HỌC VIỆN NGÂN HÀNG

KHOA CÔNG NGHỆ THÔNG TIN VÀ KINH TẾ SỐ

BÁO CÁO BÀI TẬP LỚN

MÔN KHO DỮ LIỆU VÀ KINH DOANH THÔNG MINH

ĐỀ TÀI: NGHIÊN CỨU VÀ XÂY DỰNG KHO DỮ LIỆU CHO TẬP ĐOÀN PFIZER

Giảng viên hướng dẫn : TS. Đinh Trọng Hiếu

Nhóm sinh viên thực hiện : Nhóm 11

Hà Nội – 12/2024

HỌC VIỆN NGÂN HÀNG

KHOA CÔNG NGHỆ THÔNG TIN VÀ KINH TẾ SỐ

BÁO CÁO BÀI TẬP LỚN

MÔN KHO DỮ LIỆU VÀ KINH DOANH THÔNG MINH

ĐỀ TÀI: NGHIÊN CỨU VÀ XÂY DỰNG KHO DỮ LIỆU CHO TẬP ĐOÀN PFIZER

Giảng viên hướng dẫn : **TS. Đinh Trọng Hiếu**

Sinh viên thực hiện

Phạm Thị Thu Chan - **24A4043034**

Phạm Thị Hồng Duyên - **24A4043038**

Hoàng Phương Thanh - **24A4041437**

Nguyễn Tiến Mạnh - **24A4040033**

Hà Nội – 12/2024

BẢNG PHÂN CÔNG CHI TIẾT CÔNG VIỆC

STT	Mã sinh viên	Họ và tên	Phân công	Tỉ lệ đóng góp (%)
1	24A4043034	Phạm Thị Thu Chan	- Nội dung Chương III: Hiện Trạng Hệ Thống Thông Tin và Dữ Liệu của Doanh nghiệp Pfizer - Nội dung Chương VII: Đánh Giá và Khuyến Nghị	25%
2	24A4043038	Phạm Thị Hồng Duyên	- Nội dung Chương I: Giới Thiệu Tổng Quan Doanh Nghiệp - Nội dung Chương II: Người Dùng Hệ Thống BI (BI Users)	25%
3	24A4040033	Nguyễn Tiến Mạnh	- Nội dung Chương IV: Thiết Kế và Cài Đặt Kho Dữ Liệu - Nội dung Chương V: Mô Hình Quan Hệ Giữa Nhu Cầu Phân Tích và Khung Nhìn Dữ Liệu - Nội dung Chương VI: Đánh Giá Nền Tảng BI và Trực Quan Hóa Dữ Liệu	25%
4	24A4041437	Hoàng Phương Thanh	- Nội dung Chương II: Người Dùng Hệ Thống BI (BI Users) - Chương V: Mô Hình Quan Hệ Giữa Nhu Cầu Phân Tích và Khung Nhìn Dữ Liệu	25%

DANH MỤC HÌNH ẢNH

Hình 1. Sơ đồ tổ chức 5

Hình 2. Kiến trúc Kho dữ liệu của Kimball 14

Hình 3. Bảng DimCountry 15

Hình 4. Bảng DimDate 16

Hình 5. Bảng DimContinent 16

Hình 6. Bảng WHO_Region Dimension 17

Hình 7. Bảng FactCovid 17

Hình 8. Mô hình trên Data Mart 20

Hình 9. Kết quả báo cáo Time Trend Analysis 37

Hình 10. Biểu đồ Bar Chart 39

Hình 11. Biểu đồ Donut Chart 40

Hình 12. Power BI 42

Hình 13. Tổng quan dashboard 45

DANH MỤC BẢNG BIỂU

Bảng 1. So sáng các công cụ trực quan.....43

Bảng 2. Bảng chi tiết các biểu đồ.....45

Bảng 3. Đánh giá tính trực quan dữ liệu trên PowerBI.....47

MỤC LỤC

LỜI MỞ ĐẦU 1

Chương I: Giới Thiệu Tổng Quan Doanh Nghiệp 2

1.1 Mô tả hoạt động doanh nghiệp 2

1.2 Vai trò của hệ thống báo cáo tình hình kinh doanh 2

Chương II: Người Dùng Hệ Thống BI (BI Users) 5

2.1 Phân tích cơ cấu tổ chức của Pfizer 5

2.2 Các nhóm BI Users và đặc điểm 7

2.2.1 Nhóm cấp chiến lược 7

2.2.2 Nhóm cấp chiến thuật 7

2.2.3 Nhóm cấp tác nghiệp 8

2.3 Quyết định kinh doanh của BI Users 8

2.3.1. Quyết định của Ban Điều Hành (Executive Management) 9

2.3.2. Quyết định của Giám Đốc Tài Chính (CFO) 9

2.3.3. Quyết định của Giám Đốc Kinh Doanh 9

2.3.4. Quyết định của Phòng Phân Tích Dữ Liệu (Data Analytics Department): 9

2.3.5. Quyết định của Phòng Kế Hoạch (Planning Department) 9

2.3.6. Quyết định của Phòng Quản Lý Chuỗi Cung Ứng (Supply Chain Management Department) 10

Chương III: Khảo sát Hiện Trạng Hệ Thống Thông Tin và Dữ Liệu của Doanh nghiệp Pfizer 11

3.1 Hiện trạng hệ thống thông tin 11

3.2 Hiện trạng về dữ liệu và quản trị dữ liệu 11

3.2.1. Dữ liệu đang được lưu trữ 11

3.2.2. Quản trị dữ liệu 12

Chương IV: Thiết Kế và Cài Đặt Kho Dữ Liệu 13

4.1. Sự cần thiết của việc xây dựng Kho dữ liệu trong bối cảnh mục tiêu kinh doanh của Pfizer 13

4.2. Thiết kế các bảng chiều (bảng cắt lớp) DIM 15

4.3. Thiết kế bảng FACT 17

4.3.1. Bảng Fact Covid 17
4.3.2. Lợi ích của bảng FactCovid đối với Pfizer 18
4.4. Mô hình trên Data Mart 19
4.4.1. Mô tả 19
4.4.2 Lợi ích của mô hình Data Mart theo lược đồ hình sao đối với Pfizer (tiếp tục) 20
4.5. Tổng quan quá trình đổ dữ liệu 21
4.5.1. Giới Thiệu Về Quá Trình ETL (Extract, Transform, Load) 21
4.5.2. Phân Biệt Giữa Bảng Dimension và Bảng Fact 21
4.5.3. Các Phương Pháp Phổ Biến Để Đổ Dữ Liệu vào SQL Server 22
4.6. Chuẩn bị dữ liệu và đổ dữ liệu vào bảng 22
4.6.1. Kiểm tra và Chuẩn Hóa Dữ Liệu trong File CSV 22
4.6.2. Tạo các bảng Dim 27
4.6.3. Tạo bảng Fact 28
4.6.4. Đổ dữ liệu vào bảng 29
Chương V: Mô Hình Quan Hệ Giữa Nhu Cầu Phân Tích và Khung Nhìn Dữ Liệu 33
5.1. Phân tích nhu cầu phân tích của BI Users 33
5.1.1. Ban Điều Hành (Executive Board) 33
5.1.2. Giám Đốc Tài Chính (CFO) 33
5.1.3. Giám Đốc Kinh Doanh (Sales Director) 34
5.1.4. Phòng Phân Tích Dữ Liệu (Data Analytics Department) 34
5.1.5. Phòng Kế Hoạch (Planning Department) 35
5.1.6. Quản Lý Chuỗi Cung Ứng (Supply Chain Management) 35
5.2 Khung nhìn dữ liệu (Data Views) 36
5.2.1. Báo cáo Xu Hướng Thời Gian (Time Trend Analysis) 36
5.2.2. Báo cáo Tổng số ca tử vong theo khu vực (Deaths by Regions) 37
Chương VI: Đánh Giá Nền Tảng BI và Trực Quan Hóa Dữ Liệu 41
6.1. So sánh, lựa chọn nền tảng BI phù hợp 41
6.2. Trực quan hóa dữ liệu 43
6.2.1. Tổng quan các chỉ số chính 46

6.2.3. Biểu đồ phân loại khu vực WHO 46
6.2.4. Bản đồ thế giới 46
6.3. Chính sách đảm bảo dữ liệu.....47
Chương VII: Đánh Giá và Khuyến Nghị 48
7.1. Đánh giá hệ thống báo cáo.....48
7.2. Khuyến nghị 48
KẾT LUẬN.....50
TÀI LIỆU THAM KHẢO.....52

LỜI MỞ ĐẦU

Trong bối cảnh đại dịch COVID-19, dữ liệu và hệ thống Business Intelligence (BI) trở thành những công cụ quan trọng, không chỉ đối với các tổ chức kinh doanh mà đặc biệt là trong ngành y tế và dược phẩm. Việc quản lý và phân tích dữ liệu một cách hiệu quả đã chứng minh vai trò then chốt trong việc kiểm soát dịch bệnh, tối ưu hóa chuỗi cung ứng, và hỗ trợ ra quyết định chiến lược nhằm đối phó với những thách thức mới.

Tập đoàn Pfizer, một trong những công ty dẫn đầu trong lĩnh vực dược phẩm và công nghệ sinh học, đã phải xử lý một lượng lớn dữ liệu đa dạng, bao gồm thông tin về sản xuất vaccine, phân phối thuốc, và nghiên cứu y khoa. Đại dịch đã làm nổi bật tầm quan trọng của việc xây dựng một hệ thống BI mạnh mẽ, giúp Pfizer khai thác giá trị từ dữ liệu để đưa ra các quyết định kịp thời, chính xác và hiệu quả.

Nhằm đáp ứng nhu cầu này, mục tiêu của bài tập lớn là xây dựng một kho dữ liệu và hệ thống báo cáo hỗ trợ ra quyết định cho Pfizer. Kho dữ liệu này sẽ tích hợp dữ liệu từ nhiều nguồn khác nhau, từ doanh số bán hàng, thông tin nghiên cứu, đến dữ liệu dịch tễ học. Hệ thống BI đi kèm sẽ cung cấp các báo cáo phân tích chuyên sâu, giúp Pfizer cải thiện hiệu quả kinh doanh, hiểu rõ xu hướng thị trường, và tăng cường khả năng ứng phó với những biến động trong ngành.

Thông qua việc nghiên cứu và thực hiện dự án này, bài tập không chỉ minh họa tầm quan trọng của dữ liệu và BI trong ngành y tế mà còn đóng góp vào việc phát triển các giải pháp thực tiễn nhằm nâng cao năng lực quản lý và ra quyết định chiến lược cho Pfizer.

Chương I: Giới Thiệu Tổng Quan Doanh Nghiệp

1.1 Mô tả hoạt động doanh nghiệp

Pfizer được thành lập vào năm 1849 bởi hai nhà sáng lập người Mỹ gốc Đức là Charles Pfizer và Charles Erhart, với mục tiêu ban đầu là sản xuất và cung cấp các sản phẩm hóa học và dược phẩm. Trải qua hơn 170 năm phát triển, Pfizer đã từng bước mở rộng quy mô hoạt động, từ một công ty nhỏ tại Brooklyn, New York, trở thành một tập đoàn toàn cầu hàng đầu trong lĩnh vực dược phẩm và chăm sóc sức khỏe. Trong suốt lịch sử hoạt động, Pfizer đã không ngừng đổi mới và đầu tư mạnh mẽ vào nghiên cứu và phát triển (R&D). Công ty nổi tiếng với những phát minh mang tính cách mạng trong y học, bao gồm kháng sinh Penicillin, thuốc giảm đau Advil, và Viagra.

Pfizer hiện có mặt tại hơn 100 quốc gia và sở hữu hơn 80 cơ sở sản xuất, đảm bảo nguồn cung cấp các sản phẩm y tế chất lượng cao đến mọi khu vực trên thế giới. Công ty tuyển dụng hàng trăm nghìn nhân viên, bao gồm đội ngũ chuyên gia nghiên cứu, kỹ sư và các nhà quản lý, tạo thành một mạng lưới hoạt động rộng lớn và phức tạp. Các cơ sở nghiên cứu của Pfizer được đặt tại nhiều quốc gia, tập trung vào việc phát triển các sản phẩm tiên tiến đáp ứng nhu cầu đa dạng của thị trường. Hằng năm, Pfizer đầu tư hàng tỷ USD vào hoạt động R&D, với mục tiêu mang lại các giải pháp y tế tiên tiến nhất.

Pfizer không chỉ là một nhà sản xuất thuốc và vaccine mà còn đóng vai trò quan trọng trong việc định hình xu hướng của ngành dược phẩm toàn cầu. Những nỗ lực trong nghiên cứu y học và đổi mới công nghệ đã giúp Pfizer trở thành đối tác chiến lược của nhiều tổ chức y tế và chính phủ trên thế giới. Trong bối cảnh đại dịch COVID-19, Pfizer nổi bật với sự hợp tác cùng BioNTech để phát triển vaccine ngừa COVID-19 sử dụng công nghệ mRNA. Đây là một bước đột phá y học khi công nghệ này không chỉ giúp kiểm soát đại dịch mà còn mở ra tiềm năng trong việc điều trị nhiều loại bệnh lý khác.

Không dừng lại ở việc phát triển vaccine, Pfizer còn tích cực tham gia vào các chương trình y tế công cộng nhằm đảm bảo sự công bằng trong tiếp cận vaccine, đặc biệt tại các quốc gia đang phát triển. Pfizer cũng phát triển các loại thuốc kháng virus để hỗ trợ điều trị bệnh nhân COVID-19, góp phần giảm thiểu tác động của đại dịch lên hệ thống y tế. Ngoài lĩnh vực vaccine, Pfizer còn là đơn vị tiên phong trong nghiên cứu và phát triển thuốc điều trị các bệnh nghiêm trọng khác như ung thư, tim mạch, và các bệnh lý tự miễn. Các sản phẩm của công ty không chỉ cứu sống hàng triệu người mà còn nâng cao chất lượng cuộc sống, đặc biệt là ở các khu vực khó khăn.

Nhằm thích nghi với xu thế công nghệ, Pfizer đã mở rộng hoạt động sang lĩnh vực y tế kỹ thuật số. Công ty đã phát triển các nền tảng chăm sóc sức khỏe từ xa, ứng dụng quản lý bệnh nhân, và các giải pháp dựa trên trí tuệ nhân tạo (AI) nhằm cá nhân hóa các dịch vụ y tế.

Những đổi mới này không chỉ cải thiện hiệu quả điều trị mà còn mở rộng khả năng tiếp cận dịch vụ y tế, đặc biệt trong bối cảnh đại dịch khi nhu cầu chăm sóc sức khỏe trực tuyến tăng cao.

1.2 Vai trò của hệ thống báo cáo tình hình kinh doanh

Hỗ trợ chiến lược kinh doanh

Hệ thống báo cáo tình hình kinh doanh của Pfizer là công cụ quan trọng để đưa ra các quyết định chiến lược, đảm bảo công ty duy trì vị thế cạnh tranh trong ngành. Với khối lượng dữ liệu khổng lồ thu thập được từ các hoạt động trên toàn cầu, Pfizer sử dụng các công cụ phân tích dữ liệu tiên tiến để:

Xác định các xu hướng thị trường, dự đoán nhu cầu sản phẩm và cải thiện hiệu suất kinh doanh.

Đánh giá hiệu quả của các chương trình tiếp thị và phân phối, từ đó tối ưu hóa chiến lược phát triển sản phẩm.

Phân tích hiệu quả tài chính, giúp công ty đưa ra các quyết định đầu tư và cắt giảm chi phí hợp lý.

Các báo cáo cũng cung cấp cái nhìn chi tiết về các thị trường tiềm năng, từ đó giúp Pfizer định hướng mở rộng và phát triển. Trong giai đoạn đại dịch, hệ thống báo cáo đã giúp công ty điều chỉnh chiến lược phân phối vaccine, ưu tiên các khu vực có nhu cầu cấp bách.

Quản lý chuỗi cung ứng

Pfizer sở hữu một trong những chuỗi cung ứng phức tạp nhất trong ngành y tế, với mạng lưới bao gồm hàng nghìn nhà cung cấp và đối tác trên toàn cầu. Hệ thống báo cáo tình hình kinh doanh giúp:

Theo dõi chặt chẽ quá trình vận chuyển nguyên liệu và sản phẩm.

Dự báo chính xác nhu cầu tại các thị trường khác nhau, từ đó tối ưu hóa hàng tồn kho.

Đảm bảo các sản phẩm quan trọng, đặc biệt là vaccine, được phân phối nhanh chóng và hiệu quả ngay cả trong điều kiện khẩn cấp.

Trong bối cảnh đại dịch COVID-19, hệ thống này đã chứng minh giá trị vượt trội khi giúp Pfizer duy trì sự liên tục trong chuỗi cung ứng toàn cầu, bất chấp các rào cản về vận chuyển và các yếu tố bất ổn khác.

Nâng cao trải nghiệm khách hàng

Hệ thống báo cáo không chỉ phục vụ nội bộ mà còn giúp Pfizer cải thiện trải nghiệm khách hàng. Bằng cách phân tích phản hồi từ bác sĩ, bệnh nhân và các cơ sở y tế, Pfizer có thể:

Cải tiến chất lượng sản phẩm, đáp ứng nhu cầu ngày càng cao của thị trường.

Tăng cường tính cá nhân hóa trong các dịch vụ hỗ trợ khách hàng, giúp xây dựng lòng tin và sự hài lòng.

Đổi mới dựa trên dữ liệu

Pfizer không chỉ sử dụng hệ thống báo cáo để quản lý hoạt động hiện tại mà còn để thúc đẩy đổi mới trong tương lai. Các dữ liệu từ hệ thống này giúp công ty xác định các lĩnh vực cần tập trung nghiên cứu, phát triển các sản phẩm mới, và tối ưu hóa quy trình sản xuất.

→ Hệ thống báo cáo tình hình kinh doanh là một phần không thể thiếu trong hoạt động của Pfizer, giúp công ty duy trì sự ổn định, đổi mới và phát triển bền vững trong một ngành công nghiệp đầy cạnh tranh và biến động. Kết hợp giữa lịch sử lâu đời, năng lực đổi mới, và sự hỗ trợ từ công nghệ hiện đại, Pfizer tiếp tục đóng vai trò tiên phong trong việc cải thiện sức khỏe cộng đồng toàn cầu.

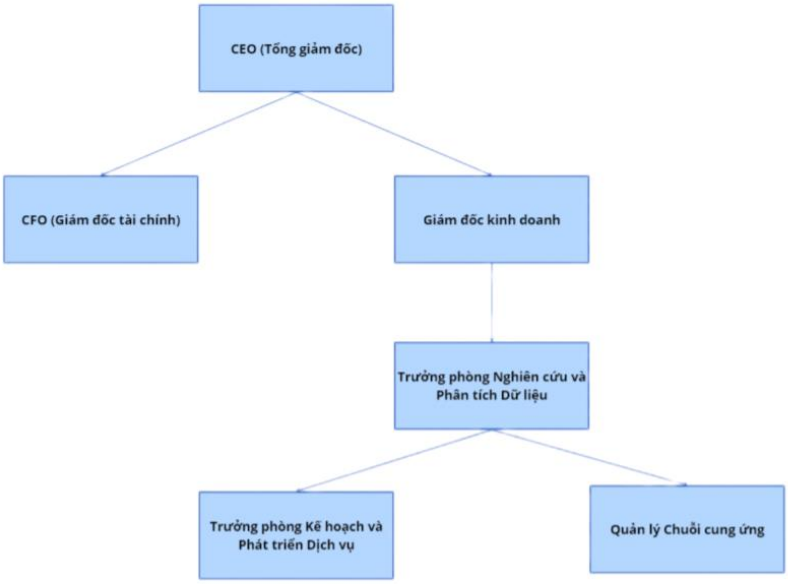
Chương II: Người Dùng Hệ Thống BI (BI Users)

2.1 Phân tích cơ cấu tổ chức của Pfizer

Cơ cấu tổ chức của Pfizer được xây dựng theo mô hình quản lý tập trung, trong đó **Tổng Giám đốc (CEO)** đóng vai trò lãnh đạo và định hướng chiến lược. Dưới CEO là các vị trí quản lý cấp cao bao gồm các giám đốc chức năng và trưởng bộ phận, đảm bảo mọi hoạt động của công ty được điều hành một cách hiệu quả.

Pfizer tổ chức các phòng ban theo chức năng và khu vực địa lý để dễ dàng quản lý, tối ưu hóa hoạt động nghiên cứu, sản xuất, kinh doanh và phân phối dược phẩm trên toàn cầu.

Sơ đồ tổ chức:



Hình 1. Sơ đồ tổ chức

Tổng giám đốc (CEO)

Vai trò:

CEO chịu trách nhiệm điều hành toàn bộ hoạt động của Pfizer, đưa ra các quyết định chiến lược lớn nhằm duy trì vị thế cạnh tranh trong ngành dược phẩm toàn cầu. CEO cũng chịu trách nhiệm đảm bảo rằng công ty tuân thủ các quy định pháp lý, duy trì giá trị cổ đông và đáp ứng mục tiêu kinh doanh.

Mối quan hệ:

Tất cả các giám đốc cấp cao (CFO, Giám đốc Nghiên cứu & Phát triển, Giám đốc Kinh doanh...) báo cáo trực tiếp lên CEO.

Giám đốc Tài chính (CFO)

Vai trò:

CFO quản lý các vấn đề tài chính, bao gồm kế hoạch ngân sách, kiểm soát tài chính, và báo cáo kết quả hoạt động tài chính. CFO đảm bảo rằng các nguồn lực tài chính được phân bổ hợp lý cho nghiên cứu, sản xuất và phát triển sản phẩm.

Mối quan hệ:

CFO làm việc chặt chẽ với CEO và các trưởng bộ phận khác để đảm bảo sự ổn định tài chính và hỗ trợ chiến lược dài hạn.

Giám đốc Kinh doanh (CCO - Chief Commercial Officer)

Vai trò:

Giám đốc Kinh doanh xây dựng và thực hiện các chiến lược kinh doanh toàn cầu, quản lý các thị trường, và giám sát hiệu quả kinh doanh của các sản phẩm. CCO cũng phụ trách mở rộng các kênh phân phối và tối ưu hóa trải nghiệm khách hàng.

Mối quan hệ:

Phối hợp với các bộ phận như chuỗi cung ứng, marketing và nghiên cứu để đảm bảo sản phẩm đến tay khách hàng một cách hiệu quả.

Trưởng phòng Nghiên cứu và Phát triển (R&D)

Vai trò:

Lãnh đạo các hoạt động nghiên cứu và phát triển sản phẩm mới, tập trung vào đổi mới công nghệ và thử nghiệm lâm sàng. Bộ phận này chịu trách nhiệm phát triển các loại dược phẩm và vắc-xin tiên tiến.

Mối quan hệ:

Là bộ phận then chốt, R&D làm việc với tài chính để đảm bảo ngân sách nghiên cứu và liên kết với marketing để đưa sản phẩm ra thị trường.

Quản lý Chuỗi cung ứng

Vai trò:

Đảm bảo sản phẩm được sản xuất, lưu trữ, vận chuyển và phân phối hiệu quả trên toàn cầu. Bộ phận này cũng tối ưu hóa quy trình kho vận và giảm chi phí sản xuất.

Mối quan hệ:

Làm việc chặt chẽ với bộ phận kinh doanh và các nhà cung cấp bên ngoài để đảm bảo nguồn nguyên liệu ổn định và đáp ứng nhu cầu của khách hàng.

Trưởng phòng Kế hoạch và Phát triển Dịch vụ

Vai trò:

Chịu trách nhiệm lập kế hoạch chiến lược phát triển dịch vụ và mở rộng danh mục sản phẩm y tế kỹ thuật số, bao gồm các giải pháp chăm sóc sức khỏe trực tuyến.

Mối quan hệ:

Phối hợp với nghiên cứu và kinh doanh để xây dựng các dịch vụ chăm sóc khách hàng và mở rộng thị trường.

Cơ cấu tổ chức của Pfizer mang tính chuyên môn hóa cao và tập trung vào sự phối hợp liên phòng ban để tối đa hóa hiệu quả hoạt động. Điều này cho phép Pfizer nhanh chóng phát triển các sản phẩm đột phá và duy trì vị trí dẫn đầu trong ngành dược phẩm toàn cầu.

2.2 Các nhóm BI Users và đặc điểm

Tại Pfizer, các nhóm người dùng Business Intelligence (BI) được phân chia theo vai trò và chức năng cụ thể của từng bộ phận trong tổ chức. Việc phân loại này giúp đảm bảo mỗi nhóm nhận được thông tin chính xác và kịp thời để hỗ trợ tối ưu việc ra quyết định và nâng cao hiệu quả kinh doanh. Dưới đây là các nhóm người dùng BI tại Pfizer và đặc điểm của từng nhóm.

2.2.1 Nhóm cấp chiến lược

Nhóm cấp chiến lược bao gồm Ban Điều Hành và Giám Đốc Tài Chính (CFO). Đây là những người chịu trách nhiệm xây dựng các chiến lược phát triển dài hạn và đưa ra các quyết định quan trọng mang tính định hướng cho toàn bộ hoạt động của công ty.

Ban Điều Hành yêu cầu các báo cáo tổng hợp và phân tích sâu về tình hình tài chính, thị trường và hoạt động của công ty. Thông qua BI, họ có thể nhận diện các xu hướng thị trường và các yếu tố ảnh hưởng đến hoạt động kinh doanh, từ đó đưa ra các chiến lược phát triển phù hợp. Ví dụ, họ có thể sử dụng BI để phân tích tình hình dịch bệnh tại các khu vực, giúp định hướng các chiến lược cung ứng và phân phối sản phẩm hiệu quả.

Giám Đốc Tài Chính cần các báo cáo chi tiết về dòng tiền, lợi nhuận, chi phí và các chỉ số tài chính quan trọng. Thông qua BI, CFO có thể phân bổ nguồn lực tài chính hiệu quả, chẳng hạn như tập trung nguồn tiền vào các khu vực có tỷ lệ nhiễm bệnh cao, hoặc điều chỉnh sản xuất vacxin theo nhu cầu cụ thể của từng khu vực. BI cũng giúp dự báo các rủi ro tài chính và tối ưu hóa chi phí trong quá trình sản xuất và phân phối.

Nhóm cấp chiến lược yêu cầu BI với khả năng phân tích toàn diện và hỗ trợ dự báo dài hạn để đảm bảo ra quyết định chính xác và kịp thời.

2.2.2 Nhóm cấp chiến thuật

Nhóm cấp chiến thuật bao gồm Phòng Kế Hoạch và Giám Đốc Kinh Doanh. Nhóm này chịu trách nhiệm triển khai chiến lược tổ chức thành các kế hoạch và hành động cụ thể trong ngắn hạn và trung hạn.

Phòng Kế Hoạch sử dụng BI để phân tích dữ liệu về nhu cầu thị trường, sản xuất và các yếu tố chuỗi cung ứng. Dựa trên các báo cáo từ BI, phòng Kế Hoạch có thể xây dựng các dự báo về nhu cầu sản phẩm, điều chỉnh kế hoạch sản xuất và phân phối một cách linh hoạt. Ví dụ, họ có thể sử dụng BI để điều chỉnh kế hoạch cung cấp vacxin cho các khu vực có số lượng ca bệnh gia tăng.

Giám Đốc Kinh Doanh sử dụng BI để phân tích hành vi khách hàng, hiệu quả marketing và hiệu suất bán hàng. Các phân tích này giúp họ tối ưu hóa chiến lược tiếp thị, điều chỉnh các chương trình quảng bá và tập trung vào các khu vực tiềm năng. Thông qua BI, Giám Đốc Kinh Doanh có thể đưa ra các quyết định chiến lược chính xác, tối đa hóa doanh thu và hiệu quả chiến dịch marketing.

Nhóm cấp chiến thuật yêu cầu BI cung cấp các phân tích chi tiết và hỗ trợ quyết định chiến lược ngắn hạn và trung hạn, đặc biệt là trong việc tối ưu hóa quy trình sản xuất và phân phối.

2.2.3 Nhóm cấp tác nghiệp

Nhóm cấp tác nghiệp bao gồm Phòng Quản Lý Chuỗi Cung Ứng và Phòng Phân Tích Dữ Liệu. Đây là nhóm chịu trách nhiệm đảm bảo các hoạt động hàng ngày của công ty diễn ra hiệu quả và liền mạch.

Phòng Quản Lý Chuỗi Cung Ứng sử dụng BI để giám sát và tối ưu hóa các quy trình vận chuyển, kho bãi và phân phối sản phẩm. Ví dụ, BI giúp phòng Quản Lý Chuỗi Cung Ứng xác định các điểm nghẽn trong chuỗi cung ứng, điều phối các nguồn lực để tránh tình trạng thiếu hụt hoặc dư thừa nguyên vật liệu. Thông qua BI, họ có thể theo dõi và dự báo nhu cầu của các khu vực có tỷ lệ ca nhiễm cao, từ đó đưa ra các chiến lược phân phối kịp thời và hiệu quả.

Phòng Phân Tích Dữ Liệu thu thập, xử lý và phân tích dữ liệu từ các nguồn khác nhau, cung cấp thông tin chi tiết cho các phòng ban khác. Các báo cáo và dự báo từ BI giúp phòng này hỗ trợ các bộ phận khác trong việc tối ưu hóa sản phẩm, cải thiện chất lượng dịch vụ và dự báo nhu cầu thị trường. Ví dụ, BI có thể giúp phân tích các xu hướng dịch bệnh và hỗ trợ các quyết định về sản xuất và phân phối vacxin.

Nhóm cấp tác nghiệp yêu cầu BI cung cấp phân tích chi tiết về các hoạt động hàng ngày, giúp tối ưu hóa các quy trình vận hành và đảm bảo việc phân phối và cung cấp sản phẩm diễn ra đúng thời điểm và hiệu quả.

2.3 Quyết định kinh doanh của BI Users

Các nhóm người dùng BI tại Pfizer sử dụng thông tin từ các báo cáo và phân tích dữ liệu để hỗ trợ quá trình ra quyết định và tối ưu hóa các chiến lược kinh doanh. Dưới đây là các quyết định chủ yếu của từng nhóm trong tổ chức Pfizer.

2.3.1. Quyết định của Ban Điều Hành (Executive Management)

Ban Điều Hành sử dụng BI để hỗ trợ các quyết định chiến lược quan trọng, bao gồm:

Phân bổ tài chính và nguồn lực: Ban Điều Hành dựa vào BI để phân tích các số liệu tài chính và chi phí, từ đó phân bổ nguồn lực một cách hiệu quả cho các bộ phận quan trọng như nghiên cứu và phát triển (R&D), sản xuất và tiếp thị.

Điều chỉnh chiến lược kinh doanh trong bối cảnh dịch bệnh: BI cung cấp thông tin chi tiết về tình hình dịch bệnh, nhu cầu vắc xin và thuốc đặc trị, giúp Ban Điều Hành điều chỉnh chiến lược kinh doanh, ưu tiên sản xuất các sản phẩm y tế thiết yếu như vắc xin mRNA.

Định hướng mở rộng thị trường quốc tế: Ban Điều Hành sử dụng BI để đánh giá tiềm năng thị trường ở các khu vực quốc tế, giúp Pfizer mở rộng hoạt động và tiếp cận các thị trường mới.

2.3.2. Quyết định của Giám Đốc Tài Chính (CFO)

Giám Đốc Tài Chính sử dụng BI để đưa ra các quyết định tài chính và chiến lược chi tiêu:

Tối ưu hóa quản lý tài chính và chi phí: CFO sử dụng BI để theo dõi và điều chỉnh chi phí và lợi nhuận, từ đó đưa ra các quyết định phân bổ ngân sách hợp lý, đặc biệt trong các tình huống khẩn cấp như đại dịch.

Dự đoán doanh thu: BI cung cấp các mô hình dự đoán doanh thu, giúp CFO xây dựng kế hoạch tài chính dài hạn và ứng phó với biến động tài chính.

2.3.3. Quyết định của Giám Đốc Kinh Doanh

Giám Đốc Kinh Doanh sử dụng BI để tối ưu hóa các chiến lược bán hàng và tiếp thị:

Tối ưu hóa chiến lược kinh doanh: BI giúp Giám Đốc Kinh Doanh phân tích xu hướng thị trường và hành vi khách hàng, từ đó điều chỉnh chiến lược để tối ưu hóa doanh thu.

Phân tích đối thủ cạnh tranh: Dữ liệu từ BI giúp Giám Đốc Kinh Doanh đánh giá các đối thủ cạnh tranh, đưa ra các chiến lược để duy trì vị thế cạnh tranh của Pfizer trên thị trường.

2.3.4. Quyết định của Phòng Phân Tích Dữ Liệu (Data Analytics Department):

Phòng Phân Tích Dữ Liệu sử dụng BI để cung cấp các phân tích chi tiết và dự báo cho các bộ phận khác:

Dự đoán xu hướng lây nhiễm và ảnh hưởng đến chuỗi cung ứng y tế: BI giúp phòng phân tích dự báo sự lây lan của dịch bệnh và tác động của nó đến chuỗi cung ứng y tế, giúp Pfizer chủ động điều chỉnh sản xuất và phân phối.

Tối ưu hóa phát triển sản phẩm: Phòng Phân Tích Dữ Liệu sử dụng BI để đánh giá hiệu quả của các sản phẩm hiện có, giúp Pfizer quyết định việc cải tiến hoặc phát triển các sản phẩm mới.

2.3.5. Quyết định của Phòng Kế Hoạch (Planning Department)

Phòng Kế Hoạch sử dụng BI để tối ưu hóa các chiến lược phát triển sản phẩm và dịch vụ:

Dự đoán nhu cầu thị trường: BI cung cấp dữ liệu giúp phòng kế hoạch dự báo nhu cầu thị trường, từ đó điều chỉnh kế hoạch sản xuất để tránh tình trạng dư thừa hoặc thiếu hụt sản phẩm.

Tối ưu hóa phân phối dịch vụ y tế: BI giúp phòng kế hoạch đánh giá nhu cầu phân phối sản phẩm, giúp Pfizer phân phối đúng sản phẩm vào đúng thời điểm và đúng khu vực có nhu cầu cao.

Điều chỉnh chiến lược marketing: BI cung cấp thông tin về hiệu quả của các chiến dịch marketing, giúp phòng kế hoạch điều chỉnh chiến lược quảng bá để tối đa hóa hiệu quả.

2.3.6. Quyết định của Phòng Quản Lý Chuỗi Cung Ứng (Supply Chain Management Department)

Phòng Quản Lý Chuỗi Cung Ứng sử dụng BI để tối ưu hóa hoạt động chuỗi cung ứng:

Quản lý kho hàng và cung ứng vật tư y tế: BI giúp phòng quản lý chuỗi cung ứng theo dõi tình trạng kho hàng và điều chỉnh lượng nguyên liệu đầu vào, giúp đảm bảo sản xuất không bị gián đoạn và tránh lãng phí.

Tối ưu hóa quy trình vận chuyển: BI giúp phòng quản lý chuỗi cung ứng tối ưu hóa các quy trình vận chuyển, giảm thiểu sự chậm trễ và chi phí vượt quá kế hoạch.

Quản lý rủi ro chuỗi cung ứng: Dựa trên dữ liệu từ BI, phòng quản lý chuỗi cung ứng có thể dự báo các rủi ro tiềm ẩn trong chuỗi cung ứng, chẳng hạn như sự gián đoạn cung cấp nguyên liệu hoặc biến động giá cả, từ đó có các chiến lược dự phòng hiệu quả.

Chương III: Khảo sát Hiện Trạng Hệ Thống Thông Tin và Dữ Liệu của Doanh nghiệp Pfizer

3.1 Hiện trạng hệ thống thông tin

Pfizer hiện là một trong những doanh nghiệp hàng đầu trong lĩnh vực dược phẩm toàn cầu, với hệ thống thông tin hiện đại nhằm hỗ trợ quản lý và vận hành hiệu quả. Dưới đây là phân tích về hiện trạng hệ thống thông tin của Pfizer:

Hệ thống ERP (Enterprise Resource Planning):

Pfizer triển khai hệ thống ERP (như **SAP S/4HANA**) để tích hợp và quản lý toàn bộ hoạt động của doanh nghiệp, bao gồm:

Quản lý tài chính, chuỗi cung ứng, sản xuất, và nhân sự.

Tối ưu hóa nguồn lực, cải thiện hiệu quả vận hành, và hỗ trợ giám sát hoạt động theo thời gian thực.

Hệ thống CRM (Customer Relationship Management):

Pfizer sử dụng hệ thống CRM (**Salesforce**) để quản lý dữ liệu khách hàng, đặc biệt là các đối tác như bệnh viện, nhà thuốc và tổ chức y tế. Hệ thống này giúp:

Phân tích hành vi khách hàng, dự đoán nhu cầu và tăng cường khả năng chăm sóc khách hàng.

Hỗ trợ đội ngũ đại diện y khoa (Medical Representatives) trong việc triển khai chiến lược bán hàng và tiếp thị.

Khả năng tích hợp dữ liệu giữa các hệ thống:

Pfizer xây dựng nền tảng tích hợp dữ liệu toàn diện, kết nối các hệ thống ERP, CRM và các công cụ phân tích dữ liệu khác thông qua:

Sử dụng **API** (Application Programming Interface) để đồng bộ hóa dữ liệu.

Hệ thống tích hợp đảm bảo thông tin luôn cập nhật và sẵn sàng phục vụ công tác ra quyết định nhanh chóng, đặc biệt trong chuỗi cung ứng và nghiên cứu phát triển.

3.2 Hiện trạng về dữ liệu và quản trị dữ liệu

3.2.1. Dữ liệu đang được lưu trữ

Dữ liệu nội bộ:

Tài chính: Báo cáo doanh thu, chi phí và ngân sách.

Nhân sự: Hồ sơ nhân viên, lương thưởng và đánh giá hiệu suất.

Nghiên cứu & phát triển: Thông tin về quy trình sản xuất thuốc, kết quả nghiên cứu lâm sàng.

Chuỗi cung ứng: Quản lý tồn kho, lịch trình giao hàng và nguồn cung cấp nguyên liệu.

Dữ liệu bên ngoài:

Thị trường: Thông tin từ các nghiên cứu thị trường, dự báo nhu cầu, và các đối thủ cạnh tranh.

Y tế: Thông tin bệnh lý, xu hướng điều trị và các nghiên cứu từ tổ chức y tế quốc tế (WHO, FDA).

Nguồn dữ liệu:

Nội bộ: Dữ liệu được thu thập từ hệ thống ERP, CRM và các công cụ vận hành khác.

Bên ngoài: Báo cáo từ tổ chức y tế, dữ liệu nghiên cứu thị trường, và thông tin bệnh nhân từ các đối tác nghiên cứu.

Chiều sâu dữ liệu:

Dữ liệu được lưu trữ có chiều sâu từ cấp độ tổng quan (báo cáo doanh nghiệp) đến chi tiết (dữ liệu từng sản phẩm, khách hàng và nghiên cứu).

3.2.2. Quản trị dữ liệu

a. Chính sách đảm bảo chất lượng dữ liệu

Pfizer triển khai các công cụ và quy trình tự động để đảm bảo:

Chất lượng dữ liệu:

Tính chính xác, đầy đủ và kịp thời của dữ liệu.

Kiểm tra chất lượng dữ liệu định kỳ tại tất cả các giai đoạn, từ nhập liệu đến phân tích.

Đồng bộ dữ liệu: Sử dụng các công cụ đồng bộ để tránh xung đột và sai lệch trong dữ liệu.

b. Chính sách an toàn, bảo mật dữ liệu

Pfizer áp dụng các chính sách bảo mật nghiêm ngặt nhằm bảo vệ dữ liệu doanh nghiệp và thông tin khách hàng:

Tuân thủ quy định quốc tế:

GDPR (General Data Protection Regulation): Bảo vệ dữ liệu cá nhân của khách hàng và đối tác tại châu Âu.

HIPAA (Health Insurance Portability and Accountability Act): Bảo vệ thông tin y tế nhạy cảm.

Công nghệ bảo mật:

Dữ liệu được mã hóa toàn diện khi lưu trữ và truyền tải.

Hệ thống giám sát an ninh 24/7 để phát hiện và ngăn chặn các cuộc tấn công mạng.

Đào tạo nhân viên: Pfizer tổ chức các khóa huấn luyện định kỳ để nâng cao nhận thức về bảo mật dữ liệu và an toàn thông tin cho toàn bộ nhân viên.

Pfizer đã xây dựng nền tảng hệ thống thông tin và quản trị dữ liệu hiện đại, đáp ứng đầy đủ các yêu cầu quản lý và bảo mật, hỗ trợ doanh nghiệp duy trì vị thế cạnh tranh trong ngành dược phẩm toàn cầu.

Chương IV: Thiết Kế và Cài Đặt Kho Dữ Liệu

4.1. Sự cần thiết của việc xây dựng Kho dữ liệu trong bối cảnh mục tiêu kinh doanh của Pfizer

Trong bối cảnh đại dịch Covid-19, mục tiêu kinh doanh của Pfizer không chỉ tập trung vào việc phát triển và phân phối vắc xin, mà còn nhằm tối ưu hóa quá trình ra quyết định dựa trên phân tích dữ liệu. Việc xây dựng một Kho dữ liệu (Data Warehouse) là cần thiết vì các lý do sau:

Tích hợp và quản lý dữ liệu hiệu quả

Kho dữ liệu cho phép tích hợp dữ liệu từ nhiều nguồn khác nhau như số ca nhiễm, tiêm chủng, và khả năng lây lan của virus. Điều này giúp Pfizer có một cái nhìn toàn diện và chính xác về tình hình dịch bệnh.

Hỗ trợ phân tích và dự đoán

Với Kho dữ liệu, Pfizer có thể sử dụng các công cụ phân tích và mô hình dự đoán để dự báo nhu cầu dịch vụ y tế trên quy mô toàn cầu. Điều này giúp công ty đưa ra các quyết định nhanh chóng và kịp thời, đảm bảo dịch vụ chăm sóc sức khỏe được phân bổ hợp lý.

Tối ưu hóa quá trình ra quyết định

Dữ liệu được tổ chức và lưu trữ trong Kho dữ liệu giúp các nhà quản lý và chuyên gia của Pfizer dễ dàng truy cập và phân tích thông tin. Điều này tối ưu hóa quá trình ra quyết định, giúp công ty đưa ra các biện pháp quản lý dịch vụ y tế hiệu quả.

Đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng

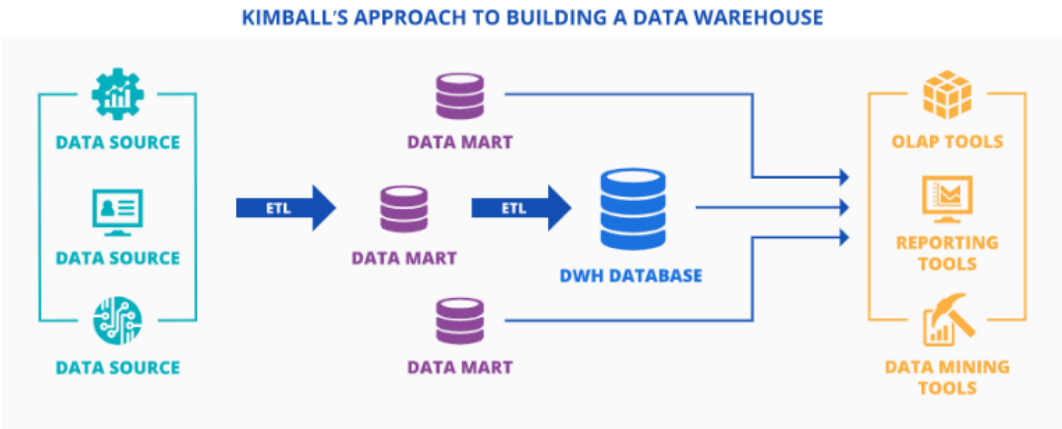
Thông qua việc phân tích dữ liệu, Pfizer có thể dự đoán chính xác hơn về nhu cầu y tế trong tương lai. Điều này giúp công ty và các đối tác của mình đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng, đặc biệt là trong các khu vực chịu ảnh hưởng nặng nề.

Hỗ trợ hợp tác và chia sẻ thông tin

Kho dữ liệu cũng tạo điều kiện thuận lợi cho việc hợp tác và chia sẻ thông tin giữa Pfizer và các đối tác. Điều này giúp tăng cường hiệu quả của các biện pháp phòng chống dịch bệnh và đảm bảo sự phối hợp nhịp nhàng trong việc phân phối vắc xin và dịch vụ y tế.

Việc xây dựng Kho dữ liệu không chỉ giúp Pfizer tối ưu hóa quá trình ra quyết định mà còn đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng trong thời điểm khủng hoảng. Đây là một bước quan trọng để công ty có thể thực hiện mục tiêu kinh doanh và đóng góp vào việc kiểm soát đại dịch Covid-19.

Kiến trúc kho dữ liệu của nhóm sẽ dùng phương pháp bottom-up của Ralph Kimball



Hình 2. Kiến trúc Kho dữ liệu của Kimball

Kiến trúc hệ thống Data Warehouse theo kiểu bottom-up của Kimball tập trung vào việc xây dựng từng phần nhỏ, gọi là Data marts, rồi kết hợp chúng để tạo thành một Data Warehouse lớn hơn. Đây là cách tiếp cận phổ biến và linh hoạt, cho phép tạo ra các phần nhỏ có khả năng đáp ứng nhu cầu kinh doanh cụ thể trước khi tích hợp chúng thành một hệ thống toàn diện hơn.

Đầu tiên từ data source sau đó sẽ thiết kế các data mart cần thiết cho nghiệp vụ phân tích và dự đoán nhu cầu dịch vụ y tế trên quy mô toàn cầu như:

Data Mart về Quốc gia (Country Data Mart):

Tập trung vào dữ liệu liên quan đến các quốc gia, bao gồm số ca nhiễm, số ca tử vong, số ca hồi phục, và các chỉ số khác theo từng quốc gia.

Data Mart về Ngày tháng (Date Data Mart):

Tập trung vào dữ liệu theo ngày tháng, bao gồm số ca nhiễm, số ca tử vong, số ca hồi phục, và các chỉ số khác theo từng ngày.

Data Mart về Châu lục (Continent Data Mart):

Tập trung vào dữ liệu theo châu lục, bao gồm số ca nhiễm, số ca tử vong, số ca hồi phục, và các chỉ số khác theo từng châu lục.

Data Mart về Khu vực WHO (WHO Region Data Mart):

Tập trung vào dữ liệu theo khu vực của WHO, bao gồm số ca nhiễm, số ca tử vong, số ca hồi phục, và các chỉ số khác theo từng khu vực của WHO.

Lợi ích của kiến trúc này

Tích hợp dữ liệu từ nhiều nguồn: Giúp Pfizer có cái nhìn toàn diện về tình hình dịch bệnh.

Hỗ trợ phân tích và dự đoán: Giúp công ty đưa ra các quyết định nhanh chóng và kịp thời.

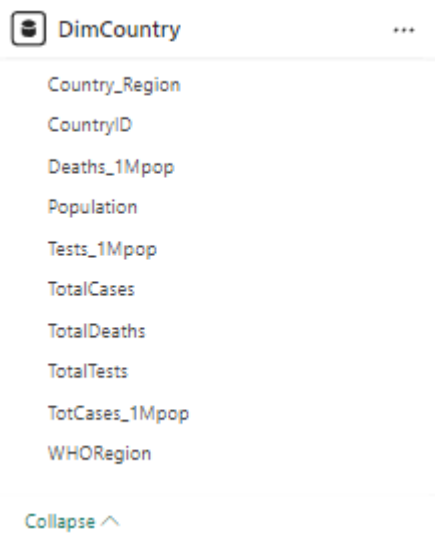
Tối ưu hóa quá trình ra quyết định: Dữ liệu được tổ chức và lưu trữ một cách khoa học, dễ dàng truy cập và phân tích.

Đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng: Giúp dự đoán chính xác hơn về nhu cầu y tế trong tương lai

4.2. Thiết kế các bảng chiều (bảng cắt lớp) DIM

Bảng Country Dimension

Bảng Country Dimension là một phần quan trọng trong mô hình dữ liệu Star Schema mà nhóm đang hướng tới. Được sử dụng để lưu trữ thông tin về các quốc gia, bảng này giúp cải thiện khả năng phân tích, theo dõi và hiểu rõ hơn về dữ liệu liên quan đến các quốc gia trong hệ thống thông tin.

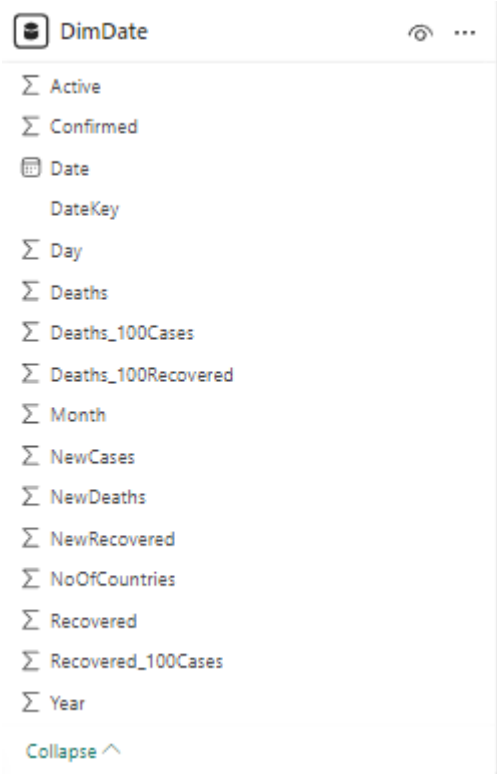


Hình 3. Bảng DimCountry

Bảng này sẽ giúp phân tích dữ liệu theo nhiều khía cạnh quốc gia khác nhau, cho phép phân tích dữ liệu theo các chỉ số như dân số, số ca nhiễm, số ca tử vong, và số lượng xét nghiệm. Điều này giúp hiểu rõ hơn về tình hình dịch bệnh và các chỉ số y tế theo từng quốc gia.

Bảng Date Dimension

Bảng DimDate được thiết kế để quản lý thời gian, bao gồm các trường như ngày, tháng, năm, và các thông tin tổng hợp liên quan như số ca nhiễm, tử vong, và hồi phục mỗi ngày. Đây là bảng dimension quan trọng, cho phép phân tích dữ liệu theo chiều thời gian.

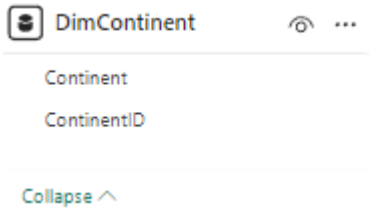


Hình 4. Bảng DimDate

Với bảng này, các xu hướng theo thời gian như tốc độ lây lan, mức độ phục hồi, hoặc tác động của các biện pháp phòng chống dịch có thể được phân tích và hiển thị trực quan trên các dashboard.

Bảng Continent Dimension

Bảng DimContinent lưu trữ thông tin về các châu lục và các quốc gia thuộc mỗi châu lục. Bảng này giúp cung cấp cái nhìn tổng thể về tình hình dịch bệnh trên từng châu lục, hỗ trợ phân tích dữ liệu ở mức cao hơn.

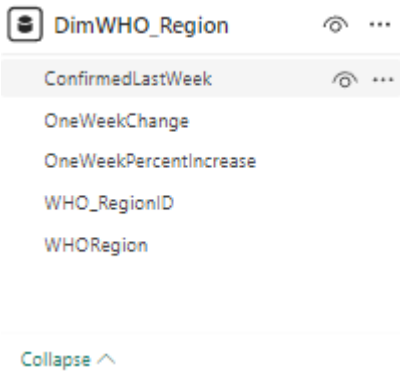


Hình 5. Bảng DimContinent

Phân tích từ bảng **DimContinent** cho phép người dùng so sánh tỷ lệ lây nhiễm hoặc tỷ lệ tử vong giữa các châu lục, giúp các cơ quan y tế đưa ra chính sách toàn cầu hiệu quả hơn.

Bảng WHO_Region Dimension

Bảng **DimWHO_Region** lưu trữ thông tin về các khu vực quản lý của Tổ chức Y tế Thế giới (WHO), bao gồm tên vùng, quốc gia, số ca nhiễm trong tuần qua, thay đổi theo tuần, và tỷ lệ phần trăm tăng trưởng.



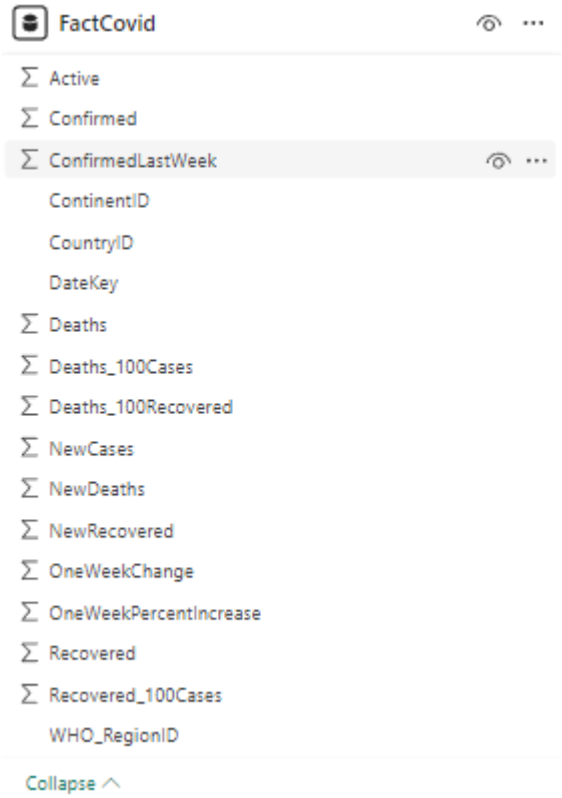
Hình 6. Bảng WHO_Region Dimension

Bảng này hỗ trợ các báo cáo đánh giá hiệu quả chính sách theo từng khu vực WHO và theo dõi tốc độ lây lan trong từng khu vực, là cơ sở cho các chiến lược kiểm soát dịch bệnh.

4.3. Thiết kế bảng FACT

4.3.1. Bảng Fact Covid

Bảng FactCovid là một phần quan trọng trong mô hình dữ liệu Star Schema mà nhóm đang hướng tới. Được sử dụng để lưu trữ các dữ liệu thực tế liên quan đến Covid-19, bảng này giúp cải thiện khả năng phân tích, theo dõi và hiểu rõ hơn về tình hình dịch bệnh trên toàn cầu.



Hình 7. Bảng FactCovid

Bảng này sẽ giúp phân tích dữ liệu theo nhiều khía cạnh khác nhau, cho phép phân tích dữ liệu theo quốc gia, ngày tháng, châu lục, và khu vực WHO. Điều này giúp hiểu rõ hơn về xu hướng và biểu đồ thay đổi của dịch bệnh theo thời gian và không gian.

Chi tiết các cột khoá chính trong bảng FactCovid

CountryID: Mã định danh duy nhất cho mỗi quốc gia, liên kết với bảng DimCountry.

DateKey: Mã định danh duy nhất cho mỗi ngày, liên kết với bảng DimDate.

ContinentID: Mã định danh duy nhất cho mỗi châu lục, liên kết với bảng DimContinent.

WHO_RegionID: Mã định danh duy nhất cho mỗi khu vực của WHO, liên kết với bảng DimWHO_Region.

Confirmed: Số ca nhiễm được xác nhận.

Deaths: Số ca tử vong.

Recovered: Số ca hồi phục.

Active: Số ca đang hoạt động.

NewCases: Số ca nhiễm mới.

NewDeaths: Số ca tử vong mới.

NewRecovered: Số ca hồi phục mới.

Deaths_100Cases: Tỷ lệ tử vong trên 100 ca nhiễm.

Recovered_100Cases: Tỷ lệ hồi phục trên 100 ca nhiễm.

Deaths_100Recovered: Tỷ lệ tử vong trên 100 ca hồi phục.

4.3.2. Lợi ích của bảng FactCovid đối với Pfizer

Tích hợp dữ liệu từ nhiều nguồn

Bảng FactCovid giúp tích hợp dữ liệu từ nhiều nguồn khác nhau, bao gồm các quốc gia, ngày tháng, châu lục, và khu vực WHO. Điều này giúp Pfizer có một cái nhìn toàn diện và chính xác về tình hình dịch bệnh.

Hỗ trợ phân tích và dự đoán

Với bảng FactCovid, Pfizer có thể sử dụng các công cụ phân tích và mô hình dự đoán để dự báo nhu cầu dịch vụ y tế trên quy mô toàn cầu. Điều này giúp công ty đưa ra các quyết định nhanh chóng và kịp thời, đảm bảo dịch vụ chăm sóc sức khỏe được phân bổ hợp lý.

Tối ưu hóa quá trình ra quyết định

Dữ liệu được tổ chức và lưu trữ trong bảng FactCovid giúp các nhà quản lý và chuyên gia của Pfizer dễ dàng truy cập và phân tích thông tin. Điều này tối ưu hóa quá trình ra quyết định, giúp công ty đưa ra các biện pháp quản lý dịch vụ y tế hiệu quả.

Đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng

Thông qua việc phân tích dữ liệu, Pfizer có thể dự đoán chính xác hơn về nhu cầu y tế trong tương lai. Điều này giúp công ty và các đối tác của mình đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng, đặc biệt là trong các khu vực chịu ảnh hưởng nặng nề.

Hỗ trợ hợp tác và chia sẻ thông tin

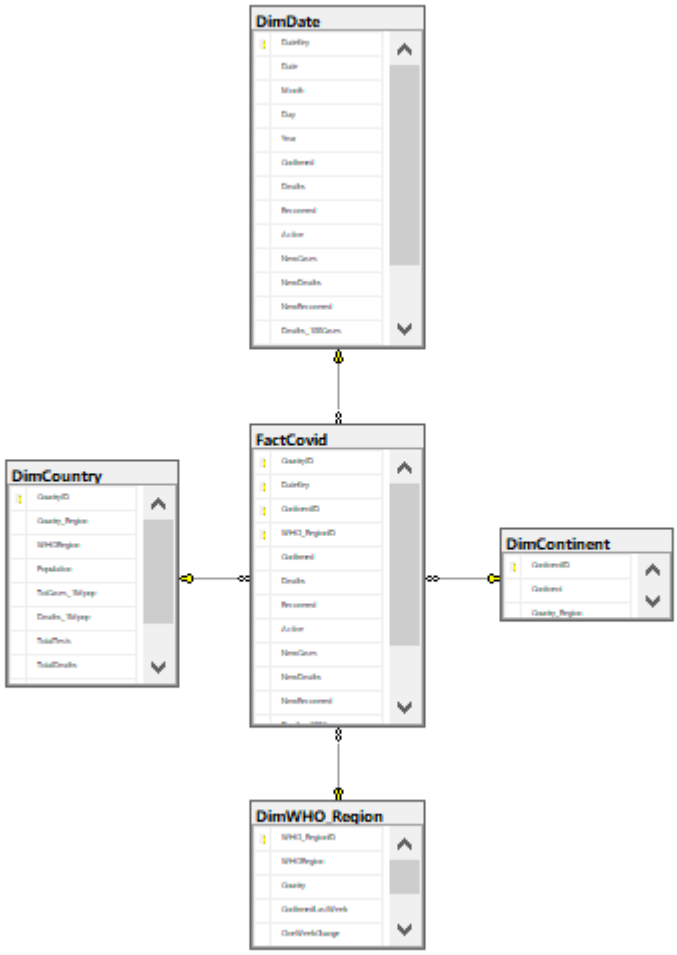
Bảng FactCovid cũng tạo điều kiện thuận lợi cho việc hợp tác và chia sẻ thông tin giữa Pfizer và các đối tác. Điều này giúp tăng cường hiệu quả của các biện pháp phòng chống dịch bệnh và đảm bảo sự phối hợp nhịp nhàng trong việc phân phối vắc xin và dịch vụ y tế.

Việc xây dựng và sử dụng bảng FactCovid không chỉ giúp Pfizer tối ưu hóa quá trình ra quyết định mà còn đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng trong thời điểm khủng hoảng. Đây là một bước quan trọng để công ty có thể thực hiện mục tiêu kinh doanh và đóng góp vào việc kiểm soát đại dịch Covid-19.

4.4. Mô hình trên Data Mart

4.4.1. Mô tả

Mô hình lược đồ của nhóm sẽ hướng tới mô hình sao (Star Schema). Lược đồ hình sao (Star Schema) là một cấu trúc dữ liệu phổ biến trong kho dữ liệu, giúp tối ưu hóa quá trình truy vấn và phân tích dữ liệu. Bảng FactCovid là trung tâm của lược đồ hình sao, chứa các dữ liệu thực tế liên quan đến Covid-19. Bảng này bao gồm các chỉ số chính như số ca nhiễm, số ca tử vong, số ca hồi phục, và các chỉ số khác. Các khóa ngoại trong bảng này liên kết với các bảng Dimension để cung cấp ngữ cảnh cho dữ liệu.



Hình 8. Mô hình trên Data Mart

4.4.2 Lợi ích của mô hình Data Mart theo lược đồ hình sao đối với Pfizer (tiếp tục)

Hỗ trợ phân tích và dự đoán

Các bảng Dimension cung cấp thông tin chi tiết về các thực thể như quốc gia, ngày tháng, châu lục, và khu vực WHO. Điều này giúp Pfizer có thể thực hiện các phân tích và dự đoán chính xác hơn về tình hình dịch bệnh và nhu cầu y tế.

Tăng hiệu suất truy vấn

Sử dụng mô hình lược đồ hình sao giúp tăng hiệu suất truy vấn bằng cách giảm thiểu số lượng dữ liệu cần xử lý trong mỗi truy vấn. Điều này giúp Pfizer có thể truy xuất và phân tích dữ liệu nhanh chóng, hỗ trợ quá trình ra quyết định kịp thời.

Hỗ trợ ra quyết định chiến lược

Với dữ liệu được tổ chức và phân tích một cách khoa học, Pfizer có thể đưa ra các quyết định chiến lược dựa trên dữ liệu, chẳng hạn như phân bổ nguồn lực y tế, phát triển và phân phối vắc xin, và quản lý dịch vụ y tế hiệu quả hơn.

Đáp ứng nhu cầu sức khỏe cộng đồng

Thông qua việc phân tích dữ liệu từ mô hình này, Pfizer có thể dự đoán chính xác hơn về nhu cầu y tế trong tương lai. Điều này giúp công ty và các đối tác của mình đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng, đặc biệt là trong các khu vực chịu ảnh hưởng nặng nề bởi đại dịch.

Việc xây dựng và sử dụng mô hình Data Mart theo lược đồ hình sao không chỉ giúp Pfizer tối ưu hóa quá trình ra quyết định mà còn đảm bảo khả năng đáp ứng nhu cầu sức khỏe của cộng đồng trong thời điểm khủng hoảng. Đây là một bước quan trọng để công ty có thể thực hiện mục tiêu kinh doanh và đóng góp vào việc kiểm soát đại dịch Covid-19.

4.5. Tổng quan quá trình đổ dữ liệu

Trong xây dựng kho dữ liệu (Data Warehouse), quá trình đổ dữ liệu được thực hiện thông qua ETL (Extract, Transform, Load). Đây là một quy trình quan trọng giúp chuyển đổi dữ liệu thô từ các nguồn khác nhau vào kho dữ liệu để phục vụ cho các hoạt động phân tích và ra quyết định trong doanh nghiệp.

4.5.1. Giới Thiệu Về Quá Trình ETL (Extract, Transform, Load)

Quy trình ETL bao gồm ba bước chính:

Extract (Trích xuất):

Trích xuất dữ liệu từ nhiều nguồn khác nhau, có thể là file CSV, cơ sở dữ liệu, API hoặc các nguồn dữ liệu khác.

Trong bài toán về COVID-19, dữ liệu được trích xuất từ các file CSV chứa thông tin thống kê về số ca nhiễm, tử vong, phục hồi, và các yếu tố liên quan theo ngày, quốc gia, và khu vực.

Transform (Chuyển đổi):

Dữ liệu thô sau khi trích xuất thường không ở trạng thái phù hợp cho phân tích.

Giai đoạn này bao gồm các thao tác như: loại bỏ dữ liệu lỗi hoặc trùng lặp; chuyển đổi định dạng ngày tháng hoặc kiểu dữ liệu (như chuyển chuỗi thành số); tổng hợp hoặc chuẩn hóa dữ liệu để đồng nhất giữa các nguồn. Ví dụ như gộp dữ liệu từ các quốc gia hoặc chuẩn hóa tên quốc gia để nhất quán giữa các file.

Load (Nạp vào kho dữ liệu):

Dữ liệu sau khi được chuyển đổi sẽ được nạp vào các bảng dimension và fact trong kho dữ liệu. Quá trình này có thể được thực hiện bằng nhiều phương pháp khác nhau, như Import Wizard của SQL Server hoặc sử dụng câu lệnh SQL BULK INSERT.

4.5.2. Phân Biệt Giữa Bảng Dimension và Bảng Fact

Bảng Dimension (Dim)

Chứa thông tin mô tả hoặc tham chiếu giúp phân loại và cung cấp bối cảnh cho dữ liệu trong bảng Fact. Dữ liệu trong bảng Dimension có tính ổn định cao, ít thay đổi thường xuyên.

Ví dụ trong bài toán dữ liệu COVID-19 của nhóm:

DimCountry chứa thông tin mô tả về các quốc gia, như Country/Region, Continent, Population, WHO Region.

DimDate chứa dữ liệu về thời gian, như ngày tháng (Date).

Bảng Fact

Chứa dữ liệu giao dịch hoặc dữ liệu thực tế cần được phân tích. Bảng Fact thường bao gồm các giá trị số (measures) và các khóa ngoại liên kết với bảng Dimension.

Sự khác biệt chính giữa 2 bảng

Tiêu chí	Bảng Dimension	Bảng Fact
Mục đích	Cung cấp bối cảnh, thông tin mô tả	Lưu trữ dữ liệu thực tế cần phân tích
Loại dữ liệu	Dữ liệu ổn định, ít thay đổi	Dữ liệu biến động, thay đổi liên tục

4.5.3. Các Phương Pháp Phổ Biến Để Đổ Dữ Liệu vào SQL Server

Sử dụng Import Wizard

SQL Server Management Studio (SSMS) cung cấp công cụ Import and Export Wizard giúp nhập dữ liệu từ các nguồn như CSV, Excel, và cơ sở dữ liệu khác. Phù hợp cho việc nhập dữ liệu đơn giản từ file CSV, XLSX hoặc các nguồn tương tự vào bảng SQL.

Ưu điểm: Dễ sử dụng, không cần nhiều kiến thức kỹ thuật.

Nhược điểm: Thiếu tính linh hoạt khi cần xử lý dữ liệu phức tạp

Sử dụng BULK INSERT

BULK INSERT là câu lệnh SQL dùng để nhập dữ liệu từ file vào bảng nhanh chóng và hiệu quả.

Ưu điểm: Tốc độ cao, phù hợp cho lượng dữ liệu lớn.

Nhược điểm: Yêu cầu cấu hình chính xác, dễ gặp lỗi nếu file nguồn có vấn đề.

4.6. Chuẩn bị dữ liệu và đổ dữ liệu vào bảng

Quá trình chuẩn bị dữ liệu và cơ sở dữ liệu là bước quan trọng để đảm bảo quá trình ETL diễn ra suôn sẻ.

4.6.1. Kiểm tra và Chuẩn Hóa Dữ Liệu trong File CSV

Khi xử lý dữ liệu từ nhiều nguồn, việc sử dụng Python cho EDA (Exploratory Data Analysis) giúp kiểm tra và chuẩn hóa dữ liệu nhanh chóng và hiệu quả.

Mục tiêu của EDA và Chuẩn Hóa Dữ Liệu

Xử lý dữ liệu thiếu và lỗi định dạng.

Kiểm tra tính nhất quán của dữ liệu giữa các file.

Loại bỏ dữ liệu không hợp lệ hoặc trùng lặp.

Đảm bảo tất cả dữ liệu đã sẵn sàng cho quá trình ETL.

- Đầu tiên, nhóm thực hiện đọc dữ Liệu từ các file CSV

```
df_country = pd.read_csv('country_wise_latest.csv')
df_covid = pd.read_csv('covid_19_clean_complete.csv')
df_day = pd.read_csv('day_wise.csv')
df_fgroup = pd.read_csv('full_grouped.csv')
df_world = pd.read_csv('worldometer_data.csv')
```

```
all_dfs = [df_country, df_covid, df_day, df_fgroup, df_world]
for df in all_dfs:
    print(f'-------{df.head()}----------\n\n\n\n\n')
```

```
-------  Country/Region  Confirmed  Deaths  Recovered  Active  New cases  New deaths  \
0    Afghanistan      36263    1269      25198    9796        106          10
1        Albania       4880     144       2745    1991        117           6
2        Algeria      27973    1163      18837    7973        616           8
3        Andorra        907      52        803      52         10           0
4         Angola        950      41        242     667         18           1

   New recovered  Deaths / 100 Cases  Recovered / 100 Cases  \
0             18                3.50                  69.49
1             63                2.95                  56.25
2            749                4.16                  67.34
3              0                5.73                  88.53
4              0                4.32                  25.47

   Deaths / 100 Recovered  Confirmed last week  1 week change  \
0                    5.04                35526            737
1                    5.25                 4171            709
2                    6.17                23691           4282
3                    6.48                  884             23
4                   16.94                  749            201

   1 week % increase             WHO Region
0               2.07  Eastern Mediterranean
1              17.00                 Europe
2              18.07                 Africa
```

- Xử lý dữ liệu khuyết thiếu (Missing Data) bằng cách kiểm tra và xử lý các gí trị **NULL**

```
df_fgroup.isnull().sum()
```

```
Date              0
Country/Region    0
Confirmed         0
Deaths            0
Recovered         0
Active            0
New cases         0
New deaths        0
New recovered     0
WHO Region        0
dtype: int64
```

```
df_country.isnull().sum()
```

```
Country/Region            0
Confirmed                 0
Deaths                    0
Recovered                 0
Active                    0
New cases                 0
New deaths                0
New recovered             0
Deaths / 100 Cases        0
Recovered / 100 Cases     0
Deaths / 100 Recovered    0
Confirmed last week       0
1 week change             0
1 week % increase         0
WHO Region                0
dtype: int64
```

```
df_covid.isnull().sum()
```

```
Province/State    34404
Country/Region        0
Lat                   0
Long                  0
Date                  0
Confirmed             0
Deaths                0
Recovered             0
Active                0
WHO Region            0
dtype: int64
```

```
df_world.isnull().sum()
```

```
Country/Region        0
Continent             1
Population            1
TotalCases            0
NewCases            205
TotalDeaths          21
NewDeaths           206
TotalRecovered        4
NewRecovered        206
ActiveCases           4
Serious,Critical     87
Tot Cases/1M pop      1
Deaths/1M pop        22
TotalTests           18
Tests/1M pop         18
WHO Region           25
dtype: int64
```

```
df_day.isnull().sum()
```

```
Date                      0
Confirmed                 0
Deaths                    0
Recovered                 0
Active                    0
New cases                 0
New deaths                0
New recovered             0
Deaths / 100 Cases        0
Recovered / 100 Cases     0
Deaths / 100 Recovered    0
No. of countries          0
dtype: int64
```

- Kiểm tra dữ liệu trùng lặp

```
df_covid.duplicated().sum()
```

```
np.int64(0)
```

```
df_country.duplicated().sum()
```

```
np.int64(0)
```

```
df_day.duplicated().sum()
```

```
np.int64(0)
```

```
df_fgroup.duplicated().sum()
```

```
np.int64(0)
```

```
df_world.duplicated().sum()
```

```
np.int64(0)
```

Với việc sử dụng Python cho EDA, nhóm đã kiểm tra dữ liệu một cách có hệ thống và loại bỏ những vấn đề tiềm ẩn trước khi đổ vào SQL Server. Các thao tác trên giúp tối ưu hóa quá trình ETL, tránh lỗi trong khi import dữ liệu và đảm bảo dữ liệu sẵn sàng cho phân tích trong kho dữ liệu. Nhóm chuyển đổi các **file csv -> file xlsx** và tạo các cột ID tự động cho các bảng và thời gian định dạng dữ liệu Date MMDDYYYY; sau đó, nhóm tạo cơ sở dữ liệu và các bảng Dim, Fact phù hợp trước khi đẩy dữ liệu vào bảng.

```
create database Covid19_dataa
Go
```

4.6.2. Tạo các bảng Dim

Bảng DimCountry

```sql
CREATE TABLE DimCountry (
    CountryID INT PRIMARY KEY,
    Country_Region VARCHAR(255),
    WHORegion VARCHAR(255),
    Population BIGINT,
    TotCases_1Mpop INT,
    Deaths_1Mpop INT,
    TotalTests BIGINT,
    TotalDeaths BIGINT,
    Tests_1Mpop INT,
    TotalCases BIGINT
);
```

Bảng DimDate

```sql
CREATE TABLE DimDate (
    DateKey INT PRIMARY KEY,
    Date DATE,
    Month INT,
    Day INT,
    Year INT,
    Confirmed BIGINT,
    Deaths BIGINT,
    Recovered BIGINT,
    Active BIGINT,
    NewCases BIGINT,
    NewDeaths BIGINT,
    NewRecovered BIGINT,
    Deaths_100Cases DECIMAL(5, 2),
    Recovered_100Cases DECIMAL(5, 2),
    Deaths_100Recovered DECIMAL(5, 2),
    NoOfCountries INT
);
```

Bảng DimContinent

```sql
CREATE TABLE DimContinent (
    ContinentID INT PRIMARY KEY,
    Continent VARCHAR(255),
```

Bảng DimWHO_Region

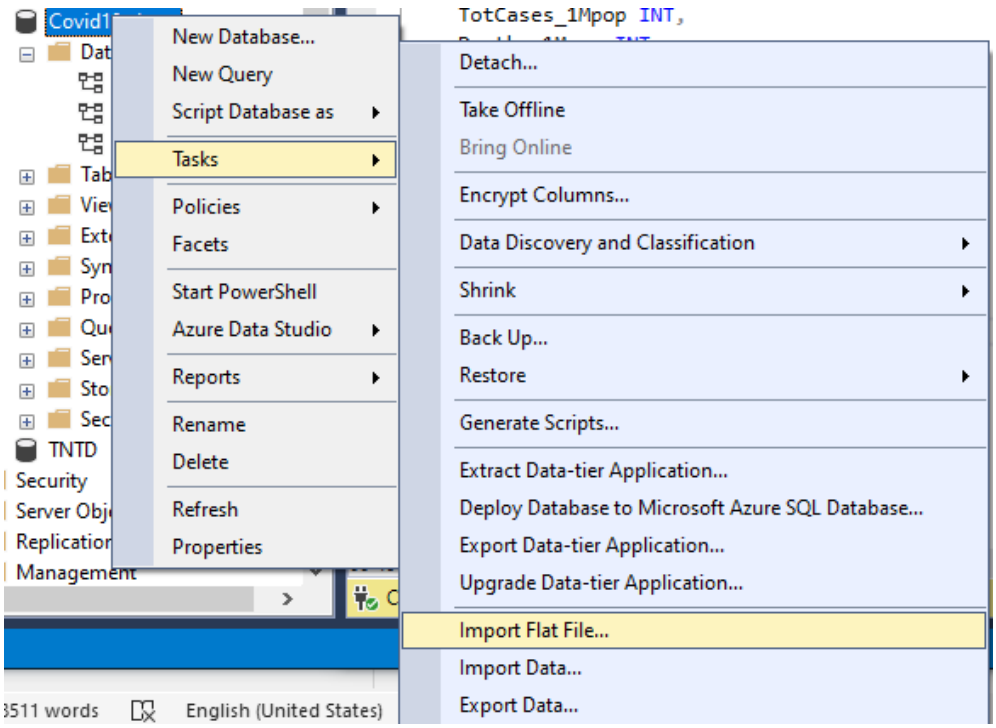
```sql
CREATE TABLE DimWHO_Region (
    WHO_RegionID INT PRIMARY KEY,
    WHORegion VARCHAR(255),
    Country VARCHAR(255),
    ConfirmedLastWeek BIGINT,
    OneWeekChange BIGINT,
    OneWeekPercentIncrease DECIMAL(5, 2)
);
```

4.6.3. Tạo bảng Fact

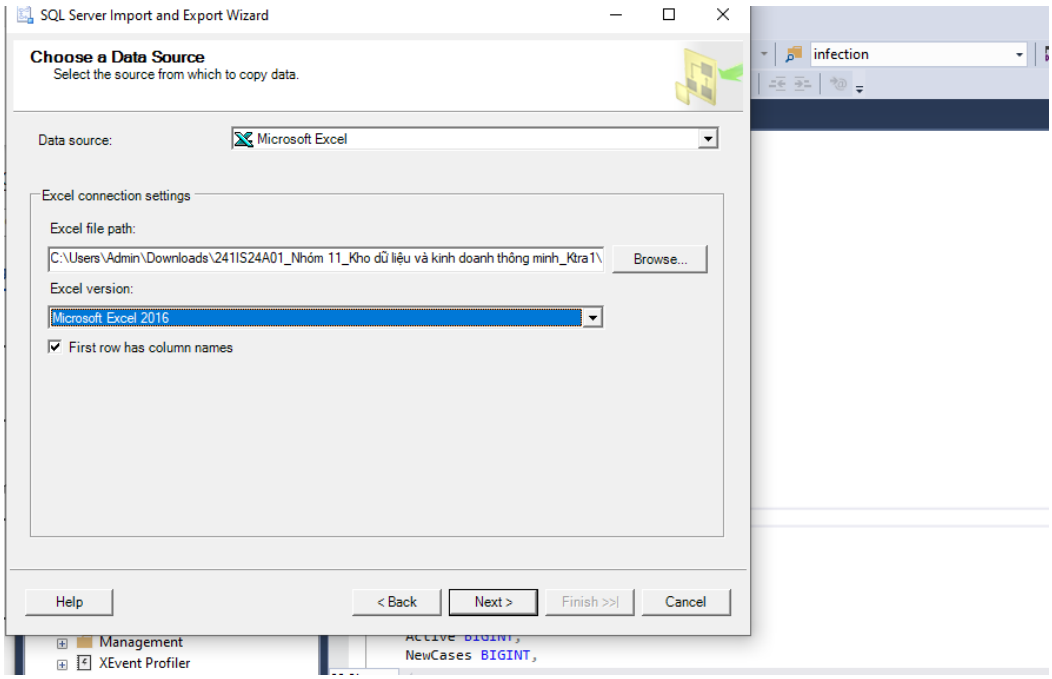
```sql
CREATE TABLE FactCovid (
    CountryID INT,
    DateKey INT,
    ContinentID INT,
    WHO_RegionID INT,
    Confirmed BIGINT,
    Deaths BIGINT,
    Recovered BIGINT,
    Active BIGINT,
    NewCases BIGINT,
    NewDeaths BIGINT,
    NewRecovered BIGINT,
    Deaths_100Cases DECIMAL(5, 2),
    Recovered_100Cases DECIMAL(5, 2),
    Deaths_100Recovered DECIMAL(5, 2),
    PRIMARY KEY (CountryID, DateKey, ContinentID, WHO_RegionID),
    FOREIGN KEY (CountryID) REFERENCES DimCountry(CountryID),
    FOREIGN KEY (DateKey) REFERENCES DimDate(DateKey),
    FOREIGN KEY (ContinentID) REFERENCES DimContinent(ContinentID),
    FOREIGN KEY (WHO_RegionID) REFERENCES DimWHO_Region(WHO_RegionID)
);
```

4.6.4. Đổ dữ liệu vào bảng

Trong SQL Server Management Studio (SSMS), **chuột phải vào cơ sở dữ liệu đã tạo "Covid19_dataa".** Chọn **Task -> Import Data.**



Trong cửa sổ SQL Server Import and Export Wizard, chọn Microsoft Excel. Nhấn Browse và chọn file cần nhập. Đảm bảo đánh dấu vào **First row has column names -> Next.**

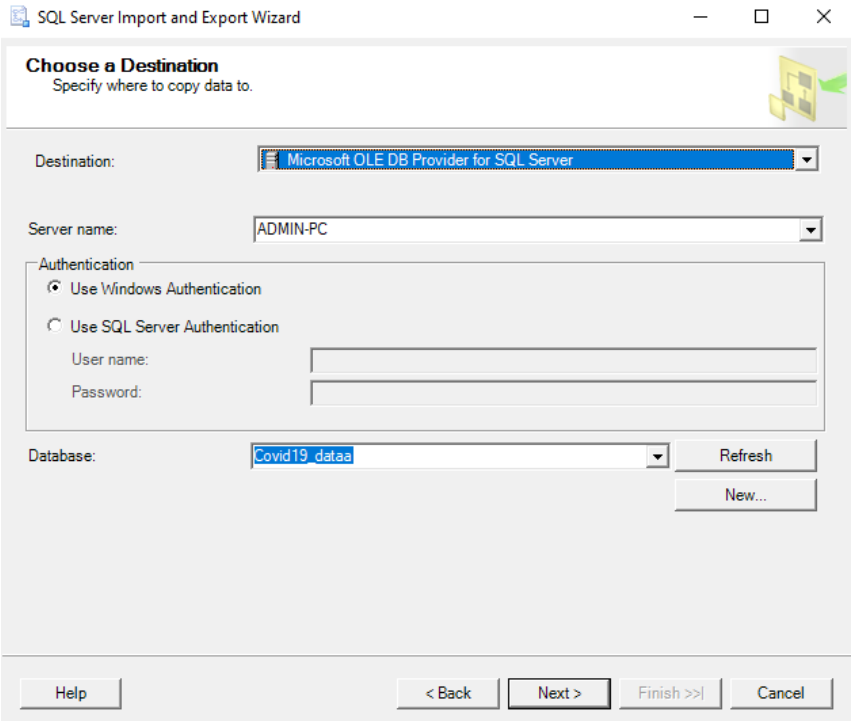


Tiếp theo, chọn Destination (Đích Đến) cần mapping dữ liệu với nhau.

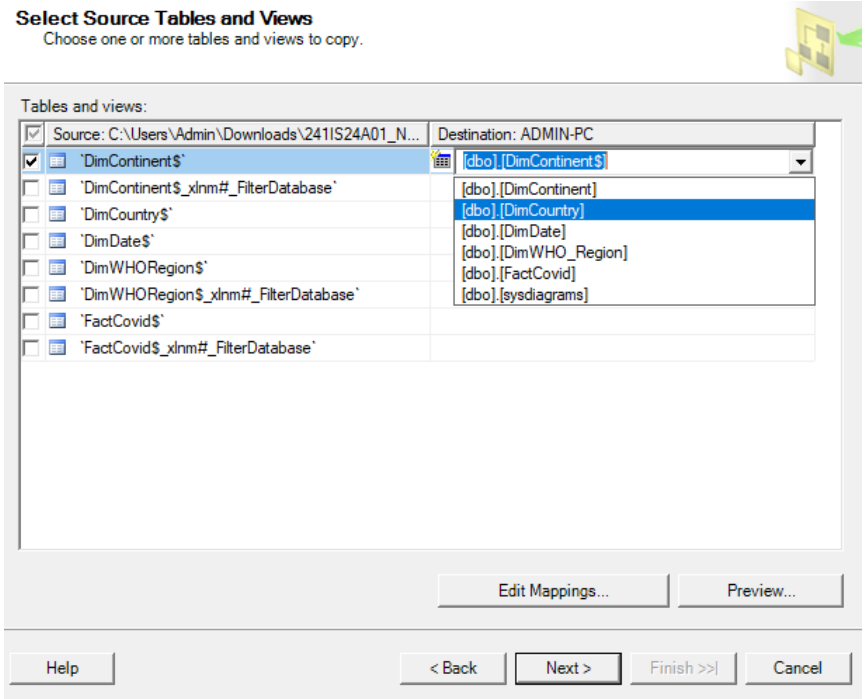
Server name: Nhập tên máy chủ SQL Server.

Authentication: Chọn Windows Authentication

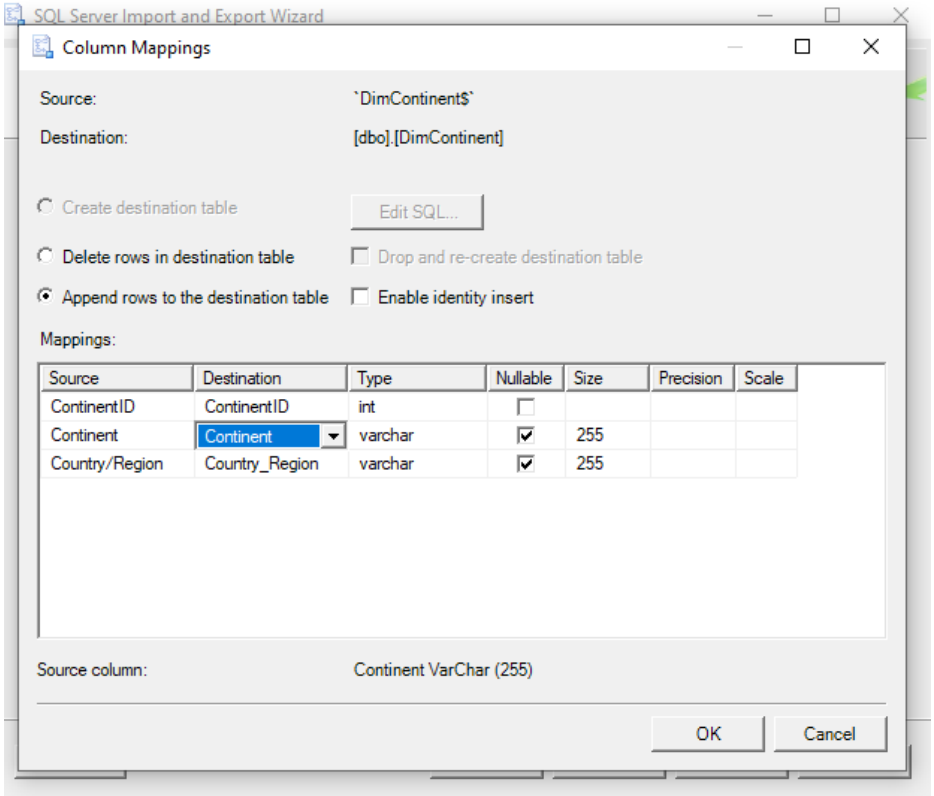
Database: Chọn cơ sở dữ liệu Covid19_dataa



Bước thứ 4, chọn Bảng Excel và Ánh Xạ vào Bảng SQL đã tạo. Chọn bảng hoặc sheet trong file Excel muốn nhập (VD: DimContinent) -> Chọn tên bảng SQL đích đến tương ứng đã tạo.



Sau đó, tiến hành kiểm tra ánh xạ các cột. Nhấp chọn Edit Mappings, kiểm tra từng cột từ file Excel đã được ánh xạ đúng với bảng SQL. Chỉnh sửa kiểu dữ liệu nếu cần (VD: Đảm bảo ContinentID là INT). Nhấp OK để tiếp tục.



Cuối cùng, xem lại và thực thi đổ dữ liệu. Nhấp **Next** và kiểm tra lại cấu hình. Chọn **Run immediately** và nhấp **Finish** để bắt đầu quá trình nhập. Import Wizard cung cấp cách thức nhanh chóng và dễ dàng để nhập dữ liệu từ Excel vào SQL Server. Với quy trình này, dữ liệu từ các file Excel đã được

đưa vào đúng bảng Dim và Fact, giúp nhóm hoàn thành bước Load trong quá trình ETL và sẵn sàng cho việc phân tích.

Chương V: Mô Hình Quan Hệ Giữa Nhu Cầu Phân Tích và Khung Nhìn Dữ Liệu

5.1. Phân tích nhu cầu phân tích của BI Users

5.1.1. Ban Điều Hành (Executive Board)

Mục tiêu phân tích:

Ban Điều Hành cần các báo cáo tổng quan để đưa ra các quyết định chiến lược. Các báo cáo này giúp họ theo dõi tình hình dịch bệnh trên toàn cầu, đánh giá hiệu quả chiến lược và phân bổ nguồn lực cho các khu vực quan trọng. Báo cáo cũng hỗ trợ trong việc quản lý tài chính và đưa ra các kế hoạch dài hạn về nghiên cứu và sản xuất vaccine.

Nhu cầu báo cáo:

Báo cáo tình hình dịch bệnh toàn cầu: Cung cấp cái nhìn tổng quan về số ca nhiễm, tử vong, hồi phục trên toàn cầu, phân bổ theo khu vực và quốc gia. Báo cáo giúp nhận diện các khu vực cần hỗ trợ và đánh giá mức độ ảnh hưởng của dịch bệnh đối với các chiến lược của công ty.

Dữ liệu sử dụng: FactCovid: confirmed, deaths, recovered; dimcontinent: continent

Mô hình báo cáo: Biểu đồ cột hoặc đường theo thời gian, thể hiện tổng số ca và xu hướng thay đổi qua các ngày/tuần/tháng.

Báo cáo phân bổ tài nguyên toàn cầu: Tổng hợp thông tin về phân bổ vaccine và thuốc, giúp ban điều hành đưa ra quyết định về việc phân phối nguồn lực. Báo cáo giúp theo dõi sự phân bố tài chính và hiệu quả chi tiêu.

Dữ liệu sử dụng: dimcountry: country, FactCovid: confirmed, deaths, recovered

Mô hình báo cáo: Biểu đồ phân bổ theo khu vực quốc gia hoặc bản đồ nhiệt thể hiện tình trạng phân bổ vaccine.

5.1.2. Giám Đốc Tài Chính (CFO)

Mục tiêu phân tích:

CFO sử dụng BI để theo dõi và quản lý ngân sách liên quan đến các hoạt động chống dịch, từ đó tối ưu hóa chi phí và tài chính. Báo cáo giúp CFO đánh giá hiệu quả tài chính của các chiến dịch chống dịch và dự báo các khoản chi phí trong tương lai.

Nhu cầu báo cáo:

Báo cáo tình hình tài chính toàn cầu: Giám sát chi phí và các khoản đầu tư vào các khu vực có tỷ lệ nhiễm bệnh cao. Báo cáo cung cấp thông tin để CFO điều chỉnh ngân sách và đảm bảo tài chính được phân bổ hợp lý.

Dữ liệu sử dụng: dimcountry: country, FactCovid: confirmed, deaths, recovered

Mô hình báo cáo: Biểu đồ cột hoặc tròn cho phép phân bổ chi phí, đầu tư và tài chính theo các khu vực.

Báo cáo dự báo tài chính: Sử dụng các mô hình dữ liệu để dự báo tình hình tài chính dựa trên sự phát triển của dịch bệnh, giúp CFO chuẩn bị ngân sách cho các chiến dịch tương lai.

Dữ liệu sử dụng: FactCovid: newcases, newdeaths; dimdate: date

Mô hình báo cáo: Biểu đồ dự báo tài chính với các số liệu về số ca nhiễm và tử vong trong các tháng tới.

5.1.3. Giám Đốc Kinh Doanh (Sales Director)

Mục tiêu phân tích:

Giám đốc kinh doanh sử dụng BI để đánh giá tình hình bán hàng của vaccine và thuốc, từ đó đưa ra các chiến lược tăng trưởng doanh thu. Các báo cáo giúp theo dõi thị trường và cải thiện các chiến lược tiếp cận khách hàng.

Nhu cầu báo cáo:

Báo cáo doanh thu theo khu vực: Theo dõi doanh thu từ các khu vực và quốc gia, giúp đánh giá sự hiệu quả của chiến lược kinh doanh. Báo cáo này cũng giúp xác định các khu vực có tiềm năng lớn và các khu vực cần sự can thiệp thêm.

Dữ liệu sử dụng: dimcountry: country, FactCovid: confirmed, deaths, recovered

Mô hình báo cáo: Biểu đồ cột hoặc đường thể hiện doanh thu theo từng quốc gia và khu vực.

Báo cáo hiệu quả chiến lược tiếp thị: Cung cấp cái nhìn về hiệu quả của các chiến lược tiếp thị và quảng bá vaccine, thuốc trong các thị trường mục tiêu.

Dữ liệu sử dụng: dimcountry: country, FactCovid: newcases, newdeaths, newrecovered

Mô hình báo cáo: Biểu đồ sự thay đổi theo thời gian giúp theo dõi tác động của các chiến dịch marketing đối với doanh thu và số ca sử dụng vaccine.

5.1.4. Phòng Phân Tích Dữ Liệu (Data Analytics Department)

Mục tiêu phân tích:

Phòng phân tích dữ liệu là nơi thu thập, xử lý và phân tích dữ liệu từ các nguồn khác nhau. Các báo cáo giúp cung cấp cái nhìn chi tiết về tình hình dịch bệnh, hỗ trợ các phòng ban khác trong tổ chức.

Nhu cầu báo cáo:

Báo cáo tình hình dịch bệnh theo thời gian: Phân tích số ca nhiễm, tử vong và hồi phục theo ngày, tuần, tháng để theo dõi sự thay đổi của dịch bệnh và đưa ra các dự báo về xu hướng dịch bệnh trong tương lai.

Dữ liệu sử dụng: FactCovid: newcases, newdeaths, newrecovered; dimdate: date

Mô hình báo cáo: Biểu đồ đường hoặc cột thể hiện sự thay đổi qua các ngày, tuần, tháng.

Báo cáo dự báo xu hướng dịch bệnh: Sử dụng các mô hình dữ liệu để dự báo các xu hướng dịch bệnh trong các tháng tiếp theo, từ đó hỗ trợ các phòng ban khác trong việc lập kế hoạch ứng phó.

Dữ liệu sử dụng: FactCovid: confirmed, newcases, newdeaths; dimdate: year, dimWHO_region: whoregion

Mô hình báo cáo: Biểu đồ dự báo xu hướng sử dụng các mô hình phân tích tăng trưởng dịch bệnh.

5.1.5. Phòng Kế Hoạch (Planning Department)

Mục tiêu phân tích:

Phòng kế hoạch chịu trách nhiệm phát triển các chiến lược dài hạn dựa trên tình hình dịch bệnh và các yếu tố bên ngoài. Các báo cáo giúp phòng này đưa ra các kế hoạch ứng phó kịp thời và điều chỉnh chiến lược.

Nhu cầu báo cáo:

Báo cáo tình hình dịch bệnh theo khu vực: Phân tích số ca mắc và tử vong theo các quốc gia, khu vực để hỗ trợ việc lên kế hoạch chiến lược cho các hoạt động ứng phó.

Dữ liệu sử dụng: FactCovid: confirmed, deaths; dimcountry: country

Mô hình báo cáo: Bản đồ nhiệt hoặc biểu đồ phân bổ theo khu vực.

Báo cáo triển khai chiến lược ứng phó: Đánh giá tình hình thực hiện các chiến lược ứng phó tại các khu vực và quốc gia, giúp điều chỉnh kế hoạch trong thời gian thực.

Dữ liệu sử dụng: FactCovid: newcases, newdeaths; dimWHO_region: whoregion

Mô hình báo cáo: Biểu đồ thanh hoặc cột theo khu vực để đánh giá mức độ triển khai chiến lược.

5.1.6. Quản Lý Chuỗi Cung Ứng (Supply Chain Management)

Mục tiêu phân tích:

Quản lý chuỗi cung ứng cần theo dõi tình trạng tồn kho và phân phối vaccine, thuốc để đảm bảo sự cung ứng ổn định và kịp thời, đáp ứng nhu cầu của các khu vực có tình hình dịch bệnh phức tạp.

Nhu cầu báo cáo:

Báo cáo tồn kho vaccine và thuốc: Theo dõi tình trạng tồn kho của vaccine và thuốc tại các khu vực khác nhau, giúp điều phối cung ứng đúng thời điểm và tránh tình trạng thiếu hụt.

Dữ liệu sử dụng: dimcountry: country, FactCovid: confirmed, deaths

Mô hình báo cáo: Biểu đồ cột hoặc biểu đồ phân bổ thể hiện tình trạng tồn kho theo từng khu vực.

Báo cáo phân phối vaccine và thuốc: Cung cấp thông tin về tiến độ phân phối vaccine và thuốc cho các khu vực cần thiết, theo dõi thời gian và mức độ sẵn có của nguồn cung.

Dữ liệu sử dụng: FactCovid: confirmed, deaths; dimWHO_region: whoregion

Mô hình báo cáo: Biểu đồ phân phối thể hiện tiến độ và số lượng vaccine đã phân phối, giúp đánh giá hiệu quả của chuỗi cung ứng.

5.2 Khung nhìn dữ liệu (Data Views)

Trong mô hình Business Intelligence (BI), việc xây dựng các khung nhìn dữ liệu (Data Views) đóng vai trò quan trọng, giúp tối ưu hóa cách thức trình bày và phân tích thông tin. Khung nhìn dữ liệu được thiết kế dựa trên nhu cầu báo cáo của từng nhóm người dùng, tập trung vào việc trình bày thông tin một cách trực quan, dễ hiểu và hỗ trợ ra quyết định hiệu quả.

5.2.1. Báo cáo Xu Hướng Thời Gian (Time Trend Analysis)

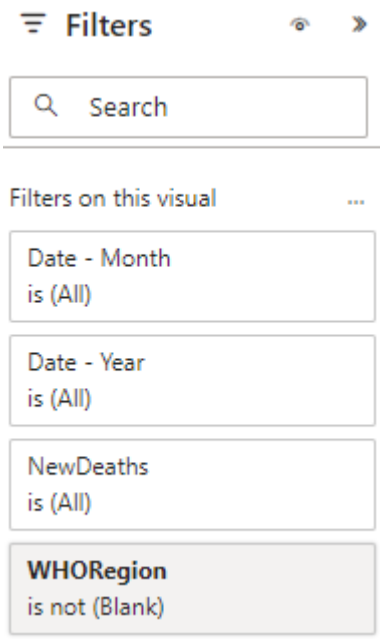
Mục tiêu:

Cung cấp cái nhìn toàn diện về sự thay đổi của các chỉ số quan trọng như số ca nhiễm, tử vong, và hồi phục trong từng giai đoạn. Báo cáo giúp theo dõi xu hướng dịch bệnh, đánh giá tác động của các biện pháp can thiệp và dự đoán tình hình trong tương lai.

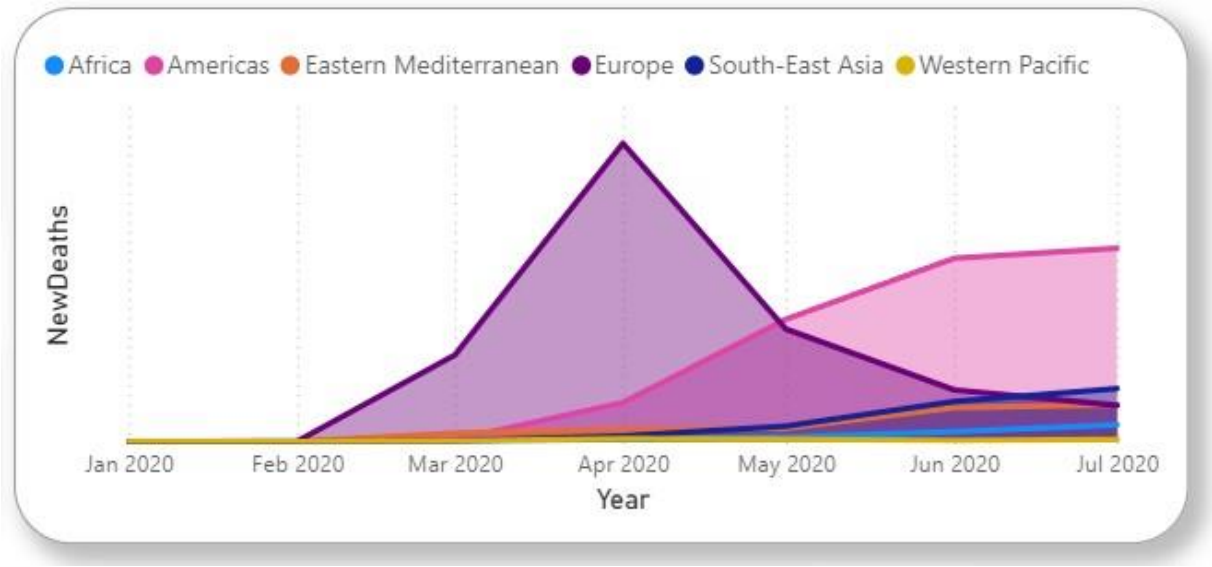
Thiết kế khung nhìn: Biểu đồ đường số ca tử vọng mới (Line Chart)

Nội dung hiển thị:

Trục X: Các mốc thời gian (ngày, tháng, năm).	
Trục Y: Số lượng ca tử vong mới	
Legend	

Tính năng bổ sung: Bộ lọc thời gian (Time Filter) để người dùng chọn giai đoạn cụ thể.	

Kết quả:



Hình 9. Kết quả báo cáo Time Trend Analysis

5.2.2. Báo cáo Tổng số ca tử vong theo khu vực (Deaths by Regions)

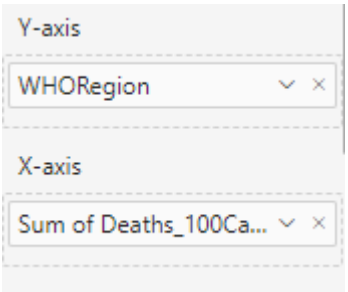
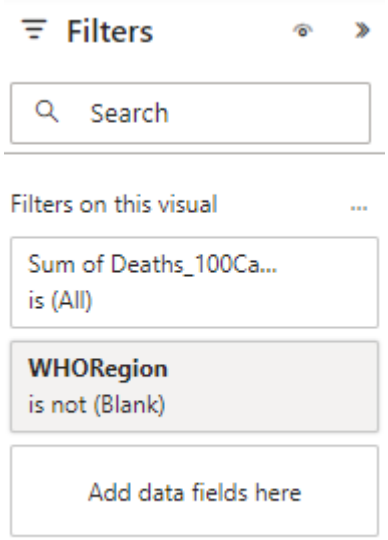
Mục tiêu:

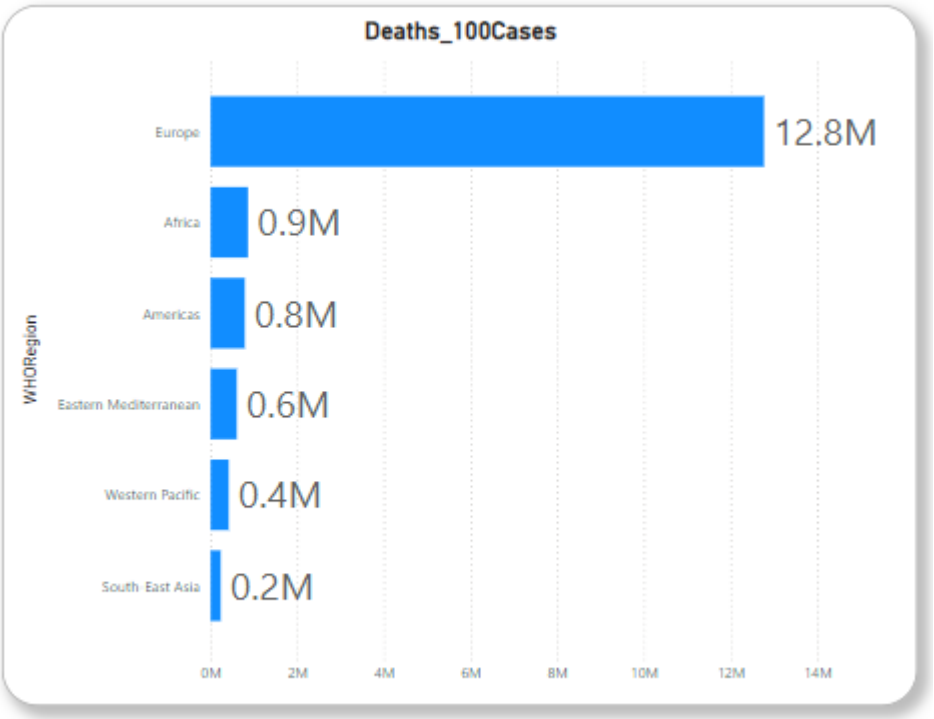
Theo dõi và so sánh số lượng ca tử vong do Covid-19 ở các khu vực trên toàn cầu; từ đó, giúp Giám Đốc Kinh Doanh (Sales Director) đưa ra nhận định về tình trạng, nhu cầu của các quốc gia theo khu vực, đánh giá hiệu quả chiến lược kinh doanh và xác định các thị trường tiềm năng để phân phối thuốc vaccine đến các khu vực đó.

Thiết kế khung nhìn: Biểu đồ so sánh (Pie Chart, Bar Chart):

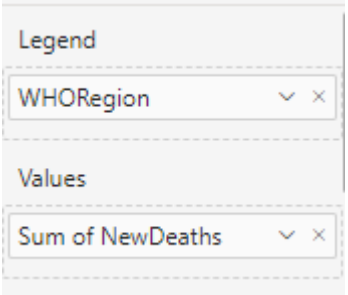
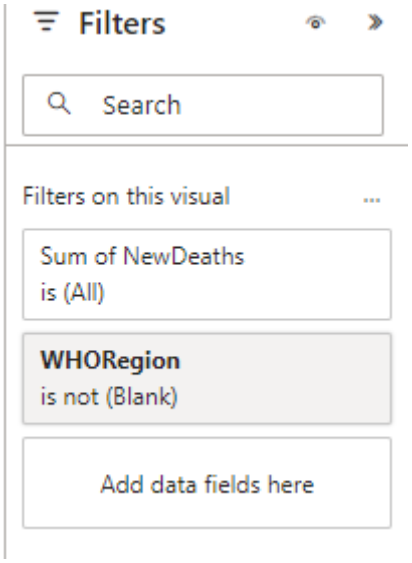
Nội dung hiển thị: Tỷ lệ tổng số ca tử vong trên 1 triệu dân, tỷ lệ tổng số ca tử vong trên 1 triệu ca mắc mới

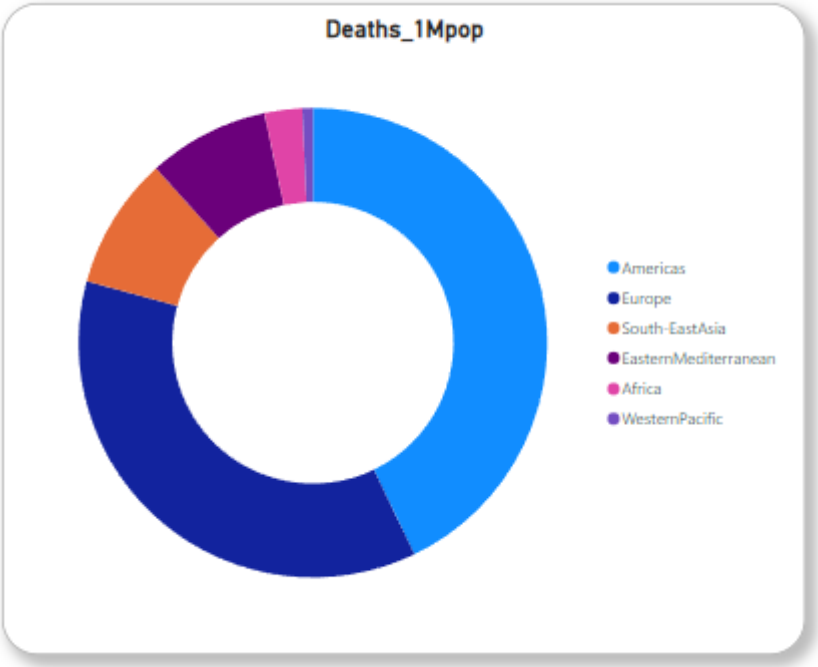
Ứng dụng: Giúp định hình tình thình tỷ lệ lây nhiễm các khu vực để đưa rhiến lược tiếp thị và phân phối sản phẩm vaccine phù hợp với từng thị trường.

Trục X: Tổng số ca tử vong trên 1 triệu dân Trục Y: Phân theo khu vực WHO	
Tính năng bổ sung: Bộ lọc các khu vực	



Hình 10. Biểu đồ Bar Chart

Legend: Theo các khu vực WHO Giá trị so sánh: Các ca tử cọng mới	
Tính năng bổ sung: Bộ lọc theo khu vực	



Hình 11. Biểu đồ Donut Chart

Chương VI: Đánh Giá Nền Tảng BI và Trực Quan Hóa Dữ Liệu

6.1. So sánh, lựa chọn nền tảng BI phù hợp

Pfizer, một tập đoàn dược phẩm hàng đầu thế giới, đang đối mặt với nhu cầu tối ưu hóa quy trình quản trị kinh doanh thông qua việc sử dụng nền tảng kinh doanh thông minh (BI). Trong bối cảnh đó, việc lựa chọn nền tảng BI phù hợp đóng vai trò quan trọng để nâng cao hiệu quả hoạt động.

Power BI, Tableau, và Oracle BI là ba công cụ hàng đầu trong lĩnh vực trực quan hóa dữ liệu và phân tích kinh doanh. Mỗi công cụ có những điểm mạnh và điểm yếu riêng, phù hợp với từng loại hình doanh nghiệp và mục đích sử dụng.

Power BI nổi bật với chi phí thấp và giao diện thân thiện, dễ dàng sử dụng ngay cả với người dùng mới làm quen với các công cụ phân tích dữ liệu. Được tích hợp chặt chẽ với hệ sinh thái Microsoft, Power BI là lựa chọn lý tưởng cho doanh nghiệp đã sử dụng Excel, Teams, hoặc Azure. Ngoài ra, công cụ này còn hỗ trợ các tính năng phân tích nâng cao, như AI và Machine Learning, giúp cải thiện hiệu quả ra quyết định. Tuy nhiên, Power BI có thể không phải là lựa chọn tối ưu nếu cần tạo các biểu đồ trực quan phức tạp hoặc xử lý lượng dữ liệu cực lớn.

Tableau lại là sự lựa chọn mạnh mẽ cho các doanh nghiệp muốn tập trung vào trực quan hóa dữ liệu. Với giao diện trực quan và khả năng hỗ trợ nhiều biểu đồ sáng tạo, Tableau cho phép người dùng phân tích dữ liệu chi tiết và thể hiện kết quả một cách trực quan nhất. Dù vậy, chi phí cao và yêu cầu kỹ năng sử dụng tốt hơn khiến Tableau phù hợp hơn với các doanh nghiệp vừa và lớn.

Oracle BI, trong khi đó, mạnh mẽ trong việc xử lý dữ liệu quy mô lớn và tích hợp sâu với hệ thống Oracle. Công cụ này thường được sử dụng bởi các doanh nghiệp lớn, cần báo cáo và phân tích phức tạp. Tuy nhiên, mức chi phí cao, mô hình license phức tạp, và giao diện kém thân thiện khiến Oracle BI trở thành lựa chọn không phù hợp cho các doanh nghiệp nhỏ hoặc vừa.

Qua đây, nhóm đưa ra đánh giá cho thấy Power BI là lựa chọn lý tưởng cho các doanh nghiệp tìm kiếm sự cân bằng giữa chi phí và hiệu quả, đặc biệt nếu đã quen thuộc với hệ sinh thái Microsoft. Trong khi đó, Tableau lại thích hợp cho những doanh nghiệp chú trọng vào trực quan hóa dữ liệu chuyên sâu, đem lại giao diện đẹp mắt và khả năng phân tích sâu. Oracle BI chỉ thực sự phát huy hiệu quả với các tổ chức quy mô lớn, có ngân sách và hạ tầng IT mạnh, cụ thể đồng bộ và thống nhất khi tích hợp với các hệ thống ERP lớn như Oracle Cloud, nhưng chi phí cao hơn nhiều so với hai nền tảng kia.

Xét đến nhu cầu và quy mô của Pfizer, nhóm đánh giá Power BI là lựa chọn tích hợp nhất. Sự kết hợp giữa tính linh hoạt, chi phí hợp lý, và khả năng tích hợp tốt giúp Power BI trở thành công cụ đắc lực trong quy trình quản trị dữ liệu của doanh nghiệp.

Hình 12. Power BI

Tiêu chí	Power BI	Tableau	Oracle BI
Chi phí	Giá khởi điểm thấp	Trung bình - Cao	Rất cao: Tốn kém với mô hình license phức tạp.
Giao diện người dùng	Thân thiện: Trực quan, dễ học với người quen dùng Excel và sản phẩm MS.	Rất trực quan, nhưng cần thời gian học với người mới.	Cổ điển, giao diện phức tạp, khó làm quen.
Hiệu năng	Tốt: Phù hợp với báo cáo, phân tích dữ liệu vừa và lớn.	Rất mạnh: Xử lý dữ liệu lớn nhanh và hiệu quả, tối ưu cho trực quan hóa.	Trung bình: Tập trung vào quy mô lớn nhưng tốc độ chậm với tập nhỏ.
Tích hợp hệ sinh thái	Xuất sắc: Tích hợp chặt với Microsoft (Excel, Azure, Teams...).	Tốt: Hỗ trợ nhiều nguồn dữ liệu, nhưng không tối ưu với MS.	Trung bình: Tốt với các công cụ Oracle, nhưng kém linh hoạt bên ngoài.
Trực quan hóa dữ liệu	Mạnh mẽ: Cung cấp các biểu đồ thông dụng và AI-driven analytics.	Xuất sắc: Hỗ trợ biểu đồ nâng cao, trực quan sáng tạo và đa dạng.	Hạn chế: Chủ yếu tập trung vào biểu đồ cơ bản.
Khả năng phân tích	Tốt: Hỗ trợ phân tích dự đoán với AI, ML tích hợp sẵn.	Mạnh: Phân tích chi tiết với tính linh hoạt cao.	Trung bình: Thiên về xử lý OLAP và báo cáo doanh nghiệp hơn phân tích.
Khả năng mở rộng	Linh hoạt: Phù hợp với cả doanh nghiệp nhỏ và lớn.	Tốt: Phù hợp doanh nghiệp vừa và lớn.	Tập trung lớn: Tốt với tổ chức quy mô rất lớn.
Khả năng sử dụng mobile	Tốt: Có ứng dụng di động đơn giản, dễ sử dụng.	Xuất sắc: Mobile dashboard linh hoạt, tương tác tốt.	Hạn chế: Ít tính năng cho thiết bị di động.

Hỗ trợ kỹ thuật	Mạnh mẽ: Cộng đồng lớn, tài liệu phong phú, hỗ trợ từ Microsoft.	Tốt: Hỗ trợ từ Tableau, cộng đồng người dùng rộng rãi.	Tốt: Hỗ trợ chuyên sâu nhưng chi phí cao.
Phí ẩn	Thấp: Không có phí triển khai phức tạp.	Trung bình: Tăng cao khi mở rộng số lượng người dùng.	Rất cao: Phí triển khai, duy trì, và nâng cấp tốn kém.

Bảng 1. So sáng các công cụ trực quan

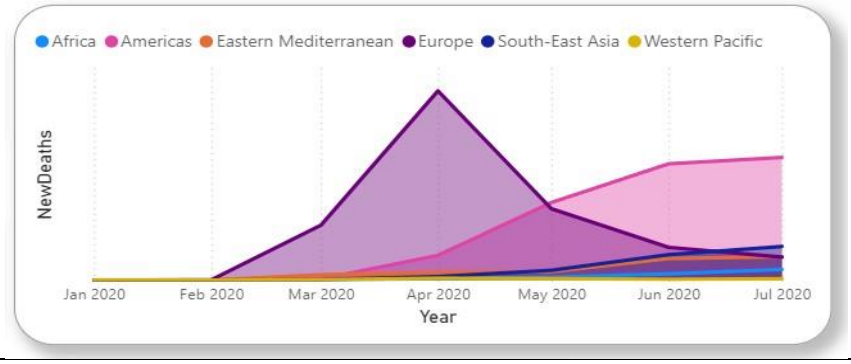
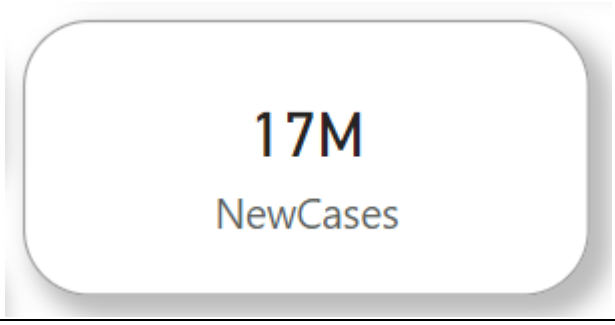
6.2. Trực quan hóa dữ liệu

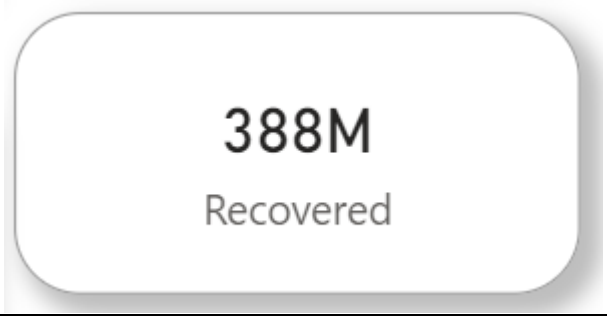
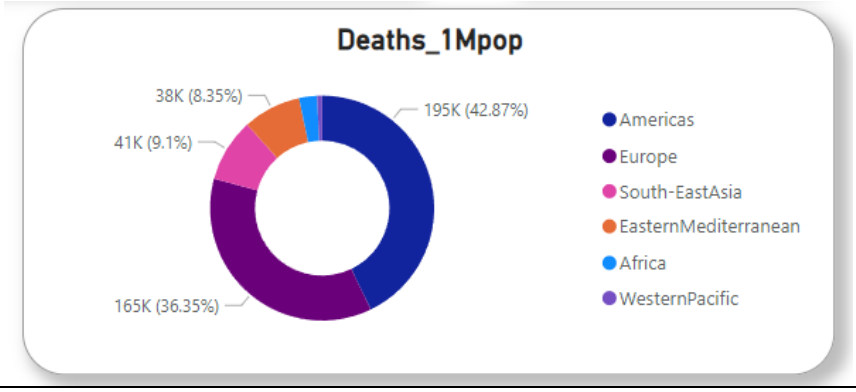
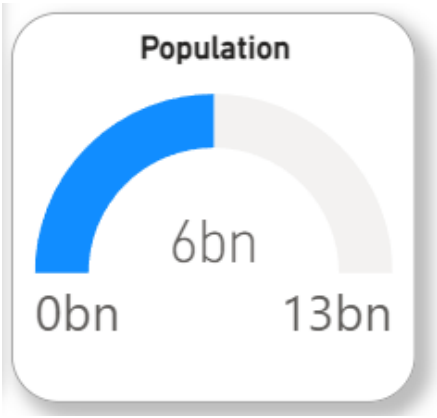
Trực quan hoá dữ liệu là bước quan trọng giúp truyền tải thông tin dễ hiểu và nhanh chóng đến người dùng. Với Power BI, Pfizer có thể thiết kế những dashboard mạnh mẻ, linh hoạt, phù hợp với nhu cầu quản trị và theo dõi của tập đoàn.

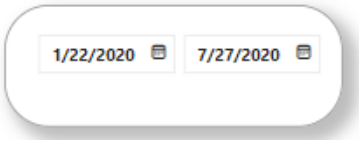
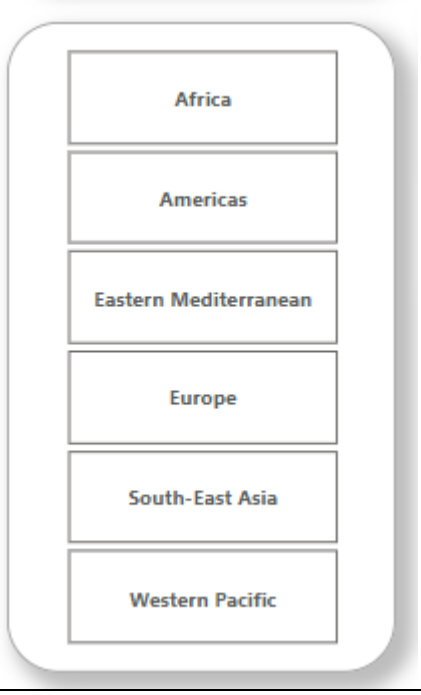
Dashboard doanh thu giúp hiển thị thông tin doanh thu theo khu vực và thời gian, sử dụng các biểu đường và cột để phân tích xu hướng và so sánh doanh thu giữa các khu vực. Bản đồ nhiệt, một biểu đồ trực quan hiện đại, phân bố số ca nhiễm theo quốc gia giúp nhìn nhận ngay lập tức tình hình lây lan dịch bệnh. Cuối cùng, dashboard hiệu quả sản xuất cung cấp thông tin tồn kho, chi phí sản xuất, và năng suất, cho phép quản lý hiệu quả nguồn lực.

Nhờ sự linh hoạt và sâu sát trong khả năng trực quan hoá, Power BI đáp ứng mọi nhu cầu của Pfizer về phân tích và quản trị

Thiết kế các dashboard:

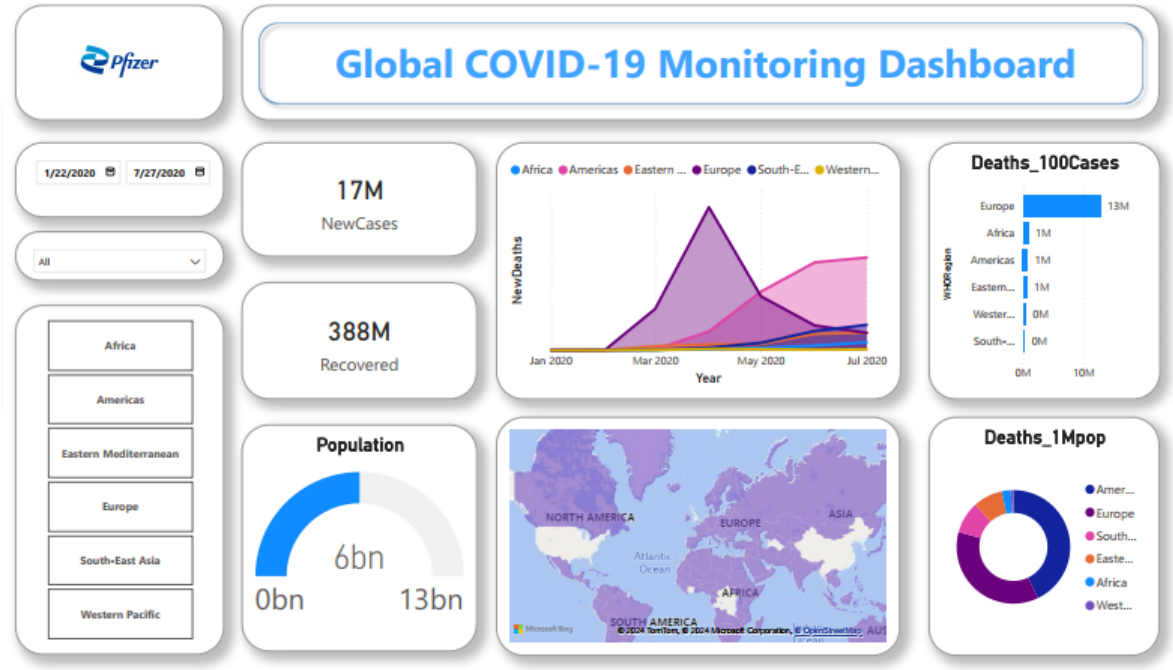
Xu hướng dịch bệnh theo khu vực và thời gian	
Tổng số ca nhiễm mới toàn cầu trong 7 tháng năm 2020	

Tổng số ca phục hồi trong 7 tháng đầu năm 2020	
Tỷ lệ tử vọng: Số lượng ca tử vong trên 1 triệu dân	
Bản đồ toàn cầu hiển thị số ca tử vong theo màu sắc, kích thước vùng.	
Tổng dân số toàn cầu ở các quốc gia được khảo sát	
Lọc theo thời gian	

	
Lọc theo khu vực	

Bảng 2. Bảng chi tiết các biểu đồ

Tổng quan dashboard



Hình 13. Tổng quan dashboard

Dashboard được nhóm thiết kế trên Power BI nhằm theo dõi tình hình COVID-19 toàn cầu từ tháng 1/2020 đến tháng 7/2020, với sự phân loại theo khu vực WHO (World Health Organization).

6.2.1. Tổng quan các chỉ số chính

New Cases (Ca nhiễm mới): Chỉ số này hiển thị tổng số ca nhiễm mới được báo cáo trên toàn cầu. Đây là một thước đo quan trọng để đánh giá tốc độ lây lan trong từng giai đoạn.

Recovered (Ca hồi phục): Số lượng ca hồi phục cho thấy khả năng ứng phó và chữa trị của hệ thống y tế toàn cầu..

New Deaths (Ca tử vong mới): Biểu đồ thể hiện xu hướng số ca tử vong trong từng giai đoạn. Sự tăng giảm của chỉ số này cần được so sánh với số ca nhiễm để đánh giá mức độ nghiêm trọng

6.2.2. Biểu đồ xu hướng theo thời gian

Biểu đồ đường (line chart) hiển thị số ca nhiễm mới và tử vong theo thời gian từ tháng 1/2020 đến tháng 7/2020.

Giai đoạn từ tháng 3 đến tháng 5/2020 chứng kiến sự gia tăng mạnh số ca nhiễm ở nhiều khu vực, đặc biệt là châu Âu và châu Mỹ. Sau tháng 5/2020, số ca hồi phục tăng lên, cho thấy hiệu quả của các biện pháp kiểm soát và điều trị.

6.2.3. Biểu đồ phân loại khu vực WHO

Dashboard chia dữ liệu theo các khu vực WHO như Africa, Americas, Europe, Eastern Mediterranean, South-East Asia, và Western Pacific.

Europe và Americas: Hai khu vực có số ca nhiễm và tử vong cao nhất, cho thấy mức độ bùng phát nghiêm trọng.

Africa và South-East Asia: Số liệu tương đối thấp hơn, có thể do dân số trẻ hoặc hạn chế trong việc xét nghiệm.

Các chỉ số như Deaths per 100 Cases hoặc Deaths per 1M Population cung cấp bức tranh tổng quan về mức độ ảnh hưởng so với quy mô dân số.

6.2.4. Bản đồ thế giới

Dashboard sử dụng bản đồ để hiển thị tình hình toàn cầu. Đây là cách trình bày trực quan, giúp người dùng dễ dàng xác định vị trí các khu vực bị ảnh hưởng nặng nề nhất.

Bản đồ sử dụng các chấm màu để biểu thị số lượng ca nhiễm, nhưng không làm nổi bật các tỷ lệ tương đối (như số ca nhiễm trên 1 triệu dân). Các khu vực như châu Mỹ và châu Âu có mật độ ca nhiễm cao, được thể hiện rõ trên bản đồ.

Đánh giá tính trưc quan dữ liệu trên công cụ Power BI

Chỉ tiêu đánh giá	Nội dung
Tính trực quan cao	Dashboard tích hợp bản đồ, giúp người dùng hình dung nhanh về tình hình các khu vực khác nhau. Các biểu đồ đường (line charts) biểu diễn dữ liệu thời gian, giúp theo dõi xu hướng trong giai đoạn từ tháng 1/2020 đến tháng 7/2020.
Phân loại chi tiết	Dữ liệu được nhóm theo từng khu vực WHO như Africa, Americas, Europe,..tạo điều kiện dễ dàng cho việc phân tích theo từng khu vực
Tích hợp nguồn dữ liệu đáng tin cậy	Nguồn đảm bảo tính chính xác và độ tin cậy của dữ liệu.

Bảng 3. Đánh giá tính trực quan dữ liệu trên PowerBI

6.3. Chính sách đảm bảo dữ liệu

Quyền truy cập theo cấp bậc: Áp dụng phân quyền dựa trên vai trò (Role-Based Access Control). Đảm bảo chỉ những nhân viên được ủy quyền mới có quyền truy cập các thông tin nhạy cảm.

Bảo mật dữ liệu nhạy cảm: Mã hóa dữ liệu (Encryption) trong quá trình lưu trữ và truyền tải. Ẩn danh các dữ liệu cá nhân để tuân thủ quy định về quyền riêng tư.

Tuân thủ quy định (như GDPR): Đảm bảo dữ liệu cá nhân được xử lý đúng mục đích và lưu trữ trong thời gian cần thiết. Cung cấp công cụ cho phép khách hàng kiểm soát và yêu cầu xóa dữ liệu cá nhân.

Cấu trúc trên sẽ giúp Pfizer xây dựng một hệ thống kho dữ liệu toàn diện và tuân thủ các tiêu chuẩn quốc tế về bảo mật dữ liệu.

Chương VII: Đánh Giá và Khuyến Nghị

7.1. Đánh giá hệ thống báo cáo

Hệ thống báo cáo được nhóm xây dựng đã đạt được những kết quả đáng khích lệ trong việc hỗ trợ các nhà quản lý (BI Users) của Pfizer đưa ra quyết định kịp thời và tối ưu hóa hoạt động doanh nghiệp. Cụ thể:

Hiệu quả hỗ trợ ra quyết định: Các báo cáo cung cấp dữ liệu chi tiết, đáng tin cậy và phân tích chuyên sâu, giúp các nhà quản lý đánh giá chính xác xu hướng thị trường, hiệu quả chuỗi cung ứng và tài chính. Điều này không chỉ giúp tăng cường hiệu suất kinh doanh mà còn hỗ trợ Pfizer định hướng các chiến lược phát triển dài hạn.

Tối ưu hóa hoạt động: Nhờ tích hợp dữ liệu từ nhiều nguồn khác nhau, hệ thống báo cáo đã cải thiện hiệu quả hoạt động chuỗi cung ứng và phân phối sản phẩm, đặc biệt là trong giai đoạn đại dịch COVID-19, khi việc đáp ứng nhanh nhu cầu thị trường là yếu tố sống còn.

Bài học thu được:

Sự quan trọng của việc tích hợp dữ liệu và phân tích chuyên sâu trong việc đưa ra quyết định.

Quy trình xây dựng hệ thống báo cáo cần có sự phối hợp chặt chẽ giữa các nhóm chức năng và quản lý.

Ý nghĩa đối với doanh nghiệp: Hệ thống báo cáo không chỉ đóng vai trò là công cụ hỗ trợ ra quyết định, mà còn là nền tảng để cải tiến liên tục và đổi mới chiến lược trong môi trường kinh doanh biến động.

7.2. Khuyến nghị

Để nâng cao hơn nữa giá trị của hệ thống báo cáo, nhóm đưa ra các khuyến nghị sau:

Nâng cao chất lượng dữ liệu với AI và tự động hóa:

Sử dụng trí tuệ nhân tạo (AI) để làm sạch và chuẩn hóa dữ liệu, giảm thiểu lỗi sai và tăng độ chính xác.

Tự động hóa quy trình thu thập và phân tích dữ liệu, giúp tiết kiệm thời gian và tài nguyên.

Khai thác tốt hơn giá trị dữ liệu:

Ứng dụng các công cụ phân tích dự đoán để dự báo xu hướng thị trường và nhu cầu sản phẩm.

Phát triển thêm các báo cáo động, cho phép nhà quản lý tương tác với dữ liệu theo thời gian thực.

Tăng cường an toàn và bảo mật:

Áp dụng mã hóa dữ liệu toàn diện, bảo vệ thông tin nhạy cảm của doanh nghiệp và khách hàng.

Xây dựng cơ chế giám sát hệ thống chặt chẽ, phát hiện sớm các hành vi truy cập bất thường.

Những khuyến nghị trên không chỉ giúp hệ thống báo cáo phát huy tối đa tiềm năng, mà còn đảm bảo đáp ứng tốt các yêu cầu bảo mật và hiệu quả trong bối cảnh công nghệ đang thay đổi nhanh chóng.

KẾT LUẬN

Hệ thống Business Intelligence (BI) và kho dữ liệu đóng vai trò then chốt trong việc nâng cao hiệu quả kinh doanh và hỗ trợ Pfizer đáp ứng các nhu cầu y tế toàn cầu. Với sự phát triển mạnh mẽ của công nghệ phân tích dữ liệu, Pfizer đã triển khai một hệ thống BI toàn diện, giúp tối ưu hóa quá trình ra quyết định và hỗ trợ các nhóm người dùng ở các cấp độ khác nhau có thể tiếp cận và phân tích thông tin một cách chính xác và kịp thời. Điều này không chỉ giúp công ty cải thiện hiệu quả kinh doanh mà còn nâng cao khả năng thích ứng với những biến động nhanh chóng của thị trường y tế toàn cầu.

Hệ thống BI tại Pfizer đã giúp các nhóm quản lý chiến lược, chiến thuật và tác nghiệp có thể đưa ra các quyết định nhanh chóng và chính xác hơn. Các giám đốc điều hành, nhà quản lý tài chính và các bộ phận khác trong công ty có thể sử dụng các báo cáo tổng hợp và chi tiết về tình hình tài chính, doanh thu, chi phí và thị trường để xác định chiến lược phát triển dài hạn và phân bổ nguồn lực hợp lý. Đặc biệt, giám đốc tài chính có thể dựa vào các dữ liệu phân tích tài chính từ BI để tối ưu hóa ngân sách, tối đa hóa lợi nhuận và dự đoán các rủi ro tài chính trong tương lai, từ đó giúp công ty duy trì sự ổn định và phát triển bền vững.

Ngoài ra, hệ thống BI của Pfizer còn hỗ trợ trong việc quản lý chuỗi cung ứng và tối ưu hóa quy trình sản xuất. Nhờ vào các báo cáo và phân tích dự báo từ BI, Pfizer có thể dự đoán các tình huống thiếu hụt hoặc dư thừa nguyên liệu, đồng thời theo dõi hiệu quả của các chiến lược phân phối sản phẩm, giúp cải thiện quá trình sản xuất và cung cấp các sản phẩm chất lượng cao cho khách hàng trên toàn thế giới. Điều này cũng góp phần vào việc giảm thiểu chi phí và tăng hiệu quả hoạt động trong toàn bộ chuỗi cung ứng.

Đặc biệt trong bối cảnh dịch bệnh toàn cầu, việc sử dụng hệ thống BI giúp Pfizer đưa ra các quyết định nhanh chóng và chính xác hơn trong việc phân bổ nguồn lực và sản xuất vaccine. Với khả năng theo dõi tình hình dịch bệnh tại các khu vực khác nhau, BI cho phép Pfizer đưa ra các chiến lược marketing và phân phối sản phẩm một cách linh hoạt, giúp công ty đáp ứng kịp thời nhu cầu y tế của các quốc gia bị ảnh hưởng bởi đại dịch. Việc phân tích số liệu về số ca nhiễm, số ca tử vong và tình trạng bệnh tật tại từng khu vực không chỉ giúp Pfizer điều chỉnh kế hoạch sản xuất mà còn hỗ trợ trong việc tối ưu hóa quy trình phân phối vaccine đến các khu vực có nhu cầu cao, đảm bảo sự công bằng và hiệu quả trong việc triển khai các chiến dịch phòng chống dịch bệnh.

Hệ thống BI của Pfizer cũng đóng vai trò quan trọng trong việc hỗ trợ nghiên cứu và phát triển (R&D) các sản phẩm mới. Các nhóm R&D có thể sử dụng BI để phân tích xu hướng thị trường, đánh giá phản hồi từ khách hàng và theo dõi kết quả nghiên cứu để phát triển các sản phẩm đáp ứng nhu cầu ngày càng tăng của người tiêu dùng. Dữ liệu từ các nghiên cứu y tế, phản hồi khách hàng và thông tin về các bệnh tật mới nổi giúp Pfizer cải tiến và phát triển các sản phẩm đáp ứng nhu cầu y tế toàn cầu một cách chính xác và hiệu quả.

Hệ thống BI và kho dữ liệu của Pfizer không chỉ nâng cao hiệu quả quản lý và ra quyết định mà còn là yếu tố then chốt giúp công ty duy trì sự cạnh tranh và phát triển bền vững trong ngành công nghiệp dược phẩm. Việc sử dụng các công cụ phân tích dữ liệu mạnh mẽ giúp Pfizer không chỉ cải thiện hiệu quả tài chính và quản lý chuỗi cung ứng mà còn đáp ứng kịp thời các nhu cầu y tế toàn cầu. Hệ thống này giúp Pfizer phát triển các sản phẩm y tế chất lượng cao, đồng thời tăng cường khả năng ứng phó với các tình huống khẩn cấp, như đại dịch, và đóng góp vào công cuộc chăm sóc sức khỏe cộng đồng trên toàn thế giới.

TÀI LIỆU THAM KHẢO

Charles H. Jones, Subha Madhavan, Kannan Natarajan, Michael Corbo, Jane M,True, Mikael Dolsten. (2024, September). Rewriting the textbook for pharma: how to adapt and thrive in a digital, personalized and collaborative world. trang 1-11.

edu, C. Được truy lục từ csc.edu.vn: https://csc.edu.vn/data-analysis/Phan-tich-va-truc-quan-hoa-du-lieu-voi-Power-BI-co-ban-va-nang-cao_223

Hanh, L. H. *Base.vn*. Được truy lục từ https://base.vn/blog/power-bi-la-gi/

Pfizer. *"Data and AI are Helping to Get Medicines to Patients Faster.* Được truy lục từ Pfizer: https://www.pfizer.com/sites/default/files/investors/financial_reports/annual_reports/2022/story/data-and-ai-are-helping-to-get-medicines-to-patients-faster/?form=MG0AV3

Pfizer. *mRNA and Artificial Intelligence for Advanced Vaccine Innovation*. Được truy lục từ Pfizer: https://www.pfizer.com/news/articles/how_a_novel_incubation_sandbox_helped_speed_up_data_analysis_in_pfizer_s_covid_19_vaccine_trial?form=MG0AV3

Pfizer. *Pfizer*. Được truy lục từ pfizer.com: https://www.pfizer.com.vn/vi

Pfizer. *Pfizer Is Using AI to Discover Breakthrough Medicines*. Được truy lục từ Pfizer: https://insights.pfizer.com/pfizer-is-using-ai-to-discover-breakthrough-medicines/?form=MG0AV3