khung1khung3khung4**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP.HCM**

**KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO CUỐI KÌ**

**môn học: KHO DỮ LIỆU**

**Đề tài: THIẾT KẾ DATA WAREHOUSE CHO DOANH NGHIỆP BÁN LẺ XE HƠI**

**Giảng viên hướng dẫn:** ThS. Quách Đình Hoàng

**Nhóm thực hiện: Nhóm 2 MSSV**

Trần Trọng Nghĩa 17133040

Nguyễn Quang Nhật 17133043

Đỗ Viết Hải 17133017

Vũ Anh Thái Dương 17133010

Nguyễn Quý Trọng Hiếu 17133019

khung2**TP. HCM, tháng 6 năm 2020**

**MỤC LỤC**

[**MỞ ĐẦU** 4](#_Toc44843628)

[**Chương 1: MÔ TẢ VÀ XÂY DỰNG DATA WAREHOUSE** 6](#_Toc44843629)

[**1.1. Dữ liệu** 6](#_Toc44843631)

[**1.1.1. Giới thiệu về cơ sở dữ liệu** 6](#_Toc44843632)

[**1.1.2. Lý do cần phải có một data warehouse** 7](#_Toc44843633)

[**1.2. Business process** 7](#_Toc44843634)

[**1.2.1. Business process mà nhóm mô hình** 7](#_Toc44843635)

[**1.2.2. Business questions** 8](#_Toc44843636)

[**1.2.3. Granularity** 8](#_Toc44843637)

[**1.2.4. Mô tả mỗi dòng trong fact table và các function áp dụng trên fact** 8](#_Toc44843638)

[**1.2.5. Mô tả các dimension** 9](#_Toc44843639)

[**1.2.6. Measures** 9](#_Toc44843640)

[**1.3. Thiết kế** 10](#_Toc44843641)

[**1.3.1. Vẽ schema cho các business process** 10](#_Toc44843642)

[**Chương 2: ETL CHO SALES REPORT BUSINESS PROCESS** 12](#_Toc44843643)

[**2.1. Tạo ETL project và solution, thiết lập các Connection** 12](#_Toc44843645)

[**2.2. Import Date Dimension** 12](#_Toc44843646)

[**2.2.1. Tạo Cơ sở dữ liệu ExternalSources\_P** 12](#_Toc44843647)

[**2.2.2. Load dữ liệu vào bảng stgDate của CSDL buycarStage** 13](#_Toc44843648)

[**2.2.3. Ánh xạ các cột từ bảng stgDate sang bảng DimDate** 13](#_Toc44843649)

[**2.2.4. Hoàn thành quá trình Import Date Dimension** 14](#_Toc44843650)

[**2.3.** **Staging Sales Report Data Mart** 15](#_Toc44843651)

[**2.3.1. Staging Products** 15](#_Toc44843652)

[**2.3.2. Staging Offices** 16](#_Toc44843653)

[**2.3.3. Staging Customers** 17](#_Toc44843654)

[**2.3.4. Staging Employees** 18](#_Toc44843655)

[**2.3.5. Staging Sales** 19](#_Toc44843656)

[**2.3.6. Hoàn tất việc Staging cho BuycarDW** 21](#_Toc44843657)

[**2.4. Load dữ liệu từ BuycarStage vào BuycarDW** 21](#_Toc44843658)

[**2.4.1. Data Flow Stage to DimProduct** 21](#_Toc44843659)

[**2.4.2. Data Flow Stage to DimOffices** 23](#_Toc44843660)

[**2.4.3. Data Flow Stage to DimCustomers** 24](#_Toc44843661)

[**2.4.4. Data Flow Stage to DimEmployees** 25](#_Toc44843662)

[**2.4.5. Data Flow Stage to Fact\_Sales\_Report** 27](#_Toc44843663)

[**2.4.6. Hoàn thành việc load dữ liệu vào BuycarDW** 28](#_Toc44843664)

[**Chương 3: THIẾT KẾ DATA CUBE CarSalesOLAP** 29](#_Toc44843665)

[**3.1. Tạo Data Source View** 29](#_Toc44843667)

[**3.2. Tạo Cubes** 29](#_Toc44843668)

[**3.3. Tạo các Dimensions và các hierarchy** 30](#_Toc44843669)

[**3.4. Xem Data Cube trong Excel với Pivot Tables** 32](#_Toc44843670)

[**3.5. Excel Power Pivot** 33](#_Toc44843671)

[**Chương 4: MỘT SỐ TRUY VẤN VÀ BÁO CÁO** 34](#_Toc44843672)

[**4.1. Truy vấn thống kê số lượng sản phẩm đặt hàng theo từng khu vực qua các năm** 34](#_Toc44843674)

[**4.2. Truy vấn tỷ lệ doanh thu của các chi nhánh qua các năm** 36](#_Toc44843675)

[**4.3. Truy vấn thống kê sản phẩm được ưa chuộng nhất năm 2003** 38](#_Toc44843676)

[**4.4. Truy vấn doanh thu và tỷ lệ doanh thu từng quý năm 2004** 40](#_Toc44843677)

[**KẾT LUẬN** 42](#_Toc44843678)

[**1. Những việc nhóm đã làm được** 42](#_Toc44843679)

[**2. Những việc nhóm chưa làm được và dự định thực hiện trong tương lai** 42](#_Toc44843680)

[**TÀI LIỆU THAM KHẢO** 43](#_Toc44843681)

# **MỞ ĐẦU**

**1. Tính quan trọng của đề tài**

Ngày nay đối với các doanh nghiệp kinh doanh dù là lớn hay nhỏ thì việc phát triển và gia tăng doanh số là một điều vô cùng thiết yếu. Để làm được những điều này thì chúng ta cần có một báo cáo bán hàng. Báo cáo bán hàng là nơi cung cấp cái nhìn sâu sắc về quy trình và nền tảng của các hoạt động, chiến lược của doanh nghiệp. Đó là quá trình theo dõi thông tin ở mọi bước của quy trình bán hàng và phân tích dữ liệu, để dự đoán sự phát triển của doanh nghiệp nằm ở đâu và xác định các yếu tố ảnh hưởng đến sự phát triển. Báo cáo bán hàng cũng hỗ trợ đưa ra quyết định về số lượng sản xuất, thời gian, nhu cầu việc làm, giá cả và các yếu tố cần thiết khác giúp doanh nghiệp phát triển. Nhận thấy được sự quan trọng của báo cáo bán hàng nên nhóm chúng em đã quyết định lựa chọn cơ sở dữ liệu của doanh nghiệp bán lẻ xe hơi (BuyCar) để thiết kế, xây dựng một kho dữ liệu.

**2. Nhiệm vụ nghiên cứu**

* Thiết kế data warehouse dựa trên tập dữ liệu BuyCarDW
* Xây dựng các business process và các bus matrix tương ứng
* Xây dựng các bảng fact, dimension, các cây phân cấp hierachy thích hợp
* ETL dữ liệu vào data warehouse bằng tool SSIS
* Thiết kế các data cube ứng với bảng fact bằng tool SSAS
* Viết các truy vấn, report cần thiết cho việc buôn bán, thống kê

**3. Kết cấu đề tài**

Ngoài phần mở đầu, kết luận và danh mục tài liệu tham khảo ra báo cáo còn gồm 4 chương cụ thể:

- Chương 1: Mô tả và xây dựng Data Warehouse

- Chương 2: ETL cho Sales Report Business Process

- Chương 3: Thiết kế data cube CarSalesOLAP

- Chương 4: Một số truy vấn và báo cáo

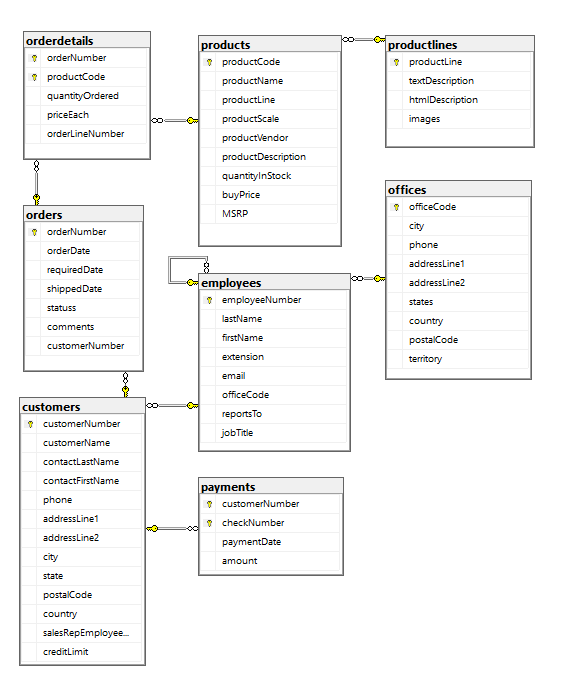
# **Chương 1**

# **MÔ TẢ VÀ XÂY DỰNG DATA WAREHOUSE**

## **1.1. Dữ liệu**

### **1.1.1. Giới thiệu về cơ sở dữ liệu**

Cơ sở dữ liệu được sử dụng là hệ thống bán lẻ mô hình xe hơi. Nó chứa dữ liệu kinh doanh điển hình như khách hàng, sản phẩm, đơn đặt hàng, chi tiết đơn hàng, nhân viên, thanh toán, loại sản phẩm.



### **1.1.2. Lý do cần phải có một data warehouse**

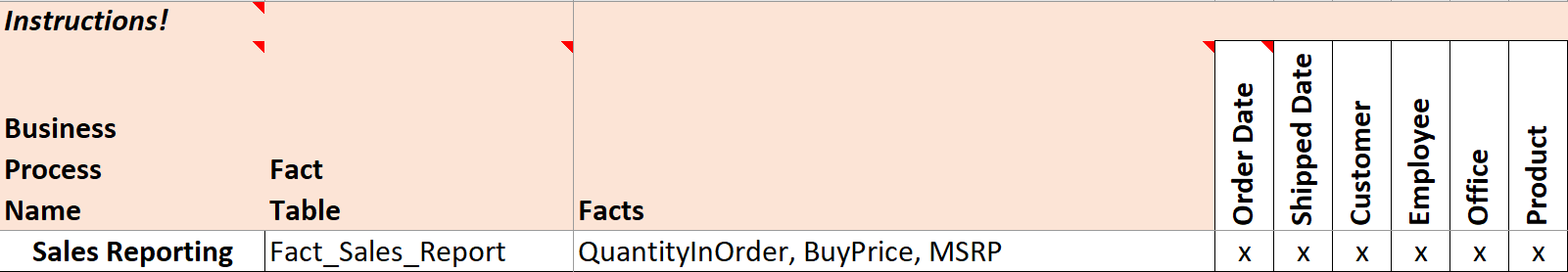
Data warehouse giúp công ty (hay tổ chức) tối ưu business và hỗ trợ vấn đề ra quyết định. Data warehouse thực hiện các truy vấn, report phức tạp để đưa ra các quyết định như:

* Thống kê tổng doanh thu, tỷ lệ doanh thu, phân hoạch theo năm của các chi nhánh (từ đó quyết định có nên thêm/giảm số lượng hàng hóa và nhân viên).
* Thống kê biết được số sản phẩm đã bán được của các nhân viên (từ đó đưa ra quyết định khen thưởng nếu bán được nhiều).
* Thống kê xem tình hình bán hàng ở các khu vực (từ đó đưa ra quyết định có nên mở thêm/ đóng chi nhánh nếu bán nhiều/ít).
* Thống kê biết được số lượng của từng sản phẩm bán ra (từ đó đưa ra quyết định có nên nhập thêm/ giảm số lượng sản phẩm đó).
* Thống kê biết được khách hàng nào đã mua nhiều sản phẩm nhất (từ đó quyết định giảm giá tặng thưởng tri ân khách hàng).

## **1.2. Business process**

### **1.2.1. Business process mà nhóm mô hình**

Do nhu cầu muốn thống kê doanh thu, lợi nhuận bán hàng của các chi nhánh (Office) và theo dõi khách hàng (customer), nhân viên (employee), sản phẩm (product) để đưa ra những chính sách kinh doanh tối ưu, hiệu quả nhằm thúc đẩy chi nhánh gia tăng lợi nhuận, phát triển.



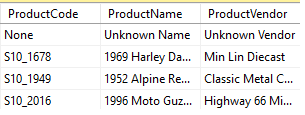
### **1.2.2. Business questions**

* Sản phẩm nào là bán chạy nhất.
* Khách hàng nào đã mua nhiều sản phẩm nhất.
* Chi nhánh bạn chạy nhất sẽ bán chạy nhất những sản phẩm nào.
* Tổng doanh thu và lợi nhuận từ năm 2003 đến 2005 tại các chi nhánh.
* Xếp hạng các chi nhánh theo tổng doanh thu, tỷ lệ doanh thu của từng chi nhánh, phân hoạch theo năm.

### **1.2.3. Granularity**

Nhóm chọn mức độ chi tiết thấp nhất (càng mịn thì kích cỡ fact càng lớn., từ đó dùng các sum avg funtion từ đó tính được ở mức độ chi tiết cao hơn).

Ví dụ: Sản phẩm thì lưu trữ ở mức Mã Code.



### **1.2.4. Mô tả mỗi dòng trong fact table và các function áp dụng trên fact**

* **Mô tả mỗi dòng trong fact table**
* ID đơn đặt hàng.
* Mã sản phẩm.
* Mã khách hàng.
* Mã nhân viên.
* Mã chi nhánh.
* Ngày đặt.
* Ngày yêu cầu giao.
* Ngày giao hàng.
* Số lượng sản phẩm của đơn hàng đó.
* Giá nhập vào.
* Giá bán ra.
* **Các function áp dụng trên fact**
* Sum
* Analytic functions
* Count
* Min
* Max

### **1.2.5. Mô tả các dimension**

* Dim product: Mã sản phẩm, tên sản phẩm, nhà cung cấp, mô tả về sản phẩm.
* Dim customer: Mã khách hàng, tên khách hàng, khu vực, thành phố, địa chỉ, mã bưu điện.
* Dim offices: Mã chi nhánh, đất nước, khu vực, thành phố, địa chỉ.
* Dim employee: Mã nhân viên, địa chỉ email, mã chi nhánh, chức vụ.
* Dim date: Thứ, ngày, tháng, quý, năm.

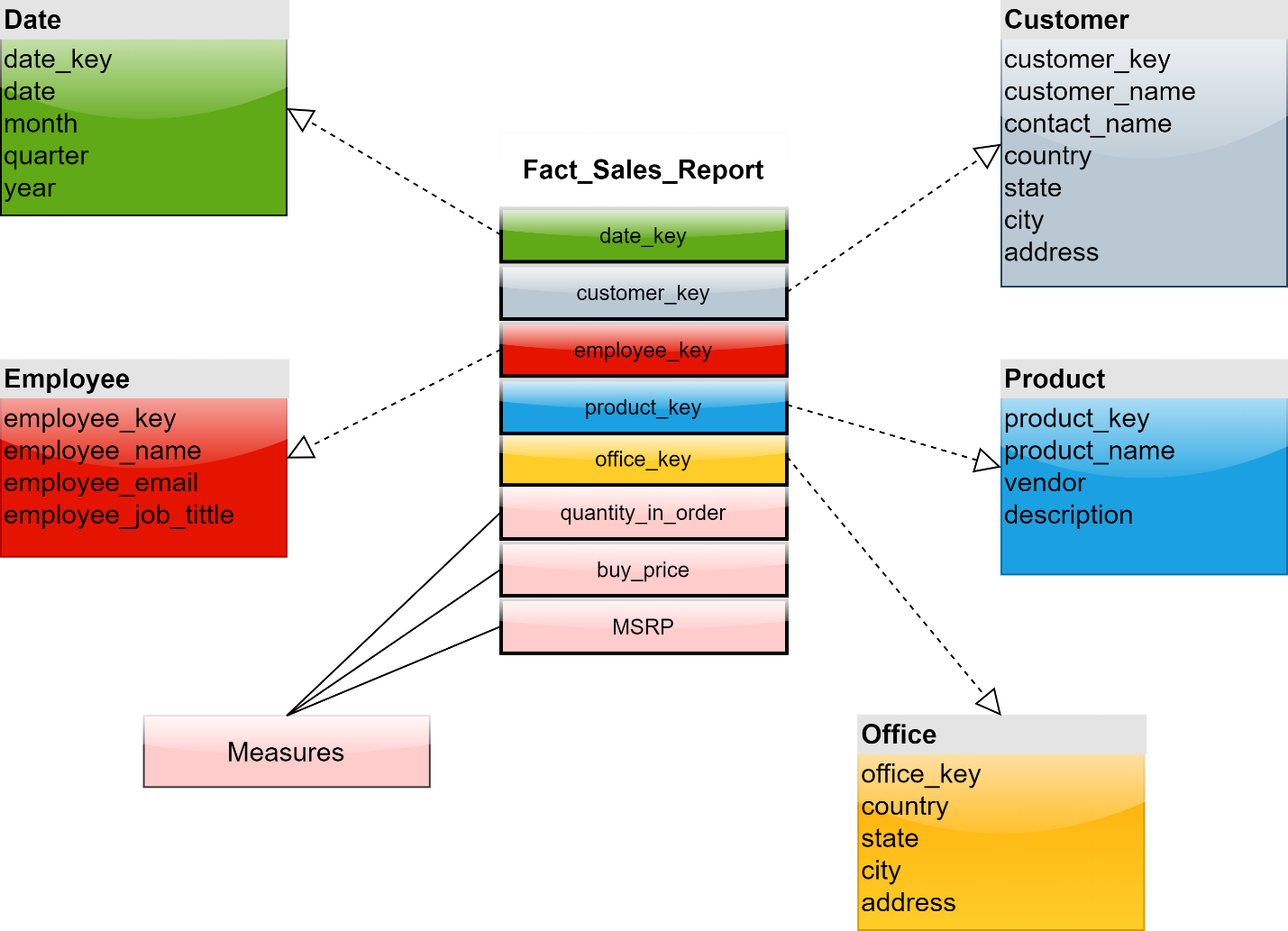
### **1.2.6. Measures**

* **Mô tả các measure**
* QuantityInOrder: Chứa thông tin về số lượng sản phẩm trong một đơn đặt hàng.
* BuyPrice: Chứa thông tin về giá nhập sản phẩm về.
* MSRP: Chứa thông tin về giá mà sản phẩm bán ra.
* **Phân loại các measure**
* QuantityInOrder: Additive (sum theo Product, Date, Office, Customer, Employee).
* BuyPrice: Non-additive.
* MSRP: Non-additive.

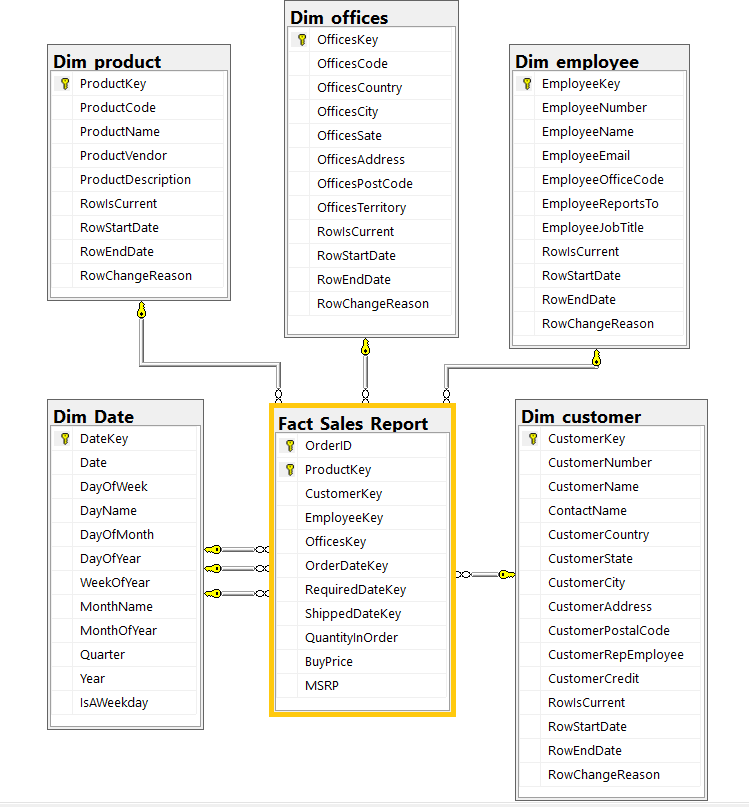
## **1.3. Thiết kế**

### **1.3.1. Vẽ schema cho các business process**

* **Thiết kế mô hình ở mức ý niệm**

****

* **Thiết kế mô hình ở mức logic**

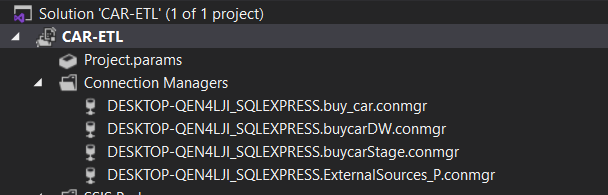
****

# **Chương 2**

# **ETL CHO SALES REPORT BUSINESS PROCESS**

## **2.1. Tạo ETL project và solution, thiết lập các Connection**

* Tạo CAR-ETL để đưa dữ liệu vào các bảng dữ liệu tương ứng với Sales Report business process.
* Để có thể truy xuất dữ liệu cho quá trình ETL, ta cần thiết lập các connection tới nguồn (source) và đích (target). Ta có thể tạo các connection này trong mỗi package.

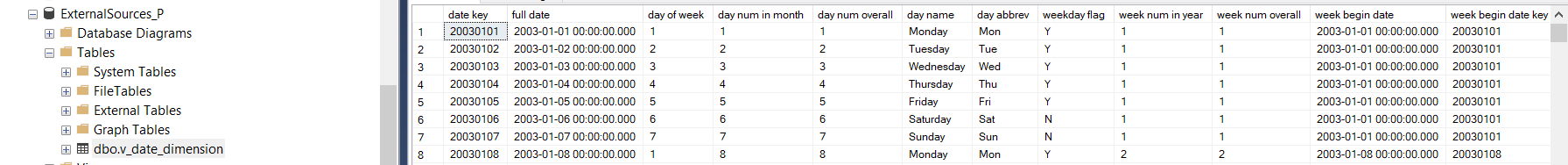


## **2.2. Import Date Dimension**

Do dữ liệu về thời gian là rất cần thiết cho Sales Report business process để phân tích và đánh giá chi tiết.

**2.2.1. Tạo Cơ sở dữ liệu ExternalSources\_P**

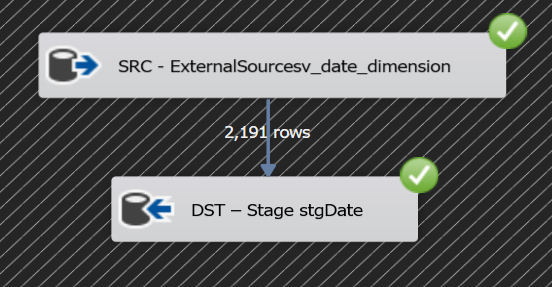
Cơ sở dữ liệu ExternalSources\_P sẽ load dữ liệu về Datetime từ file Excel (nơi chưa các thông tin về chi tiết về thời gian) vào SQL Server. Sau khi đã load xong thì ta được cơ sở dữ liệu như sau.



**2.2.2. Load dữ liệu vào bảng stgDate của CSDL buycarStage**

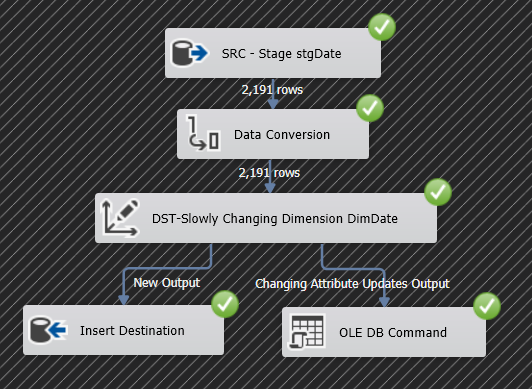
CSDL buycarStage được tạo ra với mục đích để tăng hiệu quả của các quy trình ETL, đảm bảo tính toàn vẹn dữ liệu và làm sạch dữ liệu (nhận dạng và loại bỏ dữ liệu không hợp lệ khỏi CSDL).

Tạo DateDimensionImport.dtsx package để chép dữ liệu từ bảng v\_date\_dimension của CSDL ExternalSources\_P sang bảng stgDate CSDL buycarStage. Sử dụng SSIS để tạo destination table, truncate table trước khi staging dữ liệu.

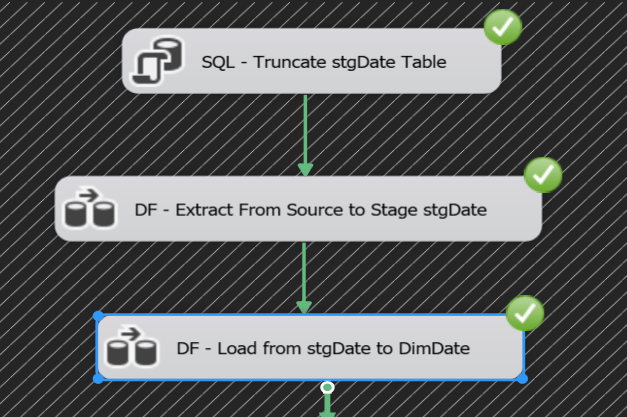


**2.2.3. Ánh xạ các cột từ bảng stgDate sang bảng DimDate**

* Load dữ liệu từ stgDate.
* Thực hiện Data Conversion (chuyển đổi các kiểu dữ liệu cho phù hợp để mapping).
* Dùng SCD (Slowly Changing Dimension) Type 1(Changing). SCD trong trường hợp này được sử dụng với mục đích cập nhật vào trong bảng DimDate.



**2.2.4. Hoàn thành quá trình Import Date Dimension**

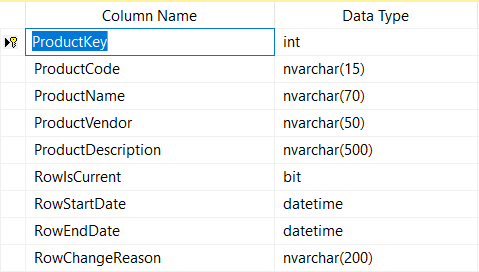


**2.3.** **Staging Sales Report Data Mart**

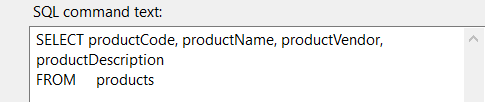
Vào SQL Server tạo một CSDL trống với tên là buycarStage. Sau đó, tạo Stage\_Fact\_Sales\_Report.dtsx package để load dữ liệu từ CSDL Buy\_car (CSDL gốc) sang BuycarStage (CSDL lưu trữ tạm thời).

**2.3.1. Staging Products**

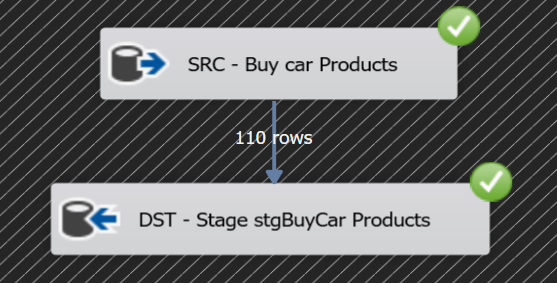
Bảng Dim\_product của buycarDW như sau:



Do đó ta cần các thuộc tính ProductCode, ProductName, ProductVendor, ProductDescription cho bảng Dim\_product. Ta thực hiện truy vấn sau để lấy được các trường này.

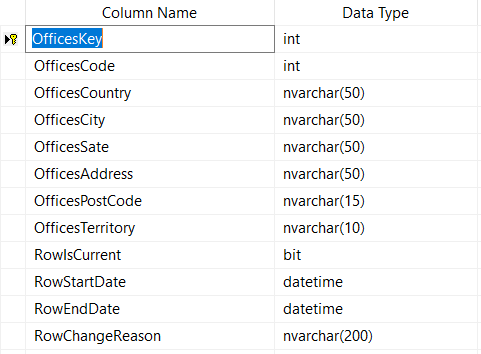


Sau khi thực hiện việc Staging cho bảng Products ta được kết quả như hình bên dưới.

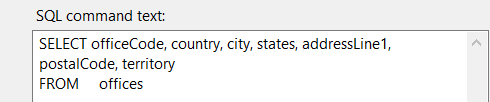


**2.3.2. Staging Offices**

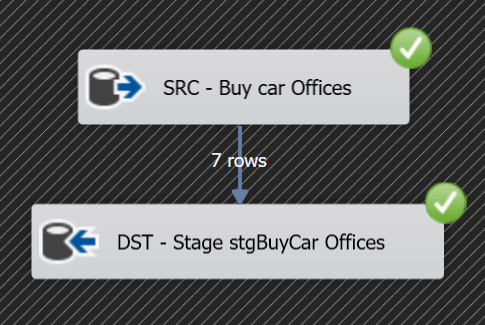
Bảng Dim\_offices của buycarDW như sau:



Do đó ta cần các thuộc tính OfficesCode, OfficesCountry, OfficesCity, OfficesSate, OfficesAddress, OfficesPostCode, OfficesTerritory cho bảng Dim\_offices. Ta thực hiện truy vấn sau để lấy được các trường này.

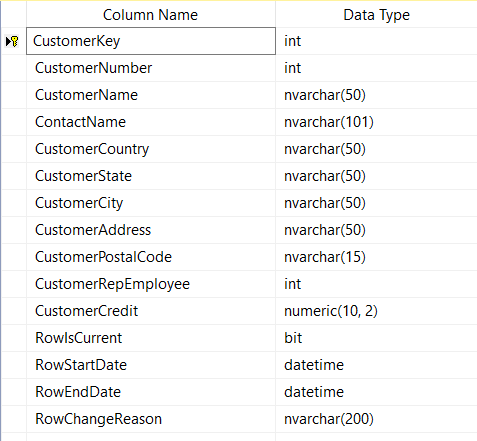


Sau khi thực hiện việc Staging cho bảng Offices ta được kết quả như hình bên dưới.

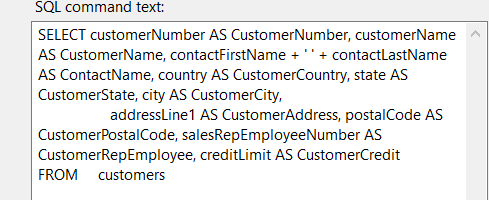


**2.3.3. Staging Customers**

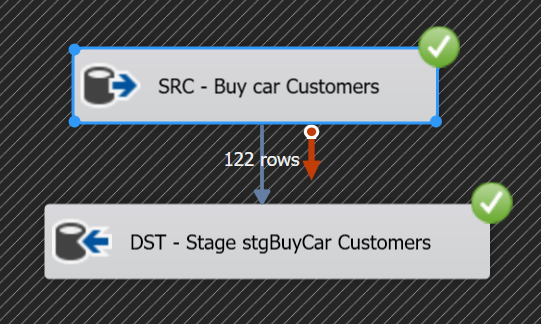
Bảng Dim\_customers của buycarDW như sau:



Do đó ta cần các thuộc tính CustomerNumber, CustomerName, ContactName, CustomerCountry, CustomerState, CustomerCity, CustomerAddress, CustomerPostalCode, CustomerRepEmployee, CustomerCredit cho bảng Dim\_customers. Ta thực hiện truy vấn sau để lấy được các trường này.

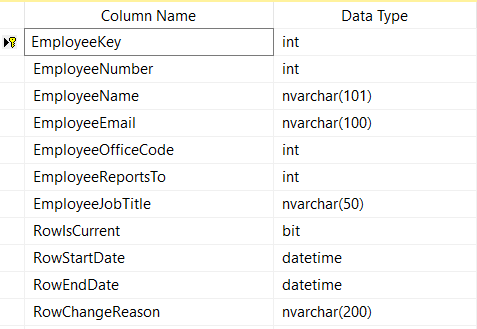


Sau khi thực hiện việc Staging cho bảng Customers ta được kết quả như hình bên dưới.

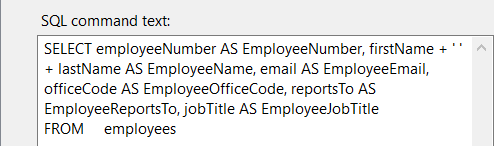


**2.3.4. Staging Employees**

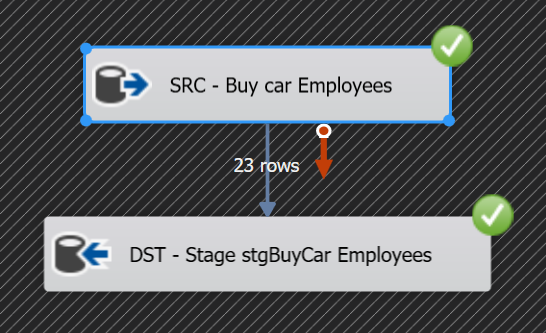
Bảng Dim\_employees của buycarDW như sau:



Do đó ta cần các thuộc tính EmployeeNumber, EmployeeName, EmployeeEmail, EmployeeOfficeCode, EmployeeReportsTo, EmployeeJobTitle cho bảng Dim\_employees. Ta thực hiện truy vấn sau để lấy được các trường này.

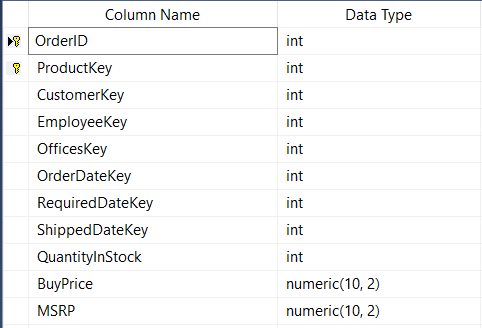


Sau khi thực hiện việc Staging cho bảng Employees ta được kết quả như hình bên dưới.

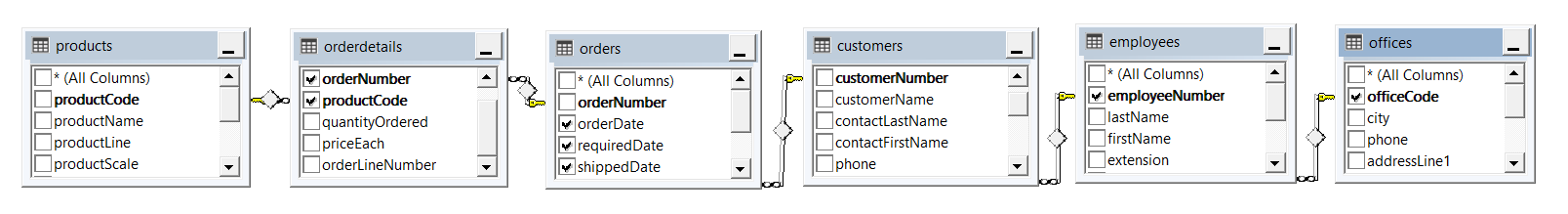


**2.3.5. Staging Sales**

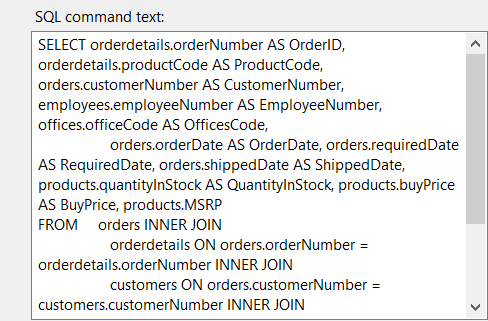
Bảng Fact\_Sales\_Report của buycarDW như sau:



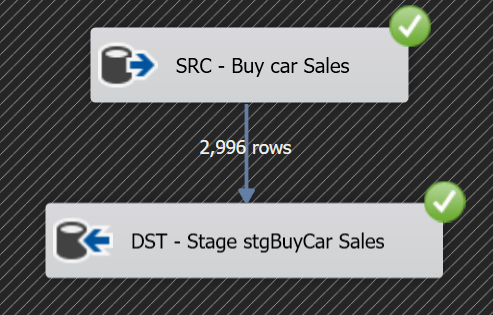
Do đó ta cần các thuộc tính OrderID, ProductCode, CustomerNumber, EmployeeNumber, OfficesCode, OrderDate, RequiredDate, ShippedDate, QuantityInStock, BuyPrice, MSRP cho bảng Fact\_Sales\_Report. Các thuộc tính này không thuộc cùng một bảng nên ta thực hiện việc kết các bảng.



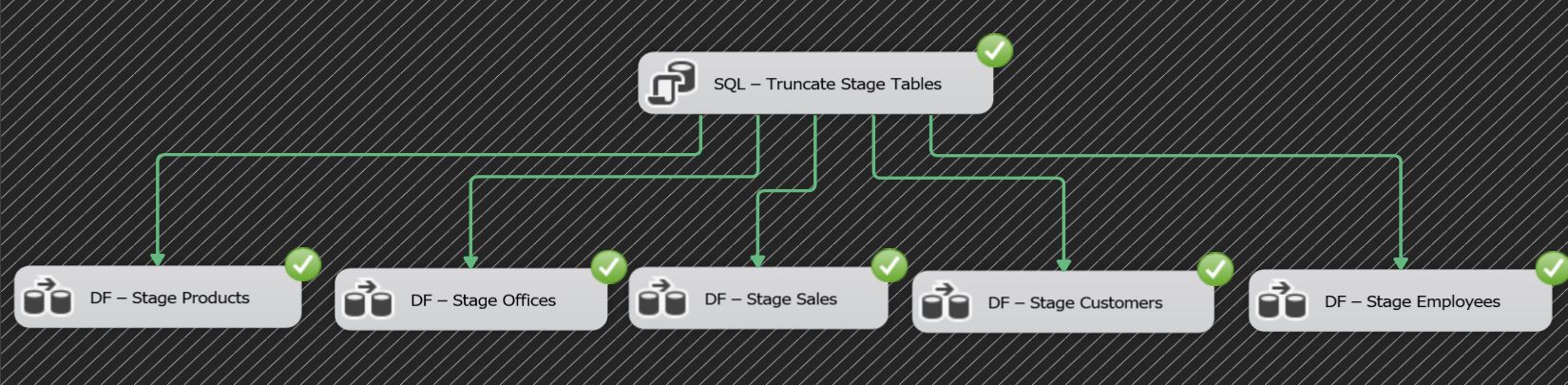
Sau đó, ta thực hiện truy vấn sau để lấy được các trường này.



Sau khi thực hiện việc Staging cho bảng Fact\_Sales\_Report ta được kết quả như hình bên dưới.



**2.3.6. Hoàn tất việc Staging cho BuycarDW**

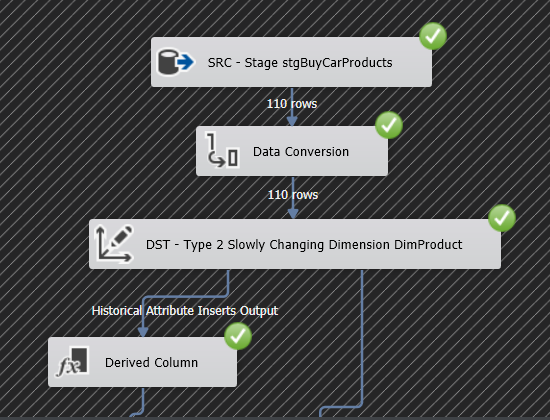


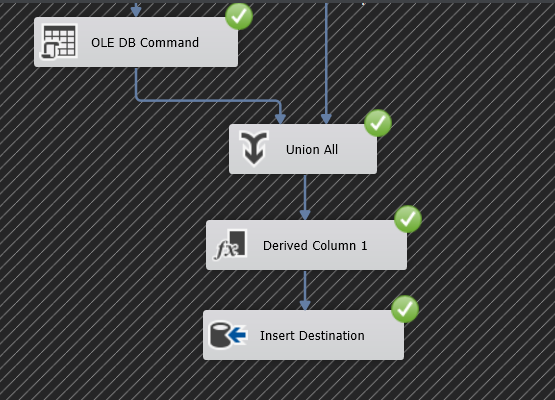
**2.4. Load dữ liệu từ BuycarStage vào BuycarDW**

Sau khi tiến hành tạo và load dữ liệu vào CSDL BuycarStage thành công. Ta tiến hành load các dữ liệu đó vào kho dữ liệu BuycarDW. Tạo DW\_Sales.dtsx package để load dữ liệu từ CSDL BuycarStage sang CSDL BuycarDW.

**2.4.1. Data Flow Stage to DimProduct**

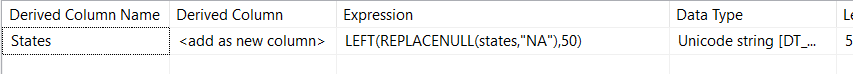
* Load dữ liệu từ stgBuyCarProducts.
* Thực hiện Data Conversion (chuyển đổi các kiểu dữ liệu cho phù hợp để mapping).
* Dùng SCD (Slowly Changing Dimension) Type 2 (Historical). SCD trong trường hợp này được sử dụng với mục đích thêm dòng mới, đánh dấu dòng cũ không hoạt động trong bảng DimProduct.



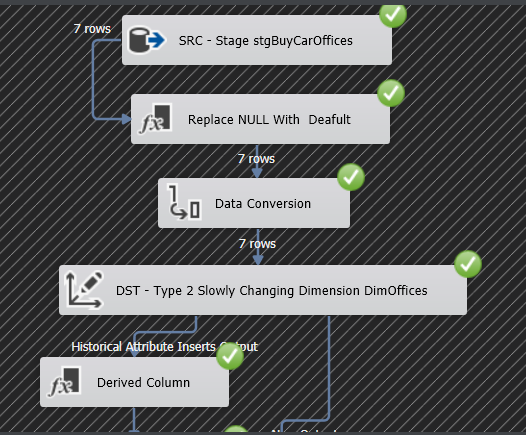


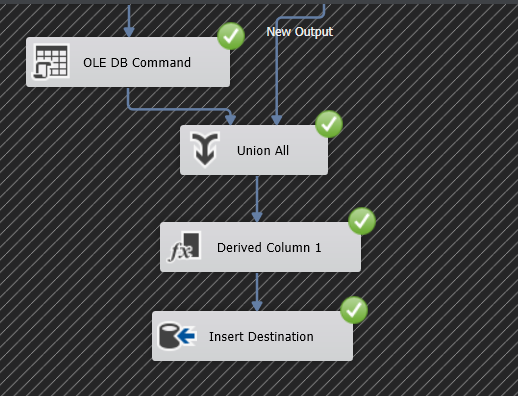
**2.4.2. Data Flow Stage to DimOffices**

* Load dữ liệu từ stgBuyCarOffices.
* Replace các giá trị NULL (trong trường hợp này là dữ liệu ở cột States).



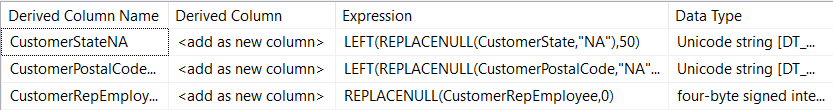
* Thực hiện Data Conversion.
* Dùng SCD (Slowly Changing Dimension) Type 2 (Historical). SCD trong trường hợp này được sử dụng với mục đích thêm dòng mới, đánh dấu dòng cũ không hoạt động trong bảng DimOffices.



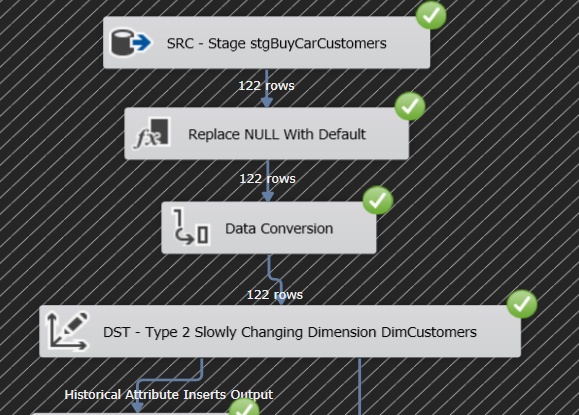


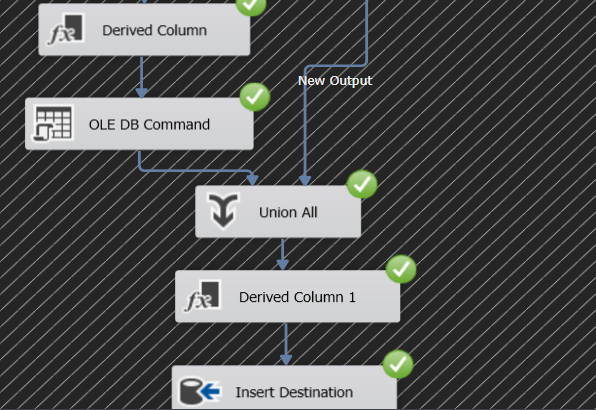
**2.4.3. Data Flow Stage to DimCustomers**

* Load dữ liệu từ stgBuyCarCustomers.
* Replace các giá trị NULL (trong trường hợp này là dữ liệu ở cột CustomerState, CustomerPostalCode, CustomerRepEmployee).



* Thực hiện Data Conversion.
* Dùng SCD (Slowly Changing Dimension) Type 2 (Historical). SCD trong trường hợp này được sử dụng với mục đích thêm dòng mới, đánh dấu dòng cũ không hoạt động trong bảng DimCustomers.



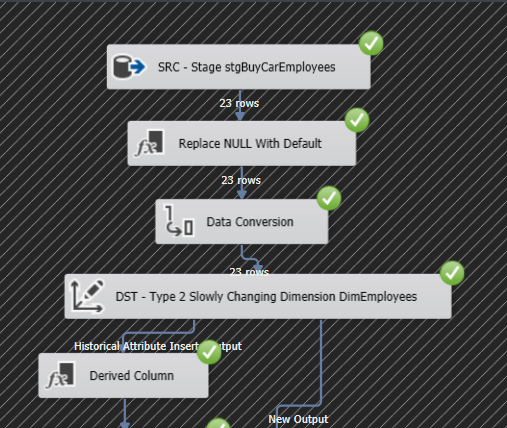


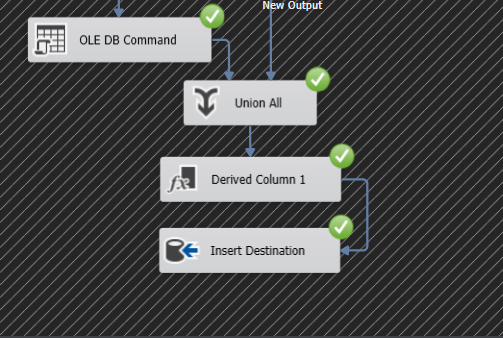
**2.4.4. Data Flow Stage to DimEmployees**

* Load dữ liệu từ stgBuyCarEmployees.
* Replace các giá trị NULL (trong trường hợp này là dữ liệu ở cột CustomerState, CustomerPostalCode, CustomerRepEmployee).



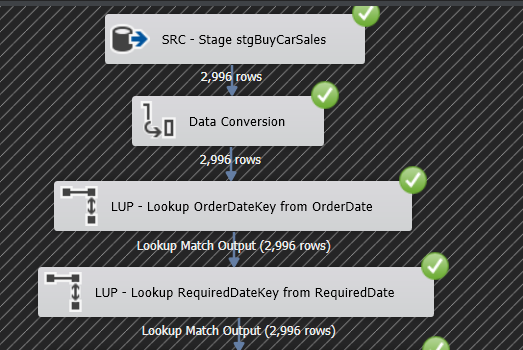
* Thực hiện Data Conversion.
* Dùng SCD (Slowly Changing Dimension) Type 2 (Historical). SCD trong trường hợp này được sử dụng với mục đích thêm dòng mới, đánh dấu dòng cũ không hoạt động trong bảng DimEmployees.

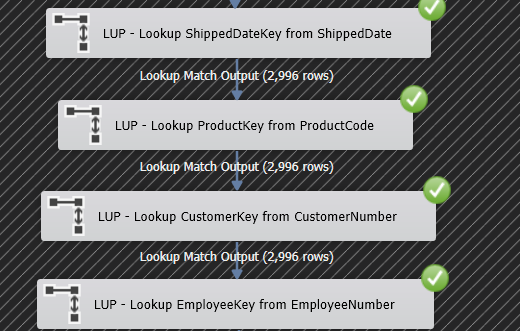


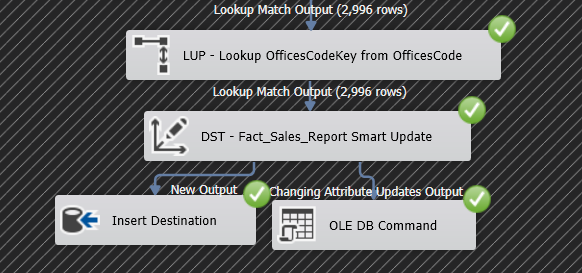


### **2.4.5. Data Flow Stage to Fact\_Sales\_Report**

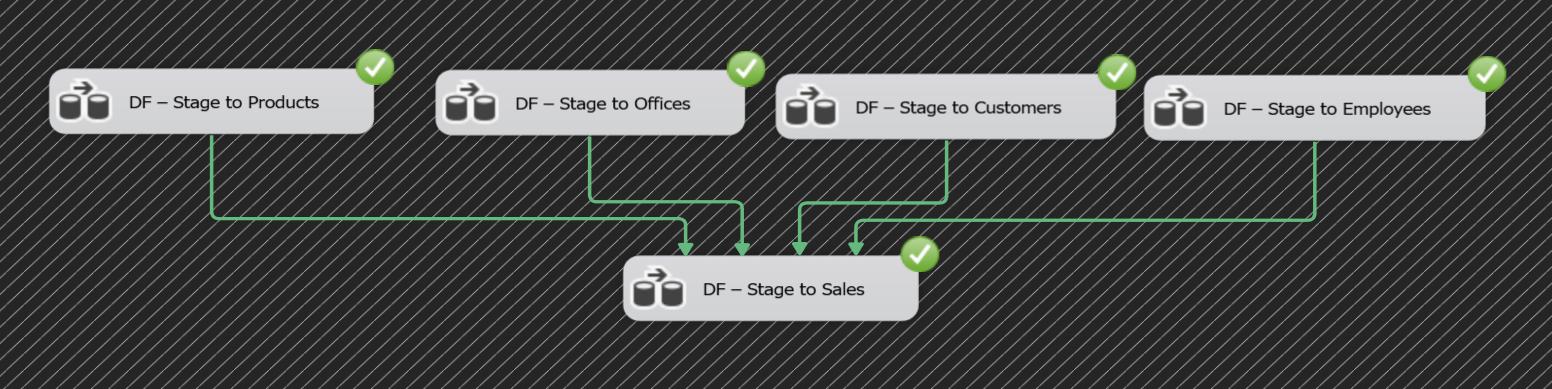
* Load dữ liệu từ stgBuyCarSales.
* Thực hiện Data Conversion.
* Mapping các khóa.
* Dùng SCD (Slowly Changing Dimension) Type 2 (Historical).







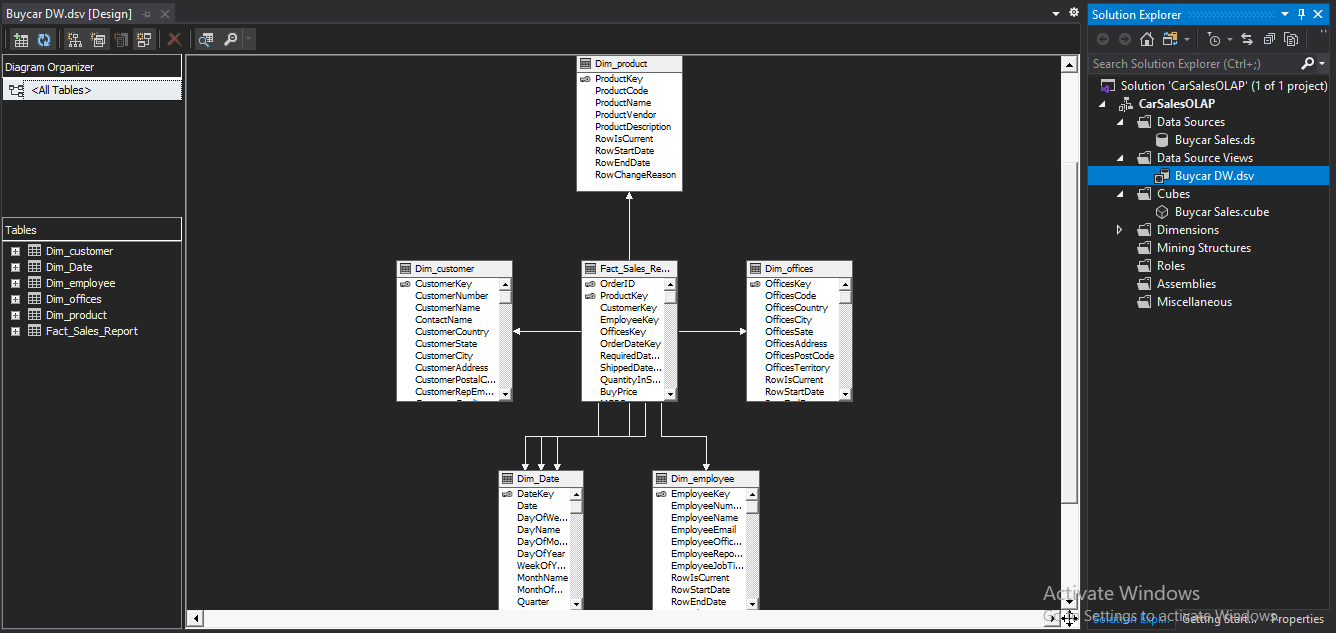
**2.4.6. Hoàn thành việc load dữ liệu vào BuycarDW**



# **Chương 3**

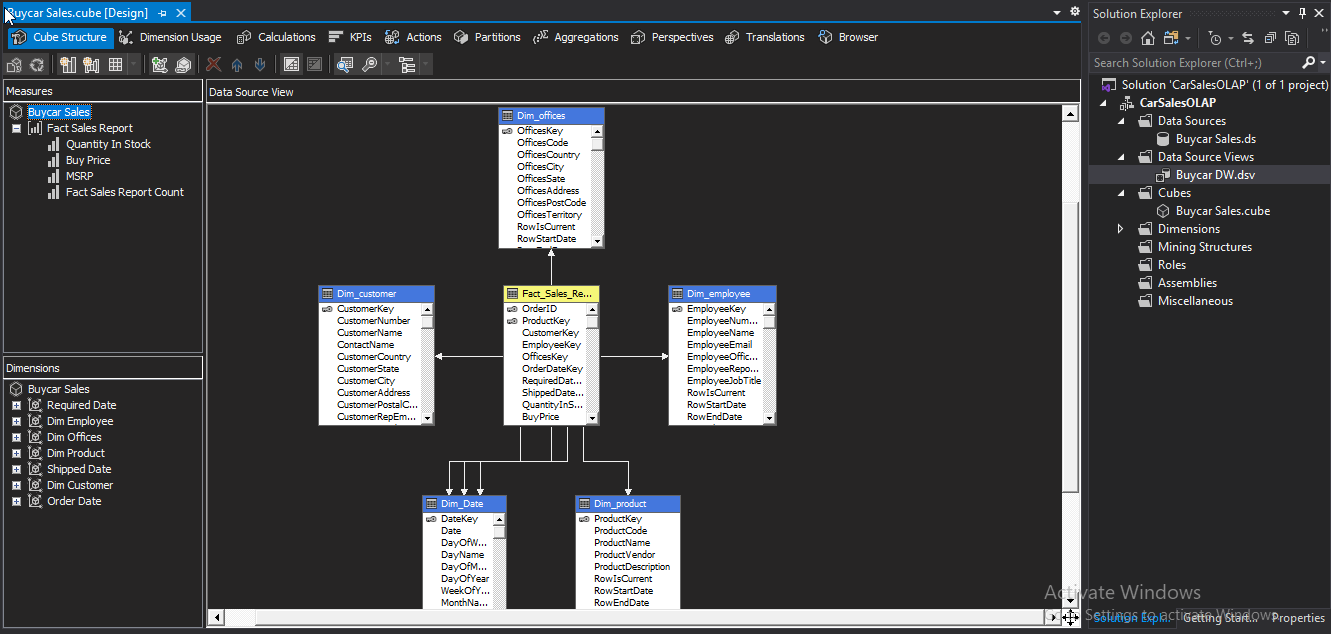
# **THIẾT KẾ DATA CUBE CarSalesOLAP**

## **3.1. Tạo Data Source View**



## **3.2. Tạo Cubes**

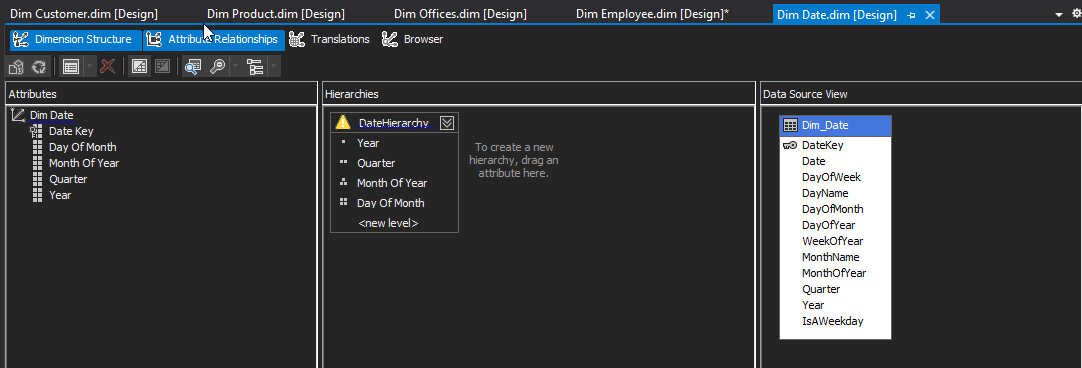
Buycar Sales.cube với các report



## **3.3. Tạo các Dimensions và các hierarchy**

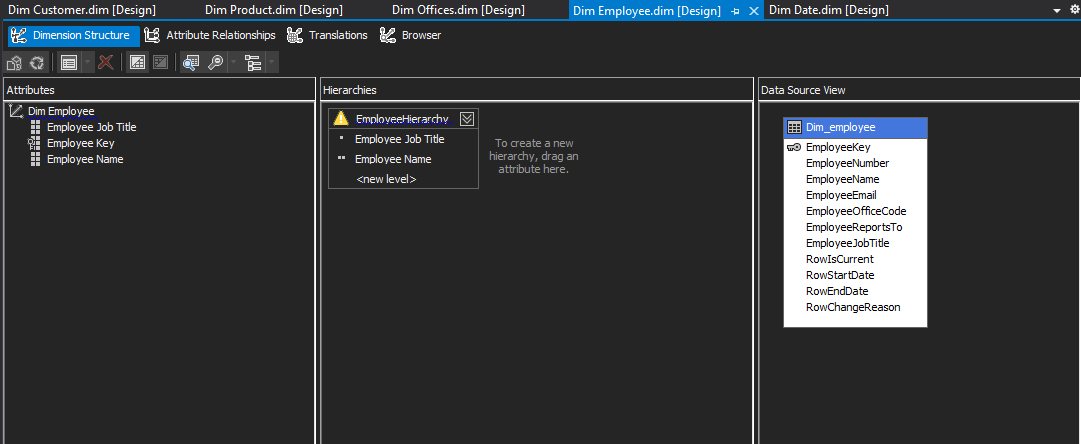
* Dim Date.dim

Datehierarchy: Year → Quarter → MonthOfYear → DayOfMonth



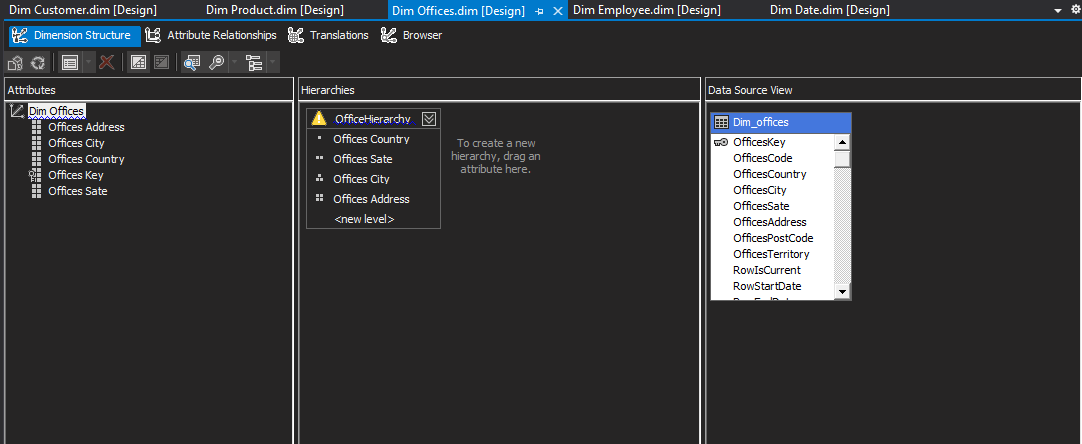
* Dim Employee.dim

EmployeeHierarchy: EmployeeJobTitle → EmployeeName



* Dim Offices.dim

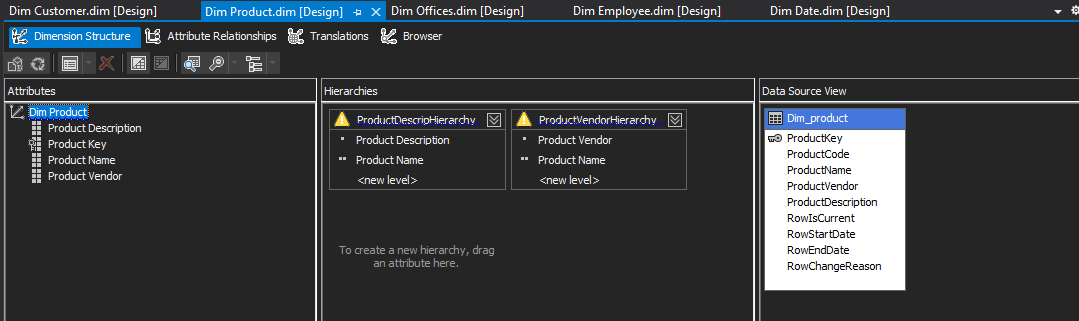
OfficeHierarchy: OfficesCountry → OfficesSate → OfficeCity → OfficeAddress



* Dim Product.dim

ProductDescripHierarchy: ProductDescription → ProductName

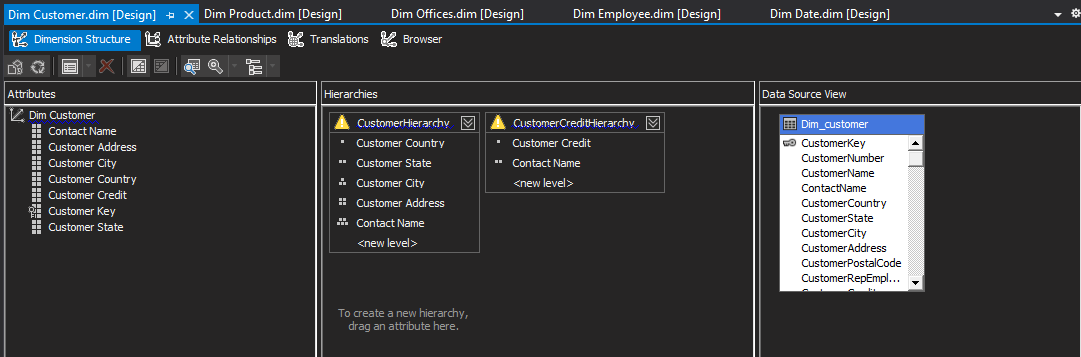
ProductVendorHierarchy: ProductVendor → ProductName



* Dim Customer.dim

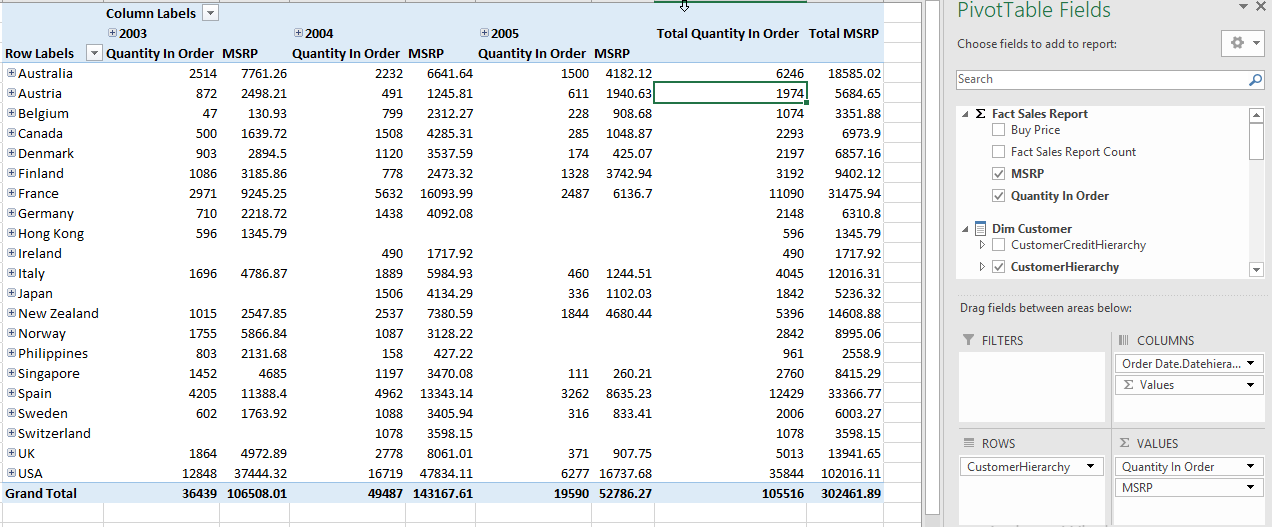
CustomerHierarchy: Customer → Country → CustomerState → CustomerCity → CustomerAddress → CustomerName

CustomerCreditHierarchy: CustomerCredit → CustomerName



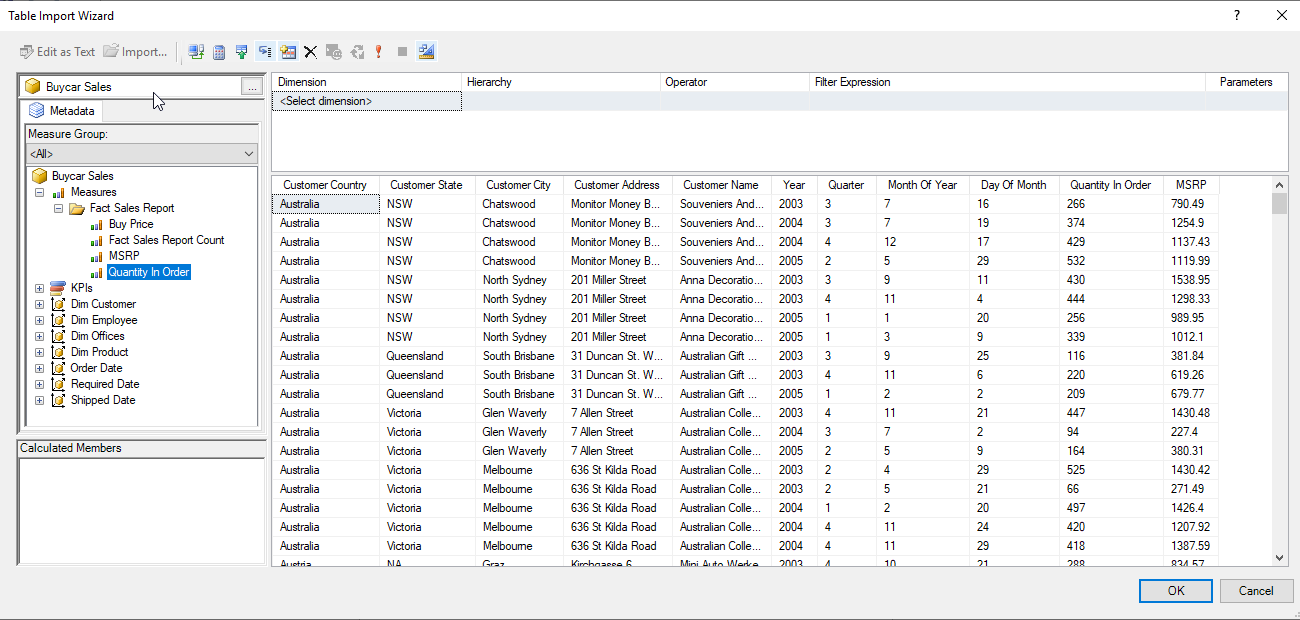
## **3.4. Xem Data Cube trong Excel với Pivot Tables**

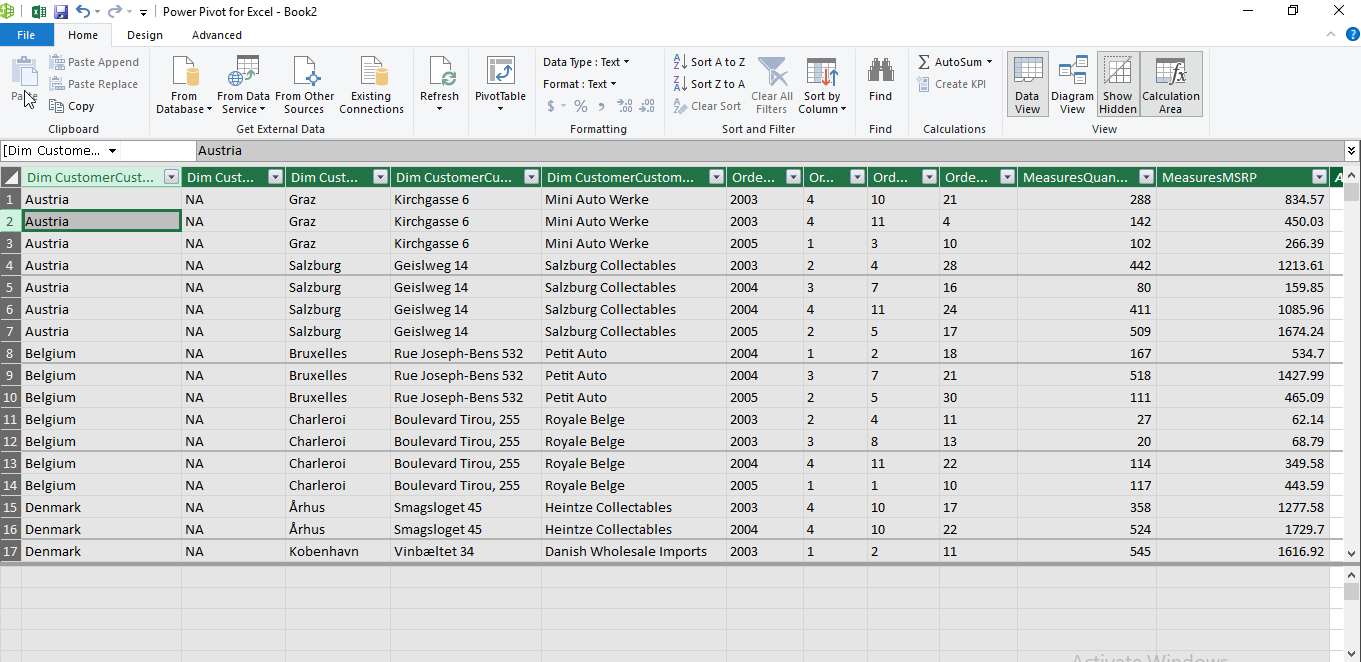
Ví dụ: Tổng số lượng và tổng số tiền bán được theo từng khu vực:



## **3.5. Excel Power Pivot**

Tạo file exel Customer by TimeOder để xem số lượng xe đã bán được trong từng khu vực (Quantity In Oder) và tổng số tiền thu được (MSRP).

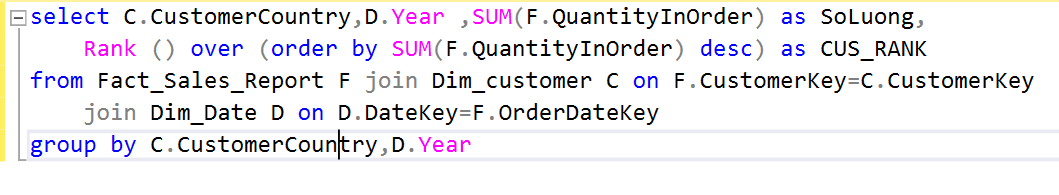




# **Chương 4**

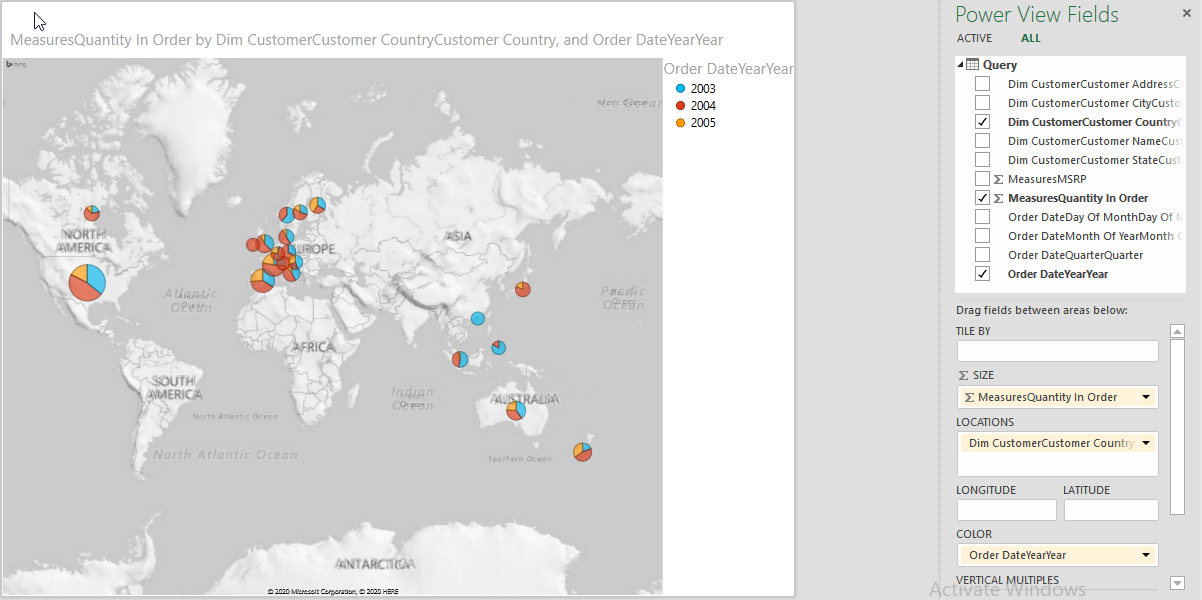
# **MỘT SỐ TRUY VẤN VÀ BÁO CÁO**

## **4.1. Truy vấn thống kê số lượng sản phẩm đặt hàng theo từng khu vực qua các năm**

****

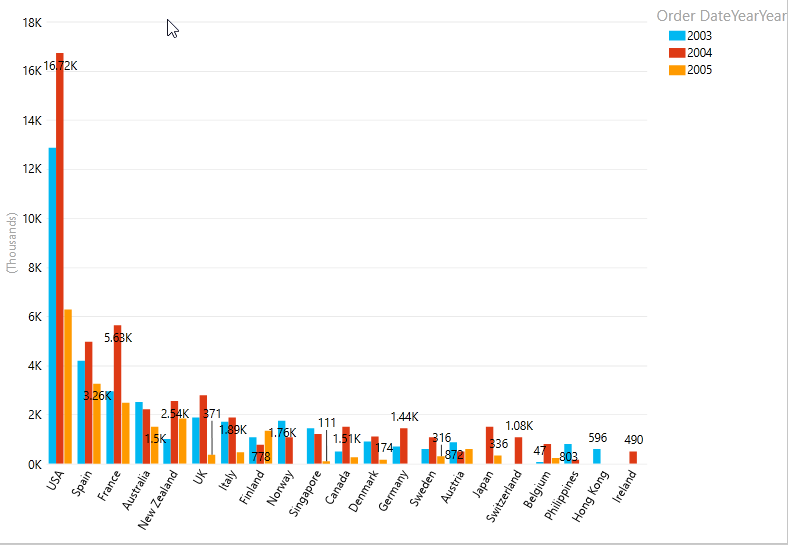
****

* **Biểu đồ Map thể hiện sự phân bố của số lượng sản phẩm đặt hàng theo từng khu vực qua các năm**



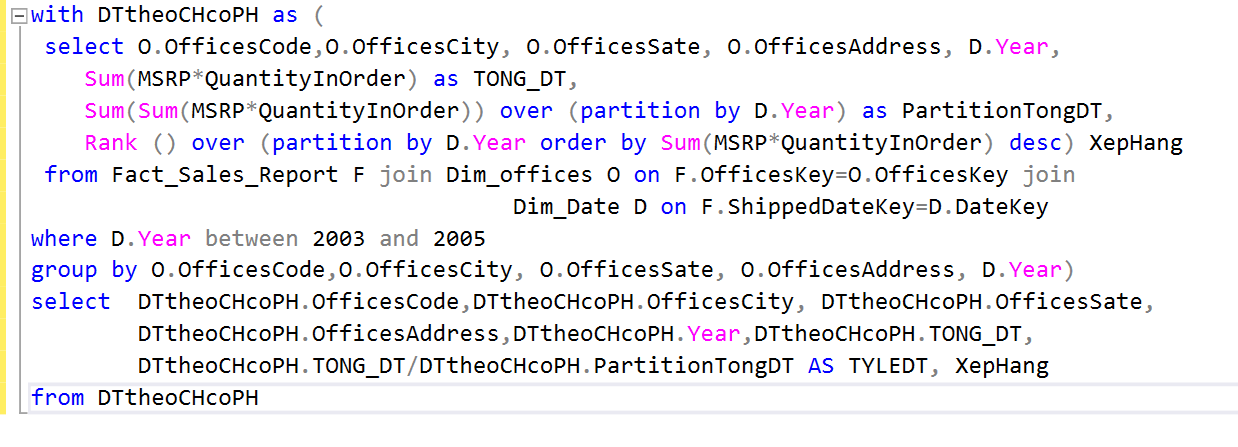
Qua biểu đồ Map ta thấy được sự phân bố của số lượng sản phẩm đặt hàng theo từng khu vực qua các năm. Tập trung nhiều nhất là ở khu vực Bắc Mỹ và Châu Âu, nằm rải rác ở các khu vực Đông Nam Á và Australia. Qua đây ta có thể đánh giá được tiềm năng phát triển ở từng khu vực một cách khách quan.

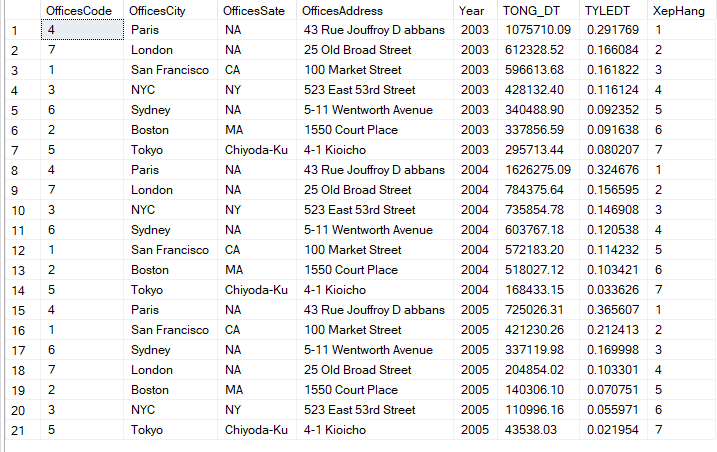
* **Biểu đồ cột thống kê số lượng sản phẩm đặt hàng theo từng khu vực qua các năm**



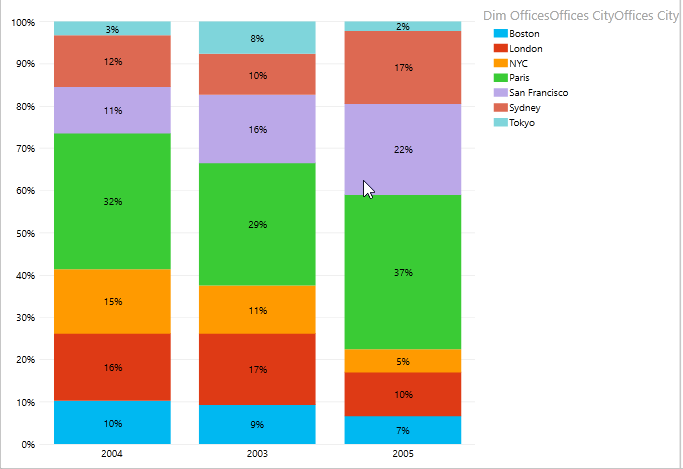
Biểu đồ cho ta cái nhìn chi tiết hơn về số lượng sản phẩm đặt hàng theo từng khu vực qua các năm. Điển hình, cao nhất là ở khu vực USA với số lượng sản phẩm đặt hàng của khách hàng rất cao và cao nhất trong cả 3 năm cụ thể là năm 2003 với hơn 12,8k đơn hàng, đỉnh điểm là vào năm 2004 với hơn 16,7k đơn hàng. Qua số liệu ta thấy đây là một thị trường lớn cần đẩy mạnh khai thác.

## **4.2. Truy vấn tỷ lệ doanh thu của các chi nhánh qua các năm**





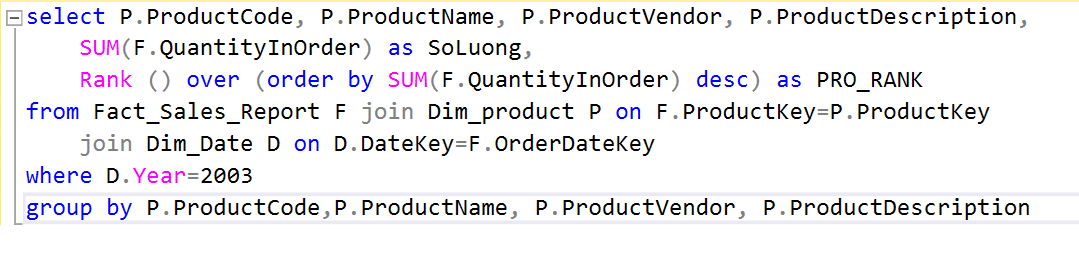
* **Biểu đồ tỷ lệ doanh thu của các chi nhánh qua các năm**



Biểu đồ cho ta cái nhìn rõ nét về doanh thu của các chi nhánh qua các năm, cụ thể chi nhánh ở Paris là chi nhánh mang lại doanh thu cao nhất, từ đây ta có thể xem xét những ưu điểm của chi nhánh này và phát triển nó ở các chi nhánh khác.

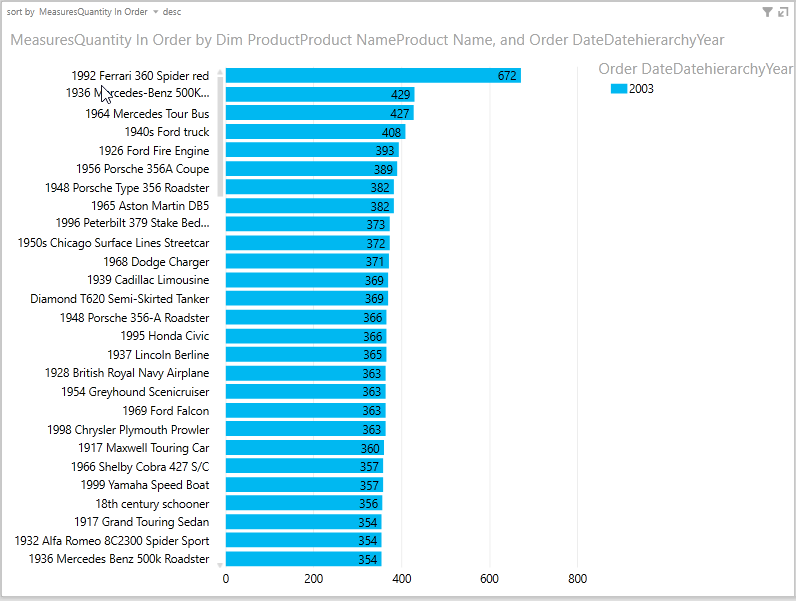
Đồng thời ta cũng thấy một số chi nhánh có doanh thu khá thấp là chi nhánh ở Tokyo thấp nhất là 2% vào năm 2005 và chi nhánh ở Boston cao nhất là 10% vào năm 2003 và giảm mạnh ở các năm sau đó. Từ đây ta có thể tư vấn cho doanh nghiệp các phương án phát triển như: Tìm các hướng marketing sản phẩm và đảy mạnh mở rộng phát triển, hoặc là cắt giảm nhân viên và chi nhánh để tập trung phát triển ở khu vực khác.

## **4.3. Truy vấn thống kê sản phẩm được ưa chuộng nhất năm 2003**



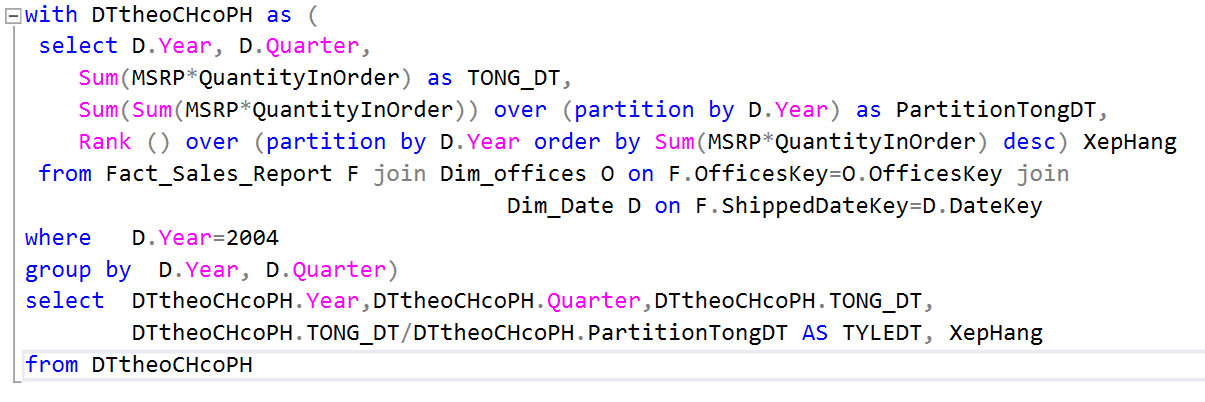


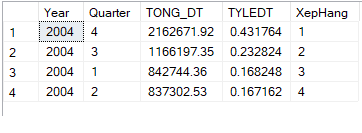
* **Biểu đồ thống kê sản phẩm được ưa chuộng nhất năm 2003**



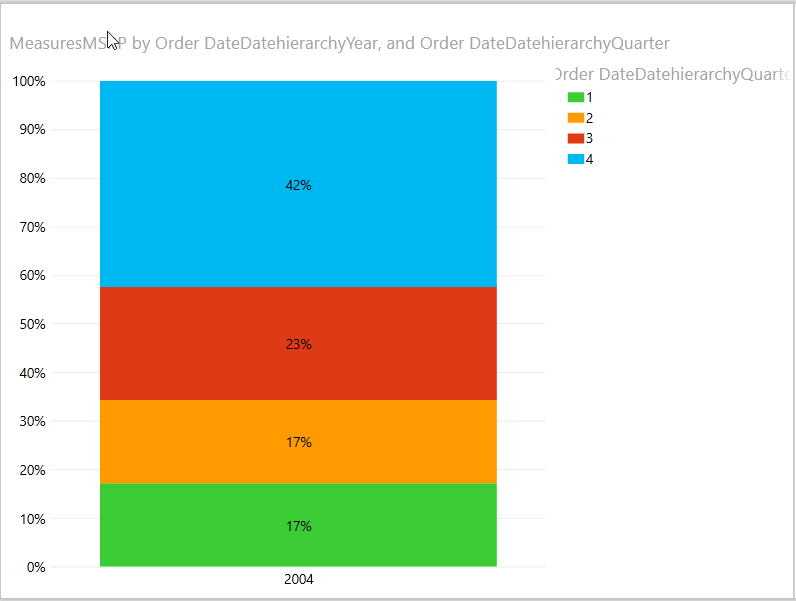
Biểu đồ cho ta thấy sản phẩm 1992 Ferrari 360 spider red là sản phẩm có số lượng order từ khách hàng nhiều nhất, cao gấp rưỡi so với các mặt hàng khác. Tiếp đến là các mặt hàng của hãng Mercedes và Ford. Từ đây doanh nghiệp có thể đẩy mạnh hợp tác phát triển các sản phẩm mới hoặc nhập số lượng hàng với giá ưu đãi với các hãng xe nổi tiếng này để nâng cao sự phát triển.

## **4.4. Truy vấn doanh thu và tỷ lệ doanh thu từng quý năm 2004**





* **Biểu đồ tỷ lệ phần trăm doanh thu từng quý năm 2004**



Dựa vào biểu đồ tỷ lệ doanh thu từng quý ta thấy các tháng đầu năm là các có sự phát triển chậm về doanh thu, chỉ chiếm 17%, và ở các tháng cuối năm (Quý 4) thường khách hàng sẽ có nhu cầu mua hàng rất lớn (42%). Từ đây doanh nghiệp có kế hoạch phát triển và nhập hàng cho từng chi nhánh của mình vào từng thời gian một cách hợp lý, tránh để cho số lượng hàng tồn kho quá lớn cũng như tránh thiếu hụt nguồn cung cho khách hàng.

# **KẾT LUẬN**

## **1. Những việc nhóm đã làm được**

* Đã thiết kế được Data Warehouse.
* Biết sử dụng các công cụ quá trình ETL (SSIS) và công cụ thống kê phân tích (SSAS).
* Sử dụng các hàm của Analytic Funtions để trả lời các Business Questions.

**2. Những việc nhóm chưa làm được và dự định thực hiện trong tương lai**

* Tạo thêm một data mart Phân tích hàng tồn kho (Product Inventory Analytic).
* Cở sở dữ liệu còn giới hạn nên việc xây dựng thêm các Business Process gặp khó khăn.

**TÀI LIỆU THAM KHẢO**

1. [**https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/books/data-warehouse-dw-toolkit/**](https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/books/data-warehouse-dw-toolkit/?fbclid=IwAR0ZhzkkLvYwqnDbzDA_Da4GmkMdBy7aUJMNripZJ6dnxOZ33g-issnHTAg)
2. [**https://www.microsoft.com/en-us/research/academic-program/give-great-research-talk/**](https://www.microsoft.com/en-us/research/academic-program/give-great-research-talk/?fbclid=IwAR0YLSMXZ7x4NjznEocHUAr6Hek1F1tp2AGkykEbM4wbs1Clxx72BghofqY)
3. [**https://www.1keydata.com/datawarehousing/datawarehouse.html**](https://www.1keydata.com/datawarehousing/datawarehouse.html?fbclid=IwAR3-qt9byb_aCb0ofApkmUL_lVEe8atTJIuMBvyu7g7T-4idXlfYFHWqXHk)
4. [**https://www.mysqltutorial.org/mysql-sample-database.aspx**](https://www.mysqltutorial.org/mysql-sample-database.aspx?fbclid=IwAR3ILEZ2p1psaXfKsAatetuPpWyCIDl8M1Ac9fQc8RQMksAJ3ugL2VR8O4c)