**BÀI TẬP 2: CÀI ĐẶT BỔ SUNG CÔNG CỤ THỰC HIỆN**

**KHAI PHÁ DỮ LIỆU**

Lời cảm ơn

Đầu tiên, nhóm chúng em xin gửi lời cảm ơn đến Thầy – Th.S Nguyễn Văn Thành – Giảng viên phụ trách môn Kho Dữ Liệu – Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh vì những đóng góp của thầy trong thời gian qua. Thầy đã cung cấp những kiến thức, hướng dẫn, đóng góp, xây dựng trong quá trình học để nhóm em có thể hoàn thành đồ án môn học một cách tốt nhất.

Kho dữ liệu là một trong những kiến thức đang ngày càng phổ biến cùng với OLAP, nhóm chúng em đã lên ý tưởng thực hiện xây dựng một kho dữ liệu cùng với các quy trình nghiệp vụ. Nhóm chúng em quyết định xây dựng đồ án “Xây dựng và khai thác kho dữ liệu về lĩnh vực hàng không”. Trong quá trình thực hiện đồ án, vận dụng những kiến thức học được từ lý thuyết đến thực hành trong quá trình học tập, cùng với sự tìm hiểu các kiến thức, công cụ mới, nhóm chúng em đã cố gắng hoàn thiện đồ án một cách tốt nhất.

Nếu có những chỗ chưa hoàn thiện, nhóm chúng em mong nhận được sự đóng góp từ thầy nhằm cải thiện và coi đó là một kinh nghiệm để có thể hoàn thành tốt hơn.

Nhóm chúng em một lần nữa xin chân thành cảm ơn thầy!

Lời Nhận Xét Của Giảng Viên

…………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………….………………………………………………………………………………..………

Chương 1: Lý do chọn đề tài và giới thiệu tổng quan Dataset

* 1. Lý do hình thành đồ án

1.1.1 Vấn đề đặt ra

Ngành hàng không là một trong những ngành trọng điểm đối với thế giới cũng như Việt Nam. Trải qua lịch sử hơn 100 năm đầy biến động, ngành hàng không đã có những biến động cả về công nghệ lẫn dịch vụ.

Từ năm 1994 đến năm 2002 nổi lên với sự phát triển của mô hình hàng không giá rẻ, có ý nghĩa quan trọng với ngành hàng không ngày nay. Các công ty con và trung tâm phân phối được mở rộng toàn thế giới cùng với sự phát triển của thương mại điện tử cho các hệ thống bán vé máy bay trực tuyến.

Tháng 12/2019, tâm dịch Covid-19 nhanh chóng lan rộng ra toàn thế giới khiến tất cả các nước phải áp dụng các biện pháp hạn chế di chuyển, điểu này đặc biệt nghiêm trọng cho ngành hàng không thế giới. American Airlines cắt giảm 75% số chuyến bay quốc tế, Korean Air giảm 80% lịch trình,… tóm lại, hầu như các chuyến bay trên thế giới đều bị cắt giảm đáng kể.

Trong 2 năm vừa qua, ngành hàng không quốc tế cũng như Việt Nam đang có dấu hiệu phục hồi trở lại

Các kiến thức về ngành hàng không rất đa dạng như thông tin về các hãng hàng không trên toàn thế giới, các sân bay, chuyến bay,… và phải kể đến quy trình đặt vé máy bay hay còn gọi là booking. Những data kể trên rất phù hợp để xây dựng một kho dữ liệu về lịch trình chuyến bay cũng như booking.

Vì vậy, nhóm chúng em quyết định xây dựng một kho dữ liệu về các chuyến bay và đặt vé máy bay của ngành hàng không thế giới

<https://thinksmart.com.vn/hon-100-nam-ra-doi-va-phat-trien-nganh-hang-khong/>

1.1.2 Giải pháp

Dựa trên nhu cầu phân tích, thống kê và khai thác dữ liệu về ngành hàng không thế giới. Giải pháp được đưa ra là xây dựng một kho dữ liệu nhằm mục đích khai thác, phân tích và tạo báo cáo tổng về các chuyến bay và booking.

1.1.3 Mục tiêu và ý nghĩa dự án

Đầu tiêu là xây dựng một kho dữ liệu và OLAP

Data Warehouse hay kho dữ liệu lưu trữ data từ nhiều nguồn, giúp tăng cường hiệu suất báo cáo và phân tích. Dựa vào điều đó cũng như các kiến thức đã học được trong quá trình học tập, nhóm chúng em quyết định xây dựng một kho dữ liệu OLAP để thấy được hiệu quả của nó.

Thứ hai là khai thác và phân tích data về hàng không trong thực tế. Khai thác các thông tin của hàng không quốc tế nhằm có cái nhìn chung về kiến thức cũng như quy trình vận hành cơ bản của một hệ thống đặt vé máy bay.

1.2 Giới thiệu tổng quan dataset