Commercial Airline Flights Project

# 1. Mục tiêu, câu hỏi và nhiệm vụ nghiên cứu

## 1.1. Mục tiêu nghiên cứu

Dựa trên bộ dữ liệu liên quan đến các chuyến bay, địa điểm, giá vé của các hãng bay, việc phân tích bộ dữ liệu nhằm đưa ra các xu hướng, đặc điểm nổi bật của các hãng bay cũng như sở thích khách hàng nhằm đưa ra giải pháp cho các hãng bay.

## 1.2. Câu hỏi nghiên cứu

* Bộ dữ liệu bao gồm các thông tin gì, được phục vụ cho mục đích gì và có thể được khai thác như thế nào, giúp ích như thế nào?
* Bộ dữ liệu đã được làm sạch và đủ điều kiện để đưa vào phân tích hay chưa, đồng thời cần sử dụng thêm các biến số đo lường nào khác hay không?
* Bộ dữ liệu trên có thể được triển khai dưới dạng các bảng như thế nào và các bảng này có mối quan hệ với nhau như thế nào (mối quan hệ một-một, một-nhiều hay nhiều-nhiều)?
* Các dữ liệu sau khi được phân tích nên được trực quan hóa như thế nào để có thể được khai thác một cách hiệu quả nhất và có cần sử dụng thêm các measurements nào nữa hay không?
* Từ các bảng đã được vẽ lên, các thông tin gì và có những gợi ý gì được đưa ra dành cho các hãng bay về nhu cầu và sở thích của khách hàng?

## 1.3. Nhiệm vụ nghiên cứu

# 2. Mô tả giải pháp

## 2.1. Làm sạch bộ dữ liệu

- Bộ dữ liệu ban đầu gồm một file excel duy nhất, chưa được tách thành các bảng, chưa có các relationships, một số cột bị sai format (booking date). Các bước làm sạch dữ liệu bao gồm như sau:

* Xóa cột Traveler FirstName, TravelerLastName, Return Airport
* Chuẩn hóa các cột ngày sang dạng Date
* Gộp dữ liệu hai cột Return và One-way thành Ticket Type
* Chuẩn hóa lại dữ liệu các cột (xóa khoảng trắng, đổi tên)

## 2.2. Xây dựng mô hình

- Bộ dữ liệu được xử lý và tách thành 3 bảng, lần lượt là: DimOrigin, DimDestination và FactFlights. Bảng DimDate được tạo thêm để thuận tiện trong việc nghiên cứu dữ liệu theo thời gian (năm, quý, tháng). Các relationships giữa các bảng được tạo như sau:

* DimOrigin đến FactFlights: khóa chung Origin Airport, mối quan hệ một – nhiều.
* DimDestination đến FactFlights: khóa chung Destination Airport, mối quan hệ một – nhiều.
* DimDate đến FactFlights đến: khóa chung Booking Date – Date, mối quan hệ một-nhiều. Bên cạnh đó hai mối quan hệ được thiết lập thêm là Travel Start Date – Date và Return Date – Date, mối quan hệ một nhiều, không active để tạo các measure phân tích sau này.

- Các measure về doanh thu, tăng trưởng, số lượng vé theo năm, loại vé được tạo nhằm thuận tiện trong việc trực quan hóa dữ liệu

## 2.3. Trực quan hóa và phân tích bộ dữ liệu

### 2.3.1. Phân tích tổng quan

*Các biểu đồ phân tích*

* Tổng quan hoạt động bay
* Phân bổ hạng ghế
* Số lượng chuyến bay theo quý
* Chi tiêu hàng tháng cho vé máy bay

*Insight nổi bật*

* Số chuyến bay và chi phí giữa hai năm gần như không thay đổi, thể hiện sự ổn định trong hoạt động vận chuyển.
* Tỷ lệ phân bổ hạng ghế giữa hai năm gần như không đổi, duy trì sự cân bằng nhu cầu giữa các phân khúc khách hàng.
* Chi tiêu trung bình hàng tháng năm 2019 giảm nhẹ so với 2018, phù hợp với xu hướng giảm nhẹ 0.49% về tổng chi phí vé.
* Số chuyến bay quốc tế và nội địa đều giảm nhẹ trong năm 2019, tuy nhiên chi phí quốc tế lại tăng nhẹ, gợi ý giá vé quốc tế có thể tăng hoặc các chuyến bay quốc tế có chi phí cao hơn.

*Kiến nghị & Đề xuất*

* Hoạt động bay và chi phí khá ổn định qua hai năm, không có biến động lớn.
* Phân bổ hạng ghế đồng đều, chưa có xu hướng dịch chuyển mạnh về một phân khúc cụ thể.
* Nên tiếp tục theo dõi chi phí quốc tế vì có dấu hiệu tăng nhẹ dù số chuyến giảm.
* Có thể cân nhắc các chương trình tối ưu chi phí cho các tháng có chi tiêu đột biến để kiểm soát ngân sách tốt hơn.

### 2.3.2. Phân tích hành vi hành khách

*Các biểu đồ phân tích*

* Xu hướng chuyến bay theo tháng & loại vé
* Phân bổ mục đích chuyến đi
* Thời gian đặt vé trước khi bay
* Phân bổ hạng ghế
* Mục đích chuyến đi

*Insight nổi bật*

* Tỷ lệ các mục đích chuyến đi khá đồng đều, không có mục đích nào chiếm ưu thế vượt trội.
* Hành khách chủ yếu đặt vé sát ngày, ưu tiên chuyến bay thẳng và phân bổ đều giữa các hạng ghế.
* Mục đích cá nhân vẫn chiếm ưu thế, tuy nhiên số lượng chuyến công tác có xu hướng tăng nhẹ trong năm 2019.
* Không có biến động lớn về hành vi đặt vé hoặc lựa chọn hạng ghế giữa hai năm, thể hiện sự ổn định trong hành vi tiêu dùng dịch vụ hàng không của nhóm khách hàng này.

*Kiến nghị & Đề xuất*

* Khách hàng chủ yếu đặt vé sát ngày bay và ưu tiên chuyến bay thẳng, cần tăng cường vé linh hoạt và dịch vụ tiện lợi.
* Phân bổ hạng ghế cân bằng giữa các phân khúc, nên đa dạng hóa sản phẩm và nâng cao trải nghiệm khách hàng cá nhân và công tác.
* Các chương trình dịch vụ và ưu đãi nên được thiết kế linh hoạt để phục vụ đa dạng nhóm khách hàng với các mục đích khác nhau.
* Có thể tập trung phát triển các gói dịch vụ phù hợp cho nhóm cá nhân vì đây là nhóm chiếm tỷ trọng lớn nhất (51%).

### 2.3.3. Phân tích phân khúc hãng hàng không

*Các biểu đồ phân tích*

* So sánh chi tiêu vé máy bay theo tháng (2018 vs 2019)
* Top hãng hàng không theo số chuyến bay
* Phân cụm hãng hàng không theo chi tiêu trung bình & số chuyến bay
* Doanh thu vé theo hãng hàng không

*Insight nổi bật*

* Năm 2018 có một số tháng đột biến chi tiêu (tháng 6, tháng 10), trong khi năm 2019 chi tiêu ổn định hơn, ít biến động lớn.
* **Hawaiian** là hãng có hiệu suất tốt nhất cả về số chuyến bay lẫn doanh thu.
* Các hãng **Premium** có doanh thu tốt nhờ giá vé cao, trong khi các hãng **Core** và **Budget** cạnh tranh về số lượng chuyến.
* Hạng ghế **First Class** và **Premium Economy** có mức tăng trưởng cao nhất trong khi **Business Class** sụt giảm mạnh nhất, đặc biệt ở phân khúc **Premium**.

*Kiến nghị & Đề xuất*

* Chi tiêu năm 2019 thấp hơn hoặc tương đương 2018 ở hầu hết các tháng, phản ánh xu hướng kiểm soát chi phí tốt hơn hoặc giảm nhu cầu.
* Các hãng **Premium Leaders** nên tiếp tục duy trì chiến lược giá vé cao kết hợp tăng số chuyến để tối ưu doanh thu.
* Các hãng **Core Competitors** và **Budget Regionals** có thể cân nhắc điều chỉnh chiến lược giá hoặc tăng chất lượng dịch vụ để cạnh tranh với nhóm dẫn đầu.

### 2.3.4. Phân tích chuyến bay nội địa

*Các biểu đồ phân tích*

* Số lượng chuyến bay xuất phát nội địa theo bang
* Số lượng chuyến bay và doanh thu vé nội địa theo tháng
* Doanh thu và tăng trưởng nội địa của các hãng hàng không trong 2 năm (2018 – 2019)
* Phân bổ mục đích chuyến đi nội địa

*Insight nổi bật*

* Hành khách nội địa giảm sử dụng **Business Class**, chuyển sang các hạng cao cấp khác như **Premium Economy** và **First Class** – phản ánh xu hướng tối ưu hóa chi phí theo chất lượng dịch vụ.
* Doanh thu tăng mạnh vào mùa hè (tháng 6–8), đặc biệt trong tháng 6/2018 và 8/2019; doanh thu thường giảm trong tháng 2, 4 và giảm đột ngột vào tháng 9/2019.
* Phân khúc **Premium** có sự suy giảm tăng trưởng ở nhiều hãng bay, **Budget** tăng mạnh (trừ Ryanair), còn nhóm **Core** có mức tăng trưởng không đồng đều.

*Kiến nghị & Đề xuất*

* Tập trung đẩy mạnh các sản phẩm Premium Economy và First Class, phù hợp với xu hướng nâng cấp trải nghiệm của khách hàng nội địa.
* Tối ưu chính sách giá và khuyến mãi theo mùa vụ, đặc biệt vào các tháng thấp điểm như tháng 2, 4 và 9.
* Rà soát hiệu quả từng phân khúc thị trường và điều chỉnh chiến lược cạnh tranh phù hợp theo từng nhóm đối thủ.

*2.3.5. Phân tích chuyến bay quốc tế*

*Các biểu đồ phân tích*

* Số lượng chuyến bay quốc tế theo các quốc gia
* Số lượng chuyến bay quốc tế theo hãng hàng không
* Doanh thu chuyến bay quốc tế theo tháng
* Phân phối số lượng chuyến bay khứ hồi theo ngày khứ hồi
* Doanh thu và tăng trưởng nội địa của các hãng hàng không trong 2 năm (2018 – 2019)

*Insight nổi bật*

* Trung Quốc, Đức và Anh là 3 quốc gia được khai thác chuyến bay quốc tế nhiều nhất.
* Doanh thu quốc tế ghi nhận ở mức cao nhất vào tháng 6/2018 và tháng 10 ở cả 2 năm trong khi tháng 2/2019 có sự suy giảm rõ rệt so với cùng kì.
* Phân khúc Premium dẫn đầu về số chuyến bay quốc tế; Ryanair tăng trưởng mạnh nhất; các hãng còn lại thể hiện sự phân hóa rõ rệt.
* Khách hàng quốc tế thường chọn thời gian khứ hồi từ 6–9 ngày.

*Kiến nghị & Đề xuất*

* Tăng cường khai thác các đường bay đến Trung Quốc, Đức, Anh và thiết kế sản phẩm phù hợp với thời gian lưu trú trung bình (6–9 ngày).
* Tập trung nâng cao chất lượng và tối ưu hoá lợi nhuận từ phân khúc Premium, đồng thời theo dõi kỹ động thái cạnh tranh của các hãng tăng trưởng nhanh như Ryanair.
* Lập kế hoạch doanh thu theo mùa, đẩy mạnh marketing và ưu đãi trong các tháng thấp điểm như tháng 2.