

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THỰC PHẨM TP. HỒ CHÍ MINH**  
**KHOA CÔNG NGHỆ THÔNG TIN**

-----oOo-----



**BÁO CÁO BÀI TẬP NHÓM**

**ĐỀ TÀI: ÁP DỤNG KHO DỮ LIỆU VÀ KỸ THUẬT OLAP**  
**PHÂN TÍCH VÀ KHAI THÁC KẾT QUẢ KHÁM CHỮA BỆNH**  
**CỦA CÁC CƠ SỞ Y TẾ**

**Giáo viên hướng dẫn: Đinh Thị Mận**

**Sinh viên thực hiện:**

**Đỗ Thị Hồng Thắm – 2001170165**

**Nguyễn Phi Nhung - 2001170124**

TP. Hồ Chí Minh, tháng 7 năm 2020



## NHẬN XÉT ĐÁNH GIÁ CỦA GIẢNG VIÊN HƯỚNG DẪN

### ĐỀ TÀI: ÁP DỤNG KHO DỮ LIỆU VÀ KỸ THUẬT OLAP PHÂN TÍCH VÀ KHAI THÁC KẾT QUẢ KHÁM CHỮA BỆNH CỦA CÁC CƠ SỞ Y TẾ

Bảng phân công nhiệm vụ:

Thành viên	Công việc
Đỗ Thị Hồng Thắm	Tìm hiểu kho dữ liệu ,tạo cơ sở dữ liệu,làm word, sử dụng ssis vào kho dữ liệu
Nguyễn Phi Nhung	Sử dụng ssas vào kho dữ liệu, tạo MDX ,truy vấn MDX, tạo phần mềm.

Nhận xét của giáo viên: .....

.....

.....

.....

.....

.....

.....

.....

.....

.....

TP.HCM, ngày      tháng      năm 2020

Giảng viên

(Ký và ghi rõ họ tên)

## MỤC LỤC

### I. TÌM HIỂU VỀ KHO DỮ LIỆU VÀ OLAP

1. Khái niệm Data Warehouse.....	1
2. Khái niệm về kho dữ liệu .....	1
3. Tính chất.....	1
4. Cấu trúc .....	2
5. Tiến trình xử lý.....	3
6. Ưu nhược điểm của kho dữ liệu .....	3

### II. TÌM HIỂU VÀ PHÂN TÍCH NGHIỆP VỤ ( TRONG Y TẾ)

1. Vì sao cần sử dụng kho dữ liệu vào y tế?.....	5
2. Quy trình khám và chữa bệnh .....	6
3. Phân tích nghiệp vụ khám chữa bệnh trong y tế .....	6

### III. THIẾT KẾ VÀ CÀI ĐẶT: DATA WAREHOUSE, MÔ HÌNH NẠP DỮ LIỆU TỪ HỆ THỐNG QUẢN LÝ VÀO DATAWAREHOUSE .....

7

### IV. CUBE .....

1.Cài đặt.....	14
2.Truy vấn MDX .....	15

### V.ỨNG DỤNG .....

### TÀI LIỆU THAM KHẢO .....

## **I. TÌM HIỂU VỀ KHO DỮ LIỆU VÀ OLAP:**

### **1. Khái niệm về Data Warehouse**

**Data Warehouse** là kho dữ liệu tập trung chuyên dùng cho tạo báo cáo và phân tích dữ liệu. Data Warehouse hỗ trợ các truy vấn phức tạp, vừa là điểm tập trung dữ liệu từ nhiều nguồn khác nhau để có được thông tin phân tích đầy đủ nhất

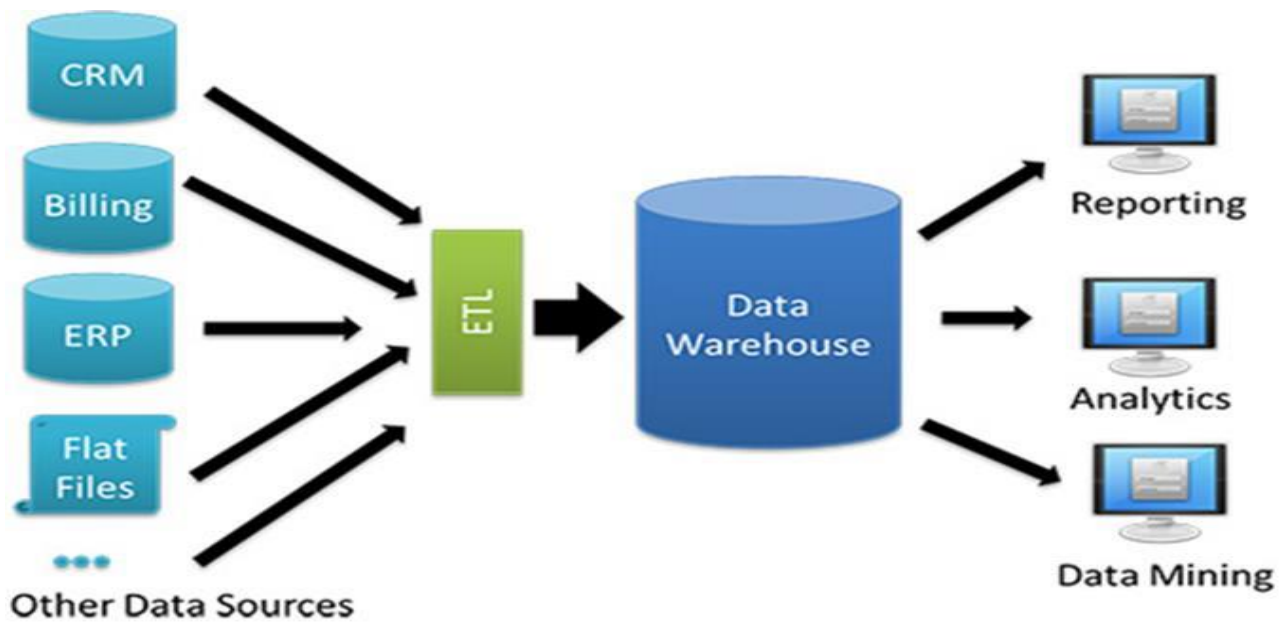
### **2. Khái niệm về kho dữ liệu**

- Kho: Chứa dữ liệu hiện tại và quá khứ, hỗ trợ việc đưa ra quyết định mà nhà quản lý quan tâm
- Dữ liệu: Được cấu trúc luôn sẵn sàng cho các hoạt động phân tích, xử lý
- Kho dữ liệu: là một cơ sở dữ liệu tập hợp từ nguồn của một tổ chức và chủ yếu được sử dụng cho việc báo cáo, phân tích

### **3. Tính chất**

- **Tính hướng chủ đề:** nghĩa là Data Warehouse tập trung vào việc phân tích các yêu cầu quản lý ở nhiều cấp độ khác nhau trong quy trình ra quyết định. Các yêu cầu phân tích này thường rất cụ thể, và xoay quanh loại hình kinh doanh của doanh nghiệp, ví dụ các công ty phân phối sẽ quan tâm đến tình hình kinh doanh, doanh nghiệp viễn thông quan tâm đến lưu lượng dịch vụ... Tuy nhiên một doanh nghiệp thường quan tâm đến vài chủ đề khác nhau, như công ty phân phối còn phải quan tâm đến kho bãi, chuỗi cung ứng...
- **Tính toàn vẹn (Integrated):** Data Warehouse giải quyết các khó khăn trong việc kết hợp dữ liệu từ nhiều nguồn dữ liệu khác nhau, giải quyết các sai khác về tên trường dữ liệu (dữ liệu khác nhau nhưng tên giống nhau), ý nghĩa dữ liệu (tên giống nhau nhưng dữ liệu khác nhau), định dạng dữ liệu (tên và ý nghĩa giống nhau nhưng kiểu dữ liệu khác nhau).
- **Tính hướng chủ đề bất biến (Nonvolatile):** quy định rằng dữ liệu phải thống nhất theo thời gian (bằng cách hạn chế tối đa sửa đổi hoặc xóa dữ liệu), từ đó làm tăng quy mô dữ liệu lên đáng kể so với hệ thống nghiệp vụ (5-10 năm so với 2 đến 6 tháng như Database thông thường).

#### 4. Cấu trúc:



Hình 4.1: Orther Data Sources

Data Warehouse có kiến trúc cơ bản rất đơn giản. Nó dựa trên nguyên tắc là xây dựng một kho dữ liệu thống nhất từ nhiều nguồn dữ liệu khác nhau để phục vụ truy vấn gồm 3 lớp chính:

- **Lớp dữ liệu nguồn (Data Sources):** Data Warehouse tổng hợp dữ liệu từ rất nhiều nguồn khác nhau. Điều này thể hiện đặc tính tích hợp của Data Warehouse. Dữ liệu nguồn có thể là dữ liệu có sẵn trong hệ thống tác nghiệp của doanh nghiệp và cũng có thể là dữ liệu lấy từ nguồn bên ngoài doanh nghiệp. Dữ liệu có thể là loại có cấu trúc chặt chẽ như dữ liệu quan hệ đã được chuẩn hoá hay có thể là loại phi cấu trúc như các văn bản thông thường.
- **Lớp Warehouse:** Lớp này chứa đựng dữ liệu đã được tổng hợp cùng siêu dữ liệu mô tả chúng cũng như các tiến trình tổng hợp, phân bổ dữ liệu. Hai thành phần quan trọng nhất của lớp Warehouse là dữ liệu và siêu dữ liệu. Chúng phải được lưu giữ và mô tả nhất quán về nội dung (ý nghĩa dữ liệu) cũng như hình thức (khuôn dạng dữ liệu). Để tăng tốc cho việc đáp ứng trả lời truy vấn thì dữ liệu trong Data Warehouse thường được để sẵn dưới dạng tổng kết.
- **Lớp Users:** có nhiệm vụ tương tác với người dùng cuối. Data Warehouse có đặc điểm quan trọng là cung cấp thông tin cho rất nhiều người sử dụng với những yêu cầu không thể dự đoán trước, vì vậy nhìn chung lớp người dùng có cấu trúc phức tạp. Hơn

nữa, người dùng ở đây có thể sử dụng nhiều công cụ khai thác và truy xuất dữ liệu khác nhau nên giao diện giữa lớp Warehouse và người dùng cũng đa dạng.

Do tính đa dạng của lớp ứng dụng và lớp người dùng nên chúng không thể giao tiếp với Data Warehouse theo các giao diện chuẩn mà thường cần phải có những công cụ được thiết kế đặc biệt chuyên thực hiện công việc này. Dữ liệu khi được đưa vào hoặc đưa ra đòi hỏi phải có những tiến trình xử lý phức tạp. Các ứng dụng đảm nhiệm công việc này thường tạo ra các kho trung gian và phải thực thi nhiều bước chuyển tiếp.

## **5. Tiến trình xử lý**

- **Tập hợp dữ liệu đưa vào Warehouse:** ngoài việc đọc hiểu các cấu trúc dữ liệu, tiến trình này còn phải thực hiện nhiều chức năng khác để bảo đảm tính nhất quán của dữ liệu trong Warehouse.
- **Phân bổ dữ liệu đến người dùng cuối:** có nhiều công cụ để thực hiện việc này, nhưng nói chung dữ liệu thường được xử lý trước rồi mới hiển thị tới người dùng cuối.

## **6. Ưu nhược điểm của kho dữ liệu**

- **Ưu điểm của kho dữ liệu:**
  - Kho dữ liệu cho phép người dùng doanh nghiệp nhanh chóng truy cập dữ liệu quan trọng từ một số nguồn ở một nơi.
  - Kho dữ liệu cung cấp thông tin phù hợp về các hoạt động đa chức năng khác nhau. Nó cũng hỗ trợ báo cáo và truy vấn đặc biệt.
  - Kho dữ liệu giúp tích hợp nhiều nguồn dữ liệu để giám căng thẳng cho hệ thống sản xuất.
  - Kho dữ liệu giúp giảm tổng thời gian quay vòng để phân tích và báo cáo. Tái cấu trúc và tích hợp giúp người dùng dễ sử dụng hơn để báo cáo và phân tích.
  - Kho dữ liệu cho phép người dùng truy cập dữ liệu quan trọng từ số lượng nguồn ở một nơi duy nhất. Do đó, nó giúp tiết kiệm thời gian lấy dữ liệu của người dùng từ nhiều nguồn.



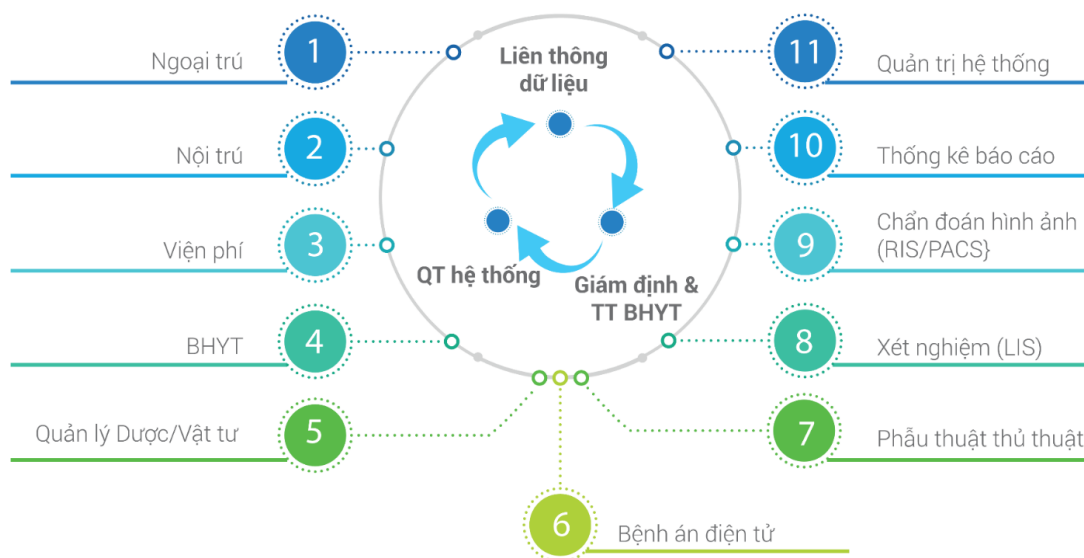
- Kho dữ liệu lưu trữ một lượng lớn dữ liệu lịch sử. Điều này giúp người dùng phân tích các khoảng thời gian và xu hướng khác nhau để đưa ra dự đoán trong tương lai.
- **Nhược điểm của Kho dữ liệu:**
  - Không phải là một lựa chọn lý tưởng cho dữ liệu phi cấu trúc.
  - Kho dữ liệu có thể bị lỗi thời tương đối nhanh. Khó thực hiện thay đổi về kiểu và phạm vi dữ liệu, lược đồ nguồn dữ liệu, chỉ mục và truy vấn.
  - Kho dữ liệu có vẻ dễ dàng, nhưng thực sự, nó quá phức tạp đối với người dùng trung bình. Mặc dù có những nỗ lực tốt nhất trong quản lý dự án, phạm vi dự án kho dữ liệu sẽ luôn tăng.
  - Đôi khi người dùng kho sẽ phát triển các quy tắc kinh doanh khác nhau. Các tổ chức cần dành nhiều nguồn lực cho mục đích đào tạo và thực hiện. Tương lai của kho dữ liệu
  - Thay đổi các ràng buộc quy định có thể hạn chế khả năng kết hợp nguồn dữ liệu khác nhau. Những nguồn khác nhau này có thể bao gồm dữ liệu phi cấu trúc rất khó lưu trữ.
  - Khi kích thước của cơ sở dữ liệu tăng lên, các ước tính về những gì tạo nên một cơ sở dữ liệu rất lớn tiếp tục phát triển. Việc xây dựng và chạy các hệ thống kho dữ liệu luôn tăng kích thước là rất phức tạp.
  - Các tài nguyên phần cứng và phần mềm có sẵn ngày hôm nay không cho phép giữ một lượng lớn dữ liệu trực tuyến.
  - Dữ liệu đa phương tiện không thể dễ dàng thao tác dưới dạng dữ liệu văn bản, trong khi thông tin văn bản có thể được truy xuất bằng phần mềm quan hệ hiện có.

## II. TÌM HIỂU VÀ PHÂN TÍCH NGHIỆP VỤ ( TRONG Y TẾ)

### 1. Vì sao cần sử dụng kho dữ liệu vào y tế?

- **Chính xác và nhất quán trong dữ liệu:** các thông tin chứa trong cơ sở dữ liệu phải được chính xác và nhất quán nhằm đảm bảo thu được kết quả chính xác và đáng tin cậy nhất cho người dùng
- **Quyền riêng tư và bảo mật dữ liệu:** thông tin dữ liệu, bệnh án của bệnh nhân được được bảo đảm, tôn trọng quyền riêng tư của mỗi người. Bảo mật, an ninh, tránh bị xâm phạm bởi người xấu
- **Lưu trữ thông tin:** đảm bảo hồ sơ, bệnh án của bệnh nhân, không bị mất, hư hại theo thời gian . Thông tin được bảo mật tại mọi thời điểm
- **Chuẩn hóa mã chẩn đoán, mã quy trình can thiệp, chữ viết tắt, kí hiệu, định nghĩa dùng trong bệnh viện:** Việc chuẩn hóa thuật ngữ, định nghĩa, từ vựng và danh pháp giúp đảm bảo sự tương thích dữ liệu và thông tin trong nội bộ bệnh viện và với các bệnh viện khác. Ngoài ra, chuẩn hóa giúp ngăn ngừa việc hiểu sai thông tin và các sai sót có thể xảy ra. Việc thống nhất mã chẩn đoán và mã quy trình can thiệp tạo điều kiện cho việc tổng hợp và phân tích dữ liệu.
- **Ứng dụng công nghệ thông tin vào trong quản lý:** được kiểm tra, kiểm thử trước khi áp dụng vào bệnh viện và bệnh nhân, đảm bảo được chất lượng an toàn sử dụng cho người bệnh

## 2. Quy trình khám và chữa bệnh



Hình 2.1: Quy trình khám và chữa bệnh

## 3. Phân tích nghiệp vụ khám chữa bệnh trong y tế:

Phân loại bệnh nhân theo:

- Khu vực bệnh nhân sinh sống
- Giới tính của bệnh nhân
- Độ tuổi của bệnh nhân
- Tiền sử khám và chữa bệnh của bệnh nhân

### III. THIẾT KẾ VÀ CÀI ĐẶT: DATA WAREHOUSE, MÔ HÌNH NẠP DỮ LIỆU TỪ HỆ THỐNG QUẢN LÝ VÀO DATAWAREHOUSE.

\*BHYTe

Thuộc tính	Kiểu giá trị	Giải thích	Ghi chú
SotheBH	Char	Số thẻ bảo hiểm	Khóa chính
TenChuThe	nvarchar	Tên chủ thẻ	

\*Khoa

Thuộc tính	Kiểu giá trị	Giải thích	Ghi chú
MaKhoa	char	Mã khoa	Khóa chính
TenKhoa	nvarchar	Tên khoa	

\*Thuoc

Thuộc Tính	Kiểu Giá Trị	Giải thích	Ghi chú
TenThuoc	nvarchar	Tên thuốc	Khóa chính
DVT	Nvarchar	Đơn vị tính	
DonGia	float	Đơn giá	

\*BenhNhan

Thuộc Tính	Kiểu Giá Trị	Giải thích	Ghi chú
MaBN	Char	Mã bệnh nhân	Khóa chính
TenBN	nvarchar	Tên bệnh nhân	
GioiTinh	nvarchar	Giới tính	
DiaChi	nvarchar	Địa chỉ	
SoTheBH	nvarchar	Số thẻ bảo hiểm	

\*BenhLy

Thuộc tính	Kiểu giá trị	Giải thích	Ghi chú
MaBL	char	Mã bệnh lý	Khóa chính

TenBL	nvarchar	Tên bệnh lý	
MaKhoa	char	Mã khoa	

**\*NhanVien**

Thuộc tính	Kiểu giá trị	Giải thích	Ghi chú
MaNV	Char	Mã nhân viên	Khóa chính
TenNV	nvarchar	Tên nhân viên	
MaKhoa	Char	Mã Khoa	
KVCongTac	nvarchar	Khu vực công tác	

**\*HoaDon**

Thuộc tính	Kiểu giá trị	Giải thích	Ghi chú
MaHD	Char	Mã hóa đơn	Khóa chính
NgayLap	Date	Ngày lập	
MaBN	Char	Mã bệnh nhân	
MaPK	Char	Mã phiếu khám	

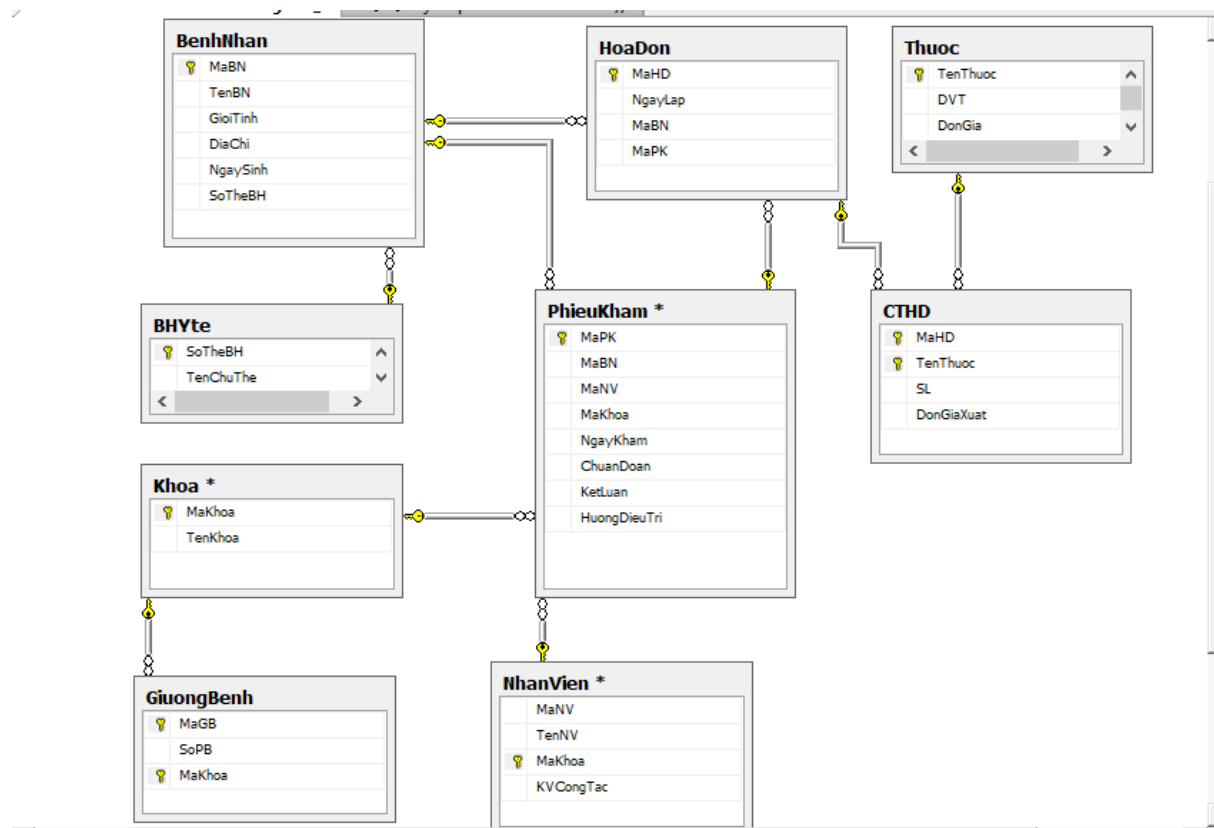
**\*PhieuKham**

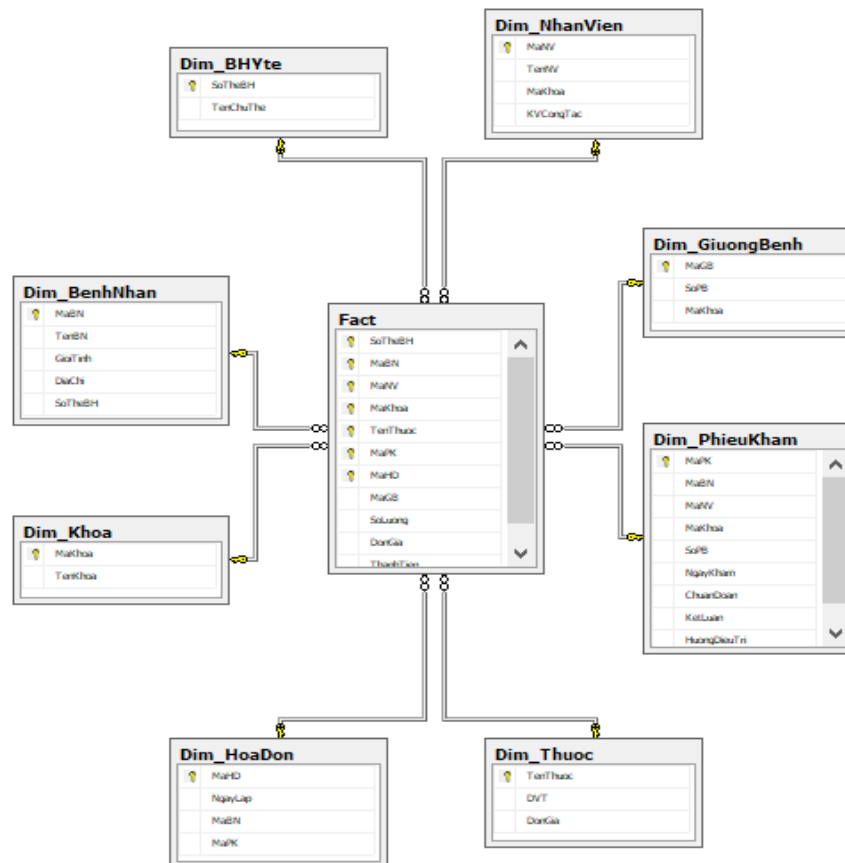
Thuộc tính	Kiểu giá trị	Giải thích	Ghi chú
MaPK	Char		Khóa chính
MaBN	Char	Mã bệnh nhân	
MaNV	char	Mã nhân viên	
MaKhoa	char	Mã khoa	
SoPB	Char	Số phòng bệnh	
NgayKham	Date	Ngày khám	
ChuanDoan	nvarchar	Chuẩn đoán	
KetLuon	nvarchar	Kết luận	
HuongDieuTri	nvarchar	Hướng điều trị	

\*CTHD

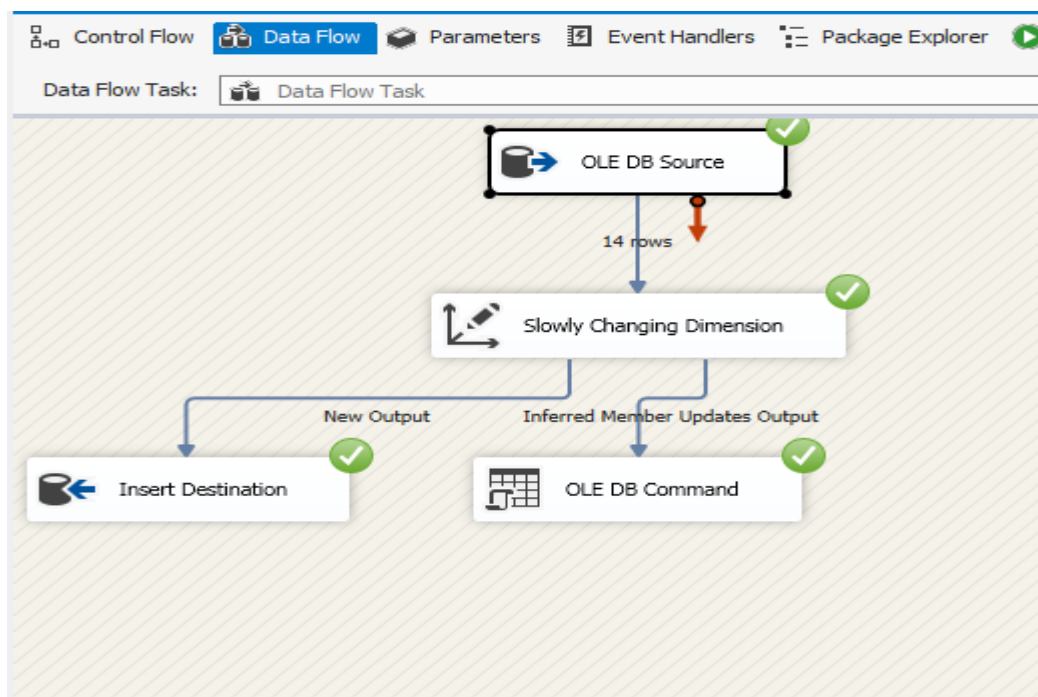
Thuộc tính	Kiểu giá trị	Giải thích	Ghi chú
MaHD	Char	Mã hóa đơn	Khóa chính
TenThuoc	Nvarchar	Tên thuốc	Khóa chính
SL	Int	Số lượng	
DonGiaXuat	Float	Đơn giá xuất	

## CSDLSQL:





Load từ cơ sở dữ liệu bảng dim

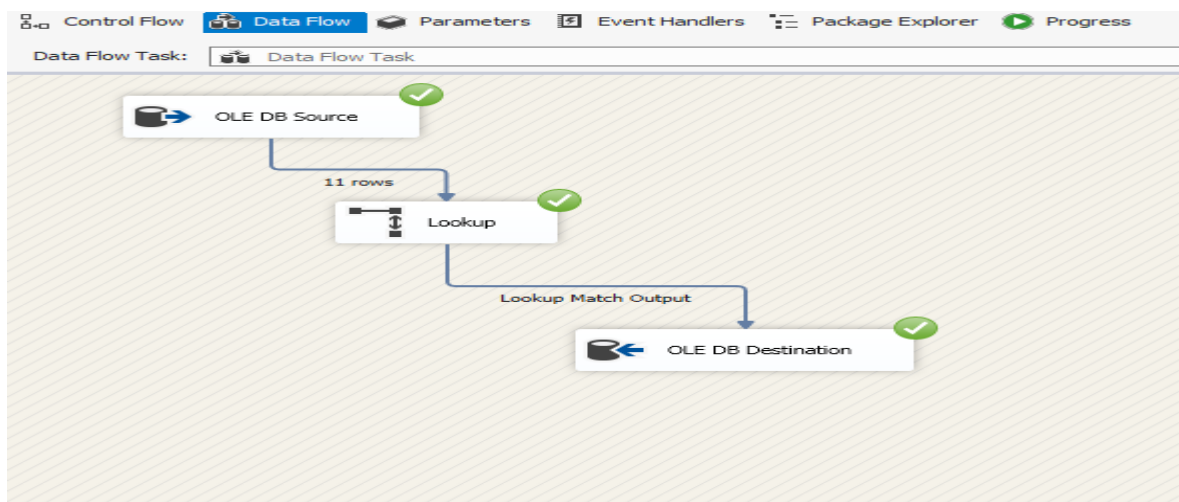




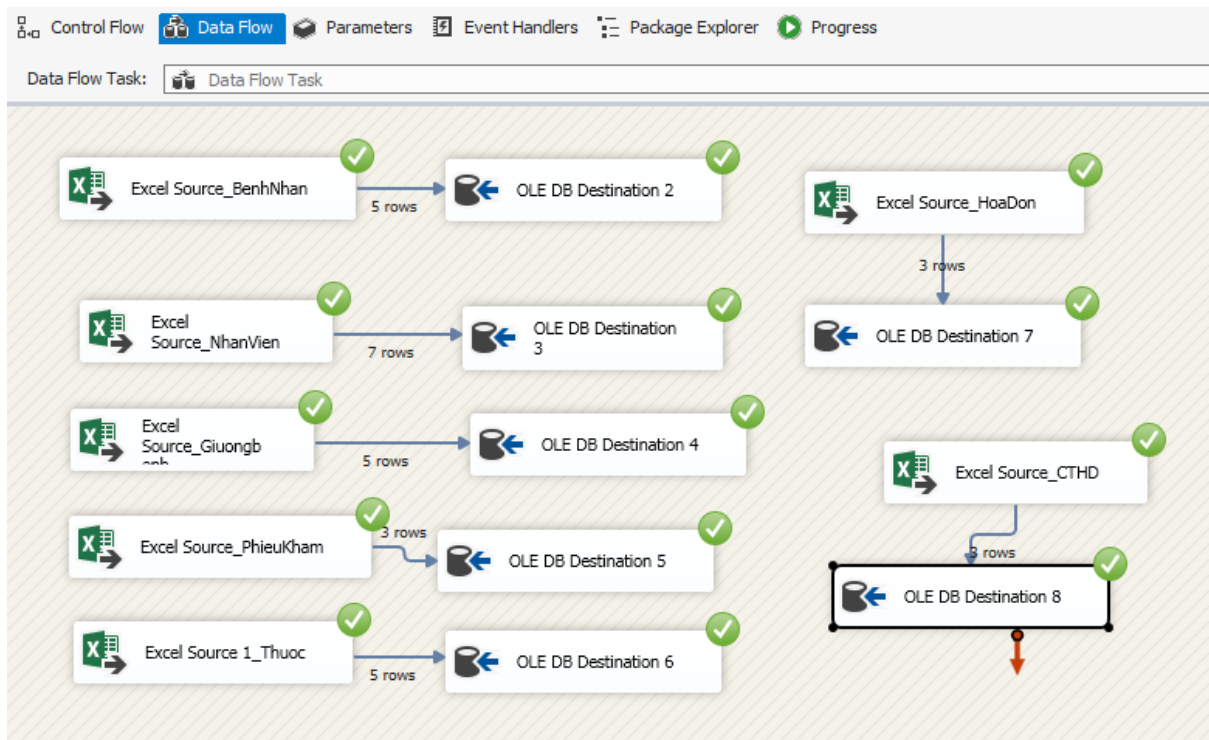
Dữ liệu được insert vào bảng Dim\_xong khi chạy thành công

	MaHD	NgayLap	MaBN	MaPK
▶	HD001	2020-05-25	BN001	PK1
	HD002	2020-05-25	BN003	PK3
	HD003	2020-05-26	BN004	PK2
	HD004	2020-05-26	BN006	PK4
	HD005	2020-05-27	BN002	PK5
	HD006	2020-05-27	BN010	PK6
*	NULL	NULL	NULL	NULL

Load dữ liệu bảng Fact

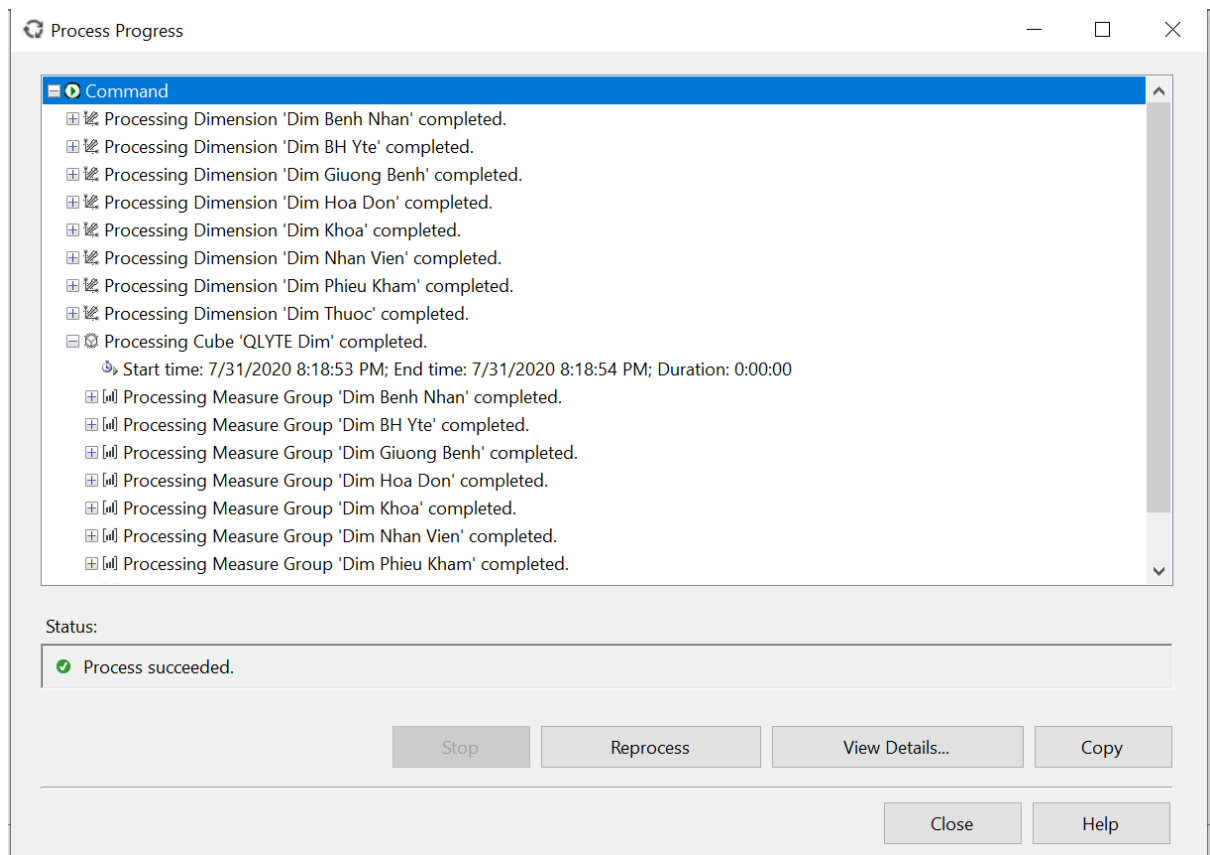
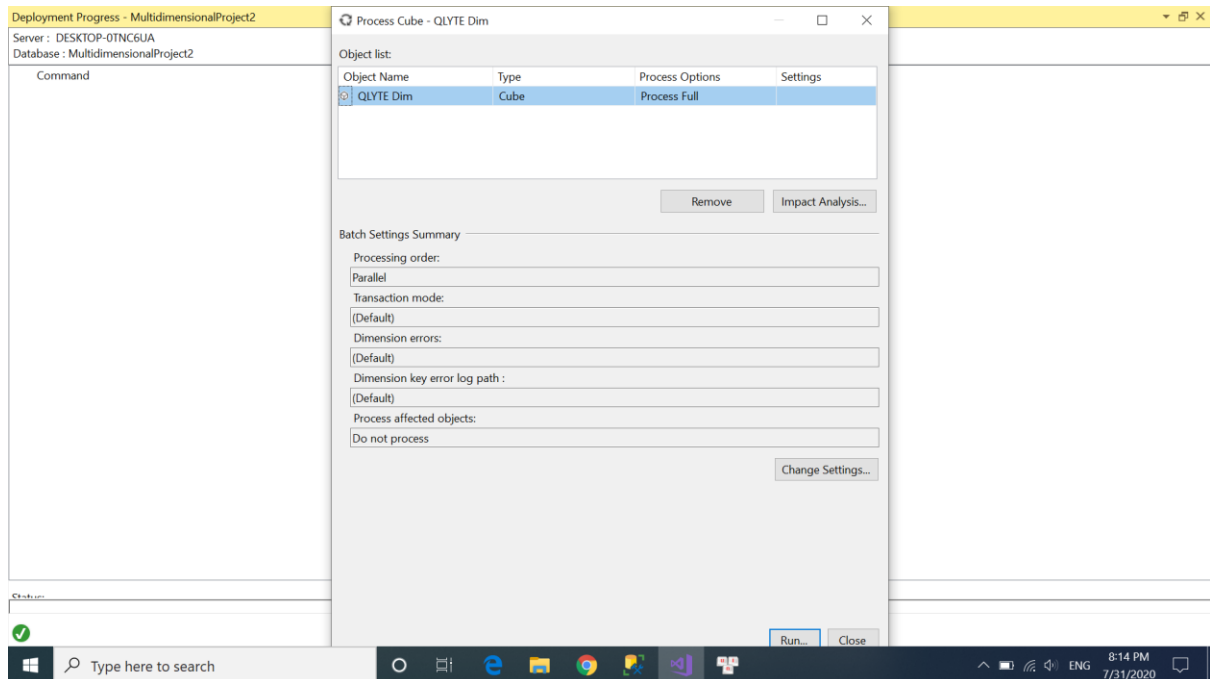


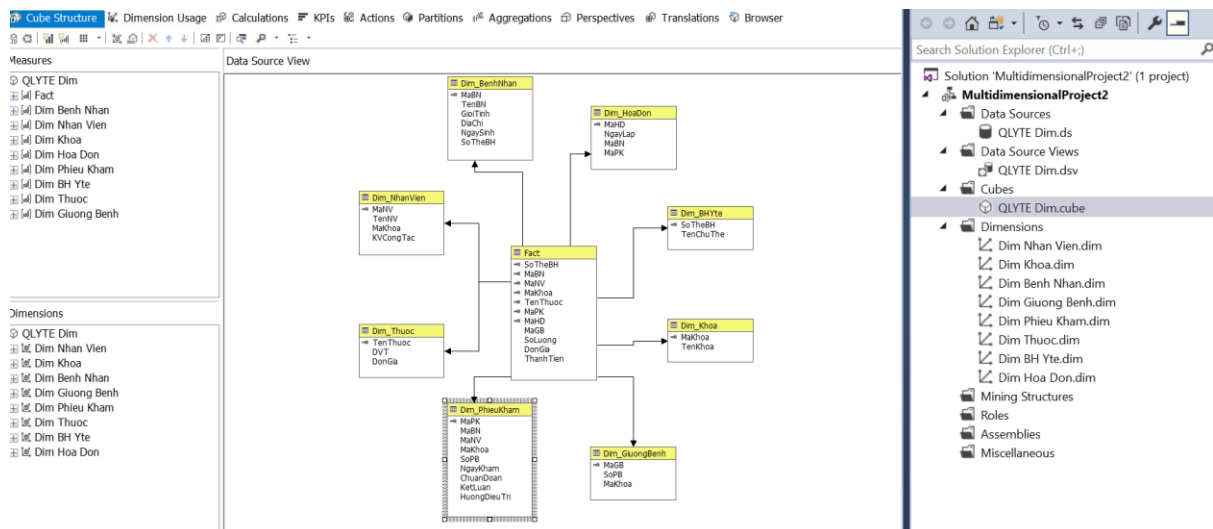
--Load dữ liệu từ excel



## IV CUBE

### 1: CÀI ĐẶT





## 2 TRUY VẤN DỮ LIỆU

Dimension	Hierarchy	Operator	Filter Expression
<Select dimension>			
Ten Thuoc	Don Gia - Dim Thuoc	Dim Thuoc Count	
SUNCUR...	22000	1	
Aspirin	22000	1	
Augmentin	20000	1	
Bocalex C...	21000	1	
Chlorpro...	20000	1	
Diamicro...	20000	1	
Ether	20000	1	
Fellaini 2...	18000	1	
Insulin	20000	1	
Klenzit - C	20000	1	
Klenzit MS	18000	1	
Morphin	18000	1	
Penicillin	20000	1	
Timolol ...	16000	1	

SELECT { [Measures].[SL] } ON COLUMNS,  
NON EMPTY { ([Dim Time].[Quy Ban].[Quy Ban].ALLMEMBERS ) }  
ON ROWS FROM [QL YTE Dim]

00 %

Messages Results

SL

2 45

SELECT { [Measures].[Dim BH Yte Count] } ON COLUMNS FROM [QLYTE Dim] ;

100 %

Messages Results

Dim BH Yte Count
8

SELECT { [Measures].[Dim Nhan Vien Count] } ON COLUMNS FROM [QLYTE Dim] ;

100 %

Messages Results

Dim Nhan Vien Count
7

## IV. ỨNG DỤNG:

frmMain

DS Khoa  
DS BHYT  
DS Thuốc  
DS Bệnh nhân  
DS Nhân Viên  
DS Hóa Đơn  
DS Giường Bệnh  
DS Phiếu Khám  
Chi Tiết Hóa Đơn  
Chuyển sang kho di

**Số Thẻ BHYT** BH01

**Tên Chủ Thẻ** Nguyễn Thanh An

THÊM SỬA XÓA LƯU

Số Thẻ BHYT	Tên Chủ Thẻ
BH01	Nguyễn Thanh An
BH02	Nguyễn Thanh
BH03	Trần Đại Long
BH04	Hoàng Thùy Chi
BH05	Huỳnh Nhân
BH06	Trần Minh Đức
BH07	Đinh Linh
BH08	Huỳnh Chí Khanh

frmMain

DS Khoa  
DS BHYT  
DS Thuốc  
DS Bệnh nhân  
DS Nhân Viên  
DS Hóa Đơn  
DS Giường Bệnh  
DS Phiếu Khám  
Chi Tiết Hóa Đơn  
Chuyển sang kho di

**Mã Khoa** CTCH

**Tên Khoa** Chấn Thương chỉnh hình

THÊM SỬA XÓA LƯU

Mã Khoa	Tên Khoa
CTCH	Chấn Thương chỉnh hình
DL	Bệnh ngoài da
HH	khoa Hô Hấp
HS	Khoa Hồi Súc
NS	Khoa Nội Soi
RHM	Khoa Răng Hàm Mặt
TK	Khoa Thần Kinh

frmMain

DS Khoa  
DS BHYT  
**DS Thuốc**  
DS Bệnh nhân  
DS Nhân Viên  
DS Hóa Đơn  
DS Giường Bệnh  
DS Phiếu Khám  
Chi Tiết Hóa Đơn  
Chuyển sang kho d

Tên Thuốc  
DVT  
Đơn Giá  
THÊM SỬA XÓA LƯU

Tên Thuốc	Đơn Vị Tính	Đơn Giá
SUNCURMIN	Lọ	22000
Aspirin	Viên	22000
Augmentin	Gói	20000
Bocalex C 1000	Viên	21000
Chlorpromazin	Viên	20000
Diamicon MR	Viên	20000
Ether	Viên	20000
Fellaini 25mg Medisun	Viên	18000



## TÀI LIỆU THAM KHẢO

<http://qpsolutions.vn/newsdetail.asp?newsID=102&cat1id=7&cat2id=16&title=quan-ly-thong-tin-trong-benh-vien>

<https://www.youtube.com/watch?v=PHckRz8vpuM&fbclid=IwAR0if7HP17H9mc8H3-Jg5ykORU1S0JKdZmd3-LXelrzQgxeEIDEJ2Pc5NXE>