**PROJECT XÂY DỰNG DATABASE VÀ DATA WAREHOUSE CHO MỘTCHUỖI CỬA HÀNG**

1. **Xây dựng Database.  
   1. Xác định các table và các thành phần dữ liệu.**- **City** (City\_Id, Name, CreatedDate, UpdatedDate)  
   - **Store** (Store\_Id, Code, Name, CreatedDate, UpdatedDate)  
   - **Employee** (Employee\_Id, Code, Name, DOB, HireDate, IdentityNumber, Address, CreatedDate, UpdatedDate)  
   - **Customer** (Customer\_Id, Name, Phone, CreatedDate, UpdatedDate)  
   - **Source\_Online** (Source\_Online\_Id, Name, CreatedDate, UpdatedDate)  
   - **Brand** (Brand\_Id, Name, CreatedDate, UpdatedDate)  
   - **Category** (Category\_Id, Name, CreatedDate, UpdatedDate)  
   - **Product** (Product\_Id, Code, Name, Description, Size, Color, StandardPrice, CreatedDate, UpdatedDate)  
   - **Order\_Header** (Order\_Header\_Id, Code, OrderDate, Total, StandardCost, Status, IsOnline, CreatedDate, UpdatedDate)  
   - **Order\_Detail** (Order\_Detail\_Id, Qty, Price, Total, Code, CreatedDate, UpdatedDate)  
   **2. Data Type của từng thành phần dữ liệu cho từng table.**

|  |  |
| --- | --- |
| **Brand** | |
| ***Tên cột*** | ***Data type*** |
| Brand\_Id | Nvarchar(50) |
| Name | Nvarchar(50) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

|  |  |
| --- | --- |
| **Category** | |
| ***Tên cột*** | ***Data type*** |
| Category\_Id | Nvarchar(50) |
| Name | Nvarchar(250) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

|  |  |
| --- | --- |
| **City** | |
| ***Tên cột*** | ***Data type*** |
| City\_Id | Nvarchar(50) |
| Name | Nvarchar(250) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

|  |  |
| --- | --- |
| **Customer** | |
| ***Tên cột*** | ***Data type*** |
| Customer\_Id | Nvarchar(50) |
| Name | Nvarchar(250) |
| Phone | Nvarchar(50) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

|  |  |
| --- | --- |
| **Employee** | |
| ***Tên cột*** | ***Data type*** |
| Employee\_Id | Nvarchar(50) |
| Code | Nvarchar(10) |
| Name | Nvarchar(250) |
| DOB | Date |
| HireDate | Date |
| IdentityNumber | Nvarchar(50) |
| Address | Nvarchar(500) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

|  |  |
| --- | --- |
| **Order\_Detail** | |
| ***Tên cột*** | ***Data type*** |
| Order\_Detail\_Id | Nvarchar(50) |
| Qty | Numeric(8,0) |
| Price | Numeric(18,2) |
| Total | Numeric(18,2) |
| Code | Nvarchar(50) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

|  |  |
| --- | --- |
| **Order\_Header** | |
| ***Tên cột*** | ***Data type*** |
| Order\_Header\_Id | Nvarchar(50) |
| Code | Nvarchar(50) |
| OrderDate | Date |
| Total | Decimal(18,2) |
| StandardCost | Decimal(18,2) |
| Status | Int |
| IsOnline | Bit |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

|  |  |
| --- | --- |
| **Product** | |
| ***Tên cột*** | ***Data type*** |
| Product\_Id | Nvarchar(50) |
| Code | Nvarchar(50) |
| Name | Nvarchar(255) |
| Description | Nvarchar(50) |
| Size | Nvarchar(50) |
| Color | Nvarchar(50) |
| StandardPrice | Numeric(18,2) |
| Price | Numeric(18,2) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

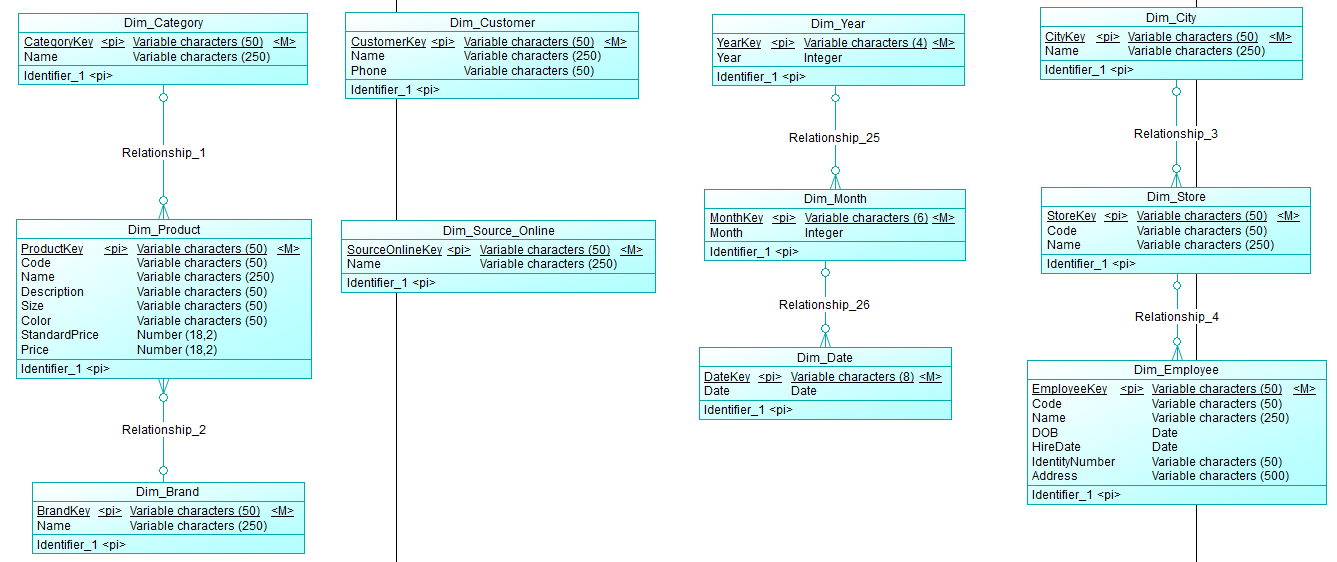
|  |  |
| --- | --- |
| **Source\_Online** | |
| ***Tên cột*** | ***Data type*** |
| Source\_Online\_Id | Nvarchar(50) |
| Name | Nvarchar(255) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

|  |  |
| --- | --- |
| **Store** | |
| ***Tên cột*** | ***Data type*** |
| Store\_Id | Nvarchar(50) |
| Code | Nvarchar(50) |
| Name | Nvarchar(250) |
| CreatedDate | Datetime |
| UpdatedDate | Datetime |

**3. ERD.  
Ảnh có chứa biểu đồ, sơ đồ

Mô tả được tạo tự động  
  
4. EER.  
Ảnh có chứa biểu đồ

Mô tả được tạo tự động**

1. **Xây dựng Data Warehouse.  
   1. Yêu cầu.**- Doanh số, lợi nhuận, số đơn theo các tiêu chí sản phẩm, cửa hàng, nhân viên, khu vực, khách hàng.  
   - Số lượng sản phẩm bán ra theo nhóm, cửa hàng, khu vực, thương hiệu.  
   - Tỉ lệ doanh số, số đơn online/offline theo thời gian.  
   - Với các đơn online cơ cấu số đơn, doanh số theo nguồn.  
   - Tổng doanh số theo cửa hàng, nhân viên, sản phẩm.  
   - Top 10 sản phẩm bán chạy nhất, chậm nhất. => Doanh số theo sản phẩm  
   - Top 10 nhân viên, của hàng có doanh số lớn nhất, nhỏ nhất. => Doanh số theo nhân viên, cửa hàng.  
     
   **2. Phân tích.**  
   - Dim: sản phẩm, cửa hàng, nhân viên, nhóm, khu vực, thời gian, nguồn  
   - Fact: doanh số, lợi nhuận, số đơn, số lượng sản phẩm bán ra, tỉ lệ doanh số onl/ off, số đơn online/offline.  
   => FactSaleOrder(doanh số, doanh số onl, doanh số off, lợi nhuận, số đơn onl, số đơn off)  
   FK: [Nhân viên <= cửa hàng <= khu vực], [khách hàng], [thời gian], [nguồn]  
   => FactProduct(số lượng sản phẩm, doanh số sản phẩm)  
   FK: [Sản phẩm <= Category, Brand], [cửa hàng <= khu vực], [thời gian].  
     
   **3. Lựa chọn lược đồ.**Lựa chọn lược đồ bông tuyết (phân cấp)  
    **4. Xây dựng các bảng Dimension.**- Dim\_Product (ProductKey, Code, Name, Description, Size, Color, StandardPrice).  
   - Dim\_Store (StoreKey, Code, Name).  
   - Dim\_Employee (EmployeeKey, Code, Name, DOB, HireDate, IdentityNumber, Address).  
   - Dim\_Category (CategoryKey, Name).  
   - Dim\_City (CityKey, Name).  
   - Dim\_Date (DateKey, Date).  
   - Dim\_Month (MonthKey, Month).   
   - Dim\_Year (YearKey, Year).  
   - Dim\_Source\_Online (SourceOnlineKey, Name).  
   - Dim\_Customer (CustomerKey, Name, Phone).  
   - Dim\_Brand (BrandKey, Name).  
     
     
     
     
   \* ERD Dim table  
   ****  
     
   \* EER Dim table  
   **Ảnh có chứa văn bản, ảnh chụp màn hình, trong nhà

   Mô tả được tạo tự động**  
    **4. Xây dựng các bảng Fact.**- Fact\_Sales\_Order(Id, Revenue, RevenueOnline, RevenueOffline, Profit, NumberOrder, NumberOrderOnline, NumberOrderOffline).  
   - Fact\_Product(Id, Qty, Revenue).\* ERD Fact table **Ảnh có chứa văn bản

   Mô tả được tạo tự động  
   5. Data Warehouse.**\* ERD DW **Ảnh có chứa biểu đồ, sơ đồ

   Mô tả được tạo tự động**  
   \* EER DW **Ảnh có chứa biểu đồ, sơ đồ

   Mô tả được tạo tự động**
2. **ETL.  
   1. Tạo vùng Staging.**- Stag Dim\_Date.  
   - Stag Dim\_Product.  
   - Stag Dim\_Store.  
   - Stag Dim\_City.  
   - Stag Dim\_Brand.  
   - Stag Dim\_Category.  
   - Stag Fact\_Product.  
   - Stag Dim\_Source\_Online.  
   - Stag Dim\_Customer.  
   - Stag Dim\_Employee.  
   - Stag Fact\_SalesOrder.  
   - Stag Category.  
   - Stag Product.  
   - Stag Brand.  
   - Stag Order\_Detail.  
   - Stag Order\_Header.  
   - Stag Customer.  
   - Stag Employee.  
   - Stag Store.  
   - Stag City.  
   - Stag Source\_Online.  
     
   **2. Extract.**Lấy source data từ các bảng [Category, Product, Brand, Order\_Detail, Order\_Header, Customer, Employee, Store, City, Source\_Online.] đổ vào các bảng staging tương ứng.  
   VD:  
   A screenshot of a computer

   Description automatically generated  
     
   **3. Tranform.**=> FactSaleOrder(doanh số, doanh số onl, doanh số off, lợi nhuận, số đơn onl, số đơn off)  
   FK: Nhân viên <= cửa hàng <= khu vực, khách hàng, thời gian, nguồn  
   OrderHeader(count(\*) => NumberOrder, sum(Total) => Revenue, sum(Total-StandardCost) => Profit, OrderDate => date, [IsOnline =True => NumberOrderOnline, RevenueOnline], [IsOnline=False => NumberOrderOffline, RevenueOffline])  
     
   => FactProduct(số lượng sản phẩm, doanh số sản phẩm)  
   FK: Sản phẩm <= nhóm, thương hiệu, cửa hàng <= khu vực, thời gian.  
   OrderDetail(sum(Qty) => NumberProduct, sum(Total) => RevenueProduct)  
     
   A computer screen shot of a computer

   Description automatically generated with low confidence  
   A screenshot of a computer

   Description automatically generated  
   A computer screen shot of a computer

   Description automatically generated with low confidence  
     
   **4. Load  
     
   a. Load DW Dim Lv1**  
   DimYear, DimSourceOnline, DimCustomer, DimCity, DimBranch  
   A screenshot of a computer

   Description automatically generated  
     
   **b. Load DW Dim Lv2**  
   DimMonth, DimStore, DimProduct  
   A screenshot of a computer

   Description automatically generated  
     
   **c. Load DW Dim Lv3**  
   DimDate, DimEmployee  
   A screenshot of a computer

   Description automatically generated  
     
   **d. Load DW Fact**  
   FactSalesOrder, FactProduct  
   A computer screen shot of a computer program

   Description automatically generated with low confidence  
     
   **5. Truncate staging**  
   A screenshot of a computer

   Description automatically generated
3. **Create reports (Power BI)**A screenshot of a computer

   Description automatically generated with medium confidence
4. **COPY DATA FROM MSSQL TO POSTGRES**A screenshot of a computer

   Description automatically generated