**Automated ETL Data Warehouse using SSIS and visualize data using PowerBI**

* **Report vẫn đang cập nhật**
* **Nhóm: 2 người**

**Đặc tả bộ dữ liệu nguồn:**

Bộ dữ liệu AdventureWorks2012 là một cơ sở dữ liệu mẫu của Microsoft gồm một số bảng mô tả như sau:

1. Bảng về Khách hàng và Bán hàng

| Bảng | Mô tả |
| --- | --- |
| Sales.Customer | Chứa thông tin về khách hàng cá nhân và doanh nghiệp. |
| Sales.SalesOrderHeader | Chứa thông tin chung về đơn hàng, như khách hàng, ngày đặt hàng, trạng thái, tổng tiền. |
| Sales.SalesOrderDetail | Chi tiết từng sản phẩm trong mỗi đơn hàng. |
| Sales.SalesPerson | Thông tin về nhân viên kinh doanh. |
| Sales.SalesTerritory | Thông tin khu vực bán hàng. |

2. Bảng về Sản phẩm và Danh mục

| Bảng | Mô tả |
| --- | --- |
| Production.Product | Chứa thông tin về sản phẩm như tên, giá, số lượng tồn kho. |
| Production.ProductCategory | Danh mục sản phẩm (Ví dụ: Xe đạp, Phụ tùng, Quần áo). |
| Production.ProductSubcategory | Nhóm con trong danh mục sản phẩm. |
| Production.ProductModel | Mô hình sản phẩm, chứa thông tin mô tả chi tiết hơn về sản phẩm. |

3. Bảng về Nhân sự

| Bảng | Mô tả |
| --- | --- |
| HumanResources.Employee | Chứa thông tin về nhân viên. |
| HumanResources.EmployeeDepartmentHistory | Lịch sử làm việc của nhân viên theo phòng ban. |
| HumanResources.Department | Danh sách phòng ban trong công ty. |
| HumanResources.JobCandidate | Thông tin về ứng viên xin việc. |

4. Bảng về Mua hàng và Nhà cung cấp

| Bảng | Mô tả |
| --- | --- |
| Purchasing.Vendor | Chứa thông tin về nhà cung cấp. |
| Purchasing.PurchaseOrderHeader | Thông tin chung về đơn hàng mua từ nhà cung cấp. |
| Purchasing.PurchaseOrderDetail | Chi tiết từng sản phẩm trong đơn hàng mua. |

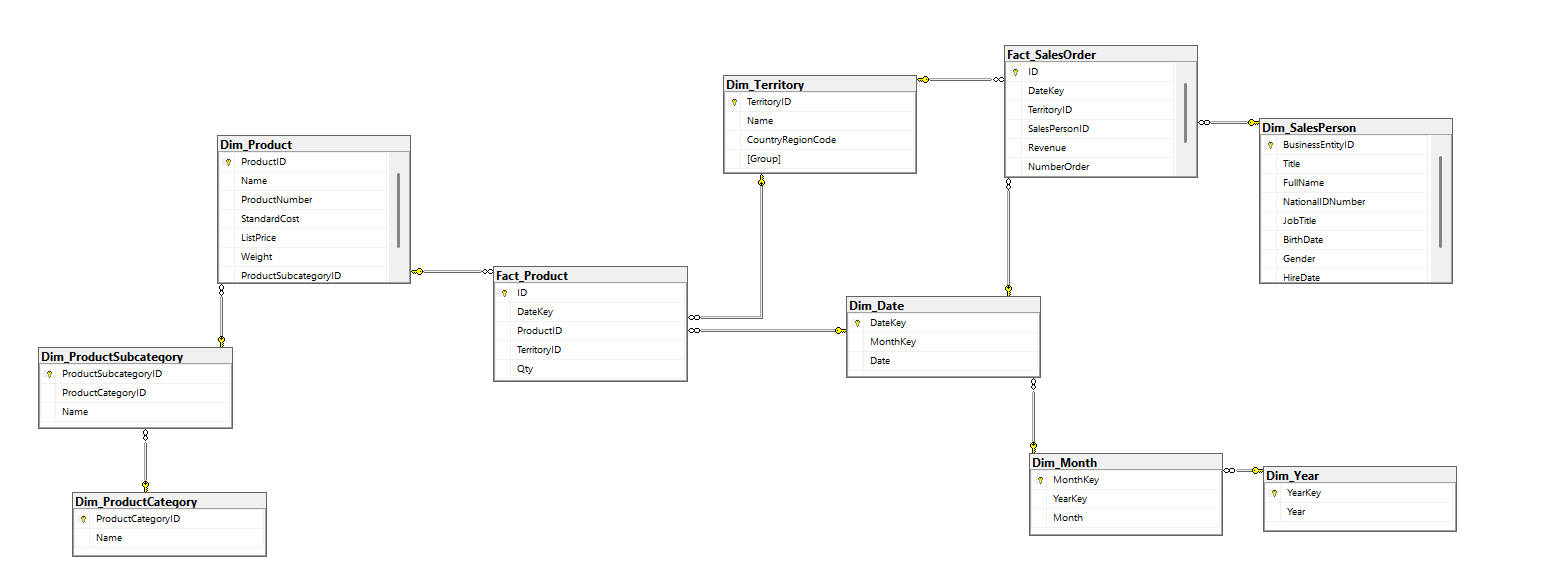
5. Bảng về Sản xuất và Tồn kho

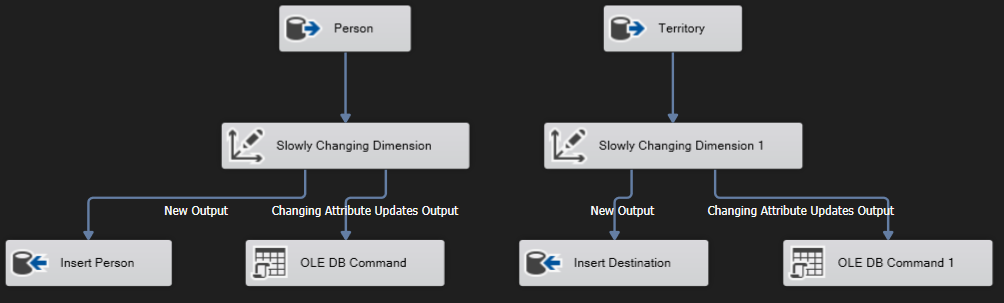
| Bảng | Mô tả |
| --- | --- |
| Production.WorkOrder | Thông tin về các lệnh sản xuất. |
| Production.WorkOrderRouting | Lộ trình sản xuất cho từng lệnh sản xuất. |
| Production.ProductInventory | Tồn kho sản phẩm tại các vị trí lưu trữ khác nhau. |

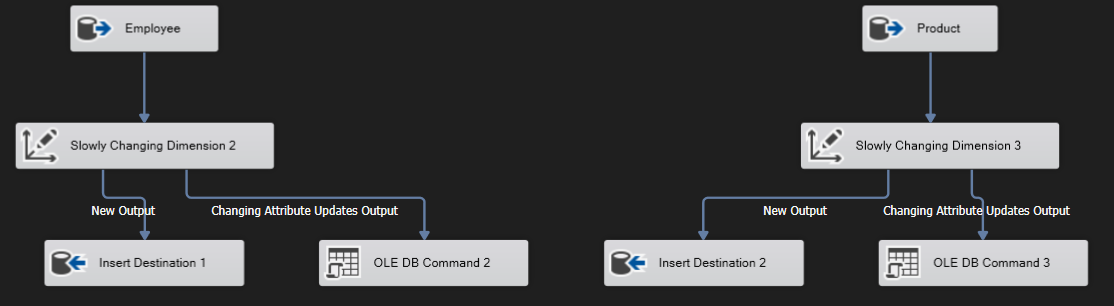
6. Bảng hệ thống và Quản lý tài khoản

| Bảng | Mô tả |
| --- | --- |
| Person.Person | Thông tin cá nhân chung (cả nhân viên và khách hàng). |
| Person.Address | Chứa thông tin địa chỉ của khách hàng, nhân viên. |
| Person.EmailAddress | Email của từng cá nhân. |
| Person.PhoneNumber | Số điện thoại liên hệ. |
| Person.StateProvince | Thông tin về các tỉnh, bang. |

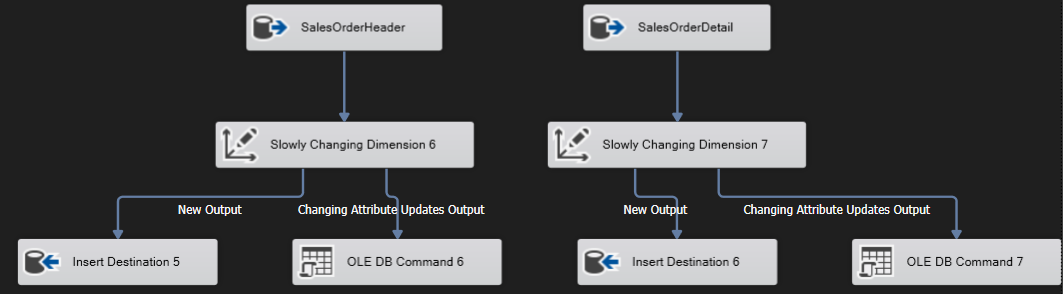
1. **Các báo cáo:**1. Báo cáo doanh số theo từng nhân viên.  
   2. Báo cáo doanh số theo từng khu vực.  
   3. Báo cáo số đơn hàng theo nhân viên.  
   4. Báo cáo số đơn theo khu vực.  
   5. Báo cáo số lượng bán ra theo nhóm sản phẩm.  
   6. Báo cáo số lượng bán ra theo sản phẩm và khu vực.
2. **Lựa chọn lược đồ:**Thiết kế theo lược đồ snowflake ( có phân cấp)



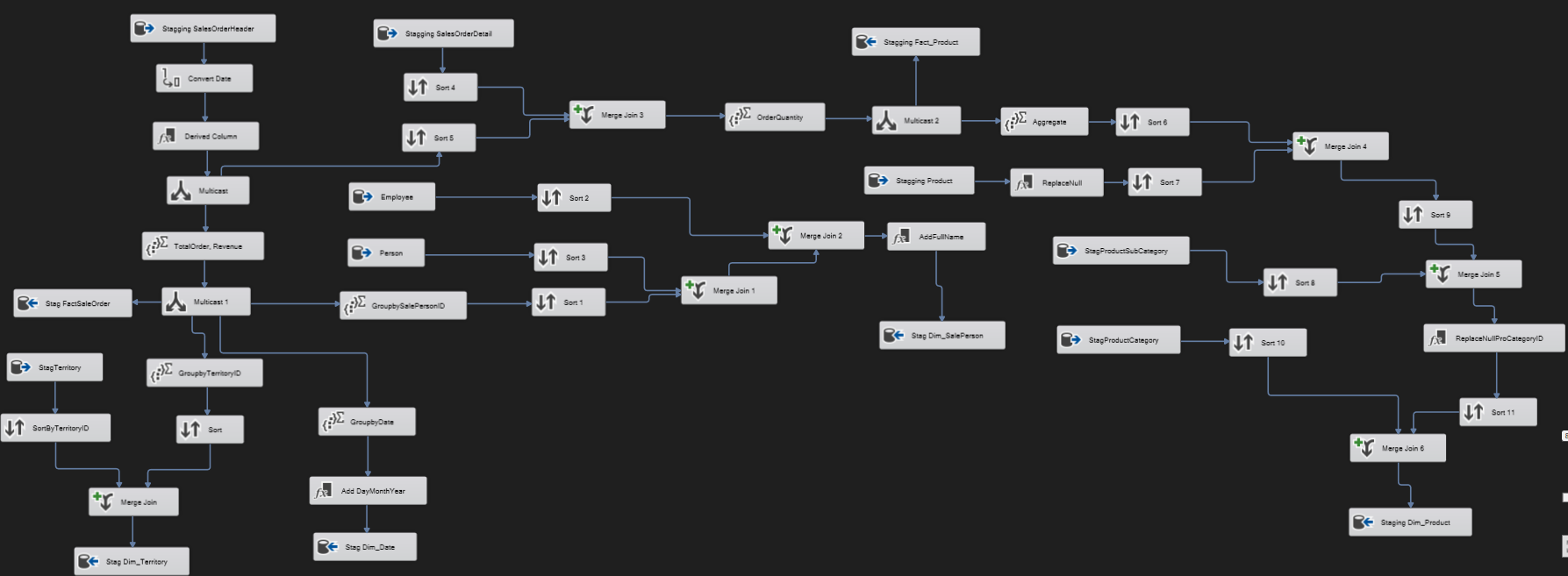
1. **Thiết kế bảng dimension**1. Báo cáo doanh số theo từng nhân viên. => nhân viên  
   2. Báo cáo doanh số theo từng khu vực. => khu vực  
   3. Báo cáo số đơn hàng theo nhân viên. => nhân viên  
   4. Báo cáo số đơn theo khu vực. => khu vực  
   5. Báo cáo số lượng bán ra theo nhóm sản phẩm. => nhóm sản phẩm  
   6. Báo cáo số lượng bán ra theo sản phẩm và khu vực. => sản phẩm, khu vực **=> Các bảng dimension cần tạo: nhân viên, khu vực, sản phẩm, nhóm sản phẩm, thời gian (ngày, tháng, năm).**
2. **Thiết kế bảng FACT**1. Báo cáo doanh số theo từng nhân viên. => doanh số  
   2. Báo cáo doanh số theo từng khu vực. => doanh số  
   3. Báo cáo số đơn hàng theo nhân viên. => số đơn hàng  
   4. Báo cáo số đơn theo khu vực. => số đơn hàng  
   5. Báo cáo số lượng bán ra theo nhóm sản phẩm. => số lượng sản phẩm bán  
   6. Báo cáo số lượng bán ra theo sản phẩm và khu vực. => số lượng sản phẩm bán  
    **=> Các bảng Fact cần tạo:**- Báo cáo 1->4: Doanh số + Số đơn hàng => Tạo bảng Fact Fact\_SalesOrder có các thuộc tính [Id, revenue, NumberOrder] với các khoá ngoại từ bảng Dim\_SalesPerson, Dim\_Territory, Dim\_Date.  
   - Báo cáo 5, 6: Số lượng sản phẩm => Tạo bảng Fact Fact\_Product có các thuộc tính [Id, Qty] với các khoá ngoại từ bảng Dim\_Product, Dim\_Territory, Dim\_Date.
3. **ETL  
   1. Extract**  
   - Lấy data từ các bảng: Employee, Person, SalesTerritory, Product, ProductSubCategory, ProductCategory, SalesOrderHeader, SalesOrderDetail.  
   - Tạo table staging từ các bảng trên



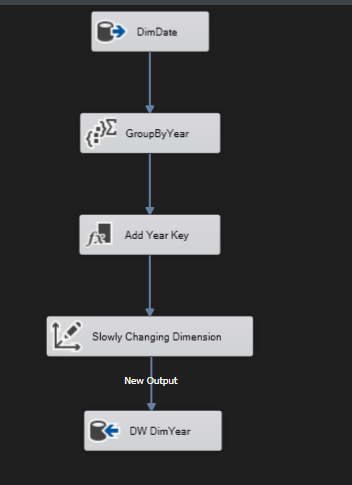




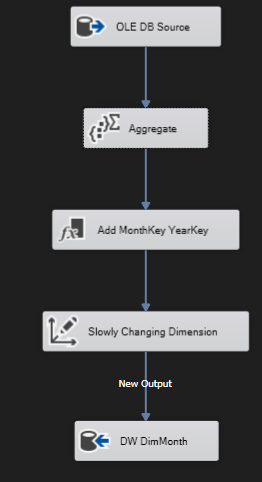
**2. Tranform  
Tổng hợp các bước transform ( datetime, replaceNull, chuẩn hóa)**



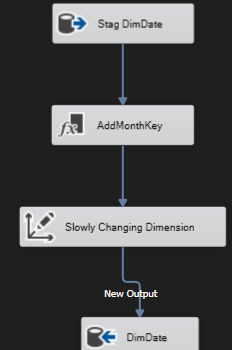
**3.** Load  
a. Load Dim Year



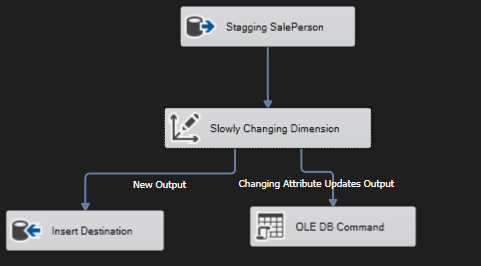
b. Load Dim Month



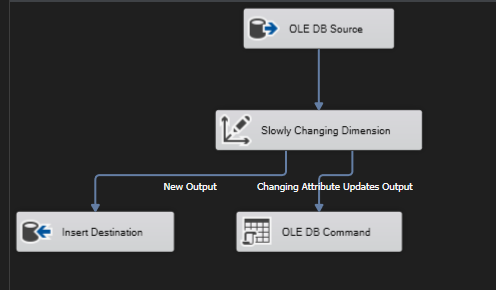
c. Load Dim Date

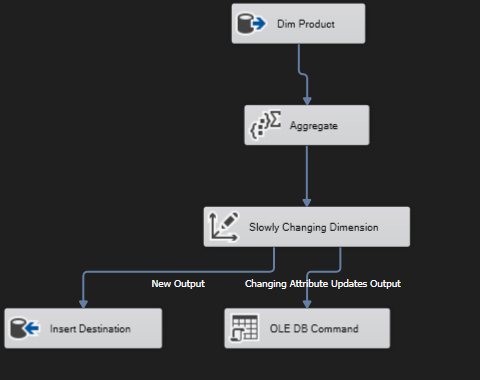


d. Load Dim SalePerson:



e. Load Dim Territory

  
f. Load Dim Product Category



**g. Load Dim ProductSubCategory**

**h. Load Dim Product**

**i. Load Fact**

(phần g,h,i vẫn đang update tiếp )  
**4. Tổng quát**  


1. **Report BI**

**Đang update**

1. **Lập lịch sử dụng SQL Server Agent**

**Đag update**