**BỘ MÔN HỆ THỐNG THÔNG TIN – KHOA CÔNG NGHỆ THÔNG TIN**

**ĐẠI HỌC KHOA HỌC TỰ NHIÊN, ĐẠI HỌC QUỐC GIA TP HCM**

Nhóm sinh viên thực hiện:

19120559 Hà Duy Lãm

19120545 Lê Ngọc Khoa

19120677 Nguyễn Diệp Minh Tiến

Giáo viên phụ trách: Hồ Thị Hoàng Vy

Giáo viên thực hành: Tiết Gia Hồng

ĐỒ ÁN MÔN HỌC  - HỆ THỐNG THÔNG TIN PHỤC VỤ TRÍ TUỆ KINH DOANH

HỌC KỲ I – NĂM HỌC 2022-2023

**MÔN HỆ THỐNG THÔNG TIN PHỤC VỤ TRÍ TUỆ KINH DOANH**

**BẢNG PHÂN CÔNG CHI TIẾT**

|  |  |  |
| --- | --- | --- |
| **MSSV** | **Họ tên** | **Phân công công việc** |
| 19120559 | Hà Duy Lãm | * Thực hiện thiết kế DDS, NDS, Stage * Thực hiện MDX * Báo cáo * Data mining |
| 19120545 | Lê Ngọc Khoa | * ETL dữ liệu từ Stage vào NDS * ETL dữ liệu từ NDS vào DDS * Thực hiện refresh data |
| 19120677 | Nguyễn Diệp Minh Tiến | * Thực hiện thiết kế DDS, NDS, Stage * Visualize dữ liệu * Thực hiện OLAP |

**Mục lục**

[**I.** **THÔNG TIN ĐỒ ÁN VÀ LINK KẾT QUẢ:** 3](#_Toc123240962)

[1. Yêu cầu đồ án: 3](#_Toc123240963)

[**II.** **MÔ TẢ DỮ LIỆU:** 4](#_Toc123240964)

[1. Bảng Compiled\_COVID-19\_Case\_Deatails: 4](#_Toc123240965)

[2. Bảng Cases Report: 5](#_Toc123240966)

[3. Bảng Vaccines\_by\_age\_phu: 6](#_Toc123240967)

[4. Bảng Public health unit: 7](#_Toc123240968)

[5. Bảng ongoing\_outbreadks\_phu: 8](#_Toc123240969)

[6. Bảng Public Health Units GROUP: 8](#_Toc123240970)

[7. Bảng METADATA: 9](#_Toc123240971)

[III. **THIẾT KẾ STAGE:** 9](#_Toc123240972)

[IV. **THIẾT KẾ NDS:** 10](#_Toc123240973)

[**V.** **THIẾT KẾ DDS:** 11](#_Toc123240974)

# **THÔNG TIN ĐỒ ÁN VÀ LINK KẾT QUẢ:**

Link drive của nhóm: https://drive.google.com/drive/folders/11kQJ6yOkRNACNLVR0ZRsAitpwPNQIig-?usp=share\_link

Link GitHub của nhóm: <https://github.com/hduylam31/TTKD---CQ2019.git>

## Yêu cầu đồ án:

Thiết kế kho dữ liệu (KDL), tổng hợp, nạp dữ liệu các nguồn vào KDL và thiết kế, xây dựng Cube:

Gợi ý:

* Mapping các nguồn dữ liệu trên và đề xuất giải pháp xây dựng Geography dimension với phân cấp: City > PHU\_Group > PHU
* Chuyển đổi dữ liệu ngày tháng sao cho có thể tạo được Date dimension với phân cấp chiều: Year > Quarter > Month > Day
* Xác định và thiết kế các phân cấp chiều khác để đáp ứng yêu cầu OLAP và report - OLAP và Report.

1. Thống kê Số ca nhiễm, số ca tử vong, số ca phục hồi của dịch Covid-19 theo từng PHU trong từng năm.

2. Thống kê Mức Độ Nghiêm Trọng (tiêu chí nghiêm trọng sinh viên tự định nghĩa) của dịch Covid-19 theo PHU và theo các Quý trong từng năm.

3. Thống kê tổng số người tử vong theo Giới Tính và Nhóm Tuổi theo các năm.

4. Thống kê số ca nhiễm, tử vong theo Mức Độ Nghiêm Trọng theo Ngày Trong Tháng của các năm.

5. Thống kê số ca nhiễm, tử vong theo Mức Độ Nghiêm Trọng, khu vực (PHU\_Group, City), và số người đã được tiêm vaccin trong các năm.

6. Thống kê số ca nhiễm theo Mức Độ Nghiêm Trọng, nhóm bùng phát của từng khu vực trong các năm

7. Sinh viên tự thiết kế những bảng thống kê khác để có thêm nhiều chiều đánh giá số ca nhiễm và tử vong ở Ontario.

8. Xây dựng đồ thị/ biểu đồ cho các bảng thống kê ở trên.

9. [Data Visualization] Dùng regional map để biểu diễn trực quan (bằng màu sắc) số lượng ca nhiễm, số ca tử vong ở các vùng trong năm. Ví dụ tham khảo:

# **MÔ TẢ DỮ LIỆU:**

## Bảng Compiled\_COVID-19\_Case\_Deatails:

Dữ liệu ca nhiễm của tất cả các tỉnh bang ở Canada.

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Tên thuộc tính** | **Kiểu dữ liệu** | **Ý nghĩa** |
| 1 | ObjectID | int | Khóa tự tăng |
| 2 | rowID | int | Mã dòng |
| 3 | Date\_report | datetime | Mức độ nghiêm trọng của tai nạn  (1: Fatal, 2: Serious, 3: Slight) |
| 4 | Health\_region | Varchar(100) | Số lượng phương tiện |
| 5 | Age\_group | Varchar(100) | Nhóm tuổi:  - 5 to 11 years old  - 12 to 17 year olds  - 18 to 29 years old  - 30 to 39 years old  - 40 to 49 years old  - 50 to 59 years old  - 60 to 69 years old  - 70 to 79 years old  - 80 years and older  - Adults\_18plus  - Ontario\_12plus  - Ontario\_5plus  - Undisclosed\_or\_missing |
| 6 | Gender | Varchar(50) | Giới tính của bệnh nhân |
| 7 | exposure | Varchar(50) | Phơi nhiễm:  - Outbreak - Bùng phát  - Close Contact - Liên hệ chặt chẽ  - Not Reported - Không được báo cáo  - Travel-Related - Du lịch |
| 8 | Case\_status | Varchar(50) | Trạng thái của ca nhiễm  - Recovered - Phục hồi  - Deceased - Tử vong  - Active - Điều trị tích cực |
| 9 | province | Varchar(50) | Vùng diễn ra dịch bệnh |

## Bảng Cases Report:

Dữ liệu ca nhiễm của tỉnh bang Ontario

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Tên thuộc tính** | **Kiểu dữ liệu** | **Ý nghĩa** |
| 1 | ObjectID | int | Khóa tự tăng |
| 2 | Outcome | varchar(100) | Kết quả:  - Resolved - Điều trị  - Fatal - Tử vong |
| 3 | Age | varchar(100) | Tham chiếu từ Compiled\_COVID-19\_Case\_Deatails |
| 4 | Gender | varchar(100) | Giới tính của bệnh nhân |
| 5 | Reporting\_PHU | varchar(100) | Tham chiếu từ Public health unit, Compiled\_COVID-19\_Case\_Deatails |
| 6 | SpecimenDate | Datetime | Ngày cập nhật |
| 7 | PHU\_City | varchar(100) | Tham chiếu từ Public health unit |
| 8 | TestReportedDate | Datetime | Ngày trả kết quả |
| 9 | AccurateEpisodeDt | varchar(50) | Ngày khởi phát |
| 10 | CaseAccquistitionInfo | varchar(100) | Thông tin ca nhiễm:  - CC: dương tính xác định được nguồn lây (closed contact)  - No known Epi-link: dương tính không rõ nguồn lây  - OB: bùng phát (Outbreak) - Travel |
| 11 | PHU\_Address | varchar(100) | Tham chiếu từ Public health unit |
| 12 | PHU\_Website | varchar(100) | Tham chiếu từ Public health unit |
| 13 | OutbreakRelated | varchar(100) | Có liên quan đến đợt bùng phát |
| 14 | PHU\_Latitude | char(20) | Tham chiếu từ Public health unit |
| 15 | PHU\_Longitude | char(20) | Tham chiếu từ Public health unit |
| 16 | PHU\_Postal\_Code | varchar(10) | Tham chiếu từ Public health unit |
| 17 | CaseReport\_Date | Datetime | Tham chiếu từ Compiled\_COVID-19\_Case\_Deatails |

## Bảng Vaccines\_by\_age\_phu:

Dữ liệu tiêm vắc-xin tại các đơn vị chăm sóc sức khỏe của Ontario

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Tên thuộc tính** | **Kiểu dữ liệu** | **Ý nghĩa** |
| 1 | Date | datetime | Ngày tiêm |
| 2 | Agegroup | Varchar(50) | Mã nhóm tuổi |
| 3 | PhuID | int | Tham chiếu từ Public health unit |
| 4 | AtLeastOneDoesCumulative | int | Số người tiêm được ít nhất 1 mũi |
| 5 | Second\_dose\_cumulative | int | Số người tiêm được 1 mũi |
| 6 | Fully\_vaccinated\_cumulative | int | Số người tiêm đủ vaccin. Tiêm đầy đủ nghĩa là:  - Tiêm 1 mũi Janssen (Johnson & Johnson)  - Tiêm 2 mũi trong danh mục vaccin được Bộ y tế Canada phê duyệt  - Tiêm 1 mũi trong danh mục được Bộ ý tế phê duyệt + 1 mũi trong danh mục không được phê duyệt  - Tiêm 3 mũi vaccin thuộc loại bất kỳ |

## Bảng Public health unit:

Dữ liệu các đơn vị chăm sóc sức khỏe của Ontario

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Tên thuộc tính** | **Kiểu dữ liệu** | **Ý nghĩa** |
| 1 | PHU\_ID | Int | Mã của các PHU |
| 2 | Reporting\_PHU\_Address | Varchar(100) | Địa chỉ PHU được báo cáo |
| 3 | Reporting\_PHU\_City | Varchar(30) | Thành phố của các PHU được báo cáo |
| 4 | Reporting\_PHU\_Postal\_Code | Varchar(100) | Mã bưu điện của PHU được báo cáo |
| 5 | Reporting\_PHU\_Website | Varchar(10) | Địa chỉ website của PHU |
| 6 | Reporting\_PHU\_Latitude | Char(20) | Vĩ tuyến PHU |
| 7 | Reporting\_PHU\_Longitude | Char(20) | Kinh tuyến PHU |
| 8 | Reporting\_PHU | Varchar(100) | Địa chỉ PHU được báo cáo |

## Bảng ongoing\_outbreadks\_phu:

Dữ liệu về việc bùng phát dịch tại các đơn vị chăm sóc sức khỏe của Ontario

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Tên thuộc tính** | **Kiểu dữ liệu** | **Ý nghĩa** |
| 1 | Date | int | Ngày bùng phát dịch |
| 2 | Phu\_num | int | Tham chiếu từ Public health unit |
| 3 | Outbreak\_group | Varchar(50) | Cơ sở bùng phát dịch:  - 1 Congregate Care - Chăm sóc cộng đồng  - 2 Congregate Living - Lưu trú cộng đồng  - 3 Education - Giáo dục  - 4 Workplace - Nơi làm việc  - 5 Recreational - Cơ sở giải trí  - 6 Other/Unknown - Không xác định |
| 4 | Number\_ongoing\_outbreaks | int | Số đợt bùng phát đang diễn ra |

## Bảng Public Health Units GROUP:

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Tên thuộc tính** | **Kiểu dữ liệu** | **Ý nghĩa** |
| 1 | PHU\_Group | Varchar(50) | Nhóm các PHU |
| 2 | PHU\_City | Varchar(30) | Thành phố của các PHU được báo cáo. |
| 3 | PHU\_Region | Varchar(100) | Tham chiếu từ Public health unit |

## Bảng METADATA:

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Tên thuộc tính** | **Kiểu dữ liệu** | **Ý nghĩa** |
| 1 | [index] | int | Khóa tự tăng |
| 2 | CreatedDate | Datetime | Ngày thêm |
| 3 | UpdatedDate | Datetime | Ngày cập nhật |

# **THIẾT KẾ STAGE:**

Diagram, schematic

Description automatically generated

# **THIẾT KẾ NDS:** Diagram Description automatically generated

# **THIẾT KẾ DDS:**

Diagram, engineering drawing

Description automatically generated

# 