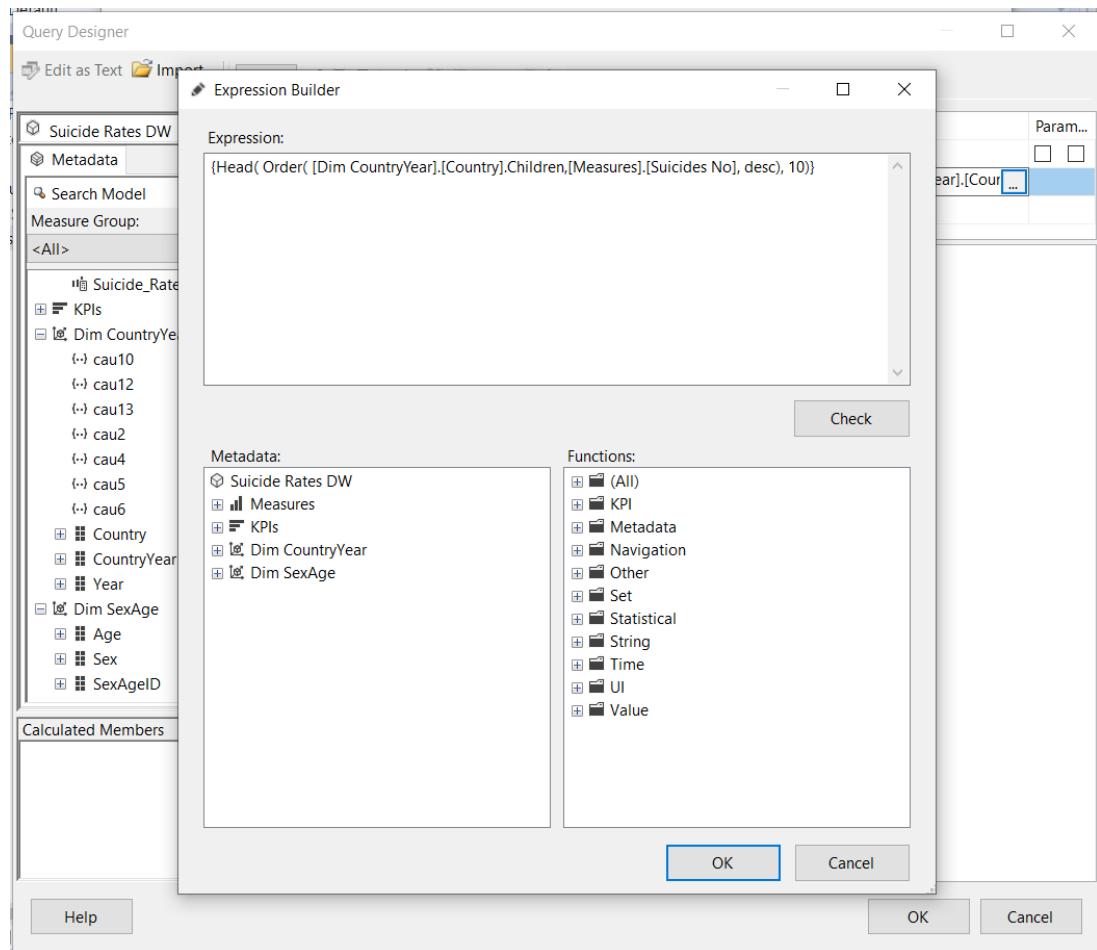
Below the query is a table showing the results:

Country	Sex	Suicides No
Brazil	female	13268
Brazil	male	48585
France	female	12368
France	male	36643
Germany	female	15576
Germany	male	44899
Japan	female	47820
Japan	male	110135
Mexico	female	6541
Mexico	male	27379
Poland	female	4784
Poland	male	31626
Republic of Korea	female	27235
Republic of Korea	male	60158
Russian Federation	female	31019
Russian Federation	male	143681
Ukraine	female	7922
Ukraine	male	34740
United States	female	54370
United States	male	192197

At the bottom right are 'OK' and 'Cancel' buttons.

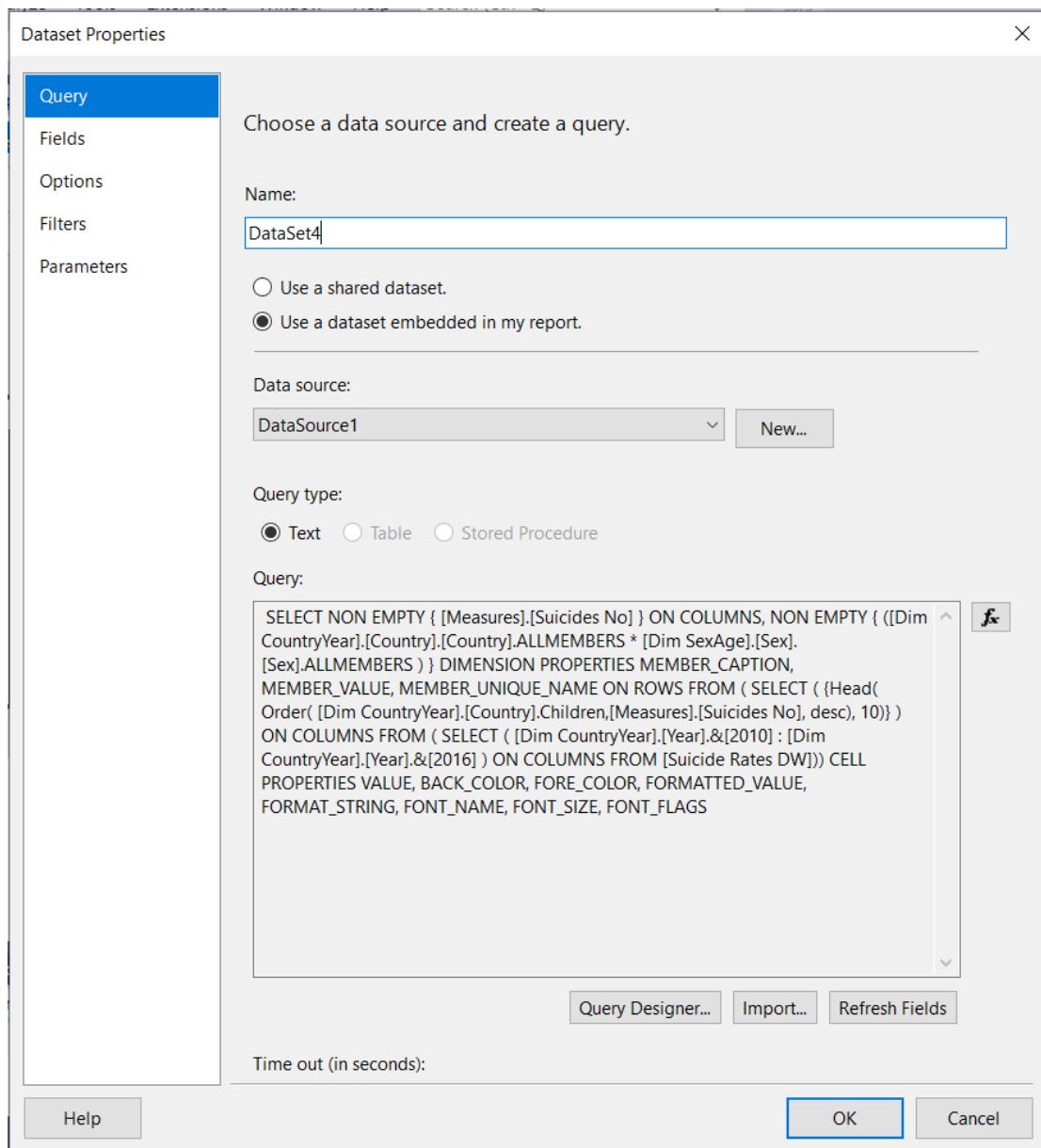
Cửa sổ Expression Builder hiển thị, sau đó thêm câu truy vấn MDX lấy Top10 quốc gia:

```
{Head( Order( [Dim CountryYear].[Country].Children, [Measures].[Suicides No], desc), 10)}
```



Nháy Chọn OK.

Trở về cửa sổ Dataset Properties nháy chọn OK.

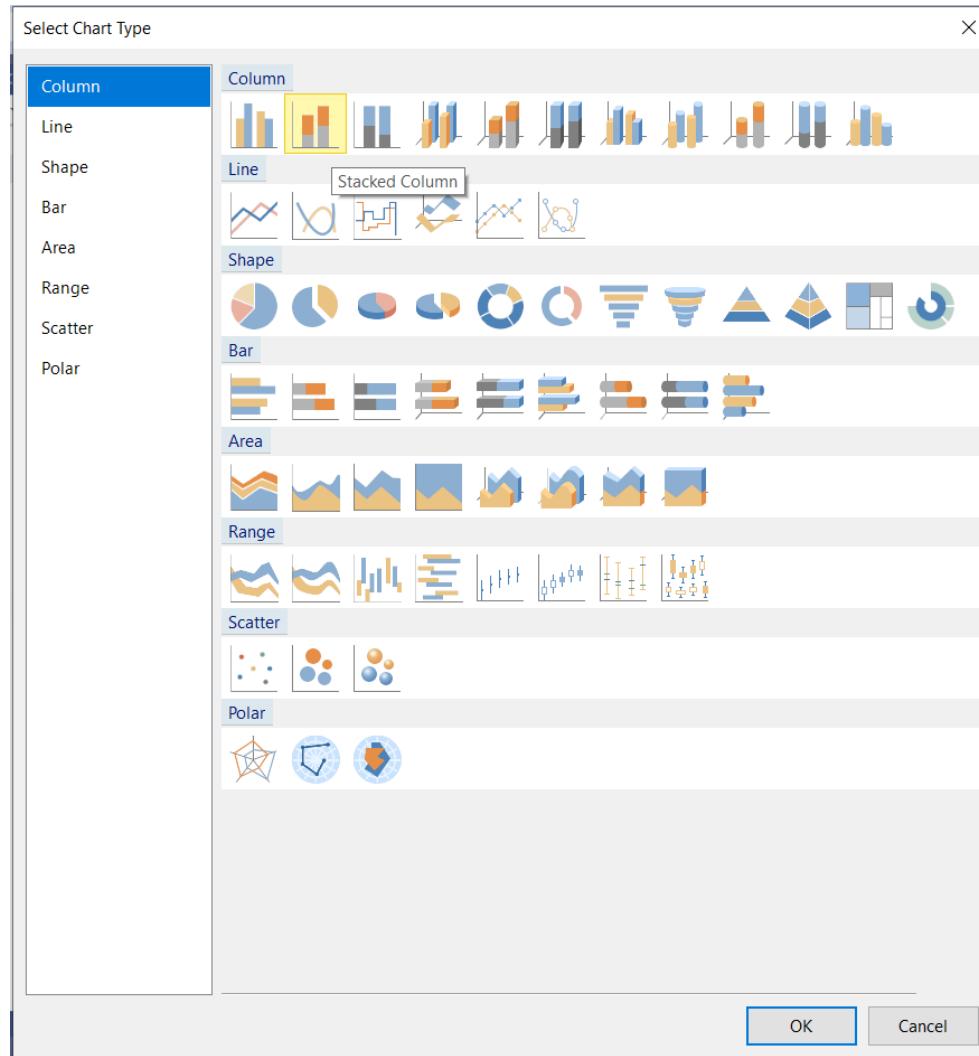


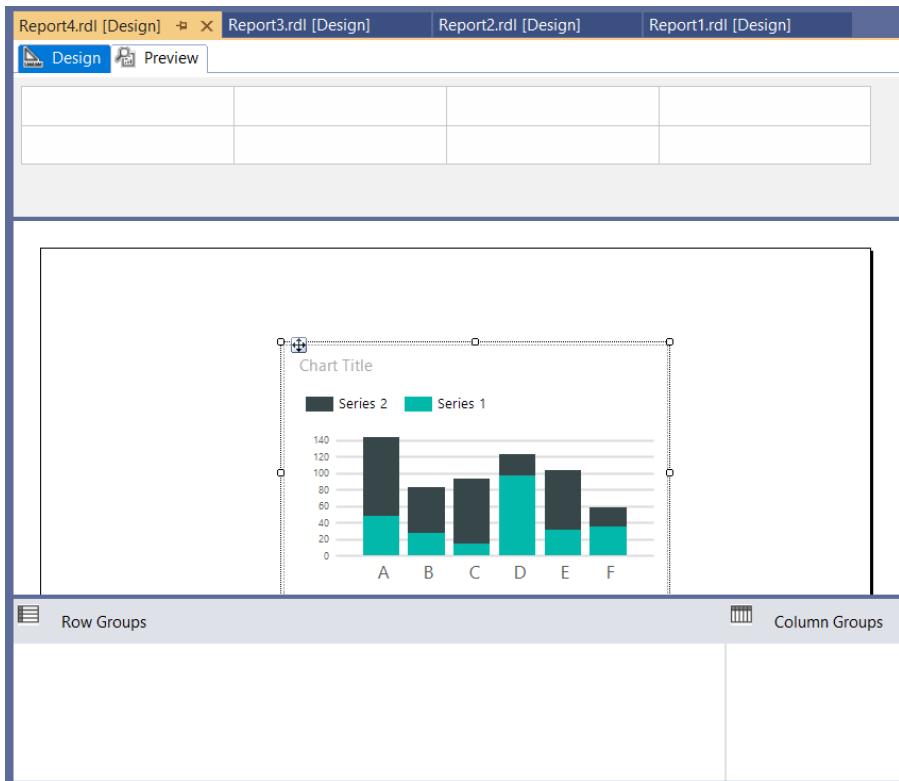
Dataset mới đã được thêm vào mục Dataset.

- Thêm dữ liệu vào Report

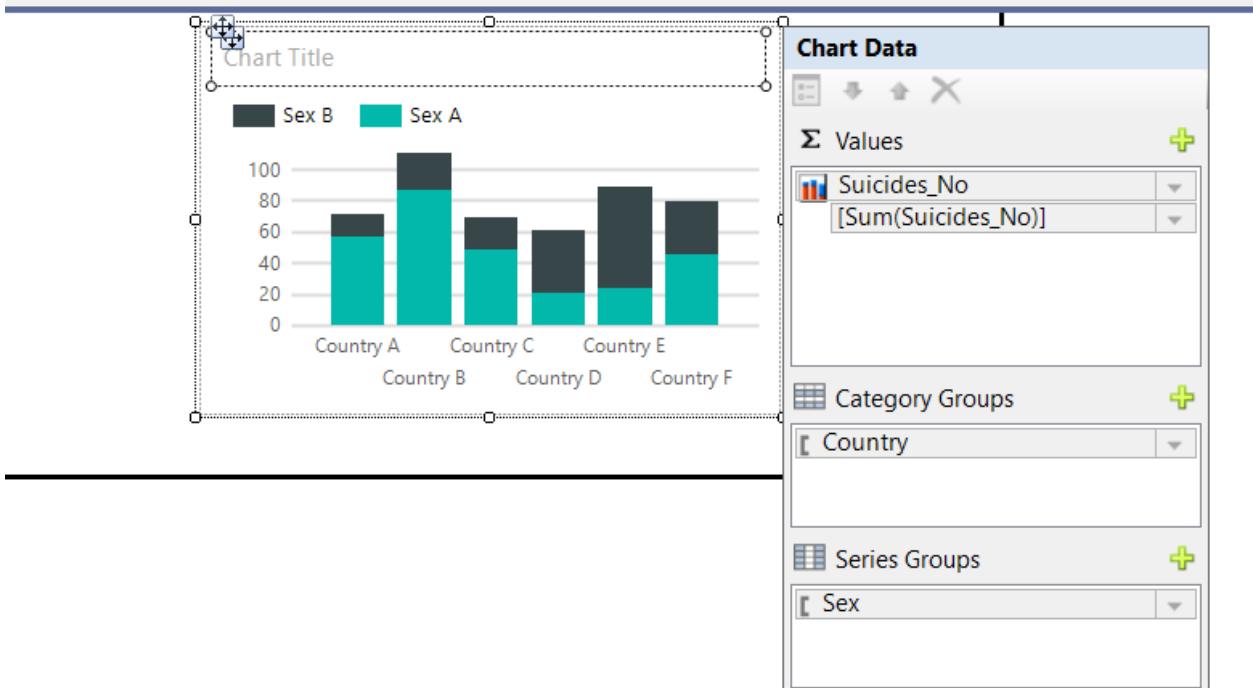
Kéo một Chart mới từ Toolbox sang giao diện Design

Tại cửa sổ Select Chart Type, chọn đồ thị Stacked Column, sau đó nháy chọn OK.





Nháy chọn vào đồ thị Chart Data xuất hiện, kéo thả những thuộc tính cần thiết vào Chart Data như hình minh họa:

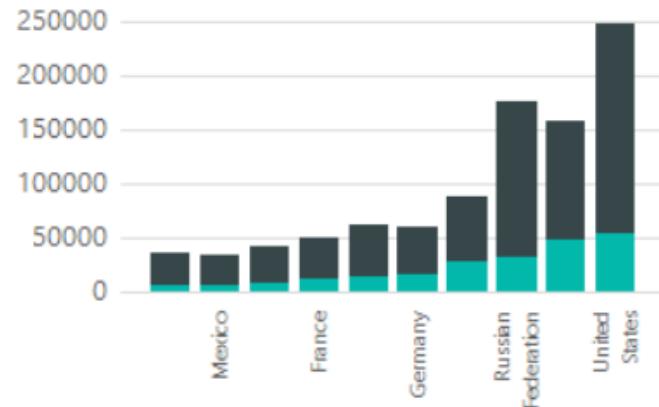


Chuyển sang tab Preview để xem Report

Top 10 quốc gia có tổng số vụ tự tử cao nhất từ 2010-2016, thống kê theo giới tính.

Suicides No by Country and Sex

male female



Kéo một Table mới từ Toolbox sang giao diện Design

Kéo những thuộc tính cần thiết vào Table như hình minh họa:

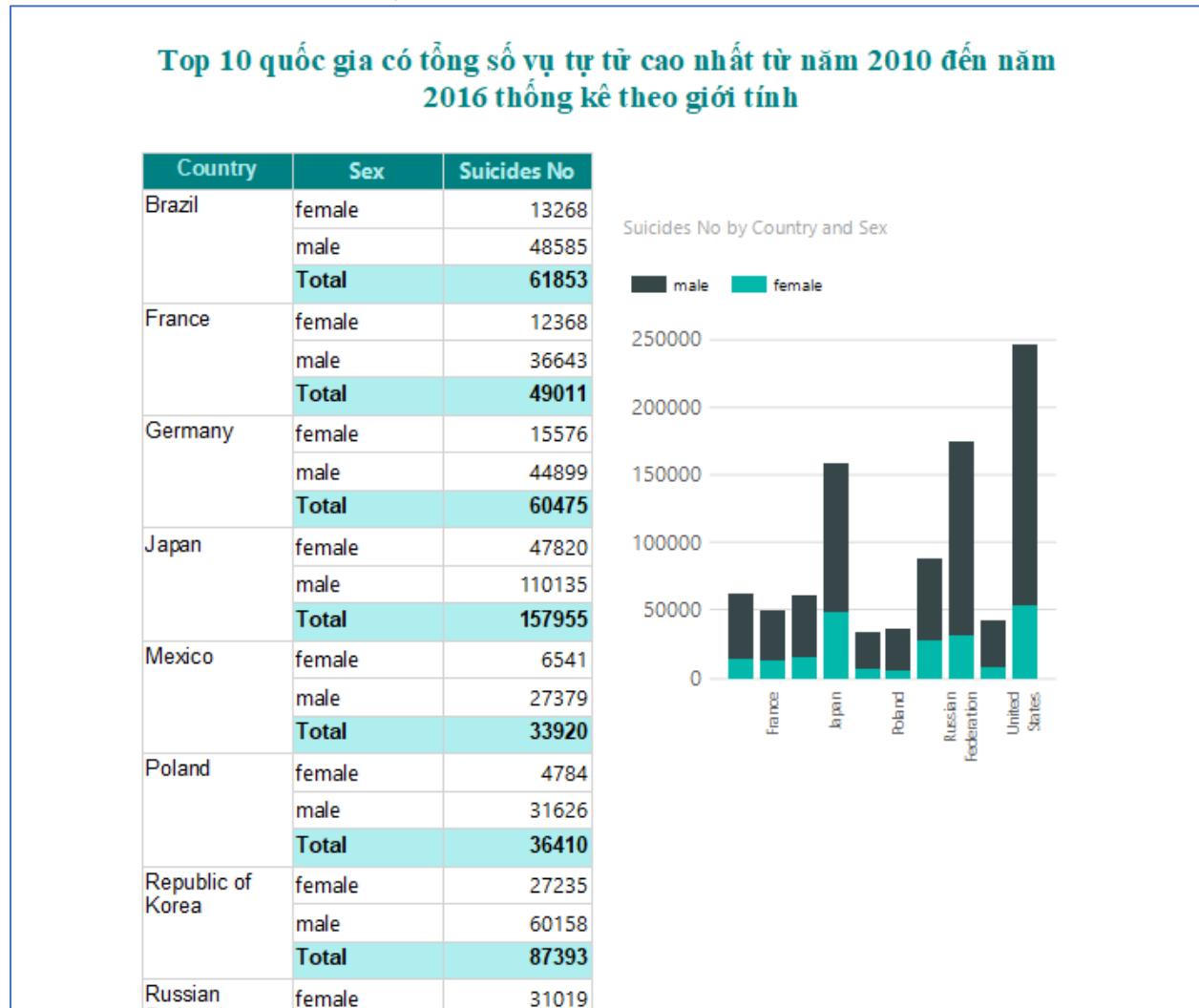
The screenshot shows the Report Designer interface with the following details:

- Report Data View:** Shows built-in fields, parameters, images, data sources (DataSource1), and datasets (DataSet4 with columns: Country, Sex, Suicides_No).
- Report Structure View:** Shows five reports: Report4.rdl [Design]*, Report3.rdl [Design], Report2.rdl [Design], Report1.rdl [Design]. The current report is Report4.rdl [Design].
- Report Preview View:** Shows a blank table structure.
- Report Content View:** Contains:
 - A title: "thống kê theo giới tính."
 - A table:

Country	Sex	Suicides No
[Country]	[Sex]	[Suicides_No]
		[Sum(Suicides_No)]
 - A chart titled "Suicides No by Country and Sex". The legend indicates "Sex B" (dark grey) and "Sex A" (teal). The chart shows data for three countries: Country A (~100 total), Country C (~65 total), and Country F (~85 total). The chart is currently labeled with "Sex B" and "Sex A" instead of "male" and "female".
- Report Filter View:** Shows filters for "Country" and "(Details)".

Tiếp tục định dạng lại Report

Tại Preview, chọn Print Layout để hiển thị báo cáo trước khi in.



7 Quá trình lập báo biểu bằng công cụ Power BI

Microsoft Power BI là một phần mềm sử dụng dữ liệu để đưa ra các báo cáo và hiển thị thông tin chi tiết của doanh nghiệp.

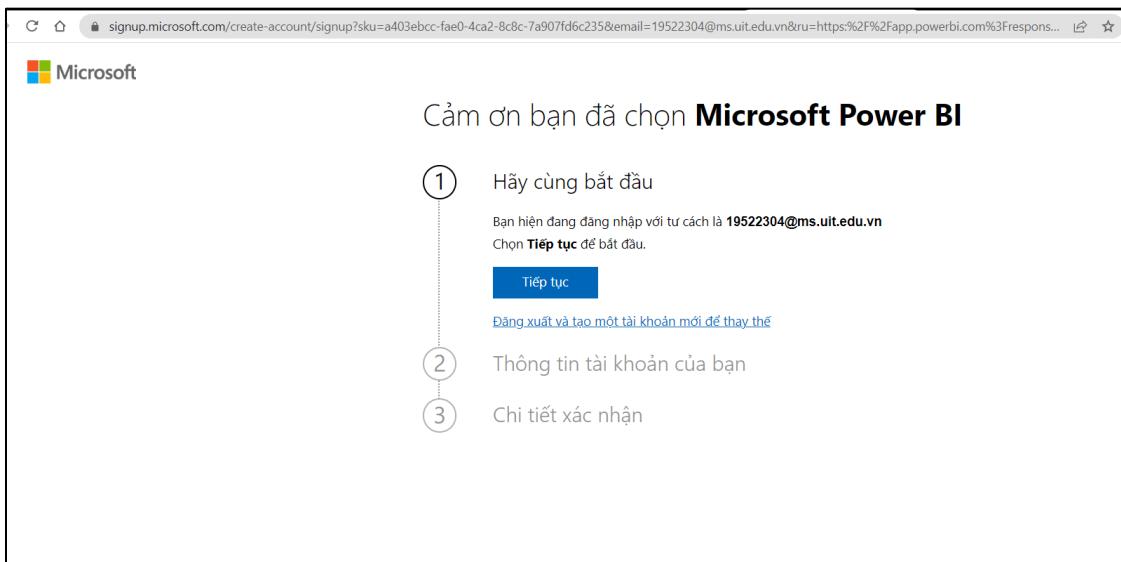
Power BI có khả năng kết nối nhiều loại tập dữ liệu lại với nhau và báo cáo lại cho bạn những dữ liệu được sắp xếp gọn gàng và trực quan nhất.

7.1. Đăng ký dịch vụ Power BI

Truy cập vào đường dẫn: <https://docs.microsoft.com/en-us/power-bi/fundamentals/service-self-service-signup-for-power-bi>

Tiếp tục truy cập vào đường dẫn tại Step 1: <https://powerbi.com/>

Màn hình đăng ký xuất hiện, sau đó tiến hành điền các thông tin cần thiết



Cảm ơn bạn đã chọn Microsoft Power BI

1 Thiết lập tài khoản

2 Thông tin tài khoản của bạn

Chào **Thư** !
Trước khi bắt đầu, bạn cần cung cấp cho chúng tôi thêm một chút thông tin

19522304@ms.uit.edu.vn

Quốc gia hoặc khu vực của bạn
Việt Nam

Số điện thoại

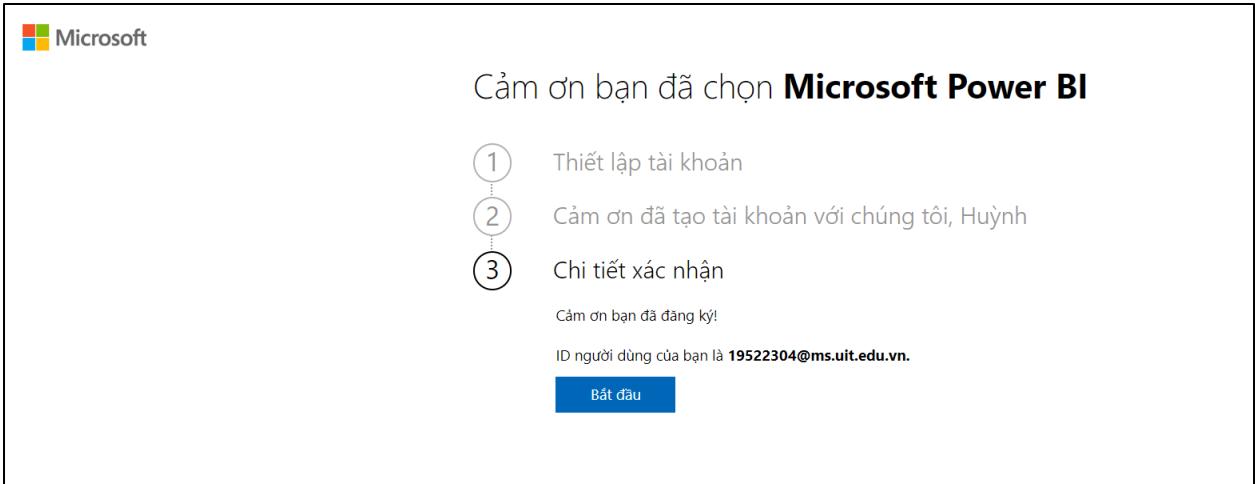
Tôi hiểu rằng Microsoft có thể liên hệ với tôi để trao đổi thông tin về bản dùng thử.

Tôi muốn nhận thông tin, mẹo và ưu đãi được cá nhân hóa về Power BI, Giải pháp cho doanh nghiệp và tổ chức, cũng như các sản phẩm và dịch vụ khác của Microsoft. [Điều khoản về quyền riêng tư](#).

Tôi muốn Microsoft chia sẻ thông tin của tôi với một số đối tác để tôi có thể nhận thông tin liên quan đến sản phẩm và dịch vụ của họ. Để tìm hiểu thêm, hãy xem [Điều khoản về quyền riêng tư](#).

Bằng cách chọn **Bắt đầu**, bạn đồng ý với [các điều khoản và điều kiện của chúng tôi](#) và [điều khoản về quyền riêng tư](#).

Bắt đầu



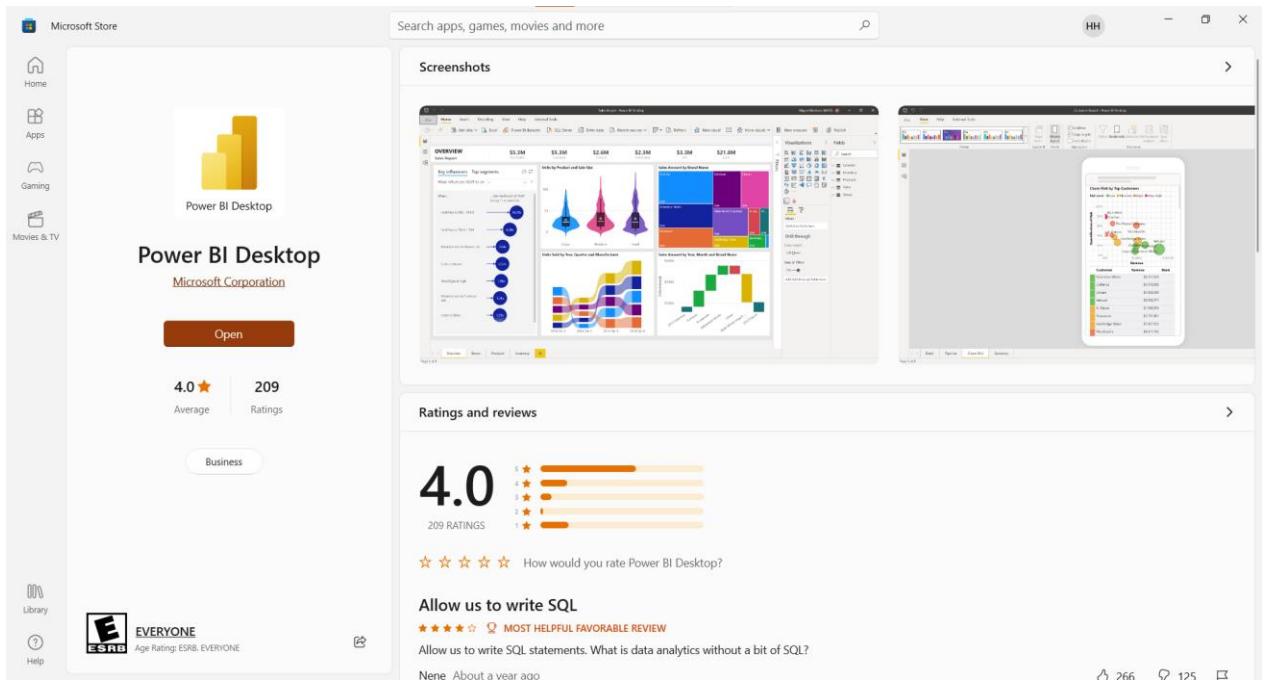
Màn hình sau khi đăng ký thành công.

A screenshot of the Microsoft Power BI homepage. The left sidebar shows navigation options like Trang chủ, Mục yêu thích, Gần đây, Tạo, Tập dữ liệu, Ứng dụng, Được chia sẻ với tôi, Tìm hiểu, Không gian làm việc, and Không gian làm việc. The main area has a greeting "Chào bạn, Thư" and a search bar. Below is a section titled "Được đề xuất" with five cards: "Khám phá khái niệm Power BI cơ bản", "Giới thiệu—Power BI là gì?", "Bắt đầu nhanh—Khám phá dịch vụ", "Xem và hiểu báo cáo Power BI", and "Cách để biết nó". At the bottom are filters for "Gần đây", "Mục yêu thích", "Ứng dụng của tôi", a search bar, and a "Lọc theo từ khóa" button.

7.2. Tải Power BI Desktop trên Microsoft Store

Sau khi đăng ký thành công, mở Microsoft Store tìm Power BI Desktop

Hoàn tất quá trình tải nhấn Open để mở phần mềm



Đăng nhập bằng tài khoản đã đăng ký trước đó

Untitled - Power BI Desktop

File Home Insert Modeling View Help

Clipboard

Data

Queries

Insert

Calculated columns

Sensitivity

Publish

Collaborate and share

Sign in

Enter your email address

Once I sign in, I can:

- Import data from Excel
- Get data from another source →

19522304@ms.uit.edu.vn

Continue Cancel

Build visual

Filters

Values

Add data fields here

Drill through

Cross-report

Keep all filters

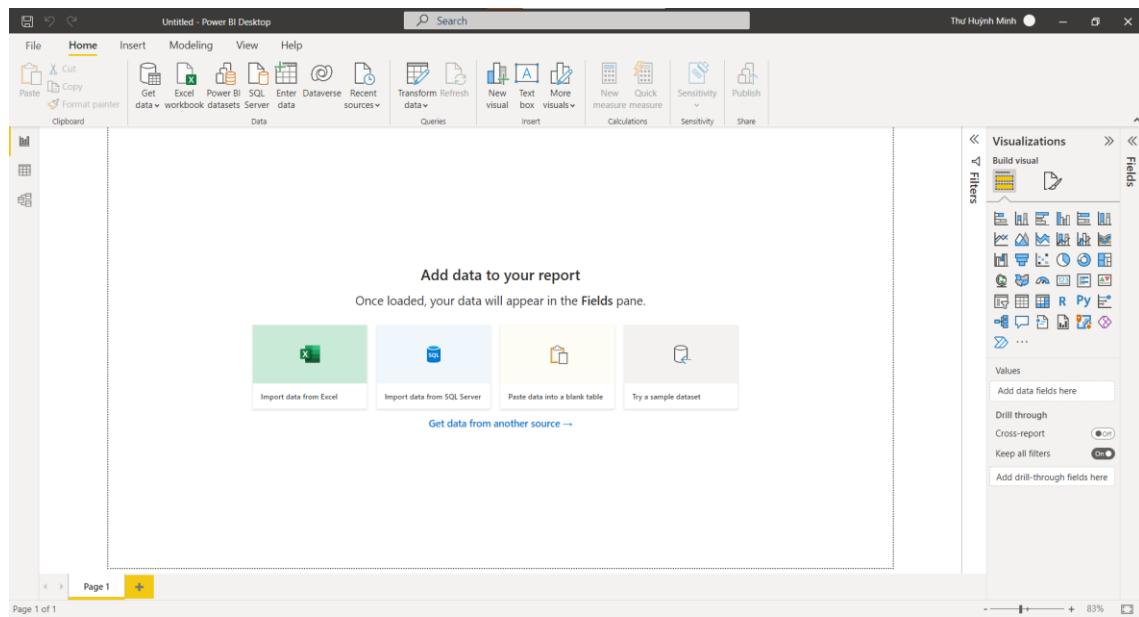
Add drill-through fields here

Page 1

Page 1 of 1

83%

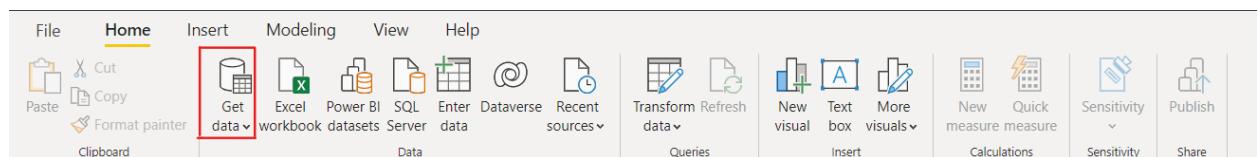
Màn hình giao diện chính xuất hiện



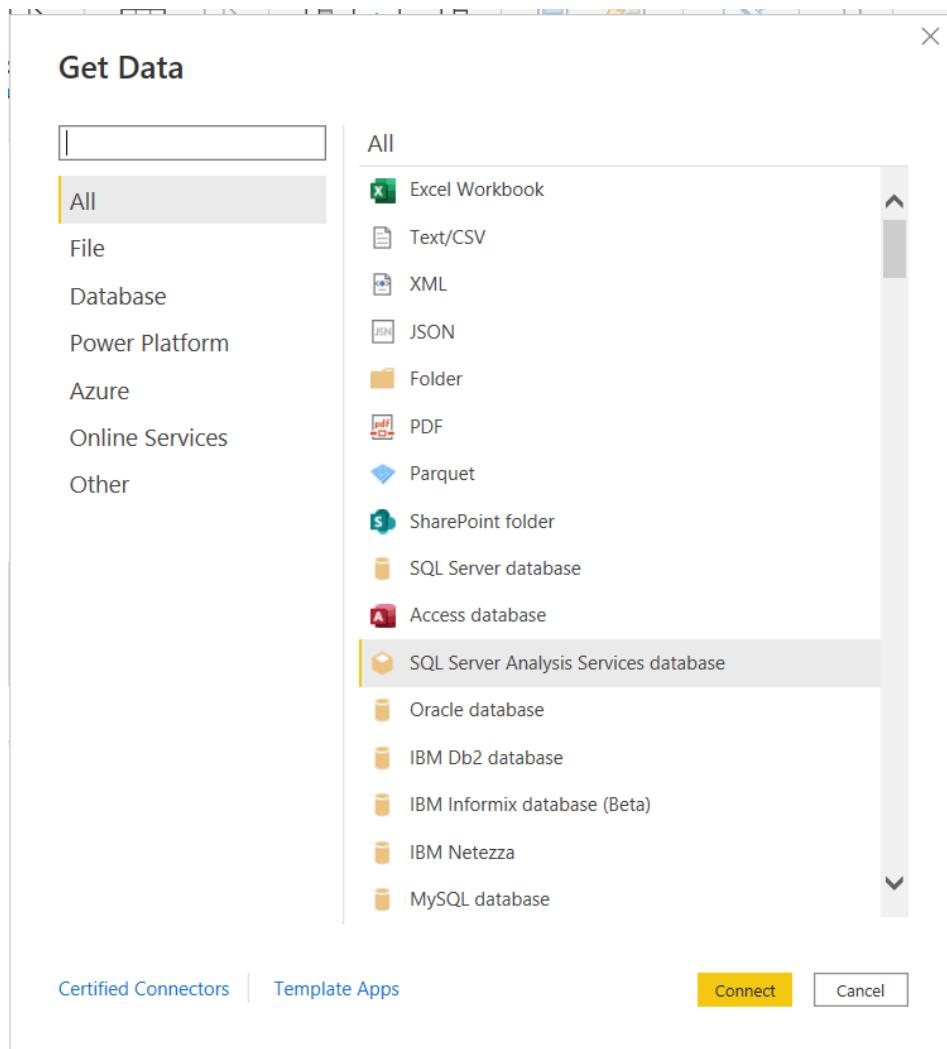
7.3. Tạo Report trên Power Bi Desktop

7.3.1. Report1 (Report Totals) Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính.

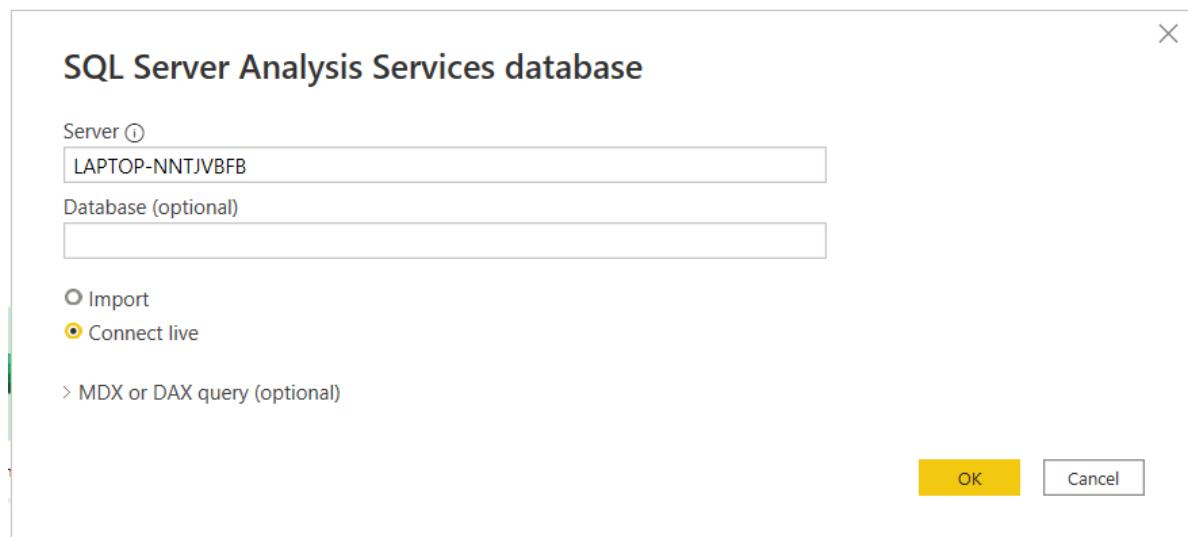
Chọn Get Data → Chọn nguồn dữ liệu để import



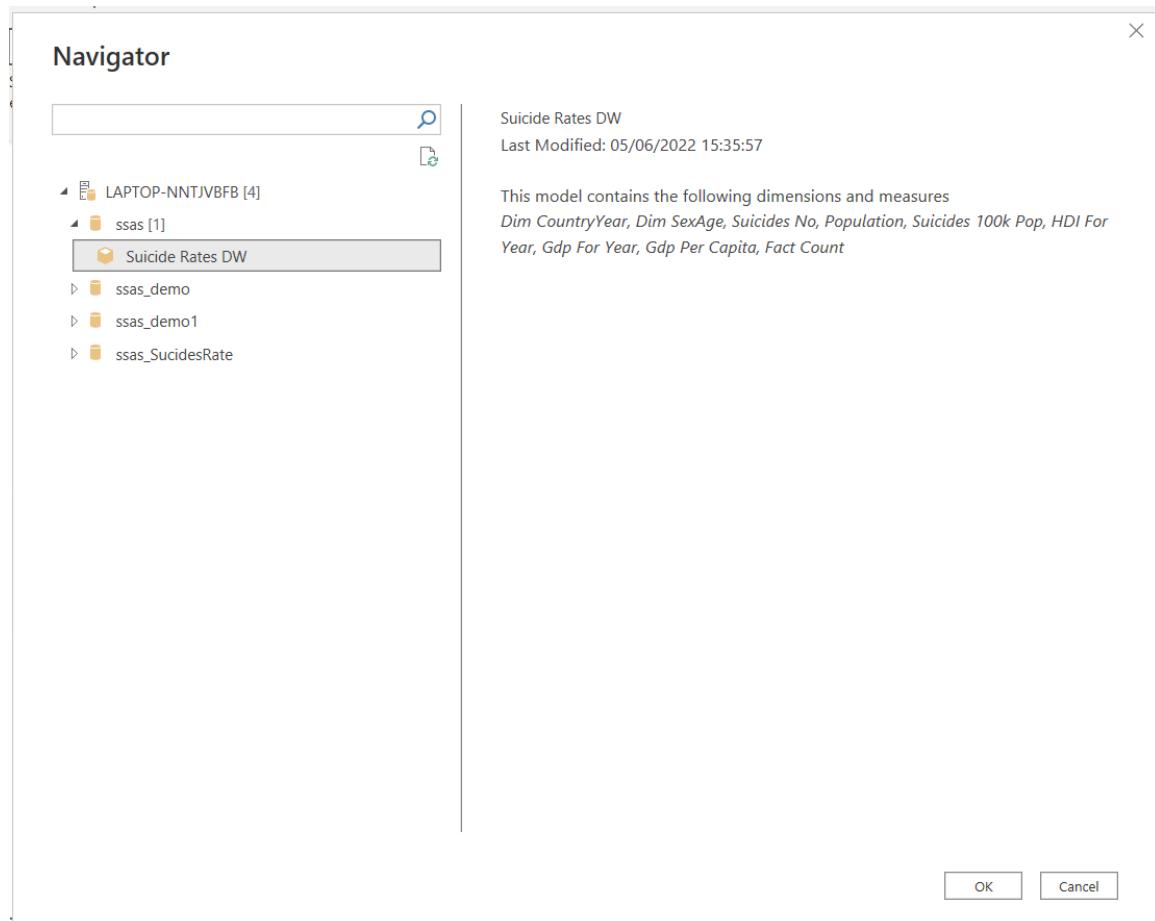
Chọn nguồn dữ liệu là SQL Server Analysis Services database



Nhập tên Server và chọn Connect live → chọn OK



Chọn khôi dữ liệu cần import



Kết nối thành công, các Measure và Dimension sẽ xuất ở Fields

Thực hiện kéo thả các các Dimension và Measure cần thiết cho câu truy vấn

The screenshot shows the Power BI Desktop interface. On the left, there is a table visual titled "SexAgeID_Age_Sex_Suicides No" with the following data:

SexAgeID	Age	Sex	Suicides No
1	15-24 years	female	5940
2	25-34 years	female	7080
3	35-44 years	female	19396
4	45-54 years	female	698
5	55-74 years	female	14071
6	75+ years	female	7326
7	15-24 years	male	20340
8	25-34 years	male	29545
9	35-44 years	male	64496
10	45-54 years	male	1144
11	55-74 years	male	44372
12	75+ years	male	17252
Total			230160

The "Filters" pane on the right shows filters applied to the visual: "Year is 2012". The "Fields" pane lists various dimensions and measures, including "Fact", "Dim CountryYear", "Dim SexAge", and "Measures".

Thiết kế lại Report bằng chức năng Visualizations

The screenshot shows the "Visualizations" pane in Power BI Desktop. It includes sections for "Format visual", "Search", and "Visual". Under "Visual", the "General" tab is selected, showing options for "Legend" (On), "Slices", "Detail labels" (On), and "Rotation". The "Fields" pane on the right lists the same dimensions and measures as the previous screenshot.

Trong năm 2012, tính tổng số vụ tự tử ở từng nhóm tuổi và giới tính

Year	Age	Sex	Suicides No
2012	75+ years	female	7326
2012	75+ years	male	17252
2012	55-74 years	female	14071
2012	55-74 years	male	44372
2012	5-14 years	female	608
2012	5-14 years	male	1144
2012	35-54 years	female	17986
2012	35-54 years	male	64496
2012	25-34 years	female	7080
2012	25-34 years	male	29545
2012	15-24 years	female	5940
2012	15-24 years	male	20340
	Total		230160

- Tạo biểu đồ biểu diễn dữ liệu đơn giản
Thực hiện kéo thả biểu đồ hình tròn từ trường Visualizations

Visualizations > Fields

Build visual

Fact

Dim CountryYear

Dim SexAge

Kéo thả các thuộc tính cho câu truy vấn như hình minh họa:

The screenshot shows the Power BI interface with three main panes: Filters, Visualizations, and Fields.

Filters Pane:

- Filters on this visual:**
 - SexAgeID is (All)
 - Suicides No is (All)
 - Year** is 2012
- Add data fields here**
- Filters on this page:**
 - Add data fields here
- Filters on all pages:**
 - Add data fields here

Visualizations Pane:

Build visual icons: Bar chart, Line chart, Map, Gauge, etc.

Fields Pane:

- Fact:**
 - Fact Count
 - Gdp For Year
 - Gdp Per Capita
 - HDI For Year
 - Population
 - Suicides 100k Pop
 - Suicides No
- Dim CountryYear:**
 - Country
 - CountryYear
 - Year
- Dim SexAge:**
 - Age
 - Sex
 - SexAgeID

report1 - Power BI Desktop

File Home Insert Modeling View Help

Cut Copy Format painter Paste Get data Data Transform Refresh data Queries New visual Text box insert More visuals Calculations Sensitivity Publish Share

Trong năm 2012, tổng số vụ tự tử ở từng nhóm tuổi và giới tính

Year	Age	Sex	Suicides No
2012	75+ years	female	7326
2012	75+ years	male	17252
2012	55-74 years	female	14071
2012	55-74 years	male	44372
2012	5-14 years	female	608
2012	5-14 years	male	1144
2012	35-54 years	female	17986
2012	35-54 years	male	64496
2012	25-34 years	female	7080
2012	25-34 years	male	29545
2012	15-24 years	female	5940
2012	15-24 years	male	20340
Total			230160

Suicides No by SexAgeID and Year

Filters

Visualizations

Fields

Fact

- Fact Count
- Gdp For Year
- Gdp Per Capita
- HDI For Year
- Population
- Suicides 100k Pop
- Suicides No

Dim CountryYear

- Country
- CountryYear
- Year

Dim SexAge

- Age
- Sex
- SexAgeID

Values

Drill through

Cross-report

Keep all filters

Add drill-through fields here

Page 1 Page 2

Page 1 of 2

Live connection: Connected 100%

Report thu được

Trong năm 2012 tổng số vụ tự tử ở từng nhóm tuổi và giới tính

SexAgeID	Age	Sex	Suicides No
1	15-24 years	female	5940
2	25-34 years	female	7080
3	35-54 years	female	17986
4	5-14 years	female	608
5	55-74 years	female	14071
6	75+ years	female	7326
7	15-24 years	male	20340
8	25-34 years	male	29545
9	35-54 years	male	64496
10	5-14 years	male	1144
11	55-74 years	male	44372
12	75+ years	male	17252
Total			230160

Suicides No by SexAgeID and Year

Ngày 19, tháng 05, năm 2022
Người tạo

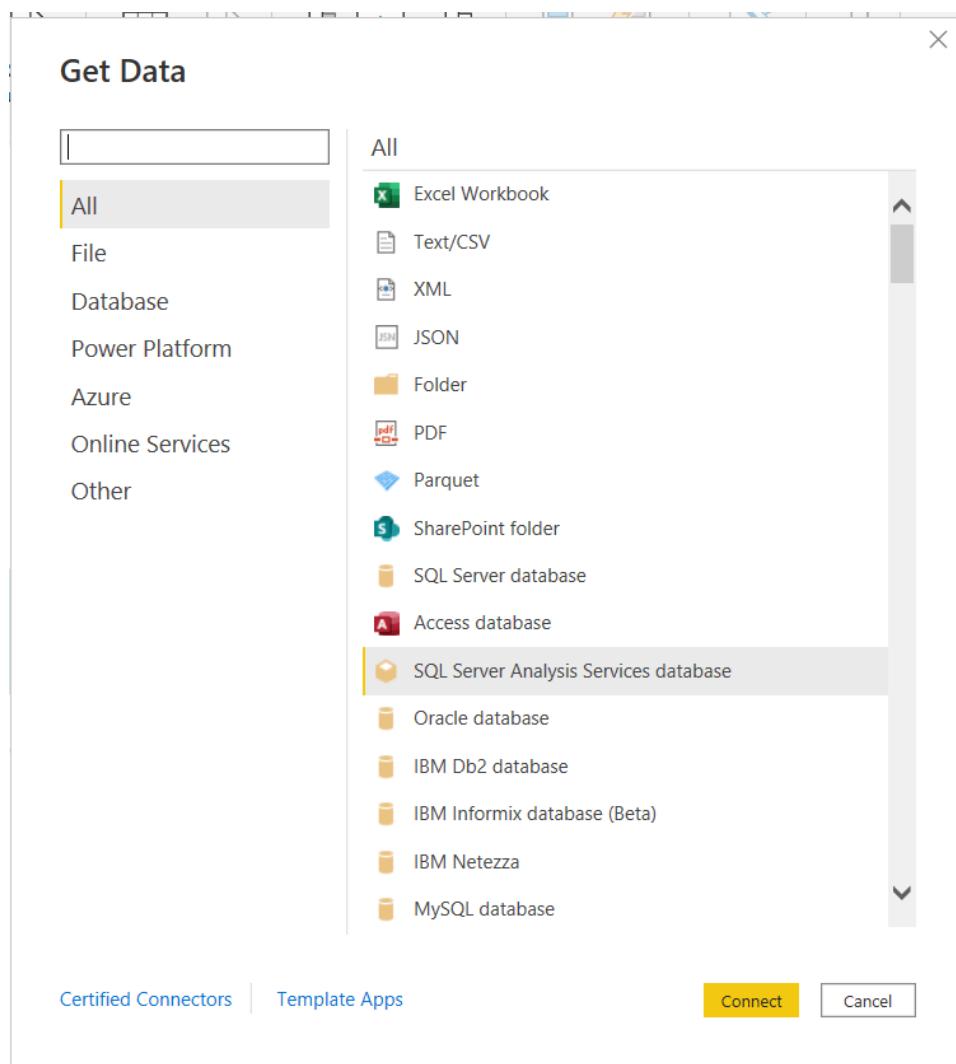
7.3.2. Report2 (Report Grouping) Mỗi quốc gia, cho biết nhóm tuổi có tỷ lệ tử tử cao nhất và thấp nhất.

- Tạo report bằng câu truy vấn MDX

Chọn Get Data → Chọn nguồn dữ liệu để import



Chọn nguồn dữ liệu là SQL Server Analysis Services database

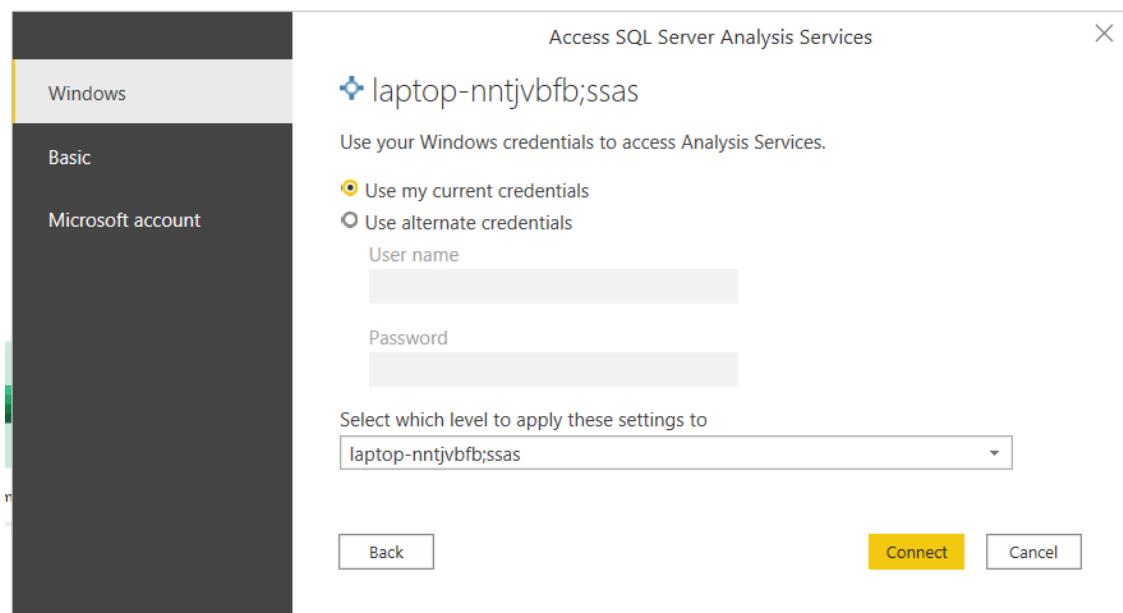
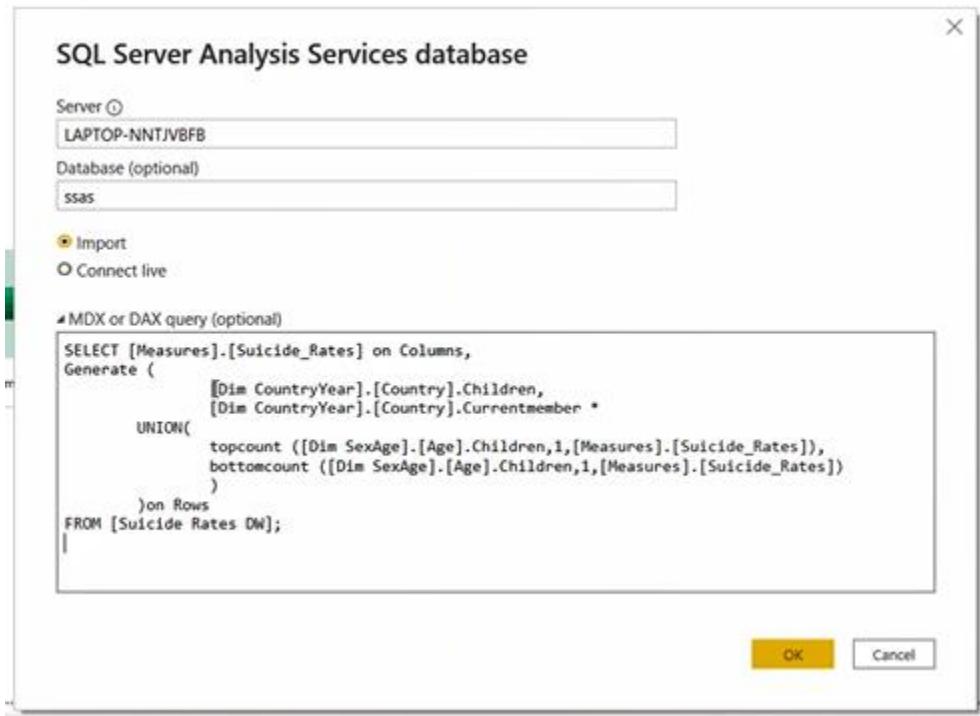


Nhập tên Server

Tại mục Database nhập tên database “ssas” đã hoàn thành tại phần 5. Quá trình SSAS. Chọn Import

Nhập câu truy vấn MDX vào khung MDX query sau đó nhấn OK

```
SELECT [Measures].[Suicide_Rates] on Columns,
Generate (
    [Dim CountryYear].[Country].Children,
    [Dim CountryYear].[Country].Currentmember *
UNION(
    topcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates]),
    bottomcount ([Dim SexAge].[Age].Children,1,[Measures].[Suicide_Rates])
)
)on Rows
FROM [Suicide Rates DW];
```



Kết nối thành công bảng dữ liệu sẽ xuất hiện, sau đó chọn Transform Data để chỉnh sửa dữ liệu.

The screenshot shows the Power BI Data Load Preview window titled "LAPTOP-NNTJVBFB: ssas". It displays a table with three columns: [Dim CountryYear].[Country].[Country].[MEMBER_CAPT], [Dim SexAge].[Age].[Age].[MEMBER_CAPTION], and [Measures].[Suicide_Rates]. The data includes rows for various countries and age groups, with some values being truncated due to size limits. At the bottom, there are buttons for Load, Transform Data, and Cancel.

[Dim CountryYear].[Country].[Country].[MEMBER_CAPT]	[Dim SexAge].[Age].[Age].[MEMBER_CAPTION]	[Measures].[Suicide_Rates]
Albania	75+ years	4.746484964
Albania	5-14 years	0.433199624
Antigua and Barbuda	55-74 years	1.886578878
Antigua and Barbuda	75+ years	0
Argentina	75+ years	17.56056176
Argentina	5-14 years	0.692214508
Armenia	75+ years	6.940139321
Armenia	5-14 years	0.176467279
Aruba	75+ years	14.58470065
Aruba	5-14 years	0
Australia	25-34 years	17.05145094
Australia	5-14 years	0.401032925
Austria	75+ years	46.13795067
Austria	5-14 years	0.510993525
Azerbaijan	75+ years	2.822454191
Azerbaijan	5-14 years	0.231448443
Bahamas	35-54 years	2.392345798
Bahamas	5-14 years	0.154885075
Bahrain	25-34 years	4.807942633
Bahrain	75+ years	0

The data in the preview has been truncated due to size limits.

The screenshot shows the Power Query Editor window titled "Untitled - Power Query Editor". It displays a table with three columns: [Dim CountryYear].[Country].[Country].[MEMBER_CAPT], [Dim SexAge].[Age].[Age].[MEMBER_CAPTION], and [Measures].[Suicide_Rates]. The data is identical to the one shown in the previous screenshot. On the right side, there are sections for Query Settings, Properties (Name: Query1), and Applied Steps (Query1). The status bar at the bottom indicates 3 COLUMNS, 202 ROWS, and PREVIEW DOWNLOADED AT 1:10 AM.

[Dim CountryYear].[Country].[Country].[MEMBER_CAPT]	[Dim SexAge].[Age].[Age].[MEMBER_CAPTION]	[Measures].[Suicide_Rates]
1 Albania	75+ years	4.746484964
2 Albania	5-14 years	0.433199624
3 Antigua and Barbuda	55-74 years	1.886578878
4 Antigua and Barbuda	75+ years	0
5 Argentina	75+ years	17.56056176
6 Argentina	5-14 years	0.692214508
7 Armenia	75+ years	6.940139321
8 Armenia	5-14 years	0.176467279
9 Aruba	75+ years	14.58470065
10 Aruba	5-14 years	0
11 Australia	25-34 years	17.05145094
12 Australia	5-14 years	0.401032925
13 Austria	75+ years	46.13795067
14 Austria	5-14 years	0.510993525
15 Azerbaijan	75+ years	2.822454191
16 Azerbaijan	5-14 years	0.231448443
17 Bahamas	35-54 years	2.392345798
18 Bahamas	5-14 years	0.154885075
19 Bahrain	25-34 years	4.807942633
20 Bahrain	75+ years	0
21 Barbados	35-54 years	4.121890181

Thực hiện thay đổi tên cột cho phù hợp

The screenshot shows the Power Query Editor interface. A query named "Query1" is open, displaying a table with three columns: "Country", "Age", and "Suicide_Rates". The "Country" column contains country names like Albania, Argentina, Australia, Azerbaijan, Austria, Bahrain, Bahamas, Armenia, Aruba, and Barbados. The "Age" column contains age ranges such as "75+ years", "5-14 years", and "35-54 years". The "Suicide_Rates" column contains numerical values ranging from 0 to 4.746484964. The "APPLIED STEPS" pane on the right shows the step "Renamed Columns".

Chọn Close & Apply để đỗ dữ liệu vào Power BI

The screenshot shows the Power Query ribbon. The "Close & Apply" button is highlighted with a red box.

Kết nối thành công, các Measure và Dimension sẽ xuất ở Fields

The screenshot shows the Power BI Desktop interface. The "Fields" pane on the right lists three fields from the "Query1" table: "Age", "Country", and "Suicide_Rates". These fields are highlighted with a red box. The main canvas area is empty, showing the placeholder "Build visuals with your data".

Thực hiện kéo thả và Visualizations thích hợp với câu truy vấn

Mỗi quốc gia, thống kê nhóm tuổi có tỷ lệ tử vong cao nhất và thấp nhất

Age	Country	Suicide_Rates
15-24 years	Colombia	7.534628639
15-24 years	Dominica	0
15-24 years	El Salvador	10.97958554
15-24 years	El Salvador	15.7384131
15-24 years	Guatemala	4.094306422
15-24 years	Kiribati	14.89575493
15-24 years	New Zealand	20.53762489
15-24 years	Nicaragua	9.70517586
15-24 years	Paraguay	6.230716325
15-24 years	Saint Kitts and Nevis	0
25-34 years	Australia	17.02914594
25-34 years	Bahrain	4.827954019
25-34 years	Ireland	15.41077202
25-34 years	Kuwait	2.697565349
25-34 years	Mauritius	17.54461588
25-34 years	Mongolia	24.361017923
25-34 years	South Africa	1.309331167
25-34 years	Thailand	9.7161441
25-34 years	Trinidad and Tobago	18.91466917
25-34 years	Turkmenistan	10.39644908
35-44 years	Albania	2.02351798
35-44 years	Barbados	4.121890181
35-44 years	Belarus	43.19516867
35-44 years	Cabo Verde	20.7997354
35-44 years	Canada	17.22314262
35-44 years	Costa Rica	9.04785587
35-44 years	Cyprus	4.605118515
35-44 years	Finland	32.39270505

Mỗi quốc gia, thống kê nhóm tuổi có tỷ lệ tử vong cao nhất và thấp nhất

Country	Age	Suicide_Rates
Albania	5-14 years	0.433199624
Albania	75+ years	4.746484964
Antigua and Barbuda	55-74 years	1.886578878
Antigua and Barbuda	75+ years	0
Argentina	5-14 years	0.692214508
Argentina	75+ years	17.56051376
Armenia	5-14 years	0.176467279
Armenia	75+ years	6.940139321
Aruba	5-14 years	0
Aruba	75+ years	14.58470065
Australia	25-34 years	17.05145094
Australia	5-14 years	0.401032952
Austria	5-14 years	0.510999523
Austria	75+ years	46.137995067
Azerbaijan	5-14 years	0.231446444
Azerbaijan	75+ years	2.822454191
Bahamas	35-54 years	2.392345798
Bahamas	5-14 years	0.154885075
Bahrain	25-34 years	4.807942633
Bahrain	75+ years	0
Barbados	35-54 years	4.121890181
Barbados	5-14 years	0.193994196
Belarus	35-54 years	43.19516867
Belarus	5-14 years	1.20156033
Belgium	5-14 years	0.56982815

Report thu được

Mỗi quốc gia, thống kê nhóm tuổi có tỷ lệ tử tự cao nhất và thấp nhất

Country	Age	Suicide_Rates
Albania	5-14 years	0.433199624
Albania	75+ years	4.746484964
Antigua and Barbuda	55-74 years	1.886578878
Antigua and Barbuda	75+ years	0
Argentina	5-14 years	0.692214508
Argentina	75+ years	17.56056176
Armenia	5-14 years	0.176467279
Armenia	75+ years	6.940139321
Aruba	5-14 years	0
Aruba	75+ years	14.58470065
Australia	25-34 years	17.05145094
Australia	5-14 years	0.401032925
Austria	5-14 years	0.510993525
Austria	75+ years	46.13795067
Azerbaijan	5-14 years	0.231448443
Azerbaijan	75+ years	2.822454191
Bahamas	35-54 years	2.392345798
Bahamas	5-14 years	0.154885075
Bahrain	25-34 years	4.807942633
Bahrain	75+ years	0
Barbados	35-54 years	4.121890181
Barbados	5-14 years	0.193994136
Belarus	35-54 years	43.19516867
Belarus	5-14 years	1.20156033
Belgium	5-14 years	0.56983815

Ngày 19, tháng 05, năm 2022

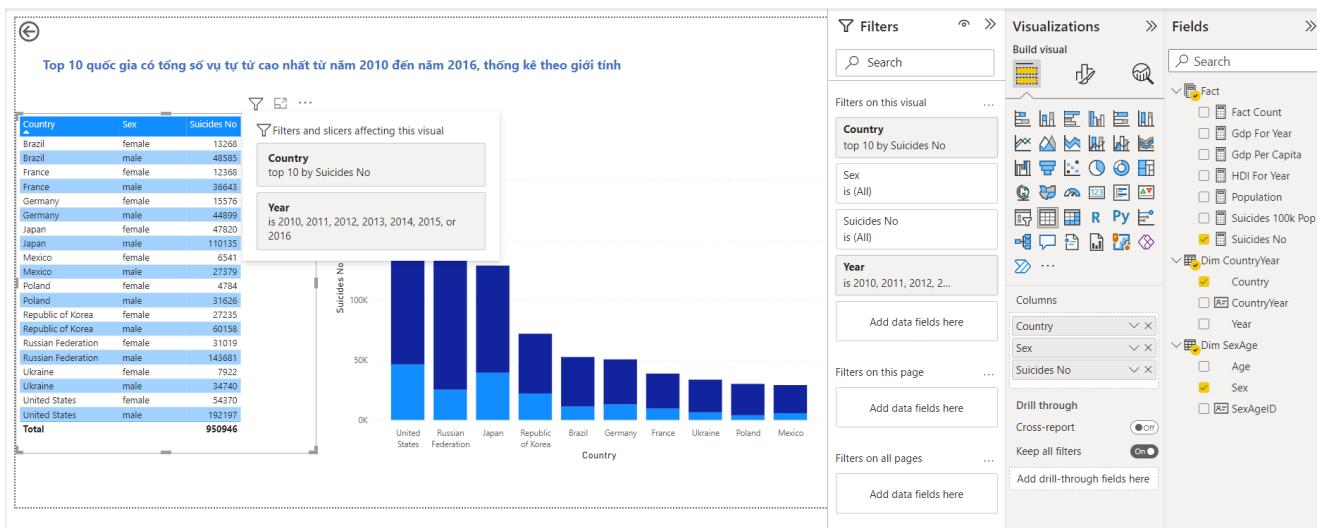
Người tạo

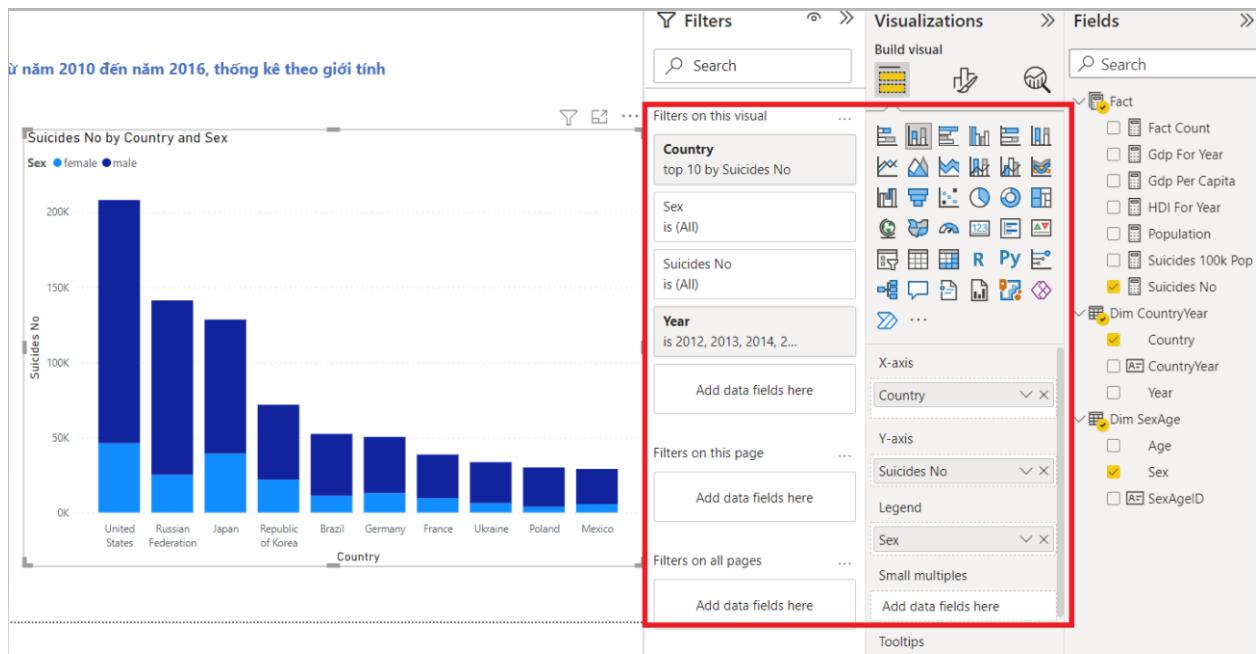


7.3.3. Report3 (Report Totals + Grouping) Top 10 quốc gia có tổng số vụ tự tử cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính.

Thực hiện tương tự phần 7.3.1

Kết quả thu được report như hình minh họa:





Kết quả thu được:

Top 10 quốc gia có tổng số vụ tự tử cao nhất từ năm 2010 đến năm 2016, thống kê theo giới tính

Country	Sex	Suicides No
United States	female	48585
United States	male	13268
Russian Federation	female	36643
Russian Federation	male	143681
Japan	female	27379
Germany	female	15576
Germany	male	44899
Brazil	female	12368
Brazil	male	31019
France	female	60158
France	male	34740
Poland	female	34740
Poland	male	31626
Republic of Korea	female	27235
Republic of Korea	male	13268
Ukraine	female	7922
Ukraine	male	34740
United States	female	54370
United States	male	192197
Total		950946

Suicides No by Country and Sex

Sex ● female ● male

Ngày 19, tháng 05, năm 2022
Người tạo

8 Quá trình Data mining

8.1. Chọn lại kho dữ liệu Marketing Analytics

8.1.1. Mô tả dữ liệu

Tên bộ dữ liệu: Marketing Analytics.

Kho dữ liệu thu thập thông tin dữ liệu về: Hồ sơ khách hàng, Sở thích tiêu dùng sản phẩm, Chiến dịch thành công / thất bại, Hiệu suất kênh.

Kho dữ liệu gốc gồm: 2205 dòng, 39 thuộc tính.

Link dataset: <https://www.kaggle.com/datasets/jackdaoud/marketing-data>

8.1.2. Thuộc tính dữ liệu gốc

STT	Tên thuộc tính	Kiểu dữ liệu	Ý nghĩa
1	AcceptedCmp1	int	Khách hàng có chấp nhận đề nghị trong chiến dịch thứ nhất
2	AcceptedCmp2	int	Khách hàng có chấp nhận đề nghị trong chiến dịch thứ 2
3	AcceptedCmp3	int	Khách hàng có chấp nhận đề nghị trong chiến dịch thứ 3
4	AcceptedCmp4	int	Khách hàng có chấp nhận đề nghị trong chiến dịch thứ 4
5	AcceptedCmp5	int	Khách hàng có chấp nhận đề nghị trong chiến dịch thứ 5
6	Response	int	Khách hàng có chấp nhận đề nghị trong chiến dịch cuối cùng

7	Complain	int	Khách hàng có phàn nàn trong 2 năm gần đây
8	DtCustomer	date	Ngày khách hàng đăng ký
9	Education	Nvarchar	Trình độ học vấn khách hàng
10	Marital	Nvarchar	Tình trạng hôn nhân khách hàng
11	Kidhome	int	Số trẻ nhỏ trong gia đình khách hàng
12	Teenhome	int	Số lượng thanh thiếu niên trong gia đình khách hàng
13	Income	int	Thu nhập hàng năm hộ gia đình khách hàng
14	MntFishProducts	int	Khoản tiền dành cho các sản phẩm cá trong vòng 2 năm
15	MntMeatProducts	int	Khoản tiền dành cho các sản phẩm thịt trong vòng 2 năm
16	MntFruits	int	Khoản tiền mua trái cây trong vòng 2 năm
17	MntSweetProducts	int	Khoản tiền dành cho các sản phẩm đồ ngọt trong vòng 2 năm
18	MntWines	int	Khoản tiền mua rượu trong vòng 2 năm
19	MntGoldProds	int	Khoản tiền mua vàng trong vòng 2 năm

20	NumDealsPurchases	int	Số lần mua hàng được giảm giá
21	NumCatalogPurchases	int	Số lần mua hàng sử dụng danh mục
22	NumStorePurchases	int	Số lần mua hàng tại cửa hàng
23	NumWebPurchases	int	Số lần mua hàng qua website
24	NumWebVisitsMonth	int	Số lần truy cập website trong tháng trước
25	Recency	int	Số ngày kể từ lần mua hàng cuối cùng

8.2. Chuẩn bị dữ liệu gốc, Import dữ liệu

- Dữ liệu gốc

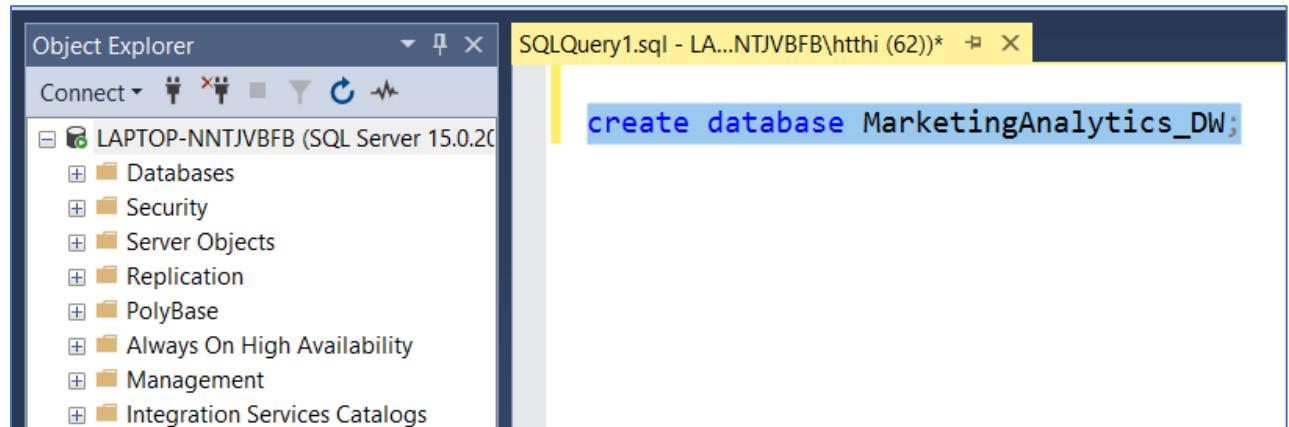
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W		
1	Income	Kidhome	Teenhome	Recency	MntWines	MntFruits	MntMeatP	MntFishPr	MntSweetP	MntGoldP	NumDeals	NumWebV	NumCatal	NumStore	AcceptedC_AcceptedC	AcceptedC_AcceptedC	AcceptedC_AcceptedC	Complain	Z_CostCor	Z_Revenue	Re			
2	58138	0	0	58	635	88	546	172	88	88	3	8	10	4	7	0	0	0	0	0	0	3	11	
3	46344	1	1	38	11	1	6	2	1	6	2	1	1	2	5	0	0	0	0	0	0	0	3	11
4	71613	0	0	26	426	49	127	111	21	42	1	8	2	10	4	0	0	0	0	0	0	0	3	11
5	26646	1	0	26	11	4	20	10	3	5	2	2	0	4	6	0	0	0	0	0	0	0	3	11
6	58293	1	0	94	173	43	118	46	27	15	5	5	3	6	5	0	0	0	0	0	0	0	3	11
7	65213	0	1	16	520	42	98	0	42	14	2	6	4	10	6	0	0	0	0	0	0	0	3	11
8	55635	0	1	34	235	65	164	50	49	27	4	7	3	7	6	0	0	0	0	0	0	0	3	11
9	33454	1	0	32	76	10	56	3	1	23	2	4	0	4	8	0	0	0	0	0	0	0	3	11
10	30351	1	0	19	14	0	24	3	3	2	1	3	0	2	9	0	0	0	0	0	0	0	3	11
11	5648	1	1	68	28	0	6	1	1	13	1	1	0	0	20	1	0	0	0	0	0	0	3	11
12	7500	0	0	59	6	16	11	11	1	16	1	2	0	3	8	0	0	0	0	0	0	0	3	11
13	63033	0	0	82	194	61	480	225	112	30	1	3	4	8	2	0	0	0	0	0	0	0	3	11
14	59354	1	1	53	233	2	53	3	5	14	3	6	1	5	6	0	0	0	0	0	0	0	3	11
15	17323	0	0	38	3	14	17	6	1	5	1	1	0	3	8	0	0	0	0	0	0	0	3	11
16	82800	0	0	23	1006	22	115	59	68	45	1	7	6	12	3	0	0	1	1	0	0	0	3	11
17	41850	1	1	51	53	5	19	2	13	4	3	3	0	3	8	0	0	0	0	0	0	0	3	11
18	37760	0	0	20	84	5	38	150	12	28	2	4	1	6	7	0	0	0	0	0	0	0	3	11
19	76995	0	1	91	1012	80	498	0	16	176	2	11	4	9	5	0	0	0	1	0	0	0	3	11
20	33812	1	0	86	4	17	19	30	24	39	2	2	1	3	6	0	0	0	0	0	0	0	3	11
21	37040	0	0	41	86	2	73	69	38	48	1	4	2	5	8	0	0	0	0	0	0	0	3	11
22	2447	1	0	42	1	1	1725	1	1	1	15	0	28	0	1	0	0	0	0	0	0	0	3	11
23	58607	0	1	63	867	0	86	0	0	19	3	2	3	9	8	0	1	0	0	0	0	0	3	11
24	65324	0	1	0	384	0	102	21	32	5	3	6	2	9	4	0	0	0	0	0	0	0	3	11
25	40689	0	1	69	270	3	27	39	6	99	7	7	1	5	8	0	0	0	0	0	0	0	3	11
26	18589	0	0	89	6	4	25	15	12	13	2	2	1	3	7	0	0	0	0	0	0	0	3	11
27	53359	1	1	4	173	4	30	3	6	41	4	5	1	4	7	0	0	0	0	0	0	0	3	11

Thêm cột Cus_id có giá trị tự tăng

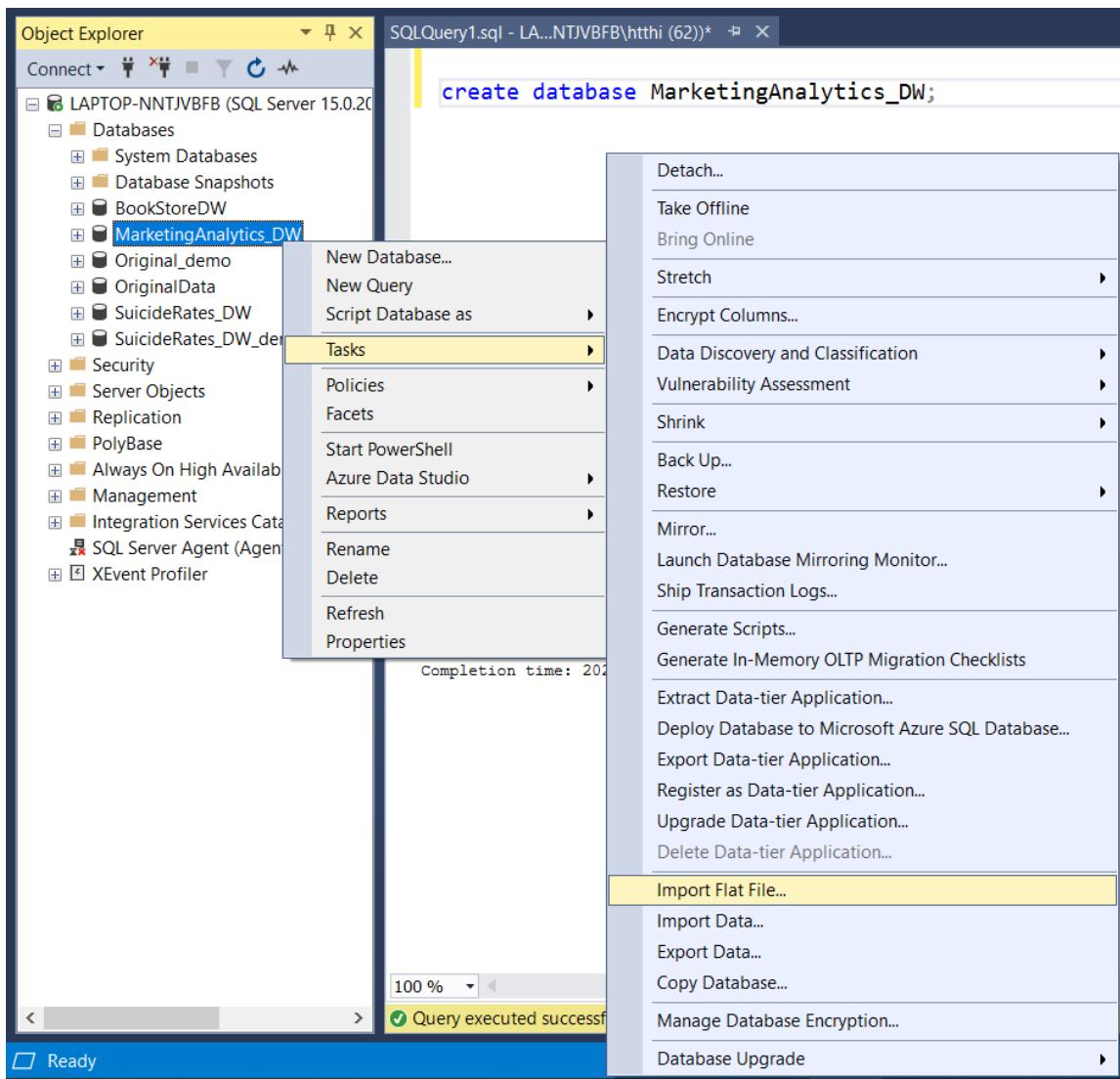
A	B	C	D	E	F	G	H	I	J	K	L	M
Cus_id	Income	Kidhome	Teenhome	Recency	MntWines	MntFruits	MntMeatP	MntFishPr	MntSweet	MntGoldP	NumDeals	NumI
1	1	58138	0	0	58	635	88	546	172	88	88	3
2	2	46344	1	1	38	11	1	6	2	1	6	2
3	3	71613	0	0	26	426	49	127	111	21	42	1
4	4	26646	1	0	26	11	4	20	10	3	5	2
5	5	58293	1	0	94	173	43	118	46	27	15	5
6	6	62513	0	1	16	520	42	98	0	42	14	2
7	7	55635	0	1	34	235	65	164	50	49	27	4
8	8	33454	1	0	32	76	10	56	3	1	23	2
9	9	30351	1	0	19	14	0	24	3	3	2	1
10	10	5648	1	1	68	28	0	6	1	1	13	1
11	11	7500	0	0	59	6	16	11	11	1	16	1
12	12	63033	0	0	82	194	61	480	225	112	30	1
13	13	59354	1	1	53	233	2	53	3	5	14	3
14	14	17323	0	0	38	3	14	17	6	1	5	1
15	15	82800	0	0	23	1006	22	115	59	68	45	1
16	16	41850	1	1	51	53	5	19	2	13	4	3
17	17	37760	0	0	20	84	5	38	150	12	28	2
18	18	76995	0	1	91	1012	80	498	0	16	176	2
19	19	33812	1	0	86	4	17	19	30	24	39	2
20	20	37040	0	0	41	86	2	73	69	38	48	1
21	21	2447	1	0	42	1	1	1725	1	1	1	15
22	22	58607	0	1	63	867	0	86	0	0	19	3
23	23	65324	0	1	0	384	0	102	21	32	5	3
24	24	40689	0	1	69	270	3	27	39	6	99	7
25	25	18589	0	0	89	6	4	25	15	12	13	2
26	26	53359	1	1	4	173	4	30	3	6	41	4
27	27	28260	1	0	26	26	2	42	20	21	10	2

- Import dữ liệu vào SQL Server

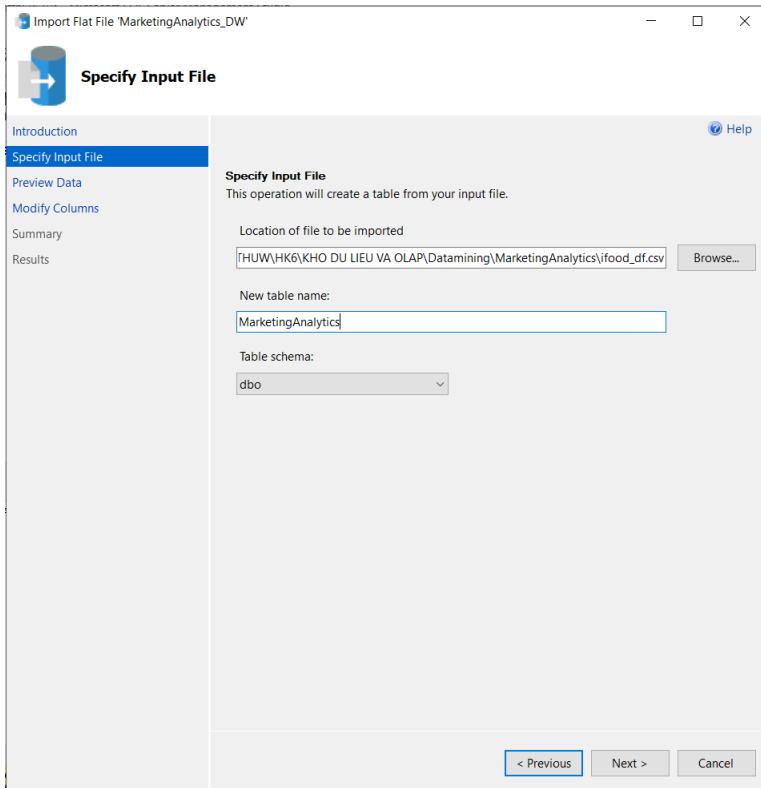
Tạo cơ sở dữ liệu MarketingAnalytics_DW



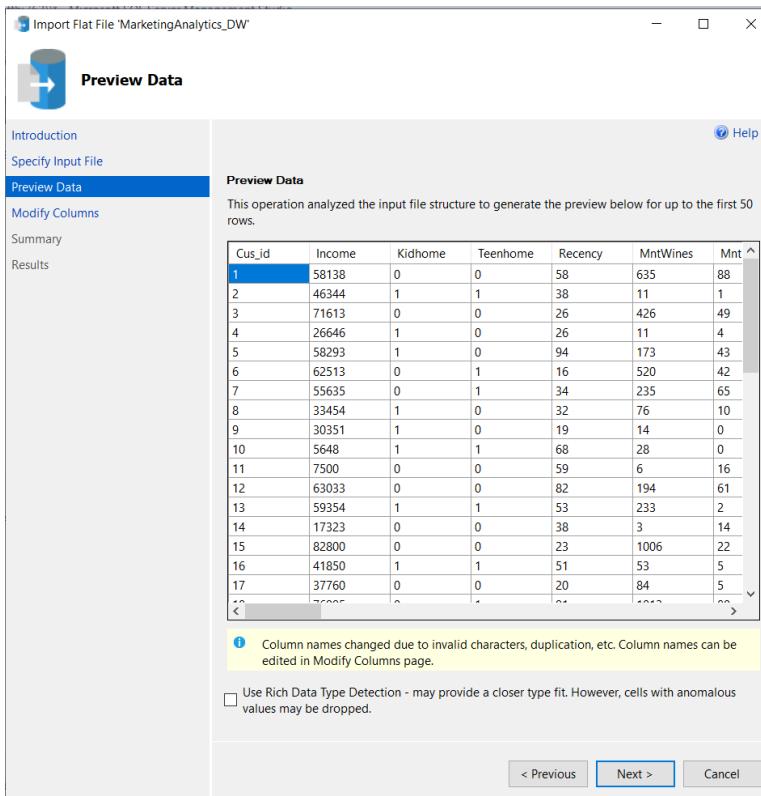
Click phải vào database MarketingAnalytics_DW → chọn Tasks → Import Flat File..



Chọn đường dẫn lưu trữ và đặt tên bảng. Chọn Next.



Chọn Next



Điều chỉnh kiểu dữ liệu cho phù hợp

Import Flat File 'MarketingAnalytics_DW'

Modify Columns

Introduction | Specify Input File | Preview Data | **Modify Columns**

Help

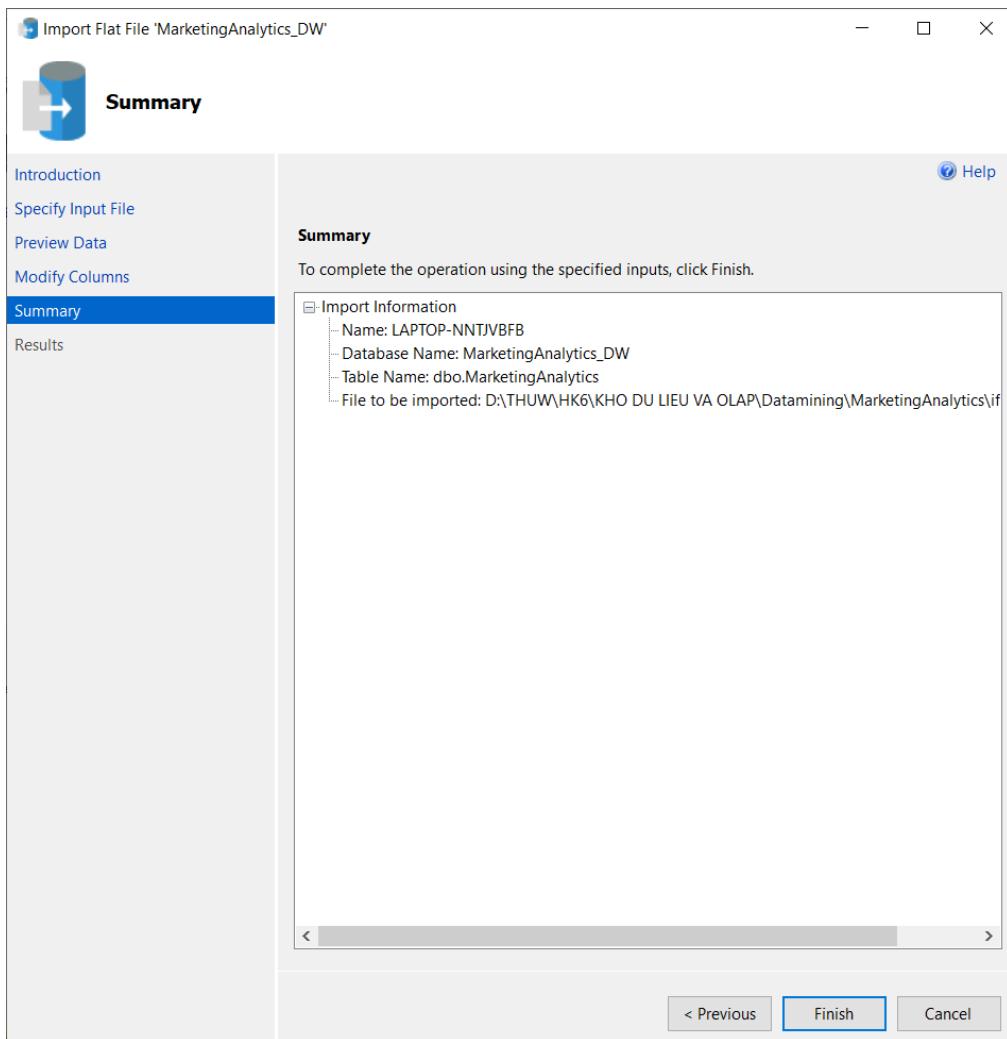
Modify Columns

This operation generated the following table schema. Please verify if schema is accurate, and if not, please make any changes.

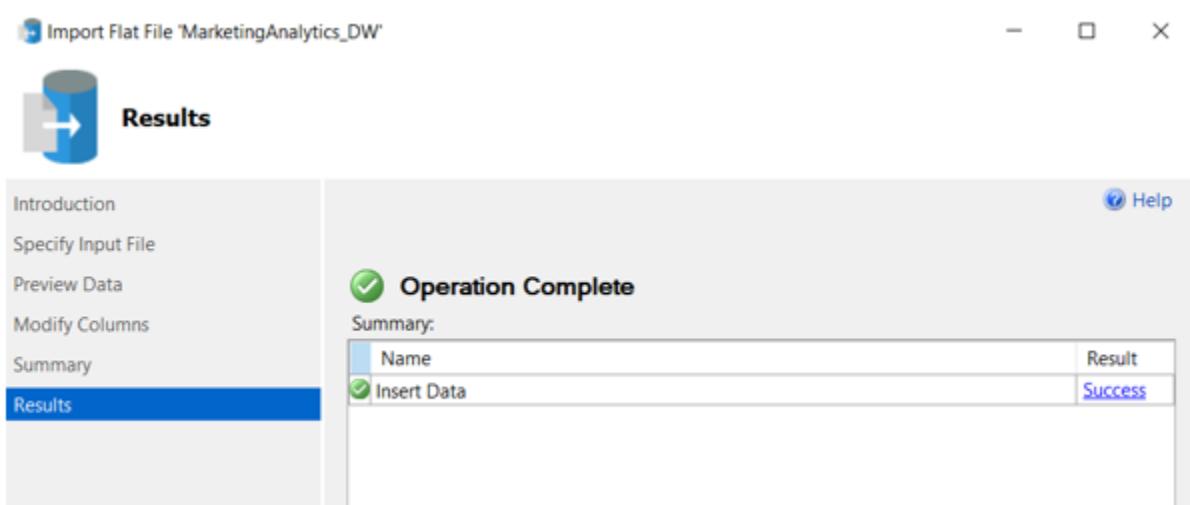
Column Name	Data Type	Primary Key	Allow Nulls
Cus_id	int	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Income	int	<input type="checkbox"/>	<input type="checkbox"/>
Kidhome	int	<input type="checkbox"/>	<input type="checkbox"/>
Teenhome	int	<input type="checkbox"/>	<input type="checkbox"/>
Recency	int	<input type="checkbox"/>	<input type="checkbox"/>
MntWines	int	<input type="checkbox"/>	<input type="checkbox"/>
MntFruits	int	<input type="checkbox"/>	<input type="checkbox"/>
MntMeatProducts	int	<input type="checkbox"/>	<input type="checkbox"/>
MntFishProducts	int	<input type="checkbox"/>	<input type="checkbox"/>
MntSweetProducts	int	<input type="checkbox"/>	<input type="checkbox"/>
MntGoldProds	int	<input type="checkbox"/>	<input type="checkbox"/>
NumDealsPurchases	int	<input type="checkbox"/>	<input type="checkbox"/>
NumWebPurchases	int	<input type="checkbox"/>	<input type="checkbox"/>
NumCatalogPurchases	int	<input type="checkbox"/>	<input type="checkbox"/>
NumStorePurchases	int	<input type="checkbox"/>	<input type="checkbox"/>
NumWebVisitsMonth	int	<input type="checkbox"/>	<input type="checkbox"/>
AcceptedCmp3	int	<input type="checkbox"/>	<input type="checkbox"/>
AcceptedCmp4	int	<input type="checkbox"/>	<input type="checkbox"/>
AcceptedCmp5	int	<input type="checkbox"/>	<input type="checkbox"/>
AcceptedCmp1	int	<input type="checkbox"/>	<input type="checkbox"/>
AcceptedCmp2	int	<input type="checkbox"/>	<input type="checkbox"/>
Complain	int	<input type="checkbox"/>	<input type="checkbox"/>
Z_CostContact	int	<input type="checkbox"/>	<input type="checkbox"/>

Error Reporting - Selecting a smaller range may have a performance impact No Range

< Previous | Next > | Cancel



Import dữ liệu thành công.

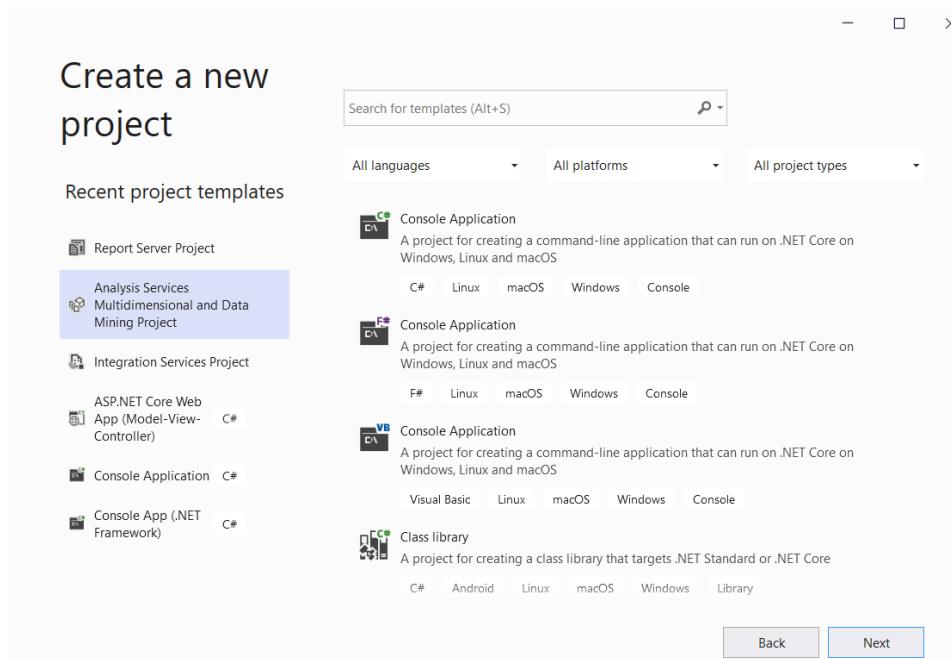


8.3. Quá trình Data mining sử dụng thuật toán Cây quyết định (Microsoft Decision Trees)

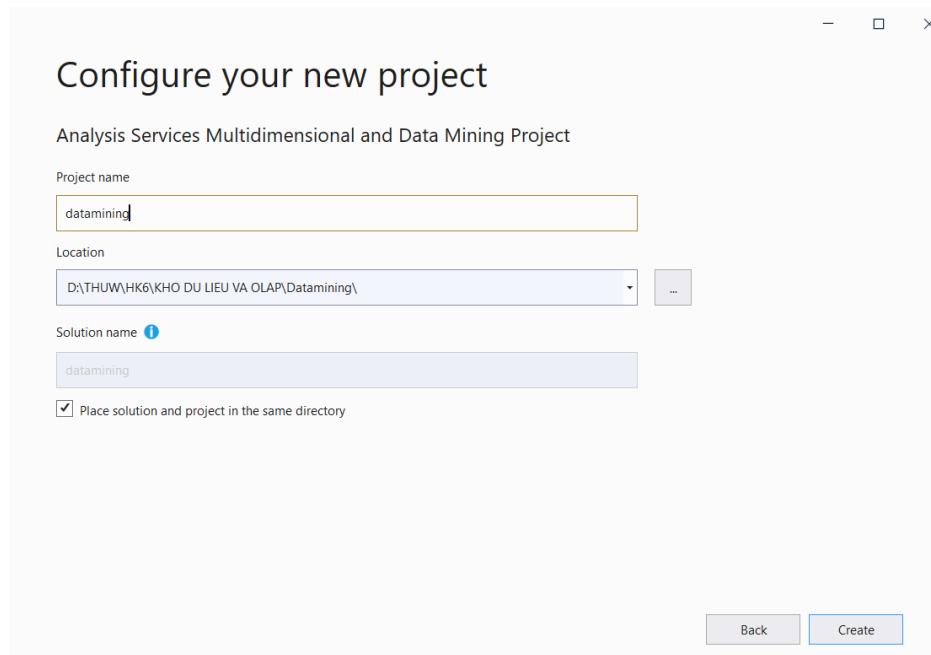
8.3.1. Tạo project SSAS

Mở Visual Studio (phiên bản 2019) và tạo mới project.

Chọn Analysis Services Multidimensional and Data Mining Project.



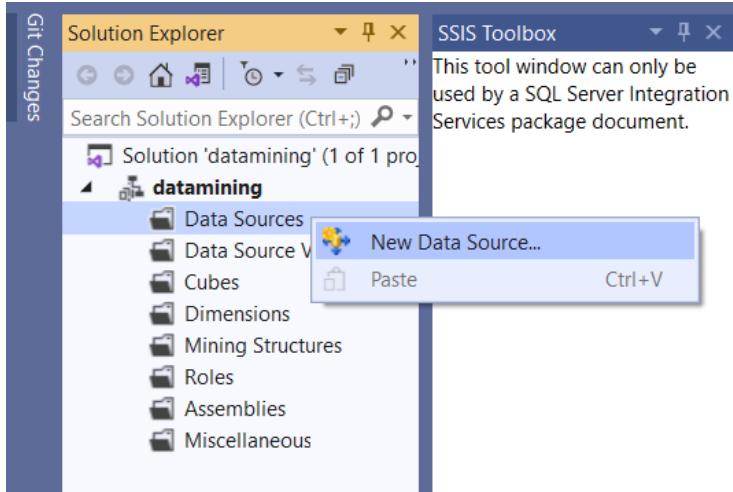
Đặt tên project và chọn đường dẫn lưu trữ.



8.3.2. Tạo Data Source

Tại phần Solution Explorer nhấn chuột phải chọn Data Source.

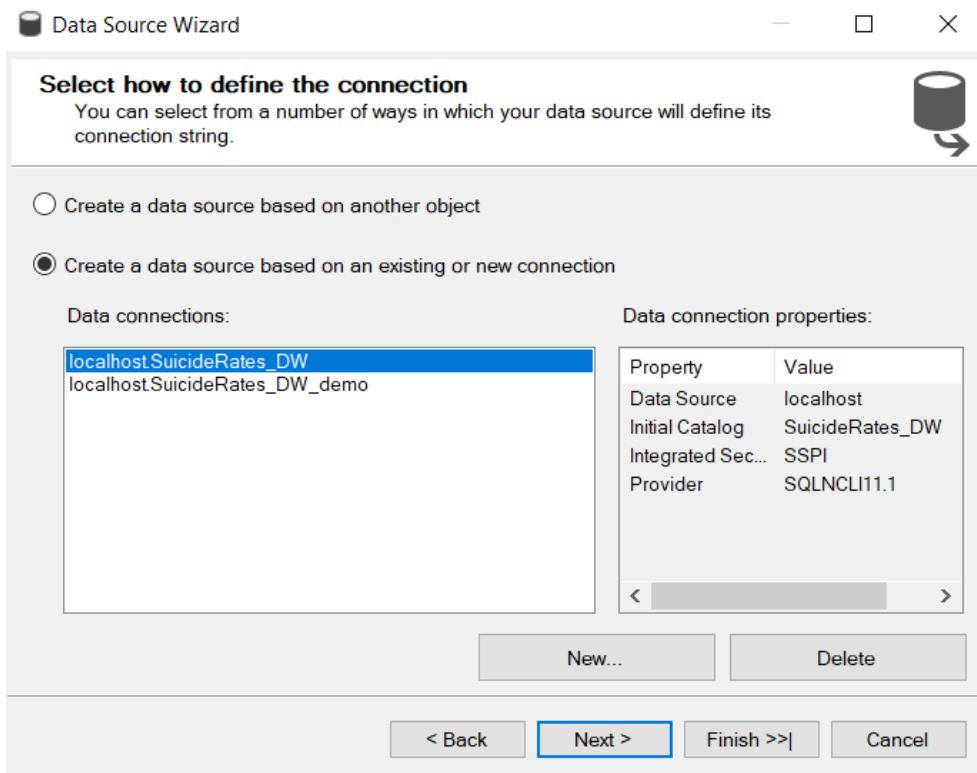
Chọn New Data Source để tiến hành tạo một Data Source.



Cửa sổ Select how to define the connection được hiển thị.

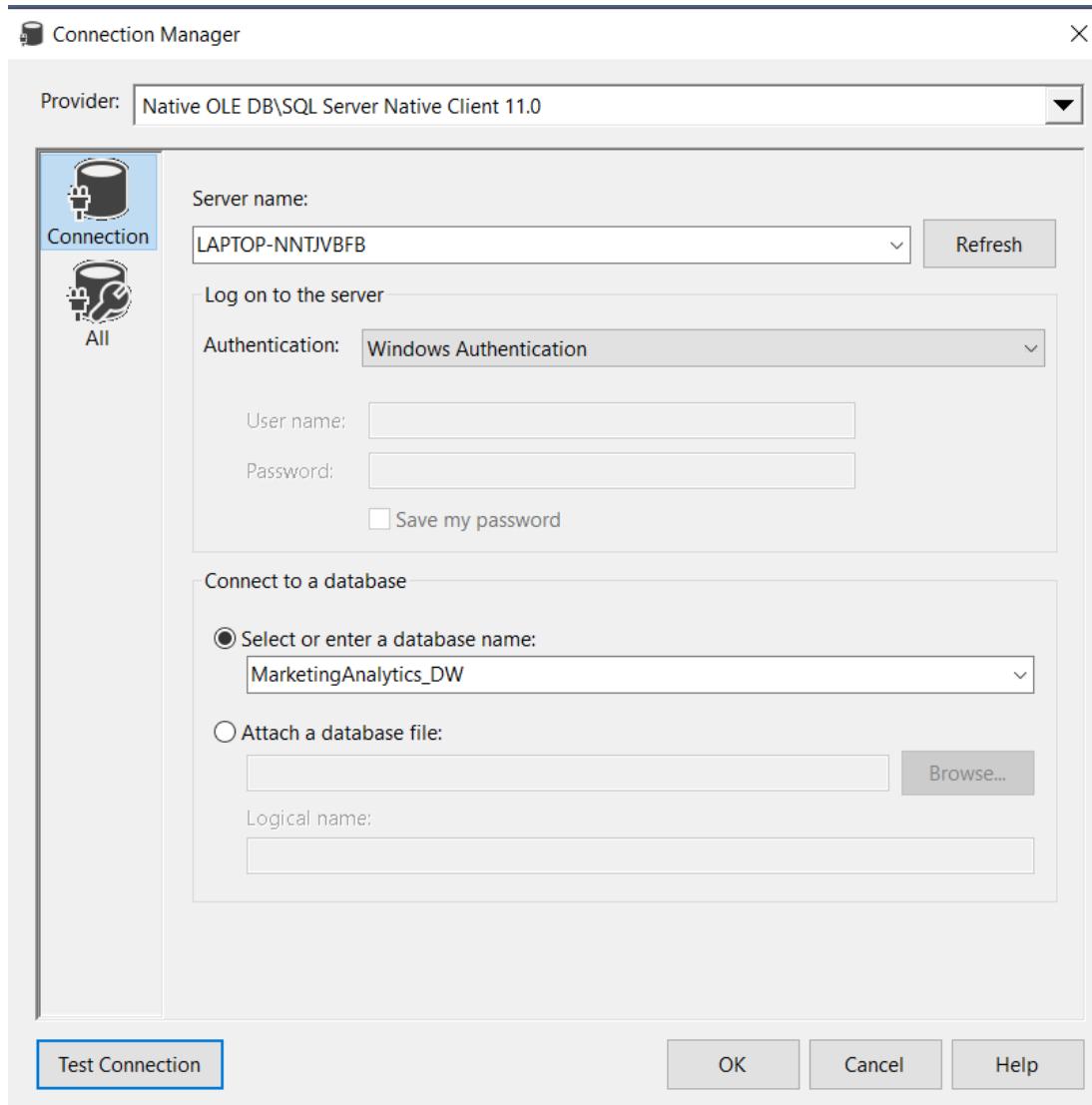
Tại đây chọn Create a data source based on an existing or new connection để tạo 1 data source dựa trên 1 kết nối có sẵn.

Chọn New để tạo kết nối mới

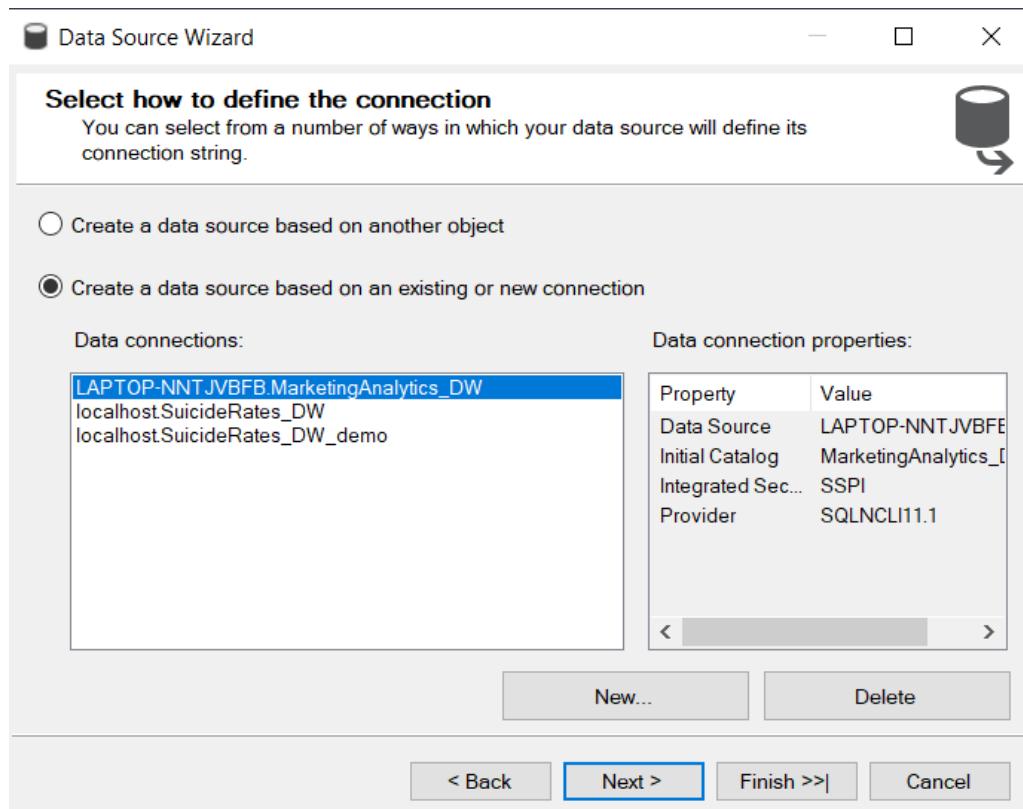


Cửa sổ Connection manager hiển thị.

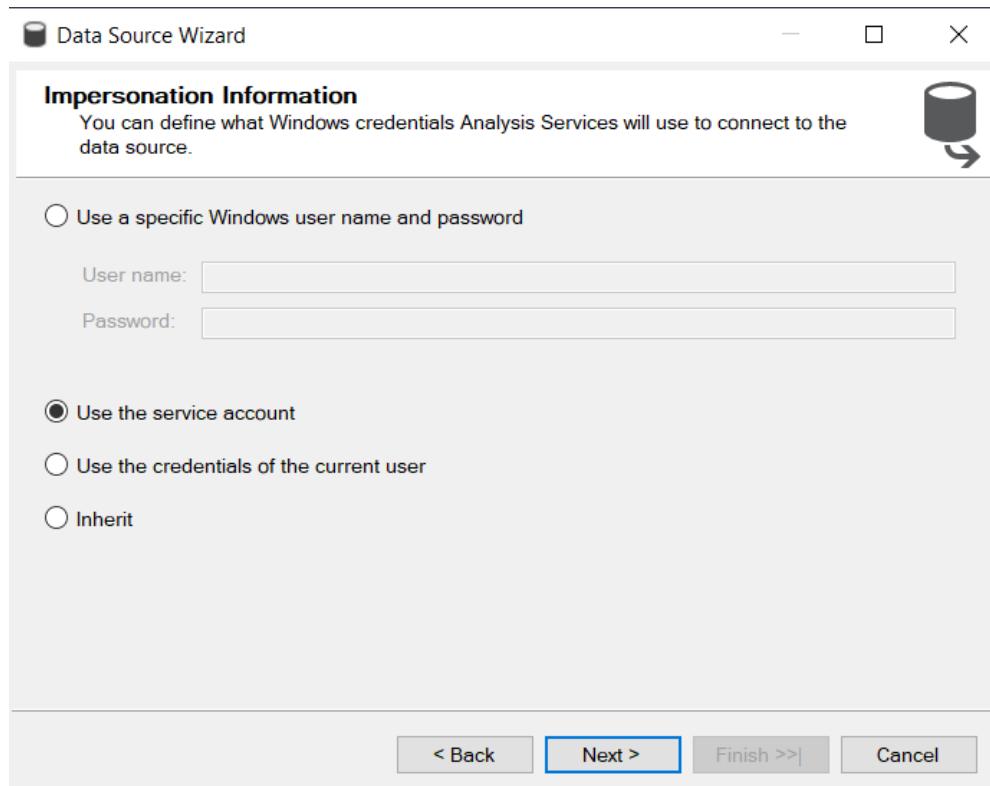
Chọn database cần kết nối, click OK.



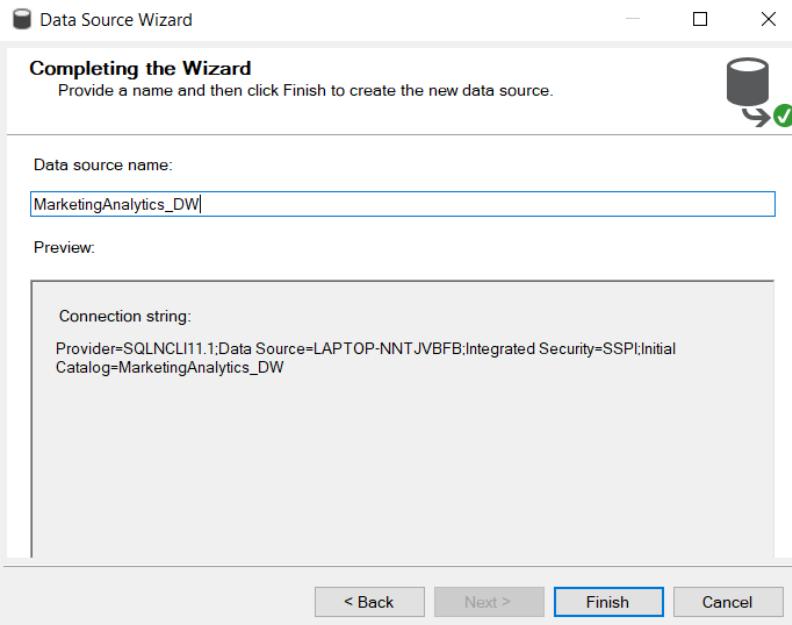
Chọn data connections vừa tạo → Next.



Chọn Use the service account, sau đó chọn Next



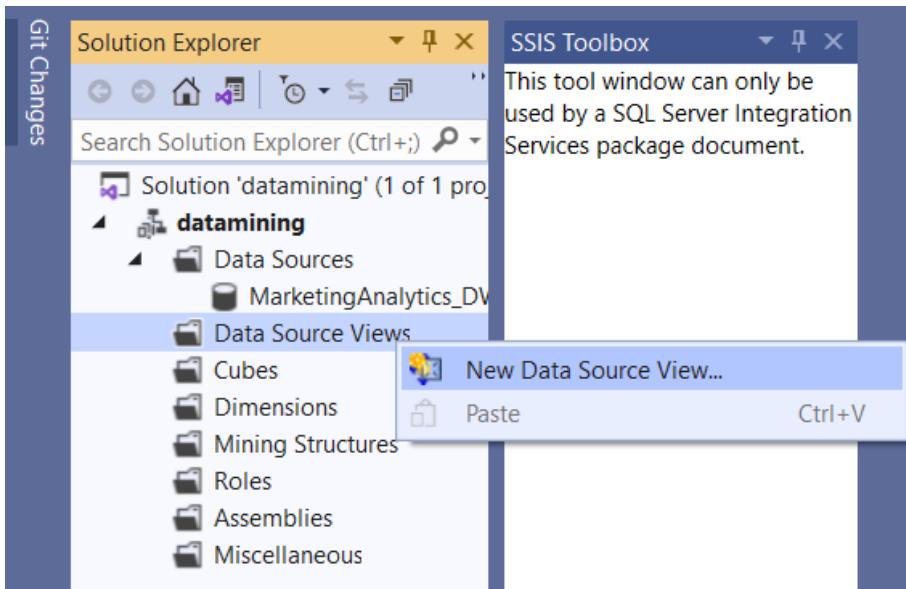
Nhấn Finish để kết thúc



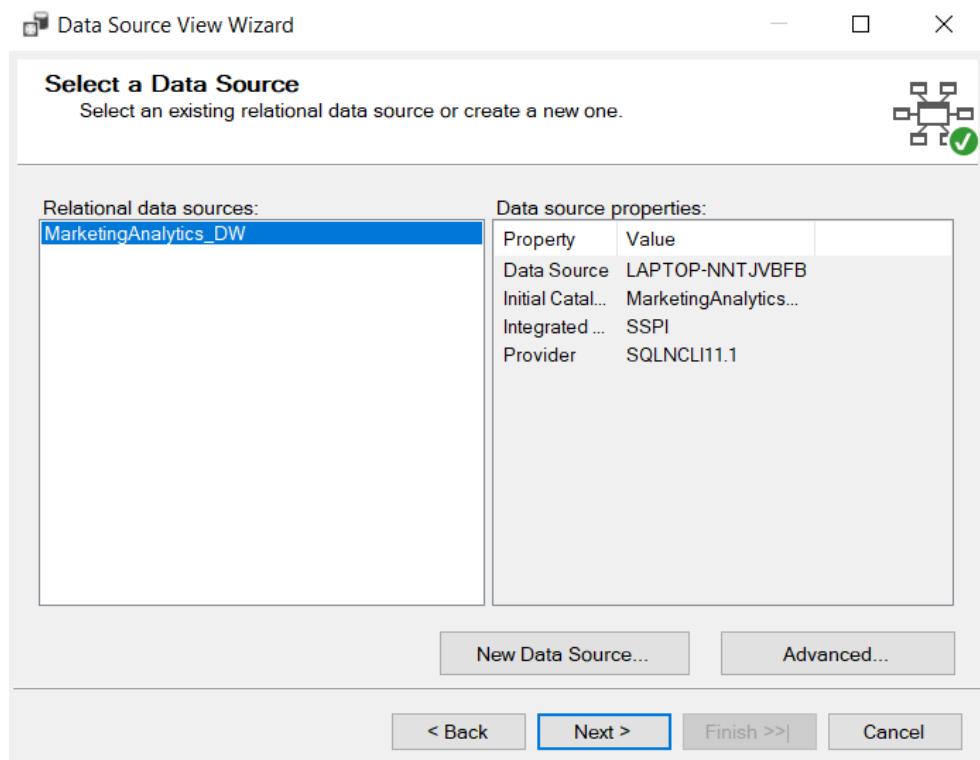
8.3.3. Tạo Data Source Views

Tại phần Solution Explorer nhấn chuột phải chọn Data Source View.

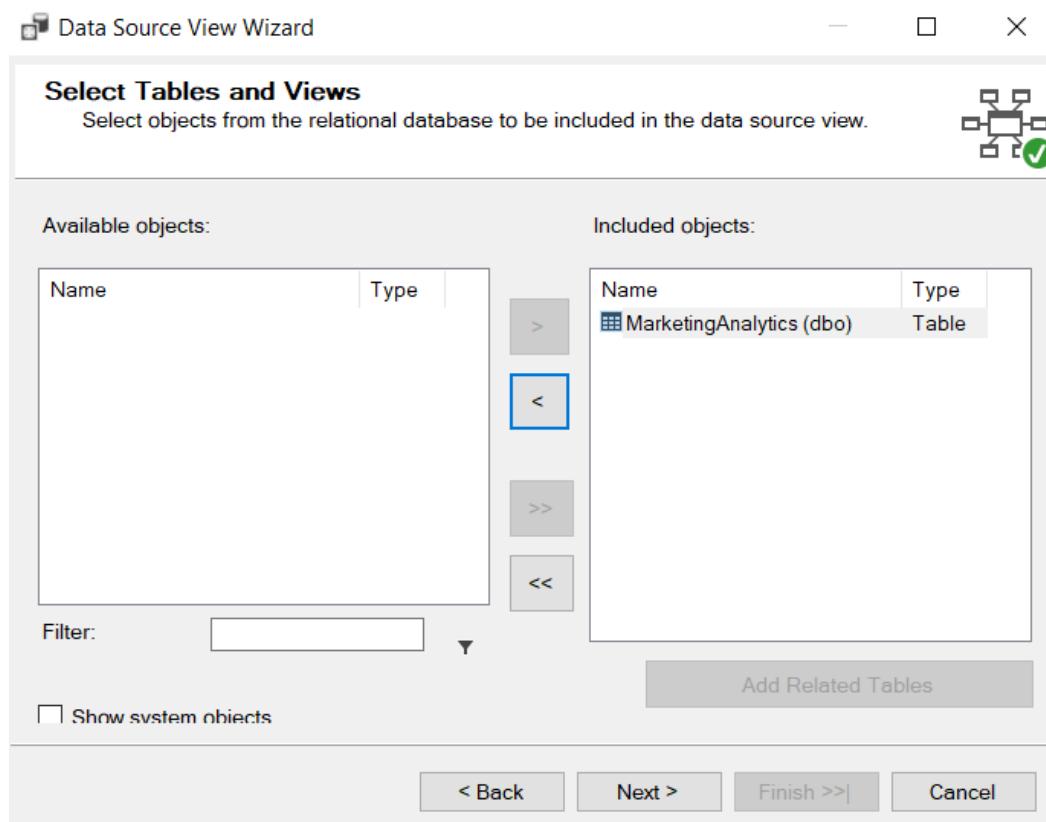
Chọn New Data Source View...



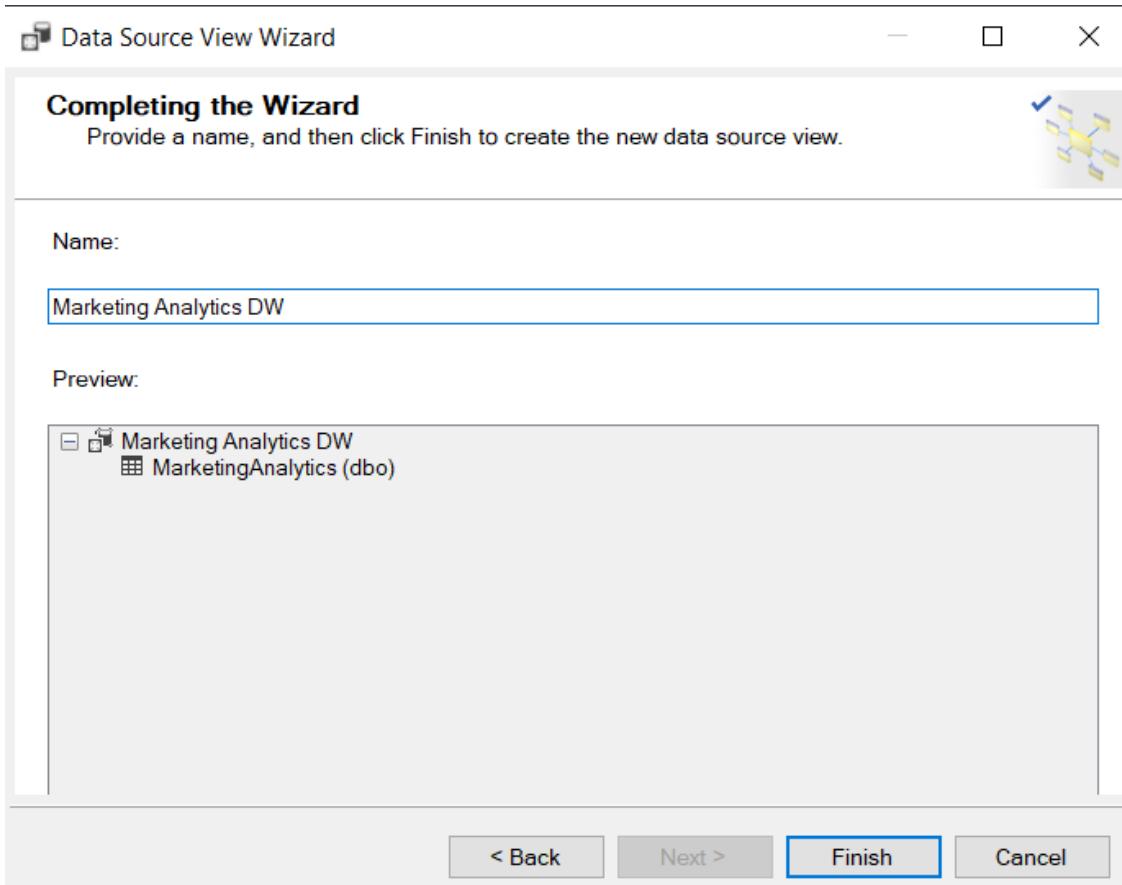
Chọn kho dữ liệu MarketingAnalytics_DW, nhấn Next để tiếp tục



Chọn bảng cần sử dụng cho quá trình Mining → Next



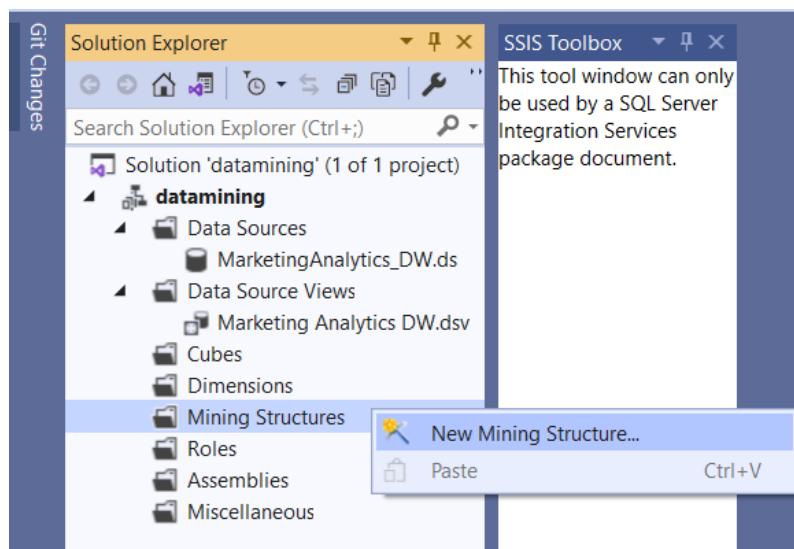
Nhấn Finish để kết thúc quá trình tạo Data Source View



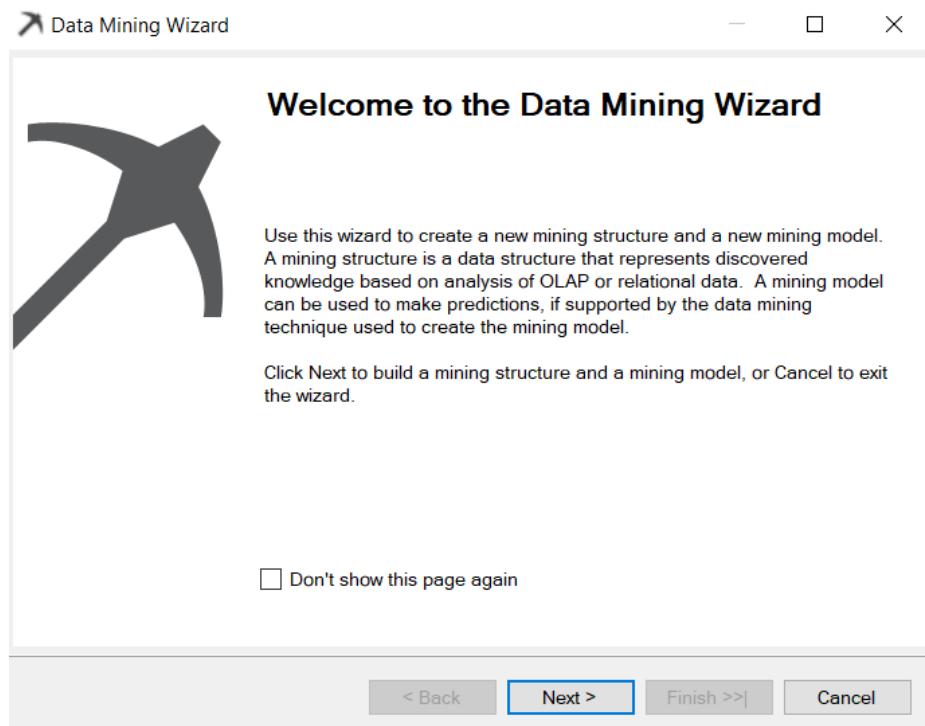
8.3.4. Tạo cấu trúc data mining (Mining Structures)

Tại phần Solution Explorer nhấn chuột phải chọn Mining Structures.

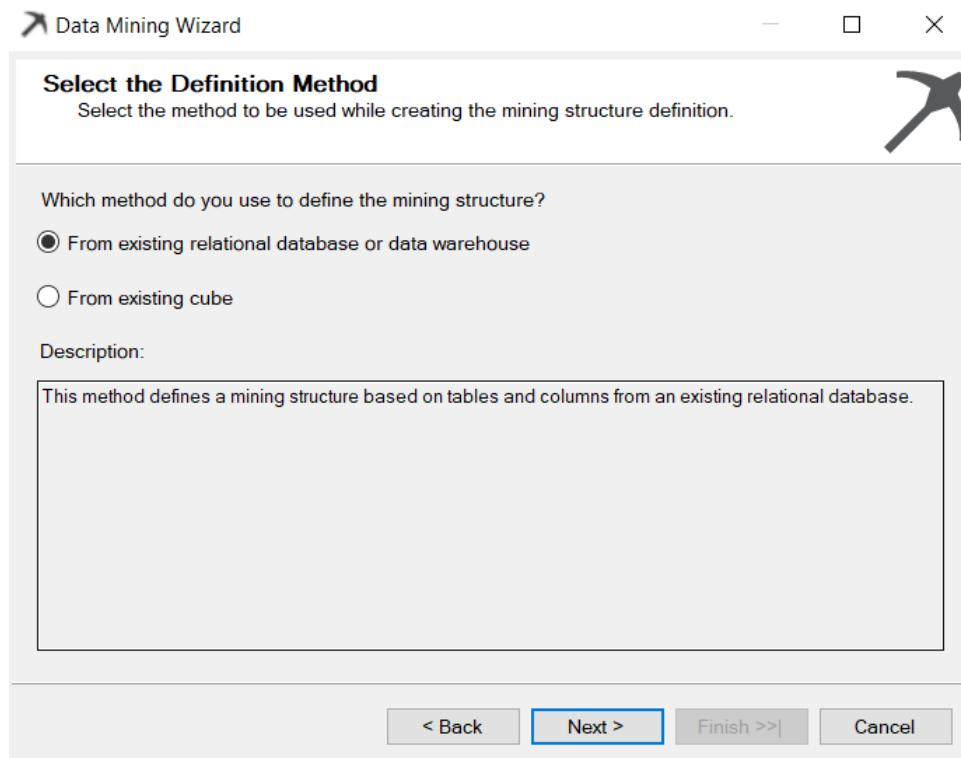
Chọn New Mining Structure...



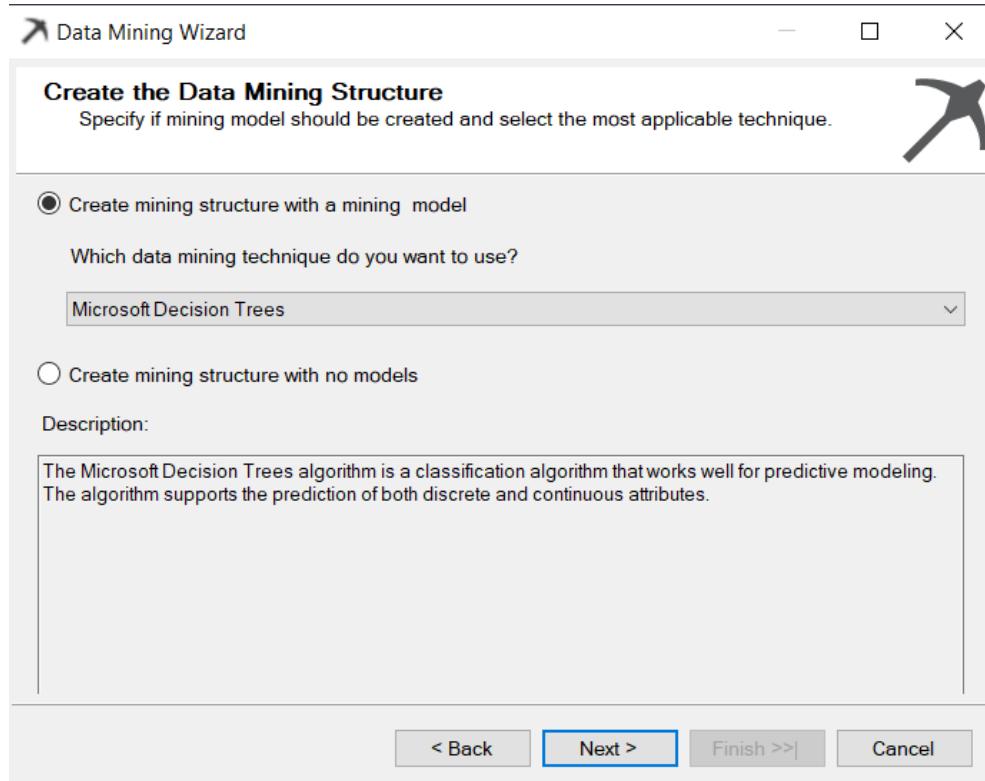
Cửa sổ Data Mining Wizard hiện ra, nhấn Next để tiếp tục.



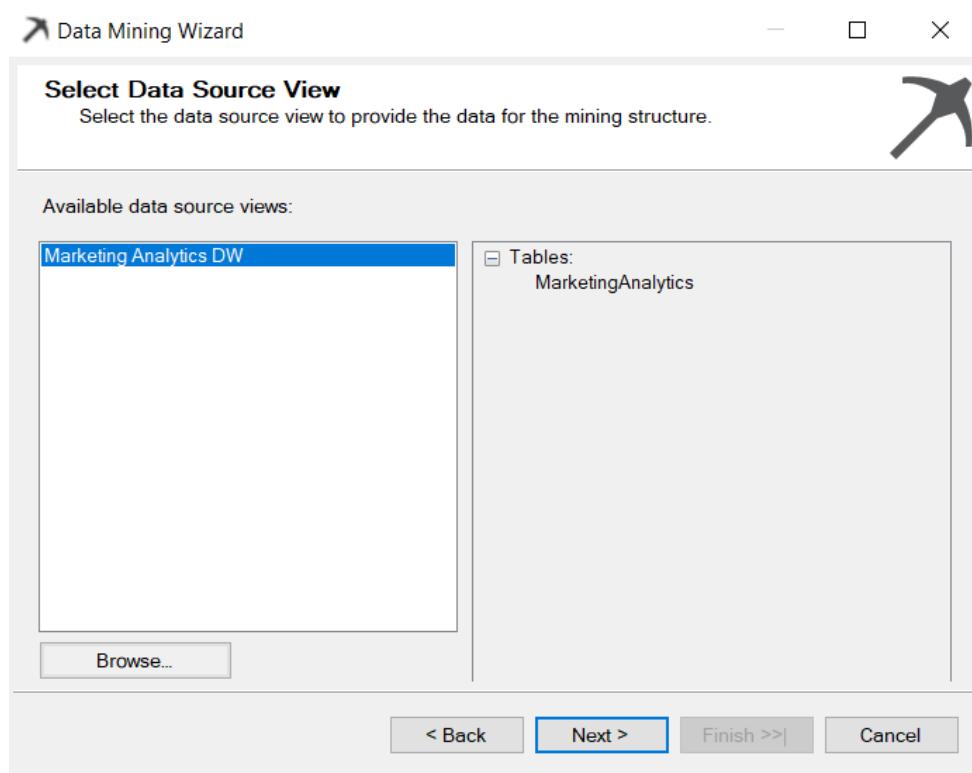
Chọn From Existing relational database for data warehouse → Nhấn Next

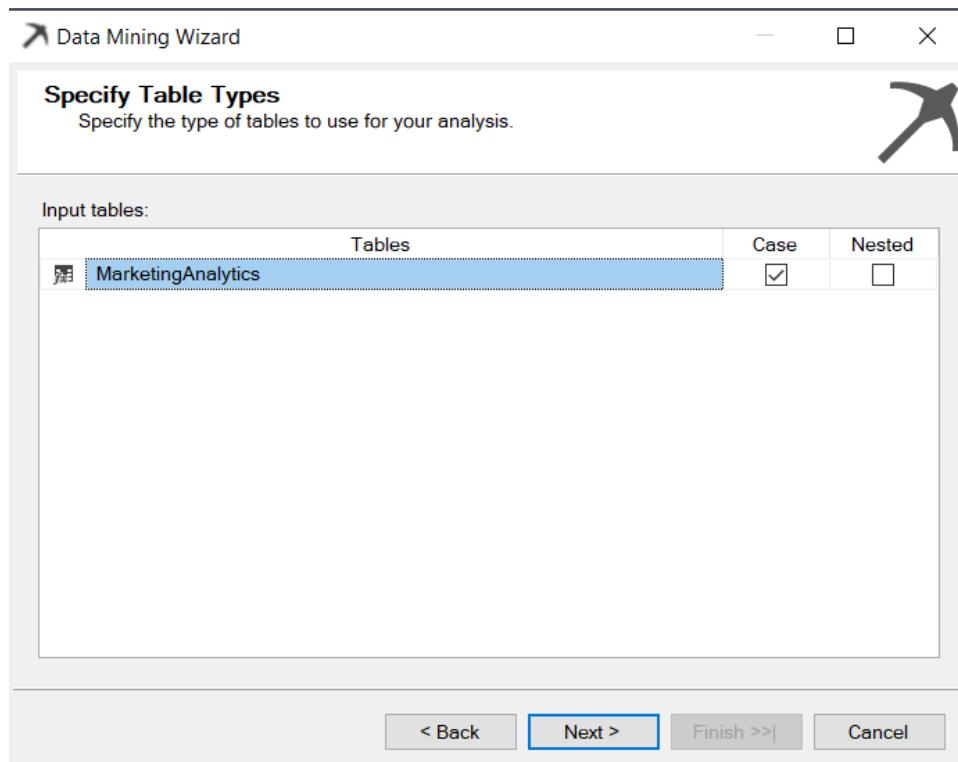


Chọn Create mining structure with a mining model → Chọn Microsoft Decision Trees → Nhấn Next



Chọn data source view → nhấn Next





Chọn cột Cus_id là Key Column, cột Response là Predict Column và các cột Input Column bao gồm: AcceptedCmpOverall, Complain, MntFishProducts, MntFruit, MntFishProds, MntMeatProducts, MntSweetProducts, MntWines, NumDealsPurchases, NumCatalogPurchases, NumStorePurchases, NumWebPurchases, NumWebPurchases, Recency cho bài toán phân loại

→ Chọn Next

Data Mining Wizard

Specify the Training Data

Specify the columns used in your analysis.

Mining model structure:

	Tables/Columns	Key	<input type="checkbox"/> Input	<input type="checkbox"/> Predictive
-	MarketingAnalytics			
	AcceptedCmp1			
	AcceptedCmp2			
	AcceptedCmp3			
	AcceptedCmp4			
	AcceptedCmp5			
<input checked="" type="checkbox"/>	AcceptedCmpOverall		<input checked="" type="checkbox"/>	
	Age			
<input checked="" type="checkbox"/>	Complain		<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	Cus_id	<input checked="" type="checkbox"/>		
	Customer_Days			
	education_2n_Cycle			
	education_Basic			
	education_Graduation			
	education_Master			
	education_PhD			
	Income			
	Kidhome			
	marital_Divorced			
	marital_Married			
	marital_Single			
	marital_Together			
	marital_Widow			
<input checked="" type="checkbox"/>	MntFishProducts		<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	MntFruits		<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/>	MntGoldProds		<input checked="" type="checkbox"/>	

Recommend inputs for currently selected predictable:

	Tables/Columns				
	MntFishProducts		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	MntFruits		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	MntGoldProds		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	MntMeatProducts		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	MntRegularProds		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	MntSweetProducts		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	MntTotal		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	MntWines		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	NumCatalogPurchases		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	NumDealsPurchases		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	NumStorePurchases		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	NumWebPurchases		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	NumWebVisitsMonth		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	Recency		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	Response		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	Teenhome		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	Z_CostContact		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	Z_Revenue		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	

Chọn Detect để tự động chuyển đổi kiểu dữ liệu các column

↗ Data Mining Wizard

Specify Columns' Content and Data Type
Specify mining structure columns' content and data type.

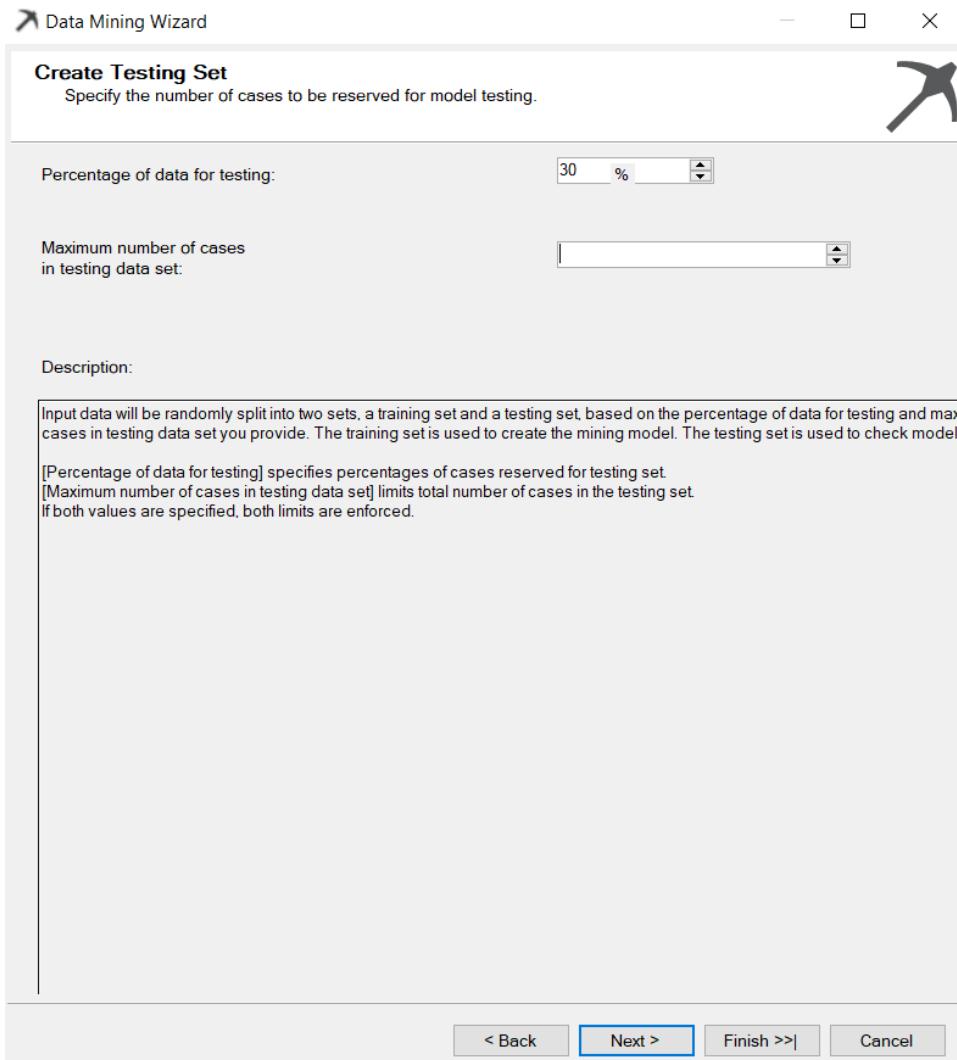
Mining model structure:

Columns	Content Type	Data Type
Accepted Cmp Overall	Continuous	Long
Complain	Discrete	Long
Cus Id	Key	Long
Mnt Fish Products	Continuous	Long
Mnt Fruits	Continuous	Long
Mnt Gold Prods	Continuous	Long
Mnt Meat Products	Continuous	Long
Mnt Sweet Products	Continuous	Long
Mnt Wines	Continuous	Long
Num Catalog Purchases	Continuous	Long
Num Deals Purchases	Continuous	Long
Num Store Purchases	Continuous	Long
Num Web Purchases	Continuous	Long
Num Web Visits Month	Continuous	Long
Recency	Discrete	Long
Response	Discrete	Long

Detect continuous or discrete for numeric columns:

< Back Cancel

Percentage of data for testing đặt mặc định là 30% → Chọn Next



Nhấn Finish để kết thúc quá trình

 Data Mining Wizard

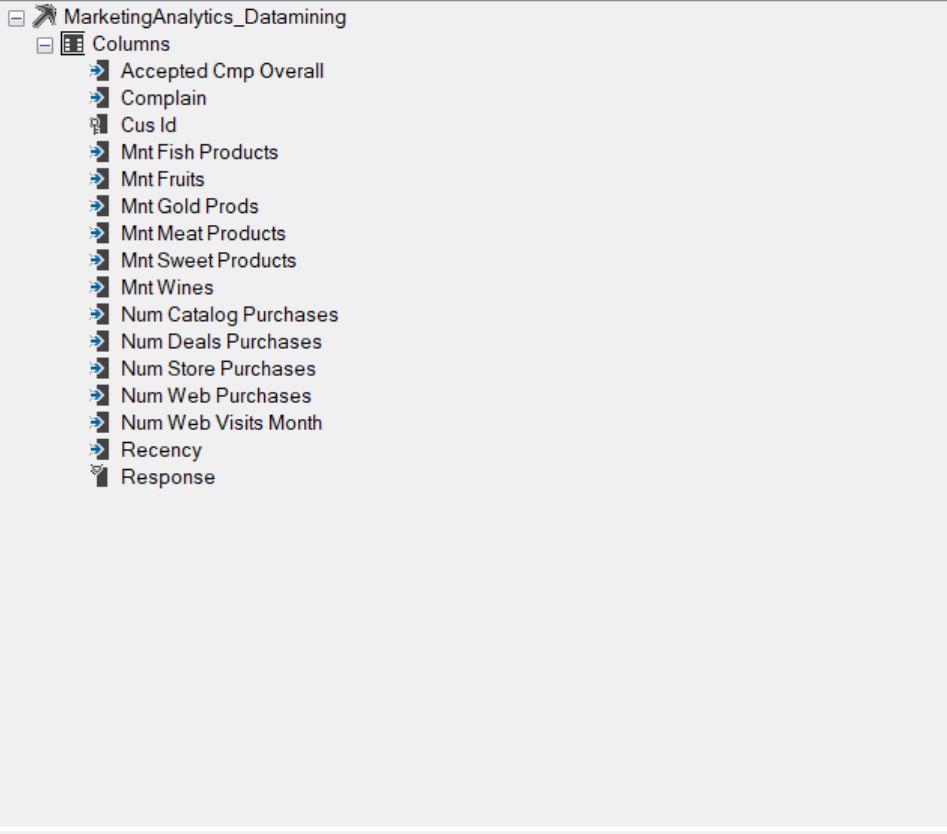
Completing the Wizard
Completing the Data Mining Wizard by providing a name for the mining structure.



Mining structure name:

Mining model name:
 Allow drill through

Preview:



-  MarketingAnalytics_Datamining
 -  Columns
 - ↳ Accepted Cmp Overall
 - ↳ Complain
 - ▢ Cus Id
 - ↳ Mnt Fish Products
 - ↳ Mnt Fruits
 - ↳ Mnt Gold Prods
 - ↳ Mnt Meat Products
 - ↳ Mnt Sweet Products
 - ↳ Mnt Wines
 - ↳ Num Catalog Purchases
 - ↳ Num Deals Purchases
 - ↳ Num Store Purchases
 - ↳ Num Web Purchases
 - ↳ Num Web Visits Month
 - ↳ Recency
 - ▢ Response

Kết quả quá trình tạo Mining Structures sử dụng thuật toán Decision Trees

Solution Explorer

MarketingAnalytics...ining.dmm [Design]*

Mining Structure Mining Models Mining Model Viewer Mining Accuracy Ch... Mining Model Prediction

MarketingAnalytics_Datamining

Columns

- Accepted Cmp Overall
- Complain
- Cus Id
- Mnt Fish Products
- Mnt Fruits
- Mnt Gold Prods
- Mnt Meat Products
- Mnt Sweet Products
- Mnt Wines
- Num Catalog Purchases
- Num Deals Purchases
- Num Store Purchases
- Num Web Purchases
- Num Web Visits Month
- Recency
- Response

Data Source View

MarketingAnalytics

- Cus_Id
- Income
- Kidhome
- Teenhome
- Recency
- MntWines
- MntFruits
- MntMeatProducts
- MntFishProducts
- MntSweetProducts
- MntGoldProd
- NumDealsPurchases
- NumWebVisitsMonth

Solution Explorer

MarketingAnalytics...ining.dmm [Design]*

Mining Structure Mining Models Mining Model Viewer Mining Accuracy Ch... Mining Model Prediction

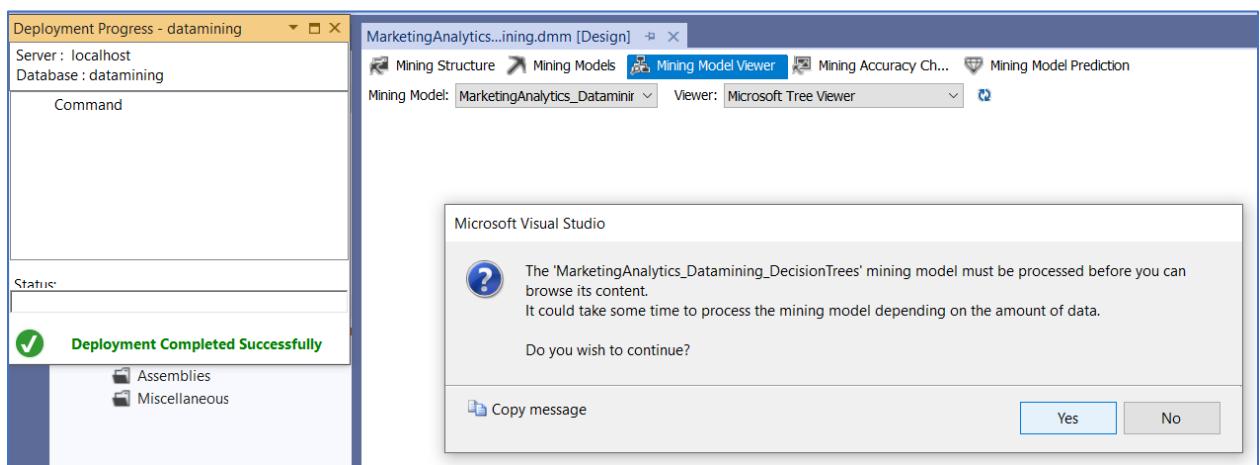
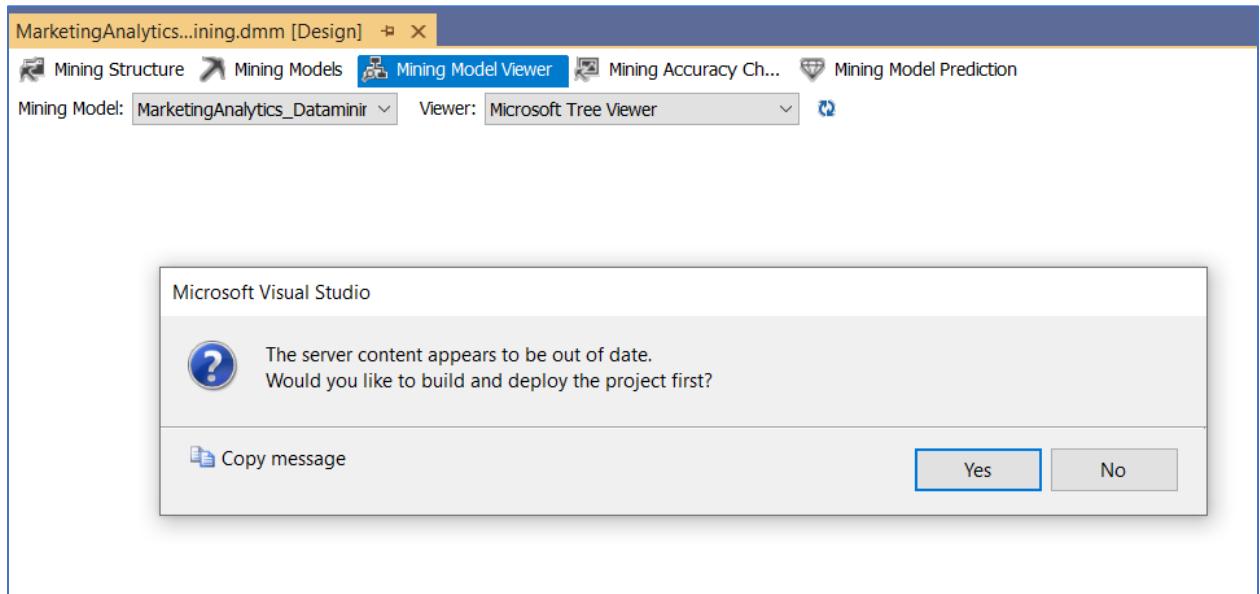
MarketingAnalytics_Datamining_Dec...

Structure

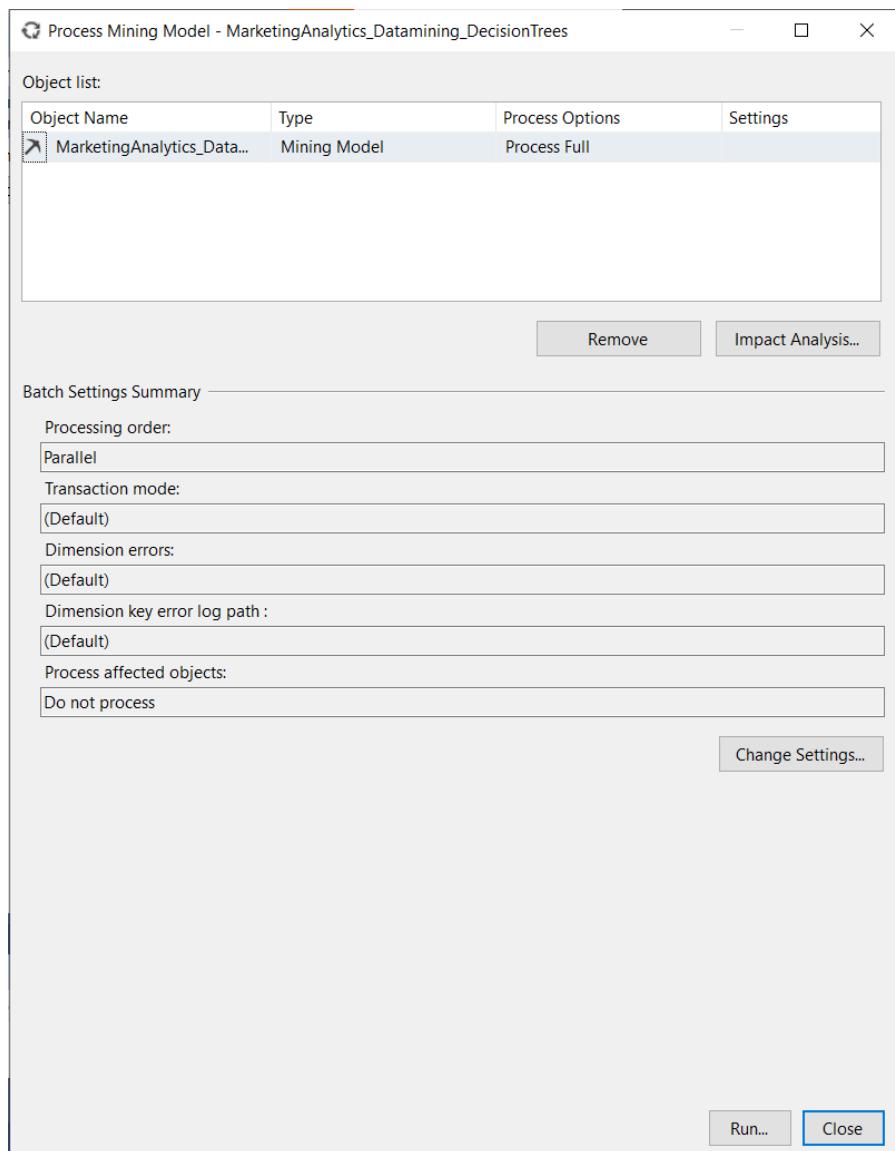
Structure	
Accepted Cmp Overall	Microsoft_Decision_Trees
Complain	Input
Cus Id	Key
Mnt Fish Products	Input
Mnt Fruits	Input
Mnt Gold Prods	Input
Mnt Meat Products	Input
Mnt Sweet Products	Input
Mnt Wines	Input
Num Catalog Purchases	Input
Num Deals Purchases	Input
Num Store Purchases	Input
Num Web Purchases	Input
Num Web Visits Month	Input
Recency	Input
Response	PredictOnly

Chuyển sang Tab Mining Model Viewer

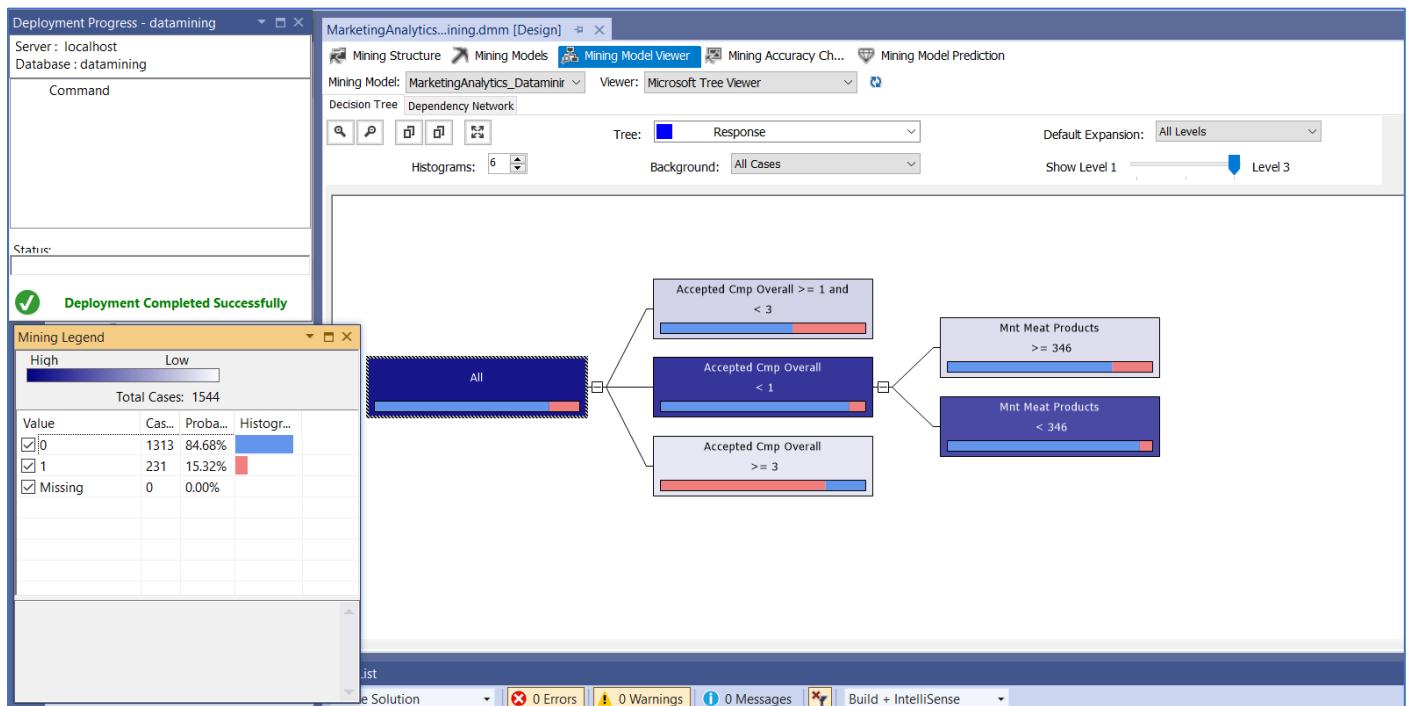
Chọn Yes để tiến hành Deploy project



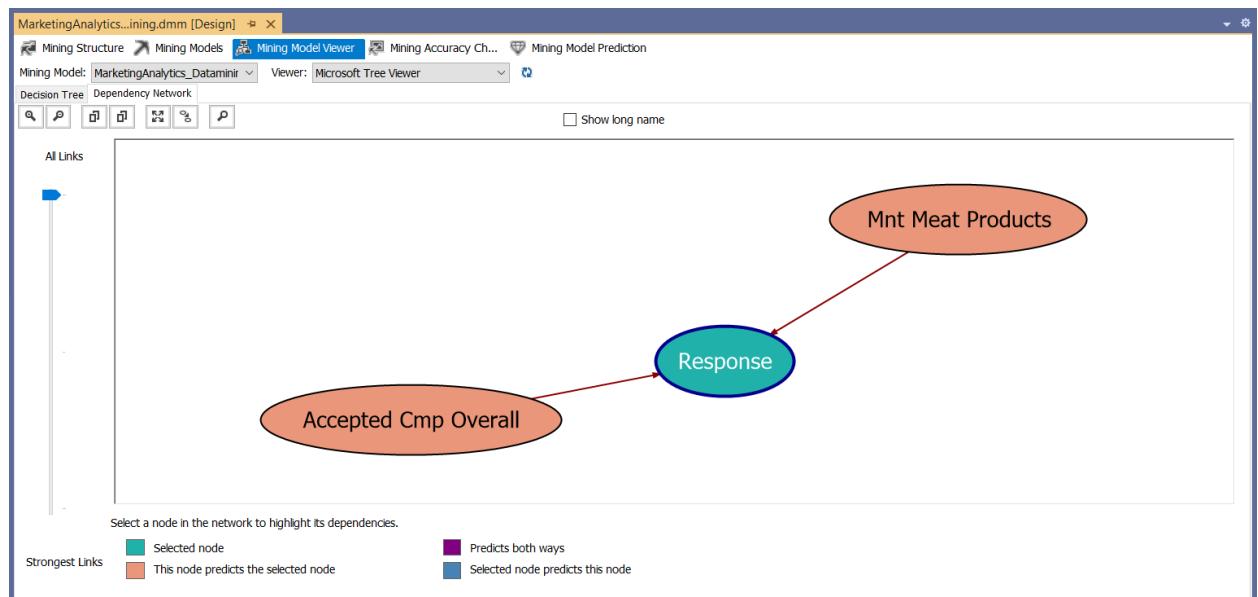
Chọn Run



Kết quả thu được mô hình Decision Trees



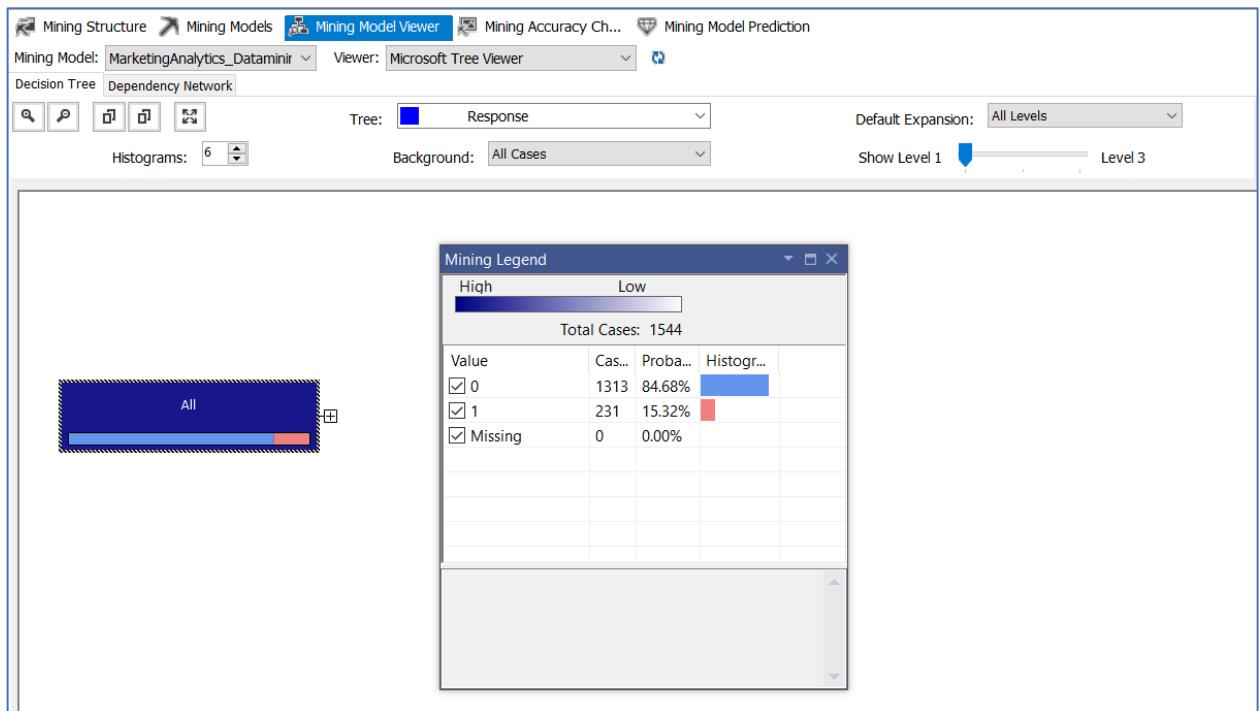
Chuyển sang tab Mining Model Viewer để xem các cấp phụ thuộc



8.4. Phân tích kết quả thuật toán Cây quyết định(Decision Trees) và đưa ra tập luật

Phản hồi của khách hàng (Response) đối với chiến dịch Marketing.

- **Ở mức 1:**

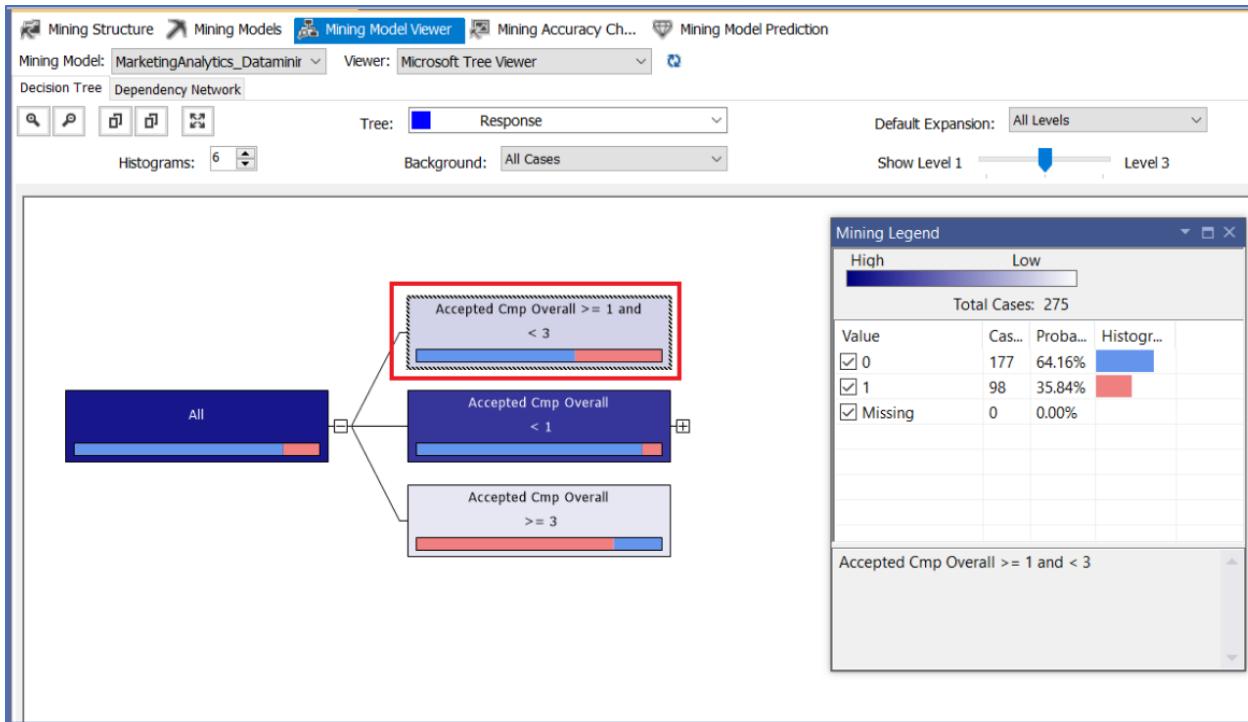


Dựa vào những features đầu vào, tổng quan dự đoán:

- Xác suất khách hàng từ chối chiến dịch Marketing là 84.68%
- Xác suất khách hàng chấp nhận chiến dịch Marketing là 15.32%

- **Ở mức 2:**

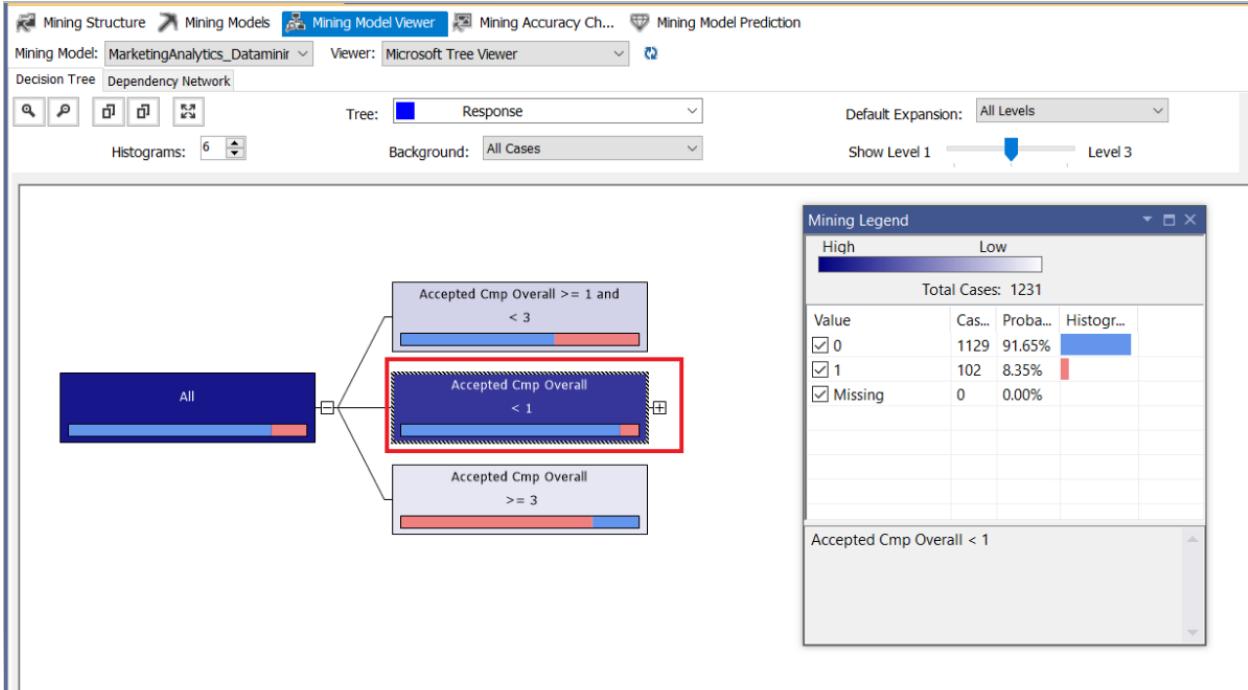
- *Trường hợp 1:* Accepted Cmp Overall ≥ 1 and < 3



Xác suất khách hàng từ chối chiến dịch Marketing là 64.16%

Xác suất khách hàng chấp nhận chiến dịch Marketing là 35.84%

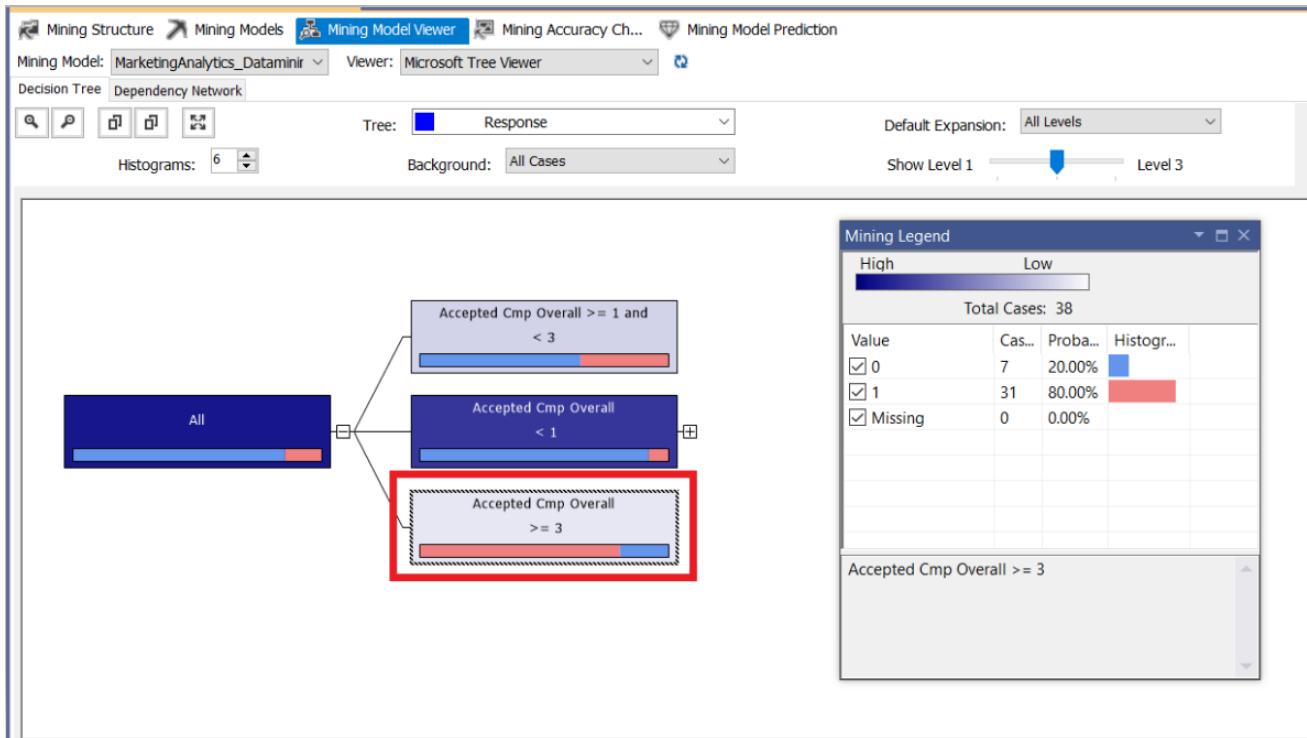
- Trường hợp 2: Accepted Cmp Overall < 1



Xác suất khách hàng từ chối chiến dịch Marketing là 91.65%

Xác suất khách hàng chấp nhận chiến dịch Marketing là 8.35%

- *Trường hợp 3: Accepted Cmp Overall >= 3*



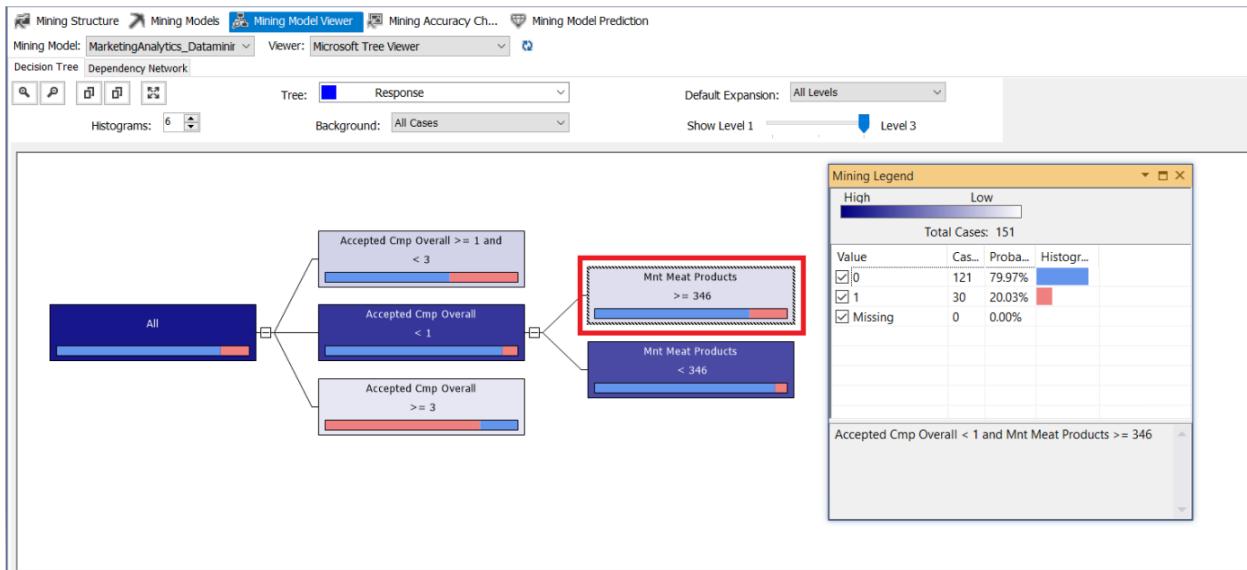
Xác suất khách hàng từ chối chiến dịch Marketing là 20%

Xác suất khách hàng chấp nhận chiến dịch Marketing là 80%

→ Ở mức 2, khách hàng có số lần chấp nhận chiến dịch Marketing Accepted Cmp Overall ≥ 3 được dự đoán có khả năng chấp nhận chiến dịch Marketing cao hơn trong tương lai.

- **Ở mức 3:**

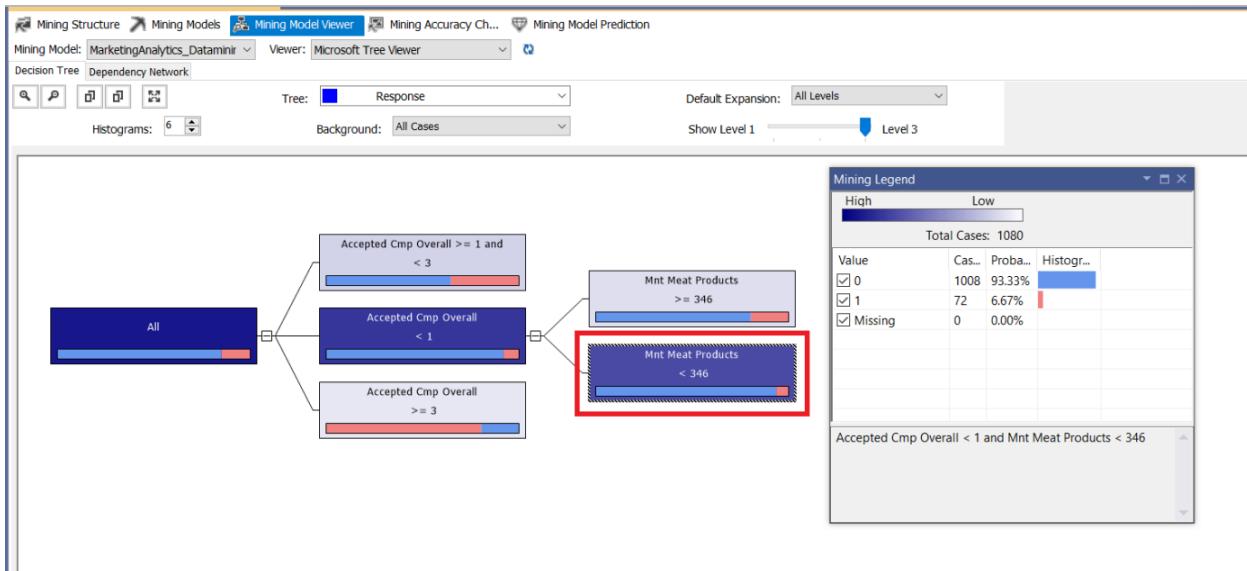
- *Trường hợp 1: Accepted Cmp Overall < 1 and Mnt Meat Products >= 346*



Xác suất khách hàng từ chối chiến dịch Marketing là 79.97%

Xác suất khách hàng chấp nhận chiến dịch Marketing là 20.03%

- *Trường hợp 2: Accepted Cmp Overall < 1 and Mnt Meat Products < 346*



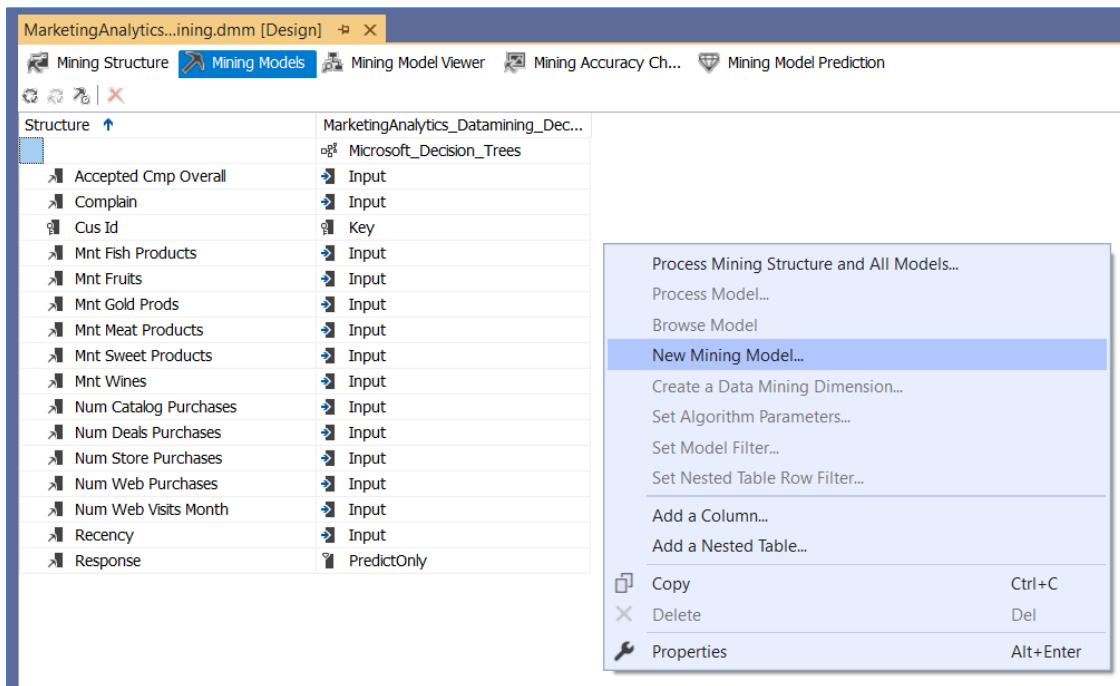
Xác suất khách hàng từ chối chiến dịch Marketing là 93.33%

Xác suất khách hàng chấp nhận chiến dịch Marketing là 6.67%

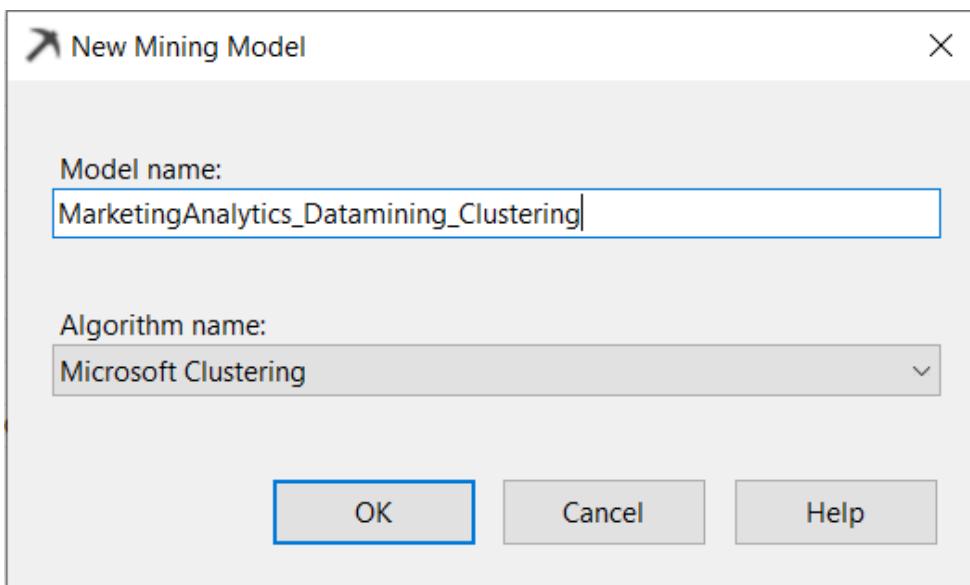
→ Ở mức 3, khách hàng có số lần chấp nhận chiến dịch Marketing Accepted Cmp Overall < 1 và tổng tiền chi trả cho mặt hàng về thịt Mnt Meat Products ≥ 346 được dự đoán có khả năng chấp nhận chiến dịch Marketing cao hơn trong tương lai.

8.5. So sánh độ chính xác với các thuật toán khác

Tại tab Mining Models, nhấp chuột phải chọn New Mining Model



Chọn thuật toán Clustering



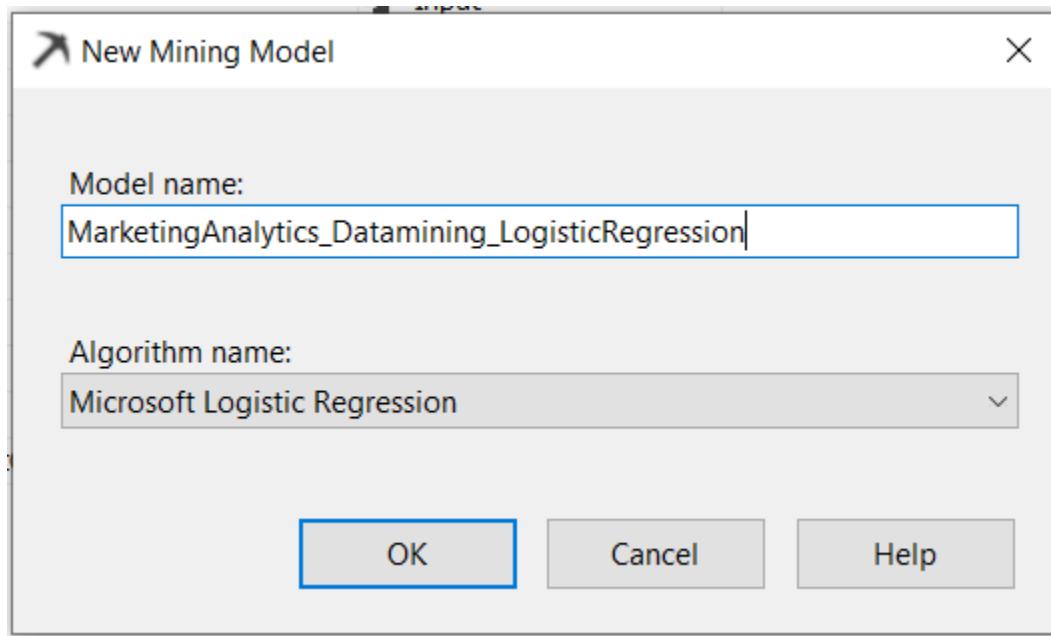
MarketingAnalytics...ining.dmm [Design] * X

Mining Structure Mining Models Mining Model Viewer Mining Accuracy Ch... Mining Model Prediction

Structure ↑ MarketingAnalytics_Datamining_De... MarketingAnalytics_Data...

Accepted Cmp Overall	Microsoft_Decision_Trees	Microsoft_Clustering
Complain	Input	Input
Cus Id	Key	Key
Mnt Fish Products	Input	Input
Mnt Fruits	Input	Input
Mnt Gold Prods	Input	Input
Mnt Meat Products	Input	Input
Mnt Sweet Products	Input	Input
Mnt Wines	Input	Input
Num Catalog Purchases	Input	Input
Num Deals Purchases	Input	Input
Num Store Purchases	Input	Input
Num Web Purchases	Input	Input
Num Web Visits Month	Input	Input
Recency	Input	Input
Response	PredictOnly	PredictOnly

Tương tự, tạo thêm thuật toán Logistic Regression



Structure	MarketingAnalytics_Datamining_De...	MarketingAnalytics_Data...	MarketingAnalytics_Data...
Accepted Cmp Overall	Microsoft_Decision_Trees	Microsoft_Clustering	Microsoft_Logistic_Re...
Complain	Input	Input	Input
Cus Id	Key	Key	Key
Mnt Fish Products	Input	Input	Input
Mnt Fruits	Input	Input	Input
Mnt Gold Prods	Input	Input	Input
Mnt Meat Products	Input	Input	Input
Mnt Sweet Products	Input	Input	Input
Mnt Wines	Input	Input	Input
Num Catalog Purchases	Input	Input	Input
Num Deals Purchases	Input	Input	Input
Num Store Purchases	Input	Input	Input
Num Web Purchases	Input	Input	Input
Num Web Visits Month	Input	Input	Input
Recency	Input	Input	Input
Response	PredictOnly	PredictOnly	PredictOnly

Tiến hành Deploy và Process

Process Progress

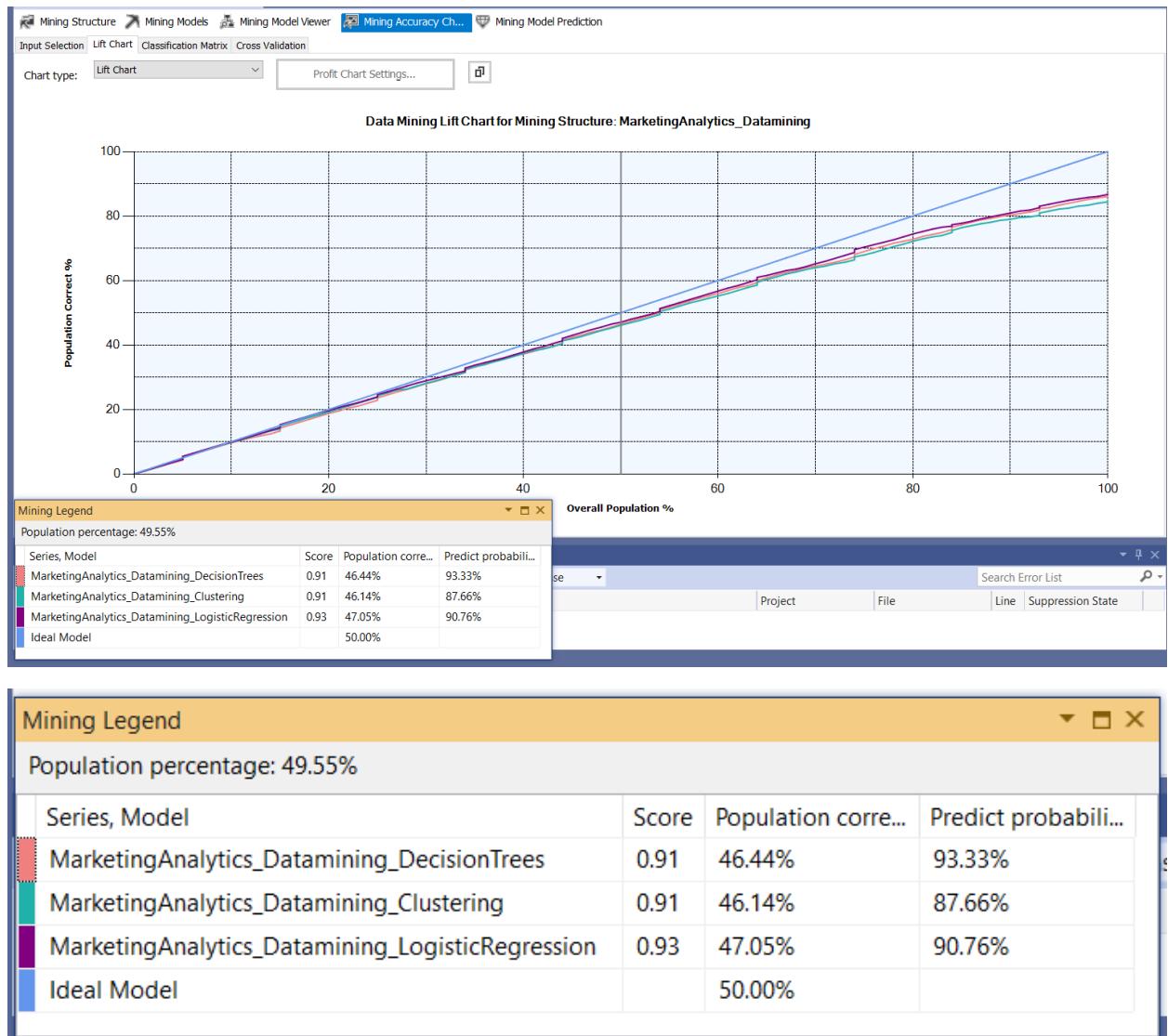
- Command
 - Processing Mining Structure 'MarketingAnalytics_Datamining'.
 - Start time 5/21/2022 2:51:33 PM
 - Processing Mining Model 'MarketingAnalytics_Clustering' - In Progress - 29 of 1000.
 - Processing Mining Model 'MarketingAnalytics_DecisionTrees' - In Progress - 1000 of 1544.
 - Processing Mining Model 'MarketingAnalytics_LogisticRegression' - Training the neural network: 0 subnets left to train.
 - Processing Cube 'MarketingAnalytics_Datamining ~MC'.
 - Start time 5/21/2022 2:51:33 PM
 - Processing Measure Group '~CaseDetail ~MG'.
 - Processing Dimension 'MarketingAnalytics_Datamining ~MC-Cus Id'.

Status:

Process succeeded with warnings. Trace information is still being transferred. If you do not want to wait for all of the information to transfer, click Close.

Stop Reprocess View Details... Copy Close Help

- Sử dụng đồ thị Lift cho so sánh và đánh giá 3 thuật toán
Tại tab Mining Accuracy Chart → Chart type chọn Lift chart



Thuật toán Decision Trees có xác suất dự đoán Predict probability 93.99% với số điểm 0.91

Thuật toán Clustering có xác suất dự đoán Predict probability 87.66% với số điểm 0.91

Thuật toán Logistic Regression có xác suất dự đoán Predict probability 90.76% với số điểm 0.93

→ Ưu tiên chọn thuật toán Decision Trees

Tài liệu tham khảo

- [1] Slide bài giảng và tài liệu thực hành do giảng viên cung cấp.
- [2] <https://docs.microsoft.com/en-us/power-bi/fundamentals/service-self-service-signup-for-power-bi>

Link Đồ án cuối kỳ OLAP IS217M22HTCL_19522304_19521520

<https://drive.google.com/drive/folders/1RgUW1DwmuvnwSMzRhD3q25ch5qlbA0yE?usp=sharing>