TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

KHOA HỆ THỐNG THÔNG TIN



ĐỒ ÁN MÔN HỌC

KHO DỮ LIỆU VÀ OLAP

ĐỀ TÀI GIẤY PHÉP XÂY DỰNG

SINH VIÊN:

17520973 – Lê Trần Duy Sang

17521284 – Huỳnh Hữu Ý

LỜI CẢM ƠN

Lời đầu tiên, nhóm xin đặc biệt gửi lời cảm ơn đến Cô Đỗ Thị Minh Phụng (Giáo viên lý thuyết và thực hành môn Phân tích thiết kế hệ thống thông tin). Cô đã cung cấp kiến thức, chỉ bảo và đóng góp những ý kiến quý báu giúp nhóm hoàn thành được đồ án môn học của mình.

Xuất phát từ mục đích học tập phương pháp phân tích, thiết kế hệ thống hệ thống thông tin, cũng như tìm hiểu thêm về quy trình nghiệp vụ của các khách sạn, nhóm chúng em đã thực hiện đồ án “Thiết kế hệ thống quản khách sạn Tam Ngư”. Trong quá trình thực hiện đồ án, dựa trên những kiến thức được Cô cung cấp trên trường kết hợp với việc tự tìm hiểu những công cụ và kiến thức mới, nhóm đã cố gắng thực hiện đồ án một cách tốt nhất. Tuy nhiên, đồ án chưa hoàn thiện và còn nhiều sai sót nhưng nó là kết quả của sự nỗ lực của các thành viên trong nhóm, sự giúp đỡ của tất cả bạn bè và Cô.

Nhóm rất mong nhận sự góp ý từ phía cô nhằm rút ra những kinh nghiệm quý báu và hoàn thiện vốn kiến thức để nhóm có thể tiếp tục hoàn thành những đồ án khác trong tương lai.

Xin chân thành cảm ơn quý Cô!

NHẬN XÉT GIÁO VIÊN HƯỚNG DẪN

CHƯƠNG 1: TỔNG QUAN ĐỀ TÀI

1 Lý do chọn đề tài

2 Mô tả dữ liệu

2.1. Giới thiệu chung

2.2. Kiểu dữ liệu

3. Tiền xử lý dữ liệu

4. Thiết kế kho dữ liệu

4.1. FACT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Tên thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | FACT\_ID | INT | Khóa chính | Mã FACT |
| 2 | START\_DATE\_ID | INT | Khóa ngoại | Mã ngày nộp giấy phép |
| 3 | ISSUE\_DATE\_ID | INT | Khóa ngoại | Mã ngày giấy phép được duyệt |
| 4 | PERMIT\_TYPE | INT | Khóa ngoại | Mã loại giấy phép |
| 5 | REVIEW\_TYPE | INT | Khóa ngoại | Mã qui trình xin giấy phép |
| 6 | CONTACT\_UNIT\_ID | INT | Khóa ngoại | Mã đơn vị xin giấy phép |
| 7 | STREET\_ID | INT | Khóa ngoại | Mã nơi xin giấy phép xây dựng |
| 8 | PROCESSING\_TIME | INT |  | Thời gian từ ngày nộp cho đến ngày được duyệt |
| 9 | BUILDING\_FEE\_PAID | FLOAT |  | Chi phí giấy phép xây dựng phải trả |
| 10 | ZONING\_FEE\_PAID | FLOAT |  | Chi phí |
| 11 | OTHER\_FEE\_PAID | FLOAT |  | Các loại chi phí giấy phép khác phải trả |
| 12 | SUBTOTAL\_PAID | FLOAT |  | Tổng chi phí phải trả |
| 13 | BUILDING\_FEE\_UNPAID | FLOAT |  | Chi phí giấy phép xây dựng được miễn |
| 14 | ZONING\_FEE\_UNPAID | FLOAT |  | Chi phí |
| 15 | OTHER\_FEE\_UNPAID | FLOAT |  | Các loại phí giấy phép khác được miễn |
| 16 | SUBTOTAL\_UNPAID | FLOAT |  | Tổng chi phí được miễn |
| 17 | BUILDING\_FEE\_WAIVED | FLOAT |  | Chi phí giấy phéo xây dựng còn nợ |
| 18 | ZONING\_FEE\_WAIVED | FLOAT |  |  |
| 19 | OTHER\_FEE\_WAIVED | FLOAT |  | Các loại phí giấy phép khác còn nợ |
| 20 | SUBTOTAL\_WAIVED | FLOAT |  | Tổng chi phí còn nợ |
| 21 | SUBTOTAL | FLOAT |  | Tổng tất cả các chi phí phải trả, được miển, còn nợ |
| 22 | REPORT\_COST | FLOAT |  | Chi phí phải trả do đơn vị xin giấy phép dữ đoán |

4.2 DIM\_START\_DATE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | START\_DATE\_ID | INT | Khóa chính | Mã ngày nộp giấy phép |
| 2 | START\_DATE | DATE |  | Ngày nộp giấy phép |
| 3 | DAY | INT |  | Ngày trong tháng |
| 4 | MONTH | INT |  | Tháng |
| 5 | QUARTER | INT |  | Quý |
| 6 | YEAR | INT |  | Năm |

4.3 DIM\_ISSUE\_DATE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | START\_ISSUE\_ID | INT | Khóa chính | Mã ngày giấy phép được duyệt |
| 2 | ISSUE\_DATE | DATE |  | Ngày giấy phép được duyệt |
| 3 | DAY | INT |  | Ngày trong tháng |
| 4 | MONTH | INT |  | Tháng |
| 5 | QUARTER | INT |  | Quý |
| 6 | YEAR | INT |  | Năm |

4.4 DIM\_PERMIT\_TYPE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | PERMIT\_TYPE\_ID | INT | Khóa chính | Mã loại giấy phép xây dựng |
| 2 | PERMIT\_TYPE\_NAME | VARCHAR(300) |  | Tên loại giấy phép xây dựng |

4.5 DIM\_REVIEW\_TYPE

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | REVIEW\_TYPE\_ID | INT | Khóa chính | Mã loại quy trình xin giấy phép xây dựng |
| 2 | REVIEW\_TYPE\_NAME | VARCHAR(300) |  | Tên loại quy trình |

4.6 DIM\_STREET

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | STREET\_ID | INT | Khoá chính | Mã đường |
| 2 | STREET\_DIRECTION\_ID | INT | Khóa ngoại | Mã hướng đường |
| 3 | WARD\_ID | INT | Khóa ngoại | Mã quận |
| 4 | STREET\_NAME | VARCHAR(300) |  | Tên đừng |

4.7 DIM\_STREET\_DIRECTION

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | STREET\_DIRECTION\_ID | INT | Khóa chính | Mã hướng đường |
| 2 | STREET\_DIRECTION\_NAME | VARCHAR(300) |  | Tên hướng đường |

4.8 DIM\_WARD

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | WARD\_ID | INT | Khóa chính | Mã quận |
| 2 | WARD\_NAME | VARCHAR(300)\_ |  | Tên quận |

4.9 DIM\_CONTACT\_UNIT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| STT | Thuộc tính | Kiểu dữ liệu | Ràng buộc | Ý nghĩa / Ghi chú |
| 1 | CONTACT\_TYPE\_ID | INT | Khóa chính | Mã đơn vị xin giấy phép |
| 2 | CONTACT\_TYPE\_NAME | VARCHAR(500) |  | Tên đơn vị xin giấy phép |

CHƯƠNG 2: XÂY DỰNG KHO DỮ LIỆU (SSIS)

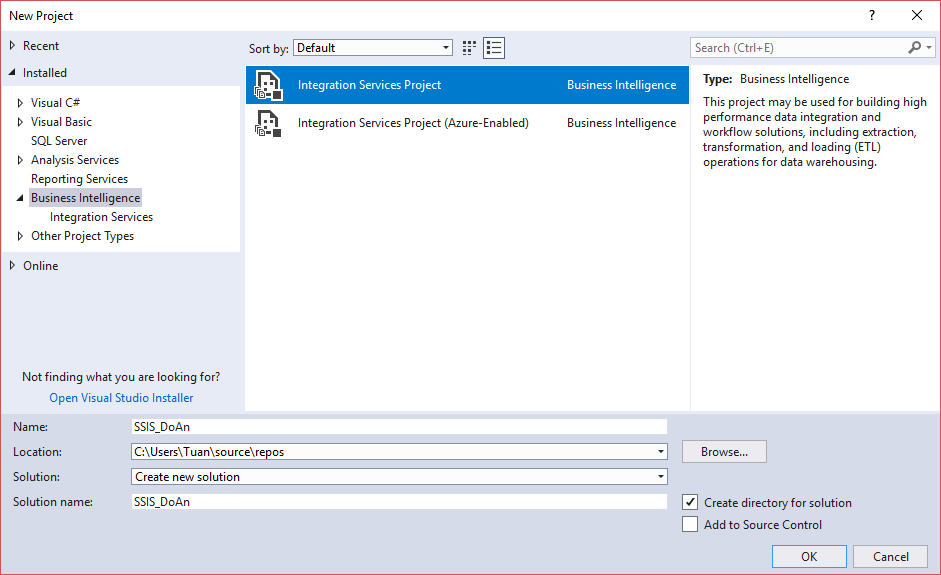
1. Tạo mới Project và thiết lập kết nối

1.1 Tạo mới “Integration Services Project”

Mở Visual Studio 2017(SSDT). Chọn File ->New->Project.

Ở Installed, chọn tab Business Intelligence -> Integration Services Project.

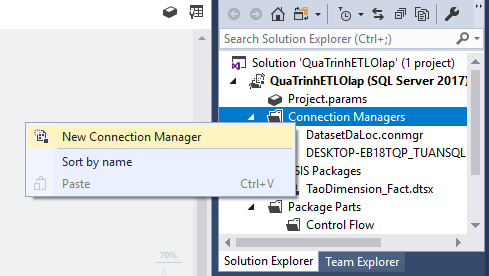
Sau khi đặt tên file, chọn nơi lưu trữ chọn OK để hoàn tất.



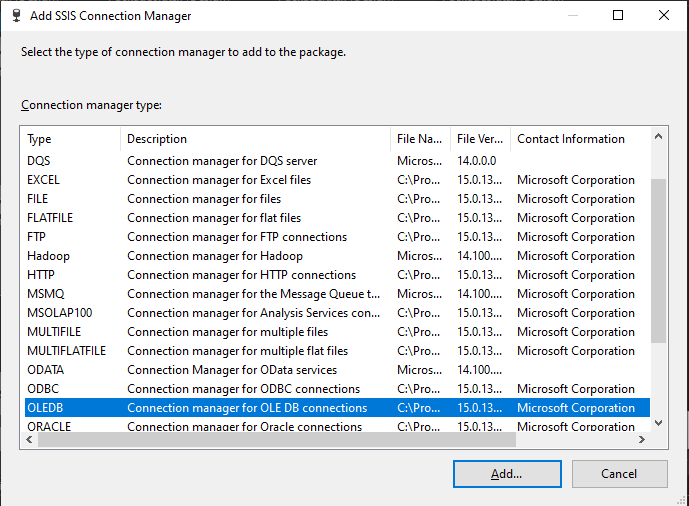
1.2 Thêm và cấu hình kết nối

1.2.1 Kết nối tới Database “DATAWAREHOUSE”

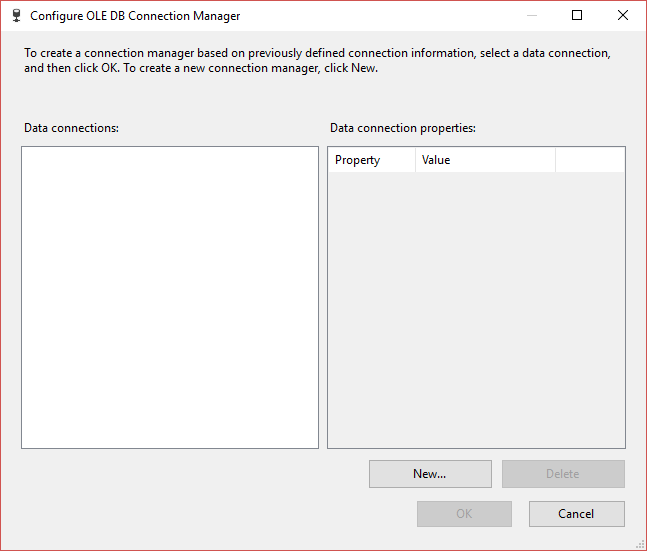
Tại phần Solution Explorer, click chuột phải vào folder Connection Managers, sau đó chọn New Connection Manager để tiến hành tạo kết nối mới.



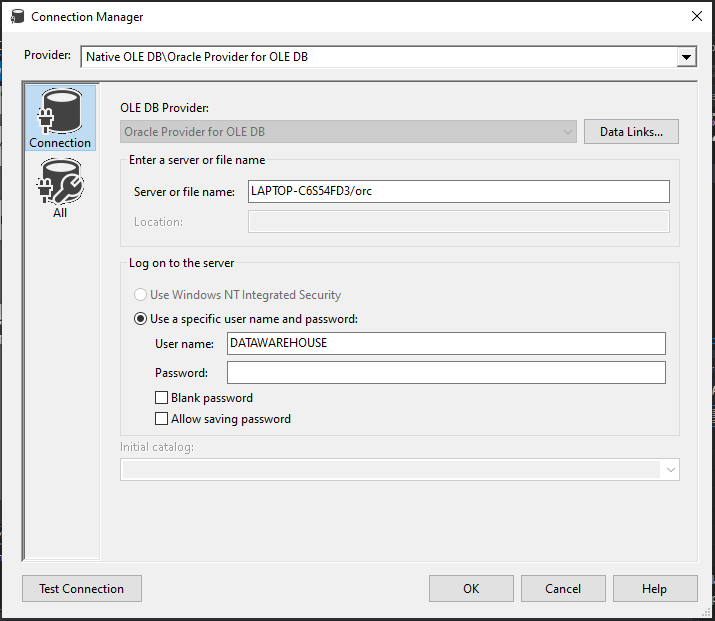
Chọn OLEDB, sau đó chọn Add để kết nối tới hệ quản trị Oracle:



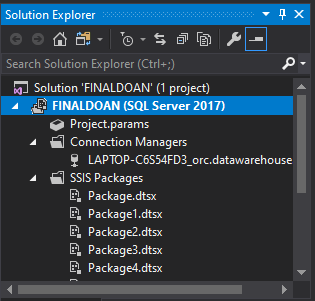
Chọn New để tiến hành thêm Connection:



Điền tên Oracle vào ô Server name, điển Username và Password của Connection. Sau đó chọn 2 lần OK để hoàn tất.



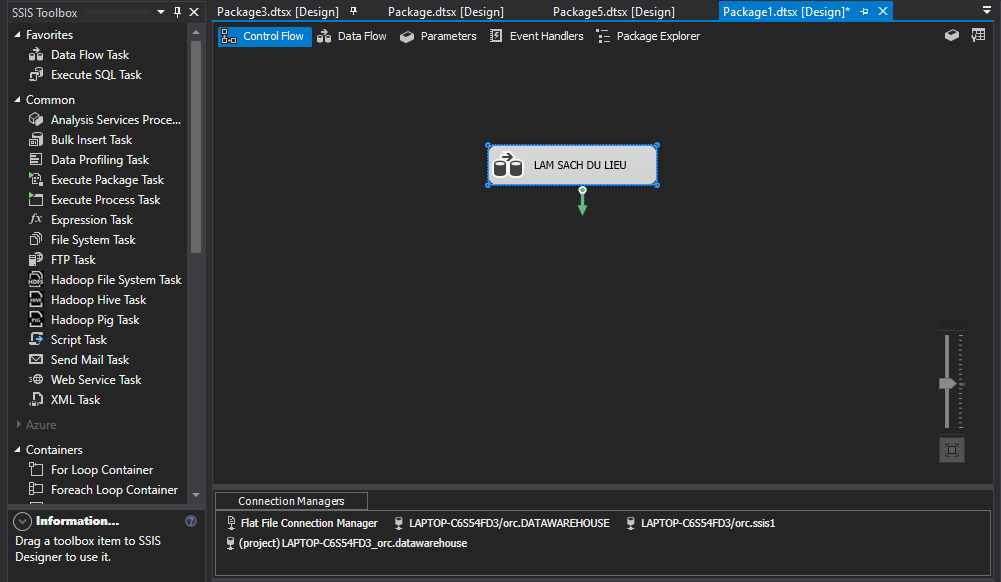
Kết nối thành công tới Database DATAWAREHOUSE:



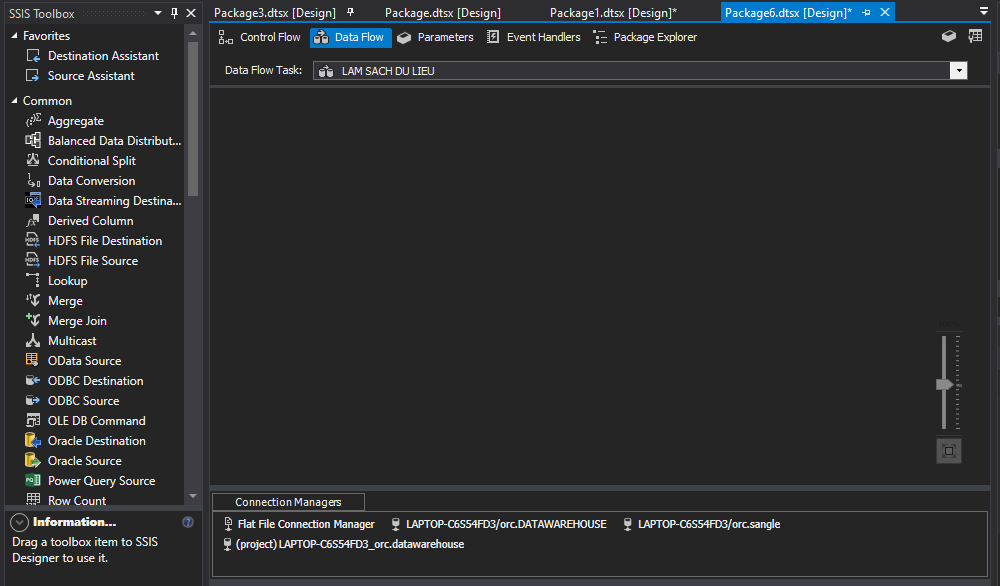
2. Quá trình làm sạch dữ liệu

Tạo “Data Flow Task”:

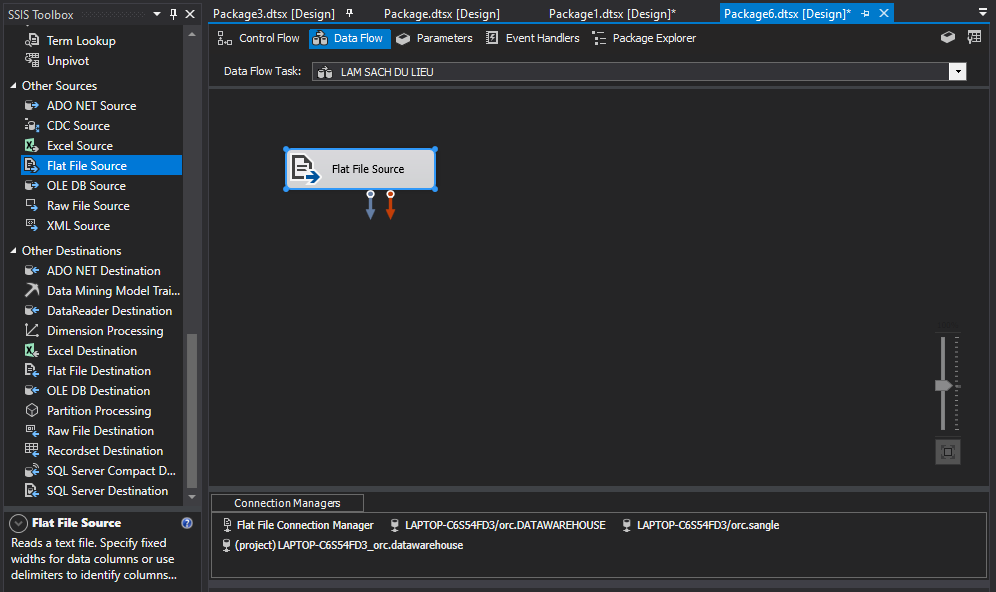
* Kéo thả công cụ Data Flow Task ở SSIS Toolbox sang vùng làm việc Control Flow. Sau đó tiến hành đổi tên “Quá trình làm sạch dữ liệu”:



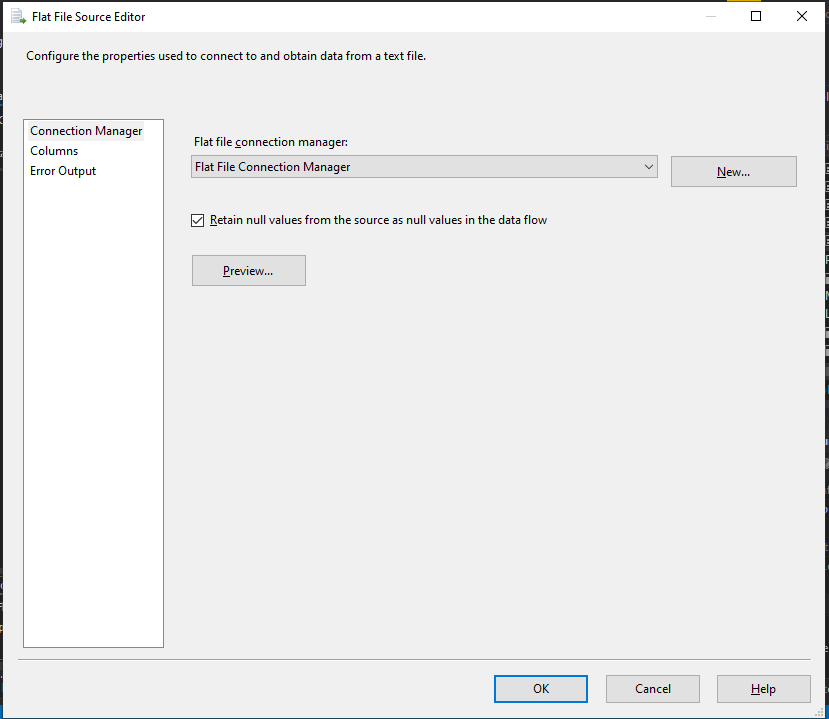
* + Click vào tab Data Flow để tiến hành làm sạch dữ liệu:



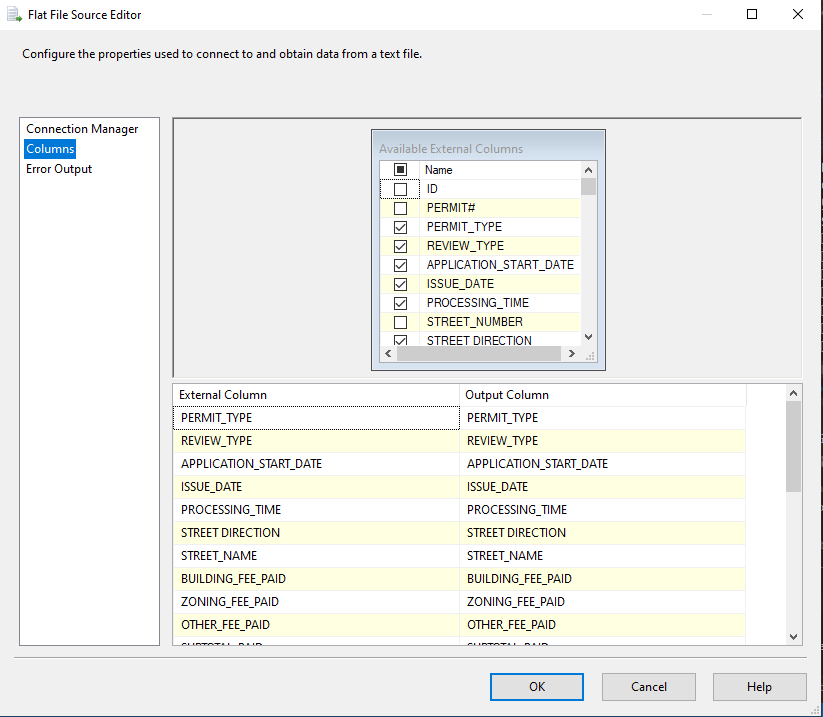
* Tạo và cấu hình “Flat File Source” kết nối tới fille CSV của dữ liệu gốc
  + Kéo thả công cụ “OLE DB Source” ở “SSIS ToolBox”:



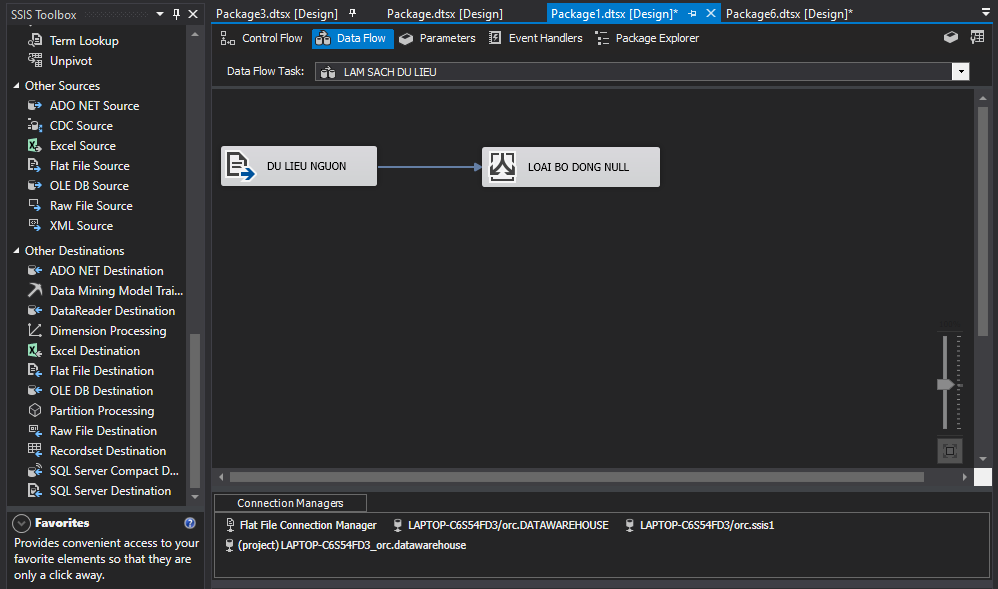
* + Ở mục “Flat file connection manager”, chọn Flat File Connection Manager



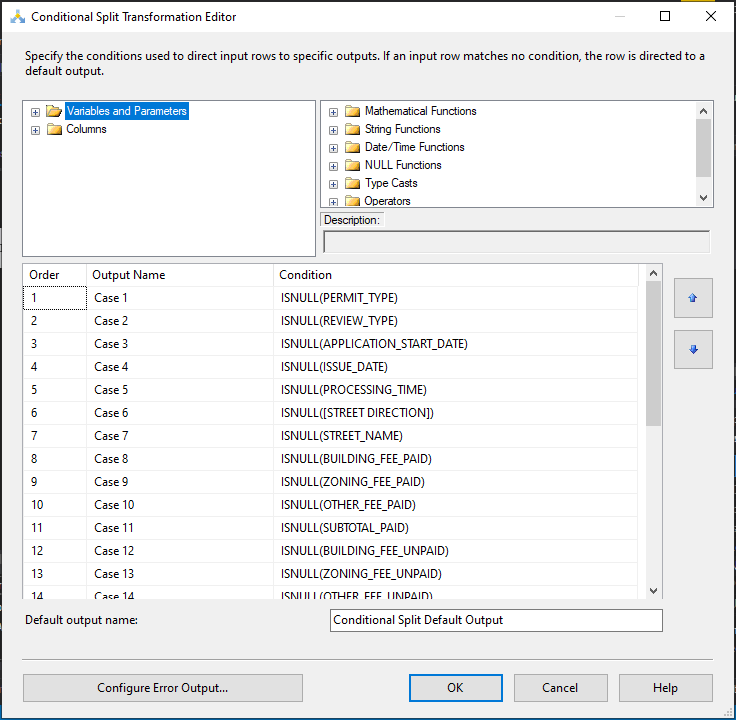
* + Chọn dữ liệu cần để đưa vào bảng FACT



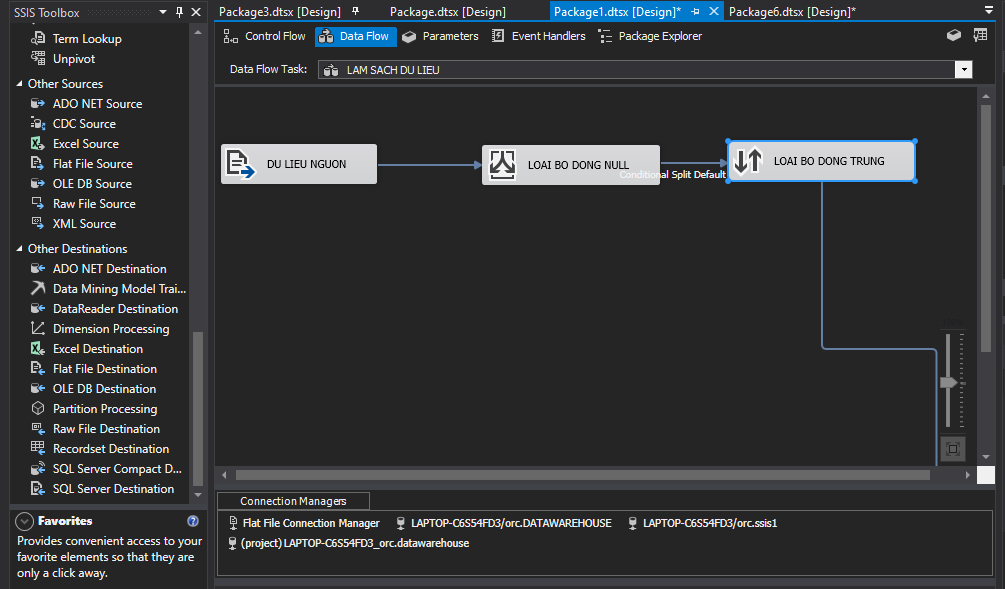
* Tạo và cấu hình “Condition Split” để lọc dữ liệu NULL
  + Kéo thả công cụ “Conditional Split” ở “SSIS Toolbox”:



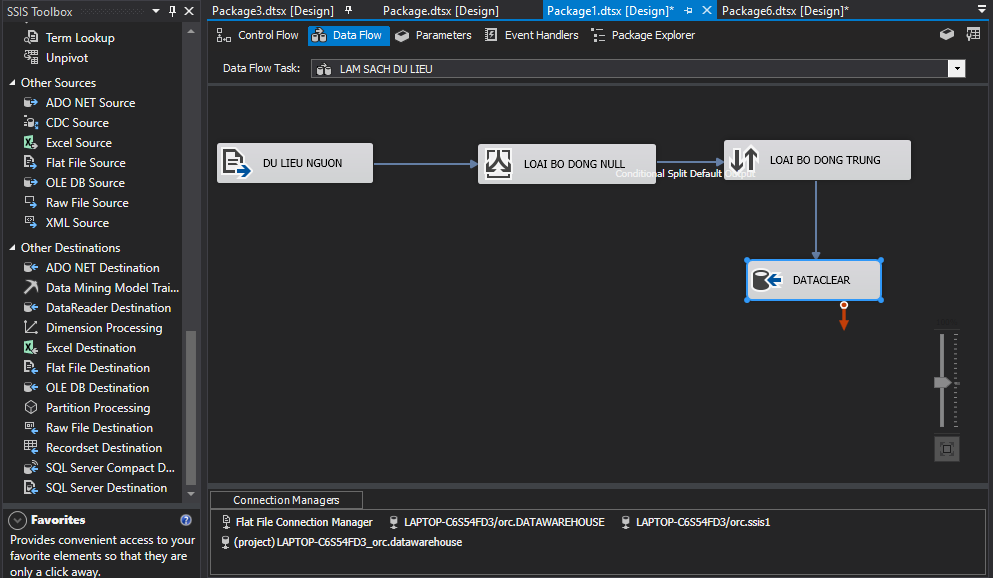
Click đúp vào “Conditional Split” để tiến hành cấu hình. Sử dụng hàm ISNULL( «expression» ) để lọc NULL:



* Tạo và cấu hình “Data Conversion”
  + Kéo thả công cụ “Data Conversion” từ “SSIS Toolbox”:



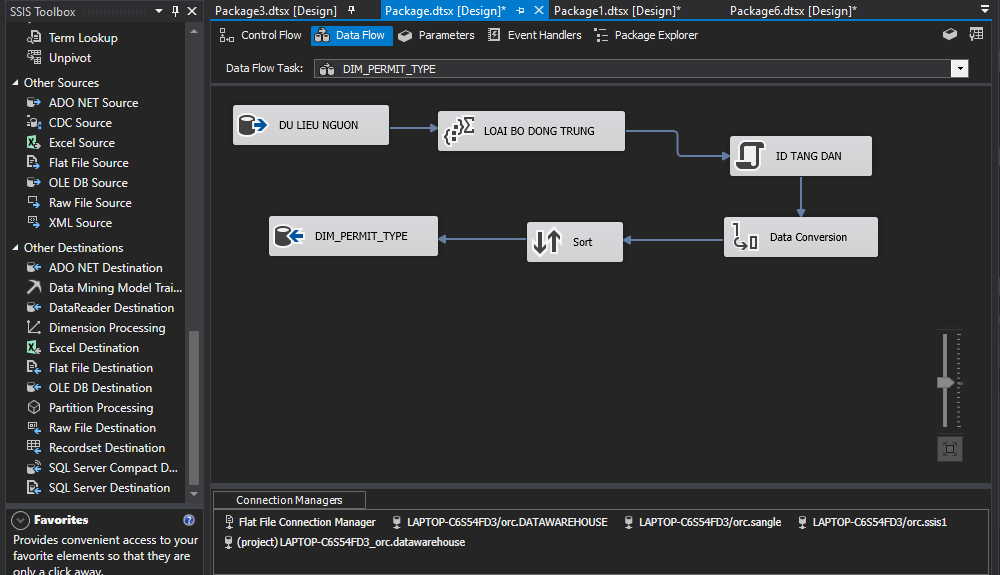
* Tạo và cấu hình “OLE DB Destination” kết nối tới bảng DATA\_CLEAR của Database
  + Kéo thả công cụ “OLE DB Destination” từ “SSIS Toolbox”:



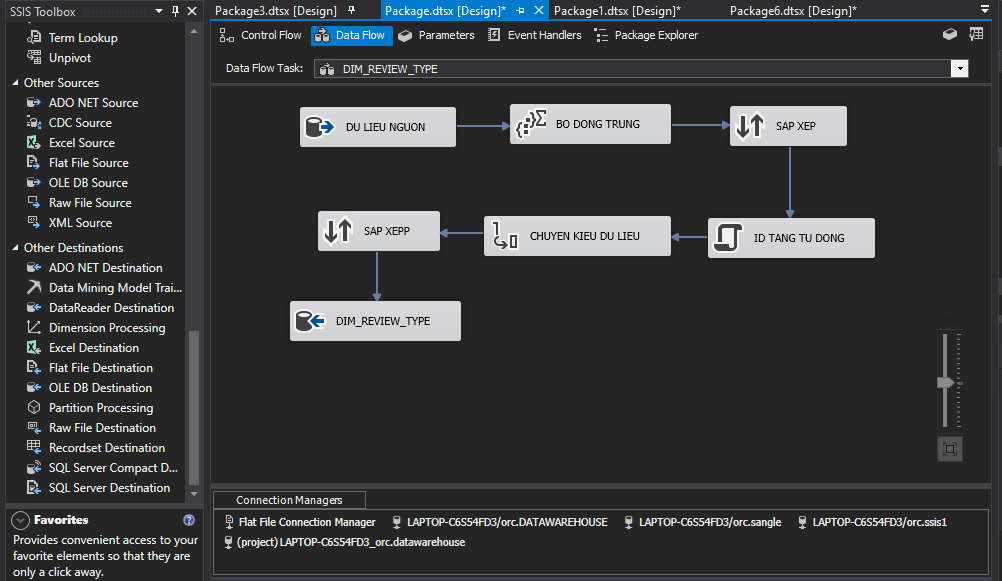
* + Click đúp vào “OLE DB Destination” vừa tạo ở trên, chọn New để tiến hành tạo 1 bảng “DATA\_CLEAR” trên Connection “DATAWAREHOUSE”:

4. Quá trình tạo các bảng “Dimension”

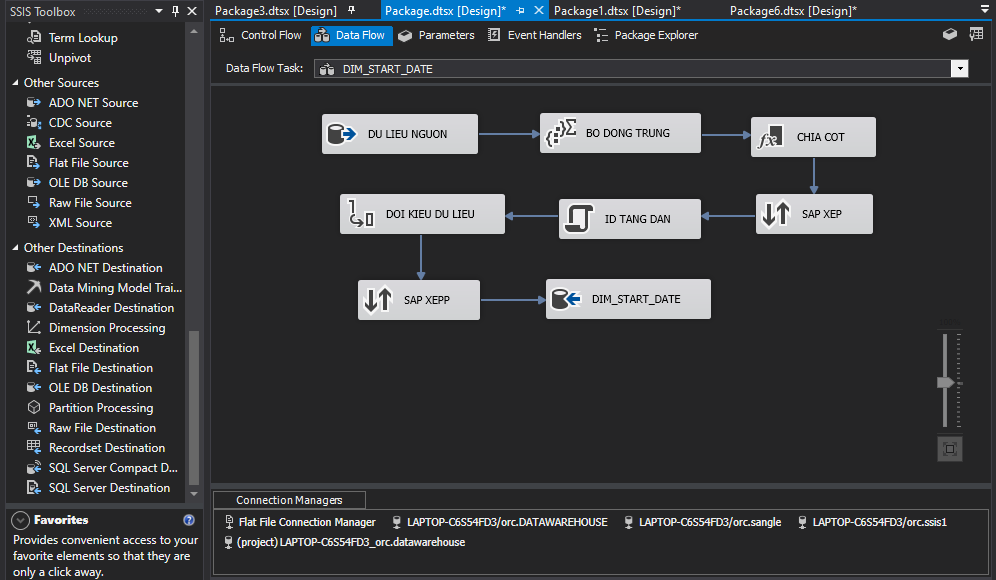
4.1 Quá trình tạo bảng “DIM\_PERMIT\_TYPE”



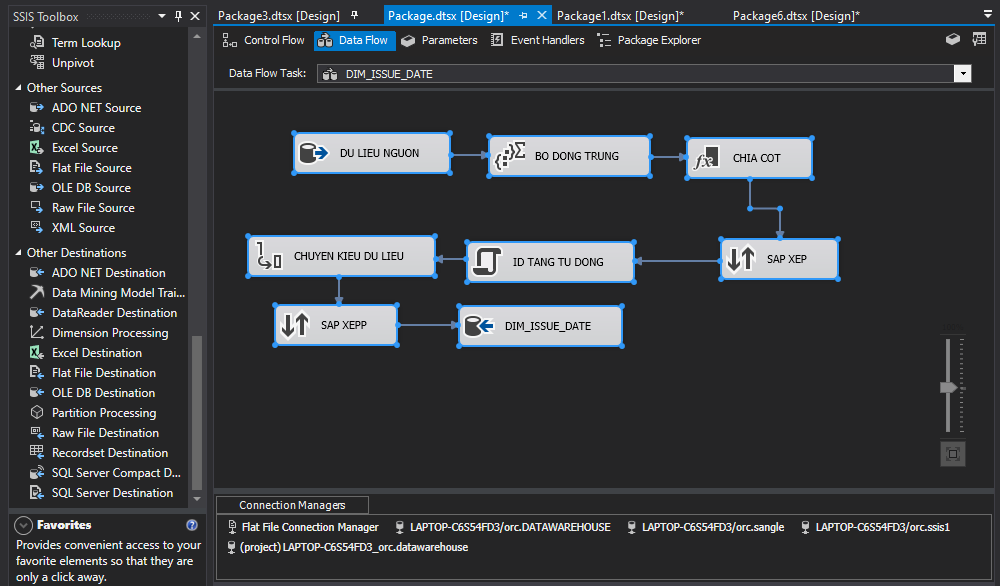
4.2 Quá trình tạo bảng “DIM\_REVIEW\_TYPE”



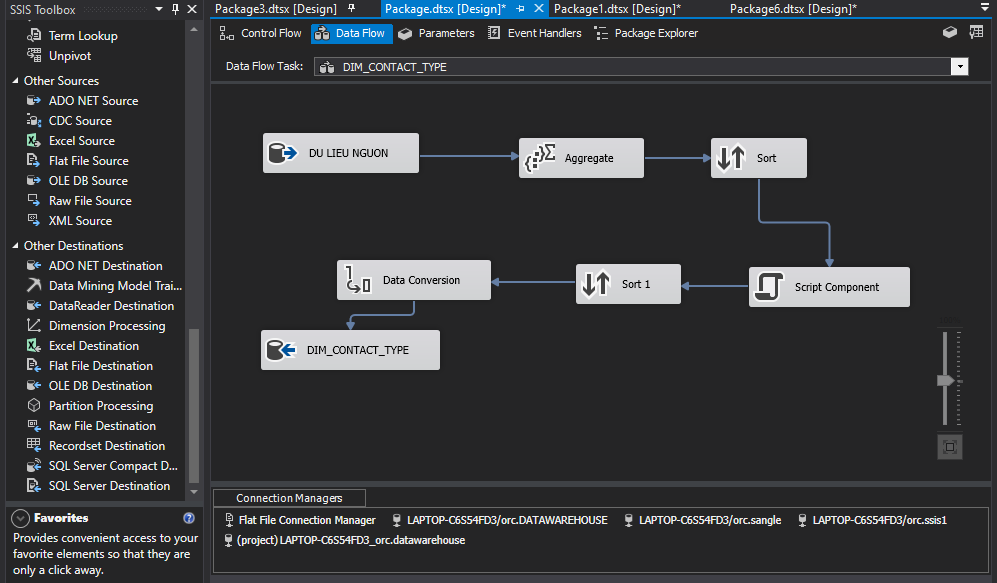
4.3 Quá trình tạo bảng “DIM\_START\_DATE”



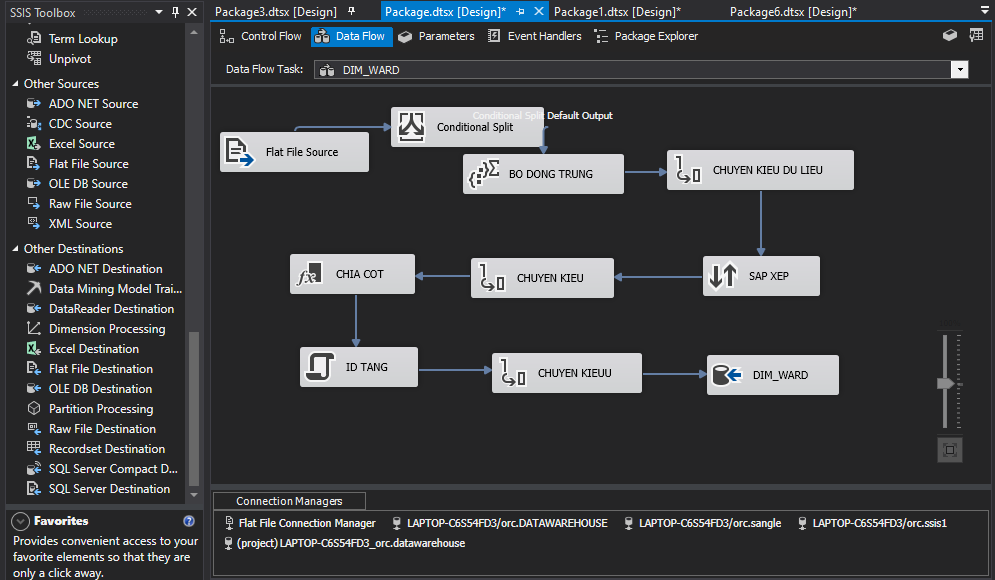
4.4 Quá trình tạo bảng “DIM\_ISSUE\_DATE”



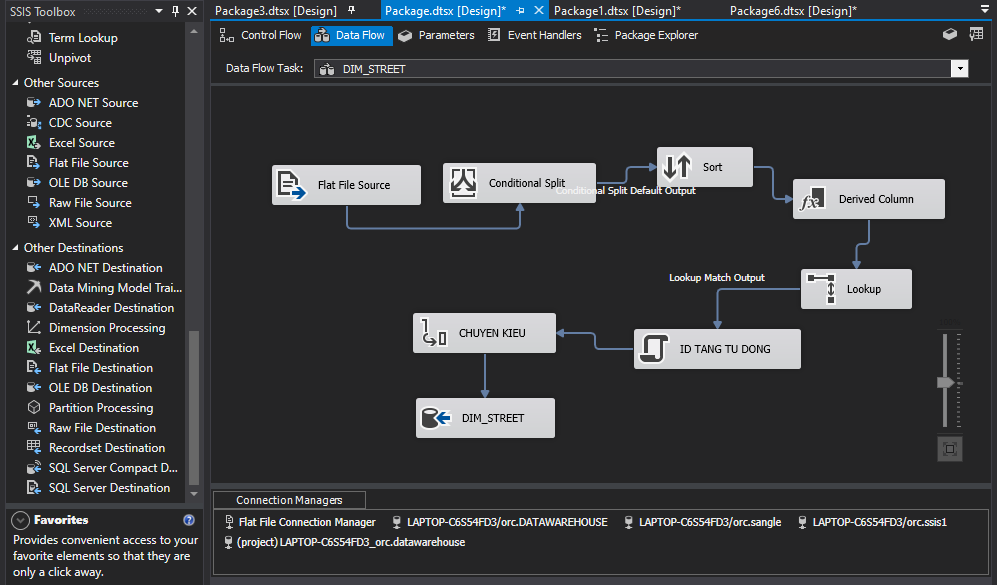
4.5 Quá trình tạo bảng “DIM\_CONTACT\_TYPE”



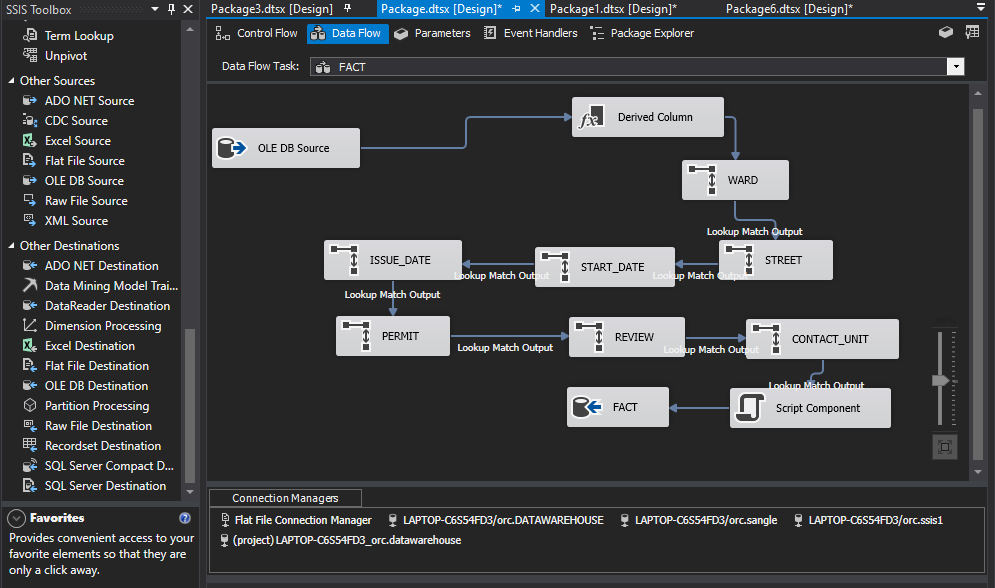
4.6 Quá trình tạo bảng “DIM\_WARD”



4.7 Quá trình tạo bảng “DIM\_STREET”



5. Quá trình tạo bảng FACT



CHƯƠNG 3: PHÂN TÍCH KHO DỮ LIỆU (SSAS)

1. Tạo Project SSAS

2. Tạo Data Source

3. Tạo Data Source View

4. Tạo Cube

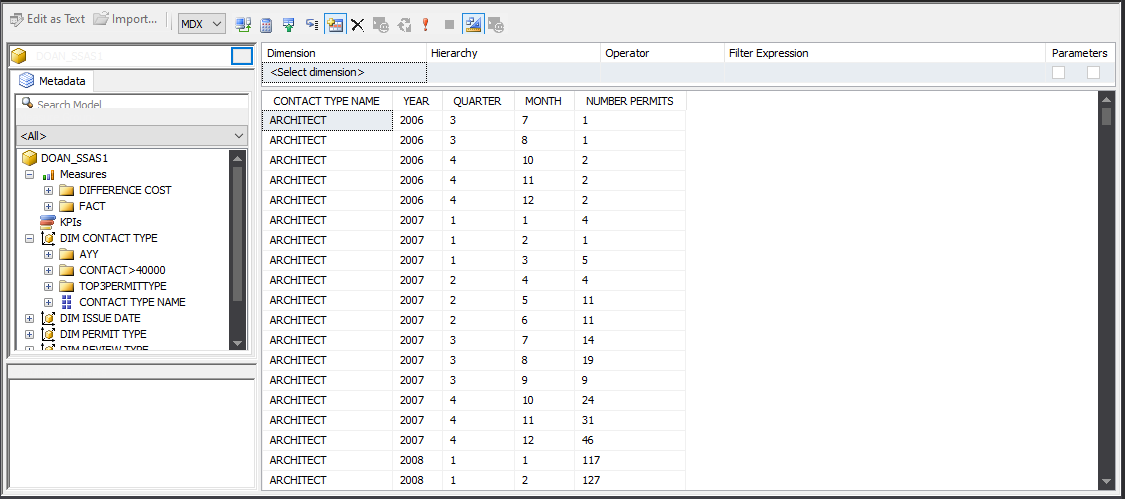
5. Chỉnh sửa các Dimension

6. Chạy tiến trình để deployment

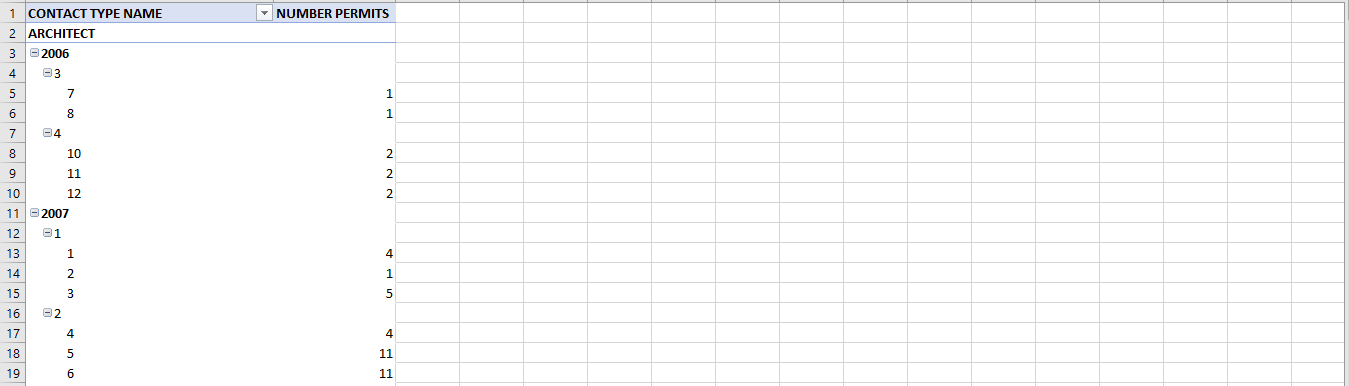
**CHƯƠNG 4: TRUY VẤN**

**1. Thống kê số lượng giấy phép xây dựng được nộp qua từng tháng, quý, năm của mỗi loại đơn vị xin giấy phép xây dựng.**

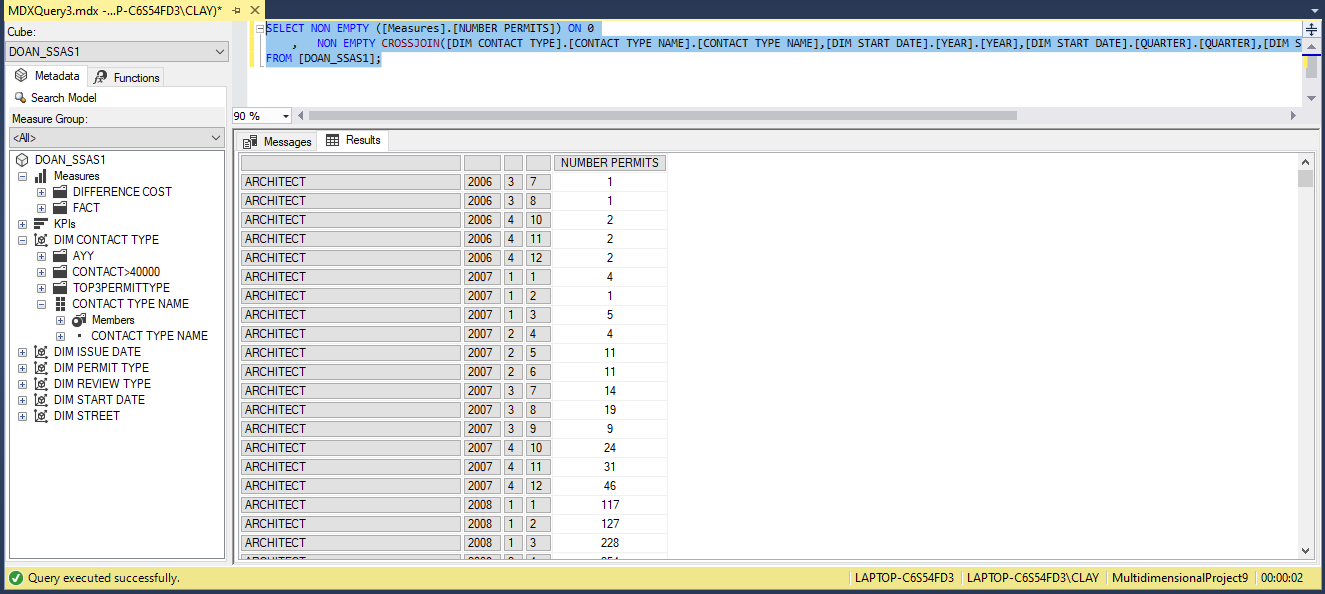
1.1 Truy vấn dữ liệu sử dụng sử dụng BI



1.2 Truy vấn dữ liệu sử dụng Pivot Table



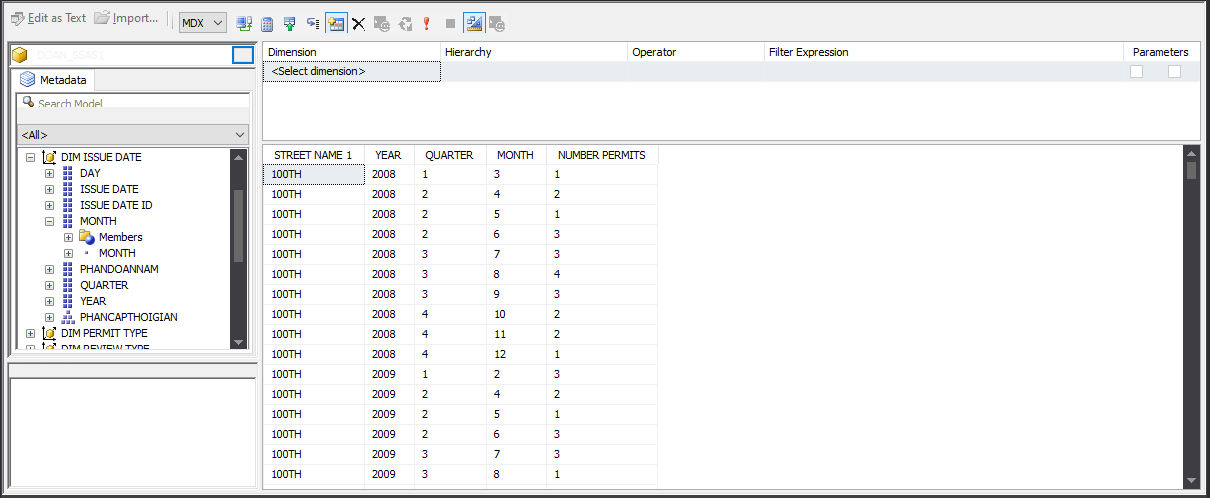
1.3 Truy vấn sử dụng MDX



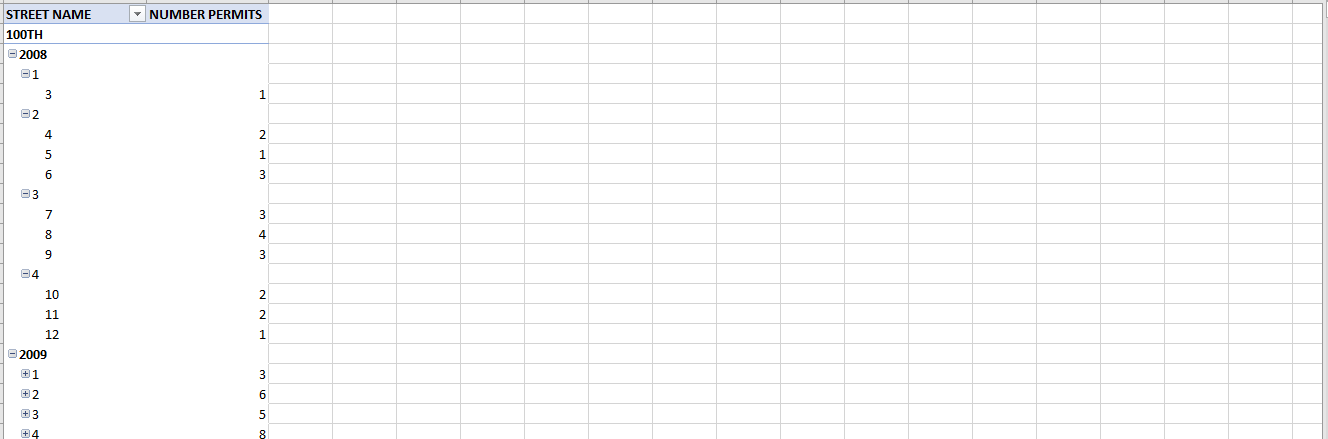
|  |
| --- |
| SELECT NON EMPTY ([Measures].[NUMBER PERMITS]) ON 0  , NON EMPTY CROSSJOIN([DIM CONTACT TYPE].[CONTACT TYPE NAME].[CONTACT TYPE NAME],[DIM START DATE].[YEAR].[YEAR],[DIM START DATE].[QUARTER].[QUARTER],[DIM START DATE].[MONTH].[MONTH]) ON 1  FROM [DOAN\_SSAS1]; |

**2. Thống kê số lượng giấy phép xây dựng được duyệt của mỗi con đường theo từng năm, quý và tháng**

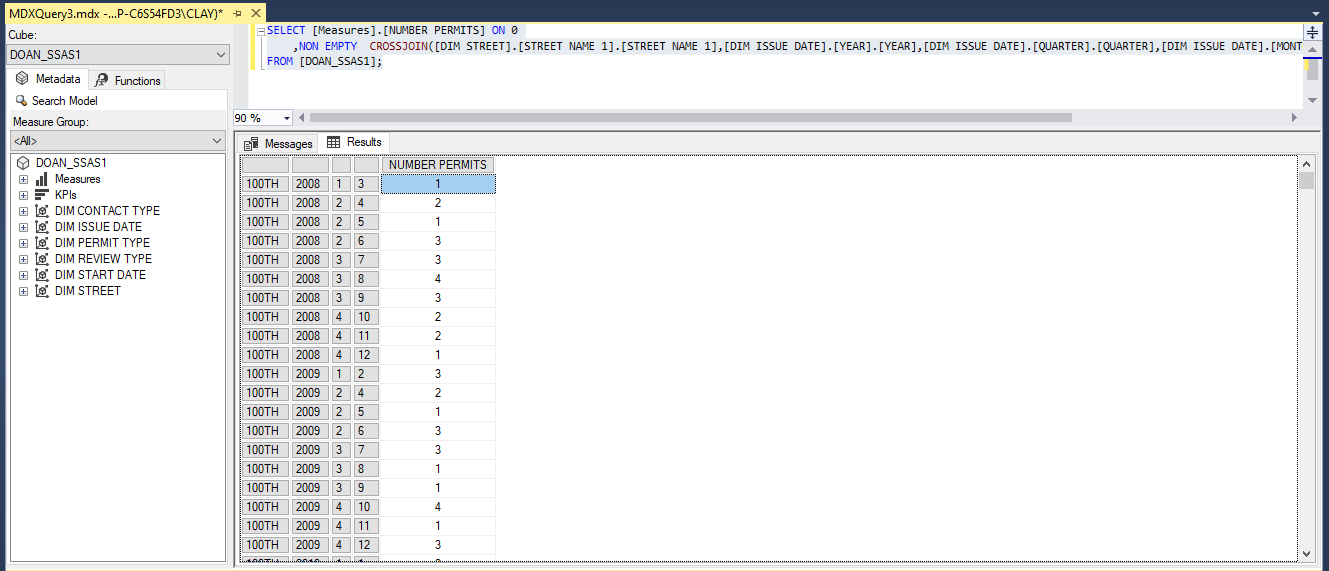
2.1 Truy vấn dữ liệu sử dụng sử dụng BI



2.2 Truy vấn dữ liệu sử dụng Pivot Table



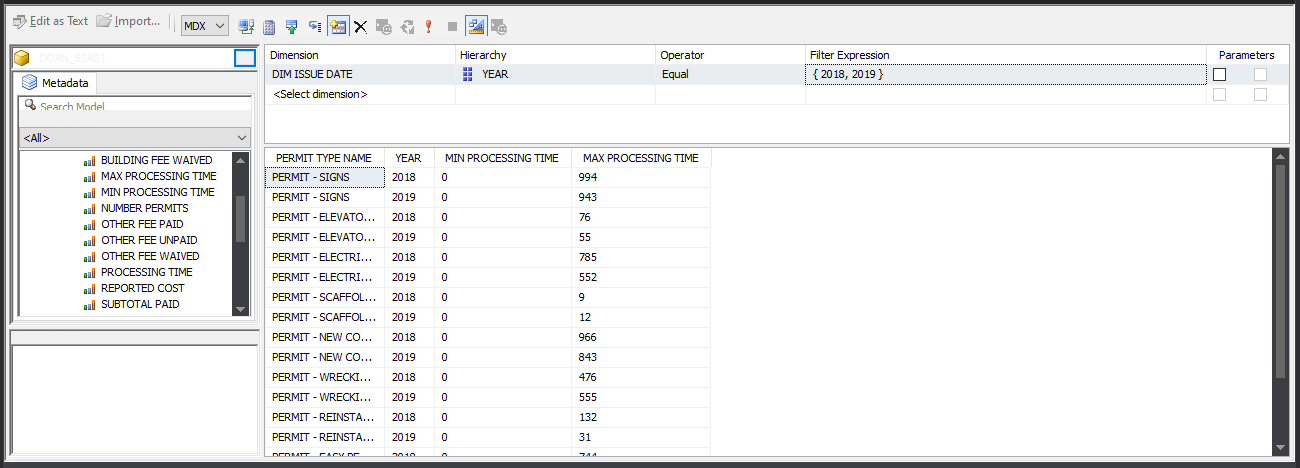
2.3 Truy vấn sử dụng MDX



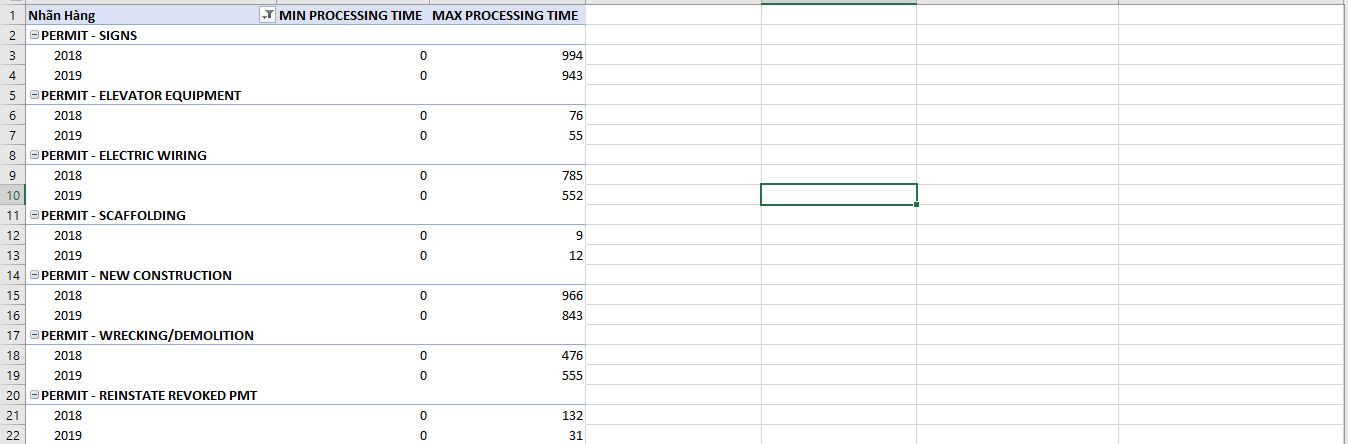
|  |
| --- |
| SELECT [Measures].[NUMBER PERMITS] ON 0  ,NON EMPTY CROSSJOIN([DIM STREET].[STREET NAME 1].[STREET NAME 1],[DIM ISSUE DATE].[YEAR].[YEAR],[DIM ISSUE DATE].[QUARTER].[QUARTER],[DIM ISSUE DATE].[MONTH].[MONTH]) ON 1  FROM [DOAN\_SSAS1]; |

**3. Thống kê số ngày giấy phép được xét duyệt lâu nhất và ngắn nhất của từng loại giấy phép xây theo trong năm 2018 và 2019**

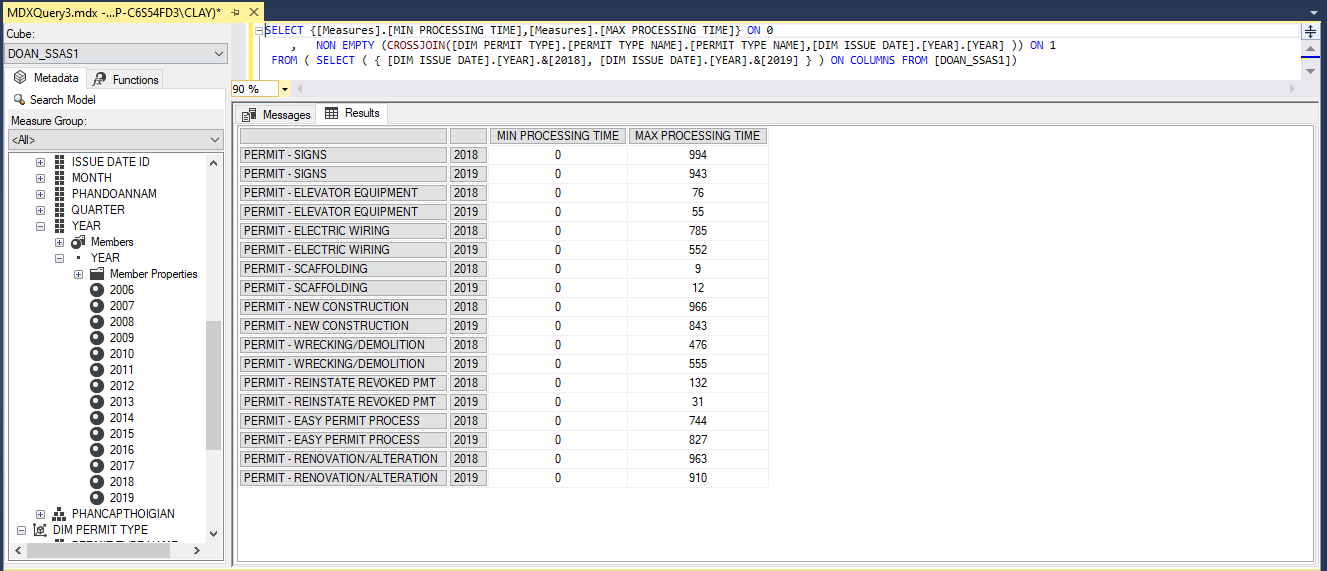
3.1 Truy vấn dữ liệu sử dụng sử dụng BI



3.2 Truy vấn dữ liệu sử dụng Pivot Table



3.3 Truy vấn sử dụng MDX



|  |
| --- |
| SELECT {[Measures].[MIN PROCESSING TIME],[Measures].[MAX PROCESSING TIME]} ON 0  , NON EMPTY (CROSSJOIN([DIM PERMIT TYPE].[PERMIT TYPE NAME].[PERMIT TYPE NAME],[DIM ISSUE DATE].[YEAR].[YEAR] )) ON 1  FROM ( SELECT ( { [DIM ISSUE DATE].[YEAR].&[2018], [DIM ISSUE DATE].[YEAR].&[2019] } ) ON COLUMNS FROM [DOAN\_SSAS1]) |

SELECT {[Measures].[MIN PROCESSING TIME],[Measures].[MAX PROCESSING TIME]} ON 0

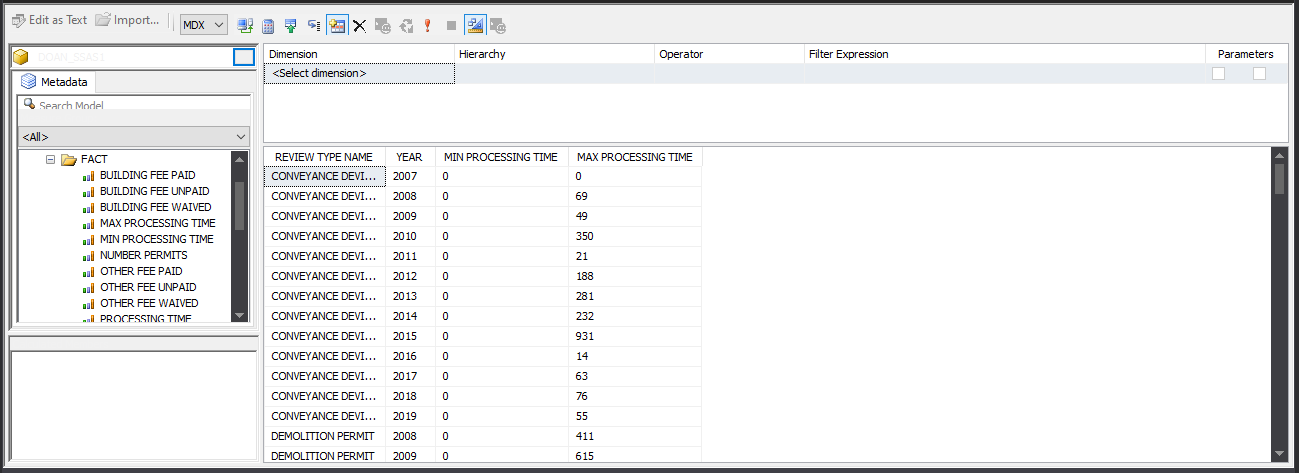
, NON EMPTY (CROSSJOIN([DIM PERMIT TYPE].[PERMIT TYPE NAME].[PERMIT TYPE NAME],[DIM ISSUE DATE].[YEAR].[YEAR] )) ON 1

FROM [DOAN\_SSAS1]

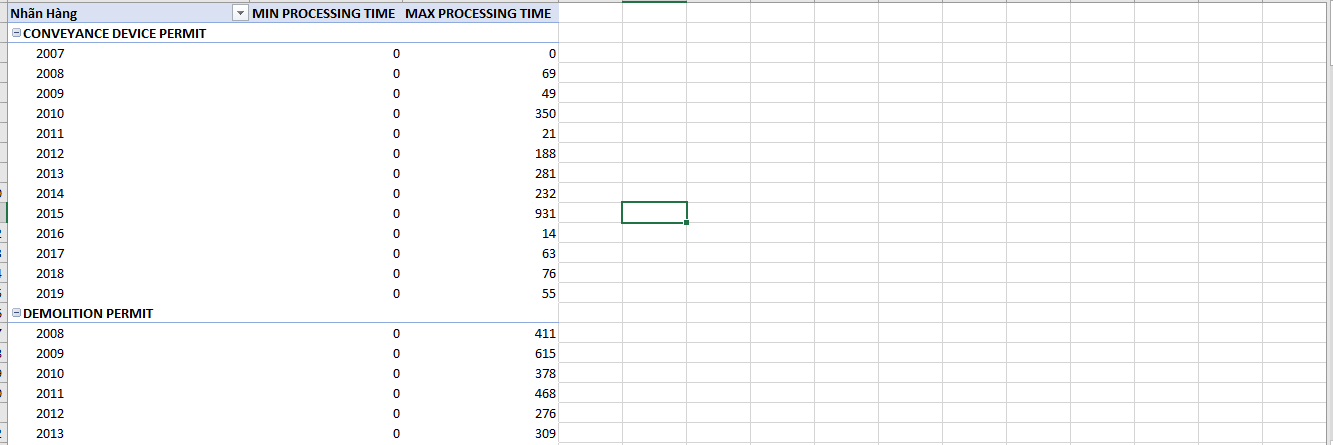
WHERE {[DIM ISSUE DATE].[YEAR].&[2018],[DIM ISSUE DATE].[YEAR].&[2019]};

**4. Thống kê số ngày giấy phép được xét duyệt lâu nhất và nhanh nhất của mỗi loại quy trình xin giấy phép xây dựng theo từng năm**

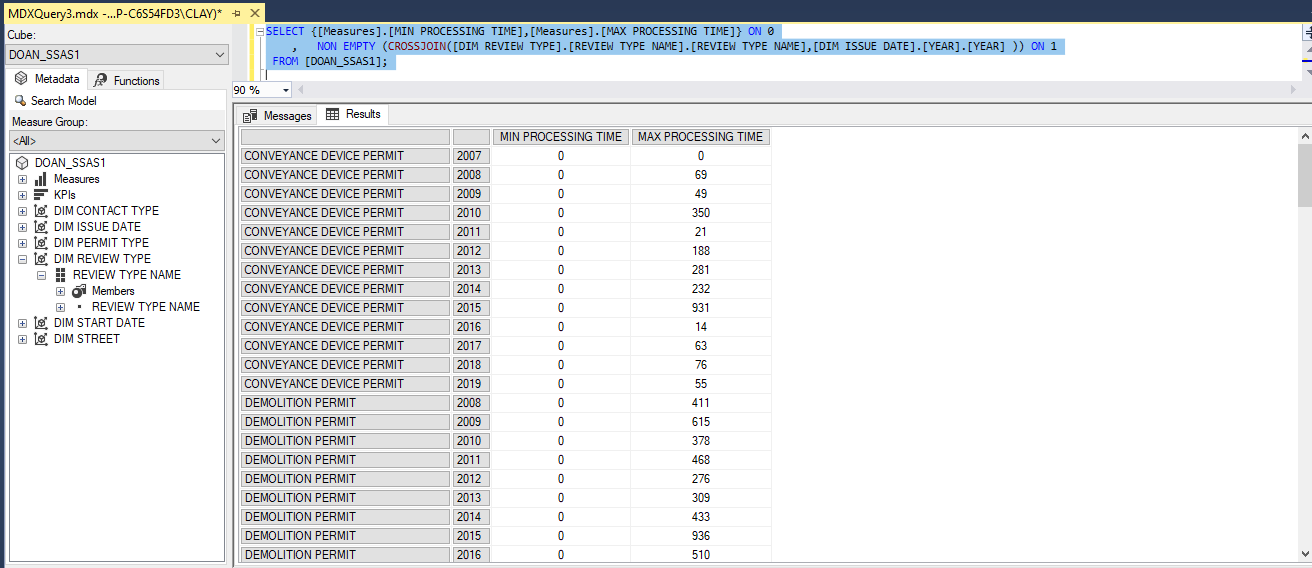
4.1 Truy vấn dữ liệu sử dụng sử dụng BI



4.2 Truy vấn dữ liệu sử dụng Pivot Table



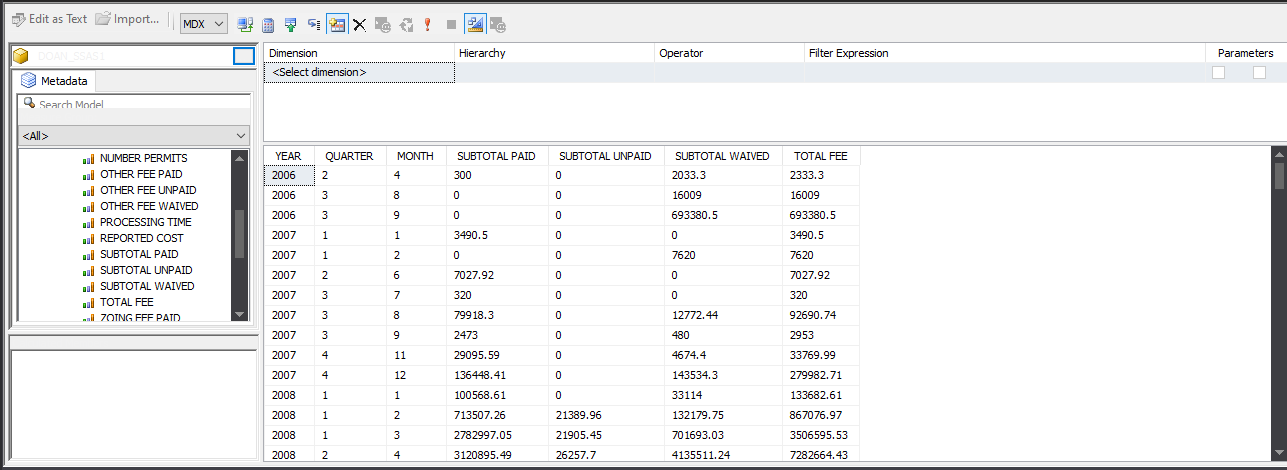
4.3 Truy vấn sử dụng MDX



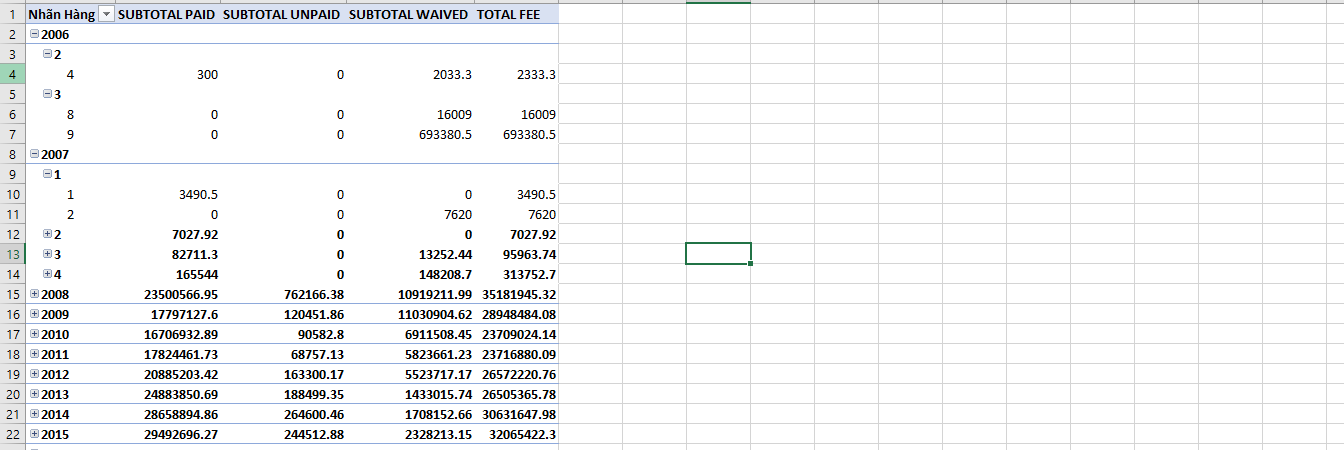
|  |
| --- |
| SELECT {[Measures].[MIN PROCESSING TIME],[Measures].[MAX PROCESSING TIME]} ON 0  , NON EMPTY (CROSSJOIN([DIM REVIEW TYPE].[REVIEW TYPE NAME].[REVIEW TYPE NAME],[DIM ISSUE DATE].[YEAR].[YEAR] )) ON 1  FROM [DOAN\_SSAS1]; |

**5. Thống kê tổng chi phí đã trả, chi phí được miễn và còn nợ tổng phí thanh toán của theo từng năm, quý và tháng**

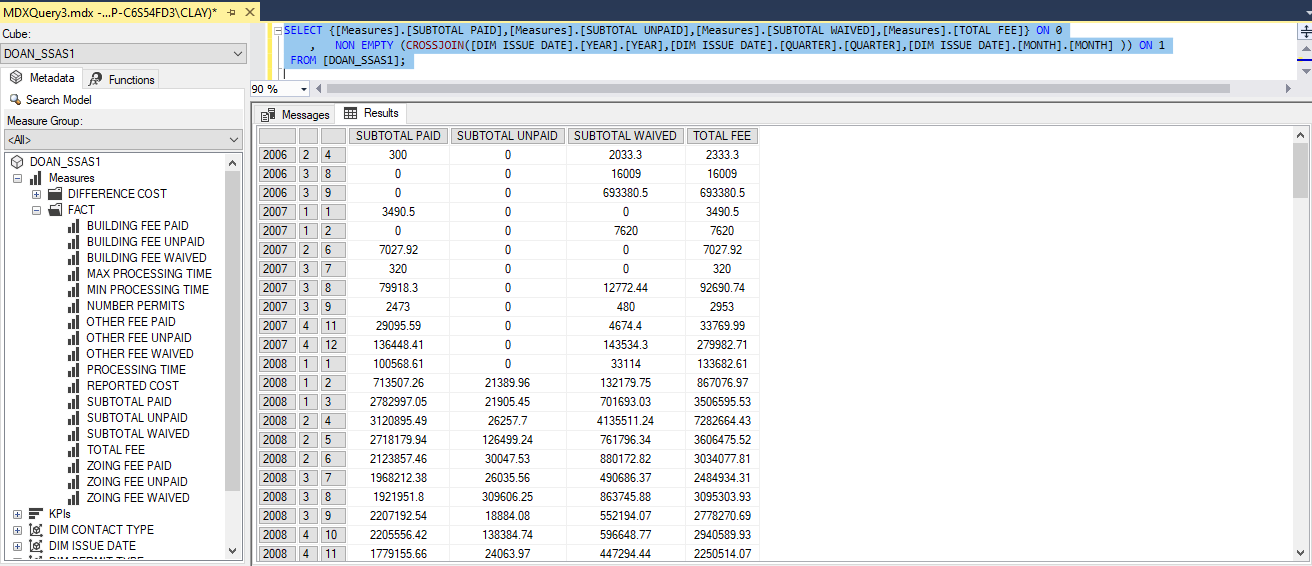
5.1 Truy vấn dữ liệu sử dụng sử dụng BI



5.2 Truy vấn dữ liệu sử dụng Pivot Table



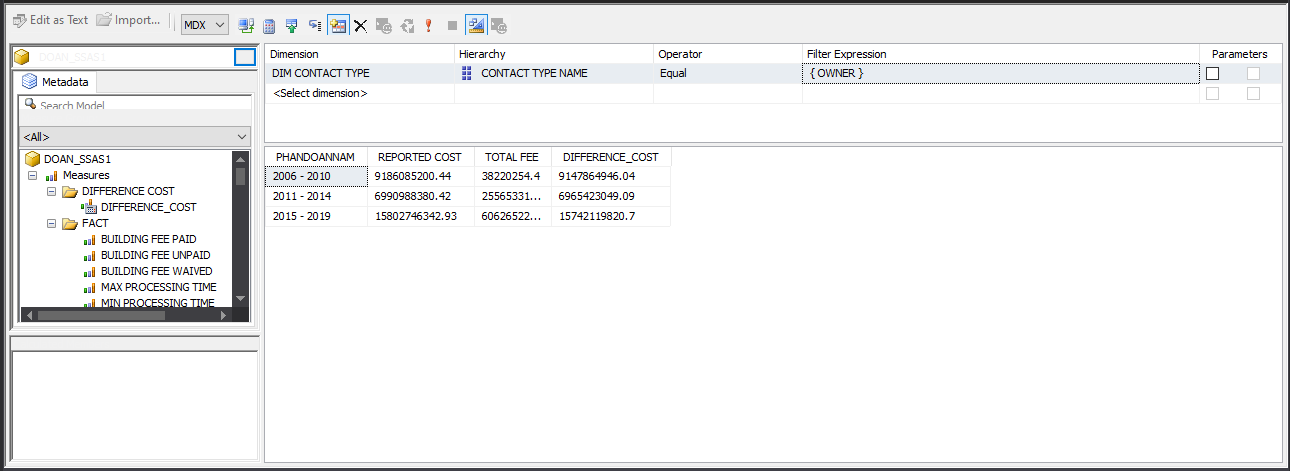
5.3 Truy vấn sử dụng MDX



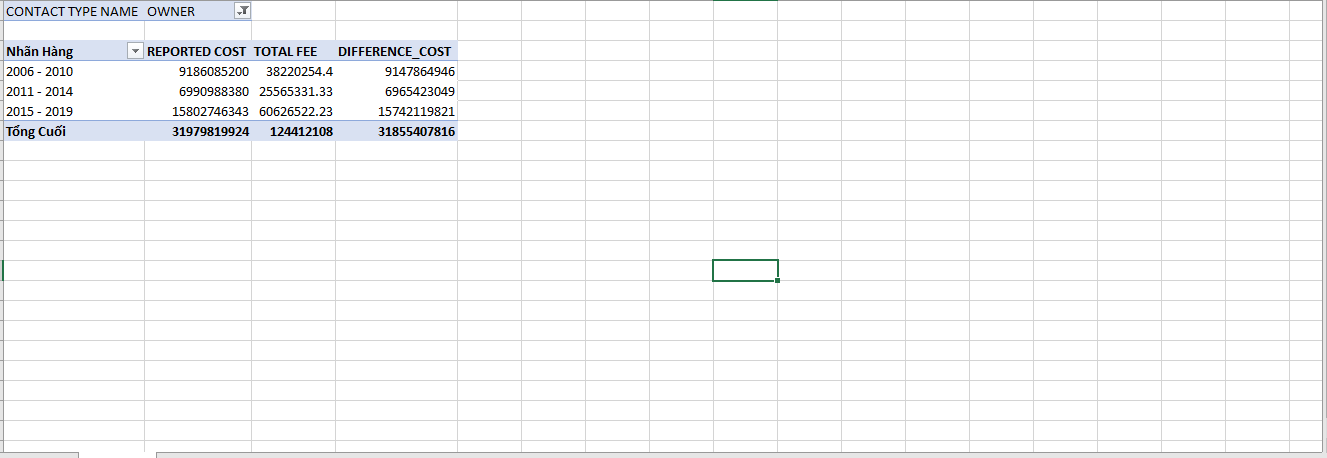
|  |
| --- |
| SELECT {[Measures].[SUBTOTAL PAID],[Measures].[SUBTOTAL UNPAID],[Measures].[SUBTOTAL WAIVED],[Measures].[TOTAL FEE]} ON 0  , NON EMPTY (CROSSJOIN([DIM ISSUE DATE].[YEAR].[YEAR],[DIM ISSUE DATE].[QUARTER].[QUARTER],[DIM ISSUE DATE].[MONTH].[MONTH] )) ON 1  FROM [DOAN\_SSAS1]; |

**6. Thống kê tổng chi phí dự đoán phải trả,chi phí thực tế phải trả và sự chênh lệnh của chi phí dự đoán và thực tế của từng loại đơn vị xin giấy phép là chủ sở hữu công trình “OWNER” theo từng giai đoạn của năm**

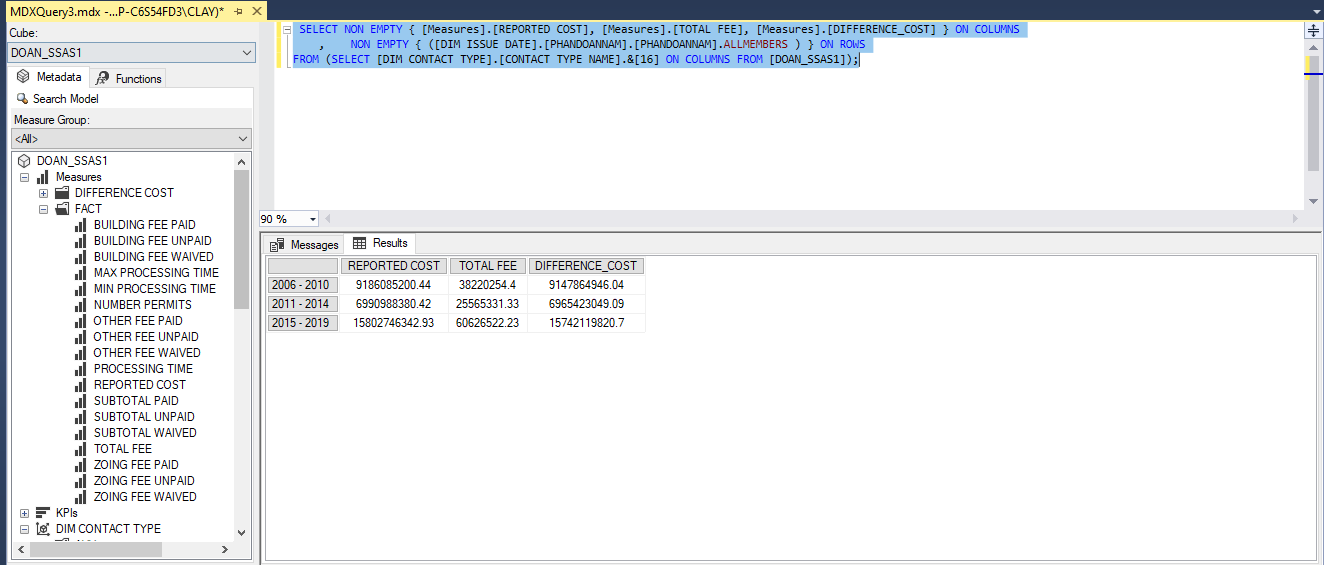
6.1 Truy vấn dữ liệu sử dụng sử dụng BI



6.2 Truy vấn dữ liệu sử dụng Pivot Table



6.3 Truy vấn sử dụng MDX



|  |
| --- |
| SELECT NON EMPTY { [Measures].[REPORTED COST], [Measures].[TOTAL FEE], [Measures].[DIFFERENCE\_COST] } ON COLUMNS  , NON EMPTY { ([DIM ISSUE DATE].[PHANDOANNAM].[PHANDOANNAM].ALLMEMBERS ) } ON ROWS  FROM (SELECT [DIM CONTACT TYPE].[CONTACT TYPE NAME].&[16] ON COLUMNS FROM [DOAN\_SSAS1]); |

**7. Top 5 quận có số lượng giấy phép được xin nhiều nhất**

7.1 Truy vấn dữ liệu sử dụng sử dụng BI

7.2 Truy vấn dữ liệu sử dụng Pivot Table

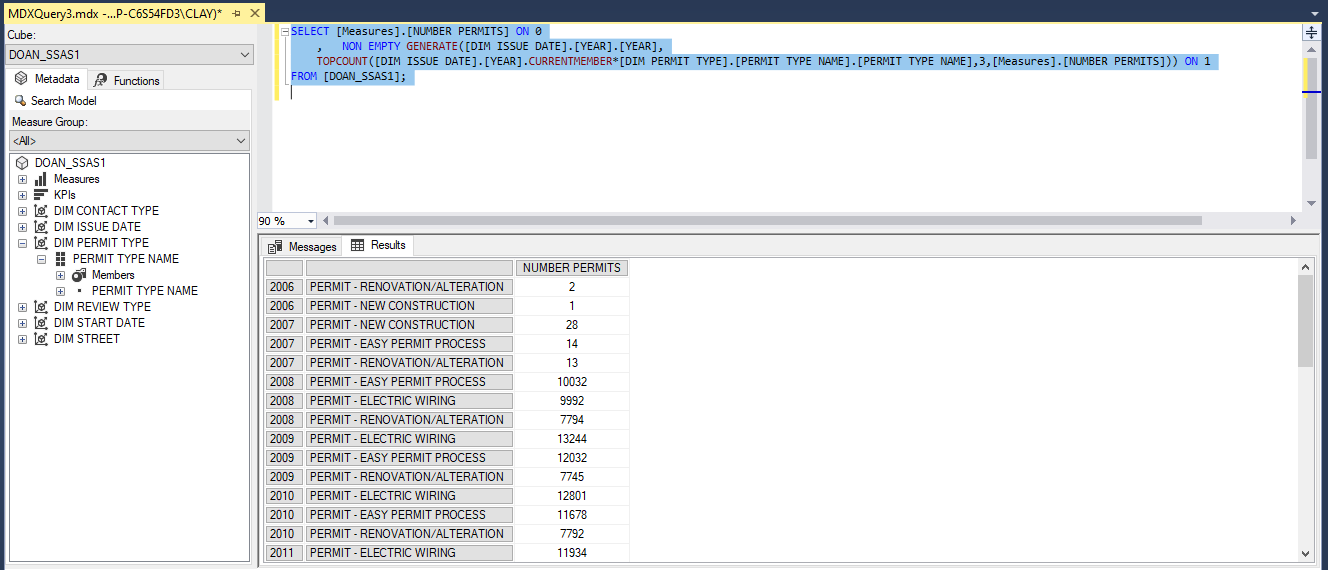
7.3 Truy vấn sử dụng MDX

**8. Trong mỗi năm thống kê top 3 loại giấy phép xây dựng có số lượng giấy phép được duyệt nhiều nhất**

8.1 Truy vấn dữ liệu sử dụng sử dụng BI

8.2 Truy vấn dữ liệu sử dụng Pivot Table

8.3 Truy vấn sử dụng MDX



SELECT [Measures].[NUMBER PERMITS] ON 0

, NON EMPTY GENERATE([DIM ISSUE DATE].[YEAR].[YEAR],

TOPCOUNT([DIM ISSUE DATE].[YEAR].CURRENTMEMBER\*[DIM PERMIT TYPE].[PERMIT TYPE NAME].[PERMIT TYPE NAME],3,[Measures].[NUMBER PERMITS])) ON 1

FROM [DOAN\_SSAS1];

**9**

9.1 Truy vấn dữ liệu sử dụng sử dụng BI

9.2 Truy vấn dữ liệu sử dụng Pivot Table

9.3 Truy vấn sử dụng MDX

**10.**

10.1 Truy vấn dữ liệu sử dụng sử dụng BI

10.2 Truy vấn dữ liệu sử dụng Pivot Table

10.3 Truy vấn sử dụng MDX

1.

CHƯƠNG 5: BÁO CÁO (SSRS)

-- TOP 3

SELECT NON EMPTY { [Measures].[NUMBERS PERMIT] } ON COLUMNS,

TOPCOUNT( [DIM PERMIT TYPE].[PERMIT TYPE ID].[PERMIT TYPE ID].ALLMEMBERS

,3

, [Measures].[NUMBERS PERMIT]

) ON ROWS

FROM [LAPTOP-C6S54FD3orc];

SELECT NON EMPTY { [Measures].[NUMBERS PERMIT] } ON COLUMNS,

NON EMPTY { ([DIM PERMIT TYPE].[PERMIT TYPE ID].[PERMIT TYPE ID].ALLMEMBERS ) } ON ROWS

FROM [LAPTOP-C6S54FD3orc];

SELECT NON EMPTY {[Measures].[NUMBERS PERMIT]} ON COLUMNS

, NON EMPTY {[DIM PERMIT TYPE].[PERMIT TYPE ID].[PERMIT TYPE ID].ALLMEMBERS} ON ROWS

FROM [LAPTOP-C6S54FD3orc];

SELECT NON EMPTY {[Measures].[TOTAL FEE]} ON COLUMNS

, NON EMPTY {([DIM STREET].[STREET DIRECTION].[STREET DIRECTION].ALLMEMBERS\*[DIM STREET].[WARD NAME].[WARD NAME].ALLMEMBERS)} ON ROWS

FROM [LAPTOP-C6S54FD3orc];

SELECT NON EMPTY { [Measures].[MIN PROCESSING TIME], [Measures].[MAX PROCESSING TIME] } ON COLUMNS,

NON EMPTY { ([DIM PERMIT TYPE].[PERMIT TYPE ID].[PERMIT TYPE ID].ALLMEMBERS \* [DIM ISSUE DATE].[YEAR].[YEAR].ALLMEMBERS ) } ON ROWS

FROM [LAPTOP-C6S54FD3orc];

SELECT NON